UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

PROCESOS DE WISHART

T E S I S

QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS

PRESENTA:
JORGE IGNACIO GONZÁLEZ CÁZARES

DR. GERÓNIMO FRANCISCO URIBE BRAVO
FACULTAD DE CIENCIAS, UNAM

ABRIL 2017 CDMX, MÉXICO

Dedicado a mis padres y hermanos que me han sido una invaluable fuente de apoyo, comprensión y motivación y que me han llevado a superarme constantemente.

Agradecimientos

Estoy profundamente agradecido con mis padres que me han apoyado y guiado a lo largo de mi vida, académica y no académica. También le debo esto en parte a mis hermanos que de una u otra forma me motivan a esforzarme y dar lo mejor de mí. Mis amigos y profesores me han guiado en más de un sentido, no sólo esclareciendo mis dudas y respondiendo mis preguntas, sino planteándome otras y permitiéndome desarrollarme más allá de lo que tenía originalmente planeado. También quiero agradecer a CONACyT por ayudarme a financiar estos estudios y este trabajo en especial.

Finalmente quiero hacer mención especial de mi tutor, el Dr. Gerónimo, quien me hizo ver que el significado de ser un matemático. Este debe ir más allá de hacer conocimiento y comunicarlo; es importante motivarlo, interpretarlo y hacerlo apetecible a los demás.

UNAM, Ciudad de México

Índice general

Introducción

La presente tesis tiene como objetivo dar una presentación clara, ordenada y comprensiva de la teoría concerniente a los llamados procesos de Wishart. De este mismo modo se busca dar soluciones y pruebas simples y poco obscuras a los resultados que plantearemos al respecto, motivando las ideas exploradas. Contrario a lo que se acostumbra en la literatura disponible, pretendemos evitar mandar al lector a buscar pruebas relacionadas en otros textos, volviendo al trabajo relativamente autocontenido salvo por un conocimiento básico de probabilidad, procesos estocásticos, álgebra lineal y demás.

Los procesos de Wishart son relativamente nuevos comparados con otro tipo de procesos, principalmente porque se requirió de muchos resultados modernos sobre semimartingalas en tiempo continuo y sobre distribuciones matriciales para que se diera su motivación histórica. No obstante, estos procesos han resultado de gran interés para una gran cantidad de disciplinas. Principalmente porque en algún sentido son generalizaciones matriciales de otros populares e importantes procesos, como el proceso cuadrado de Bessel y el modelo de Cox-Ingersoll-Ross (CIR).

Se han explorado aplicaciones del tema, principalmente a finanzas y a física. En el primero usualmente se plantean modelos de volatilidad de varios activos riesgosos no independientes, mientras que en el segundo se exploran sus valores propios y se modelan partículas repulsivas que viajan a una misma velocidad constante horizontalmente en un plano. También se tienen algunos métodos de estimación de parámetros y simulación en la bibliografía disponible (e.g., [1, 2, 3]).

Se estudiará la existencia de las distribuciones y procesos de Wishart con una diversidad de técnicas y resultados del cálculo estocástico y otros resultados avanzados sobre el cono de las matrices simétricas y positivas semidefinidas. Se investigarán los diversos procesos escalares asociados, las distribuciones límite e invariantes de los procesos de Wishart.

Las diferentes secciones están dedicadas al estudio de los procesos de Wishart como procesos de Markov, como semimartingalas y movimientos brownianos cambiados de tiempo y mezclados. Se presentarán los pocos casos en los que se tienen construcciones explícitas de procesos de Wishart.

De manera general, seguiremos el planteamiento de [1, 3, 4, 5, 6, 7, 8], recopilando sus ideas principales, tomando también técnicas y resultados importantes de [2, 9]. Sin embargo, en la medida de lo posible daremos pruebas distintas y más detalladas, corrigiendo los posibles errores encontrados en los mismos. El último capítulo está dedicado a concretar algunos resultados generales, para poder darle forma explícita a los mismos y vislumbrar la complejidad con la que se tiene que enfrentar el interesado, al buscar representaciones concretas en los casos generales. También se hablará un poco sobre cómo simular dichos procesos con bastante precisión, a costa de una alta carga computacional.

Capítulo 1

Preliminares y notación

Notación 1.1. Seguiremos la notación usual en probabilidad, específicamente, la que sigue [10]. En particular, consideramos que $\inf \emptyset = \infty = -\sup \emptyset$. Denotamos por S_d al conjunto de todas las matrices simétricas de $d \times d$ con valores reales. También consideramos a $\bar{S}_d^+$, S_d^+ y $\hat{S}_d^+$ como los conjuntos de matrices en S_d que son positivas semidefinidas, positivas definidas y positivas definidas con valores propios distintos respectivamente.

En lo que sigue ∂A denota la frontera de una conjunto A, con lo que definimos $\partial S_d^+ = \bar{S}_d^+ \backslash S_d^+$. Los conjuntos $\bar{S}_d^+$ y S_d^+ son conos, es decir que son convexos y cerrados ante multiplicación por reales positivos, donde $\bar{S}_d^+$ también es cerrado ante multiplicación por el 0 y si $a \in \bar{S}_d^+$ y $b \in S_d^+$, entonces $a + b \in S_d^+$.

Denotemos por $\lambda_i(x)$ al i-ésimo valor propio de $x \in S_d$, donde

$$\lambda_1(x) \leq \cdots \leq \lambda_d(x).$$

Más aún, en $M_{d,d}$ consideraremos el producto interior de Frobenius $\langle \cdot, \cdot \rangle : (x, y) \mapsto \operatorname{tr}(x'y)$ donde tr es el funcional traza y cuyo valor coincide con $\operatorname{tr}(xy)$ en S_d. Este producto interior induce una norma $\|\cdot\|_F$, conocida norma de Frobenius o norma de Hilbert-Schmidt, la cuál es equivalente a la norma de operadores $\|\cdot\|$ y además es invariante ante multiplicación por matrices ortogonales: si $R' = R^{-1}$, entonces

$$\begin{aligned}
\|AR\|_F &= \operatorname{tr}(R'A'AR) = \operatorname{tr}(RR'A'A) = \operatorname{tr}(A'A) = \|A\|_F \\
\|RA\|_F &= \operatorname{tr}(A'R'RA) = \operatorname{tr}(A'A) = \|A\|_F.
\end{aligned}$$

El producto interior también induce una noción de ortogonalidad en $M_{d,d}$, donde $U^\perp$ denota el anulador de un conjunto U, $\overline{U}$ denota su cerradura y U° su interior. De este modo notamos que justamente $\bar{S}_d^+$ es la cerradura de S_d^+. También consideramos a $\{c^{ij} : i \leq j \leq d\}$ como la base canónica de S_d, donde

$$c_{kl}^{ij} = \delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il}\left(1 - \delta_{ij}\right) = \begin{cases} 1 & \text{si } \{k,l\} = \{i,j\} \\ 0 & \text{e.o.c.} \end{cases},$$

y en ocasiones también usaremos a la base $\{e^{ij} : i \leq j \leq d\}$ dada por

$$e_{kl}^{ij} = \delta_{ik}\delta_{jl},$$

donde δ_{ij} es la delta de Kronecker. $M_{n,m}$ es el espacio de matrices de $n \times m$ mientras que I_d denota la matriz identidad en $M_{d,d}$. Para todo $x \in \bar{S}_d^+$, denotaremos por $\sqrt{x} = x^{\frac{1}{2}}$ la única matriz $y \in \bar{S}_d^+$ tal que $y^2 = x$. De manera general, también podemos definir $\sigma(A)$ como el espectro de A, es decir, el conjuntos de valores propios de una matriz $A \in M_{d,d}$.

Para un espacio localmente compacto segundo numerable y de Hausdorff S, denotamos por $\mathcal{S} = \mathcal{B}(S)$ su σ-álgebra de Borel, en tanto que $b(S)$ es el espacio de Banach de funciones reales Borel medibles y acotadas sobre S con la norma $\|\cdot\| : f \mapsto \sup_{x \in U} |f(x)|$. Por otro lado $C(S)$ es el espacio de funciones continuas, $C_b(S)$ es el espacio de funciones continuas y acotadas, $C_K(S)$ denota el espacio de funciones continuas con soporte compacto y $C_0(S)$ es el espacio de funciones continuas que se desvanecen al infinito.

Cuando además $S \subset S_d$ es medible, entonces definimos por $C^k(S)$ al espacio de funciones k veces diferenciable en S°, tal que todas sus derivadas parciales de orden $\leq k$ pertenecen a $C(S)$. Finalmente definimos $C^\infty(S) = \bigcap_{k \geq 1} C^k(S)$ y similarmente se definen $C_b^k(S)$, $C_K^k(S)$ y $C_0^k(S)$ para $0 \leq k \leq \infty$.

Uno de los resultados de importancia sobre estos espacios es que $\bar{S}_d^+$ es su propio dual, i.e.,

$$\bar{S}_d^+ = \left\{x \in S_d : \langle x, y\rangle \geq 0, \forall y \in \bar{S}_d^+\right\},$$

mientras que $\bar{S}_d^+$ y S_d^+ inducen un orden parcial en S_d, donde denotamos $x \preceq y$ o $x \prec y$ si $y - x \in \bar{S}_d^+$ o $y - x \in S_d^+$ respectivamente.

Esta característica es especialmente útil puesto que nos ayudará a poder definir la transformada de Laplace, tomando la esperanza de una exponencial de una forma $\exp(-\langle\cdot, u\rangle)$. Gracias a esto, no tendremos que preocuparnos, en principio, por problemas de integrabilidad de las distribuciones de Wishart.

También se le recomienda al lector que se familiarice con el cálculo matricial, especialmente el cálculo diferencial. Muchos de los resultados usados al respecto se pueden encontrar en [11, 12, 13, 14, 15]. Para amortiguar este problema, en la medida de lo posible y cuando se tenga una prueba rápida de esta clase de resultados, se presentará la misma en este trabajo. Esto tiene como objetivo el evitar que el lector tenga que recurrir constantemente a textos que son, relativamente, ajenos al tema.

Como preparativo, presentaremos una manera de calcular la derivada de $\sqrt{\cdot}$, la única raíz cuadrada simétrica y positiva semidefinida en $\bar{S}_d^+$. Para esto, definamos a la función vectorización vec, que toma una matriz y devuelve un vector con las columnas de la matriz apiladas una sobre la otra iniciando con la de el extremo izquierdo; explícitamente,

$$\text{vec}: \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix} \mapsto \begin{pmatrix} a_{1,1} & \cdots & a_{n,1} & a_{1,2} & \cdots & a_{n,m} \end{pmatrix}'.$$

En repetidas ocasiones usaremos el producto de Kronecker, dado por

$$\begin{aligned} \otimes : M_{n,m} \times M_{p,q} &\to M_{np,mq} \\ (A \otimes B)_{p(w-1)+y,q(x-1)+z} &:= A_{wx} B_{yz}. \end{aligned}$$

Este producto usualmente va de la mano con la vectorización, puesto que tienen varias propiedades deseables y compatibles entre ellos. Para ejemplificar este fenómeno, tomemos $A, B, C \in M_{n,n}$, entonces podemos conseguir con algunas manipulaciones de índices

$$\begin{aligned} & (ACB')_{ji} = \sum_{k,l=1}^{d} A_{jk} B_{il} C_{kl} \\ = & \sum_{k,l=1}^{d} (A \otimes B)_{i+n(j-1),l+n(k-1)} C_{kl} \\ = & \sum_{k,l=1}^{n} (A \otimes B)_{i+n(j-1),l+n(k-1)} \text{vec}\,(C')_{l+n(k-1)} \\ = & \left((A \otimes B)\,\text{vec}\,(C')\right)_{i+n(j-1)}, \end{aligned}$$

lo que también se puede escribir como

$$\begin{aligned} & (ACB')_{ji} \\ = & \sum_{k,l=1}^{d} (B \otimes A)_{j+n(i-1),k+n(l-1)} C_{kl} \\ = & \left((B \otimes A)\,\text{vec}\,(C)\right)_{j+n(i-1)}. \end{aligned}$$

Esto es útil si queremos despejar a C de un producto ACB'. Los cálculos anteriores en particular implican que

$$\operatorname{vec}(ACB) = \left(B' \otimes A\right)\operatorname{vec}(C).$$

También haremos uso del operador binario $\oplus$, la suma de Kronecker, para facilitar los cálculos, donde para $A \in M_{n,n}$ y $B \in M_{m,m}$ tenemos

$$A \oplus B := A \otimes I_m + I_n \otimes B.$$

Lema 1.2. *La función $\sqrt{\cdot}$ es continuamente diferenciable en S_d^+ y tenemos la siguiente representación para el jacobiano,*

$$D\sqrt{A} = \left(\sqrt{A} \oplus \sqrt{A}\right)^{-1},$$

donde para una función $F : M_{p,q} \to M_{m,n}$, el jacobiano está dado por $DF(A) = \frac{\partial vec(F(A))}{\partial vec(A)'}$.

Demostración. Derivando la ecuación $\sqrt{A}\sqrt{A} = A$, obtenemos una ecuación de Sylvester

$$\left(d\sqrt{A}\right)\sqrt{A} + \sqrt{A}\left(d\sqrt{A}\right) = dA,$$

cuya solución está dada por

$$\operatorname{vec}\left(d\sqrt{A}\right) = \left(\sqrt{A}' \oplus \sqrt{A}\right)^{-1}\operatorname{vec}(dA).$$

Cabe aclarar que como $A \in S_d^+$, entonces $\sqrt{A}' \oplus \sqrt{A} = \sqrt{A} \oplus \sqrt{A}$ es positiva definida (y por lo tanto no singular).

Finalmente, la regla de identificación del jacobiano (ver p. 198, capítulo 9, sección 5 en [15]) nos permite obtener el resultado deseado. En particular tenemos

$$\begin{aligned}
\frac{\partial}{\partial a_{ij}}\left(\sqrt{A}\right)_{kl} &= \left(\left(\sqrt{A} \oplus \sqrt{A}\right)^{-1}\operatorname{vec}\left(e^{ij}\right)\right)_{k+d(l-1)} \\
&= \left(\left(\sqrt{A} \oplus \sqrt{A}\right)^{-1}\right)_{k+d(l-1),i+d(j-1)} \\
&= \left(\sqrt{A}^{-1} \oplus \sqrt{A}^{-1}\right)_{k+d(l-1),i+d(j-1)} \\
&= \left(\sqrt{A}^{-1}\right)_{l,j}\left(\sqrt{A}^{-1}\right)_{k,i},
\end{aligned}$$

que nos será de gran utilidad más adelante. □

En preparativo para el lector en lo que se refiere a los cálculos matriciales, recordemos que

$$\operatorname{tr}\left(A^1 \cdots A^n\right) = \sum_{r_1,\ldots,r_n} A^1_{r_1 r_2} A^2_{r_2 r_3} \cdots A^n_{r_n r_1},$$

y por lo tanto, cuando nos encontremos con sumas como estas, en las que los índices de alguna manera forman un ciclo (inician y terminan en el mismo índice, o se puede llevar a este caso con algunas transposiciones), hemos de identificar trazas.

Si no es el caso, entonces seguramente se trata de la entrada (i, j) de algún producto de matrices. Es muy importante identificar apropiadamente a cuál de los dos casos nos enfrentamos en cada suma, puesto que la cantidad de índices puede confundir con facilidad a muchos lectores.

Capítulo 2

Procesos de Wishart

2.1. Distribución Wishart

En 1928, Wishart introdujo la distribución de las matrices de covarianza muestrales (estimadores de las matrices de covarianza) de las muestras de variables aleatorias independientes con distribución normal en su texto Biometrika. Las distribuciones actualmente conocidas como distribuciones de Wishart son múltiplos escalares de las que Wishart presentó en su trabajo.

Su introducción y estudio inicial detonaron un gran interés por su estudio en una variedad de disciplinas, entre ellas, estadística multivariada, finanzas, matemáticas financieras, economía y teoría de la probabilidad. Una forma simple de construir distribuciones Wishart es a través de vectores aleatorios normales, tal y como lo hizo Wishart en su trabajo.

Esta construcción es completamente análoga al caso real unidimensional, en donde surgen las distribuciones χ^2. En el transcurso de la tesis, es posible que el lector reconozca un parecido entre la transformada de Laplace (2.1), de las distribuciones de Wishart, y la de un las distribuciones χ^2 y gamma no centrales. Este resultado es completamente natural, y se junta con otras similitudes entre ambas distribuciones y los procesos asociados, que en este caso son el proceso cuadrado de Bessel o el modelo Cox-Ingersoll-Ross (CIR), con el proceso de Wishart.

Estas son algunas de las razones por las cuales se puede considerar a las distribuciones de Wishart como una generalización matricial de las distribuciones gamma no centrales, mientras que los procesos de Wishart son generalizaciones matriciales del modelo Cox-Ingersoll-Ross.

En esos casos, es fácil probar que se puede tomar cualquier parámetro de forma $k \geq 0$. Se puede conjeturar que el lado derecho de (2.1) corresponde a la transformada de Laplace de una distribución, también llamada Wishart, para $d \geq 2$ y cualquier elección de parámetros con $p \geq 0$, $\omega, \sigma \in \bar{S}_d^+$. No obstante, esta conjetura es falsa; de hecho, si consideramos al conjunto de Gindikin

$$\Delta_d = \left\{0, \frac{1}{2}, \ldots, \frac{d-1}{2}\right\} \cup \left(\frac{d-1}{2}, \infty\right),$$

se puede probar que el lado derecho de (2.1) es una transformada de Laplace de una distribución si y sólo si $p \in \Delta_d$ y $\operatorname{rank}(\omega) \leq 2p$, donde $\operatorname{rank}(A)$ es el rango de una matriz A. Este es uno de los principales objetivos de esta tesis.

Denotemos por $N_d(\mu, \Sigma)$ a la distribución normal en $\mathbb{R}^d$ con media μ y matriz de varianza $\Sigma \in S_d^+$ y por x' a la transpuesta de la matriz x. Sean $\xi_1, \ldots, \xi_k$ vectores columna aleatorios independientes con distribución $\xi_i \sim N_d(\mu_i, \Sigma)$ para $i \leq k$; entonces

$$\Xi = \sum_{i=1}^{k} \xi_i \xi_i',$$

es una matriz aleatoria con distribución Wishart con parámetro de escala $p = \frac{k}{2}$, parámetro de forma $\sigma = 2\Sigma$ y parámetro de no centralidad $\omega = \sum_{i=1}^{k} \mu_i \mu_i'$. Denotamos por $W_d(k, \omega, \sigma)$ a la distribución de Ξ. Notemos que Ξ es semidefinida positiva al igual que ω y σ.

Con un cálculo simple podemos obtener la transformada de Laplace de una distribución Wishart. Este cálculo, como es común en esta clase de problemas, se basa principalmente en sacar todos los términos que se puedan de las integrales y posteriormente completar densidades conocidas. En este caso, tendremos que «completar el cuadrado» para reconocer una densidad normal en la integral.

Para el siguiente resultado, y en lo que sigue, denotaremos por $\det(A)$ al determinante de una matriz cuadrada A.

Lema 2.1. *Sean $\xi_1, \ldots, \xi_k$ independientes con distribución $\xi_i \sim N_d(\mu_i, \Sigma)$ para $i \leq p$. Sea $\Xi = \sum_{i=1}^{k} \xi_i \xi_i'$, entonces la transformada de Laplace de Ξ es*

$$E \exp\{-\langle u, \Xi \rangle\} = \frac{\exp\left\{-tr\left(u\left(I_d + \sigma u\right)^{-1} \omega\right)\right\}}{\det\left(I_d + \sigma u\right)^p}, \quad u \in \bar{S}_d^+, \tag{2.1}$$

donde $p = \frac{k}{2}$, $\sigma = 2\Sigma$ y $\omega = \sum_{i=1}^{k} \mu_i \mu_i'$.

Demostración. Consideremos $u \in \bar{S}_d^+$ y notemos que los vectores $\zeta_i = \Sigma^{-\frac{1}{2}}(\xi_i - \mu_i)$ son iid con distribución $N(0, I_d)$ donde tomamos a $\Sigma^{\frac{1}{2}}$ como la única raíz cuadrada en S_d^+. Luego, como la traza permite rotaciones y la transpuesta de un real es él mismo,

$$\begin{aligned}
\langle u, \Xi \rangle &= \operatorname{tr}\left(u \sum_{i=1}^{k} \left(\Sigma^{\frac{1}{2}}\zeta_i + \mu_i\right)\left(\Sigma^{\frac{1}{2}}\zeta_i + \mu_i\right)'\right) = \sum_{i=1}^{k} \operatorname{tr}\left(\left(\zeta_i'\Sigma^{\frac{1}{2}} + \mu_i'\right) u \left(\Sigma^{\frac{1}{2}}\zeta_i + \mu_i\right)\right) \\
&= \sum_{i=1}^{k} \operatorname{tr}\left(\left(\zeta_i'\Sigma^{\frac{1}{2}} + \mu_i'\right) u \left(\Sigma^{\frac{1}{2}}\zeta_i + \mu_i\right)\right) = \sum_{i=1}^{k} \zeta_i'\Sigma^{\frac{1}{2}} u \Sigma^{\frac{1}{2}}\zeta_i + 2\zeta_i'\Sigma^{\frac{1}{2}} u\mu_i + \mu_i' u \mu_i \\
&= \sum_{i=1}^{k} \zeta_i' v \zeta_i + 2\sum_{i=1}^{k} \mu_i' u \Sigma^{\frac{1}{2}}\zeta_i + \operatorname{tr}\left(u \sum_{i=1}^{k} \mu_i\mu_i'\right) = \sum_{i=1}^{k} \zeta_i' v \zeta_i + \sum_{i=1}^{k} w_i'\zeta_i + \operatorname{tr}(u\omega),
\end{aligned}$$

donde $v = \Sigma^{\frac{1}{2}} u \Sigma^{\frac{1}{2}} \in \bar{S}_d^+$ y $w_i = 2\Sigma^{\frac{1}{2}} u \mu_i$. De aquí, por independencia obtenemos

$$E \exp\left\{-\langle u, \Xi \rangle\right\} = \exp\left\{-\langle u, \omega \rangle\right\} \prod_{i=1}^{k} E \exp\left\{-\zeta_i' v \zeta_i - w_i'\zeta_i\right\}.$$

Para calcular la esperanza de la derecha, supongamos que $z = (z_i)_{i=1}^{d}$ y $v = (v_{ij})_{i,j=0}^{d}$ y escribamos

$$\begin{aligned}
& E \exp\left\{-\zeta_i' v \zeta_i - w_i'\zeta_i\right\} = \int_{\mathbb{R}^d} (2\pi)^{-\frac{d}{2}} \exp\left\{-\frac{1}{2} z'z - z'vz - w_i'z\right\} dz \\
=\ & \frac{e^{\frac{1}{2} w_i' (I_d + 2v)^{-1} w_i}}{\sqrt{\det(I_d + 2v)}} \int_{\mathbb{R}^d} \sqrt{\frac{\det(I_d + 2v)}{(2\pi)^d}} e^{-\frac{1}{2}\left(z + (I_d + 2v)^{-1} w_i\right)' (I_d + 2v)\left(z + (I_d + 2v)^{-1} w_i\right)} dz \\
=\ & \exp\left\{\frac{1}{2} w_i' (I_d + 2v)^{-1} w_i\right\} \det(I_d + 2v)^{-\frac{1}{2}},
\end{aligned}$$

ya que el integrando en la segunda línea es la densidad de una distribución normal de media $(I_d + 2v)^{-1} w_i$ y varianza $(I_d + 2v)^{-1}$. Ahora usemos la fórmula de Sylvester para ver que

$$\det(I_d + 2v) = \det\left(I_d + 2\Sigma^{\frac{1}{2}} u \Sigma^{\frac{1}{2}}\right) = \det\left(I_d + 2\Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}} u\right) = \det(I_d + \sigma u),$$

de la misma forma, como la traza de una matriz y su transpuesta es la misma, tenemos

$$\begin{aligned}
& \sum_{i=1}^{k} \frac{1}{2} w_i' (I_d + 2v)^{-1} w_i \\
=\ & \sum_{i=1}^{k} \operatorname{tr}\left(\frac{1}{2} w_i' (I_d + 2v)^{-1} w_i\right) = \sum_{i=1}^{k} \operatorname{tr}\left(2\mu_i' u \Sigma^{\frac{1}{2}} \left(I_d + 2\Sigma^{\frac{1}{2}} u \Sigma^{\frac{1}{2}}\right)^{-1} \Sigma^{\frac{1}{2}} u \mu_i\right) \\
=\ & \operatorname{tr}\left(2u\left(\Sigma^{-\frac{1}{2}}\Sigma^{-\frac{1}{2}} + 2u\right)^{-1} u \left(\sum_{i=1}^{k} \mu_i\mu_i'\right)\right) = \operatorname{tr}\left(2u\left(\Sigma^{-1}(I_d + 2\Sigma u)\right)^{-1} u\omega\right) \\
=\ & \operatorname{tr}\left(u (I_d + \sigma u)^{-1} \sigma u \omega\right) = \operatorname{tr}\left(u (I_d + \sigma u)^{-1} (I_d + \sigma u - I_d)\, \omega\right) \\
=\ & \operatorname{tr}\left(u\left(I_d - (I_d + \sigma u)^{-1}\right)\omega\right) = \operatorname{tr}\left(u\omega - u (I_d + \sigma u)^{-1} \omega\right).
\end{aligned}$$

Finalmente tenemos

$$Ee^{-\langle u,\Xi\rangle} = \frac{e^{-\operatorname{tr}(u\omega)+\operatorname{tr}\left(u\omega-u(I_d+\sigma u)^{-1}\omega\right)}}{\det\left(I_d+\sigma u\right)^p} = \frac{e^{-\operatorname{tr}\left(u(I_d+\sigma u)^{-1}\omega\right)}}{\det\left(I_d+\sigma u\right)^p}.$$

□

Observación 2.2. Del mismo modo es posible probar que $Ee^{\langle u,\Xi\rangle}$ existe, es finito e igual a $\det\left(I_d-\sigma u\right)^{-p} e^{\operatorname{tr}\left(u(I_d-\sigma u)^{-1}\omega\right)}$ en tanto $I_d-\sigma u$ sea invertible y positivo semidefinido, i.e., los valores propios de $u\sigma$ son menores a 1. En particular, Ξ tiene momentos de todos los órdenes.

En general, se podría pensar que existen matrices aleatorias cuya transformada de Laplace tiene la misma forma que la de estas distribuciones Wishart.

Definición 2.3. Decimos que una matriz aleatoria Ξ con valores en S_d tiene distribución Wishart no central de parámetros de escala $p\geq 0$, forma $\sigma\in\bar{S}_d^+$ y no centralidad $\omega\in\bar{S}_d^+$ y lo denotamos por $\Xi\sim W_d\left(p,\omega,\sigma\right)$, si su transformada de Laplace, cuando existe, está dada por

$$Ee^{-\langle u,\Xi\rangle} = \frac{\exp\left\{-\operatorname{tr}\left(u\left(I_d+\sigma u\right)^{-1}\omega\right)\right\}}{\det\left(I_d+\sigma u\right)^p}, \quad u\in\bar{S}_d^+.$$

Por supuesto, es necesario probar que $\left(I_d+\sigma u\right)^{-1}$ siempre está definido para $u,\sigma\in\bar{S}_d^+$, de este modo, tenemos el siguiente resultado.

Lema 2.4. *Si $A,B\in\bar{S}_d^+$, entonces su producto tiene todos sus valores propios no negativos e iguales a los de $\sqrt{A}B\sqrt{A}$. En particular, I_d+AB es invertible.*

Demostración. Sea $A_t=A+tI_d$ para $t\geq 0$ y notemos que es invertible si $t>0$. Veamos ahora que A_tB es similar a

$$A_t^{-\frac{1}{2}}\left(A_tB\right)A_t^{\frac{1}{2}} = A_t^{\frac{1}{2}}BA_t^{\frac{1}{2}} = \left(A_t^{\frac{1}{2}}\right)'BA_t^{\frac{1}{2}}\in\bar{S}_d^+,$$

luego, A_tB tiene valores propios no negativos e iguales a los de $\sqrt{A_t}B\sqrt{A_t}$. Como A_t es continua, sus valores propios dependen continuamente de t también, así que conforme tomamos $t\to 0$, obtenemos que los valores propios de AB son todos no negativos e iguales a los de $\sqrt{A}B\sqrt{A}$ como queríamos probar. □

Daremos algunas propiedades importantes de la distribución Wishart que nos permitirán hacer simplificaciones más adelante, pero requeriremos de unos resultados técnicos.

Definición 2.5. Sea $A \in M_{n,m}$, una pseudoinversa de Moore-Penrose de A es una matriz $B \in M_{m,n}$ que satisface:

(I) $ABA = A$,

(II) $BAB = B$,

(III) $(AB)' = AB$,

(IV) $(BA)' = BA$.

Observación 2.6. Es conocido que existe una única pseudoinversa de A denotada A^g y que si A es invertible, entonces $A^g = A^{-1}$. Podemos notar que $P = AA^g$ y $Q = A^g A$ son proyecciones ortogonales sobre $\operatorname{Im}(A)$ y $\ker(A)^{\perp}$ respectivamente porque, de las primeras dos propiedades obtenemos

$$P^2 = AA^g AA^g = AA^g = P, \quad Q^2 = A^g AA^g A = A^g A = Q$$

y porque de las últimas dos propiedades obtenemos que $P' = P$ y $Q' = Q$. Además, podemos verificar que $(A^g)'$ satisface las propiedades de una pseudoinversa para A', lo que implica que $(A^g)' = (A')^g$:

(I) $A'(A^g)' A' = (AA^g A)' = A'$,

(II) $(A^g)' A' (A^g)' = (A^g AA^g)' = (A^g)'$,

(III) $\left(A'(A^g)'\right)' = A^g A = (A^g A)' = A'(A^g)'$,

(IV) $\left((A^g)' A'\right)' = AA^g = (AA^g)' = (A^g)' A'$.

En particular, si $n = m$ y A es simétrica, entonces A^g también lo es, en cuyo caso, la tercer propiedad implica que $AA^g = (AA^g)' = A^g A$, i.e., conmutan. En el otro caso particular en que $m = 1$, entonces una multiplicación directa muestra que $A^g = 1_{A \neq 0} \frac{A'}{A'A}$. Del mismo modo, para el caso general de n y m, se ve que $(A^g)^g = A$ y que bajo cualquiera de las siguientes condiciones se tiene que $(AB)^g = B^g A^g$, para $A \in M_{n,m}$ y $B \in M_{m,k}$,

(I) $A'A = I_m$,

(II) $BB' = I_m$,

(iii) $\operatorname{rank}(A) = m = \operatorname{rank}(B)$,

(iv) $B = A'$.

Además, si tenemos una sucesión de matrices $A_k \to A$, entonces $\operatorname{rank}(A_k) \to \operatorname{rank}(A)$ (constante a partir de un punto) si y sólo si $A_k^g \to A^g$ (ver [16]). En particular, si tenemos $\Sigma : \mathbb{R}_+ \to M_{n,m}$ es continua, entonces sus valores singulares (i.e., la raíz de los valores propios de $\Sigma\Sigma'$) son funciones continuas (ver [13]). Luego, el conjunto de puntos donde cada valor singular se anula es un conjunto cerrado. Por lo tanto, la unión de estos conjuntos sigue siendo cerrado y así, su complemento es un abierto.

Como todo abierto en $\mathbb{R}$ es unión numerable de intervalos abiertos disjuntos, entonces existe una cantidad de puntos a lo más numerable N donde algún valor propio se anula o deja de ser 0. Esto hace que $t \mapsto \Sigma(t)^g = \Sigma^g(t)$ tenga a lo más una cantidad numerable de discontinuidades y entonces podemos definir a $\tilde{\Sigma}(t) := \operatorname{lím\,sup}_{s\downarrow t} \Sigma(s)^g$ (donde el límite se toma entrada a entrada) es càdlàg. En particular, si Σ es adaptado, entonces podemos argumentar como en [17] para ver que $\tilde{\Sigma}$ es predecible y progresivamente medible.

Para nuestras aplicaciones, el conjunto de tiempos t donde $\tilde{\Sigma}(t) \neq \Sigma(t)^g$ tiene medida de Lebesgue 0 y por lo tanto cuando integremos respecto de semimartingalas continuas, no tendremos problemas si asumimos que en realidad son iguales (y por lo tanto, cumple las propiedades de la pseudoinversa de Moore-Penrose).

De hecho, es posible probar que el mapeo $A \mapsto A^g$ es medible. Para ver esto, simplemente notemos que si nos restringimos a los subespacios correspondientes a cada rango, el mapeo es continuo. Como estos subespacios son medibles, entonces podemos ver que la imagen inversa de cualquier abierto es entonces unión finita de conjuntos relativamente abiertos en cada uno de los subespacios medibles considerados. Esto hace que cada uno de dichos conjuntos sea medible y prueba que $A \mapsto A^g$ efectivamente es medible. Luego, si $\Sigma(t)$ es progresivamente medible, entonces $\Sigma^g(t)$ también lo es.

Lema 2.7. *(identidad de Woodbury) Supongamos que $B \in M_{m.m}$ es invertible y sean $A \in M_{n,n}$, $U \in M_{n,m}$ y $V \in M_{m,n}$. Si $B^{-1} + VA^{-1}U$ es invertible, entonces*

$$(A + AA^gUBVA^gA)^g = A^g - A^gU\left(B^{-1} + VA^gU\right)^{-1}VA^g,$$

además, si C^g es el lado izquierdo, entonces $CC^g = AA^g$ y $C^gC = A^gA$. Si A es invertible, entonces podemos reemplazar todas las pseudoinversas por verdaderas inversas y la expresión se simplifica.

Demostración. Realicemos la multiplicación por la izquierda:

$$
\begin{aligned}
& (A + AA^gUBVA^gA)\left(A^g - A^gU\left(B^{-1} + VA^gU\right)^{-1}VA^g\right) \\
= \; & AA^g - AA^gU\left(B^{-1} + VA^gU\right)^{-1}VA^g \\
- \; & AA^gUBVA^gU\left(B^{-1} + VA^gU\right)^{-1}VA^g + AA^gUBVA^g.
\end{aligned}
$$

Agregando un $BB^{-1} = I_m$ apropiadamente y agrupando los términos adecuados, obtenemos

$$
\begin{aligned}
& AA^g - AA^gUBB^{-1}\left(B^{-1} + VA^gU\right)^{-1}VA^g \\
- \; & AA^gUBVA^gU\left(B^{-1} + VA^gU\right)^{-1}VA^g + AA^gUBVA^g \\
= \; & AA^g - AA^gUB\left(B^{-1} + VA^gU\right)\left(B^{-1} + VA^gU\right)^{-1}VA^g + AA^gUBVA^g \\
= \; & AA^g - AA^gUBVA^g + AA^gUBVA^g = AA^g.
\end{aligned}
$$

La multiplicación por el otro lado da un resultado análogo:

$$
\begin{aligned}
& \left(A^g - A^gU\left(B^{-1} + VA^gU\right)^{-1}VA^g\right)(A + AA^gUBVA^gA) \\
= \; & A^gA - A^gU\left(B^{-1} + VA^gU\right)^{-1}VA^gA \\
- \; & A^gU\left(B^{-1} + VA^gU\right)^{-1}VA^gUBVA^gA + A^gUBVA^gA \\
= \; & A^gA - A^gU\left(B^{-1} + VA^gU\right)^{-1}B^{-1}BVA^gA \\
- \; & A^gU\left(B^{-1} + VA^gU\right)^{-1}VA^gUBVA^gA + A^gUBVA^gA \\
= \; & A^gA - A^gU\left(B^{-1} + VA^gU\right)^{-1}\left(B^{-1} + VA^gU\right)BVA^gA + A^gUBVA^gA \\
= \; & A^gA - A^gUBVA^gA + A^gUBVA^gA = A^gA.
\end{aligned}
$$

De las propiedades de A^g se siguen las que buscamos. □

Observación 2.8. El uso más común de la fórmula anterior es para $B = I_m$, en cuyo caso, se le conoce como la identidad de Kailath. También lo usaremos casi exclusivamente para cuando A es invertible.

Otra transformación de gran importancia nos ayudará a poder manipular simetrizaciones de la forma

$$S_{A,B} : X \mapsto AXB + B'X'A', \quad A, B, X \in M_{d,d}.$$

De hecho, las propiedades del producto de Kronecker nos ayudan a obtener

$$S_{A,B}(X) = (B' \otimes A)\,\mathrm{vec}(X) + (A \otimes B')\,\mathrm{vec}(X').$$

Considerando las siguientes funciones auxiliares inducidas por vec y vec^{-1} respectivamente:

$$\begin{array}{rcl} \upsilon(i,j) & \mapsto & i + d(j-1), \quad i,j \in \{1,\ldots,d\} \\ \upsilon^{-1}(i) & \mapsto & \left(i - d\left(\left\lceil \frac{i}{d} \right\rceil - 1\right), \left\lceil \frac{i}{d} \right\rceil\right), \quad i \in \left\{1,\ldots,d^2\right\}, \end{array}$$

podemos usar los cálculos hechos en el primer capítulo, los que implican que si definimos la matriz

$$\begin{array}{rcl} (S_{A,B})_{\upsilon(i,j),\upsilon(k,l)} & = & (B' \otimes A)_{\upsilon(i,j),\upsilon(k,l)} + (B' \otimes A)_{\upsilon(j,i),\upsilon(k,l)} \\ & = & B_{lj}A_{ik} + B_{li}A_{jk}, \end{array}$$

entonces tenemos

$$S_{A,B}(X) = S_{A,B}\text{vec}(X).$$

También se puede deducir fácilmente que

$$\begin{array}{rcl} \left(S'_{A,B}\right)_{\upsilon(i,j),\upsilon(k,l)} & = & (S_{A,B})_{\upsilon(k,l),\upsilon(i,j)} \\ & = & (B \otimes A')_{\upsilon(i,j),\upsilon(k,l)} + (A' \otimes B)_{\upsilon(j,i),\upsilon(k,l)} \\ & = & B_{jl}A_{ki} + B_{jk}A_{li}, \end{array}$$

y por lo tanto,

$$\begin{array}{rcl} S'_{A,B}(X) & = & A'(X + X')B' \\ & = & A' S_{I_d,I_d}(X) B'. \end{array}$$

Es evidente que si A y B son invertibles, entonces

$$\text{rank}\left(S'_{A,B}\right) = \text{rank}\left(S'_{I_d,I_d}\right) = \frac{d(d+1)}{2},$$

puesto que este último tiene imagen S_d y por kernel tiene a las matrices antisimétricas. De hecho, la proyección ortogonal sobre S_d es claramente $\frac{1}{2}S_{I_d,I_d}$ puesto que tiene imagen S_d, es idempotente y autoadjunto.

Consecuentemente, las siguientes deducciones, aunque aparentemente intuitivas, son falsas:

$$\begin{array}{rcl} S_{A,B}\left(A^{-1}\right) = B + B' & \Longrightarrow & S^g_{A,B}(B + B') = A^{-1} \\ S_{A,B}\left(B^{-1}\right) = A + A' & \Longrightarrow & S^g_{A,B}(A + A') = B^{-1}. \end{array}$$

Esto se debe a que $S^g_{A,B}S_{A,B} \neq \frac{1}{2}S_{I_d,I_d}$ en general. No obstante, podemos conseguir conclusiones similares en otro contexto. Notemos que $S_{A,B}S'_{A,B} = S_{AA',B'B} + S_{B'B,AA'}$ cuya imagen nuevamente está en S_d y desde luego preserva su rango, de donde concluimos que

$$\left(S_{A,B}S'_{A,B}\right)^g S_{A,B}S'_{A,B} = \left(S_{A,B}S'_{A,B}\left(S_{A,B}S'_{A,B}\right)^g\right)' = \frac{1}{2}S'_{I_d,I_d} = \frac{1}{2}S_{I_d,I_d}.$$

Aquí sí se siguen las siguientes implicaciones:

$$\begin{aligned} S_{A,B}\left(S'_{A,B}(X)\right) &= AA'(X+X')B'B + B'B(X+X')AA' \\ S_{A,B}\left(S'_{A,B}\left((AA')^{-1}\right)\right) &= 4B'B \implies \left(S_{A,B}S'_{A,B}\right)^g \operatorname{vec}(B'B) = \frac{1}{4}(AA')^{-1} \\ S_{A,B}\left(S'_{A,B}\left((B'B)^{-1}\right)\right) &= 4AA' \implies \left(S_{A,B}S'_{A,B}\right)^g \operatorname{vec}(AA') = \frac{1}{4}(B'B)^{-1}. \quad (2.2) \end{aligned}$$

Estas relaciones nos serán de utilidad en el último capítulo.

A continuación probaremos un resultado importante que usaremos en una sección mucho más adelante, pero cuya prueba se puede dar desde ahora y cuya idea en sí es bastante interesante. Para esto, recordando lo probado de la transformada de Laplace de una distribución Wishart, definamos para $\nu = W_d(p,\omega,\sigma)$,

$$D(\nu) = \left\{u \in S_d : \mathcal{L}_\nu(u) := \int_{\bar{S}_d^+} e^{-\langle u,\xi\rangle}\nu(d\xi) < \infty\right\},$$

cuya forma está perfectamente descrita en 2.2. En particular, cuando $\sigma \in S_d^+$, tenemos que $D(\nu) = -\sigma^{-1} + S_d^+$. Para cada $u \in D(\nu)$ le podemos asociar la distribución exponencial $F(\nu,u)$ dada por

$$F(\nu,u)(A) = \int_A \frac{\exp\{-\langle u,\xi\rangle\}\nu(d\xi)}{\mathcal{L}_\nu(u)},$$

la cuál es claramente una medida de probabilidad con $\frac{dF(\nu,u)}{d\nu} = \frac{\exp\{-\langle u,\cdot\rangle\}}{\mathcal{L}_\nu(u)} > 0$.

Proposición 2.9. *Sean $p \geq 0$, $\omega,\sigma \in \bar{S}_d^+$ y supongamos que existe $\Xi \sim W_d(p,\omega,\sigma)$. Supongamos que rank$(\sigma) = d_0 \leq d$ y sea $I_{d_0,d}$ la matriz diagonal que tiene en sus primeras d_0 entradas de la diagonal un 1 y 0 en el resto. Entonces existe $\rho \in M_{d,d}$ tal que*

$$\rho\Xi\rho' \sim W_d\left(p,\rho\omega\rho',I_{d_0,d}\right),$$

en consecuencia, si $\sigma \in S_d^+$, una distribución existe si y sólo si la otra existe. En general, tenemos para todo $\rho \in M_{n,d}$ que

$$\rho\Xi\rho' \sim W_n\left(p,\rho\omega\rho',\rho\sigma\rho'\right).$$

Si existe $W_d(p,\omega,I)$ y $\sigma \in S_d^+$, entonces $W_d(p,\sigma\omega\sigma,\sigma)$ también existe y la implicación inversa también es cierta. En el caso en que $\sigma \in S_d^+$, si $W_d(p,\omega,\sigma)$ existe, entonces

$W_d(p,\omega,t\sigma)$ también existe para todo $t>0$.

Demostración. Probemos primero la tercer afirmación. Para esto, analicemos la transformada de Laplace de $\rho\Xi\rho'$ usando la propiedad cíclica de la traza y la fórmula de Sylvester. Con esta idea, algunas manipulaciones algebraicas y matriciales podemos deducir que para todo $u\in\bar{S}_n^+$,

$$\begin{aligned} Ee^{-\langle u,\rho\Xi\rho'\rangle} &= Ee^{-\operatorname{tr}(u\rho\Xi\rho')}=Ee^{-\operatorname{tr}(\rho' u\rho\Xi)}=Ee^{-\langle\rho' u\rho,\Xi\rangle} \\ &= \frac{e^{-\operatorname{tr}\left(\rho' u\rho(I_d+\sigma\rho' u\rho)^{-1}\omega\right)}}{\det\left(I_d+\sigma\rho' u\rho\right)^p} \\ &= \frac{e^{-\operatorname{tr}\left(u\rho(I_d+\sigma\rho' u\rho)^{-1}\omega\rho'\right)}}{\det\left(I_n+\rho\sigma\rho' u\right)^p}. \end{aligned}$$

Usando ahora la identidad de Kailath (lema 2.7) para $A=I_d$, $U=\sigma\rho' u$ y $V=\rho$ y obtengamos, gracias a la invertibilidad de $I_d+I_{d_0,d}u$,

$$\left(I_d+\sigma\rho' u\rho\right)^{-1} = I_d-\sigma\rho' u\left(I_n+\rho\sigma\rho' u\right)^{-1}\rho,$$

de dónde

$$\begin{aligned} u\rho\left(I_d+\sigma\rho' u\rho\right)^{-1}\omega\rho' &= u\rho\left(I_d-\sigma\rho' u\left(I_n+\rho\sigma\rho' u\right)^{-1}\rho\right)\omega\rho' \\ &= u\rho\omega\rho'-u\rho\sigma\rho' u\left(I_n+\rho\sigma\rho' u\right)^{-1}\rho\omega\rho' \\ &= u\rho\omega\rho'+u\left(I_n-\left(I_n+\rho\sigma\rho' u\right)\right)\left(I_n+\rho\sigma\rho' u\right)^{-1}\rho\omega\rho' \\ &= u\rho\omega\rho'+u\left(\left(I_n+\rho\sigma\rho' u\right)^{-1}-I_n\right)\rho\omega\rho' \\ &= u\left(I_n+\rho\sigma\rho' u\right)^{-1}\rho\omega\rho'. \end{aligned}$$

Al juntar esta igualdad con la expresión que encontramos para $Ee^{-\langle u,\rho\Xi\rho'\rangle}$ obtenemos la última afirmación.

Para probar la segunda afirmación, sólo hace falta multiplicar por ρ^{-1} (la cuál existe porque σ es invertible) de ambos lados y seguir el proceso de vuelta, notando que en este caso el parámetro de no centralidad es $\rho^{-1}\rho\omega\rho'\left(\rho^{-1}\right)'=\omega$. Para la primera afirmación, como $\sqrt{\sigma}\sqrt{\sigma}^g=\sqrt{\sigma}^g\sqrt{\sigma}$ es una proyección ortogonal simétrica, tiene un1a diagonalización $R\sqrt{\sigma}\sqrt{\sigma}^gR'=\Lambda$.

Aquí debemos tener $\Lambda=I_{d_0,d}$ porque como proyección ortogonal, todo elemento en su imagen (la cuál tiene dimensión $\operatorname{rank}(\sqrt{\sigma})=\operatorname{rank}(\sigma)=d_0$) es vector propio del valor propio 1, y todo elemento en $\operatorname{Im}(\sqrt{\sigma})^{\perp}=\ker(\sigma)$ (de dimensión $d-d_0$) es vector propio del valor propio 0. Si tomamos $\rho = R\sqrt{\sigma}^g$, entonces obtenemos la primera afirmación de la proposición.

Ahora notemos que $\sigma^{-1} - I_d + u \in D(\nu)$ donde $\nu = W_d(p, \omega, I)$ para cualquier $u \in \bar{S}_d^+$. Si tomamos $v = \sigma^{-1} - I_d$ entonces $I_d + u + v = u + \sigma^{-1} = \sigma^{-1}(I_d + \sigma u)$, por lo que

$$\int_{\bar{S}_d^+} e^{-\langle u+v,\xi\rangle}\nu(d\xi) = \frac{e^{-\operatorname{tr}\left((u+v)(I_d+u+v)^{-1}\omega\right)}}{\det(I_d+u+v)^p} = \frac{e^{-\operatorname{tr}\left(\left(-I_d+u+\sigma^{-1}\right)\left(u+\sigma^{-1}\right)^{-1}\omega\right)}}{\det(\sigma)^{-p}\det(I_d+\sigma u)^p},$$

donde el exponente se simplifica a

$$\begin{aligned}
-\operatorname{tr}\left(\omega - \left(u+\sigma^{-1}\right)^{-1}\omega\right) &= -\operatorname{tr}\left((I_d-\sigma)\omega - \left(\left(u+\sigma^{-1}\right)^{-1} - \sigma\left(u+\sigma^{-1}\right)\left(u+\sigma^{-1}\right)^{-1}\right)\omega\right) \\
&= -\operatorname{tr}\left((I_d-\sigma)\omega - (I_d - \sigma u - I_d)\left(u+\sigma^{-1}\right)^{-1}\omega\right) \\
&= -\operatorname{tr}\left((I_d-\sigma)\omega\right) - \operatorname{tr}\left(\sigma u (I_d+\sigma u)^{-1}\sigma\omega\right) \\
&= -\operatorname{tr}\left((I_d-\sigma)\omega\right) - \operatorname{tr}\left(u (I_d+\sigma u)^{-1}\sigma\omega\sigma\right).
\end{aligned}$$

Por lo tanto,

$$\mathcal{L}_{F(\nu,\sigma^{-1}-I_d)}(u) \propto \mathcal{L}_{W_d(p,\sigma\omega\sigma,\sigma)}(u),$$

lo que prueba la cuarta afirmación puesto que el regreso se sigue de que la derivada de Radon-Nikodym $\frac{dF\left(\nu,\sigma^{-1}-I_d\right)}{d\nu}$ sea estrictamente positiva, lo que hace que ambas medidas sean equivalentes. Para la última afirmación deberemos usar repetidamente la tercera y cuarta, concretamente, sólo debemos seguir la siguiente cadena de existencias distribucionales para $t > 0$ en el caso en que $\sigma \in S_d^+$,

$$\begin{aligned}
W_d(p,\omega,\sigma) &\ \exists \implies \\
W_d\left(p, \sigma^{-1}\omega\sigma^{-1}, I_d\right) &\ \exists \implies \\
W_d\left(p, \frac{\sigma}{t}\sigma^{-1}\omega\sigma^{-1}\frac{\sigma}{t}, \frac{\sigma}{t}\right) &\ \exists \implies \\
W_d\left(p, tI_d\frac{\omega}{t^2}tI_d, tI_d\frac{\sigma}{t}tI_d\right) &\ \exists \implies \\
W_d(p,\omega,t\sigma) &\ \exists \ ,
\end{aligned}$$

como queríamos probar. □

Observación 2.10. Una característica interesante de las distribuciones Wishart es que sí son consistentes en el sentido siguiente: si $\Xi \sim W_d(p,\omega,\sigma)$, entonces $I_{d_0,d}\Xi I_{d_0,d}$ es distribución Wishart. Esto es, si nos restringimos a las primeras d_0 dimensiones, sí recuperamos la distribución Wishart. De hecho, los parámetros corresponden a p y a las submatrices de

ω y σ que consisten en las primeras d_0 entradas. Esta propiedad no se preserva para los procesos de Wishart en general, como veremos más adelante.

También cabe aclarar que más adelante daremos una caracterización para las distribuciones Wishart cuando $\sigma \in S_d^+$, pero si σ no es de rango completo, entonces tiene rango $d_0 < d$. En este caso, podemos encontrar $\rho \in M_{d_0,d}$ tal que $\rho\sigma\rho' = I_{d_0}$ y $\sigma = \rho^g (\rho')^g$, por lo que en cierto sentido podemos reducir el problema a un contexto en el que σ es de rango completo. El único problema es que sólo caracterizaríamos a las distribuciones donde la imagen de ω está contenida en la imagen de σ.

Corolario 2.11. *Cualquier distribución Wishart $W_d(p, \omega, \sigma)$ tiene soporte en $\bar{S}_d^+$.*

Demostración. Basta ver que para toda $x \in \mathbb{R}^d$ tenemos que $x'\Xi x$ tiene soporte $\mathbb{R}_+$. Para esto, usemos la proposición 2.9 para ver que para todo $u \geq 0$, tenemos

$$Ee^{-ux\Xi x'} = \frac{e^{-\mathrm{tr}\left(u(1+x'\sigma xu)^{-1}x'\omega x\right)}}{\det(1+x'\sigma xu)^p} = (1+x'\sigma xu)^{-p}\exp\left\{-\frac{x'\omega xu}{1+x'\sigma xu}\right\},$$

que es la transformada de Laplace de una distribución χ^2 no central con $2p$ grados de libertad, y parámetro de no centralidad dado por $2\frac{x'\omega x}{x'\sigma x}$ evaluado en $\frac{1}{2}x'\sigma xu$. Es decir que $x'\Xi x$ es una variable aleatoria con distribución χ^2 no central y escalada, lo que en particular implica que tiene soporte en $\mathbb{R}_+$ como queríamos. □

Los procesos de Wishart se pueden entender en muchos sentidos. Su construcción original fue a base de resolver EDE (ecuaciones diferenciales estocásticas), sin embargo recientemente han surgido otros enfoques de semimartingalas o procesos afines. Como es de esperarse, se verá que sus distribuciones marginales son Wishart.

2.2. Construcción de los procesos de Wishart

A lo largo de este capítulo, supondremos que nuestros procesos y variables aleatorias están todos definidos en un espacio de probabilidad filtrado $(\Omega, \mathscr{F}, \mathcal{F}, \mathbb{P})$ donde $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$ es continuo por la derecha y completo. La notación y los resultado aquí usados de la teoría general de procesos estocásticos deben entenderse de acuerdo a sus versiones en [17, 18, 10], a donde referimos al lector ante cualquier complicación en la lectura de este trabajo. Denotaremos por Id al mapeo identidad en $\mathbb{R}$, i.e., $Id : t \mapsto t$.

Nuestro principal objetivo es construir los procesos de Wishart más elementales.

Esto se logra al contruir semimartingalas a base de movimientos brownianos matriciales. Posteriormente, calculamos su transformada de Laplace en el teorema (2.42), ecuación (2.15) y la EDE que satisface en los lemas (2.26) y (2.28), lo que motiva la definicion más general de procesos de Wishart.

Para un proceso con valores matriciales X, denotaremos por X^{ij} a su entrada (i,j), mientras que para una matriz constante a las denotaremos por a_{ij}. Para todo proceso de variación finita en compactos A en $\mathbb{R}$, denotamos por $L(A)$ al conjunto de procesos predecibles y progresivamente medibles V tal que $(V \cdot A)_t := \int_0^t V_s dA_s$ existe en el sentido de la integral de Stieltjes. Si $B = M + A$ es una semimartingala continua con M una martingala local y A un proceso de variación finita en compactos, denotamos por

$$L(B) = \left\{ V \in L(A) : V^2 \in L([M]) \right\}.$$

De este modo $V \cdot U$ denota la integral $(V \cdot U)_t = \int_0^t V_s dU_s$ (estocástica o no) siempre que $V \in L(U)$ mientras que $[V,U]$ y $[U]$ son la covariación cuadrática de dos semimartingalas U y V y la variación cuadrática de U. También adoptamos la notación $dUdV = d[U,V]$ cuando sea necesario.

Cuando U, V y B toman valores en $M_{n,m}$, $M_{l,k}$ y $M_{m,l}$ respectivamente, diremos que B es semimartingala si lo es por cada entrada y entenderemos que $(U,V) \in L(B)$ si $U^{ir}V^{sj} \in L(B^{rs})$ para todo $i \le n$, $r \le m$, $s \le l$ y $j \le k$. También lo denotaremos a veces por $U \in lL(B)$ cuando $(U, I_l) \in L(B)$ y $V \in rL(B)$ cuando $(I_m, V) \in L(B)$. Así, $\int_0^t U_u dB_u V_u$ es un proceso con valores en $M_{n,k}$ dado por

$$\left(\int_0^t U_u dB_u V_u \right)^{ij} = \sum_{r=1}^{m} \sum_{s=1}^{l} \int_0^t U_u^{ir} dB_u^{rs} V_u^{sj}, \quad i \le n, \quad j \le k.$$

Intuitivamente, lo que sucede es que estamos tratando a dB como la matriz $\left(dB^{ij}\right)_{ij}$ y realizamos las integrales de acuerdo al producto matricial. Mientras tanto, si hablamos de vectores, podemos entender a $\mathbb{R}^k$ como el espacio $M_{k,1}$ o como $M_{1,k}$ de acuerdo a como sea conveniente para evitar problemas de dimensionalidad.

De esta forma, cuando integremos respecto de Id, lo que interpretamos es que a cada entrada la integramos con respecto a la medida de Lebesgue, en este sentido, es como integrar respecto de $I_n Id$ para alguna $n \in \mathbb{N}$.

Si B y C son semimartingalas con valores en $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente, consideraremos a $[B,C]$, su covariación cuadrática, como el proceso con valores en $M_{n,m}$ dado por $[B,C]^{ij} = [B^i, C^j]$ para $i \le n$ y $j \le m$.

Una noción similar se puede entender para matrices, sin embargo, nuestra covariación ya no toma valores matriciales y evitaremos a toda costa darle alguna interpretación más allá del proceso multidimensional que a cada entrada de B y C le asigna su covariación usual. Por ejemplo, si B y C son semimartingalas en $\mathbb{R}^k$ y $\mathbb{R}^l$ mientras que $U \in lL(B)$ y $V \in lL(C)$ toman valores en $M_{n,k}$ y $M_{m,l}$, entonces para $i \leq n$ y $j \leq m$,

$$[UdB, VdC]^{ij} = \left[\sum_{r=1}^{k} U^{ir} dB^r, \sum_{s=1}^{l} V^{js} dC^s\right] = \sum_{r,s} U^{ir} d[B^r, C^s](V')^{sj} = (Ud[B,C]V')^{ij}.$$

Como en la situación anterior, en ocasiones es más conveniente usar la notación diferencial que la integral.

De este modo se pueden extender nociones y resultados que ya se conocen bastante bien en el caso real e incluso vectorial. Usaremos constantemente la fórmula de integración por partes y la fórmula de Itô, por lo tanto, vamos a enunciarlas para el caso matricial.

Teorema 2.12. *(integración por partes y fórmula de Itô)* Sean A y B semimartingalas continuas con valores en $M_{n,k}$ y $M_{k,m}$ respectivamente y sea f una función real de clase C^2 sobre $M_{n,k}$. Si $A \in lL(B)$ y $B \in rL(A)$, entonces tenemos

$$d(AB) = AdB + dAB + dAdB.$$

Además,

$$df(A) = \operatorname{tr}\left(Df(A)' dA\right) + \frac{1}{2}\sum_{i,j=1}^{n}\sum_{r,s=1}^{k} \partial_{ir}\partial_{js} f(A)\, d\left[A^{ir}, A^{js}\right],$$

donde $D = (\partial_{ij})$.

Demostración. La primera fórmula se sigue de la fórmula de integración por partes usual: para todo $t \geq 0$, $i \leq n$ y $j \leq m$ tenemos precisamente

$$d(AB)^{ij} = \sum_{l=1}^{k} d\left(A^{il}B^{lj}\right) = \sum_{l=1}^{k} A^{il} dB^{lj} + dA^{il} B^{lj} + dA^{il} dB^{lj}.$$

Para la segunda fórmula, usemos la fórmula de Itô usual. Con esto, sólo hace falta notar que

$$\operatorname{tr}\left(Df(A)' dA\right) = \sum_{i=1}^{n}\sum_{l=1}^{k} (Df(A))^{il} dA^{li} = \sum_{i=1}^{n}\sum_{l=1}^{k} \partial_{li} f(A)\, dA^{li},$$

para obtener el resultado deseado. □

También resultará conveniente usar la integral estocástica de Stratonovich porque la fórmula de Itô y la de integración por partes adquieren formas más simples. Para esto, consideramos semimartingalas U y V en $M_{k,n}$ y $M_{n,m}$ respectivamente, entonces definimos a la integral de Stratonovich dada por la forma diferencial

$$\begin{aligned} U_t \circ dV_t &= U_t dV_t + \frac{1}{2} dU_t dV_t, \\ dU_t \circ V_t &= dU_t V_t + \frac{1}{2} dU_t dV_t, \\ dU_t \circ dV_t &= dU_t dV_t. \end{aligned}$$

Con esta consideración, la fórmula de integración por partes establece simplemente que $d\left(U_t V_t\right) = U_t \circ dV_t + dU_t \circ V_t$. Asimismo tenemos para una semimartingala W en $M_{m,l}$,

$$dU_t \circ \left(V_t W_t\right) = \left(dU_t \circ V_t\right) \circ W_t, \quad \left(dU_t \circ V_t\right)' = V_t' \circ dU_t',$$

y otras relaciones análogas se pueden concluir con facilidad.

De ahora en adelante, denotaremos por $B \sim BM_{n,m}$ a un proceso B a tiempo continuo con valores en $M_{n,m}$ cuyas entradas son movimientos brownianos estándar independientes. El interés en los procesos de Wishart surgió con [5], cuando Marie-France Bru por primera vez construyó procesos de Wishart a través de movimientos brownianos matriciales. Ella inició proponiendo a los procesos de Wishart como generalizaciones matriciales de los procesos cuadrado de Bessel. Posteriormente, el problema se llevó al estudio de procesos matriciales que satisficieran una EDE particular. No obstante, en años más recientes se ha encontrado una manera alterna de ver el problema, que en ocasiones es más sencilla de manejar.

A los procesos cuadrado de Bessel se les puede construir (para dimensiones enteras) como el cuadrado de la norma euclidiana de un movimiento browniano, i.e., si $B \sim BM_{n,1}$, entonces $B'B = \|B\|_F$ es un proceso cuadrado de Bessel n-dimensional. Su extensión natural sería considerar ahora al producto de tal forma que nos resultara una matriz en vez de un real. De esta manera, iniciemos considerando $S \sim BM_{d,p}$, una matriz arbitraria $s \in M_{d,p}$ y construyamos al proceso $X = \left(\Sigma S + s\right)\left(\Sigma S + s\right)'$, conocido como proceso de Wishart de dimensión d, índice p, forma $\Sigma \in \bar{S}_d^+$ y estado inicial $ss' \in \bar{S}_d^+$.

Algo inmediato es que X siempre es positivo semidefinido. Además, este proceso satisface una variedad de EDEs bastante importantes, en tanto que su distribución a cada tiempo es Wishart, como probaremos próximamente. Antes de probar esta propiedad, requerimos de un resultado técnico necesario que vincula a las variaciones cuadráticas absolutamente continuas y al movimiento browniano.

Lema 2.13. *(representación integral de Doob) Supongamos que M es una martingala local continua en $\mathbb{R}^n$ con $M_0 = 0$ y con variación cuadrática dada por $[M] = A \cdot I_n Id$ donde $A = \sigma\sigma'$ para un proceso progresivamente medible σ en $M_{n,m}$. Entonces existe, posiblemente en una extensión del espacio de probabilidad, un movimiento browniano B en $\mathbb{R}^m$ que satisface $M = \sigma \cdot B$.*

Demostración. Extendamos el espacio de probabilidad para incorporar un movimiento browniano independiente W en $\mathbb{R}^n$. De acuerdo al hecho 2.6, podemos definir un proceso σ^g progresivamente medible en $M_{m,n}$, el cuál nos ayudará a encontrar a B. Con este fin, tomamos $\rho = I_m - \sigma^g \sigma$ y definimos

$$B = \sigma^g \cdot M + \rho \cdot W,$$

lo que resta es probar que B es un movimiento browniano y que $M = \sigma \cdot B$. Para esto, notemos que B es una martingala local por construcción, gracias a que M y W lo son; además, su variación cuadrática está dada por (gracias a la independencia entre M y W),

$$\begin{aligned} [B] &= (\sigma^g, \rho) \begin{pmatrix} A \cdot Id & 0 \\ 0 & Id \end{pmatrix} \begin{pmatrix} (\sigma^g)' \\ \rho' \end{pmatrix} \\ &= \Big(\sigma^g \sigma \left(\sigma^g \sigma\right)' + \left(I_n - \sigma^g \sigma\right)\left(I_n - \sigma^g \sigma\right)\Big) \cdot Id \\ &= \left(\sigma^g \sigma \sigma^g \sigma + I_n - 2\sigma^g \sigma + \sigma^g \sigma \sigma^g \sigma\right) \cdot Id \\ &= \left(\sigma^g \sigma + I_n - 2\sigma^g \sigma + \sigma^g \sigma\right) \cdot Id = I_m Id. \end{aligned}$$

Luego, el teorema de caracterización de Lévy muestra que B es un movimiento browniano y además,

$$\sigma \cdot B = \sigma\sigma^g \cdot M + \sigma\left(I_m - \sigma^g \sigma\right) \cdot W = \sigma\sigma^g \cdot M + \left(\sigma - \sigma\sigma^g \sigma\right) \cdot W = \sigma\sigma^g \cdot M.$$

Así que sólo resta probar que la martingala local $M - \sigma\sigma^g \cdot M$ se anula idénticamente. Para esto, calculemos su variación cuadrática

$$\begin{aligned} [M - \sigma\sigma^g \cdot M] &= A - 2\sigma\sigma^g \cdot A + \sigma\sigma^g \sigma\sigma^g \cdot A \\ &= \sigma\sigma' \cdot Id - 2\sigma\sigma^g \cdot \left(\sigma\sigma' \cdot Id\right) + \sigma\sigma^g \cdot \left(\sigma\sigma' \cdot Id\right) \\ &= \sigma\sigma' \cdot Id - 2\sigma\sigma^g \sigma\sigma' \cdot Id + \sigma\sigma^g \sigma\sigma' \cdot Id \\ &= \sigma\sigma' \cdot Id - 2\sigma\sigma' \cdot Id + \sigma\sigma' \cdot Id = 0, \end{aligned}$$

lo que prueba que es constante, y como ambos procesos inician en 0, debe anularse idénticamente. Por lo tanto concluimos que $\sigma \cdot B = M$ como se quería. □

Observación 2.14. Es importante notar que la extensión del espacio de probabilidad sólo fue necesaria para introducir a W, que a su vez sólo fue necesario para completar la variación cuadrática de un browniano por parte de B. Si supiéramos, por ejemplo, que σ es c.s. invertible en toda t fuera de un conjunto de medida de Lebesgue 0 (posiblemente aleatorio), entonces podríamos prescindir de W del todo y no sería necesario extender al espacio de probabilidad para definir a B. Además, una forma de verificar la condición sobre A y σ es probando que $[\sigma \cdot N] = A \cdot Id$ para algún movimiento browniano N en $\mathbb{R}^m$ (posiblemente en alguna extensión del espacio de probabilidad). Esta forma de verificar la condición es mucho más sencilla de visualizar cuando lidiamos con matrices en vez de vectores.

Algo que nos ayudará a traducir estos resultados para el caso matricial es tener algunas consideraciones que nos lleve del espacio de matrices $M_{d,d}$ a $\mathbb{R}^n$ con $n = d^2$. Como amos espacios tienen la misma dimensión, a cada elemento de $M_{d,d}$ le podemos asociar uno de $\mathbb{R}^n$ a través de la función vec, la cuál acomoda las columnas de una matriz en $M_{d,d}$ una encima de la otra, colocando iniciando por la izquierda. Esta función es entonces una isometría (bajo la norma euclidiana en $\mathbb{R}^n$ y la de Frobenius en $M_{d,d}$).

El teorema de Knight nos permite representar a martingalas locales continuas fuertemente ortogonales (con covariación nula) como movimientos brownianos independientes cambiados de tiempo. El siguiente resultado nos permitirá extender ligeramente ese resultado, pero primero debemos lidiar con algunas consideraciones para obtener una función medible π_d que a cada $A \in S_d$ le asigne alguna matriz $\pi_d(A)$ tal que $\pi_d(A) A \pi_d(A)' = \Lambda_d(A)$, la matriz diagonal con los valores propios de A ordenados de forma no decreciente.

Una vez que tengamos eso, para un arreglo $A \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$, le podemos asignar una matriz en $M_{n,n}$ donde $n = d^2$ (también denotada por A) de tal forma que para todo $i, j, k, l \in \{1, \ldots, d\}$, su coordenada $(i + (j-1)d, k + (l-1)d)$ está dada por A_{ijkl}. Si esta nueva matriz es simétrica o incluso un elemento de $\bar{S}_n^+$ (por ejemplo, A podría ser el proceso de covariación cuadrática de una semimartingala en $M_{d,d}$), entonces podemos definir a $\Lambda_n(A)$ como antes. Si reacomodamos las entradas diagonales de $\Lambda_n(A)$ a través de vec^{-1} (la rama inversa que nos regresa matrices cuadradas), podemos obtener a $\Lambda_{d,d}(A) \in M_{d,d}$.

Consideremos $A \in \bar{S}_d^+$, entonces la resolvente $R(A, \zeta) = (A - \zeta I_d)^{-1}$ es conjuntamente continua en sus dos argumentos cuando hacemos $R(A, \zeta) = \Delta$ (el punto al infinito en $\bar{S}_d^+$, donde la continuidad se da con respecto a la topología del espacio compactificado). Así, para cada valor propio λ, podemos calcular las proyecciones propias (proyecciones al espacio de vectores propios asociados al valor propio correspondiente) a través de la integral $\int_{\Gamma_\lambda} R(A, \zeta)\, d\zeta$, donde Γ_λ se puede tomar como un pequeño círculo en $\mathbb{C}$ que encierre a λ y no contenga a ningún otro valor propio. Luego, podemos considerar a los subespacios $T_n \subset \bar{S}_d^+$ donde tenemos exactamente $1 \leq r \leq d$ valores propios distintos ordenados crecientemente $\mu = (\mu_1, \ldots, \mu_r)$ con multiplicidades $n = (n_1, \ldots, n_r)$ (donde $n_1 + \ldots + n_r = d$)

respectivamente.

Como en [13], podemos ver que si nos restringimos a cada T_n, entonces podemos tomar un mapeo $A \mapsto P$ tal que $PAP' = diag\,(\lambda_1, \ldots, \lambda_d)$ de manera localmente continua. Podemos tomar una colección de subconjuntos $\{T_n^m\}$ de T_n tal que cada T_n^m es compacto en T_n, su unión es todo T_n y sus interiores son ajenos. En cada uno de dichos compactos, la continuidad de los valores propios como función de A, nos dice que de hecho tienen gráficas compactas y por lo tanto están a distancia positiva unas de otras. Luego, existe $\delta_m > 0$ tal que podemos tomar a los círculos Γ_λ (obviamente dependientes de A) centrados en λ y de radio δ_m de forma que estos no se toquen para ninguna A y los cuáles van a variar continuamente respecto de A.

Consecuentemente, la continuidad de la resolvente en la unión de las gráficas de los Γ_λ prueba que podemos conseguir a las proyecciones propias como funciones continuas de A. Por lo tanto podemos construir un mapeo $f_n^m : A \mapsto P$ continuo. Definiendo $f_n = f_n^m$ en $T_n^m \setminus \bigcup_{k \le m} T_n^k$, obtenemos el mapeo medible con las condiciones deseadas en T_n. Definiendo finalmente $\pi_d = f_n$ en T_n, obtenemos una rama el mapeo $A \mapsto P$ como función medible.

Este mismo resultado se puede probar de una manera más limpia usando fuertes resultados de geometría algebraica real. Con esto podríamos definir a π_d de manera que sea, no sólo medible, sino además semialgebraica. De hecho, para notar esto, basta ver que $\bar{S}_d^+$ es un subconjunto semialgebraico del espacio vectorial de matrices simétricas. Podemos definir al conjunto

$$\mathscr{P} = \left\{ (S, \lambda, v) \in \bar{S}_d^+ \times \mathbb{R} \times \mathbb{R}^d \ : \ Sv = \lambda v, |v| = 1 \right\},$$

el cuál también es semialgebraico y tiene una proyección natural $\pi : \mathscr{P} \to \bar{S}_d^+$. Luego, un resultado no trivial de geometría algebraica muestra que $\bar{S}_d^+$ admite una estratificación cuyos estratos son variedades topológicas semialgebraicas contráctiles $X_1, \ldots, X_N \subset \bar{S}_d^+$ tales que el mapeo inducido $\pi^{-1}(X_k) \to X_k$ es localmente un haz fibrado trivial. El fibrado de π en una matriz específica A, consiste en una unión de esferas disjuntas, a saber, las esferas unitarias de los espacios vectoriales asociados a cada valor propio de S. Estos haces fibrados son triviales puesto que tienen una base contráctil. De esta forma obtenemos justamente diagonalizaciones semialgebraicas, en particular, medibles. Al lector interesado con esta teoría se le refiere a [19].

De ahora en adelante, cuando tengamos un proceso X en $\mathbb{R}^n$ y un proceso Y diagonal en $\bar{S}_n^+$ (o bien, un proceso en $\mathbb{R}_+^n$), definimos al proceso $X \circ Y$ en $\mathbb{R}^n$ dado por $(X \circ Y)^i = (X^i \circ Y^{ii})$. Si por otro lado, ambos toman valores en $M_{n,n}$ y las entradas de Y son todas no negativas, entonces definimos al proceso $X \circ Y$ en $M_{n,n}$ dado por $(X \circ Y)^{ij} =$

$X^{ij} \circ Y^{ij}$.

Lema 2.15. *(representación como cambio de tiempo) Supongamos que M es una martingala local continua en $\mathbb{R}^n$ con variación cuadrática dada por $[M] = A \cdot V$ donde A es un proceso progresivamente medible en $\bar{S}_n^+$ y V es un proceso progresivamente medible con valores diagonales en $M_{n,n}$ cuyas entradas son no decrecientes. Entonces existe, posiblemente en una extensión del espacio de probabilidad, un movimiento browniano B en $\mathbb{R}^n$ tal que*

$$M = \pi_n (A)' \cdot (B \circ (\Lambda_n (A) \cdot V)),$$

donde $\Lambda_n (A) = \pi_n (A) A \pi_n (A)'$ es la matriz diagonal que contiene a los valores propios A (podemos suponer que están en orden ascendente de ser necesario).

Demostración. La idea principal es justamente transformar el problema a un contexto en el que podamos hacer uso del teorema de Knight. Considerando a π_n como antes, podemos ver que la medibilidad de π_n implica que $\pi_n (A)$ es progresivamente medible (y de hecho toma valores en el espacio de matrices ortogonales y por lo tanto es acotado). Entonces podemos definir a la martingala local continua

$$N = \pi_n (A) \cdot M,$$

y ver que

$$[N] = [\pi_n (A) \cdot M] = \pi_n (A) A \pi_n (A)' \cdot V = \Lambda_n (A) \cdot V.$$

Luego, las entradas de N son fuertemente ortogonales y podemos usar el teorema de Knight para encontrar, posiblemente en una extensión del espacio de probabilidad (dependiendo de si $\left[N^i\right]_\infty$ es infinito para toda $i = 1, \ldots, n$ o no), n movimientos brownianos independientes B^i tales que

$$N^i = B^i \circ \left[N^i\right], \quad i = 1, \ldots, n,$$

o lo que es equivalente, $N = B \circ [N]$. Dado que $\pi_n (A)' \pi_n (A) = I_n$, entonces $M = \pi_n (A)' \cdot N$ de donde se sigue que

$$M = \pi_n (A)' \cdot (B \circ [N]) = \pi_n (A)' \cdot (B \circ (\Lambda (A) \cdot Id)),$$

como se quería probar. □

Observación 2.16. Uno podría, en principio, definir a N como el producto $\pi_n (A) M$ (o con alguna expresión similar) para que podamos recuperar a M de una manera sencilla. No obstante, esto es poco útil, porque N no necesariamente sería una martingala local, y al calcular su variación cuadrática, tendríamos que usar la fórmula de integración por partes, donde no conocemos la relación directa entre $\pi_n (A)$ y M, ni sabemos si $\pi_n (A)$ es de variación finita en compactos o no.

Luego, la representación anterior no es la más explícita, sobre todo porque desconocemos la forma precisa de $\pi_n(A)$. No obstante, este puede resultar de utilidad porque nuestros problemas de EDE se pueden traducir a un contexto de cambio de tiempo, que en ocasiones es más fácil de resolver.

Por ejemplo, uno puede verificar que si $M = U \cdot W$ para algún proceso progresivamente medible $U \in L(W)$ y un movimiento browniano W en $\mathbb{R}^n$ (como en el lema anterior), entonces se cumplen las condiciones del presente lema. Así, si estuviera en duda la existencia de una solución débil para M, podríamos traducir este problema a encontrar una solución débil a la formulación equivalente respecto de cambios de tiempo.

De hecho, como en el caso del lema 2.13, el principal uso que le daremos al lema es cuando $V = I_n Id$. Además, es fácil notar que los valores propios se encargarán de «eliminar» toda información en exceso, es decir que no usaremos a todos los movimientos brownianos, sino sólo los necesarios.

Por ejemplo, en el caso de matrices simétricas en S_d, a lo más requerimos $\frac{d(d+1)}{2}$ movimientos brownianos distintos, porque habrá $\frac{d(d-1)}{2}$ valores propios idénticamente nulos. Esto también complicará ligeramente los cálculos en el modelo de estimación que se presenta en el capítulo que sigue.

Lema 2.17. *Sea $2p \in \mathbb{N}$, consideremos $S \sim BM_{2p,d}$, $s \in M_{2p,d}$, $Q \in M_{d,d}$ y definamos $X = (SQ + s)'(SQ + s)$. Entonces X satisface*

$$dX_t = Q' dS_t' (S_t Q + s) + (Q' S_t' + s') dS_t Q + 2pQ'Qdt, \quad X_0 = x = s's.$$

Más aún, existe $B \sim BM_{d,d}$, posiblemente en una extensión del espacio de probabilidad, tal que

$$dX_t = \sqrt{X_t} dB_t Q + Q' dB_t' \sqrt{X_t} + 2p\alpha dt,$$

donde $\alpha = Q'Q$.

Demostración. Veamos que la primera fórmula se obtiene de usar la fórmula de integración por partes y el que

$$Q' dS_t' dS_t Q = 2pQ' I_d Q dt = 2pQ'Qdt.$$

Para probar la segunda EDE, vamos a usar el lema 2.13, para lo cuál basta calcular la variación cuadrática de X y probar que coincide con la de $Y_t = \int_0^t \left(\sqrt{X_s} dB_s Q + Q' dB_s' \sqrt{X_s}\right)$ para algún $B \sim MB_{d,d}$.

En este sentido, estamos considerando al mapeo lineal y simétrico (también referido como la simetrización $S_{\sqrt{X_t},Q}$ en el lema 2.2)

$$\sigma_t : A \mapsto \sqrt{X}_t AQ + Q'A'\sqrt{X}_t,$$

como una matriz en $M_{n,n}$ multiplicando a un vector $A \in \mathbb{R}^n$ para $n = d^2$. σ es continuo y adaptado porque $\sqrt{X}$ lo es y gracias a la continuidad de la multiplicación matricial.

En este contexto, $Y = \sigma \cdot B$, donde B se piensa como en un movimiento browniano en $\mathbb{R}^n$ y sólo hace falta comparar la variación cuadrática de X (en realidad, de su componente martingala local) y Y. Así, para $1 \leq i, i', j, j' \leq d$, podemos calcular,

$$\begin{aligned}
& \left[Y^{ij}, Y^{i'j'}\right] \\
= & \sum_{k,l,k',l' \leq d} \left[\sqrt{X}^{ik} Q_{lj} \cdot B^{kl} + Q_{ki}\sqrt{X}^{lj} \cdot B^{lk}, \sqrt{X}^{i'k'} Q_{l'j'} \cdot B^{k'l'} + Q_{k'i'}\sqrt{X}^{l'j'} \cdot B^{l'k'}\right] \\
= & \sum_{k,l,k',l' \leq d} \sqrt{X}^{ik} Q_{lj} \sqrt{X}^{i'k'} Q_{l'j'} \cdot \left[B^{kl}, B^{k'l'}\right] + \sqrt{X}^{ik} Q_{lj} Q_{k'i'} \sqrt{X}^{l'j'} \cdot \left[B^{kl}, B^{l'k'}\right] \\
+ & Q_{ki}\sqrt{X}^{lj}\sqrt{X}^{i'k'} Q_{l'j'} \cdot \left[B^{lk}, B^{k'l'}\right] + Q_{ki}\sqrt{X}^{lj} Q_{k'i'}\sqrt{X}^{l'j'} \cdot \left[B^{lk}, B^{l'k'}\right].
\end{aligned}$$

Aquí debemos ahora separar por los que se anulan o no en los casos $(k, l) = (k', l')$ y $(k, l) = (l', k')$ y usar la simetría de $\sqrt{X}$, obtenemos finalmente

$$\begin{aligned}
& \left(\sum_{l \leq d} \sqrt{X}^{ik}\sqrt{X}^{ki'}\right)\left(\sum_{k \leq d} Q_{lj}Q_{lj'}\right) \cdot Id + \left(\sum_{l \leq d} \sqrt{X}^{jl}\sqrt{X}^{lj'}\right)\left(\sum_{k \leq d} Q_{ki}Q_{ki'}\right) \cdot Id \\
+ & \left(\sum_{l \leq d} \sqrt{X}^{jl}\sqrt{X}^{li'}\right)\left(\sum_{k \leq d} Q_{ki}Q_{kj'}\right) \cdot Id + \left(\sum_{l \leq d} \sqrt{X}^{ik}\sqrt{X}^{kj'}\right)\left(\sum_{k \leq d} Q_{lj}Q_{li'}\right) \cdot Id \\
= & \left(X^{ii'}\left(Q'Q\right)_{jj'} + X^{ij'}\left(Q'Q\right)_{ji'} + X^{ji'}\left(Q'Q\right)_{ij'} + X^{jj'}\left(Q'Q\right)_{ii'}\right) \cdot Id.
\end{aligned}$$

Por otro lado, de la primera EDE se puede deducir fácilmente que para $1 \leq i, i', j, j' \leq d$,

$$\begin{aligned}
& d\left[X^{ij}, X^{i'j'}\right]_t = dX_t^{ij} dX_t^{i'j'} \\
= & \left(Q'dS_t'\left(S_tQ + s\right) + \left(Q'S_t' + s'\right)dS_tQ\right)_{ij} \\
\times & \left(Q'dS_t'\left(S_tQ + s\right) + \left(Q'S_t' + s'\right)dS_tQ\right)_{i'j'} \\
= & \sum_{l,l' \leq 2p} \sum_{m,m' \leq d} \left(Q_{mi} dS_t^{lm}\left(S_tQ + s\right)_{lj} + \left(S_tQ + s\right)_{li} dS_t^{lm} Q_{mj}\right) \\
\times & \left(Q_{m'i'} dS_t^{l'm'}\left(S_tQ + s\right)_{l'j'} + \left(S_tQ + s\right)_{l'i'} dS_t^{l'm'} Q_{m'j'}\right).
\end{aligned}$$

Ahora notemos que los sumandos no se anulan si y sólo si $(m,l)=(m',l')$, obteniendo

$$\begin{aligned}
& \sum_{l=1}^{2p}\sum_{m=1}^{d}\left(Q_t^{mi}\left(\int_0^t dS_uQ_u+s\right)_{lj}Q_t^{mi'}\left(\int_0^t dS_uQ_u+s\right)_{lj'}\right.\\
+ \quad & Q_t^{mi}\left(\int_0^t dS_uQ_u+s\right)_{lj}\left(\int_0^t dS_uQ_u+s\right)_{li'}Q_t^{mj'}\\
+ \quad & \left(\int_0^t dS_uQ_u+s\right)_{li}Q_t^{mj}Q_t^{mi'}\left(\int_0^t dS_uQ_u+s\right)_{lj'}\\
+ \quad & \left.\left(\int_0^t dS_uQ_u+s\right)_{li}Q_t^{mj}\left(\int_0^t dS_uQ_u+s\right)_{li'}Q_t^{mj'}\right)dt\\
= \quad & \sum_{l=1}^{2p}\sum_{m=1}^{d}\left(\left(Q_t'Q_t\right)^{ii'}Z_t^{jj'}+\left(Q_t'Q_t\right)^{ij'}Z_t^{ji'}+\left(Q_t'Q_t\right)^{ji'}Z_t^{ij'}+\left(Q_t'Q_t\right)^{jj'}Z_t^{ii'}\right)dt.
\end{aligned}$$

como queríamos probar. Esto ya nos permite usar el lema 2.13 y concluir la prueba. □

Observación 2.18. Notemos que el lema 2.13 nos provee la existencia de un movimiento browniano B en $\mathbb{R}^n$ que cumple nuestras condiciones. Justamente habría que reagrupar sus entradas hasta poder considerarlo como una matriz en $M_{d,d}$. En relación a la observación del mismo lema, es posible ver que si $2p>d-1$, entonces c.s. X es invertible para todo $t>0$ (fuera de un conjunto de medida de Lebesgue 0 si $2p=d-1$). Si Q es además invertible, entonces no hace falta extender el espacio de probabilidad para introducir a B. Mientras que si $2p<d$, en general, la extensión del espacio de probabilidad es inevitable.

Del mismo modo en que generalizamos la construcción de procesos cuadrados de Bessel, podemos construir al proceso de Cox-Ingersoll-Ross a través de la suma de los cuadrados de procesos de Ornstein-Uhlenbeck reales independientes. Luego, el siguiente paso para generalizar esta noción es considerar el caso matricial. Denotaremos de ahora en adelante denotamos la distribución del proceso CIR por $CIR\,(a,b,\sigma,x)$ donde a es el parámetro de reversión a la media, b es la media, $\sigma\sqrt{\cdot}$ es el coeficiente de difusión y tiene estado inicial x, mientras que denotamos por $BESQ\,(p,x)$ a la distribución de un proceso cuadrado de Bessel con deriva p y estado inicial es x

Definición 2.19. *(proceso de Ornstein-Uhlenbeck)* Decimos que un proceso estocástico S con valores en $M_{n,m}$, es un proceso de Ornstein-Uhlenbeck matricial, y lo denotamos $S\sim OUP_{n,m}\,(A,\beta,s_0)$, si satisface

$$dS_t=S_t\beta dt+dW_tA,\quad S_0=s_0, \tag{2.3}$$

para algunas $A,\beta\in M_{m,m}$, $s_0\in M_{n,m}$ y $W\sim BM_{n,m}$.

La existencia y unicidad fuerte de los procesos de Ornstein-Uhlenbeck se sigue de los teoremas de Yamada-Watanabe para coeficientes globalmente Lipschitz; específicamente, se sigue del teorema 6.7.3 en [20]. Por completitud, enunciaremos el mismo junto con algunas definiciones auxiliares.

Definición 2.20. *(continuidad Lipschitz)* Sean U y V espacios normados y sea $W \subset U$ un abierto. Decimos que una función $f : W \to V$ es localmente Lipschitz si para todo $x \in W$ existe una vecindad $A \subset W$ de x y una constante $K > 0$ tal que

$$\|f(z) - f(y)\|_V \leq K \|z - y\|_U, \quad \forall z, y \in A.$$

Si en particular se puede tomar K no dependiente de x, entonces se dice que f es gobalmente Lipschitz.

Teorema 2.21. *(Yamada-Watanabe multidimensional, Stelzer)* Sea U un subconjunto abierto de $M_{n,m}$ tal que existe una sucesión creciente $\{U_n\}$ de subconjuntos cerrados y convexos de U tales que $\bigcup_n U_n = U$. Si $f : U \to M_{n,k}$ es localmente Lipschitz y Z es una semimartingala continua en $M_{k,m}$. Entonces para todo valor inicial $X_0 \in U$ que sea $\mathcal{F}_0$-medible, existe un tiempo de paro $T > 0$ c.s. y una semimartingala X en U con valor inicial X_0 que es trayectorialmente la única solución fuerte a la EDE

$$dX_t = f(X_t)\, dZ_t, \tag{2.4}$$

en el intervalo aleatorio $[0, T)$. En el evento $T < \infty$ se tiene que $X_t \to \partial U$ conforme $t \to T$ o hay explosión, i.e., $\limsup_{t<T} \|X_t\| = \infty$. Si f satisface

$$\|f(x)\|^2 \leq K\left(1 + \|x\|^2\right), \tag{2.5}$$

para alguna constante $K > 0$, entonces no hay explosiones.

Aunque, como mencionamos antes, la existencia fuerte de procesos Ornstein-Uhlenbeck se sigue del teorema anterior, en el caso particular de estos procesos, podemos dar una construcción explícita de X en términos de sus parámetros y de W. Este resultado es casi idéntico a su análogo escalar.

Teorema 2.22. *Para cada $W \sim BM_{n,m}$, existe una única solución fuerte a (2.3) dada por*

$$S_t = s_0 e^{\beta t} + \left(\int_0^t dW_s A e^{-\beta s}\right) e^{\beta t}. \tag{2.6}$$

Demostración. Definiendo a X de acuerdo a la ecuación anterior y usando la fórmula de integración por partes, tenemos que

$$\begin{aligned} dS_t &= s_0 e^{\beta t}\beta dt + \left(\int_0^t dW_s A e^{-\beta s}\right) e^{\beta t}\beta dt + \left(dW_t A e^{-\beta t}\right) e^{\beta t} \\ &+ \left(dW_t A e^{-\beta t}\right) e^{\beta t}\beta dt = S_t \beta dt + dW_t A, \end{aligned}$$

Esto prueba la existencia fuerte, mientras que la evidente continuidad Lipschitz global prueba la unicidad y la condición de crecimiento (2.5) se satisface y nos garantiza la ausencia de explosiones. Aquí, implícitamente usamos parte del teorema 2.21, con $f(x) = (x, I_d)$ y $Z_t = (\beta t, W_t A)'$. □

A continuación daremos una construcción de procesos de Wishart a través de procesos de Ornstein-Uhlenbeck matriciales. La razón por la que aún consideramos a estos como procesos es Wishart es porque la distribución marginal de S_t es normal multivariada, cuya prueba es bastante sencilla.

Lema 2.23. *Sea $B \sim BM_{n,m}$ y consideremos un proceso determinista y cuadrado integrable Y con valores en $M_{m,k}$, entonces el proceso*

$$H_t = \int_0^t dB_s Y_s,$$

tiene renglones normales independientes e idénticamente distribuidas con media cero y

$$E\left(H_t^{li} H_t^{lj}\right) = \int_0^t \left(Y_s' Y_s\right)^{ij} ds.$$

Demostración. La construcción de la integral de Itô para el movimiento browniano prueba que H_t es el límite en L^2 de combinaciones lineales de las entradas de B a diferentes tiempos y por lo tanto, H tiene distribuciones marginales gaussianas.

Por otro lado, la independencia e igualdad distribucional de los renglones de H se sigue de que

$$H^{ij} = \sum_{l=1}^{m} \int_0^t dB_s^{il} Y_s^{lj},$$

y por lo tanto los movimientos brownianos involucrados con $\left(H^{ij}\right)_j$ y con $\left(H^{i'j}\right)_j$ son distintos para $i \neq i'$. Finalmente tenemos

$$\begin{aligned} E\left(H_t^{1i} H_t^{1j}\right) &= \sum_{l_1, l_2=1}^{m} E\left(\int_0^t dB_s^{1l_1} Y_s^{l_1 i}, \int_0^t dB_s^{1l_2} Y_s^{l_2 j}\right) \\ &= \sum_{l=1}^{m} \int_0^t Y_s^{li} Y_s^{lj} ds = \int_0^t \left(Y_s' Y_s\right)^{ij} ds, \end{aligned}$$

lo que termina la prueba. □

Observación 2.24. El resultado se puede escribir como

$$H_t \sim N_{n,k}\left(0, \int_0^t Y_s' Y_s ds \otimes I_n\right),$$

donde $\otimes$ denota el producto de Kronecker y $N_{n,k}$ denota a la distribución normal multivariada matricial en $M_{n,k}$, esto es, $A \sim N_{n,k}(\mu, \Sigma)$ si y sólo si A toma valores en $M_{n,k}$, $\operatorname{vec}(A)$ tiene distribución normal con esperanza μ y

$$E\left(\operatorname{vec}(H_t)\operatorname{vec}(H_t)'\right) = \Sigma.$$

Luego, esta nueva escritura se debe a que

$$\begin{aligned} E\left(\left(\operatorname{vec}(H_t)\operatorname{vec}(H_t)'\right)_{v(k,i),v(l,j)}\right) &= E\left(H_t^{lj} H_t^{ki}\right) = \delta_{kl}\left(\int_0^t Y_s' Y_s ds\right)^{ij} \\ &= \left(\int_0^t Y_s' Y_s ds \otimes I_n\right)_{v(k,i),v(l,j)}. \end{aligned}$$

Corolario 2.25. *Sea $S \sim OUP_{n,m}(A, \beta, s_0)$, entonces*

$$S_t \sim N_{n,k}\left(s_0 e^{\beta t}, e^{\beta' t}\left(\int_0^t e^{-\beta' s} A' A e^{-\beta s} ds\right) e^{\beta t} \otimes I_n\right).$$

Demostración. El lema anterior aplicado a $Ae^{-\beta Id}$ nos da el resultado casi directamente, puesto que sólo requerimos perturbar esto por $e^{\beta Id}$ y luego sumar la nueva media. Lo más resaltable de esto es que tenemos otra representación en ciertos escenarios; sea

$$\begin{aligned} \Upsilon : X &\mapsto \beta' X + X\beta \\ \operatorname{vec}(\Upsilon(X)) &= (I_d \otimes \beta' + \beta' \otimes I_d)\operatorname{vec}(X) \\ &= (\beta' \oplus \beta')\operatorname{vec}(X). \end{aligned}$$

Supongamos que Υ es invertible, lo que es equivalente a que $0 \notin \sigma(\beta) + \sigma(\beta) = \sigma(\beta' \oplus \beta')$. Entonces tenemos

$$\frac{d}{dt}\left(e^{\beta' t} A' A e^{\beta t}\right) = \Upsilon\left(e^{\beta' t} A' A e^{\beta t}\right),$$

lo que implica que

$$\begin{aligned} e^{\beta' t}\left(\int_0^t e^{-\beta' s} A' A e^{-\beta s} ds\right) e^{\beta t} &= \int_0^t e^{\beta' s} A' A e^{\beta s} ds = \Upsilon^{-1}\left(e^{\beta' s} A' A e^{\beta s}\right)\Big|_0^t \\ &= \Upsilon^{-1}\left(e^{\beta' t} A' A e^{\beta t}\right) - \Upsilon^{-1}(A'A), \end{aligned}$$

por lo que

$$S_t \sim N_{n,k}\left(s_0 e^{\beta t}, I_n \otimes \left[\Upsilon^{-1}\left(e^{\beta' t} A' A e^{\beta t}\right) - \Upsilon^{-1}\left(A'A\right)\right]\right),$$

como queríamos probar. □

Con esto, justo tenemos distribuciones normales a cada tiempo y es con estas distribuciones que inicialmente construimos a las distribuciones Wishart, lo que motiva el siguiente resultado.

Lema 2.26. *Sea $2p$ un entero no negativo, $Q, \beta \in M_{d,d}$, $x \in \bar{S}_d^+$ y $S \sim OUP_{2p,d}(Q, \beta', s)$ donde $s's = x$. Entonces existe $B \sim BM_{d,d}$, posiblemente en una extensión del espacio de probabilidad, tal que $X = S'S$ satisface*

$$dX_t = \sqrt{X_t} dB_t Q + Q' dB_t' \sqrt{X_t} + \left(\beta X_t + X_t \beta' + 2p\alpha\right) dt, \quad X_0 = x,$$

donde $\alpha = Q'Q$.

Demostración. Supongamos que S satisface (2.3) con respecto a $W \sim BM_{2p,d}$. Usando la fórmula de integración por partes, podemos verificar que

$$\begin{aligned} dX_t &= dS_t' S_t + S_t' dS_t + dS_t' dS_t \\ &= \left(\beta S_t' dt + Q' dW_t'\right) S_t + S_t'\left(S_t \beta' dt + dW_t Q\right) + \left(Q' dW_t'\right)\left(dW_t Q\right) \\ &= \beta X_t dt + Q' dW_t' S_t + X_t \beta' dt + S_t' dW_t Q + 2pQ'Q dt, \end{aligned}$$

donde $Q'Q = \alpha$. Por lo tanto, la parte de variación finita en compactos ya tiene la forma deseada y basta ver que la variación cuadrática de X satisface las condiciones del lema 2.13, i.e., coincide con la de $Y_t = \int_0^t \left(\sqrt{X_s} dB_s Q + Q' dB_s' \sqrt{X_s}\right)$ para algún $B \sim MB_{d,d}$. Ya hemos calculado esta última en la prueba del lema 2.17, así que sólo hace falta calcular la de X:

$$\begin{aligned} & d\left[X^{ij}, X^{i'j'}\right]_t = dX_t^{ij} dX_t^{i'j'} = \left(Q' dW_t' S_t\right)_{ij} \left(S_t' dW_t Q\right)_{i'j'} \\ = & \sum_{l,l' \leq 2p} \sum_{m,m' \leq d} \left(Q_{mi} dW_t^{lm} S_t^{lj} + S_t^{li} dW_t^{lm} Q_{mj}\right)\left(Q_{m'i'} dW_t^{l'm'} S_t^{l'j'} + S_t^{l'i'} dW_t^{l'm'} Q_{m'j'}\right), \end{aligned}$$

donde los únicos términos que no se cancelan son aquellos donde $(l, m) = (l', m')$. De este modo se simplifica a

$$\begin{aligned} & \sum_{l \leq 2p} \sum_{m \leq d} \left(Q_{mi} S_t^{lj} Q_{mi'} S_t^{lj'} + S_t^{li} Q_{mj} Q_{mi'} S_t^{lj'} + Q_{mi} S_t^{lj} S_t^{li'} Q_{mj'} + S_t^{li} Q_{mj} S_t^{li'} Q_{mj'}\right) dt \\ = & \left(X^{jj'} \left(Q'Q\right)_{ii'} + X^{ij'} \left(Q'Q\right)_{ji'} + X^{ji'} \left(Q'Q\right)_{ij'} + X^{ii'} \left(Q'Q\right)_{jj'}\right) dt, \end{aligned}$$

entonces las variaciones cuadráticas coinciden y podemos usar el lema 2.13 para finalizar la prueba. $\square$

Observación 2.27. De la misma manera que en el caso anterior, si Q es invertible y $2p \geq d$, entonces no es necesario extender el espacio de probabilidad para definir a B; de lo contrario, en general es inevitable. Además, aunque parece no haber ninguna restricción sobre los parámetros salvo por $2p \in \mathbb{N}$, veamos que como $x = s's$, entonces el rango de x está acotado por $2p$. Esta última aclaración es importante para la prueba de existencia de distribuciones Wishart, puesto que el parámetro de no centralidad cumple justamente una condición de rango.

Aunque la motivación de la siguiente construcción vendrá por parte del teorema 2.42, podemos probar el resultado principal y ver desde otro enfoque constructivo al proceso de Wishart. Aquí, lo que planeamos hacer es iniciar con la idea del lema 2.17, donde construimos una primera versión del proceso de Wishart, y agregar un factor de mezcla que nos llevará a obtener una semimartingala con las mismas propiedades que la obtenida en el lema 2.26.

Lo más importante de esta construcción es que abre el camino a procesos más generales, usando las mismas ideas que en el caso de los procesos de Wishart. De hecho, más adelante veremos las características markovianas del proceso de Wishart y en particular encontraremos su función de transición. Con esta a su vez se pueden construir nuevos procesos de Markov con transiciones Wishart (homogéneos o no), haciendo uso de la caracterización de la existencia de las distribuciones Wishart.

En cierto sentido esto se convierte en un análogo al uso que se dan en las transiciones brownianas (que son normales o gaussianas) para construir otros procesos gaussianos de Markov.

Lema 2.28. *Sea $2p \in \mathbb{N}$ y consideremos $S \sim BM_{2p,d}$, $s \in M_{2p,d}$, y dos procesos con valores en $M_{d,d}$: $\beta \in L(Id)$ y $Q \in rL(S)$. Si definimos,*

$$Z_t = \left(\int_0^t dS_u Q_u + s\right)' \left(\int_0^t dS_u Q_u + s\right),$$

entonces existe $W \sim BM_{d,d}$, posiblemente en una extensión del espacio de probabilidad, tal que $Z_0 = x := s's$ y para toda $t \geq 0$ se tiene

$$dZ_t = \sqrt{Z_t} dW_t Q_t + Q_t' dW_t' \sqrt{Z_t} + 2p Q_t' Q_t dt.$$

Más aún, cada vez que se tiene la EDE anterior, si definimos $X = e^{\beta \cdot Id} Z e^{\beta' \cdot Id}$, entonces existe otro movimiento browniano $B \sim BM_{d,d}$, que también puede estar definido en una *extensión del espacio de probabilidad, tal que para toda t tenemos*

$$dX_t = \sqrt{X_t}dB_t\delta_t + \delta_t' dB_t'\sqrt{X_t} + \left(2p\alpha_t + \beta_t X_t + X_t\beta_t'\right) dt,$$

donde $\alpha = e^{\beta\cdot Id}Q'Qe^{\beta'\cdot Id}$ y δ satisface $\delta'\delta = \alpha$, por ejemplo ,$\delta = \sqrt{\alpha}$. En particular, si β es constante y Q satisface $Q'Q = e^{-\beta Id}\alpha e^{-\beta' Id}$ para alguna $\alpha \in \bar{S}_d^+$ constante, por ejemplo, $Q = \sqrt{e^{-\beta Id}\alpha e^{-\beta' Id}}$, entonces $Q_0'Q_0 = \delta'\delta = \alpha$ y se satisface

$$dX_t = \sqrt{X_t}dB_tQ_0 + Q_0'dB_t'\sqrt{X_t} + \left(2p\alpha + \beta X_t + X_t\beta'\right) dt,$$

para toda t y $X_0 = x$.

Demostración. Procedamos como en el lema 2.17 para ver que al usar la fórmula de integración por partes y el que

$$Q_t'dS_t'dS_tQ_t = 2pQ_t'I_dQ_tdt = 2pQ_t'Q_tdt,$$

obtenemos que la deriva de Z es $2pQ'Q \cdot Id$. De este mismo modo podemos calcular su variación y covariación cuadrática, la cual satisface las siguientes identidades para cualquier elección de índices $1 \leq i, i', j, j' \leq d$,

$$\begin{aligned}
& d\left[Z^{ij}, Z^{i'j'}\right]_t = dZ_t^{ij}dZ_t^{i'j'} \\
= \quad & \left(Q_t'dS_t'\left(\int_0^t dS_uQ_u + s\right) + \left(\int_0^t Q_u'dS_u' + s'\right)dS_tQ_t\right)_{ij} \\
& \left(Q_t'dS_t'\left(\int_0^t dS_uQ_u + s\right) + \left(\int_0^t Q_u'dS_u' + s'\right)dS_tQ_t\right)_{i'j'}, \\
= \quad & \sum_{l,l'=1}^{2p}\sum_{m,m'=1}^{d}\left(Q_t^{mi}dS_t^{lm}\left(\int_0^t dS_uQ_u + s\right)_{lj} + \left(\int_0^t dS_uQ_u + s\right)_{li}dS_t^{lm}Q_t^{mj}\right) \\
& \left(Q_t^{m'i'}dS_t^{l'm'}\left(\int_0^t dS_uQ_u + s\right)_{l'j'} + \left(\int_0^t dS_uQ_u + s\right)_{l'i'}dS_t^{l'm'}Q_t^{m'j'}\right).
\end{aligned}$$

Ahora notemos que los sumandos no se anulan si y sólo si $(m, l) = (m', l')$, por lo tanto se simplifica a

$$\sum_{l=1}^{2p}\sum_{m=1}^{d}\left(\left(Q_t'Q_t\right)^{ii'}Z_t^{jj'} + \left(Q_t'Q_t\right)^{ij'}Z_t^{ji'} + \left(Q_t'Q_t\right)^{ji'}Z_t^{ij'} + \left(Q_t'Q_t\right)^{jj'}Z_t^{ii'}\right)dt.$$

Esto nos permite usar la representación integral de Doob (lema 2.13) para introducir un movimiento browniano W en $M_{d,d}$ tal que

$$dZ_t = \sqrt{Z_t}dW_tQ_t + Q_t'dW_t'\sqrt{Z_t} + 2pQ_t'Q_tdt.$$

Ahora supongamos que se tiene la EDE anterior para alguna combinación de parámetros (p, Q) y definamos a la martingala $N_u = \int_0^u \sqrt{Z_t} dW_t Q_t + \int_0^u Q_t' dW_t' \sqrt{Z_t}$ y usemos la fórmula de integración por partes para obtener

$$\begin{aligned} dX_t &= e^{\beta t}\left(dN_t + 2pQ_t'Q_t dt\right) e^{\beta' t} + \left(\beta e^{\beta t} Z_t e^{\beta' t} + e^{\beta t} Z_t e^{\beta' t}\beta'\right) dt \\ &= e^{\beta t} dN_t e^{\beta' t} + \left(2p\alpha_t + \beta X_t + X_t \beta'\right) dt, \end{aligned}$$

y sólo resta calcular la variación cuadrática de X para corroborar la legitimidad del uso de la representación integral de Doob. Esto es muy simple porque para todo $1 \leq i, i', j, j' \leq d$,

$$\begin{aligned} d\left[X^{ij}, X^{i'j'}\right]_t &= \sum_{klk'l'} \left(e^{\beta t}\right)_{ik} dN_t^{kl} \left(e^{\beta' t}\right)_{lj} \left(e^{\beta t}\right)_{i'k'} dN_t^{k'l'} (e)_{l'j'} \\ &= \sum_{klk'l'} \left(e^{\beta t}\right)_{ik} \left(e^{\beta' t}\right)_{lj} \left(e^{\beta t}\right)_{i'k'} \left(e^{\beta' t}\right)_{l'j'} \left(Z_t^{kk'} \left(Q_t'Q_t\right)^{ll'}\right. \\ &+ \left. Z_t^{kl'} \left(Q_t'Q_t\right)^{lk'} + Z_t^{lk'} \left(Q_t'Q_t\right)^{kl'} + Z_t^{ll'} \left(Q_t'Q_t\right)^{kk'}\right) dt \\ &= \left(X_t^{ii'}\alpha_t^{jj'} + X_t^{ij'}\alpha_t^{ji'} + X_t^{ji'}\alpha_t^{ij'} + X_t^{jj'}\alpha_t^{ii'}\right) \cdot Id, \end{aligned}$$

como queríamos probar. La última afirmación se sigue de esto último, notando que en ese caso α es constante y Q_0 es cualquier matriz que satisfaga $Q_0'Q_0 = \alpha$. $\square$

Observación 2.29. Es importante notar cómo esta idea transforma el problema de resolver el problema de una EDE con deriva linealmente dependiente del estado actual en otro con deriva determinista y con carácter espiral (divergente, convergente o periódico) de acuerdo al comportamiento de $e^{\beta Id}$. Esto se debe a la biyección que se genera entre X y Z, teniendo esta propiedad incluso cuando $2p$ no es entero. También podríamos considerar el resolver una EDE para X, donde debemos dar por adelantado el proceso α, tomar $Q = \sqrt{e^{-\beta \cdot Id}\alpha e^{-\beta' \cdot Id}}$ y resolver la EDE correspondiente para Z.

2.3. Semimartingalas y EDE de Wishart

Uno se puede preguntar sobre la existencia de procesos que satisfagan la EDE del lema anterior y sobre qué tanto más se pueden generalizar los parámetros involucrados. También puede ser que, para evitarnos problemas respecto de soluciones débiles contra soluciones fuertes, podríamos tratar de extraer lo más importante del proceso.

Esto es, dar una definición de procesos de Wishart a omitiendo a los movimientos brownianos en la medida de lo posible y en vista del lema 2.13. Con esto, lograremos poder concenctrarnos en propiedades intrínsecas de estos procesos, sobre todo aquellas de particular utilidad en $\bar{S}_d^+$. Esto a su vez nos llevará a poder caracterizar mejor su comportamiento, en otras secciones.

Esto nos lleva a la introducción de las siguientes definiciones y caracterizaciones, como herramientas para lidiar con un problema nada trivial. La razón por la que estos procesos son llamados «de Wishart», es porque su construcción (las presentadas anteriormente) son un análogo, en procesos, a la construcción de las distribuciones de Wishart. Además, estos procesos presentan distribuciones marginales Wishart, por lo que es aún más natural este nombre.

Definición 2.30. *(semimartingala de Wishart)* Para $p \geq 0$, $\beta \in M_{d,d}$ y $\alpha \in \bar{S}_d^+$, una semimartingala continua X con valores en $\bar{S}_d^+$ se dice semimartingala de Wishart con parámetros (p, α, β, x) si podemos escribir $X = x + D + M$ donde $X_0 = x$,

$$D = \left(2p\alpha + \beta X + X\beta'\right) \cdot Id,$$

y M es una martingala local con variación cuadrática

$$\left[M^{ij}, M^{kl}\right] = \left(X^{ik}\alpha_{jl} + X^{il}\alpha_{jk} + X^{jk}\alpha_{il} + X^{jl}\alpha_{ik}\right) \cdot Id, \quad 0 \leq i, j, k, l \leq d. \tag{2.7}$$

En dicho caso, lo denotamos por $X \sim WP_d\left(p, \alpha, \beta, x\right)$.

Es inmediato que tanto D como M son continuas y que ambas inician en 0. Además, podemos observar que la deriva, tiene un comportamiento afín con respecto al estado actual. A continuación veremos un pequeño resultado sobre la aditividad de estos procesos.

Lema 2.31. *Sean $X^i \sim WP_d\left(p_i, \alpha, \beta, x_i\right)$ para $i = 1, \ldots, n$ fuertemente ortogonales, i.e.,*

$$\left\langle \left(X^i\right)^{kl}, \left(X^i\right)^{k'l'} \right\rangle = 0,$$

para todo $i \leq n$, $k, l, k', l' \leq d$. Entonces $X = \sum_{i=1}^n X^i \sim WP_d\left(\sum_{i=1}^n p_i, \alpha, \beta, \sum_{i=1}^n x_i\right)$.

Demostración. Notemos que $X_0 = \sum_{i=1}^n X_0^i = \sum_{i=1}^n x_i$ mientras que

$$\begin{aligned} D &:= \sum_{i=1}^n D^i = \sum_{i=1}^n \left(2p_i\alpha + \beta X^i + X^i\beta'\right) \cdot Id \\ &= \left(2\left(\sum_{i=1}^n p_i\right)\alpha + \beta X + X\beta'\right) \cdot Id, \end{aligned}$$

define un proceso continuo de variación finita en compactos, donde $X^i = D^i + M^i + x_i$es la descomposición de X^i para $i \leq n$. Si ahora tomamos a la martingala local continua

$M=\sum_{i=1}^{n}M^{i}$, entonces podemos verificar por la ortogonalidad fuerte que,

$$\begin{aligned}\left[M^{kl},M^{k'l'}\right] &= \left[\sum_{i=1}^{n}\left(M^{i}\right)^{kl},\sum_{i=1}^{n}\left(M^{i}\right)^{k'l'}\right]=\sum_{i=1}^{n}\left[\left(M^{i}\right)^{kl},\left(M^{i}\right)^{k'l'}\right]\\ &= \sum_{i=1}^{n}\left(\left(X^{i}\right)^{kk'}\alpha_{ll'}+\left(X^{i}\right)^{lk'}\alpha_{kl'}+\left(X^{i}\right)^{kl'}\alpha_{lk'}+\left(X^{i}\right)^{ll'}\alpha_{kk'}\right)\cdot Id\\ &= \left(X^{kk'}\alpha_{ll'}+X^{lk'}\alpha_{kl'}+X^{kl'}\alpha_{lk'}+X^{ll'}\alpha_{kk'}\right)\cdot Id \quad 0\leq i,j,k,l\leq d,\end{aligned}$$

con lo que concluye la prueba. □

Observación 2.32. Una de las aplicaciones más importantes es cuando los procesos son independientes por parejas, donde trivialmente se satisface la hipótesis de ortogonalidad.

Definición 2.33. *(EDE de Wishart)* Sean $Q,\beta\in M_{d,d}$, $p\geq 0$ y $x\in\bar{S}_{d}^{+}$, entonces definimos la ecuación diferencial estocástica de Wishart como la ecuación

$$X_{t}=x+\int_{0}^{t}\left(\sqrt{X}_{s}dB_{s}Q+Q'dB_{s}'\sqrt{X}_{s}\right)+\int_{0}^{t}\left(2pQ'Q+\beta X_{s}+X_{s}\beta'\right)ds,\quad t\geq 0,\tag{2.8}$$

donde B es un movimiento browniano estándar en $M_{d,d}$.

Decimos que una solución a la EDE de Wishart es una solución *débil* si dados los parámetros p, Q, β y x existe un espacio de probabilidad filtrado donde se satisfaga la ecuación en términos de la integral de Itô. Sin embargo, la entenderemos como una solución *fuerte* si el espacio de probabilidad y B también están dados y ahí encontramos una solución, en cuyo caso lo denotaremos por $X\sim WP_{d}\left(p,\alpha,\beta,x\right)$ donde $\alpha=Q'Q$.

Cuando no queramos dar un estado inicial explícito x, denotaremos $X\sim WP_{d}\left(p,\alpha,\beta\right)$ y se entiende que estamos considerando a la colección de distribuciones $\left\{WP_{d}\left(p,\alpha,\beta,x\right)\right\}_{x\in\bar{S}_{d}^{+}}$, i.e., el estado inicial es arbitrario pero no aleatorio. En este último caso, es crucial que existan todas las distribuciones $WP_{d}\left(p,\alpha,\beta,x\right)$. Esta distinción será de gran importancia cuando exploremos las consecuencias de la forma de su generador infinitesimal.

Mientras que una solución fuerte siempre nos da una débil, el converso no siempre es cierto. Una observación interesante es que cuando $Q=0$ (y por lo tanto $\alpha=0$), entonces tenemos una evolución completamente determinista dada por $X_{t}=e^{\beta t}xe^{\beta' t}$. Más adelante veremos que las dos definiciones de procesos de Wishart son equivalentes en un cierto sentido; específicamente, veremos que las soluciones débiles a la EDE de Wishart son las semimartingalas de Wishart.

Proposición 2.34. *(invariancia) Supongamos que X satisface la EDE de Wishart (2.8) y sea $P\in M_{d,d}$ una matriz ortogonal. Entonces el proceso $U=P'XP$ satisface la EDE de*

Wishart para los parámetros

$$\begin{aligned} Q_P &= P'QP \\ \beta_P &= P'\beta P \\ \alpha_P &= Q_P'Q_P = P'\alpha P. \end{aligned}$$

Demostración. Notemos que $P'B$ es un movimiento browniano en $M_{d,d}$ porque cada vector columna mantiene su ley browniana y se preserva la independencia entre columnas. Similarmente podemos concluir que $W = P'BP$ es un movimiento browniano en $M_{d,d}$. Si ahora multiplicamos por P y P' a (2.8), obtenemos

$$\begin{aligned} P'dX_tP &= P'\sqrt{X_t}dB_tQ + P'QdB_t'\sqrt{X_t}P + P'\left(2p\alpha + \beta X_t + X_t\beta'\right)Pdt \\ dU_t &= P'\sqrt{X_t}PP'dB_tPP'QP + P'QPP'dB_t'PP'\sqrt{X_t}P \\ &+ \left(2pP'\alpha P + P'\beta PP'X_tP + P'X_tPP'\beta'P\right)dt \\ &= \sqrt{U_t}dW_tQ_P + Q_PdW_t'\sqrt{U_t} + \left(2p\alpha_P + \beta_PU_t + U_t\beta_P\right)dt, \end{aligned}$$

como se quería probar. □

Observación 2.35. De hecho, si $\rho \in M_{n,d}$, entonces el proceso $Z = \rho X \rho'$ satisface

$$dZ_t = \rho\sqrt{X}_tdB_tQ\rho' + \rho Q'dB_t'\sqrt{X}_t\rho' + \left(2p\rho\alpha\rho' + \rho\beta X_t\rho' + \rho X_t\beta'\rho'\right)dt,$$

en particular, Z no necesariamente es un proceso de Wishart, aunque sí tiene marginales Wishart gracias a la proposición 2.9. En particular, los procesos de Wishart son una familia cerrada bajo cambios de coordenadas ortogonales, las cuáles no deforman las percepciones de tamaños y no lo es para otro tipo de simetrizaciones.

Particularmente, con respecto a un cambio de coordenadas apropiado, a saber, el correspondiente a la matriz ortogonal P en la diagonalización de α, podemos reducir el análisis del proceso de Wishart al caso en que α es diagonal, puesto que α_P sería diagonal en nuestro caso. Si además consideramos el resultado siguiente, donde se enfatiza que el papel de Q es relativamente superfluo porque sólo es de importancia el que $Q'Q = \alpha$, entonces podemos también reducir el estudio a cuando Q_P es diagonal y lo mismo se podría decir de β_P si supiéramos que es simétrico.

Lema 2.36. *(equivalencias entre definiciones)* Supongamos que X es una solución a la EDE de Wishart (2.8), entonces X es una semimartingala de Wishart. Por el contrario, si X es una semimartingala de Wishart y consideremos cualquier $Q \in M_{d,d}$ tal que $Q'Q = \alpha$. Entonces existe un movimiento browniano B en $M_{d,d}$ (posiblemente en una extensión del espacio de probabilidad) tal que X es solución a la EDE de Wishart (2.8). En este sentido, una semimartingala de Wishart es una solución débil a la EDE de Wishart.

Demostración. ($\Longrightarrow$) Supongamos que X satisface (2.8). Es inmediato es un proceso de Itô con deriva $b(X_t) = 2pQ'Q + \beta X_t + X_t\beta'$, mientras que

$$M_t = \int_0^t \left(\sqrt{X_s} dB_s Q + Q' dB_s' \sqrt{X_s}\right),$$

define una martingala local. Por lo tanto, si definimos $D_t = \int_0^t b(X_t)\, dt$, entonces $X = x + D + M$ donde x y D satisfacen las condiciones requeridas para $\alpha = Q'Q$. Finalmente, notemos que

$$M_t^{ij} = \sum_{k,l \le d} \int_0^t \left(\sqrt{X_s}\right)^{ik} dB_s^{kl} Q_{lj} + \int_0^t Q_{ki} dB_s^{lk} \left(\sqrt{X_s}\right)^{lj}, \quad i, j \le d,$$

y que $\left[B^{ij}, B^{i'j'}\right] = \delta_{(i,j)(i',j')} Id$ para $i, i', j, j' \le d$. Luego, con el mismo cálculo que se hizo en el lema 2.17, concluimos que para todo $1 \le i, i', j, j' \le d$,

$$\left[X^{ij}, X^{i'j'}\right] = \left(X^{ii'}\alpha_{jj'} + X^{ij'}\alpha_{ji'} + X^{ji'}\alpha_{ij'} + X^{jj'}\alpha_{ii'}\right) \cdot Id,$$

lo que prueba que X es una semimartingala de Wishart.

($\Longleftarrow$) Ahora supongamos que X es una semimartingala de Wishart como en la definición 2.30 y tomemos una matriz Q tal que $Q'Q = \alpha$, por ejemplo, $Q = \sqrt{\alpha}$. Ahora queremos tratar a $M_{d,d}$ como si fuera $\mathbb{R}^n$ para $n = d^2$ y usar el lema 2.13. Esto se logra a través de la función vec definida anteriormente. De este modo, sólo debemos probaremos que se cumplen las hipótesis del lema. Como en el lema 2.17, notemos que el mapeo (también denotado por $S_{\sqrt{X_t},Q}$ en el lema 2.2)

$$\sigma_t : A \mapsto \sqrt{X}_t A Q + Q' A' \sqrt{X}_t,$$

es lineal y simétrico para cada $t \ge 0$ porque $\sqrt{X}_t$ es simétrica. Además, σ es continuo y adaptado porque $\sqrt{X}$ lo es y gracias a la continuidad de la multiplicación matricial. Ahora sólo hace falta probar que si consideramos a σ como una matriz en $M_{n,n}$, entonces $[M] = \sigma\sigma' \cdot Id = [\sigma \cdot B]$ para algún movimiento browniano B en $\mathbb{R}^n$. Sin embargo, este cálculo ya se hizo, con lo cuál podemos usar el lema 2.13 para concluir la demostración. $\square$

Observación 2.37. La continuidad (y noción de convergencia, en general) que conseguimos en $\mathbb{R}^n$ para $n = d^2$ con respecto a la norma euclidiana es equivalente a la continuidad que nos da la norma de Frobenius en $M_{d,d}$ (porque de hecho vec es isometría), así que podemos pasar de un enfoque a otro sin tener problemas técnicos de esta índole.

Además de que ya conocemos la variación cuadrática de un proceso de Wishart X, la verdadera importancia del lema anterior es que justamente nos traduce el problema de buscar soluciones a la EDE de Wishart en otro que, en ocasiones, es más sencillo. El que el problema de las semimartingalas sea más sencillo se puede apreciar a plena vista porque tuvimos que hacer uso de herramientas técnicas muy fuertes para probar el regreso en la demostración anterior y este habría sido un recurso bastante recurrente a lo largo de este texto de no ser por el lema anterior.

Lema 2.38. *(semimartingalas de Wishart como cambio de tiempo) Sea $X = x + D + M$ una semimartingala de Wishart y definamos $\alpha_{ijkl}(x) = x_{ik}\alpha_{jl} + x_{il}\alpha_{jk} + x_{jk}\alpha_{il} + x_{jl}\alpha_{ik}$. Entonces existe un movimiento browniano B en $M_{d,d}$, posiblemente en una extensión del espacio de probabilidad, tal que*

$$X = x + D + vec^{-1}\pi_{d^2}(\alpha(X))' \cdot (vecB \circ (\Lambda_{d,d}(\alpha(X)) \cdot Id)).$$

Similarmente, si existe un proceso X en $\bar{S}_d^+$ que satisfaga la ecuación anterior donde $D = (2p\alpha + \beta X + X\beta') \cdot Id$, entonces X es una semimartingala de Wishart.

Demostración. ($\Longrightarrow$) Consideremos a la representación de $\alpha(X)$ como matriz en $M_{n,n}$ para $n = d^2$ (denotado de la misma forma) construido de tal forma que para todo $i, j, k, l \in \{1, \ldots, d\}$, su coordenada $(i + (j-1)d, k + (l-1)d)$ está dada por $\alpha_{ijkl}(X)$. A esta representación le podemos calcular $\pi_n(\alpha(X))$ y $\Lambda_{d,d}(\alpha(X))$. Luego, como vimos en el lema 2.15, a M lo podemos expresar de la siguiente forma

$$M = \mathrm{vec}^{-1}\pi_n(\alpha(X))' \cdot (\mathrm{vec}B \circ (\Lambda_{d,d}(\alpha(X)) \cdot Id)),$$

donde B es un movimiento browniano en $M_{d,d}$; obteniendo así, la representación deseada.

($\Longleftarrow$) Esta dirección es trivial, puesto que $N = B \circ (\Lambda_{d,d}(\alpha(X)) \cdot Id)$ sería una martingala local y por lo tanto $M = \mathrm{vec}^{-1}\pi_n(\alpha(X))' \cdot \mathrm{vec}N$ también lo es. Además, podemos calcular la variación cuadrática de M y comprobar que efectivamente

$$\begin{aligned} d[M] &= \pi_n(\alpha(X))' \, d[B \circ (\Lambda_{d,d}(\alpha(X)) \cdot Id)] \, \pi_n(\alpha(X)) \\ &= \pi_n(\alpha(X))' \, d([B] \circ (\Lambda_{d,d}(\alpha(X)) \cdot Id)) \, \pi_n(\alpha(X)) \\ &= \pi_n(\alpha(X))' \, d(\Lambda_{d,d}(\alpha(X)) \cdot Id) \, \pi_n(\alpha(X)) \\ &= \pi_n(\alpha(X))' \, \Lambda_{d,d}(\alpha(X)) \, \pi_n(\alpha(X)) \cdot Id = \alpha(X) \cdot Id, \end{aligned}$$

donde a $[M]$ lo estamos pensando como una matriz en $M_{n,n}$. Consecuentemente, $X = x + D + M$ es una semimartingala de Wishart. $\square$

El problema del resultado anterior es que apela a diagonalizaciones, que en general son muy complejas, de la matriz de variaciones cuadráticas de la vectorización de nuestro proceso. Esto es, anida cálculos complejos y rápidamente se convierte en un problema casi imposible de resolver en general, puesto que incluso la simetría y positividad de X se pierden en el proceso de vectorización. Por otro lado, el enfoque de las EDE nos deja el problema del cálculo de la única raíz cuadrada simétrica y positiva semidefinida, que también tiene su grado de complejidad. Más adelante daremos una representación explícita para un caso particular.

Ahora veremos que la distribución marginal de un proceso de Wishart es justamente la distribución de Wishart. Además, la distribución Wishart exhibe momentos exponenciales y por lo tanto M, la parte martingala local de X, debe ser en realidad una martingala en L^2 en particular. En lo que sigue nos proponemos encontrar la distribución de cada X_t, para esto, primero probaremos un par de resultados sobre el comportamiento afín de X.

Por ahora, nos dedicaremos a construir la transformada de Fourier-Laplace de X_t para todo $t \geq 0$. Para esto, es importante recordar que para una variable aleatoria Z en $\mathbb{R}^d$, si existe $E \exp\{v'Z\}$ para v en un vecindario del 0 en $\mathbb{C}^d$, entonces todos sus momentos existen y tenemos

$$E \exp\{v'Z\} = \sum_{n_1,\ldots,n_d \geq 0} \frac{1}{\left(\sum_i n_i\right)!} E \prod_{i \leq d} v_i^{n_i} Z_i^{n_i}.$$

Esto es, la transformada de Fourier-Laplace es una función analítica. Por lo tanto, en muchas ocasiones basta probar que $E\left|\exp\{v'Z\}\right| < \infty$ en un vecindario del 0 y caracterizar su forma en algún subconjunto con al menos un punto de acumulación en el dominio de convergencia conveniente en términos de una función analítica Φ_Z que esté definida en todo el dominio de convergencia. Luego, el principio de permanencia mostraría que la forma de la transformada de Fourier-Laplace sigue la misma expresión que la de Φ_Z.

Definición 2.39. Sean $p \geq 0$, $\beta \in M_{d,d}$ y $\alpha \in \bar{S}_d^+$ cualesquiera y definamos las siguientes funciones en $\mathbb{R}_+ \times \bar{S}_d^+$ dadas por,

$$\begin{aligned}
\omega_t^\beta(x) &= e^{\beta t} x e^{\beta' t}, \\
\sigma_t^\beta(x) &= 2\int_0^t \omega_s^\beta(x)\, ds, \\
\phi_u(t) &= p \log\left(\det\left(I_d + \sigma_t^\beta(\alpha)\, u\right)\right) \geq 0, && (2.9) \\
\psi_u(t) &= e^{\beta' t} u \left(I_d + \sigma_t^\beta(\alpha)\, u\right)^{-1} e^{\beta t}. && (2.10)
\end{aligned}$$

Todas las funciones son claramente continuas en sus dos argumentos porque los valores propios de $I_d + \sigma_t^\beta(\alpha)\, x$ son todos mayores o iguales a 1, lo que hace a su inversa continua y acotada con valores propios acotados por 1 y al logaritmo de su determinante no negativo. En particular, ψ_u tiene trayectorias en S_d^+ mientras $u \in S_d^+$, puesto que en ese caso

$$\psi_u(t) = e^{\beta' t}\left(u^{-1} + \sigma_t^\beta(\alpha)\right)^{-1} e^{\beta t},$$

la cuál claramente es simétrica, positiva semidefinida (porque sus valores propios son no negativos) e invertible.

Luego, como es una función continua de u, las trayectorias de ψ_u están en $\bar{S}_d^+$ para todo $u \in \bar{S}_d^+$. De hecho, si $\lambda^\beta = \inf \Re(\sigma)$ y $\mu^\beta = \sup \Re(\sigma)$ y $u \in S_d^+$, entonces los valores propios de $u^{-1} + \sigma_t^\beta(\alpha)$ están acotados superiormente por $\lambda_1(u)^{-1} + e^{t\mu^\beta}$ y por lo tanto las trayectorias de los valores propios de ψ_u están acotados inferiormente por

$$\left(\lambda_1(u)^{-1} + e^{2t\mu^\beta}\right)^{-1} e^{2t\lambda^\beta}.$$

Como consecuencia, la imágen de $\left(\bar{S}_d^+ + u\right) \times [0, t]$ bajo ψ está contenida en $S_d^+ + vI_d$ para algún $v > 0$, esto es, se encuentra a distancia positiva de ∂S_d^+ con respecto a la norma inducida por $\langle \cdot, \cdot \rangle$.

A las últimas dos funciones (ϕ_u, ψ_u), se les llamará *cumulantes de Wishart* asociados a p, β y α. Una propiedad bastante importante de ϕ y ψ es que cumplen las llamadas *ecuaciones de semiflujo* y también unas *ecuaciones diferenciales tipo Riccati.*

Algo bastante impresionante es que cualquier proceso càdlàg de Markov cuya transformada de Laplace (o Fourier) depende de manera afín de su estado inicial (como en (2.15)) tiene funciones asociadas ϕ y ψ que cumplen no sólo las propiedades de semiflujo y las ecuaciones de Riccati, sino también la relación (2.15) (con algunas variaciones para el caso de Fourier).

A esta clase de procesos se les conoce como *procesos afines* y al lector interesado en estos resultados, técnicas implementadas y otras caracterizaciones se le refiere a los artículos y libros [21, 22, 23, 24, 25]. Es de especial interés que los procesos de Wishart sean procesos afines, porque su estudio actual se ha concentrado en procesos afines en espacios de la forma $\mathbb{R}^n \times \mathbb{R}_+^m$.

En estos espacios, todos los procesos afines son infinitamente divisibles, es decir que para cualquier n, existen n procesos independientes e idénticamente distribuidos tal que el proceso afín tiene la misma ley que la suma de estos procesos procesos. En contraste, las propiedades geométricas de $\bar{S}_d^+$ permiten la existencia de procesos afines que no sean infinitamente divisibles. El ejemplo más claro siendo el de los procesos de Wishart, puesto que el parámetro p sólo admite valores disctretos cerca del 0.

Lema 2.40. *(propiedad de semiflujo) ϕ y ψ satisfacen, para $s, t \geq 0$ y $u \in \bar{S}_d^+$,*

$$\phi_u(s+t) = \phi_u(t) + \phi_{\psi_u(t)}(s) \tag{2.11}$$

$$\psi_u(s+t) = \psi_{\psi_u(t)}(s). \tag{2.12}$$

Demostración. Calculemos usando el teorema de Sylvester repetidamente,

$$\begin{aligned}
& \det\left(I_d + \sigma_s^\beta(\alpha) e^{\beta' t} u \left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1} e^{\beta t}\right) \\
= & \det\left(I_d + e^{\beta t}\left(\int_0^s e^{\beta r} \alpha e^{\beta' r} dr\right) e^{\beta' t} u \left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1}\right) \\
= & \det\left(I_d + \left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1}\left(I_d + \sigma_{t+s}^\beta(\alpha) u - I_d - \sigma_t^\beta(\alpha) u\right)\right) \\
= & \det\left(\left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1}\left(I_d + \sigma_{t+s}^\beta(\alpha) u\right)\right).
\end{aligned}$$

Se sigue que

$$\begin{aligned}
\exp\left\{p^{-1}\left(\phi_u(t) + \phi_{\psi_u(t)}(s)\right)\right\} &= \det\left(I_d + \sigma_t^\beta(\alpha) u\right) \det\left(\left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1}\left(I_d + \sigma_{t+s}^\beta(\alpha) u\right)\right) \\
&= \det\left(I_d + \sigma_{t+s}^\beta(\alpha) u\right) = \exp\left\{p^{-1}\phi_u(t+s)\right\},
\end{aligned}$$

lo que prueba la primera ecuación. Por otro lado, tenemos la siguiente cadena de identidades

$$\begin{aligned}
& e^{-\beta(t+s)} \psi_{\psi_u(t)}(s) e^{-\beta'(t+s)} \\
= & \left(u\left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1} e^{\beta t}\right)\left(I_d + \sigma_s^\beta(\alpha)\left(e^{\beta' t} u \left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1} e^{\beta t}\right)\right)^{-1} e^{-\beta t} \\
= & u\left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1} e^{\beta t}\left(e^{\beta t} + e^{\beta t}\left(2\int_0^s e^{\beta r} \alpha e^{\beta' r} dr\right) e^{\beta' t} u \left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1} e^{\beta t}\right)^{-1}.
\end{aligned}$$

Unos cálculos algebraicos nos permiten ver que esto es igual a $u\left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1}$ multiplicado por la derecha con

$$\begin{aligned}
& \left(I_d + \left(I_d + \sigma_{t+s}^\beta(\alpha) u - I_d - \sigma_t^\beta(\alpha) u\right)\left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1}\right)^{-1} \\
= & \left(\left(I_d + \sigma_{t+s}^\beta(\alpha) u\right)\left(I_d + \sigma_t^\beta(\alpha) u\right)^{-1}\right)^{-1} \\
= & \left(I_d + \sigma_t^\beta(\alpha) u\right)\left(I_d + \sigma_{t+s}^\beta(\alpha) u\right)^{-1}.
\end{aligned}$$

Es decir que

$$e^{-\beta(t+s)} \psi_{\psi_u(t)}(s) e^{-\beta'(t+s)} = u\left(I_d + \sigma_{t+s}^\beta(\alpha) u\right)^{-1} = e^{-\beta(t+s)} \psi_u(s+t) e^{-\beta'(t+s)},$$

lo que prueba la segunda ecuación. □

Proposición 2.41. *(ecuaciones de Riccati) Las funciones* ϕ *y* ψ *satisfacen un sistema de ecuaciones diferenciales generalizadas de Riccati, a saber, para todo* $t \geq 0$ *y* $u \in S_d^+$ *tenemos,*

$$\frac{\partial}{\partial t}\phi_u = 2ptr\left(\alpha\psi_u\right), \quad \phi_u\left(0\right) = 0, \tag{2.13}$$

$$\frac{\partial}{\partial t}\psi_u = -2\psi_u\alpha\psi_u + \psi_u\beta + \beta'\psi_u, \quad \psi_u\left(0\right) = u. \tag{2.14}$$

Demostración. Las condiciones de frontera son evidentes de las definiciones. Si a es una función con valores en el espacio de matrices invertibles en $M_{n,n}$ diferenciable en un abierto de$\mathbb{R}$, tenemos la identidad $\frac{d}{dt}a^{-1} = -a^{-1}\left(\frac{d}{dt}a\right)a^{-1}$, que se obtiene de derivar $a^{-1}a = I_n$ y despejar adecuadamente (cf. [14]). Entonces podemos derivar (2.10) para obtener

$$\begin{aligned}\frac{\partial}{\partial t}\psi_u\left(t\right) &= \frac{\partial}{\partial t}\left(e^{\beta' t}u\left(I_d + \sigma_t^\beta\left(\alpha\right)u\right)^{-1}e^{\beta t}\right) = \beta'\psi_u\left(t\right) + \psi_u\left(t\right)\beta \\ &- e^{\beta' t}u\left(I_d + \sigma_t^\beta\left(\alpha\right)u\right)^{-1}\left(2e^{\beta t}xe^{\beta' t}u\right)\left(I_d + \sigma_t^\beta\left(\alpha\right)u\right)^{-1}e^{\beta t} \\ &= \beta'\psi_u\left(t\right) + \psi_u\left(t\right)\beta - 2\psi_u\left(t\right)\alpha\psi_u\left(t\right).\end{aligned}$$

La fórmula de Jacobi dice que $\frac{d}{dt}\det\left(a\right) = \operatorname{tr}\left(\operatorname{co}\left(a\right)'\frac{d}{dt}a\right)$ donde $\operatorname{co}\left(A\right)$ es la matriz de cofactores de A. Si factorizamos un $\det\left(a\right)$, se puede simplificar a

$$\frac{d}{dt}\log\left(\det\left(a\right)\right) = \operatorname{tr}\left(a^{-1}\frac{d}{dt}a\right),$$

cuando a es invertible. Usando esta identidad para derivar (2.9) y usar la identidad de Kailath para obtener finalmente para $u \in S_d^+$,

$$\begin{aligned}\frac{\partial}{\partial t}\phi_u\left(t\right) &= p\frac{\partial}{\partial t}\log\left(\det\left(I_d + \sigma_t^\beta\left(\alpha\right)u\right)\right) = p\operatorname{tr}\left(\left(I_d + \sigma_t^\beta\left(\alpha\right)u\right)^{-1}2e^{\beta t}\alpha e^{\beta' t}u\right) \\ &= 2p\operatorname{tr}\left(e^{\beta' t}u\left(I_d + \sigma_t^\beta\left(\alpha\right)u\right)^{-1}e^{\beta t}\alpha\right) = 2p\operatorname{tr}\left(\psi_u\left(t\right)\alpha\right) \geq 0,\end{aligned}$$

donde la última desigualdad prueba que $\phi_u\left(t\right)$ es no decreciente en t. □

Después de todos estos preliminares, finalmente estamos preparados para dar uno de los resultados más importantes del capítulo, lo que motiva justamente al nombre de los procesos de Wishart.

Teorema 2.42. *Supongamos que* $X \sim WP_d\left(p, \alpha, \beta, x_0\right)$*, entonces para todo* $t \geq 0$ *tenemos* $X_t \sim W_d\left(p, \omega_t^\beta\left(x_0\right), \sigma_t^\beta\left(\alpha\right)\right)$ *y*

$$Ee^{-\langle u, X_t\rangle} = e^{-\phi_u(t) - \langle\psi_u(t), x_0\rangle}, \tag{2.15}$$

para todo $t \geq 0$ y $u \in \bar{S}_d^+$. De hecho, la fórmula anterior existe y es finita para toda matriz u hermitiana tal que el espectro de $\sigma_t^\beta(\alpha)u$ está contenido en $(-\infty, 1)$. Además, no hay explosión en tiempo finito.

Demostración. Denotemos por M y D a su parte martingala local y a su proceso de variación finita en compactos respectivamente. En esencia, lo que haremos será construir una semimartingala con una ligera modificación del lado derecho de (2.15) y probar que este es martingala, lo que nos dará una forma de calcular su esperanza y concluir el resultado.

Sean $t > 0$, $u \in S_d^+$ fijos y consideremos los coeficientes laplacianos de Wishart (ϕ, ψ) correspondientes a p, α y β. Construyamos a la semimartingala J^t dada por $J_s^t = f(s, X_s)$ para toda $t > s \geq 0$, donde

$$f(s,x) = e^{-\phi_u(t-s)-\mathrm{tr}(\psi_u(t-s)x)} = e^{-\phi_u(t-s)-\sum_{ij}\psi_u(t-s)_{ij}x_{ji}},$$

para todo $0 \leq s \leq t$ y $x \in \bar{S}_d^+$. Denotemos por $\partial_{ij} = \frac{\partial}{\partial x_{ij}}$ y notemos que

$$\partial_{ij}\partial_{kl}f(s,x) = -\partial_{ij}f(s,x)\,\psi_u(t-s)_{lk} = f(s,x)\,\psi_u(t-s)_{ji}\,\psi_u(t-s)_{lk}.$$

Al usar esto junto con la simetría de ψ, α y X, obtenemos

$$\begin{aligned}
& \sum_{i,j,k,l=1}^{d} \frac{\partial_{ij}\partial_{kl}f(s,X_s)}{J_s^t} d\left[X^{ij}, X^{kl}\right]_s \\
= & \sum_{i,j,k,l=1}^{d} \psi_u(t-s)_{ij}\,\psi_u(t-s)_{lk}\left(X_s^{ik}\alpha_{jl} + X_s^{il}\alpha_{jk} + X_s^{jk}\alpha_{il} + X_s^{jl}\alpha_{ik}\right) ds \\
= & \; 4\sum_{i,j,k,l=1}^{d} \psi_u(t-s)_{ij}\,\alpha_{jl}\psi_u(t-s)_{lk}\,X_s^{ki} ds = 4\mathrm{tr}\left(\psi_u(t-s)\,\alpha\psi_u(t-s)\,X_s\right) ds.
\end{aligned}$$

Ahora usamos la fórmula de Itô junto con la proposición 2.41 para obtener

$$\begin{aligned}
\frac{dJ_s^t}{J_s^t} = & -\left(\frac{\partial}{\partial s}\phi_u(t-s) + \mathrm{tr}\left(\left(\frac{\partial}{\partial s}\psi_u(t-s)\right)X_s\right)\right) ds - \mathrm{tr}\left(\psi_u(t-s)\,dX_s\right) \\
+ & \; \frac{1}{2J_s^t}\sum_{i,j,k,l=1}^{d} \partial_{ij}\partial_{kl}f(s,X_s)\,d\left[X_s^{ij}, X_s^{kl}\right] \\
= & \; (2p\mathrm{tr}\left(\alpha\psi_u(t-s)\right) + \mathrm{tr}\left(\left(\beta'\psi_u(t-s) + \psi_u(t-s)\,\beta - 2\psi_u(t-s)\,\alpha\psi_u(t-s)\right)X_s\right)) ds \\
- & \; \mathrm{tr}\left(\psi_u(t-s)\,dM_s\right) - \mathrm{tr}\left(\psi_u(t-s)\left(2p\alpha + \beta X_s + X_s\beta'\right)\right) ds \\
+ & \; 2\mathrm{tr}\left(\psi_u(t-s)\,\alpha\psi_u(t-s)\,X_s\right) ds = -\mathrm{tr}\left(\psi_u(t-s)\,dM_s\right),
\end{aligned}$$

donde en las cancelaciones usamos la invariancia ante transposiciones y rotaciones de la traza:

$$\begin{aligned}\operatorname{tr}\left(\psi_u\left(t-s\right)X_s\beta'\right) &= \operatorname{tr}\left(\beta X_s'\psi_u\left(t-s\right)'\right)=\operatorname{tr}\left(\beta X_s\psi_u\left(t-s\right)\right)\\ &= \operatorname{tr}\left(\psi_u\left(t-s\right)\beta X_s\right).\end{aligned}$$

Concluimos así que $dJ_s^t=-\operatorname{tr}\left(J_s^t\psi_u\left(t-s\right)dM_s\right)$ es una combinación lineal de martingalas locales y por lo tanto, es una martingala local. Como esta martingala local es no negativa, en particular tenemos una supermartingala y concluimos que si $\sigma_t^{\beta}\left(\alpha\right)u$ tiene a todos valores propios estrictamente menores a 1 para $u\in S_d$, entonces $\phi_u\left(t\right)$ y $\psi_u\left(t\right)$ existen y son finitos, por lo que

$$Ee^{-\langle u,X_t\rangle} \;=\; EJ_t^t\leq EJ_0^t=e^{-\phi_u(t)-\langle\psi_u(t),x_0\rangle}<\infty.$$

Como además para todo $s\geq 0$ y $u\in\bar{S}_d^+$ se satisface $\phi_u\left(s\right)\geq 0$ y $\psi_u\left(s\right)\in\bar{S}_d^+$, entonces concluimos que

$$\log\left(J_s^t\right)=-\phi_u\left(t-s\right)-\operatorname{tr}\left(\sqrt{X_s}\psi_u\left(t-s\right)\sqrt{X_s}\right)\leq 0,$$

lo que implica que J está uniformemente acotado por 1 en $[0,t]$ y por lo tanto es una verdadera martingala. De esta forma, las condiciones de frontera de ϕ y ψ nos dan

$$Ee^{-\langle u,X_t\rangle}=EJ_t^t=J_0^t=e^{-\phi_u(t)-\langle\psi_u(t),x_0\rangle}.$$

Con esto probamos la segunda afirmación y podemos regresar a las definiciones de ϕ, ψ y la distribución Wishart para verificar que X_t efectivamente la tiene distribución Wishart deseada. La extensión de la fórmula para valores de u hermitianos como antes describimos se sigue de las propiedades analíticas de la transformada de Fourier-Laplace cuando esta existe en una vecindad del 0. La útlima afirmación se sigue de que $\mathbb{P}\left(X_t\in\bar{S}_d^+\right)=\lim_{u\to 0}Ee^{-\langle u,X_t\rangle}=1$. $\square$

Observación 2.43. En (2.15) podemos observar una dependencia afín de la transformada de Laplace respecto del estado inicial. Esta clase de comportamientos dio luz a los procesos afines, los cuáles satisfacen (2.15) para algunas funciones ϕ y ψ (y posiblemente para la transformada de Fourier, si es que no nos restringimos a matrices positivas semidefinidas).

Si, por ejemplo, se supiera de antemano que la fórmula (2.15) se satisface, entonces la propiedad de Markov nos daría

$$\begin{aligned}e^{-\phi_u(t+s)-\langle\psi_u(t+s),x\rangle} &= Ee^{-\langle u,X_{t+s}\rangle}=E\left(E\left(e^{-\langle u,X_{t+s}\rangle}\middle|\mathcal{F}_t\right)\right)\\ &= E\left(e^{-\phi_u(s)-\langle\psi_u(s),X_t\rangle}\right)\\ &= \exp\left\{-\phi\left(s,u\right)-\phi\left(t,\psi_u\left(s,u\right)\right)-\left\langle\psi_{\psi_u(s)}\left(t\right),x\right\rangle\right\},\end{aligned}$$

de donde se siguen las ecuaciones de semiflujo

$$\begin{aligned}
\phi_u(t+s) &= \phi_u(s) + \phi_{\psi_u(s)}(t) \\
\psi_u(t+s) &= \psi_{\psi_u(s)}(t).
\end{aligned}$$

Si añadimos la regularidad y otros argumentos fuertes concernientes a la teoría de distribuciones infinitamente divisibles, uno obtiene ecuaciones tipo Riccati para ϕ y ψ. De esta manera, uno puede extender algunas nociones y resultados que tenemos para los procesos de Wishart, hacia los procesos afines. Gracias a este resultado, ya sabemos que M es una verdadera martingala en L^p para todo $p < \infty$.

El siguiente resultado nos será de utilidad más adelante y es de interés en sí mismo, puesto que nos permite tener un criterio simple sobre los momentos exponenciales de $X \cdot Id$.

Lema 2.44. *Supongamos que $X \sim WP_d(p, \alpha, \beta, x)$, entonces para cada $v \in \bar{S}_d^+$ tenemos*

$$Ee^{\left\langle v, \int_0^t X_s ds \right\rangle} < \infty, \quad \forall t < T_v,$$

donde $T_v > 0$ está dado por

$$T_v = \text{ínf}\left\{t > 0 : \ \sup \sigma\left(\sigma_t^\beta(\alpha) v\right) \geq 1\right\},$$

esto es, la primera vez que $\sigma_t^\beta(\alpha) v$ tiene un valor propio mayor o igual a 1.

Demostración. Las desigualdades de Weyl y la continuidad de $\sigma_t^\beta(\alpha)$ y de los valores propios, nos ayudan a ver que de hecho $\sup \sigma\left(\sigma_t^\beta(\alpha) v\right) = \lambda_d\left(\sigma_t^\beta(\alpha) v\right)$ es una función no decreciente del tiempo si $v \in \bar{S}_d^+$. Fijemos $t > 0$ y supongamos de momento que $\sup \sigma\left(\sigma_t^\beta(\alpha) v\right) < 1$, de tal forma que el resultado anterior garantice la existencia de la función generadora de momentos. Se sigue entonces del lema de Fatou y la desigualdad generalizada de Hölder que

$$\begin{aligned}
Ee^{\left\langle v, \int_0^t X_s ds \right\rangle} &= E \lim_{n\to\infty} \exp\left\{\sum_{k=0}^{n-1}\left\langle \frac{v}{n}, X_{t\frac{k}{n}}\right\rangle\right\} \leq \liminf_{n\to\infty} E \prod_{k=0}^{n-1} \exp\left\{\left\langle \frac{v}{n}, X_{t\frac{k}{n}}\right\rangle\right\} \\
&\leq \liminf_{n\to\infty} \prod_{k=0}^{n-1} E\left(\exp\left\{\left\langle v, X_{t\frac{k}{n}}\right\rangle\right\}\right)^{\frac{1}{n}} \\
&= \liminf_{n\to\infty} \prod_{k=0}^{n-1} \exp\left\{-\frac{1}{n}\phi_{-v}\left(t\frac{k}{n}\right) - \frac{1}{n}\left\langle \psi_{-v}\left(t\frac{k}{n}\right), x_0\right\rangle\right\} \\
&= \exp\left\{-\int_0^t \left(\phi_{-v}(s) + \langle \psi_{-v}(s), x_0\rangle\right) ds\right\},
\end{aligned}$$

en particular, $Ee^{\left\langle v,\int_0^t X_s ds\right\rangle} < \infty$ cada vez que $\int_0^t \left(\phi_{-v}(s) + \langle \psi_{-v}(s), x_0\rangle\right) ds > -\infty$. Ahora bien, recordemos que

$$\begin{aligned}\phi_{-v}(t) &= p\log\left(\det\left(I_d - 2\int_0^t e^{\beta t}\alpha e^{\beta' t} v ds\right)\right)\\ \psi_{-v}(t) &= -e^{\beta' t} v\left(I_d - 2\int_0^t e^{\beta t}\alpha e^{\beta' t} v ds\right)^{-1} e^{\beta t},\end{aligned}$$

así que la integral antes mencionada no diverge en tanto $\sigma_t^\beta(\alpha) v$ tenga sus valores propios (los cuáles siempre son no negativos) estrictamente menores a 1. Como la matriz anterior inicia en 0 y es continua, es claro que esta propiedad es cierta hasta el tiempo determinista (posiblemente infinito) estrictamente positivo T_v. □

Observación 2.45. El resultado anterior es igualmente válido para $v \in S_d$ y su demostración se sigue del mismo modo, suponiendo ahora que $\sup_{t_0 \le t} \sup \sigma\left(\sigma_{t_0}^\beta(\alpha) v\right) < 1$. De hecho, si $\sigma_t^\beta(\alpha)$ es una función acotada del tiempo, entonces este tiempo es infinito para algunas elecciones de $v \in S_d^+$ (para un análisis más detallado de $\sigma_t^\beta(\alpha)$, referimos al lector al teorema 2.62). De particular interés es cuando $v = re^{ii}$ para algún $r \ge 0$, en cuyo caso, los valores propios de la matriz coinciden con los de

$$\sqrt{v}\sigma_t^\beta(\alpha)\sqrt{v} = 2r\int_0^t e^{ii}e^{\beta t}\alpha e^{\beta' t}e^{ii} ds = 2re^{ii}\int_0^t \left(e^{\beta t}\alpha e^{\beta' t}\right)_{ii} ds,$$

así que basta ver si $r\sigma_t^\beta(\alpha)_{ii} < 1$ o equivalentemente, basta ver si

$$t < T_{re^{ii}} = \text{ínf}\left\{t > 0 : \ r\sigma_t^\beta(\alpha)_{ii} \ge 1\right\}.$$

Un análisis más específico del tiempo de explosión no es posible a menos que se explore el tiempo de explosión de $\psi_{-u,-v}$ (o particularmente en $\psi_{0,-v}$).

Inspirados en [26], consideremos $v \in \bar{S}_d^+$ y definamos a la matriz $A_v \in M_{2d,2d}$ dada por

$$A_v = \begin{pmatrix} -\beta & 2\alpha \\ v & \beta' \end{pmatrix},$$

consideremos a la función suave $P_{u,v} = \left(S'_{u,v}, T'_{u,v}\right)'$ con valores en $M_{2d,d}$, donde $S_{u,v}$ y $T_{u,v}$ toman valores en $M_{d,d}$, dada por

$$P_{u,v}(t) = e^{tA_v}\begin{pmatrix} I_d \\ u \end{pmatrix}.$$

Se sigue que $S_{u,v}(0) = I_d$ y $T_{u,v}(0) = u$ y por lo tanto podemos definir a la función suave

$$\psi_{u,v}(t) = T_{u,v}(t) S_{u,v}(t)^{-1}, \quad \psi_{u,v}(0) = u,$$

al menos en $[0,T)$, para un tiempo fijo $T > 0$ (posiblemente infinito) en el que $S_{u,v}$ se vuelve singular (no invertible). Luego, como $\frac{\partial}{\partial t} P_{u,v} = A_v P_{u,v}$, se sigue que

$$\begin{aligned}
\frac{\partial}{\partial t} S_{u,v} &= -\beta S_{u,v} + 2\alpha T_{u,v} \\
\frac{\partial}{\partial t} T_{u,v} &= v S_{u,v} + \beta' T_{u,v},
\end{aligned}$$

y por lo tanto en $(0,T)$ tenemos

$$\begin{aligned}
\frac{\partial}{\partial t} \psi_{u,v} &= \left(\frac{\partial}{\partial t} T_{u,v}\right) S_{u,v}^{-1} - T_{u,v} S_{u,v}^{-1} \left(\frac{\partial}{\partial t} S_{u,v}\right) S_{u,v}^{-1} \\
&= \left(v S_{u,v} + \beta' T_{u,v}\right) S_{u,v}^{-1} - T_{u,v} S_{u,v}^{-1} \left(-\beta S_{u,v} + 2\alpha T_{u,v}\right) S_{u,v}^{-1} \\
&= v + \beta' T_{u,v} S_{u,v}^{-1} + T_{u,v} S_{u,v}^{-1} \left(\beta - 2\alpha T_{u,v} S_{u,v}^{-1}\right) \\
&= -2\psi_{u,v} \alpha \psi_{u,v} + \psi_{u,v} \beta + \beta' \psi_{u,v} + v.
\end{aligned}$$

Luego, si definimos

$$\phi_{u,v}(t) = \int_0^t 2p \operatorname{tr}\left(\alpha \psi_{u,v}(s)\right) ds, \quad t \in [0,T),$$

evidentemente tendremos $\frac{\partial}{\partial t} \phi_{u,v} = 2p \operatorname{tr}\left(\alpha \psi_{u,v}\right)$, es decir que en $[0,T)$ se satisfacen las ecuaciones diferenciales tipo Riccati con condición inicial:

$$\begin{aligned}
\tfrac{\partial}{\partial t} \phi_{u,v} &= 2p \operatorname{tr}\left(\alpha \psi_{u,v}\right), & \phi_{u,v}(0) &= 0, \\
\tfrac{\partial}{\partial t} \psi_{u,v} &= -2\psi_{u,v} \alpha \psi_{u,v} + \psi_{u,v} \beta + \beta' \psi_{u,v} + v, & \psi_{u,v}(0) &= u.
\end{aligned}$$

En esencia, estas son las mismas que las que satisfacen (ϕ_u, ψ_u) pero con una perturbación en la dirección v por parte de $\psi_{u,v}$. Sea

$$\begin{aligned}
\Upsilon : S_d &\to S_d \\
x &\mapsto -2x\alpha x + x\beta + \beta' x,
\end{aligned}$$

y a cada partición $0 = t_0 < t_1 = \frac{1}{n} < \cdots < t_n = \frac{nt}{n} = t$ con $n \geq 1$, asociemos las matrices $\psi_{u,v}^{n,k}$ definidas recursivamente para $0 \leq k < n$,

$$\begin{aligned}
\psi_{u,v}^{n,k+1}(t) &= \psi_{u,v}^{n,k}(t) + \left(-2\Upsilon\left(\psi_{u,v}^{n,k}(t)\right) + v\right) \frac{1}{n}, \\
\psi_{u,v}^{n,0}(t) &= u.
\end{aligned}$$

Así, para ver que $\psi_{u,v}$ tiene trayectorias en S_d, basta ver que inductivamente $\psi_{u,v}^{n,k}(t) \in S_d$

y por lo tanto, $\psi_{u,v}(t) = \lim_n \psi_{u,v}^{n,k}(t) \in S_d$. Fijemos algún $t \in (0,T)$ y consideremos $Y_s = \int_0^s X_r dr$. Sea J^t la semimartingala dada por $J_s^t = f(s, X_s, Y_s)$ donde

$$f(s,x,y) = e^{-\phi_{u,v}(t-s) - \mathrm{tr}(\psi_{u,v}(t-s)x + vy)},$$

y adoptemos la notación $\partial_{ij}^x = \frac{\partial}{\partial x_{ij}}$ y $\partial_{ij}^y = \frac{\partial}{\partial y_{ij}}$. Al igual que en el caso marginal, se puede deducir sin problemas que

$$\sum_{i,j,k,l=1}^{d} \frac{\partial_{ij}^x \partial_{kl}^x f(s, X_s, Y_s)}{J_s^t} d\left[X^{ij}, X^{kl}\right]_s = 4\mathrm{tr}(\psi_{u,v}(t-s)\alpha\psi_{u,v}(t-s)X_s)\, ds.$$

Ahora sólo tenemos que probar que la parte de variación finita de J^t se desvanece. Para probar esto, notemos que

$$\begin{aligned} \frac{dJ_s^t}{J_s^t} &= -\left(\frac{\partial}{\partial s}\phi_{u,v}(t-s) + \mathrm{tr}\left(\left(\frac{\partial}{\partial s}\psi_{u,v}(t-s)\right)X_s\right)\right) ds - \mathrm{tr}(v dY_s) \\ &- \mathrm{tr}(\psi_{u,v}(t-s)\, dX_s) + \frac{1}{2J_s^t}\sum_{i,j,k,l=1}^{d} \partial_{ij}^x \partial_{kl}^x f(s, X_s)\, d\left[X_s^{ij}, X_s^{kl}\right]. \end{aligned}$$

Si ahora identificamos los términos que simplificamos anteriormente, podemos reducir esta expresión a

$$\begin{aligned} \frac{dJ_s^t}{J_s^t} &= -\mathrm{tr}(\psi_{u,v}(t-s)\, dM_s) + 2p\mathrm{tr}(\alpha\psi_{u,v})\, ds \\ &+ \mathrm{tr}((-2\psi_{u,v}(t-s)\alpha\psi_{u,v}(t-s) + \psi_{u,v}(t-s)\beta + \beta'\psi_{u,v}(t-s) + v)X_s)\, ds \\ &+ \mathrm{tr}((2\psi_{u,v}(t-s)\alpha\psi_{u,v}(t-s) - v)X_s - \psi_{u,v}(t-s)(2p\alpha + \beta X_s + X_s\beta'))\, ds \\ &= -\mathrm{tr}(\psi_{u,v}(t-s)\, dM_s), \end{aligned}$$

es decir que J^t es una martingala local no negativa y de hecho una exponencial de Doléans-Dade. Aquí, la condición de Novikov se convierte en

$$Ee^{-\frac{1}{2}\left[\int_0^\cdot \mathrm{tr}(\psi_{u,v}(t-s)dM_s)\right]_t} = E\exp\left(-2\int_0^t \mathrm{tr}(\psi_{u,v}(t-s)\alpha\psi_{u,v}(t-s)X_s)\, ds\right) < \infty.$$

Esto es trivialmente cierto mientras $\psi_{u,v}$ viva en $\bar{S}_d^+$ (que es justo el caso por un tiempo puesto por continuidad y porque $\psi_{u,v}(0) = u$) e incluso por un poco más de tiempo más allá de que deja el conjunto si es que lo hace, gracias al resultado anterior y la continuidad de $\psi_{u,v}$. Se sigue que J^t es una verdadera martingala y volvemos a obtener como antes

$$Ee^{-\langle u, X_t\rangle - \left\langle v, \int_0^t X_s ds\right\rangle} = e^{-\phi_{u,v}(t) - \langle\psi_{u,v}(t), x\rangle}.$$

Si $\psi_{u,v}$ en algún momento sale de $\bar{S}_d^+$, entonces podemos variar $x \in \bar{S}_d^+$ (incluso restringiéndonos a $\operatorname{rank}(x) = 1$ si es necesario, de acuerdo a las condiciones de existencia que tendremos después) en la ecuación anterior para ver que el lado izquierdo está acotado por 1 pero el derecho no, por lo que $\psi_{u,v}$ jamás sale de $\bar{S}_d^+$. Como este es el caso, si $T < \infty$ entonces para algún x (que nuevamente podemos restringir a que tenga rango 1) podemos usar la continuidad de X y el teorema de convergencia acotada para ver que

$$\begin{aligned} 0 &= \lim_{t \to \sigma} e^{-\phi_{u,v}(t) - \langle \psi_{u,v}(t), x \rangle} \\ &= \lim_{t \to T} E e^{-\langle u, X_t \rangle - \left\langle v, \int_0^t X_s ds \right\rangle} \\ &= E \lim_{t \to T} e^{-\langle u, X_t \rangle - \left\langle v, \int_0^t X_s ds \right\rangle} \\ &= E e^{-\langle u, X_T \rangle - \left\langle v, \int_0^T X_s ds \right\rangle}, \end{aligned}$$

Esto a su vez implica que c.s.,

$$\langle u, X_T \rangle + \left\langle v, \int_0^T X_s ds \right\rangle = \infty,$$

i.e., tenemos explosión en tiempo finito por parte de X, cosa que en realidad sucede con probabilidad 0. Luego, $T = \infty$ y por lo tanto, casi hemos demostrado el siguiente resultado.

Teorema 2.46. *(distribución conjunta) Supongamos que $X \sim WP_d(p, \alpha, \beta, x_0)$, entonces para todo $u, v \in \bar{S}_d^+$ y $t \geq 0$ tenemos*

$$E \exp\left\{ -\langle u, X_t \rangle - \left\langle v, \int_0^t X_s ds \right\rangle \right\} = \exp\left\{ -\phi_{u,v}(t) - \langle \psi_{u,v}(t), x_0 \rangle \right\}. \tag{2.16}$$

Además, $\phi_{u,v}$ y $\psi_{u,v}$ satisfacen las ecuaciones de semiflujo

$$\begin{aligned} \phi_{u,v}(t+s) &= \phi_{u,v}(s) + \phi_{\psi_{u,v}(s),v}(t) \\ \psi_{u,v}(t+s) &= \psi_{\psi_{u,v}(s),v}(t). \end{aligned}$$

Además, la fórmula anterior se satisface para u, v hermitianas hasta el primer tiempo de explosión de $\psi_{u,v}$.

Demostración. El resultado ya está probado para $u \in S_d^+$, pero tomando límites en (2.16) y usando el teorema de convergencia acotada conseguimos extender el resultado a $u \in \partial S_d^+$.

Si ahora usamos la propiedad de Markov, obtenemos para $s,t\geq 0$,

$$\begin{aligned}
& E\left(E\left(\exp\left\{-\left\langle u,X_{t+s}\right\rangle-\left\langle v,\int_0^{t+s}X_rdr\right\rangle\right\}\middle|\mathcal{F}_t\right)\right)\\
= & E\left(E\left(\exp\left\{-\left\langle u,X_{t+s}\right\rangle-\left\langle v,\int_t^{t+s}X_rdr\right\rangle\right\}\middle|X_t\right)\exp\left\{-\left\langle v,\int_0^{t}X_rdr\right\rangle\right\}\right)\\
= & E\left(\exp\left\{-\phi_{u,v}\left(s\right)-\left\langle \psi_{u,v}\left(s\right),X_t\right\rangle-\left\langle v,\int_0^{t}X_rdr\right\rangle\right\}\right)\\
= & \exp\left\{-\phi_{u,v}\left(s\right)-\phi_{\psi_{u,v}(s),v}\left(t\right)-\left\langle \psi_{\psi_{u,v}(s),v}\left(t\right),x_0\right\rangle\right\},
\end{aligned}$$

de donde se sigue la penúltima afirmación. La última afirmación se sigue nuevamente de las propiedades analíticas de la transformada de Fourier-Laplace cuando esta existe en una vecindad del 0 y de que J^t es una supermartingala, lo que muestra que para $u,v\in S_d$, tenemos hasta antes del tiempo de explosión de $\psi_{u,v}$:

$$\begin{aligned}
& E\exp\left\{-\left\langle u,X_t\right\rangle-\left\langle v,\int_0^t X_sds\right\rangle\right\}\\
= & EJ_t^t\leq EJ_0^t\\
= & \exp\left\{-\phi_{u,v}\left(t\right)-\left\langle\psi_{u,v}\left(t\right),x_0\right\rangle\right\}<\infty,
\end{aligned}$$

como se deseaba probar. □

Observación 2.47. Es fácil transferir estos resultados al caso no homogéneo. Específicamente, con referente a la dualidad del problema de resolver la EDE de Wishart y resolver la EDE que resuelve $Z=e^{-\beta Id}Xe^{-\beta' Id}$ en el lema 2.28, podemos observar que

$$\begin{aligned}
& Ee^{-\langle u,Z_t\rangle}=E\exp\left\{-\mathrm{tr}\left(ue^{-\beta t}X_te^{-\beta' t}\right)\right\}\\
= & E\exp\left\{-\mathrm{tr}\left(e^{-\beta' t}ue^{-\beta t}X_t\right)\right\}\\
= & \exp\left\{-\phi\left(t,\omega_t^{-\beta'}\left(u\right)\right)-\left\langle\psi\left(t,\omega_t^{-\beta'}\left(u\right)\right),x_0\right\rangle\right\}\\
= & \det\left(I_d+\sigma_t^{\beta}\left(\alpha\right)e^{-\beta' t}ue^{-\beta t}\right)^{-p}\exp\left\{-\mathrm{tr}\left(e^{-\beta' t}ue^{-\beta t}\left(I_d+\sigma_t^{\beta}\left(\alpha\right)e^{-\beta' t}ue^{-\beta t}\right)^{-1}e^{\beta t}x_0e^{\beta' t}\right)\right\}\\
= & \det\left(I_d+e^{-\beta t}\sigma_t^{\beta}\left(\alpha\right)e^{-\beta' t}u\right)^{-p}\exp\left\{-\mathrm{tr}\left(u\left(I_d+e^{-\beta t}\sigma_t^{\beta}\left(\alpha\right)e^{-\beta' t}u\right)^{-1}x_0\right)\right\}\\
= & \det\left(I_d+\sigma_t^{-\beta}\left(\alpha\right)u\right)^{-p}e^{-\mathrm{tr}\left(u\left(I_d+\sigma_t^{-\beta}(\alpha)u\right)^{-1}x_0\right)},
\end{aligned}$$

lo que implica que $Z_t\sim W_d\left(p,x,\sigma_t^{-\beta}\left(\alpha\right)\right)$. De hecho esta metodología es fácilmente generalizable al contexto del lema 2.28.

Proposición 2.48. *Supongamos que Q es una función determinística en $M_{d,d}$ tal que $\alpha:=Q'Q\in L\left(Id\right)$ y que $Z=z+N+2p\alpha\cdot Id$ para alguna martingala local N con variación cuadrática*

$$\left[N^{ij},N^{kl}\right]=\left(Z^{ik}\alpha^{jl}+Z^{il}\alpha^{jk}+Z^{jk}\alpha^{il}+Z^{jl}\alpha^{ik}\right)\cdot Id.$$

Entonces $Z_t \sim W_d\left(p, z, 2\int_0^t \alpha_s ds\right)$, *es decir,*

$$Ee^{-\langle u, Z_t\rangle} = \frac{\exp\left\{-tr\left(u\left(I_d + 2\left(\int_0^t \alpha_s ds\right)u\right)^{-1} z\right)\right\}}{\det\left(I_d + 2\left(\int_0^t \alpha_s ds\right)u\right)^p}, \quad u \in \bar{S}_d^+.$$

Si además $\beta \in L\left(Id\right)$ *es una función determinística en* $M_{d,d}$, *entonces*

$$X_t = \omega_t^\beta\left(Z\right) \sim W_d\left(p, \omega_t^\beta\left(z\right), 2\omega_t^\beta\left(\int_0^t \alpha_s ds\right)\right),$$

donde $\omega_t^\beta\left(x\right) = e^{\int_0^t \beta_s ds} x e^{\int_0^t \beta_s' ds}$.

Demostración. Procederemos como en la prueba del teorema 2.42; definamos $A_t = 2\int_0^t \alpha_s ds$, fijemos $t > 0$ y consideremos al proceso $J_s^t = f\left(s, Z_s\right)$, donde

$$\begin{aligned} f\left(s,x\right) &= \exp\left\{-p\log\left(\det\left(I_d + A_{t-s}u\right)\right) - \operatorname{tr}\left(u\left(I_d + A_{t-s}u\right)^{-1} x\right)\right\} \\ &= \exp\left\{-p\log\left(\det\left(I_d + A_{t-s}u\right)\right) - \sum_{ij}\left(u\left(I_d + A_{t-s}u\right)^{-1}\right)_{ij} x_{ji}\right\}, \end{aligned}$$

para todo $0 \leq s \leq t$ y $x \in \bar{S}_d^+$ y $u \in S_d^+$. Denotemos por $\partial_{ij} = \frac{\partial}{\partial x_{ij}}$ y notemos que

$$\partial_{ij}\partial_{kl} f\left(s, Z_s\right) = f\left(s, Z_s\right)\left(u\left(I_d + A_{t-s}u\right)^{-1}\right)_{ji}\left(u\left(I_d + A_{t-s}u\right)^{-1}\right)_{lk}.$$

Al usar esto junto con la simetría de $u\left(I_d + A_{t-s}u\right)^{-1} = \left(u^{-1} + A_{t-s}\right)^{-1}$, α y Z, obtenemos

$$\begin{aligned} &\sum_{ijkl} \frac{\partial_{ij}\partial_{kl} f\left(s, Z_s\right)}{J_s^t} d\left[Z_s^{ij}, Z_s^{kl}\right] \\ =\ &\sum_{ijkl}\left(u\left(I_d + A_{t-s}u\right)^{-1}\right)_{ji}\left(u\left(I_d + A_{t-s}u\right)^{-1}\right)_{lk} \\ \times\ &\left(Z_s^{ik}\alpha_s^{jl} + Z_s^{il}\alpha_s^{jk} + Z_s^{jk}\alpha_s^{il} + Z_s^{jl}\alpha_s^{ik}\right) ds \\ =\ &4\operatorname{tr}\left(u\left(I_d + A_{t-s}u\right)^{-1}\alpha_s u\left(I_d + A_{t-s}u\right)^{-1} Z_s\right) ds. \end{aligned}$$

Ahora usamos la fórmula de Itô junto con las identidades $\frac{d}{dt}a^{-1} = -a^{-1}\left(\frac{d}{dt}a\right)a^{-1}$ y $\frac{d}{dt}\log\left(\det\left(a\right)\right) = \operatorname{tr}\left(a^{-1}\frac{d}{dt}a\right)$ para obtener,

$$\begin{aligned} \frac{dJ_s^t}{J_s^t} &= \left(-p\frac{\partial}{\partial s}\log\left(\det\left(I_d + A_{t-s}u\right)\right) + \operatorname{tr}\left(u\left(\frac{\partial}{\partial s}\left(I_d + A_{t-s}u\right)^{-1}\right) Z_s\right)\right) ds \\ &- \operatorname{tr}\left(u\left(I_d + A_{t-s}u\right)^{-1} dZ_s\right) + \frac{1}{2J_s}\sum_{i,j,k,l=1}^{d}\partial_{ij}\partial_{kl} f\left(s, Z_s\right) d\left[Z_s^{ij}, Z_s^{kl}\right], \end{aligned}$$

que se puede desarrollar para obtener

$$\begin{aligned}
&\Big(2p\mathrm{tr}\left(\left(I_d + A_{t-s}u\right)^{-1}\alpha_{t-s}u\right) + 2\mathrm{tr}\left(u\left(I_d + A_{t-s}u\right)^{-1}\alpha_{t-s}u\left(I_d + A_{t-s}u\right)^{-1}Z_s\right)\Big)\\
&-\mathrm{tr}\left(u\left(I_d + A_{t-s}u\right)^{-1}dN_s\right) - 2p\mathrm{tr}\left(u\left(I_d + A_{t-s}u\right)^{-1}\alpha_s\right)ds\\
&+2\mathrm{tr}\left(u\left(I_d + A_{t-s}u\right)^{-1}\alpha_s u\left(I_d + A_{t-s}u\right)^{-1}Z_s\right)ds = -\mathrm{tr}\left(u\left(I_d + A_{t-s}u\right)^{-1}dN_s\right),
\end{aligned}$$

donde en las cancelaciones usamos la rotación de la traza. Concluimos así que $dJ_s^t = -\mathrm{tr}\left(J_s^t u\left(I_d + A_{t-s}u\right)^{-1}dN_s\right)$ es una combinación lineal de martingalas locales y por lo tanto, es una martingala local. Además, como $\det\left(I_d + A_{t-s}u\right) \geq 1$ y

$$\mathrm{tr}\left(u\left(I_d + A_{t-s}u\right)^{-1}Z_s\right) \geq 0,$$

concluimos que

$$\log\left(J_s^t\right) = -p\log\left(\det\left(I_d + A_{t-s}u\right)\right) - \mathrm{tr}\left(u\left(I_d + A_{t-s}u\right)^{-1}Z_s\right) \leq 0.$$

Esto implica que J está uniformemente acotado por 1 en $[0,t]$ y por lo tanto, es una verdadera martingala. De esta forma obtenemos

$$Ee^{-\langle u, Z_t\rangle} = EJ_t^t = J_0^t = \frac{\exp\left\{-\mathrm{tr}\left(u\left(I_d + 2\left(\int_0^t \alpha_s ds\right)u\right)^{-1}z\right)\right\}}{\det\left(I_d + 2\left(\int_0^t \alpha_s ds\right)u\right)^p},$$

como queríamos probar. La última afirmación se sigue de forma inmediata después de unas pocas manipulaciones algebraicas como en la observación del teorema anterior. $\square$

Vamos ahora a estudiar algunas semimartingalas relacionadas al proceso de Wishart. Con ellas, podremos obtener resultados interesantes más adelante, incolucrando su rango, polaridad de la frontera y resultados asintóticos.

En lo que sigue, para cada matriz compleja A, denotaremos por A^* a la matriz adjunta de A y por $\bar{a}$ al conjugado de $a \in \mathbb{C}$ mientras que $\Re a$ y $\Im a$ son sus partes real e imaginaria.

Proposición 2.49. *(semimartingalas escalares) Supongamos que $X \sim WP_d(p, \alpha, \beta, x)$, entonces existen movimientos brownianos B^1 y B^2 en $\mathbb{R}$ tales que*

$$\begin{aligned}
d\left(tr\left(X_t\right)\right) &= 2\sqrt{tr\left(X_t\alpha\right)}dB_t^1 + 2\left(ptr\left(\alpha\right) + tr\left(\beta X_t\right)\right)dt\\
d\left(\det\left(X_t\right)\right) &= 2\sqrt{\det\left(X_t\right)tr\left(co\left(X_t\right)\alpha\right)}dB_t^2 + 2\left(\left(p - \frac{d-1}{2}\right)tr\left(co\left(X_t\right)\alpha\right) + \det\left(X_t\right)tr\left(\beta\right)\right)dt\\
d\left(\log\left(\det\left(X_t\right)\right)\right) &= 2\sqrt{tr\left(X_t^{-1}\alpha\right)}dB_t^2 + 2\left(\left(p - \frac{d+1}{2}\right)tr\left(X_t^{-1}\alpha\right) + tr\left(\beta\right)\right)dt,
\end{aligned}$$

donde la última fórmula es válida para $x \in S_d^+$ *y sólo en* $[0, T_0)$*, donde*

$$T_0 = \text{ínf}\left\{t > 0 : X_t \in \partial S_d^+\right\}.$$

Además, si $\gamma \neq 0$ *es un valor propio de* β' *asociado a* $u \in \mathbb{C}^d$ *y definimos* $\mu = u^*\alpha u$*, entonces* $Y = u^*Xu$ *es una semimartingala no negativa con estado inicial* $y = u^*xu$ *y proceso de variación finita en compactos dado por* $(2p\mu + 2\Re\gamma Y) \cdot Id$ *y variación cuadrática dada por*

$$[Y] = 2\left(\mu Y + \Re\left(\nu \overline{Z}\right)\right) \cdot Id,$$

donde $Z = u'Xu$ *y* $\nu = u'\alpha u$*. Similarmente,* Z *es una semimartingala con estado inicial* $z = u'xu$*, proceso de variación finita en compactos* $2(p\nu + \gamma Z) \cdot Id$ *y con*

$$\begin{aligned} \left[Z, \bar{Z}\right] &= 4\mu Y \cdot Id, \\ [Z, Z] &= 4\nu Z \cdot Id. \end{aligned}$$

En particular si $\mu = 0$ *y por lo tanto* $\nu = 0$*, entonces la evolución de* Z *y* Y *es determinística y dada por* $Z_t = ze^{2\gamma t}$ *y* $Y_t = ye^{2\Re\gamma t}$ *mientras que si* $\mu > 0 = \nu$*, entonces la parte real e imaginaria de* Z *son fuertemente ortogonales. Si* $\gamma \in \mathbb{R}$*, entonces* $Y = Z$*, de hecho, si* $\mu \neq 0$ *y* $\gamma < 0$ *entonces se trata de un proceso* $CIR\left(-2\gamma, -\frac{p\mu}{\gamma}, 2\sqrt{\mu}, y\right)$*, mientras que si* $\gamma = 0$*, entonces* $\frac{1}{\mu}Y$ *es un proceso* $BESQ\left(p, \frac{y}{\mu}\right)$*.*

Demostración. Primero notemos que

$$\begin{aligned} \operatorname{tr}(X) &= \operatorname{tr}(x) + \operatorname{tr}\left(2p\alpha + \beta X + X\beta'\right) \cdot Id + \operatorname{tr}(M) \\ &= \operatorname{tr}(x) + 2p\operatorname{tr}(\alpha)\, Id + \operatorname{tr}\left(\beta X + \beta X'\right) \cdot Id + \operatorname{tr}(M) \\ &= \operatorname{tr}(x) + \left(2p\operatorname{tr}(\alpha) + 2\operatorname{tr}(\beta X)\right) \cdot Id + \operatorname{tr}(M). \end{aligned}$$

Como las entradas de M son martingalas, entonces $\operatorname{tr}(M)$, al ser suma de algunas de ellas, también es martingala. Inmediatamente podemos calcular su variación cuadrática:

$$\begin{aligned} [\operatorname{tr}(M)] &= \left[\sum_{i=1}^{d} M^{ii}\right] = \sum_{i,j=1}^{d}\left[M^{ii}, M^{jj}\right] = 4\sum_{i,j=1}^{d} X^{ij}\alpha_{ij} \cdot Id, \\ &= 4\sum_{i,j=1}^{d} X^{ij}\alpha_{ji} \cdot Id = 4\operatorname{tr}(X\alpha) \cdot Id, \end{aligned}$$

y por lo tanto es posible usar la representación integral de Doob (lema 2.13) introducir un movimiento browniano B^1 tal que

$$\operatorname{tr}(X) = \operatorname{tr}(x) + 2\sqrt{\operatorname{tr}(X\alpha)} \cdot B^1 + 2p\operatorname{tr}(\alpha)\, Id + 2\operatorname{tr}(\beta X) \cdot Id,$$

como queríamos probar. Usando la fórmula de Jacobi,

$$\partial_{ij}\det(X_t)=\operatorname{tr}\left(\operatorname{co}(X)'e^{ij}\right)=\operatorname{co}(X)_{ij},$$

por lo tanto, de la fórmula de Itô,

$$\begin{aligned}\det(X)-\det(x) &= \operatorname{tr}\left(D(\det(X))'\cdot X\right)+\frac{1}{2}\sum_{i,j,k,l=1}^{d}\partial_{ij}\partial_{kl}\det(X)\cdot\left[X^{ij},X^{kl}\right]\\ &= \operatorname{tr}\left(\operatorname{co}(X)'\cdot(D+M)\right)+\frac{1}{2}\sum_{i,j,k,l=1}^{d}\left(\partial_{ij}\operatorname{co}(X)\right)_{kl}\alpha_{ijkl}(X)\cdot Id.\end{aligned}$$

Para calcular $\frac{1}{2}\sum_{ijkl}\alpha_{ijkl}(a)\left(\partial_{ij}\operatorname{co}(a)\right)_{kl}$ para alguna función a con valores en $M_{d,d}$, notemos que al aplicar el operador ∂_{ji} a la identidad $\operatorname{co}(a)'a=\det(a)I_d$, obtenemos

$$\begin{aligned}\left(\partial_{ji}\operatorname{co}(a)'\right)a &= \operatorname{tr}\left(\operatorname{co}(a)'\partial_{ji}a\right)I_d-\operatorname{co}(a)'\partial_{ji}a\\ \left(\partial_{ij}\operatorname{co}(a)\right)a &= \operatorname{tr}\left(\operatorname{co}(a)'e^{ji}\right)I_d-\operatorname{co}(a)e^{ji}\\ &= \operatorname{co}(a)_{ji}I_d-\operatorname{co}(a)e^{ji},\end{aligned}$$

y una fórmula análoga para $a\left(\partial_{ij}\operatorname{co}(a)\right)$. Si a es simétrica (y por lo tanto $\operatorname{co}(a)$ también), como $\left(e^{ij}A\right)_{kl}=0$ cada vez que $i\neq k$, podemos obtener

$$\begin{aligned}&\sum_{ijkl}\left(\partial_{ij}\operatorname{co}(a)\right)_{kl}\left(a_{ik}\alpha_{jl}+a_{il}\alpha_{jk}+a_{jk}\alpha_{il}+a_{jl}\alpha_{ik}\right)\\ =\ &\sum_{ij}\left(a\left(\partial_{ij}\operatorname{co}(a)\right)\alpha\right)_{ij}+\left(\alpha\left(\partial_{ij}\operatorname{co}(a)\right)a\right)_{ji}+\left(a\left(\partial_{ij}\operatorname{co}(a)\right)\alpha\right)_{ji}+\left(\alpha\left(\partial_{ij}\operatorname{co}(a)\right)a\right)_{ij}\\ =\ &\sum_{ij}\left(\operatorname{co}(a)_{ji}\alpha-e^{ji}\operatorname{co}(a)\alpha\right)_{ij}+\left(\operatorname{co}(a)_{ji}\alpha-\alpha\operatorname{co}(a)e^{ji}\right)_{ji}\\ +\ &\left(\operatorname{co}(a)_{ji}\alpha-e^{ji}\operatorname{co}(a)\alpha\right)_{ji}+\left(\operatorname{co}(a)_{ji}\alpha-\alpha\operatorname{co}(a)e^{ji}\right)_{ij}\\ =\ &\sum_{ij}4\operatorname{co}(a)_{ij}\alpha_{ji}-\left(\operatorname{co}(a)\alpha\right)_{jj}-\left(\alpha\operatorname{co}(a)\right)_{ii}-\sum_{i=j}\left(e^{ji}\operatorname{co}(a)\alpha\right)_{ij}+\left(\alpha\operatorname{co}(a)e^{ji}\right)_{ij}\\ =\ &4\operatorname{tr}\left(\operatorname{co}(a)\alpha\right)-2(d+1)\operatorname{tr}\left(\alpha\operatorname{co}(a)\right)=2(1-d)\operatorname{tr}\left(\operatorname{co}(a)\alpha\right).\end{aligned}$$

Por otro lado,

$$\begin{aligned}&\operatorname{tr}\left(\operatorname{co}(X)'\cdot(D+M)\right)\\ =\ &\operatorname{tr}\left(\operatorname{co}(X)'\left(2p\alpha+\beta X+X\beta'\right)\cdot I_d Id\right)+\operatorname{tr}\left(\operatorname{co}(X)'\cdot M\right)\\ =\ &\left(2p\operatorname{tr}\left(\operatorname{co}(X)\alpha\right)+\operatorname{tr}\left(X\operatorname{co}(X)'\beta\right)+\operatorname{tr}\left(\operatorname{co}(X)'X\beta'\right)\right)\cdot Id+\operatorname{tr}\left(\operatorname{co}(X)'\cdot M\right),\\ =\ &\left(2p\operatorname{tr}\left(\operatorname{co}(X)\alpha\right)+2\det(X)\operatorname{tr}(\beta)\right)\cdot Id+\operatorname{tr}\left(\operatorname{co}(X)'\cdot M\right).\end{aligned}$$

En la fórmula anterior, el último término es una martingala con variación cuadrática dada por

$$\begin{aligned}
\left[\operatorname{tr}\left(\operatorname{co}(X)'\cdot M\right)\right] &= \left[\sum_{ij}\operatorname{co}(X)^{ij}\cdot M^{ij}\right] = \sum_{ijkl}\operatorname{co}(X)^{ij}\operatorname{co}(X)^{kl}\cdot\left[M^{ij},M^{kl}\right]\\
&= \sum_{ijkl}\operatorname{co}(X)^{ij}\operatorname{co}(X)^{kl}\left(X^{ik}\alpha_{jl}+X^{il}\alpha_{jk}+X^{jk}\alpha_{il}+X^{jl}\alpha_{ik}\right)\cdot Id\\
&= 2\sum_{ij}\operatorname{co}(X)^{ij}\left(\left(X\operatorname{co}(X)'\alpha\right)_{ij}+\left(X\operatorname{co}(X)'\alpha\right)_{ji}\right)\cdot Id\\
&= 4\det(X)\sum_{ij}\operatorname{co}(X)^{ij}\alpha_{ji}\cdot Id = 4\det(X)\operatorname{tr}(\operatorname{co}(X)\alpha)\cdot Id.
\end{aligned}$$

Luego, al combinar todos los resultados y usar 2.13, introducimos un movimiento browniano B^2 tal que

$$d(\det(X_t)) = 2\sqrt{\det(X_t)\operatorname{tr}(\operatorname{co}(X_t)\alpha)}dB_t^2+2\left(\left(p-\frac{d-1}{2}\right)\operatorname{tr}(\operatorname{co}(X_t)\alpha)+\det(X_t)\operatorname{tr}(\beta)\right)dt.$$

Ahora definamos $Z=\det(X)$ y supongamos que $x\in S_d^+$, en cuyo caso, podemos usar la fórmula de Itô para obtener, de la ecuación anterior,

$$\begin{aligned}
& d(\log(Z_t)) = Z_t^{-1}dZ_t-\frac{1}{2}Z_t^{-2}d[Z]_t\\
=\ & Z^{-1}\left(2\sqrt{Z_t^2\operatorname{tr}\left(X_t^{-1}\alpha\right)}dB_t^2+2Z_t\left(\left(p-\frac{d-1}{2}\right)\operatorname{tr}\left(X_t^{-1}\alpha\right)+\operatorname{tr}(\beta)\right)dt\right)\\
-\ & \frac{1}{2}Z^{-2}\left(4Z_t^2\operatorname{tr}\left(X_t^{-1}\alpha\right)\right)dt\\
=\ & 2\sqrt{\operatorname{tr}\left(X_t^{-1}\alpha\right)}dB_t^2+2\left(\left(p-\frac{d+1}{2}\right)\operatorname{tr}\left(X_t^{-1}\alpha\right)+\operatorname{tr}(\beta)\right)dt.
\end{aligned}$$

Finalmente, definamos $Y=u^*Xu$ y $\mu=u^*\alpha u$ para calcular

$$\begin{aligned}
dY_t &= u^*dX_tu = u^*dM_tu+u^*\left(2p\alpha+\beta X_t+X_t\beta'\right)udt\\
&= u^*dM_tu+\left(2pu^*\alpha u+u^*\beta X_tu+u^*X_t\beta'u\right)dt\\
&= u^*dM_tu+2\left(p\mu+\Re\gamma u^*X_tu\right)dt = u^*dM_tu+2\left(p\mu+\Re\gamma Y_t\right)dt
\end{aligned}$$

donde la martingala de la izquierda tiene variación cuadrática

$$\begin{aligned}
[u^*Mu] &= \left[\sum_{ij}\bar{u}_iM^{ij}u_j\right] = \sum_{ijkl}\left[\bar{u}_iM^{ij}u_j,\bar{u}_kM^{kl}u_l\right]\\
&= \sum_{ijkl}\bar{u}_iu_j\bar{u}_ku_l\left(X^{ik}\alpha_{jl}+X^{il}\alpha_{jk}+X^{jk}\alpha_{il}+X^{jl}\alpha_{ik}\right)\cdot Id\\
&= \left(2u^*Xuu^*\alpha u+\overline{u'Xu}u'\alpha u+u'Xu\overline{u'\alpha u}\right)\cdot Id\\
&= \left(2\mu Y+2\Re\left(\nu\overline{Z}\right)\right)\cdot Id,
\end{aligned}$$

donde $\nu = u'\alpha u$. Similarmente,

$$\begin{aligned} dZ_t &= u'dX_tu = u'dM_tu + u'\left(2p\alpha + \beta X_t + X_t\beta'\right) udt \\ &= u'dM_tu + \left(2pu'\alpha u + u'\beta X_tu + u'X_t\beta'u\right) dt \\ &= u'dM_tu + \left(2p\nu + 2\gamma u'X_tu\right) dt = u'dM_tu + 2\left(p\nu + \gamma Z_t\right) dt \end{aligned}$$

con variación cuadrática

$$\begin{aligned} [u'Mu] &= \left[\sum_{ij} u_iM^{ij}u_j\right] = \sum_{ijkl}\left[u_iM^{ij}u_j, \overline{u_kM^{kl}u_l}\right] \\ &= \sum_{ijkl} u_iu_j\bar{u}_k\bar{u}_l\left(X^{ik}\alpha_{jl} + X^{il}\alpha_{jk} + X^{jk}\alpha_{il} + X^{jl}\alpha_{ik}\right)\cdot Id \\ &= 4u^*Xuu^*\alpha u\cdot Id \\ &= 4\mu Y\cdot Id, \end{aligned}$$

y similarmente vemos que $[Z, Z] = 4\nu Z\cdot Id$. Así, podemos usar la representación integral de Doob (lema 2.13) para introducir un movimiento browniano B^3 tal que

$$\begin{aligned} Y &= y + \left(2p\mu + 2\Re\gamma Y\right)\cdot Id + \sqrt{2\mu Y + 2\Re\left(\nu\overline{Z}\right)}\cdot B^3 \\ &= y + \left(2p\mu + 2\gamma Y\right)\cdot Id + 2\sqrt{\mu Y}\cdot B^3 \\ &= y - 2\gamma\left(-\frac{p\mu}{\gamma} - Y\right)\cdot Id + 2\sqrt{\mu Y}\cdot B^3, \end{aligned}$$

donde la segunda igualdad sólo es válida (en general) si $\gamma\in\mathbb{R}$ y última igualdad se da sólo si $\gamma\neq 0$, de donde se siguen las afirmaciones del CIR y $BESQ$. Si por otro lado $\mu = 0$ entonces u es vector propio de α con valor propio 0 y claramente $\nu = 0$, en cuyo caso verificamos que $[Z] = 0$ y por lo tanto, Z y Y satisfacen ecuaciones diferenciales determinísticas

$$dZ_t = 2\gamma Z_tdt, \quad dY_t = 2\Re\gamma Y_tdt,$$

cuyas únicas soluciones son $Z_t = ze^{2\gamma t}$ y $Y_t = ye^{2\Re\gamma t}$. Si $\mu > 0 = \nu$, entonces evidentemente

$$[\Re Z, \Re Z] - [\Im Z, \Im Z] + 2\left[\Re Z, \Im Z\right] i = [Z, Z] = 0,$$

lo que implica que su parte real e imaginaria son fuertemente ortogonales y tienen la misma variación cuadrática. Además,

$$2\left(\Re u\right)'\alpha u = \left(u + \bar{u}\right)'\alpha u = \mu > 0,$$

lo que implica que $(\Re u)' \alpha \Im u = 0$ y similarmente obtenemos $(\Im u)' \alpha \Re u = 0$ y

$$0 < \frac{\mu}{2} = (\Re u)' \alpha \Re u = (\Im u)' \alpha \Im u.$$

De hecho, se puede decir más porque si $Z = A + Bi$ para algunas semimartingalas reales A y B, entonces

$$\begin{aligned} dA_t &= 2(\Re\gamma A_t - \Im\gamma B_t)\,dt + \Re(u' dM_t u) \\ dB_t &= 2(\Re\gamma B_t + \Im\gamma A_t)\,dt + \Im(u' dM_t u), \end{aligned}$$

donde la ortogonalidad de A y B implica que $N = \Re(u' dM_t u)$ y $L = \Im(u' dM_t u)$ son martingalas locales fuertemente ortogonales con misma variación cuadrática dada por

$$[N] = \frac{[N] + [L]}{2} = \frac{[\Re Z, \Re Z] + [\Im Z, \Im Z]}{2} = \frac{\left[Z, \bar{Z}\right]}{2} = 2\mu Y \cdot Id,$$

concluyendo la prueba. $\square$

Un resultado muy importante que se puede obtener de estas semimartingalas escalares es un criterio para la inexistencia de la distribución estacionaria. Posteriormente veremos un resultado relacionado, dando condiciones suficientes para la existencia.

Proposición 2.50. *Con la misma notación que en la proposición anterior, si existe un valor propio γ de β' con $\Re\gamma \geq 0$ y $p > 0$, entonces X no tiene distribución ergódica en los siguientes casos:*

(I) $\gamma \neq 0$ y $z \neq 0 = \mu$,

(II) $\gamma \neq 0$ y $y > 0 = \mu$,

(III) $\mu > 0$.

Sin embargo, si $\gamma \in i\mathbb{R}$ y $\mu = 0$, entonces Z y Y tienen distribución invariante.

Demostración. La prueba se basará en ver que Z o Y no pueden converger en distribución y por lo tanto X tampoco puede converger en distribución puesto que Y y Z son funciones continuas de X. Si $\mu = 0$, entonces $Z_t = ze^{2\gamma t}$ no converge en distribución en el primer caso; sin embargo Z sí puede tener distribución invariante si $\gamma \in i\mathbb{R}$, que consiste en la distribución uniforme en la circunferencia en el plano complejo de norma $|z|$ (que, si $z = 0$, se trata de δ_0).

Si $\mu = 0$ y estamos en el segundo caso, entonces $Y_t = ye^{2\Re\gamma t}$ no converge en distribución, lo que cubre todos los dos primeros casos y de ahora en adelante supondremos que $\mu > 0$. De manera general tenemos

$$Y = y + (2p\mu + 2\Re\gamma Y) \cdot Id + \sqrt{2\mu Y + 2\Re\left(\nu \bar{Z}\right)} \cdot B^3,$$

Notemos que como $\Re\gamma Y \geq 0$, entonces

$$Y \geq 2p\mu Id + W_{\left(2\mu Y + 2\Re\left(\nu \bar{Z}\right)\right)\cdot Id} \geq W_{\left(2\mu Y + 2\Re\left(\nu \bar{Z}\right)\right)\cdot Id},$$

para algún movimiento browniano W. En el evento $A := \int_0^\infty \left(2\mu Y_t + 2\Re\left(\nu \bar{Z}_t\right)\right) dt < \infty$, tenemos que

$$\limsup_{t\to\infty} Y \geq \lim_{t\to\infty} 2p\mu t + W_A = \infty,$$

mientras que si $A = \infty$, entonces simplemente

$$\limsup_{t\to\infty} Y \geq \limsup_{t\to\infty} W_t = \infty,$$

lo que prueba que $\limsup_{t\to\infty} Y = \infty$ y por lo tanto no converge en distribución. Para un caso más particular del caso 3, podemos dar otra prueba: si $\nu = 0$, entonces obtenemos

$$Y = y + (2p\mu + 2\Re\gamma Y) \cdot Id + W \circ (2\mu Y \cdot Id),$$

donde W es un movimiento browniano. cambiado de tiempoy en particular, si $\Re\gamma = 0$, entonces $\frac{2}{\mu} Y$ es un proceso de $BESQ\left(2p, \frac{2y}{\mu}\right)$, que no converge en distribución y de hecho $\limsup_t Y_t = \infty$. $\square$

2.4. Propiedades markovianas de los procesos de Wishart

Otro enfoque de este concepto de semimartingalas y procesos de Wishart es el de Markov. Aquí se estudian algunas de las propiedades más importantes, que después nos motivarán a obtener algunas caracterizaciones y resultados importantes. Dentro de las propiedades que se estudian, está el cálculo del generador infinitesimal. Este se calcula en el lema 2.53 y se dan dos distintas pruebas: una basada puramente en argumentos analíticos y de la transformada de Laplace, mientras que el otro enfoque explota las técnicas de semimartingalas continuas.

El cálculo del generador, como se puede ver en el teorema 2.56, nos da consecuencias que la deriva de los procesos de Wishart deben satisfacer para que $WP_d(p, \alpha, \beta)$ exista.

Finalmente, se tiene un estudio sobre el comportamiento asintótico del proceso en el que se encuentran las distribuciones límites, cuando estas existen.

Definición 2.51. Cuando exista, a una familia de kernels $\mu = \left\{\mu_t(x,\cdot) : t \geq 0, x \in \bar{S}_d^+\right\}$ tales que $\mu_t(x,\cdot) = W_d\left(p, \omega_t^\beta(x), \sigma_t^\beta(\alpha)\right)$, se dice transiciones de Wishart con parámetro de deriva constante $p \geq 0$, parámetro de deriva lineal β y coeficiente de difusión $\alpha \in \bar{S}_d^+$.

De la definición es evidente que la transformada de Laplace de $\mu_t(x,\cdot)$ es la misma que vemos en (2.15). Sin embargo, no es inmediato que μ induce un proceso de Markov, pero se puede probar sin mucha dificultad. Además, en este caso será muy importante notar que requerimos la existencia de los $\mu_t(x,\cdot)$ para todo valor inicial $x \in \bar{S}_d^+$.

Como es usual en el estudio de procesos de Markov, denotaremos por E_ν al operador esperanza para la distribución inicial ν con la consideración particular: $E_x = E_{\delta_x}$ para todo $x \in \bar{S}_d^+$, donde δ_x es la medida de Dirac que sólo tiene un átomo en x.

Lema 2.52. *Cualquier familia de kernels de Wishart μ es markoviana (i.e., satisface las ecuaciones de Chapman-Kolmogorov) y se concentra en $\bar{S}_d^+$. Además, el semigrupo de transición asociado $T = (T_t)_{t\geq 0}$, dado por*

$$f \mapsto T_t f(x) = \int_{\bar{S}_d^+} f(y)\, \mu_t(x, dy),$$

es de Feller.

Demostración. La primera afirmación ya ha sido parcialmente probada anteriormente para cualquier distribución Wishart, así que sólo resta ver que se cumple la propiedad de Markov y la de Feller. Para probar estas dos, veremos que basta con la información que ya hemos deducido de nuestras funciones ϕ y ψ. Para probar la propiedad de Markov, basta probar que T es un semigrupo.

Motivados por esto, definamos a $f_u = e^{-\langle u,\cdot\rangle}$ para todo $u \in S_d^+$ y a $L\left(\bar{S}_d^+\right)$ como el espacio vectorial generado por la familia

$$\left\{f_u : u \in S_d^+\right\} \subset C_0\left(\bar{S}_d^+\right).$$

Luego, por el teorema de Stone-Weierstrass para espacios localmente compactos, como $L\left(\bar{S}_d^+\right)$ separa puntos y se anula en ningún lado, entonces es denso en $C_0\left(\bar{S}_d^+\right)$. Luego, basta probar la propiedad de semigrupo y la de Feller para dicha familia. Tomemos $s, t \geq 0$,

$x \in \bar{S}_d^+$ y calculemos usando la propiedad de semiflujo

$$\begin{aligned} T_{t+s} f_u(x) &= \int_{\bar{S}_d^+} f_u(y) \mu_{t+s}(x, dy) = e^{-\phi(t+s,u) - \langle \psi(t+s,u), x \rangle} \\ &= e^{-\phi(t,u) - \phi(s, \psi(t,u)) - \langle \psi(s, \psi(t,u)), x \rangle} \\ &= e^{-\phi(t,u)} \int_{\bar{S}_d^+} f_{\psi(t,u)}(y) \mu_s(x, dy) \\ &= \int_{\bar{S}_d^+} e^{-\phi(t,u) - \langle \psi(t,u), x \rangle} \mu_s(x, dy) = T_s T_t f(x), \end{aligned}$$

lo que prueba que T es un semigrupo. Para la propiedad de Feller, basta ver que $T_t f_u \in C_0\left(\bar{S}_d^+\right)$ y que $T_t f_u(x) \to f_u(x)$ conforme $t \to 0$ para cada $x \in \bar{S}_d^+$. Sin embargo, tenemos que $T_t f_u(x) = \exp\{-\phi(t,u) - \langle \psi(t,u), x \rangle\}$, que claramente es continuo y converge a cero conforme $x \to \infty$ (recordemos que $u \in S_d^+$), lo que implica la primera propiedad. Para la segunda, basta tomar límite conforme $t \downarrow 0$ y usar la continuidad de ϕ y ψ y sus condiciones de frontera, con lo que termina la prueba. $\square$

A los procesos de Markov cuya transformada de Laplace (o Fourier) depende de manera afín del estado inicial como en (2.15) se les conoce como *procesos afínes*. Por esta razón, los procesos de Wishart son una subfamilia importante de los procesos afines continuos en $\bar{S}_d^+$. Además, a diferencia de los procesos afines en $\mathbb{R}^n \times \mathbb{R}_+^m$, los procesos de Wishart no son infinitamente divisibles, cosa que tiene que ver con la estructura geométrica-algebraica del cono $\bar{S}_d^+$. Recordemos que, cuando escribimos $X \sim WP_d(p, \alpha, \beta)$, nos referimos a que X es un proceso de Wishart con estado inicial arbitrario en $\bar{S}_d^+$. Esto es, existen todas las distribuciones $WP_d(p, \alpha, \beta, x)$ y X tiene transiciones de Wishart.

Lema 2.53. *Si $X \sim WP_d(p, \alpha, \beta)$, entonces el semigrupo de transición asociado T tiene un generador infinitesimal $\mathcal{A}$ dado por*

$$\mathcal{A} f(x) = \frac{1}{2} \sum_{i,j,k,l=1}^{d} \alpha_{ijkl}(x) \partial_{ij} \partial_{kl} f(x) + tr\left(\left(\beta x + x\beta' + 2p\alpha\right) Df(x)\right), \tag{2.17}$$

donde, como antes, $\partial_{ij} = \frac{\partial}{\partial x_{ij}}$, $D = (\partial_{ij})$ y

$$\alpha_{ijkl}(x) = x_{ik}\alpha_{jl} + x_{il}\alpha_{jk} + x_{jk}\alpha_{il} + x_{jl}\alpha_{ik}.$$

Primera demostración. Nuestro objetivo es usar el criterio de Watanabe en el núcleo $L\left(\bar{S}_d^+\right)$. El que este sea núcleo se sigue de $T_t f_u = e^{-\phi(t,u)} f_{\psi(t,u)} \in L\left(\bar{S}_d^+\right)$ (lo que prueba la invariancia) y de la densidad del mismo en $C_0\left(\bar{S}_d^+\right)$.

Luego, sólo hace falta probar que $\mathcal{A}$ tiene la forma deseada en este núcleo, y por linealidad, basta probarlo para cada f_u. Fijemos entonces $u \in S_d^+$ y $x \in \bar{S}_d^+$ y notemos que $\phi(t,u) + \langle \psi(t,u) - u, x\rangle \to 0$ conforme $t \to 0$, se sigue que

$$
\begin{aligned}
& \lim_{t \downarrow 0} t^{-1} \left(T_t f_u(x) - f_u(x)\right) = \lim_{t \downarrow 0} t^{-1} \left(e^{-\phi(t,u)} f_{\psi(t,u)}(x) - f_u(x)\right) \\
= \;& f_u(x) \lim_{t \downarrow 0} \frac{e^{-\phi(t,u) - \langle \psi(t,u) - u, x\rangle} - 1}{t} = f_u(x)\, \partial_t e^{-\phi(t,u) - \langle \psi(t,u) - u, x\rangle}\Big|_{t=0} \\
= \;& -f_u(x)\, e^0 \left(\partial_t \phi(t,u) + \operatorname{tr}\left(\partial_t \psi(t,u)\, x\right)\right)\big|_{t=0} \\
= \;& -f_u(x) \left(2p\operatorname{tr}(\alpha u) + \operatorname{tr}\left(\left(-2u\alpha u + u\beta + \beta' u\right) x\right)\right).
\end{aligned}
$$

De hecho, gracias a la tercera línea, podemos ver que la convergencia se da incluso en el sentido fuerte (con respecto a $\|\cdot\|_\infty$). Sólo queda ver que este resultado coincide con lo que resultaría al sustituir a f_u en el lado derecho de (2.17). No obstante, efectivamente coinciden. De hecho tenemos

$$
\begin{aligned}
& \frac{1}{2} \sum_{i,j,k,l=1}^{d} \alpha_{ijkl}(x)\, \partial_{ij} \partial_{kl} f_u(x) + \operatorname{tr}\left(\left(\beta x + x\beta' + 2p\alpha\right) D f_u(x)\right) \\
= \;& \frac{1}{2} \sum_{i,j,k,l=1}^{d} \alpha_{ijkl}(x)\, f_u(x)\, u_{ji} u_{lk} - \sum_{i,j,k=1}^{d} \left(\beta_{ik} x_{kj} + x_{ik} \beta'_{kj}\right) u_{ij} f_u(x) - \sum_{i,j=1}^{d} 2p\alpha_{ij} u_{ij} f_u(x) \\
= \;& f_u(x) \left(\frac{1}{2} \sum_{i,j,k,l=1}^{d} \alpha_{ijkl}(x)\, u_{ji} u_{lk} - \operatorname{tr}\left(2p\alpha u + \left(u\beta + \beta' u\right) x\right) \right).
\end{aligned}
$$

mientras que

$$
\begin{aligned}
\sum_{i,j,k,l=1}^{d} \alpha_{ijkl}(x)\, u_{ji} u_{lk} &= \sum_{i,j,k,l=1}^{d} \left(x_{ik}\alpha_{jl} + x_{il}\alpha_{jk} + x_{jk}\alpha_{il} + x_{jl}\alpha_{ik}\right) u_{ji} u_{lk} \\
&= 4 \sum_{i,j,k,l=1}^{d} u_{ij} \alpha_{jl} u_{lk} x_{ki} \\
&= 4\operatorname{tr}(u\alpha u x),
\end{aligned}
$$

lo que termina la prueba. □

Esta demostración hizo uso de las herramientas sobre la transformada de Laplace del proceso que hemos ido desarrollando junto con argumentos puramente analíticos sobre conjuntos densos en $C_0\left(\bar{S}_d^+\right)$. No obstante, es posible dar otra demostración usando la fórmula de Itô. Para esto, deberemos intercambiar la derivada con la esperanza mútliples veces, de modo que requeriremos un pequeño resultado preliminar.

Lema 2.54. *Sea $f : \mathbb{R}^n \times S \to \mathbb{R}$ acotado tal que $\frac{\partial}{\partial x_i} f$ existe y es dominado uniformemente en alguna vecindad de x por una función integrable $F_i : S \to \mathbb{R}_+$ para $i = 1, \ldots, n$.*

Consideremos a μ una medida finita en $(S, \mathcal{S})$, entonces

$$\frac{d}{dx_i}\int f(x,s)\,\mu(ds) = \int \frac{\partial}{\partial x_i} f(x,s)\,\mu(ds).$$

Demostración. Lo que queremos es hacer uso del teorema de convergencia dominada. Para esto, fijemos $x \in \mathbb{R}^n$ e $i \in \{1, \ldots, n\}$ y definamos e_i como el vector i-ésimo de la base canónica en $\mathbb{R}^n$. Entonces tenemos para $g(x) = \int f(x,s)\,\mu(ds)$, que el teorema del valor medio implica que para cada $h > 0$, existe $\theta(h) \in (0,h)$ tal que

$$h^{-1}\left(f(x+he_i,s) - f(x,s)\right) = \frac{\partial}{\partial x_i} f(x+\theta(h)\,e_i,s)$$

el cuál está dominado por F_i. Luego, tenemos por convergencia acotada,

$$\begin{aligned}
\frac{d}{dx_i} g(x) &= \lim_{h\downarrow 0}\int \frac{f(x+he_i,s)-f(x,s)}{h}\mu(ds)\\
&= \lim_{h\downarrow 0}\int \frac{\partial}{\partial x_i} f(x+\theta(h)\,e_i,s)\,\mu(ds)\\
&= \int \lim_{h\downarrow 0}\frac{f(x+he_i,s)-f(x,s)}{h}\mu(ds)\\
&= \int \frac{\partial}{\partial x_i} f(x,s)\,\mu(ds),
\end{aligned}$$

como queríamos probar. □

Segunda demostración del lema 2.53. Como $L\left(\bar{S}_d^+\right) \subset C_0^\infty\left(\bar{S}_d^+\right)$, es claro que este último es denso en $C_0\left(\bar{S}_d^+\right)$ y probaremos en lo que sigue que es invariante (y por lo tanto un núcleo) y calcularemos además la forma que $\mathcal{A}$ tiene ahí. Para ver esto, tomemos $f \in C_0^\infty\left(\bar{S}_d^+\right)$ y notemos que como f tiene derivadas acotadas de todos los órdenes, podemos usar lema anterior repetidamente para ver que los operadores de derivadas parciales conmutan con E_x, el operador esperanza dado que el proceso inicia en x.

Por lo tanto, $T_t f(x) = E_x f(X_t)$ está en $C^\infty\left(\bar{S}_d^+\right)$, además de que, para cualquier derivada f_k de orden $k \geq 0$ de f, tenemos que $f_k \in C_0\left(\bar{S}_d^+\right)$ y nuevamente por convergencia acotada tenemos que $T_t f_k(x) \to 0$ conforme $x \to \infty$. Consideremos lu descomposición $X = M + D + X_0$ donde M es martingala y D es de variación finita en compactos. Ahora usemos la fórmula de Itô para obtener

$$df(X_t) = \operatorname{tr}\left(Df(X_t)'\,(dM_t + dD_t)\right) + \frac{1}{2}\sum_{i,j,kl=1}^{d} \partial_{ij}\partial_{kl} f(X_t)\, d\left[M^{ij}, M^{kl}\right]_t,$$

o equivalentemente,

$$\begin{aligned} df\left(X_t\right) &= \operatorname{tr}\left(Df\left(X_t\right)' dM_t\right) + \operatorname{tr}\left(Df\left(X_t\right)'\left(\beta X_t + X_t\beta' + 2p\alpha\right)\right) dt \\ &+ \frac{1}{2}\sum_{i,j,kl=1}^{d} \partial_{ij}\partial_{kl} f\left(X_t\right) A_{ijkl}\left(X_t\right) dt. \end{aligned}$$

Notemos que como M es una martingala en L^2 y Df es continuo y acotado, entonces la primera integral es una combinación lineal de martingalas y por lo tanto, es una martingala que inicia en 0. Se sigue que

$$\begin{aligned} &T_t f\left(x\right) - f\left(x\right) = E_x f\left(X_t\right) - f\left(x\right) \\ = \; &E_x\left[\int_0^t \left(\operatorname{tr}\left(Df\left(X_s\right)'\left(\beta X_s + X_s\beta' + 2p\alpha\right)\right) + \frac{1}{2}\sum_{i,j,kl=1}^{d}\int_0^t \partial_{ij}\partial_{kl} f\left(X_s\right) A_{ijkl}\left(X_s\right)\right) ds\right]. \end{aligned}$$

Al dividir entre t y llevar $t \downarrow 0$, obtenemos, por el teorema fundamental del cálculo podemos calcular el límite de los integrandos y sólo debemos justificar el intercambio de la esperanza con el límite. Aquí queremos acotar el integrando, para esto, usando que las derivadas de todos los órdenes de f son acotados y el que X es positivo semidefinido, podemos acotar al integrando por $t^{-1}C_1\int_0^t\left(\operatorname{tr}\left(X_s\right) + C_2\right) ds$ para algunas $C_1, C_2 > 0$.

De acuerdo al teorema del valor medio, esta expresión es igual a $C_1\left(\operatorname{tr}\left(X_{\theta(t)}\right) + C_2\right)$, para algún $\theta\left(t\right) \in \left(0, t\right)$ aleatorio. Entonces, bastaría probar que $Y_t = \sup_{s\leq t}\operatorname{tr}\left(X_s\right)$ tiene primer momento para alguna $t > 0$ para poder usar el teorema de convergencia dominada como en el lema anterior. Una vez que lo hagamos, podremos ver que efectivamente

$$\mathcal{A}f\left(x\right) = \operatorname{tr}\left(Df\left(x\right)'\left(\beta x + X\beta' + 2p\alpha\right)\right) + \frac{1}{2}\sum_{i,j,kl=1}^{d}\partial_{ij}\partial_{kl} f\left(x\right) A_{ijkl}\left(x\right),$$

porque $X_0 = x$, lo que terminaría la prueba. Veamos primero que

$$\operatorname{tr}\left(X\right) = \operatorname{tr}\left(M\right) + \operatorname{tr}\left(x\right) + \operatorname{tr}\left(\beta X + X\beta' + 2p\alpha\right) \cdot Id,$$

tomando supremos, y usando el que X es semidefinido positivo, obtenemos

$$Y_t \leq \sup_{s\leq t}\operatorname{tr}\left(M_s\right) + Lt + \operatorname{tr}\left(x\right) + \int_0^t KY_s dt,$$

para algunas constantes $K, L > 0$. Como $N_t = \sup_{s\leq t}\operatorname{tr}\left(M_s\right) + Lt + \operatorname{tr}\left(x\right)$ es no decreciente, podemos usar la desigualdad de Grönwall para ver que

$$Y_t \leq N_t e^{Kt}.$$

Por otro lado, la desigualdad maximal de Doob muestra que N tiene primer momento finito, ya que $\operatorname{tr}(M_t)$ es martingala en L^p para todo $p \geq 1$. Consecuentemente, lo mismo es cierto para Y_t también, con lo que concluye la prueba. $\square$

Observación 2.55. Como en la situación anterior, dependiendo del problema al que nos enfrentemos, puede ser más favorable abordarlo con un enfoque u otro. En particular, podíamos ver claramente la forma del generador de $\mathcal{A}$ con el enfoque de semimartingalas, mientras que la demostración fue más limpia desde el enfoque de transformadas de Laplace. Por supuesto que la demostración más markoviana hizo uso de varias propiedades que tuvimos que probar con anterioridad, entre ellas, el cálculo de la transformada de Laplace.

Teorema 2.56. *(condiciones de deriva) Sea $X \sim WP_d(p,\alpha,\beta)$ con $\alpha \neq 0$, entonces debemos tener $p \geq \frac{d-1}{2}$. Esto es, la condición $p \geq \frac{d-1}{2}$ es necesaria para la existencia de $WP_d(p,\alpha,\beta)$ cuando $\alpha \neq 0$.*

Demostración. Notemos que para toda $f \in \mathcal{D}(\mathcal{A})$ (el dominio de $\mathcal{A}$) que alcanza su mínimo en x_0, debemos tener

$$\mathcal{A}f(x_0) = \lim_{t \downarrow 0} \frac{E_{x_0}(f(X_t) - f(x_0))}{t} \geq 0.$$

En particular, para cada $x_0 \in \bar{S}_d^+$ podemos tomar una función f_{x_0} de clase $C_0^\infty\left(\bar{S}_d^+\right)$ tal que su comportamiento local en x_0 sea como el de $x \mapsto \det(x)$ (el cuál depende polinomialmente de sus entradas). Dicha función satisface se encuentra en $\mathcal{D}(\mathcal{A})$ y por lo tanto, de la proposición 2.49 obtenemos para cualquier $x_0 \in \partial S_d^+$,

$$\begin{aligned}
0 &\leq \mathcal{A}f_x(x_0) = \mathcal{A}\det(x_0) = (1-d)\operatorname{tr}(\operatorname{co}(x_0)\alpha) + \operatorname{tr}((\beta x_0 + x_0\beta' + 2p\alpha)\operatorname{co}(x_0)) \\
&= (1-d+2p)\operatorname{tr}(\operatorname{co}(x_0)\alpha) + \operatorname{tr}(\beta x \operatorname{co}(x_0) + \operatorname{co}(x_0)x\beta') \\
&= (1-d+2p)\operatorname{tr}(\operatorname{co}(x_0)\alpha) + 2\det(x_0)\operatorname{tr}(\beta) = (1-d+2p)\operatorname{tr}(\operatorname{co}(x_0)\alpha).
\end{aligned}$$

Ahora notemos que como $0 \neq \alpha \in \bar{S}_d^+$, entonces existe $i \in \{1,\ldots,d\}$ tal que $\alpha_{ii} > 0$, por lo tanto, tomando $x_0 = I_d - \operatorname{diag}(e_i)$, obtenemos

$$0 \leq (1-d+2p)\operatorname{tr}(\operatorname{diag}(e_i)\alpha) = (1-d+2p)\alpha_{ii},$$

lo que implica que $2p \geq d-1$. $\square$

Observación 2.57. Como comentamos anteriormente, si $\alpha = 0$ entonces no hay condiciones sobre los demás parámetros puesto que la solución fuerte siempre existe, es determinística y está dada por $X_t = e^{\beta t} x e^{\beta' t}$. Sin embargo, estos casos, al ser determinísticos, son de poco interés. Es importante hacer una aclaración, aunque parece que esto limita la existencia de semimartingalas de Wishart, en realidad no lo hace, simplemente se asegura que si existieran todos los kernels $\mu_t(x,\cdot)$, entonces forzosamente se debe tener $p \geq \frac{d-1}{2}$. Es decir que si tenemos semimartingalas de Wishart para $p < \frac{d-1}{2}$, entonces debemos pedir condiciones extra sobre el estado inicial x, por ejemplo, condiciones sobre el rango de x, como fue el caso del lema 2.26.

Encontrar una distribución estacionaria para el proceso de Wishart no es del todo trivial; sin embargo, de existir, entonces los procesos $Y = u'Xu$ deben también tener distribución invariante. En particular esto se debe satisfacer para el caso en que u sea vector propio de β y en vista de la proposición 2.49, si $\Re\sigma(\beta) \cap [0,\infty)$, donde $\Re z$ es la parte real del complejo z, entonces podemos encontrar un vector propio u tal que Y no tenga distribución invariante.

Este es precisamente el caso del CIR con coeficiente de reversión a la media negativo o del proceso cuadrado de Bessel. Por fortuna, se puede probar que en el caso contrario, si existen ciertas distribuciones, entonces sí existe una distribución estacionaria. Para todo esto, entenderemos por distribución ergódica a la distribución límite, de existir, y no a la distribución ergódica promedio. Es importante la diferenciación puesto que esta última puede existir incluso si el proceso tiene un comportamiento periódico y la primera no, además de que este comportamiento periódico es justo lo que deseamos resaltar.

Para probar este resultado, requeriremos primero un lema concerniente al crecimiento de las matrices exponenciales. El resultado, a grandes razgos, muestra que el crecimiento de dicha matriz es de orden a lo más exponencial, y esta velocidad está fuertemente relacionada con los valores propios de la misma.

Lema 2.58. *Sea $A \in M_{d,d}$ tal que $\Re\sigma(A) \subset (-\infty, r]$, entonces para todo $\epsilon > 0$, existe una constante $M > 0$ tal que*

$$\left\| e^{At} \right\| \leq M e^{(r+\epsilon)t}, \quad t \geq 0.$$

Demostración. Supongamos que $A = P\Lambda P^{-1}$ es una descomposición en bloques de Jordan para A, digamos,

$$\Lambda = \begin{pmatrix} \lambda_1 I_{d_1} + N_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{d_2} + N_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n I_{d_n} + N_n \end{pmatrix},$$

para algunos valores propios λ_k de dimensiones algebraicas d_k y matrices nilpotentes N_k para $k = 1, \ldots, n$ tales que $d_1 + \cdots + d_n = d$. Denotemos por Λ_k por la matriz que resulta de Λ si quitamos todos los bloques excepto el k-ésimo y D_k a la parte correspondiente al I_{d_k} de Λ_k y finalmente $M_k = \Lambda_k - \lambda_k D_k$ para $k = 1, \ldots, n$.

Podemos verificar sencillamente que $\Lambda_k \Lambda_j = 0$ si $k \neq j$ y por lo tanto conmutan. Con esto en mente y usando el que $Be^C B^{-1} = e^{BCB^{-1}}$ para matrices $B, C \in M_{d,d}$ tales que B es invertible (cf. [27]), obtenemos,

$$\begin{aligned} Pe^{At}P^{-1} = e^{PAP^{-1}t} = e^{\Lambda t} = \prod_{k=1}^{n} e^{\Lambda_k t} = \prod_{k=1}^{n} e^{\lambda_k D_k t} e^{M_k t} \\ = \prod_{k=1}^{n} e^{\lambda_k t} D_k \sum_{m=0}^{d-1} \frac{t^m}{m!} M_k^m = \sum_{k=1}^{n} e^{\lambda_k t} D_k + \sum_{k=1}^{n} \sum_{m=1}^{d-1} e^{\lambda_k t} D_k \frac{t^m}{m!} M_k^m, \end{aligned}$$

puesto que todos los productos cruzados se anulan. Se sigue que para todo $\epsilon > 0$,

$$\begin{aligned} \left\| e^{At} \right\| &= \left\| Pe^{At}P^{-1} \right\| \leq \left\| \sum_{k=1}^{n} e^{\lambda_k t} D_k \right\| + \sum_{k=1}^{n} e^{\Re \lambda_k t} \sum_{m=1}^{d-1} \frac{t^m}{m!} \left\| D_k \right\| \left\| M_k \right\|^m \\ &\leq e^{rt} p\left(t\right) = e^{(r+\epsilon)t} p\left(t\right) e^{-\epsilon t}, \end{aligned}$$

para algún polinomio p de grado a lo más $d-1$. Luego, la función $p\left(t\right) e^{-\epsilon t}$ está acotada en $\mathbb{R}_+$ por alguna constante $M > 0$, con lo que finalmente obtenemos el resultado deseado. $\square$

El siguiente resultado complementa lo que se probó en la proposición 2.50, dando condiciones necesarias para la existencia de la distribución ergódica.

Proposición 2.59. *(distribución ergódica) Supongamos que* $\Re\sigma\left(\beta\right) \subset \left(-\infty, 0\right)$*, entonces* $\Sigma = 2\int_0^\infty e^{\beta s} \alpha e^{\beta' s} ds \in \bar{S}_d^+$ *existe. Si además existen* $\nu = W_d\left(p, 0, \Sigma\right)$ *y* $WP_d\left(p, \alpha, \beta\right)$ *entonces* ν *es una distribución estacionaria y límite para* $X \sim WP_d\left(p, \alpha, \beta\right)$*.*

Demostración. Si $\Re\sigma\left(\beta\right) \subset \left(-\infty, 0\right)$, entonces existe $r < 0$ tal que $\Re\sigma\left(\beta\right) \subset \left(-\infty, r\right]$. De aquí, podemos tomar $|r| > \epsilon > 0$ y usar el lema anterior para concluir que existe una constante $M > 0$ tal que para todo $T > t > 0$,

$$\begin{aligned} \left\| \int_t^T e^{\beta s} \alpha e^{\beta' s} ds \right\| &\leq \int_t^T \left\| e^{\beta s} \right\| \left\| \alpha \right\| \left\| e^{\beta' s} \right\| ds \\ &\leq M \int_t^\infty e^{2(r+\epsilon)s} ds \\ &= M \left(1 + \frac{e^{2(r+\epsilon)t}}{2\left(r+\epsilon\right)} \right) \to 0, \end{aligned}$$

conforme $t \to \infty$, garantizando la existencia de Σ. Del mismo modo, podemos usar el lema anterior para ver que $e^{\beta t}, e^{\beta' t} \to 0$ conforme $t \to \infty$. De la convexidad del cerrado $\bar{S}_d^+$, se sigue que $\int_n^{n+1} e^{\beta s}\alpha e^{\beta' s} ds \in \bar{S}_d^+$ para todo $n \geq 0$, luego, de su cerradura ante sumas con coeficientes positivos y su cerradura topológica, obtenemos que $\Sigma \in \bar{S}_d^+$ cada vez que existe.

Ahora notemos que la estacionariedad es equivalente a tener $E_\nu e^{-\langle u, X_0\rangle} = E_\nu e^{-\langle u, X_t\rangle}$ para todo $t \geq 0$ y $u \in \bar{S}_d^+$, esto es, tenemos

$$\det (I_d + \Sigma u)^{-p} = e^{-\phi(t,u)} \int_{\bar{S}_d^+} e^{-\langle \psi(t,u), x\rangle} \nu (dx) = \frac{\det \left(I_d + \sigma_t^\beta (\alpha) u\right)^{-p}}{\det (I_d + \Sigma \psi (t,u))^p}$$

si y sólo si

$$\begin{aligned}
\det (I_d + \Sigma u) &= \det \left(I_d + \sigma_t^\beta (\alpha) u\right) \det \left(I_d + e^{\beta t}\Sigma e^{\beta' t} u \left(I_d + \sigma_t^\beta (\alpha) u\right)^{-1}\right) \\
&= \det \left(\left(I_d + e^{\beta t}\Sigma e^{\beta' t} u \left(I_d + \sigma_t^\beta (\alpha) u\right)^{-1}\right) \left(I_d + \sigma_t^\beta (\alpha) u\right)\right) \\
&= \det \left(I_d + \left(\sigma_t^\beta (\alpha) + e^{\beta t}\Sigma e^{\beta' t}\right) u\right).
\end{aligned}$$

Para que esto suceda, es suficiente que $\Sigma = \sigma_t^\beta (\alpha) + e^{\beta t}\Sigma e^{\beta' t}$, cosa que se puede verificar sencillamente de la definición de $\sigma_t^\beta (\alpha)$ pues

$$\begin{aligned}
\sigma_t^\beta (\alpha) + e^{\beta t}\Sigma e^{\beta' t} &= 2\int_0^t e^{\beta s}\alpha e^{\beta' s} ds + e^{\beta t} \left(2\int_0^\infty e^{\beta s}\alpha e^{\beta' s} ds\right) e^{\beta' t} \\
&= 2\int_0^t e^{\beta s}\alpha e^{\beta' s} ds + 2\int_t^\infty e^{\beta s}\alpha e^{\beta' s} ds = \Sigma.
\end{aligned}$$

La ergodicidad es inmediato la transformada de Laplace, puesto que $E_x e^{-\langle u, X_t\rangle} = e^{-\phi(t,u) - \langle \psi(t,u), x\rangle}$, donde $\sigma_t^\beta (\alpha) \to \Sigma$ conforme $t \to \infty$ y por lo tanto

$$\begin{aligned}
\phi (t,u) &\to p \log (\det (I_d + \Sigma u)) \\
\psi (t,u) &\to \lim_{t\to\infty} e^{\beta' t} u (I_d + \Sigma u)^{-1} e^{\beta t} = 0,
\end{aligned}$$

de donde se sigue que la transformada de Laplace converge a la de $W_d (p, 0, \Sigma)$, como queríamos probar. □

Observación 2.60. Como veremos más adelante, las condiciones de existencia de $W_d (p, 0, \Sigma)$ y $WP_d (p, \alpha, \beta)$ se van a reducir a pedir que $d - 1 \leq 2p$ si $\alpha \neq 0$.

De manera general, la existencia de Σ no está restringida a que $\Re\sigma (\beta) \subset (-\infty, 0)$. De hecho, como siempre tenemos $\Re\sigma \left(e^{t\beta'}\alpha e^{t\beta}\right) \subset \mathbb{R}_+$, lo único necesario es que

$$\Re\sigma \left(e^{t\beta'}\alpha e^{t\beta}\right) \subset [0, o(1)]$$

donde $o(1) \to 0$ conforme $t \to \infty$. En dicho caso, es fácil probar que la decaída es geométrica y por lo tanto la integral existe.

Un ejemplo claro no trivial es considerar a $\alpha = I_d - e^{11}$ y $\beta = -e^{11}$, cuyos valores propios son todos 0 salvo 1 que es -1. Se sigue que $e^{t\beta} = e^{t\beta'} = e^{11}e^{-t}$ y evidentemente se cumplen las condiciones.

A pesar de este problema, es posible obtener una caracterización de este comportamiento. Sin embargo, los parámetros y las condiciones relacionadss dependen fuertemente del comportamiento espectral de β, su descomposición de Jordan y un tanto en su relación con α y x. Estas dependencias quedarán más claras adelante, aunque el ejemplo anterior es suficiente para mostrar su existencia.

Para esto, requeriremos primero de un resultado preliminar que nos da cotas bastante útiles respecto a los valores propios del producto de matrices con respecto a los valores propios de las mismas. Desde luego, será necesario recurrir al principio minimax para valores propios, cuya prueba no sólo no es del todo difícil, sino que es bastante útil al momento de querer deducir el comportamiento de los valores propios de una matriz al sufrir ciertas perturbaciones.

Lema 2.61. *Consideremos dos matrices $A, B \in \bar{S}_n^+$, entonces tenemos las siguientes desigualdades*

$$\lambda_1 (B) \lambda_k (A) \leq \lambda_k (AB) \leq \lambda_n (B) \lambda_k (A),$$

para cualquier $1 \leq k \leq n$.

Demostración. Procediendo como en el lema 2.4, podemos asumir que son positivas definidas. Recordemos del mismo lema que $\sqrt{A}B\sqrt{A}$ y AB tienen los mismos valores propios. Ahora recurrimos al principio minimax para valores propios para así poder obtener las siguientes igualdades

$$\begin{aligned}
\lambda_k (AB) &= \lambda_k \left(\sqrt{A}B\sqrt{A}\right) \\
&= \min_{\substack{V \subset \mathbb{R}^n \\ \dim (V) = k}} \left(\max_{x \in V \setminus \{0\}} \frac{x'\sqrt{A}B\sqrt{A}x}{x'x} \right) \\
&= \min_{\substack{V \subset \mathbb{R}^n \\ \dim (V) = k}} \left(\max_{x \in V \setminus \{0\}} \frac{\left(x\sqrt{A}\right)' B\sqrt{A}x}{\left(x\sqrt{A}\right)' \sqrt{A}x} \frac{x'Ax}{x'x} \right).
\end{aligned}$$

Nuevamente por el principio minimax obtenemos

$$\begin{aligned}
\lambda_1(B)\,\lambda_k(A) &= \min_{\substack{V\subset\mathbb{R}^n\\ \dim(V)=k}} \left(\min_{y\in\mathbb{R}^n\setminus\{0\}} \frac{y'By}{y'y} \max_{x\in V\setminus\{0\}} \frac{x'Ax}{x'x} \right) \\
&\leq \min_{\substack{V\subset\mathbb{R}^n\\ \dim(V)=k}} \left(\max_{x\in V\setminus\{0\}} \frac{\left(x\sqrt{A}\right)' B\sqrt{A}x}{\left(x\sqrt{A}\right)'\sqrt{A}x} \frac{x'Ax}{x'x} \right) \\
&\leq \min_{\substack{V\subset\mathbb{R}^n\\ \dim(V)=k}} \left(\max_{y\in\mathbb{R}^n\setminus\{0\}} \frac{y'By}{y'y} \max_{x\in V\setminus\{0\}} \frac{x'Ax}{x'x} \right) = \lambda_n(B)\,\lambda_k(A),
\end{aligned}$$

lo que prueba el resultado. □

Finalmente estamos preparados para dar un resultado muy fuerte concerniente a las distribuciones invariantes y ergódicas de los procesos de Wishart. Como siempre, asumimos que existe $X \sim WP_d(p,\alpha,\beta,x)$ para alguna elección de parámetros adecuada.

Teorema 2.62. *Si $\Re\sigma(\beta) \not\subset (-\infty,0]$, entonces no existe distribución invariante. Si $\Re\sigma(\beta)\subset(-\infty,0]$, existe una medida invariante si $\sigma_t^\beta(\alpha)$ y $\omega_t^\beta(x)$ son funciones acotadas del tiempo. De hecho, $\sigma_t^\beta(\alpha)$ y $\omega_t^\beta(x)$ admiten la representación*

$$\begin{aligned}
\sigma_t^\beta(\alpha) &= R_0 + \sum_{\lambda\in S,} p_\lambda(t)\,e^{\lambda t}, \\
\omega_t^\beta(x) &= R_1 + \sum_{\lambda\in S,} q_\lambda(t)\,e^{\lambda t},
\end{aligned}$$

para algunas matrices complejas R_0 y R_1, donde p_λ y q_λ son polinomios en t con coeficientes matriciales y donde $S=\{a+b : a,b\in\sigma(\beta)\}$. Existe la distribución invariante si y sólo si las órbitas de $\sigma_t^\beta(\alpha)$ y $\omega_t^\beta(\alpha)$ son cerradas. Más aún, X tiene distribución límite si y sólo si $p_\lambda, q_\lambda \equiv 0$ cada vez que $\Re\lambda = 0$ para algún $\lambda\in S$.

Demostración. La primera parte ya fue probada en la proposición 2.50. Supongamos que $\beta = P\Lambda P^{-1}$ es una descomposición en bloques de Jordan para β, donde supondremos que Λ tiene la forma

$$\Lambda = \begin{pmatrix} \lambda_1 I_{d_1}+N_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{d_2}+N_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n I_{d_n}+N_n \end{pmatrix},$$

para algunos valores propios λ_k de dimensiones algebraicas d_k y matrices nilpotentes N_k para $k=1,\ldots,n$ tales que $d_1+\cdots+d_n=d$. Entonces podemos escribir como antes:

$$\begin{aligned} e^{\beta t} &= \sum_{k=1}^{n}\sum_{m=0}^{d-1} e^{\lambda_k t} P^{-1} D_k \frac{t^m}{m!} M_k^m P, \quad \text{y} \\ e^{\beta' t} &= \sum_{k=1}^{n}\sum_{m=0}^{d-1} e^{\lambda_k t} P' D_k \frac{t^m}{m!} \left(M_k^m\right)' \left(P^{-1}\right)'. \end{aligned}$$

Integrando por partes obtenemos,

$$\begin{aligned} \int_0^t e^{\lambda s} s^k ds &= \lambda^{-1} t^k e^{\lambda t} + \lambda^{-1} k \int_0^t s^{k-1} e^{\lambda s} ds \\ &\vdots \\ &= \sum_{j=0}^{k} \frac{k!\lambda^{-j-1} t^{k-j}}{(k-j)!} e^{\lambda t} - k!\lambda^{-k-1} \\ &= \lambda^{-k-1} k! \left(p_k\left(\frac{t}{\lambda}\right) e^{\lambda t} - 1 \right), \end{aligned}$$

para el polinomio

$$p_k(t) = \sum_{j=0}^{k} \frac{t^{k-j}}{(k-j)!} = \sum_{j=0}^{k} \frac{t^j}{j!}.$$

Se sigue que

$$\begin{aligned} & \omega_t^\beta(x) = e^{\beta t} x e^{\beta' t} \\ = & \sum_{k,j=1}^{n}\sum_{m,l=0}^{d-1} P^{-1} D_k M_k^m P x P' D_j \left(M_j^l\right)' \left(P^{-1}\right)' e^{(\lambda_k+\lambda_j)t} \frac{t^{m+l}}{m!l!} \\ = & R_1 + \sum_{\lambda\in S,} q_\lambda(t) e^{\lambda t}, \end{aligned}$$

mientras que

$$\begin{aligned} \sigma_t^\beta(\alpha) &= 2\int_0^t e^{\beta s} \alpha e^{\beta' s} ds = 2\sum_{k,j=1}^{n}\sum_{m,l=0}^{d-1} P^{-1} D_k M_k^m P \alpha P' D_j \left(M_j^l\right)' \left(P^{-1}\right)' \\ &\times \left(\delta_0(\lambda_k+\lambda_j) \int_0^t \frac{s^{m+l}}{m!l!} ds + (1-\delta_0(\lambda_k+\lambda_j)) \int_0^t e^{(\lambda_k+\lambda_j)s} \frac{s^{m+l}}{m!l!} ds \right) \\ &= 2\sum_{k,j=1}^{n}\sum_{m,l=0}^{d-1} P^{-1} D_k M_k^m P \alpha P' D_j \left(M_j^l\right)' \left(P^{-1}\right)' \left(\delta_0(\lambda_k+\lambda_j) \left(\frac{t^{m+l+1}}{m!l!(m+l+1)} \right) \right. \\ &+ \left. (1-\delta_0(\lambda_k+\lambda_j)) (\lambda_k+\lambda_j)^{-m-l-1} \binom{m+l}{m} \left(p_{m+l}\left(\frac{t}{\lambda_k+\lambda_j}\right) e^{(\lambda_k+\lambda_j)t} - 1 \right) \right) \\ &= R_0 + \sum_{\lambda\in S,} p_\lambda(t) \left((1-\delta_0(\lambda)) e^{\lambda t} + \delta_0(\lambda) \right). \end{aligned}$$

Para evitar una divergencia, que cada vez que $\Re\lambda = 0$, tengamos polinomios constantes, en cuyo caso es claro que $\sigma_t^{\beta}(\alpha)$ y $\omega_t^{\beta}(x)$ son acotados. Esto provoca en $\sigma_t^{\beta}(\alpha)$ una rotación en la dirección $p_\lambda(0)$ con la orientación $\operatorname{sgn}(\Im\lambda)$ y a la velocidad $\Im\lambda$, donde sgn es la función signo, con la convención $\operatorname{sgn}(0) = 0$. Algo equivalente sucede con $\omega_t^{\beta}(x)$. Considerando al conjunto $\Gamma \subset \bar{S}_d^+ \times \bar{S}_d^+$ dado por

$$\Gamma = (R_0, R_1) + \left(\sum_{\lambda\in S,\Re\lambda=0} (p_\lambda(0), q_\lambda(0)) \exp\{\lambda\cdot\}\right)(\mathbb{R}),$$

y a la medida uniforme ν en Γ, obtenemos como medida invariante a $W_d(p,\omega,\sigma)\,\nu(d\omega, d\sigma)$ cuya prueba es idéntica a la de la proposición anterior. En particular, podemos tomar una distribución invariante si ν es finita, i.e., si las órbitas son cerradas, comportamiento ampliamente estudiado en la teoría de sistemas dinámicos y es conocido como *rotación irracional.* No obstante, esta distribución no es límite a menos que todos estos polinomios $p_\lambda, q_\lambda \equiv 0$ cuando $\Re\lambda = 0$ y por lo tanto, este comportamiento también depende de α y x, a diferencia de los otros escenarios. En caso de divergencia en $\sigma_t^{\beta}(\alpha) \in \bar{S}_d^+$, entonces podemos notar las siguientes desigualdades

$$\frac{1}{d}\left\|\sigma_t^{\beta}(\alpha)\right\|_F \leq \operatorname{tr}\left(\sigma_t^{\beta}(\alpha)\right) \leq \det\left(I_d + \sigma_t^{\beta}(\alpha)\right).$$

De manera que, como

$$\begin{aligned}
\phi_u(t) &= p\log\left(\det\left(I_d + \sigma_t^{\beta}(\alpha)\,u\right)\right)\\
\langle\psi_u(t), x\rangle &= \operatorname{tr}\left(u\left(I_d + \sigma_t^{\beta}(\alpha)\,u\right)^{-1}\omega_t^{\beta}(x)\right)\\
&= \operatorname{tr}\left(\left(I_d + \sigma_t^{\beta}(\alpha)\,u\right)^{-1}\omega_t^{\beta}(x)\,u\right),
\end{aligned}$$

entonces debemos tener $\phi_{I_d}(t) \to \infty$. Por lo tanto, obtenemos directamente

$$0 \leq E_x e^{-\langle I_d, X_t\rangle} \leq e^{-\phi_{I_d}(t)} \to 0.$$

Esto implica directamente que X_t diverge en probabilidad, gracias a la desigualdad de Markov:

$$P_x\left(e^{-\langle I_d, X_t\rangle} > \epsilon\right) \leq \frac{E_x e^{-\langle I_d, X_t\rangle}}{\epsilon} \xrightarrow{t\to\infty} 0, \quad \epsilon > 0.$$

Si la divergencia la tenemos en $\omega_t^{\beta}(x)$, mientras que $\sigma_t^{\beta}(\alpha)$ está acotado, entonces el valor propio

$$\lambda_1\left(\left(I_d + \sigma_t^{\beta}(\alpha)\,u\right)^{-1}\right)$$

se mantiene a distancia positiva del 0 como función de t. Se sigue del lema anterior que

$$\begin{aligned}\langle \psi_{I_d}(t), x \rangle &= \operatorname{tr}\left(\left(I_d + \sigma_t^{\beta}(\alpha) \right)^{-1} \omega_t^{\beta}(x) \right) \\ &\geq \lambda_1 \left(\left(I_d + \sigma_t^{\beta}(\alpha) \right)^{-1} \right) \operatorname{tr}\left(\omega_t^{\beta}(x) \right) \to \infty,\end{aligned}$$

lo que igualmente implica que

$$0 \leq E_x e^{-\langle I_d, X_t \rangle} \leq e^{-\left\langle \psi_{I_d}(t), x \right\rangle} \overset{t \to \infty}{\longrightarrow} 0,$$

y por lo tanto X_t diverge en probabilidad. □

Observación 2.63. El problema de encontrar una distribución estacionaria que no sea ergódica es mucho más complejo en el caso introducido en el lema 2.28, donde β y α son procesos estocásticos (deterministas o no). Esto es bastante evidente de lo que obtuvimos del estudio anterior en donde su comportamiento conjunto determinaba este comportamiento. No obstante, se puede ver de su distribución marginal (proposición 2.48) y procediendo como en la proposición 2.59, que si β y α son ambos determinísticos y existen

$$\varpi = \lim_{t \to \infty} \omega_t^{\beta}(x) \quad \text{y} \quad \Sigma = \lim_{t \to \infty} 2\omega_t^{\beta}\left(\int_0^t \alpha_s ds \right),$$

entonces la distribución ergódica de X existe y es $W_d(p, \varpi, \Sigma)$.

2.5. Polaridad de la frontera ∂S_d^+ y EDE espectrales

Además de los casos muy particulares en los que ya probamos existencia para el proceso de Wishart, hay otras situaciones en las que podemos probar existencia fuerte. Supongamos que $X_0 = x \in S_d^+$, como la función $y \mapsto \sqrt{y}$ es analítica (y por lo tanto, localmente Lipschitz) en S_d^+, entonces los coeficientes de la EDE de Wishart son localmente Lipschitz. Por lo tanto existe una única solución fuerte a la EDE en $[0, T_x)$, donde $T_x = \inf\{t > 0 : \det(X_t) = 0\}$ es el tiempo de paro del primer arribo de X a la frontera ∂S_d^+. De aquí que sea de interés encontrar criterios para la polaridad de la frontera, i.e., para determinar cuándo $T_x = \infty$ c.s.

Para poder deducir este resultado, vamos a tener que trabajar un poco, en particular, vamos requerir que $T_x > 0$ c.s. Sin embargo, esta propiedad es consecuencia directa de la continuidad de X y de que nuestro valor inicial x esté a distancia positiva de ∂S_d^+. En todo este capítulo seguiremos las ideas de [6, 2, 9, 4]. Las técnicas que mostramos a continuación están basadas principalmente en las representaciones de las martingalas locales continuas como movimientos brownianos cambiados de tiempo. Esto nos permite explorar sus comportamientos asintóticos, así como sus comportamientos casi singular en las fronteras de sus espacios de estados.

Lema 2.64. *Sea T tiempo de paro y M una martingala local continua en $[0,T)$ donde $T>0$ c.s. Entonces c.s. existe $\lim_{t\uparrow T} M_t$ o $\limsup_{t\uparrow T} M_t = \infty = -\liminf_{t\uparrow T} M_t$, en cuyo caso no tiene límite.*

Demostración. Por el teorema de Dambis-Dubins-Schwarz, M es un movimiento browniano B cambiado de tiempo a través de $[M]$. El resultado se sigue de que el límite existe si y sólo si $\lim_{t\uparrow T} [M]_t < \infty$ y en caso contrario, i.e., cuando $[M]_t \uparrow \infty$, entonces el browniano oscila infinitamente por la recurrencia de sus estados, lo que implica que $\limsup_{t\uparrow T} M_t = \infty = -\liminf_{t\uparrow T} M_t$. □

El resultado siguiente es una consecuencia directa, donde además transferimos el comportamiento del punto 0 al infinito a través de una función adecuada.

Proposición 2.65. *(argumento de McKean) Sea Z un proceso adaptado càdlàg en $[0,T)$ con valores en $\mathbb{R}_+ \setminus \{0\}$ donde $Z_0 > 0$ c.s. y $T = \inf\{t>0 : Z_{t-} = 0\}$. Supongamos además que $h : \mathbb{R}_+ \setminus \{0\} \to \mathbb{R}$ es continua, $\lim_{z\downarrow 0} h(z) = -\infty$ y para todo $t \in [0,T)$, tenemos que $h(Z_t) = h(Z_0) + M_t + R_t$ donde*

(I) R es un proceso càdlàg adaptado tal que $R_0 = 0$ y $\inf_{t\in[0,T\wedge n)} R_t > -\infty$ para todo $n \in \mathbb{N}$ y

(II) M es una martingala local continua con $M_0 = 0$.

Entonces $T = \infty$ c.s.

Demostración. Como Z es continuo por la derecha, entonces $T > 0$ c.s. Además, R_{t-} está acotado inferiormente en compactos y $h(Z_{t-}) = h(Z_0) + M_{t-} + R_{t-}$, por lo tanto $T = \inf\{t > 0 : M_{t-} = -\infty\}$. Es decir que en el conjunto $\{T < \infty\}$, el límite $\lim_{t\uparrow T} M_t$ existe y es igual a $-\infty$. Se sigue del lema anterior que esto sucede con probabilidad 0 y así $T = \infty$ c.s. □

Teorema 2.66. *Supongamos que $2p \geq d+1$ y $x \in S_d^+$, entonces existe una única solución fuerte a la EDE de Wishart tal que c.s. $X_t \in S_d^+$ para todo $t \geq 0$.*

Demostración. Con las observaciones que hicimos anteriormente, dado que $x \in S_d^+$, la solución fuerte ya existe hasta el tiempo T_x. Luego, basta probar que $T_x = \infty$ c.s. Definamos para $t \in [0, T_x)$

$$Z_t = \det(X_t), \quad h(z) = \log(z),$$

y usemos la proposición 2.49 para ver que

$$\begin{aligned} h(Z) &= \log(\det(x)) + Y + R, \quad \text{y} \\ R &= 2\left(\left(p - \frac{d+1}{2}\right)\operatorname{tr}\left(X^{-1}\alpha\right) + \operatorname{tr}(\beta)\right) \cdot Id, \end{aligned}$$

donde Y es una martingala local continua con $Y_0 = 0$ y $[Y] = 4\operatorname{tr}\left(X^{-1}\alpha\right)$. Luego, las hipótesis implican que $R \geq 2\operatorname{tr}(\beta)\, Id$ y así, es acotado inferiormente en compactos. Por lo tanto podemos usar el argumento de McKean y concluir que $T_x = \infty$ c.s. $\square$

En resumen, juntando los resultados anteriores y la construcción explícita de procesos de Wishart hecha en el lema 2.26, podemos concluir que si $2p \geq d+1$ y $x \in S_d^+$ o bien, si $2p \in \mathbb{N}$, entonces existe (en el sentido fuerte y débil respectivamente) una semimartingala de Wishart en todo $\mathbb{R}_+$ con estos parámetros sin restricciones sobre α o β, pero cuyo estado inicial tiene rango a lo más $2p$. Asímismo, si $\alpha = 0$, entonces también tenemos existencia, sin embargo, la evolución es determinística y dada por $X_t = e^{\beta t} x e^{\beta' t}$ de forma que $T_x = \infty 1\left\{x \in S_d^+\right\}$. En ese mismo caso, si tenemos $x \in \partial S_d^+$, entonces $X_t \in \partial S_d^+$ para todo $t \geq 0$.

Teorema 2.67. *Si $\alpha \neq 0$ y $p < \frac{d-1}{2}$, entonces existe $x \in S_d^+$ tal que $T_x < \infty$ con probabilidad positiva. De hecho, para todo $s > 0$ existen estados iniciales $x \in S_d^+$ tales que $\mu_t(x, \cdot)$ no existe para $t \geq s$.*

Demostración. Si suponemos que $T_x = \infty$ c.s. para todo $x \in S_d^+$, entonces en particular estarían bien definidos los kernels $\mu_t(x, \cdot)$ para todo $x \in S_d^+$. Luego, fijando $x_0 \in \partial S_d^+$ y tomando $x_1, x_2, \ldots \in S_d^+$ tales que $x_n \to x_0$, tenemos que sus transformada de Laplace cumplen

$$\lim_{n\to\infty} \int e^{-\langle u,y\rangle} \mu_t(x_n, dy) = \lim_{n\to\infty} e^{-\phi(t,u) - \langle \psi(t,u), x_n\rangle} = e^{-\phi(t,u) - \langle \psi(t,u), x_0\rangle}.$$

Luego, como la función de la derecha es continua, el teorema de continuidad de Lévy para transformadas de Laplace implica que el término de la derecha es una transformada de Laplace, la cuál corresponde por definición a $\mu_t(x_0, \cdot)$. Esto es, $\mu_t(x_0, \cdot)$ está bien definida para todo $x_0 \in \partial S_d^+$ y por lo tanto existe $WP_d(p, \alpha, \beta)$.

Se sigue del teorema 2.56 que $p \geq \frac{d-1}{2}$, lo que nos lleva a una contradicción. Si ahora nos restringiéramos al intervalo temporal $[0, s)$ y suponemos que los kernels $\mu_t(x, \cdot)$ existen para $0 \leq t < s$ y $x \in S_d^+$, a través de los mismos argumentos podemos llegar a una contradicción, lo que prueba la segunda afirmación. □

En preparación para el siguiente resultado, consideremos para cada $n \leq d$ al polinomio simétrico elemental de una matriz simétrica A, dado por

$$e_n(A) = \sum_{1 \leq i_1 < \ldots < i_n \leq d} \lambda_{i_1}(A) \cdots \lambda_{i_n}(A),$$

el cuál es claramente no negativo si $A \in \bar{S}_d^+$, puesto que todos los sumandos son no negativos. De hecho, la importancia de estos polinomios yace en que si $A \in \bar{S}_d^+$, entonces $e_n(A) = 0$ implica que hay al menos $d - n + 1$ valores propios de A que son cero, porque todos los sumandos deben anularse. Esto nos permite concluir inmediatamente que $\mathrm{rnk}(A) \leq n-1$.

Estos polinomios de hecho surgen de considerar los coeficientes del polinomio característico ya que,

$$\det(A + rI_d) = \prod_{i=1}^{d} (\lambda_i(A) + r) = \sum_{n=0}^{d} r^{d-n} e_n(A),$$

donde usamos la convención $e_0(A) = 1$. En particular, tenemos que $e_d(A) = \det(A)$. Finalmente, denotaremos por $e_n^i(A)$ al polinomio simétrico de grado n que no incluye a $\lambda_i(A)$ y similarmente se define a e_n^{ij}.

Proposición 2.68. *Sea $X \sim WP_d(p, \alpha, \beta, x)$, entonces*

$$\begin{aligned} \xi &= \xi_0 + tr\left(\xi Y^{-1} \cdot M\right) + \xi\left(2tr(\beta) - 2rtr\left(Y^{-1}\beta\right)\right) \cdot Id \\ &+ \xi\left(tr\left(Y^{-1}\alpha\right)\left(2p - d + 1 + rtr\left(Y^{-1}\right)\right) - rtr\left(Y^{-2}\alpha\right)\right) \cdot Id, \end{aligned}$$

donde $Y = X + rI_d \in \bar{S}_d^+$ y $\xi = \det(Y) \geq r^d$ para alguna $r > 0$. Si definimos $Z = X + sI_d$ para alguna $s > 0$ y $\eta = \det(Z)$, entonces obtenemos

$$[\xi, \eta] = 4\xi\eta \mathrm{tr}\left(Y^{-1} X Z^{-1} \alpha\right) \cdot Id.$$

En particular, si $(\beta, \alpha) = (0, I_d)$, obtenemos que los polinomios simétricos elementales $e_n(X)$ satisfacen el sistema de ecuaciones diferenciales estocásticas

$$d(e_n(X_t)) = V_t^n dW_t^n + (2p - n + 1)(d - n + 1) e_{n-1}(X_t) dt, \quad 1 \leq n \leq d,$$

donde los procesos W^n son todos movimientos brownianos no necesariamente independientes *y donde* $V^n = V^n(e_1, \ldots, e_d)$ *para algúna función simétrica* V^n *tal que*

$$V^n V^m \cdot [W^n, W^m] = \sum_{i=1}^{d} \lambda_i(X)\, e^i_{n-1}(X)\, e^i_{m-1}(X) \cdot Id.$$

Similarmente, si $(\beta, \alpha) = (bI_d, aI_d)$ *con* $a \geq 0$*, obtenemos, para la misma colección de procesos* $\{V^n\}$ *y* $\{W^n\}$*,*

$$d(e_n(X_t)) = V^n_t dW^n_t + a(2p - n + 1)(d - n + 1)\, e_{n-1}(X_t)\, dt + 2bne_n(X_t)\, dt.$$

Demostración. Procediendo como en la prueba de la proposición 2.49, usaremos las fórmulas de derivadas matriciales para $\partial_{ij} = \frac{\partial}{\partial x_{ij}}$ junto con la fórmula de Itô, específicamente la fórmula de Jacobi, la identidad $\partial_{ij}\xi = \xi \mathrm{tr}(Y^{-1}e^{ij})$ y también el que $\partial_{ij}Y^{-1} = -Y^{-1}e^{ij}Y^{-1}$, para obtener

$$\begin{aligned}
& \xi - \xi_0 - \mathrm{tr}\left(\xi Y^{-1} \cdot M\right) \\
= \; & \mathrm{tr}\left(\xi Y^{-1} \cdot D\right) + \frac{1}{2}\sum_{ijkl} \partial_{ij}\partial_{kl}\xi \cdot \left[X^{ij}, X^{kl}\right] \\
= \; & \mathrm{tr}\left(\xi Y^{-1} \cdot D\right) + \frac{1}{2}\sum_{ijkl} \partial_{ij}\left(\xi \mathrm{tr}\left(Y^{-1}e^{kl}\right)\right)\alpha_{ijkl}(X) \cdot Id \\
= \; & \mathrm{tr}\left(\xi Y^{-1} \cdot D\right) + \frac{1}{2}\sum_{ijkl} \left(\xi \mathrm{tr}\left(Y^{-1}e^{ij}\right)\mathrm{tr}\left(Y^{-1}e^{kl}\right) - \xi \mathrm{tr}\left(Y^{-1}e^{ij}Y^{-1}e^{kl}\right)\right)\alpha_{ijkl}(X) \cdot Id.
\end{aligned}$$

Esto a su vez se puede simplificar más con algunas manipulaciones algebraicas, de hecho, el proceso es absolutamente continuo respecto de la medida de Lebesgue. La densidad consta de ξ multiplicada por la suma de

$$\begin{aligned}
& \mathrm{tr}\left(Y^{-1}(2p\alpha + \beta X + X\beta')\right) \\
= \; & \mathrm{tr}\left(Y^{-1}(2p\alpha - r(\beta + \beta') + \beta Y + Y\beta')\right) \\
= \; & \mathrm{tr}\left(2pY^{-1}\alpha - 2rY^{-1}\beta + 2\beta\right)
\end{aligned}$$

y

$$\begin{aligned}
& \frac{1}{2}\sum_{ijkl}\left(\left(Y^{-1}\right)^{ji}\left(Y^{-1}\right)^{lk} - \left(Y^{-1}\right)^{li}\left(Y^{-1}\right)^{jk}\right)\left(\alpha_{ik}X^{jl} + \alpha_{il}X^{jk} + \alpha_{jk}X^{il} + \alpha_{jl}X^{ik}\right) \\
= \; & \frac{1}{2}\left(4\mathrm{tr}\left(Y^{-1}\alpha Y^{-1}X\right) - 2\mathrm{tr}\left(Y^{-1}\alpha Y^{-1}X\right) - 2\sum_{ij}\left(Y^{-1}\alpha\right)^{ii}\left(Y^{-1}X\right)^{jj}\right) \\
= \; & \left(\mathrm{tr}\left(Y^{-1}\alpha Y^{-1}(Y - rI_d)\right) - \mathrm{tr}\left(Y^{-1}\alpha\right)\mathrm{tr}\left(Y^{-1}(Y - rI_d)\right)\right).
\end{aligned}$$

Esto finalmente nos lleva a obtener que la parte martingala local continua de ξ es $\mathrm{tr}(\xi Y^{-1} \cdot M)$

y su parte de variación finita en compactos es

$$\xi\left[2\mathrm{tr}\left(\beta\right)-2r\mathrm{tr}\left(Y^{-1}\beta\right)+\mathrm{tr}\left(Y^{-1}\alpha\right)\left(2p-d+1+r\mathrm{tr}\left(Y^{-1}\right)\right)-r\mathrm{tr}\left(Y^{-2}\alpha\right)\right]\cdot Id.$$

En particular, si $(\beta,\alpha)=(0,I_d)$, entonces las cantidades anteriores son mucho más simples porque los valores propios de Y^{-1} son todos de la forma $\left(\lambda_i\left(X\right)+r\right)^{-1}$ para $1\leq i\leq d$. De esta forma, algunos cálculos nos ayudan a obtener que la densidad de la parte de variación finita en compactos es

$$\begin{aligned}
&\xi\left(\mathrm{tr}\left(Y^{-1}\right)\left(2p-d+1+r\mathrm{tr}\left(Y^{-1}\right)\right)-r\mathrm{tr}\left(Y^{-2}\right)\right)\\
=&\left(\prod_{i=1}^{d}\left(\lambda_i\left(X\right)+r\right)\right)\left(\left(\sum_{i=1}^{d}\frac{1}{\lambda_i\left(X\right)+r}\right)\left(2p-d+1+r\sum_{i=1}^{d}\frac{1}{\lambda_i\left(X\right)+r}\right)-r\sum_{i=1}^{d}\frac{1}{\left(\lambda_i\left(X\right)+r\right)^2}\right)\\
=&\left(\prod_{i=1}^{d}\left(\lambda_i\left(X\right)+r\right)\right)\left(\left(2p-d+1\right)\sum_{i=1}^{d}\frac{1}{\lambda_i\left(X\right)+r}+r\sum_{i\neq j}\frac{1}{\left(\lambda_i\left(X\right)+r\right)\left(\lambda_i\left(X\right)+r\right)}\right)\\
=&\left(2p-d+1\right)\sum_{i=1}^{d}\prod_{k\neq i}\left(\lambda_k\left(X\right)+r\right)+r\sum_{i\neq j}\prod_{k\neq i,j}\left(\lambda_k\left(X\right)+r\right).
\end{aligned}$$

Ahora notemos de la definición de los polinomios simétricos e_n que, si agrupamos por potencias de r, esto lo podemos reescribir como

$$\begin{aligned}
&\left(2p-d+1\right)\sum_{i=1}^{d}\sum_{n=1}^{d}r^{d-n}e_{n-1}^{i}\left(X\right)+\sum_{i\neq j}\sum_{n=1}^{d-1}r^{d-n}e_{n-1}^{ij}\left(X\right)\\
=&\sum_{n=1}^{d-1}r^{d-n}\left(\left(2p-d+1\right)\sum_{i=1}^{d}e_{n-1}^{i}\left(X\right)+\sum_{i\neq j}e_{n-1}^{ij}\left(X\right)\right)+\left(2p-d+1\right)\sum_{i=1}^{d}e_{d-1}^{i}\left(X\right).
\end{aligned}$$

Observemos que para todo $n\leq d$ tenemos $\sum_{i=1}^{d}e_{n-1}^{i}\left(X\right)=\left(d-n+1\right)e_{n-1}\left(X\right)$ puesto que cada combinación de productos $\lambda_{i_1}\left(A\right)\cdots\lambda_{i_n}\left(A\right)$ de valores propios de X en $e_{n-1}\left(X\right)$ aparece exactamente una vez en cada $e_{n-1}^{i}\left(X\right)$ para todo $i\notin\{i_1,\ldots,i_n\}$. Similarmente obtenemos $\sum_{i\neq j}e_{n-1}^{ij}\left(X\right)=\left(d-n+1\right)\left(d-n\right)e_{n-1}\left(X\right)$ para todo $n\leq d-1$ y por lo tanto,

$$\begin{aligned}
&\xi-\xi_0-\mathrm{tr}\left(\xi Y^{-1}\cdot M\right)=\left(2p-d+1\right)e_{d-1}\left(X\right)\cdot Id\\
+&\sum_{n=1}^{d-1}r^{d-n}\left(\left(2p-d+1\right)\left(d-n+1\right)+\left(d-n+1\right)\left(d-n\right)\right)e_{n-1}\left(X\right)\cdot Id\\
=&\sum_{n=1}^{d}r^{d-n}\left(2p-n+1\right)\left(d-n+1\right)e_{n-1}\left(X\right)\cdot Id.
\end{aligned}$$

Si ahora agrupamos de acuerdo a las potencias de r, obtenemos el resultado deseado, salvo por la parte martingala local. Análogamente, si tenemos $(\beta,\alpha)=(bI_d,aI_d)$, entonces la expresión de arriba se modifica ligeramente para obtener que la deriva de ξ es lo único que cambia. De hecho ahora tenemos (aunque dividimos por a, que posiblemente es 0, lo hicimos para identificar la forma que adquiere y evitarnos cálculos, sin embargo, esto no altera los resultados),

$$\frac{d\xi_t}{\xi_t} = \mathrm{tr}\left(Y_t^{-1} dM_t\right) + a\left[2\frac{b}{a}d + \mathrm{tr}\left(Y_t^{-1}\right)\left(2p - 2r\frac{b}{a} - d + 1 + r\mathrm{tr}\left(Y_t^{-1}\right)\right) - r\mathrm{tr}\left(Y_t^{-2}\right)\right] dt,$$

y como $2bd\xi = \sum_{n=0}^{d} r^{d-n}\left(2bde_n\left(X\right)\right)$, obtenemos análogamente que

$$\begin{aligned}
& \xi - \xi_0 - \mathrm{tr}\left(\xi Y^{-1} \cdot M\right) \\
= & \sum_{n=1}^{d} r^{d-n}\left(2bde_n\left(X\right) + a\left(2p - 2r\frac{b}{a} - n + 1\right)\left(d - n + 1\right) e_{n-1}\left(X\right)\right) \cdot Id + r^d 2bde_0\left(X\right) \\
= & \sum_{n=1}^{d} r^{d-n}\left(2bde_n\left(X\right) + a\left(2p - n + 1\right)\left(d - n + 1\right) e_{n-1}\left(X\right)\right) \cdot Id \\
+ & r^d 2bde_0\left(X\right) - 2b\sum_{n=1}^{d} r^{d-n+1}\left(d - n + 1\right) e_{n-1}\left(X\right) \cdot Id \\
= & \sum_{n=1}^{d} r^{d-n}\left(2bde_n\left(X\right) + a\left(2p - n + 1\right)\left(d - n + 1\right) e_{n-1}\left(X\right)\right) \cdot Id \\
+ & r^d 2bde_0\left(X\right) - 2b\sum_{n=0}^{d} r^{d-n}\left(d - n\right) e_n\left(X\right) \cdot Id \\
= & \sum_{n=1}^{d} r^{d-n}\left(2bne_n\left(X\right) + a\left(2p - n + 1\right)\left(d - n + 1\right) e_{n-1}\left(X\right)\right) \cdot Id.
\end{aligned}$$

La última afirmación ahora se sigue de comparar los coeficientes de cada potencia de r junto con el análisis que presentaremos respecto de la parte matringala local. Para lidiar con la parte que es martingala local, bastará utilizar el lema 2.13, comprobando que las variaciones y covariaciones cuadráticas son absolutamente continuas. Se sigue que

$$\begin{aligned}
& [\xi, \eta] = \left[\mathrm{tr}\left(\xi Y^{-1} \cdot M\right), \mathrm{tr}\left(\eta Z^{-1} \cdot M\right)\right] = \sum_{j,l=1}^{d}\left[\left(\xi Y^{-1} \cdot M\right)^{jj}, \left(\eta Z^{-1} \cdot M\right)^{ll}\right] \\
= & \sum_{j,l=1}^{d}\left[\sum_{i=1}^{d} \xi\left(Y^{-1}\right)^{ji} \cdot M^{ij}, \sum_{k=1}^{d} \eta\left(Z^{-1}\right)^{lk} \cdot M^{kl}\right] \\
= & \sum_{ijkl} \xi\eta\left(Y^{-1}\right)^{ji}\left(Z^{-1}\right)^{lk} \cdot \left[M^{ij}, M^{kl}\right] \\
= & \sum_{ijkl} \xi\eta\left(Y^{-1}\right)^{ji}\left(Z^{-1}\right)^{lk}\left(X^{ik}\alpha_{jl} + X^{il}\alpha_{jk} + X^{jk}\alpha_{il} + X^{jl}\alpha_{ik}\right) \cdot Id \\
= & 4\xi\eta\mathrm{tr}\left(Y^{-1} X Z^{-1} \alpha\right) \cdot Id,
\end{aligned}$$

lo cuál se puede simplificar más en el caso especial en que $\alpha = aI_d$, justo como hicimos anteriormente. En ese caso, si $X = P\Lambda P'$ es una diagonalización de X, entonces

$$Y^{-1}XZ^{-1} = P\left(\Lambda + rI_d\right)^{-1}P'P\Lambda P'P\left(X + sI_d\right)^{-1}P' = P\left(\Lambda + rI_d\right)^{-1}\Lambda\left(\Lambda + sI_d\right)^{-1}P',$$

y por lo tanto

$$\operatorname{tr}\left(Y^{-1}XZ^{-1}\right) = \operatorname{tr}\left(\left(\Lambda + rI_d\right)^{-1}\Lambda\left(\Lambda + sI_d\right)^{-1}\right) = \sum_{i=1}^{d}\frac{\lambda_i\left(X\right)}{\left(\lambda_i\left(X\right) + r\right)\left(\lambda_i\left(X\right) + s\right)}.$$

De aquí se sigue que

$$\begin{aligned}
[\xi, \eta] &= 4\sum_{i=1}^{d}\lambda_i\left(X\right)\left(\prod_{j\neq i}\left(\lambda_j\left(X\right) + r\right)\right)\left(\prod_{j\neq i}\left(\lambda_j\left(X\right) + s\right)\right)\cdot Id\\
&= 4\sum_{i=1}^{d}\lambda_i\left(X\right)\left(\sum_{n=1}^{d}r^{d-n}e_{n-1}^{i}\left(X\right)\right)\left(\sum_{m=1}^{d}s^{d-n}e_{m-1}^{i}\left(X\right)\right)\cdot Id\\
&= 4\sum_{n,m=1}^{d}r^{d-n}s^{d-m}\sum_{i=1}^{d}\lambda_i\left(X\right)e_{n-1}^{i}\left(X\right)e_{m-1}^{i}\left(X\right)\cdot Id.
\end{aligned}$$

Por otro lado,

$$[\xi, \eta] = \sum_{n=0}^{d}r^{d-n}s^{d-m}\left[e_n\left(X\right), e_m\left(X\right)\right],$$

y por lo tanto

$$\left[e_n\left(X\right), e_m\left(X\right)\right] = \sum_{i=1}^{d}\lambda_i\left(X\right)e_{n-1}^{i}\left(X\right)e_{m-1}^{i}\left(X\right)\cdot Id,$$

lo que prueba que las covariaciones son absolutamente continuas y termina la prueba. $\square$

Observación 2.69. En particular, en el caso $(\beta, \alpha) = (0, I_d)$ podemos observar que $\left[e_n\left(X\right)\right] = \sum_i\lambda_i\left(X\right)e_{n-1}^{i}\left(X\right)^2\cdot Id$, i.e., $V^n = \sqrt{\sum_i\lambda_i\left(X\right)e_{n-1}^{i}\left(X\right)^2}$ y que la deriva de $e_n\left(X\right)$ es un múltiplo de $e_{n-1}\left(X\right)\cdot Id$. Por lo tanto e_n se mantiene constante si y sólo si el rango de X se mantiene menor o igual a $d-n+1$, es decir, si y sólo si e_{n-1} se anula. Además, aunque podemos obtener representaciones relativamente simples para el caso $(\beta, \alpha) = (bI_d, aI_d)$, en el caso general esto no es tan sencillo puesto que la multiplicación por las matrices α y β, en general distorsionan la dinámica de los valores propios y es complicado seguirla cuando la matriz Y se mueve en el tiempo.

Además, estos mismos casos ilustran apropiadamente cómo las condiciones de deriva de los polinomios simétricos elementales pueden cambiar drásticamente de acuerdo a β, agregando factores positivos o negativos a la deriva, de acuerdo con el signo de b, que parecen propiciar el que el 0 sea o no polar para algunos de estos polinomios. Sin embargo, tomando logaritmos uno puede percatarse de que este cambio en la deriva se vuelve lineal y por lo tanto no puede, por sí solo, causar emplosión en tiempo finito en el logaritmo; i.e., no puede hacer que el polinomio se anule. Finalmente hacemos la observación de que β sigue sin perturbar la dinámica del proceso más allá del papel que juega en la deriva, incluso en el caso más general.

Es un resultado conocido que los valores propios dependen suavemente de las perturbaciones en las matrices simétricas hasta el primer punto de contacto entre ellos, en cuyo caso se pierde la diferenciabilidad aunque se preserva la continuidad Lipschitz como función de las matrices (cf. [13]). Es natural pensar que entonces los valores propios $\lambda_i(X)$ también siguen un sistema de ecuaciones diferenciales estocásticas hasta el primer tiempo de colisión entre dos de ellos. Esto efectivamente es cierto, y lo mismo es cierto para el proceso de vectores propios correspondientes. A continuación presentaremos una serie de resultados relacionados con las EDE que satisfacen los valores y vectores propios que mencionamos, así como criterios para determinar cuándo hay o no colisión entre los valores propios en tiempo finito.

Como requeriremos adelante algunas simplificaciones, haremos algunos de los cálculos algebraicos de manera anticipada. Por ahora supongamos que $A = P\Lambda P'$ es una diagonalización de la matriz A que tiene todos sus valores propios distintos, entonces tenemos que

$$\begin{aligned} \det(A - rI_d)\, I_d &= \operatorname{co}(A - rI_d)(A - rI_d) \\ &= P\operatorname{co}(A - rI_d)\, P'P(A - rI_d)\, P \\ &= P\operatorname{co}(A - rI_d)\, P'(\Lambda - rI_d), \end{aligned}$$

lo que significa que cuando r no es un valor propio, $P\operatorname{co}(A - rI_d)\, P'$ es una matriz diagonal y su entrada (i,i) está dada por

$$\frac{\det(A - rI_d)}{\lambda_i(A) - r} = \frac{\prod_{j=1}^{d}(\lambda_j(A) - r)}{\lambda_i(A) - r} = \prod_{j\neq i}(\lambda_j(A) - r).$$

Este resultado de hecho implica que

$$\operatorname{tr}(\operatorname{co}(A - rI_d)) = \operatorname{tr}\left(P\operatorname{co}(A - rI_d)\, P'\right) = \sum_{i=1}^{d}\prod_{j\neq i}(\lambda_j(A) - r),$$

de tal forma que si perturbamos a A de forma continua, la continuidad del determinante (y por lo tanto, de la matriz de cofactores) y de la traza implican que

$$\operatorname{tr}(\operatorname{co}(A - \lambda_i(A)\, I_d)) = \sum_{j=1}^{d}\prod_{k\neq j}(\lambda_k(A) - \lambda_i(A)) = \prod_{k\neq i}(\lambda_k(A) - \lambda_i(A)).$$

Similarmente obtenemos para $r, s \notin \sigma(A)$,

$$\begin{aligned}\operatorname{tr}\left(\operatorname{co}\left(A-rI_d\right)\operatorname{co}\left(A-sI_d\right)\right) &= \operatorname{tr}\left(P\operatorname{co}\left(A-rI_d\right)P'P\operatorname{co}\left(A-sI_d\right)P'\right)\\ &= \sum_{i=1}^{d}\prod_{k\neq i}\left(\lambda_k(A)-r\right)\left(\lambda_k(A)-s\right),\end{aligned}$$

y así obtenemos

$$\begin{aligned}\operatorname{tr}\left(\operatorname{co}\left(A-\lambda_i(A)I_d\right)\operatorname{co}\left(A-\lambda_j(A)I_d\right)\right) &= \sum_{k=1}^{d}\prod_{l\neq k}\left(\lambda_l(A)-\lambda_i(A)\right)\left(\lambda_l(A)-\lambda_j(A)\right)\\ &= \delta_{ij}\prod_{l\neq i}\left(\lambda_l(A)-\lambda_i(A)\right)^2.\end{aligned}$$

De hecho, si $i \neq j$, entonces $P\operatorname{co}\left(A-\lambda_i(A)I_d\right)\operatorname{co}\left(A-\lambda_j(A)I_d\right)P'$ es diagonal con entradas diagonales 0, es decir que $\operatorname{co}\left(A-\lambda_i(A)I_d\right)\operatorname{co}\left(A-\lambda_j(A)I_d\right)=0$, mientras que si $i=j$, entonces esta matriz diagonal tiene exactamente una entrada distinta de 0, a saber, la i-ésima, con el valor de $\prod_{l\neq i}\left(\lambda_l(A)-\lambda_i(A)\right)^2$. Otros resultados muy similares se pueden derivar con facilidad extendiendo esta idea.

En lo que sigue, denotaremos al primer tiempo en que dos valores propios se igualan por $\tau = \inf\left\{t\geq 0:\ \lambda_t^i=\lambda_t^j,\quad i\neq j\right\}$. Como se puede verificar en [13], la diagonalización y los valores propios (ordenados) de matrices simétricas, son funciones suaves como funciones de éstas, es decir que por la fórmula de Itô, si X inicia en un estado con valores propios distintos, entonces hasta el tiempo $\tau>0$, tanto sus valores propios como su matriz de vectores propios son semimartingalas. Para averiguar las EDE que satisfacen estos últimos, dejamos el siguiente resultado.

Proposición 2.70. *Supongamos que $x\in S_d^+$ tiene todos sus valores propios distintos y denotemos por $X=P\Lambda P'$ a la diagonalización de X con $\Lambda^{ii}=\lambda^i=\lambda_i(X)$. Entonces en $[0,\tau)$ tenemos*

$$\begin{aligned}\lambda^i-\lambda_0^i &= 2\sqrt{\lambda^i\left(P'\alpha P\right)^{ii}}\cdot W^i+2\left(p\left(P'\alpha P\right)^{ii}+\lambda^i\left(P'\beta'P\right)^{ii}\right)\cdot Id\\ &+ \sum_{k\neq i}\frac{\lambda^i\left(P'\alpha P\right)^{kk}+\lambda^k\left(P'\alpha P\right)^{ii}}{\lambda^i-\lambda^k}\cdot Id,\end{aligned}$$

para movimientos brownianos independientes $\left(W^i\right)$ y

$$\begin{aligned}dP_t^{ij} &= \sum_{k\neq j}P_t^{ik}\frac{\sqrt{\lambda_t^k\left(P_t'\alpha P_t\right)^{jj}+\lambda^j\left(P_t'\alpha P_t\right)^{kk}}}{\lambda_t^j-\lambda_t^k}dV_t^{kj}\\ &- \frac{1}{2}\sum_k P_t^{ik}\sum_{l\neq k,j}\frac{\delta_{kj}\lambda_t^k\left(P_t'\alpha P_t\right)^{ll}+\lambda_t^l\left(P_t'\alpha P_t\right)^{kj}}{\left(\lambda_t^l-\lambda^k\right)\left(\lambda^l-\lambda^j\right)}dt.\end{aligned}$$

para movimientos brownianos $\left(V^{ij}\right)_{i<j}$ *con* $V^{ji}=V^{ij}$ *que satisfacen (2.21) y (2.22). En el caso particular en que* $\alpha = aI_d$ *para* $a \geq 0$*, tenemos*

$$dP_t^{ij} = \sqrt{a}\sum_{k\neq j} P_t^{ik}\frac{\sqrt{\lambda_t^k+\lambda^j}}{\lambda_t^j-\lambda_t^k}dV_t^{kj} - \frac{1}{2}P_t^{ij}a\sum_{k\neq j}\frac{\lambda_t^j+\lambda_t^k}{\left(\lambda_t^k-\lambda_t^j\right)^2}dt,$$

donde los movimientos brownianos$\left(V^{ij}\right)_{i<j}$ *son independientes entre sí e independientes de* $\left(W^i\right)$*. En ese mismo caso*

$$\lambda^i-\lambda_0^i = 2\sqrt{a\lambda^i}\cdot W^i + 2\left(ap+\lambda^i\left(P'\beta'P\right)^{ii}\right)\cdot Id + a\sum_{k\neq i}\frac{\lambda^i+\lambda^k}{\lambda^i-\lambda^k}\cdot Id,$$

y si además tenemos $\beta = bI_d$*, entonces la identidad*

$$\lambda^i-\lambda_0^i = 2\sqrt{a\lambda^i}\cdot W^i + \left(2ap+b\lambda^i\right)\cdot Id + a\sum_{k\neq i}\frac{\lambda^i+\lambda^k}{\lambda^i-\lambda^k}\cdot Id,$$

se satisface para cada $i=1,\ldots,d$.

Demostración. En general seguiremos el procedimiento usual de la prueba de las fórmulas de segunda variación de Hadamard. Todos los resultados que se presentarán se satisfacen sólo hasta el tiempo aleatorio τ, que por continuidad de los valores propios, resulta ser un tiempo de paro. Si definimos al proceso A como el logaritmo estocástico de P, esto es, al proceso

$$A_t = \int_0^t P_s'\circ dP_s = \int_0^t P_s'dP_s + \frac{1}{2}\int_0^t dP_s'dP_s$$

y observemos que la fórmula de integración por partes aplicada a $I_d = P'P$ nos da

$$\begin{aligned} 0 &= dP_t'\circ P_t + P_t'\circ dP_t \\ -dA_t &= dP_t'\circ P_t = dA_t', \end{aligned}$$

es decir que A es antisimétrica y por lo tanto tiene cero en la diagonal. Definiendo $N_t = \int_0^t P_s'\circ dX_s\circ P_s$ y usando la fórmula de integración por partes para $\Lambda = P'XP$, obtenemos

$$\begin{aligned} d\Lambda_t &= dP_t'\circ(X_tP_t) + P_t'\circ dX_t\circ P_t + \left(P_t'X_t\right)\circ dP_t = dP_t'\circ(P_t\Lambda_t) + dN_t + \left(\Lambda_tP_t'\right)\circ dP_t \\ &= dA_t'\circ\Lambda_t + \Lambda_t\circ dA_t + dN_t = -dA_t\circ\Lambda_t + \Lambda_t\circ dA_t + dN_t, \end{aligned}$$

donde $\int_0^t \Lambda_s\circ dA_s - dA_s\circ\Lambda_s$ es un proceso que tiene 0 en la diagonal. Se sigue que

$$\begin{aligned} d\lambda_t^i &= d\Lambda_t^{ii} = dN_t^{ii} = \left(P_t'\circ dX_t\circ P_t\right)^{ii}, \\ 0 &= dN_t^{ij} + \left(\lambda_t^i-\lambda_t^j\right)\circ dA_t^{ij}, \quad i\neq j, \end{aligned}$$

y por lo tanto, tenemos que

$$dA_t^{ij} = \frac{1}{\lambda_t^j - \lambda_t^i} dN_t^{ij}, \quad i \neq j. \tag{2.18}$$

Además, como la integral de Stratonovich no altera la parte martingala local, tenemos que

$$\begin{aligned} d\left[N^{ij}, N^{i'j'}\right]_t &= \left(P_t' dM_t P_t\right)^{ij} \left(P_t' dM_t P_t\right)^{i'j'} = \sum_{klk'l'} P_t^{ki} dM_t^{kl} P_t^{lj} P_t^{k'i'} dM_t^{k'l'} P_t^{l'j'} \\ &= \sum_{klk'l'} P_t^{ki} P_t^{lj} P_t^{k'i'} P_t^{l'j'} \left(X_t^{kk'} \alpha_{ll'} + X_t^{kl'} \alpha_{lk'} + X_t^{lk'} \alpha_{kl'} + X_t^{ll'} \alpha_{kk'}\right) dt \\ &= \Big(\left(P_t' X_t P_t\right)^{ii'} \left(P_t' \alpha P_t\right)^{jj'} + \left(P_t' X_t P_t\right)^{ij'} \left(P_t' \alpha P_t\right)^{ji'} \\ &+ \left(P_t' X_t P_t\right)^{ji'} \left(P_t' \alpha P_t\right)^{ij'} + \left(P_t' X_t P_t\right)^{jj'} \left(P_t' \alpha P_t\right)^{ii'}\Big) dt \\ &= \left(\Lambda_t^{ii'} \left(P_t' \alpha P_t\right)^{jj'} + \Lambda_t^{ij'} \left(P_t' \alpha P_t\right)^{ji'} + \Lambda_t^{ji'} \left(P_t' \alpha P_t\right)^{ij'} + \Lambda_t^{jj'} \left(P_t' \alpha P_t\right)^{ii'}\right) dt, \end{aligned} \tag{2.19}$$

en particular tenemos

$$d\left[\lambda^i, \lambda^j\right]_t = d\left[N^{ii}, N^{jj}\right]_t = 4\Lambda_t^{ij} \left(P_t' \alpha P_t\right)^{ij} dt = 4\delta_{ij} \lambda_t^i \left(P_t' \alpha P_t\right)^{ii} dt. \tag{2.20}$$

Por otro lado, la parte de variación finita en compactos de N satisface

$$\begin{aligned} & P_t' dD_t P_t + \frac{1}{2}\left(dP_t' dM_t P_t + P_t' dM_t dP_t\right) \\ = \; & P_t' \left(2p\alpha + \beta X_t + X_t \beta'\right) P_t dt + \frac{1}{2}\left(dP_t' P_t P_t' dM_t P_t + P_t' dM_t P_t P_t' dP_t\right) \\ = \; & \left(2p P_t' \alpha P_t + P_t' \beta P_t P_t' X_t P_t + P_t' X_t P_t P_t' \beta' P_t\right) dt + \frac{1}{2}\left(dP_t' P_t P_t' dM_t P_t + P_t' dM_t P_t P_t' dP_t\right) \\ = \; & \left(2p P_t' \alpha P_t + P_t' \beta P_t \Lambda_t + \Lambda_t P_t' \beta' P_t\right) dt + \frac{1}{2}\Big(\left(dN_t dA_t\right)' + dN_t dA_t\Big). \end{aligned}$$

Se sigue de la antisimetría de A, el que Λ es diagonal y de las ecuaciones (2.18) y (2.19) que

$$\begin{aligned} \left(dN_t dA_t\right)^{ij} &= \sum_{k \neq j} dN_t^{ik} dA_t^{kj} = \sum_{k \neq j} \frac{dN_t^{ik} dN_t^{kj}}{\lambda_t^j - \lambda_t^k} \\ &= \sum_{k \neq j} \frac{\Lambda_t^{ik} \left(P_t' \alpha P_t\right)^{kj} + \Lambda_t^{ij} \left(P_t' \alpha P_t\right)^{kk} + \Lambda_t^{kk} \left(P_t' \alpha P_t\right)^{ij} + \Lambda_t^{kj} \left(P_t' \alpha P_t\right)^{ik}}{\lambda_t^j - \lambda_t^k} dt \\ &= \left(\frac{\lambda_t^i \left(P_t' \alpha P_t\right)^{ij}}{\lambda_t^j - \lambda_t^i} \left(1 - \delta_{ij}\right) + \delta_{ij} \lambda_t^i \sum_{k \neq i} \frac{\left(P_t' \alpha P_t\right)^{kk}}{\lambda_t^i - \lambda_t^k} + \sum_{k \neq j} \frac{\lambda_t^k}{\lambda_t^j - \lambda_t^k} \left(P_t' \alpha P_t\right)^{ij}\right) dt, \end{aligned}$$

donde δ_{ij} es la delta de Krónecker. Así concluimos que la deriva de λ^i es

$$2\left(p \left(P_t' \alpha P_t\right)^{ii} + \lambda_t^i \left(P_t' \beta' P_t\right)^{ii}\right) dt + \sum_{k \neq i} \frac{\lambda_t^i \left(P_t' \alpha P_t\right)^{kk} + \lambda_t^k \left(P_t' \alpha P_t\right)^{ii}}{\lambda_t^i - \lambda_t^k} dt.$$

Así, gracias a (2.20) y a la ecuación anterior, podemos usar la representación integral de Doob (lema 2.13) para argumentar la existencia de movimientos brownianos reales independientes W^i tales que

$$\begin{aligned}\lambda^i - \lambda_0^i &= 2\sqrt{\lambda^i \left(P'\alpha P\right)^{ii}} \cdot W^i + 2\left(p\left(P'\alpha P\right)^{ii} + \lambda^i \left(P'\beta' P\right)^{ii}\right) \cdot Id \\ &+ \sum_{k\neq i} \frac{\lambda^i \left(P'\alpha P\right)^{kk} + \lambda^k \left(P'\alpha P\right)^{ii}}{\lambda^i - \lambda^k} \cdot Id.\end{aligned}$$

Ahora bien, de la definición de A, está claro que

$$\begin{aligned}P_t \circ dA_t &= P_t dA_t + \frac{1}{2} dP_t dA_t \\ &= P_t P_t' dP_t + \frac{1}{2} P_t dP_t' dP_t + \frac{1}{2} dP_t dA_t \\ &= dP_t + \frac{1}{2} dA_t' dP_t + \frac{1}{2} dP_t dA_t = dP_t,\end{aligned}$$

además de que para $i \neq j$ tenemos

$$\left[N^{ij}, N^{ij}\right] = \left(\lambda^i \left(P'\alpha P\right)^{jj} + \lambda^j \left(P'\alpha P\right)^{ii}\right) \cdot Id.$$

Si además asumimos que $\alpha = aI_d$, entonces se sigue de (2.19) que $\left[N^{ij}, N^{kl}\right] = 0$ cuando $(i,j) \neq (k,l)$. Concluimos de (2.18) y de la representación integral de Doob que existen movimientos brownianos $\left(V^{ij}\right)_{i<j}$ tales que si definimos $V^{ij} := V^{ji}$ para $i > j$, entonces para $i \neq j$ tenemos

$$A^{ij} = \frac{\sqrt{\lambda^i \left(P'\alpha P\right)^{jj} + \lambda^j \left(P'\alpha P\right)^{ii}}}{\lambda^j - \lambda^i} \cdot V^{ij}.$$

En el caso particular en que $\alpha = aI_d$, inferimos independencia en los movimientos brownianos $\left(V^{ij}\right)_{i<j}$ de (2.18), además, la expresión se simplifica a $A^{ij} = \sqrt{a}\frac{\sqrt{\lambda^i+\lambda^j}}{\lambda^j-\lambda^i} \cdot V^{ij}$; sino, simplemente sabemos que

$$\left[V^{ij}, V^{kl}\right] = \frac{\left(\Lambda^{ik} \left(P'\alpha P\right)^{jl} + \Lambda^{il} \left(P'\alpha P\right)^{jk} + \Lambda^{jk} \left(P'\alpha P\right)^{il} + \Lambda^{jl} \left(P'\alpha P\right)^{ik}\right)}{\sqrt{\lambda^i \left(P'\alpha P\right)^{jj} + \lambda^j \left(P'\alpha P\right)^{ii}}\sqrt{\lambda^k \left(P'\alpha P\right)^{ll} + \lambda^l \left(P'\alpha P\right)^{kk}}} \cdot Id. \quad (2.21)$$

Por otro lado, por (2.19) tenemos para $j \neq k$,

$$\begin{aligned}\left[\lambda^i, A^{jk}\right] &= \frac{\left[N^{ii}, N^{jk}\right]}{\lambda^k - \lambda^j} = \frac{2\lambda^i \left(\delta_{ij} \left(P'\alpha P\right)^{ik} + \delta_{ik} \left(P'\alpha P\right)^{ij}\right)}{\lambda^k - \lambda^j} \cdot Id \\ \left[\lambda^i, A^{jk}\right] &= 2\frac{\sqrt{\lambda^i \left(P'\alpha P\right)^{ii}}\sqrt{\lambda^j \left(P'\alpha P\right)^{kk} + \lambda^k \left(P'\alpha P\right)^{jj}}}{\lambda^k - \lambda^j} \cdot \left[W^i, V^{jk}\right],\end{aligned}$$

es decir que

$$\left[W^i, V^{jk}\right] = \frac{\sqrt{\lambda^i}\left(\delta_{ij}\left(P'\alpha P\right)^{ik} + \delta_{ik}\left(P'\alpha P\right)^{ij}\right)}{\sqrt{\left(P'\alpha P\right)^{ii}}\sqrt{\lambda^j\left(P'\alpha P\right)^{kk} + \lambda^k\left(P'\alpha P\right)^{jj}}} \cdot Id. \tag{2.22}$$

Esto implica que los movimientos brownianos $\left(V^{ij}\right)_{i<j}$ y $\left(W^i\right)$ son independientes cuando $\alpha = aI_d$. En este mismo caso, la independencia de $\left(V^{ij}\right)_{i<j}$ implica que $\int_0^t dA_s dA_s$ es diagonal y la antisimetría de A nos dice que

$$\left(dA_t dA_t\right)^{ii} = -\sum_{k\neq i} dA_t^{ik} dA_t^{ik} = -a\sum_{k\neq i}\frac{\lambda^i + \lambda^k}{\left(\lambda^k - \lambda^i\right)^2} \cdot Id.$$

En el caso contrario sólo obtenemos que

$$\begin{aligned}
\left(dA_t dA_t\right)^{ij} &= -\sum_{k\neq i,j} dA_t^{ik} dA_t^{jk} = -\sum_{k\neq i,j}\frac{1}{\left(\lambda_t^k - \lambda_t^i\right)\left(\lambda_t^k - \lambda_t^j\right)} d\left[N^{ik}, N^{jk}\right]_t \\
&= -\sum_{k\neq i,j}\frac{\delta_{ij}\lambda^i\left(P_t'\alpha P_t\right)^{kk} + \lambda_t^k\left(P_t'\alpha P_t\right)^{ij}}{\left(\lambda_t^k - \lambda_t^i\right)\left(\lambda_t^k - \lambda_t^j\right)} dt.
\end{aligned}$$

Así podemos usar el que

$$dP_t = P_t dA_t + \frac{1}{2} dP_t dA_t = P_t dA_t + \frac{1}{2} P_t dA_t dA_t,$$

para calcular

$$\begin{aligned}
dP_t^{ij} &= \sum_{k\neq j} P_t^{ik} dA_t^{kj} + \frac{1}{2}\sum_k P_t^{ik}\left(dA_t dA_t\right)^{kj} \\
&= \sum_{k\neq j} P_t^{ik}\frac{\sqrt{\lambda_t^k\left(P_t'\alpha P_t\right)^{jj} + \lambda^j\left(P_t'\alpha P_t\right)^{kk}}}{\lambda_t^j - \lambda_t^k} dV_t^{kj} \\
&- \frac{1}{2}\sum_k P_t^{ik}\sum_{l\neq k,j}\frac{\delta_{kj}\lambda_t^k\left(P_t'\alpha P_t\right)^{ll} + \lambda_t^l\left(P_t'\alpha P_t\right)^{kj}}{\left(\lambda_t^l - \lambda^k\right)\left(\lambda^l - \lambda^j\right)} dt,
\end{aligned}$$

que en el caso $\alpha = aI_d$ se simplifica a

$$dP_t^{ij} = \sqrt{a}\sum_{k\neq j} P_t^{ik}\frac{\sqrt{\lambda_t^k + \lambda^j}}{\lambda_t^j - \lambda_t^k} dV_t^{kj} - \frac{1}{2} P_t^{ij} a\sum_{k\neq j}\frac{\lambda_t^j + \lambda_t^k}{\left(\lambda_t^k - \lambda_t^j\right)^2} dt.$$

Las simplificaciones de las EDE de los valores propios se sigue directamente. $\square$

Observación 2.71. Algo verdaderamente interesante es que el factor de deriva β no altera directamente la dinámica estocástica de P, sino sólo a través de la deriva de los valores propios, exactamente en el término $2\lambda^i\left(P'\beta' P\right)^{ii}$. Además, como es usual, $\alpha = aI_d$ nos permite inferir una estructura de independencias en todos lados, lo que facilita la simulación y comprensión de la situación.

Definamos al operador matricial E_n para $n = 1, \dots, \frac{d(d-1)}{2}$ que a cada matriz en $M_{d,d}$ le asigna el polinomio simétrico elemental e_n aplicado al vector compuesto por todas las entradas en o por encima de la diagonal principal. De esta forma, E_n^{ij} corresponde a quitar sólo el término (i,j) de la matriz antes de formar el vector y aplicar e_n y similarmente se define E_n^{ijkl}.

Supongamos que x tiene todos sus valores propios distintos y que $X \sim WP_d\left(p, I_d, 0, x\right)$. Definamos al proceso A dado por $A^{ij} = \left(\lambda^i - \lambda^j\right)^2$ y a los procesos $L^n = E_n\left(A\right)$, entonces es claro que

$$L^{\frac{d(d-1)}{2}} = \prod_{1<j}\left(\lambda^i - \lambda^j\right)^2, \quad L^1 = \sum_{i<j}\left(\lambda^i - \lambda^j\right)^2,$$

y en particular tenemos $\tau = \inf\left\{t > 0 : L_t^{\frac{d(d-1)}{2}} = 0\right\}$. Podemos obtener la EDE que satisfacen los L^n, que son los responsables de llevar la cuenta del número de colisiones entre valores propios.

El que los valores propios sean funciones suaves de X cuando no nos preocupamos por el orden que preservan los valores propios (cf. [13]) muestra que la simetría (y por lo tanto invariancia ante permutaciones) en la definición de los procesos L^n implicará que las EDE que estos satisfagan no van a depender de si hay colisiones o no, puesto que estas serán funciones suaves de X. Sin embargo, para encontrar estas, es mucho más sencillo partir del resultado anterior.

Como $E_n\left(A\right) = A^{ij}E_{n-1}^{ij}\left(A\right) + E_n^{ij}\left(A\right)$ y $E_n^{ij}\left(A\right) = A^{kl}E_{n-1}^{ijkl}\left(A\right) + E_n^{ijkl}\left(A\right)$ mientras $\{i,j\} \neq \{k,l\}$, entonces la fórmula de Itô implica que

$$dL_t^n = \sum_{i<j} E_{n-1}^{ij}\left(A_t\right) dA_t^{ij} + \frac{1}{2}\sum_{i<j} \sum_{\substack{k<l \\ \{i,j\}\neq\{k,l\}}} E_{n-2}^{ijkl}\left(A_t\right) d\left[A^{ij}, A^{kl}\right]_t.$$

Ahora observemos que $\left[A^{ij}, A^{kl}\right] \not\equiv 0$ si y sólo si $\{i,j\} \cap \{k,l\} \neq \emptyset$. De hecho, si $i < j$

$$\begin{aligned} dA_t^{ij} &= 2\left(\lambda_t^i - \lambda_t^j\right) d\lambda_t^i + 2\left(\lambda_t^j - \lambda_t^i\right) d\lambda_t^j + d\left[\lambda^i\right]_t + d\left[\lambda^j\right]_t \\ &= 4\left(\lambda_t^i - \lambda_t^j\right)\sqrt{\lambda_t^i} dW_t^i + 4\left(\lambda_t^j - \lambda_t^i\right)\sqrt{\lambda_t^j} dW_t^j \\ &+ \left(8\left(\lambda_t^i + \lambda_t^j\right) + 2\left(\lambda_t^i - \lambda_t^j\right)\sum_{k\neq i,j}\frac{\lambda_t^i + \lambda_t^k}{\lambda_t^i - \lambda_t^k} - \frac{\lambda_t^j + \lambda_t^k}{\lambda_t^j - \lambda_t^k}\right) dt, \end{aligned}$$

lo que nos da

$$
\begin{aligned}
d\left[A^{ij},A^{ik}\right]_t &= 16\lambda_t^i\left(\lambda_t^i-\lambda_t^j\right)\left(\lambda_t^i-\lambda_t^k\right)dt, \quad i\neq j\neq k\neq i,\\
d\left[A^{ij},A^{ij}\right]_t &= 16\left(\lambda_t^i+\lambda_t^j\right)\left(\lambda_t^j-\lambda_t^i\right)^2 dt, \quad i\neq j.
\end{aligned}
$$

En consecuencia, la parte martingala de L^n está dada por

$$
\begin{aligned}
dM_t^n &= 4\sum_{i<j}\left(\lambda_t^i-\lambda_t^j\right)E_{n-1}^{ij}\left(A_t\right)\left(\sqrt{\lambda_t^i}dW_t^i-\sqrt{\lambda_t^j}dW_t^j\right)\\
&= 4\sum_{i=1}^{d}\sqrt{\lambda_t^i}\sum_{j\neq i}\left(\lambda_t^i-\lambda_t^j\right)E_{n-1}^{ij}\left(A_t\right)dW_t^i,\\
d\left[M^n\right]_t = d\left[L^n\right]_t &= 16\sum_{i=1}^{d}\lambda_t^i\left(\sum_{j\neq i}\left(\lambda_t^i-\lambda_t^j\right)E_{n-1}^{ij}\left(A_t\right)\right)^2 dt.
\end{aligned}
$$

Del mismo modo, la parte de variación finita en compactos de L^n es

$$
\begin{aligned}
dN_t^n &= \sum_{i<j}E_{n-1}^{ij}\left(A_t\right)\left(8\left(\lambda_t^i+\lambda_t^j\right)+2\left(\lambda_t^i-\lambda_t^j\right)\sum_{k\neq i,j}\frac{\lambda_t^i+\lambda_t^k}{\lambda_t^i-\lambda_t^k}-\frac{\lambda_t^j+\lambda_t^k}{\lambda_t^j-\lambda_t^k}\right)dt\\
&+ 8\sum_{i<j}\left(\lambda_t^i-\lambda_t^j\right)\left(\sum_{k\neq i,j}E_{n-2}^{ijik}\left(A_t\right)\lambda_t^i\left(\lambda_t^i-\lambda_t^k\right)-E_{n-2}^{ijkj}\left(A_t\right)\lambda_t^j\left(\lambda_t^j-\lambda_t^k\right)\right)dt,
\end{aligned}
$$

donde fácilmente observamos que la deriva no explota ni se indetermina porque cada factor $\lambda_t^i-\lambda_t^k$ y $\lambda_t^j-\lambda_t^l$ está contenido en $E_{n-1}^{ij}\left(A_t\right)$ cuando $k\neq j$ y $l\neq i$ respectivamente o en el coeficiente $\lambda_t^i-\lambda_t^j$ en otro caso. El que estas propiedades no dependan de los valores iniciales posiblemente degenerados se sigue de que V_n es una función suave y analítica de X, lo que implica qu podemos extender analíticamente las derivadas de todos los órdenes a estos puntos excepcionales. Por lo tanto, la fórmula de Itô nos ayuda a completar el argumento, por unicidad de la parte martingala local y de variación finita en compactos.

Para estudiar este comportamiento en el caso en que los valores propios de x no son distintos, veremos que de hecho tenemos una difracción inmediata, es decir que hay un momento inmediatamente después del tiempo inicial en el que todos son distintos y así, el lema anterior implicará que nunca más se volverán a tocar. Esto es, los valores propios son distintos para todo tiempo $t>0$.

Proposición 2.72. *Supongamos que $X\sim WP_d\left(p,I_d,0,x\right)$ donde $x\in\bar{S}_d^+$ satisface $rnk\left(x\right)\geq d-1$ y definamos $\tau_n=\inf\left\{t>0:L_t^n>0\right\}$ para $n=1,\ldots,\frac{d(d-1)}{2}$, entonces tenemos $\tau_n=0$ c.s. para cada n. De este modo, los resultados de la proposición anterior se extienden al caso en que $x\in S_d^+$ o incluso $rnk\left(x\right)\geq d-1$.*

Demostración. Como $(0,\infty)$ es abierto, si $L_0^{\frac{d(d-1)}{2}} > 0$ entonces el resultado es claro por continuidad de las trayectorias de cada L^n, así que supondremos que $L_0^{\frac{d(d-1)}{2}} = 0$. Notemos que las colisiones no pueden iniciar en el estado 0 (i.e., no podemos tener $\lambda_0^i = \lambda_0^j = 0$ para $i \neq j$) porque x es de rango al menos $d-1$. Dicho de otra forma, estamos asumiendo que $\Lambda = \left(\lambda^1, \ldots, \lambda^d\right)$ inicia en el conjunto semiabierto

$$C = \bigcup_{i \leq d} (0,\infty)^{i-1} \times [0,\infty) \times (0,\infty)^{d-i}.$$

La continuidad de trayectorias implica que existe un tiempo de paro T_C tal que en $[0,T_C)$ el proceso $\Lambda = \Lambda(X)$ se mantiene en C. Inductivamente mostraremos que si $\tau_n > 0$ con probabilidad positiva, lo mismo es cierto para τ_{n-1}. Esto es claro si notamos que $L^n = 0$ en $[0,\tau_n)$ y por lo tanto $\left(\lambda^i - \lambda^j\right) E_{n-1}^{ij}(A) = 0 = \left(\lambda^i - \lambda^j\right)\left(\lambda^k - \lambda^l\right) E_{n-2}^{ijkl}(A)$, ya que todos los sumandos en L^n se anulan. Además, como su parte de variación finita se debe desvanecer, obtenemos

$$N^n \;=\; \sum_{i<j} 8\left(\lambda^i + \lambda^j\right) E_{n-1}^{ij}(A) \cdot Id = 0.$$

Como todas las sumas satisfacen $\lambda_t^i + \lambda_t^j > 0$ para $t < T_C$, entonces concluimos que $E_{n-1}^{ij}(A) = 0$ para todo $i < j$ en $[0,T_C)$ y por lo tanto, lo mismo es cierto para L^{n-1}. Esto muestra que si $\tau_n > 0$ con probabilidad positiva, lo mismo es cierto para τ_{n-1} y por lo tanto, si inicialmente suponemos que $\tau_{\frac{d(d-1)}{2}} > 0$ con probabilidad positiva, entonces $\tau_1 > 0$ con probabilidad positiva.

Esto es, si suponemos que al menos dos partículas permanecen pegadas por un tiempo, entonces todas las partículas permanecen pegadas. Sin embargo, para $n=1$ obtenemos

$$N^1 \;=\; \sum_{i<j} 8\left(\lambda^i + \lambda^j\right) \cdot Id = 0 \implies \lambda^i = 0 \;\forall 1 \leq i \leq d,$$

que es contradictorio por construcción de C, lo que termina la prueba. □

Esto último muestra que las colisiones entre valores propios son bastante raras, salvo en el caso en que éstas se dan en 0, en cuyo caso, como veremos después, estas pueden ser frecuentes (o incluso podemos tener $\lambda^i \equiv 0$ para algunos i) dependiendo del valor que toma p.

Proposición 2.73. *Para $p \in \Delta_d$ y $x \in \bar{S}_d^+$ con valores propios distintos o con $rnk(x) \geq d-1$, el proceso de valores propios $\left(\lambda^1, \ldots, \lambda^d\right)$ de $X \sim WP_d(p, I_d, 0, x)$ es la trayectorialmente* *única solución fuerte del sistema de EDE*

$$d\lambda_t^i = 2\sqrt{\lambda_t^i}dW_t^i + \left(2p + \sum_{k\neq i} \frac{\lambda_t^i + \lambda_t^k}{\lambda_t^i - \lambda_t^k}\right) dt, \quad i = 1, \ldots, d,$$

donde $\left(W^i\right)$ *son movimientos brownianos independientes. Además, los valores propios nunca colisionan en* $t > 0$.

Demostración. La EDE ya la hemos probado anteriormente, así que basta probar la unicidad trayectorial y la no colisión. Para la última afirmación, lo que haremos será primero probarlo para el caso en que x tiene todos sus valores propios distintos y luego para el caso general, así que inicialmente supondremos el primer caso.

Definamos ahora $U = \log L^n$ para $n = \frac{d(d-1)}{2}$ y usemos la fórmula de Itô para obtener que la parte martingala de U es

$$4\sum_{i<j} \frac{\sqrt{\lambda_t^i}dW_t^i - \sqrt{\lambda_t^j}dW_t^j}{\lambda_t^i - \lambda_t^j}.$$

Observemos que en este caso, $\left(\lambda^i - \lambda^j\right)^2 E_{n-1}^{ij}(A) = L^n$ y algo similar sucede con $E_{n-2}^{ijkl}(A)$. Usando estas identidades, se deduce que la parte de variación finita en compactos se puede expresar como

$$\begin{aligned}
& \sum_{i<j} \left(8\frac{\lambda_t^i + \lambda_t^j}{\left(\lambda_t^i - \lambda_t^j\right)^2} + \frac{2}{\lambda_t^i - \lambda_t^j} \sum_{k\neq i,j} \frac{\lambda_t^i + \lambda_t^k}{\lambda_t^i - \lambda_t^k} - \frac{\lambda_t^j + \lambda_t^k}{\lambda_t^j - \lambda_t^k}\right) dt \\
+ \quad & 8\sum_{i<j}\sum_{k\neq i,j} \left(\frac{\lambda_t^i}{\left(\lambda_t^i - \lambda_t^j\right)\left(\lambda_t^i - \lambda_t^k\right)} + \frac{\lambda_t^j}{\left(\lambda_t^j - \lambda_t^i\right)\left(\lambda_t^j - \lambda_t^k\right)}\right) dt - 8\sum_{i=1}^d \lambda_t^i \left(\sum_{j\neq i} \frac{1}{\lambda_t^i - \lambda_t^j}\right)^2 dt,
\end{aligned}$$

donde tenemos

$$\begin{aligned}
& \sum_{i=1}^d \lambda_t^i \left(\sum_{j\neq i} \frac{1}{\lambda_t^i - \lambda_t^j}\right)^2 \\
= \quad & \sum_{i<j} \frac{\lambda_t^i + \lambda_t^j}{\left(\lambda_t^i - \lambda_t^j\right)^2} + 2\sum_{i<j}\sum_{k\neq i,j} \left(\frac{\lambda_t^i}{\left(\lambda_t^i - \lambda_t^j\right)\left(\lambda_t^i - \lambda_t^k\right)} + \frac{\lambda_t^j}{\left(\lambda_t^j - \lambda_t^i\right)\left(\lambda_t^j - \lambda_t^k\right)}\right),
\end{aligned}$$

y también

$$\frac{\lambda_t^i}{\left(\lambda_t^i - \lambda_t^j\right)\left(\lambda_t^i - \lambda_t^k\right)} + \frac{\lambda_t^j}{\left(\lambda_t^j - \lambda_t^i\right)\left(\lambda_t^j - \lambda_t^k\right)} = \frac{\lambda_t^i\left(\lambda_t^j - \lambda_t^k\right) - \left(\lambda_t^i - \lambda_t^k\right)\lambda_t^j}{\left(\lambda_t^i - \lambda_t^j\right)\left(\lambda_t^i - \lambda_t^k\right)\left(\lambda_t^j - \lambda_t^k\right)}.$$

Se sigue de algunas cancelaciones y factorizaciones que la parte de variación finita en compactos se puede representar de la siguiente forma

$$\begin{aligned} & 2\sum_{i<j}\sum_{k\neq i,j}\frac{\left(\lambda_t^i+\lambda_t^k\right)\left(\lambda_t^j-\lambda_t^k\right)-\left(\lambda_t^j+\lambda_t^k\right)\left(\lambda_t^i-\lambda_t^k\right)}{\left(\lambda_t^i-\lambda_t^j\right)\left(\lambda_t^i-\lambda_t^k\right)\left(\lambda_t^j-\lambda_t^k\right)}dt \\ - & \ 8\sum_{i<j}\sum_{k\neq i,j}\left(\frac{\lambda_t^i\left(\lambda_t^j-\lambda_t^k\right)-\left(\lambda_t^i-\lambda_t^k\right)\lambda_t^j}{\left(\lambda_t^i-\lambda_t^j\right)\left(\lambda_t^i-\lambda_t^k\right)\left(\lambda_t^j-\lambda_t^k\right)}\right)dt \\ = & \ -4\sum_{i<j}\sum_{k\neq i,j}\frac{\lambda_t^k}{\left(\lambda_t^i-\lambda_t^k\right)\left(\lambda_t^j-\lambda_t^k\right)}dt+8\sum_{i<j}\sum_{k\neq i,j}\frac{\lambda_t^k}{\left(\lambda_t^i-\lambda_t^k\right)\left(\lambda_t^j-\lambda_t^k\right)}dt \\ = & \ 4\sum_{i<j}\sum_{k\neq i,j}\frac{\lambda_t^k}{\left(\lambda_t^i-\lambda_t^k\right)\left(\lambda_t^j-\lambda_t^k\right)}dt. \end{aligned}$$

Esto se puede factorizar de la siguiente forma

$$\begin{aligned} & 4\sum_{i<j<k}\left(\frac{\lambda_t^i}{\left(\lambda_t^k-\lambda_t^i\right)\left(\lambda_t^j-\lambda_t^i\right)}+\frac{\lambda_t^j}{\left(\lambda_t^i-\lambda_t^j\right)\left(\lambda_t^k-\lambda_t^j\right)}+\frac{\lambda_t^k}{\left(\lambda_t^i-\lambda_t^k\right)\left(\lambda_t^j-\lambda_t^k\right)}\right)dt \\ = & \ 4\sum_{i<j<k}\frac{\lambda_t^i\left(\lambda_t^k-\lambda_t^j\right)-\lambda_t^j\left(\lambda_t^k-\lambda_t^i\right)+\lambda_t^k\left(\lambda_t^j-\lambda_t^i\right)}{\left(\lambda_t^k-\lambda_t^i\right)\left(\lambda_t^j-\lambda_t^i\right)\left(\lambda_t^k-\lambda_t^j\right)}dt=0. \end{aligned}$$

Entonces, además de obtener el que U es una martingala local, claramente podemos usar el argumento de McKean (proposición 2.65) para $U=\log L^n$ y obtenemos que los valores propios jamás colisionan, es decir, $\tau=\infty$ c.s. Si x satisface $\operatorname{rnk}(x)\geq d-1$, entonces la proposición anterior muestra que inmediatamente después nos encontramos en el caso que ya probamos, por lo que la propiedad de Markov termina la prueba.

Para la primera afirmación, consideremos otra solución $\tilde{\Lambda}=\left(\tilde{\lambda}^1,\ldots,\tilde{\lambda}^n\right)$ al sistema de EDE con el mismo valor inicial y los mismos movimientos brownianos y evidentemente esta solución tampoco tiene colisiones.

Ahora el lema 3.3 de [28], p. 389 (o bien, lema 23.4 de [10], p.454), implica que el tiempo local de $z^i=\lambda^1-\tilde{\lambda}^i$ en 0 es nulo. Notemos que podemos usar la fórmula de Tanaka con z^i porque incluso si la integral de la deriva de z^i es impropia en el tiempo 0, podemos escribir la fórmula en $[s,t]$ y tomar $s\to 0$ ya que el tiempo local del proceso $\{z_u^i\}_{s\leq u\leq t}$ converge c.s. al de $\{z_u^i\}_{0\leq u\leq t}$.

En consecuencia, si definimos a los procesos F^{ij} dados por

$$F^{ij}=\frac{\lambda^i+\lambda^j}{\lambda^i-\lambda^j}-\frac{\tilde{\lambda}^i+\tilde{\lambda}^j}{\tilde{\lambda}^i-\tilde{\lambda}^j},$$

entonces

$$\begin{aligned}\sum_{i=1}^{d} E\left|\lambda_t^i-\tilde{\lambda}_t^i\right| &= E\int_0^t \sum_{i=1}^{d} \operatorname{sgn}\left(\lambda_s^i-\tilde{\lambda}_s^i\right)\sum_{j\neq i}\left(\frac{\lambda_s^i+\lambda_s^j}{\lambda_s^i-\lambda_s^j}-\frac{\tilde{\lambda}_s^i+\tilde{\lambda}_s^j}{\tilde{\lambda}_s^i-\tilde{\lambda}_s^j}\right)ds\\ &= E\int_0^t \sum_{i<j}\left(\operatorname{sgn}\left(\lambda_s^i-\tilde{\lambda}_s^i\right)F_s^{ij}+\operatorname{sgn}\left(\lambda_s^j-\tilde{\lambda}_s^j\right)F_s^{ji}\right)ds.\end{aligned}$$

Como $F^{ij}+F^{ji}=0$, si $\operatorname{sgn}\left(\lambda^i-\tilde{\lambda}^i\right)=\operatorname{sgn}\left(\lambda^j-\tilde{\lambda}^j\right)$, entonces el sumando se anula, de lo contrario, el sumando correspondiente es igual a

$$\begin{aligned}2\operatorname{sgn}\left(\lambda^i-\tilde{\lambda}^i\right)F^{ij} &= 2\operatorname{sgn}\left(\lambda^i-\tilde{\lambda}^i\right)\left(\frac{\left(\lambda^i+\lambda^j\right)\left(\tilde{\lambda}^i-\tilde{\lambda}^j\right)-\left(\tilde{\lambda}^i+\tilde{\lambda}^j\right)\left(\lambda^i-\lambda^j\right)}{\left(\lambda^i-\lambda^j\right)\left(\tilde{\lambda}^i-\tilde{\lambda}^j\right)}\right)\\ &= 4\operatorname{sgn}\left(\lambda^i-\tilde{\lambda}^i\right)\left(\frac{\lambda^j\tilde{\lambda}^i-\lambda^i\tilde{\lambda}^j}{\left(\lambda^i-\lambda^j\right)\left(\tilde{\lambda}^i-\tilde{\lambda}^j\right)}\right)\leq 0,\end{aligned}$$

porque $i<j$ implica que $\lambda^i-\lambda^j$ y $\tilde{\lambda}^i-\tilde{\lambda}^j$ son ambos negativos y tenemos

$$\begin{aligned}\lambda^i-\tilde{\lambda}^i\leq 0<\lambda^j-\tilde{\lambda}^j &\implies \lambda^i\leq\tilde{\lambda}^i<\tilde{\lambda}^j\leq\lambda^j \implies \lambda^j\tilde{\lambda}^i\geq\lambda^i\tilde{\lambda}^j\\ \lambda^i-\tilde{\lambda}^i> 0\geq\lambda^j-\tilde{\lambda}^j &\implies \tilde{\lambda}^j\geq\lambda^j>\lambda^i>\tilde{\lambda}^i \implies \lambda^j\tilde{\lambda}^i\leq\lambda^i\tilde{\lambda}^j.\end{aligned}$$

Consecuentemente, $\sum_{i=1}^{d} E\left|\lambda_t^i-\tilde{\lambda}_t^i\right|=0$ y tenemos que un proceso es modificación del otro. Finalmente, la continuidad de las trayectorias nos da la unicidad deseada. □

El siguiente paso natural es preguntarse cuáles de las propiedades anteriores se preservan en casos más generales. De acuerdo con la proposición 2.70, lo mejor a lo que podemos asprirar sin tener que involucrar a los procesos de vectores porpios, es al caso en que $\alpha=aI_d$ para $a>0$ y $\beta=bI_d$ para $b\in\mathbb{R}$. En este caso podemos dar ligeras extensiones a los resultados anteriores, en particular, al resultado de no colisión de los valores propios. Para esto, primero analicemos el subconjunto de $[0,\infty)^2$ donde se tiene

$$2a\left(x+y\right)+b\left(x-y\right)^2>0.$$

Si $b\geq 0$, entonces evidentemente se trata del conjunto semiabierto $[0,\infty)\times(0,\infty)\cup(0,\infty)\times[0,\infty)$, de otra forma, se trata de un conjunto no acotado determinado por ser la componente conexa de $[0,\infty)^2$ que contiene a la recta identidad (salvo el 0) definida por la ecuación cuadrática $2a\left(x+y\right)+b\left(x-y\right)^2=0$, que corresponde a la gráfica de una parábola de la forma $z=-\frac{b}{2a}w^2$ donde (z,w) es una rotación (cambio de coordenadas) de los ejes (x,y) en un ángulo de $\frac{\pi}{4}$. Este conjunto es semiabierto puesto que el conjunto

$$\left\{(x,y)\in\mathbb{R}^2 : 4a\left(x+y\right)+2b\left(x-y\right)^2>0\right\}$$

es abierto, y al intersectarse con $[0,\infty)^2$, esta propiedad se pierde porque el primer conjunto contiene una parte no vacía de los ejes, a saber, contine a $\{0\}\times\left(0,\frac{-2a}{b}\right)\cup\left(0,\frac{-2a}{b}\right)\times\{0\}$.

De esta forma, podemos definir a $C^{a,b}$ como el subconjunto semiabierto de $[0,\infty)^d$ tal que para todo $(\lambda_1,\ldots,\lambda_d)\in C^{a,b}$ se tenga que $2a\left(\lambda_i+\lambda_j\right)+b\left(\lambda_i-\lambda_j\right)^2>0$. En particular, si $b\geq 0$, recuperamos el conjunto descrito en el caso $(a,b)=(1,0)$.

Proposición 2.74. *Supongamos que $X\sim WP_d\left(p,aI_d,bI_d,x\right)$ donde $x\in\bar{S}_d^+$ satisface $rnk(x)\geq d-1$ y es tal que $\left(\lambda_1(x),\ldots,\lambda_d(x)\right)\in C^{a,b}$, entonces tenemos $\tau_n=0$ c.s. para cada n. Esto es, tenemos una repulsión inmediata de los valores propios. Más aún, para $p\in\Delta_d$ y x como antes o con valores propios distintos, el proceso de valores propios $\left(\lambda^1,\ldots,\lambda^d\right)$ de X es la trayectorialmente única solución fuerte del sistema de EDE*

$$d\lambda_t^i=2\sqrt{a\lambda_t^i}dW_t^i+\left(2ap+b\lambda_t^i+a\sum_{k\neq i}\frac{\lambda_t^i+\lambda_t^k}{\lambda_t^i-\lambda_t^k}\right)dt,\quad i=1,\ldots,d,$$

donde $\left(W^i\right)$ son movimientos brownianos independientes. Además, los valores propios nunca colisionan en $t>0$.

Demostración. Repasando sobre las pruebas de los resultados análogos para $(a,b)=(1,0)$, observamos que el argumento recaía en que si $\tau_n>0$ con probabilidad positiva, entonces podíamos decir lo mismo de τ_{n-1} y esto era consecuencia de que la deriva de L^n se anulaba y los coeficientes de los términos $E_{n-1}^{ij}(A)$ eran estrictamente positivos y por lo tanto estos términos debían ser nulos. La contradicción, a su vez, se seguía de que N_t^1, la deriva de L^1, era estrictamente positiva para $t>0$.

Todo esto era cierto justo por la construcción del conjunto C y nuevamente es cierto en nuestro contexto debido a que

$$2a\left(\lambda_t^i+\lambda_t^j\right)+b\left(\lambda_t^i-\lambda_t^i\right)^2>0,$$

para todo t en una vecindad no vacía (y posiblemente aleatoria, en la que los valores propios no salen de $C^{a,b}$) del 0 en $[0,\infty)$, con lo que concluye la prueba de la difracción instantánea de los valores propios.

El sistema de EDEs ya fue probado, así que sólo resta probar la unicidad fuerte y la ausencia de colisiones. Para probar la unicidad, procedemos como en la prueba de su proposición respectiva, donde ahora vemos que los argumentos mencionados ahí nos ayudan a concluir junto con el teorema de Tonelli que

$$\begin{aligned}\sum_{i=1}^{d} E\left|\lambda_t^i-\tilde{\lambda}_t^i\right| &\leq bE\int_0^t\sum_{i=1}^{d}\operatorname{sgn}\left(\lambda_s^i-\tilde{\lambda}_s^i\right)\left(\lambda_s^i-\tilde{\lambda}_s^i\right)ds\\ &= b\int_0^t\sum_{i=1}^{d}E\left|\lambda_s^i-\tilde{\lambda}_s^i\right|ds.\end{aligned}$$

Luego, la desigualdad de Grönwall implica que $\sum_{i=1}^{d} E\left|\lambda_t^i-\tilde{\lambda}_t^i\right|=0$, por lo que un proceso es modificación del otro y la continuidad implica que en realidad son indistinguibles. Finalmente, para probar la no colisión, observemos que la mayoría de las fórmulas para L^n se mantienen bajo ligeras modificaciones con respecto al escalar $\sqrt{a}$ o con un término extra a la deriva. Específicamente tenemos,

$$\begin{aligned}dA_t^{ij} &= 4\left(\lambda_t^i-\lambda_t^j\right)\sqrt{a\lambda_t^i}dW_t^i+4\left(\lambda_t^j-\lambda_t^i\right)\sqrt{a\lambda_t^j}dW_t^j\\ &+ \left(8a\left(\lambda_t^i+\lambda_t^j\right)+2b\left(\lambda_t^i-\lambda_t^j\right)^2+2a\left(\lambda_t^i-\lambda_t^j\right)\sum_{k\neq i,j}\frac{\lambda_t^i+\lambda_t^k}{\lambda_t^i-\lambda_t^k}-\frac{\lambda_t^j+\lambda_t^k}{\lambda_t^j-\lambda_t^k}\right)dt,\end{aligned}$$

por lo que, aunque $U=\log L^n$ ya no es una martignala local, su parte de variación finita en compactos está dada por

$$2b\sum_{i<j}Id=bd\left(d-1\right)Id,$$

la cuál está acotada en compactos y por lo tanto aún podemos usar el argumento de McKean (proposición 2.65) para concluir la prueba. □

Observación 2.75. Aunque a primera vista pueda parecer que la no colisión se puede deducir de la misma forma incluso en el caso general de (α,β) o al menos en (aI_d,β), esto no es del todo cierto ya que, a pesar de que se pueda llegar a algunas simplificaciones, en la parte de variación finita en compactos de U aparece el término

$$\sum_{i<j}\frac{2}{\lambda^i-\lambda^j}\left(p\left(\left(P'\alpha P\right)^{ii}-\left(P'\alpha P\right)^{jj}\right)+\lambda^i\left(P'\beta'P\right)^{ii}-\lambda^j\left(P'\beta'P\right)^{jj}\right)\cdot Id,$$

el cuál puede, en general, ir a $-\infty$ si tenemos una colisión entre dos valores propios.

2.6. Existencia de distribuciones y procesos de Wishart

Tal y como hicimos en la sección anterior, podemos comprobar que la evolución de $e_d\left(X\right)$ es la que controla la llegada y salida de X a ∂S_d^+. Similarmente, vamos a utilizar la información de $e_n\left(X\right)$ para inferir la existencia de distribuciones Wishart. Si conseguimos probar que $X \sim WP_d(p,\alpha,0,x)$ existe, entonces

$$X_t \sim W_d(p,x,t\alpha),$$

existirá también para cada $t>0$. Similarmente, si existe $W_d(p,x,\alpha)$ y $\alpha \in S_d^+$, entonces la proposición 2.9 muestra que existen todas las distribuciones $W_d(p,x,t\alpha)$ y por lo tanto podemos crear un proceso $WP_d(p,\alpha,0,x)$ a través de los kernels de transición de Wishart. Esta dualidad de existencia entre los procesos y las distribuciones de Wishart es lo que vamos a explotar para poder dar una caracterización de la existencia de ambos. Vamos a asumir de momento que $\alpha \in S_d^+$, y de hecho, gracias a la misma proposición 2.9, podemos reducir nuestro problema al caso $\alpha = I_d$. Posteriormente abordaremos el problema de $\alpha \in \partial S_d^+$, utilizando los recursos obtenidos del caso $\alpha \in S_d^+$.

Proposición 2.76. *Supongamos que $p \geq 0$ y $x \in \bar{S}_d^+$.*

(I) Si $0 < 2p < d-1$ y $p \notin \Delta_d$. Entonces $W_d(p,x,I_d)$ no existe.

(II) Si $p \in \Delta_d$, $rank(x) < d$ y además $W_d(p,x,I_d)$ existe, entonces $rank(x) \leq 2p$ y cualquier proceso $X \sim WP_d(p,I_d,0,x)$ satisface c.s. que para todo $t \geq 0$, $rank(X_t) \leq 2p$.

En particular, en ambos casos $[0,T_x)$ no contiene ningún intervalo determinista para $x \in S_d^+$, esto es, T_x puede ser arbitrariamente pequeño con probabilidad positiva.

Demostración. Supongamos que la primera afirmación es falsa y por lo tanto existen $W_d(p,x,I_d)$ y $X \sim WP_d(p,I_d,0,x)$. Aquí, la condición $\operatorname{rnk}(x) < d$ es equivalente a tener $e_d(x) = 0$. Recordando que

$$e_n(X) = \sqrt{\sum_i \lambda_i(X)\, e_{n-1}^i(X)^2} dW_t^n + (2p-n+1)(d-n+1)\, e_{n-1}(X)\cdot Id,$$

$$[e_n(X), e_m(X)] = \sum_{i=1}^d \lambda_i(X)\, e_{n-1}^i(X)\, e_{m-1}^i(X)\cdot Id,$$

en la EDE de $e_d(X)$ observamos que

$$\sqrt{\sum_i \lambda_i(X)\, e_{d-1}^i(X)^2} = \sqrt{e_d(X)\, e_{d-1}(X)},$$

lo que muestra que este es un múltiplo del proceso $BESQ(2p-d+1,0)$ con el cambio de tiempo $A = e_{d-1}(X)\cdot Id$. Luego, como la dimensión del proceso cuadrado de Bessel es negativo, esto equivale a tener un $-BESQ(|2p-d+1|,0)$. Es decir que si $A_t > 0$ con probabilidad positiva, entonces $e_d(X)$ se vuelve negativo inmediatamente (esto requiere que agreguemos $|\cdot|$ a todos los argumentos dentro de las raíces cuadradas en las EDEs encontradas anteriormente) y permanece negativo, es decir, $e_d(X_t) < 0$ c.s. en el evento $\{A_t > 0\}$. Como $e_n(X) \geq 0$ c.s., esto es imposible, por lo que debemos tener $A \equiv 0$ c.s. y en consecuencia $e_{d-1}(X) \equiv 0$ c.s. y lo mismo se puede decir de $e_d(X)$.

Como $2p-n+1$ jamás se anula, dado que $p \notin \Delta_d$, podemos proceder inductivamente sobre n descendiendo desde $n = d-1$, observando que como $e_n(X) \equiv 0$ c.s. entonces su deriva debe ser nula y por lo tanto $e_{n-1}(X) \equiv 0$. Esto prueba que $e_n(X) \equiv 0$ para todo $n \geq 0$, lo cuál es imposible porque $e_0(X) \equiv 1$ por definición y por lo tanto la primera afirmación es cierta.

La prueba de la segunda afirmación es idéntica, con la diferencia de que la última vez que el argumento funciona es cuando $n = 2p+2$, y por lo tanto sólo podemos concluir que $e_{2p+1}(X) \equiv 0$, lo que implica que $\operatorname{rank}(x) \leq 2p$ y además muestra que si $X \sim WP_d(p, I_d, 0, x)$, entonces $\operatorname{rank}(X_t) \leq 2p$ para todo $t \geq 0$.

La razón por la que $[0, T_x)$ no contiene ningún intervalo determinista es porque μ_t no existe y por lo tanto hay probabilidad positiva de que la solución fuerte a la EDE de Wishart hasta el tiempo $T_x > 0$ deje de existir antes del tiempo t para todo $t > 0$. $\square$

Observación 2.77. Si ahora contrastamos la segunda afirmación con la construcción de procesos con estado inicial de rango a lo más $2p \in \mathbb{N}$ con $2p < d-1$, podemos ver que en efecto existen procesos de Markov, entendido como que existen los kernels de transición para todos los valores iniciales del proceso, si nos restringimos a

$$\mathcal{D}_p = \left\{ x \in \bar{S}_d^+ : \operatorname{rank}(x) \leq 2p \right\} \subset \bar{S}_d^+.$$

Para concluir con la caracterización de distribuciones Wishart, requeriremos de unos resultados preliminares debido a que esta técnica no puede ser usada cuando x es de rango completo. Para ello, tendremos que enfocarnos un poco más en el análisis de sus valores propios.

Lema 2.78. *Supongamos que $2p < d$ y supongamos que $x^1 \leq \ldots \leq x^d$ es una solución sin colisiones del sistema de EDEs*

$$dx_t^i = 2\sqrt{|x_t^i|}dW_t^i + \left(2p + \sum_{k \neq i} \frac{|x_t^i| + \left|x_t^k\right|}{x_t^i - x_t^k} \right) dt, \quad i = 1, \ldots, d, \tag{2.23}$$

donde $x_0^1 \geq 0$ y (W^i) son movimientos brownianos independientes. Entonces $\mathbb{P}(x_t^1 < 0) > 0$

para toda $t > 0$.

Demostración. Sea y^1 una solución a

$$dy_t^1 = 2\sqrt{|y_t^1|}dW_t^1 + (2p - d + 1)\,dt, \quad y_0^1 = x_0^1.$$

Usando las técnicas de tiempos locales de Le Gall (ver [29]); más específicamente, usando los lemas 3.3 y 3.4 de [28], p. 389-390, vemos que el tiempo local $L^0\left(x^1 - y^1\right)$ se anula idénticamente. Luego, la fórmula de Tanaka (tal y como se hace en el teorema 3.7 de [28], p. 394) nos da

$$E\left(x_t^1 - y_t^1\right)^+ = E\int_0^t 1\left\{x_t^1 > y_t^1\right\}\left(d - 1 + \sum_{k=2}^{d}\frac{|x_t^1| + \left|x_t^k\right|}{x_t^1 - x_t^k}\right)ds \leq 0,$$

ya que $|x_t^1| + \left|x_t^k\right| \geq x_t^k - x_t^1$ para cada $d \geq k \geq 2$. Esto y la continuidad trayectorial nos da que $x^1 \leq y^1$ c.s. Ahora, como y^1 es un proceso de Bessel de dimensión negativa, este cruza el 0 y permanece no positivo, concluyendo la prueba. $\square$

Proposición 2.79. *Supongamos que $2p < d-1$ y que $x \in S_d^+$, entonces no existe $W_d\left(p, x, I_d\right)$.*

Demostración. Supongamos que sí existe y por lo tanto podemos considerar un proceso $X \sim WP_d\left(p, I_d, 0, x\right)$. Luego, sus valores propios satisfacen (2.23) y por lo tanto λ^1 tiene probabilidad positiva de volverse negativo, lo cuál contradice el que $X \in \bar{S}_d^+$, lo que termina la prueba. $\square$

Dada una matriz $A \in S_k$, denotamos por $\det_n(A) = \det\left(\left(A_{ij}\right)_{1\leq i,j\leq n}\right)$ y para un vector de entradas enteras $\kappa = (\kappa_1, \ldots, \kappa_k)$, también llamado multi-índice, con $\kappa_1 \geq \cdots \geq \kappa_k \geq 0$, denotamos por $|\kappa|$ a $\sum_i \kappa_i$ y denotamos por E_n al conjunto de tales vectores κ que satisfacen $|\kappa| = n$. Definamos también

$$\det_\kappa(A) = \prod_{i=1}^{k}\det_i(A)^{\kappa_i - \kappa_{i+1}},$$

donde $\kappa_{k+1} = 0$; en particular tenemos $\det_\kappa(A) > 0$ si $\left(A_{ij}\right)_{1\leq i,j\leq k} \in S_k^+$. Además, para $z = (z_1, \ldots, z_k) \in \mathbb{R}^k$ tal que $z_i > \frac{i-1}{2}$, $i = 1, \ldots, k$, definimos

$$\Gamma_k(z) = \pi^{\frac{k(k-1)}{4}}\prod_{i=1}^{k}\Gamma\left(z_i - \frac{i-1}{2}\right),$$

en particular tenemos para $r > \frac{k-1}{2}$ (ver [11], teorema 2.1.12),

$$\Gamma_k (r\mathbf{1}) = \int_{S_k^+} e^{-\mathrm{tr}(x)} \det (x)^{r-\frac{k+1}{2}} \, dx.$$

Esto da pie a la definición de

$$(r)_\kappa = \frac{\Gamma_k (\kappa + r\mathbf{1})}{\Gamma_k (r\mathbf{1})} = \prod_{i=1}^{k} \left(r - \frac{i-1}{2} \right)_{\kappa_i},$$

donde $(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}$ es el símbolo de Pochhammer.

El polinomio de zonal $C_\kappa (A)$, definido como la siguiente integral sobre el grupo $\mathbb{O}(k)$ de matrices ortogonales de orden k respecto de la medida de Haar du (normalizada para que tenga masa total 1):

$$C_\kappa (x) = C_\kappa^0 \int_{\mathbb{O}(k)} \det_\kappa \left(u^{-1} x u \right) du,$$

donde la constante de normalización C_κ^0 (cuya forma no es relevante para nosotros) es bastante complicada y se puede encontrar en [12], p. 234, en la última línea o en la fórmula (18) de [11], p. 237.

Un análisis profundo de estos polinomios homogéneos con respecto a las entradas de la matriz simétrica A se puede encontrar en el capítulo 7 de [11] o en [30]. Lo más importante al respecto es que $C_\kappa (A)$ es que toma valores positivos en S_d^+ y que $C_\kappa (A) = C_\kappa (u' A u)$ para $u \in \mathbb{O}(k)$, lo que implica que el polinomio sólo depende de los valores propios de su argumento. Estos polinomios satisfacen varias propiedades que nos serán útiles, algunas de ellas son:

$$e^{\mathrm{tr}(x)} = \sum_{n=0}^{\infty} \sum_{\kappa \in E_n} \frac{1}{n!} C_\kappa (x), \tag{2.24}$$

$$\det (I_k - x)^{-p} = \sum_{n=0}^{\infty} \sum_{\kappa \in E_n} \frac{(p)_\kappa}{n!} C_\kappa (x), \tag{2.25}$$

$$(p)_\kappa \det (u)^{-p} C_\kappa \left(u^{-1} \right) = \int_{S_d^+} e^{-\mathrm{tr}(ux)} C_\kappa (x) \frac{\det (x)^{p-\frac{k+1}{2}}}{\Gamma_k (p\mathbf{1})} dx, \tag{2.26}$$

$$\frac{C_\kappa (x) \, C_\kappa (y)}{C_\kappa (I_k)} = \int_{\mathbb{O}(u)} C_\kappa \left(\sqrt{y} u' x u \sqrt{y} \right) du. \tag{2.27}$$

La constante C_κ^0 juega un papel importante en el que estas identidades se puedan dar. De hecho, a (2.24) lo podemos reemplazar por la identidad equivalente

$$\operatorname{tr}(x)^n = \sum_{\kappa\in E_n} \frac{1}{n!} C_\kappa(x),$$

de la cuál podemos deducir que $0 \le C_\kappa(x) \le \operatorname{tr}(x)^n$ para $x \in \bar{S}_d^+$. Finalmente estamos listos para dar uno de los principales teoremas de esta sección.

Teorema 2.80. *$W_d(p, x, I_d)$ existe, y por lo tanto $W_d(p, x, \sigma)$ y $WP_d(p, \alpha, \beta, x)$ también para $\sigma, \alpha \in S_d^+$, si y sólo si $p \in \Delta_d$ y se da alguno de los siguientes casos:*

(I) $2p < d - 1$ y $rank(x) \le 2p$, o

(II) $2p \ge d - 1$,

Si además $2p > d - 1$, entonces $W_d(p, x, I_d)$ tiene densidad. Si σ o α no son invertibles, pero se dan los casos anteriores, entonces también existen $W_d(p, x, \sigma)$ y $WP_d(p, \alpha, \beta, x)$. Más aún, si existen matrices $\rho \in M_{d,d_0}$ y $\sigma_0, x_0 \in \bar{S}_{d_0}^+$ donde $d_0 < d$ tales que $\sigma = \rho\sigma_0\rho'$ y $x = \rho x_0 \rho$ y donde $W_{d_0}(p, x_0, \sigma_0)$ existe, entonces $W_d(p, x, \sigma)$ también existe.

Demostración. La última afirmación se sigue directamente de la proposición 2.9. Una vez que probemos la existencia de $W_d(p, x, I_d)$, la de $W_d(p, x, \sigma)$ se seguirá de la proposición 2.9 y la de $WP_d(p, \alpha, \beta, x)$ se seguirá de las construcciones de soluciones dadas en las secciones anteriores para $2p < d-1$ y $2p \ge d+1$, mientras que para $2p \in [d-1, d+1)$ lo justificamos usando el que todas las distribuciones $W_d(p, x, \sigma)$ existen para cualesquiera $x, \sigma \in \bar{S}_d^+$ y por lo tanto todos los kernels de transición de Wishart correspondientes existen.

La necesidad de las distintas condiciones es consecuencia de todos los resultados anteriores en la sección presente. Para la suficiencia del primer caso, basta seguir la construcción dada en el lema 2.17 y fijarnos en el tiempo 1 para probar el caso general de $\sigma, \alpha \in \bar{S}_d^+$ directamente. Si $2p > d-1$, entonces podemos definir a la medida $\gamma(p, x, \cdot)$ sobre S_d^+dada por la derivada de Radon-Nikodym

$$\frac{\gamma(p, x, dy)}{dy} = \frac{e^{-\operatorname{tr}(y+x)}}{\Gamma_d(p)} \det(y)^{p-\frac{d+1}{2}} \left(\sum_{n=0}^{\infty} \sum_{\kappa\in E_n} \frac{C_\kappa(\sqrt{x}y\sqrt{x})}{n!\,(p)_\kappa} \right).$$

Es fácil verificar que esta derivada de Radon-Nikodym es no negativa porque $x+y, \sqrt{x}y\sqrt{x} \in \bar{S}_d^+$. Supongamos de momento que $x \in S_d^+$, y notemos que el cambio de variable $y \mapsto z =$

$\sqrt{x}y\sqrt{x}$ tiene por Jacobiano $dz = \det(x)^{\frac{d+1}{2}}\, dy$, por lo que tenemos para cada $u \in \bar{S}_d^+$,

$$\begin{aligned}
I_\kappa(x) &:= e^{-\mathrm{tr}(x)} \int_{S_d^+} \frac{e^{-\mathrm{tr}((I_d+u)y)}}{\Gamma_d(p)} \det(y)^{p-\frac{d+1}{2}} \frac{C_\kappa(\sqrt{x}y\sqrt{x})}{n!\,(p)_\kappa} dy \\
&= \frac{e^{-\mathrm{tr}(x)}}{\det(x)^p\, n!\,(p)_\kappa} \int_{S_d^+} \exp\left\{-\mathrm{tr}\left(x^{-\frac{1}{2}}(I_d+u)\,x^{-\frac{1}{2}}z\right)\right\} \frac{C_\kappa(z)\det(z)^{p-\frac{d+1}{2}}}{\Gamma_d(p)} dz \\
&= \frac{e^{-\mathrm{tr}(x)}}{\det(x)^p\, n!} \det\left(x^{-\frac{1}{2}}(I_d+u)\,x^{-\frac{1}{2}}\right)^{-p} C_\kappa\left(\sqrt{x}\,(I_d+u)^{-1}\sqrt{x}\right) \\
&= \frac{e^{-\mathrm{tr}(x)}}{n!} \det(I_d+u)^{-p}\, C_\kappa\left(\sqrt{x}\,(I_d+u)^{-1}\sqrt{x}\right),
\end{aligned}$$

donde en la tercera línea usamos (2.26). Si $x \in \partial S_d^+$, entonces $x_m := x + \frac{1}{m}I_d \in S_d^+$ y tenemos

$$0 \le C_\kappa(\sqrt{x_m}y\sqrt{x_m}) \le \mathrm{tr}(\sqrt{x_m}y\sqrt{x_m})^n \le \mathrm{tr}(\sqrt{x_1}y\sqrt{x_1})^n = \mathrm{tr}(yx_1)^n,$$

así que el teorema de convergencia dominada implica que $I_\kappa(x_m) \to I_\kappa(x)$, por lo que también admite la representación que dimos. Si ahora usamos (2.24), conseguimos que

$$\begin{aligned}
\int_{S_d^+} e^{-\mathrm{tr}(uy)}\gamma(p,x,dy) &= \sum_{n=0}^{\infty}\sum_{\kappa\in E_n} I_\kappa(x) = \frac{e^{-\mathrm{tr}(x)}e^{\mathrm{tr}\left(\sqrt{x}(u+I_d)^{-1}\sqrt{x}\right)}}{\det(I_d+u)^p} \\
&= \frac{e^{-\mathrm{tr}\left(\left(I_d-(I_d+u)^{-1}\right)x\right)}}{\det(I_d+u)^p} = \frac{e^{-\mathrm{tr}\left(((I_d+u)-I_d)(I_d+u)^{-1}x\right)}}{\det(I_d+u)^p} \\
&= \frac{e^{-\mathrm{tr}\left(u(I_d+u)^{-1}x\right)}}{\det(I_d+u)^p},
\end{aligned}$$

lo que muestra que su transformada de Laplace es la correspondiente a la de una distribución $W_d(p, I_d, x)$. Si sustituimos $u = 0$ verificamos que $\gamma(p,x,\cdot)$ es medida de probabilidad y por lo tanto $\gamma(p,x,\cdot) = W_d(p,I_d,x)$. Esto muestra la existencia de la densidad cuando $2p > d-1$. Por otro lado, observando la transformada de Laplace, podemos ver que si tomamos $\frac{d-1}{2} < p_n \downarrow p = \frac{d-1}{2}$, entonces

$$\int_{S_d^+} e^{-\mathrm{tr}(uy)}\gamma(p_n,x,dy) \to \frac{e^{-\mathrm{tr}\left(u(I_d+u)^{-1}x\right)}}{\det(I_d+u)^p}, \quad x \in \bar{S}_d^+,$$

donde el término de la derecha es continuo y vale 1 en $u = 0$. Esto nos permite usar el teorema de continuidad de Lévy para transformadas de Laplace y conseguir la existencia para $2p = d-1$. Si ahora suponemos que σ no es invertible, en cualquiera de los casos, basta ver que

$$\frac{e^{-\mathrm{tr}\left(u(I_d+\sigma_n u)^{-1}x\right)}}{\det(I_d+\sigma_n u)^p} \to \frac{e^{-\mathrm{tr}\left(u(I_d+\sigma u)^{-1}x\right)}}{\det(I_d+\sigma u)^p},$$

donde $\sigma_n = \sigma + \frac{1}{n} I_d \in S_d^+$, por lo que el teorema de continuidad de Lévy una vez más nos da el resultado faltante. $\square$

Observación 2.81. La única importancia de la última afirmación yace en el hecho de que nos permite más elecciones de parámetros de p si σ no es de rango completo y mientras σ y x sean debidamente compatibles.

Resumiendo los resultados de algunos de estos teoremas, proposiciones, lemas y hechos como el que $A \mapsto e^{\beta \cdot Id} A e^{\beta' \cdot Id}$ es biyectivo (y por lo tanto no altera el rango). En especial, usaremos el lema 2.28, los teoremas 2.66, 2.80 y las proposiciones2.48 y 2.76 para sintetizar el siguiente resultado directo:

Corolario 2.82. *Consideremos $x \in \bar{S}_d^+$ y $p \in \Delta_d$ tales que rank$(x) \leq 2p$ y supongamos que $X \sim WP_d(p, \alpha, \beta, x)$ con $\alpha \in S_d^+$, entonces X se mantiene c.s. en S_d^+ si $2p \geq d+1$ o en ∂S_d^+ si $2p < d-1$ (de hecho, rank$(X) \leq 2p$ c.s.). Además, si $\Xi \sim W_d(p, \omega, \sigma)$ entonces* $\ker(\sigma) \subset \ker(\Xi - \omega)$ *c.s., por lo que rank$(\Xi - \omega) \leq$ rank(σ) c.s.*

Si $2p = d-1$, entonces en cualquier conjunto con interior no vacío de $\mathbb{R}_+$, el proceso pasa tiempo positivo (respecto de la medida de Lebesgue) en ∂S_d^+ con probabilidad positiva. Para $2p \in (d-1, d+1)$ las transiciones de X tienen densidad y por lo tanto c.s. está en S_d^+ para todo $t \in \mathbb{Q}_+$, pero puede ser posible que $\tau_0 < \infty$ con probabilidad positiva. No obstante, el tiempo que pasa en ∂S_d^+ tiene medida de Lebesgue 0 c.s. si $2p > d-1$. Finalmente, la solución a una EDE de Wishart tiene unicidad trayectorial si $2p \geq d+1$ y $x \in S_d^+$, mientras que tenemos unicidad en ley si $2p \geq d-1$.

Demostración. Tomemos cualquier $v \in \ker(\sigma)$ como vector columna, si $r > 0$, entonces $v'\sigma = 0$, obteniendo

$$Ee^{-\left\langle r^2 vv', \Xi - \omega \right\rangle} \equiv e^{-\left\langle r^2 vv', \omega - \omega \right\rangle} = 1,$$

de forma que conforme $r \to \infty$ obtenemos que $v \in \ker(\Xi - \omega)$ c.s., probando la afirmación correspondiente. Lo único que falta probar es la unicidad de las soluciones y que si $2p > d-1$, entonces el tiempo que pasa en ∂S_d^+ tiene medida de Lebesgue 0 y que esta medida es positiva con probabilidad positiva cuando $2p = d-1$. Esto se sigue directamente del teorema de Tonelli y de que $\mu_t\left(x, \partial S_d^+\right) = 0$ para todo $t > 0$ porque μ_t tiene densidad cuando $2p > d-1$: denotemos por m a la medida de Lebesgue y observemos que $\left\{t \geq 0 : X_t \in \partial S_d^+\right\}$ es cerrado (y por lo tanto medible) por continuidad de X, de tal forma que

$$\begin{aligned} E_x m\left(\left\{t \geq 0 : X_t \in \partial S_d^+\right\}\right) &= E_x \int_{\mathbb{R}_+} 1_{\left\{X_t \in \partial S_d^+\right\}} dt = \int_{\mathbb{R}_+} \mathbb{P}_x\left(X_t \in \partial S_d^+\right) dt \\ &= \int_{\mathbb{R}_+} \mu_t\left(x, \partial S_d^+\right) dt = 0. \end{aligned}$$

Como $m\left(\left\{t\geq 0:X_t\in\partial S_d^+\right\}\right)$ es una variable aleatoria no negativa, el que tenga esperanza nula implica que esta es 0 c.s. Si ahora tomamos $2p=d-1$, entonces la continuidad de Lévy antes usada muestra que para $t>0$ fijo, conforme $p\downarrow\frac{d-1}{2}$, las leyes de transición, que ahora denotaremos μ_t^p, convergen débilmente a $\mu_t^{\frac{d-1}{2}}$.

Luego, la degeneración de las densidades de μ_t^p en ∂S_d^+ conforme $p\downarrow\frac{d-1}{2}$ implica que $\mu_t^{\frac{d-1}{2}}$ asigna probabilidad positiva a ∂S_d^+. Notemos que $\mu_t^{\frac{d-1}{2}}$ es continuo en t porque X tiene trayectorias continuas y por lo tanto para cada conjunto medible $C\subset\mathbb{R}_+$ con $m(C)>0$ tenemos

$$\begin{aligned} E_x m\left(\left\{t\in C:X_t\in\partial S_d^+\right\}\right) &= E_x\int_C 1_{\left\{X_t\in\partial S_d^+\right\}}dt \\ &= \int_C \mu_t^{\frac{d-1}{2}}\left(x,\partial S_d^+\right)dt>0, \end{aligned}$$

de lo contrario tendríamos $\mu_t^{\frac{d-1}{2}}\left(x,\partial S_d^+\right)=0$ casi en todos lados con respecto a la medida de Lebesgue, una contradicción. Luego, $m\left(\left\{t\in C:X_t\in\partial S_d^+\right\}\right)$ es positivo con probabilidad positiva, como queríamos probar. Para la última afirmación, notemos que en la sección anterior probamos la polaridad de ∂S_d^+ y por lo tanto la existencia fuerte de soluciones si $x\in S_d^+$ y $2p\geq d+1$. Si ahora tomamos $2p\geq d-1$, la existencia de todos los kernels de transición implica directamente la existencia débil y unicidad en ley de las soluciones. □

Observación 2.83. El que $\ker(\sigma)\subset\ker(\Xi-\omega)$ muestra directamente que toda la aleatoriedad recae sobre el espacio de matrices positivas semidefinidas cuyo núcleo contiene al núcleo de σ.

Otra aplicación de estas semimartingalas escalares que fueron descritas anteriormente es la prueba de que las distribuciones Wishart presentan momentos exponenciales en la mayoría de los casos de interés.

Proposición 2.84. *(función generadora de momentos) Las distribuciones Wishart $\Xi\sim W_d(p,x,\sigma)$ tienen momentos exponenciales bajo las condiciones del teorema 2.80. De hecho, en tanto $u\sigma$ tenga todos sus valores propios estrictamente menores a 1 para alguna $u\in\bar{S}_d^+$, tenemos $Ee^{\langle u,X\rangle}<\infty$. Además, esta esperanza tiene la misma forma que en (2.1) con los signos de u cambiados.*

Demostración. Aunque esto es consecuencia del teorema 2.42, daremos una prueba alternativa. Inicialmente supongamos que $\sigma\in S_d^+$ y tomemos $u\in S_d^+$ como en el enunciado, veamos que $e^{\langle u,\Xi\rangle}=e^{\left\langle I_d,\sqrt{u}\Xi\sqrt{u}\right\rangle}$ y recordemos de la proposición 2.9 que

$$\sqrt{u}\Xi\sqrt{u}\sim W_d\left(p,\sqrt{u}\omega\sqrt{u},\sqrt{u}\sigma\sqrt{u}\right),$$

donde los valores propios de $\sqrt{u}\sigma\sqrt{u}$ son (los mismos que los de $u\sigma$) estrictamente menores a 1. Luego, podemos suponer sin pérdida de generalidad que $u = I_d$ y que σ tiene valores propios en $(0,1)$. Luego, bajo un cambio de coordenadas ortogonales (una rotación), podemos asumir que σ es diagonal. Finalmente, podemos notar que

$$e^{\langle I_d, \Xi\rangle} \leq e^{r\left\langle \sigma^{-1}, \Xi\right\rangle},$$

para cualquier $\lambda_d(\sigma) \leq r < 1$ ya que σ^{-1} es diagonal, con entradas (en la diagonal) mayores o iguales a $\lambda_1\left(\sigma^{-1}\right) = \lambda_d(\sigma)^{-1} > 1$. Ahora tomemos $X \sim WP_d(p, \sigma, 0, \omega)$ y observemos que la traza de $X\sigma^{-1}$ tiene a su parte de variación finita en compactos dada por $2pdId$ y con

$$\begin{aligned} d\left[\operatorname{tr}\left(X\sigma^{-1}\right)\right]_t &= \left(\sum_i dX_t^{ii}\sigma_{ii}^{-1}\right)^2 = \sum_i d\left[X^{ii}, X^{jj}\right]_t \sigma_{ii}^{-1}\sigma_{jj}^{-1} \\ &= 4\sum_{ij}\left(X^{ij}\sigma_{ij}\right)\sigma_{ii}^{-1}\sigma_{jj}^{-1}dt = 4\sum_{ii} X^{ii}\sigma_{ii}^{-1}dt \\ &= 4\operatorname{tr}\left(X_t\sigma^{-1}\right)dt. \end{aligned}$$

Se sigue que $\operatorname{tr}\left(X\sigma^{-1}\right)$ es un proceso cuadrado de Bessel y por lo tanto $\operatorname{tr}\left(X\sigma^{-1}\right)$ tiene distribuciones marginales χ^2 no centrales. Concluimos que, como $r < 1$ y $X_{\frac{1}{2}} \overset{d}{=} \Xi$, tenemos

$$\begin{aligned} Ee^{\langle I_d, \Xi\rangle} &\leq Ee^{r\left\langle \sigma^{-1}, \Xi\right\rangle} = E\exp\left\{r\operatorname{tr}\left(X_{\frac{1}{2}}\sigma^{-1}\right)\right\} \\ &= (1-r)^{-pd}\exp\left\{\frac{r\operatorname{tr}\left(\omega\sigma^{-1}\right)}{1-r}\right\} < \infty, \end{aligned}$$

como se quería probar. Además, la forma de esta función generadora de momentos debe ser la que afirmamos porque esta es la única extensión analítica. Finalmente, para el caso en que σ no es invertible, simplemente seguimos el mismo procedimiento que en la prueba del teorema 2.80, donde el teorema de continuidad de Lévy nos da el resultado deseado. □

El teorema de Girsanov es de gran utilidad para mostrar la existencia débil de ecuaciones diferenciales estocásticas semejantes pero con distintas derivas. En la situación vectorial, este teorema adquiere un carácter un tanto más complicado y por lo tanto, lo mismo es cierto para el caso matricial. De hecho, las complicaciones están relacionadas a comprobar que el proceso de derivadas de Radon-Nikodym efectivamente es una martingala, puesto que incluso el criterio de Novikov es difícil de comprobar. Para futuras referencias, enunciaremos el teorema de Girsanov para procesos matriciales y lo probaremos usando el conocido caso vectorial.

Teorema 2.85. *(cambio de deriva, Girsanov) Supongamos que ξ es un proceso continuo y adaptado con valores en $M_{d,d}$ y B un movimiento browniano en $M_{d,d}$ bajo $\mathbb{P}$ tales que*

$\int_0^T tr\left(\xi_t'\xi_t\right)dt < \infty$ c.s. Si además se tiene que $Z_t = \mathcal{E}\left(-\int_0^t tr\left(\xi_s dB_s\right)\right)$ es una martingala en $[0,T]$, donde $\mathcal{E}$ denota la exponencial estocástica de Dóleans-Dade, entonces podemos definir una medida $\tilde{\mathbb{P}}$ equivalente a $\mathbb{P}$ a través de $\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = Z_T$ de tal forma que $W = B + \xi \cdot Id$ es un $\tilde{\mathbb{P}}$-movimiento browniano.

Demostración. Basta ver que las condiciones anteriores equivalen a (en su versión vectorial) que c.s. tengamos

$$\infty > \int_0^T \sum_{ij} \xi_t^{ij}\xi_t^{ij} dt = \int_0^T \operatorname{vec}\left(\xi_t\right)' \operatorname{vec}\left(\xi_t\right) dt,$$

a que $Z_t = \mathcal{E}\left(-\int_0^t \operatorname{vec}\left(\xi_s\right)' \operatorname{vec}\left(dB_s\right)\right)$ sea martingala en $[0,T]$ y a que

$$\operatorname{vec}(W) = \operatorname{vec}(B) + \operatorname{vec}(\xi) \cdot Id,$$

sea un $\tilde{\mathbb{P}}$-movimiento browniano, que es justo lo que afirma el teorema clásico de Girsanov.

$\square$

Ahora estamos preparados para enunciar un fuerte resultado sobre el cambio de deriva de los procesos de Wishart bajo cambios de medidas adecuadas.

Teorema 2.86. *(cambio de deriva para procesos de Wishart) Supongamos que $X \sim WP_d\left(p, \alpha, \beta, x\right)$ es una solución débil a la EDE de Wishart para $B \sim BM_{d,d}$ donde $p \geq \frac{d-1}{2}$ y consideremos $q \geq \frac{d-1}{2}$ y $\gamma \in M_{d,d}$. Supongamos además que*

(I) *si $q \neq p$, entonces $p \wedge q \geq \frac{d+1}{2}$ y $x \in S_d^+$,*

(II) *si $\gamma \neq \beta$, entonces Q es invertible, o lo que es equivalente, α es invertible.*

Definamos también a los procesos

$$\begin{aligned} \xi &= \sqrt{X}\left(\beta - \gamma\right)' Q^{-1} 1_{\{\beta \neq \gamma\}} + (p - q)\sqrt{X}^{-1} Q' 1_{\{p \neq q\}}, \\ Z_t &= \mathcal{E}\left(-\int_0^t tr\left(\xi_s dB_s\right)\right), \end{aligned}$$

Entonces Z es una martingala no negativa en cualquier compacto $[0,T]$ para cualquier $T > 0$ fijo. Además, existe medida de probabilidad $\tilde{\mathbb{P}}$, equivalente a $\mathbb{P}$, cuyo proceso de derivadas de Radon-Nikodym es Z, bajo la cuál $W = B + \xi \cdot Id$ es un $\tilde{\mathbb{P}}$-movimiento browniano y $X \sim WP_d\left(q, \alpha, \gamma, x\right)$ respecto de $\tilde{\mathbb{P}}$ y W.

Demostración. Si $p = q$ y $\gamma = \beta$, entonces no hay nada que probar, así que supondremos que al menos uno es distinto. La necesidad de las condiciones es clara puesto que debemos garantizar la existencia de $\sqrt{X}^{-1}$ y de Q^{-1} respectivamente. Observemos que

$$\sqrt{X_t}\xi_t Q + Q'\xi_t'\sqrt{X_t} = 2\left(p-q\right)Q'Q + \left(\beta-\gamma\right)X_t + X_t\left(\beta-\gamma\right)',$$

lo que implica que

$$\begin{aligned} dX_t &= \sqrt{X_t}dB_tQ + Q'dB_t'\sqrt{X_t} + \left(2pQ'Q + \beta X_t + X_t\beta'\right)dt \\ &= \sqrt{X_t}\left(\xi_t dt + dB_t\right)Q + Q'\left(\xi_t' dt + dB_t'\right)\sqrt{X_t} \\ &+ \left(2pQ'Q + \beta X_t + X_t\beta' - \sqrt{X_t}\xi_t Q - Q'\xi_t'\sqrt{X_t}\right)dt \\ &= \sqrt{X_t}dW_tQ + Q'dW_t'\sqrt{X_t} + \left(2qQ'Q + \gamma X_t + X_t\gamma'\right)dt. \end{aligned}$$

Sólo resta probar que Z es una martingala, puesto que del teorema de Girsanov (ver [10]), se seguirán el resto de las conclusiones.

Observemos que Z es una martingala local no negativa por construcción y por lo tanto es una supermartingala. Luego, si mostramos que $E\left(Z_T\right) = E\left(Z_0\right) = 1$, habremos probado que Z es una martingala, de lo contrario, la esperanza habría decrecido.

Si $p = q$ y $\gamma \neq \beta$ o bien, $p \neq q$ y $\gamma = \beta$, entonces los resultados de esta sección y la anterior garantizan la existencia de $WP_d\left(q, \alpha, \gamma, x\right)$ y por lo tanto, de soluciones débiles a la EDE de Wishart correspondiente. Consideremos a $\tilde{X}$ como una solución a la EDE de Wishart con respecto a B y con parámetros (q, α, γ, x) (cuya existencia está garantizada posiblemente en una extensión del espacio de probabilidad, a través del teorema *transfer* en [10]) y definamos a

$$\tilde{\xi}_t = \sqrt{\tilde{X}_t}\left(\beta-\gamma\right)'Q^{-1}1_{\{\beta\neq\gamma\}} + \left(p-q\right)\sqrt{\tilde{X}_t}^{-1}Q'1_{\{p\neq q\}}.$$

Ahora introducimos a dos sucesiones no decrecientes de tiempos de paro:

$$\begin{aligned} \tau_n &= \text{ínf}\left\{t > 0 : \|\xi_t\| \geq n\right\} \wedge T \quad \text{y} \\ \tilde{\tau}_n &= \text{ínf}\left\{t > 0 : \left\|\tilde{\xi}_t\right\| \geq n\right\} \wedge T. \end{aligned}$$

La polaridad del 0 cuando $p \neq q$ y la continuidad de X y $\tilde{X}$ prueban, junto con la continuidad de la probabilidad y la no explosión de las soluciones, que

$$1 = \lim_{n\to\infty}\mathbb{P}\left(\tau_n = T\right) = \lim_{n\to\infty}\mathbb{P}\left(\tilde{\tau}_n = T\right).$$

Para cada $n \in \mathbb{N}$ definimos al proceso $\xi_t^n = \xi_t 1_{\{t\leq\tau_n\}}$ para $t \in [0, T]$, de tal forma que

$\int_0^t \|\xi_s^n\|^2 \, ds \leq n^2 t$ y por lo tanto se satisface la condición de Novikov:

$$Ee^{\frac{1}{2}\int_0^T \|\xi_t^n\|^2 dt} \leq e^{\frac{1}{2}n^2 T} < \infty.$$

Consecuentemente $Z^n = \mathcal{E}\left(-\int_0^{Id} \operatorname{tr}\left(\xi_t^n dB_t\right)\right)$ es una martingala que induce una medida $\tilde{\mathbb{P}}^n$ equivalente a $\mathbb{P}$ para la cuál $B^n = B + \xi^n \cdot Id$ es un movimiento browniano en $[0,T]$. Más aún, el proceso $\tilde{X}^n := \tilde{X}^{\tilde{\tau}_n}$ tiene la misma ley bajo $\mathbb{P}$ que el proceso $X^n := X^{\tau_n}$ bajo $\tilde{\mathbb{P}}^n$. Como $Z_T > 0$, tenemos que $\left\{Z_T^n 1_{\{\tau_n = T\}}\right\}_{n\geq 1}$ es una sucesión no decreciente (de hecho, constante por pedazos) lo que nos permite usar el teorema de convergencia monótona para calcular

$$\begin{aligned} E\left(Z_T\right) &= \lim_{n\to\infty} E\left(Z_T^n 1_{\{\tau_n = T\}}\right) = \lim_{n\to\infty} \int_\Omega 1_{\{\tau_n = T\}} d\tilde{\mathbb{P}}^n \\ &= \lim_{n\to\infty} \tilde{\mathbb{P}}^n\left(\tau_n = T\right) = \lim_{n\to\infty} \mathbb{P}\left(\tilde{\tau}_n = T\right) = 1, \end{aligned}$$

con lo que termina la prueba. □

Observación 2.87. Uno puede probar fácilmente que el resultado se extiende si quitamos la segunda hipótesis y en lugar pedimos que $Q^g Q(\beta - \gamma) = \beta - \gamma$ cuando $\beta \neq \gamma$. A pesar de su mayor generalidad, esta es una hipótesis un poco más complicada de verificar.

Capítulo 3

Representaciones explícitas y aplicaciones

3.1. Representación como cambio de tiempo en $d = 2$

Considerémonos en el contexto $X \sim WP_d\,(p, \alpha, \beta, x)$ donde $d = 2$. Inspirados en el lema 2.15, es claro que podemos considerar la siguiente representación de $[X]$,

$$[\mathrm{vec}\,(X)] \quad = \quad \begin{pmatrix} [X^{11}] & [X^{11}, X^{12}] & [X^{11}, X^{12}] & [X^{11}, X^{22}] \\ [X^{11}, X^{12}] & [X^{12}] & [X^{12}] & [X^{12}, X^{22}] \\ [X^{11}, X^{12}] & [X^{12}] & [X^{12}] & [X^{12}, X^{22}] \\ [X^{11}, X^{22}] & [X^{12}, X^{22}] & [X^{12}, X^{22}] & [X^{22}] \end{pmatrix} = \begin{pmatrix} A & B & B & E \\ B & C & C & F \\ B & C & C & F \\ E & F & F & D \end{pmatrix} \cdot Id,$$

donde,

$$\begin{aligned} A &= 4X^{11}\alpha_{11} & D &= 4X^{22}\alpha_{22} \\ B &= 2\left(X^{11}\alpha_{12} + X^{12}\alpha_{11}\right) & E &= 4X^{12}\alpha_{12} \\ C &= X^{11}\alpha_{22} + X^{22}\alpha_{11} + 2X^{12}\alpha_{12} & F &= 2\left(X^{22}\alpha_{21} + X^{21}\alpha_{22}\right). \end{aligned}$$

Para obtener sus valores propios, escribimos

$$\det \begin{pmatrix} A-\lambda & B & B & E \\ B & C-\lambda & C & F \\ B & C & C-\lambda & F \\ E & F & F & D-\lambda \end{pmatrix} = \det \begin{pmatrix} A-\lambda & B & B & E \\ B & C-\lambda & C & F \\ 0 & \lambda & -\lambda & 0 \\ E & F & F & D-\lambda \end{pmatrix},$$

lo cuál se reduce a

$$(A-\lambda)\det\begin{pmatrix} C-\lambda & C & F \\ \lambda & -\lambda & 0 \\ F & F & D-\lambda \end{pmatrix} - B\det\begin{pmatrix} B & B & E \\ \lambda & -\lambda & 0 \\ F & F & D-\lambda \end{pmatrix} - E\det\begin{pmatrix} B & B & E \\ C-\lambda & C & F \\ \lambda & -\lambda & 0 \end{pmatrix}$$

al desarrollar por menores. Esto fácilmente se puede ir desarrollando y simplificando para encontrar el polinomio característico. De hecho, el polinomio característico es

$$\begin{aligned} & (A-\lambda)\left[-\lambda(C-\lambda)(D-\lambda) - \lambda C(D-\lambda) + 2\lambda F^2\right] \\ - \quad & B\left[-2\lambda B(D-\lambda) + 2\lambda EF\right] - E\left[2\lambda BF - \lambda E(2C-\lambda)\right] \\ = \quad & \lambda^2 A(D+2C) - \lambda^3 A + 2\lambda A\left(F^2 - CD\right) + \lambda^4 - \lambda^3(D+2C) + 2\lambda^2\left(CD - F^2\right) \\ - \quad & 2\lambda^2 B^2 + 2\lambda B(BD - EF) + 2\lambda E(CE - BF) - \lambda^2 E^2, \end{aligned}$$

y al agrupar por potencias de λ, se puede escribir como

$$\begin{aligned} & \lambda\left(\lambda^3 - \lambda^2(A+D+2C) + \lambda\left(AD + 2AC + 2CD - 2F^2 - 2B^2 - E^2\right)\right. \\ & \left. 2\left(AF^2 - ACD + B^2 D - 2BEF + CE^2\right)\right). \end{aligned}$$

Se sigue que un valor propio es idénticamente nulo; mientras que los tres restantes son bastante más complejos. Por el momento vamos a asumir que α es diagonal, lo que simplificará las cuentas considerablemente, de hecho

$$\begin{aligned} A &= 4X^{11}\alpha_1 & D &= 4X^{22}\alpha_2 \\ B &= 2X^{12}\alpha_1 & E &= 0 \\ C &= X^{11}\alpha_2 + X^{22}\alpha_1 & F &= 2X^{12}\alpha_2, \end{aligned}$$

donde $\alpha_i = \alpha_{ii}$. Así, el polinomio se reduce a,

$$\lambda^3 - \lambda^2(A+D+2C) + \lambda\left(AD + 2AC + 2CD - 2F^2 - 2B^2\right) + 2\left(AF^2 - ACD + B^2 D\right),$$

lo que equivale a

$$\begin{aligned} 0 \quad = \quad & \lambda^3 - \lambda^2\left(\left(4X^{11} + 2X^{22}\right)\alpha_1 + \left(2X^{11} + 4X^{22}\right)\alpha_2\right) \\ + \quad & \lambda\left(16X^{11}X^{22}\alpha_1\alpha_2 + 8\left(X^{11}\right)^2\alpha_1\alpha_2 + 8X^{11}X^{22}\alpha_1^2\right. \\ + \quad & \left. 8\left(X^{22}\right)^2\alpha_1\alpha_2 + 8X^{11}X^{22}\alpha_2^2 - 8\left(X^{12}\right)^2\left(\alpha_1^2 + \alpha_2^2\right)\right) \\ + \quad & 32\alpha_1\alpha_2\left(X^{11}\left(X^{12}\right)^2\alpha_2 - X^{11}X^{22}\left(X^{11}\alpha_2 + X^{22}\alpha_1\right) + X^{22}\left(X^{12}\right)^2\alpha_1\right). \end{aligned}$$

Si definimos $Y = X^{11}\alpha_2$, $Z = X^{22}\alpha_1$, $W = X^{12}\sqrt{\alpha_1\alpha_2}$ y $\mu = \frac{\alpha_1}{\alpha_2}$, entonces se simplifica a

$$\begin{aligned} 0 &= \lambda^3 - \lambda^2\left(2Y(2\mu+1) + 2Z\left(2\mu^{-1}+1\right)\right) \\ &+ 8\lambda\left(2YZ + Y^2\mu + Z^2\mu^{-1} + \left(YZ - W^2\right)\left(\mu+\mu^{-1}\right)\right) \\ &+ 32\left(W^2 - YZ\right)(Y+Z). \end{aligned}$$

Más aún, si $V = YZ - W^2 = \det(X)$, entonces esto se simplifica a

$$\begin{aligned} 0 &= \lambda^3 - \lambda^2\left(4\left(Y\sqrt{\mu} + Z\sqrt{\mu}^{-1}\right) - 2(Y+Z)\right) \\ &+ \lambda\left(8\left(Y\sqrt{\mu} + Z\sqrt{\mu}^{-1}\right)^2 + 8V\left(\mu+\mu^{-1}\right)\right) \\ &- 32V(Y+Z), \end{aligned}$$

que de cierta forma sólo depende del determinante de X y de trazas ponderadas $T = Y\sqrt{\mu} + Z\sqrt{\mu}^{-1}$ y $S = Y + Z$, obteniendo

$$\lambda^3 - \lambda^2(4T - 2S) + \lambda\left(8T^2 + 8VR\right) - 32VS,$$

para $R = \left(\mu + \mu^{-1}\right)$. Como la solución sigue siendo demasiado complicada, nos veremos forzados a concentrarnos en el caso especial en que α es un múltiplo de I_d, i.e., $\alpha_1 = \alpha_2$. En este caso podemos ver fácilmente que nuestra ecuación original equivale a

$$\begin{aligned} 0 &= \lambda^3 - 6\lambda^2(Y+Z) + 8\lambda\left((Y+Z)^2 + 2\left(YZ - W^2\right)\right) + 32\left(W^2 - YZ\right)(Y+Z) \\ &= (\lambda - 2(Y+Z))\left(\lambda^2 - 4\lambda(Y+Z) + 16\left(YZ - W^2\right)\right), \end{aligned}$$

cuyas raíces son, para $i = -1, 0, 1$,

$$\lambda_{3+i} = 2\left(Y + Z + 2i\sqrt{(Y-Z)^2 + 4W^2}\right) = 2\alpha_1\left(\operatorname{tr}(X) + 2i\sqrt{\operatorname{tr}(X)^2 - 4\det(X)}\right),$$

dependientes únicamente de un múltiplo de la traza y el determinante de X. De hecho, tenemos 2 valores propios distintos (en cuyo caso son 0 y $4Y$) o 4 dependiendo de si $(Y-Z)^2 + 4W^2$ se anula o no respectivamente, i.e., sólo depende de si X es un múltiplo de la identidad o no. En busca de los vectores propios, queremos encontrar vectores $v^i = \left(v_1^i, v_2^i, v_3^i, v_4^i\right)'$ tales que

$$\lambda_i v^i = [X]\, v^i = \alpha_1 \begin{pmatrix} 4X^{11} & 2X^{12} & 2X^{12} & 0 \\ 2X^{12} & X^{11} + X^{22} & X^{11} + X^{22} & 2X^{12} \\ 2X^{12} & X^{11} + X^{22} & X^{11} + X^{22} & 2X^{12} \\ 0 & 2X^{12} & 2X^{12} & 4X^{22} \end{pmatrix} v^i$$

es decir,

$$\lambda_i v^i = \alpha_1 \begin{pmatrix} 4X^{11}v_1^i \\ 2X^{12}\left(v_1^i + v_4^i\right) \\ 2X^{12}\left(v_1^i + v_4^i\right) \\ 4X^{22}v_4^i \end{pmatrix} + \alpha_1\left(v_2^i + v_3^i\right)\begin{pmatrix} 2X^{12} \\ X^{11} + X^{22} \\ X^{11} + X^{22} \\ 2X^{12} \end{pmatrix}.$$

En particular podemos comprobar que los siguientes vectores satisfacen las ecuaciones anteriores, además de ser ortogonales entre sí,

$$v^1 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \quad v^3 = \begin{pmatrix} -2X^{12} \\ X^{11} - X^{22} \\ X^{11} - X^{22} \\ 2X^{12} \end{pmatrix},$$

$$v^{3\pm 1} = \begin{pmatrix} 2\left(X^{12}\right)^2 + \left(X^{11} - X^{22}\right)\left(X^{11} - X^{22} \pm \sqrt{\left(X^{11} - X^{22}\right)^2 + 4\left(X^{12}\right)^2}\right) \\ X^{12}\left(X^{11} - X^{22} \pm \sqrt{\left(X^{11} - X^{22}\right)^2 + 4\left(X^{12}\right)^2}\right) \\ X^{12}\left(X^{11} - X^{22} \pm \sqrt{\left(X^{11} - X^{22}\right)^2 + 4\left(X^{12}\right)^2}\right) \\ 2\left(X^{12}\right)^2 \end{pmatrix}.$$

Sin embargo, cuando sólo tenemos dos valores propios o más generalmente, cuando $X^{12} = 0$ (y nuestros valores propios son $\lambda_{3+i} = 2\left((Y+Z) + i\,(Y-Z)\right)$ para $i = -1, 0, 1$), nuestros vectores propios propuestos v^2, v^3 y v^4 coinciden o bien, $v^2 = v^4 = 0$ respectivamente. Afortunadamente, en este caso podemos simplemente tomar a los siguientes vectores propios

$$v^1 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \quad v^2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

$$v^3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad v^4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Se sigue de la continuidad de X que los procesos v^i y λ^i son todos progresivamente medibles. De hecho, con un cuidado más delicado sobre los v^i, uno puede considerar a los λ^i ordenados siempre (teniendo nuestro caso un único posible problema cuando $X^{12} = 0$) para que estos sean continuos, sin perder la medibilidad progresiva de los v^i. En cualquiera de los dos casos, consideraremos una normalización de los vectores (que, con un abuso de notación seguiremos llamando v^i) para obtener al proceso progresivamente medible en $M_{4,4}$

$$\pi = \begin{pmatrix} v^1 & v^2 & v^3 & v^4 \end{pmatrix}'.$$

Lema 3.1. *Supongamos que $X \sim WP_2(p, \alpha I_d, \beta, x)$ para algún real $\alpha > 0$ y consideremos al proceso progresivamente medible π con valores en $\mathbb{O}(4)$ descrito anteriormente y al proceso diagonal y progresivamente medible Λ, descrito por los valores propios de $\frac{d[vec(X)]_t}{dt}$. Entonces X satisface*

$$vec(X_t) = x + \int_0^t \pi_s' dB_{\int_0^s \Lambda_u du} + \int_0^t vec(2p\alpha I_d + \beta X_s + X_s\beta')\, ds,$$

donde $B = (0, B^2, B^3, B^4)'$ para movimientos brownianos independientes B^2, B^3 y B^4.

Demostración. El conjunto de resultados y cálculos que hicimos anteriormente junto con el lema 2.15 nos dan la prueba requerida. De hecho, otra forma de representar a la parte de martingala local de manera más explícita es:

$$\left(\sum_{j=2}^{4} \int_0^t \pi_s^{ji} dB^j_{\int_0^s \Lambda_u^{jj} du} \right)_{i=1}^{4}.$$

□

Observación 3.2. Aunque el resultado obtenido es bastante particular, el procedimiento nos deja ver que en general el problema no es nada trivial.

3.2. Simulación

Por nuestra construcción, cuando $2p \in \mathbb{N}$ (y más que nada $2p < d-1$) es relativamente fácil dar una simulación exacta del proceso de Wishart. Esto se sigue directamente de nuestra construcción a través de procesos Ornstein-Uhlenbeck matriciales puesto que conocemos sus transiciones, las cuáles son normales y sencillas de simular exactamente. Además, estos métodos superan a la aproximación de Euler-Maruyama porque es exacto y no hace falta hacer una partición muy fina ni calcular varias raíces cuadradas matriciales.

Por lo tanto, nuestro objetivo ahora será describir algoritmos que nos ayuden a simular el proceso cuando $2p \geq d-1$. Por esta razón asumiremos a lo largo de esta sección que este es el caso. Por supuesto, la relevancia de este capítulo yace en el caso $d > 1$, puesto que para el caso escalar ya se tienen bastantes resultados sobre simulaciones exacta con técnicas muy diversas, e.g. [31].

De hecho, si se recurre a simular las transiciones (i.e., las distribuciones Wishart), uno puede recuperar métodos de simulación para $2p \geq d-1$ de [23] donde también hay una discusión sobre los errores, las velocidades de convergencia y el problema de Cauchy en las secciones 5.5, 5.6 y 5.7. Para el cálculo de esperanzas con respecto a los procesos de Wishart, existe muy poca literatura, entre ella, [23, 3, 7].

Recordemos de la proposición 2.9 que podemos asumir sin pérdida de generalidad que la distribución a simular es de la forma $\Xi \sim W_d\left(p, x, I_{k,d}\right)$ para alguna $k \leq d$, ya que si $\Xi \sim W_d\left(p, x, \sigma\right)$ para alguna matriz $\sigma \in \bar{S}_d^+$ con rango $\operatorname{rank}\left(\sigma\right) = k$, entonces

$$\sqrt{\sigma}^g \Xi \sqrt{\sigma}^g \sim W_d\left(p, \sqrt{\sigma}^g x \sqrt{\sigma}^g, I_{k,d}\right).$$

Lo que haremos ahora es ver que hay una forma de partir el generador infinitesimal en términos de operadores que conmutan. Esta técnica se basa en descomponer la simulación de la distribución Wishart en sólo simular procesos más sencillos como el movimiento browniano multidimensional y el proceso cuadrado de Bessel (BESQ). En efecto, lo que se obtendrá al final es un algoritmo de simulación en el que se realizan ciertas permutaciones para reacomodar los índices e ir rellenando las entradas de nuestra distribución con los resultados de simulaciones más sencillas.

Como antes, tomemos $X \sim WP_d\left(p, I_{k,d}, 0, x\right)$ y consideraremos a $\mathcal{A}$ como el generador infinitesimal de X. Similarmente, denotaremos por $\mathcal{A}_i$ al generador infinitesimal de $W_d\left(p, e^i, 0, x\right)$ donde $e^i = e^{ii}$ para cada $i = 1, \ldots, d$. Notemos de nuevo que todos estos procesos existen porque $2p \geq d-1$.

Lema 3.3. *Sean $\mathcal{A}$ y $\mathcal{A}_i$ para $i = 1, \ldots, d$ como antes, entonces*

$$\mathcal{A} = \sum_{i=1}^{k} \mathcal{A}_i, \quad y \quad \mathcal{A}_i \mathcal{A}_j = \mathcal{A}_j \mathcal{A}_i.$$

Demostración. Directamente del lema 2.53 podemos observar la primer identidad puesto que $I_{k,d} = \sum_{i=1}^{k} e^i$. La conmutatividad se prueba con la misma facilidad pero con muchos más cálculos. □

El lema anterior junto con el corolario 2.82 (y sobre todo la observación del corolario) justifican la metodología que estamos por explicar para obtener una simulación exacta de las matrices Wishart. Con este objetivo en mente, definamos iterativamente para alguna matriz $x \in \bar{S}_d^+$ fija y $\Xi^0 \equiv x$,

$$\Xi^i \Big| \Xi^{i-1}, \ldots, \Xi^0 \sim W_d \left(p, \Xi^{i-1}, e^i\right), \quad i = 1, \ldots, k.$$

Proposición 3.4. *Consideremos a las variables aleatorias* (Ξ^i) *definidas como antes, entonces*

$$\Xi^k \sim W_d \left(p, x, I_{k,d}\right).$$

Demostración. Notemos recursivamente que

$$\begin{aligned} Ee^{-\left\langle u, \Xi^k \right\rangle} &= E\left(E\left(e^{-\left\langle u, \Xi^k \right\rangle} \Big| \Xi^{k-1}\right)\right) \\ &= \frac{Ee^{-\left\langle u\left(I_d + e^k u\right)^{-1}, \Xi^{k-1}\right\rangle}}{\det\left(I_d + e^k u\right)^p} \\ &\vdots \\ &= \frac{\exp\left(-\left\langle u\left(\prod_{i\leq k}\left(I_d + e^i u\right)\right)^{-1}, x\right\rangle\right)}{\det\left(\prod_{i\leq k}\left(I_d + e^i u\right)\right)^p}, \end{aligned}$$

donde uno puede notar que $\prod_{i\leq k}\left(I_d + e^i u\right)$ es igual a

$$I_d + \sum_{i\leq k} e^i u = I_d + I_{k,d} u$$

más una columna extra que es combinación lineal de las columnas de $I_{k,d}u$. Luego de unas cuentas, se puede ver que en realidad no afecta a la traza ni al determinante y concluir. Otra forma de hacerlo es tomar procesos

$$X^i \Big| X^{i-1}, \ldots, X^0 \sim WP_d\left(p, e^i, 0, X_1^{i-1}\right),$$

con $X^0 \equiv x$ de forma que $\Xi^i \stackrel{d}{=} X_1^i$ y $X \sim WP_d\left(p, I_{k,d}, 0, x\right)$. Definamos

$$C = \sup_{\substack{\deg(p)\leq m \\ t\leq 1}} E\left|p\left(X_t\right)\right|,$$

donde el supremo se toma sobre el espacio de polinomios mónicos de un término y de grado a lo más m.

Consideremos cualquier función polinomial $f : \bar{S}_d^+ \to \mathbb{R}$ de grado a lo más m y denotemos por $\|f\| = \sup_{t\leq 1} E\left|f\left(X_t\right)\right|$, el cuál es finito porque los procesos de Wishart tienen momentos de todos los órdenes. Luego, $\mathcal{A}f$ también es un polinomio de grado a lo más m, lo que implica que si

$$\|\mathcal{A}\| = \sup_{\substack{\deg(p)\leq m \\ t\leq 1}} \|\mathcal{A}p\|,$$

entonces $\|\mathcal{A}^n f\| \leq \|\mathcal{A}\|^n C$ muestra que el crecimiento de los términos es a lo más geométrico. Iterando la fórmula de Itô N veces, obtenemos

$$Ef(X_t) = \sum_{n=0}^{N} \frac{t^n}{n!}\mathcal{A}^n f(x) + \int_0^t \frac{(t-s)^N}{N!} E\mathcal{A}^{N+1} f(X_s)\, ds.$$

Concluimos que la siguiente serie converge y es igual a

$$Ef(X_t) = \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathcal{A}^n f(x),$$

y un resultado completamente análogo se tiene para cada $\mathcal{A}_i$. Luego, la conmutatividad de los $\mathcal{A}_i$ muestra que iterativamente

$$\begin{aligned}
Ef\left(X_1^k\right) &= E\left(E\left(f\left(X_1^k\right)\middle| X^{k-1}\right)\right) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!}\mathcal{A}_k^n E f\left(X_1^{k-1}\right) \\
&\vdots \\
&= \sum_{n_1,\ldots,n_k=0}^{\infty} \frac{1}{\prod_{i\leq k} n_i!} \prod_{i\leq k} \mathcal{A}_i^{n_i} f(x) \\
&= \sum_{n=0}^{\infty} \left(\sum_{i\leq k} \mathcal{A}_i\right)^n f(x) \\
&= \sum_{n=0}^{\infty} \mathcal{A}^n f(x) = Ef(X_1),
\end{aligned}$$

como se quería mostrar. □

Lo único que nos hace falta para este punto es poder hacer simulaciones del tipo $W_d(p,\omega,e^i)$, las cuáles se pueden asumir que son $W_d(p,\omega,e^1)$ con un pequeño cambio de coordenadas. Escribamos nuevamente (2.17) para nuestro caso de interés actual

$$\mathcal{A}_1 f = 2p\partial_{11} f(x) + 2x_{11}\partial_{11}^2 f(x) + 2\sum_{1<i\leq d} x_{1i}\partial_{1i}\partial_{11} f(x) + \frac{1}{2}\sum_{1<i,j\leq d} x_{ij}\partial_{1i}\partial_{1j} f(x).$$

Ahora nos dedicaremos a construir una EDE con el mismo generador, la cuál puede ser fácilmente resuelta explícitamente. Para esto, primero definamos

$$r = \operatorname{rank}\left((x_{ij})_{i,j\geq 2}\right),$$

donde claramente $0 \leq r \leq d-1$. Por el momento supongamos que existe una matriz c de la forma

$$c = \begin{pmatrix} c_r & 0 \\ k_r & 0 \end{pmatrix},$$

donde $c_r \in M_{r,r}$ es invertible y triangular inferior y donde $k_r \in M_{d-1-r,r}$, tal que $x = c_r c_r'$, donde si $r = 0$, tomamos $c = 0$.

En el caso particular $r = d-1$, tenemos $c = c_r$, en cuyo caso se trata de la descomposición de Cholesky usual. Aunque esta descomposición no siempre es posible, para nuestros fines, este caso es suficientemente general. De hecho, tenemos el siguiente resultado, cuya prueba se puede encontrar en [32] junto con un algoritmo para su cálculo, en el algoritmo 4.2.4.

Lema 3.5. *Sea $x \in \bar{S}_d^+$ de rango r, entonces existe una matriz de permutación π, una matriz invertible triangular inferior $c_r \in M_{r,r}$ y una matriz $k_r \in M_{d-1-r,r}$ tales que*

$$\pi x \pi' = cc', \quad c = \begin{pmatrix} c_r & 0 \\ k_r & 0 \end{pmatrix}.$$

A la tripleta (π, c_r, k_r) se le conoce como la descomposición extendida de Cholesky de x. Además,

$$x = (\tilde{c}'\pi)' I_{r,d} \tilde{c}' \pi, \quad \tilde{c} = \begin{pmatrix} c_r & 0 \\ k_r & I_{d-r} \end{pmatrix},$$

donde $\tilde{c}$ es de rango completo.

Con esto en mente, nos damos cuenta de que realmente no estamos en un caso restrictivo. Luego del siguiente resultado, estableceremos la metodología general de simulación.

Proposición 3.6. *Sea $x = cc'$ como arriba. Sean W y B movimientos brownianos independientes en $\mathbb{R}^r$ y $\mathbb{R}$ respectivamente, entonces existe una trayectorialmente única solución con condición inicial $X_0 = x$ al sistema de EDEs*

$$\begin{aligned} dX_t^{11} &= 2\sqrt{X_t^{11} - \sum_{i=1}^{r}\left(\sum_{j=1}^{r}\left(c_r^{-1}\right)_{ij} X_t^{1,j+1}\right)^2}\, dB_t + 2\sum_{i,j=1}^{r}\left(c_r^{-1}\right)_{ij} X_t^{1,j+1} dW_t^i + 2p\,dt, \qquad (3.1) \\ dX_t^{1,i+1} &= \left(c_r dW_t\right)^i, \quad i \leq d-1, \\ X^{ij} &\equiv x_{ij}, \qquad 2 \leq i,j \leq d. \end{aligned}$$

La solución además toma valores en $\bar{S}_d^+$, tiene el generador infinitesimal $\mathcal{A}_1$ y está dada *por*

$$\begin{aligned} X_t &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r & 0 \\ 0 & k_r & I_{d-r-1} \end{pmatrix} \begin{pmatrix} U_t + V_t'V_t & V_t' & 0 \\ V_t & I_r & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r' & k_r' \\ 0 & 0 & I_{d-r-1} \end{pmatrix}, \qquad (3.2) \\ dU_t &= 2\sqrt{U_t}dB_t + (2p-r)\,dt, \quad u = x_{11} - v'v \geq 0, \\ V &= W + v, \quad v = c_r^{-1}\left(x_{1,i+1}\right)_{i\leq r}, \end{aligned}$$

donde V se toma como vector columna en $\mathbb{R}^r$.

Demostración. Iniciemos probando que (3.1) tiene una trayectorialmente única solución fuerte y que está dada por (3.2). Para esto, supongamos que X es una solución y definamos

$$\begin{aligned} V &= c_r^{-1}\left(X^{1,i+1}\right)_{i\leq r}, \\ U &= X^{11} - \sum_{i=1}^{r}\left(\sum_{j=1}^{r}\left(c_r^{-1}\right)_{ij} X^{1,j+1}\right)^2, \\ &= X^{11} - V'V. \end{aligned}$$

Ahora usemos el que $x = cc'$ junto con las definiciones anteriores para ver que

$$\begin{aligned} &\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r & 0 \\ 0 & k_r & I_{d-r-1} \end{pmatrix} \begin{pmatrix} U_t + V_t'V_t & V_t' & 0 \\ V_t & I_r & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r' & k_r' \\ 0 & 0 & I_{d-r-1} \end{pmatrix} \\ = &\begin{pmatrix} U_t + V_t'V_t & V_t'c_r' & V_t'k_r' \\ c_rV_t & c_rc_r' & c_rk_r' \\ k_rV_t & k_rc_r' & 0 \end{pmatrix} = X_t, \end{aligned}$$

es decir que la primera parte de (3.2) se satisface. Como la primer matriz en la ecuación anterior es invertible, tenemos que $X_t \in \bar{S}_d^+$ si y sólo si para todo $z \in \mathbb{R}^d$ se tiene

$$z' \begin{pmatrix} U_t + V_t'V_t & V_t' & 0 \\ V_t & I_r & 0 \\ 0 & 0 & 0 \end{pmatrix} z = z_1^2 U_t + \sum_{i=1}^{r}\left(z_{i+1} + V_i z_1\right)^2 \geq 0,$$

lo cuál es cierto si y sólo si $U_t \geq 0$.

Ahora, notemos que $U_0 \geq 0$ porque $x \in \bar{S}_d^+$ mientras que $dV_t = c_r^{-1}c_r dW_t = dW_t$, que junto con la condición inicial, prueban la última parte de (3.2). Finalmente, notemos que por como

$$dX_t^{11} = 2\sqrt{U_t}dB_t + 2V_t'dV_t + 2pdt$$

y por la fórmula de integración por partes,

$$\begin{aligned} dU_t &= dX_t^{11} - dV_t'V_t - V_t'dV_t - dV_t'dV_t \\ &= 2\sqrt{U_t}dB_t + 2pdt - dW_t'dW_t \\ &= 2\sqrt{U_t}dB_t + (2p-r)\,dt, \end{aligned}$$

lo que prueba (3.2) y en particular muestra que $U \geq 0$ porque $2p - r \geq 0$ y por lo tanto $X_t \in \bar{S}_d^+$ c.s., como queríamos probar. Probar que (3.2) efectivamente es una solución es simple una vez que se sustituye todo.

Ahora probaremos que su generador infinitesimal es el deseado. Para esto, podemos repetir los argumentos que usamos en la segunda prueba del lema 2.53 y sólo enfocarnos en la parte de la difusión y en su variación cuadrática, verificando que coincidan con los de $\mathcal{A}_1$. Como la deriva es $2pe^1Id$, claramente coincide con la deseada, por lo que sólo calcularemos las variaciones y covariaciones cuadráticas:

$$\begin{aligned} d\left[X^{11}\right]_t &= 4U_tdt + 4V_t'V_tdt = 4X_t^{11}dt, \\ d\left[X^{1,i+1}, X^{1,j+1}\right]_t &= (c_rdW_t)^i\,(c_rdW_t)^j = \left(c_rc_r'\right)_{ij}dt = X_t^{ij}dt, \quad i,j \leq d \\ d\left[X^{11}, X^{1,i+1}\right]_t &= (c_rdW_t)^i\, 2\sum_{j,l=1}^{r}\left(c_r^{-1}\right)_{jl}X_t^{1,l+1}dW_t^j = 2X_t^{1,i}dt, \quad i \leq r \\ d\left[X^{11}, X^{1,i+1}\right]_t &= 2\sum_{j,l=1}^{r}k_r^{i-r,j}\left(c_r^{-1}\right)_{jl}X_t^{1,l+1}dt \\ &= 2\sum_{l=1}^{r}\left(k_rc_r^{-1}\right)_{i-r,l}X_t^{1,l+1}dt = 2X_t^{1,i}dt, \quad i > r, \end{aligned}$$

donde en la segunda igualdad usamos la tercera parte de (3.1). Se sigue de inmediato el que su generador es $\mathcal{A}_1$, concluyendo la prueba. □

Observación 3.7. Notemos que para nuestros fines, sólo hace falta hacer la simulación a un tiempo específico que podemos suponer sin pérdida de generalidad, es 1. Para esto, sólo requerimos simular a U_1 y V_1. El segundo se logra por medio de vaiid $N\,(0,1)$, mientras que $U_1 \sim \chi_{2p-r}\,(u)$ ya que U es un proceso cuadrado de Bessel.

Algoritmo 3.8. *(simulación Wishart exacta $2p \geq d-1$) Sean $\omega, \sigma \in \bar{S}_d^+$ y $2p \geq d-1$. Suponga que se quiere simular $\Xi \sim W_d\,(p, \omega, \sigma)$.*

(I) Calcule $\rho \in M_{d,d}$ tales que $\rho\sigma\rho^{-1} = I_{k,d}$ donde $k = \operatorname{rank}(\sigma)$ y definamos $x = \rho\omega\rho^{-1}$.

(II) Para cada $m = 0, \ldots, k-1$ repita:

a*)* Calcule a (π, c_r, k_r), la descomposición extendida de Cholesky de $(x_{ij})_{i,j \geq 2}$ (ver algoritmo en [32])

b*)* Defina $\tilde{\pi} = \begin{pmatrix} 1 & 0 \\ 0 & \pi \end{pmatrix}$, $\tilde{x} = \tilde{\pi} x \tilde{\pi}'$, $v = c_r^{-1} (\tilde{x}_{1,l+1})_{l \leq r}$ y $u = \tilde{x}_{11} - v'v \geq 0$.

c*)* Simule W_1 como r vaiid $N(0,1)$ y a $U_1 \sim \chi_{2p-r}(u)$ (por ejemplo, ver [31]) y defina $V_1 = v + W_1$, así como

$$X_1 = \tilde{\pi}' \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r & 0 \\ 0 & k_r & I_{d-r-1} \end{pmatrix} \begin{pmatrix} U_1 + V_1'V_1 & V_1' & 0 \\ V_1 & I_r & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_r' & k_r' \\ 0 & 0 & I_{d-r-1} \end{pmatrix} \tilde{\pi}.$$

d*)* Si $m < k$, tome a la matriz de permutación q dada por $q_{k-m,1} = q_{1,k-m} = q_{ii}$ para todo $i \notin \{k-m, 1\}$ y con $q_{ij} = 0$ en otro caso (permutación de las coordenadas 1 y $k-m$) y redefina $x = qX_1q$.

(III) Tome $\Xi = \rho^{-1} x \rho$.

3.3. Estimación de los parámetros

Para poder tener aplicaciones, es importante poder hacer estimación de los parámetros del modelo. El método más comúnmente utilizado debido a sus fuertes resultados asintóticos (y posiblemente uno de los más populares) es el de máxima verosimilitud. No obstante, las transiciones Wishart sólo cuentan con densidad cuando $2p > d-1$ e incluso entonces, esta tiene una forma composicional muy complicada.

De hecho, en el caso $d = 1$ (en el que nos encontramos con el modelo CIR), podemos comprobar en [33] (y en el paquete de R que ofrece el mismo autor) que las densidades de transición son complejas. Estas densidades están dadas en términos de sumas infinitas, lo que convierte en un problema muy complicado maximizar la verosimilitud. El camino que se toma es truncar la suma y optimizar numéricamente.

Esto da buenos resultados; sin embargo, estos son computacionalmente costosos y no son del todo excepcionales, después de todo, el autor mismo declara en https://goo.gl/7e1XYv, al presentar su paquete de R: «(...) for CIR you might want to use the exact law of the increments (i.e. the exact transition density)».

En nuestro contexto con $d \in \mathbb{N}$, incluso teniendo densidades, la optimización de la verosimilitud es considerablemente compleja, bastante más que en el caso $d = 1$. Esto nos lleva a tomar el otro camino usual, a saber, discretizar y establecer un modelo lineal o algo semejante. Supongamos que tenemos observaciones en un conjunto de tiempos finito $N \subset \mathbb{Q}_+$ con separación constante Δn, esto es,

$$N = \{k\Delta n : k = 0, \ldots, |N| - 1\}.$$

Recordando que si tenemos una semimartingala de Wishart, entonces para cualquier Q con $Q'Q = \alpha$ podíamos conseguir la existencia del movimiento browniano B con respecto al cual se satisface la EDE de Wishart.

Para evitar este problema de identificabilidad de Q, lo que haremos es considerar siempre $Q = \sqrt{\alpha}$ i.e., que Q es simétrica y positiva definida. De esta forma, la discretización de la EDE de Wishart es:

$$\Delta X_n = \sqrt{X_n}\Delta B_n Q + Q\Delta B_n' \sqrt{X_n} + \left(2p\alpha + \beta X_n + X_n \beta'\right) \Delta n.$$

Lo que debemos conseguir ahora es aislar a los incrementos del browniano ΔB_n, para que el estimador máximo verosímil coincida con el de mínimos cuadrados (cuyo cálculo, *a priori,* no es tan complicado). Para esto, haremos uso del producto de Kronecker y más específicamente, de la simetrización $S_{A,B}$ antes introducida justo antes de (2.2):

$$\operatorname{vec}(\Delta X_n) \quad = \quad S_{\sqrt{X_n},Q}\operatorname{vec}(\Delta B_n) + (2p\operatorname{vec}(\alpha) + S_{I_d,X_n}\operatorname{vec}(\beta))\,\Delta n.$$

Abreviamos $Z_n = S_{I_d,X_n}$ y $Y_n = S_{\sqrt{X_n},Q}$, con las que podemos reescribir:

$$Y_n^g\left[\operatorname{vec}(\Delta X_n) - (2p\operatorname{vec}(\alpha) + Z_n\operatorname{vec}(\beta))\,\Delta n\right] = Y_n^g Y_n \operatorname{vec}(\Delta B_n)\,. \tag{3.3}$$

Aquí es importante notar que $Y_n^g Y_n$ (la cuál vale $\frac{1}{2}S_{I_d,I_d}$ si X_n es invertible) es una proyección ortogonal, es decir que la podemos ver como $P_n' I_{d_n,d} P_n$ donde P_n es una matriz ortogonal y donde $I_{d_n,d}$ es una matriz diagonal con d_n unos y $d - d_n$ ceros.

Como $P_n\operatorname{vec}(\Delta B_n) \overset{d}{=} \operatorname{vec}(\Delta B_n)$, entonces $Y_n^g Y_n \operatorname{vec}(\Delta B_n)$ consta de una colección variables aleatorias gaussianas con matriz de covarianzas $Y_n^g Y_n$ (obteniendo el mismo efecto que en la representación como cambio de tiempo, donde observamos que a lo más requerimos de $\frac{d(d+1)}{2}$ movimientos brownianos escalares independientes) que suman a d_n

vaiid gaussianas.

Luego, si queremos optimizar la verosimilitud tendríamos que seguir un camino relativamente complejo puesto que Y_n depende de Q. No obstante, dado que el lado izquierdo se encuentra en la imágen de Y_n^g, este sistema de ecuaciones se puede reescribir en términos de d_n ecuaciones igualadas a ruidos gaussianos independientes e idénticamente distribuidos.

Por esta razón queremos minimizar la suma de cuadrados de $I_{d.,d}P_n\text{vec}\left(\Delta B_n\right)$, que, como P_n' es ortogonal, esto equivale a minimizar la suma de cardados del lado izquierdo de (3.3). Sea $V_n = Y_n^g Z_n$, entonces la ecuación (3.3) se convierte en

$$Y_n^g\text{vec}\left(\frac{\Delta}{\Delta n}X_n\right) - 2pY_n^g\text{vec}\left(\alpha\right) - V_n\text{vec}\left(\beta\right) = Y_n^g Y_n\text{vec}\left(\frac{\Delta}{\Delta n}B_n\right).$$

Denotemos $U_n = 2pY_n^g\text{vec}\left(\alpha\right) - Y_n^g\text{vec}\left(\frac{\Delta}{\Delta n}X_n\right)$ y notemos que debemos minimizar

$$\begin{aligned} f\left(p, Q, \beta\right) &:= \sum_n \left\|U_n + V_n\text{vec}\left(\beta\right)\right\|^2 \\ &= \sum_n \sum_{i=1}^{d^2}\left(U_n^i + \sum_{k,l} V_n^{i,\upsilon(k,l)}\beta_{k,l}\right)^2. \end{aligned} \tag{3.4}$$

Derivando parcialmente respecto de cada $\beta_{r,s}$ e igualando a 0 obtenemos,

$$\begin{aligned} 0 &= \frac{\partial}{\partial\beta_{r,s}} f\left(p, Q, \beta\right) \\ &= \sum_n \sum_{i=1}^{d^2} \frac{\partial}{\partial\beta_{r,s}}\left(U_n^i + \sum_{k,l} V_n^{i,\upsilon(k,l)}\beta_{k,l}\right)^2 \\ &= \sum_n \sum_{i=1}^{d^2} 2V_n^{i,\upsilon(r,s)}\left(U_n^i + \sum_{k,l} V_n^{i,\upsilon(k,l)}\beta_{k,l}\right) \\ &= 2\sum_n \sum_{i=1}^{d^2} V_n^{i,\upsilon(r,s)}U_n^i + 2\sum_{k,l}\sum_n\sum_{i=1}^{d^2} V_n^{i,\upsilon(r,s)}V_n^{i,\upsilon(k,l)}\beta_{k,l} \\ &= 2\sum_n \left(V_n'U_n\right)^{\upsilon(r,s)} + 2\sum_{k,l}\left(\sum_n \left(V_n'V_n\right)^{\upsilon(r,s),\upsilon(k,l)}\right)\beta_{k,l} \\ &= 2\sum_n \left(V_n'U_n\right)^{\upsilon(r,s)} + 2\left(\left(\sum_n V_n'V_n\right)\text{vec}\left(\beta\right)\right)^{\upsilon(r,s)}. \end{aligned}$$

Matricialmente esto se traduce en

$$-\sum_n V_n'U_n = \left(\sum_n V_n'V_n\right)\text{vec}\left(\beta\right),$$

cuya solución es

$$\beta\left(p,Q\right):=-\mathrm{vec}^{-1}\left(S_1\left(Q\right)S_2\left(p,Q\right)\right),$$

donde $S_1\left(Q\right)=\left(\sum_n V_n'V_n\right)^{-1}$ y

$$\begin{aligned} S_2\left(p,Q\right) &= \sum_n V_n'U_n \\ &= \sum_n V_n'\left(2pY_n^g\mathrm{vec}\left(\alpha\right)-Y_n^g\mathrm{vec}\left(\frac{\Delta}{\Delta n}X_n\right)\right) \\ &= \sum_n 2p\left(\sum_n V_n'Y_n^g\right)\mathrm{vec}\left(\alpha\right)-V_n'Y_n^g\mathrm{vec}\left(\frac{\Delta}{\Delta n}X_n\right) \\ &= 2pS_3\left(Q\right)\mathrm{vec}\left(\alpha\right)-S_4\left(Q\right), \end{aligned}$$

donde $S_3\left(Q\right)=\sum_n V_n'Y_n^g$ y $S_4\left(Q\right)=\sum_n V_n'Y_n^g\mathrm{vec}\left(\frac{\Delta}{\Delta n}X_n\right)$. Ahora hagamos lo mismo respecto de p en $\left[\frac{d-1}{2},\infty\right)$ y comparemos con las evaluaciones de f para $p\in\left\{\frac{1}{2},\ldots,\frac{d-2}{2}\right\}$. Procedamos con algunos cálculos (recordando que sólo U_n depende de p),

$$\begin{aligned} 0 &= \left.\frac{\partial}{\partial p}f\left(p,Q,\beta\right)\right|_{\beta=\beta(p,Q)} \\ &= \sum_n\sum_{i=1}^{d^2}\left.\frac{\partial}{\partial p}\left(U_n^i+\left(V_n\mathrm{vec}\left(\beta\left(p,Q\right)\right)\right)^i\right)^2\right|_{\beta=\beta(p,Q)} \\ &= \sum_n\sum_{i=1}^{d^2}2\left(2Y_n^g\mathrm{vec}\left(\alpha\right)\right)^i\left(U_n+V_n\mathrm{vec}\left(\beta\left(p,Q\right)\right)\right)^i \\ &= 4\sum_n\left(Y_n^g\mathrm{vec}\left(\alpha\right)\right)'\left(U_n-V_nS_1\left(Q\right)S_2\left(p,Q\right)\right). \end{aligned}$$

Si ahora definimos $S_5\left(Q\right)=\sum_n Y_n^{g\prime}Y_n^g\mathrm{vec}\left(\frac{\Delta}{\Delta n}X_n\right)$ y $S_6\left(Q\right)=\sum_n Y_n^{g\prime}Y_n^g$, obtenemos fácilmente

$$\begin{aligned} 0 &= \sum_n\mathrm{vec}\left(\alpha\right)'Y_n^{g\prime}\left(2pY_n^g\mathrm{vec}\left(\alpha\right)-Y_n^g\mathrm{vec}\left(\frac{\Delta}{\Delta n}X_n\right)-V_nS_1\left(Q\right)S_2\left(p,Q\right)\right) \\ &= \mathrm{vec}\left(\alpha\right)'\left(2p\left(\sum_n Y_n^{g\prime}Y_n^g\right)\mathrm{vec}\left(\alpha\right)\right. \\ &- \left.\sum_n Y_n^{g\prime}Y_n^g\mathrm{vec}\left(\frac{\Delta}{\Delta n}X_n\right)-\left(\sum_n V_n'Y_n^g\right)'S_1\left(Q\right)S_2\left(p,Q\right)\right) \\ &= \mathrm{vec}\left(\alpha\right)'\left(2pS_6\left(Q\right)\mathrm{vec}\left(\alpha\right)-S_5\left(Q\right)+S_3\left(Q\right)'S_1\left(Q\right)\left(S_4\left(Q\right)-2pS_3\left(Q\right)\mathrm{vec}\left(\alpha\right)\right)\right) \\ &= \mathrm{vec}\left(\alpha\right)'\left(S_3\left(Q\right)'S_1\left(Q\right)S_4\left(Q\right)-S_5\left(Q\right)+2p\left(S_6\left(Q\right)-S_3\left(Q\right)'S_1\left(Q\right)S_3\left(Q\right)\right)\mathrm{vec}\left(\alpha\right)\right). \end{aligned}$$

Es decir que si definimos a

$$p_0(Q) := \frac{\operatorname{vec}(\alpha)' \left(S_3(Q)' S_1(Q) S_4(Q) - S_5(Q)\right)}{2\operatorname{vec}(\alpha)' \left(S_3(Q)' S_1(Q) S_3(Q) - S_6(Q)\right) \operatorname{vec}(\alpha)},$$

entonces el mínimo se da en $p(Q)$, dado por $p_0(Q)$ si $p_0(Q) \geq \frac{d-1}{2}$ o en algún $p \in \left\{\frac{1}{2}, \ldots, \frac{d-2}{2}\right\}$ en caso contrario.

De hecho, como $f(p, Q, \beta(p, Q))$ es una suma de términos de la forma $(a + bp)^2$, es claro que es convexa respecto de p y por lo tanto en $p_0(Q)$ se obtiene el mínimo global. Así, f es de la forma $ap^2 + bp + c$ y es por lo tanto simétrica; de tal suerte que $p(Q) = \arg\operatorname{ínf}_{q \in \Delta_d} |p_0(Q) - q|$.

La optimización respecto de Q es más complicada por el papel que juegan V_n, Y_n, $p(Q)$ y $\beta(p(Q), Q)$ en él. De resulta extremadamente complicado derivar respecto de las entradas de Q y obtener algo con representación de donde podamos despejar a Q fácilmente, por lo tanto optimizaremos numéricamente a $f(Q) := f(p(Q), Q, \beta(p(Q), Q))$.

Por supuesto, podemos calcular la derivada e intentar encontrar el cero numéricamente o realizar evaluaciones hasta encontrar la que minimice f. Los cálculos computacionalmente complicados (i.e., posiblemente largos, dependiendo de N) recaen principalmente sobre:

$$\begin{aligned}
S_1(Q) &= \left(\sum_n V_n' V_n\right)^{-1} = \left(\sum_n Z_n' Y_n^{g\prime} Y_n^g Z_n\right)^{-1} \\
S_3(Q) &= \sum_n V_n' Y_n^g = \sum_n Z_n' Y_n^{g\prime} Y_n^g \\
S_4(Q) &= \sum_n V_n' Y_n^g \operatorname{vec}\left(\frac{\Delta}{\Delta n} X_n\right) = \sum_n Z_n' Y_n^{g\prime} Y_n^g \operatorname{vec}\left(\frac{\Delta}{\Delta n} X_n\right) \\
S_5(Q) &= \sum_n Y_n^{g\prime} Y_n^g \operatorname{vec}\left(\frac{\Delta}{\Delta n} X_n\right) \\
S_6(Q) &= \sum_n Y_n^{g\prime} Y_n^g,
\end{aligned}$$

que, si definimos $W_n = Y_n^{g\prime} Y_n^g = (Y_n Y_n')^g$, se pueden simplificar a

$$\begin{aligned}
S_1(Q) &= \left(\textstyle\sum_n Z_n' W_n Z_n\right)^{-1} & S_5(Q) &= \textstyle\sum_n W_n \operatorname{vec}\left(\frac{\Delta}{\Delta n} X_n\right) \\
S_3(Q) &= \textstyle\sum_n Z_n' W_n & S_6(Q) &= \textstyle\sum_n W_n. \\
S_4(Q) &= \textstyle\sum_n Z_n' W_n \operatorname{vec}\left(\frac{\Delta}{\Delta n} X_n\right)
\end{aligned}$$

De esta forma, toda la dependencia de las funciones S_k para $k = 1, \ldots, 6$ en Q se absorbe por las W_n. Se hacen todas estas consideraciones para la implementación numérica en la optimización de $f(Q) := f(p(Q), Q, \beta(p(Q), Q))$ dada por:

$$\begin{aligned} f(Q) &= \sum_n \left\| 2p(Q) Y_n^g \mathrm{vec}(\alpha) - Y_n^g \mathrm{vec}\left(\frac{\Delta}{\Delta n} X_n\right) \right. \\ &+ \left. Y_n^g Z_n S_1(Q)(S_4(Q) - 2p(Q) S_3(Q)\mathrm{vec}(\alpha)) \right\|^2, \end{aligned}$$

y en el cálculo de $p(Q)$ y $\beta(p(Q), Q)$.

Como cualquier método numérico requiere de una estimación inicial, queremos una semilla que a priori sea buena, para que el algoritmo sea más eficiente. Para esto requerimos una estimación inicial de α y lo que haremos será discretizar la identidad

$$d\left[X^{ij}, X^{kl}\right]_t = \left(X_t^{ik}\alpha_{jl} + X_t^{il}\alpha_{jk} + X_t^{jk}\alpha_{il} + X_t^{jl}\alpha_{ik}\right) dt.$$

Esto es, quisiéramos que

$$g(\alpha) = \sum_n \sum_{ijkl} \left(X_n^{ik}\alpha_{jl} + X_n^{il}\alpha_{jk} + X_n^{jk}\alpha_{il} + X_n^{jl}\alpha_{ik} - \frac{\Delta X_n^{ij} \Delta X_n^{kl}}{\Delta n}\right)^2 = 0,$$

y por lo tanto, el camino natural es buscar minimizar a g. Aunque en general se puede usar cualquier otra distancia, como los cálculos salen bastante sencillos con la euclidiana, esta es la que se presenta en este trabajo.

Para minimizar g, derivemos respecto de cada α_{rs} e igualemos a cero para obtener

$$\begin{aligned} 0 &= \frac{\partial}{\partial \alpha_{rs}} g(\alpha) \\ &= \sum_n \sum_{ijkl} 2\left(X_n^{ik}\alpha_{jl} + X_n^{il}\alpha_{jk} + X_n^{jk}\alpha_{il} + X_n^{jl}\alpha_{ik} - \frac{\Delta X_n^{ij}\Delta X_n^{kl}}{\Delta n}\right) \\ &\times \left(X_n^{ik}\delta_{jr}\delta_{ls} + X_n^{il}\delta_{jr}\delta_{ks} + X_n^{jk}\delta_{ir}\delta_{ls} + X_n^{jl}\delta_{ir}\delta_{ks}\right) \\ &= 8\sum_n \sum_{ik} X_n^{ik}\left(X_n^{ik}\alpha_{rs} + X_n^{is}\alpha_{rk} + X_n^{rk}\alpha_{is} + X_n^{rs}\alpha_{ik} - \frac{\Delta X_n^{ir}\Delta X_n^{ks}}{\Delta n}\right) \\ &= 8\sum_n \left(\mathrm{tr}\left(X_n^2\right)\alpha_{rs} + \left(\alpha X_n^2\right)^{rs} + \left(X_n^2\alpha\right)^{rs} + X_n^{rs}\mathrm{tr}(X_n\alpha) - \frac{(\Delta X_n X_n \Delta X_n)^{rs}}{\Delta n}\right), \end{aligned}$$

cuya expresión matricial es

$$0 = \sum_n \mathrm{tr}\left(X_n^2\right)\alpha + \alpha X_n^2 + X_n^2\alpha + X_n\mathrm{tr}(X_n\alpha) - \frac{\Delta X_n X_n \Delta X_n}{\Delta n}.$$

donde $\mathrm{tr}(X_n\alpha) = \sum_{ik} X_n^{ik}\alpha_{ik} = \mathrm{vec}(X_n)'\,\mathrm{vec}(\alpha)$. Si ahora vectorizamos y hacemos uso de las propiedades del producto de Kronecker, podemos reescribirlo como

$$\operatorname{vec}\left(\sum_n \frac{\Delta X_n X_n \Delta X_n}{\Delta n}\right) = \left(\sum_n \operatorname{tr}\left(X_n^2\right) I_{d^2} + \left(X_n^2 \otimes I_d\right) + \left(I_d \otimes X_n^2\right) + \operatorname{vec}\left(X_n\right) \operatorname{vec}\left(X_n\right)'\right) \operatorname{vec}\left(\alpha\right).$$

Ahora observemos que

$$\operatorname{tr}\left(X_n^2\right) I_{d^2}, \left(X_n^2 \otimes I_d\right), \left(I_d \otimes X_n^2\right), \operatorname{vec}\left(X_n\right) \operatorname{vec}\left(X_n\right)' \in \bar{S}_d^+,$$

y por lo tanto

$$\sum_n \operatorname{tr}\left(X_n^2\right) I_{d^2} + \left(X_n^2 \otimes I_d\right) + \left(I_d \otimes X_n^2\right) + \operatorname{vec}\left(X_n\right) \operatorname{vec}\left(X_n\right)' \in \bar{S}_d^+,$$

donde sabemos que está en S_d^+ si al menos uno de los términos es positivo definido. En caso contrario, es bastante complicado determinar si es o no positivo definido, pero para fines de encontrar una única solución, asumiremos que lo es y que por lo tanto podemos calcular la estimación inicial $\hat{\alpha}$ a través de

$$\begin{aligned} \operatorname{vec}\left(\hat{\alpha}\right) &= \left(\sum_n \operatorname{tr}\left(X_n^2\right) I_{d^2} + \left(X_n^2 \otimes I_d\right) + \left(I_d \otimes X_n^2\right) + \operatorname{vec}\left(X_n\right) \operatorname{vec}\left(X_n\right)'\right)^{-1} \\ &\times \operatorname{vec}\left(\sum_n \frac{\Delta X_n X_n \Delta X_n}{\Delta n}\right). \end{aligned}$$

Asimismo, podemos tomar $\hat{Q} := \sqrt{\hat{\alpha}}$ como estimador inicial y proceder con una optimización numérica de $f\left(Q\right)$. Alternativamente, si el programa es muy lento, uno puede simplemente tomar como parámetros estimados a la terna $\left(\hat{p}, \hat{Q}, \hat{\beta}\right)$ donde $\hat{p} = p\left(\hat{Q}\right)$ y $\hat{\beta} = \beta\left(p\left(\hat{Q}\right), \hat{Q}\right)$.

Algunas consideraciones se pueden tener en el caso en que estemos seguros de que $2p \geq d-1$, es decir, cuando nuestras observaciones X_n son todas positivas definidas (i.e., invertibles). Usando (2.2), vemos que

$$\begin{aligned} W_n \operatorname{vec}\left(\alpha\right) &= \left(Y_n Y_n'\right)^g \operatorname{vec}\left(\alpha\right) = \frac{1}{4} \operatorname{vec}\left(X_n^{-1}\right) \\ \operatorname{vec}\left(\alpha\right)' W_n &= \left(\left(Y_n Y_n'\right)^g \operatorname{vec}\left(\alpha\right)\right)' = \frac{1}{4} \operatorname{vec}\left(X_n^{-1}\right)'. \end{aligned}$$

Algunos cálculos similares muestran que entonces

$$\begin{aligned} \left(Z_n' \left(Y_n Y_n'\right)^g \operatorname{vec}\left(\alpha\right)\right)^{\upsilon(i,j)} &= \frac{1}{4} \sum_{kl} \left(X_n^{jl} \delta_{ik} + X_n^{jk} \delta_{il}\right) \left(X_n^{-1}\right)^{kl} \\ &= \frac{2}{4} \delta_{ij} = \left(\frac{1}{2} \operatorname{vec}\left(I_d\right)\right)_{\upsilon(i,j)}, \end{aligned}$$

es decir que $Z_n' \left(Y_n Y_n'\right)^g \operatorname{vec}(\alpha) = \frac{1}{2}\operatorname{vec}(I_d)$ y por lo tanto tenemos

$$\begin{aligned} S_3(Q)\operatorname{vec}(\alpha) &= \frac{|N|}{2}\operatorname{vec}(I_d) \\ \operatorname{vec}(\alpha)' S_6(Q)\operatorname{vec}(\alpha) &= \frac{1}{4}\operatorname{tr}\left(\left(\sum_n X_n^{-1}\right)\alpha\right). \end{aligned}$$

Esto puede simplificar las formas

$$\begin{aligned} p_0(Q) &= \frac{|N|\operatorname{vec}(I_d)' S_1(Q) S_4(Q) + \frac{|N|}{2}\sum_n \operatorname{tr}\left(X_n^{-1}\frac{\Delta}{\Delta n}X_n\right)}{|N|^2\operatorname{vec}(I_d)' S_1(Q)\operatorname{vec}(I_d) + \operatorname{tr}\left(\left(\sum_n X_n^{-1}\right)\alpha\right)}, \quad \text{y} \\ \beta(Q) &= \operatorname{vec}^{-1}\left(S_1(Q)\left(S_4(Q) - |N|\, p(Q)\operatorname{vec}(I_d)\right)\right), \end{aligned}$$

con el fin de tener menor cantidad de errores por truncamientos de la computadora.

3.4. El modelo Heston-Wishart

Uno podría, en principio, plantear modelar un conjunto de volatilidades interactuando con la ayuda de procesos de Wishart, siguiendo la idea de Cox, Ingersoll y Ross (ver [23]) para el caso de una única volatilidad. Esto tendría sentido en el caso en que se tienen posiciones en activos o tasas de distintas nacionalidades o en diferentes divisas.

La interpretación de esto es que las entradas diagonales simbolizan las volatilidades de cada uno de los activos, mientras que las entradas fuera de la diagonal miden los efectos entre estas volatilidades, de la misma forma en que las covarianzas relacionan las varianzas de diferentes variables aleatorias.

Por ejemplo, si $X_t \sim WP_d(p, \alpha, \beta, x)$ está modelando una colección de d volatilidades y sus interacciones y en nuestro portafolio contamos con una ponderación en diferentes activos de valores S_t^i, entonces la volatilidad y covolatilidad del proceso están dados por la variación y covariación cuadrática del proceso logaritmo de los precios.

Siguiendo las ideas del modelo Black-Scholes y Heston (ver [23]), querríamos modelar d activos a través de un proceso estocástico S con valores en $\mathbb{R}_+^d$ cuya dinámica estocástica satisface una EDE. Para especificar la misma, lo que haremos es dar la dinámica del proceso logaritmo de los precios. La motivación es clara: el proceso logaritmo de los precios satisface propiedades mucho más sencillas de analizar.

Como en muchas otras situaciones, la extensión de los modelos al caso multidimensional nunca está dada de manera única. Aquí sólo presentaremos una posible generalización del modelo de Heston al caso vectorial. Otros enfoques (equivalentes a casos particulares de este) los introdujeron Gourieroux y Sufana en [34, 35], Gnoatto en [36], Ahdida, Alfonsi y Palidda en [37] y Alfonsi en [23].

Aunque este enfoque está aún es su etapa experimental, los procesos de Wishart ya han sido usados para modelar la volatilidad de un único activo. Comúnmente se usa la traza del proceso en lugar del modelo CIR para modelar la volatilidad, lo que preserva el carácter afín pero introduce efectos diversos (ver 2.49). Esta ligera extensión al modelo de Heston unidimensional ha sido planteado y descrito por Da Fonseca, Grasselli y Tebaldi en [38], donde también hay experimentos numéricos con datos del mercado, y por Benabid, Bensusan y El Karoui en [39].

La existencia fuerte y unicidad trayectorial de las soluciones viene dada por el teorema 2.21 y no nos concentraremos tanto en ello. En lo que sigue, denotaremos por $\operatorname{diag}(A)$ al vector columna cuyas entradas son los elementos de la diagonal de $A \in M_{n,n}$. Similarmente, para $x \in \mathbb{R}^n$, denotaremos por $\operatorname{diag}^{-1}(x)$ a la matriz diagonal en $M_{n,n}$ tal que $\operatorname{diag}(A) = x$. Observemos que diag^{-1} no es una verdadera inversa, a pesar de la notación que usamos, porque siempre regresa matrices diagonales.

Para $X \sim WP_d(p, \alpha, \beta, x_0)$, un vector $\nu \in \mathbb{R}^d$ y una matriz $\mu \in M_{d,d}$, consideraremos el sistema de EDE dado por

$$\begin{aligned} dZ_t &= \left(\mu Z_t + \nu - \frac{1}{2}\operatorname{diag}(X_t)\right) dt + \sqrt{X_t}\, dW_t, \quad Z_0 = z_0, \\ dX_t &= \sqrt{X_t}\, dB_t Q + Q' dB_t' \sqrt{X_t} + \left(2p\alpha + \beta X_t + X_t \beta'\right) dt, \end{aligned} \tag{3.5}$$

donde W es un movimiento browniano en $\mathbb{R}^d$ y donde

$$[W, B] = [W, \operatorname{vec}(B)] = \rho I d,$$

para una matriz $\rho \in M_{d,d^2}$ tal que

$$\begin{pmatrix} I_d & \rho \\ \rho' & I_{d^2} \end{pmatrix} \in S^+_{d^2+d}.$$

Nos referiremos a (3.5) como el sistema de EDEs de Heston-Wishart. Como mencionamos, la existencia el proceso se sigue de la existencia de (W, X, B) y del teorema 2.21. Esto se consigue notando que el coeficiente f es globalmente Lipschitz en

$$dZ_t' = f(Z_t')\, dU_t$$

donde

$$\begin{aligned} U_t &= \left(\mu' t, \operatorname{diag}^{-1}\int_0^t (\nu - \operatorname{diag}(X_s))\, ds, \operatorname{diag}^{-1}\int_0^t dW_s'\sqrt{X_s}\right)' \\ f(Z_t') &= (Z_t', \mathbf{1}'), \end{aligned}$$

donde $\mathbf{1}$ es el vector en $\mathbb{R}^n$ (para una n apropiada, en este caso $n = 2d$) que tiene un 1 en cada entrada; esto es directo puesto que $dZ_t' = \left(Z_t'\mu' + \nu' - \operatorname{diag}(X_t)'\right) dt + dW_t'\sqrt{X_t}$. No obstante, en nuestro caso podemos conseguir una fórmula explícita para la solución; definamos

$$Z_t = e^{t\mu}\left(z_0 + \int_0^t e^{-s\mu}\left(\nu - \frac{1}{2}\operatorname{diag}(X_s)\right) ds + \int_0^t e^{-s\mu}\sqrt{X_s}dW_s\right),$$

y verifiquemos con la fórmula de integración por partes que

$$\begin{aligned} dZ_t &= \mu e^{t\mu}\left(z_0 + \int_0^t e^{-s\mu}\left(\nu - \frac{1}{2}\operatorname{diag}(X_s)\right) ds + \int_0^t e^{-s\mu}\sqrt{X_s}dW_s\right) dt \\ &+ e^{t\mu}\left(e^{-t\mu}\left(\nu - \frac{1}{2}\operatorname{diag}(X_t)\right) dt + e^{-t\mu}\sqrt{X_t}dW_t\right) \\ &= \mu Z_t dt + \left(\nu - \frac{1}{2}\operatorname{diag}(X_t)\right) dt + \sqrt{X_t}dW_t, \end{aligned}$$

además de que trivialmente $Z_0 = z_0$. Ahora, es inmediato el que

$$d\left[Z^i, Z^j\right]_t = \sum_k \sqrt{X_t}^{ik}\sqrt{X_t}^{jk} dt = X_t^{ij} dt,$$

de forma que X efectivamente captura la información de la volatilidad en el sentido de que modela la covariación cuadrática del proceso de logprecios o rendimientos. Alternativamente, podemos considerar la primer EDE como

$$dZ_t = \left(\mu Z_t + \nu - \frac{1}{2}\operatorname{diag}(X_t)\right) dt + \sqrt{X_t}\left(\rho \operatorname{vec}(dB_t) + \sqrt{I_d - \rho\rho'}dV_t\right),$$

donde V es un movimiento browniano en $\mathbb{R}^d$ independiente de B. Es claro también, del argumento de McKean (proposición 2.65) que Z no explota en tiempo finito y por lo tanto existe en todo $\mathbb{R}_+$. También podemos calcular

$$\begin{aligned} d\left[X^{ij}, Z^k\right]_t &= \sum_{lrs}\left(\sqrt{X_t}^{kl} dW_t^l\right)\left(\sqrt{X_t}^{ri} dB_t^{rs} Q^{sj} + Q^{si} dB_t^{rs}\sqrt{X_t}^{rj}\right) \\ &= \sum_{lrs}\sqrt{X_t}^{kl}\rho_{l,\upsilon(r,s)}\left(\sqrt{X_t}^{ri} Q^{sj} + Q^{si}\sqrt{X_t}^{rj}\right) dt \\ &= \sum_{lrs}\sqrt{X_t}^{kl}\rho_{l,\upsilon(r,s)}\left(\left(Q\otimes\sqrt{X_t}\right)_{\upsilon(r,s),\upsilon(i,j)} + \left(Q\otimes\sqrt{X_t}\right)_{\upsilon(r,s),\upsilon(j,i)}\right) dt \\ &= \left(\left(\sqrt{X_t}\rho\left(Q\otimes\sqrt{X_t}\right)\right)_{k,\upsilon(i,j)} + \left(\sqrt{X_t}\rho\left(Q\otimes\sqrt{X_t}\right)\right)_{k,\upsilon(j,i)}\right) dt. \end{aligned}$$

De hecho, la expresión anterior se puede simplificar si ρ tiene una estructura más simple. Nuestro ejemplo de mayor interés es cuando los efectos de correlación entre brownianos vienen dados por dos fuentes.

Primero, cada entrada de W por columna de B aportan a la estructura y entre dos columnas de B el efecto es en esencia el mismo, i.e., sólo cambia de acuerdo a un factor constante. En segundo lugar, hay otro efecto que relaciona a todas las entradas de B con cada una de W y esta estructura sólo varía de acuerdo a una constante de una entrada de W a otra. Esto se resume en pensar que $\rho = \rho^{1,1} \otimes \rho^{1,2} + \rho^{2,1} \otimes \mathrm{vec}\left(\rho^{2,2}\right)$ para algunos $\rho^{1,1}, \left(\rho^{2,1}\right)' \in M_{1,d}$, $\rho^{1,2} \in M_{d,d}$ y $\rho^{2,2} \in M_{d.d}$. En este caso, la expresión anterior se simplifica a

$$
\begin{aligned}
d\left[X^{ij}, Z^{k}\right]_t &= \sum_{lrs} \sqrt{X_t}^{kl} \left(\rho_s^{1,1}\rho_{l,r}^{1,2} + \rho_l^{2,1}\rho_{r,s}^{2,2}\right)\left(\sqrt{X_t}^{ri} Q^{sj} + Q^{si}\sqrt{X_t}^{rj}\right) dt \\
&= \left(\left(\sqrt{X_t}\rho^{1,2}\sqrt{X_t}\right)^{ki}\left(\rho^{1,1}Q\right)^{j} + \left(\sqrt{X_t}\rho^{1,2}\sqrt{X_t}\right)^{kj}\left(\rho^{1,1}Q\right)^{i}\right) dt \\
&+ \left(\rho^{2,1}\sqrt{X_t}\right)^{k}\left(\left(\sqrt{X_t}\rho^{2,2}Q\right)^{ij} + \left(\sqrt{X_t}\rho^{2,2}Q\right)^{ji}\right) dt.
\end{aligned}
$$

De hecho tenemos

$$
\begin{aligned}
\rho \mathrm{vec}\left(B\right) &= \left(\rho^{1,1} \otimes \rho^{1,2} + \rho^{2,1} \otimes \mathrm{vec}\left(\rho^{2,2}\right)\right) \mathrm{vec}\left(B\right) \\
&= \mathrm{vec}\left(\rho^{1,2} B \left(\rho^{1,1}\right)' + \mathrm{vec}\left(\rho^{2,2}\right) B \left(\rho^{2,1}\right)'\right),
\end{aligned}
$$

por lo que si $\rho = \rho^{1,1} \otimes I_d$ y $\mu = 0$, obtenemos justamente el modelo de Gourieroux y Sufana descrito en [23] (donde la primer condición va a ser necesaria para poder calcular la transformada de Fourier-Laplace).

Además, fácilmente podemos calcular para $S^i = \exp\left\{Z^i\right\}$,

$$
\begin{aligned}
\frac{dS_t^i}{S_t^i} &= \left(\sum_j \mu_{ij} Z_t^j + \nu_i\right) dt + \sum_j \sqrt{X_t}^{ij} dW_t^j, \\
d\left[S^i, S^j\right]_t &= \sum_k \sqrt{X_t}^{ik}\sqrt{X_t}^{jk} S_t^i S_t^j dt = S_t^i X_t^{ij} S_t^j dt.
\end{aligned}
$$

Así, podemos notar cómo el proceso de covariaciones cuadráticas de S es proporcional a cada activo y a su covolatilidad X^{ij}.

Con todo esto, podemos entender el papel que juega cada parámetro: μ establece una conexión de deriva entre los rendimientos de los activos, esto podría verse como una influencia direccional entre los logaritmos de los activos (comúnmente interpretado como sus rendimientos instantáneos pues $\Delta \log(x) = \frac{\Delta x}{x} + o(\Delta x))$ porque, de acuerdo al signo de μ_{ij}, si Z^j toma valores grandes, este influenciará a Z^i a tomar valores grandes también si $\mu_{ij} \geq 0$ o a tomar valores más pequeños si $\mu_{ij} < 0$. Similarmente ν determina la tendencia media de Z, tal como en el caso del modelo de Heston clásico (escalar).

Como en el caso escalar (en el modelo de Heston clásico) existe una transformada de Fourier para el proceso (Z, X) y nos permite hacer valuaciones exactas de opciones europeas, buscaremos extender este resultado al modelo Heston-Wishart. Antes de poder lograrlo, deberemos probar un pequeño lema técnico.

Lema 3.9. *Supongamos que A es una función analítica del tiempo con valores en $M_{n,n}$ de la forma*

$$A(t) = \sum_{n \geq 0} A_n t^n,$$

entonces la matriz fundamental Φ, solución a

$$\frac{\partial}{\partial t}\Phi(t) = A(t)\Phi(t), \quad \Phi(0) = I_n,$$

es analítica con coeficientes Φ_n que satisfacen $\Phi_0 = I_n$ y

$$n\Phi_n = \sum_{k=0}^{k-1} \Phi_k A_{n-k-1},$$

para cada $n > 0$.

Demostración. Basta probar que si tenemos coeficientes que satisfagan la recursión anterior, entonces $\Phi(t) := \sum_{n \geq 0} \Phi_n t^n$ es solución a la ecuación diferencial. Claramente tenemos $\Phi(0) = I_n$, además

$$\begin{aligned}
\frac{\partial}{\partial t}\Phi(t) &= \sum_{n \geq 1} n\Phi_n t^{n-1} \\
&= \sum_{n \geq 1} \sum_{k=0}^{k-1} A_{n-k-1}\Phi_k t^{n-1} \\
&= \sum_{n \geq 1} \sum_{k=0}^{k-1} A_{n-k-1} t^{n-k-1} \Phi_k t^k \\
&= \left(\sum_{n \geq 0} A_n t^n\right)\left(\sum_{n \geq 0} \Phi_n t^n\right) \\
&= A(t)\Phi(t),
\end{aligned}$$

con lo que concluye la prueba. $\square$

Supongamos de momento que $\rho^{2,1} = 0$ y que $\rho^{1,2}$ es un múltiplo escalar de la matriz identidad. Equivalentemente (permitiendo que $\rho^{1,1}$ absorba la constante), asumimos que $\rho = \rho^{1,1} \otimes I_d$. Similarmente a cuando conseguimos la transformada de Fourier-Laplace conjunta de $(X, X \cdot Id)$, para cada $w \in \mathbb{R}^d$ y $v \in S_d$, consideremos a la matriz

$$A_{v,w}(t) = \begin{pmatrix} -\beta - \frac{1}{2}Q'\left(\rho^{1,1}\right)' w' e^{t\mu} & -2\alpha \\ -\frac{1}{2}\mathrm{diag}^{-1}\left(e^{t\mu'}w\right) + \frac{1}{2}e^{t\mu'}ww'e^{t\mu} + v & \beta' + \frac{1}{2}e^{t\mu'}w\rho^{1,1}Q \end{pmatrix}$$

$$= \begin{pmatrix} -\beta & -2\alpha \\ v & \beta' \end{pmatrix} + \sum_{n\geq 0} t^n \frac{1}{n!} \begin{pmatrix} -\frac{1}{2}Q'\left(\rho^{1,1}\right)' w'\mu^n & 0 \\ -\frac{1}{2}\mathrm{diag}^{-1}\left(\left(\mu^n\right)' w\right) + \frac{1}{2}\sum_{k=0}^{n}\binom{n}{k}\mu^k ww'\mu^{n-k} & \frac{1}{2}\mu^n w\rho^{1,1}Q \end{pmatrix},$$

y sea $\Phi_{v,w}$ la matriz fundamental dada por el lema anterior. Fijemos $u \in S_d$ y consideremos a la función suave $P_{u,v,w} = \left(S'_{u,v,w}, T'_{u,v,w}\right)'$ con valores en $M_{2d,d}$, donde $S_{u,v,w}$ y $T_{u,v,w}$ toman valores en $M_{d,d}$, dada por

$$P_{u,v,w} = \Phi_{v,w}\begin{pmatrix} I_d \\ u \end{pmatrix}.$$

Se sigue que $S_{u,v}(0) = I_d$ y $T_{u,v}(0) = u$ y por lo tanto podemos definir a la función suave

$$\psi_{u,v,w}(t) = T_{u,v,w}(t)\, S_{u,v,w}(t)^{-1}, \quad \psi_{u,v,w}(0) = u,$$

al menos en $[0, T)$, para un tiempo fijo $T > 0$ en el que $S_{u,v,w}$ se vuelve singular (no invertible). Para relajar la notación, quitaremos los subíndices de P, S y T a menos que la dependencia en u, v o w se desee enfatizar. Luego, como

$$\frac{\partial}{\partial t}P = \frac{\partial}{\partial t}\Phi_{v,w}\begin{pmatrix} I_d \\ u \end{pmatrix} = A_{v,w}\Phi_{v,w}\begin{pmatrix} I_d \\ u \end{pmatrix} = A_{v,w}P,$$

deducimos que

$$\begin{aligned} \frac{\partial}{\partial t}S &= \left(-\beta - \frac{1}{2}Q'\left(\rho^{1,1}\right)' w'e^{t\mu}\right)S - 2\alpha T \\ \frac{\partial}{\partial t}T &= \left(-\frac{1}{2}\mathrm{diag}^{-1}\left(e^{t\mu'}w\right) + \frac{1}{2}e^{t\mu'}ww'e^{t\mu} + v\right)S + \left(\beta' + \frac{1}{2}e^{t\mu'}w\rho^{1,1}Q\right)T. \end{aligned}$$

Así, podemos deducir que

$$\begin{aligned}
\frac{\partial}{\partial t}\psi_{u,v,w} &= \left(\left(-\frac{1}{2}\mathrm{diag}^{-1}\left(e^{t\mu'}w\right)+\frac{1}{2}e^{t\mu'}ww'e^{t\mu}+v\right)S+\left(\beta'+\frac{1}{2}e^{t\mu'}w\rho^{1,1}Q\right)T\right)S^{-1}\\
&- TS^{-1}\left(\left(-\beta-\frac{1}{2}Q'\left(\rho^{1,1}\right)'w'e^{t\mu}\right)S-2\alpha T\right)S^{-1}\\
&= -\frac{1}{2}\mathrm{diag}^{-1}\left(e^{t\mu'}w\right)+\frac{1}{2}e^{t\mu'}ww'e^{t\mu}+v+\left(\beta'+\frac{1}{2}e^{t\mu'}w\rho^{1,1}Q\right)\psi_{u,v,w}\\
&+ \psi_{u,v,w}\left(\beta+\frac{1}{2}Q'\left(\rho^{1,1}\right)'w'e^{t\mu}\right)+2\psi_{u,v,w}\alpha\psi_{u,v,w}.
\end{aligned}$$

De manera completamente análoga al teorema 2.46, es fácil verificar que $\psi_{u,v,w}$ es simétrica y positiva definida al menos hasta un tiempo fijo $\sigma > 0$. Definamos ahora a las funciones ϑ_w y $\phi_{u,v,w}$ con valores en $\mathbb{R}^d$ y $\mathbb{R}$ respectivamente, dadas por

$$\begin{aligned}
\vartheta_w(t) &= e^{t\mu'}w\\
\phi_{u,v,w}(t) &= \int_0^t \left(2p\mathrm{tr}\left(\psi_{u,v,w}(s)\,\alpha\right)+w'e^{s\mu}\nu\right)ds.
\end{aligned}$$

Entonces claramente tenemos que estas funciones satisfacen el sistema de ecuaciones tipo Riccati dadas por

$$\begin{array}{rclrcl}
\frac{\partial}{\partial t}\vartheta_w &=& \mu'\vartheta_w & \vartheta_w(0) &=& w\\
\frac{\partial}{\partial t}\phi_{u,v,w} &=& 2p\mathrm{tr}\left(\psi_{u,v,w}\alpha\right)+\vartheta_w'\nu & \phi_{u,v,w}(0) &=& 0\\
\frac{\partial}{\partial s}\psi_{u,v,w} &=& \frac{1}{2}\vartheta_w\vartheta_w'-\frac{1}{2}\mathrm{diag}^{-1}\left(\vartheta_w\right)+\psi_{u,v,w}\left(\beta+\frac{1}{2}Q'\left(\rho^{1,1}\right)'\vartheta_w'\right) & \psi_{u,v,w}(0) &=& u.\\
&+& \left(\beta'+\frac{1}{2}\vartheta_w\rho^{1,1}Q\right)\psi_{u,v,w}+2\psi_{u,v,w}\alpha\psi_{u,v,w}+v & & &
\end{array} \tag{3.6}$$

Se sigue directamente que $\frac{\partial}{\partial t}\vartheta_w' z = \vartheta_w'\mu z$ y además podemos escribir

$$\begin{aligned}
&\mathrm{tr}\left(\vartheta_w\rho^{1,1}Q\psi_{u,v,w}x\right)\\
&= \rho^{1,1}Q\psi_{u,v,w}\sqrt{x}\left(\rho^{1,2}\right)'\sqrt{x}\vartheta_w\\
&= \sum_{lrs}\left(Q\psi_{u,v,w}\sqrt{x}\right)^{sr}\left(\rho_s^{1,1}\rho_{l,r}^{1,2}\right)\left(\sqrt{x}\vartheta_w\right)^l\\
&= \sum_{lrs}\left(\sqrt{x}\vartheta_w\right)^l\rho_{l,v(r,s)}\left(\sqrt{x}\psi_{u,v,w}Q'\right)^{rs},
\end{aligned}$$

entonces también tenemos para todo $x \in \bar{S}_d^+$, por la invariancia cíclica de la traza y su invariancia ante transposiciones,

$$\begin{aligned}
\mathrm{tr}\left(\frac{\partial}{\partial t}\psi_{u,v,w}x\right) &= \frac{1}{2}\vartheta_w'x\vartheta_w+\mathrm{tr}\left(vx\right)+2\mathrm{tr}\left(\psi_{u,v,w}\left(\alpha\psi_{u,v,w}+\beta\right)x\right)\\
&+ \vartheta_w'\sqrt{x}\rho\left(Q\otimes\sqrt{x}\right)\mathrm{vec}\left(\psi_{u,v,w}\right)-\frac{1}{2}\vartheta_w'\mathrm{diag}\left(x\right).
\end{aligned}$$

Luego, sumando las tres igualdades podemos deducir que para $t \in (0,\sigma)$ fijo,

$$\begin{aligned} 0 &= \frac{\partial}{\partial s}\phi_{u,v,w}(t-s) + \mathrm{tr}\left(\frac{\partial}{\partial s}\psi_{u,v,w}(t-s)\,x\right) \\ &+ \frac{\partial}{\partial s}\vartheta_w(t-s)'\,z + \frac{1}{2}\vartheta_w(t-s)'\,x\vartheta_w(t-s) \\ &+ 2\mathrm{tr}\left(\psi_{u,v,w}(t-s)\left(\alpha\psi_{u,v,w}(t-s)+\beta\right)x\right) \\ &+ \vartheta_w(t-s)'\left(\sqrt{x}\rho\left(Q\otimes\sqrt{x}\right)\right)\mathrm{vec}\left(\psi_{u,v,w}(t-s)\right) \\ &+ 2p\mathrm{tr}\left(\psi_{u,v,w}(t-s)\,\alpha\right) \\ &+ \vartheta_w(t-s)'\left(\mu z+\nu-\frac{1}{2}\mathrm{diag}(x)\right)+\mathrm{tr}(vx)\,. \end{aligned}$$

Equivalentemente, lo podemos escribir como

$$\begin{aligned} 0 &= \frac{\partial}{\partial s}\phi_{u,v,w}(t-s) + \mathrm{tr}\left(\frac{\partial}{\partial s}\psi_{u,v,w}(t-s)\,x\right) + \frac{\partial}{\partial s}\vartheta_w(t-s)'\,z \\ &+ \frac{1}{2}\sum_{ij}\vartheta_w(t-s)_i\,\vartheta_w(t-s)_j\,x_{ij} + \frac{1}{2}\sum_{ijkl}\psi_{u,v,w}(t-s)_{ji}\,\psi_{u,v,w}(t-s)_{lk}\,\alpha_{ijkl}(x) \\ &+ \frac{1}{2}\sum_{ijk}\psi_{u,v,w}(t-s)_{ji}\,\vartheta_w(t-s)_k\left(\left(\sqrt{x}\rho\left(Q\otimes\sqrt{x}\right)\right)_{k,v(i,j)}+\left(\sqrt{x}\rho\left(Q\otimes\sqrt{x}\right)\right)_{k,v(j,i)}\right) \\ &+ \mathrm{tr}\left(\psi_{u,v,w}(t-s)\left(2p\alpha+\beta x+x\beta'\right)\right)+\sum_i\vartheta_w(t-s)_i\left((\mu z)_i+\nu_i-\frac{1}{2}x_{ii}\right)+\mathrm{tr}(vx)\,. \end{aligned}$$

De aquí se sigue directamente que si tomamos a la semimartingala $J_s^t = f\left(s, X_s, \int_0^s X_r dr, Z_s\right)$ donde

$$f(s,x,y,z) = \exp\left\{\phi_{u,v,w}(t-s) + \langle\psi_{u,v,w}(t-s)\,,x\rangle + \langle v,y\rangle + \vartheta_w(t-s)'\,z\right\},$$

entonces, puesto que todos los términos de variación finita en compactos desaparecen al usar la fórmula de Itô, tenemos que J^t es una martingala local no negativa y por lo tanto, una supermartingala.

Teorema 3.10. *Consideremos u, v hermitianas con $\Re u, \Re v \in -\bar{S}_d^+$ y $w \in i\mathbb{R}^d$ y supongamos que $\rho = \rho^{1,1}\otimes I_d$. Entonces para todo $t \geq 0$ tenemos*

$$E\exp\left\{\langle u, X_t\rangle + \left\langle v, \int_0^t X_s ds\right\rangle + w'Z_t\right\} = \exp\left\{\phi_{u,v,w}(t) + \langle\psi_{u,v,w}(t)\,,x_0\rangle + \vartheta_w(t)'\,z_0\right\}.$$

Además, para $s, t \geq 0$ se tienen las siguientes ecuaciones de semiflujo

$$\begin{aligned} \phi_{u,v,w}(t+s) &= \phi_{u,v,w}(s) + \phi_{\varsigma_{u,v}(s),v,\vartheta_w(s)}(t) \\ \psi_{u,v,w}(t+s) &= \psi_{\psi_{u,v,w}(s),v,\vartheta_w(s)}(t) \\ \vartheta_w(t+s) &= \vartheta_{\vartheta_w(s)}(t)\,. \end{aligned} \tag{3.7}$$

De hecho, la fórmula es válida para toda elección de matrices hermitianas u, v y $w \in \mathbb{C}^d$ hasta antes de la explosión de $(\phi_{u,v,w}, \psi_{u,v,w}, \vartheta_w)$.

Demostración. Para la fórmula general de la característica, observemos que la construcción de $(\phi_{u,v,w}, \psi_{u,v,w}, \vartheta_w)$ es válida para cualesquiera u, v hermitianas y $v \in \mathbb{C}^d$ y las conclusiones son las mismas. Como vimos, J_s^t es una martingala local compleja para $t \in (0, \sigma)$ con (recordemos que los términos de variación finita en compactos desaparecieron)

$$\frac{dJ_s^t}{J_s^t} = -\vartheta_w (t-s)' \sqrt{X_s} dW_s - \operatorname{tr}\left(\psi_{u,v,w} (t-s)\, dM_s\right), \quad 0 \le s \le t,$$

donde M es la parte martingala de X. Como $J^t = f(Id, X, Y, Z)$ es una exponencial cuyo exponente tiene parte real no positiva, entonces $|J^t| \le 1$, de donde podemos concluir que J^t es una verdadera martingala y concluir la primera afirmación (incluyendo la conclusión implícita de que $\sigma = \infty$) de la misma manera que en los teoremas 2.42 y 2.46. Para la segunda parte, nuevamente usemos la propiedad de Markov:

$$\begin{aligned}
& E \exp\left\{\langle u, X_{t+s}\rangle + \left\langle v, \int_0^{t+s} X_u du\right\rangle + w' Z_{t+s}\right\} \\
= \; & E\left(E\left(\exp\left\{\langle u, X_{t+s}\rangle + \left\langle v, \int_0^{t+s} X_u du\right\rangle + w' Z_{t+s}\right\}\middle| \mathcal{F}_t\right)\right) \\
= \; & E \exp\left\{\phi_{u,v,w}(s) + \langle \psi_{u,v,w}(s), X_t\rangle + \left\langle v, \int_0^t X_u du\right\rangle + \vartheta_w(s)' Z_t\right\} \\
& \exp\left\{\phi_{u,v,w}(s) + \phi_{\psi_{u,v,w}(s), v, \vartheta_w(s)}(t) + \left\langle \psi_{\psi_{u,v,w}(s), v, \vartheta_w(s)}(t), x_0\right\rangle + \vartheta_{\vartheta_w(s)}(t)' z_0\right\},
\end{aligned}$$

concluyendo la prueba de estas afirmaciones. La última afirmación una vez más se sigue de las propiedades analíticas de la transformada de Fourier-Laplace cuando esta existe en una vecindad del 0 y de que, para $u, v \in S_d$ y $w \in \mathbb{R}^d$, J^t es una martingala local no negativa y por lo tanto una supermartingala, lo que implica que

$$E \exp\left\{\langle u, X_t\rangle + \left\langle v, \int_0^t X_s ds\right\rangle + w' Z_t\right\} = E J_t^t \le E J_0^t < \infty,$$

hasta antes del tiempo de explosión (posiblemente infinito). $\square$

Un resultado de especial interés en finanzas es el teorema de Girsanov. Este teorema nos permite modificar la deriva del modelo con una medida equivalente $\mathbb{Q}$ hasta obtener una semimartingala que, al descontarse con la tasa libre de riesgo (i.e., multiplicando porexp $\{-RId\}$ para alguna R) nos deja una martingala local.

Por esta razón, a la medida $\mathbb{Q}$ se le conoce como *medida libre de riesgo* o *neutral al riesgo*. El caso general en el modelo de Heston-Wishart, por ser de interés, lo presentaremos a continuación.

Teorema 3.11. *(cambio de deriva para el modelo Heston-Wishart) Consideremos* $\tilde{\mu}, \tilde{\beta} \in M_{d,d}$, $\tilde{\nu} \in \mathbb{R}^d$, $p, \tilde{p} \geq \frac{d-1}{2}$ *y definamos* $A = \{p \neq \tilde{p}\} \cup \{\mu \neq \tilde{\mu}\} \cup \{\nu \neq \tilde{\nu}\}$. *Supongamos adicionalmente que* $1 \notin \sigma(\rho\rho')$ *y que, según sea el caso,*

(I) *si se tiene* A, *entonces* $p \wedge \tilde{p} \geq \frac{d+1}{2}$ *y* $x \in S_d^+$,

(II) *si* $\beta \neq \tilde{\beta}$, *entonces* Q *es invertible.*

Entonces existe una medida de probabilidad equivalente $\tilde{\mathbb{P}}$, *tal que su proceso de derivadas de Radon-Nikodym está dado por la martingala* ζ *donde*

$$\zeta_t = \mathcal{E}\left(-\int_0^t tr(\xi_s dB_s) - \int_0^t \eta_s' dV_s\right),$$

para los procesos

$$\begin{aligned}
\eta &= \sqrt{I_d - \rho\rho'}^{-1}\sqrt{X}^{-1}((\mu - \tilde{\mu})Z + \nu - \tilde{\nu}) \\
&- \sqrt{I_d - \rho\rho'}^{-1}\rho\left(\left(\left(Q^{-1}\right)' \otimes \sqrt{X}\right) vec\left(\beta' - \tilde{\beta}'\right) 1_{\{\beta \neq \tilde{\beta}\}} + (p - \tilde{p})\left(Q \otimes \sqrt{X}^{-1}\right) vec(I_d) 1_A\right) \\
\xi &= \sqrt{X}\left(\beta - \tilde{\beta}\right)' Q^{-1} 1_{\{\beta \neq \tilde{\beta}\}} + (p - \tilde{p})\sqrt{X}^{-1} Q' 1_A.
\end{aligned}$$

Además, $\tilde{V} = V + \eta \cdot Id$ *y* $\tilde{B} = B + \xi \cdot Id$ *son movimientos brownianos independientes bajo* $\tilde{\mathbb{P}}$ *para los cuáles se tiene*

$$\begin{aligned}
dZ_t &= \left(\tilde{\mu} Z_t + \tilde{\nu} - \frac{1}{2} diag(X_t)\right) dt + \sqrt{X_t}\left(\rho vec\left(d\tilde{B}_t\right) + \sqrt{I_d - \rho\rho'} d\tilde{V}_t\right) \\
dX_t &= \sqrt{X_t} d\tilde{B}_t Q + Q' d\tilde{B}_t' \sqrt{X_t} + \left(2\tilde{p}\alpha + \tilde{\beta} X_t + X_t \tilde{\beta}'\right) dt.
\end{aligned}$$

Demostración. La condición $1 \notin \sigma(\rho\rho')$ sólo es necesaria para que exista $\sqrt{I_d - \rho\rho'}^{-1}$. Observemos que con los procesos ξ, η, $\tilde{B}$ y $\tilde{V}$, efectivamente podemos modificar la deriva de X del mismo modo que en el teorema 2.86, mientras que para Z tenemos

$$\begin{aligned} dZ_t &= \left(\mu Z_t + \nu - \frac{1}{2}\operatorname{diag}(X_t)\right) dt \\ &+ \sqrt{X_t}\rho\left(\operatorname{vec}\left(d\tilde{B}_t - \sqrt{X_t}\left(\beta - \tilde{\beta}\right)' Q^{-1} 1_{\{\beta\neq\tilde{\beta}\}} dt + (p-\tilde{p})\sqrt{X_t}^{-1} Q' 1_A dt\right)\right) \\ &+ \sqrt{X_t}\sqrt{I_d - \rho\rho'}\left(d\tilde{V}_t + \sqrt{I_d-\rho\rho'}^{-1}\sqrt{X}^{-1}\left((\mu-\tilde{\mu})Z + \nu - \tilde{\nu}\right) dt\right. \\ &- \sqrt{I_d-\rho\rho'}^{-1}\rho\left(\left(\left(\left(Q^{-1}\right)'\otimes\sqrt{X}\right)\operatorname{vec}\left(\beta'-\tilde{\beta}'\right) 1_{\{\beta\neq\tilde{\beta}\}} + (p-\tilde{p})\left(Q\otimes\sqrt{X}^{-1}\right)\operatorname{vec}(I_d)\, 1_A\right) dt\right) \\ &= \left(\mu Z_t + \nu - \frac{1}{2}\operatorname{diag}(X_t)\right) dt + \sqrt{X_t}\rho\operatorname{vec}\left(d\tilde{B}_t\right) + \sqrt{X_t}\sqrt{I_d-\rho\rho'}d\tilde{V}_t - \left((\mu-\tilde{\mu})Z + \nu - \tilde{\nu}\right) dt \\ &= \left(\tilde{\mu} Z_t + \tilde{\nu} - \frac{1}{2}\operatorname{diag}(X_t)\right) dt + \sqrt{X_t}\left(\rho\operatorname{vec}\left(d\tilde{B}_t\right) + \sqrt{I_d-\rho\rho'}d\tilde{V}_t\right). \end{aligned}$$

Sólo resta ver que ζ es una verdadera martingala en cualquier compacto $[0,T]$ para $T>0$. El resto de la demostración es idéntica a la prueba que dimos en el teorema 2.86, donde se usa la polaridad de ∂S_d^+ cuando $p,\tilde{p}\geq\frac{d+1}{2}$ y $x\in S_d^+$, i.e., cuando es necesario invertir a $\sqrt{X_t}$ y la unicidad en ley de soluciones de Wishart para $p,\tilde{p}\geq\frac{d-1}{2}$; la única diferencia es que ahora es necesario también usar la existencia fuerte y unicidad trayectorial (y en particular en ley) de las soluciones de Z, así como su no explosión. Esta última característica es inmediata de la continuidad Lipschitz de la deriva y difusión de Z (ver proposición 20.6 de [10]), así como la no explosión de X. $\square$

Reiterando lo que mencionamos anteriormente, el caso de mayor interés es cuando los nuevos parámetros convierten a S en una martingala local al ser descontado con la tasa libre se riesgo. Si fijamos a la tasa libre de riesgo como $R>0$ y definamos al proceso de activos $\tilde{S}=e^{-RId}S$, entonces la dinámica estocástica que satisface $\tilde{S}$ es

$$\begin{aligned} d\tilde{S}_t^i &= e^{-Rt}dS_t^i - Re^{-Rt}S_t^i dt \\ &= \tilde{S}_t^i\left(\left(\sum_j \tilde{\mu}_{ij}Z_t^j + \tilde{\nu}_i\right)dt + \sum_j \sqrt{X_t}^{ij} d\tilde{W}_t^j\right) - R\tilde{S}_t^i dt \\ &= \tilde{S}_t^i\left(\left(\sum_j \tilde{\mu}_{ij}Z_t^j + \tilde{\nu}_i - R\right)dt + \sum_j \sqrt{X_t}^{ij} d\tilde{W}_t^j\right), \end{aligned}$$

donde $\tilde{W}=\rho\operatorname{vec}\left(\tilde{B}\right)+\sqrt{I_d-\rho\rho'}\tilde{V}$. Luego, la dinámica que queremos que satisfaga $\tilde{S}$ bajo la medida libre de riesgo $\mathbb{Q}$, es la de una martingala local, i.e., buscamos que bajo $\mathbb{Q}$ la parte de variación finita en compactos sea nula (usando la notación del teorema anterior):

$$\sum_j \tilde{\mu}_{ij}Z^j + \tilde{\nu}_i - R = 0, \quad i=1,\ldots,d,$$

cosa que sólo es posible si $\tilde{\mu} = 0$ y $\tilde{\nu} = R\mathbf{1}$, en cuyo caso $d\tilde{S}_t^i = \tilde{S}_t^i \left(\sqrt{X_t} d\tilde{W}_t\right)^i$. Aunque hay un problema con la unicidad de $\mathbb{Q}$ (debido a que no tenemos ninguna especificación sobre p y β), no nos concentraremos en eso y supondremos que en el cambio de medida estos dos parámetros no sufrieron cambios.

En este caso, buscaremos probar que S tiene momentos finitos y que por lo tanto $\tilde{S}$ es una verdadera martingala. Para esto, notemos que la dinámica estocástica de Z está dada por

$$dZ_t^i = \left(R - \frac{1}{2} X_t^{ii}\right) dt + \sum_j \sqrt{X_t}^{ij} d\tilde{W}_t^j,$$

De hecho, si definimos $\tilde{Z} = \log\left(\tilde{S}\right) = Z - RId$, entonces

$$d\tilde{Z}_t^i = -\frac{1}{2} X_t^{ii} dt + \sum_j \sqrt{X_t}^{ij} d\tilde{W}_t^j.$$

En este caso especial, es claro que la terna de funciones que usamos para describir la transformada de Fourier de (X, Z) se simplifica. De hecho, para construir la transformada de Fourier-Laplace de $\left(X, X \cdot Id, \tilde{Z}\right)$ (bajo la misma hipótesis sobre ρ), basta reemplazar en las fórmulas $\mu = 0$ y $\nu = 0$. De esta forma tenemos

$$A_{v,w}(t) = \begin{pmatrix} -\beta - \frac{1}{2} Q' \left(\rho^{1,1}\right)' w' & -2\alpha \\ -\frac{1}{2} \mathrm{diag}^{-1}(w) + \frac{1}{2} w w' + v & \beta' + \frac{1}{2} w \rho^{1,1} Q \end{pmatrix},$$

es decir que $A_{v,w}$ es contante y tiene una forma particularmente similar a la matriz $-A_v$ que presentamos en el (previo al) teorema 2.46. Como consecuencia, tenemos una forma explícita (no recursiva, a diferencia de lo que nos da el lema 3.9 y mucho más sencilla) de encontrar la forma que adquieren $\psi_{u,v,w}$, $\phi_{u,v,w}$ y ϑ_w en este caso particular.

Ahora, aunque sería deseable que $\tilde{S}$ fuera una verdadera martingala, es conocido (ver [**?**, 40, 23]) que el modelo de Heston presenta momentos finitos hasta un tiempo fijo (dependiente de sus parámetros y posiblemente infinito) y explota hacia infinito después de este tiempo. Este fenómeno también ocurre en nuestro contexto, como podemos verificar de la transformada de Fourier-Laplace de $\tilde{Z}$. Observemos que si buscamos probar que $E \exp\left\{\tilde{Z}_t^i\right\} = \exp\left\{z_0^i\right\}$ (esto es, $\tilde{S}^i$ es una martingala hasta el tiempo t), basta corroborar la condición de Novikov:

$$E \exp\left\{\frac{1}{2} \int_0^t X_s^{ii} ds\right\} < \infty.$$

Esta condición resulta moderadamente fácil de verificar a gracias a la proposición 2.44. Una forma sencilla de verificarlo es pedir que $\frac{1}{2}\sigma_t^{\beta}(\alpha)_{ii} < 1$, es decir,

$$\int_0^t \left(e^{\beta s}\alpha e^{\beta' s}\right)_{ii} ds < 1.$$

De esta forma, tal y como sucede en el caso escalar, para cada i podemos dar un tiempo fijo T_i (estrictamente positivo y posiblemente infinito) en términos de (α, β) hasta el cuál sabemos que la condición de Novikov se satisface y por lo tanto $\tilde{S}^i$ es una martingala; de hecho,

$$T_i = \inf\left\{t > 0 : \int_0^t \left(e^{\beta s}\alpha e^{\beta' s}\right)_{ii} ds \geq 1\right\}.$$

En particular, para tiempos menores a $T = \min_{i\leq d}\{T_i\}$, el proceso $\tilde{S}$ es una martingala y por lo tanto tenemos (la desigualdad de Hölder generalizada y la monotonía de los momentos)

$$E\prod_{i\leq d}\left(\tilde{S}_t^i\right)^{\lambda_i} < \infty,$$

si $\sum_i \lambda_i \leq 1$ y $\lambda_i \geq 0$ para cada $i \leq d$. Una consecuencia directa es que si $\alpha \in S_d^+$, $\Sigma = \int_0^\infty e^{\beta s}\alpha e^{\beta' s} ds$ existe y satisface $\Sigma_{ii} \leq 1$, entonces $\tilde{S}$ es una martingala en todo $\mathbb{R}_+$.

Otra forma de determinar si hay explosión en los momentos, es verificar que la solución a las ecuaciones tipo Riccati no explote y que evaluados en $u = v = 0$ y $w = e^i$, recuperemos $ES_t^i = e^{z_0^i}$. En este caso, podemos calcular la transformada de Laplace de $\left(X, \tilde{Z}^i\right)$, dada por la misma fórmula que en el teorema 2.46 para $v = 0$ (que corresponde a $A_{v,w}$ cuando $v = 0$ y $w = e^i$) pero en las entradas de la diagonal correspondería a tener el parámetro $\beta + \frac{1}{2}Q'\left(\rho^{1,1}\right)' w'$ en vez de β.

El único posible problema con el teorema 2.46 es que la submatriz correspondiente a $\frac{1}{2}\mathrm{diag}^{-1}(w) - \frac{1}{2}ww'$ era positiva semidefinida (lo que bastaba para probar la no explosión). Para encontrar los valores de v para los cuales esta propiedad es cierta, requeriremos de un pequeño lema.

Teorema 3.12. *(círculo de Gershgorin) Sea A una matriz compleja de $n \times n$. Sea $R_i = \sum_{j\neq i}|A_{ij}|$, entonces cada valor propio de A está en alguno de los discos*

$$\bar{B}(A_{ii}, R_i) = \{x \in \mathbb{C} : |A_{ii} - x| \leq R_i\}, \quad i = 1, \ldots, n.$$

Demostración. Sea λ un valor propio y x su vector propio no nulo asociado. Sea $i = \arg\max_j |x_j|$, entonces claramente $x_i \neq 0$ y además, en

$$\sum_{j\neq i} A_{ij}x_j = \lambda x_i - A_{ii}x_i,$$

podemos dividir entre x_i para obtener

$$|\lambda - A_{ii}| = \left| \sum_{j \neq i} A_{ij} \frac{x_j}{x_i} \right| \leq \sum_{j \neq i} |A_{ij}| \left| \frac{x_j}{x_i} \right| \leq \sum_{j \neq i} |A_{ij}| = R_i,$$

como queríamos demostrar. $\square$

Ahora observemos que

$$M_w := \frac{1}{2} \mathrm{diag}^{-1}(w) - \frac{1}{2} w w' = \frac{1}{2} \begin{pmatrix} w_1 - w_1^2 & -w_1 w_2 & \cdots \\ -w_1 w_2 & w_2 - w_2^2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix},$$

por lo tanto, usando el teorema del círculo de Gershgorin, obtenemos que los valores propios de M_w están en

$$\bigcup_{i \leq d} \bar{B} \left(\frac{1}{2} \left(w_i - w_i^2 \right), \frac{1}{2} |w_i| \sum_{j \neq i} |w_j| \right).$$

La simetría de M_w nos garantiza que los valores propios son reales. Así que para asegurarnos de que este conjunto esté contenido en $\mathbb{R}_+$ (y por lo tanto $M_w \in \bar{S}_d^+$), basta garantizar el que

$$w_i - w_i^2 - |w_i| \sum_{j \neq i} |w_j| \geq 0, \quad i = 1, \ldots, d.$$

Tomando $w_i \geq 0$ para todo i y dividiendo entre w_i cuando este sea no nulo (en otro caso, es decir, cuando $w_i = 0$ su desigualdad respectiva se satisface trivialmente), podemos ver que esto se reduce a conseguir

$$1 - w_i \geq \sum_{j \neq i} w_j, \quad i = 1, \ldots, d.$$

Esta condición es trivialmente satisfecha si $w \in \left[0, \frac{1}{d}\right]^d$ o si $w = e^i$ para alguna $i \leq d$. Luego, esto basta para poder replicar los resultados obtenidos en el teorema 2.46.

De hecho, si $w = e^i$, entonces $M_w = 0$ y si recordamos la forma que tenía A_v y que A_0 justo nos ayudaba a recuperar el caso de la transformada de Laplace marginal, el problema se reduce al caso tratado en el teorema 2.42, cuya solución es mucho más explícita. Lo que quiere decir esto es que $\psi_{0,0,e^i} = -\psi_0 = 0$, $\phi_{0,0,e^i} = -\phi_0 = 0$ y $\vartheta_w = w = e^i$. Con todo esto finalmente obtenemos

$$E^{\mathbb{Q}} e^{\left(e^i\right)' \tilde{Z}_t} = E^{\mathbb{Q}} \tilde{S}_t^i = e^{\left(e^i\right)' z_0} = e^{z_0^i}.$$

Justo hemos dado la prueba del siguiente resultado.

Corolario 3.13. *Si $\rho = \rho^{1,1} \otimes I_d$, el proceso de activos descontados $\tilde{S}$ siempre es una martingala respecto de $\mathbb{Q}$. De hecho, si $w \in \mathbb{R}^d$ satisface $sgn(w_i) - |w_i| \geq |w_i| \sum_{j \neq i} |w_j|$ para cada $i = 1, \ldots, d$, entonces para toda $t \geq 0$ se tiene*

$$E \prod_{i \leq d} \left(\tilde{S}_t^i\right)^{w_i} < \infty.$$

De manera más general, $Ee^{w' \tilde{Z}_t} < \infty$ para $w \in \mathbb{R}^d$ si

$$\sup_{t_0 \leq t} \sup \sigma \left(\sigma_{t_0}^{\beta + \frac{1}{2}\left(w \rho^{1,1} Q\right)'} (\alpha) \left(diag^{-1}(w) - ww' \right) \right) < 2,$$

en particular, $E\left(\tilde{S}_t^i\right)^r < \infty$ si

$$\left(\sigma_t^{\beta + \frac{1}{2} r \left(e^i \rho^{1,1} Q\right)'} (\alpha) \right)_{ii} < \frac{2}{r - r^2}.$$

Demostración. Lo único que falta probar son las últimas dos afirmaciones, pero esto se sigue de la misma discusión anterior, donde usamos la proposición 2.44 y el que $A_{v,w}$ es igual a $-A_v$ con la nueva matriz de deriva $\beta + \frac{1}{2} Q' \left(\rho^{1,1}\right)' w'$. □

Para saber si $\tilde{S}$ tiene momentos de orden mayor, en el caso en que no se satisface la condición del corolario anterior, hay que recurrir al teorema 3.10 y usar el que $\tilde{S}^i = \exp\left\{\tilde{Z}^i\right\}$.

3.4.1. Estimación

Para la estimación de parámetros, lo que haremos es seguir la misma metodología que cuando sólo contábamos con las observaciones del proceso de Wishart X. Primero, de acuerdo a las características de nuestra muestra, si no contamos con observaciones directas de X, podemos usar las técnicas de [41] u otras referencias, para estimar la variación cuadrática de Z (de manera vectorial, en nuestro caso). De aquí podemos tomar estimadores de X puesto que $\left[Z^i, Z^j\right] = X^{ij} \cdot Id$.

Ahora bien, para estimar a los parámetros (p, α, β) seguimos el procedimiento especificado en la sección anterior. Por otro lado, para estimar a (μ, ν), discreticemos la

EDE que satisface Z y usemos la misma notación que en la sección anterior:

$$\Delta Z_n = \left(\mu Z_n + \nu - \frac{1}{2}\mathrm{diag}\left(X_n\right)\right)\Delta n + \sqrt{X_n}\Delta W_n.$$

Para despejar a ΔW_n, supondremos que $2p \geq d+1$; si no es el caso, aún se puede proceder usando únicamente inversas generalizadas, que es la mejor manera de recuperar la información que tenemos de W_n y, como argumentamos antes, la minimización de cuadrados también coincide con la estimación máximo verosímil (dada la discretización) en este caso. Luego, con esta hipótesis podemos deducir que

$$X_n^{-\frac{1}{2}}\left(\frac{\Delta}{\Delta n}Z_n - \mu Z_n - \nu + \frac{1}{2}\mathrm{diag}\left(X_n\right)\right) = \frac{\Delta}{\Delta n}W_n,$$

de forma que debemos minimizar

$$\begin{aligned} g\left(\mu,\nu\right) &= \sum_n \left\| X_n^{-\frac{1}{2}}\left(\frac{\Delta}{\Delta n}Z_n - \mu Z_n - \nu + \frac{1}{2}\mathrm{diag}\left(X_n\right)\right)\right\|^2 \\ &= \sum_n\sum_i\left(\sum_j\left(X_n^{-\frac{1}{2}}\right)^{ij}\left(\frac{\Delta}{\Delta n}Z_n^j - \sum_k \mu_{jk}Z_n^k - \nu_j + \frac{1}{2}X_n^{jj}\right)\right)^2. \end{aligned}$$

Derivando parcialmente respecto de cada ν_l e igualando a cero obtenemos

$$\begin{aligned} 0 &= \frac{\partial}{\partial\nu_l}g\left(\mu,\nu\right) \\ &= -2\sum_n\sum_i\left(X_n^{-\frac{1}{2}}\right)^{il}\left(\sum_j\left(X_n^{-\frac{1}{2}}\right)^{ij}\left(\frac{\Delta}{\Delta n}Z_n^j - \sum_k\mu_{jk}Z_n^k - \nu_j + \frac{1}{2}X_n^{jj}\right)\right) \\ &= -2\sum_n\left(\left(X_n^{-1}\frac{\Delta}{\Delta n}Z_n\right)^l - \left(X_n^{-1}\mu Z_n\right)^l - \left(X_n^{-1}\nu\right)^l + \frac{1}{2}\left(X_n^{-1}\mathrm{diag}\left(X_n\right)\right)^l\right). \end{aligned}$$

Usando la representación vectorial, esto se simplifica a tener

$$\left(\sum_n X_n^{-1}\right)\nu = \sum_n X_n^{-1}\left(\frac{\Delta}{\Delta n}Z_n + \frac{1}{2}\mathrm{diag}\left(X_n\right) - \mu Z_n\right),$$

donde $X_n^{-1} \in S_d^+$ (o $\bar{S}_d^+$ si $2p < d+1$) y por lo tanto $\sum_n X_n^{-1} \in S_d^+$ (probablemente también es cierto si $2p < d-1$, dependiendo de β, α y del tamaño de la muestra). Consecuentemente, g se minimiza si la evaluamos en $\nu\left(\mu\right)$ dada por

$$\begin{aligned} \nu\left(\mu\right) &= \left(\sum_n X_n^{-1}\right)^{-1}\sum_n X_n^{-1}\left(\frac{\Delta}{\Delta n}Z_n + \frac{1}{2}\mathrm{diag}\left(X_n\right) - \mu Z_n\right) \\ &= \left(\sum_n X_n^{-1}\right)^{-1}\sum_n X_n^{-1}\left(\frac{\Delta}{\Delta n}Z_n + \frac{1}{2}\mathrm{diag}\left(X_n\right)\right) - \left(\sum_n X_n^{-1}\right)^{-1}\left(\sum_n Z_n' \otimes X_n^{-1}\right)\mathrm{vec}\left(\mu\right). \end{aligned}$$

Similarmente, derivamos a g parcialmente respecto de μ_{rs}, igualamos a cero y sustituimos a $\nu(\mu)$ de la siguiente forma

$$\begin{aligned}
0 &= \frac{\partial}{\partial \mu_{rs}} g(\mu,\nu) \\
&= -2\sum_n \sum_i \left(X_n^{-\frac{1}{2}}\right)^{ir} Z_n^s \left(\sum_j \left(X_n^{-\frac{1}{2}}\right)^{ij} \left(\frac{\Delta}{\Delta n} Z_n^j - \sum_k \mu_{jk} Z_n^k - \nu_j + \frac{1}{2} X_n^{jj}\right)\right) \\
&= -2\sum_n \left(\left(X_n^{-1}\frac{\Delta}{\Delta n} Z_n\right)^r - \left(X_n^{-1}\mu Z_n\right)^r - \left(X_n^{-1}\nu\right)^r + \frac{1}{2}\left(X_n^{-1}\mathrm{diag}(X_n)\right)^r\right) Z_n^s.
\end{aligned}$$

Si usamos su representación matricial, vectorizamos y sustituimos $\nu(\mu)$, se obtiene que

$$\sum_n \left(Z_n \otimes X_n^{-1}\right)\frac{\Delta}{\Delta n} Z_n + \frac{1}{2}\sum_n \left(Z_n \otimes X_n^{-1}\right)\mathrm{diag}(X_n)$$

es igual a

$$\begin{aligned}
& \left(\sum_n Z_n \otimes X_n^{-1}\right)\nu(\mu) + \left(\sum_n Z_n Z_n' \otimes X_n^{-1}\right)\mathrm{vec}(\mu) \\
= & \left(\sum_n Z_n \otimes X_n^{-1}\right)\left(\sum_n X_n^{-1}\right)^{-1} \sum_n X_n^{-1}\left(\frac{\Delta}{\Delta n} Z_n + \frac{1}{2}\mathrm{diag}(X_n)\right) \\
+ & \left(\left(\sum_n Z_n Z_n' \otimes X_n^{-1}\right) - \left(\sum_n Z_n \otimes X_n^{-1}\right)\left(\sum_n X_n^{-1}\right)^{-1}\left(\sum_n Z_n' \otimes X_n^{-1}\right)\right)\mathrm{vec}(\mu).
\end{aligned}$$

De aquí, debemos suponer una vez más que la muestra es lo suficientemente grande como para que la matriz que multiplica a $\mathrm{vec}(\mu)$ es no singular y por lo tanto podamos recuperar al estimador

$$\begin{aligned}
\hat{\mu} &= \left(\left(\sum_n Z_n Z_n' \otimes X_n^{-1}\right) - \left(\sum_n Z_n \otimes X_n^{-1}\right)\left(\sum_n X_n^{-1}\right)^{-1}\left(\sum_n Z_n' \otimes X_n^{-1}\right)\right)^{-1} \\
&\times \left[\sum_n \left(Z_n \otimes X_n^{-1}\right)\frac{\Delta}{\Delta n} Z_n + \frac{1}{2}\sum_n \left(Z_n \otimes X_n^{-1}\right)\mathrm{diag}(X_n)\right. \\
&- \left.\left(\sum_n Z_n \otimes X_n^{-1}\right)\left(\sum_n X_n^{-1}\right)^{-1} \sum_n X_n^{-1}\left(\frac{\Delta}{\Delta n} Z_n + \frac{1}{2}\mathrm{diag}(X_n)\right)\right],
\end{aligned}$$

y $\hat{\nu} = \nu(\hat{\mu})$. Finalmente, para estimar a ρ, podemos usar una discretización de

$$\left[X^{ij}, Z^k\right] = \left(\left(\sqrt{X}\rho\left(Q \otimes \sqrt{X}\right)\right)_{k,\upsilon(i,j)} + \left(\sqrt{X}\rho\left(Q \otimes \sqrt{X}\right)\right)_{k,\upsilon(j,i)}\right) \cdot Id,$$

y nuevamente minimizar cuadrados en el error, esto es, minimizar:

$$h(\rho) = \sum_{ijkn} \left(\frac{\Delta}{\Delta n} X_n^{ij} \frac{\Delta}{\Delta n} Z_n^k - \sum_{lrs} \sqrt{X_n}^{kl} \rho_{l,\upsilon(r,s)} \left(\hat{Q}_{sj} \sqrt{X_n}^{ri} + \hat{Q}_{si} \sqrt{X_n}^{rj} \right) \right)^2 .$$

Para esto, nuevamente derivamos parcialmente respecto de cada $\rho_{t,\upsilon(u,v)}$ para obtener

$$\begin{aligned}
0 &= \frac{\partial}{\partial \rho_{t,\upsilon(u,v)}} h(\rho) \\
&= 2 \sum_{ijkn} \sqrt{X_n}^{kt} \left(\hat{Q}_{vj} \sqrt{X_n}^{ui} + \hat{Q}_{vi} \sqrt{X_n}^{uj} \right) \\
&\times \left(\frac{\Delta}{\Delta n} X_n^{ij} \frac{\Delta}{\Delta n} Z_n^k - \sum_{lrs} \sqrt{X_n}^{kl} \rho_{l,\upsilon(r,s)} \left(\hat{Q}_{sj} \sqrt{X_n}^{ri} + \hat{Q}_{si} \sqrt{X_n}^{rj} \right) \right) \\
&= 4 \sum_n \left(\frac{\Delta}{\Delta n} Z_n \sqrt{X_n} \right)^t \left(\sqrt{X_n} \frac{\Delta}{\Delta n} X_n \hat{Q} \right)^{uv} \\
&- 2 \sum_{ijlrsn} X_n^{tl} \rho_{l,\upsilon(r,s)} \left(\hat{Q}_{vj} \sqrt{X_n}^{ui} + \hat{Q}_{vi} \sqrt{X_n}^{uj} \right) \left(\hat{Q}_{sj} \sqrt{X_n}^{ri} + \hat{Q}_{si} \sqrt{X_n}^{rj} \right) .
\end{aligned}$$

Desarrollando,

$$\begin{aligned}
0 &= 4 \sum_n \left(\frac{\Delta}{\Delta n} Z_n \sqrt{X_n} \right)^t \left(\sqrt{X_n} \frac{\Delta}{\Delta n} X_n \hat{Q} \right)^{uv} \\
&- 4 \sum_{lrsn} (X_n \rho)^{t,\upsilon(r,s)} \left(\left(\hat{Q}^2 \right)_{vs} X_n^{ur} + \left(\sqrt{X_n} \hat{Q} \right)_{rv} \left(\hat{Q} \sqrt{X_n} \right)_{su} \right) \\
&= 4 \sum_n \left(\frac{\Delta}{\Delta n} Z_n \sqrt{X_n} \otimes \operatorname{vec} \left(\sqrt{X_n} \frac{\Delta}{\Delta n} X_n \hat{Q} \right)' \right)^{t,\upsilon(u,v)} \\
&- 4 \sum_{rsn} (X_n \rho)^{t,\upsilon(r,s)} \left(\left(\hat{Q}^2 \otimes X_n \right)_{\upsilon(r,s),\upsilon(u,v)} + \left(\hat{Q} \sqrt{X_n} \otimes \hat{Q} \sqrt{X_n} \right)_{\upsilon(r,s),\upsilon(u,v)} \right) \\
&= 4 \sum_n \left(\left(\frac{\Delta}{\Delta n} Z_n \sqrt{X_n} \otimes \operatorname{vec} \left(\sqrt{X_n} \frac{\Delta}{\Delta n} X_n \hat{Q} \right)' \right) - \left(X_n \rho \left(\hat{Q}^2 \otimes X_n + \left(\hat{Q} \sqrt{X_n} \right) \otimes \left(\hat{Q} \sqrt{X_n} \right) \right) \right) \right)^{t,\upsilon(u,v)} ,
\end{aligned}$$

donde usamos repetidamente que $\hat{Q}$ es simétrica. En su representación matricial (usando una extensión natural de vec que asocia a cada elemento en M_{d,d^2}, uno en $\mathbb{R}^{d^3}$) tenemos

$$\left(\sum_n X_n \otimes \left(\hat{Q}^2 \otimes X_n + \left(\hat{Q} \sqrt{X_n} \right) \otimes \left(\hat{Q} \sqrt{X_n} \right) \right) \right) \operatorname{vec}(\rho) = \operatorname{vec} \left(\sum_n \frac{\Delta}{\Delta n} Z_n \sqrt{X_n} \otimes \operatorname{vec} \left(\sqrt{X_n} \frac{\Delta}{\Delta n} X_n \hat{Q} \right)' \right) ,$$

de donde, asumiendo que la matriz del lado izquierdo (en M_{d^3,d^3}) es invertible, podemos despejar a nuestro estimador

$$\operatorname{vec}(\hat{\rho}) = \left(\sum_n X_n \otimes \left(\hat{Q}^2 \otimes X_n + \left(\hat{Q} \sqrt{X_n} \right) \otimes \left(\hat{Q} \sqrt{X_n} \right) \right) \right)^{-1} \operatorname{vec} \left(\sum_n \frac{\Delta}{\Delta n} Z_n \sqrt{X_n} \otimes \operatorname{vec} \left(\sqrt{X_n} \frac{\Delta}{\Delta n} X_n \hat{Q} \right)' \right) .$$

Bibliografía

[1] O. Pfaffel, "Wishart processes," *arXiv preprint arXiv:1201.3256*, 2012.

[2] P. Graczyk and J. Małecki, "Strong solutions of non-colliding particle systems," *Electron. J. Probab.*, vol. 19, pp. no. 119, 1–21, 2014.

[3] G. Letac and H. Massam, "The noncentral wishart as an exponential family, and its moments," *Journal of Multivariate Analysis*, vol. 99, no. 7, pp. 1393 – 1417, 2008. Special Issue: Multivariate Distributions, Inference and Applications in Memory of Norman L. Johnson.

[4] E. Mayerhofer, "Wishart processes and wishart distributions: an affine processes point of view," *arXiv preprint arXiv:1201.6634*, 2012.

[5] M.-F. Bru, "Wishart processes," *J. Theoret. Probab.*, vol. 4, no. 4, pp. 725–751, 1991.

[6] P. Graczyk and J. Małecki, "Wallach sets and squared bessel particle systems," *arXiv preprint arXiv:1602.01597v1*, p. 5, 2016.

[7] P. Graczyk and L. Vostrikova, "The moments of Wishart processes via Itô calculus," *Teor. Veroyatn. Primen.*, vol. 51, no. 4, pp. 732–751, 2006.

[8] E. Mayerhofer, "On the existence of non-central Wishart distributions," *J. Multivariate Anal.*, vol. 114, pp. 448–456, 2013.

[9] P. Graczyk and J. Małecki, "Multidimensional yamada-watanabe theorem and its applications to particle systems," *Journal of Mathematical Physics*, vol. 54, no. 2, 2013.

[10] O. Kallenberg, *Foundations of modern probability.* Probability and its Applications (New York), Springer-Verlag, New York, second ed., 2002.

[11] R. Muirhead, *Aspects of Multivariate Statistical Theory.* Wiley Series in Probability and Statistics, Wiley, 1982.

[12] J. Faraut and A. Korányi, *Analysis on symmetric cones.* Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1994. Oxford

Science Publications.

[13] T. Kato, *Perturbation Theory for Linear Operators.* Classics in Mathematics, Springer Berlin Heidelberg, 1995.

[14] K. B. Petersen and M. S. Pedersen, “The matrix cookbook,” nov 2012. Version 20121115.

[15] J. R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics.* Chichester, New York, Weinheim: J. Wiley & Sons, 1999.

[16] V. Rakočević, “On continuity of the moore-penrose and drazin inverses.,” *Matematichki Vesnik*, vol. 49, no. 3-4, pp. 163–172, 1997.

[17] C. Dellacherie and P. Meyer, *A Probabilities and Potential.* North-Holland Mathematics Studies, Elsevier Science, 1979.

[18] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes, and Martingales*, vol. 1. Cambridge University Press, second ed., 2000. Cambridge Books Online.

[19] M. Coste, M., *An Introduction to O-minimal Geometry.* Dottorato di ricerca in matematica / Università di Pisa, Dipartimento di Matematica, Istituti Editoriali e Poligrafici Internazionali, 2000.

[20] R. J. Stelzer, *Multivariate Continuous time stochastic volatility models driven by a Lévy process.* Dissertation, Technische Universität München, MÃŒnchen, 2007.

[21] D. Duffie, D. Filipović, and W. Schachermayer, “Affine processes and applications in finance,” *Ann. Appl. Probab.*, vol. 13, pp. 984–1053, 08 2003.

[22] C. Cuchiero, D. Filipović, E. Mayerhofer, and J. Teichmann, “Affine processes on positive semidefinite matrices,” *Ann. Appl. Probab.*, vol. 21, pp. 397–463, 04 2011.

[23] A. Alfonsi, *Affine Diffusions and Related Processes: Simulation, Theory and Applications.* Bocconi & Springer Series, Springer International Publishing, 2015.

[24] E. Mayerhofer, “Positive semidefinite affine processes have jumps of finite variation,” Apr 2011. Comments: 8 pages.

[25] M. Keller-Ressel, W. Schachermayer, and J. Teichmann, “Affine processes are regular,” *Probability Theory and Related Fields*, vol. 151, no. 3, pp. 591–611, 2011.

[26] C. Kristopher Garrett and R.-C. Li, “Gip integrators for matrix riccati differential equations,” *Appl. Math. Comput.*, vol. 241, pp. 283–297, Aug. 2014.

[27] B. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction.*

Graduate Texts in Mathematics, Springer International Publishing, 2015.

[28] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion.* Grundlehren der mathematischen Wissenschaften, Springer Berlin Heidelberg, 2004.

[29] J.-F. Le Gall, "Applications du temps local aux équations différentielles stochastiques unidimensionnelles," *Séminaire de probabilités de Strasbourg*, vol. 17, pp. 15–31, 1983.

[30] A. Takemura, *Chapter 5. Complex Zonal Polynomials*, vol. Volume 4 of *Lecture Notes–Monograph Series*, pp. 83–97. Hayward, CA: Institute of Mathematical Statistics, 1984.

[31] S. G. Walker and R. H. Mena, "On a construction of markov models in continuous time," *Metron*, vol. 67, no. 3, pp. 303–323, 2009.

[32] G. H. Golub and C. F. Van Loan, *Matrix computations.* Johns Hopkins studies in the mathematical sciences, Baltimore, London: The Johns Hopkins University Press, 1996.

[33] S. M. Iacus, *Simulation and Inference for Stochastic Differential Equations: With R Examples (Springer Series in Statistics).* Springer Publishing Company, Incorporated, 1 ed., 2008.

[34] R. S. Christian Gourieroux, "Derivative pricing with wishart multivariate stochastic volatility," *Journal of Business & Economic Statistics*, vol. 28, no. 3, pp. 438–451, 2010.

[35] C. Gourieroux and R. Sufana, "Derivative pricing with multivariate stochastic volatility: Application to credit risk," Working Papers 2004-31, Centre de Recherche en Economie et Statistique, 2004.

[36] A. Gnoatto, "The wishart short rate model," *International Journal of Theoretical and Applied Finance (IJTAF)*, vol. 15, no. 08, 2012.

[37] A. Ahdida, A. Alfonsi, and E. Palidda, "Smile with the gaussian term structure model," papers, arXiv.org, 2015.

[38] J. Da Fonseca, M. Grasselli, and C. Tebaldi, "A multifactor volatility heston model," *Quantitative Finance*, vol. 8, no. 6, pp. 591–604, 2008.

[39] A. Benabid, H. Bensusan, and N. El Karoui, "Wishart Stochastic Volatility: Asymptotic Smile and Numerical Framework." 48 pages, June 2008.

[40] L. B. G. Andersen and V. V. Piterbarg, "Moment explosions in stochastic volatility models," *Finance and Stochastics*, vol. 11, no. 1, pp. 29–50, 2007.

[41] Y. Zu and H. Peter Boswijk, "Estimating spot volatility with high-frequency financial data," *Journal of Econometrics*, vol. 181, no. 2, pp. 117–135, 2014.