



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
MAESTRÍA EN CIENCIAS DE LA TIERRA
CAMPO I. GEOFÍSICA DE LA TIERRA SÓLIDA

CUANTIFICACIÓN DE LA INCERTIDUMBRE EN LA INVERSIÓN DEL DESLIZAMIENTO SISMICO

TESIS

QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS DE LA TIERRA

PRESENTA:

ISAAC RUBÉN VALVERDE GUERRERO

DR. JOSUÉ TAGO PACHECO
FACULTAD DE INGENIERIA

DRA. URSULA ITURRARÁN VIVEROS
FACULTAD DE CIENCIAS

DR. VICTOR MANUEL CRUZ ATIENZA
INSTITUTO DE GEOFÍSICA

CIUDAD UNIVERSITARIA, CD. MX. OCTUBRE 2023



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice general

Resumen/Abstract	iv
Agradecimientos	vii
1 Introducción	1
2 Marco teórico	4
2.1 Inferencia bayesiana	4
2.2 Teorema de Bayes	4
2.3 Cadenas de Markov-Monte Carlo	6
2.4 Ejemplo de aplicación: Regresión lineal	7
3 Cuantificación de la incertidumbre usando IB en la inversión estocástica del deslizamiento estático	14
3.1 Problema directo	14
3.2 PDF posterior conjunta y notación	15
3.3 PDF marginal	19
3.3.1 Ejemplo descomposición 1D	22
3.3.2 Ejemplo descomposición 2D	23
3.4 Aproximación de la apriori	25
3.5 Media, varianza y coeficiente de variación	25
4 Casos de estudio	27
4.1 Caso sintético	27
4.2 Caso real	29
4.2.1 Metodologías propuestas	30
4.2.2 Evaluación de coeficientes de variación	42
4.3 Alternativas al coeficiente de variación	45
5 Conclusiones y trabajo a futuro	49
Bibliografía	51
Apéndice	54

Índice de figuras

2.1	Datos sintéticos generados a partir del modelo de la recta, las barras verticales indican la desviación estándar de cada observación.	8
2.2	Curva de convergencia para las muestras de la pendiente utilizando MCMC.	10
2.3	Curva de convergencia para las muestras de la ordenada al origen utilizando MCMC.	11
2.4	Histograma de las muestras de la cadena generada con MCMC para la pendiente después de retirar las primeras 450 muestras.	11
2.5	Histograma de las muestras de la cadena generada con MCMC para la pendiente después de retirar las primeras 450 muestras.	12
2.6	Rectas obtenidas en función de los parámetros ajustados con inferencia bayesiana (IB) y con mínimos cuadrados (MC).	13
3.1	Aproximación a una distribución uniforme mediante distribuciones normales truncadas en función de α	26
4.1	PDF posterior conjunta calculada mediante la estrategia semianalítica (izquierda) y con MCMC (derecha).	28
4.2	PDFs marginales obtenidas con ambos métodos para el parámetro m_1 (izquierda) y para el parámetro m_2 (derecha).	28
4.3	Serie de tiempo del desplazamiento y la velocidad de un pulso de duración τ en campo lejano (izquierda) y sus espectros (derecha). Imagen modificada de (Shearer, 2009).	30
4.4	Variación del resultado de la función <i>mvnun</i> para diferentes valores de α	33
4.5	Coefficientes de variación para el deslizamiento invertido de cada subfalla obtenidos a partir de la aproximación despreciando el término de asimetría, $\alpha = 3$ (los valores en cero se resaltan en color blanco).	33
4.6	Acumulación de masa en la campana de una gaussiana contra número de dimensiones. Imagen modificada de (Giraud, 2021)	34
4.7	Matriz de covarianza de la posterior, $\alpha = 3$	35
4.8	Coefficientes de variación para el deslizamiento invertido de cada subfalla a partir de la simulación de la posterior, $\alpha = 3$ (los valores en cero se resaltan en color blanco).	36
4.9	Comparación entre marginales usando una secuencia de datos fija contra una variable con QMC, $\alpha = 3$	38

4.10 Prueba de robustez, secuencia de datos fija contra una variable con QMC, $\alpha = 3$	39
4.11 Comparación de resultados obtenidos con mvnun contra QMC usando una secuencia fija, $\alpha = 3$	39
4.12 Análisis del efecto del tamaño de muestra, parámetro 1 $\alpha = 5$	41
4.13 Análisis del efecto del valor de α , discretización 50, 16 muestras.	42
4.14 Análisis del efecto de la discretización del vector de recorrido de la marginal, $\alpha = 8$, 16 muestras.	43
4.15 Coeficientes de variación para el deslizamiento invertido de cada subfalla utilizando QMC, $\alpha = 8$ (los valores en cero se resaltan en color blanco).	44
4.16 Comparación entre los métodos revisados, $\alpha = 8$	44
4.17 Comparación entre los métodos revisados, $\alpha = 8$ (los valores en cero se resaltan en color blanco).	45
4.18 Coeficientes de variación calculados respecto al MAP.	46
4.19 Mapa de los índices de restitución obtenidos usando la estrategia de MOCs.	47
4.20 Izquierda: Mapa de la desviación estándar de la marginal para cada subfalla. Derecha: Mapa del deslizamiento invertido mediante ELADIN proporcionado por el Dr.Cruz Atienza.	48

Cuantificación de la Incertidumbre en la Inversión del Deslizamiento Sísmico

Isaac Rubén Valverde Guerrero

Resumen

La cuantificación de la incertidumbre en la resolución de problemas de inversión es de suma importancia, ya que nos permite estimar la confiabilidad de los resultados obtenidos. Por lo general, estos problemas se resuelven mediante optimización numérica. Sin embargo, dentro del marco teórico de la optimización, no existe una forma directa de cuantificar la incertidumbre.

Por otro lado, la inferencia bayesiana ofrece una alternativa en la cual se puede estimar la incertidumbre de forma inherente, ya que implica determinar distribuciones de probabilidad posteriores cuya varianza está relacionada con la incertidumbre, siendo una medida de la dispersión del valor más probable.

Normalmente, los problemas de inferencia bayesiana se resuelven mediante el muestreo de la distribución posterior utilizando cadenas de Markov-Monte Carlo (MCMC). Sin embargo, este enfoque puede resultar computacionalmente costoso. En el caso de la inversión del deslizamiento ocasionado por sismos a partir de datos de desplazamiento GPS, Noquet (2018) propone una estrategia semianalítica que permite estimar las distribuciones posteriores sin la necesidad de utilizar MCMC.

La dificultad de la estrategia semianalítica de Noquet radica en el cálculo de integrales multivariadas de funciones exponenciales. Al utilizar la estrategia numérica propuesta, la metodología se vuelve ineficiente para problemas de 100 o más dimensiones. Como alternativa, se evalúan dos propuestas: la primera consiste en simplificar el problema para evitar el cálculo de las integrales, y la segunda propuesta implica el cálculo de las integrales mediante un estimador cuasi-Monte Carlo. Ambas propuestas se comparan con una eficiente MCMC que muestrea la distribución posterior. Estas estrategias propuestas se utilizaron para analizar la incertidumbre asociada a la inversión del deslizamiento del sismo de Acapulco del 7 de septiembre de 2021, considerando todas las estaciones GPS disponibles.

Abstract

The uncertainty quantification in inverse problems solving is extremely important as it allows us to estimate the reliability of the obtained results. Typically, such problems are solved using numerical optimization. However, within the theoretical framework of numerical optimization, there is no direct way to quantify uncertainty.

On the other hand, Bayesian inference offers an alternative approach in which uncertainty can be inherently estimated. This is achieved by determining posterior probability distributions whose variance is proportional to the uncertainty, as it measures the dispersion of the most probable value.

Commonly, Bayesian inference problems are resolved through sampling from the posterior distribution using Markov Chain Monte Carlo (MCMC) methods. Unfortunately, this approach can be computationally expensive. For the inversion of earthquake-induced slip based on GPS displacement data, Noquet (2018) proposes a semi-analytical strategy that allows estimation of posterior distributions without the need for MCMC.

The challenge with Noquet's semi-analytical strategy lies in the need to compute multivariate integrals of exponential functions. By employing the proposed numerical strategy, the methodology becomes inefficient for problems with 100 or more dimensions. As an alternative, two approaches are evaluated: the first involves simplifying the problem to avoid integral calculations, while the second approach involves estimating the integrals using a quasi-Monte Carlo estimator. Both proposals are compared against an efficient MCMC method that samples the posterior distribution. These proposed strategies were used to analyze the uncertainty associated with the slip inversion of the September 7, 2021, Acapulco earthquake, including all available GPS stations.

Agradecimientos

Al Posgrado en Ciencias de la Tierra-UNAM por la formación académica.

Al Consejo Nacional de Ciencia y Tecnología por la beca otorgada para la realización de estudios de Maestría.

Agradezco a los sinodales que forman parte del jurado para mi examen de grado; Dr. Víctor Manuel Cruz Atienza, Dr. Arturo Iglesias Mendoza, Dr. Francisco Sánchez Sesma y Dra. Ursula Xiomara Iturrarán Viveros, por el tiempo que dedicaron a leer mi tesis, realizando observaciones y comentarios que contribuyeron a mejorarla significativamente.

Estoy agradecido de haber tenido la oportunidad de trabajar con mi asesor, el Dr. Josué Tago, quien, en mi opinión, es el claro ejemplo de la definición de la palabra 'tutor'. Siempre mostró un genuino interés y compromiso en apoyarme en todo lo necesario para realizar este trabajo de la mejor forma posible, teniendo la mejor disposición para orientarme de manera clara y oportuna cada vez que recurrí a él. Por todo lo anterior, además de ser una persona en quien confío plenamente, es alguien a quien admiro como investigador y docente.

Agradezco la valiosa colaboración del M.C. Isaias Bañales, quien amablemente me facilitó un código computacional que contribuyó ampliamente a este trabajo.

También agradezco la ayuda de mi compañera, la Ing. Deni López, y de la persona que curiosamente fue quien más me ayudó cuando me encontré con algún obstáculo en tareas o proyectos, aún cuando ni siquiera tomaba las mismas clases que yo, el Ing. Gustavo López. Estoy convencido de que es el geofísico más capaz de mi generación y creo firmemente que tendrá éxito en cualquier cosa a la que se dedique.

Considero que el hecho de que alguien sea de mi familia no significa necesariamente que tenga que ser valioso para mí. Sin embargo, estoy agradecido de tener personas que lo son. Han estado cerca de mí y cada vez que nos vemos, lo hacemos con gusto. Mis padrinos, Elisa y Óscar, así como mis primas Karla, Adriana, Araceli y Evelyn, que son como mis hermanas.

Me siento sumamente afortunado de haber podido experimentar el amor más puro y sincero que pude recibir de una mujer. El tiempo que conviví con ella me dejó una huella tan profunda que siento que su recuerdo sigue tan presente en mí como el día en que se fue. Agradezco que mi madre sigue siendo una luz que me guía por el buen camino.

Estoy muy agradecido con mi hermano, Alberto Valverde. Desde que tengo memoria, él ha sido un ejemplo para mí, ya que no solo es una gran persona, sino que es alguien a quien siempre tengo la confianza de recurrir. Disfruto cada momento que

comparto con él y, aunque no siempre se lo demuestre, le tengo un gran cariño y respeto. Los hermanos no se eligen, pero si tuviera la oportunidad, lo escogería sin ninguna duda.

Y finalmente, con quien estoy más agradecido es con mi padre, Marco Polo Valverde. La mayor parte de lo que soy y lo que he logrado hasta ahora es gracias a él. Ha estado para mí en momentos felices, en momentos tristes, momentos de frustración y momentos de satisfacción, siempre dando lo mejor de sí para procurar mi bienestar. No tengo manera de agradecer todo lo que ha hecho por mí; sin embargo, espero poder mostrarle mi gratitud siendo una persona que lo haga sentir tan orgulloso como yo lo estoy de él.

Capítulo 1

Introducción

La caracterización de una zona sísmicamente activa involucra, entre otras cosas, el estudio de la distribución del deslizamiento en las fallas existentes, el cual es ocasionado por el movimiento relativo entre dos bloques a ambos lados de dichas fallas ([Wei, Chen, y Meng, 2023](#)).

La velocidad del deslizamiento determina si se trata de un sismo lento o de un sismo regular. El proceso de ruptura de un sismo regular tiene una duración de segundos, mientras que la de un sismo lento puede durar varios meses. En cualquier caso, estos deslizamientos provocan deformaciones en la superficie que pueden ser medidas con estaciones GPS.

El papel de los eventos de deslizamiento lento (SSE, siglas en inglés de slow slip event) en el ciclo sísmico ha sido considerado en trabajos como [Cruz-Atienza et al. \(2021\)](#) y [Radiguet et al. \(2016\)](#) debido a que se cree que pueden estar estrechamente relacionados con la ocurrencia de sismos de magnitud importante que han ocasionado múltiples afectaciones.

Este tipo de eventos muestra una variabilidad importante en términos de duración (desde días hasta años), de desplazamientos superficiales registrados (desde unos pocos milímetros hasta unos pocos centímetros) y tiempo de recurrencia ([Radiguet et al., 2011](#)). Aunque el desarrollo de instrumentos GPS cada vez más precisos ha permitido el estudio de los SSE, el mecanismo responsable de su ocurrencia aun no es del todo claro ([Radiguet et al., 2011](#)).

Lograr un mejor entendimiento del deslizamiento de sismos regulares o lentos es fundamental para explicar el comportamiento de las zonas de subducción, especialmente la acumulación de esfuerzos en zonas fuertemente acopladas que pueden producir sismos de gran magnitud ([Ito et al., 2013](#)).

La cuantificación de la incertidumbre representa uno de los más grandes retos en la estimación de parámetros geofísicos en general. Esta consiste en poder medir la confianza que tenemos en un modelo geofísico que se ha inferido a partir de observaciones indirectas. Debido a que los modelos geofísicos se utilizan para tomar decisiones empresariales, que pueden ser muy costosas, la misma industria de exploración geofísica ha impulsado esta área de investigación.

La principal herramienta utilizada para los problemas de inversión en geofísica ha sido la teoría de optimización numérica (ON) ([Nocedal y Wright, 2006](#)). Lamenta-

blemente, bajo el marco teórico de la ON no existe una forma directa de evaluar la incertidumbre. Generalmente se realizan pruebas de tablero de ajedrez que utilizan modelos sintéticos con distintos tamaños de casilla para estimar qué tan bien pueden reconstruirse. Recientemente, [Tago et al. \(2021\)](#) proponen una estrategia que involucra tableros de ajedrez móviles (MOCs, por sus siglas en inglés *Mobile Checkerboards*) para poder cuantificar la incertidumbre. Sin embargo, esto es muy costoso computacionalmente debido al número de evaluaciones que se requieren.

Otra manera sencilla de evaluar la incertidumbre es a través de la matriz de resolución, pero esta no siempre puede calcularse porque requiere la inversión de una matriz que puede ser muy grande o estar mal condicionada ([Tarantola y Valette, 1982](#)). De manera ingeniosa, [Fichtner y Leeuwen \(2015\)](#) proponen una estrategia de sondeo aleatorio para inferir la resolución a través del Hessiano usando modelos aleatorios. Al final, todas estas estrategias, por sofisticadas que sean, intentan subsanar la falta de un marco teórico en ON.

Por otro lado, la inferencia bayesiana (IB) ofrece una alternativa al déficit en ON debido a que consiste en buscar no solo un modelo, sino toda una distribución de probabilidad cuyo soporte es proporcional a la incertidumbre de la inversión. Esta distribución se conoce como “posterior” (PDF, por sus siglas en inglés *Probability Density Function*) y se deriva a partir de la combinación de una distribución *a priori* que considera el espacio de modelos y una distribución que considera el espacio de datos.

Se han desarrollado múltiples métodos para determinar la posterior en aplicaciones para el problema inverso de deslizamiento finito. [Minson, Simons, y Beck \(2013\)](#) utilizan el método de cadenas de Markov-Montecarlo (MCMC), basado en un algoritmo de Metropolis en cascada para muestrear la posterior conjunta. El inconveniente es que para obtener una buena aproximación es necesario determinar una gran cantidad de muestras, lo cual requiere una gran cantidad de cálculo computacional.

Una forma rápida de obtener la posterior, a costa de reducir la generalidad del método, es seleccionando una distribución específica para la *a priori*. [Tarantola y Valette \(1982\)](#) demuestran que, suponiendo un problema inverso lineal con una distribución normal como *a priori*, la posterior conjunta también es una distribución normal y además proporcionan ecuaciones para calcular la expectativa y la covarianza posterior del modelo.

En el caso de la inversión del deslizamiento, un inconveniente de usar una distribución normal como *a priori* es que implica considerar la posible ocurrencia de deslizamiento negativo en zonas donde se podría tener un acoplamiento sísmico mayor al 100 por ciento ([Savage, 1983](#)), lo cual puede resultar físicamente incongruente. Por un lado, [Yagi y Fukahata \(2011\)](#) indican que si el deslizamiento negativo es significativamente mayor que los errores de estimación, esto revela problemas en el modelo físico y/o un modelo de error inadecuado. Por otro lado, autores como [Minson et al. \(2014\)](#) o [Nocquet et al. \(2014\)](#) sostienen que el deslizamiento positivo es información que se basa en una suposición física razonable que debe tomarse en cuenta como *a priori* en la inversión y no usarse como validación en el análisis posterior.

Según [Nocquet \(2018\)](#), las restricciones de no negatividad o de delimitación son razonables en un sentido físico debido a que durante un sismo se libera el esfuerzo acumulado y se debe inducir deslizamiento en la dirección opuesta al movimiento relativo

de la placa.

La estrategia que propone [Nocquet \(2018\)](#) permite añadir restricciones mediante la *a priori*, utilizando distribuciones normales truncadas multivariadas (TMVN, por sus siglas en inglés *Truncated Multivariate Normal*), aprovechando los avances que se han hecho en este campo. Siguiendo el enfoque de [Tarantola y Valette \(1982\)](#) pero en un espacio de modelos acotado, se considera que si la *a priori* es una normal truncada multivariada, la posterior también es una normal truncada multivariada. Esta propuesta es atractiva debido a que se plantea una estrategia semianalítica que permite obtener distribuciones posteriores, así como los parámetros que las caracterizan necesarios para cuantificar la incertidumbre de manera eficiente, evitando el uso de MCMC.

El principal obstáculo al emplear la estrategia de [Nocquet \(2018\)](#) es el cálculo de integrales multidimensionales. En el presente trabajo se exploran dos alternativas para lidiar con este inconveniente, la primera consiste en una simplificación para evitar el cálculo de dichas integrales, mientras que en la segunda se utiliza el estimador de cuasi-Monte Carlo (QMC), el cual comúnmente se aplica en la resolución de integrales de varias variables. Los resultados obtenidos se comparan con un método eficiente de simulación de la posterior, con la finalidad de evaluar qué tan bien se pueden calcular las PDF a partir de las propuestas antes mencionadas.

Capítulo 2

Marco teórico

2.1 Inferencia bayesiana

La inferencia bayesiana es un proceso en el que una creencia, que se expresa como una distribución de probabilidad llamada “*distribución a priori*”, se actualiza a partir de datos observados para obtener otra distribución de probabilidad que se conoce como “*distribución posterior*”. El Teorema de Bayes nos permite calcular la distribución posterior a partir de la distribución *a priori* y los datos observados que se tengan disponibles (Kruschke, 2015).

Este método se va a emplear para resolver un problema de inversión, por lo tanto, la creencia que se menciona anteriormente se refiere a una propuesta inicial de los parámetros de un modelo que describe un sistema físico, que en este caso corresponde al deslizamiento entre las placas tectónicas. Posteriormente, se actualiza el modelo buscando aquel que reproduzca mejor las observaciones que se tengan disponibles. El teorema de Bayes y la forma en la que se interpreta para un problema de inversión se describen a continuación.

2.2 Teorema de Bayes

Para comenzar a adentrarse en la inferencia bayesiana es necesario entender algunos conceptos básicos, lo primero y quizá más importante es el teorema de Bayes. Partiendo del concepto de probabilidad condicional, la cual se denota como $P(B|A)$ ¹ y se expresa matemáticamente de la siguiente forma

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \quad P(A) > 0, \quad (2.2.1)$$

donde A y B son subespacios de un espacio muestral S , $P(A)$ es la probabilidad de que ocurra el evento A y $P(B \cap A)$ es la probabilidad de la intersección de los eventos B y A . De la misma forma en la que expresa $P(B|A)$ en la ec. (2.2.1), se puede escribir $P(A|B)$ como

¹Esta expresión se lee como “la probabilidad de que ocurra B, dado que ocurrió A”, o simplemente, “la probabilidad de B, dado A” (Walpole, Myers, Myers, y Ye, 2012).

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad P(B) > 0. \quad (2.2.2)$$

Sabiendo que $A \cap B$ es equivalente a $B \cap A$, se pueden despejar estos términos en sus respectivas ecuaciones e igualarlas, con lo que se obtiene

$$\begin{aligned} P(A|B)P(B) &= P(B|A)P(A), \\ P(A|B) &= \frac{P(B|A)P(A)}{P(B)}. \end{aligned} \quad (2.2.3)$$

La ec. (2.2.3) se conoce como regla de Bayes o Teorema de Bayes y se puede considerar como la base de la inferencia bayesiana. El Teorema de Bayes en términos de distribuciones de probabilidad puede escribirse como

$$\sigma(\theta|\mathbf{x}) = \frac{\rho(\mathbf{x}|\theta)\rho(\theta)}{\rho(\mathbf{x})}, \quad (2.2.4)$$

donde $\sigma(\theta|\mathbf{x})$ se denomina como “*posterior*”², $\rho(\mathbf{x}|\theta)$ se denomina “*verosimilitud*”, $\rho(\theta)$ es la “*distribución a priori*” y $\rho(\mathbf{x})$ es la verosimilitud marginal o “*evidencia*” (Lee y Wagenmakers, 2014).

Recordando que el propósito de emplear inferencia bayesiana es resolver un problema de inversión, es conveniente identificar qué representa cada término de la ec. (2.2.4). La posterior, $\sigma(\theta|\mathbf{x})$, corresponde a la distribución de probabilidad de los parámetros θ del modelo, dados los datos observados \mathbf{x} , *i.e.* representa la distribución de la estimación de los parámetros actualizada a partir de las observaciones del fenómeno.

La verosimilitud, $\rho(\mathbf{x}|\theta)$, es la probabilidad de que el modelo genere los datos observados, \mathbf{x} , a partir del valor que se tiene de los parámetros, θ , o dicho de otra forma, qué tan probable es obtener los datos observados con estos parámetros del modelo.

La distribución *a priori*, $\rho(\theta)$, es la credibilidad de los parámetros, θ , independiente de las observaciones. Finalmente, la evidencia, $\rho(\mathbf{x})$, representa la probabilidad general de los datos a partir del modelo, determinada por el promedio de todos los valores posibles de los parámetros, ponderados por la credibilidad que tienen dichos parámetros (Kruschke, 2015).

El cálculo de la posterior a partir del Teorema de Bayes implica la resolución de una integral para calcular la verosimilitud marginal que, en el caso habitual de parámetros continuos, puede ser imposible de resolver analíticamente (Gelman, Carlin, Stern, y Rubin, 2014). Una alternativa, dada la dificultad de la integración, es restringir los modelos a funciones de verosimilitud relativamente simples con fórmulas para distribuciones *a priori* llamadas “*a priori conjugadas*”, que van acordes a ciertas funciones de verosimilitud para obtener integrales manejables (Kruschke, 2015).

Cuando la propuesta de la *a priori* conjugada no resulta conveniente, otra opción es aproximar las funciones reales con otras funciones que sean más fáciles de trabajar y posteriormente mostrar que la aproximación es razonablemente buena en condiciones

²En la literatura normalmente se menciona únicamente como *posterior*, se refiere a la distribución de densidad de probabilidad posterior.

típicas. A este método se le conoce como “*aproximación variacional*” (Gelman et al., 2014).

Existe otra clase de métodos en los que no se resuelve la integral de forma analítica, sino que se resuelve con una aproximación numérica. El espacio de parámetros se discretiza usando una cuadrícula de puntos y posteriormente se hace una suma a través de toda la malla para obtener el valor de la integral. Este método es eficaz para problemas que no involucren un gran número de parámetros, debido a que la discretización de cada parámetro puede generar una cantidad de combinaciones que puede ser demasiado grande para manejar, incluso con computadoras de última generación (Kruschke, 2015).

Otro tipo de aproximación implica el muestreo aleatorio de un gran número de combinaciones representativas de valores de los parámetros de la posterior. Los algoritmos que se emplean normalmente para realizar el muestreo se denominan métodos de Cadenas Markov-Monte Carlo (MCMC). La ventaja de esta clase de métodos es que se puede trabajar con modelos complejos sin resolver la integral involucrada en el teorema de Bayes (Lee y Wagenmakers, 2014). En la siguiente sección se habla de las MCMC con más detalle.

2.3 Cadenas de Markov-Monte Carlo

Las cadenas de Markov-Monte Carlo (MCMC) son una familia de algoritmos de muestreo que permite caracterizar una distribución de probabilidad sin conocer todas sus propiedades matemáticas, esto mediante el muestreo aleatorio de valores de dicha distribución.

El nombre “cadenas de Markov-Monte Carlo” proviene de la combinación de dos procesos; el primero denominado Monte Carlo, consiste en estimar las propiedades de una distribución examinando muestras aleatorias de dicha distribución; el segundo se denomina cadena de Markov, que consiste en generar muestras aleatorias de modo secuencial en el que cada muestra aleatoria se utiliza para generar la siguiente muestra aleatoria, es decir, cada muestra depende únicamente de la muestra inmediata anterior.

Quizá el algoritmo más simple de MCMC es el de Metrópolis, que comienza con un valor inicial plausible de acuerdo con el rango de valores que determinan el soporte de la distribución.

El segundo paso es generar una nueva propuesta a partir de la anterior agregando ruido aleatorio, el cual se genera a partir de una distribución propuesta que debe ser simétrica y centrada en cero.

El tercer paso es comparar el valor de la muestra evaluada en una distribución de probabilidad propuesta para la posterior (la altura de la función) con el valor de la muestra anterior evaluada en esta misma distribución. En este punto se utiliza un criterio de aceptación en el que, si la evaluación de la nueva muestra tiene un valor más alto que la anterior, se acepta; de lo contrario se decide si se acepta o se rechaza aleatoriamente con una probabilidad que se asigna en función de la relación entre la propuesta y la muestra anterior.

El criterio de aceptación consiste en dividir el valor de la posterior evaluada con

Ejemplo de aplicación: Regresión lineal

la propuesta entre el valor de la distribución evaluada con la muestra anterior y posteriormente se toma esa relación como porcentaje de probabilidad de aceptación. En caso de aceptación, la propuesta se vuelve el nuevo eslabón de la cadena, es decir la nueva muestra, pero si se rechaza se repite el valor previo. El proceso que se acaba de describir corresponde a una iteración, para obtener la siguiente se regresa al segundo paso hasta obtener la cantidad de muestras suficientes que permitan la reconstrucción de la posterior (Van Ravenzwaaij, Cassey, y Brown, 2018). El algoritmo (1) muestra el pseudocódigo del algoritmo de Metrópolis.

Algoritmo 1 Algoritmo de Metrópolis

Entrada: *propuesta inicial*

Salida: *cadena*

$i \leftarrow 0$

$y \leftarrow \text{cadena}$

$x \leftarrow \text{propuesta inicial}$

$N \leftarrow \text{Elementos de la cadena}$

mientras $N \neq 0$ **hacer**

si $\text{posterior}(x + \text{ruido aleatorio}) > \text{posterior}(x)$ **entonces**

$x \leftarrow x + \text{ruido aleatorio}$

si no si $\text{posterior}(x + \text{ruido aleatorio}) / \text{posterior}(x) > \text{num aleatorio}$ **entonces**

$x \leftarrow x + \text{ruido aleatorio}$

si no

$x \leftarrow x$

fin si

$i \leftarrow i + 1$

$y(i) \leftarrow x$

$N \leftarrow N - 1$

fin mientras

Un aspecto importante a considerar es que las primeras muestras de la cadena pueden no ser representativas de la posterior, esto debido a que las muestras se empiezan a generar a partir de una propuesta inicial, la cual puede ser poco probable que provenga de la distribución que se está aproximando. Lo anterior implica que antes de alcanzar la convergencia, las muestras anteriores no deben ser consideradas y por lo tanto se deben eliminar, lo cual se conoce como convergencia o quemado (*convergence* o *burn-in*) (Van Ravenzwaaij et al., 2018).

En la siguiente sección se muestra el método de MCMC usando el algoritmo de Metrópolis con un ejemplo simple.

2.4 Ejemplo de aplicación: Regresión lineal

Para ejemplificar el uso de la inferencia bayesiana en un problema de inversión, se propone un caso sencillo que se resuelve usando el algoritmo de Metrópolis para MCMC.

Ejemplo de aplicación: Regresión lineal

Consiste en ajustar el modelo de una recta utilizando datos sintéticos generados a partir de la evaluación de una serie de puntos (dominio de la función) generados de forma aleatoria de una distribución uniforme definida de 0 a 2.

Posteriormente, a los datos sintéticos se les agrega ruido aleatorio con una distribución normal, cuya media es el valor de la recta evaluada en el respectivo punto del dominio y la desviación estándar es un número aleatorio de una distribución uniforme de 0.1 a 0.2 (ver Fig. 2.1).

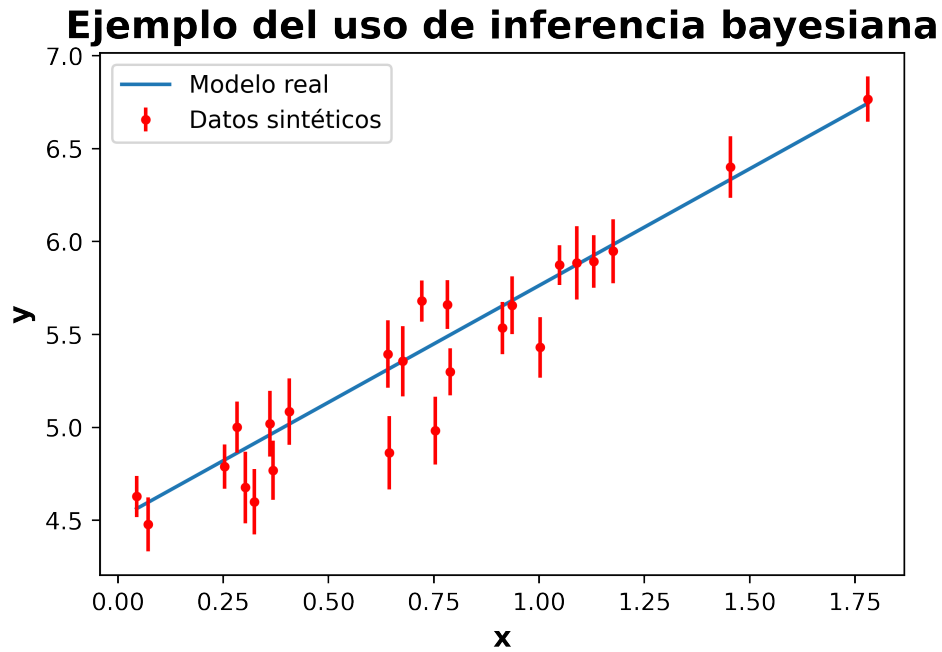


Figura 2.1: Datos sintéticos generados a partir del modelo de la recta, las barras verticales indican la desviación estándar de cada observación.

El modelo de la recta depende de dos parámetros; la pendiente, cuyo valor real se definió como $m = 1.256$ y la ordenada al origen, que se definió como $b = 4.507$.

En este caso, se sabe que los errores (variaciones) en los datos se distribuyen normalmente, por lo tanto, para la verosimilitud se propone una distribución normal, la cual se define como

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2}\right), \quad (2.4.1)$$

donde x es el dominio de la función, μ es la media y corresponde al punto en el que se centra la distribución y σ es la desviación estándar, que corresponde al parámetro de forma que indica el ancho de la curva.

Para este ejemplo, se obtiene el logaritmo natural de la ec. (2.4.1) y se escribe en términos de las variables involucradas en el problema

$$\begin{aligned} \ln(\mathcal{N}(y|\sigma^2)) &= \sum_{i=1}^N \ln\left(\frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2} \frac{(d_i - y_i)^2}{\sigma_i^2}\right)\right), \\ &= -\frac{1}{2} \left[N \cdot \ln(2\pi) + 2 \sum_{i=1}^N \ln(\sigma_i) + \sum_{i=1}^N (d_i - y_i)^2 \cdot \sigma_i^{-2} \right], \end{aligned} \quad (2.4.2)$$

donde d_i es la i -ésima observación, y_i es el i -ésimo valor del modelo en función de los parámetros y σ_i corresponde a la desviación estándar de la i -ésima observación.

El motivo por el que se usa el logaritmo natural en la distribución es que se logra una mejor estabilidad numérica, esto debido a que una gran cantidad de multiplicaciones de valores pequeños puede causar un desbordamiento en la memoria del sistema, además de que los cálculos pasan de involucrar multiplicaciones y números positivos pequeños a sumas y números negativos menos pequeños.

Para la *a priori* (que expresa nuestro conocimiento previo acerca de los parámetros) se va a definir un intervalo en el que se pueden encontrar los valores de ambos parámetros, ya que no tiene sentido considerar valores que claramente no se acercan a los que se están buscando.

En la Fig. 2.1 se puede ver claramente que la pendiente m es positiva, considerando un rango amplio de valores posibles se va a suponer que se encuentra entre 0 y 100.

De la misma manera, si sólo se pudieran ver las observaciones y pensando en una recta imaginaria que pasara lo más cerca posible de todos los puntos, difícilmente alguien podría proponer un valor que saliera de un rango entre -10 y 10, sin embargo se va a considerar un margen amplio y se define el rango de valores entre -50 y 50.

Suponiendo que no se cuenta con mayor información acerca de los parámetros, se va a dar la misma probabilidad a cualquier valor dentro de los rangos definidos anteriormente, lo cual en términos de probabilidad implica que se va a considerar una distribución uniforme que se define como

$$f(x) = \frac{1}{b-a}, \forall x \in [a, b], \quad (2.4.3)$$

donde a y b son el límite inferior y el límite superior del intervalo respectivamente.

De acuerdo con la ec. (2.2.4), la posterior se obtiene multiplicando la verosimilitud y la *a priori*. Recordando que de la ec. (2.4.2) se obtiene el logaritmo natural de la verosimilitud, es sencillo obtener el logaritmo natural de la *a priori* (ec. (2.4.3)), por lo tanto, al tomar en cuenta las propiedades de los logaritmos, esta multiplicación se vuelve una suma.

Seguramente no habrá pasado inadvertido que en el párrafo anterior no se menciona la *evidencia* $\rho(\mathbf{x})$. Como se menciona anteriormente, obtener la *evidencia* implica el cálculo de una integral que no siempre es fácil de resolver, sin embargo, para el método utilizado en este ejemplo no es necesario.

La razón por la cual se puede evitar el cálculo de la evidencia es que el resultado de la integral divide a la multiplicación de la *a priori* y la verosimilitud, por lo tanto este se puede ver como un factor de normalización que escala la posterior sin alterar la forma, haciendo que cumpla con la definición de una distribución de probabilidad

Ejemplo de aplicación: Regresión lineal

sin modificar la posición de la media (que, para la distribución normal, corresponde al parámetro más probable y por lo tanto es el valor que se está buscando), lo cual explica también el por qué en la práctica comúnmente se omite.

Teniendo ya la forma de obtener la posterior, las cadenas se generan siguiendo el algoritmo 1 con 9000 muestras cada una, los valores obtenidos para cada parámetro se muestran en las Figs. 2.2 y 2.3.

Al observar las muestras de la cadena para cada variable, el efecto de convergencia que se mencionó anteriormente es bastante notorio. Estas muestras se deben retirar para poder obtener una estimación confiable del valor que deben tener los parámetros para la aproximación. Las gráficas de las muestras pueden servir como referencia para determinar aproximadamente cuántas muestras se deben retirar.

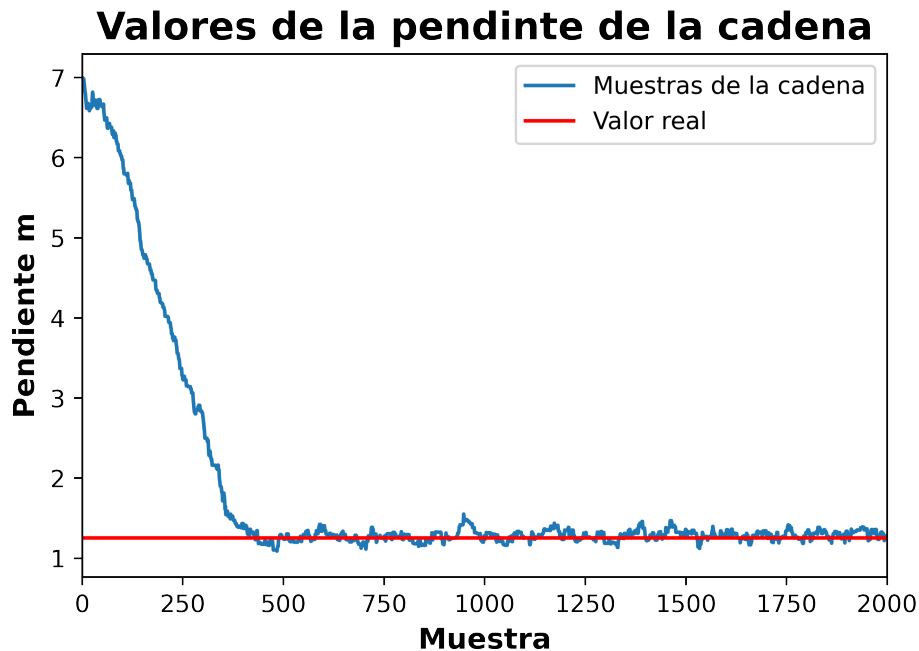


Figura 2.2: Curva de convergencia para las muestras de la pendiente utilizando MCMC.

Luego de cortar las primeras 450 muestras de cada cadena, se realizan las gráficas de los histogramas para cada una, los cuales se muestran en las Figs. 2.4 y 2.5. Además de los histogramas, las gráficas muestran una línea roja vertical que indican la mediana³ de los datos.

Como se puede observar, las distribuciones obtenidas se ven bastante simétricas y, por consiguiente, las medidas de tendencia central como la media y la mediana deben ser similares. El punto en el que están centradas las distribuciones corresponde al valor más alto y, por tanto, al valor más probable para cada parámetro, es decir, el valor que se está buscando estimar.

El valor de la mediana obtenido de la posterior para la pendiente fue de 1.27, mientras que el valor de la media fue de 1.2694. Como se anticipó anteriormente, son

³Se utiliza la mediana debido a que es menos susceptible a valores atípicos que otras medidas de tendencia central como la media o la moda.

Valores de la ordenada al origen de la cadena

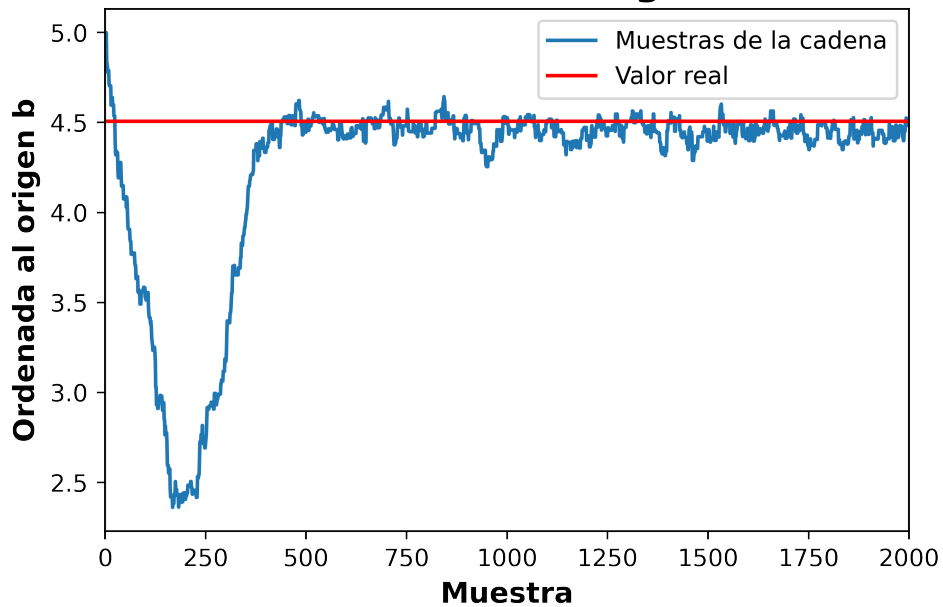


Figura 2.3: Curva de convergencia para las muestras de la ordenada al origen utilizando MCMC.

Histograma de la posterior para la pendiente

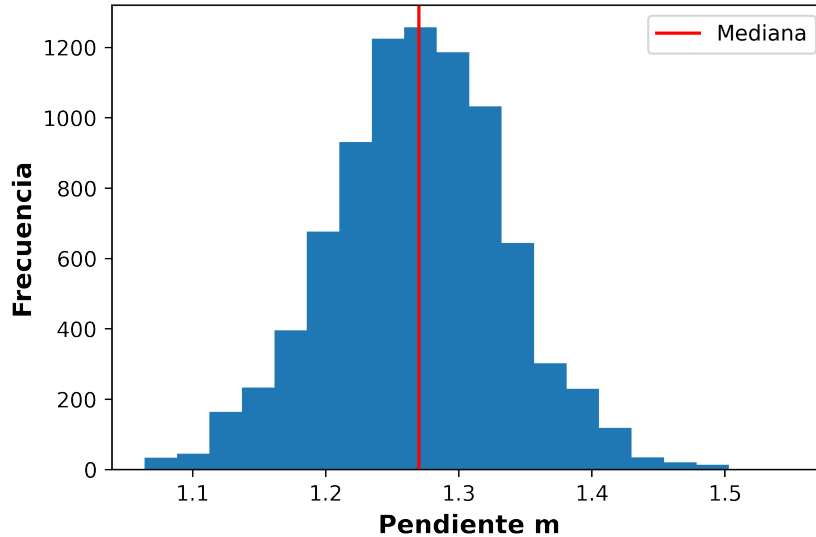


Figura 2.4: Histograma de las muestras de la cadena generada con MCMC para la pendiente después de retirar las primeras 450 muestras.

valores bastante similares entre sí y son también bastante cercanos al valor real de 1.256. Así mismo, para la mediana obtenida de la posterior para la ordenada al origen se obtiene un valor de 4.4714, mientras que el valor de la media fue de 4.4697 que, comparando con el valor real de 4.507 son considerablemente similares.

Histograma de la posterior para la ordenada al origen

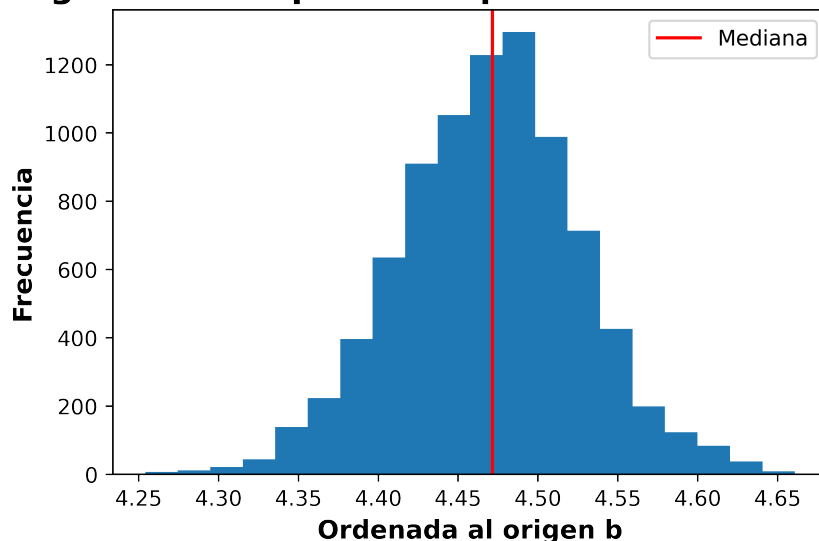


Figura 2.5: Histograma de las muestras de la cadena generada con MCMC para la pendiente después de retirar las primeras 450 muestras.

Resolviendo este mismo ejemplo con el método de mínimos cuadrados que se utiliza frecuentemente en problemas de inversión lineales se obtienen valores de 1.2952 para la pendiente y 4.4169 para la ordenada al origen. Esto se realiza utilizando el método “*polyfit*” incluido en la paquetería “*numpy*” (Harris et al., 2020).

Los modelos obtenidos en función de los parámetros ajustados con ambos métodos se muestran en la Fig. 2.6. Como se puede observar, la recta cuyos parámetros se calculan usando inferencia bayesiana se acerca ligeramente más a la recta real que aquella cuyos parámetros se calculan usando mínimos cuadrados. La peculiaridad de la inferencia bayesiana es que se puede obtener la desviación estándar ligada al parámetro obtenido, la cual a su vez esta relacionada con la incertidumbre de dicho parámetro.

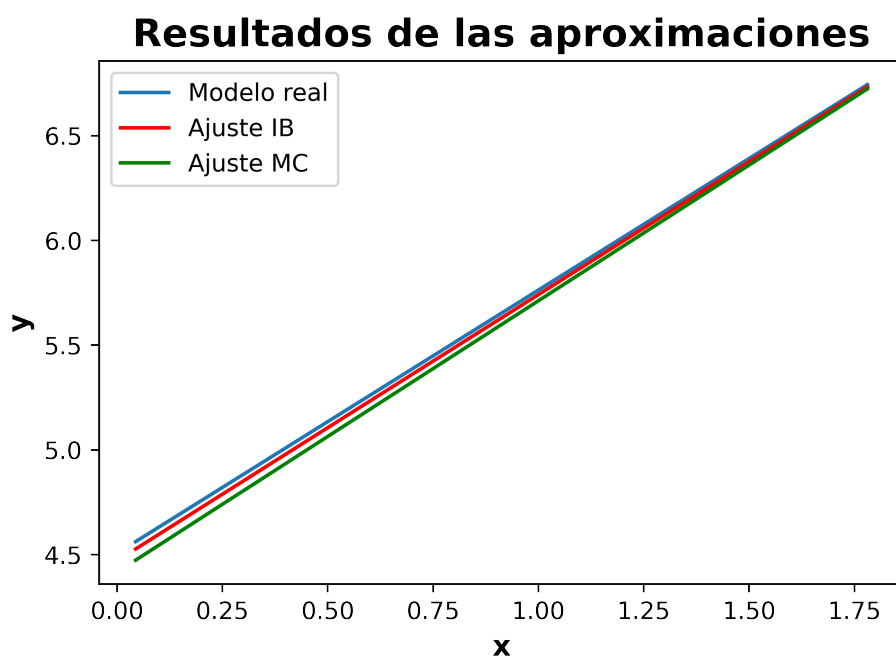


Figura 2.6: Rectas obtenidas en función de los parámetros ajustados con inferencia bayesiana (IB) y con mínimos cuadrados (MC).

Capítulo 3

Cuantificación de la incertidumbre usando IB en la inversión estocástica del deslizamiento estático

Como se menciona anteriormente, esta tesis esta basada en el trabajo de [Nocquet \(2018\)](#), por lo tanto todas las ecuaciones que se muestran a continuación fueron extraídas de este artículo a no ser que se indique lo contrario.

3.1 Problema directo

Para calcular el campo de desplazamiento, $D(x)$, debido a un deslizamiento, $M(\xi)$, producido en una falla, Σ , se puede utilizar el Teorema de representación de elastostática dado por

$$D_j(x) = \int_{\Sigma} T_k(S_{ij}(\xi, x))M_k(\xi)d\Sigma, \quad i, j, k \in \{x, y, z\}, \quad (3.1.1)$$

donde $T_i(\cdot)$ es el i -ésimo componente del vector de tracciones calculado sobre la falla, a partir del tensor de Somigliana, $S_{ij}(\cdot, \cdot)$, que representa el desplazamiento en superficie generado por fuentes puntuales impulsivas ([Udías, Vallina, Madariaga, y Buforn, 2014](#)).

[Okada \(1985\)](#) obtuvo la solución analítica de la ec. (3.1.1) para el caso de un semiespacio homogéneo lineal elástico. Utilizando las ecuaciones de Okada, se puede modelar de manera lineal al desplazamiento, $D(x)$, debido a un deslizamiento, $M(\xi)$, sobre una falla plana como

$$\begin{bmatrix} D_1(x) \\ D_2(x) \\ D_3(x) \end{bmatrix} = \begin{bmatrix} D_1(x|s_p(\xi) = 1) & D_1(x|s_c(\xi) = 1) \\ D_2(x|s_p(\xi) = 1) & D_2(x|s_c(\xi) = 1) \\ D_3(x|s_p(\xi) = 1) & D_3(x|s_c(\xi) = 1) \end{bmatrix} \begin{bmatrix} M_p(\xi) \\ M_c(\xi) \end{bmatrix}, \quad (3.1.2)$$

donde las columnas de la matriz son soluciones analíticas de Okada para deslizamientos unitarios en direcciones perpendiculares sobre la falla, p y c respectivamente. La elección

de las direcciones perpendiculares p y c sobre la falla, Γ , son arbitrarias. Sin embargo, resulta conveniente elegir las de acuerdo al problema que se quiera estudiar. Por ejemplo, si se desea estudiar el acoplamiento tectónico, es conveniente elegir a la dirección de convergencia como una de las direcciones sobre la falla.

Para modelar deslizamientos complicados sobre geometrías de falla complejas, resulta conveniente discretizar a la falla en pequeñas subfallas planas. Agrupando la respuesta de todas las subfallas para todas las estaciones, utilizando para cada par de subfalla-estación la ec. (3.1.2), se puede formar el sistema lineal

$$d = Gm, \tag{3.1.3}$$

donde $d \in \mathbb{R}^{3 \cdot \#estaciones}$, $G \in \mathbb{R}^{3 \cdot \#estaciones \times 2 \cdot \#subfallas}$ y $m \in \mathbb{R}^{2 \cdot \#subfallas}$.

3.2 PDF posterior conjunta y notación

Al resolver un problema de inversión, es importante tener una idea del rango de valores que pueden tomar las variables que se busca determinar, esto debido a que no tiene sentido explorar magnitudes que no son físicamente coherentes.

Tomando en cuenta lo anterior, para la inversión del deslizamiento se define un espacio de modelos lineal acotado \mathbb{M}_b , el cual esta asociado con el modelado intersísmico con $dim(\mathbb{M}_b) = p$. El hecho de que sea un espacio acotado implica que hay límites, tanto inferiores como superiores para cada una de las p dimensiones del espacio paramétrico.

Dicho espacio paramétrico se representa mediante una función de densidad de probabilidad que se define como

$$\rho_{\mathbb{M}_b} = K_m^b \exp\left(-\frac{1}{2}(m - m_0)^T C_m^{-1}(m - m_0)\right), \tag{3.2.1}$$

donde m es el vector de datos que pertenece al espacio \mathbb{M}_b , es decir $m \in \mathbb{M}_b$.

La función en la ec. (3.2.1) es muy similar a una PDF (*Probability Density Function*) normal con media m_0 y covarianza asociada C_m , la diferencia es que el factor de normalización K_m^b garantiza que la integral sobre \mathbb{M}_b es igual a 1, es decir, hace que cumpla con la definición de una PDF considerando un espacio acotado y por lo tanto es una TMVN (*Truncated Multivariate Normal*).

La verosimilitud se define de la siguiente forma

$$\rho_{\mathbb{D}} = K_d \exp\left(-\frac{1}{2}(d - d_{obs})^T C_d^{-1}(d - d_{obs})\right), \tag{3.2.2}$$

donde el término $d - d_{obs}$ corresponde a la diferencia entre los datos sintéticos calculados a partir del modelo directo y los datos observados. Tomando en cuenta lo anterior, la ec. (3.2.2) se trata de una PDF normal multivariada de los residuos con media cero, covarianza C_d y factor de normalización $K_d = ((2\pi)^n |C_d|)^{-1/2}$. Cabe recalcar que esto implica que se esta asumiendo que los residuos se distribuyen siguiendo una normal.

Considerando que el desplazamiento se relaciona de forma lineal con el deslizamiento, se sustituye la ec. (3.1.3) en d_{obs} de la ec. (3.2.2), con lo cual se obtiene

$$\rho_{\mathbb{D}} = K_d \exp\left(-\frac{1}{2}(Gm - d_{obs})^T C_d^{-1}(Gm - d_{obs})\right). \quad (3.2.3)$$

La forma general de la PDF posterior conjunta es el producto de la *a priori* y la verosimilitud multiplicado por un factor de normalización de la siguiente forma

$$\sigma_{\mathbb{M}} = k \rho_{\mathbb{M}}(m) \rho_{\mathbb{D}}(Gm), \quad (3.2.4)$$

sustituyendo por las distribuciones explícitas

$$\begin{aligned} \sigma_{\mathbb{M}} &= k K_m \exp\left(-\frac{1}{2}(m - m_0)^T C_m^{-1}(m - m_0)\right) K_d \exp\left(-\frac{1}{2}(Gm - d_{obs})^T C_d^{-1}(Gm - d_{obs})\right), \\ &= k K_m K_d \exp\left(-\frac{1}{2}\left[(m - m_0)^T C_m^{-1}(m - m_0) + (Gm - d_{obs})^T C_d^{-1}(Gm - d_{obs})\right]\right). \end{aligned}$$

Tarantola y Valette (1982) demuestran que si la *a priori* es una PDF normal multivariada, la posterior también sera una normal multivariada. Siguiendo esta idea, dado que la *a priori* definida en (3.2.3) es una TMVN, la posterior también será un TMVN.

Si además se define $S(m) = \frac{1}{2}\left[(m - m_0)^T C_m^{-1}(m - m_0) + (Gm - d_{obs})^T C_d^{-1}(Gm - d_{obs})\right]$, la posterior conjunta queda como

$$\sigma_{\mathbb{M}_b} = k_b \exp(-S(m)), \quad (3.2.5)$$

donde $k_b = \left[\int_{\mathbb{M}_b} \exp(-S(m)) dm\right]^{-1}$.

Se puede desarrollar $S(m)$ de tal manera que se llegue la forma

$$S(m) = \frac{1}{2}\left[(m - \tilde{m})^T C_{\tilde{m}}^{-1}(m - \tilde{m}) + K_0\right], \quad (3.2.6)$$

donde $K_0 = (Gm_0 - d_{obs})^T (GC_m G^T + C_d)^{-1}(Gm_0 - d_{obs})$, \tilde{m} y $C_{\tilde{m}}$ se conocen como *expectativa* y *matriz de covarianza posterior* respectivamente y se determinan usando las expresiones de Tarantola y Valette (1982)

$$\tilde{m} = (G^T C_d^{-1} G + C_m^{-1})^{-1}(G^T C_d^{-1} d_{obs} + C_m^{-1} m_0), \quad (3.2.7)$$

$$C_{\tilde{m}} = (G^T C_d^{-1} G + C_m^{-1})^{-1}. \quad (3.2.8)$$

Para llegar a la ec. (3.2.6) se empieza por desarrollar $2S(m)$ de la siguiente forma

$$\begin{aligned} 2S(m) &= (m - m_0)^T C_m^{-1}(m - m_0) + (Gm - d_{obs})^T C_d^{-1}(Gm - d_{obs}) \\ &= (m^T - m_0^T) C_m^{-1}(m - m_0) + ((Gm)^T - d_{obs}^T) C_d^{-1}(Gm - d_{obs}) \\ &= (m^T - m_0^T) C_m^{-1}(m - m_0) + (m^T G^T - d_{obs}^T) C_d^{-1}(Gm - d_{obs}) \\ &= (m^T C_m^{-1} - m_0^T C_m^{-1})(m - m_0) + (m^T G^T C_d^{-1} - d_{obs}^T C_d^{-1})(Gm - d_{obs}) \\ &= m^T C_m^{-1} m - m^T C_m^{-1} m_0 - m_0^T C_m^{-1} m + m_0^T C_m^{-1} m_0 + m^T G^T C_d^{-1} Gm \\ &\quad - m^T G^T C_d^{-1} d_{obs} - d_{obs}^T C_d^{-1} Gm + d_{obs}^T C_d^{-1} d_{obs}, \end{aligned} \quad (3.2.9)$$

tanto el primer término como el quinto termino tienen a m^T multiplicando por la izquierda y m multiplicando por la derecha, de forma similar, tanto el segundo término como el sexto tienen multiplicado a m^T por la izquierda y finalmente el tercer término y el séptimo tienen multiplicando a m por la derecha, por lo tanto se pueden agrupar de la siguiente forma

$$\begin{aligned} 2S(m) &= 2S(m) = m^T(C_m^{-1} + G^T C_d^{-1} G)m - m^T(G^T C_d^{-1} d_{obs} + C_m^{-1} m_0) \\ &\quad - (G^T C_d^{-1} d_{obs} + C_m^{-1} m_0)^T m + m_0^T C_m^{-1} m_0 + d_{obs}^T C_d^{-1} d_{obs}. \end{aligned} \quad (3.2.10)$$

Observando la ec. 3.2.10, se puede notar que lo que se encuentra entre paréntesis en el primer término es igual a la inversa de 3.2.8, haciendo esta sustitución y definiendo $a = G^T C_d^{-1} d_{obs} + C_m^{-1} m_0$ se obtiene

$$2S(m) = m^T C_{\tilde{m}}^{-1} m - m^T a - a^T m + m_0^T C_m^{-1} m_0 + d_{obs}^T C_d^{-1} d_{obs}. \quad (3.2.11)$$

Poniendo atención ahora en (3.2.7) se puede ver que el primer paréntesis es igual a (3.2.8) y el segundo es igual a lo que se definió anteriormente como a , haciendo estas sustituciones se obtiene

$$\tilde{m} = C_{\tilde{m}} a,$$

restando m en ambos lados y multiplicando por -1

$$(\tilde{m} - m)(-1) = (C_{\tilde{m}} a - m)(-1),$$

$$m - \tilde{m} = m - C_{\tilde{m}} a$$

factorizando $C_{\tilde{m}}$

$$m - \tilde{m} = C_{\tilde{m}}(C_{\tilde{m}}^{-1} m - a). \quad (3.2.12)$$

Utilizando la ec. 3.2.12 para obtener el primer término dentro de los corchetes en 3.2.6 se obtiene

$$\begin{aligned} (m - \tilde{m})^T C_{\tilde{m}}^{-1} (m - \tilde{m}) &= (C_{\tilde{m}}(C_{\tilde{m}}^{-1} m - a))^T C_{\tilde{m}}^{-1} (C_{\tilde{m}}(C_{\tilde{m}}^{-1} m - a)) \\ &= ((C_{\tilde{m}}^{-1} m - a)^T C_{\tilde{m}}^T) C_{\tilde{m}}^{-1} (C_{\tilde{m}}(C_{\tilde{m}}^{-1} m - a)) \\ &= ((C_{\tilde{m}}^{-1} m)^T - a^T) C_{\tilde{m}} (C_{\tilde{m}}^{-1} m - a) \\ &= (m^T C_{\tilde{m}}^{-1} - a^T) C_{\tilde{m}} (C_{\tilde{m}}^{-1} m - a) \\ &= (m^T - a^T C_{\tilde{m}}) (C_{\tilde{m}}^{-1} m - a) \\ &= m^T C_{\tilde{m}}^{-1} m - m^T a - a^T m + a^T C_{\tilde{m}} a. \end{aligned} \quad (3.2.13)$$

La ec. (3.2.13) es muy similar a los primeros 3 términos del lado derecho de la igualdad en la ec. (3.2.11) a diferencia del término $a^T C_{\tilde{m}} a$, si se suma y se resta este término a (3.2.11) se obtiene

$$2S(m) = m^T C_{\tilde{m}}^{-1} m - m^T a - a^T m + m_0^T C_m^{-1} m_0 + a^T C_{\tilde{m}} a - a^T C_{\tilde{m}} a + d_{obs}^T C_d^{-1} d_{obs}. \quad (3.2.14)$$

Sustituyendo (3.2.13) en (3.2.14)

$$\begin{aligned} 2S(m) &= (m - \tilde{m})^T C_{\tilde{m}}^{-1} (m - \tilde{m}) - a^T C_{\tilde{m}} a + m_0^T C_m^{-1} m_0 + d_{obs}^T C_d^{-1} d_{obs} \\ &= (m - \tilde{m})^T C_{\tilde{m}}^{-1} (m - \tilde{m}) + K, \end{aligned} \quad (3.2.15)$$

donde $K = -a^T C_{\tilde{m}} a + m_0^T C_m^{-1} m_0 + d_{obs}^T C_d^{-1} d_{obs}$.

Para llegar de K al término K_0 en (3.2.6), se sustituye el término a y se desarrolla la ecuación de manera que se obtiene

$$\begin{aligned} K &= -(G^T C_d^{-1} d_{obs} + C_m^{-1} m_0)^T C_{\tilde{m}} (G^T C_d^{-1} d_{obs} + C_m^{-1} m_0) + m_0^T C_m^{-1} m_0 + d_{obs}^T C_d^{-1} d_{obs} \\ &= -((G^T C_d^{-1} d_{obs})^T + (C_m^{-1} m_0)^T) C_{\tilde{m}} (G^T C_d^{-1} d_{obs} + C_m^{-1} m_0) + m_0^T C_m^{-1} m_0 + d_{obs}^T C_d^{-1} d_{obs} \\ &= -(d_{obs}^T C_d^{-1} G + m_0^T C_m^{-1}) C_{\tilde{m}} (G^T C_d^{-1} d_{obs} + C_m^{-1} m_0) + m_0^T C_m^{-1} m_0 + d_{obs}^T C_d^{-1} d_{obs} \\ &= -(d_{obs}^T C_d^{-1} G C_{\tilde{m}} + m_0^T C_m^{-1} C_{\tilde{m}}) (G^T C_d^{-1} d_{obs} + C_m^{-1} m_0) + m_0^T C_m^{-1} m_0 + d_{obs}^T C_d^{-1} d_{obs} \\ &= -d_{obs}^T C_d^{-1} G C_{\tilde{m}} G^T C_d^{-1} d_{obs} - d_{obs}^T C_d^{-1} G C_{\tilde{m}} C_m^{-1} m_0 - m_0^T C_m^{-1} C_{\tilde{m}} G^T C_d^{-1} d_{obs} \\ &\quad - m_0^T C_m^{-1} C_{\tilde{m}} C_m^{-1} m_0 + m_0^T C_m^{-1} m_0 + d_{obs}^T C_d^{-1} d_{obs}. \end{aligned} \quad (3.2.16)$$

tanto el primer término como el sexto termino tienen a d_{obs}^T multiplicando por la izquierda y d_{obs} multiplicando por la derecha, los términos cuarto y quinto tienen a m_0^T multiplicando por la izquierda y m_0 multiplicando por la derecha, agrupando estos términos y además reescribiendo el tercer término como su transpuesta se obtiene

$$\begin{aligned} K &= d_{obs}^T (C_d^{-1} - C_d^{-1} G C_{\tilde{m}} G^T C_d^{-1}) d_{obs} + m_0^T (C_m^{-1} - C_m^{-1} C_{\tilde{m}} C_m^{-1}) m_0 \\ &\quad - d_{obs}^T C_d^{-1} G C_{\tilde{m}} C_m^{-1} m_0 - (d_{obs}^T C_d^{-1} G C_{\tilde{m}} C_m^{-1} m_0)^T. \end{aligned} \quad (3.2.17)$$

A partir de las siguientes identidades

$$(C_d^{-1} - C_d^{-1} G C_{\tilde{m}} G^T C_d^{-1}) = (G C_m G^T + C_d)^{-1}, \quad (3.2.18)$$

$$(C_m^{-1} - C_m^{-1} C_{\tilde{m}} C_m^{-1}) = G^T (G C_m G^T + C_d)^{-1} G, \quad (3.2.19)$$

$$d_{obs}^T C_d^{-1} G C_{\tilde{m}} C_m^{-1} m_0 = d_{obs}^T (G C_m G^T + C_d)^{-1} G m_0, \quad (3.2.20)$$

se sustituye (3.2.18), (3.2.19) y (3.2.20) en (3.2.17), con lo cual se obtiene

$$\begin{aligned} K &= d_{obs}^T (G C_m G^T + C_d)^{-1} d_{obs} + m_0^T (G^T (G C_m G^T + C_d)^{-1} G) m_0 \\ &\quad - d_{obs}^T (G C_m G^T + C_d)^{-1} G m_0 - (d_{obs}^T (G C_m G^T + C_d)^{-1} G m_0)^T. \end{aligned} \quad (3.2.21)$$

Observando el tercer y cuarto término en la ec. (3.2.21) se puede identificar que uno es la transpuesta del otro, además, por la izquierda se multiplica un vector columna transpuesto (dimensión $1 \times n$) y por la derecha un vector columna ($m \times 1$), lo cual implica que el resultado siempre será una matriz de $[1 \times 1]$, entonces $(d_{obs}^T(GC_m G^T + C_d)^{-1}Gm_0) = (d_{obs}^T(GC_m G^T + C_d)^{-1}Gm_0)^T$, de manera que se obtiene

$$\begin{aligned} K &= d_{obs}^T(GC_m G^T + C_d)^{-1}d_{obs} + m_0^T(G^T(GC_m G^T + C_d)^{-1}G)m_0 \\ &- 2d_{obs}^T(GC_m G^T + C_d)^{-1}Gm_0. \end{aligned} \quad (3.2.22)$$

En primera instancia, quizá no sea fácil notar que la ec. (3.2.22) corresponde a una ecuación cuadrática matricial de la forma $[X - Y]^T \Sigma [X - Y] = X^T \Sigma X - 2X^T \Sigma Y + Y^T \Sigma Y$, donde X corresponde a Gm_0 , Σ sería $(GC_m G^T + C_d)^{-1}$ y Y es d_{obs} , por lo tanto queda como

$$\begin{aligned} K &= (Gm_0 - d_{obs})^T (GC_m G^T + C_d)^{-1} (Gm_0 - d_{obs}) \\ &= K_o. \end{aligned} \quad (3.2.23)$$

Una vez que se ha comprobado que se puede llegar de $S(m)$ como la suma de los argumentos de la exponencial de la *a priori* y la verosimilitud a la ec. (3.2.6), esta se sustituye en (3.2.5), con lo cual se obtiene la PDF posterior conjunta, que queda como

$$\sigma_{\mathbb{M}_b} = K_b \exp\left(-\frac{1}{2}(m - \tilde{m})^T C_{\tilde{m}}^{-1} (m - \tilde{m})\right), \quad (3.2.24)$$

donde $K_b = k_b \exp\left(-\frac{K_0}{2}\right)$. Cabe destacar que la ec. (3.2.24) muestra que al utilizar un distribución normal truncada multivariada como *a priori* se obtiene una posterior que también es una normal truncada multivariada.

3.3 PDF marginal

Para obtener las PDFs marginales de una distribución normal truncada se realiza una descomposición, en la que los vectores m y \tilde{m} se subdividen en dos subvectores $m = [m'_1, m'_2]^T$ y $\tilde{m} = [\tilde{m}'_1, \tilde{m}'_2]^T$, de tal manera que la matriz de covarianza posterior se subdivide en cuatro submatrices $C_{\tilde{m}} = \begin{bmatrix} C_{\tilde{m}'_{11}} & C_{\tilde{m}'_{21}} \\ C_{\tilde{m}'_{12}} & C_{\tilde{m}'_{22}} \end{bmatrix}$ y su inversa $C_{\tilde{m}}^{-1} = \begin{bmatrix} \Sigma_{\tilde{m}'_{11}} & \Sigma_{\tilde{m}'_{21}} \\ \Sigma_{\tilde{m}'_{12}} & \Sigma_{\tilde{m}'_{22}} \end{bmatrix}$.

A partir de la descomposición descrita anteriormente, se busca llegar a una relación de la siguiente forma

$$(m - \tilde{m})^T C_{\tilde{m}}^{-1} (m - \tilde{m}) = Q_1(m'_1) + Q_2(m'_1, m'_2). \quad (3.3.1)$$

Se parte de considerar que $(m - \tilde{m})^T C_{\tilde{m}}^{-1} (m - \tilde{m})$ es una función $Q(m_1, m_2)$ que a partir de la descomposición definida se expresa como

$$\begin{aligned}
Q(m_1, m_2) &= [(m'_1 - \tilde{m}'_1)^T, (m'_2 - \tilde{m}'_2)^T] \begin{bmatrix} \Sigma_{\tilde{m}'_{11}} & \Sigma_{\tilde{m}'_{12}} \\ \Sigma_{\tilde{m}'_{21}} & \Sigma_{\tilde{m}'_{22}} \end{bmatrix} \begin{bmatrix} m'_1 - \tilde{m}'_1 \\ m'_2 - \tilde{m}'_2 \end{bmatrix}, \\
&= (m'_1 - \tilde{m}'_1)^T \Sigma_{\tilde{m}'_{11}} (m'_1 - \tilde{m}'_1) + (m'_2 - \tilde{m}'_2)^T \Sigma_{\tilde{m}'_{21}} (m'_1 - \tilde{m}'_1) \\
&+ (m'_1 - \tilde{m}'_1)^T \Sigma_{\tilde{m}'_{12}} (m'_2 - \tilde{m}'_2) + (m'_2 - \tilde{m}'_2)^T \Sigma_{\tilde{m}'_{22}} (m'_2 - \tilde{m}'_2), \tag{3.3.2}
\end{aligned}$$

considerando que $C_{\tilde{m}}$ es simétrica, $C_{\tilde{m}'_{12}} = C_{\tilde{m}'_{21}}^T$ y $\Sigma_{\tilde{m}'_{12}} = \Sigma_{\tilde{m}'_{21}}^T$, por lo tanto $[(m'_2 - \tilde{m}'_2)^T \Sigma_{\tilde{m}'_{21}} (m'_1 - \tilde{m}'_1)]^T = (m'_1 - \tilde{m}'_1)^T \Sigma_{\tilde{m}'_{12}} (m'_2 - \tilde{m}'_2)$ y como $(m'_2 - \tilde{m}'_2)^T \Sigma_{\tilde{m}'_{21}} (m'_1 - \tilde{m}'_1)$ siempre va a ser un escalar, $(m'_2 - \tilde{m}'_2)^T \Sigma_{\tilde{m}'_{21}} (m'_1 - \tilde{m}'_1) + (m'_1 - \tilde{m}'_1)^T \Sigma_{\tilde{m}'_{12}} (m'_2 - \tilde{m}'_2)$ se puede agrupar como $2(m'_1 - \tilde{m}'_1)^T \Sigma_{\tilde{m}'_{12}} (m'_2 - \tilde{m}'_2)$, de manera que

$$\begin{aligned}
Q(m'_1, m'_2) &= (m'_1 - \tilde{m}'_1)^T \Sigma_{\tilde{m}'_{11}} (m'_1 - \tilde{m}'_1) + 2(m'_1 - \tilde{m}'_1)^T \Sigma_{\tilde{m}'_{12}} (m'_2 - \tilde{m}'_2) \\
&+ (m'_2 - \tilde{m}'_2)^T \Sigma_{\tilde{m}'_{22}} (m'_2 - \tilde{m}'_2). \tag{3.3.3}
\end{aligned}$$

Para la descomposición propuesta, las matrices inversas cumplen las siguientes identidades

$$\Sigma_{\tilde{m}'_{11}} = C_{\tilde{m}'_{11}}^{-1} + C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}} (C_{\tilde{m}'_{22}} - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}})^{-1} C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} \tag{3.3.4}$$

$$\Sigma_{\tilde{m}'_{22}} = (C_{\tilde{m}'_{22}} - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}})^{-1} \tag{3.3.5}$$

$$\Sigma_{\tilde{m}'_{12}} = -C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}} (C_{\tilde{m}'_{22}} - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}})^{-1} \tag{3.3.6}$$

Con la finalidad de hacer la ecuación más compacta y fácil de seguir paso a paso, se definen $\Lambda = (m'_1 - \tilde{m}'_1)$ y $\Gamma = (m'_2 - \tilde{m}'_2)$, sustituyendo (3.3.4), (3.3.5) y (3.3.6) en (3.3.3),

$$\begin{aligned}
Q(m'_1, m'_2) &= \Lambda^T (C_{\tilde{m}'_{11}}^{-1} + C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}} (C_{\tilde{m}'_{22}} - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}})^{-1} C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1}) \Lambda \\
&+ 2\Lambda^T (-C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}} (C_{\tilde{m}'_{22}} - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}})^{-1}) \Gamma \\
&+ \Gamma^T (C_{\tilde{m}'_{22}} - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}})^{-1} \Gamma \\
&= \Lambda^T C_{\tilde{m}'_{11}}^{-1} \Lambda + \Lambda^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}} (C_{\tilde{m}'_{22}} - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}})^{-1} C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} \Lambda \\
&- 2\Lambda^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}} (C_{\tilde{m}'_{22}} - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}})^{-1} \Gamma \\
&+ \Gamma^T (C_{\tilde{m}'_{22}} - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}})^{-1} \Gamma, \tag{3.3.7}
\end{aligned}$$

definiendo $A = C_{\tilde{m}'_{22}} - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}}$ y reescribiendo $\Lambda^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}}$ como $[C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} \Lambda]^T$

$$\begin{aligned}
Q(m'_1, m'_2) &= \Lambda^T C_{\tilde{m}'_{11}}^{-1} \Lambda + [C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} \Lambda]^T A^{-1} C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} \Lambda \\
&- 2[C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} \Lambda]^T A^{-1} \Gamma + \Gamma^T A^{-1} \Gamma. \tag{3.3.8}
\end{aligned}$$

Analizando la ec. (3.3.8), se puede ver que el segundo termino del primer renglón y ambos términos del segundo corresponden a una ecuación cuadrática matricial de la forma $[X - Y]^T \Sigma [X - Y] = X^T \Sigma X - 2X^T \Sigma Y + Y^T \Sigma Y$ donde X corresponde a $C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} \Lambda$, Σ sería A y Y es Γ , por lo tanto queda como

$$Q(m'_1, m'_2) = \Lambda^T C_{\tilde{m}'_{11}}^{-1} \Lambda + [\Gamma - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} \Lambda]^T A^{-1} [\Gamma - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} \Lambda], \quad (3.3.9)$$

regresando Λ y Γ a los términos que se estaban manejando inicialmente se obtiene

$$\begin{aligned} Q(m'_1, m'_2) &= (m'_1 - \tilde{m}'_1)^T C_{\tilde{m}'_{11}}^{-1} (m'_1 - \tilde{m}'_1) \\ &+ (m'_2 - b(m'_1))^T A^{-1} (m'_2 - b(m'_1)), \end{aligned} \quad (3.3.10)$$

donde

$$Q_1(m'_1) = (m'_1 - \tilde{m}'_1)^T C_{\tilde{m}'_{11}}^{-1} (m'_1 - \tilde{m}'_1)$$

$$Q_2(m'_1, m'_2) = (m'_2 - b(m'_1))^T A^{-1} (m'_2 - b(m'_1))$$

$$A = C_{\tilde{m}'_{22}} - C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} C_{\tilde{m}'_{12}}$$

$$b(m'_1) = \tilde{m}'_2 + C_{\tilde{m}'_{12}}^T C_{\tilde{m}'_{11}}^{-1} (m'_1 - \tilde{m}'_1).$$

Sustituyendo (3.3.1) en (3.2.8) se obtiene

$$\begin{aligned} \sigma_{\mathbb{M}_b} &= K_b \exp\left(-\frac{1}{2} [Q_1(m'_1) + Q_2(m'_1, m'_2)]\right), \\ &= K_b \exp\left(-\frac{1}{2} Q_1(m'_1)\right) \exp\left(-\frac{1}{2} Q_2(m'_1, m'_2)\right). \end{aligned} \quad (3.3.11)$$

Integrando (3.3.11) respecto a m_2 sobre todo \mathbb{M}_{b2} y sustituyendo $Q_1(m'_1)$ y $Q_2(m'_1, m'_2)$

$$\begin{aligned} \sigma_{\mathbb{M}_{b1}} &= K_b \exp\left(-\frac{1}{2} Q_1(m'_1)\right) \int_{\mathbb{M}_{b2}} \exp\left(-\frac{1}{2} Q_2(m'_1, m'_2)\right) dm'_2, \\ &= K_b \exp\left(-\frac{1}{2} (m'_1 - \tilde{m}'_1)^T C_{\tilde{m}'_{11}}^{-1} (m'_1 - \tilde{m}'_1)\right) \\ &\quad \int_{\mathbb{M}_{b2}} \exp\left(-\frac{1}{2} (m'_2 - b(m'_1))^T A^{-1} (m'_2 - b(m'_1))\right) dm'_2. \end{aligned} \quad (3.3.12)$$

La ec. (3.3.12) corresponde a la PDF marginal debido a que depende únicamente del o los parámetros agrupados en m'_1 . Cabe recalcar que, aunque tiene una forma bastante similar a la de una PDF normal truncada, no lo es debido a que la integral no es una constante, ya que involucra al término b que depende de m'_1 . Este término se conoce como *asimetría*.

3.3.1 Ejemplo descomposición 1D

El procedimiento que se mostró anteriormente es válido para obtener marginales de cualquier dimensión (desde 1D hasta $(p - 1)$ D) y para cualquier parámetro o parámetros, sin embargo, la forma de definir los subvectores y las submatrices necesarias de m , \tilde{m} y $C_{\tilde{m}}$ puede no ser trivial.

Con la finalidad de que la forma de realizar la descomposición sea lo más clara posible, se plantea un ejemplo en el que se supone un problema con 5 parámetros (es decir $\dim(\mathbb{M}_b) = 5$), por lo tanto m y \tilde{m} son vectores de $[5 \times 1]$ y $C_{\tilde{m}}$ es una matriz de $[5 \times 5]$ como se muestra a continuación

$$m = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \end{bmatrix}, \tilde{m} = \begin{bmatrix} \tilde{m}_1 \\ \tilde{m}_2 \\ \tilde{m}_3 \\ \tilde{m}_4 \\ \tilde{m}_5 \end{bmatrix}, C_{\tilde{m}} = \begin{bmatrix} C_{\tilde{m}_{11}} & C_{\tilde{m}_{12}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{21}} & C_{\tilde{m}_{22}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{31}} & C_{\tilde{m}_{32}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{41}} & C_{\tilde{m}_{42}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \\ C_{\tilde{m}_{51}} & C_{\tilde{m}_{52}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \end{bmatrix}.$$

Suponiendo que se busca obtener la PDF marginal unidimensional del parámetro m_3 los subvectores m'_1 y \tilde{m}'_1 van a ser de $[1 \times 1]$ y su componente será aquella asociada con el parámetro de interés, mientras que en m'_2 y \tilde{m}'_2 se van a agrupar el resto de los parámetros de la siguiente forma

$$m'_1 = [m_3], m'_2 = \begin{bmatrix} m_1 \\ m_2 \\ m_4 \\ m_5 \end{bmatrix}, \tilde{m}'_1 = [\tilde{m}_3], \tilde{m}'_2 = \begin{bmatrix} \tilde{m}_1 \\ \tilde{m}_2 \\ \tilde{m}_4 \\ \tilde{m}_5 \end{bmatrix}.$$

Para definir las submatrices de la matriz de covarianza posterior $C_{\tilde{m}}$ es conveniente hacer un paso previo, en el que tanto la fila como la columna asociadas con el parámetro m_3 se mueven a la posición de la fila y columna del primer parámetro de la siguiente forma

$$C_{\tilde{m}} = \begin{bmatrix} C_{\tilde{m}_{11}} & C_{\tilde{m}_{12}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{21}} & C_{\tilde{m}_{22}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{31}} & C_{\tilde{m}_{32}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{41}} & C_{\tilde{m}_{42}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \\ C_{\tilde{m}_{51}} & C_{\tilde{m}_{52}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \end{bmatrix} R1 \longleftrightarrow R3 \begin{bmatrix} C_{\tilde{m}_{31}} & C_{\tilde{m}_{32}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{21}} & C_{\tilde{m}_{22}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{11}} & C_{\tilde{m}_{12}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{41}} & C_{\tilde{m}_{42}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \\ C_{\tilde{m}_{51}} & C_{\tilde{m}_{52}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \end{bmatrix}$$

$$\begin{bmatrix} C_{\tilde{m}_{31}} & C_{\tilde{m}_{32}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{21}} & C_{\tilde{m}_{22}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{11}} & C_{\tilde{m}_{12}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{41}} & C_{\tilde{m}_{42}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \\ C_{\tilde{m}_{51}} & C_{\tilde{m}_{52}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \end{bmatrix} C1 \longleftrightarrow C3 \begin{bmatrix} C_{\tilde{m}_{33}} & C_{\tilde{m}_{32}} & C_{\tilde{m}_{31}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{23}} & C_{\tilde{m}_{22}} & C_{\tilde{m}_{21}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{13}} & C_{\tilde{m}_{12}} & C_{\tilde{m}_{11}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{43}} & C_{\tilde{m}_{42}} & C_{\tilde{m}_{41}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \\ C_{\tilde{m}_{53}} & C_{\tilde{m}_{52}} & C_{\tilde{m}_{51}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \end{bmatrix}$$

Una vez realizado el cambio de renglones y columnas correspondiente, se definen las submatrices necesarias. En este caso $C_{\tilde{m}'_1}$ es una matriz de $[1 \times 1]$, cuyo elemento

corresponde al de la primer fila y la primer columna de la matriz $C_{\tilde{m}}$ reordenada. La matriz $C_{\tilde{m}'_{12}}$ es de dimensión $[1 \times 4]$ y corresponde al renglón 1 exceptuando al primer elemento. De forma similar, la matriz $C_{\tilde{m}'_{21}}$ es de dimensión $[4 \times 1]$ y corresponde a la primer columna exceptuando también al primer elemento. Finalmente, la matriz $C_{\tilde{m}'_{22}}$ es de dimensión $[4 \times 4]$ y corresponde a la matriz de covarianza posterior reordenada sin considerar el primer renglón ni la primer columna. Las submatrices para este ejemplo se muestran a continuación

$$C_{\tilde{m}'_{11}} = [C_{\tilde{m}_{33}}], C_{\tilde{m}'_{12}} = \begin{bmatrix} C_{\tilde{m}_{32}} \\ C_{\tilde{m}_{31}} \\ C_{\tilde{m}_{34}} \\ C_{\tilde{m}_{35}} \end{bmatrix}, C_{\tilde{m}'_{21}} = \begin{bmatrix} C_{\tilde{m}_{23}} \\ C_{\tilde{m}_{13}} \\ C_{\tilde{m}_{43}} \\ C_{\tilde{m}_{53}} \end{bmatrix}, C_{\tilde{m}'_{22}} = \begin{bmatrix} C_{\tilde{m}_{22}} & C_{\tilde{m}_{21}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{12}} & C_{\tilde{m}_{11}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{42}} & C_{\tilde{m}_{41}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \\ C_{\tilde{m}_{52}} & C_{\tilde{m}_{51}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \end{bmatrix}.$$

Para ilustrar visualmente qué elementos de la matriz $C_{\tilde{m}}$ corresponden a cada submatriz, se hace una asociación de color, en la que a la matriz $C_{\tilde{m}'_{11}}$ se le asocia el color azul, a $C_{\tilde{m}'_{12}}$ se le asocia el color violeta, a $C_{\tilde{m}'_{21}}$ se le asocia el color magenta y a $C_{\tilde{m}'_{22}}$ se le asocia el color rojo, los elementos correspondientes son los siguientes

$$C_{\tilde{m}} = \begin{bmatrix} C_{\tilde{m}_{33}} & C_{\tilde{m}_{32}} & C_{\tilde{m}_{31}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{23}} & C_{\tilde{m}_{22}} & C_{\tilde{m}_{21}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{13}} & C_{\tilde{m}_{12}} & C_{\tilde{m}_{11}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{43}} & C_{\tilde{m}_{42}} & C_{\tilde{m}_{41}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \\ C_{\tilde{m}_{53}} & C_{\tilde{m}_{52}} & C_{\tilde{m}_{51}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \end{bmatrix}.$$

3.3.2 Ejemplo descomposición 2D

Aprovechando el planteamiento del ejemplo anterior, se va a suponer el mismo problema, solo que en esta ocasión se busca obtener la marginal bidimensional de los parámetros m_2 y m_5 .

Los subvectores m'_1 y \tilde{m}'_1 quedan de $[2 \times 1]$ con los elementos asociados a los parámetros de interés, mientras que de nuevo el resto se agrupan en m'_2 y \tilde{m}'_2 de la siguiente forma

$$m'_1 = \begin{bmatrix} m_2 \\ m_5 \end{bmatrix}, m'_2 = \begin{bmatrix} m_1 \\ m_3 \\ m_4 \end{bmatrix}, \tilde{m}'_1 = \begin{bmatrix} \tilde{m}_2 \\ \tilde{m}_5 \end{bmatrix}, \tilde{m}'_2 = \begin{bmatrix} \tilde{m}_1 \\ \tilde{m}_3 \\ \tilde{m}_4 \end{bmatrix}.$$

La matriz de covarianza posterior también se reordena para poder definir las submatrices de manera más sencilla, para lo cual la fila y la columna asociadas con el parámetro m_2 se mueven a la posición del parámetro m_1 y la fila y la columna asociadas con el parámetro m_5 se mueven a la posición del parámetro m_2 , de manera que la matriz $C_{\tilde{m}}$ reordenada queda de la siguiente forma

$$C_{\tilde{m}} = \begin{bmatrix} C_{\tilde{m}_{11}} & C_{\tilde{m}_{12}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{21}} & C_{\tilde{m}_{22}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{31}} & C_{\tilde{m}_{32}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{41}} & C_{\tilde{m}_{42}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \\ C_{\tilde{m}_{51}} & C_{\tilde{m}_{52}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \end{bmatrix} R1 \longleftrightarrow R2 \begin{bmatrix} C_{\tilde{m}_{21}} & C_{\tilde{m}_{22}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{11}} & C_{\tilde{m}_{12}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{31}} & C_{\tilde{m}_{32}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{41}} & C_{\tilde{m}_{42}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \\ C_{\tilde{m}_{51}} & C_{\tilde{m}_{52}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \end{bmatrix}$$

$$\begin{bmatrix} C_{\tilde{m}_{21}} & C_{\tilde{m}_{22}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{11}} & C_{\tilde{m}_{12}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{31}} & C_{\tilde{m}_{32}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{41}} & C_{\tilde{m}_{42}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \\ C_{\tilde{m}_{51}} & C_{\tilde{m}_{52}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \end{bmatrix} R5 \longrightarrow R2 \begin{bmatrix} C_{\tilde{m}_{21}} & C_{\tilde{m}_{22}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{51}} & C_{\tilde{m}_{52}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \\ C_{\tilde{m}_{11}} & C_{\tilde{m}_{12}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{31}} & C_{\tilde{m}_{32}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{41}} & C_{\tilde{m}_{42}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \end{bmatrix}$$

$$\begin{bmatrix} C_{\tilde{m}_{21}} & C_{\tilde{m}_{22}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{51}} & C_{\tilde{m}_{52}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \\ C_{\tilde{m}_{11}} & C_{\tilde{m}_{12}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{31}} & C_{\tilde{m}_{32}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{41}} & C_{\tilde{m}_{42}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \end{bmatrix} C1 \longleftrightarrow C2 \begin{bmatrix} C_{\tilde{m}_{22}} & C_{\tilde{m}_{21}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{52}} & C_{\tilde{m}_{51}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \\ C_{\tilde{m}_{12}} & C_{\tilde{m}_{11}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{32}} & C_{\tilde{m}_{31}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{42}} & C_{\tilde{m}_{41}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \end{bmatrix}$$

$$\begin{bmatrix} C_{\tilde{m}_{22}} & C_{\tilde{m}_{21}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{52}} & C_{\tilde{m}_{51}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} & C_{\tilde{m}_{55}} \\ C_{\tilde{m}_{12}} & C_{\tilde{m}_{11}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{32}} & C_{\tilde{m}_{31}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{42}} & C_{\tilde{m}_{41}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} & C_{\tilde{m}_{45}} \end{bmatrix} C5 \longrightarrow C2 \begin{bmatrix} C_{\tilde{m}_{22}} & C_{\tilde{m}_{25}} & C_{\tilde{m}_{21}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} \\ C_{\tilde{m}_{52}} & C_{\tilde{m}_{55}} & C_{\tilde{m}_{51}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} \\ C_{\tilde{m}_{12}} & C_{\tilde{m}_{15}} & C_{\tilde{m}_{11}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} \\ C_{\tilde{m}_{32}} & C_{\tilde{m}_{35}} & C_{\tilde{m}_{31}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} \\ C_{\tilde{m}_{42}} & C_{\tilde{m}_{45}} & C_{\tilde{m}_{41}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} \end{bmatrix}$$

De esta matriz $C_{\tilde{m}}$ se van a determinar las submatrices. $C_{\tilde{m}'_{11}}$ va a quedar de dimensión $[2 \times 2]$ y corresponde a los dos primeros elementos de las primeras dos primeras filas y los dos primeros elementos de las dos primeras columnas. La submatriz $C_{\tilde{m}'_{12}}$ va a quedar de dimensión $[2 \times 3]$ y corresponde a las primeras dos filas sin los elementos que pertenecen a las dos primeras columnas. La submatriz $C_{\tilde{m}'_{21}}$ va a quedar de dimensión $[3 \times 2]$ y corresponde a las primeras dos columnas sin los elementos que pertenecen a las dos primeras filas. La submatriz $C_{\tilde{m}'_{22}}$ va a quedar de dimensión $[3 \times 3]$ y corresponde a los elementos que están fuera de las dos primeras filas y las dos primeras columnas. Las submatrices se muestran a continuación

$$C_{\tilde{m}'_{11}} = \begin{bmatrix} C_{\tilde{m}_{22}} & C_{\tilde{m}_{25}} \\ C_{\tilde{m}_{52}} & C_{\tilde{m}_{55}} \end{bmatrix}, C_{\tilde{m}'_{12}}^T = \begin{bmatrix} C_{\tilde{m}_{21}} & C_{\tilde{m}_{51}} \\ C_{\tilde{m}_{23}} & C_{\tilde{m}_{53}} \\ C_{\tilde{m}_{24}} & C_{\tilde{m}_{54}} \end{bmatrix} C_{\tilde{m}'_{21}} = \begin{bmatrix} C_{\tilde{m}_{12}} & C_{\tilde{m}_{15}} \\ C_{\tilde{m}_{32}} & C_{\tilde{m}_{35}} \\ C_{\tilde{m}_{42}} & C_{\tilde{m}_{45}} \end{bmatrix} C_{\tilde{m}'_{22}} = \begin{bmatrix} C_{\tilde{m}_{11}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} \\ C_{\tilde{m}_{31}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} \\ C_{\tilde{m}_{41}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} \end{bmatrix}$$

De la misma forma que el caso 1D se ilustra visualmente qué elementos pertenecen a cada una de las submatrices, nuevamente se asigna el color azul a $C_{\tilde{m}'_{11}}$, el color violeta a $C_{\tilde{m}'_{12}}$, el color magenta a $C_{\tilde{m}'_{21}}$ y el color rojo a $C_{\tilde{m}'_{22}}$ de tal forma que los colores quedan de la siguiente forma

$$C_{\tilde{m}} = \begin{bmatrix} C_{\tilde{m}_{22}} & C_{\tilde{m}_{25}} & C_{\tilde{m}_{21}} & C_{\tilde{m}_{23}} & C_{\tilde{m}_{24}} \\ C_{\tilde{m}_{52}} & C_{\tilde{m}_{55}} & C_{\tilde{m}_{51}} & C_{\tilde{m}_{53}} & C_{\tilde{m}_{54}} \\ C_{\tilde{m}_{12}} & C_{\tilde{m}_{15}} & C_{\tilde{m}_{11}} & C_{\tilde{m}_{13}} & C_{\tilde{m}_{14}} \\ C_{\tilde{m}_{32}} & C_{\tilde{m}_{35}} & C_{\tilde{m}_{31}} & C_{\tilde{m}_{33}} & C_{\tilde{m}_{34}} \\ C_{\tilde{m}_{42}} & C_{\tilde{m}_{45}} & C_{\tilde{m}_{41}} & C_{\tilde{m}_{43}} & C_{\tilde{m}_{44}} \end{bmatrix}.$$

3.4 Aproximación de la apriori

Como se mencionó en la Sección 3.2, la estrategia propuesta por Nocquet (2018) considera que la distribución *a priori* es una PDF normal truncada multivariada, sin embargo, al no tener información adicional, la distribución más adecuada sería la uniforme.

Tomando en cuenta lo anterior, lo que se busca es modelar una uniforme mediante una normal truncada sobre el intervalo de truncamiento. Para esto, se empieza por parametrizar la desviación estándar en términos de una variable α de la siguiente forma

$$\sigma = \alpha \cdot \max\left(\frac{b_i - a_i}{2}\right), \quad (3.4.1)$$

donde a_i es el límite inferior y b_i es el límite superior del intervalo de truncamiento del i -ésimo parámetro.

El valor de la desviación estándar está asociado con el ancho de la PDF, mientras mayor sea la magnitud de la desviación, el ancho de la normal también será mayor y por lo tanto va a tomar una forma cada vez más plana.

En la ec. (3.4.1) se puede notar que α es proporcional a σ , esto implica que al incrementar el valor de α , la normal se va asemejar cada vez más a una uniforme como se muestra en la Fig. 3.1.

Matemáticamente, conforme $\alpha \rightarrow \infty$, la normal truncada va a tender a la uniforme en el intervalo de truncamiento, por lo tanto, entre mayor sea el valor de α , se obtendría una mejor aproximación, no obstante, el incrementar demasiado este valor induce errores numéricos. Lo ideal es buscar un valor de α tal que se logre un equilibrio entre la estabilidad numérica y una aproximación aceptable.

Para el caso de una PDF normal multivariada, la matriz de covarianza de los parámetros se define a partir de la parametrización mostrada en (3.4.1) multiplicada por la identidad de la siguiente forma

$$C_m = \sigma^2 I_p. \quad (3.4.2)$$

3.5 Media, varianza y coeficiente de variación

En la Sección 3.3 se describe la forma de obtener distribuciones marginales, no obstante el objetivo principal es adquirir información cuantitativa de cada parámetro y, en el caso

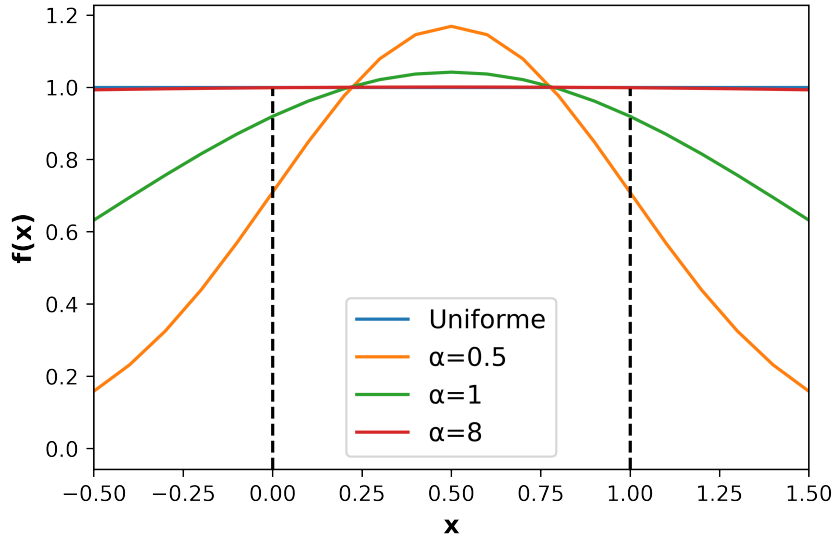


Figura 3.1: Aproximación a una distribución uniforme mediante distribuciones normales truncadas en función de α .

particular de este trabajo, se busca estimar la incertidumbre, para lo cual se requiere determinar la media y la varianza de las PDFs.

La forma que se emplea para cuantificar la incertidumbre es mediante el coeficiente de variación, C_V , que se define de la siguiente forma

$$C_V = \frac{\sigma}{\mu} \cdot 100. \tag{3.5.1}$$

Se considera que el C_V es un buen indicador de la incertidumbre debido a que la desviación estándar, σ , esta relacionada con qué tanto se acercan los posibles valores de la distribución al más probable.

Mientras mayor sea el valor de la desviación estándar, hay más dispersión y, por lo tanto, una mayor incertidumbre. Al estar normalizada respecto a la media, μ , y multiplicada por 100, se obtiene un valor adimensional porcentual, lo cual permite evitar manejar una magnitud subjetiva.

Para calcular μ y σ se utiliza la definición del primer y segundo momento de una función aleatoria, que se definen de la siguiente forma

$$\mu = \int_{-\infty}^{\infty} x f(x) dx, \tag{3.5.2}$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx. \tag{3.5.3}$$

Debido a que la ec. (3.3.3) (que correspondería a $f(x)$ en (3.5.2) y (3.5.3)) no es fácil de integrar analíticamente, se puede evaluar la marginal y posteriormente realizar la integral numérica.

Capítulo 4

Casos de estudio

4.1 Caso sintético

La metodología que se describe en el Capítulo anterior se utilizará para resolver un problema sencillo de inferencia bayesiana, propuesto por [Nocquet \(2018\)](#). Este caso supone una falla con dos componentes de deslizamiento, lo cual implica que el problema

incluye dos parámetros, la matriz del modelo es $G = \begin{bmatrix} -7 & -4 \\ 1 & 10 \\ 2 & -11 \end{bmatrix}$. Se supone también

una estación GPS que registra 3 componentes del desplazamiento, las cuales se incluyen en el vector $d = \begin{bmatrix} 10 \\ 3 \\ -5 \end{bmatrix}$ y la matriz de covarianza de los datos se define como $C_d = 5^2 I_3$.

El intervalo de truncamiento se define de 0 a 1 para ambos parámetros, es decir $\mathbb{M}_b = [0, 1] \times [0, 1]$.

Para aproximar una *a priori* uniforme se utiliza un valor de $\alpha = 8$, usando la ec. (3.4.1) y posteriormente con la ec. (3.4.2) se obtiene la matriz de covarianza de los parámetros C_m .

Además de resolver el ejemplo con la estrategia semianalítica, se resuelve también con el método de MCMC, explicado en el Capítulo 2, con la finalidad de hacer una comparación entre ambos métodos.

Para considerar el hecho de que el espacio de modelos está truncado, el muestreo se realiza tomando en cuenta en qué posición se encuentra la muestra dentro de \mathbb{M}_b y la perturbación se obtiene de una normal truncada cuyos límites se establecen de tal manera que no se pueda generar una propuesta fuera de los límites para cada parámetro. Se genera un total de 500000 muestras y a continuación se hace un quemado de 1000. En la Fig. 4.1 se muestra la PDF (*Probability Density Function*) posterior conjunta obtenida.

Posteriormente se obtienen las marginales para ambos parámetros, de igual forma empleando ambos métodos; para la estrategia semianalítica se utiliza el procedimiento que se describe en la Sección 3.3, mientras que para MCMC se hace una discretización del espacio de modelos y se hace un conteo del número de muestras que entran en cada subespacio, luego se hace la normalización correspondiente para garantizar que sea una

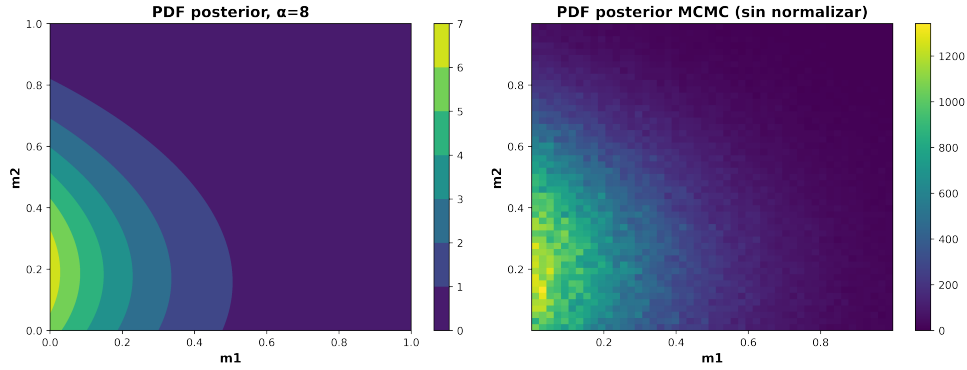


Figura 4.1: PDF posterior conjunta calculada mediante la estrategia semianalítica (izquierda) y con MCMC (derecha).

PDF.

Las marginales obtenidas se muestran en la Fig. 4.2.

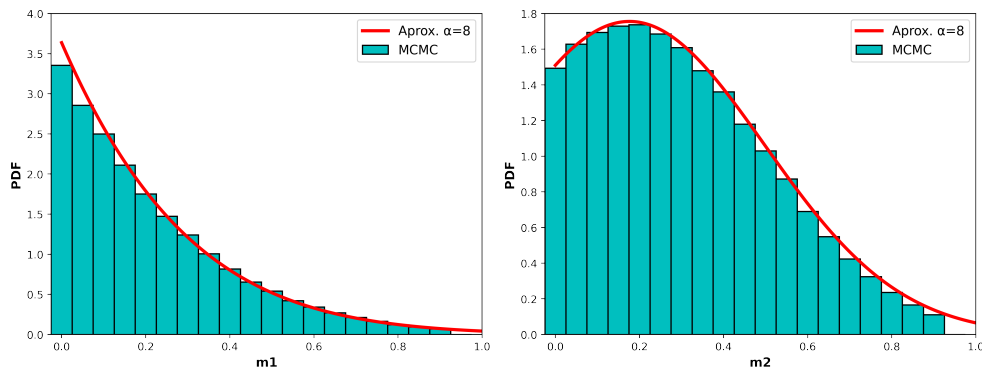


Figura 4.2: PDFs marginales obtenidas con ambos métodos para el parámetro m_1 (izquierda) y para el parámetro m_2 (derecha).

Se registró el tiempo de ejecución en una CPU con un procesador intel Core i3-6006U con 4 Gb en RAM. En el caso de MCMC, para determinar las 500000 muestras de la posterior tomó 240.12 s, mientras que para obtener la posterior conjunta y las marginales de ambos parámetros con la estrategia semianalítica de [Nocquet \(2018\)](#) se llevó 4.97 s.

La diferencia en el tiempo de ejecución para ambos métodos es considerable aún para dos variables y, aunque el algoritmo empleado para obtener los resultados con MCMC no esta optimizado, al trabajar con un mayor número de parámetros sería mucho más practico utilizar la estrategia semianalítica.

La media y la desviación estándar se muestran en la tabla 4.1, así como el coeficiente de variación calculado usando la ec. 3.5.1.

Método	Media (m_1/m_2)	Sd (m_1/m_2)	C_V % (m_1/m_2)
Nocquet	0.229/0.328	0.200/0.219	87.33/66.77
MCMC	0.234/0.332	0.200/0.217	85.47/65.36

Tabla 4.1: Media, desviación estándar (Sd) σ y coeficiente de variación C_V de los parámetros m_1 y m_2 del ejemplo sintético.

4.2 Caso real

Los datos utilizados corresponden al sismo de Acapulco del 7 de septiembre de 2021, magnitud 7 Mw. Se hace una discretización del plano de falla de 18 por 18, con lo cual se obtiene un total de 324 subfallas cuadradas de 5 km por lado. Se considera que se conoce la dirección del deslizamiento (a partir del mecanismo focal del sismo) y, por lo tanto, se invierte únicamente la magnitud para cada subfalla, de manera que se trabaja con un total de 324 parámetros.

El intervalo de truncamiento para todos los parámetros se define de 0 a 4.9315 m . El deslizamiento máximo que se permite en la solución se puede estimar a partir del momento sísmico escalar mediante la ecuación

$$M_0 = \mu S \bar{D}, \quad (4.2.1)$$

donde M_0 es el momento escalar, μ es el módulo de rigidez de la roca en la ruptura, S es el área de la ruptura y \bar{D} es el deslizamiento promedio que se usa para definir el rango de los parámetros.

El momento sísmico estático se puede determinar a partir de registros sísmicos usando el espectro de desplazamiento de campo lejano debido a que, a bajas frecuencias, la amplitud de dicho espectro presenta un comportamiento plano cuyo valor, Ω_0 , es proporcional a M_0 (ver Fig. 4.3) (Shearer, 2009).

El factor de proporcionalidad entre M_0 y Ω_0 depende de ciertos aspectos como la densidad del medio, la velocidad de la onda, la distancia desde la fuente y el patrón de radiación. La relación entre el momento sísmico estático y la magnitud de momento, M_W , se da mediante la siguiente ecuación

$$M_W = \frac{\log M_0}{1.5} - 10.73. \quad (4.2.2)$$

Los datos se obtienen de 11 estaciones que registran 3 componentes de desplazamiento (norte, este y vertical), dando un total de 33 datos, los cuales se ordenan en un vector columna, d_{obs} , colocando las 3 componentes de una estación, posteriormente las 3 componentes de la siguiente estación y así sucesivamente hasta completar los datos registrados por todas las estaciones.

La matriz de covarianza, C_d , se define como una matriz identidad de $[33 \times 33]$, cuyas columnas se multiplican elemento a elemento por un vector columna de $[33 \times 1]$ conformado por las varianzas de cada componente de las estaciones GPS, siguiendo el mismo orden del vector de datos.

Se considera que todas las estaciones tienen las mismas varianzas para sus 3 componentes, es decir 2.1 km^2 , 2.5 km^2 y 5.1 km^2 para las componentes norte, este y

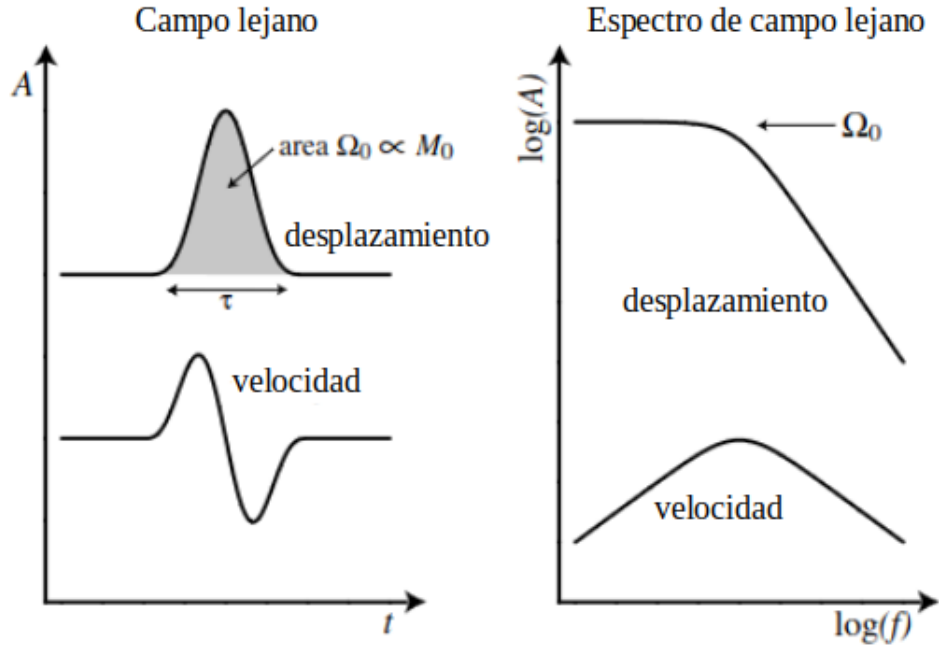


Figura 4.3: Serie de tiempo del desplazamiento y la velocidad de un pulso de duración τ en campo lejano (izquierda) y sus espectros (derecha). Imagen modificada de (Shearer, 2009).

vertical respectivamente, por lo que el vector de varianzas se construye repitiendo 11 veces esta terna de valores.

Tomando en cuenta que el vector de parámetros, m , es de dimensión $[324 \times 1]$ y el vector de datos, d_{obs} , es de dimensión $[33 \times 1]$, la matriz G será de dimensión $[33 \times 324]$ de manera que se tenga una multiplicación conformable de acuerdo con la ec. 3.1.3.

4.2.1 Metodologías propuestas

El cálculo de las marginales mediante la ec. (3.3.12) requiere de la resolución de una gran cantidad de integrales de funciones multivariadas, esto debido a que, para cada valor de m'_1 , se va a tener un valor distinto de $b(m'_1)$ y por lo tanto una integral distinta que calcular. En otras palabras, para cada punto del recorrido de la marginal m'_1 , será necesario calcular una integral de la forma $\exp(\frac{1}{2}(m-\mu)^T \Sigma^{-1}(m-\mu))$, también conocida como integral gaussiana.

Considerando que la solución analítica de integrales de funciones de este tipo no es trivial, la elección obvia sería optar por algún método numérico. No obstante dado el gran número de parámetros que se están manejando, no resulta una tarea sencilla.

A continuación se muestran metodologías propuestas para obtener las marginales de cada parámetro y posteriormente los C_V .

Factor de asimetría despreciable

Como primera alternativa, se siguió la propuesta de [Nocquet \(2018\)](#) que utilizó la función *mvnun* incluida en la librería *SciPy*, que permite estimar la integral numérica de una PDF normal multivariada. Sin embargo, su aplicación para problemas de altas dimensiones, como el que quiere resolver para el sismo de Acapulco, implica dificultades numéricas que no habían sido discutidas por [Nocquet \(2018\)](#). A continuación se explican las estrategias exploradas cuando se trabajan con altas dimensiones.

Las funciones que se necesitan integrar, aunque son muy similares, no son normales multivariadas. La diferencia es el factor de normalización que multiplica la exponencial, que tiene la siguiente forma

$$K = \frac{1}{\sqrt{((2\pi)^n |C|)}}, \quad (4.2.3)$$

donde $|C|$ es el determinante de la matriz de covarianza de la distribución normal multivariada y n la dimensión. Dado que el valor que se necesita calcular no incluye este factor, es necesario calcularlo para dividirlo al resultado obtenido de *mvnun*.

Como se necesita el determinante de la matriz de covarianza para calcular el factor (4.2.3), es conveniente obtener el logaritmo natural de las ecuaciones debido a que, al ser matrices de dimensión grande, dicho determinante se puede volver tan grande que Python no puede manejarlo (tomándolo como *inf*), lo cual no sucede con el logaritmo. Se ejemplifica este procedimiento aplicado en el cálculo del término k_b , empezando por definir el resultado que se obtiene al usar la función *mvnun*

$$I_f = \frac{1}{\sqrt{((2\pi)^n |C_{\tilde{m}}|)}} \int_{\mathbb{M}_b} \exp\left(-\frac{1}{2}(m - \tilde{m})^T C_{\tilde{m}}^{-1}(m - \tilde{m})\right) dm, \quad (4.2.4)$$

despejando la integral se obtiene

$$\int_{\mathbb{M}_b} \exp\left(-\frac{1}{2}(m - \tilde{m})^T C_{\tilde{m}}^{-1}(m - \tilde{m})\right) dm = I_f \sqrt{((2\pi)^n |C_{\tilde{m}}|)}. \quad (4.2.5)$$

Ahora, escribiendo la ecuación para k_b con $S(m)$ en términos de \tilde{m} y $C_{\tilde{m}}$

$$\begin{aligned} k_b &= \frac{1}{\int_{\mathbb{M}_b} \exp\left(-\frac{1}{2}\left[(m - \tilde{m})^T C_{\tilde{m}}^{-1}(m - \tilde{m}) + K_0\right]\right) dm}, \\ &= \frac{1}{\exp\left(-\frac{K_0}{2}\right) \int_{\mathbb{M}_b} \exp\left(-\frac{1}{2}(m - \tilde{m})^T C_{\tilde{m}}^{-1}(m - \tilde{m})\right) dm}. \end{aligned} \quad (4.2.6)$$

Sustituyendo (4.2.5) en (4.2.6) se obtiene

$$k_b = \frac{1}{\exp\left(-\frac{K_0}{2}\right) I_f \sqrt{((2\pi)^n |C_{\tilde{m}}|)}}. \quad (4.2.7)$$

Aplicando logaritmo natural a ambos lados de (4.2.7) y utilizando las propiedades de los logaritmos

$$\begin{aligned}
 \ln(k_b) &= \ln \left(\frac{1}{\exp\left(-\frac{K_0}{2}\right) I_f \sqrt{((2\pi)^n |C_{\tilde{m}}|)}} \right), \\
 &= \ln(1) - \ln \left(\exp\left(-\frac{K_0}{2}\right) I_f \sqrt{((2\pi)^n |C_{\tilde{m}}|)} \right), \\
 &= - \left[\ln \left(\exp\left(-\frac{K_0}{2}\right) \right) + \ln(I_f) + \ln \left(\sqrt{((2\pi)^n |C_{\tilde{m}}|)} \right) \right], \\
 &= - \left[-\frac{K_0}{2} + \ln(I_f) + \frac{1}{2} \ln((2\pi)^n |C_{\tilde{m}}|) \right], \\
 &= \frac{K_0}{2} - \ln(I_f) - \frac{1}{2} \ln((2\pi)^n) - \frac{1}{2} \ln(|C_{\tilde{m}}|). \tag{4.2.8}
 \end{aligned}$$

Como se puede observar, la ec. (4.2.8) ahora depende de I_f y del logaritmo natural del determinante de la matriz $C_{\tilde{m}}$, el cual se puede calcular usando la función *linalg.slogdet* de *numpy*, lo cual permite trabajar con un número mayor de parámetros sin inconvenientes en este aspecto.

Desafortunadamente, al incrementar el valor de α buscando aproximar la *a priori* a una uniforme de manera satisfactoria, la integral de la normal multivariada para la cantidad de parámetros que se esta manejando se vuelve demasiado pequeña y la función *mvnun* proporciona valores inestables, es decir, con cada corrida usando la misma configuración se obtienen valores diferentes.

Para evaluar qué tanto varían los valores de la integral numérica y desde qué valor de α inicia el comportamiento inestable, se hizo una prueba que consiste en utilizar valores de α entre 0.1 y 3 a cada 0.1 (es decir 30 valores en total) para determinar \tilde{m} y $C_{\tilde{m}}$. Para cada valor se calcula la integral numérica 20 veces.

Teniendo las 20 muestras de la integral numérica para un valor de α , se normalizan respecto al orden de magnitud del resultado más pequeño y se calcula la desviación estándar. El procedimiento se repite para cada uno de los 30 valores de α .

Los resultados obtenidos se muestran en la Fig. 4.4. Como se puede observar, para $\alpha = 1$ el orden de magnitud de la desviación estándar es de decenas. Tener tanta variación en la integral numérica no es aceptable debido a que existe una gran incertidumbre de que el valor sea correcto.

Dada la complicación de obtener el valor de las integrales con *mvnun*, se optó en primera instancia por considerar que el factor de asimetría $b(m_1)$ es despreciable. Esto implica que la integral en la ec. (3.3.12) sea una constante, de modo que ahora la marginal sí se trata de una normal truncada. Lo anterior simplifica el cálculo de la marginal porque se reduce a calcular la integral de una curva 1D para obtener el factor de normalización, lo cual puede hacerse utilizando métodos numéricos convencionales.

En la Fig. 4.5 se muestran los C_V obtenidos mediante la metodología descrita en el párrafo anterior, los cuales son un indicio del grado de incertidumbre del valor de deslizamiento invertido para cada subfalla. Se utilizó un valor de $\alpha = 3$ como propuesta inicial y los valores de C_V fueron normalizados y se muestran de manera porcentual.

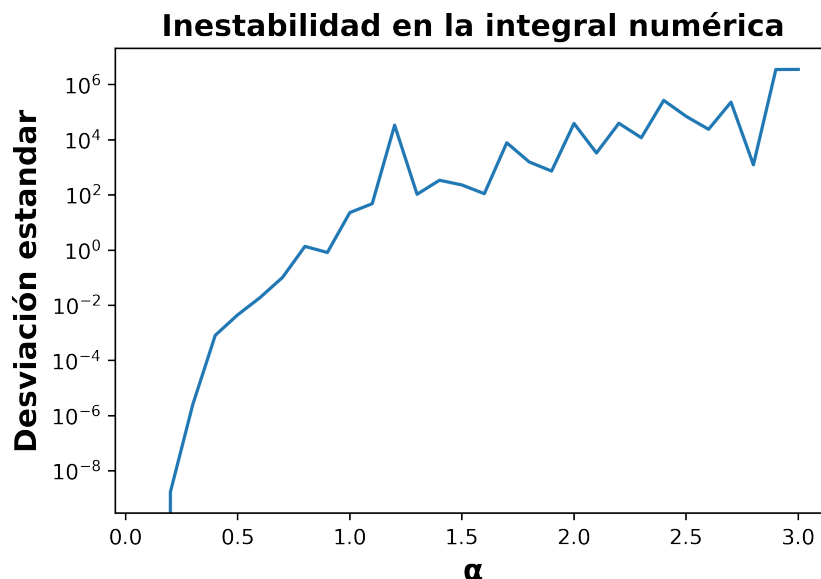


Figura 4.4: Variación del resultado de la función *munun* para diferentes valores de α .

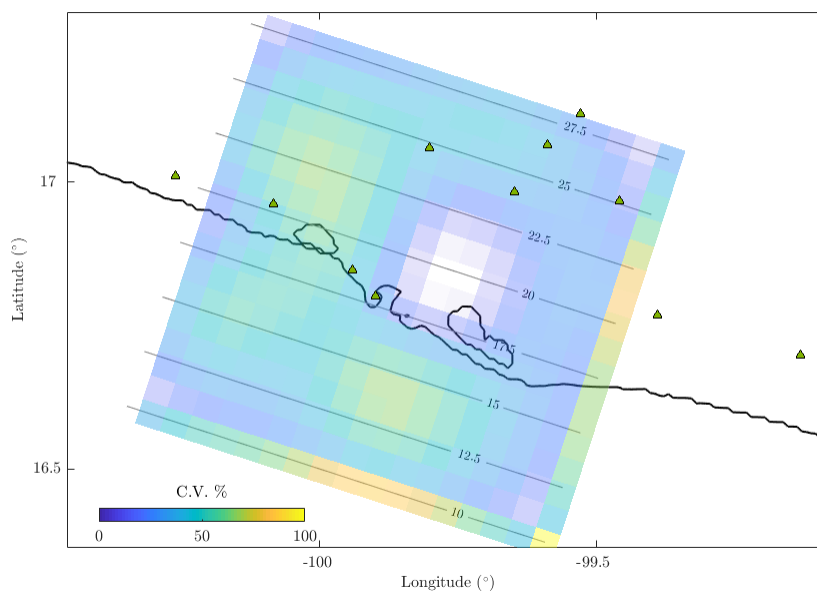


Figura 4.5: Coeficientes de variación para el deslizamiento invertido de cada subfalla obtenidos a partir de la aproximación despreciando el término de asimetría, $\alpha = 3$ (los valores en cero se resaltan en color blanco).

Simulación de la posterior

Dada la complicación del cálculo de las integrales que dependen del factor de asimetría, se explora otra alternativa para obtener las marginales mediante la simulación de la posterior.

La ventaja de esta metodología es que, una vez teniendo las muestras de la posterior, obtener las marginales es inmediato, ya que sólo se deben tomar las componentes del parámetro correspondiente y estimar la media y la desviación estándar muestral para calcular el C_V .

Considerando que se tiene la ventaja de conocer el vector de media y la matriz de covarianza de la distribución posterior conjunta, es decir, \tilde{m} y $C_{\tilde{m}}$ existe la posibilidad de simular muestras de una distribución normal multivariada y posteriormente tomar las muestras que caen en la región de truncamiento.

Desafortunadamente, no es eficiente simular una normal truncada multivariada a partir de una normal multivariada haciendo aceptación y rechazo de muestras para altas dimensiones, esto debido a que para grandes valores de dimensión n , la mayor parte de la masa de la distribución normal estándar se encuentra en las colas (Giraud, 2021), razón por la cual se obtienen bajas tasas de aceptación, especialmente cuando la región de truncamiento se encuentra cerca de la media.

La forma en que la masa en la campana se reduce conforme aumenta la dimensión de una normal se puede ver claramente en la Fig. 4.6, en la que se observa que a partir de 15 dimensiones aproximadamente, la mitad de la masa se encuentra en las colas.

Masa en la campana

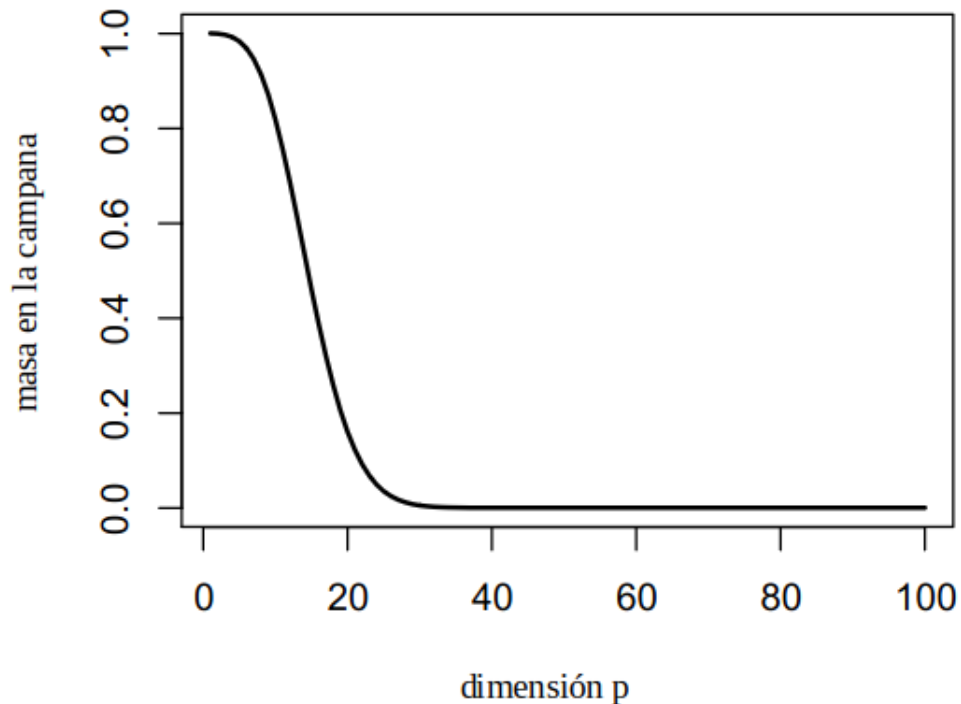


Figura 4.6: Acumulación de masa en la campana de una gaussiana contra número de dimensiones. Imagen modificada de (Giraud, 2021) .

Existen artículos en los que se simula directamente la normal truncada multivariada como el de Montesinos-López, Capella, Christen, y Tago (2023) publicado reciente-

mente. Para este trabajo se recibió la valiosa colaboración del M. en C. Isaías Manuel Ramírez Bañales, quien desarrolló un algoritmo que permite generar muestras de la posterior mediante muestreo de Gibbs (Robert, 1995).

Realizar un muestreo de Gibbs consiste en proponer un valor inicial para cada uno de los parámetros, posteriormente se genera una nueva propuesta de manera aleatoria únicamente para el primero de los parámetros, se utiliza un criterio de aceptación o rechazo para decidir si la propuesta será el valor del primer parámetro o se conserva el anterior y luego de esto se genera una propuesta para el segundo parámetro, dejando el resto como están y realizando nuevamente un criterio de aceptación o rechazo, esto se repite hasta llegar al último parámetro, lo cual completa una iteración del algoritmo.

Al muestrear la posterior es conveniente analizar la matriz de covarianza de los datos, esto debido a que la forma en la que se explora el espacio de parámetros mediante el muestreo de Gibbs puede ser ineficiente para ciertos casos. Sin embargo, observando la Fig. 4.7 se puede apreciar que se tiene poca dependencia debido a que los valores fuera de la diagonal principal (covarianzas) son pequeñas respecto a los valores de la diagonal principal (varianzas) que son similares entre sí, lo cual es indicio de dicha clase de muestreo funciona bien para este problema.

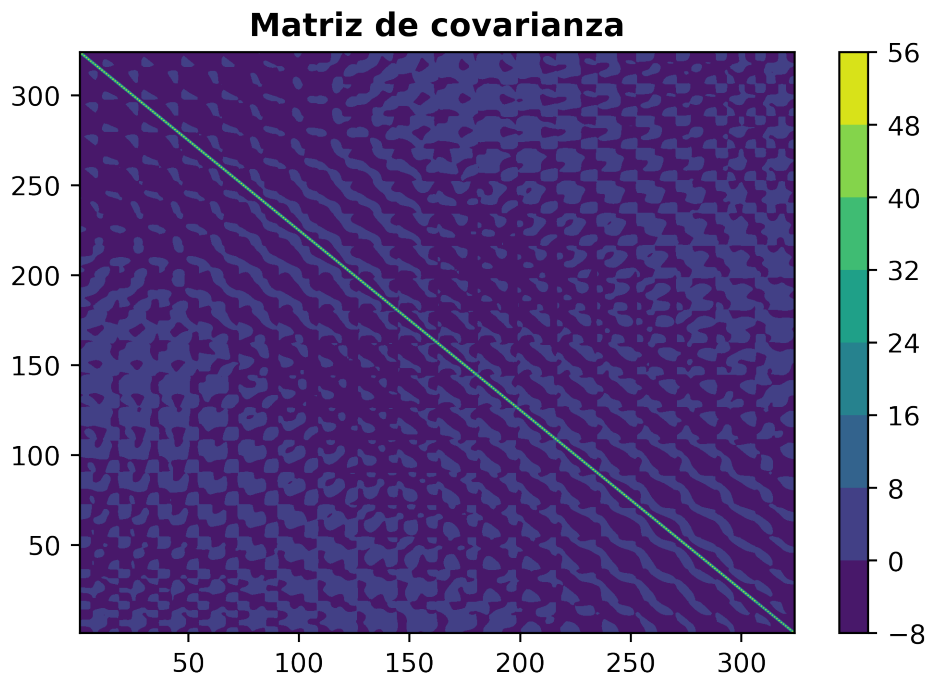


Figura 4.7: Matriz de covarianza de la posterior, $\alpha = 3$.

Tomando en cuenta que el muestreo de Gibbs se trata de una MCMC cuya distribución ergódica es una TMVN *Truncated Multivariate Normal*, es necesario realizar un quemado de muestras y hacer una selección cada cierto número de muestras, lo cual se conoce como *lag*.

El descarte de muestras que se hace a partir del *lag* antes mencionado se debe a que, dado que el proceso de muestreo implica modificar el valor de un parámetro a la

vez, se requiere haber recorrido las propuestas para todos los parámetros para llegar a una muestra completamente actualizada, debido a lo anterior, el *lag* debe ser mayor o igual al número de parámetros buscando reducir la correlación entre las muestras.

El procedimiento descrito anteriormente se realiza observando la gráfica de la log PDF de las muestras e identificando a partir de qué número de muestra se aprecia un comportamiento estable para definir el quemado y el *lag* se define tomando como referencia el tiempo de correlación integrado o IAT (por sus siglas en inglés *Integrated Correlation Time*) y, como se menciona anteriormente, que sea un valor mayor o igual a la dimensión del problema.

Como consecuencia de lo anterior, del número de muestras total solo se obtiene una pequeña fracción, debido a que de dicho total se resta el quemado y se divide entre el *lag*. Esto representa una desventaja debido a que para obtener un número de muestras aceptable para estimar los C_V , se necesitan arreglos sumamente grandes que pueden ser difíciles de manejar y tardados de estimar.

En la Fig. 4.8 se muestran los coeficientes de variación a partir de la simulación de la posterior utilizando un valor de $\alpha = 3$.

De las muestras simuladas se obtuvo un IAT de 264.26, al ser menor que el número de parámetros se toma este como referencia y se da un margen de error, por lo que el *lag* se toma como 400. Para el quemado de las muestras se desechan las primeras 150 obteniendo un total de 3448.

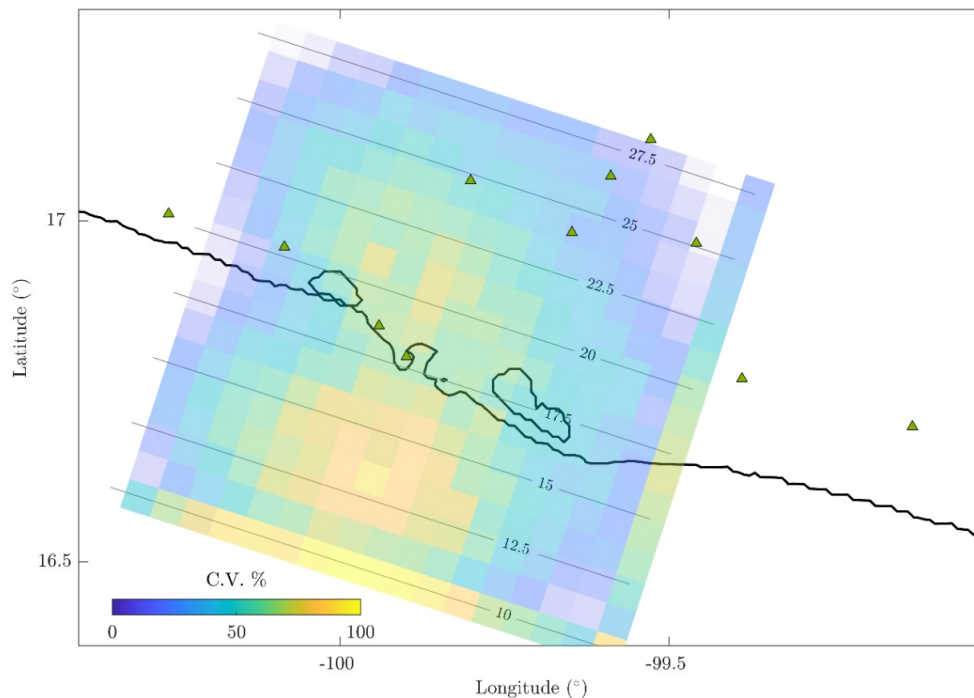


Figura 4.8: Coeficientes de variación para el deslizamiento invertido de cada subfalla a partir de la simulación de la posterior, $\alpha = 3$ (los valores en cero se resaltan en color blanco).

Integración de Monte Carlo y cuasi-Monte Carlo

Buscando alternativas para resolver numéricamente las integrales que se necesitan para calcular las marginales, quizá las primeras opciones que llegan a la mente son métodos como el de cuadratura gaussiana o la regla del trapecio, sin embargo, considerando que las funciones que se deben integrar son de gran dimensión, los métodos basados en funciones de interpolación resultan inconvenientes debido a que el número de puntos que se requiere evaluar crece exponencialmente conforme aumenta el número de dimensiones. En la literatura este problema se conoce como *curse of dimensionality* (Wang y Fang, 2003).

Por otro lado, el método de Monte Carlo se usa frecuentemente en la integración de funciones multivariadas debido a que se trata de un método simple y robusto en el que la precisión sólo depende de la complejidad de la función (Cafisch, 1998).

Para formular al estimador de Monte Carlo se utiliza la ley del estadístico inconsciente (Schervish y DeGroot, 2012), la cual establece que la esperanza de una función $f(X)$ se obtiene mediante la siguiente ecuación

$$E|f(X)| = \int_{-\infty}^{\infty} f(x)g_X(x)dx, \quad (4.2.9)$$

donde X es una variable aleatoria continua y $g_X(x)$ es una función de densidad de probabilidad conocida. Si definimos que X es una variable aleatoria continua con distribución uniforme, es decir $X \sim U(a, b)$, la ec. 4.2.9 queda como

$$E|f(X)| = \int_a^b f(x)\frac{1}{b-a}dx. \quad (4.2.10)$$

Por otra parte, de la Ley fuerte de los grandes números, la esperanza $E|f(x)|$ se puede escribir como

$$E|f(X)| \approx \frac{1}{N} \sum_{i=1}^N f(X_i). \quad (4.2.11)$$

Igualando la ec. (4.2.10) con (4.2.11) y despejando la integral

$$\int_a^b f(x)\frac{1}{b-a}dx \approx \frac{1}{N} \sum_{i=1}^N f(X_i).$$

$$\int_a^b f(x)dx \approx \frac{b-a}{N} \sum_{i=1}^N f(X_i). \quad (4.2.12)$$

Desafortunadamente, este no es un método infalible debido a que la integral converge con una tasa de $O(N^{-1/2})$, lo cual puede ser sumamente lento. Para subsanar la lentitud de la convergencia, se ha propuesto una alternativa que consiste en utilizar secuencias de números cuasi-aleatorios (también conocidas como secuencias de baja discrepancia) como muestra, que, a diferencia de números pseudo-aleatorios, no tratan de asemejarse al comportamiento de secuencias aleatorias (Cafisch, 1998).

Debido a que las muestras de una secuencia cuasi-aleatoria están correlacionadas de tal forma que se distribuyen de manera más uniforme que secuencias aleatorias o

pseudo-aleatorias, el método cuasi-Monte Carlo virtualmente converge más rápidamente, a un tasa de $O(N^{-1}(\log N)^k)$, para alguna constante k . No obstante, el hecho de que los números de la secuencia se obtienen de forma determinista hace que no se pueda obtener su varianza, lo cuál dificulta el cálculo del error.

Al momento de evaluar el desempeño de QMC se observó que al utilizar una misma secuencia de números cuasi-aleatorios para estimar cada una de las integrales que se necesitan en cada punto del dominio de la marginal, se obtienen resultados suaves. Por el contrario, al utilizar una secuencia distinta para cada punto, se obtienen resultados con aspecto ruidoso.

Para ilustrar lo anterior, se muestra un ejemplo en el que se obtiene la marginal m_1 de un problema de 150 variables, el cual se define tomando la sub-matriz de los primeros 150 por 150 valores de la matriz de covarianza y el sub-vector de los primeros 150 valores del vector de media de la posterior obtenida con los datos del caso real que se está trabajando.

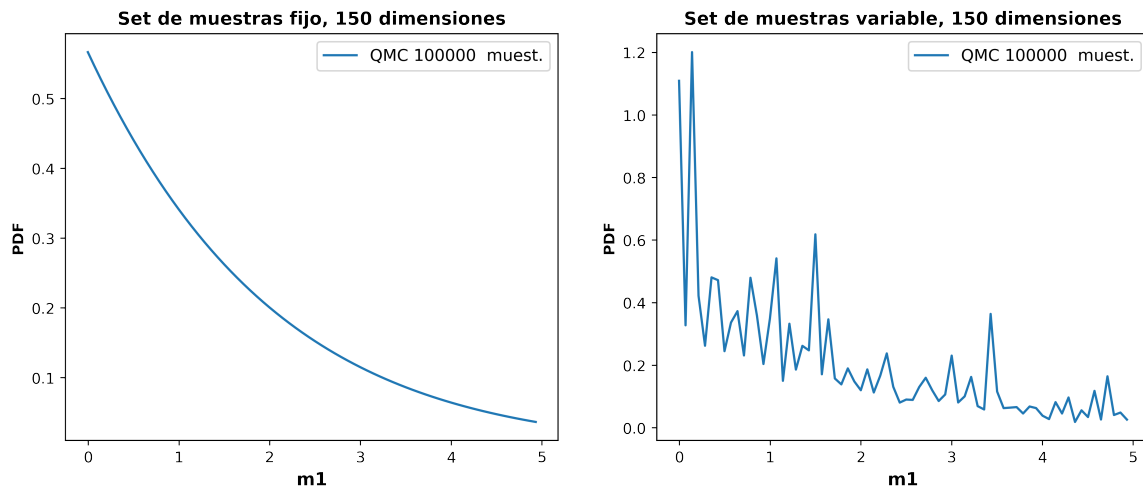


Figura 4.9: Comparación entre marginales usando una secuencia de datos fija contra una variable con QMC, $\alpha = 3$.

Como se puede observar en la Fig. 4.9, al calcular la marginal usando una secuencia diferente para cada m_1 , se obtienen variaciones que provienen del error de la estimación de cada integral en cada punto, lo cual no ocurre al utilizar una misma secuencia para todos los puntos.

Para verificar que este resultado no haya sido fortuito, se realiza el mismo ejercicio calculando 5 veces la misma marginal, cambiando la secuencia de muestras de una marginal a otra en el caso fijo.

Se puede ver claramente en la Fig. 4.10 que, aunque se alcanza a distinguir una pequeña diferencia entre cada marginal calculada utilizando una secuencia fija, la forma de todas las curvas es sumamente similar, lo cual revela que además de obtener resultados suaves también son robustos.

Con la finalidad de evaluar ahora la exactitud de los resultados, se comparan marginales obtenidas resolviendo las integrales usando tanto *mvnun* como con QMC utilizando una secuencia fija. Las comparaciones se hacen resolviendo problemas con distinto

Caso real

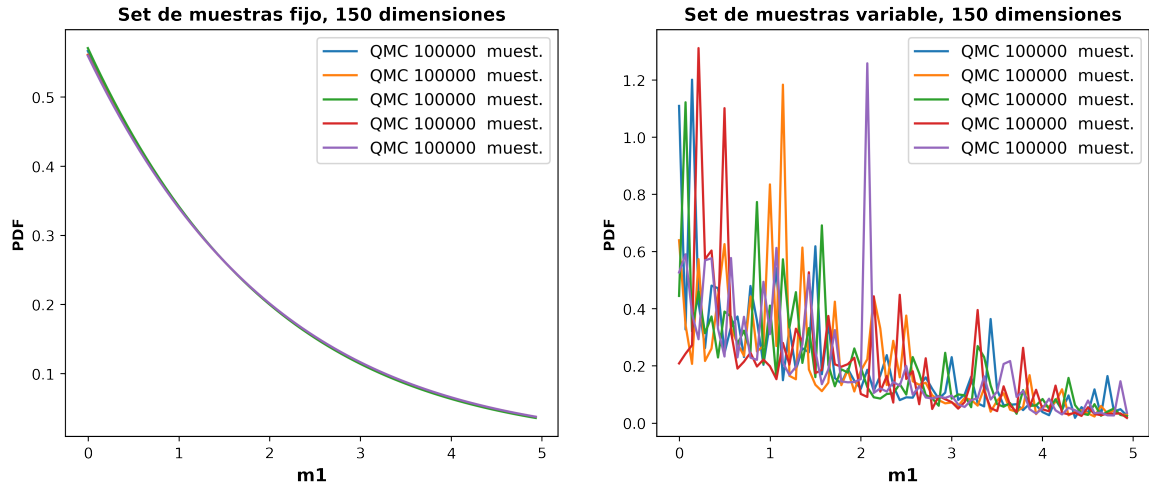


Figura 4.10: Prueba de robustez, secuencia de datos fija contra una variable con QMC, $\alpha = 3$.

número de dimensiones, los cuales se definen de forma similar a la que se explico anteriormente para el de 150 dimensiones.

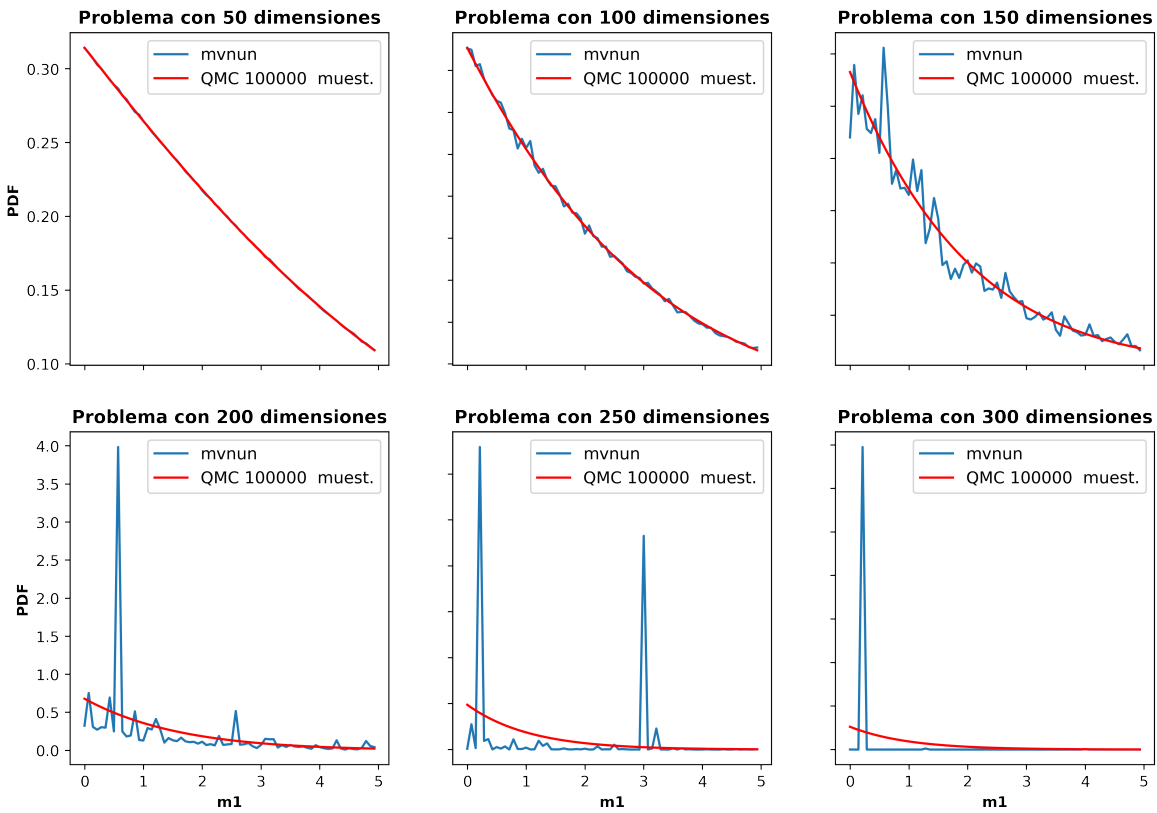


Figura 4.11: Comparación de resultados obtenidos con mvnun contra QMC usando una secuencia fija, $\alpha = 3$.

La Fig. 4.11 muestra que para 50 dimensiones, las marginales obtenidas con ambos métodos son prácticamente idénticas. Para 100 dimensiones la marginal obtenida con *mvnun* comienza a presentar un aspecto ligeramente ruidoso, lo cual es más notorio para 150 dimensiones. Aun así, resulta claro que para 100 como para 150 dimensiones las marginales con QMC se ajustan a la forma que tienen los resultados de *mvnun* pero con un aspecto suave.

La comparación para 200, 250 y 300 no tiene sentido puesto que los resultados obtenidos con *mvnun* son demasiado ruidosos; sin embargo cabe señalar que en estos casos los resultados con QMC mantienen un aspecto suave.

La hipótesis es que, al calcular las integrales que se evalúan utilizando una misma secuencia, estas conllevan un error asociado a dicha secuencia, de tal forma que

$$\int_{\mathbb{M}_{b_2}} f_I(m'_1) dm'_2 = \frac{\mathbb{M}_{b_2}}{n} \sum_{i=1}^n f_I(X_i) + e(\{X_j : 1, \dots, n\}), \quad (4.2.13)$$

donde

$$f_I(m'_1) = \left(-\frac{1}{2} (m'_2 - b(m'_1))^T A^{-1} (m'_2 - b(m'_1)) \right). \quad (4.2.14)$$

Se puede notar fácilmente que para cada valor de m'_1 se tendrá una integral de la forma de la ec. (4.2.13) diferente. No obstante, suponiendo que el error depende principalmente de la secuencia que se utiliza para estimar la integral, i.e. $e(\{X_j : 1, \dots, n\}) \perp\!\!\!\perp m'_1$, al usar la misma secuencia se tiene un error similar de una integral a otra. De modo que estaríamos sumando el mismo error para todo valor de m'_1 , tal que las diferencias relativas de las integrales entre distintos valores de m'_1 se conservan. Por lo tanto, este error aditivo se anula al momento de normalizar la distribución marginal.

A partir de las observaciones anteriores el siguiente paso es evaluar el efecto del tamaño de la muestra que se utiliza para estimar las integrales. Para esto se trabaja con el problema completo, es decir, con los 324 parámetros y para fines ilustrativos se muestra el resultado obtenido del parámetro 1.

Dado que aparentemente el utilizar valores más grandes de α no genera inestabilidad en QMC, se decide utilizar $\alpha = 5$ usando una discretización de 50 elementos para el vector de recorrido de la integral. Para evaluar los distintos tamaños de muestra se obtiene un secuencia de Sobol de 2^{15} , la razón por la que se trabaja con este tipo de secuencia es que de acuerdo a [Palluotto, Dumont, Rodrigues, Gicquel, y Vicquelin \(2019\)](#), esta presenta mejores propiedades de convergencia para integrales de altas dimensiones.

A partir de la secuencia descrita anteriormente se van tomando submuestras cuyos tamaños son potencias de 2 desde 2^3 , esto debido a que es un requisito para que se cumplan las propiedades de dicho tipo de secuencia.

El resultado obtenido de la prueba descrita anteriormente se muestra en la Fig. 4.12. El primer aspecto destacable es que, a pesar de que se puede notar una diferencia entre las integrales obtenidas con 2^3 muestras y con 2^{15} , la forma de ambas es considerablemente similar. El segundo aspecto es que al incrementar el tamaño de muestra, el resultado converge a una forma determinada. Finalmente, el tercer aspecto a destacar es que el valor de C_V calculado a partir de las integrales varía aproximadamente un

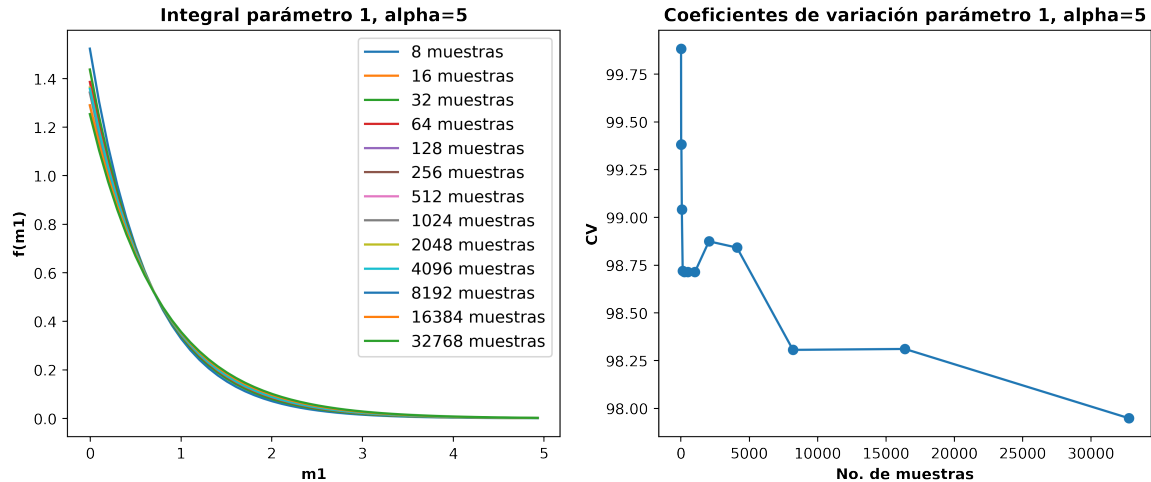


Figura 4.12: Análisis del efecto del tamaño de muestra, parámetro 1 $\alpha = 5$.

2% entre los valores extremos, lo cuál, para fines prácticos, se puede considerar una variación pequeña, tomando en cuenta también los tamaños de muestra.

Se puede identificar entonces que para utilizar QMC en la estrategia de [Nocquet \(2018\)](#) se requiere determinar 3 parámetros fundamentales; el valor de α , la discretización del vector de recorrido de las marginales y el número de muestras utilizado para estimar las integrales.

El primer parámetro que se analiza para obtener los C_V del problema con el que se está trabajando es α , la razón de esto es que se busca aprovechar el hecho de que se puede explorar de manera rápida dicho valor con un número de muestras pequeño, considerando que se obtiene una buena aproximación como se observa en la Fig. 4.12. Por lo tanto se usa una secuencia de Sobol de 2^4 muestras y se define una discretización de 50 para el vector en el que se evalúan las marginales. Se utilizan 12 valores enteros desde 1 hasta 12 y se calculan los C_V obtenidos con cada configuración.

En la Fig. 4.13 se puede ver que desde un valor de $\alpha = 5$ se obtiene un resultado estable, comportamiento favorable debido a que teóricamente se obtiene cuando la *a priori* es lo suficientemente plana como para simular una uniforme en la región de truncamiento, no obstante se decide utilizar un valor ligeramente mayor de $\alpha = 8$, aprovechando que no hay indicios de que al usar un valor mayor se ocasione inestabilidad.

El segundo parámetro que se evalúa es la discretización del vector con el que se evalúan las marginales, aprovechando nuevamente que se puede obtener una buena idea de la forma de la integral aún con un número pequeño de muestras. Una vez más se utilizan 2^4 muestras y el valor de $\alpha = 8$ que se determinó anteriormente. Se utilizan valores de discretización desde 20 hasta 180 a cada 20 muestras.

Los resultados obtenidos se muestran en la Fig. 4.14. A partir de una discretización de 80 se puede apreciar un cambio más sutil hasta 180 respecto a las primeras 3 gráficas. Aunque de 140 a 180 es más difícil notar una diferencia, se considera que una discretización de 100 es adecuada debido a que, por un lado, aunque hay una diferencia perceptible esta no es muy grande y, por otro lado, la diferencia en el tiempo

Coefficientes de variación, discretización 50, 16 muestras

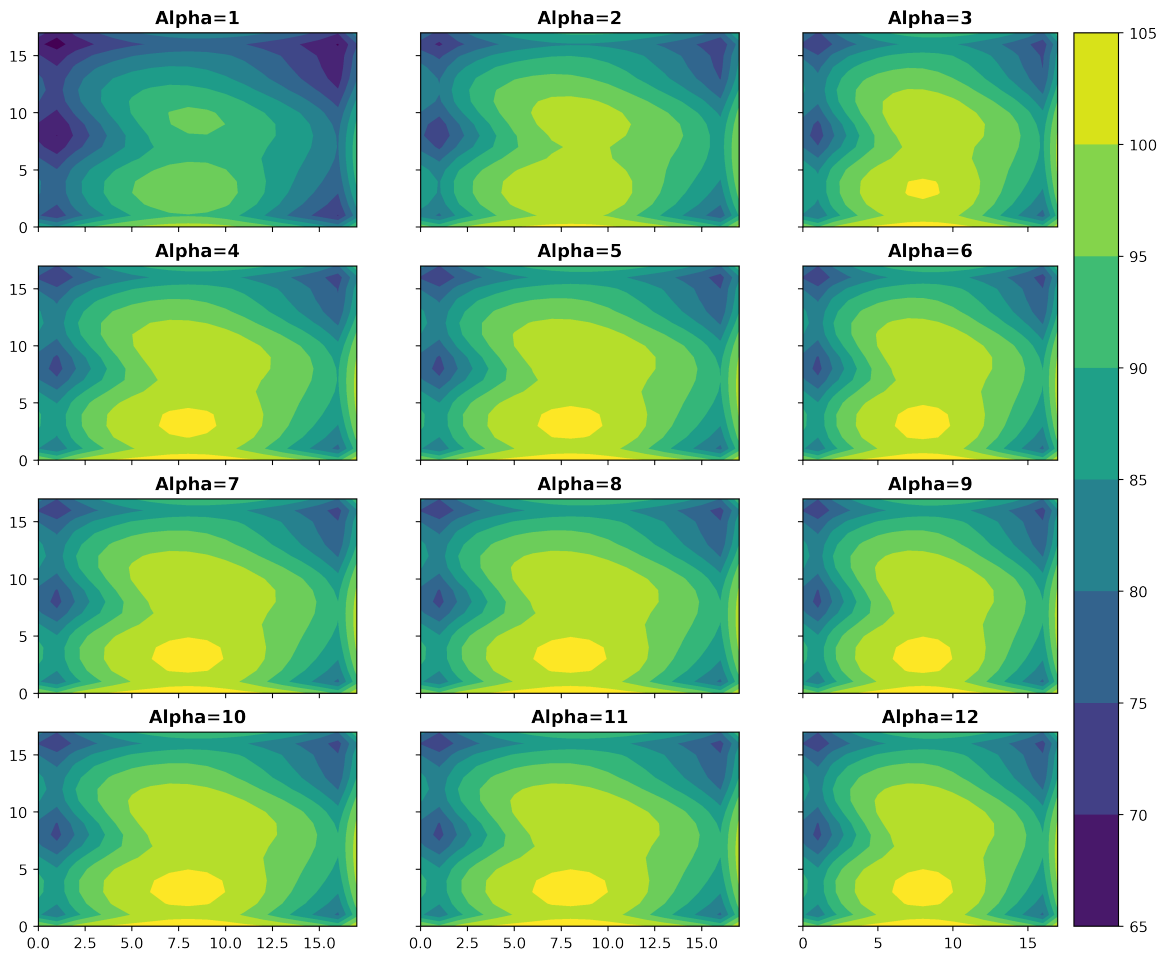


Figura 4.13: Análisis del efecto del valor de α , discretización 50, 16 muestras.

de computo de las integrales adicionales al escoger una discretización más fina puede ser significativa.

Finalmente para estimar las integrales de las marginales se utilizan 2^{20} muestras obtenidas a partir de una secuencia de Sobol. Este número de muestras se seleccionó buscando que fuera lo mayor posible antes de que se vuelva demasiado pesado computacionalmente, tanto para obtener la secuencia como evaluar las muestras para estimar cada integral.

Los C_V obtenidos de la configuración descrita anteriormente se muestran respecto a las estaciones en la Fig. 4.15.

4.2.2 Evaluación de coeficientes de variación

Tomando en cuenta la exploración de α que se hizo mediante QMC, se calculan los C_V con un valor de $\alpha = 8$ con las tres estrategias propuestas para comparar los resultados.

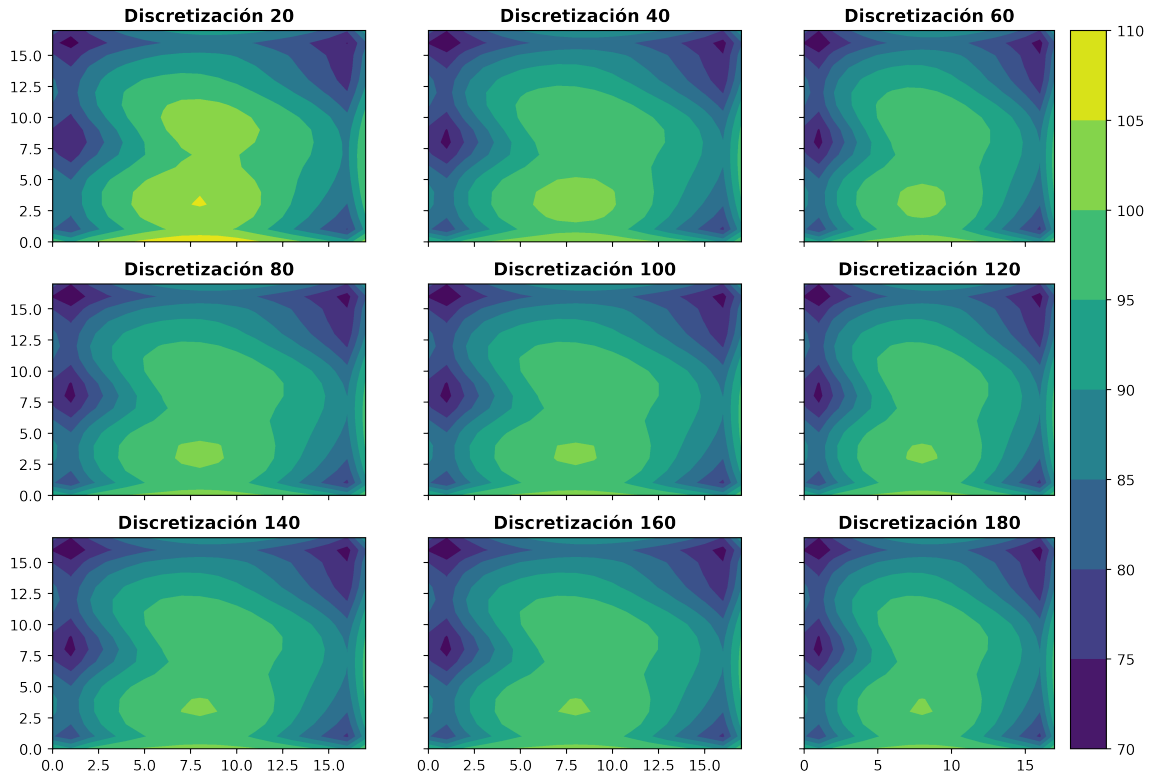
Coeficientes de variación, $\alpha=8$, 16 muestras

Figura 4.14: Análisis del efecto de la discretización del vector de recorrido de la marginal, $\alpha = 8$, 16 muestras.

Para muestrear la posterior, se simularon 1 400 000 muestras de las cuales se obtuvo un IAT de 311.041. Por lo tanto, para definir el *lag* se toma como referencia el número de parámetros (324) y se da un margen de error, por lo que se define un valor de *lag* de 400. Para el quemado de las muestras se desechan las primeras 150 dando un total de 3500 muestras para estimar los C_V .

Los resultados de los 3 métodos comparados están normalizados y definidos en un intervalo de cero a 100. Dicha comparación se muestra en la Fig. 4.16.

En términos generales, se puede notar que el perímetro de las 3 gráficas es considerablemente similar.

Para el caso de la aproximación en la que se desprecia el factor de asimetría, se observa una zona de valores bajos de C_V ligeramente a la derecha del centro de la gráfica que no se observa en las otras dos.

Por otro lado, los resultados obtenidos mediante QMC y de la simulación presentan un parecido significativo entre sí, no obstante, se puede apreciar que se obtiene un resultado menos uniforme de la simulación, lo cual se debe a que el número de muestras no es suficiente para un muestrear adecuadamente de la posterior.

En la Fig. 4.17 se presenta la comparación de los resultados obtenidos con las 3 estrategias evaluadas, esta vez mostrando la posición de la falla discretizada sobre un

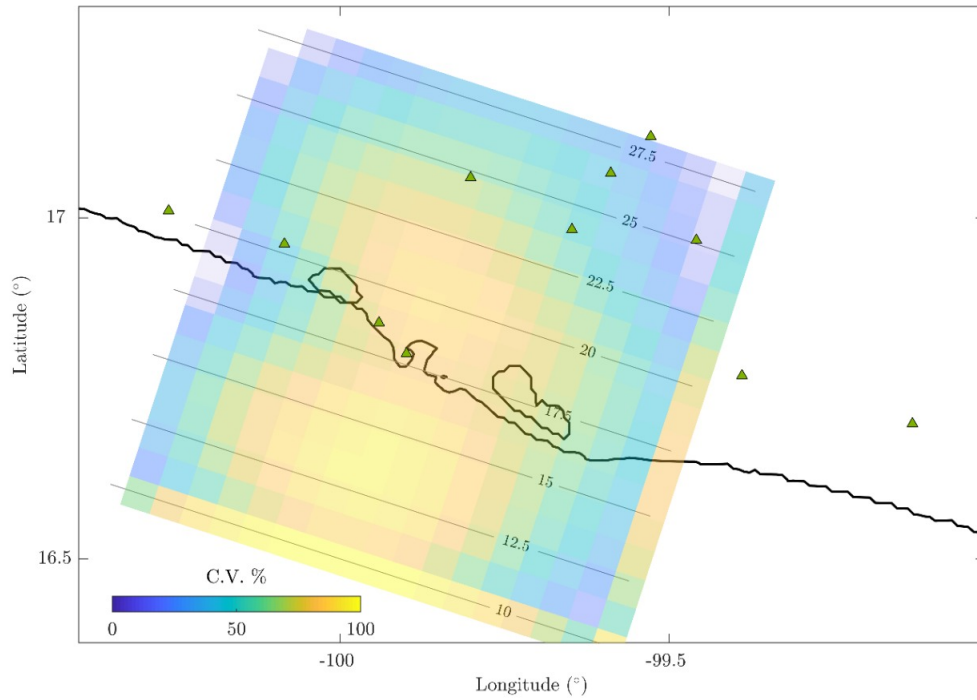


Figura 4.15: Coeficientes de variación para el deslizamiento invertido de cada subfalla utilizando QMC, $\alpha = 8$ (los valores en cero se resaltan en color blanco).

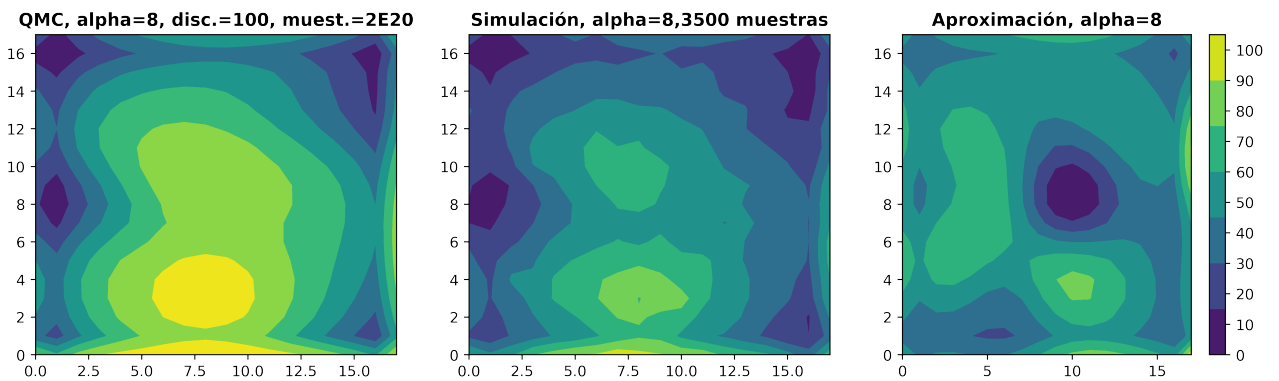


Figura 4.16: Comparación entre los métodos revisados, $\alpha = 8$.

mapa en el que también se muestra la posición de las estaciones GPS, con el propósito de analizar los valores de C_V de cada subfalla respecto a dichas estaciones.

Es posible ver en los primeros dos mapas correspondientes al resultado obtenido de la simulación y con QMC respectivamente, que hacia la esquina superior derecha de la falla se tienen los valores más bajos de C_V , lo cual coincide con la zona en la que se tiene una mayor cobertura de estaciones.

Se puede observar también que, hacia la parte media de ambos mapas mencionados anteriormente, la zona en la que se encuentran las estaciones más cercanas a la línea

Alternativas al coeficiente de variación

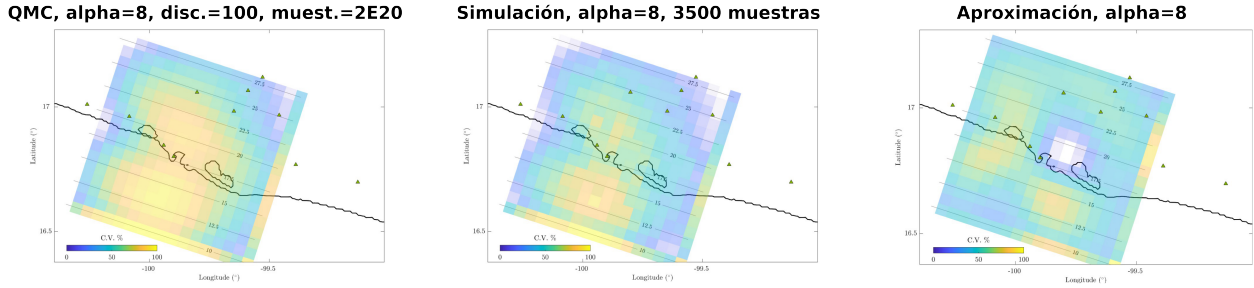


Figura 4.17: Comparación entre los métodos revisados, $\alpha = 8$ (los valores en cero se resaltan en color blanco).

de costa también ocurre una reducción de los valores de C_V .

El mapa correspondiente a la aproximación despreciando el coeficiente de asimetría presenta los valores más bajos cerca del centro. En primera instancia esto parecería coherente puesto que esta región se encuentra entre las estaciones de la esquina superior y las dos estaciones junto a la bahía. No obstante, no se tiene una cobertura por la parte sureste de dicha región. Aunado a lo anterior, los otros dos mapas no presentan valores bajos en esta zona, por lo que se considera que la aproximación subestima los valores de C_V en esa parte.

Como era de esperar, se observa claramente que, tanto para QMC como para la simulación, se obtienen los valores más altos de C_V hacia el océano, lo cual tiene sentido ya que no se tiene cobertura de estaciones en esta región.

Si bien, en el mapa correspondiente a la simulación también se observa un aumento en los valores de C_V en la zona sobre el océano, este no es tan claro como en los otros dos mapas, ya que presenta valores bajos en la parte inferior izquierda.

4.3 Alternativas al coeficiente de variación

Un inconveniente de utilizar el coeficiente de variación como medida de la incertidumbre es que la media que se utiliza para dividir la desviación estándar no está asociada directamente con el *maximum a posteriori* (MAP), i.e. la magnitud del deslizamiento más probable. Por lo cual se obtiene una relación que es complicada de interpretar en un sentido físico, sobre todo cuando la media y el MAP son muy diferentes, debido a que su magnitud no representa la dispersión que se tiene con respecto al valor más probable.

Tomando en cuenta la explicación del párrafo anterior, se propone utilizar el MAP en lugar de la media, i.e. el deslizamiento obtenido de la resolución del problema de inversión. De esta forma se obtiene un valor que normaliza la desviación estándar respecto al valor de deslizamiento obtenido para cada parámetro, lo cual se puede analizar de manera más intuitiva.

Para poner en práctica esta forma de obtener el coeficiente de variación se recibió la valiosa colaboración del Dr. Víctor Manuel Cruz Atienza, quien proporcionó datos

de desplazamiento para el mismo evento de Acapulco pero con una mayor cobertura, contando con 15 estaciones GPS. Además de esto, proporcionó una solución para el problema de inversión obtenida mediante el programa ELADIN (por sus siglas en inglés *ELastodinamic ADjoint INversion* (Tago et al., 2021)) lo cual resulta sumamente conveniente ya que, aunque la IB permite obtener una solución al problema de inversión encontrando el máximo de la posterior conjunta (Nocquet, 2018), no es el propósito de esta tesis.

En este caso la falla se divide en 15 (dip) por 17 (strike) subfallas, dando un total de 255. Como nuevamente se asume que se conoce la dirección del deslizamiento, se considera únicamente la magnitud como parámetro por subfalla, i.e. un total de 255 parámetros. Las marginales se obtienen usando el estimador cuasi-Monte Carlo usando la configuración que se menciona en la Subsección 4.2.1, es decir con un valor de $\alpha = 8$, una secuencia de Sobol de 2^{20} muestras y una discretización de 100 muestras para evaluar las marginales.

La Fig. 4.18 muestra los C_V calculados con la desviación estándar de las marginales y divididas por el MAP.

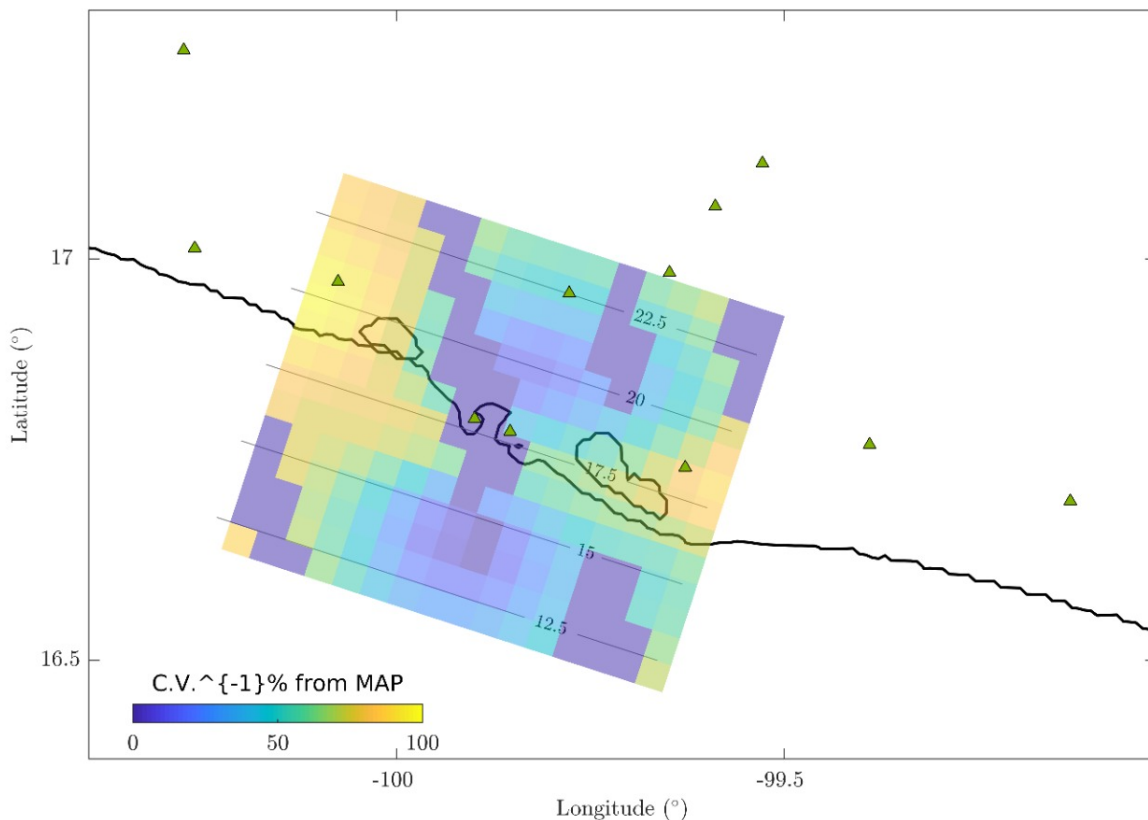


Figura 4.18: Coeficientes de variación calculados respecto al MAP.

En el mapa de C_V respecto al MAP se pueden distinguir dos regiones con valores pequeños, una al norte de la línea de costa en una zona rodeada de estaciones y una sobre el océano. De acuerdo con la forma en la que se calculan, los valores pequeños de C_V implican que se tiene un valor pequeño de desviación estándar respecto al

Alternativas al coeficiente de variación

deslizamiento invertido, lo cual se traduce en una incertidumbre baja.

Tago et al. (2021) cuantifican la incertidumbre mediante MOCs (por sus siglas en inglés *MOBILE Checkerboards*), que utiliza tableros de ajedrez (checkerboards), de un tamaño de parche determinado, considerando todas las configuraciones posibles. Lo anterior, para evitar el sesgo causado por la elección de una configuración en específico y eliminar todo rastro de los tableros de ajedrez. Teniendo las inversiones de todos los tableros, las soluciones de deslizamiento se comparan con el checkerboard objetivo correspondiente para determinar el Índice de Restitución (RI), una métrica que cuantifica qué tan cercana es la solución del modelo objetivo y que se promedia para el conjunto de checkerboards invertidos. El RI desprendido del MOC permite estimar el error nominal esperado en las inversiones para un tamaño de aspereza dado (i.e., el tamaño unitario del checkerboard) considerando la geometría de la falla y la disposición de las estaciones. El índice de restitución para parches de deslizamiento de 20 km se muestra en la Fig. 4.19.

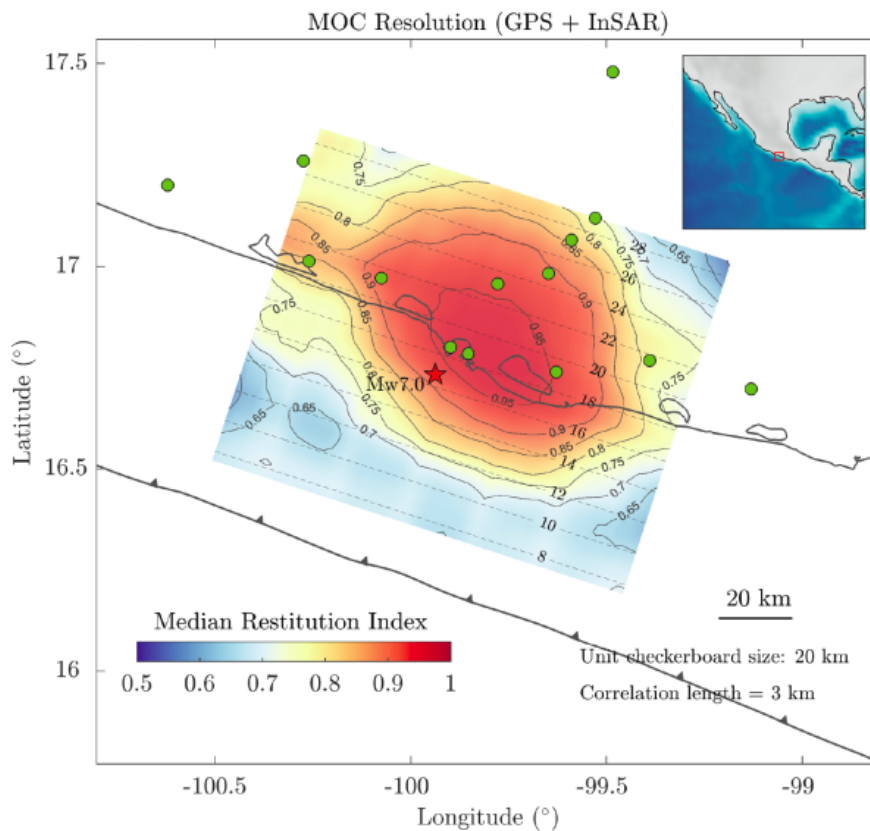


Figura 4.19: Mapa de los índices de restitución obtenidos usando la estrategia de MOCs.

Comparando las Figs. 4.18 y 4.19 se puede observar que tanto el alto índice de restitución como el bajo coeficiente de variación respecto al MAP, indican una región de baja incertidumbre en la zona rodeada de estaciones, al norte de la línea de costa.

La principal diferencia entre ambos mapas es que en la Fig. 4.18 se observa una región con valores bajos de C_V sobre el océano, al sur de la línea de costa. Una posible explicación a esto es que, mientras que los MOCs únicamente consideran la geometría

de la falla y la posición de las estaciones, el C_V considera los datos de desplazamiento, por lo que los valores pequeños de desviación estándar pueden ser información proporcionada por estos. Podemos inferir que en esa región se esperan valores de deslizamiento cercanos a cero, lo cual es consecuencia del desplazamiento registrado por las estaciones GPS.

Al utilizar el C_V , ya sea dividiendo la desviación estándar entre la media o entre el MAP, es importante considerar que es un cociente y que al tener un denominador pequeño respecto al numerador se va a obtener un valor sumamente grande que no precisamente implica que se tenga una alta incertidumbre. Para lidiar con lo anterior, se propone comparar directamente las desviaciones estándar de cada parámetro con los deslizamientos invertidos.

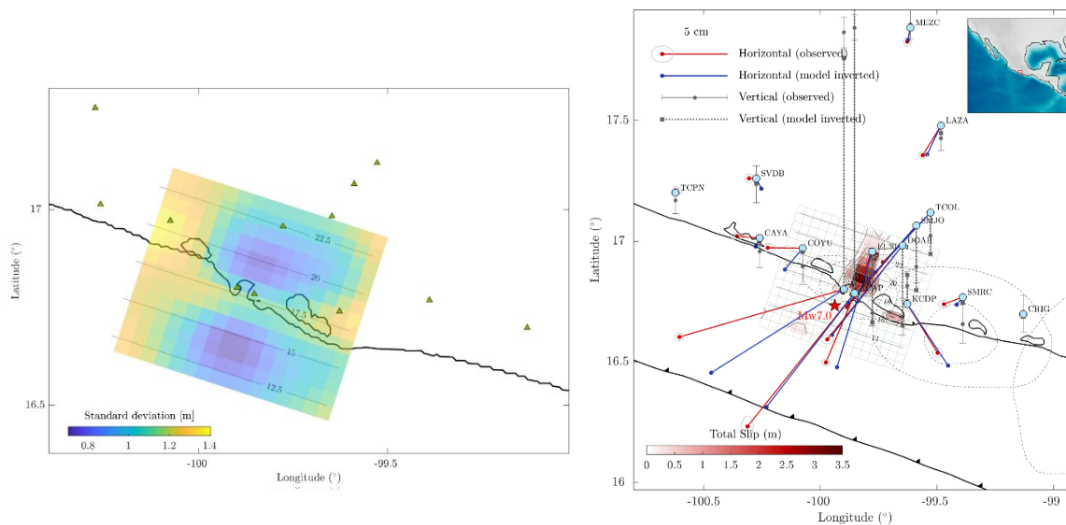


Figura 4.20: Izquierda: Mapa de la desviación estándar de la marginal para cada sub-falla. Derecha: Mapa del deslizamiento invertido mediante ELADIN proporcionado por el Dr. Cruz Atienza.

En la Fig. 4.20 se puede observar en el mapa del lado derecho que los deslizamientos de mayor magnitud se ubican en una región rodeada de estaciones, que coincide con valores pequeños de desviación estándar en el mapa del lado izquierdo. Lo anterior se traduce en una baja incertidumbre, i.e. se puede confiar en la inversión obtenida.

Capítulo 5

Conclusiones y trabajo a futuro

La cuantificación de la incertidumbre en la resolución de un problema de inversión es sumamente importante durante todas las etapas. Previo a realizar la inversión, es prudente considerar si los parámetros obtenidos serán lo suficientemente confiables y, por ende, determinar si vale la pena invertir tiempo en resolver el problema.

Durante la interpretación de los resultados de la inversión, es fundamental tener en cuenta la incertidumbre de cada parámetro obtenido, ya que considerar valores de los cuales no se tiene certeza que hayan sido bien estimados puede llevar a realizar conclusiones erróneas.

Además de lo anterior, cuando se tiene la posibilidad de instalar más instrumentos de medición, es necesario analizar un mapa de incertidumbres para saber dónde puede ser conveniente colocarlos con la finalidad de lograr una mejor cobertura y poder registrar datos que permitan hacer estimaciones más confiables.

La estrategia propuesta por [Nocquet \(2018\)](#) resulta conveniente debido a que ofrece un método para obtener la posterior conjunta y sus marginales sin la necesidad de realizar un muestreo mediante MCMC. No obstante, esta metodología tiene sus limitantes, siendo el principal obstáculo el cálculo de integrales gaussianas multivariadas.

La función *mvnun* permite obtener el valor de las integrales gaussianas eficazmente para problemas de dimensiones relativamente bajas; sin embargo, a partir de cerca de 100 parámetros, dicha función comienza a proporcionar resultados inestables, efecto que se magnifica además al incrementar el valor de α .

Por otro lado, al calcular las integrales necesarias para obtener las marginales a partir del método QMC utilizando una misma secuencia de números aleatorios proporciona resultados consistentes y de aspecto suave.

Para el uso de la estrategia propuesta por [Nocquet \(2018\)](#) utilizando QMC se pueden distinguir 3 parámetros fundamentales, estos son; el valor de α , el número de elementos del vector de recorrido de las marginales y el número de muestras para estimar las integrales.

Si bien, el inconveniente de QMC es que no hay una forma directa de cuantificar el error de la integral estimada, se observa que aún con un número sumamente pequeño de muestras, se obtienen resultados muy similares al aumentar la cantidad de dichas muestras, lo cual también se refleja en variaciones pequeñas en los C_V . Aunado a lo anterior, se observó que al incrementar el número de muestras se tiende a converger a

un resultado.

Aunque la simulación de la posterior se hace con un número de muestras relativamente pequeño para obtener los C_V , se considera que se obtiene una resolución suficiente para usar el resultado como punto de comparación.

Se puede distinguir una similitud en los resultados de la simulación y de QMC, lo cual indica que con QMC se obtienen resultados fidedignos. Caso contrario, la aproximación despreciando el factor de asimetría arroja resultados diferentes a los de los otros dos métodos, por lo cuál no es un procedimiento confiable.

Otra ventaja de poder calcular las funciones integrales representativas usando pocas muestras es que se puede explorar de manera rápida los valores adecuados de α y de la discretización.

Dado lo anterior, es posible definir una metodología simple para aplicar QMC en otros problemas similares, la cual consiste en establecer una discretización razonable (70 por ejemplo) y una secuencia con pocas muestras (2^4 por ejemplo) para calcular los C_V de varios valores de α buscando a partir de cuál se obtienen resultados estables.

Una vez definido un valor de α , se calculan nuevamente los C_V utilizando una secuencia con pocas muestras utilizando distintos valores de discretización, una vez más en busca de aquel en el que se comiencen a obtener resultados que no muestren cambios significativos.

Aunque el coeficiente de variación es una buena forma de evaluar la dispersión relativa de cada parámetro, puede llegar a ser complicado darle una interpretación física, por lo que utilizar el deslizamiento invertido en vez de la media puede ser una buena alternativa para explicar los resultados obtenidos de manera más clara.

De manera complementaria, el comparar un mapa de desviación estándar contra el deslizamiento obtenido puede ser útil en casos en los que el C_V no es efectivo, como cuando la media o el MAP tienden a cero.

Los valores pequeños de desviación estándar (que se relacionan con baja incertidumbre) en zonas donde el RI del MOC es pequeño (e.g., bajo el mar cerca de la costa) se atribuye a que el método propuesto toma en cuenta los datos reales, sugiriendo que en esas zonas los valores de deslizamiento necesarios para explicar los datos están bien resueltos (i.e., la incertidumbre en la inversión es pequeña).

La metodología propuesta a partir de la estrategia semianalítica de [Nocquet \(2018\)](#) utilizando QMC permite estimar marginales de problemas con un mayor número de parámetros de lo que se puede resolver con la función *mvnun*.

Un aspecto que vale la pena explorar es la optimización del proceso de evaluación de las muestras en las funciones para estimar las integrales, lo cual puede ser mediante la paralelización del código utilizado o buscando aprovechar librerías como *Cython*. Esto permitiría reducir el tiempo de computo y estimar las integrales con un mayor número de muestras.

Como se menciona anteriormente, los análisis de incertidumbre como el que se presenta en este trabajo se pueden usar como referencia para decidir dónde puede ser oportuno colocar nuevos instrumentos de medición, en caso de tener proyectado a futuro ampliar la red de estaciones GPS, los resultados presentados pueden ser de gran utilidad.

Bibliografía

- Caffisch, R. E. (1998). Monte carlo and quasi-monte carlo methods. *Acta numerica*, 7, 1–49. [37](#)
- Cruz-Atienza, V. M., Tago, J., Villafuerte, C., Wei, M., Garza-Girón, R., Dominguez, L. A., ... others (2021). Short-term interaction between silent and devastating earthquakes in mexico. *Nature Communications*, 12(1), 1–14. [1](#)
- Fichtner, A., y Leeuwen, T. v. (2015). Resolution analysis by random probing. *Journal of Geophysical Research: Solid Earth*, 120(8), 5549–5573. [2](#)
- Gelman, A., Carlin, J. B., Stern, H. S., y Rubin, D. B. (2014). *Bayesian data analysis*. Chapman and Hall/CRC. [5](#), [6](#)
- Giraud, C. (2021). *Introduction to high-dimensional statistics*. CRC Press. [iii](#), [34](#)
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... Oliphant, T. E. (2020, September). Array programming with NumPy. *Nature*, 585(7825), 357–362. Retrieved from <https://doi.org/10.1038/s41586-020-2649-2> doi: 10.1038/s41586-020-2649-2 [12](#)
- Ito, Y., Hino, R., Kido, M., Fujimoto, H., Osada, Y., Inazu, D., ... others (2013). Episodic slow slip events in the japan subduction zone before the 2011 tohoku-oki earthquake. *Tectonophysics*, 600, 14–26. [1](#)
- Kruschke, J. (2015). *Doing bayesian data analysis: A tutorial with r, jags, and stan* (2nd ed.). Academic Press. [4](#), [5](#), [6](#)
- Lee, M. D., y Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press. [5](#), [6](#)
- Minson, S., Simons, M., y Beck, J. (2013). Bayesian inversion for finite fault earthquake source models i—theory and algorithm. *Geophysical Journal International*, 194(3), 1701–1726. [2](#)
- Minson, S., Simons, M., Beck, J., Ortega, F., Jiang, J., Owen, S., ... Sladen, A. (2014). Bayesian inversion for finite fault earthquake source models–ii: the 2011 great tohoku-oki, japan earthquake. *Geophysical Journal International*, 198(2), 922–940. [2](#)
- Montesinos-López, J. C., Capella, A., Christen, J. A., y Tago, J. (2023). Bayesian estimation of fault slip distribution for slow slip events based on an efficient hybrid optimal directional gibbs sampler and its application to the guerrero 2006 event. *Mathematical Geosciences*, 1–28. [34](#)
- Nocedal, J., y Wright, S. (2006). *Numerical optimization*. Springer Science & Business Media. [1](#)
- Nocquet, J.-M. (2018). Stochastic static fault slip inversion from geodetic data

- with non-negativity and bound constraints. *Geophysical Journal International*, 214(1), 366–385. [2](#), [3](#), [14](#), [25](#), [27](#), [28](#), [31](#), [41](#), [46](#), [49](#), [50](#), [54](#)
- Nocquet, J.-M., Villegas-Lanza, J. C., Chlieh, M., Mothes, P., Rolandone, F., Jarrin, P., ... others (2014). Motion of continental slivers and creeping subduction in the northern andes. *Nature Geoscience*, 7(4), 287–291. [2](#)
- Okada, Y. (1985). Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the seismological society of America*, 75(4), 1135–1154. [14](#)
- Palluotto, L., Dumont, N., Rodrigues, P., Gicquel, O., y Vicquelin, R. (2019). Assessment of randomized quasi-monte carlo method efficiency in radiative heat transfer simulations. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 236, 106570. [40](#)
- Radiguet, M., Cotton, F., Vergnolle, M., Campillo, M., Valette, B., Kostoglodov, V., y Cotte, N. (2011). Spatial and temporal evolution of a long term slow slip event: the 2006 guerrero slow slip event. *Geophysical Journal International*, 184(2), 816–828. [1](#)
- Radiguet, M., Perfettini, H., Cotte, N., Gualandi, A., Valette, B., Kostoglodov, V., ... Campillo, M. (2016). Triggering of the 2014 mw7. 3 papanaoa earthquake by a slow slip event in guerrero, mexico. *Nature Geoscience*, 9(11), 829–833. [1](#)
- Robert, C. P. (1995). Simulation of truncated normal variables. *Statistics and computing*, 5, 121–125. [35](#)
- Savage, J. C. (1983). A dislocation model of strain accumulation and release at a subduction zone. *Journal of Geophysical Research: Solid Earth*, 88(B6), 4984–4996. [2](#)
- Schervish, M. J., y DeGroot, M. H. (2012). *Probability and statistics*. Pearson Education. [37](#)
- Shearer, P. M. (2009). *Introduction to seismology*. Cambridge university press. [iii](#), [29](#), [30](#)
- Tago, J., Cruz-Atienza, V. M., Villafuerte, C., Nishimura, T., Kostoglodov, V., Real, J., y Ito, Y. (2021). Adjoint slip inversion under a constrained optimization framework: Revisiting the 2006 guerrero slow slip event. *Geophysical Journal International*. [2](#), [46](#), [47](#)
- Tarantola, A., y Valette, B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics*, 20(2), 219–232. [2](#), [3](#), [16](#)
- Udías, A., Vallina, A. U., Madariaga, R., y Buforn, E. (2014). *Source mechanisms of earthquakes: Theory and practice*. Cambridge University Press. [14](#)
- Van Ravenzwaaij, D., Cassey, P., y Brown, S. D. (2018). A simple introduction to markov chain monte-carlo sampling. *Psychonomic bulletin & review*, 25(1), 143–154. [7](#)
- Walpole, R. E., Myers, R. H., Myers, S. L., y Ye, K. (2012). *Probabilidad y estadística para ingeniería y ciencias* (Vol. 162). [4](#)
- Wang, X., y Fang, K.-T. (2003). The effective dimension and quasi-monte carlo integration. *Journal of Complexity*, 19(2), 101–124. [37](#)
- Wei, G., Chen, K., y Meng, H. (2023). Bayesian inversion of finite-fault earthquake slip model using geodetic data, solving for non-planar fault geometry, variable slip, and data weighting. *Journal of Geophysical Research: Solid Earth*, 128(2),

e2022JB025225. [1](#)

Yagi, Y., y Fukahata, Y. (2011). Introduction of uncertainty of green's function into waveform inversion for seismic source processes. *Geophysical Journal International*, 186(2), 711–720. [2](#)

Apéndice

Los principales códigos empleados para obtener los resultados de esta tesis se encuentran en un repositorio de GitHub, cuya liga de internet es la siguiente: <https://github.com/Isaac-Valverde/Tesis-CIIS.git>

En este repositorio se encuentran 8 notebooks de Jupyter, el primero llamado “Programa1.ipynb” contiene la resolución de un problema de inversión mediante inferencia bayesiana.

Dicho problema consiste en ajustar el modelo de una recta a una serie de datos sintéticos generados a partir perturbaciones aleatorias a una serie de datos obtenidos de la evaluación de la función de una recta con una pendiente y ordenada al origen dadas.

El segundo es nombrado “Programa2.ipynb” y resuelve el caso sintético presentado por [Nocquet \(2018\)](#) mediante dos metodologías, la primera utilizando la estrategia semianalítica que él mismo propone y la segunda simulando muestras de la posterior mediante MCMC.

El tercer notebook llamado “Programa3.ipynb” resuelve el problema de aplicación real correspondiente al sismo de Acapulco del 7 de septiembre de 2021, a partir de la estrategia semianalítica con una simplificación para evitar el cálculo de integrales multivariadas.

El notebook llamado “Programa4_1.ipynb” fue escrito con la ayuda del M.C. Isaias Manuel Ramirez Bañales. Este simula muestras de la TMVN *Truncated Multivariate Normal* posterior de forma eficiente.

El notebook nombrado “Programa4_2.ipynb” carga las muestras generadas con “Programa4_1.ipynb” y despliega una gráfica para determinar el quemado y el IAT como referencia para determinar el *lag*. Posteriormente se obtiene una muestra efectiva que se utiliza para calcular los C_V de cada parámetro.

Los notebooks “Programa5_1.ipynb” y “Programa5_2.ipynb” fueron escritos con el propósito de explorar los valores de α y discretización respectivamente, involucrados en la resolución de la estrategia semianalítica de [Nocquet \(2018\)](#) utilizando QMC para el problema de aplicación.

El notebook nombrado “Programa6.ipynb” calcula los C_V del problema de aplicación usando la estrategia semianalítica de [Nocquet \(2018\)](#) para obtener las marginales, utilizando QMC para resolver las integrales involucradas.