

**EL FILTRO DE KALMAN
EN LA PLANEACION DEMOGRAFICA**

Manuel Ordorica Mellado

TESIS DOCTORAL

PRESENTADA A LA DIVISION DE ESTUDIOS

DE POSGRADO DE LA

FACULTAD DE INGENIERIA

DE LA

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FALLA DE ORIGEN

COMO REQUISITO PARA OBTENER

EL GRADO DE

DOCTOR EN INGENIERIA

DIRECTOR DE TESIS

DR. FRANCISCO VENEGAS MARTINEZ

CIUDAD UNIVERSITARIA

1995

01190
3
24



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.

A B S T R A C T

IN THIS THESIS, WE APPLY THE KALMAN FILTER TO GENERATE POPULATION PROJECTIONS, TO EVALUATE POPULATION POLICIES, AS WELL AS TO ANALYSE THE EFFECT ON THE POPULATION ESTIMATES WHEN THE VARIANCE OF THE ERROR CHANGE IN THE STATE-SPACE EQUATIONS.

THE KALMAN FILTER PRODUCES AN ITERATIVE AND UNBIASED ESTIMATOR WHICH GENERALIZES THE ORDINARY LEAST SQUARES METHOD WHEN THE PARAMETERS VARY IN TIME.

USUALLY, IT IS ASSUMED THAT ERRORS IN THE STATE-SPACE EQUATIONS ARE UNCORRELATED. WE EXTEND THE PROCEDURE TO THE CASE OF CORRELATED ERRORS.

INDICE

	página
Introducción	
I. Marco general: Métodos estadístico demográficos de proyecciones de la población de México	5
II. Aspectos teóricos del Filtro de Kalman	22
III. Extensión, divergencia y criterio de inicio del Filtro de Kalman. Sistemas dinámicos lineales	42
IV. Aplicación del Filtro de Kalman a estimaciones y proyecciones de población	56
V. Información imperfecta, control y políticas de población	64
VI. Resultados	69
Bibliografía	86
Apéndice I. Glosario de términos demográficos	91
Apéndice II. Recuento histórico	95

Introducción

El Filtro de Kalman ha sido aplicado exitosamente a problemas de ingeniería de control. Muchos de los artículos publicados en la literatura sobre el Filtro de Kalman aparecen en revistas de ingeniería con un lenguaje y estilo propio ajeno a la estadística. Recientemente se ha utilizado fuera del campo de la ingeniería; en pronósticos de corto plazo, en el análisis de experimentos dosis-respuesta y en la evaluación de políticas macroeconómicas.

La metodología del Filtro de Kalman es potencialmente de gran interés en el análisis de series de tiempo con parámetros dependientes del tiempo. El Filtro de Kalman es un estimador recursivo e insesgado que generaliza el método clásico de mínimos cuadrados cuando los parámetros cambian con el tiempo. La dinámica de las variables observables y no observables en el Filtro de Kalman están gobernadas por dos ecuaciones: una ecuación de observación y una ecuación de transición y ambas contienen un error estocástico. Se supone que dichos errores no están correlacionados.

Los objetivos del presente trabajo son de dos tipos: uno de orden aplicativo y otro de carácter teórico.

En primer lugar, se realizará una aplicación novedosa del Filtro de Kalman en la estimación de poblaciones a nivel nacional y en el análisis de políticas de población. Esta técnica resulta adecuada en este caso, debido a la dinámica demográfica, la cual es determinada por las condiciones sociales y económicas de la población, mismas que han sufrido importantes transformaciones en el tiempo.

En segundo lugar, se extiende la metodología del Filtro de Kalman cuando se presenta correlación entre los errores. En el modelo demográfico que se presenta, las fuentes de información que se utilizan para su implementación

son las mismas para la ecuación de observación que para la ecuación de transición por lo que se podría esperar que los errores de ambas ecuaciones se encuentren correlacionados. Los propósitos específicos son:

i) Aplicar el Filtro de Kalman en la elaboración de proyecciones de población a nivel nacional.

En nuestro país los censos se levantan cada diez años, por lo cual no se dispone de información para los años no censales. El Filtro de Kalman puede ser una técnica útil para proyectar poblaciones anualmente, usando pocos datos demográficos. Se harán comparaciones con resultados obtenidos con otras técnicas de proyección. A la fecha sólo se han utilizado metodologías restringidas al campo de los especialistas en población.

ii) Aplicar el Filtro de Kalman al análisis de políticas de población.

En México se han establecido metas de crecimiento demográfico, sobre todo en el campo de la natalidad, y se ha avanzado poco en la evaluación de las mismas. Es de especial interés comparar la trayectoria de la población sin y con programas en materia de regulación de la fecundidad, es decir, comparar resultados sin política con resultados con política. Esta comparación permite analizar el esfuerzo que el Estado Mexicano tendría que hacer para cumplir con los objetivos planteados. Este análisis nos conduce a un problema de optimización cuadrático, sujeto a las ecuaciones de observación y de transición.

iii) Extender la metodología del Filtro de Kalman en el caso de que los ruidos de las ecuaciones de observación y de transición estén correlacionados, es decir,

$$E\{W_k \quad V_k^T\} = C_k.$$

La teoría desarrollada respecto al Filtro de Kalman supone que los errores son no correlacionados. En este sentido, resulta de interés estudiar como se transforman las ecuaciones fundamentales del Filtro de Kalman al introducir este supuesto más general.

iv) Analizar el efecto en las proyecciones de población en el caso de que se adopten diferentes niveles de las varianzas de los errores, tanto en el caso de la ecuación de transición como en el de la ecuación de observación.

Hasta la fecha, todas las proyecciones de población que se han realizado en nuestro país proporcionan resultados puntuales. En ningún caso se dan estimaciones en rangos de variabilidad. Esta técnica permite dar estimaciones de la población con sus intervalos de confianza.

Uno de los problemas que se presenta en la planeación social y económica es la falta de información futura sobre la población, y sobre los componentes del crecimiento de la población, debido a que las técnicas que se usan requieren de mucha información demográfica. El Filtro de Kalman permite disponer de pronósticos con menor esfuerzo. La información sobre el número de personas es de gran utilidad en la planeación de la educación, de la salud, de la vivienda, del empleo, sólo por mencionar algunos campos. Aproximarse al conocimiento del número de habitantes que tendrá este país, podría apoyar a los encargados en la toma de decisiones sobre cuáles son y serán las mejores políticas a adoptar. Dichas políticas minimizarán el esfuerzo en la obtención de metas. Si los políticos de hace medio siglo hubieran tenido un instrumental técnico para estimar y proyectar población, es probable que el tema demográfico no se hubiera convertido en un problema fundamental del desarrollo de México.

El desarrollo de este trabajo es como sigue:

En el capítulo I, se presentan las diversas técnicas de pronóstico de la población de México que se han realizado desde hace medio siglo, a fin de establecer un marco comparativo. En el capítulo II, se estudian los aspectos teóricos del Filtro de Kalman. En primer lugar, se analiza el Filtro de Kalman a través de la minimización del error cuadrático medio, luego se introduce un enfoque alternativo del Filtro de Kalman en el marco de la estadística Bayesiana y finalmente se establece el vínculo entre el Filtro de Kalman y la programación dinámica. En el capítulo III, se proporciona una extensión del Filtro de Kalman cuando los errores están correlacionados. En el capítulo IV, se introduce un nuevo modelo de estimaciones de población utilizando el Filtro de Kalman. El capítulo V plantea un modelo para el establecimiento de escenarios, los cuales son de utilidad en la evaluación de políticas de población. Finalmente, se presentan un conjunto de resultados en donde se analiza el efecto de diferentes niveles de varianza de los errores de las ecuaciones de observación y de transición sobre las proyecciones de población y las conclusiones sobre este trabajo. Asimismo, se presenta un apéndice sobre los términos demográficos usados en el trabajo y un apéndice

que describe en forma breve la evolución histórica de la teoría de estimación hasta llegar al Filtro de Kalman.

I. Marco general: métodos estadístico demográficos de proyecciones de la población de México

Bien sea por medio de la magia o de la ciencia, los seres humanos han intentado, desde siempre, adivinar lo que el futuro les tiene reservado. Para los mexicanos, por ejemplo, el mundo por ellos habitado era uno más de los muchos que los habían precedido y ellos mismos serían remplazados por otros habitantes. El futuro estaba determinado por esta evolución cíclica; de ahí que la existencia de los individuos estuviera supeditada a vivir el presente y el pasado. El interés por dominar el tiempo estaba dirigido, por un lado, a encontrar la clave para entender y prever los fenómenos naturales y, por el otro, a definir el destino de los seres humanos. "Nada nuevo hay bajo el sol", dice el Eclesiastés. Este es el anuncio profético de la repetición de todas las cosas.

A diferencia de la antigüedad, en la época actual los individuos toman al futuro como punto de referencia y organizan sus instituciones y actividades de acuerdo con esa visión del mundo. Aproximarse al futuro de las sociedades y en particular conocer el número de personas que habrá en los próximos años es una necesidad, aun cuando se sabe que no es posible predecir el porvenir con precisión. Las proyecciones de población son importantes en la programación económica y social e indispensables para calcular los requerimientos futuros en materia de educación, empleo, vivienda, salud y seguridad social. Dichas proyecciones ayudan también a entender mejor la dinámica de los fenómenos demográficos, pues permiten estudiar los efectos

de las variaciones de las tasas de natalidad, mortalidad y migración en la estructura por edades, así como predecir el futuro y comprender el pasado. Por su parte, la proyección de escenarios deseables permite proponer alternativas a la evolución de fenómenos demográficos que pueden provocar efectos negativos en el desarrollo del país.

Las primeras proyecciones demográficas fueron las de la población total, pues no se contaba con datos de los nacimientos, de la inmigración y de la emigración. Las perspectivas de población realizadas por Gregory King en Inglaterra hacia fines del siglo XVII se basaban en el tiempo de duplicación del número de habitantes. King realizó sus pronósticos de la población del Reino Unido al final de cada siglo hasta el año 2030. Malthus también hizo proyecciones utilizando la tasa geométrica de crecimiento, aunque pronto advirtió que las perspectivas demográficas, basadas en las tasas de crecimiento o decrecimiento existentes, tenían poca confiabilidad debido a la irregularidad de epidemias y hambrunas que afectaban el nivel de los nacimientos. Suponiendo decía Malthus- que la cantidad de alimentos producidos en la etapa inicial se pudiera multiplicar por dos en 25 años y que la población se duplicara cada cuarto de siglo, ocurriría lo siguiente:

“Supongamos que la población de nuestra isla es de 11 millones de habitantes- se refería a Inglaterra- y que la producción actual basta para sostener bien a ese número de personas. Al cabo de los 25 primeros años la población sería de 22 millones de habitantes y, habiéndose doblado la producción de alimentos, los medios de subsistencia seguirían bastando para la población. En los 25 años siguientes, la duplicación sería de 44 millones y los medios de subsistencia sólo bastarían para mantener a 33 millones de habitantes. En el siguiente período de 25 años la población sería de 88 millones y los alimentos sólo bastarían para mantener a la mitad de ese número de personas. Y al finalizar el primer siglo la población sería de 176 millones de habitantes y los abastecimientos sólo podrían mantener a 55 millones, lo que dejaría sin medios de subsistencia a 121 millones” (Malthus,1951). Malthus consideraba la *tasa geométrica* y el tiempo de duplicación como instrumentos clave para elaborar proyecciones de población. Durante los siglos XVII y XVIII el tiempo de duplicación de la población era la metodología utilizada en esas proyecciones (Naciones Unidas,1978). La función geométrica es $P_t = P_0(1 + r)^t$. A fin de avanzar en las metodologías de proyecciones en los siglos XVIII y XIX fue preciso realizar con cierta regularidad censos

completos. Los pronósticos que se hacían a principios del siglo XIX suponían un crecimiento geométrico de la población. Más tarde se cuestionaron las hipótesis y se comenzaron a hacer supuestos de tasas de crecimiento en descenso.

Posteriormente se utilizaron funciones matemáticas para hacer proyecciones de población. La función más conocida es la logística, cuya primera versión fue realizada por Verhulst en 1848 y retomada por Pearl y Reed en 1920. Se basa en la hipótesis de que las poblaciones tienden a alcanzar un límite que luego no superan. Esta función se sigue utilizando para proyectar el comportamiento futuro de la fecundidad y la mortalidad, aunque se le critica no sólo por ser demasiado mecánica, sino por no tomar en cuenta la estructura por sexo y edad. Sin embargo, su aplicación para períodos breves produce resultados certeros. Otras funciones en que existen límites en el ritmo de crecimiento demográfico son la de Gompertz: kg^{ct} y Makeham: $A + B^{ct}$.

Con el método de los componentes se empieza a poner mayor atención en elaborar proyecciones de población. Este método se utiliza en la actualidad y consiste en aplicar hipótesis sobre la fecundidad, la mortalidad y la migración a la composición por edad y sexo. Cada grupo de edades se proyecta con probabilidades de supervivencia. Los nacimientos se obtienen a partir de las tasas específicas de fecundidad proyectadas y se proyectan en el tiempo. A la estructura resultante se le suman o restan los migrantes. En 1895, Cannan aplicó un procedimiento que en esencia es el de los componentes para preparar las proyecciones de Inglaterra y Gales.

Después de la segunda guerra mundial se estimuló la planificación del desarrollo a fin de reconstruir las economías devastadas por el conflicto bélico. Esto motivó que se realizaran proyecciones demográficas de la población total, por grupos de edades, medios rural y urbano, hogares y familias, entre otros. En los años veinte las proyecciones se realizaban utilizando curvas, en especial la función exponencial; en los treinta, se usaba la logística, y en los cincuenta, el citado método de los componentes. A partir del siguiente decenio y hasta la fecha se utilizan variaciones del mismo método, que se ha perfeccionado. La calidad de los resultados de la técnica de los componentes depende de la confiabilidad primordialmente de las hipótesis. Al final del capítulo se presentan las técnicas de proyección utilizadas en diversos momentos.

I.1 Las proyecciones demográficas de México en la perspectiva histórica

México cuenta con una amplia historia en la realización de proyecciones demográficas. La División de Población de las Naciones Unidas publicó en 1954 proyecciones de población por sexo y grupos de edad para el período 1950-1980 correspondientes a Centroamérica y México. Se establecieron tres hipótesis sobre el número de nacimientos futuros: una máxima, una media y una mínima. Según la clasificación del país en relación con la natalidad se determinó la forma de la transición. Por lo que se refiere a la mortalidad, se trazaron curvas representativas, cada una de las cuales permite determinar la evolución de las tasas de mortalidad por edad (Naciones Unidas, 1954).

En 1958, Coale y Hoover realizaron proyecciones de población mediante el método de los componentes con base en la población de 1950 y analizaron las perspectivas demográficas y su influencia en el desarrollo económico. Estudiaron en especial los efectos de los cambios en los niveles de fecundidad sobre los procesos económicos. Analizaron el caso de la India e incluyeron a México como ejemplo de un país de baja densidad demográfica en contraste con la India (Coale y Hoover, 1958).

En 1960 Louis Ducoff preparó para la CEPAL un estudio sobre los problemas demográficos de Centroamérica y México y sus efectos en el desarrollo de la región. En especial se intentó establecer la relación entre el crecimiento demográfico y los problemas referentes a la integración económica de la región. Para ello se corrigieron las proyecciones existentes, pues se observó que los niveles de natalidad se habían subestimado. Asimismo, debido a que se calcularon niveles de mortalidad superiores a los del trabajo de las Naciones Unidas en 1954, las cifras de población obtenidas son menores que las calculadas en ese año (Naciones Unidas, 1960).

Gustavo Cabrera realizó en 1960 estimaciones de la población del país, para lo cual estableció tres hipótesis de la dinámica futura de la fecundidad: una constante y dos que suponían un descenso. Respecto a la mortalidad estableció un incremento de dos años en la esperanza de vida al nacer por cada quinquenio. Según esta proyección la población pasaría de 35.1 millones en 1960 a una cifra de entre 60.7 y 65.2 millones en 1980 (Cabrera, 1960).

En 1962 se publicaron las proyecciones de población realizadas por Julio Durán Ochoa. Ese trabajo tomó como base los datos censales de 1960 y adoptó el supuesto de que la fecundidad permanecería invariable de 1960 a 1980. Con respecto a la mortalidad se consideró un incremento en la esperanza de vida al nacer de 2.5 años por cada cinco (Durán, 1962). Un año más tarde Zulma Recchini elaboró proyecciones de población por grupos de edad y sexo para el período 1960-1980. Supuso una fecundidad constante y un aumento de la esperanza de vida al nacer de 2.5 años cada quinquenio (Recchini, 1963). En 1964, Zulma Recchini y Miguel Chavira, y en 1966 Miguel Chavira realizan proyecciones demográficas. Sus resultados son semejantes a los de Recchini en 1963.

En 1966 se publicaron las proyecciones de población para el período 1960-1980 de Benítez y Cabrera. En ellas se utilizó el método de los componentes y la población base fue la del censo de 1960. Se establecieron tres hipótesis respecto a la fecundidad: 1) constante; 2) constante hasta 1970 y con descensos de 5 por ciento de 1970 a 1975 y de 10 por ciento de 1975 a 1980, y 3) constante hasta 1965, con descensos de 5 por ciento de 1965 a 1970, de 10 por ciento de 1970 a 1975 y de 15 por ciento de 1975 a 1980 (Benítez y Cabrera, 1966). Los resultados de esas proyecciones condujeron a modificar los objetivos de la política de población: de pronatalista se convirtió en reguladora del ritmo de crecimiento demográfico del país.

En 1978 la Secretaría de Programación y Presupuesto (SPP) elaboró proyecciones de 1970 al año 2000. La población base fue la del censo de 1970, corregida y proyectada al 30 de junio. Se establecieron cuatro hipótesis de fecundidad y una de mortalidad. Los resultados son superiores a los de las otras perspectivas demográficas. Con base en la hipótesis de fecundidad constante se estimó una población de 151.8 millones de habitantes, cifra muy elevada si se le compara con las otras proyecciones demográficas (Secretaría de Programación y Presupuesto, 1978). En todas las proyecciones no se consideró la migración internacional.

En 1980 Sergio Camposortega elaboró perspectivas demográficas para el período 1970-2040. Planteó dos hipótesis de mortalidad, cinco de fecundidad y cuatro de migración. Seleccionó ocho proyecciones de población que cubren la banda probable de la dinámica demográfica futura y dos que definen los límites máximo y mínimo. Dichas proyecciones son: 1) límite máximo; 2) límite máximo probable; 3) variación 1: límite máximo probable;

4) variación 2: límite máximo probable; 5) variación 1: alternativa media; 6) hipótesis recomendada; 7) variación 2: alternativa media; 8) variación 1: límite mínimo probable; 9) límite mínimo probable; 10) límite mínimo. Los resultados varían para el año 2000 entre 99.1 millones según el límite mínimo y 133 millones de personas a partir del límite máximo (Camposortega, 1980).

La SPP, el Consejo Nacional de Población (Conapo) y el Centro Latinoamericano de Demografía (Celade) publicaron en 1983 las proyecciones de población para el período 1980-2000. La población base fue la de 1980. Se planteó una sola hipótesis de mortalidad suponiendo que la ganancia en la esperanza de vida se hace más lenta conforme este indicador es mayor. Para obtener los niveles futuros de la fecundidad se utilizó un modelo logístico. Respecto a la migración internacional se supuso que los saldos netos migratorios del período 1960-1970 se mantenían constantes para el período de la proyección. Se realizaron dos proyecciones de población: una de ellas, la programática, tomó en cuenta la meta de crecimiento demográfico de 1 por ciento al año 2000; la otra, denominada alternativa, presentó un panorama futuro al que se llegaría en el supuesto de que la disminución de la fecundidad fuera menor que la establecida en la política de población (Secretaría de Programación y Presupuesto, Consejo Nacional de Población y Centro Latinoamericano de Demografía, 1983).

A mediados de los ochenta, el INEGI y el Conapo elaboraron "Las proyecciones de la población de México y de las entidades federativas: 1980-2010" (INEGI y Conapo, 1985). Para proyectar la mortalidad se aplicó una función logística a las esperanzas de vida entre 1950 y 1980. Respecto a la fecundidad se plantearon dos hipótesis: la programática y la alternativa. La primera se aproxima a la meta de crecimiento demográfico de 1 por ciento al año 2000 planteada por el Conapo en 1977. La alternativa supone que se llegará a una cifra más elevada en la tasa de crecimiento demográfico: 1.6 por ciento anual. La migración internacional se mantuvo constante en el período de la proyección a la observada de 1975 a 1980.

Un año después, Núñez y Moreno efectuaron proyecciones de población urbanas y rurales para 1980-2010 con base en el modelo multirregional. Este incorpora de manera explícita los componentes demográficos como el método clásico, pero además considera simultáneamente la interrelación de las variables demográficas.

En 1992 la División de Población de la ONU publicó el documento World Population Prospects, donde se revisan las proyecciones demográficas realizadas hasta esa fecha. Según la hipótesis intermedia, la población calculada para el 2000 es de 102.6 millones (Naciones Unidas, 1992).

En 1993 el Conapo elaboró proyecciones para el período 1990-2025 con base en los resultados del censo de 1990. Para el año 2000 calcula una población de 100.5 millones de habitantes, y para el período 2000-2005, una esperanza de vida al nacer de 73 años y una tasa global de fecundidad de 2.5. La población estimada para el 2025 es de 130.8 millones de personas, casi el doble de la censada en 1980 (66.8 millones). Esto significa que antes del 2025 se debe construir otro México como el de 1980, aun cuando la tasa de crecimiento demográfico en el 2025 se encontrará por debajo de 1 por ciento anual. En el cuadro se presentan los resultados de las proyecciones de población hasta el año 2000 según diversos autores e instituciones.

La exactitud de una proyección de población depende en gran medida del período que se considere: a 5 o 10 años son confiables y de utilidad para planes quinquenales; a 15-20 años son menos ciertas, pero reflejan la dinámica general, y a 30-35 años indican cómo sería la estructura de la población en un período lejano si operan los procesos de la actualidad.

Al comparar las diversas proyecciones de población con los datos censales se observa una diferencia significativa si se toma como base la población inicial de la proyección. En este análisis sólo se consideran las proyecciones siguientes: media, programática o única. Se observa que las estimaciones futuras son más certeras en plazos breves. El intervalo que se forma entre la proyección más baja y la más elevada incluye a la población de los censos de 1960, 1970 y 1980. El rango de variación en 1960 va de 33.3 a 36.0 millones de personas, mientras el censo enumeró a 34.9 millones. En 1970 la banda osciló de 41.8 a 51.1 millones, y la población censal fue de 48.2 millones y se censaron 66.8 millones de personas. En 1990 varía de 84.5 a 99.7 millones. El intervalo de variación es el siguiente: 1960, 2.7 millones; 1970, 9.3 millones; 1980, 9.3, y 1990, 15.2 millones, lo que muestra el aumento del diferencial en el tiempo. En el último año señalado el intervalo de variación de las proyecciones no incluye la censal, en gran medida debido a que las hipótesis de la migración internacional no fueron adecuadas (salieron del país más ciudadanos de los que se había previsto).

Al analizar los resultados del cuadro que aparece al final de este apartado se distinguen tres etapas bien definidas: en la primera las proyecciones realizadas en los cincuenta y los sesenta presentan rangos de variación pequeños en el número de habitantes; en la segunda, formada por las proyecciones de 1978 y 1980, los intervalos de variación son grandes, y en la tercera, que incluye las proyecciones realizadas de 1983 a la fecha, los rangos se vuelven a reducir. Una proyección es útil cuando las hipótesis planteadas se traducen en rangos de población pequeños. Un intervalo de variación muy amplio no es de utilidad, puesto que es muy probable que caigan los resultados de población de los censos.

Pese a que hoy se dispone de mayor información, super computadoras y técnicas matemáticas más avanzadas, no se ha logrado avanzar en las proyecciones de población. Se ha dedicado más tiempo a las técnicas y a la computación que a la formulación cuidadosa y analítica de hipótesis. En este sentido, es preferible detenerse a reflexionar sobre el futuro demográfico que aplicar complejas metodologías matemáticas. Por tanto, las perspectivas demográficas sólo se deben tomar como una guía para establecer objetivos y metas en la planeación del desarrollo.

I.2 Situación demográfica actual y futura

En 1995 la población del mundo llegó a 5 600 millones de habitantes. antes de finalizar el siglo seremos 6 000 millones, 7 000 un poco antes del 2010 y 8 000 en el 2020. Para ilustrar la velocidad del ritmo de crecimiento de la población mundial y la historia demográfica desde la aparición del Homo erectus hasta 1987 -año en que se llegó a 5 000 millones de habitantes- se comprime en un solo año un período que abarca un millón de años.

El Homo erectus aparece a las cero horas del 1 de enero del año demográfico. Transcurre casi todo el año y a las 23:00h del último día del mes de diciembre (1850) 1 000 millones de habitantes habitan el planeta. A las 23:30h (1930) la cantidad asciende a 2 000 millones; quince minutos después (1960), a 3 000 millones, y a las 24:00h del último día del año (1987), el número es de 5 000 millones. Los últimos 4 000 millones se alcanzaron en sólo una hora del año demográfico. A las 00:05h del primero de enero del siguiente año (1999) habrá 6 000 millones; 7 000 a las 00:11h (2010); 8 000 a

las 00:17h (2020) y 10 000 millones a las 00:45 hrs (2075). De 1987 al 2075, es decir, en 88 años se tendría que construir otro mundo como el de 1987.

Testigo de grandes transformaciones económicas, sociales y tecnológicas, el siglo XX presencié también los cambios demográficos más intensos de la historia. En la tercera década de este siglo la población del planeta ascendió a 2 000 millones. Se necesitó por lo menos un millón de años para que esto ocurriera, mientras que en prácticamente seis décadas la población pasó de esa cifra a los actuales 5 600 millones de habitantes que habitan el planeta. El crecimiento demográfico mundial es de aproximadamente mil millones de personas cada 12 años.

En este orden de ideas, ¿cuál es la situación demográfica de México? Según el censo, el país tenía 81.2 millones de habitantes en 1990 -el undécimo más poblado del mundo-; se estima que en el 2000 la cantidad ascenderá a 100 millones.

La tasa anual de crecimiento demográfico pasó de 3.4 por ciento en 1970 a 2 por ciento en 1990. A pesar de la importancia de esa disminución, el ritmo es aún elevado, pues de mantenerse la población se duplicaría en períodos de 35 años. Cabe señalar que dicho descenso se ha presentado en mayor medida en las zonas urbanas y en los estratos de población medios y altos.

En 1990, 38.4 por ciento de la población pertenecía al grupo de 0 a 14 años; 57.4 por ciento al de 15 a 64 años, y 4.2 por ciento al de 65 años y más. Se prevé que en el 2025, 23 por ciento tendrá menos de 15 años; 68 por ciento, entre 15 y 64, y 9 por ciento mayor de 65 (CONAPO,1993).

De 1960 a 1990 se registraron cambios importantes en la tasa global de fecundidad. A fines de los sesenta el promedio de hijos por mujer al terminar su vida fértil era de 7 y la de 1990 se calculó en 3.4. El descenso de la fecundidad no fue homogéneo en todos los grupos; difiere según la escolaridad y el lugar de residencia. La Encuesta Nacional sobre Fecundidad y Salud (ENFES), realizada en 1987, muestra que las mujeres que habitan en las áreas metropolitanas tienen tres hijos menos que las que viven en localidades menores a 20 000 habitantes. Con respecto a la escolaridad se observa un hecho semejante: quienes cursaron la primaria completa tienen dos hijos menos que las que no la cursaron.

La mortalidad ha descendido de manera importante. La esperanza de vida al nacer aumentó de 58 años en 1960 a 71 años en la actualidad.

Las cifras globales esconden las acentuadas diferencias entre las entidades federativas: en 1988 en Nuevo León la esperanza de vida al nacer era de 74 años, mientras que en Oaxaca era ligeramente mayor de 60: una brecha de 14 años.

La tasa de mortalidad infantil, considerada como un indicador del desarrollo, ha presentado una continua disminución: de 132 en 1950 a 40.9 defunciones de niños menores de un año por cada 1 000 nacidos vivos en 1990.

A comienzos del siglo XX la natalidad y la mortalidad eran elevadas. La esperanza de vida al nacer era de alrededor de 30 años. Un alto porcentaje de los niños que nacían en esta época no conocían a sus padres y abuelos como resultado de la reducida esperanza de vida. Incluso, debido a que la fecundidad y la mortalidad infantil y en la niñez eran elevadas, muchos infantes no llegaban a conocer ni siquiera a sus hermanos. Las personas se casaban jóvenes y debido a la elevada mortalidad, las uniones conyugales duraban poco.

La migración interna ha desempeñado un papel importante en el crecimiento de la población. Esta se ha concentrado principalmente en el centro del país, integrado por 13 entidades federativas, donde radican seis de cada diez personas ocupando sólo 20 por ciento del territorio nacional. La población también se concentra en cuatro zonas metropolitanas. En 1990 la Ciudad de México concentraba a 16 millones de personas, casi la quinta parte de la población total. Del censo de 1990 se desprende que la fuerza de atracción de las ciudades más grandes ha cedido terreno a las ciudades intermedias.

La baja de la fecundidad traerá importantes efectos en la estructura por edades: se reducirá la población en edad escolar y preescolar en términos relativos y absolutos y aumentará la proporción de los habitantes en edades activas y avanzadas. En tanto que ese descenso tiene efectos directos en la población que demandará escuelas y atención a la salud materno-infantil, las repercusiones son más lentas y tardan más tiempo en afectar la estructura y el número de la población en edad activa, pues ésta ya ha nacido.

¿Qué nos espera en el futuro?

Los escenarios demográficos del futuro, así como las demandas y presiones de orden social que van a generarse, son previsibles. La observación de las

tendencias de la población y sus posibles modificaciones brindan los elementos necesarios para apreciar la magnitud de los esfuerzos que se sugieren en educación, salud, empleo, vivienda y nutrición. El análisis comprende el período 1990-2025.

En el supuesto de que el ritmo anual de crecimiento demográfico descendiera de 1.9 por ciento en 1993 a 0.7 por ciento en el período 2020-2025, la población total al último año ascendería a 130.8 millones de personas. Los efectos del decremento serían los siguientes: i) la población en edad preescolar y escolar (0-14) pasaría de 32.1 millones en 1990 a 30.1 millones en el 2025. ii) El volumen de la población en edad activa (15-64) pasaría de 48.1 millones en 1990 a 88.9 en el 2025. iii) La población en edad de jubilación (65 y más) se elevaría de 3.6 millones en 1990 a 11.8 millones en el 2025. Esto significa que ya han nacido o están por nacer muchos de los que requerirán empleo en el 2025 y su número sería casi el doble de la población en edad activa de 1990. La población en edad avanzada del 2025 se habría multiplicado por tres respecto a 1990. Este fenómeno, en cambio, no se presentaría en el grupo de edades de 0 a 14 años, ya que este grupo se reduciría en 2 millones de 1990 al 2025. El descenso del ritmo de crecimiento de la población dará como resultado que las políticas de bienestar social orientadas a determinados sectores (educación, salud) tendrán la posibilidad de trascender la mera prestación de servicios atada por el problema cuantitativo, a una política en que lo cualitativo sea el eje central de la planificación del desarrollo. Es importante destacar, por ejemplo, que el ritmo de formación de profesores debe reducirse, pues podría ocurrir que hubiera maestros sin alumnos.

Si bien es necesaria, la regulación de la tasa de crecimiento es insuficiente para adecuar los procesos demográficos con los del desarrollo. Es urgente, además, regular los movimientos migratorios, sea estimulando la permanencia de la población en su lugar de origen o reorientando las migraciones hacia ciudades de tamaño intermedio a zonas con recursos naturales y productivos mediante la creación de empleos. La regulación de los movimientos migratorios y la disminución del crecimiento demográfico harán posible que la población de las zonas metropolitanas se reduzca.

La alarma sobre el proceso de urbanización ha crecido con las proyecciones realizadas para las poblaciones futuras de algunas de las zonas metropolitanas más grandes del país. Se proyecta que la Ciudad de México, por ejemplo, se hubiera podido transformar en una población con 31 millones de personas en

el año 2000 si hubiera permanecido invariable el ritmo de crecimiento de la población del decenio de los setenta.

Es importante destacar que aún se tendrán que hacer esfuerzos importantes para intensificar la salud reproductiva en el marco de los programas sociales, sobre todo en las zonas marginales y las áreas rurales que muestran un lento descenso de la natalidad y una mortalidad elevada. Estos esfuerzos se deben enmarcar en los programas económicos y sociales tendientes a mejorar las condiciones de vida. Al mismo tiempo, se deben fortalecer las bases para elevar la calidad de la vida en las zonas que expulsan población.

En este sentido, las acciones que se emprendan tendrán mejor aceptación y mayores posibilidades de éxito si se sustentan en la amplia participación ciudadana. Esta sólo podrá lograrse mediante la consolidación de los programas de educación y comunicación, que pueden ser la base de sustento de programas orientados a reducir la mortalidad y la morbilidad infantiles, regular la fecundidad mediante programas de salud reproductiva y racionalizar la distribución espacial de la población.

¿Cómo será la fecundidad y la mortalidad al finalizar el primer cuarto de siglo?

La edad media de la fecundidad en el período 2020-2025 se ubicará alrededor de 32 años como resultado de la mayor participación de la mujer en la actividad económica. El promedio de hijos por mujer al final de su vida fértil será cercano a dos, es decir, la población se encontrará en un crecimiento demográfico de remplazo. Como resultado de la ampliación del período fértil, las parejas podrán decidir el espaciamiento entre sus hijos. Debido a los avances de la ciencia, disminuirá el alto riesgo de los embarazos después de los 35 o 40 años de edad. Los niños del siglo XXI tendrán uno o dos hermanos.

El aumento en la esperanza de vida hasta los 76 años en el período 2020-2025 permitirá que la mayoría de los niños que nazcan en esa época conozcan a sus padres, abuelos e incluso a sus bisabuelos. Un niño podrá convivir con su padre hasta que éste cumpla 75 años o más, cuando a principios del siglo sólo lo hacían por uno o dos decenios.

Es importante resaltar que a partir del primer quinquenio del siglo XXI empezará a incrementarse la tasa bruta de mortalidad como resultado del proceso de envejecimiento de la población mexicana. Esto redundará en una

aceleración en el descenso de la tasa de crecimiento demográfico. La inercia demográfica conducirá a una tasa de crecimiento demográfico cercana a cero poco antes de la mitad del siglo XXI.

En los primeros años del próximo siglo se iniciaría una nueva etapa demográfica con un descenso lento en la tasa bruta de natalidad y un incremento también lento en la de mortalidad. Esto provocará que se tienda más rápido hacia el crecimiento demográfico cero.

En conclusión, las proyecciones de población no deben considerarse como una bola de cristal con la que podríamos adivinar el porvenir, más bien deben verse como una brújula que nos permita orientar los programas en materia socioeconómica.

Además, los pronósticos que se han realizado se centran en técnicas de orden demográfico principalmente, pero no se ha hecho nada usando técnicas dinámicas como es el caso del Filtro de Kalman que pueden, además de proporcionar intervalos de confianza en las estimaciones de población, aprovechar el potencial matemático de orden teórico que subyace a este tipo de herramientas. También el Filtro de Kalman puede ser utilizado en la evaluación permanente de los objetivos planteados en materia poblacional, ya que los componentes del crecimiento demográfico cambian con el tiempo, no son fijos, ya que dependen de las condiciones sociales y económicas.

RESULTADOS DE LAS PROYECCIONES DE LA POBLACION DE MEXICO
CALCULADAS POR DIVERSOS AUTORES
(Millones de habitantes)

CALCULO	HIPOTESIS	1950	1960	1970	1980	1990	2000
Naciones Unidas (1)	Alta	25.6	34.1	46.3	64.4	-	-
	Media	25.6	33.3	43.0	55.5	-	-
	Baja	25.6	32.5	40.0	48.2	-	-
Naciones Unidas (2) (revisión)	Alta	25.8	33.6	44.9	62.0	-	-
	Media	25.8	32.8	41.8	53.3	-	-
	Baja	25.8	32.0	38.9	46.4	-	-
Zulma Recchini y Miguel Chavira (3)	Unica	-	36.0	50.7	72.4	-	-
Coale y Hoover (4)	Alta	26.6	36.0	50.1	61.9	-	-
	Media	26.6	36.0	49.3	70.9	-	-
	Baja	26.6	35.7	46.4	57.5	-	-
Gustavo Cabrera (5)	Alta	-	35.1	47.5	65.2	-	-
	Media	-	35.1	47.1	63.4	-	-
	Baja	-	35.1	46.6	60.7	-	-
Miguel Chavira (6)	Unica	-	36.0	50.7	72.4	-	-
Julio Durón Ochoa (7)	Unica	-	34.7	46.8	63.3	-	-
Zulma Recchini (8)	Unica	-	36.0	50.7	72.6	-	-
Benitez y Cabrera (9)	Alta	-	36.0	51.1	73.6	-	-
	Media	-	36.0	51.1	71.9	-	-
	Baja	-	36.0	50.8	69.3	-	-
Secretaría de Programación y Presupuesto (10)	Constante	-	-	50.7	71.9	104.2	151.8
	Hip. 1	-	-	50.7	71.6	102.3	143.9
	Hip. 2	-	-	50.7	71.1	99.7	135.8
	Hip. 3	-	-	50.7	70.4	95.5	126.1

Sergio Camposortega (11)	Límite Máximo	-	-	50.1	69.0	95.5	133.6
	Recomendada	-	-	50.1	68.1	86.3	107.3
	Límite mínimo	-	-	50.1	68.3	84.3	99.1
S.P.P., CONAPO, CELADE (12)	Alternativa Programativa	27.4	37.0	51.2	69.4	89.0	109.2
		27.4	37.0	51.2	69.4	86.2	99.6
Leopoldo Nuñez y Lorenzo Moreno (13)	Unica	-	-	-	69.7	84.5	100.1
INEGI-CONAPO (14)	Alternativa Programativa	-	-	-	69.7	86.1	104.0
		-	-	-	69.7	85.8	100.0
CELADE Y NACIONES UNIDAS (15)	UNICA	27.3	36.5	50.3	67.0	84.5	102.6
CONAPO (16)	UNICA	-	-	-	-	83.8	100.5

FUENTES

- (1) NACIONES UNIDAS. La población de la América Central y México en el período 1950 a ST/SIA/SERIE A. Estudios sobre población No. 16, Nueva York, 1954.
- (2) NACIONES UNIDAS. Los recursos humanos de Centro América, Panamá y México en algunos aspectos del desarrollo económico. Estudio preparado por Louis J. Ducoff, CEPAL, 1960.
- (3) RECCHINI, Zulma L. y CHAVIRA O. Miguel. Proyección de la población de México, por sexo y grupos de edad, 1960-1980. CELADE, Santiago de Chile, 1964.
- (4) COALE, Ansley J. HOOVER, Edgar M. Population Growth and Economic Development in Low-Income Countries. Princeton University Press, 1958.
- (5) CABRERA ACEVEDO, Gustavo. La futura población de México, 1960-1980. Reproducción fotográfica, 1960
- (6) CHAVIRA OLIVOS, Miguel. Proyecciones demográficas de la República Mexicana. Secretaría de Industria y Comercio. D.G. Estadística, México, 1966
- (7) DURON OCHOA, Julio. La explosión demográfica, México. Cincuenta años de Revolución. II La vida social, Fondo de Cultura Económica, México, 1962

- (8) RECCHINI L., Zulma. Proyección de la población de México por sexo y grupos de edades, 1960-1980, CELADE, Edición mimeográfica, Chile, 1963
- (9) BENITEZ ZENTENO, Raúl y CABRERA ACEVEDO, Gustavo. Proyección de la población de México, 1960-1980, Banco de México S. A., México, 1966
- (10) SECRETARIA DE PROGRAMACION Y PRESUPUESTO. Proyecciones de la población, México, 1970-2000 (Nivel Nacional. S.P.P., México, 1978)
- (11) CAMPOSORTEGA CRUZ, Sergio. Proyecciones de la población mexicana, 1970-2040. El Colegio de México, Tesis de Maestría, México, 1989. En el cuadro sólo se presentan tres de las diez proyecciones que elabora en su tesis. La hipótesis máxima, la denominada y la hipótesis mínima
- (12) SECRETARIA DE PROGRAMACION Y PRESUPUESTO, CONSEJO NACIONAL DE POBLACION Y CENTRO LATINOAMERICANO DE DEMOGRAFIA. México, Estimaciones y Proyecciones de población, 1950-2000. Secretaría de Programación y presupuesto, México, 1983.
- (13) NUÑEZ F. Leopoldo, MORENO N., Lorenzo. México. Proyecciones de población urbana y rural, 1880-2010, Academia Mexicana de Investigación en Demografía Médica, México, 1986
- (14) INSTITUTO NACIONAL DE ESTADISTICA, GEOGRAFIA E INFORMATICA (INEGI) y CONSEJO NACIONAL DE POBLACION (CONAPO). Proyecciones de la población de México y de las Entidades Federativas: 1980-2010 Secretaría de Programación y Presupuesto, México, 1985
- (15) CENTRO LATINOAMERICANO DE DEMOGRAFIA (CELADE). Proyecciones de población vigentes, y NACIONES UNIDAS, World Population Prospects: The 1992 Revision, Nueva York, 1992
- (16) CONSEJO NACIONAL DE POBLACION (CONAPO). Proyecciones de la población de México, inédita, México, 1993

Técnicas de estimación y proyecciones de población

Gregory-King (1696), Petty y Graunt	Utilizaban el tiempo de duplicación de la población
Malthus (1798)(Geométrica y exponencial)	$P(t) = P(0)(1+r)^t$ y $P(t) = P(0)e^{rt}$
Verhulst (1838) y Pearl y Reed (1920)(logística)	$P(t) = (k/(1 + e^{b-at}))$
Farr (1873) y Cannan (1895)(Método de los componentes)	Consiste en aplicar hipótesis sobre la fecundidad, la mortalidad y la migración. Cada grupo de edades se proyecta con probabilidades de supervivencia.
Pritchett (1891)(Polinomio de tercer grado)	$P(t) = a + bt + ct^2 + dt^3$
Whelpton (1936) y Leslie (1945)(Método de los componentes perfeccionado)	Leslie lo desarrolla matricialmente
Lineal	$P(t) = P(0) + (P(n) - P(0))(t/n)$
Mattelart (1964)(Métodos numéricos)	$P(t) = P(0) + C_2^t \Delta_0 + \dots + C_k^t \Delta_0^k + \dots$
Mattelart (1964)(Regresión lineal)	$P(t) = a + bt + u$
Mattelart (1964) (Regresión polinomial)	$P(t) = a + bt + ct^2 + dt^3 + \dots + u$
Mattelart (1964) (Gompertz (1825))	$P(t) = a + b^{c^t}$
Mattelart (1964) (Makeham (1860))	$P(t) = a + bc^t$
Rogers (1968)(Modelo de crecimiento multirregional)	Considera las 4 variables demográficas que determinan el crecimiento de una región y toma en cuenta en forma conjunta la población del país dividida en regiones.
División de Población de Naciones Unidas (1989) (Método de los componentes computarizado.)	Se desarrollan las etapas de computación para ponerlo a disposición de los países.

II. Aspectos teóricos del Filtro de Kalman

En este capítulo se presentan los principales resultados teóricos del Filtro de Kalman a través de los enfoques de minimización del error cuadrático medio, estadística Bayesiana y programación dinámica.

II.1 Filtro de Kalman vía la minimización del error cuadrático medio

Considere el sistema dinámico descrito por la siguiente ecuación en diferencias de una cantidad no observable, X_k , la cual se conoce como ecuación del sistema

$$X_{k+1} = \phi_k X_k + W_k . \quad (\text{II-1})$$

La observación (o medición) del proceso se supone que ocurre en tiempos discretos de acuerdo a la siguiente relación lineal:

$$Z_k = H_k X_k + V_k , \quad (\text{II-2})$$

donde X_k es un vector de dimensión $n \times 1$ del proceso de estado en el tiempo t_k , ϕ_k es una matriz (de $n \times n$) que relaciona X_k con X_{k+1} , W_k es un vector

(de $n \times 1$) ruido blanco¹, Z_k es un vector (de $m \times 1$) de observaciones en t_k , H_k es una matriz (de $m \times n$) que establece la conexión entre la medición y el vector de estado en el tiempo t_k y V_k es un error de medición.

El estado del sistema en t_k está dado por el vector X_k de dimensión n , con $t_0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, las t 's no se encuentran necesariamente equidistantes.

Se tienen los siguientes supuestos:

$$E\{W_k W_i^T\} = \begin{cases} Q_k, & \text{para } i = j, \\ 0, & \text{para } i \neq j, \end{cases} \quad (\text{II-3})$$

y

$$E\{V_k V_i^T\} = \begin{cases} R_k, & \text{para } i = j, \\ 0, & \text{para } i \neq j. \end{cases} \quad (\text{II-4})$$

Las matrices Q_k y R_k son no negativas definidas. Además, se supone que los procesos V_k y W_k son no correlacionados, es decir,

$$E\{V_k W_i^T\} = 0, \text{ para toda } k \text{ e } i. \quad (\text{II-5})$$

Sea X_0 un estado inicial el cual es considerado un vector aleatorio. Se supone que se tiene una estimación inicial del proceso en algún punto del tiempo, digamos t_k , y esta estimación se basa en nuestro conocimiento acerca del proceso a priori a t_k . Esta estimación a priori se denota por \hat{X}_k^- . El signo menos indica que es nuestra mejor estimación a priori a la medición en t_k . Definimos el error de estimación como

$$e_k^- = X_k - \hat{X}_k^-. \quad (\text{II-6})$$

Por tanto, la matriz de covarianza es:

$$P_k^- = E(e_k^- e_k^{-T}) = E[(X_k - \hat{X}_k^-)(X_k - \hat{X}_k^-)^T]. \quad (\text{II-7})$$

¹Ruido blanco está definido como un proceso estocástico estacionario el cual tiene una función de densidad espectral constante. El término "blanco" proviene de la óptica donde la luz blanca contiene todas las frecuencias visibles, con estructura de covarianza conocida.

Con base en la estimación a priori \hat{X}_k^- se busca ahora utilizar la observación Z_k para mejorar la estimación previa. Es posible escribir \hat{X}_k en función de \hat{X}_k^- y del error de medición de acuerdo a la siguiente relación

$$\hat{X}_k = \hat{X}_k^- + K_k(Z_k - H_k\hat{X}_k^-), \quad (\text{II-8})$$

donde \hat{X}_k es igual a la estimación actualizada. La matriz K_k es conocida como la matriz de ganancia, todavía a ser determinada.

Ahora formamos la matriz de covarianza asociada a la estimación a posteriori (o estimación actualizada)

$$P_k = E[e_k e_k^T] = E[(X_k - \hat{X}_k)(X_k - \hat{X}_k)^T]. \quad (\text{II-9})$$

Sustituyendo la ecuación (II-2) en la (II-8) y luego sustituyendo en la (II-9) resulta lo siguiente:

$$P_k = E[(X_k - \hat{X}_k^-) - K_k(H_k X_k + V_k - H_k \hat{X}_k^-) \times \\ [(X_k - \hat{X}_k^-) - K_k(H_k X_k + V_k - H_k \hat{X}_k^-)]^T]. \quad (\text{II-10})$$

Calculando la esperanza y observando que el error de estimación a priori no está correlacionado con el error de medición, esto es $E[(X_k - \hat{X}_k^-) V_k] = 0$, se tiene que

$$P_k = (I - K_k H_k) P_k^- (I - K_k H_k)^T + K_k R_k K_k^T. \quad (\text{II-11})$$

Esta expresión es la matriz del error de covarianza actualizada y se aplica para cualquier matriz de ganancia K_k .

Regresando al problema de optimización, interesa ahora minimizar los términos de la diagonal de P_k , puesto que estos representan las estimaciones de las varianzas de los errores de los elementos del vector de estado. Vamos a expandir la ecuación (II-11), eliminando los subíndices tenemos que

$$P = P^- - K H P^- - P^- H^T K^T + K (H P^- H^T + R) K^T. \quad (\text{II-12})$$

Deseamos minimizar la traza de P porque es la suma del error cuadrático medio. Por lo tanto, procederemos ahora a diferenciar la traza de P con respecto a K y es importante hacer notar que la traza de $P^{-1}H^TK^T$ es igual a la traza de su transpuesta KHP^{-1} , así que

$$(d(\text{traza } P^{-1}H^TK^T)/dK) = -2(HP^{-1})^T + 2K(HP^{-1}H^T + R). \quad (\text{II-13})$$

Después de derivar igualamos a cero la expresión, resolvemos para el valor de K y obtenemos que

$$K_k = P_k^{-1}H_k^T(H_kP_k^{-1}H_k^T + R_k)^{-1}. \quad (\text{II-14})$$

Esta particular K_k , que minimiza el error cuadrático medio es llamada la matriz de ganancia (en inglés Kalman gain).

La matriz de covarianza P_k asociada con la estimación óptima puede calcularse de la siguiente forma:

$$P_k = (I - K_kH_k)P_k^{-1}(I - K_kH_k)^T + K_kR_kK_k^T, \quad (\text{II-15})$$

o equivalentemente

$$P_k = P_k^{-1} - K_kH_kP_k^{-1} - P_k^{-1}H_k^TK_k^T + K_k(H_kP_k^{-1}H_k^T + R_k)K_k^T, \quad (\text{II-16})$$

Sustituyendo (II-14) en (II-16) se tienen tres formas diferentes de P_k :

$$P_k = P_k^{-1} - P_k^{-1}H_k^T(H_kP_k^{-1}H_k^T + R_k)^{-1}H_kP_k^{-1}, \quad (\text{II-17})$$

$$P_k = P_k^{-1} - K_k(H_kP_k^{-1}H_k^T + R_k)K_k^T, \quad (\text{II-18})$$

y

$$P_k = (I - K_k H_k) P_k^- , \quad (\text{II-19})$$

esta última ecuación es la más común de todas. La ecuación (II-15) es válida para cualquier matriz de ganancia, mientras que las ecuaciones (II-17), (II-18) y (II-19) son válidas únicamente para la matriz de ganancia óptima.

Ahora consideramos la extrapolación de la estimación de la variable de estado

$$\hat{X}_{k+1} = \phi_k \hat{X}_k , \quad (\text{II-20})$$

El error de la matriz de covarianza asociado a \hat{X}_{k+1} se obtiene de la siguiente forma:

$$e_{k+1} = X_{k+1} - \hat{X}_{k+1} \quad (\text{II-21})$$

$$= (\phi_k X_k + W_k) - \phi_k \hat{X}_k \quad (\text{II-22})$$

$$= \phi_k e_k - W_k \quad (\text{II-23})$$

Es importante hacer notar que W_k y e_k tienen correlaciones cruzadas iguales a cero. Así podemos escribir la expresión para P_{k+1} como:

$$P_{k+1} = E[e_{k+1}^- e_{k+1}^{-T}] \quad (\text{II-24})$$

$$= E[(\phi_k X_k + W_k)(\phi_k X_k + W_k)^T] \quad (\text{II-25})$$

$$= \phi_k P_k \phi_k^T + Q_k \quad (\text{II-26})$$

Las ecuaciones (II-8), (II-14), (II-19), (II-20) y (II-26) comprenden las ecuaciones recursivas del Filtro de Kalman.

Las etapas básicas del procedimiento del cálculo del filtro de Kalman en tiempo discreto son:

1. Calcular P_k usando P_{k-1} , ϕ_{k-1} y Q_{k-1}
2. Calcular K_k usando P_k^- (calculado en la etapa 1), H_k y R_k .
3. Calcular P_k usando K_k (calculado en la etapa 2) y P_k^- (de la etapa 1).

4. Calcular los valores sucesivos de X_k recursivamente, usando los valores calculados de K_k (de etapa 3), el estimador inicial dado \hat{X}_0 , y el insumo de datos Z_k .

Las ecuaciones que se presentan a continuación resumen la metodología.

METODOLOGIA DEL FILTRO DE KALMAN EN TIEMPO DISCRETO

ECUACION DE TRANSICION:

$$X_{k+1} = \phi_k X_k + W_k$$

$$W_k \sim N(0, Q_k)$$

ECUACION DE MEDICION U OBSERVACION:

$$Z_k = H_k X_k + V_k$$

$$V_k \sim N(0, R_k)$$

CONDICIONES INICIALES:

$$E(X_0) = \hat{X}_0$$

SUPUESTO DE NO CORRELACION:

$$E(W_k V_j^T) = 0 \quad \forall k \neq j$$

EXTRAPOLACION DE LA ESTIMACION DE ESTADO:

$$\hat{X}_{k+1}^- = \phi_k \hat{X}_k$$

EXTRAPOLACION DE LA COVARIANZA DEL ERROR:

$$P_{k+1}^- = \phi_k P_k \phi_k^T + Q_k$$

ESTIMACION DE ESTADO ACTUAL:

$$\hat{X}_k = \hat{X}_k^- + K_k (Z_k - H_k \hat{X}_k^-)$$

COVARIANZA DEL ERROR ACTUAL:

$$P_k = (I - K_k H_k) P_k^-$$

MATRIZ DE GANANCIA DE KALMAN:

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1}$$

Las últimas 4 ecuaciones es lo que se conoce comúnmente como el Filtro de Kalman, y generalizan las ecuaciones normales del método clásico de mínimos cuadrados cuando los parámetros son funciones del tiempo.

II.2 Idea intuitiva del Filtro de Kalman

A fin de presentar una explicación sencilla del Filtro de Kalman, consideremos un problema de estimación de la media de alguna constante aleatoria basada en la secuencia de mediciones de ruido. Supongamos que se quiere estimar la media muestral. Como deseamos refinar nuestra estimación después de cada nueva medición, la sucesión de mediciones se denota por X_1, X_2, \dots, X_n , donde el subíndice representa el tiempo en que fue tomada la medición. Un método es el siguiente:

1. Primera medición X_1 ; dejo X_1 y estimo la medición
 $\mu_1 = X_1$
2. Segunda medición X_2 : teniendo X_2 y X_1 se estima la medición
 $\mu_2 = (X_1 + X_2)/2$
3. Tomo la tercera medición X_3 : teniendo X_3 y además X_1 y X_2 estimo la medición como sigue:
 $\mu_3 = (X_1 + X_2 + X_3)/3$
4. Etcétera.

Este proceso es correcto. Conforme pasa el tiempo se incrementa la memoria y por tanto, el número de operaciones aritméticas se incrementan con el tiempo. Consideremos una simple variación al procedimiento anterior, que incorpora la metodología del Filtro de Kalman en donde cada nueva estimación se obtenga a partir de la vieja estimación y de la medición actual. El algoritmo es el siguiente:

1. Primera medición X_1 ; cálculo de la estimación
 $\mu_1 = X_1$
Se deja μ_1 y se elimina X_1 .
2. Segunda medición X_2 ; Calcular la estimación como un promedio ponderado de la suma de la estimación anterior de μ_2 y de la medición actual X_2 :
 $\mu_2 = 1/2\mu_1 + 1/2X_2$
Guardar μ_2 y desechar X_2 y μ_1 .

3. Tercera medición X_3 . Calcular la estimación como un promedio ponderado de la suma de μ_2 y X_3 :

$$\mu_3 = 2/3\mu_2 + 1/3X_3$$

Guardar μ_3 y desechar X_3 y μ_2 .

4. Etcétera. En la etapa n ésima la suma ponderada es:

$$\mu_n = ((n-1)/n)\mu_{n-1} + 1/nX_n$$

Este algoritmo es un procedimiento de tipo recursivo. En este tipo de procedimiento se utilizan los resultados de la etapa previa. Esta es una característica del Filtro de Kalman.

En los modelos de espacio de estado, el objetivo central es estimar la señal en presencia del ruido. Deseamos estimar un vector X_k . El Filtro de Kalman proporciona un conjunto de ecuaciones que permiten estimar X_k y actualizar dicha estimación, cuando se cuenta con una nueva información.

II.3 Filtro de Kalman vía el enfoque Bayesiano

Sean Z_k, Z_{k-1}, \dots, Z_1 los datos observados en los tiempos $k, k-1, \dots, 1$, respectivamente, los cuales pueden ser escalares o vectores. Supongamos que Z_k depende de una cantidad no observable X_k . Nuestro propósito es hacer inferencia acerca de X_k , el cual puede ser un escalar o un vector.

La relación entre Z_k y X_k es lineal y se especifica por la ecuación de observación:

$$Z_k = H_k X_k + V_k, \quad (\text{II-27})$$

donde H_k es una matriz conocida. El vector V_k llamado error de observación se supone normalmente distribuido con media cero y varianza conocida R_k , y se denota de la siguiente manera: $V_k \sim N(0, R_k)$.

Se incorpora una ecuación de transición del sistema,

$$X_k = \phi_k X_{k-1} + W_k, \quad (\text{II-28})$$

donde ϕ_k es conocido y el error de la ecuación del sistema se distribuye normalmente: $W_k \sim N(0, Q_k)$ con Q_k conocida.

Las matrices H_k, ϕ_k, V_k y W_k pueden cambiar con el tiempo. Además, se supone que V_k y W_k son no correlacionados entre sí.

La idea del Filtro de Kalman es hacer inferencias acerca de X_k lo cual se puede realizar utilizando el Teorema de Bayes

$$\Pr(\text{estado natural}(X) | \text{datos}(Z)) \propto \Pr(\text{datos}(Z) | \text{estado natural}(X)) \quad (\text{II-29})$$

La fórmula anterior se puede escribir de la siguiente manera:

$$\Pr(X_k | Z_k) \propto \Pr(Z_k | X_k, Z_{k-1}) \Pr(X_k | Z_{k-1}) \quad (\text{II-30})$$

donde el lado izquierdo de la ecuación se refiere a la distribución a posteriori para X en el tiempo k , mientras que la primera y la segunda expresión

del lado derecho denotan la verosimilitud y la distribución a priori para X , respectivamente.

Lo anterior puede ser explicado en forma simple si nos ubicamos en $k - 1$. En $k - 1$, nuestro estado de conocimiento acerca de X_{k-1} lo establecemos en la siguiente proposición

$$(X_{k-1}|Z_{k-1}) \sim N(\hat{X}_{k-1}, P_{k-1}) \quad (\text{II-31})$$

donde \hat{X}_{k-1} y P_{k-1} son la media y la varianza de $(X_{k-1}|Z_{k-1})$, esto es,

$$E(X_{k-1}|Z_{k-1}) = \hat{X}_{k-1} \quad (\text{II-32})$$

y

$$\text{Var}(X_{k-1}|Z_{k-1}) = P_{k-1} \quad (\text{II-33})$$

La ecuación (II-31) representa a la distribución a posteriori de X_{k-1} . El análisis de k se realiza en dos etapas:

- 1) Antes de observar Z_k y
- 2) Después de observar Z_k .

Etapas 1

Antes de observar Z_k , nuestra mejor elección para X_k está gobernada por la ecuación (II-28) y esta dada por $\phi_k X_{k-1} + W_k$. Puesto que X_{k-1} esta descrito por (II-31) nuestro estado de conocimiento acerca de X_k esta basado en

$$(X_k|Z_{k-1}) \sim N(\phi_k \hat{X}_{k-1}, P_k = \phi_k P_{k-1} \phi_k^T + Q_k) \quad (\text{II-34})$$

que es nuestra distribución a priori.

Etapas 2

Al observar Z_k , nuestro objetivo es calcular la distribución a posteriori de X_k usando (II-30). Para eso necesitamos conocer la verosimilitud $L(X_k, Z_k)$ o su equivalente $\Pr(Z_k|X_k, Z_{k-1})$.

Sea e_k el error de predicción de Z_k , en el punto $k - 1$, as:

$$e_k = Z_k - \hat{Z}_k = Z_k - H_k \phi_k \hat{X}_{k-1}, \quad (\text{II-35})$$

donde H_k , ϕ_k y \tilde{X}_{k-1} son matrices conocidas. Observar Z_k es equivalente a observar e_k . As pues (II-30) puede reescribirse de la siguiente forma:

$$P(X_k|Z_k, Z_{k-1}) = P(X_k|e_k, Z_{k-1})\alpha P(e_k|X_k, Z_{k-1})P(X_k, |Z_{k-1}) \quad (\text{II-36})$$

donde $P(X_k|e_k, Z_{k-1})$ es la verosimilitud.

Usando el hecho de que $Z_k = H_k X_k + V_k$, el error e_k puede reescribirse como:

$$e_k = H_k(X_k - \phi_k \tilde{X}_{k-1}) + V_k. \quad (\text{II-37})$$

Así que

$$E(e_k|X_k, Z_{k-1}) = H_k(X_k - \phi_k \tilde{X}_{k-1}), \quad (\text{II-38})$$

entonces

$$(e_k|X_k, Z_{k-1}) \sim N(H_k(X_k - \phi_k \tilde{X}_{k-1}), R_k). \quad (\text{II-39})$$

Usando el Teorema de Bayes se tiene

$$P(X_k|Z_k, Z_{k-1}) = (P(e_k|X_k, Z_{k-1})P(X_k|Z_{k-1})) / \left(\int P(e_k, X_k|Z_{k-1})dX \right). \quad (\text{II-40})$$

Es bastante tedioso obtener $P(X_k|Z_k)$ usando (II-40). Esto puede ser evitado usando algunos resultados de estadística multivariada. Usando el resultado que aparece al final de la sección, tenemos

$$(X_k|Z_{k-1}) \sim N(\phi_k \tilde{X}_{k-1}, R_k).$$

Puesto que $(e_k|X_k, Z_{k-1}) \sim N(H_k(X_k - \phi_k \hat{X}_{k-1}), R_k)$ y si reemplazamos Y_1 , Y_2 , μ_1 y Σ_{22} por e_k , X_k , $\phi_k \hat{X}_{k-1}$ y P_k respectivamente, entonces tenemos que:

$$(e_k|X_k, Z_{k-1}) \sim N(H_k(X_k - \phi_k \hat{X}_{k-1}), R_k) \Rightarrow$$

$$\mu_1 = \Sigma_{12}^{-1}(X_k - \phi_k \hat{X}_{k-1}) \Leftrightarrow H_k(X_k - \phi_k \hat{X}_{k-1}) \quad (\text{II-41})$$

as que $\mu_1 = 0$ y $\Sigma_{12} = H_k P_k$. En forma semejante tenemos para la varianza lo siguiente

$$\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} = \Sigma_{11} - H_k P_k H_k^T = R_k \Rightarrow \Sigma_{11} = R_k + H_k P_k H_k^T. \quad (\text{II-42})$$

Por lo tanto

$$\left(\begin{pmatrix} X_k \\ e_k \end{pmatrix} \middle| Z_{k-1} \right)$$

se distribuye normalmente con la siguiente media

$$\begin{pmatrix} \phi_k \hat{X}_{k-1} \\ 0 \end{pmatrix}$$

y con matriz de varianza-covarianza

$$\begin{pmatrix} P_k & P_k H_k^T \\ H_k P_k & R_k + H_k P_k H_k^T \end{pmatrix}.$$

Utilizando el resultado que se presenta al final de la sección tenemos que

$$(X_k|e_k, Z_{k-1}) \sim$$

$$N(\phi_k \hat{X}_{k-1} + P_k H_k^T (R_k + H_k P_k H_k^T)^{-1} e_k, P_k - P_k H_k^T (R_k + H_k P_k H_k^T)^{-1} H_k P_k) \quad (\text{II-43})$$

Es decir, $(X_k|e_k, Z_{k-1})$ tiene media

$$\hat{X}_k = \phi_k \hat{X}_{k-1} + P_k H_k^T (R_k + H_k P_k H_k^T)^{-1} e_k, \quad (\text{II-44})$$

y varianza

$$P_k - P_k H_k^T (R_k + H_k P_k H_k^T)^{-1} H_k P_k. \quad (\text{II-45})$$

Estas ecuaciones son las mismas que las obtenidas en la sección anterior.

A la matriz

$$P_k H_k^T (R_k + H_k P_k H_k^T)^{-1} = K_k$$

se le conoce como la matriz de ganancia.

En el análisis anterior se ha usado el siguiente resultado:

Si el vector

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$

se distribuye normalmente con media

$$\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

y matriz de varianza-covarianza

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

entonces,

$$(Y_1|Y_2) \sim N(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(Y_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}). \quad (\text{II-46})$$

También, se sabe que (II-46) y $Y_2 \sim N(\mu_2, \Sigma_{22})$ implican que

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$

se distribuye normalmente con media

$$\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

y matriz de varianza-covarianza

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

y por lo tanto

$$(Y_2|Y_1) \sim N(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(Y_1 - \mu_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}).$$

Hemos utilizado estas últimas implicaciones en el desarrollo de esta sección.

II.4 El Filtro de Kalman y su relación con la programación dinámica

La metodología del Filtro de Kalman puede ser aplicada en problemas donde la información de estado es imperfecta, el cual puede trabajarse a través de la programación dinámica. Cuando las variables de estado son inaccesibles o los elementos que se usan para medirlas son inexactos nos encontramos ante la posibilidad de utilizar el Filtro de Kalman. Además, para representar de mejor manera la realidad, el modelo se trabaja en forma probabilística. Esta situación se modela suponiendo que en cada etapa se reciben algunas observaciones acerca del valor del estado actual, el cual puede ser afectado por una variable estocástica. Matemáticamente la observación Z_k obtenida en la etapa k es una función de la forma:

$$Z_k = h_k(X_k, u_{k-1}, V_k),$$

donde V_k es un error aleatorio y u_{k-1} es un control no restringido.

Considere un modelo conformado con un sistema lineal que cuenta con información de estado imperfecta y con costos cuadráticos.

El sistema lineal es el siguiente:

$$X_{k+1} = \phi_k X_k + \Gamma_k u_k + W_k, \quad (\text{II-47})$$

y el costo cuadrático es

$$E \left\{ X_N^T S_N X_N + \sum_{k=0}^{N-1} (X_k^T S_k X_k + u_k^T T_k u_k) \right\}. \quad (\text{II-48})$$

Cuando se incorpora el término de control $\Gamma_k u_k$, es natural preguntarse, cómo el sistema compuesto por el control y por la estimación puede ser conjuntamente óptimo.

En estas expresiones X_k y u_k son vectores de dimensión m y n , respectivamente. Las matrices ϕ_k , Γ_k , S_k y T_k son conocidas y tienen una dimensión apropiada.

Se supone que S_k es una matriz simétrica, positiva semidefinida, al igual que T_k . Los errores W_k son vectores aleatorios no correlacionados y no dependen de X_k ni del control u_k .

Se considera que al inicio del período k , se recibe una información de la forma

$$Z_k = H_k X_k + V_k, \quad (\text{II-49})$$

donde $k = 0, 1, 2, \dots, N - 1$, $Z_k \in R^S$, H_k es una matriz conocida de $s \times m$ y $V_k \in R^S$ es el vector de error de la ecuación de observación. En este modelo, como en los capítulos anteriores, V_k y W_k son no correlacionados y las matrices ϕ_k y Γ_k son conocidas.

Las ecuaciones de la programación dinámica están dadas por:

Condición de frontera

$$J_{N-1}(I_{N-1}) = \min_{u_{N-1}} \{E \{g_N(f_{N-1}(X_{N-1}, u_{N-1}, W_{N-1})) + g_{N-1}(X_{N-1}, u_{N-1}, W_{N-1}) | I_{N-1}, u_{N-1}\}\} \quad (\text{II-50})$$

Ecuaciones recursivas

$$J_k(I_k) = \min_{u_k} \{E \{g_k(X_k, u_k, W_k) + J_{k+1}(I_k, Z_{k+1}, u_k) | I_k, u_k\}\}, \quad (\text{II-51})$$

donde I_k es un vector de la información disponible,

$$I_k = (Z_0, Z_1, \dots, Z_k, u_0, u_1, u_2, \dots, u_{k-1}).$$

Para el valor inicial se tiene lo siguiente:

$$I_0 = Z_0.$$

Tenemos por definición que

$$I_{k+1} = (I_k, Z_{k+1}, u_k).$$

La ley de control óptimo es la siguiente:

$\{\mu_0^*, \mu_1^*, \mu_2^*, \dots, \mu_{N-1}^*\}$ y se resuelve minimizando la ecuación (II-50). Se calcula en forma recursiva hasta $J_0(I_0) = J_0(Z_0)$ donde el costo óptimo está dado por:

$$J^* = E \{J_0(Z_0)\}$$

De la ecuaciones (II-48) y (II-50) se tiene:

$$J_{N-1}(I_{N-1}) = \min_{u_{N-1}} \left\{ E \left\{ (\phi_{N-1}X_{N-1} + \Gamma_{N-1}u_{N-1} + W_{N-1})^T S_N \right. \right. \\ \left. \left. (\phi_{N-1}X_{N-1} + \Gamma_{N-1}u_{N-1} + W_{N-1}) + X_{N-1}^T S_{N-1} X_{N-1} \right. \right. \\ \left. \left. + u_{N-1}^T T_{N-1} u_{N-1} | I_{N-1} \right\} \right\}. \quad (\text{II-52})$$

Por hipótesis sabemos que:

$$E \{W_{N-1} | I_{N-1}\} = E \{W_{N-1}\} = 0.$$

Por tanto se tiene

$$J_{N-1}(I_{N-1}) = E \left\{ X_{N-1}^T (\phi_{N-1}^T S_N \phi_{N-1} + S_{N-1}) X_{N-1} | I_{N-1} \right\} \\ + E \left\{ W_{N-1}^T S_N W_{N-1} \right\} + \min \left\{ u_{N-1}^T (\Gamma_{N-1}^T S_N \Gamma_{N-1} + T_{N-1}) u_{N-1} \right. \\ \left. + 2E(X_{N-1} | I_{N-1})^T \phi_{N-1} S_N \Gamma_{N-1} u_{N-1} \right\}. \quad (\text{II-53})$$

La minimización produce la ley de control óptimo para la última etapa, la cual se obtiene derivando respecto a u_{N-1} ,

$$u_{N-1}^* = \mu_{N-1}^*(I_{N-1}) \\ = -(\Gamma_{N-1}^T S_N \Gamma_{N-1} + T_{N-1})^{-1} \Gamma_{N-1}^T S_N \phi_{N-1} E(X_{N-1} | I_{N-1}). \quad (\text{II-54})$$

Sustituyendo la última ecuación en (II-53), tenemos que

$$J_{N-1}(I_{N-1}) = E(X_{N-1}^T K_{N-1} X_{N-1} | I_{N-1}) + \\ E \left\{ [X_{N-1} - E(X_{N-1} | I_{N-1})]^T P_{N-1} [X_{N-1} - E(X_{N-1} | I_{N-1})] \right\} \quad (\text{II-55})$$

$$+E(W_{N-1}^T S_N W_{N-1})$$

donde las matrices K_{N-1} y P_{N-1} estan dadas por

$$P_{N-1} = \phi_{N-1}^T S_N \Gamma_{N-1} (T_{N-1} + \Gamma_{N-1}^T S_N \Gamma_{N-1})^{-1} \Gamma_{N-1}^T S_N \phi_{N-1} \quad (\text{II-56})$$

y

$$K_{N-1} = \phi_{N-1}^T S_N \phi_{N-1} - P_{N-1} + S_{N-1} . \quad (\text{II-57})$$

Ahora bien, la ecuación recursiva de programación dinámica para el período $N - 2$ es igual a

$$J_{N-2}(I_{N-2}) = \min_{u_{N-2}} \left\{ E \left\{ X_{N-2}^T S_{N-2} X_{N-2} + u_{N-2}^T T_{N-2} u_{N-2} + J_{N-1}(I_{N-1} | I_{N-2}, u_{N-2}) \right\} \right\} \quad (\text{II-58})$$

o equivalentemente

$$\begin{aligned} J_{N-2}(I_{N-2}) = & \min_{u_{N-2}} \left\{ E \left\{ X_{N-2}^T S_{N-2} X_{N-2} + u_{N-2}^T T_{N-2} u_{N-2} \right. \right. \\ & + (\phi_{N-2} X_{N-2} + \Gamma_{N-2} u_{N-2} + W_{N-2})^T K_{N-1} \\ & \left. \left. + (\phi_{N-2} X_{N-2} + \Gamma_{N-2} u_{N-2} + W_{N-2} | I_{N-2}) \right\} \right\} \\ & + E \left\{ [X_{N-1} - E(X_{N-1} | I_{N-1})]^T P_{N-1} [X_{N-1} - E(X_{N-1} | I_{N-1}) | I_{N-2}, u_{N-2}] \right. \\ & \left. + E(W_{N-1}^T S_N W_{N-1}) \right\} . \quad (\text{II-59}) \end{aligned}$$

Resolviendo para u_{N-2} tenemos

$$u_{N-2}^* = \mu_{N-2}^*(I_{N-2})$$

$$= -(T_{N-2} + \Gamma_{N-2}^T K_{N-2} \Gamma_{N-2})^{-1} \Gamma_{N-2}^T K_{N-1} \phi_{N-2} E(X_{N-2} | I_{N-2}) \quad (\text{II-60})$$

La ley de control óptimo para la etapa k es

$$\mu_k^*(I_k) = L_k E(X_k | I_k), \quad (\text{II-61})$$

donde

$$L_k = -(T_k + \Gamma_k^T K_{k+1} \Gamma_k)^{-1} \Gamma_k^T K_{k+1} \phi_k. \quad (\text{II-62})$$

La matriz L_k es llamada la matriz de ganancia. La matriz K_k se obtiene en forma recursiva mediante la ecuación de Riccati.

$$K_N = S_N$$

y

$$K_k = \phi_k^T (K_{k+1} - K_{k+1} \Gamma_k (T_k + \Gamma_k^T K_{k+1} \Gamma_k)^{-1} \Gamma_k^T K_{k+1}) \phi_k + S_k \quad (\text{II-63})$$

Uno de los problemas está en construir $E(X_k | I_k)$. Una alternativa es considerar que W_k y V_k y el vector inicial X_0 son vectores Gaussianos. La estimación de estado presenta la relación siguiente

$$\hat{X}_k = X_k^- + K_k (Z_k - H_k \hat{X}_k^-), \quad (\text{II-64})$$

donde

$$\hat{X}_k^- = (\phi_k + \Gamma_k L_{k-1}) \hat{X}_{k-1}. \quad (\text{II-65})$$

Si

$$\Omega_k = \phi_k + \Gamma_k L_{k-1} , \quad (\text{II-66})$$

la ecuación de estimación puede reescribirse como

$$\hat{X}_k = \Omega_k X_{k-1} + K_k \{ Z_k - H_k(\Omega_k \hat{X}_{k-1}) \} . \quad (\text{II-67})$$

Observe que las ecuaciones obtenidas en esta sección son las mismas a las obtenidas en las secciones anteriores.

III. Extensión, divergencia y criterio de inicio en el Filtro de Kalman. Sistemas dinámicos lineales

III.1 Procesos correlacionados

A lo largo de este trabajo hemos considerado que el ruido W_k de la ecuación de transición y V_k de la ecuación de observación no se encuentran correlacionados. Este es un supuesto razonable en muchos problemas, pero no siempre se cumple. ¿Cómo se modificarían los resultados en el caso de que los errores de ambas ecuaciones estuvieran correlacionados? En este capítulo daremos respuesta a esta pregunta.

Presentemos las ecuaciones siguientes, ya vistas en capítulos anteriores:

$$X_{k+1} = \phi_k X_k + W_k \quad (\text{III-1})$$

y

$$Z_k = H_k X_k + V_k, \quad (\text{III-2})$$

donde

$$E\{W_k W_i^T\} = \begin{cases} Q_k, & \text{para } i = k, \\ 0, & \text{para } i \neq k, \end{cases} \quad (\text{III-3})$$

y

$$E\{V_k V_i^T\} = \begin{cases} R_k, & \text{para } i = k, \\ 0, & \text{para } i \neq k. \end{cases} \quad (\text{III-4})$$

Supongamos que W y V están correlacionados de la siguiente manera

$$E\{W_{k-1} V_k^T\} = C_k. \quad (\text{III-5})$$

Reescribiendo la ecuación (III-1) en función de $k - 1$, se tiene

$$X_k = \phi_{k-1} X_{k-1} + W_{k-1}. \quad (\text{III-6})$$

Empezaremos con la ecuación actualizada

$$\hat{X}_k = \hat{X}_k^- + K_k(Z_k - H_k \hat{X}_k^-). \quad (\text{III-7})$$

Ahora utilizamos la ecuación del error

$$e_k = X_k - \hat{X}_k \quad (\text{III-8})$$

$$= X_k - \hat{X}_k = X_k - (\hat{X}_k^- + K_k(Z_k - H_k \hat{X}_k^-)) \quad (\text{III-9})$$

$$= (I - K_k H_k) e_k^- - K_k V_k. \quad (\text{III-10})$$

La esperanza entre e_k^- y V_k está dada por

$$E(e_k^- V_k^T) = E((X_k - \hat{X}_k^-) V_k^T), \quad (\text{III-11})$$

o equivalentemente

$$E(e_k^- V_k^T) = E((\phi_{k-1} X_{k-1} + W_{k-1} - \phi_{k-1} \hat{X}_k^-) V_k^T) . \quad (\text{III-12})$$

Por lo tanto,

$$E(e_k^- V_k^T) = E(e_k^- V_k^T) = C_k , \quad (\text{III-13})$$

y tenemos que

$$P_k = E(e_k e_k^T) \quad (\text{III-14})$$

$$= E((I - K_k H_k) e_k^- - K_k V_k)((I - K_k H_k) e_k^- - K_k V_k)^T , \quad (\text{III-15})$$

ó

$$P_k = (I - K_k H_k) P_k^- (I - K_k H_k)^T + K_k V_k K_k^T \quad (\text{III-16})$$

$$- (I - K_k H_k) C_k K_k^T - K_k C_k^T (I - K_k H_k)^T ,$$

y

$$K_k = (P_k^- H_k^T + C_k)(H_k P_k^- H_k^T + R_k + H_k C_k + C_k^T H_k^T)^{-1} , \quad (\text{III-17})$$

donde P_k se puede reescribir de la siguiente manera:

$$P_k = P_k^- - K_k (H_k P_k^- H_k^T + R_k + H_k C_k + C_k^T H_k^T) K_k^T , \quad (\text{III-18})$$

y por tanto

$$P_k = (I - K_k H_k) P_k^- - K_k C_k^T . \quad (\text{III-19})$$

Esto significa que las ecuaciones fundamentales no se modifican cuando hay una correlación cruzada entre W y V . Las ecuaciones de proyección se repiten como antes.

$$\hat{X}_{k+1}^- = \phi_k \hat{X}_k \quad (\text{III-20})$$

y

$$P_{k+1}^- = \phi_k P_k \phi_k^T + Q_k . \quad (\text{III-21})$$

Las ecuaciones presentadas en este apartado comprenden el conjunto de ecuaciones recursivas en el caso de que los errores de las ecuaciones de observación y de transición estén correlacionados.

III.2 Divergencia

Puesto que el Filtro de Kalman es recursivo, es posible entrar en un ciclo y permanecer allí indefinidamente. Existen algunos problemas prácticos por los que puede surgir la divergencia. Son tres las principales fuentes de dificultad.

Errores de redondeo

Como en cualquier procedimiento numérico, los errores de redondeo nos pueden llevar a problemas, de tal forma que el número de etapas sea muy largo. No existe una solución sencilla para esto, y cada caso tiene que ser

analizado según sus características. Afortunadamente, si el sistema es observable y el proceso de ruido conduce a las variables de estado, el Filtro de Kalman tiene un grado natural de estabilidad. En el caso de la estabilidad, una solución para el estado estable existirá para la matriz P , aun cuando el proceso no sea estacionario. Si la matriz P es afectada de su estado estable, no pierde el que sea positiva definida y por tanto tiende a regresar a la misma solución de estado estable.

Algunas técnicas que se han encontrado para prevenir los problemas de redondeo son:

- 1) Usar alta precisión aritmética.
- 2) Si los datos de medición están dispersos, formar pequeñas etapas entre mediciones para obtener P .
- 3) En la medida de lo posible es importante evitar procesos determinísticos en el modelo del Filtro de Kalman. Un ejemplo de este caso es una constante aleatoria. Esto usualmente lleva a una situación donde la matriz P se aproxima a una condición semidefinida, cuando el número de etapas se hace cada vez más grande, lo que puede llevar a la matriz P a que diverja. Una buena solución es añadir pequeñas cantidades positivas a los términos de la diagonal de la matriz Q . Esto lleva a un grado de suboptimalidad, pero que es preferible a que el filtro diverja.
- 4) Dar simetría a P y P^- en cada etapa recursiva. Esto es muy importante, sabemos que la matriz de covarianza debe ser simétrica, si es asimétrica debe deberse a un error de redondeo.

Errores de modelado (Problema de identificación)

Otro tipo de divergencia surge porque se utiliza un modelo que no es el adecuado. Si decimos que el proceso es una constante aleatoria con pendiente cero y el proceso actual tiene una pendiente diferente de cero, el filtro trataría de ajustar una curva equivocada a los datos observados. Esto puede ocurrir tanto en procesos no determinísticos como determinísticos.

Problema de observabilidad (Omisión de variables)

Existe una tercera clase de problema de divergencia, esto puede ocurrir cuando el sistema es no observable. Esto significa que hay una o más variables de estado que están ocultas a la vista del investigador. Como resultado de esto, si los procesos son no observables entonces son inestables y los errores de estimación correspondientes son también inestables. En el sentido estricto, este no es un problema del modelo, debido a que el ruido genera la mejor estimación posible bajo condiciones adversas.

Una buena solución a esta clase de divergencia es mejorar la situación de observabilidad añadiendo mediciones apropiadas para hacer el sistema completamente observable.

III.3 Criterio de inicio

Antes de presentar el criterio de inicio del proceso, se incorporarán dos conceptos importantes: el de máxima entropía y el de mínima entropía cruzada. El principio de máxima entropía proporciona un método general para inferir sobre una densidad desconocida, $\pi(\theta)$, cuando existe nueva información acerca de $\pi(\theta)$ en términos de sus esperanzas. El principio establece que de todas las densidades compatibles con la nueva información deberíamos elegir como estimador de $\pi(\theta)$, aquella que tenga la mayor entropía. Este principio es equivalente al de mínima entropía cruzada en el caso especial de espacios discretos y estimaciones iniciales uniformes. El principio de mínima entropía cruzada además de la nueva información en términos de sus esperanzas considera una estimación inicial, $p(\theta)$ de $\pi(\theta)$, y en este caso, deberíamos escoger como estimación final de $\pi(\theta)$, aquella que tenga la menor entropía cruzada.

Máxima entropía

Para encontrar un estimador a posteriori de una función de densidad desconocida, π , cuando existe información a priori $I(\theta)$, el principio de máxima entropía nos lleva a resolver el siguiente problema:

$$\max H(\pi) = - \int_{\Theta} \pi(\theta) \log \pi(\theta) d\theta,$$

sujeto a $I(\theta) = \{\Theta; 1, a_1(\theta), a_2(\theta), \dots, a_m(\theta); 1, \bar{a}_1, \bar{a}_2, \dots, \bar{a}_m\}$.

Una condición necesaria para que un estimador π^* sea máximo es que

$$\pi^*(\theta) = \exp \left\{ \lambda_0 + \sum_{k=1}^m \lambda_k a_k(\theta) \right\},$$

$$1 - \int_{\Theta} \pi^*(\theta) d(\theta) = 0,$$

$$\int_{\Theta} (\bar{a}_k - a_k(\theta)) \pi^*(\theta) d(\theta) = 0, \quad k = 1, 2, \dots, m,$$

donde $\lambda_0, \lambda_1, \dots, \lambda_m$ son los multiplicadores de Lagrange asociados a las restricciones dadas.

Sustituyendo π^* en las tres condiciones anteriores se encuentra que

$$0 = \lambda_0 - \log \left\{ \int_{\Theta} \prod_{k=1}^m e^{\lambda_k a_k(\theta)} d\theta \right\},$$

$$0 = \int_{\Theta} (a_k(\theta) - \bar{a}_k) \prod_{k=1}^m e^{\lambda_k a_k(\theta)} d\theta, \quad k = 1, 2, \dots, m,$$

el cual es un sistema lineal homogéneo con variables $\lambda_0, \lambda_1, \dots, \lambda_m$. Además, si es posible resolver la integral que determina a λ_0 en el sistema anterior, entonces el resto de los multiplicadores se puede encontrar a partir de las siguientes relaciones

$$\frac{\partial \lambda_0}{\partial \lambda_k} = -a_k, \quad k = 1, 2, \dots, m,$$

Mínima entropía cruzada

Para encontrar un estimador a posteriori de una función de densidad desconocida, π , cuando se tiene un estimador a priori, $p(\theta)$ y existe información a priori $I(\theta)$, nos lleva a resolver el siguiente problema de minimización de la entropía cruzada:

$$\min H(\pi, p) = \int_{\Theta} \pi(\theta) \log(\pi(\theta)/p(\theta)) d\theta,$$

sujeto a $I(\theta) = \{\Theta; 1, a_1(\theta), a_2(\theta), \dots, a_m(\theta); 1, \bar{a}_1, \bar{a}_2, \dots, \bar{a}_m\}$.
 Una condición necesaria para que un estimador π^* sea mínimo es que

$$\pi^*(\theta) = p(\theta) \exp \left\{ \lambda_0 + \sum_{k=1}^m \lambda_k a_k(\theta) \right\},$$

$$1 - \int_{\Theta} \pi^*(\theta) d(\theta) = 0,$$

$$\int_{\Theta} (\bar{a}_k - a_k(\theta)) \pi^*(\theta) d(\theta) = 0, \quad k = 1, 2, \dots, m.$$

Procediendo de manera semejante al caso de máxima entropía, se encuentra que

$$0 = \lambda_0 - \log \left\{ \int_{\Theta} p(\theta) \prod_{k=1}^m e^{\lambda_k a_k(\theta)} d\theta \right\},$$

$$0 = \int_{\Theta} (a_k(\theta) - \bar{a}_k) p(\theta) \prod_{k=1}^m e^{\lambda_k a_k(\theta)} d\theta, \quad k = 1, 2, \dots, m.$$

Relaciones similares al de máxima entropía se presentan en este caso.

Inicialización

Suponga que el proceso inicia en el tiempo $t = 0$ y se conocen \hat{X}_0 y $\hat{\Sigma}_0$, esto es, la media y la varianza de X_0 . Es posible usar el principio de máxima entropía a fin de encontrar una estimación de la distribución a priori de X_0 , a la que se llamará $\pi(X_0)$, que tome en cuenta información inicial. El principio de máxima entropía conduce a resolver el problema variacional

$$\max \int_{-\infty}^{\infty} \log(\pi(X_0)) \pi(X_0) dX_0, \quad (\text{III-22})$$

sujeto a

$$\int_{-\infty}^{\infty} \pi(X_0) dX_0 = 1, \quad (\text{III-23})$$

$$\int_{-\infty}^{\infty} X_0 \pi(X_0) dX_0 = \hat{X}_0, \quad (\text{III-24})$$

$$\int_{-\infty}^{\infty} (X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T \pi(X_0) dX_0 = \hat{\Sigma}_0. \quad (\text{III-25})$$

Se puede demostrar resolviendo por multiplicadores de Lagrange que

$$X_0 \sim N(\hat{X}_0, \hat{\Sigma}_0).$$

III.4 Sistemas dinámicos lineales

En esta sección estudiaremos algunas características de los sistemas dinámicos lineales.

Un sistema lineal invariante y discreto en la forma espacio-estado, es especificado por las matrices (Φ, C, D, F) y las ecuaciones:

$$X_{k+1} = \Phi X_k + D \xi_k, \quad (\text{III-26})$$

y

$$Y_k = C X_k + F \xi_k, \quad (\text{III-27})$$

donde $\Phi \in R^{n \times n}$, $C \in R^{s \times n}$, $D \in R^{n \times r}$ y $F \in R^{s \times r}$. El proceso ξ consiste de entradas no observables al sistema, ξ_k son variables aleatorias independientes e idénticamente distribuidas, con vector de media cero y matriz de covarianzas I , Y es un proceso observable o de salida, X es el estado del sistema en el tiempo k , es una variable no observable, C es una matriz de observación, que sirve como ventana para X , Φ es la matriz de transición para el sistema, define la manera como evolucionaría el sistema en ausencia de ruido y F, D son matrices que correlacionan a ξ .

REVISIÓN DE CONCEPTOS

Existen sistemas lineales en los que el vector de estados evoluciona de manera continua y el proceso se observa de manera continua o discreta.

Las dos últimas ecuaciones forman un sistema que no es el más general, pues carecen de variables de control, que pueden ser variables exógenas al sistema.

A continuación se definen los conceptos de sistema observable y controlable, que como se verá son de gran importancia en el análisis del Filtro de Kalman.

Observabilidad

El sistema formado por las ecuaciones (III-26) y (III-27) es observable, si la matriz Θ^T definida como sigue:

$$\Theta^T = (C^T, \Phi C^T, \Phi^2 C^T, \dots, \Phi^{n-1} C^T) \quad (\text{III-28})$$

es de rango completo.

El significado de la definición de observabilidad es el siguiente, si Θ^T es de rango completo, es posible resolver las n ecuaciones.

$$Y_k = C\Phi^k X_0 \text{ para } k = 0, 1, 2, \dots, n-1, \quad (\text{III-29})$$

es decir, en ausencia de ξ es posible determinar X_0 (y en consecuencia a X_n) a partir de Y_k , para $k = 0, \dots, n-1$.

Cuando el sistema no es observable, Θ^T tiene rango n_1 con $n_1 < n$, entonces existe una matriz T tal que $(T^T)^{-1}\Theta^T$ tiene $n - n_1$ renglones de ceros, los primeros n_1 renglones de T son linealmente independientes y T satisface lo siguiente

$$T\Phi T^{-1} = \begin{pmatrix} \Phi_{11} & 0 \\ \Phi_{21} & \Phi_{22} \end{pmatrix} C^T^{-1} \quad (\text{III-30})$$

y

$$= (C_1 \ 0). \quad (\text{III-31})$$

Al multiplicar por T al sistema de ecuaciones original se tiene

$$TX_{k+1} = T\Phi T^{-1}TX_k + TD\xi_k \quad (\text{III-32})$$

y

$$Y_k = CT^{-1}TX_k + F\xi_k, \quad (\text{III-33})$$

y reordenando adecuadamente el sistema, éste se transforma en

$$\begin{aligned} X_{1,k+1} &= \Phi_{11}X_{1,k} + D_1\xi_k, \\ X_{2,k+1} &= \Phi_{21}X_{1,k} + \Phi_{22}X_{2,k} + D_2\xi_k, \end{aligned} \quad (\text{III-34})$$

y

$$Y_k = C_1X_{1,k} + F\xi_k,$$

donde (Φ_{11}, C_1) es observable. De la nueva representación del sistema, se deduce que puede omitirse a $X_{2,k}$, pues no aparece en la ecuación para Y_k , es decir, el efecto de ξ sólo se observa a través de $X_{1,k}$.

Controlabilidad

El sistema de ecuaciones (III-26) y (III-27) es controlable si la matriz Γ es igual a

$$\Gamma = (D, \ \Phi D, \ \Phi^2 D, \ \dots, \ \Phi^{n-1} D) \quad (\text{III-35})$$

y tiene rango completo.

La intuición de la definición de controlabilidad se deduce de la igualdad

$$X_{k+1} = \sum_{j=1}^{t+1} \Phi^{j-1} D \xi(k+1-j) + \Phi^{k+1} X_0. \quad (\text{III-36})$$

Si pudiéramos manipular los valores de ξ_k, \dots, ξ_0 , podríamos hacer que el sistema en el tiempo $k+1$ tome cualquier valor, mientras X_0 sea finito y conocido.

Cuando el sistema no es controlable, podemos encontrar una matriz no singular T de manera que el sistema transformado sea el siguiente:

$$\begin{aligned} X_{1,k+1} &= \Phi_{11} X_{1,k} + \Phi_{12} X_{2,k} + D_1 \xi_k, \\ X_{2,k+1} &= \Phi_{22} X_{2,k}, \end{aligned} \quad (\text{III-37})$$

donde el subvector de estados X_k no es controlable, porque su valor no puede ser modificado a través de ξ .

Otras definiciones importantes se dan a continuación.

Estabilidad

Se dice que una matriz D de $n \times n$ es estable si el $\lim_{k \rightarrow \infty} D^k = 0$.

Sistemas construibles

El sistema de ecuaciones original (III-26) y (III-27) es construible si

$$\text{rango}(\Theta^T) \supseteq \text{rango}(\Phi^{T^n}),$$

lo que significa que el conocimiento de ξ_t , nos permite conocer a X_{k+1} , donde $0 \leq t \leq k$ y $t \geq n-1$.

Detectabilidad

El sistema de ecuaciones original (III-26) y (III-27) es detectable si:

a) existe una matriz T tal que,

$$T\Phi T^{-1} = \begin{pmatrix} \Phi_{11} & 0 \\ \Phi_{21} & \Phi_{22} \end{pmatrix} CT^{-1} \quad (\text{III-38})$$

$$= (C_1 \quad 0) \quad (\text{III-39})$$

con (Φ_{11}, C_1) construible y

b) los eigenvalores de Φ_{22} son menores que 1 en valor absoluto, es decir : $\lambda(\Phi_{22}) < 1$.

La propiedad de detectabilidad se interpreta del modo siguiente:

La parte del vector que no observamos, no dependerá después de mucho tiempo de $X_{2,0}$, porque los eigenvalores de Φ_{22} son menores que 1 en valor absoluto.

Controlabilidad al origen

El sistema (III-26) y (III-27) es controlable al origen si (Φ^T, D^T) es construible, y se interpreta así:

Para sistemas controlables con un valor conocido de X_0 , existen ξ_k , tal que $X_{k+1} = 0$, donde $0 \leq t \leq k$ y $k \geq n - 1$.

Finalmente, un sistema es estabilizable si (Φ^T, D^T) es detectable.

A continuación se presentan tres teoremas, que muestran la importancia de los conceptos anteriores.

El primer teorema muestra que un sistema lineal e invariante, puede ser expresado en las formas entrada-salida o espacio-estado.

Teorema

Cualquier función de transferencia racional en z

$$Z(z) = \sum_{j=0}^{\infty} L_j z^j, \quad (\text{III-40})$$

puede ser expresada como

$$C(Iz - \Phi)^{-1}D + F,$$

donde $F = L_0$, $L_j = C\Phi^j D$, $\Phi \in R^{n \times n}$, $C \in R^{s \times n}$, $D \in R^{n \times r}$ y $F \in R^{s \times r}$, y se dice que (Φ, C, D, F) es una realización algebraica de Z . Todo esto está definido para $L_0 < \infty$.

Un concepto importante es la minimalidad de un sistema lineal.

Definición. El sistema lineal (Φ, C, D, F) es minimal, si Φ es la matriz de menor dimensión posible entre todas las realizaciones posibles de la función de transferencia $Z(z)$.

A continuación se presentan dos teoremas que nos dicen que es suficiente trabajar con los sistemas observables y controlables.

Teorema

Un sistema lineal (Φ, C, D, F) es minimal si y sólo si es observable y controlable.

Teorema

Si los sistemas lineales minimales (Φ, C, D, F) y (Φ^*, C^*, D^*, F^*) son realizaciones de la misma función de transferencia, entonces existe una matriz no singular T tal que

$$\Phi^* = T\Phi T^{-1},$$

$$C^* = CT^{-1},$$

$$D^* = TD$$

y

$$F = F^*.$$

IV. Aplicación del Filtro de Kalman a estimaciones y proyecciones de población

Las estimaciones de población son de especial relevancia en los estudios demográficos y en la planeación social y económica. El propósito de este trabajo es aplicar el Filtro de Kalman a la estimación de poblaciones a nivel de entidad federativa.

El Filtro de Kalman puede ser utilizado para estimar la población de cada una de las 32 entidades federativas de la nación, dada la población nacional año con año y dadas las poblaciones estatales en forma intermitente, digamos cada diez años, es posible estimar las poblaciones estatales anualmente. El Filtro de Kalman tiene numerosas ventajas sobre otros métodos demográficos. Estas son:

i) El modelo proporciona una adecuada descripción de la realidad, ya que los parámetros en este tipo de modelos son funciones de política y por tanto del tiempo. Las estimaciones también son óptimas en el sentido que minimizan el error cuadrático medio.

ii) El Filtro de Kalman proporciona una medida de la exactitud de las estimaciones, debido a que es posible obtener intervalos de confianza de las estimaciones de población.

iii) Una vez que los parámetros desconocidos se han estimado, los cálculos de las poblaciones se producen en forma automática e instantánea.

iv) Las estimaciones siempre satisfacen la condición de que la suma proporciona el total nacional.

v) Las estimaciones se producen con menos información que las derivadas con métodos tradicionales.

IV.1 Una primera aproximación para estimar subpoblaciones

Sea $X_{i,t}$ la población del estado i en el tiempo t . El modelo que se presenta a continuación es válido para todas las entidades federativas del país:

$$X_{i,t} = X_{i,t-1} + b_i + c_{i,t} + \xi_{i,t}^{(1)}, \quad i = 1, 2, \dots, k, \quad (\text{IV-1})$$

donde $c_{i,t}$ es el crecimiento natural (nacimientos menos defunciones) y b_i representa la participación de la entidad federativa i en la migración neta nacional.

$\xi_{i,t}^{(1)}$ es una variable aleatoria con media cero, la cual toma en consideración efectos no sistemáticos.

En este modelo se supone que la migración permanece invariable en el tiempo, esto es, que la migración neta en números absolutos en cada entidad federativa es constante. La estimación del parámetro b_i se realiza con base en las técnicas tradicionales de regresión lineal.

IV.2 Un modelo general para estimar subpoblaciones

En las secciones que se presentan a continuación se utiliza la metodología del Filtro de Kalman en las estimaciones de subpoblaciones. Los resultados de los capítulos anteriores juegan un papel esencial en el desarrollo del modelo.

Sea $X_{i,t}$ la población del estado i en el tiempo t . La dinámica de la población para las 32 entidades federativas del país esta dada por

$$X_{i,t} = X_{i,t-1} + b_{i,t-1} + c_{i,t} + \xi_{i,t}^{(1)}, \quad i = 1, 2, \dots, k, \quad (\text{IV-2})$$

donde $c_{i,t}$ es el crecimiento natural (nacimientos menos defunciones) y $b_{i,t-1}$ representa la participación de la entidad federativa i en la migración neta nacional.

La variable aleatoria $\xi_{i,t}^{(1)}$ tiene media cero y toma en consideración efectos no sistemáticos. Además, $b_{i,t}$ se desarrolla de acuerdo a un proceso estocástico de caminata aleatoria. Se ha supuesto un modelo autorregresivo de orden 1, debido a que se ha observado que la migración ocurre en etapas, es decir, primeramente migra el jefe de familia y posteriormente, migran las demás personas que integran dicha familia, es decir,

$$b_{i,t} = \mu_i + b_{i,t-1} + \xi_{i,t}^{(2)}, \quad i = 1, 2, \dots, k, \quad (\text{IV-3})$$

donde μ_i es un parámetro que representa el aumento de la migración de un período al otro.

Si se definen los siguientes vectores:

$$\alpha_t = (X_{1t}, X_{2t}, \dots, X_{kt}, b_{1t}, b_{2t}, \dots, b_{kt}), \quad (\text{IV-4})$$

$$c_t = (c_{1t}, c_{2t}, \dots, c_{kt}, 0, 0, \dots, 0), \quad (\text{IV-5})$$

$$\mu = (0, 0, \dots, 0, \mu_1, \mu_2, \dots, \mu_k), \quad (\text{IV-6})$$

y

$$\xi_t = (\xi_{1t}^{(1)}, \xi_{2t}^{(1)}, \dots, \xi_{kt}^{(1)}, \xi_{1t}^{(2)}, \xi_{2t}^{(2)}, \dots, \xi_{kt}^{(2)}), \quad (\text{IV-7})$$

entonces las ecuaciones (IV-2) y (IV-3) pueden escribirse de manera matricial de la siguiente forma:

$$\alpha_t = A\alpha_{t-1} + c_t + \mu + \xi_t, \quad (\text{IV-8})$$

donde

$$A = \begin{pmatrix} I_k & I_k \\ 0 & I_k \end{pmatrix}.$$

La ecuación (IV-8) es conocida como la ecuación de transición del sistema. Las variables X_{it} y b_{it} , y en consecuencia α_t , generalmente son variables no observadas, pero la variable c_t si lo es.

Este modelo puede generalizarse aún más si se supone que la migración neta sigue el comportamiento de un modelo autorregresivo de orden p ($AR(p)$). El modelo para la migración sera el siguiente:

$$b_{i,t} = \mu_i + \phi_1 b_{i,t-1} + \phi_2 b_{i,t-2} + \dots + \phi_p b_{i,t-p} + \xi_{it}^{(2)}.$$

Esto puede explicarse debido a que la migración neta es un fenómeno que puede ocurrir en etapas, esto es, una vez que se instala el jefe de familia, los restantes miembros que la integran migran después de un determinado período. Esto puede originar que el modelo de series de tiempo más adecuado para explicar la migración sea un autorregresivo de orden superior a 1.

A continuación estudiamos dos situaciones: el modelo para años censales y para años no censales.

(a) Años censales

Durante los años censales las poblaciones de las entidades federativas son datos observados. Utilizaremos el vector Y_t como el vector de observaciones. De esta forma, en los años censales, Y_t es un vector de k componentes, dado por

$$Y_t = [X_{1t}, X_{2t}, \dots, X_{kt}]^T. \quad (IV-9)$$

De la definición de α_t en la ecuación (IV-4), se tiene

$$Y_t = Z_t \alpha_t, \quad (IV-10)$$

donde

$$Z_t = [I_k, 0] \quad (IV-11)$$

es una matriz de $k \times 2k$.

(b) Años no censales

En los años no censales, no se cuenta con la población por entidad federativa. Solamente se conoce la población total nacional. Para estos años la ecuación esta dada por

$$Y_t = \sum_{i=1}^k X_{it} , \quad (\text{IV-12})$$

Y_t es un escalar y esta dado por la siguiente relación

$$Y_t = Z_t \alpha_t , \quad (\text{IV-13})$$

donde ahora

$$Z_t = [J_k^T, 0] \quad (\text{IV-14})$$

y J_k^T es un vector de unos. En este caso Z_t es una matriz de $1 \times 2k$.

En resumen, el sistema consiste de dos ecuaciones:

$$\alpha_t = A\alpha_{t-1} + c_t + \mu + \xi_t$$

junto con la ecuación de observación:

$$Y_t = Z_t \alpha_t,$$

donde Z_t adopta la forma de la ecuación (IV-11) en el caso de los años censales y la forma de la ecuación (IV-14) en el caso de los años no censales.

IV.3 Implementación del modelo

Considere el sistema lineal dinámico

$$\alpha_t = A_t \alpha_{t-1} + c_t + \mu + \xi_t \quad (\text{IV-15})$$

y

$$Y_t = Z_t \alpha_t + \eta_t , \quad (\text{IV-16})$$

donde (IV-15) representa la ecuación de transición y α_t es un vector de estado de $m \times 1$ y la ecuación (IV-16) relaciona la variable de observación Y_t de $n_t \times 1$ con α_t . En la ecuaciones antes señaladas c_t y Y_t son datos observados, para toda t , μ es un vector de parámetros, mientras que α_t no es información observada. Las variables aleatorias ξ_t y η_t tienen media cero y

no están mutuamente correlacionadas. Las matrices de covarianza son Q_t y H_t , respectivamente. Las matrices A_t , Z_t , Q_t , H_t y μ se suponen conocidas para toda t .

Las dimensiones de los vectores y de las matrices son las siguientes:

- α_t , c_t , μ y ξ_t son de $m \times 1$.
- Y_t y n son de $n_t \times 1$.
- A_t y Q_T son de $m \times m$.
- Z_t es de $n_t \times m$, y
- H_t es de $n_t \times n_t$.

El Filtro de Kalman da un óptimo en el sentido que minimiza el error cuadrático medio del estimador lineal $\hat{\alpha}_t$ del vector no observado α_t . Además, se obtiene su error de covarianza P_t . Se dan como conocidos $\hat{\alpha}_0$ y P_0 . Las observaciones c_1, c_2, \dots, c_t y Y_1, Y_2, \dots, Y_t son datos conocidos.

Los errores de predicción juegan un papel importante en el Filtro de Kalman, estos son:

- i) $\alpha_t - \hat{\alpha}_t$, el error de predicción óptimo. Este error tiene una matriz de covarianza P_t .
- ii) $\theta_t = \alpha_t - \hat{\alpha}_{t|t-1}$ donde $\hat{\alpha}_{t|t-1}$ es el error óptimo de predicción dadas las observaciones hasta $t-1$. Se denotará a la matriz de covarianza θ_t como $P_{t|t-1}$.
- iii) $\nu_t = Y_t - \hat{Y}_{t|t-1}$, donde $\hat{Y}_{t|t-1}$ es la predicción de Y_t que utiliza la información disponible en el tiempo $t-1$. De la ecuación (IV-16) se desprende que

$$Y_{t|t-1} = Z_t \hat{\alpha}_{t|t-1} . \quad (\text{IV-17})$$

A la matriz de covarianza de ν_t se le denotará por F_t .

Suponiendo que se conocen $\hat{\alpha}_{t-1}$ y P_{t-1} , el filtro trabaja de manera recursiva de la siguiente forma

$$P_{t-1} \Rightarrow P_{t|t-1} \Rightarrow F_t \Rightarrow P_t$$

Observe que $\hat{\alpha}_{t|t-1} = \hat{\alpha}_t^-$. Esta última es la notación utilizada en los capítulos anteriores.

$$\hat{\alpha}_{t-1} \Rightarrow \hat{\alpha}_{t|t-1} \Rightarrow \nu_t \Rightarrow \hat{\alpha}_t$$

Asimismo, $P_{t|t-1} = P_t^-$. Esta última es la notación utilizada en los capítulos anteriores.

Es importante hacer notar que en la recursividad de donde surgen P_t y P_{t-1} sólo se utilizan las matrices conocidas A_t , Q_t , Z_t y H_t . No se requiere de los valores Y_t y c_t .

Por otro lado, la recursividad de donde se obtiene $\hat{\alpha}_t$ a partir de $\hat{\alpha}_{t|t-1}$ sólo utilizan las observaciones Y_t y c_t , junto con las matrices de covarianza $P_{t|t-1}$ y F_t .

Las ecuaciones que conforman el Filtro de Kalman son las siguientes: Las matrices recursivas de covarianza:

$$P_{t|t-1} = A_t P_{t-1} A_t^T + Q_t, \quad (\text{IV-18})$$

$$F_t = Z_t P_{t|t-1} Z_t^T + H_t, \quad (\text{IV-19})$$

$$P_t = P_{t|t-1} - P_{t|t-1} Z_t^T F_t^{-1} Z_t P_{t|t-1}, \quad (\text{IV-20})$$

y las ecuaciones recursivas de predicción

$$\hat{\alpha}_{t|t-1} = A_t \hat{\alpha}_{t-1} + c_t + \mu, \quad (\text{IV-21})$$

$$\nu_t = Y_t - Z_t \hat{\alpha}_{t|t-1}, \quad (\text{IV-22})$$

$$\hat{\alpha}_t = \hat{\alpha}_{t|t-1} + G_t \nu_t, \quad (\text{IV-23})$$

donde

$$G_t = P_{t|t-1} Z_t^T F_t^{-1}. \quad (\text{IV-24})$$

Es de especial importancia el caso en que la ecuación de observación no tiene ruido, lo que implica que $H_t = 0$. Esto significa que para las estimaciones estatales estimadas para los años no censales siempre su suma es igual a la población total, y en los años censales siempre son iguales a las poblaciones estatales observadas.

Las ecuaciones recursivas de covarianza y de predicción del Filtro de Kalman presentadas en este capítulo, se pueden generalizar si se utiliza el resultado del capítulo 3 en que se hace una extensión cuando se supone que

$$E(\xi_{t-1}\eta_t^T) = C_t,$$

es decir, cuando los errores de las ecuaciones de observación y de transición están correlacionados.

V. Información imperfecta, control y políticas de población

El Filtro de Kalman resulta ser una técnica interesante para evaluar y redefinir objetivos en materia de políticas públicas. Es posible obtener la trayectoria de la población tal que se minimice la suma de las diferencias al cuadrado entre la población estimada y la población meta, sujeto a las restricciones de la ecuación de observación y la ecuación de transición del modelo del Filtro de Kalman. Antes de entrar al modelo realizaré una breve descripción de la evolución de las Políticas de Población en nuestro país.

V.1 Breve historia de las Políticas de Población en México

Es posible distinguir varias etapas en las Políticas de Población en México, las cuales cubren un período de medio siglo.

El auge poblacionista (los años cuarenta).- Hace medio siglo, la población del país era de 20 millones de habitantes. La tasa de crecimiento demográfico en 1940 era, al igual que hoy, de alrededor de 2 por ciento anual; como resultado del moderado crecimiento demográfico observado en esta época, la preocupación central de políticos y científicos sociales fue la escasez de población en el territorio nacional. En consecuencia, se promovieron acciones dirigidas a incrementar el número de habitantes, el ritmo de crecimiento de la población y la densidad demográfica. Se consideró que México era un país subpoblado, en donde el crecimiento demográfico era un requisito indispensable para el desarrollo social y económico.

Esta preocupación surgió de la gran disparidad de población existente entre México y Estados Unidos: México contaba con una población de alrededor de 20 millones de habitantes, mientras que Estados Unidos tenía poco más de 130 millones de personas.

Dado el moderado crecimiento de la población y la inadecuada distribución de ésta, se fomentó la inmigración de extranjeros, de preferencia aquellos que se adaptarían con mayor facilidad a nuestra cultura. Las acciones se dirigieron a la disminución de la mortalidad, pero conservando el nivel elevado de la natalidad, así como a la regulación de la migración. Este planteamiento se manifestó en la promulgación de la Ley General de Población de 1947, la cual tenía un carácter eminentemente pronatalista.

La incubación del cambio (los años cincuenta).- En el año de 1950 la población estimada era de 26.5 millones de habitantes y la tasa de crecimiento demográfico se incrementó a casi 3 por ciento anual.

Aunque se tenía conciencia del rápido crecimiento de la población que se venía presentando, se era optimista con respecto al desarrollo económico sustentado en los avances de la tecnología y la explotación de recursos naturales. En las esferas políticas y académicas se pensaba que así se podrían superar los problemas derivados del elevado crecimiento demográfico, por lo que, en este período, el acelerado aumento de la población no se consideraba como un obstáculo al desarrollo. Es posible que en estos años se estuviera gestando una etapa futura en la que lo demográfico sería motivo de discusión amplia.

La institucionalización de la Política Demográfica (los años sesenta).- En 1960 la población del país era ya de 36 millones de habitantes. La tasa de crecimiento se estimaba en 3.2 por ciento anual y en la mitad de los sesenta se presentaba la tasa de crecimiento más elevada del siglo.

Durante la década de los sesenta se empezó a percibir una polarización de las posiciones respecto a la conveniencia o no de un elevado crecimiento de la población. Al inicio de la década había consenso respecto a la necesidad de promover el ritmo de crecimiento demográfico, planteamiento que al poco tiempo fue discutido en diversos ámbitos; se ponía en duda la existencia de un problema de subpoblación y se empezaba a considerar la posibilidad de establecer programas orientados a regular el crecimiento de la población.

Es importante destacar que algunas personas que en el decenio de los treinta se manifestaron a favor de una política pronatalista de población, cambiaron su posición en este periodo y señalaron la necesidad de realizar acciones para enfrentar el acelerado crecimiento de la población, regular la migración y reducir la mortalidad, con el fin de que estas medidas condujeran a un mejoramiento de los niveles de vida.

El impulso a la Política de Población (los años sesenta).- En el año de 1970 la población estimada fue de 50 millones de habitantes y la tasa de crecimiento demográfico se calculó en 3.3 por ciento anual.

En 1974 fue promulgada la Ley General de Población y se creó el Consejo Nacional de Población (Conapo), con el fin de establecer las directrices de la política de población del país, la cual tiene como objetivo: "regular los fenómenos que afectan a la población en cuanto a su número, estructura, dinámica y distribución en el territorio con el fin de lograr que participe justa y equitativamente de los beneficios del desarrollo".

En 1977 el Conapo estableció objetivos y metas en materia de crecimiento demográfico. Se planteó la meta de 2.5 por ciento de crecimiento demográfico para 1982 y de 1 por ciento al año 2000, buscando establecer la congruencia entre los objetivos demográficos y los aspectos de la política de desarrollo.

La interdisciplinaria demográfica (los años ochenta).-En 1980 la población del país se estimó en casi 70 millones de habitantes y la tasa de crecimiento de la población bajó a 2.7 por ciento anual en 1980.

Frente a las transformaciones observadas en los últimos años, se suscitó un interés por examinar las características de la evolución de la población de México y de América Latina y sus relaciones con el proceso de desarrollo.

En el caso de población y políticas públicas se avanzó en la discusión del papel de la población en la planificación del desarrollo, haciéndose evidente la necesidad de entender las políticas de población en el contexto histórico de la economía y de la sociedad.

La necesidad de actualizar la Política demográfica (los años noventa).-En 1990 la población censal de México fue de 81.1 millones de habitantes y la tasa de crecimiento demográfico fue de alrededor de 2 por ciento anual, como hace 5 decenios.

Las metas que se plantearon en términos del crecimiento demográfico en 1977 no se han cumplido y este momento es oportuno para corregir el rumbo. Nuevas metas se han propuesto a la luz de nueva información demográfica. Sin embargo es necesario evaluar dichas metas a partir de técnicas que analicen su viabilidad.

V.2 Un modelo de Planeación Demográfica

Sea X_t la población nacional en el momento t , \tilde{X}_t la población meta en el momento t y Y_t la desviación entre la población nacional y la población meta en el momento t . Las matrices S y T ya han sido definidas en el capítulo 2. Es interesante analizar cuál es la trayectoria de población que minimiza la siguiente función cuadrática:

$$\min E \left\{ (X_N - \tilde{X}_N)^T S_N (X_N - \tilde{X}_N) + \sum_{k=0}^{N-1} ((X_k - \tilde{X}_k)^T S_k (X_k - \tilde{X}_k) + u_k^T T_k u_k) \right\} \quad (V-1)$$

La función objetivo expresa el deseo de que el sistema se mantenga cerca de una trayectoria deseada de población ($\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_N$) sujeta a la ecuación de transición y a la ecuación de observación, las cuales son:

$$X_{k+1} = A_k X_k + B_k u_k + W_k \quad (V-2)$$

y

$$Z_k = C_k (X_k - \tilde{X}_k) + V_k, \quad (V-3)$$

las cuales tienen las mismas características ya presentadas en capítulos anteriores, sólo que ahora se incluye un control u_k .

La función objetivo de tipo cuadrático es razonable, puesto que induce a fuertes penalizaciones en el caso que haya grandes desviaciones respecto a

la población meta y pequeñas penalizaciones si las desviaciones también son pequeñas. El propósito es encontrar un óptimo conjunto para la trayectoria de las poblaciones y de los controles.

La variable que ha estado sujeta a una política de reducción por parte del Estado Mexicano ha sido la natalidad, puesto que la migración es difícil de regular debido a que las variables sociales y económicas que la determinan no han sido explicadas en su totalidad, por lo que supondré que la población es cerrada, es decir, consideraré que la emigración y la inmigración internacional es nula. La variable mortalidad que es otro de los componentes de la dinámica demográfica ha descendido continuamente, y ahora su nivel se encuentra casi constante. Por tanto, los controles óptimos significan las medidas de planificación familiar que se tendrían que realizar para cumplir con las metas de población establecidas, es decir, indica el número de nacimientos por evitar, lo que se traduce en el número de usuarios y usuarias a cubrir en los programas de planificación familiar para alcanzar las metas de población.

La ley de control óptimo sería la siguiente:

$$\mu_k^* = L_k E(X_k),$$

donde la matriz L_k ya se ha presentado en el capítulo 2.

VI. Resultados

Antes de establecer los supuestos del modelo, es necesario conocer cuales son las acciones que el Estado Mexicano busca emprender en materia de Política de Población en lo que resta de este siglo. Entre los objetivos demográficos del Plan Nacional de Desarrollo (1995-2000) se busca impulsar la reducción del crecimiento demográfico. Se destaca que el bienestar no puede fincarse sobre bases sólidas si el rápido crecimiento de la población erosiona los beneficios del desarrollo. Se señala también, que la política de población persigue una importante intensificación de la planificación familiar. Para el año 2000 se espera un país de 100 millones de mexicanos. Entre 1995 y el año 2000 se esperan anualmente un promedio de 1.8 millones de nuevos mexicanos.

Tomando en cuenta estos propósitos, en este capítulo se presentan los resultados de las proyecciones de población utilizando el Filtro de Kalman. Se han simulado diversos escenarios bajo el supuesto de que se presenten diferentes valores de la varianza de los errores, tanto de la ecuación de observación como de la ecuación de transición. El establecimiento de las diversas hipótesis sobre las varianzas de los errores de las dos ecuaciones mencionadas, busca analizar el efecto que éstas tienen sobre las proyecciones de población, lo que permite apreciar la robustez del modelo.

Los principios en que se basan las hipótesis son:

- En primer lugar se analizan los casos en donde la varianza del error de la ecuación de observación toma diferentes valores; uno por ciento, dos por ciento, tres por ciento, cinco por ciento y diez por ciento de los nacimientos evitados calculados en el momento inicial. Las cifras obtenidas se mantienen constantes a lo largo de los diez años. Al principio del período de análisis se estiman 100 mil nacimientos evitados. La varianza del error en la ecuación de transición se mantiene fija y se supone igual al uno por ciento de la población

total en 1990. Esta combinación de supuestos permite analizar la robustez de la varianza de la ecuación de observación.

- En segundo lugar se analizan los casos en donde la varianza del error de la ecuación de transición toma diferentes valores; uno, dos, tres, cuatro y cinco por ciento de la población total inicial de 1990, que es de 81.7 millones de habitantes. La varianza de la ecuación de observación se mantiene fija en el tiempo. En todos los casos es igual al uno por ciento de los nacimientos evitados de 1991. Las diversas simulaciones muestran la robustez de la varianza del error de la ecuación de transición.

En el modelo se dan como conocidos el valor de ϕ , el de H y el de Z , que representan las tasas de crecimiento demográfico meta, las tasas de nacimientos evitados y los nacimientos evitados, respectivamente. Las cifras presentadas en el modelo si bien son hipótéticas, muestran el esfuerzo adicional que el estado tendría que realizar para llegar a una población cercana a la meta planteada de 100 millones de personas.

A continuación se detallan las hipótesis:

Hipótesis 1-R.- Se supone que la varianza del error de la ecuación de transición es de 1 por ciento respecto a la población de 1990 (81.7 millones de personas) y se supone que la varianza del error de la ecuación de observación es de 1 por ciento sobre los nacimientos evitados de 1991 (100 mil).

Hipótesis 2-R.- Se supone que la varianza del error de la ecuación de transición es de 1 por ciento respecto a la población de 1990 (81.7 millones de personas) y se supone que la varianza del error de la ecuación de observación es de 2 por ciento sobre los nacimientos evitados de 1991 (100 mil).

Hipótesis 3-R.- Se supone que la varianza del error de la ecuación de transición es de 1 por ciento respecto a la población de 1990 (81.7 millones de personas) y se supone que la varianza del error de la ecuación de observación es de 3 por ciento sobre los nacimientos evitados de 1991 (100 mil).

Hipótesis 4-R.- Se supone que la varianza del error de la ecuación de transición es de 1 por ciento respecto a la población de 1990 (81.7 millones de personas) y se supone que la varianza del error de la ecuación de observación es de 5 por ciento sobre los nacimientos evitados de 1991 (100 mil).

Hipótesis 5-R.- Se supone que la varianza del error de la ecuación de transición es de 1 por ciento respecto a la población de 1990 (81.7 millones de

personas) y se supone que la varianza del error de la ecuación de observación es de 10 por ciento sobre los nacimientos evitados de 1991 (100 mil).

Hipótesis 2-Q.- Se supone que la varianza del error de la ecuación de transición es de 2 por ciento respecto a la población de 1990 (81.7 millones de personas) y se supone que la varianza del error de la ecuación de observación es de 1 por ciento sobre los nacimientos evitados de 1991 (100 mil).

Hipótesis 3-Q.- Se supone que la varianza del error de la ecuación de transición es de 3 por ciento respecto a la población de 1990 (81.7 millones de personas) y se supone que la varianza del error de la ecuación de observación es de 1 por ciento sobre los nacimientos evitados de 1991 (100 mil).

Hipótesis 4-Q.- Se supone que la varianza del error de la ecuación de transición es de 5 por ciento respecto a la población de 1990 (81.7 millones de personas) y se supone que la varianza del error de la ecuación de observación es de 1 por ciento sobre los nacimientos evitados de 1991 (100 mil).

También se plantean 2 hipótesis extremas, que son:

Hipótesis E-1.- Se supone que la varianza del error de la ecuación de transición es de 5 por ciento respecto a la población de 1990 (81.7 millones de personas) y se supone que la varianza del error de la ecuación de observación es de 10 por ciento sobre los nacimientos evitados de 1991 (100 mil).

Hipótesis E-2.- Se supone que la varianza del error de la ecuación de transición es de 0 por ciento respecto a la población de 1990 (81.7 millones de personas) y se supone que la varianza del error de la ecuación de observación es de 10 por ciento sobre los nacimientos evitados de 1991 (100 mil).

Las ecuaciones que resumen la metodología del Filtro de Kalman en el caso univariado son las siguientes:

$$P_K^- = \phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2$$

$$K_K = (P_K^- H_K) / (H_K^2 P_K^- + \sigma_R^2)$$

$$\sigma_K^2 = P_K = (1 - K_K H_K) P_K^-$$

$$\rho_K = (1 / (1 + (\sigma_R^2 / H_K^2) (\phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2)))$$

$$\hat{x}_K = (1 - \rho_K) \phi_{K-1} \hat{x}_{K-1} + \rho_K (Z_K / H_K)$$

Los resultados aparecen en los cuadros y en las gráficas.

La población proyectada al año 2000 según las hipótesis de la I-R a la V-R varía entre 95.7 millones para la más baja y 96.7 millones para la más alta, esto es, 1.0 millones de individuos de diferencia. Mientras que la población proyectada al año 2000 según las hipótesis de la I-R o I-Q a la IV-Q varía entre 96.7 millones para la más baja y 97.2 millones para la más alta, esto es, 500 mil personas de diferencia. La diferencia entre la más baja de todas la hipótesis y la más alta de todas ellas, es de 1.5 millones de habitantes.

La mayor varianza de los errores de la ecuación de transición, esto es, la que se refiere a la población, le da una mayor ponderación a la ecuación de observación, mientras que la mayor varianza de los errores de la ecuación de observación, esto es, la que se refiere a los nacimientos evitados, le da una mayor ponderación a la ecuación de transición. Esto se puede ver con los valores de ρ . Los valores de las estimaciones de población de las hipótesis extremas se encuentran cerca del intervalo de variación antes mencionado.

Es importante destacar la notable similitud en los resultados obtenidos.

A fin de comparar estos resultados con otros, se ha proyectado la población de México al año 2000, a través de una función lineal, una geométrica y una exponencial. Los resultados aparecen al final de este capítulo.

Como puede observarse, las diferencias no son tan grandes entre los diversos resultados de la proyecciones demográficas, sin embargo, la ventaja del uso del Filtro de Kalman es que se pueden establecer intervalos de confianza de las poblaciones, ya que se conoce la media y la varianza. Además, una vez que se conoce nueva información, se actualiza el sistema para proyectar la población.

HIPOTESIS I-R

Año	σ_Q^2	σ_R^2	ϕ_K	H_K	Z_K	$\sigma_K^2 = P_K$	P_K^-	K_K	ρ_K	\hat{x}_K
1990	817 000	1000	1.02	0	0	0	0	0	-	81 700 000
1991	817 000	1000	1.019	.001200	100 000	816 040	817 000	0.979248	0.001175	83 334 000
1992	817 000	1000	1.018	.002358	200 000	1 651 533	1 664 344	3.264316	0.007697	84 916 579
1993	817 000	1000	1.017	.003468	300 000	2 453 859	2 528 523	8.510059	0.029513	86 446 852
1994	817 000	1000	1.016	.004540	400 000	3 138 036	3 355 041	14.246759	0.064680	87 928 691
1995	817 000	1000	1.015	.005574	500 000	3 592 658	4 056 256	20.504503	0.114292	89 377 454
1996	817 000	1000	1.014	.006572	600 000	3 780 486	4 518 246	24.845489	0.163285	90 812 542
1997	817 000	1000	1.013	.007534	700 000	3 712 740	4 704 081	27.971935	0.210741	92 258 456
1998	817 000	1000	1.012	.008466	800 000	3 474 615	4 626 899	29.416519	0.249040	93 716 273
1999	817 000	1000	1.011	.009365	900 000	3 162 061	4 375 506	29.613117	0.277327	95 190 755
2000	817 000	1000	1.010	.010235	1000 000	2 843 112	4 049 009	29.098693	0.297825	96 674 496

$$P_K^- = \phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2$$

$$K_K = (P_K^- H_K) / (H_K^2 P_K^- + \sigma_R^2)$$

$$\sigma_K^2 = P_K = (1 - K_K H_K) P_K^-$$

$$\rho_K = (1 / (1 + (\sigma_R^2 / H_K^2) (\phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2)))$$

$$\hat{x}_K = (1 - \rho_K) \phi_{K-1} \hat{x}_{K-1} + \rho_K (Z_K / H_K)$$

HIPOTESIS II-R

Año	σ_Q^2	σ_R^2	ϕ_K	H_K	Z_K	$\sigma_K^2 = P_K$	P_K^-	K_K	ρ_K	\hat{x}_K
1990	817 000	2300	1.02	0	0	0	0	0	-	81 700 000
1991	817 000	2000	1.019	.001200	100 000	816 520	817 000	0.489912	0.000588	83 289 900
1992	817 000	2000	1.018	.002358	200 000	1 657 173	1 664 344	1.953807	0.004607	84 872 156
1993	817 000	2000	1.017	.003468	300 000	2 496 323	2 528 523	4.328608	0.015012	86 401 436
1994	817 000	2000	1.016	.004540	400 000	3 320 606	3 355 041	7.537796	0.034222	87 878 319
1995	817 000	2000	1.015	.005574	500 000	3 979 085	4 056 256	11.226944	0.062579	89 310 519
1996	817 000	2000	1.014	.006572	600 000	4 444 475	4 518 246	14.604587	0.095981	90 712 203
1997	817 000	2000	1.013	.007534	700 000	4 672 462	4 704 081	17.601218	0.132608	92 105 494
1998	817 000	2000	1.012	.008466	800 000	4 672 136	4 626 899	19.7777	0.17435	93 502 576
1999	817 000	2000	1.011	.009365	900 000	4 455 869	4 375 506	21.85555	0.21585	94 926 964
2000	817 000	2000	1.010	.010235	1000 000	4 192 048	4 349 009	21.452533	0.219567	96 351 626

$$P_K^- = \phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2$$

$$K_K = (P_K^- H_K) / (H_K^2 P_K^- + \sigma_R^2)$$

$$\sigma_K^2 = P_K = (1 - K_K H_K) P_K^-$$

$$\rho_K = (1 / (1 + (\sigma_R^2 / H_K^2 (\phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2))))$$

$$\hat{x}_K = (1 - \rho_K) \phi_{K-1} \hat{x}_{K-1} + \rho_K (Z_K / H_K)$$

HIPOTESIS III-R

Año	σ_Q^2	σ_R^2	ϕ_K	H_K	Z_K	$\sigma_K^2 = P_K$		K_K	ρ_K	\hat{x}_K
1990	817 000	3000	1.02	0	0	0	0	0	-	81 700 000
1991	817 000	3000	1.019	.001200	100 000	816 680	817 000	0.326672	0.000392	83 334 000
1992	817 000	3000	1.018	.002358	200 000	1 659 387	1 665 009	1.304671	0.003076	84 917 039
1993	817 000	3000	1.017	.003468	300 000	2 511 634	2 537 181	2.903441	0.010069	86 446 146
1994	817 000	3000	1.016	.004540	400 000	3 336 477	3 414 755	5.049211	0.022923	87 920 086
1995	817 000	3000	1.015	.005574	500 000	4 079 581	4 261 098	7.642381	0.042599	89 342 798
1996	817 000	3000	1.014	.006572	600 000	4 681 543	5 019 886	10.255720	0.067401	90 724 288
1997	817 000	3000	1.013	.007534	700 000	5 088 459	5 630 544	12.778844	0.096276	92 082 781
1998	817 000	3000	1.012	.008466	800 000	5 277 268	6 038 619	14.892545	0.126080	93 433 142
1999	817 000	3000	1.011	.009365	900 000	5 264 187	6 221 682	16.433148	0.153896	94 792 597
2000	817 000	3000	1.010	.010235	1000 000	5 095 025	6 197 636	17.382354	0.177908	96 167 762

$$P_K^- = \phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2$$

$$K_K = (P_K^- H_K) / (H_K^2 P_K^- + \sigma_R^2)$$

$$\sigma_K^2 = P_K = (1 - K_K H_K) P_K^-$$

$$\rho_K = (1 / (1 + (\sigma_R^2 / H_K^2 (\phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2))))$$

$$\hat{x}_K = (1 - \rho_K) \phi_{K-1} \hat{x}_{K-1} + \rho_K (Z_K / H_K)$$

HIPOTESIS IV-R

Año	σ_Q^2	σ_R^2	ϕ_K	H_K	Z_K	$\sigma_K^2 = P_K$	P_K	K_K	ρ_K	\hat{x}_K
1990	817 000	5000	1.02	0	0	0	0	0		81 700 000
1991	817 000	5000	1.019	.001200	100 000	816 808	817 000	0.196034	0.000235	83 334 000
1992	817 000	5000	1.018	.002358	200 000	1 662 064	1 665 142	0.783830	0.001848	84 917 162
1993	817 000	5000	1.017	.003468	300 000	2 524 019	2 539 437	1.750657	0.006071	86 446 032
1994	817 000	5000	1.016	.004540	400 000	3 379 810	3 427 565	3.068871	0.013933	87 918 263
1995	817 000	5000	1.015	.005574	500 000	4 193 080	4 305 829	4.697746	0.026185	89 334 833
1996	817 000	5000	1.014	.006572	600 000	4 918 564	5 136 816	6.464969	0.042488	90 701 264
1997	817 000	5000	1.013	.007534	700 000	5 507 008	5 874 248	8.297971	0.062517	92 029 913
1998	817 000	5000	1.012	.008466	800 000	5 919 293	6 468 121	10.022588	0.084851	93 334 006
1999	817 000	5000	1.011	.009365	900 000	6 138 498	6 879 208	11.497458	0.107674	94 631 514
2000	817 000	5000	1.010	.010235	1000 000	6 174 021	7 091 288	12.638135	0.129351	95 935 237

$$P_K^- = \phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2$$

$$K_K = (P_K^- H_K) / (H_K^2 P_K^- + \sigma_R^2)$$

$$\sigma_K^2 = P_K = (1 - K_K H_K) P_K^-$$

$$\rho_K = (1 / (1 + (\sigma_R^2 / H_K^2) (\phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2)))$$

$$\hat{x}_K = (1 - \rho_K) \phi_{K-1} \hat{x}_{K-1} + \rho_K (Z_K / H_K)$$

HIPOTESIS V-R

Año	σ_Q^2	σ_R^2	ϕ_K	H_K	Z_K	$\sigma_K^2 = P_K$	P_K^-	K_K	ρ_K	\hat{x}_K
1990	817 000	10 000	1.02	0	0	0	0	0	-	81 700 000
1991	817 000	10 000	1.019	.001200	100 000	816 963	817 000	.098028	.000118	83 334 000
1992	817 000	10 000	1.018	.002358	200 000	1 663 700	1 665 240	.392300	.006925	84 917 254
1993	817 000	10 000	1.017	.003468	300 000	2 533 389	2 541 132	.878579	.003047	86 445 945
1994	817 000	10 000	1.016	.004540	400 000	3 413 075	3 437 256	1.549536	.007035	87 916 864
1995	817 000	10 000	1.015	.005574	500 000	4 282 420	4 340 167	2.387020	.013305	89 328 572
1996	817 000	10 000	1.014	.006572	600 000	5 113 375	5 228 856	3.360510	.022085	90 682 368
1997	817 000	10 000	1.013	.007534	700 000	5 872 083	6 074 552	4.424028	.033331	91 983 926
1998	817 000	10 000	1.012	.008466	800 000	6 522 842	6 842 749	5.522238	.046751	93 241 237
1999	817 000	10 000	1.011	.009365	900 000	7 034 765	7 497 329	6.588057	.061697	94 467 631
2000	817 000	10 000	1.010	.010235	1000 000	7 387 690	8 007 361	7.561301	.077390	95 676 814

$$P_K^- = \phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2$$

$$K_K = (P_K^- H_K) / (H_K^2 P_K^- + \sigma_R^2)$$

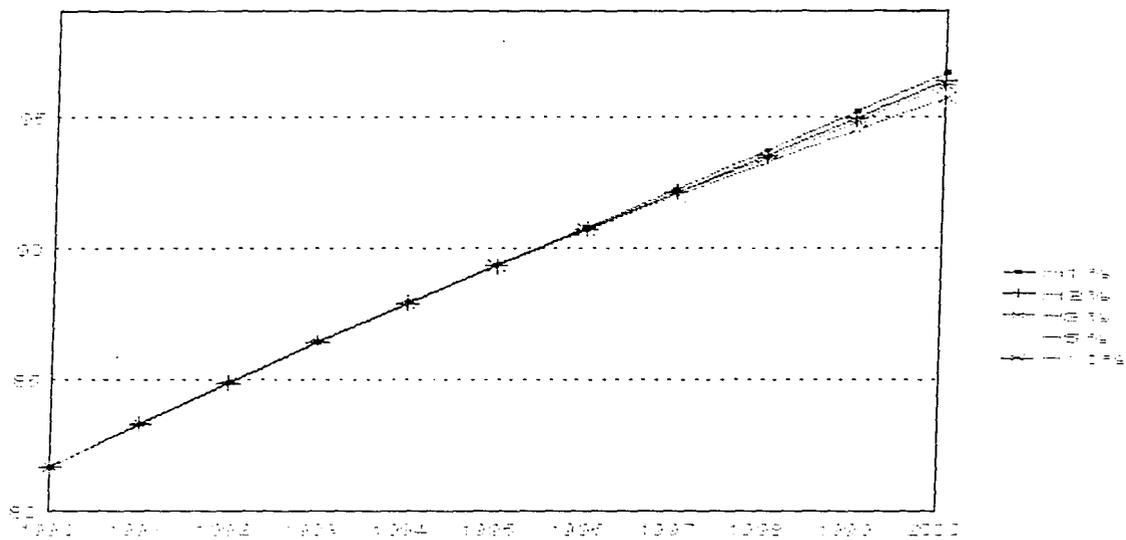
$$\sigma_K^2 = P_K = (1 - K_K H_K) P_K^-$$

$$\rho_K = (1 / (1 + (\sigma_R^2 / H_K^2 (\phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2))))$$

$$\hat{x}_K = (1 - \rho_K) \phi_{K-1} \hat{x}_{K-1} + \rho_K (Z_K / H_K)$$

PROYECCIONES DE POBLACION SEGUN DIVERSAS HIPOTESIS DE LA VARIANZA DE LA ECUACION DE OBSERVACION

Mujeres



HIPOTESIS II-Q

Año	σ_Q^2	σ_R^2	ϕ_K	H_K	Z_K	$\sigma_K^2 = P_K$	P_K^-	K_K	ρ_K	\hat{x}_K
1990	1 634 000	1000	1.02	0	0	0	0	0		000
1991	1 634 000	1000	1.019	.001200	100 000	1 630 164	1 634 000	1.956197	0.002347	83 333 498
1992	1 634 000	1000	1.018	.002358	200 000	3 266 283	3 326 699	7.701898	0.018161	84 915 533
1993	1 634 000	1000	1.017	.003468	300 000	4 733 221	5 018 927	16.414565	0.056926	86 447 495
1994	1 634 000	1000	1.016	.004540	400 000	5 754 982	6 529 518	26.127893	0.118621	87 939 477
1995	1 634 000	1000	1.015	.005574	500 000	6 078 782	7 574 615	35.428737	0.197480	89 416 748
1996	1 634 000	1000	1.014	.006572	600 000	5 888 249	7 896 513	38.697937	0.254323	90 894 929
1997	1 634 000	1000	1.013	.007534	700 000	5 352 463	7 688 274	40.325817	0.303815	92 393 701
1998	1 634 000	1000	1.012	.008466	800 000	4 717 082	7 126 532	39.935713	0.338096	93 899 380
1999	1 634 000	1000	1.011	.009365	900 000	4 125 671	6 464 971	38.637719	0.361842	95 415 636
2000	1 634 000	1000	1.010	.010235	1000 000	3 627 589	5 850 935	37.127341	0.379998	96 935 930

$$P_K^- = \phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2$$

$$K_K = (P_K^- H_K) / (H_K^2 P_K^- + \sigma_R^2)$$

$$\sigma_K^2 = P_K = (1 - K_K H_K) P_K^-$$

$$\rho_K = (1 / (1 + (\sigma_R^2 / H_K^2) (\phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2)))$$

$$\hat{x}_K = (1 - \rho_K) \phi_{K-1} \hat{x}_{K-1} + \rho_K (Z_K / H_K)$$

HIPOTESIS III-Q

Año	σ_Q^2	σ_R^2	ϕ_K	H_K	Z_K	$\sigma_K^2 = P_K$	P_K	K_K	ρ_K	\bar{x}_K
1990	2 451 000	1000	1.02	0	0	0	0	0	0	81 700 000
1991	2 451 000	1000	1.019	.001200	100 000	2 442 380	2 451 000	2.930856	0.003517	83 333 998
1992	2 451 000	1000	1.018	.002358	200 000	4 852 517	4 987 072	11.442244	0.026981	84 914 654
1993	2 451 000	1000	1.017	.003468	300 000	6 862 451	7 479 780	23.798451	0.082533	86 448 241
1994	2 451 000	1000	1.016	.004540	400 000	7 978 453	9 548 758	36.222771	0.164451	87 948 756
1995	2 451 000	1000	1.015	.005574	500 000	7 893 871	10 686 806	46.886335	0.261344	89 446 427
1996	2 451 000	1000	1.014	.006572	600 000	7 263 285	10 583 463	47.734905	0.313714	90 947 580
1997	2 451 000	1000	1.013	.007534	700 000	6 346 980	9 919 081	47.811913	0.360215	92 469 858
1998	2 451 000	1000	1.012	.008466	800 000	5 457 232	8 963 151	46.202230	0.391148	93 994 141
1999	2 451 000	1000	1.011	.009365	900 000	4 715 118	8 039 991	44.158230	0.413542	95 527 523
2000	2 451 000	1000	1.010	.010235	1000 000	4 127 197	7 270 421	42.240393	0.432330	97 064 070

$$P_K^- = \phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2$$

$$K_K = (P_K^- H_K) / (H_K^2 P_K^- + \sigma_R^2)$$

$$\sigma_K^2 = P_K = (1 - K_K H_K) P_K^-$$

$$\rho_K = (1 / (1 + (\sigma_R^2 / H_K^2) (\phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2)))$$

$$\bar{x}_K = (1 - \rho_K) \phi_{K-1} \bar{x}_{K-1} + \rho_K (Z_K / H_K)$$

HIPOTESIS IV-Q

AÑO	σ_Q^2	σ_R^2	ϕ_K	H_K	Z_K	$\sigma_K^2 = P_K$	P_h	K_K	ρ_K	\hat{x}_K
1990	4 085 000	1000	1.02	0	0	0	0	0	0	81 700 000
1991	4 085 000	1000	1.019	.001200	100 000	4 061 111	4 085 000	4.873333	0.005848	83 333 996
1992	4 085 000	1000	1.018	.002358	200 000	7 935 592	8 301 899	18.712152	0.044123	84 912 943
1993	4 085 000	1000	1.017	.003468	300 000	10 721 678	12 308 844	37.181418	0.128945	86 449 604
1994	4 085 000	1000	1.016	.004540	400 000	11 558 965	15 174 314	52.478974	0.238255	87 963 677
1995	4 085 000	1000	1.015	.005574	500 000	10 345 407	16 016 811	63.525421	0.354091	89 488 333
1996	4 085 000	1000	1.014	.006572	600 000	9 007 362	14 743 097	59.197419	0.389045	91 011 856
1997	4 085 000	1000	1.013	.007534	700 000	7 593 650	13 346 334	57.211443	0.431031	92 555 895
1998	4 085 000	1000	1.012	.008466	800 000	6 415 620	11 877 368	54.316675	0.459845	94 097 801
1999	4 085 000	1000	1.011	.009365	900 000	5 507 993	10 655 519	51.584131	0.483085	95 649 952
2000	4 085 000	1060	1.010	.010235	1000 000	4 814 967	9 714 835	49.278906	0.504370	97 207 397

$$P_K^- = \phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2$$

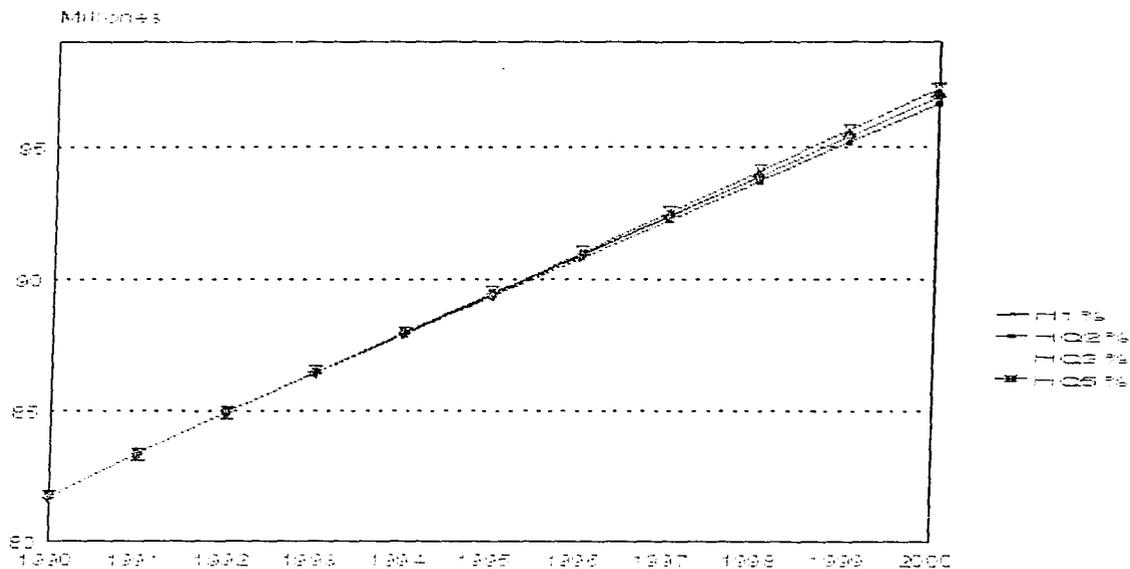
$$K_K = (P_K^- H_K) / (H_K^2 P_K^- + \sigma_R^2)$$

$$\sigma_K^2 = P_K = (1 - K_K H_K) P_K^-$$

$$\rho_K = (1 / (1 + (\sigma_R^2 / H_K^2) (\phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2)))$$

$$\hat{x}_K = (1 - \rho_K) \phi_{K-1} \hat{x}_{K-1} + \rho_K (Z_K / H_K)$$

PROYECCIONES DE POBLACION SEGUN DIVERSAS HIPOTESIS DE LA VARIANZA DE LA ECUACION DE TRANSICION



HIPOTESIS E-1

Año	σ_Q^2	σ_R^2	ϕ_K	H_K	Z_K	$\sigma_K^2 = P_K$	P_K^-	K_K	ρ_K	\hat{x}_K
1990	4 085 000	10 000	1.02	0	0	0	0	0		81 700 000
1991	4 085 000	10 000	1.019	.001200	100 000	4 082 598	4 085 000	0.489912	0.000598	83 334 000
1992	4 085 000	10 000	1.018	.002358	200 000	8 285 861	8 324 211	1.953806	0.004607	84 916 887
1993	4 085 000	10 000	1.017	.003468	300 000	12 481 612	12 671 837	4.328607	0.015012	86 446 289
1994	4 085 000	10 000	1.016	.004540	400 000	16 419 443	16 994 594	7.454447	0.033843	87 922 301
1995	4 085 000	10 000	1.015	.005574	500 000	19 728 937	21 034 069	11.131766	0.062048	89 352 210
1996	4 085 000	10 000	1.014	.006572	600 000	22 082 105	24 410 244	14.512401	0.095376	90 750 092
1997	4 085 000	10 000	1.013	.007534	700 000	23 253 722	26 789 732	17.519409	0.131991	92 138 268
1998	4 085 000	10 000	1.012	.008466	800 000	23 283 388	27 947 249	19.711893	0.166881	93 529 575
1999	4 085 000	10 000	1.011	.009365	900 000	22 434 854	27 930 542	21.010430	0.196763	94 937 350
2000	4 085 000	10 000	1.010	.010235	1000 000	21 056 930	27 016 135	21.551489	0.220579	96 361 563

$$P_K^- = \phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2$$

$$K_K = (P_K^- H_K) / (H_K^2 P_K^- + \sigma_R^2)$$

$$\sigma_K^2 = P_K = (1 - K_K H_K) P_K^-$$

$$\rho_K = (1 / (1 + (\sigma_R^2 / H_K^2 (\phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2))))$$

$$\hat{x}_K = (1 - \rho_K) \phi_{K-1} \hat{x}_{K-1} + \rho_K (Z_K / H_K)$$

HIPOTESIS E-2

Año	σ_Q^2	σ_R^2	ϕ_K	H_K	Z_K	$\sigma_K^2 = P_K$	P_K^-	K_K	ρ_K	\hat{x}_K
1990	0	10 000	1.02	0	0	0	0	0	0	81 700 000
1991	0	10 000	1.019	.001200	100 000	0	0	0	0	83 334 000
1992	0	10 000	1.018	.002358	200 000	0	0	0	0	84 917 346
1993	0	10 000	1.017	.003468	300 000	0	0	0	0	86 445 858
1994	0	10 000	1.016	.004540	400 000	0	0	0	0	87 915 437
1995	0	10 000	1.015	.005574	500 000	0	0	0	0	89 322 085
1996	0	10 000	1.014	.006572	600 000	0	0	0	0	90 661 916
1997	0	10 000	1.013	.007534	700 000	0	0	0	0	91 931 182
1998	0	10 000	1.012	.008466	800 000	0	0	0	0	93 126 288
1999	0	10 000	1.011	.009365	900 000	0	0	0	0	94 243 804
2000	0	10 000	1.010	.010235	1000 000	0	0	0	0	95 280 486

$$P_K^- = \phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2$$

$$K_K = (P_K^- H_K) / (H_K^2 P_K^- + \sigma_R^2)$$

$$\sigma_K^2 = P_K = (1 - K_K H_K) P_K^-$$

$$\rho_K = (1 / (1 + (\sigma_R^2 / H_K^2) (\phi_{K-1}^2 \sigma_{K-1}^2 + \sigma_Q^2)))$$

$$\hat{x}_K = (1 - \rho_K) \phi_{K-1} \hat{x}_{K-1} + \rho_K (Z_K / H_K)$$

PROYECCIONES DE LA POBLACION NACIONAL Y POR ENTIDAD FEDERATIVA 1990-2000

AÑO	1990 ¹	1992 ²	1995 ³	1995 ⁴	1995 ⁵	2000 ⁶	2000 ⁷	2000 ⁸
TOTAL	812 496 645	85 627 971	90 302 867	90 648 739	90 649 535	98 103 806	99 911 897	99 910 442
Aguascalientes	719 659	770 972	825 809	829 695	829 695	917 349	937 932	937 932
Baja California Norte	1 660 855	1 908 434	2 173 063	2 213 588	2 214 086	2 614 518	2 835 246	2 836 574
Baja California Sur	317 764	351 690	387 895	391 994	391 898	448 333	469 719	469 496
Campeche	535 185	569 417	606 043	608 345	608 505	667 162	679 310	679 685
Coahuila	1 972 340	2 040 046	2 112 179	2 115 137	2 114 743	2 232 492	2 246 298	2 245 706
Colima	428 510	458 607	490 730	493 172	493 172	544 336	556 720	556 677
Chiapas	3 210 496	3 274 574	3 678 586	3 694 960	3 694 960	4 081 825	4 171 076	4 170 755
Chihuahua	2 441 811	2 515 15	2 569 583	2 570 804	2 571 536	2 679 467	2 687 525	2 688 502
Distrito Federal	8 235 744	8 276 345	8 317 278	8 317 278	8 317 278	8 387 282	8 388 105	8 388 929
Durango	1 349 378	1 394 571	1 442 620	1 444 509	1 444 239	1 523 043	1 531 949	1 531 409
Guanajuato	3 982 593	4 170 885	4 371 294	4 382 047	4 382 047	4 706 230	4 758 800	4 760 393
Guanajuato	2 620 637	2 732 699	2 852 039	2 857 281	2 857 543	3 051 470	3 078 462	3 078 724
Hidalgo	1 888 366	1 945 514	2 006 767	2 008 844	2 008 277	2 108 927	2 118 747	2 117 802
Jalisco	5 302 689	5 693 177	6 109 758	6 141 044	6 142 105	6 804 941	6 969 854	6 971 445
México	9 815 795	10 705 862	1 165 238	11 746 562	11 745 580	13 244 452	13 712 666	13 709 721
Michoacán	3 548 199	3 723 543	3 910 825	3 920 050	3 922 115	4 223 421	4 271 322	4 273 451
Morelos	1 195 059	1 259 170	1 327 591	1 331 535	1 331 415	1 441 839	1 461 677	1 461 318
Nayarit	824 643	871 710	921 868	925 167	925 167	1 005 652	1 021 568	1 021 568
Nuevo León	3 098 736	3 336 044	3 590 196	3 609 098	3 610 027	4 013 483	4 116 051	4 118 220
Oaxaca	3 019 560	3 207 147	3 407 272	3 420 558	3 421 161	3 741 235	3 808 571	3 809 779
Puebla	4 126 101	4 406 652	4 706 231	4 727 274	4 726 861	5 205 902	5 315 656	5 314 005
Querétaro	1 051 235	1 126 143	1 206 292	1 211 969	1 211 969	1 340 009	1 370 075	1 370 075
Quintana Roo	493 277	577 419	667 453	683 287	683 287	817 607	904 769	904 818
San Luis Potosí	2 003 187	2 088 544	2 180 068	2 184 075	2 184 275	2 332 511	2 353 144	2 353 344
Sinaloa	2 204 054	2 341 346	2 488 157	2 498 075	2 497 193	2 733 247	2 782 839	2 780 855
Sonora	1 823 606	1 866 757	1 912 598	1 914 422	1 913 692	1 989 189	1 996 301	1 994 843
Tabasco	1 501 744	1 595 487	1 695 319	1 702 077	1 702 527	1 862 313	1 896 102	1 896 853
Tamaulipas	2 249 581	2 351 663	2 461 267	2 465 541	2 465 766	2 643 708	2 668 003	2 668 678
Tlaxcala	761 277	812 749	867 856	871 738	871 586	959 666	979 687	979 383
Veracruz	6 228 239	6 405 478	6 593 837	6 602 556	6 600 688	6 908 363	6 943 241	6 939 504
Yucatán	1 362 940	1 390 318	1 419 911	1 420 047	1 420 320	1 469 113	1 471 294	1 471 703
Zacatecas	1 276 323	1 309 493	1 345 244	1 346 010	1 345 627	1 404 721	1 409 188	1 408 295

NOTAS: ¹ Población censal al 12 de marzo de 1990.

² Población al 31 de diciembre de 1992 según la Encuesta Nacional de la Dinámica Demográfica.

³ Población proyectada a 1995 a partir de una función lineal.

⁴ Población proyectada a 1995 a partir de una función geométrica.

⁵ Población proyectada a 1995 a partir de una función exponencial.

⁶ Población proyectada al año 2000 a partir de una función lineal.

⁷ Población proyectada al año 2000 a partir de una función geométrica.

⁸ Población proyectada al año 2000 a partir de una función exponencial.

Bibliografía

LIBROS

Anderson, B.D.O. and Moore, J.B. (1979), *Optimal Filtering*, Prentice Hall, Englewood Cliffs, N.J.

Balakrishnan, A.V. (1987), *Kalman Filtering Theory*, Optimization Software, New York.

Benítez, R. y Cabrera, G. (1966), *Proyección de la Población de México, 1960-1980*, Banco de México, S.A., México.

Bertsekas, D.P. (1987), *Dynamic Programming, Deterministic and Stochastic Models*, Prentice-Hall., Englewood Cliffs, N.Y.

Bozic, S.M. (1979), *Digital and Kalman Filtering*, Department of electronic and electrical Engineering, University of Birmingham, Edward Arnold.

Brammer, K. and Siffing, G. (1989), *Kalman-Bucy Filters*, Artech House, Norwood, Mass.

Brown, R.G. (1983), *Introduction to Random Signal Analysis and Kalman Filtering*, Wiley, New York.

Bucy, R.S. and Joseph, P.D. (1968), *Filtering for Stochastic Processes, with Applications to Guidance*, Wiley, New York.

Cabrera, G. (1960), *La futura población de México, 1960-1980*, reproducción fotográfica.

Camposortega, S. (1980), *Proyecciones de la Población Mexicana, 1970-2040*, El Colegio de México, tesis de maestría, México.

Catlin, D.E. (1989), *Estimation, Control, and the Discrete Kalman Filter*, Springer-Verlag, New York.

Coale, A.J. and Hoover, E.M. (1958), *Population Growth and Economic Development in Low-income Countries*, Princeton University Press.

Conapo (1993), *Proyecciones de la población de México*, México, inédito.

- Chavira, M. (1966), *Proyecciones demográficas de la República Mexicana*, Dirección General de Estadística, Secretaría de Industria y Comercio, México.
- Durán, J. (1962), "La explosión demográfica en México". Cincuenta años de revolución. II: La vida social. Fondo de Cultura Económica, México.
- INEGI y Conapo (1985), *Proyecciones de la población de México y de las entidades federativas: 1980-2010*, S.P.P., México.
- Jazwinski, A.H. (1970), *Stochastic Processes and Filtering Theory*, Academic Press, San Diego.
- Kailath, T. (1981), *Lectures on Wiener and Kalman Filtering*, Springer-Verlag, New York.
- Kailath, T. (1977), *Linear Least Squares Estimation*, Dowden, Hutchinson and Ross, Inc., Stroudsburg, Pennsylvania.
- Lewis, F.H. (1986), *Optimal Estimation with and Introduction to Stochastic Control Theory*, Wiley, New York.
- Malthus, T. (1951), *Ensayos sobre el principio de la población*, Fondo de Cultura Económica, México.
- Mendel, J.M. (1987), *Lessons in digital Estimation Techniques*, Prentice Hall, Englewood Cliffs, N. J.
- Naciones Unidas (1978), *Factores determinantes y consecuencias de las tendencias demográficas*. vol. 1, Estudios sobre población, núm. 50, Nueva York.
- Naciones Unidas (1960), *Los recursos humanos de Centroamérica, Panamá México en algunos aspectos del desarrollo económico*, estudio preparado por Louis Ducoff, CEPAL, 1960.
- Naciones Unidas (1954), *La población de América Central y México en el período 1950 a 1980*, ONU, Nueva York.
- Núñez, L. y Moreno, L. (1986), *México: proyecciones de población urbana y rural, 1980-2010*, Academia Mexicana de Investigación en Demografía Médica, México.
- ONU (1992), *World Population Prospects: The 1992 Revision*, Nueva York.
- Papoulis, A. (1988), *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York.
- Parzen, E. (1962), *Stochastic Processes*, Holden-Day, San Francisco.
- Recchini, Z. (1963), *Proyección de la población de México por sexo y grupos de edades, 1960-1980*, mimeo., Celade, Santiago de Chile.

Recchini, Z. y Chavira, M. (1964), *Proyección de la población de México por sexo y grupos de edades, 1960-1980*, Celade, Santiago de Chile.

Secretaría de Programación y Presupuesto (1978), *Proyecciones de la población, México, 1970-2000*, México.

Secretaría de Programación y Presupuesto, Consejo Nacional de Población y Centro Latinoamericano de Demografía (1983), *México: estimaciones y proyecciones de población. 1950-2000*, S.P.P., México.

Smith, D. and Keyfitz, N. (1977), *Mathematical Demography*, Berlin: Springer-Verlag.

Sorenson, H.W. (1985), *Kalman Filtering: Theory and Application*, IEEE Press, Piscataway, N. J.

Stengel, R.F. (1986), *Stochastic Optimal Control: Theory and Application*, Wiley, New York.

Young, P.C. (1984), *Recursive Estimation and Time Series*, Springer-Verlag, New York.

FUNDAMENTACION MATEMATICA DEL FILTRO DE KALMAN

Arnold, L. (1974), *Stochastic Differential Equations: Theory and Applications*, Wiley, New York.

Coddington, E.A. and Levinson, N. (1955), *Theory of Ordinary Differential Equations*, McGraw-Hill, New York.

Davenport, W.D. and Root, W.L. (1958), *Random Signals and Noise*, McGraw-Hill, New York.

Harris, B. (1966), *Theory of Probability*, Addison-Wesley, Reading, Mass.

Kailath, T. (1980), *Linear Systems*, Prentice Hall, Englewood Cliffs, N.J.

Loève, M. (1963), *Probability Theory*, 3rd., Van Nostrand Reinhold, New York.

Searle, S.R. (1971), *Linear Models*, Wiley, New York.

Gauss, C.F. (1963), *Theory of Motion of the Heavenly Bodies* (English Transl.), Dover, Mineola, N. Y.

TRABAJOS y ARTICULOS

Bryson, A.E. Jr., and Johansen, D.E. (1965), "Linear filtering for time-varying systems using measurements containing colored noise," *IEEE Transactions on Automatic Control*, Vol. AC-10, pp. 4-10.

Bucy, R.S. (1967), "Optimal filtering for correlated noise," *Journal of Mathematical the Analysis Applications*, Vol. 20, pp. 1-8.

- Heffes, H. (1966), "The effects of erroneous models on the Kalman filter response," IEEE Transactions on Automatic Control, Vol. AC-11, pp. 541-543.
- Kailath, T. (1974), "A view of three decades of linear filtering theory," IEEE Transaction on Information Theory, Vol. IT-20, No. 2 pp. 146-181.
- Kailath, T. (1968), "An innovations approach to least squares estimation, Part I: Linear filtering in additive white noise," IEEE Transactions on Automatic Control, Vol. AC-13, pp. 646-655.
- Kalman, R.E. (1960), "A new approach to linear filtering and prediction problems," ASME Journal of Basic Engineering, series D Vol. 82, pp. 34-45.
- Kalman, R.E. (1961), "New methods and results in linear prediction and filtering theory," Proc. Symposium on Engineering Applications of Random Function Theory and Probability, Wiley, New York.
- Kalman, R.E. and Bucy, R.S. (1961), "New results in linear filtering and prediction theory," ASME Journal of Basic Engineering, Series D. Vol. 83, pp. 95-108.
- Kalman, R.E. (1963), "New methods in Wiener filtering," chapter 9, Proc. First Symposium on Engineering Applications of Random Function Theory and Probability, Wiley, New York.
- Kao, M.H. and Eller, K.H. (1983), "Multiconfiguration Kalman filter design for high-performance GPS navigation," IEEE Transactions on Automatic Control, Vol. AC-28.
- Meditch, J.S. (1973), "A survey of data smoothing for linear and nonlinear systems," Automatica, vol. 9, pp. 151-162.
- Riccati, J.F. (1724), Animadversiones in aequationes differentiales secundi gradus," Acta Eruditorum Quae Lipsiae Publicantur Supplementa, Vol. VIII, pp. 66-73.
- Schlee, F.H., Standish, C.J. and Toda, N.F. (1967), "Divergence in the Kalman filter", AIAA Journal, Vol. 5, pp. 1114-1120.
- Sorenson, H.W. (1967), "On the error behavior in linear minimum variance estimation problems," IEEE Transactions on Automatic Control, Vol. Ac-12, pp. 557-562.
- Sorenson, H.W. (1970), "Least-squares estimation: from Gauss to Kalman," IEEE Spectrum, pp. 63-68.
- Sorenson, H.W. (1974), "On the development of practical nonlinear filters," Information Sciences, vol. 7, pp. 253-270.

Swerling, P. (1959), "First order error propagation in a stagewise differential smoothing procedure for satellite observations," *Journal of Astronautical Sciences*, vcl. 6, pp. 46-52.

Apéndice I. Glosario de términos demográficos

AUMENTO DE LA POBLACION. El aumento total de la población resultante de la interacción de los nacimientos, las defunciones y la migración en una población, en un determinado periodo de tiempo.

CENSO. Estudio de una zona determinada que da como resultado la enumeración de toda la población y la recopilación de la información demográfica, social y económica concerniente a dicha población en un momento dado.

CRECIMIENTO NATURAL. El excedente (o déficit) de nacimientos sobre las defunciones en una población, durante un determinado periodo.

CRECIMIENTO NEGATIVO DE LA POBLACION. Una disminución neta en la dimensión de la población.

DEMOGRAFIA. Del griego demos [pueblo] + graphie [estudio]. El estudio científico de las poblaciones humanas y de su dimensión, composición, distribución, densidad, crecimiento y otras características demográficas y socioeconómicas y de las causas y consecuencias de los cambios experimentados por esos factores.

DISTRIBUCION DE LA POBLACION. El régimen de asentamiento y dispersión de una población.

ECUACION COMPENSADORA. Una fórmula demográfica básica utilizada para estimar el cambio total de población entre dos fechas dadas, o para estimar cualquier componente desconocido del movimiento de población a partir de otros componentes conocidos. La ecuación compensadora abarca todos los componentes del movimiento de población: nacimientos, defunciones, inmigración y emigración.

EMIGRACION. El proceso de dejar un país para adoptar residencia en otro.

ESPERANZA DE VIDA. El número de años de vida que le restan en promedio a una persona de no variar la tendencia de la mortalidad. Se cita comúnmente

como esperanza de vida al nacer.

ESTADISTICAS VITALES. Datos demográficos sobre nacimientos, defunciones, muertes fatales, casamientos y divorcios.

ESTRUCTURA DE LA POBLACION POR EDAD Y SEXO. La composición de una población de acuerdo con el número o proporción de varones y mujeres en cada categoría de edades. La estructura de la población por edad y sexo de una población es el resultado acumulativo de las tendencias retrospectivas de la fecundidad, mortalidad y migración. Para describir y analizar muchas de las otras clases de datos demográficos es esencial disponer antes de información sobre la composición de la población por edad y sexo.

FECUNDIDAD. Procreación real de un individuo, pareja, grupo o población.

INMIGRACION. El proceso de entrar a un país para adoptar residencia permanente los que residían en otro.

MIGRACION. El movimiento de personas a través de una división política para establecer una nueva residencia permanente. Se divide en migración internacional (migración entre países) y migración interna (migración dentro de un país).

MIGRACION INTERNA. El proceso de pasar de una a otra subdivisión administrativa de un país (por ejemplo, municipio o estado) para adoptar residencia en ella.

MIGRACION NETA. El efecto neto de la inmigración y la emigración sobre la población de una zona en un determinado período de tiempo, expresado como aumento o disminución.

MORTALIDAD. Defunciones como componentes del cambio de población.

NATALIDAD. Nacimientos como componentes del cambio de población.

PERIODO DE REPRODUCCION. En la mujer, la edad en que es capaz de procrear y que, según se ha supuesto arbitrariamente para fines estadísticos, está comprendida entre los 15 y los 49 años de edad en la mayoría de los países.

PIRAMIDE DE POBLACION. Una clase especial de gráfico de barras que presenta la distribución de una población por edad y sexo. La mayoría de los países caen dentro de una de las tres categorías generales de pirámides siguientes: 1) expansiva -con una base amplia, lo cual indica una proporción elevada de niños y una rápida tasa de crecimiento de la población; 2) constrictiva -con una base más estrecha que el centro de la pirámide, lo que generalmente ilustra un rápido descenso de la fecundidad; 3) estacionaria -con una base estrecha y un número de

personas aproximadamente igual en cada grupo de edades, que se estrecha con los grupos de más edad, lo que indica una proporción moderada de niños y una tasa de crecimiento lenta o nula.

POLITICA DE POBLACION. Medidas explícitas o implícitas instituidas por un gobierno para influir en la dimensión, crecimiento, distribución o composición de la población.

PROBABILIDAD DE SUPERVIVENCIA. La proporción de personas en un grupo específico (por edad, sexo, estado de salud, etc.) que vivían al principio de un intervalo (por ejemplo, un período de 5 años) y que sobreviven al final de dicho intervalo.

PROYECCIONES DE LA POBLACION. Cálculo de los cambios futuros en el número de personas, con base a ciertas hipótesis acerca de las tendencias futuras en las tasas de fecundidad, mortalidad y migración. Se presentan frecuentemente proyecciones bajas, medias y altas de la misma población, basándose en diferentes hipótesis sobre cómo cambiarán estas tasas en el futuro.

RAZON DE DEPENDENCIA (POR EDAD). La razón entre las personas que por su edad se definen como dependientes (menores de 15 años y mayores de 64) y las que se definen como económicamente productivas (15-64 años) dentro de una población.

TASA DE AUMENTO NATURAL. La tasa de aumento (o disminución) de una población en un determinado año debido a un excedente (o déficit) de nacimientos frente a las defunciones, expresada como porcentaje de la población base.

TASA BRUTA. La tasa de cualquier evento demográfico calculada para toda una población.

TASA BRUTA DE REPRODUCCION (TBR). El número medio de hijas que nacerían vivas durante la vida de una mujer (o grupo de mujeres), si sus años reproductivos transcurrieran conforme a las tasas de fecundidad por edad de un determinado año.

TASA DE CRECIMIENTO. La tasa a la que una población aumenta (o disminuye) en un determinado año debido al incremento natural o a la migración, expresada como porcentaje de la población.

TASA POR EDAD. La tasa obtenida para grupos de edades específicos (por ejemplo la tasa de fecundidad por edad, la tasa de mortalidad por edad, la tasa de nupcialidad por edad, la tasa de analfabetismo por edad, la tasa de inscripción escolar por edad, etc.).

TASA DE EMIGRACION. El número de emigrantes que salen de una zona de origen por 1,000 habitantes de dicha zona.

TASA GLOBAL DE FECUNDIDAD (TGF). El número promedio de niños que habría tenido una mujer (o grupo de mujeres) durante su vida, si sus años de reproducción transcurrieran conforme a las tasas de fecundidad por edad de un determinado año.

TASA DE MORTALIDAD. (o tasa bruta de mortalidad). El número de defunciones por 1,000 habitantes en un determinado año.

TASA DE MORTALIDAD INFANTIL. El número de defunciones de menores de un año en un determinado año por 1.000 nacidos vivos en dicho año.

TASA DE NATALIDAD. (tasa bruta de natalidad). El número de nacimientos por 1.000 habitantes en un determinado año. No confundir con tasa de crecimiento.

TASA NETA DE REPRODUCCION (TNR). El número medio de hijas que tendrá una mujer (o grupo de mujeres) si, desde que nace, su vida se comporta a las tasas de fecundidad y mortalidad por edad de un determinado año. Esta tasa es análoga a la tasa bruta de reproducción, pero toma en cuenta el hecho de que algunas mujeres morirán antes de transcurridos sus años de reproducción. Una TNR de 1.00 significa que cada generación de madres tiene exactamente la cantidad suficiente de hijas para reemplazarlas en la población.

Apéndice II. Recuento histórico

¿Qué es el Filtro de Kalman¹?

El Filtro de Kalman es un estimador inesgado, recursivo y de mínimos cuadrados con una señal aleatoria Gaussiana. Es posible señalar que esta técnica es uno de los mayores descubrimientos en la historia de la Teoría de la Estimación. Tiene importantes aplicaciones en el control de sistemas dinámicos tal es el caso de procesos de producción, aviación, navegación marítima, aviación espacial, entre otras. El Filtro de Kalman se usa en la predicción del curso futuro probable de sistemas dinámicos que las personas no están en posibilidad de controlar, como por ejemplo, el flujo de agua en los ríos, las trayectorias de cuerpos celestes, el precio de comercialización de ciertos bienes o la población que puede tener una región o un país.

El Filtro de Kalman es una forma moderna de discutir la Teoría de estimación de mínimos cuadrados. La diferencia esencial entre el Filtro de Kalman y el modelo lineal convencional es que el parámetro de estado en el Filtro de Kalman no es constante, sino que puede cambiar con el tiempo.

¿Por qué se llama Filtro de Kalman?

Comúnmente se dice que un filtro es un artefacto físico para remover fracciones no deseadas de ciertas mezclas. La palabra filtro tiene sus raíces latinas en la

¹Rudolf Emil Kalman nació el 19 de mayo de 1930 en Budapest. Hijo de Otto y Ursula Kalman. La familia emigró de Hungría a los Estados Unidos durante la Segunda Guerra Mundial. En 1944 residieron en Youngstown, Ohio. Terminó sus estudios de licenciatura y maestría en Ingeniería Eléctrica en el MIT, en 1953 y 1954, respectivamente. En 1956, Kalman trabajó en un laboratorio de investigación de la International Business Machines Corporation en Poughkeepsy y posteriormente en el Research Institute of Advance Studies. En 1985 recibió el premio Kyoto, considerado para algunos como equivalente al premio Nobel.

Edad Media, y tenía que ver con el material que era usado para filtrar líquidos. En sus orígenes, un filtro solucionaba el problema de separación de componentes no deseados de mezclas de -gas, líquidos y sólidos.

Posteriormente el término fue aplicado en el filtrado de señales electrónicas. Estas señales son mezclas de diferentes componentes de frecuencia. El concepto se extendió en el decenio de los 30's y de los 40's al analizar la separación de "señales" de las de "ruido". Kolmogorov y Wiener usaron la caracterización estadística de su distribución de probabilidad para formar estimaciones óptimas de la "señal", dada una suma de señales y de ruido. Con el Filtro de Kalman el término fue más allá de la idea original de separación de componentes de una mezcla. También incluye la solución al problema inverso, en el cual uno conoce como representar la medición de las variables como función de las variables de interés principal, es decir, se invierte esta relación funcional y se estima a las variables independientes como una función inversa de las variables dependientes. Estas variables de interés pueden ser dinámicas.

Uno se enfrenta a problemas de filtrado en la vida cotidiana. Cuando uno prende un radio en AM y le sube el volumen, ocurren dos cosas; o hay señal buena o hay ruido. El filtro apropiado para escuchar nuestra melodía favorita es aquel que la frecuencia pase por un pequeño rango de kilohertz.

El Filtro de Kalman es la generalización de otras disciplinas de la teoría estadística, sobre todo del campo de la Teoría de la Estimación. Se basa en los conceptos del análisis de los mínimos cuadrados, en los de procesos estocásticos, en los de sistemas dinámicos y en la Teoría de probabilidad.

El desarrollo de la Teoría de la Estimación cubre un período que abarca cinco siglos, desde Cardano hasta Bierman.²

²Gerónimo Cardano(1501-1576). Matemático, astrónomo y médico italiano. En el libro "Book on Games of Chance" presenta de manera sistemática el cálculo de probabilidades, un siglo antes de Pascal y Fermat.

Galileo Galilei(1564-1642). Matemático, astrónomo y médico, considerado el fundador del método experimental.

Fermat, Pierre(1601-1665). Matemático francés. Por su correspondencia con Pascal es cofundador de la Teoría de la probabilidad.

Pascal, Blaise(1623-1662). Matemático francés, fundador de la moderna Teoría de probabilidad.

Huygens, Christian(1629-1695). Matemático y astrónomo alemán. Hizo importantes aportaciones a la ciencia de la dinámica.

Newton, Isaac(1643-1727). Matemático inglés, quien presentó los fundamentos del cálculo.

Un misterio de la historia de la ciencia es por qué tomó tanto tiempo a los matemáticos la formalización de la incertidumbre.

Los romanos vendían seguros antes que el riesgo fuera un concepto tratado seriamente por matemáticos. Al principio del Renacimiento los italianos tenían políticas de seguros. El italiano Girolamo Cardano (1501-1576) presenta de manera sistemática el cálculo de probabilidades en su libro "Book on Games of Chance". Supuso que el lanzamiento de dados son eventos estadísticamente independientes.

Es importante señalar que Galileo Galilei (1564-1642) ya reconoce los errores de medición.

Tratamientos en el campo de la probabilidad realizados con mayor detalle fueron desarrollados por Blaise, Pascal (1623-1662), Pierre de Fermat (1601-1655) y Christian Huygens (1629-1695). El trabajo de Fermat fue retomado por Bernoulli (1654-1705), quien es considerado por los historiadores como el padre de la Teoría de Probabilidad. El realizó una prueba rigurosa de la Ley de los Grandes Números para ensayos repetidos e independientes. Thomas Bayes (1702-1761) fue el primero en usar la probabilidad en forma inductiva y estableció las bases de la inferencia estadística. El Teorema de Bayes usa la probabilidad "a posteriori", el cual presenta en su libro "Essay Towards Solving a Problem in the doctrine of Chances",

Bernoulli, Daniel (1700-1782). Matemático suizo. Hizo importantes aportaciones a la Teoría de la probabilidad. Además desarrolló la famosa distribución de Bernoulli.

Bayes (1702-1761). Teólogo y matemático inglés. Fue el primero en usar la probabilidad en forma inductiva y estableció la base de la inferencia estadística. Se tiene el teorema de Bayes usando la probabilidad a posteriori, el cual presenta en su libro: "Essay Towards Solving a Problem in the Doctrine of Chances".

Laplace (1749-1827). Matemático y astrónomo francés. Estudió la estabilidad del Sistema Solar. En 1814 se publicó "A philosophical essay on probability" que fue una introducción a la "Analytic Theory of Probability".

Legendre, Adrien (1752-1833). Matemático francés. Determinó nuevas órbitas de los cometas.

Gauss, Friedrich (1777-1855). Matemático alemán. Determinó las órbitas de asteroides a través del método de los mínimos cuadrados.

Markov, Andrei (1856-1922). Matemático ruso quien ayudó a desarrollar la teoría de procesos estocásticos, en el caso especial de las cadenas de Markov.

Maxwell, James (1831-1879). Físico inglés. Realizó un enfoque estocástico de la mecánica estadística.

Wiener, Norbert (1894-1964). Matemático de Estados Unidos. Trabajó sobre problemas no lineales en Teoría estocástica, interpretación y suavizamiento de series de tiempo.

Kolmogorov, Andrei (1903-1987). Desarrolló los aspectos teóricos de la Teoría de la Probabilidad desde un punto de vista axiomático. Trabajó en la Teoría de la medida, y en los modelos dinámicos.

Abraham de Moivre (1667-1754), Pierre Simon Marquis de Laplace (1749-1827), Adrien Marie Legendre (1752-1833) y Carl Friederich Gauss (1777-1855) continuaron estos desarrollos en la Teoría de Estimación en el siglo XIX. Laplace escribe en 1814 su popular trabajo: "A philosophical essay on probability". Legendre, por su parte, determinó nuevas órbitas de los cometas.

Entre el principio del siglo XIX y mediados del siglo XX, las probabilidades tomaron gran importancia. La idea de que las leyes de la naturaleza son de carácter aleatorio y que podrían tratarse con modelos probabilísticos provocó que rápidamente se expandiera el desarrollo y aplicación de estos modelos. El trabajo de James Clerk Maxwell (1831-1879) realizó un enfoque probabilístico de la mecánica estadística.

Una figura importante en la Teoría de probabilidad y en procesos estocásticos fue el ruso Andrei Nikolaevich Kolmogorov (1903-1987). Redefinió los fundamentos de la Teoría de probabilidad. Con Norbert Wiener (1894-1964) trabajó en predicción, suavizamiento y filtrado de procesos Markovianos, y en la teoría general de procesos ergódicos.

¿Cuándo empezó la Teoría de la Estimación?

El importante desarrollo de la Teoría de la Estimación se debió en gran medida a los estudios de astronomía, principalmente aquéllos que se dirigieron a la determinación de las órbitas de los planetas y de los cometas, mediante el uso de información obtenida a partir de telescopios.

En 1795, Friederick Gauss desarrolló el método de los mínimos cuadrados, aplicándolo al movimiento de los planetas y los cometas. El movimiento de los cuerpos se podía describir completamente, utilizando 6 parámetros. El problema consistía en estimar dichos parámetros a partir de los datos observados. Fue hasta 1809 cuando Gauss publicó su libro: "Theoria motus corporum coelestium" en el que describe el método de los mínimos cuadrados. En el año de 1806, Legendre publica, en Paris, el trabajo "Nouvelles méthodes pour la détermination des orbites des comètes", en donde también presenta la técnica de los mínimos cuadrados.

En el siglo XIX se presentó una controversia sobre quién era el inventor de los mínimos cuadrados. Los historiadores tienen indicios de que este método es obra de Gauss.

Gauss explicaba la técnica de la siguiente manera: si las observaciones astronómicas sobre las que se basan los cálculos de las órbitas fueran absolutamente correctas, entonces los elementos que se deducirían a partir de 3 ó 4 observaciones serían exactos. Sin embargo, la realidad no opera de esa manera, ya que las observaciones y medidas que se realizan no son más que aproximaciones a la verdad,

lo mismo puede decirse acerca de los cálculos que se derivan de ellas, y la idea es acercarse lo más que se pueda a la verdad.

Algunos de los principales desarrollos en la Teoría de Estimación que precedieron al Filtro de Kalman y que resultan ser fundamentales son los siguientes: el método de máxima verosimilitud desarrollado por R.A. Fisher en 1912. Kolmogorov en 1941 y Wiener en 1942 desarrollaron una técnica de estimación del mínimo cuadrado medio.

Una nueva versión de los mínimos cuadrados es el Filtro de Kalman. En términos intuitivos es posible señalar que esta técnica es una forma moderna de analizar el método de los mínimos cuadrados, y que este último es un caso particular del Filtro de Kalman. La diferencia esencial entre el Filtro de Kalman y el modelo lineal convencional es que el estado natural (o parámetros) no es constante, sino que puede cambiar con el tiempo.

Existen algunas diferencias importantes entre las técnicas modernas y las clásicas. Como la teoría moderna de probabilidad no existía al iniciar el siglo XIX, las consideraciones de Gauss eran necesariamente de naturaleza determinística. La llegada de modernas tecnologías de computación permitieron la realización de modelos matemáticos más complejos. En la era de Gauss las técnicas que se desarrollaron sólo se apoyaban en lápiz y papel, y en su inteligencia.

El Filtro de Kalman ha sido utilizado por ingenieros y físicos en temas como: la determinación de órbitas, en la navegación, en exploración geofísica, en procesos industriales, en aplicaciones oceanográficas, entre otros. Más recientemente se han realizado aplicaciones en campos como la economía, a fin de realizar pronósticos de corto plazo y en la regulación de tiempos muertos en procesos industriales.

¿Cuándo se descubrió el Filtro de Kalman?

En noviembre de 1958 Kalman regresaba en tren a Baltimore después de haber estado en Princeton, cuando se le ocurrió aplicar la noción de variables de estado al Filtro de Wiener. Su teoría fue vista con escepticismo, y para la publicación de su artículo escogió una revista de Ingeniería Eléctrica. Su segundo artículo, donde trata el caso continuo, fue rechazado la primera vez.

Es importante mencionar que a Kalman le ocurrió algo semejante que a Gauss. Kalman publicó su primer trabajo en 1960 y Peter Swerling en 1958, publicó un memorándum en RAND CORPORATION describiendo un método recursivo para la determinación de órbitas. El método de Swerling es en esencia similar al de Kalman. Mientras que la controversia de quién inventó la técnica entre Gauss-Legendre fue analizada y discutida en el ámbito académico, la controversia Kalman-Swerling no ha trascendido.

En el año de 1960 en ocasión de la visita de Kalman a Stanley F. Schmidt a la NASA (National Aeronautics and Space Administration) se analizó el potencial de aplicación del Filtro de Kalman a la estimación de la trayectoria del Proyecto Apolo, el cual tenía como propósito ir a la luna y regresar.