

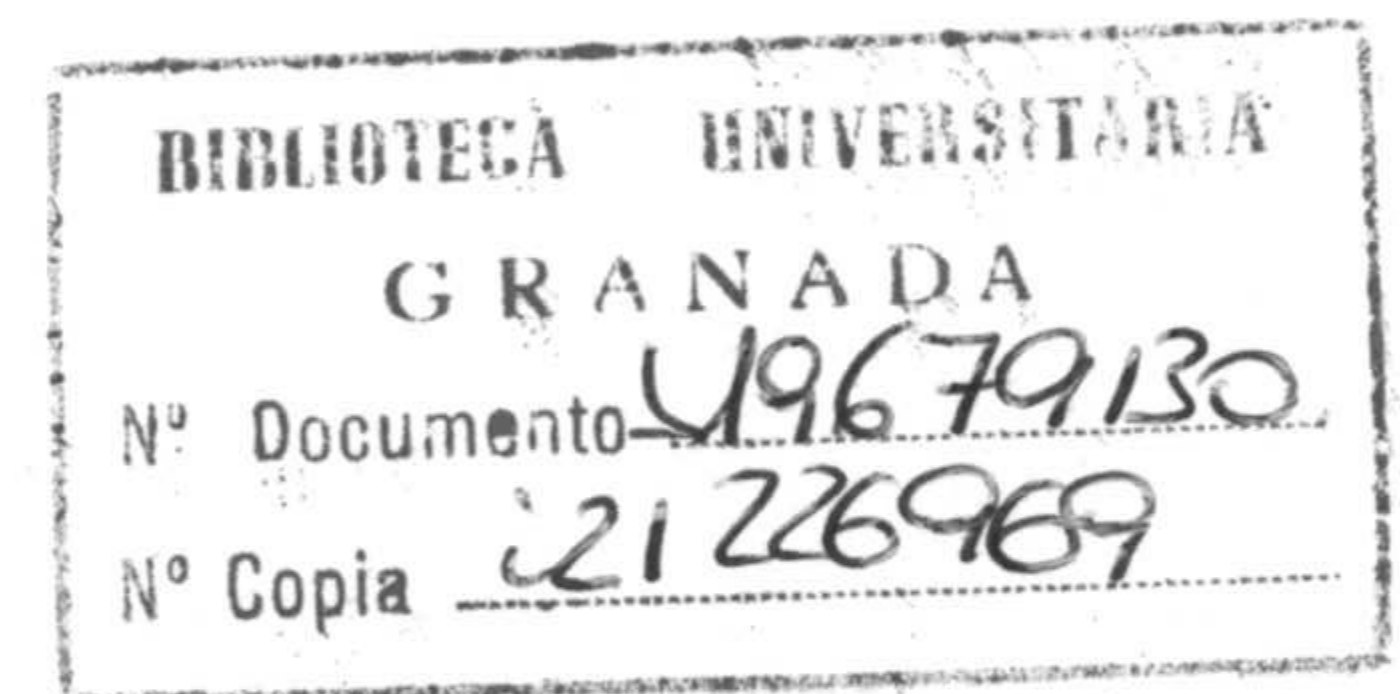
UNIVERSIDAD DE GRANADA
FACULTAD DE CIENCIAS



Departamento de Estadística e Investigación Operativa

**FILTRADO POLINOMIAL EN SISTEMAS CON
OBSERVACIONES INCIERTAS**

Raquel Caballero Águila



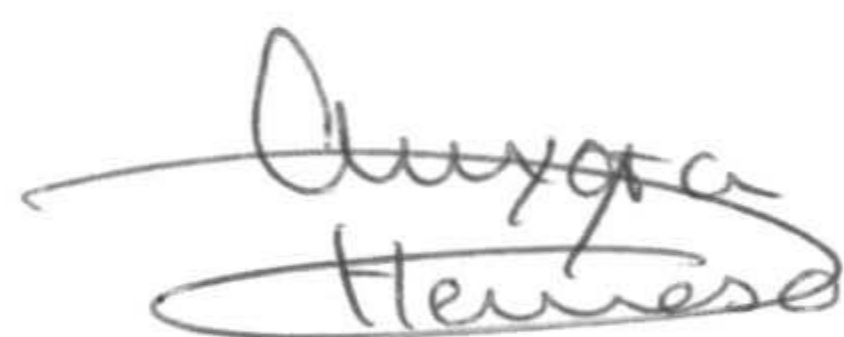
TESIS DOCTORAL

Granada, 1999

FILTRADO POLINOMIAL EN SISTEMAS CON
OBSERVACIONES INCIERTAS

Memoria presentada por
Raquel Caballero Águila para
optar al grado de Doctor por la
Universidad de Granada

Vº Bº
Directoras de Tesis



Dra. D^a Aurora Hermoso Carazo



Dra. D^a Josefa Linares Pérez

DEPARTAMENTO DE ESTADÍSTICA E INVESTIGACIÓN OPERATIVA
FACULTAD DE CIENCIAS
UNIVERSIDAD DE GRANADA



FILTRADO POLINOMIAL EN SISTEMAS CON
OBSERVACIONES INCIERTAS

El doctorando



Raquel Caballero Águila

DEPARTAMENTO DE ESTADÍSTICA E INVESTIGACIÓN OPERATIVA
FACULTAD DE CIENCIAS
UNIVERSIDAD DE GRANADA

Deseo manifestar mi más sincero agradecimiento a todas las personas que, con su ayuda y estímulo, han hecho posible la realización de esta memoria. En primer lugar, a las Profesoras Dra. D^a Aurora Hermoso Carazo y Dra. D^a Josefa Linares Pérez, directoras de la misma, por su excelente labor y constante apoyo. Asimismo, al Profesor Dr. D. Francisco Torres Ruíz por su asesoramiento y por el interés mostrado. No quisiera concluir sin agradecer a mi familia, amigos y compañeros del Departamento de Estadística e Investigación Operativa de la Universidad de Jaén las continuas muestras de ánimo y confianza recibidas.

Índice General

Introducción General	v
1 Estimación en Sistemas Lineales Discretos	1
1.1 Antecedentes Históricos	1
1.2 Estimación Óptima en Sistemas Discretos	4
1.2.1 Formulación y Solución del Problema de Estimación	4
1.2.2 Estimación Lineal y Proyecciones Ortogonales	7
1.3 Estimación en Sistemas Lineales Discretos Gaussianos	9
1.3.1 Descripción del Sistema	9
1.3.2 Filtro de Kalman	10
1.4 Estimación en Sistemas con Observaciones Inciertas	13
1.4.1 Sistemas con Observaciones Inciertas	14
1.4.2 Filtro Óptimo en Sistemas con Observaciones Inciertas	16
1.4.3 Filtro Lineal en Sistemas con Observaciones Inciertas	18
2 Estimación Cuadrática en Sistemas con Observaciones Inciertas	25
2.1 Introducción	25
2.2 Estimación Cuadrática y Proyecciones Ortogonales	26

2.3	Sistemas con Incertidumbre Modelizada por Variables Aleatorias Independientes	29
2.3.1	Hipótesis sobre el Sistema	30
2.3.2	Filtro Cuadrático	32
2.3.3	Filtro Steady-State	34
2.4	Sistemas con Incertidumbre Modelizada por Variables Aleatorias Arbitrarias	35
2.4.1	Descripción del Sistema y Planteamiento del Problema	36
2.4.2	Sistema Aumentado	38
2.4.3	Algoritmo de Filtrado	44
2.5	Sistemas Estacionarios	47
2.5.1	Sistema Aumentado Asintóticamente Estacionario	48
2.5.2	Existencia del Filtro Steady-State	55
3	Estimación Polinomial en Sistemas con Observaciones Inciertas	69
3.1	Introducción	69
3.2	Estimación Polinomial y Proyecciones Ortogonales	71
3.3	Descripción del Sistema y Planteamiento del Problema de Estimación	72
3.4	Resultados Previos	74
3.5	Sistema Aumentado. Algoritmo de Filtrado Lineal	81
3.5.1	Caso de Ruidos Aditivos Independientes	83
3.5.2	Caso de Ruidos Aditivos Correlados	90
3.6	Sistemas Estacionarios	96
3.7	Filtro Polinomial Steady-State	105
	Ejemplo Numérico	111

A	Algunos Resultados de Cálculo Matricial	125
A.1	Producto Kronecker de Matrices	125
A.2	Operador <i>vec</i> o <i>stack</i>	126
A.3	Matriz Conmutación o Permutación	127
A.4	Potencia Kronecker de un Binomio	128
B	Algoritmo de Predicción y Filtrado para Sistemas con Ruido de la Observación Correlado	129
B.1	Predictor en Función del Filtro	130
B.2	Proceso Innovación	131
B.3	Filtro en Función del Predictor	133
B.4	Predictor del Ruido de la Observación	134
B.5	Algoritmo para el Problema de Filtrado	137
C	Algoritmo de Filtrado para Sistemas con Observaciones Inciertas y Ruidos Correlados	139
C.1	Predictor en Función del Filtro	143
C.2	Proceso Innovación	144
C.3	Filtro en Función del Predictor	146
C.4	Filtro del Ruido del Estado	147
C.5	Algoritmo Recursivo	149
	Bibliografía	151

Introducción General

El objetivo de la Teoría de Sistemas Dinámicos es el estudio de los cambios que experimentan los sistemas físicos a lo largo del tiempo debido a la influencia de factores externos y/o internos. En la mayor parte de las situaciones reales, las perturbaciones que afectan a un sistema son aleatorias, por lo que tanto el estado como la observación del mismo tienen también carácter aleatorio. Estos sistemas se denominan *Sistemas Estocásticos*.

Para explicar el comportamiento de un sistema estocástico ha de conocerse el modelo matemático que lo rige, así como las condiciones iniciales que posee al comienzo del estudio. Determinado dicho modelo, pueden utilizarse diversas técnicas de estimación para predecir valores, en el transcurso del tiempo, de la variable en estudio. Dicha variable vendrá representada matemáticamente por el vector estado que, en general, no será accesible directamente, sino que dispondremos de unas medidas u observaciones perturbadas, relacionadas con él, pero que no proporcionan un valor exacto del mismo. Por tanto, se planteará la necesidad de estudiar distintos criterios de optimalidad bajo los cuales el estimador del estado sea el mejor posible y asociar a éste una medida que refleje, en cada instante, la bondad del proceso de estimación.

El problema de estimar el vector estado de un sistema dinámico tiene una larga

historia. El interés del hombre por realizar predicciones futuras sobre fenómenos naturales se remonta a los tiempos de los babilonios, si bien el inicio de la Teoría de la Estimación, propiamente dicha, puede situarse en el siglo XVII, en el que Galileo Galilei estudia la forma de minimizar varias funciones de error.

A finales del siglo XVIII, Gauss aplicó el método de Mínimos Cuadrados a la determinación de la órbita del asteroide Ceres y, en 1821, desarrolló una variante recursiva de este método que permitía corregir un estimador previo, tras una nueva observación, sin necesidad de repetir todos los cálculos. Las primeras aplicaciones en el campo de los procesos comienzan sobre los años cuarenta del siglo XX, con Kolmogorov, Krein y Wiener, que desarrollaron una técnica de estimación lineal mínimo cuadrática para procesos estocásticos.

A partir de los años cincuenta, los esfuerzos se dirigen hacia la eliminación de las restricciones impuestas por la teoría de Wiener y Kolmogorov. El primer intento significativo de extender la teoría se debe a Zadeh y Ragazzini y tuvo lugar en 1950. La necesidad de obtener algoritmos recursivos fue inminente y Follin, en 1955, sugirió una aproximación recursiva para un sistema específico. Esta aproximación fue el fundamento para desarrollos posteriores, entre los que cabe destacar los trabajos de Swerling en 1958, que presentan algoritmos recursivos para la determinación de órbitas que pronto serían aplicados y extendidos.

En 1960, Kalman [24] publicó un algoritmo algo más restringido que el de Swerling y modificó la formulación convencional del problema, utilizando lo que se conoce como Modelo de Espacio de Estados. De forma más concreta, Kalman obtuvo el estimador de mínimos cuadrados (la esperanza condicionada del estado dadas las observaciones) en sistemas lineales discretos gaussianos; dicho estimador, debido a la propiedad de gaussianidad, es una función lineal de las observaciones. En sistemas no gaussianos el cálculo de la esperanza condicionada, en general, no puede realizarse de forma sencilla

y, en este caso, el Filtro de Kalman proporciona únicamente el estimador lineal de mínimos cuadrados. Más tarde, junto a Bucy, Kalman dedujo resultados similares para el caso continuo.

Numerosas publicaciones aparecen, posteriormente, sobre extensiones del filtro de Kalman y se proponen distintas formas de obtener las ecuaciones del filtro; citemos, por ejemplo, Anderson y Moore [1], Brammer y Siffing [3], Caines [5], Catlin [6], Grewal y Andrews [13], Kailath [20], [23], Tanizaki [34], Valderrama y Ruiz [37], etc. También se han obtenido algoritmos para el problema de suavizamiento (Kailath y Frost [21], Meditch [27], etc.). Además, se han establecido conexiones con otros tipos de estimación (bayesiano, máxima verosimilitud, etc.), como puede verse en Jazwinski [19].

Un resumen histórico más completo sobre los antecedentes y el desarrollo de la Teoría de Estimación en sistemas dinámicos puede encontrarse en Kailath [22] o en Sorenson [33].

En la Teoría de Estimación de Sistemas Estocásticos Lineales se supone que la señal que se desea estudiar está siempre presente en las observaciones y que la perturbación de éstas se debe, únicamente, a un ruido aditivo. En este caso, los estimadores tienen una estructura recursiva bastante simple, debido a la linealidad, tanto de la ecuación del estado como de la ecuación de la observación. Sin embargo, en muchas aplicaciones, el vector estado interviene en la observación de forma aleatoria, de modo que la probabilidad de que las observaciones contengan únicamente ruido es positiva (*probabilidad de falsa alarma*). En estas situaciones, las condiciones necesarias para aplicar el Filtro de Kalman no se satisfacen, al estar las observaciones perturbadas por un ruido escalar multiplicativo. Los sistemas que describen dichas situaciones se denominan *Sistemas con Ruido Multiplicativo* y, aún bajo hipótesis de gaussianidad sobre los ruidos y el estado inicial, no son sistemas gaussianos. Cuando el ruido que multiplica a la matriz de las observaciones es una sucesión de variables aleatorias de Bernoulli, se habla de

Sistemas con Observaciones Inciertas.

Nahi [31] obtuvo, en 1969, el filtro lineal de mínimos cuadrados en sistemas con observaciones inciertas, suponiendo que el ruido multiplicativo era una sucesión de variables aleatorias independientes de Bernoulli y que los ruidos del estado y la observación eran incorrelados entre sí. Dicho estimador responde a una estructura similar al Filtro de Kalman. Monzingo [29] completó los resultados de Nahi estudiando el problema de suavizamiento y Tugnait [35] analizó la estabilidad del estimador. Hermoso y Linares [15], [16] trataron los problemas de filtrado y suavizamiento en el caso en que los ruidos aditivos del estado y la observación son correlados.

En 1971, Jaffer y Gupta [18] establecieron un algoritmo recursivo para el estimador óptimo (en el sentido de mínimos cuadrados) en sistemas con observaciones inciertas. Este estimador que, como antes indicamos, es la esperanza condicionada, requiere para su obtención un crecimiento exponencial de memoria, por lo que, a partir de entonces, los esfuerzos se dirigen hacia la búsqueda de estimadores subóptimos.

En 1979, Hadidi y Schwartz [14] generalizaron los resultados de Nahi, eliminando la hipótesis de independencia del ruido multiplicativo. Probaron además que, en general, el filtro lineal de mínimos cuadrados no tiene estructura recursiva y establecieron una condición necesaria y suficiente para que el estimador fuera recursivo. Bajo esta condición, la forma del filtro recursivo resulta ser una generalización de los filtros de Kalman y Nahi. En 1984, Wang [38] obtuvo los mismos resultados que Hadidi y Schwartz utilizando un tratamiento distinto y Monzingo [30] analizó el problema de suavizamiento en 1981.

Como indicaron Chow y Birkemeier [7], [8], el caso en que el ruido multiplicativo tiene rango de valores continuo ha sido menos tratado en la literatura. Rajasekaran et al. [32] obtuvieron estimadores lineales de mínimo error cuadrático medio recursivos para el caso de ruido blanco y Tugnait [36] estudió la estabilidad de los mismos. Her-

moso y Linares [17] generalizaron el trabajo de Rajasekaran et al. considerando el caso en que los ruidos aditivos que afectan al sistema son correlados. Chow y Birkemeier [7], [8] consideraron un modelo con ruido multiplicativo descrito por una ecuación dinámica y dedujeron una nueva estructura recursiva para el estimador distinta a la de Kalman.

Como ya se ha indicado, en sistemas no gaussianos el estimador de mínimos cuadrados (esto es, la esperanza condicionada) no es fácil de obtener. En sistemas con certidumbre en las observaciones el Filtro de Kalman proporciona únicamente el estimador lineal de mínimos cuadrados y, en sistemas con observaciones inciertas, el estudio también se ha centrado en la estimación lineal de mínimos cuadrados. Recientemente, la investigación se ha dirigido hacia la búsqueda de estimadores subóptimos (en el sentido de mínimos cuadrados) que mejoren al estimador lineal, concretamente, estimadores polinomiales.

En 1995, De Santis et al. [9] trataron el problema de estimación polinomial de segundo grado en sistemas lineales con certidumbre en las observaciones. Mas tarde, en 1996, Carravetta et al. [4] generalizaron este estudio considerando el problema de estimación polinomial de cualquier grado.

Para sistemas con ruido multiplicativo, García-Ligero et al. [11], [12] abordaron el problema de estimación polinomial de segundo grado, bajo la hipótesis de independencia mutua de los ruidos y la condición inicial del sistema, tanto en el caso de sistemas con ruido multiplicativo blanco, como en el caso de sistemas con observaciones inciertas en los que la incertidumbre de las observaciones está modelizada por variables aleatorias (de Bernoulli) independientes.

El problema de estimación polinomial en sistemas lineales con observaciones inciertas constituye, en términos generales, el objetivo de esta memoria. En primer lugar, haremos un resumen de los resultados, ya establecidos, sobre el problema de estimación óptima y lineal en dichos sistemas. Seguidamente, abordaremos el problema

de estimación polinomial de segundo grado, presentando los resultados obtenidos por García-Ligero et al. [12] para el caso en que la incertidumbre de las observaciones está modelizada por variables aleatorias independientes y generalizando dichos resultados al caso en que dichas variables no son necesariamente independientes. El objetivo último es el estudio del problema de estimación polinomial de grado arbitrario en sistemas con observaciones inciertas en los que la incertidumbre está modelizada por variables aleatorias independientes. Abordaremos este problema bajo la hipótesis de independencia mutua de los ruidos aditivos que intervienen en las ecuaciones del estado y la observación y, también, cuando dichos ruidos son correlados entre sí.

A continuación describimos, de forma más precisa, el contenido de la presente memoria que está estructurada de la siguiente forma:

Capítulo 1 (Estimación en Sistemas Lineales Discretos). En este Capítulo realizamos un breve resumen sobre el problema de estimación en sistemas lineales discretos. Comenzamos planteando el problema general de estimación óptima, su solución y la interpretación geométrica del mismo. A continuación, realizamos la descripción matemática de un sistema lineal discreto gaussiano, planteamos el problema de estimación del estado a partir de una serie de observaciones y presentamos el algoritmo recursivo para la obtención del predictor en una etapa y del filtro. A lo largo de este estudio se observará que, bajo condiciones de gaussianidad, el estimador óptimo coincide con el estimador lineal de mínimos cuadrados proporcionado por el Filtro de Kalman.

Finalizamos este Capítulo analizando el problema de estimación de menor error cuadrático medio en sistemas con observaciones inciertas. En primer lugar, presentamos el algoritmo para el problema de estimación óptima, establecido por Jaffer y Gupta [18]. Este algoritmo requiere un crecimiento exponencial de memoria, por lo que nos centramos en el estudio de estimadores subóptimos. Concretamente, presentamos el algoritmo de filtrado lineal, tanto en el caso en que la incertidumbre de las observaciones

está modelizada por variables aleatorias de Bernoulli independientes (Nahi [31]), como en el caso en que dicha sucesión de variables aleatorias de Bernoulli es arbitraria (Hadidi y Schwartz [14]).

Capítulo 2 (Estimación Cuadrática en Sistemas con Observaciones Inciertas). El objetivo de este Capítulo es obtener un algoritmo recursivo para el problema de estimación polinomial de segundo grado de menor error cuadrático medio en sistemas con observaciones inciertas. Comenzamos presentando la interpretación geométrica del problema general de estimación polinomial de segundo grado o cuadrática. Seguidamente, realizamos un breve resumen del problema de estimación cuadrática en sistemas en los que la incertidumbre de las observaciones está modelizada por variables aleatorias independientes, estudio realizado por García-Ligero et al. [12]. A continuación, extendemos los resultados obtenidos en el mencionado trabajo al caso de sistemas con observaciones inciertas suprimiendo la hipótesis de independencia del ruido que modeliza la incertidumbre, obteniendo un algoritmo para el problema de filtrado polinomial de segundo grado. Finalmente, se considera la versión estacionaria de los sistemas en estudio, probando que, bajo la hipótesis de estabilidad asintótica del sistema estacionario original, el filtro polinomial de segundo grado admite una forma *steady-state*, con las ventajas computacionales que ello conlleva.

Capítulo 3 (Estimación Polinomial en Sistemas con Observaciones Inciertas). En este Capítulo analizamos el problema de estimación polinomial de grado arbitrario en sistemas con observaciones inciertas en los que la incertidumbre de las observaciones está modelizada por variables aleatorias independientes. Comenzamos presentando un enfoque geométrico del problema general de estimación polinomial de menor error cuadrático medio. Seguidamente, planteamos el problema de estimación polinomial en los sistemas bajo estudio y abordamos dicho problema, distinguiendo dos casos: en primer lugar, supondremos que los ruidos y el estado inicial del sistema son

mutuamente independientes (es decir, consideramos los sistemas estudiados por Nahi [31] y García-Ligero et al. [12]) y, a continuación, extendemos el estudio al caso en que los ruidos aditivos que intervienen en la ecuación del estado y la observación son correlados entre sí. Este estudio generaliza el trabajo de García-Ligero et al. [12] en una doble dirección: por una parte, se consideran estimadores polinomiales de grado arbitrario y, por otra, se debilita la hipótesis de independencia sobre los ruidos del sistema. Para este último caso, proponemos un algoritmo para la obtención del filtro polinomial de grado arbitrario. Para concluir, consideramos la versión estacionaria de los sistemas en estudio, concluyendo que, tanto en el caso en que los ruidos y la condición inicial son mutuamente independientes, como en el caso de ruidos aditivos correlados, si el sistema estacionario original es asintóticamente estable, el filtro polinomial admite una versión *steady-state*.

Seguidamente, con el fin de ilustrar los resultados establecidos en el Capítulo 3, presentamos un **Ejemplo Numérico** que pone de manifiesto que, tanto bajo la hipótesis de independencia mutua del estado inicial y los ruidos del sistema, como en el caso en que los ruidos aditivos son correlados entre sí, la varianza del error cometido al estimar el estado mediante una función polinómica de las observaciones disminuye al aumentar el grado de dicha función.

Por último, se incluyen tres apéndices donde se presentan los resultados sobre Álgebra Matricial básicos para el desarrollo de la memoria (**Apéndice A**), así como la deducción de los algoritmos de filtrado establecidos en los Capítulos 2 y 3 (**Apéndice B** y **Apéndice C**, respectivamente).

Capítulo 1

Estimación en Sistemas Lineales Discretos

1.1 Antecedentes Históricos

Para explicar el comportamiento de un sistema dinámico, ha de conocerse el modelo matemático que lo rige, así como las condiciones iniciales que posee al comienzo del estudio. La variable bajo consideración vendrá representada a través del vector estado que, en general, no será accesible directamente, sino que dispondremos de una serie de medidas u observaciones perturbadas, relacionadas con él, pero que no proporcionan un valor exacto del mismo. En la predicción de valores de dicha variable pueden utilizarse diversas técnicas, por lo que debemos enunciar criterios de optimalidad para el estimador y acompañar a éste de una medida de la bondad de la estimación.

El problema de estimar el vector estado de un sistema dinámico tiene una larga historia. El interés del hombre por realizar predicciones futuras sobre fenómenos naturales, a través de la experimentación y la observación, se remonta a los tiempos de los babilonios, si bien el inicio de la Teoría de la Estimación, propiamente dicha, puede

situarse en el siglo XVII, en el que Galileo Galilei estudia la forma de minimizar varias funciones de error.

La aparición de los primeros modelos dinámicos tiene lugar en el campo de la Física, cuando Newton enuncia sus conocidas leyes de la Mecánica y de la Gravitación Universal. Posteriormente, estos modelos matemáticos se van perfeccionando y los procedimientos de estimación y predicción se hacen cada vez más rigurosos. A finales del siglo XVIII, Gauss aplica el método de Mínimos Cuadrados a la determinación de la órbita de cuerpos celestes. En 1821, Gauss desarrolló también una variante recursiva de este método, que permitía corregir un estimador previo tras una nueva observación, sin tener que repetir todos los cálculos.

Las primeras aplicaciones de dicha técnica en el campo de Procesos Estocásticos comienzan sobre los años cuarenta del siglo XX, con Kolmogorov, Krein y Wiener. En 1939, Kolmogorov consideró procesos estacionarios en tiempo discreto y en 1945, Krein extendió los resultados al caso continuo. De manera independiente, Wiener estudió, en 1942, el método de estimación de Mínimos Cuadrados en procesos estacionarios continuos. Wiener utilizó técnicas variacionales para determinar una fórmula explícita del estimador óptimo y estableció una condición necesaria y suficiente de optimalidad dada mediante una ecuación integral del tipo Wiener-Hopf, que había sido resuelta por él mismo y por Hopf en 1931.

En 1950, Bode y Shanon publicaron una aproximación (distinta a la de Wiener) para la solución de la ecuación de Wiener-Hopf.

Hasta entonces, los estudios se centraban en procesos estacionarios con intervalos de observación infinitos o semiinfinitos. A partir de los años cincuenta, los esfuerzos se dirigen hacia la eliminación de las restricciones impuestas por la teoría de Wiener y Kolmogorov. El primer intento significativo de extender la teoría se debe a Zadeh y Ragazzini y tuvo lugar en 1950. En 1952, Booton generalizó la ecuación integral de

Wiener-Hopf al caso de procesos no estacionarios, pero su solución era más teórica que práctica.

Hacia 1955, Follin y otros consideraron el problema de filtrado en el caso de procesos estacionarios con intervalo de observación finito y establecieron ecuaciones diferenciales para el estimador en este caso. Además, probaron que la solución de estas ecuaciones diferenciales coincidía con la de Wiener si el intervalo de observación se hacía tender a infinito. Bucy mostró, en 1959, que el caso de procesos no estacionarios también podía resolverse usando estas ecuaciones. Un estudio similar para el caso discreto fue llevado a cabo por Swerling en 1958. Además, en 1959, Swerling desarrolló algunos algoritmos recursivos que pronto serían aplicados y extendidos.

En 1960, Kalman [24] publicó un algoritmo algo más restringido que el de Swerling y modificó la formulación convencional del problema, utilizando lo que se conoce como Modelo de Espacio de Estados. En principio, Kalman obtuvo resultados para el caso discreto y, más tarde, junto a Bucy, dedujo resultados similares para el caso continuo; aplicando un argumento límite, obtuvieron una ecuación recursiva en tiempo continuo, a partir de la ecuación en tiempo discreto, para la estimación mínimo cuadrática. La aportación de Bucy fue la obtención del filtro a través de la ecuación de Wiener-Hopf.

Cabe destacar, también, a otros investigadores que, con sus estudios, colaboraron en la obtención del Filtro de Kalman. Tal es el caso de Wold, que en 1938 utilizó la idea, propuesta por Fréchet en 1937, de considerar las variables aleatorias como elementos de un espacio métrico, siendo la distancia entre dos variables la varianza de su diferencia. Con esta interpretación geométrica, el estimador de mínimos cuadrados es la proyección ortogonal sobre un subespacio.

Numerosas publicaciones aparecen, posteriormente, sobre extensiones del Filtro de Kalman y se proponen distintas formas de obtener las ecuaciones del filtro; citemos, por ejemplo, Anderson y Moore [1], Brammer y Siffing [3], Caines [5], Catlin [6], Gre-

wal y Andrews [13], Kailath [20], Korbicz y Bidyuk [25], Meditch [28], etc. También se han obtenido, utilizando distintas técnicas, algoritmos para el problema de suavizado (Anderson y Moore [1], Kailath y Frost [21], Meditch [27], etc.). Asimismo, se han establecido conexiones con otros procedimientos de estimación (bayesiano, máxima verosimilitud, etc.), como puede verse en Jazwinski [19].

Actualmente, el Filtro de Kalman se usa en diversos campos, como son sistemas aeroespaciales y de navegación, ingeniería química, control de procesos industriales, teoría de comunicaciones, meteorología y un largo etcétera.

Un resumen histórico más completo sobre los antecedentes y el desarrollo de la Teoría de Estimación en sistemas dinámicos puede encontrarse en Kailath [22] o Sorenson [33].

1.2 Estimación Óptima en Sistemas Discretos

1.2.1 Formulación y Solución del Problema de Estimación

Consideremos un sistema cuyo estado se manifiesta a través del tiempo mediante un proceso estocástico n -dimensional, $\{x(k); k \in I\}$, definido sobre un espacio probabilístico (Ω, A, P) , donde I es un conjunto de índices discreto. Supongamos que se está interesado en conocer el valor de $x(k)$ para algún k fijo, pero $x(k)$ no es directamente accesible para el investigador, sino que sólo se dispone de una serie de medidas u observaciones $z(0), \dots, z(j)$ relacionadas de algún modo con $x(k)$ y se quiere utilizar estas observaciones para estimar el valor de $x(k)$. Supondremos que $\{z(i); i = 0, \dots, j\}$ es un proceso estocástico m -dimensional, definido sobre (Ω, A, P) .

Puesto que sólo disponemos de las medidas $z(0), \dots, z(j)$ para estimar $x(k)$, un estimador de $x(k)$ basado en dichas medidas, que notaremos mediante $\hat{x}(k/j)$, será

una función vector valuada n -dimensional de las mismas, esto es

$$\hat{x}(k/j) = \Phi_k [z(i); i = 0, \dots, j].$$

Por consiguiente, el problema se reduce a determinar Φ_k de forma que el estimador sea óptimo en algún sentido. Si $k > j$, el problema se denomina de *predicción*; si $k = j$, de *filtrado*; y, si $k < j$, de *suavizamiento*. Para abordar estos problemas, consideraremos el error de estimación, definido como

$$\tilde{x}(k/j) = x(k) - \hat{x}(k/j).$$

En el caso ideal, $\tilde{x}(k/j) = 0$, el estimador sería óptimo, pero, generalmente, $\tilde{x}(k/j) \neq 0$, por lo que debemos asignar una penalización o pérdida al error de estimación. Con este fin, se considera una función $L = L[\tilde{x}(k/j)]$ con las siguientes propiedades:

1. L es una función real de n variables.
2. $L(0) = 0$.
3. $L[\tilde{x}^b(k/j)] \geq L[\tilde{x}^a(k/j)]$, siempre que $\rho[\tilde{x}^b(k/j)] \geq \rho[\tilde{x}^a(k/j)]$, siendo ρ una función real, no negativa y convexa.
4. $L[\tilde{x}(k/j)] = L[-\tilde{x}(k/j)]$.

Toda función que verifique estas cuatro propiedades se denomina *función de pérdida admisible*.

Debido al carácter aleatorio de $x(k)$ y $\hat{x}(k/j)$, se sigue que $\tilde{x}(k/j)$ es también un vector aleatorio y, por tanto, L será una variable aleatoria. De esta forma, una medida útil de la pérdida global es la pérdida media

$$J[\tilde{x}(k/j)] = E\{L[\tilde{x}(k/j)]\}$$

y, por tanto, el estimador óptimo bajo dicha función de pérdida es aquel que minimiza la pérdida esperada del error de estimación.

Bajo determinadas condiciones sobre la distribución condicionada de $x(k)$ a $\{z(0), \dots, z(j)\}$ o sobre la función de pérdida, el estimador óptimo puede ser caracterizado de forma simple, como se establece en los siguientes resultados.

Teorema 1.2.1 (Sherman) *Si la función de distribución condicionada de $x(k)$, dadas las observaciones $\{z(0), \dots, z(j)\}$, es simétrica respecto de su media y convexa para todo valor menor o igual que la media, entonces, para cualquier función de pérdida admisible, el estimador óptimo es*

$$\hat{x}(k/j) = E[x(k) / z(0), \dots, z(j)].$$

Una consecuencia inmediata que se deduce del Teorema 1.2.1 es que, si los procesos $\{x(k); k \in I\}$ y $\{z(i); i = 0, \dots, j\}$ verifican las hipótesis de este teorema y son independientes, entonces el estimador óptimo es

$$\hat{x}(k/j) = E\{x(k)\}.$$

Por otra parte, si la distribución condicionada de $x(k)$, dadas las observaciones $\{z(0), \dots, z(j)\}$, es gaussiana, se verifican las hipótesis del Teorema 1.2.1 y podemos enunciar el siguiente resultado.

Corolario 1.2.2 (Doob) *Si los procesos $\{x(k); k \in I\}$ y $\{z(i); i = 0, \dots, j\}$ son conjuntamente gaussianos, entonces, para cualquier función de pérdida admisible, el estimador óptimo es*

$$\hat{x}(k/j) = E[x(k) / z(0), \dots, z(j)].$$

Como consecuencia de este corolario, y teniendo en cuenta la expresión de la esperanza condicionada bajo hipótesis de normalidad, se tiene que, en el caso gaussiano, el estimador óptimo es lineal.

Consideremos ahora una función de pérdida específica, concretamente, una función de pérdida cuadrática

$$L[\tilde{x}(k/j)] = \tilde{x}^T(k/j) \tilde{x}(k/j).$$

El estimador que minimiza la pérdida esperada se denomina *estimador de mínimos cuadrados* o de *menor error cuadrático medio*.

Para este tipo de funciones de pérdida se tiene el siguiente resultado.

Teorema 1.2.3 (Doob) *Si la función de pérdida es cuadrática, entonces el estimador óptimo es*

$$\hat{x}(k/j) = E[x(k) / z(0), \dots, z(j)].$$

Comparando el Teorema 1.2.1 y el Teorema 1.2.3, se observa que el segundo, a pesar de ser más restrictivo respecto a la clase de funciones de pérdida, es más general que el primero, en lo que a la clase de procesos estocásticos considerados se refiere. Ambos teoremas se complementan en el sentido de que, si un problema satisface las hipótesis de Teorema 1.2.1, pero no las del Teorema 1.2.3, o viceversa, el estimador óptimo es la esperanza condicionada, en cualquiera de los casos.

1.2.2 Estimación Lineal y Proyecciones Ortogonales

Se ha indicado en la Sección anterior (Corolario 1.2.2) que, en el caso gaussiano, la esperanza condicionada es una función lineal de las observaciones y, por tanto, el estimador óptimo para cualquier función de pérdida admisible es lineal. Teniendo ahora en cuenta el Teorema 1.2.3, podemos concluir que, en el caso gaussiano, el estimador óptimo es el estimador lineal de menor error cuadrático medio. En esta Sección presentamos una interpretación geométrica de dicho estimador, que se obtiene mediante el Teorema de la Proyección Ortogonal.

Consideremos los vectores observación $z(0), \dots, z(j)$, y denotemos por $\mathcal{Y}(j)$ el subespacio lineal del espacio de Hilbert $\mathcal{L}_{\mathbb{R}^n}^2(\Omega, A, P)$ ¹ generado por dichos vectores. En virtud del Teorema de la Proyección Ortogonal, todo vector $x \in \mathcal{L}_{\mathbb{R}^n}^2(\Omega, A, P)$ se puede descomponer de forma única como suma de dos vectores, un vector $\bar{x} \in \mathcal{Y}(j)$, y un vector \tilde{x} ortogonal a $\mathcal{Y}(j)$. El vector \bar{x} se denomina *proyección ortogonal* de x sobre $\mathcal{Y}(j)$ y es aquel vector de $\mathcal{Y}(j)$ que minimiza la pérdida cuadrática media.

Por tanto, la proyección ortogonal, \bar{x} , es el estimador lineal de mínimos cuadrados de x basado en $\{z(0), \dots, z(j)\}$ y \tilde{x} es el error de estimación.

Para el problema de estimación formulado en la Sección anterior, estos resultados se resumen en el siguiente teorema.

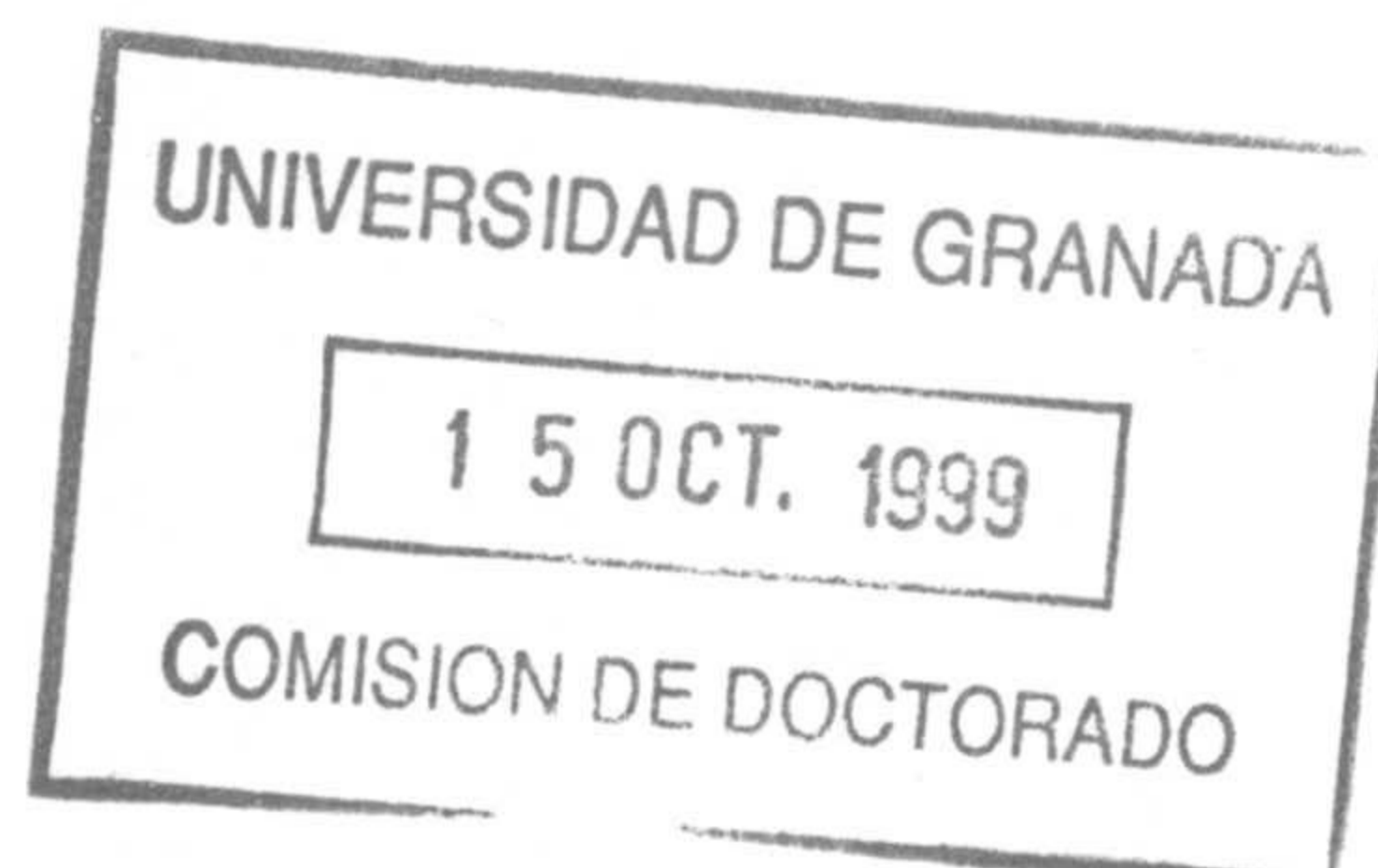
Teorema 1.2.4 Sean $\{x(k); k \in I\}$ y $\{z(k); k \in I\}$ procesos estocásticos con media cero. Supongamos que observamos $z(0), \dots, z(j)$ y que se da alguna de las condiciones siguientes:

- (i) Los procesos $\{x(k); k \in I\}$ y $\{z(k); k \in I\}$ son conjuntamente gaussianos.
- (ii) La función de pérdida considerada es cuadrática y nos restringimos a la clase de funciones lineales de las observaciones.

Entonces el estimador óptimo de $x(k)$ dadas $z(0), \dots, z(j)$ viene dado por la proyección ortogonal de $x(k)$ sobre $\mathcal{Y}(j)$.

Las siguientes propiedades, consecuencias inmediatas del Teorema de la Proyección Ortogonal, son básicas para la obtención del estimador lineal de mínimos cuadrados:

¹ $\mathcal{L}_{\mathbb{R}^n}^2(\Omega, A, P)$ denota el conjunto de clases de equivalencia de vectores aleatorios n -dimensionales, definidas sobre el espacio probabilístico (Ω, A, P) , con momento de segundo orden finito. Como es conocido, $\mathcal{L}_{\mathbb{R}^n}^2(\Omega, A, P)$ es un espacio de Hilbert con el producto escalar $\langle Z, W \rangle = E\{Z^T W\}$.



- El estimador lineal de mínimos cuadrados, $\hat{x}(k/j)$, es el único elemento del subespacio generado mediante transformaciones lineales de $z(0), \dots, z(j)$ que verifica

$$E \{ \hat{x}(k/j) z^T(\alpha) \} = 0, \quad \alpha = 0, \dots, j \quad (1.1)$$

- $\hat{x}(k/j)$ puede expresarse como

$$\hat{x}(k/j) = \hat{x}(k/j-1) + F(k, j) [z(j) - \hat{z}(j/j-1)] \quad (1.2)$$

siendo $F(k, j)$ una matriz a determinar, denominada *ganancia del estimador*.

1.3 Estimación en Sistemas Lineales Discretos Gaussianos

1.3.1 Descripción del Sistema

Consideramos un sistema lineal en tiempo discreto cuya descripción matemática se realiza mediante las siguientes ecuaciones del estado y la observación.

Ecuación del Estado

Es una ecuación vectorial en diferencias, también denominada ecuación de transición

$$\begin{aligned} x(k+1) &= \Phi(k) x(k) + \Gamma(k) w(k), \quad k \geq 0 \\ x(0) &= x_0 \end{aligned} \quad (1.3)$$

donde $\{x(k); k \geq 0\}$ es un proceso estocástico n -dimensional que representa el *estado* del sistema. $\{w(k); k \geq 0\}$ es un proceso estocástico r -dimensional; el vector $w(k)$ se denomina *vector de perturbación aleatoria del estado* o *ruido del estado*. $\Phi(k)$ y $\Gamma(k)$ son matrices determinísticas conocidas de dimensiones apropiadas, denominadas *matriz de transición del estado* y *matriz de transición* o de *ponderación de la perturbación aleatoria*, respectivamente.

Ecuación de la Observación

La ecuación de la observación es una ecuación vectorial,

$$z(k) = H(k)x(k) + v(k), \quad k \geq 0 \quad (1.4)$$

donde $\{z(k); k \geq 0\}$ es un proceso estocástico m -dimensional; $z(k)$ representa el *vector de medidas u observaciones* en el instante k . $\{v(k); k \geq 0\}$ es un proceso estocástico m -dimensional; el vector $v(k)$ se llama *vector del error de medidas o ruido de las medidas*. $H(k)$ es una matriz determinística conocida de dimensión apropiada, llamada *matriz de medida u observación*.

Sobre los ruidos y el estado inicial del sistema se imponen las siguientes hipótesis

1. El estado inicial, x_0 , es un vector aleatorio n -dimensional gaussiano con media cero y matriz de covarianzas P_0 .
2. $\{w(k); k \geq 0\}$ es un proceso ruido blanco gaussiano, centrado, con matriz de covarianzas $Q(k)$.
3. El proceso $\{v(k); k \geq 0\}$ es una sucesión ruido blanco gaussiana, centrada, con covarianza $R(k)$.
4. El estado inicial, x_0 , y los ruidos aditivos, $\{w(k)\}$ y $\{v(k)\}$, son mutuamente independientes.

Las matrices $\Phi(k)$, $\Gamma(k)$, $Q(k)$, $R(k)$ se denominan *matrices del sistema*. Cuando estas matrices no dependen de k , se dice que el sistema es *invariante en el tiempo*.

1.3.2 Filtro de Kalman

Consideremos el sistema lineal en tiempo discreto (1.3) y (1.4), definido en el apartado anterior. Debido a las hipótesis de gaussianidad e independencia mutua (hipótesis 1-

4) y teniendo en cuenta los resultados presentados en el apartado 1.2.1, el estimador óptimo, $\hat{x}(k/j)$, para cualquier función de pérdida admisible, es

$$\hat{x}(k/j) = E[x(k) / z(0), \dots, z(j)]$$

y tiene la siguiente expresión

$$\hat{x}(k/j) = E\{x(k) Z_j^T\} (E\{Z_j Z_j^T\})^{-1} Z_j$$

donde Z_j denota el vector $(j+1)m$ -dimensional

$$Z_j = (z^T(0), \dots, z^T(j))^T$$

Por tanto, el estimador óptimo, para cualquier función de pérdida admisible, es el estimador lineal de menor error cuadrático medio.

Desde el punto de vista teórico, el problema de estimación está resuelto, sin embargo, la utilidad práctica de este estimador es bastante limitada ya que como puede verse en su expresión, es necesario invertir una matriz de dimensión $(j+1)m \times (j+1)m$, siendo $j+1$ el número de medidas y m el número de componentes del vector de medidas. Si j varía, y se desea realizar estimaciones "on-line", esto es, procesar la información conforme ésta se va haciendo disponible, la aplicación de la expresión anterior para generar el estimador óptimo se vuelve impracticable.

Desde el punto de vista de las aplicaciones, lo deseable sería disponer de algoritmos eficientes y prácticos para los distintos problemas de estimación, que procesen secuencialmente las observaciones. El más significativo de estos algoritmos es el Filtro de Kalman. Se trata de un algoritmo recursivo para el cálculo del estimador óptimo del vector estado de un sistema dinámico discreto en un instante determinado, basado en la información disponible hasta ese instante. Este filtro permite realizar estimaciones "on-line", es decir, actualizar el valor del estimador conforme se va haciendo disponible la información, sin necesidad de repetir todos los cálculos. Además, proporciona una

medida de la bondad de la estimación, a través de la matriz de covarianzas del error de estimación.

Kalman [24] obtuvo el algoritmo recursivo para el problema de estimación mínimo-cuadrática haciendo uso de la técnica de proyecciones ortogonales. Concretamente, utilizó la propiedad (1.2) que pone de manifiesto la naturaleza recursiva del estimador de mínimos cuadrados. Inicialmente, Kalman consideró un sistema en el que la ecuación de la observación no estaba perturbada por ruido; no obstante, aquí vamos a presentar el algoritmo de predicción y filtrado para los sistemas descritos por (1.3) y (1.4) bajo hipótesis de gaussianidad del estado inicial y de los ruidos. Señalemos que, si se suprime la hipótesis de normalidad, no hay garantía de que el filtro proporcione la esperanza condicionada del vector estado dadas las observaciones, si bien podemos asegurar que el filtro es el óptimo, en el sentido de mínimos cuadrados, dentro de la clase de todos los estimadores lineales.

Teorema 1.3.1 (Filtro de Kalman) *El algoritmo para los problemas de predicción y filtrado está dado por*

$$\begin{aligned}\hat{x}(k/k) &= \hat{x}(k/k-1) + F(k) [z(k) - H(k) \hat{x}(k/k-1)], \quad k \geq 0 \\ \hat{x}(k+1/k) &= \Phi(k) \hat{x}(k/k) \\ \hat{x}(0/-1) &= 0\end{aligned}$$

donde $F(k)$, matriz de ganancia del filtro, satisface

$$F(k) = P(k/k-1) H^T(k) [H(k) P(k/k-1) H^T(k) + R(k)]^{-1}$$

y las matrices de covarianzas de los errores de filtrado y predicción verifican

$$\begin{aligned}P(k/k) &= [I - F(k) H(k)] P(k/k-1), \quad k \geq 0 \\ P(k+1/k) &= \Phi(k) P(k/k) \Phi^T(k) + \Gamma(k) Q(k) \Gamma^T(k) \\ P(0/-1) &= P_0.\end{aligned}$$

La expresión que multiplica a la matriz de ganancia en la ecuación del filtro, diferencia entre la observación actual $z(k)$ y su estimador basado en las observaciones anteriores, $\{z(0), \dots, z(k-1)\}$, representa la nueva información que aporta la última observación; por este motivo, se denomina *innovación*. El proceso innovación es un ruido blanco gaussiano, centrado, con covarianzas $H(k)P(k/k-1)H^T(k) + R(k)$.

Bajo la hipótesis de que $R(k)$ sea definida positiva, la no singularidad de estas matrices de covarianzas está garantizada. Sin esta hipótesis, dichas matrices podrían ser singulares, en cuyo caso, el algoritmo puede aplicarse sin más que sustituir la inversa por la pseudoinversa (véase Anderson y Moore [1] o Catlin [6]).

1.4 Estimación en Sistemas con Observaciones Inciertas

Como se ha indicado en la Sección 1.1, el problema de estimar el estado de un sistema lineal a partir de las observaciones ha sido ampliamente estudiado. En los sistemas que se han considerado en las secciones previas se supone que la señal que se desea estimar está siempre presente en dichas observaciones y que la perturbación de éstas se debe, únicamente, a un ruido aditivo. En este caso, el Filtro de Kalman proporciona el estimador lineal de mínimos cuadrados, que coincide con el estimador óptimo para cualquier función de pérdida admisible bajo condiciones de gaussianidad e independencia mutua de la condición inicial y los ruidos del estado y las observaciones.

Sin embargo, en muchas aplicaciones prácticas, la señal interviene en la ecuación de la observación de manera aleatoria, de modo que la probabilidad de que las observaciones contengan únicamente ruido es positiva (probabilidad de falsa alarma). Esto ocurre, por ejemplo, en diversos problemas relativos a la teoría de la telecomunicación o al procesamiento de imágenes, en los que la señal está sujeta a atenuaciones aleatorias. En estas situaciones y, en general, en cualquier sistema con interrupciones aleatorias en el mecanismo de medidas, las condiciones necesarias para aplicar el Filtro de Kalman

no se satisfacen, al estar las observaciones perturbadas por un ruido escalar multiplicativo. Los sistemas que describen dichas situaciones se denominan *Sistemas con Ruido Multiplicativo*. Cuando el ruido que multiplica a la matriz de observaciones es una sucesión de variables aleatorias de Bernoulli, se habla de *Sistemas con Observaciones Inciertas*.

El objetivo de esta Sección es el estudio del problema de estimación en sistemas con observaciones inciertas. En primer lugar (apartado 1.4.1), se presenta la descripción matemática de los sistemas en estudio y se especifican las hipótesis sobre los procesos que intervienen. A continuación (apartado 1.4.2), se aborda el problema de estimación óptima en sistemas con observaciones inciertas, presentando los resultados obtenidos por Jaffer y Gupta [18]. Este estudio pone de manifiesto que el estimador óptimo no es una función lineal de las observaciones y su cálculo requiere un crecimiento exponencial de memoria. Por este motivo, los esfuerzos han de dirigirse hacia la búsqueda de estimadores subóptimos. En el apartado 1.4.3 se analiza el problema de estimación lineal, tanto en el caso de sistemas con incertidumbre en las observaciones modelizada por variables independientes (caso 1), como en el caso en que se suprime dicha hipótesis de independencia (caso 2).

1.4.1 Sistemas con Observaciones Inciertas

Como ya hemos comentado, en muchos problemas prácticos, la observación en un instante puede contener sólo ruido o señal más ruido. Estas situaciones se modelizan mediante un sistema lineal cuya observación no sólo está perturbada por un ruido aditivo, sino que incluye también una componente (ruido) multiplicativa que hace que la señal intervenga en la observación de forma aleatoria. Por tanto, la diferencia entre los sistemas que aquí trataremos y los descritos en la Sección 1.3.1 radica en la ecuación de la observación, ya que el problema planteado sólo afecta a ésta.

Ecuación del Estado

En virtud de los comentarios anteriores, la ecuación del estado está dada por (1.3), es decir

$$\begin{aligned}x(k+1) &= \Phi(k)x(k) + \Gamma(k)w(k), & k \geq 0 \\x(0) &= x_0\end{aligned}\tag{1.5}$$

donde los procesos y matrices que intervienen están definidos en la Sección 1.3.1.

Ecuación de la Observación

En la ecuación de la observación (1.4), al ser la matriz de observación $H(k)$ determinística, el estado $x(k)$ está siempre presente en la observación $z(k)$. Ahora, para reflejar la posibilidad de que la observación en un instante k pueda contener únicamente ruido con probabilidad $1 - p(k)$, definimos la ecuación de la observación de la siguiente forma

$$z(k) = \begin{cases} H(k)x(k) + v(k), & \text{con probabilidad } p(k) \\ v(k), & \text{con probabilidad } 1 - p(k) \end{cases}$$

Si denotamos mediante $u(k)$ una variable aleatoria de Bernoulli con

$$P[u(k) = 1] = p(k),$$

la ecuación de la observación se puede reescribir como

$$z(k) = u(k)H(k)x(k) + v(k), \quad k \geq 0.\tag{1.6}$$

El ruido multiplicativo $\{u(k); k \geq 0\}$ modeliza la incertidumbre de las observaciones. La probabilidad de que la observación $z(k)$ contenga únicamente ruido es $1 - p(k)$ y se denomina *probabilidad de falsa alarma*.

El resto de los procesos y matrices que intervienen en la ecuación (1.6) están definidos en la Sección 1.3.1.

1.4.2 Filtro Óptimo en Sistemas con Observaciones Inciertas

En 1971, Jaffer y Gupta [18] establecieron un algoritmo recursivo para la obtención del estimador óptimo (en el sentido de mínimos cuadrados) en el caso de sistemas con observaciones inciertas cuando la incertidumbre está modelizada por una sucesión markoviana de variables aleatorias. A continuación presentamos los resultados obtenidos por Jaffer y Gupta, poniendo de manifiesto que, en estos sistemas, la esperanza condicionada que proporciona el estimador óptimo, no es una función lineal de las observaciones y su cálculo requiere un crecimiento exponencial de memoria.

Consideremos el sistema lineal en tiempo discreto con observaciones inciertas (1.5) y (1.6) descrito en el apartado anterior. Para la obtención de su algoritmo, Jaffer y Gupta impusieron las siguientes hipótesis

- El estado inicial, x_0 , y los ruidos blancos $\{w(k); k \geq 0\}$, $\{v(k); k \geq 0\}$ son gaussianos y centrados.
- El ruido multiplicativo $\{u(k); k \geq 0\}$ es una sucesión de variables aleatorias de Bernoulli verificando la propiedad de Markov, con probabilidades de transición y distribución inicial conocidas.
- El estado inicial, x_0 , y los ruidos $\{w(k); k \geq 0\}$, $\{v(k); k \geq 0\}$, $\{u(k); k \geq 0\}$ son mutuamente independientes.

Denotemos mediante $f(x(k)/Z_k)$ la función de densidad condicionada del vector $x(k)$ dadas las observaciones $Z_k = (z^T(0), \dots, z^T(k))^T$. El filtro óptimo de mínimos cuadrados viene dado por

$$\hat{x}(k/k) = E[x(k)/Z_k] = \int x(k) f(x(k)/Z_k) dx(k).$$

Como las observaciones $z(0), \dots, z(k)$ son función de $u(0), \dots, u(k)$, tenemos que

$$\hat{x}(k/k) = \int \sum_{U_k \in S^{k+1}} x(k) f(x(k)/Z_k, U_k) P[U_k/Z_k] dx(k), \quad (1.7)$$

donde S representa el espacio de estados de $u(k)$, $S = \{0, 1\}$, S^{k+1} es el producto cartesiano $S \times \dots \times S$ y $U_k = (u(0), \dots, u(k))$ es un elemento de S^{k+1} .

Notando

$$\hat{x}^*(k/k) = E[x(k)/Z_k, U_k] = \int x(k) f(x(k)/Z_k, U_k) dx(k)$$

podemos reescribir (1.7) como sigue

$$\hat{x}(k/k) = \sum_{U_k \in S^{k+1}} \hat{x}^*(k/k) P[U_k/Z_k]. \quad (1.8)$$

Observemos que $\hat{x}^*(k/k)$ es el estimador de $x(k)$ condicionado a un valor particular $U_k \in S^{k+1}$ y, por tanto, puede obtenerse aplicando el Filtro de Kalman.

Para obtener $\hat{x}(k/k)$ necesitamos, también, calcular $P[U_k/Z_k]$. En virtud de la fórmula de Bayes,

$$P[U_k/Z_k] = \frac{f(z(k)/U_k, Z_{k-1}) P[U_k/Z_{k-1}]}{\sum_{U'_k \in S^{k+1}} f(z(k)/U'_k, Z_{k-1}) P[U'_k/Z_{k-1}]}.$$

Teniendo en cuenta las hipótesis del sistema y las ecuaciones del mismo, deducimos que la densidad $f(z(k)/U_k, Z_{k-1})$ es gaussiana con

$$\begin{aligned} E[z(k)/U_k, Z_{k-1}] &= u(k) H(k) \hat{x}^*(k/k-1), \\ Cov[z(k)/U_k, Z_{k-1}] &= u^2(k) H(k) P^*(k/k-1) H^T(k) + R(k). \end{aligned}$$

Por hipótesis, $\{u(k); k \geq 0\}$ es un proceso de Markov, independiente de los procesos $\{w(k); k \geq 0\}$ y $\{v(k); k \geq 0\}$. Por tanto,

$$\begin{aligned} P[U_k/Z_{k-1}] &= P[U_k/U_{k-1}, Z_{k-1}] P[U_{k-1}/Z_{k-1}] \\ &= P[u(k)/u(k-1)] P[U_{k-1}/Z_{k-1}]. \end{aligned}$$

La ecuación anterior permite calcular, de manera recursiva, las probabilidades condicionadas $P[U_k/Z_k]$. El cómputo de $\hat{x}^*(k/k)$ se realiza, como ya se ha comentado,

mediante el Filtro de Kalman. Por último, sustituyendo en (1.8), se obtiene el filtro óptimo, $\hat{x}(k/k)$.

Destaquemos que, debido al término $P[U_k/Z_k]$ que aparece en (1.8), el estimador óptimo no es función lineal de las observaciones y su cómputo requiere un crecimiento exponencial de memoria pues, en el instante k , se necesita haber almacenado 2^{k+1} estimadores $\hat{x}^*(k/k)$, 2^{k+1} matrices de covarianzas $P^*(k/k)$, y 2^{k+1} probabilidades $P[U_k/Z_k]$. Esta dificultad computacional motiva la búsqueda de estimadores subóptimos.

1.4.3 Filtro Lineal en Sistemas con Observaciones Inciertas

Nuestro objetivo en este apartado es el estudio del problema de estimación lineal de mínimos cuadrados en sistemas con observaciones inciertas. Nahi [31] fue el primero en abordar este problema; modelizando la incertidumbre de las observaciones mediante una sucesión de variables aleatorias independientes de Bernoulli y suponiendo que los ruidos del estado y las observaciones eran incorrelados, obtuvo el algoritmo para los problemas de filtrado y predicción lineal de mínimos cuadrados. Dicho algoritmo responde a una estructura recursiva similar al Filtro de Kalman. Monzingo [29] completó los resultados de Nahi estudiando el problema de suavizamiento y Tugnait [35] analizó la estabilidad del estimador. Más tarde, Hermoso y Linares [15], [16] extendieron los resultados de Nahi y Monzingo al caso en que los ruidos aditivos del estado y la observación son correlados.

En algunos problemas prácticos la incertidumbre de las observaciones no puede representarse por variables aleatorias independientes. En estos casos, el problema de estimación lineal del estado ha sido tratado por diversos autores. En 1979, Hadidi y Schwartz [14] estudiaron el problema de filtrado y predicción lineal de menor error cuadrático medio cuando la incertidumbre está modelizada por una sucesión arbitraria

de variables aleatorias de Bernoulli. Probaron que, en general, el estimador lineal de mínimos cuadrados no tiene estructura recursiva y establecieron una condición necesaria y suficiente para su recursividad. Bajo esta condición, la forma del algoritmo resulta ser una generalización de los algoritmos de Kalman y Nahi. En 1984, Wang [38] estableció un algoritmo para los problemas de filtrado y predicción lineal, sustancialmente igual al de Hadidi y Schwartz, pero utilizando un tratamiento distinto. Monzingo [30] analizó el problema de suavizamiento en 1981 y sus resultados, junto con los de Hadidi y Schwartz, dan un tratamiento completo del problema de estimación lineal mínimo cuadrática para estos sistemas.

En este apartado se presenta la solución al problema de estimación lineal de menor error cuadrático medio en sistemas con observaciones inciertas, tanto en el caso en que la incertidumbre está modelizada por variables independientes (caso 1), como en el caso en que se suprime dicha hipótesis de independencia (caso 2).

Caso 1. Incertidumbre Modelizada por Variables Independientes

Consideremos el sistema con observaciones inciertas (1.5) y (1.6) descrito en el apartado 1.4.1. Sobre los ruidos del sistema y la condición inicial se imponen las siguientes hipótesis

- El estado inicial, x_0 , y los ruidos blancos $\{w(k); k \geq 0\}$, $\{v(k); k \geq 0\}$ son centrados.
- El ruido multiplicativo $\{u(k); k \geq 0\}$ es una sucesión de variables aleatorias independientes de Bernoulli, con $P[u(k) = 1] = p(k)$.
- El estado inicial, x_0 , y los ruidos $\{w(k); k \geq 0\}$, $\{v(k); k \geq 0\}$, $\{u(k); k \geq 0\}$ son mutuamente independientes.

En 1969, Nahi [31] estudió el problema de estimación lineal de menor error cuadrático medio para este tipo de sistemas. Concretamente, utilizando la técnica de proyecciones ortogonales, dedujo el siguiente algoritmo recursivo.

Teorema 1.4.1 *El algoritmo para los problemas de predicción y filtrado viene dado por*

$$\begin{aligned}\hat{x}(k/k) &= \hat{x}(k/k-1) + F(k) \tilde{z}(k/k-1), \quad k \geq 0 \\ \hat{x}(k+1/k) &= \Phi(k) \hat{x}(k/k) \\ \hat{x}(0/-1) &= 0\end{aligned}$$

donde $\tilde{z}(k/k-1)$ es el proceso innovación

$$\tilde{z}(k/k-1) = z(k) - p(k) H(k) \hat{x}(k/k-1), \quad k \geq 0.$$

La ganancia $F(k)$ satisface

$$F(k) = p(k) P(k/k-1) H^T(k) \Pi^-(k), \quad k \geq 0^1$$

siendo $\Pi(k)$ la matriz de covarianzas del proceso innovación, dada por

$$\begin{aligned}\Pi(k) &= p(k)(1-p(k)) H(k) D(k) H^T(k) \\ &\quad + p^2(k) H(k) P(k/k-1) H^T(k) + R(k)\end{aligned}$$

Las matrices de covarianzas de los errores de filtrado y predicción verifican

$$\begin{aligned}P(k/k) &= [I - p(k) F(k) H(k)] P(k/k-1), \quad k \geq 0 \\ P(k+1/k) &= \Phi(k) P(k/k) \Phi^T(k) + \Gamma(k) Q(k) \Gamma^T(k) \\ P(0/-1) &= P_0\end{aligned}$$

donde $D(k) = E\{x(k) x^T(k)\}$.

Observemos que este algoritmo generaliza al de Kalman pues, si la probabilidad de falsa alarma, $1 - p(k) = P[u(k) = 0]$, es cero, el sistema bajo consideración es un sistema con certidumbre en las observaciones y el algoritmo de Nahi coincide con el de Kalman.

Caso 2. Incertidumbre Modelizada por Variables Arbitrarias

El objetivo de este apartado es presentar los resultados establecidos por Hadidi y Schwartz [14] sobre el problema de estimación lineal en sistemas con observaciones inciertas cuando la incertidumbre de las observaciones está modelizada por una sucesión arbitraria de variables aleatorias de Bernoulli.

Consideremos el sistema con observaciones inciertas (1.5) y (1.6) descrito en el apartado 1.4.1. Sobre el estado inicial y los ruidos del sistema imponemos las siguientes hipótesis

- El estado inicial, x_0 , y los ruidos blancos $\{w(k); k \geq 0\}$, $\{v(k); k \geq 0\}$ son centrados.
- El ruido multiplicativo $\{u(k); k \geq 0\}$ es una sucesión de variables aleatorias (de Bernoulli, con $P[u(k) = 1] = p(k)$) con probabilidades de transición y distribución inicial conocidas.
- El estado inicial, x_0 , y los ruidos $\{u(k); k \geq 0\}$, $\{v(k); k \geq 0\}$, $\{w(k); k \geq 0\}$ son mutuamente independientes.

Como ya hemos comentado, en 1979 Hadidi y Schwartz demostraron que, en general, bajo estas hipótesis, no existen estimadores recursivos y obtuvieron una condición necesaria y suficiente para su existencia. Concretamente, probaron que el estimador lineal óptimo tiene estructura recursiva si, y sólo si, la probabilidad de transición desde

un instante j a otro instante k es independiente de j , es decir,

$$p_{11}(k/j) = P[u(k) = 1 / u(j) = 1] = p_{11}(k), \quad j = 0, \dots, k-1.$$

Bajo esta condición, Hadidi y Schwartz establecieron el siguiente algoritmo recursivo.

Teorema 1.4.2 *El algoritmo para los problemas de predicción y filtrado está dado por*

$$\begin{aligned} \hat{x}(k/k) &= \hat{x}(k/k-1) + F(k) \tilde{z}(k/k-1), \quad k \geq 0 \\ \hat{x}(k+1/k) &= \Phi(k) \hat{x}(k/k) \\ \hat{x}(0/-1) &= 0 \end{aligned}$$

siendo $\tilde{z}(k/k-1)$ la innovación

$$\tilde{z}(k/k-1) = z(k) - p_{11}(k) H(k) \hat{x}(k/k-1), \quad k \geq 0$$

y $F(k)$ la matriz de ganancia dada por

$$F(k) = [p_{11}(k) P(k/k-1) H^T(k) + (p(k) - p_{11}(k)) D(k) H^T(k)] \Pi^-(k)$$

donde $\Pi(k)$ es la matriz de covarianzas del proceso innovación,

$$\begin{aligned} \Pi(k) &= (p(k) - p_{11}^2(k)) H(k) D(k) H^T(k) \\ &\quad + p_{11}^2(k) H(k) P(k/k-1) H^T(k) + R(k) \end{aligned}$$

Las matrices de covarianzas de los errores de filtrado y predicción verifican

$$\begin{aligned} P(k/k) &= (I - p_{11}(k) F(k) H(k)) P(k/k-1) \\ &\quad - (p(k) - p_{11}(k)) F(k) H(k) D(k) \\ P(k+1/k) &= \Phi(k) P(k/k) \Phi^T(k) + \Gamma(k) Q(k) \Gamma^T(k) \\ P(0/-1) &= P_0 \end{aligned}$$

donde $D(k) = E\{x(k) x^T(k)\}$.

Nota 1.4.1 Para completar los algoritmos dados en los teoremas 1.4.1 y 1.4.2, es necesario dar una fórmula recursiva para la matriz $D(k) = E\{x(k)x^T(k)\}$. Utilizando la ecuación del estado (1.5) y teniendo en cuenta las hipótesis impuestas sobre los procesos que en ella intervienen, se deduce de forma inmediata que

$$\begin{aligned}D(k) &= \Phi(k-1)D(k-1)\Phi^T(k-1) \\ &\quad + \Gamma(k-1)Q(k-1)\Gamma^T(k-1) \\ D(0) &= P_0.\end{aligned}$$

Capítulo 2

Estimación Cuadrática en Sistemas con Observaciones Inciertas

2.1 Introducción

Como se ha indicado en el Capítulo 1, la esperanza condicionada del estado dadas las observaciones proporciona el estimador óptimo (de menor error cuadrático medio) del estado de un sistema lineal y coincide, bajo hipótesis de gaussianidad e independencia mutua de los ruidos y la condición inicial, con el estimador lineal de menor error cuadrático medio proporcionado por el Filtro de Kalman. Si se suprime la hipótesis de gaussianidad, el cálculo de la esperanza condicionada en general no se puede llevar a cabo de forma fácil. Surge, así, la necesidad de encontrar estimadores subóptimos más sencillos desde el punto de vista computacional.

Dentro de la estimación subóptima, el problema de estimación lineal ha sido analizado en el Capítulo 1, tanto en sistemas con certidumbre en las observaciones (Sección

1.3), como en sistemas con observaciones inciertas (Sección 1.4.3). El objetivo de este Capítulo es el estudio del problema de estimación polinomial de segundo grado o estimación cuadrática en sistemas con observaciones inciertas.

En la Sección 2.2 se presenta una interpretación geométrica de dicho problema de estimación. Seguidamente (Sección 2.3), se exponen los principales resultados acerca del problema de estimación polinomial de segundo grado en sistemas con observaciones inciertas cuando la incertidumbre de las observaciones está modelizada por variables aleatorias de Bernoulli independientes, estudio realizado por García-Ligero, Hermoso y Linares [12]. A continuación, en la Sección 2.4, se generalizan estos resultados al caso de sistemas con observaciones inciertas en los que la incertidumbre de las observaciones está modelizada por variables aleatorias no necesariamente independientes. Finalmente, en la Sección 2.5, se considera la versión estacionaria de los sistemas estudiados en la Sección 2.4, con el fin de analizar si el filtro lineal obtenido en dicha sección admite una forma steady-state.

2.2 Estimación Cuadrática y Proyecciones Ortogonales

En esta Sección se presenta un enfoque geométrico del problema de estimación óptima y subóptima de menor error cuadrático medio. Concretamente, el estimador de un vector aleatorio parcialmente observable se identificará con la proyección ortogonal de dicho vector sobre un espacio de Hilbert apropiado.

Consideremos un espacio probabilístico (Ω, A, P) y sea $\mathcal{L}_{\mathbb{R}^n}^2(\Omega, A, P)$ el espacio de Hilbert de clases de equivalencia de vectores aleatorios n -dimensionales con momento de segundo orden finito.

Como se recordaba en el Capítulo 1, dados dos vectores aleatorios, $X \in \mathcal{L}_{\mathbb{R}^n}^2(\Omega, A, P)$ e $Y \in \mathcal{L}_{\mathbb{R}^m}^2(\Omega, A, P)$, el estimador de mínimos cuadrados de X basado en Y es la proyección ortogonal de X sobre el subespacio de Hilbert $\mathcal{L}_{\mathbb{R}^n}^2(\Omega, \sigma(Y), P) (\subset \mathcal{L}_{\mathbb{R}^n}^2(\Omega, A, P))$

de funciones $\sigma(Y)$ -medibles \mathbb{R}^n -valuadas de cuadrado integrable ($\sigma(Y)$ denota la σ -álgebra generada por Y) y coincide con la esperanza condicionada $E[X/Y]$.

En el caso gaussiano (esto es, si X e Y son conjuntamente gaussianos) dicha esperanza condicionada es una función lineal del vector Y . En el caso no gaussiano, la esperanza condicionada, en general, no puede calcularse fácilmente, por lo que el interés se centra en la obtención de estimadores subóptimos que tengan una estructura matemática simple, de forma que permitan el tratamiento de datos reales.

El estimador subóptimo de X basado en Y con estructura más simple es el estimador lineal de menor error cuadrático medio. Teniendo en cuenta los resultados presentados en la Sección 1.2.2 del Capítulo 1, y notando $\mathcal{H}_n(Y_1)$ al subespacio lineal cerrado de $\mathcal{L}_{\mathbb{R}^n}^2(\Omega, \sigma(Y), P)$ de vectores aleatorios de dimensión n obtenidos mediante transformaciones lineales de $Y_1 = (1, Y^T)^T$, el estimador lineal insesgado de menor error cuadrático medio de X basado en Y es la proyección ortogonal del vector X sobre $\mathcal{H}_n(Y_1)$. En lo que sigue, este estimador será notado mediante

$$\hat{X} = \pi(X)|_{\mathcal{H}_n(Y_1)}.$$

Entre el estimador lineal óptimo y la esperanza condicionada, puede definirse otros estimadores mediante proyecciones en subespacios mayores que $\mathcal{H}_n(Y_1)$. Concretamente, en este Capítulo consideraremos proyecciones en el espacio de transformaciones polinómicas de segundo grado (transformaciones cuadráticas) del vector Y .

Con este fin, consideremos el vector aleatorio

$$Y^{[2]} = Y \otimes Y \in \mathbb{R}^{m^2}$$

donde \otimes denota el producto Kronecker (véase Apéndice A). Bajo la hipótesis de que $E\{Y^{[2]T}Y^{[2]}\} < \infty$, el vector $Y_2 = (1, Y^T, Y^{[2]T})^T \in \mathcal{L}_{\mathbb{R}^{m^2+m+1}}^2(\Omega, \sigma(Y), P)$.

El estimador polinomial de segundo grado insesgado de mínima varianza de X basado en Y , que notaremos $\hat{\hat{X}}$, es la proyección ortogonal de X sobre el subespacio

lineal $\mathcal{H}_n(Y_2)$,

$$\widehat{X} = \pi(X)|_{\mathcal{H}_n(Y_2)}.$$

Puesto que $\mathcal{H}_n(Y_1) \subset \mathcal{H}_n(Y_2)$, se sigue que

$$E \left\{ \left\| X - \widehat{X} \right\|^2 \right\} \leq E \left\{ \left\| X - \hat{X} \right\|^2 \right\} \quad (2.1)$$

y, por tanto, el error cometido al estimar el vector X mediante \widehat{X} es siempre menor o igual que el cometido al utilizar el estimador lineal \hat{X} .

Estudiamos la relación entre estos estimadores. Para ello, expresamos el subespacio lineal $\mathcal{H}_n(Y_2)$ como suma directa de $\mathcal{H}_n(Y_1)$ y otro subespacio. Observemos que

$$\mathcal{H}_n(Y_2) = \mathcal{H}_n(1) + \mathcal{H}_n(Y) + \mathcal{H}_n(Y^{[2]}).$$

Si definimos

$$\tilde{Y} = Y - \pi(Y)|_{\mathcal{H}_n(1)} = Y - E\{Y\}$$

y notamos $\tilde{Y}_1 = (1, \tilde{Y}^T)^T$, entonces, puesto que $\mathcal{H}_n(\tilde{Y})$ es ortogonal a $\mathcal{H}_n(1)$, se tiene que

$$\mathcal{H}_n(Y_1) = \mathcal{H}_n(\tilde{Y}_1) = \mathcal{H}_n(1) \oplus \mathcal{H}_n(\tilde{Y})$$

donde \oplus denota la suma directa de subespacios ortogonales.

Análogamente, definimos

$$\tilde{Y}^{[2]} = Y^{[2]} - \pi(Y^{[2]})|_{\mathcal{H}_n(\tilde{Y}_1)}$$

y notando

$$\tilde{Y}_2 = \begin{pmatrix} 1 \\ \tilde{Y} \\ \tilde{Y}^{[2]} \end{pmatrix} = \begin{pmatrix} \tilde{Y}_1 \\ \tilde{Y}^{[2]} \end{pmatrix},$$

podemos escribir

$$\mathcal{H}_n(Y_2) = \mathcal{H}_n(\tilde{Y}_2) = \mathcal{H}_n(\tilde{Y}_1) \oplus \mathcal{H}_n(\tilde{Y}^{[2]})$$

donde se ha utilizado que $\mathcal{H}_n(\tilde{Y}^{[2]})$ es ortogonal a $\mathcal{H}_n(\tilde{Y}_1)$. Por consiguiente, podemos expresar $\hat{\hat{X}}$ como sigue

$$\begin{aligned}\hat{\hat{X}} &= \pi(X)|_{\mathcal{H}_n(Y_2)} = \pi(X)|_{\mathcal{H}_n(\tilde{Y}_1)} + \pi(X)|_{\mathcal{H}_n(\tilde{Y}^{[2]})} \\ &= \hat{X} + \pi(X)|_{\mathcal{H}_n(\tilde{Y}^{[2]})}.\end{aligned}$$

En consecuencia, el estimador polinomial de segundo grado, $\hat{\hat{X}}$, es la suma del estimador lineal óptimo, \hat{X} , y un término ortogonal a \hat{X} que proporciona información adicional a la obtenida mediante transformaciones lineales. A partir de esta expresión del estimador es fácil comprobar que

$$E \left\{ \|X - \hat{X}\|^2 \right\} = E \left\{ \|X - \hat{\hat{X}}\|^2 \right\} + E \left\{ \|\pi(X)|_{\mathcal{H}_n(\tilde{Y}^{[2]})}\|^2 \right\}.$$

Por tanto, siempre que $E \left\{ \|\pi(X)|_{\mathcal{H}_n(\tilde{Y}^{[2]})}\|^2 \right\}$ sea no nula, la desigualdad (2.1) es estricta y el estimador polinomial de segundo grado tiene menor error que el lineal.

2.3 Estimación Cuadrática en Sistemas con Incertidumbre Modelizada por Variables Aleatorias Independientes

Como se ha indicado en el Capítulo 1 (Sección ??), en sistemas gaussianos la esperanza condicionada, que proporciona el estimador óptimo del estado basado en las observaciones, es una función lineal de las mismas. En sistemas con observaciones inciertas, incluso bajo las hipótesis de gaussianidad e independencia mutua de la condición inicial y los ruidos, la distribución conjunta del estado y las observaciones no es gaussiana y, por tanto, el estimador de mínimos cuadrados no es una función lineal de dichas observaciones. Además, su obtención de forma recursiva no es simple, ya que su cómputo requiere un crecimiento exponencial de memoria (Capítulo 1, Sección 1.4.2).

Por este motivo, el estudio del problema de estimación en estos sistemas se ha dirigido fundamentalmente hacia la obtención de estimadores subóptimos.

Dentro de la estimación subóptima, los principales resultados sobre el problema de estimación lineal de menor error cuadrático medio en sistemas con observaciones inciertas han sido resumidos en el Capítulo 1 (Sección 1.4.3), tanto bajo la hipótesis de que la incertidumbre en las observaciones está modelizada por una sucesión de variables aleatorias independientes (Nahi [31]), como por variables de Bernoulli arbitrarias (Hadidi y Schwartz [14]).

El objetivo de esta Sección es presentar los resultados obtenidos por García-Ligero et al. [12] acerca del problema de estimación polinomial de segundo grado de menor error cuadrático medio en sistemas con observaciones inciertas cuando la incertidumbre de las observaciones está modelizada por variables independientes. El procedimiento que utilizaron en su estudio consistió en construir un nuevo sistema (*Sistema Aumentado*) cuyo estado y observación se determinan añadiendo a los correspondientes vectores del sistema original sus potencias de segundo grado definidas mediante el producto Kronecker. De esta forma, el filtro polinomial de segundo grado óptimo del estado del sistema original se obtiene a partir del filtro lineal óptimo del estado del sistema aumentado.

2.3.1 Hipótesis sobre el Sistema

Consideremos el siguiente sistema con observaciones inciertas

$$x(k+1) = \Phi(k)x(k) + w(k), \quad k \geq 0 \quad (2.2)$$

$$x(0) = x_0$$

$$z(k) = u(k)H(k)x(k) + v(k), \quad k \geq 0 \quad (2.3)$$

con vector estado n -dimensional, $x(k)$, y vector observación m -dimensional, $z(k)$.

Sobre el estado inicial x_0 y los ruidos del sistema $\{w(k); k \geq 0\}$, $\{v(k); k \geq 0\}$ y $\{u(k); k \geq 0\}$ se imponen las siguientes hipótesis:

1. x_0 es un vector aleatorio n -dimensional con

$$\begin{aligned} E\{x_0\} &= \bar{x}_0 \\ E\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\} &= P_0 \\ E\{(x_0 - \bar{x}_0)(x_0^{[2]} - E\{x_0^{[2]}\})^T\} &= P_0^{(3)} \\ E\{(x_0^{[2]} - E\{x_0^{[2]}\})(x_0^{[2]} - E\{x_0^{[2]}\})^T\} &= P_0^{(4)} \end{aligned}$$

2. $\{w(k); k \geq 0\}$ es un proceso ruido blanco n -dimensional centrado con

$$\begin{aligned} E\{w(k)w^T(k)\} &= Q(k) \\ E\{w(k)w^{[2]T}(k)\} &= Q^{(3)}(k) \\ E\{w^{[2]}(k)w^{[2]T}(k)\} &= Q^{(4)}(k) \end{aligned}$$

3. $\{v(k); k \geq 0\}$ es un proceso ruido blanco m -dimensional centrado con

$$\begin{aligned} E\{v(k)v^T(k)\} &= R(k) \\ E\{v(k)v^{[2]T}(k)\} &= R^{(3)}(k) \\ E\{v^{[2]}(k)v^{[2]T}(k)\} &= R^{(4)}(k) \end{aligned}$$

4. $\{u(k); k \geq 0\}$ es una sucesión de variables aleatorias independientes de Bernoulli con

$$P[u(k) = 1] = p(k), \quad P[u(k) = 0] = 1 - p(k)$$

5. El estado inicial x_0 y los ruidos $\{w(k); k \geq 0\}$, $\{v(k); k \geq 0\}$, $\{u(k); k \geq 0\}$ son mutuamente independientes.

Bajo las hipótesis anteriores, se plantea el problema de obtener el estimador polinomial de segundo grado óptimo del estado $x(k)$ basado en las observaciones hasta el instante k . Este estimador es una función lineal de $z(0), z^{[2]}(0), \dots, z(k), z^{[2]}(k)$.

Para abordar este problema, se definen los vectores estado y observación aumentados,

$$X_0 = \begin{pmatrix} x_0 \\ x_0^{[2]} \end{pmatrix}, \quad X(k) = \begin{pmatrix} x(k) \\ x^{[2]}(k) \end{pmatrix}, \quad Z(k) = \begin{pmatrix} z(k) \\ z^{[2]}(k) \end{pmatrix}$$

de modo que el vector constituido por las n primeras componentes del estimador lineal de menor error cuadrático medio de $X(k)$ basado en las observaciones $\{Z(0), \dots, Z(k)\}$ proporcionará el estimador polinomial de segundo grado de menor error cuadrático medio del estado, $x(k)$, del sistema original.

2.3.2 Filtro Cuadrático

Analizando la evolución de $X(k)$ y $Z(k)$ se demuestra que los vectores centrados en esperanzas, $\bar{X}(k) = X(k) - E\{X(k)\}$ y $\bar{Z}(k) = Z(k) - E\{Z(k)\}$, satisfacen

$$\bar{X}(k+1) = \bar{\Phi}(k) \bar{X}(k) + \bar{W}(k), \quad k \geq 0 \quad (2.4)$$

$$\bar{X}(0) = X_0 - E\{X_0\} = \bar{X}_0$$

$$\bar{Z}(k) = u(k) \bar{H}(k) \bar{X}(k) + \bar{V}(k), \quad k \geq 0. \quad (2.5)$$

donde

$$\bar{\Phi}(k) = \begin{pmatrix} \Phi(k) & 0 \\ 0 & \Phi^{[2]}(k) \end{pmatrix}, \quad \bar{H}(k) = \begin{pmatrix} H(k) & 0 \\ 0 & H^{[2]}(k) \end{pmatrix}$$

$$\bar{W}(k) = \begin{pmatrix} w(k) \\ F_1(k) - E\{F_1(k)\} \end{pmatrix}$$

$$\bar{V}(k) = \begin{pmatrix} v(k) \\ F_2(k) - E\{F_2(k)\} \end{pmatrix} + (u(k) - p(k)) \bar{H}(k) E\{X(k)\}.$$

Las ecuaciones (2.4) y (2.5) constituyen un nuevo sistema lineal con observaciones inciertas en el que la incertidumbre está modelizada por el mismo ruido multiplicativo, $\{u(k)\}$, del sistema original, y al que se denomina Sistema Aumentado.

García-Ligero et al. [12] probaron que este sistema verifica las condiciones necesarias para aplicar el Filtro de Nahi. Concretamente, el estado inicial y los ruidos del sistema aumentado verifican las siguientes propiedades.

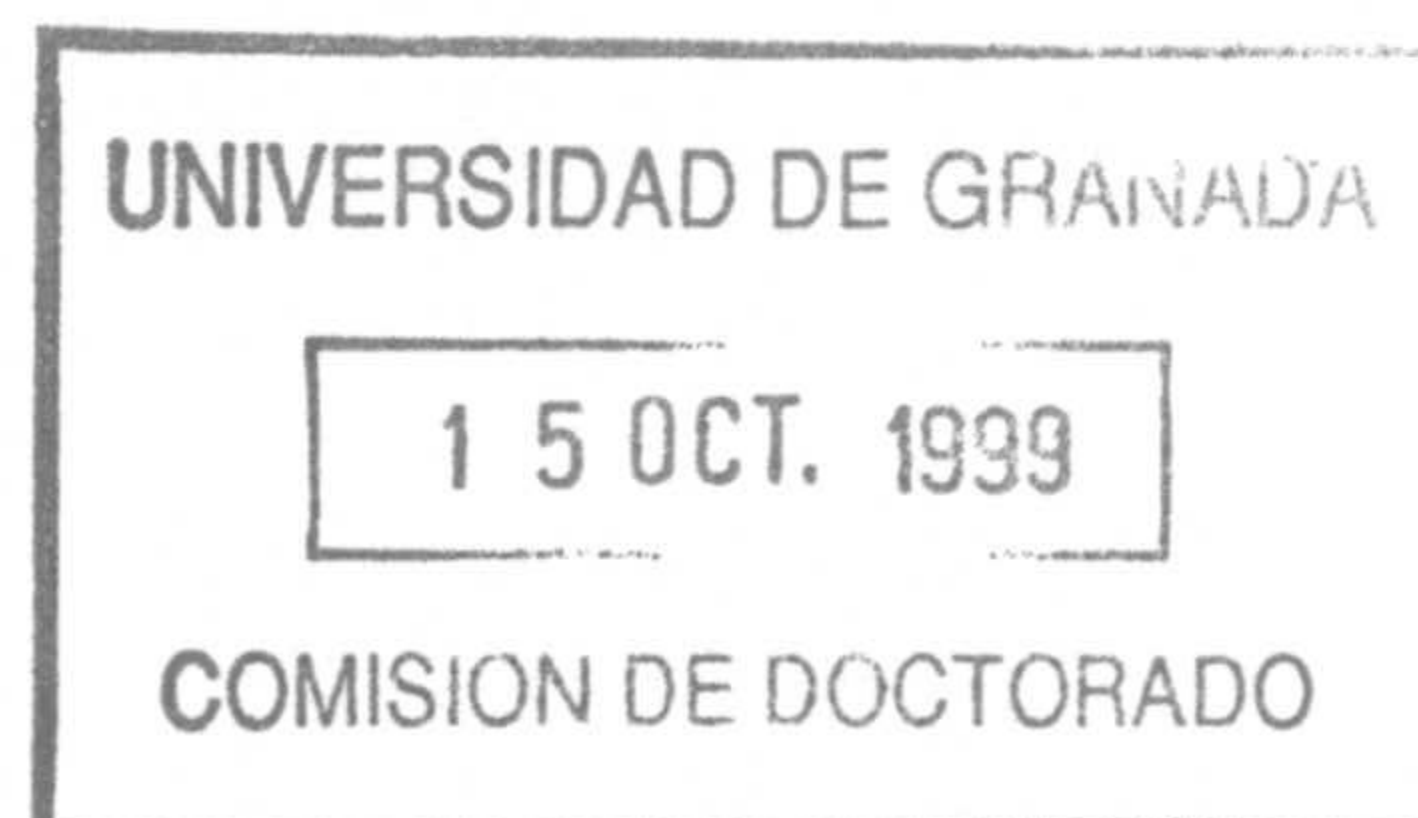
El estado inicial \bar{X}_0 es un vector aleatorio centrado con momento de segundo orden finito.

Los ruidos aditivos $\{\bar{W}(k)\}$ y $\{\bar{V}(k)\}$ son procesos blancos centrados con momentos de segundo orden finitos. Además, dichos ruidos son incorrelados entre sí e incorrelados con el estado inicial. Teniendo en cuenta este resultado, es inmediato que $\bar{W}(k)$ es incorrelado con $\bar{Z}(j)$ para $j \leq k$.

Por último, aplicando la hipótesis de independencia mutua impuesta sobre el sistema original (2.2)-(2.3), se deduce que $u(k)$ y $(\bar{X}(k), \bar{Z}(0), \dots, \bar{Z}(k-1))$ son independientes para todo k .

Como consecuencia de estas propiedades, el estimador lineal de menor error cuadrático medio, $\hat{\bar{X}}(k/k)$, del estado del sistema aumentado, $\bar{X}(k)$, basado en las observaciones $\{\bar{Z}(0), \dots, \bar{Z}(k)\}$, se puede obtener mediante el Filtro de Nahi (Capítulo 1, Sección 1.4.3).

Finalmente, las n primeras componentes del estimador lineal $\hat{\bar{X}}(k/k)$ proporcionan el filtro polinomial de segundo grado de menor error cuadrático medio del estado, $x(k)$, del sistema original.



2.3.3 Filtro Steady-State

Una de las propiedades más interesantes, tanto del Filtro de Kalman, como del Filtro de Nahi, es el comportamiento asintótico de los estimadores y de las matrices de covarianzas de los errores de estimación. Para sistemas estacionarios se puede utilizar la forma *steady-state* del Filtro de Nahi (Tugnait [35]), que presenta grandes ventajas a la hora de calcular los estimadores.

En este apartado, se considera la versión estacionaria del sistema con observaciones inciertas (2.2)-(2.3), descrito en el apartado 2.3.1, esto es

$$\begin{aligned}x(k+1) &= \Phi x(k) + w(k), & k \geq 0 \\x(0) &= x_0 \\z(k) &= u(k) H x(k) + v(k), & k \geq 0\end{aligned}$$

siendo Φ y H matrices reales constantes y $\{w(k)\}$, $\{v(k)\}$ procesos centrados, para los que las matrices definidas en el apartado 2.3.1 son constantes, en cualquier instante de tiempo k ,

$$\begin{aligned}Q(k) &= Q, & Q^{(3)}(k) &= Q^{(3)}, & Q^{(4)}(k) &= Q^{(4)} \\R(k) &= R, & R^{(3)}(k) &= R^{(3)}, & R^{(4)}(k) &= R^{(4)}.\end{aligned}$$

Se supone, además, que el ruido multiplicativo, $\{u(k)\}$, que modeliza la incertidumbre de las observaciones, es una sucesión de variables aleatorias de Bernoulli independientes e idénticamente distribuidas con

$$P[u(k) = 1] = p, \quad \forall k \geq 0.$$

Como en el apartado 2.3.2, se define el sistema aumentado correspondiente al anterior sistema estacionario

$$\bar{X}(k+1) = \bar{\Phi} \bar{X}(k) + \bar{W}(k), \quad k \geq 0$$

$$\begin{aligned}\bar{X}(0) &= \bar{X}_0 \\ \bar{Z}(k) &= u(k)\bar{H}\bar{X}(k) + \bar{V}(k), \quad k \geq 0.\end{aligned}$$

Hay que notar que este sistema aumentado no es un sistema estacionario, puesto que los ruidos aditivos $\{\bar{W}(k)\}$ y $\{\bar{V}(k)\}$ no lo son, ya que sus matrices de covarianzas dependen de la evolución del estado.

Sin embargo, bajo la hipótesis de que el sistema original sea asintóticamente estable, es decir, la matriz Φ sea estrictamente estable (sus autovalores tengan módulo menor que la unidad), García-Ligero et al. [12] establecieron que el sistema aumentado es asintóticamente estacionario. Concretamente, demostraron que las matrices de covarianzas de los procesos ruido $\{\bar{W}(k)\}$, $\{\bar{V}(k)\}$ y del proceso vector estado $\{\bar{X}(k)\}$ del sistema aumentado convergen a matrices constantes y, por tanto, dichos procesos son asintóticamente estacionarios de segundo orden.

Por consiguiente, si la matriz Φ es estrictamente estable, se puede utilizar el filtro steady-state de Nahi para estimar el vector estado del sistema aumentado y, a partir de él, obtener el filtro polinomial de segundo orden steady-state para el estado del sistema estacionario original.

2.4 Estimación Cuadrática en Sistemas con Incertidumbre Modelizada por Variables Aleatorias Arbitrarias

El objetivo de esta Sección es generalizar el estudio realizado por García-Ligero et al. [12] sobre el problema de estimación polinomial de segundo grado en sistemas con observaciones inciertas en los que la incertidumbre de las observaciones está modelizada por variables aleatorias (de Bernoulli) independientes, al caso en que la incertidumbre está modelizada por variables aleatorias arbitrarias.

Seguiremos, para ello, un procedimiento similar al utilizado por dichos autores,

construyendo un nuevo sistema lineal (sistema aumentado), de modo que el estimador polinomial de segundo grado de menor error cuadrático medio del estado del sistema original se deducirá a partir del estimador lineal óptimo del estado de dicho sistema aumentado.

2.4.1 Descripción del Sistema y Planteamiento del Problema

Consideremos el siguiente sistema con observaciones inciertas

$$x(k+1) = \Phi(k)x(k) + w(k), \quad k \geq 0 \quad (2.6)$$

$$x(0) = x_0$$

$$z(k) = u(k)H(k)x(k) + v(k), \quad k \geq 0 \quad (2.7)$$

donde el estado $x(k)$ es n -dimensional, la observación $z(k)$ es m -dimensional.

Sobre los ruidos del sistema y el estado inicial imponemos las siguientes hipótesis:

1. x_0 es un vector aleatorio n -dimensional centrado con

$$E\{x_0\} = 0$$

$$E\{x_0 x_0^T\} = P_0$$

$$E\left\{x_0 \left(x_0^{[2]} - E\{x_0^{[2]}\}\right)^T\right\} = P_0^{(3)}$$

$$E\left\{\left(x_0^{[2]} - E\{x_0^{[2]}\}\right) \left(x_0^{[2]} - E\{x_0^{[2]}\}\right)^T\right\} = P_0^{(4)}.$$

Notaremos

$$\bar{P}_0 = \begin{pmatrix} P_0 & P_0^{(3)} \\ P_0^{(3)T} & P_0^{(4)} \end{pmatrix}.$$

2. $\{w(k); k \geq 0\}$ es un proceso ruido blanco n -dimensional centrado con

$$E\{w(k)w^T(k)\} = Q(k)$$

$$E\{w(k)w^{[2]T}(k)\} = Q^{(3)}(k)$$

$$E\{w^{[2]}(k)w^{[2]T}(k)\} = Q^{(4)}(k).$$

Notaremos

$$Q_w(k) = \begin{pmatrix} Q(k) & Q^{(3)}(k) \\ Q^{(3)T}(k) & \bar{Q}^{(4)}(k) \end{pmatrix}$$

siendo $\bar{Q}^{(4)}(k) = Q^{(4)}(k) - (\text{vec } Q(k))(\text{vec } Q(k))^T$.

3. $\{v(k); k \geq 0\}$ es un proceso ruido blanco m -dimensional centrado con

$$\begin{aligned} E\{v(k)v^T(k)\} &= R(k) \\ E\{v(k)v^{[2]T}(k)\} &= R^{(3)}(k) \\ E\{v^{[2]}(k)v^{[2]T}(k)\} &= R^{(4)}(k). \end{aligned}$$

Notaremos

$$R_v(k) = \begin{pmatrix} R(k) & R^{(3)}(k) \\ R^{(3)T}(k) & \bar{R}^{(4)}(k) \end{pmatrix}$$

siendo $\bar{R}^{(4)}(k) = R^{(4)}(k) - (\text{vec } R(k))(\text{vec } R(k))^T$.

4. $\{u(k); k \geq 0\}$ es una sucesión de variables aleatorias de Bernoulli con vector de probabilidades iniciales $(1 - p_0, p_0)^T$. Las probabilidades de transición serán notadas mediante $p_{mn}(k/j) = P[u(k) = n / u(j) = m]$, para $m, n \in \{0, 1\}$ y $k, j \geq 0$.

Supondremos, además, que la probabilidad $p_{11}(k/j)$ es independiente de j para $j \leq k - 1$. Notaremos $p_{11}(k) = p_{11}(k/j)$ para $j \leq k - 1$.

5. El estado inicial, x_0 , y los ruidos $\{u(k); k \geq 0\}$, $\{w(k); k \geq 0\}$ y $\{v(k); k \geq 0\}$ son mutuamente independientes.

²Puesto que el estimador polinomial de segundo orden de menor error cuadrático medio es una generalización del estimador lineal de menor error cuadrático medio, suponemos que $p_{11}(k/j)$ es independiente de j , condición necesaria y suficiente para la recursividad del estimador lineal (Capítulo 1, Sección 1.4.3).

Como antes se ha indicado, el objetivo de esta Sección es obtener el estimador polinomial de segundo grado de menor error cuadrático medio del estado $x(k)$ basado en las observaciones hasta el instante k , $\{z(0), \dots, z(k)\}$. En virtud de los resultados presentados en la Sección 2.2, este estimador es la proyección ortogonal de $x(k)$ sobre el subespacio $\mathcal{H}_n \left((1, z^T(0), z^{[2]T}(0), \dots, z^T(k), z^{[2]T}(k))^T \right)$ de vectores n -dimensionales obtenidos mediante transformaciones lineales de las observaciones hasta el instante k y de sus potencias de segundo grado.

Con el fin de obtener dicho estimador, definimos los vectores estado y observación aumentados, añadiendo al estado y observación originales sus potencias de segundo grado,

$$X_0 = \begin{pmatrix} x_0 \\ x_0^{[2]} \end{pmatrix}, \quad X(k) = \begin{pmatrix} x(k) \\ x^{[2]}(k) \end{pmatrix}, \quad Z(k) = \begin{pmatrix} z(k) \\ z^{[2]}(k) \end{pmatrix}.$$

El filtro polinomial de segundo grado de menor error cuadrático medio de $x(k)$ se obtendrá a partir del estimador lineal de menor error cuadrático medio de $X(k)$ basado en las observaciones $\{Z(0), \dots, Z(k)\}$.

2.4.2 Sistema Aumentado

Analicemos la evolución de los vectores $X(k)$ y $Z(k)$ con el fin de comprobar si verifican un sistema lineal al que podamos aplicar un algoritmo recursivo para obtener el filtro lineal de menor error cuadrático medio. Para ello, necesitamos estudiar la evolución de $x^{[2]}(k)$ y $z^{[2]}(k)$.

Puesto que para la obtención del sistema aumentado no interviene la hipótesis de independencia de las variables $\{u(k)\}$, se tiene que $X(k)$ y $Z(k)$ verifican el sistema con observaciones inciertas descrito en la Sección 2.3, en el que la incertidumbre está modelizada por la misma sucesión $\{u(k)\}$ del sistema original. Sin embargo, hemos comprobado que el ruido aditivo de la observación de este nuevo sistema no es blanco,

hipótesis requerida para aplicar el Filtro de Hadidi y Schwartz [14]. Para salvar esta dificultad, construiremos un nuevo sistema aumentado, en este caso con certidumbre en las observaciones, y estudiaremos si se verifican las condiciones necesarias para poder aplicar el Filtro de Kalman.

Teniendo en cuenta las ecuaciones del estado (2.6) y de la observación (2.7), y haciendo uso de las propiedades del producto Kronecker (véase Apéndice A), obtenemos que la evolución de $x^{[2]}(k)$ y $z^{[2]}(k)$ está descrita por las siguientes ecuaciones

$$\begin{aligned} x^{[2]}(k+1) &= \Phi^{[2]}(k) x^{[2]}(k) + w^{[2]}(k) + f_1(k), \quad k \geq 0 \\ z^{[2]}(k) &= u(k) H^{[2]}(k) x^{[2]}(k) + v^{[2]}(k) + f_2(k), \quad k \geq 0 \end{aligned}$$

con

$$\begin{aligned} f_1(k) &= (I_{n^2} + K_{n^2}) (\Phi(k) x(k) \otimes w(k)) \\ f_2(k) &= u(k) (I_{m^2} + K_{m^2}) (H(k) x(k) \otimes v(k)) \end{aligned}$$

donde I_p y K_p son la matriz identidad y conmutación de dimensión $p \times p$, respectivamente (véase Apéndice A).

Notando

$$\begin{aligned} \bar{\Phi}(k) &= \begin{pmatrix} \Phi(k) & 0 \\ 0 & \Phi^{[2]}(k) \end{pmatrix}, & \bar{H}(k) &= \begin{pmatrix} H(k) & 0 \\ 0 & H^{[2]}(k) \end{pmatrix} \\ \mathcal{F}_1(k) &= \begin{pmatrix} 0 \\ f_1(k) \end{pmatrix}, & \mathcal{F}_2(k) &= \begin{pmatrix} 0 \\ f_2(k) \end{pmatrix} \\ W(k) &= \begin{pmatrix} w(k) \\ w^{[2]}(k) \end{pmatrix}, & V(k) &= \begin{pmatrix} v(k) \\ v^{[2]}(k) \end{pmatrix} \end{aligned}$$

es fácil comprobar que los vectores $X(k)$ y $Z(k)$ satisfacen las ecuaciones siguientes

$$X(k+1) = \bar{\Phi}(k) X(k) + \bar{W}(k), \quad k \geq 0$$

$$\begin{aligned} X(0) &= X_0 \\ Z(k) &= p(k) \bar{H}(k) X(k) + \tilde{V}(k), \quad k \geq 0 \end{aligned}$$

donde

$$\begin{aligned} \tilde{W}(k) &= W(k) + \mathcal{F}_1(k) \\ \tilde{V}(k) &= V(k) + \mathcal{F}_2(k) + \mathcal{G}_2(k) \end{aligned}$$

con

$$\mathcal{G}_2(k) = (u(k) - p(k)) \bar{H}(k) X(k).$$

Claramente, el sistema anterior es un sistema con certidumbre en las observaciones, ya que la matriz de la observación, $p(k) \bar{H}(k)$, es determinística. También es evidente que los ruidos y el estado inicial de este nuevo sistema no son centrados. No obstante, podemos centrarlos y transformar el sistema anterior en el siguiente *Sistema Aumentado*

$$\bar{X}(k+1) = \bar{\Phi}(k) \bar{X}(k) + \bar{W}(k), \quad k \geq 0 \quad (2.8)$$

$$\bar{X}(0) = \bar{X}_0$$

$$\bar{Z}(k) = p(k) \bar{H}(k) \bar{X}(k) + \bar{V}(k), \quad k \geq 0 \quad (2.9)$$

donde

$$\begin{aligned} \bar{X}(k) &= X(k) - E\{X(k)\}, & \bar{W}(k) &= \tilde{W}(k) - E\{W(k)\} \\ \bar{Z}(k) &= Z(k) - E\{Z(k)\}, & \bar{V}(k) &= \tilde{V}(k) - E\{V(k)\} \end{aligned}$$

Comprobamos, a continuación, si este sistema aumentado (con certidumbre en las observaciones) verifica las hipótesis necesarias para aplicar el Filtro de Kalman. Para ello estudiamos las propiedades de los ruidos $\{\bar{W}(k)\}$, $\{\bar{V}(k)\}$ y del estado inicial \bar{X}_0 .

Claramente, el estado inicial \bar{X}_0 es un vector aleatorio centrado, con momento de segundo orden finito. Teniendo en cuenta la hipótesis 1 realizada sobre el sistema original, su matriz de covarianzas, $E\{\bar{X}_0 \bar{X}_0^T\}$, es

$$\bar{P}_0 = \begin{pmatrix} P_0 & P_0^{(3)} \\ P_0^{(3)T} & P_0^{(4)} \end{pmatrix}.$$

Las propiedades de los procesos ruido $\{\bar{W}(k)\}$ y $\{\bar{V}(k)\}$ se establecen en las siguientes proposiciones.

Proposición 2.4.1 $\{\bar{W}(k); k \geq 0\}$ es un proceso ruido blanco, centrado y con matrices de covarianzas, $\bar{Q}(k) = E\{\bar{W}(k)\bar{W}^T(k)\}$, dadas por

$$\bar{Q}(k) = Q_w(k) + \begin{pmatrix} 0 & 0 \\ 0 & E\{f_1(k)f_1^T(k)\} \end{pmatrix}$$

con

$$E\{f_1(k)f_1^T(k)\} = (I_{n^2} + K_{n^2})(\Phi(k)D(k)\Phi^T(k) \otimes Q(k))(I_{n^2} + K_{n^2})$$

donde $D(k) = E\{x(k)x^T(k)\}$ se puede obtener de forma recursiva mediante las relaciones siguientes

$$D(k+1) = \Phi(k)D(k)\Phi^T(k) + Q(k), \quad k \geq 0$$

$$D(0) = E\{x_0x_0^T\} = P_0.$$

-Demostración-

Observemos que el proceso $\{\bar{W}(k)\}$ tiene media cero, ya que se puede expresar como suma de dos procesos centrados, concretamente,

$$\bar{W}(k) = \begin{pmatrix} w(k) \\ w^{[2]}(k) - E\{w^{[2]}(k)\} \end{pmatrix} + \begin{pmatrix} 0 \\ f_1(k) \end{pmatrix}, \quad k \geq 0$$

donde

$$f_1(k) = (I_{n^2} + K_{n^2})(\Phi(k)x(k) \otimes w(k)).$$

Por otro lado, la hipótesis de que el ruido $\{w(k)\}$ del sistema original es blanco e independiente del estado inicial x_0 , garantiza que ambos procesos son blancos con matrices de covarianzas $Q_w(k)$ y $\begin{pmatrix} 0 & 0 \\ 0 & E\{f_1(k)f_1^T(k)\} \end{pmatrix}$, respectivamente. Por

consiguiente, $\{\bar{W}(k)\}$ es un proceso blanco y sus matrices de covarianzas, $\bar{Q}(k) = E\{\bar{W}(k)\bar{W}^T(k)\}$, son

$$\bar{Q}(k) = Q_w(k) + \begin{pmatrix} 0 & 0 \\ 0 & E\{f_1(k)f_1^T(k)\} \end{pmatrix}$$

con

$$E\{f_1(k)f_1^T(k)\} = (I_{n^2} + K_{n^2})(\Phi(k)D(k)\Phi^T(k) \otimes Q(k))(I_{n^2} + K_{n^2})$$

y, teniendo en cuenta la ecuación del estado (2.6), obtenemos

$$\begin{aligned} D(k+1) &= \Phi(k)D(k)\Phi^T(k) + Q(k), \quad k \geq 0 \\ D(0) &= E\{x_0x_0^T\} = P_0 \end{aligned}$$

donde se ha utilizado, de nuevo, que $\{w(k)\}$ es un ruido blanco independiente del estado inicial x_0 . ■

Proposición 2.4.2 *El proceso ruido $\{\bar{V}(k); k \geq 0\}$ tiene media cero y sus matrices de covarianzas, $\bar{R}(k, j) = E\{\bar{V}(k)\bar{V}^T(j)\}$, están dadas por*

$$\begin{aligned} \bar{R}(k, j) &= R_v(k)\delta_{kj} + \begin{pmatrix} 0 & 0 \\ 0 & E\{f_2(k)f_2^T(k)\} \end{pmatrix} \delta_{kj} \\ &+ p(j)[\delta_{kj} + (1 - \delta_{kj})p_{11}(k) - p(k)]\bar{H}(k)E\{X(k)X^T(j)\}\bar{H}^T(j), \quad j \leq k \end{aligned}$$

siendo

$$E\{f_2(k)f_2^T(k)\} = p(k)(I_{m^2} + K_{m^2})(H(k)D(k)H^T(k) \otimes R(k))(I_{m^2} + K_{m^2}).$$

-Demostración-

Claramente $\{\bar{V}(k)\}$ tiene media cero, ya que se puede expresar como suma de tres procesos centrados, concretamente,

$$\bar{V}(k) = \begin{pmatrix} v(k) \\ v^{[2]}(k) - E\{v^{[2]}(k)\} \end{pmatrix} + \begin{pmatrix} 0 \\ f_2(k) \end{pmatrix} + (u(k) - p(k))\bar{H}(k)X(k), \quad k \geq 0$$

donde

$$f_2(k) = u(k) (I_{m^2} + K_{m^2}) (H(k) x(k) \otimes v(k)).$$

Por otro lado, en virtud de las hipótesis de independencia realizadas sobre el sistema original, estos procesos son incorrelados entre sí. Los dos primeros sumandos son procesos blancos con matrices de covarianzas $R_v(k)$ y $\begin{pmatrix} 0 & 0 \\ 0 & E\{f_2(k) f_2^T(k)\} \end{pmatrix}$, respectivamente. Calculemos las matrices de covarianzas del tercer sumando,

$$\begin{aligned} & E\{(u(k) - p(k)) \bar{H}(k) X(k) X^T(j) \bar{H}^T(j) (u(j) - p(j))\} \\ &= E\{(u(k) - p(k)) (u(j) - p(j))\} \bar{H}(k) E\{X(k) X^T(j)\} \bar{H}^T(j) \\ &= \begin{cases} p(j) (p_{11}(k) - p(k)) \bar{H}(k) E\{X(k) X^T(j)\} \bar{H}^T(j), & j < k \\ p(k) (1 - p(k)) \bar{H}(k) E\{X(k) X^T(k)\} \bar{H}^T(k), & j = k \end{cases} \end{aligned}$$

Por consiguiente, las matrices de covarianzas, $\bar{R}(k, j) = E\{\bar{V}(k) \bar{V}^T(j)\}$, del ruido $\{\bar{V}(k)\}$ vienen dadas por

$$\begin{aligned} \bar{R}(k, j) &= R_v(k) \delta_{kj} + \begin{pmatrix} 0 & 0 \\ 0 & E\{f_2(k) f_2^T(k)\} \end{pmatrix} \delta_{kj} \\ &+ p(j) [\delta_{kj} + (1 - \delta_{kj}) p_{11}(k) - p(k)] \bar{H}(k) E\{X(k) X^T(j)\} \bar{H}^T(j), \quad j \leq k \end{aligned}$$

con

$$E\{f_2(k) f_2^T(k)\} = p(k) (I_{m^2} + K_{m^2}) (H(k) D(k) H^T(k) \otimes R(k)) (I_{m^2} + K_{m^2})$$

donde se ha utilizado, de nuevo, la independencia mutua de los ruidos y la condición inicial. ■

Las hipótesis de independencia realizadas sobre el sistema original y las propiedades del producto Kronecker permiten establecer, también, el siguiente resultado.

Proposición 2.4.3 *Los procesos ruido $\{\bar{W}(k); k \geq 0\}$ y $\{\bar{V}(k); k \geq 0\}$ son incorrelados entre sí e incorrelados con el estado inicial \bar{X}_0 .*

-Demostración-

El resultado es inmediato, teniendo en cuenta que los ruidos $\{w(k)\}$ y $\{v(k)\}$ del sistema original son procesos blancos centrados, y que el estado inicial x_0 y los ruidos $\{w(k)\}$, $\{v(k)\}$, $\{u(k)\}$ son mutuamente independientes. ■

Puesto que $E\{\bar{V}(k)\bar{V}^T(j)\} \neq 0$ para $k \neq j$, no podemos aplicar el filtro de Kalman al sistema aumentado, ya que este filtro requiere que el ruido de la observación sea blanco. Por otro lado, tampoco podemos aplicar los filtros que hemos encontrado en la bibliografía de que disponemos para el caso en que el ruido de la observación, $\{v(k)\}$, es correlado, ya que suponen que dicho ruido verifica una ecuación dinámica del tipo

$$v(k+1) = A(k)v(k) + \beta(k)$$

siendo $\beta(k)$ un ruido blanco, condición que tampoco se satisface en el caso que nos ocupa.

Teniendo en cuenta estas consideraciones, nuestro siguiente propósito es obtener un algoritmo para el problema de filtrado lineal de menor error cuadrático medio del estado del sistema aumentado (2.8)-(2.9).

2.4.3 Algoritmo de Filtrado

Para la obtención del estimador lineal de menor error cuadrático medio del estado del sistema aumentado, nos basamos en la interpretación geométrica de estos estimadores, presentada en la Sección 2.2: si $y(k)$ es un vector aleatorio t -dimensional de cuadrado integrable y $s(0), \dots, s(j)$ son vectores r -dimensionales de cuadrado integrable, el estimador lineal de menor error cuadrático medio del vector $y(k)$ basado en $\{s(0), \dots, s(j)\}$, que notamos $\hat{y}(k/j)$, es la proyección ortogonal de $y(k)$ sobre el subespacio $\mathcal{H}_t \left((1, s^T(0), \dots, s^T(j))^T \right)$ de vectores t -dimensionales obtenidos mediante transformaciones lineales de $s(0), \dots, s(j)$.

Concretamente, utilizando las propiedades (1.1) y (1.2) establecidas en el Capítulo 1 (Sección 1.2.2), obtenemos el siguiente algoritmo para el estimador lineal de menor error cuadrático medio, $\widehat{X}(k/k)$, del estado aumentado $\bar{X}(k)$, basado en las observaciones $\{\bar{Z}(0), \dots, \bar{Z}(k)\}$ (la deducción puede verse en el Apéndice B).

Teorema 2.4.4 *El algoritmo para los problemas de predicción y filtrado viene dado por*

$$\begin{aligned}\widehat{X}(k/k) &= \widehat{X}(k/k-1) + \bar{F}(k) \widetilde{Z}(k/k-1), \quad k \geq 0 \\ \widehat{X}(k+1/k) &= \bar{\Phi}(k) \widehat{X}(k/k), \quad k \geq 0 \\ \widehat{X}(0/-1) &= 0\end{aligned}$$

siendo $\widetilde{Z}(k/k-1)$ la innovación

$$\widetilde{Z}(k/k-1) = \bar{Z}(k) - p(k) \bar{H}(k) \widehat{X}(k/k-1) - \widehat{V}(k/k-1), \quad k \geq 0.$$

Para cada $k > 0$ fijo, el estimador del proceso ruido, $\widehat{V}(k/k-1)$, se calcula mediante la relación

$$\begin{aligned}\widehat{V}(k/l) &= \widehat{V}(k/l-1) + \bar{F}_{\bar{V}}(k,l) \widetilde{Z}(l/l-1), \quad l < k \\ \widehat{V}(k/-1) &= 0.\end{aligned}$$

La ganancia del estimador del proceso ruido, $\bar{F}_{\bar{V}}(k,l)$, viene dada por

$$\bar{F}_{\bar{V}}(k,l) = [\bar{P}_{\bar{V}}(k,l) p(l) \bar{H}^T(l) + \bar{P}_{\bar{V}\bar{V}}(k,l/l-1)] \bar{\Pi}^-(l), \quad l < k$$

siendo

$$\begin{aligned}\bar{P}_{\bar{V}}(k,l) &= \bar{P}_{\bar{V}}(k,l-1) \bar{\Phi}^T(l-1) \\ &\quad - \bar{F}_{\bar{V}}(k,l-1) \bar{\Pi}(l-1) \bar{F}^T(l-1) \bar{\Phi}^T(l-1), \quad 0 < l \leq k \\ \bar{P}_{\bar{V}}(k,0) &= 0\end{aligned}$$

y

$$\begin{aligned}\bar{P}_{\bar{V}\bar{V}}(k, l/j) &= \bar{P}_{\bar{V}\bar{V}}(k, l/j-1) - \bar{F}_{\bar{V}}(k, j) \bar{\Pi}(j) \bar{F}_{\bar{V}}^T(l, j), \quad 0 \leq j < l \leq k \\ \bar{P}_{\bar{V}\bar{V}}(k, l/-1) &= \bar{R}(k, l).\end{aligned}$$

La ganancia del estimador del estado, $\bar{F}(k)$, está dada por

$$\bar{F}(k) = [\bar{P}(k/k-1) p(k) \bar{H}^T(k) + \bar{P}_{\bar{V}}^T(k, k)] \bar{\Pi}^{-1}(k), \quad k \geq 0.$$

Las matrices de covarianzas de los errores de estimación verifican

$$\begin{aligned}\bar{P}(k/k) &= \bar{P}(k/k-1) - \bar{F}(k) \bar{\Pi}(k) \bar{F}^T(k), \quad k \geq 0 \\ \bar{P}(k+1/k) &= \bar{\Phi}(k) \bar{P}(k/k) \bar{\Phi}^T(k) + \bar{Q}(k), \quad k \geq 0 \\ \bar{P}(0/-1) &= \bar{P}_0\end{aligned}$$

siendo $\bar{\Pi}(k)$ la matriz de covarianzas del proceso innovación,

$$\begin{aligned}\bar{\Pi}(k) &= p^2(k) \bar{H}(k) \bar{P}(k/k-1) \bar{H}^T(k) + \bar{P}_{\bar{V}\bar{V}}(k, k/k-1) \\ &\quad + p(k) \bar{H}(k) \bar{P}_{\bar{V}}^T(k, k) + \bar{P}_{\bar{V}}(k, k) p(k) \bar{H}^T(k), \quad k \geq 0.\end{aligned}$$

Nota 2.4.1 Las matrices $\bar{R}(k, j)$, establecidas en la Proposición 2.4.2, dependen de $\mathcal{D}(k, j) = E\{X(k)X^T(j)\}$, por tanto, para poder aplicar el algoritmo, es necesario dar una expresión recursiva para obtener $\mathcal{D}(k, j)$, $0 \leq j \leq k$. Teniendo en cuenta que $\{\bar{W}(k)\}$ es un proceso blanco centrado (Proposición 2.4.1) e incorrelado con el estado inicial \bar{X}_0 (Proposición 2.4.3), se puede comprobar que

$$\begin{aligned}\mathcal{D}(k, j) &= \mathcal{D}(k, j-1) \bar{\Phi}^T(j-1) + \bar{\Phi}(k-1) \cdots \bar{\Phi}(j) \bar{Q}(j-1) \\ &\quad + \bar{\Phi}(k-1) \cdots \bar{\Phi}(0) \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec} P_0) (\text{vec} Q(j-1))^T \end{pmatrix} \\ &\quad + \bar{\Phi}(k-1) \cdots \bar{\Phi}(1) \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec} Q(0)) (\text{vec} Q(j-1))^T \end{pmatrix}\end{aligned}$$

$$\begin{aligned}
& + \dots + \bar{\Phi}(k-1) \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec}Q(k-2))(\text{vec}Q(j-1))^T \end{pmatrix} \\
& + \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec}Q(k-1))(\text{vec}Q(j-1))^T \end{pmatrix}, \quad 0 < j \leq k
\end{aligned}$$

con

$$\begin{aligned}
\mathcal{D}(k,0) &= E\{X(k)X^T(0)\} \\
&= \bar{\Phi}(k)\mathcal{D}(k-1,0) + \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec}Q(k-1))(\text{vec}P_0)^T \end{pmatrix}, \quad k > 0 \\
\mathcal{D}(0,0) &= E\{X_0X_0^T\} = \begin{pmatrix} E\{x_0x_0^T\} & E\{x_0x_0^{[2]T}\} \\ E\{x_0^{[2]}x_0^T\} & E\{x_0^{[2]}x_0^{[2]T}\} \end{pmatrix}.
\end{aligned}$$

Finalmente, la aplicación del Teorema 2.4.4 nos permite obtener el estimador lineal $\hat{X}(k/k)$, y las n primeras componentes del mismo proporcionan el filtro polinomial de segundo grado de menor error cuadrático medio del estado, $x(k)$, del sistema original (2.6)-(2.7).

2.5 Sistemas Estacionarios

En la Sección 2.3.3 se ha considerado la versión estacionaria de los sistemas con incertidumbre en las observaciones modelizada por variables aleatorias independientes, poniendo de manifiesto que, bajo la hipótesis de que el sistema original sea asintóticamente estable, se tiene que el sistema aumentado es asintóticamente estacionario. Esto permitía utilizar la forma steady-state del filtro de Nahi para calcular el estimador lineal del estado aumentado y, a partir de él, obtener el filtro polinomial de segundo grado steady-state del estado del sistema original.

En esta Sección consideramos la versión estacionaria del sistema con observaciones inciertas (2.6)-(2.7) descrito en la Sección 2.4.1. Hay que señalar que, en general, el

correspondiente sistema aumentado no es estacionario. Nuestro objetivo es, en primer lugar (apartado 2.5.1), establecer condiciones bajo las cuales este sistema aumentado sea asintóticamente estacionario (es decir, en el límite, las matrices del sistema sean constantes y los ruidos sean estacionarios de segundo orden) con el fin de analizar (apartado 2.5.2) si el filtro lineal que se ha obtenido en la Sección anterior para dicho sistema aumentado admite una forma steady-state.

2.5.1 Sistema Aumentado Asintóticamente Estacionario

La versión estacionaria del sistema con observaciones inciertas (2.6)-(2.7) viene descrita mediante las siguientes ecuaciones

$$x(k+1) = \Phi x(k) + w(k), \quad k \geq 0 \quad (2.10)$$

$$x(0) = x_0$$

$$z(k) = u(k) H x(k) + v(k), \quad k \geq 0 \quad (2.11)$$

donde Φ y H son matrices determinísticas constantes y $\{w(k)\}$, $\{v(k)\}$ procesos blancos centrados, para los que las matrices definidas en la Sección 2.4.1 son constantes en cualquier instante de tiempo k ,

$$Q(k) = Q, \quad Q^{(3)}(k) = Q^{(3)}, \quad Q^{(4)}(k) = Q^{(4)}$$

$$R(k) = R, \quad R^{(3)}(k) = R^{(3)}, \quad R^{(4)}(k) = R^{(4)}.$$

Además, suponemos que el ruido, $\{u(k)\}$, que modeliza la incertidumbre de las observaciones, es una sucesión de variables aleatorias de Bernoulli con

$$p(k) = P[u(k) = 1] = p, \quad k \geq 0$$

$$p_{11}(k/j) = P[u(k) = 1/u(j) = 1] = p_{11}, \quad 0 \leq j \leq k-1.$$

Construimos el sistema aumentado correspondiente al sistema estacionario anterior

$$\bar{X}(k+1) = \bar{\Phi}\bar{X}(k) + \bar{W}(k), \quad k \geq 0 \quad (2.12)$$

$$\bar{X}(0) = \bar{X}_0$$

$$\bar{Z}(k) = p\bar{H}\bar{X}(k) + \bar{V}(k), \quad k \geq 0 \quad (2.13)$$

donde los vectores y matrices que intervienen están definidos en la Sección 2.4.2. Observemos que las matrices $\bar{\Phi}$ y $p\bar{H}$ no dependen de k , en virtud de las hipótesis impuestas sobre el sistema (2.10)-(2.11).

Hay que notar que este sistema aumentado no es estacionario, puesto que los ruidos aditivos $\{\bar{W}(k)\}$ y $\{\bar{V}(k)\}$ no son procesos estacionarios de segundo orden, ya que sus matrices de covarianzas, $\bar{Q}(k)$ y $\bar{R}(k, j)$, respectivamente, dependen de las matrices $D(k) = E\{x(k)x^T(k)\}$ y de $\mathcal{D}(k, j) = E\{X(k)X^T(j)\}$ (Proposiciones 2.4.1 y 2.4.2) y, por tanto, de la evolución del estado.

No obstante, bajo la hipótesis de que el sistema original sea asintóticamente estable, es decir, la matriz Φ sea estrictamente estable (sus autovalores tengan módulo menor que la unidad), vamos a demostrar que, cuando k tiende a infinito, $D(k)$ converge a una matriz constante (Lema 2.5.1). Aplicando este resultado, se obtendrá que las matrices $\bar{Q}(k)$ tienen límite constante (Teorema 2.5.2) y $\mathcal{D}(k, k-r)$ converge a una matriz que sólo depende del incremento r (Lema 2.5.3). Finalmente, estableceremos que también $\bar{R}(k, k-r)$ converge a una matriz que depende únicamente de r (Teorema 2.5.4). Consecuentemente, los procesos $\{\bar{W}(k)\}$ y $\{\bar{V}(k)\}$ serán asintóticamente estacionarios de segundo orden.

Lema 2.5.1 *Si la matriz Φ es estrictamente estable, entonces*

$$\lim_{k \rightarrow \infty} D(k) = D = \sum_{i=0}^{\infty} \Phi^i Q (\Phi^T)^i$$

-Demostración-

Véase Anderson y Moore [1]. ■

Teorema 2.5.2 *Si la matriz Φ es estrictamente estable, el ruido $\{\bar{W}(k); k \geq 0\}$ del estado del sistema aumentado (2.12)-(2.13) es asintóticamente estacionario de segundo orden, con*

$$\lim_{k \rightarrow \infty} \bar{Q}(k) = \bar{Q} = \begin{pmatrix} Q & Q^{(3)} \\ Q^{(3)T} & \bar{Q}_{22} \end{pmatrix}$$

siendo

$$\bar{Q}_{22} = Q^{(4)} - (\text{vec } Q)(\text{vec } Q)^T + (I_{n^2} + K_{n^2})(\Phi D \Phi^T \otimes Q)(I_{n^2} + K_{n^2}).$$

-Demostración-

La convergencia de las matrices $\bar{Q}(k)$ se obtiene de forma inmediata utilizando el Lema 2.5.1. En consecuencia, el ruido $\{\bar{W}(k)\}$ es estacionario de segundo orden, ya que es centrado y sus matrices de covarianzas tienen por límite \bar{Q} . ■

Lema 2.5.3 *Si la matriz Φ es estrictamente estable entonces, para r fijo, $0 \leq r < k$, se tiene que*

$$\begin{aligned} \lim_{\substack{k \rightarrow \infty \\ 0 \leq r < k}} \mathcal{D}(k, k-r) &= \mathcal{D}(r) \\ &= \bar{\Phi}^r \sum_{i=0}^{\infty} \bar{\Phi}^i \bar{Q} (\bar{\Phi}^T)^i + \left(\sum_{i=0}^{\infty} \bar{\Phi}^i \right) \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec } Q)(\text{vec } Q)^T \end{pmatrix} \sum_{i=0}^{\infty} (\bar{\Phi}^T)^i \end{aligned}$$

-Demostración-

Para r fijo, $0 \leq r < k$, se tiene que (véase Nota 2.4.1)

$$\mathcal{D}(k, k-r) = \mathcal{D}(k, 0) (\bar{\Phi}^T)^{k-r} + \bar{\Phi}^r \sum_{i=0}^{k-r-1} \bar{\Phi}^i \bar{Q} (k-r-i-1) (\bar{\Phi}^T)^i$$

$$\begin{aligned}
& + \bar{\Phi}^k \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec}P_0)(\text{vec}Q)^T \end{pmatrix} \sum_{i=0}^{k-r-1} (\bar{\Phi}^T)^i \\
& + \left(\sum_{i=0}^{k-1} \bar{\Phi}^i \right) \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec}Q)(\text{vec}Q)^T \end{pmatrix} \sum_{i=0}^{k-r-1} (\bar{\Phi}^T)^i.
\end{aligned}$$

Reiterando la expresión de $\mathcal{D}(k, 0)$, y teniendo en cuenta que $\bar{\Phi}$ es estrictamente estable³, es fácil ver que

$$\lim_{k \rightarrow \infty} \mathcal{D}(k, 0) = \left(\sum_{i=0}^{\infty} \bar{\Phi}^i \right) \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec}Q)(\text{vec}P_0)^T \end{pmatrix},$$

de donde se deduce que el primer sumando que aparece en la expresión de $\mathcal{D}(k, k-r)$ converge a cero cuando $k \rightarrow \infty$, puesto que $\lim_{k \rightarrow \infty} (\bar{\Phi}^T)^{k-r} = 0$ por ser $\bar{\Phi}$ estrictamente estable.

El tercer sumando converge también a cero cuando $k \rightarrow \infty$, en virtud de la estabilidad estricta de $\bar{\Phi}$.

Claramente, cuando $k \rightarrow \infty$, el último sumando converge a

$$\left(\sum_{i=0}^{\infty} \bar{\Phi}^i \right) \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec}Q)(\text{vec}Q)^T \end{pmatrix} \sum_{i=0}^{\infty} (\bar{\Phi}^T)^i.$$

En cuanto al segundo sumando, tenemos que

$$\sum_{i=0}^{k-r-1} \bar{\Phi}^i \bar{Q}(k-r-i-1) (\bar{\Phi}^T)^i = \sum_{i=0}^{k-r-1} \bar{\Phi}^{k-r-i-1} \bar{Q}(i) (\bar{\Phi}^T)^{k-r-i-1}.$$

Puesto que $\lim_{i \rightarrow \infty} \bar{Q}(i) = \bar{Q}$ (Teorema 2.5.2), para cada $\varepsilon > 0$, existe i_ε tal que

$$\bar{Q} - \varepsilon I \leq \bar{Q}(i) \leq \bar{Q} + \varepsilon I, \quad i > i_\varepsilon$$

³La estabilidad estricta de Φ implica que $\bar{\Phi}$ también es estrictamente estable (véase Magnus-Neudecker [26])

y, consecuentemente, para $k - r > i_\varepsilon + 1$, tenemos que

$$\begin{aligned} & \sum_{i=0}^{i_\varepsilon} \bar{\Phi}^{k-r-i-1} \bar{Q}(i) (\bar{\Phi}^T)^{k-r-i-1} + \sum_{i=i_\varepsilon+1}^{k-r-1} \bar{\Phi}^{k-r-i-1} (\bar{Q} - \varepsilon I) (\bar{\Phi}^T)^{k-r-i-1} \\ & \leq \sum_{i=0}^{k-r-1} \bar{\Phi}^{k-r-i-1} \bar{Q}(i) (\bar{\Phi}^T)^{k-r-i-1} \\ & \leq \sum_{i=0}^{i_\varepsilon} \bar{\Phi}^{k-r-i-1} \bar{Q}(i) (\bar{\Phi}^T)^{k-r-i-1} + \sum_{i=i_\varepsilon+1}^{k-r-1} \bar{\Phi}^{k-r-i-1} (\bar{Q} + \varepsilon I) (\bar{\Phi}^T)^{k-r-i-1}. \end{aligned}$$

Por ser $\bar{\Phi}$ estrictamente estable, el primer sumando del miembro de la izquierda de esta expresión converge a cero cuando $k \rightarrow \infty$. En cuanto al segundo sumando, se tiene que

$$\begin{aligned} & \sum_{i=i_\varepsilon+1}^{k-r-1} \bar{\Phi}^{k-r-i-1} (\bar{Q} - \varepsilon I) (\bar{\Phi}^T)^{k-r-i-1} \\ & = \sum_{i=0}^{k-r-1} \bar{\Phi}^{k-r-i-1} (\bar{Q} - \varepsilon I) (\bar{\Phi}^T)^{k-r-i-1} - \sum_{i=0}^{i_\varepsilon} \bar{\Phi}^{k-r-i-1} (\bar{Q} - \varepsilon I) (\bar{\Phi}^T)^{k-r-i-1} \\ & = \sum_{i=0}^{k-r-1} \bar{\Phi}^i (\bar{Q} - \varepsilon I) (\bar{\Phi}^T)^i - \sum_{i=0}^{i_\varepsilon} \bar{\Phi}^{k-r-i-1} (\bar{Q} - \varepsilon I) (\bar{\Phi}^T)^{k-r-i-1} \end{aligned}$$

que converge a $\sum_{i=0}^{\infty} \bar{\Phi}^i (\bar{Q} - \varepsilon I) (\bar{\Phi}^T)^i$ cuando $k \rightarrow \infty$.

Razonando de forma análoga para el miembro de la derecha, podemos escribir

$$\sum_{i=0}^{\infty} \bar{\Phi}^i (\bar{Q} - \varepsilon I) (\bar{\Phi}^T)^i \leq \lim_{k \rightarrow \infty} \sum_{i=0}^{k-r-1} \bar{\Phi}^i \bar{Q}(k-r-i-1) (\bar{\Phi}^T)^i \leq \sum_{i=0}^{\infty} \bar{\Phi}^i (\bar{Q} + \varepsilon I) (\bar{\Phi}^T)^i$$

y, haciendo $\varepsilon \rightarrow 0$, concluimos que

$$\lim_{k \rightarrow \infty} \sum_{i=0}^{k-r-1} \bar{\Phi}^i \bar{Q}(k-r-i-1) (\bar{\Phi}^T)^i = \sum_{i=0}^{\infty} \bar{\Phi}^i \bar{Q} (\bar{\Phi}^T)^i.$$

Queda así demostrado que, para cada r fijo, $0 \leq r < k$,

$$\begin{aligned} \mathcal{D}(r) & = \lim_{\substack{k \rightarrow \infty \\ 0 \leq r < k}} \mathcal{D}(k, k-r) \\ & = \bar{\Phi}^r \sum_{i=0}^{\infty} \bar{\Phi}^i \bar{Q} (\bar{\Phi}^T)^i + \left(\sum_{i=0}^{\infty} \bar{\Phi}^i \right) \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec} Q) (\text{vec} Q)^T \end{pmatrix} \sum_{i=0}^{\infty} (\bar{\Phi}^T)^i. \end{aligned}$$

Como consecuencia de estos resultados, podemos establecer el siguiente teorema. ■

Teorema 2.5.4 Si la matriz Φ es estrictamente estable, el ruido $\{\bar{V}(k); k \geq 0\}$ de la observación del sistema aumentado (2.12)-(2.13) es asintóticamente estacionario de segundo orden, con

$$\lim_{\substack{k \rightarrow \infty \\ 0 \leq r < k}} \bar{R}(k, k-r) = \bar{R}(r) = \begin{pmatrix} R & R^{(3)} \\ R^{(3)T} & \bar{R}_{22} \end{pmatrix} \delta_{0,r+p} [\delta_{0,r} + (1 - \delta_{0,r}) p_{11} - p] \bar{H} \mathcal{D}(r) \bar{H}^T$$

siendo

$$\bar{R}_{22} = R^{(4)} - (\text{vec } R)(\text{vec } R)^T + p(I_{m^2} + K_{m^2})(HDH^T \otimes R)(I_{m^2} + K_{m^2}).$$

-Demostración-

La convergencia de las matrices $\bar{R}(k, k-r)$ se obtiene de forma inmediata aplicando los lemas 2.5.1 y 2.5.3. En consecuencia, el ruido $\{\bar{V}(k)\}$ es asintóticamente estacionario de segundo orden, ya que es un proceso centrado y sus matrices de covarianzas tienen límite $\bar{R}(r)$. ■

Para completar el estudio asintótico del sistema aumentado, veamos finalmente que el proceso estado $\{\bar{X}(k)\}$ de dicho sistema aumentado es también asintóticamente estacionario de segundo orden.

Teorema 2.5.5 Si la matriz Φ es estrictamente estable, entonces el proceso estado $\{\bar{X}(k); k \geq 0\}$ del sistema aumentado (2.12) y (2.13) es asintóticamente estacionario de segundo orden, con

$$\lim_{\substack{k \rightarrow \infty \\ 0 \leq r < k}} E \{ \bar{X}(k) \bar{X}^T(k-r) \} = \bar{\Phi}^r \sum_{i=0}^{\infty} \bar{\Phi}^i \bar{Q} (\bar{\Phi}^T)^i.$$

-Demostración-

Basta observar que $\{\bar{X}(k)\}$ es un proceso centrado y que su matriz de covarianzas está dada por

$$E\{\bar{X}(k)\bar{X}^T(k-r)\} = \mathcal{D}(k, k-r) - E\{X(k)\}E\{X^T(k-r)\}$$

para todo r fijo, $0 \leq r < k$.

En virtud del Lema 2.5.3, $\mathcal{D}(k, k-r)$ tiende a $\mathcal{D}(r)$ cuando $k \rightarrow \infty$. En cuanto al segundo sumando, se tiene que

$$\begin{aligned} E\{X(k)\} &= \bar{\Phi}E\{X(k-1)\} + \begin{pmatrix} 0 \\ E\{w^{[2]}(k-1)\} \end{pmatrix} \\ &= \bar{\Phi}E\{X(k-1)\} + \begin{pmatrix} 0 \\ \text{vec}Q \end{pmatrix}. \end{aligned}$$

Reiterando esta expresión, se llega a que

$$\begin{aligned} E\{X(k)\} &= \bar{\Phi}^k E\{X(0)\} + \sum_{i=0}^{k-1} \bar{\Phi}^i \begin{pmatrix} 0 \\ \text{vec}Q \end{pmatrix} \\ &= \bar{\Phi}^k \begin{pmatrix} 0 \\ \text{vec}P_0 \end{pmatrix} + \sum_{i=0}^{k-1} \bar{\Phi}^i \begin{pmatrix} 0 \\ \text{vec}Q \end{pmatrix} \end{aligned}$$

y, puesto que $\bar{\Phi}$ es estrictamente estable, tenemos que

$$\lim_{k \rightarrow \infty} E\{X(k)\} = \lim_{\substack{k \rightarrow \infty \\ 0 \leq r < k}} E\{X(k-r)\} = \sum_{i=0}^{\infty} \bar{\Phi}^i \begin{pmatrix} 0 \\ \text{vec}Q \end{pmatrix}.$$

Concluimos, así, que

$$\begin{aligned} \lim_{\substack{k \rightarrow \infty \\ 0 \leq r < k}} E\{\bar{X}(k)\bar{X}^T(k-r)\} &= \mathcal{D}(r) - \left(\sum_{i=0}^{\infty} \bar{\Phi}^i \right) \begin{pmatrix} 0 & 0 \\ 0 & (\text{vec}Q)(\text{vec}Q)^T \end{pmatrix} \sum_{i=0}^{\infty} (\bar{\Phi}^T)^i \\ &= \bar{\Phi}^r \sum_{i=0}^{\infty} \bar{\Phi}^i \bar{Q} (\bar{\Phi}^T)^i \end{aligned}$$

es decir que, en el límite, la matriz de covarianzas $E \{ \bar{X}(k) \bar{X}^T(k-r) \}$ del estado aumentado depende únicamente del incremento r y, en consecuencia, el proceso $\{ \bar{X}(k) \}$ es asintóticamente estacionario de segundo orden. ■

2.5.2 Existencia del Filtro Steady-State

En este apartado consideramos el sistema estacionario con observaciones inciertas (2.10)-(2.11); nuestro objetivo es establecer condiciones bajo las cuales el algoritmo de filtrado que se ha obtenido en la Sección 2.4.3, para el sistema aumentado correspondiente a dicho sistema estacionario, admite una forma steady-state. En caso de existir, esta forma steady-state del filtro podría utilizarse para obtener el estimador lineal de menor error cuadrático medio del estado del sistema aumentado y, a partir de él, el filtro polinomial de segundo grado steady-state para el estado del sistema estacionario original, con las ventajas computacionales que ello conlleva.

Para sistemas invariantes en el tiempo, esto es, para sistemas con matrices de transición y observación constantes, una condición necesaria para la existencia del filtro steady-state es la estacionariedad de los procesos asociados al sistema.

Consideremos el sistema aumentado (2.12)-(2.13)

$$\begin{aligned} \bar{X}(k+1) &= \bar{\Phi} \bar{X}(k) + \bar{W}(k), \quad k \geq 0 \\ \bar{X}(0) &= \bar{X}_0 \\ \bar{Z}(k) &= p\bar{H} \bar{X}(k) + \bar{V}(k), \quad k \geq 0. \end{aligned}$$

Claramente, este sistema es invariante en el tiempo, ya que las matrices $\bar{\Phi}$ y $p\bar{H}$ son constantes, pero, como se comentó en el apartado anterior, los ruidos aditivos $\{ \bar{W}(k) \}$ y $\{ \bar{V}(k) \}$ no son estacionarios, ya que sus matrices de covarianzas, $\bar{Q}(k)$ y $\bar{R}(k, j)$, respectivamente, dependen de la evolución del estado. Sin embargo, bajo la hipótesis de que el sistema original sea asintóticamente estable, esto es, la matriz Φ sea estrictamente estable, los procesos $\{ \bar{W}(k) \}$ y $\{ \bar{V}(k) \}$ son asintóticamente estacionarios de

segundo orden. Teniendo en cuenta los resultados obtenidos en dicho apartado sobre el comportamiento asintótico de los ruidos del sistema, para el estudio de la existencia del filtro steady-state, vamos a suponer las siguientes condiciones sobre el estado inicial y los ruidos:

1. $\{\bar{W}(k); k \geq 0\}$ es un proceso ruido blanco centrado con $E\{\bar{W}(k)\bar{W}^T(k)\} = \bar{Q}$, para todo $k \geq 0$.
2. $\{\bar{V}(k); k \geq 0\}$ es un proceso ruido centrado con $E\{\bar{V}(k)\bar{V}^T(j)\} = \bar{R}(k-j)$, para $0 \leq j \leq k$.
3. El estado inicial \bar{X}_0 es un vector aleatorio centrado con $E\{\bar{X}_0\bar{X}_0^T\} = \bar{P}_0$.
4. $\{\bar{W}(k); k \geq 0\}$ y $\{\bar{V}(k); k \geq 0\}$ son incorrelados entre sí e incorrelados con el estado inicial \bar{X}_0 .

Bajo estas condiciones, el algoritmo de predicción y filtrado obtenido en la Sección 2.4.3 (Teorema 2.4.4) se puede reescribir de la forma siguiente

$$\begin{aligned} \hat{X}(k/k) &= \hat{X}(k/k-1) + \bar{F}(k)\tilde{Z}(k/k-1), & k \geq 0 \\ \hat{X}(k+1/k) &= \bar{\Phi}\hat{X}(k/k), & k \geq 0 \\ \hat{X}(0/-1) &= 0 \end{aligned}$$

$$\tilde{Z}(k/k-1) = \bar{Z}(k) - p\bar{H}\hat{X}(k/k-1) - \hat{V}(k/k-1), \quad k \geq 0$$

$$\begin{aligned} \hat{V}(k/l) &= \hat{V}(k/l-1) + \bar{F}_{\bar{V}}(k,l)\tilde{Z}(l/l-1), & 0 \leq l < k \\ \hat{V}(k/-1) &= 0 \end{aligned}$$

$$\bar{F}_{\bar{V}}(k,l) = [\bar{P}_{\bar{V}}(k,l)p\bar{H}^T + \bar{P}_{\bar{V}\bar{V}}(k,l/l-1)]\bar{\Pi}^-(l), \quad l < k$$

$$\bar{P}_{\bar{V}}(k, l) = \bar{P}_{\bar{V}}(k, l-1) \bar{\Phi}^T - \bar{F}_{\bar{V}}(k, l-1) \bar{\Pi}(l-1) \bar{F}^T(l-1) \bar{\Phi}^T, \quad 0 < l \leq k$$

$$\bar{P}_{\bar{V}}(k, 0) = 0$$

$$\bar{P}_{\bar{V}\bar{V}}(k, l/j) = \bar{P}_{\bar{V}\bar{V}}(k, l/j-1) - \bar{F}_{\bar{V}}(k, j) \bar{\Pi}(j) \bar{F}_{\bar{V}}^T(l, j), \quad 0 \leq j < l \leq k$$

$$\bar{P}_{\bar{V}\bar{V}}(k, l/-1) = \bar{R}(k-l)$$

$$\boxed{\bar{F}(k) = [\bar{P}(k/k-1) p \bar{H}^T + \bar{P}_{\bar{V}}^T(k, k)] \bar{\Pi}^-(k), \quad k \geq 0}$$

$$\bar{P}(k/k) = \bar{P}(k/k-1) - \bar{F}(k) \bar{\Pi}(k) \bar{F}^T(k), \quad k \geq 0 \quad (2.14)$$

$$\bar{P}(k+1/k) = \bar{\Phi} \bar{P}(k/k) \bar{\Phi}^T + \bar{Q}, \quad k \geq 0 \quad (2.15)$$

$$\bar{P}(0/-1) = \bar{P}_0$$

$$\bar{\Pi}(k) = p^2 \bar{H} \bar{P}(k/k-1) \bar{H}^T + \bar{P}_{\bar{V}\bar{V}}(k, k/k-1) + p \bar{H} \bar{P}_{\bar{V}}^T(k, k) + \bar{P}_{\bar{V}}(k, k) p \bar{H}^T, \quad k \geq 0$$

Como hemos indicado, el objetivo de esta Sección es establecer condiciones bajo las cuales este algoritmo admite una forma steady-state, en otras palabras, estamos interesados en encontrar condiciones que garanticen que las matrices de ganancia del filtro, $\{\bar{F}(k)\}$, convergen a un límite finito (un valor steady-state), cuando $k \rightarrow \infty$.

Notemos asimismo que, puesto que en cada iteración del algoritmo debemos obtener el predictor del ruido de la observación, $\hat{V}(k/k-1)$, es necesario también estudiar el comportamiento asintótico de las matrices de ganancia, $\bar{F}_{\bar{V}}(k, l)$, asociadas al mismo.

Límite de las Ganancias del Filtro del Estado

La existencia de $\lim_{k \rightarrow \infty} \bar{F}(k)$ va a venir determinada por la convergencia de la sucesión de matrices de covarianzas de los errores de predicción. Teniendo en cuenta las expresiones de las matrices de covarianzas de los errores de filtrado y predicción, (2.14) y (2.15), respectivamente, podemos escribir

$$\bar{P}(k+1/k) = \bar{\Phi} [\bar{P}(k/k-1) - \bar{F}(k) \bar{\Pi}(k) \bar{F}^T(k)] \bar{\Phi}^T + \bar{Q}$$

$$\bar{P}(0/-1) = \bar{P}_0.$$

Por tanto, si demostramos que las matrices de covarianzas de los errores de predicción $\{\bar{P}(k+1/k)\}$ tienen límite finito, podremos afirmar que el producto $\bar{F}(k) \bar{\Pi}(k) \bar{F}^T(k)$ converge cuando $k \rightarrow \infty$. Con este resultado, estableciendo que $\lim_{k \rightarrow \infty} \bar{\Pi}(k) = \bar{\Pi}$, concluiremos que también la sucesión $\{\bar{F}(k)\}$ tiene límite finito cuando $k \rightarrow \infty$.

Comenzamos demostrando que, si la condición inicial $\bar{P}(0/-1)$ es nula, entonces la sucesión de matrices de covarianzas $\{\bar{P}(k+1/k)\}$ es monótona no decreciente y acotada (Lemas 2.5.6 y 2.5.7); en consecuencia, dicha sucesión tendrá límite finito (Teorema 2.5.8).

Lema 2.5.6 *Si la covarianza inicial $\bar{P}(0/-1)$ es nula, entonces la sucesión de matrices de covarianzas $\{\bar{P}(k+1/k)\}$ es monótona no decreciente.*

-Demostración-

Comencemos estableciendo una expresión alternativa para las matrices $\bar{P}(k+1/k)$ que será básica para demostrar su monotonía. Teniendo en cuenta la expresión de $\bar{F}(k)$, claramente

$$\begin{aligned} \bar{F}(k) \bar{\Pi}(k) \bar{F}^T(k) &= p\bar{F}(k) \bar{H} \bar{P}(k/k-1) + \bar{F}(k) \bar{P}_{\bar{V}}(k, k) \\ &= p\bar{P}(k/k-1) \bar{H}^T \bar{F}^T(k) + \bar{P}_{\bar{V}}^T(k, k) \bar{F}^T(k). \end{aligned} \quad (2.16)$$

Sustituyendo (2.14) en (2.15), y utilizando esta expresión se deduce que

$$\begin{aligned} \bar{P}(k+1/k) &= \bar{\Phi} \left[\bar{P}(k/k-1) - p\bar{F}(k) \bar{H} \bar{P}(k/k-1) - \bar{F}(k) \bar{P}_{\bar{V}}(k, k) \right] \bar{\Phi}^T + \bar{Q} \\ &= [\bar{\Phi} - p\bar{\Phi} \bar{F}(k) \bar{H}] \bar{P}(k/k-1) [\bar{\Phi} - p\bar{\Phi} \bar{F}(k) \bar{H}]^T \\ &\quad + \bar{\Phi} \left[p\bar{P}(k/k-1) \bar{H}^T \bar{F}^T(k) - p^2 \bar{F}(k) \bar{H} \bar{P}(k/k-1) \bar{H}^T \bar{F}^T(k) \right. \\ &\quad \left. - \bar{F}(k) \bar{P}_{\bar{V}}(k, k) \right] \bar{\Phi}^T + \bar{Q}. \end{aligned}$$

Haciendo, de nuevo, uso de (2.16), se obtiene

$$\bar{P}(k+1/k) = [\bar{\Phi} - p\bar{\Phi} \bar{F}(k) \bar{H}] \bar{P}(k/k-1) [\bar{\Phi} - p\bar{\Phi} \bar{F}(k) \bar{H}]^T$$

$$\begin{aligned}
& +\bar{\Phi} \left[\bar{F}(k) \bar{\Pi}(k) \bar{F}^T(k) - p^2 \bar{F}(k) \bar{H} \bar{P}(k/k-1) \bar{H}^T \bar{F}^T(k) \right. \\
& \left. - \bar{P}_{\bar{V}}^T(k, k) \bar{F}^T(k) - \bar{F}(k) \bar{P}_{\bar{V}}(k, k) \right] \bar{\Phi}^T + \bar{Q}
\end{aligned}$$

y, teniendo en cuenta la expresión de $\bar{\Pi}(k)$, concluimos

$$\begin{aligned}
\bar{P}(k+1/k) &= [\bar{\Phi} - p\bar{\Phi}\bar{F}(k)\bar{H}] \bar{P}(k/k-1) [\bar{\Phi} - p\bar{\Phi}\bar{F}(k)\bar{H}]^T \\
& + \bar{\Phi}\bar{F}(k) [p\bar{H}\bar{P}_{\bar{V}}^T(k, k) + \bar{P}_{\bar{V}}(k, k)p\bar{H}^T + \bar{P}_{\bar{V}\bar{V}}(k, k/k-1)] \bar{F}^T(k) \bar{\Phi}^T \\
& - \bar{\Phi} [\bar{P}_{\bar{V}}^T(k, k) \bar{F}^T(k) + \bar{F}(k) \bar{P}_{\bar{V}}(k, k)] \bar{\Phi}^T + \bar{Q}. \tag{2.17}
\end{aligned}$$

Para demostrar la monotonía de $\{\bar{P}(k+1/k)\}$ consideremos una nueva sucesión $\{\bar{P}^*(k+1/k)\}$, con condición inicial $\bar{P}^*(0/-1) = \bar{P}(0/-1) = 0$, obtenida tomando

$$\begin{aligned}
\bar{F}^*(k) &= \bar{F}(k+1), \quad k \geq 0 \\
\bar{P}_{\bar{V}}^*(k, k) &= \bar{P}_{\bar{V}}(k+1, k+1), \quad k \geq 0 \\
\bar{P}_{\bar{V}\bar{V}}^*(k, k/k-1) &= \bar{P}_{\bar{V}\bar{V}}(k+1, k+1/k), \quad k \geq 0.
\end{aligned}$$

Veamos que $\bar{P}^*(k/k-1) \leq \bar{P}(k+1/k)$ para todo $k \geq 0$. Razonamos por inducción.

- Para $k=0$, es claro que $\bar{P}^*(0/-1) = 0 \leq \bar{P}(1/0)$, ya que $\bar{P}(1/0)$ es una matriz de covarianzas y, por tanto, definida no negativa.
- Supongamos que $\bar{P}^*(i/i-1) \leq \bar{P}(i+1/i)$ para todo $i < k$.
- Veamos que es cierto para k . Razonando de forma análoga a como se ha hecho para obtener (2.17), y teniendo en cuenta las expresiones de $\bar{F}^*(k)$, $\bar{P}_{\bar{V}}^*(k, k)$, $\bar{P}_{\bar{V}\bar{V}}^*(k, k/k-1)$, se obtiene la siguiente expresión para la matriz $\bar{P}^*(k/k-1)$

$$\begin{aligned}
\bar{P}^*(k/k-1) &= [\bar{\Phi} - \bar{\Phi}\bar{F}(k)p\bar{H}] \bar{P}^*(k-1/k-2) [\bar{\Phi} - \bar{\Phi}\bar{F}(k)p\bar{H}]^T \\
& + \bar{\Phi}\bar{F}(k) [p\bar{H}\bar{P}_{\bar{V}}^T(k, k) + \bar{P}_{\bar{V}}(k, k)p\bar{H}^T + \bar{P}_{\bar{V}\bar{V}}(k, k/k-1)] \bar{F}^T(k) \bar{\Phi}^T \\
& - \bar{\Phi} [\bar{P}_{\bar{V}}^T(k, k) \bar{F}^T(k) + \bar{F}(k) \bar{P}_{\bar{V}}(k, k)] \bar{\Phi}^T + \bar{Q}
\end{aligned}$$

de donde concluimos que la diferencia

$$\begin{aligned} \bar{P}(k+1/k) - \bar{P}^*(k/k-1) \\ = [\bar{\Phi} - \bar{\Phi}\bar{F}(k)p\bar{H}] (\bar{P}(k/k-1) - \bar{P}^*(k-1/k-2)) [\bar{\Phi} - \bar{\Phi}\bar{F}(k)p\bar{H}]^T \end{aligned}$$

es definida no negativa, en virtud de la hipótesis de inducción.

Por consiguiente, $\bar{P}^*(k-1/k-2) \leq \bar{P}(k/k-1)$ para todo k . Además, por la optimalidad de $\bar{P}(k+1/k)$, se verifica que $\bar{P}(k/k-1) \leq \bar{P}^*(k/k-1)$, para todo k y, por tanto,

$$0 \leq \bar{P}(k/k-1) \leq \bar{P}^*(k/k-1) \leq \bar{P}(k+1/k).$$

Luego, la sucesión $\{\bar{P}(k+1/k)\}$, con condición inicial $\bar{P}(0/-1) = 0$, es no decreciente. ■

Lema 2.5.7 *Si la matriz $\bar{\Phi}$ es estrictamente estable, la sucesión $\{\bar{P}(k+1/k)\}$ está acotada.*

-Demostración-

Definimos el siguiente estimador subóptimo

$$\hat{\bar{X}}_e(k/k-1) = 0, \quad k \geq 0.$$

Claramente, el error asociado coincide con el estado del sistema

$$\tilde{\bar{X}}_e(k/k-1) = \bar{X}(k), \quad k \geq 0,$$

con lo que su matriz de covarianzas, $\bar{P}_e(k+1/k) = E\{\bar{X}(k+1)\bar{X}^T(k+1)\}$, tiene límite constante al ser $\bar{\Phi}$ estrictamente estable (Teorema 2.5.5), concretamente,

$$\lim_{k \rightarrow \infty} \bar{P}_e(k+1/k) = \sum_{i=0}^{\infty} \bar{\Phi}^i \bar{Q} (\bar{\Phi}^T)^i,$$

en particular, la sucesión $\{\bar{P}_e(k+1/k)\}$ está acotada.

Por otro lado, debido a la suboptimalidad del estimador $\hat{X}_e(k/k-1)$, tenemos que

$$\bar{P}(k+1/k) \leq \bar{P}_e(k+1/k), \quad k \geq 0$$

y, por consiguiente, también $\{\bar{P}(k+1/k)\}$ está acotada. ■

Como consecuencia directa de los Lemas 2.5.6 y 2.5.7, podemos establecer el siguiente resultado.

Teorema 2.5.8 *Si la matriz Φ es estrictamente estable y $\bar{P}(0/-1) = 0$, entonces la sucesión $\{\bar{P}(k+1/k)\}$ tiene límite finito, \bar{P} , cuando $k \rightarrow \infty$.*

Como antes indicamos, el Teorema 2.5.8, nos permite afirmar que el producto $\bar{F}(k) \bar{\Pi}(k) \bar{F}^T(k)$ converge cuando $k \rightarrow \infty$. Si demostramos ahora que $\lim_{k \rightarrow \infty} \bar{\Pi}(k) = \bar{\Pi}$, concluiremos que también la sucesión $\{\bar{F}(k)\}$ tiene límite finito cuando $k \rightarrow \infty$.

Lema 2.5.9 *Si la matriz Φ es estrictamente estable, la sucesión $\{\bar{\Pi}(k)\}$ está acotada y es monótona no decreciente.*

-Demostración-

Para demostrar la acotación, consideremos los estimadores subóptimos

$$\begin{aligned} \hat{X}_e(k/k-1) &= 0, \quad k \geq 0 \\ \hat{V}_e(k/k-1) &= 0, \quad k \geq 0. \end{aligned}$$

El proceso innovación viene dado por

$$\begin{aligned} \tilde{Z}_e(k+1/k) &= \bar{Z}(k+1) - p\bar{H}\hat{X}_e(k+1/k) - \hat{V}_e(k+1/k) \\ &= \bar{Z}(k+1). \end{aligned}$$

Por tanto, sus matrices de covarianzas, $\bar{\Pi}_e(k) = E\{\bar{Z}(k)\bar{Z}^T(k)\}$, satisfacen

$$\bar{\Pi}_e(k) = p\bar{H}E\{\bar{X}(k)\bar{X}^T(k)\}\bar{H}^T + \bar{R}(0)$$

y, al estar $E \{ \bar{X}(k) \bar{X}^T(k) \}$ acotada (Teorema 2.5.5), la sucesión $\{ \bar{\Pi}_e(k) \}$ también está acotada. Por otro lado, en virtud de la suboptimalidad del estimador $\widehat{\bar{X}}_e(k/k-1)$, tenemos que

$$\bar{\Pi}(k) \leq \bar{\Pi}_e(k), \quad k \geq 0$$

y, consecuentemente, la sucesión $\{ \bar{\Pi}(k) \}$ está acotada.

Para demostrar que $\{ \bar{\Pi}(k) \}$ es monótona no decreciente, consideremos las matrices de covarianzas del proceso innovación asociado a las sucesiones $\{ \bar{P}^*(k+1/k) \}$, $\{ \bar{P}_{\bar{V}}^*(k, k) \}$ y $\{ \bar{P}_{\bar{V}\bar{V}}^*(k, k/k-1) \}$ definidas en el Lema 2.5.6,

$$\begin{aligned} \bar{\Pi}_*(k) &= p^2 \bar{H} \bar{P}_*(k/k-1) \bar{H}^T + \bar{P}_{\bar{V}\bar{V}}(k+1, k+1/k) \\ &\quad + p \bar{H} \bar{P}_{\bar{V}}^T(k+1, k+1) + \bar{P}_{\bar{V}}(k+1, k+1) p \bar{H}^T. \end{aligned}$$

Se tiene que

$$\bar{\Pi}(k+1) - \bar{\Pi}_*(k) = p^2 \bar{H} [\bar{P}(k+1/k) - \bar{P}^*(k/k-1)] \bar{H}^T \geq 0,$$

ya que $\bar{P}^*(k/k-1) \leq \bar{P}(k+1/k)$, según se ha demostrado en el Lema 2.5.6. Por consiguiente, $\bar{\Pi}(k+1) \geq \bar{\Pi}_*(k)$ para todo k .

Por otro lado, $\bar{\Pi}(k) \leq \bar{\Pi}_*(k)$ en virtud de la optimalidad del estimador asociado a $\{ \bar{P}(k+1/k) \}$. Uniendo ambas desigualdades, se concluye que

$$\bar{\Pi}(k) \leq \bar{\Pi}_*(k) \leq \bar{\Pi}(k+1), \quad k \geq 0$$

y, por tanto, la sucesión $\{ \bar{\Pi}(k) \}$ es monótona no decreciente. ■

Como consecuencia inmediata del Teorema 2.5.8 y del Lema anterior, podemos enunciar el siguiente resultado.

Teorema 2.5.10 *Si la matriz Φ es estrictamente estable y $\bar{P}(0/-1) = 0$, entonces existen los siguientes límites*

$$(i) \bar{\Pi} = \lim_{k \rightarrow \infty} \bar{\Pi}(k)$$

$$(ii) \bar{F} = \lim_{k \rightarrow \infty} \bar{F}(k).$$

Límite de las Ganancias del Predictor del Ruido

Como antes hemos indicado, la existencia de la matriz \bar{F} y la utilización, en cada iteración del algoritmo, de dicha matriz como ganancia del estimador del estado, presenta considerables ventajas computacionales.

Ahora bien, ello no simplifica demasiado los cálculos ya que, en cada iteración k , debemos determinar las matrices de ganancia $\bar{F}_{\bar{V}}(k, 0), \dots, \bar{F}_{\bar{V}}(k, k-1)$ para obtener el predictor del ruido de la observación, $\hat{\bar{V}}(k/k-1)$.

Por este motivo, con objeto nuevamente de simplificar los cálculos, nuestro siguiente propósito es establecer condiciones que garanticen que las ganancias $\{\bar{F}_{\bar{V}}(k, l); 0 \leq l < k\}$ convergen a un límite finito cuando $k \rightarrow \infty$ y utilizar dicho límite para la obtención de $\hat{\bar{V}}(k/k-1)$.

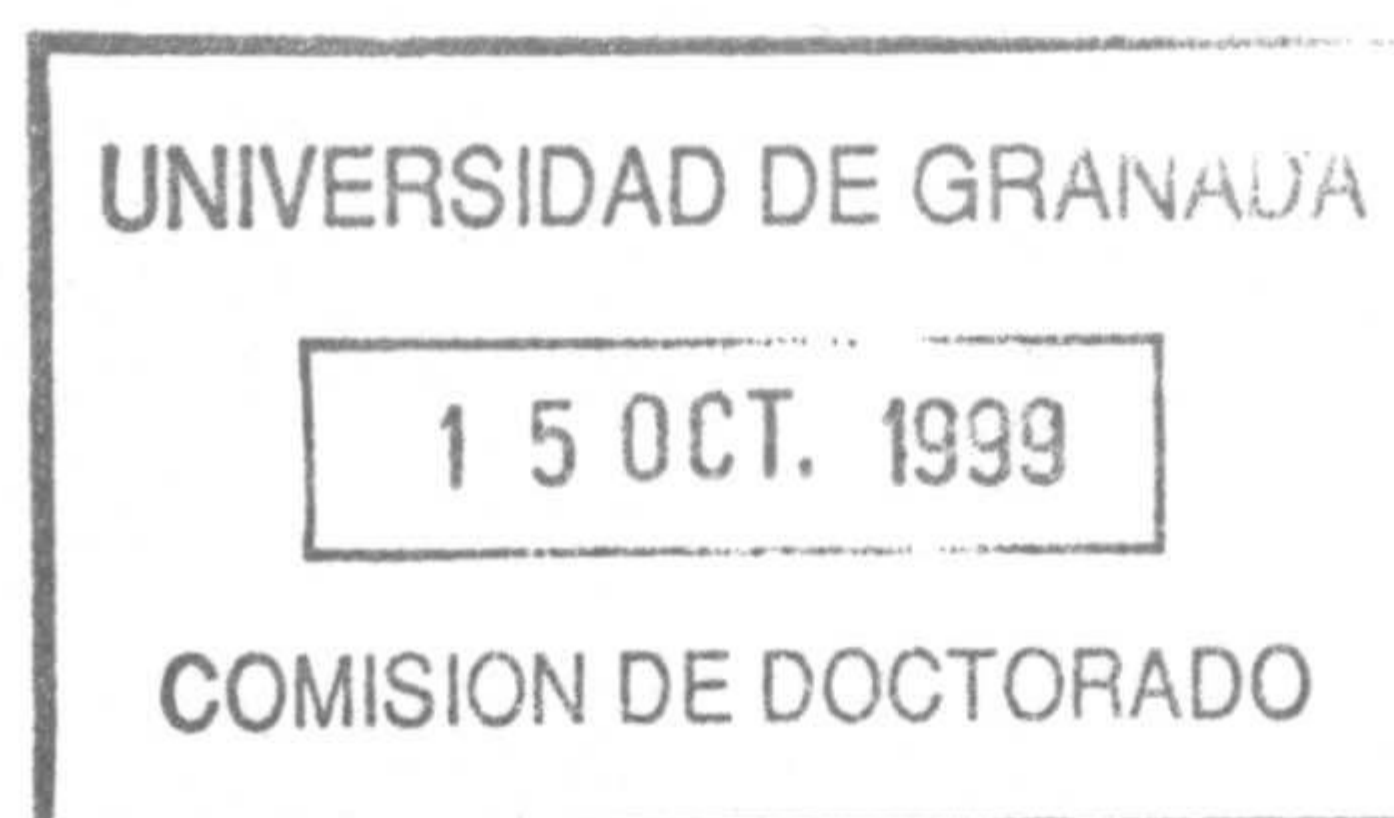
Puesto que las ganancias $\bar{F}_{\bar{V}}(k, l)$ dependen de las matrices $\bar{P}_{\bar{V}}(k, l)$ y $\bar{P}_{\bar{V}\bar{V}}(k, l/l-1)$ debemos analizar también la evolución de dichas matrices.

Proposición 2.5.11 *Si la matriz Φ es estrictamente estable, entonces, para l fijo, existen los siguientes límites*

$$\begin{aligned} \bar{P}_{\bar{V}\bar{V}}(l/l-1) &= \lim_{\substack{k \rightarrow \infty \\ 0 \leq l < k}} \bar{P}_{\bar{V}\bar{V}}(k, l/l-1) \\ \bar{P}_{\bar{V}}(l) &= \lim_{\substack{k \rightarrow \infty \\ 0 \leq l < k}} \bar{P}_{\bar{V}}(k, l) \\ \bar{F}_{\bar{V}}(l) &= \lim_{\substack{k \rightarrow \infty \\ 0 \leq l < k}} \bar{F}_{\bar{V}}(k, l) \end{aligned}$$

-Demostración-

Razonamos por inducción sobre l .



Para $l = 0$, se tiene que

$$\begin{aligned}\bar{P}_{\bar{V}\bar{V}}(k, 0/-1) &= \bar{R}(k) \\ \bar{P}_{\bar{V}}(k, 0) &= 0 \\ \bar{F}_{\bar{V}}(k, 0) &= \bar{R}(k)\bar{\Pi}^-(0)\end{aligned}$$

y, por tanto, tomando límite cuando $k \rightarrow \infty$, obtenemos

$$\begin{aligned}\bar{P}_{\bar{V}\bar{V}}(0/-1) &= \bar{R} \\ \bar{P}_{\bar{V}}(0) &= 0 \\ \bar{F}_{\bar{V}}(0) &= \bar{R}\bar{\Pi}^-(0),\end{aligned}$$

siendo

$$\bar{R} = \lim_{k \rightarrow \infty} \bar{R}(k) = p(p_{11} - p)\bar{H}\mathcal{D}\bar{H}^T$$

$$\text{con } \mathcal{D} = \lim_{k \rightarrow \infty} \mathcal{D}(k) = \begin{pmatrix} \sum_{i=0}^{\infty} \bar{\Phi}^i & 0 \\ 0 & (\text{vec}Q)(\text{vec}Q)^T \end{pmatrix} \sum_{i=0}^{\infty} (\bar{\Phi}^T)^i$$

(esta expresión es inmediata a partir de la expresión de $\mathcal{D}(k)$ dada en el Lema 2.5.3 y teniendo en cuenta la estabilidad estricta de $\bar{\Phi}$).

Sea $l > 0$ fijo, $0 < l < k$, y supongamos que para todo $i < l$ existen los siguientes límites

$$\bar{P}_{\bar{V}\bar{V}}(k, i/i-1) \xrightarrow{k \rightarrow \infty} \bar{P}_{\bar{V}\bar{V}}(i/i-1), \quad \bar{P}_{\bar{V}}(k, i) \xrightarrow{k \rightarrow \infty} \bar{P}_{\bar{V}}(i), \quad \bar{F}_{\bar{V}}(k, i) \xrightarrow{k \rightarrow \infty} \bar{F}_{\bar{V}}(i).$$

Teniendo en cuenta las expresiones

$$\begin{aligned}\bar{P}_{\bar{V}\bar{V}}(k, l/l-1) &= \bar{R}(k-l) - \sum_{i=0}^{l-1} \bar{F}_{\bar{V}}(k, i)\bar{\Pi}(i)\bar{F}_{\bar{V}}^T(l, i) \\ \bar{P}_{\bar{V}}(k, l) &= \bar{P}_{\bar{V}}(k, l-1)\bar{\Phi}^T - \bar{F}_{\bar{V}}(k, l-1)\bar{\Pi}(l-1)\bar{F}^T(l-1)\bar{\Phi}^T,\end{aligned}$$

en virtud de la hipótesis de inducción, cuando $k \rightarrow \infty$ estas matrices convergen a los siguientes límites

$$\begin{aligned}\bar{P}_{\bar{V}\bar{V}}(l/l-1) &= \bar{R} - \sum_{i=0}^{l-1} \bar{F}_{\bar{V}}(i) \bar{\Pi}(i) \bar{F}_{\bar{V}}^T(l,i) \\ \bar{P}_{\bar{V}}(l) &= \bar{P}_{\bar{V}}(l-1) \bar{\Phi}^T - \bar{F}_{\bar{V}}(l-1) \bar{\Pi}(l-1) \bar{F}^T(l-1) \bar{\Phi}^T.\end{aligned}$$

En consecuencia,

$$\bar{F}_{\bar{V}}(k,l) = \left[\bar{P}_{\bar{V}\bar{V}}(k,l/l-1) + \bar{P}_{\bar{V}}(k,l) p \bar{H}^T \right] \bar{\Pi}^-(l)$$

también converge, para $k \rightarrow \infty$, a la matriz

$$\bar{F}_{\bar{V}}(l) = \left[\bar{P}_{\bar{V}\bar{V}}(l/l-1) + \bar{P}_{\bar{V}}(l) p \bar{H}^T \right] \bar{\Pi}^-(l),$$

concluyendo así la demostración. ■

En una primera tentativa, podríamos utilizar como ganancia para el cálculo del predictor del ruido la matriz límite $\bar{F}_{\bar{V}}(l)$, es decir

$$\hat{V}(k/l) = \hat{V}(k/l-1) + \bar{F}_{\bar{V}}(l) \tilde{Z}(l/l-1), \quad l < k.$$

Reiterando esta expresión, obtendríamos

$$\hat{V}(k/k-1) = \hat{V}(k/-1) + \sum_{i=0}^{k-1} \bar{F}_{\bar{V}}(i) \tilde{Z}(i/i-1) = \sum_{i=0}^{k-1} \bar{F}_{\bar{V}}(i) \tilde{Z}(i/i-1),$$

de donde se deduce de forma inmediata la siguiente relación recursiva

$$\hat{V}(k/k-1) = \hat{V}(k-1/k-2) + \bar{F}_{\bar{V}}(k-1) \tilde{Z}(k-1/k-2).$$

Ahora bien, si probamos que la sucesión de matrices $\{\bar{F}_{\bar{V}}(l)\}$ tiene límite cuando $l \rightarrow \infty$, podremos utilizar dicho límite en la expresión anterior del predictor del ruido. De este modo, obtendríamos una forma steady-state para el mismo, que sería utilizada en la obtención del filtro steady-state del estado.

La convergencia de la sucesión $\{\bar{F}_{\bar{V}}(l)\}$ va a venir determinada por la existencia del límite

$$\lim_{l \rightarrow \infty} \left(\lim_{k \rightarrow \infty} \bar{P}_{\bar{V}\bar{V}}(k, k/l) \right)$$

que se establece en los lemas siguientes.

Lema 2.5.12 *Si la matriz Φ es estrictamente estable, entonces, para l fijo, existe el siguiente límite*

$$\bar{P}_{\bar{V}\bar{V}}(l) = \lim_{\substack{k \rightarrow \infty \\ -1 \leq l < k}} \bar{P}_{\bar{V}\bar{V}}(k, k/l) = \bar{R}(0) - \sum_{i=0}^l \bar{F}_{\bar{V}}(i) \bar{\Pi}(i) \bar{F}_{\bar{V}}^T(i).$$

-Demostración-

Para $l = -1$, es claro que

$$\bar{P}_{\bar{V}\bar{V}}(-1) = \lim_{k \rightarrow \infty} \bar{P}_{\bar{V}\bar{V}}(k, k/-1) = \lim_{k \rightarrow \infty} \bar{R}(0) = \bar{R}(0).$$

Si $l \geq 0$, reiterando la expresión de $\bar{P}_{\bar{V}\bar{V}}(k, k/l)$, obtenemos

$$\bar{P}_{\bar{V}\bar{V}}(k, k/l) = \bar{R}(0) - \sum_{i=0}^l \bar{F}_{\bar{V}}(k, i) \bar{\Pi}(i) \bar{F}_{\bar{V}}^T(k, i).$$

Puesto que Φ es estrictamente estable, aplicando la Proposición 2.5.11, se tiene que para $0 \leq i \leq l < k$,

$$\lim_{k \rightarrow \infty} \bar{F}_{\bar{V}}(k, i) = \bar{F}_{\bar{V}}(i)$$

y, por tanto,

$$\lim_{k \rightarrow \infty} \bar{P}_{\bar{V}\bar{V}}(k, k/l) = \bar{P}_{\bar{V}\bar{V}}(l) = \bar{R}(0) - \sum_{i=0}^l \bar{F}_{\bar{V}}(i) \bar{\Pi}(i) \bar{F}_{\bar{V}}^T(i), \quad l \geq 0.$$

■

Lema 2.5.13 *Si la matriz Φ es estrictamente estable, entonces la sucesión $\{\bar{P}_{\bar{V}\bar{V}}(l)\}$ es monótona decreciente y acotada, por tanto, dicha sucesión es convergente. Además,*

$$\bar{P}_{\bar{V}\bar{V}} = \lim_{l \rightarrow \infty} \bar{P}_{\bar{V}\bar{V}}(l) = \bar{R}(0) - \sum_{i=0}^{\infty} \bar{F}_{\bar{V}}(i) \bar{\Pi}(i) \bar{F}_{\bar{V}}^T(i).$$

-Demostración-

- Monotonía. De la expresión de $\bar{P}_{\bar{V}\bar{V}}(l)$ obtenida en el Lema anterior, es inmediato que

$$\begin{aligned}\bar{P}_{\bar{V}\bar{V}}(l) &= \bar{P}_{\bar{V}\bar{V}}(l-1) - \bar{F}_{\bar{V}}(l) \bar{\Pi}(l) \bar{F}_{\bar{V}}^T(l) \\ \bar{P}_{\bar{V}\bar{V}}(-1) &= \bar{R}(0).\end{aligned}$$

Al ser $\bar{\Pi}(l)$ definida no negativa, podemos afirmar que

$$\bar{F}_{\bar{V}}(l) \bar{\Pi}(l) \bar{F}_{\bar{V}}^T(l) \geq 0, \quad l \geq 0$$

y, por tanto,

$$\bar{P}_{\bar{V}\bar{V}}(l) \leq \bar{P}_{\bar{V}\bar{V}}(l-1), \quad l \geq 0$$

lo que garantiza el decrecimiento de la sucesión $\{\bar{P}_{\bar{V}\bar{V}}(l)\}$.

- Acotación. Puesto que $\bar{P}_{\bar{V}\bar{V}}(l)$ es límite de matrices de covarianzas, se tiene que

$$\bar{P}_{\bar{V}\bar{V}}(l) \geq 0, \quad l \geq 0.$$

Por otro lado, al ser decreciente, podemos escribir

$$\bar{P}_{\bar{V}\bar{V}}(l) \leq \bar{P}_{\bar{V}\bar{V}}(-1) = \bar{R}(0), \quad l \geq 0$$

y, por consiguiente,

$$0 \leq \bar{P}_{\bar{V}\bar{V}}(l) \leq \bar{R}(0), \quad l \geq 0.$$

de donde se deduce que la sucesión $\{\bar{P}_{\bar{V}\bar{V}}(l)\}$ está acotada.

En consecuencia, la sucesión $\{\bar{P}_{\bar{V}\bar{V}}(l)\}$ es convergente. Para calcular su límite, basta hacer $l \rightarrow \infty$ en la expresión dada en el Lema anterior, obteniendo

$$\bar{P}_{\bar{V}\bar{V}} = \lim_{l \rightarrow \infty} \bar{P}_{\bar{V}\bar{V}}(l) = \bar{R}(0) - \sum_{i=0}^{\infty} \bar{F}_{\bar{V}}(i) \bar{\Pi}(i) \bar{F}_{\bar{V}}^T(i)$$

como queríamos demostrar. ■

La convergencia de la serie $\sum_{i=0}^{\infty} \bar{F}_{\bar{V}}(i) \bar{\Pi}(i) \bar{F}_{\bar{V}}^T(i)$ nos permite afirmar que

$$\lim_{i \rightarrow \infty} \bar{F}_{\bar{V}}(i) \bar{\Pi}(i) \bar{F}_{\bar{V}}^T(i) = 0.$$

Por otro lado sabemos que, si $\bar{P}(0/-1) = 0$, entonces $\lim_{i \rightarrow \infty} \bar{\Pi}(i) = \bar{\Pi}$. Uniendo ambos resultados, podemos enunciar el siguiente resultado.

Teorema 2.5.14 *Si la matriz Φ es estrictamente estable y $\bar{P}(0/-1) = 0$, entonces la sucesión $\{\bar{F}_{\bar{V}}(l)\}$ tiene límite finito, $\bar{F}_{\bar{V}}$, cuando $l \rightarrow \infty$.*

El estudio realizado sobre el comportamiento asintótico de las ganancias se puede resumir en el siguiente Teorema.

Teorema 2.5.15 *Bajo las hipótesis de estabilidad estricta de Φ y condición inicial $\bar{P}(0/-1) = 0$, el algoritmo de predicción y filtrado obtenido en la Sección 2.4.3 admite una forma steady-state, en el sentido de que las matrices de ganancia, $\bar{F}(k)$ y $\bar{F}_{\bar{V}}(k, l)$, tienen límites finitos, $\bar{F} = \lim_{k \rightarrow \infty} \bar{F}(k)$ y $\bar{F}_{\bar{V}} = \lim_{l \rightarrow \infty} \left(\lim_{k \rightarrow \infty} \bar{F}_{\bar{V}}(k, l) \right)$, respectivamente.*

Sustituyendo las ganancias del filtro del estado y del predictor del ruido por sus límites, obtenemos la versión steady-state del algoritmo presentado en la Sección 2.4.3

$$\hat{X}(k/k) = \hat{X}(k/k-1) + \bar{F} \tilde{Z}(k/k-1), \quad k \geq 0$$

$$\hat{X}(k+1/k) = \Phi \hat{X}(k/k), \quad k \geq 0$$

$$\hat{X}(0/-1) = 0$$

$$\tilde{Z}(k/k-1) = \bar{Z}(k) - p \bar{H} \hat{X}(k/k-1) - \hat{V}(k/k-1), \quad k \geq 0$$

$$\hat{V}(k/k-1) = \hat{V}(k-1/k-2) + \bar{F}_{\bar{V}} \tilde{Z}(k-1/k-2)$$

$$\hat{V}(0/-1) = 0.$$

Capítulo 3

Estimación Polinomial en Sistemas con Observaciones Inciertas

3.1 Introducción

Como se ha indicado en los capítulos anteriores en sistemas con observaciones inciertas, aún bajo hipótesis de gaussianidad e independencia mutua del estado inicial y los ruidos, la distribución conjunta del estado y las observaciones no es gaussiana; por tanto, la esperanza condicionada, que proporciona el estimador de mínimos cuadrados, no es una función lineal de las observaciones y su cómputo requiere un crecimiento exponencial de memoria (Capítulo 1, Sección 1.4.2). Por este motivo, el estudio del problema de estimación en estos sistemas se ha dirigido fundamentalmente hacia la obtención de estimadores subóptimos.

Dentro del problema de estimación subóptima en sistemas con observaciones inciertas cuya incertidumbre está modelizada por variables aleatorias (de Bernoulli) indepen-

dientes, bajo la hipótesis de independencia mutua de los ruidos y la condición inicial, el caso lineal fue tratado por Nahi [31] (Capítulo 1, Sección 1.4.3) y, posteriormente, García-Ligero et al. [12] obtuvieron el filtro cuadrático (Capítulo 2, Sección 2.3).

El objetivo de este Capítulo es el estudio del problema de estimación polinomial de grado ν arbitrario ($\nu \geq 1$) en sistemas con observaciones inciertas con incertidumbre modelizada por variables aleatorias independientes. Abordaremos dicho problema de estimación polinomial, tanto en el caso en que los ruidos sean mutuamente independientes (esto es, para los sistemas considerados por Nahi y García-Ligero et al.), como la extensión del mismo suponiendo que los ruidos aditivos que afectan al sistema son correlados entre sí. Este estudio generaliza el trabajo de García-Ligero et al. [12] en dos direcciones: por una parte, se consideran estimadores polinomiales de grado arbitrario y, por otra, se debilita la hipótesis de independencia sobre los ruidos del sistema.

El procedimiento que seguiremos, similar al utilizado en el estudio del problema de estimación cuadrática (Capítulo 2), consiste en construir un nuevo sistema lineal (sistema aumentado), de modo que el estimador lineal óptimo del estado de dicho sistema aumentado se obtendrá aplicando un algoritmo de filtrado lineal y, a partir de él, se determinará el estimador polinomial de grado ν del estado del sistema original.

La estructura del Capítulo se puede desglosar de la siguiente forma.

Siguiendo la línea de los capítulos anteriores, en la Sección 3.2 se presenta un enfoque geométrico del problema de estimación polinomial de menor error cuadrático medio.

A continuación (Sección 3.3), se describen las ecuaciones que modelizan los sistemas bajo estudio, se especifican las hipótesis sobre los procesos que en ellas intervienen y se plantea el problema de estimación polinomial de menor error cuadrático medio en dichos sistemas.

En la Sección 3.4 se expone una serie de resultados preliminares que serán básicos para determinar el sistema aumentado y analizar sus propiedades. Dicho estudio sobre

el sistema aumentado se realiza en la Sección 3.5. En esta misma Sección se establece el algoritmo de filtrado lineal para el estado del sistema aumentado, a partir del cual se obtendrá el estimador polinomial del estado del sistema original. En el análisis de las propiedades de los ruidos del sistema aumentado y, consecuentemente, en la obtención del algoritmo de filtrado lineal, distinguimos dos casos: en primer lugar, supondremos que el estado inicial y los ruidos que afectan al sistema original son mutuamente independientes y, a continuación, extenderemos el estudio al caso en que los ruidos aditivos del sistema son correlados entre sí.

Seguidamente, en la Sección 3.6, se considera la versión estacionaria de los sistemas en estudio y se demuestra que el correspondiente sistema aumentado es asintóticamente estacionario si el sistema original es asintóticamente estable.

Por último, en la Sección 3.7, se estudia el comportamiento asintótico del filtro lineal presentado en la Sección 3.5, concluyendo que, tanto en el caso en que los ruidos y la condición inicial son mutuamente independientes, como en el caso de ruidos aditivos correlados, si el sistema original es asintóticamente estable, dicho filtro admite una versión steady-state.

3.2 Estimación Polinomial y Proyecciones Ortogonales

Como se recordaba en los capítulos anteriores, dados dos vectores aleatorios, $X \in \mathcal{L}_{\mathbb{R}^n}^2(\Omega, A, P)$ e $Y \in \mathcal{L}_{\mathbb{R}^m}^2(\Omega, A, P)$, el estimador de mínimos cuadrados de X basado en Y es la esperanza condicionada $E[X/Y]$ y, a excepción del caso gaussiano, esta esperanza condicionada, en general, no puede calcularse fácilmente, por lo que el interés se centra en la obtención de estimadores subóptimos.

También es conocido que el estimador subóptimo con estructura más simple es el lineal (Capítulo 1, Sección 1.2.2), seguido del estimador cuadrático o polinomial de segundo grado (Capítulo 2, Sección 2.2). En esta Sección consideraremos estimado-

res definidos mediante proyecciones del vector X en subespacios de transformaciones polinómicas de grado ν arbitrario ($\nu \geq 1$) del vector Y .

Con este fin, definimos los vectores aleatorios

$$Y^{[j]} = Y \otimes \dots \otimes Y \in \mathbb{R}^{m^j}, \quad j = 1, \dots, \nu$$

donde \otimes denota el producto Kronecker (véase Apéndice A). Bajo la hipótesis de que $E\{Y^{[2\nu]}\} < \infty$, el vector $Y_\nu = (1, Y^T, \dots, Y^{[\nu]T})^T \in \mathcal{L}_{\mathbb{R}^{m\nu+\dots+m+1}}^2(\Omega, \sigma(Y), P)$, siendo $\sigma(Y)$ la σ -álgebra generada por Y .

El estimador polinomial de grado ν insesgado de mínima varianza de X basado en Y , que notaremos \hat{X}^ν , es la proyección ortogonal de X sobre el subespacio lineal $\mathcal{H}_n(Y_\nu)$ de vectores n -dimensionales obtenidos mediante transformaciones lineales del vector Y_ν .

Puesto que $\mathcal{H}_n(Y_1) \subset \dots \subset \mathcal{H}_n(Y_{\nu-1}) \subset \mathcal{H}_n(Y_\nu)$, se sigue que

$$E\left\{\|X - \hat{X}^\nu\|^2\right\} \leq E\left\{\|X - \hat{X}^{\nu-1}\|^2\right\} \leq \dots \leq E\left\{\|X - \hat{X}^1\|^2\right\} = E\left\{\|X - \hat{X}\|^2\right\}.$$

En consecuencia, el error cometido al estimar X mediante una función polinómica del vector Y disminuye al aumentar el grado y, por tanto, dicho error es siempre menor o igual que el cometido al utilizar el estimador lineal \hat{X} .

3.3 Descripción del Sistema y Planteamiento del Problema de Estimación

Consideremos el siguiente sistema lineal en tiempo discreto

$$x(k+1) = A(k)x(k) + F(k)w(k), \quad k \geq 0 \quad (3.1)$$

$$x(0) = x_0$$

$$z(k) = u(k)C(k)x(k) + G(k)v(k), \quad k \geq 0 \quad (3.2)$$

donde $x(k)$ es un vector aleatorio n -dimensional que define el estado del sistema en el instante k ; $z(k) \in \mathbb{R}^m$ es el vector observación; $\{w(k); k \geq 0\}$ y $\{v(k); k \geq 0\}$ son procesos estocásticos con valores en \mathbb{R}^u y \mathbb{R}^t , respectivamente; $\{u(k); k \geq 0\}$ es una sucesión ruido escalar y $A(k)$, $F(k)$, $C(k)$, $G(k)$ son matrices determinísticas conocidas de dimensiones apropiadas.

Sobre el estado inicial x_0 y los ruidos $\{w(k); k \geq 0\}$, $\{v(k); k \geq 0\}$, $\{u(k); k \geq 0\}$ imponemos las siguientes hipótesis

1. x_0 es un vector aleatorio centrado.
2. $\{w(k); k \geq 0\}$, $\{v(k); k \geq 0\}$ son ruidos blancos centrados.
3. Existe $\nu \geq 1$ tal que

$$\begin{aligned} E \left\{ w^{[2\nu]}(k) \right\} &< \infty, & k \geq 0 \\ E \left\{ v^{[2\nu]}(k) \right\} &< \infty, & k \geq 0 \\ E \left\{ x_0^{[2\nu]} \right\} &< \infty \end{aligned}$$

4. $\{u(k); k \geq 0\}$ es una sucesión de variables aleatorias independientes de Bernoulli, con $P[u(k) = 1] = p(k)$.

Nuestro objetivo es obtener el estimador polinomial de grado ν óptimo del estado del sistema, $x(k)$, basado en las observaciones hasta el instante k . En virtud de los resultados presentados en la Sección 3.2, este estimador es la proyección ortogonal de $x(k)$ sobre el subespacio $\mathcal{H}_n \left((1, z^T(0), \dots, z^{[\nu]T}(0), \dots, z^T(k), \dots, z^{[\nu]T}(k)) \right)^T$ de vectores n -dimensionales obtenidos mediante transformaciones lineales de las observaciones $z(0), \dots, z(k)$ y sus potencias de Kronecker $z^{[2]}(0), \dots, z^{[\nu]}(0), \dots, z^{[2]}(k), \dots, z^{[\nu]}(k)$.

Con el fin de obtener dicho estimador, definimos el *vector estado aumentado*

$$X(k) = \begin{pmatrix} x(k) \\ x^{[2]}(k) \\ \vdots \\ x^{[\nu]}(k) \end{pmatrix}, \quad k > 0; \quad X_0 = \begin{pmatrix} x_0 \\ x_0^{[2]} \\ \vdots \\ x_0^{[\nu]} \end{pmatrix} \quad (3.3)$$

y el *vector observación aumentado*

$$Z(k) = \begin{pmatrix} z(k) \\ z^{[2]}(k) \\ \vdots \\ z^{[\nu]}(k) \end{pmatrix}, \quad k \geq 0. \quad (3.4)$$

El vector constituido por las n primeras componentes del estimador lineal de menor error cuadrático medio de $X(k)$ basado en las observaciones $\{Z(0), \dots, Z(k)\}$ nos proporcionará el filtro polinomial óptimo de grado ν del estado $x(k)$ del sistema original.

3.4 Resultados Previos

Nuestro siguiente objetivo es analizar la evolución de los vectores $X(k)$ y $Z(k)$ definidos por (3.3) y (3.4), respectivamente, con el fin de comprobar si dicha evolución puede ser modelizada mediante un sistema lineal en tiempo discreto con las propiedades requeridas para poder aplicar alguno de los algoritmos de filtrado lineal conocidos.

En esta Sección presentamos una serie de resultados generales que nos permitirán establecer, en la Sección 3.5, el sistema aumentado y analizar sus propiedades.

El siguiente Lema será básico para determinar la evolución de los vectores $X(k)$ y $Z(k)$ y, con ello, definir el sistema aumentado.

Lema 3.4.1 Sean $\{\beta(k); k \geq 0\}$, $\{\mu(k); k \geq 0\}$ procesos estocásticos con valores en \mathbb{R}^n y \mathbb{R}^t , respectivamente, y sea $\{\gamma(k); k \geq 0\}$ una sucesión de variables aleatorias

independientes de Bernoulli con $P[\gamma(k) = 1] = p(k)$. Definimos

$$\alpha(k) = \gamma(k) \Gamma(k) \beta(k) + \Psi(k) \mu(k)$$

donde $\alpha(k) \in \mathbb{R}^m$ y $\Gamma(k)$, $\Psi(k)$ son matrices determinísticas de dimensiones apropiadas.

Consideremos las potencias de Kronecker de los vectores y matrices anteriores y definamos

$$\Lambda(k) = \begin{pmatrix} \alpha(k) \\ \alpha^{[2]}(k) \\ \vdots \\ \alpha^{[\nu]}(k) \end{pmatrix} \quad B(k) = \begin{pmatrix} \beta(k) \\ \beta^{[2]}(k) \\ \vdots \\ \beta^{[\nu]}(k) \end{pmatrix}$$

$$\Theta(k) = \begin{pmatrix} \Gamma(k) & 0 & \dots & 0 \\ \theta_{21}(k) & \Gamma^{[2]}(k) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \theta_{\nu 1}(k) & \theta_{\nu 2}(k) & \dots & \Gamma^{[\nu]}(k) \end{pmatrix} \quad \Upsilon(k) = \begin{pmatrix} \Psi(k) E\{\mu(k)\} \\ \Psi^{[2]}(k) E\{\mu^{[2]}(k)\} \\ \vdots \\ \Psi^{[\nu]}(k) E\{\mu^{[\nu]}(k)\} \end{pmatrix}$$

donde

$$\theta_{jl}(k) = M_{j-1}^j(m) \left(\Psi^{[j-1]}(k) \otimes \Gamma^{[l]}(k) \right) \left(E\{\mu^{[j-1]}(k)\} \otimes I_{n,l} \right)$$

y $M_{j-1}^j(m)$ son los coeficientes matriciales dados en el Lema A.4.1 (véase Apéndice A).

Entonces, el vector $\Lambda(k)$ admite la representación

$$\Lambda(k) = \gamma(k) \Theta(k) B(k) + \Upsilon(k) + \Delta(k) \quad (3.5)$$

siendo

$$\Delta(k) = \begin{pmatrix} h_1(k) \\ h_2(k) \\ \vdots \\ h_\nu(k) \end{pmatrix} \quad (3.6)$$

con

$$h_j(k) = \gamma(k) \sum_{l=1}^{j-1} M_{j-l}^j(m) \left(\Psi^{[j-l]}(k) \otimes \Gamma^{[l]}(k) \right) \left(\left(\mu^{[j-l]}(k) - E \left\{ \mu^{[j-l]}(k) \right\} \right) \otimes I_{n,l} \right) \beta^{[l]}(k) + \Psi^{[j]}(k) \left(\mu^{[j]}(k) - E \left\{ \mu^{[j]}(k) \right\} \right) \quad (3.7)$$

siempre que todos los momentos que intervienen sean finitos.

-Demostración-

Elevemos $\alpha(k)$ a la j -sima potencia de Kronecker

$$\alpha^{[j]}(k) = (\gamma(k) \Gamma(k) \beta(k) + \Psi(k) \mu(k))^{[j]}.$$

Utilizando el Lema A.4.1, las propiedades del producto Kronecker (véase Apéndice A) y teniendo en cuenta que $\gamma(k) = \gamma^2(k) = \dots = \gamma^j(k)$, podemos escribir

$$\begin{aligned} \alpha^{[j]}(k) &= \sum_{r=0}^j M_r^j(m) \left((\Psi(k) \mu(k))^{[r]} \otimes (\gamma(k) \Gamma(k) \beta(k))^{[j-r]} \right) \\ &= \gamma(k) \Gamma^{[j]}(k) \beta^{[j]}(k) + \Psi^{[j]}(k) \mu^{[j]}(k) \\ &\quad + \sum_{r=1}^{j-1} M_r^j(m) \left((\Psi(k) \mu(k))^{[r]} \otimes (\gamma(k) \Gamma(k) \beta(k))^{[j-r]} \right) \\ &= \gamma(k) \Gamma^{[j]}(k) \beta^{[j]}(k) + \Psi^{[j]}(k) \mu^{[j]}(k) \\ &\quad + \gamma(k) \sum_{r=1}^{j-1} M_r^j(m) \left(\Psi^{[r]}(k) \otimes \Gamma^{[j-r]}(k) \right) \left(\mu^{[r]}(k) \otimes \beta^{[j-r]}(k) \right) \\ &= \gamma(k) \Gamma^{[j]}(k) \beta^{[j]}(k) + \Psi^{[j]}(k) \mu^{[j]}(k) \\ &\quad + \gamma(k) \sum_{l=1}^{j-1} M_{j-l}^j(m) \left(\Psi^{[j-l]}(k) \otimes \Gamma^{[l]}(k) \right) \left(\mu^{[j-l]}(k) \otimes I_{n,l} \right) \beta^{[l]}(k) \\ &= \gamma(k) \Gamma^{[j]}(k) \beta^{[j]}(k) + \gamma(k) \sum_{l=1}^{j-1} \theta_{jl}(k) \beta^{[l]}(k) + \Psi^{[j]}(k) E \left\{ \mu^{[j]}(k) \right\} + h_j(k) \end{aligned}$$

que es la componente j -sima de (3.5). ■

Nota 3.4.1 Para $i \in \{1, 2\}$, sean $\{\beta_i(k); k \geq 0\}$, $\{\mu_i(k); k \geq 0\}$ procesos estocásticos con valores en \mathbb{R}^{n_i} y \mathbb{R}^{t_i} , respectivamente, y sean $\{\gamma_i(k); k \geq 0\}$ dos sucesiones de variables aleatorias independientes de Bernoulli con $P[\gamma_i(k) = 1] = p_i(k)$. Definimos los procesos

$$\alpha_i(k) = \gamma_i(k) \Gamma_i(k) \beta_i(k) + \Psi_i(k) \mu_i(k), \quad i \in \{1, 2\}$$

siendo $\alpha_i(k) \in \mathbb{R}^{m_i}$ y $\Gamma_i(k)$, $\Psi_i(k)$ matrices determinísticas de dimensiones apropiadas.

Aplicando el Lema 3.4.1, obtenemos que los vectores $\Lambda_i(k)$ admiten la representación

$$\Lambda_i(k) = \gamma_i(k) \Theta_i(k) B_i(k) + \Upsilon_i(k) + \Delta_i(k), \quad i \in \{1, 2\}.$$

En el siguiente Lema, estudiamos las propiedades de los procesos $\{\Delta_i(k); k \geq 0\}$ definidos mediante (3.6) y (3.7). Para ello, introducimos la siguiente notación simplificada

$$\begin{aligned} M_{r,l}^i(k) &= M_{r-l}^r(m_i) \left(\Psi_i^{[r-l]}(k) \otimes \Gamma_i^{[l]}(k) \right), \quad i \in \{1, 2\} \\ \mu_{r,l}^i(k) &= \mu_i^{[r-l]}(k) - E \left\{ \mu_i^{[r-l]}(k) \right\}, \quad i \in \{1, 2\} \\ P_{l,j}^{r,s}(\mu_1, \mu_2) &= E \left\{ (\mu_{r,l}^1(k) \otimes I_{n_1,l}) \beta_1^{[l]}(k) \beta_2^{[j]T}(k) (\mu_{s,j}^2(k) \otimes I_{n_2,j})^T \right\}. \end{aligned}$$

Lema 3.4.2 Para $i \in \{1, 2\}$, sean $\{\gamma_i(k); k \geq 0\}$, $\{\beta_i(k); k \geq 0\}$ y $\{\mu_i(k); k \geq 0\}$ procesos estocásticos tales que

(i) $(\{\gamma_1(k)\}, \{\gamma_2(k)\})$ es independiente de $(\{\beta_1(k)\}, \{\beta_2(k)\}, \{\mu_1(k)\}, \{\mu_2(k)\})$

(ii) $(\mu_1(k), \mu_2(k))$ es independiente de

$$(\{\beta_1(j), j \leq k\}, \{\beta_2(j), j \leq k\}, \{\mu_1(j), j < k\}, \{\mu_2(j), j < k\})$$

Bajo estas hipótesis, los procesos $\{\Delta_1(k)\}, \{\Delta_2(k)\}$ definidos por (3.6) y (3.7) son centrados y verifican que

$$E \{ \Delta_1(k) \Delta_2^T(j) \} = S(k) \delta_{kj}$$

siendo δ_{kj} la función delta de Kronecker.

El bloque (r, s) de $\mathcal{S}(k)$, utilizando la notación introducida en la Nota 3.4.1, viene dado por

$$\begin{aligned}
 \mathcal{S}_{rs}(k) &= E \{ h_r^1(k) h_s^{2T}(k) \} \\
 &= E \{ \gamma_1(k) \gamma_2(k) \} \sum_{l=1}^{r-1} \sum_{j=1}^{s-1} M_{r,l}^1(k) P_{l,j}^{r,s}(\mu_1, \mu_2) (M_{s,j}^2(k))^T \\
 &\quad + p_1(k) \sum_{l=1}^{r-1} M_{r,l}^1(k) P_{l,0}^{r,s}(\mu_1, \mu_2) \Psi_2^{[s]T}(k) \\
 &\quad + p_2(k) \Psi_1^{[r]}(k) \sum_{j=1}^{s-1} P_{0,j}^{r,s}(\mu_1, \mu_2) (M_{s,j}^2(k))^T \\
 &\quad + \Psi_1^{[r]}(k) P_{0,0}^{r,s}(\mu_1, \mu_2) \Psi_2^{[s]T}(k)
 \end{aligned} \tag{3.8}$$

siendo

$$\begin{aligned}
 P_{l,j}^{r,s}(\mu_1, \mu_2) &= \text{vec}^{-1} \left[\left(I_{t_2, s-j} \otimes K_{t_1^{r-l}, n_2^j}^T \otimes I_{n_1, l} \right) \right. \\
 &\quad \times \left(\left(E \{ \mu_2^{[s-j]}(k) \otimes \mu_1^{[r-l]}(k) \} - E \{ \mu_2^{[s-j]}(k) \} \otimes E \{ \mu_1^{[r-l]}(k) \} \right) \right. \\
 &\quad \left. \left. \otimes K_{1, n_2^j} \otimes I_{n_1, l} \right) E \{ \beta_2^{[j]}(k) \otimes \beta_1^{[l]}(k) \} \right]
 \end{aligned} \tag{3.9}$$

siempre que todos los momentos que intervienen sean finitos.

-Demostración-

Los procesos $\{\Delta_1(k)\}$, $\{\Delta_2(k)\}$ son centrados, ya que

$$\begin{aligned}
 E \{ h_r^i(k) \} &= p_i(k) \sum_{l=1}^{r-1} M_{r,l}^i(k) E \{ (\mu_{r,l}^i(k) \otimes I_{n_i, l}) \beta_i^{[l]}(k) \} + \Psi_i^{[r]}(k) E \{ \mu_{r,0}^i(k) \} \\
 &= p_i(k) \sum_{l=1}^{r-1} M_{r,l}^i(k) (E \{ \mu_{r,l}^i(k) \} \otimes I_{n_i, l}) E \{ \beta_i^{[l]}(k) \} \\
 &= 0
 \end{aligned}$$

donde se ha utilizado que $\gamma_i(k)$ es independiente de $(\mu_i(k), \beta_i(k))$ y que $\mu_i(k)$ es independiente de $\beta_i(k)$.

Para demostrar que $\Delta_1(k)$ y $\Delta_2(q)$ son incorrelados para $k \neq q$, observemos que

$$\begin{aligned} E \{h_r^1(k) h_s^{2T}(q)\} &= E \{\gamma_1(k) \gamma_2(q)\} \sum_{l=1}^{r-1} \sum_{j=1}^{s-1} M_{r,l}^1(k) \\ &\times E \left\{ \left(\mu_{r,l}^1(k) \otimes I_{n_1,l} \right) \beta_1^{[l]}(k) \beta_2^{[j]T}(q) \left(\mu_{s,j}^2(q) \otimes I_{n_2,j} \right)^T \right\} \left(M_{s,j}^2(q) \right)^T \\ &+ E \{\gamma_1(k)\} \sum_{l=1}^{r-1} M_{r,l}^1(k) E \left\{ \left(\mu_{r,l}^1(k) \otimes I_{n_1,l} \right) \beta_1^{[l]}(k) \left(\mu_{s,0}^2(q) \right)^T \right\} \Psi_2^{[s]T}(q) \\ &+ E \{\gamma_2(q)\} \Psi_1^{[r]}(k) \sum_{j=1}^{s-1} E \left\{ \mu_{r,0}^1(k) \beta_2^{[j]T}(q) \left(\mu_{s,j}^2(q) \otimes I_{n_2,j} \right)^T \right\} \left(M_{s,j}^2(q) \right)^T \\ &+ \Psi_1^{[r]}(k) E \left\{ \mu_{r,0}^1(k) \mu_{s,0}^{2T}(q) \right\} \Psi_2^{[s]T}(q) \end{aligned}$$

donde se ha utilizado que $(\{\gamma_1(k)\}, \{\gamma_2(k)\})$ y $(\{\beta_1(k)\}, \{\beta_2(k)\}, \{\mu_1(k)\}, \{\mu_2(k)\})$ son independientes. Teniendo en cuenta que, para $a, b \in \{1, 2\}$, $a \neq b$, el vector $\mu_a(k)$ es independiente de $(\beta_a(k), \{\beta_b(q), q < k\}, \{\mu_b(q), q < k\})$ y que $E \left\{ \mu_{r,l}^1(k) \right\} = 0$, para $l = 0, \dots, r-1$, concluimos que

$$E \{h_r^1(k) h_s^{2T}(q)\} = 0, \quad \forall k \neq q.$$

Para obtener (3.8), observemos que

$$\begin{aligned} E \{h_r^1(k) h_s^{2T}(k)\} &= E \{\gamma_1(k) \gamma_2(k)\} \sum_{l=1}^{r-1} \sum_{j=1}^{s-1} M_{r,l}^1(k) P_{l,j}^{r,s}(\mu_1, \mu_2) \left(M_{s,j}^2(k) \right)^T \\ &+ p_1(k) \sum_{l=1}^{r-1} M_{r,l}^1(k) P_{l,0}^{r,s}(\mu_1, \mu_2) \Psi_2^{[s]T}(k) \\ &+ p_2(k) \Psi_1^{[r]}(k) \sum_{j=1}^{s-1} P_{0,j}^{r,s}(\mu_1, \mu_2) \left(M_{s,j}^2(k) \right)^T \\ &+ \Psi_1^{[r]}(k) P_{0,0}^{r,s}(\mu_1, \mu_2) \Psi_2^{[s]T}(k). \end{aligned}$$

Para concluir la demostración basta obtener la expresión (3.9). El argumento de la esperanza que aparece en la definición de $P_{l,j}^{r,s}(\mu_1, \mu_2)$ se expresa como

$$\text{vec}^{-1} \left[\text{vec} \left(\left(\mu_{r,l}^1(k) \otimes I_{n_1,l} \right) \beta_1^{[l]}(k) \beta_2^{[j]T}(k) \left(\mu_{s,j}^2(k) \otimes I_{n_2,j} \right)^T \right) \right].$$

Por otro lado, haciendo uso de las propiedades del operador vec y del producto Kronecker (ver Apéndice A), se tiene que

$$\begin{aligned}
& vec\left(\left(\mu_{r,l}^1(k) \otimes I_{n_1,l}\right) \beta_1^{[l]}(k) \beta_2^{[j]T}(k) \left(\mu_{s,j}^2(k) \otimes I_{n_2,j}\right)^T\right) \\
&= \left(\left(\mu_{s,j}^2(k) \otimes I_{n_2,j}\right) \otimes \left(\mu_{r,l}^1(k) \otimes I_{n_1,l}\right)\right) vec\left(\beta_1^{[l]}(k) \beta_2^{[j]T}(k)\right) \\
&= \left(\mu_{s,j}^2(k) \otimes \left(K_{t_1^{r-l}, n_2^j}^T \left(\mu_{r,l}^1(k) \otimes I_{n_2,j}\right) K_{1,n_2^j}\right) \otimes I_{n_1,l}\right) \left(\beta_2^{[j]}(k) \otimes \beta_1^{[l]}(k)\right) \\
&= \left[\left(\left(I_{t_2, s-j} \otimes K_{t_1^{r-l}, n_2^j}^T\right) \left(\mu_{s,j}^2(k) \otimes \left(\left(\mu_{r,l}^1(k) \otimes I_{n_2,j}\right) \left(1 \otimes K_{1,n_2^j}\right)\right)\right)\right) \otimes I_{n_1,l}\right] \\
&\quad \times \left(\beta_2^{[j]}(k) \otimes \beta_1^{[l]}(k)\right) \\
&= \left[\left(I_{t_2, s-j} \otimes K_{t_1^{r-l}, n_2^j}^T\right) \left(\mu_{s,j}^2(k) \otimes \mu_{r,l}^1(k) \otimes K_{1,n_2^j}\right) \otimes I_{n_1,l}\right] \left(\beta_2^{[j]}(k) \otimes \beta_1^{[l]}(k)\right) \\
&= \left[\left(I_{t_2, s-j} \otimes K_{t_1^{r-l}, n_2^j}^T \otimes I_{n_1,l}\right) \left(\mu_{s,j}^2(k) \otimes \mu_{r,l}^1(k) \otimes K_{1,n_2^j} \otimes I_{n_1,l}\right)\right] \left(\beta_2^{[j]}(k) \otimes \beta_1^{[l]}(k)\right).
\end{aligned}$$

Por tanto,

$$\begin{aligned}
P_{l,j}^{r,s}(\mu_1, \mu_2) &= vec^{-1}\left[\left(I_{t_2, s-j} \otimes K_{t_1^{r-l}, n_2^j}^T \otimes I_{n_1,l}\right) \right. \\
&\quad \times \left.\left(\left(E\left\{\mu_2^{[s-j]}(k) \otimes \mu_1^{[r-l]}(k)\right\} - E\left\{\mu_2^{[s-j]}(k)\right\} \otimes E\left\{\mu_1^{[r-l]}(k)\right\}\right) \right. \\
&\quad \left.\left.\otimes K_{1,n_2^j} \otimes I_{n_1,l}\right) E\left\{\beta_2^{[j]}(k) \otimes \beta_1^{[l]}(k)\right\}\right]
\end{aligned}$$

donde se ha utilizado que $(\mu_1(k), \mu_2(k))$ es independiente de $(\beta_1(k), \beta_2(k))$. \blacksquare

Nota 3.4.2 De la demostración del Lema 3.4.2, se deduce que la hipótesis de independencia (ii) se puede debilitar, basta suponer que

- $\mu_a(k)$ es independiente de $(\beta_a(k), \{\beta_b(j), j < k\}, \{\mu_b(j), j < k\})$, $a, b \in \{1, 2\}$, $a \neq b$.
- $(\mu_1(k), \mu_2(k))$ es independiente de $(\beta_1(k), \beta_2(k))$.

Nota 3.4.3 Si, además de las condiciones del Lema 3.4.2, se verifica que $\mu_1(k)$ es independiente de $\mu_2(k)$ para cada k , entonces

$$E\left\{\mu_2^{[s-j]}(k) \otimes \mu_1^{[r-l]}(k)\right\} = E\left\{\mu_2^{[s-j]}(k)\right\} \otimes E\left\{\mu_1^{[r-l]}(k)\right\}.$$

Por tanto, sustituyendo en (3.9), $P_{l,j}^{r,s}(\mu_1, \mu_2) = 0$ y, consecuentemente, $S(k) = 0$.

3.5 Sistema Aumentado. Algoritmo de Filtrado Lineal

El Lema 3.4.1 nos permite estudiar la evolución de los vectores $X(k)$ y $Z(k)$ definidos en (3.3) y (3.4), respectivamente. Dicha evolución puede ser descrita mediante un sistema con observaciones inciertas, como establece el siguiente resultado.

Teorema 3.5.1 *Los procesos $\{X(k); k \geq 0\}$ y $\{Z(k); k \geq 0\}$ definidos en (3.3) y (3.4), respectivamente, satisfacen las siguientes ecuaciones*

$$X(k+1) = \mathcal{A}(k)X(k) + U_k + \mathcal{F}(k), \quad k \geq 0 \quad (3.10)$$

$$X(0) = X_0$$

$$Z(k) = u(k)\mathcal{C}(k)X(k) + V_k + \mathcal{G}(k), \quad k \geq 0 \quad (3.11)$$

donde

$$\mathcal{A}(k) = \begin{pmatrix} A(k) & 0 & \cdots & 0 \\ H_{21}(k) & A^{[2]}(k) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ H_{\nu 1}(k) & H_{\nu 2}(k) & \cdots & A^{[\nu]}(k) \end{pmatrix} \quad U_k = \begin{pmatrix} 0 \\ F^{[2]}(k) E\{w^{[2]}(k)\} \\ \vdots \\ F^{[\nu]}(k) E\{w^{[\nu]}(k)\} \end{pmatrix}$$

$$\mathcal{C}(k) = \begin{pmatrix} C(k) & 0 & \cdots & 0 \\ L_{21}(k) & C^{[2]}(k) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ L_{\nu 1}(k) & L_{\nu 2}(k) & \cdots & C^{[\nu]}(k) \end{pmatrix} \quad V_k = \begin{pmatrix} 0 \\ G^{[2]}(k) E\{v^{[2]}(k)\} \\ \vdots \\ G^{[\nu]}(k) E\{v^{[\nu]}(k)\} \end{pmatrix}$$

con

$$H_{jl}(k) = M_{j-l}^j(n) \left(F^{[j-l]}(k) \otimes A^{[l]}(k) \right) \left(E\{w^{[j-l]}(k)\} \otimes I_{n,l} \right)$$

$$L_{jl}(k) = M_{j-l}^j(n) \left(G^{[j-l]}(k) \otimes C^{[l]}(k) \right) \left(E\{v^{[j-l]}(k)\} \otimes I_{n,l} \right).$$

Los ruidos $\{\mathcal{F}(k); k \geq 0\}$ y $\{\mathcal{G}(k); k \geq 0\}$ vienen dados por

$$\mathcal{F}(k) = \begin{pmatrix} f_1(k) \\ f_2(k) \\ \vdots \\ f_\nu(k) \end{pmatrix} \quad \mathcal{G}(k) = \begin{pmatrix} g_1(k) \\ g_2(k) \\ \vdots \\ g_\nu(k) \end{pmatrix}$$

con

$$f_j(k) = \sum_{l=0}^{j-1} M_{j-l}^j(n) \left(F^{[j-l]}(k) \otimes A^{[l]}(k) \right) \left(\left(w^{[j-l]}(k) - E \{ w^{[j-l]}(k) \} \right) \otimes I_{n,l} \right) x^{[l]}(k)$$

$$g_j(k) = u(k) \sum_{l=1}^{j-1} M_{j-l}^j(m) \left(G^{[j-l]}(k) \otimes C^{[l]}(k) \right) \left(\left(v^{[j-l]}(k) - E \{ v^{[j-l]}(k) \} \right) \otimes I_{n,l} \right) x^{[l]}(k) \\ + G^{[j]}(k) \left(v^{[j]}(k) - E \{ v^{[j]}(k) \} \right).$$

-Demostración-

La ecuación del estado aumentado (3.10) se obtiene de forma inmediata, sin más que aplicar el Lema 3.4.1 a la ecuación (3.1), es decir, tomando

$$\begin{aligned} \alpha(k) &= x(k+1) & \beta(k) &= x(k) \\ \gamma(k) &= 1 & \Psi(k) &= F(k) \\ \Gamma(k) &= A(k) & \mu(k) &= w(k) \\ & (m = n; \quad n = n; \quad t = u). \end{aligned}$$

Para obtener la ecuación de la observación aumentada (3.11), aplicamos el Lema 3.4.1 a la ecuación (3.2), tomando

$$\begin{aligned} \alpha(k) &= z(k) & \beta(k) &= x(k) \\ \gamma(k) &= u(k) & \Psi(k) &= G(k) \\ \Gamma(k) &= C(k) & \mu(k) &= v(k) \\ & (m = m; \quad n = n; \quad t = t). \end{aligned}$$

■

Las ecuaciones (3.10) y (3.11) representan un sistema con observaciones inciertas en el que la incertidumbre está modelizada por el mismo ruido multiplicativo, $\{u(k); k \geq 0\}$, del sistema original.

Con el fin de determinar si puede aplicarse a este sistema algún algoritmo recursivo para la obtención del filtro lineal, estudiamos a continuación las propiedades del estado inicial y los ruidos.

Claramente, el estado inicial, X_0 , es un vector aleatorio con momento de segundo orden finito, en virtud de la hipótesis 3 impuesta sobre el modelo.

Las propiedades de los ruidos $\{\mathcal{F}(k)\}$ y $\{\mathcal{G}(k)\}$ y el algoritmo que nos permitirá obtener el filtro lineal se establecen a continuación. En primer lugar, supondremos que los ruidos aditivos del estado y la observación del sistema original son independientes (Sección 3.5.1) y, posteriormente, generalizaremos el estudio al caso en que dichos ruidos son correlados entre sí (Sección 3.5.2).

3.5.1 Caso de Ruidos Aditivos Independientes

En esta Sección vamos a tratar el caso en que la condición inicial y los ruidos del sistema original son independientes. Como acabamos de indicar, nuestro objetivo es el estudio de las propiedades de los procesos ruido del correspondiente sistema aumentado, con el fin de establecer un algoritmo de filtrado lineal para el mismo.

Propiedades de los Ruidos del Sistema Aumentado

Consideremos el sistema (3.1)-(3.2) y supongamos que se verifica la siguiente hipótesis de independencia:

$x_0, \{w(k); k \geq 0\}, \{v(k); k \geq 0\}, \{u(k); k \geq 0\}$ son mutuamente independientes.

En este caso, el sistema aumentado (3.10)-(3.11) verifica las hipótesis requeridas para aplicar el filtro de Nahi (Capítulo 1, Sección 1.4.3), como se establece en los

siguientes resultados.

Teorema 3.5.2 *Los ruidos $\{\mathcal{F}(k); k \geq 0\}$ y $\{\mathcal{G}(k); k \geq 0\}$ son procesos blancos, centrados, con matrices de covarianzas $\mathcal{Q}(k)$ y $\mathcal{R}(k)$, respectivamente, cuyos bloques (r, s) vienen expresados por*

$$(I) \quad \mathcal{Q}_{rs}(k) = E \{f_r(k) f_s^T(k)\}$$

$$= \sum_{l=0}^{r-1} \sum_{i=0}^{s-1} M_{r-l}^r(n) (F^{[r-l]}(k) \otimes A^{[l]}(k)) P_{l,i}^{r,s}(w, w) (F^{[s-i]}(k) \otimes A^{[i]}(k))^T (M_{s-i}^s(n))^T \quad (3.12)$$

donde

$$P_{l,i}^{r,s}(w, w) = \text{vec}^{-1} \left[\left(I_{u, s-i} \otimes K_{ur-l, ni}^T \otimes I_{n, l} \right) \times \left(\left(E \{w^{[r+s-l-i]}(k)\} - E \{w^{[s-i]}(k)\} \otimes E \{w^{[r-l]}(k)\} \right) \otimes K_{1, ni} \otimes I_{n, l} \right) E \{x^{[l+i]}(k)\} \right].$$

$$(II) \quad \mathcal{R}_{rs}(k) = E \{g_r(k) g_s^T(k)\}$$

$$= p(k) \left[\sum_{l=1}^{r-1} \sum_{i=0}^{s-1} M_{r-l}^r(m) (G^{[r-l]}(k) \otimes C^{[l]}(k)) P_{l,i}^{r,s}(v, v) (G^{[s-i]}(k) \otimes C^{[i]}(k))^T (M_{s-i}^s(m))^T \right. \\ \left. + G^{[r]}(k) \sum_{i=1}^{s-1} P_{0,i}^{r,s}(v, v) (G^{[s-i]}(k) \otimes C^{[i]}(k))^T (M_{s-i}^s(m))^T \right] \\ + G^{[r]}(k) P_{0,0}^{r,s}(v, v) G^{[s]T}(k). \quad (3.13)$$

donde

$$P_{l,i}^{r,s}(v, v) = \text{vec}^{-1} \left[\left(I_{t, s-i} \otimes K_{tr-l, ni}^T \otimes I_{n, l} \right) \times \left(\left(E \{v^{[r+s-l-i]}(k)\} - E \{v^{[s-i]}(k)\} \otimes E \{v^{[r-l]}(k)\} \right) \otimes K_{1, ni} \otimes I_{n, l} \right) E \{x^{[l+i]}(k)\} \right].$$

Además, $E \{ \mathcal{F}(k) \mathcal{G}^T(j) \} = 0$, para todo k y todo j .

-Demostración-

Para obtener (3.12) aplicamos el Lema 3.4.2 a los vectores

$$\begin{aligned} \alpha_1(k) &= \alpha_2(k) = x(k+1) & \beta_1(k) &= \beta_2(k) = x(k) \\ \gamma_1(k) &= \gamma_2(k) = 1 & \Psi_1(k) &= \Psi_2(k) = F(k) \\ \Gamma_1(k) &= \Gamma_2(k) = A(k) & \mu_1(k) &= \mu_2(k) = w(k) \\ & & (m_i &= n, \quad n_i = n, \quad t_i = u; \quad i \in \{1, 2\}). \end{aligned}$$

Previamente, comprobamos que se verifican las hipótesis de dicho Lema:

- (i) es cierta de forma trivial, al ser $\gamma_1(k) = \gamma_2(k) = 1$.
- (ii) $w(k)$ es independiente de $(\{x(j), j \leq k\}, \{w(j), j < k\})$, ya que $\{w(k)\}$ es blanco y $x_0, \{w(k)\}$ son independientes.

Aplicando el Lema 3.4.2 a los vectores anteriores se deduce que $\{\mathcal{F}(k)\}$ es blanco y centrado. La expresión (3.12) se obtiene de (3.8) teniendo en cuenta que $E \{ \gamma_1(k) \gamma_2(k) \} = p_1(k) = p_2(k) = 1$.

Para obtener (3.13), aplicamos el Lema 3.4.2 a los vectores

$$\begin{aligned} \alpha_1(k) &= \alpha_2(k) = z(k) & \beta_1(k) &= \beta_2(k) = x(k) \\ \gamma_1(k) &= \gamma_2(k) = u(k) & \Psi_1(k) &= \Psi_2(k) = G(k) \\ \Gamma_1(k) &= \Gamma_2(k) = C(k) & \mu_1(k) &= \mu_2(k) = v(k) \\ & & (m_i &= m, \quad n_i = n, \quad t_i = t; \quad i \in \{1, 2\}). \end{aligned}$$

comprobando, en primer lugar, que se verifican las hipótesis requeridas para aplicar dicho lema:

- (i) $\{u(k)\}$ es independiente de $(\{x(k)\}, \{v(k)\})$, ya que, por hipótesis, $\{u(k)\}$ es independiente de $(x_0, \{w(k)\}, \{v(k)\})$.

- (ii) $v(k)$ es independiente de $(\{x(j), j \leq k\}, \{v(j), j < k\})$, ya que $\{v(k)\}$ es blanco y $x_0, \{w(k)\}, \{v(k)\}$ son mutuamente independientes.

Aplicando el Lema 3.4.2 a los vectores anteriores se deduce que $\{\mathcal{G}(k)\}$ es blanco y centrado. La expresión (3.13) se obtiene de forma inmediata, teniendo en cuenta que $E\{\gamma_1(k)\gamma_2(k)\} = p_1(k) = p_2(k) = p(k)$ y agrupando los sumandos de (3.8) de forma conveniente.

Por último, veamos que $E\{\mathcal{F}(k)\mathcal{G}^T(j)\} = 0$, para todo k y todo j . Para ello utilizamos la Nota 3.4.3, comprobando las hipótesis del Lema 3.4.2 para los vectores

$$\begin{aligned} \alpha_1(k) &= x(k+1) & \beta_1(k) &= x(k) \\ \gamma_1(k) &= 1 & \Psi_1(k) &= F(k) \\ \Gamma_1(k) &= A(k) & \mu_1(k) &= w(k) \\ & (m_1 = n; \quad n_1 = n; \quad t_1 = u) \\ \alpha_2(k) &= z(k) & \beta_2(k) &= x(k) \\ \gamma_2(k) &= u(k) & \Psi_2(k) &= G(k) \\ \Gamma_2(k) &= C(k) & \mu_2(k) &= v(k) \\ & (m_2 = m; \quad n_2 = n; \quad t_2 = t) \end{aligned}$$

- (i) $\{u(k)\}$ es independiente de $(\{x(k)\}, \{w(k)\}, \{v(k)\})$, en virtud de la hipótesis de independencia mutua impuesta sobre el modelo.
- (ii) $(w(k), v(k))$ es independiente de $(\{x(j), j \leq k\}, \{w(j), j < k\}, \{v(j), j < k\})$, por ser $(x_0, \{w(k)\}, \{v(k)\})$ mutuamente independientes.

Además, $w(k)$ es independiente de $v(k)$ y, por tanto, en virtud del Lema 3.4.2 y la Nota 3.4.3, se tiene que $E\{\mathcal{F}(k)\mathcal{G}^T(j)\} = 0$, para todo k y todo j . ■

Teorema 3.5.3 *El estado inicial, X_0 , del sistema aumentado es incorrelado con los ruidos $\{\mathcal{F}(k); k \geq 0\}$ y $\{\mathcal{G}(k); k \geq 0\}$.*

-Demostración-

Observemos que

$$\begin{aligned} E \left\{ f_r(k) x_0^{[s]T} \right\} &= \sum_{l=0}^{r-1} M_{r-l}^r(n) \left(F^{[r-l]}(k) \otimes A^{[l]}(k) \right) \\ &\quad \times E \left\{ \left(w^{[r-l]}(k) - E \left\{ w^{[r-l]}(k) \right\} \right) \otimes I_{n,l} \right\} E \left\{ x^{[l]}(k) x_0^{[s]T} \right\} \\ &= 0 \end{aligned}$$

ya que $w(k)$ es independiente de $(x_0, w(0), \dots, w(k-1))$ y, por consiguiente, el vector $(w^{[r-l]}(k) - E \{w^{[r-l]}(k)\})$ es independiente de $x^{[l]}(k) x_0^{[s]T}$.

Por otro lado,

$$\begin{aligned} E \left\{ g_r(k) x_0^{[s]T} \right\} &= p(k) \sum_{l=1}^{r-1} M_{r-l}^r(m) \left(G^{[r-l]}(k) \otimes C^{[l]}(k) \right) \\ &\quad \times E \left\{ \left(v^{[r-l]}(k) - E \left\{ v^{[r-l]}(k) \right\} \right) \otimes I_{n,l} \right\} E \left\{ x^{[l]}(k) x_0^{[s]T} \right\} \\ &\quad + G^{[r]}(k) E \left\{ \left(v^{[r]}(k) - E \left\{ v^{[r]}(k) \right\} \right) \right\} E \left\{ x_0^{[s]T} \right\} \\ &= 0 \end{aligned}$$

donde, además de la independencia de $\{u(k)\}$ y $(x_0, \{w(k)\}, \{v(k)\})$, se ha utilizado que el vector $v(k)$ es independiente de $(x_0, w(0), \dots, w(k-1))$ y, por tanto, $(v^{[r]}(k) - E \{v^{[r]}(k)\})$ es independiente de $x_0^{[s]T}$ y $(v^{[r-l]}(k) - E \{v^{[r-l]}(k)\})$ es independiente de $x^{[l]}(k) x_0^{[s]T}$. ■

Solución del Problema de Filtrado Lineal para el Sistema Aumentado

Como ya se ha indicado, el estado inicial X_0 del sistema aumentado es un vector aleatorio con momento de segundo orden finito, en virtud de la hipótesis 3 impuesta sobre el modelo original. Por otra parte, según se ha demostrado en los teoremas anteriores, los ruidos aditivos $\{\mathcal{F}(k)\}$ y $\{\mathcal{G}(k)\}$ son procesos blancos centrados con momentos de segundo orden finitos. Además, dichos ruidos son incorrelados entre sí

(Teorema 3.5.2) e incorrelados con el estado inicial (Teorema 3.5.3). A partir de estos resultados y de las hipótesis impuestas al sistema original, se deducen las siguientes propiedades:

- (a) Para cada k , $\mathcal{F}(k)$ es incorrelado con $Z(j)$ para $j \leq k$, ya que $\mathcal{F}(k)$ es incorrelado con $X_0, \mathcal{F}(0), \dots, \mathcal{F}(k-1), u(0), \dots, u(k-1), \mathcal{G}(0), \dots, \mathcal{G}(k-1)$.
- (b) La hipótesis de independencia mutua impuesta sobre el sistema original (3.1)-(3.2), garantiza que $u(k)$ y $(X(k), Z(0), \dots, Z(k-1))$ son independientes para todo k .

Como consecuencia de estas propiedades, el estimador lineal de menor error cuadrático medio, $\hat{X}(k/k)$, del estado del sistema aumentado, $X(k)$, basado en las observaciones $\{Z(0), \dots, Z(k)\}$, se puede obtener mediante el Filtro de Nahi (Capítulo 1, Sección 1.4.3) generalizado, al caso en que las ecuaciones del estado y la observación están afectadas por los vectores determinísticos U_k y V_k , respectivamente. Esto afectará a la expresión de la innovación y del predictor, que ahora vendrán dados por

$$\begin{aligned}\tilde{Z}(k/k-1) &= Z(k) - p(k)C(k)\hat{X}(k/k-1) - V_k, \quad k \geq 0 \\ \hat{X}(k+1/k) &= A(k)\hat{X}(k/k) + U_k, \quad k \geq 0.\end{aligned}$$

Observemos, además, que el estado inicial no es centrado ($E\{X_0\} = \bar{X}_0$), lo cual afectará a las condiciones iniciales del algoritmo. Concretamente, la solución al problema de filtrado lineal para el sistema aumentado en el caso que nos ocupa, se establece en el siguiente Teorema.

Teorema 3.5.4 *El algoritmo para los problemas de filtrado y predicción del estado, $X(k)$, del sistema aumentado está dado por*

$$\hat{X}(k/k) = \hat{X}(k/k-1) + K(k)\tilde{Z}(k/k-1), \quad k \geq 0$$

$$\begin{aligned}\hat{X}(k+1/k) &= \mathcal{A}(k)\hat{X}(k/k) + U_k, \quad k \geq 0 \\ \hat{X}(0/-1) &= \bar{X}_0\end{aligned}$$

siendo $\tilde{Z}(k/k-1)$ la innovación

$$\tilde{Z}(k/k-1) = Z(k) - p(k)C(k)\hat{X}(k/k-1) - V_k, \quad k \geq 0.$$

La matriz de ganancia, $K(k)$, está dada por

$$K(k) = p(k)P(k/k-1)C^T(k)\Pi^-(k), \quad k \geq 0.$$

Las matrices de covarianzas de los errores de filtrado y predicción se calculan mediante la siguiente relación recursiva

$$\begin{aligned}P(k/k) &= P(k/k-1) - K(k)\Pi(k)K^T(k), \quad k \geq 0 \\ P(k+1/k) &= \mathcal{A}(k)P(k/k)\mathcal{A}^T(k) + \mathcal{Q}(k) \\ P(0/-1) &= P_0 = E\{(X_0 - \bar{X}_0)(X_0 - \bar{X}_0)^T\}.\end{aligned}$$

siendo $\Pi(k)$ la matriz de covarianzas del proceso innovación,

$$\Pi(k) = p(k)(1-p(k))C(k)\mathcal{D}(k)C^T(k) + p^2(k)C(k)P(k/k-1)C^T(k) + \mathcal{R}(k),$$

donde $\mathcal{D}(k) = E\{X(k)X^T(k)\}$ se puede obtener de forma recursiva mediante las relaciones siguientes

$$\begin{aligned}\mathcal{D}(k) &= \mathcal{A}(k-1)\mathcal{D}(k-1)\mathcal{A}^T(k-1) + U_{k-1}U_{k-1}^T + \mathcal{Q}(k-1) \\ &\quad + \mathcal{A}(k-1)E\{X(k-1)\}U_{k-1}^T + U_{k-1}E\{X^T(k-1)\}\mathcal{A}^T(k-1) \\ \mathcal{D}(0) &= E\{X_0X_0^T\} = P_0 + \bar{X}_0\bar{X}_0^T\end{aligned}$$

con

$$\begin{aligned}E\{X(k)\} &= \mathcal{A}(k-1)E\{X(k-1)\} + U_{k-1} \\ E\{X(0)\} &= \bar{X}_0.\end{aligned}$$

El Teorema 3.5.7 permite calcular el estimador lineal de menor error cuadrático medio del estado, $X(k)$, del sistema aumentado. Una vez obtenido dicho estimador lineal, las n primeras componentes del mismo proporcionan el filtro polinomial de grado ν de menor error cuadrático medio del estado, $x(k)$, del sistema original, estableciendo de esta forma la solución del problema planteado.

3.5.2 Caso de Ruidos Aditivos Correlados

El objetivo de esta Sección es el estudio de las propiedades del sistema aumentado en el caso en que los ruidos aditivos del sistema original son correlados entre sí, con el fin de establecer un algoritmo de filtrado lineal para dicho sistema aumentado.

Propiedades de los Ruidos del Sistema Aumentado

Consideremos el sistema (3.1)-(3.2) y supongamos que, además de las hipótesis 1-4 impuestas en la Sección 3.3, se verifica que

- (a) Para cada k , $w(k)$ es correlado con $v(k)$.
- (b) $\{u(k); k \geq 0\}$ es independiente de $(x_0, \{w(k); k \geq 0\}, \{v(k); k \geq 0\})$.
- (c) $(w(k), v(k))$ es independiente de $(x_0, w(0), \dots, w(k-1), v(0), \dots, v(k-1))$.

En este caso, las propiedades de los ruidos $\{\mathcal{F}(k); k \geq 0\}$ y $\{\mathcal{G}(k); k \geq 0\}$ del sistema aumentado se establecen en los siguientes resultados.

Teorema 3.5.5 *Los ruidos $\{\mathcal{F}(k); k \geq 0\}$ y $\{\mathcal{G}(k); k \geq 0\}$ son procesos blancos, centrados, con matrices de covarianzas $\mathcal{Q}(k)$ y $\mathcal{R}(k)$, respectivamente, cuyos bloques (r, s) vienen dados por las expresiones (3.12) y (3.13), establecidas en el Teorema 3.5.2.*

Además, $E \{ \mathcal{F}(k) \mathcal{G}^T(j) \} = \mathcal{S}(k) \delta_{kj}$, y el bloque (r, s) de $\mathcal{S}(k)$ está dado por

$$\begin{aligned} \mathcal{S}_{rs}(k) &= E \{ f_r(k) g_s^T(k) \} \\ &= p(k) \sum_{l=0}^{r-1} \sum_{i=1}^{s-1} M_{r-l}^r(n) (F^{[r-l]}(k) \otimes A^{[l]}(k)) P_{l,i}^{r,s}(w, v) (G^{[s-i]}(k) \otimes C^{[i]}(k))^T (M_{s-i}^s(m))^T \\ &\quad + \sum_{l=0}^{r-1} M_{r-l}^r(n) (F^{[r-l]}(k) \otimes A^{[l]}(k)) P_{l,0}^{r,s}(w, v) G^{[s]T}(k) \end{aligned} \quad (3.14)$$

donde

$$\begin{aligned} P_{l,i}^{r,s}(w, v) &= \text{vec}^{-1} \left[\left(I_{t,s-i} \otimes K_{u^{r-l}, n^i}^T \otimes I_{n,l} \right) \right. \\ &\quad \times \left(\left(E \{ v^{[s-i]}(k) \otimes w^{[r-l]}(k) \} - E \{ v^{[s-i]}(k) \} \otimes E \{ w^{[r-l]}(k) \} \right) \right. \\ &\quad \left. \left. \otimes K_{1, n^i} \otimes I_{n,l} \right) E \{ x^{[l+i]}(k) \} \right]. \end{aligned}$$

-Demostración-

Para obtener (3.12) y (3.13), se procede como en la demostración del Teorema 3.5.2, teniendo en cuenta que las condiciones necesarias para aplicar el Lema 3.4.2 se verifican también en este caso, como consecuencia de las hipótesis (b) y (c) impuestas sobre el sistema original.

Para completar la demostración del teorema, debemos obtener la expresión (3.14). Con tal fin, tomamos

$$\begin{aligned} \alpha_1(k) &= x(k+1) & \beta_1(k) &= x(k) \\ \gamma_1(k) &= 1 & \Psi_1(k) &= F(k) \\ \Gamma_1(k) &= A(k) & \mu_1(k) &= w(k) \\ & & (m_1 = n; \quad n_1 = n; \quad t_1 = u) & \end{aligned}$$

$$\begin{aligned}
 \alpha_2(k) &= z(k) & \beta_2(k) &= x(k) \\
 \gamma_2(k) &= u(k) & \Psi_2(k) &= G(k) \\
 \Gamma_2(k) &= C(k) & \mu_2(k) &= v(k) \\
 & & (m_2 = m; \quad n_2 = n; \quad t_2 = t)
 \end{aligned}$$

y aplicamos el Lema 3.4.2 a estos vectores. Previamente, comprobamos que se dan las condiciones necesarias:

- (i) $\{u(k)\}$ es independiente de $(\{x(k)\}, \{w(k)\}, \{v(k)\})$, en virtud de la hipótesis (b) impuesta sobre el modelo.
- (ii) $(w(k), v(k))$ es independiente de $(\{x(j), j \leq k\}, \{w(j), j < k\}, \{v(j), j < k\})$, por ser $(w(k), v(k))$ y $(x_0, w(0), \dots, w(k-1), v(0), \dots, v(k-1))$ independientes (hipótesis (c)).

Aplicando el Lema 3.4.2 a los vectores anteriores se deduce que $E\{\mathcal{F}(k)\mathcal{G}^T(j)\} = 0$, para todo $k \neq j$.

Por último, la expresión (3.14) se obtiene de forma inmediata a partir de (3.8), sin más que agrupar los sumandos de forma conveniente y tener en cuenta que $p_1(k) = 1$, $p_2(k) = p(k)$, $E\{\gamma_1(k)\gamma_2(k)\} = p(k)$, concluyendo así la demostración del teorema. ■

Nota 3.5.1 En virtud de la Nota 3.4.2, la hipótesis (c) de independencia impuesta sobre el sistema (3.1)-(3.2) se puede debilitar, basta suponer que

- $w(k)$ y $v(k)$ son independientes de $(x_0, w(0), \dots, w(k-1), v(0), \dots, v(k-1))$
- $(w(k), v(k))$ es independiente de $(x_0, w(0), \dots, w(k-1))$.

Finalmente, observemos que las propiedades de independencia utilizadas en la demostración del Teorema 3.5.3, siguen siendo ciertas en este caso. Por tanto, podemos enunciar el siguiente resultado.

Teorema 3.5.6 *El estado inicial, X_0 , del sistema aumentado es incorrelado con los ruidos $\{\mathcal{F}(k); k \geq 0\}$ y $\{\mathcal{G}(k); k \geq 0\}$.*

Solución del Problema de Filtrado Lineal para el Sistema Aumentado

Como se ha indicado anteriormente, el estado inicial X_0 del sistema aumentado es un vector aleatorio con momento de segundo orden finito. Además, según acabamos de demostrar, los ruidos aditivos $\{\mathcal{F}(k)\}$ y $\{\mathcal{G}(k)\}$ son procesos blancos centrados con momentos de segundo orden finitos y dichos ruidos son incorrelados con el estado inicial.

Sin embargo, $E\{\mathcal{F}(k)\mathcal{G}^T(k)\} \neq 0$ y, por consiguiente, no podemos aplicar el filtro de Nahi (Capítulo 1, Sección 1.4.3) al sistema aumentado, ya que este filtro requiere que los ruidos del estado y la observación sean incorrelados entre sí. Por otra parte, el filtro para sistemas con observaciones inciertas en presencia de ruidos correlados propuesto por Hermoso y Linares [15] tampoco es aplicable en este caso, ya que supone que los ruidos del estado y la observación están correlados en instantes consecutivos, y no en el mismo instante, como ocurre en el caso que nos ocupa.

Por esta razón, nos planteamos obtener un algoritmo recursivo para el problema de filtrado lineal de menor error cuadrático medio que pueda aplicarse al sistema aumentado (3.10)-(3.11), suponiendo que los ruidos y la condición inicial satisfacen las propiedades establecidas en los Teoremas 3.5.5 y 3.5.6.

En el Apéndice C se presenta el algoritmo que proponemos, así como su deducción basada en la interpretación geométrica de los estimadores lineales que fue presentada en el Capítulo 2 (Sección 2.2). Hay que indicar que, si bien el sistema aumentado no satisface la hipótesis de independencia impuesta al sistema (C.1)-(C.2) considerado en el Apéndice C, dicho sistema aumentado sí verifica las propiedades establecidas en las Proposiciones C.1-C.3, que son suficientes para aplicar el algoritmo. En efecto:

1. Teniendo en cuenta los teoremas 3.5.5, 3.5.6 y las hipótesis impuestas al sistema

original, es inmediato que, para cada k , los ruidos $\mathcal{F}(k)$ y $\mathcal{G}(k)$ son incorrelados con $X(0), \dots, X(k)$ y $Z(0), \dots, Z(k-1)$.

2. Para cada k , $\mathcal{G}(k)$ es incorrelado con el producto $u(j)X(j)$ con $j \leq k$, ya que

$$\begin{aligned} E\{u(j)g_r(k)x^{[s]T}(j)\} &= E\{u(j)u(k)\} \sum_{l=1}^{r-1} M_{r-l}^r(m) (G^{[r-l]}(k) \otimes C^{[l]}(k)) \\ &\quad \times E\{(v^{[r-l]}(k) - E\{v^{[r-l]}(k)\}) \otimes I_{n,l}\} E\{x^{[l]}(k)x^{[s]T}(j)\} \\ &\quad + p(k)G^{[r]}(k)E\{(v^{[r]}(k) - E\{v^{[r]}(k)\})\} E\{x^{[s]T}(j)\} \\ &= 0 \end{aligned}$$

donde se ha utilizado que $\{u(k)\}$ es independiente de $(x_0, \{w(k)\}, \{v(k)\})$ y que $v(k)$ es independiente de $(x_0, w(0), \dots, w(k-1))$.

3. Para cada k , la variable $u(k)$ es independiente de $(X(k), \mathcal{F}(k))$ e incorrelada con $X(k)Z^T(j)$ para $j < k$. En efecto, al ser $\{u(k)\}$ una sucesión de variables aleatorias independientes y, en virtud de la hipótesis (b) impuesta sobre el sistema original, se deduce que $u(k)$ es independiente de $(X_0, \{\mathcal{F}(k)\})$ e incorrelada con $\mathcal{G}(0), \dots, \mathcal{G}(k-1)$ para todo k .

Por consiguiente, teniendo en cuenta los resultados presentados en el Apéndice C, obtenemos el siguiente algoritmo para el estimador lineal de menor error cuadrático medio, $\hat{X}(k/k)$, del estado aumentado $X(k)$, basado en las observaciones $\{Z(0), \dots, Z(k)\}$.

Teorema 3.5.7 *El algoritmo para los problemas de filtrado y predicción del estado, $X(k)$, del sistema aumentado está dado por*

$$\begin{aligned} \hat{X}(k/k) &= \hat{X}(k/k-1) + K(k)\tilde{Z}(k/k-1), \quad k \geq 0 \\ \hat{X}(k+1/k) &= \mathcal{A}(k)\hat{X}(k/k-1) + U_k + [\mathcal{A}(k)K(k) + \mathcal{S}(k)\Pi^-(k)]\tilde{Z}(k/k-1) \\ \hat{X}(0/-1) &= \bar{X}_0 \end{aligned}$$

siendo $\tilde{Z}(k/k-1)$ la innovación

$$\tilde{Z}(k/k-1) = Z(k) - p(k)C(k)\hat{X}(k/k-1) - V_k, \quad k \geq 0.$$

La matriz de ganancia, $K(k)$, está dada por

$$K(k) = p(k)P(k/k-1)C^T(k)\Pi^-(k), \quad k \geq 0.$$

Las matrices de covarianzas de los errores de filtrado y predicción se calculan mediante la siguiente relación recursiva

$$\begin{aligned} P(k/k) &= P(k/k-1) - K(k)\Pi(k)K^T(k), \quad k \geq 0 \\ P(k+1/k) &= A(k)P(k/k)A^T(k) + Q(k) - S(k)\Pi^-(k)S^T(k) \\ &\quad - A(k)K(k)S^T(k) - S(k)K^T(k)A^T(k) \\ P(0/-1) &= P_0. \end{aligned}$$

siendo $\Pi(k)$ la matriz de covarianzas del proceso innovación,

$$\Pi(k) = p(k)(1-p(k))C(k)D(k)C^T(k) + p^2(k)C(k)P(k/k-1)C^T(k) + \mathcal{R}(k),$$

donde $D(k) = E\{X(k)X^T(k)\}$ se puede obtener de forma recursiva mediante las relaciones siguientes

$$\begin{aligned} D(k) &= A(k-1)D(k-1)A^T(k-1) + U_{k-1}U_{k-1}^T + Q(k-1) \\ &\quad + A(k-1)E\{X(k-1)\}U_{k-1}^T + U_{k-1}E\{X^T(k-1)\}A^T(k-1) \\ D(0) &= E\{X_0X_0^T\} = P_0 + \bar{X}_0\bar{X}_0^T \end{aligned}$$

con

$$\begin{aligned} E\{X(k)\} &= A(k-1)E\{X(k-1)\} + U_{k-1} \\ E\{X(0)\} &= \bar{X}_0. \end{aligned}$$

Una vez obtenido el estimador lineal de menor error cuadrático medio del estado, $X(k)$, del sistema aumentado, las n primeras componentes del mismo proporcionan el filtro polinomial de grado ν óptimo del estado, $x(k)$, del sistema original.

3.6 Sistemas Estacionarios

Como se ha comentado en los capítulos anteriores, una de las propiedades más interesantes, tanto del filtro de Kalman como del filtro de Nahi, es el comportamiento asintótico de los estimadores y de las matrices de covarianza de los errores de estimación. Para sistemas estacionarios se puede utilizar la versión steady-state de estos filtros, que presenta enormes ventajas a la hora de calcular los estimadores.

En esta Sección consideramos la versión estacionaria del sistema con observaciones inciertas (3.1)-(3.2), tanto en el caso de ruidos independientes, como en el caso en que los ruidos del estado y la observación son correlados entre sí. Hay que indicar que, en general, los correspondientes sistemas aumentados no son estacionarios. Nuestro objetivo es establecer condiciones bajo las cuales estos sistemas aumentados sean asintóticamente estacionarios (esto es, en el límite, las matrices del sistema sean constantes y los ruidos sean estacionarios de segundo orden) con el fin de analizar, en la Sección 3.7, si los filtros lineales que se han obtenido para dichos sistemas aumentados admiten una forma steady-state.

La versión estacionaria del sistema con observaciones inciertas (3.1)-(3.2) viene descrita por las ecuaciones siguientes

$$x(k+1) = Ax(k) + Fw(k), \quad k \geq 0 \quad (3.15)$$

$$x(0) = x_0$$

$$z(k) = u(k)Cx(k) + Gv(k), \quad k \geq 0 \quad (3.16)$$

donde A, F, C, G , son matrices determinísticas constantes y $\{w(k); k \geq 0\}, \{v(k); k \geq 0\}$ procesos blancos centrados verificando

$$\begin{aligned} E \left\{ w^{[i]}(k) w^{[j]T}(k) \right\} &= Q^{(i+j)}, \quad i \leq j, \quad j = 1, \dots, \nu \\ E \left\{ v^{[i]}(k) v^{[j]T}(k) \right\} &= R^{(i+j)}, \quad i \leq j, \quad j = 1, \dots, \nu. \end{aligned}$$

Además, suponemos que el ruido, $\{u(k); k \geq 0\}$, que modeliza la incertidumbre de las observaciones, es una sucesión de variables aleatorias de Bernoulli independientes e idénticamente distribuidas, con

$$P[u(k) = 1] = p, \quad k \geq 0.$$

Consideramos, como en la Sección anterior, tanto el caso de ruidos independientes, como el caso en que los ruidos aditivos del estado y la observación son correlados entre sí. En este último caso supondremos, además de las hipótesis previas, que las covarianzas cruzadas,

$$E \{w^{[i]}(k) v^{[j]T}(k)\} = S^{(i,j)}, \quad i, j = 1, \dots, \nu$$

no dependen de k .

Como en la Sección 3.5, definimos el sistema aumentado correspondiente al sistema estacionario anterior

$$X(k+1) = \mathcal{A}X(k) + U + \mathcal{F}(k), \quad k \geq 0 \quad (3.17)$$

$$X(0) = X_0$$

$$Z(k) = u(k)CX(k) + V + \mathcal{G}(k), \quad k \geq 0 \quad (3.18)$$

donde los vectores y matrices que intervienen están definidos en el Teorema 3.5.1. Observemos que las matrices \mathcal{A} , \mathcal{C} y los vectores U , V no dependen de k , en virtud de las hipótesis impuestas sobre el sistema (3.15)-(3.16).

Hay que notar que, tanto en el caso de ruidos independientes como en el caso de ruidos correlados, este sistema aumentado no es estacionario puesto que los ruidos aditivos $\{\mathcal{F}(k)\}$ y $\{\mathcal{G}(k)\}$ no lo son, ya que sus matrices de covarianzas, $\mathcal{Q}(k)$ y $\mathcal{R}(k)$, respectivamente, dependen de los vectores $E \{x^{[l+i]}(k)\}$, $l, i = 0, 1, \dots, \nu - 1$ (Teoremas 3.5.2 y 3.5.5) y, por tanto, de la evolución del estado.

No obstante, bajo la hipótesis de que el sistema original sea asintóticamente estable, es decir, la matriz A sea estrictamente estable (sus autovalores tengan módulo menor que la unidad), vamos a demostrar que estos vectores convergen a vectores constantes (Lema 3.6.1). Aplicando este resultado obtendremos que las matrices $\mathcal{Q}(k)$ y $\mathcal{R}(k)$ tienen límites constantes (Teorema 3.6.3) y, consecuentemente, los procesos $\{\mathcal{F}(k)\}$ y $\{\mathcal{G}(k)\}$ son asintóticamente estacionarios de segundo orden.

Lema 3.6.1 *Si la matriz A es estrictamente estable, el proceso estado $\{x(k); k \geq 0\}$ verifica*

$$\lim_{k \rightarrow \infty} E \{x^{[i]}(k)\} = m_i, \quad i = 1, \dots, 2\nu$$

siendo m_i la componente i -sima del vector

$$\mathcal{M}_{2\nu} = \left(\sum_{j=0}^{\infty} A_{2\nu}^j \right) U_{2\nu}$$

donde

$$\mathcal{A}_{2\nu} = \begin{pmatrix} A & 0 & \dots & 0 \\ H_{21} & A^{[2]} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ H_{2\nu-1} & H_{2\nu-2} & \dots & A^{[2\nu]} \end{pmatrix}, \quad U_{2\nu} = \begin{pmatrix} 0 \\ F^{[2]} \text{vec } Q^{(2)} \\ \vdots \\ F^{[2\nu]} \text{vec } Q^{(2\nu)} \end{pmatrix} \quad (3.19)$$

y las matrices H_{il} vienen dadas por

$$H_{il} = M_{i-l}^i(n) \left(F^{[i-l]} \otimes A^{[l]} \right) \left(\left(\text{vec } Q^{(i-l)} \right) \otimes I_{n,l} \right)$$

para $i = 2, \dots, 2\nu$, $l = 1, \dots, i-1$.

-Demostración-

Teniendo en cuenta la ecuación del estado (3.15), elevamos $x(k+1)$ a la i -sima potencia de Kronecker, $i = 2, \dots, 2\nu$, obteniendo

$$x^{[i]}(k+1) = A^{[i]} x^{[i]}(k) + \sum_{l=1}^{i-1} M_{i-l}^i(n) \left(F^{[i-l]} \otimes A^{[l]} \right) \left(w^{[i-l]}(k) \otimes I_{n,l} \right) x^{[l]}(k)$$

$$+F^{[i]}w^{[i]}(k) \quad (3.20)$$

donde se ha hecho uso del Lema A.4.1 y de las propiedades del producto Kronecker (ver Apéndice A).

Definiendo los vectores

$$m_i(k) := E \left\{ x^{[i]}(k) \right\}, \quad k \geq 0, \quad i = 2, \dots, 2\nu$$

y utilizando la expresión (3.20), obtenemos

$$m_i(k+1) = A^{[i]}m_i(k) + \sum_{l=1}^{i-1} H_{il}m_l(k) + F^{[i]}vec Q^{(i)}$$

donde se ha utilizado que $w(k)$ es independiente de $(x_0, w(0), \dots, w(k-1))$ (hipótesis que se verifica tanto en el caso de ruidos independientes, como en el caso de ruidos correlados) y que $u \otimes v = vec(vu^T)$ (ver Apéndice A).

Definiendo, ahora, el vector

$$m(k) := \begin{pmatrix} 0 \\ m_2(k) \\ \vdots \\ m_{2\nu}(k) \end{pmatrix},$$

podemos escribir

$$m(k+1) = \mathcal{A}_{2\nu}m(k) + U_{2\nu}$$

donde $\mathcal{A}_{2\nu}$ y $U_{2\nu}$ están definidos en (3.19).

Reiterando la expresión anterior, tenemos que

$$m(k+1) = \mathcal{A}_{2\nu}^{k+1}m(0) + \left(\sum_{i=0}^k \mathcal{A}_{2\nu}^i \right) U_{2\nu}$$

y, haciendo $k \rightarrow \infty$, concluimos que

$$\lim_{k \rightarrow \infty} m(k) = \left(\sum_{i=0}^{\infty} \mathcal{A}_{2\nu}^i \right) U_{2\nu}$$

donde se ha utilizado que la matriz $\mathcal{A}_{2\nu}$ es estrictamente estable, debido a su estructura triangular por cajas y a la estabilidad estricta de la matriz A . ■

Teniendo ahora en cuenta la expresión de las matrices $P_{l,i}^{r,s}(w, w)$, $P_{l,i}^{r,s}(v, v)$, $P_{l,i}^{r,s}(w, v)$ definidas en los Teoremas 3.5.2 y 3.5.5, observamos que dichas matrices convergen a límites finitos si los $2(\nu - 1)$ primeros momentos de $x(k)$ son convergentes. Podemos entonces establecer, como consecuencia inmediata del Lema 3.6.1, el siguiente resultado.

Corolario 3.6.2 *Si la matriz A es estrictamente estable, entonces*

- (i) *Las matrices $P_{l,i}^{r,s}(w, w)$, $P_{l,i}^{r,s}(v, v)$ que intervienen en las expresiones (3.12), (3.13), respectivamente, convergen a matrices constantes, concretamente*

$$\begin{aligned} \bar{P}_{l,i}^{r,s}(w, w) &:= \lim_{k \rightarrow \infty} P_{l,i}^{r,s}(w, w) = \text{vec}^{-1} \left[\left(I_{u, s-i} \otimes K_{u^{r-l}, n^i}^T \otimes I_{n, l} \right) \right. \\ &\quad \times \left. \left(\left(\text{vec } Q^{(r+s-l-i)} - \left(\text{vec } Q^{(s-i)} \right) \otimes \left(\text{vec } Q^{(r-l)} \right) \right) \otimes K_{1, n^i} \otimes I_{n, l} \right) m_{l+i} \right] \\ \bar{P}_{l,i}^{r,s}(v, v) &:= \lim_{k \rightarrow \infty} P_{l,i}^{r,s}(v, v) = \text{vec}^{-1} \left[\left(I_{t, s-i} \otimes K_{t^{r-l}, n^i}^T \otimes I_{n, l} \right) \right. \\ &\quad \times \left. \left(\left(\text{vec } R^{(r+s-l-i)} - \left(\text{vec } R^{(s-i)} \right) \otimes \left(\text{vec } R^{(r-l)} \right) \right) \otimes K_{1, n^i} \otimes I_{n, l} \right) m_{l+i} \right] \end{aligned}$$

para $r, s = 1, \dots, \nu$; $l = 0, \dots, r - 1$; $i = 0, \dots, s - 1$.

- (ii) *Además, en el caso de ruidos correlados, las matrices $P_{l,i}^{r,s}(w, v)$ que intervienen en la expresión (3.14) convergen también a la siguiente matriz*

$$\begin{aligned} \bar{P}_{l,i}^{r,s}(w, v) &:= \lim_{k \rightarrow \infty} P_{l,i}^{r,s}(w, v) = \text{vec}^{-1} \left[\left(I_{t, s-i} \otimes K_{u^{r-l}, n^i}^T \otimes I_{n, l} \right) \right. \\ &\quad \times \left. \left(\left(\text{vec } S^{(r-l, s-i)} - \left(\text{vec } R^{(s-i)} \right) \otimes \left(\text{vec } Q^{(r-l)} \right) \right) \otimes K_{1, n^i} \otimes I_{n, l} \right) m_{l+i} \right] \end{aligned}$$

para $r, s = 1, \dots, \nu$; $l = 0, \dots, r - 1$; $i = 0, \dots, s - 1$.

Como consecuencia de las propiedades anteriores, obtenemos el siguiente resultado.

Teorema 3.6.3 *Si la matriz A es estrictamente estable, entonces*

(i) Las matrices de covarianzas, $\mathcal{Q}(k) = E \{ \mathcal{F}(k) \mathcal{F}^T(k) \}$, $\mathcal{R}(k) = E \{ \mathcal{G}(k) \mathcal{G}^T(k) \}$, de los ruidos del sistema aumentado convergen a matrices constantes, \mathcal{Q} , \mathcal{R} , respectivamente, cuyos bloques (r, s) vienen expresados por

$$\begin{aligned} \mathcal{Q}_{rs} &= \sum_{l=0}^{r-1} \sum_{i=0}^{s-1} M_{r-l}^r(n) \left(F^{[r-l]} \otimes A^{[l]} \right) \bar{P}_{l,i}^{r,s}(w, w) \left(F^{[s-i]} \otimes A^{[i]} \right)^T \left(M_{s-i}^s(n) \right)^T \\ \mathcal{R}_{rs} &= p \left[\sum_{l=1}^{r-1} \sum_{i=0}^{s-1} M_{r-l}^r(m) \left(G^{[r-l]} \otimes C^{[l]} \right) \bar{P}_{l,i}^{r,s}(v, v) \left(G^{[s-i]} \otimes C^{[i]} \right)^T \left(M_{s-i}^s(m) \right)^T \right. \\ &\quad \left. + G^{[r]} \sum_{i=1}^{s-1} \bar{P}_{0,i}^{r,s}(v, v) \left(G^{[s-i]} \otimes C^{[i]} \right)^T \left(M_{s-i}^s(m) \right)^T \right] + G^{[r]} \bar{P}_{0,0}^{r,s}(v, v) G^{[s]T}. \end{aligned}$$

En consecuencia, los procesos $\{ \mathcal{F}(k) \}$ y $\{ \mathcal{G}(k) \}$ son asintóticamente estacionarios de segundo orden.

(ii) Además, en el caso de ruidos correlados, las matrices de covarianzas cruzadas, $\mathcal{S}(k) = E \{ \mathcal{F}(k) \mathcal{G}^T(k) \}$, convergen a una matriz constante, \mathcal{S} , cuyo bloque (r, s) está dado por

$$\begin{aligned} \mathcal{S}_{rs} &= p \sum_{l=0}^{r-1} \sum_{i=1}^{s-1} M_{r-l}^r(n) \left(F^{[r-l]} \otimes A^{[l]} \right) \bar{P}_{l,i}^{r,s}(w, v) \left(G^{[s-i]} \otimes C^{[i]} \right)^T \left(M_{s-i}^s(m) \right)^T \\ &\quad + \sum_{l=0}^{r-1} M_{r-l}^r(n) \left(F^{[r-l]} \otimes A^{[l]} \right) \bar{P}_{l,0}^{r,s}(w, v) G^{[s]T}. \end{aligned}$$

-Demostración-

La convergencia de las matrices $\mathcal{Q}(k)$ y $\mathcal{R}(k)$ se obtiene de forma inmediata teniendo en cuenta las expresiones (3.12) y (3.13), y utilizando el apartado (i) del Corolario 3.6.2. Consecuentemente, los ruidos $\{ \mathcal{F}(k) \}$ y $\{ \mathcal{G}(k) \}$ son asintóticamente estacionarios de segundo orden, ya que son procesos centrados y sus matrices de covarianzas tienen límite constante.

Por último, teniendo en cuenta la expresión (3.14), la convergencia de las matrices $\mathcal{S}(k)$ se deduce como consecuencia directa del apartado (ii) del Corolario 3.6.2. ■

Para completar el estudio asintótico del sistema aumentado veamos, finalmente, que el proceso estado $\{X(k)\}$ de dicho sistema aumentado es también asintóticamente estacionario de segundo orden.

Teorema 3.6.4 *Si la matriz A es estrictamente estable, entonces*

$$\lim_{k \rightarrow \infty} E \{X(k)\} := \mathcal{M} = \left(\sum_{i=0}^{\infty} \mathcal{A}^i \right) U$$

$$\lim_{k \rightarrow \infty} E \{X(k) X^T(k)\} := \mathcal{D} = \sum_{i=0}^{\infty} \mathcal{A}^i (UU^T + \mathcal{Q} + \mathcal{A}MU^T + UM^T \mathcal{A}^T) (\mathcal{A}^T)^i.$$

En consecuencia, el proceso estado $\{X(k)\}$ del sistema aumentado es asintóticamente estacionario de segundo orden.

-Demostración-

Aplicando el Lema 3.6.1 tenemos que

$$\lim_{k \rightarrow \infty} E \{x^{[i]}(k)\} = m_i, \quad i = 1, \dots, \nu$$

y, por tanto, teniendo en cuenta las expresiones de m_i y de la matriz \mathcal{A} , se obtiene de forma inmediata que

$$\mathcal{M} := \lim_{k \rightarrow \infty} E \{X(k)\} = \left(\sum_{i=0}^{\infty} \mathcal{A}^i \right) U.$$

Por otra parte, sabemos (Teoremas 3.5.4 y 3.5.7) que la matriz $\mathcal{D}(k) = E \{X(k) X^T(k)\}$ admite la siguiente expresión recursiva

$$\mathcal{D}(k) = \mathcal{A} \mathcal{D}(k-1) \mathcal{A}^T + UU^T + \mathcal{Q} + \mathcal{A} E \{X(k-1)\} U^T + U E \{X^T(k-1)\} \mathcal{A}^T.$$

Reiterando esta expresión, obtenemos que

$$\mathcal{D}(k) = \mathcal{A}^k \mathcal{D}(0) (\mathcal{A}^T)^k + \sum_{i=0}^{k-1} \mathcal{A}^i UU^T (\mathcal{A}^T)^i + \sum_{i=0}^{k-1} \mathcal{A}^i \mathcal{Q} (k-i-1) (\mathcal{A}^T)^i$$

$$\begin{aligned}
& + \sum_{i=0}^{k-1} \mathcal{A}^{i+1} E \{X(k-i-1)\} U^T (\mathcal{A}^T)^i + \sum_{i=0}^{k-1} \mathcal{A}^i U E \{X^T(k-i-1)\} (\mathcal{A}^T)^{i+1} \\
& = \mathcal{A}^k \mathcal{D}(0) (\mathcal{A}^T)^k + \sum_{i=0}^{k-1} \mathcal{A}^i U U^T (\mathcal{A}^T)^i + \sum_{i=0}^{k-1} \mathcal{A}^{k-i-1} \mathcal{Q}(i) (\mathcal{A}^T)^{k-i-1} \\
& + \sum_{i=0}^{k-1} \mathcal{A}^{k-i} E \{X(i)\} U^T (\mathcal{A}^T)^{k-i-1} + \sum_{i=0}^{k-1} \mathcal{A}^{k-i-1} U E \{X^T(i)\} (\mathcal{A}^T)^{k-i}.
\end{aligned}$$

Estudiamos, a continuación, el comportamiento asintótico de cada uno de los sumandos de esta expresión.

- El primer sumando converge a cero cuando $k \rightarrow \infty$, ya que \mathcal{A} es estrictamente estable debido a su estructura triangular por cajas y a la estabilidad estricta de la matriz A .
- El segundo sumando converge a la matriz

$$\sum_{i=0}^{\infty} \mathcal{A}^i U U^T (\mathcal{A}^T)^i$$

que es finita, en virtud de la estabilidad estricta de \mathcal{A} .

- Veamos que el tercer sumando converge también a una matriz constante. En efecto, sabemos que

$$\lim_{k \rightarrow \infty} \mathcal{Q}(k) = \mathcal{Q}$$

luego, para cada $\varepsilon > 0$, existe i_ε tal que

$$\mathcal{Q} - \varepsilon I \leq \mathcal{Q}(i) \leq \mathcal{Q} + \varepsilon I, \quad i > i_\varepsilon.$$

En consecuencia, para $k > i_\varepsilon + 1$, podemos escribir

$$\sum_{i=0}^{i_\varepsilon} \mathcal{A}^{k-i-1} \mathcal{Q}(i) (\mathcal{A}^T)^{k-i-1} + \sum_{i=i_\varepsilon+1}^{k-1} \mathcal{A}^{k-i-1} (\mathcal{Q} - \varepsilon I) (\mathcal{A}^T)^{k-i-1}$$

$$\begin{aligned} &\leq \sum_{i=0}^{k-1} \mathcal{A}^{k-i-1} \mathcal{Q}(i) (\mathcal{A}^T)^{k-i-1} \\ &\leq \sum_{i=0}^{i_\varepsilon} \mathcal{A}^{k-i-1} \mathcal{Q}(i) (\mathcal{A}^T)^{k-i-1} + \sum_{i=i_\varepsilon+1}^{k-1} \mathcal{A}^{k-i-1} (\mathcal{Q} + \varepsilon I) (\mathcal{A}^T)^{k-i-1}. \end{aligned}$$

El primer sumando del miembro de la derecha de esta desigualdad converge a cero, para $k \rightarrow \infty$, por ser \mathcal{A} estrictamente estable. El segundo sumando se puede expresar de la siguiente forma

$$\begin{aligned} &\sum_{i=0}^{k-1} \mathcal{A}^{k-i-1} (\mathcal{Q} + \varepsilon I) (\mathcal{A}^T)^{k-i-1} - \sum_{i=0}^{i_\varepsilon} \mathcal{A}^{k-i-1} (\mathcal{Q} + \varepsilon I) (\mathcal{A}^T)^{k-i-1} \\ &= \sum_{i=0}^{k-1} \mathcal{A}^i (\mathcal{Q} + \varepsilon I) (\mathcal{A}^T)^i - \sum_{i=0}^{i_\varepsilon} \mathcal{A}^{k-i-1} (\mathcal{Q} + \varepsilon I) (\mathcal{A}^T)^{k-i-1} \end{aligned}$$

que converge, para $k \rightarrow \infty$, a la matriz

$$\sum_{i=0}^{\infty} \mathcal{A}^i (\mathcal{Q} + \varepsilon I) (\mathcal{A}^T)^i.$$

Razonando de forma análoga con el miembro de la izquierda de la desigualdad, podemos escribir

$$\sum_{i=0}^{\infty} \mathcal{A}^i (\mathcal{Q} - \varepsilon I) (\mathcal{A}^T)^i \leq \lim_{k \rightarrow \infty} \sum_{i=0}^{k-1} \mathcal{A}^{k-i-1} \mathcal{Q}(i) (\mathcal{A}^T)^{k-i-1} \leq \sum_{i=0}^{\infty} \mathcal{A}^i (\mathcal{Q} + \varepsilon I) (\mathcal{A}^T)^i$$

y, haciendo $\varepsilon \rightarrow 0$, concluimos que

$$\lim_{k \rightarrow \infty} \sum_{i=0}^{k-1} \mathcal{A}^{k-i-1} \mathcal{Q}(i) (\mathcal{A}^T)^{k-i-1} = \sum_{i=0}^{\infty} \mathcal{A}^i \mathcal{Q} (\mathcal{A}^T)^i.$$

En cuanto a los sumandos cuarto y quinto, puesto que

$$\lim_{k \rightarrow \infty} E \{X(k)\} = \mathcal{M},$$

para cada $\varepsilon > 0$, existe i_ε tal que

$$\mathcal{M}U^T - \varepsilon I \leq E \{X(i)\} U^T \leq \mathcal{M}U^T + \varepsilon I, \quad i > i_\varepsilon$$

y, procediendo de forma totalmente análoga a como se ha hecho para el tercer sumando, obtenemos que

$$\lim_{k \rightarrow \infty} \sum_{i=0}^{k-1} \mathcal{A}^{k-i} E \{X(i)\} U^T (\mathcal{A}^T)^{k-i-1} = \sum_{i=0}^{\infty} \mathcal{A}^{i+1} \mathcal{M} U^T (\mathcal{A}^T)^i.$$

Queda, así, demostrado que

$$\begin{aligned} \mathcal{D} &:= \lim_{k \rightarrow \infty} \mathcal{D}(k) \\ &= \sum_{i=0}^{\infty} \mathcal{A}^i (U U^T + \mathcal{Q}) (\mathcal{A}^T)^i + \sum_{i=0}^{\infty} \mathcal{A}^{i+1} \mathcal{M} U^T (\mathcal{A}^T)^i + \sum_{i=0}^{\infty} \mathcal{A}^i U \mathcal{M}^T (\mathcal{A}^T)^{i+1} \\ &= \sum_{i=0}^{\infty} \mathcal{A}^i (U U^T + \mathcal{Q} + \mathcal{A} \mathcal{M} U^T + U \mathcal{M}^T \mathcal{A}^T) (\mathcal{A}^T)^i. \end{aligned}$$

En consecuencia, el proceso estado $\{X(k)\}$ es asintóticamente estacionario de segundo orden, ya que, en el límite, sus momentos de primer y segundo orden son constantes. ■

3.7 Filtro Polinomial Steady-State

El objetivo de esta Sección es analizar si el filtro lineal obtenido tanto en la Sección 3.5.1 (caso de ruidos independientes), como en la Sección 3.5.2 (caso de ruidos correlados), para el sistema aumentado correspondiente a un sistema original estacionario, admite una forma steady-state. Esto permitiría calcular, a partir de él, el filtro polinomial steady-state para el estado del sistema original, con las ventajas computacionales que ello conlleva.

Como es conocido, para sistemas invariantes en el tiempo, es decir, sistemas con matrices de transición y observación constantes, una condición necesaria para la existencia del filtro steady-state es la estacionariedad de los procesos asociados al sistema.

Consideremos, por tanto, el sistema estacionario (3.15)-(3.16) y su correspondiente sistema aumentado (3.17)-(3.18).

Según hemos probado en la Sección 3.6, tanto en el caso de ruidos independientes como en el caso de ruidos correlados, si el sistema estacionario original es asintóticamente estable (su matriz de transición, A , es estrictamente estable), los procesos ruido $\{\mathcal{F}(k)\}$, $\{\mathcal{G}(k)\}$ y el proceso estado $\{X(k)\}$ del sistema aumentado son asintóticamente estacionarios de segundo orden, es decir, en el límite sus momentos de primer y segundo orden son constantes.

Por consiguiente, en el caso de independencia mutua de la condición inicial y los ruidos del sistema estacionario original, si la matriz A es estrictamente estable, podemos utilizar el filtro steady-state de Nahi (ver Tugnait [35]) para estimar el vector estado del sistema aumentado, lo que nos proporcionará el filtro polinomial steady-state para el estado del sistema estacionario original.

Supongamos ahora que los ruidos aditivos, $\{w(k)\}$ y $\{v(k)\}$, que afectan al sistema estacionario original son correlados entre sí. Con el fin de estudiar la existencia del filtro steady-state en este caso, consideremos el siguiente sistema

$$X(k+1) = \mathcal{A}X(k) + U + \mathcal{F}(k), \quad k \geq 0 \quad (3.21)$$

$$X(0) = X_0$$

$$Z(k) = \mathcal{C}_e X(k) + V + \mathcal{G}_e(k), \quad k \geq 0 \quad (3.22)$$

donde la ecuación del estado coincide con (3.17), la matriz de observación está dada por

$$\mathcal{C}_e = p\mathcal{C}$$

y el proceso $\{\mathcal{G}_e(k); k \geq 0\}$ es un ruido blanco centrado con

$$\mathcal{R}_e(k) = E\{\mathcal{G}_e(k)\mathcal{G}_e^T(k)\} = \mathcal{R}(k) + p(1-p)\mathcal{C}\mathcal{D}(k)\mathcal{C}^T, \quad k \geq 0$$

siendo

$$\mathcal{D}(k) = E\{X(k)X^T(k)\}, \quad k \geq 0$$

las matrices dadas en el Teorema 3.5.7.

Supondremos, además, que el estado inicial, X_0 , es incorrelado con los ruidos $\{\mathcal{F}(k)\}$, $\{\mathcal{G}_e(k)\}$ y que

$$E \{ \mathcal{F}(k) \mathcal{G}_e^T(j) \} = \mathcal{S}(k) \delta_{kj}$$

siendo $\mathcal{S}(k) = E \{ \mathcal{F}(k) \mathcal{G}_e^T(k) \}$ (Teorema 3.5.5).

El sistema (3.21)-(3.22) es un sistema invariante en el tiempo con certidumbre en las observaciones, ya que la matriz de observación, \mathcal{C}_e , es determinística. Además, dicho sistema verifica las hipótesis requeridas para aplicar el Filtro de Kalman para sistemas con ruidos correlados (ver Balakrishnan [2]). Concretamente, el algoritmo para los problemas de predicción y filtrado lineales del estado $X(k)$ del sistema (3.21)-(3.22) viene dado mediante las siguientes ecuaciones

$$\hat{X}_e(k+1/k) = \mathcal{A}\hat{X}_e(k/k-1) + U + (\mathcal{A}K_e(k) + \mathcal{S}(k)\Pi_e^-(k))\tilde{Z}_e(k/k-1)$$

$$\hat{X}_e(0/-1) = \bar{X}_0$$

$$\hat{X}_e(k/k) = \hat{X}_e(k/k-1) + K_e(k)\tilde{Z}_e(k/k-1), \quad k \geq 0$$

siendo $\tilde{Z}_e(k/k-1)$ la innovación

$$\tilde{Z}_e(k/k-1) = Z(k) - \mathcal{C}_e\hat{X}_e(k/k-1) - V, \quad k \geq 0.$$

La ganancia $K_e(k)$ está dada por

$$K_e(k) = P_e(k/k-1)\mathcal{C}_e^T\Pi_e^-(k), \quad k \geq 0$$

y las matrices de covarianzas de los errores de filtrado y predicción se calculan mediante la siguiente relación recursiva

$$P_e(k+1/k) = \mathcal{A}P_e(k/k)\mathcal{A}^T + Q(k) - \mathcal{S}(k)\Pi_e^-(k)\mathcal{S}^T(k) \\ - \mathcal{A}K_e(k)\mathcal{S}^T(k) - \mathcal{S}(k)K_e^T(k)\mathcal{A}^T$$

$$P_e(0/-1) = P_0$$

$$P_e(k/k) = P_e(k/k-1) - K_e(k)\Pi_e(k)K_e^T(k), \quad k \geq 0.$$

siendo $\Pi_e(k)$ la matriz de covarianzas del proceso innovación,

$$\Pi_e(k) = C_e P_e(k/k-1) C_e^T + \mathcal{R}_e(k), \quad k \geq 0.$$

Comparando este algoritmo con el establecido en el Teorema 3.5.7 y teniendo en cuenta la definición de las matrices C_e y $\mathcal{R}_e(k)$, observamos que el estimador lineal $\hat{X}_e(k/k)$ del estado del sistema (3.21)-(3.22) coincide con el estimador lineal $\hat{X}(k/k)$ del estado del sistema aumentado, siempre que las observaciones $Z(0), \dots, Z(k)$ sean las mismas para ambos sistemas. En este sentido diremos que el sistema (3.21)-(3.22) es equivalente al sistema aumentado (3.17)-(3.18).

Como hemos demostrado, si la matriz de transición, A , del sistema estacionario original es estrictamente estable, entonces el sistema aumentado (3.17)-(3.18) es asintóticamente estacionario de segundo orden. En consecuencia, el sistema equivalente (3.21)-(3.22) es también asintóticamente estacionario de segundo orden, ya que los procesos que en él intervienen coinciden con los del sistema aumentado, a excepción del ruido de la observación $\{\mathcal{G}_e(k)\}$, cuyas matrices de covarianzas verifican

$$\lim_{k \rightarrow \infty} \mathcal{R}_e(k) = \mathcal{R} + p(1-p)CDC^T =: \mathcal{R}_e$$

siendo

$$\mathcal{D} = \lim_{k \rightarrow \infty} \mathcal{D}(k) \quad (\text{Teorema 3.6.4})$$

Por consiguiente, tanto los ruidos $\{\mathcal{F}(k)\}$, $\{\mathcal{G}_e(k)\}$, como el proceso estado $\{X(k)\}$ asociados al sistema (3.21)-(3.22) son asintóticamente estacionarios de segundo orden y, además, en virtud del Teorema 3.6.3 se tiene que

$$\lim_{k \rightarrow \infty} \mathcal{S}(k) = \mathcal{S}.$$

Podemos, por tanto, utilizar el Filtro de Kalman steady-state para sistemas con ruidos correlados (ver Balakrishnan [2]), el cual nos conduce al siguiente algoritmo que

permite calcular el estimador lineal del estado del sistema equivalente

$$\hat{X}_e(k+1/k) = \mathcal{A}\hat{X}_e(k/k-1) + U + (\mathcal{A}K_e + S\Pi_e^-)\tilde{Z}_e(k/k-1), \quad k \geq 0$$

$$\hat{X}_e(0/-1) = \bar{X}_0$$

$$\hat{X}_e(k/k) = \hat{X}_e(k/k-1) + K_e\tilde{Z}_e(k/k-1), \quad k \geq 0$$

$$\tilde{Z}_e(k/k-1) = Z(k) - C_e\hat{X}_e(k/k-1) - V, \quad k \geq 0.$$

$$K_e = P_e C_e^T \Pi_e^-$$

$$P_e = \mathcal{A} [P_e - K_e \Pi_e K_e^T] \mathcal{A}^T + Q - S \Pi_e^- S^T - \mathcal{A} K_e S^T - S K_e^T \mathcal{A}^T$$

$$\Pi_e = C_e P_e C_e^T + \mathcal{R}_e.$$

Como ya hemos indicado, si las observaciones $Z(0), \dots, Z(k)$ son las mismas para ambos sistemas, se tiene que $\hat{X}_e(k/k) = \hat{X}(k/k)$. Teniendo esto en cuenta y sustituyendo las matrices C_e y \mathcal{R}_e por sus valores, obtenemos el algoritmo steady-state para los problemas de predicción y filtrado del estado aumentado, que se recoge en el siguiente teorema.

Teorema 3.7.1 *Si la matriz de transición del sistema estacionario original es estrictamente estable, entonces el algoritmo de predicción y filtrado presentado en la Sección 3.5.2 admite la siguiente versión steady-state*

$$\hat{X}(k+1/k) = \mathcal{A}\hat{X}(k/k-1) + U + (\mathcal{A}K + S\Pi^-)\tilde{Z}(k/k-1), \quad k \geq 0$$

$$\hat{X}(0/-1) = \bar{X}_0$$

$$\hat{X}(k/k) = \hat{X}(k/k-1) + K\tilde{Z}(k/k-1), \quad k \geq 0$$

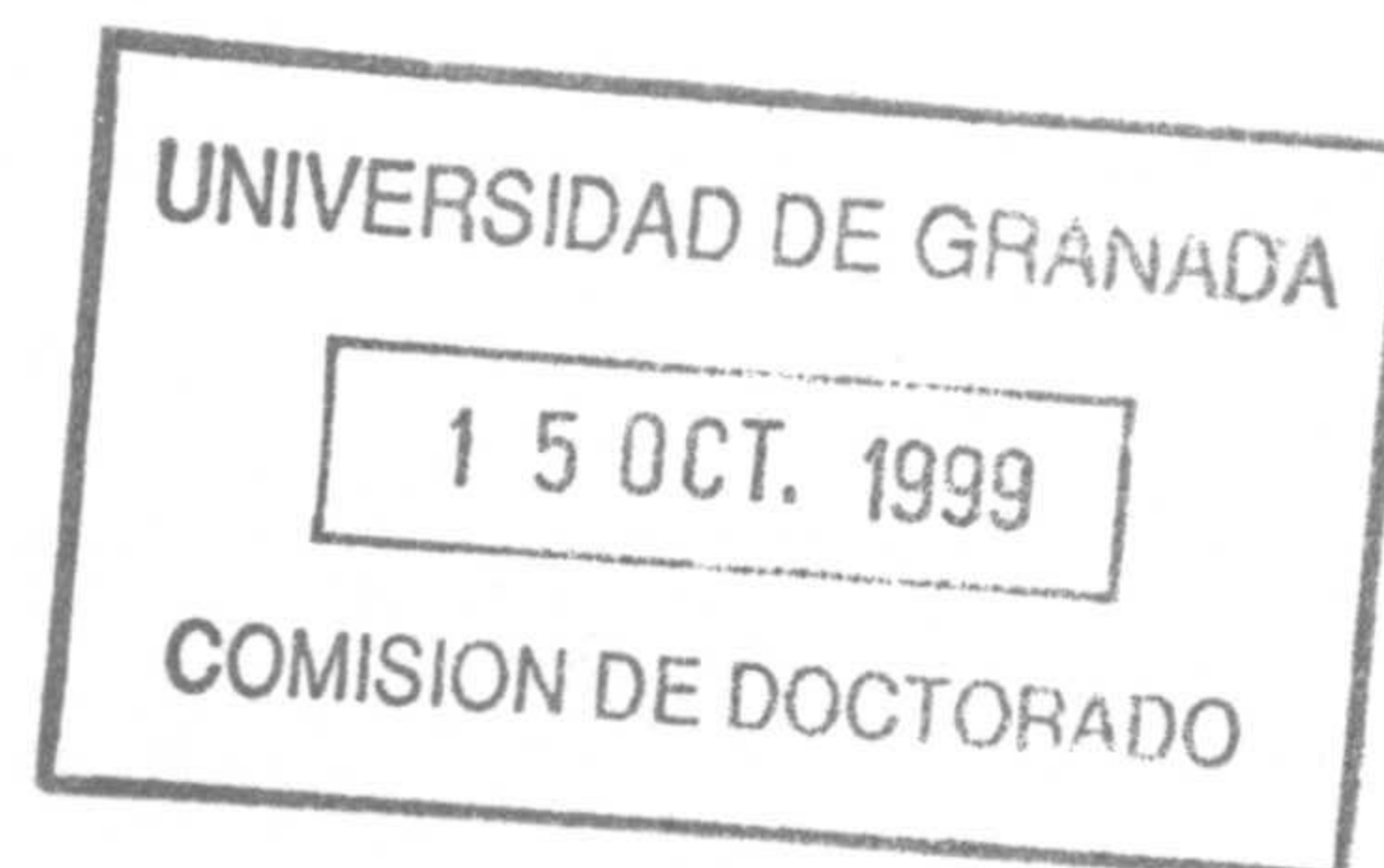
$$\tilde{Z}(k/k-1) = Z(k) - pC\hat{X}(k/k-1) - V, \quad k \geq 0$$

$$K = pPC^T\Pi^-$$

$$P = \mathcal{A} [P - K\Pi K^T] \mathcal{A}^T + Q - S\Pi^- S^T - \mathcal{A}K S^T - S K^T \mathcal{A}^T$$

$$\Pi = p(1-p)CDC^T + p^2CPC^T + \mathcal{R}.$$

Finalmente, las n primeras componentes de $\hat{X}(k/k)$ proporcionarán el estimador polinomial steady-state de grado ν del estado del sistema original, $x(k)$.



Ejemplo Numérico

Con el fin de ilustrar los resultados obtenidos en esta memoria vamos a analizar, en un ejemplo numérico, el comportamiento de los estimadores que proporcionan los algoritmos de predicción y filtrado propuestos en la Sección 3.5 del Capítulo 3 para sistemas con observaciones inciertas, tanto en el caso de ruidos mutuamente independientes (Teorema 3.5.4), como en el caso de ruidos correlados (Teorema 3.5.7).

Hay que indicar que, para el estudio que nos proponemos, no es necesario obtener los estimadores de forma explícita, ya que el cálculo de las matrices de covarianza de los errores de estimación es independiente de la obtención del filtro y el predictor. Esto permite determinar una medida de la bondad de la estimación sin necesidad de calcular explícitamente los estimadores.

Para obtener, en cada iteración del algoritmo, las matrices de covarianza del error asociado al estimador polinomial de grado $\nu \geq 1$ arbitrario, hemos realizado un programa con "MATLAB" que ha sido aplicado, considerando $\nu = 1$, $\nu = 2$ y $\nu = 3$, al ejemplo que describimos a continuación.

Consideremos el siguiente sistema escalar estacionario con observaciones inciertas

$$x(k+1) = 1/2x(k) + w(k), \quad k \geq 0$$

$$x(0) = x_0$$

$$z(k) = u(k)x(k) + v(k), \quad k \geq 0$$

donde $\{w(k); k \geq 0\}$ y $\{v(k); k \geq 0\}$ son sucesiones de variables aleatorias independientes e idénticamente distribuidas, centradas, con función masa de probabilidad

$$P[w(k) = -1] = \frac{15}{18}, \quad P[w(k) = 3] = \frac{2}{18}, \quad P[w(k) = 9] = \frac{1}{18}$$

$$P[v(k) = 1] = \frac{15}{18}, \quad P[v(k) = -3] = \frac{2}{18}, \quad P[v(k) = -9] = \frac{1}{18}.$$

Para el estudio del problema de estimación polinomial de grado $\nu \leq 3$ necesitamos los momentos $Q^{(i)} = E\{w^i(k)\}$ y $R^{(i)} = E\{v^i(k)\}$ ($i = 1, \dots, 6$) de estas variables, que vienen dados por

$$Q^{(1)} = 0, \quad Q^{(2)} = \frac{19}{3}, \quad Q^{(3)} = \frac{128}{3}, \quad Q^{(4)} = \frac{1123}{3}, \quad Q^{(5)} = \frac{9920}{3}, \quad Q^{(6)} = \frac{88819}{3}$$

$$R^{(1)} = 0, \quad R^{(2)} = \frac{19}{3}, \quad R^{(3)} = \frac{-128}{3}, \quad R^{(4)} = \frac{1123}{3}, \quad R^{(5)} = \frac{-9920}{3}, \quad R^{(6)} = \frac{88819}{3}.$$

El ruido multiplicativo $\{u(k); k \geq 0\}$ es una sucesión de variables aleatorias de Bernoulli independientes e idénticamente distribuidas con parámetro p .

El estado inicial, x_0 , es una variable aleatoria centrada con momentos

$$E\{x_0^2\} = 1, \quad E\{x_0^3\} = 0, \quad E\{x_0^4\} = 3, \quad E\{x_0^5\} = 0, \quad E\{x_0^6\} = 15.$$

Este sistema fue considerado por García-Ligero et al. [12] para el estudio del problema de estimación polinomial de segundo grado en sistemas con observaciones inciertas bajo la hipótesis de independencia mutua del estado inicial y los ruidos. Nuestro objetivo es ahora, por una parte, y bajo la misma hipótesis de independencia mutua,

analizar el problema de estimación polinomial de tercer grado (Caso 1) y, por otra, bajo hipótesis más débiles sobre la independencia de los ruidos del sistema, analizar los problemas de estimación lineal, cuadrática o polinomial de segundo grado y cúbica o polinomial de tercer grado (Caso 2).

Caso 1: Ruidos Aditivos Independientes

Supongamos, en primer lugar, de acuerdo con el desarrollo teórico llevado a cabo en la Sección 3.5.1, que el estado inicial y los ruidos son mutuamente independientes.

Para este sistema, hemos realizado cincuenta iteraciones del algoritmo propuesto para los problemas de filtrado lineal, polinomial de segundo grado y polinomial de tercer grado. Las Tablas 1-4 muestran, para distintos valores del parámetro p , las varianzas de los errores de estimación obtenidas en cada una de las iteraciones del algoritmo.

Indicamos, seguidamente, algunos comentarios sobre los resultados reflejados en las Tablas 1-4.

- Confirmando los resultados teóricos, en cada una de las tablas se aprecia que la varianza del error cometido al estimar el estado mediante una función polinómica de las observaciones disminuye al aumentar el grado de dicha función. Además, la utilización de un estimador polinomial de grado $\nu > 1$ puede mejorar, de forma considerable, al ampliamente utilizado estimador lineal.
- Si comparamos ahora las distintas tablas, el estudio realizado para diferentes valores de p pone de manifiesto que, como intuitivamente podía pensarse, a medida que aumenta la probabilidad de falsa alarma, las varianzas de los errores de estimación (lineal, cuadrática y cúbica) también aumentan.

Si $p = 1$, tenemos un sistema con certidumbre en las observaciones y el algoritmo propuesto coincide con el Filtro de Kalman. En este caso extremo, las varianzas

de los errores de estimación son más pequeñas que para cualquier otro valor de $p < 1$. Por consiguiente, si a un sistema con observaciones inciertas le aplicásemos el Filtro de Kalman, las conclusiones obtenidas serían poco fiables ya que, dependiendo del valor de p , las varianzas de los errores de estimación podrían ser hasta cuatro veces mayores que las obtenidas ignorando la incertidumbre de las observaciones.

El otro valor extremo, $p = 0$, no ha sido considerado ya que, en este caso, las matrices de ganancia se reducen a la matriz nula y, por tanto, las observaciones (que en este caso serían únicamente ruido) no son tenidas en cuenta para el cálculo de los estimadores.

- Como antes indicamos, se han realizado cincuenta iteraciones del algoritmo para los problemas de estimación lineal, cuadrática y cúbica. Notemos no obstante que, como puede observarse en cada una de las tablas, los valores de las varianzas de los errores de estimación se estabilizan a partir de una determinada iteración.

También hemos considerado, para $p = 1/4$, $p = 1/2$, $p = 3/4$ y $p = 1$, los valores steady-state de las varianzas de los errores del estimador lineal, \bar{p} , polinomial de segundo grado, \bar{P}_{11}^2 , y polinomial de tercer grado, \bar{P}_{11}^3 , del estado $x(k)$ del sistema estacionario propuesto (notemos que este sistema es asintóticamente estable).

La Tabla 5 muestra dichas varianzas steady-state; la relación entre \bar{p} , \bar{P}_{11}^2 y \bar{P}_{11}^3 pone de nuevo de manifiesto que los estimadores polinomiales que proporciona el método propuesto mejoran al lineal.

Notemos, por último, que estos resultados obtenidos en el análisis steady-state coinciden con los valores establecidos en las Tablas 1-4, ya que, como antes señalamos, a partir de una determinada iteración, los valores de las varianzas de los errores de estimación se estabilizan. Por tanto, la Tabla 5 nos conduce de nuevo a las mismas conclusiones que hemos indicado anteriormente.

$p = 1/4$			
k	Lineal $p(k/k)$	Cuadrático $P_{11}^2(k/k)$	Cúbico $P_{11}^3(k/k)$
0	0.990506329114	0.967039203673	0.966913692677
1	6.241718360211	5.419092376429	5.397192491204
2	7.425834548711	6.330992002089	6.311959377084
3	7.691138048843	6.524903715516	6.507614936999
4	7.750587334910	6.567132775062	6.550395167887
5	7.763929886852	6.576482740700	6.559892273295
6	7.766930522291	6.578580017955	6.562025859497
7	7.767606876944	6.579055794053	6.562510177421
8	7.767759707698	6.579164851028	6.562621175187
9	7.767794334109	6.579190093757	6.562646846196
10	7.767802201835	6.579195990353	6.562652834729
11	7.767803994994	6.579197379608	6.562654243154
12	7.767804405006	6.579197709509	6.562654576919
13	7.767804499077	6.579197788413	6.562654656568
14	7.767804520737	6.579197807406	6.562654675696
15	7.767804525743	6.579197812004	6.562654680316
16	7.767804526905	6.579197813122	6.562654681438
17	7.767804527175	6.579197813396	6.562654681711
18	7.767804527238	6.579197813463	6.562654681778
19	7.767804527253	6.579197813479	6.562654681794
20	7.767804527257	6.579197813483	6.562654681798
21	7.767804527258	6.579197813484	6.562654681799
≥ 22	7.767804527258	6.579197813485	6.562654681800

Tabla 1 Probabilidad de falsa alarma: $1 - p = 3/4$

$p(k/k)$ varianza del error del filtro lineal

$P_{11}^2(k/k)$ varianza del error del filtro polinomial de segundo grado

$P_{11}^3(k/k)$ varianza del error del filtro polinomial de tercer grado

$p = 1/2$			
k	Lineal $p(k/k)$	Cuadrático $P_{11}^2(k/k)$	Cúbico $P_{11}^3(k/k)$
0	0.963414634146	0.908045365031	0.900717278692
1	5.451324532453	4.074714855283	4.072384033166
2	6.251812443155	4.601689671596	4.601118721031
3	6.395932179435	4.701568221955	4.700797044576
4	6.422716763398	4.721940555710	4.721043486481
5	6.427896612267	4.726312044702	4.725381339894
6	6.428943063248	4.727287256750	4.726349834960
7	6.429163948175	4.727511554897	4.726573108422
8	6.429212494605	4.727564374744	4.726625847054
9	6.429223537291	4.727577035448	4.726638528066
10	6.429226118682	4.727580109585	4.726641616014
11	6.429226734648	4.727580862898	4.726642374568
12	6.429226883824	4.727581048682	4.726642562017
13	6.429226920329	4.727581094703	4.726642608523
14	6.429226929325	4.727581106138	4.726642620092
15	6.429226931553	4.727581108985	4.726642622974
16	6.429226932107	4.727581109694	4.726642623693
17	6.429226932245	4.727581109871	4.726642623873
18	6.429226932279	4.727581109916	4.726642623918
19	6.429226932288	4.727581109927	4.726642623929
20	6.429226932290	4.727581109929	4.726642623932
21	6.429226932290	4.727581109930	4.726642623933
≥ 22	6.429226932291	4.727581109930	4.726642623933

Tabla 2 Probabilidad de falsa alarma: $1 - p = 1/2$

$p(k/k)$ varianza del error del filtro lineal

$P_{11}^2(k/k)$ varianza del error del filtro polinomial de segundo grado

$P_{11}^3(k/k)$ varianza del error del filtro polinomial de tercer grado

$p = 3/4$			
k	Lineal $p(k/k)$	Cuadrático $P_{11}^2(k/k)$	Cúbico $P_{11}^3(k/k)$
0	0.920588235294	0.835319323454	0.808275337315
1	4.411365756456	2.687307201861	2.681406469136
2	4.846790991366	2.935863770566	2.927159835761
3	4.907096966113	2.980754799532	2.971132894904
4	4.917028784420	2.989899490885	2.980176826295
5	4.918967636863	2.991903944400	2.982189686796
6	4.919394787721	2.992365464889	2.982659759627
7	4.919495485467	2.992475220504	2.982772780852
8	4.919520015636	2.992501858991	2.982800424359
9	4.919526080025	2.992508404877	2.982807252208
10	4.919527588913	2.992510025228	2.982808947951
11	4.919527965372	2.992510428038	2.982809370378
12	4.919528059406	2.992510528419	2.982809475783
13	4.919528082906	2.992510553469	2.982809502107
14	4.919528088780	2.992510559725	2.982809508684
15	4.919528090249	2.992510561288	2.982809510328
16	4.919528090616	2.992510561679	2.982809510739
17	4.919528090708	2.992510561777	2.982809510842
18	4.919528090731	2.992510561801	2.982809510868
19	4.919528090736	2.992510561807	2.982809510874
≥ 20	4.919528090738	2.992510561809	2.982809510876

Tabla 3 Probabilidad de falsa alarma: $1 - p = 1/4$

$p(k/k)$ varianza del error del filtro lineal

$P_{11}^2(k/k)$ varianza del error del filtro polinomial de segundo grado

$P_{11}^3(k/k)$ varianza del error del filtro polinomial de tercer grado

$p = 1$			
k	Lineal $p(k/k)$	Cuadrático $P_{11}^2(k/k)$	Cúbico $P_{11}^3(k/k)$
0	0.863636363636	0.751511163652	0.690833015793
1	3.219739292365	1.268196928910	1.246786686097
2	3.355876559422	1.291277161683	1.260557605346
3	3.363379775742	1.293774801377	1.261373025291
4	3.363792216698	1.294063001874	1.261439104948
5	3.363814884660	1.294096457876	1.261445113609
6	3.363816130492	1.294100344758	1.261445683091
7	3.363816198963	1.294100796384	1.261445737861
8	3.363816202727	1.294100848860	1.261445743156
9	3.363816202933	1.294100854957	1.261445743669
10	3.363816202945	1.294100855666	1.261445743718
11	3.363816202945	1.294100855748	1.261445743723
12	3.363816202945	1.294100855758	1.261445743724
≥ 13	3.363816202945	1.294100855759	1.261445743724

Tabla 4 Probabilidad de falsa alarma: $1 - p = 0$

$p(k/k)$ varianza del error del filtro lineal

$P_{11}^2(k/k)$ varianza del error del filtro polinomial de segundo grado

$P_{11}^3(k/k)$ varianza del error del filtro polinomial de tercer grado

p	Lineal \bar{p}	Cuadrático \bar{P}_{11}^2	Cúbico \bar{P}_{11}^3
1/4	7.767804527258	6.579197813485	6.562654681800
1/2	6.429226932291	4.727581109930	4.726642623933
3/4	4.919528090738	2.992510561809	2.982809510876
1	3.363816202945	1.294100855759	1.261445743724

Tabla 5 \bar{p} varianza steady-state del error del filtro lineal

\bar{P}_{11}^2 varianza steady-state del error del filtro polinomial de segundo grado

\bar{P}_{11}^3 varianza steady-state del error del filtro polinomial de tercer grado

Caso 2: Ruidos Aditivos Correlados

Supongamos ahora, de acuerdo con el desarrollo teórico llevado a cabo en la Sección 3.5.2, que los ruidos aditivos, $\{w(k); k \geq 0\}$ y $\{v(k); k \geq 0\}$, que afectan al sistema no son independientes entre sí, sino que su función masa de probabilidad conjunta está dada por

$$P[w(k) = w, v(k) = v] = \begin{cases} 14/18, & \text{si } w = -1, v = 1 \\ 1/18, & \text{si } w = -1, v = -9 \\ 1/18, & \text{si } w = 3, v = 1 \\ 1/18, & \text{si } w = 3, v = -3 \\ 1/18, & \text{si } w = 9, v = -3 \\ 0, & \text{en otro caso} \end{cases}$$

de modo que las distribuciones marginales de $w(k)$ y $v(k)$ coinciden con las establecidas en el Caso 1. Para $i, j \leq 3$, los momentos bidimensionales $S^{(i,j)} = E\{w^i(k)v^j(k)\}$ vienen dados por

$$\begin{aligned} S^{(1,1)} &= -38/18 & S^{(2,1)} &= -256/18 & S^{(3,1)} &= -2246/18 \\ S^{(1,2)} &= 16/18 & S^{(2,2)} &= 914/18 & S^{(3,2)} &= 6736/18 \\ S^{(1,3)} &= 394/18 & S^{(2,3)} &= -3136/18 & S^{(3,3)} &= -19670/18. \end{aligned}$$

También en este caso hemos implementado el algoritmo propuesto para los problemas de filtrado lineal, polinomial de segundo grado y polinomial de tercer grado. Las Tablas 6-9 muestran, para diferentes valores del parámetro p , las varianzas de los errores de estimación obtenidas en cada iteración del algoritmo. Nuevamente, hemos realizado cincuenta iteraciones del algoritmo, si bien, como ocurría en el caso anterior, los valores de las varianzas de los errores de estimación se estabilizan a partir de una determinada iteración.

Hay que indicar que todos los comentarios realizados en el Caso 1, acerca de los resultados reflejados en las tablas, pueden efectuarse también en este caso.

$p = 1/4$			
k	Lineal $p(k/k)$	Cuadrático $P_{11}^2(k/k)$	Cúbico $P_{11}^3(k/k)$
0	0.990506329114	0.967039203673	0.966913692677
1	5.702328947137	4.117175374773	3.479212258148
2	7.161214853975	6.004948078018	4.500581372394
3	7.585109184746	6.565350250427	4.868804973698
4	7.706456901835	6.722083828858	4.999142907009
5	7.741044216893	6.765315812569	5.044560870499
6	7.750886393493	6.777183964222	5.060240095797
7	7.753684653795	6.780435436946	5.065620759018
8	7.754479748904	6.781325358217	5.067459753337
9	7.754705554135	6.781568803186	5.068086472042
10	7.754769655006	6.781635380552	5.068299610506
11	7.754787845005	6.781653585081	5.068371986422
12	7.754793005129	6.781658562291	5.068396536113
13	7.754794468530	6.781659922986	5.068404856541
14	7.754794883444	6.781660294959	5.068407674835
15	7.754795001057	6.781660396641	5.068408629028
16	7.754795034389	6.781660424436	5.068408951985
17	7.754795043834	6.781660432033	5.068409061268
18	7.754795046510	6.781660434110	5.068409098241
19	7.754795047268	6.781660434678	5.068409110748
20	7.754795047483	6.781660434833	5.068409114978
21	7.754795047544	6.781660434875	5.068409116409
22	7.754795047561	6.781660434887	5.068409116893
23	7.754795047566	6.781660434890	5.068409117057
24	7.754795047567	6.781660434891	5.068409117112
25	7.754795047568	6.781660434891	5.068409117131
26	7.754795047568	6.781660434891	5.068409117137
≥ 27	7.754795047568	6.781660434891	5.068409117140

Tabla 6 Probabilidad de falsa alarma: $1 - p = 3/4$

$p = 1/2$			
k	Lineal $p(k/k)$	Cuadrático $P_{11}^2(k/k)$	Cúbico $P_{11}^3(k/k)$
0	0.963414634146	0.908045365031	0.900717278692
1	5.104606345744	3.689028696690	3.335551262015
2	6.330361641887	4.696630201960	4.134665534477
3	6.648480506303	4.919386765992	4.374026791362
4	6.728738282835	4.968104504144	4.444930365818
5	6.748848205849	4.978932453469	4.465870744634
6	6.753878528664	4.981370336840	4.472061661955
7	6.755136286991	4.981924094677	4.473893681932
8	6.755450737674	4.982050677478	4.474436064267
9	6.755529351039	4.982079757293	4.474596667782
10	6.755549004424	4.982086466631	4.474644225199
11	6.755553917772	4.982088020826	4.474658307381
12	6.755555146110	4.982088382249	4.474662477049
13	6.755555453194	4.982088466619	4.474663711606
14	6.755555529965	4.982088486390	4.474664077117
15	6.755555549158	4.982088491041	4.474664185328
16	6.755555553956	4.982088492139	4.474664217363
17	6.755555555156	4.982088492399	4.474664226846
18	6.755555555456	4.982088492461	4.474664229654
19	6.755555555531	4.982088492476	4.474664230485
20	6.755555555549	4.982088492480	4.474664230731
21	6.755555555554	4.982088492481	4.474664230803
22	6.755555555555	4.982088492481	4.474664230825
23	6.755555555555	4.982088492481	4.474664230831
24	6.755555555556	4.982088492481	4.474664230833
≥ 25	6.755555555556	4.982088492481	4.474664230834

Tabla 7 Probabilidad de falsa alarma: $1 - p = 1/2$

$p = 3/4$			
k	Lineal $p(k/k)$	Cuadrático $P_{11}^2(k/k)$	Cúbico $P_{11}^3(k/k)$
0	0.920588235294	0.835319323454	0.808275337315
1	4.224324901530	2.630824792828	2.474391381554
2	5.027107221591	3.031280927760	2.880056914994
3	5.193114781050	3.098445198067	2.978712152145
4	5.226555094042	3.110959148435	3.001984973966
5	5.233299812426	3.113513274006	3.007399124302
6	5.234671377013	3.114067597159	3.008661686352
7	5.234953359956	3.114193014483	3.008957570490
8	5.235012079548	3.114222226911	3.009027215666
9	5.235024483272	3.114229173089	3.009043668399
10	5.235027144277	3.114230849323	3.009047567428
11	5.235027724502	3.114231258108	3.009048494114
12	5.235027853126	3.114231358547	3.009048714968
13	5.235027882106	3.114231383355	3.009048767744
14	5.235027888737	3.114231389505	3.009048780388
15	5.235027890276	3.114231391033	3.009048783425
16	5.235027890638	3.114231391414	3.009048784156
17	5.235027890724	3.114231391509	3.009048784333
18	5.235027890745	3.114231391533	3.009048784376
19	5.235027890750	3.114231391539	3.009048784386
20	5.235027890751	3.114231391540	3.009048784389
21	5.235027890751	3.114231391540	3.009048784389
≥ 22	5.235027890751	3.114231391541	3.009048784389

Tabla 8 Probabilidad de falsa alarma: $1 - p = 1/4$

$p = 1$			
k	Lineal $p(k/k)$	Cuadrático $P_{11}^2(k/k)$	Cúbico $P_{11}^3(k/k)$
0	0.863636363636	0.751511163652	0.690833015793
1	3.140462550393	1.266121478249	1.218888312376
2	3.497391194192	1.293300750027	1.278778328865
3	3.546234378688	1.296985087723	1.292321849082
4	3.552787665981	1.297721962229	1.294768324493
5	3.553664576139	1.297852933684	1.295153407389
6	3.553781875510	1.297873746555	1.295210684321
7	3.553797565240	1.297876833037	1.295219040349
8	3.553799663854	1.297877272289	1.295220254587
9	3.553799944558	1.297877333278	1.295220431286
10	3.553799982104	1.297877341621	1.295220457067
11	3.553799987126	1.297877342752	1.295220460837
12	3.553799987798	1.297877342904	1.295220461389
13	3.553799987888	1.297877342925	1.295220461470
14	3.553799987900	1.297877342928	1.295220461482
≥ 15	3.553799987902	1.297877342928	1.295220461484

Tabla 9 Probabilidad de falsa alarma: $1 - p = 0$

Para completar el estudio, hemos considerado los valores steady-state de las varianzas de los errores asociados al filtro lineal, cuadrático y cúbico. La Tabla 10 muestra dichos valores para diferentes probabilidades de falsa alarma, poniendo de manifiesto, una vez más, que la utilización de un estimador polinomial de grado $\nu > 1$ puede mejorar, considerablemente, al tradicional estimador lineal. Notemos que, al igual que en el caso anterior, estos resultados obtenidos en el análisis steady-state coinciden con los establecidos en las Tablas 6-9, ya que como indicamos antes, a partir de una cierta iteración, los valores de las varianzas de los errores de estimación se estabilizan.

p	Lineal \bar{p}	Cuadrático \bar{P}_{11}^2	Cúbico \bar{P}_{11}^3
1/4	7.754795047568	6.781660434891	5.068409117140
1/2	6.755555555556	4.982088492481	4.474664230834
3/4	5.235027890751	3.114231391541	3.009048784389
1	3.553799987902	1.297877342928	1.295220461484

Tabla 10 \bar{p} varianza steady-state del error del filtro lineal

\bar{P}_{11}^2 varianza steady-state del error del filtro polinomial de segundo grado

\bar{P}_{11}^3 varianza steady-state del error del filtro polinomial de tercer grado

Para concluir señalemos que, aunque los Casos 1 y 2 se comportan de forma análoga en lo que se refiere al comportamiento de los filtros (lineal, cuadrático y cúbico) y también a los distintos valores considerados para la probabilidad de falsa alarma, los resultados obtenidos no permiten hacer ningún comentario respecto a la relación entre ambos casos.

En efecto, comparando las Tablas 5 y 10 observamos que, sólo para el filtro cuadrático, las varianzas de los errores del Caso 1 son menores que las del Caso 2 para cualquier valor de p . Para los filtros lineal y cúbico, estos valores son menores o mayores, dependiendo de la probabilidad de falsa alarma que se considere.

Apéndice A

Algunos Resultados de Cálculo Matricial

A.1 Producto Kronecker de Matrices

Sea A una matriz $n \times m$ y B una matriz $p \times q$. Se define el *producto Kronecker* de las matrices A y B como una matriz $np \times mq$ dada por

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1m}B \\ \dots & \dots & \dots \\ a_{n1}B & \dots & a_{nm}B \end{pmatrix}$$

donde a_{ij} son las componentes de la matriz A . Para $\nu > 2$, notaremos mediante $A^{[\nu]}$ la matriz $A^{[\nu-1]} \otimes A$.

El producto Kronecker de dos vectores $u \in \mathbb{R}^n$ y $v \in \mathbb{R}^p$ se define de forma análoga, sin más que considerar los vectores como matrices de dimensiones $n \times 1$ y $p \times 1$,

respectivamente. De este modo, se tiene que

$$u \otimes v = \begin{pmatrix} u_1 v \\ \vdots \\ u_n v \end{pmatrix} \in \mathbb{R}^{np}$$

siendo u_1, \dots, u_n las componentes del vector u .

Si A, B, C y D son matrices de dimensiones adecuadas y $\alpha \in \mathbb{R}$, se pueden demostrar las siguientes propiedades (Díaz y Gutiérrez [10], Magnus y Neudecker [26])

- (i) $A \otimes B \otimes C = A \otimes (B \otimes C) = (A \otimes B) \otimes C$
- (ii) $(A + B) \otimes (C + D) = A \otimes C + A \otimes D + B \otimes C + B \otimes D$
- (iii) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$
- (iv) $\alpha \otimes A = \alpha A = A\alpha = A \otimes \alpha$
- (v) $(A \otimes B)^T = (A^T \otimes B^T)$
- (vi) $(A \otimes B)^{-1} = (A^{-1} \otimes B^{-1})$

(si las matrices A y B son singulares esta propiedad sigue siendo válida sin más que utilizar la pseudoinversa de Moore-Penrose).

- (vii) Si $A \in \mathbb{R}^{n \times n}$ y $B \in \mathbb{R}^{p \times p}$ son matrices con autovalores $\lambda_1, \dots, \lambda_n$ y $\delta_1, \dots, \delta_p$, respectivamente, entonces los np autovalores de $A \otimes B$ están dados por

$$\lambda_i \delta_j, \quad i = 1, \dots, n, \quad j = 1, \dots, p.$$

A.2 Operador *vec* o *stack*

Se define la operación *vec* o *stack* de una matriz $A = ((a_{ij}))_{\substack{i=1, \dots, n \\ j=1, \dots, m}}$ como

$$\text{vec}(A) = (a_{11}, \dots, a_{n1}, \dots, a_{1m}, \dots, a_{nm})^T.$$

La relación entre el producto Kronecker de vectores y el operador vec es la siguiente

$$u \otimes v = vec(vu^T)$$

siendo u y v vectores de dimensiones arbitrarias. Esta propiedad se puede generalizar al caso matricial, teniéndose que

$$vec(ABC) = (C^T \otimes A) vec(B)$$

siendo A, B, C matrices de dimensiones apropiadas.

A.3 Matriz Conmutación o Permutación

Dada una matriz A de dimensión $n \times m$, los vectores $vec(A)$ y $vec(A^T)$ se pueden obtener uno a partir del otro mediante una permutación adecuada de las componentes. Existe, por tanto, una única matriz que notaremos $K_{n,m}$, de dimensión $nm \times nm$, llamada *matriz conmutación o permutación*, tal que

$$K_{n,m}vec(A) = vec(A^T)$$

(si $m = n$, suele utilizarse la notación K_{n^2} en lugar de $K_{n,n}$).

Se puede demostrar que

$$K_{n,m}^T = K_{n,m}^{-1} = K_{m,n}.$$

Además, es claro que

$$K_{n,1} = K_{1,n} = I_{n,1}$$

donde $I_{n,l}$ es la matriz identidad $n^l \times n^l$.

La propiedad clave de la matriz conmutación es que permite “permutar” el producto Kronecker (de ahí el nombre de matriz conmutación o permutación). Concretamente,

se verifica que

$$\begin{aligned} K_{p,n}(A \otimes B) &= (B \otimes A) K_{q,m} \\ K_{p,n}(A \otimes B) K_{m,q} &= B \otimes A \\ K_{p,n}(A \otimes v) &= v \otimes A \\ K_{n,p}(v \otimes A) &= A \otimes v \\ K_{n,p}(v \otimes u) &= u \otimes v \end{aligned}$$

siendo A, B matrices de dimensión $n \times m$ y $p \times q$, respectivamente, y u, v vectores de dimensión $n \times 1$ y $p \times 1$, respectivamente.

Un estudio más completo sobre la matriz conmutación y otros aspectos de cálculo matricial puede encontrarse en Díaz y Gutiérrez [10] o en Magnus y Neudecker [26].

A.4 Potencia Kronecker de un Binomio

Sean a, b vectores de dimensión $n \times 1$ y $h \geq 0$ un entero cualquiera. Se puede demostrar (véase Carravetta et al. [4], páginas 1669-1670) el siguiente resultado.

Lema A.4.1 *Para cualquier entero $h \geq 0$, los coeficientes matriciales de la potencia del binomio*

$$(a + b)^{[h]} = \sum_{k=0}^h M_k^h(n) (a^{[k]} \otimes b^{[h-k]})$$

constituyen un conjunto de matrices $\{M_0^h(n), M_1^h(n), \dots, M_h^h(n)\}$ tal que

$$\begin{aligned} M_0^h(n) &= M_h^h(n) = I_{n,h} \\ M_k^h(n) &= (M_k^{h-1}(n) \otimes I_{n,1}) + (M_{k-1}^{h-1}(n) \otimes G_{h-k}), \quad 1 \leq k \leq h-1 \end{aligned}$$

donde

$$G_l = (I_{n,1} \otimes G_{l-1})(G_1 \otimes I_{n,l-1}), \quad G_1 = K_{n^2}^T.$$

Apéndice B

Algoritmo de Predicción y Filtrado para Sistemas con Ruido de la Observación Correlado

Consideremos el siguiente sistema lineal con vector estado n -dimensional, $x(k)$, y vector observación m -dimensional, $z(k)$

$$x(k+1) = \Phi(k)x(k) + w(k), \quad k \geq 0 \quad (\text{B.1})$$

$$x(0) = x_0$$

$$z(k) = H(k)x(k) + v(k), \quad k \geq 0. \quad (\text{B.2})$$

Sobre los ruidos, $\{w(k); k \geq 0\}$ y $\{v(k); k \geq 0\}$, y el estado inicial, x_0 , de este sistema imponemos las siguientes hipótesis:

1. $\{w(k); k \geq 0\}$ es un proceso ruido blanco n -dimensional centrado con

$$E\{w(k)w^T(k)\} = Q(k).$$

2. $\{v(k); k \geq 0\}$ es un proceso m -dimensional centrado con

$$E \{v(k) v^T(j)\} = R(k, j), \quad k, j \geq 0.$$

3. El estado inicial x_0 es un vector aleatorio centrado con

$$E \{x_0 x_0^T\} = P_0.$$

4. Los ruidos $\{w(k); k \geq 0\}$ y $\{v(k); k \geq 0\}$ son incorrelados entre sí e incorrelados con el estado inicial x_0 .

Nuestro objetivo es obtener un algoritmo de filtrado lineal para el sistema (B.1)-(B.2) bajo las hipótesis 1-4 antes mencionadas. Para ello haremos uso de las siguientes propiedades (consecuencias directas del Teorema de la Proyección Ortogonal): si $y(k) \in \mathcal{L}_{\mathbb{R}^n}^2(\Omega, A, P)$, y notamos mediante $\hat{y}(k/j)$ el estimador lineal de menor error cuadrático medio de $y(k)$ basado en $z(0), \dots, z(j)$, entonces

- $\hat{y}(k/j)$ es el único elemento del subespacio generado mediante transformaciones lineales de $z(0), \dots, z(j)$ que verifica

$$E \{\tilde{y}(k/j) z^T(\alpha)\} = 0, \quad \alpha = 0, \dots, j$$

siendo $\tilde{y}(k/j) = y(k) - \hat{y}(k/j)$.

- $\hat{y}(k/j) = \hat{y}(k/j-1) + F_y(k, j)[z(j) - \hat{z}(j/j-1)]$, siendo $F_y(k, j)$ una matriz a determinar.

B.1 Predictor en Función del Filtro

Como ya hemos señalado, el Teorema de la Proyección Ortogonal garantiza que $\hat{x}(k+1/k)$ es la única combinación lineal de las observaciones $z(0), \dots, z(k)$ tal que

$$E \{x(k+1) z^T(j)\} = E \{\hat{x}(k+1/k) z^T(j)\}, \quad j = 0, \dots, k.$$

Ahora bien, utilizando la ecuación del estado (B.1) y la incorrelación de los ruidos y la condición inicial, tenemos que

$$\begin{aligned} E \{x(k+1) z^T(j)\} &= \Phi(k) E \{x(k) z^T(j)\} + E \{w(k) z^T(j)\} \\ &= \Phi(k) E \{\hat{x}(k/k) z^T(j)\}, \quad j = 0, \dots, k \end{aligned}$$

y, por tanto,

$$\begin{aligned} \hat{x}(k+1/k) &= \Phi(k) \hat{x}(k/k), \quad k \geq 0 \\ \hat{x}(0/-1) &= E \{x_0\} = 0. \end{aligned}$$

De esta relación entre el predictor y el filtro y de la ecuación del estado, se deduce la siguiente expresión para el error de predicción

$$\begin{aligned} \tilde{x}(k+1/k) &= \Phi(k) \tilde{x}(k/k) + w(k), \quad k \geq 0 \\ \tilde{x}(0/-1) &= x_0. \end{aligned}$$

Su matriz de covarianzas, $P(k+1/k) = E \{\tilde{x}(k+1/k) \tilde{x}^T(k+1/k)\}$, teniendo en cuenta las hipótesis de incorrelación impuestas al sistema, viene dada por

$$\begin{aligned} P(k+1/k) &= \Phi(k) P(k/k) \Phi^T(k) + Q(k), \quad k \geq 0 \\ P(0/-1) &= P_0 \end{aligned}$$

siendo

$$P(k/k) = E \{\tilde{x}(k/k) \tilde{x}^T(k/k)\}, \quad k \geq 0.$$

B.2 Proceso Innovación

En virtud del Teorema de la Proyección Ortogonal, $\hat{z}(k/k-1)$ es la única combinación lineal de las observaciones $z(0), \dots, z(k-1)$ que verifica

$$E \{z(k) z^T(j)\} = E \{\hat{z}(k/k-1) z^T(j)\}, \quad j = 0, \dots, k-1.$$

Teniendo en cuenta la ecuación de la observación (B.2) y, aplicando de nuevo el Teorema de la Proyección Ortogonal, se tiene que

$$\begin{aligned} E \{z(k) z^T(j)\} &= H(k) E \{x(k) z^T(j)\} + E \{v(k) z^T(j)\} \\ &= H(k) E \{\hat{x}(k/k-1) z^T(j)\} + E \{\hat{v}(k/k-1) z^T(j)\} \end{aligned}$$

y concluimos que

$$\begin{aligned} \hat{z}(k/k-1) &= H(k) \hat{x}(k/k-1) + \hat{v}(k/k-1), \quad k > 0 \\ \hat{z}(0/-1) &= E \{z(0)\} = 0. \end{aligned}$$

Así, el proceso innovación del problema que nos ocupa viene dado por

$$\begin{aligned} \tilde{z}(k/k-1) &= H(k) \tilde{x}(k/k-1) + \tilde{v}(k/k-1), \quad k > 0 \\ \tilde{z}(0/-1) &= z(0). \end{aligned}$$

Su matriz de covarianzas, $\Pi(k) = E \{\tilde{z}(k/k-1) \tilde{z}^T(k/k-1)\}$, está dada por

$$\begin{aligned} \Pi(k) &= H(k) P(k/k-1) H^T(k) + P_{vv}(k, k/k-1) \\ &\quad + H(k) P_v^T(k, k) + P_v(k, k) H^T(k), \quad k \geq 0 \end{aligned}$$

siendo

$$\begin{aligned} P_v(k, l) &= E \{\tilde{v}(k/l-1) \tilde{x}^T(l/l-1)\}, \quad l \leq k \\ P_{vv}(k, l/j) &= E \{\tilde{v}(k/j) \tilde{v}^T(l/j)\}, \quad j < l \leq k. \end{aligned}$$

En efecto, puesto que los estimadores son ortogonales a los errores de estimación, tenemos que

$$E \{\hat{z}(k/k-1) \tilde{z}^T(k/k-1)\} = 0,$$

por tanto,

$$\Pi(k) = E \{z(k) \tilde{z}^T(k/k-1)\}$$

y, utilizando la ecuación de la observación, obtenemos que

$$\Pi(k) = H(k) E \{x(k) \tilde{z}^T(k/k-1)\} + E \{v(k) \tilde{z}^T(k/k-1)\}.$$

Sustituyendo ahora la expresión obtenida para $\tilde{z}(k/k-1)$, podemos escribir

$$\begin{aligned} \Pi(k) &= H(k) E \{x(k) \tilde{x}^T(k/k-1)\} H^T(k) + E \{v(k) \tilde{v}^T(k/k-1)\} \\ &\quad + H(k) E \{x(k) \tilde{v}^T(k/k-1)\} + E \{v(k) \tilde{x}^T(k/k-1)\} H^T(k). \end{aligned}$$

Por último, aplicando de nuevo que los estimadores y los errores de estimación son ortogonales, concluimos que

$$\begin{aligned} \Pi(k) &= H(k) E \{\tilde{x}(k/k-1) \tilde{x}^T(k/k-1)\} H^T(k) + E \{\tilde{v}(k/k-1) \tilde{v}^T(k/k-1)\} \\ &\quad + H(k) E \{\tilde{x}(k/k-1) \tilde{v}^T(k/k-1)\} + E \{\tilde{v}(k/k-1) \tilde{x}^T(k/k-1)\} H^T(k) \\ &= H(k) P(k/k-1) H^T(k) + P_{vv}(k, k/k-1) + H(k) P_v^T(k, k) + P_v(k, k) H^T(k). \end{aligned}$$

B.3 Filtro en Función del Predictor

Como antes señalábamos, el Teorema de la Proyección Ortogonal garantiza que

$$\hat{x}(k/k) = \hat{x}(k/k-1) + F_x(k, k) \tilde{z}(k/k-1), \quad k \geq 0.$$

Para determinar la matriz de ganancia $F_x(k, k)$, consideremos el error de estimación

$$\tilde{x}(k/k) = \tilde{x}(k/k-1) - F_x(k, k) \tilde{z}(k/k-1), \quad k \geq 0.$$

Puesto que $\tilde{x}(k/k)$ es ortogonal a $z(k)$, se tiene que

$$\begin{aligned} 0 &= E \{\tilde{x}(k/k-1) z^T(k)\} - F_x(k, k) E \{\tilde{z}(k/k-1) z^T(k)\} \\ &= E \{\tilde{x}(k/k-1) z^T(k)\} - F_x(k, k) \Pi(k). \end{aligned}$$

Por tanto,

$$F_x(k, k) = E \{ \tilde{x}(k/k-1) z^T(k) \} \Pi^-(k), \quad k \geq 0^4.$$

Por otra parte, teniendo en cuenta la ecuación (B.2) y la ortogonalidad de los estimadores y los errores de estimación, se tiene que

$$\begin{aligned} E \{ \tilde{x}(k/k-1) z^T(k) \} &= E \{ \tilde{x}(k/k-1) x^T(k) \} H^T(k) + E \{ \tilde{x}(k/k-1) v^T(k) \} \\ &= P(k/k-1) H^T(k) + P_v^T(k, k). \end{aligned}$$

Por consiguiente,

$$F_x(k, k) = [P(k/k-1) H^T(k) + P_v^T(k, k)] \Pi^-(k), \quad k \geq 0.$$

Por último, teniendo en cuenta la expresión del error de filtrado, $\tilde{x}(k/k)$, se deduce fácilmente que su matriz de covarianzas, $P(k/k) = E \{ \tilde{x}(k/k) \tilde{x}^T(k/k) \}$, verifica

$$P(k/k) = P(k/k-1) - F_x(k, k) \Pi(k) F_x^T(k, k), \quad k \geq 0.$$

B.4 Predictor del Ruido de la Observación

Puesto que la innovación $\tilde{z}(k/k-1)$ depende del predictor del ruido de la observación, $\hat{v}(k/k-1)$, para establecer el algoritmo de filtrado, es necesario obtener una expresión que nos permita obtener $\hat{v}(k/k-1)$.

Apliquemos, de nuevo el Teorema de la Proyección Ortogonal, para afirmar que

$$\begin{aligned} \hat{v}(k/l) &= \hat{v}(k/l-1) + F_v(k, l) \tilde{z}(l/l-1), \quad l < k \\ \hat{v}(k/-1) &= E \{ v(k) \} = 0. \end{aligned}$$

Para determinar la ganancia $F_v(k, l)$, basta considerar el error de estimación

$$\tilde{v}(k/l) = \tilde{v}(k/l-1) - F_v(k, l) \tilde{z}(l/l-1), \quad l < k$$

⁴Dada una matriz A , arbitraria, denotamos mediante A^- su pseudoinversa de Moore-Penrose (véase Magnus y Neudecker [26]).

y, procediendo como en el apartado anterior, obtenemos que

$$F_v(k, l) = E \{ \tilde{v}(k/l-1) z^T(l) \} \Pi^-(l)$$

con

$$\begin{aligned} E \{ \tilde{v}(k/l-1) z^T(l) \} &= E \{ \tilde{v}(k/l-1) x^T(l) \} H^T(l) + E \{ \tilde{v}(k/l-1) v^T(l) \} \\ &= E \{ \tilde{v}(k/l-1) \tilde{x}^T(l/l-1) \} H^T(l) + E \{ \tilde{v}(k/l-1) \tilde{v}^T(l/l-1) \} \\ &= P_v(k, l) H^T(l) + P_{vv}(k, l/l-1). \end{aligned}$$

Por consiguiente,

$$F_v(k, l) = [P_v(k, l) H^T(l) + P_{vv}(k, l/l-1)] \Pi^-(l), \quad l < k.$$

Expresión para $P_v(k, l)$ y $P_{vv}(k, l/j)$

Para completar el algoritmo de filtrado es necesario dar una expresión para calcular las matrices

$$\begin{aligned} P_v(k, l) &= E \{ \tilde{v}(k/l-1) \tilde{x}^T(l/l-1) \}, \quad l \leq k \\ P_{vv}(k, l/j) &= E \{ \tilde{v}(k/j) \tilde{v}^T(l/j) \}, \quad j < l \leq k. \end{aligned}$$

Haciendo uso de la ortogonalidad de los estimadores y los errores de estimación, y teniendo en cuenta que

$$\begin{aligned} \tilde{x}(l/l-1) &= \Phi(l-1) \tilde{x}(l-1/l-2) + w(l-1) \\ &\quad - \Phi(l-1) F_x(l-1, l-1) \tilde{z}(l-1/l-2) \end{aligned}$$

podemos escribir

$$\begin{aligned} P_v(k, l) &= E \{ v(k) \tilde{x}^T(l/l-1) \} \\ &= E \{ v(k) \tilde{x}^T(l-1/l-2) \} \Phi^T(l-1) \end{aligned}$$

$$\begin{aligned}
& -E \{v(k) \tilde{z}^T(l-1/l-2)\} F_x^T(l-1, l-1) \Phi^T(l-1) \\
= & P_v(k, l-1) \Phi^T(l-1) \\
& -E \{\tilde{v}(k/l-2) \tilde{z}^T(l-1/l-2)\} F_x^T(l-1, l-1) \Phi^T(l-1)
\end{aligned}$$

y, puesto que $F_v(k, l-1) = E \{\tilde{v}(k/l-2) \tilde{z}^T(l-1/l-2)\} \Pi^-(l-1)$, tenemos que

$$\begin{aligned}
P_v(k, l) = & P_v(k, l-1) \Phi^T(l-1) \\
& -F_v(k, l-1) \Pi(l-1) F_x^T(l-1, l-1) \Phi^T(l-1), \quad 0 < l \leq k.
\end{aligned}$$

La condición inicial es

$$P_v(k, 0) = E \{\tilde{v}(k/k-1) \tilde{x}^T(0/-1)\} = E \{v(k) x_0\} = 0.$$

Por otra parte, utilizando que

$$\tilde{v}(l/j) = \tilde{v}(l/j-1) - F_v(l, j) \tilde{z}(j/j-1)$$

tenemos que

$$\begin{aligned}
P_{vv}(k, l/j) &= E \{\tilde{v}(k/j) \tilde{v}^T(l/j-1)\} - E \{\tilde{v}(k/j) \tilde{z}^T(j/j-1)\} F_v^T(l, j) \\
&= E \{\tilde{v}(k/j-1) \tilde{v}^T(l/j-1)\} - F_v(k, j) E \{\tilde{z}(j/j-1) \tilde{v}^T(l/j-1)\} \\
&\quad - E \{\tilde{v}(k/j-1) \tilde{z}^T(j/j-1)\} F_v^T(l, j) \\
&\quad + F_v(k, j) E \{\tilde{z}(j/j-1) \tilde{z}^T(j/j-1)\} F_v^T(l, j)
\end{aligned}$$

y, utilizando de nuevo que $F_v(k, j) = E \{\tilde{v}(k/j-1) \tilde{z}^T(j/j-1)\} \Pi^-(j)$, concluimos que

$$P_{vv}(k, l/j) = P_{vv}(k, l/j-1) - F_v(k, j) \Pi(j) F_v^T(l, j), \quad 0 \leq j < l \leq k.$$

La condición inicial es

$$P_{vv}(k, l/-1) = E \{\tilde{v}(k/-1) \tilde{v}^T(l/-1)\} = E \{v(k) v^T(l)\} = R(k, l).$$

B.5 Algoritmo para el Problema de Filtrado

Los resultados obtenidos se resumen en el siguiente Teorema.

Teorema B.5.1 *El algoritmo para los problemas de predicción y filtrado del estado $x(k)$ del sistema (B.1)-(B.2) viene dado por*

$$\begin{aligned}\hat{x}(k/k) &= \hat{x}(k/k-1) + F_x(k,k) \tilde{z}(k/k-1), & k \geq 0 \\ \hat{x}(k+1/k) &= \Phi(k) \hat{x}(k/k), & k \geq 0 \\ \hat{x}(0/-1) &= 0\end{aligned}$$

siendo $\tilde{z}(k/k-1)$ la innovación

$$\tilde{z}(k/k-1) = z(k) - H(k) \hat{x}(k/k-1) - \hat{v}(k/k-1), \quad k \geq 0.$$

Para cada $k > 0$ fijo, $\hat{v}(k/k-1)$ se calcula mediante

$$\begin{aligned}\hat{v}(k/l) &= \hat{v}(k/l-1) + F_v(k,l) \tilde{z}(l/l-1), & l < k \\ \hat{v}(k/-1) &= 0\end{aligned}$$

donde la ganancia $F_v(k,l)$ viene dada por

$$F_v(k,l) = [P_v(k,l) H^T(l) + P_{vv}(k,l/l-1)] \Pi^-(l), \quad l < k$$

siendo

$$\begin{aligned}P_v(k,l) &= P_v(k,l-1) \Phi^T(l-1) \\ &\quad - F_v(k,l-1) \Pi(l-1) F_x^T(l-1,l-1) \Phi^T(l-1), \quad 0 < l \leq k \\ P_v(k,0) &= 0\end{aligned}$$

y

$$\begin{aligned}P_{vv}(k,l/j) &= P_{vv}(k,l/j-1) - F_v(k,j) \Pi(j) F_v^T(l,j), \quad 0 \leq j < l \leq k \\ P_{vv}(k,l/-1) &= R(k,l).\end{aligned}$$

La matriz de ganancia $F_x(k, k)$ está dada por

$$F_x(k, k) = [P(k/k-1)H^T(k) + P_v^T(k, k)]\Pi^-(k), \quad k \geq 0.$$

La matriz de covarianzas del error de estimación se calcula mediante la siguiente relación recursiva

$$P(k/k) = P(k/k-1) - F_x(k, k)\Pi(k)F_x^T(k, k), \quad k \geq 0$$

$$P(k+1/k) = \Phi(k)P(k/k)\Phi^T(k) + Q(k), \quad k \geq 0$$

$$P(0/-1) = P_0$$

siendo $\Pi(k)$ la matriz de covarianzas del proceso innovación,

$$\Pi(k) = H(k)P(k/k-1)H^T(k) + P_{vv}(k, k/k-1) + H(k)P_v^T(k, k) + P_v(k, k)H^T(k).$$

Apéndice C

Algoritmo de Filtrado para Sistemas con Observaciones Inciertas y Ruidos Correlados

Consideremos el siguiente sistema lineal con observaciones inciertas

$$x(k+1) = \Phi(k)x(k) + U_k + w(k), \quad k \geq 0 \quad (\text{C.1})$$

$$x(0) = x_0$$

$$z(k) = u(k)H(k)x(k) + V_k + v(k), \quad k \geq 0 \quad (\text{C.2})$$

con vector estado n -dimensional, $x(k)$, y vector observación m -dimensional, $z(k)$. $\Phi(k)$, $H(k)$, U_k , V_k son matrices determinísticas conocidas de dimensiones apropiadas.

Imponemos las siguientes hipótesis sobre los ruidos, $\{w(k); k \geq 0\}$, $\{v(k); k \geq 0\}$, $\{u(k); k \geq 0\}$, y el estado inicial, x_0 , de este sistema:

1. El estado inicial x_0 es un vector aleatorio con

$$\begin{aligned} E\{x_0\} &= \bar{x}_0 \\ E\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\} &= P_0. \end{aligned}$$

2. $\{w(k); k \geq 0\}$ es un proceso ruido blanco n -dimensional centrado con

$$E\{w(k)w^T(k)\} = Q(k), \quad k \geq 0.$$

3. $\{v(k); k \geq 0\}$ es un proceso ruido blanco m -dimensional centrado con

$$E\{v(k)v^T(k)\} = R(k), \quad k \geq 0.$$

4. $\{u(k); k \geq 0\}$ es una sucesión de variables aleatorias independientes de Bernoulli, con $P[u(k) = 1] = p(k)$.

5. Los ruidos $\{w(k); k \geq 0\}$ y $\{v(k); k \geq 0\}$ son incorrelados con el estado inicial x_0 y verifican

$$E\{w(k)v^T(j)\} = S(k)\delta_{kj}, \quad k, j \geq 0.$$

6. $\{u(k); k \geq 0\}$ es independiente de $(x_0, \{w(k); k \geq 0\}, \{v(k); k \geq 0\})$.

A continuación se recogen algunas propiedades interesantes, que serán básicas en desarrollos posteriores.

Proposición C.1 *Bajo las hipótesis 1-6 impuestas sobre el modelo, el ruido $\{w(k); k \geq 0\}$ verifica las siguientes propiedades:*

(i) Para cada k , $E\{w(k)x^T(j)\} = 0$, $j \leq k$.

(ii) Para cada k , $E\{w(k)z^T(j)\} = 0$, $j < k$.

-Demostración-

- (i) Puesto que $x(j)$ es función lineal de $x_0, w(0), \dots, w(j-1)$, el apartado (i) es inmediato teniendo en cuenta que, para cada k , $w(k)$ es centrado e incorrelado con $x_0, w(0), \dots, w(k-1)$.
- (ii) Para probar este apartado, observemos que, si $j < k$, la hipótesis 6 garantiza que $u(j)$ es independiente de $(w(k), x(j))$ y, por tanto,

$$E \{w(k) z^T(j)\} = E \{u(j)\} E \{w(k) x^T(j)\} H^T(j) + E \{w(k)\} V_j^T + E \{w(k) v^T(j)\}.$$

Utilizando ahora el apartado (i) y que $w(k)$ es centrado e incorrelado con $v(j)$, concluimos que $E \{w(k) z^T(j)\} = 0$. ■

Proposición C.2 *Bajo las hipótesis de la Proposición anterior, el ruido $\{v(k); k \geq 0\}$ verifica:*

- (i) *Para cada k , $E \{v(k) x^T(j)\} = 0$, $j \leq k$.*
- (ii) *Para cada k , $E \{v(k) u(j) x^T(j)\} = 0$, $j \leq k$.*
- (iii) *Para cada k , $E \{v(k) z^T(j)\} = 0$, $j < k$.*

-Demostración-

- (i) Este apartado se deduce de forma inmediata teniendo en cuenta que $x(j)$ es función lineal de $x_0, w(0), \dots, w(j-1)$ y utilizando que, para cada k , $v(k)$ es centrado e incorrelado con $x_0, w(0), \dots, w(k-1)$.
- (ii) Para demostrar este apartado, basta observar que, para $j \leq k$, $u(j)$ es independiente de $(v(k), x(j))$ y aplicar el apartado (i).

(iii) Por último, (iii) se demuestra fácilmente, teniendo en cuenta el apartado (ii) y que el ruido $\{v(k); k \geq 0\}$ es blanco y centrado. ■

Proposición C.3 *Bajo las hipótesis impuestas sobre el modelo, para cada k , la variable $u(k)$ es independiente de $(x(k), w(k))$ e incorrelada con $x(k) z^T(j)$ para $j < k$.*

-Demostración-

Claramente, $u(k)$ es independiente de $(x(k), w(k))$ para cada k , en virtud de la hipótesis 6 de independencia impuesta sobre el modelo.

Por otro lado, para $j < k$,

$$\begin{aligned} E\{u(k) x(k) z^T(j)\} &= E\{u(k) u(j)\} E\{x(k) x^T(j)\} H^T(j) \\ &\quad + E\{u(k)\} E\{x(k)\} V_j^T + E\{u(k)\} E\{x(k) v^T(j)\} \end{aligned}$$

ya que $\{u(k); k \geq 0\}$ es independiente de $(x_0, \{w(k); k \geq 0\}, \{v(k); k \geq 0\})$. Utilizando ahora que $u(k)$ y $u(j)$ son independientes y, a su vez, $u(j)$ es independiente de $\{x(k); k \geq 0\}$, concluimos que

$$\begin{aligned} E\{u(k) x(k) z^T(j)\} &= E\{u(k)\} E\{u(j) x(k) x^T(j) H^T(j) + x(k) V_j^T + x(k) v^T(j)\} \\ &= E\{u(k)\} E\{x(k) z^T(j)\} \end{aligned}$$

y, por tanto, $u(k)$ es incorrelado con el producto $x(k) z^T(j)$. ■

Nuestro objetivo es obtener un algoritmo de filtrado lineal para el sistema (C.1)-(C.2) bajo las hipótesis 1-6 antes mencionadas. Para ello haremos uso de las Proposiciones C.1-C.3 y de las siguientes propiedades (consecuencias directas del Teorema de la Proyección Ortogonal): si $y(k) \in \mathcal{L}_{\mathbb{R}^n}^2(\Omega, A, P)$, y notamos mediante $\hat{y}(k/j)$ el estimador lineal de menor error cuadrático medio de $y(k)$ basado en $z(0), \dots, z(j)$, entonces

- $\hat{y}(k/j)$ es el único elemento del subespacio generado mediante transformaciones lineales de $z(0), \dots, z(j)$ que verifica

$$E \{ \tilde{y}(k/j) z^T(\alpha) \} = 0, \quad \alpha = 0, \dots, j$$

siendo $\tilde{y}(k/j) = y(k) - \hat{y}(k/j)$.

- $\hat{y}(k/j) = \hat{y}(k/j-1) + F_y(k, j) [z(j) - \hat{z}(j/j-1)]$, siendo $F_y(k, j)$ una matriz a determinar.

C.1 Predictor en Función del Filtro

Como ya hemos señalado, el Teorema de la Proyección Ortogonal garantiza que $\hat{x}(k+1/k)$ es la única combinación lineal de las observaciones $z(0), \dots, z(k)$ tal que

$$E \{ x(k+1) z^T(j) \} = E \{ \hat{x}(k+1/k) z^T(j) \}, \quad j = 0, \dots, k.$$

Teniendo en cuenta la ecuación del estado (C.1), y aplicando de nuevo el Teorema de la Proyección Ortogonal a $\hat{x}(k/k)$ y $\hat{w}(k/k)$, podemos escribir

$$\begin{aligned} E \{ x(k+1) z^T(j) \} &= \Phi(k) E \{ x(k) z^T(j) \} + U_k E \{ z^T(j) \} + E \{ w(k) z^T(j) \} \\ &= \Phi(k) E \{ \hat{x}(k/k) z^T(j) \} + U_k E \{ z^T(j) \} + E \{ \hat{w}(k/k) z^T(j) \} \end{aligned}$$

concluyendo que

$$\hat{x}(k+1/k) = \Phi(k) \hat{x}(k/k) + U_k + \hat{w}(k/k), \quad k \geq 0$$

$$\hat{x}(0/-1) = E \{ x_0 \} = \bar{x}_0.$$

Haciendo uso de esta relación y de la ecuación del estado, obtenemos la siguiente expresión para el error de predicción

$$\tilde{x}(k+1/k) = \Phi(k) \tilde{x}(k/k) + \tilde{w}(k/k), \quad k \geq 0$$

$$\tilde{x}(0/-1) = x_0 - \bar{x}_0.$$

Su matriz de covarianzas, $P(k+1/k) = E\{\tilde{x}(k+1/k)\tilde{x}^T(k+1/k)\}$, verifica

$$\begin{aligned} P(k+1/k) &= \Phi(k)P(k/k)\Phi^T(k) + P_w(k/k) \\ &\quad + \Phi(k)E\{\tilde{x}(k/k)\tilde{w}^T(k/k)\} + E\{\tilde{w}(k/k)\tilde{x}^T(k/k)\}\Phi^T(k) \\ P(0/-1) &= E\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\} = P_0 \end{aligned}$$

donde

$$\begin{aligned} P(k/k) &= E\{\tilde{x}(k/k)\tilde{x}^T(k/k)\}, \quad k \geq 0 \\ P_w(k/k) &= E\{\tilde{w}(k/k)\tilde{w}^T(k/k)\}, \quad k \geq 0. \end{aligned}$$

C.2 Proceso Innovación

En primer lugar, vamos a obtener una expresión para el predictor de la observación, $\hat{z}(k/k-1)$; aplicando el Teorema de la Proyección Ortogonal dicho predictor es la única combinación lineal de las observaciones $z(0), \dots, z(k-1)$ que verifica

$$E\{z(k)z^T(j)\} = E\{\hat{z}(k/k-1)z^T(j)\}, \quad j = 0, \dots, k-1.$$

Teniendo en cuenta la ecuación de la observación (C.2) y que $u(k)$ es incorrelada con $x(k)z^T(j)$ para $j = 0, \dots, k-1$ (Proposición C.3), se tiene que

$$\begin{aligned} E\{z(k)z^T(j)\} &= p(k)H(k)E\{x(k)z^T(j)\} + V_kE\{z^T(j)\} + E\{v(k)z^T(j)\} \\ &= p(k)H(k)E\{\hat{x}(k/k-1)z^T(j)\} + V_kE\{z^T(j)\} \end{aligned}$$

donde, en la última igualdad, se ha aplicado de nuevo el Teorema de la Proyección Ortogonal y que $v(k)$ es centrado e incorrelado con $z(0), \dots, z(k-1)$ (Proposición C.2, apartado (iii)).

Concluimos, así, que

$$\hat{z}(k/k-1) = p(k)H(k)\hat{x}(k/k-1) + V_k, \quad k \geq 0.$$

Por tanto, utilizando la ecuación de la observación y la expresión anterior del predictor, deducimos que el proceso innovación $\tilde{z}(k/k-1) = z(k) - \hat{z}(k/k-1)$ viene dado por

$$\tilde{z}(k/k-1) = (u(k) - p(k)) H(k) x(k) + p(k) H(k) \tilde{x}(k/k-1) + v(k), \quad k \geq 0.$$

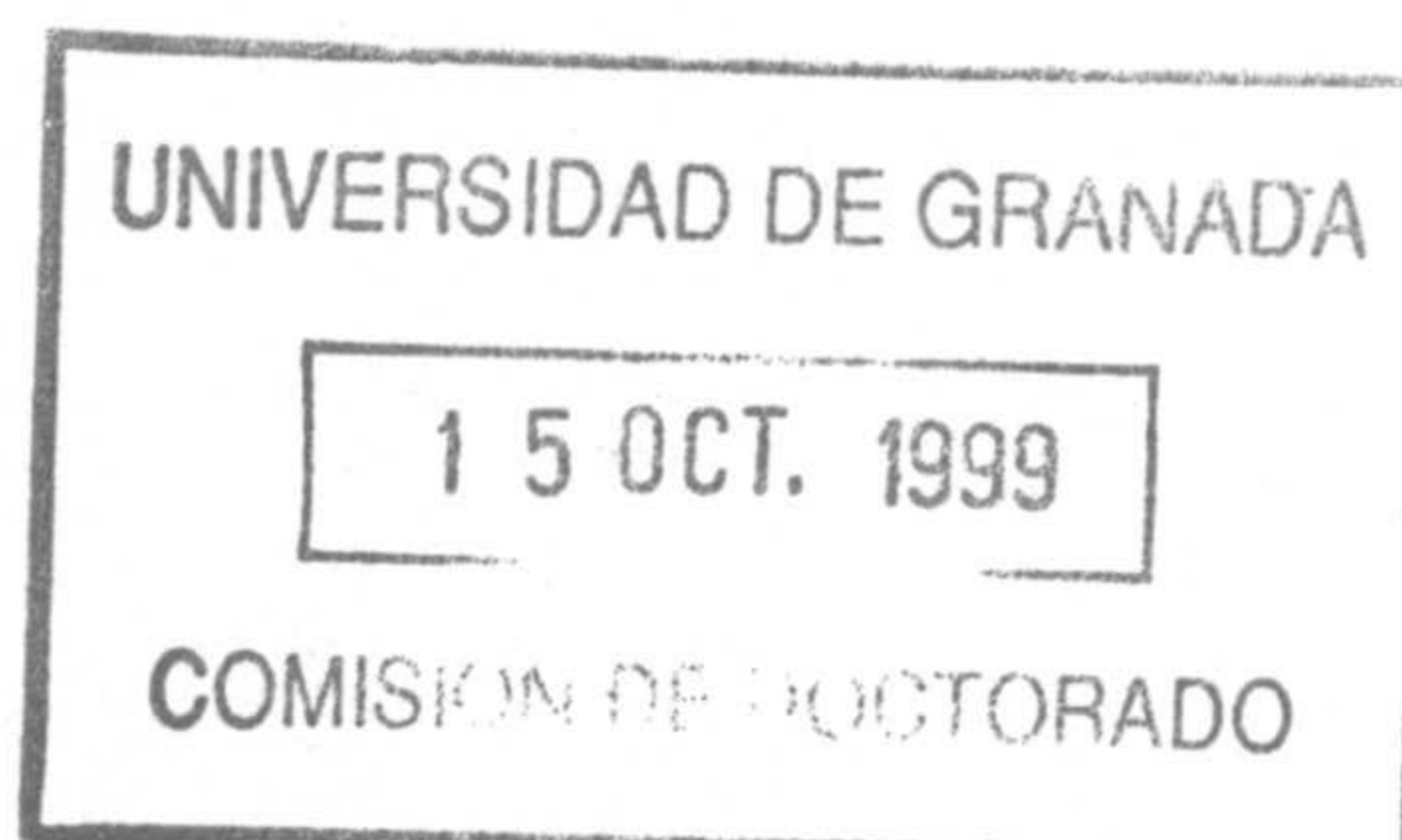
Para obtener su matriz de covarianzas, $\Pi(k) = E\{\tilde{z}(k/k-1)\tilde{z}^T(k/k-1)\}$, utilizamos que $u(k)$ es independiente de $x(k)$ e incorrelada con $x(k)z^T(j)$, para $j < k$ (Proposición C.3), de modo que

$$\begin{aligned} \Pi(k) = & E\{(u(k) - p(k))^2\} H(k) E\{x(k)x^T(k)\} H^T(k) \\ & + p^2(k) H(k) P(k/k-1) H^T(k) + R(k) \\ & + p(k) E\{u(k) - p(k)\} H(k) E\{x(k)\tilde{x}^T(k/k-1)\} H^T(k) \\ & + p(k) E\{u(k) - p(k)\} H(k) E\{\tilde{x}(k/k-1)x^T(k)\} H^T(k) \\ & + E\{(u(k) - p(k)) H(k) x(k) v^T(k)\} + p(k) H(k) E\{\tilde{x}(k/k-1) v^T(k)\} \\ & + E\{(u(k) - p(k)) v(k) x^T(k) H^T(k)\} + p(k) E\{v(k) \tilde{x}^T(k/k-1)\} H^T(k). \end{aligned}$$

Teniendo en cuenta ahora que $v(k)$ es centrado e incorrelado con $u(k)x(k)$ (Proposición C.2, apartado (ii)) y con $x(k), z(0), \dots, z(k-1)$ (Proposición C.2, apartados (i) y (iii)), obtenemos

$$\begin{aligned} \Pi(k) = & p(k)(1 - p(k)) H(k) E\{x(k)x^T(k)\} H^T(k) \\ & + p^2(k) H(k) P(k/k-1) H^T(k) + R(k). \end{aligned}$$

Observemos que, bajo la hipótesis de que $R(k)$ sea definida positiva, la no singularidad de la matriz $\Pi(k)$ está garantizada. Sin esta hipótesis, $\Pi(k)$ podría ser singular; en lo que sigue, notaremos mediante $\Pi^-(k)$ la pseudoinversa de Moore-Penrose de la matriz $\Pi(k)$ (véase Magnus y Neudecker [26]).



C.3 Filtro en Función del Predictor

Como hemos señalado anteriormente, el Teorema de la Proyección Ortogonal garantiza que

$$\hat{x}(k/k) = \hat{x}(k/k-1) + F_x(k, k) \tilde{z}(k/k-1), \quad k \geq 0.$$

Para determinar la matriz de ganancia $F_x(k, k)$, consideremos el error de filtrado

$$\tilde{x}(k/k) = \tilde{x}(k/k-1) - F_x(k, k) \tilde{z}(k/k-1), \quad k \geq 0.$$

Aplicando que $\tilde{x}(k/k)$ es ortogonal a $z(k)$, tenemos que

$$\begin{aligned} 0 &= E \{ \tilde{x}(k/k-1) z^T(k) \} - F_x(k, k) E \{ \tilde{z}(k/k-1) z^T(k) \} \\ &= E \{ \tilde{x}(k/k-1) z^T(k) \} - F_x(k, k) \Pi(k). \end{aligned}$$

Por tanto,

$$F_x(k, k) = E \{ \tilde{x}(k/k-1) z^T(k) \} \Pi^{-1}(k), \quad k \geq 0.$$

Teniendo en cuenta la ecuación de la observación y que $u(k)$ es independiente de $x(k)$ e incorrelada con $x(k) z^T(j)$ para $j < k$, podemos escribir

$$\begin{aligned} E \{ \tilde{x}(k/k-1) z^T(k) \} &= p(k) E \{ \tilde{x}(k/k-1) x^T(k) \} H^T(k) \\ &\quad + E \{ \tilde{x}(k/k-1) \} V_k^T + E \{ \tilde{x}(k/k-1) v^T(k) \}. \end{aligned}$$

Utilizando ahora que $v(k)$ es incorrelado con $x(k), z(0), \dots, z(k-1)$ y que el error de estimación es centrado e incorrelado con las observaciones, en virtud del Teorema de la Proyección Ortogonal, obtenemos que

$$E \{ \tilde{x}(k/k-1) z^T(k) \} = p(k) P(k/k-1) H^T(k)$$

y, por consiguiente,

$$F_x(k, k) = p(k) P(k/k-1) H^T(k) \Pi^{-1}(k), \quad k \geq 0.$$

Teniendo en cuenta la expresión del error $\tilde{x}(k/k)$ y, puesto que

$$E \{ \tilde{x}(k/k-1) \tilde{z}^T(k/k-1) \} = E \{ \tilde{x}(k/k-1) z^T(k) \} = F_x(k, k) \Pi(k),$$

se deduce fácilmente que la matriz de covarianzas $P(k/k) = E \{ \tilde{x}(k/k) \tilde{x}^T(k/k) \}$ verifica

$$P(k/k) = P(k/k-1) - F_x(k, k) \Pi(k) F_x^T(k, k), \quad k \geq 0.$$

C.4 Filtro del Ruido del Estado

Puesto que el predictor $\hat{x}(k+1/k)$ depende del estimador del ruido del estado, $\hat{w}(k/k)$, para establecer el algoritmo de filtrado es necesario obtener una expresión recursiva que nos permita calcular $\hat{w}(k/k)$.

Utilizando la expresión del filtro en función del predictor, podemos escribir

$$\hat{w}(k/k) = \hat{w}(k/k-1) + F_w(k, k) \tilde{z}(k/k-1).$$

Puesto que $w(k)$ es centrado e incorrelado con las observaciones $z(0), \dots, z(k-1)$ (Proposición C.1, apartado (ii)), se tiene que $\hat{w}(k/k-1) = 0$ y, por tanto,

$$\hat{w}(k/k) = F_w(k, k) \tilde{z}(k/k-1).$$

Para determinar la ganancia $F_w(k, k)$, basta considerar el error de filtrado

$$\tilde{w}(k/k) = w(k) - F_w(k, k) \tilde{z}(k/k-1)$$

y, procediendo como en el apartado anterior, se obtiene que

$$F_w(k, k) = E \{ w(k) z^T(k) \} \Pi^-(k).$$

Teniendo en cuenta que $u(k)$ es independiente de $(x(k), w(k))$ (Proposición C.3), podemos escribir

$$\begin{aligned} E \{ w(k) z^T(k) \} &= p(k) E \{ w(k) x^T(k) \} H^T(k) + E \{ w(k) \} V_k^T \\ &\quad + E \{ w(k) v^T(k) \} \end{aligned}$$

y, utilizando que $w(k)$ es centrado e incorrelado con $x(k)$ (Proposición C.1, apartado (i)), obtenemos que $E\{w(k)z^T(k)\} = E\{w(k)v^T(k)\} = S(k)$. Por tanto,

$$F_w(k, k) = S(k)\Pi^-(k).$$

Puesto que la matriz de covarianzas del error de predicción, $P(k+1/k)$, depende de la matriz $P_w(k/k)$ y de $E\{\tilde{x}(k/k)\tilde{w}^T(k/k)\}$, para completar el algoritmo es necesario obtener una expresión para estas matrices.

Teniendo en cuenta la expresión del error

$$\tilde{w}(k/k) = w(k) - S(k)\Pi^-(k)\tilde{z}(k/k-1)$$

se deduce que

$$\begin{aligned} P_w(k/k) &= Q(k) + S(k)\Pi^-(k)S^T(k) \\ &\quad - E\{w(k)\tilde{z}^T(k/k-1)\}\Pi^-(k)S^T(k) \\ &\quad - S(k)\Pi^-(k)E\{\tilde{z}(k/k-1)w^T(k)\}. \end{aligned}$$

Al ser $w(k)$ centrado e incorrelado con $z(0), \dots, z(k-1)$,

$$E\{w(k)\tilde{z}^T(k/k-1)\} = E\{w(k)z^T(k)\} = S(k)$$

y, por tanto,

$$P_w(k/k) = Q(k) - S(k)\Pi^-(k)S^T(k).$$

Por otro lado, teniendo en cuenta la expresión del error de filtrado, $\tilde{x}(k/k)$, podemos escribir

$$\begin{aligned} E\{\tilde{x}(k/k)\tilde{w}^T(k/k)\} &= E\{\tilde{x}(k/k)w^T(k)\} \\ &= E\{\tilde{x}(k/k-1)w^T(k)\} - F_x(k, k)E\{\tilde{z}(k/k-1)w^T(k)\} \\ &= -F_x(k, k)S^T(k) \end{aligned}$$

donde se ha utilizado que $w(k)$ es centrado e incorrelado con $x(k), z(0), \dots, z(k-1)$.

C.5 Algoritmo Recursivo

Los resultados obtenidos en los apartados anteriores se resumen en el siguiente teorema.

Teorema C.5.1 *El algoritmo para los problemas de predicción y filtrado del estado, $x(k)$, del sistema (C.1)-(C.2) está dado por*

$$\begin{aligned}\hat{x}(k/k) &= \hat{x}(k/k-1) + F_x(k, k) \tilde{z}(k/k-1), \quad k \geq 0 \\ \hat{x}(k+1/k) &= \Phi(k) \hat{x}(k/k-1) + U_k + [\Phi(k) F_x(k, k) + S(k) \Pi^-(k)] \tilde{z}(k/k-1) \\ \hat{x}(0/-1) &= \bar{x}_0\end{aligned}$$

siendo $\tilde{z}(k/k-1)$ la innovación

$$\tilde{z}(k/k-1) = z(k) - p(k) H(k) \hat{x}(k/k-1) - V_k, \quad k \geq 0.$$

La matriz de ganancia, $F_x(k, k)$, viene dada por

$$F_x(k, k) = p(k) P(k/k-1) H^T(k) \Pi^-(k), \quad k \geq 0.$$

Las matrices de covarianzas de los errores de filtrado y predicción se calculan mediante la siguiente relación recursiva

$$\begin{aligned}P(k/k) &= P(k/k-1) - F_x(k, k) \Pi(k) F_x^T(k, k), \quad k \geq 0 \\ P(k+1/k) &= \Phi(k) P(k/k) \Phi^T(k) + Q(k) - S(k) \Pi^-(k) S^T(k) \\ &\quad - \Phi(k) F_x(k, k) S^T(k) - S(k) F_x^T(k, k) \Phi^T(k) \\ P(0/-1) &= P_0.\end{aligned}$$

donde $\Pi(k)$ representa la matriz de covarianzas del proceso innovación,

$$\Pi(k) = p(k) (1 - p(k)) H(k) D(k) H^T(k) + p^2(k) H(k) P(k/k-1) H^T(k) + R(k),$$

siendo $D(k) = E \{x(k) x^T(k)\}$.

Nota C.5.1 Observemos que, para poder aplicar el algoritmo, es necesario dar una expresión recursiva para calcular $D(k)$. De la ecuación del estado (C.1), y teniendo en cuenta que $w(k)$ es centrado e incorrelado con el estado $x(k)$, se deduce de forma inmediata que

$$\begin{aligned} D(k) &= \Phi(k-1) D(k-1) \Phi^T(k-1) + U_{k-1} U_{k-1}^T + Q(k-1) \\ &\quad + \Phi(k-1) E\{x(k-1)\} U_{k-1}^T + U_{k-1} E\{x^T(k-1)\} \Phi^T(k-1) \\ D(0) &= E\{x_0 x_0^T\} = P_0 + \bar{x}_0 \bar{x}_0^T \end{aligned}$$

con

$$\begin{aligned} E\{x(k)\} &= \Phi(k-1) E\{x(k-1)\} + U_{k-1} \\ E\{x(0)\} &= \bar{x}_0. \end{aligned}$$

Bibliografía

- [1] ANDERSON, B. Y MOORE, J. (1979). *Optimal Filtering*. Prentice Hall. Englewood Cliffs. New Jersey.
- [2] BALAKRISHNAN, A. V. (1984). *Kalman Filtering Theory*. Optimization Software, Inc., Publications Division, New York.
- [3] BRAMMER, K. Y SIFFLING, G. (1989). *Kalman-Bucy Filters*. Artech House.
- [4] CARRAVETTA, F., GERMANI, A. Y RAIMONDI, M. (1996). Polynomial Filtering for Linear Discrete Time Non-Gaussian Systems. *SIAM J. Control and Optimization*. Vol. 34, No. 5, pp.1666-1690.
- [5] CAINES, P.E. (1988). *Linear Stochastic Systems*. John Wiley & Sons, New York.
- [6] CATLIN, D.E. (1989). *Estimation, Control and the Discrete Kalman Filter*. Springer-Verlag, New York.
- [7] CHOW, B.S. Y BIRKEMEIER, W.P. (1989). A New Structure of Recursive Estimator. *IEEE Transactions on Automatic Control*. Vol. AC-34, No. 5, pp. 568-572.
- [8] CHOW, B.S. Y BIRKEMEIER, W.P. (1990). A New Recursive Filter for Systems with Multiplicative Noise. *IEEE Transactions on Automatic Control*. Vol. AC-36, No. 6, pp. 1430-1435.

- [9] DE SANTIS A., GERMANI A. Y RAIMONDI M. (1995). Optimal Quadratic Filtering of Linear Discrete-Time Non-Gaussian Systems. *IEEE Transactions on Automatic Control*. Vol. 40, No. 7, pp. 1274-1278.
- [10] DÍAZ GARCÍA, J.A. Y GUTIÉRREZ JAÍMEZ, R. (1996). *Cálculo Diferencial Matricial y Momentos de Matrices Aleatorias Elípticas*. Servicios de Reprografía. Facultad de Ciencias. Universidad de Granada.
- [11] GARCÍA-LIGERO, M.J., HERMOSO, A. Y LINARES, J. (1997). Second Order Estimation of Signals in Presence of Multiplicative Noise. *Conference Control of Industrial Systems*. IFAC, pp. 653-658.
- [12] GARCÍA-LIGERO, M.J., HERMOSO, A. Y LINARES, J. (1997). Second Order Polynomial Filtering for Discrete Systems with Uncertain Observation. *VIII International Symposium on ASMDA*, pp. 157-162.
- [13] GREWAL, M.S. Y ANDREWS, A.P. (1993). *Kalman Filtering. Theory and Practice*. Prentice Hall, New Jersey.
- [14] HADIDI, M.T. Y SCHWARTZ, S.C. (1979). Linear Recursive State Estimators under Uncertain Observations. *IEEE Transactions on Automatic Control* Vol. AC-24, No. 6, pp. 944-948.
- [15] HERMOSO, A. Y LINARES, J. (1994). Linear Estimation for Discrete-Time Systems in the Presence of Time-Correlated Disturbances and Uncertain Observations. *IEEE Transactions on Automatic Control*. Vol. AC-39, No. 8, pp. 1636-1638.
- [16] HERMOSO, A. Y LINARES, J. (1995). Linear Smoothing for Discrete-Time Systems in the Presence of Correlated Disturbances and Uncertain Observations. *IEEE Transactions on Automatic Control*. Vol. 40, No. 8, pp. 1486-1488.

- [17] HERMOSO, A. Y LINARES, J. (1995). Recursive Algorithms for LMSE Linear Filter and Predictor under Correlated Disturbances and Multiplicative Noise. *VII International Symposium on ASMDA*, pp. 305-311.
- [18] JAFFER, A.G. Y GUPTA, S.C. (1971). Recursive Bayesian Estimation with Uncertain Observation. *IEEE Transactions on Information Theory*. Vol. IT-17, pp. 614-616.
- [19] JAZWINSKI, A.H. (1970). *Stochastic Processes and Filtering Theory*. Academic Press.
- [20] KAILATH, T. (1968). An Innovation Approach to Least Squares Estimation. Part I. Linear Filtering in Additive White Noise. *IEEE Transactions on Automatic Control*. Vol. AC-13, pp. 646-655.
- [21] KAILATH, T. Y FROST, P. (1968). An Innovation Approach to Least Squares Estimation. Part II. Linear Smoothing in Additive White Noise. *IEEE Transactions on Automatic Control*. Vol. AC-13, pp. 655-660.
- [22] KAILATH, T. (1974). A View of Three Decades of Linear Filtering Theory. *IEEE Transactions on Information Theory*. Vol. IT-20, pp. 146-181.
- [23] KAILATH, T. (1980). *Linear Systems*. Prentice Hall. Englewood Cliffs. New Jersey.
- [24] KALMAN, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME. Journal of Basic Engineering*. Ser. D. Vol. 82, pp. 35-45.
- [25] KORBICZ, J. Y BIDYUK, P. (1993). *State and Parameter Estimation*. Higher College Engineering Press. Poland.

- [26] MAGNUS, J.R. Y NEUDECKER, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley & Sons. New York.
- [27] MEDITCH, J.S. (1967). Orthogonal Projection and Discrete Optimal Linear Smoothing. *Journal SIAM on Control*. Vol. 5, No. 1, pp. 74-89.
- [28] MEDITCH, J.S. (1969). *Stochastic Optimal Linear Estimation and Control*. McGraw-Hill, New York.
- [29] MONZINGO, R.A. (1975). Discrete Optimal Linear Smoothing for Systems with Uncertain Observations. *IEEE Transactions on Information Theory*. Vol. IT-21, No. 3, pp. 271-275.
- [30] MONZINGO, R.A. (1981). Discrete Linear Recursive Smoothing for Systems with Uncertain Observations. *IEEE Transactions on Automatic Control*. Vol. AC-26, No. 3, pp. 754-757.
- [31] NAHI, N.E. (1969). Optimal Recursive Estimation with Uncertain Observation. *IEEE Transactions on Information Theory*. Vol. IT-15, No. 4, pp. 457-462.
- [32] RAJASEKARAN, P.K., SATYANARAYANA, N. Y SRINATH, M.D. (1971). Optimum Linear Estimation of Stochastic Signals in the Presence of Multiplicative Noise. *IEEE Transactions on Aerospace and Electronic Systems*. Vol. AES-7, No. 3, pp. 462-468.
- [33] SORENSON, H.W. (1970). Least Squares Estimation: from Gauss to Kalman. *IEEE Spectrum*. Vol. 7, pp. 63-68.
- [34] TANIZAKI, H. (1996). *Nonlinear Filters. Estimation and Applications*. Springer-Verlag. New York.

- [35] TUGNAIT, J.K. (1981). Asymptotic Stability of the MMSE Linear Filter for Systems with Uncertain Observations. *IEEE Transactions on Information Theory*. Vol. IT-27, No. 2, pp. 247-250.
- [36] TUGNAIT, J.K. (1981). Stability of Optimum Linear Estimators of Stochastic Signals in White Multiplicative Noise. *IEEE Transactions on Automatic Control*. Vol. AC-26, No. 3, pp. 757-761.
- [37] VALDERRAMA, M.J. Y RUIZ, J.C. (1996). *Filtrado de Kalman. Aplicaciones en Economía e Ingeniería*. EUB, Barcelona.
- [38] WANG, X. (1984). Recursive Algorithms for Linear LMSE Estimators Under Uncertain Observations. *IEEE Transactions on Automatic Control*. Vol. AC-29, No. 9, pp. 853-854.