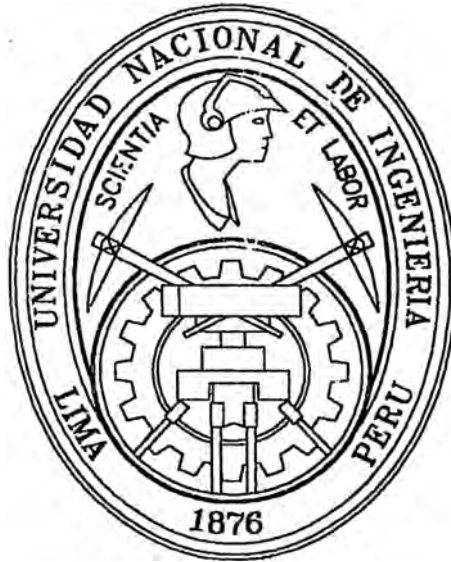


UNIVERSIDAD NACIONAL DE INGENIERIA

**FACULTAD DE INGENIERIA ECONOMICA
Y CIENCIAS SOCIALES**



**Aproximación Dinámica Bayesiana para
el Consumo Privado en el Perú**

**PARA OPTAR EL TÍTULO PROFESIONAL DE
LICENCIADO EN ESTADÍSTICA**

POR LA MODALIDAD DE TESIS

ELABORADO POR

Hernán Belisario Garrafa Aragón

**LIMA - PERÚ
2002**

Aproximación Dinámica Bayesiana para el Consumo Privado en el Perú

por

Hernán Garrafa Aragón

Disertación sometida al cuerpo docente de la Escuela Profesional de Ingeniería Estadística de la Universidad Nacional de Ingeniería como parte de los requisitos necesarios para obtener el título profesional de Licenciado en Estadística

Area de concentración: Estadística.

Universidad Nacional de Ingeniería

Escuela Profesional de Ingeniería Estadística

Junio del 2002

*A mis padres: Margarita y Braulio a ellos mi eterna gratitud por el esfuerzo,
dedicación y aliento pilares para mi desarrollo
profesional y personal*

Agradecimientos

- **A mi asesor, Dr. Luis Navarro por la dedicación de tiempo que me brindó para poder darme el soporte teorico-práctico de esta tesis y la difícil tarea que tuvo que realizar en multiples formas de corrección tratando de darle mayor fluidez a los temas aqui tratados.**
- **A mis hermanos, Ines y José por las motivaciones que siempre me dieron para que culminara esta tesis.**
- **A mi enamorada, Jessica Aricoché por su colaboración que me brindó en la corrección de la gramática.**
- **Al Docente, Msc. Carlos Abanto por ser el principal gestor para iniciar esta tesis, luego coordinar y desarrollar el tema aqui expuesto.**
- **Al Docente, Anibal Serrano, por su ayuda en lo que al área de computo se refiere y las sugerencias dadas en la presentación de la tesis.**
- **Y así también expreso mi aprecio a todos los colegas docentes por su orientación en el desarrollo del presente trabajo.**

RESUMEN

El análisis de Modelos Dinámicos Lineales involucra una secuencia de distribuciones a posteriori y predictivas, correspondientes a las diferentes etapas de aplicación del modelo. La variable macro-económica consumo privado en el Perú es explicada por un Modelo Dinámico Bayesiano Lineal por componentes, donde la tendencia es modelada por un polinomio de segundo orden, y la estacionalidad trimestral por un modelo de formato libre. Este proceso de inferencia es completada por una secuencia de pronósticos que determina la validez de la utilización de este modelo propiamente dicho.

ABSTRACT

The analysis of Linear Dynamic Models involves a sequence of posterior and forecast distribution corresponding to the subsequent stages stage of model application. the macroeconomic variable domestic consumer in the Perú is explain for a Dynamic Linear Bayesian Model for component, where the trends is modeling for a second-order polynomial, and the quarter seasonal for a format free model. This process of inference is complete by forecasting sequence that determine the vality in the use of this model say properly.

Indice General

Agradecimientos	ii
Resumen	iii
Lista de Figuras	v
Lista de Tabelas	vi
1 Introducción	1
2 Inferencia Estadística Bayesiana	7
2.1 Naturaleza de la Inferencia Bayesiana	8
2.1.1 Teorema de Bayes	8
3 El Modelo Lineal Dinámico Univariado	14
3.1 Introducción	14
3.2 Definiciones y Notación .	14
3.3 Ecuaciones de Evolución: El MLD univariado	16
3.4 Distribuciones Predictivas	19
3.5 Varianza Observacional . .	26
3.6 Recurrencias para Filtración	36
3.7 Análisis de referencia para un MLD	43

3.7.1	Caso especial de $W_t = 0$	52
3.7.2	Filtración	54
3.8	Factores de descuento y especificación del modelo componente	56
3.8.1	Descuento por componente	58
3.8.2	Práctica de descuento estratégico	61
4	Consumo Privado en el Perú	64
4.1	Introducción	64
4.2	Importancia del Consumo Privado en el Perú	65
4.3	Análisis de datos	65
4.4	Modelo Lineal Dinámico para el Consumo Privado	68
4.4.1	Análisis de los resultados	70
4.4.2	Modelos Clásicos	81
5	Conclusiones y Sugerencias	83
A	Distribuciones Multivariadas	85
A.1	La Distribución Normal Multivariada	85
A.2	Distribución T-Student	86
B	Conjunto de Datos	87
	Bibliografía	89

Indice de Figuras

2.1	Densidad de la distribución Normal estandarizada	13
4.1	Consumo Privado a soles constantes de 1979	67
4.2	Evolución del nivel de la serie del consumo privado	74
4.3	Evolución del crecimiento para la serie del consumo privado	75
4.4	Evolución de la Estacionalidad	76
4.5	Evolución de los pronósticos un paso al frente	77
4.6	Filtración de la función de respuesta media, $F'_{T-k}\theta_{T-k}$	78
4.7	Actualización de la función de respuesta media, $F'_t\theta_t$	79
4.8	Ploteo probabilidad Normal para los Residuales	80
4.9	Regresión Lineal	81
4.10	Regresión polinomial de 2do order	82
4.11	Regresión polinomial de 3er order	82

Índice de Tablas

4.1	Pronósticos para el Consumo Privado (MLD)	72
B.1	Consumo Privado 1990.1 - 1999.1	88

Capítulo 1

Introducción

El caso real a tratar es desarrollar una aproximación Dinámica Bayesiana a la variable macro-económica consumo privado en el Perú. En nuestro país el consumo privado es el 70% del PBI. Lo que nos indica una importante dependencia del mercado local, de ahí la importancia del consumo en la economía. Nuestro consumo privado depende del consumo privado de las potencias mundiales, viendo las expectativas en EE.UU. definitivamente nos va a afectar aunque menos que a otros países donde las exportaciones son más importantes. Por ejemplo la caída de precios en las materias primas nos va a afectar y eventualmente en todo lo que es productos finales (textiles por ejemplo). El mayor impacto sobre el consumo privado se da por la propia situación local. Actualmente no se da medidas retroactivadoras en el empleo que pueda mejorar el consumo privado. Muestra de ello son las recientes iniciativas en materia laboral. La reactivación del empleo pasa por un incremento en la inversión que, en un mundo globalizado, es algo por lo que se compite. En ese sentido, el énfasis del Gobierno debería estar puesto en volver al Perú en un lugar muy atractivo para la inversión.

El modelamiento estadístico de series de tiempo, inferencia y pronóstico, está basado en procesos Dinámicos. En este contexto el término Dinámico se refiere a los

cambios que se producen debido al paso del tiempo. Predicción Bayesiana y modelamiento Dinámico tiene una historia que puede remontarse hacia fines de los años 50, con aplicaciones en monitoreo y pronóstico en el área comercial. A la par, desarrollos técnicamente similares, surgieron en las áreas de ingeniería de sistemas e ingeniería de control, con un enfoque centrado en la estimación adaptativa y teoría de filtros para control automático.

En las últimas tres décadas, el desarrollo y aplicaciones en el ámbito de los Modelos Dinámicos ha crecido de manera considerable y de forma vertiginosa. Así, para tener una idea de las diferentes áreas de aplicación de los Modelos Dinámicos, éstas abarcan campos tan tradicionales como el pronóstico de series comerciales e industriales, incluyendo pronóstico de ventas y de demanda (por ejemplo Green e Harrison [9]; Johnston e Harrison [11]; Queen y Smith [14]; West y Harrison [21]), demanda de consumo de energía, Investigación de mercados y publicidad (Migon y Harrison [13]), gerenciamiento de stocks/inventarios (Harrison [10]), entre otros. Esto es complementado con aplicaciones en econometría y macroeconomía con énfasis en el modelamiento en finanzas, cambios estructurales, agregación (por ejemplo Quintana y West [15], [16]; Quintana [17]), en bioestadística y economía (Gamerman y West [6]; Gamerman [7]; West [18]), análisis espectral (Kleiner et al [12]; West [19]).

La clase de Modelos Dinámicos más conocida y usada es la de los Modelos Lineales Dinámicos (**MLD**), en los cuales el supuesto de Normalidad en los datos está sobre entendida, y es precisamente en esta clase de modelos que el trabajo de tesis tratará, conjuntamente con un estudio de caso real.

Un punto de vista de modelamiento científico es que un modelo es “una descripción simplificada de un sistema, que permite hacer cálculos y pronósticos”. Más claramente, un modelo es un esquema de descripción y explicación que organiza información y experiencias proporcionando un medio de análisis y pronóstico. Una de las razones para

modelar, es preveer un proceso de análisis eficiente, el cual mejorará y permitirá tomar decisiones oportunas. De igual manera, toda la operación de alguna organización se puede ver como la comprensión de alguna secuencia de decisiones basadas sobre un flujo continuo de información en el tiempo. Consecuentemente, hay una acumulación de conocimientos los cuales en principio, conducirían a mejorar la comprensión y la toma de decisiones.

En tal sentido, la estructuración de los problemas de pronóstico y modelamiento Dinámico Bayesiano se fundamenta en:

- Definición de un modelo secuencial;
- Estructuración mediante el uso de modelos paramétricos;
- Representación probabilística de la información acerca de los parámetros;
- Las inferencias y pronósticos son obtenidos a partir de las distribuciones posteriori y predictiva, respectivamente.

En el contexto de series de tiempo el uso de modelos secuenciales y estructurados son naturales. A cada momento, toda la información relevante para prever el futuro es recepcionada y puede ser usada en la revisión y crítica del modelo. Supongamos que en tiempo $t = 0$, D_0 representa toda la información relevante y disponible, la cual será usada por el modelador para hacer los pronósticos iniciales. Similarmente, se supone que para cualquier $t > 0$, la información disponible y relevante es denotada por D_t . Cualquier afirmación sobre el futuro estará condicionada a esta información D_t . Una vez que Y_t fue observado en el tiempo t , D_t contiene la información anterior D_{t-1} e Y_t . Si esta es toda la información relevante, se tendrá que $D_t = \{Y_t, D_{t-1}\}$.

El enfoque secuencial se centra en el uso de modelos probabilísticos para el desarrollo de las series en el futuro, condicional a la información existente. Esto quiere decir, que

si nuestro interés principal es Y_t , las inferencias serán hechas basadas en la información contenida en \mathbf{D}_{t-1} , y toda la información relevante acerca de θ_t , a través de un modelo paramétrico definido por:

$$p(Y_t | \theta_t, \mathbf{D}_{t-1}), \quad (1.1)$$

donde, θ_t es el vector de parámetros en el tiempo t . Entonces, condicionado a estos argumentos esta notación reconoce explícitamente la dependencia de Y_t al modelo paramétrico y la información pasada.

Para ilustrar las ideas sobre la importancia de los Modelos Lineales Dinámicos, consideraremos dos ejemplos simples.

Ejemplo 1.1.-Un prototipo de modelo estadístico en el que las observaciones son independientes y idénticamente distribuidas Normalmente, $(Y_t | \mu) \sim N[\mu, V]$, ($t = 1, 2, \dots$), y en el que la media varía en el tiempo (y algunas veces la varianza) para este modelo son consideraciones naturales y características típicamente inevitables cuando las observaciones son hechas en un proceso o sistema determinado. La variación en el tiempo puede ser lenta y gradual, reflejando continuos cambios en las condiciones del medio ambiente, y más abruptamente debido a los cambios del mercado. Por ejemplo, se puede asumir un modelo Normal como una representación conveniente para el tratamiento de la variación aleatoria en una demanda del consumo “estable” para un determinado producto, en donde el nivel de la demanda raramente permanecerá absolutamente constante sobre el tiempo. Substituyendo μ por t , se tiene que $(Y_t | \mu_t) \sim N[\mu_t, V]$, o también:

$$Y_t = \mu_t + \nu_t,$$

donde μ_t representa el nivel de la serie en el tiempo t , y $\nu_t \sim N[0, V]$, error observacional o ruido. Una de las formas más simples de modelar μ_t , será usar el “camino aleatorio”. En tal caso, la evolución del nivel μ_t está representado por $\mu_t = \mu_{t-1} + \omega_t$,

donde, ω_t representa el error o ruido asociado al nivel μ_t . Por ejemplo MLD estandar asume $\omega_t \sim N[0, W]$ para una alguna varianza W . Se podrá suponer que los niveles subyacentes son constantes localmente, pero cambios esperados pueden ser significativa sobre largos periodos de tiempo. Las suposiciones media-cero e independencia para las series de ω_t son consistentes con lo observado en periodos largos de tiempo.

Con referencia a la discusión anterior, si las varianzas V y W son asumidas conocidas, entonces el modelo del parámetro θ_t representa solo el nivel μ_t . De otro lado θ_t puede tener varianzas inciertas para uno o el otro, o ambos ν_t y ω_t . Cuando las varianzas son conocidas, sus valores están incluidas en D_0 . Entonces, con $\theta_t = \mu_t \forall t$, $p(Y_t | \theta_t, D_{t-1}) = p(Y_t | \mu_t, D_0)$ es la densidad Normal de $(Y_t | \mu_t, D_0) \sim N[\mu_t, V]$. La información pasada hasta el tiempo $t - 1$ conduce a la distribución $\theta_t = \mu_t$ dado D_{t-1} .

Ejemplo 1.2.-Regresión está orientado con la construcción de relaciones entre dos series de tiempo representadas por las observaciones X_t , ($t = 1, 2, \dots$), y observaciones contemporáneas Y_t . En los modelos de regresión a menudo se piensa que el cambio en la media μ_t de Y_t , puede ser explicado por X_t , y posibles valores pasados X_s , para $s < t$. Comunmente nos referimos a Y_t como la respuesta o la serie de variable dependiente, y a X_t como el regresor o serie de variable independiente. Entonces, μ_t es la respuesta media relativa a la variable regresora. Luego, la función de respuesta media estaría dado por $\mu_t = \mu_t(X_t, X_{t-1}, \dots)$ definiendo así la regresión. Por ejemplo, un Modelo Lineal Simple con un efecto de X_t en la media es $\mu_t = \alpha + \beta X_t$, donde los parámetros α y β toman valores adecuados. Modelos de esta clase encuentran un amplio uso en predicción, interpolación y estimación. Para algunos propósitos, el modelo lineal simple puede ser satisfactorio localmente, pero son inadecuados para describir la relación global como el tiempo y como cambios de X_t . Una manera de flexibilizar el modelo puede ser, permitir la simple posibilidad de que los coeficientes varíen en el

tiempo,

$$\mu_t = \alpha_t + \beta_t \mathbf{X}_t.$$

De este modo aunque la forma del modelo de regresión es lineal en \mathbf{X}_t para todo t , el modelo anterior puede tener diferentes parámetros definidos en diferentes tiempos. La distinción entre una forma de modelo local apropiado y el Modelo Dinámico anterior es crítica. Frecuentemente los valores de la variable independiente \mathbf{X}_t cambia de manera suave en el tiempo y un modelo apropiado fue descrita anteriormente por $\mu_t = \alpha_t + \beta_t \mathbf{X}_t$, en donde los parámetros varían ligeramente desde un punto del tiempo hasta el siguiente. Estos pueden ser modelados usando caminos aleatorios definiendo parámetros, como:

$$\alpha_t = \alpha_{t-1} + \delta\alpha_t,$$

$$\beta_t = \beta_{t-1} + \delta\beta_t,$$

donde $\delta\alpha_t$ y $\delta\beta_t$ son los errores con media-cero. Estas ecuaciones de evolución expresan el concepto de constancia local en los parámetros, sujeto a la variación controlada y modelada mediante la distribución de la evolución de los términos de error $\delta\alpha_t$ y $\delta\beta_t$. Pequeños grados de variación aquí implica una función de regresión lineal estable sobre el tiempo, valores grandes modela mayor volatilidad y sugiere precaución a la hora de explorar o a la hora de hacer pronósticos para horizontes de tiempo demasiado lejos.

El presente trabajo está organizado de la siguiente forma, en el Capítulo 2 se hace una revisión de los principales aspectos de la Estadística Bayesiana. en el Capítulo 3 se presenta la Teoría de los Modelos Dinámicos Bayesianos Lineales sobre la cual versará la presente tesis. En el Capítulo 4 detallaremos una aplicación de los Modelos Dinámicos Bayesianos Lineales para la variable macroeconómica **Consumo Privado**. En el Capítulo 5, finaliza con conclusiones y sugerencias sobre este trabajo de tesis.

Capítulo 2

Inferencia Estadística Bayesiana

La Inferencia Estadística Bayesiana es una propuesta de Inferencia que se basa en el Teorema de Bayes [1]. Para mayor información acerca de la Estadística Bayesiana ver Bernardo y Smith [4], Box-Tiao [5], etc. El objetivo fundamental de todo análisis estadístico es hacer inferencia estadística, datos son coleccionados y necesitamos hacer inferencias sobre uno o más factores desconocidos relacionados al sistema generador de estos datos. En las próximas secciones anunciaremos brevemente los principales puntos de esta teoría.

Aunque la Teoría frecuentista produce resultados simples en casos en donde suposiciones especiales, tales como Normalidad e independencia son hechas, en otros casos las soluciones no son satisfactorias, particularmente cuando ninguna **estadística suficiente** existe. Aunque sea verdad que estas suposiciones especiales involucren un número de situaciones de interés científico, es ideal pretender que un conjunto de problemas estadísticos cuya solución fué colocada como necesaria por el investigador coincida con el conjunto de problemas sujetos a un tratamiento padrón. La metodología Bayesiana dá mayor énfasis al interés científico que conveniencia matemática.

Una distribución a priori, que representa toda la información sobre los parámetros

desconocidos antes que los datos sean disponibles, asume un papel esencial en la aproximación Bayesiana. Tal distribución puede ser usada para representar el conocimiento a priori o ignorancia “relativa”. En problemas de inferencia estadística deberíamos, llevar a cabo un análisis como si un estado de ignorancia relativa existiese a priori, siempre que fuera posible. El temor por parte de algunos estadísticos acerca del uso de distribuciones a priori está asociado con el temor que la distribución a priori domine y distorciona la información dada por los datos a través de la verosimilitud.

Una ventaja del paradigma Bayesiano es que él puede ser usado para explorar las consecuencias de cualquier tipo de modelo probabilístico, sin restricción para aquellos que tiene formas matemáticas especiales o convenientes

2.1 Naturaleza de la Inferencia Bayesiana

2.1.1 Teorema de Bayes

Suponga que $\mathbf{y} = (y_1, \dots, y_n)$ es un vector de n observaciones cuya función de densidad de probabilidades f.d.p $p(\mathbf{y}|\theta)$ depende del vector de parámetros $\theta' = (\theta_1, \dots, \theta_k)$. A priori, θ tiene una f.p.d $p(\theta)$. Entonces, la f.d.p de θ e \mathbf{y} es:

$$p(\mathbf{y}, \theta) = p(\mathbf{y}|\theta)p(\theta) = p(\theta|\mathbf{y})p(\mathbf{y}), \quad (2.1)$$

dado los datos observados \mathbf{y} , la f.d.p de θ es dada por:

$$p(\theta|\mathbf{y}) = \frac{p(\mathbf{y}|\theta)p(\theta)}{p(\mathbf{y})} \quad (2.2)$$

Note que la f.d.p marginal de los datos es:

$$p(\mathbf{y}) = E [p(\mathbf{y}|\theta)] = c^{-1} = \begin{cases} \int p(\mathbf{y}|\theta)p(\theta)d\theta; & \theta \text{ continuo} \\ \sum p(\mathbf{y}|\theta)p(\theta); & \theta \text{ discreto} \end{cases} \quad (2.3)$$

donde la utilización de los operadores suma o integral dependerá de las características del dominio de θ , y $E[f(\theta)]$ es el operador esperanza matemática de $f(\theta)$ con respecto a $p(\theta)$. Así (2.2) puede ser escrita como:

$$p(\theta|\mathbf{y}) = c p(\mathbf{y}|\theta)p(\theta). \quad (2.4)$$

El enunciado (2.2), o equivalentemente (2.4), es comumente conocido como el *Teorema de Bayes*. En esta expresión, $p(\theta)$ representa toda la información acerca de θ sin conocimiento de los datos, y es llamada la distribución *a priori* de θ . Una vez conocidos los datos $p(\theta|\mathbf{y})$, representa la distribución actualizada de θ y será llamada distribución *a posteriori* de θ dado \mathbf{y} . La cantidad c es una constante de normalización necesaria para garantizar que la integral (o suma) de $p(\theta|\mathbf{y})$ sea igual a 1.

Teorema de Bayes y la Función de Verosimilitud

Dada la información \mathbf{y} , $p(\mathbf{y}|\theta)$ en (2.4) puede ser vista como una función únicamente de θ , y no de \mathbf{y} . Vista de esta forma, ella es llamada de *función de verosimilitud* de θ dada la información \mathbf{y} , y puede ser escrita como $l(\theta|\mathbf{y})$. Luego es posible escribir el Teorema de Bayes como:

$$\begin{aligned} p(\theta|\mathbf{y}) &= c l(\theta|\mathbf{y})p(\theta) \\ p(\theta|\mathbf{y}) &\propto l(\theta|\mathbf{y})p(\theta) \end{aligned} \quad (2.5)$$

O sea, el Teorema de Bayes dice que la f.d.p de θ dado \mathbf{y} , es proporcional al producto de la f.d.p. a priori de θ y de la verosimilitud de θ dado \mathbf{y} . Es decir,

distribución a posteriori \propto verosimilitud \times distribución a priori.

La función de verosimilitud $l(\theta|\mathbf{y})$ destaca en el Teorema de Bayes. A través de ella los datos \mathbf{y} modifican el conocimiento a priori de θ . Así, ella representa toda la

información actualizada cerca de θ .

Naturaleza secuencial del Teorema de Bayes

La ecuación (2.5) nos dá una formulación matemática de como la información anterior puede ser combinada con nuevas informaciones. El Teorema de Bayes permite una actualización continua de información acerca de θ a medida que nuevas informaciones aparezcan. Para ilustrar la idea anterior, suponga que se tiene inicialmente \mathbf{y}_1 , por el Teorema de Bayes se tiene que:

$$p(\theta|\mathbf{y}_1) \propto p(\theta)l(\theta|\mathbf{y}_1), \quad (2.6)$$

a continuación, suponga ahora que observamos \mathbf{y}_2 obtenido independientemente de \mathbf{y}_1 entonces:

$$\begin{aligned} p(\theta|\mathbf{y}_1, \mathbf{y}_2) &\propto p(\theta)l(\theta|\mathbf{y}_1)l(\theta|\mathbf{y}_2) \\ &\propto p(\theta|\mathbf{y}_1)l(\theta|\mathbf{y}_2) \end{aligned} \quad (2.7)$$

la expresión (2.7) tiene la misma forma que (2.6), en donde $p(\theta|\mathbf{y}_1)$ asume el papel de la distribución a priori en la obtención de $p(\theta|\mathbf{y}_1, \mathbf{y}_2)$ a través del Teorema de Bayes. Obviamente este proceso puede ser repetido muchas veces. Para extender estas ideas, imagine que tenemos n observaciones independientes, la distribución a posteriori puede ser re-calculada después de cada nueva observación, así en el m -ésimo estado la verosimilitud asociada a la m -ésima observación es combinada con la distribución posteriori de θ dado $m-1$ observaciones, para obtener la nueva distribución a posteriori:

$$p(\theta|\mathbf{y}_1, \dots, \mathbf{y}_m) \propto p(\theta|\mathbf{y}_1, \dots, \mathbf{y}_{m-1})l(\theta|\mathbf{y}_m) \quad (2.8)$$

donde:

$$p(\theta|\mathbf{y}_1) \propto p(\theta)l(\theta|\mathbf{y}_1)$$

Estimación Puntual

Un problema de estimación puntual Bayesiana puede ser colocada como un problema de decisión, ver detalles en Berger [3].

Definición 2.1. La pérdida esperada a posteriori de una acción \mathbf{a} , cuando la f.d.p posteriori es $p(\boldsymbol{\theta} | x)$, es dado por:

$$\rho(p(\boldsymbol{\theta} | x), \mathbf{a}) = \int_{\Theta} L(\boldsymbol{\theta}, \mathbf{a})p(\boldsymbol{\theta} | x)dx$$

con $L(\boldsymbol{\theta}, \mathbf{a})$, función de pérdida al asumir la acción \mathbf{a} cuando la naturaleza selecciona $\boldsymbol{\theta}$. ($\boldsymbol{\theta} \in \Theta$: espacio disponible de estados de la naturaleza). La acción de Bayes \mathbf{a} es cualquier $\mathbf{a} \in \mathbf{A}$ (espacio de acciones disponibles), tal que minimize $\rho(p(\boldsymbol{\theta} | x), \mathbf{a})$.

En estimación Bayesiana de un parámetro de valor real $\boldsymbol{\theta}$, la función de pérdida usada es $L(\boldsymbol{\theta}, \mathbf{a}) = (\boldsymbol{\theta} - \mathbf{a})^2$. La pérdida esperada a posteriori es entonces:

$$\int_{\Theta} (\boldsymbol{\theta} - \mathbf{a})^2 p(\boldsymbol{\theta} | x) d\boldsymbol{\theta},$$

el valor de \mathbf{a} el cual minimiza la anterior expresión es hallado, expandiendo la expresión cuadrática y diferenciando con respecto a \mathbf{a} , e igualando a cero. El resultado es:

$$\begin{aligned} 0 &= \frac{d}{dx} \left[\int_{\Theta} \boldsymbol{\theta}^2 p(\boldsymbol{\theta} | x) d\boldsymbol{\theta} - 2\mathbf{a} \int_{\Theta} \boldsymbol{\theta} p(\boldsymbol{\theta} | x) d\boldsymbol{\theta} + \mathbf{a}^2 \int_{\Theta} p(\boldsymbol{\theta} | x) d\boldsymbol{\theta} \right], \\ 0 &= -2E^{p(\boldsymbol{\theta}|x)}(\boldsymbol{\theta}) + 2\mathbf{a}. \end{aligned}$$

Lema.- Sea $L(\boldsymbol{\theta}, \mathbf{a}) = (\boldsymbol{\theta} - \mathbf{a})^2$, la acción de Bayes \mathbf{a}^p es dado por

$$\mathbf{a}^p = E^{p(\boldsymbol{\theta}|x)}(\boldsymbol{\theta}),$$

el cual es la media a posteriori de $\boldsymbol{\theta}$ dado x .

A veces $\mathbf{a}^p = \delta(x)$, dado que esta acción depende del valor observado x .

Estimación por intervalos

Definición 2.2. C es un intervalo de confianza Bayesiano o una región de credibilidad de $100(1 - \alpha)\%$ para, θ si $pr(\theta \in C) \geq 1 - \alpha$. En este caso, $1 - \alpha$ es llamado nivel de confianza o credibilidad.

La probabilidad usada en el enunciado de la definición 2.2, es calculado usando la distribución actualizada de θ .

Ejemplo Sea $\underline{X} = (X_1, \dots, X_n)$ una muestra con $X_i \sim N(\theta, \sigma^2)$ con $\sigma^2 < \infty$ conocido. Entonces la a priori no informativa para θ será $p(\theta) \propto cte$, y la verosimilitud está dado por:

$$l(\theta; \underline{x}) \propto \left\{ -\frac{n}{2\sigma^2}(\theta - \bar{x}) \right\},$$

se sabe que $p(\theta | \underline{x}) \propto l(\theta; \underline{x})p(\theta) \propto l(\theta; \underline{x})$ por lo tanto $(\theta | \underline{x}) \sim N(\bar{x}, \frac{\sigma^2}{n})$ o equivalentemente $(\frac{\sqrt{n}(\theta - \bar{x})}{\sigma} | \underline{x}) \sim N(0, 1)$. A partir de aqui se podrá construir varios intervalos de confianza $100(1 - \alpha)\%$ para θ . Definiendo $\phi(x) = P(X \leq x)$ si $X \sim N(0, 1)$, entonces z_c es tal que $\phi(z_c) = 1 - c$, $0 < c < 1$ y los intervalos pueden ser construidos a partir de:

usando la simetría de la distribución Normal, se tiene

$$\phi(-z_\beta) = P(X \leq -z_\beta) = P(X \geq z_\beta) = 1 - P(X < z_\beta) = \beta,$$

y la probabilidad del intervalo que se quiere construir está dado por

$$\phi(z_\gamma) - \phi(-z_\beta) = 1 - (\gamma + \beta),$$

donde $\gamma + \beta = \alpha$. Luego, el intervalo de confianza Bayesiano de $100(1 - \alpha)\%$ será de la forma

$$-z_\beta \leq \frac{\sqrt{n}(\theta - \bar{x})}{\sigma} \leq z_\gamma,$$

el cual se puede expresar

$$-\frac{\sigma}{\sqrt{n}}z_\beta + \bar{x} \leq \theta \leq z_\gamma \frac{\sigma}{\sqrt{n}} + \bar{x},$$

entonces

$$C = \left[\bar{x} - \frac{\sigma}{\sqrt{n}}z_\beta, \bar{x} + z_\gamma \frac{\sigma}{\sqrt{n}} \right],$$

es un intervalo de confianza Bayesiano $100(1 - \alpha)\%$ para θ . Observe que la longitud de C es $[(z_\gamma + z_\beta) \cdot \sigma \mid \sqrt{n}]$. Permanece la cuestión de como minimizar esta longitud.

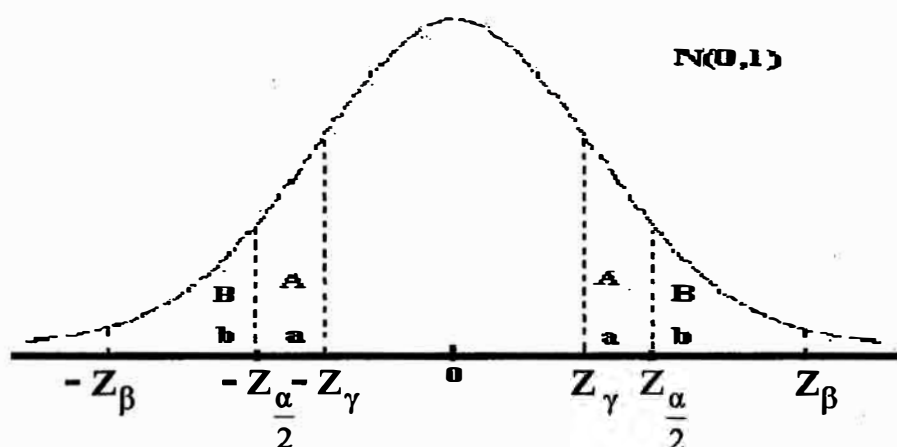


Figura 2.1: Densidad de la distribución Normal estandarizada

Considere por un momento que $z_\gamma < z_{\frac{\alpha}{2}} < z_\beta$ y defina $a = z_{\frac{\alpha}{2}} - z_\gamma > 0$, $b = z_\beta - z_{\frac{\alpha}{2}} > 0$, A y B como las áreas comprendidas entre z_γ y $z_{\frac{\alpha}{2}}$ y entre $z_{\frac{\alpha}{2}}$ y z_β respectivamente. La longitud del intervalo C es $[2z_{\frac{\alpha}{2}} + b - a]$ siempre que $A = B$. Se puede ver en la figura que la densidad sobre el primer intervalo es estrictamente mayor que sobre el segundo intervalo, se tiene que $b > a$. Luego, un intervalo de confianza obtenido con extremos simétricos $-z_{\frac{\alpha}{2}}$ y $z_{\frac{\alpha}{2}}$ es el de menor longitud.

Definición 2.3.- Un intervalo de confianza Bayesiano $100(1 - \alpha)\%$ de máxima densidad a posteriori (MDP) para θ es el intervalo de confianza Bayesiano $100(1 - \alpha)\%$ de forma que $C = \{\theta \in \Theta : p(\theta \mid x) \geq k(\alpha)\}$ donde $k(\alpha)$ es la mayor constante tal que $P(\theta \in C \mid x) \geq 1 - \alpha$.

Capítulo 3

El Modelo Lineal Dinámico

Univariado

3.1 Introducción

En este capítulo se presentan la clase general de los Modelos Normales Lineales Dinámicos y sus principales características, referidos de aquí en adelante como los Modelos Lineales Dinámicos (MLD) donde la normalidad está sobre entendida. Esta clase de modelos es discutida y analizada basados en la propuesta de West y Harrison, [21].

3.2 Definiciones y Notación

Se definirá el MLD general para un vector de observaciones \mathbf{Y}_t . Sin embargo, toda la discusión que se presentará a continuación y en lo que sigue del presente trabajo, estará referido estrictamente al caso particular en que \mathbf{Y}_t es escalar para todo t . Aquí, \mathbf{Y}_t es un vector ($r \times 1$) de observaciones de una serie de tiempo $\mathbf{Y}_1, \mathbf{Y}_2, \dots$.

Definición 3.1. El Modelo Lineal Normal Dinámico general está caracterizado por la cuádrupla:

$$\{\mathbf{F}, \mathbf{G}, \mathbf{V}, \mathbf{W}\}_t = \{\mathbf{F}_t, \mathbf{G}_t, \mathbf{V}_t, \mathbf{W}_t\}$$

para cada tiempo t , donde:

- (a) \mathbf{F}_t es una matriz conocida ($n \times r$);
- (b) \mathbf{G}_t es una matriz conocida ($n \times n$);
- (c) \mathbf{V}_t es una matriz de varianza conocida ($r \times r$);
- (d) \mathbf{W}_t es una matriz de varianza conocida ($n \times n$);

esta cuádrupla define el modelo relacionando \mathbf{Y}_t al vector de parámetros $\boldsymbol{\theta}_t$, de dimensión ($n \times 1$), en el tiempo t y la secuencia $\boldsymbol{\theta}_t$ a través del tiempo, vía la especificación secuencial de las distribuciones:

$$(\mathbf{Y}_t | \boldsymbol{\theta}_t) \sim N(\mathbf{F}'_t \boldsymbol{\theta}_t, \mathbf{V}_t), \quad (3.1)$$

$$(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1}) \sim N(\mathbf{G}_t \boldsymbol{\theta}_{t-1}, \mathbf{W}_t), \quad (3.2)$$

estas ecuaciones están implícitamente condicionadas a \mathbf{D}_{t-1} , la información disponible al tiempo $t - 1$. Esto incluye, en particular los valores de las varianzas \mathbf{V}_t y \mathbf{W}_t y las observaciones anteriores $\mathbf{Y}_{t-1}, \mathbf{Y}_{t-2}, \dots$ así como la información inicial \mathbf{D}_0 . Una representación alternativa para las ecuaciones anteriores está dado por:

$$\mathbf{Y}_t = \mathbf{F}'_t \boldsymbol{\theta}_t + \boldsymbol{\nu}_t \quad \boldsymbol{\nu}_t \sim N(\mathbf{0}, \mathbf{V}_t) \quad (3.3)$$

$$\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t \quad \boldsymbol{\omega}_t \sim N(\mathbf{0}, \mathbf{W}_t) \quad (3.4)$$

donde la secuencia de errores $\boldsymbol{\nu}_t$ y $\boldsymbol{\omega}_t$ son independientes y mutuamente independientes. La ecuación (3.3) es denominada **ecuación observacional** para el modelo, la misma

que define la distribución muestral de Y_t condicional al vector de parámetros θ_t . Se asume que dado este vector θ_t , Y_t es condicionalmente independiente de los valores pasados de la serie. F_t juega el rol de la matriz de regresión con valores conocidos y θ_t es el vector de parámetros de regresión considerado dinámico y al que se hará referencia como el **vector de estados** o **vector del sistema**. La función de respuesta media en t es $\mu_t = F_t' \theta_t$, simplemente el valor esperado de Y_t en (3.3), la cual define el nivel de la serie en el instante t . ν_t es el error observacional en el tiempo t . La ecuación (3.4) es la **ecuación del sistema** o **ecuación de evolución**, la misma que define la evolución en el tiempo del vector de estado θ_t . G_t es la **matriz de evolución** y ω_t es el **error de evolución** con varianza conocida W_t .

Definición 3.2. Para cada t , el **MLD univariado** general está definido por:

$$Y_t = F_t' \theta_t + \nu_t \quad \nu_t \sim N(0, V_t) \quad (3.5)$$

$$\theta_t = G_t \theta_{t-1} + \omega_t \quad \omega_t \sim N(0, W_t) \quad (3.6)$$

La información inicial supone que $(\theta_0 | D_0) \sim N(m_o, C_o)$, donde m_o y C_o son valores conocidos. La sucesión de errores ν_t y ω_t son asumidas independientes y mutuamente independientes entre sí, e independientes de $(\theta_0 | D_0)$.

3.3 Ecuaciones de Evolución: El MLD univariado

Considérese el **MLD univariado** cerrado a información externa en el tiempo $t \geq 1$. De este modo, dada la información inicial D_0 en $t = 0$, la información disponible en cualquier tiempo t es simplemente:

$$D_t = \{Y_t, D_{t-1}\}$$

donde Y_t es el valor observado de la serie en el tiempo t ; se supone que D_0 contiene la información necesaria para definir completamente la cuádrupla $\{F, G, V, W\}_t$

Teorema 3.1. En el MLD univariado de la definición 3.2 las distribuciones para cualquier $t \geq 1$ están dadas por:

(a) Posteriori en $t - 1$: $(\boldsymbol{\theta}_{t-1} | \mathbf{D}_{t-1}) \sim N[m_{t-1}, C_{t-1}]$

(b) Priori en t : $(\boldsymbol{\theta}_t | \mathbf{D}_{t-1}) \sim N[\mathbf{a}_t, R_t]$

(c) Pronóstico un paso al frente : $(\mathbf{Y}_t | \mathbf{D}_{t-1}) \sim N[f_t, Q_t]$

(d) Posteriori en t : $(\boldsymbol{\theta}_t | \mathbf{D}_t) \sim N[m_t, C_t]$

donde:

$$m_t = \mathbf{a}_t + A_t(\mathbf{Y}_t - f_t), \quad A_t = R_t F_t' Q_t^{-1}, \quad C_t = R_t - A_t A_t' Q_t, \quad \mathbf{a}_t = G_t m_{t-1}, \quad R_t = G_t C_{t-1} G_t' + W_t, \quad f_t = F_t' \mathbf{a}_t \quad \text{y} \quad Q_t = F_t' R_t F_t + V_t$$

Demostración.

La prueba se hará utilizando el método de inducción. Para ello se supone que (a) es cierto en $t - 1$, luego $E(\boldsymbol{\theta}_{t-1} | \mathbf{D}_{t-1}) = m_{t-1}$ y $V(\boldsymbol{\theta}_{t-1} | \mathbf{D}_{t-1}) = C_{t-1}$. La prueba de (b) se obtiene del hecho que $(\boldsymbol{\theta}_t | \mathbf{D}_{t-1})$ es la suma de dos variables aleatorias independientes Normales; por lo tanto, su distribución será Normal con media \mathbf{a}_t y matriz de varianza y covarianza R_t . Las formas funcionales de \mathbf{a}_t y R_t se obtienen a partir de:

$$\begin{aligned} E(\boldsymbol{\theta}_t | \mathbf{D}_{t-1}) &= E((G_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t) | \mathbf{D}_{t-1}), \\ &= G_t E(\boldsymbol{\theta}_{t-1} | \mathbf{D}_{t-1}) + E(\boldsymbol{\omega}_t | \mathbf{D}_{t-1}), \\ &= G_t m_{t-1}, \\ &= \mathbf{a}_t, \end{aligned} \tag{3.7}$$

la expresión para $Var(\boldsymbol{\theta}_t | \mathbf{D}_{t-1})$ está dado por:

$$\begin{aligned}
Var(\boldsymbol{\theta}_t | \mathbf{D}_{t-1}) &= Var((G_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t) | \mathbf{D}_{t-1}), \\
&= Var(G_t \boldsymbol{\theta}_{t-1} | \mathbf{D}_{t-1}) + Var(\boldsymbol{\omega}_t | \mathbf{D}_{t-1}), \\
&= G_t Var(\boldsymbol{\theta}_{t-1} | \mathbf{D}_{t-1}) G_t' + W_t, \\
&= G_t C_{t-1} G_t' + W_t, \\
&= \mathbf{R}_t,
\end{aligned} \tag{3.8}$$

para probar la parte (c), se partirá de (3.5). Dada la información en el instante \mathbf{D}_{t-1} se tiene que $(\mathbf{Y}_t | \mathbf{D}_{t-1})$ es la combinación lineal de dos variables aleatorias Normales independientes. Así, su distribución será también Normal, esto es $(\mathbf{Y}_t | \mathbf{D}_{t-1}) \sim N(f_t, Q_t)$. Las expresiones de f_t y Q_t son dados por:

$$\begin{aligned}
E(\mathbf{Y}_t | \mathbf{D}_{t-1}) &= E((F_t' \boldsymbol{\theta}_t + \boldsymbol{\nu}_t) | \mathbf{D}_{t-1}), \\
&= F_t' E(\boldsymbol{\theta}_t | \mathbf{D}_{t-1}) + E(\boldsymbol{\nu}_t | \mathbf{D}_{t-1}), \\
&= F_t' \mathbf{a}_t, \\
&= f_t,
\end{aligned} \tag{3.9}$$

la expresión para $Var(\mathbf{Y}_t | \mathbf{D}_{t-1})$ está dado por:

$$\begin{aligned}
Var(\mathbf{Y}_t | \mathbf{D}_{t-1}) &= Var((F_t' \boldsymbol{\theta}_t + \boldsymbol{\nu}_t) | \mathbf{D}_{t-1}), \\
&= F_t' Var(\boldsymbol{\theta}_t | \mathbf{D}_{t-1}) F_t' + Var(\boldsymbol{\nu}_t | \mathbf{D}_{t-1}), \\
&= F_t' R_t F_t' + V_t, \\
&= Q_t,
\end{aligned} \tag{3.10}$$

Para demostrar (d) se usará el hecho que la distribución condicional de \mathbf{Y}_t y $\boldsymbol{\theta}_t$ dado

D_{t-1} se distribuye Normalmente. Luego se tiene que:

$$\begin{aligned}
Cov[\mathbf{Y}_t, \boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}] &= Cov[F_t' \boldsymbol{\theta}_t + \boldsymbol{\nu}_t, \boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}], \\
&= F_t' Var[\boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}] + Cov[\boldsymbol{\nu}_t, \boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}], \\
&= F_t' R_t, \\
&= A_t Q_t,
\end{aligned}$$

donde $A_t = R_t F_t' Q_t^{-1}$, luego se usarán los resultados (b) y(c) del teorema 3.1, se tiene que:

$$\begin{pmatrix} \mathbf{Y}_t \\ \boldsymbol{\theta}_t \end{pmatrix} \mid \mathbf{D}_{t-1} \sim N \left[\begin{pmatrix} f_t \\ \mathbf{a}_t \end{pmatrix}, \begin{pmatrix} Q_t & Q_t' A_t' \\ A_t Q_t & R_t \end{pmatrix} \right]$$

Se utilizará las propiedades de la distribución Normal Multivariada (ver apéndice A1) se tiene que:

$$\left(\boldsymbol{\theta}_t \mid \mathbf{Y}_t, \mathbf{D}_{t-1} \right) \sim N \left(\mathbf{a}_t + A_t Q_t Q_t^{-1} (\mathbf{Y}_t - f_t), R_t - A_t Q_t Q_t^{-1} Q_t' A_t' \right)$$

$$\left(\boldsymbol{\theta}_t \mid \mathbf{D}_t \right) \sim N \left(\mathbf{a}_t + A_t (\mathbf{Y}_t - f_t), R_t - A_t Q_t A_t' \right).$$

donde:

$$m_t = \mathbf{a}_t + A_t (\mathbf{Y}_t - f_t),$$

$$C_t = R_t - A_t A_t' Q_t.$$

3.4 Distribuciones Predictivas

Definición 3.3. Para cualquier tiempo t , la función predictiva $f_t(k)$ está definida para todo entero $k \geq 0$ como:

$$f_t(k) = E[\mu_{t+k} \mid \mathbf{D}_t] = E[F_{t+k}' \boldsymbol{\theta}_{t+k} \mid \mathbf{D}_t]$$

donde:

$$\mu_{t+k} = F'_{t+k} \boldsymbol{\theta}_{t+k}$$

es la función de respuesta media.

Para $k > 0$, la función predictiva proporciona el valor esperado de las observaciones futuras \mathbf{Y}_{t+k} , dado \mathbf{D}_t . Es decir:

$$f_t(k) = E[\mathbf{Y}_{t+k} | \mathbf{D}_t], \quad (k \geq 1),$$

Para completar, la definición está dada en términos del valor esperado de la función de respuesta media μ_{t+k} en vez de $\mathbf{Y}_{t+k} = \mu_{t+k} + \boldsymbol{\nu}_{t+k}$. Al hacer esto la definición incluye el caso $k = 0$, dando $f_t(0) = E[\mu_t | \mathbf{D}_t]$.

Teorema 3.2. Para cada tiempo t y $k \geq 1$, la distribución de k -pasos al frente para $\boldsymbol{\theta}_{t+k}$ y \mathbf{Y}_{t+k} dado \mathbf{D}_t está dado por :

(a) Distribución de estados:

$$(\boldsymbol{\theta}_{t+k} | \mathbf{D}_t) \sim N[\mathbf{a}_t(k), \mathbf{R}_t(k)],$$

(b) Distribución predictiva:

$$(\mathbf{Y}_{t+k} | \mathbf{D}_t) \sim N[f_t(k), \mathbf{Q}_t(k)],$$

con momentos definidos recursivamente por:

$$f_t(k) = F'_t \mathbf{a}_t(k),$$

y

$$\mathbf{Q}_t(k) = F'_t \mathbf{R}_t(k) F'_t + \mathbf{V}_{t+k},$$

donde:

$$\mathbf{a}_t(k) = \mathbf{G}_{t+k} \mathbf{a}_t(k-1),$$

y

$$R_t(k) = G_{t+k}R_t(k-1)G'_{t+k} + W_{t+k},$$

con valores iniciales $\mathbf{a}_t(0) = m_t$, y $R_t(0) = C_t$.

Demostración.

Para todo t y entero $r < t$, se define las matrices de dimensión $n \times n$ $H_t(r) = G_t G_{t-1} \dots G_{t-r+1}$, con $H_t(0) = I$. De la aplicación repetitiva de la ecuación de evolución de estado se puede expresar a θ_{t+k} como:

$$\theta_{t+k} = H_{t+k}(k)\theta_t + \sum_{r=1}^k H_{t+k}(k-r)\omega_{t+r}.$$

Se verificará la anterior ecuación utilizando el método de inducción, luego para $k = 1$ se tiene que:

$$\begin{aligned} \theta_{t+1} &= H_{t+1}(1)\theta_t + H_{t+1}(0)\omega_{t+1}, \\ \theta_{t+1} &= G_{t+1}\theta_t + \omega_{t+1}, \end{aligned}$$

se supone que se cumple para $k = s - 1$, entonces:

$$\theta_{t+s-1} = H_{t+s-1}(s-1)\theta_t + \sum_{r=1}^{s-1} H_{t+s-1}(s-1-r)\omega_{t+r}.$$

Se tiene que verificar que se cumple para $k = s$, es decir que:

$$\theta_{t+s} = H_{t+s}(s)\theta_t + \sum_{r=1}^s H_{t+s}(s-r)\omega_{t+r}.$$

Se utilizará el hecho que se puede expresar las siguientes igualdades

$$H_{t+s}(s) = G_{t+s}H_{t+s-1}(s-1) \quad \text{y} \quad H_{t+s}(s-1) = G_{t+s}H_{t+s-1}(s-2),$$

luego se tiene que:

$$\begin{aligned}\boldsymbol{\theta}_{t+s-1} &= H_{t+s-1}(s-1)\boldsymbol{\theta}_t + \sum_{r=1}^{s-1} H_{t+s-1}(s-1-r)\boldsymbol{\omega}_{t+r}, \\ G_{t+s}\boldsymbol{\theta}_{t+s-1} + \boldsymbol{\omega}_{t+s} &= G_{t+s}H_{t+s-1}(s-1)\boldsymbol{\theta}_t + \sum_{r=1}^{s-1} G_{t+s}H_{t+s-1}(s-1-r)\boldsymbol{\omega}_{t+r} \\ &\quad + \boldsymbol{\omega}_{t+s} \\ \boldsymbol{\theta}_{t+s} &= H_{t+s}(s)\boldsymbol{\theta}_t + \sum_{r=1}^s H_{t+s}(s-r)\boldsymbol{\omega}_{t+r}.\end{aligned}$$

Con lo cual quedaría verificado. De este modo, por linealidad e independencia de suma de variables aleatorias Normales se puede expresar que:

$$(\boldsymbol{\theta}_{t+k} \mid \mathbf{D}_t) \sim N[\mathbf{a}_t(k), R_t(k)],$$

Se tiene que demostrar que:

$$\mathbf{a}_t(k) = E[\boldsymbol{\theta}_{t+k} \mid \mathbf{D}_t] = G_{t+k}\mathbf{a}_t(k-1), \quad (3.11)$$

y

$$R_t(k) = \text{Var}[\boldsymbol{\theta}_{t+k} \mid \mathbf{D}_t] = G_{t+k}R_t(k-1)G'_{t+k} + W_{t+k}. \quad (3.12)$$

Se sabe que:

$$\begin{aligned}\boldsymbol{\theta}_{t+k} &= H_{t+k}(k)\boldsymbol{\theta}_t + \sum_{r=1}^k H_{t+k}(k-r)\boldsymbol{\omega}_{t+r} \\ &= G_{t+k}G_{t+k-1}\cdots G_{t+1}\boldsymbol{\theta}_t + \sum_{r=1}^k H_{t+k}(k-r)\boldsymbol{\omega}_{t+r}.\end{aligned}$$

Se sabe por el Teorema 3.1.d que:

$$E[\boldsymbol{\theta}_t \mid \mathbf{D}_t] = \mathbf{m}_t,$$

además $E[\omega_t] = 0$, luego por independencia $E[\omega_{t+r} | \mathbf{D}_t] = 0$ entonces:

$$\begin{aligned}
E[\boldsymbol{\theta}_{t+k} | \mathbf{D}_t] &= G_{t+k}G_{t+k-1} \dots G_{t+2}G_{t+1}E[\boldsymbol{\theta}_t | \mathbf{D}_t] + \sum_{r=1}^k H_{t+k}(k-r)E[\omega_{t+r} | \mathbf{D}_t], \\
&= G_{t+k}G_{t+k-1} \dots G_{t+2}G_{t+1}m_t + 0, \\
&= G_{t+k}G_{t+k-1} \dots G_{t+2}G_{t+1}\mathbf{a}_t(0), \\
&= G_{t+k}G_{t+k-1} \dots G_{t+2}\mathbf{a}_t(1), \\
&= G_{t+k}\mathbf{a}_t(k-1).
\end{aligned}$$

Con lo que quedaría demostrado (3.11).

Se utilizará nuevamente el Teorema 3.1.d

$$Var[\boldsymbol{\theta}_t | \mathbf{D}_t] = C_t,$$

Por independencia afirmamos que $Var[\omega_{t+r} | \mathbf{D}_t] = W_{t+r}$ entonces, también se puede expresar a $Var[\boldsymbol{\theta}_{t+k} | \mathbf{D}_t]$ como:

$$\begin{aligned}
Var[\boldsymbol{\theta}_{t+k} | \mathbf{D}_t] &= H_{t+k}(k)Var[\boldsymbol{\theta}_t | \mathbf{D}_t]H'_{t+k}(k) \\
&\quad + \sum_{r=1}^k H_{t+k}(k-r)Var[\omega_{t+r} | \mathbf{D}_t]H'_{t+k}(k-r), \\
R_t(k) &= H_{t+k}(k)C_tH'_{t+k}(k) \\
&\quad + \sum_{r=1}^k H_{t+k}(k-r)W_{t+r}H'_{t+k}(k-r), \\
&= G_{t+k}H_{t+k-1}(k-1)C_tH'_{t+k-1}(k-1)G'_{t+k} + W_{t+k} \\
&\quad + G_{t+k} \left[\sum_{r=1}^{k-1} H_{t+k-1}(k-1-r)W_{t+r}H'_{t+k-1}(k-1-r) \right] G'_{t+k}, \\
&= G_{t+k} \left[H_{t+k-1}(k-1)C_tH'_{t+k-1}(k-1) \right. \\
&\quad \left. + \sum_{r=1}^{k-1} H_{t+k-1}(k-1-r)W_{t+r}H'_{t+k-1}(k-1-r) \right] G'_{t+k} + W_{t+k}, \\
&= G_{t+k}R_t(k-1)G'_{t+k} + W_{t+k}.
\end{aligned}$$

Para la parte (b) al tiempo $t+k$, se tiene que:

$$\mathbf{Y}_{t+k} = F'_{t+k}\boldsymbol{\theta}_{t+k} + \nu_{t+k} \quad \nu_{t+k} \sim N(0, V_{t+k})$$

por lo demostrado en la parte (a) del presente Teorema, se tiene que:

$$E[\boldsymbol{\theta}_{t+k} | \mathbf{D}_t] = \mathbf{a}_t(k)$$

entonces:

$$\begin{aligned} E[\mathbf{Y}_{t+k} | \mathbf{D}_t] &= \mathbf{F}'_{t+k} E[\boldsymbol{\theta}_{t+k} | \mathbf{D}_t] + E[\nu_{t+k} | \mathbf{D}_t], \\ &= \mathbf{F}'_{t+k} \mathbf{a}_t(k), \\ &= f_t(k). \end{aligned} \tag{3.13}$$

También, se sabe que:

$$\text{Var}[\boldsymbol{\theta}_{t+k} | \mathbf{D}_t] = \mathbf{R}_t(k)$$

entonces:

$$\begin{aligned} \text{Var}[\mathbf{Y}_{t+k} | \mathbf{D}_t] &= \text{Var}[\mathbf{F}'_{t+k} \boldsymbol{\theta}_{t+k} | \mathbf{D}_t] + \text{Var}[\nu_{t+k} | \mathbf{D}_t], \\ &= \mathbf{F}'_{t+k} \text{Var}[\boldsymbol{\theta}_{t+k} | \mathbf{D}_t] \mathbf{F}_{t+k} + V_{t+k}, \\ &= \mathbf{F}'_{t+k} \mathbf{R}_t(k) \mathbf{F}_{t+k} + V_{t+k}, \\ &= \mathbf{Q}_t(k). \end{aligned} \tag{3.14}$$

Colorario 3.1.- Para cualquier entero j y k con $1 \leq j < k$,

$$\text{Cov}[\boldsymbol{\theta}_{t+k}, \boldsymbol{\theta}_{t+j} | \mathbf{D}_t] = \mathbf{C}_t(k, j)$$

luego

$$\text{Cov}[\mathbf{Y}_{t+k}, \mathbf{Y}_{t+j} | \mathbf{D}_t] = \mathbf{F}'_{t+k} \mathbf{C}_t(k, j) \mathbf{F}_{t+j},$$

donde, $\mathbf{C}_t(k, j)$ puede también ser definido recursivamente vía:

$$\mathbf{C}_t(r, j) = \mathbf{G}_{t+r} \mathbf{C}_t(r-1, j), \quad (r = j+1, \dots, k),$$

con valor inicial:

$$C_t(j, j) = R_t(j), \quad \forall t,$$

Demostración:

Para $r = j + 1, \dots, k$ se tiene que:

$$\begin{aligned} C_t(r, j) &= Cov[\boldsymbol{\theta}_{t+r}, \boldsymbol{\theta}_{t+j} \mid \mathbf{D}_t], \\ &= Cov[(G_{t+r}\boldsymbol{\theta}_{t+r-1} + \boldsymbol{w}_{t+r}, \boldsymbol{\theta}_{t+j}) \mid \mathbf{D}_t], \\ &= G_{t+r}C[\boldsymbol{\theta}_{t+r-1}, \boldsymbol{\theta}_{t+j} \mid \mathbf{D}_t] + Cov[\boldsymbol{w}_{t+r}, \boldsymbol{\theta}_{t+j} \mid \mathbf{D}_t], \\ &= G_{t+r}C_t(r-1, j). \end{aligned}$$

Se sabe que ν_t es el error de observación en el tiempo t , simplemente una perturbación aleatoria independiente de $\boldsymbol{\theta}_t$, luego:

$$Cov[\nu_{t+k}, \nu_{t+j} \mid \mathbf{D}_t] = 0; \quad Cov[\nu_{t+k}, \boldsymbol{\theta}_{t+j} \mid \mathbf{D}_t] = 0; \quad Cov[\boldsymbol{\theta}_{t+k}, \nu_{t+j} \mid \mathbf{D}_t] = 0,$$

por otra parte,

$$\begin{aligned} Cov[Y_{t+k}, Y_{t+j} \mid \mathbf{D}_t] &= Cov[(F'_{t+k}\boldsymbol{\theta}_{t+k} + \nu_{t+k}, F'_{t+k}\boldsymbol{\theta}_{t+j} + \omega_{t+j}) \mid \mathbf{D}_t], \\ &= Cov[F'_{t+k}\boldsymbol{\theta}_{t+k}, F'_{t+j}\boldsymbol{\theta}_{t+j} \mid \mathbf{D}_t] + Cov[\nu_{t+k}, F'_{t+j}\boldsymbol{\theta}_{t+j} \mid \mathbf{D}_t] \\ &\quad + Cov[F'_{t+k}\boldsymbol{\theta}_{t+k}, \nu_{t+j} \mid \mathbf{D}_t] + Cov[\nu_{t+k}, \nu_{t+j} \mid \mathbf{D}_t], \\ &= F'_{t+k}Cov[\boldsymbol{\theta}_{t+k}, \boldsymbol{\theta}_{t+k} \mid \mathbf{D}_t]F_{t+j} + F'_{t+j}Cov[\nu_{t+k}, \boldsymbol{\theta}_{t+j} \mid \mathbf{D}_t] \\ &\quad + F'_{t+k}Cov[\boldsymbol{\theta}_{t+k}, \nu_{t+j} \mid \mathbf{D}_t] + Cov[\nu_{t+k}, \nu_{t+j} \mid \mathbf{D}_t], \\ &= F'_{t+k}Cov[\boldsymbol{\theta}_{t+k}, \boldsymbol{\theta}_{t+j} \mid \mathbf{D}_t]F_{t+j}, \\ &= F'_{t+k}C_t(k, j)F_{t+j}. \end{aligned}$$

Con lo que quedaría verificado el corolario 3.1.

Corolario 3.2.- En el caso especial que la matriz de evolución G_t sea constante, $G_t = G$ para todo t , entonces para $k \geq 0$,

$$\mathbf{a}_t(k) = G^k \mathbf{m}_t,$$

además:

$$f_t(k) = F'_{t+k} G^k m_t,$$

Demostración:

Se sabe del Teorema 3.2, que:

$$a_t(k) = G_{t+k} a_t(k-1),$$

luego:

$$\begin{aligned} a_t(k) &= G_{t+k} G_{t+k-1} \dots G_{t+2} G_{t+1} a_t(0), \\ &= GG \dots G m_t, \\ &= G^k m_t. \end{aligned}$$

Por la definición (3.3) se tiene que:

$$f_t(k) = E[F'_{t+k} \theta_{t+k} | D_t] = F'_{t+k} E[\theta_{t+k} | D_t],$$

del Teorema 3.2 $E[\theta_{t+k} | D_t] = a_t(k)$ entonces:

$$f_t(k) = F'_{t+k} a_t(k) = F'_{t+k} G^k m_t$$

Adicionalmente, se $F_t = F$ para todo t , entonces la función predictiva tiene la forma:

$$f_t(k) = F' G^k m_t$$

3.5 Varianza Observacional

Hasta ahora se ha asumido la cuádrupla del MLD univariado conocido para todo tiempo t . Las matrices de regresión y evolución están definidas por el modelador en concordancia con el modelo designado. La matriz de varianza-covarianza de evolución

es también escogida por el modelador. El elemento final de la cuádrupla, la varianza observacional V_t , está también, completamente sujeto a incertidumbre. Acorde, con esto, se presentará el procedimiento Bayesiano para varianzas observacionales desconocidas. Luego se tratará el caso especial en el cual la varianza es constante pero desconocida $V_t = V$. Sin embargo para facilitar la identificación de las posteriores se trabajará en términos del parámetro de precisión $\phi = V^{-1}$. El análisis aquí está basado en las familias conjugadas. El hecho clave del análisis es que todas las varianzas-covarianza en la definición y análisis del modelo en cada tiempo es escalada por V .

Definición 3.4. Para cada t , el modelo está definido por:

Ecuación de observación:

$$\mathbf{Y}_t = F_t' \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim N[0, V]. \quad (3.15)$$

Ecuación de sistema:

$$\boldsymbol{\theta}_t = G_t \boldsymbol{\theta}_{t-1} + \omega_t, \quad \omega_t \sim N[0, V W_t^*]. \quad (3.16)$$

Con información inicial:

$$\begin{aligned} (\boldsymbol{\theta}_0 | \mathbf{D}_0, \phi) &\sim N[m_0, V C_0^*] \\ (\phi | \mathbf{D}_0) &\sim G[n_0/2, d_0/2], \end{aligned}$$

donde $\phi = V^{-1}$

Aquí las cantidades iniciales m_0 , C_0^* , n_0 y d_0 deben ser especificadas, así como las matrices W_t^* . Note que todas las varianzas y covarianzas en el modelo tienen a V como un multiplicador, o factor de escala, proporcionando un modelo a escala-libre en términos de varianzas de escala-libre dadas por C_0^* y el W_t^* . Para V conocido el modelo coincide con la definición 3.2 pues V es absorbido en estas matrices.

La suposición usual de independencia se cumple, ahora condicionada a V ó ϕ . La priori para ϕ tiene media dado por,

$E[\phi | \mathbf{D}_{t-1}] = n_0/d_0 = 1/S_0$, donde S_0 es la estimativa puntual a priori para la varianza observacional V .

Teorema 3.3. Con el modelo especificado en la definición 3.4, resultan las siguientes distribuciones obtenidas en cada tiempo $t \geq 1$.

(a) Condicional a V :

$$\begin{aligned}(\boldsymbol{\theta}_{t-1} | \mathbf{D}_{t-1}, V) &\sim N[m_{t-1}, VC_{t-1}^*], \\(\boldsymbol{\theta}_t | \mathbf{D}_{t-1}, V) &\sim N[\mathbf{a}_t, VR_t^*], \\(\mathbf{Y}_t | \mathbf{D}_{t-1}, V) &\sim N[f_t, VQ_t^*], \\(\boldsymbol{\theta}_t | \mathbf{D}_t, V) &\sim N[m_t, VC_t^*],\end{aligned}$$

con $\mathbf{a}_t = G_t m_{t-1}$ y $R_t^* = G_t C_{t-1}^* G_t' + W_t^*$ y $f_t = F_t' \mathbf{a}_t$, y $Q_t^* = 1 + F_t' R_t^* F_t$. Aquí se define m_t y C_t^* como:

$$\begin{aligned}m_t &= a_t + A_t e_t, \\C_t^* &= R_t^* - A_t A_t' Q_t^*,\end{aligned}$$

con $e_t = \mathbf{Y}_t - f_t$ y $A_t = R_t^* F_t' / Q_t^*$.

(b) Para la precisión $\phi = V^{-1}$

$$\begin{aligned}(\phi | \mathbf{D}_{t-1}) &\sim G[n_{t-1}/2, d_{t-1}/2], \\(\phi | \mathbf{D}_t) &\sim G[n_t/2, d_t/2],\end{aligned}$$

donde $n_t = n_{t-1} + 1$ y $d_t = d_{t-1} + e_t^2 / Q_t^*$.

(c) Incondicional a V:

$$\begin{aligned}(\boldsymbol{\theta}_{t-1} \mid \mathbf{D}_{t-1}) &\sim T_{n_{t-1}}[m_{t-1}, C_{t-1}], \\(\boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}) &\sim T_{n_t}[a_t, R_t], \\(\mathbf{Y}_t \mid \mathbf{D}_{t-1}) &\sim T_{n_t}[f_t, Q_t], \\(\boldsymbol{\theta}_t \mid \mathbf{D}_t) &\sim T_{n_t}[m_t, C_t],\end{aligned}$$

donde $C_{t-1} = S_{t-1}C_{t-1}^*$, $R_t = S_{t-1}R_t^*$, $Q_t = S_{t-1}Q_t^*$ y $C_t = S_tC_t^*$, con $S_{t-1} = d_{t-1}/n_{t-1}$ y $S_t = d_t/n_t$.

(d) Definición operacional de las ecuaciones de actualización:

$$\begin{aligned}m_t &= a_t + A_t e_t, \\C_t &= (S_t/S_{t-1})[R_t - A_t A_t' a_t], \\S_t &= d_t/n_t,\end{aligned}$$

con $n_t = n_{t-1} + 1$ y $d_t = d_{t-1} + S_{t-1}e_t^2/Q_t$,

donde $Q_t = S_{t-1} + F_t' R_t F_t$ y $A_t = R_t F_t / Q_t$.

Demostración.

Para demostrar la parte (a) del Teorema (3.3), se asume que se cumple:

$$(\boldsymbol{\theta}_{t-1} \mid \mathbf{D}_{t-1}, V) \sim N[m_{t-1}, VC_{t-1}^*].$$

Además se sabe que

$$\boldsymbol{\theta}_t = G_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t \quad \boldsymbol{\omega}_t \sim N(0, V\mathbf{W}_t^*).$$

Ahora $\omega_t \sim N(0, V\mathbf{W}_t^*)$ y por independencia de ω_t y ν_t , se tiene que $(\omega_t \mid \mathbf{D}_{t-1}, V) \sim N(0, V\mathbf{W}_t^*)$. Así $(\theta_t \mid \mathbf{D}_{t-1}, V)$ es la suma de dos variables aleatorias independientes Normales con media.

$$\begin{aligned}
E(\theta_t \mid \mathbf{D}_{t-1}, V) &= E((G_t\theta_{t-1} + \omega_t) \mid \mathbf{D}_{t-1}, V), \\
&= E(G_t\theta_{t-1} \mid \mathbf{D}_{t-1}, V) + E(\omega_t \mid \mathbf{D}_{t-1}, V), \\
&= G_tE(\theta_{t-1} \mid \mathbf{D}_{t-1}, V) + 0, \\
&= G_tm_{t-1}, \\
&= \mathbf{a}_t.
\end{aligned} \tag{3.17}$$

La expresión para $Var(\theta_t \mid \mathbf{D}_{t-1}, V)$ está dado por:

$$\begin{aligned}
Var(\theta_t \mid \mathbf{D}_{t-1}, V) &= Var((G_t\theta_{t-1} + \omega_t) \mid \mathbf{D}_{t-1}, V), \\
&= Var(G_t\theta_{t-1} \mid \mathbf{D}_{t-1}, V) + Var(\omega_t \mid \mathbf{D}_{t-1}, V), \\
&= G_tVar(\theta_{t-1} \mid \mathbf{D}_{t-1}, V)G_t' + V\mathbf{W}_t^*, \\
&= G_tVC_{t-1}^*G_t' + V\mathbf{W}_t^*, \\
&= V(G_tC_{t-1}^*G_t' + \mathbf{W}_t^*), \\
&= VR_t^*.
\end{aligned} \tag{3.18}$$

Por lo tanto, se concluye que:

$$(\theta_t \mid \mathbf{D}_{t-1}, V) \sim N(\mathbf{a}_t, VR_t^*).$$

Usando esta última expresión, se puede también concluir que

$$(F_t'\theta_t \mid \mathbf{D}_{t-1}, V) \sim N(g_t, P_t),$$

donde:

$$\begin{aligned}
 g_t &= E(F_t' \boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}, V), \\
 &= F_t' E(\boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}, V), \\
 &= F_t' \mathbf{a}_t.
 \end{aligned} \tag{3.19}$$

También:

$$\begin{aligned}
 P_t &= Var(F_t' \boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}, V), \\
 &= F_t' Var(\boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}, V) F_t, \\
 &= F_t' V R_t^* F_t.
 \end{aligned} \tag{3.20}$$

Además se sabe que $(\boldsymbol{\nu}_t \mid \mathbf{D}_{t-1}, V) \sim N(0, V)$, así

$$(F_t' \boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}, V) + (\boldsymbol{\nu}_t \mid \mathbf{D}_{t-1}, V) \sim N(F_t' \mathbf{a}_t, F_t' V R_t^* F_t + V),$$

esto es,

$$(\mathbf{Y}_t \mid \mathbf{D}_{t-1}, V) \sim N(F_t' \mathbf{a}_t, V(F_t' R_t^* F_t + 1)),$$

luego

$$f_t = F_t' \mathbf{a}_t \qquad y \qquad Q_t^* = F_t' R_t^* F_t + 1.$$

Por otra parte, como se sabe que $(\mathbf{Y}_t \mid \mathbf{D}_{t-1}, \boldsymbol{\theta}_t, \mathbf{V}) = (\mathbf{Y}_t \mid \boldsymbol{\theta}_t, \mathbf{V})$ se distribuye Normalmente, y que $(\boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}, \mathbf{V})$ también tiene una distribución Normal, es posible concluir que la distribución conjunta $(\mathbf{Y}_t, \boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}, \mathbf{V})$ es también Normal con matriz de varianza-covarianza dado por:

$$\begin{aligned}
 Cov[\mathbf{Y}_t, \boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}, V] &= Cov[F_t' \boldsymbol{\theta}_t + \boldsymbol{\nu}_t, \boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}, V] \\
 &= F_t' V [\boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}, V] + 0' \\
 &= F_t' V R_t^*,
 \end{aligned}$$

esto es,

$$\begin{pmatrix} \mathbf{Y}_t \\ \boldsymbol{\theta}_t \end{pmatrix} | V, \mathbf{D}_{t-1} \sim N \left[\begin{pmatrix} \mathbf{f}_t \\ \mathbf{a}_t \end{pmatrix} \begin{pmatrix} VQ_t^* & F_t'VR_t^* \\ R_t^*VF_t & VR_t^* \end{pmatrix} \right],$$

a partir de esta última expresión se puede obtener $(\boldsymbol{\theta}_t | V, \mathbf{D}_t)$ usando resultados del (apéndice A1) es decir

$$\begin{aligned} (\boldsymbol{\theta}_t | V, \mathbf{Y}_t, \mathbf{D}_{t-1}) &\sim N \left(\mathbf{a}_t + R_t^*VF_t(VQ_t^*)^{-1}(\mathbf{Y}_t - \mathbf{f}_t), V(R_t^* - R_t^*F_tQ_t^{*-1}F_t'R_t^*) \right), \\ (\boldsymbol{\theta}_t | V, \mathbf{D}_t) &\sim N \left(\mathbf{a}_t + A_t(\mathbf{Y}_t - \mathbf{f}_t), V(R_t^* - A_tA_t'Q_t^*) \right), \end{aligned}$$

si:

$$A_t = R_t^*F_t/Q_t^* \quad y \quad e_t = \mathbf{Y}_t - \mathbf{f}_t,$$

se tiene que:

$$m_t = \mathbf{a}_t + A_t e_t$$

y

$$C_t^* = R_t^* - A_tA_t'Q_t^*$$

es decir,

$$(\boldsymbol{\theta}_t | \mathbf{D}_t, V) \sim N(m_t, VC_t^*).$$

Con lo que finalmente quedaría demostrado la parte (a) del Teorema 3.3.

Para la parte (b) del Teorema 3.3, se supone que:

$$(\phi | \mathbf{D}_{t-1}) \sim G\left[\frac{n_{t-1}}{2}, \frac{d_{t-1}}{2}\right],$$

en términos de f.d.p se tiene:

$$p(\phi | \mathbf{D}_{t-1}) \propto \phi^{\frac{n_{t-1}}{2}-1} \exp(-0.5\phi d_{t-1}), \quad (3.21)$$

de la parte (a) del Teorema 3.3 se tiene que:

$$(\mathbf{Y}_t | \mathbf{D}_{t-1}, \phi) \sim N[f_t, \mathbf{Q}_t^* \phi^{-1}].$$

Lo anterior se puede expresar como:

$$p(\mathbf{Y}_t | \mathbf{D}_{t-1}, \phi) = (2\pi\mathbf{Q}_t^*/\phi)^{-\frac{1}{2}} \exp(-0.5(\mathbf{Y}_t - f_t)^2/(\mathbf{Q}_t^*/\phi)),$$

De lo anterior se concluye que:

$$p(\mathbf{Y}_t | \mathbf{D}_{t-1}, \phi) \propto \phi^{-\frac{1}{2}} \exp(-0.5\phi e_t^2/\mathbf{Q}_t^*). \quad (3.22)$$

Una forma de expresar el Teorema de Bayes es:

$$p(\phi | \mathbf{D}_t) \propto P(\phi | \mathbf{D}_{t-1})P(\mathbf{Y}_t | \phi). \quad (3.23)$$

Se puede expresar $p(\mathbf{Y}_t | \phi)$ como:

$$p(\mathbf{Y}_t | \phi) = p(\mathbf{Y}_t | \phi, \mathbf{D}_{t-1}). \quad (3.24)$$

Luego se tiene:

$$p(\phi | \mathbf{D}_t) \propto p(\phi | \mathbf{D}_{t-1})P(\mathbf{Y}_t | \phi, \mathbf{D}_{t-1}),$$

reemplazando en la anterior ecuación se tiene que:

$$\begin{aligned} p(\phi | \mathbf{D}_t) &\propto \phi^{\frac{n_t-1}{2}-1} \exp(-0.5\phi d_{t-1}) \phi^{1/2} \exp(-0.5\phi e_t^2/\mathbf{Q}_t^*), \\ &\propto \phi^{\frac{n_t-1+1}{2}-1} \exp(-(d_{t-1} + e_t^2/\mathbf{Q}_t^*)\phi/2), \end{aligned}$$

luego se puede deducir que $(\phi | \mathbf{D}_t) \sim \mathbf{G}[n_t/2, d_t/2]$ donde los parámetros son $n_t = n_{t-1} + 1$ y $d_t = d_{t-1} + e_t^2/\mathbf{Q}_t^*$.

Con lo cual quedaría demostrado la 2da parte del Teorema 3.3 parte(b).

En el caso de la 1ra parte del Teorema 3.3 parte (c) se usará la 1ra parte del Teorema 3.3 parte(a), haciendo $\phi = V^{-1}$ luego se tiene que:

$$(\theta_{t-1} \mid \mathbf{D}_{t-1}, \phi) \sim N[m_t, C_{t-1}^* \mid \phi]$$

De la parte (b) del Teorema 3.3:

$$(\phi \mid \mathbf{D}_{t-1}) \sim G[n_{t-1}/2, d_{t-1}/2],$$

entonces se puede afirmar que $(\theta_{t-1} \mid \mathbf{D}_{t-1})$ tiene una distribución marginal T ver (apéndice A2)

$$\begin{aligned} (\theta_{t-1} \mid \mathbf{D}_{t-1}) &\sim T_{n_{t-1}}[m_{t-1}, C_{t-1}^* d_{t-1}/n_{t-1}] \\ &\sim T_{n_{t-1}}[m_{t-1}, C_{t-1}^* S_{t-1}^*]. \end{aligned}$$

Si $S_{t-1} = d_{t-1}/n_{t-1}$ y $C_{t-1} = S_{t-1} C_{t-1}^*$, se tiene que:

$$(\theta_{t-1} \mid \mathbf{D}_{t-1}) \sim T_{n_{t-1}}[m_{t-1}, C_{t-1}].$$

Con lo que quedaría demostrado la 1ra parte del Teorema 3.3 parte(c).

En el caso de la 2da parte del Teorema 3.3 parte(c) se usará la 2da parte del Teorema 3.3 parte(a), haciendo $\phi = V^{-1}$ se tiene que:

$$(\theta_t \mid \mathbf{D}_{t-1}, \phi) \sim N[a_t, R_t^* / \phi],$$

de la 1ra parte del Teorema 3.3 parte(b)

$$(\phi \mid \mathbf{D}_{t-1}) \sim G[n_{t-1}/2, d_{t-1}/2],$$

entonces se puede afirmar que $(\theta_t \mid \mathbf{D}_{t-1})$ tiene una distribución marginal T de la forma:

$$\begin{aligned} (\theta_t \mid \mathbf{D}_{t-1}) &\sim T_{n_{t-1}}[a_t, R_t^* d_{t-1}/n_{t-1}], \\ &\sim T_{n_{t-1}}[a_t, R_t^* S_{t-1}^*], \end{aligned}$$

si $\mathbf{R}_t = \mathbf{S}_{t-1}\mathbf{R}_t^*$, se tiene que:

$$(\boldsymbol{\theta}_{t-1} \mid \mathbf{D}_{t-1}) \sim T_{n_{t-1}}[\mathbf{a}_t, \mathbf{R}_t].$$

Con lo que quedaría demostrado la 2da parte del Teorema 3.3 parte (c).

En el caso de la 3ra parte del Teorema 3.3 parte(c) se usará la 3ra parte del Teorema 3.3 parte(a), haciendo $\boldsymbol{\phi} = \mathbf{V}^{-1}$, se tiene que:

$$(\mathbf{Y}_t \mid \mathbf{D}_{t-1}, \boldsymbol{\phi}) \sim N[\mathbf{f}_t, \mathbf{Q}_t^*/\boldsymbol{\phi}].$$

De la 1ra parte del Teorema 3.3 parte(b)

$$(\boldsymbol{\phi} \mid \mathbf{D}_{t-1}) \sim \mathbf{G}[n_{t-1}/2, d_{t-1}/2].$$

Similar al caso anterior se puede afirmar que:

$$\begin{aligned} (\mathbf{Y}_t \mid \mathbf{D}_{t-1}) &\sim T_{n_{t-1}}[\mathbf{f}_t, \mathbf{Q}_t^* d_{t-1}/n_{t-1}], \\ &\sim T_{n_{t-1}}[\mathbf{f}_t, \mathbf{Q}_t^* \mathbf{S}_{t-1}^*], \end{aligned}$$

donde si $\mathbf{Q}_t = \mathbf{S}_{t-1}\mathbf{Q}_t^*$, tiene que:

$$(\mathbf{Y}_t \mid \mathbf{D}_{t-1}) \sim T_{n_{t-1}}[\mathbf{f}_t, \mathbf{Q}_t].$$

Con lo que quedaría demostrado la 3ra parte del Teorema 3.3 parte(c).

En el caso de la 4ta parte del Teorema 3.3 parte(c) se usará la 4ta parte del Teorema 3.3 parte(a), haciendo $\boldsymbol{\phi} = \mathbf{V}^{-1}$ se tiene que:

$$(\boldsymbol{\theta}_t \mid \mathbf{D}_t, \boldsymbol{\phi}) \sim N[\mathbf{m}_t, \mathbf{C}_t^*/\boldsymbol{\phi}]$$

De la 2da parte del Teorema 3.3 parte(b)

$$(\boldsymbol{\phi} \mid \mathbf{D}_t) \sim \mathbf{G}[n_t/2, d_t/2].$$

En forma similar al caso anterior se puede afirmar que $(\theta_t | D_t)$ tiene una distribución marginal T

$$\begin{aligned}(\theta_t | D_t) &\sim T_{n_t}[m_t, C_t^* d_t/n_t], \\ &\sim T_{n_t}[m_t, C_t^* S_t^*],\end{aligned}$$

donde si $Q_t = S_t C_t^*$, luego se tiene que:

$$(\theta_t | D_t) \sim T_{n_t}[m_t, C_t].$$

Con lo que quedaría demostrado la 4ta parte del Teorema 3.3 parte(c).

3.6 Recurrencias para Filtración

La actualización secuencial y componentes de pronóstico del MLD están dirigidas a hacer inferencias acerca de los estados θ_t del proceso de las Series de Tiempo, y acerca de observaciones en $t, t + 1, t + 2, \dots, t + k$, ($k > 0$) basado sobre la información disponible en el tiempo t . Sin embargo, existe requerimientos complementarios para hacer inferencias acerca de los estados del proceso en el **pasado**, en tiempos $t-k$ para $k > 0$. En procesos de monitoramiento y control, por ejemplo, es de interés comparar el nivel actual de la Serie μ_t con los estados anteriores μ_{t-k} , para evaluar los cambios en nivel sobre el tiempo. De igual forma, para series mensuales con una estructura estacional durante el año, las componentes estacionales del MLD sobre los pasados 12 meses proporcionan información de interés en reportar formas estacionales, mientras que valores de los restantes componentes del modelo dan una Serie desestacionalizadas de interés para evaluar tendencias. En estos casos, las inferencias requeridas se dirigen a θ_{t-k} , $k = 0, 1, \dots, 11$, condicional sobre la información actual D_t .

Para la terminología usada en el presente trabajo, la distribución de $(\boldsymbol{\theta}_{t-k} \mid \mathbf{D}_t)$ para $k \geq 1$ y cualquier t fijo, es llamada Distribución Filtrada $k - pasos$ para el vector de estados en el tiempo t , análogo a la Distribución predictiva $k - pasos$ al frente. El acto de usar datos recientes para revisar información acerca de los valores anteriores, del vector de estados es llamado Filtración; La información recientemente obtenida es filtrada hacia atrás. Un concepto relacionado a Filtración es el de suavisar una Serie de Tiempo. La estimación retrospectiva de la evolución histórica de la función respuesta media μ_t de la Serie de Tiempo usando Distribuciones Filtradas $(\mu_{t-k} \mid \mathbf{D}_t)$ para $k \geq 1$ es llamada suavización de la serie.

En cualquier tiempo t , las Distribuciones Filtradas son derivadas recursivamente hacia atrás en el tiempo, usando relaciones entre ellas, que serán demostradas a continuación, que son similares en estructuras a las ecuaciones de actualización secuencial padrón. Para $k \geq 1$, se extiende la definición de Distribuciones predictivas $k - pasos$ al frente con momentos $\boldsymbol{\theta}_t(k)$ y $R_t(k)$ para argumentos negativos $\boldsymbol{a}_t(-k)$ y $R_t(-k)$, y demostraremos los siguientes resultados.

Teorema 3.4. En el MLD univariado $\{F_t, G_t, V_t, W_t\}$, defina:

$$\mathbf{B}_t = \mathbf{C}_t \mathbf{G}'_{t+1} \mathbf{R}_{t+1}^{-1}$$

para todo t . Entonces, para todo k , ($1 \leq k < t$), las distribuciones filtradas están definidas por:

$$(\boldsymbol{\theta}_{t-k} \mid \mathbf{D}_t) \sim N[\boldsymbol{a}_t(-k), R_t(-k)],$$

donde:

$$\boldsymbol{a}_t(-k) = m_{t-k} + \mathbf{B}_{t-k}[\boldsymbol{a}_t(-k+1) - \boldsymbol{a}_{t-k+1}]$$

y

$$R_t(-k) = C_{t-k} - \mathbf{B}_{t-k}[R_{t-k+1} - R_t(-k+1)]\mathbf{B}'_{t-k},$$

con valores iniciales dados por $\mathbf{a}_t(0) = \mathbf{m}_t$, y $R_t(0) = C_t$. También $\mathbf{a}_{t-k+1} = \mathbf{a}_{t-k}(1)$ y $R_{t-k+1} = R_{t-k}(1)$.

Demostración.

Considere la siguiente expresión:

$$\begin{aligned} p(\boldsymbol{\theta}_{t-k} \mid \mathbf{D}_t) &= \int p(\boldsymbol{\theta}_{t-k}, \boldsymbol{\theta}_{t-k+1} \mid \mathbf{D}_t) d\boldsymbol{\theta}_{t-k+1}, \\ &= \int p(\boldsymbol{\theta}_{t-k} \mid \boldsymbol{\theta}_{t-k+1}, \mathbf{D}_t) p(\boldsymbol{\theta}_{t-k+1} \mid \mathbf{D}_t) d\boldsymbol{\theta}_{t-k+1}, \end{aligned} \quad (3.25)$$

la anterior expresión sugiere trabajar por inducción, luego se asumirá como cierto para $k - 1$ el resultado del Teorema, así se tiene que el segundo término en el integrando de (3.25) es:

$$(\boldsymbol{\theta}_{t-k+1} \mid \mathbf{D}_t) \sim N[\mathbf{a}_t(-k + 1), R_t(-k + 1)].$$

Se sabe que $\mathbf{D}_t = \{\mathbf{Y}_t, \mathbf{Y}_{t-1}, \dots, \mathbf{Y}_{t-k+1}, \mathbf{D}_{t-k}\}$, y por el Teorema de Bayes

$$p(\boldsymbol{\theta}_{t-k} \mid \boldsymbol{\theta}_{t-k+1}, \mathbf{D}_t) = \frac{p(\boldsymbol{\theta}_{t-k} \mid \boldsymbol{\theta}_{t-k+1}, \mathbf{D}_{t-k}) p(\mathbf{Y}_{t-k+1}, \dots, \mathbf{Y}_t \mid \boldsymbol{\theta}_{t-k}, \boldsymbol{\theta}_{t-k+1}, \mathbf{D}_{t-k})}{p(\mathbf{Y}_{t-k+1}, \dots, \mathbf{Y}_t \mid \boldsymbol{\theta}_{t-k+1}, \mathbf{D}_{t-k})}.$$

Ahora, dado $\boldsymbol{\theta}_{t-k+1}$, los datos $\{\mathbf{Y}_{t-k+1}, \dots, \mathbf{Y}_t\}$ es independiente del valor previo $\boldsymbol{\theta}_{t-k}$ luego entonces se puede cancelar los dos últimos términos quedando la siguiente expresión:

$$p(\boldsymbol{\theta}_{t-k} \mid \boldsymbol{\theta}_{t-k+1}, \mathbf{D}_t) = p(\boldsymbol{\theta}_{t-k} \mid \boldsymbol{\theta}_{t-k+1}, \mathbf{D}_{t-k}).$$

por el Teorema de Bayes se tiene que:

$$p(\boldsymbol{\theta}_{t-k} \mid \boldsymbol{\theta}_{t-k+1}, \mathbf{D}_{t-k}) \propto p(\boldsymbol{\theta}_{t-k} \mid \mathbf{D}_{t-k}) p(\boldsymbol{\theta}_{t-k+1} \mid \boldsymbol{\theta}_{t-k}, \mathbf{D}_{t-k}). \quad (3.26)$$

Por otro lado tenemos de la definición (3.4) que:

$$\boldsymbol{\theta}_t = G_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t \quad \forall \quad t.$$

$$\begin{aligned}
E(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1}, \mathbf{D}_{t-1}) &= E(G_t \boldsymbol{\theta}_{t-1} | \boldsymbol{\theta}_{t-1}, \mathbf{D}_{t-1}) + E(\boldsymbol{\omega}_t | \boldsymbol{\theta}_{t-1}, \mathbf{D}_{t-1}), \\
&= G_t E(\boldsymbol{\theta}_{t-1} | \boldsymbol{\theta}_{t-1}, \mathbf{D}_{t-1}) + 0, \\
&= G_t \boldsymbol{\theta}_{t-1}.
\end{aligned}$$

En particular cuando $t = t - k + 1$ se tiene que:

$$E(\boldsymbol{\theta}_{t-k+1} | \boldsymbol{\theta}_{t-k}, \mathbf{D}_{t-k}) = G_{t-k+1} \boldsymbol{\theta}_{t-k}.$$

Además:

$$\begin{aligned}
Var(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1}, \mathbf{D}_{t-1}) &= Var(G_t \boldsymbol{\theta}_{t-1} | \boldsymbol{\theta}_{t-1}, \mathbf{D}_{t-1}) + Var(\boldsymbol{\omega}_t | \boldsymbol{\theta}_{t-1}, \mathbf{D}_{t-1}), \\
&= 0 + \mathbf{W}_t, \\
&= \mathbf{W}_t,
\end{aligned}$$

en particular cuando $t = t - k + 1$ se tiene que:

$$Var(\boldsymbol{\theta}_{t-k+1} | \boldsymbol{\theta}_{t-k}, \mathbf{D}_{t-k}) = \mathbf{W}_{t-k+1}.$$

Se sabe que la suma de dos v.a. Normales da una v.a. Normal, entonces se puede concluir que:

$$(\boldsymbol{\theta}_{t-k+1} | \boldsymbol{\theta}_{t-k}, \mathbf{D}_{t-k}) \sim N[G_{t-k+1} \boldsymbol{\theta}_{t-k}, \mathbf{W}_{t-k+1}].$$

Además como

$$(\boldsymbol{\theta}_{t-k} | \mathbf{D}_{t-k}) \sim N[m_{t-k}, C_{t-k}]. \quad (3.27)$$

(3.26) y (3.27) define una Distribución conjunta Normal para $(\boldsymbol{\theta}_{t-k}, \boldsymbol{\theta}_{t-k+1} | \mathbf{D}_{t-k})$,

$$\begin{aligned}
&\left(\begin{array}{c} \boldsymbol{\theta}_{t-k} \\ \boldsymbol{\theta}_{t-k+1} \end{array} \middle| \mathbf{D}_{t-k} \right) \sim \\
&N \left[\left(\begin{array}{c} E(\boldsymbol{\theta}_{t-k} | \mathbf{D}_{t-k}) \\ E(\boldsymbol{\theta}_{t-k+1} | \mathbf{D}_{t-k}) \end{array} \right) \left(\begin{array}{cc} Var(\boldsymbol{\theta}_{t-k} | \mathbf{D}_{t-k}) & Cov(\boldsymbol{\theta}_{t-k}, \boldsymbol{\theta}_{t-k+1} | \mathbf{D}_{t-k}) \\ Cov(\boldsymbol{\theta}_{t-k+1}, \boldsymbol{\theta}_{t-k} | \mathbf{D}_{t-k}) & Var(\boldsymbol{\theta}_{t-k+1} | \mathbf{D}_{t-k}) \end{array} \right) \right]
\end{aligned}$$

con

$$E(\boldsymbol{\theta}_{t-k} \mid \mathbf{D}_{t-k}) = \mathbf{m}_{t-k} \quad ; \quad \text{Var}(\boldsymbol{\theta}_{t-k} \mid \mathbf{D}_{t-k}) = \mathbf{C}_{t-k},$$

$$E(\boldsymbol{\theta}_{t-k+1} \mid \mathbf{D}_{t-k}) = \mathbf{a}_{t-k}(1) = \mathbf{a}_{t-k+1} \quad ; \quad \text{Var}(\boldsymbol{\theta}_{t-k+1} \mid \mathbf{D}_{t-k}) = \mathbf{R}_{t-k}(1) = \mathbf{R}_{t-k+1},$$

$$\text{Cov}(\boldsymbol{\theta}_{t-k}, \boldsymbol{\theta}_{t-k+1} \mid \mathbf{D}_t) = [\text{Cov}(\boldsymbol{\theta}_{t-k+1}, \boldsymbol{\theta}_{t-k} \mid \mathbf{D}_t)]',$$

donde

$$\begin{aligned} \text{Cov}(\boldsymbol{\theta}_{t-k}, \boldsymbol{\theta}_{t-k+1} \mid \mathbf{D}_{t-k}) &= \text{Cov}((\boldsymbol{\theta}_{t-k}, \mathbf{G}_{t-k+1}\boldsymbol{\theta}_{t-k} + \boldsymbol{\omega}_{t-k+1}) \mid \mathbf{D}_{t-k}), \\ &= \text{Var}(\boldsymbol{\theta}_{t-k} \mid \mathbf{D}_{t-k})\mathbf{G}'_{t-k+1} + \text{Cov}(\boldsymbol{\theta}_{t-k}, \boldsymbol{\omega}_{t-k+1} \mid \mathbf{D}_{t-k}), \\ &= \mathbf{C}_{t-k}\mathbf{G}'_{t-k+1}, \end{aligned}$$

luego

$$(\boldsymbol{\theta}_{t-k} \mid \boldsymbol{\theta}_{t-k+1}, \mathbf{D}_{t-k}) \sim N[h_t(k), H_t(k)]$$

donde

$$h_t(k) = \mathbf{m}_{t-k} + \mathbf{C}_{t-k}\mathbf{G}'_{t-k+1}\mathbf{R}_{t-k+1}^{-1}(\boldsymbol{\theta}_{t-k+1} - \mathbf{a}_{t-k+1}),$$

y

$$H_t(k) = \mathbf{C}_{t-k} - \mathbf{C}_{t-k}\mathbf{G}'_{t-k+1}\mathbf{R}_{t-k+1}^{-1}\mathbf{G}_{t-k+1}\mathbf{C}_{t-k},$$

como $\mathbf{B}_t = \mathbf{C}_t\mathbf{G}'_{t+1}\mathbf{R}_{t+1}^{-1}$, se tiene que

$$h_t(k) = \mathbf{m}_{t-k} + \mathbf{B}_{t-k}(\boldsymbol{\theta}_{t-k+1} - \mathbf{a}_{t-k+1}),$$

y

$$H_t(k) = \mathbf{C}_{t-k} - \mathbf{B}_{t-k}\mathbf{R}_{t-k+1}\mathbf{B}'_{t-k}.$$

De igual forma; como $(\boldsymbol{\theta}_{t-k+1} \mid \mathbf{D}_t) \sim N[\mathbf{a}_t(-k+1), \mathbf{R}_t(-k+1)]$ y

$(\boldsymbol{\theta}_{t-k} \mid \boldsymbol{\theta}_{t-k+1}, \mathbf{D}_{t-k}) \sim N[h_t(k), H_t(k)]$, esto define una distribución conjunta Normal

para $(\boldsymbol{\theta}_{t-k}, \boldsymbol{\theta}_{t-k+1} | \mathbf{D}_t)$:

$$\begin{pmatrix} \boldsymbol{\theta}_{t-k} \\ \boldsymbol{\theta}_{t-k+1} \end{pmatrix} | \mathbf{D}_t \sim N \left[\begin{pmatrix} a_t(-k) \\ a_t(-k+1) \end{pmatrix} \begin{pmatrix} R_t(-k) & Cov(\boldsymbol{\theta}_{t-k}, \boldsymbol{\theta}_{t-k+1} | \mathbf{D}_t) \\ Cov(\boldsymbol{\theta}_{t-k+1}, \boldsymbol{\theta}_{t-k} | \mathbf{D}_t) & R_t(-k+1) \end{pmatrix} \right]$$

donde

$$\begin{aligned} Cov(\boldsymbol{\theta}_{t-k}, \boldsymbol{\theta}_{t-k+1} | \mathbf{D}_t) &= Cov(\boldsymbol{\theta}_{t-k}, \mathbf{G}'_{t-k+1} \boldsymbol{\theta}_{t-k} + \omega_{t-k+1} | \mathbf{D}_t). \\ &= Var(\boldsymbol{\theta}_{t-k} | \mathbf{D}_t) \mathbf{G}'_{t-k+1} + Cov(\boldsymbol{\theta}_{t-k}, \omega_{t-k+1} | \mathbf{D}_t), \\ &= R_t(-k) \mathbf{G}'_{t-k+1}, \end{aligned}$$

de aqui fácilmente se determina que:

$$a_t(-k) = E_{\boldsymbol{\theta}_{t-k+1} | \mathbf{D}_t} [E(\boldsymbol{\theta}_{t-k} | \boldsymbol{\theta}_{t-k+1} \mathbf{D}_t)],$$

desarrollando esta igualdad se tiene que:

$$\begin{aligned} a_t(-k) &= E_{\boldsymbol{\theta}_{t-k+1} | \mathbf{D}_t} [h_t(k) | \mathbf{D}_t], \\ &= E_{\boldsymbol{\theta}_{t-k+1} | \mathbf{D}_t} [m_{t-k} + B_{t-k} [\boldsymbol{\theta}_{t-k+1} - a_{t-k+1}] | \mathbf{D}_t], \\ &= m_{t-k} + B_{t-k} [E(\boldsymbol{\theta}_{t-k+1} | \mathbf{D}_t) - a_{t-k+1}], \\ &= m_{t-k} + B_{t-k} [a_t(-k+1) - a_{t-k+1}], \end{aligned}$$

de la misma forma se tiene que $R_t(-k) = Var(\boldsymbol{\theta}_{t-k} | \mathbf{D}_t)$ por propiedad de varianza se puede expresar como

$$R_t(-k) = E_{\boldsymbol{\theta}_{t-k+1} | \mathbf{D}_t} [Var(\boldsymbol{\theta}_{t-k} | \boldsymbol{\theta}_{t-k+1} \mathbf{D}_t)] + Var_{\boldsymbol{\theta}_{t-k+1} | \mathbf{D}_t} [E(\boldsymbol{\theta}_{t-k} | \boldsymbol{\theta}_{t-k+1} \mathbf{D}_t)],$$

desarrollando esta igualdad se tiene que:

$$\begin{aligned}
R_t(-k) &= E[H_t(k) \mid \mathbf{D}] + Var[h_t(k) \mid \mathbf{D}], \\
&= E[C_{t-k} - B_{t-k}R_{t-k+1}B'_{t-k}] + Var[m_{t-k} + B_{t-k}[\boldsymbol{\theta}_{t-k+1} - a_{t-k+1}] \mid \mathbf{D}_t], \\
&= C_{t-k} - B_{t-k}R_{t-k+1}B'_{t-k} + B_{t-k}Var(\boldsymbol{\theta}_{t-k+1} \mid \mathbf{D}_t)B'_{t-k}, \\
&= C_{t-k} - B_{t-k}R_{t-k+1}B'_{t-k} + B_{t-k}R_t(-k+1)B'_{t-k}, \\
&= C_{t-k} - B_{t-k}[R_{t-k+1} - R_t(-k+1)]B'_{t-k},
\end{aligned}$$

los cuales son los valores establecidos en el Teorema.

Para completar la demostración inductiva se tiene que considerar el caso $k = 1$ para esto $t-k+1 = t$, el cual implica que $a_t(-k+1) = a_t(0) = m_t$ y $R_t(-k+1) = R_t(0) = C_t$, con lo cual quedaría demostrado el Teorema.

Colorario 3.3.- Si $V_t = V = \phi^{-1}$ es desconocido y el análisis de la conjugada de la Sección 3.5 es aplicado, entonces

$$(\boldsymbol{\theta}_{t-k} \mid \mathbf{D}_t) \sim T_{n_t}[a_t(-k), (S_t/S_{t-k})R_t(-k)].$$

Note que, como con la ecuación secuencial actualizada, un cambio de escala esta implícita cuando las varianzas observacionales es desconocida, estimado por el análisis conjugado Normal/Gamma.

Corolario 3.4.- Las correspondientes distribuciones suavizadas para la función respuesta media de las series están dados por

$$(\mu_{t-k} \mid \mathbf{D}_t) \sim T_{n_t}[f_t(-k), (S_t/S_{t-k})F'_{t-k}R_t(-k)F_{t-k}],$$

donde, es una extensión de la notación para la función pronóstico con argumentos negativos,

$$f_t(-k) = F'_{t-k}a_t(-k).$$

3.7 Análisis de referencia para un MLD

Definición 3.5. En los modelos de las definiciones 3.2 y 3.4 (varianzas observacionales conocidas y desconocidas respectivamente), secuencialmente define las siguientes cantidades:

$$\begin{aligned} H_t &= W_t^{-1} - W_t^{-1}G_tP_t^{-1}G_t'W_t^{-1}, \\ P_t &= G_t'W_t^{-1}G_t + K_{t-1}, \\ h_t &= W_t^{-1}G_tP_t^{-1}k_{t-1}, \end{aligned}$$

y

$$K_t = \begin{cases} H_t + F_tF_t' & \text{si } V_t = V \text{ es desconocido,} \\ H_t + F_tF_t'/V_t & \text{si } V_t \text{ es conocido,} \end{cases}$$

con:

$$k_t = \begin{cases} h_t + F_tY_t & \text{si } V_t = V \text{ es desconocido,} \\ h_t + F_tY_t/V_t & \text{si } V_t \text{ es conocido,} \end{cases}$$

Teniendo valores iniciales $H_1 = 0$ y $h_1 = 0$.

en el caso de $V_t = V$, desconocido, W_t en la anterior definición es reemplazada por la matriz escalar-libre W_t^* . En adición, a este caso, se define:

$$\begin{aligned} \gamma_t &= \gamma_{t-1} + 1, \\ \lambda_t &= \delta_{t-1} - k_{t-1}'P_{t-1}^{-1}k_{t-1}, \\ \delta_t &= \lambda_t + Y_t^2, \end{aligned}$$

con valores iniciales $\lambda_1 = 0$ y $\gamma_0 = 0$.

Teorema 3.5.

(1).- Caso de varianza conocida

En el modelo de definición 3.2 se tiene que:

$$p(\boldsymbol{\theta}_1 | \mathbf{D}_0) \propto \text{constante},$$

Las distribuciones a priori y posteriori del vector estado en el tiempo t están dadas por:

$$\begin{aligned} p(\boldsymbol{\theta}_t | \mathbf{D}_{t-1}) &\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}'_t H_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t h_t)\right\}, \\ p(\boldsymbol{\theta}_t | \mathbf{D}_t) &\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}'_t K_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t k_t)\right\}, \end{aligned}$$

con los parámetros definidos en definición 3.4.

(2).- Caso de varianza desconocida

En el modelo de definición 3.4 con:

$$p(\boldsymbol{\theta}_1, V | \mathbf{D}_t) \propto V^{-1}$$

las distribuciones conjuntas a priori y posteriori del vector estado y la varianza de observación en el tiempo $t = 1, 2, \dots$ están dados por:

$$\begin{aligned} p(\boldsymbol{\theta}_t, V | \mathbf{D}_{t-1}) &\propto V^{-(1+\frac{\gamma_t-1}{2})} \exp\left\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}'_t H_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t h_t + \lambda_t)\right\}, \\ p(\boldsymbol{\theta}_t, V | \mathbf{D}_t) &\propto V^{-(1+\frac{\gamma_t}{2})} \exp\left\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}'_t K_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t k_t + \delta_t)\right\}. \end{aligned}$$

Demostración.

(1).-Caso de varianza conocida

La demostración es similar al caso de varianza desconocida por esta razón solamente se demostrará este caso.

(2).-Caso de varianza desconocida

La demostración es por inducción para el caso cuando $t = 1$ se tiene que:

$$\begin{aligned} p(\boldsymbol{\theta}_1, V | \mathbf{D}_0) &\propto V^{-(1+\frac{\gamma_0-1}{2})} \exp\left\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}'_0 H_0 \boldsymbol{\theta}_0 - 2\boldsymbol{\theta}'_0 h_0 + \lambda_0)\right\} \\ &\propto V^{-1}. \end{aligned}$$

Cuando $V_t = V$ es desconocido, y $t = t - 1$ se asume que:

$$p(\boldsymbol{\theta}_t, V | \mathbf{D}_{t-1}) \propto V^{-(1+\frac{t-1}{2})} \exp\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}'_t H_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t h_t + \lambda_t)\}.$$

Tambi3n, se tiene que la f.d.p de \mathbf{Y}_t es:

$$p(\mathbf{Y}_t | \boldsymbol{\theta}_t, V, \mathbf{D}_{t-1}) \propto V^{-\frac{1}{2}} \exp\{-\frac{1}{2}V^{-1}(\mathbf{Y}_t - F_t \boldsymbol{\theta}_t)^2\},$$

por el Teorema de Bayes se tiene que:

$$\begin{aligned} p(\boldsymbol{\theta}_t, V | \mathbf{D}_t) &= p(\boldsymbol{\theta}_t, V | \{\mathbf{Y}_t, \mathbf{D}_{t-1}\}), \\ &= \frac{p(\mathbf{Y}_t, \boldsymbol{\theta}_t, V, \mathbf{D}_{t-1})}{p(\mathbf{Y}_t, \mathbf{D}_{t-1})}, \\ &= p(\mathbf{Y}_t | \boldsymbol{\theta}_t, V, \mathbf{D}_{t-1}) \frac{p(\boldsymbol{\theta}_t, V, \mathbf{D}_{t-1})}{p(\mathbf{Y}_t, \mathbf{D}_{t-1})}, \\ &= p(\mathbf{Y}_t | \boldsymbol{\theta}_t, V, \mathbf{D}_{t-1}) p(\boldsymbol{\theta}_t, V | \mathbf{D}_{t-1}) \frac{p(\mathbf{D}_{t-1})}{p(\mathbf{Y}_t, \mathbf{D}_{t-1})}, \\ &\propto p(\mathbf{Y}_t | \boldsymbol{\theta}_t, V, \mathbf{D}_{t-1}) p(\boldsymbol{\theta}_t, V | \mathbf{D}_{t-1}), \\ &\propto V^{-1/2} \exp\{-\frac{1}{2}V^{-1}(\mathbf{Y}_t - F'_t \boldsymbol{\theta}_t)^2\} \\ &\quad \times V^{-(1+\frac{t-1}{2})} \exp\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}'_t H_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t h_t + \lambda_t)\}, \\ &\propto V^{-(1+\frac{t}{2})} \exp\{-\frac{1}{2}V^{-1}[(\mathbf{Y}_t - F'_t \boldsymbol{\theta}_t)^2 + (\boldsymbol{\theta}'_t H_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t h_t + \lambda_t)]\}. \end{aligned}$$

La expresi3n que est3 dentro del exponencial se puede expresar de la siguiente manera:

$$\begin{aligned} (\mathbf{Y}_t - F'_t \boldsymbol{\theta}_t)^2 + (\boldsymbol{\theta}'_t H_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t h_t + \lambda_t) &= \boldsymbol{\theta}'_t H_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t h_t + \lambda_t \\ &\quad + \boldsymbol{\theta}'_t F_t F'_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t F_t \mathbf{Y}_t + \mathbf{Y}_t^2, \\ &= \boldsymbol{\theta}'_t H_t \boldsymbol{\theta}_t + \boldsymbol{\theta}'_t F_t F'_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t h_t - 2\boldsymbol{\theta}'_t F_t \mathbf{Y}_t + \lambda_t + \mathbf{Y}_t^2, \\ &= \boldsymbol{\theta}'_t (H_t + F_t F'_t) \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t (h_t + F_t \mathbf{Y}_t) + \lambda_t + \mathbf{Y}_t^2, \\ &= \boldsymbol{\theta}'_t K_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t k_t + \delta_t. \end{aligned}$$

Luego entonces se puede expresar $p(\boldsymbol{\theta}_t, V | \mathbf{D}_t)$ como:

$$\begin{aligned} p(\boldsymbol{\theta}_t, V | \mathbf{D}_t) &\propto V^{-(1+\frac{\gamma_t}{2})} \exp\left\{-\frac{1}{2}V^{-1}[(\mathbf{Y}_t - \mathbf{F}'_t\boldsymbol{\theta}_t)^2 + (\boldsymbol{\theta}'_t\mathbf{H}_t\boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t\mathbf{h}_t + \lambda_t)]\right\}, \\ &\propto V^{-(1+\frac{\gamma_t}{2})} \exp\left\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}'_t\mathbf{K}_t\boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t\mathbf{k}_t + \delta_t)\right\}. \end{aligned}$$

Considere la a priori implicada para $t + 1$:

$$\begin{aligned} p(\boldsymbol{\theta}_{t+1}, V | \mathbf{D}_t) &= \int p(\boldsymbol{\theta}_{t+1}, V | \boldsymbol{\theta}_t, \mathbf{D}_t)p(\boldsymbol{\theta}_t | \mathbf{D}_t)d\boldsymbol{\theta}_t, \\ &= \int p(\boldsymbol{\theta}_{t+1} | \boldsymbol{\theta}_t, V, \mathbf{D}_t)p(V | \boldsymbol{\theta}_t, \mathbf{D}_t)p(\boldsymbol{\theta}_t | \mathbf{D}_t)d\boldsymbol{\theta}_t, \\ &= \int p(\boldsymbol{\theta}_{t+1} | \boldsymbol{\theta}_t, V, \mathbf{D}_t)p(\boldsymbol{\theta}_t, V | \mathbf{D}_t)d\boldsymbol{\theta}_t. \end{aligned}$$

Se puede afirmar que el primer término de la integral es la distribución Normal de $(\boldsymbol{\theta}_{t+1} | \boldsymbol{\theta}_t, V, \mathbf{D}_t) \sim N[G_{t+1}\boldsymbol{\theta}_t, VW_{t+1}^*]$. El segundo término de la integral tiene la forma demostrada inicialmente, esto es:

$$p(\boldsymbol{\theta}_t, V | \mathbf{D}_t) \propto V^{-(1+\frac{\gamma_t}{2})} \exp\left\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}'_t\mathbf{K}_t\boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t\mathbf{k}_t + \delta_t)\right\}.$$

Luego la f.d.p. $p(\boldsymbol{\theta}_{t+1}, V | \mathbf{D}_t)$ tiene la siguiente expresión:

$$\begin{aligned} p(\boldsymbol{\theta}_{t+1}, V | \mathbf{D}_t) &\propto \int V^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}_{t+1} - G_{t+1}\boldsymbol{\theta}_t)' \times W_{t+1}^{*-1}(\boldsymbol{\theta}_{t+1} - G_{t+1}\boldsymbol{\theta}_t)\right\} \\ &\quad \times V^{-(1+\frac{\gamma_t}{2})} \exp\left\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}'_t\mathbf{K}_t\boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t\mathbf{k}_t + \delta_t)\right\}d\boldsymbol{\theta}_t, \\ &\propto V^{-(1+\frac{\gamma_t+n}{2})} \int \exp\left\{-\frac{1}{2}V^{-1}[(\boldsymbol{\theta}_{t+1} - G_{t+1}\boldsymbol{\theta}_t)'W_{t+1}^{*-1}(\boldsymbol{\theta}_{t+1} - G_{t+1}\boldsymbol{\theta}_t) \right. \\ &\quad \left. + (\boldsymbol{\theta}'_t\mathbf{K}_t\boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t\mathbf{k}_t + \delta_t)]\right\}d\boldsymbol{\theta}_t. \end{aligned}$$

Se sabe que $\boldsymbol{\theta}'_t\mathbf{G}'_{t+1}W_{t+1}^{*-1}\boldsymbol{\theta}_{t+1}$ y $\boldsymbol{\theta}'_t\mathbf{k}_t$ son escalares y W_{t+1}^{*-1} , P_{t+1} son matrices de varianza-covarianza; es decir $(W_{t+1}^{*-1})' = W_{t+1}^{*-1}$ y $(P_{t+1})' = P_{t+1}$ luego entonces se tiene que:

$$\boldsymbol{\theta}'_t\mathbf{k}_t = (\boldsymbol{\theta}'_t\mathbf{k}_t)' = \mathbf{k}'_t\boldsymbol{\theta}_t,$$

y

$$\boldsymbol{\theta}'_t G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} = (\boldsymbol{\theta}'_t G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1})' = \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} G_{t+1} \boldsymbol{\theta}_t.$$

Entonces la expresión contenida dentro del exponencial se puede expresar como:

$$\begin{aligned} & (\boldsymbol{\theta}_{t+1} - G_{t+1} \boldsymbol{\theta}_t)' W_{t+1}^*{}^{-1} (\boldsymbol{\theta}_{t+1} - G_{t+1} \boldsymbol{\theta}_t) + (\boldsymbol{\theta}'_t K_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t k_t + \delta_t) \\ &= \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} G_{t+1} \boldsymbol{\theta}_t \\ &\quad - \boldsymbol{\theta}'_t G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \boldsymbol{\theta}'_t G'_{t+1} W_{t+1}^*{}^{-1} G_{t+1} \boldsymbol{\theta}_t + \boldsymbol{\theta}'_t K_t \boldsymbol{\theta}_t - [\boldsymbol{\theta}'_t k_t + k'_t \boldsymbol{\theta}] + \delta_t, \\ &= [\boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \delta_t] + [\boldsymbol{\theta}'_t K_t \boldsymbol{\theta}_t + \boldsymbol{\theta}'_t G'_{t+1} W_{t+1}^*{}^{-1} G_{t+1} \boldsymbol{\theta}_t] \\ &\quad - [\boldsymbol{\theta}'_t k_t + \boldsymbol{\theta}'_t G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1}] - [k'_t \boldsymbol{\theta}_t + \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} G_{t+1} \boldsymbol{\theta}_t], \\ &= \boldsymbol{\theta}'_t (K_t + G'_{t+1} W_{t+1}^*{}^{-1} G_{t+1}) \boldsymbol{\theta}_t - \boldsymbol{\theta}'_t (k_t + G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1}) \\ &\quad - (k'_t + \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} G_{t+1}) \boldsymbol{\theta}_t + (\boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \delta_t), \\ &= \boldsymbol{\theta}'_t P_{t+1} \boldsymbol{\theta}_t - \boldsymbol{\theta}'_t P_{t+1} P_{t+1}^{-1} (k_t + G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1}) \\ &\quad - (P_{t+1}^{-1} (k_t + G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1}))' (P_{t+1})' \boldsymbol{\theta}_t + (\boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \delta_t), \\ &= \boldsymbol{\theta}'_t P_{t+1} \boldsymbol{\theta}_t - \boldsymbol{\theta}'_t P_{t+1} \boldsymbol{\alpha}_{t+1} + \boldsymbol{\alpha}'_{t+1} P_{t+1} \boldsymbol{\theta}_t + \boldsymbol{\alpha}'_{t+1} P_{t+1} \boldsymbol{\alpha}_{t+1} \\ &\quad + \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \delta_t - \boldsymbol{\alpha}'_{t+1} P_{t+1} \boldsymbol{\alpha}_{t+1}, \\ &= (\boldsymbol{\theta}_t - \boldsymbol{\alpha}_{t+1})' P_{t+1} (\boldsymbol{\theta}_t - \boldsymbol{\alpha}_{t+1}) + R_{t+1}. \end{aligned}$$

Donde:

$$\begin{aligned} P_{t+1} &= K_t + G'_{t+1} W_{t+1}^*{}^{-1} G_{t+1}, \\ \boldsymbol{\alpha}_{t+1} &= P_{t+1}^{-1} (k_t + G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1}), \\ R_{t+1} &= \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \delta_t - \boldsymbol{\alpha}'_{t+1} P_{t+1} \boldsymbol{\alpha}_{t+1}, \end{aligned}$$

luego la $p(\boldsymbol{\theta}_{t+1}, V \mid D_t)$ será:

$$\begin{aligned} & p(\boldsymbol{\theta}_{t+1}, V \mid D_t) \\ & \propto V^{-(1+\frac{\tau_t+n}{2})} \int \exp\left\{-\frac{1}{2} V^{-1} [(\boldsymbol{\theta}_t - \boldsymbol{\alpha}_{t+1})' P_{t+1} (\boldsymbol{\theta}_t - \boldsymbol{\alpha}_{t+1}) + R_{t+1}]\right\} d\boldsymbol{\theta}_t. \end{aligned}$$

Y de la integral anterior se tiene que:

$$\begin{aligned}
p(\boldsymbol{\theta}_{t+1}, V \mid \mathbf{D}_t) &\propto V^{-(1+\frac{\gamma_t+n}{2})} \exp\{-\frac{1}{2}V^{-1}R_{t+1}\} \\
&\quad \times \int \exp\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}_t - \boldsymbol{\alpha}_t)'P_{t+1}(\boldsymbol{\theta}_t - \boldsymbol{\alpha}_t)\} d\boldsymbol{\theta}_t, \\
&\propto V^{-(1+\frac{\gamma_t+n}{2})} \exp\{-\frac{1}{2}V^{-1}R_{t+1}\} \times V^{\frac{n}{2}}, \\
&\propto V^{-(1+\frac{\gamma_t}{2})} \exp\{-\frac{1}{2}V^{-1}R_{t+1}\}.
\end{aligned}$$

Por otro lado, usando el hecho que $\boldsymbol{\theta}'_{t+1}(W_{t+1}^*)^{-1}'G_{t+1}(P_{t+1}^{-1})'k_t$ es un escalar, se tiene que:

$$\begin{aligned}
\boldsymbol{\theta}'_{t+1}(W_{t+1}^*)^{-1}'G_{t+1}(P_{t+1}^{-1})'k_t &= (\boldsymbol{\theta}'_{t+1}(W_{t+1}^*)^{-1}'G_{t+1}(P_{t+1}^{-1})'k_t)', \\
&= k_t'(P_{t+1}^{-1})'G_{t+1}W_{t+1}^*{}^{-1}\boldsymbol{\theta}_{t+1}.
\end{aligned}$$

Adicionalmente se tiene que $W_{t+1}^*{}^{-1}$ y P_{t+1} son matrices de covarianza por lo tanto, son matrices simétricas. Además por definición:

$$\lambda_{t+1} = \delta_t - k_t'P_{t+1}^{-1}k_t$$

Entonces, expandiendo R_{t+1} se puede expresar como:

$$\begin{aligned}
R_{t+1} &= \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \delta_t - \alpha'_{t+1} P_{t+1} \alpha_{t+1}, \\
&= \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \delta_t - \alpha'_{t+1} P_{t+1} P_{t+1}^{-1} (k_t + G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1}), \\
&= \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \delta_t - \alpha'_{t+1} (k_t + G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1}), \\
&= \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \delta_t - [k'_t + \boldsymbol{\theta}'_{t+1} (W_{t+1}^*{}^{-1})' G_{t+1}] [(P_{t+1}^{-1})' (k_t + G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1})], \\
&= \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \delta_t - [k'_t (P_{t+1}^{-1})' k_t + \boldsymbol{\theta}'_{t+1} (W_{t+1}^*{}^{-1})' G_{t+1} (P_{t+1}^{-1})' k_t \\
&\quad + k'_t (P_{t+1}^{-1})' G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \boldsymbol{\theta}'_{t+1} (W_{t+1}^*{}^{-1})' G_{t+1} (P_{t+1}^{-1})' G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1}], \\
&= \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} + \delta_t - [k'_t (P_{t+1}^{-1})' k_t + \boldsymbol{\theta}'_{t+1} (W_{t+1}^*{}^{-1})' G_{t+1} (P_{t+1}^{-1})' k_t \\
&\quad + \boldsymbol{\theta}'_{t+1} (W_{t+1}^*{}^{-1})' G_{t+1} (P_{t+1}^{-1})' k_t + \boldsymbol{\theta}'_{t+1} (W_{t+1}^*{}^{-1})' G_{t+1} (P_{t+1}^{-1})' G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1}], \\
&= \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} G_{t+1} P_{t+1}^{-1} G'_{t+1} W_{t+1}^*{}^{-1} \boldsymbol{\theta}_{t+1} \\
&\quad - \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} G_{t+1} P_{t+1}^{-1} k_t - \boldsymbol{\theta}'_{t+1} W_{t+1}^*{}^{-1} G_{t+1} P_{t+1}^{-1} k_t + [\delta_t - k'_t (P_{t+1}^{-1})' k_t].
\end{aligned}$$

Se tiene que reordenar esta igualdad para luego poder expresarlo:

$$\begin{aligned}
R_{t+1} &= \boldsymbol{\theta}'_{t+1} [W_{t+1}^*{}^{-1} - W_{t+1}^*{}^{-1} G_{t+1} P_{t+1}^{-1} G'_{t+1} W_{t+1}^*{}^{-1}] \boldsymbol{\theta}_{t+1} \\
&\quad + 2\boldsymbol{\theta}'_{t+1} [W_{t+1}^*{}^{-1} G_{t+1} P_{t+1}^{-1} k_t] + [\delta_t - k'_t P_{t+1}^{-1} k_t], \\
&= \boldsymbol{\theta}'_{t+1} H_{t+1} \boldsymbol{\theta}_{t+1} - 2\boldsymbol{\theta}'_{t+1} h_{t+1} + \lambda_{t+1}.
\end{aligned}$$

Por lo tanto:

$$p(\boldsymbol{\theta}_{t+1}, V \mid D_t) \propto V^{-(1+\frac{n}{2})} \exp\left\{-\frac{1}{2} V^{-1} (\boldsymbol{\theta}'_{t+1} H_{t+1} \boldsymbol{\theta}_{t+1} - 2\boldsymbol{\theta}'_{t+1} h_{t+1} + \lambda_{t+1})\right\}.$$

Se recalca que el modelo tiene n parámetros en el vector de estados en cada tiempo t y así, partiendo desde una a priori referencial, las a posteriori serán impropias hasta por lo menos un tiempo n , en el caso de varianza conocida, y tiempo $n + 1$ en el caso de varianza desconocida. Después que observaciones suficientes han sido procesadas las distribuciones a posteriori impropias que caracteriza el modelo llegan a ser propias Aunque las anteriores recursiones permanecen válidas es más usual revertir a las formas

padrones involucrando directamente actualizaciones de la media y varianza-covarianza a posteriori del vector de estados una vez que estos no requieran matrices inversas. El número de observaciones requeridas para alcanzar distribuciones propias depende sobre la forma del modelo y los datos; como se dijo anteriormente el menor número el cual es suficiente es el número de parámetros desconocidos en el modelo, incluyendo 1 para la varianza observacional si fuese desconocido. Esto requiere que no halla datos perdidos en estas primeras observaciones y por lo contrario que no exista problema de colinearidad si el modelo incluye regresión . En la práctica, más que el mínimo raramente es requerido. En general se define un n tal que:

$$[n] = \min\{t : \text{distribuciones a posteriori son propias}\}$$

La relación necesaria entre las cantidades que define la distribución a posteriori representada en el Teorema (3.5) y aquellas en la representación original son fácilmente obtenidos del modo siguiente:

Corolario 3.5.-

(1) Caso de varianza conocida

para $t \geq [n]$, la distribución a posteriori de $(\theta_t | D_t)$ es como en la sección 3.3 Teorema 3.1, con:

$$C_t = K_t^{-1} \quad \text{y} \quad m_t = K_t^{-1} k_t.$$

Demostración.

Por el Teorema 3.1 se tiene que $(\theta_t | D_t) \sim N[m_t, C_t]$ luego se puede expresar:

$$\begin{aligned} p(\theta_t | D_t) &\propto \exp\left\{-\frac{1}{2}(\theta - m_t)'C_t^{-1}(\theta - m_t)\right\}, \\ &\propto \exp\left\{-\frac{1}{2}[\theta_t' C_t^{-1} \theta_t - \theta_t' C_t^{-1} m_t - m_t' C_t^{-1} \theta_t + m_t' C_t^{-1} m_t]\right\}, \\ &\propto \exp\left\{-\frac{1}{2}[\theta_t' C_t^{-1} \theta_t - 2\theta_t' C_t^{-1} m_t]\right\}. \end{aligned}$$

Del Teorema 3.5 caso(1) se tiene que:

$$p(\boldsymbol{\theta}_t | \mathbf{D}_t) \propto \exp\left\{-\frac{1}{2}[\boldsymbol{\theta}'_t K_t^{-1} \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t k_t]\right\}.$$

Igualando términos entre las dos últimas ecuaciones se tiene que $K_t = C_t^{-1}$ y $k_t = C_t^{-1} m_t$ el cual se puede expresar como $m_t = K_t^{-1} k_t$ con lo que quedaría demostrado.

(2) Caso de varianza desconocida

para $t \geq [n]$, la distribución a posteriori de $(\boldsymbol{\theta}_t | \mathbf{D}_t)$ es como en la sección 3.5, con:

$$C_t = S_t K_t^{-1} \quad \text{y} \quad m_t = K_t^{-1} k_t,$$

donde $S_t = d_t/n_t$, con $n_t = \gamma_t - n$ y $d_t = \delta_t - k'_t m_t$. En el caso usual que $[n] = n + 1$, entonces $n_{n+1} = 1$ y es fácilmente demostrable que $d_{n+1} = S_{n+1} = e_{n+1}^2/Q_{n+1}^*$.

Del Teorema 3.3 parte(a) se tiene que: $(\boldsymbol{\theta}_t | \mathbf{D}_t, V) \sim N[m_t, V C_t^*]$ y

del Teorema 3.3 parte(b) se tiene que: $(\phi | \mathbf{D}_t) \sim G[n_t/2, d_t/2]$, donde $\phi = V^{-1}$, $n_t = n_{t-1} + 1$, $d_{t-1} = d_{t-1} + e_t^2/Q_t^*$, $C_t = S_t C_t^*$ y $S_t = d_t/n_t$. También,

$$p(\boldsymbol{\theta}_t, V | \mathbf{D}_t) = p(\boldsymbol{\theta}_t | \mathbf{D}_t, V) p(V | \mathbf{D}_t)$$

observe que

$$p(\phi | \mathbf{D}_t) \propto \phi^{\frac{n_t}{2}-1} \exp\{-\phi d_t/2\},$$

además se sabe que $\phi = V^{-1}$ entonces $V = \phi^{-1}$ y con esto se puede hallar la f.d.p de V , con jacobiano dado por $|J| = |d\phi/dv| = V^{-2}$, así

$$p(V | \mathbf{D}_t) \propto V^{-\frac{n_t}{2}+1} \exp\left\{-\frac{1}{2}V^{-1}d_t\right\} V^{-2},$$

también

$$p(\boldsymbol{\theta}_t, V | \mathbf{D}_t) \propto V^{-\frac{n_t}{2}} \exp\left\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}_t - m_t)' C_t^{*-1} (\boldsymbol{\theta}_t - m_t)\right\} V^{-\frac{n_t}{2}+1} \exp\left\{-\frac{1}{2}V^{-1}d_t\right\} V^{-2}.$$

Reordenando esta expresión se tiene que:

$$p(\boldsymbol{\theta}_t, V | \mathbf{D}_t) \propto V^{-(1+\frac{n_t+n_t}{2})} \exp\left\{-\frac{1}{2}V^{-1}[\boldsymbol{\theta}'_t C_t^{*-1} \boldsymbol{\theta}_t - 2\boldsymbol{\theta}'_t C_t^{*-1} m_t + m'_t C_t^{*-1} m_t + d_t]\right\}$$

como $C_t = \frac{d_t}{n_t} C_t^*$ entonces $C_t^{*-1} = \frac{d_t}{n_t} C_t^{-1}$, y reemplazando en la expresión anterior:

$$p(\boldsymbol{\theta}_t, V | \mathbf{D}_t) \propto V^{-(1+\frac{n+n_t}{2})} \exp\left\{-\frac{1}{2}V^{-1}\left[\boldsymbol{\theta}_t' \frac{d_t}{n_t} C_t^{-1} \boldsymbol{\theta}_t - 2\boldsymbol{\theta}_t' \frac{d_t}{n_t} C_t^{-1} \mathbf{m}_t + d_t + \mathbf{m}_t' \frac{d_t}{n_t} C_t^{-1} \mathbf{m}_t\right]\right\}.$$

Del Teorema 3.5 parte(2) se tiene que:

$$p(\boldsymbol{\theta}_t, V | \mathbf{D}_t) \propto V^{-(1+\frac{\gamma_t}{2})} \exp\left\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}_t' K_t \boldsymbol{\theta}_t - 2\boldsymbol{\theta}_t' k_t + \delta_t)\right\}.$$

Comparando éstas dos últimas ecuaciones fácilmente se verifica que $\gamma_t = n_t + n$, $K_t = \frac{d_t}{n_t} C_t^{-1}$, $k_t = \frac{d_t}{n_t} C_t^{-1} \mathbf{m}_t$ y $\delta_t = d_t + \frac{d_t}{n_t} \mathbf{m}_t' (C_t^{-1})' \mathbf{m}_t$, además esta última expresión se puede expresar como $\delta_t = d_t + \frac{d_t}{n_t} (C_t^{-1} \mathbf{m}_t)' \mathbf{m}_t$.

Luego, lo anterior se puede expresar como $C_t = S_t K_t^{-1}$ y $\mathbf{m}_t = K_t^{-1} k_t$

donde $S_t = \frac{d_t}{n_t}$, $n_t = \gamma_t - n$ y $d_t = \delta_t - k_t' \mathbf{m}_t$. En el usual caso que $[n] = n + 1$, entonces $n_{n+1} = 1$ y fácilmente se demuestra que $d_{n+1} = S_{n+1} = e_{n+1}^2 / Q_{n+1}^*$

3.7.1 Caso especial de $W_t = 0$

Consideremos ahora el caso de modelos con ecuaciones de evolución determinísticas, en los cuales $W_t = W_t^* = 0$. Este caso especial es discutido primeramente por razones prácticas. Las motivaciones derivan de la necesidad para especificar W_t en las ecuaciones de recurrencias detalladas anteriormente. Los métodos mostrados anteriormente no se puede aplicar en el análisis referencial para $t < [n]$ porque las covarianzas posteriores pueden no existir. Una aproximación alternativa es requerida, y la práctica usual de covarianza inicial, $W_t = 0$ para $t = 1, 2, \dots, [n]$ es recomendada.

En el análisis referencial de un modelo con $n+1$ parámetros (incluyendo V), se necesita $[n]$ (al menos $n+1$) observaciones para obtener una distribución conjunta a posteriori completamente especificada: una observación para cada parámetro. Más generalmente, en el tiempo $t = [n]$ se tiene una observación para cada parámetro. De este modo es imposible detectar o estimar cualquier cambio en los parámetros durante las primeras

(n+1) observaciones sobre la cual el análisis referencial es ejecutado. Consecuentemente, el uso de matrices W_t diferente de cero es irrelevante puesto que ello básicamente permitiría cambios que no puede ser estimado, y no se perdería nada por direccionarlo a cero para $t = 1, 2, \dots, [n]$. En el tiempo $t = [n]$. Las a posteriores estan completamente especificadas y futuros cambios paramétricos pueden ser identificadas. En este modo, se revierte a un modelo Dinámico completo de matrices de covarianza de evolución diferente de cero.

Teorema 3.6.

En el Teorema 3.5, se supone que G_t es no-singular y $W_t = W_t^* = \mathbf{0}$. Entonces las distribuciones a priori y a posteriori de θ_t y V tienen las formas del Teorema 3.5 con recurrencias definidas como sigue.

(1).- Caso de varianza conocida

$$H_t = G_t'^{-1} K_{t-1} G_t^{-1}$$

$$h_t = G_t'^{-1} k_{t-1}.$$

(2).- Caso de varianza desconocida

$$H_t = G_t'^{-1} K_{t-1} G_t^{-1}$$

$$h_t = G_t'^{-1} k_{t-1}$$

$$\lambda_t = \delta_{t-1}.$$

Detalles de esta demostración puede ser vista en West y Harrison [21].

3.7.2 Filtración

Filtración en el caso de una a priori referencial usa exactamente el mismo resultado como en el caso de las prioris conjugadas para tiempos $t \geq [n]$, desde que todas las distribuciones en este rango sean propias. En particular, la usuales recurrencias en las filtraciones de la sección 3.6 son válidas. Sin embargo, para $t < [n]$ estas recurrencias no se pueden aplicar una vez que las medias y varianzas-covarianzas a posteriori requeridas no existen. El siguiente Teorema proporciona la solución para este caso.

Teorema 3.7.

En el Teorema 3.5, las distribuciones de filtración en el MLD para tiempos $t - r$, $r = 0, 1, 2, \dots, [n] - 1$ están definidas como sigue.

(1).- Caso de varianza conocida

$$p(\boldsymbol{\theta}_{t-r} | \mathbf{D}_t) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}'_{t-r} K_t(-r) \boldsymbol{\theta}_{t-r} - 2\boldsymbol{\theta}'_{t-r} k_t(-r))\right\},$$

(2).- Caso de varianza desconocida

$$p(\boldsymbol{\theta}_{t-r}, V | \mathbf{D}_t) \propto V^{-(1+\frac{n}{2})} \exp\left\{-\frac{1}{2}V^{-1}(\boldsymbol{\theta}'_{t-r} K_t(-r) \boldsymbol{\theta}_{t-r} - 2\boldsymbol{\theta}'_{t-r} k_t(-r))\right\},$$

donde las cantidades definidas son calculadas recursivamente de acuerdo a:

$$\begin{aligned}
K_t(-r) &= G'_{t-r+1} W_{t-r+1}^{-1} G_{t-r+1} + K_{t-r} \\
&\quad - G'_{t-r+1} W_{t-r+1}^{-1} P_t^{-1}(-r+1) W_{t-r+1}^{-1} G_{t-r+1} \\
P_t(-r+1) &= W_{t-r+1}^{-1} + K_t(-r+1) - H_{t-r+1} \\
k_t(-r) &= k_{t-r} + G'_{t-r+1} W_{t-r+1}^{-1} P_t^{-1}(-r+1) \\
&\quad \times [k_t(-r+1) - h_{t-r+1}] \\
\delta_t(-r) &= \delta_t(-r+1) - \lambda_{t-r+1} \\
&\quad - [k_t(-r+1) - h_{t-r+1}]' P_t^{-1}(-r+1) \\
&\quad \times [k_t(-r+1) - h_{t-r+1}]
\end{aligned}$$

y H_t , h_t , K_t , k_t , λ_t , δ_t son como las definidas en 3.5 (Otra vez note que, en el caso de V desconocido, VW_t^* reemplaza W_t , para todo t). Inicializando valores para estas recurrencias con $K_t(0) = K_t$, $k_t(0) = k_t$, $\delta_t(0) = \delta_t$.

Colorario 3.6.- Para $t > [n]$, las distribuciones definidas en el teorema 3.7 son propias, como las dadas en la sección 3.6, con

(1).- Caso de varianza conocida

$$(\theta_{t-r} \mid \mathbf{D}_t) \sim N[a_t(-r), R_t(-r)]$$

donde $a_t(-r) = K_t(-r)^{-1} k_t(-r)$ y $R_t(-r) = K_t(-r)^{-1}$.

(2).- Caso de varianza desconocida

$$(\theta_{t-r} \mid \mathbf{D}_t) \sim T_{\gamma_t - [n]}[a_t(-r), R_t(-r)]$$

donde $a_t(-r) = K_t(-r)^{-1} k_t(-r)$ y $R_t(-r) = S_t K_t(-r)^{-1}$.

Teorema 3.8.

En el caso de varianza de evolución cero como en el Teorema 3.6, los resultados del Teorema 3.7 son válidas pero con los siguientes cambios para las recurrencias.

$$K_t(-r) = G'_{t-r+1} K_{-r+1} G_{t-r+1}$$

$$k_t(-r) = G'_{t-r+1} k_t(-r+1)$$

$$\delta_t(-r) = \delta_t(-r+1)$$

3.8 Factores de descuento y especificación del modelo componente

La matriz de sistema es bloque diagonal con sub-matrices individuales suministrando contribuciones de simples modelos componentes. El vector de regresión está particionado dentro del encadenamiento del correspondiente sub-vector. Para completar la especificación del modelo, se necesita tres componentes más. Estos son la secuencia del estado de evolución de las matrices de varianza-covarianza W_t , ($t = 1, \dots$); la secuencia de la varianza observacional V_t , ($t = 1, \dots$); y los parámetros determinando las a priori iniciales para el vector estado, y/o varianza para el error observacional, en $t = 0$ dado D_0 . En este trabajo solamente nos ocuparemos de analizar el caso de W_t . los detalles sobre las especificaciones de V_t y otras consideraciones puede ser encontrada en West y Harrison [21]. La estructura del componente del MLD conduce a una a priori inicial siendo especificado en términos de una colección de a priori, una por cada sub-vector de θ_0 correspondiente a los modelos de componente individual, con independencia entre los componentes. En esta sección se considera, la especificación de la secuencia de evolución de las matrices W_t , usando el modelo general de la definición 3.2, con pequeños cambios puramente técnicos para el caso de varianzas observacionales desconocidas como en la definición 3.4.

La especificación de estructuras convenientes y magnitudes de W_t es importante para lograr un buen modelamiento y pronósticos adecuados, los valores controlan, la extensión de variación estocástica en la evolución del modelo y en consecuencia determina la estabilidad en el tiempo. En la ecuación del sistema W_t conduce a un crecimiento en incertidumbre o, equivalentemente, a pérdida de información acerca del vector estados entre los tiempos $t - 1$ y t . Más precisamente considera la información secuencial de las informaciones de actualización resumidas en la sección 3.5. En $t - 1$ la a posteriori para el vector estado actual tiene matriz varianza-covarianza $Var[\theta_{t-1} | D_{t-1}] = C_{t-1}$ el cual conduce, a través de la ecuación de evolución, para la matriz de varianza-covarianza de θ_t dado por $Var[\theta_t | D_{t-1}] = G_t C_{t-1} G_t' + W_t$. Sea P_t dado por:

$$P_t = G_t C_{t-1} G_t' = Var[G_t \theta_{t-1} | D_{t-1}].$$

P_t debería ser visto como la matriz varianza-covarianza a priori apropiado en el modelo $\{F_t, G_t, V_t, 0\}$, que es el modelo estándar sin ningún error de evolución al tiempo t . Esto debería por tanto la matriz de varianza-covarianza a priori requerido, de un vector de estados estable ideal con ningun cambio estocástico, adicionando el error de evolución ω_t a $G_t \theta_{t-1}$ dá el vector estado verdadero θ_t , que tiene el efecto de incrementar la incertidumbre desde P_t hacia el $R_t = P_t + W_t$ actual.

El término de error de evolución dá a estos modelos un incremento aditivo en incertidumbre o pérdida de información acerca del vector estado entre las observaciones, para alcanzar un grado de información conveniente decayendo sobre tiempo, es claro que las magnitudes relativas de W_t y P_t son importantes. Esto conduce a pensar en términos de una tasa de decaimiento de información natural que sugiere un incremento multiplicativo en vez de aditivo. Para el caso que el vector de estado es univariado, con: $P_t = C_{t-1}$, $W_t = W_t$, $R_t = C_{t-1} + W_t$ y $C_t = C_t$. En el caso de un factor de descuento δ , ($0 < \delta \leq 1$) enlaza W_t a $P_t = C_{t-1}$ via $W_t = P_t(1 - \delta)/\delta$ tal que $R_t = P_t/\delta$. Esto implica en este caso un incremento en varianza o pérdida de información del

$100(1 - \delta)/\delta\%$. Hay varias formas en las cuales esta idea puede ser generalizada para el caso multiparámetros, para ayudar en la elección de W_t . Aquí se discutirá la aproximación práctica más importante basada sobre el descuento por componente.

Una forma parsimoniosa para la secuencia de la varianza de evolución es obtenida directamente extendiendo el descuento escalar para el caso de la matriz, definiendo:

$$R_t = P_t/\delta,$$

para algun factor de descuento escalar δ , como es usual. Esto implica que:

$$R_t = P_t + W_t,$$

donde:

$$W_t = P_t(1 - \delta)/\delta.$$

De este modo W_t tiene formalmente la misma estructura interna, en términos de correlación, como P_t . La magnitud de las varianzas y covarianzas es controlado por el factor de descuento en el caso escalar. La implicación es que la información se reduce en la misma tasa para cada uno de los elementos del vector estado; esto es apropiado, en particular cuando la cuantificación de la función pronóstico entera del modelo es vista como sujeto a cambio en una tasa constante sin referencia a los componentes de esa función. Frecuentemente ésta es una suposición deseable en la práctica. Notar sin embargo, que cualquier de tales aproximaciones de descuento no es apropiado en casos cuando alguno o todos los parámetros de estados son conocidos al tiempo t . En tales circunstancias las varianzas y covarianzas correspondientes de P_t serían cero tal que la construcción del descuento no se pueden aplicar.

3.8.1 Descuento por componente

Adaptando a un MLD general la superposición de varios componentes, la idea se extiende naturalmente sugiriendo un factor de descuento para cada componente. Se

considera un **MLD** conteniendo la superposición de $h \geq 1$ sub-modelos

$$M_i : \quad \{F_i, G_i, V_i, W_i\},$$

con el vector estado θ_{it} , y error de evolución ω_{it} , de dimensión n_i , para $i = 1, \dots, h$. El **MLD**

$$\{F, G, V, W\}_t$$

con el vector estado, de dimensión $n = n_1 + \dots + n_h$, dado por:

$$\theta'_t = (\theta_{1t}, \dots, \theta_{ht}),$$

donde:

$$F'_t = (F'_{1t}, \dots, F'_{ht}),$$

$$\begin{aligned} G_t &= \text{block diag}[G_{1t}, \dots, G_{ht}] \\ &= \begin{pmatrix} G_{1t} & 0 & 0 & \dots & 0 \\ 0 & G_{2t} & 0 & \dots & 0 \\ 0 & 0 & G_{3t} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & G_{ht} \end{pmatrix}, \end{aligned}$$

y

$$\begin{aligned} W_t &= \text{block diag}[W_{1t}, \dots, W_{ht}] \\ &= \begin{pmatrix} W_{1t} & 0 & 0 & \dots & 0 \\ 0 & W_{2t} & 0 & \dots & 0 \\ 0 & 0 & W_{3t} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & W_{ht} \end{pmatrix}. \end{aligned}$$

Al tiempo t , la matriz de varianza:

$$P_t = \text{Var}[\boldsymbol{\theta}_t \mid \mathbf{D}_{t-1}] = G_t C_{t-1} G_t'$$

representa incertidumbre acerca de $G_t \boldsymbol{\theta}_{t-1}$ antes de la adición de la evolución del término ruido, denota los componentes del bloque diagonal correspondiente a P_{it} llamado:

$$P_{it} = \text{Var}[G_{it} \boldsymbol{\theta}_{it-1} \mid \mathbf{D}_{t-1}] \quad (i = 1, \dots, h).$$

Note que P_t no será un bloque diagonal, pero estos componentes proporcionan fuentes de información acerca de los componentes correspondientes de $G_t \boldsymbol{\theta}_{t-1}$. Ahora, adicionando la evolución ruido de ω_t con bloque diagonal de matriz de varianza W_t anterior, la matriz de varianza a priori R_t para $\boldsymbol{\theta}_t$ tiene componentes fuera de la diagonal, los mismos de P_t , pero bloques diagonales

$$R_{it} = P_{it} + W_{it}$$

para cada i , El concepto descuento se aplica naturalmente al modelo estructurado mediante la siguiente manera.

Definición 3.6.

Sea $\delta_1, \dots, \delta_h$ son cualquier factor de descuento, ($0 < \delta_i \leq 1; i = 1, \dots, h$), con δ_i siendo el factor de descuento asociado con el modelo componente M_i . Se supone que los componentes de evolución de las matrices de varianza W_{it} están definidas como en la sección anterior 3.8, a través de:

$$W_{it} = P_{it}(1 - \delta_i)/\delta_i, \quad (i = 1, \dots, h).$$

Entonces el modelo es llamado como **MLD** descuento por componente.

El efecto del descuento por componente es para modelar la disminución de la información en el tiempo, en una posible tasa diferente para cada modelo componente. El

modelador escoge el factor de descuento, algunos de los cuales serán iguales, para reflejar su creencia acerca de la estabilidad en el tiempo de los componentes individuales. Note que, desde un punto de vista operacional en actualizar, la evolución desde P_t hacia R_t no necesita hacer referencia a la construcción de la secuencia de W_t . Esto es simplemente logrado tomando los componentes de la covarianza como invariables y dividiendo los elementos del bloque diagonal por el factor de descuento apropiado:

$$R_{it} = P_{it}/\delta_i,$$

para cada i . El Bloque de descuento es la aproximación recomendada para estructurar la secuencia de la evolución de la varianza en casi todas las aplicaciones. La aproximación es parsimoniosa, aplicando un único factor de descuento para cada componente de un modelo más largo. Algunas veces un único factor de descuento aplicando a todo un modelo visto como un simple componente sería adecuado, pero la flexibilidad nos permite modelar hasta n componentes separados, cada uno con descuentos individuales. También, es importante la matriz W_t derivada es naturalmente escalada, los factores de descuento comienza siendo cantidades de menor dimensión sobre una escala estandarizada. Con o sin varianza, la construcción de descuentos aplica directamente.

3.8.2 Práctica de descuento estratégico

El descuento deberá ser observado como una técnica para ser usada en formular la estructura de evolución la varianza un paso al frente. Dando C_{t-1} desde P_t , hacia una noción de la relativa durabilidad de componentes cuantificado en términos de los factores de descuento, W_t está identificado como una matriz de varianza-covarianza apropiada para la evolución del error en el tiempo t . Por ejemplo, con un único modelo componente teniendo factor de descuento δ , aplicaciones repetidas nos llevará al uso de δ^k como un factor de descuento k -pasos al frente. La información disminuye

exponencialmente en el futuro es obviamente inconsistente con el MLD en el cual la información disminuye aritméticamente a través de los errores de evolución futuros de las matrices de varianza-covarianza. En consecuencia, el descuento aproximado debería ser obtenido mediante extrapolación (y también, cuando se busquen los valores perdidos en las series de tiempo). discutir este punto, y usar descuentos un paso al frente desde $t = 0$ para determinar la secuencia W_t implicada para todo el tiempo futuro t . Esto es posible desde que, dado los otros modelos componentes, estas matrices son simplemente funciones de cantidades asumidas inicialmente conocidas. Se puede ver que esto es también posible en modelos donde la varianza observacional esta siendo estimada.

El resultado práctico del descuento estratégico es como sigue:

(1) Considere la posición a posteriori para D_t , con la varianza a posteriori usual C_t , y la varianza a priori del paso siguiente: $R_{t+1} = P_{t+1} + W_{t+1}$.

(2) Pronosticando k -pasos al frente, sería usualmente adecuado adoptar una matriz condicional varianza constante, tomando:

$$Var[\omega_{t+k} | D_t] = W_t(k) = W_{t+1}, \quad (k = 1, \dots).$$

De este modo distribuciones de pronóstico serían basados en la adición de la evolución de errores con la misma matriz de varianza W_{t+1} para todo k .

(3) Una vez que Y_{t+1} es observado, la a posteriori para θ_{t+1} es calculado, también como es P_{t+2} y W_{t+2} es deducido usando descuento. De este modo pronosticando desde el tiempo $t + 1$, hacia adelante se tiene:

$$Var[\omega_{t+k+1} | D_{t+1}] = W_{t+1}(k) = W_{t+2}, \quad (k = 1, \dots).$$

(4) Proceder de esta forma al tiempo $t + 2$.

La simplicidad computacional de esta estrategia es evidente; en cualquier tiempo, una única evolución de la matriz de varianza es calculada y usada k -pasos al frente para

cualquier k . Notar que una modificación importante de los análisis estandar **MLD** está implicado aquí. Hasta ahora, la evolución de errores fueron asumidos por tener matrices de varianzas conocidas para todo tiempo, y también independencia del pasado de las series. Con la estrategia de descuento esta suposición ha sido modificada para permitir a las matrices de varianza en el futuro depender del estado corriente de la información. Matemáticamente, la suposición que, para cualquier $k = 1, \dots$, $Var[\omega_{t+k} | \mathbf{D}_t] = Var[\omega_{t+k} | \mathbf{D}_0]$ ha sido revisado; esto es ahora el caso que $Var[\omega_{t+k} | \mathbf{D}_t] = W_t(k)$ depende de t en adición a $t + k$. Por ejemplo, al tiempo t , el segundo paso al frente de la matriz de varianza-covarianza es:

$$Var[\omega_{t+2} | \mathbf{D}_t] = W_t(2) = W_{t+1}.$$

Obteniendo una nueva observación, ésta es revisada

$$Var[\omega_{t+2} | \mathbf{D}_{t+1}] = W_{t+1}(1) = W_{t+2}.$$

Está modificación no tiene consecuencia complicada en la práctica.

Capítulo 4

Consumo Privado en el Perú

4.1 Introducción

En este capítulo se analizará el consumo privado en los últimos diez años, considerando que el principal componente de la demanda es el gasto del consumidor por el sector de las economías domésticas. El consumo privado comprende desde los alimentos hasta los gastos administrativos; Comprende también gasto de consumo de bienes duraderos, como automóviles, gastos que podría concebirse más como inversión que como consumo privado. En el Perú se basa básicamente en las exportaciones de bienes y servicios. La explicación y el pronóstico del consumo privado será hecha a través de un modelo particular propuesto en el Capítulo 3, usando Análisis Referencial. El software que se ha empleado es el BATS ¹.

¹BATS significa Bayesian Analysis of Time Series y es de la autoría de Pole, West y Harrison. Esta disponible libremente vía ftp en el site [ftp.stat.duke.edu/Bats/](ftp://ftp.stat.duke.edu/Bats/)

4.2 Importancia del Consumo Privado en el Perú

Durante los últimos años, el consumo privado en el Perú ha mostrado una tasa de crecimiento insuficiente que se ha reflejado en una menor actividad económica y empresarial. Por su importancia en la economía representa más del 70% del PBI su evolución marcará no solo la pauta en los principales indicadores macroeconómicos, sino también en la propia actividad empresarial con sus consiguientes efectos sobre el empleo. Para mayor información acerca de Producto Bruto Interno y otros indicadores económicos se puede ver Blanchard Oliver [2]. En términos macro-económicos el estudio del Consumo Privado es de suma importancia ya que dinamiza la economía en conjunción con la Inversión Privada.

4.3 Análisis de datos

En este sentido analizando los últimos 10 años, se puede apreciar que se pasa de un consumo deprimido a inicio de los 90 como consecuencia de la hiperinflación y el shock de aquel entonces, para luego pasar a un crecimiento sostenido a partir del año 1993, dadas las condiciones en el año 1992 como por ejemplo el acceso fácil y rápido a la denominada banca de consumo, privatizaciones, etc. (ver figura 4.1). Se sabe que a finales del año 1999 y a inicios del año 2000, el estancamiento del consumo se debió principalmente a la aguda recesión, consecuencia de los problemas políticos suscitados en nuestro país y de la que aún se está tratando de salir en estos momentos. Al ver esta gráfica se observa el crecimiento sostenido, luego se puede **modelar** el consumo privado Y_t . en una función de un nivel μ_t adicionando a una componente estacional ψ_t . Un componente de tendencia β_t , pues según esta figura se puede observar una estructura de crecimiento a partir del año 1993 hacia adelante modelamos esta componente μ_t para una posible “quiebra” de tendencias dado por un crecimiento o decrecimiento a

través de la inducción del término β_t en μ_t ; es decir:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \delta\mu_t.$$

Una componente de estacionalidad ψ_t , pues observamos un comportamiento periódico trimestral en cada año. Utilizamos el formato Free-Form para explicar la estacionalidad trimestral.

Finalmente, postulamos el siguiente modelo para el consumo privado Y_t ,

$$\begin{aligned}y_t &= \mu_t + \psi_{t,0} + \nu_t, \\ \mu_t &= \mu_{t-1} + \beta_{t-1} + \delta\mu_t, \\ \beta_t &= \beta_{t-1} + \delta\beta_t, \\ \psi_{t,0} &= \psi_{t-1,1} + \omega_{t,0}, \\ \psi_{t,1} &= \psi_{t-1,2} + \omega_{t,1}, \\ \psi_{t,2} &= \psi_{t-1,3} + \omega_{t,2}, \\ \psi_{t,3} &= \psi_{t-1,0} + \omega_{t,3}.\end{aligned}\tag{4.1}$$

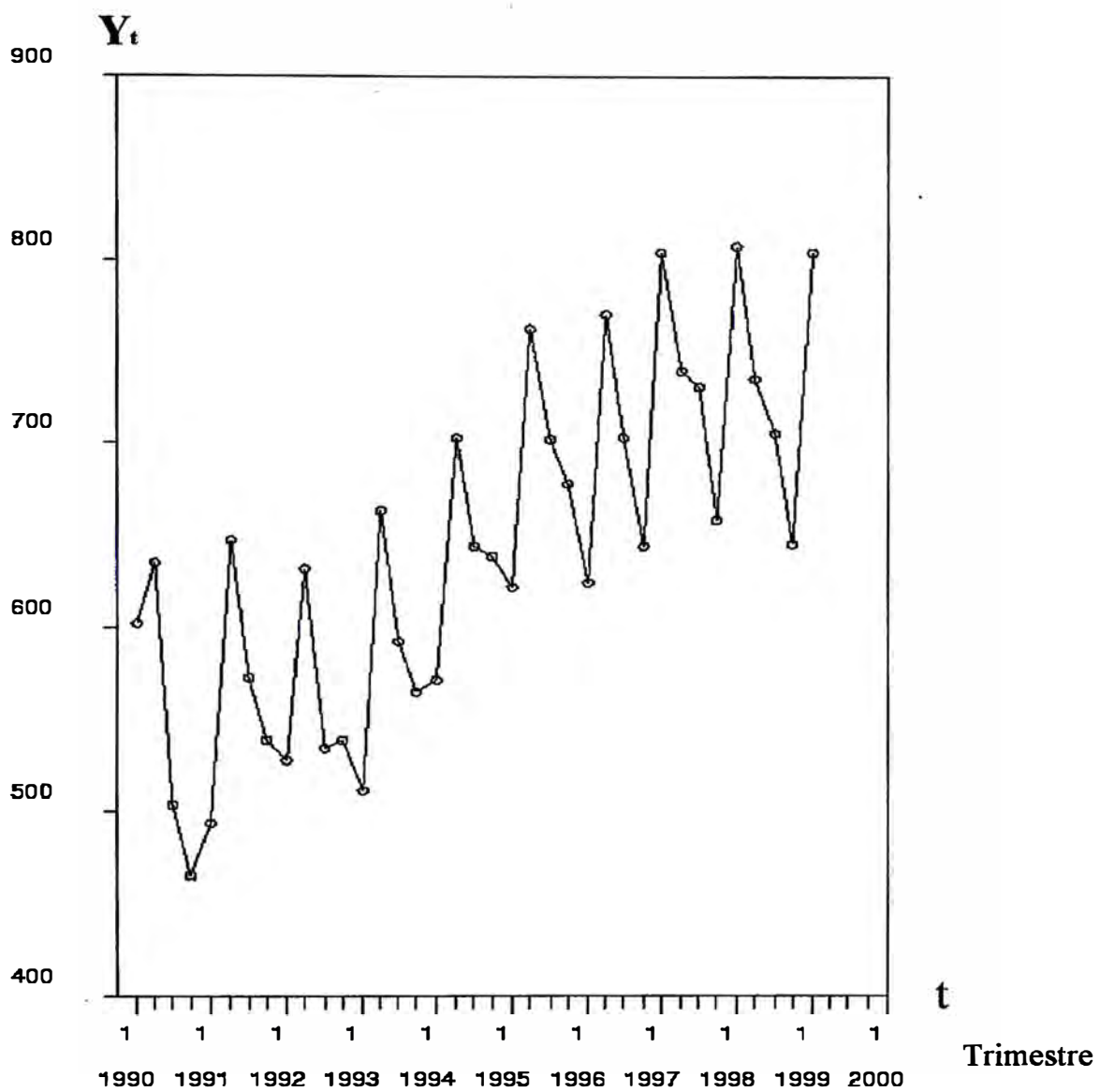


Figura 4.1: Consumo Privado a soles constantes de 1979

4.4 Modelo Lineal Dinámico para el Consumo Privado

Para usar la teoría desarrollada en el Capítulo 3, coloquemos la ecuación (4.1) en un formato MLD, según las ecuaciones (3.15) y (3.16). En realidad (4.1) es un caso particular representado en la forma de una combinación de un modelo polinomial de segundo-orden y un modelo estacional, para mayores detalles sobre estos casos particulares ver cap. 7 y 8 de West and Harrison [21].

Brevemente, un modelo polinomial de **Segundo – Orden** es caracterizado por la cuádrupla $\{E_2, J_2(1), V_t, W_t\}$ donde:

$$E_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{y} \quad J_2(1) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

se puede escribir el modelo en términos de las ecuaciones usuales como:

Ecuación de Observación: $Y_t = \mu_t + \nu_t,$

Ecuación sistema: $\mu_t = \mu_{t-1} + \omega_{t1},$

$$\beta_t = \beta_{t-1} + \omega_{t2},$$

si se define:

$$\theta_t^p = \begin{pmatrix} \theta_{t1} \\ \theta_{t2} \end{pmatrix} = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} \quad \text{y} \quad \omega_t^p = \begin{pmatrix} \omega_{t1} \\ \omega_{t2} \end{pmatrix} = \begin{pmatrix} \delta\mu_t \\ \delta\beta_t \end{pmatrix}$$

Se puede escribir el modelo polinomial de primer-orden como:

Ecuación de Observación: $Y_t = E_2' \theta_t^p + \nu_t,$

Ecuación sistema: $\theta_t^p = J_2(1) \theta_{t-1}^p + \omega_t^p.$

De igual forma, un modelo estacional con periodo **trimestral** es caracterizado por

la cuádrupla $\{E_4, P, V_t, W_t\}$ donde:

$$E_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad y \quad P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & I \\ 1 & 0' \end{pmatrix}$$

Luego la forma del modelo estacional usual puede ser escrito como:

Ecuación de observación: $Y_t = E_p' \psi_t + \nu_t,$

Ecuación sistema: $\psi_t = P\psi_{t-1} + \omega_t^e,$

si se define:

$$\psi_t = \begin{pmatrix} \psi_{t0} \\ \psi_{t1} \\ \psi_{t2} \\ \psi_{t3} \end{pmatrix} \quad y \quad \omega_t^e = \begin{pmatrix} \omega_{t0} \\ \omega_{t1} \\ \omega_{t2} \\ \omega_{t3} \end{pmatrix}$$

Se puede escribir el modelo estacional con $p = 4$ como:

Ecuación de observación: $Y_t = E_4' \psi_t + \nu_t,$

Ecuación sistema: $\psi_t = P\psi_{t-1} + \omega_t^e.$

Finalmente, nuestro modelo (4.1) puede ser expresado como un MLD, ecuaciones (3.3) (3.4), con matrices:

$$F_t = \begin{pmatrix} E_2 \\ E_4 \end{pmatrix} \quad G_t = \begin{pmatrix} J_2(1) & 0 \\ 0 & P \end{pmatrix} \quad \theta_t = \begin{pmatrix} \theta_t^p \\ \psi_t \end{pmatrix} \quad \omega_t = \begin{pmatrix} \omega_t^p \\ \omega_t^e \end{pmatrix}$$

En donde $\delta\mu_t$, $\delta\beta_t$ y $\omega_{t,r}$ son los errores de evolución. En este modelo β_t tiene la interpretación del crecimiento en el nivel de la serie producido entre el tiempo $t - 1$ a t , evolucionando en dicho intervalo de manera suave y perturbada por una cantidad estocástica $\delta\beta_t$ μ_t es el nivel en si mismo en el instante t , evolucionando sistemáticamente

por la adición de β_t y estocásticamente por $\delta\mu_t$, y finalmente $\psi_{t,r}$ representa la componente estacional t , debido a la naturaleza trimestral de la información.

El modelo se ajustó usando una a priori de referencia no informativa y un factor de descuento de 0.9 para la tendencia, 0.95 para el factor estacional y 0.99 para la varianza, el empleo de factores de descuento es una forma de modelar la pérdida de información que se dá con la dinámica de los datos.

4.4.1 Análisis de los resultados

Figura 4.2.- lo que nos muestra son las tendencias actualizadas, esto es $(\mu_t | D_t)$. Ahora, como era de esperarse, μ_t describe la trayectoria de la tendencia del conjunto de observaciones Y_t . El hecho de considerar una estructura de crecimiento y/o decrecimiento β_t en el componente de la tendencia μ_t , nos permitió “capturar” un decrecimiento de la tendencia μ_t en el año 1990, seguido de un crecimiento sostenido hasta el año 1996. A partir de este año se observa una especie de saturación en la tendencia μ_t del consumo privado hasta el primer trimestre del año 1999.

Figura 4.3.- Recuerde que modelamos β_t por un camino aleatorio, es decir $\beta_t = \beta_{t-1} + \delta\beta_t$. La intención es “capturar” crecimiento o decrecimiento en la tendencia μ_t moderadamente. Sin embargo, se observa en esta figura que hay un camino “brusco” entre el cuarto trimestre de 1990 y el primer trimestre de 1991, lo que sugiere averiguar que pasó en esa fecha para ver la posibilidad de un análisis de intervención. Por otro lado, esta gráfica configura lo expuesto en la figura de μ_t . Es decir, hasta el año 1990 el crecimiento es negativo (decrecimiento en la tendencia del consumo privado). Es a partir del segundo trimestre de 1991 que el crecimiento es positivo manteniendose hasta el segundo trimestre de 1993. A partir del tercer trimestre de 1993 hasta 1996

vemos un moderado aumento en el crecimiento de la tendencia μ_t , para en los últimos tres años 1997,1998 y primer trimestre 1999 disminuyendo levemente pero aún positivo.

Figura 4.4.- Nos detalla como las componentes del vector de factores estacionales ψ_t se comporta en el tiempo ($\psi_t = (\psi_{0,t}, \psi_{1,t}, \psi_{2,t}, \psi_{3,t})$ si t corresponde al primer trimestre de cada año). Las regiones de credibilidad al 95% inicialmente son mayores pues poca información Y_t esta siendo procesada. A medida que disponemos de más información estas regiones de credibilidad son mas precisas.

Según lo sugerido en el modelo, cada componente del vector de factores estacionales varía levemente segun un “camino aleatorio”. Así se puede observar que toda nueva información será capturada por un factor estacional, en particular, después de cuatro periodos y que la matriz de evolución W_t inducirá correlación entre los factores estacionales. (Aquí se esta usando un único factor de descuento para esta matriz de evolución W_t).

Se debe anotar también que, no hay un cambio “brusco” en la estacionalidad entre el cuarto trimestre de 1991 y el primer trimestre de 1992, como fue detectado en la componente de tendencia μ_t . Esto indica o sugiere que este cambio se debe particularmente a factores no estacionales.

Figura 4.5.- según puede observarse en esta figura, las regiones de credibilidad al 95% para las distribuciones predictivas un paso al frente son menos precisas, pues se tiene poca información Y_t . A medida que se dispone de nuevas observaciones estas regiones de credibilidad son más precisas. Se puede notar también que los pronósticos un paso al frente para el año de 1991 son ”perturbadas” por el cambio “brusco” dado entre el cuarto trimestre de 1990 y el primer trimestre de 1991. Es a partir de este año 1992 que el Modelo Lineal Dinámico (MLD) usado en esta tesis, los pronósticos un

paso al frente son coherentes comparados con los valores reales.

Tabla 4.1: Pronósticos para el Consumo Privado (MLD)
Millones de Nuevos Soles

Trimestre	Pronóstico	Valor Real	Error de Pronóstico
1999.2	811.2	804.23	6.97
1999.3	760.8	738.59	22.21
1999.4	718.1	739.72	21.62
2000.1	794.4	680.20	114.2

De esta tabla se desprende que a medida que el horizonte de pronósticos es mayor nuestros pronósticos son menos precisos y esto es por la propuesta del (MLD), lo recomendable sería dar pronóstico un paso al frente.

Figura 4.6.- Muestra un análisis retrospectivo del “nivel” del Modelo Lineal Dinámico (MLD), $F_t' \theta_t$, (donde $F_t' = [(0,1), (1,0,0,0)]$, $\theta_t = [(\mu_t, \beta_t)'; (\psi_{t0}, \psi_{t1}, \psi_{t2}, \psi_{t3})']$; 't' correspondiente al primer trimestre.) 'T' representa el número total de datos observacionales Y_t . Esto nos permite comparar los niveles actuales frente a los niveles anteriores, a la luz de toda la información disponible. Si comparamos esta figura 4.6 con la figura 4.7 se puede concluir que las estimaciones retrospectivas del nivel son más “suaves”. esto es explicado por el hecho de disponer más información en la estimación del “nivel” $F_t' \theta_t$. También que no existe grandes cambios en el comportamiento de la estacionalidad trimestral, así como de la tendencia.

Figura 4.7.- Para un t correspondiente al primer trimestre, se tiene que:

$$\begin{aligned}
 \mathbf{Y}_t &= (1, 0, 0, 0)(\psi_{t0}, \psi_{t1}, \psi_{t2}, \psi_{t3})' + (1, 0)(\mu_t, \beta_t)' + \nu_t, \\
 \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \beta_{t-1} \end{pmatrix} + \begin{pmatrix} \omega_{t4} \\ \omega_{t5} \end{pmatrix}, \\
 \psi_{t,j} &= \psi_{t-1,j+1} + \omega_{t,j}; \quad j=0,1,2 \\
 \psi_{t,3} &= \psi_{t-1,0} + \omega_{t,3}.
 \end{aligned}$$

Se puede observar en la figura que los componentes de la tendencia μ_t y de estacionalidad $\psi_{t,j}$ ($j = 0, 1, 2, 3$) describe que satisfactoriamente el comportamiento del consumo privado \mathbf{Y}_t . La naturaleza dinámica impuesta a los datos permite esta flexibilidad en el control de ambas componentes en conjunto y por separado.

Las regiones de credibilidad generadas son al 95%. De 37 observaciones, 34 de ellas caen dentro de estas regiones de credibilidad haciendo coherente la aproximación.

Figura 4.8.- Como se sabe los residuales son la diferencia entre los valores observados y los valores predcidos o valores ajustados. Esta es la parte de la observación que no está explicada por el modelo. La figura muestra los cuantiles empíricos residuales vs. cuántiles teóricos si los pares de puntos se sitúan alrededor de una línea recta , esto sugiere que los residuales se comportan como una distribución Normal. Caso contrario los residuales sugieren que el (MLD) 4.1 no está explicando completamente la variable macro-económica consumo privado.

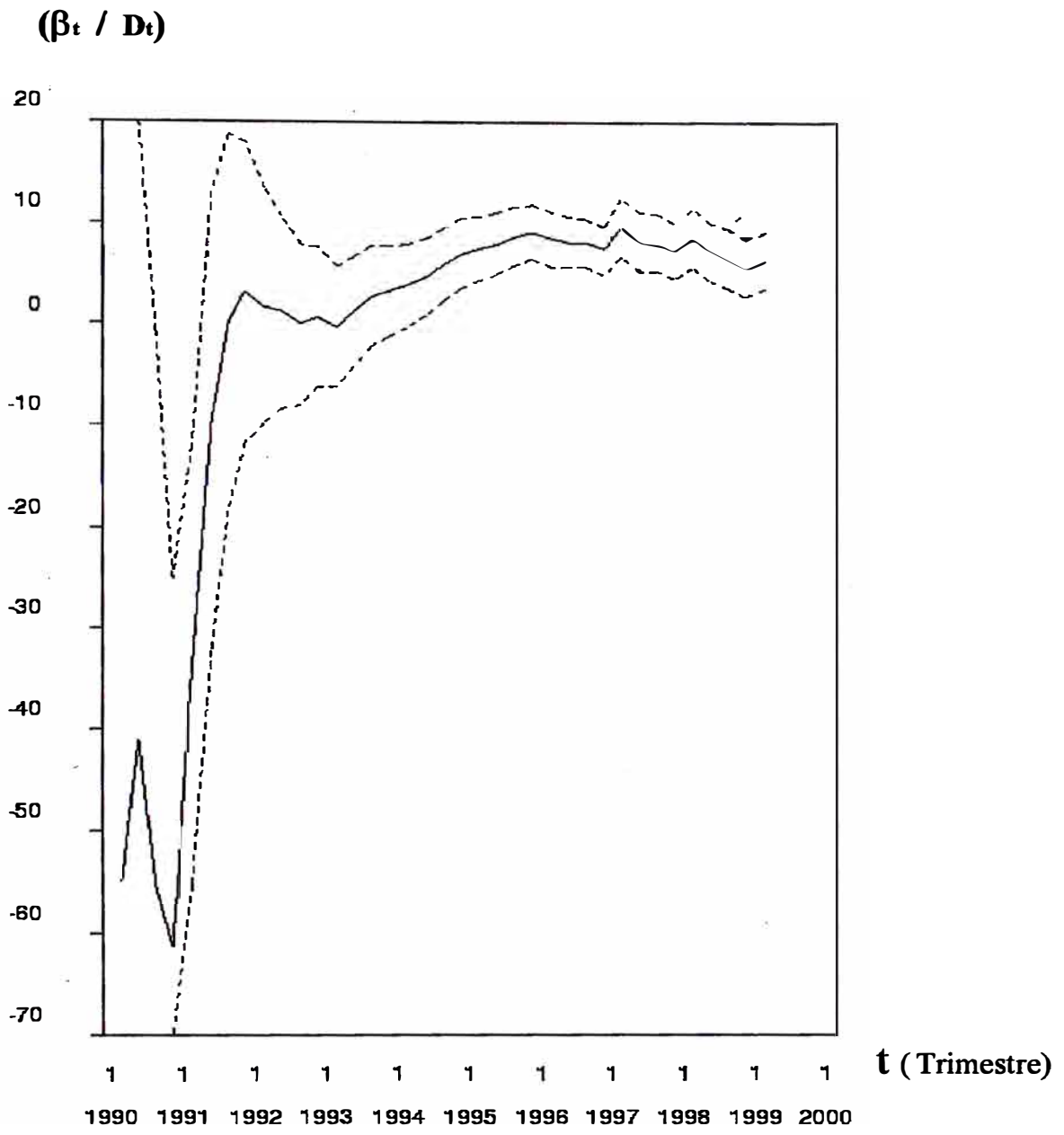


Figura 4.3: Evolución del crecimiento para la serie del consumo privado

$$(\mu_{T-k} + \psi_{T-k,J} / DT)$$

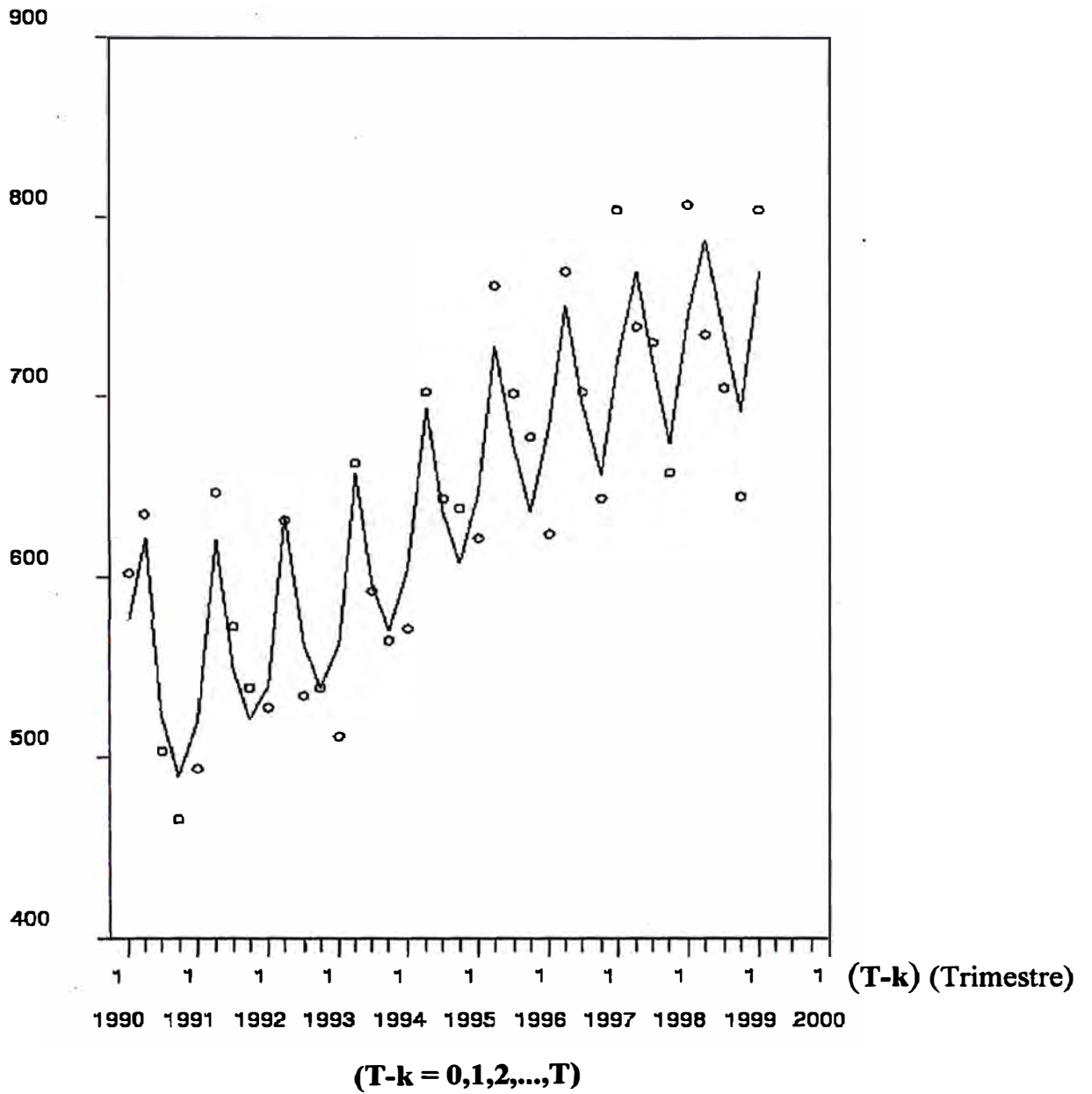


Figura 4.6: Filtración de la función de respuesta media, $F'_{T-k} \theta_{T-k}$

$$(\mu_t + \psi_{tj} / D_t)$$

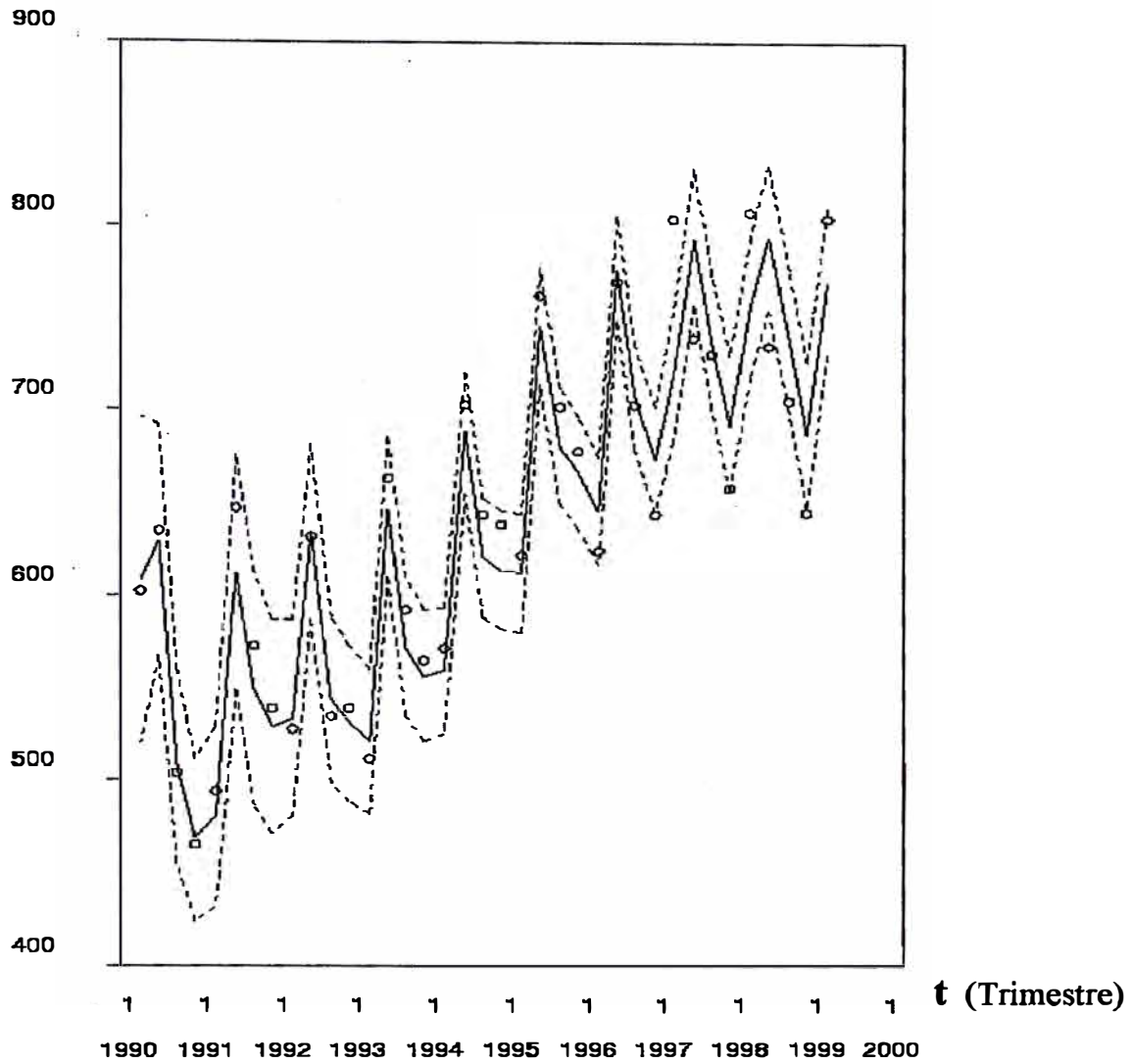


Figura 4.7: Actualización de la función de respuesta media, $F_t'\theta_t$

Ploteo QQ

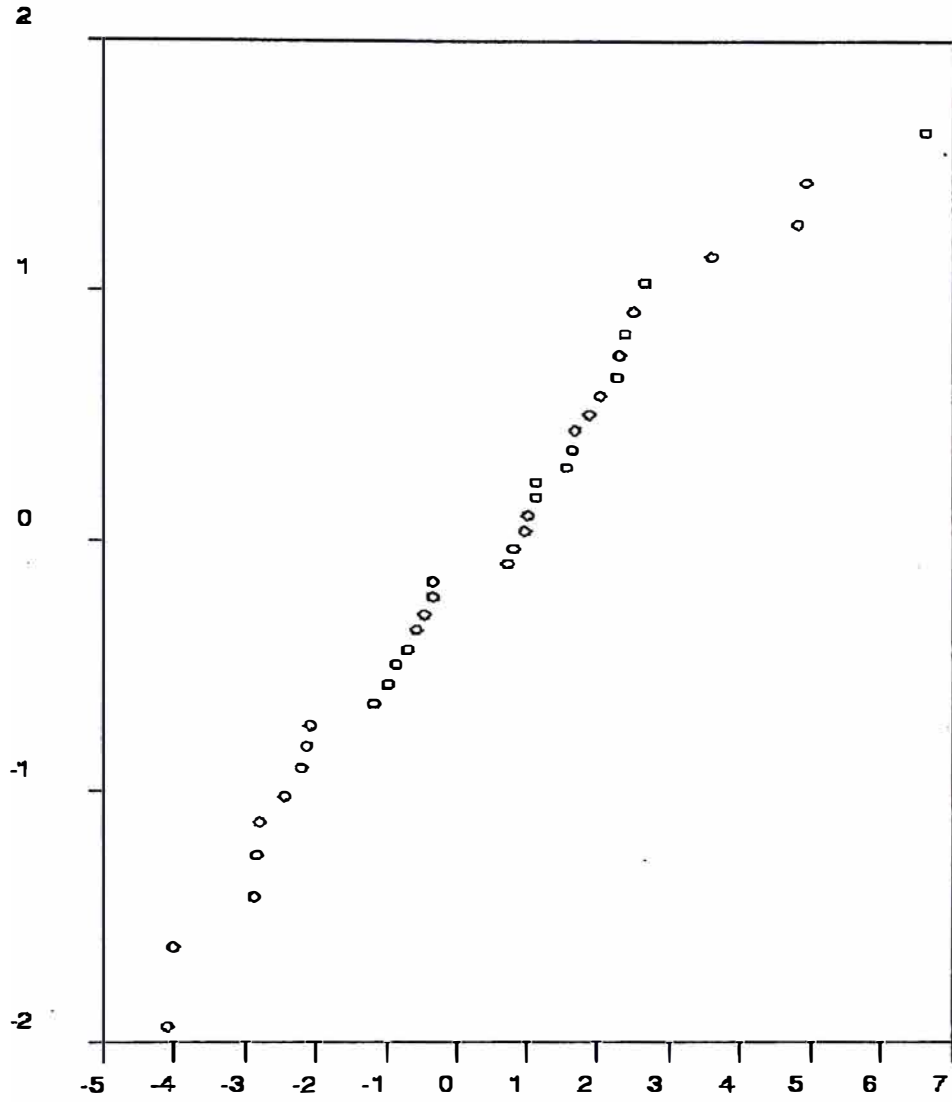


Figura 4.8: Ploteo probabilidad Normal para los Residuales

4.4.2 Modelos Clásicos

El modelo que se ha propuesto es más elaborado que los modelos antiguamente utilizados. Por que, son modelos Dinámicos, por ejemplo, mientras que anteriormente se utilizaba modelo clásico como $Y_t = \alpha_0 + \beta_0 t + \epsilon_0$. Ahora, utilizando (MLD) para cada tiempo t existe un modelo que torna el modelo más flexible que el anterior que solo usa un modelo, una vez que para cada tiempo t se usa un modelo finalmente se tiene una secuencia de modelos implicando ello una mejora sustantiva en el modelamiento a continuación se puede ver los diferentes modelos clásicos usando la metodología frecuentista. En los años 80 se utilizaba los siguientes modelos para explicar el consumo privado del Perú ver Hiep y John Kuiper [22].

Lineal: $Y_t = \alpha_0 + \beta_0 t + \epsilon_0$,

Cuadrática: $Y_t = \alpha_1 + \beta_1 t + \gamma_1 t^2 + \epsilon_1$,

Cúbica: $Y_t = \alpha_2 + \beta_2 t + \gamma_2 t^2 + \delta_2 t^3 + \epsilon_2$.

Bajo una aproximación clásica los parámetros eran estimados vía Máximo Verosimilitud o Mínimos Cuadrados. Por ejemplo, si se utiliza Métodos Mínimos Cuadrados para los modelos anteriormente descritos con los datos del consumo privado, se tiene los siguientes modelos estimados:

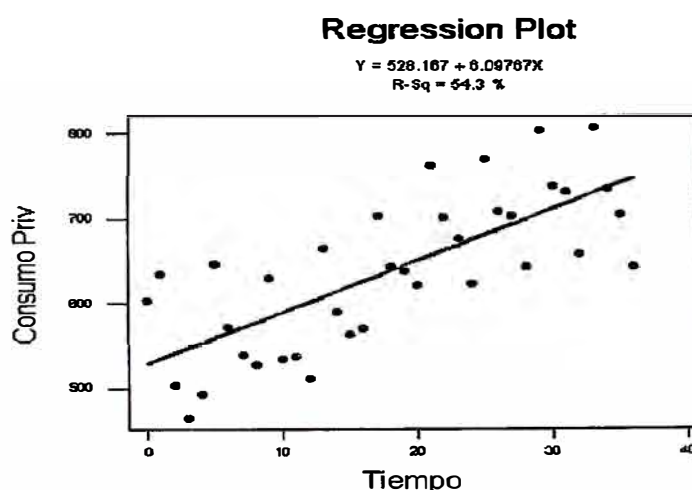


Figura 4.9: Regresión Lineal

Regression Plot

$$Y = 525.293 + 6.59046X - 1.37E-02X^2$$

R-Sq = 54.4 %

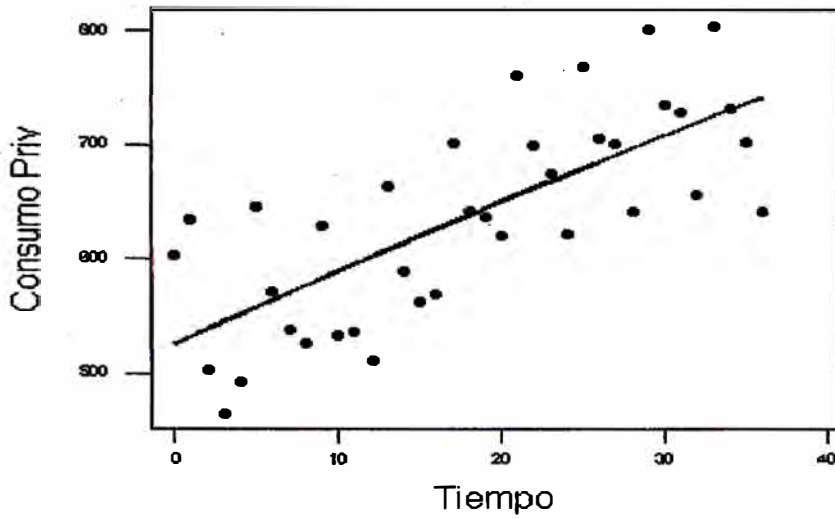


Figura 4.10: Regresión polinomial de 2do order

Regression Plot

$$Y = 594.908 - 14.7564X + 1.48922X^2 - 2.78E-02X^3$$

R-Sq = 63.4 %

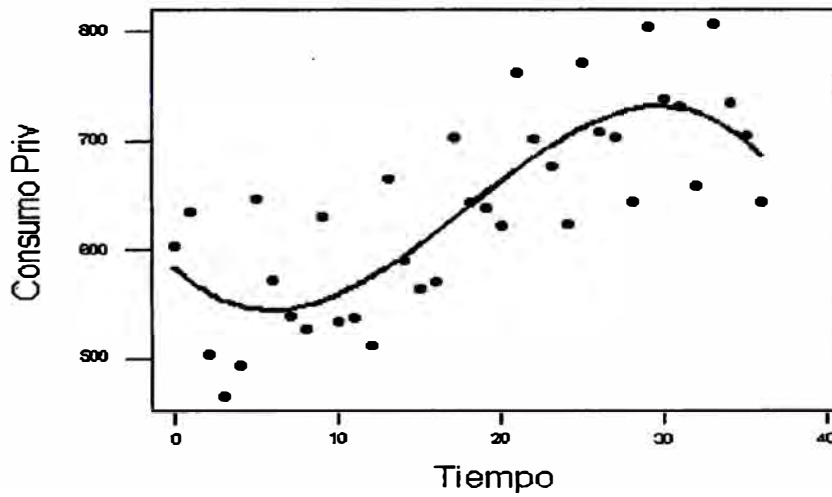


Figura 4.11: Regresión polinomial de 3er order

Capítulo 5

Conclusiones y Sugerencias

La principal contribución del presente trabajo de tesis, es mostrar cómo los Modelos Dinámicos Bayesianos Lineales pueden ser usados en el tratamiento de problemas reales de naturaleza dinámica, y en particular para la variable macroeconómica **consumo privado** usada en este trabajo.

El Modelo Dinámico Bayesiano particular usado para modelar la variable macroeconómica **consumo privado** fue un Modelo Dinámico por componentes, en donde modelamos el nivel medio por un modelo polinomial de segundo-orden, y la parte estacional por un modelo de forma libre con $p = 4$ (datos trimestrales). Los resultados obtenidos en estimación y pronósticos fueron aceptables. Esto es, los residuales informan que toda la información sobre el **consumo privado** fue explicado por este Modelo Dinámico Bayesiano particular, a no ser por algunos eventos relevantes “no esperados”, como por ejemplo el cambio “brusco” entre el cuarto trimestre de 1990 y el primer trimestre de 1991, en donde deberíamos incluir un análisis de intervención por tratarse de una economía inestable. Esta fase se recomienda ser usada como trabajo futuro, otros eventos son citados y explicados oportunamente en el capítulo anterior, que no fue absorbida por el modelo. Esto era de esperarse pues el Modelo Dinámico

Bayesiano tratada en la tesis no considera estos casos. Sin embargo, los eventos no inducen a no usar estos modelos.

En lo que se refiere a los pronósticos, se puede decir que son aceptables después de comparar algunos pronósticos, de horizonte un “paso al frente”, con datos reales.

Se puede concluir entonces, que después de la implementación de estos Modelos Dinámicos Bayesianos a nuestra variable macro-económica **consumo privado**, sugerimos la introducción de esta nueva metodología a todo ambiente que envuelva Dinamismo en el tiempo, y que esta puedan reemplazar las metodologías “estáticas” que tiene limitaciones en el tratamiento de problemas complejos. Algunos de los puntos que quedan pendientes para temas de interés son por ejemplo aplicaciones incluyendo variables regresoras.

Apéndice A

Distribuciones Multivariadas

A.1 La Distribución Normal Multivariada

Para toda Distribución Normal Multivariada de la forma:

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N \left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right]$$

Puedo expresar la Distribución Condicional como:

$$(X_1 | X_2) \sim N[m_1(X_2), V_1(X_2)],$$

donde

$$m_1(X_2) = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2),$$

y

$$V_1(X_2) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.$$

donde $\mu_1 = E[X_1]$ y $\mu_2 = E[X_2]$ además

$$\Sigma_{11} = \text{Var}(X_1),$$

$$\Sigma_{22} = \text{Var}(X_2),$$

$$\Sigma_{12} = \text{Cov}(X_1, X_2),$$

$$\Sigma_{21} = \text{Cov}(X_2, X_1).$$

También

$$(X_2 | X_1) \sim N[m_2(X_1), V_2(X_1)],$$

donde

$$m_2(X_1) = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1),$$

y

$$V_2(X_1) = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}.$$

A.2 Distribución T-Student

La distribución multivariada Normal/gamma dice que si $\phi \sim G[n/2, d/2]$ y $(\mathbf{X} | \phi) \sim N[m, C\phi^{-1}]$ entonces se puede afirmar que \mathbf{X} tiene una distribución marginal T.

$$\mathbf{X} = T_n[m, Cd/n].$$

Apéndice B

Conjunto de Datos

Los datos (Fuente BCR) del Perú que se analizan corresponden a la serie del Consumo Privado trimestral entre los años 1990 y 1999 a soles constantes del año base 1979. Esta información servirá para el contraste del pronóstico de los valores para los trimestres del año 2000 con los valores reales de los mismos y ver cuan efectivo puede ser el uso de los modelos propuestos.

Tabla B.1: Consumo Privado 1990.1 - 1999.1**Millones de Nuevos Soles**

Trimestre	Consumo Privado Real	Trimestre	Consumo Privado Real
1990.1	602.94	1995.1	621.52
1990.2	635.66	1995.2	761.98
1990.3	502.57	1995.3	701.64
1990.4	464.60	1995.4	677.12
1991.1	492.65	1996.1	623.60
1991.2	647.38	1996.2	770.31
1991.3	572.06	1996.3	707.11
1991.4	538.79	1996.4	703.23
1992.1	527.39	1997.1	643.73
1992.2	630.30	1997.2	803.99
1992.3	533.67	1997.3	738.16
1992.4	537.63	1997.4	730.34
1993.1	510.71	1998.1	658.61
1993.2	664.28	1998.2	807.06
1993.3	591.06	1998.3	734.13
1993.4	564.36	1998.4	704.83
1994.1	570.59	1999.1	644.09
1994.2	703.61		
1994.3	643.35		
1994.4	638.19		

Bibliografía

- [1] Bayes (1763). An essay towards solving a problem in the doctrine of chances. *philos. trans. R. Soc. London*, 53 370-418.
- [2] Blanchard Oliver. *Macroeconomía*. Prentice Hall International (UK) Ltd.
- [3] Berger, James(1985). *Statistical Decision Theory and Bayesian Analysis*, Second Edition. Springel-Verlang, New York.
- [4] Bernardo, J.M and Smith, A.F. (1994) *Bayesian Theory*. 1st Edition, Wiley, London.
- [5] Box, G.E. and Tiao, G.C (1973) *Bayesian Inference in Statistical Analysis*. Reading, Mass Adissolo-Wesley.
- [6] Gamerman, D., and West, M. (1987) *A Time series application of dynamic survival models in unemployment studies*. *The Statistician*, **36**, 269-174.
- [7] Gamerman, D. (1991) *Dynamic Bayesian models for survival data*. *Applied Statistics*, **40**, 63-79.
- [8] Geweke,J. (1994) *Bayesian comparison of econometrics models*. Working Paper 532, Research Department, Federal Reserve Bank of Minneapolis.
- [9] Green, M., and Harrison, P.J. (1973) *Fashion forecasting for a mail order company*. *Operations Research Quarterly*, **24**, 193-205.

- [10] Harrison, P.J. (1988) *Bayesian forecasting in O.R.* Developments in Operational Research 1988, N.B. Cook and A.M. Jonson, eds. Pergamon Press, Oxford.
- [11] Johnston, F.R., and Harrison, P.J. (1980) *An application of forecasting in the alcoholic drinks industry.* Journal of Operational Research Society, **31**, 699-709.
- [12] Kleiner, B., Martin, R.D., and Thompson, D.J. (1979) *Robust estimation of power spectra* (with discussion). Journal of the Royal Statistical Society, Ser. B, **41**, 313-351.
- [13] Migon, H.S., and Harrison, P.J. (1985) *An Application of non-linear Bayesian forecasting to television advertising.* Bayesian Statistics 2, J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith, eds. North Holland, Amsterdam, and Valencia University Press.
- [14] Queen, C.M., e Smith, J.Q. (1992) *Dynamic graphical models.* Bayesian Statistic 4, J.O. Berger, J.M. Bernardo, A.P. Dawid and A.F.M. Smith, eds. Oxford University Press. pp 741-762
- [15] Quintana, J.M., and West, M. (1987) *Multivariate time series analysis: new techniques applied to international exchange rate data.* The Statistician, **36**, 275-281.
- [16] Quintana, J.M., and West, M. (1988) *Time series analysis of compositional data.* Bayesian Statistics 3, J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith, eds. Oxford University Press. pp 747-756.
- [17] Quintana, J.M. (1992) *Optimal portfolios of forward currency contracts.* Bayesian Statistics 4, J.O. Berger, J.M. Bernardo, A.P. Dawid and A.F.M. Smith, eds. Oxford University Press. pp 753-762
- [18] West, M. (1992) *Modelling time-varying hazards and covariate effects* (with discussion). Survival Analysis: State of the Art, J.P. Klein and P.K. Goel, eds. Kluwer.

- [19] West, M. (1995a) *Bayesian inference in cyclical component dynamic linear models*. Journal of the American Statistical Association.
- [20] West, M. (1995b) *Bayesian Forecasting*. Discussion paper 95-11, ISDS, Duke University.
- [21] West, M., and Harrison, P.J. (1997) *Bayesian Forecasting and Dynamic Models*. 2nd edition. Springer-Verlag: New York.
- [22] Working paper Perú 80/1 Fuente Hiep Cao John Kuiper (1979) *El uso de Modelos Econométricos y la Planificación del desarrollo en el Perú*