5. Modelos dinámicos

- Los modelos lineales dinámicos son un caso particular de una clase más grande de modelos dinámicos. En general, los modelos dinámicos se caracterizan por tener una "dinámica" en los parámetros del modelo, i.e., los parámetros no son fijos sino cambiantes o dependientes del tiempo.

- Los modelos dinámicos tienen su principal aplicación en el análisis de series de tiempo o en modelos de regresión con errores auto-correlacionados. En general son muy útiles en análisis secuenciales debido a que la actualización de los parámetros debe de hacerse con base en datos obtenidos secuencialmente.

- MODELO LINEAL: Considera el modelo de regresión lineal $Y_t = x_t'\beta + \varepsilon_t$, donde x_t contiene p variables predictoras, siendo la primera el intercepto. Suponga ahora que el vector de coeficientes β cambia con el tiempo, i.e.,

  Observación: $Y_t = x_t'\beta_t + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, V_t^{-1}\right),$

  Evolución: $\beta_t = G_t\beta_{t-1} + \omega_t, \quad \omega_t \sim N\left(0, W_t^{-1}\right)$

  para t=1,2,..., donde V_t^{-1} y W_t^{-1} son precisiones y G_t es una matriz.

- *Explicación*:
  1) La primera ecuación define el modelo observacional para una respuesta Y_t y su relación con p covariables o variables explicativas x_t. La primera variable explicativa es por lo general una constante o intercepto que representa el nivel de la serie. Como Y_t es univariado, entonce β_t es un

vector de la forma $(\beta_{0t}, \beta_{1t}, \ldots, \beta_{p-1t})'$. Es posible considerar a Y_t como multivariado, en cuyo caso β_t es una matriz de dimensión m×p.

2) La segunda ecuación, conocida como la ecuación de estado o de sistema, especifica la evolución de los parámetros a través del tiempo. Si el modelo incluye p coeficientes cambiantes, entonces la evolución se define mediante una matriz de transición G_t de dimensión p×p.

3) La forma general del modelo lineal dinámico tiene errores ε_t y ω_t con varianzas dependientes del tiempo V_t y W_t que denotan la varianza observacional y la varianza del sistema respectivamente.

- CASOS PARTICULARES:

1) Si G_t=I y W_{jt}=0 para j=0,...,p-1 entonces, el modelo lineal dinámico se reduce al modelo usual de regresión lineal múltiple con coeficientes fijos.

2) En general W_t es una matriz de varianzas y covarianzas de dimensión p×p, sin embargo para simplificar la estimación, es común suponer que los parámetros varían con el tiempo independientemente unos de otros (covarianzas cero), por lo que en este caso W_t sería una matriz diagonal $(W_{0t}, W_{1t}, \ldots, W_{p-1t})$.

3) Más aún, es común también suponer que las varianzas V_t y W_t son constantes en el tiempo. En este caso la dinámica sólo se representa en los coeficientes de regresión.

- MODELO SIMPLIFICADO.

  Observación: $Y_t = x_t'\beta_t + \varepsilon_t$, $\varepsilon_t \sim N(0, V^{-1})$,

  Evolución: $\beta_{jt} = G_j \beta_{jt-1} + \omega_{jt}$, $\omega_{jt} \sim N(0, W^{-1})$, j=0,1,...,p-1

  para t=1,2,....

- ANÁLISIS DEL MODELO: Consideremos el modelo general lineal dinámico. El análisis Bayesiano requiere de especificar una distribución inicial sobre los parámetros del modelo, que en este caso son $\beta_1, \beta_2, \ldots$ donde $\beta_t' = (\beta_{0t}, \beta_{1t}, \ldots, \beta_{p-1t})$. Debido a la relación de evolución de los coeficientes β_t, es suficiente con establecer una distribución inicial sobre el vector β_0.

  Sea D_t la información disponible al tiempo t, de manera que:

  D_0 es la información inicial y

  $D_t = D_{t-1} \cup \{y_t\}$

  La forma de realizar la actualización es de manera recursiva, mediante el *Teorema de Bayes secuencial*, i.e.,

  Inicial (al tiempo t): $f(\theta_t | D_{t-1}) = \int f(\theta_t | \theta_{t-1}) f(\theta_{t-1} | D_{t-1}) d\theta_{t-1}$

  Final (al tiempo t): $f(\theta_t | D_t) \propto f(y_t | \theta_t) f(\theta_t | D_{t-1})$

  Predictiva (a un paso): $f(y_t | D_{t-1}) = \int f(y_t | \theta_t) f(\theta_t | D_{t-1}) d\theta_t$

En el caso particular del modelo lineal dinámico normal tenemos,

Inicial (al tiempo 0): $f\left(\beta_0 \middle| D_0\right) = N\left(\beta_0 \middle| m_0, C_0^{-1}\right)$

Final (al tiempo t–1): $f\left(\beta_{t-1} \middle| D_{t-1}\right) = N\left(\beta_{t-1} \middle| m_{t-1}, C_{t-1}^{-1}\right)$

Inicial (al tiempo t): $f\left(\beta_t \middle| D_{t-1}\right) = N\left(\beta_t \middle| a_t, R_t^{-1}\right)$

Predictiva (a un paso): $f\left(y_t \middle| D_{t-1}\right) = N\left(y_t \middle| q_t, Q_t^{-1}\right)$

donde, $a_t = G_t m_{t-1}$,

$$R_t = G_t C_{t-1} G_t' + W_t$$

$$q_t = x_t' a_t$$

$$Q_t = x_t' R_t x_t + V_t$$

$$m_t = a_t + R_t x_t \left(V_t + x_t' R_t x_t\right)^{-1} \left(y_t - q_t\right)$$

$$C_t = R_t - R_t x_t \left(V_t + x_t' R_t x_t\right)^{-1} x_t' R_t$$

Las ecuaciones para a_t, R_t, m_t y C_t son conocidas como las recursiones del *filtro de Kalman*.

- Las ecuaciones recursivas de actualización Bayesiana se complican, o simplemente no existen en forma explícita, si se asigna una distribución inicial a las varianzas V, W y C_0. Excepto en un caso, cuando

$$W_t = V W_t^*, \text{ y}$$

$$C_0 = V C_0^*$$

$$\text{con } \tau_y = V^{-1} \sim Ga\left(\frac{n_0}{2}, \frac{n_0 S_0}{2}\right).$$

En este caso, condicional en V, las ecuaciones recursivas de actualización son iguales a las anteriores junto con una ecuación más de actualización para V, ó τ_y, de la forma

$$\tau_y | D_t \sim \text{Ga}\left(\frac{n_t}{2}, \frac{n_t S_t}{2}\right), \text{ con}$$

$$n_t = n_{t-1} + 1, \text{ y}$$

$$S_t = S_{t-1} + \frac{S_{t-1}}{n_t}\left(\frac{e_t^2}{Q_t} - 1\right).$$

- SUAVIZAMIENTO. Otro caso interesante en las especificaciones del modelo lineal simplificado, es cuando las precisiones de las ecuaciones de observación y de evolución están relacionadas de la siguiente manera. Denotemos por $\tau_y = V^{-1}$ y $\tau_\beta = W^{-1}$ las respectivas precisiones, y sea $\lambda>0$ un parámetro. Si

$$\tau_\beta = \lambda \tau_y,$$

entonces λ juega el papel de *factor de suavizamiento*.

  - Para valores grandes de λ la serie predicha por el modelo producirá valores mucho más suaves que la serie observada. Por el contrario, si λ es pequeño la serie predicha será más parecida a la serie de datos observada.

- MODELOS CON TENDENCIA LINEAL. Existen varias formas de definir una tendencia lineal en un modelo dinámico lineal. Aquí se presentan dos opciones:

- Modelo con pendiente θ. Para t=1,2,...

$$Y_t = \beta_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, V^{-1}),$$

$$\beta_t = \theta\beta_{t-1} + \omega_t, \quad \omega_t \sim N(0, W^{-1}),$$

$$\beta_0 \sim N(0, C_0) \quad y \quad \theta \sim N(0, C_\theta)$$

- Modelo dinámico de segundo orden. Para t=1,2,…

$$Y_t = \beta_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, V^{-1}),$$

$$\beta_t = \beta_{t-1} + \alpha_{t-1} + \omega_{t1}, \quad \omega_{t1} \sim N(0, W_1^{-1}),$$

$$\alpha_t = \alpha_{t-1} + \omega_{t2}, \quad \omega_{t2} \sim N(0, W_2^{-1})$$

$$\beta_0 \sim N(0, C_{01}) \quad y \quad \alpha_0 \sim N(0, C_{02})$$

➢ EJERCICIO 5. Producción de leche (Congdon, 2001). Se tienen los datos anuales de producción de leche (Y_t) en lbs$\times 10^9$, y número de vacas (x_t) en unidades$\times 10^6$, en el período de 1970 a 1982.

El modelo sugerido para estos datos es: para t=1,2,...,13

Observación: $Y_t = \beta_t x_t + \varepsilon_t$, $\varepsilon_t \sim N(0, V^{-1}) \Rightarrow Y_t \sim N(\beta_t x_t, V^{-1})$

Evolución: $\beta_t = \beta_{t-1} + \omega_t$, $\omega_t \sim N(0, W^{-1}) \Rightarrow \beta_t \sim N(\beta_{t-1}, W^{-1})$

Además sugieren una varianza constante V=1 y W=0.05.

➢ EJERCICIO 6. Participación de mercado, promoción y precios (Congdon, 2001). Se tiene una serie de tiempo semanal de la participación de mercado (S_t) de un producto de consumo, durante un período de dos años (1990-1991). El número total de observaciones es N=104. Las fluctuaciones en la participación de mercado están relacionadas con (P_t) precio del producto relativo al promedio para esos productos, ($OPROM_t$) índice del nivel de promoción del producto y ($CPROM_t$) índice de productos competidores. El impacto en la participación de mercado debido a un aumento del precio y a un aumento de la promoción de la competencia debería de ser negativo, mientras que un aumento en la promoción del propio producto debería aumentar la participación del mercado. Se sugieren dos modelos para estos datos:

a) Modelo estático:

$$S_t = \beta_0 + \beta_1 P_t + \beta_2 OPROM_t + \beta_3 CPROM_t + \varepsilon_t, \quad \varepsilon_t \sim N(0,\tau)$$

Las distribuciones iniciales sugeridas por Congdon (2001) son:

$$\beta_0 \sim N(42,0.04),\ \beta_1 \sim N(0,0.25),\ \beta_2 \sim N(0,0.25) \text{ y } \beta_3 \sim N(0,0.25)$$

b) Modelo dinámico:

$$S_t = \beta_{0t} + \beta_{1t} P_t + \beta_{2t} OPROM_t + \beta_{3t} CPROM_t + \varepsilon_t,\ \ \varepsilon_t \sim N\left(0, V_t^{-1}\right)$$

$$\beta_{jt} = \beta_{jt-1} + \omega_{jt},\ \ \omega_{jt} \sim N\left(0, W_{jt}^{-1}\right),\ \text{j=0,1,2,3}$$

$$V_t^{-1} = \delta^{t-1} V^{-1},\ \text{t=1,2,...}$$

$$W_{jt}^{-1} = \left(\delta_j\right)^{t-1} W_j^{-1},\ \text{j=0,...,3 y t=1,2,...}$$

donde δ y δ_j son factores de descuento de las varianzas observacional y de la evolución de los parámetros, respectivamente.

Los valores de los factores de descuento sugeridos son:

$$\delta = 0.99,\ \delta_0 = 0.99,\ \delta_1 = 1,\ \delta_2 = 1 \text{ y } \delta_3 = 1.$$

Las distribuciones iniciales son:

$$\beta_{00} \sim N(42,0.04),\quad \beta_{10} \sim N(0,0.25) I(\beta_{10} < 0),\quad \beta_{20} \sim N(0,0.25) I(\beta_{20} > 0),$$

$$\beta_{03} \sim N(0,0.25) I(\beta_{30} < 0),\ V^{-1} \sim Ga(0.5,0.5) \text{ y } W_j^{-1} \sim Ga(1,1),\ \text{j=0,...,3.}$$

Modelos lineales generalizados dinámicos

- Los modelos lineales generalizados dinámicos son otro caso particular de la clase de modelos dinámicos. Esta clase general de modelos dinámicos se puede representar por las siguientes ecuaciones:

  Observación: $Y_t \mid x_t, \theta_t \sim f(y_t \mid x_t, \theta_t)$

  Evolución: $\theta_t \mid \theta_{t-1} \sim f(\theta_t \mid \theta_{t-1})$

- Si tomamos al modelo "f(·)" en la *ecuación de observación* como un modelo normal, obtenemos una ecuación de regresión normal como en el modelo estudiado anteriormente. Otras opciones son considerar miembros de la familia exponencial, produciendo una ecuación de regresión lineal generalizado.

- Por otro lado, si tomamos al modelo "f(·)" en la ecuación de evolución como un modelo normal más una relación lineal entre θ_t y θ_{t-1}, se obtiene una ecuación de observación lineal (normal). Es posible también considerar otras opciones no lineales para la ecuación de evolución, produciéndose modelos dinámicos no lineales.

- Cuando la serie de tiempo de interés está representada por una variable con soporte acotado o discreto, es necesario recurrir a los modelos de regresión lineal generalizados dinámicos.

➢ MODELO LINEAL GENERALIZADO: Considera un modelo de regresión lineal generalizado $Y_t | x_t = f(y_t | \mu_t, \phi_t)$, donde f(·) es un miembro de la familia exponencial, $\mu_t = E(Y_t | x_t)$ y ϕ_t es un parámetro de escala, con $g(\mu_t) = x_t'\beta$, y β un vector de coeficientes. Suponga ahora que el vector de coeficientes β cambia con el tiempo, i.e.,

Observación: $Y_t | x_t = f\left(y_t \middle| g^{-1}(x_t'\beta_t), \phi_t\right)$,

Evolución: $\beta_t = G_t\beta_{t-1} + \omega_t$, $\omega_t \sim N\left(0, W_t^{-1}\right)$,

para t=1,2,..., donde W_t^{-1} es una precisión y G_t es una matriz.

- En el modelo anterior, lo único que cambia con respecto al modelo dinámico lineal previamente visto, es la ecuación de observación. La ecuación de evolución no cambia, sigue siendo lineal–normal.

- ANÁLISIS DEL MODELO: El análisis Bayesiano requiere de especificar una distribución inicial sobre los parámetros desconocidos del modelo, que en este caso son el valor inicial de las β's, $\beta_0' = (\beta_{00}, \beta_{01}, \ldots, \beta_{0,p-1})$ (y en algunos casos ϕ_t y W_t^{-1}).

  Recordemos que D_t es la información disponible al tiempo t, tal que:

  D_0 es la información inicial y

  $D_t = D_{t-1} \cup \{y_t\}$

  Al igual que en el caso anterior, la forma de realizar la actualización es de manera recursiva mediante el *Teorema de Bayes secuencial*, i.e.,

Inicial (al tiempo t): $f(\theta_t \mid D_{t-1}) = \int f(\theta_t \mid \theta_{t-1}) f(\theta_{t-1} \mid D_{t-1}) d\theta_{t-1}$
Final (al tiempo t): $f(\theta_t \mid D_t) \propto f(y_t \mid \theta_t) f(\theta_t \mid D_{t-1})$
Predictiva (a un paso): $f(y_t \mid D_{t-1}) = \int f(y_t \mid \theta_t) f(\theta_t \mid D_{t-1}) d\theta_t$

- Sin embargo, estas distribuciones inicial, final y predictiva, por lo general no tienen una forma analítica estándar y dependen de la familia exponencial elegida y de la distribución inicial para β_0 (y ϕ_t y W_t^{-1}).

- Las distribuciones iniciales, al tiempo t=0, comúnmente usadas son:

$$\beta_0 \sim N(0, C_0),$$

$$\phi_t \sim Ga(a_0, b_0),$$

Aunque por lo general es común suponer que $\phi_t = \phi$ para todo t, y

$$W_t^{-1} \sim \text{Wishart}(p, Q_0, \nu).$$

- *Nota 1*: Vale la pena aclarar que la forma analítica de las distribuciones inicial, final y predictiva no es indispensable, ya que el análisis se puede hacer a través de simulación en un paquete Bayesiano como OpenBugs.

- *Nota 2*. Ligado a la nota anterior, debido a la flexibilidad para realizar análisis de este tipo de modelos con OpenBugs, se puede pensar incluso en modelos más generales como modelos de regresión de escala dinámicos, o algo aún más complejo.

- EJERCICIO 6 (continuación…). Consideremos el ejercicio 6 de los mineros. Realmente el número total de desastres (Y_t) es una serie de tiempo para una variable con soporte en los enteros positivos. Entonces una tercer alternativa de modelado es:

  c) Modelo Poisson dinámico:

  Observación: $Y_t | \mu_t \sim Po(\mu_t)$

  Evolución: $\mu_t = \mu_{t-1} + \omega_t$, $\omega_t \sim N(0, W^{-1}) \Rightarrow \mu_t | \mu_{t-1} \sim N(\mu_{t-1}, W^{-1})$

  con $\mu_1 \sim N(0,0.0001)$ y $W^{-1} \sim Ga(0.01,0.01)$.