

CAPÍTULO 11. MODELOS DINÁMICOS LINEALES

Para leer

Harrison y Stevens (1976), West y Harrison (1999), Pole et al (1994).

Se consideran variables dinámicas cuyos valores cambian con el tiempo. Se supone que los valores de la serie observada son relacionados con unos parámetros desconocidos que también cambian con el tiempo.

Definición 14 *El modelo lineal, dinámico o MLD, general (normal) es*

$$\begin{aligned} Y_t &= \mathbf{F}_t^T \boldsymbol{\theta}_t + \nu_t & \nu_t &\sim \mathcal{N}(\mathbf{0}, \mathbf{V}_t) \\ \boldsymbol{\theta}_t &= \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \omega_t & \omega_t &\sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t) \end{aligned}$$

Se representa el modelo como el conjunto de cuádruples $\{\mathbf{F}, \mathbf{G}, \mathbf{V}, \mathbf{W}\}_t$ que, en principio, se supone conocido.

La primera ecuación está relacionada con las observaciones y la segunda describe el sistema.

Este modelo está relacionado con muchos modelos conocidos, por ejemplo

- $\{1, 1, V, 0\}$ (independiente de t) es el modelo de datos (i.i.d.) normales univariadas.

$$Y_t = \theta + \nu_t \Rightarrow Y_t | \theta \sim \mathcal{N}(\theta, V)$$

- $\{1, 1, 0, W\}$ es un paseo aleatorio.
- $\{\mathbf{F}_t, \mathbf{I}, \mathbf{V}, \mathbf{0}\}$ es el modelo de regresión multivariante normal:

$$\mathbf{Y}_t = \mathbf{F}_t \boldsymbol{\theta} + \boldsymbol{\nu}_t$$

Definición 15 *Cuando la pareja*

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

es independiente de t , se dice que el modelo es un MLD series temporal (time series DLM).

Si además $\{\mathbf{V}, \mathbf{W}\}_t = \{\mathbf{V}, \mathbf{W}\}$ se dice que el MLD es constante.

El MLD constante es definido por el cuadruple $\{\mathbf{F}, \mathbf{G}, \mathbf{V}, \mathbf{W}\}$. La clase de MLDs constantes incluye la mayoría de los modelos clásicos para series temporales.

Consideramos unos modelos específicos. Para hacer el análisis, necesitamos un resultado para la distribución normal.

Un teorema sobre la distribución normal

Teorema 16 *Sea \mathbf{X}_1 y \mathbf{X}_2 distribuidas con una distribución normal multivariante*

$$\begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}; \begin{pmatrix} \boldsymbol{\sigma}_{11} & \boldsymbol{\sigma}_{12} \\ \boldsymbol{\sigma}_{21} & \boldsymbol{\sigma}_{22} \end{pmatrix} \right)$$

Entonces, dada $\mathbf{X}_2 = \mathbf{x}_2$, la distribución condicional de \mathbf{X}_1 es

$$\mathbf{X}_1 | \mathbf{x}_2 \sim \mathcal{N}(\boldsymbol{\mu}_1^*, \boldsymbol{\sigma}_{11}^*) \quad \text{donde}$$

$$\begin{aligned} \boldsymbol{\mu}_1^* &= \boldsymbol{\mu}_1 + \boldsymbol{\sigma}_{12} \boldsymbol{\sigma}_{22}^{-1} (\mathbf{x}_2 - \boldsymbol{\mu}_2) \\ \boldsymbol{\sigma}_{11}^* &= \boldsymbol{\sigma}_{11} - \boldsymbol{\sigma}_{12} \boldsymbol{\sigma}_{22}^{-1} \boldsymbol{\sigma}_{21} \end{aligned}$$

El modelo constante, cerrado $\{1, 1, V, W\}$

En este caso,

$$\begin{aligned}Y_t &= \theta_t + \nu_t \\ \theta_t &= \theta_{t-1} + \omega_t\end{aligned}$$

Se utiliza el modelo para representar series temporales sin tendencia.

Supongamos que a tiempo $t-1$, nuestros conocimientos (datos previos etcétera) son D_{t-1} y que nuestra distribución para θ_{t-1} es

$$\theta_{t-1}|D_{t-1} \sim \mathcal{N}(m_{t-1}, C_{t-1})$$

Entonces, a través de las ecuaciones del MLD, se tiene

$\theta_t|D_{t-1} \sim \mathcal{N}(m_{t-1}, R_t)$ donde $R_t = C_{t-1} + W$ es la distribución a priori para $\theta_t|D_{t-1}$.

$Y_t|D_{t-1} \sim \mathcal{N}(m_{t-1}, Q_t)$ donde $Q_t = R_t + V$
 es la distribución predictiva de Y_t .

$$\begin{pmatrix} Y_t \\ \theta_t \end{pmatrix} \sim \mathcal{N} \left(\begin{matrix} m_{t-1} \\ m_{t-1} \end{matrix}; \begin{pmatrix} Q_t & R_t \\ R_t & R_t \end{pmatrix} \right)$$

es la distribución conjunta.

Dada la información $D_t = \{D_{t-1}, y_t\}$, se tiene

$$\theta_t|D_t \sim \mathcal{N}(m_t, C_t) \quad \text{donde}$$

$$m_t = m_{t-1} + A_t(y_t - m_{t-1})$$

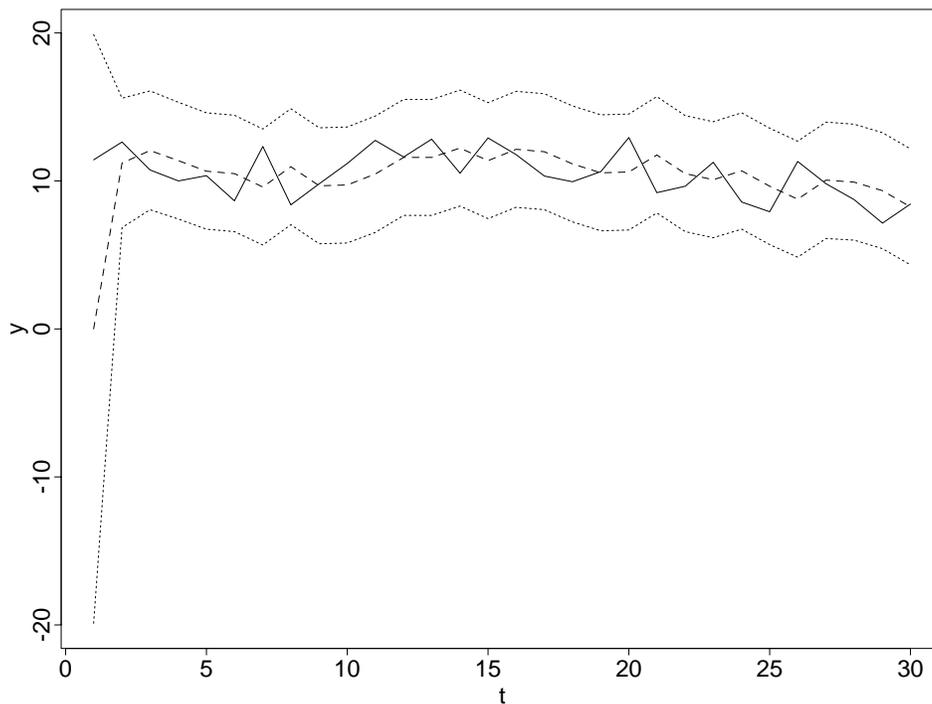
$$C_t = A_t V$$

donde $A_t = \frac{R_t}{Q_t} = \frac{(C_{t-1} + W)}{C_{t-1} + W + V}$.

La media a posteriori es una media ponderada de la media a priori y y_t , con pesos proporcionales a V y $C_{t-1} + W$.

Observación 93 Las ecuaciones utilizadas para hallar la distribución a posteriori son esencialmente equivalentes al filtro de Kalman (1960,1963).

Ejemplo 116 Se generaron 30 datos del modelo constante, cerrado con $V = 2$ y $W = 1$ y $\theta_0 = 10$. Se utilizó el mismo modelo para predecir empezando con la distribución a priori $\theta_0 \sim \mathcal{N}(0, 100)$.



Se ve que la anchura de los intervalos predictivos se acerca a (aproximadamente) 8.

Convergencia cuando $t \rightarrow \infty$

Teorema 17 Cuando $t \rightarrow \infty$, entonces $A_t \rightarrow A$ donde

$$A = \frac{r}{2} \left(\sqrt{1 + \frac{4}{r}} - 1 \right)$$

y $r = \frac{W}{V}$ es la razón de señal a ruido (signal to noise ratio).

Luego $C_t \rightarrow C = AV$, $R_t \rightarrow R = \frac{AV}{1-A}$, $Q_t \rightarrow Q = \frac{V}{1-A}$ y $W = A^2Q$.

Demostración parcial

Suponiendo que un límite existe, se tiene

$$\begin{aligned}C &= AV \\ &= \frac{C + W}{C + W + V}V \\ C^2 + CW - VW &= 0 \\ C &= \frac{-W + \sqrt{W^2 + 4VW}}{2} \\ &= \frac{-rV + \sqrt{r^2V^2 + 4rV^2}}{2} \\ &= \frac{rV}{2} \left(\sqrt{1 + \frac{4}{r}} - 1 \right) \Rightarrow \\ A &= \frac{r}{2} \left(\sqrt{1 + \frac{4}{r}} - 1 \right)\end{aligned}$$

Para completar la demostración, es suficiente ver que C_t está limitada y monótona. Ver West y Harrison (1999).

Ejemplo 117 En el Ejemplo 116 se tiene $r = \frac{W}{V} = \frac{1}{2}$ y entonces,

$$A_t \rightarrow \frac{1}{4} (\sqrt{1+8} - 1) = \frac{1}{2}.$$

Luego $C_t \rightarrow AV = 1$ y la varianza predictiva de $Y_t|D_{t-1}$ es

$$Q_t \rightarrow \frac{V}{1-A} = \frac{2}{1/2} = 4.$$

Entonces, la anchura de un intervalo predictiva de 95 % se aproxima a $2 \times 1,96 \times \sqrt{4} \approx 8$.

Selección de $W(r)$

Supongamos que los verdaderos valores de V y r son V_0 y r_0 . Si elegimos $r < r_0$, subestimaremos mucho la varianza predictiva Q en el límite y por lo general es mejor sobreestimar esta varianza.

El método habitual es fijar el factor de descuento (*discount factor*) $\delta = 1 - A$ en algún valor alto como $A > 0,8$ y de allí sacar el valor de r , y luego W .

Ver Ameen y Harrison (1985) y West y Harrison (1999) para más detalles.

La función predictiva

Se pueden calcular las distribuciones predictivas varias etapas hacía adelante. En el modelo constante, cerrado:

$$\begin{aligned} Y_{t+k} &= \theta_{t+k} + \nu_{t+k} \\ &= \theta_{t+k-1} + \omega_{t+k} + \nu_{t+k} \\ &= \dots \\ &= \theta_t + \sum_{i=1}^k \omega_{t+i} + \nu_{t+k} \end{aligned}$$

Entonces, $Y_{t+k}|D_t \sim \mathcal{N}(m_t, C_t + kW + V)$.

Para un MLD general, se define

$$f_t(k) = E[Y_{t+k}|D_t]$$

es la función predictiva de k etapas (*k step ahead forecast function*).

Relación del modelo $\{1, 1, V, W\}$ con algunos modelos clásicos

Cuando $t \rightarrow \infty$, el modelo está relacionada con el modelo clásico ARIMA (Box y Jenkins 1976).

Sea $Z_t = Y_t - Y_{t-1}$. Entonces

$$Z_t = \nu_t + \omega_t - \nu_{t-1}$$

y se puede representar $\{Z_t\}$ como un modelo ARIMA(0,1,1).

$$Z_t = a_t - (1 - A)a_{t-1}$$

donde $a_t \sim \mathcal{N}(0, V/(1 - A))$ son variables i.i.d.

También los modelos EWMA (Holt 1957) y EWR (Brown 1962) dan predicciones similares a las predicciones bayesianas asintóticas. Ver West y Harrison (1999).

Varianzas desconocidas

Es poco probable que V y W sean conocidas. Existen dos posibilidades de modelización.

1 $W = rV$ y r es conocido.

Sea

$$\begin{aligned} V^{-1}|D_{t-1} &\sim \mathcal{G}\left(\frac{a_{t-1}}{2}, \frac{b_{t-1}}{2}\right) \quad \text{y} \\ \theta_{t-1}|D_{t-1}, V &\sim \mathcal{N}(m_{t-1}, C_{t-1}V). \end{aligned}$$

Entonces, se tiene un modelo normal gamma

$$\theta_{t-1}, V^{-1}|D_{t-1} \sim \mathcal{NG}\left(m_{t-1}, C_{t-1}^{-1}, \frac{a_{t-1}}{2}, \frac{b_{t-1}}{2}\right)$$

y mediante la ecuación del sistema,

$$\theta_t, V^{-1}|D_{t-1} \sim \mathcal{NG}\left(m_{t-1}, (C_{t-1} + r)^{-1}, \frac{a_{t-1}}{2}, \frac{b_{t-1}}{2}\right)$$

La distribución a posteriori

La ecuación para la observación implica que $Y_t|\theta_t, V \sim \mathcal{N}(\theta_t, V)$ y ahora, mediante el teorema 5, se tiene

$$\theta_t, V^{-1}|D_t = D_{t-1}, y_t \sim \mathcal{NG}\left(m_t, C_t^{-1}, \frac{a_t}{2}, \frac{b_t}{2}\right)$$

donde

$$\begin{aligned} m_t &= \frac{\frac{m_{t-1}}{C_{t-1}+r} + y_t}{\frac{1}{C_{t-1}+r} + 1} \\ &= \frac{m_{t-1} + (C_{t-1} + r)y_t}{C_{t-1} + r + 1} \\ &= m_{t-1} + A_t(y_t - m_{t-1}) \quad \text{donde} \end{aligned}$$

$$A_t = \frac{C_{t-1} + r}{C_{t-1} + r + 1}$$

$$C_t^{-1} = (C_{t-1} + r)^{-1} + 1$$

$$C_t = \frac{C_{t-1} + r}{C_{t-1} + r + 1}$$

$$a_t = a_{t-1} + 1$$

$$b_t = b_{t-1} + \frac{1}{C_{t-1} + r + 1}(y_t - m_{t-1})^2$$

Observación 94 *A menudo, se elige el valor de r mediante el factor de descuento, fijando el valor de $\delta = 1 - \lim_{t \rightarrow \infty} A_t$.*

- 2 W y V ambos desconocidos. Se puede usar un algoritmo Gibbs porque se conocen todas las distribuciones condicionales a posteriori.

Modelos de regresión dinámica

$\{\mathbf{F}, \mathbf{I}, V, \mathbf{W}\}$ es el MLD constante de regresión

$$\begin{aligned} Y_t &= \mathbf{F}^T \boldsymbol{\theta} + \nu_t \\ \boldsymbol{\theta}_t &= \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t \end{aligned}$$

Dada la información a priori $\boldsymbol{\theta}_{t-1} \sim \mathcal{N}(\mathbf{m}_{t-1}, \mathbf{C}_{t-1})$ se tiene

- $\boldsymbol{\theta}_t | D_{t-1} \sim \mathcal{N}(\mathbf{m}_{t-1}, \mathbf{C}_{t-1} + \mathbf{W})$

- $Y_t | D_{t-1} \sim \mathcal{N}(\mathbf{F}^T \mathbf{m}_{t-1}, Q_t)$ donde

$$Q_t = \mathbf{F}^T (\mathbf{C}_{t-1} + \mathbf{W}) \mathbf{F} + V$$

- $Cov[\boldsymbol{\theta}_t | D_{t-1}, Y_t | D_{t-1}] = \mathbf{F}^T (\mathbf{C}_{t-1} + \mathbf{W})$

- $\theta_t | D_t \sim \mathcal{N}(\mathbf{m}_t, \mathbf{C}_t)$ donde

$$\mathbf{m}_t = \mathbf{m}_{t-1} + (\mathbf{C}_{t-1} + \mathbf{W})\mathbf{F}Q_t^{-1}(y_t - \mathbf{F}^T\mathbf{m}_{t-1})$$

$$\mathbf{C}_t = \mathbf{C}_{t-1} - (\mathbf{C}_{t-1} + \mathbf{W})\mathbf{F}Q_t^{-1}\mathbf{F}^T(\mathbf{C}_{t-1} + \mathbf{W})$$

Ejemplo 118 (*West y Harrison 1999*)

Producción total (Y_t) de leche y números de vacas (F_t).

<i>año</i>	1970	71	72	73	74	75	
<i>leche</i>	117,0	118,6	120,0	115,5	115,6	115,4	
<i>vacas</i>	12,0	11,8	11,7	11,4	11,2	11,1	
<i>año</i>	76	77	78	79	80	81	82
<i>leche</i>	120,2	122,7	121,5	123,4	128,5	130,0	135,8
<i>vacas</i>	11,0	11,0	10,8	10,7	10,8	10,9	11,0

Ajustamos el modelo dinámico de regresión con $V = 1$ y $W = 0,05$.



El modelo general

Supongamos que $\boldsymbol{\theta}_{t-1}|D_{t-1} \sim \mathcal{N}(\mathbf{m}_{t-1}, \mathbf{C}_{t-1})$.
Luego tenemos las ecuaciones de aprendizaje:

- $\boldsymbol{\theta}_t|D_{t-1} \sim \mathcal{N}(\mathbf{a}_t, \mathbf{R}_t)$ donde $\mathbf{a}_t = \mathbf{G}_t\mathbf{m}_{t-1}$ y

$$\mathbf{R}_t = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t^T + \mathbf{W}_t.$$

- $\mathbf{Y}_t|D_{t-1} \sim \mathcal{N}(\mathbf{f}_t, \mathbf{Q}_t)$ donde $\mathbf{f}_t = \mathbf{F}_t^T\mathbf{a}_t$ y

$$\mathbf{Q}_t = \mathbf{F}_t^T\mathbf{R}_t\mathbf{F}_t + \mathbf{V}_t.$$

- $\text{Cov}[\boldsymbol{\theta}_t, \mathbf{Y}_t|D_{t-1}] = \mathbf{F}_t^T\mathbf{R}_t.$

- $\boldsymbol{\theta}_t|D_t \sim \mathcal{N}(\mathbf{m}_t, \mathbf{C}_t)$ donde

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t(\mathbf{y}_t - \mathbf{f}_t)$$

$$\mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t\mathbf{Q}_t\mathbf{A}_t^T$$

$$\text{y } \mathbf{A}_t = \mathbf{R}_t\mathbf{F}_t\mathbf{Q}_t^{-1}.$$

Para las situaciones en las que Y_t es univariable, definimos unos modelos para los rasgos típicos de series temporales.

Tendencia

Un MLD con función predictiva de forma

$$f_t(k) = \sum_{i=0}^{n-1} a_{ti} k^i$$

se llama un MLD polinómico de orden n . Modelos así son parecidos al modelo $\{\mathbf{E}_n, \mathbf{L}_n, \cdot, \cdot\}$ donde

$$\mathbf{E}_n = (1, 0, \dots, 0)^T \quad \text{y} \quad \mathbf{L}_n = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Estacionalidad

Decimos que una función $g(t)$ es cíclica si

$$g(t + np) = g(t).$$

Los factores de $g(\cdot)$ son los valores de un ciclo $\psi_j = g(j)$ por $j = 1, \dots, p$.

Por estación t definimos el vector de factores estacionales

$$\psi_t = (\psi_j, \dots, \psi_p, \psi_1, \dots, \psi_{p-1})$$

donde el factor de estación t es ψ_j .

Un modelo estacional sin forma es $\{\mathbf{E}_p, \mathbf{P}, V_t, \mathbf{W}_t\}$ donde $\mathbf{E}_p = (1, 0, \dots, 0)$ y

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{I}_{p-1} \\ 1 & \mathbf{0}^T \end{pmatrix}$$

con la ecuación de observación

$$Y_t = \mathbf{E}_p^T \psi_t + \nu_t$$

y la ecuación del sistema

$$\psi_t = \mathbf{P}\psi_{t-1} + \omega_t$$

Notas

- A menudo, se modifica la definición para incluir la restricción $\sum \psi_j = 0$.
- Se pueden representar estos modelos en la forma Fourier. Ver West y Harrison (1999).

Superposición de Modelos

A menudo una serie contiene rasgos distintos, por ejemplo, una tendencia, una parte estacional etcétera.

Se modelizan series así usando el principio de superposición.

Supongamos dos modelos para Y_{1t} y Y_{2t} :

$$M_1 : \{ \mathbf{F}_1, \mathbf{G}_1, V_1, \mathbf{W}_1 \}$$

$$M_2 : \{ \mathbf{F}_2, \mathbf{G}_2, V_2, \mathbf{W}_2 \}$$

Entonces, si $Z_t = Y_{1t} + Y_{2t}$, Z_t cumple con el modelo $\{ \mathbf{F}, \mathbf{G}, V, \mathbf{W} \}$ donde $\mathbf{F} = \begin{pmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \end{pmatrix}$,

$$\mathbf{G} = \text{block diag}(\mathbf{G}_1, \mathbf{G}_2)$$

$$V = V_1 + V_2 \text{ y}$$

$$\mathbf{W} = \text{block diag}(\mathbf{W}_1, \mathbf{W}_2).$$

Luego modelizamos independientemente los rasgos distintos.

Nota: la función predictiva $f_t(k)$ de Z_{t+k} será

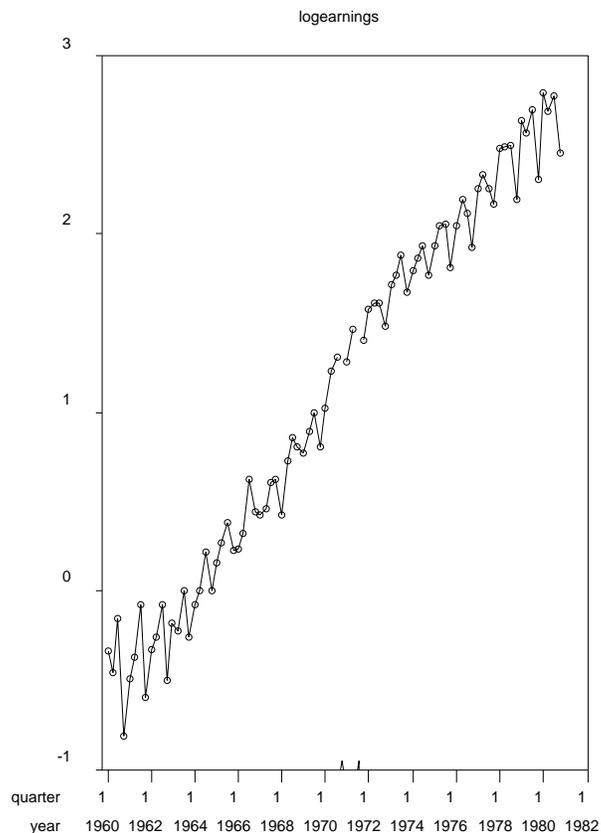
$$\begin{aligned} f_t(k) &= E[Z_{t+k}|D_t] \\ &= E[Y_{1,t+k} + Y_{2,t+k}|D_t] \\ &= f_{1t}(k) + f_{2t}(k) \end{aligned}$$

donde $f_{1t}(k)$ y $f_{2t}(k)$ son las funciones predictivas de cada parte.

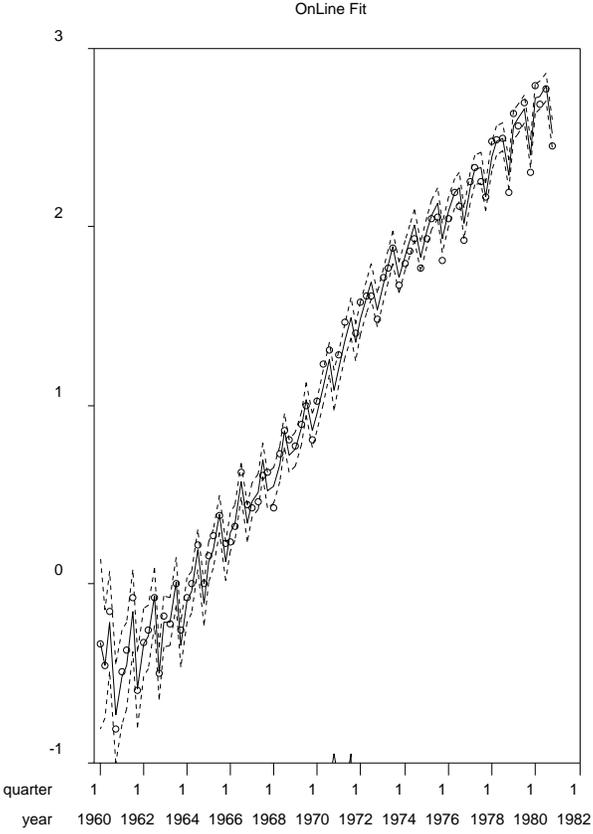
El siguiente ejemplo está hecho mediante el software *BATS*. Ver Pole et al (1994) y

<http://www.stat.duke.edu/~mw/bats.html>

Ejemplo 119 *Los datos (Shumway y Stoffer 2000) son el logaritmo de los ganancias de activos de Johnson & Johnson entre 1960 y 1982.*



Ajustamos estos datos con un modelo incluyendo una tendencia lineal y un efecto estacional.



El siguiente dibujo muestra los efectos debidos a la tendencia y a la estacionalidad, tanto como los residuos.

Estacionariedad (Stationarity)

Podemos usar el concepto clásico de estacionariedad para comprobar el modelo.

Definición 16 Decimos que un proceso $\{Z_t\}$ es estacionario normal si

- Por $\mathbf{Z}_{t,n} = (Z_t, \dots, Z_{t+n})^T$, $E[\mathbf{Z}_{t,n}]$ y $V[\mathbf{Z}_{t,n}]$ son independientes de t y
- $\mathbf{Z}_{t,n} \sim \mathcal{N}(\cdot, \cdot)$

Ejemplo 120 Supongamos que $a_t \sim \mathcal{N}(0, \sigma^2)$ y definimos $Z_t = 3a_t - a_{t-1}$.

Entonces $E[Z_t] = 0$ y

$$\text{Cov}[Z_t, Z_{t+k}] = \begin{cases} 10\sigma^2 & \text{si } k = 0 \\ -3\sigma^2 & \text{si } k = \pm 1 \\ 0 & \text{si no} \end{cases}$$

Entonces $\{Z_t\}$ es estacionario normal.

Definición 17 Sea $Z_{t,n}$ estacionario normal (con media μ). Se define la autocovarianza de retraso k es

$$\gamma_k = \text{Cov}[Z_t, Z_{t+k}].$$

La autocorrelación de retraso k es

$$\rho_k = \text{Corr}[Z_t, Z_{t+k}].$$

La función generadora de autocovarianzas es

$$\Gamma(\beta) = \sum_{k=-\infty}^{\infty} \gamma_k \beta^k,$$

La función generadora de correlaciones es

$$R(\beta) = \sum_{k=-\infty}^{\infty} \rho_k \beta^k.$$

Escribimos $Z \sim \mathcal{SN}(\mu, \Gamma(\beta))$

Ejemplo 121 Volviendo al Ejemplo 120 tenemos $Z \sim \mathcal{SN}(\mu, \Gamma(\beta))$ donde $\Gamma(\beta) = (-3\beta^{-1} + 10 - 3\beta)\sigma^2$.

Estacionariedad y comprobación del modelo

Definición 18 *Por la serie $\{Z_t\}$ definimos el operador de retardo B :*

$$BZ_t = Z_{t-1} \quad \text{y} \quad B^k Z_t = Z_{t-k}$$

Tenemos el siguiente teorema.

Teorema 18 *Dado el MLD cerrado y constante $\{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\}$, supongamos que \mathbf{G} de dimensión $r \times r$ tiene r eigenvalores λ_i , $i = 1, \dots, r$ donde $|\lambda_i| \geq 1 \forall i$. Se define*

$$Z_t = \phi(B)Y_{t+r}$$

donde B es el operador de retardo y

$$\phi(B) = \prod_{i=1}^r (1 - \lambda_i B)$$

Entonces $Z_t \sim \mathcal{SN}(0, \Gamma(\beta))$ donde es fácil hallar la función de autocovarianza.

Podemos utilizar el teorema para comprobar la suficiencia de un MLD.

Ejemplo 122 Consideramos el modelo $\{1, 1, V, W\}$. Aquí $G = 1$ y tiene un eigenvalor igual a 1. Entonces definimos $Z_t = (1 - B)Y_t = Y_t - Y_{t-1}$. Entonces

$$\begin{aligned} Z_t &= \theta_t + \nu_t - \theta_{t-1} - \nu_{t-1} \\ &= \theta_{t-1} + \omega_t + \nu_t - \theta_{t-1} - \nu_{t-1} \\ &= \omega_t + \nu_t - \nu_{t-1} \end{aligned}$$

Luego, dadas las varianzas V y W , se tiene $Z_t|V, W \sim \mathcal{SN}(0, \Gamma(\beta))$ donde

$$\Gamma(\beta) = -V\beta^{-1} + (W + 2V) - V\beta$$

Dada una muestra $\mathbf{Z} = (Z_1, \dots, Z_n)$, luego,

$$\bar{Z}|V, W \sim \mathcal{N}\left(0, \frac{2V + nW}{n^2}\right).$$

Dados los datos $\mathbf{z} = (z_1, \dots, z_n)$ es posible comprobar el modelo.

Ejemplo 123 *Supongamos que tenemos los datos de la serie $\{Y_t\}$*

117 118,6 120 115,5 115,6 115,4 120,2
122,7 121,5 123,4 128,5 130,0 135,8

y queremos decidir si el modelo $\{1, 1, 1, 0,05\}$ será razonable.

Calculamos las diferencias

$$z_t = (1,6, 1,4, -4,5, ,1, -,2, 4,8, 2,5, -1,2, 1,9, 5,1, 1,5, 5,8)$$

Ahora la media de los datos es $\bar{z} = 1,507$ y si el modelo es el verdadero,

$$\bar{Z} \sim \mathcal{N}\left(0, \frac{2 \times 1 + 12 * 0,05}{12^2}\right) \approx \mathcal{N}(0, ,17)$$

la probabilidad de generar una media tan larga de esta distribución es minúscula. Entonces, el modelo no parece adecuado.

Además, se han desarrollado métodos para comparación de modelos. Ver Frühwith Schnatter (1995).

Más recientemente

- ARMA y ARIMA bayesiana. Marriot y Newbold (1998).
- Modelos de espacio de estado no lineales y no normales. Carlin et al (1992).
- Estructura latente. Aguilar et al (1998).
- Volatilidad estocástica. Jacquier et al (1994).
- GARCH. Bauwens y Lubrano (1998).
- Series econométricas. Bauwens et al (2000).
- Wavelets. Nason et al (1999).

Y mucho, mucho más.