

# 1 Estimación recursiva.

Um modelo de regresión multiple:

$$\begin{aligned} y_t &= x_t\beta + u_t \\ x_t &= \begin{bmatrix} 1 & x_{2t} & x_{3t} & \cdots & x_{kt} \end{bmatrix} & k \times 1 \\ \beta' &= \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \cdots & \beta_k \end{bmatrix} \end{aligned} \quad (1)$$

puede ser estimado de forma recursiva utilizando:

$$\hat{\beta}_t = (\mathbf{X}_t' \mathbf{X}_t)^{-1} \mathbf{X}_t' \mathbf{y}_t \quad (2)$$

donde  $\mathbf{X}_t$  es la matriz de regresores de orden  $t \times k$  que incluye información de las variables explicativas hasta el periodo  $t$ . De forma equivalente  $\mathbf{y}_t$  es un vector de orden  $t \times 1$  que contiene la información de la variable dependiente  $y_t$  hasta el periodo  $t$ . Por lo tanto,  $\hat{\beta}_t$  es el estimador MCO que recoge información de las variables del modelo hasta el periodo  $t$ .

La ventaja de la estimación recursiva es que no es necesario estimar el modelo cada vez que hay una nueva

observación. Es decir, si disponemos de la estimación  $\hat{\beta}_t$  es decir la estimación de los parametros del modelo (1) con información hasta el periodo  $t$  y se actualiza nuestra información y pasamos a disponer de la información correspondiente al periodo  $t + 1$  ( es decir conocemos  $y_{t+1}, x_{2t+1}, \dots, x_{kt+1}$ ), no es necesario aplicar  $\hat{\beta}_{t+1} = (\mathbf{X}'_{t+1}\mathbf{X}_{t+1})^{-1} \mathbf{X}'_{t+1}\mathbf{y}_{t+1}$  y para obtener la estimación actualizada de  $\beta$  con información hasta  $t + 1$ . sino que basado en el siguiente resultado (ver pdf de A. Sansó):

$$\begin{aligned}
 (\mathbf{X}'_{t+1}\mathbf{X}_{t+1})^{-1} &= (\mathbf{X}'_t\mathbf{X}_t)^{-1} - \frac{(\mathbf{X}'_t\mathbf{X}_t)^{-1} x'_{t+1}x_{t+1} (\mathbf{X}'_t\mathbf{X}_t)^{-1}}{1 + x_{t+1} (\mathbf{X}'_t\mathbf{X}_t)^{-1} x'_{t+1}} \\
 x_{t+1} &= \begin{bmatrix} 1 & x_{2t+1} & x_{3t+1} & \cdots & x_{kt+1} \end{bmatrix} \\
 \mathbf{P}_{t+1} &= (\mathbf{X}'_{t+1}\mathbf{X}_{t+1})^{-1} \\
 \mathbf{P}_{t+1} &= \mathbf{P}_t - \frac{\mathbf{P}_t x'_{t+1} x_{t+1} \mathbf{P}_t}{1 + x_{t+1} \mathbf{P}_t x'_{t+1}} \\
 \mathbf{P}_{t+1} &= \mathbf{P}_t - g_{t+1} x_{t+1} \mathbf{P}_t \\
 g_{t+1} &= \mathbf{P}_t x'_{t+1} \left(1 + x_{t+1} \mathbf{P}_t x'_{t+1}\right)^{-1}.
 \end{aligned}$$

La estimación recursiva, de los estimadores se puede obtener a partir de (ver pdf A. Sansó):

$$\begin{aligned}
 \hat{\beta}_{t+1} &= \hat{\beta}_t + \frac{(\mathbf{X}'_t \mathbf{X}_t)^{-1} x'_{t+1}}{1 + x_{t+1} (\mathbf{X}'_t \mathbf{X}_t)^{-1} x'_{t+1}} (y_{t+1} - x_{t+1} \hat{\beta}_t) \\
 &= \hat{\beta}_t + g_{t+1} (y_{t+1} - x_{t+1} \hat{\beta}_t) \\
 &= \hat{\beta}_t + g_{t+1} e_{t+1|t}.
 \end{aligned}$$

Entonces llegamos al siguiente sistema de ecuaciones que se aplica cada vez que se dispone una nueva observación:

$$e_{t+1|t} = y_{t+1} - x_{t+1} \hat{\beta}_t \quad (3)$$

$$f_t = (1 + x_{t+1} \mathbf{P}_t x'_{t+1}) \quad (4)$$

$$\begin{aligned}
 g_{t+1} &= \mathbf{P}_t x'_{t+1} (1 + x_{t+1} \mathbf{P}_t x'_{t+1})^{-1} \\
 &= \mathbf{P}_t x'_{t+1} f_t^{-1}
 \end{aligned} \quad (5)$$

$$\hat{\beta}_{t+1} = \hat{\beta}_t + g_{t+1} e_{t+1|t} \quad (6)$$

$$\begin{aligned}
 \mathbf{P}_{t+1} &= \mathbf{P}_t - g_{t+1} x_{t+1} \mathbf{P}_t \\
 &= (I - g_{t+1} x_{t+1}) \mathbf{P}_t.
 \end{aligned} \quad (7)$$

Donde  $e_{t+1|t}$  (3) es el error de predicción cuando estimamos el modelo con información hasta el periodo  $t$ . A

$g_{t+1}$  (5) la llamamos ganancia del filtro, y permite actualizar la información correspondiente a las variables explicativas. Una vez que disponemos de  $g_{t+1}$  podemos actualizar la estimación de los parámetros del modelo  $\hat{\beta}_{t+1}$  mediante (6) y finalmente actualizamos la estimación de  $\mathbf{P}_{t+1}$ . Cada vez que se dispone de la información correspondiente a un nuevo periodo se actualiza la estimación siguiendo el algoritmo (3)-(7).

**Error de predicción:**  $e_{t+1|t} = y_{t+1} - x_{t+1}\hat{\beta}_t$

**Dispersión de  $e_{t+1|t}$ :**  $f_t = (1 + x_{t+1}\mathbf{P}_t x'_{t+1})$

**Ganancia del filtro:**  $g_{t+1} = \mathbf{P}_t x'_{t+1} f_t^{-1}$

**Estimación  $\beta$ :**  $\hat{\beta}_{t+1} = \hat{\beta}_t + g_{t+1} e_{t+1|t}$

**Dispersión de  $\hat{\beta}_{t+1}$ :**  $\mathbf{P}_{t+1} = (I - g_{t+1} x_{t+1}) \mathbf{P}_t$

En el filtro de Kalman existe un sistema de ecuaciones muy similar al descrito anteriormente tal y como veremos a continuación.

## 2 Filtro del Kalman y modelos lineal en el espacio de los estados

El modelo lineal de espacio de los estados (lineal state space model) responde a:

$$y_t = Z\alpha_t + \varepsilon_t \quad \varepsilon_t \sim N(0, H) \quad (8)$$

$$\alpha_t = T\alpha_{t-1} + R\eta_t \quad \eta_t \sim N(0, Q) \quad (9)$$

$$t = 1, 2, \dots, T$$

$$E[\varepsilon_t \eta_t'] = 0$$

$$E[\alpha_0 \eta_t'] = 0$$

$$E[\alpha_0 \varepsilon_t'] = 0$$

donde  $y_t$  puede ser un vector de orden  $p \times 1$ , que recoge todas las variables relevantes que son observables. Viene determinada mediante la ecuación (8), que recibe el nombre de **ecuación de medida**. El vector  $\alpha_t$  de orden  $m \times 1$

recoge las variables estado (state variables) que son un conjunto de variables no observables o latentes que determinan el comportamiento de las variables  $y_t$ . La matriz  $Z$  es de orden  $p \times m$  y finalmente  $\varepsilon_t$  es un vector de innovaciones de la ecuación de medida (8) con valor esperado zero y matriz de varianzas y covarianzas  $H$  de orden  $p \times p$ . También tenemos una **ecuación de transición** (9) que recoge la evolución temporal del vector  $\alpha_t$  de variables estado que adopta un comportamiento autorregresivo, donde las matrices  $T$  y  $R$  de orden  $m \times m$  y  $m \times g$ , finalmente las innovaciones de la ecuación de transición vienen recogidas en el vector  $\eta_t$  de orden  $g \times 1$  y tiene valor esperado zero y matrix de varianzas y covarianzas  $Q$  de orden  $g \times g$ .

La version más general del modelo lineal de espacio de los estados permite que las matrices del sistema  $Z$ ,  $T$  y  $R$  y las matrices de varianzas y covarianzas puedan ir variando a lo largo del tiempo. La versión más sencilla

de (8)-(9) se dá cuando  $y_t$  recoge una única variable y entonces tenemos:

$$y_t = Z\alpha_t + \varepsilon_t \quad \varepsilon_t \sim N(0, h) \quad (10)$$

$$\alpha_t = T\alpha_{t-1} + R\eta_t \quad \eta_t \sim N(0, Q) \quad (11)$$

$$t = 1, 2, \dots, T$$

entonces tanto  $y_t$  como  $\varepsilon_t$  pasan a ser escalares ( $1 \times 1$ ) y tenemos que  $VAR[\varepsilon_t] = h$ . El comportamiento de las innovaciones puede quedar resumido en:

$$\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} h & 0 \\ 0 & Q \end{bmatrix} \right).$$

A partir de la ecuación de transición (11) podemos ver que la predicción del vector estado  $\alpha_t$   $h$  periodos hacia adelante basada en la información hasta el periodo  $t$  responde a:

$$\begin{aligned} \hat{\alpha}_t(h) &= E[\alpha_{t+h} | \alpha_j, j \leq t] = a_{t+h|t} = Ta_{t+h-1|t} = \\ &= T^h a_t \end{aligned}$$

entonces:

$$\begin{aligned}\hat{y}_t(h) &= E[y_{t+h}|y_j, j \leq t] = y_{t+h|t} = \\ &= HT^h a_t \\ a_t &= E[\alpha_t|\alpha_j, j \leq t].\end{aligned}$$

De aquí se desprende que la calidad de las predicciones que se puedan obtener para  $y_{t+h}$  con información disponible hasta el periodo  $t$   $\hat{y}_t(h) = y_{t+h|t}$  dependerán de la estimación que realizemos del vector estado  $\hat{\alpha}_t$ . Esta estimación del vector estado resume toda la información de la que disponemos hasta el momento  $t$ . Si sustituimos recursivamente en (11):

$$\begin{aligned}\alpha_t &= T^t \alpha_0 + \sum_{j=0}^{t-1} T^j R \eta_{t-j} \\ &= T^t \alpha_0 + R \eta_t + T^j R \eta_{t-j} + \dots + T^{t-1} R \eta_1.\end{aligned}$$

Vemos que el vector estado  $\alpha_0$  para el periodo inicial  $t = 0$  juega un papel importante en como se distribuirá el vector estado  $\alpha_t$  en un periodo cualquiera  $t$ . Para determinar la distribución de  $\alpha_t$  se supone que la distribución del vector estado inicial es conocida y responde



a:

$$\alpha_0 \sim N(a_0, P_0) \quad (12)$$

A esto lo llamaremos información inicial  $\mathcal{I}_0 = \{a_0, P_0\}$ . Y a la información disponible hasta el periodo  $t$  la llamamos  $\mathcal{I}_t = \{y_t, y_{t-1}, \dots, y_1, \mathcal{I}_0\}$ .

El filtro de Kalman está formado por siguiente sistema de ecuaciones:

**Predicción del vector estado:**

$$E[\alpha_{t+1}|\mathcal{I}_t] = a_{t+1|t} = Ta_t \quad (13)$$

**Dispersión de  $a_{t+1|t}$  :**

$$VAR[a_{t+1|t}] = P_{t+1|t} = TP_tT' + RQR' \quad (14)$$

**Error de predicción:**

$$e_{t+1|t} = y_{t+1} - Za_{t+1|t} \quad (15)$$

**Dispersión de  $e_{t+1|t}$  :**

$$VAR [e_{t+1|t}] = f_t = Z P_{t+1|t} Z' + h \quad (16)$$

**Ganancia de KALMAN:**

$$G_{t+1} = P_{t+1|t} Z' f_t^{-1} \quad (17)$$

**Estimación del vector estado:**

$$E [\alpha_{t+1} | \mathcal{I}_{t+1}] = a_{t+1} = a_{t+1|t} + G_{t+1} e_{t+1|t} \quad (18)$$

**Dispersión de  $a_{t+1}$  :**

$$P_{t+1} = (I - G_{t+1} Z) P_{t+1|t}. \quad (19)$$

El anterior sistema se puede deducir siguiendo el siguiente razonamiento. Asumimos que la distribución del vector

estado  $a_t$  condicionada a la información disponible hasta el periodo  $t$  ( $\mathcal{I}_t$ ) responde a:

$$(\alpha_t|\mathcal{I}_t) \sim N(a_t, P_t) \quad (20)$$

de donde se puede deducir que:

$$\begin{aligned} E[\alpha_{t+1}|\mathcal{I}_t] &= a_{t+1|t} = Ta_t \\ E[y_{t+1}|\mathcal{I}_t] &= Za_{t+1|t} = ZTa_t. \end{aligned}$$

En cuanto a  $VAR[\alpha_{t+1}|\mathcal{I}_t]$ , teniendo en cuenta la ecuación de transición (11) podemos deducir que:

$$\begin{aligned} VAR[\alpha_{t+1}|\mathcal{I}_t] &= VAR[T\alpha_t + R\eta_{t+1}|\mathcal{I}_t] = \\ &= TVAR[\alpha_t|\mathcal{I}_t]T' + RVAR[\eta_{t+1}]R' \\ &= TP_tT' + RQR' \end{aligned}$$

dado que  $\alpha_t$  depende  $\eta_{t-j} \forall j \geq 0$  pero no de  $\eta_{t+1}$  y podemos escribir:

$$(\alpha_{t+1}|\mathcal{I}_t) \sim N(Ta_t, TP_tT' + RQR').$$

Con lo que tenemos las dos primeras ecuaciones del filtro de Kalman (13) y (14).

Al disponer de la información correspondiente al periodo  $t + 1$  ( $\mathcal{I}_{t+1}$ ) podemos calcular el error de predicción (15) como:

$$\begin{aligned}
 e_{t+1} &= y_{t+1} - y_{t+1|t} & (21) \\
 &= y_{t+1} - Z a_{t+1|t} \\
 &= y_{t+1} - Z T a_t \\
 &= Z \alpha_{t+1} + \varepsilon_{t+1} - Z T a_t \\
 &= Z (\alpha_{t+1} - T a_t) + \varepsilon_{t+1}
 \end{aligned}$$

Conocer  $y_{t+1}$  es equivalente a conocer el error  $e_{t+1|t}$  un periodo hacia adelante. Por lo tanto la distribución condicionada de  $(\alpha_{t+1}|\mathcal{I}_{t+1})$  es equivalente a  $(\alpha_{t+1}|y_{t+1}, \mathcal{I}_t)$  y a  $(\alpha_{t+1}|e_{t+1|t}, \mathcal{I}_t)$  que permite obtener las ecuaciones (18) y (19) del filtro del Kalman. Para poder hacer esto obtendremos la distribución conjunta de  $\alpha_{t+1}$  y  $e_{t+1|t}$  condicionada a  $\mathcal{I}_t$   $(\alpha_{t+1}, e_{t+1|t}|\mathcal{I}_t)$  y a partir de un resultado de la distribución normal multivariante obtendremos la distribución de  $(\alpha_{t+1}|e_{t+1|t}, \mathcal{I}_t)$  y la  $VAR[e_{t+1|t}]$  que son los únicos resultados del filtro del Kalman que no faltan por deducir.

La distribución de  $e_{t+1}$  condicionada a  $\mathcal{I}_{t+1}$  ( $e_{t+1}|\mathcal{I}_{t+1}$ ) será igual a  $(e_{t+1}|y_{t+1}, \mathcal{I}_t) = (e_{t+1}|\alpha_{t+1}, \mathcal{I}_t)$  y teniendo en cuenta (21) tendremos:

$$(e_{t+1}|\alpha_{t+1}, \mathcal{I}_t) \sim N(Z(\alpha_{t+1} - Ta_t), h).$$

Si tenemos dos vectores aleatorios  $\mathbf{X}_1$  y  $\mathbf{X}_2$  tales que:

$$P(\mathbf{X}_1, \mathbf{X}_2) = P(\mathbf{X}_2) P(\mathbf{X}_1|\mathbf{X}_2)$$

su distribución conjunta multivariante responde a:

$$\begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)$$

si:

$$\mathbf{X}_2 \sim N(\boldsymbol{\mu}_2, \Sigma_{22}) \quad (22)$$

$$(\mathbf{X}_1|\mathbf{X}_2) \sim N\left(\boldsymbol{\mu}_1 + \Sigma_{12} \Sigma_{22}^{-1} (\mathbf{X}_2 - \boldsymbol{\mu}_2), \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}\right) \quad (23)$$

Si asignamos  $\alpha_{t+1}$  a  $\mathbf{X}_2$  y  $e_{t+1}$  a  $\mathbf{X}_1$ . Pasamos a tener:

$$(\alpha_{t+1}|\mathcal{I}_t) \sim N(Ta_t, TP_tT' + RQR')$$

$$\boldsymbol{\mu}_2 = Ta_t$$

$$\Sigma_{22} = TP_tT' + RQR' = P_{t+1|t}.$$

Y en el caso de:

$$\begin{aligned}(e_{t+1}|\alpha_{t+1}, \mathcal{I}_t) &\sim N(Z(\alpha_{t+1} - Ta_t), h) \\ Z(\alpha_{t+1} - Ta_t) &= \mu_1 + \sum_{12} \sum_{22}^{-1} (\mathbf{X}_2 - \mu_2) \\ h &= \sum_{11} - \sum_{12} \sum_{22}^{-1} \sum_{21}.\end{aligned}$$

Como:

$$\begin{aligned}Z(\alpha_{t+1} - Ta_t) &= \mu_1 + \sum_{12} \sum_{22}^{-1} (\mathbf{X}_2 - \mu_2) \\ Z(\alpha_{t+1} - Ta_t) &= \mu_1 + \sum_{12} P_{t+1|t}^{-1} (\alpha_{t+1} - Ta_t) \\ \text{Por lo que :}\end{aligned}$$

$$\begin{aligned}\mu_1 &= 0 \\ \sum_{12} &= ZP_{t+1|t}.\end{aligned}$$

También tendremos:

$$\begin{aligned}h &= \sum_{11} - \sum_{12} \sum_{22}^{-1} \sum_{21} \\ h &= \sum_{11} - ZP_{t+1|t} P_{t+1|t}^{-1} P_{t+1|t} Z' \\ h &= \sum_{11} - ZP_{t+1|t} Z' \\ \sum_{11} &= h + ZP_{t+1|t} Z' .\end{aligned}$$

Entonces obtenemos:

$$\begin{bmatrix} \alpha_{t+1} \\ e_{t+1} \end{bmatrix} \Big| \mathcal{I}_t \sim N \left( \begin{bmatrix} T a_t \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} P_{t+1|t} & P_{t+1|t} Z' \\ Z P_{t+1|t} & h + Z P_{t+1|t} Z \end{bmatrix} \right).$$

Y finalmente, llegamos a los resultados requeridos:

$$(\alpha_{t+1} | e_{t+1|t}, \mathcal{I}_t) \sim N(a_{t+1}, P_{t+1})$$

$$\begin{aligned} a_{t+1} &= a_{t+1|t} + P_{t+1|t} Z' (h + Z P_{t+1|t} Z)^{-1} e_{t+1|t} \\ P_{t+1} &= P_{t+1|t} - P_{t+1|t} Z' (h + Z P_{t+1|t} Z)^{-1} Z P_{t+1|t}. \end{aligned}$$

El filtro de Kalman asume que los valores que el valor esperado y varianza del vector estado inicial son conocidos ( $\alpha_0 \sim N(a_0, P_0)$ ), es decir, que conocemos  $a_0$  y  $P_0$ . Los valores iniciales de  $a_0$  y  $P_0$  en algunos casos se obtienen de la experiencia previa, por ejemplo a partir de estudios anteriores. Otra práctica habitual es utilizar el criterio asumir que  $a_0 = 0$  y que  $P_0$  es una matrix diagonal con valores muy grandes en su diagonal principal, indicando nuestro desconocimiento a priori sobre

el problema. Esta solución es válida para tamaños muestrales elevados en los que la información muestral acabara dominando a los valores iniciales y siempre que trabajemos con series estacionarias. Para series temporales no estacionarias el enfoque anterior ha sido criticado mucho y la práctica habitual es utilizar el enfoque de Durbin Koopman llamado "Exact initia Kalman filter".

En filtro de Kalman puede ser utilizado en gran variedad de situaciones. Nosotros veremos los modelos estructurales de series temporales, los modelos ARIMA y los metodos de alisado exponencial ode Holt-Winters. Pero, por ejemplo puede ser utilizado para trabajar con modelos de regresión en los que los párametros son cambiantes como:

$$y_t = X_t \beta_t + \varepsilon_t \quad \varepsilon_t \sim N(0, h) \quad (24)$$

$$\beta_t = T \beta_{t-1} + R \eta_t \quad \eta_t \sim N(0, Q) \quad (25)$$

$$t = 1, 2, \dots, T$$

Donde en vez de tener una matriz  $Z$  constante en la ecuación de medida tenemos la matriz  $X_t$  que recoge



los regresores o variables explicativas. Y en la ecuación de medida  $\beta_t$  tenemos un vector de parámetros que van evolucionado a lo largo del tiempo de acuerdo a la ecuación de transición. Es decir, tratamos a los parámetros del modelo como el vector estado.

### **3 Modelos estructurales de Series temporales**

Los modelos estructurales de series temporales, son modelos en los que se determina cual el PGD que siguen los componentes no observables de la serie temporal. Llamamos componentes no observables de una serie temporales a la tendencia  $T_t$ , ciclo  $C_t$  y estacionalidad  $S_t$ . Se les llama así por pueden ser observados en la evolución temporal de la serie temporal, pero no podemos determinar para cada datos de la serie que parte corresponde a

cada componente. En este enfoque se supone que una serie  $y_t$  se puede expresar como:

$$y_t = T_t + C_t + S_t + I_t$$

donde  $I_t$  es el componente irregular.

Tradicionalmente los economistas han considerado como componentes de interés el tendencia y el ciclo. Y los componentes estacional e irregular como provocados por factores exógenos e independientes a los factores económicos.

- Tendencia y tendencia ciclo: El comportamiento tendencial de las series temporales está asociado a oscilaciones con período de repetición superior a 5 años  $5S$  periodos (60 meses). Oscilaciones asociadas a frecuencias comprendidas entre la frecuencia 0 (de período de repetición infinito) y  $2\pi/5S$ . Por lo que respecta a al componente cíclico se identifica con aquellas oscilaciones asociadas a períodos de repetición comprendidos entre 5 años  $5S$  periodos (60

meses) y un año  $S$  (12 meses), es decir asociadas a las frecuencias comprendidas entre  $2\pi/5S$  y  $2\pi/S$ . El límite entre componente tendencial y cíclico es arbitrario, por lo que en la práctica se suele hablar de componente tendencia-ciclo, que se identifica con oscilaciones que se repiten con un período de repetición infinito y un año  $S$  periodos (12 meses), entre la frecuencia 0 y  $2\pi/S$ .

- Componente estacional: Se identifica con las oscilaciones de carácter periódico o cuasi-periódico de duración anual o inferior a la anual, estando asociadas a las frecuencias  $\omega_j = 2\pi j/S$  con  $j = 1, 2, \dots, [S/2]$  (es decir, con período de repetición de  $\lambda_j = S/j$  con  $j = 1, 2, \dots, [S/2]$ ) y a las frecuencias adyacentes a las anteriores.
- Componente irregular: Se identifica a oscilaciones no asociadas a los componentes tendencia-ciclo y estacional.

**Local level trend** El modelo estructural mas sencillo es aquel en el que sólo tenemos componente tendencial y se le llama Local level trend:

$$\begin{aligned}
 y_t &= T_t + \varepsilon_t \\
 T_t &= T_{t-1} + \beta_{t-1} + e_t \\
 \beta_t &= \beta_{t-1} + v_t \\
 \varepsilon_t &\sim iid(0, \sigma_\varepsilon^2) \\
 e_t &\sim iid(0, \sigma_e^2) \\
 v_t &\sim iid(0, \sigma_v^2).
 \end{aligned} \tag{26}$$

Cuando  $\sigma_v^2 = 0$ ,  $\beta_t$  pasa ser una constante  $\beta$ , y entonces obtenemos:

$$\begin{aligned}
 T_t &= \beta + T_{t-1} + e_t \\
 e_t &\sim iid(0, \sigma_e^2).
 \end{aligned}$$

Y si  $\beta = 0$  se reduce a un simple paseo aleatorio:

$$\begin{aligned} T_t &= T_{t-1} + e_t \\ e_t &\sim iid(0, \sigma_e^2). \end{aligned}$$

Finalmente es facil ver que (26) también puede escribirse como:

$$\begin{aligned} \Delta T_t &= \beta_{t-1} + e_t \\ \Delta \beta_t &= v_t \\ \Delta T_t &= \frac{v_{t-1}}{\Delta} + e_t \\ \Delta^2 T_t &= v_{t-1} + (1 - L) e_t. \end{aligned}$$

Es decir, un proceso IMA(2,1) en el que el proceso MA(1) es no invertible.

El anterior modelo recibe el nombre de "Local Level Trend" y se puede escribir en forma de espacio de los estados como:

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix}^Z \begin{bmatrix} T_t \\ \beta_t \end{bmatrix} + \varepsilon_t$$

$$\begin{bmatrix} T_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} T_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^R \begin{bmatrix} e_t \\ v_t \end{bmatrix}$$

En cuanto a las matrices de varianzas y covarianzas tenemos:

$$h = \sigma_\varepsilon^2$$

$$Q = \begin{bmatrix} \sigma_e^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}.$$

**Estacionalidad** Existen dos posibles especificaciones para el componente estacional el primero responde:

$$\begin{aligned}
S_t &= - \sum_{j=1}^{S-1} S_{t-j} + w_t \\
w_t &\sim iid(0, \sigma_w^2).
\end{aligned} \tag{27}$$

Que también puede expresarse como:

$$\begin{aligned}
S_t + S_{t-1} + \cdots + S_{t-s+1} &= w_t \\
(1 + L + \cdots + L^{s-1}) S_t &= w_t \\
S(L) S_t &= w_t.
\end{aligned}$$

También existe la especificación trigonométrica:

$$\begin{aligned}
S_t &= \sum_{j=1}^{[s/2]} S_{jt} \\
\left. \begin{aligned} S_{jt} &= S_{j,t-1} \cos(\omega_j) + S_{j,t-1}^* \sin(\omega_j) + w_{jt} \\ S_{jt}^* &= -S_{j,t-1} \sin(\omega_j) + S_{j,t-1}^* \cos(\omega_j) + w_{jt}^* \end{aligned} \right\} \\
\omega_j &= 2\pi j/s \quad j = 1, 2, \dots, [s/2]
\end{aligned}$$

$w_{jt}$  y  $w_{jt}^*$  sob procesos ruido blanco incorrelacionados uno con otro y con varianza común  $\sigma_w^2$  para  $j = 1, 2, \dots, [s/2]$ . El término  $S_{jt}^*$  aparece en la expresión por motivos de construcción y su interpretación puede ser ignorada.

**Modelo estructural básico** El modelo estructural básico con el primer modelo estacional, responde a:

$$\begin{aligned}x_t &= T_t + S_t + \varepsilon_t \\T_t &= T_{t-1} + \beta_{t-1} + e_t \\\beta_t &= \beta_{t-1} + v_t \\S_t &= - \sum_{j=1}^{s-1} S_{t-j} + w_t.\end{aligned}$$

Se puede comprobar que responde al siguiente modelo ARIMA:



$$x_t = T_t + S_t + \varepsilon_t$$

$$T_t = \frac{v_{t-1}}{\Delta^2} + \frac{e_t}{\Delta}$$

$$S_t = \frac{w_t}{S(L)}$$

$$x_t = \frac{v_{t-1}}{\Delta^2} + \frac{e_t}{\Delta} + \frac{w_t}{S(L)} + \varepsilon_t$$

$$\Delta_S \Delta x_t = S(L)v_{t-1} + \Delta^s e_t + \Delta^2 w_t + \Delta^s \Delta \varepsilon_t.$$

$$\Delta_S = (1 - L^S) = \Delta S(L)$$

Que es un modelo muy parecido al modelo de lineas aéreas  $\Delta_S \Delta x_t = (1 - \Theta L^S) (1 - \theta L) \varepsilon_t$ . Y que puede

ser escrito en forma de espacio de los estados:

$$y_t = \begin{bmatrix} 1 & 0 & 1 & 0 & \dots & 0 & 0 \end{bmatrix} \begin{bmatrix} T_t \\ \beta_t \\ S_t \\ S_{t-1} \\ \vdots \\ S_{t-S+2} \end{bmatrix} + \varepsilon_t$$

$$\begin{bmatrix} T_t \\ \beta_t \\ S_t \\ S_{t-1} \\ \vdots \\ S_{t-S+2} \end{bmatrix} = \text{diag}[T_T, T_S] \begin{bmatrix} T_{t-1} \\ \beta_{t-1} \\ S_{t-1} \\ S_{t-2} \\ \vdots \\ S_{t-S+1} \end{bmatrix} + R \begin{bmatrix} e_t \\ v_t \\ w_t \end{bmatrix}$$

$$T_T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$T_S = \begin{bmatrix} -1 & -1 & -1 & \dots & -1 & -1 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix}$$

En cuanto a las matrices de varianzas y covarianzas tenemos:

$$h = \sigma_{\varepsilon}^2$$
$$Q = \begin{bmatrix} \sigma_e^2 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_w^2 \end{bmatrix}.$$

En el caso de tener la estacionalidad trigonométrica el modelo se puede escribir en el espacio de los estados como:

$$\begin{aligned}
y_t &= Z\alpha_t + \varepsilon_t \\
con &: \\
Z &= \begin{cases} \begin{bmatrix} 1 & 0 & 1 & 0 & \cdots & 1 & 0 & 1 & 0 \end{bmatrix} & t \text{ impar} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & \cdots & 1 & 0 & 1 \end{bmatrix} & t \text{ par} \end{cases} \\
\alpha'_t &= \begin{cases} \begin{bmatrix} T_t & \beta_t & S_{1t} & S_{1t}^* & \cdots \\ \cdots & S_{(S-1)/2,t} & S_{(S-1)/2,t}^* \end{bmatrix} & t \text{ impar} \\ \begin{bmatrix} T_t & \beta_t & S_{1t} & S_{1t}^* & \cdots \\ \cdots & S_{[S/2]-1,t} & S_{[S/2]-1,t}^* & S_{S/2t} \end{bmatrix} & t \text{ par} \end{cases} \\
\alpha_t &= diag [T_T, T_S] \alpha_{t-1} + R \begin{bmatrix} e_t \\ v_t \\ w_{1t} \\ w_{1t}^* \\ \vdots \\ \vdots \end{bmatrix} \\
T_T &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \\
T_S &= \begin{cases} diag \begin{bmatrix} C_1 & C_2 & C_3 & \cdots & C_{(S-1)/2} \end{bmatrix} & t \text{ impar} \\ diag \begin{bmatrix} C_1 & C_2 & C_3 & \cdots & C_{(S-1)/2} & -1 \end{bmatrix} & t \text{ par} \end{cases} \\
C_j &= \begin{bmatrix} \cos(\omega_j) & \sin(\omega_j) \\ -\sin(\omega_j) & \cos(\omega_j) \end{bmatrix} \\
R &= I_{2+S-1}.
\end{aligned}$$

En cuanto a las matrices de varianzas y covarianzas tenemos:

$$\begin{aligned} h &= \sigma_{\varepsilon}^2 \\ Q &= \text{diag} [\sigma_e^2, \sigma_v^2, \sigma_w^2 I] . \end{aligned}$$

En los modelos estructurales de series temporales lo que se estima son las varianzas de las innovaciones. Si las varianzas de las innovaciones son iguales a cero, el proceso se vuelve determinista.

La estimación se hace mediante maxima verosimilitud y consiste en maximizar:

$$\ln L = -\frac{T}{2} \ln (\sigma^2) - \frac{1}{2} \sum_{t=1}^T \ln (f_t) - \frac{1}{2\sigma^2} \sum_{t=1}^T \frac{e_{t+1|t}^2}{f_t} \quad (28)$$