Apuntes de suavizamiento exponencial

Rubén Miranda F.

Valparaíso, 07 de Octubre de 2008

APUNTES DE SERIES DE TIEMPO

Suavizamiento exponencial

Los modelos clásicos (o ingenuos) tienen la desventaja que no se adaptan a lo largo del tiempo.

Hace algunas décadas apareció una clase de modelos que satisface la adaptabilidad del tiempo y son los de **suavizamiento exponencial**.

Estos modelos consideran que el pasado reciente sea más importante que el pasado remoto.

Notación:

$\widehat{X}_n(1)$ estimación de X_{n+1}.
$\widehat{X}_n(h)$ estimación h pasos adelante basados en la historia hasta el instante n.
Se puede determinar $\widehat{X}_n(1)$ de la siguiente forma:

$$\widehat{X}_n(1) = w_1 X_n + w_2 X_{n-1} + \cdots + w_{n-1} X_2 + w_n X_1 \tag{1}$$

Siendo w_i los pesos para cada observación (1) debe satisfacer lo siguiente

- w_i decrece cuando t aumenta
- $\sum_{t=1}^{\infty} w_t = 1$

Objetivos: Predecir la serie en valores futuros y suavizar la serie (eliminar el ruido). Se considerarán los métodos de suavizamiento exponencial simple, doble, triple y de Holt-Winter

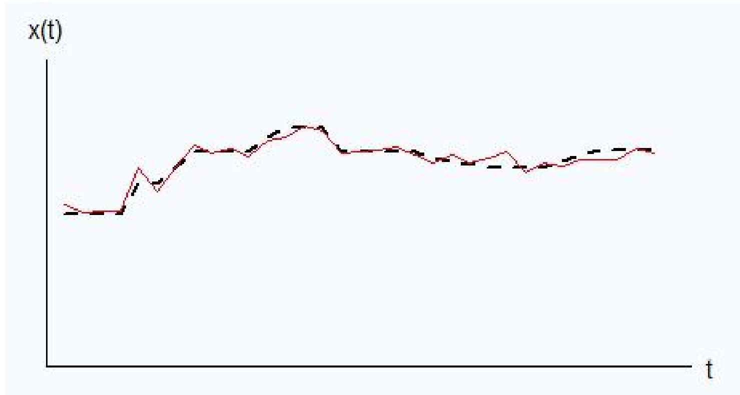
§ Suavizamiento exponencial simple

Supuestos: La serie de tiempo es localmente constante, es decir, no tiene clara tendencia ni estacionalidad y sus valores oscilan en torno a una constante que varía levemente en el tiempo, es decir $X_t = u_t + a_t, \;\; t = 1 \ldots n$ donde u_t es un parámetro que varía lentamente en el tiempo y a_t es ruido blanco.

En este caso $w_i = \alpha(1-\alpha)^i$ satisface las condiciones propuestas en (1). La ecuación queda de la siguiente manera:

$$\widehat{X}_n(1) = \alpha X_n + \alpha(1-\alpha) X_{n-1} + \alpha(1-\alpha)^2 X_{n-2} + \cdots \quad 0 < \alpha < 1 \tag{2}$$

Esto se ilustra en la siguiente gráfica



Para predecir a un paso contando con la información anterior, se procede de la siguiente manera:

$$\widehat{X}_n(1) = \alpha X(n) + (1-\alpha)\widehat{X}_{n-1}(1) \tag{3}$$

Esto requiere iteración a un paso, para ello se utiliza como semilla $\widehat{X}_1(1) = X(1)$

¿Cómo elegir α?. Se considera de (3) $\widehat{X}_n(1) = \underbrace{\alpha[X_n - \widehat{X}_{n-1}(1)] + \widehat{X}_{n-1}(1)}_{\alpha\widehat{\epsilon}_n + \widehat{X}_{n-1}(1)}$.

Donde $\widehat{\epsilon}_n$ es el error de predicción. Para solucionar el mejor α, se debe resolver el problema de $min_\alpha\left(\sum_{i=1}^{t}\widehat{\epsilon}_t^{\,2}\right)$, donde $\widehat{\epsilon}_n$ depende de α.

Predicción h pasos adelante basados en la información hasta el instante n

$$\widehat{X}_n(h) = \alpha X(n) + (1-\alpha)\widehat{X}_{n-1}(1) \tag{4}$$

Desventaja: Las predicciones h pasos adelante dependen únicamente de la información $t = n$. Algunos autores sugieren la siguiente modificación:

$$\widehat{X}_n(h) = \alpha X(n) + (1-\alpha)\widehat{X}_n(h-1) \tag{5}$$

§ **Suavizamiento exponencial doble**

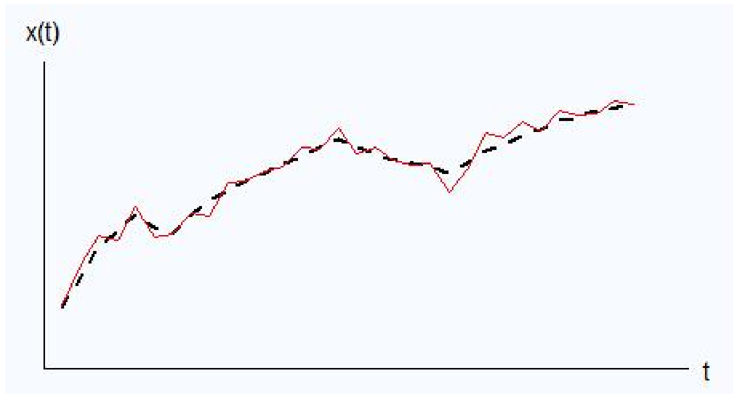
Supuestos: Conocer la situación en que el nivel medio de la serie cambia en el tiempo de forma lineal, es decir $X_t = u_t + T_t + a_t,\ \ t = 1\dots n$, donde $T_t = \beta_0 + \beta_1 t$.

La idea es compensar el suavizado $\overline{X}_t$ utilizando un nuevo suavizado $\overline{\overline{X}}_t$ a la serie X_t, puesto que el primero no logrará eliminar la componente de tendencia.

Los suavizados son de la siguiente forma:

$$
\begin{aligned}
\overline{X}_t &= \alpha X_t + (1-\alpha)\overline{X}_{t-1} \\
\overline{\overline{X}}_t &= \alpha \overline{X}_t + (1-\alpha)\overline{\overline{X}}_{t-1}
\end{aligned} \tag{6}
$$

Esto se muestra en la siguiente figura:



Puesto que la tendencia es localmente lineal, se usan los valores obtenidos en $\overline{X}_t$ y $\overline{\overline{X}}_t$ para estimar la pendiente y el intercepto en cada instante.

$$
\begin{aligned}
\beta_0 &= 2\overline{X}_t - \overline{\overline{X}}_t \\
\beta_1 &= \tfrac{\alpha}{1-\alpha}(\overline{X}_t - \overline{\overline{X}}_t)
\end{aligned} \tag{7}
$$

Predicción h pasos adelante basados en la información hasta el instante n.

$$
\widehat{X}_n(h) = \left(2 + \frac{\alpha h}{1-\alpha}\right)\overline{X}_n - \left(1 + \frac{\alpha h}{1-\alpha}\right)\overline{\overline{X}}_n \tag{8}
$$

Obervaciones: Para calcular la constante se procede de manera similar al suavizamiento exponencial simple, es decir, se minimiza el error cuadrático medio.

El suavizamiento exponencial doble se inicializa de la forma $X_1 = \overline{X}_1 = \overline{\overline{X}}_1$

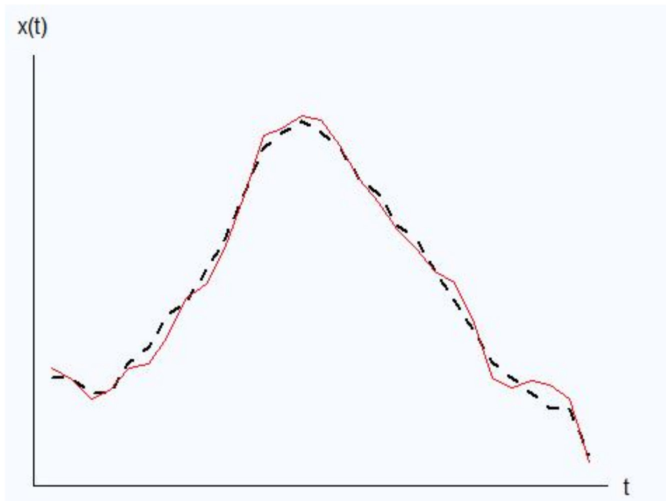
§ **Suavizamiento exponencial triple**

Supuestos: Se concidera la situación en que el nivel medio de la serie cambia en el tiempo de forma cuadrática, es decir $X_t = u_t + T_t + a_t,\;\; t = 1\ldots n$, donde $T_t = \beta_0 + \beta_1 t + \beta_2 t^2$ y estimar mediante mínimos cuadrados.

Se utilizan tres suavizados y son los siguientes:

$$\begin{aligned}\overline{X}_t &= \alpha X_t + (1-\alpha)\overline{X}_{t-1}\\ \overline{\overline{X}}_t &= \alpha \overline{X}_t + (1-\alpha)\overline{\overline{X}}_{t-1}\\ \overline{\overline{\overline{X}}}_t &= \alpha \overline{\overline{X}}_t + (1-\alpha)\overline{\overline{\overline{X}}}_{t-1}\end{aligned} \tag{9}$$

Esto se muestra en la siguiente figura:



Predicción h pasos adelante basados en la información hasta el instante n.

$$\begin{aligned}&X_n(h) = \beta_0 + \beta_1 h + \beta_2 h^2 \quad h \geq 1, donde\\ &\beta_0 = 3\overline{X}_n - 3\overline{\overline{X}}_n + \overline{\overline{\overline{X}}}_n\\ &\beta_1 = \tfrac{\alpha}{2(1-\alpha)^2}((6-5\alpha)\overline{X}_n - 2(5-4\alpha)\overline{\overline{X}}_n + (4-3\alpha)\overline{\overline{\overline{X}}}_n)\\ &\beta_2 = \tfrac{\alpha^2}{2(1-\alpha)^2}(\overline{X}_n - 2\overline{\overline{X}}_n + \overline{\overline{\overline{X}}}_n)\end{aligned} \tag{10}$$

Obervaciones: Para calcular la constante se procede de manera similar al suavizamiento exponencial simple, es decir, se minimiza el error cuadrático medio.

El suavizamiento exponencial triple se inicializa de la forma $X_1 = \overline{X}_1 = \overline{\overline{X}}_1 = \overline{\overline{\overline{X}}}_n$

§ **Suavizamiento de Holt-Winters (Caso no estacionario)**

Este método supone que la serie de tiempo X_t se comporta localmente como una suma de un nivel y una tendencia central

Se anota $\overline{x}_t$ y m_t las estimaciones del nivel y de la pensiente de la pendiente de la recta en el instante t, entonces se propone:

$$\begin{array}{ll} \overline{x}_t = \alpha x_t + (1-\alpha)[\overline{x}_{t-1} + m_{t-1}] & 0 < \alpha < 1 \\ m_t = \beta(\overline{x}_t - \overline{x}_{t-1}) + (1-\beta) m_{t-1} & 0 < \beta < 1 \end{array} \tag{11}$$

Obervaciones: Para calcular la constante se procede de manera similar al suavizamiento exponencial, es decir, se minimiza el error cuadrático medio de acuerdo a los valores de α y β.

El suavizamiento de Holt-Winters (no períodico) se inicializa de la forma

$$m_2 = x_2 - x_1 \qquad \overline{x}_2 = x_2 \qquad \text{se itera para } x = 3, 4 \ldots, n.$$

§ **Suavizamiento de Holt-Winters (Caso estacionario)**

Las suposiciones de este método son las mismas que en el caso no estacional, pero se agrega un período que llamaremos s. Este período puede ser multiplicativo o aditivo respecto a la tendencia,

$\overline{x}_t$ se interpreta como un nivel descentralizado y $\widehat{E}_t$ la estimación del factor estacional en el instante t.

Las ecuaciones en el caso multiplicativo son

$$\begin{array}{ll} \overline{x}_t = \alpha(\frac{x_t}{\widehat{E}_{t-s}}) + (1-\alpha)[\overline{x}_{t-1} + m_{t-1}] & 0 < \alpha < 1 \\ m_t = \beta(\overline{x}_t - \overline{x}_{t-1}) + (1-\beta) m_{t-1} & 0 < \beta < 1 \\ \overline{E}_t = \delta\left(\frac{x_t}{\overline{x}_t}\right) + (1-\delta)\widehat{E}_{t-s} & 0 < \delta < 1 \end{array} \tag{12}$$

Y en el caso aditivo

$$\begin{array}{ll} \overline{x}_t = \alpha(x_t - \widehat{E}_{t-s}) + (1-\alpha)[\overline{x}_{t-1} + m_{t-1}] & 0 < \alpha < 1 \\ m_t = \beta(\overline{x}_t - \overline{x}_{t-1}) + (1-\beta) m_{t-1} & 0 < \beta < 1 \\ \widehat{E}_t = \delta(x_t - \overline{x}_t) + (1-\delta)\widehat{E}_{t-s} & 0 < \delta < 1 \end{array} \tag{13}$$

Predicción en el caso multiplicativo h pasos adelante basados en la información hasta el instante n

$$\begin{array}{ll} \widehat{X}_n(h) = (\overline{X}_n + h m_n)\widehat{E}_{n+h-s} & h = 1, 2, \ldots, s \\ \widehat{X}_n(h) = (\overline{X}_n + h m_n)\widehat{E}_{n+h-2s} & h = s+1, s+2, \ldots, 2s \\ \vdots & \end{array} \tag{14}$$

Predicción en el caso aditivo h pasos adelante basados en la información hasta el instante n

$$\begin{array}{ll} \widehat{X}_n(h) = (\overline{X}_n + hm_n) + \widehat{E}_{n+h-s} & h = 1, 2, \ldots, s \\ \widehat{X}_n(h) = (\overline{X}_n + hm_n) + \widehat{E}_{n+h-2s} & h = s+1, s+2, \ldots, 2s \\ \vdots & \end{array} \tag{15}$$

Obervaciones: Para calcular la constante se procede de manera similar al suavizamiento exponencial simple, es decir, se minimiza el error cuadrático medio donde $\epsilon_j^2 = (x_j - \widehat{X}_{j-1}(1))^2 = x_j - [m_{j-1} + \overline{x}_{j-1}]\widehat{E}_{j-s}$ para el caso multiplicativo y $\epsilon_j^2 = (x_j - \widehat{X}_{j-1}(1))^2 = x_j - [m_{j-1} + \overline{x}_{j-1} + \widehat{E}_{j-s}]$ en el caso aditivo.

El suavizamiento de Holt-Winters para el caso períodico debe iniciarse en

$$m_s = 0 \qquad \overline{x}_s = (1/s)\textstyle\sum_{k=1}^{s} x_k$$

$$\begin{array}{ll} \widehat{E}_j = \frac{x_j}{(1/s)\sum_{k=1}^{s} x_k} & h = 1, 2, \ldots, s \text{ en el caso multiplicativo} \\ \widehat{E}_j = x_j - (1/s)\sum_{k=1}^{s} x_k & h = 1, 2, \ldots, s \text{ en el caso aditivo} \end{array} \tag{16}$$

Con estas condiciones inciales se puede calcular $\overline{x}_t \quad m_t \quad \widehat{E}_t \quad \forall t = s+1, s+2, \ldots, n$