Análisis de Series de Tiempo

Camilo Enrique Argoty Pulido

Especialización en Inteligencia Artificial

25 de marzo de 2025

Modelos de series estacionarias

El proceso lineal general

Sea X_t una serie de tiempo observada, y ϵ_t una serie no observable de tiempo de ruido blanco, es decir, una serie de v.a.'s iid, de media cero.

El proceso lineal general

Sea X_t una serie de tiempo observada, y ϵ_t una serie no observable de tiempo de ruido blanco, es decir, una serie de v.a.'s iid, de media cero.

Se dice que X_t es un **proceso lineal general** si:

El proceso lineal general

Sea X_t una serie de tiempo observada, y ϵ_t una serie no observable de tiempo de ruido blanco, es decir, una serie de v.a.'s iid, de media cero.

Se dice que X_t es un **proceso lineal general** si:

$$X_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \cdots$$

El proceso lineal general

Sea X_t una serie de tiempo observada, y ϵ_t una serie no observable de tiempo de ruido blanco, es decir, una serie de v.a.'s iid, de media cero.

Se dice que X_t es un **proceso lineal general** si:

$$X_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \cdots$$

O, en otras palabras:

El proceso lineal general

Sea X_t una serie de tiempo observada, y ϵ_t una serie no observable de tiempo de ruido blanco, es decir, una serie de v.a.'s iid, de media cero.

Se dice que X_t es un **proceso lineal general** si:

$$X_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \cdots$$

O, en otras palabras:

$$X_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$$

El proceso lineal general

Sea X_t una serie de tiempo observada, y ϵ_t una serie no observable de tiempo de ruido blanco, es decir, una serie de v.a.'s iid, de media cero.

Se dice que X_t es un **proceso lineal general** si:

$$X_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \cdots$$

O, en otras palabras:

$$X_t = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$$

Dado que el lado derecho es una serie infinita, es necesario que dicha serie sea convergente.

El proceso lineal general

Para ello, se asume que:

$$\sum_{i=1}^{\infty} \psi_i^2 < \infty$$

El proceso lineal general

Para ello, se asume que:

$$\sum_{i=1}^{\infty} \psi_i^2 < \infty$$

Una forma de conseguir esto es asumir:

El proceso lineal general

Para ello, se asume que:

$$\sum_{i=1}^{\infty} \psi_i^2 < \infty$$

Una forma de conseguir esto es asumir:

$$\psi_i = \phi^i$$

El proceso lineal general

Para ello, se asume que:

$$\sum_{i=1}^{\infty} \psi_i^2 < \infty$$

Una forma de conseguir esto es asumir:

$$\psi_i = \phi^i$$

donde $-1 < \phi < 1$.

El proceso lineal general

Para ello, se asume que:

$$\sum_{i=1}^{\infty} \psi_i^2 < \infty$$

Una forma de conseguir esto es asumir:

$$\psi_i = \phi^i$$

donde $-1 < \phi < 1$.

De aquí que:

El proceso lineal general

Para ello, se asume que:

$$\sum_{i=1}^{\infty} \psi_i^2 < \infty$$

Una forma de conseguir esto es asumir:

$$\psi_i = \phi^i$$

donde $-1 < \phi < 1$.

De aquí que:

$$X_t = \epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \cdots$$

Esperanza y varianza de un proceso lineal general

Para calcular la esperanza tenemos que:

Esperanza y varianza de un proceso lineal general

Para calcular la esperanza tenemos que:

$$
\begin{aligned}
E(X_t) &= E(\epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \cdots) && (1)\\
&= E(\epsilon_t) + \phi E(\epsilon_{t-1}) + \phi^2 E(\epsilon_{t-2}) + \cdots && (2)\\
&= 0 && (3)
\end{aligned}
$$

Esperanza y varianza de un proceso lineal general

Para calcular la esperanza tenemos que:

$$\begin{aligned} E(X_t) &= E(\epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \cdots) && (1) \\ &= E(\epsilon_t) + \phi E(\epsilon_{t-1}) + \phi^2 E(\epsilon_{t-2}) + \cdots && (2) \\ &= 0 && (3) \end{aligned}$$

Para la varianza se tiene que:

Esperanza y varianza de un proceso lineal general

Para calcular la esperanza tenemos que:

$$\begin{aligned} E(X_t) &= E(\epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \cdots) && (1)\\ &= E(\epsilon_t) + \phi E(\epsilon_{t-1}) + \phi^2 E(\epsilon_{t-2}) + \cdots && (2)\\ &= 0 && (3) \end{aligned}$$

Para la varianza se tiene que:

$$\begin{aligned} Var(X_t) &= Var(\epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \cdots) && (4)\\ &= Var(\epsilon_t) + \phi^2 Var(\epsilon_{t-1}) + \phi^4 Var(\epsilon_{t-2}) + \cdots && (5)\\ &= \sigma_e^2(1 + \phi^2 + \phi^4 + \cdots) && (6)\\ &= \frac{\sigma_e^2}{1-\phi^2} && (7) \end{aligned}$$

Autocovarianza y autocorrelación de un proceso general lineal

Para la autocovarianza se tiene:

Autocovarianza y autocorrelación de un proceso general lineal

Para la autocovarianza se tiene:

$$Cov(X_t, X_{t-\tau}) = Cov(\sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}, \sum_{i=0}^{\infty} \phi^i \epsilon_{t-\tau-i}) \tag{8}$$

$$= \sigma_e^i \sum_{i=0} \phi^{\tau+k} \tag{9}$$

$$= \frac{\phi^\tau \sigma_e^2}{1-\phi^2} \tag{10}$$

Autocovarianza y autocorrelación de un proceso general lineal

Para la autocovarianza se tiene:

$$Cov(X_t, X_{t-\tau}) = Cov(\sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}, \sum_{i=0}^{\infty} \phi^i \epsilon_{t-\tau-i}) \tag{8}$$

$$= \sigma_e^i \sum_{i=0} \phi^{\tau+k} \tag{9}$$

$$= \frac{\phi^\tau \sigma_e^2}{1-\phi^2} \tag{10}$$

Por lo tanto, para la autocorrelación se tiene:

$$Corr(X_t, X_{t-\tau}) = \frac{\left(\frac{\phi^\tau \sigma_e^2}{1-\phi^2}\right)}{\left(\frac{\sigma_e^2}{1-\phi^2}\right)} \tag{11}$$

$$= \phi^\tau \tag{12}$$

Estacionaridad de un proceso lineal general

De lo anterior se sigue que un proceso lineal general es débilmente estacionario.

Estacionaridad de un proceso lineal general

De lo anterior se sigue que un proceso lineal general es débilmente estacionario.

Más aún, es de aquí que proviene la definición de estacionaridad.

Procesos de media móvil (MA)

Ocurre cuando, en un proceso general lineal todos los coeficientes son nulos a partir de un rezago q, es decir:

$$\psi_i = 0$$

para todo $i > q$.

Procesos de media móvil (MA)

Ocurre cuando, en un proceso general lineal todos los coeficientes son nulos a partir de un rezago q, es decir:

$$\psi_i = 0$$

para todo $i > q$.
En este caso, a X_t se lo denomina **proceso de media móvil de orden** q, y se denota por $MA(q)$.

Procesos de media móvil (MA)

Ocurre cuando, en un proceso general lineal todos los coeficientes son nulos a partir de un rezago q, es decir:

$$\psi_i = 0$$

para todo $i > q$.
En este caso, a X_t se lo denomina **proceso de media móvil de orden** q, y se denota por $MA(q)$.

En este caso, se suele cambiar la notación de los coeficientes y se suele expresar el proceso de la siguiente forma:

Procesos de media móvil (MA)

Ocurre cuando, en un proceso general lineal todos los coeficientes son nulos a partir de un rezago q, es decir:

$$\psi_i = 0$$

para todo $i > q$.
En este caso, a X_t se lo denomina **proceso de media móvil de orden** q, y se denota por $MA(q)$.

En este caso, se suele cambiar la notación de los coeficientes y se suele expresar el proceso de la siguiente forma:

$$X_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} + \cdots - \theta_q \epsilon_{t-q}$$

Procesos MA(1)

En este caso:

$$X_t = \epsilon_t - \theta\epsilon_{t-1}$$

Procesos MA(1)

En este caso:

$$X_t = \epsilon_t - \theta\epsilon_{t-1}$$

La esperanza es:

$$E(X_t) = 0$$

Procesos MA(1)

En este caso:

$$X_t = \epsilon_t - \theta\epsilon_{t-1}$$

La esperanza es:

$$E(X_t) = 0$$

La varianza es:

Procesos MA(1)

En este caso:

$$X_t = \epsilon_t - \theta\epsilon_{t-1}$$

La esperanza es:

$$E(X_t) = 0$$

La varianza es:

$$Var(X_t) = \sigma_e^2(1 + \theta^2)$$

Procesos MA(1)

En este caso:

$$X_t = \epsilon_t - \theta \epsilon_{t-1}$$

La esperanza es:

$$E(X_t) = 0$$

La varianza es:

$$Var(X_t) = \sigma_e^2 (1 + \theta^2)$$

La autocovarianza es:

Procesos MA(1)

En este caso:

$$X_t = \epsilon_t - \theta\epsilon_{t-1}$$

La esperanza es:

$$E(X_t) = 0$$

La varianza es:

$$Var(X_t) = \sigma_e^2(1 + \theta^2)$$

La autocovarianza es:

$$\begin{aligned} Cov(X_t, X_{t-1}) &= Cov(\epsilon_t - \theta\epsilon_{t-1}, \epsilon_{t-1} - \theta\epsilon_{t-2}) \quad (13) \\ &= -\theta Cov(\epsilon_{t-1}, \epsilon_{t-1}) = -\theta\sigma_e^2 \quad (14) \end{aligned}$$

Procesos MA(1)

En este caso:

$$X_t = \epsilon_t - \theta\epsilon_{t-1}$$

La esperanza es:

$$E(X_t) = 0$$

La varianza es:

$$Var(X_t) = \sigma_e^2(1 + \theta^2)$$

La autocovarianza es:

$$\begin{aligned} Cov(X_t, X_{t-1}) &= Cov(\epsilon_t - \theta\epsilon_{t-1}, \epsilon_{t-1} - \theta\epsilon_{t-2}) && (13) \\ &= -\theta Cov(\epsilon_{t-1}, \epsilon_{t-1}) = -\theta\sigma_e^2 && (14) \end{aligned}$$

y,

Procesos MA(1)

En este caso:

$$X_t = \epsilon_t - \theta\epsilon_{t-1}$$

La esperanza es:

$$E(X_t) = 0$$

La varianza es:

$$Var(X_t) = \sigma_e^2(1 + \theta^2)$$

La autocovarianza es:

$$\begin{aligned} Cov(X_t, X_{t-1}) &= Cov(\epsilon_t - \theta\epsilon_{t-1}, \epsilon_{t-1} - \theta\epsilon_{t-2}) && (13) \\ &= -\theta Cov(\epsilon_{t-1}, \epsilon_{t-1}) = -\theta\sigma_e^2 && (14) \end{aligned}$$

y,

$$Cov(X_t, X_{t-2}) = Cov(\epsilon_t - \theta\epsilon_{t-1}, \epsilon_{t-2} - \theta\epsilon_{t-3}) = 0 \qquad (15)$$

Procesos MA(1)

En este caso:

$$X_t = \epsilon_t - \theta\epsilon_{t-1}$$

La esperanza es:

$$E(X_t) = 0$$

La varianza es:

$$Var(X_t) = \sigma_e^2(1+\theta^2)$$

La autocovarianza es:

$$Cov(X_t, X_{t-1}) = Cov(\epsilon_t - \theta\epsilon_{t-1}, \epsilon_{t-1} - \theta\epsilon_{t-2}) \quad (13)$$

$$= -\theta Cov(\epsilon_{t-1}, \epsilon_{t-1}) = -\theta\sigma_e^2 \quad (14)$$

y,

$$Cov(X_t, X_{t-2}) = Cov(\epsilon_t - \theta\epsilon_{t-1}, \epsilon_{t-2} - \theta\epsilon_{t-3}) = 0 \quad (15)$$

La autocorrelación es:

$$\rho_1 = \frac{-\theta\sigma_e^2}{\sigma_e^2(1+\theta^2)} = \frac{-\theta}{1+\theta^2}$$

Procesos MA(q)

En este caso:

$$X_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q \epsilon_{t-q}$$

Procesos MA(q)

En este caso:

$$X_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q \epsilon_{t-q}$$

La esperanza es:

$$E(X_t) = 0$$

Procesos MA(q)

En este caso:

$$X_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q \epsilon_{t-q}$$

La esperanza es:

$$E(X_t) = 0$$

La varianza es:

Procesos MA(q)

En este caso:

$$X_t = \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - \cdots - \theta_q\epsilon_{t-q}$$

La esperanza es:

$$E(X_t) = 0$$

La varianza es:

$$Var(X_t) = \sigma_e^2(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)$$

La autocorrelación es:

Procesos MA(q)

En este caso:

$$X_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q \epsilon_{t-q}$$

La esperanza es:

$$E(X_t) = 0$$

La varianza es:

$$Var(X_t) = \sigma_e^2 (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)$$

La autocorrelación es:

$$r_i = \frac{-\theta_i + \theta_1 \theta_{i+1} - \theta_2 \theta_{i+2} + \cdots + \theta_{q-i} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2}$$

Para $i > q$

Procesos autorregresivos

Una serie de tiempo se denomina **un proceso autorregresivo** de orden p si:

Procesos autorregresivos

Una serie de tiempo se denomina **un proceso autorregresivo** de orden p si:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

$\mu+$

El modelo AR(1)

Para $p = 1$, se tiene que:

El modelo AR(1)

Para $p = 1$, se tiene que:

$$X_t = \phi X_{t-1} + \epsilon_t$$

El modelo AR(1)

Para $p = 1$, se tiene que:

$$X_t = \phi X_{t-1} + \epsilon_t$$

La autocovarianza es:

El modelo AR(1)

Para $p = 1$, se tiene que:

$$X_t = \phi X_{t-1} + \epsilon_t$$

La autocovarianza es:

$$R(0) = \frac{\sigma_e^2}{1 - \phi^2},$$

El modelo AR(1)

Para $p = 1$, se tiene que:

$$X_t = \phi X_{t-1} + \epsilon_t$$

La autocovarianza es:

$$R(0) = \frac{\sigma_e^2}{1-\phi^2},$$

$$R(\tau) = \phi^\tau \frac{\sigma_e^2}{1-\phi^2}$$

El modelo AR(1)

Para $p = 1$, se tiene que:

$$X_t = \phi X_{t-1} + \epsilon_t$$

La autocovarianza es:

$$R(0) = \frac{\sigma_e^2}{1 - \phi^2},$$

$$R(\tau) = \phi^\tau \frac{\sigma_e^2}{1 - \phi^2}$$

De aquí:

El modelo AR(1)

Para $p = 1$, se tiene que:

$$X_t = \phi X_{t-1} + \epsilon_t$$

La autocovarianza es:

$$R(0) = \frac{\sigma_e^2}{1 - \phi^2},$$

$$R(\tau) = \phi^\tau \frac{\sigma_e^2}{1 - \phi^2}$$

De aquí:

$$r_\tau = \phi^\tau$$

El modelo AR(1)

Para $p = 1$, se tiene que:

$$X_t = \phi X_{t-1} + \epsilon_t$$

La autocovarianza es:

$$R(0) = \frac{\sigma_e^2}{1 - \phi^2},$$

$$R(\tau) = \phi^\tau \frac{\sigma_e^2}{1 - \phi^2}$$

De aquí:

$$r_\tau = \phi^\tau$$

Siempre y cuando $\phi^2 < 1$

El modelo autorregresivo general $AR(p)$

Es de la forma:

El modelo autorregresivo general $AR(p)$

Es de la forma:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

Estacionariedad del proceso AR y otras variantes

Si un modelo AR no es estacionario, significa que los valores anteriores del término de error tendrán un efecto no decreciente en el valor actual de la variable dependiente.

Esto implica que los coeficientes en el proceso MA no convergen a cero a medida que aumenta la longitud de lag (rezago).

$\therefore$ Para que un modelo AR para sea estacionario, los coeficientes del correspondiente proceso MA disminuyen con la longitud del lag, convergiendo a 0.

El test de estacionariedad en un modelo AR (con p rezagos) e que las raíces de la ecuación característica caen fuera del círculo unitario (es decir > 1), donde la ecuación característica es:

$$1 - \Phi_1 z - \Phi_2 z^2 - \ldots - \Phi_p z^p = 0.$$

Test de Dickey-Fuller - Test de Dickey-Fuller Aumentado

- Esta prueba supone que el término de error (u_t) siguen los supuestos Gauss-Markov.
- El estadístico de prueba no sigue la distribución t, sus valores críticos se han producido específicamente para la prueba (pero son los de la distribución t).
- Una constante y una tendencia también podrían incluirse en esta prueba.
- La prueba es para una raíz unitaria contra ninguna raíz unitaria, es decir, la variable debe ser diferenciada una vez para inducir estacionariedad.

Test Dickey-Fuller Aumentado (ADF)

- El término de error en la prueba de Dickey-Fuller tiene generalmente autocorrelación, que necesita ser eliminado para que el resultado sea válido. La principal vía de quitar esta autocorrelación es añadir variables dependientes rezagadas hasta que la autocorrelación haya desaparecido.

La prueba es como sigue, donde el número de variables dependientes rezagadas se determina mediante un criterio de información:

$$\Delta y_t = \beta y_{t-1} + \sum_{i=0}^{N} \Delta y_{t-1} + \varepsilon_t,$$

Pruebas para raíz unitaria

Las pruebas más utilizadas incluyen:

- Prueba de Dickey-Fuller aumentada, válida en muestras grandes.
- Prueba de Phillips-Perron
- Prueba KPSS (Kwiatkowski–Phillips–Schmidt–Shin),la hipótesis nula es la estacionariedad de la tendencia en lugar de la presencia de una raíz unitaria .
- Prueba ADF-GLS

Las pruebas de raíz unitaria están estrechamente vinculadas a la correlación de la serie. Sin embargo, si bien todos los procesos con una raíz unitaria exhibirán correlación serial, no todas las series temporales correlacionadas serialmente tendrán una raíz unitaria. Las pruebas de correlación serial más populares incluyen:

- Prueba de Ljung-Box
- Prueba de Breusch-Godfrey
- Prueba de Durbin-Watson

¿Cómo estimar los parámetros de un modelo?

En general no conocemos los valores de las constantes del modelo y debemos estimarlos a partir de las observaciones de la serie de tiempo.
Existen distintos enfoques:

- Usando las ecuaciones de Yule-Walker.
- Cuadrados mínimos.
- Método de máxima verosimilitud basado en el modelo de estados.

Estimación del modelo AR mediante las ecs. de Y-W

Las ecuaciones de Yule-Walker son:

$$\gamma_m = \sum_{k=1}^{p} \varphi_k \gamma_{m-k} + \sigma_\varepsilon^2 \delta_{m,0}$$

donde $m = 0, ..., p$, produciendo $p + 1$ ecuaciones, donde γ_m es la función de autocovarianza de X_t, σ_ε es la desviación estándar del proceso de ruido de entrada, y $\delta_{m,0}$ es la función delta de Kronecker.
Debido a que la última parte de una ecuación individuo es distinto de cero sólo si m $= 0$, el conjunto de ecuaciones se puede resolver mediante la representación de las ecuaciones para m> 0 en forma de matriz

Introducción a los Modelos ARIMA

Auto-**R**egresive **I**ntegrated **M**oving **A**verage.

Los modelos ARIMA(p,d,q) tiene tres parámetros: p (rezagos de la propia variable), d (orden de diferenciación hasta que se haga estacionaria la serie), y q (rezagos del error de estimación).

1. p es el número de términos autorregresivos.
2. d es el número de diferencias que tomo sobre la serie original.
3. q es el número de términos moving-average

Si la serie necesita ser diferenciada (calcular primeras o segundas diferencias) para hacerla estacionaria se dice que es "integrada" de orden d. Es decir, es I(d). Si la serie temporal que me dan ya es estacionaria, entonces la serie es I(0). Si la tengo que diferenciar una vez para que se haga estacionaria, entonces es I(1), y así sucesivamente.

El objetivo es definir los valores de p, d, q que separen la serie temporal entre una "señal" que nos permita hacer forecasts, y ruido blanco.

Otras extensiones de los modelos ARMA

1. SARIMA: Modelo de media móvil autorregresiva integrada estacional, refleja la característica de variación estacional en series de tiempo.
2. ARFIMA: la F indica que el parámetro d no es entero.
3. ARIMAX: Modelo de media móvil autorregresiva integrada con entradas exógenas.
4. NARMA, WARIMAX,...

Metodología Box-Jenkins

Metodología Box-Jenkins

Es un método para estimar modelos ARIMA, basado en la ACF y la PACF como medios de determiner la estacionariedad de la variable en cuestión, y el número de rezagos del modelo ARIMA.

El enfoque Box-Jenkins típicamente comprende cuatro partes:

1. Identificación.
2. Estimación.
3. Verificación del diagnóstico (principalmente para autocorrelación).
4. Predicción/Forecasting.

A pesar de que los métodos usando ACF y PACF se utilizan frecuentemente, hay otros métodos basados en ‘criterios de información’ que también pueden ser usados.

1. Identificación del modelo

Es la parte más importante del proceso. Esto es tanto 'arte' como 'ciencia'.

- El primer paso es determinar si la variable es estacionaria. Esto se puede hacer con el correlograma. Si no es estacionario la serie se tiene que diferenciar hasta inducir estacionariedad.
- La siguiente etapa es determinar p y q en el ARIMA(p, d, q).

Identificación del modelo

- Para determinar la estructura de rezago apropiado en la parte AR del modelo, se utiliza el PACF o correlograma parcial, donde la cantidad de puntos no nulos de la PACF determina dónde los rezagos del AR necesitan ser incluidos.
- Para determinar la estructura de rezagos del MA, se utiliza el ACF o correlograma. De nuevo los puntos distintos de cero indican donde se deben incluir los rezagos.
- También se pueden necesitar variables ficticias (dummies) estacionales, si los datos exhiben efectos estacionales.

¿Cómo chequeamos?

- Con este enfoque sólo prueba (por lo general) autocorrelación, usando la Q o el estadístico de Ljung-Box.
- Si hay evidencia de autocorrelación, tenemos que volver a la etapa de identificación y especificar de nuevo el modelo, añadiendo más retardos.
- Una problema de este enfoque es que no identifica si el modelo es demasiado grande o sobre-parametrizado, sólo nos dice si es demasiado pequeño.

Modelo parsimonioso El objetivo de este tipo de modelado es producir un modelo parsimonioso, o tan pequeño como sea posible, y que al mismo tiempo que pase las comprobaciones de diagnóstico. Un modelo parsimonioso es deseable porque incluyendo rezagos irrelevantes en el modelo, aumentan los coeficientes de los errores coeficiente estándar y por lo tanto reduce los estadísticos t. Los modelos que incorporan un gran número de rezagos, no tienden a pronosticar bien, ya que se ajustan demasiado a las características específicas de los datos, explicando demasiado el ruido o características aleatorias en los datos (overfitting).