Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Universidad Nacional de Rosario

Identificación de Sistemas

Título: Predictor Optimo para Estructura de Modelo Lineal
Autor: Dr. Juan Carlos Gómez **Fecha: Abril de 2002**

Definición: Sea $\hat{y}(n/n-1,\theta)$ un predictor de la salida $y(n)$ dados los datos hasta el instante *n-1* y dado el vector de parámetros θ. Se dice que el predictor es **óptimo**, en el sentido de media cuadrática, si minimiza la varianza del error de predicción.

A continuación, derivaremos la expresión del **predictor óptimo** para algunas estructuras de modelo particulares [1].
Consideremos la estructura de **modelo lineal** general dada por:

$$M(\theta): \quad y(n) = G(q^{-1},\theta)u(n) + H(q^{-1},\theta)e(n)$$

$$E\{e(n)e^T(s)\} = \Lambda\,\delta_{n,s}$$

donde:

$y(n) \in \Re^n$ vector de salida

$u(n) \in \Re^m$ vector de entrada

$e(n) \in \Re^n$ ruido blanco con media cero

$\theta \in D \subset \Re^p$ vector de parámetros

$$D = \left\{ \begin{array}{c} \theta \,/\, H^{-1}(q^{-1},\theta) \text{ y } H^{-1}(q^{-1},\theta)G(q^{-1},\theta) \text{ son asintóticamente estables;} \\ G(0,\theta) = 0;\ H(0,\theta) = I;\ \ \Lambda \text{ es definida no negativa} \end{array} \right\}$$

NB:

- La notación $M(\theta)$ denota un modelo particular, dentro de la estructura de modelo, correspondiente al valor del parámetro θ.
- Notar que *e(n)* es ruido temporalmente blanco (lo cual significa que *e(n)* y *e(s)* no están correlacionados, como lo pone de manifiesto la función delta que multiplica a la matriz de covarianza). Sin embargo, *e(n)* no necesariamente es espacialmente blanco, ya que la matriz de covarianza Λ no es en general una matriz diagonal.
- La forma en que se definió *D* se justificará más adelante.

Se asume que *u(n)* y *e(s)* no están correlacionados para $n < s$ ($\Rightarrow$ lazo abierto). A partir de la estructura de modelo lineal general:

$$y(n) = G(q^{-1},\theta)u(n) + H(q^{-1},\theta)e(n)$$

se puede obtener una expresión del ruido *e(n)*, como se muestra a continuación:

$$H(q^{-1},\theta)e(n) = y(n) - G(q^{-1},\theta)u(n)$$
$$e(n) = H^{-1}(q^{-1},\theta)\left[y(n) - G(q^{-1},\theta)u(n)\right] \quad (*)$$

Luego, el predictor óptimo sale de plantear:

$$\begin{aligned} y(n) &= G(q^{-1},\theta)u(n) + H(q^{-1},\theta)e(n) = \\ &= G(q^{-1},\theta)u(n) + \left[H(q^{-1},\theta) - I\right]e(n) + e(n) \end{aligned}$$

y reemplazando (*) en la expresión anterior:

$$\begin{aligned} y(n) &= G(q^{-1},\theta)u(n) + \left[H(q^{-1},\theta) - I\right]H^{-1}(q^{-1},\theta)\left[y(n) - G(q^{-1},\theta)u(n)\right] + e(n) = \\ &= H^{-1}(q^{-1},\theta)G(q^{-1},\theta)u(n) + \left[I - H^{-1}(q^{-1},\theta)\right]y(n) + e(n) = \\ &= z(n) + e(n) \end{aligned}$$

donde se ha definido:

$$z(n) = H^{-1}(q^{-1},\theta)G(q^{-1},\theta)u(n) + \left[I - H^{-1}(q^{-1},\theta)\right]y(n)$$

Notar que *z(n)* y *e(n)* **no están correlacionados**, ya que la hipótesis $\theta \in D$ implica que:

$$H(0,\theta) = I \Rightarrow H^{-1}(0,\theta) = I$$

por lo que el término $\left[I - H^{-1}(q^{-1},\theta)\right]y(n)$ depende sólo de valores pasados de la salida $y(n-1), y(n-2),\ldots$ que por lo tanto no están correlacionados con *e(n)*.

Sea $y^*(n)$ un predictor arbitrario de *y(n)* basado en los datos hasta el instante *n-1*. La matriz de covarianza del error de predicción resulta:

$$\begin{aligned} E\left\{\varepsilon(n)\varepsilon^T(n)\right\} &= E\left\{\left[y(n) - y^*(n)\right]\left[y(n) - y^*(n)\right]^T\right\} = \\ &= E\left\{\left[z(n) + e(n) - y^*(n)\right]\left[z(n) + e(n) - y^*(n)\right]^T\right\} = \\ &= E\left\{\left[z(n) - y^*(n)\right]\left[z(n) - y^*(n)\right]^T\right\} + \Lambda \geq \Lambda \end{aligned}$$

Puede verse entonces que si $y^*(n) = z(n)$, el predictor resulta óptimo (en el sentido de media cuadrática) ya que la covarianza del error de predicción resulta mínima.

Además, notar que el error de predicción resulta:

$$\varepsilon(n,\theta) = e(n)$$

Finalmente, el **predictor óptimo** resulta:

$$\boxed{\hat{y}(n/n-1,\theta) = H^{-1}(q^{-1},\theta)G(q^{-1},\theta)u(n) + \left[I - H^{-1}(q^{-1},\theta)\right]y(n)}$$

y es denominado **one step ahead predictor** (predictor de un paso en avance).

El **error de predicción** resulta:

$$\boxed{\varepsilon(n,\theta) = H^{-1}(q^{-1},\theta)\left[y(n) - G(q^{-1},\theta)u(n)\right]}$$

Notar que la hipótesis planteada anteriormente sobre que θ esté restringida a *D*, implica que el predictor es estable (ya que específicamente, se pedía que $H^{-1}(q^{-1},\theta)$ sea asintóticamente estable), lo cual justifica la definición del conjunto *D*.

Ejemplo:

Consideremos el siguiente modelo ARMAX:

$$y(n) + a\,y(n-1) = bu(n-1) + e(n) + ce(n-1)$$

donde *e(n)* es ruido blanco con media cero y varianza λ^2.

Esto se puede escribir de la forma:

$$(1 + aq^{-1})y(n) = bq^{-1}u(n) + (1 + cq^{-1})e(n)$$

lo cual nos conduce a la siguiente expresión:

$$y(n) = \frac{bq^{-1}}{(1 + aq^{-1})}u(n) + \frac{(1 + cq^{-1})}{(1 + aq^{-1})}e(n)$$

Luego:

$$G(q^{-1},\theta) = \frac{bq^{-1}}{(1 + aq^{-1})} \qquad\qquad H(q^{-1},\theta) = \frac{(1 + cq^{-1})}{(1 + aq^{-1})}$$

Para poder obtener una expresión del **predictor óptimo**, primero se deben imponer ciertas condiciones sobre $G(q^{-1},\theta)$ y $H(q^{-1},\theta)$:

$$G(q^{-1},\theta) \textit{ debe ser asintóticamente estable } \Rightarrow |a| < 1$$
$$H(q^{-1},\theta) \textit{ y } H^{-1}(q^{-1},\theta) \textit{ deben ser asintóticamente estables} \Rightarrow |a| < 1 \textit{ y } |c| < 1$$

Finalmente, el **predictor óptimo** resulta:

$$\hat{y}(n/n-1,\theta) = \left[1 - \frac{1 + aq^{-1}}{1 + cq^{-1}}\right]y(n) + \frac{bq^{-1}}{1 + cq^{-1}}u(n) =$$
$$= \left[\frac{(c-a)q^{-1}}{1 + cq^{-1}}\right]y(n) + \left[\frac{bq^{-1}}{1 + cq^{-1}}\right]u(n)$$

Pasando a una ecuación en diferencias, el predictor óptimo para el modelo ARMAX resulta:

$$\hat{y}(n/n-1,\theta) = -c\,\hat{y}(n-1/n-2,\theta) + (c-a)\,y(n-1) + bu(n-1)$$

Referencias

[1] Söderström, T. and Stoica, P.. *System Identification*, Prentice Hall, N.J., 1989.