Identificación de SISTemas

Métodos de Error de Predicción

Métodos de Error de Predicción (PEM)

- Seleccionar una **estructura de modelos** ***M*,** con modelos particulares $M(\theta)$ parametrizados por el vector $\theta \in \mathbf{D_M} \subset \Re^p$

$$\mathbf{M} = \{\mathbf{M}(\theta) | \theta \in \mathbf{D_M}\}$$

- Se asume que se dispone de *N* pares de datos de entrada-salida para la estimación

$$Z^N = \{y(n), u(n) : n = 1, \cdots, N\}$$

- Para predecir la salida basados en datos pasados se construye un predictor $\hat{y}(n \mid \theta)$. Un predictor lineal típico es

$$\hat{y}(n \mid \theta) = F_1(q,\theta)y(n) + F_2(q,\theta)u(n)$$ **Predictor Lineal General**

Por ejemplo, para la estructura de modelo

$$y(n) = G(q,\theta)u(n) + H(q,\theta)e(n)$$

un predictor es el **one-step-ahead predictor,** dado por

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

el cual es **óptimo** en el sentido de media cuadrática.

- El **Problema de Identificación** consiste en seleccionar un valor apropiado del vector de parámetros, *i.e.,* un modelo particular $M(\theta)$ que sea el que mejor reproduzca los datos.

$$Z^N \rightarrow \hat{\theta}_N \in \mathbf{D_M}$$

mapeo $\rightarrow$ estimación de parámetros

- Para evaluar las capacidades predictivas del modelo se define un **error de predicción** como

$$\varepsilon(n,\theta_*) = y(n) - \hat{y}(n \mid \theta_*)$$

donde θ_* corresponde a un modelo particular $M(\theta_*)$.

- Se define una **función de costo**, o **criterio** que es una función escalar, definida positiva, de los errores de predicción $\{\varepsilon(n,\theta)\}_{n=1}^{N}$

$$V_N(\theta) = \frac{1}{N}\sum_{n=1}^{N} \ell(\varepsilon(n,\theta))$$

$\ell(\bullet)$ Función escalar a valores reales positivos

- La **estima de parámetros** se obtiene minimizando el criterio, i.e.

$$\hat{\theta}_N = \underset{\theta \in \mathbf{D_M}}{\operatorname{argmin}} V_N(\theta)$$

- Un ejemplo típico de Método de Error de Predicción es la estima de mínimos cuadrados para una estructura de regresor lineal.

Análisis Asintótico de las Estimas

Si se modelan las perturbaciones como un **proceso estocástico**, entonces las estimas resultan **variables aleatorias**, y tiene sentido darles una caracterización estadística, por ejemplo determinando su **valor medio** y su **matriz de covarianza** (momentos de primero y segundo orden).

Definición: Una estima $\hat{\theta}$ (o estimador) se dice **no desviada** (**unbiased**) si se verifica

$$\mathbf{E}[\hat{\theta}] = \theta_0$$

donde θ_0 es el verdadero valor del parámetro.

Definición: Una estima se dice **consistente** si se verifica

$$\hat{\theta} \to \theta_0 \quad \text{cuando} \quad N \to \infty$$

donde N es el número de datos, y donde la convergencia es con **probabilidad 1**.

Para analizar la consistencia de un estimador se realiza generalmente un **análisis asintótico** con el número de datos N tendiendo a infinito, y en algunos casos es posible determinar la **estima asintótica** (es decir determinar si es consistente) y también la **distribución asintótica** de las estimas.

Análisis Asintótico para los Métodos PEM

- Es de interés analizar las **propiedades estadísticas** de la estima $\hat{\theta}_N$ cuando el número de datos $N \to \infty$.
- Bajo ciertas condiciones puede probarse que

$$\sup_{\theta \in \mathbf{D_M}} \left| V_N(\theta) - V_\infty(\theta) \right| \xrightarrow{\text{as}} 0 \qquad \text{cuando} \qquad N \to \infty$$

donde $V_\infty(\theta) = \lim_{N\to\infty} V_N(\theta)$. Además

$$\boxed{\hat{\theta}_N \xrightarrow{\text{as}} \theta_*} \qquad \text{cuando} \qquad N \to \infty \qquad \textbf{consistencia}$$

donde $\theta_* \in \mathbf{D_C}$, que es el conjunto de parámetros que minimiza el criterio límite $V_\infty(\theta)$

Distribución asintótica de las estimas

La distribución de la variable aleatoria $\sqrt{N}\left(\hat{\theta}_N - \theta_*\right)$ converge a una **distribución Gaussiana** con media cero y matriz de covarianza

$$P = \left[V_\infty''(\theta_*)\right]^{-1} \left[\lim_{N\to\infty} N \mathbf{E}\left\{V_N'(\theta_*)^T V_N'(\theta_*)\right\}\right] \left[V_\infty''(\theta_*)\right]^{-1}$$

cuando $N \to \infty$, donde $V_N'(\theta_*)$ es el gradiente de $V_N(\theta)$, y $V_\infty''(\theta_*)$ es el Hessiano del criterio límite.

$$\boxed{\sqrt{N}\left(\hat{\theta}_N - \theta_*\right) \xrightarrow{\text{dist}} N(0, P)}$$