Estimación paramétrica

Planteamiento del problema

- **Parámetro a estimar:** θ
- **Vector de observaciones:** $\mathbf{x} = [\, x[0]\; x[1]\; \cdots\; x[N-1]\,]^T$

  Un vector de observaciones es una realización de un vector aleatorio $\mathbf{X}$ que depende, de alguna forma, del parámetro a estimar θ.

- **Modelo**

  Define la dependencia de $\mathbf{X}$ con θ. Dos posibilidades

  - Ecuación de medida: Muestra explícitamente la dependencia de las observaciones con θ.
  - Función densidad de probabilidad de $\mathbf{X}$ dependiente de θ: $f_{\mathbf{X}}(\mathbf{x};\theta)$

- **Estimador**

  Función que proporciona una estima de θ en función del vector de observaciones: $\hat{\theta}(\mathbf{x})$. Como las observaciones son realizaciones de un vector aleatorio $\mathbf{X}$, un estimador es una variable aleatoria:

  - Media: $E[\hat{\theta}]$
  - Varianza: $\mathrm{Var}[\hat{\theta}]$

- **Prestaciones de un estimador**

  - Sesgo: $b(\hat{\theta}) = E[\hat{\theta}] - \theta$
  - Varianza: $\mathrm{Var}(\hat{\theta}) = E[(\hat{\theta} - E[\hat{\theta}])^2]$
  - MSE: $MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] = \mathrm{Var}(\hat{\theta}) + b(\hat{\theta})^2$

  En general son función del valor de θ.

- **Generalización a la estima de varios parámetros**

$$\boldsymbol{\theta} = [\, \theta_1\; \theta_2\; \cdots\; \theta_p\,]^T \quad \Longrightarrow \quad \hat{\boldsymbol{\theta}}(\mathbf{x}) = [\, \hat{\theta}_1(\mathbf{x})\; \hat{\theta}_2(\mathbf{x})\; \cdots\; \hat{\theta}_p(\mathbf{x})\,]^T$$

Jesús Pérez Arriaga

Estimadores de máxima verosimilitud (ML)

- **Función de verosimilitud**

  Sea el modelo definido por la f.d.p. del vector de observaciones: $f_{\mathbf{X}}(\mathbf{x};\theta)$.
  Dado un vector de observaciones $\mathbf{x}$, la función de verosimilitud se obtiene sustituyendo el vector de observaciones en la f.d.p.:

$$L(\boldsymbol{\theta}) = f_{\mathbf{X}}(\mathbf{x};\boldsymbol{\theta})$$

  Es una función de los parámetros a estimar $\boldsymbol{\theta}$. Distintos vectores de observaciones dan lugar a distintas funciones de verosimilitud.

- **Estimador ML**

  Dada una observación $\mathbf{x}$, la estima de máxima verosimilitud es el valor de $\boldsymbol{\theta}$ que maximiza la función de verosimilitud

$$\hat{\boldsymbol{\theta}}_{ML}(\mathbf{x}) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\ L(\boldsymbol{\theta})$$

- **Propiedades de los estimadores ML**

  - Son asintóticamente insesgados: $\lim_{N\to\infty} E[\hat{\boldsymbol{\theta}}_{ML}] = \boldsymbol{\theta}$
  - Son asintóticamente Gaussianos: $\hat{\boldsymbol{\theta}}_{ML} \sim \mathcal{N}(\boldsymbol{\theta}, I^{-1}(\boldsymbol{\theta}))$, $\qquad [I((\boldsymbol{\theta}))]_{i,j} = -E\left[\frac{\partial^2 \log f_{\mathbf{X}}(\mathbf{x};\theta)}{\partial\theta_i\partial\theta_j}\right]$
  - Propiedad de invarianza: $\boldsymbol{\alpha} = h(\boldsymbol{\theta}) \;\Rightarrow\; \hat{\boldsymbol{\alpha}}_{ML} = h(\hat{\boldsymbol{\theta}_{ML}})$

Estimadores de mínimos cuadrados (LS)

- **Modelo:**
$$x[n] = s(n;\boldsymbol{\theta}) + r[n], \quad n = 0, 1, \cdots, N-1,$$
donde $r[n]$ es ruido blanco y $s(n;\boldsymbol{\theta})$ es una función determinista.

  En forma vectorial, $\mathbf{x} = \mathbf{s}(\boldsymbol{\theta}) + \mathbf{r}$, donde
$$\mathbf{x} = \begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ x[N-1] \end{bmatrix}, \quad \mathbf{s}(\boldsymbol{\theta}) = \begin{bmatrix} s(0;\boldsymbol{\theta}) \\ s(1;\boldsymbol{\theta}) \\ \vdots \\ s(N-1;\boldsymbol{\theta}) \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} r[0] \\ r[1] \\ \vdots \\ r[N-1] \end{bmatrix}$$

- **Función de coste:**
$$J(\boldsymbol{\theta}) = \sum_{n=0}^{N-1} |x[n] - s(n;\boldsymbol{\theta})|^2 = ||\mathbf{x} - \mathbf{s}(\boldsymbol{\theta}))||^2.$$

- **Estimador LS:** $$\hat{\boldsymbol{\theta}}_{LS}(\mathbf{x}) = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \; J(\boldsymbol{\theta})$$

- **Propiedades:**
  - No es necesario connocer la distribución del ruido.
  - Si $r[n]$ es WGN con varianza conocida, $\hat{\boldsymbol{\theta}}_{LS}(\mathbf{x}) = \hat{\boldsymbol{\theta}}_{ML}(\mathbf{x})$.

- **Estimadores LS en modelo lineal**
  - Modelo lineal: $\mathbf{s}(\boldsymbol{\theta}) = \mathbf{S}\ \boldsymbol{\theta}$
  - $\hat{\boldsymbol{\theta}}_{LS}(\mathbf{x}) = (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{x} = \mathbf{S}^{\#}\mathbf{x}$.
  - Casos particulares:
    - Si las columnas de $\mathbf{S}$ son ortogonales con norma α, $\mathbf{S}^{\#} = \frac{1}{\alpha^2}\mathbf{S}^T$.
    - Si $p = 1$, $\mathbf{S}^{\#} = \frac{1}{\alpha^2}\mathbf{s}^T$