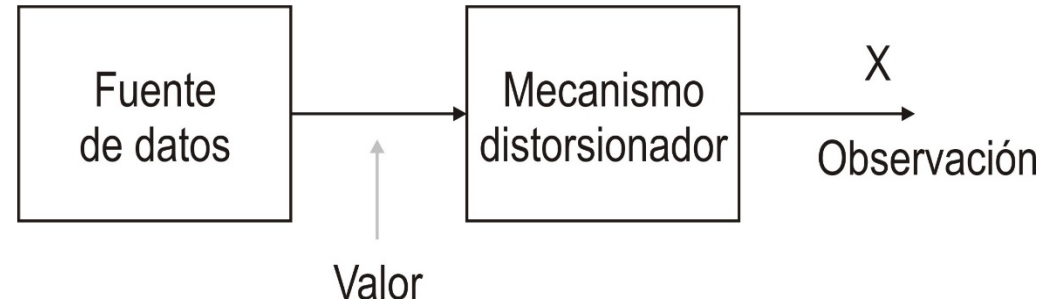


Estimación

Estimador de Máxima Verosimilitud

- Problema genérico



- La fuente puede generar un conjunto de valores o incluso cualquier valor.
- **Objetivo:** determinar una **estimación** del valor original a partir de la observación realizada de forma que se cometa el menor error
 - Basado en:
 - Conocimiento de las características de la fuente de datos
 - Conocimiento de las características del mecanismo distorsionador (ej.: canal, ruido, etc.)

- Al tratar el tema de decisión existía una variante denominada de “máxima verosimilitud” (ML, maximum likelihood).
- En la decisión ML se supone desconocida la información a priori de la fuente de información cuyo valor actual se quiere estimar/decidir.
- En el **problema de decisión**, la información a priori es la probabilidad de cada una de las posibles hipótesis.
- En el **problema de estimación**, la información a priori es la función densidad de probabilidad de la variable aleatoria mediante la que se modela la fuente de información, $f(s)$



El problema de estimación de máxima verosimilitud se va a resolver de modo paralelo al de detección de máxima verosimilitud.

- El problema de estimación se ha planteado como el estudio, según un cierto criterio de error, de la función densidad de probabilidad condicional:

$$f(s|x) = \frac{f(x|s)f(s)}{f(x)}$$

- En un caso multidimensional, se dispone de un conjunto de observaciones:

$$\vec{x} = [x_1, x_2, \dots, x_N]$$

- Entonces, la expresión anterior se reformula como:

$$f(s|\vec{x}) = \frac{f(\vec{x}|s)f(s)}{f(\vec{x})}$$

$$f(\vec{x}|s) = \prod_{i=1}^N f(x_i|s)$$

Planteamiento

$f(x|s)$: relación estadística entre X y S

$f(s)$: información a priori sobre S

$f(x)$: interviene exclusivamente como un factor de escala

- La forma de $f(s|x)$, que es la que nos interesa estudiar, dependerá del problema:

$$f(s|x) = K \cdot f(x|s) \cdot f(s)$$

- Cuando se emplea el criterio MAP se pretende proponer como estima el valor que sea más probable, lo que comprobamos que respondía a realizar la estima como

$$\hat{s}_{MAP}(x) = \operatorname{argmax}\{f(s|x)\}$$

- Ejemplo: $f(s)$ uniforme y ruido gaussiano

Planteamiento

- Ejemplo: $f(s)$ uniforme y ruido gaussiano

Estamos tomando como estima el valor que hace máximo $f(x|s)$

Esto es lo mismo que seleccionar el valor más probable del ruido n (la esperanza matemática del ruido) y restarlo de la observación x :

$$x = s + n$$

$$\hat{s}_{ML} = x - E[n]$$

- Esta estimación es la que se conoce como de **máxima verosimilitud** (**maximum likelihood, ML**)
- Obviamente, esta es una aproximación sencilla. Puede formalizarse. Para ello consideramos de nuevo el caso MAP general:

$$f(s|\vec{x}) = \frac{f(\vec{x}|s)f(s)}{f(\vec{x})}$$

- Partiendo del Teorema de Bayes y considerando el caso MAP:

$$f(s|\vec{x}) = \frac{f(\vec{x}|s)f(s)}{f(\vec{x})}$$

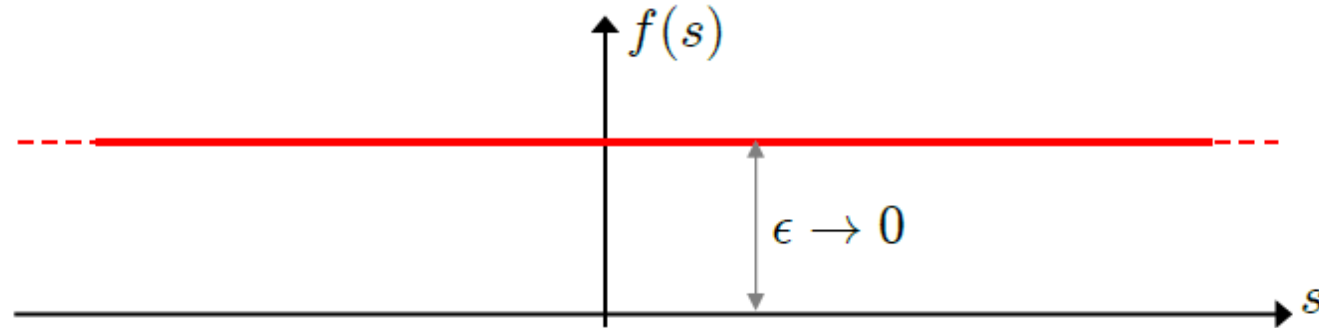
- Tomando logaritmos neperianos y minimizando se obtiene la llamada **ecuación MAP**:

$$\left. \frac{\partial \ln f(\vec{x}|s)}{\partial s} \right|_{s=\hat{s}(\vec{x})} + \left. \frac{\partial \ln f(s)}{\partial s} \right|_{s=\hat{s}(\vec{x})} = 0$$

- La primera parte de la ecuación MAP es la función de log-verosimilitud (log likelihood), y la segunda parte corresponde a la información a priori
- ¿En qué consistiría no conocer la información a priori en este caso?

Formalización

- La fdp $f(s)$ sería



- De esta manera, la ecuación MAP se convierte en

$$\left. \frac{\partial \ln f(\vec{x}|s)}{\partial s} \right|_{s=\hat{s}(\vec{x})} = 0$$

- Esta es la **ecuación de verosimilitud** o **ML**
- El caso ML corresponde matemáticamente al caso límite MAP en el que el conocimiento a priori se aproxima a cero

- Si se descarta la información a priori (o se considera como se ha visto), el máximo de $f(s|\vec{x})$ coincide con el de $f(\vec{x}|s)$

$$\hat{s}_{ML}(x) = \operatorname{argmax}\{f(x|s)\}$$

- Diferente del MAP:

$$\hat{s}_{MAP}(x) = \operatorname{argmax}\{f(s|x)\}$$

- El estimador de máxima verosimilitud de un valor escalar se toma como el valor de x que maximiza la función de verosimilitud. Esta maximización se realiza sobre todo el rango posible de valores de s .

Estimación

Estimador de Máxima Verosimilitud