TRATAMIENTO DIGITAL DE SEÑALES

Ingeniería de Telecomunicación (4º, 2º c)

Unidad 4ª: Los problemas generales gaussianos

Aníbal R. Figueiras Vidal
Jesús Cid Sueiro
Ángel Navia Vázquez

Área de Teoría de la Señal y Comunicaciones
Universidad Carlos III de Madrid

io de Ampliación

a. independientes, s: $G(0, v_s)$ y r: $G(0, v_r)$, dan lugar a la observación:
r (señal gaussiana en ruido gaussiano).

ínese el estimador ms de s.

e está buscando $\hat{s}_{ms} = E\{s \mid x\}$

conoce directamente $p(s \mid x)$; pero sí

$p(s)$: $G(0, v_s)$

ás

$p(x)$: $G(0, v_s + v_r)$

$p(x \mid s)$: $G(s, v_r)$

que se puede determinar

$$p(s \mid x) = \frac{p(x \mid s)\, p(s)}{p(x)}$$

á gaussiana, al serlo las demás:

$$p(s\,|\,x) = \frac{1}{\sqrt{2\pi}}\,\frac{\dfrac{1}{\sqrt{v_r}}\,\dfrac{1}{\sqrt{v_s}}}{\dfrac{1}{\sqrt{v_s+v_r}}}\,exp\left[-\frac{1}{2}\left(\frac{(x-s)^2}{v_r}+\frac{s^2}{v_s}-\frac{x^2}{v_s+v_r}\right)\right]$$

tificando exponentes:

$$\frac{\left(s-E\{s\,|\,x\}\right)^2}{v_{s|x}} = \frac{(x-s)^2}{v_r}+\frac{s^2}{v_s}-\frac{x^2}{v_s+v_r}$$

minos en s^2*:*

$$\frac{1}{v_{s|x}} = \frac{1}{v_r}+\frac{1}{v_s}; \qquad v_{s|x} = \frac{v_s v_r}{v_s+v_r}$$

minos en -2s:

$$\frac{E\{s\,|\,x\}}{v_{s|x}} = \frac{x}{v_r}; \qquad E\{s\,|\,x\} = \frac{v_{s|x}}{v_r}x = \frac{v_s}{v_s+v_r}x$$

preciso seguir)

odo que

$$\hat{s}_{ms} = \frac{v_s}{v_s + v_r} x$$

e lo que puede observarse que:

lineal

ribuye a la señal una parte de la observación de acuerdo con los valores e potencia de señal y ruido.

Debe recordarse también que, dada la gaussianidad, el valor encontrado es mbién $\hat{s}_{ma}$ y $\hat{s}_{map}$.

io de Ampliación

no problema en el caso multidimensional toma la forma general

$$\mathbf{x} = \mathbf{Hs}+\mathbf{r}$$

$\mathbf{s}$ (N×1) es G($\mathbf{0}$,$\mathbf{V}_s$), $\mathbf{r}$ (M×1) es G($\mathbf{0}$,$\mathbf{V}_r$), y $\mathbf{H}$ es una matriz M×N, que se rá conocida.

Establecer $\hat{\mathbf{s}}_{ms}$

lo que

$\mathbf{x}$: $G(\mathbf{0}, \mathbf{HV}_s\mathbf{H}^T+\mathbf{V}_r)$

$\mathbf{x}|\mathbf{s}$: $G(\mathbf{Hs}, \mathbf{V}_r)$

diendo, como en el ejercicio anterior, a identificar exponentes:

$$E\{\mathbf{s}\,|\,\mathbf{x}\}\big]^T \mathbf{V}_{s|x}^{-1}\big[\mathbf{s} - E\{\mathbf{s}\,|\,\mathbf{x}\}\big] =$$

$$= \left(\mathbf{x}\text{-}\mathbf{Hs}\right)^T \mathbf{V}_r^{-1}\left(\mathbf{x}\text{-}\mathbf{Hs}\right) + \mathbf{s}^T\mathbf{V}_s^{-1}\mathbf{s} - \mathbf{x}^T\left(\mathbf{HV}_s\mathbf{H}^T + \mathbf{V}_r\right)^{-1}\mathbf{x}$$

inos cuadráticos:

$$\mathbf{s}^T\mathbf{V}_{s|x}^{-1}\mathbf{s} = (\mathbf{H}\mathbf{s})^T\mathbf{V}_r^{-1}\mathbf{H}\mathbf{s} + \mathbf{s}^T\mathbf{V}_s^{-1}\mathbf{s}$$

onde:

$$\mathbf{V}_{s|x}^{-1} = \mathbf{H}^T\mathbf{V}_r^{-1}\mathbf{H} + \mathbf{V}_s^{-1}$$

$$\mathbf{V}_{s|x} = \left(\mathbf{H}^T\mathbf{V}_r^{-1}\mathbf{H} + \mathbf{V}_s^{-1}\right)^{-1}$$

inos lineales (sin -2):

$$E^T\{\mathbf{s}|\mathbf{x}\}\mathbf{V}_{s|x}^{-1} = \mathbf{x}^T\mathbf{V}_r^{-1}\mathbf{H}$$

onde:

$$E^T\{\mathbf{s}\,|\,\mathbf{x}\} = \mathbf{x}^T\mathbf{V}_r^{-1}\mathbf{H}\mathbf{V}_{s|x}$$

$$E\{\mathbf{s}\,|\,\mathbf{x}\} = \mathbf{V}_{s|x}\mathbf{H}^T\mathbf{V}_r^{-1}\mathbf{x}$$

tituyendo $\mathbf{V}_{s|x}$*:*

$$\hat{s}_{ms} = \left(\mathbf{H}^T\mathbf{V}_r^{-1}\mathbf{H} + \mathbf{V}_s^{-1}\right)^{-1}\mathbf{H}^T\mathbf{V}_r^{-1}\mathbf{x}$$

le notarse lo mismo que en el caso unidimensional: en particular, que el lor es una arquitectura lineal (forma $\mathbf{w}^T\mathbf{x}$*).*

laro que, si **H=I** *(implica M=N), la solución tiene el mismo sentido de to" que el en caso unidimensional.*

<u>*io*</u>

e y discuta el caso de (M) observaciones unidimensionales directas de es independientes en ruido de valores independientes.

*a de un caso con **H=I***

con matrices de autocovarianza diagonales: $[\mathbf{D}_s]_{ii} = v_{si}$

$[\mathbf{D}_r]_{ii} = v_{ri}$

$\mathbf{V}_{s|x}$: $\quad \mathbf{D}_{s|x} = (\mathbf{D}_s^{-1} + \mathbf{D}_r^{-1})^{-1}$

$$[\mathbf{D}_{s|x}]_{ii} = \left(\frac{1}{v_{si}} + \frac{1}{v_{ri}}\right)^{-1} = \frac{v_{si} v_{ri}}{v_{si} + v_{ri}}$$

que

$$\hat{\mathbf{s}}_{ms} = \mathbf{D}_{s|x} \mathbf{D}_r^{-1} \mathbf{x}$$

omo elementos

$$\hat{s}_{ms\,i} = \frac{v_{si}}{v_{si} + v_{ri}} x_i$$

abía esperar, extensión directa del caso unidimensional.

o

y discuta el caso de (M) observaciones de la misma variable en ruido de valores dientes.

caso, $\mathbf{H} = \mathbf{1}\ (M\times 1)$

$\mathbf{V}_s = v_s$

$[\mathbf{D}_r]_{ii} = v_{ri}$

$$V_{s|x} = \left(\mathbf{1}^T \mathbf{D}_r^{-1} \mathbf{1} + v_s^{-1}\right)^{-1} = \frac{1}{v_s^{-1} + \sum_i v_{ri}^{-1}}$$

$$\hat{s}_{ms} = \frac{1}{v_s^{-1} + \sum_i v_{ri}^{-1}} \mathbf{1}^T \mathbf{D}_r^{-1} \mathbf{x} = \frac{1}{v_s^{-1} + \sum_i v_{ri}^{-1}} \sum_i v_{ri}^{-1} x_i$$

na media muestral ponderada; si los ruidos son i. d., $v_{ri} = v_r$

$$\hat{s}_{ms} = \frac{1}{v_s^{-1} + M v_r^{-1}} v_r^{-1} \sum_i x_i = \frac{1}{M + v_r v_s^{-1}} \sum_i x_i$$

na media muestral modificada (según potencias: teniendo en cuenta la cia de la señal y la incoherencia del ruido).

io de Ampliación

na general del problema de decidir entre dos señales deterministas idas) en presencia de ruido gaussiano (conocido) es

$$H_0 : \mathbf{x} = \mathbf{m}_0 + \mathbf{n}_0$$

$$H_1 : \mathbf{x} = \mathbf{m}_1 + \mathbf{n}_1$$

ruidos $\mathbf{n}_i : G(\mathbf{0}, \mathbf{V}_i)$

Formule el correspondiente LRT.

ando que $$p(\mathbf{x} \mid H_i) = \frac{1}{(2\pi)^{N/2} |\mathbf{V}_i|^{1/2}} \exp\left[-\frac{1}{2} (\mathbf{x} - \mathbf{m}_i)^T \mathbf{V}_i^{-1} (\mathbf{x} - \mathbf{m}_i) \right]$$

lo neperianos y eliminando términos comunes

$$\ldots_1) \ldots_0) = \ln\left(\frac{|\mathbf{V}_0|}{|\mathbf{V}_1|} \right)^{1/2} - \frac{1}{2}\left[(\mathbf{x} - \mathbf{m}_1)^T \mathbf{V}_1^{-1} (\mathbf{x} - \mathbf{m}_1) - (\mathbf{x} - \mathbf{m}_0)^T \mathbf{V}_0^{-1} (\mathbf{x} - \mathbf{m}_0) \right] \underset{D_0}{\overset{D_1}{\gtrless}}$$

$$\underset{D_0}{\overset{D_1}{\gtrless}} \ln \frac{(C_{10} - C_{00}) Pr(H_0)}{(C_{01} - C_{11}) Pr(H_1)} = \ln \eta$$

…uede reescribir

$$\ldots m_0)^T V_0^{-1}(\mathbf{x}-\mathbf{m}_0)-(\mathbf{x}-\mathbf{m}_1)^T V_1^{-1}(\mathbf{x}-\mathbf{m}_1) \underset{D_0}{\overset{D_1}{\gtrless}} 2\ln\eta-\ln\frac{|\mathbf{V}_0|}{|\mathbf{V}_1|} = \eta_1$$

…ne una frontera (general) ***cuadrática*** *(e igual tipo de* ***discriminantes****).*

…s formas ponderadas $(\mathbf{x}-\mathbf{m}_i)^T \mathbf{V}_i^{-1}(\mathbf{x}-\mathbf{m}_i)$*, obviamente generalizaciones …(cuadrados de) las distancias euclídeas entre* $\mathbf{x}$ *y* $\mathbf{m}_i$*, se denominan …***as de Mahalanobis***: la decisión se realiza en función de ellas.*

io de Ampliación

e y discuta el problema anterior si el ruido no depende de la señal presente.

$_1$ $= \boldsymbol{V_0} = \boldsymbol{V}$; *operando, desaparecen los términos cuadráticos y queda*

$$-2\mathbf{m}_0^T\mathbf{V}^{-1}\mathbf{x} + \mathbf{m}_0^T\mathbf{V}^{-1}\mathbf{m}_0 + 2\mathbf{m}_1^T\mathbf{V}^{-1}\mathbf{x} - \mathbf{m}_1^T\mathbf{V}^{-1}\mathbf{m}_1 \underset{D_0}{\overset{D_1}{\gtrless}} \eta_1$$

$$(\mathbf{m}_1 - \mathbf{m}_0)^T\mathbf{V}^{-1}\mathbf{x} \underset{D_0}{\overset{D_1}{\gtrless}} \frac{1}{2}\left(\eta_1 + \mathbf{m}_1^T\mathbf{V}^{-1}\mathbf{m}_1 - \mathbf{m}_0^T\mathbf{V}^{-1}\mathbf{m}_0\right) = \eta_2$$

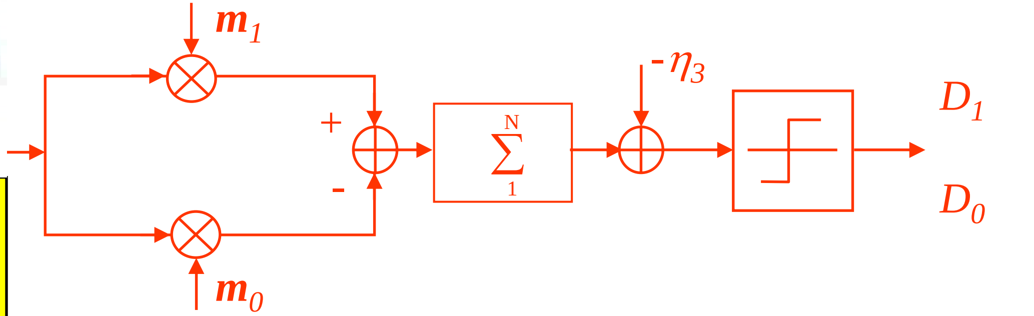
un <u>decisor lineal</u> $\quad \mathbf{w}^T\mathbf{x} \underset{D_0}{\overset{D_1}{\gtrless}} \eta_2$

ás las componentes del ruido son iid: $\quad \mathbf{V} = v\mathbf{I}$

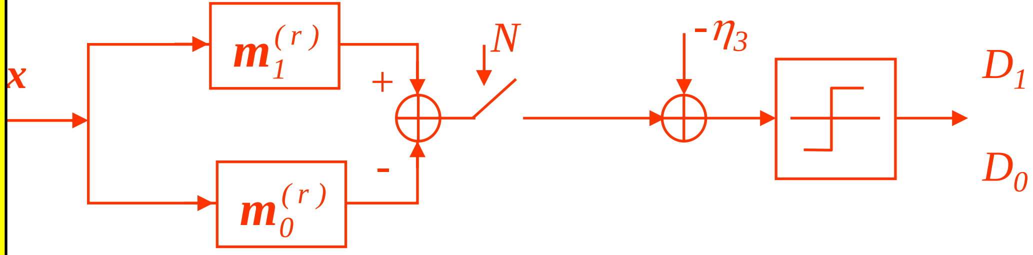
$$(\mathbf{m}_1 - \mathbf{m}_0)^T v^{-1}\mathbf{I}\mathbf{x} \underset{D_0}{\overset{D_1}{\gtrless}} \eta_2$$

$$\ _0)^T\mathbf{x} \underset{D_0}{\overset{D_1}{\gtrless}} v\eta_2 = \eta_3 \qquad \left(= \frac{1}{2}\left[2v\ln\eta + \|m_1\|_2^2 - \|m_0\|_2^2\right]\right)$$

mo se sabe, es realizable como "decisor de correlación"



iltro adaptado"



i, además $\mathbf{m}_0 = \mathbf{0}$ (<u>detección</u> de $\mathbf{m}_1$): desaparecen las ramas

pondientes; el umbral pasa a $\eta_4 = \frac{1}{2}\left(2\nu \ln \eta + \|\mathbf{m}_1\|_2^2\right)$

o ML, queda $\frac{1}{2}\|\mathbf{m}_1\|_2^2$ *: se pide que* $\mathbf{m}_1^T \mathbf{x}$ *llegue a la mitad de energía de*

io de Ampliación

e y discuta la decisión (binaria) gaussiana cuando se trata de la misma n distintos ruidos.

*$_0$ = **m**: siguen empleándose las distancias de Mahalanobis*

$$(\mathbf{x}-\mathbf{m})^T\left(\mathbf{V}_0^{-1}-\mathbf{V}_1^{-1}\right)(\mathbf{x}-\mathbf{m}) \underset{D_0}{\overset{D_1}{\gtrless}} \eta_1$$

demás, los ruidos son iid: $\mathbf{V}_i = v_i\mathbf{I}$

$$(\mathbf{x}-\mathbf{m})^T\left(\frac{1}{v_0}-\frac{1}{v_1}\right)\mathbf{I}(\mathbf{x}-\mathbf{m})$$

nitiendo $v_1>v_0$ (si no, voltean las desigualdades)

$$\|\mathbf{x}-\mathbf{m}\|_2^2 \underset{D_0}{\overset{D_1}{\gtrless}} \frac{v_0 v_1}{v_1-v_0}\eta_1$$

isor por distancia euclídea

*demás **m** = **0***   : $$\|\mathbf{x}\|_2^2 \underset{D_0}{\overset{D_1}{\gtrless}} \frac{v_0 v_1}{v_1-v_0}\eta_1$$

decide por energía entre las dos opciones de varianza del ruido.

ticos suficientes

han visto casos en los que no es preciso el conocimiento del conjunto de ervaciones (o de las variables) para realizar la estimación o la decisión:

imar m y v de una gaussiana requiere sólo $\sum_k x^{(k)}, \sum_k x^{(k)2}$

imar una gaussiana en otra gaussiana requiere formas del tipo $\mathbf{w}^T\mathbf{x}$: eales también

cidir qué señal determinista es la que se encuentra en ruido gaussiano quiere formas cuadráticas (o lineales)

Estas formas se llaman **estadísticos suficientes**: conocerlos permite nplificar procesos de estimación o decisión.

ıa de Factorización

ltado de una transformación de las observaciones $\mathbf{t} = \mathbf{t}(\mathbf{x})$ es un estadístico nte para estimar (ML) el parámetro determinista $\mathbf{s}$ si y sólo si

$$p(\mathbf{x} \mid \mathbf{s}) = g(\mathbf{t}, \mathbf{s})\, h(\mathbf{x}, \mathbf{t})$$

scomposición no es única: $g'(\mathbf{t}, \mathbf{s}) = g(\mathbf{t}, \mathbf{s})\, f(\mathbf{t})$

$h'(\mathbf{x}, \mathbf{t}) = h(\mathbf{x}, \mathbf{t}) / f(\mathbf{t})$)

/a que

$$\arg[\max_{\mathbf{s}} p(\mathbf{x} \mid \mathbf{s})] = \arg[\max_{\mathbf{s}} g(\mathbf{t}, \mathbf{s})]$$

omprende fácilmente la ventaja si g es sencilla); y no ocurre en otro caso.

rmulación equivalente es:

$$\mathrm{p}(\mathbf{x} \mid \mathbf{s}, \mathbf{t}) = \mathrm{p}(\mathbf{x} \mid \mathbf{t})$$

nte: si se da la equivalente,

$$\mathrm{p}(\mathbf{x} \mid \mathbf{s}) = \mathrm{p}(\mathbf{x} \mid \mathbf{s}, \mathbf{t})\, \mathrm{p}(\mathbf{t} \mid \mathbf{s}) = \mathrm{p}(\mathbf{x} \mid \mathbf{t})\, \mathrm{p}(\mathbf{t} \mid \mathbf{s})$$

eden identificar los factores con h y g, respectivamente.

ria: suponemos que se cumple el Teorema de Factorización; como

$$\mathrm{p}(\mathbf{x} \mid \mathbf{s}, \mathbf{t}) = \frac{\mathrm{p}(\mathbf{x}, \mathbf{t} \mid \mathbf{s})}{\mathrm{p}(\mathbf{t} \mid \mathbf{s})} = \frac{\mathrm{p}(\mathbf{x}, \mathbf{t} \mid \mathbf{s})}{\int \mathrm{p}(\mathbf{x}, \mathbf{t} \mid \mathbf{s})\, d\mathbf{x}}$$

p($\mathbf{x}$, $\mathbf{t}$ | $\mathbf{s}$) es p($\mathbf{x}$ | $\mathbf{s}$), al ser $\mathbf{t}$ una función de $\mathbf{x}$, se tiene

$$p(\mathbf{x}\,|\,\mathbf{s},\mathbf{t}) \;=\; \frac{p(\mathbf{x}\,|\,\mathbf{s})}{\displaystyle\int p(\mathbf{x}\,|\,\mathbf{s})\,d\mathbf{x}}$$

do el Teorema de Factorización

$$\mathbf{x}\,|\,\mathbf{s},\mathbf{t}) \;=\; \frac{g(\mathbf{t},\mathbf{s})\,h(\mathbf{x},\mathbf{t})}{\displaystyle\int g(\mathbf{t},\mathbf{s})h(\mathbf{x},\mathbf{t})\,d\mathbf{x}} \;=\; \frac{h(\mathbf{x},\mathbf{t})}{\displaystyle\int h(\mathbf{x},\mathbf{t})\,d\mathbf{x}}$$

depende de **s**, y que, por identificación, resulta ser p(**x** | **t**).

En ocasiones **t** se obtiene de una transformación uno a uno de los datos:

i; el conjunto **i** es un **estadístico irrelevante**.

e estimación de parámetro aleatorio

e siendo válido el Teorema de Factorización, ya que de p(**x** | **s**, **t**) = p(**x** | **t**)
e

$$p(\mathbf{s}\,|\,\mathbf{x}) = p(\mathbf{s}\,|\,\mathbf{x},\mathbf{t}) = \frac{p(\mathbf{s},\mathbf{x}\,|\,\mathbf{t})}{p(\mathbf{x}\,|\,\mathbf{t})} = \frac{p(\mathbf{x}\,|\,\mathbf{s},\mathbf{t})p(\mathbf{s}\,|\,\mathbf{t})}{p(\mathbf{x}\,|\,\mathbf{t})} = p(\mathbf{s}\,|\,\mathbf{t})$$

indica claramente que, dado **t**, los datos originales no se necesitan para
r cualquier estimación Bayesiana

$$\min_{\hat{\mathbf{s}}} \int C\left(\mathbf{s},\hat{\mathbf{s}}\right) p\left(\mathbf{s}\,|\,\mathbf{x}\right) d\mathbf{s}$$

e decisión

Es obvio que la condición equivalente es

$$p(\mathbf{x} \mid H_i) = g(\mathbf{t}, H_i)\, h(\mathbf{x}, \mathbf{t})$$

ién

$$\Pr(H_i \mid \mathbf{x}) \; (= \Pr(H_i \mid \mathbf{x}, \mathbf{t})) \; = \; \Pr(H_i \mid \mathbf{t})$$

basta $g(\mathbf{t}, H_i)$ o $\Pr(H_i \mid \mathbf{t})$ para la decisión Bayesiana.

io

inar un estadístico suficiente para la estimación del parámetro de una
ución exponencial unilateral (Maxwell), $p(x) = a \exp(-ax)\, u(x)$ $(a > 0)$,
te K observaciones independientes.

$$\mathbf{x} \mid a) = \prod_{k=1}^{K} a \exp\left(-ax^{(k)}\right) , \qquad \{ \mathbf{x} \} \geq 0$$

$$(\mathbf{x} \mid a) = \sum_{k=1}^{K} \left[\ln a - ax^{(k)} \right] = K \ln a - a \sum_{k=1}^{K} x^{(k)}$$

observa que $\sum_{k=1}^{K} x^{(k)}$ es un estadístico suficiente...

$$= \frac{K}{a} - \sum_{k=1}^{K} x^{(k)}; \quad \left. \frac{\partial \ln}{\partial a} \right|_{a=\hat{a}_{ml}} = 0 = \frac{K}{\hat{a}_{ml}} - \sum_{k=1}^{K} x^{(k)} \quad \left(\frac{\partial^2 \ln}{\partial a^2} = -\frac{K}{a^2} < 0 \right)$$

$$= \frac{K}{\sum_{k=1}^{K} x^{(k)}}$$

ese que en este caso la propia p(x | a) tiene la forma de g(t, a), con

$$t = \sum_{k=1}^{K} x^{(k)}$$

se ve claramente en las expresiones que resultan de tomar logaritmos
la factorización se convierte en separación de sumandos):

$$\ln p(\mathbf{x} \mid a) = K \ln a - a \sum_{k=1}^{K} x^{(k)} = K \ln a - at \qquad (= g(t, a))$$