OPTIMAL FILTERING, 2012-09-03
DUE, 2012-09-10
MATS BENGTSSON/BJÖRN OTTERSTEN

Homework # 2

1. Let X and Y be two jointly distributed scalar random variables with Y taking on the value y. Let $\hat{x}$ be an estimate chosen so that

$$\mathrm{E}\{|X - \hat{x}| \ \mid Y = y\} \leq \mathrm{E}\{|X - z| \ \mid Y = y\} \, .$$

In other words, $\hat{x}$ is chosen to minimize the average value of the *absolute error* between $\hat{x}$ and the actual value taken by X. Show that $\hat{x}$ is the median of the conditional density $f_{X|Y}(x|y)$. The median, α, of a continuous density $p_A(a)$ satisfies $P(A \leq \alpha) = P(A \geq \alpha)$.

2. Given a sequence of zero-mean random variables $\{y_0, \ y_1, \ \ldots\}$ let

$$\begin{aligned} e_i &= y_i - \hat{y}_{i|i-1}, \qquad \hat{y}_{0|-1} = 0 \\ \hat{y}_{i|i-1} &= \text{ l.l.s.e. of } \ y_i \ \text{ given } \ y_0, \ y_1, \ldots, y_{i-1} \, . \end{aligned}$$

a) Show that the e_i are orthogonal random variables (Hint: recall Gram-Schmidt).

b) Show that if $\mathrm{E}e_i^2 > 0$ for all i, then the vectors

$$\boldsymbol{y}_n = \begin{bmatrix} y_n \\ \vdots \\ y_1 \\ y_0 \end{bmatrix} , \qquad \boldsymbol{e}_n = \begin{bmatrix} e_n \\ \vdots \\ e_1 \\ e_0 \end{bmatrix}$$

can be related for all n by a nonsingular triangular matrix

$$\boldsymbol{y}_n = \boldsymbol{T}_n \boldsymbol{e}_n \, .$$

c) If $\boldsymbol{H}_n$ is a linear operation that yields $\hat{\boldsymbol{y}}_n$ from $\boldsymbol{y}_n$ show that

$$\boldsymbol{H}_n = \boldsymbol{I} - \boldsymbol{T}_n^{-1} \, , \qquad \hat{\boldsymbol{y}}_n = \boldsymbol{y}_n - \boldsymbol{e}_n \, .$$

d) Let $r_{ye}(i,j) = \mathrm{E}\{y_i e_j^*\}$. Show that

$$\boldsymbol{T}_n = \begin{bmatrix} 1 & r_{ye}(n, n-1) r_{ee}^{-1}(n-1, n-1) & \cdots & r_{ye}(n,0) r_{ee}^{-1}(0,0) \\ 0 & 1 & \cdots & \\ & \ddots & \ddots & \\ & & 0 & 1 \end{bmatrix} .$$

e) Show that

$$\boldsymbol{R}_{\boldsymbol{y}} = \boldsymbol{T}_n \boldsymbol{R}_{\boldsymbol{e}} \boldsymbol{T}_n^*$$

where $\boldsymbol{R}_{\boldsymbol{y}} = \mathrm{E}\{\boldsymbol{y}_n \boldsymbol{y}_n^*\}$, $\boldsymbol{R}_{\boldsymbol{e}} = \mathrm{E}\{\boldsymbol{e}_n \boldsymbol{e}_n^*\}$.

f) Let $\boldsymbol{x}$ be a zero-mean stochastic vector related to the random process y_i and let $\hat{\boldsymbol{x}}_n$ denote the l.l.s.e. of $\boldsymbol{x}$ based on the observations of y_i up to $i = n$. Show that

$$\hat{\boldsymbol{x}}_n = \boldsymbol{R}_{\boldsymbol{xy}} \boldsymbol{R}_{\boldsymbol{y}}^{-1} \boldsymbol{y}_n = \boldsymbol{R}_{\boldsymbol{x}_n \boldsymbol{e}} \boldsymbol{R}_{\boldsymbol{e}}^{-1} \boldsymbol{e}_n \ .$$

g) Show that

$$\hat{\boldsymbol{x}}_{n+1} = \hat{\boldsymbol{x}}_n + \mathrm{E}\{\boldsymbol{x} e_{n+1}^*\} r_{ee}^{-1}(n+1, n+1) e_{n+1} \ .$$

3. Let $\boldsymbol{y} = \boldsymbol{h}\boldsymbol{x} + \boldsymbol{v}$ where

$$\begin{aligned} \mathrm{E}\{\boldsymbol{x}\boldsymbol{x}^*\} &= \boldsymbol{\Pi} \\ \mathrm{E}\{\boldsymbol{x}\boldsymbol{v}^*\} &= 0 \\ \mathrm{E}\{\boldsymbol{v}\boldsymbol{v}^*\} &= \boldsymbol{R} \\ \mathrm{E}\{\boldsymbol{y}\boldsymbol{y}^*\} &= \boldsymbol{R}_y, \qquad |\boldsymbol{R}_y| \neq 0 \ . \end{aligned}$$

a) Show that the l.l.s.e. of $\boldsymbol{x}$ can be written

$$\hat{\boldsymbol{x}} = \boldsymbol{\Pi} \boldsymbol{h}^* (\boldsymbol{R} + \boldsymbol{h} \boldsymbol{\Pi} \boldsymbol{h}^*)^{-1} \boldsymbol{y} \ .$$

b) If $|\boldsymbol{R}| \neq 0$ and $|\boldsymbol{\Pi}| \neq 0$ show that

$$\hat{\boldsymbol{x}} = (\boldsymbol{\Pi}^{-1} + \boldsymbol{h}^* \boldsymbol{R}^{-1} \boldsymbol{h})^{-1} \boldsymbol{h}^* \boldsymbol{R}^{-1} \boldsymbol{y} \ .$$

4. Let x be a zero-mean (non-Gaussian) random variable with moments

$$\mathrm{E} x^n = \mu_n \ .$$

a) Find the l.l.s.e. of x^3 given x.

b) What is the l.s.e. of x^3 given x?

Note that in the above, when it says *linear* we even include *affine* mappings.

5. Let X and Y be jointly distributed stochastic variables and assume that we know the prior distribution $f_X(x)$ of X, and the conditional distribution $f_{Y|X}(y|x)$ of Y given X. Show that the (non-linear) l.s.e. of X given an observation y of Y is

$$\hat{x} = \frac{\int x f_{Y|X}(y|x) f_X(x)\, dx}{\int f_{Y|X}(y|x) f_X(x)\, dx}$$