Mean-Square Estimation

Estimation of a Random Variable by a Constant

It is some of interest to estimate a random variable by a constant. That is, we want to find a constant such that the error,

$$I = E\{(X-\alpha)^2\}$$

is minimum. It then follows that

$$I = E[(X-\alpha)^2] = E\{X^2\} - 2\alpha E\{X\} + \alpha^2 .$$

Minimizing I,

$$\frac{\partial I}{\partial \alpha} = 0$$

lead to

$$\alpha = E\{X\},$$

where $E\{X\}$ is the expected value of X .

i) Nonlinear Mean-Square Estimation

It is some of interest to estimate a random variable as function of another random variable. That is, we want to estimate random variable Y by a function $g(X)$ such tha the error,

$$E\{[Y-g(X)]^2\} \text{ is minimum.}$$

Theorem: It may be shown (see page 217 of Papoulis for proof), minimizing the error leads to

$$g(X) = E\{Y \mid X\}.$$

ii) Linear Mean-Square Estimation

Assume $g(X) = aX + b$.

Theorem: When the joint statistics of *X* and *Y* are known, it may be shown that the parameters of the linear mean-square estimation are given by

$$a = \frac{r\sigma_Y}{\sigma_X}, \qquad b = E\{Y\} - aE\{X\}$$

and the minimum error e_m is given by

$$e_m = \sigma_Y^2\left(1 - r^2\right).$$

Here r is the correlation coefficient defined by

$$r = E\frac{\{(X-\eta_X)(Y-\eta_Y)\}}{\sigma_X\sigma_Y}.$$

If $\eta_X = \eta_Y = 0$, then $b = 0$,

$$a = \frac{E\{XY\}}{E\{X^2\}}$$

and

$$e_m = E\{Y^2\} - E\{(aX)^2\}$$

Note that a minimizes $E\{(Y-aX)^2\}$. i.e., $E\{(Y-aX)X\} = 0$. This means X is orthogonal to $Y - aX$ and $e_m = E\{(Y-aX)Y\}$.

Theorem:

If X and Y are jointly normal, then nonlinear and linear mean-square estimation of Y in terms of X leads to identical solution. i.e.,

$$E\{Y \mid X\} = aX, \qquad a = E\frac{\{XY\}}{E\{X^2\}}.$$

iii) Mean-Square Estimation (Several Random Variables)

Find estimate of random variable X_0 in terms of X_1, X_2,… X_n.

Minimizing the error

$$E\{[X_0 - g(X_1,\dots,X_n)]^2\},$$

it follows that

$$X_0 = g(X_1,\dots,X_n) = E\{X_0 \mid X_1,\dots,X_n\}.$$

iv) Linear Mean-Square Estimation

Assuming is a linear function. That is,

$$g = a_1 X_1 + \ldots + a_n X_n = \sum_i a_i X_i \ .$$

Then minimizing the estimation error leads to

$$E\left\{\left(X_0 - \sum_i a_i X_i\right) X_j\right\} = 0$$

and

$$R_{0j} = \sum_i R_{ji} a_i \quad \text{with} \quad j = 1,\ldots,n \,,$$

which can be solved for finding a_i.

v) Jointly Normal Random Variables

If X_0, X_1, … X_n are jointly normal, then the linear mean-square estimation becomes identical to the best nonlinear mean-square estimation. That is,

$$E\{X_0 \mid X_1,\ldots,X_n\} = \sum_i a_i X_i \ .$$