PHY2506 Lecture 7 21/10/15

Estimation Theory (Polavarapu Ch 5.1-5.2)

Consider an observation model

$$z = Hx + v$$

v = obs error
H = obs operator

the joint probability

$$p_{x,z}(x, z) = p_{z|x}(z|x)\, p_x(x)$$
$$= p_{x|z}(x|z)\, p_z(z)$$

In DA the goal is to obtain information about the model state given observations z.

This information is given by:

$$p_{x|z}(x|z) = \frac{p_{x,z}(x,z)}{p_z(z)}$$

Bayes Theorem

How do we use $p_{x|z}(x|z)$ to estimate x?

(we don't have full knowledge of $p(x|z)$)

Let's define $\hat{x} = \hat{x}(z)$, the estimate of the the model state that depends on z.

$\hat{x}$ = estimator

If unbiased $E[x^t - \hat{x}] = 0$

If biased $b(\hat{x}) = E[x^t - \hat{x}] = E[x^t] - \hat{x}$

The expectation (mean) of the estimator is the estimator since it depends only on z (the expectation is with respect to x).

How do you choose the best estimator?

If we have two estimators with bias, the one with the least variance is not necessarily the one with the least bias. In this case we want the one with the lowest mean square error (MSE).

$$MSE = E[(x^t - \hat{x})^2]$$

$$= E[(x^t - E[x^t] + E[x^t] - \hat{x})^2]$$

$$= E[(x^t - E[x^t] + b(\hat{x}))^2]$$

$$= E[(x^t - E[x^t])^2] + b^2(\hat{x}) + 2E[x^t - E[x^t]]\, b(\hat{x})$$

$$= E[(x^t - E[x^t])^2 + b^2(\hat{x})$$

$\Rightarrow$ $MSE = \text{Variance}(x^t) + b^2(\hat{x})$

So we want the smallest MSE.

For an unbiased estimator, MSE = Variance

What is "best" with respect to the estimator?

Define a risk function J and minimize the risk (i.e., the expected value of the risk function)

$$F(\hat{x}) = E[J(\tilde{x})] = \int_{-\infty}^{\infty} J(\tilde{x})\, p_x(x)\, dx$$

$J(\tilde{x})$ = the risk function (cost function)

$\tilde{x} = x - \hat{x}$ the error of the estimator

recall that

$$p_x(x) = \int_{-\infty}^{\infty} p_{x,z}(x,z)\, dz$$

$$\Rightarrow \quad F(\hat{x}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J(\tilde{x})\, p_{x,z}(x,z)\, dz\, dx$$

minimize WRT $\hat{x}$

We want small J for $\tilde{x} \approx 0$ and large J for large $\tilde{x}$

Common Cost Functions

1) Quadratic cost function $J(\tilde{x}) = \tilde{x}^T S \tilde{x}$

S = positive definite, symmetric

2) Uniform cost function

$$J(\tilde{x}) = \begin{cases} 0 & |\tilde{x}| < \epsilon \\ \frac{1}{2\epsilon} & |\tilde{x}| \geq \epsilon \end{cases}$$

The minimum variance estimator minimizes a quadratic cost function $\tilde{x} = x - \hat{x}$ and $J(\tilde{x}) = (x-\hat{x})^T S (x-\hat{x})$

We want to minimize

$$F(\hat{x}) = \int_{-\infty}^{\infty} \Bigg[\underbrace{\int_{-\infty}^{\infty} (x-\hat{x})^T S (x-\hat{x}) \, p_{x|z}(x|z)\, dx}_{\text{has only } \hat{x} \text{ dependance}} \Bigg] p(z)\, dz$$

So we minimize WRT $\hat{x}$

$$F_{MV}(\hat{x} \mid z) = \int_{-\infty}^{\infty} (x-\hat{x})^T S (x-\hat{x})\, p_{x|z}(x|z)\, dx$$

conditional Bayes risk

$$\left.\frac{\partial F}{\partial \hat{x}}\right|_{\hat{x}=\hat{x}_{MV}} = -2\left(S \int_{-\infty}^{\infty} (x-\hat{x})\, p_{x|z}(x|z)\, dx \right) = 0$$

$$\text{so} \quad \int_{-\infty}^{\infty} x\, p_{x|z}(x|z)\, dx = \int_{-\infty}^{\infty} \hat{x}_{MV}\, p_{x|z}(x|z)\, dx$$

but 1) $$\int_{-\infty}^{\infty} \hat{x}_{MV}\, p_{x|z}(x|z)\, dx = \hat{x}_{MV} \underbrace{\int_{-\infty}^{\infty} p_{x|z}(x|z)\, dx}_{=1}$$

$$= \hat{x}_{MV}$$

2) $$\int_{-\infty}^{\infty} x\, p_{x|z}(x|z)\, dx = E[x|z]$$

conditional expectation

$$\Rightarrow \boxed{\hat{x}_{MV} = E[x|z]}$$

the minimum variance estimate is the mean of $P_{x|z}(x|z)$

$P_{x|z}(x|z)$ is the a posteriori pdf since it is obtained after the observations are known.

Note

1) expectation $E[E[x|z]] = E[x]$

$$E(x) = \int_{-\infty}^{\infty} x\, P_x(x)\, dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, P_{x,z}(x,z)\, dz\, dx$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, P_{x|z}(x|z)\, P(z)\, dz\, dx$$

$$= \int_{-\infty}^{\infty} P(z)\, dz \int_{-\infty}^{\infty} x\, P_{x|z}(x|z)\, dx$$

$$= \int_{-\infty}^{\infty} E[x|z]\, P_z(z)\, dz$$

$$= E[E[x|z]]$$

2) The conditional variance is given by

$$= E\left[\left(x - E[x|z]\right)\left(x - E[x|z]\right) \middle| z\right]$$

3) The MV estimator is unbiased

$$E(\tilde{x}) = E(x - \hat{x}_{MV}) = E\left[x - E[x|z]\right]$$
$$= E[x] - E\left[E[x|z]\right]$$
$$= E[x] - E[x] = 0$$

4) $$\frac{\partial^2 F_{MV}(\hat{x}|z)}{\partial \hat{x}^2} = 2S$$

and since S is positive definite $\Rightarrow$ the solution is a minimum

5) The MV estimator is equivalent to taking the conditional mean without having to make any assumptions about the nature of the pdf.

6) The conditional mean is independent of S.