Lecture 6

Optimal Estimation Theory and Its Applications

Prof. Yan LIANG

liangyan@nwpu.edu.cn

School of Automation,

Northwestern Polytechnical University

Autumn,2019

Outline

- Bayesian estimate
- Minimum Mean Square Error estimate

Bayesian estimate

Example3.4:

Further Consideration

If Fish 1 is sorted as Fish 2, the faced quality problem is "sell seconds at best quality prices". It is much severe with the cost being C1.

If Fish 2 is sorted as Fish 1, the faced problem is economic loss with the cost being C2.

In general, C1>C2

Requirement

Please design a fish sorting system to minimize the total sorting cost.

Bayesian estimate

Suppose $\mathrm{Z} = \mathrm{Z}_{\mathrm{BE}}$ is the classification threshold, i.e.,

If z$<Z_{\mathrm{BE}}$, then Fish 1;Fish 2 otherwise.

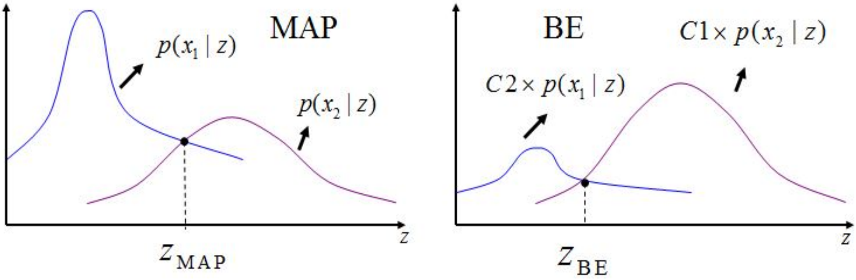
The total cost is

$$F(Z_{\mathrm{BE}}) = \int_{\mathrm{z} \geq Z_{\mathrm{BE}}} C2 \times p(x_1|z)p(z)dz + \int_{\mathrm{z} < Z_{\mathrm{BE}}} C1 \times p(x_2|z)p(z)dz$$

To minimize the total cost, Z_{BE} should satisfy

$$C2 \times p(X = x_1|z = Z_{\mathrm{BE}}) = C1 \times p(X = x_2|z = Z_{\mathrm{BE}})$$

Bayesian estimate



$Z_{\mathrm{BE}} < Z_{\mathrm{MAP}}$ means that the proposed estimate more prefers to Fish 1, compared with the MAP estimate.

Bayesian estimate

Consider the vector to be estimated X, its measurement Z, the estimate $\hat{X}(Z)$ and the estimate error $\tilde{X} = X - \hat{X}(Z)$

Construct a scalar function!

$$L(\tilde{X}) = L\left[X - \hat{X}(Z)\right] \geq 0$$

L is called as the loss function, and its expectation

$$B(\tilde{X}) = E\left[L(\tilde{X})\right] = \iint L\left[x - \hat{X}(z)\right] p(x,z)\,dxdz$$

is called as the Bayesian risk. The Bayesian estimate aims at minimizing the Bayesian risk.

Bayesian estimate

$$L = \begin{cases} C2 & \text{if} \quad \mathrm{z} > \mathrm{Z_{BE}} \\ C1 & \text{otherwise} \end{cases}$$ in the fish sorting example considered the cost

If X belongs to the R^n space, the Bayesian cost is always defined satisfying

$$\begin{cases} 1. L(X) = 0, & \text{if} \quad X = 0 \\ 2. L(X_2) \geq L(X_1), & \text{if} \quad \|X_2\| \geq \|X_1\| \\ 3. L(X) = L(-X) \end{cases}$$

L is a convex function

$L(X)$ is nonnegative.

Bayesian estimate

If $L\left(\tilde{X}\right) = \tilde{X}^T \tilde{X} = \left(X - \hat{X}\right)^T \left(X - \hat{X}\right)$, the Bayesian estimate is the minimum-mean-square- error estimate.

If

$$L\left[X - \hat{X}(Z)\right] = \begin{cases} 0 & \text{if } \left\|X - \hat{X}(Z)\right\| < \frac{\varepsilon}{2} \\ \frac{1}{\varepsilon} & \text{if } \left\|X - \hat{X}(Z)\right\| \geq \frac{\varepsilon}{2} \end{cases}$$

then we have the following Bayesian risk.

Bayesian estimate

$$B\left(\tilde{X}\right) = E\left\{L\left[X - \hat{X}\left(Z\right)\right]\right\} = \int_{-\infty}^{+\infty} \int\limits_{\left\|X - \hat{X}\right\| \geq \frac{\varepsilon}{2}} \frac{1}{\varepsilon} p\left(x|z\right) dx p\left(z\right) dz$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\varepsilon} \left[1 - \int\limits_{\left\|X - \hat{X}\right\| < \frac{\varepsilon}{2}} p\left(x|z\right) dx\right] p\left(z\right) dz$$

$$\int\limits_{\|X\| \geq \frac{\varepsilon}{2}} p\left(x|z\right) dx + \int\limits_{\|X\| < \frac{\varepsilon}{2}} p\left(x|z\right) dx = 1$$

suppose $\hat{X}_B\left(Z\right)$ is the Bayesian estimate, i.e.,

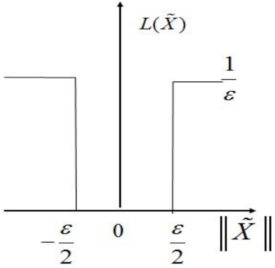
$$\hat{X}_B(Z) = \arg\min_{\hat{X}} B(\hat{X})$$

Bayesian estimate

$$\hat{X}_B(Z) = \arg\min_{\hat{X}} B(\hat{X})$$
$$\Leftrightarrow C_B(Z) = \arg\max_{\hat{X}} \int_{\|X-\hat{X}\|<\frac{\varepsilon}{2}} \frac{1}{\varepsilon}\, p(x|z)dx$$

As ε approaches zero, we have

$$\hat{X}_B(Z) = \arg\max_{x} p(x|z) \qquad \leftarrow \text{MAP estimate}$$



Minimum Mean Square Error estimate

Sometimes also called as Minimum Variance estimation

The estimate is the optimal in the sense for any a function of the measurement Z, $\hat{X}(Z)$, the following inequality holds

$$E\left\{\left[X-\hat{X}_{\mathrm{MMSE}}(Z)\right]^T\left[X-\hat{X}_{\mathrm{MMSE}}(Z)\right]\right\} \leq E\left\{\left[X-\hat{X}(Z)\right]^T\left[X-\hat{X}(Z)\right]\right\}$$

or the following matrix inequality holds

$$E\left\{\left[X-\hat{X}_{\mathrm{MV}}(Z)\right]\left[X-\hat{X}_{\mathrm{MV}}(Z)\right]^T\right\} \leq E\left\{\left[X-\hat{X}(Z)\right]\left[X-\hat{X}(Z)\right]^T\right\}$$

Minimum Mean Square Error estimate

On one hand, we have

$$
\begin{aligned}
&E\left\{\left[X-\hat{X}_{\mathrm{MV}}(Z)\right]\left[X-\hat{X}_{\mathrm{MV}}(Z)\right]^T\right\} \leq E\left\{\left[X-\hat{X}(Z)\right]\left[X-\hat{X}(Z)\right]^T\right\}\\
&\Rightarrow \mathrm{trace}\left\{E\left\{\left[X-\hat{X}_{\mathrm{MV}}(Z)\right]\left[X-\hat{X}_{\mathrm{MV}}(Z)\right]^T\right\}\right\} \leq \mathrm{trace}\left\{E\left\{\left[X-\hat{X}(Z)\right]\left[X-\hat{X}(Z)\right]^T\right\}\right\}\\
&\Rightarrow \mathrm{trace}\left\{E\left\{\left[X-\hat{X}_{\mathrm{MV}}(Z)\right]^T\left[X-\hat{X}_{\mathrm{MV}}(Z)\right]\right\}\right\} \leq \mathrm{trace}\left\{E\left\{\left[X-\hat{X}(Z)\right]^T\left[X-\hat{X}(Z)\right]\right\}\right\}\\
&\Rightarrow E\left\{\left[X-\hat{X}_{\mathrm{MV}}(Z)\right]^T\left[X-\hat{X}_{\mathrm{MV}}(Z)\right]\right\} \leq E\left\{\left[X-\hat{X}(Z)\right]^T\left[X-\hat{X}(Z)\right]\right\}
\end{aligned}
$$

Conclusion:the MV estimate is the MMSE estimate.

$\boxed{\text{trace\{AB\}=trace\{BA\}}}$

Minimum Mean Square Error estimate

On the other hand, we have

For any a estimate $\hat{X}(Z)$, its covariance matrix is

$$E(\tilde{X}\tilde{X}^T) = E\left[(X-\hat{X})(X-\hat{X})^T\right]$$
$$= \iint (x-\hat{X})(x-\hat{X})^T p(x,z)\,dxdz$$

$\hat{X}(z)$ is the function of z, instead of x.

Minimum Mean Square Error estimate

$$= \int_{-\infty}^{+\infty} p(z) \int_{-\infty}^{+\infty} \left[x - E(x|z) + \underline{E(x|z) - \hat{X}}\right] \times \left[x - E(x|z) + \underline{E(x|z) - \hat{X}}\right]^T p(x|z) dx dz$$

Notice $\int_{-\infty}^{+\infty} \underline{\left[E(x|Z) - \hat{X}\right]} [x - E(x|Z)]^T p(x|z) dx = 0 \quad \int_{-\infty}^{+\infty} p(x|z) dz = 1$

$$= \int_{-\infty}^{+\infty} \left\{ \mathrm{var}(x|z) + \left[E(x|z) - \hat{X}\right] \left[E(x|z) - \hat{X}\right]^T \right\} p(z) dz$$
$$\geq \int_{-\infty}^{+\infty} \{\mathrm{var}(x|z)\} p(z) dz = var(\hat{X}_{\mathrm{MMSE}})$$

For any a estimate $\hat{X}$, we have $var(\hat{X}_{\mathrm{MMSE}}) \leq var(\hat{X})$.

Conclusion:the MMSE estimate is the MV estimate.

according to definition

Minimum Mean Square Error estimate

$$E\left\{\left[X-\hat{X}\right]^T\left[X-\hat{X}\right]\right\}$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[x-\hat{X}\right]^T\left[x-\hat{X}\right]p(x,z)dxdz$$

$$=\int_{-\infty}^{+\infty}p(z)\left\{\int_{-\infty}^{+\infty}\left[x-\hat{X}(z)\right]^T\left[x-\hat{X}(z)\right]p(x|z)dx\right\}dz$$

$\because p(z)\geq 0 \therefore$ The MMSE estimate requires that for each $Z=z$,$g(z)$ should be minimized.

Minimum Mean Square Error estimate

$$\left.\frac{\partial g(\hat{X})}{\partial \hat{X}}\right|_{\hat{X}=\hat{X}_{\mathrm{MMSE}}} = 0$$

$$= -2\int_{-\infty}^{+\infty}\left[x - \hat{X}(Z)\right]p(x|z)dx\Big|_{\hat{X}=\hat{X}_{\mathrm{MMSE}}} = 0$$

$$\Rightarrow \hat{X}_{\mathrm{MMSE}} = \int_{-\infty}^{+\infty} xp(x|z)dx = E(X|Z)$$

The MMSE estimate is also called the conditional expectation.

Question: is the MMSE estimate is linear estimate?

Minimum Mean Square Error estimate

Performance Analysis

(1) Unbiased analysis

$$E\left[\hat{X}_{MS}\right] = E\left[E(X|Z)\right]$$

$$= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} xp(x|z)dx\right] p(z)dz$$

$$= \int_{-\infty}^{+\infty} x \left[\int_{-\infty}^{+\infty} p(x,z)dz\right] dx$$

$$= \int_{-\infty}^{+\infty} xp(x)dx = E(X)$$

Performance Analysis

(2) Covariance analysis

$$
\begin{aligned}
&\operatorname{var}(\hat{X}_{\mathrm{MMSE}}) = E(\tilde{X}_{\mathrm{MMSE}}\tilde{X}^T_{\mathrm{MMSE}}) \\
&= \iint \left[x - \hat{X}_{\mathrm{MMSE}}\right]\left[x - \hat{X}_{\mathrm{MMSE}}\right]^T p(x,z)\,dx\,dz \\
&= \int_{-\infty}^{+\infty} \underline{\underline{\left\{\int_{-\infty}^{+\infty} [x - E(x\mid z)][x - E(x\mid z)]^T p(x\mid z)\,dx\right\}}} p(z)\,dz \\
&= \int_{-\infty}^{+\infty} \underline{\underline{\operatorname{var}(x\mid z)}}\, p(z)\,dz
\end{aligned}
$$

Minimum Mean Square Error estimate

Consider the case that the measurement is not available,but the a priori statistics about X is known. How to choose a constant as the MMSE?

$$J\!=\!E\left\{\left(X-\hat{X}\right)^T\left(X-\hat{X}\right)\right\}=E\left\{X^TX\right\}-\hat{X}^TEX-(EX)^T\hat{X}+E\left\{\hat{X}^T\hat{X}\right\}$$

$$=E\left\{X^TX\right\}-2\hat{X}^TEX+E\left\{\hat{X}^T\hat{X}\right\}$$

$$\boxed{\hat{X}^TEX=\ E(X\hat{X})^T}$$

$$\frac{\partial J}{\partial \hat{X}}=-2EX-2\hat{X}\Rightarrow\hat{X}_{MS}=EX$$

And they are scalar and thus equal each other.

Conclusion: without the measurement, the mean of X is its MMSE estimate.

Minimum Mean Square Error estimate

Example3.5: Let unknown X be distributed uniformly according to

$$f_X(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & otherwise \end{cases}$$

Suppose a Z is measured, which is related to X by

$$Z = \ln\left(\frac{1}{X}\right) + V$$

where V is noise with exponential distribution

$$f_V(v) = \begin{cases} e^{-v}, & v \geq 0 \\ 0, & v < 0 \end{cases}$$

and X and V are independent. Find the MS estimate.

Minimum Mean Square Error estimate

Solution

$$f_{Z|X}(z|x) = f_V(z - \ln(\tfrac{1}{x}))$$

$$= \begin{cases} e^{-(z-\ln(1/x))} & z - \ln(\frac{1}{x}) \geq 0 \\ 0, & z - \ln(\frac{1}{x}) < 0 \end{cases} = \begin{cases} \frac{1}{x}e^{-z} & x \geq e^{-z} \\ 0, & x < e^{-z} \end{cases}$$

$$\hat{X}_{MS} = \frac{\int_{e^{-z}}^{1} e^{-z}\mathrm{d}x}{\int_{e^{-z}}^{1} \frac{1}{x}e^{-z}\mathrm{d}x} = \frac{xe^{-z}\big|_{e^{-z}}^{1}}{e^{-z}\ln x|_{e^{-z}}^{1}} = \frac{1-e^{-z}}{z}$$