ECSE 6520: Estimation and Detection Theory

Linear Estimation

Class Notes - 11

March 20th, 2013

Contents

1 Motivations

Linear estimators

- Linear estimators are simple in structure, therefore it is easy to implement them, particularly for real-time data processing.
- Linear estimators only require first and second order moment a priori information on the parameter to be estimated.
- MAP or MVUE lead to intractable mathematical models.
- MVUE is often not commputable.
- The randomness in the data can be modeled up to first and second order moments.

In such cases, we must content with linear and often times, suboptimal estimators. We will study the following linear estimators:

- The estimators is a constant: $\hat{\theta}(x) = c$.
- The estimator is a linear transformation: $\hat{\theta}(x) = Hx$.
- The estimator is an affine transformation: $\hat{\theta}(x) = Hx + c$.

2 Linear Minimum Mean Squared Error Estimation

- The optimal linear estimator depends only on the first and second order moments of X and θ.
- The optimal linear estimator always exist in the Bayesian setting.
- Bayesian linear estimation has important geometric interpretations in terms of orthogonal projections in Hilbert space.

Definition 1 ***Linear Minimum Mean Squared Error (LMMSE) Estimator -***
$\hat{\theta}(x) = Hx$ is the LMMSE estimator if H minimizes

$$\hat{\theta}(x) = E_{X,\theta}[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T].$$

Notation -

- Auto-covariance matrix of θ $(p \times p)$:

$$R_{\theta\theta} = E[(\theta - E[\theta])(\theta - E[\theta])^T]$$

- Cross-covariance matrix of θ and X $(p \times N)$:

$$R_{\theta X} = E[(\theta - E[\theta])(X - E[X])^T]$$

- Auto-covariance matrix of X $(N \times N)$:

$$R_{XX} = E[(X - E[X])(X - E[X])^T]$$

Theorem 1 *If $E[\theta] = 0$ and $E[X] = 0$ then the LMMSE is*

$$\hat{\theta}(x) = R_{\theta X} R_{XX}^{-1} x$$

provided R_{XX} is positive definite.

3 Affine Estimation and Non-zero Mean

- Linear estimators can be generalized by affine estimators:

$$\hat{\theta}(x) = Hx + b.$$

Definition 2 ***Affine MMSE -*** *We can define the affine MMSE estimator by minimizing*

$$MSE(H, b) = E_{X,\theta}[(\theta - HX - b)(\theta - HX - b)^T]$$

with respect to H and b.

Theorem 2 ***Affine MMSE*** **-** *The affine MMSE estimator is*

$$\hat{\theta}(x) = E[\theta] + R_{\theta X} R_{XX}^{-1} (x - E[X]).$$

Remarks -

- The affine MMSE estimator can be derived from the linear MMSE estimator by setting:

  $$\theta' = \theta - E[\theta]$$

  and

  $$X' = X - E[X].$$

  Then,

  $$\hat{H} = R_{\theta' X'} R_{X' X'}^{-1} = R_{\theta X} R_{XX}^{-1}$$

  and

  $$\hat{b} = E[\theta] - \hat{H} E[X].$$

  Thus, affine estimation, non-zero mean reduces to the case covered by the LMMSE.

- The algorithm for the non-zero mean is as follows:

  - Subtract off the mean of X.
  - Apply the LMMSE.
  - Add back the mean of θ.

Corrolary 1 ***Optimal Constant Estimator*** **-** *The optimal constant estimator is*

$$\hat{\theta}(x) = \hat{b} = \underbrace{E[\theta]}_{prior\ mean}.$$

Theorem 3 ***Bayesian Gauss-Markov Theorem*** **-** *Assume*

$$X = A\theta + \omega$$

where

A is full rank $N \times p$, known.

$E[\theta]$, $R_{\theta\theta}$ are known

$E[\omega] = 0$, $R_{\omega\omega}$ known.

$R_{\theta\omega} = 0_{p\times N}$, i.e, θ and ω uncorrelated.

Then the affine MMSE estimator is

$$\begin{aligned}\hat{\theta} &= E[\theta] + R_{\theta\theta}A^T(AR_{\theta\theta}A^T + R_{\omega\omega})^{-1}(x - AE[\theta]) \\ &= E[\theta] + (R_{\theta\theta}^{-1} + A^T R_{\omega\omega}^{-1} A)^{-1} A^T R_{\omega\omega}^{-1}(x - AE[\theta])\end{aligned}$$

4 The Orthogonality Principle

- The LMMSE estimator satisfies

$$R_{\theta X} = H^T R_{XX}$$

or

$$\begin{aligned}0 &= R_{\theta X} - H^T R_{XX} \\ &= E[\theta X^T - H^T X X^T] \\ &= E[(\theta - HX)X^T] \\ &= E[(\theta - \hat{\theta})X^T]\end{aligned}$$

- Thus,$(\theta - \hat{\theta})$ is orthogonal to X.
- As a consequence $(\theta - \hat{\theta})$ is orthogonal to the entire subspace of linear estimators. This is the manifestation of the orthogonality principle.
- Recall that in the deterministic least squares the difference between the orthogonal projection $\Pi_{\mathcal{A}}$ of x and x itself is also orthogonal to the signal subspace $\mathcal{A}$.

Definition 3 ***Hilbert Space of Finite Variance Random Variables -*** *Let $\mathcal{H}$ be the space of all random variables with zero mean and finite variance. Then, $\mathcal{H}$ forms a Hilbert space with the inner product*

$$< V_1,\ V_2 >= E[V_1 \bar{V_2}].$$

Definition 4 ***Wiener-Hopf Equation -*** *The equations*

$$R_{XX} H = R_{X\theta}$$

is called the Wiener-Hopf equation and the optimal H is called the Wiener estimator.

5 Best Linear Unbiased Estimation

LMMSE estimator is a Bayesian estimator, i.e., the unknown is regarded as random. BLUE is a linear estimator in classical approach where the unknown is non-random.

Definition 5 ***Best Linear Unbiased Estimator (BLUE)-*** *The best linear unbiased estimator is the linear estimator*

$$\hat{\theta}(x) = Hx, \qquad H \in \mathbb{R}^{p \times N}$$

with smallest variance among all linear, unbiased estimators.

- Note that for $\hat{\theta}(x)$ to be unbiased, we must have

  $$\theta = E[\hat{\theta}] = E[HX] = HE[X].$$

  Therefore the mean of the data must obey a linear relationship with the true parameter. This relationship will not always be true, so even the suboptimal BLUE may not be feasible. However there is an important class of problems where BLUE is feasible.

- Consider the following model:

  $$X = A\theta + \omega$$

  where

θ is fixed and unknown.

A is $N \times p$ known and full rank.

$E[\omega] = 0$

$E[\omega^T \omega] = R$ is known, positive definite.

Theorem 4 ***Gauss-Markov Theorem*** **-** *In the linear model described above, the best linear unbiased estimator is*

$$\hat{\theta} = (A^T R^{-1} A)^{-1} A^T R^{-1} x$$

and its covariance matrix is

$$R_{\hat{\theta}} = E[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] = (A^T R^{-1} A)^{-1}.$$

- Remark -

- When ω is Gaussian, the BLUE is MVUE for the linear model, i.e., the BLUE is optimal.