EE 574 Detection and Estimation Theory

Lecture Presentation 5

Aykut HOCANIN

Dept. of Electrical and Electronic Engineering
Eastern Mediterranean University

☞ Real (Non-random) Parameter Estimation

The unknown parameter may not always be a random variable. We would still like to design estimators which will estimate the parameter with respect to some criteria.

In order to evaluate the performance of non-random parameter estimation other measures of quality are needed. The first measure of quality is the expectation of the estimate;

$$E[\hat{a}(\mathbf{R})] = \int_{-\infty}^{\infty} \hat{a}(\mathbf{R}) p_{\mathbf{r}|a}(\mathbf{R}|A)\, dA\, d\mathbf{R} \tag{1}$$

The possible values of the expectation are classified into three classes:

1. If $E[\hat{a}(\mathbf{R})] = A$ for all values of A, the estimate is said to be unbiased.
2. If $E[\hat{a}(\mathbf{R})] = A + B$ where B is not a function of A, the estimate is said to have a known bias.
3. If $E[\hat{a}(\mathbf{R})] = A + B(A)$, the estimate is said to have an unknown bias.

Even an unbiased estimate may give a bad result on a particular trial. This is shown in Figure 1, where the probability density of the estimate is centered around A, but the variance of this density is large enough that big errors are possible.

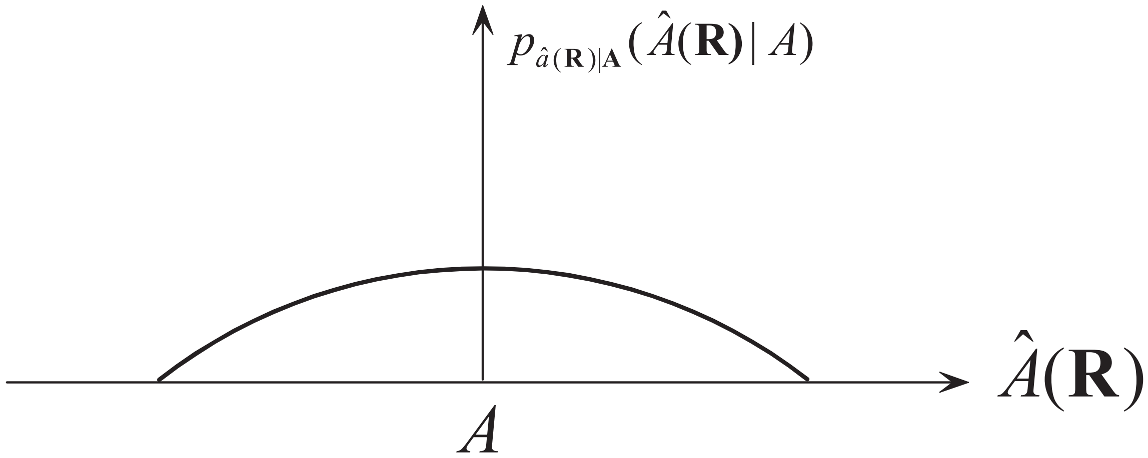

Figure 1: Probability density for an estimate.

A second measure of quality is the variance of the estimation error:

$$\mathrm{Var}[\hat{a}(\mathbf{R}) - A] = E\{[\hat{a}(\mathbf{R}) - A]^2\} - B^2(A) \tag{2}$$

This provides a measure of the spread of the error. In general, it is desired to have unbiased estimates with small variances.

☞ Maximum Likelihood Estimation

In general, we will choose the estimate to be the value of A that most likely caused a given value of R to occur.

The function $p_{\mathbf{r}|a}(\mathbf{R}|A)$ is a function of A and is denoted as the likelihood function.

The logarithm $\ln p_{\mathbf{r}|a}(\mathbf{R}|A)$ is denoted as the log likelihood function.

The maximum likelihood estimate $\hat{a}_{\mathrm{ml}}(\mathbf{R})$ is that value of A at which the likelihood function is a maximum. If the maximum is interior to the range of A, and $\ln p_{\mathbf{r}|a}(\mathbf{R}|A)$ has a continuous first derivative, then a necessary condition on $\hat{a}_{\mathrm{ml}}(\mathbf{R})$ is obtained by the following:

$$\left.\frac{\partial \ln p_{\mathbf{r}|a}(\mathbf{R}|A)}{\partial A}\right|_{A=\hat{a}_{\mathrm{ml}}(\mathbf{R})} = 0. \tag{3}$$

Equation (3) is called the likelihood equation. From discussion of the MAP estimate, we see that ML estimate corresponds to the limiting case of MAP estimate in which the a priori information approaches zero.

☞ Cramer-Rao Inequality: Non-random Parameters

If $\hat{a}_{\mathrm{ml}}(\mathbf{R})$ is any unbiased estimate of A, then

$$\mathrm{Var}[\hat{a}(\mathbf{R}) - A] \geq \left(E\left\{ \left[\frac{\partial \ln p_{\mathbf{r}|a}(\mathbf{R}|A)}{\partial A} \right]^2 \right\} \right)^{-1} \tag{4}$$

or

$$\mathrm{Var}[\hat{a}(\mathbf{R}) - A] \geq \left(-E\left\{ \left[\frac{\partial^2 \ln p_{\mathbf{r}|a}(\mathbf{R}|A)}{\partial A^2} \right] \right\} \right)^{-1} \tag{5}$$

where the first and second partial derivatives exist and are absolutely integrable.

Any estimate that satisfies the above bound with an equality is called an efficient estimate.

Some important observations regarding the Cramer-Rao Bound are:

1. Any unbiased estimate must have a variance greater than a certain number.
2. If an efficient estimate exists, it is the ML estimate $\hat{a}_{\mathrm{ml}}(\mathbf{R})$, and can be obtained as a unique solution of the likelihood equation.
3. If an efficient estimate does not exist, we do not know how good $\hat{a}_{\mathrm{ml}}(\mathbf{R})$ is.

Further, we do not know how close the variance of any estimate will approach the bound.

4. In order to use the bound, we must make sure that the estimate of concern is unbiased.

The properties of the ML estimate which are valid when the error is small are generally referred to as asymptotic.

A procedure for developing them formally is to study the behavior of the estimate as the number of independent observations N approaches infinity.

1. The solution of the likelihood equation Equation (3) converges in probability to the correct value of A as $N \rightarrow \infty$. Any estimate with this property is called consistent.
2. The ML estimate is asymptotically efficient:

$$\lim_{N \rightarrow \infty} \frac{\mathrm{Var}[\hat{a}_{\mathrm{ml}}(\mathbf{R}) - A]}{\left(-E\left\{\left[\frac{\partial^2 \ln p_{\mathbf{r}|a}(\mathbf{R}|A)}{\partial A^2}\right]\right\}\right)^{-1}} = 1$$

3. The ML estimate is asymptotically Gaussian, $G(A, \sigma_{a_\epsilon})$.

☞ Lower Bound on the Minimum Mean-Square Error in Estimating a Random Parameter

Let a be a random variable and $\mathbf{r}$, the observation vector. The mean-square error of any estimate $\hat{a}(\mathbf{R})$ satisfies the inequality:

$$E\{[\hat{a}(\mathbf{R}) - a]^2\} \geq \left(E\left\{ \left[\frac{\partial \ln p_{\mathbf{r},a}(\mathbf{R}, A)}{\partial A} \right]^2 \right\} \right)^{-1} \tag{6}$$

or

$$E\{[\hat{a}(\mathbf{R}) - a]^2\} \geq \left(-E\left\{ \left[\frac{\partial^2 \ln p_{\mathbf{r},a}(\mathbf{R}, A)}{\partial A^2} \right] \right\} \right)^{-1} \tag{7}$$

It should be noted that the probability density is a joint density and that the expectation is over both a and $\mathbf{r}$.

We assume that the first and second derivatives of the joint density exist and are absolutely integrable with respect to both $\mathbf{R}$ and A.

The conditional expectation of the error, given A is

$$B(A) = \int_{-\infty}^{\infty} [\hat{a}(\mathbf{R}) - A] p_{\mathbf{r}|a}(\mathbf{R}|A) \, d\mathbf{R} \tag{8}$$

We assume that

$$\lim_{A \to \infty} B(A) p_a(A) = 0 \tag{9}$$

$$\lim_{A \to -\infty} B(A) p_a(A) = 0 \tag{10}$$

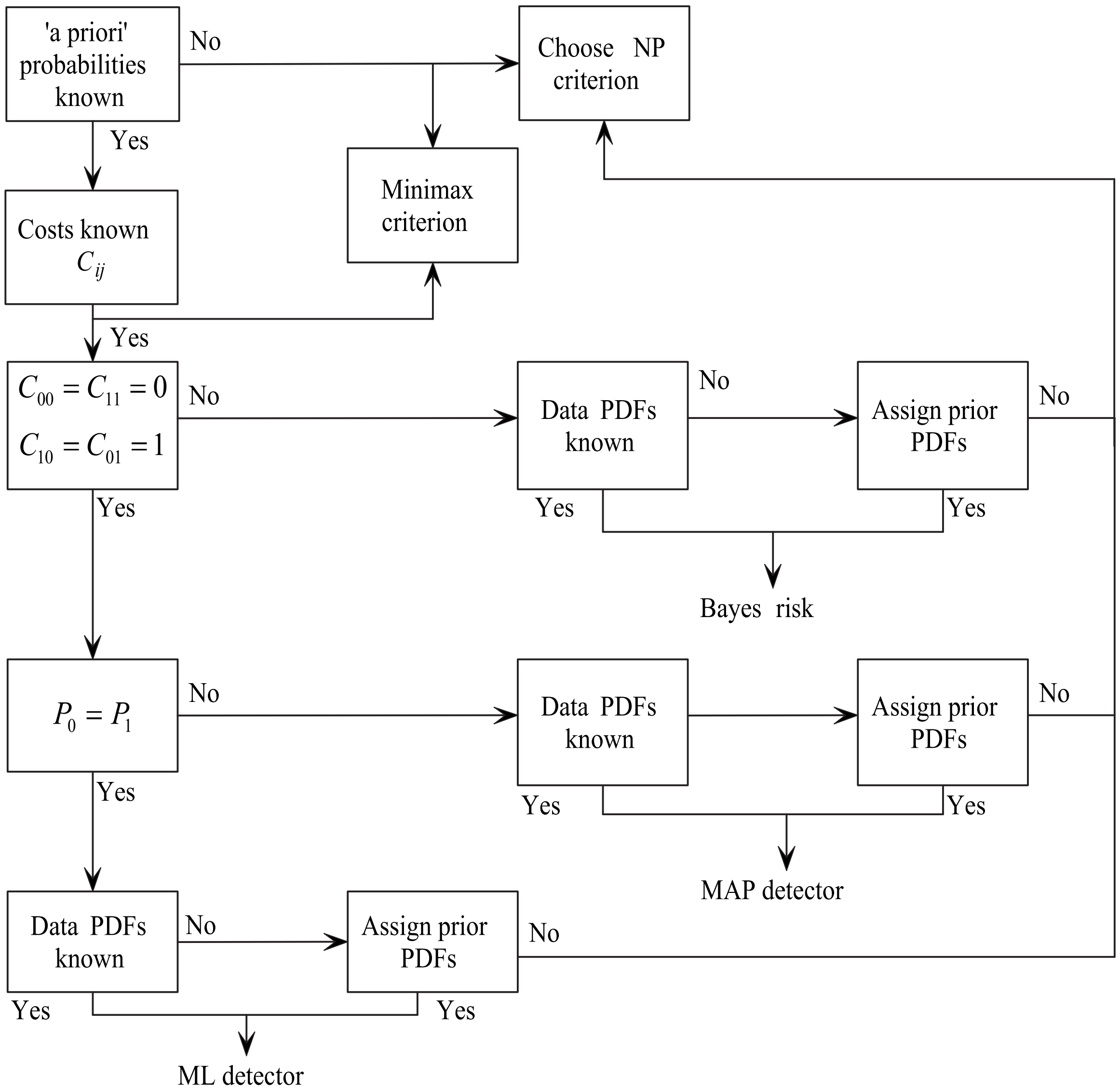

Figure 2: Summary of binary hypothesis testing

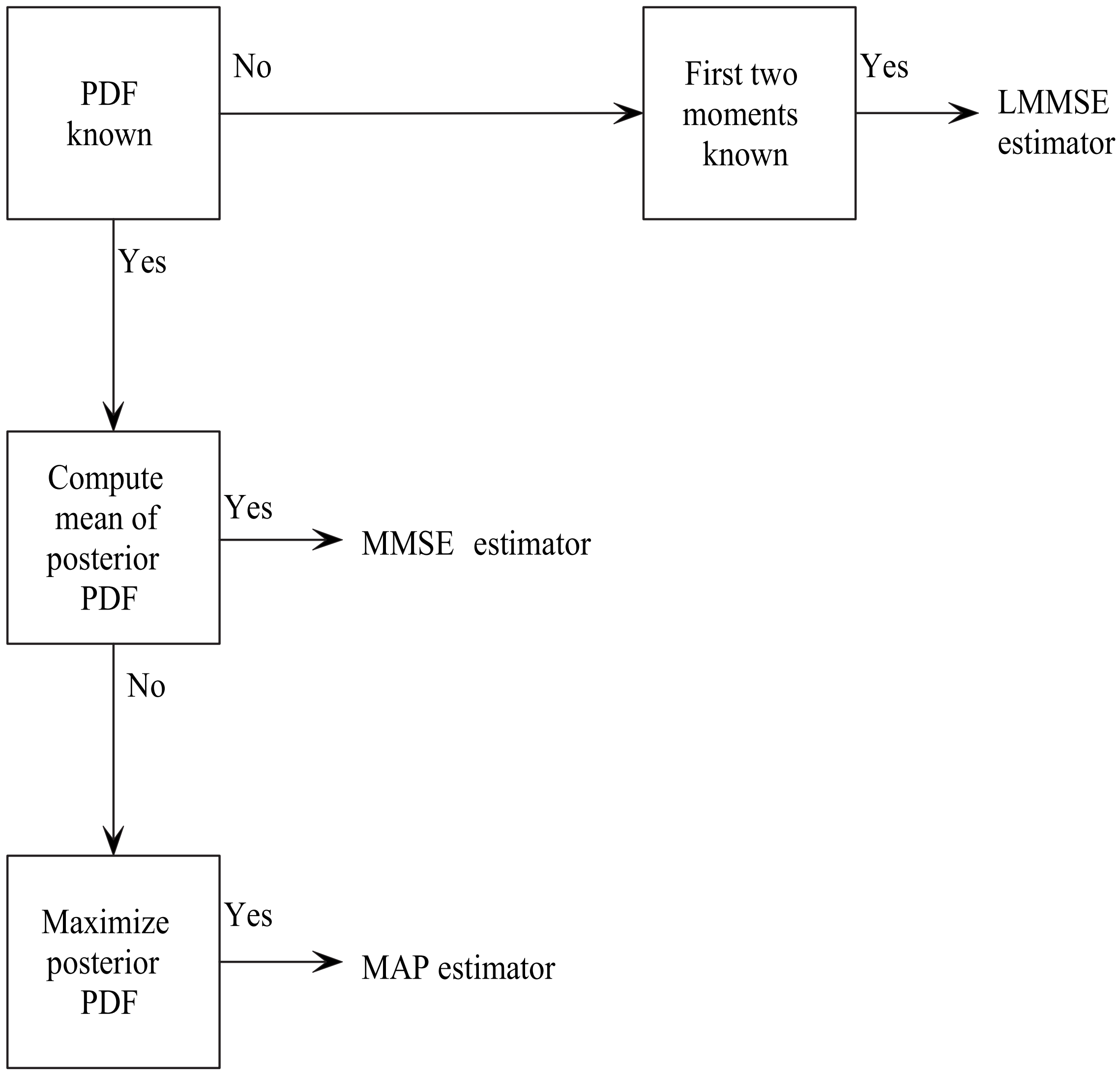

Figure 3: Bayesian approach to estimation

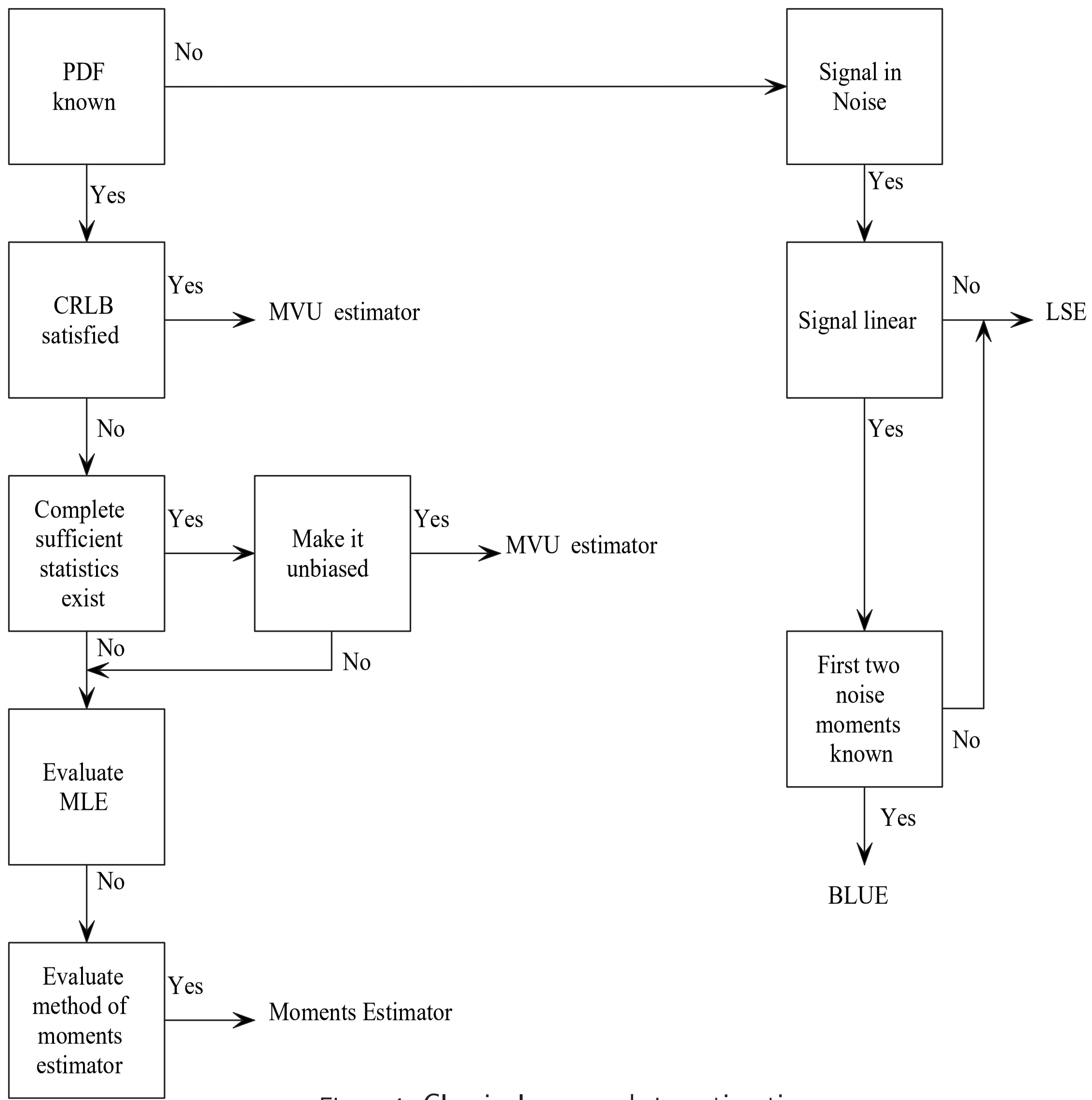

Figure 4: Classical approach to estimation