EE 631: Estimation and Detection
Part 7

Dr. Mehrdad Soumekh

Linear MMSE estimation

Consider an unknown message $m(t)$ which is estimated from the measured signal $r(t)$. Let the estimator be a linear system given by $h(t,u)$). The output of the estimator is given by:

$$\hat{m}(t) = \int_T h(t,u) r(u) du$$

Estimator error is given by:

$$e(t) = m(t) - \hat{m}(t)$$

Our aim is to design an optimum filter that minimizes

$$\xi_I = \int_T e^2(t) dt$$

and satisfies the orthogonality condition

$$E[e(t).r^*(u)] = 0 \qquad \forall \quad t, u \in [0, T]$$

<u>Estimator equation</u>
Substitute for the error in the orthogonality condition via:

$$\begin{aligned} e(t) &= m(t) - \hat{m}(t) \\ &= m(t) - \textstyle\int h(t,\lambda) r(\lambda) d\lambda \end{aligned}$$

This yields:

$$E\Bigg[[m(t) - \textstyle\int_T h(t,\lambda) r(\lambda) d\lambda] r^*(u)\Bigg] = 0$$

$$\Rightarrow E[m(t) r(u)] = \textstyle\int_T h(t,\lambda) E[r(\lambda) r^*(u)] d\lambda$$

$$\Rightarrow R_{mr}(t,u) = \textstyle\int_T h(t,\lambda) R_r(\lambda, u) d\lambda$$

The construction of the filter is based on solving for $h(t,u)$ from the above using the knowledge of $R_{mr}(.,.)$ and $R_r(.,.)$.
<u>Special case:</u>

$$r(t) = m(t) + n(t)$$

where $m(t)$ and $n(t)$ are orthogonal to each other. In this case:

$$\begin{aligned} R_{mr}(t,u) &= E\Bigg[m(t) r^*(u)\Bigg] \\ &= E\Bigg[m(t)[m^*(u) + n^*(u)]\Bigg] \\ &= R_m(t,u) \end{aligned}$$

Also

$$R_r(t,u) = R_m(t,u) + R_n(t,u)$$

Therefore, the estimator equation becomes:

$$R_m(t,u) = \textstyle\int_T h(t,\lambda) \big[R_m(\lambda,u) + R_n(\lambda,u)\big] d\lambda$$

Spectral solution via KL representation

Model:

$$r(t) = m(t) + n(t)$$

where $m(t)$ and $n(t)$ are general non-stationary signals that are orthogonal and uncorrelated to each other. Suppose the autocorrelation of both signals possess a common set of eigenfunctions that are denoted by:

$$\Phi = \{\phi_i(t); \quad ; i = 1, 2, \cdots\}$$

Thus,

$$\begin{aligned} R_m(t,u) &= \textstyle\sum_i \lambda_{mi}\phi_i(t)\phi_i^*(u) \\ R_n(t,u) &= \textstyle\sum_i \lambda_{ni}\phi_i(t)\phi_i^*(u) \end{aligned}$$

We consider the following decomposition for the optimal filter:

$$h(t,u) \quad = \textstyle\sum_i hi\phi_i(t)\phi_i^*(u)$$

Also, we have:

$$\begin{aligned} R_r(t,u) &= \textstyle\sum_i \lambda_{ri}\phi_i(t)\phi_i^*(u) \\ &= \textstyle\sum_i(\lambda_{mi} + \lambda_{ni})\phi_i(t)\phi_i^*(u) \end{aligned}$$

We now use the spectral representations in the estimator equation:

$$\begin{aligned} R_m(t,u) &= \textstyle\int_T h(t,\lambda)R_r(\lambda,u)d\lambda \\ \textstyle\sum \lambda_{mi}\phi_i(t)\phi_i^*(u) &= \textstyle\int_T \sum_i h_i\phi_i(t)\phi_i^*(\lambda)\sum_j(\lambda_{mj} + \lambda nj)\phi_j(\lambda)\phi_j^*(u)d\lambda \\ &= \textstyle\sum_i\sum_j h_i(\lambda_{mj} + \lambda_{nj})\phi_i(t)\phi_j^*(u)\underbrace{\left[\int_T \phi_i^*(\lambda)\phi_j(\lambda)d\lambda\right]}_{\delta_{ij}} \\ \therefore \textstyle\sum_i \lambda_{mi}\phi_i(t)\phi_i^*(u) &= \textstyle\sum_i h_i(\lambda_{mi} + \lambda_{ni})\phi_i(t)\phi_i^*(u) \end{aligned}$$

Due to the uniqueness of the spectral coefficients,

$$\lambda_{mi} = h_i(\lambda_{mi} + \lambda_{ni}) \qquad \forall \quad i = 1, 2, ...$$

$$\Rightarrow h_i = \frac{\lambda_{mi}}{\lambda_{mi} + \lambda_{ni}}$$

which is known as the general Weiner filter. The stationary equivalent of the above is given by:

$$H(\omega) = \frac{S_m(\omega)}{S_m(\omega) + S_n(\omega)}$$

Note that:

a) $\lambda_{mi} \gg \lambda_{ni} \Rightarrow h_i \approx 1$

b) $\lambda_{mi} \ll \lambda_{ni} \Rightarrow h_i \approx 0$

<u>The estimate:</u>
The estimate is constructed via the following:

$$\hat{m}(t) = \int_T h(t,u)r(u)du$$

We consider the spectral decomposition for the measurement:

$$\begin{aligned} m(t) &= \textstyle\sum_i m_i\phi_i(t) \\ n(t) &= \textstyle\sum_i n_i\phi_i(t) \\ \Rightarrow r(t) &= \textstyle\sum_i r_i\phi_i(t) \qquad ; r_i = m_i + n_i \end{aligned}$$

Using this and the spectral representation of for $h(t,u)$ in $\hat{m}(t)$, we get:

$$\begin{aligned} \hat{m}(t) &= \textstyle\int_T \sum_i h_i\phi_i(t)\phi_i^*(u) \sum_j r_j\phi_j(u)du \\ &= \textstyle\sum_i \sum_j h_i r_j \phi_i(t) \underbrace{\left[\int_T \phi_i^*(u)\phi_j(u)du\right]}_{\delta_{ij}} \\ &= \textstyle\sum_i h_i r_i \phi_i(t) \\ \rightarrow \hat{m}(t) &= \textstyle\sum_i \hat{m}_i\phi_i(t) \end{aligned}$$

where $\hat{m}_i = h_i r_i = \frac{\lambda_{mi} r_i}{\lambda_{mi}+\lambda_{ni}}$

<u>Estimator error energy:</u>

Point error:

$$\begin{aligned} \xi_p(t) &= E\Big[[m(t)-\hat{m}(t)][m^*(t)-\hat{m}^*(t)]\Big] \\ &= E\Big[e(t)[m^*(t) - \textstyle\int_T h(t,u)r^*(u)]du\Big] \\ &= E[e(t)m^*(t)] - \textstyle\int_T h(t,u) \underbrace{E[e(t)r^*(u)]}_{=0 \text{ due to orthogonality}} du \\ \Rightarrow \xi_p(t) &= E[e(t)m^*(t)] \\ &= E\Big[[m(t)-\hat{m}(t)]m^*(t)\Big] \\ &= R_m(t,t) - E\Big[\textstyle\int_T h(t,u)r(u)du m^*(t)\Big] \\ &= R_m(t,t) - \textstyle\int_T h(t,u) \underbrace{E[r(u)m^*(t)]}_{=R_m(u,t)} du \\ &= R_m(t,t) - \textstyle\int_T h(t,u)R_m(u,t)du \\ &= R_m(t,t) - \textstyle\int_T \sum_i h_i\phi_i(t)\phi_i^*(u)\sum_j \lambda_{mj}\phi_j(u)\phi_j^*(t)du \\ &= R_m(t,t) - \textstyle\sum_i\sum_j h_i\lambda_{mj}\phi_i(t)\phi_j^*(t) \underbrace{\left[\int_T \phi_i^*(u)\phi_j(u)du\right]}_{\delta_{ij}} \end{aligned}$$

$$\begin{aligned}\Rightarrow \xi_p(t) &= \sum_i \lambda_{mi}\phi_i(t)\phi_i^*(t) - \sum_i h_i\lambda_{mi}\phi_i(t)\phi_i^*(t) \\ &= \sum_i \left(\lambda_{mi} - \frac{\lambda_{mi}}{\lambda_{mi}+\lambda_{ni}}\lambda_{mi}\right)\phi_i(t)\phi_i^*(t) \\ &= \sum_i \left(\frac{\lambda_{mi}\lambda_{ni}}{\lambda_{mi}+\lambda_{ni}}\right)\left|\phi_i(t)\right|^2\end{aligned}$$

Interval error:

$$\begin{aligned}\xi_I &= \int_T \xi_p(t)dt \\ &= \int_T \sum_i \left(\frac{\lambda_{mi}\lambda_{ni}}{\lambda_{mi}+\lambda_{ni}}\right)\left|\phi_i(t)\right|^2 dt \\ &= \sum_i \left(\frac{\lambda_{mi}\lambda_{ni}}{\lambda_{mi}+\lambda_{ni}}\right)\underbrace{\int_T \left|\phi_i(t)\right|^2 dt}_{=1} \\ \Rightarrow \xi_I &= \sum_i \left(\frac{\lambda_{mi}\lambda_{ni}}{\lambda_{mi}+\lambda_{ni}}\right)\end{aligned}$$

Detection in additive white Gaussian noise

Under the mth hypothesis, the received signal is :

$$H_m: \qquad r(t) = s_m(t) + n(t)$$

where $n(t)$ is the white Gaussian noise and the transmitted signals are:

$$\vec{S} \triangleq \{s_m(t); \quad m = 0, 1, \cdots, M-1\}$$

The linear signal subspace can be represented by (e.g using the Gram Schmidt procedure) by $N \leq M$ orthonormal basis functions given by:

$$\vec{S} = \vec{\Phi} = \{\phi_i(t); \quad i = 1, \cdots, N \leq M\}$$

<u>Irrelevant data:</u>
The finite signal subspace $\vec{S} = \vec{\Phi}$ cannot be a CON set. Therefore, we identify the set:

$$\vec{S^c} \triangleq \{\phi_i(t); \quad i = N+1, \cdots, \infty\}$$

such that $[\vec{S}, \vec{S^c}]$ forms a CON set.
Once this signal set is identified, we can construct:

1. Projection of $S_m(t)$ into $\vec{S}$:

   $$\vec{S_m} = \begin{bmatrix} S_{m1} \\ S_{m2} \\ \vdots \\ S_{mN} \end{bmatrix}$$

   where $S_{mi} = < S_m, \phi_i > \quad \forall \;\; i = 1, 2, \cdots, N.$
   Clearly, $< S_m, \phi_i > = 0 \quad \forall \;\; i = N+1, \cdots, \infty$

2. Projection of $r(t)$ into $\vec{S}$:

   $$\vec{R_1} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix}$$

where $r_i =< R, \phi_i > \quad \forall \ \ i = 1, 2, \cdots, N$.

Under H_m : $\quad r_i = S_{mi} + n_i \quad ; i = 1, 2, ...N$,
where $n_i =< n, \phi_i >$.
Furthermore, $\forall \ \ i = N+1, \cdots, \infty$, we also have $r_i = n_i$, i.e. no signal component.

$$\vec{R_2} = \begin{bmatrix} r_{N+1} \\ r_{N+2} \\ \vdots \\ r_\infty \end{bmatrix}$$

Therefore, the total measurement vector is given by:

$$\vec{R} \quad = \left[\begin{array}{c} r_1 \\ r_2 \\ \vdots \\ r_N \\ \hline r_{N+1} \\ r_{N+2} \\ \vdots \\ r_\infty \end{array}\right] = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$$

Under H_m:

$$\vec{R} \quad = \left[\begin{array}{c} s_{m1} + n_1 \\ s_{m2} + n_2 \\ \vdots \\ s_{mN} + n_N \\ \hline n_{N+1} \\ n_{N+2} \\ \vdots \\ n_\infty \end{array}\right]$$

Since R_2 has only noise components, it is called *redundant data.*
Noise variance:

$$E\left[|n_i|^2\right] \quad = \frac{N_0}{2}$$

$$E[n_i n_j^*] \quad = 0 \qquad ; i \neq j$$

$\Rightarrow n_i$'s are i.i.d. $\sim N(0, \frac{N_0}{2})$. R_1 and R_2 are independent of each other. i.e.

$$p(\vec{R}|H_m) = p(\vec{R_1}|H_m) p(\vec{R_2}|H_m)$$

Since, R_2 does not depend on H_m,

$$p(\vec{R_2}|H_m) = p(R_2)$$

The likelihood ratio test:

$$\Lambda_m(\vec{R}) = \frac{p(\vec{R}|H_m)}{p(\vec{R}|H_0)} = \frac{p(\vec{R_1}|H_m)}{p(\vec{R_1}|H_0)}$$

i.e. $\vec{R_1}$ is sufficient statistic. R_2 on the other hand is irrelevant data and does not influence decision.
Statistic on R_1:
R_1 is multivariate normal. Under H_m, its mean is S_m and its covariance matrix is $\frac{N_0}{2} \times \vec{I}_{N \times N}$.