Section 6

EE278: Introduction to Statistical Signal Processing (Fall 2020)

Monday, Oct 26, 2020 - 3 to 4 pm

1. Singular Covariance Matrix. If $X, \mathbf{Y}$ are jointly Gaussian, the MMSE estimator of X given vector $\mathbf{Y}$ is given by:

$$\hat{X}(\mathbf{Y}) = K_{X\mathbf{Y}} K_{\mathbf{Y}}^{-1}(\mathbf{Y} - \bar{\mathbf{Y}}) + \bar{X}, \tag{1}$$

where $\bar{\mathbf{Y}}$ and $\bar{X}$ are the mean of $\mathbf{Y}$ and X.

However, this formula would not work if the covariance matrix $K_{\mathbf{Y}}$ is not invertible. In this problem, you are required to investigate this situation and obtain the corresponding solutions.

(a) Prove that if the covariance matrix of a d-dimensional zero-mean random vector $\mathbf{Y}$ is not invertible, then there must exist at least one component of $\mathbf{Y}$ which is expressible as a linear combination of the others.

(b) Suppose $X \sim \mathcal{N}(0, \sigma_X^2)$ (a scalar random variable) and $\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, K_{\mathbf{Y}})$ (a random vector) are jointly Gaussian, where

$$K_{\mathbf{Y}} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 5 \end{bmatrix}, \tag{2}$$

and the covariance between X and $\mathbf{Y}$ is

$$K_{X\mathbf{Y}} = [3, 4, 7]. \tag{3}$$

Derive the MMSE estimator of X based on $\mathbf{Y}$.

Solution:

(a) If $K_{\mathbf{Y}}$ is not invertible, then there exists a non-zero vector $\mathbf{a}$ such that $\mathbf{a}^T K_{\mathbf{Y}} \mathbf{a} = 0$. This means that $\mathrm{Var}(\mathbf{a}^T\mathbf{Y}) = 0$, hence $\mathbf{a}^T\mathbf{Y} = 0$ with probability 1. At least one component of $\mathbf{a}$ is non-zero, so suppose that $a_k \neq 0$. Then we can write

$$Y_k = \sum_{i \neq k} \frac{-a_i}{a_k} Y_i$$

Thus, we have shown that Y_k could be expressed as linear combination of other components in $\mathbf{Y}$.

Conversely, if there exists a component Y_i such that it can be expressed as a linear combination of other components, then there must exist a non-zero d-dimensional vector $\mathbf{a}$ such that

$$\mathbf{a}^T\mathbf{Y} \equiv 0. \tag{4}$$

We have

$$\mathbb{E}[(\mathbf{a}^T\mathbf{Y})^2] = \mathbb{E}[(\mathbf{Y}^T a)^2] = \mathbf{a}^T \mathbb{E}[\mathbf{Y}\mathbf{Y}^T]\mathbf{a} = \mathbf{a}^T K_{\mathbf{Y}} \mathbf{a} = 0. \tag{5}$$

The existence of a non-zero vector $\mathbf{a}$ such that $\mathbf{a}^T K_{\mathbf{Y}} \mathbf{a}$ implies that $K_{\mathbf{Y}}$ is not invertible.

(b) This covariance matrix is not invertible. Observe that with $\mathbf{a} = [1, 1, -1]^T$, $\mathbf{a}^T K_{\mathbf{Y}} \mathbf{a} = 0$. Hence, we have

$$Y_3 = Y_1 + Y_2. \tag{6}$$

Thus, we can abandon Y_3 and use $\mathbf{Y}' = (Y_1, Y_2)^T$ to construct the optimal linear estimator of X given $\mathbf{Y}$.

We have

$$\hat{X}(\mathbf{Y}) = \hat{X}(\mathbf{Y}') = K_{X\mathbf{Y}'}K_{\mathbf{Y}'}^{-1}\mathbf{Y}', \tag{7}$$

since we have assumed $\mathbb{E}X = 0, \mathbb{E}\mathbf{Y} = 0$.

We have

$$K_{X\mathbf{Y}'} = [3, 4], \tag{8}$$

and

$$K_{\mathbf{Y}'}^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}. \tag{9}$$

Hence we have

$$\hat{X}(\mathbf{Y}) = [2, 1]\mathbf{Y}' = 2Y_1 + Y_2. \tag{10}$$

2. Noise cancellation A classic problem in statistical signal processing involves estimating a weak signal (e.g., the heart beat of a fetus) in the presence of a strong interference (the heart beat of its mother) by making two observations—one with the weak signal present and one without (by placing one microphone on the mother's belly and another close to her heart). The observations can then be combined to estimate the weak signal by "canceling out" the interference. The following is a simple version of this application.

Let the weak signal $X \sim \mathcal{N}(\mu, P)$ Let the observations be $Y_1 = X + Z_1$ and $Y_2 = Z_1 + Z_2$, where Z_1 is the strong interference and Z_2 is measurement noise. Assume that $Z_1 \sim \mathcal{N}(0, N_1)$ and $Z_2 \sim \mathcal{N}(0, N_2)$ Further assume that X, Z_1, and Z_2 are independent. Find the MMSE estimate of X given Y_1 and Y_2, and the corresponding MSE. Interpret the results.

Solution 1:

This is a vector MSE linear estimation problem. Since Z_1 and Z_2 are zero mean, $\mu_{Y_1} = \mu_X + \mu_{Z_1} = \mu$ and $\mu_{Y_2} = \mu_{Z_1} + \mu_{Z_2} = 0$. We first normalize the random variables by subtracting their means to get

$$X' = X - \mu \quad \text{and} \quad \mathbf{Y}' = \begin{bmatrix} Y_1 - \mu \\ Y_2 \end{bmatrix}.$$

Formulating this in the form of Example 3 in class,

$$\begin{aligned} \mathbf{Y}' &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} X' + \begin{bmatrix} 1 \\ 1 \end{bmatrix} Z_1 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} Z_2 \\ &= \mathbf{h}X' + \mathbf{U} \end{aligned} \tag{11}$$

Let $K_{\mathbf{U}}$ covariance matrix of $\mathbf{U}$.

$$K_{\mathbf{U}} = N_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} + N_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} N_1 & N_1 \\ N_1 & N_1 + N_2 \end{bmatrix}$$

The first step we do is to whiten the noise. Multiplying (11) by $K_{\mathbf{U}}^{-\frac{1}{2}}$,

$$\begin{aligned} \widetilde{\mathbf{Y}} = K_{\mathbf{U}}^{-\frac{1}{2}}\mathbf{Y}' &= K_{\mathbf{U}}^{-\frac{1}{2}}\mathbf{h}X' + K_{\mathbf{U}}^{-\frac{1}{2}}\mathbf{U} \\ &= \widetilde{\mathbf{h}}X' + \widetilde{\mathbf{U}} \end{aligned}$$

Here, the covariance matrix of $\widetilde{\mathbf{U}}$ is the identity matrix. The next step is to project on to the direction $\widetilde{\mathbf{h}}$, which gives a sufficient statistic.

$$\mathbf{Y}'' = \widetilde{\mathbf{h}}^T\widetilde{\mathbf{Y}} = \widetilde{\mathbf{h}}^T\widetilde{\mathbf{h}}X' + \widetilde{\mathbf{h}}^T\widetilde{\mathbf{U}}$$

Using $\widetilde{\mathbf{h}} = K_{\mathbf{U}}^{-\frac{1}{2}}\mathbf{h}$ and $\widetilde{\mathbf{U}} = K_{\mathbf{U}}^{-\frac{1}{2}}\mathbf{U}$, we have

$$\mathbf{Y}'' = \mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{Y}' = \mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h}X' + \mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{U} \tag{12}$$
$$= X'' + Z'' \tag{13}$$

Thus we have converted our problem to the scalar estimation problem (Example 1 in class). The variance of X'' is $P(\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h})^2$, and the variance of Z'' is $\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h}$. Thus the MMSE estimator for X'' from Y'' is

$$\widehat{X''} = \frac{\mathrm{Var}(X'')}{\mathrm{Var}(X'') + \mathrm{Var}(Z'')}Y'' = \frac{P(\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h})}{1 + P(\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h})}Y''$$

Working backwards, we form the estimator

$$\widehat{X'} = \frac{P}{1 + P(\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h})}\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{Y}'$$

Using the values of $\mathbf{h}$ and $K_{\mathbf{U}}$, we get

$$K_{\mathbf{U}}^{-1} = \frac{1}{N_1N_2}\begin{bmatrix} N_1 + N_2 & -N_1 \\ -N_1 & N_1 \end{bmatrix}, \qquad \mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h} = \frac{N_1 + N_2}{N_1N_2}, \qquad \mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{Y}' = \frac{N_1 + N_2}{N_1N_2}Y_1 - \frac{1}{N_2}Y_2$$

Thus the final estimator is

$$\begin{aligned}\widehat{X'} &= \frac{P(N_1 + N_2)Y_1' - PN_1Y_2'}{P(N_1 + N_2) + N_1N_2} \\ \Longrightarrow\ \widehat{X} &= \frac{P(N_1 + N_2)(Y_1 - \mu) - PN_1Y_2}{P(N_1 + N_2) + N_1N_2} + \mu \\ &= \frac{P(N_1 + N_2)Y_1 - PN_1Y_2 + N_1N_2\mu}{P(N_1 + N_2) + N_1N_2}\end{aligned}$$

From the scalar problem (12),

$$\begin{aligned}\mathrm{Var}(X'' \mid Y'') &= \mathrm{Var}(X'') - \frac{\mathrm{Var}(X'')^2}{\mathrm{Var}(X'') + \mathrm{Var}(Z'')} \\ &= \frac{\mathrm{Var}(X'')\mathrm{Var}(Z'')}{\mathrm{Var}(X'') + \mathrm{Var}(Z'')} \\ &= \frac{P(\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h})^2(\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h})}{P(\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h})^2 + (\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h})}\end{aligned}$$

The final MSE is

$$\begin{aligned}\mathrm{MSE} = \mathrm{Var}(X \mid \mathbf{Y}) &= \frac{\mathrm{Var}(X'' \mid Y'')}{(\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h})^2} \\ &= \frac{P}{P(\mathbf{h}^T K_{\mathbf{U}}^{-1}\mathbf{h}) + 1} \\ &= \frac{PN_1N_2}{P(N_1 + N_2) + N_1N_2}\end{aligned}$$

Solution 2:

We can also solve this using the formula in (1). We first find the covariance matrices

$$\Sigma_{\mathbf{Y}} = \begin{bmatrix} P+N_1 & N_1 \\ N_1 & N_1+N_2 \end{bmatrix} \quad \text{and} \quad \Sigma_{\mathbf{Y}X} = \begin{bmatrix} P \\ 0 \end{bmatrix}.$$

Therefore

$$\begin{aligned}
\hat{X}' &= \Sigma_{\mathbf{Y}X}^T \Sigma_{\mathbf{Y}}^{-1} \mathbf{Y}' \\
&= \begin{bmatrix} P & 0 \end{bmatrix} \frac{1}{P(N_1{+}N_2) + N_1 N_2} \begin{bmatrix} N_1{+}N_2 & -N_1 \\ -N_1 & P{+}N_1 \end{bmatrix} \mathbf{Y}' \\
&= \frac{P}{P(N_1{+}N_2) + N_1 N_2} \begin{bmatrix} N_1{+}N_2 & -N_1 \end{bmatrix} \begin{bmatrix} Y_1 - \mu \\ Y_2 \end{bmatrix} \\
&= \frac{P(N_1 + N_2)(Y_1 - \mu) - P N_1 Y_2}{P(N_1{+}N_2) + N_1 N_2} .
\end{aligned}$$

Replacing the means of X and $\mathbf{Y}$,

$$\begin{aligned}
\hat{X} &= \frac{P(N_1 + N_2)(Y_1 - \mu) - P N_1 Y_2}{P(N_1{+}N_2) + N_1 N_2} + \mu \\
&= \frac{P(N_1 + N_2) Y_1 - P N_1 Y_2 + N_1 N_2 \mu}{P(N_1 + N_2) + N_1 N_2} .
\end{aligned}$$

The MSE can be calculated by

$$\begin{aligned}
\text{MSE} &= \sigma_X^2 - \Sigma_{\mathbf{Y}X}^T \Sigma_{\mathbf{Y}}^{-1} \Sigma_{\mathbf{Y}X} \\
&= P - \frac{P}{P(N_1 + N_2) + N_1 N_2} \begin{bmatrix} N_1{+}N_2 & -N_1 \end{bmatrix} \begin{bmatrix} P \\ 0 \end{bmatrix} \\
&= P - \frac{P^2 (N_1 + N_2)}{P(N_1 + N_2) + N_1 N_2} \\
&= \frac{P N_1 N_2}{P(N_1 + N_2) + N_1 N_2} = \frac{P \frac{N_1 N_2}{N_1 + N_2}}{P + \frac{N_1 N_2}{N_1 + N_2}} .
\end{aligned}$$

Interpretation:

If either N_1 or N_2 go to 0, the MSE also goes to 0. This is because the estimator will then use the measurement with zero noise variance (that is, the one with no noise) to perfectly reconstruct X.

If we use only the observation Y_1, then our estimator would be $\widehat{X} = \frac{P}{P+N_1} Y_1$, whose MSE would be $\frac{PN_1}{P+N_1}$. Another naive approach is to use only $Y_1 - Y_2 = X - Z_2$. In this case, the estimator would be $\widehat{X} = \frac{P}{P+N_2}(Y_1 - Y_2)$, whose MSE would be $\frac{PN_2}{P+N_2}$. The MMSE estimator's MSE is smaller than both of these!