5. Linear Estimation

5.1 LMMSE Estimator, MSE and SNR

Consider a MIMO point-to-point channel where the transmitter is equipped with N_T transmit antennas and the receiver is equipped with N_R receive antennas. Using the linear MIMO model and assuming additive white Gaussian noise $\boldsymbol{v} \in \mathbb{R}^{N_\mathrm{R}}$ with $\boldsymbol{v} \sim \mathcal{N}(\mathbf{0}, \sigma_{\boldsymbol{v}}^2 \boldsymbol{I})$, the receive signal $\boldsymbol{y} \in \mathbb{R}^{N_\mathrm{R}}$ at the receiver can be expressed as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{v},$$

where $\boldsymbol{x} \in \mathbb{R}^{N_\mathrm{T}}$ with $\boldsymbol{x} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{x}})$ denotes the transmit signal and $\boldsymbol{H} \in \mathbb{R}^{N_\mathrm{R} \times N_\mathrm{T}}$ denotes the channel matrix. In the following, assume that $\boldsymbol{H}$ is known to the transmitter and the receiver.

a) Derive the linear minimum mean square error (LMMSE) estimator for an estimation of the transmit signal $\boldsymbol{x}$ based on the receive signal $\boldsymbol{y}$. Why is the LMMSE estimator the minimum mean square error (MMSE) estimator as well?

Hint: Following derivatives may be helpful:

$$\frac{\partial \mathrm{tr}\left\{\boldsymbol{X}\boldsymbol{A}\right\}}{\partial \boldsymbol{X}} = \boldsymbol{A}^\mathrm{T},$$
$$\frac{\partial \mathrm{tr}\left\{\boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^\mathrm{T}\right\}}{\partial \boldsymbol{X}} = \boldsymbol{X}\boldsymbol{A}^\mathrm{T} + \boldsymbol{X}\boldsymbol{A}.$$

In the following, we consider a SIMO channel, i.e., we have $N_\mathrm{T} = 1$.

b) Determine the filter which maximizes the receive SNR. Which SNR does this filter achieve?

c) Show that in this case the LMMSE estimator actually is a matched filter, i.e., a filter that maximizes the signal to noise power ratio as well.

Hint: Given the matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ and the vectors $\boldsymbol{b} \in \mathbb{R}^{n \times 1}$ and $\boldsymbol{c} \in \mathbb{R}^{n \times 1}$, it follows that

$$\left(\boldsymbol{A} + \boldsymbol{b}\boldsymbol{c}^\mathrm{T}\right)^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{b}\boldsymbol{c}^\mathrm{T}\boldsymbol{A}^{-1}}{1 + \boldsymbol{c}^\mathrm{T}\boldsymbol{A}^{-1}\boldsymbol{b}}.$$

d) Express the MSE achieved by LMMSE estimator as a function of the optimal SNR.

5.2 ”Linear” Models and LMMSE

The random vectors $\boldsymbol{x} \in \mathbb{R}^m$ and $\boldsymbol{y} \in \mathbb{R}^n$ model the inputs and outputs of an unknown noisy operator as depicted in Figure 5.1. The input vector and the output vector are stacked into the random vector

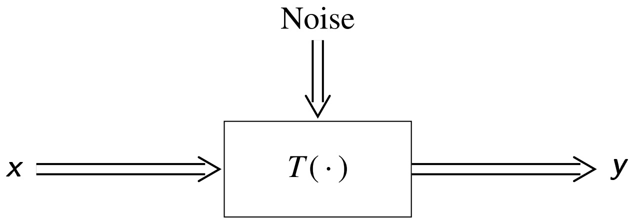

Fig. 5.1: Unknown and noisy operator.

$\boldsymbol{z} = \begin{bmatrix} \boldsymbol{x}^{\mathrm{T}} & \boldsymbol{y}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$. The distribution of $\boldsymbol{z}$ as well as the marginal distributions of $\boldsymbol{x}$ and $\boldsymbol{y}$ are unknown. In contrast, the first and second order moment of $\boldsymbol{z}$, i.e.,

$$\boldsymbol{\mu}_{\boldsymbol{z}} = \mathrm{E}\left[\boldsymbol{z}\right] = \begin{bmatrix} \boldsymbol{\mu}_{\boldsymbol{x}} \\ \boldsymbol{\mu}_{\boldsymbol{y}} \end{bmatrix} = \mathbf{0} \quad \text{and} \quad \boldsymbol{C}_{\boldsymbol{z}} = \mathrm{E}\left[\boldsymbol{z}\boldsymbol{z}^{\mathrm{T}}\right] = \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{x}} & \boldsymbol{C}_{\boldsymbol{yx}}^{\mathrm{T}} \\ \boldsymbol{C}_{\boldsymbol{yx}} & \boldsymbol{C}_{\boldsymbol{y}} \end{bmatrix},$$

are available. In the following, assume that $\boldsymbol{C}_{\boldsymbol{x}} \succ \mathbf{0}$ and $\boldsymbol{C}_{\boldsymbol{y}} \succ \mathbf{0}$.

a) Determine $\boldsymbol{T}$, $\boldsymbol{\mu}_{\boldsymbol{v}}$, and $\boldsymbol{C}_{\boldsymbol{v}} \succeq \mathbf{0}$ in order to formulate the ”linear” model

$$\boldsymbol{y}' = \boldsymbol{T}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{v},$$

where $\boldsymbol{v}$ is independent of $\boldsymbol{x}$, such that

$$\mathrm{E}\left[\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y}' \end{bmatrix}\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y}' \end{bmatrix}^{\mathrm{T}}\right] = \boldsymbol{C}_{\boldsymbol{z}} \quad \text{and} \quad \mathrm{E}\left[\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y}' \end{bmatrix}\right] = \boldsymbol{\mu}_{\boldsymbol{z}}.$$

Determine $\boldsymbol{S}$, $\boldsymbol{\mu}_{\boldsymbol{n}}$, and $\boldsymbol{C}_{\boldsymbol{n}} \succeq \mathbf{0}$ for the equivalent ”linear” model

$$\boldsymbol{x}' = \boldsymbol{S}^{\mathrm{T}}\boldsymbol{y} + \boldsymbol{n}$$

as well.

b) Compare the linear models to the LMMSE estimators $\boldsymbol{T}_{\mathrm{LMMSE}}^{\mathrm{T}}$ with $\hat{\boldsymbol{y}}_{\mathrm{LMMSE}} = \boldsymbol{T}_{\mathrm{LMMSE}}^{\mathrm{T}}\boldsymbol{x}$ and $\boldsymbol{S}_{\mathrm{LMMSE}}^{\mathrm{T}}$ with $\hat{\boldsymbol{x}}_{\mathrm{LMMSE}} = \boldsymbol{S}_{\mathrm{LMMSE}}^{\mathrm{T}}\boldsymbol{y}$.

c) Determine the error covariance matrix and the covariance matrix of the estimates for both LMMSE estimators. What do you observe?

d) Let $\boldsymbol{z} \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{z}}, \boldsymbol{C}_{\boldsymbol{z}})$. Show that in this case $\boldsymbol{T}_{\mathrm{LMMSE}}$ and $\boldsymbol{S}_{\mathrm{LMMSE}}$ are minimum mean square error (MMSE) estimators as well.

5.3 Least Squares MIMO Channel Estimation

Consider a transmitter which is equipped with N_T transmit antennas and a receiver which is equipped with N_R receive antennas. In order to estimate the channel between transmitter and receiver, the transmitter transmits N globally known training symbols $\{\boldsymbol{x}_i\}_{i=1}^N \in \mathbb{R}^{N_\mathrm{T}}$. The corresponding receive signals at the receiver are given by N vectors $\{\boldsymbol{y}_i\}_{i=1}^N \in \mathbb{R}^{N_\mathrm{R}}$. The channel between transmitter and receiver is constant during the transmission of the N training symbols.

a) Formulate a linear model $\hat{\boldsymbol{Y}}_\mathrm{LS} = \boldsymbol{X}\boldsymbol{T}_\mathrm{LS} \in \mathbb{R}^{N\times N_\mathrm{R}}$ in order to determine the linear least squares (LS) estimator $\boldsymbol{T}_\mathrm{LS}^\mathrm{T} : \mathbb{R}^{N_\mathrm{T}} \to \mathbb{R}^{N_\mathrm{R}},\ \boldsymbol{x} \mapsto \hat{\boldsymbol{y}}_\mathrm{LS} = \boldsymbol{T}_\mathrm{LS}^\mathrm{T}\boldsymbol{x}$.

b) State the orthogonality condition for the estimator $\boldsymbol{T}_\mathrm{LS}$. What is the span of the subspace of the errors $\boldsymbol{Y} - \hat{\boldsymbol{Y}}_\mathrm{LS}$ and what is the span of the subspace of the estimates $\hat{\boldsymbol{Y}}_\mathrm{LS}$?

c) Using the result from sub-problem b), determine the linear least squares estimator $\boldsymbol{T}_\mathrm{LS}$ and the least squares estimate $\hat{\boldsymbol{Y}}_\mathrm{LS}$.

For those who can't get enough of optimization problems:

d) Determine $\boldsymbol{T}_\mathrm{LS}$ without exploiting the orthogonality principle.

Hint 1: For $\boldsymbol{A} \in \mathbb{R}^{n\times m}$, the square of the Frobenius norm is given as

$$\|\boldsymbol{A}\|_\mathrm{F}^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} |a_{ij}|^2 = \mathrm{tr}\left\{\boldsymbol{A}\boldsymbol{A}^\mathrm{T}\right\}.$$

Hint 2: Following derivatives may be helpful:

$$\frac{\partial \mathrm{tr}\left\{\boldsymbol{X}\boldsymbol{A}\right\}}{\partial \boldsymbol{X}} = \boldsymbol{A}^\mathrm{T},$$

$$\frac{\partial \mathrm{tr}\left\{\boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^\mathrm{T}\right\}}{\partial \boldsymbol{X}} = \boldsymbol{X}\boldsymbol{A}^\mathrm{T} + \boldsymbol{X}\boldsymbol{A}.$$

5.4 Least Squares Estimation–A Different Perspective

Consider two random vectors $\boldsymbol{x} \in \mathbb{R}^M$ and $\boldsymbol{y} \in \mathbb{R}^d$. In this problem, no knowledge about the marginal distributions and joint distribution of $\boldsymbol{x}$ and $\boldsymbol{y}$ is available. However, we have a set of samples/realizations $\{(\boldsymbol{x}_1, \boldsymbol{y}_1), \ldots, (\boldsymbol{x}_N, \boldsymbol{y}_N)\}$ with $N > d$. We define the matrices

$$\boldsymbol{Y} = \begin{bmatrix} \text{——} \boldsymbol{y}_1^{\mathrm{T}} \text{——} \\ \vdots \\ \text{——} \boldsymbol{y}_N^{\mathrm{T}} \text{——} \end{bmatrix}, \quad \hat{\boldsymbol{Y}}_{\mathrm{LS}} = \begin{bmatrix} \text{——} \hat{\boldsymbol{y}}_1^{\mathrm{T}} \text{——} \\ \vdots \\ \text{——} \hat{\boldsymbol{y}}_N^{\mathrm{T}} \text{——} \end{bmatrix}, \quad \text{and} \quad \boldsymbol{X} = \begin{bmatrix} \text{——} \boldsymbol{x}_1^{\mathrm{T}} \text{——} \\ \vdots \\ \text{——} \boldsymbol{x}_N^{\mathrm{T}} \text{——} \end{bmatrix}. \tag{5.1}$$

Using a linear model $\hat{\boldsymbol{Y}}_{\mathrm{LS}} = \boldsymbol{X}\boldsymbol{T}_{\mathrm{LS}} \in \mathbb{R}^{N\times d}$ we can formulate the least squares (LS) estimator $\boldsymbol{T}_{\mathrm{LS}}^{\mathrm{T}} : \mathbb{R}^M \to \mathbb{R}^d,\ \boldsymbol{x} \mapsto \hat{\boldsymbol{y}}_{\mathrm{LS}} = \boldsymbol{T}_{\mathrm{LS}}^{\mathrm{T}}\boldsymbol{x}$. It can be shown that $\boldsymbol{T}_{\mathrm{LS}}^{\mathrm{T}}$ is given as $\boldsymbol{T}_{\mathrm{LS}}^{\mathrm{T}} = \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}$.

a) Express the LS estimator and the LS estimate in terms of the sample correlation matrices

$$\hat{\boldsymbol{R}}_{\boldsymbol{xx}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}} \quad \text{and} \quad \hat{\boldsymbol{R}}_{\boldsymbol{yx}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{y}_i\boldsymbol{x}_i^{\mathrm{T}}. \tag{5.2}$$

b) Assume that $\mathrm{E}[\boldsymbol{x}] = \boldsymbol{0}$ and $\mathrm{E}[\boldsymbol{y}] = \boldsymbol{0}$. What does this imply for your result of sub-problem a)? Give an interpretation of your result!

c) Assume that $\mathrm{E}[\boldsymbol{x}] = \boldsymbol{0}$ and $\mathrm{E}[\boldsymbol{y}] = \boldsymbol{0}$. What happens to the LS estimator and the LS estimate if $N \to \infty$?

Advanced: From now on, we consider the affine case

$$\hat{\boldsymbol{y}}_{\mathrm{LS}} = \boldsymbol{T}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{m} = \begin{bmatrix} \boldsymbol{T}^{\mathrm{T}} & \boldsymbol{m} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} = \boldsymbol{T}_{\mathrm{LS}}^{\prime,\mathrm{T}}\boldsymbol{x}', \tag{5.3}$$

where $\boldsymbol{T}_{\mathrm{LS}}^{\prime,\mathrm{T}} = \begin{bmatrix} \boldsymbol{T}^{\mathrm{T}} & \boldsymbol{m} \end{bmatrix}$ is given as $\boldsymbol{T}_{\mathrm{LS}}^{\prime,\mathrm{T}} = \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}'(\boldsymbol{X}'^{,\mathrm{T}}\boldsymbol{X}')^{-1}$ using $\boldsymbol{X}' = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{1} \end{bmatrix}$ with $\boldsymbol{X}$ defined above.

d) Express the LS estimator and the LS estimate in terms of the sample covariance matrices

$$\hat{\boldsymbol{C}}_{\boldsymbol{xx}} = \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_{\boldsymbol{x}})(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_{\boldsymbol{x}})^{\mathrm{T}} \quad \text{and} \quad \hat{\boldsymbol{C}}_{\boldsymbol{yx}} = \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{y}_i - \hat{\boldsymbol{\mu}}_{\boldsymbol{y}})(\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_{\boldsymbol{x}})^{\mathrm{T}}. \tag{5.4}$$

where $\hat{\boldsymbol{\mu}}_{\boldsymbol{x}}$ and $\hat{\boldsymbol{\mu}}_{\boldsymbol{y}}$ are the sample means

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{x}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_i \quad \text{and} \quad \hat{\boldsymbol{\mu}}_{\boldsymbol{y}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{y}_i. \tag{5.5}$$

Give an interpretation of your result!

Hint 1: Recall that

$$\hat{\boldsymbol{C}}_{\boldsymbol{xx}} = \hat{\boldsymbol{R}}_{\boldsymbol{xx}} - \hat{\boldsymbol{\mu}}_{\boldsymbol{x}}\hat{\boldsymbol{\mu}}_{\boldsymbol{x}}^{\mathrm{T}} \quad \text{and} \quad \hat{\boldsymbol{C}}_{\boldsymbol{yx}} = \hat{\boldsymbol{R}}_{\boldsymbol{yx}} - \hat{\boldsymbol{\mu}}_{\boldsymbol{y}}\hat{\boldsymbol{\mu}}_{\boldsymbol{x}}^{\mathrm{T}}.$$

Hint 2: It can be shown that

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}^{-1} = \begin{bmatrix} (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1} & -(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} \\ -\boldsymbol{D}^{-1}\boldsymbol{C}(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1} & \boldsymbol{D}^{-1} + \boldsymbol{D}^{-1}\boldsymbol{C}(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} \end{bmatrix}$$

if $\boldsymbol{D}$ and $\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}$ are regular.

e) What happens to the LS estimator and the LS estimate if $N \to \infty$?