Georgia Institute of Technology

ECE 6555 - Assignment 2

due Thursday September 22, 2022 - v1.0

- There are 8 problems over 9 pages (including the cover page).
- The problems are not necessarily in order of difficulty.
- Every question in a problem is worth 2 points, so problems with many questions are worth more than problems with few questions.
- Each question is graded as follows: no credit without meaningful work, half credit for partial work, full credit if essentially correct.
- Unless otherwise specified, you should concisely indicate your reasoning and show all relevant work.
- The grade on each question is based on our judgment of your level of understanding as reflected by what you have written. If we cannot read it, we cannot grade it.
- Please use a pen and not a pencil if you handwrite your solution.
- **You must submit your assignment on Gradescope**.

Problem 1: A separation principle

All variables in this problem are zero-mean. Consider the linear model $\mathbf{y} = \mathbf{Hx} + \mathbf{v}$ where $\mathbf{x}$ and $\mathbf{v}$ are uncorrelated with variances $\mathbf{R}_x$ and $\mathbf{R}_v$, respectively. Consider a random variable $\mathbf{z}$ uncorrelated to $\mathbf{v}$ but otherwise correlated to $\mathbf{x}$. Let $\hat{\mathbf{z}}_{|\mathrm{x}}$ denote the Linear Least Mean Square (LLMS) estimator of $\mathbf{z}$ given $\mathbf{x}$. Let $\widehat{\hat{\mathbf{z}}_{|\mathrm{x}}}$ denote the LLMS estimator of $\hat{\mathbf{z}}_{|\mathrm{x}}$ given $\mathbf{y}$. Let $\hat{\mathbf{z}}_{|\mathrm{y}}$ denote the LLMS estimator of $\mathbf{z}$ given $\mathbf{y}$.

Show that $\hat{\mathbf{z}}_{\mathrm{y}} = \widehat{\hat{\mathbf{z}}_{|\mathrm{x}}}$.

SOLUTION We avoid the case where R_y is singular for now for simplicity. In that case we know that the LLMS $\widehat{\hat{\mathbf{z}}_{|\mathrm{x}}}$ is

$$\widehat{\hat{\mathbf{z}}_{|\mathrm{x}}} = R_{\hat{\mathrm{z}}_{|\mathrm{x}} y} R_y^{-1} \mathbf{y}$$

where

$$R_{\hat{\mathrm{z}}_{|\mathrm{x}} y} = \left\langle \hat{\mathbf{z}}_{|\mathrm{x}}, \mathbf{y} \right\rangle = \left\langle \hat{\mathbf{z}}_{|\mathrm{x}} - \mathbf{z} + \mathbf{z}, \mathbf{y} \right\rangle = \left\langle \hat{\mathbf{z}}_{|\mathrm{x}} - \mathbf{z}, \mathbf{y} \right\rangle + \langle \mathbf{z}, \mathbf{y} \rangle.$$

Note that

$$\left\langle \hat{\mathbf{z}}_{|\mathrm{x}} - \mathbf{z}, \mathbf{y} \right\rangle = \left\langle \hat{\mathbf{z}}_{|\mathrm{x}} - \mathbf{z}, \mathbf{Hx} + \mathbf{v} \right\rangle = \left\langle \hat{\mathbf{z}}_{|\mathrm{x}} - \mathbf{z}, \mathbf{x} \right\rangle \mathbf{H}^\dagger + \left\langle \hat{\mathbf{z}}_{|\mathrm{x}} - \mathbf{z}, \mathbf{v} \right\rangle.$$

Finally, $\left\langle \hat{\mathbf{z}}_{|\mathrm{x}} - \mathbf{z}, \mathbf{x} \right\rangle = 0$ since $\hat{\mathbf{z}}_{|\mathrm{x}}$ is the LLMS estimate and by orthogonality of the error, and $\left\langle \hat{\mathbf{z}}_{|\mathrm{x}} - \mathbf{z}, \mathbf{v} \right\rangle = 0$ because both $\mathbf{z}$ and $\mathbf{x}$ are uncorrelated to $\mathbf{v}$. Hence,

$$R_{\hat{\mathrm{z}}_{|\mathrm{x}} y} = \langle \mathbf{z}, \mathbf{y} \rangle = R_{zy},$$

so that $\hat{\mathbf{z}}_{|\mathrm{x}} = R_{\hat{\mathrm{z}}_{|\mathrm{x}} y} R_y^{-1} \mathbf{y} = R_{zy} R_y^{-1} \mathbf{y} = \hat{\mathbf{z}}_{|\mathrm{y}}$.

Looking back at the analysis, note that we never really used the assumption that R_y was not singular, except for writing the estimates. Our analysis really shows that the normal equations for finding $\hat{\mathbf{z}}_{|\mathrm{y}}$ and $\widehat{\hat{\mathbf{z}}_{|\mathrm{x}}}$ are the same. □

Problem 2: Multiplicative noise

Consider the noisy measurement $\mathbf{y} = (1 + v)\mathbf{x}$ where $\mathbf{x}$ and v are zero mean independent random variables. The variance of v is σ_v^2. Determine the LLMS estimator of $\mathbf{x}$ given $\mathbf{y}$. Show that the Minimum Mean-Square Error (MMSE) is smaller than the variance of $\mathbf{x}$.

SOLUTION Let us assume that R_x is not singular. Note that

$$R_y = \langle \mathbf{x} + v\mathbf{x}, \mathbf{x} + v\mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + 2\mathbb{E}\left[v\right] \langle \mathbf{x}, \mathbf{x} \rangle + \mathbb{E}\left[v^2\right] \langle \mathbf{x}, \mathbf{x} \rangle = R_x(1 + \sigma_v^2).$$

Note that we have used the fact that for any function $f(\cdot)$ of v, $\langle \mathbf{x}, f(v)\mathbf{x} \rangle = \mathbb{E}\left[\mathbf{x} f(v) \mathbf{x}^\intercal\right] = \mathbb{E}\left[f(v)\right] \mathbb{E}\left[\mathbf{x}\mathbf{x}^\intercal\right]$ since v and $\mathbf{x}$ are independent and that $\mathbb{E}\left[v\right] = 0$. Similarly,

$$R_{xy} = \langle \mathbf{x}, (1 + v)\mathbf{x} \rangle = R_x + \mathbb{E}\left[v\right] R_x = R_x.$$

Hence $\hat{x} = R_{xy} R_y^{-1} \mathbf{y} = \frac{1}{1+\sigma_v^2} \mathbf{y}$.

Next note that the the MMSE is given by

$$
\begin{aligned}
\|\mathbf{x}-\hat{\mathbf{x}}\|^2 &= \left\langle \mathbf{x} - \frac{1}{1+\sigma_v^2}\mathbf{y}, \mathbf{x} - \frac{1}{1+\sigma_v^2}\mathbf{y} \right\rangle \\
&= \left\langle \frac{\sigma_v^2}{1+\sigma_v^2}\mathbf{x} - \frac{v}{1+\sigma_v^2}\mathbf{x}, \frac{\sigma_v^2}{1+\sigma_v^2}\mathbf{x} - \frac{v}{1+\sigma_v^2}\mathbf{x} \right\rangle \\
&= \left(\frac{\sigma_v^2}{1+\sigma_v^2}\right)^2 R_x + \frac{\sigma_v^2}{(1+\sigma_v^2)^2} R_x - 2\frac{\sigma_v^2}{(1+\sigma_v^2)^2} \underbrace{\langle v\mathbf{x}, \mathbf{x}\rangle}_{=0} \\
&= \frac{\sigma_v^4+\sigma_v^2}{(1+\sigma_v^2)^2} R_x \\
&= \frac{\sigma_v^2}{1+\sigma_v^2} R_x
\end{aligned}
$$

Note that $\sigma_v^2 \leq 1+\sigma_v^2$ so that $\|\mathbf{x}-\hat{\mathbf{x}}\|^2 \preceq R_x$. □

Problem 3: Defective measurement sensors

Consider a zero-mean random variable $\mathbf{x}$ with variance Π_0 and two possible measurements for $\mathbf{x}$:

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x} + \mathbf{v}_1 \qquad \mathbf{y}_2 = \mathbf{H}_2\mathbf{x} + \mathbf{v}_2.$$

where $(\mathbf{v}_1, \mathbf{v}_2)$ are zero-mean uncorrelated sensor noise with variance $\mathbf{R}_1$ and $\mathbf{R}_2$, respectively, also uncorrelated with $\mathbf{x}$. One of the measurements is defective and is either sensor 1 with probability $1-p$ or sensor 2 with probability p. Denote this measurement by $\mathbf{z}$.

[Q1] Find the LLMS estimator of $\mathbf{x}$ given $\mathbf{z}$

SOLUTION Let $s \in \{1,2\}$ denote the random variable that indicates the position of the switch. We $s=1$ with probability p and $s=2$ with probability $1-p$. The trick is to use the towering property of expectation, which says that $\mathbb{E}[f(\mathbf{z})] = \mathbb{E}_s\left[\mathbb{E}_{z|s}[f(\mathbf{z})]\right]$. This will make our life incredibly easier because conditioned on s, we know if $\mathbf{z}$ is $\mathbf{y}_1$ or $\mathbf{y}_2$.

Specifically here,

$$
\begin{aligned}
R_{xz} &= \mathbb{E}[\mathbf{x}\mathbf{z}^\intercal] = p\mathbb{E}[\mathbf{x}\mathbf{y}_1^\intercal] + (1-p)\mathbb{E}[\mathbf{x}\mathbf{y}_2^\intercal] = p\Pi_0\mathbf{H}_1^\dagger + (1-p)\Pi_0\mathbf{H}_2^\dagger \quad = \Pi_0(p\mathbf{H}_1 + (1-p)\mathbf{H}_2)^\dagger \\
R_z &= \mathbb{E}[\mathbf{z}\mathbf{z}^\intercal] = p(\mathbf{H}_1\Pi_0\mathbf{H}_1^\dagger + \mathbf{R}_1) + (1-p)(\mathbf{H}_2\Pi_0\mathbf{H}_2^\dagger + \mathbf{R}_2)
\end{aligned}
$$

We can simplify things a bit using The estimator is therefore

$$\hat{x} = R_{xz}R_z^{-1}\mathbf{z}.$$

□

[Q2] Find the corresponding MMSE

SOLUTION The corresponding MMSE □

[Q3] How does your answer change if $\mathbf{v}_1$ and $\mathbf{v}_2$ were correlated?

SOLUTION The answer does *not* change, the correlation between $\mathbf{v}_1$ and $\mathbf{v}_2$ never shows up in the proof. This is not surprising since we only get to observe either one of the sensor measurements, not both at the same time. □

[Q4] What can you say about the special case $\mathbf{H}_1 = \mathbf{H}_2$.

SOLUTION The solution the reduces to our standard LLMS estimate with a linear model seen in class! $\square$

Problem 4: Linear estimator of x^2

Consider $y = x + v$ where v and x are independent real-valued independent zero-mean Gaussian scalar random variables with variances σ_v^2 and σ_x^2, respectively. Find the LLMS estimator of the random variable x^2 using y and y^2. (*Hint:* recall that for a real-valued zero-mean Gaussian random variable z with variance σ^2, $\mathbb{E}\left[z^3\right] = 0$ ad $\mathbb{E}\left[z^4\right] = 3\sigma^4$.)

SOLUTION Note that the random variables x^2 and y^2 are *not* centered, therefore we introduce the variables $z \triangleq x^2 - \sigma_x^2$ and $\tilde{y} \triangleq y^2 - (\sigma_x^2 + \sigma_v^2)$, which are centered. We also form the vector $\mathbf{y} \triangleq \begin{bmatrix} y & \tilde{y} \end{bmatrix}^\intercal$. Our LLMS estimator of z given $\mathbf{y}$ is then $\hat{z} = R_{\tilde{x}\mathrm{y}} R_{\mathrm{y}}^{-1}\mathbf{y}$ with

$$R_{\tilde{x}\mathrm{y}} = \begin{bmatrix} \mathbb{E}\left[\tilde{x}y\right] \\ \mathbb{E}\left[\tilde{x}\tilde{y}\right] \end{bmatrix}^\intercal = \begin{bmatrix} \mathbb{E}\left[\tilde{x}x\right] + \mathbb{E}\left[\tilde{x}v\right] \\ \mathbb{E}\left[\tilde{x}x^2\right] + 2\mathbb{E}\left[\tilde{x}xv\right] + \mathbb{E}\left[\tilde{x}v^2\right] - (\sigma_x^2 + \sigma_v^2)\mathbb{E}\left[\tilde{x}\right] \end{bmatrix}^\intercal .$$

Note that

$$\begin{aligned}
\mathbb{E}\left[\tilde{x}x\right] &= \mathbb{E}\left[x^3\right] - \sigma_x^2\mathbb{E}\left[x\right] = 0 \\
\mathbb{E}\left[\tilde{x}v\right] &= \mathbb{E}\left[\tilde{x}\right]\mathbb{E}\left[v\right] = 0 \\
\mathbb{E}\left[\tilde{x}x^2\right] &= \mathbb{E}\left[x^4\right] - \sigma_x^2\mathbb{E}\left[x^2\right] = 3\sigma_x^4 - \sigma_x^4 = 2\sigma_x^4. \\
\mathbb{E}\left[\tilde{x}xv\right] &= \mathbb{E}\left[\tilde{x}x\right]\mathbb{E}\left[v\right] = 0 \\
\mathbb{E}\left[\tilde{x}v^2\right] &= \mathbb{E}\left[\tilde{x}\right]\mathbb{E}\left[v^2\right] = 0
\end{aligned}$$

so that

$$R_{\tilde{x}\mathrm{y}} = \begin{bmatrix} 0 & 2\sigma_x^4 \end{bmatrix}$$

Next, note that

$$R_{\mathrm{y}} = \mathbb{E}\left[\mathbf{y}\mathbf{y}^\intercal\right] = \begin{bmatrix} \mathbb{E}\left[y^2\right] & \mathbb{E}\left[y\tilde{y}\right] \\ \mathbb{E}\left[y\tilde{y}\right] & \mathbb{E}\left[\tilde{y}^2\right] \end{bmatrix}$$

with (since y is Gaussian as the linear combination of two Gaussian variables)

$$\begin{aligned}
&\mathbb{E}\left[\tilde{y}y\right] = \mathbb{E}\left[y^3\right] - (\sigma_x^2 + \sigma_v^2)\mathbb{E}\left[y\right] = 0 \\
&\mathbb{E}\left[y^2\right] = \sigma_x^2 + \sigma_v^2 \\
&\mathbb{E}\left[\tilde{y}^2\right] = \mathbb{E}\left[y^4\right] - 2(\sigma_x^2 + \sigma_v^2)\mathbb{E}\left[y^2\right] + (\sigma_x^2 + \sigma_v^2)^2 = 2(\sigma_x^2 + \sigma_v^2)^2
\end{aligned}$$

so that

$$R_{\mathrm{y}} = \begin{bmatrix} (\sigma_x^2 + \sigma_v^2) & 0 \\ 0 & 2(\sigma_x^2 + \sigma_v^2)^2 \end{bmatrix}$$

Putting everything together, we have

$$\begin{aligned}
\hat{z} &= \begin{bmatrix} 0 & 2\sigma_x^4 \end{bmatrix} \begin{bmatrix} (\sigma_x^2 + \sigma_v^2) & 0 \\ 0 & 2(\sigma_x^2 + \sigma_v^2)^2 \end{bmatrix}^{-1} \begin{bmatrix} y \\ \tilde{y} \end{bmatrix} \\
&= \frac{\sigma_x^4}{(\sigma_x^2 + \sigma_v^2)}\tilde{y}.
\end{aligned}$$

Finally, we obtain

$$\hat{x} = \sigma_x^2 + \frac{\sigma_x^4}{(\sigma_x^2+\sigma_v^2)}y^2 - \frac{\sigma_x^4}{\sigma_x^2+\sigma_v^2} = \frac{\sigma_x^4}{(\sigma_x^2+\sigma_v^2)}y^2 - \frac{\sigma_x^2\sigma_v^2}{\sigma_x^2+\sigma_v^2}.$$

As a sanity check, if $\sigma_v^2 = 0$, we should estimate x^2 as y^2, which is the case. Similarly, if $\sigma_v^2 \to \infty$, we should estimate x^2 as σ_x^2, which is the case.

□

Problem 5: Separation of signal and structured noise

Consider the model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{S}\boldsymbol{\theta} + \mathbf{v}$$

where $\mathbf{v}$ is a zero mean additive noise random vector with unit variance and $\mathbf{x}, \boldsymbol{\theta}$ are *unknown* constant vectors. The matrices $\mathbf{H} \in \mathbb{C}^{m\times n}$ and $\mathbf{S} \in \mathbb{C}^{m\times p}$ are known and such that $\begin{bmatrix} \mathbf{H} & \mathbf{S} \end{bmatrix}$ is full rank and $m \geq n+p$. The term $\mathbf{S}\boldsymbol{\theta}$ is interpreted as a *perturbation* while the term $\mathbf{H}\mathbf{x}$ is the useful signal we wish to separate.

[Q1] Define the vector $\mathbf{z} = \begin{bmatrix} \mathbf{x} & \boldsymbol{\theta} \end{bmatrix}^\mathsf{T}$. Determine the optimal *unbiased* estimator $\hat{\mathbf{z}}$ of $\mathbf{z}$ given $\mathbf{y}$.

SOLUTION This is exactly the setting of the Gauss-Markov Theorem, the optimal unbiased estimator is

$$\begin{aligned}\hat{\mathbf{z}} &= (\begin{bmatrix} \mathbf{H} & \mathbf{S} \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{H} & \mathbf{S} \end{bmatrix})^{-1} \begin{bmatrix} \mathbf{H} & \mathbf{S} \end{bmatrix}^\dagger \mathbf{y} \\ &= \begin{bmatrix} \mathbf{H}^\dagger\mathbf{H} & \mathbf{H}^\dagger\mathbf{S} \\ \mathbf{S}^\dagger\mathbf{H} & \mathbf{S}^\dagger\mathbf{S} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{H}^\dagger\mathbf{y} \\ \mathbf{S}^\dagger\mathbf{y} \end{bmatrix}\end{aligned}$$

□

[Q2] Write $\hat{\mathbf{z}} = \begin{bmatrix} \hat{\mathbf{x}} & \hat{\boldsymbol{\theta}} \end{bmatrix}$ to separate the estimation of $\mathbf{x}$ and $\boldsymbol{\theta}$. Let $\hat{\mathbf{s}} \triangleq \mathbf{H}\hat{\mathbf{x}}$ denote the estimator of $\mathbf{s} \triangleq \mathbf{H}\mathbf{x}$. Show that

$$\hat{\mathbf{s}} = \mathbf{E}\mathbf{y} \text{ with } \mathbf{E} = P_\mathrm{H}\left[\mathbf{I} - \mathbf{S}(\mathbf{S}^\dagger P_\mathrm{H}^\perp \mathbf{S})^{-1}\mathbf{S}^\dagger P_\mathrm{H}^\perp\right] = \mathbf{H}(\mathbf{H}^\dagger P_\mathrm{S}^\perp \mathbf{H})^{-1}\mathbf{H}^\dagger P_\mathrm{S}^\perp.$$

with $P_\mathrm{H}^\perp = \mathbf{I} - P_\mathrm{H}$, $P_\mathrm{S}^\perp = \mathbf{I} - P_\mathrm{S}$, and P_H and P_S are the orthogonal projection matrices on the space spanned by the rows of their respective matrices.

SOLUTION This requires a bit of work playing with matrix identities. Using the block matrix inversion lemma, note that the two blocks on the first line of the inverse matrix are

$$\begin{aligned}&(\mathbf{H}^\dagger\mathbf{H} - \mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger\mathbf{S})^{-1}\mathbf{S}^\dagger\mathbf{H})^{-1} \\ &\qquad \triangleq (\mathbf{H}^\dagger P_\mathrm{S}^\perp \mathbf{H})^{-1} \\ &\qquad = (\mathbf{H}^\dagger\mathbf{H})^{-1} + (\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger\mathbf{S} - \mathbf{S}^\dagger\mathbf{H}(\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S})^{-1}\mathbf{S}^\dagger\mathbf{H}(\mathbf{H}^\dagger\mathbf{H})^{-1} \\ &\qquad = (\mathbf{H}^\dagger\mathbf{H})^{-1} + (\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger P_\mathrm{H}^\perp \mathbf{S})^{-1}\mathbf{S}^\dagger\mathbf{H}(\mathbf{H}^\dagger\mathbf{H})^{-1}\end{aligned}$$

where we have used the matrix inversion lemma and

$$
\begin{aligned}
-&(\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger\mathbf{S}-\mathbf{S}^\dagger\mathbf{H}(\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S})^{-1}\\
&\triangleq -(\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger P_\mathrm{H}^\perp\mathbf{S})\\
&= -(\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}((\mathbf{S}^\dagger\mathbf{S})^{-1}-(\mathbf{S}^\dagger\mathbf{S})^{-1}\mathbf{S}^\dagger\mathbf{H}(-\mathbf{H}^\dagger\mathbf{H}+\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger\mathbf{S})^{-1}\mathbf{S}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger\mathbf{S})^{-1})\\
&= -(\mathbf{H}^\dagger\mathbf{H})^{-1}(\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger\mathbf{S})^{-1}-\mathbf{H}^\dagger P_\mathrm{S}\mathbf{H}(-\mathbf{H}^\dagger\mathbf{H}+\mathbf{H}^\dagger P_\mathrm{S}\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger\mathbf{S})^{-1})\\
&= -(\mathbf{H}^\dagger\mathbf{H})^{-1}(I-\mathbf{H}^\dagger P_\mathrm{S}\mathbf{H}(-\mathbf{H}^\dagger\mathbf{H}+\mathbf{H}^\dagger P_\mathrm{S}\mathbf{H})^{-1})\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger\mathbf{S})^{-1}\\
&= -(\mathbf{H}^\dagger\mathbf{H})^{-1}((-\mathbf{H}^\dagger\mathbf{H}+\mathbf{H}^\dagger P_\mathrm{S}\mathbf{H})-\mathbf{H}^\dagger P_\mathrm{S}\mathbf{H})(-\mathbf{H}^\dagger\mathbf{H}+\mathbf{H}^\dagger P_\mathrm{S}\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger\mathbf{S})^{-1}\\
&= -(\mathbf{H}^\dagger P_\mathrm{S}^\perp\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger\mathbf{S})^{-1}
\end{aligned}
$$

where we have again used the matrix inversion lemma and banged our head against the wall long enough to rearrange terms in a pleasing way. Combining the two forms of the terms, we obtain

$$
\begin{aligned}
\hat{s} &= \mathbf{H}(\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{y}+\mathbf{H}(\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger P_\mathrm{H}^\perp\mathbf{S})^{-1}\mathbf{S}^\dagger\mathbf{H}(\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{y}\\
&\quad -\mathbf{H}(\mathbf{H}^\dagger\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger P_\mathrm{H}^\perp\mathbf{S})\mathbf{S}^\dagger\mathbf{y}\\
&= P_\mathrm{H}\mathbf{y}+P_\mathrm{H}\mathbf{S}(\mathbf{S}^\dagger P_\mathrm{H}^\perp\mathbf{S})^{-1}\mathbf{S}^\dagger P_\mathrm{H}\mathbf{y}-P_\mathrm{H}\mathbf{S}(\mathbf{S}^\dagger P_\mathrm{H}^\perp\mathbf{S})^{-1}\mathbf{S}^\dagger\mathbf{y}\\
&= P_\mathrm{H}[I-\mathbf{S}(\mathbf{S}^\dagger P_\mathrm{H}^\perp\mathbf{S})^{-1}\mathbf{S}^\dagger P_\mathrm{H}^\perp]\mathbf{y}.
\end{aligned}
$$

and combining the two other forms, we obtain

$$
\begin{aligned}
\hat{s} &= \mathbf{H}(\mathbf{H}^\dagger P_\mathrm{S}^\perp\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{y}-\mathbf{H}(\mathbf{H}^\dagger P_\mathrm{S}^\perp\mathbf{H})^{-1}\mathbf{H}^\dagger\mathbf{S}(\mathbf{S}^\dagger\mathbf{S})^{-1}\mathbf{S}^\dagger\mathbf{y}\\
&= \mathbf{H}(\mathbf{H}^\dagger P_\mathrm{S}^\perp\mathbf{H})^{-1}\mathbf{H}^\dagger P_\mathrm{S}^\perp\mathbf{y}
\end{aligned}
$$

□

[Q3] Conclude that $\mathbf{ES}=0$ and provide a geometric interpretation.

SOLUTION Since $P_\mathrm{S}^\perp\mathbf{S}=0$, the second form of $\mathbf{E}$ ensures that $\mathbf{ES}=0$. Geometrically, this means that the estimate $\hat{s}$ of $\mathbf{Hx}$ is projecting on the subspace orthogonal to the columns of $\mathbf{S}$. □

[Q4] Let $\tilde{\mathbf{s}}=\mathbf{s}-\hat{\mathbf{s}}$. Show that the mean square error $\mathbb{E}\left[\tilde{\mathbf{s}}\tilde{\mathbf{s}}^\dagger\right]$ is $\mathbf{EE}^\dagger$.

SOLUTION Using the first form of $\mathbf{E}$, note that $\mathbf{EHx}=P_\mathrm{H}\mathbf{Hx}=\mathbf{Hx}$, so that $\mathbf{s}-\mathbf{EHx}=0$. Hence

$$\tilde{\mathbf{s}}=-\mathbf{Ev}\text{ and }\mathbb{E}\left[\tilde{\mathbf{s}}\tilde{\mathbf{s}}^\dagger\right]=\mathbf{E}R_v\mathbf{E}^\dagger=\mathbf{EE}^\dagger.$$

□

[Q5] Assume now that $\mathbf{x}$ is a zero mean random variable with known variance $\Pi_0>0$. Show that the LLMS of $\mathbf{s}=\mathbf{Hx}$ is now

$$\hat{\mathbf{s}}=\mathbf{Fy}\text{ with }\mathbf{F}=P_\mathrm{H}\left[\mathbf{I}-\mathbf{S}(\mathbf{S}^\dagger P_\mathrm{H}^\perp\mathbf{S})^{-1}\mathbf{S}^\dagger P_\mathrm{H}^\perp\right]$$

with $P_\mathrm{H}^\perp=\mathbf{I}-P_\mathrm{H}$ and $P_\mathrm{H}=\mathbf{H}(\mathbf{H}^\dagger\mathbf{H}+\Pi_0)^{-1}\mathbf{H}$. Make sure you understand the difference with the previous situation where $\mathbf{x}$ was modeled as a constant.

SOLUTION We treat $\boldsymbol{\theta}$ as a random variable with covariance Π_1 infinite to capture the fact that it is actually a constant. We also treat it as uncorrelated to $\mathbf{x}$ since $\mathbb{E}\left[\mathbf{x}\boldsymbol{\theta}\right]=0$. Consequently, the LLMS estimate of the vector $\begin{bmatrix}\mathbf{x} & \boldsymbol{\theta}\end{bmatrix}^\mathbf{T}$ is

$$\hat{\mathbf{z}}=\left(\begin{bmatrix}\Pi_0 & 0\\ 0 & \Pi_1\end{bmatrix}^{-1}+\begin{bmatrix}\mathbf{H}^\dagger\\ \mathbf{S}^\dagger\end{bmatrix}\begin{bmatrix}\mathbf{H} & \mathbf{S}\end{bmatrix}\right)^{-1}\begin{bmatrix}\mathbf{H}^\dagger\\ \mathbf{S}^\dagger\end{bmatrix}\mathbf{y}$$

where we have used the alternative form for the LLMS for the linear model obtained after using the matrix inversion lemma. Since we want θ to be deterministic, we can take $\Pi_1^{-1} = 0$. This is essentially the same calculation as earlier, plus some bookkeeping to keep track of Π_0^{-1}. Hence we obtain immediately

$$\mathbf{F} = P_{\mathrm{H}} \left[\mathbf{I} - \mathbf{S}(\mathbf{S}^{\dagger} P_{\mathrm{H}}^{\perp} \mathbf{S})^{-1} \mathbf{S}^{\dagger} P_{\mathrm{H}}^{\perp} \right]$$

with P_{H} modified to account for Π_0^{-1}. □

[Q6] Verify that $\mathbf{FS} = 0$.

SOLUTION This follows by inspection since

$$\left[\mathbf{I} - \mathbf{S}(\mathbf{S}^{\dagger} P_{\mathrm{H}}^{\perp} \mathbf{S})^{-1} \mathbf{S}^{\dagger} P_{\mathrm{H}}^{\perp} \right] \mathbf{S} = \mathbf{S} - \mathbf{S} = 0.$$

□

[Q7] Compute the new MMSE and compare with the previous result.

SOLUTION Note that $\mathbf{FHx} = P_{\mathrm{H}} \mathbf{Hx} = \mathbf{Hx}$ so that $\mathbf{s} - \mathbf{Fy} = -\mathbf{Fv}$ and

$$\mathbb{E}\left[(\mathbf{s} - \mathbf{Fy})(\mathbf{s} - \mathbf{Fy})^{\dagger} \right] = \mathbf{FF}^{\dagger}$$

□

Problem 6: General combined estimator

Let $\mathbf{y}_1$ and $\mathbf{y}_2$ be two separate observations of a zero-mean random variable $\mathbf{x}$ using the linear models

$$\mathbf{y}_1 = \mathbf{H}_1 \mathbf{x} + \mathbf{v}_1 \qquad \mathbf{y}_2 = \mathbf{H}_2 \mathbf{x} + \mathbf{v}_2$$

and

$$\left\langle \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{x} \end{bmatrix}, \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{x} \end{bmatrix} \right\rangle = \begin{bmatrix} R_1 & 0 \\ 0 & M_1 \end{bmatrix} \qquad \left\langle \begin{bmatrix} \mathbf{v}_2 \\ \mathbf{x} \end{bmatrix}, \begin{bmatrix} \mathbf{v}_2 \\ \mathbf{x} \end{bmatrix} \right\rangle = \begin{bmatrix} R_2 & 0 \\ 0 & M_2 \end{bmatrix}.$$

Note that this is a bit different from what we saw in class because the covariance matrix of $\mathbf{x}$ *is different in both experiments.*

Let $\hat{\mathbf{x}}_1$ be the LLMS estimate of $\mathbf{x}$ from $\mathbf{y}_1$ with corresponding error covariance matrix P_1. Let $\hat{\mathbf{x}}_2$ be the LLMS estimate of $\mathbf{x}$ from $\mathbf{y}_2$ with corresponding error covariance matrix P_2. Let $\hat{\mathbf{x}}$ be the LLMS estimate of $\mathbf{x}$ from $\mathbf{y}_1$ and $\mathbf{y}_2$ assuming $\langle \mathbf{x}, \mathbf{x} \rangle = \Pi$ with corresponding error covariance matrix P.

Show that $P^{-1}\hat{\mathbf{x}} = P_1^{-1}\hat{\mathbf{x}}_1 + P_2^{-1}\hat{\mathbf{x}}_2$ and

$$P^{-1} = P_1^{-1} + P_2^{-1} + \Pi^{-1} - M_1^{-1} - M_2^{-1}.$$

Make sure your provide enough details to justify your answer, but you can of course use any result seen in class. Just make sure that you clearly indicate which results you use!

Problem 7: Optimal estimation for exponential distribution

Suppose $y = x + v$ where x and v are independent real-valued random variables with exponential distribution of parameters $\lambda > 0$ and $\mu > 0$, respectively. Recall that that an exponential distribution is of the form $\lambda e^{-\lambda x}$ for $x \geq 0$, with mean λ^{-1} and variance λ^{-2}.

[Q1] Show that $p(x, y) = \lambda \mu e^{-(\lambda - \mu)x} e^{-\mu y}$ for $x \leq y$ and 0 else.

[Q2] Show that $p(y) = \frac{\lambda\mu}{\lambda-\mu}(e^{-\mu y} - e^{-\lambda y})$

[Q3] Show that the optimal (non linear) least mean square estimate of x given y is

$$\hat{x} = \frac{1}{\lambda - \mu} - \frac{e^{-\lambda y}}{e^{-\mu y} - e^{-\lambda y}} y.$$

[Q4] Calculate the LLMS (*Hint:* x and y are not centered!) and compare with the result above.

Problem 8: An optimal nonlinear estimator for binary signals

Consider the observations $y_i = x + v_i$ where x and v_i are independent real-valued random variables and $\{v_i\}_{i\geq 0}$ is a white noise Gaussian process with zero mean and unit variance.

[Q1] Assume that the variable x takes the values ± 1 with equal probability. Show that the optimal nonlinear Least Mean Square (LMS) estimator of x given n observations $\{y_i\}_{i=0}^{n-1}$ is

$$\hat{x}_n = \tanh\left(\sum_{i=0}^{n-1} y_i\right)$$

[Q2] Now assume that the variable x takes the value $+1$ with probability p and -1 with probability $1-p$. Show that the optimal nonlinear LMS estimator of x given n observations $\{y_i\}_{i=0}^{n-1}$ is now

$$\hat{x}_n = \tanh\left(\frac{1}{2}\ln\left(\frac{p}{1-p}\right) + \sum_{i=0}^{n-1} y_i\right).$$

SOLUTION We need to compute $\mathbb{E}[x|\mathbf{y}]$ where $\mathbf{y} = \{y_i\}_{i=0}^{n-1}$. Note that $p(x|\mathbf{y})$ is given by

$$p(x|\mathbf{y}) = \frac{p(\mathbf{y}|x)p(x)}{p(\mathbf{y})} = \frac{p_v(\mathbf{y} - x)p(x)}{p(\mathbf{y})}.$$

Note that by definition

$$p_v(\mathbf{v}) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{\mathbf{v}^\intercal \mathbf{v}}{2}\right).$$

So that

$$p(\mathbf{y}) = p(\mathbf{y}|x = +1)p + p(\mathbf{y}|x = -1)(1-p) = p_v(\mathbf{y} - \mathbf{1})p + p_v(\mathbf{y} + \mathbf{1})(1-p).$$

Hence,

$$p(x|\mathbf{y}) = \frac{p_v(\mathbf{y} - x\mathbf{1})p(x)}{p_v(\mathbf{y} - \mathbf{1})p + p_v(\mathbf{y} + \mathbf{1})(1-p)} = \frac{\exp\left(-\frac{(\mathbf{y}-x\mathbf{1})^\intercal(\mathbf{y}-x\mathbf{1})}{2}\right)p(x)}{\exp\left(-\frac{(\mathbf{y}-\mathbf{1})^\intercal(\mathbf{y}-\mathbf{1})}{2}\right)p + \exp\left(-\frac{(\mathbf{y}+\mathbf{1})^\intercal(\mathbf{y}+\mathbf{1})}{2}\right)(1-p)}$$

Now,

$$
\begin{aligned}
\mathbb{E}\left[x|\mathbf{y}\right] &= p(1|\mathbf{y}) - p(-1|\mathbf{y}) \\
&= \frac{\exp\left(-\frac{(\mathrm{y}-1)^\intercal(\mathrm{y}-1)}{2}\right)p - \exp\left(-\frac{(\mathrm{y}+1)^\intercal(\mathrm{y}+1)}{2}\right)(1-p)}{\exp\left(-\frac{(\mathrm{y}-1)^\intercal(\mathrm{y}-1)}{2}\right)p + \exp\left(-\frac{(\mathrm{y}+1)^\intercal(\mathrm{y}+1)}{2}\right)(1-p)} \\
&= \frac{\exp\left(\mathbf{y}^\intercal\mathbf{1}\right)p - \exp\left(-\mathbf{y}^\intercal\mathbf{1}\right)(1-p)}{\exp\left(\mathbf{y}^\intercal\mathbf{1}\right)p + \exp\left(-\mathbf{y}^\intercal\mathbf{1}\right)(1-p)} \\
&= \frac{1 - \exp\left(-2\mathbf{y}^\intercal\mathbf{1} + \ln\frac{1-p}{p}\right)}{1 + \exp\left(-2\mathbf{y}^\intercal\mathbf{1} + \ln\frac{1-p}{p}\right)} \\
&= \tanh\left(\frac{1}{2}\ln\frac{p}{1-p} + \mathbf{y}^\intercal\mathbf{1}\right).
\end{aligned}
$$

$\square$