Solution to Problem Set 6

Problem 1:

1. Notice that:

$$Y_1 = X_1 + N_1$$
$$Y_2 = X_2 + N_2$$

$$h_{MAP,\underline{Y}\to X_1}\left(\underline{y}\right) = arg \max_{x_1\in\{-1,1\}} Pr\left\{X_1 = x_1 | \underline{Y} = \underline{y}\right\} =$$

$$= arg \max_{x_1\in\{-1,1\}} \frac{f_{\underline{Y}|X_1}\left(\underline{y}|x_1\right) Pr\{X_1 = x_1\}}{f_{\underline{Y}}\left(\underline{y}\right)} =$$

$$= arg \max_{x_1\in\{-1,1\}} \left(\sum_{x_2\in\{-1,1\}} f_{\underline{Y}|X_1X_2}\left(\underline{y}|x_1,x_2\right) Pr\{X_2 = x_2 | X_1 = x_1\}\right) Pr\{X_1 = x_1\} \underset{(@)}{=}$$

$$= arg \max_{x_1\in\{-1,1\}} \left(\sum_{x_2\in\{-1,1\}} f_{\underline{N}|X_1X_2}(y_1 - x_1, y_2 - x_2 | x_1, x_2) Pr\{X_2 = x_2\}\right) Pr\{X_1 = x_1\} =$$

$$= arg \max_{x_1\in\{-1,1\}} \left(\sum_{x_2\in\{-1,1\}} f_{N_1}(y_1 - x_1) f_{N_2}(y_2 - x_2) Pr\{X_2 = x_2\}\right) Pr\{X_1 = x_1\} =$$

$$= arg \max_{x_1\in\{-1,1\}} \left(\sum_{x_2\in\{-1,1\}} f_{N_2}(y_2 - x_2) Pr\{X_2 = x_2\}\right) Pr\{X_1 = x_1\} f_{N_1}(y_1 - x_1) =$$

$$= arg \max_{x_1\in\{-1,1\}} Pr\{X_1 = x_1\} f_{N_1}(y_1 - x_1) = arg \max_{x_1\in\{-1,1\}} Pr\{X_1 = x_1 | Y_1 = y_1\}$$

The optimal estimator is given according to the scalar MAP rule, that is:

$$Pr\{X_1 = 1 | Y_1 = y_1\} \underset{-1}{\overset{1}{\gtrless}} Pr\{X_1 = -1 | Y_1 = y_1\}$$

$$\frac{f_{Y_1|X_1}(y_1|1) Pr\{X_1 = 1\}}{f_{Y_1}(y_1)} \underset{-1}{\overset{1}{\gtrless}} \frac{f_{Y_1|X_1}(y_1|-1) Pr\{X_1 = -1\}}{f_{Y_1}(y_1)}$$

$$\frac{1}{\sqrt{2\sigma^2}} e^{-\frac{(y_1+1)^2}{2\sigma^2}} \cdot (1-p) \underset{-1}{\overset{1}{\gtrless}} \frac{1}{\sqrt{2\sigma^2}} e^{-\frac{(y_1-1)^2}{2\sigma^2}} \cdot p$$

$$e^{-\frac{(y_1+1)^2}{2\sigma^2} + \frac{(y_1-1)^2}{2\sigma^2}} \underset{-1}{\overset{1}{\gtrless}} \frac{p}{1-p}$$

$$\Rightarrow h_{MAP,\underline{Y}\to X_1}\left(\underline{y}\right) := y_1 \underset{-1}{\overset{1}{\gtrless}} \underbrace{\frac{\sigma^2}{2} ln\left(\frac{p}{1-p}\right)}_{Th(\sigma^2,p)}$$

Similarly:

$$h_{MAP,\underline{Y}\to X_2}\left(\underline{y}\right) := \quad y_2 \underset{-1}{\overset{1}{\gtrless}} \underbrace{\frac{\sigma^2}{2} \ln\left(\frac{p}{1-p}\right)}_{Th\left(\sigma^2, p\right)}$$

The probability of error in X_i:

$$\begin{aligned}
\Pr\{E\} &= \Pr\{E \mid X_i = 1\}\Pr\{X_i = 1\} + \Pr\{E \mid X_i = -1\}\Pr\{X_i = -1\} = \\
&= (1-p)\Pr\{Y < Th(\sigma^2, p) \mid X_i = 1\} + p\Pr\{Y > Th(\sigma^2, p) \mid X_i = -1\} = \\
&= (1-p)\cdot\left(1 - Q\left(\frac{Th(\sigma^2, p) - 1}{\sigma}\right)\right) + p\cdot Q\left(\frac{Th(\sigma^2, p) + 1}{\sigma}\right) = \\
&= (1-p)\cdot Q\left(\frac{1 - Th(\sigma^2, p)}{\sigma}\right) + p\cdot Q\left(\frac{Th(\sigma^2, p) + 1}{\sigma}\right)
\end{aligned}$$

For $p = \frac{1}{2}$: $\Pr\{E\} = Q\left(\frac{1}{\sigma}\right)$ as we saw in class.

2. For the above case, the estimation function is $sign(\bullet)$. The probability of error:

$$\begin{aligned}
&\Pr\{E\} = \Pr\{E \mid X_1 = 1\}\frac{1}{2} + \Pr\{E \mid X_1 = -1\}\frac{1}{2} = \\
&= \Pr\{Y_1 < 0 \mid X_1 = 1\}\frac{1}{2} + \Pr\{Y_1 > 0 \mid X_1 = -1\}\frac{1}{2} = \\
&= \frac{1}{2}\sum_{x_2 \in \{-1,1\}} \Pr\{Y_1 < 0 \mid X_1 = 1, X_2 = x_2\}\Pr\{X_2 = x_2 \mid X_1 = 1\} + \\
&+ \frac{1}{2}\sum_{x_2 \in \{-1,1\}} \Pr\{Y_1 > 0 \mid X_1 = -1, X_2 = x_2\}\Pr\{X_2 = x_2 \mid X_1 = -1\} = \\
&= \frac{1}{4}\left(Q\left(\frac{1+\alpha}{\sigma}\right) + Q\left(\frac{1-\alpha}{\sigma}\right) + Q\left(\frac{1+\alpha}{\sigma}\right) + Q\left(\frac{1-\alpha}{\sigma}\right)\right) = \\
&= \frac{1}{2}\left(Q\left(\frac{1+\alpha}{\sigma}\right) + Q\left(\frac{1-\alpha}{\sigma}\right)\right)
\end{aligned}$$

3. First, let us find the inverse matrix of H:

$$H^{-1} = \frac{1}{1-\alpha^2}\begin{bmatrix} 1 & -\alpha \\ -\alpha & 1 \end{bmatrix}$$

$$\underline{W} = [X_1 \ X_2]^T + \underbrace{H^{-1}\underline{N}}_{\underline{\tilde{N}}}$$

$\underline{\tilde{N}}$ is also a normal random vector:

$$E\{\underline{\tilde{N}}\} = \underline{0}; \ C_{\tilde{N}\tilde{N}} = H^{-1}\begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}(H^{-1})^T$$

$$= \frac{\sigma^2}{(1-\alpha^2)^2}\begin{bmatrix} 1 & -\alpha \\ -\alpha & 1 \end{bmatrix}\begin{bmatrix} 1 & -\alpha \\ -\alpha & 1 \end{bmatrix} = \frac{\sigma^2}{(1-\alpha^2)^2}\begin{bmatrix} 1+\alpha^2 & -2\alpha \\ -2\alpha & 1+\alpha^2 \end{bmatrix}$$

For $p = \frac{1}{2}$, the estimator is the $sign(\bullet)$ function. Furthermore, based on the probability of error from section 1:

$$\Pr\{E\} = Q\left(\frac{(1-\alpha^2)}{\sigma\sqrt{(1+\alpha^2)}}\right)$$

Problem 2:

1. From the linearity property of the expected value:

$$\hat{W}^{\text{MMSE}} = E[W \mid Y] = E[a^T X \mid Y] = a^T E[X \mid Y] = a^T \hat{X}^{\text{MMSE}}$$

2.

$$MSE = Var(W) - Var(\hat{W}^{\text{MMSE}}) = Var(a^T X) - Var(a^T \hat{X}^{\text{MMSE}})$$
$$= a^T C_X a - a^T C_{\hat{X}^{\text{MMSE}}} a = a^T \left[C_X - C_{\hat{X}^{\text{MMSE}}} \right] a = a^T C_{\text{MMSE}} a$$

3. We will show this from the perpendicularity principle: (this solution is correct also for the case of non-invertible covariance matrix of *Y*). We are interested in checking the "candidate" $a^T \hat{X}^{BLE}$. Let *Z* be a random variable that is a **linear** function of *Y*. Let us check whether the estimation error is orthogonal to Z:

$$E\left[\left(W - a^T \hat{X}^{\text{BLE}}\right) Z\right] = E\left[\left(a^T X - a^T \hat{X}^{\text{BLE}}\right) Z\right] = a^T \cdot \underbrace{E\left[\left(X - \hat{X}^{\text{BLE}}\right) Z\right]}_{0} = 0$$

We got that the estimation error of $a^T \hat{X}^{BLE}$ is orthogonal to all linear functions of the measurement Y, and from here we get that this is the optimal linear estimator (in the sense of MMSE) of X from Y.

4.

$$MSE = Var(W) - Var(\hat{W}^{\text{BLE}}) = Var(a^T X) - Var(a^T \hat{X}^{\text{BLE}})$$
$$= a^T C_X a - a^T C_{\hat{X}^{\text{BLE}}} a = a^T \left[C_X - C_{\hat{X}^{\text{BLE}}} \right] a = a^T C_{\text{BLE}} a$$

5.

$$C_{\underline{e}} = E\left\{\left(\underline{e} - \underline{\eta}_e\right)\left(\underline{e} - \underline{\eta}_e\right)^T\right\} = E\left\{\left(\underline{X} - \underline{\hat{X}} - \underline{\eta}_e\right)\left(\underline{X} - \underline{\hat{X}} - \underline{\eta}_e\right)^T\right\} =$$
$$E\left\{\left(\underline{X} - \underline{\hat{X}}_{MMSE} + \underline{\hat{X}}_{MMSE} - \underline{\hat{X}} - \underline{\eta}_e\right)\left(\underline{X} - \underline{\hat{X}}_{MMSE} + \underline{\hat{X}}_{MMSE} - \underline{\hat{X}} - \underline{\eta}_e\right)^T\right\} =$$
$$= E\left\{\left(\left(\underline{X} - \underline{\hat{X}}_{MMSE}\right) + \left(\underline{\hat{X}}_{MMSE} - \underline{\hat{X}} - \underline{\eta}_e\right)\right)\left(\left(\underline{X} - \underline{\hat{X}}_{MMSE}\right) + \left(\underline{\hat{X}}_{MMSE} - \underline{\hat{X}} - \underline{\eta}_e\right)\right)^T\right\} =$$
$$= \underbrace{E\left\{\underline{e}_{MMSE}\, \underline{e}_{MMSE}^T\right\}}_{= C_{\underline{e}_{MMSE}}} + 0 + 0 + PSD \Rightarrow C_{\underline{e}} - C_{\underline{e}_{MMSE}} \;\; is \;\; PSD$$

6. We need to prove that:

$$E\left\{\underline{e}\,\underline{e}^T\right\} - E\left\{\underline{e}_{MMSE}\,\underline{e}_{MMSE}{}^T\right\} \geq 0$$
$$\Leftrightarrow \underline{a}^T \left(E\left\{\underline{e}\,\underline{e}^T\right\} - E\left\{\underline{e}_{MMSE}\,\underline{e}_{MMSE}{}^T\right\}\right) \underline{a} \geq 0 \qquad for\ all\ \underline{a}$$
$$\Leftrightarrow \underline{a}^T E\left\{\underline{e}\,\underline{e}^T\right\} \underline{a} \geq \underline{a}^T E\left\{\underline{e}_{MMSE}\,\underline{e}_{MMSE}{}^T\right\} \underline{a}$$

But we saw in section 1 that for MMSE estimator it holds that $MSE_{MMSE} = \underline{a}^T E\left\{\underline{e}_{MMSE}\,\underline{e}_{MMSE}{}^T\right\} \underline{a}$ and, similarly, $MSE_e = \underline{a}^T E\left\{\underline{e}\,\underline{e}^T\right\} \underline{a}$, and we know that the MSE of all estimators is at least the MSE of the optimal estimator.

Problem 3:

1. For $m<n$:

$$\hat{Y}_n = E\{Y_n \mid Y_m\} = E\left\{\sum_{i=0}^{n} X_i \mid Y_m\right\} = E\left\{\sum_{i=0}^{m} X_i + \sum_{i=m+1}^{n} X_i \mid Y_m\right\} = E\left\{Y_m + \sum_{i=m+1}^{n} X_i \mid Y_m\right\} =$$

$$= E\{Y_m \mid Y_m\} + E\left\{\sum_{i=m+1}^{n} X_i \mid \sum_{i=0}^{m} X_i\right\} \underset{(1)}{=} E\{Y_m \mid Y_m\} + E\left\{\sum_{i=m+1}^{n} X_i\right\} = Y_m + \left\{\sum_{i=m+1}^{n} \underbrace{E\{X_i\}}_{0\forall i}\right\} = Y_m$$

where (1) follows from the fact that $\{X_i\}$ are i.i.d.

For $m=n$: $\hat{Y}_n = E\{Y_n \mid Y_m\} = E\{Y_m \mid Y_m\} = Y_m$

And overall: $\hat{Y}_n = Y_m$

2.

$$\hat{Y}_n = E\{Y_n \mid Y_1, Y_2, \ldots, Y_m\} = E\left\{\sum_{i=0}^{n} X_i \mid Y_1, Y_2, \ldots, Y_m\right\} = E\left\{\sum_{i=0}^{m} X_i + \sum_{i=m+1}^{n} X_i \mid Y_1, Y_2, \ldots, Y_m\right\} =$$

$$= E\left\{Y_m + \sum_{i=m+1}^{n} X_i \mid Y_1, Y_2, \ldots, Y_m\right\} = E\{Y_m \mid Y_1, Y_2, \ldots, Y_m\} + E\left\{\sum_{i=m+1}^{n} X_i \mid Y_1, Y_2, \ldots, Y_m\right\} \underset{(2)}{=}$$

$$= E\{Y_m \mid Y_1, Y_2, \ldots, Y_m\} + E\left\{\sum_{i=m+1}^{n} X_i\right\} = Y_m + \left\{\sum_{i=m+1}^{n} \underbrace{E\{X_i\}}_{0\forall i}\right\} = Y_m$$

where (2) is of the same reason: $\{X_i\}$ are i.i.d.

Therefore, again we got that $\hat{Y}_n = Y_m$. Namely, $Y_1, Y_2, \ldots, Y_{m-1}$ do not contribute information beyond what is already found in Y_m for the estimation of Y_n.

3. The question may be extended as follows: repeat section 1 where, now, m is a random variable that takes the values $0,1,\ldots,M$ with some probabilities (their sum must be 1 of course) when $M \le n$:

$$\hat{Y}_n = E\{Y_n \mid Y_m\} = E\left\{\underbrace{E\{Y_n \mid Y_m, m\}}_{(*)\ \ =Y_m} \mid Y_m\right\} = E\{Y_m \mid Y_m\} = Y_m$$

where (*) follows from section 1.

Problem 4:

1. From the fact that the MMSE estimator is unbiased we get:

$$E\{X\} = E\left\{\hat{X}\right\} = 1\cdot \Pr\{Y>0\} + (-1)\cdot \Pr\{Y \le 0\} = \frac{1}{2} - \frac{1}{2} = 0$$

2. From the perpendicularity principle:

$$0 = E\left\{\left(X - \hat{X}\right)Y\right\} = E\{XY\} - E\left\{\hat{X}Y\right\}$$

$$\Rightarrow E\{XY\} = E\{\hat{X}Y\} = E\{\text{sign}(Y)\cdot Y\} = \int_{-1}^{1} \frac{1}{2}\text{sgn}(y)\cdot y\,dy = \int_{-1}^{1} \frac{1}{2}|y|\,dy = 2\int_{0}^{1} \frac{1}{2} y dy = \frac{1}{2}$$

$$\sigma_{XY} = E\{XY\} - \underbrace{E\{X\}E\{Y\}}_{0} = \frac{1}{2}$$

3.

$$\hat{X}_{BLE}(Y) = \eta_X + \frac{\sigma_{XY}}{\sigma_Y^2}(Y - \eta_Y) = 0 + \frac{\sigma_{XY}}{\sigma_Y^2}(Y - 0) = \frac{\sigma_{XY}}{\sigma_Y^2} Y$$

In the previous section, we found that $\sigma_{XY} = 1/2$, and we also know that for $Y \sim U(-1,1)$ it holds $\sigma_Y^2 = 1/3$. From here:

$$\hat{X}_{BLE}(Y) = \frac{\sigma_{XY}}{\sigma_Y^2} Y = \frac{1/2}{1/3} Y = \frac{3}{2} Y$$

4. The estimation error of the linear estimator is given by:

$$E\left\{\left(X - \hat{X}_{BLE}\right)^2\right\} = \sigma_X^2 - \frac{\sigma_{XY}^2}{\sigma_Y^2}$$

All the values except σ_X^2 have already been calculated. For the calculation of σ_X^2, notice that from the given:

$$\frac{1}{2} = E\left\{\left(X - \hat{X}_{MMSE}\right)^2\right\} = \sigma_X^2 - \sigma_{\hat{X}_{MMSE}}^2$$

$$\sigma_{\hat{X}_{MMSE}}^2 = \text{var}(\text{sign}(Y)) = 1$$

$$\Rightarrow \sigma_X^2 - 1 = \frac{1}{2} \Rightarrow \sigma_X^2 = \frac{3}{2}$$

Let us substitute in the error formula and get:

$$E\left\{\left(X - \hat{X}_{BLE}\right)^2\right\} = \sigma_X^2 - \frac{\sigma_{XY}^2}{\sigma_Y^2} = \frac{3}{2} - \frac{(1/2)^2}{1/3} = \frac{3}{4}$$

Notice that:

$$E\left\{\left(X - \hat{X}_{MMSE}\right)^2\right\} \le E\left\{\left(X - \hat{X}_{BLE}\right)^2\right\} \le \text{var}(X)$$

Problem 5:

1. N_1, N_2 are jointly normal, X is normal and independent of them, and, thus, X, N_1, N_2 are jointly normal.

   X, N_1, N_2, Y_1, Y_2 result from a linear transformation on X, N_1, N_2, and, therefore, they are, too, jointly normal.

2. Since X, Y_1 are jointly Gaussian, the conditional distribution is also normal:

$$E[Y_1 \mid X] = E[X \mid X] + E[N_1 \mid X] = X$$
$$Var(Y_1 \mid X) = Var(N_1 \mid X) = Var(N_1) = \sigma_N{}^2$$
$$\Rightarrow Y_1 \mid X \sim N\left(X, \sigma_N{}^2\right)$$

3.

$$\hat{X}_{BLE} = \frac{\mathrm{Cov}(X,Y_1)}{\mathrm{Var}(Y_1)} Y_1 = \frac{\mathrm{Var}(X)}{\mathrm{Var}(X)+\mathrm{Var}(N_1)} Y_1 = \frac{\sigma_X{}^2}{\sigma_X{}^2+\sigma_N{}^2} Y_1$$

4. The same answer as in section 3, since X, Y_1 are jointly normal.

5.

$$\mathrm{MSE} = \mathrm{Var}(X) - \mathrm{Var}(\hat{X}_{BLE}) = \sigma_X{}^2 - \left(\frac{\sigma_X{}^2}{\sigma_X{}^2+\sigma_N{}^2}\right)^2 \mathrm{Var}(Y_1)$$
$$= \sigma_X{}^2 - \left(\frac{\sigma_X{}^2}{\sigma_X{}^2+\sigma_N{}^2}\right)^2 \left(\sigma_X{}^2+\sigma_N{}^2\right) = \frac{\sigma_X{}^2\sigma_N{}^2}{\sigma_X{}^2+\sigma_N{}^2}$$

6. <u>Option 1</u>: for symmetry reasons, the estimator will be a function of $Y_1 + Y_2$. Since $X, Y_1 + Y_2$ are jointly Gaussian:

$$\hat{X}_{BLE} = \frac{\mathrm{Cov}(X, Y_1+Y_2)}{\mathrm{Var}(Y_1+Y_2)}(Y_1+Y_2) = \frac{\mathrm{Cov}(X, 2X+N_1+N_2)}{\mathrm{Var}(2X+N_1+N_2)}(Y_1+Y_2)$$
$$= \frac{2\mathrm{Var}(X)}{4\mathrm{Var}(X)+\mathrm{Var}(N_1)+\mathrm{Var}(N_2)+2\mathrm{Cov}(N_1,N_2)}(Y_1+Y_2)$$
$$= \frac{\sigma_{\mathrm{X}}{}^2}{2\sigma_{\mathrm{X}}{}^2+(1+\rho)\sigma_{\mathrm{N}}{}^2}(Y_1+Y_2)$$

<u>Option 2</u>: Direct calculation

$$\hat{X}_{BLE} = C_{XY} C_Y{}^{-1} Y$$
$$= (\mathrm{Cov}(X, Y_1) \quad \mathrm{Cov}(X, Y_2)) \begin{pmatrix} \mathrm{Var}(Y_1) & \mathrm{Cov}(Y_1, Y_2) \\ \mathrm{Cov}(Y_1, Y_2) & \mathrm{Var}(Y_2) \end{pmatrix}^{-1} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$
$$= \frac{\sigma_X{}^2}{2\sigma_X{}^2 + (1+\rho)\sigma_N{}^2}(Y_1 + Y_2)$$

7.

$$\mathrm{MSE} = \mathrm{Var}(X) - \mathrm{Var}(\hat{X}_{BLE}) = \sigma_X{}^2 - \left(\frac{\sigma_X{}^2}{2\sigma_X{}^2 + (1+\rho)\sigma_N{}^2}\right)^2 \mathrm{Var}(Y_1 + Y_2)$$
$$= \frac{(1+\rho)\sigma_N{}^2\sigma_X{}^2}{2\sigma_X{}^2 + (1+\rho)\sigma_N{}^2}$$

8. The minimum is 0, and it is obtained for $\rho = -1$. In this case, the sum of the two noises is zero, and, thus, X is equal in a deterministic way to the average of the two measurements – namely, it can be reproduced with zero error.

9. For $\sigma_X^2 << \sigma_N^2$: MSE $\approx \sigma_X{}^2$. For $\sigma_X^2 >> \sigma_N^2$: MSE $\approx \left(\frac{1+\rho}{2}\right)\sigma_N{}^2$