HOMEWORK 2

(1) Suppose that $\boldsymbol{z} = \boldsymbol{s} + \boldsymbol{v}$, where $\boldsymbol{s}$ and $\boldsymbol{v}$ are independent, jointly distributed RVs with $\boldsymbol{s} \sim \mathcal{N}(\eta, \sigma^2)$ and $\boldsymbol{v} \sim \mathcal{N}(0, V^2)$.
   (a) Derive an expression for $E[\boldsymbol{s}|\boldsymbol{z} = z]$.
   (b) Derive an expression for $E[\boldsymbol{s}^2|\boldsymbol{z} = z]$.
   - Solution: as

$$f_{\mathbf{s}}(s|\mathbf{z} = z) = \frac{1}{\sqrt{2\pi}\sqrt{\frac{\sigma^2 V^2}{\sigma^2+V^2}}} \exp\left[-\frac{(s - \eta - \frac{\sigma^2}{V^2+\sigma^2}(z-\eta))^2}{2\frac{\sigma^2 V^2}{\sigma^2+V^2}}\right],$$

   we have

$$E[\boldsymbol{s}|\boldsymbol{z} = z] = \eta + \frac{\sigma^2}{V^2+\sigma^2}(z - \eta)$$

$$\begin{aligned} E[\boldsymbol{s}^2|\boldsymbol{z} = z] &= \int_{-\infty}^{\infty} s^2 f_{\mathbf{s}}(s|\mathbf{z} = z) \\ &= D[\boldsymbol{s}|\boldsymbol{z} = z] + E[\boldsymbol{s}|\boldsymbol{z} = z]^2 \\ &= \frac{\sigma^2 V^2}{\sigma^2 + V^2} + \left[\eta + \frac{\sigma^2}{V^2+\sigma^2}(z-\eta)\right]^2 \end{aligned}$$

(2) Suppose that $\boldsymbol{z} = \boldsymbol{s} + \boldsymbol{v}$, where $\boldsymbol{s}$ and $\boldsymbol{v}$ are independent, jointly distributed RVs with $\boldsymbol{s} \sim \mathcal{N}(\eta_s, \sigma_s^2)$ and $\boldsymbol{v} \sim \mathcal{N}(0, \sigma_v^2)$. Assume we have measurements $\boldsymbol{z}(1), \ldots, \boldsymbol{z}(n)$,
   (a) Derive the maximum likelihood estimate for $\boldsymbol{s}$;
   (b) Derive the maximum *a posteriori* estimate for $\boldsymbol{s}$;
   (c) Derive the minimum mean square estimate for $\boldsymbol{s}$;
   (d) Derive the linear minimum mean square estimate for $\boldsymbol{s}$;

Solution:
(a) See Example 4.1 in handouts
(b) Similar to (a), add the prior distribution of s
(c) We first demonstrate that $\boldsymbol{s}, \boldsymbol{z}(1), \ldots, \boldsymbol{z}(n)$ are jointly Gaussian, which is true as the linear combination of $\boldsymbol{x}, \boldsymbol{z}(1), \ldots, \boldsymbol{z}(n)$ are Gaussian, i.e.,

$$Y = a_0\boldsymbol{s} + a_1\boldsymbol{z}(1) + \ldots + a_n\boldsymbol{z}(n) = \left(\sum_{i=0}^{n} a_i\right) s + \sum_{i=1}^{n} a_i v(i)$$

is Gaussian with mean $\sum_{i=0}^{n} a_i\eta_s$ and variance $(\sum_{i=0}^{n} a_i)^2\sigma_s^2 + \sum_{i=1}^{n} a_i^2\sigma_v^2$. Similarly, $\boldsymbol{z}(1), \ldots, \boldsymbol{z}(n)$ are also jointly Gaussian.

Assume $\boldsymbol{z} = [z(1), \ldots, z(n)]^T$, and as $\boldsymbol{s}$ and $\boldsymbol{z}$ are jointly Gaussian, we have

$$(s, \boldsymbol{z}) \sim \mathcal{N}\left(\left[\begin{array}{c} \mu_s \\ \mu_z \end{array} \right], \left[\begin{array}{cc} \Sigma_{ss} & \Sigma_{sz} \\ \Sigma_{zs} & \Sigma_{zz} \end{array} \right] \right)$$

According to Schur complement, we have

$$\left[\begin{array}{cc} \Sigma_{ss} & \Sigma_{sz} \\ \Sigma_{zs} & \Sigma_{zz} \end{array} \right] = \left[\begin{array}{cc} I & \Sigma_{sz}\Sigma_{zz}^{-1} \\ 0 & I \end{array} \right] \left[\begin{array}{cc} \Sigma_{ss} - \Sigma_{sz}\Sigma_{zz}^{-1}\Sigma_{zs} & 0 \\ 0 & \Sigma_{zz} \end{array} \right] \left[\begin{array}{cc} I & 0 \\ \Sigma_{zz}^{-1}\Sigma_{zs} & I \end{array} \right]$$

and the inversion gives,

$$\begin{aligned} &\left[\begin{array}{cc} \Sigma_{ss} & \Sigma_{sz} \\ \Sigma_{zs} & \Sigma_{zz} \end{array} \right]^{-1} \\ &= \left[\begin{array}{cc} I & 0 \\ -\Sigma_{zz}^{-1}\Sigma_{zs} & I \end{array} \right] \left[\begin{array}{cc} (\Sigma_{ss} - \Sigma_{sz}\Sigma_{zz}^{-1}\Sigma_{zs})^{-1} & 0 \\ 0 & \Sigma_{zz}^{-1} \end{array} \right] \left[\begin{array}{cc} I & -\Sigma_{sz}\Sigma_{zz}^{-1} \\ 0 & I \end{array} \right] \end{aligned}$$

the joint distribution $p(s, \boldsymbol{z})$ is

$$p(s, \boldsymbol{z}) = \frac{1}{\sqrt{(2\pi)^{n+1} \det \Sigma}} \exp\left(-\frac{1}{2}(X - \mu_X)^T \Sigma^{-1} (X - \mu_X) \right)$$

in which $X = [s, \boldsymbol{z}^T]^T$, $\Sigma = \left[\begin{array}{cc} \Sigma_{ss} & \Sigma_{sz} \\ \Sigma_{zs} & \Sigma_{zz} \end{array} \right]$, and the quadratic part is

$$\begin{aligned} &(X - \mu_X)^T \Sigma^{-1} (X - \mu_X) \\ &= [(s - \eta_s)^T, (\boldsymbol{z} - \mu_z)^T] \cdot \left[\begin{array}{cc} I & 0 \\ -\Sigma_{zz}^{-1}\Sigma_{zs} & I \end{array} \right] \left[\begin{array}{cc} (\Sigma_{ss} - \Sigma_{sz}\Sigma_{zz}^{-1}\Sigma_{zs})^{-1} & 0 \\ 0 & \Sigma_{zz}^{-1} \end{array} \right] \left[\begin{array}{cc} I & -\Sigma_{sz}\Sigma_{zz}^{-1} \\ 0 & I \end{array} \right] \cdot \left[\begin{array}{c} s - \eta_s \\ \boldsymbol{z} - \mu_z \end{array} \right] \\ &= \left[(s - \eta_s)^T - (\boldsymbol{z} - \mu_z)^T \Sigma_{zz}^{-1}\Sigma_{zs} \;\; \boldsymbol{z} - \mu_z \right] \cdot \left[\begin{array}{cc} (\Sigma_{ss} - \Sigma_{sz}\Sigma_{zz}^{-1}\Sigma_{zs})^{-1} & 0 \\ 0 & \Sigma_{zz}^{-1} \end{array} \right] \cdot \left[\begin{array}{c} s - \eta_s - \Sigma_{ss}\Sigma_{zz}^{-1}(\boldsymbol{z} - \mu_z) \\ \boldsymbol{z} - \mu_z \end{array} \right] \\ &= [(s - \eta_s) - \Sigma_{sz}\Sigma_{zz}^{-1}(\boldsymbol{z} - \mu_z)]^T (\Sigma_{ss} - \Sigma_{sz}\Sigma_{zz}^{-1}\Sigma_{zs})^{-1} [\cdots] + (\boldsymbol{z} - \mu_z)^T \Sigma_{zz}^{-1} (\boldsymbol{z} - \mu_z) \end{aligned}$$

the determinant

$$\det\left(\left[\begin{array}{cc} \Sigma_{ss} & \Sigma_{sz} \\ \Sigma_{zs} & \Sigma_{zz} \end{array} \right] \right) = \det(\Sigma_{zz}) \cdot \det(\Sigma_{ss} - \Sigma_{zs}\Sigma_{zz}^{-1}\Sigma_{sz})$$

As

$$p(s, \boldsymbol{z}) = p(s|\boldsymbol{z})p(\boldsymbol{z})$$

we then have

$$p(s|\boldsymbol{z}) = \mathcal{N}\left(\eta_s + \Sigma_{sz}\Sigma_{zz}^{-1}(\boldsymbol{z} - \mu_z), \Sigma_{ss} - \Sigma_{sz}\Sigma_{zz}^{-1}\Sigma_{zs} \right)$$

and

$$p(\boldsymbol{z}) = \mathcal{N}(\mu_z, \Sigma_{zz}).$$

Hence the MMSE estimate is

$$E(s|\boldsymbol{z}) = \eta_s + \Sigma_{sz}\Sigma_{zz}^{-1}(\boldsymbol{z} - \mu_z)$$

(d) The linear MMSE estimate can be expressed as follows:

$$\hat{s}_{\text{LMMSE}} = E[\boldsymbol{s}\boldsymbol{z}^T][E(\boldsymbol{z}\boldsymbol{z}^T)]^{-1}\boldsymbol{z},$$

in which

$$E[\boldsymbol{s}\boldsymbol{z}^T] = E[\boldsymbol{s}\boldsymbol{z}(1), \ldots, \boldsymbol{s}\boldsymbol{z}(n)] = \left[\begin{array}{ccc} \eta_s^2 + \sigma_s^2 & \cdots & \eta_s^2 + \sigma_s^2 \end{array} \right]$$

and

$$E[\boldsymbol{z}\boldsymbol{z}^T] = \begin{bmatrix} E[\boldsymbol{z}(1)^2] & \cdots & E[\boldsymbol{z}(1)\boldsymbol{z}(n)] \\ \vdots & \ddots & \vdots \\ E[\boldsymbol{z}(n)\boldsymbol{z}(1)] & \cdots & E[\boldsymbol{z}(n)^2] \end{bmatrix}$$

$$= \begin{bmatrix} \eta_s^2+\sigma_s^2+\sigma_v^2 & \eta_s^2+\sigma_s^2 & \cdots & \eta_s^2+\sigma_s^2 \\ \eta_s^2+\sigma_s^2 & \eta_s^2+\sigma_s^2+\sigma_v^2 & \cdots & \eta_s^2+\sigma_s^2 \\ \vdots & \vdots & \ddots & \vdots \\ \eta_s^2+\sigma_s^2 & \eta_s^2+\sigma_s^2 & \cdots & \eta_s^2+\sigma_s^2+\sigma_v^2 \end{bmatrix}$$

In order to calculate the inversion of $E[\boldsymbol{z}\boldsymbol{z}^T]$, we represent it as

$$E[\boldsymbol{z}\boldsymbol{z}^T] = \begin{bmatrix} \sigma_v^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_v^2 \end{bmatrix} + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} (\sigma_s^2+\eta_s^2) \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}$$

According to the matrix inversion lemma, i.e.,

$$(A+BD^{-1}C)^{-1} = A^{-1} - A^{-1}B(D+CA^{-1}B)^{-1}CA^{-1}$$

we have

$$\{E[\boldsymbol{z}\boldsymbol{z}^T]\}^{-1} = \frac{1}{\sigma_v^2[\sigma_v^2+n(\sigma_s^2+\eta_s^2)]} \begin{bmatrix} \sigma_v^2+(n-1)(\sigma_s^2+\eta_s^2) & \cdots & -(\sigma_s^2+\eta_s^2) \\ \vdots & \ddots & \vdots \\ -(\sigma_s^2+\eta_s^2) & \cdots & \sigma_v^2+(n-1)(\sigma_s^2+\eta_s^2) \end{bmatrix}$$

Therefore, the LMMSE estimate is

$$\hat{\boldsymbol{s}}_{\text{LMMSE}} = \frac{\sigma_s^2+\eta_s^2}{\sigma_v^2+n(\sigma_s^2+\eta_s^2)} \sum_{i=1}^{n} z(i)$$