Exercise 2 b p. 171

$\underline{Y} = \underline{\underline{H}}\,\underline{\Theta} + \underline{V}$ $\qquad$ $\underline{V}$ is zero-mean RV with covariance matrix $\underline{\underline{\Sigma}}$

$$\Rightarrow \hat{\theta}(\underline{Y}) = \underline{\underline{\Sigma}}_{\Theta}\,\underline{\underline{H}}^T\left(\underline{\underline{H}}\,\underline{\underline{\Sigma}}_{\Theta}\,\underline{\underline{H}}^T + \underline{\underline{\Sigma}}_V\right)^{-1}\left(\underline{y} - \underline{\underline{H}}\,\underline{m}_{\Theta}\right) + \underline{m}_{\Theta}$$

Apparently, we can write for a given Θ

$$Y = \frac{1}{2}\,\Theta + V$$

uniformly distributed

$$\Rightarrow \Sigma_{\Theta} = \mathrm{Var}\,\Theta = E\left\{\left(\Theta - \cancelto{0}{E\Theta}\right)^2\right\} = \frac{1}{2}\left\{1 + 1\right\} = 1$$

$$H = \frac{1}{2}\;;\quad \Sigma_V = \mathrm{Var}\,V = \frac{1}{12}$$

$$\Rightarrow \boxed{\hat{\Theta}_{LMMSE}(y) = \frac{1}{2}\left(\frac{1}{4} + \frac{1}{12}\right)^{-1} y = \frac{3}{2}\,y}$$

MSE: $\quad E\left[\left(\hat{\Theta}_{LMMSE}(Y) - \Theta\right)^2\right]$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} \left[\frac{3}{2}y - \theta\right]^2 \mathrm{rect}\left(y - \frac{\theta}{2}\right)\frac{1}{2}\left[\delta(\theta-1) + \delta(\theta+1)\right] d\theta\, dy$$

$$= \frac{1}{2}\left[\int_{\mathbb{R}}\left(\frac{3}{2}y - 1\right)^2 \mathrm{rect}\left(y-\frac{1}{2}\right)dy + \int_{\mathbb{R}}\left(\frac{3}{2}+1\right)^2 \mathrm{rect}\left(y+\frac{1}{2}\right)dy\right]$$

$$= \frac{1}{2}\left[\int_0^1\left(\frac{3}{2}y-1\right)^2 dy + \int_{-1}^{0}\left(\frac{3}{2}+1\right)^2 dy\right]$$

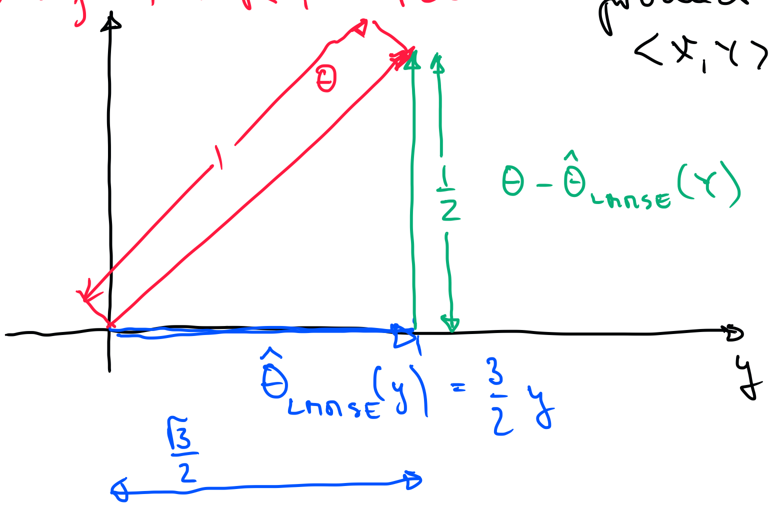
$$= \underline{\underline{\frac{1}{4}}}$$

Exercise 2 (c) p. 171

$$\int_{\mathbb{R}} \theta^2 \frac{1}{2}\{\delta(\theta-1)+\delta(\theta+1)\}d\theta = \|\theta\| = \sqrt{\langle\theta,\theta\rangle} = \sqrt{E\theta^2}$$

$$= 1$$

Consider inner product $\langle X, Y\rangle = E\,XY$.



$$\|\hat{\theta}_{LMMSE}\| = \sqrt{E\left[\frac{3}{2}Y\right]^2}$$

$$\|\hat{\theta}_{LMMSE}(Y)\|^2 = \frac{9}{4}EY^2 = \frac{9}{4}E\left[\frac{\theta}{2}+V\right]^2$$

$$= \frac{9}{4}\left[\frac{1}{4}\underbrace{E\theta^2}_{1} + \underbrace{EV^2}_{1/12}\right]$$

$$= \frac{9}{\cancel{4}}\left[\frac{\cancel{4}}{12}\right] = \frac{3}{4}$$

$$\Rightarrow \|\hat{\Theta}_{LMMSE}(Y)\| = \frac{\sqrt{3}}{2}$$

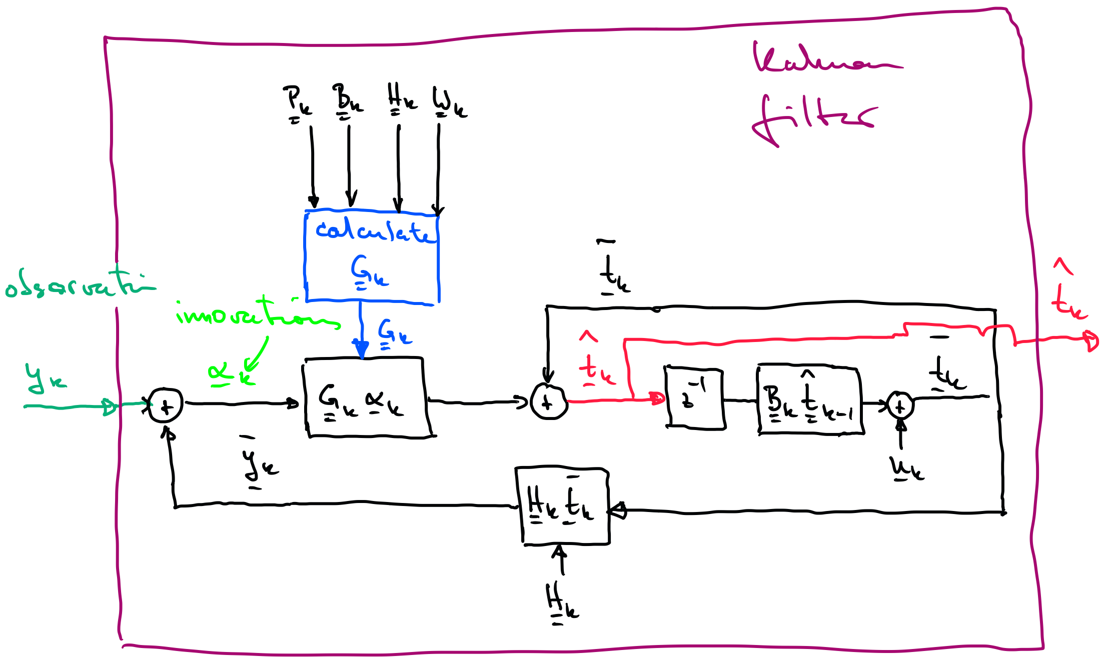
$$E\left[(\Theta - \hat{\Theta}_{LMMSE}(Y))^2\right] = \|\Theta - \hat{\Theta}_{LMMSE}(Y)\|^2$$

$$= \underbrace{E\Theta^2}_{} - 2\underbrace{E\left[\Theta\,\hat{\Theta}_{LMMSE}(Y)\right]}_{} + \underbrace{E\left[\hat{\Theta}_{LMMSE}(Y)\right]^2}_{}$$

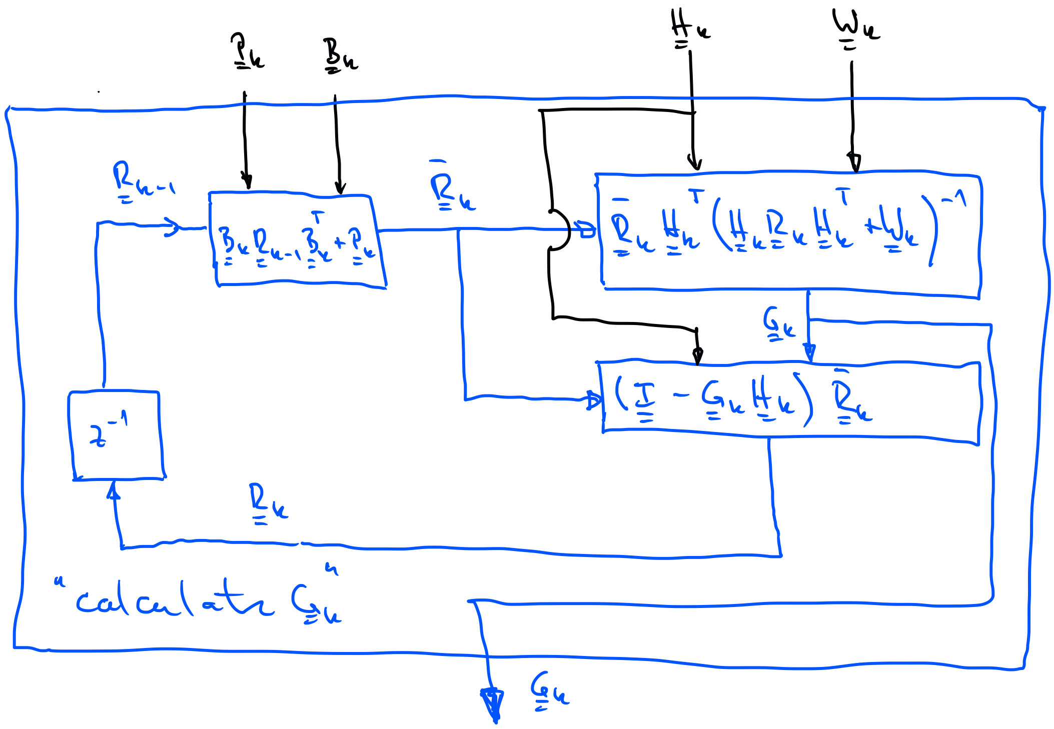
$$= 1 - 2 \cdot \frac{3}{4} \qquad\qquad \frac{3}{4}$$

$$= \frac{1}{4} \qquad \Rightarrow \|\Theta - \hat{\Theta}_{LMMSE}(Y)\| = \frac{1}{2}$$

Before we consider the problem on p. 172, some more comments on the structure of the Kalman filter (in terms of the signal flow graph) should be provided.



Block "calculate $\underline{G}_k$"



Exercise 3 on p.172

$\underline{\alpha}_k = \underline{y}_k - \bar{\underline{y}}_k$ is to be shown as being an uncorrelated process containing the "new information" for estimating $\underline{t}_k$ in the sense that, as opposed to $\underline{y}_k$, $\underline{\alpha}_k$ is uncorrelated in time so that only "new" information is provided for estimating $\underline{t}_k$ by $\hat{\underline{t}}_k$. Note that we slightly correct the assumptions in the problem to be valid for $k = 2, 3, \ldots$

(a) Note that $E[\underline{\alpha}_k \bar{\underline{y}}_n^T]$ is a special case of the more general affine mapping expression

$$E\left[\underline{\alpha}_k \underline{f}_a^T(\underline{y}_1,\dots,\underline{y}_{k-1})\right]$$

$$= E\left[(\underline{y}_k - \bar{\underline{y}}_k)\underline{f}_a^T(\cdot)\right] = E\left[(\underline{\underline{H}}_k \underline{t}_k + \underline{v}_k - \underline{\underline{H}}_k \bar{\underline{t}}_k)\underline{f}_a^T(\cdot)\right]$$

$$= \underline{\underline{H}}_k E\left[(\underline{t}_k - \bar{\underline{t}}_k)\underline{f}_a^T(\cdot)\right] + \underbrace{E\left[\underline{v}_k \underline{f}_a^T(\cdot)\right]}_{\underline{\underline{0}}}$$

$$= \underline{\underline{H}}_k \underline{\underline{B}}_k E\left[(\underline{t}_{k-1} - \hat{\underline{t}}_{k-1})\underline{f}_a^T(\cdot)\right]$$

$$= -\underline{\underline{H}}_k \underline{\underline{B}}_k \underbrace{E\left[(\hat{\underline{t}}_{k-1} - \underline{t}_{k-1})\underline{f}_a^T(\cdot)\right]}_{=\underline{\underline{0}}\ \text{(cf. p.160 ; orthogonality principle)}}$$

$\square$

(b) $E\left[\underline{\alpha}_n \underline{\alpha}_n^T\right] = E\left[\underline{\alpha}_n(\underline{y}_n - \bar{\underline{y}}_n)^T\right]$

$$= \underbrace{E\left[\underline{\alpha}_n \underline{y}_n^T\right]}_{0\ \text{(cf. 3 (a))}} - E\left[\underline{\alpha}_n \bar{\underline{y}}_n^T\right]$$

$$= -E\left[\underline{\alpha}_n \bar{\underline{y}}_n^T\right]$$

$$= -E\left[\underline{\alpha}_n \bar{\underline{t}}_n^T\right]\underline{\underline{H}}_n^T \qquad (*)$$

Consider the blue expression:

$$E\left[\underline{\alpha}_n \bar{\underline{t}}_n^T\right] = E\left[\underline{\alpha}_n(\underline{\underline{B}}_n \hat{\underline{t}}_{n-1} + \underline{u}_n)^T\right]$$

$$= E\left[\underline{\alpha}_n \hat{\underline{t}}_{n-1}^T\right]\underline{\underline{B}}_n^T + \underbrace{E\left[\underline{\alpha}_n\right]}_{=\underline{0}}\underline{u}_n^T$$

since $\hat{\underline{t}}_{n-1} = \bar{\underline{t}}_{n-1} + \underline{\underline{G}}_n(\underline{y}_{n-1} - \bar{\underline{y}}_{n-1})$

$$= \bar{\underline{t}}_{n-1} + \underline{\underline{G}}_k \underline{\alpha}_{n-1}$$

$$\Rightarrow E\left[\underline{\alpha}_k \bar{\underline{t}}_n^T\right] = E\left[\underline{\alpha}_k \bar{\underline{t}}_{n-1}^T\right]\underline{\underline{B}}_n^T + \underbrace{E\left[\underline{\alpha}_k \underline{\alpha}_{n-1}^T\right]}_{\text{the expression under consideration for } n \text{ replaced by } n-1\text{; so with (*)}} \underline{\underline{G}}_k^T \underline{\underline{B}}_n^T$$

$$= E\left[\underline{\alpha}_k \underline{t}_{n-1}^T\right]\underline{\underline{B}}_n^T + \left[-E\left[\underline{\alpha}_k \bar{\underline{t}}_{n-1}^T\right]\underline{\underline{H}}_{n-1}\right]\underline{\underline{G}}_k^T \underline{\underline{B}}_n^T$$

$$= E\left[\underline{\alpha}_k \hat{\underline{t}}_{n-1}^T\right]\left(\underline{\underline{B}}_n^T - \underline{\underline{H}}_{n-1}\underline{\underline{G}}_k^T \underline{\underline{B}}_n^T\right) \quad \left(\text{recursion in } \underline{\underline{z}}_n = E\left[\underline{\alpha}_k \hat{\underline{t}}_n^T\right]\right)$$

$$= E\left[\underline{\alpha}_k \bar{\underline{t}}_{n-2}^T\right]\left(\underline{\underline{B}}_{n-1}^T - \underline{\underline{H}}_{n-2}\underline{\underline{G}}_k \underline{\underline{B}}_{n-1}^T\right)\left(\underline{\underline{B}}_n^T - \underline{\underline{H}}_{n-1}\underline{\underline{G}}_k^T \underline{\underline{B}}_n^T\right)$$

$= \ldots$ (recursively count down the index of $\underline{\underline{z}}_n$ until $\underline{\underline{z}}_1$)

For $n=1$, $\bar{\underline{t}}_1 = \underline{u}_1$ is deterministic and we can directly calculate the right-hand side of the above equation to be

$$E\left[\underline{\alpha}_k \bar{\underline{t}}_1^T\right] = E\left[\underline{\alpha}_k\right]\bar{\underline{t}}_1^T = E\left[\underline{\alpha}_k\right]\underline{u}_1^T = \underline{0}$$

For $n>1$, we count down until we reach on the right-hand side $E\left[\underline{\alpha}_k \bar{\underline{t}}_1^T\right] = \underline{0}$

Thus, $E\left[\underline{\alpha}_k \underline{\alpha}_n^T\right] = \underline{0}$ $\square$.

Exercise 1 on p. 206

There are different way to construct an MVUE.

① The most straightforward way is to exploit the so-called Rao-Blackwell theorem. It is based on the concept of a "sufficient statistic" for estimating a parameter θ with minimum variance. The theorem essentially says that, if an unbiased estimator indeed exists, one should condition this estimator on a so-called "complete sufficient statistic". This step is sometimes called "Rao-Blackwellization".

② Another way (used here) is a bit more pragmatic, namely to use a kind of a trial-and-error approach: just use a maximum-likelihood estimator and check

Ⓐ whether it is unbiased and, if so,

Ⓑ " " reaches the Cramér-Rao lower bound.

$\Rightarrow$ Therefore, we use exercises (b) and (c).

(b) The density of the observed vector $\underline{Y} \in \{0,1\}^n$ with $\theta = \Pr\{Y_k = 1\}$ is given

$$p(\underline{y} \mid \theta) = \prod_{k=1}^{n} \theta^{y_k} (1-\theta)^{1-y_k}$$

$$= \theta^{\sum_{k=1}^{n} y_k} (1-\theta)^{n-\sum y_k} = (1-\theta)^n \left(\frac{\theta}{1-\theta}\right)^{\sum_{k=1}^{n} y_k}$$

From slide 178, the solution of the likelihood equation CAN be an MVUE.

Likelihood equation:

$$0 = \frac{\partial}{\partial \theta} \ln p(\underline{y} \mid \theta) \Big|_{\theta = \hat{\theta}_{ML}}$$

$$= \frac{\partial}{\partial \theta} \left[n \ln(1-\hat{\theta}_{ML}) + \left[\ln(\hat{\theta}_{ML}) - \ln(1-\hat{\theta}_{ML})\right] \sum_{k=1}^{n} y_k \right]$$

$$= -\frac{n}{1-\hat{\theta}_{ML}} + \left(\frac{1}{\hat{\theta}_{ML}} + \frac{1}{1-\hat{\theta}_{ML}} \right) \sum_k y_k$$

$$= \frac{1}{\hat{\theta}_{ML}} \sum_k y_k + \frac{1}{1-\hat{\theta}_{ML}} \left[\sum_k y_k - n \right]$$

$$\Rightarrow \frac{1}{\hat{\theta}_{ML}} \sum_k y_k = \frac{1}{\hat{\theta}_{ML}-1} \left[\sum_k y_k - n \right] \qquad \Big| \cdot (\hat{\theta}_{ML} - 1)$$

$$\Leftrightarrow \left(\cancel{1} - \frac{1}{\hat{\theta}_{ML}} \right) \sum y_k = \cancel{\sum y_k} - n$$

$$\Leftrightarrow \boxed{\hat{\theta}_{ML} = \frac{\sum_{k=1}^{n} y_k}{n}} \qquad \Rightarrow \boxed{\hat{\theta}_{ML}(\underline{Y}) = \frac{1}{n} \sum_{k=1}^{n} Y_k}$$

Since Y_k are i.i.d. Bernoulli-distributed with $E Y_k = 1 \cdot \Pr\{Y_k = 1\} + 0 \cdot \Pr\{Y_k = 0\}$

$$= 1 \cdot \theta = \theta$$

we obtain for the expected value of $\hat{\theta}_{ML}(\underline{Y})$

$$E\,\hat{\Theta}_{ML}(\underline{Y}) = \frac{1}{n}\sum_{k=1}^{n} E Y_k = \frac{1}{n}\sum_{k=1}^{n}\Theta = \frac{n\Theta}{n} = \Theta$$

$\Rightarrow \hat{\Theta}_{ML}(\underline{Y}) = \frac{1}{n}\sum_{k=1}^{n} Y_k$ is unbiased.

Its variance is given by

$$\mathrm{Var}\,\hat{\Theta}_{ML}(\underline{Y}) = \mathrm{Var}\left[\frac{1}{n}\sum_{k=1}^{n} Y_k\right]$$

$$= \frac{1}{n^2}\mathrm{Var}\sum_{k=1}^{n} Y_k \overset{\text{i.i.d.}}{=} \frac{1}{n^2}\sum_{k=1}^{n}\mathrm{Var}\,Y_k$$

$$= \frac{1}{n^2}\,n\,\mathrm{Var}\,Y_1 = \frac{\mathrm{Var}\,Y_1}{n} = \frac{E Y_1^2 - (E Y_1)^2}{n}$$

$$= \frac{1\cdot\Theta - \Theta^2}{n} = \frac{\Theta(1-\Theta)}{n}$$

To check whether this estimator is an MVUE, first find the CR lower bound in problem 1(c) and compare the aforementioned variance $\frac{\Theta(1-\Theta)}{n}$ with it.

(c) Fisher information of i.i.d. samples (cf. slide 182)

$$I_\Theta(\underline{y}) = E\left[\left(\frac{\partial \ln p_{\underline{Y}}(\underline{y}\,|\,\Theta)}{\partial\Theta}\right)^2\right]$$

$$= n\,E\left[\left(\frac{\partial \ln p_Y(Y_1\,|\,\Theta)}{\partial\Theta}\right)^2\right]$$

$$= n\, i_\Theta(y_1) \qquad \text{with}$$

$$i_\theta(y_1) = -E\left[\frac{\partial^2 \ln p_Y(y_1|\theta)}{\partial\theta^2}\right]$$

$$= -E\left[\frac{\partial^2}{\partial\theta^2}\left[\ln\left(\theta^{Y_1}(1-\theta)^{1-Y_1}\right)\right]\right]$$

$$= \frac{1}{\theta(1-\theta)}$$

Thus, we have $I_\theta(\underline{y}) = \dfrac{n}{\theta(1-\theta)}$ and the CR lower bound is given by

$\dfrac{1}{I_\theta(\underline{y})} = \dfrac{\theta(1-\theta)}{n}$, so that the ML estimator

$\hat{\theta}_{ML}(\underline{Y}) = \frac{1}{n}\sum_{k=1}^{n} Y_k$ is unbiased <u>and</u> reaches the CR lower bound $\Rightarrow$ it <u>is</u> an MVUE.

<u>Exercise 2 p. 206</u>

Results:

(a) $s^3 - s^2 y_1 y_2 + s(y_1^2 + y_2^2 - 1) - y_1 y_2 = 0$

(b) CR lower bound

$$\frac{1}{I_\theta} = \frac{(1-s^2)^2}{1+s^2}$$

<u>Exercise 3 p. 207</u>

(a) $\hat{\theta}_{ML} = \frac{1}{n}\sum_{k=1}^{n} y_k = \bar{y}$

$$E\,\hat{\theta}_{ML}(\underline{Y}) = \theta$$

$$\mathrm{Var}\,\hat{\theta}_{ML}(\underline{Y}) = \frac{\theta^2}{n}$$

(b) $I_\theta = n\, i_\theta = n \dfrac{1}{\theta^2}$