Solutions 6

Exercise 11.7

For Bayesian linear model, MMSE estimation is identical to MAP estimation since $p(\boldsymbol{\theta}|\mathbf{x})$ is Gaussian. But MAP estimation maximizes $p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})$ with no prior information, equivalent to maximizing $p(\mathbf{x}|\boldsymbol{\theta})$. In the Bayesian model, $p(\mathbf{x}|\boldsymbol{\theta}) = p(\mathbf{x};\boldsymbol{\theta})$. Thus, maximizing $p(\mathbf{x};\boldsymbol{\theta})$, which yields the MLE or MVUE, also yields the MMSE.

Exercise 11.11

$$\begin{aligned} R &= \mathbb{E}[C(\boldsymbol{\epsilon})] \\ &= \int\int C(\boldsymbol{\epsilon})p(\mathbf{x},\boldsymbol{\theta})d\mathbf{x}d\boldsymbol{\theta} \\ &= \int\left(\int C(\boldsymbol{\epsilon})p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}\right)p(\mathbf{x})d\mathbf{x} \end{aligned}$$

$$\int C(\boldsymbol{\epsilon})p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta} = \int_{||\boldsymbol{\epsilon}||>\delta} p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta} = 1 - \int_{||\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}||<\delta} p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}$$

as $\delta \to 0$, we minimize the above by choosing $\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\mathbf{x})$.

Exercise 11.12

If $\boldsymbol{\alpha} = \mathbf{A}\boldsymbol{\theta}$, then $\partial\boldsymbol{\alpha}/\partial\boldsymbol{\theta} = \mathbf{A}$

$$p(\mathbf{x},\boldsymbol{\alpha}) = \frac{p(\mathbf{x},\boldsymbol{\theta})}{|\det\frac{\partial\boldsymbol{\alpha}/}{\partial\boldsymbol{\theta}}|} = \frac{p(\mathbf{x},\boldsymbol{\theta})}{|\det\mathbf{A}|}$$

However, $\mathbf{A}$ does not depend on $\boldsymbol{\alpha}$ and $\boldsymbol{\theta} = \mathbf{A}^{-1}\boldsymbol{\alpha}$, so that

$$p(\mathbf{x},\boldsymbol{\alpha}) = \frac{p_{x,\theta}(\mathbf{x},\mathbf{A}^{-1}\boldsymbol{\alpha})}{|\det\mathbf{A}|}$$

The MAP estimator of $\boldsymbol{\alpha}$ maximizes $p_{x,\theta}(\mathbf{x},\mathbf{A}^{-1}\boldsymbol{\alpha})$, equivalent to maximizing $p(\mathbf{x},\boldsymbol{\theta})$ because $\boldsymbol{\theta} = \mathbf{A}^{-1}\boldsymbol{\alpha}$ is invertible. Thus, $\hat{\boldsymbol{\alpha}} = \mathbf{A}\hat{\boldsymbol{\theta}}$

Exercise 12.2

From (12.27) in page 391, we can get

$$\hat{A} = \mu_A + (\frac{1}{\sigma_A^2} + \frac{\mathbf{h}^T\mathbf{h}}{\sigma^2})^{-1}\frac{\mathbf{h}^T}{\sigma^2}$$

where $\mathbf{h} = [1, r, ..., r^{N-1}]^T$. Thus,

$$\hat{A} = \mu_A + \frac{\sum_{n=0}^{N-1} r^n (x[n] - r^n \mu_A)}{\frac{\sigma^2}{\sigma_A^2} + \sum_{n=0}^{N-1} r^{2n}}$$

From (12.29) and (12.30), we get

$$\text{Bmse}(\hat{A}) = \frac{1}{\frac{1}{\sigma_A^2} + \frac{1}{\sigma^2} \sum_{n=0}^{N-1} r^{2n}}$$

Exercise 12.14

To minimize $\mathbb{E}[(x[n] - \hat{x}[n])^2]$, we use the orthogonality principle, i.e.

$$\mathbb{E}[(x[n] - \hat{x}[n])x[n-l]] = 0, \qquad l = -M, ..., M (l \neq 0)$$

$$r_{xx}(l) = \mathbb{E}[\sum_k a_k x[n-k]x[n-l]] = \sum_k a_k r_{xx}(l-k)$$

To show that $a_{-k} = a_k$, we let $k^{'} = -k$

$$r_{xx}(l) = \sum_{k^{'}=-M, k^{'} \neq 0}^{M} a_{-k^{'}} r_{xx}(l + k^{'})$$

Let $l^{'} = -l$

$$r_{xx}(-l^{'}) = \sum_{k^{'}=-M, k^{'} \neq 0}^{M} a_{-k^{'}} r_{xx}(-l^{'} + k^{'})$$

$$r_{xx}(l^{'}) = \sum_{k^{'}=-M, k^{'} \neq 0}^{M} a_{-k^{'}} r_{xx}(l^{'} - k^{'})$$

Hence $r_{xx}(-k) = r_{xx}(k)$. But these are the same set of equation for which there is a unique solution. Hence $a_{-k} = a_k$. This must be true since the correlation of $x[n]$ with $x[n+k]$ is the same as that with $x[n-k]$, due to the even symmetry.

Exercise 12.19

$$\hat{x}[n] = \sum_{k=1}^{N} h(k)x[n-k]$$

$$\mathbb{E}[(x[n] - \hat{x}[n])x[n-l]] = 0$$

$$r_{xx}(l) = \sum_{k=1}^{N} h(k)\mathbb{E}(x[n-k]x[n-l]) = \sum_{k=1}^{N} h(k) r_{xx}(l-k)$$

The equations are independent of n since in deriving (12.65) we assumed $n = N$ was the index of the sample to be predicted. Hence the ACF does not depend on n

$$\begin{aligned} M_{\hat{x}} &= \mathbb{E}[(x[n]-\hat{x}[n])x[n]] - \mathbb{E}[(x[n]-\hat{x}[n])\hat{x}[n]] \\ &= \mathbb{E}[x^2[n]] - \sum_{k=1}^{N} h(k)\mathbb{E}[x[n-k]x[n]] \\ &= r_{xx}(0) - \sum_{k=1}^{N} h(k) r_{xx}(k) \end{aligned}$$

Exercise 12.20

From the previous problem, we have

$$r_{xx}(l) = \sum_{k=1}^{N} h(k) r_{xx}(l-k)$$

must be solved for the optimal one-step prediction. But for an AR(N) process, we know that

$$r_{xx}(l) = -\sum_{k=1}^{N} a[k] r_{xx}(l-k)$$

which are the Yale-Walker equations. Hence the solution for the $h(k)$ is unique,

$$h(k) = -a[k]$$

so that $\hat{x}[n] = -\sum_{k=1}^{N} a[k]x[n-k]$ and the MMSE is

$$\begin{aligned} M_{\hat{x}} &= r_{xx}(0) - \sum_{k=1}^{N} h(k) r_{xx}(k) \\ &= r_{xx}(0) + \sum_{k=1}^{N} a[k] r_{xx}(k) \\ &= \sigma_u^2 \end{aligned}$$