02471 Machine Learning for Signal Processing

Solution

Exercise 4: Adaptive Linear Filtering with LMS

4.1 Wiener Filter review

Exercise 4.1.1

Using the formulas (4.5)-(4.6) and (4.43) we get

$$\begin{bmatrix} r_u(0) & r_u(1) \\ r_u(1) & r_u(0) \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{bmatrix} r_{du}(0) \\ r_{du}(1) \end{bmatrix}$$

Exercise 4.1.2

We insert into the expression for $r_x(k)$

$$\begin{aligned} r_u(k) &= \mathbb{E}[\mathrm{u}_n \mathrm{u}_{n-k}] \\ &= \mathbb{E}[(\mathrm{s}_n + \mathrm{v}_n)(\mathrm{s}_{n-k} + \mathrm{v}_{n-k})] \\ &= \mathbb{E}[\mathrm{s}_n \mathrm{s}_{n-k}] + \mathbb{E}[\mathrm{v}_n \mathrm{v}_{n-k}] \\ &= r_s(k) + r_v(k) \end{aligned}$$

where we have used that s_n and v_n are uncorrelated and v_n is a zero-mean noise signal (the cross-terms vanish).

For the cross-correlation between the desired signal and input signal, we get (setting the desired signal to s_n)

$$\begin{aligned} r_{du}(k) &= \mathbb{E}[\mathrm{d}_n \mathrm{u}_{n-k}] \\ &= \mathbb{E}[\mathrm{s}_n(\mathrm{s}_{n-k} + \mathrm{v}_{n-k})] \\ &= \mathbb{E}[\mathrm{s}_n \mathrm{s}_{n-k}] + \mathbb{E}[\mathrm{s}_n \mathrm{v}_{n-k}] \\ &= r_s(k) \end{aligned}$$

Exercise 4.1.3

Since we are using a filter of length 2, we only need to compute two coefficients. They are computed using the formula $\boldsymbol{w} = \Sigma_u^{-1} \boldsymbol{p}$:

$$\begin{aligned} \boldsymbol{w} &= \begin{bmatrix} r_u(0) & r_u(1) \\ r_u(1) & r_u(0) \end{bmatrix}^{-1} \begin{bmatrix} r_{du}(0) \\ r_{du}(1) \end{bmatrix} \\ &= \begin{bmatrix} r_s(0) + r_v(0) & r_s(1) + r_v(1) \\ r_s(1) + r_v(1) & r_s(0) + r_v(0) \end{bmatrix}^{-1} \begin{bmatrix} r_s(0) \\ r_s(1) \end{bmatrix} \\ &= \begin{bmatrix} 1.2 & -0.4 \\ -0.4 & 1.2 \end{bmatrix}^{-1} \begin{bmatrix} 0.6 \\ -0.3 \end{bmatrix} \\ &= \begin{bmatrix} 0.469 \\ -0.094 \end{bmatrix} \end{aligned}$$

Exercise 4.1.4

Using formula (4.9) from the book we get the MMSE for the filter to

$$\begin{aligned}\text{MMSE}_2 &= \sigma_d^2 - \boldsymbol{p}^T \boldsymbol{w} \\ \text{MMSE}_2 &= 0.9 - \begin{bmatrix} 0.6 \\ -0.3 \end{bmatrix}^T \begin{bmatrix} 0.469 \\ -0.094 \end{bmatrix} \\ &= 0.591\end{aligned}$$

4.2 System identification using a Wiener filter

Exercise 4.2.1

The filter length is l which is directly read from the Model. The input sequence is u_n, the desired sequence is $d_n = y_n + \eta_n = H * u_n + \eta_n$. The error sequence is e_n, and the output from the model is $\hat{d} = \boldsymbol{w} * u_n$. The symbols are italic case in the figure, since the figure assumes realizations of the random processes.

Exercise 4.2.2

Using the definition of convolution (eq. 2.126), we get

$$\begin{aligned}\mathbb{E}[(\mathbf{u}_n * H)\mathrm{u}_{n-k}] &= \mathbb{E}\left[\mathrm{u}_{n-k} \sum_{i=0}^{l-1} u_{n-i} h_i\right] \\ &= \sum_{i=0}^{l-1} \mathbb{E}[\mathrm{u}_{n-k}\mathrm{u}_{n-i}] h_i\end{aligned}$$

Since u_n is a white noise sequence, we have $\mathbb{E}[\mathrm{u}_n \mathrm{u}_{n-k}] = \delta_k \sigma_u^2$, hence we have

$$\mathbb{E}[\mathrm{u}_{n-k}\mathrm{u}_{n-i}] = \begin{cases} \sigma_u^2 & k = i \\ 0 & k \neq i \end{cases}$$

Using this result we obtain

$$\mathbb{E}[(\mathbf{u}_n * H)\mathrm{u}_{n-k}] = \sigma_{\mathrm{u}}^2 h_k$$

Exercise 4.2.3

First we determine the input correlation matrix Σ_u. Since the input sequence is a white noise sequence, $r_u(k) = \sigma_u^2$ for $k = 0$ and $r_u(k) = 0$ for $k \neq 0$, therefore the input covariance matrix is:

$$\Sigma_u = \sigma_u^2 I$$

Next we identify the cross-correlation vector $\boldsymbol{p}$ whose elements are $p_k = r_{du}(k)$. We have for d_u

$$\begin{aligned}\mathrm{d}_n &= \mathrm{y}_n + \eta_n \\ &= H * \mathbf{u}_n + \eta_n\end{aligned}$$

For the cross-correlation we then get

$$\begin{aligned} r_{du}(k) &= \mathbb{E}[\mathrm{d}_n \mathrm{u}_{n-k}] \\ &= \mathbb{E}[(H * \mathbf{u}_n + \eta_n)\mathrm{u}_{n-k}] \\ &= \mathbb{E}[(H * \mathbf{u}_n)\mathrm{u}_{n-k}] + \mathbb{E}[\eta_n \mathrm{u}_{n-k}] \end{aligned}$$

Since η_n is uncorrelated with u_n, and $\mathbb{E}[\mathrm{u}_n] = 0$ we get:

$$r_{du}(k) = \mathbb{E}[(H * \mathbf{u}_n)\mathrm{u}_{n-k}]$$

Using the result from 4.2.2 we get:

$$\begin{aligned} r_{du}(k) &= h_k \sigma_u^2 \quad \Rightarrow \\ \boldsymbol{p} &= H\sigma_u^2 \end{aligned}$$

From the normal equation we finally get:

$$\begin{aligned} \boldsymbol{w}_* &= \Sigma_u^{-1}\boldsymbol{p} \\ &= (\sigma_u^2 I)^{-1} H \sigma_u^2 \\ &= H \end{aligned}$$

Exercise 4.2.4

First we determine an expression for the error sequence

$$\begin{aligned} \mathrm{e}_n &= \mathrm{d}_n - \hat{\mathrm{d}}_n \\ &= H * \mathbf{u}_n + \eta_n - \boldsymbol{w} * \mathbf{u}_n \\ &= (H - \boldsymbol{w}) * \mathbf{u}_n + \eta_n \\ &= \boldsymbol{g}^T \mathbf{u}_n + \eta_n \end{aligned}$$

Where we have used that the convolution is distributive, ie. $f * (g + h) = f * g + f * h$, and defined the vector $\boldsymbol{g} = H - \boldsymbol{w}$. We now get the mean squared error to

$$\begin{aligned} \mathbb{E}\left[\mathrm{e}_n^2\right] &= \mathbb{E}\left[(\boldsymbol{g}^T \mathbf{u}_n + \eta_n)^2\right] \\ &= \mathbb{E}\left[\boldsymbol{g}^T \mathbf{u}_n \mathbf{u}_n^T \boldsymbol{g} + 2\boldsymbol{g}^T \mathbf{u}_n \eta_n + \eta_n^2\right] \\ &= \boldsymbol{g}^T \mathbb{E}\left[\mathbf{u}_n \mathbf{u}_n^T\right] \boldsymbol{g} + 2\boldsymbol{g}^T \mathbb{E}[\mathbf{u}_n \eta_n] + \mathbb{E}\left[\eta_n^2\right] \end{aligned}$$

where we have used that $\boldsymbol{g}$ is a deterministic vector.

From exercise 4.2.2 we found $\mathbb{E}[\mathrm{u}_n \mathrm{u}_{n-k}] = \delta_k \sigma_u^2$, hence $\mathbb{E}\left[\mathbf{u}_n \mathbf{u}_n^T\right] = \sigma_u^2 I$.

Additionally, since $\mathbf{u}_n$ and η_n are two uncorrelated white noise sequences we have $\mathbb{E}[\mathbf{u}_n \eta_n] = \mathbb{E}[\mathbf{u}_n]\mathbb{E}[\eta_n] = 0$ (since $\mathbb{E}[\mathbf{u}_n] = \mathbb{E}[\eta_n] = 0$).

Using these results, our expression reduce to

$$\mathbb{E}\left[\mathrm{e}_n^2\right] = \boldsymbol{g}^T \sigma_u^2 I \boldsymbol{g} + \sigma_\eta^2$$

And from linear algebra we know that: $\boldsymbol{g}^T \boldsymbol{g} = \|\boldsymbol{g}\|_2^2$, hence:

$$\mathbb{E}\left[e_n^2\right] = \sigma_u^2 \|H - \boldsymbol{w}\|_2^2 + \sigma_\eta^2$$

So the lowest MSE possible is σ_η^2, and σ_u^2 is adding to the MSE proportionally to the filter differences.