TD2: Optimal and Adaptive Filtering, Equalization Solutions

1 Wiener Filtering

Problem 1. A Wiener filtering problem

In this problem we want to estimate the signal x_k from a measurement y_k using a filter $H(z)$ such that the estimate $\hat{x}_k$ minimizes $\mathrm{E}(x_k - \hat{x}_k)^2$. This is clearly a Wiener filter problem, but it can also be solved directly by realizing that $X(z) = (G(z) + z^{-d})Y(z)$. This means that by choosing $H(z) = G(z) + z^{-d}$ the estimation error is zero. In terms of a Wiener filter, the optimal $H(z)$ is given by $S_{yy}^{-1}(z)S_{xy}(z) = S_{yy}^{-1}(z)(G(z) + z^{-d})S_{yy}(z) = G(z) + z^{-d}$.

Problem 2. Constrained and Unconstrained Two-Channel Wiener Filtering

(a) Using the two measurements, y_{1k} and y_{2k} of the signal x_k, we want to choose two filters $H_1(z)$ and $H_2(z)$ satisfying

$$H_1(z) + H_2(z) = 1$$

such that the sum of their outputs is a MMSE estimate of x_k. The noise components are assumed independent, to have zero-mean and have power spectra $S_{v_1v_1}(z)$ and $S_{v_2v_2}(z)$. Before finding the exact solution, we can find the form of the solution by inspection. Assume for a particular frequency $S_{v_1v_1}(z) \gg S_{v_2v_2}(z)$, so that the measurement y_{2k} is more reliable at that frequency. In this case $|H_1(z)| \approx 1$ and $|H_2(z)| \approx 0$. The same must hold for the opposite case.

Writing the output in the time–domain

$$\begin{aligned} \hat{x}_k &= x_k * (h_{1k} + h_{2k}) + v_{1k} * h_1(k) + v_{2k} * h_2(k) \\ &= x_k + (v_{1k} - v_{2k}) * h_{1k} + v_{2k} \end{aligned} \tag{1}$$

we have the simple Wiener filter problem: find the filter $H_1(z)$ which when given an input $v_{1k} - v_{2k}$ yields a MMSE estimate of $-v_{2k}$. The optimal $H_1(z)$ is then given by

$$H_1(z) = S_{v_1-v_2,v_1-v_2}^{-1}(z)\; S_{-v_2,v_1-v_2}(z) = \frac{S_{v_2,v_2}(z)}{S_{v_1,v_1}(z) + S_{v_2,v_2}(z)}$$

From the constraint equation, the other filter is given by

$$H_2(z) = \frac{S_{v_1,v_1}(z)}{S_{v_1,v_1}(z) + S_{v_2,v_2}(z)}$$

We see therefore that these filters have the form that we indicated from the outset.

2 Equalization and Wiener Filtering

Problem 3. Equalization of a First-Order FIR Channel

Here we want to equalize a channel with response $C(z) = 1 - az^{-1}$. The zero forcing equalizer is simply $H_{\mathrm{ZF-LE}}(z) = 1/C(z) = z/(z-a)$. The associated MSE is given by

$$\mathrm{MSE}_{\mathrm{ZF-LE}} = \frac{\sigma_v^2}{2\pi j}\oint_{|z|=1}\frac{dz}{zC^\dagger(z)C(z)} \tag{2}$$

$$= -\frac{\sigma_v^2}{2\pi j a^*}\oint_{|z|=1}\frac{dz}{(z-a)(z-1/a^*)} \tag{3}$$

Assuming $|a| < 1$ this is simply $-\sigma_v^2/a^*$ times the residue at $z = a$ which is $1/(a - 1/a^*)$ so that

$$\mathrm{MSE}_{\mathrm{ZF-LE}} = \frac{\sigma_v^2}{1-|a|^2}$$

We see that as $|a| \to 1$, the MSE tends to infinity. This is because the channel response has a zero close to $|z| = 1$ so that the equalizer has a pole close to $|z| = 1$. This has the effect of amplifying the noise at the corresponding frequency by a large amount.

We now consider the MMSE equalizer which is simply the Wiener filter

$$H_{\mathrm{MMSE-LE}}(z) = \frac{C^\dagger(z)}{C(z)C^\dagger(z)+1/\gamma} = \frac{1-a^*z}{(1-a^*z)(1-a/z)+1/\gamma} = \frac{z(z-1/a^*)}{z^2-(a+(1+\gamma)/a^*)z+a/a^*} \tag{4}$$

where $\gamma = \sigma_x^2/\sigma_v^2$ is the SNR (signal–to–noise ratio). The MSE for this case is given by

$$\mathrm{MSE}_{\mathrm{MMSE-LE}} = \frac{\sigma_v^2}{2\pi j}\oint_{|z|=1}\frac{dz}{z(C(z)C^\dagger(z)+1/\gamma)} = -\frac{\sigma_v^2}{2\pi j a^*}\oint_{|z|=1}\frac{dz}{z^2-(a+(1+1/\gamma)/a^*)z+a/a^*} \tag{5}$$

The poles are given by $p_{1,2} = \frac{1}{2a^*}\left(|a|^2+1+1/\gamma \pm \sqrt{(|a|^2+1+1/\gamma)^2-4|a|^2}\right)$ so that

$$\mathrm{MSE}_{\mathrm{MMSE-LE}} = \frac{\sigma_v^2}{\sqrt{\left(|a|^2+1+1/\gamma\right)^2-4|a|^2}}$$

We see that as the SNR tends to infinity, $\mathrm{MSE}_{\mathrm{MMSE-LE}} = \mathrm{MSE}_{\mathrm{ZF-LE}}$ which is to be expected.

The UMMSE equalizer is simply the MMSE equalizer scaled by the factor

$$L = \frac{1}{\sigma_x^2}\underbrace{\left(\frac{1}{2\pi j}\oint_{|z|=1}\frac{dz}{z}C^\dagger(z)S_{yy}^{-1}(z)C(z)\right)^{-1}}_{K}$$

which can be expressed in terms of $\mathrm{MSE}_{\mathrm{MMSE-LE}}$ as

$$L = \frac{1}{1-\mathrm{MSE}_{\mathrm{MMSE-LE}}/\sigma_x^2}$$

so that $H_{\mathrm{UMMSE-LE}} = LH_{\mathrm{MMSE-LE}}$. The MSE is given by

$$MSE_{\mathrm{UMMSE-LE}} = K - \sigma_x^2 = \sigma_x^2(L-1) = \frac{MSE_{\mathrm{MMSE-LE}}}{1-\frac{MSE_{\mathrm{MMSE-LE}}}{\sigma_x^2}}$$

Problem 4. Wiener Filtering and Zero-Forcing Linear Equalization of a Second-Order FIR Channel

(a) We get from the course notes

$$
\begin{aligned}
\text{MMSE} &= E\,\tilde{x}_k^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{S_{xx}(f)S_{vv}(f)}{S_{xx}(f)+S_{vv}(f)}df \\
&= \sigma_v^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{S_{xx}(f)}{S_{xx}(f)+S_{vv}(f)}df \;=\; \sigma_v^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} H(f)df \;=\; \sigma_v^2\, h_0\ .
\end{aligned}
$$

(b) We have $\mathrm{H}_{ZF}(z) = \dfrac{1}{\mathrm{C}(z)}$. As stated in the course notes, we have to factor $\mathrm{C}(z)$ into its minimum-phase and maximum-phase factors since we need to take the causal inverse for the minimum-phase factor and the anticausal inverse for the maximum-phase factor in order to have a stable inverse. Now,

$$
\mathrm{C}(z) = 1 - \frac{5}{2}z^{-1} + z^{-2} = (1-\frac{1}{2}z^{-1})(1-2z^{-1}) = -2z^{-1}(1-\frac{1}{2}z^{-1})(1-\frac{1}{2}z)\ .
$$

So we get

$$
\begin{aligned}
\mathrm{H}_{ZF}(z) &= \frac{1}{1-\frac{5}{2}z^{-1}+z^{-2}} = \frac{-\frac{1}{2}z}{(1-\frac{1}{2}z^{-1})(1-\frac{1}{2}z)} \\
&= \frac{A}{1-\frac{1}{2}z^{-1}} + \frac{Bz}{1-\frac{1}{2}z} = \frac{(A-\frac{1}{2}B)+(B-\frac{1}{2}A)z}{(1-\frac{1}{2}z^{-1})(1-\frac{1}{2}z)}
\end{aligned}
$$

from which we find

$$
\left\{ \begin{array}{l} A - \frac{1}{2}B = 0 \\ B - \frac{1}{2}A = -\frac{1}{2} \end{array} \right. \quad \Rightarrow \quad \left\{ \begin{array}{l} A = -\frac{1}{3} \\ B = -\frac{2}{3} \end{array} \right.
$$

Hence

$$
\mathrm{H}_{ZF}(z) = -\frac{1}{3}\frac{1}{1-\frac{1}{2}z^{-1}} - \frac{2}{3}\frac{z}{1-\frac{1}{2}z}
$$

from which we can find the impulse response

$$
h_k^{ZF} = \left\{ \begin{array}{ll} -\frac{1}{3}\,(\frac{1}{2})^k & ,\ k \geq 0 \\ -\frac{4}{3}\,(\frac{1}{2})^{-k} & ,\ k < 0 \end{array} \right.
$$

(c) As the term Zero-Forcing tells us, there is no ISI at the equalizer output, thus we have the following scheme :

```
           ________       v_k       ________
 x_k      |        |       |+      |        |      x_k+w_k
 _______| C(z)   |-----O------|  H(z) |_____
          |        |               |        |
          |________|               |_______|
```

where $w_k = \mathrm{H}(q)v_k$, hence $\widetilde{x}_k = w_k$ and MSE$=Ew_k^2$. From $S_{ww}(f) = |H(f)|^2 S_{vv}(f) = \sigma_v^2 |H(f)|^2$, we get

$$
\begin{aligned}
MSE_{ZF} &= E\,w_k^2 = \int_{-1/2}^{1/2} S_{ww}(f)df = \sigma_v^2 \int_{-1/2}^{1/2} |H(f)|^2 df = \sigma_v^2 \sum_{k=-\infty}^{\infty} h_k^2 \\
&= \sigma_v^2 \left(\frac{1}{9} \frac{1}{1-\frac{1}{4}} + \frac{4}{9} \frac{1}{1-\frac{1}{4}} \right) = \sigma_v^2 \, \frac{5}{9} \frac{1}{1-\frac{1}{4}} = \sigma_v^2 \, \frac{20}{27} \,.
\end{aligned}
$$

(d) With the MSE, we find immediately the $SNR_{ZF} = \sigma_x^2 / MSE_{ZF} = \dfrac{27}{20} \dfrac{\sigma_x^2}{\sigma_v^2} = 1.35 \dfrac{\sigma_x^2}{\sigma_v^2}$.
For the MFB on the other hand,

$$
MFB = \frac{\sigma_x^2}{\sigma_v^2} \int_{-1/2}^{1/2} |C(f)|^2 df = \frac{\sigma_x^2}{\sigma_v^2} \sum_{k=-\infty}^{\infty} c_k^2 = \frac{33}{4} \frac{\sigma_x^2}{\sigma_v^2} = 8.25 \, \frac{\sigma_x^2}{\sigma_v^2} \,.
$$

So the MFB is $\dfrac{8.25}{1.35} = 6.11$ times better than the ZF-LE SNR.

Problem 5. FIR MMSE Linear Equalization of a FIR Channel

(a) $R_{YY} = C\,R_{SS}C^T + C\,R_{SV} + R_{VS}C^T + R_{VV} = \sigma_s^2 CC^T + \sigma_v^2 I_N$.
Note $x_k^{(d)} = \underline{e}_d^T S_k$, hence $R_{Yx^{(d)}} = R_{YS}\underline{e}_d = C\,R_{SS}\underline{e}_d = \sigma_s^2 C\underline{e}_d = \sigma_s^2 \underline{c}_d$.

(b)
$$
\begin{aligned}
\sigma^2_{\widetilde{x}^{(d)}_{MMSE}} &= R_{x^{(d)}x^{(d)}} - R_{x^{(d)}Y} R_{YY}^{-1}\, R_{Yx^{(d)}} = \sigma_s^2 - \sigma_s^2 \underline{c}_d^T (\sigma_s^2 CC^T + \sigma_v^2 I_N)^{-1} \underline{c}_d \sigma_s^2 \;. \\
&= \sigma_s^2 \, [1 - \underline{c}_d^T (CC^T + \tfrac{\sigma_v^2}{\sigma_s^2} I_N)^{-1} \underline{c}_d]
\end{aligned}
$$

(c) Since C^T has the structure of a Toeplitz pre- and post-windowed matrix, $C\,C^T$ is a banded Toeplitz matrix and so is $R_{YY} = \sigma_s^2 CC^T + \sigma_v^2 I_N$. Element $(1, i+1)$ of R_{YY} is $\sigma_v^2 \delta_{i0} + \sigma_s^2 \sum_{k=0}^{L-1} c_{k+i} c_k$ where the last term contains the correlation sequence of the channel response, which is zero beyond lag $L-1$.

(d) Since as remarked earlier, $\underline{c}_d = C\underline{e}_d$, we get in absence of noise
$\sigma^2_{\widetilde{x}^{(d)}_{MMSE}} = \sigma_s^2 \, [1 - \underline{c}_d^T (CC^T)^{-1} \underline{c}_d] = \sigma_s^2 \; \underline{e}_d^T [I - C^T (CC^T)^{-1} C] \underline{e}_d = \sigma_s^2 \; \underline{e}_d^T \; P_{C^T}^{\perp} \; \underline{e}_d$.

(e) A tall Vandermonde matrix has full column rank if all roots are different. Furthermore

$$
\begin{bmatrix}
c_0 & c_1 & \cdots & c_{L-1} & 0 & \cdots & 0 \\
0 & c_0 & \cdots & c_{L-2} & c_{L-1} & \ddots & 0 \\
\vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & c_0 & \cdots & \cdots & c_{L-1}
\end{bmatrix}
\begin{bmatrix}
1 & 1 & \cdots & 1 \\
z_1 & z_2 & \cdots & z_{L-1} \\
\vdots & \vdots & & \vdots \\
z_1^{N+L-2} & z_2^{N+L-2} & \cdots & z_{L-1}^{N+L-2}
\end{bmatrix}
$$

$$
= \begin{bmatrix}
1 & 1 & \cdots & 1 \\
z_1 & z_2 & \cdots & z_{L-1} \\
\vdots & \vdots & & \vdots \\
z_1^{N-1} & z_2^{N-1} & \cdots & z_{L-1}^{N-1}
\end{bmatrix}
\begin{bmatrix}
C(z_1) & 0 & \cdots & 0 \\
0 & C(z_2) & \ddots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & & C(z_{L-1})
\end{bmatrix}
= 0 \,.
$$

Hence $[C^T\ V]$ is a square matrix of full (column) rank since it is composed of 2 orthogonal blocks, each of full column rank. So we get

$$I = [C^T\ V]([C^T\ V]^T[C^T\ V])^{-1}[C^T\ V]^T = [C^T\ V]\begin{bmatrix} (CC^T)^{-1} & 0 \\ 0 & (V^TV)^{-1} \end{bmatrix}[C^T\ V]^T = P_{C^T}+P_V\ .$$

or hence $P^{\perp}_{C^T} = P_V$ and $\sigma^2_{\tilde{x}^{(d)}_{MMSE}} = \sigma^2_s\ \underline{e}^T_d\ P_V\ \underline{e}_d$, which should be an expression of interest when $\sigma^2_{\tilde{x}^{(d)}_{MMSE}}$ needs to be computed for many values of d when L is small.

(f) $\sum_{d=0}^{N+L-2} \sigma^2_{\tilde{x}^{(d)}_{MMSE}} = \sigma^2_s \sum_{d=0}^{N+L-2} \underline{e}^T_d\ P_V\ \underline{e}_d = \sigma^2_s\,\mathrm{tr}\{P_V\} = \sigma^2_s\,\mathrm{tr}\{(V^TV)^{-1}V^TV\} = \sigma^2_s\,\mathrm{tr}\{I_{L-1}\} = (L-1)\,\sigma^2_s$. Hence $\frac{1}{N+L-1}\sum_{d=0}^{N+L-2} \sigma^2_{\tilde{x}^{(d)}_{MMSE}} = \sigma^2_s\,\frac{L-1}{N+L-1}$ which tends to zero as $N/L \to \infty$ (convergence to a ZF equalizer with zero MSE in absence of noise).

3 Steepest-Descent and Adaptive Filtering Algorithms

Problem 6. Steepest-descent algorithm

Here, we consider the covariance matrix $R_Y = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. The characteristic polynomial is $\mathcal{P}(\lambda) = |R_Y - \lambda I|$ where $|A|$ stands for the determinant of the matrix A:

$$\mathcal{P}(\lambda) = (1-\lambda)^2 - \rho^2 = (1-\lambda-|\rho|)(1-\lambda+|\rho|)\ ,$$

which leads to $\lambda_1 = 1+|\rho|$ and $\lambda_2 = 1-|\rho|$, so that:
a) maximum stepsize: $0 < \mu < \dfrac{2}{1+|\rho|}$
b) fastest convergence obtained for $\mu = 1$ with corresponding mode $|\rho|$.

Problem 7. An application of the LMS algorithm

(a) $\mathrm{E}y_k = \mathrm{E}(a) + \mathrm{E}(w_k) = 0$
(b) $R_{YY} = \mathrm{E}Y_kY_k^T = \mathrm{E}(a\mathbf{1} + W_k)(a\mathbf{1}^T + W_k^T) = \sigma^2_a\mathbf{1}\mathbf{1}^T + \sigma^2_w I$
(c) One has $\mathbf{1}$ as eigenvector of R_{YY}, with $(\sigma^2_a\mathbf{1}\mathbf{1}^T + \sigma^2_w I)\mathbf{1} = (N\sigma^2_a + \sigma^2_w)\mathbf{1}$, we have the corresponding eigenvalue which is $\lambda_1 = N\sigma^2_a + \sigma^2_w$. Note $\mathbf{1}^{\perp}_i$, one of the $N-1$ orthogonal vectors to $\mathbf{1}$ ($\mathbf{1}^T\mathbf{1}^{\perp}_i = 0$). Every $\mathbf{1}^{\perp}_i$ is an eigenvector with corresponding eigenvalue $\lambda_i = \sigma^2_w\ ,\ i = 2\cdots N$.
(d) $0 < \mu < \dfrac{2}{N\sigma^2_a + \sigma^2_w}$.
(e) For fastest convergence: $\mu = \dfrac{2}{N\sigma^2_a + 2\sigma^2_w}$.