Institut EURECOM 11/06/2007
SSP Statistical Signal Processing

TD2: Optimal and Adaptive Filtering, Equalization Partial 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 $S_{v_1}(z)$ and $S_{v_2}(z)$. Before finding the exact solution, we can find the form of the solution by inspection. Assume for a particular frequency $S_{v_1(z)} \gg S_{v_2(z)}$, 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

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

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

$$\begin{aligned} \mathrm{MSE}_{\mathrm{ZF-LE}} &= \frac{\sigma_v^2}{2\pi j}\oint_{|z|=1} \frac{dz}{zC^\dagger(z)C(z)} && (2) \\ &= -\frac{\sigma_v^2}{2\pi j a^*}\oint_{|z|=1} \frac{dz}{(z-a)(z-1/a^*)} && (3) \end{aligned}$$

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

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

(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

$$\begin{cases} A - \frac{1}{2}B = 0 \\ B - \frac{1}{2}A = -\frac{1}{2} \end{cases} \Rightarrow \begin{cases} A = -\frac{1}{3} \\ B = -\frac{2}{3} \end{cases}$$

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

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

(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 $\tilde{x}_k = w_k$ 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.

3 Steepest-Descent and Adaptive Filtering Algorithms

Problem 5. 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 6. An application of the LMS algorithm

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

Problem 7. Stochastic Newton Algorithm

1. We must find the optimum form for H to minimize the error ξ_k^{SN}. Taking the gradient w.r.t. H and setting it to zero yields

$$\nabla_H \xi_k^{\mathrm{SN}}(H) = -2(x_k - Y_k^{\mathrm{T}} H)Y_k + 2A_k(H - H_{k-1}) = 0$$

which gives

$$H_k = (x_k Y_k + A_k H_{k-1})(A_k + Y_k Y_k^{\mathrm{T}})^{-1} = x_k (A_k + Y_k Y_k^{\mathrm{T}})^{-1} Y_k + (A_k + Y_k Y_k^{T})^{-1} A_k H_{k-1} \quad (6)$$

For the third part of the problem and for implementation reasons it is more appropriate to expand the terms $(A_k + Y_k Y_k^T)^{-1}$ using the matrix inversion Lemma as

$$(A_k + Y_k Y_k^T)^{-1} = A_k^{-1} - A_k^{-1} Y_k (1 + Y_k^{\mathrm{T}} A_k^{-1} Y_k)^{-1} Y_k^{\mathrm{T}} A_k^{-1}$$

Inserting this into (6) we get

$$\begin{aligned} H_k &= x_k A_k^{-1} Y_k - x_k A_k^{-1} Y_k (1 + Y_k^{\mathrm{T}} A_k^{-1} Y_k)^{-1} Y_k^{\mathrm{T}} A_k^{-1} Y_k + H_{k-1} - A_k^{-1} Y_k (1 + Y_k^{\mathrm{T}} A_k^{-1} Y_k)^{-1} Y_k^{\mathrm{T}} H_{k-1} \\ &= H_{k-1} + (1 + Y_k^{\mathrm{T}} A_k^{-1} Y_k)^{-1} \left[x_k A_k^{-1} Y_k (1 + Y_k^{\mathrm{T}} A_k^{-1} Y_k) - x_k A_k^{-1} Y_k Y_k^{\mathrm{T}} A_k^{-1} Y_k - A_k^{-1} Y_k Y_k^{\mathrm{T}} H_{k-1} \right] \\ &= H_{k-1} + (1 + Y_k^{\mathrm{T}} A_k^{-1} Y_k)^{-1} (x_k - Y_k^{\mathrm{T}} H_{k-1}) A_k^{-1} Y_k \\ &= H_{k-1} + (1 + Y_k^{\mathrm{T}} A_k^{-1} Y_k)^{-1} \epsilon_k^p \, A_k^{-1} Y_k \end{aligned} \quad (7)$$

2. We see that by setting $A_k = R_{k-1}$ we obtain the RLS algorithm exactly. For LMS we need

$$\frac{A_k^{-1}}{1 + Y_k^{\mathrm{T}} A_k^{-1} Y_k} = \mu I$$

which after some manipulation yields

$$A_k = \frac{1}{\mu} I - Y_k Y_k^{\mathrm{T}}$$

3. To show this relationship we can either replace the expression for A_k into (9) and verify both sides or

$$\begin{aligned} Y_k^{\mathrm{T}} A_k^{-1} Y_k &= \frac{Y_k^{\mathrm{T}} R_{YY}^{-1} Y_k}{1/\mu - Y_k^{\mathrm{T}} R_{YY}^{-1} Y_k} \\ \Leftrightarrow \quad (1/\mu) Y_k^{\mathrm{T}} A_k^{-1} Y_k - (Y_k^{\mathrm{T}} A_k^{-1} Y_k \left(Y_k^{\mathrm{T}} R_{YY}^{-1} Y_k \right) &= Y_k^{\mathrm{T}} R_{YY}^{-1} Y_k \\ \Leftrightarrow \quad Y_k^{\mathrm{T}} R_{YY}^{-1} Y_k \left(1 + Y_k^{\mathrm{T}} A_k^{-1} Y_k \right) &= Y_k^{\mathrm{T}} \left(\tfrac{1}{\mu} A_k^{-1} \right) Y_k \\ \Leftrightarrow \quad \mu R_{YY}^{-1} &= \frac{A_k^{-1}}{1 + Y_k^T A_k^{-1} Y_k} \end{aligned} \quad (8)$$

Using this relationship in (7) we get the update equation

$$H_k = H_{k-1} + \mu \, \epsilon_k^p \, R_{\mathrm{YY}}^{-1} Y_k$$

Defining

$$\widetilde{H}_k = H^0 - H_k$$

where $H^0 = R_{YY}^{-1} R_{Yx}$ is the Wiener filter solution, we obtain

$$\begin{aligned}
\widetilde{H}_k &= H^0 - H_{k-1} - \mu\, \epsilon_k^p\, R_{YY}^{-1} Y_k \\
&= \widetilde{H}_{k-1} - \mu(x_k + Y_k^{\mathrm{T}} \widetilde{H}_{k-1} - Y_k^{\mathrm{T}} H^0) R_{YY}^{-1} Y_k \\
&= \widetilde{H}_{k-1} - \mu(\widetilde{x}_k - Y_k^T \widetilde{H}_{k-1}) R_{YY}^{-1} Y_k \\
&= (I - \mu R_{YY}^{-1} Y_k Y_k^{\mathrm{T}}) \widetilde{H}_{k-1} - \mu\, \widetilde{x}_k\, R_{YY}^{-1}\, Y_k
\end{aligned} \tag{9}$$

Taking the expected value of both sides yields $\mathrm{E}\widetilde{H}_k = (1-\mu)\,\mathrm{E}\widetilde{H}_{k-1}$ so that we have convergence for $|1-\mu| < 1$ or $\mu \in (0,2)$. The steady state value is H^0. By choosing $\mu = 1$ we have the fastest convergence which occurs in 1 update!