SGN 21006 Advanced Signal Processing: Lecture 3: Optimal Wiener Filters

Ioan Tabus

Department of Signal Processing
Tampere University of Technology
Finland

Complete descriptions of random signals

- A random signal is a sequence of N random variables $X_1, X_2, \ldots, X_N$, where each random variable may have a different probability distribution function, e.g., X_i is distributed according to $p(X_i)$.
- The complete description of a random signal is given by the joint distribution of the N random variables. If variables are discrete (i.e. they can take only a finite number of values), denote the set of values taken by the X_i as $\mathcal{X}_i$; then $Prob(X_1 = x_1; X_2 = x_2, \ldots, X_N = x_N)$ is enough for determining the marginal distributions $Prob(X_i = x_i)$, for any $i \leq N$, by summing with respect to all values of the r.v. $X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots X_N$

$$Prob(X_i = x_i) = \sum_{x_1 \in \mathcal{X}_1} \ldots \sum_{x_{i-1} \in \mathcal{X}_{i-1}} \sum_{x_{i+1} \in \mathcal{X}_{i+1}} \ldots \sum_{x_N \in \mathcal{X}_N} Prob(X_1 = x_1; X_2 = x_2, \ldots, X_N = x_N)$$

- Such a description is rarely available in practice, but for simulations it is possible to use it when studying statistical properties of signal models.
- In practice one has a particular sequence $x_1, x_2, \ldots, x_N$ which is called one "realization" of the random signal. If one has many realizations of the signal available he can extract the random signal model out of these realizations.

Descriptions of random signals: simplifications

- *Mean removal*: The mean value of a random variable is most often available. Making the substitution $X' = X - E[X]$ the mean of the new random variable X' is 0. In the rest of the lecture the implicit assumption is that the mean has been taken out of the random variables, so that all random variables under discussion have zero mean.
- The "second order" description of the signal operates with the expected product of two random variables, say X_k and X_m, which for zero mean variables coincides with both correlation and covariance

  $$R(k, m) = EX_k X_m$$

  It tells how much the sample k influences the signal at sample m.
- for *wide-sense stationary* signals the function $R(k, m)$ depends only on the difference between k and m and we then denote

  $$R(i) \quad = \quad EX_k X_{k+i} = EX_m X_{m+i} = EX_m X_{m-i} = R(-i)$$

- for *wide-sense stationary* signals the value

  $$R(0) \quad = \quad EX_k^2 = \sigma^2$$

  is the variance of the signal.

Purely random signals

- If the variables X_k and X_m are independent for all values $k \neq m$, then also $R(i) = 0$ for all $i \neq 0$.
- A pure random signal has no memory, the value at sample k does not influence the value of the signal at $m \neq k$.
- White noise assumes only un-correlatedness of X_k and X_m, i.e., $E[X_k X_m] = 0$. If additionally we assume that the white noise is jointly Gaussian, independence for all X_k and X_m, with $k \neq m$ is obtained.
- Most often we assume white noise of zero mean, which is a sequence $e_1, \ldots, e_N$ where $E[e_i] = 0$ and $E[e_i e_j] = 0$ for all $i \neq j$.
- We can generate in Matlab, for the needs of simulations, such sequences using $e = randn(N,1)$ if we want to generate continuous valued samples.
- if we need a binary sequence $e_1, \ldots, e_N$ where $E[e_i] = 0.5$ and $E[e_i e_j] = 0$, we can generate in matlab $e = rand(N, 1) > 0.5$, where the intermediate value v=rand(1,1) is uniformly distributed between 0 and 1, and where $e = 0$ if the uniform variable $v <= 0.5$, and $e = 1$ if $v > 0.5$.

Temporal average versus statistical average

- Consider the sequence of N random variables $X_1, \ldots, X_N$.
  - The expectation, or statistical average, $R(i) = E[X_k X_{k+i}]$ can be computed if one knows the joint probability distribution of (X_k, X_{k+i}).
  - For continuous pdf $p(x_i, x_{k+i})$, the correlation is

$$R(i) = E[X_k X_{k+i}] = \int_{x_k=-\infty}^{\infty} \int_{x_{k+i}=-\infty}^{\infty} x_k x_{k+i} p(x_k, x_{k+i}) dx_k dx_{k+i}$$

  - For discrete valued random variables X, with probability mass function $Prob(X_i = x_i, X_{k+i} = x_{k+i}) = p(x_k, x_{k+i})$, the correlation is

$$R(i) = E[X_k X_{k+i}] = \sum_{x_k=-\infty}^{\infty} \sum_{x_{k+i}=-\infty}^{\infty} x_k x_{k+i} p(x_k, x_{k+i})$$

- Consider now a single realization, $x_1, \ldots, x_N$, of the sequence of N random variables $X_1, \ldots, X_N$
  - The corresponding time averages along the sequence are denoted with a hat mark

$$\hat{R}(i) = \frac{1}{N-i} \sum_{k=1}^{N-i} x_k x_{k+i}$$

  - In the ergodic theory the assumption is that statistical average is equal to the limit of the temporal average, when the number of samples N goes to infinity.
  - Hence, for a finite sample of size N, one can approximate $R(i)$ with $\hat{R}(i)$, the degree of approximation being better as N grows larger.
- In the first part of the course the ergodic assumption will be used for deriving optimal filtering algorithms, by first manipulating expressions involving statistical averages, and after arriving at desired results, replacing the statistical averages by their approximations, namely the temporal averages.

Optimal Wiener filters: Problem statement

- Given the set of input samples $\{u(0), u(1), u(2), \ldots\}$ and the set of desired response $\{d(0), d(1), d(2), \ldots\}$
- In the family of filters computing their output according to

$$y(n) = \sum_{k=0}^{\infty} w_k u(n-k), \quad n = 0, 1, 2, \ldots \tag{1}$$

- Find the parameters $\{w_0, w_1, w_2, \ldots\}$ such as to minimize the mean square error defined as

$$J = E[e(n)^2]$$

where the error signal is

$$e(n) = d(n) - y(n) = d(n) - \sum_{l=0}^{\infty} w_l u(n-l)$$

□
The family of filters (1) is the family of linear discrete time filters (IIR or FIR).

Optimal Wiener filters: Principle of orthogonality

Define the gradient operator ∇, having its k-th entry

$$\nabla_k = \frac{\partial}{\partial w_k}$$

and thus, the $k-$th entry of the gradient of criterion J is (remember, $e(n) = d(n) - \sum_{l=0}^{\infty} w_l u(n-l)$)

$$\nabla_k J = \frac{\partial J}{\partial w_k} = 2E\left[e(n)\frac{\partial e(n)}{\partial w_k}\right] = -2E\left[e(n)u(n-k)\right]$$

For the criterion to attain its minimum, the gradient of the criterion must be identically zero, that is

$$\nabla_k J = 0, \quad k = 0, 1, 2, \dots$$

resulting in the fundamental

$$\textbf{Principle of orthogonality:} \quad E\left[e_o(n)u(n-k)\right] = 0, \quad k = 0, 1, 2, \dots$$

Stated in words:

- The criterion J attains its minimum *iff*
- the estimation error $e_o(n)$ is orthogonal to the samples $u(i)$ which are used to compute the filter output.

We will index with o all the variables e.g. e_o, y_o computed using the optimal parameters $\{w_{o0}, w_{o1}, w_{o2}, \dots\}$.
Let us compute the cross- correlation

$$E\left[e_o(n)y_o(n)\right] = E\left[e_o(n)\sum_{k=0}^{\infty} w_{ok}u(n-k)\right] = \sum_{k=0}^{\infty} w_{ok}E\left[u(n-k)e_o(n)\right] = 0$$

Otherwise stated, in words, we have the following **Corollary of Orthogonality Principle:**

- The estimation error $e_0(n)$ is orthogonal to the filter output $y_0(n)$ (when the filter has optimal parameters).

Optimal Wiener filters: Wiener – Hopf equations

From the orthogonality *estimation error – input window samples* we have

$$E\left[u(n-k)e_0(n)\right] = 0, \quad k = 0, 1, 2, \ldots$$

$$E\left[u(n-k)(d(n) - \sum_{i=0}^{\infty} w_{oi}u(n-i))\right] = 0, \quad k = 0, 1, 2, \ldots$$

$$\sum_{i=0}^{\infty} w_{oi}E\left[u(n-k)u(n-i)\right] = E\left[u(n-k)d(n)\right], \quad k = 0, 1, 2, \ldots$$

But
* $E\left[u(n-k)u(n-i)\right] = r(i-k)$ is the autocorrelation function of input signal $u(n)$ at lag $i-k$
* $E\left[u(n-k)d(n)\right] = p(-k)$ is the cross-correlation between the filter input $u(n-k)$ and the desired signal $d(n)$
and therefore

$$\sum_{i=0}^{\infty} w_{oi}r(i-k) = p(-k), \quad k = 0, 1, 2, \ldots \quad WIENER - HOPF$$

Solution of the Wiener – Hopf equations for linear transversal filters (FIR)

A linear transversal filter (FIR) has a finite number of weights (taps), M,

$$y(n) = \sum_{k=0}^{M-1} w_k u(n-k), \quad n = 0, 1, 2, \ldots$$

and since only $w_0, w_1, w_2, \ldots \; w_{M-1}$ are nonzero, Wiener-Hopf equations become

$$\sum_{i=0}^{M-1} w_{oi} r(i-k) = p(-k), \quad k = 0, 1, 2, \ldots, M-1 \quad WIENER - HOPF$$

which is a system of M equations with M unknowns: $\{w_{o,0}, w_{o,1}, w_{o,2}, \ldots, w_{o,M-1}\}$.

Matrix formulation of Wiener – Hopf equations

Let us denote

$$\underline{u}(n) = \begin{bmatrix} u(n) & u(n-1) & u(n-2) & \dots & u(n-M+1) \end{bmatrix}^T$$

$$R = E[\underline{u}(n)\underline{u}^T(n)] = E \begin{bmatrix} u(n) \\ u(n-1) \\ u(n-2) \\ u(n-M+1) \end{bmatrix} \begin{bmatrix} u(n) & u(n-1) & u(n-2) & \dots & u(n-M+1) \end{bmatrix} \quad (2)$$

$$= \begin{bmatrix} Eu(n)u(n) & Eu(n)u(n-1) & \dots & Eu(n)u(n-M+1) \\ Eu(n-1)u(n) & Eu(n-1)u(n-1) & \dots & Eu(n-1)u(n-M+1) \\ \vdots & \vdots & \vdots & \vdots \\ Eu(n-M+1)u(n) & Eu(n-M+1)u(n-1) & \dots & Eu(n-M+1)u(n-M+1) \end{bmatrix}$$

$$= \begin{bmatrix} r(0) & r(1) & \dots & r(M-1) \\ r(1) & r(0) & \dots & r(M-2) \\ \vdots & \vdots & \vdots & \vdots \\ r(M-1) & r(M-2) & \dots & r(0) \end{bmatrix}$$

$$\underline{p} = E[\underline{u}(n)d(n)] = \begin{bmatrix} p(0) & p(-1) & p(-2) & \dots & p(1-M) \end{bmatrix}^T \quad (3)$$

$$\underline{w}_0 = \begin{bmatrix} w_{o,0} & w_{o,1} & \dots & w_{o,M-1} \end{bmatrix}^T \quad (4)$$

Now Wiener – Hopf equations can be written in a compact form

$$R\underline{w}_0 = \underline{p} \quad \text{with solution} \quad \underline{w}_o = R^{-1}\underline{p} \quad (5)$$

Optimal Wiener filters: Mean square error surface

Let us define

$$e_{\underline{w}}(n) = d(n) - \sum_{k=0}^{M-1} w_k u(n-k) = d(n) - \underline{w}^T \underline{u}(n)$$

Then the cost function can be written as

$$\begin{aligned}
J_{\underline{w}} &= E[e_{\underline{w}}(n)e_{\underline{w}}(n)] = E[(d(n) - \underline{w}^T \underline{u}(n))(d(n) - \underline{u}^T(n)\underline{w})] \\
&= E[d^2(n) - d(n)\underline{u}^T(n)\underline{w} - \underline{w}^T \underline{u}(n)d(n) + \underline{w}^T \underline{u}(n)\underline{u}^T(n)\underline{w}] \\
&= E[d^2(n)] - E[d(n)\underline{u}^T(n)]\underline{w} - \underline{w}^T E[\underline{u}(n)d(n)] + \underline{w}^T E[\underline{u}(n)\underline{u}^T(n)]\underline{w} \\
&= E[d^2(n)] - 2E[d(n)\underline{u}^T(n)]\underline{w} + \underline{w}^T E[\underline{u}(n)\underline{u}^T(n)]\underline{w} \\
&= \sigma_d^2 - 2\underline{p}^T \underline{w} + \underline{w}^T R\underline{w} \\
&= \sigma_d^2 - 2\sum_{i=0}^{M-1} p(-i)w_i + \sum_{l=0}^{M-1}\sum_{i=0}^{M-1} w_l w_i R_{i,l}
\end{aligned}$$

Optimal Wiener filters: Minimum Mean square error

Minimum Mean square error
Using the form of the criterion

$$J_{\underline{w}} = \sigma_d^2 - 2\underline{p}^T\underline{w} + \underline{w}^T R\underline{w}$$

one can find the value of the minimum criterion (remember, $R\underline{w}_0 = \underline{p}$ and $\underline{w}_o = R^{-1}\underline{p}$):

$$\begin{aligned}
J_{\underline{w}_o} &= \sigma_d^2 - 2\underline{p}^T\underline{w}_o + \underline{w}_o^T R\underline{w}_o = \sigma_d^2 - 2\underline{w}_o^T R\underline{w}_o + \underline{w}_o^T R\underline{w}_o \\
&= \sigma_d^2 - \underline{w}_o^T R\underline{w}_o \\
&= \sigma_d^2 - \underline{w}_o^T \underline{p} \\
&= \sigma_d^2 - \underline{p}^T R^{-1}\underline{p}
\end{aligned}$$

Optimal Wiener filters: Quadratic form of the Error - performance surface

(Parenthesis: How to compute a scalar out of a vector $\underline{w}$, containing the entries of $\underline{w}$ at power one (linear combination) or at power two (quadratic form):

* linear combination (first order form) $\underline{a}^T \underline{w} = \sum_{l=0}^{M-1} a_l w_l$;

* quadratic form $\underline{w}^T R \underline{w} = \sum_{l=0}^{M-1} \sum_{i=0}^{M-1} w_l w_i R_{i,l} = w_0^2 R_{0,0} + w_0 w_1 R_{1,0} + \ldots w_{M-1}^2 R_{M-1,M-1}$)

How can we rewrite the criterion

$$J_{\underline{w}} \quad = \quad \sigma_d^2 - 2\underline{p}^T \underline{w} + \underline{w}^T R \underline{w} \tag{6}$$

in a quadratic form (how to complete a perfect "square", encompassing $-2\underline{p}^T \underline{w}$)?
Consider first the case when $\underline{w}$ is simply a scalar (resulting also in scalars R, $\underline{r}$, $\underline{p}$)

$$J_w = Rw^2 - 2pw + \sigma_d^2 = R(w^2 - 2w\frac{p}{R}) + \sigma_d^2 = R(w^2 - 2w\frac{p}{R} + \frac{p^2}{R^2}) - \frac{p^2}{R} + \sigma_d^2 = R(w - \frac{p}{R})^2 - \frac{p^2}{R} + \sigma_d^2$$

In the case when $\underline{w}$ is a vector, the term corresponding to the one-dimensional $\frac{p^2}{R}$ is $\underline{p}^T R^{-1} \underline{p}$

$$\begin{aligned} J_{\underline{w}} &= \underline{w}^T R \underline{w} - 2\underline{p}^T \underline{w} + \underline{p}^T R^{-1} \underline{p} - \underline{p}^T R^{-1} \underline{p} + \sigma_d^2 = (\underline{w} - R^{-1}\underline{p})^T R (\underline{w} - R^{-1}\underline{p}) - \underline{p}^T R^{-1} \underline{p} + \sigma_d^2 \\ &= J_{\underline{w}_0} + (\underline{w} - R^{-1}\underline{p})^T R (\underline{w} - R^{-1}\underline{p}) \\ &= J_{\underline{w}_0} + (\underline{w} - \underline{w}_0)^T R (\underline{w} - \underline{w}_0) \end{aligned}$$

□

Optimal Wiener filters: Canonical form of the Error - performance surface

Let $\lambda_1, \lambda_2, \ldots, \lambda_M$ be the eigenvalues and (generally the complex) eigenvectors $\mu_1, \mu_2, \ldots, \mu_M$ of the matrix R, thus satisfying

$$R\mu_i = \lambda_i \mu_i \tag{7}$$

Then the matrix $Q = [\mu_1 \ \mu_2 \ \ldots \ \mu_M]$ can transform R to a diagonal form Λ as follows

$$R = Q\Lambda Q^H \tag{8}$$

where the superscript H means complex conjugation and transposition. Then

$$J_{\underline{w}} = J_{\underline{w}_0} + (\underline{w} - \underline{w}_0)^T R(\underline{w} - \underline{w}_0) = J_{\underline{w}_0} + (\underline{w} - \underline{w}_0)^T Q\Lambda Q^H(\underline{w} - \underline{w}_0)$$

Introduce now the transformed version of the tap vector w as

$$\underline{\nu} = Q^H(\underline{w} - \underline{w}_o) \tag{9}$$

Now the quadratic form can be put into its canonical form

$$\begin{aligned} J &= J_{\underline{w}_o} + \underline{\nu}^H \Lambda \nu \\ &= J_{\underline{w}_o} + \sum_{i=1}^{M} \lambda_i \nu_i \nu_i^* \\ &= J_{\underline{w}_o} + \sum_{i=1}^{M} \lambda_i |\nu_i|^2 \end{aligned}$$

Optimal Wiener Filter Design for a Given Application

Application: channel equalization

- *(Useful) Signal Generating Model* The model is given by the transfer function

$$H_1(z) = \frac{D(z)}{V_1(z)} = \frac{1}{1 + az^{-1}} = \frac{1}{1 + 0.8458z^{-1}}$$

or the difference equation

$$d(n) + ad(n-1) = v_1(n) \qquad d(n) + 0.8458d(n-1) = v_1(n)$$

where $\sigma_{v_1}^2 = r_{v_1}(0) = 0.27$

- *The channel (perturbation) model* is more complex. It involves a low pass filter with a transfer function

$$H_2(z) = \frac{X(z)}{D(z)} = \frac{1}{1 + bz^{-1}} = \frac{1}{1 - 0.9458z^{-1}}$$

leading for the variable $x(n)$ to the difference equation

$$x(n) = 0.9458x(n-1) + d(n)$$

and a white noise corruption ($x(n)$ and $v_2(n)$ are uncorrelated)

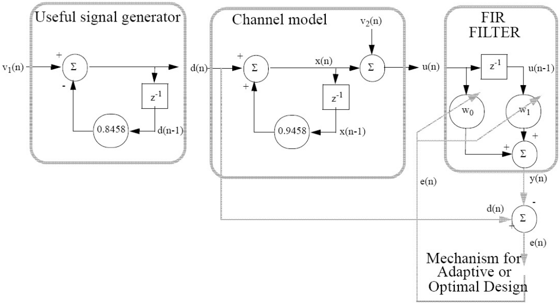
$$u(n) = x(n) + v_2(n)$$

with $\sigma_{v_2}^2 = r_{v_2}(0) = 0.1$ resulting in the final measurable signal $u(n)$.

- *FIR Filter* The signal $u(n)$ will be filtered in order to recover the original (useful) $d(n)$ signal, using the filter

$$y(n) = w_0 u(n) + w_1 u(n-1)$$

Application: channel equalization



Optimal Wiener filters: Needed expectations

We plan to apply the Wiener – Hopf equations

$$\begin{bmatrix} r_u(0) & r_u(1) \\ r_u(1) & r_u(0) \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{bmatrix} Ed(n)u(n) \\ Ed(n)u(n-1) \end{bmatrix}$$

The signal $x(n)$ obeys the generation model

$$H(z) = \frac{X(z)}{V_1(z)} = H_1(z)H_2(z) = \frac{1}{1+az^{-1}}\frac{1}{1+bz^{-1}} = \frac{1}{1+a_1z^{-1}+a_2z^{-2}} = \frac{1}{1-0.1z^{-1}-0.8z^{-2}}$$

and thus

$$x(n) + a_1x(n-1) + a_2x(n-2) = v_1(n)$$

Using the fact that $x(n)$ and $v_2(n)$ are uncorrelated and

$$u(n) = x(n) + v_2(n)$$

it results

$$r_u(k) = r_x(k) + r_{v_2}(k)$$

and consequently, since for white noise $r_{v_2}(0) = \sigma_{v_2}^2 = 0.1$ and $r_{v_2}(1) = 0$ it follows

$$r_u(0) = r_x(0) + 0.1, \text{ and } \quad r_u(1) = r_x(1)$$

AR process: autocorrelations and crosscorrelations

Now we concentrate to find $r_x(0)$, $r_x(1)$ for the AR process

$$x(n) + a_1 x(n-1) + a_2 x(n-2) = v(n)$$

First multiply in turn the equation with $x(n)$, $x(n-1)$ and $x(n-2)$ and then take the expectation

$$Ex(n)\times \rightarrow \quad Ex(n)x(n) + Ex(n)a_1 x(n-1) + Ex(n)a_2 x(n-2) = Ex(n)v(n)$$
$$\text{resulting in} \quad r_x(0) + a_1 r_x(1) + a_2 r_x(2) = Ex(n)v(n) = \sigma_v^2$$

$$Ex(n-1)\times \rightarrow \quad Ex(n-1)x(n) + Ex(n-1)a_1 x(n-1) + Ex(n-1)a_2 x(n-2) = Ex(n-1)v(n)$$
$$\text{resulting in} \quad r_x(1) + a_1 r_x(0) + a_2 r_x(1) = Ex(n-1)v(n) = 0$$

$$Ex(n-2)\times \rightarrow \quad Ex(n-2)x(n) + Ex(n-2)a_1 x(n-1) + Ex(n-2)a_2 x(n-2) = Ex(n-2)v(n)$$
$$\text{resulting in} \quad r_x(2) + a_1 r_x(1) + a_2 r_x(0) = Ex(n-2)v(n) = 0$$

The equality $Ex(n)v(n) = \sigma_v^2$ can be obtained multiplying the AR model difference equation with $v(n)$ and then taking expectations

$$Ev(n)\times \rightarrow \quad Ev(n)x(n) + Ev(n)a_1 x(n-1) + Ev(n)a_2 x(n-2) = Ev(n)v(n)$$
$$\text{resulting in} \quad Ev(n)x(n) = \sigma_v^2$$

since $v(n)$ is uncorrelated with older values, $x(n-\tau)$.

AR process: Yule Walker equations

We obtained the most celebrated Yule Walker equations:

$$\begin{aligned} r_x(0) + a_1 r_x(1) + a_2 r_x(2) &= \sigma_v^2 \\ r_x(1) + a_1 r_x(0) + a_2 r_x(1) &= 0 \\ r_x(2) + a_1 r_x(1) + a_2 r_x(0) &= 0 \end{aligned}$$

or as usually given in matrix form

$$\begin{bmatrix} r_x(0) & r_x(1) & r_x(2) \\ r_x(1) & r_x(0) & r_x(1) \\ r_x(2) & r_x(1) & r_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sigma_v^2 \\ 0 \\ 0 \end{bmatrix}$$

But we need to use the equations differently:

$$\begin{bmatrix} 1 & a_1 & a_2 \\ a_1 & 1 + a_2 & 0 \\ a_2 & a_1 & 1 \end{bmatrix} \begin{bmatrix} r_x(0) \\ r_x(1) \\ r_x(2) \end{bmatrix} = \begin{bmatrix} \sigma_v^2 \\ 0 \\ 0 \end{bmatrix}$$

Channel equalization: finding the necessary expectations

Solving for $r_x(0)$, $r_x(1)$, $r_x(2)$ we obtain

$$r_x(0) = \left(\frac{1+a_2}{1-a_2}\right)\frac{\sigma_v^2}{(1+a_2)^2 - a_1^2}$$

$$r_x(1) = \frac{-a_1}{1+a_2} r_x(0)$$

$$r_x(2) = \left(-a_2 + \frac{a_1^2}{1+a_2}\right) r_x(0)$$

In our example we need only the first two values, $r_x(0)$, $r_x(1)$, which result to be $r_x(0) = 1, \; r_x(1) = 0.5$.
Now we will solve for the cross-correlations $Ed(n)u(n)$, $Ed(n)u(n-1)$. First observe

$$\begin{aligned} Eu(n)d(n) &= E(x(n) + v_2(n))d(n) = Ex(n)d(n) \\ Eu(n-1)d(n) &= E(x(n-1) + v_2(n-1))d(n) = Ex(n-1)d(n) \end{aligned}$$

and now take as a "master" difference equation

$$x(n) + bx(n-1) = d(n)$$

and multiply it in turn with $x(n)$ and $x(n-1)$ and then take the expectation

$$\begin{aligned} Ex(n) \to \qquad & Ex(n)x(n) + Ex(n)bx(n-1) = Ex(n)d(n) \\ & Ex(n)d(n) = r_x(0) + br_x(1) \end{aligned}$$

Channel equalization: final solution

$$Ex(n-1) \rightarrow \qquad \begin{aligned} Ex(n-1)x(n) + Ex(n-1)bx(n-1) &= Ex(n-1)d(n) \\ Ex(n-1)d(n) &= r_x(1) + br_x(0) \end{aligned}$$

Using the numerical values, one obtains

$$Eu(n)d(n) = Ex(n)d(n) = 0.5272 \quad Eu(n-1)d(n) = Ex(n-1)d(n) = -0.4458$$

Now we have all necessary variables needed to write the Wiener – Hopf equations

$$\begin{bmatrix} r_u(0) & r_u(1) \\ r_u(1) & r_u(0) \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{bmatrix} Ed(n)u(n) \\ Ed(n)u(n-1) \end{bmatrix}$$

$$\begin{bmatrix} 1.1 & 0.5 \\ 0.5 & 1.1 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{bmatrix} 0.5272 \\ -0.4458 \end{bmatrix}$$

resulting in

$$\begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{bmatrix} 0.8360 \\ -0.7853 \end{bmatrix}$$