Chapter 2
Fundamentals of Adaptive Filter Theory

In this chapter we will treat some fundamentals of the adaptive filtering theory highlighting the system identification problem. We will introduce a signal and system model that will be used throughout this book.

2.1 Signal and System Model

In this book, we will assume that the acoustic echo path from the loudspeaker to the microphone is linear and can be modeled by a finite impulse response (FIR) filter.

There are two possibilities for the matrix representation of a MIMO FIR-system which are equivalent with respect to the output of the system. These two models are introduced in the following subsections. The first representation which we will call the standard representation has a direct correspondence to the underlying physics that will be treated in detail in Sect. 4.1.1. The second representation is the compact representation which is preferable in the absence of prior knowledge about the system, as we will show later on in this monograph.

2.1.1 Standard Representation

The convolution of an input signal of length M with an FIR filter of the order L can be expressed by a matrix vector multiplication where the matrix represents the filter and exhibits a Toeplitz structure with M columns [1]

K. Helwani, *Adaptive Identification of Acoustic Multichannel Systems Using Sparse Representations*, T-Labs Series in Telecommunication Services,
DOI 10.1007/978-3-319-08954-6_2

$$\overset{\circ}{\mathbf{H}}_{[M],pq} := \begin{bmatrix} h_{pq,0} & 0 & \dots & 0 \\ h_{pq,1} & h_{pq,0} & & \vdots \\ \vdots & h_{pq,1} & \ddots & 0 \\ h_{pq,L-1} & \vdots & & h_{pq,0} \\ 0 & h_{pq,L-1} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & h_{pq,L-1} \end{bmatrix}. \tag{2.1}$$

In the MIMO case, the system representing the $P \times Q$ acoustic paths is composed as

$$\overset{\circ}{\mathbf{H}} := \begin{bmatrix} \overset{\circ}{\mathbf{H}}_{[M],11} & \dots & \overset{\circ}{\mathbf{H}}_{[M],P1} \\ \vdots & & \vdots \\ \overset{\circ}{\mathbf{H}}_{[M],1Q} & \dots & \overset{\circ}{\mathbf{H}}_{[M],PQ} \end{bmatrix}. \tag{2.2}$$

For a finite input signal block with length PM

$$\begin{aligned} \mathbf{x}(n) &= [\mathbf{x}_1^{\mathrm{T}}(n), \mathbf{x}_2^{\mathrm{T}}(n), \dots, \mathbf{x}_P^{\mathrm{T}}(n)]^{\mathrm{T}}, \\ \mathbf{x}_p(n) &= [x_p(n), x_p(n+1), \dots, x_p(n+M-1)]^{\mathrm{T}}, \end{aligned} \tag{2.3}$$

one obtains a block of the output signal $\mathbf{y}$

$$\mathbf{y}(n) = \overset{\circ}{\mathbf{H}}\, \mathbf{x}(n), \tag{2.4}$$

with

$$\begin{aligned} \mathbf{y}(n) &= [\mathbf{y}_1^{\mathrm{T}}(n), \mathbf{y}_2^{\mathrm{T}}(n), \dots, \mathbf{y}_Q^{\mathrm{T}}(n)]^{\mathrm{T}}, \\ \mathbf{y}_q(n) &= [y_q(n), y_q(n+1), \dots, y_q(n+(L+M)-1)]^{\mathrm{T}}. \end{aligned} \tag{2.5}$$

2.1.2 Compact Representation

An equivalent matrix representation of the MIMO system w.r.t. the output can be achieved by employing the Toeplitz structure, which results from the convolution operation, in the input signal. In this case the MIMO system can be expressed by a non redundant matrix $\mathbf{H}$ with the dimensions $P \cdot L \times Q$. This is composed by $P \cdot Q$ subfilters, $\mathbf{h}_{pq} = [h_{pq,0}, h_{pq,1}, \dots, h_{pq,L-1}]^{\mathrm{T}}$.

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_{11} & \cdots & \mathbf{h}_{1Q} \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{P1} & \cdots & \mathbf{h}_{PQ} \end{bmatrix}, \tag{2.6}$$

a vector with the output samples at the time instant n is obtained by

$$\mathbf{y}(n) = \mathbf{H}^{\mathrm{T}} \cdot \mathbf{x}(n), \tag{2.7}$$

with $\mathbf{x}(n)$ as input signal vector with the length-PL (loudspeaker signals in the near-end) interpreted as one column of a block-Toeplitz matrix that represents the signal

$$\begin{aligned} \mathbf{x}(n) &= [\mathbf{x}_1^{\mathrm{T}}(n), \mathbf{x}_2^{\mathrm{T}}(n), \ldots, \mathbf{x}_P^{\mathrm{T}}(n)]^{\mathrm{T}}, \\ \mathbf{x}_p(n) &= [x_p(n), x_p(n-1), \ldots, x_p(n-L+1)]^{\mathrm{T}}. \end{aligned} \tag{2.8}$$

Please note that the output vector in Eq. (2.7) contains only the current output sample for each output channel.

2.2 Optimal System Identification in Least-Squares Sense

2.2.1 The Wiener–Hopf Equation

As stated in Sect. 1.1 the most popular optimization criterion is the LSE. Typically, in the scenario given by an MC-AEC setup, the MIMO identification problem is considered as series of independent MISO systems for each microphone channel [2]. The echo paths from the P loudspeakers to a single microphone with the index q are identified by minimizing the following cost function

$$J(\widehat{\mathbf{h}}_q) = \widehat{\mathcal{E}}\{|e_q(n)|^2\} = \hat{\mathcal{E}}\{|y_q(n) - \widehat{\mathbf{h}}_q^{\mathrm{H}}\mathbf{x}(n)|^2\}, \tag{2.9}$$

here the definitions from Sect. 2.1.2 are used. Determining the minimum of the quadratic cost function in Eq. (2.9) requires a gradient calculation w.r.t. $\widehat{\mathbf{h}}_q$. By applying the chain rule and Eq. (A.3) we obtain

$$\begin{aligned} \nabla_{\widehat{\mathbf{h}}_q} J &= -2\hat{\mathcal{E}}\{\mathbf{x}(n)[y_q^*(k) - \widehat{\mathbf{h}}_q^{\mathrm{T}}\mathbf{x}^*(n)]\}, \\ &= -2\mathbf{r}_{\mathbf{x}y_q}(n) + 2\mathbf{R}_{\mathbf{xx}}(n)\widehat{\mathbf{h}}_q, \end{aligned} \tag{2.10}$$

where $\mathbf{R}_{\mathbf{xx}}$ denotes the $PL \times PL$ correlation matrix containing all inter- and intra-channel correlations

$$\mathbf{R}_{\mathbf{xx}}(n) := \mathcal{E}\{\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)\}, \tag{2.11}$$

and $\mathbf{r}_{\mathbf{x}y}$ is the $PL \times 1$ crosscorrelation vector

$$\mathbf{r}_{\mathbf{x}y_q}(n) := \mathcal{E}\{\mathbf{x}(n)y_q^*(n)\}. \tag{2.12}$$

At the minimum of the cost function we set

$$\nabla_{\widehat{\mathbf{h}}_q} J = \mathbf{0}, \tag{2.13}$$

where $\mathbf{0}$ is a $PL \times 1$ zero vector. Hence, by substitution into Eq. (2.10) we obtain the optimal estimation of $\widehat{\mathbf{h}}_q$ in the LS-sense by the so-called Wiener-Hopf equation [3]

$$\widehat{\mathbf{h}}_{q,\mathrm{opt}} = \mathbf{R}_{\mathbf{xx}}^{-1}(n)\mathbf{r}_{\mathbf{x}y_q}(n). \tag{2.14}$$

For the optimal estimation of the complete MIMO system, it can be easily seen that

$$\widehat{\mathbf{H}}_{\mathrm{opt}} = \mathbf{R}_{\mathbf{xx}}^{-1}(n)\mathbf{R}_{\mathbf{xy}}(n), \tag{2.15}$$

with

$$\mathbf{R}_{\mathbf{xy}}(n) := \mathcal{E}\{\mathbf{x}(n)\mathbf{y}^{\mathrm{T}}(n)\}. \tag{2.16}$$

2.2.2 Derivation of Iterative Estimation Approaches

2.2.2.1 The Recursive Least-Squares Algorithm Based on the Matrix Inversion Lemma

Preferably, the optimal solution is estimated iteratively. The derivation of an iterative solution of the given estimation problem that can be found in the literature is based typically, either on the matrix inversion lemma (Woodbury matrix identity) [3] or on the Newton method [4–6].

The starting point for the derivation based on the matrix inversion lemma is the iterative estimation of the autocorrelation matrix

$$\mathbf{R}_{\mathbf{xx}}(n) = \alpha\mathbf{R}_{\mathbf{xx}}(n) + (1-\alpha)\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n), \tag{2.17}$$

where α denotes the forgetting factor. The crosscorrelation vector is estimated by

$$\mathbf{r}_{\mathbf{x}y_q}(n) = \alpha\mathbf{R}_{\mathbf{xx}}(n) + (1-\alpha)\mathbf{x}(n)y_q^*(n). \tag{2.18}$$

Applying the matrix inversion lemma Eq. (A.7) and substituting in the normal equation leads to the recursive least-squares update equation [3, 7, 8]

$$\hat{\mathbf{h}}_q(n) = \hat{\mathbf{h}}_q(n-1) + \mathbf{k}(n)\underbrace{\left(y_q^*(n) - \mathbf{x}^{\mathrm{H}}(n)\hat{\mathbf{h}}_q(n-1)\right)}_{:=e_q^*(n)}, \tag{2.19}$$

with $\mathbf{k}(n)$ denoting the so-called Kalman-gain defined as

$$\mathbf{k}(n) = \frac{\alpha^{-1}(1-\alpha)\mathbf{R}_{\mathbf{xx}}^{-1}(n-1)\mathbf{x}(n)}{1+\alpha^{-1}(1-\alpha)\mathbf{x}^{\mathrm{H}}(n)\mathbf{R}_{\mathbf{xx}}^{-1}(n-1)\mathbf{x}(n)}. \tag{2.20}$$

2.2.2.2 Newton-Method Based Derivation of the Recursive Least Squares Algorithm

More generally, the Newton-based iterative estimation offers a method that is not restricted to a particular correlation estimation approach. The cost function in Eq. (2.14) can be approximated by the Taylor series [see Eq. (A.4)] and accordingly, its roots can be determined using the Newton method. This reads in the multidimensional case [5]

$$\widehat{\mathbf{h}}_q(n) = \widehat{\mathbf{h}}_q(n-1) - (\nabla_{\widehat{\mathbf{h}}_q}\nabla_{\widehat{\mathbf{h}}_q}^{\mathrm{T}} J(\widehat{\mathbf{h}}(n-1)))^{-1}\nabla_{\widehat{\mathbf{h}}_q} J(\widehat{\mathbf{h}}_q(n-1)). \tag{2.21}$$

Using Eqs. (2.10) and (2.11) we derive for the Hessian matrix

$$\nabla_{\widehat{\mathbf{h}}}\nabla_{\widehat{\mathbf{h}}}^{\mathrm{T}} J(\widehat{\mathbf{h}}(n-1)) = \mathbf{R}_{\mathbf{xx}}(n). \tag{2.22}$$

Finally, we obtain

$$\hat{\mathbf{h}}_q(n) = \hat{\mathbf{h}}_q(n-1) + \mathbf{R}_{\mathbf{xx}}^{-1}(n)\mathbf{x}(n)e_q^*(n). \tag{2.23}$$

The equivalence of the Eqs. (2.19) and (2.23) can be directly obtained by the identity

$$\mathbf{k}(n) = \mathbf{R}_{\mathbf{xx}}^{-1}(n)\mathbf{x}(n). \tag{2.24}$$

In the general MIMO case, we can write

$$\hat{\mathbf{H}}(n) = \hat{\mathbf{H}}(n-1) + \mathbf{R}_{\mathbf{xx}}^{-1}(n)\mathbf{x}(n)\mathbf{e}^{\mathrm{H}}(n) \tag{2.25}$$

with length-Q vector of the error signals

$$\mathbf{e}(n) := \mathbf{y}(n) - \hat{\mathbf{H}}^{\mathrm{H}}(n-1)\mathbf{x}(n). \tag{2.26}$$

With Eq. (2.25), it becomes apparent that the RLS algorithm takes the nonwhiteness of the input signal into account since all crosscorrelations need to be computed. This leads typically to high computational complexity. Ignoring the nonwhiteness of the input signal leads to a less accurate but efficient algorithm. This will be briefly discussed in the next section.

2.2.2.3 The Normalized Least Mean Square Algorithm

For white noise input we can approximate $\mathbf{R}_{\mathbf{xx}} \approx \|\mathbf{x}\|^2 \mathbf{I}$ leading to the simplified update equation

$$\hat{\mathbf{h}}_q(n) = \hat{\mathbf{h}}_q(n-1) + \mu \frac{e_q^*(n)}{\mathbf{x}^{\mathrm{H}}\mathbf{x}} \mathbf{x}, \tag{2.27}$$

with μ denoting a step size. This algorithm is known as the normalized least mean squares (NLMS) algorithm. Hence, the NLMS algorithm can be obtained from RLS by neglecting the nonwhiteness of the input signal [3] and it meets a trade-off between the computational complexity and taking into account all correlations.

References

1. Buchner H, Aichner R, Kellermann W (2007) TRINICON-based blind system identification with application to multiple-source localization and separation. In: Blind speech separation. Springer, pp 101–147
2. Huang Y, Benesty J, Chen J (2006) Acoustic MIMO signal processing. Signals and communication technology. Springer, New York
3. Haykin S (1991) Adaptive filter theory. Prentice Hall, Englewood Cliff
4. Buchner H (2010) A systematic approach to incorporate deterministic prior knowledge in broadband adaptive MIMO systems. In: Proceedings of 44th Asilomar conference on signals, systems, and computers, IEEE, Pacific Grove, pp 461–468
5. Buchner H, Benesty J, Gansler T, Kellermann W (2006) Robust extended multidelay filter and double-talk detector for acoustic echo cancellation. IEEE Trans Audio Speech Language Process 14(5):1633–1644
6. Kay SM (1993) Fundamentals of statistical signal processing: estimation theory. Prentice Hall PTR, Upper Saddle River
7. Hayes M (1996) Statistical digital signal processing and modeling. Wiley, New York
8. Swanson DC (2000) Signal Processing for Intelligent Sensor Systems with MATLAB. CRC Press, Boca Raton

http://www.springer.com/978-3-319-08953-9

Adaptive Identification of Acoustic Multichannel Systems
Using Sparse Representations
Helwani, K.
2015, XIV, 113 p. 39 illus., 10 illus. in color., Hardcover
ISBN: 978-3-319-08953-9