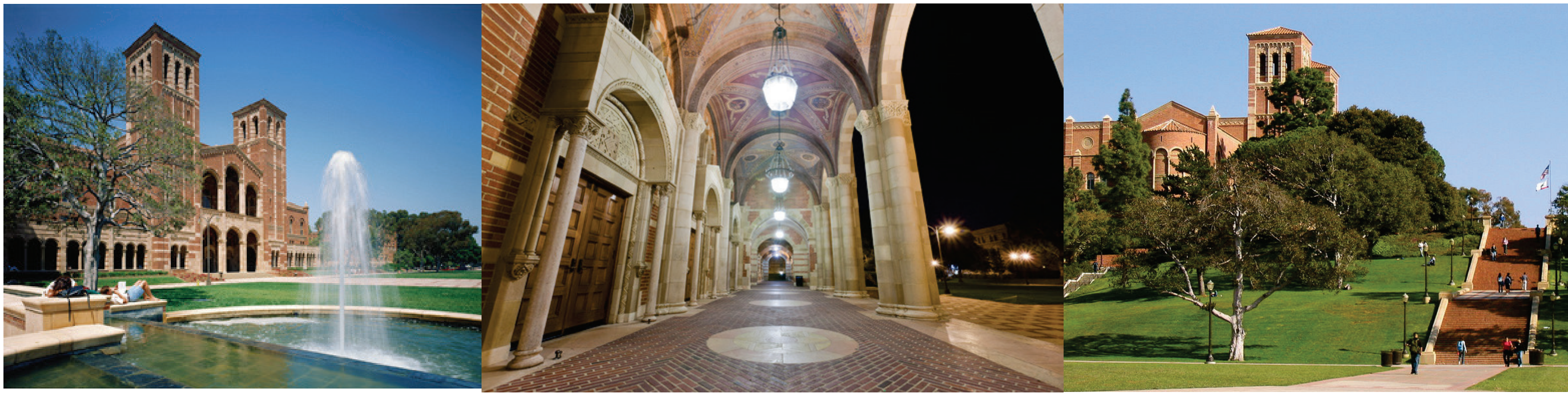

EE210A: Adaptation and Learning

Professor Ali H. Sayed

TRANSIENT PERFORMANCE OF LMS

Sections in order: 23.2-23.5, 24.1

ENERGY & VARIANCE RELATIONS

Theorem 22.4 (Weighted variance relation with independence) For adaptive filters of the form (22.2), Hermitian nonnegative-definite matrices Σ, and for data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ satisfying model (22.1) and the independence assumption (22.23), it holds that:

$$\mathsf{E}\,\|\tilde{\boldsymbol{w}}_i\|^2_{\Sigma} = \mathsf{E}\,\|\tilde{\boldsymbol{w}}_{i-1}\|^2_{\Sigma'} + \mu^2\sigma_v^2\mathsf{E}\left(\frac{\|\boldsymbol{u}_i\|^2_{\Sigma}}{g^2[\boldsymbol{u}_i]}\right)$$

$$\Sigma' = \Sigma - \mu\Sigma\mathsf{E}\left(\frac{\boldsymbol{u}_i^*\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}\right) - \mu\mathsf{E}\left(\frac{\boldsymbol{u}_i^*\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}\right)\Sigma + \mu^2\mathsf{E}\left(\frac{\|\boldsymbol{u}_i\|^2_{\Sigma}}{g^2[\boldsymbol{u}_i]}\boldsymbol{u}_i^*\boldsymbol{u}_i\right)$$

Theorem 22.5 (Mean weight error recursion) For any adaptive filter of the form (22.2), and for any data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ satisfying (22.1) and (22.23), it holds that

$$\mathsf{E}\,\tilde{\boldsymbol{w}}_i = \left[\mathrm{I} - \mu\mathsf{E}\left(\frac{\boldsymbol{u}_i^*\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}\right)\right]\mathsf{E}\,\tilde{\boldsymbol{w}}_{i-1} \qquad (22.29)$$

CHANGE OF COORDINATES

Evaluation of the moments (22.28), and the subsequent analysis, can at times be simplified if we introduce a convenient change of coordinates by appealing to the eigen-decomposition of $R_u = \mathsf{E}\, \boldsymbol{u}_i^* \boldsymbol{u}_i$. So let

$$\boxed{R_u = U \Lambda U^*} \qquad (22.30)$$

where Λ is diagonal with the eigenvalues of R_u, $\Lambda = \text{diag}\{\lambda_k\}$, and U is unitary (i.e., it satisfies $UU^* = U^*U = \mathrm{I}$). Then define the transformed quantities:

$$\boxed{\overline{\boldsymbol{w}}_i \triangleq U^* \tilde{\boldsymbol{w}}_i, \qquad \overline{\boldsymbol{u}}_i \triangleq \boldsymbol{u}_i U, \qquad \overline{\Sigma} \triangleq U^* \Sigma U} \qquad (22.31)$$

Since U is unitary, it is easy to see that

$$\|\tilde{\boldsymbol{w}}_i\|_\Sigma^2 = \|\overline{\boldsymbol{w}}_i\|_{\overline{\Sigma}}^2 \quad \text{and} \quad \|\boldsymbol{u}_i\|_\Sigma^2 = \|\overline{\boldsymbol{u}}_i\|_{\overline{\Sigma}}^2 \qquad (22.32)$$

TRANSFORMED RELATIONS

Theorem 22.6 (Transformed weighted-variance relation) For any adaptive filter of the form (22.2), any Hermitian nonnegative-definite matrix Σ, and for data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ satisfying model (22.1) and the independence assumption (22.23), it holds that:

$$\begin{aligned}
\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{\Sigma}} &= \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{\Sigma}'} + \mu^2\sigma_v^2\mathsf{E}\left[\frac{\|\overline{\boldsymbol{u}}_i\|^2_{\overline{\Sigma}}}{g^2[\overline{\boldsymbol{u}}_i]}\right] \\
\overline{\Sigma}' &= \overline{\Sigma} - \mu\overline{\Sigma}\mathsf{E}\left[\frac{\overline{\boldsymbol{u}}_i^*\overline{\boldsymbol{u}}_i}{g[\overline{\boldsymbol{u}}_i]}\right] - \mu\mathsf{E}\left[\frac{\overline{\boldsymbol{u}}_i^*\overline{\boldsymbol{u}}_i}{g[\overline{\boldsymbol{u}}_i]}\right]\overline{\Sigma} + \mu^2\mathsf{E}\left[\frac{\|\overline{\boldsymbol{u}}_i\|^2_{\overline{\Sigma}}}{g^2[\overline{\boldsymbol{u}}_i]}\overline{\boldsymbol{u}}_i^*\overline{\boldsymbol{u}}_i\right]
\end{aligned}$$

where the transformed variables $\{\overline{\boldsymbol{w}}_i, \overline{\boldsymbol{u}}_i, \overline{\Sigma}, \overline{\Sigma}'\}$ are related to the original variables $\{\tilde{\boldsymbol{w}}_i, \boldsymbol{u}_i, \Sigma, \Sigma'\}$ via (22.31) and (22.34), so that $\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{\Sigma}} = \mathsf{E}\,\|\tilde{\boldsymbol{w}}_i\|^2_{\Sigma}$.

$$\boxed{\mathsf{E}\,\overline{\boldsymbol{w}}_i = \left[\mathrm{I} - \mu\mathsf{E}\left(\frac{\overline{\boldsymbol{u}}_i^*\overline{\boldsymbol{u}}_i}{g[\overline{\boldsymbol{u}}_i]}\right)\right]\mathsf{E}\,\overline{\boldsymbol{w}}_{i-1}} \qquad (22.36)$$

LMS WITH GAUSSIAN REGRESSORS

$$\boldsymbol{w}_i = \boldsymbol{w}_{i-1} + \mu \boldsymbol{u}_i^* \boldsymbol{e}(i) \qquad (23.1)$$

for which the data normalization in (22.2) is given by

$$g[\boldsymbol{u}_i] = 1 \qquad (23.2)$$

$$\begin{aligned}
\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{\Sigma}} &= \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{\Sigma}'} + \mu^2\sigma_v^2 \mathsf{Tr}(\Lambda\overline{\Sigma}) && (23.12)\\
\overline{\Sigma}' &= \overline{\Sigma} - \mu\overline{\Sigma}\Lambda - \mu\Lambda\overline{\Sigma} + \mu^2\left[\Lambda \mathsf{Tr}(\overline{\Sigma}\Lambda) + \Lambda\overline{\Sigma}\Lambda\right] && (23.13)\\
\mathsf{E}\,\overline{\boldsymbol{w}}_i &= (\mathrm{I} - \mu\Lambda)\,\mathsf{E}\,\overline{\boldsymbol{w}}_{i-1} && (23.14)
\end{aligned}$$

$$\boxed{\bar{\sigma} \triangleq \text{diag}\{\overline{\Sigma}\} \quad \text{and} \quad \lambda \triangleq \text{diag}\{\Lambda\}} \qquad (23.15)$$

$$\boxed{\overline{\Sigma} = \text{diag}\{\bar{\sigma}\} \quad \text{and} \quad \Lambda = \text{diag}\{\lambda\}} \qquad (23.16)$$

LMS ENERGY & VARIANCE RELATIONS

$$\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{\sigma}} = \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}\overline{\sigma}} + \mu^2\sigma_v^2(\lambda^{\mathsf{T}}\overline{\sigma}) \qquad (23.20)$$

$$\overline{\sigma}' = \overline{F}\overline{\sigma} \qquad (23.21)$$

$$\overline{F} = \left(\mathrm{I} - 2\mu\Lambda + \mu^2\Lambda^2\right) + \mu^2\lambda\lambda^{\mathsf{T}} \qquad (23.22)$$

$$\mathsf{E}\,\overline{\boldsymbol{w}}_i = [\mathrm{I} - \mu\Lambda]\,\mathsf{E}\,\overline{\boldsymbol{w}}_{i-1} \qquad (23.23)$$

23.2 MEAN BEHAVIOR

The behavior of $\mathsf{E}\,\tilde{\boldsymbol{w}}_i$ can be inferred from (23.23). Thus note that since, by assumption, the initial condition is zero, $\boldsymbol{w}_{-1} = 0$, we get

$$\tilde{\boldsymbol{w}}_{-1} = w^o - \boldsymbol{w}_{-1} = w^o$$

or, equivalently,

$$\overline{\boldsymbol{w}}_{-1} = U^* w^o \stackrel{\Delta}{=} \overline{w}^o$$

The vector w^o is modeled as an unknown constant so that $\mathsf{E}\,\overline{\boldsymbol{w}}_{-1} = \overline{w}^o$. Therefore, iterating (23.23) we find that

$$\mathsf{E}\,\overline{\boldsymbol{w}}_i = (\mathrm{I} - \mu\Lambda)^{i+1}\,\overline{w}^o$$

We can now derive a condition on the step-size in order to guarantee convergence in the mean, i.e., in order to ensure that $\mathsf{E}\,\overline{\boldsymbol{w}}_i \to 0$ as $i \to \infty$, which is equivalent to $\mathsf{E}\,\tilde{\boldsymbol{w}}_i \to 0$. Indeed, since $\mathrm{I} - \mu\Lambda$ is a diagonal matrix, the condition on μ is easily seen to be:

$$|1 - \mu\lambda_k| < 1 \quad \text{for } k = 1, 2, \ldots, M$$

MEAN BEHAVIOR

where the $\{\lambda_k\}$ are the entries of Λ (i.e., the eigenvalues of R_u). In other words, μ should satisfy

$$\boxed{\mu < 2/\lambda_{\max}} \qquad (23.24)$$

where $\lambda_{\max}$ is the largest eigenvalue of R_u. For such step-sizes, it follows that

$$\boxed{\lim_{i\to\infty} \mathsf{E}\,\boldsymbol{w}_i \;=\; w^o}$$

and we say that the LMS filter is convergent in the mean and, hence, asymptotically unbiased.

23.3 MEAN-SQUARE BEHAVIOR

The study of the mean-square behavior of the filter is more demanding and also more interesting. We start by noting that the desired quantity $\mathsf{E}\,\|\tilde{\boldsymbol{w}}_i\|^2$, which is also equal to $\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2$, can be obtained from the variance recursion (23.20) if $\overline{\Sigma}$ is chosen as $\overline{\Sigma} = \mathrm{I}$ (or, equivalently, $\Sigma = \mathrm{I}$ in view of (22.31)). This corresponds to choosing $\overline{\sigma}$ as the column vector with unit entries, i.e.,

$$\overline{\sigma} \;=\; \boxed{\mathrm{col}\{1, 1, \ldots, 1, 1\} \;\overset{\Delta}{=}\; q} \tag{23.25}$$

which we denote by q. Then (23.20) gives

$$\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2 = \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}q} \;+\; \mu^2\sigma_v^2\left(\lambda^{\mathsf{T}} q\right) \tag{23.26}$$

MEAN-SQUARE BEHAVIOR

$$\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2 = \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}q} \;+\; \mu^2\sigma_v^2\left(\lambda^{\mathsf{T}} q\right) \qquad (23.26)$$

This recursion shows that in order to evaluate $\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2$ we need to know $\mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}q}$, with a weighting matrix whose diagonal entries are $\overline{F}q$. Now the quantity $\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}q}$ can be inferred from the variance relation (23.20) by writing it for the choice $\overline{\sigma} = \overline{F}q$, namely,

$$\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}q} = \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}^2 q} \;+\; \mu^2\sigma_v^2\left(\lambda^{\mathsf{T}} \overline{F} q\right) \qquad (23.27)$$

We now find from (23.27) that in order to evaluate the term

$$\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}q}$$

we need to know

$$\mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}^2 q}$$

with a weighting matrix defined by the vector $\overline{F}^2 q$. This term can again be inferred from (23.20) by writing it for the choice $\overline{\sigma} = \overline{F}^2 q$:

MEAN-SQUARE BEHAVIOR

$$\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^2 q} = \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}^3 q} \;+\; \mu^2\sigma_v^2\left(\lambda^{\mathsf{T}}\overline{F}^2 q\right) \tag{23.28}$$

and a new term with weighting matrix determined by $\overline{F}^3 q$ appears. The natural question is whether this procedure terminates, and whether weighting matrices that correspond to increasing powers of $\overline{F}$ keep coming up? The procedure does terminate. This is because once we write (23.20) for the choice $\overline{\sigma} = \overline{F}^{(M-1)} q$ we get:

$$\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^{(M-1)} q} = \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}^M q} \;+\; \mu^2\sigma_v^2\left(\lambda^{\mathsf{T}}\overline{F}^{(M-1)} q\right) \tag{23.29}$$

where the weighting matrix on the right-hand side is now $\overline{F}^M q$. However, we do not need to write (23.20) for the choice $\overline{\sigma} = \overline{F}^M q$ in order to determine $\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^M q}$. This is because this last term can be deduced from the already available weighted factors:

$$\left\{ \mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2,\; \mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F} q},\; \mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^2 q},\; \ldots,\; \mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^{M-1} q} \right\} \tag{23.30}$$

CAYLEY-HAMILTON

This fact can be seen as follows. With any matrix $\overline{F}$ we associate its characteristic polynomial, defined by

$$p(x) = \det(x\mathrm{I} - \overline{F})$$

It is an $M-$th order polynomial in x,

$$p(x) = x^M + p_{M-1}x^{M-1} + p_{M-2}x^{M-2} + \ldots + p_1 x + p_0$$

with coefficients $\{p_k, p_M = 1\}$ and whose roots coincide with the eigenvalues of $\overline{F}$. Now a famous result in matrix theory, known as the Cayley-Hamilton theorem, states that every matrix satisfies its characteristic equation, i.e., $p(\overline{F}) = 0$. In other words, $\overline{F}$ satisfies

$$\overline{F}^M + p_{M-1}\overline{F}^{M-1} + p_{M-2}\overline{F}^{M-2} + \ldots + p_1\overline{F} + p_0\mathrm{I}_M = 0$$

which means that the $M-$th power of $\overline{F}$ can be expressed as a linear combination of its lower order powers.

CAYLEY-HAMILTON

Using this fact we can write

$$\begin{aligned}\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^M q} &= \mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{(-p_{M-1}\overline{F}^{M-1}-p_{M-2}\overline{F}^{M-2}-\ldots-p_1\overline{F}-p_0\mathrm{I}_M)q} \\ &= -p_0\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2 \;-\; p_1\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}q} \;-\; \ldots \;-\; p_{M-1}\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^{(M-1)}q}\end{aligned}$$

That is,

$$\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^M q} = \sum_{k=0}^{M-1} -p_k\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^k q} \qquad (23.31)$$

which expresses $\mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^M q}$ as a linear combination of the terms in (23.30), as desired. We can collect the above results into a single self-contained recursion by writing (23.26)–(23.31) as:

STATE-SPACE MODEL

$$
\underbrace{\begin{bmatrix} \mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2 \\ \mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}q} \\ \mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^2 q} \\ \vdots \\ \mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^{(M-2)} q} \\ \mathsf{E}\,\|\overline{\boldsymbol{w}}_i\|^2_{\overline{F}^{(M-1)} q} \end{bmatrix}}_{\triangleq\, \mathcal{W}_i}
=
\underbrace{\begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ 0 & 0 & 0 & 1 & \\ \vdots & & & & \\ 0 & 0 & 0 & & 1 \\ -p_0 & -p_1 & -p_2 & \ldots & -p_{M-1} \end{bmatrix}}_{\triangleq\, \mathcal{F}}
\begin{bmatrix} \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2 \\ \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}q} \\ \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}^2 q} \\ \vdots \\ \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}^{(M-2)} q} \\ \mathsf{E}\,\|\overline{\boldsymbol{w}}_{i-1}\|^2_{\overline{F}^{(M-1)} q} \end{bmatrix}
$$

$$
+\ \mu^2\sigma_v^2 \underbrace{\begin{bmatrix} \lambda^{\mathsf{T}} q \\ \lambda^{\mathsf{T}}\overline{F} q \\ \lambda^{\mathsf{T}}\overline{F}^2 q \\ \vdots \\ \lambda^{\mathsf{T}}\overline{F}^{M-1} q \end{bmatrix}}_{\triangleq\, \mathcal{Y}}
$$

STATE-SPACE MODEL

If we introduce the vector and matrix quantities $\{\mathcal{W}_i, \mathcal{F}, \mathcal{Y}\}$ indicated above, then this recursion can be rewritten more compactly as

$$\boxed{\mathcal{W}_i = \mathcal{F}\mathcal{W}_{i-1} + \mu^2\sigma_v^2\mathcal{Y}} \tag{23.32}$$

This is a first-order recursion with a constant coefficient matrix $\mathcal{F}$. In the language of linear system theory, a recursion of the form (23.32) is called a *state-space* recursion with the vector $\mathcal{W}_i$ denoting the state vector. We therefore find that the mean-square behavior of LMS is described by the $M-$dimensional state-space recursion (23.32) with coefficient matrix $\mathcal{F}$. To be more precise, the transient behavior of LMS is described by the combination of both (23.32) and recursion (23.23) for the mean weight-error vector. These two recursions can be grouped together into a single $2M$-dimensional state-space model with a block diagonal coefficient matrix as follows

$$\begin{bmatrix} \mathsf{E}\,\overline{\boldsymbol{w}}_i \\ \mathcal{W}_i \end{bmatrix} = \begin{bmatrix} \mathrm{I}-\mu\Lambda & \\ & \mathcal{F} \end{bmatrix} \begin{bmatrix} \mathsf{E}\,\overline{\boldsymbol{w}}_{i-1} \\ \mathcal{W}_{i-1} \end{bmatrix} + \mu^2\sigma_v^2 \begin{bmatrix} 0 \\ \mathcal{Y} \end{bmatrix}$$

COMPANION MATRIX

The matrix $\mathcal{F}$ in (23.32) has a special structure in this case; it is a companion matrix (namely, a matrix with ones on the upper diagonal and the negatives of the coefficients of $p(x)$ in the last row). A well-known property of such matrices is that their eigenvalues coincide with the roots of $p(x) = 0$, which in turn are the eigenvalues of $\overline{F}$ from (23.22), i.e.,

$$\{\text{ eigenvalues of } \mathcal{F} \} \; = \; \{ \text{ roots of } \; p(x) \} \; = \; \{ \text{ eigenvalues of } \overline{F} \}$$

Recursion (23.32) shows that the transient behavior of LMS is the combined result of the time evolutions of the M variables in (23.30), which are the entries of $\mathcal{W}_i$. For this reason, these variables are called *state* variables; since they determine the state of the filter at any particular time instant.

MEAN-SQUARE BEHAVIOR

Lemma 23.1 (Mean-square behavior of complex LMS) Consider the LMS recursion (23.1) and assume the data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ satisfy model (22.1) and the independence assumption (22.23). Assume further that the regressor sequence is circular Gaussian. Then the mean and mean-square behaviors of the filter are characterized by (23.23) and (23.32), namely, by the recursion

$$\begin{bmatrix} \mathsf{E}\,\overline{\boldsymbol{w}}_i \\ \mathcal{W}_i \end{bmatrix} = \begin{bmatrix} \mathrm{I} - \mu\Lambda & \\ & \mathcal{F} \end{bmatrix} \begin{bmatrix} \mathsf{E}\,\overline{\boldsymbol{w}}_{i-1} \\ \mathcal{W}_{i-1} \end{bmatrix} + \mu^2\sigma_v^2 \begin{bmatrix} 0 \\ \mathcal{Y} \end{bmatrix}$$

23.4 MEAN-SQUARE STABILITY

The LMS filter will be said to be mean-square stable if, and only if, the state vector $\mathcal{W}_i$ remains bounded and tends to a steady-state value, regardless of the initial condition $\mathcal{W}_{-1}$. A necessary and sufficient condition for this to hold can be found as follows. For any state-space model of the form

$$x_i = Cx_{i-1} + b$$

a well-known result from linear system theory states that the sequence $\{x_j\}$ remains bounded, regardless of the initial state vector $\{x_{-1}\}$, and that the sequence $\{x_j\}$ will tend to a finite steady-state value as well, if, and only if, all the eigenvalues of C lie inside the open unit disc. Applying this result to (23.32), we conclude that the LMS filter will be mean-square stable if, and only if, all eigenvalues of $\mathcal{F}$ are inside the unit disc or, equivalently, the eigenvalues of $\overline{F}$ from (23.22) should satisfy

$$-1 < \lambda\left(\overline{F}\right) < 1 \qquad (23.33)$$

That is, we need $\overline{F}$ to be a stable matrix; here we are writing $\lambda(\overline{F})$ to refer to the eigenvalues (spectrum) of $\overline{F}$.

BOUNDS FOR STABILITY

The lower bound on $\lambda(\overline{F})$ is automatically satisfied because $\overline{F}$ is at least nonnegative-definite (and, hence, its eigenvalues are all nonnegative). This fact can be seen by writing $\overline{F}$ from (23.22) as the sum of two nonnegative-definite matrices:

$$\overline{F} \;=\; \left(\mathrm{I} - 2\mu\Lambda + \mu^2\Lambda^2\right) + \mu^2\lambda\lambda^{\mathsf{T}} \;=\; \left(\mathrm{I} - \mu\Lambda\right)^2 + \mu^2\lambda\lambda^{\mathsf{T}} \tag{23.34}$$

To find a condition on μ for the upper bound on the eigenvalues of $\overline{F}$ to be satisfied, we start by expressing $\overline{F}$ as

$$\overline{F} = \mathrm{I} - \mu A + \mu^2 B \tag{23.35}$$

where, in this case, the matrices A and B are both positive-definite and given by

$$A \;\triangleq\; 2\Lambda, \qquad B \;\triangleq\; \Lambda^2 + \lambda\lambda^{\mathsf{T}} \tag{23.36}$$

It is shown in Prob. V.3 that for nonnegative-definite matrices $\overline{F}$ of the form (23.35), its eigenvalues will be upper bounded by one if, and only if, the parameter μ satisfies

$$0 < \mu < 1/\lambda_{\max}\left(A^{-1}B\right) \tag{23.37}$$

BOUNDS FOR STABILITY

It is also shown in that problem that the eigenvalues of $A^{-1}B$ are real and positive.

Let

$$\eta^o = 1/\lambda_{\max}(A^{-1}B)$$

The bound (23.37) on μ is simply the smallest positive scalar η that makes the matrix $\mathrm{I} - \eta A^{-1}B$ singular, i.e., it is the smallest η such that

$$\det\left(\mathrm{I} - \eta A^{-1}B\right) = 0 \tag{23.38}$$

Combining (23.37) with (23.24) we find that the condition on μ for the filter to converge in both the mean and mean-square senses is

$$0 < \mu < \min\{2/\lambda_{\max}, \eta^o\} \tag{23.39}$$

It turns out that we can be more explicit and show that $\eta^o < 2/\lambda_{\max}$ so that the upper bound on μ is ultimately determined by η^o alone. Indeed, using the definitions (23.36) for A and B we have

BOUNDS FOR STABILITY

$$\begin{aligned}\det\left(\mathrm{I}-\eta A^{-1}B\right) &= \det\left(\mathrm{I}-\eta\frac{\Lambda^{-1}}{2}\left[\Lambda^2+\lambda\lambda^{\mathsf{T}}\right]\right)\\ &= \det\left(\mathrm{I}-\frac{\eta}{2}\left[\Lambda+q\lambda^{\mathsf{T}}\right]\right)\end{aligned}$$

We are interested in the smallest value of η that makes $\mathrm{I}-\eta A^{-1}B$ singular. Actually, in view of condition (23.39) on μ, we are only interested in those values of η that lie within the open interval $(0, 2/\lambda_{\max})$. If any such η can be found, then $\eta^o < 2/\lambda_{\max}$. Since over the interval $\eta \in (0, 2/\lambda_{\max})$ the matrix $(\mathrm{I}-\frac{\eta}{2}\Lambda)$ is invertible, we can write

$$\begin{aligned}\det\left(\mathrm{I}-\eta A^{-1}B\right) &= \det\left(\mathrm{I}-\frac{\eta}{2}\Lambda\right)\cdot\det\left(\mathrm{I}-\left[\mathrm{I}-\frac{\eta}{2}\Lambda\right]^{-1}\frac{\eta}{2}q\lambda^{\mathsf{T}}\right)\\ &= \left(1-\lambda^{\mathsf{T}}\left[2\eta^{-1}\mathrm{I}-\Lambda\right]^{-1}q\right)\cdot\det\left(\mathrm{I}-\frac{\eta}{2}\Lambda\right) \qquad (23.40)\end{aligned}$$

BOUNDS FOR STABILITY

where in the last step we used the determinant identity

$$\det(\mathrm{I} - XY) \;=\; \det(\mathrm{I} - YX)$$

for any matrices $\{X, Y\}$ of compatible dimensions, so that when X is a column and Y is a row we have

$$\det(\mathrm{I} - xy^{\mathsf{T}}) \;=\; 1 - y^{\mathsf{T}}x$$

The values of $\eta \in (0, 2/\lambda_{\max})$ that result in $\det\left(\mathrm{I} - \eta A^{-1}B\right) = 0$ should therefore satisfy

$$\lambda^{\mathsf{T}}\left(2\eta^{-1}\mathrm{I} - \Lambda\right)^{-1} q = 1$$

i.e.,

$$\sum_{k=1}^{M} \frac{\lambda_k \eta}{2 - \lambda_k \eta} = 1 \qquad (23.41)$$

BOUNDS FOR STABILITY

Introduce the function

$$f(\eta) \triangleq \sum_{k=1}^{M} \frac{\lambda_k \eta}{2 - \lambda_k \eta} \tag{23.42}$$

and observe that $f(0) = 0$ and $f(\eta)$ is monotonically increasing between 0 and $2/\lambda_{\max}$; this latter claim follows by noting that the derivative of $f(\cdot)$ with respect to η is positive,

$$\frac{\mathrm{d}f}{\mathrm{d}\eta} = \sum_{k=1}^{M} \frac{2\lambda_k}{(2 - \lambda_k \eta)^2} > 0 \qquad \text{for} \quad \eta \in [0, 2/\lambda_{\max})$$

Observe further that $f(\eta)$ has a singularity (i.e., a pole) at $\eta = 2/\lambda_{\max}$, so that $f(\eta) \to \infty$ as $\eta \to 2/\lambda_{\max}$ — Fig. 23.1.

FUNCTION BEHAVIOR

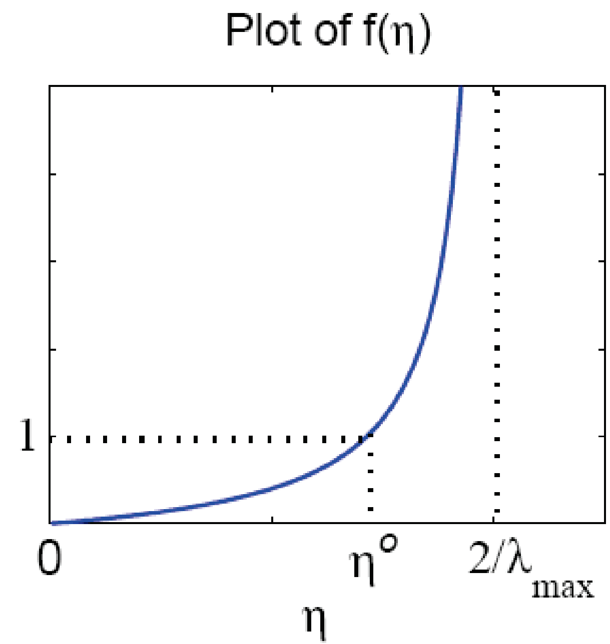

FIGURE 23.1 Behavior of the function $f(\eta)$ defined by (23.42) over the semi-open interval $[0, 2/\lambda_{\max})$.

STABILITY BOUNDS

Therefore, we conclude that there exists a unique positive η within the interval $(0, 2/\lambda_{\max})$ where the function $f(\eta)$ crosses one, i.e., for which

$$\sum_{k=1}^{M} \frac{\lambda_k \eta^o}{2 - \lambda_k \eta^o} = 1$$

This value of η is the smallest η that makes $\mathrm{I} - \eta A^{-1} B$ singular and, therefore, it coincides with the desired η^o and is smaller than $2/\lambda_{\max}$. In conclusion, we find that condition (23.39) is equivalent to requiring μ to satisfy

$$\sum_{k=1}^{M} \frac{\lambda_k \mu}{2 - \lambda_k \mu} < 1$$

When this happens, the LMS filter will be mean-square stable.

COMPLEX LMS

Theorem 23.1 (Stability of complex LMS) Consider the LMS recursion (23.1) and assume the data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ satisfy model (22.1) and the independence assumption (22.23). Assume further that the regressor sequence is circular Gaussian. Then the LMS filter is mean-square stable (i.e., the state vector $\mathcal{W}_i$ remains bounded and tends to a finite steady-state value) if, and only if, the positive step-size μ satisfies

$$\sum_{k=1}^{M} \frac{\lambda_k \mu}{2 - \lambda_k \mu} < 1$$

where the $\{\lambda_k\}$ are the eigenvalues of R_u. The above condition on μ also guarantees convergence in the mean, i.e., $\mathsf{E}\,\boldsymbol{w}_i \to w^o$.

REAL LMS

Theorem 23.2 (Mean-square behavior of real LMS) Consider the LMS recursion (23.1) and assume the data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ satisfy model (22.1) and the independence assumption (22.23). Assume further that the regressor sequence is real-valued Gaussian. Then the LMS filter is mean-square stable (i.e., the state vector $\mathcal{W}_i$ remains bounded and tends to a finite steady-state value) if, and only if, the step-size μ satisfies

$$\frac{1}{2}\sum_{k=1}^{M}\frac{\lambda_k\mu}{1-\lambda_k\mu} < 1$$

where the $\{\lambda_k\}$ are the eigenvalues of R_u. The above condition on μ also guarantees convergence in the mean, i.e., $\mathsf{E}\,\boldsymbol{w}_i \rightarrow w^o$. Moreover, the mean and mean-square behaviors of LMS are characterized by recursions (23.23) and (23.32), namely,

$$\begin{bmatrix} \mathsf{E}\,\overline{\boldsymbol{w}}_i \\ \mathcal{W}_i \end{bmatrix} = \begin{bmatrix} I-\mu\Lambda & \\ & \mathcal{F} \end{bmatrix}\begin{bmatrix} \mathsf{E}\,\overline{\boldsymbol{w}}_{i-1} \\ \mathcal{W}_{i-1} \end{bmatrix} + \mu^2\sigma_v^2\begin{bmatrix} 0 \\ \mathcal{Y} \end{bmatrix}$$

(or, more generally, by (23.48) further ahead for other choices of $\overline{\sigma}$ in the mean-square case). The coefficients $\{p_k\}$ that define the matrix $\mathcal{F}$ are now obtained from the characteristic polynomial of the matrix $\overline{F}$ in (23.45).

23.5 STEADY-STATE PERFORMANCE

In this section, we shall re-examine the excess mean-square error,

$$\text{EMSE} \triangleq \lim_{i\to\infty} \mathsf{E}\,|\boldsymbol{e}_a(i)|^2$$

as well as study the so-called mean-square deviation of the filter, which is defined as

$$\text{MSD} \triangleq \lim_{i\to\infty} \mathsf{E}\,\|\tilde{\boldsymbol{w}}_i\|^2$$

To begin with, let us first note that if model (23.32) is stable, then it will remain stable if q is replaced by any other choice for $\overline{\sigma}$. Indeed, it is straightforward to see from the arguments that led to (23.32) that had we started with any other choice for $\overline{\sigma}$, a similar state-space recursion would have resulted with the same coefficient matrix $\mathcal{F}$, namely

STEADY-STATE PERFORMANCE

$$\underbrace{\begin{bmatrix} \mathsf{E}\,\|\overline{w}_i\|^2_{\overline{\sigma}} \\ \mathsf{E}\,\|\overline{w}_i\|^2_{\overline{F}\overline{\sigma}} \\ \mathsf{E}\,\|\overline{w}_i\|^2_{\overline{F}^2\overline{\sigma}} \\ \vdots \\ \mathsf{E}\,\|\overline{w}_i\|^2_{\overline{F}^{(M-2)}\overline{\sigma}} \\ \mathsf{E}\,\|\overline{w}_i\|^2_{\overline{F}^{(M-1)}\overline{\sigma}} \end{bmatrix}}_{\triangleq\, \mathcal{W}_i} = \underbrace{\begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ 0 & 0 & 0 & 1 & \\ \vdots & & & & \\ 0 & 0 & 0 & & 1 \\ -p_0 & -p_1 & -p_2 & \cdots & -p_{M-1} \end{bmatrix}}_{\triangleq\, \mathcal{F}} \begin{bmatrix} \mathsf{E}\,\|\overline{w}_{i-1}\|^2_{\overline{\sigma}} \\ \mathsf{E}\,\|\overline{w}_{i-1}\|^2_{\overline{F}\overline{\sigma}} \\ \mathsf{E}\,\|\overline{w}_{i-1}\|^2_{\overline{F}^2\overline{\sigma}} \\ \vdots \\ \mathsf{E}\,\|\overline{w}_{i-1}\|^2_{\overline{F}^{(M-2)}\overline{\sigma}} \\ \mathsf{E}\,\|\overline{w}_{i-1}\|^2_{\overline{F}^{(M-1)}\overline{\sigma}} \end{bmatrix}$$

$$+\,\mu^2\sigma_v^2 \underbrace{\begin{bmatrix} \lambda^{\mathsf{T}}\overline{\sigma} \\ \lambda^{\mathsf{T}}\overline{F}\overline{\sigma} \\ \lambda^{\mathsf{T}}\overline{F}^2\overline{\sigma} \\ \vdots \\ \lambda^{\mathsf{T}}\overline{F}^{M-1}\overline{\sigma} \end{bmatrix}}_{\triangleq\, \mathcal{Y}}$$

with q replaced by $\overline{\sigma}$. In other words, it would still hold that

$$\boxed{\mathcal{W}_i = \mathcal{F}\mathcal{W}_{i-1} + \mu^2\sigma_v^2\mathcal{Y}} \qquad (23.48)$$

STEADY-STATE PERFORMANCE

With this issue settled, we can now explain how the freedom in selecting $\overline{\sigma}$ can be useful. To see this, consider the setting of Thm. 23.1 and assume the step-size μ has been chosen to guarantee filter stability. Then recursion (23.20) becomes in steady-state

$$\mathsf{E}\,\|\overline{\boldsymbol{w}}_\infty\|^2_{\overline{\sigma}} \;=\; \mathsf{E}\,\|\overline{\boldsymbol{w}}_\infty\|^2_{\overline{F}\overline{\sigma}} \;+\; \mu^2\sigma_v^2(\lambda^{\mathsf{T}}\overline{\sigma}) \tag{23.49}$$

which is equivalent to

$$\boxed{\mathsf{E}\,\|\overline{\boldsymbol{w}}_\infty\|^2_{(\mathrm{I}-\overline{F})\overline{\sigma}} \;=\; \mu^2\sigma_v^2(\lambda^{\mathsf{T}}\overline{\sigma})} \tag{23.50}$$

This expression can be used to examine the mean-square performance of LMS in an interesting way.

MSD PERFORMANCE

For example, in order to evaluate the MSD of LMS we would need to evaluate the expectation $\mathsf{E}\,\|\overline{w}_\infty\|^2$. Thus assume that we select the weighting vector $\overline{\sigma}$ in (23.50) as the solution to the linear system of equations $(\mathrm{I}-\overline{F})\overline{\sigma}_{\rm msd} = q$, i.e., as $\overline{\sigma}_{\rm msd} = (\mathrm{I}-\overline{F})^{-1}q$. In this way, the weighting vector that appears in (23.50) will become q. Then the left-hand side of (23.50) will coincide with the filter MSD and, therefore, we would be able to conclude that

$$\boxed{\mathsf{MSD} \;=\; \mu^2\sigma_v^2\lambda^{\mathsf{T}}(\mathrm{I}-\overline{F})^{-1}q} \qquad (23.51)$$

A more explicit expression for the MSD can be found by evaluating the product $\lambda^{\mathsf{T}}(\mathrm{I}-\overline{F})^{-1}q$. Using expression (23.22) for $\overline{F}$ we have

$$\mathrm{I}-\overline{F} \;=\; 2\mu\Lambda - \mu^2\Lambda^2 - \mu^2\lambda\lambda^{\mathsf{T}} \;\stackrel{\Delta}{=}\; D - \mu^2\lambda\lambda^{\mathsf{T}}$$

where we introduced, for convenience, the matrix

$$D = 2\mu\Lambda - \mu^2\Lambda^2 \qquad (23.52)$$

MSD PERFORMANCE

Then, using the matrix inversion formula (5.4),

$$\begin{aligned}
\lambda^{\mathsf{T}}(\mathrm{I}-\overline{F})^{-1}q &= \lambda^{\mathsf{T}}(D-\mu^2\lambda\lambda^{\mathsf{T}})^{-1}q \\
&= \lambda^{\mathsf{T}}\left(D^{-1}+\frac{\mu^2}{1-\mu^2\lambda^{\mathsf{T}}D^{-1}\lambda}D^{-1}\lambda\lambda^{\mathsf{T}}D^{-1}\right)q \\
&= \frac{\lambda^{\mathsf{T}}D^{-1}q}{1-\mu^2\lambda^{\mathsf{T}}D^{-1}\lambda} \qquad (23.53)
\end{aligned}$$

Substituting (23.53) into (23.51), we get

$$\boxed{\mathsf{MSD}=\frac{\mu\sigma_v^2\sum_{k=1}^{M}\frac{1}{2-\mu\lambda_k}}{1-\mu\sum_{k=1}^{M}\frac{\lambda_k}{2-\mu\lambda_k}}} \qquad (23.54)$$

EMSE PERFORMANCE

In a similar vein, we can evaluate the EMSE of LMS. Thus recall from (22.24) that $\boldsymbol{u}_i$ and $\tilde{\boldsymbol{w}}_{i-1}$ are independent random variables. Then it follows that

$$\begin{aligned}
\mathsf{E}\,|\boldsymbol{e}_a(i)|^2 &= \mathsf{E}\,\tilde{\boldsymbol{w}}_{i-1}^*\boldsymbol{u}_i^*\boldsymbol{u}_i\tilde{\boldsymbol{w}}_{i-1} \\
&= \mathsf{E}\left[\mathsf{E}\left(\tilde{\boldsymbol{w}}_{i-1}^*\boldsymbol{u}_i^*\boldsymbol{u}_i\tilde{\boldsymbol{w}}_{i-1} \mid \tilde{\boldsymbol{w}}_{i-1}\right)\right] \\
&= \mathsf{E}\left[\tilde{\boldsymbol{w}}_{i-1}^*\ \mathsf{E}\left(\boldsymbol{u}_i^*\boldsymbol{u}_i \mid \tilde{\boldsymbol{w}}_{i-1}\right)\tilde{\boldsymbol{w}}_{i-1}\right] \\
&= \mathsf{E}\left(\tilde{\boldsymbol{w}}_{i-1}^* R_u \tilde{\boldsymbol{w}}_{i-1}\right) \\
&= \|\tilde{\boldsymbol{w}}_{i-1}\|_{R_u}^2 \\
&= \|\overline{\boldsymbol{w}}_{i-1}\|_{\lambda}^2 \qquad (23.55)
\end{aligned}$$

In other words, in order to determine the EMSE we need to evaluate $\mathsf{E}\,\|\overline{\boldsymbol{w}}_\infty\|_\lambda^2$, with weighting factor $\lambda = \text{diag}\{\Lambda\}$. Therefore, assume that we now select $\overline{\sigma}$ in (23.50) as the solution to the linear system of equations $(\mathrm{I}-\overline{F})\overline{\sigma}_{\text{emse}} = \lambda$, i.e., as $\overline{\sigma}_{\text{emse}} = (\mathrm{I}-\overline{F})^{-1}\lambda$. In this way, the weighting quantity that appears in (23.50) will become λ. Then the left-hand side of (23.50) will coincide with the filter EMSE and we would get

EMSE PERFORMANCE

$$\boxed{\text{EMSE} = \mu^2\sigma_v^2\lambda^{\mathsf{T}}(\mathrm{I}-\overline{F})^{-1}\lambda} \tag{23.56}$$

Again a more explicit expression for the EMSE can be found by evaluating the product $\lambda^{\mathsf{T}}(\mathrm{I}-\overline{F})^{-1}\lambda$ in much the same way as we did for the MSD above, leading to

$$\lambda^{\mathsf{T}}(\mathrm{I}-\overline{F})^{-1}\lambda = \frac{\lambda^{\mathsf{T}}D^{-1}\lambda}{1-\mu^2\lambda^{\mathsf{T}}D^{-1}\lambda}$$

so that

$$\boxed{\text{EMSE} = \frac{\mu\sigma_v^2\sum_{k=1}^{M}\frac{\lambda_k}{2-\mu\lambda_k}}{1-\mu\sum_{k=1}^{M}\frac{\lambda_k}{2-\mu\lambda_k}}}$$

MSD AND EMSE EXPRESSIONS

Theorem 23.3 (MSD and EMSE of LMS) Consider the LMS recursion (23.1) and assume the data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ satisfy model (22.1) and the independence assumption (22.23). Then the MSD and EMSE are given by

$$\text{EMSE} = \frac{\mu\sigma_v^2 \sum_{k=1}^{M} \frac{\lambda_k}{2-s\mu\lambda_k}}{1-\mu\sum_{k=1}^{M}\frac{\lambda_k}{2-s\mu\lambda_k}} \qquad \text{MSD} = \frac{\mu\sigma_v^2 \sum_{k=1}^{M} \frac{1}{2-s\mu\lambda_k}}{1-\mu\sum_{k=1}^{M}\frac{\lambda_k}{2-s\mu\lambda_k}}$$

where $s=1$ if the regressors are circular Gaussian and $s=2$ if the regressors are real-valued Gaussian. Moreover, the $\{\lambda_k\}$ are the eigenvalues of R_u.

24.1 MEAN AND VARIANCE RELATIONS

Thus refer again to the transformed recursion (23.7), which characterizes the transient performance of LMS. When the regressors $\boldsymbol{u}_i$ were Gaussian, we were able to evaluate the three moments below (see (23.9)–(23.11)):

$$\mathsf{E}\,\overline{\boldsymbol{u}}_i^*\overline{\boldsymbol{u}}_i, \qquad \mathsf{E}\,\|\overline{\boldsymbol{u}}_i\|^2_{\overline{\Sigma}} \qquad \text{and} \qquad \mathsf{E}\,\|\overline{\boldsymbol{u}}_i\|^2_{\overline{\Sigma}}\overline{\boldsymbol{u}}_i^*\overline{\boldsymbol{u}}_i \tag{24.1}$$

In particular, we found that $\mathsf{E}\,\overline{\boldsymbol{u}}_i^*\overline{\boldsymbol{u}}_i$ and $\mathsf{E}\,\|\overline{\boldsymbol{u}}_i\|_{\overline{\Sigma}}\overline{\boldsymbol{u}}_i^*\overline{\boldsymbol{u}}_i$ were simultaneously diagonal and that the weighting matrices $\{\overline{\Sigma}, \overline{\Sigma}'\}$ themselves could be made diagonal as well — see (23.13).

However, when the regressors $\boldsymbol{u}_i$ are non-Gaussian, it is generally not possible to express the last moment in (24.1) in closed-form any longer (as we did in (23.11); see though Prob. V.11). In addition, and more importantly perhaps, the moments $\mathsf{E}\,\overline{\boldsymbol{u}}_i^*\overline{\boldsymbol{u}}_i$ and $\mathsf{E}\,\|\overline{\boldsymbol{u}}_i\|_{\overline{\Sigma}}\overline{\boldsymbol{u}}_i^*\overline{\boldsymbol{u}}_i$ need not be simultaneously diagonal anymore. In this way, the weighting matrix $\overline{\Sigma}'$ need not be diagonal even if $\overline{\Sigma}$ is.

VEC NOTATION

Nevertheless, the transient analysis of LMS can still be pursued in much the same way as we did in the Gaussian case if we replace the diag$\{\cdot\}$ notation in (23.15) by an alternative vec$\{\cdot\}$ notation. Before doing so, we remark that since the weighting matrices $\{\overline{\Sigma}', \overline{\Sigma}\}$ are not necessarily diagonal anymore, we shall pursue our analysis by working with the original variance and mean relations (23.3), instead of the transformed variance and mean relations (23.7), namely, we shall now work with

$$\left\{\begin{array}{ccl} \mathsf{E}\,\|\tilde{\boldsymbol{w}}_i\|^2_{\Sigma} & = & \mathsf{E}\,\|\tilde{\boldsymbol{w}}_{i-1}\|^2_{\Sigma'} \;+\; \mu^2\sigma_v^2\mathsf{E}\,\|\boldsymbol{u}_i\|^2_{\Sigma} \\ \Sigma' & = & \Sigma - \mu\Sigma\mathsf{E}\,[\boldsymbol{u}_i^*\boldsymbol{u}_i] - \mu\mathsf{E}\,[\boldsymbol{u}_i^*\boldsymbol{u}_i]\,\Sigma \;+\; \mu^2\mathsf{E}\,\left[\|\boldsymbol{u}_i\|^2_{\Sigma}\boldsymbol{u}_i^*\boldsymbol{u}_i\right] \\ \mathsf{E}\,\tilde{\boldsymbol{w}}_i & = & (\mathrm{I} - \mu R_u)\,\mathsf{E}\,\tilde{\boldsymbol{w}}_{i-1} \end{array}\right. \tag{24.2}$$

Vector Notation

The diag$\{\cdot\}$ notation allowed us to replace an $M \times M$ matrix by an $M \times 1$ column vector whose entries were the diagonal entries of the matrix. More generally, we shall use the vec$\{\cdot\}$ notation to replace an $M \times M$ arbitrary matrix by an $M^2 \times 1$ column vector whose entries are formed by stacking the successive columns of the matrix on top of each other.

VEC NOTATION

We shall therefore write

$$\boxed{\sigma = \text{vec}(\Sigma) \qquad (\sigma \text{ is now } M^2 \times 1)} \tag{24.3}$$

with σ denoting the vectorized version of Σ. Likewise, we shall write r to denote the vectorized version of R_u,

$$\boxed{r = \text{vec}(R_u) \qquad (r \text{ is } M^2 \times 1)} \tag{24.4}$$

and r' to denote to the vectorized version of R_u^{T},

$$\boxed{r' = \text{vec}(R_u^{\mathsf{T}})} \tag{24.5}$$

When the regressors are real-valued, so that $R_u = R_u^{\mathsf{T}}$, the vectors $\{r, r'\}$ will coincide. However, when the regressors are complex-valued, we need to distinguish between r and r'.

VEC NOTATION

In addition, we shall use the notation $\text{vec}^{-1}\{\cdot\}$ to recover a matrix from its vec representation. Thus writing $\text{vec}^{-1}\{a\}$ for an $M^2 \times 1$ column vector a, results in an $M \times M$ matrix whose entries are obtained by unstacking the elements of a. This choice of notation is in contrast to the diag$\{\cdot\}$ operation, which is generally accepted as a two-directional operation: it maps diagonal matrices to vectors and vectors into diagonal matrices. Therefore, we shall write

$$\boxed{\Sigma = \text{vec}^{-1}\{\sigma\} \qquad \text{and} \qquad R_u = \text{vec}^{-1}\{r\}} \tag{24.6}$$

to recover $\{\Sigma, R_u\}$ from $\{\sigma, r\}$.

KRONECKER PRODUCTS

The vec$\{\cdot\}$ notation is most convenient when working with Kronecker products (see App. B.7). The Kronecker product of two matrices A and B, say of dimensions $m_a \times n_a$ and $m_b \times n_b$, respectively, is denoted by $A \otimes B$ and is defined as the $m_a m_b \times n_a n_b$ matrix

$$A \otimes B \triangleq \begin{bmatrix} a_{1,1}B & a_{1,2}B & \ldots & a_{1,n_a}B \\ a_{2,1}B & a_{2,2}B & \ldots & a_{2,n_a}B \\ \vdots & \vdots & & \vdots \\ a_{m_a,1}B & a_{m_a,2}B & \ldots & a_{m_a,n_a}B \end{bmatrix}$$

This operation has several useful properties (see Lemma B.8), but the one that most interests us here is the following. For any matrices $\{P, \Sigma, Q\}$ of compatible dimensions, it holds that

$$\boxed{\text{vec}(P\Sigma Q) = (Q^{\mathsf{T}} \otimes P)\text{vec}(\Sigma)} \qquad (24.7)$$

This property tells us how the vec of the product of three matrices is related to the vec of the center matrix.

LINEAR VECTOR RELATION

With the above notations, we can now verify that expression (24.2) for Σ' in terms of Σ still amounts to a linear relation between the corresponding vectors $\{\sigma, \sigma'\}$; just like it was the case for $\{\Sigma, \Sigma'\}$ in (23.17). Indeed, applying (24.7) to some of the terms in the expression for Σ' in (24.2) we find that

$$\begin{aligned}
\text{vec}(\Sigma \mathsf{E}\,[\boldsymbol{u}_i^*\boldsymbol{u}_i]) &= ([\mathsf{E}\,\boldsymbol{u}_i^*\boldsymbol{u}_i]^{\mathsf{T}} \otimes \mathrm{I}_M)\sigma \;=\; (R_u^{\mathsf{T}} \otimes \mathrm{I}_M)\sigma \\
\text{vec}(\mathsf{E}\,[\boldsymbol{u}_i^*\boldsymbol{u}_i]\,\Sigma) &= (\mathrm{I}_M \otimes [\mathsf{E}\,\boldsymbol{u}_i^*\boldsymbol{u}_i])\sigma \;=\; (\mathrm{I}_M \otimes R_u)\sigma \\
\text{vec}(\mathsf{E}\,\|\boldsymbol{u}_i\|_\Sigma^2\boldsymbol{u}_i^*\boldsymbol{u}_i) &= \text{vec}(\mathsf{E}\,[\boldsymbol{u}_i^*\boldsymbol{u}_i\Sigma\boldsymbol{u}_i^*\boldsymbol{u}_i]) \;=\; \mathsf{E}\left([\boldsymbol{u}_i^*\boldsymbol{u}_i]^{\mathsf{T}} \otimes [\boldsymbol{u}_i^*\boldsymbol{u}_i]\right)\sigma
\end{aligned}$$

Taking the vec of both sides of (24.2), and using the above equalities, we find that the weighting vectors $\{\sigma, \sigma'\}$ satisfy a relation similar to (23.21), albeit one that is M^2-dimensional,

$$\boxed{\sigma' = F\sigma} \tag{24.8}$$

where F is $M^2 \times M^2$ and given by

$$\boxed{F \stackrel{\Delta}{=} \mathrm{I}_{M^2} - \mu(\mathrm{I}_M \otimes R_u) - \mu(R_u^{\mathsf{T}} \otimes \mathrm{I}_M) + \mu^2\mathsf{E}\,([\boldsymbol{u}_i^*\boldsymbol{u}_i]^{\mathsf{T}} \otimes [\boldsymbol{u}_i^*\boldsymbol{u}_i])} \tag{24.9}$$

or, more compactly, in factored form:

$$F \triangleq \mathsf{E}\left[(\mathrm{I}_M - \mu \boldsymbol{u}_i^* \boldsymbol{u}_i)^{\mathsf{T}} \otimes (\mathrm{I}_M - \mu \boldsymbol{u}_i^* \boldsymbol{u}_i)\right] \tag{24.10}$$

Variance Relation

We can further rewrite recursion (24.2) for $\mathsf{E}\,\|\tilde{\boldsymbol{w}}_i\|_{\Sigma}^2$ by using the weighting vectors $\{\sigma, \sigma'\}$ instead of the matrices $\{\Sigma, \Sigma'\}$. Using (24.8) and the notation (24.6), recursion (24.2) becomes

$$\mathsf{E}\,\|\tilde{\boldsymbol{w}}_i\|^2_{\mathrm{vec}^{-1}\{\sigma\}} = \mathsf{E}\,\|\tilde{\boldsymbol{w}}_{i-1}\|^2_{\mathrm{vec}^{-1}\{F\sigma\}} + \mu^2 \sigma_v^2 (r'^{\mathsf{T}} \sigma)$$

where, for the last term, we used the fact that $\mathsf{E}\,\|\boldsymbol{u}_i\|_{\Sigma}^2 = \mathsf{Tr}(R_u \Sigma) = r'^{\mathsf{T}} \sigma$. For compactness of notation, we shall drop the $\mathrm{vec}^{-1}\{\cdot\}$ notation from the subscripts and keep the vectors, so that the above can be rewritten more compactly as

$$\boxed{\mathsf{E}\,\|\tilde{\boldsymbol{w}}_i\|_{\sigma}^2 = \mathsf{E}\,\|\tilde{\boldsymbol{w}}_{i-1}\|_{F\sigma}^2 + \mu^2 \sigma_v^2 (r'^{\mathsf{T}} \sigma)} \tag{24.11}$$

VARIANCE RELATION

The vector weighting factors $\{\sigma, F\sigma\}$ in this expression should be understood as compact representations for the actual weighting matrices $\{\text{vec}^{-1}\{\sigma\}, \text{vec}^{-1}\{F\sigma\}\}$. In other words, if σ is any column vector, the compact notation $\|x\|_\sigma^2$ denotes

$$\|x\|_\sigma^2 \stackrel{\Delta}{=} \|x\|^2_{\text{vec}^{-1}\{\sigma\}} = x^*\Sigma x, \quad \text{where} \quad \Sigma = \text{vec}^{-1}\{\sigma\}$$

In summary, starting from (24.2), we argued that the weighting matrices $\{\Sigma, \Sigma'\}$ can be vectorized, so that (24.2) can be equivalently expressed more compactly as in (24.8)–(24.10) and (24.11), namely,

$$\begin{aligned}
\mathsf{E}\,\|\tilde{\boldsymbol{w}}_i\|_\sigma^2 &= \mathsf{E}\,\|\tilde{\boldsymbol{w}}_{i-1}\|^2_{F\sigma} + \mu^2\sigma_v^2(r'^{\mathsf{T}}\sigma) && (24.12)\\
\sigma' &= F\sigma && (24.13)\\
F &= \mathrm{I}_{M^2} - \mu(\mathrm{I}_M \otimes R_u) - \mu(R_u^{\mathsf{T}} \otimes \mathrm{I}_M) + \mu^2\mathsf{E}\left([\boldsymbol{u}_i^*\boldsymbol{u}_i]^{\mathsf{T}} \otimes [\boldsymbol{u}_i^*\boldsymbol{u}_i]\right) && (24.14)\\
\mathsf{E}\,\tilde{\boldsymbol{w}}_i &= (\mathrm{I} - \mu R_u)\,\mathsf{E}\,\tilde{\boldsymbol{w}}_{i-1} && (24.15)
\end{aligned}$$

STABILITY CONDITIONS

Although the coefficient matrix in the recursion for $\mathsf{E}\,\tilde{\boldsymbol{w}}_i$ is $(\mathrm{I} - \mu R_u)$, this recursion is equivalent under the unitary transformation (22.31) to (23.23), so that the same condition on μ from (23.24) guarantees convergence in the mean, i.e.,

$$\mu < 2/\lambda_{\max} \tag{24.16}$$

guarantees $\mathsf{E}\,\boldsymbol{w}_i \rightarrow w^o$.

$$F = \mathrm{I}_{M^2} - \mu A + \mu^2 B \tag{24.20}$$

with Hermitian matrices $\{A, B\}$ given by

$$A = (\mathrm{I}_M \otimes R_u) \;+\; (R_u^{\mathsf{T}} \otimes \mathrm{I}_M), \qquad B = \mathsf{E}\left([\boldsymbol{u}_i^* \boldsymbol{u}_i]^{\mathsf{T}} \otimes [\boldsymbol{u}_i^* \boldsymbol{u}_i]\right) \tag{24.21}$$

Actually, A is positive-definite and B is nonnegative-definite. We shall assume that the distribution of the regression sequence is such that B is a finite matrix. For mean-square stability we need to find conditions on μ in order to guarantee $-1 < \lambda(F) < 1$. However, contrary to the Gaussian case in (23.34), the matrix F is no longer guaranteed to be nonnegative-definite in general — see Prob. V.4 for an example.

STABILITY CONDITION

It is shown in App. 25.A that the matrix F will be stable for values of μ in the range:

$$0 < \mu < \min\left\{ \frac{1}{\lambda_{\max}(A^{-1}B)},\ \frac{1}{\max\left\{\lambda(H) \in \mathbb{R}^{+}\right\}} \right\} \tag{24.22}$$

where the second condition is in terms of the largest positive real eigenvalue of the following block matrix,

$$\boxed{H \triangleq \begin{bmatrix} A/2 & -B/2 \\ \mathrm{I}_{M^2} & 0 \end{bmatrix}} \qquad (2M \times 2M) \tag{24.23}$$

when it exists. If H does not have any real positive eigenvalue, then the corresponding condition is removed from (24.22) and we only require $\mu < 1/\lambda_{\max}(A^{-1}B)$. Conditions (24.16) and (24.22) can be grouped together into a single condition as follows:

$$\boxed{0 \ < \ \mu \ < \min\left\{ \frac{2}{\lambda_{\max}},\ \frac{1}{\lambda_{\max}(A^{-1}B)},\ \frac{1}{\max\left\{\lambda(H) \in \mathbb{R}^{+}\right\}} \right\}} \tag{24.24}$$

MSD AND EMSE

Theorem 24.1 (Stability of LMS for non-Gaussian regressors) Assume the data $\{d(i), u_i\}$ satisfy model (22.1) and the independence assumption (22.23). The regressors need not be Gaussian. Then the LMS filter (23.1) is convergent in the mean and is mean-square stable if the step-size μ is chosen to satisfy (24.24), where the matrices A and B are defined by (24.21) and B is finite. Moreover, the transient behavior of LMS is characterized by the M^2-dimensional state-space recursion (24.17)–(24.19), and the mean-square deviation and excess mean-square error are given by

$$\text{MSD} = \mu^2\sigma_v^2 r'^{\mathsf{T}}(\mathrm{I}-F)^{-1}q, \qquad \text{EMSE} = \mu^2\sigma_v^2 r'^{\mathsf{T}}(\mathrm{I}-F)^{-1}r$$

where $r = \text{vec}(R_u)$, $r' = \text{vec}(R_u^{\mathsf{T}})$, and $q = \text{vec}(\mathrm{I})$.

MSD AND EMSE

The expressions for the MSD and EMSE in the statement of the theorem are derived in a manner similar to (23.51) and (23.56). They can be rewritten as

$$\mathsf{MSD} = \mu^2\sigma_v^2 \mathsf{E}\,\|\boldsymbol{u}_i\|^2_{(\mathrm{I}-F)^{-1}q} \quad \text{and} \quad \mathsf{EMSE} = \mu^2\sigma_v^2 \mathsf{E}\,\|\boldsymbol{u}_i\|^2_{(\mathrm{I}-F)^{-1}r}$$

or, equivalently, as

$$\boxed{\mathsf{MSD} = \mu^2\sigma_v^2 \mathsf{Tr}(R_u\Sigma_{\mathrm{msd}}) \quad \text{and} \quad \mathsf{EMSE} = \mu^2\sigma_v^2 \mathsf{Tr}(R_u\Sigma_{\mathrm{emse}})}$$

where $\{\Sigma_{\mathrm{msd}}, \Sigma_{\mathrm{emse}}\}$ are the weighting matrices that correspond to the vectors $\sigma_{\mathrm{msd}} = (\mathrm{I}-F)^{-1}q$ and $\sigma_{\mathrm{emse}} = (\mathrm{I}-F)^{-1}r$. That is,

$$\Sigma_{\mathrm{msd}} = \mathrm{vec}^{-1}(\sigma_{\mathrm{msd}}) \quad \text{and} \quad \Sigma_{\mathrm{emse}} = \mathrm{vec}^{-1}(\sigma_{\mathrm{emse}})$$

DATA NORMALIZED FILTERS

$$\boxed{w_i = w_{i-1} + \mu \frac{u_i^*}{g[u_i]} e(i), \quad i \geq 0} \tag{22.2}$$

$$\boxed{e(i) = d(i) - u_i w_{i-1}} \tag{22.3}$$

DATA NORMALIZED FILTERS

Theorem 25.2 (Stability of data-normalized filters) Consider data normalized adaptive filters of the form (22.2)–(22.3), and assume the data $\{d(i), u_i\}$ satisfy model (22.1) and the independence assumption (22.23). Then the filter is convergent in the mean and is mean-square stable for step-sizes satisfying

$$0 < \mu < \min\{2/\lambda_{\max}(P),\ 1/\lambda_{\max}(A^{-1}B),\ 1/\max\{\lambda(H) \in \mathbb{R}^+\}\}$$

where the matrices $\{A, B, P, H\}$ are defined by (25.15)–(25.17) and (25.20) and B is assumed finite. Moreover, the transient behavior of the filter is characterized by the M^2-dimensional state-space recursion $\mathcal{W}_i = \mathcal{F}\mathcal{W}_{i-1} + \mu^2\sigma_v^2\mathcal{Y}$, where $\mathcal{F}$ is the companion matrix

$$\mathcal{F} = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ 0 & 0 & 0 & 1 & \\ \vdots & & & & \\ 0 & 0 & 0 & & 1 \\ -p_0 & -p_1 & -p_2 & \ldots & -p_{M^2-1} \end{bmatrix} \quad (M^2 \times M^2)$$

with

$$p(x) \triangleq \det(x\mathrm{I} - F) = x^{M^2} + \sum_{k=0}^{M^2-1} p_k x^k$$

denoting the characteristic polynomial of F in (25.19). Also,

$$\mathcal{W}_i \triangleq \begin{bmatrix} \mathsf{E}\,\|\tilde{w}_i\|^2_\sigma \\ \mathsf{E}\,\|\tilde{w}_i\|^2_{F\sigma} \\ \mathsf{E}\,\|\tilde{w}_i\|^2_{F^2\sigma} \\ \vdots \\ \mathsf{E}\,\|\tilde{w}_i\|^2_{F^{(M^2-1)}\sigma} \end{bmatrix}, \quad [\mathcal{Y}]_k = \mathsf{E}\left[\frac{\|u_i\|^2_{F^k\sigma}}{g^2[u_i]}\right], \quad k = 0, \ldots, M^2-1$$

for any σ of interest, e.g., $\sigma = q$ or $\sigma = r$. In addition, the mean-square deviation and the excess mean-square error are given by

$$\text{MSD} = \mu^2\sigma_v^2\mathsf{E}\left[\frac{\|u_i\|^2_{(\mathrm{I}-F)^{-1}q}}{g^2[u_i]}\right], \qquad \text{EMSE} = \mu^2\sigma_v^2\mathsf{E}\left[\frac{\|u_i\|^2_{(\mathrm{I}-F)^{-1}r}}{g^2[u_i]}\right]$$

where $r = \text{vec}(R_u)$ and $q = \text{vec}(\mathrm{I})$.

COMPUTER PROJECT

Project V.1 (Transient behavior of LMS) In this project we examine the transient behavior of LMS and verify some of the results derived in the chapter for both cases of Gaussian and non-Gaussian data.

Thus consider a real-valued regression sequence $\{\boldsymbol{u}_i\}$ with covariance matrix R_u whose eigenvalue spread we set at $\rho = 5$. Let the noise variance be $\sigma_v^2 = 0.001$ and fix the filter order at $M = 5$.

(a) Generate a covariance matrix R_u whose eigenvalue spread is 5. This can be achieved, for example, by choosing R_u to be diagonal with smallest eigenvalue at 1 and largest eigenvalue at 5. Generate independent and identically distributed regression vectors $\{\boldsymbol{u}_i\}$ from a Gaussian distribution with covariance matrix R_u. For each time instant i, generate also the reference signal as follows:

$$\boldsymbol{d}(i) = \boldsymbol{u}_i w^o + \boldsymbol{v}(i)$$

where $\boldsymbol{v}(i)$ is white Gaussian noise with variance 0.001, and w^o is some arbitrary weight vector that we wish to estimate. Adjust the norm of w^o to unity.

COMPUTER PROJECT

(b) Fix initially the step-size at $\mu = 0.01$ and train LMS for 10000 iterations. Average the squared-weight-error curve, $\|\tilde{w}_i\|^2$, over 30 experiments and generate an ensemble-average curve for $\mathsf{E}\,\|\tilde{w}_i\|^2$. Use recursion (23.32) to generate the theoretical curve for $\mathsf{E}\,\|\tilde{w}_i\|^2$. Observe that since in this project we are choosing R_u to be diagonal and, hence, $R_u = \Lambda$, it holds that $\overline{w}_i = \tilde{w}_i$. Compare the simulated and theoretical curves. Use the simulated curve to estimate the MSD. Compare this value with the one predicted by theory through expression (23.54).

(c) Repeat the simulation of part (b) and generate a curve of the MSD performance as a function of the step-size. Simulate for the values

$$\mu \in \{\, 0.03,\ 0.05,\ 0.07,\ 0.08,\ 0.09,\ 0.095,\ 0.1,\ 0.125\}$$

At what value of the step-size does the filter become unstable? Now recall that according to Thm. 23.2, the LMS filter remains mean-square stable for all step-sizes that satisfy the condition

$$f(\mu) \triangleq \frac{1}{2}\sum_{k=1}^{M} \frac{\lambda_k \mu}{1-\mu\lambda_k} < 1$$

Plot the function $f(\mu)$. At what value of μ does $f(\mu)$ exceed one? Compare this theoretical value with the simulated value.

COMPUTER PROJECT

(d) Repeat the simulation of part (b) except that now the regression vectors $\{\boldsymbol{u}_i\}$ are selected as independent and identically distributed realizations of a *uniform* distribution with covariance matrix R_u. More specifically, each $\boldsymbol{u}_i$ is generated as follows. Select first random vectors $\boldsymbol{s}_i$, of the same size as $\boldsymbol{u}_i$, but with i.i.d. entries that are uniformly distributed within the interval $[-1, 1]$. In this way, each entry of $\boldsymbol{s}_i$ has variance $1/12$. Then set

$$\boldsymbol{u}_i \;=\; \sqrt{12} \cdot \boldsymbol{s}_i \cdot R_u^{1/2}$$

where, since R_u is diagonal, $R_u^{1/2}$ is the diagonal matrix with the positive square-roots of the entries of R_u. Verify analytically that the vectors $\boldsymbol{u}_i$ generated in this manner have covariance matrix R_u.

Generate also a uniform noise sequence $\boldsymbol{v}(i)$ with variance $\sigma_v^2 = 0.001$. Use the data $\{\boldsymbol{d}(i), \boldsymbol{u}_i\}$ so obtained to simulate the operation of LMS with step-size $\mu = 0.05$. Use the recursion in Prob. V.7 to generate the theoretical curve for $\mathsf{E}\,\|\tilde{\boldsymbol{w}}_i\|^2$. Compare the simulated and theoretical curves. Construct also a curve for the MSD performance of the filter for the same range of step-sizes as in part (b). At what value of the step-size does the filter become unstable? Now recall that according to Thm. 24.1, the LMS filter remains mean-square stable for step-sizes that satisfy condition (24.24), namely,

$$0 \;<\; \mu \;<\min\left\{\frac{2}{\lambda_{\max}},\; \frac{1}{\lambda_{\max}(A^{-1}B)}\,\frac{1}{\max\{\lambda(H)\in\mathbb{R}^+\}}\right\}$$

COMPUTER PROJECT

where $\{A, B\}$ are given by (24.21); for real-valued data, these matrices reduce to

$$A = (\mathrm{I}_M \otimes R_u) \; + \; (R_u \otimes \mathrm{I}_M), \qquad B = \mathsf{E}\left(\left[\boldsymbol{u}_i^\mathsf{T} \boldsymbol{u}_i\right] \otimes \left[\boldsymbol{u}_i^\mathsf{T} \boldsymbol{u}_i\right]\right)$$

and

$$H = \begin{bmatrix} A/2 & -B/2 \\ \mathrm{I}_{M^2} & 0 \end{bmatrix}$$

Estimate the matrices A and B via ensemble-averaging and evaluate the upper bound on μ for mean-square stability. Compare this result with the one obtained from the simulated MSD curve.

COMPUTER PROJECT

Project V.1 (Transient behavior of LMS) The programs that solve this project are the following.

1. partA.m This program solves parts A, B, and C. It generates data $\{\boldsymbol{u}_i, \boldsymbol{d}(i)\}$ with the desired specifications. Figure 1 shows a typical ensemble-average learning curve for the weight-error vector, when the step-size is fixed at $\mu = 0.01$. The ensemble-average curve is superimposed onto the theoretical curve. It is seen that there is a good match between theory and practice. Figure 2 plots the MSD as a function of the step-size. Observe how the filter starts to diverge for step-sizes close to 0.1. The rightmost plot in Fig. 2 shows the behavior of the function $f(\mu)$ over the same range of step-sizes. The plot shows that $f(\mu)$ exceeds one around $\mu \approx 0.092$. We thus find that the simulation results are consistent with theory.

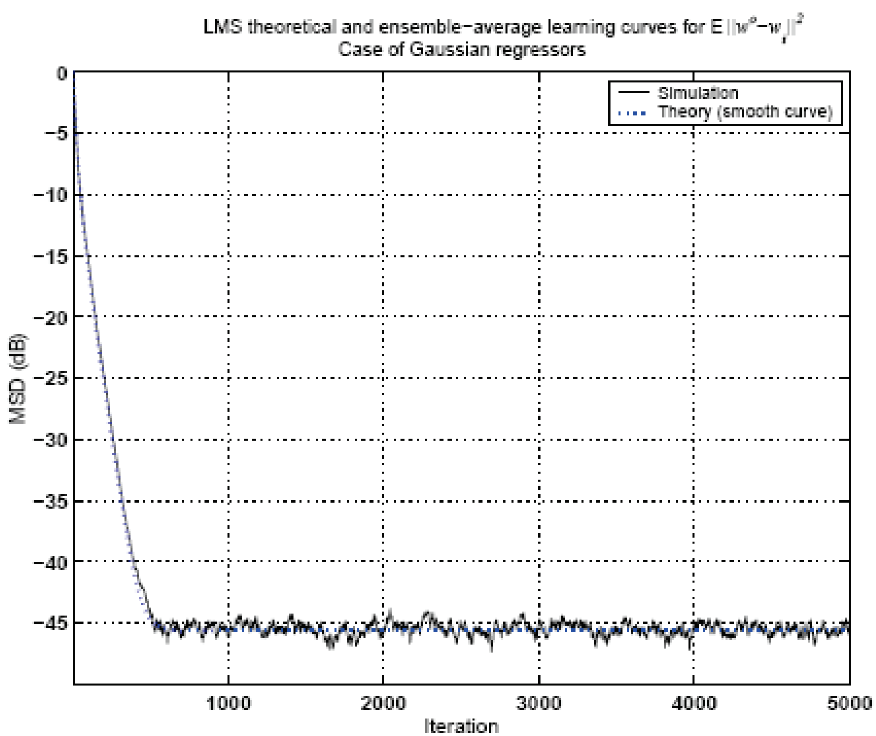

Figure V.1. Theoretical and simulated learning curves for the weight-error vector of LMS operating with a step-size $\mu = 0.01$ and using Gaussian regressors.

COMPUTER PROJECT

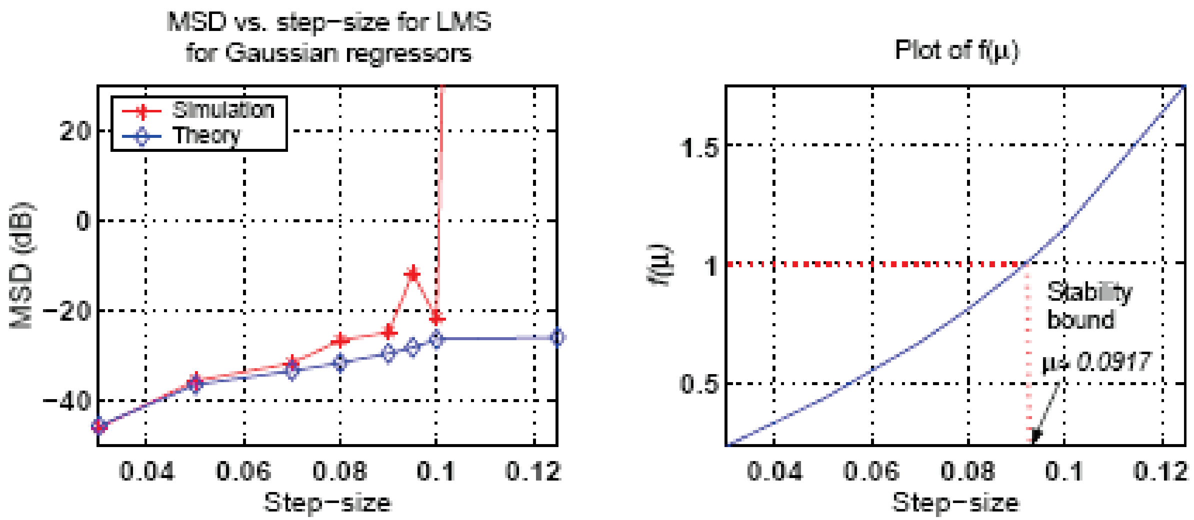

Figure V.2. Theoretical and simulated MSD of LMS as a function of the step-size for Gaussian regressors (*leftmost plot*). The rightmost plot shows the behavior of the function $f(\mu)$ over the same range of step-sizes.

2. partD.m This program generates two plots. Figure 3 shows a typical ensemble-average learning curve for the weight-error vector, when the step-size is fixed at $\mu = 0.05$. The ensemble-average curve is superimposed onto the theoretical curve that follows from Prob. V.7. It is seen that there is also a good match between theory and practice for non-Gaussian regressors.

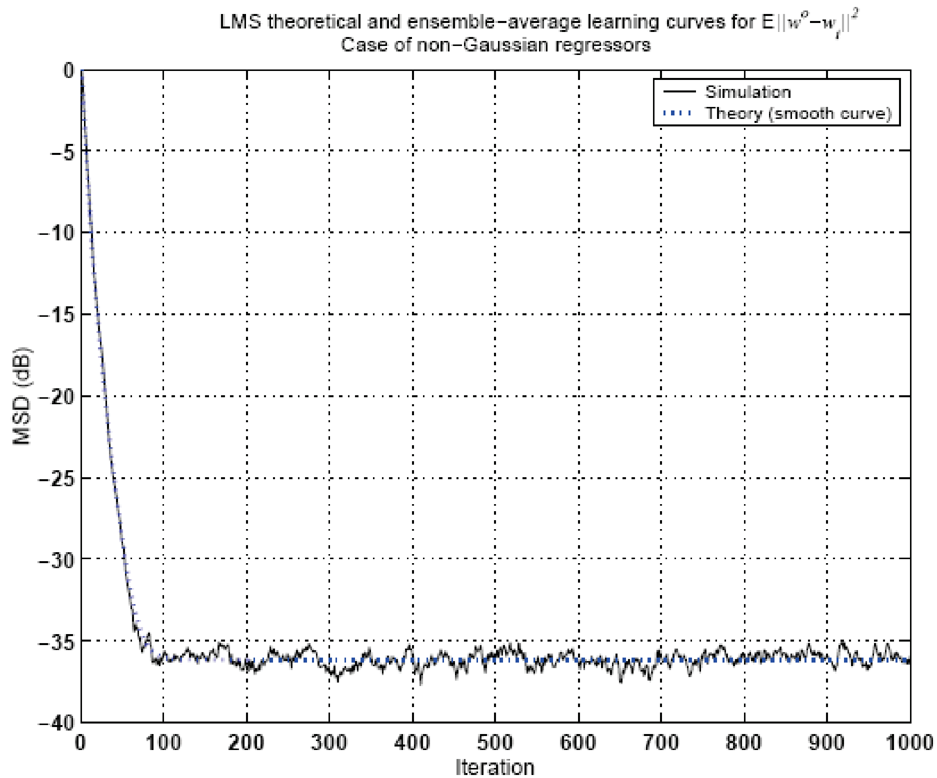

Figure V.3. Theoretical and simulated learning curves for the weight-error vector of LMS operating with a step-size $\mu = 0.05$ and using non-Gaussian regressors.

COMPUTER PROJECT

Figure 4 plots the MSD as a function of the step-size. The theoretical values are obtained by using the following expression

$$\text{MSD} = \mu^2\sigma_v^2 r'^{\mathsf{T}}(\mathrm{I}-F)^{-1}q, \qquad \text{EMSE} = \mu^2\sigma_v^2 r'^{\mathsf{T}}(\mathrm{I}-F)^{-1}r$$

where $\{r, r', q, F\}$ are defined in the statement of the theorem. Observe how the filter starts to diverge for step-sizes close to 0.1. Actually, using the estimated values for A and B that are generated by the program, it is found that

$$\frac{1}{\lambda_{\max}(A^{-1}B)} \approx 0.103, \qquad \frac{1}{\max\{\lambda(H) \in \mathbb{R}^+\}} \approx 0.204$$

so that we again find that the simulation results are consistent with theory.

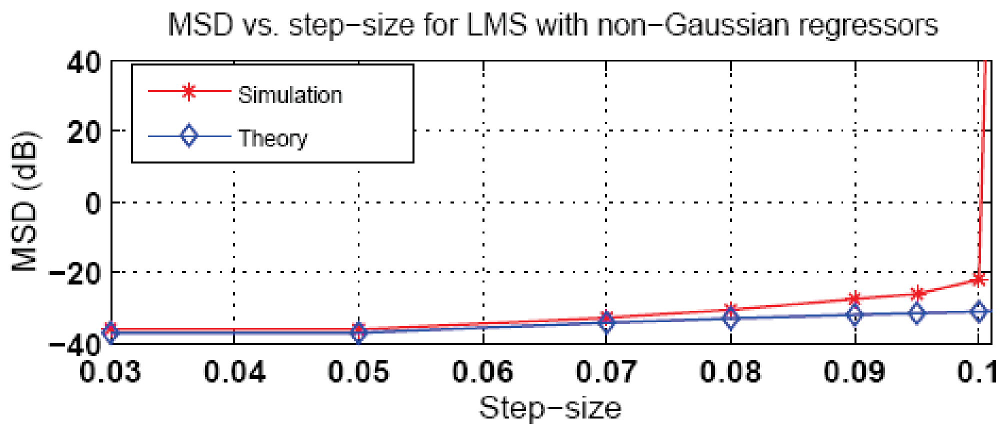

Figure V.4. Theoretical and simulated MSD of LMS as a function of the step-size for non-Gaussian regressors.