9. Transient Performance of Adaptive Filters

Adaptive filters are time-variant and non-linear stochastic systems with inherent learning and tracking abilities.

If they meet underlying requirements for steady-state performance and tracking, they should eventually approach the defined performance levels given enough time.

But what is enough time for an ensemble of adaptations? This is what this section on transient performance will be about.

Most important sections (from preamble)
9.4 and 9.5.

See Chapters 22 to 25 of the on-line text book.

9.1 Data Model

Using the data model from section 6 …

We assume the data satisfies the following conditions

(a) There exists a vector w^o such that $d(i) = u_i \cdot w^o + v(i)$
(b) The noise sequence { v(i) } is i.i.d. with variance $\sigma_v^2 \equiv E|v(i)|^2$
(c) The sequence v(i) is independent of all u_j for all i and j
(d) The initial condition w_{-1} is independent of all { d(j), u_j, v(j) }
(e) The regressor covariance matrix is $R_{uu} \equiv E\left[u_i^{\,h} \cdot u_i\right] > 0$
(f) The random variables { d(i), u_i, v(i) } have zero mean.

9.2 Data-Normalized Adaptive Filters

This section is concerned with a data normalized filter with weight updates of the form:

$$w_i = w_{i-1} + \mu \cdot \frac{u_i^H}{g[u_i]} \cdot e(i)$$

with

$$e(i) = d(i) - u_i \cdot w_{i-1}$$

and where g[u_i] is some positive-valued function of u_i.

For LMS $g[u_i] = 1$

For e-NLMS $g[u_i] = \varepsilon + \|u_i\|^2$

Note: *this is a significant change* from the previous structure and the definition of g[e(i)] which used to be

$$w_i = w_{i-1} + \mu \cdot u_i^H \cdot g[e(i)]$$

9.3 Weighted Energy-Conservation Relation

We shall be dealing with weighted vector norms, as previously seen with the RLS algorithm derivations. Where

$$\|x\|_\Sigma^2 = x^H \cdot \Sigma \cdot x$$

for some Hermitian positive-definite weighting matrix Σ. The choice of Σ =I resultsin the standard Euclidean norm of x.

$$\|x\|_I^2 = x^H \cdot I \cdot x = x^H \cdot x = \|x\|^2$$

To study the transient performance we are interested in the theoretical adaptive performance of

$$\|\tilde{w}_i\|^2 \quad \text{and} \quad |e_a(i)|^2$$

where the first term relates

$$\tilde{w}_i = w^o - w_i$$

and the second

$$e_a(i) \equiv u_i \cdot \tilde{w}_{i-1} = u_i \cdot \left(w^o - w_{i-1}\right)$$

The evaluation of the second term will require the weighted norm

$$\|\tilde{w}_{i-1}\|_{R_{uu}}^2$$

Forming a weighted-norm version of the energy relation, using the approach of section 6.3. Using the following definitions:

$$\tilde{w}_i = w^o - w_i$$

$$e_a^{\Sigma}(i) \equiv u_i \cdot \Sigma \cdot \tilde{w}_{i-1} = u_i \cdot \Sigma \cdot \left(w^o - w_{i-1}\right)$$

$$e_p^{\Sigma}(i) \equiv u_i \cdot \Sigma \cdot \tilde{w}_i = u_i \cdot \Sigma \cdot \left(w^o - w_i\right)$$

Notice that

$$e_a^{I}(i) \equiv u_i \cdot I \cdot \tilde{w}_{i-1} = e_a(i)$$

$$e_p^{I}(i) \equiv u_i \cdot I \cdot \tilde{w}_i = e_p(i)$$

For the derivation, start with new normalized weight update

$$w_i = w_{i-1} + \mu \cdot \frac{u_i^{H}}{g[u_i]} \cdot e(i)$$

$$w^o - w_i = w^o - w_{i-1} - \mu \cdot \frac{u_i^{H}}{g[u_i]} \cdot e(i)$$

$$\tilde{w}_i = \tilde{w}_{i-1} - \mu \cdot \frac{u_i^{H}}{g[u_i]} \cdot e(i)$$

multiply both sides by $u_i \cdot \Sigma$

$$u_i \cdot \Sigma \cdot \tilde{w}_i = u_i \cdot \Sigma \cdot \tilde{w}_{i-1} - \mu \cdot \frac{u_i \cdot \Sigma \cdot u_i^{H}}{g[u_i]} \cdot e(i)$$

Substituting equivalent expressions based on the error estimates

$$e_p^{\Sigma}(i) = e_a^{\Sigma}(i) - \mu \cdot \frac{\left\|u_i\right\|_{\Sigma}^{2}}{g[u_i]} \cdot e(i)$$

For $\left\|u_i\right\|_{\Sigma}^{2} \neq 0$, we derive the energy relation by solving and substituting for the term in g[*]

$$\frac{e(i)}{g[u_i]} = \frac{1}{\mu \cdot \left\|u_i\right\|_{\Sigma}^{2}} \cdot \left(e_a^{\Sigma}(i) - e_p^{\Sigma}(i)\right)$$

and in terms of the weight norms

$$\tilde{w}_i = \tilde{w}_{i-1} - \mu \cdot \frac{u_i^{H}}{\mu \cdot \left\|u_i\right\|_{\Sigma}^{2}} \cdot \left(e_a^{\Sigma}(i) - e_p^{\Sigma}(i)\right)$$

$$\tilde{w}_i + \frac{u_i^{H}}{\left\|u_i\right\|_{\Sigma}^{2}} \cdot e_a^{\Sigma}(i) = \tilde{w}_{i-1} + \frac{u_i^{H}}{\left\|u_i\right\|_{\Sigma}^{2}} \cdot e_p^{\Sigma}(i)$$

Again using a weighted norm, we form the weighted energy terms of both sides

$$\left(\tilde{w}_i + \frac{u_i^{H}}{\left\|u_i\right\|_{\Sigma}^{2}} \cdot e_a^{\Sigma}(i)\right)^{H} \cdot \Sigma \cdot \left(\tilde{w}_i + \frac{u_i^{H}}{\left\|u_i\right\|_{\Sigma}^{2}} \cdot e_a^{\Sigma}(i)\right) = \left(\tilde{w}_{i-1} + \frac{u_i^{H}}{\left\|u_i\right\|_{\Sigma}^{2}} \cdot e_p^{\Sigma}(i)\right)^{H} \cdot \Sigma \cdot \left(\tilde{w}_{i-1} + \frac{u_i^{H}}{\left\|u_i\right\|_{\Sigma}^{2}} \cdot e_p^{\Sigma}(i)\right)$$

Expanding the terms of both sides

$$\tilde{w}_i^{\,H} \cdot \Sigma \cdot \tilde{w}_i + \frac{e_a^\Sigma(i)^H \cdot u_i}{\|u_i\|_\Sigma^{\,2}} \cdot \Sigma \cdot \tilde{w}_i + \tilde{w}_i^{\,H} \cdot \Sigma \cdot \frac{u_i^{\,H}}{\|u_i\|_\Sigma^{\,2}} \cdot e_a^\Sigma(i) + \frac{e_a^\Sigma(i)^H \cdot u_i}{\|u_i\|_\Sigma^{\,2}} \cdot \Sigma \cdot \frac{u_i^{\,H}}{\|u_i\|_\Sigma^{\,2}} \cdot e_a^\Sigma(i) =$$

$$\tilde{w}_{i-1}^{\,H} \cdot \Sigma \cdot \tilde{w}_{i-1} + \frac{e_p^\Sigma(i)^H \cdot u_i}{\|u_i\|_\Sigma^{\,2}} \cdot \Sigma \cdot \tilde{w}_{i-1} + \tilde{w}_{i-1}^{\,H} \cdot \Sigma \cdot \frac{u_i^{\,H}}{\|u_i\|_\Sigma^{\,2}} \cdot e_p^\Sigma(i) + \frac{e_p^\Sigma(i)^H \cdot u_i}{\|u_i\|_\Sigma^{\,2}} \cdot \Sigma \cdot \frac{u_i^{\,H}}{\|u_i\|_\Sigma^{\,2}} \cdot e_p^\Sigma(i)$$

Rewriting to recognized weighted norms

$$\|\tilde{w}_i\|_\Sigma^2 + \frac{e_a^\Sigma(i)^H}{\|u_i\|_\Sigma^{\,2}} \cdot u_i \cdot \Sigma \cdot \tilde{w}_i + \tilde{w}_i^{\,H} \cdot \Sigma \cdot u_i^{\,H} \cdot \frac{e_a^\Sigma(i)}{\|u_i\|_\Sigma^{\,2}} + \frac{1}{\|u_i\|_\Sigma^{\,2}} \cdot \left|e_a^\Sigma(i)\right|^2 =$$

$$\|\tilde{w}_{i-1}\|_\Sigma^2 + \frac{e_p^\Sigma(i)^H}{\|u_i\|_\Sigma^{\,2}} \cdot u_i \cdot \Sigma \cdot \tilde{w}_{i-1} + \tilde{w}_{i-1}^{\,H} \cdot \Sigma \cdot u_i^{\,H} \cdot \frac{e_p^\Sigma(i)}{\|u_i\|_\Sigma^{\,2}} + \frac{1}{\|u_i\|_\Sigma^{\,2}} \cdot \left|e_p^\Sigma(i)\right|^2$$

Based on the Orthogonality of terms (errors and weights), this result in

$$\|\tilde{w}_i\|_\Sigma^2 + \frac{1}{\|u_i\|_\Sigma^{\,2}} \cdot \left|e_a^\Sigma(i)\right|^2 = \|\tilde{w}_{i-1}\|_\Sigma^2 + \frac{1}{\|u_i\|_\Sigma^{\,2}} \cdot \left|e_p^\Sigma(i)\right|^2$$

or equivalently we have the ***weighted energy-conservation relation***

$$\|u_i\|_\Sigma^{\,2} \cdot \|\tilde{w}_i\|_\Sigma^2 + \left|e_a^\Sigma(i)\right|^2 = \|u_i\|_\Sigma^{\,2} \cdot \|\tilde{w}_{i-1}\|_\Sigma^2 + \left|e_p^\Sigma(i)\right|^2$$

Note that for $\|u_i\|_\Sigma^{\,2} = 0$, the final equation is also correct.

A weight recursion relation can be defined as

$$e(i) \equiv d(i) - u_i \cdot w_{i-1}$$

$$e(i) = v(i) + u_i \cdot (\tilde{w}_{i-1})$$

Substituting into the weight update equation

$$\tilde{w}_i = \tilde{w}_{i-1} - \mu \cdot \frac{u_i^{\,H}}{g[u_i]} \cdot e(i)$$

$$\tilde{w}_i = \tilde{w}_{i-1} - \mu \cdot \frac{u_i^{\,H}}{g[u_i]} \cdot [v(i) + u_i \cdot (\tilde{w}_{i-1})]$$

$$\tilde{w}_i = \left[I - \mu \cdot \frac{u_i^{\,H} \cdot u_i}{g[u_i]} \right] \cdot \tilde{w}_{i-1} - \mu \cdot \frac{u_i^{\,H}}{g[u_i]} \cdot v(i)$$

This is a generalized weight recursion that exists for all adaptive system meeting the data model.

9.4 Weighted Variance Relation

The weighted variance relationship no longer starts with the assumption of equal weight norms,

$$\|u_i\|_\Sigma^{\,2}\cdot\|\tilde{w}_i\|_\Sigma^2+\left|e_a^\Sigma(i)\right|^2=\|u_i\|_\Sigma^{\,2}\cdot\|\tilde{w}_{i-1}\|_\Sigma^2+\left|e_p^\Sigma(i)\right|^2$$

but still substitutes for the a-posteriori weighted estimation error

$$e_p^\Sigma(i)=e_a^\Sigma(i)-\mu\cdot\frac{\|u_i\|_\Sigma^{\,2}}{g[u_i]}\cdot e(i)$$

$$\|u_i\|_\Sigma^{\,2}\cdot\|\tilde{w}_i\|_\Sigma^2+\left|e_a^\Sigma(i)\right|^2=\|u_i\|_\Sigma^{\,2}\cdot\|\tilde{w}_{i-1}\|_\Sigma^2+\left|e_a^\Sigma(i)-\mu\cdot\frac{\|u_i\|_\Sigma^{\,2}}{g[u_i]}\cdot e(i)\right|^2$$

Expanding

$$\|u_i\|_\Sigma^{\,2}\cdot\|\tilde{w}_i\|_\Sigma^2+\left|e_a^\Sigma(i)\right|^2=\|u_i\|_\Sigma^{\,2}\cdot\|\tilde{w}_{i-1}\|_\Sigma^2+$$

$$\left|e_a^\Sigma(i)\right|^2-\mu\cdot e_a^\Sigma(i)^H\cdot\frac{\|u_i\|_\Sigma^{\,2}}{g[u_i]}\cdot e(i)-\mu\cdot\frac{\|u_i\|_\Sigma^{\,2}}{g[u_i]}\cdot e(i)^H\cdot e_a^\Sigma(i)+\mu^2\cdot\frac{\|u_i\|_\Sigma^{\,4}}{g[u_i]^2}\cdot|e(i)|^2$$

Eliminating common terms and dividing through by the weighted norm of u_i

$$\|\tilde{w}_i\|_\Sigma^2=\|\tilde{w}_{i-1}\|_\Sigma^2-\mu\cdot\frac{1}{g[u_i]}\cdot e_a^\Sigma(i)^H\cdot e(i)-\mu\cdot\frac{1}{g[u_i]}\cdot e(i)^H\cdot e_a^\Sigma(i)+\mu^2\cdot\frac{\|u_i\|_\Sigma^{\,2}}{g[u_i]^2}\cdot|e(i)|^2$$

Using

$$e(i)=e_a(i)+v(i)$$

Substituting to eliminate e(i)

$$\|\tilde{w}_i\|_\Sigma^2=\|\tilde{w}_{i-1}\|_\Sigma^2-\mu\cdot e_a^\Sigma(i)^H\cdot\frac{1}{g[u_i]}\cdot\left(e_a(i)+v(i)\right)+$$

$$-\mu\cdot\frac{1}{g[u_i]}\cdot\left(e_a(i)+v(i)\right)^H\cdot e_a^\Sigma(i)+\mu^2\cdot\frac{\|u_i\|_\Sigma^{\,2}}{g[u_i]^2}\cdot\left|e_a(i)+v(i)\right|^2$$

and expanding

$$\|\tilde{w}_i\|_\Sigma^2=\|\tilde{w}_{i-1}\|_\Sigma^2-\frac{\mu}{g[u_i]}\cdot e_a^\Sigma(i)^H\cdot e_a(i)-\frac{\mu}{g[u_i]}\cdot e_a^\Sigma(i)^H\cdot v(i)-\frac{\mu}{g[u_i]}\cdot e_a(i)^H\cdot e_a^\Sigma(i)+$$

$$-\frac{\mu}{g[u_i]}\cdot v(i)^H\cdot e_a^\Sigma(i)+\frac{\mu^2\cdot\|u_i\|_\Sigma^{\,2}}{g[u_i]^2}\cdot|e_a(i)|^2+\frac{\mu^2\cdot\|u_i\|_\Sigma^{\,2}}{g[u_i]^2}\cdot e_a(i)^H\cdot v(i)+$$

$$+\frac{\mu^2\cdot\|u_i\|_\Sigma^{\,2}}{g[u_i]^2}\cdot v(i)^H\cdot e_a(i)+\frac{\mu^2\cdot\|u_i\|_\Sigma^{\,2}}{g[u_i]^2}\cdot|v(i)|^2$$

There are some terms to deal with … weighted and unweighted a-priori errors

Analyzing specific terms

$$e_a^{\Sigma}(i)^H \cdot e_a(i) = \tilde{w}_{i-1}{}^H \cdot \Sigma \cdot u_i{}^H \cdot u_i \cdot \tilde{w}_{i-1} = \tilde{w}_{i-1}{}^H \cdot \left[\Sigma \cdot u_i{}^H \cdot u_i\right] \cdot \tilde{w}_{i-1}$$

$$e_a(i)^H \cdot e_a^{\Sigma}(i) = \tilde{w}_{i-1}{}^H \cdot u_i{}^H \cdot u_i \cdot \Sigma \cdot \tilde{w}_{i-1} = \tilde{w}_{i-1}{}^H \cdot \left[u_i{}^H \cdot u_i \cdot \Sigma\right] \cdot \tilde{w}_{i-1}$$

or

$$e_a^{\Sigma}(i)^H \cdot e_a(i) + e_a(i)^H \cdot e_a^{\Sigma}(i) = \tilde{w}_{i-1}{}^H \cdot \left[\Sigma \cdot u_i{}^H \cdot u_i + u_i{}^H \cdot u_i \cdot \Sigma\right] \cdot \tilde{w}_{i-1}$$

and using

$$e_a(i)^H \cdot e_a(i) = \tilde{w}_{i-1}{}^H \cdot u_i{}^H \cdot u_i \cdot \tilde{w}_{i-1} = \tilde{w}_{i-1}{}^H \cdot \left[u_i{}^H \cdot u_i\right] \cdot \tilde{w}_{i-1}$$

Substituting

$$\begin{aligned}
\|\tilde{w}_i\|_{\Sigma}^2 = \|\tilde{w}_{i-1}\|_{\Sigma}^2 &- \frac{\mu}{g[u_i]} \cdot \tilde{w}_{i-1}{}^H \cdot \left[\Sigma \cdot u_i{}^H \cdot u_i + u_i{}^H \cdot u_i \cdot \Sigma\right] \cdot \tilde{w}_{i-1} \\
&+ \frac{\mu^2 \cdot \|u_i\|_{\Sigma}{}^2}{g[u_i]^2} \cdot \tilde{w}_{i-1}{}^H \cdot \left[u_i{}^H \cdot u_i\right] \cdot \tilde{w} + \frac{\mu^2 \cdot \|u_i\|_{\Sigma}{}^2}{g[u_i]^2} \cdot |v(i)|^2 \\
&- \frac{\mu}{g[u_i]} \cdot e_a^{\Sigma}(i)^H \cdot v(i) - \frac{\mu}{g[u_i]} \cdot v(i)^H \cdot e_a^{\Sigma}(i) \\
&+ \frac{\mu^2 \cdot \|u_i\|_{\Sigma}{}^2}{g[u_i]^2} \cdot e_a(i)^H \cdot v(i) + \frac{\mu^2 \cdot \|u_i\|_{\Sigma}{}^2}{g[u_i]^2} \cdot v(i)^H \cdot e_a(i)
\end{aligned}$$

and recognizing the orthogonality of an expected value of v(i) and the a-priori errors

$$\begin{aligned}
E\|\tilde{w}_i\|_{\Sigma}^2 = E\|\tilde{w}_{i-1}\|_{\Sigma}^2 &- E\frac{\mu}{g[u_i]} \cdot \tilde{w}_{i-1}{}^H \cdot \left[\Sigma \cdot u_i{}^H \cdot u_i + u_i{}^H \cdot u_i \cdot \Sigma\right] \cdot \tilde{w}_{i-1} + \\
&+ E\frac{\mu^2 \cdot \|u_i\|_{\Sigma}{}^2}{g[u_i]^2} \cdot \tilde{w}_{i-1}{}^H \cdot \left[u_i{}^H \cdot u_i\right] \cdot \tilde{w}_{i-1} + \sigma_v^2 \cdot E\frac{\mu^2 \cdot \|u_i\|_{\Sigma}{}^2}{g[u_i]^2}
\end{aligned}$$

or as defined in the text

$$E\|\tilde{w}_i\|_{\Sigma}^2 = \left\{E\|\tilde{w}_{i-1}\|_{\Sigma}^2 - E\frac{\mu}{g[u_i]} \cdot \|\tilde{w}_{i-1}\|_{\Sigma \cdot u_i{}^H \cdot u_i + u_i{}^H \cdot u_i \cdot \Sigma}^2 + E\frac{\mu^2 \cdot \|u_i\|_{\Sigma}{}^2}{g[u_i]^2} \cdot \|\tilde{w}_{i-1}\|_{u_i{}^H \cdot u_i}^2\right\} + \mu^2 \cdot \sigma_v^2 \cdot E\frac{\|u_i\|_{\Sigma}{}^2}{g[u_i]^2}$$

The scalar quantities can be included in the weighted norm functions and the common weights can be combined. First, define

$$\Sigma' = \Sigma - \frac{\mu}{g[u_i]} \cdot \Sigma \cdot u_i{}^H \cdot u_i - \frac{\mu}{g[u_i]} \cdot u_i{}^H \cdot u_i \cdot \Sigma + \frac{\mu^2 \cdot \|u_i\|_{\Sigma}{}^2}{g[u_i]^2} \cdot u_i{}^H \cdot u_i$$

Then

$$E\|\tilde{w}_i\|_{\Sigma}^2 = E\|\tilde{w}_{i-1}\|_{\Sigma'}^2 + \mu^2 \cdot \sigma_v^2 \cdot E\left[\frac{\|u_i\|_{\Sigma}{}^2}{g[u_i]^2}\right]$$

This is the new weighted variance relation.
Again, it is an exact relationship without assumptions.

Independence Assumption

This relation is difficult to propagate due to it's dependence on the regression vector u_i.

The independence assumption is used as an addition to the model in order to evaluate the iterations. Therefore, assume that:

The sequence { u_i } is independent and identically distributed.

The results have been shown to be reasonable using this assumption, but in general, it is not valid.

This is a new assumption, and that is why transient analysis is separate from the previous work in Sections 6, 7, and 8.

With this condition, we can say that

The estimation error, $\tilde{w}_{i-1}$, at i-1 is independent of both u_i and Σ'

This follows as the estimation error, $\tilde{w}_{i-1}$, at i-1 is dependent upon the previous regression vectors u_j and v(j) for j<i, while Σ' is a function of u_i. Then we can say that

$$E\|\tilde{w}_{i-1}\|^2_{\Sigma'} = E\|\tilde{w}_{i-1}\|^2_{E[\Sigma']}$$

$$E\|\tilde{w}_{i-1}\|^2_{\Sigma'} = E\left[\tilde{w}_{i-1}{}^H \cdot E[\Sigma'] \cdot \tilde{w}_{i-1}\right]$$

Now, the weighting matrices become deterministic and no longer must be treated as a random variables.

The weight variance relation can then be computed as

$$E\|\tilde{w}_i\|^2_{\Sigma} = E\|\tilde{w}_{i-1}\|^2_{\Sigma'} + \mu^2 \cdot \sigma_v^2 \cdot E\left[\frac{\|u_i\|_{\Sigma}{}^2}{g[u_i]^2}\right]$$

$$\Sigma' = \Sigma - \mu \cdot \Sigma \cdot E\left[\frac{u_i{}^H \cdot u_i}{g[u_i]}\right] - \mu \cdot E\left[\frac{u_i{}^H \cdot u_i}{g[u_i]}\right] \cdot \Sigma + \mu^2 \cdot E\left[\frac{\|u_i\|_{\Sigma}{}^2}{g[u_i]^2} \cdot u_i{}^H \cdot u_i\right]$$

Using the generalized weight recursion, previous developed and the independence assumption,

$$\tilde{w}_i = \left[I - \mu \cdot \frac{u_i{}^H \cdot u_i}{g[u_i]}\right] \cdot \tilde{w}_{i-1} - \mu \cdot \frac{u_i{}^H}{g[u_i]} \cdot v(i)$$

becomes

$$E[\tilde{w}_i] = \left[I - \mu \cdot E\left[\frac{u_i{}^H \cdot u_i}{g[u_i]}\right]\right] \cdot E[\tilde{w}_{i-1}] - \mu \cdot E\left[\frac{u_i{}^H}{g[u_i]} \cdot v(i)\right]$$

With the last term being independent (orthogonal) becomes

$$E\left[\tilde{w}_i\right] = \left[I - \mu \cdot E\left[\frac{u_i^H \cdot u_i}{g[u_i]} \right]\right] \cdot E\left[\tilde{w}_{i-1}\right]$$

We now have a recursion in the weight error and in Σ.

To summarize:
The terms of interest for solving the iterations are:

$$E\|\tilde{w}_i\|_{\Sigma}^2 = E\|\tilde{w}_{i-1}\|_{\Sigma'}^2 + \mu^2 \cdot \sigma_v^2 \cdot E\left[\frac{\|u_i\|_{\Sigma}^2}{g[u_i]^2} \right]$$

$$\Sigma' = \Sigma - \mu \cdot \Sigma \cdot E\left[\frac{u_i^H \cdot u_i}{g[u_i]} \right] - \mu \cdot E\left[\frac{u_i^H \cdot u_i}{g[u_i]} \right] \cdot \Sigma + \mu^2 \cdot E\left[\frac{\|u_i\|_{\Sigma}^2}{g[u_i]^2} \cdot u_i^H \cdot u_i \right]$$

$$E\left[\tilde{w}_i\right] = \left[I - \mu \cdot E\left[\frac{u_i^H \cdot u_i}{g[u_i]} \right]\right] \cdot E\left[\tilde{w}_{i-1}\right]$$

with the expected values that must be computed:

$$E\left[\frac{\|u_i\|_{\Sigma}^2}{g[u_i]^2} \right] \text{ and } E\left[\frac{u_i^H \cdot u_i}{g[u_i]} \right] \text{ and } E\left[\frac{\|u_i\|_{\Sigma}^2}{g[u_i]^2} \cdot u_i^H \cdot u_i \right]$$

This can be nasty. So, where possible, a change in coordinates to simplify the computations can be performed.

Convenient Change of Coordinates

We can use matrix manipulation to diagonalizable the regression covariance.

$$E\left[u_i^H \cdot u_i\right] = R_{uu} = U \cdot \Lambda \cdot U^H$$

The estimate-bars are then generated by multiplying by the unitary matrix U

$$\overline{\tilde{w}_i} = U^H \cdot \tilde{w}_i \text{ and } \overline{u_i} = u_i \cdot U \text{ and } \overline{\Sigma} = U^H \cdot \Sigma \cdot U$$

Note that

$$\left\|\overline{\tilde{w}_i}\right\|_{\overline{\Sigma}}^2 = \tilde{w}_i^H \cdot U \cdot \left(U^H \cdot \Sigma \cdot U\right) \cdot U^H \cdot \tilde{w}_i = \|\tilde{w}_i\|_{\Sigma}^2$$

$$\left\|\overline{u_i}\right\|_{\overline{\Sigma}}^2 = \left(u_i \cdot U\right) \cdot \left(U^H \cdot \Sigma \cdot U\right) \cdot \left(U^H \cdot u_i^H\right) = \left\|\overline{u_i}\right\|_{\Sigma}^2$$

This results in a transformed relation that replaces variable-bar for the desired variables.

See p. 450-452.

9.5 Transient Performance of LMS

For the LMS filter

$$e(i) = d(i) - u_i \cdot w_{i-1}$$
$$w_i = w_{i-1} + \mu \cdot u_i^H \cdot e(i)$$

The new g[] function

$$g[u_i] = 1$$

Using the non-bar relationships for the "terms of interest"

$$E\left[\|\tilde{w}_i\|_\Sigma^2\right] = E\left[\|\tilde{w}_{i-1}\|_{\Sigma'}^2\right] + \mu^2 \cdot \sigma_v^2 \cdot E\left[\|u_i\|_\Sigma^2\right]$$

$$\Sigma' = \Sigma - \mu \cdot \Sigma \cdot E\left[u_i^H \cdot u_i\right] - \mu \cdot E\left[u_i^H \cdot u_i\right] \cdot \Sigma + \mu^2 \cdot E\left[\|u_i\|_\Sigma^2 \cdot u_i^H \cdot u_i\right]$$

$$E[\tilde{w}_i] = \left[I - \mu \cdot E\left[u_i^H \cdot u_i\right]\right] \cdot E[\tilde{w}_{i-1}]$$

with the expected values that must be computed:

$$E\left[\|u_i\|_\Sigma^2\right] \text{ and } E\left[u_i^H \cdot u_i\right] \text{ and } E\left[\|u_i\|_\Sigma^2 \cdot u_i^H \cdot u_i\right]$$

By definition,

$$E\left[u_i^H \cdot u_i\right] = R_{uu}$$

and

$$E\left[\|u_i\|_\Sigma^2\right] = E\left[u_i \cdot \Sigma \cdot u_i^H\right] = E\left[Tr\left(u_i \cdot u_i^H \cdot \Sigma\right)\right] = Tr\left(R_{uu} \cdot \Sigma\right)$$

The last moment is more difficult and will be performed as two possible cases, Gaussian and Non-Gaussian.

Gaussian Regressors

Assume that the regressor arises from a circular Gaussian distribution a coordinate change will result in a diagonal covariance matrix. So that the previous evaluations become

$$E\left[\overline{u_i}^H \cdot \overline{u_i}\right] = \Lambda$$

and

$$E\left[\left\|\overline{u_i}\right\|_{\overline{\Sigma}}^2\right] = E\left[\overline{u_i} \cdot \overline{\Sigma} \cdot \overline{u_i}^H\right] = E\left[Tr\left(\overline{u_i} \cdot \overline{u_i}^H \cdot \overline{\Sigma}\right)\right] = Tr\left(\Lambda \cdot \overline{\Sigma}\right)$$

Then in computing the “more difficult” term, the scalar term can be moved in the computation as

$$E\left[\left\|\overline{u_i}\right\|_{\overline{\Sigma}}^2 \cdot \overline{u_i}^H \cdot \overline{u_i}\right] = E\left[\left(\overline{u_i} \cdot \overline{\Sigma} \cdot \overline{u_i}^H\right) \cdot \overline{u_i}^H \cdot \overline{u_i}\right] = E\left[\overline{u_i}^H \cdot \left(\overline{u_i} \cdot \overline{\Sigma} \cdot \overline{u_i}^H\right) \cdot \overline{u_i}\right]$$

Using Lemma 1.B.3 on p. 46 which requires a diagonal covariance matrix and complex regressors, u_i

$$E\left[\left\|\overline{u_i}\right\|_{\overline{\Sigma}}^2 \cdot \overline{u_i}^H \cdot \overline{u_i}\right] = \Lambda \cdot Tr\left(\overline{\Sigma} \cdot \Lambda\right) + \Lambda \cdot \overline{\Sigma} \cdot \Lambda$$

Important Note: For real data and regressor, u_i, Lemma 1.B.2 on p. 44 should be used and

$$E\left[\left\|\overline{u_i}\right\|_{\overline{\Sigma}}^2 \cdot \overline{u_i}^H \cdot \overline{u_i}\right] = \Lambda \cdot Tr\left(\overline{\Sigma} \cdot \Lambda\right) + 2 \cdot \Lambda \cdot \overline{\Sigma} \cdot \Lambda$$

The following assumes a complex regressor.

Using the results for the three terms

$$E\left\|\overline{\tilde{w}_i}\right\|_{\overline{\Sigma}}^2 = E\left\|\overline{\tilde{w}_{i-1}}\right\|_{\overline{\Sigma}'}^2 + \mu^2 \cdot \sigma_v^2 \cdot Tr\left(\Lambda \cdot \overline{\Sigma}\right)$$

$$\overline{\Sigma}' = \overline{\Sigma} - \mu \cdot \overline{\Sigma} \cdot \Lambda - \mu \cdot \Lambda \cdot \overline{\Sigma} + \mu^2 \cdot \left[\Lambda \cdot Tr\left(\overline{\Sigma} \cdot \Lambda\right) + \Lambda \cdot \overline{\Sigma} \cdot \Lambda\right]$$

$$E\left[\overline{\tilde{w}_i}\right] = \left[I - \mu \cdot \Lambda\right] \cdot E\left[\overline{\tilde{w}_{i-1}}\right]$$

Note that all elements are known based on the assumptions previously made and can be iteratively computed as a ***theoretical performance curve****.*

We can now project theoretically the performance of an adaptive system based on a-priori knowledge of the Gaussian nature of the regressor and the Gaussian statistics.

Also, if Σ is diagonal than Σ’ is as well and further simplification can be made to compute vectors instead of a diagonal matrix.

Simplifications for diagonal matrices follow …

Diagonal Notation

Allow the standard matrix to vector definitions to exist (the MATLAB def of diag)

$$\overline{\sigma} \equiv diag\left[\overline{\Sigma}\right] \text{ and } \lambda \equiv diag\left[\Lambda\right]$$

and

$$\left[\overline{\Sigma}\right] \equiv diag\left[\overline{\sigma}\right] \text{ and } \Lambda \equiv diag\left[\lambda\right]$$

Then

$$diag\left(\overline{\Sigma'}\right) = diag\left(\overline{\Sigma}\right) - diag\left(\mu \cdot \overline{\Sigma} \cdot \Lambda + \mu \cdot \Lambda \cdot \overline{\Sigma}\right) + diag\left(\mu^2 \cdot \Lambda \cdot Tr\left(\overline{\Sigma} \cdot \Lambda\right)\right) + diag\left(\mu^2 \cdot \Lambda \cdot \overline{\Sigma} \cdot \Lambda\right)$$

As both matrices are diagonal

$$diag\left(\overline{\Sigma'}\right) = diag\left(I \cdot \overline{\Sigma} - 2 \cdot \mu \cdot \Lambda \cdot \overline{\Sigma} + \mu^2 \cdot \Lambda \cdot \Lambda \cdot \overline{\Sigma}\right) + diag\left(\mu^2 \cdot \Lambda \cdot Tr\left(\overline{\Sigma} \cdot \Lambda\right)\right)$$

or

$$\overline{\sigma'} = \left(I - 2 \cdot \mu \cdot \Lambda + \mu^2 \cdot \Lambda \cdot \Lambda\right) \cdot \overline{\sigma} + \mu^2 \cdot \lambda \cdot \left(\lambda^T \cdot \overline{\sigma}\right)$$

$$\overline{\sigma'} = \left(I - 2 \cdot \mu \cdot \Lambda + \mu^2 \cdot \Lambda \cdot \Lambda\right) \cdot \overline{\sigma} + \mu^2 \cdot \left(\lambda \cdot \lambda^T\right) \cdot \overline{\sigma}$$

or finally

$$\overline{\sigma'} = \overline{F} \cdot \overline{\sigma}$$

where

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

Interesting result, there is a linear vector relation between successive sigma iterations

$$\overline{\Sigma} = diag\left[\overline{\sigma}\right] \text{ and } \overline{\Sigma'} = diag\left(\overline{\sigma'}\right)$$

The relations can then be summarized as

$$E\left\|\overline{\tilde{w}_i}\right\|^2_{\overline{\sigma}} = E\left\|\overline{\tilde{w}_{i-1}}\right\|^2_{\overline{F} \cdot \overline{\sigma}} + \mu^2 \cdot \sigma_v^2 \cdot \left(\lambda^T \cdot \overline{\sigma}\right)$$

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

$$\overline{\sigma'} = \overline{F} \cdot \overline{\sigma}$$

$$E\left[\overline{\tilde{w}_i}\right] = \left[I - \mu \cdot \Lambda\right] \cdot E\left[\overline{\tilde{w}_{i-1}}\right]$$

Properties of the equations.

Mean Behavior
Mean-Square Behavior
State-Space Stability
Steady-State Performance

Mean Behavior

By assumption, the initial condition for the iteration of the weight error is

$$\tilde{w}_{-1} \equiv w^{opt} - w_{-1} \equiv w^{opt}$$

and

$$\overline{w}_{-1} \equiv U^{H} \cdot w^{opt} \equiv \overline{w}^{opt}$$

The expected mean value is then

$$E\left[\overline{w}_{-1}\right] \equiv \overline{w}^{opt}$$

Then, based on the iterative equation

$$E\left[\overline{\tilde{w}_{i}}\right] = \left[I - \mu \cdot \Lambda\right] \cdot E\left[\overline{\tilde{w}_{i-1}}\right]$$

the mean is

$$E\left[\overline{\tilde{w}_{i}}\right] = \left[I - \mu \cdot \Lambda\right]^{i+1} \cdot \overline{w}^{opt}$$

As previously shown for the LMS algorithm, this places a bound on the step size. For

$$\left|1 - \mu \cdot \lambda_{k}\right| < 1,\ \ for\ \ k = 1 : M$$

we require

$$\mu < {}^{2}\!/\!_{\lambda_{\max}}$$

and then

$$\lim_{i \to \infty} E\left[\overline{\tilde{w}_{i}}\right] = \overline{w}^{opt}$$

Mean-Square Behavior

By assumption, let the initial value for sigma bar be the identity matrix

$$\overline{\Sigma} = I \quad \text{or} \quad \overline{\sigma} = col\{1 \quad \cdots \quad 1\} \equiv q$$

Using the variance recursion

$$E\left\|\overline{\tilde{w}_{i}}\right\|^{2}_{\overline{\sigma}} = E\left\|\overline{\tilde{w}_{i-1}}\right\|^{2}_{\overline{F} \cdot \overline{\sigma}} + \mu^{2} \cdot \sigma_{v}^{2} \cdot \left(\lambda^{T} \cdot \overline{\sigma}\right)$$

and variance propagation recursion

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

$$\overline{\sigma'} = \overline{F} \cdot \overline{\sigma}$$

Then “back projecting”

$$E\left\|\overline{\tilde{w}_{i}}\right\|^{2} = E\left\|\overline{\tilde{w}_{i}}\right\|^{2}_{q} = E\left\|\overline{\tilde{w}_{i-1}}\right\|^{2}_{\overline{F} \cdot q} + \mu^{2} \cdot \sigma_{v}^{2} \cdot \left(\lambda^{T} \cdot q\right)$$

$$E\left\|\overline{\tilde{w}_{i-1}}\right\|^{2}_{\overline{F} \cdot q} = E\left\|\overline{\tilde{w}_{i-2}}\right\|^{2}_{\overline{F}^{2} \cdot q} + \mu^{2} \cdot \sigma_{v}^{2} \cdot \left(\lambda^{T} \cdot \overline{F} \cdot q\right)$$

$$E\left\|\overline{\tilde{w}_{i-2}}\right\|^{2}_{\overline{F}^{2} \cdot q} = E\left\|\overline{\tilde{w}_{i-3}}\right\|^{2}_{\overline{F}^{3} \cdot q} + \mu^{2} \cdot \sigma_{v}^{2} \cdot \left(\lambda^{T} \cdot \overline{F}^{2} \cdot q\right)$$

continuing to the initial conditions

$$E\left\|\overline{\tilde{w}_{0}}\right\|^{2}_{\overline{F}^{i} \cdot q} = E\left\|\overline{w^{opt}}\right\|^{2}_{\overline{F}^{i+1} \cdot q} + \mu^{2} \cdot \sigma_{v}^{2} \cdot \left(\lambda^{T} \cdot \overline{F}^{i} \cdot q\right) = \left\|\overline{w^{opt}}\right\|^{2}_{\overline{F}^{i+1} \cdot q} + \mu^{2} \cdot \sigma_{v}^{2} \cdot \left(\lambda^{T} \cdot \overline{F}^{i} \cdot q\right)$$

Now that we have the back projection to the initial conditions, we can develop the forward summation of terms.

$$E\left\|\overline{\tilde{w}_i}\right\|^2 = E\left\|\overline{\tilde{w}_i}\right\|_q^2 = \left\|\overline{w^{opt}}\right\|^2_{\overline{F}^{i+1}\cdot q} + \mu^2\cdot\sigma_v^2\cdot\sum_{k=0}^{i}\left(\lambda^T\cdot\overline{F}^k\cdot q\right)$$

From this result, note that

$$E\left\|\overline{\tilde{w}_{i-1}}\right\|^2 = \left\|\overline{w^{opt}}\right\|^2_{\overline{F}^{i}\cdot q} + \mu^2\cdot\sigma_v^2\cdot\sum_{k=0}^{i-1}\left(\lambda^T\cdot\overline{F}^k\cdot q\right)$$

Then

$$E\left\|\overline{\tilde{w}_i}\right\|^2 = \left\|\overline{w^{opt}}\right\|^2_{\overline{F}^{i+1}\cdot q} + \mu^2\cdot\sigma_v^2\cdot\left(\lambda^T\cdot\overline{F}^i\cdot q\right) + \left[E\left\|\overline{\tilde{w}_{i-1}}\right\|^2 - \left\|\overline{w^{opt}}\right\|^2_{\overline{F}^{i}\cdot q}\right]$$

$$E\left\|\overline{\tilde{w}_i}\right\|^2 = E\left\|\overline{\tilde{w}_{i-1}}\right\|^2 + \mu^2\cdot\sigma_v^2\cdot\left(\lambda^T\cdot\overline{F}^i\cdot q\right) + \left[\left\|\overline{w^{opt}}\right\|^2_{\overline{F}^{i+1}\cdot q} - \left\|\overline{w^{opt}}\right\|^2_{\overline{F}^{i}\cdot q}\right]$$

and leading to

$$E\left\|\overline{\tilde{w}_i}\right\|^2 = E\left\|\overline{\tilde{w}_{i-1}}\right\|^2 + \mu^2\cdot\sigma_v^2\cdot\left(\lambda^T\cdot\overline{F}^i\cdot q\right) - \left\|\overline{w^{opt}}\right\|^2_{\overline{F}^{i}\cdot\left(I-\overline{F}\right)\cdot q}$$

This recursive relationship can be used to generate the theoretical curve desired. Given

$$\overline{F} = I - 2\cdot\mu\cdot\Lambda + \mu^2\cdot\Lambda^2 + \mu^2\cdot\left(\lambda\cdot\lambda^T\right) \quad \text{and} \quad \overline{w}_{-1} \equiv \overline{w}^{opt} \quad \text{and} \quad \overline{F}^0 = I$$

$$E\left\|\overline{\tilde{w}_0}\right\|^2 = \left\|\overline{w^{opt}}\right\|^2$$

The computations needed are, starting at i=1:

$$\overline{F}^i = \overline{F}^{i-1}\cdot\overline{F}$$

$$b = \mu^2\cdot\sigma_v^2\cdot\left(\lambda^T\cdot\overline{F}^i\cdot q\right)$$

$$c = \overline{w^{opt}}^H\cdot\left[\overline{F}^i\cdot\left(I-\overline{F}\right)\cdot q\right]\cdot\overline{w^{opt}}$$

$$E\left\|\overline{\tilde{w}_i}\right\|^2 = E\left\|\overline{\tilde{w}_{i-1}}\right\|^2 + b - c$$

State-Space, Mean-Square Stability

The system will be stable of the eigenvalues of F-bar lie between -1 and +1.

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

or

$$\overline{F} = I - 2 \cdot \mu \cdot \Lambda + \mu^2 \cdot \left[\Lambda^2 + \lambda \cdot \lambda^T\right] = I - \mu \cdot A + \mu^2 \cdot B$$

Conditions:

F-bar is positive definite. Then we only must worry about the maximum eigenvalue of F-bar. From linear algebra, it's eigenvalues will be bounded by one if and only if

$$0 < \mu < {}^{1}\!/_{\lambda_{\max}\left(of\ \ A^{-1} \cdot B\right)}$$

To converge in the mean and mean-square sense,

$$0 < \mu < \min\left({}^{2}\!/_{\lambda_{\max}\left(of\ \ \Lambda\right)}, {}^{1}\!/_{\lambda_{\max}\left(of\ \ A^{-1} \cdot B\right)} \right)$$

but it can be proven that

$${}^{1}\!/_{\lambda_{\max}\left(of\ \ A^{-1} \cdot B\right)} < {}^{2}\!/_{\lambda_{\max}\left(of\ \ \Lambda\right)}$$

therefore it is sufficient that

$$0 < \mu < {}^{1}\!/_{\lambda_{\max}\left(of\ \ A^{-1} \cdot B\right)}$$

Deriving the maximum step size ….

Letting

$$\eta_0 = {}^{1}\!/_{\lambda_{\max}\left(of\ \ A^{-1} \cdot B\right)}$$

The value we are interested in is the minimum value for which

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

$$\det\left(I - \eta \cdot \frac{\Lambda^{-1}}{2} \cdot \left[\Lambda^2 + \lambda \cdot \lambda^T\right] \right) = 0$$

$$\det\left(I - \frac{\eta}{2} \cdot \left[\Lambda + q \cdot \lambda^T\right] \right) = 0$$

$$\det\left(\left(I - \frac{\eta}{2} \cdot \Lambda \right) \cdot \left(I - \left(I - \frac{\eta}{2} \cdot \Lambda \right)^{-1} \cdot \frac{\eta}{2} \cdot q \cdot \lambda^T \right) \right) = 0$$

$$\det\left(I - \frac{\eta}{2} \cdot \Lambda \right) \cdot \det\left(I - \left(I - \frac{\eta}{2} \cdot \Lambda \right)^{-1} \cdot \frac{\eta}{2} \cdot q \cdot \lambda^T \right) = 0$$

Using the formula $\det\left(I - X \cdot Y\right) = \det\left(I - Y \cdot X\right)$

$$\det\left(I - \eta \cdot \frac{\Lambda}{2} \right) \cdot \det\left(1 - \lambda^T \cdot \left(\frac{2}{\eta} \cdot I - \Lambda \right)^{-1} \cdot q \right) = 0$$

This defines the two conditions previously stated. In addition, the second term is a scalar.

$$\det\left(I-\eta\cdot\frac{\Lambda}{2}\right)\cdot\left(1-\lambda^T\cdot\left(\frac{2}{\eta}\cdot I-\Lambda\right)^{-1}\cdot q\right)=0$$

The value that sets the determinant to zero then must come from

$$1-\lambda^T\cdot\left(\frac{2}{\eta}\cdot I-\Lambda\right)^{-1}\cdot q=0$$

Note that this is the equivalent of

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

and using q as a summation of the diagonal terms produced

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

As a function potentially complicated to compute, we form an eta function and then search for where it goes to 1 as the maximum.

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

It may be noted that at the point where

$$2-\eta\cdot\lambda_k=0$$

the function will approach infinity! But this is simply the point where

$$\eta=\frac{2}{\lambda_{\max}}$$

As stated before, this corresponds to a point where the eta function is >> 1, beyond where it would equal 1 for our bound on the step size.

Real vs. Complex regressors

Note that for real regressors, the function becomes

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

Going back further, the iterative equations must also be adjusted as

$$E\left[\left\|\overline{u_i}\right\|_{\overline{\Sigma}}^{2}\cdot\overline{u_i}^{H}\cdot\overline{u_i}\right]=\Lambda\cdot Tr\left(\overline{\Sigma}\cdot\Lambda\right)+2\cdot\Lambda\cdot\overline{\Sigma}\cdot\Lambda$$

$$\overline{\Sigma}'=\overline{\Sigma}-\mu\cdot\overline{\Sigma}\cdot\Lambda-\mu\cdot\Lambda\cdot\overline{\Sigma}+\mu^2\cdot\left[\Lambda\cdot Tr\left(\overline{\Sigma}\cdot\Lambda\right)+2\cdot\Lambda\cdot\overline{\Sigma}\cdot\Lambda\right]$$

$$diag\left(\overline{\Sigma}'\right)=diag\left(I\cdot\overline{\Sigma}-2\cdot\mu\cdot\Lambda\cdot\overline{\Sigma}+2\cdot\mu^2\cdot\Lambda\cdot\Lambda\cdot\overline{\Sigma}\right)+diag\left(\mu^2\cdot\Lambda\cdot Tr\left(\overline{\Sigma}\cdot\Lambda\right)\right)$$

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

and a change in the value B where the new functions is

$$\overline{F}=I-2\cdot\mu\cdot\Lambda+\mu^2\cdot\left[2\cdot\Lambda^2+\lambda\cdot\lambda^T\right]=I-\mu\cdot A+\mu^2\cdot B$$

for

$$0<\mu<{}^{1}\!/_{\lambda_{\max}}\left(of\ \ A^{-1}\cdot B\right)$$

Steady-State Performance

Based on the theoretical analysis, we have "conveniently" selected initial conditions and elements that simplify the solution in terms of the diagonal computations. We can also do this for the steady state responses for the EMSE and MSD or mean-squared deviation of the filter.

$$EMSE = \lim_{i\to\infty} E\left[\left\|e_a(i)\right\|^2\right]$$

$$MSD = \lim_{i\to\infty} E\left[\left\|\tilde{w}_i\right\|^2\right]$$

$$MSE = \sigma_v{}^2 + \lim_{i\to\infty} E\left[\left\|e_a(i)\right\|^2\right]$$

The values are

$$MSD = \lim_{i\to\infty} E\left[\left\|\tilde{w}_i\right\|^2\right] = \mu^2 \cdot \sigma_v^2 \cdot \lambda^T \cdot \left(I - \overline{F}\right)^{-1} \cdot q$$

$$MSD = \lim_{i\to\infty} E\left[\left\|\tilde{w}_i\right\|^2\right] = \mu^2 \cdot \sigma_v^2 \cdot \frac{\lambda^T \cdot D^{-1} \cdot q}{1 - \mu^2 \cdot \lambda^T \cdot D^{-1} \cdot \lambda}$$

where

$$D = 2 \cdot \mu \cdot \Lambda - \mu^2 \cdot \Lambda^2$$

and

$$EMSE = \lim_{i\to\infty} E\left[\left\|e_a(i)\right\|^2\right] = \mu^2 \cdot \sigma_v^2 \cdot \lambda^T \cdot \left(I - \overline{F}\right)^{-1} \cdot \lambda$$

$$EMSE = \lim_{i\to\infty} E\left[\left\|e_a(i)\right\|^2\right] = \mu^2 \cdot \sigma_v^2 \cdot \frac{\lambda^T \cdot D^{-1} \cdot \lambda}{1 - \mu^2 \cdot \lambda^T \cdot D^{-1} \cdot \lambda}$$

Appendix 9.D Convergence Time of Adaptive Filters

Appendix 9.E Learning Behavior of Adaptive Filters