Stability and Performance Analysis of the LMS Algorithm

Excess MSE and Misadjustment

Our expression for excess MSE,

$$J_{ex}(n) = \lambda^T \boldsymbol{\xi}(\nu)$$

can be approximated as follows. Assume $x_i(n+1) \approx \xi_\varphi(\nu)$ for large n. Then

$$\begin{aligned}
x_i(n) &\approx (1-2\mu\lambda_i + \alpha\mu^2\lambda_i^2)x_i(n) + \mu^2\lambda_i \sum_{j=1}^{M} \lambda_j x_j(n) + \mu^2 J_{\min}\lambda_i \\
&= \frac{\mu^2\lambda_i \sum_{j=1}^{M} \lambda_j x_j(n) + \mu^2 J_{\min}\lambda_i}{2\mu\lambda_i - \alpha\mu^2\lambda_i^2} \\
&= \frac{\mu^2\lambda_i \left[\sum_{j=1}^{M} \lambda_j x_j(n) + J_{\min}\right]}{2\mu\lambda_i - \alpha\mu^2\lambda_i^2} \\
&= \frac{\mu \left[J_{\min} + \sum_{j=1}^{M} \lambda_j x_j(n)\right]}{2-\alpha\mu\lambda_i}
\end{aligned}$$

Next assume $\sum_{j=1}^{M} \lambda_j x_j(n)$ is small compared to $J_{\min}$, Thus we have

$$x_i(n) \approx \frac{\mu\vartheta_{\mu\iota\nu}}{2-\mu\alpha\lambda_\iota}.$$

And with substitution,

$$\begin{aligned}
J_{ex}(n) &= \lambda^T \mathbf{x}(n) \\
&\approx J_{\min} \sum_{i=1}^{M} \frac{\mu\lambda_i}{2-\alpha\mu\lambda_i}.
\end{aligned}$$

Since we've assumed a large n in the approximation and our result is independent of n, we write

$$\begin{aligned}
J_{ex}(\infty) &\approx J_{\min} \sum_{i=1}^{M} \frac{\mu\lambda_i}{2-\alpha\mu\lambda_i} \\
&\approx \frac{\mu}{2_i} J_{\min} \sum_{i=1}^{M} \lambda_i \quad . \\
&= \frac{\mu}{2_i} J_{\min}\, \mathrm{tr}(\mathbf{R})
\end{aligned}$$

Property 3: The final value of the excess MSE is less than the MMSE if μ satisfies

$$\sum_{i=1}^{M} \frac{\mu\lambda_i}{2-\alpha\mu\lambda_i} < 1.$$

Clearly, the amount of excess MSE is up to the designer.

Property 4: The misadjustment is defined as

$$\begin{aligned} M &\equiv \frac{J_{ex}(\infty)}{J_{\min}} \\ &= \sum_{i=1}^{M} \frac{\mu\lambda_i}{2-\alpha\mu\lambda_i} \end{aligned}.$$

The misadjustment provides a measure of how close the LMS algorithm is operating to optimality. For example a misadjustment of 10% indicates that the actual (settled) MSE, $J(\infty)$ is 10% greater than the theoretical MMSE, $J_{\min}$.

We could further approximate the misadjustment as follows.

$$\begin{aligned} M &= \sum_{i=1}^{M} \frac{\mu\lambda_i}{2-\mu\alpha\lambda_i} \\ &\approx \mu \sum_{i=1}^{M} \frac{\lambda_i}{2} \\ &= \frac{\mu}{2}\operatorname{tr}(\mathbf{R}) \end{aligned}$$

where the last line is found in [Widrow and Sterns], [Solo and Kong], [Diniz], and [Haykin (9.95) see below]. We note that M is directly proportional to μ.

Working Rules and Comments

As a condition for convergence in the mean square, we choose μ such that

$$0 < \mu < \frac{2}{3\operatorname{tr}(\mathbf{R})}$$

but typically much, much smaller than the upper bound for colored inputs. Since in most applications we do not know **R** a priori, we must make an estimate of the upper bound.

$$\begin{aligned} \operatorname{tr}(\mathbf{R}) &= \sum_{j=1}^{M} R_{jj} = \sum_{j=1}^{M} r(0) \\ &= Mr(0) = M\sigma_u^2 \\ &= \sum_{k=0}^{M-1} E[|u(n-k)|^2] \\ &\approx \sum_{k=0}^{M-1} |u(n-k)|^2 \\ &= \mathbf{u}(n)^H \mathbf{u}(n) \\ &= \text{tap input power} \end{aligned}$$

So we have

$$0 < \mu < \frac{2}{3(\text{tap input power})}$$

and

$$M \approx \frac{\mu}{2}(\text{ταπ ινπυτ ποωερ})$$

A few comments on misadjustment.

1) Misadjustment (using our last approximation for M) increases linearly with filter length.
2) Settling time for LMS (how long it takes J_{ex} to get to a constant) is inversely proportional to μ. Thus we have a tradeoff between settling time (how fast our MSE gets to a constant) and misadjustment (how close to optimum we operate).
3) When comparing adaptive algorithms you **must** make sure misadjustment is the same (calibrated) for both algorithms in the experiments.!!!