Lecture 3

Lecture 3 includes the following:

- Eigenvalue spread of $\mathbf{R}$ and its influence on the convergence speed for the LMS.
- Variants of the LMS:
  - The Normalized LMS
  - The Leaky LMS
  - The Sign LMS
- The Echo Canceller

Example: Identification

$$\begin{bmatrix} b_0 & b_1 \end{bmatrix}^T \begin{bmatrix} \mathbf{w}_0 & \mathbf{w}_1 \end{bmatrix}^T$$

The cost function $\mathbf{J}(n)$ has its fastest growth in this direction.

Along the long axis, the convergence is the slowest since the growth of $\mathbf{J}(n)$ is minimal in this direction.

Adaptive Signal Processing 2011 Lecture 3

LMS and eigenvalue spread

The convergence speed for the LMS depends on the spread of the eigenvalues of $\mathbf{R}$. This spread is normally written as the ratio between the largest and the smallest eigenvalues.

$$\chi(\mathbf{R}) = \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}}$$

If $\chi(\mathbf{R}) > 1$, the contour plot of $\mathbf{J}(n)$ for the 2-dimensional case ($\mathbf{M} = 2$) is elliptic.

The convergence is fastest along the short axis of the ellipse since the cost function $\mathbf{J}(n)$ has its fastest growth in this direction.

Adaptive Signal Processing 2011 Lecture 3

Calculation of the correlations:

$$\begin{align*}
\mathbf{u}(n) &= \mathbf{u}(n-1) + \mathbf{v}_1(n) \\
\mathbf{E}[\mathbf{u}(n)[\mathbf{u}(n)-\mathbf{u}(n-1)]] &= \mathbf{E}[\mathbf{u}(n)\mathbf{v}_1(n)] = \sigma^2
\end{align*}$$

$$\begin{align*}
\mathbf{E}[\mathbf{u}(n-1)[\mathbf{u}(n)-\mathbf{u}(n-1)]] &= \mathbf{E}[\mathbf{u}(n-1)\mathbf{v}_1(n)] = 0
\end{align*}$$

$$\begin{bmatrix} 1 & -a & r(0) & r(1) \\
-1 & 1 & r(1) & 0 \end{bmatrix} \begin{bmatrix} \sigma^2_1 \\
r(0) \\
r(1) \\
1-\frac{1}{a^2} \end{bmatrix}$$

$$\begin{align*}
d(n) &= b_0\mathbf{u}(n) + b_1\mathbf{u}(n-1) + \mathbf{v}_2(n) \\
\mathbf{E}[\mathbf{u}(n)d(n)] &= \mathbf{E}[\mathbf{u}(n)[b_0\mathbf{u}(n) + b_1\mathbf{u}(n-1) + \mathbf{v}_2(n)]] \\
\mathbf{E}[\mathbf{u}(n-1)d(n)] &= \mathbf{E}[\mathbf{u}(n-1)[b_0\mathbf{u}(n) + b_1\mathbf{u}(n-1) + \mathbf{v}_2(n)]]
\end{align*}$$

$$\begin{bmatrix} r(0) \\
r(1) \\
r(0) \\
r(1) \end{bmatrix} \begin{bmatrix} b_0 \\
b_1 \\
b_0 \\
b_1 \end{bmatrix} = \mathbf{Rb}$$

$$\sigma^2 = \mathbf{E}[(b_0\mathbf{u}(n) + b_1\mathbf{u}(n-1) + \mathbf{v}_2(n))d(n)] = \mathbf{b}^T\mathbf{Rb} + \sigma^2$$

Adaptive Signal Processing 2011 Lecture 3
Let us assume that $b = [0.25, 0.50]^T$, and that $\sigma_b^2 = 0.1$. What we now want to achieve is identification of this filter, i.e. $\hat{w} = b$, by using the LMS algorithm.

We can vary the parameter $a$, $0 \leq |a| < 1$, in order to get different degrees of colouring of the exciting noise. For $a = 0$ the exciting noise is white, and $\chi(R) = 1$.

For this structure

$$J_{min} = \sigma_d^2 - \sigma_a^2 = b^T R b + \sigma_{\nu_2}^2 - w_d^T R w_\nu = \sigma_{\nu_2}^2$$

since the length of the adaptive filter matches the filter to be identified. Without additive noise in the form of $v_2(n)$, $J_{min}$ and thereby $J_{ex}$ becomes zero!
The Normalized LMS

For the standard LMS which we have looked at earlier, the step is proportional to \( u(n) \):

\[
\hat{w}(n+1) = \hat{w}(n) + \mu u(n)e^*(n)
\]

where \( 0 < \mu < 2 \), and \( a \) is a positive protection constant.

LMS with time-variable adaptation step

One way to reduce the misadjustment \( \mathcal{M} \) without affecting the convergence is to use a variable adaptation step,

\[
\hat{w}(n+1) = \hat{w}(n) + \alpha(n)u(n)e^*(n),
\]

where \( \alpha(n) \) for example can be

\[
\alpha(n) = \frac{1}{n+c}
\]

This method is useful for steady-state self-tuning, but cannot be used for tracking.

The comparison shows that for a coloured input the convergence speed depends on the start values of the filter.

The two choices of initial values that was investigated was \( \hat{w}(0) = [-1.0, -1.0]^T \) and \( \hat{w}(0) = [1.5, -1.0]^T \). When the input signal was uncorrelated, i.e. "white" (\( \alpha = 0 \)), there was no difference in the convergence speed between these two choices, while, when the input signal was coloured (\( \alpha = 0.8 \)), the convergence speed was dependent on the initial values of the filter.

Since \( \hat{w}(0) = [-1.0, -1.0]^T \) deviates from \( w_{\text{m}} \) mainly along the short axis of the ellips, this case converged faster.

This means that the gradient noise is amplified when the signal is strong.

One solution to this problem is to normalize the update \( \hat{w}(n) \) med \( ||u(n)||^2 = u^H(n)u(n) \):

\[
\hat{w}(n+1) = \hat{w}(n) + \frac{\tilde{\mu}}{a + ||u(n)||^2} u(n)e^*(n)
\]

where \( 0 < \tilde{\mu} < 2 \), and \( a \) is a positive protection constant.
LMS with independent adaptation steps

It is also possible to choose independent adaptation steps for each filter coefficient,
\[ \hat{w}(n+1) = \hat{w}(n) + \mu u(n)e^*(n) , \]
where \( M \) is given by
\[ M = \begin{bmatrix} \alpha_1 & 0 & 0 & \ldots & 0 \\ 0 & \alpha_2 & 0 & \ldots & 0 \\ 0 & 0 & \alpha_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \alpha_M \end{bmatrix} . \]
This method becomes more interesting when updating the filter coefficients in the frequency domain, where different frequency components can be assigned different step sizes.

The Leaky LMS

As for the standard LMS we start with the method of the Steepest descent,
\[ w(n+1) = w(n) + \frac{\mu}{2}[-\nabla J(n)] , \]
but we now instead use the cost function
\[ J(n) = E\{|e(n)|^2\} + \alpha||w(n)||^2 \]
\[ = E\{|e\gamma(n)c(n)\} + \alpha w^H(n)w(n) . \]
The gradient becomes
\[ \nabla J(n) = -2p + 2Rw(n) + 2\alpha w(n) . \]
With estimated statistics we get
\[ \nabla J(n) = -2p(n) + 2Rw(n) + 2\alpha w(n) . \]

The Leaky LMS

The update equation for the Leaky LMS is given by
\[ \hat{w}(n+1) = (1 - \mu\alpha) \hat{w}(n) + \mu u(n)e^*(n) . \]
For the Leaky LMS it can be shown that
\[ \lim_{n \to \infty} E\{\hat{w}(n)\} = (R + \alpha I)^{-1}p \]
i.e., it results in the same effect as if white noise with variance \( \alpha \) was added at the input. This increases the uncertainty in the input signal, which holds the filter coefficients back.

The phenomenon is referred to as prewhitening, and \( \alpha \) is denoted prewhitening parameter. The factor \((1 - \mu\alpha)\) is called Leakage factor, where \( \alpha \) is bounded by
\[ 0 \leq \alpha < \frac{1}{N} . \]
The Sign LMS

Again, starting from the method of the Steepest descent
\[ w(n+1) = w(n) + \mu \nabla J(n) \],
and the cost function
\[ J(n) = E\{|e(n)|\} \]
the gradient becomes
\[ \nabla J(n) = -\text{sign}[e(n)]u(n) . \]
With this gradient, the update equation becomes
\[ \hat{w}(n+1) = \hat{w}(n) + \alpha \cdot \text{sign}[e(n)]u(n) . \]

The Sign LMS is used in applications with extremely high requirements on computational complexity.

System: Echo canceller I

What to read

- Haykin chap. 5.9, 6.1–6.2
- Exercises: 4.6, 4.7, 4.8

Computer exercise, theme: Implement the NLMS in Matlab, and compare the NLMS to the LMS for echo cancellation of a speech signal (Echo canceller I)