Convergence Analysis of the RLS Algorithm

• Reference model (linear regression):

\[ d(n) = e_0(n) + w_0^H u(n) \]

where \( e_0(n) \) is white with variance \( \sigma^2 \).

• Recall the weight vector of RLS algorithm (in batch form):

\[ w(n) = \Phi^{-1}(n) z(n) \]

• Assuming \( \lambda = 1 \), we have

\[
\Phi(n) = \sum_{i=1}^{n} u(i)u^H(i) \\
z(n) = \sum_{i=1}^{n} u(i)d^*(i) \\
= \sum_{i=1}^{n} u(i)\left(e_0(i) + w_0^H u(i)\right)^* \\
= \sum_{i=1}^{n} u(i)e_0(i)^* + \Phi(n)w_0
\]

• Then, one can show:

\[ E(w(n)) = w_0 \]

(convergence in the mean value)

• **Note**: This convergence holds at any \( n \) provided \( \Phi(n) \) is invertible. If \( u(i) = 0 \) for \( i < 1 \), we need to wait until \( n \geq M \).
• To show the covariance of the weight vector, define

\[ \varepsilon(n) = w(n) - w_0 \]

\[ = \Phi^{-1}(n) \sum_{i=1}^{n} u(i)e_0^* (i) \]

• Then,

\[ K(n) \equiv E(\varepsilon(n)\varepsilon^H(n)) \]

\[ = ... \]

\[ = \sigma^2 E(\Phi^{-1}(n)) \]

• If \( u(1) \, u(2) \, \cdots \, u(n) \) are i.i.d., Gaussian of zero mean and \((M \times M)\) correlation matrix \( R \), then it can be shown (Appendix J.4) that

\[ K(n) = \sigma^2 \frac{1}{n-M-1} R^{-1}, \quad n > M + 1 \]

• Note: The covariance converges to zero linearly as \( n \) increases. The covariance is large if the smallest eigenvalue of \( R \) is small.
• To show the learning curve of the RLS algorithm, we consider the a priori estimation error:

\[ \xi(n) = d(n) - w^H(n-1)u(n) \]
\[ = e_0(n) + w_0^H u(n) - w^H(n-1)u(n) \]
\[ = e_0(n) - e^H(n-1)u(n) \]

• Assuming \( u(i) \) is i.i.d. and Gaussian, then

\[ J'(n) = E\left( \| \xi(n) \|^2 \right) \]
\[ = \sigma^2 + E\left( e^H(n-1)u(n)u^H(n)e(n-1) \right) \]
\[ = \sigma^2 + \text{tr}\{RK(n-1)\} \]
\[ = \sigma^2 + \text{tr}\left\{ \mathbf{R}\left( \frac{\sigma^2}{n-M-2} \mathbf{R}^{-1} \right) \right\} \]
\[ = \sigma^2 + \frac{M}{n-M-2} \sigma^2 \]

which is independent of the eigenvalues of \( \mathbf{R} \). It converges linearly. Zero excess error.
• Recall that for the LMS algorithm (see (9.81)):

\[ J(n) = \sum_{i=1}^{M} \gamma_i c_i^n + J_{\text{min}} + J_{\text{ex}}(\infty) \]

where

\[ c_i = (1 - \mu \lambda_i)^2 \]

\[ J_{\text{ex}}(\infty) = J_{\text{min}} \sum_{i=1}^{M} \frac{\mu \lambda_i}{2 - \mu \lambda_i} \]

which converges exponentially, but depending on the eigenvalue spread of \( R \) and the step size. There is always an excess error.