Stability and Performance Analysis of the LMS Algorithm

Convergence in the Mean Square (cont.)

Theorem (Gersgorin or Gerschgorin Circle Theorem) Let **A** be an $N \times N$ matrix with elements a_{ij} and let R_i denote the circle in the complex plan with center a_{ii} and radius $\sum_{j=1,\, j\neq i}^{N} |a_{ij}|$, i.e.

$$R_i = \left\{ z \in C \,\Big|\, |z - a_{ii}| \leq \sum_{j=1,\, j\neq i}^{N} |a_{ij}| \right\}$$

where C denotes the complex plane. The eigenvalues of A are contained within $R = \bigcup_{i=1}^{N} R_i$. Moreover, the union of any k of these circles that do not intersect the remaining $(n–k)$ must precisely contain k (counting multiplicities) of the eigenvlaues.

Proof. See *Numerical Analysis 4th ed* by Burden and Faires or *Matrix Analysis* by Horn and Johnson.

We can use Gerschgorin's theorem to bound the spectral radius of **B** (and hence eigenvalues) as follows.

Figure.

From the above figure we require for every row i

$$|b_{ii}| + \sum_{j=1,\, j\neq i}^{M} |b_{ij}| < 1$$

or equivalently

$$\sum_{j=1}^{M} |b_{ij}| < 1 \, .$$

With a bit of work, it can be shown that if

$$\boxed{\mu < \frac{2}{3\,\mathrm{tr}(\mathbf{R})}}$$

$\mathbf{x}(n)$ converges and hence

$$\lim_{n\to\infty} J(n) = \text{constant}$$

and thus we have convergence in the mean square for uncorrelated Gaussian signals.

In a number of texts a looser bound is (unfortunately) published because its derivation is simpler. In this case we have for each row i

$$\begin{aligned} 1 &> \sum_{j=1}^{M} |b_{ij}| \\ &> \left| \sum_{j=1}^{M} b_{ij} \right| \end{aligned}$$

which leads to

$$-1 < \sum_{j=1}^{M} b_{ij} < 1 \,.$$

Since

$$\sum_{j=1}^{M} b_{ij} = 1 - 2\mu\lambda_\iota + \alpha\mu^2\lambda_\iota^2 + \mu^2\lambda_\iota \sum_{\varphi=1}^{M} \lambda_\varphi$$

(where α = 2 if signals are real-valued and α = 1 if signals are complex-valued) the critical values of μ are those which the above equation approaches 1 since for any μ the expression is always positive, thus we need

$$-2\mu\lambda_i + \alpha\mu^2\lambda_i^2 + \mu^2\lambda_i \sum_{j=1}^{M} \lambda_j \approx 0 \,.$$

Simple manipulation yields

$$\alpha\mu^2\lambda_i^2 + \mu^2\lambda_i \sum_{j=1}^{M} \lambda_j \approx 2\mu\lambda_i$$

$$\alpha\mu\lambda_i + \mu \sum_{j=1}^{M} \lambda_j \approx 2$$

$$\mu \approx \frac{2}{\alpha\lambda_i + \sum_{j=1}^{M} \lambda_j} \approx \frac{2}{\sum_{j=1}^{M} \lambda_j}$$

and a looser bound

$$0 < \mu < \frac{2}{\tau\rho(\mathbf{P})} \,.$$

Again, we urge extreme caution with approximations to upper stability bounds.

Transient Behavior of the MSE

$$\begin{aligned} J(n) &= \vartheta_{\mu\iota\nu} + \vartheta_{\varepsilon\xi}(\nu) \\ &= \vartheta_{\mu\iota\nu} + \underbrace{\lambda^T \boldsymbol{\xi}(\nu)}_{\tau\rho\alpha\nu\sigma\iota\varepsilon\nu\tau} \end{aligned}$$

where

$$\mathbf{x}(n+1) = \mathbf{B}\boldsymbol{\xi}(\nu) + \mu^2 \vartheta_{\mu\iota\nu}\, \lambda$$

and

$$b_{ij} = \begin{cases} 1 - 2\mu\lambda_\iota + (\alpha+1)\mu^2\lambda_\iota^2, & \iota = \varphi \\ \mu^2\lambda_\iota\lambda_\varphi & \iota \neq \varphi \end{cases}.$$

For $x_i(n+1)$ this becomes

$$\begin{aligned} x_i(n+1) &= [1 - 2\mu\lambda_\iota + (\alpha+1)\mu^2\lambda_\iota^2]\xi_\iota(\nu) + \mu^2\lambda_\iota \sum_{\substack{\varphi=1 \\ \varphi\neq\iota}}^{M} \lambda_\varphi \xi_\varphi(\nu) + \mu^2 \vartheta_{\mu\iota\nu}\, \lambda_\iota \\ &= (1 - 2\mu\lambda_\iota + \alpha\mu^2\lambda_\iota^2)\, \xi_\iota(\nu) + \mu^2\lambda_\iota \sum_{\varphi=1}^{M} \lambda_\varphi \xi_\varphi(\nu) + \mu^2 \vartheta_{\mu\iota\nu}\, \lambda_\iota \end{aligned}$$

Property 2: The transient component of the MSE $J(n)$ dies out; that is, the LMS algorithm is convergent in the mean-square sense if and only if

$$0 < \mu < \frac{2}{3\,\mathrm{tr}(\mathbf{P})}.$$

[Compare to Haykin (9.83)]