# ECE 792-41 Homework 5

## Material Covered: Steepest Descent, Newton's Method, Least Mean-Squares (LMS)

**Problem 1** (Alternative derivation for steepest-descent) In this problem, we explore another way of deriving the steepest-descent algorithm. The inverse of a positive-definite matrix may be expanded in a series as

$$\mathbf{R}^{-1} = \mu \sum_{k=0}^{\infty} (\mathbf{I} - \mu \mathbf{R})^k$$

where $\mathbf{I}$ is the identity matrix and $\mu$ is a positive constant. To ensure that the series converges, the constant $\mu$ must lie inside the range $0 < \mu < \frac{2}{\lambda_{\max}}$, where $\lambda_{\max}$ is the largest eigenvalue of the matrix $\mathbf{R}$. By using this series expansion of the inverse of the correlation matrix in the Wiener–Hopf equation, develop the recursion

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \left[\mathbf{p} - \mathbf{R}\mathbf{w}(n)\right]$$

where

$$\mathbf{w}(n) = \mu \sum_{k=0}^{n-1} (\mathbf{I} - \mu \mathbf{R})^k \mathbf{p}$$

is the approximation to the Wiener solution for the tap-weight vector.

**Problem 2** (Modified LMS) The zero-mean output $d(n)$ of an unknown real-valued system is represented by the multiple linear regression model $d(n) = \mathbf{w}_0^T \mathbf{u}(n) + v(n)$ where $\mathbf{w}_0$ is the unknown but fixed parameter vector of the model, $\mathbf{u}(n)$ is the input vector (regressor), and $v(n)$ is the sample value of an immeasurable white-noise process of zero mean and variance $\sigma_v^2$. This real-valued system is tracked by a *modified LMS algorithm*, in which the tap-weight vector $\mathbf{w}(n)$ of the transversal filter is chosen so as to minimize the index of performance

$$J(\mathbf{w}, K) = E[e^{2K}(n)], \qquad K \in \mathbb{Z}_+.$$

**(a)** By using the instantaneous gradient vector, show that the new adaptation rule for the corresponding estimate of the tap-weight vector is

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu K \mathbf{u}(n) e^{2K-1}(n)$$

where $\mu$ is the step-size parameter and $e(n) = d(n) - \mathbf{w}^T(n)\mathbf{u}(n)$ is the estimation error.

**(b)** Assume that the weight-error vector

$$\epsilon(n) = \mathbf{w}_0 - \hat{\mathbf{w}}(n)$$

is close to zero and that $v(n)$ is independent of $\mathbf{u}(n)$. Show that

$$E[\epsilon(n+1)] = \left(\mathbf{I} - \mu K(2K-1)E[v^{2(K-1)}(n)]\mathbf{R}\right) E[\epsilon(n)]$$

where $\mathbf{R}$ is the correlation matrix of the input vector $\mathbf{u}(n)$.

**(c)** Show that the modified LMS algorithm described in part (a) coverages in the mean value if the step-size parameter $\mu$ satisfies the condition

$$0 < \mu < \frac{2}{K(2K-1)E[v^{2(K-1)}(n)]\lambda_{\max}}$$

where $\lambda_{\max}$ is the largest eigenvalue of matrix $\mathbf{R}$.

**(d)** For $K = 1$, show that the results given in parts (a), (b), and (c) reduce to those of the conventional LMS algorithm.

**Problem 3** (Leaky LMS) Consider the time-varying cost function

$$J(n) = |e(n)|^2 + \alpha\|\mathbf{w}(n)\|^2$$

where $\mathbf{w}(n)$ is the tap-weight vector of a transversal filter, $e(n)$ is the estimation error, and $\alpha$ is a constant. As usual, $e(n) = d(n) - \mathbf{w}^H(n)\mathbf{u}(n)$ where $d(n)$ is the desired response and $\mathbf{u}(n)$ is the tap-input vector. In the leaky LMS algorithm, the cost function $J(n)$ is minimized w.r.t. the weight vector $\mathbf{w}(n)$.

**(a)** Show that the time update for the tap-weight vector $\hat{\mathbf{w}}(n)$ is defined by

$$\hat{\mathbf{w}}(n+1) = (1 - \mu\alpha)\hat{\mathbf{w}}(n) + \mu\mathbf{u}(n)e^*(n).$$

**(b)** Using the small-step-size theory, show that

$$\lim_{n\to\infty} E[\hat{\mathbf{w}}(n)] = (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{p}$$

where $\mathbf{R}$ is the correlation matrix of the tap inputs and $\mathbf{p}$ is the cross-correlation vector between the tap inputs and the desired response. What is the condition for the algorithm to converge in the mean value?

**(c)** How would you modify the tap-input vector in the conventional LMS algorithm to obtain the equivalent result described in part (a)?

**Problem 4** The convergence ratio of an adaptive algorithm is defined in terms of the weight-error vector by

$$\eta(n) = \frac{E[\|\epsilon(n+1)\|^2]}{E[\|\epsilon(n)\|^2]}$$

Show that, for small $n$, the convergence ratio of the LMS algorithm for stationary inputs is given by

$$\eta(n) \approx (1 - \mu\sigma_u^2)^2, \quad n \text{ small.}$$

Assume that the correlation matrix of the tap-input vector $\mathbf{u}(n)$ is approximately equal to $\sigma_u^2\mathbf{I}$.

**Problem 5** (Adaptive step size) In much of the material presented on LMS filters, we focused attention on the use of a fixed step-size parameter. In this problem, we consider an adaptive LMS filter in which the step size is adaptively controlled. The motivation for such a modification is to improve the convergence behavior of the LMS filter.

**(a)** Define the gradient vector

$$\gamma(n) = \frac{\partial \hat{\mathbf{w}}(n)}{\partial \mu(n)}$$

where $\hat{\mathbf{w}}(n)$ is the weight vector estimated by an LMS filter with time-varying step-size parameter $\mu(n)$. Starting with the instantaneous cost function

$$J(n) = \frac{1}{2}|e(n)|^2$$

where $e(n)$ is the error signal produced by the LMS filter, show that

$$\gamma(n) = \left[\mathbf{I} - \mu(n)\mathbf{u}(n-1)\mathbf{u}^H(n-1)\right]\gamma(n-1) + \mathbf{u}(n-1)e^*(n-1)$$

where $\mathbf{I}$ is the identity matrix.

**(b)** Formulate the LMS algorithm with time-varying step size by the pair of update equations

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

and

$$\mu(n) = \mu(n-1) + \rho e(n)\gamma^H(n)\mathbf{u}(n)$$

where $\gamma(n)$ is as defined in part (a) and $\rho$ a small positive constant that controls the adaptation of $\mu(n)$.

**Problem 6** Consider an AR process $u(n)$ defined by the difference equation

$$u(n) = -a_1 u(n-1) - a_2 u(n-2) + v(n)$$

where $v(n)$ is an additive white noise of zero mean and variance $\sigma_v^2$. Assume $a_1 = 0.1$ and $a_2 = -0.8$.

**(a)** Calculate the noise variance $\sigma_v^2$ such that the AR process has unit variance. Generate different realizations of the process $u(n)$.

**(b)** Given the input $u(n)$, an LMS filter of length $M = 2$ is used to estimate the unknown AR parameters $a_1$ and $a_2$. The step-size parameter $\mu$ is assigned the value 0.05. Justify the use of this design value in the application of the small-step-size theory.

**(c)** For one realization of the LMS filter, compute the prediction error

$$f(n) = u(n) - \hat{u}(n)$$

and the two tap-weight errors

$$\epsilon_i(n) = -a_i - \hat{w}_i(n), \quad i = 1, 2.$$

Using power spectral plots of $f(n)$, $\epsilon_1(n)$, and $\epsilon_2(n)$, show that $f(n)$ behaves as white noise, whereas $\epsilon_1(n)$ and $\epsilon_2(n)$ behave as low-pass processes.

**(d)** Compute the ensemble-average learning curve of the LMS filter by averaging the squared value of the prediction error $f(n)$ over an ensemble of 100 different realizations of the filter.

**(e)** Using the small-step-size theory, compute the theoretical learning curve of the LMS filter and compare your result against the measured result of part (d).

**Problem 7** (Steepest descent method and Newton's method) Complete Problem 1 in this url: `http://www.cs.umd.edu/~elman/660.12/hw4/hw4.pdf`