Deptartment of Electrical Communication Engineering
Indian Institute of Science

E9 211 Adaptive Signal Processing

4 December 2019, 14:00–17:00, Final Exam Solutions

This exam has three questions (50 points).

One A4 cheat sheet is allowed. No other materials will be allowed.

You need to submit **Homework 3 by the deadline 10 December 2019** and should have already submitted Homeworks 1 and 2 to pass this course. If you have not yet submitted Homeworks 1 and 2, do it before 8 December 2019.

Question 1 (15 points)

Consider a filter $\mathbf{w}$, which is applied to $\mathbf{x}_k$ to estimate the desired signal d_k as

$$\hat{d}_k = \mathbf{w}^H \mathbf{x}_k.$$

Assume that the desired signal has unit power, i.e., $E\{|d_k|^2\} = 1$. In this problem, we will find the filter by optimizing the following cost functions:

$$J(\mathbf{w}) = E\{|d_k - \mathbf{w}^H \mathbf{x}_k|^2\}$$

and

$$J'(\mathbf{w}) = E\{|d_k - \mathbf{w}^H \mathbf{x}_k|^2\} + \beta \|\mathbf{w}\|_2^2,$$

where $\beta > 0$. To answer this question, use the following definitions $E\{\mathbf{x}_k \mathbf{x}_k^H\} = \mathbf{R}_x$ and $E\{\mathbf{x}_k \bar{d}_k\} = \mathbf{r}_{xd}$.

(2 pts) (a) Find the optimal filter that minimizes the cost $J(\mathbf{w})$. Compute the resulting minimum cost $J(\mathbf{w}_{\text{opt}})$.

(4 pts) (b) Find the optimal filter that minimizes the cost $J'(\mathbf{w})$. Compute the resulting minimum cost $J'(\mathbf{w}_{\text{opt}})$.

(2 pts) (c) Show that $J'(\mathbf{w}_{\text{opt}}) > J(\mathbf{w}_{\text{opt}})$.

(2 pts) (d) Give the steepest-descent update equations to find the minimizer of $J(\mathbf{w})$. What is the maximum step size μ that can be used so that the steepest-descent algorithm converges?

(4 pts) (e) Derive the steepest-descent algorithm for minimizing $J'(\mathbf{w})$ and determine the condition on the step size so that the iterations converge.

(1 pts) (f) When is the algorithm in (e) more useful that the algorithm in (d)?

Solutions

(a) The $\mathbf{w}_{\text{opt}}$ is obtained by setting the gradient of $J(\mathbf{w}) = \mathbf{w}^H\mathbf{R}_x\mathbf{w} - \mathbf{w}^H\mathbf{r}_{xd} - \mathbf{r}_{xd}^H\mathbf{w} + 1$ towards $\mathbf{w}^H$ to zero. This leads to $\mathbf{w}_{\text{opt}} = \mathbf{R}_x^{-1}\mathbf{r}_{xd}^H$.

The minimum cost is $J(\mathbf{w}_{\text{opt}}) = 1 - \mathbf{r}_{xd}\mathbf{R}_x^{-1}\mathbf{r}_{xd}$.

(b) The modified cost function

$J'(\mathbf{w}) = \mathbf{w}^H\mathbf{R}_x\mathbf{w} - \mathbf{w}^H\mathbf{r}_{xd} - \mathbf{r}_{xd}^H\mathbf{w} + 1 + \beta\mathbf{w}^H\mathbf{w} = \mathbf{w}^H(\mathbf{R}_x + \beta\mathbf{I})\mathbf{w} - \mathbf{w}^H\mathbf{r}_{xd} - \mathbf{r}_{xd}^H\mathbf{w} + 1$

has a minimum at $\mathbf{w}_{\text{opt}} = (\mathbf{R}_x + \beta\mathbf{I})^{-1}\mathbf{r}_{xd}^H$.

The minimum cost is $J'(\mathbf{w}_{\text{opt}}) = 1 - \mathbf{r}_{xd}(\mathbf{R}_x + \beta\mathbf{I})^{-1}\mathbf{r}_{xd}$.

(c) Since the eigenvalue spread of $\mathbf{R}_x + \beta\mathbf{I}$ is smaller than the eigenvalue spread of $\mathbf{R}_x$, $J'(\mathbf{w}_{\text{opt}}) > J(\mathbf{w}_{\text{opt}})$.

(d) The steepest-descent iterations are given by

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} - \mu[\mathbf{R}_x\mathbf{w}^{(k)} - \mathbf{r}_{xd}^H].$$

These iterations will converge for $0 < \mu < 2/\lambda_{\max}$, where $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{R}_x$.

(e) For the modified cost function, the steepest-descent iterations are given by

$$\begin{aligned}\mathbf{w}^{(k+1)} &= \mathbf{w}^{(k)} - \mu[(\mathbf{R}_x + \beta\mathbf{I})\mathbf{w}^{(k)} - \mathbf{r}_{xd}^H] \\ &= (1 - \mu\beta)\mathbf{w}^{(k)} - \mu[\mathbf{R}_x\mathbf{w}^{(k)} - \mathbf{r}_{xd}^H].\end{aligned}$$

The condition on the step size for the iterations to converge is $0 < \mu < 2/(\lambda_{\max} + \beta)$.

(f) The algorithm in (e) is useful when $\mathbf{R}_x$ is ill-conditioned and by diagonally loading $\mathbf{R}_x$ with $\beta > 0$ we are improving the condition number.

Question 2 (15 points)

Let us consider the problem of estimating an unknown constant x given measurements that are corrupted by uncorrelated, zero mean noise $v(n)$ that has a variance σ_v^2. The measurement equation is

$$y(n) = x(n) + v(n)$$

Since the value of x does not change with time n, we have

$$x(n) = x(n-1).$$

(2 pts) (a) Assume that we are at time step N and gathered all the measurements $\{y(1), y(2), \cdots y(N)\}$. Compute the least squares solution for x.

(5 pts) (b) Give the recursive least squares (RLS) update equations assuming that the observations arrive sequentially.

(5 pts) (c) Derive now the Kalman filter update equations.

(3 pts) (d) Let us denote $P(n-1|n-1) = P(n-1)$. Show that

$$P(n) = \frac{P(0)\sigma_v^2}{nP(0) + \sigma_v^2}.$$

Solutions

(a) We can write the measurements as

$$\mathbf{y} = \mathbf{1}x + \mathbf{v}$$

where $\mathbf{y} = [y(1), y(2), \cdots y(N)]^T$ and $\mathbf{v} = [v(1), v(2), \cdots v(N)]^T$. Then the least squares solution for $x = (\mathbf{1}^T\mathbf{1})^{-1}\mathbf{1}^T\mathbf{y} = \frac{1}{N}\sum_{n=1}^{N} y(n)$.

(b) The RLS update equations are

$$\begin{aligned} P_{n+1} &= P_n - \frac{P_n^2}{1+P_n} = \frac{P_n}{1+P_n} \\ \theta_{n+1} &= \theta_n + y(n+1) \\ \hat{x}(n+1) &= P_{n+1}\theta_{n+1} \end{aligned}$$

The iterations are initialized with $P_1 = 1$ and $\theta_1 = y(1)$.

(c) Kalman filter update equations with $\mathbf{A}(n) = \mathbf{C}(n) = 1$, $\mathbf{Q}_w(n) = 0$, and $\mathbf{Q}_v(n) = \sigma_v^2$ are

$$\begin{aligned} \hat{x}(n|n-1) &= \hat{x}(n-1|n-1) \\ P(n|n-1) &= P(n-1|n-1) = P(n-1) \\ K(n) &= P(n-1)[P(n-1) + \sigma_v^2]^{-1} \\ P(n) &= \frac{P(n-1)\sigma_v^2}{P(n-1) + \sigma_v^2}. \end{aligned}$$

(d) Using the difference equation $P(n) = \frac{P(n-1)\sigma_v^2}{P(n-1)+\sigma_v^2}$, and substituting $n = 1, 2, \ldots$, we can see that

$$P(n) = \frac{P(0)\sigma_v^2}{nP(0) + \sigma_v^2}$$

.

Question 3 (20 points)

In many signal processing applications, it is important to design a filter with *linear phase* without which the phase distortion introduced by the filter might severely degrade the estimated signal. Therefore, we would like to design an adaptive linear phase filter $\mathbf{w}_k = [w_k(0), w_k(1), \ldots, w_k(P)]^T$ whose weights at each time k satisfy the following symmetry constraint

$$w_k(n) = w_k(P-n), \quad n = 0, 1, \ldots, P.$$

(3 pts) (a) Formulate the symmetry constraint as a linear equality constraint $\mathbf{C}^H\mathbf{w}_k = \mathbf{f}$. What will be $\mathbf{C}$ and $\mathbf{f}$ for the special case with $P = 2$?

(6 pts) (b) Find the optimal solution to the constrained optimization problem

$$\text{minimize} \quad E\{|d_k - \mathbf{w}^H\mathbf{x}_k|^2\} \quad \text{subject to} \quad \mathbf{c}^H\mathbf{w} = f.$$

Notice that $\mathbf{c}$ is a vector and f is a scalar.

(6 pts) (c) Derive a linearly constrained adaptive LMS filter $\mathbf{w}_k$ that approximates the optimal solution in (b) by minimizing the instantaneous error as

$$\text{minimize} \quad |d_k - \mathbf{w}_k^H \mathbf{x}_k|^2 \quad \text{subject to} \quad \mathbf{c}^H \mathbf{w}_k = f.$$

Hint: Use the extended cost function with a Lagrange multiplier and enforce the condition that each successive weight vector, including the initial condition, satisfies the linear equality constraint.

(5 pts) (d) It is possible to eliminate the equality constraint by modifying the input signal $\mathbf{x}_k$ to $\mathbf{z}_k$. The vector $\mathbf{z}_k$ will be of smaller length as compared to $\mathbf{x}_k$. For $P = 2$, what is $\mathbf{z}_k$? Give the standard (unconstrained) LMS update equations for computing the adaptive filter $\tilde{\mathbf{w}}_k$ that will be applied on $\mathbf{z}_k$ to estimate the desired signal as $\tilde{\mathbf{w}}_k^H \mathbf{z}_k$.

Solutions

(a) In general, the size of $\mathbf{C}$ will be $\lceil \frac{P+1}{2} \rceil \times P$. For the special case of $P = 2$, we will have

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$

and $\mathbf{f} = \mathbf{0}$. This can be generalized to arbitrary P.

(b) Using the method of Lagrangian multipliers, the cost function is given by

$$J(\mathbf{w}, \lambda) = \mathbf{w}^H \mathbf{R}_x \mathbf{w} - \mathbf{r}_{xd}^H \mathbf{w} - \mathbf{w}^H \mathbf{r}_{xd} + 1 + \lambda(\mathbf{w}^H \mathbf{c} - f).$$

Setting the derivative of the cost function w.r.t. $\mathbf{w}^H$ to zero, we get

$$\mathbf{w}(\lambda) = \mathbf{R}_x^{-1} \mathbf{r}_{xs} + \lambda \mathbf{R}_x^{-1} \mathbf{c}.$$

From the constraint equation, we get $\lambda = (\mathbf{c}^H \mathbf{R}_x^{-1} \mathbf{c})^{-1} [f - \mathbf{r}_{xd}^H \mathbf{R}_x^{-1} \mathbf{c}]$. Substituting in $\mathbf{w}(\lambda)$, we get

$$\mathbf{w}_{\text{opt}} = \mathbf{R}_x^{-1} \mathbf{r}_{xs} + (\mathbf{c}^H \mathbf{R}_x^{-1} \mathbf{c})^{-1} \mathbf{R}_x^{-1} \mathbf{c} [f - \mathbf{r}_{xd}^H \mathbf{R}_x^{-1} \mathbf{c}].$$

(c) To derive a linearly constrained adaptive LMS filter, we start with the steepest descent method using the gradient of $J(\mathbf{w}, \lambda)$ towards $\mathbf{w}^H$ as

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mu[\mathbf{R}_x \mathbf{w}_k - \mathbf{r}_{xs} + \lambda_k \mathbf{c}].$$

Since each iterate should satisfy the equality constraint, we have

$$\mathbf{c}^H \mathbf{w}_{k+1} = f \Rightarrow \lambda_k = \frac{1}{\mu} \left(\mathbf{c}^H \mathbf{c} \right)^{-1} \left([\mathbf{c}^H - \mu \mathbf{c}^H \mathbf{R}_x] \mathbf{w}_k - [f - \mu \mathbf{c}^H \mathbf{r}_{xs}] \right).$$

Substituting for λ_k in $\mathbf{w}_{k+1}$, we get

$$\mathbf{w}_{k+1} = \mathbf{P}[\mathbf{w}_k - \mu(\mathbf{R}_x \mathbf{w}_k - \mathbf{r}_{xs})] + \mathbf{g} f$$

where $\mathbf{P} = [\mathbf{I} - \frac{\mathbf{c}\mathbf{c}^H}{\|\mathbf{c}\|_2^2}]$ and $\mathbf{g} = \frac{\mathbf{c}}{\|\mathbf{c}\|_2^2}$. Finally, wsing the instantaneous data we arrive at the LMS update equation

$$\mathbf{w}_{k+1} = \mathbf{P}[\mathbf{w}_k - \mu(\mathbf{x}_k \mathbf{x}_k \mathbf{w}_k - \mathbf{x}_k \bar{s}_k)] + \mathbf{g} f.$$

(d) Since the weight vector will have the structure $\mathbf{w} = [w(0), w(1), w(0)]^T$, the input $\mathbf{x}_k = [x_k(0), x_k(1), x_k(2)]^T$ can be modified as $\mathbf{z}_k = [x_k(0) + x_k(2), x_k(1)]^T$ and the standard (unconstrained) LMS may be used.