ECE/MATH 520, Spring 2008

Homework Problems 4

Solutions (version: March 25, 2008, 11:13)

12.3 Suppose that we perform an experiment to calculate the gravitational constant g as follows. We drop a ball from a certain height, and measure its distance from the original point at certain time instants. The results of the experiment are shown in the following table:

Time (seconds)	1.00	2.00	3.00
Distance (meters)	5.00	19.5	44.0

The equation relating the distance s and the time t at which s is measured is given by

$$s = \frac{1}{2} g t^2.$$

a. Find a least-squares estimate of g using the experimental results from the above table.

b. Suppose that we take an additional measurement at time 4.00, and obtain a distance of 78.5. Use the recursive least-squares algorithm to calculate an updated least-squares estimate of g.

Ans.: a. We form

$$\boldsymbol{A} = \begin{bmatrix} 1^2/2 \\ 2^2/2 \\ 3^2/2 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 5.00 \\ 19.5 \\ 44.0 \end{bmatrix}.$$

The least squares estimate of g is then given by

$$g = (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{A}^T \boldsymbol{b} = 9.776.$$

b. We start with $\boldsymbol{P}_0 = 0.040816$, and $\boldsymbol{x}^{(0)} = 9.776$. We have $\boldsymbol{a}_1 = 4^2/2 = 8$, and $\boldsymbol{b}^{(1)} = 78.5$. Using the RLS formula, we get $\boldsymbol{x}^{(1)} = 9.802$, which is our updated estimate of g.

12.9 Suppose that we take measurements of a sinusoidal signal $y(t) = \sin(\omega t + \theta)$ at times $t_1, \ldots, t_p$, and obtain values $y_1, \ldots, y_p$, where $-\pi/2 \leq \omega t_i + \theta \leq \pi/2$, $i = 1, \ldots, p$, and the t_i are not all equal. We wish to determine the values of the frequency ω and phase θ.

a. Express the problem as a system of linear equations.

b. Find the least-squares estimate of ω and θ based on part a. Use the following notation:

$$\begin{aligned}
\overline{T} &= \frac{1}{p}\sum_{i=1}^{p} t_i, \\
\overline{T^2} &= \frac{1}{p}\sum_{i=1}^{p} t_i^2, \\
\overline{TY} &= \frac{1}{p}\sum_{i=1}^{p} t_i \arcsin y_i, \\
\overline{Y} &= \frac{1}{p}\sum_{i=1}^{p} \arcsin y_i.
\end{aligned}$$

Ans.: a. We wish to find ω and θ such that

$$\begin{aligned}
\sin(\omega t_1 + \theta) &= y_1 \\
&\vdots \\
\sin(\omega t_p + \theta) &= y_p.
\end{aligned}$$

Taking $\arcsin$, we get the following system of linear equations:

$$\begin{aligned}
\omega t_1 + \theta &= \arcsin y_1 \\
&\vdots \\
\omega t_p + \theta &= \arcsin y_p.
\end{aligned}$$

b. We may write the system of linear equations in part a as $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, where

$$\boldsymbol{A} = \begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_p & 1 \end{bmatrix}, \qquad \boldsymbol{x} = \begin{bmatrix} \omega \\ \theta \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} \arcsin y_1 \\ \vdots \\ \arcsin y_p \end{bmatrix}.$$

Since the t_i are not all equal, the first column of $\boldsymbol{A}$ is not a scalar multiple of the second column. Therefore, rank $\boldsymbol{A} = 2$. Hence, the least squares solution is

$$\begin{aligned}
\boldsymbol{x} &= (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b} \\
&= \begin{bmatrix} \sum_{i=1}^{p} t_i^2 & \sum_{i=1}^{p} t_i \\ \sum_{i=1}^{p} t_i & p \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{p} t_i \arcsin y_i \\ \sum_{i=1}^{p} \arcsin y_i \end{bmatrix} \\
&= \begin{bmatrix} \overline{T^2} & \overline{T} \\ \overline{T} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \overline{TY} \\ \overline{Y} \end{bmatrix} \\
&= \frac{1}{\overline{T^2} - (\overline{T})^2} \begin{bmatrix} 1 & -\overline{T} \\ -\overline{T} & \overline{T^2} \end{bmatrix} \begin{bmatrix} \overline{TY} \\ \overline{Y} \end{bmatrix} \\
&= \frac{1}{\overline{T^2} - (\overline{T})^2} \begin{bmatrix} \overline{TY} - (\overline{T})(\overline{Y}) \\ -(\overline{T})(\overline{TY}) + (\overline{T^2})(\overline{Y}) \end{bmatrix}.
\end{aligned}$$

12.11 Consider the affine function $f : \mathbb{R}^n \to \mathbb{R}$ of the form $f(\boldsymbol{x}) = \boldsymbol{a}^T\boldsymbol{x} + c$, where $\boldsymbol{a} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

a. We are given a set of p pairs $(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_p, y_p)$, where $\boldsymbol{x}_i \in \mathbb{R}^n$, $y_i \in \mathbb{R}$, $i = 1, \ldots, p$. We wish to find the affine function of best fit to these points, where "best" is in the sense of minimizing the total square error

$$\sum_{i=1}^{p}(f(\boldsymbol{x}_i) - y_i)^2.$$

Formulate the above as an optimization problem of the form: minimize $\|\boldsymbol{A}\boldsymbol{z} - \boldsymbol{b}\|^2$ with respect to $\boldsymbol{z}$. Specify the dimensions of $\boldsymbol{A}$, $\boldsymbol{z}$, and $\boldsymbol{b}$.

b. Suppose that the points satisfy

$$\boldsymbol{x}_1 + \cdots + \boldsymbol{x}_p = \boldsymbol{0}$$

and

$$y_1\boldsymbol{x}_1 + \cdots + y_p\boldsymbol{x}_p = \boldsymbol{0}.$$

Find the affine function of best fit in this case, assuming it exists and is unique.

Ans.: a. Write

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{x}_1^T & 1 \\ \vdots & \vdots \\ \boldsymbol{x}_p^T & 1 \end{bmatrix} \in \mathbb{R}^{p\times(n+1)}, \qquad \boldsymbol{z} = \begin{bmatrix} \boldsymbol{a} \\ c \end{bmatrix} \in \mathbb{R}^{n+1}, \qquad \boldsymbol{b} = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} \in \mathbb{R}^p.$$

The objective function can then be written as $\|\boldsymbol{A}\boldsymbol{z} - \boldsymbol{b}\|^2$.

b. Let $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p]^T \in \mathbb{R}^{p\times n}$, and $\boldsymbol{e} = [1, \ldots, 1]^T \in \mathbb{R}^p$. Then we may write $\boldsymbol{A} = [\boldsymbol{X}\ \boldsymbol{e}]$. The solution to the problem is $(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$. But

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{bmatrix} \boldsymbol{X}^T\boldsymbol{X} & \boldsymbol{X}^T\boldsymbol{e} \\ \boldsymbol{e}^T\boldsymbol{X} & p \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}^T\boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{0}^T & p \end{bmatrix}$$

since $\boldsymbol{X}^T\boldsymbol{e} = \boldsymbol{x}_1 + \cdots + \boldsymbol{x}_p = \boldsymbol{0}$ by assumption. Also,

$$\boldsymbol{A}^T\boldsymbol{y} = \begin{bmatrix} \boldsymbol{X}^T\boldsymbol{y} \\ \boldsymbol{e}^T\boldsymbol{y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{e}^T\boldsymbol{y} \end{bmatrix}$$

since $\boldsymbol{X}^T\boldsymbol{y} = y_1\boldsymbol{x}_1 + \cdots + y_p\boldsymbol{x}_p = \boldsymbol{0}$ by assumption. Therefore, the solution is given by

$$\boldsymbol{z}^* = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b} = \begin{bmatrix} (\boldsymbol{X}^T\boldsymbol{X})^{-1} & \boldsymbol{0} \\ \boldsymbol{0}^T & 1/p \end{bmatrix} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{e}^T\boldsymbol{y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \frac{1}{p}\boldsymbol{e}^T\boldsymbol{y} \end{bmatrix}.$$

The affine function of best fit is the constant function $f(\boldsymbol{x}) = c$, were

$$c = \frac{1}{p}\sum_{i=1}^{p} y_i.$$

12.13 Consider a discrete time linear system $x_{k+1} = ax_k + bu_k$, where u_k is the input at time k, x_k is the output at time k, and $a, b \in \mathbb{R}$ are system parameters. Suppose that we apply a constant input $u_k = 1$ for all $k \geq 0$, and measure the first 4 values of the output to be $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $x_3 = 8$. Find the least-squares estimate of a and b based on the above data.

Ans.: We pose the problem as a least squares problem: $\text{minimize}\, \|\boldsymbol{Ax} - \boldsymbol{b}\|^2$ where $\boldsymbol{x} = [a, b]^T$, and

$$\boldsymbol{A} = \begin{bmatrix} x_0 & 1 \\ x_1 & 1 \\ x_2 & 1 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

We have

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{bmatrix} \sum_{i=0}^2 x_i^2 & \sum_{i=0}^2 x_i \\ \sum_{i=0}^2 x_i & 3 \end{bmatrix}, \qquad \boldsymbol{A}^T\boldsymbol{b} = \begin{bmatrix} \sum_{i=0}^2 x_i x_{i+1} \\ \sum_{i=0}^2 x_{i+1} \end{bmatrix}.$$

Therefore, the least squares solution is

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^2 x_i^2 & \sum_{i=0}^2 x_i \\ \sum_{i=0}^2 x_i & 3 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=0}^2 x_i x_{i+1} \\ \sum_{i=0}^2 x_{i+1} \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 18 \\ 11 \end{bmatrix} = \begin{bmatrix} 7/2 \\ 1/6 \end{bmatrix}.$$

12.14 Consider a discrete time linear system $x_{k+1} = ax_k + bu_k$, where u_k is the input at time k, x_k is the output at time k, and $a, b \in \mathbb{R}$ are system parameters. Given the first $n+1$ values of the impulse response $h_0, \ldots, h_n$, find the least squares estimate of a and b. You may assume that at least one h_k is nonzero. **Note:** The *impulse response* is the output sequence resulting from an input of $u_0 = 1$, $u_k = 0$ for $k \neq 0$, and zero initial condition $x_0 = 0$.

Ans.: We pose the problem as a least squares problem: $\text{minimize}\, \|\boldsymbol{Ax} - \boldsymbol{b}\|^2$ where $\boldsymbol{x} = [a, b]^T$, and

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ h_1 & 0 \\ \vdots & \vdots \\ h_{n-1} & 0 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}.$$

We have

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{bmatrix} \sum_{i=1}^{n-1} h_i^2 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \boldsymbol{A}^T\boldsymbol{b} = \begin{bmatrix} \sum_{i=1}^{n-1} h_i h_{i+1} \\ h_1 \end{bmatrix}.$$

The matrix $\boldsymbol{A}^T\boldsymbol{A}$ is nonsingular because we assume that at least one h_k is nonzero. Therefore, the least squares solution is

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n-1} h_i^2 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{n-1} h_i h_{i+1} \\ h_1 \end{bmatrix} = \begin{bmatrix} (\sum_{i=1}^{n-1} h_i h_{i+1})/(\sum_{i=1}^{n-1} h_i^2) \\ h_1 \end{bmatrix}.$$

Optional Write MATLAB routines to implement the RLS algorithm and various random search algorithms (e.g., the genetic algorithm). Test your routines on various examples.

22.6a Consider the problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\|\boldsymbol{x}\|^2 \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \end{array}$$

where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $m \leq n$, and rank $\boldsymbol{A} = m$. Let $\boldsymbol{x}^*$ be the solution. Suppose we solve the problem using the penalty method, with the penalty function

$$P(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2.$$

Let $\boldsymbol{x}_\gamma^*$ be the solution to the associated unconstrained problem with the penalty parameter $\gamma > 0$, that is, $\boldsymbol{x}_\gamma^*$ is the solution to

$$\text{minimize}\, \frac{1}{2}\|\boldsymbol{x}\|^2 + \gamma\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2.$$

a. Suppose

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 \end{bmatrix}, \qquad \boldsymbol{b} = [1].$$

Verify that $\boldsymbol{x}_\gamma^*$ converges to the solution $\boldsymbol{x}^*$ of the original constrained problem as $\gamma \to \infty$.

Ans.: We have

$$\frac{1}{2}\|\boldsymbol{x}\|^2 + \gamma\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2 = \frac{1}{2}\boldsymbol{x}^T \begin{bmatrix} 1+2\gamma & 2\gamma \\ 2\gamma & 1+2\gamma \end{bmatrix} \boldsymbol{x} - \boldsymbol{x}^T \begin{bmatrix} 2\gamma \\ 2\gamma \end{bmatrix} + \gamma.$$

The above is a quadratic with positive definite Hessian. Therefore, the minimizer is

$$\begin{aligned} \boldsymbol{x}_\gamma^* &= \begin{bmatrix} 1+2\gamma & 2\gamma \\ 2\gamma & 1+2\gamma \end{bmatrix}^{-1} \begin{bmatrix} 2\gamma \\ 2\gamma \end{bmatrix} \\ &= \frac{1}{2+1/2\gamma} \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \end{aligned}$$

Hence,

$$\lim_{\gamma\to\infty} \boldsymbol{x}_\gamma^* = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The solution to the original constrained problem is

$$\boldsymbol{x}^* = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1}\boldsymbol{b} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$