K. Comanor May 8, 2007
EE103

Midterm Exam Solutions

Problem 1 (20 points)

Let $p(t) = x_0 + x_1 t + x_2 t^2$, and suppose $p(1) = 0$, $p'(1) = 7$, $p''(1) = 4$.

a. Write the 3×3 linear system $Ax = b$ that solves for the coefficients x_0, x_1, and x_2 of $p(t)$.

b. Is A non-singular? Explain why or why not.

c. If A is non-singular, determine $p(t)$. If A is singular, find a non-zero vector x such that $Ax = 0$.

Solution.
a. $p'(t) = x_1 + 2x_2 t$, and $p''(t) = 2x_2$. Using these facts, we can write the linear system that solves for the coefficients x_0, x_1, and x_2 as

$$\begin{bmatrix} 1 & t & t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} p(t) \\ p'(t) \\ p''(t) \end{bmatrix}$$

Using the given values for $p(t)$, $p'(t)$, and $p''(t)$ at $t = 1$, we can write this linear system as

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 7 \\ 4 \end{bmatrix}.$$

b. The coefficient matrix A is upper triangular with all non zero entries on its diagonal. Hence, it is non-singular.

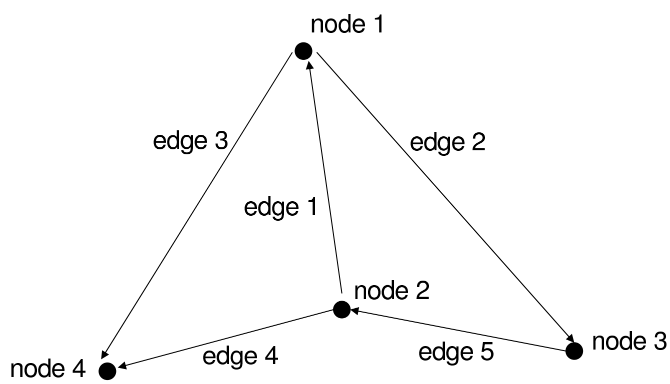
c. Use backward substitution to solve for x_0, x_1, and x_2.
$2x_2 = 4 \Rightarrow x_2 = 2$
$x_1 + 2x_2 = 7 \Rightarrow x_1 + 4 = 7 \Rightarrow x_1 = 3$
$x_0 + x_1 + x_2 = 0 \Rightarrow x_0 + 2 + 3 = 0 \Rightarrow x_0 = -5$

Thus, the polynomial is $p(t) = -5 + 3t + 2t^2$.

Problem 2 (20 points)

Below we define two functions g that we want to minimize (see equations (1) and (2)). For each of these two functions, explain how you would use a linear least-squares method to minimize g. You should express the problem in the standard form of minimizing $\|Ax - b\|^2$, and give the matrix A, the vector of variables x, and the right hand side b.
Note: The description of the problem is long, but all the information you need is summarized in the definitions of g (*i.e.*, the equations (1) and (2)). The rest of the discussion is background and can be read quickly.

The figure shows a *leveling network* used to determine the vertical height (or elevation) of three landmarks.



The nodes in the graph represent four points on a map. The variables in the problem are the heights h_1, h_2, h_3 of the first three points with respect to the fourth, which will be used as reference (in other words we take $h_4 = 0$). Each of the five edges in the graph represents a measurement of the height difference of its two end points. For example, edge 1 from node 2 to node 1 indicates a measurement of $h_1 - h_2$. We will denote the measured values of the five height differences as ρ_1, ρ_2, ρ_3, ρ_4, ρ_5. Our goal is to determine h_1, h_2, and h_3 from the five measurements. We distinguish two situations.

1. There is a small error in each measurement. The measured values are

$$\begin{aligned}
\rho_1 &= h_1 - h_2 + v_1 \\
\rho_2 &= h_3 - h_1 + v_2 \\
\rho_3 &= -h_1 + v_3 \\
\rho_4 &= -h_2 + v_4 \\
\rho_5 &= h_2 - h_3 + v_5
\end{aligned}$$

   where $v_1, \ldots, v_5$ are small and unknown measurement errors. To estimate the heights, we minimize the function

$$\begin{aligned}
g(h_1, h_2, h_3) &= (h_1 - h_2 - \rho_1)^2 + (h_3 - h_1 - \rho_2)^2 + (-h_1 - \rho_3)^2 \\
&\quad + (-h_2 - \rho_4)^2 + (h_2 - h_3 - \rho_5)^2. \qquad (1)
\end{aligned}$$

   The variables are h_1, h_2, h_3. The numbers ρ_i are given.

2. In addition to the measurement errors v_i, there is a systematic error, or offset, w in the measurement device. We actually measure

$$\begin{aligned}
\rho_1 &= h_1 - h_2 + v_1 + w \\
\rho_2 &= h_3 - h_1 + v_2 + w \\
\rho_3 &= -h_1 + v_3 + w \\
\rho_4 &= -h_2 + v_4 + w \\
\rho_5 &= h_2 - h_3 + v_5 + w.
\end{aligned}$$

   To estimate h_1, h_2, h_3, and w we minimize

$$\begin{aligned}
g(h_1, h_2, h_3, w) &= (h_1 - h_2 + w - \rho_1)^2 + (h_3 - h_1 + w - \rho_2)^2 + (-h_1 + w - \rho_3)^2 \\
&\quad + (-h_2 + w - \rho_4)^2 + (h_2 - h_3 + w - \rho_5)^2. \qquad (2)
\end{aligned}$$

   The variables are h_1, h_2, h_3, and w. The numbers ρ_i are given.

Solution.

1. Linear least squares:

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}, \quad x = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}, \quad b = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \\ \rho_5 \end{bmatrix},$$

2. Linear least squares:

$$A = \begin{bmatrix} 1 & -1 & 0 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ w \end{bmatrix}, \quad b = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \\ \rho_5 \end{bmatrix},$$

Problem 3 (20 points)

Consider the matrix A

$$A = \begin{bmatrix} 1 & 1 & 1^2 & \cdots & 1^{n-1} \\ 1 & 2 & 2^2 & \cdots & 2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & n-1 & (n-1)^2 & \cdots & (n-1)^{n-1} \\ 1 & n & n^2 & \cdots & n^{n-1} \end{bmatrix}.$$

Show that $\kappa(A) \geq n^{n-3/2}$.

Solution. We recognize the matrix A, previously seen in the context of polynomial interpolation, to be of the form

$$A = \begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{n-2} & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \cdots & t_2^{n-2} & t_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & t_{n-1} & t_{n-1}^2 & \cdots & t_{n-1}^{n-2} & t_{n-1}^{n-1} \\ 1 & t_n & t_n^2 & \cdots & t_n^{n-2} & t_n^{n-1} \end{bmatrix}.$$

where

$$t_1 = 1, \qquad t_2 = 2, \qquad t_3 = 3, \qquad \ldots, \qquad t_{n-1} = n-1, \qquad t_n = n.$$

Since the t_is are distinct, namely $t_i \neq t_j$ for $i \neq j$, we know that A is nonsingular. For $x \neq 0$,

$$\|A\| \geq \frac{\|Ax\|}{\|x\|}, \qquad \|A^{-1}\| \geq \frac{\|x\|}{\|Ax\|}.$$

Choosing $x = (0, \ldots, 0, 1)$ in the first inequality gives

$$\|A\| \geq \left(\sum_{i=1}^{n} t_i^{2(n-1)} \right)^{1/2} \geq t_n^{n-1} = n^{n-1}.$$

Choosing $x = (1, 0, \ldots, 0)$ in the second expression gives

$$\|A^{-1}\| \geq 1/\sqrt{n}.$$

Multiplying the two bounds gives

$$\kappa(A) = \|A\| \, \|A^{-1}\| \geq \frac{n^{n-1}}{\sqrt{n}} = n^{n-3/2}.$$

Problem 4 (20 points)

Let A be a square $n \times n$ nonsingular matrix, and let P be an $n \times n$ permuation matrix.

a. Explain why $\|PA\| = \|A\| = \|AP\|$.

b. The matrix A has the LU factorization $A = PLU$. Suppose $\|L\| = \alpha$, $\|U\| = \beta$, $\kappa(L) = \gamma$, and $\kappa(U) = \delta$, where α, β, γ, and δ are all positive. What are the smallest upper bounds you can derive for $\|A\|$, $\|A^{-1}\|$, $\kappa(A)$, and $\kappa(A^{-1})$?

Solution.

a. We'll first show that $\|PA\| = \|A\|$.

$$\begin{aligned} \|PA\| &= \max_{x \neq 0} \frac{\|PAx\|}{\|x\|} \\ &= \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \\ &= \|A\| \end{aligned}$$

where we used the fact that $\|PAx\| = \|Ax\|$ because permuting the elements of a vector does not change the norm of the vector.

We'll next show that $\|AP\| = \|A\|$.

$$\begin{aligned} \|APx\| &\leq \|AP\|\|x\| \\ &\leq \|A\|\|P\|\|x\| \\ &= \|A\|\|x\| \end{aligned}$$

Therefore, $\frac{\|APx\|}{\|x\|} \leq \|A\|$ for all nonzero vectors x. This implies $\|AP\| \leq \|A\|$. However, if we can find a vector x such that $\|APx\|/\|x\| = \|A\|$, then we will have shown that $\|AP\| = \|A\|$ since we already know that $\|AP\| \leq \|A\|$. Let $x^\star$ be a vector such that $\|Ax^\star\|/\|x^\star\| = \|A\|$. In other words, $x^\star$ is a vector whose gain through the system described by A is the maximum gain possible, namely the norm of A. Let $x' = P^{-1}x^\star$. Then, $\|APx'\|/\|x'\| = \|APP^{-1}x^\star\|/\|P^{-1}x^\star\| = \|Ax^\star\|/\|x^\star\| = \|A\|$. Therefore, $\|APx\|/\|x\| = \|A\|$ for $x = x'$ which proves that $\|AP\| = \|A\|$.

b. To solve this problem, we use the following property of matrix norms: $\|AB\| \leq \|A\|\|B\|$.
Thus, $\|A\| \leq \|P\|\|L\|\|U\| = \alpha\beta$.
$A^{-1} = (PLU)^{-1} = U^{-1}L^{-1}P^{-1}$. Therefore,

$$\begin{aligned} \|A^{-1}\| &\leq \|U^{-1}\|\|L^{-1}\|\|P^{-1}\| \\ &= \frac{\kappa(U)}{\|U\|}\frac{\kappa(L)}{\|L\|} \\ &= \frac{\gamma\delta}{\alpha\beta} \end{aligned}$$

Thus $\|A^{-1}\| \leq \frac{\gamma\delta}{\alpha\beta}$.

$$\begin{aligned}\kappa(A) = \kappa(A^{-1}) &= \|A\|\|A^{-1}\| \\ &\leq \alpha\beta\frac{\gamma\delta}{\alpha\beta} \\ &= \gamma\delta\end{aligned}$$

Thus, $\kappa(A) = \kappa(A^{-1}) \leq \gamma\delta$.

Problem 5 (20 points)

Consider the underdetermined set of linear equations

$$Ax + By = b$$

where the p-vector b, the $p \times p$-matrix A, and the $p \times q$-matrix B are given. The variables are the p-vector x and the q-vector y. We assume that A is nonsingular, and that B has a zero nullspace. (which implies $q \leq p$). The equations are underdetermined, so there are infinitely many solutions. For example, we can pick any y, and solve the set of linear equations $Ax = b - By$ to find x.

We will consider in this problem one possible solution, namely the one that minimizes $\|x\|^2$. Explain how to formulate this problem as a least squares problem. Describe the factorizations (QR, Cholesky, or LU) that you would use to calculate x and y. Clearly specify the matrices that you factor, the type of factorization, and the flop counts required. You do not need to worry about the flop counts required for the solve steps. Only consider flop counts needed to do the factor steps.

Solution.
x and y must be related by $x = A^{-1}b - A^{-1}By$. We can therefore minimize $\|x\|^2$ by solving the least-squares problem

$$\text{minimize } \|A^{-1}b - A^{-1}By\|^2$$

with variable y. Before we can solve this LS problem we need to calculate $A^{-1}b$ and the matrix $A^{-1}B$. The most efficient method is to take the LU factorization of A and then solve $q + 1$ sets of linear equations

$$Az_0 = b, \quad Az_1 = b_1 \quad \ldots, \quad Az_q = b_q$$

where b_i denotes the ith column of B. The result is $A^{-1}b = z_0$ and

$$A^{-1}B = [z_1 \; z_2 \; \cdots \; z_q].$$

The total cost of calculating $A^{-1}b$ and $A^{-1}B$ using this method is $(2/3)p^3$ for the factorization of A and $2(q + 1)p^2$ for the forward and back substitutions to find the vectors z_i.
We then solve the least-squares problem

$$\text{minimize } \|A^{-1}b - A^{-1}By\|^2$$

by taking the QR factorization of $A^{-1}B$, which costs $2pq^2$ flops. We know that $A^{-1}B$ has a zero nullspace, because B has a zero nullspace ($A^{-1}By = 0$ implies $By = 0$, which implies $y = 0$.) We can also use the Cholesky factorization of $(A^{-1}B)^T(A^{-1}B)$. The cost would be pq^2 to form the matrix, plus $(1/3)q^3$ for the factorization.