L. Vandenberghe
EE103

2/12/02

Midterm Solutions

Problem 1. (20 points)

Suppose you are asked to find a quadratic function

$$f(u_1, u_2) = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + r$$

that satisfies the following six conditions:

$$f(0,1) = 6, \quad f(1,0) = 6, \quad f(1,1) = 3, \quad f(-1,-1) = 7, \quad f(1,2) = 2, \quad f(2,1) = 6.$$

The variables in the problem are the parameters p_{11}, p_{12}, p_{22}, q_1, q_2 and r. Show that you can determine p_{11}, p_{12}, p_{22}, q_1, q_2, and r by solving a set of linear equations $Ax = b$. State clearly what A, x, and b are. You do not have to solve the equations, or show that they are solvable.

SOLUTION. We can write $f(u_1, u_2)$ as

$$f(u_1, u_2) = u_1^2 p_{11} + 2u_1u_2\, p_{12} + u_2^2\, p_{22} + u_1q_1 + u_2q_2 + r.$$

For given u_1 and u_2, this is a linear function of p_{11}, p_{12}, p_{22}, q_1, q_2, r. For example, $f(0,1) = 6$ means

$$p_{22} + q_2 + r = 6.$$

We therefore obtain the following set of equations:

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & -1 & -1 & 1 \\ 1 & 4 & 4 & 1 & 2 & 1 \\ 4 & 4 & 1 & 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ p_{22} \\ q_1 \\ q_2 \\ r \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ 3 \\ 7 \\ 2 \\ 6 \end{bmatrix}$$

Problem 2. (20 points)

1. For what values of $a_1, a_2, \ldots, a_n$ is the $n \times n$-matrix

$$A = \begin{bmatrix} a_1 & 1 & 0 & \cdots & 0 & 0 \\ a_2 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-2} & 0 & 0 & \cdots & 1 & 0 \\ a_{n-1} & 0 & 0 & \cdots & 0 & 1 \\ a_n & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

   nonsingular?

2. Assuming A is nonsingular, how many floating-point operations (flops) do you need to solve $Ax = b$?

3. Assuming A is nonsingular, what is the inverse A^{-1}? (In other words, express the elements of A^{-1} in terms of $a_1, a_2, \ldots, a_n$.)

SOLUTION.

1. $a_n \neq 0$. We can derive this from the definition of nonsingular matrices: A is nonsingular if and only if $Ax = 0$ implies $x = 0$.

   $Ax = 0$ means

$$\begin{aligned} x_2 &= -a_1 x_1 \\ x_3 &= -a_2 x_1 \\ &\vdots \\ x_n &= -a_{n-1} x_1 \\ a_n x_1 &= 0. \end{aligned}$$

   If $a_n \neq 0$, then from the last equation, $x_1 = 0$, hence also $x_2 = x_3 = \cdots = x_n = 0$, *i.e.*, $x = 0$, so A is nonsingular.

   If $a_n = 0$, then we can take $x_1 = 1$, $x_2 = -a_1$, $x_3 = -a_2$, etc., and obtain a nonzero x with $Ax = 0$. Therefore A is singular.

2. If we put the last equation first we obtain

$$\begin{aligned} a_n x_1 &= b_n \\ a_1 x_1 + x_2 &= b_1 \\ a_2 x_1 + x_3 &= b_2 \\ &\vdots \\ a_{n-1} x_1 + x_n &= b_{n-1}. \end{aligned}$$

We can solve these equations by forward substitution:

$$\begin{aligned} x_1 &= b_n/a_n \\ x_2 &= b_1 - a_1 x_1 \\ x_3 &= b_2 - a_2 x_1 \\ &\vdots \\ x_n &= b_{n-1} - a_{n-1} x_1, \end{aligned}$$

which takes $2n-1$ flops.

3. We can find A^{-1} by solving $AX = I$ column by column using the method of part 2:

$$A^{-1} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1/a_n \\ 1 & 0 & 0 & \cdots & 0 & -a_1/a_n \\ 0 & 1 & 0 & \cdots & 0 & -a_2/a_n \\ 0 & 0 & 1 & \cdots & 0 & -a_3/a_n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1}/a_n \end{bmatrix}.$$

Problem 3. (20 points)

Give the matrix norm $\|A\|$ of each of the following matrices A. Explain your answers.

1. $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
2. $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$
3. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -3/2 \end{bmatrix}$
4. $A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -3/2 \end{bmatrix}$

SOLUTION.

1. $\|A\| = 2.$

   We have $Ax = (x_1 + x_2, x_1 + x_2)$ and $\|Ax\| = \sqrt{2}|x_1 + x_2|$. From the identity $a^T x = \|a\| \, \|x\| \cos \angle(a, x)$ with $a = (1, 1)$, we know that

   $$|x_1 + x_2| \leq \sqrt{2}\,\sqrt{x_1^2 + x_2^2}$$

   with equality if $x_1 = x_2$. Therefore,

   $$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{x \neq 0} \sqrt{2}\frac{|x_1 + x_2|}{\sqrt{x_1^2 + x_2^2}} = 2.$$

2. $\|A\| = \sqrt{2}.$

   $Ax = (x_1 - x_2, x_1 + x_2)$ and $\|Ax\| = \sqrt{(x_1 - x_2)^2 + (x_1 + x_2)^2} = \sqrt{2(x_1^2 + x_2^2)}.$

   Therefore

   $$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{x \neq 0} \frac{\sqrt{2(x_1^2 + x_2^2)}}{\sqrt{x_1^2 + x_2^2}} = \sqrt{2}.$$

3. $\|A\| = 2.$

   $Ax = (x_1 + x_2, x_1 + x_2, -(3/2)x_3)$ and $\|Ax\| = \sqrt{2(x_1 + x_2)^2 + (9/4)x_3^2}.$

   From part 1 we know that

   $$2(x_1 + x_2)^2 + (9/4)x_3^2 \leq 4x_1^2 + 4x_2^2 + (9/4)x_3^2$$

with equality if $x_1 = x_2$. This allows us to derive the norm in the same way as for a diagonal matrix with diagonal elements $(2, 2, 3/2)$:

$$
\begin{aligned}
\|A\| &= \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} \\
&\leq \max_{x\neq 0} \frac{\sqrt{4x_1^2 + 4x_2^2 + (9/4)x_3^2}}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \\
&\leq \max_{x\neq 0} \frac{\sqrt{4x_1^2 + 4x_2^2 + 4x_3^2}}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \\
&= 2.
\end{aligned}
$$

This shows that $\|A\| \leq 2$. Moreover we have $\|Ax\|/\|x\| = 2$ for $x = (1, 1, 0)$, so $\|A\| = 2$.

4. $\|A\| = 3/2$.

   $Ax = (x_1 - x_2, x_1 + x_2, -(3/2)x_3)$ and $\|Ax\| = \sqrt{2x_1^2 + 2x_2^2 + (9/4)x_3^2}$.

   The derivation is the same as for a diagonal matrix with diagonal elements $(\sqrt{2}, \sqrt{2}, 3/2)$:

$$
\begin{aligned}
\|A\| &= \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} \\
&= \max_{x\neq 0} \frac{\sqrt{2x_1^2 + 2x_2^2 + (9/4)x_3^2}}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \\
&\leq \max_{x\neq 0} \frac{\sqrt{(9/4)x_1^2 + (9/4)x_2^2 + (9/4)x_3^2}}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \\
&= 3/2.
\end{aligned}
$$

   This shows that $\|A\| \leq 3/2$. Moreover we have $\|Ax\|/\|x\| = 3/2$ for $x = (0, 0, 1)$, so $\|A\| = 3/2$.

$\|Ax\| = \sqrt{(x_1+x_2)^2 + (x_2 - x_1)^2 + \frac{9}{4}x_3^2}$

$? \sqrt{2(x_1^2 + x_2^2)}$

Problem 4. (20 points)

You are given a nonsingular $n \times n$-matrix A and an n-vector b. You are asked to evaluate

$$x = (I + A^{-1} + A^{-2} + A^{-3})\, b$$

where $A^{-2} = (A^2)^{-1}$ and $A^{-3} = (A^3)^{-1}$.
Describe in detail how you would compute x, and give the flop counts of the different steps in your algorithm. If you know several methods, you should give the most efficient one (least number of flops for large n).

SOLUTION.

1. LU-factorization $A = PLU$ ($(2/3)n^3$ flops).
2. Calculate $y = A^{-1}b$ by solving $PLUy = b$:
   (a) Calculate $v = P^T b$ (0 flops, because $P^T b$ is a permutation of b)
   (b) Solve $Lw = v$ by forward substitution (n^2 flops)
   (c) Solve $Uy = w$ by backward substitution (n^2 flops)
3. Calculate $v = A^{-2}b = A^{-1}y$ by solving $PLUv = y$ ($2n^2$)
4. Calculate $w = A^{-3}b = A^{-1}v$ by solving $PLUw = v$ ($2n^2$)
5. $x = b + y + v + w$ ($3n$).

Total: $(2/3)n^3 + 6n^2 + 3n$.

Problem 5. (20 points)

You are given the Cholesky factorization $A = LL^T$ of a positive definite $n \times n$-matrix A, and a vector $u \in \mathbf{R}^n$.

1. What is the Cholesky factorization of the $(n+1) \times (n+1)$-matrix

$$B = \begin{bmatrix} A & u \\ u^T & 1 \end{bmatrix}?$$

   You can assume that B is positive definite.

2. What is the cost of computing the Cholesky factorization of B, if the factorization of A (*i.e.*, the matrix L) is given?

3. Suppose $\|L^{-1}\| \leq 1$. Show that B is positive definite for all u with $\|u\| < 1$.

SOLUTION.

1. We need to find $L_{11} \in \mathbf{R}^{n \times n}$, $L_{21} \in \mathbf{R}^{1 \times n}$, $L_{22} \in \mathbf{R}$ such that

$$B = \begin{bmatrix} A & u \\ u^T & 1 \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} L_{11}^T & L_{21}^T \\ 0 & L_{22} \end{bmatrix}.$$

   Moreover L_{11} must be lower triangular with positive diagonal elements, and $L_{22} > 0$. This gives the following conditions:

$$A = L_{11}L_{11}^T, \quad u = L_{11}L_{21}^T, \quad 1 = L_{21}L_{21}^T + L_{22}^2.$$

   The first equation states that L_{11} must be the Cholesky factor of A, which is given: $L_{11} = L$. The second condition allows us to compute L_{21}:

$$L_{21}^T = L^{-1}u.$$

   The last condition gives L_{22}:

$$L_{22} = \sqrt{1 - L_{21}L_{21}^T} = \sqrt{1 - u^T L^{-T} L^{-1} u} = \sqrt{1 - u^T A^{-1} u}.$$

   Putting everything together, we obtain the Cholesky factorization of B:

$$B = \begin{bmatrix} L & 0 \\ u^T L^{-T} & \sqrt{1 - u^T L^{-T} L^{-1} u} \end{bmatrix} \begin{bmatrix} L^T & L^{-1}u \\ 0 & \sqrt{1 - u^T L^{-T} L^{-1} u} \end{bmatrix}.$$

2. n^2 flops.

   We are given L, so we only need to compute L_{21} and L_{22}. We can compute $L^{-1}u$ by solving the set of equations $Lx = u$, which takes n^2 flops because L is lower triangular. Computing $\sqrt{1 - u^T L^{-T} L^{-1} u} = \sqrt{1 - x^T x}$ takes roughly $2n$ flops, so the total cost is about n^2.

3. Suppose $\|L^{-1}\| \leq 1$ and $\|u\| < 1$. It follows from one the properties of the matrix norm ($\|Ax\| \leq \|A\| \, \|x\|$) that

$$u^T L^{-T} L^{-1} u = \|L^{-1}u\|^2 \leq \|L^{-1}\|^2 \|u\|^2 < 1.$$

   Therefore the Cholesky factorization of B exists, *i.e.*, B is positive definite.