L. Vandenberghe
EE103

5/10/05

Midterm Solutions

Problem 1 (20 points). Let A be an $m \times n$ matrix with $\|A\| < 1$.

1. Show that the matrix $I - A^TA$ is positive definite.
2. Show that the matrix
$$\begin{bmatrix} I & A \\ A^T & I \end{bmatrix}.$$
is positive definite.

Solution.

1. If x is nonzero, then
$$x^T(I - A^TA)x = x^Tx - x^TA^TAx = \|x\|^2 - \|Ax\|^2 > 0$$
because $\|Ax\|/\|x\| \leq \|A\| < 1$ for all nonzero x.

2. This can be proved two ways. Applying the definition of positive definite matrix, we get
$$\begin{aligned} \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} I & A \\ A^T & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= x^Tx + x^TAy + y^TA^Tx + y^Ty \\ &= (x + Ay)^T(x + Ay) + y^Ty - y^TA^TAy \\ &= \|x + Ay\|^2 + y^T(I - A^TA)y \end{aligned}$$
and this is positive for nonzero (x, y): if $y \neq 0$, then the second term is positive because $I - A^TA$ is postive definite; if $y = 0$ and $x \neq 0$, the first term is positive.

We can also note that the matrix can be factored as
$$\begin{bmatrix} I & A \\ A^T & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ A^T & L \end{bmatrix} \begin{bmatrix} I & A \\ 0 & L^T \end{bmatrix}$$
where $I - A^TA = LL^T$ is the Cholesky factorization of $I - A^TA$ (a positive definite matrix, by the result of part 1).

Problem 2 (20 points). Let A be a nonsingular $n \times n$ matrix and b an n-vector. In each subproblem, describe an efficient method for computing the vector x and give a flop count of your algorithm, including terms of order two (n^2) and higher. If you know several methods, give the most efficient one (least number of flops for large n).

1. $x = (A^{-1} + A^{-2})b$.
2. $x = (A^{-1} + A^{-T})b$.
3. $x = (A^{-1} + JA^{-1}J)b$ where J is the $n \times n$ matrix

$$J = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

(J is the identity matrix with its columns reversed: $J_{ij} = 1$ if $i+j = n+1$ and $J_{ij} = 0$ otherwise.)

Solution.

1. We write x as $x = A^{-1}b + A^{-1}A^{-1}b = y + A^{-1}y$ where $y = A^{-1}b$.
   - LU factorization of A ($(2/3)n^3$ flops)
   - Solve $PLUy = b$
     - Solve $P\hat{y} = b$ (zero flops).
     - Solve $L\tilde{y} = \hat{y}$ by forward substitution (n^2 flops).
     - Solve $Uy = \tilde{y}$ by backward substitution (n^2 flops).
   - Solve $PLUz = y$
     - Solve $P\hat{z} = y$ (zero flops).
     - Solve $L\tilde{z} = \hat{z}$ by forward substitution (n^2 flops).
     - Solve $Uz = \tilde{z}$ by backward substitution (n^2 flops).
   - $x = y + z$ (n flops).

   Total: $(2/3)n^3 + 4n^2$.

2. We write x as $x = A^{-1}b + A^{-T}b$.
   - LU factorization of A ($(2/3)n^3$ flops)
   - Solve $PLUy = b$
     - Solve $P\hat{y} = b$ (zero flops).

- Solve $L\tilde{y} = \hat{y}$ by forward substitution (n^2 flops).
- Solve $Uy = \tilde{y}$ by backward substitution (n^2 flops).

- Solve $(PLU)^T z = U^T L^T P^T z = b$
  - Solve $U^T \hat{z} = b$ by forward substitution (n^2 flops).
  - Solve $L^T \tilde{z} = \hat{z}$ by backward substitution (n^2 flops).
  - Solve $P^T z = \tilde{z}$ (zero flops).
- $x = y + z$ (n flops).

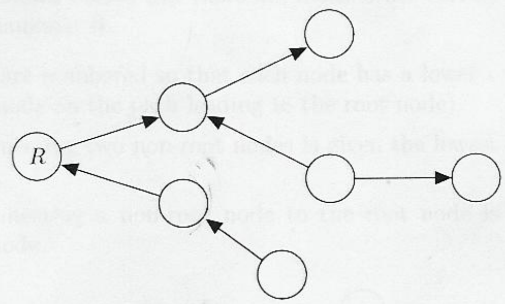
Total: $(2/3)n^3 + 4n^2$.

3. We write x as $x = A^{-1}b + JA^{-1}Jb$ and note that J is a permutation matrix with $J^{-1} = J$.

- LU factorization of A ($(2/3)n^3$ flops)
- Solve $PLUy = b$
  - Solve $P\hat{y} = b$ (zero flops).
  - Solve $L\tilde{y} = \hat{y}$ by forward substitution (n^2 flops).
  - Solve $Uy = \tilde{y}$ by backward substitution (n^2 flops).
- Solve $JPLUJz = b$.
  - Solve $Jv = b$ (zero flops, because J is a permutation matrix).
  - Solve $P\hat{z} = v$ (zero flops).
  - Solve $L\tilde{z} = \hat{z}$ by forward substitution (n^2 flops).
  - Solve $U\bar{z} = \tilde{z}$ by backward substitution (n^2 flops).
  - Solve $Jz = \bar{z}$ (zero flops, because J is a permutation matrix).
- $x = y + z$ (n flops).

Total: $(2/3)n^3 + 4n^2$.

Problem 3. (20 points). A *directed tree* is a connected graph with $n+1$ nodes and n directed arcs. The figure shows an example with $n = 6$ (7 nodes and 6 arcs).



A directed tree can be represented by a *reduced node-arc incidence matrix* defined as follows. We label one node as the *root* node. The other nodes are labeled with integers from 1 to n. The arcs are also given integer labels from 1 to n. The node-arc incidence matrix A is then defined as the $n \times n$ matrix with coefficients

$$A_{ij} = \begin{cases} +1 & \text{if arc } j \text{ ends at node } i \\ -1 & \text{if arc } j \text{ begins at node } i \\ 0 & \text{otherwise} \end{cases}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, n$.

1. Suppose we use the node labeled R as the root node in the example. Number the non-root nodes and the arcs in such a way that the node-incidence matrix is lower triangular.

2. Show that the node-incidence matrix obtained in part 1 is nonsingular.

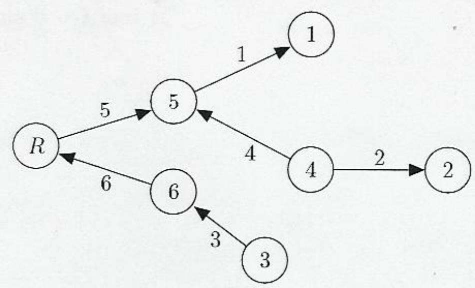
3. Show that the coefficients of the inverse of the node-incidence matrix in part 1 have values 0, +1 or −1.

4. Do properties 2 and 3 hold for any numbering of the nodes and arcs, or only if the incidence matrix is lower triangular?

Solution.

1. An example is shown below, but there are many other correct answers. In general, A will be lower triangular if

   - the nodes are numbered so that each node has a lower number than its 'parent' (the next node on the path leading to the root node),
   - an arc connecting two non-root nodes is given the lowest of the the labels of the two nodes,
   - an arc connecting a non-root node to the root node is given the label of the non-root node.



The node-arc incidence matrix for the numbering in the example is

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix}.$$

2. The inverse is

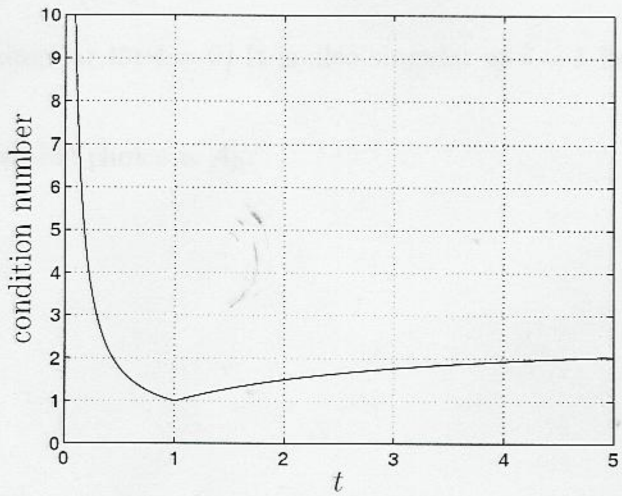
$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 \end{bmatrix}.$$

We cán find the inverse by solving $AX = I$, column by column, using forward substitution.

A faster method is based on the following interpretation. Suppose x is an n-vector, and we interpret x_k as the *flow* through arc k (positive if it follows the orientation of the arc and negative otherwise). Then $(Ax)_i$ is the total flow arriving at node i. We find th ith column of the inverse by solving the equation $Ax = e_i$, where e_i is the ith column of I. This equation can be interpreted as follows: we inject a flow of 1 into the root node and extract it from the network at node i. The total flow arriving at each of the other nodes is zero. From the graph it is easy to determine the flow vector x, *i.e.*, the ith column of A^{-1}. For example, for $i = 1$, it is clear that the solution is $(1, 0, 0, 0, 1, 0)$ becaue a unit flow injected at the root node and extracted at node 1 will pass through arcs 5 and 1. Similarly, for $i = 2$, we get $x = (0, 1, 0, -1, 1, 0)$, et cetera.

3. The properties are independent of the numbering. Choosing another numbering is equivalent to pre- and post-multiplying A with permutation matrices, *i.e.*, the new node-incidence matrix is P_1AP_2. The inverse matrix exists and is given by $(P_1AP_2)^{-1}$, so it also has elements 0, -1 and 1.

Problem 4 (20 points).



The graph shows the condition number of one of the following matrices as a function of t for $t \geq 0$.

$$A_1 = \begin{bmatrix} t & 1 \\ 1 & -t \end{bmatrix}, \qquad A_2 = \begin{bmatrix} t & t \\ -t & 1 \end{bmatrix}, \qquad A_3 = \begin{bmatrix} t & 0 \\ 0 & 1+t \end{bmatrix}, \qquad A_4 = \begin{bmatrix} t & -t \\ -t & 1 \end{bmatrix}.$$

Which of the four matrices was used in the figure? Carefully explain your answer.

Solution. We note that the condition number goes to infinity near $t = 0$, and is not defined at $t = 0$, meaning that the matrix is singular at $t = 0$. Another interesting point is $t = 1$, because the condition number there is equal to 1.

1. This matrix is not singular at $t = 0$. In fact $\kappa(A_1) = 1$ for $t = 0$ because it is a permutation matrix.

2. This matrix is singular at $t = 0$. The norm at $t = 1$ is

$$\begin{aligned} \|A_2\| &= \max_{\|x\|=1} ((x_1 + x_2)^2 + (-x_1 + x_2)^2)^{1/2} \\ &= \max_{\|x\|=1} (2x_1^2 + 2x_2^2)^{1/2} \\ &= \sqrt{2}. \end{aligned}$$

The inverse is

$$A_2^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

By the same argument as for A_2 we have $\|A_2^{-1}\| = 1/\sqrt{2}$. Therefore $\kappa(A_2) = \|A_2\| \|A_2^{-1}\| = \sqrt{2}/\sqrt{2} = 1$.

3. This matrix is singular for $t = 0$. At $t = 1$ its condition number is 2 because $\|A\| = 2$ and $\|A|^{-1}\| = 1$.

4. This matrix is singular for $t = 0$. It is also singular at $t = 1$ because the two columns are dependent.

Therefore the only possible choice is A_2.