L. Vandenberghe 4/29/15
EE133A

Midterm solutions

Problem 1. Let A be a tall $m \times n$ matrix with linearly independent columns. Define

$$P = A(A^T A)^{-1} A^T.$$

1. Show that the matrix $2P - I$ is orthogonal.
2. Use the Cauchy-Schwarz inequality to show that the inequalities

$$-\|x\|\|y\| \leq x^T(2P - I)y \leq \|x\|\|y\|$$

   hold for all m-vectors x and y.
3. Take $x = y$ in part 2. Show that the right-hand inequality implies that $\|Px\| \leq \|x\|$ for all m-vectors x.

Solution. We first note that P is symmetric and $P^2 = P$.

1. The matrix $2P - I$ is square and symmetric. We verify that $(2P - I)(2P - I) = I$:

$$(2P - I)(2P - I) = 4P^2 - 2P - 2P + I = I$$

   because $P^2 = P$.
2. We apply the Cauchy-Schwarz inequality to the vectors x and $(2P - I)y$:

$$\left|x^T(2P - I)y\right| \leq \|x\| \, \|(2P - I)y\| = \|x\| \, \|y\|.$$

   The last step uses the fact that multiplication with an orthogonal matrix preserves the norm of a vector: $\|(2P - I)y\| = \|y\|$.
3. If we plugging in $y = x$ the inequality simplifies to $x^T Px \leq \|x\|^2$. Since $P = P^2 = P^T P$, this can be written as

$$\|Px\|^2 = x^T P^T Px = x^T Px \leq \|x\|^2.$$

Problem 2. A lower triangular matrix A is *bidiagonal* if $A_{ij} = 0$ for $i > j+1$:

$$A = \begin{bmatrix} A_{11} & 0 & 0 & \cdots & 0 & 0 & 0 \\ A_{21} & A_{22} & 0 & \cdots & 0 & 0 & 0 \\ 0 & A_{32} & A_{33} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & A_{n-2,n-2} & 0 & 0 \\ 0 & 0 & 0 & \cdots & A_{n-1,n-2} & A_{n-1,n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & A_{n,n-1} & A_{nn} \end{bmatrix}.$$

Assume A is a nonsingular bidiagonal and lower triangular matrix of size $n \times n$.

1. What is the complexity of solving $Ax = b$?

2. What is the complexity of computing the inverse of A?

State the algorithm you use in each subproblem, and give the dominant term (exponent and coefficient) of the flop count. If you know several methods, consider the most efficient one.

Solution.

1. We solve $Ax = b$ by forward substitution:

$$x_1 = \frac{b_1}{A_{11}}, \qquad x_2 = \frac{b_2 - A_{21}x_1}{A_{22}}, \qquad x_3 = \frac{b_3 - A_{32}x_2}{A_{33}}, \qquad \ldots, \qquad x_n = \frac{b_n - A_{n,n-1}x_{n-1}}{A_{nn}}.$$

   This takes $3n - 2 \approx 3n$ flops.

2. We solve $AX = I$ column by column. In column i, we have $X_{1i} = \cdots = X_{i-1,i} = 0$ and

$$X_{ii} = \frac{1}{A_{ii}}, \quad X_{i+1,i} = -\frac{A_{i+1,i}X_{ii}}{A_{i+1,i+1}}, \quad X_{i+2,i} = -\frac{A_{i+2,i+1}X_{i+1,i}}{A_{i+2,i+2}}, \quad \ldots, \quad X_{n,i} = -\frac{A_{n,n-1}X_{n-1,i}}{A_{nn}}.$$

   This takes $3(n-i) - 2$ flops. Summing from $i = 1$ to n gives a total of $(3/2)n^2$.

Problem 3. Let B be an $m \times n$ matrix.

1. Prove that the matrix $I + B^T B$ is nonsingular. Since we do not impose any conditions on B, this also shows that the matrix $I + BB^T$ is nonsingular.

2. Show that the matrix
$$A = \begin{bmatrix} I & B^T \\ -B & I \end{bmatrix}$$
is nonsingular and that the following two expressions for its inverse are correct:
$$A^{-1} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -B^T \\ I \end{bmatrix} (I + BB^T)^{-1} \begin{bmatrix} B & I \end{bmatrix},$$
$$A^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} I \\ B \end{bmatrix} (I + B^T B)^{-1} \begin{bmatrix} I & -B^T \end{bmatrix}.$$

3. Now assume B has orthonormal columns. Use the result in part 2 to formulate a simple method for solving $Ax = b$. What is the complexity of your method? If you know several methods, give the most efficient one.

Solution.

1. Suppose $(I + B^T B)x = 0$. Then
$$x^T(I + BB^T)x = \|x\|^2 + \|B^T x\|^2 = 0$$
and this is only possible if $x = 0$.

2. We verify that $AA^{-1} = I$. Using the first expression
$$\begin{aligned} AA^{-1} &= \begin{bmatrix} I & B^T \\ -B & I \end{bmatrix} \left(\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -B^T \\ I \end{bmatrix} (I + BB^T)^{-1} \begin{bmatrix} B & I \end{bmatrix} \right) \\ &= \begin{bmatrix} I & 0 \\ -B & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ I + BB^T \end{bmatrix} (I + BB^T)^{-1} \begin{bmatrix} B & I \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ -B & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ B & I \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \end{aligned}$$
Using the second expression
$$\begin{aligned} AA^{-1} &= \begin{bmatrix} I & B^T \\ -B & I \end{bmatrix} \left(\begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} I \\ B \end{bmatrix} (I + B^T B)^{-1} \begin{bmatrix} I & -B^T \end{bmatrix} \right) \\ &= \begin{bmatrix} 0 & B^T \\ 0 & I \end{bmatrix} + \begin{bmatrix} I + B^T B \\ 0 \end{bmatrix} (I + B^T B)^{-1} \begin{bmatrix} I & -B^T \end{bmatrix} \\ &= \begin{bmatrix} 0 & B^T \\ 0 & I \end{bmatrix} + \begin{bmatrix} I & -B^T \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \end{aligned}$$

3. The second expression for A^{-1} simplifies because $B^TB = I$:

$$\begin{bmatrix} I & B^T \\ -B & I \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} + \frac{1}{2} \begin{bmatrix} I \\ B \end{bmatrix} \begin{bmatrix} I & -B^T \end{bmatrix}.$$

Therefore

$$\begin{bmatrix} I & B^T \\ -B & I \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 \\ b_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} I \\ B \end{bmatrix} (b_1 - B^T b_2)$$

To multiply A^{-1} with $b = (b_1, b_2)$ we take

$$x_1 = \frac{1}{2}(b_1 - B^T b_2), \qquad x_2 = b_2 + Bx_1.$$

This requires $4mn$ flops (for the multiplications with B and B^T).

Problem 4. We have defined the pseudo-inverse of a right invertible matrix B as the matrix

$$B^\dagger = B^T(BB^T)^{-1}.$$

Note that $B^\dagger B$ is a symmetric matrix. It can be shown that $B^\dagger$ is the only right inverse X of B with the property that XB is symmetric.

1. Assume A is a nonsingular $n \times n$ matrix and b is an n-vector. Show that the $n \times (n+1)$ matrix

$$B = \left[\begin{array}{cc} A & b \end{array}\right]$$

is right invertible and that

$$X = \left[\begin{array}{c} A^{-1} - A^{-1}by^T \\ y^T \end{array}\right]$$

is a right inverse of B, for any value of the n-vector y.

2. Show that XB is symmetric (hence, $X = B^\dagger$) if

$$y = \frac{1}{1 + \|A^{-1}b\|^2} A^{-T}A^{-1}b.$$

3. What is the complexity of computing the vector y in part 2 using an LU factorization of A? Give a flop count, including all cubic and quadratic terms. If you know several methods, consider the most efficient one.

Solution.

1. We verify that $BX = I$:

$$BX = \left[\begin{array}{cc} A & b \end{array}\right] \left[\begin{array}{c} A^{-1} - A^{-1}by^T \\ y^T \end{array}\right] = I - by^T + by^T = I.$$

2. We work out the product XB:

$$XB = \left[\begin{array}{c} A^{-1} - A^{-1}by^T \\ y^T \end{array}\right] \left[\begin{array}{cc} A & b \end{array}\right] = \left[\begin{array}{cc} I - A^{-1}by^TA & (1 - b^Ty)A^{-1}b \\ y^TA & y^Tb \end{array}\right].$$

We have to verify that this is symmetric for the given value of y. The 1,1 block is symmetric because A^Ty is a scalar multiple of $A^{-1}b$ and therefore $A^{-1}by^TA$ is a scalar multiple of the symmetric outer product $A^{-1}b(A^{-1}b)^T$.

The 2,1 block is the transpose of the 1,2 block because

$$A^Ty = \frac{1}{1 + \|A^{-1}b\|^2} A^{-1}b$$

and

$$(1 - b^Ty)A^{-1}b = (1 - \frac{\|A^{-1}b\|^2}{1 + \|A^{-1}b\|^2})A^{-1}b = \frac{1}{1 + \|A^{-1}b\|^2} A^{-1}b.$$

3. The LU factorization costs $(2/3)n^3$ flops. We compute $x = A^{-1}b$ using the standard method by solving $PLUx = b$ via forward and backward substitution ($2n^2$ flops). To compute $z = A^{-T}x$ we solve $U^TL^TP^Tz = x$ by forward and backward substitution, again in $2n^2$ flops.

The total is $(2/3)n^3 + 4n^2$ flops.