L. Vandenberghe 5/3/2016
EE133A

Midterm solutions

Problem 1. Suppose A is an $m \times p$ matrix with linearly independent columns and B is a $p \times n$ matrix with linearly independent rows. We are not assuming that A or B are square. Define

$$X = AB, \qquad Y = B^\dagger A^\dagger$$

where $A^\dagger$ and $B^\dagger$ are the pseudo-inverses of A and B. Show the following properties.

1. YX is symmetric.
2. XY is symmetric.
3. $YXY = Y$.
4. $XYX = X$.

Carefully explain your answers.

Solution. We use the definitions and properties

$$A^\dagger = (A^TA)^{-1}A^T, \qquad B^\dagger = B^T(BB^T)^{-1}, \qquad A^\dagger A = I, \qquad BB^\dagger = I.$$

1. Follows from
$$YX = (B^\dagger A^\dagger)(AB) = B^\dagger(A^\dagger A)B = B^\dagger B = B^T(BB^T)^{-1}B.$$
2. Follows from
$$XY = (AB)(B^\dagger A^\dagger) = A(BB^\dagger)A^\dagger = AA^\dagger = A(A^TA)^{-1}A^T.$$
3. In part 2 we have shown that $XY = AA^\dagger$. The result follows from
$$Y(XY) = (B^\dagger A^\dagger)(AA^\dagger) = B^\dagger(A^\dagger A)A^\dagger = B^\dagger A^\dagger = Y.$$
4. Similarly,
$$(XY)X = (AA^\dagger)(AB) = A(A^\dagger A)B) = AB = X.$$

Problem 2.

1. Formulate the following problem as a set of linear equations. Find a point $x \in \mathbf{R}^n$ at equal distance to $n+1$ given points $y_1, y_2, \ldots, y_{n+1} \in \mathbf{R}^n$:

$$\|x - y_1\| = \|x - y_2\| = \cdots = \|x - y_{n+1}\|.$$

   Write the equations in matrix form $Ax = b$.

2. Show that the solution x in part 1 is unique if the $(n+1) \times (n+1)$ matrix

$$\begin{bmatrix} y_1 & y_2 & \cdots & y_{n+1} \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

   is nonsingular.

Solution.

1. Squaring the equations and canceling the terms $x^T x$ we obtain n linear equations

$$\begin{aligned} 2(y_2 - y_1)^T x &= \|y_2\|^2 - \|y_1\|^2 \\ 2(y_3 - y_2)^T x &= \|y_3\|^2 - \|y_2\|^2 \\ &\vdots \\ 2(y_{n+1} - y_n)^T x &= \|y_{n+1}\|^2 - \|y_n\|^2. \end{aligned}$$

   In matrix form, $Ax = b$ with

$$A = \begin{bmatrix} (y_2 - y_1)^T \\ (y_3 - y_2)^T \\ \vdots \\ (y_{n+1} - y_n)^T \end{bmatrix}, \qquad b = \frac{1}{2} \begin{bmatrix} \|y_2\|^2 - \|y_1\|^2 \\ \|y_3\|^2 - \|y_2\|^2 \\ \vdots \\ \|y_{n+1}\|^2 - \|y_n\|^2 \end{bmatrix}.$$

2. We show that the matrix in part 1 is nonsingular. Suppose it is singular, so it has linearly dependent rows. This means there exists a nonzero z with

$$\begin{bmatrix} y_2 - y_1 & y_3 - y_2 & \cdots & y_{n+1} - y_n \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = 0.$$

   Then

$$\begin{bmatrix} y_1 & y_2 & \cdots & y_n & y_{n+1} \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \begin{bmatrix} -z_1 \\ z_1 - z_2 \\ \vdots \\ z_{n-1} - z_n \\ z_n \end{bmatrix} = 0,$$

   which contradicts the fact that the matrix on the left is nonsingular.

Problem 3. We use the notation I_n for the identity matrix of size $n \times n$ and J_n for the reversal matrix of size $n \times n$. (The reversal matrix is the identity matrix with the column order reversed.)

1. Verify that the $2n \times 2n$ reversal matrix J_{2n} can be written as

$$J_{2n} = Q \begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix} Q^T \qquad \text{where} \quad Q = \frac{1}{\sqrt{2}} \begin{bmatrix} I_n & I_n \\ J_n & -J_n \end{bmatrix}.$$

   Also show that Q is orthogonal.

2. Let A be a $2n \times 2n$ matrix with the property that

$$J_{2n}A = AJ_{2n}. \tag{1}$$

   An example is the 4×4 matrix

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 8 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{bmatrix}.$$

   Use the factorization of J_{2n} in part 1 to show that if A satisfies (1) then the matrix Q^TAQ is block-diagonal:

$$Q^TAQ = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix},$$

   where B and C are $n \times n$ matrices.

3. The complexity of solving a general linear equation $Ax = b$ of size $2n \times 2n$ is $(2/3)(2n)^3 = (16/3)n^3$. Suppose A has the property defined in part 2. By how much can the dominant term in the complexity of solving $Ax = b$ be reduced if we take advantage of the factorization property in part 2? Explain your answer.

Solution.

1. The first statement follows from

$$Q \begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix} Q^T = \frac{1}{2} \begin{bmatrix} I_n & I_n \\ J_n & -J_n \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix} \begin{bmatrix} I_n & J_n \\ I_n & -J_n \end{bmatrix} = \begin{bmatrix} 0 & J_n \\ J_n & 0 \end{bmatrix} = J_{2n}.$$

   Q is orthogonal because it is square and

$$Q^TQ = \frac{1}{2} \begin{bmatrix} I_n & J_n \\ I_n & -J_n \end{bmatrix} \begin{bmatrix} I_n & I_n \\ J_n & -J_n \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & I_n \end{bmatrix}.$$

2. Substituting the factorization of J_{2n} we find that

$$Q \begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix} Q^TA = AQ \begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix} Q^T.$$

   Multiplying on the left with Q^T and on the right with Q gives

$$\begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix} Q^TAQ = Q^TAQ \begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix}.$$

If we partition Q^TAQ as

$$Q^TAQ = \begin{bmatrix} B & D \\ E & C \end{bmatrix},$$

then this means that

$$\begin{bmatrix} B & D \\ -E & -C \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix} \begin{bmatrix} B & D \\ E & C \end{bmatrix} = \begin{bmatrix} B & D \\ E & C \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix} = \begin{bmatrix} B & -D \\ E & -C \end{bmatrix}.$$

Therefore $D = -D = 0$ and $E = -E = 0$.

3. The cost is reduced to $2 \times (2/3)n^3 = (4/3)n^3$, *i.e.*, by a factor of four. It is four times less expensive to solve two equations of size n than one equation of size $2n$.

   The details are as follows. Partition A, x, and b as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

   Note from part 2 that

$$\begin{aligned} \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix} &= Q^TAQ \\ &= \frac{1}{2} \begin{bmatrix} I_n & J_n \\ I_n & -J_n \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I_n & I_n \\ J_n & -J_n \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} A_{11} + A_{12}J_n + J_nA_{21} + J_nA_{22}J_n & 0 \\ 0 & A_{11} - A_{12}J_n - J_nA_{21} + J_nA_{22}J_n \end{bmatrix}. \end{aligned}$$

   Computing B and C costs $6n^2$ flops (n^2 for $A_{11} + J_nA_{22}J_n$, n^2 for $A_{12}J_n + J_nA_{21}$, n^2 for the sum of these two matrices, n^2 for the difference, and $2n^2$ for the scalar multiplication with $1/2$).

   To solve $Ax = b$ we use the factorization

$$A = Q \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix} Q^T.$$

   - Solve $Qu = b$. Since Q is orthogonal, the solution is

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = Q^Tb = \frac{1}{\sqrt{2}} \begin{bmatrix} b_1 + J_nb_2 \\ b_1 - J_nb_2 \end{bmatrix}.$$

     This requires $4n$ flops.
   - Compute B and C ($6n^2$ flops) and solve

$$\begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

     This is equivalent to two independent linear equations $By_1 = u_1$ and $Cy_2 = u_2$. The complexity is $(4/3)n^3$.
   - Solve $Q^Tx = y$. Since Q is orthogonal, the solution is

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = Qy = \frac{1}{\sqrt{2}} \begin{bmatrix} y_1 + y_2 \\ J_ny_1 - J_ny_2 \end{bmatrix}.$$

     This requires $4n$ flops.

   The dominant term is $(4/3)n^3$.

Problem 4. Let A be an $m \times n$ matrix with linearly independent columns. Suppose $A_{ij} = 0$ for $i > j + 1$. In other words, the elements of A below the first subdiagonal are zero:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1,n-1} & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2,n-1} & A_{2n} \\ 0 & A_{32} & \cdots & A_{3,n-1} & A_{3n} \\ 0 & 0 & \cdots & A_{4,n-1} & A_{4n} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & A_{n,n-1} & A_{nn} \\ 0 & 0 & \cdots & 0 & A_{n+1,n} \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Show that the Q-factor in the QR factorization of A has the same property: $Q_{ij} = 0$ for $i > j + 1$.

Solution. This follows from $Q = AR^{-1}$ and the fact that R^{-1} is upper triangular. Therefore the kth column of Q is a linear combination of the first k columns of A.

Problem 5. Consider a square $(n+m) \times (n+m)$-matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

with A of size $n \times n$ and D of size $m \times m$. We assume A is nonsingular. The matrix

$$S = D - CA^{-1}B$$

is called the *Schur complement* of A. Describe efficient algorithms for computing the Schur complement of each of the following types of matrices A.

1. A is diagonal.
2. A is lower triangular.
3. A is a general square matrix.

In each subproblem, give the different steps in the algorithm and their complexity. Include in the total flop count all terms that are order three (n^3, n^2m, nm^2, m^3) or higher. If you know different algorithms, choose the most efficient one.

Solution. In all three problems, we first compute $X = A^{-1}B$ and then compute $S = D - CX$. The complexity is $2m^2n$ for the product CX and m^2 for the subtraction. The complexity of computing $X = A^{-1}B$ depends on the properties of A.

1. If A is diagonal we compute the elements of $X = A^{-1}B$ as $X_{ij} = B_{ij}/A_{ii}$. This requires mn flops.
2. If A is lower triangular, we solve $AX = B$, column by column, via forward substitution. This requires mn^2 flops.
3. In the general case we compute an LU factorization of A ($(2/3)n^3$ flops) and then solve $AX = B$ column by column ($2mn^2$ flops).

Keeping only third order terms we obtain for the total complexity:

1. $2m^2n$.
2. $2m^2n + mn^2$.
3. $2m^2n + 2mn^2 + (2/3)n^3$.