L. Vandenberghe 10/31/2019
ECE133A

Midterm solutions

Problem 1. Suppose A and B are $m \times n$ matrices with linearly independent columns. Do the following matrices necessarily have linearly independent columns? If yes, explain why. If no, give a counterexample.

1. $\begin{bmatrix} A \\ B \end{bmatrix}$.
2. $\begin{bmatrix} A & B \end{bmatrix}$.
3. $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$.
4. AB^T.
5. A^TB.

Solution. To show that a matrix C has linearly independent columns, we show that $Cx = 0$ holds only if $x = 0$. To show that the columns are linearly dependent, we give a specific nonzero x for which $Cx = 0$.

1. Yes.
$$\begin{bmatrix} A \\ B \end{bmatrix} x = \begin{bmatrix} Ax \\ Bx \end{bmatrix} = 0 \quad \Longrightarrow \quad Ax = 0,\ Bx = 0 \quad \Longrightarrow \quad x = 0$$
because A and B have linearly independent columns. In fact it is sufficient that one of the two matrices has linearly independent columns.

2. No. For example, if we take $A = B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, then
$$\begin{bmatrix} A & B \end{bmatrix} x = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$
for $x = (1, -1)$.

3. Yes.
$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} Ax_1 \\ Bx_2 \end{bmatrix} = 0 \quad \Longrightarrow \quad Ax_1 = 0,\ \ Bx_2 = 0 \quad \Longrightarrow \quad x_1 = 0,\ \ x_2 = 0.$$

4. No. See the example in part 2.

5. No. For example, if we take $A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $B = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, then
$$A^TB = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 0.$$

Problem 2. Formulate the following problem as a set of linear equations. Find a polynomial

$$p(t) = x_1 + x_2 t + x_3 t^2 + \cdots + x_n t^{n-1}$$

that satisfies the following conditions on the values and the derivatives at m points $t_1, \ldots, t_m$:

$$p(t_1) = y_1, \quad p(t_2) = y_2, \quad \ldots, \quad p(t_m) = y_m, \qquad p'(t_1) = s_1, \quad p'(t_2) = s_2, \quad \ldots, \quad p'(t_m) = s_m.$$

The unknowns in the problem are the coefficients $x_1, \ldots, x_n$. The values of t_i, y_i, s_i are given.

1. Express the problem as a set of linear equations $Ax = b$. Clearly state how A and b are defined.
2. Suppose $n = 2m$ and the m points t_i are distinct. Is the matrix A in part 1 nonsingular? Explain your answer.

Solution.

1.

$$\begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{n-2} & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \cdots & t_2^{n-2} & t_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ 1 & t_m & t_m^2 & \cdots & t_m^{n-2} & t_m^{n-1} \\ 0 & 1 & 2t_1 & \cdots & (n-2)t_1^{n-3} & (n-1)t_1^{n-2} \\ 0 & 1 & 2t_2 & \cdots & (n-2)t_2^{n-3} & (n-1)t_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & 2t_m & \cdots & (n-2)t_m^{n-3} & (n-1)t_m^{n-2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \\ s_1 \\ s_2 \\ \vdots \\ s_m \end{bmatrix}.$$

2. The matrix is nonsingular. To show this, we verify that $Ax = 0$ only if $x = 0$. If $Ax = 0$, then the polynomial $p(t) = x_1 + x_2 t + \cdots + x_n t^{n-1}$ is zero and has zero derivative at the m points t_1, $\ldots$, t_m. Therefore each point t_i is a zero with multiplicity at least two. Counting roots with their multiplicity, we would have at least $n = 2m$ real roots. This is impossible for a polynomial of degree at most $n - 1$, unless the polynomial is a constant zero.

Problem 3. Let U be an $m \times n$ matrix with orthonormal columns. We do not assume that U is square. Define

$$A = I + \alpha UU^T$$

where α is a scalar.

1. For what values of α is A orthogonal?
2. For what values of α is A nonsingular?

Solution.

1. $\alpha = -2$ and $\alpha = 0$.

   $$A^T A = (I + \alpha UU^T)(I + \alpha UU^T) = I + 2\alpha UU^T + \alpha^2 UU^T UU^T = I + \alpha(2 + \alpha)UU^T.$$

   This is equal to I for $\alpha = 0$ and for $\alpha = -2$.

2. $\alpha \neq -1$.

   First consider $\alpha \neq -1$. Suppose $Ax = x + \alpha UU^T x = 0$. Multiplying with U^T on the left and using $U^T U = I$ shows that $U^T x = -\alpha U^T x$. If $\alpha \neq -1$, we must have $U^T x = 0$, but then $x = -\alpha UU^T x = 0$. This shows that A is nonsingular if $\alpha \neq -1$.

   One can also note that if $\alpha \neq -1$, the matrix has an inverse, namely

   $$I - \frac{\alpha}{1 + \alpha} UU^T.$$

   Next, consider $\alpha = -1$. In this case any x of the form $x = Uy$ satisfies

   $$Ax = x + UU^T x = Uy - UU^T Uy = Uy - Uy = 0.$$

   Moreover, if we choose $y \neq 0$, then $x = Uy \neq 0$ because $\|x\| = \|Uy\| = \|y\| \neq 0$. This shows that A is singular if $\alpha = -1$.

Problem 4. Let A be an $m \times n$ matrix with linearly independent columns. Define

$$B = \left[\begin{array}{cc} 0 & A^T \\ A & I \end{array}\right].$$

1. Show that B is invertible with inverse

$$B^{-1} = \left[\begin{array}{cc} -(A^TA)^{-1} & A^\dagger \\ (A^\dagger)^T & I - AA^\dagger \end{array}\right].$$

   Here $A^\dagger$ is the pseudo-inverse of A.

2. Substitute the QR factorization of A in the expression for B^{-1} in part 1 and simplify as much as possible.

3. Use the result of part 2 to formulate a method for computing the solution of an equation $Bx = y$ using the QR factorization of A. The right-hand side y and the variable x are $(m+n)$-vectors. Give the complexity of each step in the method and the overall complexity. Include in the total all terms that are cubic (m^3, m^2n, mn^2, n^3) and quadratic (m^2, mn, n^2) in the matrix dimensions. If you know several methods, give the most efficient one.

Solution.

1. We multiply B with the proposed inverse and verify that the product is the identity matrix:

$$\begin{aligned}
\left[\begin{array}{cc} 0 & A^T \\ A & I \end{array}\right]\left[\begin{array}{cc} -(A^TA)^{-1} & A^\dagger \\ (A^\dagger)^T & I - AA^\dagger \end{array}\right] &= \left[\begin{array}{cc} A^T(A^\dagger)^T & A^T - A^TAA^\dagger \\ -A(A^TA)^{-1} + (A^\dagger)^T & AA^\dagger + I - AA^\dagger \end{array}\right] \\
&= \left[\begin{array}{cc} I & A^T - A^TA(A^TA)^{-1}A^T \\ 0 & I \end{array}\right] \\
&= \left[\begin{array}{cc} I & 0 \\ 0 & I \end{array}\right].
\end{aligned}$$

   On line 2 we use the definition of $A^\dagger$. Since $A^\dagger$ is a left inverse of A, we get $A^T(A^\dagger)^T = (A^\dagger A)^T = I$ in the 1,1 block. For the $1,2$ and $2,1$ blocks we subsitute the definition of $A^\dagger = (A^TA)^{-1}A^T$ and obtain

$$A^T - A^TAA^\dagger = A^T - A^TA(A^TA)^{-1}A^T = A^T - A^T = 0,$$

   and

$$-A(A^TA)^{-1} + (A^\dagger)^T = -(A^\dagger)^T + (A^\dagger)^T = 0.$$

2. The QR factorization is $A = QR$ with $Q^TQ = I$ and R an upper triangular matrix with positive diagonal elements. We have

$$A^TA = (QR)^T(QR) = R^TQ^TQR = R^TR, \qquad (A^TA)^{-1} = (R^TR)^{-1} = R^{-1}R^{-T},$$

   and

$$A^\dagger = (A^TA)^{-1}A^T = R^{-1}R^{-T}R^TQ^T = R^{-1}Q^T.$$

   Also, $AA^\dagger = (QR)(R^{-1}Q^T) = QQ^T$. The inverse therefore simplifies to

$$B^{-1} = \left[\begin{array}{cc} -R^{-1}R^{-T} & R^{-1}Q^T \\ QR^{-T} & I - QQ^T \end{array}\right].$$

3. If we multiply B^{-1} with the right-hand side we obtain

$$\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{cc} -R^{-1}R^{-T} & R^{-1}Q^T \\ QR^{-T} & I - QQ^T \end{array}\right]\left[\begin{array}{c} y_1 \\ y_2 \end{array}\right] = \left[\begin{array}{c} R^{-1}(-R^{-T}y_1 + Q^Ty_2) \\ y_2 - Q(-R^{-T}y_1 + Q^Ty_2) \end{array}\right].$$

   This can be computed as follows.

- Compute $u = R^{-T} y_1$. We solve $R^T u = y_1$ by forward substitution. n^2 flops.
- Compute $v = Q^T y_2$. This is a product of an $n \times m$ matrix with an m-vector. $2mn$ flops.
- Compute $w = v - u$. n flops.
- Compute $x_1 = R^{-1} w$. We solve $Rx_1 = y_1$ by back substitution. n^2 flops.
- Compute $z = Qw$. This is a product of an $m \times n$ matrix with an n-vector. $2mn$ flops.
- Compute $x_2 = y_2 - z$. m flops.

The total, including the cost of the QR factorization, is $2mn^2 + 4mn + 2n^2$.

Problem 5. Suppose A is a nonsingular $n \times n$ matrix, and B and C are $n \times m$ matrices with $m \ll n$. Describe an efficient method for solving the matrix equation

$$AXA^T = BC^T$$

using an LU factorization of A. The variable X is an $n \times n$ matrix. Give the complexity of each step in your method and the total complexity. Include in the total complexity all terms that are cubic in m and n (n^3, n^2m, nm^2, m^3) or higher order than cubic. If you know several methods, choose the most efficient method when $m \ll n$.

Solution. The solution is

$$X = A^{-1}BC^TA^{-T} = (A^{-1}B)(A^{-1}C)^T.$$

We first compute $Y = A^{-1}B$ and $W = A^{-1}C$, *i.e.*, solve the matrix equations $AY = B$ and $AW = C$. This requires only one LU factorization of A. We then form the product $X = YW^T$.

- LU factorization $A = PLU$. $(2/3)n^3$ flops.
- Solving $AY = B$, column by column. The kth column of Y is $y_k = A^{-1}b_k = U^{-1}L^{-1}P^Tb_k$ if b_k is the kth column of B. We compute y_k in three steps, following the standard method.
  - Compute $u_k = P^Tb_k$. This is a permutation of b_k. 0 flops.
  - Compute $v_k = L^{-1}u_k$. We solve $Lv_k = u_k$ by forward substitution. n^2 flops.
  - Compute $y_k = U^{-1}v_k$. We solve $Uy_k = v_k$ by back substitution. n^2 flops.

  The total for the m columns of Y is $2mn^2$ flops.
- Compute $W = A^{-1}C$ by the same method. $2mn^2$ flops.
- Compute the product $X = YW^T$. This is a product of an $n \times m$ matrix and an $m \times n$ matrix. $2n^2m$ flops.

The total complexity is $(2/3)n^3 + 6n^2m$.

An alternative would be to first compute $D = BC^T$, then solve $AY = D$, column by column, to compute $Y = A^{-1}D$, and then solving $AX^T = Y^T$, column by column, to compute $X = YA^{-1} = A^{-1}DA^{-1}$. This would require $2n$ forward and backward substitutions, in addition to the LU factorization and the cost of the product BC^T. The total is $(2/3)n^3 + 2mn^2 + 4n^3$. When $n \gg m$ this is less efficient than the method described above because of the additional n^3 term.