Linear and Nonlinear Computation Models

(CSE 4190.313)

Midterm Exam: April 21, 2014

1. (15 points) If A has row 1 + row 2 = row 3, show that A is not invertible:

   (a) (5 points) Explain why $A\mathbf{x} = (1,0,0)^T$ cannot have a solution.

   (b) (5 points) Which right-hand sides (b_1, b_2, b_3) might allow a solution to $A\mathbf{x} = \mathbf{b}$?

   (c) (5 points) What happens to row 3 in elimination?

   Solution:

   (a) Suppose there exists a solution $\mathbf{x}$ such that $A\mathbf{x} = (1,0,0)^T$,
   where $A = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \end{bmatrix}$ and $\mathbf{a}_3^T = \mathbf{a}_1^T + \mathbf{a}_2^T$.
   Then $\mathbf{a}_1^T \cdot \mathbf{x} = 1$, $\mathbf{a}_2^T \cdot \mathbf{x} = 0$, and $\mathbf{a}_3^T \cdot \mathbf{x} = 0$.
   But, $\mathbf{a}_3^T \cdot \mathbf{x} = (\mathbf{a}_1^T + \mathbf{a}_2^T) \cdot \mathbf{x} = \mathbf{a}_1^T \cdot \mathbf{x} + \mathbf{a}_2^T \cdot \mathbf{x} = 1 + 0 = 1 \neq 0 = \mathbf{a}_3^T \cdot \mathbf{x}$,
   which is a contradiction.

   (b) To be consistent with the left-hand side, we have $b_1 + b_2 = b_3$.
   (This is because $b_1 + b_2 = \mathbf{a}_1^T \cdot \mathbf{x} + \mathbf{a}_2^T \cdot \mathbf{x} = \mathbf{a}_3^T \cdot \mathbf{x} = b_3$.)

   (c) $$\begin{bmatrix} 1 & 0 & 0 \\ -l_{21} & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & * & * \\ 0 & 0 & 0 \end{bmatrix}$$

   $$A = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ 1+l_{21} & 1 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & * & * \\ 0 & 0 & 0 \end{bmatrix} = LU$$

   Thus, the third row is reduced to a zero vector $(0,0,0)$.

2. (10 points) If $A\mathbf{x} = \mathbf{b}_1$ has infinitely many solutions, why is it impossible for $A\mathbf{x} = \mathbf{b}_2$ (new right-hand side) to have only one solution? Could $A\mathbf{x} = \mathbf{b}_2$ have no solution?

   Solution:

   (a) The columns of A are linearly dependent.
   Thus, there are infinitely many nonzero $\mathbf{x}_N (\neq \mathbf{0}) \in N(A)$ with $A\mathbf{x}_N = \mathbf{0}$.
   Now, for the unique solution $\mathbf{x}$ for $A\mathbf{x} = \mathbf{b}_2$, we have

   $$A\overline{\mathbf{x}} = A(\mathbf{x} + \mathbf{x}_N) = A\mathbf{x} + A\mathbf{x}_N = \mathbf{b}_2 + \mathbf{0} = \mathbf{b}_2,$$

   for infinitely many $\overline{\mathbf{x}} = \mathbf{x} + \mathbf{x}_N$. This is a contradiction.

   (b) $A\mathbf{x} = \mathbf{b}_2$ has no solution if and only if $\mathbf{b}_2 \notin C(A)$.

3. (10 points) If $\mathbf{v}_1, \cdots, \mathbf{v}_n$ is an orthonormal basis for R^n, show that $\mathbf{v}_1\mathbf{v}_1^T + \cdots + \mathbf{v}_n\mathbf{v}_n^T = I$.

Solution:

$$\begin{aligned} I\mathbf{x} = \mathbf{x} &= (\mathbf{v}_1^T \cdot \mathbf{x})\mathbf{v}_1 + \cdots + (\mathbf{v}_n^T \cdot \mathbf{x})\mathbf{v}_n \\ &= \mathbf{v}_1(\mathbf{v}_1^T \cdot \mathbf{x}) + \cdots + \mathbf{v}_n(\mathbf{v}_n^T \cdot \mathbf{x}) \\ &= (\mathbf{v}_1 \cdot \mathbf{v}_1^T) \cdot \mathbf{x} + \cdots + (\mathbf{v}_n \cdot \mathbf{v}_n^T) \cdot \mathbf{x} \\ &= (\mathbf{v}_1 \cdot \mathbf{v}_1^T + \cdots + \mathbf{v}_n \cdot \mathbf{v}_n^T) \cdot \mathbf{x}, \quad \text{for all } \mathbf{x} \in R^n. \end{aligned}$$

4. (15 points)

   (a) (5 points) How far is the plane $x_1 + x_2 - x_3 - x_4 = 8$ from the origin, and what point on it is nearest?

   (b) (5 points) Is there a matrix whose row space contains $(1, 1, 0)$ and whose nullspace contains $(0, 1, 1)$?

   (c) (5 points) If A is a square matrix, show that the column space of A^2 is contained in the column space of A.

   (a) The plane can be represented as $\frac{1}{2}(x_1+x_2-x_3-x_4) = 4$ with a unit normal $\mathbf{n} = \frac{1}{2}(1, 1, -1, -1)$, which is at distance 4 from the origin. This plane contains a point $\mathbf{p}_0 = (8, 0, 0, 0)$ which can be projected to a point $\hat{\mathbf{p}}_0$ on the orthogonal line passing through the origin (and thus parallel to the normal direction $\mathbf{n}$). The projection point is $\hat{\mathbf{p}}_0 = (\mathbf{p}_0 \cdot \mathbf{n})\mathbf{n} = 4 \cdot \frac{1}{2}(1, 1, -1, -1) = (2, 2, -2 - 2)$, which is the nearest point of the plane to the origin.

   (b) The row space and the nullspace are orthogonal complements each other. But the two vectors are not orthognonal as their inner product $(1, 1, 0) \cdot (0, 1, 1) = 1 \neq 0$. Thus, there is no such a matrix.

   (c) $A^2 = AA = A\,[\mathbf{a}_1 \cdots \mathbf{a}_n] = [A\mathbf{a}_1 \cdots A\mathbf{a}_n]$. Each column $A\mathbf{a}_j$ of A^2 can be represented as a linear combination of the columns of A: $A\mathbf{a}_j = \sum_{i=1}^{n} a_{ij}\mathbf{a}_i$, where $\mathbf{a}_j = (a_{1j}, \cdots, a_{n}j)^T$. Thus, we have $C(A^2) \subset C(A)$.

5. (15 points) Without computing A, find bases for the four fundamental subspaces:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 6 & 1 & 0 \\ 9 & 8 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

   (a) Basis of $C(A)$: $\{(1, 6, 9)^T, (0, 1, 8)^T, (0, 0, 1)^T\}$.

   (b) U can be reduced to $R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$,
   which means that $(0, 1, -2, 1)$ is a special solution.
   Basis of $N(A)$: $\{(0, 1, -2, 1)\}$.

   (c) Basis of $C(A^T)$: $\{(1, 2, 3, 4), (0, 1, 2, 3), (0, 0, 1, 2)\}$.

   (d) Basis of $N(A^T)$: none

6. (20 points) Find an orthonormal set $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ for which $\mathbf{q}_1, \mathbf{q}_2$ span the column space of

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

What fundamental subspace contains $\mathbf{q}_3$? What is the least-squares solution of $A\mathbf{x} = \mathbf{b}$ if $\mathbf{b} = [1\ 2\ 7]^T$?

Solution:

(a) $\mathbf{q}_1 = \mathbf{a}_1/\|\mathbf{a}_1\| = \frac{1}{\sqrt{1^2+2^2+(-2)^2}} \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix}$,

$$B = \mathbf{a}_2 - (\mathbf{a}_2^T \cdot \mathbf{q}_1)\mathbf{q}_1 = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} + 3 \begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

$$\mathbf{q}_2 = B/\|B\| = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$$

$$C = \mathbf{a}_3 - (\mathbf{a}_3^T \cdot \mathbf{q}_1)\mathbf{q}_1 - (\mathbf{a}_3^T \cdot \mathbf{q}_2)\mathbf{q}_2 = \begin{bmatrix} 0 \\ 0 \\ 9 \end{bmatrix} + 6 \begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix} - 6 \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$$

$$\mathbf{q}_3 = C/\|C\| = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$$

(b) $\mathbf{q}_3$ is contained in the left nullspace of A.

(c) $A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{a}_1^T \cdot \mathbf{q}_1 & \mathbf{a}_2^T \cdot \mathbf{q}_1 \\ 0 & \mathbf{a}_2^T \cdot \mathbf{q}_2 \end{bmatrix} = \begin{bmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \\ -2/3 & 2/3 \end{bmatrix} \begin{bmatrix} 3 & -3 \\ 0 & 3 \end{bmatrix} = QR$

$A\mathbf{x} = \mathbf{b}$, $QR\mathbf{x} = \mathbf{b}$, $R\hat{\mathbf{x}} = Q^T\mathbf{b}$; then
solving $\begin{bmatrix} 3 & -3 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} -3 \\ 6 \end{bmatrix}$ produces $\hat{\mathbf{x}} = (1, 2)^T$.

7. (15 points) The m errors e_i are independent with variance σ^2, so the average of $(\mathbf{b} - A\mathbf{x})(\mathbf{b} - A\mathbf{x})^T$ is $\sigma^2 I$. Mutiply on the left by $(A^TA)^{-1}A^T$, on the right by $A(A^TA)^{-1}$, and show that the average of $(\widehat{\mathbf{x}} - \mathbf{x})(\widehat{\mathbf{x}} - \mathbf{x})^T$ is $\sigma^2(A^TA)^{-1}$, which is the **covariance matrix** for the error in $\widehat{\mathbf{x}}$.

Solution:

$$\begin{aligned}
E((\mathbf{b} - A\mathbf{x})(\mathbf{b} - A\mathbf{x})^T) &= \sigma^2 I \\
((A^TA)^{-1}A^T)E((\mathbf{b} - A\mathbf{x})(\mathbf{b} - A\mathbf{x})^T)(A(A^TA)^{-1}) &= ((A^TA)^{-1}A^T)(\sigma^2 I)(A(A^TA)^{-1}) \\
E((A^TA)^{-1}A^T(\mathbf{b} - A\mathbf{x})(\mathbf{b} - A\mathbf{x})^T A(A^TA)^{-1}) &= \sigma^2(A^TA)^{-1}A^TA(A^TA)^{-1} \\
E(((A^TA)^{-1}A^T\mathbf{b} - (A^TA)^{-1}A^TA\mathbf{x})(\mathbf{b}^T - \mathbf{x}^TA^T)A(A^TA)^{-1}) &= \sigma^2(A^TA)^{-1} \\
E((\widehat{\mathbf{x}} - \mathbf{x})(\mathbf{b}^TA(A^TA)^{-1} - \mathbf{x}^TA^TA(A^TA)^{-1})) &= \sigma^2(A^TA)^{-1} \\
E((\widehat{\mathbf{x}} - \mathbf{x})(((A^TA)^{-1}A^T\mathbf{b})^T - \mathbf{x}^T)) &= \sigma^2(A^TA)^{-1} \\
E((\widehat{\mathbf{x}} - \mathbf{x})(\widehat{\mathbf{x}}^T - \mathbf{x}^T)) &= \sigma^2(A^TA)^{-1} \\
E((\widehat{\mathbf{x}} - \mathbf{x})(\widehat{\mathbf{x}} - \mathbf{x})^T) &= \sigma^2(A^TA)^{-1}
\end{aligned}$$