Homework 2
CS 210

Andres Calderon - SID 861243796

April 26, 2016

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
9	10	
10	10	
Total	100	

Matrix algebra

1. (Trefethen&Bau 2.6) If $\mathbf{u}$ and $\mathbf{v}$ are $m-$vectors, the matrix $A = I + \mathbf{u}\mathbf{v}^T$ is known as a rank-one pertubation of the identity. Show that if A is nonsingular, then its inverse has the form $A^{-1} = I + \alpha\mathbf{u}\mathbf{v}^T$ for some scalar α, and give an expression for α. For what $\mathbf{u}$ and $\mathbf{v}$ is A singular? If it is singular, what is null(A)?

   Answer:

$$\begin{aligned} I &= AA^{-1} \\ &= (I + \mathbf{u}\mathbf{v}^T)(I + \alpha\mathbf{u}\mathbf{v}^T) \\ &= I + \alpha\mathbf{u}\mathbf{v}^T + \mathbf{u}\mathbf{v}^T + \alpha(\mathbf{u}\mathbf{v}^T)(\mathbf{u}\mathbf{v}^T) \\ &= I + \alpha\mathbf{u}\mathbf{v}^T + \mathbf{u}\mathbf{v}^T + \alpha\mathbf{u}(\mathbf{v}^T\mathbf{u})\mathbf{v}^T \end{aligned}$$

   As $\mathbf{v}^T\mathbf{u}$ is a scalar, we can:

$$\begin{aligned} I &= I + \alpha\mathbf{u}\mathbf{v}^T + \mathbf{u}\mathbf{v}^T + \alpha(\mathbf{v}^T\mathbf{u})(\mathbf{u}\mathbf{v}^T) \\ &= I + (\alpha + 1 + \alpha(\mathbf{v}^T\mathbf{u}))\mathbf{u}\mathbf{v}^T \end{aligned}$$

   Since $\mathbf{u}\mathbf{v}^T \neq 0$ (as A is supposed to be non-singular and its inverse must exist) an expression for α is:

$$\alpha + 1 + \alpha(\mathbf{v}^T\mathbf{u}) = 0$$

$$\alpha(1 + \mathbf{v}^T\mathbf{u}) = -1$$

$$\alpha = \frac{-1}{1 + \mathbf{v}^T\mathbf{u}}$$

   We can see that for any $\mathbf{u}$ and $\mathbf{v}$ where:

$$\mathbf{v}^T\mathbf{u} = -1$$

the previous expression cannot be defined and, in those cases, A is singular.

To find the **null space** of A we have:

$$A\mathbf{x} = 0$$

From the statement we have:

$$(I + \mathbf{u}\mathbf{v}^T)\mathbf{x} = 0$$

If $\mathbf{x} = \mathbf{u}$, then:

$$(I + \mathbf{u}\mathbf{v}^T)\mathbf{u} = 0$$

$$\mathbf{u} + \mathbf{u}(\mathbf{v}^T\mathbf{u}) = 0$$

If A is singular, we know $\mathbf{v}^T\mathbf{u} = -1$, so:

$$\mathbf{u} - \mathbf{u} = 0$$

holds. Therefore, $null(A) = span\{\mathbf{u}\}$.

2. (Heath 2.8) Let A and B be any two $n \times n$ matrices.

(a) Prove that $(AB)^T = B^T A^T$.

Answer:

Let denote the i, j entry of $(AB)^T$ as $(AB)^T_{i,j}$. Note that it is the same as the j, i entry of AB. Note also that the j, i entry of AB is equal to the row j of A *dot* the column i of B.
Similarly, the i, j entry of $B^T A^T$ is equal to the row i of B^T *dot* the column j of A^T, which is the same to say the column i of B *dot* the row j of A.
So, we can see that the i, j entries of each side are the *dot* product of the same two vectors. Hence $(AB)^T = B^T A^T$.

(b) If A and B are both non-singular, prove that $(AB)^{-1} = B^{-1}A^{-1}$.

Answer:

First, we have to assume that A and B are invertible. So, A^{-1} and B^{-1} exist and:

$$AA^{-1} = A^{-1}A = I$$

and

$$BB^{-1} = B^{-1}B = I$$

Similarly, if AB is invertible, $(AB)^{-1}$ exists and similar equalities should also hold:

$$AB(AB)^{-1} = (AB)^{-1}AB = I$$

As $(AB)^{-1} = B^{-1}A^{-1}$, then:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1}$$

$$(AB)(B^{-1}A^{-1}) = AIA^{-1}$$

$$(AB)(B^{-1}A^{-1}) = AA^{-1}$$

$$(AB)(B^{-1}A^{-1}) = I$$

and

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B$$

$$(B^{-1}A^{-1})(AB) = B^{-1}IB$$

$$(B^{-1}A^{-1})(AB) = B^{-1}B$$

$$(B^{-1}A^{-1})(AB) = I$$

Since $(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I$, then $(AB)^{-1} = B^{-1}A^{-1}$.

Vector and matrix norms

3. Let $\mathbf{x} \in \mathbb{R}^n$. Two vector norms, $||\mathbf{x}||_a$ and $||\mathbf{x}||_b$, are *equivalent* if $\exists\, c, d \in \mathbb{R}$ such that

$$c||\mathbf{x}||_b \leq ||\mathbf{x}||_a \leq d||\mathbf{x}||_b.$$

Matrix norm equivalence is defined analogously to vector norm equivalence, i.e., $||\cdot||_a$ and $||\cdot||_b$ are equivalent if $\exists\, c, d$ s.t. $c||A||_b \leq ||A||_a \leq d||A||_b$.

(a) Let $\mathbf{x} \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$. For each of the following, verify the inequality and give an example of a non-zero vector or matrix for which the bound is achieved (showing that the bound is tight):

i. $||\mathbf{x}||_\infty \leq ||\mathbf{x}||_2$

$$\max_i |\mathbf{x}_i| \leq \left(\sum_{i=1}^{n} |\mathbf{x}_i|^2\right)^{\frac{1}{2}}$$

Let say $|x_r|$ is the maximum value in $\mathbf{x}$, then:

$$|x_r| \leq (\sum_{i=1}^{n} |\mathbf{x}_i|^2)^{\frac{1}{2}}$$

$$|x_r|^2 \leq |x_1|^2 + |x_2|^2 + \cdots + |x_r|^2 + \cdots + |x_n|^2$$
$$0 \leq |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2$$

It is clear that any addition of squared absolute values is positive.
i.e. $\mathbf{x} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$

$$||\mathbf{x}||_\infty \leq ||\mathbf{x}||_2$$
$$4 \leq 5$$

ii. $||\mathbf{x}||_2 \leq \sqrt{n}||\mathbf{x}||_\infty$

$$\left(\sum_{i=1}^{n} |\mathbf{x}_i|^2\right)^{\frac{1}{2}} \leq \sqrt{n}\max_i |\mathbf{x}_i|$$

Again, let say $|x_r|$ is the maximum value in $\mathbf{x}$, then:

$$\left(\sum_{i=1}^{n} |\mathbf{x}_i|^2\right)^{\frac{1}{2}} \leq \sqrt{n}|x_r|$$

$$\sum_{i=1}^{n} |\mathbf{x}_i|^2 \leq n|x_r|^2$$

$$|x_1|^2 + |x_2|^2 + \cdots + |x_r|^2 + \cdots + |x_n|^2 \leq n|x_r|^2$$

Just in the case that all values of the left hand side are x_r we will have an equality. For other cases, it is clear that the right hand side is greater.
i.e. $\mathbf{x} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$

$$||\mathbf{x}||_2 \leq \sqrt{n}||\mathbf{x}||_\infty$$
$$\sqrt{4+4+4} \leq \sqrt{3} \times 2$$
$$\sqrt{12} \leq 2\sqrt{3}$$
$$\sqrt{4 \times 3} \leq 2\sqrt{3}$$
$$2\sqrt{3} \leq 2\sqrt{3}$$

iii. $||A||_\infty \leq \sqrt{n}||A||_2$

iv. $||A||_2 \leq \sqrt{n}||A||_\infty$

For items `iii` and `iv` the verification is not straightforward. As sections 2.3.2 and 3.6.1 of the textbook state "the matrix norm corresponding to the vector 2-norm is not so easy to compute". However, based on examples 2.4 and 3.17, we can give an example using the matrix:

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 3 & -1 & 4 \end{bmatrix}$$

where $||A||_\infty = 8$ and $||A||_2 = 5,723$, so:

$$||A||_\infty \leq \sqrt{n}||A||_2$$
$$8 \leq \sqrt{3} \times 5,723$$
$$8 \leq 9,91$$

and

$$||A||_2 \leq \sqrt{n}||A||_\infty$$
$$5,723 \leq \sqrt{3} \times 8$$
$$5,723 \leq 13,85$$

This shows that $||\cdot||_\infty$ and $||\cdot||_2$ are equivalent, and that their induced matrix norms are equivalent.

(b) Prove that the equivalence of two vector norms implies the equivalence of their induced matrix norms.

Sensitivity and conditioning

4. (Heath 2.58) Suppose that the $n \times n$ matrix A is perfectly well-conditioned, i.e., cond(A) = 1. Which of the following matrices would then necessarily share this same property?

   (a) cA, where c is any nonzero scalar ✓
   (b) DA, where D is a nonsingular diagonal matrix
   (c) PA, where P is any permutation matrix ✓
   (d) BA, where B is any nonsingular matrix
   (e) A^{-1}, the inverse of A ✓
   (f) A^T, the transpose of A

Linear Systems

5. (Heath 2.4a) Show that the following matrix is singular.

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix}$$

Answer:
Let's compute the determinant of A:

$$\begin{aligned} det(A) &= 1[(2 \times 2) - (3 \times 1)] - 1[(1 \times 2) - (3 \times 0)] + 1[(1 \times 1) - (2 \times 0)] \\ &= 1 - 2 + 1 \\ &= 0 \end{aligned}$$

So, given $det(A) = 0$, A must be singular.

6. For each of the following statements, indicate whether the statement is true or false.

   T/F If a matrix A is singular, then the number of solutions to the linear system $A\mathbf{x} = \mathbf{b}$ depends on the particular choice of right-hand-side $\mathbf{b}$.

   T/F If a matrix A is nonsingular, then the number of solutions to the linear system $A\mathbf{x} = \mathbf{b}$ depends on the particular choice of right-hand-side $\mathbf{b}$.

   T/F If a matrix has a very small determinant, then the matrix is nearly singular.

   T/F If any matrix has a zero on its main diagonal, then it is necessarily singular.

7. Can a system of linear equations $A\mathbf{x} = \mathbf{b}$ have exactly two solutions? Explain your answer.

   Answer:

   No, according to section 2.2 of the textbook, a system of linear equations can have an unique solution, no solution at all or infinitely many solutions. Visually, it is easy to see that lines in a plane (or, more general, hyperplanes in higher dimensions) can only touch in one point (unique solution), be parallel and do not touch (no solution) or be the same line (hyperplane) and have multiple solutions.

   Formally, let say $\mathbf{x_1}$ and $\mathbf{x_2}$ are the two unique solutions. So, from $A\mathbf{x} = \mathbf{b}$ we have:

   $$A\mathbf{x_1} = A\mathbf{x_2} = \mathbf{b}$$

   Now, let assume $\frac{\mathbf{x_1}}{4} + \frac{3\mathbf{x_2}}{4}$ as a possible new solution. So:

   $$\begin{aligned} A(\tfrac{\mathbf{x_1}}{4} + \tfrac{3\mathbf{x_1}}{4}) &= \mathbf{b} \\ \tfrac{1}{4}(A\mathbf{x_1} + 3A\mathbf{x_2}) &= \mathbf{b} \\ \tfrac{1}{4}(\mathbf{b} + 3\mathbf{b}) &= \mathbf{b} \\ \tfrac{1}{4}(4\mathbf{b}) &= \mathbf{b} \\ \mathbf{b} &= \mathbf{b} \end{aligned}$$

   So, $\frac{\mathbf{x_1}}{4} + \frac{3\mathbf{x_2}}{4}$ is a third solution for the system. Indeed, for any λ in $(0, 1)$:

   $$\lambda\mathbf{x_1} + (1 - \lambda)\mathbf{x_2}$$

   should be a solution and the system must have infitinely many solutions. For example, in the abovementioned example $\lambda = \frac{1}{4}$.

LU Factorization and Gaussian Eliminiation

8. For each of the following statements, indicate whether the statement is true or false.

   T/F If a triangular matrix has a zero on its main diagonal, then it is necessarily singular.

   T/F The product of two upper triangular matrices is upper triangular.

   T/F If a linear system is well-conditioned, then pivoting is unnecessary in Gaussian elimination.

   T/F Once the LU factorization of a matrix has been computed to solve a linear system, then subsequent linear systems with the same matrix but different right-hand-side vectors can be solved without refactoring the matrix.

9. Consider LU factorization with partial pivoting of the matrix A which computes

   $$M_{n-1}P_{n-1} \cdots M_3P_3M_2P_2M_1P_1A = U$$

   where P_i is a row permutation matrix interchanging rows i and $j > i$.

(a) Show that the matrix $P_3P_2M_1P_2^{-1}P_3^{-1}$ has the same structure as the matrix M_1.

Answer:

Since $M_j = I - \mathbf{m_j}\mathbf{e_j^T}$, we can show that any $P_kM_jP_k^{-1}$ has the same structure as M_j by:

$$\begin{aligned} P_kM_jP_k^{-1} &= P_k(I - \mathbf{m_j}\mathbf{e_j^T})P_k^{-1} \\ &= P_kP_k^{-1} - P_k\mathbf{m_j}\mathbf{e_j^T}P_k^{-1} \end{aligned}$$

As $\mathbf{e_j^T}P_k^{-1} = \mathbf{e_j^T}$, we have:

$$\begin{aligned} P_kM_jP_k^{-1} &= P_kP_k^{-1} - (P_k\mathbf{m_j})\mathbf{e_j^T} \\ &= I - (P_k\mathbf{m_j})\mathbf{e_j^T} \\ &= M_j' \end{aligned}$$

So, $M_j' = I - (P_k\mathbf{m_j})\mathbf{e_j^T}$ has the same structure that $M_j = I - \mathbf{m_j}\mathbf{e_j^T}$. This concept can easily be applied to $P_3P_2M_1P_2^{-1}P_3^{-1}$ to show it has the same structure as the matrix M_1.

(b) Explain how the above expression is transformed into the form $PA = LU$, where P is a row permutation matrix.

Answer:

For sake of simplicity, let's take $n = 3$, so:

$$M_3P_3M_2P_2M_1P_1A = U$$

We can insert identity matrices in the form of $I_2 = P_3^{-1}P_3$ and $I_1 = P_2^{-1}P_3^{-1}P_3P_2$ after M_2 and M_1 respectively, so:

$$M_3P_3M_2(P_3^{-1}P_3)P_2M_1(P_2^{-1}P_3^{-1}P_3P_2)P_1A = U$$

or

$$M_3(P_3M_2P_3^{-1})P_3(P_2M_1P_2^{-1})P_3^{-1}P_3P_2P_1A = U$$

As we already know, $P_kM_jP_k^{-1} = M_j'$ has the same structure as M_j for $k > j$. In addition, we can group the remaining permutations as P, so:

$$M_3M_2'M_1'PA = U$$

As $M_3M_2'M_1' = L^{-1}$, we have:

$$L^{-1}PA = U$$

$$PA = LU$$

Cholesky Factorization

10. (Heath 2.37) Suppose that the symmetric $(n+1) \times (n+1)$ matrix

$$B = \begin{pmatrix} \alpha & \mathbf{a}^T \\ \mathbf{a} & A \end{pmatrix}$$

is positive definite.

(a) Show that the scalar α must be positive and the $n \times n$ matrix A must be positive definite.

Answer:

Since B is positive definite, then:

$$\mathbf{x}^T B\mathbf{x} > 0$$
$$\mathbf{x}^T \begin{pmatrix} \alpha & \mathbf{a}^T \\ \mathbf{a} & A \end{pmatrix} \mathbf{x} > 0$$

If $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, then:

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha & \mathbf{a}^T \\ \mathbf{a} & A \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} > 0$$
$$\begin{pmatrix} \alpha & \mathbf{a}^T \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} > 0$$
$$\alpha > 0$$

So, α is positive.
Similarly, if $\mathbf{x} = \begin{pmatrix} 0 \\ \mathbf{y} \end{pmatrix}$, where $\mathbf{y}$ is any non-zero vector, then:

$$\begin{pmatrix} 0 & \mathbf{y}^T \end{pmatrix} \begin{pmatrix} \alpha & \mathbf{a}^T \\ \mathbf{a} & A \end{pmatrix} \begin{pmatrix} 0 \\ \mathbf{y} \end{pmatrix} > 0$$
$$\begin{pmatrix} \mathbf{y}^T\mathbf{a} & \mathbf{y}^T A \end{pmatrix} \begin{pmatrix} 0 \\ \mathbf{y} \end{pmatrix} > 0$$
$$\mathbf{y}^T A\mathbf{y} > 0$$

So, A is positive definite.

(b) What is the Cholesky factorization of B in terms of α, $\mathbf{a}$, and the Cholesky factorization of A?

Answer:

Following section 2.5.1 of the textbook, we have $B = LL^T$, we can express LL^T in terms of α and $\mathbf{a}$:

$$B = \begin{pmatrix} \alpha & \mathbf{a}^T \\ \mathbf{a} & A \end{pmatrix} = \begin{pmatrix} \sqrt{\alpha} & 0 \\ \frac{\mathbf{a}}{\sqrt{\alpha}} & L' \end{pmatrix} \begin{pmatrix} \sqrt{\alpha} & \frac{\mathbf{a}^T}{\sqrt{\alpha}} \\ 0 & L'^T \end{pmatrix}$$

Where

$$L'L'^T = A - \frac{\mathbf{a}\mathbf{a}^T}{\sqrt{\alpha}}$$

is the Cholesky factorization of A.