Tridip Ray
ISI, Delhi

First Midterm Exam (10 September 2023)

- Answer *all* the questions. You have 3 hours to write this exam.

1. [20 marks]

Let A be an $m \times n$ matrix. *Row rank* of A is the rank of the set of row vectors, $\{A_1, A_2, ..., A_m\}$. *Column rank* of A is the rank of the set of column vectors, $\{A^1, A^2, ..., A^n\}$. Let $r =$ *row rank* of A, $c =$ *column rank* of A.

- **Rank Theorem:** *For any $m \times n$ matrix A, row rank of $A =$ column rank of A.*

- In Homework 1 you have proved that $r \geq c$ in the following way.

Suppose that $\boldsymbol{r < c}$. Reordering rows or columns of A does not affect its row or column rank. So choose a row basis for A which we may assume consists of the rows $\{A_1, A_2, ..., A_r\}$, and a column basis which we may assume consists of the columns $\{A^1, A^2, ..., A^c\}$. Let $\hat{A}_i = (a_{i1}, a_{i2}, ..., a_{ic})$ and consider the system of equations

$$\hat{A}_i \cdot y = 0, \ i = 1, 2, ..., r.$$

You can prove that this system of equations has a *nonzero* solution $\bar{y}$.

Since $\{A_1, A_2, ..., A_r\}$ is a row basis, it follows from the Basis Theorem that, for all $k = 1, 2, ..., m$, $A_k = \sum_{i=1}^{r} \mu_{ik} A_i$ for some real numbers μ_{ik}. Hence

$$\hat{A}_k = (a_{k1}, a_{k2}, ..., a_{kc}) = \sum_{i=1}^{r} \mu_{ik} \hat{A}_i,$$

and so

$$\hat{A}_k \cdot \bar{y} = \sum_{i=1}^{r} \mu_{ik} \left(\hat{A}_i \cdot \bar{y} \right) = 0, \text{ for all } k = 1, 2, ..., m.$$

You can prove that this leads to a *contradiction* which shows that our supposition $r < c$ is wrong, that is, it must be that $r \geq c$.

- **Question:** Complete the proof of the Rank Theorem by proving carefully that $r \leq c$.

2. [15 marks]

Let $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 = x_3\}$. Find a basis of S, showing your procedure clearly.

3. [15 marks]

(a) Let $y \in S \subset \mathbb{R}^n$ and $(x^1, x^2, ..., x^r)$ is a set of basis vectors of S. Prove that there exists a unique set of scalars $\lambda_1, \lambda_2, ..., \lambda_r$ such that $y = \lambda_1 x^1 + \lambda_2 x^2 + ... + \lambda_r x^r$.

(b) Let $y \in S \subset \mathbb{R}^n$ and $(z^1, z^2, ..., z^m)$ be an *arbitrary* set of vectors, $z^i \in S$, $i = 1, 2, ..., m$. Does there exist a unique set of scalars $\lambda_1, \lambda_2, ..., \lambda_m$ such that $y = \lambda_1 z^1 + \lambda_2 z^2 + ... + \lambda_m z^m$? Provide a rigorous argument for your answer. In case a unique set of scalars does exist, explain clearly how the vectors $(y, z^1, z^2, ..., z^m)$ are related.

4. [50 marks]

Recall the following theorem:

Theorem: *Let A be a symmetric $n \times n$ matrix. A is negative definite if and only if all its n leading principal minors alternate in sign, starting with negative. (That is, the r-th leading principal minor, A_r, $r = 1, 2, ..., n$, has the same sign as $(-1)^r$.)*

In the following steps we will prove this theorem. The proof has two major ingredients: the principal of induction and the theory of partitioned matrices. First a brief introduction to the theory of partitioned matrices.

- **Partitioned Matrices:** Let A be a $m \times n$ matrix. A *submatrix* of A is a matrix formed by discarding some entire rows and/or columns of A. A *partitioned matrix* is a matrix which has been partitioned into submatrices by horizontal and/or vertical lines which extend along entire rows or columns of A. For example,

$$A = \left(\begin{array}{cc|c|ccc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \end{array}\right),$$

which we can write as

$$A = \left(\begin{array}{c|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \end{array}\right).$$

Each submatrix A_{ij} is called a *block* of A.

Suppose that A and B are two $m \times n$ matrices which are partitioned in the same way, that is,

$$A = \left(\begin{array}{c|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \end{array} \right) \text{ and } B = \left(\begin{array}{c|c|c} B_{11} & B_{12} & B_{13} \\ \hline B_{21} & B_{22} & B_{23} \end{array} \right)$$

where A_{11} and B_{11} have the same dimensions, A_{12} and B_{12} have the same dimensions, and so on. Then A and B can be added as if the blocks are scalar entries:

$$A + B = \left(\begin{array}{c|c|c} A_{11} + B_{11} & A_{12} + B_{12} & A_{13} + B_{13} \\ \hline A_{21} + B_{21} & A_{22} + B_{22} & A_{23} + B_{23} \end{array} \right).$$

Similarly, two partitioned matrices A and C can be multiplied, treating the blocks as scalars, if the blocks are all of a size such that the matrix multiplication of blocks can be done. For example, if

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right) \text{ and } B = \left(\begin{array}{c|c|c} B_{11} & B_{12} & B_{13} \\ \hline B_{21} & B_{22} & B_{23} \end{array} \right),$$

then

$$AB = \left(\begin{array}{c|c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} & A_{11}B_{13} + A_{12}B_{23} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} & A_{21}B_{13} + A_{22}B_{23} \end{array} \right)$$

so long as the various matrix products $A_{ij}B_{jk}$ can be formed. For example, A_{11} must have as many columns as B_{11} has rows, and so on.

- Next we need two lemmas.
  - **Lemma 1:** *If A is a positive or negative definite matrix, then A is nonsingular.*
  - **Lemma 2:** *Suppose that A is a symmetric matrix and that Q is a nonsingular matrix. Then, $Q^T AQ$ is a symmetric matrix, and A is positive (negative) definite if and only if $Q^T AQ$ is positive (negative) definite.*

- Now we proceed to prove the theorem by using induction on the size n of A. The result is trivially true for 1×1 matrices. It is straightforward to verify the theorem (*you do not have to do it*) directly for 2×2 symmetric matrices by completing the square in the corresponding quadratic form on $\mathbb{R}^2$:

$$f(x_1, x_2) = (x_1, x_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = ax_1^2 + 2bx_1x_2 + cx_2^2.$$

So we will suppose that the theorem is true for $n \times n$ matrices and prove it to be true for $(n+1) \times (n+1)$ matrices.

- Let A be an $(n+1)\times(n+1)$ symmetric matrix. Write A_j for the $j \times j$ leading principal submatrix of A for $j = 1, 2, ..., n+1$. By the inductive hypothesis the theorem is true for $n \times n$ matrices. In part (a) we will prove that if sign of $|A_r|$ is the same as $(-1)^r$, $r = 1, 2, ..., n+1$, then A is negative definite. In part (b) we will prove the converse: A is negative definite implies that sign of $|A_r|$ is the same as $(-1)^r$, $r = 1, 2, ..., n+1$.

(a) [28 marks]

The inductive hypothesis is given, and assume that sign of $|A_r|$ is the same as $(-1)^r$, $r = 1, 2, ..., n+1$.

(i) Argue that A_n is invertible.

– Partition A as

$$A = \left(\begin{array}{c|c} A_n & a \\ \hline a^T & a_{n+1,n+1} \end{array}\right), \text{ where } a = \begin{pmatrix} a_{1,n+1} \\ \vdots \\ a_{n,n+1} \end{pmatrix}.$$

Let $d = a_{n+1,n+1} - a^T (A_n)^{-1} a$, let I_n denote the $n \times n$ identity matrix, and let 0_n denote the $n \times 1$ column vector of all 0s.

(ii) Verify that

$$A = \left(\begin{array}{c|c} I_n & 0_n \\ \hline (A_n^{-1}a)^T & 1 \end{array}\right) \left(\begin{array}{c|c} A_n & 0_n \\ \hline 0_n^T & d \end{array}\right) \left(\begin{array}{c|c} I_n & A_n^{-1}a \\ \hline 0_n^T & 1 \end{array}\right) \equiv Q^T B Q.$$

(iii) Show that $|A| = d \cdot |A_n|$, and argue that $d < 0$.

(iv) Let X be an arbitrary $(n+1)$-vector. Write $X = \begin{pmatrix} x \\ x_{n+1} \end{pmatrix}$, where x is an n-vector.
Argue that $X^T B X = x^T A_n x + d \cdot (x_{n+1})^2 < 0$.

(v) Conclude that A is negative definite.

(b) [22 marks]

The inductive hypothesis is given, and assume that A is negative definite. (Note that A is $(n+1) \times (n+1)$).

(i) Prove that A_n is negative definite.

(ii) Argue that sign of $|A_r|$ is the same as $(-1)^r$, $r = 1, 2, ..., n$.

- So we need to prove only that the sign of determinant of A itself is $(-1)^{n+1}$.
- Since A_n is invertible, we can once again write A as $Q^T BQ$ as in part (a) (ii) and conclude that $|A| = d \cdot |A_n|$ still holds.

(iii) Argue that B is negative definite.

(iv) Choose X suitably in part (a) (iv) to show that $d < 0$.

(v) Conclude that the sign of $|A|$ is $(-1)^{n+1}$, that is, sign of $|A_r|$ is the same as $(-1)^r$, $r = 1, 2, ..., n+1$.