Results done in MTH206M

Monday 30$^{\text{th}}$ September, 2024

1. Finite Dimensional Vector Spaces

(i) A non-empty subset S of $\mathbb{R}^n$ forms a subspace iff

$$[x, y \in S \text{ and } \alpha \in \mathbb{R}] \implies [\alpha x + y \in S].$$

(a) For a subset $A = \{x_1, \ldots, x_r\}$ of $\mathbb{R}^n$, a linear combination of the vectors in A is $\alpha x_1 + \cdots + \alpha_r x_r$ where $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$, and $\mathrm{Sp}(A) = \{\alpha x_1 + \cdots + \alpha_r x_r \mid \alpha_1, \ldots, \alpha_r \in \mathbb{R}\}$. Span of an infinite set is defined as the set of all linear combinations of finitely many vectors in that set.

(b) $\mathrm{Sp}(A)$ is a subspace and $A \subseteq \mathrm{Sp}(A)$. Further, $\mathrm{Sp}(A)$ is the smallest subspace containing A.

(c) $A = \mathrm{Sp}(A) \iff A$ is a subspace.

(d) $\mathrm{Sp}(A)$ is called the subspace generated by A and for a subspace S, a set A is called a generating set if $S = \mathrm{Sp}(A)$.

(e) $A \subseteq B \implies \mathrm{Sp}(A) \subseteq \mathrm{Sp}(B)$ and $\mathrm{Sp}(\mathrm{Sp}(A)) = \mathrm{Sp}(A)$.

(ii) A set of vectors $x_1, \ldots, x_r$ is linearly independent (LI) if

$$[\alpha_1, \ldots, \alpha_r = 0] \implies [\alpha_r = \cdots = \alpha_r = 0].$$

A set is linearly dependent if it is not linearly independent. An infinite set $A \subseteq \mathbb{R}^n$ is LI if every finite subset of A is linearly independent.

(a) The vectors $x_1, \ldots, x_r$ are LI iff x_j belongs to the span of $x_1, \ldots, x_{j-1}$ for some j such that $1 \leq j \leq k$.

(b) If A is LI and $y \notin A$. Then $A \cup \{y\}$ is linearly dependent iff $y \in \mathrm{Sp}(A)$.

(c) If $A = \{x_1, \ldots, x_r\}$ and $B = \{y_1, \ldots, y_{r+1}\}$ are LI subsets of $\mathbb{R}^n$, then there exists a $y_j \in B \setminus A$ such that $A \cup \{y_j\}$ is LI.

(iii) Every subspace S of $\mathbb{R}^n$ has a finite generating set with cardinality less or equal to n. Such a set is called a basis of S and every basis of S has the same cardinality. This cardinality is called the dimension of S, denoted by $d(S)$.

(a) The vectors in a basis are LI, and adding any other vector (of the same subspace) to a basis makes the set linearly dependent.

(b) Every LI subset A of S can be extended to a basis of S, and every generating set C of S contains a basis of S.

(c) For any two subspaces S and T with $S \subseteq T$, $d(S) \leq d(T)$ and the equality holds iff $S = T$.

(iv) For any two subspaces S and T,

$$\mathrm{Sp}(S \cup T) = \{x + y \mid x \in S \text{ and } y \in T\} = S + T.$$

This is called sum of S and T.

(a) $d(S + T) = d(S) + d(T) - d(S \cap T)$ (Modular law).

(b) The sum is called a direct sum if $S \cap T = \{0\}$. If the sum is direct, any vector in $S + T$ can be expressed in a unique way as $x + y$ with $x \in S$ and $y \in T$.

(c) A subspace $T \subseteq \mathbb{R}^n$ is called a complement of a subspace $S \subseteq \mathbb{R}^n$ if $S + T = \mathbb{R}^n$ and $S \cap T = \{0\}$. If T is a complement of S then clearly S is a complement of T and we say that S and T are complementary subspaces.

(d) Every subspace has a complement, but, in general, complement is not unique. It is unique if and only if the subspace is $\mathbb{R}^n$ or $\{0\}$.

(v) If S and T are complementary subspaces, then for any vector u, where $u = x + y$ with $x \in S$ and $y \in T$, x is called the projection of u into S along T.

  (a) Suppose projection into S along T is denoted by P, then

    i. $P(u) = u$ iff $u \in S$

    ii. $P(u) = 0$ iff $u \in T$

    iii. $P(P(u)) = P(u)$

    iv. $P(\alpha x + y) = \alpha P(x) + P(y)$.

2. Matrix Theory

(i) For matrices, $A_{m\times n} = ((a_{ij}))$, $B_{m\times n} = ((b_{ij}))$ and $C_{n\times p} = ((c_{ij}))$,

  (a) αA is the matrix $((\alpha a_{ij}))$,

  (b) $A + B$ is the matrix $((a_{ij} + b_{ij}))$, and

  (c) AC is the matrix $((d_{ij}))$ where $d_{ij} = \sum_{k=1}^{n} a_{ik}c_{kj}$.

  (a) A vector $x \in \mathbb{R}^n$ is a matrix of order $n \times 1$.

  (b) For a matrix $A_{m\times n}$, transpose of denoted by $A^T_{n\times m}$ is the matrix $((a_{ji}))$. A is symmetric if $A = A^T$.

  (c) We have the following results involving matrix operations:

(1) $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$ (commutativity of addition)
(2) $\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}$ (associativity of addition)
(3) $\mathbf{0}+\mathbf{A}=\mathbf{A}+\mathbf{0}=\mathbf{A}$
(4) $\mathbf{A}+(-\mathbf{A})=(-\mathbf{A})+\mathbf{A}=\mathbf{0}$
(5) $\alpha(\mathbf{A}+\mathbf{B})=\alpha\mathbf{A}+\alpha\mathbf{B}$
(6) $(\alpha+\beta)\mathbf{A}=\alpha\mathbf{A}+\beta\mathbf{A}$
(7) $\alpha(\beta\mathbf{A})=(\alpha\beta)\mathbf{A}$
(8) $1\cdot\mathbf{A}=\mathbf{A}$
(9) $\mathbf{A}(\mathbf{BC})=(\mathbf{AB})\mathbf{C}$ (associativity of matrix multiplication)
(10) $\mathbf{A}(\alpha\mathbf{B})=\alpha(\mathbf{AB})$
(11) $\mathbf{A}(\mathbf{B}+\mathbf{C})=\mathbf{AB}+\mathbf{AC}$ (distributivity)
(12) $(\mathbf{B}+\mathbf{C})\mathbf{A}=\mathbf{BA}+\mathbf{CA}$ (distributivity)
(13) $\mathbf{A}\cdot\mathbf{I}=\mathbf{A}$ where $\mathbf{A}$ is of order $m\times n$ and $\mathbf{I}$ is of order $n\times n$
(14) $\mathbf{I}\cdot\mathbf{A}=\mathbf{A}$ where $\mathbf{I}$ is of order $m\times m$ and $\mathbf{A}$ is of order $m\times n$
(15) $0\cdot\mathbf{A}=\mathbf{0}$ (0 in $0\cdot\mathbf{A}$ is a scalar)
(16) $\mathbf{0}\cdot\mathbf{A}=\mathbf{0}$
(17) $\mathbf{A}\cdot\mathbf{0}=\mathbf{0}$
(18) $(\mathbf{A}+\mathbf{B})^{\mathrm{T}}=\mathbf{A}^{\mathrm{T}}+\mathbf{B}^{\mathrm{T}}$
(19) $(\alpha\mathbf{A})^{\mathrm{T}}=\alpha\,\mathbf{A}^{\mathrm{T}}$
(20) $(\mathbf{AB})^{\mathrm{T}}=\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$
(21) $(\mathbf{A}^{\mathrm{T}})^{\mathrm{T}}=\mathbf{A}$

(d) Product of two upper triangular matrices is upper triangular and product of two lower triangular matrices is lower triangular.

(e) For a square matrix $A_{n\times n}$, $\mathrm{tr}(A)=\sum_{i=1}^{n}a_{ii}$. If AB and BA are both defined, then $\mathrm{tr}(AB)=\mathrm{tr}(BA)$.

(ii) For a matrix $A_{m\times n}$, ith row is denoted by $a_{i\cdot}$ and jth column is denoted by $a_{\cdot j}$. Further, $\mathcal{R}(A)$ is the row space of A where $\mathcal{R}(A)=\mathrm{Sp}(\{a_{1\cdot},\ldots,a_{m\cdot}\})=\{x^TA\mid x\in\mathbb{R}^m\}$. Similarly, column space of A is $\mathcal{C}(A)=\mathrm{Sp}(\{a_{\cdot 1},\ldots,a_{\cdot n}\})=\{Ax\mid x\in\mathbb{R}^n\}$. Note that $\mathcal{R}(A)\subseteq\mathbb{R}^n$ and $\mathcal{C}(A)\subseteq\mathbb{R}^m$.

(a) For two matrices $A_{m\times n}$ and $B_{n\times p}$, $\mathcal{R}(AB)\subseteq\mathcal{R}(B)$ and $\mathcal{C}(AB)\subseteq\mathcal{C}(A)$. Also, $\mathcal{C}(AB)=\{ABx\mid x\in\mathbb{R}^p\}=\{Ax\mid x\in\mathcal{C}(B)\}$.

(b) For two matrices $A_{m\times n}$ and $B_{n\times p}$, $AB=[Ab_{\cdot 1}\cdots Ab_{\cdot p}]$, i.e., the columns of AB are $Ab_{\cdot 1},\ldots,Ab_{\cdot p}$. Similarly, rows of AB are $a_{1\cdot}B,\ldots,a_{m\cdot}B$.

(c) $a_{\cdot j}=Ae_j$ and $a_{i\cdot}=e_i^TA$

(d) For two matrices A and B,

$$[\mathcal{C}(A)\subseteq\mathcal{C}(B)]\iff[A=BC\text{ for some matrix }C],\text{ and}$$

$$[\mathcal{R}(A)\subseteq\mathcal{R}(B)]\iff[A=CB\text{ for some matrix }C].$$

(e) The rank of a matrix A, denoted by $\rho(A)$, is the common value of $d(\mathcal{R}(A))$ and $d(\mathcal{C}(A))$, i.e., $\rho(A) = d(\mathcal{R}(A)) = d(\mathcal{C}(A))$.

(f) For a matrix A and every submatrix B of A, $\rho(A) \geq \rho(B)$.

(g) $\rho(AB) \leq \min\{\rho(A), \rho(B)\}$.

(iii) The inverse of a square matrix $A_{n\times n}$ is defined as a matrix $B_{n\times n}$ such that $BA = I_n$ (or $AB = I_n$).

(a) Inverse exists if and only if the matrix has full rank. Further, if exists, it is unique. For a square matrix A, we denote it's inverse (if exists) by A^{-1}.

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

(c) $\rho(A) = \rho(AB)$ if B is non-singular, i.e., inverse of B exists.

(iv) The null space of a matrix $A_{m\times n}$ is $\mathcal{N}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$. $\mathcal{N}(A)$ is a subspace and $d(\mathcal{N}(A)) = n - \rho(A)$. For matrices $A_{m\times n}$ and $B_{n\times p}$,

(a) $\mathcal{R}(A) = (\mathcal{N}(A))^{\perp}$,

(b) $\mathcal{N}(B) \subseteq \mathcal{N}(AB)$, and

(c) $\rho(AB) = \rho(B) - d(\mathcal{C}(B) \cap \mathcal{N}(A))$ and $\rho(AB) \geq \rho(A) + \rho(B) - n$.

(v) For two complementary subspaces S and T of $\mathbb{R}^n$, we call a matrix $P_{n\times n}$ a projector matrix (into S along T) if Px is the projection of x into S along T. A matrix $P_{n\times n}$ is a projection matrix if there exist two complementary subspaces S and T such that P is a projector into S along T.

(a) Given S and T, Projector matrix always exists and is unique.

(b) P is a projection matrix if and only if it is idempotent. Further, it projects into $\mathcal{C}(P)$ along $\mathcal{C}(I_n - P)$ $(= \mathcal{N}(P))$.

(c) For a subspace S, a matrix P is called the orthogonal projector if it projects into S along $S^{\perp}$. In general, an $n \times n$ matrix P is called an orthogonal projection matrix if it is the orthogonal projector into S for some subspace $S \subseteq \mathbb{R}^n$.

(d) A matrix Q is an orthogonal projection matrix if and only if $Q^TQ = Q$. Note that this is stronger than Q is idempotent.

(vi) A system of linear equations is of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

where a_{ij}'s and b_i's are known and x_j's are unknown. A solution to this system of linear equations is a point $x \in \mathbb{R}^n$ such that is satisfies all the equation.

(a) The above system can be written as $Ax = b$ where $A_{m\times n} = ((a_{ij}))$ and $b^T = (b_1, \ldots, b_m)$.

(b) A system of linear equations may have no solution, or a unique solution, or infinitely many solutions.

(c) A system of linear equations $Ax = b$ is said to be consistent if it has at least one solution, and insonsistent if there is NO solution.

(d) If $Ax = b$ is consistent, then by general solution we mean an expression which gives all possible solutions, while a specific solution is called a particular solution.

(e) A system of linear equations is called homogeneous if $b = 0_{m\times 1}$. A homogeneous system is always consistent and the set of all solutions is $\mathcal{N}(A)$. It has a unique solution if and only if $\mathcal{N}(A) = \{0\}$.

(f) A general system of equations is consistent iff $b \in \mathcal{C}(A)$. This is equivalent to $\rho(A) = \rho(A \mid b)$ where $[A \mid b]_{m\times n+1} = [a_{\cdot 1} \cdots a_{\cdot n} \;\; b]$.

(g) The set of all solutions to $Ax = b$, denoted by $\mathcal{S}_b^A = \{x \mid Ax = b\}$, is given by $u + \mathcal{N}(A)$ where $u_{n\times 1}$ is particular solution to $Ax = b$, i.e., $Au = b$.

(h) A general system has

i. no solution if $\rho(A) < \rho(A \mid b)$,

ii. a unique solution if $\rho(A) = \rho(A \mid b) = n$, and

iii. infinitely many solutions if $\rho(A) = \rho(A \mid b) < n$.

(i) A matrix $F_{m\times n}$ with $\rho(F) = r$ is said to be in Reduced Echelon Form (REF) if

i. the first r rows of F are non-null and the remaining rows are null,

ii. the first non-null element in the ith row occurs at the p_ith position for $i = 1(1)r$ with $p_1 < p_2 < \cdots < p_r$, and

iii. $f_{.p_i} = e_i^m$.

(j) Every matrix can be transformed into a matrix in REF by performing a set of row operations on the matrix. There are three types of row operations

i. R_{ij}: swapping the ith row with the jth row,

ii. $R_i(\alpha)$: multiplying the ith row by α for some $\alpha \neq 0$,

iii. $R_{ij}(\beta)$: replace the ith row by sum of ith row and β times the jth row.

(k) Each row operation is equivalent to pre-multiply the matrix by a matrix, called an elementary matrix. The elementary matrices for the row operations R_{ij}, $R_i(\alpha)$, and $R_{ij}(\beta)$ are E_{ij}, $E_i(\alpha)$, and $E_{ij}(\beta)$, respectively where

$$E_{ij} = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \end{bmatrix},$$

$$E_i(\alpha) = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \alpha & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{bmatrix},$$

$$\text{and } E_{ij}(\beta) = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \cdots & 1 & \cdots & \beta & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}.$$

(l) The elementary matrices are non-singular, and $(E_{ij})^{-1} = E_{ij}$, $(E_i)(\alpha)^{-1} = E_i(\frac{1}{\alpha})$, and $(E_{ij}(\beta))^{-1} = E_{ij}(-\beta)$.

(m) Suppose performing a set of row operations $R_1, R_2, \ldots, R_q$, a matrix A can be reduced to a matrix in REF, F. Then

$$E_q \cdots E_1 A = F$$

where E_l is the corresponding elementary matrix for R_l. Further, to solve a system of linear equations $Ax = b$ is equivalent to solving $Fx = d$ where $d = E_q \cdots E_1 b$.

(n) To solve $Fx = d$, we first find a particular solution u of $Fx = d$ and then find $\mathcal{N}(F)$. Finally, the set of all solution to $Ax = b$ is $u + \mathcal{N}(F)$.

(o) Further, the non-null rows of F will give us a basis of $\mathcal{R}(A)$, and the p_1th, p_2th,..., p_rth columns of A (**not the columns of *F***) will give us a basis of $\mathcal{C}(A)$.

(p) If $\rho(A) = n$,

$$F = \left[\begin{array}{c} I_n \\ \hline 0_{m-n\times n} \end{array}\right]$$

and if $\rho(A) = m$,

$$F = \left[\begin{array}{c|c} I_m & 0_{m\times n-m} \end{array}\right].$$

(vii) For a square matrix $A_{n\times n}$, $\det(A)$ is defined as

$$det(A) = \sum (-1)^{\sigma_n\{(1,j_1),\ldots,(n,j_n)\}} a_{1j_1} \cdots a_{nj_n},$$

where the sum is over all possible values of $\{j_1, \ldots, j_n\}$, i.e., all possible permutations of $\{1, \ldots, n\}$.

(a) In the definition, $\sigma_n\{(1, j_1), \ldots, (n, j_n)\}$ denotes the number of negative pairs out of the ${}^{n}C_2$ that can be formed by the elements of the set $\{(1, j_1), \ldots, (n, j_n)\}$. A pair $\{(i, j_i), (k, j_k)\}$ is a negative pair if either $i < k$ and $j_i > j_k$ or $i > k$ and $j_i < j_k$. Note that as $\{j_1, \ldots, j_n\}$ is a permutation of $\{1, \ldots, n\}$, $j_i \neq j_k$.

(b) $\det(A) = \det(A^T)$.

(c) $\det(U)$ where U is triangular matrix is the product of diagonal matrices.

(d) Interchanging two rows (or columns) of A reverses the sign of $\det(A)$.

(e) $\det(A)$ is linear in rows (or columns): fix $i \in \{1, \ldots, n\}$. Let ith row of A be $a_{i\cdot} = \alpha x + y$ where $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, i.e.,

$$A = \begin{bmatrix} a_{1\cdot} \\ \vdots \\ a_{i-1\cdot} \\ \alpha x + y \\ a_{i+1\cdot} \\ \vdots \\ a_{n\cdot} \end{bmatrix}.$$

Then $\det(A) = \alpha \det(B) + \det(C)$ where

$$B = \begin{bmatrix} a_{1\cdot} \\ \vdots \\ a_{i-1\cdot} \\ x \\ a_{i+1\cdot} \\ \vdots \\ a_{n\cdot} \end{bmatrix}$$

and

$$C = \begin{bmatrix} a_{1\cdot} \\ \vdots \\ a_{i-1\cdot} \\ y \\ a_{i+1\cdot} \\ \vdots \\ a_{n\cdot} \end{bmatrix}.$$

(f) Let A be a matrix such that $a_{i\cdot} = 0$ (or $a_{\cdot i} = 0$) for some $i = \{1, \ldots, n\}$. Then $\det(A) = 0$.

(g) Let A be a matrix such that $a_{i\cdot} = a_{k\cdot}$ (or $a_{\cdot i} = a_{\cdot k}$) for some $i \neq k$. Then $\det(A) = 0$.

(h) $\det(A) = 0$ iff A is non-singular.

(i) Let A and B be two square matrices of order n. Then, $\det(AB) = \det(A)\det(B)$.

(j) If A is non-singular, then $\det(A^{-1}) = (\det(A))^{-1}$.