Module – 2.2

MATRICES:

AN OVERVIEW

S. Lakshmivarahan

School of Computer Science

University of Oklahoma

Norman, Ok – 73069, USA

varahan@ou.edu

DEFINITION, BASIC OPERATION

- **An mxn real matrix A is a rectangular array of mn real numbers arranged in m rows and n columns as**

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}] \in R^{mxn}$$

- R^{mxn} – set of all real mxn matrices
- Row index i: $1 \leq i \leq m$, column index j: $1 \leq j \leq n$
- When m = n, A is called a square matrix of size or order n
- If $a_{ij} = 0$ for all i,j, then A is called a zero or null matrix

CROSS-SECTIONS OF A MATRIX

- Let $A \in R^{nxn}$
- a_{i*} - i^{th} row of A – row vector of size n
- a_{*j} – j^{th} column of A – column vector of size n

$$A = [a_{*1}, a_{*n}, \ldots, a_{*n}] = \begin{bmatrix} a_{1*} \\ a_{2*} \\ \vdots \\ a_{n*} \end{bmatrix}$$ <- column partition of A

row partition of A

- $[a_{11}, a_{22}, \ldots, a_{nn}]$ – principal diagonal
- Diagonals parallel to principal diagonal and above (below) the principal diagonal are called super (sub) diagonals

OPERATIONS ON MATRICES

- A, B, C are matrices in R^{nxn}, x, y, z ∈ R^n, a, b, c are in R

- Sum/difference: C = A ± B -> $c_{ij} = a_{ij} + b_{ij}$

  (element wise sum/difference)

- Scalar multiple: C = aA -> $c_{ij} = aa_{ij}$

- Matrix-vector produc: y = Ax, $y_i = \sum_{j=1}^{n} a_{ij} x_j \ , 1 \ \leq i \leq n$

  or $y = \sum_{j=1}^{n} a_{*j} x_j$ - Linear combination of the columns of A by elements of x

OPERATIONS ON MATRICES

- Matrix-matrix product: C = AB

  1. Inner product: $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$ $\quad 1 \leq i \leq n, 1 \leq j \leq n$

  2. Saxpy: $c_{*j} = \sum_{i=1}^{n} a_{*j} b_{ij}$

  3. Outer product: $C = \sum_{i=1}^{n} a_{*j} b_{j*}$

- AB ≠ BA – matrix product is not commutative.

OPERATIONS ON MATRICES

- 1) Transpose of $A \in R^{mxn}$ denoted by $A^T \in R^{nxm}$ - columns of A are the row of A^T

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \quad A^T = \begin{bmatrix} 1 & 3 \\ 2 & 2 \\ 3 & 1 \end{bmatrix}$$

  - a) $(A^T)^T = A$
  - b) $(A + B)^T = A^T + B^T$
  - c) $(AB)^T = B^T A^T$

- 2) $A \in R^{nxn}$ trace of A = tr(A) = $\sum_{i=1}^{n} a_{ii}$ = sum of diagonal elements
  - a) tr: R^{nxn} -> R, a function of the vector space of nxn matries
  - b) $tr(A) = tr(A^T)$
  - c) tr(A + B) = tr(A) + tr(B)
  - d) tr(αA) = αtr(A)
  - e) tr(AB) = tr(BA)
  - f) tr(ABC) = tr(BCA) = tr(CAB)
  - g) $tr(ABA^{-1}) = tr(B)$

DETERMINANT OF A, B ∈ R^{nxn}

- Determinant of A denoted by det(A)

  det(A) = $\sum_{j=1}^{n} a_{ij} A_{ij}$

  A_{ij} = cofactor of a_{ij} = $(-1)^{i+j} M_{ij}$

  M_{ij} = minor of a_{ij} =determinant of the (n-1)x(n-1) matrix obtained by deleting the i^{th} row and j^{th} column of A

a) A is nonsingular if det(A) ≠ 0 and singular otherwise
b) det(A) = det(A^T)
c) det(AB) = det(A)det(B)
d) det(A^{-1}) = $\frac{1}{\det(A)}$ if A is nonsingular

SPECIAL MATRICES

- $A \in R^{nxn}$ is symmetric if $A^T = A$
- $A = Diag(d_1, d_2, \ldots d_n)$ is a diagonal matrix of size n when

  $a_{ii} = d_i$

  $a_{ij} = 0$ if $i \neq j$
- $I_n = Diag(1,1, \ldots 1)$ is the identity matrix of size n
- A is upper triangular if $a_{ij} = 0$ if $i < j$
- A is lower triangular if $a_{ij} = 0$ if $i > j$
- A is tridiagonal is $a_{ij} \neq 0$ if $|i - j| \leq 1$

  $= 0$ otherwise
- A is orthogonal if $A^T = A^{-1}$

SPECIAL MATRICES

- $A \in R^{nxn}$ is skew symmetric if $A^T = -A$. That is

  $a_{ii} = -a_{ji}$ if $i \neq j$

  $a_{ij} = 0$ if $i = j$

- Let $A \in R^{nxn}$ $A_s = \frac{1}{2}(A + A^T)$ – symmetric part of A

  $A_{ss} = \frac{1}{2}(A - A^T)$ – skew symmetric part of A

- $A = A_s + A_{ss}$ – Additive decomposition of A
- Let $A \in R^{nxm}$. Then $AA^T \in R^{nxn}$ and $A^TA \in R^{mxm}$ are symmetric and are called the Grammian of A

RANK OF A MATRIX $A \in R^{mxn}$

- **Row(column) rank of A = number of linear independent rows(columns) of A**
- Row rank of A = column rank of A = Rank(A)
- $0 \le Rank(A) \le \min\{m,n\}$
  a) $Rank(A) = Rank(A^T)$
  b) $Rank(A + B) \le Rank(A) + Rank(B)$
  c) $Rank(AB) \le \min\{ Rank(A), Rank(B)\}$
  d) Let $A = xy^T$ – outer product matrix: Rank(A) = 1
  e) $A \in R^{nxn}$ nonsingular if
    - $\det(A) \neq 0$
    - Rank(A) = n

INVERSE OF A NONSINGULAR MATRIX $A \in R^{nxn}$

- **Inverse of A denoted by A^{-1}: $AA^{-1} = A^{-1}A = I_n$, the identity matrix**

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

a) $(A^{-1})^{-1} = A$

b) $(AB)^{-1} = B^{-1}A^{-1}$ (A, B are nonsingular)

c) $(A^T)^{-1} = (A^{-1})^T = A^{-T}$

SHERMAN-MORRISON-WOODBURY (SMW) FORMULA – INVERSE UNDER PERTURBATION

a) $(I_n + cd^T)^{-1} = I_n - \frac{cd^T}{1+d^Tc}$ $c, d \in R^n$

b) $(A + cd^T)^{-1} = A^{-1} - \frac{A^{-1}cd^TA^{-1}}{1+ d^TA^{-1}C}$ $A \in R^{nxn}$ – non singular, $c, d \in R^n$

c) $(A + CD^T)^{-1} = A^{-1} – A^{-1}C[I_k + D^TA^{-1}C]^{-1}D^TA^T$ $A \in R^{nxn}$, $B \in R^{kxk}$ are non singular $C, D \in R^{nxk}$

d) $(A + CBD^T)^{-1} = A^{-1} – A^{-1}C[B^{-1} + D^TA^{-1}C]^{-1}D^TA^{-1}$

PROOF OF SMW - FORMULA

- Let A ∈ R^{nxn}, and B ∈ R^{kxk} be non-singular. Let C, D ∈ R^{nxk}
- Let

$$\Lambda = \begin{bmatrix} A & C \\ D^T & B \end{bmatrix} \text{ and } \Lambda^{-1} = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$$

- $\Lambda\Lambda^{-1} = \begin{bmatrix} AP + CR & AQ + CS \\ D^TP + BR & D^TQ + BS \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & I_k \end{bmatrix}$

PROOF OF SMW - FORMULA

- Equating off – diagonal elements:

$$AQ + CS = 0 \quad \Rightarrow Q = -A^{-1}CS$$

$$D^TP + BR = 0 \quad \Rightarrow R = -B^{-1}D^TP$$

-> (1)

- Equating diagonal elements and using (1):

$$AP + CR = I_n \quad \Rightarrow P = (A - CB^{-1}D^T)^{-1} \quad \text{-> (2)}$$

$$D^TQ + BS = I_k \quad \Rightarrow S = (B - D^TA^{-1}C)^{-1} \quad \text{-> (3)}$$

PROOF OF SMW - FORMULA

- $\Lambda\Lambda^{-1} = \begin{bmatrix} PA + QD^T = I_n & PC + RB = 0 \\ RA + SD^T = 0 & RC + SB = I_k \end{bmatrix}$
- $PA + QD^T = I_n$

$\Rightarrow P = A^{-1} + QD^TA^{-1}$

$= A^{-1} + A^{-1}CSD^TA^{-1}$ [using (1)]

$= A^{-1} + A^{-1}C[B - D^TA^{-1}C]^{-1}D^TA^{-1}$ [using (3)]

PROOF OF SMW - FORMULA

- Using the definition of P in (2):

$$[A - CB^{-1}D^T]^{-1} = A^{-1} + A^{-1}C[B - D^TA^{-1}C]^{-1}D^TA^{-1} \quad \text{-> (4)}$$

- This proves the formula (d) in slide 12 by replacing B^{-1} in (4) by -B
- Setting $B = -I_k$, we get formula (c)
- Setting B = –1 and k = 1, we get formula (b)
- Setting B = –1 and $A = I_n$, we get formula (a)

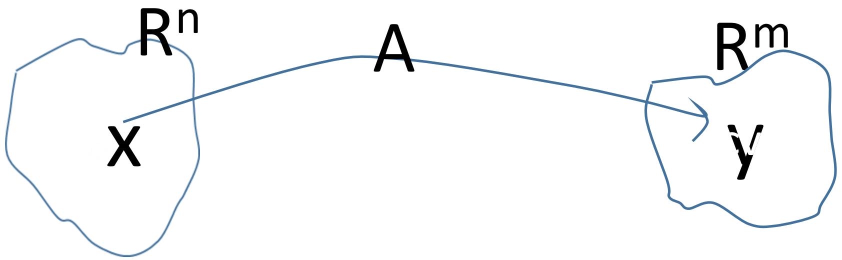
MOORE-PENROSE/GENERALIZED INVERSE

- Let $A \in R^{mxn}$ and $A^+ \in R^{nxm}$
- A^+ is called Moore-Penrose/Generalized inverse of A if
  a) $AA^+A = A$
  b) $A^+AA^+ = A^+$
  c) $(A^+A)^T = A^+A$ – A^+A is symmetric
  d) $(AA^+)^T = AA^+$ – AA^+ is symmetric
- Let A be of full rank. Then

  $A^+ = (A^TA)^{-1}A^T$ if $m > n$

  $A^+ = A^T(AA^T)^{-1}$ if $m < n$
- When n = m and A is non-singular, $A^+ = A^{-1}$, and $A^+A = AA^+ = I_n$

MATRICES AS LINEAR TRANSFORMATION



- Let $A \in R^{mxn}$
- Then $A: R^n \rightarrow R^m$ where $y = Ax \in R^m$

  when $x \in R^n$
- A is called a linear transformation of R^n to R^m with the properties:

  $A(x + y) = Ax + Ay$ for $x, y \in R^n$

  $A(ax) = aAx$ for $x \in R^n$ and $a \in R$
- $\text{Range}(A) = \{ y \in R^m \mid y = Ax \text{ for all } x \in R^n \} \subseteq R^m$
- $\text{NULL}(A) = \text{Ker}(A) = \{ x \in R^n \mid Ax = 0 \}$

  Ker(A) denotes the Kernel of A

EXAMPLES OF OPERATION

- Q matrix $\in R^{nxn}$ is called orthogonal if $Q^T = Q^{-1}$, $QQ^T = Q^TQ = I_n$
- Q: R^n -> R^n is called an orthogonal operator
- Q = $\begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix}$ is an orthogonal matrix – rotation operator
- If x $\in R^{2,}$ then Q rotates x by an angle θ in the anti-clockwise direction
- Let y = Qx and Q is orthogonal

  Then

$$||y||_2^2 = (Qx)^T(QX) = x^T(Q^TQ)x = x^Tx = ||x||_2^2$$

This is, length of a vector is invariant under orthogonal transformation

COORDINATE TRANSFORMATION

- $B_1 = \{ e_1, e_2, \dots e_n\}$ be the standard basis for R^n
- $B_2 = \{ g_1, g_2, \dots g_n\}$ be a new basis for R^n
- $E = [e_1, e_2, \dots e_n] \in R^{nxn}$ and $G = [g_1, g_2, \dots g_n] \in R^{nxn}$
- Then, for $1 \le i \le n$, express the new basis using the old basis :

$$g_i = t_{1i}e_1 + t_{2i}e_2 + \dots + t_{ni}e_n$$

Thus

$$[g_1, g_2, \dots g_n] = [e_1, e_2, \dots e_n] \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{bmatrix}$$

COORDINATE TRANSFORMATION

$G = ET$ where $T = [t_{ij}]$ – non-singular

- Let $U \in R^n$. Let

$U = Ex = e_1x_1 + e_2x_2 + \ldots + e_nx_n$ in B_1

$U = Gx^* = g_1x_1^* + g_2x_2^* + \ldots + g_nx_n^*$ in B_2

- Then $Ex = Gx^*$ -> $x = (E^{-1}G)x^* = Tx^*$

  Coordinates of U in B_1 and B_2 are related as $x = Tx^*$

SIMILARITY TRANSFORMATION

- Let x and y $\in R^n$ in the standard basis B_1 for R^n.
- Let A : R^n -> R^n be a linear operator: y = Ax
- Let T be linear transformation of the basis B_1 to a new basis B_2
- Let x* and y* be the representation of vector x and y in the new basis
- x = Tx* , y = Ty* and let y = Ax
- Then y = Ty* = Ax = ATx*

  or y* = $(T^{-1}AT)$x*
- $(T^{-1}AT)$ is the representation of A in the new basis B_2
- The transformation form A -> $T^{-1}AT$ is called similarity transformation

CONGRUENT TRANSFORMATION

- Let $A \in R^n$ and $B \in R^{nxn}$ be non-singular

- Transformation from $A \to B^TAB$ is called congruence transformation

ADJOINT OPERATOR

- Let $A: R^n \to R^n$ be the matrix that denotes the linear operator in R^n
- Define a new linear operator A^* as:

  $$<Ax, y> = <x, A^*y>$$

  Then, A^* is called the adjoint of A
- Since $<Ax, y> = (Ax)^T y = x^T A^T y = x^T (A^T y) = <x, A^T y> = <x, A^* y>$

  It follows that $A^* = A^{T.}$ Therefore, adjoint of A is given by A^T.
- If $A = A^T$ when A is symmetric, A is called self-adjoint operator
  a) $(A^*)^* = A$
  b) $(aA)^* = aA^*$
  c) $(A + B)^* = A^* + B^*$
  d) $(AB)^* = B^*A^*$
  e) if A^{-1} exists, then $(A^{-1})^* = (A^*)^{-1}$

EXISTENCE OF SOLUTION TO LINEAR SYSTEM

- Let $A \in R^{mxn}$, $b \in R^m$. Then

  $Ax = b$ has a solution only when $b \in Range(A)$

- $NULL(A^T) = \{ y \in R^m \mid A^T y = 0 \}$
- Let $b \in Range(A)$ and $y \in NULL(A^T)$. Then $b^T y = (Ax)^T y = xA^T y = 0$

  Therefore, Range(A) and $NULL(A^T)$ are mutually orthogonal.

- Fredholm's alternative: Given $A \in R^{mxn}$, then exactly one of the two statements is true:

  1) $Ax = b$ has a solution or

  2) $A^T y = 0$ has a solution such that $y^T b \neq 0$

- Let $A \in R^{nxn}$, $b \in R^n$
- Non-homogenous system $Ax = b$ has a solution only when A is non-singular and $x = A^{-1} b$
- Homogenous system $Ax = 0$ has a non-trivial solution only when A is singular

BILINEAR AND QUADRATIC FORMS

- Let $A \in R^{mxn}$ and f_A: R^m x R^n -> R given by $f_A(x, y) = x^TAy$, $x \in R^m$, $y \in R^n$ is called a bilinear form associated with A

- Let $A \in R^{nxn}$ and f_A: R^n -> R given by $Q_A(x) = x^TAx$, $x \in R^n$ is called a quadratic form associated with A

- n = 2, x = $(x_1, x_2)^T$,   A = $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

  $Q_A(x) = a_{11}x_1^2 + (a_{12} + a_{21})x_1x_2 + a_{22}x_2^2$

PROPERTY OF QUADRATIC FORM

- Since $Q_A(x)$ is a scalar,

$$Q_A(x) = x^TAx = (x^TAx)^T = x^TA^Tx = Q_{A^T}(x)$$

- Hence

$$Q_A(x) = \frac{1}{2}[x^TAx + x^TA^Tx] = x^T[\frac{A+A^T}{2}]x = x^TA_sx$$

where $A_s = \frac{A+A^T}{2}$, symmetric part of A

- Hence we are interested in $Q_A(x)$ only for a symmetric matrix

POSITIVE DEFINITE MATRIX (PD)

- Let A ∈ R^{nxn}. A is said to be PD if

$$x^TAx > 0 \text{ for all } x \neq 0$$
$$= 0 \text{ only if } x = 0$$

- An equivalent definition:
  - Principal minors of all orders are positive.
  - The eigenvalues of A are all positive.
- To get an understanding of the constraints on the elements of A:

$$\text{Let } A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

- $Q_A(x) = ax_1^2 + 2bx_1x_2 + cx_2^2$

$$= a(x_1 + \frac{b}{a}x_2)^2 + (c - \frac{b^2}{a})x_2^2$$

- This is greater than zero if $ac > b^2$
- This A is PD if $a > 0$, $c>0$ and $ac > b^2$.

EIGENVALUES AND EIGENVECTOS

- Let $A \in R^{nxn}$. If there exists $V \in R^n$ and $\lambda \in R$ such that $AV = \lambda V$, then λ is the eigenvalue and V is the eigenvector of A
- (λ, V) is the solution of the homogenous system

$$(A - \lambda I)V = 0$$

- For V to be non-trivial vector, $P(\lambda) = \det(A - \lambda I) = 0$ where $P(\lambda)$ is the n^{th} degree polynomial called the characteristic polynomial
- Let $\lambda_1, \lambda_2, \ldots \lambda_n$ be the n-roots of $P(\lambda) = 0$
- λ_i's are real or complex. Complex roots come in conjugate pairs
- When A is symmetric, λ_i's are real
- When A is symmetric and positive definite (SPD), λ_i's are real and positive

EXAMPLE OF EIGENVALUES

$$A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix} \quad \lambda_1 = 9, \lambda_2 = 4$$

- The eigenvector $V_1 = \frac{1}{\sqrt{5}}\binom{-1}{2}$, $V_2 = \frac{1}{\sqrt{5}}\binom{2}{1}$

  Clearly, $V_1 \perp V_2$
- Let A be SPD and (λ_i, V_i): $Av_i = \lambda_i V_i$
- Then $\{V_1, V_2, \dots V_n\}$ is an orthonormal system
- Let $V = [V_1, V_2, \dots V_n] \in R^n$, $V^T = V^{-1}$
- Then $AV = V\Lambda$, $\Lambda = \text{Diag}(\lambda_1, \lambda_2, \dots \lambda_n)$

  $A = V\Lambda V^T$ – Eigendecomposition of A

  $A = \sum_{i=1}^{n} \lambda_i V_i V_i^T$
- Spectral radius of A = $\rho(A) = \max_i\{|\lambda_i|\}$

SINGULAR VALUES OF A

- Let A be non-singular. The gramians A^TA and AA^T are then symmetric and positive definite.
- Let $(A^TA)V_i = \lambda_i V_i$ with

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \ldots \geq \lambda_n > 0$$

- Verify that $(AA^T)U_i = \lambda_i U_i$ where

$$U_i = \frac{1}{\sqrt{\lambda_i}} AV_i$$

- A^TA and AA^T share the same set of eigenvalues
- Define $\sigma_i = \sqrt{\lambda_i}$ for $1 \leq i \leq n$
- $\{\sigma_i\}$ are the singular values of A

MATRIX NORMS

- Let A ∈ R^{nxn}. Norm of A is a measure of the size of A

- Frobenius norm of A = $||A||_F$ = $[\sum_{i,j=1}^{n} a_{ij}^2]^{½}$

- Operator form: (Induced norm)

$$||A||_p = sup_{||x||_p \neq 0} \frac{||Ax||_p}{||x||_p} = \max_{||x||_p = 1} ||Ax||_p$$

- Setting P = 1, 2, ∞, we get various matrix norms
- Inequalities: 1) $\|Ax\| \leq \|A\|\|x\|$

  2) $\|AB\| \leq \|A\|\|B\|$

COMPUTATION OF $||A||_p$

1) $||A||_1 = \max_{\boldsymbol{j}}\{\sum_{i=1}^{n}\lceil a_{ij}\rceil\}$ – *Column norm*

2) $||A||_\infty = \max_{\boldsymbol{i}}\{\sum_{j=1}^{n}\lceil a_{ij}\rceil\}$ – *Row norm*

3) $||A||_2 = \sigma_1$ where σ_1^2 is one max eigenvalue of A^TA. σ_1 is the largest singular value of A

4) When A is symmetric, $A^TA = A^2$ and $A^2x = \lambda^2x$ if $Ax = \lambda x$

Therefore, $||A||_2 = |\lambda_{max}|$, λ_{max} = maximum eigenvalue of A

5) For A symmetric: $\rho(A) = ||A||_2 = |\lambda_{max}|$, called spectral radius

EQUIVALENCE OF MATRIX NORMS: $A \in R^{nxn}$

1) $||A||_2 \leq [||A||_1 \, ||A||_\infty]^{½}$
2) $\frac{1}{\sqrt{n}}||A||_\infty \leq ||A||_2 \leq \sqrt{n}||A||_\infty$
3) $\frac{1}{\sqrt{n}}||A||_1 \leq ||A||_2 \leq \sqrt{n}||A||_1$
4) $||A||_2 \leq ||A||_F \leq \sqrt{n}||A||_2$
5) ρ(A) $\leq ||A||$, any matrix norm

CONDITION NUMBER OF A MATRIX

- Let A ∈ R^{nxn}.
- Condition number $\mathcal{K}_p$(A) = $||A||_p||A^{-1}||_p$ and its values is norm dependent
- Since I = AA^{-1} => 1 = $||\mathrm{I}||_p \leq ||A||_p||A^{-1}||_p$ = $\mathcal{K}$(A)
- Thus, $1 \leq \mathcal{K}(A) \leq \infty$
- Spectral condition number of symmetric matrix A

$$\mathcal{K}_2(A) = ||A||_2||A^{-1}||_2 = \frac{|\lambda_{max}|}{|\lambda_{min}|}$$

- Spectral condition number of A non-singular

$$\mathcal{K}_2(A) = ||A||_2||A^{-1}||_2 = \frac{\sigma_1}{\sigma_2} = \frac{\sigma_{max}}{\sigma_{min}}$$ where

σ_i is the i[th] singular values of A with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$

RELATION BETWEEN CONDITION NUMBERS

1. $\frac{1}{n}\mathcal{K}_2(A) \leq \mathcal{K}_1(A) \leq \mathrm{n}\mathcal{K}_2(A)$

2. $\frac{1}{n}\mathcal{K}_\infty(A) \leq \mathcal{K}_2(A) \leq \mathrm{n}\mathcal{K}_\infty(A)$

3. $\frac{1}{n^2}\mathcal{K}_1(A) \leq \mathcal{K}_\infty(A) \leq n^2\mathcal{K}_1(A)$

Note: Since $||A||_1$ and $||A||_\infty$ norms are easily computed, we can estimate $\mathcal{K}_2(A)$ using the above relations.

RELATION BETWEEN det(A) AND $\mathcal{K}$(A)

- Let A = Diag($\frac{1}{2}, \frac{1}{2}, \ldots \frac{1}{2}$)

  => det(A) = $\frac{1}{2^n}$ -> 0 as n -> ∞

  =>$\mathcal{K}_p(A)$ = 1 for p = 1, 2, ∞

- Let B ∈ R^{nxn}, upper triangular:

$$a_{ij} = \begin{cases} 1\ if\ i = j \\ -1\ if\ i > j \\ 0\ if\ i < j \end{cases}$$

- det(B) = 1 and $\mathcal{K}_\infty(A)$ = n -> ∞ as n ->∞
- Thus, there is no correlation between det(A) and $\mathcal{K}$(A)

SENSITIVITY OF SOLUTION OF LINEAR SYSTEM

- Let Ax = b be the given system.
- Let (A + εB)y = (b + εf) be the perturbed system
- εB and εf are the perturbation and the vector respectively, and ε > 0 but small
- The relative error in the solution is given

$$\frac{\|y - x\|}{\|x\|} \leq \mathcal{K}(\mathrm{A})[\varepsilon \frac{\|B\|}{\|A\|} + \varepsilon \frac{\|f\|}{\|b\|}]$$

- Since $\mathcal{K}(\mathrm{A}) \geq 1$, the errors in A and b are amplified in the solution
- Larger $\mathcal{K}(\mathrm{A})$ is, more sensitive the system is to round-off error in A and b

Exercise

3.1) Give an examples of A and B where $AB \neq BA$ and $AB = BA$

3.2) Verify $(AB)^T = B^T A^T$

3.3) Verify $tr(AB) = tr(BA)$

3.4) Prove $\det(A^{-1}) = \frac{1}{\det(A)}$ (Hint: $AA^{-1} = I$)

3.5) Verify $A = \begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix}$ is an orthogonal matrix

Plot $y = Ax$ when $x = (1, 1)$ for $\theta = 30^o, 60^o, 90^o, 120^o, 150^o$

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

3.7) Verify $A^+ = (A^T A)^{-1} A^T$ and $A^+ = A^T(AA^T)^{-1}$ satisfy the definition of the generalized/ Moore – Penrose inverse

Exercise

3.8) Find the range and kernel of

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

3.9) Verify $(AB)^* = B^*A^*$ and $(A^{-1})^* = (A^*)^{-1}$ where recall that A^* is the adjoint of A

3.10) If $AV = \lambda V$, then $A^2V = \lambda^2 V$ and $A^kV = \lambda^k V$

3.11) If A is non singular, then A^TA and AA^T are SPD

3.12) If $(A^TA)V_i = \lambda_i V_i$ and $u_i = \frac{1}{\sqrt{\lambda_i}} AV_i$, verify that $(AA^T)u_i = \lambda_i u_i$

REFERENCES

1. G. H. Golub and C. F. Van Loan (1989) Matrix computations *Johns Hopkins university Press (Second edition)*
2. C. D. Meyer (2000) Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia*
3. R. A. Horn and C. R. Johnson (2013) Matrix Analysis *Cambridge university Press (Second edition)*