Inner products and Matrix Norm

CE40282-1: Linear Algebra

Hamid R. Rabiee and Maryam Ramezani
Sharif University of Technology

Inner products

$a \cdot b = \langle a, b \rangle$

- **Definition 1** (Inner product). *A function* $\langle ., . \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *is an inner product if*

  1. $\langle x, x \rangle \geq 0, \langle x, x \rangle = 0 \Leftrightarrow x = 0$ *(positivity)*
  2. $\langle x, y \rangle = \langle y, x \rangle$ *(symmetry)*
  3. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ *(additivity)*
  4. $\langle rx, y \rangle = r\langle x, y \rangle$ *for all* $r \in \mathbb{R}$ *(homogeneity)*

- using properties (2) and (4) and again (2)

$$\langle x, ry \rangle = \langle ry, x \rangle = r\langle y, x \rangle = r\langle x, y \rangle$$

- using properties (2), (3) and again (2).

$$\langle x, y + z \rangle = \langle y + z, x \rangle = \langle y, x \rangle + \langle z, x \rangle = \langle x, y \rangle + \langle x, z \rangle$$

Inner products

خارجی
outer product
$x \otimes y = xy^T$

- The standard inner product is

داخلی ضرب

$$\langle x, y \rangle = x^T y = \sum x_i y_i, \quad x, y \in \mathbb{R}^n.$$

عدد

- The standard inner product between matrices is

$$\langle X, Y \rangle = \boxed{\mathrm{Tr}(X^T Y)} = \sum_i \sum_j X_{ij} Y_{ij}$$

ماتریس

where $X, Y \in \mathbb{R}^{m \times n}$.

Example

$$\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix}$$

- $p_{(x)} = 3 - x + 2x^2$ and $q_{(x)} = 4x + x^2$

$\langle p, q \rangle = 2 + (-4) + 0 = -2$

- $U = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$, $V = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

$\langle U, V \rangle$

$= \langle vec(U), vec(V) \rangle$

$Tr(U^T V) =$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & -1 \\ 0 & 1 \end{bmatrix}$$

$1 - 2 + 1 + 0 = 0$

Norm

- **Definition** (Norm). *A function* $f : \mathbb{R}^n \to \mathbb{R}$ *is a norm if*

1. $f(x) \geq 0,\ f(x) = 0 \Leftrightarrow x = 0$ *(positivity)*
2. $f(\alpha x) = |\alpha| f(x),\ \forall \alpha \in \mathbb{R}$ *(homogeneity)*
3. $f(x+y) \leq f(x) + f(y)$ *(triangle inequality)*

"Entry-wise" matrix norms

- $\|A\|_{p,p} = \|\mathrm{vec}(A)\|_p = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^p\right)^{1/p}$

- special case $p = 2$ is the Frobenius norm, and $p = \infty$ yields the maximum norm
  - Frobenius (Euclidean) norm $\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij})^2}$
    - invariant under rotations (unitary operations)

$UU^T = I$

$$\|A\|_F = \|AU\|_F = \|UA\|_F$$

$$\|A+B\|_F^2 = \|A\|_F^2 + \|B\|_F^2 + 2\langle A, B\rangle$$

$$\|A^*A\|_F = \|AA^*\|_F \le \|A\|_F^2$$

$$\sqrt{\mathrm{trace}(A^*A)} = \sqrt{\sum_{i=1}^{\min\{m,n\}}\sigma_i^2(A)}$$

Singular value of A

  - Max norm The **max norm** is the elementwise norm with $p = q = \infty$:

$$\|A\|_{\max} = \max_{ij}|a_{ij}|.$$

  - sum-absolute-value norm: $\|A\|_{sav} = \sum_{i,j}|A_{i,j}|$

Matrix norms induced by vector norms

- $$||A||_p = \max_{\vec{x}\neq\vec{0}} \frac{||A\vec{x}||_p}{||\vec{x}||_p} = \max_{||\vec{x}||_p=1} ||A\vec{x}||_p$$

$$||\alpha x|| = \alpha ||x||$$

$$||A\alpha x|| = \alpha ||Ax||$$

$$\frac{||A\alpha x||}{||\alpha x||} = \frac{\cancel{\alpha}||Ax||}{\cancel{\alpha}||x||}$$

$$\max \begin{cases} x_1 \text{ ,,,} & x_i \neq 0 \\ x_2 \text{ ,,,} \\ \vdots \\ x_n \text{ ,,,} \end{cases}$$

- Theorem: $||Ax|| \leq ||A||\ ||x||$ for all vectors $||x||$

- Theorem: For all matrices A, B, $||AB|| \leq ||A||\ ||B||$

Matrix norms induced by vector norms

l_1 l_2 l_∞

- The norm of a matrix is a real number which is a measure of the magnitude of the matrix.
- $\|A\|_1 = \max_{1\le j\le n}\left(\sum_{i=1}^{n}|a_{ij}|\right)$
- $\|A\|_\infty = \max_{1\le i\le n}\left(\sum_{j=1}^{n}|a_{ij}|\right)$
- spectral norm (l2) is the largest singular value (the square root of the largest eigenvalue of the matrix gram A
- Example

$\|A\|_2 = \sqrt{\max\{\text{eigenvalue}(A^T A)\}} = \max\{\text{sing}(A)\}$

$l_\infty = \max(11, 6, 3) = 11$

$$B = \begin{bmatrix} 5 & -4 & 2 \\ -1 & 2 & 3 \\ -2 & 1 & 0 \end{bmatrix}$$

$l_1 = \|B\|_1 = 8$

$= \max(8, 7, 5)$

Eigenvectors and Eigenvalues-Part 2

CE40282-1: Linear Algebra

Hamid R. Rabiee and Maryam Ramezani
Sharif University of Technology

Review

- Example

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

$$|A - \lambda I| = \begin{vmatrix} 3-\lambda & -2 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 3\lambda + 2 = 0 \Rightarrow \begin{cases} \lambda_1 = 1 \\ \lambda_2 = 2 \end{cases}$$

$$\left.\begin{array}{l} \lambda_1 = 1 \\ (A - \lambda_1 I)q_1 = 0 \end{array}\right\} \Rightarrow q_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\left.\begin{array}{l} \lambda_2 = 2 \\ (A - \lambda_2 I)q_2 = 0 \end{array}\right\} \Rightarrow q_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

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

$$\Longrightarrow \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Diagonalization

- With similarity transformation Q, matrix A changed to a diagonal matrix $\mathrm{diag}(\lambda_1, \lambda_2)$
- Matrix A has n linear independent eigenvectors

$$Aq_1 = \lambda_1 q_1 = [q_1 \quad q_2 \quad \cdots \quad q_n] \begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \cdots Aq_n = \lambda_n q_n = [q_1 \quad q_2 \quad \cdots \quad q_n] \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \lambda_n \end{bmatrix}$$

$$[Aq_1 \quad Aq_2 \quad \cdots \quad Aq_n] = \underbrace{[q_1 \quad q_2 \quad \cdots \quad q_n]}_{Q} \underbrace{\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}}_{\Lambda}$$

$$A\,[q_1 \quad q_2 \quad \cdots \quad q_n] = Q\Lambda \implies AQ = Q\Lambda$$

$$\Lambda = Q^{-1}AQ \qquad A = Q\Lambda Q^{-1}$$

Example

Let $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Find a formula for A^k, given that $A = PDP^{-1}$, where

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \text{ and } D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

Diagonalizable

Definition

A matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, that is, if $A = PDP^{-1}$ for some invertible matrix P and some diagonal matrix D.

Theorem

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

The columns of P is called an **eigenvector basis** of $\mathbb{R}^n$

An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Diagonalization Process

- Find eigenvalues
- Find eigenvectors
- IF there is n linear independent eigenvectors, then matrix is diagonalizable.
- A similar transform Q can make matrix diagonalizable.
- Columns of Q are eigenvectors.
- Diagonal values of diagonal matrix is the eigenvalues (in the same order)

Non Diagonalizable Matrix

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

$$\Delta(\lambda) = \lambda^2 - 4\lambda + 4 = 0 \implies \sigma(A) = \{2, 2\}$$

$$(A - 2I)\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = 0 \implies \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \alpha_1 = \alpha_2$$

- Can write in form $\Lambda = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$?

$$Q\Lambda = AQ \Rightarrow [q_1 \quad q_2]\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} = A\,[q_1 \quad q_2] \Rightarrow \begin{cases} Aq_1 = 2q_1 \\ Aq_2 = 2q_2 + q_1 \end{cases}$$

$$q_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \implies (A - 2I)q_2 = q_1 \implies \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \implies \beta_1 + \beta_2 = 1$$

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

Jordan canonical form

- "most diagonal" representative from each family of similar matrices;

Jordan's theorem says that every square matrix A is similar to a Jordan matrix J, with Jordan blocks on the diagonal:

$$J = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & J_d \end{bmatrix}.$$

A Jordan block J_i has a repeated eigenvalue λ_i on the diagonal, zeros below the diagonal and in the upper right hand corner, and ones above the diagonal:

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & \lambda_i & 1 \\ 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix}.$$

Jordan canonical form

- Note
  - If A has n distinct eigenvalues, it is diagonalizable and its Jordan matrix is the diagonal matrix $J = \Lambda$.
  - If A has repeated eigenvalues and "missing" eigenvectors, then its Jordan matrix will have $n - d$ ones above the diagonal.
- Example: which are similar?

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & 7 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Jordan canonical form

- Example: for 4*4 matrix with four duplicated eigenvalues
- linear independent eigenvectors
  - 4
  - 3
  - 2
  - 1

$$\left[\begin{array}{c|c|c|c} \lambda & 0 & 0 & 0 \\ \hline 0 & \lambda & 0 & 0 \\ \hline 0 & 0 & \lambda & 0 \\ \hline 0 & 0 & 0 & \lambda \end{array}\right]$$

$$\left[\begin{array}{cc|c|c} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ \hline 0 & 0 & \lambda & 0 \\ \hline 0 & 0 & 0 & \lambda \end{array}\right]$$

$$\left[\begin{array}{ccc|c} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 0 \\ \hline 0 & 0 & 0 & \lambda \end{array}\right] \quad \left[\begin{array}{cc|cc} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ \hline 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{array}\right]$$

$$\left[\begin{array}{cccc} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{array}\right]$$

Conclusion

- Every matrix can convert to Jordan form by a similar transform
- Diagonal form is a special case of Jordan form where all Jordan blocks are 1*1
- Number of Jordan block=number of linear independent eigenvectors
- If a matrix has n linear independent eigenvectors, it is diagonalizable.
- If matrix has one eigenvalue with m duplicates:
  - ?
  - ?
  - ?
  - ?

Symmetric Matrix

$Q = [q_1 \cdots , q_n]$

Theorem $A = A^T$

If A is symmetric, then any two eigenvectors from different eigenspace are **orthogonal**.

$$\left.\begin{aligned} Av_1 &= \lambda_1 v_1 \\ Av_2 &= \lambda_2 v_2 \\ \lambda_1 &\neq \lambda_2 \end{aligned}\right\} \Longrightarrow v_1^T v_2 = 0$$

$\langle v_1, v_2 \rangle = 0$

$A^T = A$

$v_1, v_2 \neq 0$

$\langle v_1, v_2 \rangle = \lambda_1 v_1^T v_2 = (\lambda_1 v_1)^T v_2 = (Av_1)^T v_2 = v_1^T A^T v_2$

$= v_1^T A v_2 = v_1^T \lambda_2 v_2$

$\lambda_1 v_1^T v_2 = \lambda_2 v_1^T v_2$

$\times \; \lambda_1 = \lambda_2$

$\rightarrow (\lambda_1 - \lambda_2) v_1^T v_2 = 0$

$v_1^T v_2 = 0 \rightarrow \langle v_1, v_2 \rangle = 0$

Symmetric Matrix

nullity $(A - \lambda I) = m$ λ

m بار تکرار شده

- Eigenvalues of symmetric matrix are real.

m بار تکرار شده λ

- A symmetric matrix is always diagonalizable.

اثبات X

Q

- A similar transform that diagonalized the symmetric matrix is orthogonal. $Q^{-1} = Q^T$

$$Q^T Q = I, \quad A = Q \Lambda Q^T, \quad \Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}, \lambda_i \in \mathbb{R}$$

$Q^{-1} = Q^T$

Orthogonally Diagonalizable

Theorem

An $n \times n$ matrix A is **orthogonally diagonalizable** if and only if A is a symmetric matrix.

$$A = A^T \implies A = Q\Lambda Q^T,\ \Lambda = \mathrm{diag}\{\lambda_1, \dots, \lambda_n\}$$

$$A = A^T \impliedby A = Q\Lambda Q^T,\ \Lambda = \mathrm{diag}\{\lambda_1, \dots, \lambda_n\}$$

$$A^T = (Q\Lambda Q^T)^T = Q\Lambda^T Q^T = Q\Lambda Q^T = A$$

Spectral Theorem

4 ~ 6

The Spectral Theorem for Symmetric Matrices

An $n \times n$ symmetric matrix A has the following properties:

a. A has n real eigenvalues, counting multiplicities.

b. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation.

c. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.

d. A is orthogonally diagonalizable.

Gram matrix

$G(A) = \boxed{A^T A}$

- Eigenvalues are real. $\sigma(A^T A) \subset \mathbb{R}$
- Eigenvalues are nonnegative.

$B = A^T A$

$Bv = \lambda v$

$\boxed{A^T A} v = \lambda v$

v

$\langle v, A^T A v \rangle$

$= v^T A^T A v = \lambda v^T v$

$\boxed{(Av)^T Av} = \lambda v^T v$

$\|Av\|^2 = \lambda \|v\|^2$

$v \neq 0$

$\boxed{\lambda \geq 0}$

Quadratic form

$$Q(\boldsymbol{x}) = \boldsymbol{x}^T A \boldsymbol{x}$$

Definition

A quadratic form Q is:

a) **positive definite** if $Q(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \neq \mathbf{0}$;
b) **negative definite** if $Q(\boldsymbol{x}) < 0$ for all $\boldsymbol{x} \neq \mathbf{0}$;
c) **indefinite** if $Q(\boldsymbol{x})$ assumes both positive and negative values;
d) **positive semidefinite** if $Q(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x}$;
e) **negative semidefinite** if $Q(\boldsymbol{x}) \leq 0$ for all $\boldsymbol{x}$.

- For diagonal matrix

$x_1 x_2$ $\quad a_i > 0$

$$A = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix} \Longrightarrow x^T A x = a_1 x_1^2 + a_2 x_2^2 + \cdots + a_n x_n^2 > 0$$

Change of Variable in a QF

If $\boldsymbol{x}$ represents a variable vector in $\mathbb{R}^n$, then a **change of variable** is an equation of the form

$$\boldsymbol{x} = P\boldsymbol{y} \quad \text{or equivalently,} \quad \boldsymbol{y} = P^{-1}\boldsymbol{x}$$

where P is an **invertible matrix** and $\boldsymbol{y}$ is a new variable vector in $\mathbb{R}^n$.

- $\boldsymbol{y}$ can be regarded as the coordinate vector of $\boldsymbol{x}$ relative to the basis of $\mathbb{R}^n$ determined by the columns of P.

If the change of variable is made in a quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$, then

$$\boldsymbol{x}^T A\boldsymbol{x} = (P\boldsymbol{y})^T A(P\boldsymbol{y}) = \boldsymbol{y}^T P^T AP\boldsymbol{y} = \boldsymbol{y}^T (P^T AP)\boldsymbol{y}$$

$\boxed{\boldsymbol{y}^T D \boldsymbol{y}}$

D

- The new matrix of the quadratic form is $P^T AP$.
- A is symmetric, so there is an **orthogonal matrix** P such that $P^T AP$ is a diagonal matrix D.
- Then the quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$ becomes $\boldsymbol{y}^T D\boldsymbol{y}$, there is **no cross-product term**.

Quadratic form

Theorem

Let A be an $n \times n$ symmetric matrix. Then a quadratic form $\boldsymbol{x}^T A \boldsymbol{x}$ is:

a) **positive definite** if and only if the eigenvalues of A are **all positive**;
b) **negative definite** if and only if the eigenvalues of A are **all negative**;
c) **indefinite** if and only if A has **both positive and negative** eigenvalues.
d) How about semidefinite?