

**CSE 494 CSE/CBS 598 (Fall 2007): Numerical Linear Algebra for Data
Exploration— Basics Linear Algebra**
Instructor: Jieping Ye

1 Matrices

- Rectangular array of data: elements are real numbers.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n}$$

2 Basic concepts

- vectors
- norms and distances
- eigenvalues, eigenvectors
- linearly independent vectors, basis
- orthogonal bases
- matrices, orthogonal matrices
- orthogonal matrix decompositions: SVD

3 Review of fundamental concepts

3.1 Matrix-vector multiplication

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix} = y$$

- Alternative presentation of matrix-vector multiplication

$$y = Ax = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) x = \sum_{j=1}^n \mathbf{a}_j x_j$$

- The vector y is a linear combination of the columns of A .

3.2 Matrix-matrix multiplication

- Let $A \in \mathbb{R}^{m \times p}$ and $B \in \mathbb{R}^{p \times n}$. Then, $C = AB = (c_{ij}) \in \mathbb{R}^{m \times n}$ is defined as follows:

$$c_{ij} = \sum_{k=1}^p a_{ik}b_{kj}, \text{ for all } i = 1, \dots, m, j = 1, \dots, n.$$

- Each column vector in B is multiplied by A.

3.3 Vector norms

- Measure the “size” of a vector.

- 1-norm: $\|x\|_1 = \sum_{i=1}^n |x_i|$

- 2-norm: $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$

- max-norm: $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$

- All of the above are special cases of the L_p -norm (or p-norm):

$$\|x\|_p = \left(\sum_{i=1}^n x_i^p \right)^{1/p}$$

- Two norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ are equivalent if there exist constants C_1 and C_2 such that

$$C_1\|x\|_\alpha \leq \|x\|_\beta \leq C_2\|x\|_\alpha, \text{ for all } x.$$

- $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent with $x \in \mathbb{R}^n$:

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2.$$

- Generally, a vector norm is a mapping $\mathbb{R}^n \rightarrow \mathbb{R}$, with the properties

- $\|x\| \geq 0$, for all x

- $\|x\| = 0$, if and only if $x = 0$

- $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$

- $\|x + y\| \leq \|x\| + \|y\|$, for all x and y

- How to measure distance between vectors?

- Obvious answer: the distance between two vectors x and y is $\|x - y\|$, where $\|\cdot\|$ is some vector norm.

- Alternative: use the angle between two vectors x and y to measure the distance between them.

- How to calculate the angle between two vectors?

- Angle between vectors

- The inner product between two vectors is defined by $(x, y) = x^T y$.

- This is associated with the Euclidean norm: $\|x\|_2 = (x, x)^{1/2}$.

- The angle θ between two vectors x and y is

$$\cos(\theta) = \frac{(x, y)}{\|x\|_2 \|y\|_2}.$$

- The cosine of the angle between two vectors x and y can be used to measure the similarity between the two vectors: if x and y are close, the angle between them is small, and $\cos(\theta) \approx 1$; if x and y are orthogonal, i.e., $(x, y) = 0$, $\cos(\theta) = 0$ ($\theta = \pi/2$).

- Why not just use the Euclidean distance?

- Example: term-document matrix
- Each entry tells how many times a term appears in the document:

	Doc1	Doc2	Doc3
Term1	10	1	0
Term2	10	1	0
Term3	0	0	0

- Using the Euclidean distance Documents 1 and 2 look dissimilar, and Documents 2 and 3 look similar. This is just due to the length of the documents.
- Using the cosine of the angle between document vectors Documents 1 and 2 are similar to each other and dissimilar to Document 3.

3.4 Linear independence

- Given a set of vectors $\{v_1, v_2, \dots, v_n\} \in \mathbb{R}^m$, with $m \geq n$, consider the set of linear combinations $y = \sum_{j=1}^n \alpha_j v_j$ for arbitrary coefficients α_j 's.
- The vectors $\{v_1, v_2, \dots, v_n\}$ are linearly independent, if $\sum_{j=1}^n \alpha_j v_j = 0$, if and only if $\alpha_j = 0$ for all $j = 1, \dots, n$.
- A set of m linearly independent vectors of \mathbb{R}^m is called a basis in \mathbb{R}^m : any vector in \mathbb{R}^m can be expressed as a linear combination of the basis vectors.

3.5 Matrix rank

- The rank of a matrix is the maximum number of linearly independent column vectors.
- A square matrix $A \in \mathbb{R}^{n \times n}$ with rank n is called nonsingular.
- A nonsingular matrix A has an inverse A^{-1} satisfying

$$AA^{-1} = A^{-1}A = I_n.$$

- What is the rank of an out-product matrix $xy^T \in \mathbb{R}^{m \times n}$ with $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$?

3.6 Range and Null Space

- The range of a matrix $A \in \mathbb{R}^{m \times n}$ is defined by

$$\text{ran}(A) = \{y \in \mathbb{R}^m : y = Ax \text{ for some } x \in \mathbb{R}^n\}.$$

- The null space of A is defined by

$$\text{null}(A) = \{x \in \mathbb{R}^n : Ax = 0\}.$$

- It follows from the definition that $\text{rank}(A) = \dim(\text{ran}(A))$.
 - The dimension of a space S , denoted as $\dim(S)$ denotes the maximum number of linearly independent vectors in S .
- Show that

$$\dim(\text{null}(A)) + \text{rank}(A) = m.$$

3.7 Eigenvalues and eigenvectors

- Let A be a $n \times n$ matrix. The vector $v \neq 0$ that satisfies

$$Av = \lambda v$$

for some scalar λ is called the eigenvector of A and λ is the eigenvalue corresponding to the eigenvector v .

- An example: $A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$

$$Av = \lambda v \rightarrow (A - \lambda I_n)v = 0 \rightarrow |A - \lambda I_n| = 0 \rightarrow \left| \begin{pmatrix} 2 - \lambda & 1 \\ 1 & 3 - \lambda \end{pmatrix} \right| = 0$$

Two eigenvalues $\lambda_1 = 3.62$ and $\lambda_2 = 1.38$. and two eigenvectors:

$$v_1 = \begin{pmatrix} 0.52 \\ 0.85 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0.85 \\ -0.52 \end{pmatrix}$$

3.8 Matrix norms

- Let $\|\cdot\|$ be a vector norm and $A \in \mathbb{R}^{m \times n}$. The corresponding matrix norm is

$$\|A\| = \sup_{x \in \mathbb{R}^n : x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{x \in \mathbb{R}^n : \|x\|=1} \|Ax\|.$$

- Show that

$$\begin{aligned} \|A + B\| &\leq \|A\| + \|B\| \\ \|Ax\| &\leq \|A\| \|x\| \\ \|AB\| &\leq \|A\| \|B\| \end{aligned}$$

- $\|A\|_2 = \left(\max_i \lambda_i(A^T A)\right)^{1/2}$: square root of the largest eigenvalue of $A^T A$.
- $\|A\|_1 = \max_j \sum_{i=1}^m |a_{ij}|$: maximum over columns.
- $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$: maximum over rows.
- Frobenius norm: does not correspond to any vector norm.

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$$

- Define $\text{trace}(B) = \sum_{i=1}^n b_{ii}$ for any matrix $B = (b_{ij}) \in \mathbb{R}^{n \times n}$.
- Show that $\|A\|_F^2 = \text{trace}(AA^T)$.

3.9 Condition number

- Define the condition number of a matrix to be

$$\kappa(A) = \|A\| \|A^{-1}\|.$$

- Consider a matrix $A = \begin{pmatrix} a & 1 \\ 1 & 1 \end{pmatrix}$ with $a \neq 1$. A is nonsingular with

$$A^{-1} = \frac{1}{a-1} \begin{pmatrix} 1 & -1 \\ -1 & a \end{pmatrix}$$

- Nonsingularity is not always enough, as the norm of A^{-1} tends to infinity, as $a \rightarrow 1$.

3.10 Orthogonality

- Two vectors x and y are orthogonal, if $x^T y = 0$.
- Given a set of orthogonal vectors $\{v_1, v_2, \dots, v_n\} \in \mathbb{R}^m$, with $m \geq n$, i.e., $v_i^T v_j = 0$, for $i \neq j$, then they are linearly independent. Why?
- Let the set of orthogonal vectors $v_j, j = 1, \dots, m$ in \mathbb{R}^m be normalized, i.e., $\|v_j\| = 1$. Then they are orthonormal, and constitute an orthonormal basis in \mathbb{R}^m .
- A matrix $V = [v_1, v_2, \dots, v_m]$ is called an orthogonal matrix, if its columns are orthonormal. Prove the following properties of an orthogonal matrix:

- An orthogonal matrix $Q \in \mathbb{R}^{m \times m}$ has rank m .
- $Q^{-1} = Q^T$, that is, $Q^T Q = I_m$, and $Q Q^T = I_m$.
- The Euclidean length of a vector $x \in \mathbb{R}^m$ is invariant under an orthogonal transformation Q , that is,

$$\|Qx\|_2^2 = \|x\|_2^2.$$

- The product of two orthogonal matrices Q and P is orthogonal.

- In the previous example we determined the eigenvectors of the matrix:

$$v_1 = \begin{pmatrix} 0.52 \\ 0.85 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0.85 \\ -0.52 \end{pmatrix}$$

The vectors v_1 and v_2 are orthogonal: $v_1^T v_2 = 0$.

3.11 Eigenvalues and eigenvectors of a symmetric matrix

- The eigenvectors of a symmetric matrix are mutually orthogonal and its eigenvalues are real.

– $A \in \mathbb{R}^{n \times n}$ is a symmetric matrix, if $A = A^T$.

- A symmetric matrix $A \in \mathbb{R}^{n \times n}$ can be written in the form $A = U\Lambda U^T$, where the columns of U (U is an orthogonal matrix) are the eigenvectors of A and Λ is a diagonal matrix, the diagonal elements of Λ which are the corresponding eigenvalues of A . This is called the eigendecomposition of A .

- Example of symmetric matrices: graphs

– The adjacency matrix of an undirected graph is a symmetric matrix.

- **Courant-Fischer Min-max Theorem:**

– If $A \in \mathbb{R}^{n \times n}$ is symmetric, then

$$\lambda_k(A) = \max_{\dim(S)=k} \min_{0 \neq y \in S} \frac{y^T A y}{y^T y}.$$

- Assume both A and E are n -by- n symmetric matrix. Show that

$$\lambda_k(A) + \lambda_n(E) \leq \lambda_k(A + E) \leq \lambda_k(A) + \lambda_1(E).$$

- Assume both A and E are n -by- n symmetric matrix. Show that

$$|\lambda_k(A + E) - \lambda_k(A)| \leq \|E\|_2.$$

3.12 Positive Semi-definite Matrix and Positive Definite Matrix

- A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is positive semi-definite, if and only if $x^T A x \geq 0$, for any $x \in \mathbb{R}^n$.

– All eigenvalues of A are non-negative.

– $X^T A X$ for any $X \in \mathbb{R}^{n \times m}$ is positive semi-definite.

- A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is positive definite, if and only if $x^T A x > 0$, for any $0 \neq x \in \mathbb{R}^n$.

– All eigenvalues of A are positive.

– All principal submatrices of A are positive definite.

– All diagonal entries of A are positive.