线性代数总结

Peng Li @ DaSE ECNU

https://simplelp.github.io/

目录

一、*N*元*N*次线性方程组的矩阵表示与求解

1. 向量及其运算

向量空间

点/向量/有序数组

点积

模

垂直

夹角（余弦值）

加法和数乘

Cauchy-Schwarz不等式：

$$|v \cdot w| \leq \|v\|\|w\|$$

三角不等式：

$$\|v + w\| \leq \|v\| + \|w\|$$

2. 矩阵 (方阵) 与线性方程组

- **矩阵表示**

$$Ax = (\mathbf{u}, \mathbf{v}, \mathbf{w}) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1\mathbf{u} + x_2\mathbf{v} + x_3\mathbf{w}.$$

$$A\mathbf{x} = \begin{pmatrix} (a_{11}, a_{12}, a_{13}) \cdot (x_1, x_2, x_3) \\ (a_{21}, a_{22}, a_{23}) \cdot (x_1, x_2, x_3) \\ (a_{31}, a_{32}, a_{33}) \cdot (x_1, x_2, x_3) \end{pmatrix}$$

- ***A*可逆**：$Ax = b$对于任意b都有唯一解$\Leftrightarrow Ax = 0$只有零解$\Leftrightarrow$方阵A列向量线性无关
- **行图**：超平面交点
- **列图**：列向量线性组合

3. Gauss消元法

- Gauss消元法一般过程：
  - 行变换---消元---回代
- 消元法终止
  - 唯一解： n个方程n个主元（U是可逆上三角阵）
  - 无解或者无穷多解： $0 = c \neq 0$ 或者 $0 = 0$
- 消元法矩阵表示
  - 矩阵乘法定义：$AB = A(\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3) := (A\mathbf{b}_1 \ A\mathbf{b}_2 \ A\mathbf{b}_3)$
  - 单位矩阵/消去矩阵/置换矩阵/初等矩阵/增广矩阵
  - 左乘换行，右乘换列

4. 矩阵的计算

- 矩阵
  - 定义
  - 矩阵相等/零矩阵/方阵/主对角元/对角矩阵/单位阵
  - 矩阵的加法和数乘（满足8条性质）
  - 矩阵的乘法：四种表示；矩阵可交换；不满足消去律；方幂；
  - 分块矩阵
  - 矩阵转置：$(A\mathbf{x})\cdot\mathbf{y}=\mathbf{x}\cdot(A^T\mathbf{y})\quad(*)$；(反)对称矩阵;

矩阵乘法的四种理解

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + a_{i3} b_{3j} + \dots\dots$$

$$AB = A\begin{bmatrix} b_1 & b_2 & \dots & b_p \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_p \end{bmatrix}$$

$$AB = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_m \end{bmatrix} B = \begin{bmatrix} a_1 B \\ a_2 B \\ \vdots \\ \vdots \\ a_m B \end{bmatrix}$$

$$AB = \sum_{k=1}^{n} \underbrace{\begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix}}_{m \times 1} \underbrace{\begin{bmatrix} b_{k1} & \cdots & b_{kp} \end{bmatrix}}_{1 \times p}$$

5. 矩阵 (方阵) 的逆

定义：对方阵 A,若存在矩阵B, 满足$AB = BA = I$,则称 A是可逆的(invertible). 称B是A的逆矩阵(inverse matrix), 记作A^{-1}.

- Gauss-Jordan消元法求A^{-1}

设 A 可逆, 则

$$\left(A \,\vdots\, I_n \right) \longrightarrow \left(I_n \,\vdots\, A^{-1} \right)$$

A 可逆 $\Longrightarrow$ A可表示成一系列初等矩阵的乘积.

矩阵的逆的性质

- (4) 2×2 矩阵 $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 可逆 $\Longleftrightarrow ad-bc\neq 0$, 且 $A^{-1}=\dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$
- (2)若 n 阶方阵 A 和 B 都可逆,则 AB 可逆,且 $(AB)^{-1}=B^{-1}A^{-1}$
- (3)若 A 可逆，则 A^T 也可逆，且 $(A^T)^{-1}=(A^{-1})^T$.
- 下三角矩阵可逆 $\Longleftrightarrow$ 主对角元素都非零.
- 可逆下三角矩阵的逆也是下三角阵.
- 若原矩阵对角元素都是 1, 则逆的对角元也都是1.
- 分块上三角矩阵可逆 $\Longleftrightarrow$ 主对角线上各分块都是可逆的.
- 置换阵 P 满足 $P^{-1}=P^T$.

6. LU分解

- LU分解（LDU分解）

若$A=LU$, 则方程组 $A\mathbf{x}=\mathbf{b}$变为 $L\mathbf{c}=\mathbf{b}$ (下三角形方程组)

$U\mathbf{x}=\mathbf{c}$ (上三角形方程组)

- 高斯消元法的计算量：$O(x^3)$
- 存在性与唯一性
- 对称矩阵的LDL^T
- 设A是一个n阶可逆阵，则存在置换阵P，使得$PA=LU$

LU分解的存在性与唯一性

定理：设可逆矩阵 A 的顺序主子阵 $A_k(k=1,\cdots,n)$ 均为可逆阵，则A 有 LU 分解.

定理：设n 阶可逆阵 A 有 $A=LU$,其中 L 为下三角矩阵 ,U 为上三角矩阵，且 $l_{ii}=1, u_{ii}\neq 0(1\leq i\leq n)$,则分解唯一.

二、一般线性方程组的求解

7. 向量空间

- 如果*Ax=b*有解，则

$$b \in C(A)$$

$$C(A) = x^{*} + N(A)$$

- $N(A)$的求解：阶梯型矩阵

8. 求解齐次线性方程组

- *Ax=0* 基础解系：
  - 主元个数=A的列数-自由变量个数
- 简化行阶梯型的列变换

$$
\begin{tikzcd}
N(A) \arrow[rr, bend left, "P^{-1}"] & & N(R) \arrow[ll, bend left, "P"]
\end{tikzcd}
$$

9. 求解非齐次线性方程组

- 求特解 x^*:

  增广矩阵行变换+令自由变量为0

- 解的一般性讨论:

$$\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 1\\0 & 0\end{pmatrix}\begin{pmatrix}1 & 0 & a\\0 & 1 & b\end{pmatrix}\begin{pmatrix}1 & 0 & a\\0 & 1 & b\\0 & 0 & 0\end{pmatrix}$$

三、向量空间与线性方程组近似解

10. 线性无关、基与维数

- 线性无关
  - 零空间只有0解
- 基
  - 向量空间中的一组线性无关并且能线性表示空间内任意向量的向量
  - 任意两组基中向量一样多
- 维数
  - 基向量的个数
  - $\mathrm{r(AB)} \leq \min\left(\mathrm{r(A)}, \mathrm{r(B)}\right); \mathrm{r(A)} = \mathrm{r}(A^T); r(A+B) \leq r(A) + r(B)$

11. 四个基本子空间的基与维数

$$\boxed{PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}}$$

- 四个基本子空间的基：

基：
1. A的行阶梯矩阵主元对应的行；
2. A的行变换后的非零行；
3. A^T的列空间；

基：行阶梯矩阵主元对应的列

$C(A^T)$
$dim = r$

$C(A)$
$dim = r$

$\mathbb{R}^n$

$\mathbb{R}^m$

$N(A)$
$dim = n - r$

$N(A^T)$
$dim = m - r$

基：Ax=0的基础解系

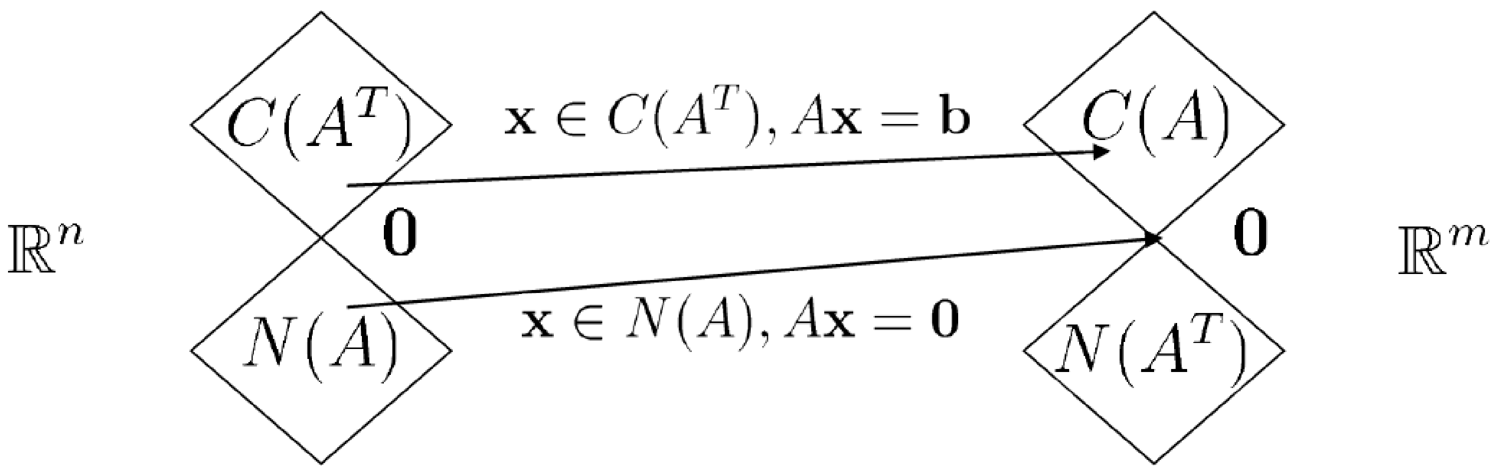
基：
1. A^T零空间的基
2. $EA=U_0$中E的后m-r行

- 维数公式：

$$\dim W_1 + dim W_2 = \dim(W_1 \cap W_2) + \dim(W_1 + W_2)$$

12. 四个基本子空间的正交关系

- S与T正交: $S \cap T = \{0\}; \dim(\mathrm{S}) + \dim(\mathrm{T}) \leq n$
- 四个子空间的正交性与正交补($V^{\perp}$):

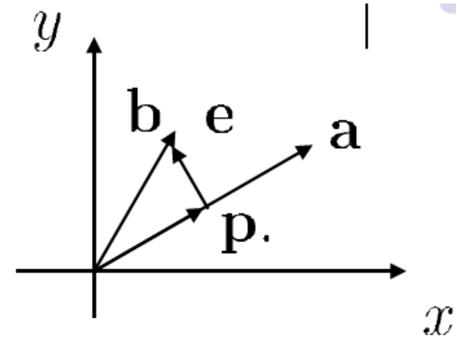


- Ax=b中x在行空间的映射唯一

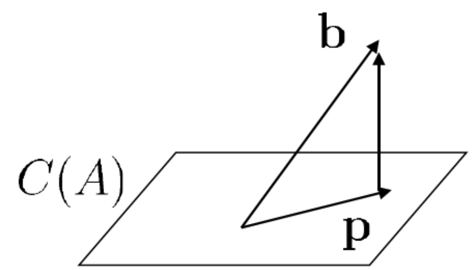
$$\boxed{N(A) = N(A^TA),\ C(A^T) = C(A^TA)}$$

13. 正交投影

- 点在直线上的投影 $S = \dfrac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}$



- 点在平面上的投影 $A(A^TA)^{-1}A^T$



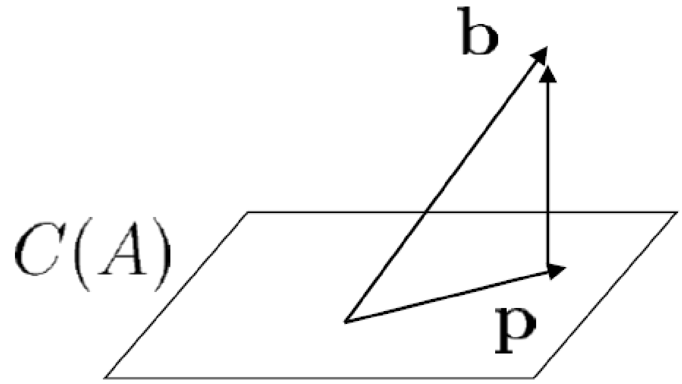
- 一般情形 $A^TA\mathbf{x} = A^T\mathbf{b}$

- 投影矩阵（关于C(P)的投影）：

$$P^2 = P, P^T = P,$$

14. 最小二乘法

- Ax=b的近似解：
  - $\min\limits_{\mathbf{x}\in\mathbb{R}^n} ||A\hat{\mathbf{x}} - \mathbf{b}||$ 的最小值点.
  - 点在平面上的投影



- 应用： 曲线拟合

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_5 & x_5^2 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} y_1 \\ \vdots \\ y_5 \end{pmatrix}, \hat{\mathbf{x}} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

15. Gram-Schmidt正交化

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

- 目的： *C(A’)=C(A)*使得$A'^T A'$更容易计算
- 正交矩阵
  - $Q^{-1} = Q$
  - $\|Qx\| = ||x||$

$$Q^TQ = \begin{pmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{pmatrix} \begin{pmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{pmatrix} = \begin{pmatrix} \mathbf{q}_1^T\mathbf{q}_1 & \cdots & \mathbf{q}_1^T\mathbf{q}_n \\ \vdots & \ddots & \vdots \\ \mathbf{q}_n^T\mathbf{q}_1 & \cdots & \mathbf{q}_n^T\mathbf{q}_n \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = I_n$$

- Gram-Schmidt正交化

$$\mathbf{e}_k = \mathbf{v}_k - (\mathbf{q}_1^T\mathbf{v}_k)\mathbf{q}_1 - \cdots - (\mathbf{q}_{k-1}^T\mathbf{v}_k)\mathbf{q}_{k-1} = \mathbf{w}_k, \ \mathbf{q}_k = \frac{\mathbf{w}_k}{||\mathbf{w}_k||} = \frac{\mathbf{e}_k}{||\mathbf{e}_k||}$$

- 矩阵形式： QR分解

https://wenku.baidu.com/view/6872ac728e9951e79b892700.html

四、行列式

16. 行列式的基本性质

- 几何意义
  - 带符号面积体积
  - 伸缩变换因子
- 行列式的若干性质
- 行列式与初等变换
  - 初等矩阵的行列式
  - 若*A*,*B*是n阶方阵，则$|AB| = |A||B|$
  - $|A + B| \neq |A| + |B|$
  - 若A可逆,$AA^{-1} = I_n \Rightarrow |A||A^{-1}| = 1$
  - 设A是一个正交阵，则|A|=±1

17. 行列式的计算

- 行列式的展开公式
- 利用代数余子式按行按列展开
- 计算方法
  - 化为上/下三角形法
  - 按行按列展开降阶法
- 典型例题
  - 范德蒙行列式
  - 凑分块矩阵

18. Cramer法则和行列式的几何意义

- 代数余子式重要性质

$$\sum_{k=1}^{n} a_{ik}C_{jk} = \begin{cases} D, & i=j, \\ 0, & i \neq j. \end{cases}$$

- 求逆矩阵公式(注意伴随矩阵有转置):

$$\text{设} A \text{ 可逆，则 } A^{-1} = \frac{adj(A)}{|A|}.$$

  - 这是A可逆的情况，A不可逆时呢？
- Cramer法则
- 几何意义：有向长度、面积、体积

五、特征值与特征向量

19. 特征值与特征向量

- 特征值/特征向量/特征方程/特征多项式
- 特征值与特征向量的计算
  - 矩阵A不可逆 $\leftrightarrow$ A有零特征值
  - 投影矩阵的特征值为0和1
  - 反射矩阵的特征值为1和-1
  - 三角矩阵的特征值为其所有对角元
  - 1是马尔可夫矩阵的特征值
  - A^2的特征值与$\mathrm{A-mI}$的特征值
  - p(A)的特征值
  - A^{-1}的特征值
  - 设 n 阶矩阵 $A=(a_{ij})$有 n 个特征值(可能重复) $\lambda_1,\cdots,\lambda_n$,

$$\lambda_1+\cdots+\lambda_n=a_{11}+\cdots+a_{nn}=:trA,$$

$$\lambda_1\cdots\lambda_n=detA.$$

20. 矩阵对角化

- 矩阵A可对角化：$\exists S, S^{-1}AS = \Lambda$
  - 充要条件：A有n个线性无关的特征向量
  - 充分条件：特征值互异则特征向量无关→ 有n个两两互异的特征值则可对角化（若矩阵有相同的特征值，也可能能对角化）
  - 矩阵可逆与矩阵可对角化没有必然的联系
- 可对角化的判断
  - 代数重数：特征值次数 ≥ 几何重数：特征值对应的特征向量的维数
  - 复方阵A可对角化 ↔ 任意$\lambda_i, GM(\lambda_i) = AM(\lambda_i)$
  - 可对角化判断https://zhidao.baidu.com/question/2056215293661831427.html
- 应用
  - 方幂的计算
  - 马尔可夫过程等

21. 实对称矩阵

- 实对称矩阵的性质
  - 实对称矩阵的特征值都是实数
  - 实对称矩阵属于不同特征值的特征向量相互正交
  - 对于实对称矩阵A, ∃正交阵$Q, Q^T AQ = A$
  - 任意实对称矩阵可以表示成秩1投影矩阵的和
  - 实对称矩阵正特征值数与正主元数相同
  - 等等

22. 正定矩阵

- 定义：特征值全为正数的实对称矩阵为正定矩阵
  - 实对称矩阵正定的充要条件（6条）
- 正定矩阵的性质
  - 正定矩阵的和、平方、逆仍然是正定矩阵
  - 正定矩阵C使得$A = C^2$
  - 等等
- 半正定矩阵：实对称矩阵的特征值非负
  - 判定条件（4条）

六、SVD与PCA

23. 谱分解

- 共轭转置矩阵
- 酉矩阵
- 正规矩阵
- 谱定理

定义 6.4　谱定理

若矩阵 $\boldsymbol{A}$ 为一个正规矩阵，则存在酉矩阵 $\boldsymbol{U}$ ，以及对角矩阵 $\boldsymbol{\Lambda}$ ，使得

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H$$

24. 奇异值分解

- 实对称矩阵的特征值分解(EVD)

$$A = U\Lambda U^{-1} = U\Lambda U^{T}$$

- 奇异值分解(SVD)
  - 奇异值
  - 奇异值分解

步骤 1：计算矩阵 $\boldsymbol{A}\boldsymbol{A}^{H}$ 和 $\boldsymbol{A}^{H}\boldsymbol{A}$；

步骤 2：分别计算矩阵 $\boldsymbol{A}\boldsymbol{A}^{H}$ 和 $\boldsymbol{A}^{H}\boldsymbol{A}$ 的特征值和对应的特征向量；

步骤 3：用矩阵 $\boldsymbol{A}\boldsymbol{A}^{H}$ 的特征向量组成矩阵 $\boldsymbol{U}$，矩阵 $\boldsymbol{A}^{H}\boldsymbol{A}$ 的特征向量组成矩阵 $\boldsymbol{V}$；

步骤 4：对矩阵 $\boldsymbol{A}\boldsymbol{A}^{H}$ 和 $\boldsymbol{A}^{H}\boldsymbol{A}$ 的非零特征值求平方根，并对应特征向量的位置，填入 $\boldsymbol{\Sigma}$ 的对角元。

25. PCA降维

- PCA降维思想
- 基于最小重构性的主成分分析
- 基于最大可分性的主成分分析
- 算法实现
  - 样本去中心化
  - 特征值分解/奇异值分解
  - 特征值排序，取前d个特征值对应的特征向量组成正交基
  - 将原来的样本点投影到新空间
  - 输出投影后的样本点

https://www.cnblogs.com/lzllovesyl/p/5243370.html

http://dase.ecnu.edu.cn/mgao/teaching/DataSci_2019_Spring/reference/BOOK_20190217.pdf