1 向量和矩阵的基础

1.1 向量和矩阵的概念

1.1.1 向量

1.1.2 矩阵

定理 1.1.1. 矩阵的乘积满足下列规律:

1. 结合律 $(AB)C = A(BC)$
2. 左分配律 $A(B + C) = AB + AC$, 右分配律 $(B + C)A = BA + CA$。

定理 1.1.2. 分块矩阵乘法, 若 A 的列的分法与 B 的行的分法一致, 就和矩阵乘法一致.

1.1.2.1 特殊结构矩阵

定义 1.1.3. 对称矩阵: $A = A^T$

定理 1.1.4. 对称矩阵 A 的性质

- 存在正交矩阵 Q 使 $Q^TAQ = \Lambda$, Λ 是对角阵
- 存在 n 个 A 的特征向量构成 $\mathbb{R}^n$ 的一个标准正交基

定义 1.1.5. 非奇异矩阵是**满秩矩阵**.

定义 1.1.6. 正交矩阵: $A^T = A^{-1} \Leftrightarrow A$ 的行向量和列向量均为单位矩阵.

定义 1.1.7. k 阶子式: 取 k 行 k 列. **k 阶主子式**: 取 k 行 k 列, 行列序号相等. **顺序主子式**: 取 $[11 : 11], [11 : 22], \ldots, [nn, nn]$.

TODO: Lec 9

定理 1.1.8.

- 行满秩矩阵 A, AA^T 可逆
- 列满秩矩阵 A, A^TA 可逆

1.2 向量空间

1.2.1 向量间线性关系

定义 1.2.1. 如果两个向量组互相可以线性表出，则称为它们**等价**。

1.2.1.1 生成集、基底和坐标

定义 1.2.2. 生成集$\{a_1, a_2, \ldots, a_r\}$ **张成的子空间** $L(a_1, a_2, \ldots, a_r)$ 或 $span(a_1, a_2, ..., a_r)$

定义 1.2.3. 基为线性无关的**生成集**.

定义 1.2.4. 维数 $dim(\mathbb{V}) =$ (**秩**) $rank\{a_1, a_2, \ldots, a_r\}$

1.2.2 线性组合

例 1.2.1 已知 $\beta=(1,2,1,1)^T$, 以及 $\alpha_1=(1,1,1,1)^T$, $\alpha_2=(1,1,-1,-1)^T$, $\alpha_3=(1,-1,1,-1)^T$, $\alpha_4=(1,-1,-1,1)^T$. 试将向量 β 表示成 $\alpha1$, $\alpha2$, $\alpha3$, $\alpha4$ 的线性组合.

解 假设 k_1, k_2, k_3, k_4 为组合系数, 将 β 表示成 $\alpha1$, $\alpha2$, $\alpha3$, $\alpha4$ 的线性组合. 即求解:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix}=\begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}$$

使用初等变换法求解线性方程组:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 2 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 \end{pmatrix}\to\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & -2 & -2 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & -2 & 2 & 0 \end{pmatrix}\to\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 1 & -\frac{1}{4} \end{pmatrix}\to$$
$$\begin{pmatrix} 1 & 1 & 0 & 0 & \frac{3}{2} \\ 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 1 & -\frac{1}{4} \end{pmatrix}\to\begin{pmatrix} 1 & 0 & 0 & 0 & \frac{5}{4} \\ 0 & 1 & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 1 & -\frac{1}{4} \end{pmatrix}$$

可得:

$$\beta=\frac{5}{4}\alpha_1+\frac{1}{4}\alpha_2-\frac{1}{4}\alpha_3-\frac{1}{4}\alpha_4$$

1.3 线性映射

定义 1.3.1. 线性映射 $\varphi:\mathbb{V}\to\mathbb{W}\Leftrightarrow\varphi(\alpha x+\beta y)=\alpha\varphi(x)+\beta\varphi(y)$

定义 1.3.2. (等价和相似是不同基的变换矩阵)

- $A=T^{-1}BS$ 成立, A, B**等价**.
- $A=T_S(B)=S^{-1}BS$ 成立, A, B**相似**. 称 T 为**相似变换**.

1.3.1 核空间和像空间

1.3.2 线性变换

定义 1.3.3. 线性变换满足 $\varphi:\mathbb{V}\to\mathbb{V}$ 的线性映射.

1.3.2.1 初等变换

定义 1.3.4. 由单位矩阵 I 经过一次初等行 (列) 变换得到的矩阵称为**初等矩阵**.

齐次方程组通解?

1.3.3 常见变换

TODO: Lec 8

1.3.3.1 旋转矩阵

1.3.3.2 反射矩阵

1.3.3.3 信号处理

1.3.3.4 逆

定义 1.3.5. n 阶矩阵 A 可逆 $\Leftrightarrow A = Q_1 Q_2 \dots Q_m$, Q_i 为初等矩阵.

性质 1.3.6. 由1.3.5可得: $(Q_1 Q_2 \dots Q_m)^{-1}(AI) = (IA^{-1})$

定义 1.3.7. **广义逆**: $(A^T A)^{-1} A^T$

1.4 行列式

定义 1.4.1. 一个排列中，如果一个大元素在小元素前，则称这两个数构成一个**逆序**。一个排列中存在的所有逆序的数目称为排列的 **逆序数**: $\tau(j_1, j_2, \dots, j_n)$.

定义 1.4.2. n 阶行列式中，划去元素 a_{ij} 所在第 i 行和第 j 列元素，剩余的元素按原来的次序组成的 $n-1$ 阶行列式称为元素 a_ij 的 **余子式**，记成 M_{ij} 。令 $A_{ij} = (-1)^{i+j} det(M_{ij})$ ，称 A_{ij} 是元素 a_{ij} 的**代数余子式**。

定义 1.4.3. **行列式**:

$$\begin{aligned} det(A) = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} &= \sum_{j_i, \dots, j_n} (-1)^{\tau(j_1, j_2, \dots, j_n)} a_{1j_1} \dots a_{nj_n} \\ &= \sum_{j=1}^{n} a_{ij} A_{ij} (\text{按行展开}) \end{aligned}$$

性质 1.4.4. 行列式即有向体积

性质 1.4.5. 设 $A, B \in R^{n \times n}$

- $det(A) = det(A^T)$

- $det(aA) = a^n det(A)$
- $det(AB) = det(BA) = det(A)det(B)$

定理 1.4.6. 克莱姆法则?(Lec5#12)

性质 1.4.7. 交换矩阵两行 (或两列) 改变矩阵行列式的符号

性质 1.4.8. 行列式关于矩阵的每行 (或每列) 是线性的。

性质 1.4.9. 如果将行列式的某一行 (列)k 倍加到另一行 (列),则行列式的值不变。

定义 1.4.10. 伴随矩阵: $A^*_{ij} = A_{ji}$

定理 1.4.11.

$$\sum_{k=1}^{n} a_{ik}A_{jk} = \begin{cases} |A| & i = j \\ 0 & i \neq j \end{cases}$$

$$AA^* = A^*A = |A|I$$

1.5 迹

定义 1.5.1. 迹: $Tr(A) = \sum_{i=0}^{n} a_{ii}$

性质 1.5.2. 迹的循环不变性

$$Tr(ABC) = Tr(CAB) = Tr(BCA)$$

同理相似矩阵的迹相等:

$$Tr(Q^{-1}AQ) = Tr(QQ^{-1}A) = Tr(A)$$

1.6 二次型

定义 1.6.1. 定义 f 为二次型, A 为二次型矩阵 (对称矩阵).

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) &= a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n \\ &+ \ldots + a_{nn}x + n^2 \\ &= x^TAx = Tr(x^TAx) \end{aligned}$$

定义 1.6.2. A**合同**于 B, $A \simeq B$: $C^TAC = B$. (即不同基下的二次型)

性质 1.6.3. 合同变换:

- 交换 A 的第 i 行和第 j 行,再交换其第 i 列和第 j 列;
- 将 A 的第 i 行乘以非零常数 k,再将其第 i 列乘以 k;

- 将 A 的第 i 行乘以 k 加到第 j 行，再将其第 i 列乘以 k 加到第 j 列。

定义 1.6.4. 设 A 是属于 K 上的非零对称矩阵, 则必存在非奇异矩阵 C，使 C^TAC 的第 1 行第 1 个元素不为零.

定义 1.6.5. 设 A 是属于 K 上的非零对称矩阵, 则必存在非奇异矩阵 C，使 A 合同于 $C^TAC = \begin{pmatrix} a'_{11} & 0_{1\times k} \\ 0_{k\times 1} & \lambda k \times k \end{pmatrix}$.

1.6.1 标准型

定义 1.6.6. 线性替换: 使用 y_i 替换 x_i, $x_{n\times 1} = C_{n\times n}y_{n\times 1}$. 当 $det(C) > 0$ 时, 线性替换非退化.

定义 1.6.7. 每一个二次型都能非退化的线性替换化为平方和, 称为**标准型**. 标准型不唯一.

定理 1.6.8. 每一个二次型矩阵都能通过合同变化为对角矩阵.

$$\begin{pmatrix} A \\ I \end{pmatrix} \xrightarrow{\text{对 A 合同变化}} \begin{pmatrix} D \\ C \end{pmatrix}$$

D 是对角矩阵, C 是非退化的线性替换矩阵.

例如将 $f(x)$ 替换为标准型 $g(y) = y^TDy$, $x = Cy$.

定理 1.6.9. 二次型矩阵的特征值即标准型的系数.

1.6.2 规范型

定义 1.6.10. **规范性**即在**标准型**的基础上把系数变成 1 和 -1. **正惯性指数** p 为 1 的数量. **负惯性指数** $r-p$ 为 -1 的数量.

1.6.3 正定型

定义 1.6.11. 令 $x \neq 0$. **正定二次型**: $x^TAx > 0$, A 为**正定矩阵**. **半正定二次型**: $x^TAx \geqslant 0$, A 为**半正定矩阵**

性质 1.6.12. 正定二次型的等价条件:

1. $x^TAx > 0$ (定义).
2. 标准型/规范型系数全部为正数, 正惯性指数等于 A 的秩.
3. 顺序主子式全大于 0 (对于负定二次型: 奇数阶全小于 0, 偶数阶全大于 0).

1.7 特征值

定义 1.7.1. $Ax = \lambda x \Rightarrow det(\lambda I - A) = 0$

1.8 内积和范数

1.8.1 范数

定义 1.8.1. 范数 $||x||$ 满足:

1. 非负性: $\begin{cases} ||x|| > 0 & x \neq 0 \\ ||x|| = 0 & x = 0 \end{cases}$
2. 齐次性: $||\lambda x|| = |\lambda| ||x||$
3. 三角不等式: $||x + y|| \leqslant ||x|| + ||y||$

定义 1.8.2. l_p **范数**:

$$||x||_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}, 1 \leqslant p < \infty$$

l_1 范数, 1 范数, Manhattan 范数

l_2 范数, 2 范数, 欧几里得范数

定义 1.8.3. $l_0 =$ 非零元素个数, l_0 并不符合范数定义

$l_\infty = \max_j |x_i|$

1.8.2 内积

定义 1.8.4. 内积: $\langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$, 并满足以下条件

1. 非负性: $\langle x, x \rangle \geqslant 0, \langle x, x \rangle = 0$ 当且仅当 $x = 0$
2. 对称性: $\langle x, y \rangle = \langle y, x \rangle$
3. 齐次性: $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$
4. 线性性: $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$

定义 1.8.5. 点积 (标准内积): $\langle x, y \rangle = x^T y$

1.8.3 距离

定义 1.8.6. 距离: $d : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$, 满足:

1. 非负性: $d(x, y) \geqslant 0$, 且 $d(x, y) = 0 \Leftrightarrow x = y$
2. 对称性: $d(x, y) = d(y, x)$
3. 三角不等式: $d(x, z) \leqslant d(x, y) + d(y, z)$

定义 1.8.7. 欧式距离: $d(x, y) = ||x - y|| = \sqrt{\langle x - y, x - y \rangle}$. 其中内积定义为点积.

1.8.4 夹角

定义 1.8.8. 夹角: $\cos\theta = \frac{\langle x,y\rangle}{||x||_2||y||_2}$

定义 1.8.9. 正角: $\langle x, y\rangle = 0$, 如果 $||x|| = ||y|| = 1$, 则为**标准正交**.

1.8.5 矩阵内积和范数

定义 1.8.10. 向量化: $vec(A)$: 将 $m \times n$ 的矩阵拉长为 $mnx1$ 向量

定义 1.8.11. 矩阵内积: $\langle A, B\rangle = \langle vec(A), vec(B)\rangle = vec(A)^T vec(B) = Tr(A^T B)$

定义 1.8.12. 广义矩阵范数: 和向量范数的条件类似.

矩阵范数: 附加相容性条件: $||AB|| \leqslant ||A||||B||$.

定义 1.8.13. 常见矩阵范数

- l_1 **范数**: $||A||_{m_1} = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|$
- l_2 **范数**, **Fiobenius 范数**: $||A||_F = (\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2)^{\frac{1}{2}} = (Tr(A^T A))^{\frac{1}{2}}$
- l_∞ **范数** (广义): $||A||_{m_\infty} = \max |a_{ij}|$

1.8.5.1 算子范数

定义 1.8.14. 相容的向量范数和矩阵范数: $||Ax||_v \leqslant ||A||_M||x||_v$

定义 1.8.15. 由向量范数 $||\cdot||_v$ 诱导出的**算子范数**: $||A|| = \max\{||Ax||_v : ||x||_v = 1\}$

定义 1.8.16. 常见的算子范数, 对于 $A \in \mathbb{R}^{m\times n}$

- 1 范数: $||A||_1 = \max_{1\leqslant j\leqslant n} \sum_{i=1}^m |a_{ij}|$, 列向量的 l_1 范数最大值
- ∞ 范数: $||A||_\infty = \max_{1\leqslant i\leqslant m} \sum_{j=1}^n |a_{ij}|$, 行向量的 l_1 范数最大值
- 2 范数: $||A||_2 = \sqrt{\lambda_{max}(A^T A)}$

1.8.6 例题

例 1.8.1 求证: $\mathbb{R}^n$ 上的一范数 $||\cdot||_1$ 和二范数 $||\cdot||_2$ 是等价的，即存在 c_1, c_2 满足不等式 $c_1||x||_1 \leqslant ||x||_2 \leqslant c_2||x||_1$

解

$$x_1^2 + x_2^2 + \ldots + x_n^2 \leqslant (|x_1| + |x_2| + \ldots + |x_n|)^2$$
$$||x||_2 \leqslant ||x||_1$$
$$\Rightarrow c_2 = 1$$

$$
\begin{aligned}
\frac{x_1^2+x_2^2+\ldots+x_n^2}{n} &\geqslant \frac{(|x_1|+|x_2|+\ldots+|x_n|)^2}{n^2} \\
n(n_1^2+x_2^2+\ldots+x_n^2) &\geqslant (|x_1|+|x_2|+\ldots+|x_n|)^2 \\
\sqrt{2}||x||_2 &\geqslant ||x||_1 \\
&\Rightarrow c_1=\frac{1}{\sqrt{n}}
\end{aligned}
$$

例 1.8.2 求证：通过向量范数 $||\cdot||_1$ 诱导得到的矩阵范数 $||\cdot||_1$ 和向量范数 $||\cdot||_1$ 是相容的，即对任意的 $A\in\mathbb{R}^{n\times n}, x\in\mathbb{R}^n$ 有 $||Ax||_1\leqslant||A||_1||x||_1$

解

$$
\begin{aligned}
||Ax||_1 &= \sum_{i=1}^{n}|a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n| \\
&\leqslant \sum_{i=1}^{n}|a_{i1}x_1|+|a_{i2}x_2|+\ldots+|a_{in}x_n| \\
&= \sum_{j=1}^{n}|x_j|||a_j||_1 \\
&\leqslant \sum_{j=1}^{n}|x_j|\max_{1\leqslant k\leqslant n}||a_k||_1 \\
&= ||x||_1||A||_1
\end{aligned}
$$

例 1.8.3 证明柯西-施瓦茨不等式 $\langle x,y\rangle\leqslant||x||_2||y||_2$

解 当 $y=0$ 显然成立.

当 $y\neq0$, 对于任意 $\lambda\in\mathbb{R}$

$$
\begin{aligned}
0 &\leqslant \langle x-\lambda y, x-\lambda y\rangle \\
&= \langle x-\lambda y, x\rangle-\lambda\langle x-\lambda y, y\rangle \\
&= \langle x,x\rangle-2\lambda\langle x,y\rangle+\lambda^2\langle y,y\rangle
\end{aligned}
$$

取 $\lambda=\langle x,y\rangle\langle y,y\rangle^{-1}$

$$
\begin{aligned}
\langle x,y\rangle^2\langle y,y\rangle^{-1} &\leqslant \langle x,x\rangle \\
\langle x,y\rangle^2 &\leqslant ||x||^2||y||^2
\end{aligned}
$$

2 子空间

2.1 向量子空间基础

2.1.1 向量子空间

定义 2.1.1. 如果 $\mathbb{Y}$ 中的每个向量 x 可唯一地表成 $x = y_1 + y_2(y_1 \in \mathbb{Y}_1, y_2 \in \mathbb{Y}_2)$ 的形式，则称 $\mathbb{Y}$ 为 $\mathbb{Y}_1$ 与 $\mathbb{Y}_2$ 的**直和**。记作 $\mathbb{Y} = \mathbb{Y}_1 + \mathbb{Y}_2$ 或 $\mathbb{Y} = \mathbb{Y}_1 \oplus \mathbb{Y}_2 \Leftrightarrow \mathbb{Y}_1 \cap \mathbb{Y}_2 = 0$

2.1.2 仿射子空间

定义 2.1.2. 仿射子空间即向量子空间偏移向量 x_0

2.1.3 正交和正交补

定义 2.1.3. $\forall v \in \mathbb{S}, \forall w \in \mathbb{T}$, 均有: $v^T w = 0$ 记做 $\mathbb{S}$**垂直**于 $\mathbb{T}$, 或**正交**.

定义 2.1.4. 正交补: $\{w \in \mathbb{R}^n | v^T w = 0, \forall v \in \mathbb{V}\}$

例 2.1.1 求 $span(1,2,0)^T, (0,1,2)^T$ 的正交补空间:

解

$$\begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 2 \end{pmatrix}^T x = 0 \Rightarrow x = k(4,-2,1)^T$$

正交补空间为 $span\{(4,-2,1)^T\}$

2.1.4 四个基本子空间

- 列空间 $Col(A) = span(a_1, a_2, \ldots, a_n)$
- 行空间 $Row(A) = Col(A^T)$
- 零空间 $Null(A) = \{x | Ax = 0\}$
- 左零空间 $Null(A^T)$

定理 2.1.5. 四个基本子空间的正交关系

- $Col(A) \cap Null(A^T) = \{0\}$
- $Col(A^T) \cap Null(A) = \{0\}$
- $Col(A)^\perp = Null(A^T)$
- $Col(A^T)^\perp = Null(A)$

性质 2.1.6. 初等变换有关的性质

1. 一系列初等行变换不改变矩阵的行空间。
2. 一系列初等行变换不改变矩阵的零空间。
3. 一系列初等列变换不改变矩阵的列空间。
4. 一系列初等列变换不改变矩阵的左零空间。

2.2 投影

定理 2.2.1. 正交投影矩阵为 $B(B^TB)^{-1}B^T$, 其中 B 为是一个有序基底

2.3 正交基

定理 2.3.1. Gram-Schmidt 正交化: $b_n = a_n - \sum_{i=1}^{n-1} \frac{\langle b_i, a_n \rangle}{\langle b_i, b_i \rangle} b_i$, $q_n = \frac{b_n}{\langle b_n, b_n \rangle}$

例 2.3.1 利用 Gram-Schmidt 正交化的过程，求下述矩阵列空间的一组正交基:

$$\begin{pmatrix} -10 & 13 & 7 & -11 \\ 2 & 1 & -5 & 3 \\ -6 & 3 & 13 & -3 \\ 16 & -16 & -2 & 5 \\ 2 & 1 & -5 & -7 \end{pmatrix}$$

解

$$\begin{aligned} b_1 &= a_1 = (-10, 2, -6, 16, 2)^T \\ b_2 &= a_2 - \frac{\langle b_1, a_2 \rangle}{\langle b_1, b_1 \rangle} b_1 = (3, 3, -3, 0, 3)^T \\ b_3 &= a_3 - \sum_{i=1}^{2} \frac{\langle b_i, a_3 \rangle}{\langle b_i, b_i \rangle} b_i = (6, 0, 6, 6, 0)^T \\ b_4 &= a_4 - \sum_{i=1}^{3} \frac{\langle b_i, a_4 \rangle}{\langle b_i, b_i \rangle} b_i = (0, 5, 0, 0, -5)^T \end{aligned}$$

单位化:

$$\begin{aligned} e_1 &= (-\frac{1}{2}, \frac{1}{10}, -\frac{3}{10}, \frac{4}{5}, \frac{1}{10})^T \\ e_2 &= (\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, 0, \frac{1}{2})^T \\ e_3 &= (\frac{1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0)^T \\ e_4 &= (0, \frac{1}{\sqrt{2}}, 0, 0, -\frac{1}{\sqrt{2}})^T \end{aligned}$$

3 矩阵分解

3.1 LU 分解

主要用于求解 $Ax = b$.

定义 3.1.1. LU 分解: $A = LU$, L 下三角方阵, U 上三角方阵.

3.2 Gauss 变换

1. 利用 Gauss 变换化矩阵 A 为上三角阶梯型矩阵 $U = QA$. 应确保 k 阶顺序主子式始终不为 0 ($\Rightarrow$ 每一步的主元均不为 0).

   Gauss 变换: $R_i \pm = kR_j, i > j$, 或可描述为 $Q = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & k_3 & 1 & \\ & k_4 & & 1 \end{pmatrix}$

2. 对 I 执行与上一步的初等变换的逆变换, 下三角矩阵 $L = Q^{-1}I$

3. $A = LU$

习题1. 判定矩阵 $C = \begin{bmatrix} 3 & 2 & -1 \\ -1 & 0 & 0 \\ -1 & 3 & 0 \end{bmatrix}$ 和 $B = \begin{bmatrix} 0 & 2 & -1 \\ -1 & 4 & -1 \\ 1 & 3 & -5 \end{bmatrix}$ 能否进行 LU 分解, 为什么? 如果能分解, 试分解之.

解 矩阵 C 的顺序主子式均不为 0, 可以进行 LU 分解.

$$\begin{bmatrix} 3 & 2 & -1 \\ -1 & 0 & 0 \\ -1 & 3 & 0 \end{bmatrix} \xrightarrow[R_3+\frac{1}{3}R_1]{R_2+\frac{1}{3}R_1} \begin{bmatrix} 3 & 2 & -1 \\ 0 & \frac{2}{3} & -\frac{1}{3} \\ 0 & \frac{11}{3} & -\frac{1}{3} \end{bmatrix} \xrightarrow{R_3-\frac{11}{2}R_1} \begin{bmatrix} 3 & 2 & -1 \\ 0 & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & \frac{3}{2} \end{bmatrix} \rightarrow U$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_3+\frac{11}{2}R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{11}{2} & 1 \end{bmatrix} \xrightarrow[R_3-\frac{1}{3}R_1]{R_2-\frac{1}{3}R_1} \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{3} & 1 & 0 \\ -\frac{1}{3} & \frac{11}{2} & 1 \end{bmatrix} \rightarrow L$$

矩阵 B 的 1 阶主子式均为 0, 因此 B 不可以进行 LU 分解.

习题2. 求矩阵 $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ 的 LU 分解.

解

$$\begin{bmatrix}2&1&1\\1&2&1\\1&1&0\end{bmatrix}\xrightarrow{R_3-R_2}\begin{bmatrix}2&1&1\\1&2&1\\0&-1&-1\end{bmatrix}\xrightarrow{R_2-\frac{1}{2}R_1}\begin{bmatrix}2&1&1\\0&\frac{3}{2}&\frac{1}{2}\\0&-1&-1\end{bmatrix}\xrightarrow{R_3+\frac{2}{3}R_2}\begin{bmatrix}2&1&1\\0&\frac{3}{2}&\frac{1}{2}\\0&0&-\frac{2}{3}\end{bmatrix}\to U$$

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\xrightarrow{R_3-\frac{2}{3}R_2}\begin{bmatrix}1&0&0\\0&1&0\\0&-\frac{2}{3}&1\end{bmatrix}\xrightarrow{R_2+\frac{1}{2}R_1}\begin{bmatrix}1&0&0\\\frac{1}{2}&1&0\\0&-\frac{2}{3}&1\end{bmatrix}\xrightarrow{R_3+R_2}\begin{bmatrix}1&0&0\\\frac{1}{2}&1&0\\\frac{1}{2}&\frac{1}{3}&1\end{bmatrix}\to L$$

3.2.1 选主元的 LU 分解

如果得到 U 不存在主元为 0 的上三角阵, 可以使用行交换 P 进行选主元 (重排). 即 $U' = PU = PQA \Rightarrow P(PQ)^{-1}PU = PA$.

3.3 QR 分解

定义 3.3.1. QR 分解 (正交三角分解): $A = Q\begin{pmatrix}R\\O\end{pmatrix}$. $A \in \mathbb{R}^{m\times n}(m \geqslant n)$, 正交矩阵$Q \in \mathbb{R}^{m\times m}$, Q 上三角矩阵$R \in \mathbb{R}^{n\times n}$.

QR 分解主要用于解决最小二乘问题, 矩阵特征值计算.

3.3.1 基于 Gram-Schmidt 正交化

使用Gram-Schmidt 对满秩矩阵 A 正交化得到正交基 $Q = (q_1, q_2, \ldots, q_m)$, 再用 Q 表示 $A = QR$.

3.3.2 Householder

TODO: After lec 8

3.3.3 Givens

TODO: After lec 8

3.4 谱分解/特征分解

3.4.1 特征分解

定义 3.4.1. 特征分解: $A = Q\Lambda Q^{-1}$, Λ 为特征值对角矩阵, Q 为特征向量

例 3.4.1 矩阵的幂 A^k

解 $A^k = (Q\Lambda Q^{-1})^k$

3.4.2 谱分解

定义 3.4.2. 对称矩阵 A, 如果 A 可以本分解为 $A = Q\Lambda Q^T$. 其中 Q 为单位化特征向量矩阵, Λ 是特征值对角矩阵. 称为**谱分解**.

3.4.3 Cholesky 分解

定义 3.4.3. 对于对称正定矩阵 A, $A = GG^T$, G 为下三角矩阵. 成为**Cholesky 分解**若把对角元素提取出来, 即 $A = LDL^T$ 分解, 称为**不带平方根的 Cholesky 分解**

习题3. 求对称正定矩阵

$$A = \begin{bmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{bmatrix}$$

的不带平方根的 Cholesky 分解.

解

$$D_1 = A_{11} = 5$$

$$L_{21} = \frac{1}{D_1} A_{21} = \frac{1}{5} \times 2 = \frac{2}{5},\ L_{31} = \frac{1}{D_1} A_{31} = \frac{1}{5} \times (-4) = -\frac{4}{5}$$

$$D_2 = A_{22} - L_{21}^2 D_1 = 1 - \left(\frac{2}{5}\right)^2 \times 5 = \frac{1}{5}$$

$$L_{32} = \frac{1}{D_2} (A_{32} - L_{31} L_{21} D_1) = 5 \left(-2 + \frac{4}{5} \times \frac{2}{5} \times 5\right) = -2$$

$$D_3 = A_{33} - \sum_{k=1}^{2} L_{3k}^2 D_k = 5 - \left(\frac{4}{5}\right)^2 \times 5 - (2)^2 \times \frac{1}{5} = 1$$

$$A = LDL^T = \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{5} & 1 & 0 \\ -\frac{4}{5} & -2 & 1 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{5} & 1 & 0 \\ -\frac{4}{5} & -2 & 1 \end{bmatrix}^T$$

3.5 奇异值分解 (SVD)

定义 3.5.1. $A \in \mathbb{R}^{m \times n}$ **SVD 分解**: $A = U\Sigma V^T$,

- $U \in \mathbb{R}^{m \times m}$ 是 AA^T 的标准化特征向量
- $\Sigma \in \mathbb{R}^{m \times n}$ 是 $A^T A$ 的特征值的根号值
- $V \in \mathbb{R}^{n \times n}$ 是 $A^T A$ 的标准化特征向量

3.6 矩阵分解应用

3.6.1 线性方程组

$$Ax = b$$

- Cholesky 法: $A = GG^T$, $Gb = d$, $G^T x = b$

3.6.2 最小二乘问题

定义 3.6.1. 最小二乘问题: $\min_x ||Ax - b||_2$

- 正则化方法: 求解 $A^T Ax = A^T b$
- QR 法: $A = QR$, $Rx = Q^T b$

4 向量和矩阵微分

定理 4.0.1. 常见导数

$$
\begin{aligned}
(x^\mu)' &= \mu x^{\mu-1} \\
(a^x)' &= a^x \ln a \\
(\log_a x)' &= \frac{1}{x \ln a} \\
(f \circ g)' &= f'(g(x))g'(x)
\end{aligned}
$$

定理 4.0.2. 行列式对矩阵偏导: $\frac{\partial |X|}{\partial X} = X^*$

如果 $|X| > 0$, $\frac{\partial |X|}{\partial X} = |X| X^{-T}$

5 优化问题

5.1 牛顿法

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$$

5.2 例题

习题4. 计算 $f(x)$ 的共轭函数, 以及共轭函数的定义域.

- $f(x) = -\log x$
- $f(x) = e^x$

解

- $f(x) = -\log x$

$$\begin{aligned} f^*(y) &= \sup_{x>0}(yx + \log x) \\ &= \begin{cases} -1 - \log(-y) & y < 0 \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

- $f(x) = e^x$

$$\begin{aligned} f^*(y) &= \sup_{x\in\mathbb{R}}(yx - e^x) \\ &= \begin{cases} \infty & y < 0 \\ 0 & y = 0 \\ y\ln(y) - y & y > 0 \end{cases} \end{aligned}$$

习题5. 求解线性规划

$$\begin{aligned} &\min\ e^T x \\ &\text{s.t. } Gx \leqslant h \\ &\qquad Ax = b \end{aligned}$$

的对偶函数，给出对偶问题。

解 Lagrange 函数:

$$
\begin{aligned}
L(x,\lambda,\nu) &= e^T x + \lambda^T(Gx-h) + \nu^T(Ax-b) \\
&= -\lambda^T h - \nu^T b + (e^T + \lambda^T G + \nu^T A)x
\end{aligned}
$$

对应的对偶函数为:

$$
g(\lambda,\nu) = \inf_x L(x,\lambda,\nu) = \begin{cases} -\lambda^T h - \nu^T b & e^T + \lambda^T G + \nu^T A = 0 \\ -\inf & \text{otherwise} \end{cases}
$$

对应的对偶问题为:

$$
\begin{aligned}
\max\ & -\lambda^T h - \nu^T b \\
s.t.\ & e^T + \lambda^T G + \nu^T A = 0 \\
& \lambda, \nu \geqslant 0
\end{aligned}
$$

习题6. 求优化问题

$$
\arg\min_{x_1,x_2,x_3} x_1 x_2 x_3
$$

当 x_1, x_2, x_3 满足 $x_1^2 + x_2^2 + x_3^2 = 1$ 的解。

解 Lagrange 函数:

$$
L(x_1,x_2,x_3,\lambda) = x_1 x_2 x_3 + \lambda(x_1^2 + x_2^2 + x_3^2 - 1)
$$

$$
\nabla L = 0 \Rightarrow \begin{cases} \frac{\partial L}{\partial x_1} = x_2 x_3 + 2\lambda x_1 = 0 \\ \frac{\partial L}{\partial x_2} = x_1 x_3 + 2\lambda x_2 = 0 \\ \frac{\partial L}{\partial x_3} = x_1 x_2 + 2\lambda x_3 = 0 \\ \frac{\partial L}{\partial \lambda} = (x_1^2 + x_2^2 + x_3^2 - 1) = 0 \end{cases} \Rightarrow \begin{cases} |x_1| = |x_2| = |x_3| = \frac{1}{\sqrt{3}} \\ \lambda = \frac{1}{2\sqrt{3}} \end{cases}
$$

因此 x_1, x_2, x_3 的解为:

$$
\begin{cases} x_1 = \frac{1}{\sqrt{3}} \\ x_2 = \frac{1}{\sqrt{3}} \\ x_3 = -\frac{1}{\sqrt{3}} \end{cases}, \begin{cases} x_1 = \frac{1}{\sqrt{3}} \\ x_2 = -\frac{1}{\sqrt{3}} \\ x_3 = -\frac{1}{\sqrt{3}} \end{cases}
$$

习题7. 已知矩阵 $A \in \mathbb{R}^{p\times q}$, $B \in \mathbb{R}^{p\times r}$, $rank(A) = min(p,q)$, 未知矩阵 $X \in \mathbb{R}^{q\times r}$, 求以下优化问题。

若 $p < q$, 求 Frobenius 范数最小的矩阵 X, 使得 $AX = B$, 即求解优化问题:

$$\begin{aligned} \min\ f(X) &= \frac{1}{2}||X||_F^2 \\ \text{s.t. } AX &= B \end{aligned}$$

解 Lagrange 函数:

$$\begin{aligned} L(X,\Lambda) &= \frac{1}{2}||X||_F^2 - \Lambda^T(AX - B) \\ &= Tr(\frac{1}{2}X^TX) - Tr(\Lambda^T(AX-B)) \end{aligned}$$

$$\nabla L = 0 \Rightarrow \begin{cases} \frac{\partial L}{\partial X} = X - A^T\Lambda = 0 \\ \frac{\partial L}{\partial \Lambda} = (AX-B)^T = 0 \end{cases}$$

因为 $p < q \Rightarrow rank(A) = p$, 所以 AA^T 可逆。

$$\begin{aligned} AX &= B \\ AX &= (AA^T)(AA^T)^{-1}B \\ X &= A^T(AA^T)^{-1}B \end{aligned}$$

习题8. 梯度下降法是最常用的优化方法之一。考虑优化问题

$$\min f(x) = x_1^2 + x_2^2 + 2x_3^2$$

证明: 在点 $x_0 = (x_1, x_2, x_3)$ 处沿梯度方向迭代的最佳步长为:

$$\lambda = \frac{x_1^2 + x_2^2 + 4x_3^2}{2x_1^2 + 2x_2^2 + 16x_3^2}$$

证 即求解 $\arg\min_\lambda f(x_0 - \lambda^T\nabla f(x_0))$

$$\begin{aligned} f(x_0 - \lambda^T\nabla f(x_0)) &= f(x_0 - \lambda^T(2x_1, 2x_2, 4x_2)^T) \\ &= (1-\lambda_1)^2x_1^2 + (1-\lambda_2)^2x_2^2 + 2(1-4\lambda)^2x_3^2 \\ \frac{\partial f(x_0 - \lambda^T\nabla f(x_0))}{\partial\lambda} &= (8x_1^2 + 8x_2^2 + 64x_3^2)\lambda - (4x_1^2 + 4x_2^2 + 16x_3^2) = 0 \\ \lambda &= \frac{x_1^2 + x_2^2 + 4x_3^2}{2x_1^2 + 2x_2^2 + 16x_3^2} \end{aligned}$$

索引