

1 前置知识

矩阵求逆

$\det(A)A^{-1} = A^* = ((-1)^{i+j}a_{ji})$

Schmidt 正交化

$y_1 = x_1$, 令 $y_2 = x_2 + \lambda_{21}y_1, \langle y_2, y_1 \rangle = 0$ 求解 y_2 , $y_3 = x_3 + \lambda_{31}y_1 + \lambda_{32}y_2, \langle y_3, y_1 \rangle = \langle y_3, y_2 \rangle = 0$ 求解 y_3 , 以此类推。

Vandemonde 行列式

$$\det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix} = \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

矩阵乘法的一些特性

- $(\beta^T a_{ij} \alpha) = \beta^T A \alpha$
- $\sum_{k_1=1}^n \sum_{k_2=1}^n c_{ij} \beta_{k_1} \alpha_{k_2} = \beta^T C \alpha$

2 空间与变换

线性空间

对 $V, +, \cdot$, 在

- $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$
- $\alpha + \beta = \beta + \alpha$
- $\exists \theta \forall \alpha, \alpha + \theta = \alpha$
- $\forall \alpha \exists \gamma, \alpha + \gamma = \theta$
- $(\alpha \cdot k_1) \cdot k_2 = \alpha \cdot (k_1 k_2)$
- $\alpha \cdot (k_1 + k_2) = \alpha \cdot k_1 + \alpha \cdot k_2$
- $(\alpha_1 + \alpha_2) \cdot k = \alpha_1 \cdot k + \alpha_2 \cdot k$
- $\alpha \cdot 1 = \alpha$

中满足前四条的 $(V, +)$ 为交换群, 满足全部的 $\langle V, +, \cdot \rangle$ 为线性空间。

维数、基底与坐标

$\forall \beta \in V, \exists !k_1 \dots k_n \in F, \beta = \sum_{i=1}^n \alpha_i k_i$ 表示

n 维空间 V 下 β 在基底 $(\alpha_1 \dots \alpha_n)$ 下的坐标为 $(k_1 \dots k_n)^T$ 。如 $V = \mathbb{R}^{2 \times 3}, F = \mathbb{R}$ 的一组基底是 将 2×3 零矩阵逐个元素代入 1 得到的六个矩阵, 按照 $\beta = (\alpha_1 \dots \alpha_n) (k_1 \dots k_n)^T$ 计算时可先将矩阵拉直为列向量。

换基底

$(\alpha_1 \dots \alpha_n) P = (\beta_1 \dots \beta_n)$ 中 P 为基 (α_i) 到 (β_i) 的过渡矩阵。于是 $\beta = (\alpha_i) X = (\beta_i) Y = (\alpha_i) P X$ 导出 $X = P Y$ 。

内积空间

对线性空间 $\langle V, +, \cdot \rangle$ 和运算 \langle , \rangle 满足

- $\langle \beta, \alpha \rangle = \overline{\langle \alpha, \beta \rangle}$
 - $\langle \alpha, \alpha \rangle \geq 0$ 当且仅当 $\alpha = \theta$ 时取等
 - $\langle \alpha, \beta_1 k_1 + \beta_2 k_2 \rangle = \langle \alpha, \beta_1 \rangle k_1 + \langle \alpha, \beta_2 \rangle k_2$
- 时 \langle , \rangle 为内积, $\langle V, +, \cdot, \langle , \rangle \rangle$ 为内积空间。注意 $\langle \beta_1 k_1 + \beta_2 k_2, \alpha \rangle = \langle \beta_1, \alpha \rangle \overline{k_1} + \langle \beta_2, \alpha \rangle \overline{k_2}$ 。

内积示例

复列向量内积 $\langle x, y \rangle = y^H x$, 方阵的 Frobenius 内积 $\langle A, B \rangle = tr(A^H B) = \sum_i \sum_j \overline{a_{ij}} b_{ij}$ 。

Gram 矩阵（度量矩阵）

$G = (\langle \alpha_i, \alpha_j \rangle)$ 是基 (α_i) 的度量矩阵。满足 Hermite 对称 $G^H = G$, Hermite 半正定 $x^H G x \geq 0$ 。使用 P 将基 (α_i) 变换为 (α'_i) 后, 后者的度量矩阵 $G' = P^H G P$ 。计算 $\langle \alpha'_i, \alpha'_j \rangle$ 即得。

子空间

若 W 是 F 上线性空间 V 的非空子集且

- $\forall \alpha + \beta \in W, \alpha + \beta \in W$
- $\forall \alpha \in W, k \in F, \alpha k \in W$

则称 W 是 V 中的一个线性子空间（线性可通过定义验证）, 记作 $W \leq V$ 最小的子空间是 θ 。

生成子空间

任取 V 中非空子集 S , 其所有线性组合组成生成子空间 $W = span S$, 这是 V 中包含 S 中全部元素的最小子空间。 V 是它的基的生成子空间。

子空间的和

$W_1 + W_2 = \{\alpha_1 + \alpha_2 | \alpha_1 \in W_1, \alpha_2 \in W_2\} \subseteq V$, 子空间的并是和的子集但不是线性空间因此不是子空间。

维数公式

$dim(W_1) + dim(W_2) = dim(W_1 + W_2) + dim(W_1 \cap W_2)$, 类似容斥原理。

直和分解

若 $W_1 + W_2 = V, W_1 \cap W_2 = \{\theta\}$ 则 $V = W_1 \oplus W_2$ (直和)。等价于 $\theta = \alpha_1 \in W_1 + \alpha_2 \in W_2$ 当且仅当 $\alpha_1 = \alpha_2 = \theta$, 证明思路是假设有两种分解后反证。 W_1 和 W_2 的基的并是 V 的一组基, 因为 W_1 的基和 W_2 的基线性无关。若给定 V 的一个子空间, 可通过求基的方式构造出与其直和得到 V 的另一个子空间。

线性映射

线性映射 $\varphi : V_1 \rightarrow V_2$ 满足

- $\varphi(\alpha + \beta) = \varphi(\alpha) + \varphi(\beta)$
- $\varphi(\alpha \cdot k) = \varphi(\alpha) \cdot k$

或写作 $\varphi((\alpha_i)^T (k_i)) = \varphi(\alpha_1 \dots \alpha_n) (k_i)$ 。 $V_1 = V_2 = V$ 时称线性自映射 (线性算子)。

线性映射的矩阵表示

$\varphi : V_1 \rightarrow V_2$, 对于 V_1 的基 α 和 V_2 的基 β 有 $\varphi(\alpha_1 \dots \alpha_n) = (T(\alpha_i)) = (\alpha_1 \dots \alpha_n) A = (\beta_1 \dots \beta_n)$

矩阵转换

令 $\varphi : V_1 \rightarrow V_2$, V_1 下基 α 到 α' 的过渡矩阵 P , V_2 下基 β 到 β' 的过渡矩阵 Q , φ 由 α 变换为 β 时矩阵表示为 A , 由 α' 变换为 β' 时矩阵表示为 B 。计算可知 $A = QB, B = Q^{-1} A P$ 。对于线性算子, 取 $\alpha = \beta, Q = P$ 即得到相似。

核

$kernel \varphi = \{\alpha \in V_1 | \varphi(\alpha) = \theta_{V_2}\} \leq V_1$ 要证子空间只需验证线性。 $kernel \varphi = \{\theta_{V_1}\}$ 对应单射 (不会有两个元素被映射到同一个元素, 反证得到)。

像

$image \varphi = \{\varphi(\alpha) | \alpha \in V_1\} \leq V_2$ 要证子空间只需验证线性。 $image \varphi = V_2$ 对应满射。

正交补

$W \leq V$, W 的正交补 $W^\perp = \{\beta \in V | \forall \alpha \in W, \langle \alpha, \beta \rangle = 0\} \leq V$ 。满足 $W + W^\perp = V$, 因为将 W 的基扩充为 V 的基时新增基的生成子空间可证明与 W^\perp 互相包含。 $W \cap W^\perp = \{\theta\}$, 故 $W \oplus W^\perp = V$, 由维数公式可知它们维数之和为 $dim V$ 。

保内积算子和酉阵

$\varphi : V \rightarrow V$ 满足 $\langle \varphi(\alpha), \varphi(\beta) \rangle = \langle \alpha, \beta \rangle$ 则称为保内积算子。取标准正交基代入定义计算, 可知其变换矩阵 $U^T U = I$ 为酉矩阵。计算两组标准正交基的 Gram 矩阵后可得到它们的过渡矩阵是酉矩阵。酉矩阵的各列向量组成一组标准正交基。

3 矩阵分解

Schur

方阵 $A = U T U^H$, 当 $A^H A = A A^H$ (正规) 时有 $A = U \Lambda U^H$ 。标准型对角线元素为 A 全部特征值。证明基本思路为取单位特征向量构造酉阵, 此后计算并归纳。正规矩阵的 Schur 求法: 求出所有特征向量并标准正交化得到 U 。

谱分解

竖切正规矩阵 Schur 分解的 U 阵得到 $A = (\xi_i) \Lambda (\overline{\xi_i})^T = \sum \xi_i \lambda_i \overline{\xi_i}$ 。

Jordan

方阵 $A = P J P^{-1}$, 算法为:

- 求 A 的特征多项式, 得到特征值
- 求 A 的最小多项式 $\prod (\lambda_i - A)^{\beta_i}$, β_i 为 λ_i 对应最大 Jordan 块大小
- 对于每个特征值, 求 $rank((\lambda_i I - A)^k)$, $k = 1 \dots \beta_i$ 。计算 $n_k = nullity(\lambda_i I - A)^k$, n_1 为属于 λ_i 的 Jordan 块总数, $n_{k+1} - n_k$ 为尺寸不小于 k 的 Jordan 块个数。原因是 $(\lambda I - J)^i$ 与 $(\lambda I - A)^i$ 相似, 秩相等, 而对于 $(\lambda I - J)$ 每自乘一次, 当前幂次阶以上的 Jordan 块就会多一个全零行。
- 求变换矩阵 P

SVD

$A^{m \times n} = U_m (\Sigma_{\begin{smallmatrix} r & r \\ O & O \end{smallmatrix}}) V_n^H, \Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}$, 奇异值 $\sigma_i = \sqrt{\lambda_i} \neq 0$, λ_i 是 $A^H A$ 的非零特征值, 编号根据值降序排列。证明时先构造 V , 由于 $A^H A$ 正规, Schur 得 $V^H A^H A V = diag(\Sigma^2, O_{n-r})$ 。竖切 $V = (V_1^{n \times r}, V_2^{n \times (n-r)})$, 代入 Schur 计算有 V_1, V_2 与 A, Σ 的关系。取 $U_1^{m \times r} = A V_1 \Sigma^{-1}$, 满足 $U_1^H U_1 = I_r$, 因

此 U_1 各列向量标准正交。将 U_1 增补为 $U^{m \times m}$ 即为所需的 U , 可代入 SVD 验证。

极分解

方阵 $A = P \tilde{U} = U \tilde{Q}$, 其中 P, Q Hermite 半正定。从 SVD 开始配凑中间的 I 为酉阵与其共轭转置乘积得到。一阶时为 $x = r e^{i\theta}$ 。

满秩分解

LR 分解

$$A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q = P \begin{pmatrix} I_r & \\ & O \end{pmatrix} (I \ O) Q = L R$$

4 矩阵分解相关

Sylvester 降幂公式

对于 $A \in \mathbb{C}^{m \times n}, B \in \mathbb{C}^{n \times m}, m \geq n$, 有 $\det(\lambda I_m - AB) = \lambda^{m-n} \det(\lambda I_n - BA)$ 。结合满秩分解可以用于降阶计算高阶特征多项式。证明需要注意到

$$\begin{pmatrix} I & B \\ O & I \end{pmatrix} \begin{pmatrix} O & O \\ A & AB \end{pmatrix} = \begin{pmatrix} BA & BAB \\ A & AB \end{pmatrix} \\ = \begin{pmatrix} BA & O \\ A & O \end{pmatrix} \begin{pmatrix} I & B \\ O & I \end{pmatrix}$$

再计算由此导出相似矩阵的特征多项式。

Carley-Hamilton 定理

对于 n 阶方阵 A , 其特征多项式 $p(\lambda) = \det(\lambda I - A)$ 是它的一个零化多项式。使用 Schur 或 Jordan 将 A 化为上三角矩阵后计算可证。

最小多项式

次数最低的首一零化多项式。是所有零化多项式的因子, 这一点可使用多项式整除求余证明。

广义特征向量

相似对角化

方阵可相似对角化等价于其特征向量线性无关, 因为 $P^{-1} A P = \Lambda, A P = P \Lambda$, 根据特征值定义可证明; 亦等价于最小多项式无重根, 因为最大 Jordan 块大小为 1。

典型正规矩阵

奇异向量

SVD 变形有 $A V = U \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix}$, 竖切 $U = (\eta_1 \dots \eta_m), V = (\xi_1 \dots \xi_n)$ 后展开计算, 于是 $A \xi_i = \eta_i \sigma_i, 1 \leq i \leq r$, 称 ξ_i 为右奇异向量, η_i 为左奇异向量。此外 $A \xi_i = O, r \leq i \leq n$ 。

基于 SVD 的矩阵压缩

对 SVD 分解式竖切 U , 横切 V^H 有 $A = (\eta_1 \dots \eta_m) \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} (\xi_1^H \dots \xi_n^H)^T$, 展开得 $A = \sum_{i=1}^r \eta_i \sigma_i \xi_i^H$, 故大奇异值项占主要成分。

5 范数理论

线性空间 v 上的范数

设 V 为 F 上的线性空间, 若映射 $\| \cdot \| : V \rightarrow \mathbb{R}$ 满足

- $\| v \| \geq 0, \| v \| = 0 \iff v = \theta$
- $\| v \lambda \| = \| v \| | \lambda |$
- $\| v_1 + v_2 \| \leq \| v_1 \| + \| v_2 \|$

则称 $\| \cdot \|$ 为 V 上的一个 F 范数, $\{V, \| \cdot \| \}$ 为赋范线性空间。 $d(v_1, v_2) = \| v_1 - v_2 \|$ 称作 v_1, v_2 的距离, $\{V, d\}$ 为度量线性空间。 $d(v_1, v_3) = \| v_1 - v_2 + v_2 - v_3 \| \leq d(v_1, v_2) + d(v_2, v_3)$ 。

内积诱导范数

$\| v \| = \sqrt{\langle v, v \rangle}$ 。对于 $\forall \alpha, \beta \in V$ 有 $\langle \alpha + \beta, \alpha + \beta \rangle = \langle \alpha, \alpha \rangle + \langle \alpha, \beta \rangle + \langle \beta, \alpha \rangle + \langle \beta, \beta \rangle$, $\langle \alpha - \beta, \alpha - \beta \rangle = \langle \alpha, \alpha \rangle - \langle \alpha, \beta \rangle - \langle \beta, \alpha \rangle + \langle \beta, \beta \rangle$, 两式相加得到 $\| \alpha + \beta \|^2 + \| \alpha - \beta \|^2 = 2 \| \alpha \|^2 + 2 \| \beta \|^2$ 。

勾股定理和 Cauchy 不等式

对于内积诱导范数, 展开计算可证明

- $\alpha \perp \beta \iff \| \alpha + \beta \| = \| \alpha \| + \| \beta \|$ 。
- $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = | \langle \alpha, \beta \rangle | \leq \| \alpha \|^2 \| \beta \|^2$

Cauchy 不等式在 $\alpha \neq \theta$ 时显然成立; 其他情况构造 $\alpha \perp \gamma = \beta - \alpha k$ 可证明, 且当且仅当 α 与 β 线性相关时等号成立。 α 和 β 的夹角满足 $\langle \alpha, \beta \rangle = \| \alpha \| \| \beta \| \cos \lambda$ 。

列向量上的 p-范数

对于列向量 v , $\| v \|_p = (\sum |x_i|^p)^{\frac{1}{p}}$:

- $p = 1 : \sum |x_i|$
- $p = 2 : \sqrt{v^H v} = \sqrt{\sum x_i^2}$
- $p = \infty : \lim_{p \rightarrow \infty} \| v \|_p = \max \{ |x_i| \}$

注意 2-范数保酉变换。

线性映射构造范数

线性单射 $\varphi : V_1 \rightarrow V_2$, $\| \cdot \|_{V_2}$ 为 V_2 上的范数, 那么 V_1 上定义 $\| v \|_\varphi = \| \varphi v \|_{V_2}$ 为 V_1 上的范数。例如, 对矩阵使用拉直变换为列向量后可使用列向量的 2-范数定义矩阵的 Frobenius 范数。

范数等价

有限维空间中序列是否为 Cauchy 列与范数选取无关。 $\exists c_1, c_2 > 0, \forall v \in V \| v \|_\alpha \leq c_1 \| v \|_\beta, \| v \|_\beta \leq c_2 \| v \|_\alpha$ 。有限维空间中 Cauchy 列一定收敛, 因为可以归结为列向量空间上的 ∞ -范数, 也就可以对每个矩阵元素取极限得到整个矩阵的极限。

线性映射的范数

赋范线性空间 $(V_2, \| \cdot \|_\beta), (V_1, \| \cdot \|_\alpha)$ 有线性映射 $\varphi : V_2 \rightarrow V_1$, 定义 $\| \varphi(v) \|_{\alpha, \beta} = \sup_{v \neq \theta} \frac{\| \varphi(v) \|_\alpha}{\| v \|_\beta} = \max_{\| v \|_\beta = 1} \| \varphi(v) \|_\alpha$, 相当于最大缩放比。 $\| \varphi(v) \|_{\alpha, \beta}$ 是 $(\varphi : V_2 \rightarrow V_1)$ 上的范数。称 $\| \varphi(v) \|_\alpha \leq \| \varphi \|_{\alpha, \beta} \| v \|_\beta$ 为相容性条件。次乘性: 若 $\psi : (V_3, \| \cdot \|_\gamma) \rightarrow (V_2, \| \cdot \|_\beta), \varphi : (V_2, \| \cdot \|_\beta) \rightarrow (V_1, \| \cdot \|_\alpha)$, 则 $\| \psi \circ \varphi \|_{\alpha, \beta} \leq \| \varphi \|_{\alpha, \beta} \cdot \| \psi \|_{\beta, \gamma}$

常用矩阵范数

1. 列向量 1-范数的诱导范数（列和范数）

$$\| A \|^{m \times n}_{1,1} = \max_{\| x \|_1 = 1} \| A x \|_1 = \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{ij}|$$

2. 列向量 2-范数的诱导范数（谱范数）
 $\|A^{m \times n}\|_{2,2} = \max \sigma$ ，SVD 可证。
3. 列向量 ∞ -范数的诱导范数（行和范数）
 $\|A^{m \times n}\|_{\infty,\infty} = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|$
4. m_1 范数，矩阵拉直为列向量后取 1-范数
 $\|A^{m \times n}\|_{m_1} = \sum \sum |a_{ij}|$
5. Frobenius 范数，矩阵拉直为列向量后取 2-范数
 $\|A^{m \times n}\|_F = \sqrt{\langle A, A \rangle_F} = \sqrt{\text{tr}(A^H A)} = \sqrt{\sum \sigma_i^2}$
6. m_∞ 范数，矩阵拉直为列向量后取 ∞ -范数再乘矩阵尺寸
 $\|A^{m \times n}\|_{m_\infty} = \max\{m, n\} \max_{i,j} |a_{ij}|$

谱半径估计

谱半径 $\rho(A)$ 是 A 全部特征值绝对值中的最大值。 $\rho(A) \leq \|A\|$ ，计算 $Ax = x\lambda$ 范数可知。 $\rho(A) \geq \|A\| + \epsilon$ ，证明思路为对 A 用 Jordan，令 $Q = \text{diag}(1, \epsilon, \epsilon^2 \dots)$ ， A 的范数 $\|Q^{-1}P^{-1}APQ\|_{\infty,\infty}$ 满足条件。因此 $\rho(A) \leq 1 \iff \exists \|\cdot\|_*, \|A\|_* \leq 1$ 。

矩阵幂收敛条件

$$\begin{aligned} \|A\| \leq 1 &\Rightarrow k \rightarrow \infty, 0 \leq \|A^k - 0\| = \\ \|A^k\| \leq \|A\|^k = 0 &\Rightarrow \lim_{k \rightarrow \infty} A^k = 0 \end{aligned}$$