

MIT 18.06

Lyshmi.Y

2022 年 9 月 26 日

木易

目录

I	$AX=B$ 和四个子空间	2
1	方程组的几何解释	3
2	矩阵消元	7
3	乘法和逆矩阵	10
4	矩阵 A 的 LU 分解	13
5	转置、置换、向量空间	15
6	列空间和零空间	17
7	求解 $Ax = 0$; 主变量和特解	19
8	可解性和解的结构	22
9	线性相关性、基和维数	25
10	四个基本子空间	27
11	矩阵空间、秩 1 矩阵和小世界图	28
12	图和网络	30
II	投影、行列式和特征值	32
13	正交向量和子空间	33

目录	2
14 子空间投影	35
15 投影矩阵和最小二乘	37
16 正交矩阵和 <i>Schmidt</i> 正交化	39
17 行列式及其性质	41
18 行列式公式和代数余子式	42
19 Cramer 法则和逆矩阵	43
20 特殊矩阵总结	45
21 特征值和特征向量	46
22 对角化和矩阵幂	48
23 微分方程 $\exp(At)$	50
24 马尔可夫矩阵和傅里叶级数	52
III 正定矩阵及其应用	54
25 对称矩阵及其正定性	55
26 复数矩阵和快速傅里叶变化	57
27 二次型、标准型、规范型和正定二次型	59
28 正定矩阵和最小值	61
29 相似矩阵和若尔当型	62
30 奇异值分解	63
31 线性变换及对应矩阵	65
32 基变换和图像压缩	68

目录	3
33 左右逆和伪逆	70

Part I

$AX=B$ 和四个子空间

Chapter 1

方程组的几何解释

我们首先来看一个方程组, 例如:

$$\begin{cases} 2x - y = 0 \\ -x + 2y = 4 \end{cases} \quad (1.1)$$

我们立马有方程组系数矩阵的概念, 这个矩阵包含方程组未知数的系数, 那么我们可以将方程组 (1.1) 改写为下面的矩阵形式 (1.2)

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \quad (1.2)$$

我们令 $\mathbf{A} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, 我们称 \mathbf{A} 为方程组 (1.1) 的系数矩阵, 令 $\mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$, 我们得到所有线性方程组的通式 $\mathbf{AX} = \mathbf{b}$.

我们首先使用行图像来说明线性方程组 (1.1), 在平面直角坐标系 xoy 中, 直线 $l_1: 2x - y = 0$ 表示方程 $2x - y = 0$ 的解的集合; 直线 $l_2: -x + 2y = 0 = 0$ 表示方程 $-x + 2y = 0$ 的解的集合; 方程组的解是这两个方程组的解集的交集, 说明两条直线 l_1 l_2 的交点 $(1, 2)$, 即 $\begin{cases} x = 1 \\ y = 2 \end{cases}$ 为方程组 (1.1) 的解.

接下来我们从另外的一个角度来看待方程组 (1.1), 我们发现方程组的

矩阵 (1.2) 可以写成如下的形式:

$$x \begin{bmatrix} 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \quad (1.3)$$

我们令向量 $\mathbf{n}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, 向量 $\mathbf{n}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, 向量 $\mathbf{b} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$; 我们只需要找到合适的 x, y , 使得向量 \mathbf{b} 由向量 $\mathbf{n}_1, \mathbf{n}_2$ 的线性组合得到.

明显的, 我们发现当 $x = 1, y = 2$ 时, 向量 $\mathbf{n}_1, \mathbf{n}_2$ 的线性组合得到向量 \mathbf{b} . 显然对于 $\forall x \in R, y \in R$; 我们得到一个向量空间, 即二维平面上的所有向量.

我们来看另一个三元一次方程组的例子:

$$\begin{cases} 2x - y = 0 \\ -x + 2y - z = -1 \\ -3y + 4z = 4 \end{cases} \quad (1.4)$$

我们写出这个方程组的矩阵形式, 如下 (1.5) 所示. $\mathbf{A}\mathbf{X} = \mathbf{b}$, 其中系数矩阵 $\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 4 \end{bmatrix}$, $\mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$.

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} \quad (1.5)$$

我们依然从行向量和列向量两个角度来看待这个方程组, 首先是行向量的角度, 我们可以理解为是在空间直角坐标系 xy 中, 每一个方程代表一个平面, 这个平面上所有的点都是方程的解, 三个平面的交点就是方程组的解集.

接下来是列向量的角度, 我们发现方程组的矩阵形式 (1.5) 还可以写为如下的形式.

$$x \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix} + z \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} \quad (1.6)$$

我们令向量 $\mathbf{n}_1 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$, 向量 $\mathbf{n}_2 = \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix}$, 向量 $\mathbf{n}_3 = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$, 向量 $\mathbf{b} = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$; 我们只需要找到合适的 x, y, z , 使得向量 \mathbf{b} 由向量 $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ 的线性组合得到.

我们也许可以很轻易的发现 $\begin{cases} x = 0 \\ y = 0 \\ z = 1 \end{cases}$ 是原方程组的一个解, 当然这个

方程组还有很多个解. 我们发现列向量 $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ 的线性组合可以覆盖掉整个三维空间. 换句话说就是, 对于 $\forall x \in R, y \in R, z \in R$; 我们得到一个向量空间, 也就是三维空间内所有向量的集合.

因此, 我们提出一个问题, 就是对于方程组 $\mathbf{A}\mathbf{X} = \mathbf{b}$, $\forall \mathbf{b} \in R^n$, 方程组是否有解? 这个说法等价于是否能找到列向量的线性组合满足向量 \mathbf{b} ? 对于方程组 (1.4), 这个答案是肯定的, 那么对于任意一个方程组呢?

对于任意一个方程组, 系数矩阵的列向量是否有线性组合满足右侧的向量 \mathbf{b} 决定方程组是否有解, 当这些列向量线性无关时, 对于任意的向量 \mathbf{b} 方程组总会有解, 反之则不一定.

矩阵和矩阵、矩阵和向量相乘

首先是矩阵和向量相乘

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix} = b_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{bmatrix} + b_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{bmatrix} + \dots + b_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{bmatrix}$$

表示的是矩阵列向量的线性组合, 我们可以类比于方程组的列向量解读.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1q} \\ b_{21} & b_{22} & \dots & b_{2q} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nq} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1q} \\ c_{21} & c_{22} & \dots & c_{2q} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mq} \end{bmatrix}$$

其中

$$c_{\mathbf{i}\mathbf{j}} = \sum_{k=1}^n a_{\mathbf{i}k} b_{k\mathbf{j}}$$

同样可以理解为矩阵和向量相乘得到, 矩阵 \mathbf{C} 的每一列都是矩阵 \mathbf{B} 的对应列和矩阵 \mathbf{A} 相乘得到, 同样是列向量的线性组合.

Chapter 2

矩阵消元

用消元法求解方程式组

消元法是所有计算机程序求解方程组使用的方法, 下面我们先来看一个例子.

$$\begin{cases} x + 2y + z = 2 \\ 3x + 8y + z = 12 \\ 4y + z = 2 \end{cases} \quad (2.1)$$

这个方程等价的矩阵形式 $AX = b$ 如下 (2.2) 所示, $A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix}, X =$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}, b = \begin{bmatrix} 2 \\ 12 \\ 2 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 12 \\ 2 \end{bmatrix} \quad (2.2)$$

为了方便, 我们使用矩阵来表示消元, 进行矩阵的初等行变换来消元.

$$\begin{bmatrix} \mathbf{1} & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} \mathbf{1} & 2 & 1 \\ 0 & \mathbf{2} & -2 \\ 0 & 4 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} \mathbf{1} & 2 & 1 \\ 0 & \mathbf{2} & -2 \\ 0 & 0 & \mathbf{5} \end{bmatrix}$$

我们将最后得到的上三角矩阵记作 \mathbf{U} , 在这个例子中, $U = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix}$.

我们消元的目的就是 $A \rightarrow U$. 在这里, 我们求得主元的个数等于方程组中方程个数; 假如主元个数小于方程组的个数时, 需要用到行交换.

我们下一步要做的是回代, 将增广矩阵进行初等行变化.

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 0 & 5 & -10 \end{bmatrix}$$

消元后最终得到的方程组如下 (2.3) 所示:

$$\begin{cases} x + 2y + z = 2 \\ 2y - 2z = 6 \\ 5z = -10 \end{cases} \quad (2.3)$$

显然我们可以得到原方程组的解为 $\begin{cases} x = 11 \\ y = -2 \\ z = -5 \end{cases}$, 但是我们关心的并不

只是方程组的解, 我们想要知道矩阵初等行变换是怎样发生的, 在第一章中我们知道矩阵右乘一个列向量可以理解为矩阵列向量的线性组合, 那么矩阵左乘一个行向量得到一个行向量, 这个行向量是矩阵行向量的线性组合. 所以在方程组 $AX = b$ 中, 系数矩阵 A 的每一次初等行变换都可以理解为系数矩阵 A 左乘一个矩阵.

所以本例子中的矩阵变化也可以写成以下的形式:

$$\begin{aligned} \text{First: } \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} &= \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix} \\ \text{Second: } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix} &= \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix} \end{aligned}$$

或者你也可以一步到位:

$$\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 6 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix}$$

我们可以知道对于 $\forall A \in R^n, \exists E \in R^n$, 使得 $EA = U$. (E 为初等矩阵)

置换矩阵

将单位矩阵 **I** 中要置换的两行置换位置得到置换矩阵 **P**.

注意点

1. $AB \neq BA$
2. 矩阵左乘代表行变换, 右乘代表列变换.

Chapter 3

乘法和逆矩阵

矩阵乘法

假设有矩阵 A 、 B 、 C , $C = AB$, 矩阵 A 是 $m \times n$ 型矩阵, 矩阵 B 是 $n \times q$ 型矩阵, 矩阵 C 是 $m \times q$ 型矩阵. 两个矩阵相乘必须满足第一个矩阵的列数等于第二个矩阵的行数, 相乘得到的矩阵行数和第一个矩阵相同, 列数和第二个矩阵相同.

简而言之: $col(A) = row(B), row(C) = row(A), col(C) = col(B)$.

且对于矩阵 C 中的任意一个元素 c_{ij} , 我们有 $c_{ij} = row(A_i) * col(B_j)$.

接下来我们从行向量和列向量的线性组合来考虑矩阵的乘法, 首先是列向量的线性组合. 矩阵 C 的每一列都可以看作是矩阵 A 和矩阵 B 对应列相乘的结果, 那么矩阵 A 和矩阵 B 的列向量相乘就等价于矩阵 A 各列的线性组合, 也就是说矩阵 C 的每一列均是矩阵 A 的列向量的线性组合.

同理, 我们分析矩阵 C 的行向量, 矩阵 C 的每行向量我们可以看作矩阵 B 左乘一矩阵 A 对应行向量的结果, 那么矩阵 C 的行向量等价于矩阵 B 行向量的线性组合.

我们最后还有一种方法, 用矩阵 A 的每一列和矩阵 B 的每一行相乘, 再将这些结果相加得到矩阵 C .

$$C = \sum_{i=1}^m row(A_i) * col(B_i)$$

这个思路其实揭示出分块矩阵相乘的原理.

我们有矩阵 $A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$, 矩阵 $B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$, 矩阵 $C = \begin{bmatrix} A_1B_1 + A_2B_3 & A_1B_2 + A_2B_4 \\ A_3B_1 + A_4B_3 & A_3B_2 + A_4B_4 \end{bmatrix}$. 其中 A_i, B_i 均为矩阵.

有了这个, 我们反过来看矩阵乘法的最后一种计算方式, 不难看出: 令

$A = \begin{bmatrix} A_1 & A_2 & \dots & A_n \end{bmatrix}$, 令 $B = \begin{bmatrix} B_1 \\ B_2 \\ \dots \\ B_n \end{bmatrix}$, 其中 A_i 是 $m \times 1$ 的列向量, B_i 是 $1 \times q$ 的行向量. 那么 $C = \begin{bmatrix} \sum_{i=1}^n A_i B_i \end{bmatrix}$.

逆矩阵

注意点:

1. 逆矩阵、可逆等说法都是建立在方阵的基础上.
2. 逆矩阵若存在则唯一.

定义: 存在方阵 A 和方阵 A^{-1} , 满足 $A^{-1}A = AA^{-1} = I$, 我们称 A^{-1} 为 A 的逆矩阵, A 为可逆矩阵.

思考: 如何判断一个方阵是否可逆呢?

方阵是否可逆取决于是否能找到非零列向量 X , 使得方程 $AX = \mathbf{0}$.

A 不可逆 $\Leftrightarrow \exists X \in R^n / \mathbf{0}$, 方程 $AX = \mathbf{0}$ 有解.

证明. 假设方阵 A 可逆, $\exists A^{-1}$, 满足 $A^{-1}A = AA^{-1} = I$.

$\because \exists X \in R^n$, 使得 $AX = \mathbf{0}$

$\therefore A^{-1}AX = A^{-1}\mathbf{0}$

又 $\because A^{-1}A = I$

\therefore 左式 $= X \neq \mathbf{0}$; 右式 $= \mathbf{0}$

明显左式 \neq 右式, 假设不成立, 原命题得证. \square

求方阵逆矩阵: 高斯-若尔当法

假设方阵 A 存在逆矩阵, 我们令 $C = \begin{bmatrix} A & I \end{bmatrix}$. 我们对 C 进行矩阵的初等行变换, 使得 $C' = \begin{bmatrix} I & E \end{bmatrix}$, 寻找到一系列初等行变换 $E = E_1E_2\dots E_n$, 使得 $EC = C'$.

根据分块矩阵的乘法原理, 我们有 $EA = I, EI = E$, 根据逆矩阵存在的唯一性, 我们知道 $E = A^{-1}$, 那么矩阵 $C' = \begin{bmatrix} I & A^{-1} \end{bmatrix}$ 后半部分即是我们所求的 A^{-1} .

重点

1. 对于任意一个矩阵 A , 我们都可以进行高斯消元, $\exists E = E_n \dots E_2 E_1$, 使得 $EA = U$. 则有 $A = E^{-1}U$, 也就是下一章中提到的 $A = LU$ 分解. 这其实是从总体上来看高斯消元, 将高斯消元的多个步骤合成了一步.

2. 假设矩阵 A 可逆, 矩阵 B 可逆, 我们有 $(AB)^{-1} = B^{-1}A^{-1}$.

证明. $\because A$ 可逆, B 可逆

$$\therefore A^{-1}A = AA^{-1} = I, BB^{-1} = B^{-1}B = I$$

$$\text{又 } \because (AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1}$$

$$\therefore (AB)(B^{-1}A^{-1}) = AA^{-1} = I$$

即 $AB(B^{-1}A^{-1}) = I$, 根据逆矩阵存在的唯一性, $B^{-1}A^{-1}$ 是 AB 的逆矩阵. \square

3. 假设 A 可逆, A 的转置 A^T 逆矩阵 $(A^T)^{-1} = (A^{-1})^T$

证明. $\because A$ 可逆

$$\therefore A^{-1}A = I$$

$$\text{等式两边同时取转置, 得到 } (A^{-1}A)^T = I^T$$

$$\therefore (A^{-1}A)^T = I \Rightarrow A^T(A^{-1})^T = I$$

根据逆矩阵存在的唯一性, 我们知道 A^T 的逆矩阵是 $(A^{-1})^T$. \square

$$4. (AB)^T = B^T A^T$$

证明. 假设 A, B 都是 $n \times n$ 型矩阵, $C = (AB)^T$, 对于矩阵 C 中任意一个元素 $c_{ij}, c_{ij} = \text{row}(A_j) * \text{col}(B_i) \Rightarrow c_{ij} = \sum_{k=1}^n a_{jk} b_{ki}$.

令 $C' = B^T A^T$, 对于矩阵 C' 中任意一个元素 $c'_{ij}, c'_{ij} = \text{row}(B_i^T) * \text{col}(A_j^T) \Rightarrow c'_{ij} = \sum_{k=1}^n b_{ik}^T a_{kj}^T$

$$\because \forall k \in n, a_{kj}^T = a_{jk}, b_{ik}^T = b_{ki}$$

对于矩阵 C 和矩阵 C' 中任意元素, 都有 $c_{ij} = c'_{ij}$.

$$\therefore (AB)^T = B^T A^T \quad \square$$

Chapter 4

矩阵 A 的 LU 分解

矩阵 A 的 LU 分解

对于矩阵 A , 假设在高斯消元的过程中没有进行交换行的操作, 我们进行一系列的初等行变换得到矩阵 U .

在每一步消元过程中, 相当于矩阵 A 左乘一个矩阵 E_i , 存在矩阵 $E = E_n \dots E_2 E_1$, 使得 $EA = U$.

对于 $EA = U$ 等式两边同时左乘 E^{-1} , 我们得到了 $A = E^{-1}U$. 这就是矩阵 A 的 LU 分解. (L 代表下三角形矩阵, U 代表上三角形矩阵)

接下来我们来看几个例子:

$$1. A = \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix}$$

对 A 进行初等行变换: $\begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$, 其中 $E_{21} =$

$\begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix}$, $U = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$, 即 $E_{21}A = U$. 我们想要得到 $A = LU$ 的

形式, 我们必须求出 E_{21}^{-1} . 在这个例子中, 明显得出 $E_{21} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}$. 即

$$L = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}.$$

$$A = LU \Leftrightarrow \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$$

上面式子最后三项可以理解为 $A = LDU$, 其中 D 是对角矩阵.

2. 假设 A 是一个 3×3 矩阵. 我们有 $E_{32}E_{31}E_{21}A = U, A = E_{32}^{-1}E_{31}^{-1}E_{21}^{-1}U$.
我们可以得出 $L = E_{32}^{-1}E_{31}^{-1}E_{21}^{-1}$.

消元的次数分析

假设一次乘法和一次减法算作一次操作, 矩阵 A 为 $n \times n$ 的非零矩阵, A 进行初等行变换时没有交换行, 那么矩阵 A 消元次数 $N_1 = \sum_{i=2}^n i(i-1)$.

$$N_1 = \sum_{i=1}^n i^2 - \sum_{i=1}^n i \Leftrightarrow N_1 = \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2} = \frac{n^3 - n}{3}$$

对于方程 $AX = b$, 右侧向量 b 在消元过程中的操作数 $N_2 = \sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$.

总的来说, 解决方程 $AX = b$, 我们需要进行高斯消元的操作数为 $N = N_1 + N_2 = \frac{2n^3 + 3n^2 - 5n}{6}$.

接下来我们考虑矩阵 A 在进行消元过程中发生行交换的情况, 对此我们首先要研究置换矩阵 P . 置换矩阵 P_{ij} 指的是 $P_{ij}A$ 得到矩阵 A 第 ij 行交换的矩阵.

对于 n 阶矩阵形式而言, 一共有 A_n^n 种不一样的置换矩阵 P_{ij}

置换矩阵满足的性质: $P_{ij}^{-1} = P_{ij}^T = P_{ji}$

所以我们改进高斯消元的矩阵形式 $A = LU \Rightarrow PA = LU$, 我们在下一章会继续讲解这些问题.

Chapter 5

转置、置换、向量空间

置换矩阵 *Permutations* 行交换

我们在求解线性方程组 $AX = b$ 时, 通常会遇到主元位置为 0 的情况, 这时候要进行行互换才能继续消元. 高斯消元的矩阵形式变为: $PA = LU$, 且对于任意可逆矩阵 A , 都有此种形式

转置 *Transpose*

对于矩阵 A , A 的转置 A^T , 假设 A 是 $m \times n$ 阶矩阵, 对于 $\forall i \in m \quad \forall j \in n$, $A_{ij} = A_{ji}^T$.

对称矩阵

存在矩阵 A , 满足 $A^T = A$, 我们称 A 为对称矩阵.

对于任意矩阵 A , A 的转置矩阵 A^T , $B = AA^T$, $C = A^T A$, 则 B 、 C 均为对称矩阵.

证明. 设矩阵 A 是 $m \times n$ 型矩阵, B 是 $m \times m$ 型矩阵, C 是 $n \times n$ 型矩阵.

对于 $\forall i \in m \quad \forall j \in m$, $b_{ij} = \text{row}(A_i) * \text{col}(A_j^T)$; $b_{ji} = \text{row}(A_j) * \text{col}(A_i^T)$
 $\because \text{row}(A_i) = \text{col}(A_i^T)^T \quad \text{row}(A_j) = \text{col}(A_j^T)^T$
 $\therefore b_{ij} = \text{col}(A_i^T)^T * \text{col}(A_j^T)$; $b_{ji} = \text{col}(A_j^T)^T * \text{col}(A_i^T)$
 $\therefore b_{ij} = b_{ji}$, B 是对称矩阵; 同理 C 也是对称矩阵. □

向量空间 *VectorSpace*

定义:

1.0 向量必须包含在向量空间中.

2. 对于任意一个向量空间 R^n , 对于 $\alpha \in R^n, \beta \in R^n, \alpha + \beta \in R^n, \alpha - \beta \in R^n$; 对于 $\forall x \in R, x * \alpha \in R^n$. (换句话说, 就是向量空间对加减法、数乘封闭).

我们并不关心整个 R^n , 我们关心的是其中的子空间. 举个简单的例子. 我们来看 R^2 向量空间中的一个子空间, 一条过原点的直线、原点、 R^2 空间本身.

存在矩阵 A 是 $m * n$ 阶矩阵, 矩阵 A 列向量组成的向量空间 (列向量的线性组合) 是 R^n 的向量子空间, 称作矩阵的列空间 $C(A)$.

Chapter 6

列空间和零空间

子空间

假设 R^3 的两个子空间: P 过原点的平面和 L 过原点的直线, $P \cup L$ 和 $P \cap L$ 是否还是向量子空间?

$P \cup L$ 不是子空间, $P \cap L$ 是子空间

证明. 不妨任意取 $\alpha \in P \cap L$, $\beta \in P \cap L$

对于任意 $x \in R$, $y \in R$, 由于 P 、 L 均为子空间

因此 $x\alpha + y\beta \in P$, $x\alpha + y\beta \in L$ 所以我们得出 $x\alpha + y\beta \in P \cap L$. 即 $P \cap L$ 是向量子空间. \square

列空间 *Column Space of A*

问题: 对于方程 $AX = b$ 是否对于任意的 b 都有解?

当 $b \in C(A)$ 时, 方程组一定有解, 右侧向量 b 在矩阵 A 的列向量线性组合的列空间内时, 等价于原方程有解.

零空间 *Null Space of A* 满足方程 $AX = 0$ 的解 X 的集合是矩阵 A 的零空间, 我们称之为 $N(A)$. 如何证明这个集合内的向量构成了一个子向量空间?

证明. 我们不妨将 $AX = 0$ 的解集称作 N .

任取 $\alpha \in N$, $\beta \in N$, $\forall x \in R$, $y \in R$, $\gamma = x\alpha + y\beta$.

由于 $A\alpha = 0$; $A\beta = 0$

$A\gamma = A(x\alpha + y\beta) = x(A\alpha) + y(A\beta) = 0$.

$N(A)$ 是子空间, 我们称之为零空间. \square

我们用具体的例子来看一下列空间和零空间.

$$\text{矩阵 } A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \text{ 方程 } AX = b.$$

显然我们可以得到 $X = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ 是方程的解, 但是方程远不止这一个解,

这是原方程的一个特解.

方程 $AX = 0$ 的一个特解为 $X = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, 零空间为过原点的一条直线 $\in R^3$. 列空间为过原点的一个平面 $\in R^4$.

Chapter 7

求解 $Ax = 0$; 主变量和特解

关于求解 $AX = b$, 我们先来看一个例子: $A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} X =$

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. 我们需要注意随着消元, 不变的是方程组的解, 零空间 $N(A)$ 不会发生变化, 但是列空间 $C(A)$ 随着消元会发生改变. 我们来写出高斯消元的过程.

$$\begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 2 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$$U = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

矩阵 U 中非零行的数被称为主元数, 也称作矩阵的秩. 我们用数学语言描述: $\text{rank}(A) = \text{rank}(U) = \text{Number}(\text{Pivot})$.

在本例子中, 当 $b = 0$ 时, 消元对于方程组的解没有影响, 我们来求出零空间首先找出主变量 (每一个非零行第一个非零元素所在的列), 主变量也被称为主元, 主元所在的列被称为主列, 非主列被称为自由列.

我们将消元后得到的矩阵 U 回代到方程组中得到消元后的方程组:

$$\begin{cases} x_1 + 2x_2 + 2x_3 + 2x_4 = 0 \\ 2x_3 + 4x_4 = 0 \end{cases}$$

尽管这样我们还是无法一眼看出结果, 其实我们可以对 U 进一步化简得到简化行阶梯矩阵 $rref(A)$, 我们称之为 R . 满足主元列除了主元的位置是 1 其余位置全为 0. 我们适当调整主元列的位置, 使得主元列全部在前, 我们得到一个矩阵 R , 它的形式形如 $\begin{bmatrix} I & F \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$. I 是一个 $r \times r$ 的矩阵, 代表有 r 个主元, F 是一个 $(n-r) \times (n-r)$ 的矩阵, 代表有 $(n-r)$ 个自由元. 我们有如下式 (7.1) 的关系:

$$AX = 0 \Leftrightarrow UX = 0 \Leftrightarrow RX = 0 \rightarrow \begin{bmatrix} I & F \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} X_{Pivot} \\ X_{Free} \end{bmatrix} = \mathbf{0} \quad (7.1)$$

我们可以得出 $X_{Pivot} = -F, X_{Free} = I$, 零空间是 $\begin{bmatrix} -F \\ I \end{bmatrix}$ 的线性组合得到的向量空间.

我们还是回到这个例子, 我们可以得到化简后的简化行阶梯矩阵 $R = \begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2 & -2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. 有两个主元, 两个自由元, 我们可以得到最简化的方程组如下:

$$\begin{cases} x_1 + 2x_2 - 2x_4 = 0 \\ x_3 + 2x_4 = 0 \end{cases}$$

我们得到方程组的两个特解 $X_1 = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$ $X_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$. A 的零空间

$$N(A) = cX_1 + dX_2, \quad c \in R, \quad d \in R.$$

我们来看另一个矩阵 $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 2 & 6 & 8 \\ 2 & 8 & 10 \end{bmatrix}$, 经过消元得到 $U = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $R =$

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \text{ 说明只有一个自由变量, 两个主元. 方程组 } AX = 0 \text{ 的特解有}$$

一个 $X_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$. A 的零空间 $N(A) = cX_1, \quad c \in R$.

注意点

1. 这一章我们主要讨论了方程组 $AX = 0$ 的解, 即 A 的零空间, 我们知道如何去用高斯消元的方法去求解零空间的结构.
2. 关于方程 $AX = b$ 的解我们下一章进行讨论, 我们已经知道要想要方程组有解必须要满足的条件是 $b \in C(A)$.

Chapter 8

可解性和解的结构

可解性 *Solvability*

方程组 $AX = b$ 有解等价于 $b \in C(A)$, 还有另一种表述: A 中行向量的线性组合得到零行, 右侧向量各行同样的线性组合必然得到 0.

如何求解方程 $AX = b$?

1. 我们首先要找到一个特解 X_p : 令所有的自由变量全部取 0
2. 我们要求出零空间中的向量 $X_n \in N(A)$

$$AX_p = b; \quad AX_n = 0 \quad \rightarrow \quad A(X_p + X_n) = b.$$

所以我们得到方程组的全部解为 $X_p + X_n$

我们还是来看一个例子:

$$A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \text{ 方程组的增广矩阵 } A_1 = \begin{bmatrix} 1 & 2 & 2 & 2 & b_1 \\ 2 & 4 & 6 & 8 & b_2 \\ 3 & 6 & 8 & 10 & b_3 \end{bmatrix}$$

对 A_1 进行消元, 我们得到以下的过程:

$$\begin{bmatrix} 1 & 2 & 2 & 2 & b_1 \\ 2 & 4 & 6 & 8 & b_2 \\ 3 & 6 & 8 & 10 & b_3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 2 & 2 & b_1 \\ 0 & 0 & 2 & 4 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 - b_1 - b_2 \end{bmatrix}$$

当 $b_3 - b_1 - b_2 = 0$ 时, 方程组有解. 我们不妨令 $b = \begin{bmatrix} 1 \\ 5 \\ 6 \end{bmatrix}$. 消元结果为

$\begin{bmatrix} 1 & 2 & 2 & 2 & 1 \\ 0 & 0 & 2 & 4 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$, 令自由变量 $x_2 = 0, x_4 = 0$, 我们得到方程组的一个特

解 $X_p = \begin{bmatrix} -2 \\ 0 \\ \frac{3}{2} \\ 0 \end{bmatrix}$. 在上一章中我们得到了零空间 $N(A) = cX_1 + dX_2, \quad c \in$

$R, \quad d \in R. X_1 = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}, X_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$

所以我们得到原方程组的解为 $X = X_p + cX_1 + dX_2; c \in R, \quad d \in R$

$X = \begin{bmatrix} -2 \\ 0 \\ \frac{3}{2} \\ 0 \end{bmatrix} + c \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix} + d \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}; c \in R, d \in R$

行满秩和列满秩

假设矩阵 A 是 $m \times n$ 型的矩阵, $\text{rank}(A) = r$, 满足 $r \leq \min(m, n)$. 我们来看一下满秩的情况.

1. 列满秩 $r = n < m$

此时方程中没有自由变量, $N(A) = \mathbf{0}$, 如果方程组有解, 那么有且只有一个特解 X_p 或者方程组没有解. 消元得到的矩阵 $R = \begin{bmatrix} I \\ \mathbf{0} \end{bmatrix}.$

2. 行满秩 $r = m < n$

此时方程组有 m 个主元, 消元过程中不会出现零行, 在方程 $AX = b$ 中, 对于任意的右侧向量 b , 方程组都有解. 消元得到矩阵 $R = \begin{bmatrix} I & F \end{bmatrix}$. 方程组有无穷多个解.

3. 满秩矩阵 $r = m = n$

此时矩阵为可逆矩阵, 消元结果 $R = I$, 零空间 $N(A) = \mathbf{0}$, 方程组有且只有 1 个解.

4. 非满秩矩阵 $r < m, r < n$

此时方程组有 r 个主元, $(n-r)$ 个自由元, 方程组中会出现零行, 消元得

到的矩阵 $R = \begin{bmatrix} I & F \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$, 方程组解的个数为 $\mathbf{0}$ 个或者无穷多个

Chapter 9

线性相关性、基和维数

线性相关性

设有向量 $x_1, x_2 \dots x_n$, 假设存在不全为 0 的 $\alpha_1, \alpha_2 \dots \alpha_n$ 使得 $\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n = \mathbf{0}$, 我们称向量 $x_1, x_2 \dots x_n$ 线性相关, 反之则称这些向量线性无关.

1. 零向量和任意向量都线性相关.
2. 向量之间线性相关, 我们一定可以找到一组不全为 0 的线性组合, 使其为零向量.
3. 矩阵 A 列向量线性相关 \rightarrow 矩阵 A 零空间内存在非零向量; 矩阵 A 列向量线性无关 \rightarrow 矩阵 A 零空间内只有零向量.

基和维数

向量 $V_1, V_2 \dots V_n$ 生成一个向量空间 \Leftrightarrow 向量空间包含 $V_1, V_2 \dots V_n$ 的所有线性组合. 我们不妨设这个生成的向量空间为 S , S 是最小的包含向量组所有线性组合的向量空间. 任意给定一个向量组, 我们都能得到一个由这个向量组线性组合生成的向量空间 S .

基

向量空间的基: 指的是一个向量组中的所有向量, 这些向量被称为向量空间的基. 他们满足以下一些性质:

1. 这个向量组中的向量都是线性无关的.
2. 他们能够生成整个向量空间.
3. 基可以有很多组, 但是基向量的向量个数相同, R^n 向量空间的基的个数为 n 个, 我们称这个个数为向量空间的维数.

维数

$$\dim(N(A)) = n - \text{rank}(A), \quad \dim(R^n) = n$$

列空间的维数等于矩阵的秩, 等价于矩阵的主元列个数. $\dim(C(A)) = \text{rank}(A) = N(\text{pivotcolumn})$

如果我们知道某个向量空间的维数, 在这个向量空间中任意 $\dim(S)$ 个线性无关的向量都可以组成这个向量空间的一组基.

Chapter 10

四个基本子空间

零空间、列空间、行空间和转置零空间

$$N(A), \quad C(A), \quad C(A^T), \quad N(A^T)$$

我们也可以理解为矩阵 A 的零空间和列空间、矩阵 A^T 的零空间和列空间, 矩阵 A 是 $m \times n$ 型矩阵, 这四个基本子空间之间有什么关系呢?

$$1. N(A) \in R^n, \quad C(A^T) \in R^n; \quad C(A) \in R^m, \quad N(A^T).$$

$$2. \dim(C(A)) = \text{rank}(A) = r, \quad \dim(N(A)) = n - \text{rank}(A) = n - r; \\ \dim(C(A^T)) = \text{rank}(A) = r, \quad \dim(N(A^T)) = m - \text{rank}(A) = m - r.$$

3. 向量空间中基的求解

(1). 由于矩阵的行变换不会改变矩阵的行空间, 所以矩阵的行空间的一组基可以是消元后的主元行.

(2). 矩阵的行变换会改变矩阵的列空间, 但是我们可以将矩阵的主元列对应的几个列向量作为列空间的一组基.

(3). 零空间的基个数是自由变量的个数 $(n - r)$, 分别取 I_{n-r} 中的每一列得到.

(4). 转置零空间 $N(A^T)$

$$A^T y = \mathbf{0}, \text{ 两边同时取转置得到 } y^T A = \mathbf{0}(\text{Row})$$

我们将矩阵 A 化简得到矩阵 R 的过程中进行初等行变化, 因此我们可以找到一个可逆矩阵 E , 使得 $EA = R$, R 中存在的零行对应 E 中的线性组合即为左零空间 $N(A^T)$ 的一组基.

Chapter 11

矩阵空间、秩 1 矩阵和小世界图

矩阵空间

由于矩阵可以进行加法和数乘运算, 我们可以将矩阵当作向量, 矩阵线性组合得到的空间成为矩阵空间. 同理矩阵空间也存在基和维数, 矩阵空间也有子空间. n 维矩阵空间的表示方法 $R^{n \times n}$.

我们以矩阵空间 M 为例, M 是 3×3 矩阵构成的矩阵空间. 所有的 3×3 的上三角型矩阵构成的矩阵空间 M_1 ; 所有 3×3 下三角型矩阵构成的矩阵空间 M_2 ; 所有的 3×3 对称矩阵构成的矩阵空间 M_3 ; 我们可以看出 M_1, M_2, M_3 都是 M 的子空间.

$$M \text{ 的一组基为 } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \dots \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$\dim(M) = 9$, M_1, M_2, M_3 的基都是 M 基的子集, M_1, M_2, M_3 的维数取决于一组基中基的个数. 我们接下来来看矩阵空间的交集和矩阵空间之和.

我们之前证明过任意两个子空间的交集仍然是一个子空间, 两个子空间的并集不一定是一个子空间. 我们以 M_1, M_2, M_3 为例, $M_1 \cap M_2$ 是对角矩阵, $M_2 \cap M_3$ 也是对角矩阵. 我们将 $M_1 + M_2$ 记作两个空间的和, 也可以写作 $\text{sum}(M_1, M_2)$.

秩 1 矩阵

存在矩阵 $A, \text{rank}(A) = 1$, 我们称矩阵 A 为秩 1 矩阵. 此时矩阵 $A =$

CR , 其中 C 为一个列向量, R 为一个行向量; 同时也分别是矩阵 A 列空间和行空间的一组基.

我们来看一个例子: S 是所有向量 $V = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$ 的集合, 其中 $v_1 + v_2 +$

$v_3 + v_4 = 0$, S 是一个子空间吗? 如果是, 求出这个空间的一组基?

S 是一个子空间, 令 $A = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, S 就是方程 $AX = 0$ 的解集, S 是矩

阵 A 的零空间. 我们通过消元后发现矩阵 A 有 3 个自由变量, S 的一组基为

$$\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}.$$

Chapter 12

图和网络

图

图是连接节点和边的集合: $Graph = \{nodes, edges\}$

我们用矩阵来表示图, 假设图有 n 个节点, m 条边, 我们用 $m * n$ 的矩阵 A 来表示这个图, 矩阵的每一列表示每一个节点, 每一行表示每一条边, 边的起点用 -1 表示, 边的终点用 1 表示. 我们可以得到一个表示图的矩阵. $dim(N(A)) = n - r$.

比如以下的这个图 G , 它的关联矩阵形式 $A = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$.

方程 $AX = b$ 的结果得到的是格格节点之间的差距. 在实际问题中, 比如在电场中各个点之间的电势差. $AX = 0$ 得到的方程组如下:

$$\begin{cases} x_2 - x_1 = 0 \\ x_4 - x_3 = 0 \\ x_3 - x_1 = 0 \\ x_4 - x_1 = 0 \\ x_4 - x_3 = 0 \end{cases}$$

关于矩阵 A 的零空间, 我们可以很容易得到 $dim(N(A)) = 3, N(A) =$

$$c \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad c \in R.$$

接下来我们来看以下 $A^T y = 0$. 我们仍然将其放在电场问题中, $y =$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}$$

是每条边上的电流, $A^T y = 0$ 就是基尔霍夫电流定律 *KCL*. 我们同

样可以得到一个方程组:

$$\begin{cases} -y_1 - y_3 - y_4 = 0 \\ y_1 - y_2 = 0 \\ y_2 + y_3 - y_5 = 0 \\ y_4 + y_5 = 0 \end{cases}$$

我们知道 A^T 的零空间的维数 $\dim(N(A^T)) = m-r$, 在本例中 $\dim(N(A^T)) =$

$$2, \text{ 我们可以找到 } A^T \text{ 的零空间的一组基 } \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

$y = Ce, e = AX, A^T y = f$, 应用数学中的三个基本方程, y 表示电流, e 表示电势差, C 表示常量, 其实就是电阻, X 表示各点的电势, f 表示电路中有电源的情况, 我们能得到以下的电路综合方程:

$$A^T C A X = f \quad (12.1)$$

树 Tree: 没有回路的图被称为树.

欧拉公式

$nodes, edges, loops$ 三者之间的关系:

$$N(nodes) + N(loops) - N(edges) = 1 \quad (12.2)$$

Part II

投影、行列式和特征值

Chapter 13

正交向量和子空间

正交 *Orthogonal*

正交向量: 列向量 $x, y, x^T y = 0$, 我们就说向量 x, y 正交, 或者说两向量正交 (垂直). 零向量和任意向量都正交.

向量长度: x 的长度 $l^2 = x^T * x \Leftrightarrow l = \sqrt{x_1^2 + x_2^2 \dots + x_n^2}$.

判断两向量是否正交: 向量 x, y 正交 $x^T y = 0 \Leftrightarrow y^T x = 0 \Leftrightarrow |x+y|^2 = |x|^2 + |y|^2$

证明. $\because |x+y| = \sqrt{(x+y)^T(x+y)}$

$$\therefore |x+y|^2 = x^T x + x^T y + y^T x + y^T y$$

当 x, y 正交时, $x^T y = y^T x = 0$

$$\therefore |x+y|^2 = x^T x + y^T y = |x|^2 + |y|^2 \quad \square$$

子空间正交

子空间正交: 存在子空间 S 和子空间 T , 对于 $\forall x \in S, \forall y \in T, x^T y = 0$ (x, y 正交), 则 $S \perp T$.

1. 四个基本子空间的正交关系

$C(A^T) \perp N(A), \quad C(A) \perp N(A^T), \quad$ 正交的两个子空间互为正交补.

首先我们看的是 A 的零空间和行空间正交, $\forall v \in N(A), Av = 0$, 对于矩阵 A 的所有行向量 $A_1, A_2 \dots A_m$, 我们有 $A_1 v = 0, A_2 v = 0 \dots A_m v = 0$, 而 A 的行空间为行向量的线性组合, $\forall w \in C(A^T), w = c_1 A_1 + c_2 A_2 \dots + c_m A_m$, $wv = c_1 A_1 v + c_2 A_2 v \dots + c_m A_m v = 0$. 因此矩阵 A 的零空间和行空间正交.

我们接下来看的是矩阵 A 的列空间和转置零空间, 设 A^T 的各行为 $A_1, A_2 \dots A_n, \forall w \in N(A^T)$, 我们有 $A_1 w = 0, A_2 w = 0 \dots A_n w = 0$, 根据矩阵转置, 我们知道 $A_1 = \text{col}_1^T, A_2 = \text{col}_2^T \dots A_n = \text{col}_n^T$, 所以我们得到 $\text{col}_1^T w = 0, \text{col}_2^T w = 0 \dots \text{col}_n^T w = 0$, A 的列空间为列向量的线性组合, $\forall v \in C(A), v = c_1 \text{col}_1 + c_2 \text{col}_2 \dots + c_n \text{col}_n, w * v^T = w * \text{col}_1^T + w * \text{col}_2^T \dots w * \text{col}_n^T = 0$. 因此矩阵 A 的列空间和转置零空间正交.

求解一个无解的方程组 $AX = b$ 的解

在现实实际问题中我们往往能得到很多数据, 他们有许多是无用的, 我们得到方程组 $AX = b$, 矩阵 A 是 $m * n$ 型矩阵, 其中 $m > n$, 我们需要剔除一些数据求出方程组的最优解.

我们来求解方程组 $A^T A \hat{X} = A^T b$, 矩阵 $A^T A$ 变成 $n * n$ 型矩阵, $A^T A$ 不一定是可逆矩阵, 具体要看矩阵 A 的列向量是否线性无关, 如果矩阵 A 列向量线性无关, 矩阵 $A^T A$ 可逆.

$$N(A^T A) = N(A), \quad \text{rank}(A^T A) = \text{rank}(A)$$

Chapter 14

子空间投影

投影 *Project*

存在向量 a, b , 向量 b 在向量 a 方向上的投影 $p = xa$, 向量 $b - p$ 满足 $(b - p) \perp a$, 所以我们得到 $(b - xa) \cdot a = 0$; 即 $x = \frac{b^T a}{a^T a}$, 向量 b 在向量 a 方向上的投影 $p = \frac{b^T a}{a^T a} a$.

我们想要向量 b 直接得到在向量 a 上的投影向量 $p, p = Pb$, 我们将 P 称之为投影矩阵, $P = \frac{aa^T}{a^T a}$.

当矩阵 A 列空间为 1 维时, 矩阵 P 有一些特殊的性质,

1. 例如 $C(A) = ca, c \in R; \text{rank}(A) = 1$
2. 投影矩阵 P 的列空间是过向量 a 的一条直线.
3. $P^T = P; \quad P^2 = P$

投影的目的

书接上章, 我们说到在实际问题中方程组不一定有解, 针对这个问题我们看方程组 $AX = b$, AX 在矩阵 A 的列空间内, 向量 b 不在 A 列空间内时方程组无解, 为了解决这个问题, 我们将向量 b 进行微调, 找到向量 b 在矩阵 A 列空间上的投影向量 p , 方程 $AX = p$ 一定有解. 这就是我们提出投影的目的, 寻找到最优解 \hat{X} .

假设向量 A 的列空间是 2 维子空间, A 的列空间的两个基为 a_1, a_2 , $A = \begin{bmatrix} a_1 & a_2 \end{bmatrix}$. 设向量 b 在矩阵 A 的列空间的投影为 $p = x_1 a_1 + x_2 a_2$, $\hat{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 是方程 $A\hat{X} = p$ 的解. 我们需要求出解 \hat{X} , 我们由 $(b - p) \perp$

$C(A)$ 得到:

$$\begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix} (b - A\hat{X}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow A^T(b - A\hat{X}) = \mathbf{0} \quad (14.1)$$

后面的式子是矩阵形式, 我们进一步可以得到 $A^T A \hat{X} = A^T b$, 最终解得 $\hat{X} = (A^T A)^{-1} A^T b$.

我们还知道投影 $p = A\hat{X} = A(A^T A)^{-1} A^T b$, $p = Pb$; 因此我们得到投影矩阵:

$$P = A(A^T A)^{-1} A^T$$

当 A 为 1 维时, $(A^T A)^{-1} = \frac{1}{A^T A} = \frac{1}{a^T a}$ $P = \frac{aa^T}{a^T a}$, 也满足上式.

一般投影矩阵的性质

1. 一般形式 $P = A(A^T A)^{-1} A^T$
2. $P^T = P$
3. $P^2 = P$

总结

我们知道投影矩阵的用途和用法了, 它出现在当方程组方程数很多时, 有许多错误的的数据, 我们需要找到一个近似的最优解, 用投影矩阵来模拟出最优的右侧向量 b . 其实这就是线性回归的回归方程的求解原理.

Chapter 15

投影矩阵和最小二乘

投影矩阵 $P = A(A^T A)^{-1} A^T$ 是矩阵 A 的投影矩阵, 存在向量 b 在矩阵 A 中的投影 Pb :

$$\text{If } b \in C(A), \quad Pb = b; \quad \text{If } b \perp C(A), \quad Pb = \mathbf{0}.$$

最小二乘法

存在数据 $(x_1, y_1), (x_2, y_2), (x_3, y_3) \dots (x_n, y_n)$, 求线性回归方程 $y = C + Dx$.

$$\text{我们得到方程组 } AX = b, \text{ 其中 } A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{bmatrix}, \quad X = \begin{bmatrix} C \\ D \end{bmatrix},$$

$$b = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}; \quad \text{显然当这 } n \text{ 个点不共线的时候, 方程组是无解的, 我们需要找}$$

到一条最近似的回归直线, 满足差距最小. 我们假设每个点与直线上对应点

$$\text{的差距为 } e_i, \text{ 我们的目的是使得 } e_1^2 + e_2^2 + \dots + e_n^2 \text{ 最小, 假设向量 } e = \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{bmatrix},$$

直线上对应点的纵坐标为 $p = \begin{bmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{bmatrix}$, 我们有 $e = b - p$, e 就是误差. 我们

的目的就是找出最优解 $\hat{X} = \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix}$ 和近似向量 p .

上一章我们提到投影矩阵 $P = A(A^T A)^{-1} A^T$, 向量 p 是向量 b 在矩阵 A 上的投影, 因此:

$$\begin{cases} p = Pb \\ A\hat{X} = p \\ P = A(A^T A)^{-1} A^T \end{cases} \quad (15.1)$$

由此方程可以解出两个未知量 \hat{X} 和 p .

注: 矩阵 A 列向量线性无关, 则 $A^T A$ 一定可逆.

证明. 若想证明 $A^T A$ 可逆, 我们只需证明 $A^T A X = 0$ 的零空间只有 $\mathbf{0}$

我们来看方程 $A^T A X = 0$, 方程左右两端同时左乘 X^T , 我们得到方程:

$$X^T A^T A X = 0 \Leftrightarrow (AX)^T (AX) = 0 \Leftrightarrow AX = 0$$

因此我们可以得到方程 $A^T A X = 0$ 的解和方程 $AX = 0$ 相同, 矩阵 A 列向量线性无关, A 的零空间只有 $\mathbf{0}$, 所以 $A^T A X = 0$ 的零空间只有 $\mathbf{0}$. \square

Chapter 16

正交矩阵和 *Schmidt* 正交化

标准正交向量组

现有一组向量 $q_1, q_2, q_3 \dots q_n$, 满足以下关系时, 我们称 $q_1, q_2, q_3 \dots q_n$ 为一组标准正交向量组.

$$q_i q_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

正交基

向量空间 S 的一组基, 这组基满足是标准正交向量组.

正交矩阵和标准正交矩阵

1. 假设矩阵 Q 的列向量为标准正交向量组, 我们有 $Q^T Q = I$, 我们称 Q 为正交矩阵, 并且我们有 $Q^T = Q^{-1}$, 任意置换矩阵都是正交矩阵.
2. 假设矩阵 Q 的列向量为正交向量组, 我们有 $Q^T Q = D$, 其中 D 为对角矩阵.

正交矩阵和投影矩阵

我们在第 14 章中提到向量投影到矩阵空间中投影矩阵 $P = A(A^T A)^{-1} A^T$, 我们换一个角度看, 假设矩阵 Q 的列空间和矩阵 A 的列空间相同, 那么投影矩阵 $P = Q(Q^T Q)^{-1} Q^T$, 我们令矩阵 Q 的列向量为矩阵 A 列空间的一组标准正交基, Q 是正交矩阵, $P = Q Q^T$.

Graham - Schmidt 正交化法

对于矩阵 A , 我们要想得到正交矩阵, 必须要让列向量成为标准正交向量组. 我们先假设矩阵 A 列空间有一组基 $\alpha_1, \alpha_2, \dots, \alpha_n$, 重新求得的正交向量组为 $\beta_1, \beta_2, \dots, \beta_n$, 我们有标准正交向量组 $q_1 = \frac{\beta_1}{|\beta_1|}, q_2 = \frac{\beta_2}{|\beta_2|}, \dots, q_n = \frac{\beta_n}{|\beta_n|}$, 我们可以得到 $\beta_1 = \alpha_1, \beta_2 = \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1, \dots$. 我们依次可以求得矩阵 A 的 n 个标准正交基 (向量减去前面几个基向量方向的投影): $\beta_n = \alpha_n - \sum_{i=1}^n \frac{(\alpha_n, \beta_i)}{(\beta_i, \beta_i)}\beta_i$.

我们可以得到以下的关系式:

$$A = QR, R \text{ is upper triangular}$$

证明. 假设 $A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}, Q = \begin{bmatrix} q_1 & q_2 & \dots & q_n \end{bmatrix}$, 那么 $R = \begin{bmatrix} a_1^T q_1 & a_2^T q_1 & \dots & a_n^T q_1 \\ a_1^T q_2 & a_2^T q_2 & \dots & a_n^T q_2 \\ \dots & \dots & \dots & \dots \\ a_1^T q_n & a_2^T q_n & \dots & a_n^T q_n \end{bmatrix}$

我们由正交矩阵知道:

$$a_i^T q_j = 0, \text{ if } i < j$$

由此我们知道 R 是上三角型矩阵.

□

Chapter 17

行列式及其性质

行列式

假设方阵 A , 我们将 A 的行列式记作: $\det A = |A|$, 我们有以下结论:

1. 矩阵可逆等价于行列式非零, 还等价于 $\text{rank}(A) \neq 0$, 用数学表达式写下来就是:

$$A^{-1} \text{ exists} \Leftrightarrow \det A \neq 0 \Leftrightarrow \text{rank}(A) = n$$

2. 交换行列式的两行, 行列式的值的符号会相反.

3. $\det I = 1$

4. $\det P = 1$ or -1 , P 表示置换矩阵.

5. 行列式中如果两行相等, 那么行列式的值为 0.

6. 高斯消元不影响行列式的值, 即 $\det A = \det U$

7. 行列式有一行为 0, 行列式的值为 0.

$$8. U = \begin{bmatrix} d_1 & a & \dots & b \\ 0 & d_2 & \dots & c \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{bmatrix} = d_1 d_2 d_3 \dots d_n.$$

9. $\det A = 0$, 当且仅当 A 是奇异矩阵.

10. A, B 都是方阵, $\det AB = (\det A)(\det B) \Rightarrow \det A^{-1} = \frac{1}{\det A}$;
 $\det A^2 = (\det A)^2 = |A|^2$; $\det mA = m^n \det A = m^n |A|$.

11. $\det A^T = \det A$.

12. $\det A^{-1} = |A|^{n-1}$

Chapter 18

行列式公式和代数余子式

行列式公式

从每一行中取: $\det A = \sum (-1)^{i+j+\dots+k} (a_{1i} a_{2j} \dots a_{nk}), i, j, \dots k$ 是 $1, 2, \dots n$ 的排列.

从每一列中取: $\det A = \sum (-1)^{i+j+\dots+k} (a_{i1} a_{j2} \dots a_{kn}), i, j, \dots k$ 是 $1, 2, \dots n$ 的排列.

代数余子式

A 是 $n * n$ 型的行列式, 把 A 按照一行展开得到行列式的值, $\det A = \sum_{j=1}^n a_{ij} C_{ij}$. 其中 C_{ij} 表示的是行列式中 a_{ij} 的代数余子式; $C_{ij} = (-1)^{i+j} \det |(n-1) \text{ matrix (with row } i \text{ col } j \text{ erased)}|$.

其实我们刚刚已经发现了两种求解行列式的方法, 分别是代数余子式和行列式公式.

范德蒙行列式

Chapter 19

Cramer 法则和逆矩阵

假设 A 是 $n \times n$ 方阵, 且假设 A^{-1} 存在, 我们可以利用高斯-若尔当消元法求解, 我们知道这是一种算法, 现在我们利用行列式求得逆矩阵的公式:

$$A^{-1} = \frac{1}{\det A} C^T \quad (19.1)$$

C 是 A 的伴随矩阵, $c_{ij} = a_{ij}A_{ij}$, 其中 $A_{ij} = (-1)^{i+j} \det |(n-1) \text{ matrix (with row } i \text{ col } j \text{ erased)}|$

证明. 假设这个公式成立, 我们只需要证明: $AA^{-1} = I \Leftrightarrow AC^T = (\det A)I$

$$AC^T = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} c_{11} & c_{21} & \dots & c_{n1} \\ c_{12} & c_{22} & \dots & c_{n2} \\ \dots & \dots & \dots & \dots \\ c_{1n} & c_{2n} & \dots & c_{nn} \end{bmatrix} \quad \text{由公式 (19.1)}$$

我们知道 AC^T 的对角线上的元素都为 $\det A$, 非对角线上的元素表示的是行列式 A 中第 i 行与第 j 行每个元素的代数余子式的乘积之和, 我们发现第 i 行与第 j 行的代数余子式之积可以理解为将第 j 行换成第 i 行的元素, 然后将行列式按照第 j 行展开, 由于行列式中由两行相同, 行列式的值为 0, 因此 AC^T 中非对角线的元素全部为 0.

因此 $AC^T = (\det A)I$ 得证.

□

Cramer 法则

假设矩阵 A 是可逆矩阵, 对于方程 $AX = b$, 我们知道这个方程的解为

$$X = A^{-1}B, \text{ 由公式 (19.1), 我们知道 } X = \frac{1}{\det A} C^T b \Leftrightarrow X = \frac{1}{\det A} (b_1 \begin{bmatrix} c_{11} \\ c_{12} \\ \dots \\ c_{1n} \end{bmatrix} +$$

$$b_2 \begin{bmatrix} c_{21} \\ c_{22} \\ \dots \\ c_{2n} \end{bmatrix} + \dots + b_n \begin{bmatrix} c_{n1} \\ c_{n2} \\ \dots \\ c_{nn} \end{bmatrix}), \text{ 我们知道 } X \text{ 的分量 } X_1 = \frac{1}{\det A} (b_1 c_{11} + b_2 c_{21} +$$

$$\dots + b_n c_{n1}) = \frac{\det B_1}{\det A}, \text{ 我们可以得到 } X = \begin{bmatrix} \frac{\det B_1}{\det A} \\ \frac{\det B_2}{\det A} \\ \dots \\ \frac{\det B_n}{\det A} \end{bmatrix}; \text{ 其中 } B_i \text{ 是将行列式}$$

的第 i 列换成向量 b .

体积 假设三维空间内存在 n_1, n_2, n_3 三个向量, 这三个向量所构成的空间平行六面体的体积 $V = \det A, A = \begin{vmatrix} n_1 \\ n_2 \\ n_3 \end{vmatrix}$.

Chapter 20

特殊矩阵总结

1. 对角矩阵: $A = \text{diag}\{d_1, d_2, \dots, d_n\}$, 除了主对角线上的原宿其余元素都为 0.
2. 上三角和下三角矩阵
3. 基本矩阵: 我们将 E_{ij} 成为基本矩阵, 这个矩阵是只有 (i, j) 元为 1 其余为 0 的 $n * n$ 矩阵.
4. 由单位矩阵经过一次初等变换得到的矩阵被称为初等矩阵.
5. 可逆矩阵: 满足 $AB = I$ 的方阵互称可逆矩阵.
6. 正交矩阵: 满足 $AA^T = I$ 的方阵称为正交矩阵.
7. 相似矩阵: $A = U^{-1}BU$, A, B, U 都是 n 阶方阵.
8. 幂零矩阵: 满足 $A^t = 0, t_{min}$ 被称为幂零指数.

Chapter 21

特征值和特征向量

假设存在方阵 A , 我们有 $AX = \lambda X$, 我们称 λ 为特征值, X 为特征向量. 那么我们该如何求解特征值和特征向量呢?

$AX = \lambda X \Leftrightarrow (\lambda I - A)X = 0$, 我们只需要保证对于方阵 $(\lambda I - A)$ 零空间内存在非零向量即可; 我们得出 $\text{rank}(\lambda I - A) \neq n \Leftrightarrow \det(\lambda I - A) = 0$, 我们解出 λ 是方阵 A 的特征值, 而方程 $(\lambda I - A)X = 0$ 对应特征值的特征向量.

$$\text{假设 } A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \text{ 那么 } \det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{n1} & -a_{n2} & \dots & \lambda - a_{nn} \end{vmatrix},$$

最后化简得到的是一个关于 λ 的多项式 $f(\lambda)$:

$$f(\lambda) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \dots + c_0$$

我们知道特征值就是方程 $f(\lambda) = 0$ 的解, 由根与系数的关系:

$$\begin{cases} \sum_{i=1}^n \lambda_i = \text{trace}(A) \\ \prod_{i=1}^n \lambda_i = (-1)^n \det A \end{cases}$$

注意

1. $n \times n$ 的矩阵 A 可对角化的等价条件为 A 有 n 个线性无关的特征向量.

我们从特征向量的定义可得:

$$\left\{ \begin{array}{l} Ax_1 = \lambda_1 x_1 \\ Ax_2 = \lambda_2 x_2 \\ \dots \\ Ax_n = \lambda_n x_n \end{array} \right. \Rightarrow AX = X * \text{diag}(\lambda_1, \lambda_2 \dots \lambda_n), X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

当矩阵 A 有 n 个线性无关的特征向量时, 矩阵 X 是可逆矩阵, 所以就有了:

$$\left\{ \begin{array}{l} X^{-1}AX = \text{diag}(\lambda_1, \lambda_2 \dots \lambda_n) \\ A = X * \text{diag}(\lambda_1, \lambda_2 \dots \lambda_n) * X^{-1} \end{array} \right.$$

2. 假设 $AX = \lambda X$ 的特征值为 $\lambda_1, \lambda_2 \dots \lambda_n$, 那么 $(A + mI)X = \alpha X$ 的特征值 $\alpha_1, \alpha_2 \dots \alpha_n$ 满足 $\alpha_i = m + \lambda_i, i = 1, 2 \dots n$, 特征向量不变.

Chapter 22

对角化和矩阵幂

在上一章中, 我们知道如何去求解一个方阵的特征值和特征向量, 那么我们求解这些有什么用呢?

假设方阵 A 有 n 个线性无关的特征向量, 那么将这些特征向量作为列向量组成新的方阵 S , 我们发现 $AS = A \begin{bmatrix} X_1 & X_2 & \dots & X_n \end{bmatrix} \Leftrightarrow \begin{bmatrix} AX_1 & AX_2 & \dots & AX_n \end{bmatrix}$; 由特征向量的性质得: $AX_i = \lambda_i X_i$, 我们得到

$$AS = \begin{bmatrix} \lambda_1 X_1 & \lambda_2 X_2 & \dots & \lambda_n X_n \end{bmatrix} \Leftrightarrow \begin{bmatrix} X_1 & X_2 & \dots & X_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \lambda_n \end{bmatrix},$$

我们称 $\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$ 为对角矩阵, 记作 $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, 由

此我们得到以下两个公式:

$$\begin{cases} S^{-1}AS = D \\ A = SDS^{-1} \end{cases} \quad (22.1)$$

我们由公式 (21.1) 知道: $A^2 = SD^2S^{-1}$; $A^n = SD^nS^{-1}$, 这也是我们求特征向量和对角化矩阵的目的, 当求方阵的高阶幂时, 可以大大简化计算量, 因为 $D^n = \text{diag}(\lambda_1^n, \lambda_2^n, \dots, \lambda_n^n)$.

我们发现: if all $|\lambda_i| < 1$, when $k \rightarrow +\infty$, $A^k \rightarrow 0$.

矩阵可对角化的条件

矩阵有 n 个线性无关的特征向量.

实际应用: 斐波那契数列

一阶线性微分方程 \Rightarrow 一阶向量方程

1. 假设向量 $u_{k+1} = Au_k$, 已知 u_0 , 求 u_{100}

我们发现 $u_k = A^k u_0$, 关于矩阵的幂, 我们用特征值和特征向量将矩阵对角化, 我们首先求出矩阵的特征值和特征向量, 特征向量构成矩阵 $S, A = S^{-1}DS, D$ 是矩阵 A 对角化的结果, 对角线上元素是特征值, 我们有 $u^k =$

$$SD^k S^{-1}u_0, u_0 = Sv, \text{ 其中 } v = \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{bmatrix}, \text{ 因此 } u_k = SD^k V = S \begin{bmatrix} c_1 \lambda_1^k \\ c_2 \lambda_2^k \\ \dots \\ c_n \lambda_n^k \end{bmatrix}.$$

我们发现 $u_k = C_1 \lambda_1^k + C_2 \lambda_2^k + \dots + C_n \lambda_n^k$, 其中 C_i 都是列向量, 我们通过首项 u_0 可以解出, 我们可以得到通项公式.

2. 设 $F_0 = 0, F_1 = 1, F_{k+2} = F_{k+1} + F_k$, 求 F_{100}

设 $u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}, u_{k+1} = \begin{bmatrix} F_{k+2} \\ F_{k+1} \end{bmatrix}, u_{k+1} = Au_k, A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$; 我们解出矩阵 A 的特征值: $\lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_2 = \frac{1-\sqrt{5}}{2}$, 根据上题, 我们发现: $u_k = C_1 \lambda_1^k + C_2 \lambda_2^k$, 又因为 $u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

我们解得通项公式 $F_k = \frac{1}{\sqrt{5}}((\frac{1+\sqrt{5}}{2})^k - (\frac{1-\sqrt{5}}{2})^k)$, 并且 $C_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ \frac{2}{1+\sqrt{5}} \end{bmatrix}, C_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 \\ \frac{2}{1-\sqrt{5}} \end{bmatrix}$.

Chapter 23

微分方程 $\exp(At)$

一阶微分方程组

我们先来看一个例子:

$$\begin{cases} \frac{du_1}{dt} = -u_1 + 2u_2 \\ \frac{du_2}{dt} = u_1 - 2u_2 \end{cases} \quad (23.1)$$

我们有 $u(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $A = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix}$, 我们发现 A 的两个特征值 $\lambda_1 =$

$0; \lambda_2 = -3$, 两个特征值对应的特征向量分别为 $x_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}; x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

原方程组的通解为 $u(t) = c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2 \Leftrightarrow u(t) = \begin{bmatrix} 2c_1 + c_2 e^{-3t} \\ c_1 - e^{-3t} \end{bmatrix}$

当 $t = 0$ 时, $u(0) = \begin{bmatrix} 2c_1 + c_2 \\ c_1 - c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} c_1 = \frac{1}{3} \\ c_2 = \frac{1}{3} \end{cases}$.

$u(+\infty) = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$, 这个例子最后是稳定的, 但是并不是所有的都是稳定

的, 我们发现:

1. 当所有的特征值 $\operatorname{Re} \lambda_i < 0, t \rightarrow +\infty, u(t) \rightarrow 0$.
2. 当 $\lambda_1 = 0, \text{others } \operatorname{Re} \lambda_i < 0$, 最后会形成一个稳态, 比如这个例子.
3. 当 $\forall \lambda_i > 0, t \rightarrow +\infty, u(t) \rightarrow +\infty$.
4. 当矩阵 A 满足 $\begin{cases} -\operatorname{trace}(A) < 0 \\ \det A > 0 \end{cases}$, $u(t)$ 是稳定的.

$$5. \frac{du}{dt} = Au, \text{ 设 } u = SV, S = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}, V(t) = e^{Dt}V(0); \frac{du}{dt} = Au \Leftrightarrow S \frac{dv}{dt} = ASV \Rightarrow S^{-1}S \frac{dv}{dt} = S^{-1}ASV = DV$$

$$\text{即: } u(t) = SV(t) = Se^{Dt}V(0) = Se^{Dt}S^{-1}u(0) = e^{At}u(0)$$

$$\text{我们只需要证明: } Se^{Dt}S^{-1} = e^{At}$$

$$\text{我们利用泰勒展开式: } e^{At} = I + \frac{(At)^2}{2} + \frac{(At)^3}{6} + \dots + \frac{(At)^n}{n!} + \dots \Rightarrow e^{At} = SS^{-1} + \frac{S(Dt)^2S^{-1}}{2} + \frac{S(Dt)^3S^{-1}}{6} + \dots + \frac{S(Dt)^nS^{-1}}{n!} + \dots \Rightarrow e^{At} = S(I + \frac{(Dt)^2}{2} + \frac{(Dt)^3}{6} + \dots + \frac{(Dt)^n}{n!} + \dots)S^{-1} \Rightarrow e^{At} = Se^{Dt}S^{-1}$$

二阶微分方程

二阶微分方程 \Rightarrow 一阶向量方程

$$\text{比如: } \ddot{y}(x) + b\dot{y}(x) + ky = 0$$

$$\text{我们设 } F_{k+1} = \begin{bmatrix} \ddot{y}(x) \\ \dot{y}(x) \end{bmatrix}, F_k = \begin{bmatrix} \dot{y}(x) \\ y(x) \end{bmatrix}, F_{k+1} = AF_k, A = \begin{bmatrix} -b & -k \\ 0 & 1 \end{bmatrix},$$

可以用解一阶线性微分方程组的方法来解这个方程组.

Chapter 24

马尔可夫矩阵和傅里叶级数

Markov Matrix

矩阵 A 满足以下两个性质, 我们称矩阵 A 为 *Markov Martix*

$$1. \forall a_{ij} \in A, a_{ij} \geq 0$$

$$2. \forall j \in n, \sum col_j = 1$$

性质

1. 矩阵 A 一定有一个特征值为 1, 其余的特征值 $|\lambda_i| \leq 1$.

2. 若 A 是 *Markov Martix*, A^k 也是 *Markov Martix*.

3. 对于第一个特征向量 X_1 中 $x_i \geq 0$.

4. 矩阵 A 和矩阵 A^T 有相同的特征值.

注意: 关于性质 1 的证明, 我们可以发现矩阵 $A - I$ 行向量之和为零向量, 所以矩阵 $A - I$ 是奇异矩阵, 也就是说 $rank(A - I) < n \Rightarrow |A - I| = 0$, 0 是矩阵 $A - I$ 的一个特征值, 所以矩阵 A 一定有一个特征值为 1.

Markov Martix 实际应用: 人口迁移问题

假设有两个小镇 A, B , 小镇 A 每年向小镇 B 的迁移率为 α , 留在小镇 A 的比率为 $1 - \alpha$; 小镇 B 每年向小镇 A 的迁移率为 β , 留在小镇 B 的比率为 $1 - \beta$, 初始两个小镇人数为 a, b , 求 k 年后两个小镇的人数?

设 $U_k = \begin{bmatrix} a_k \\ b_k \end{bmatrix}$, $U_0 = \begin{bmatrix} a \\ b \end{bmatrix}$, 表示 k 年后的 A, B 小镇人数, $U_{k+1} =$

$$AU_k, A = \begin{bmatrix} 1 - \alpha & \beta \\ \alpha & 1 - \beta \end{bmatrix}, \text{ 所以 } U_k = A^k U_0.$$

我们求出矩阵 A 的特征值 λ_1, λ_2 , $U_k = C_1 \lambda_1^k + C_2 \lambda_2^k$, 其中 C_1, C_2 根据

U_0 解出.

傅里叶级数

假设 n 维空间内存在一组标准正交基 q_1, q_2, \dots, q_n , 空间中任意的一个向量 V 都可以表示为:

$$V = x_1 q_1 + x_2 q_2 + \dots + x_n q_n \quad (24.1)$$

等式两边分别依次左乘 q_i^T , 可以得到 $x_i = q_i^T V \Rightarrow X = Q^T V$, 根据 (23.1) 式, 我们发现 $V = QX \Rightarrow X = Q^{-1}V$, 由此得到下面的两个式子:

$$\begin{cases} X = Q^T V \\ X = Q^{-1} V \end{cases}$$

我们得出: $Q^T = Q^{-1}$, 上面的这些, 尤其是 (23.1) 式是傅里叶级数建立的基础. 首先类比于向量的内积, 我们发现两个函数的内积 $f * g = \int_{min}^{max} f(x)g(x)dx$, 可以将函数看作是无穷维的向量, 那么这个内积就等价于每一个点函数值乘积之和 $\sum_{min}^{max} f * g$, 所以 $f * g = \int_{min}^{max} f(x)g(x)dx$.

对于周期函数 $f(x)$, 我们可以看作无穷维空间中的一个向量, 这个空间有一组标准正交基 $\sin(ix), \cos(ix), i \in 1, 2, 3 \dots n$. 周期函数的傅里叶展开式如下:

$$f(x) = a_0 + \sum_{i=1}^n a_i \cos(ix) + \sum_{i=1}^n b_i \sin(ix) \quad (24.2)$$

对比 (23.1) 和 (23.2) 两式, 我们可以用同样的方法求出对应基的系数, 方程两边同时在一个周期上求定积分可得 $\int_0^T f(x)dx = \frac{a_0 T}{2}$; 方程两边同时乘以 $\sin(ix)$ 后在一个周期上求定积分可得: $\int_0^T f(x) \sin(ix)dx = \int_0^T a_0 \sin(ix)dx + \int_0^T b_i (\sin(ix))^2 dx$; 方程两边同时乘以 $\cos(ix)$ 后在一个周期上求定积分可得: $\int_0^T f(x) \cos(ix)dx = \int_0^T a_0 \cos(ix)dx + \int_0^T a_i (\cos(ix))^2 dx$.

由此我们得到各个基前面的线性组合系数:

$$\begin{cases} \int_0^T f(x)dx = \frac{a_0 T}{2} \\ \int_0^T f(x) \sin(ix)dx = \int_0^T a_0 \sin(ix)dx + \int_0^T b_i (\sin(ix))^2 dx \\ \int_0^T f(x) \cos(ix)dx = \int_0^T a_0 \cos(ix)dx + \int_0^T a_i (\cos(ix))^2 dx \end{cases} \quad (24.3)$$

假设 $T = 2\pi$, 将公式 (23.2) 化简得到下式:

$$\begin{cases} a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx)dx, \quad n = 0, 1, 2 \dots \\ b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx)dx, \quad n = 1, 2 \dots \end{cases} \quad (24.4)$$

Part III

正定矩阵及其应用

Chapter 25

对称矩阵及其正定性

假设矩阵 A 是实对称矩阵, 我们有以下两个推论:

1. 矩阵 A 的特征值都是实数, 特征值的乘积等于主元乘积.
2. 矩阵 A 的特征向量可以找出一组是正交的.

对于实对称矩阵, 我们将其对角化可以得到 $A = SDS^{-1}$, 我们由性质 2 可以得到 $A = QDQ^{-1}$, 又因为 Q 是正交矩阵, $QQ^T = I$, 因此 $Q^T = Q^{-1} \Rightarrow A = QDQ^T$

关于性质 1 的证明, 我们首先由 $AX = \lambda X$, 这个方程两边同时取共轭得到 $\overline{A} \overline{X} = \overline{\lambda} \overline{X}$, 因为矩阵 A 是实对称矩阵, 我们有 $\overline{A} = A$, 因此我们得到: $A \overline{X} = \overline{\lambda} \overline{X} \Rightarrow \overline{X}^T A = \overline{X}^T \overline{\lambda}$, 我们得到如下的表达式:

$$\begin{cases} AX = \lambda X \\ \overline{X}^T A = \overline{\lambda} \overline{X}^T \end{cases} \Rightarrow \begin{cases} \overline{X}^T AX = \lambda \overline{X}^T X \\ \overline{X}^T AX = \overline{\lambda} \overline{X}^T X \end{cases} \quad (25.1)$$

由 (24.1) 得 $\lambda = \overline{\lambda}$, λ 是实数.

注意: 假设向量 A 是复数向量, $|A| = \overline{A}^T A$; 假设矩阵 A 是复数矩阵, 我们要求矩阵 A 特征值都为实数且特征向量正交, 必须满足 $A = \overline{A}^T$.

对称矩阵分解

假设矩阵 A 是对称矩阵, 那么 $A = QDQ^{-1} \Rightarrow A = \sum_{i=1}^n \lambda_i q_i q_i^T$, 在投影矩阵中我们知道对于任意一个矩阵 q , 其投影矩阵为 $P = \frac{qq^T}{q^T q}$, 特殊的当 q 为单位列向量时, $P = qq^T$, 因此对称矩阵可以分解为 n 个投影矩阵的线性组合.

对称矩阵正定性判定

1. 所有的特征值为正: $\lambda_i > 0$.
2. 所有的主元 $x_i > 0$.
3. 所有的顺序主子行列式值都为正.
4. 对任意非零向量 $X, X^TAX > 0$.

Chapter 26

复数矩阵和快速傅里叶变化

复数矩阵

假设向量 $Z = \begin{bmatrix} z_1 \\ z_2 \\ \dots \\ z_n \end{bmatrix}$ 是 n 维复空间中的一个向量, 即 $Z \in C^n$, 我们

之前提到过如何求复向量模长的方法: $|Z| = \overline{Z}^T Z$, 我们将 \overline{Z}^T 记作 Z^H , 我们就得到了复向量求模长和内积的公式:

$$\begin{cases} |X| = X^H X, X \in C^n \\ |Y| = Y^H Y, Y \in C^n \\ X * Y = X^H Y, X, Y \in C^n \end{cases}$$

同时在实数矩阵中的一些性质在复数矩阵中有一点改变, 首先是“对称”, 我们将 $A^H = A$ 的矩阵 A 称为 *Hertmitian* 矩阵, 也被称为复共轭对称矩阵, 具有实对称矩阵的某些性质: 特征值都为实数且特征值向量相互正交.

其次是正交, 复空间中的一组标准正交基 q_1, q_2, \dots, q_n 满足以下条件:

$$q_i q_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

由标准正交基组成的矩阵 Q 满足: $Q^H Q = I$

快速傅里叶变化

傅里叶矩阵是一种非常重要的复数矩阵, 而快速傅里叶变化 (*FFT*) 则是一种十分重要的手段

$$F_n = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & w & w^2 & \dots & w^{n-1} \\ 1 & w^2 & w^4 & \dots & w^{2(n-1)} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & w^{n-1} & w^{2(n-1)} & \dots & w^{(n-1)^2} \end{bmatrix}, w^n = 1$$

这就是 n 阶的傅里叶矩阵, 我们发现 $F_{ij} = w^{i*j}, i \in 0, 1, 2 \dots n-1; j \in 0, 1, 2 \dots n-1$, 矩阵的行和列都从 0 开始.

傅里叶矩阵性质

1. 矩阵的列向量正交, 且列向量为单位向量, 我们有 $F_n^H F_n = I$
2. $w = e^{\frac{2\pi}{n}} = \cos(\frac{2\pi}{n}) + i \sin(\frac{2\pi}{n})$

Chapter 27

二次型、标准型、规范型和正定二次型

$f(x_1, x_2, \dots, x_n) = \sum_{i=1, j=1}^n a_{ij} x_i x_j$, 称为 n 元二次型, 然后我们得到一个矩阵 A , 来表示多项式的系数, 如此我们有了 $f(x_1, x_2, \dots, x_n) = X^T A X$.

定义

1. 合同: 假设矩阵 A, B 都是 n 阶方阵, 存在可逆矩阵 C 使得 $C^T A C = B$, 我们称 A 与 B 合同.

2. 二次型等价: 假设存在一个可逆矩阵 C 满足 $X = CY$, 二次型 $X^T A X$ 与 $Y^T B Y$ 等价. 这说明二次型等价当且仅当系数矩阵合同.

3. 标准型: 假设二次型 $X^T A X$ 等价于一个只有平方项的二次型, 那么这个只含平方项的二次型称为 $X^T A X$ 的一个标准型 (标准型不唯一).

4. 我们说过实数对称矩阵一定可以对角化, $U^{-1} A U = D, D = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}, U$ 是正交矩阵, $U^T A U = D$, 矩阵 A 和对角矩阵 D 合同, 说明二次型 $X^T A X$ 一定等价于一个只含平方项的二次型. 特别的, $X = UY$ 被称为正交替换.

5. 规范型: 我们可以通过非退化线性变换将二次型 $X^T A X$ 等价于一个标准型 $d_1 y_1^2 + d_2 y_2^2 + \dots + d_p y_p^2 - d_{p+1} y_{p+1}^2 - \dots - d_r y_r^2$, 我们再进行一次非线性退化变换得到 $z_1^2 + z_2^2 + \dots + z_p^2 - z_{p+1}^2 - \dots - z_r^2$, 二次型中只含平方项且平方项系数为 1, -1, 0; 这个二次型被称为规范型.

6. 正负惯性指数: 二次型 $X^T A X$ 的规范型中系数为正的平方项个数是正惯性指数 p , 系数为负的平方项的个数是负惯性指数 $r - p$.

7. 正定二次型: 对于不全为 0 的自变量, 二次型 $X^T A X$ 恒正, 我们就称

之为正定二次型.

定理

1. 惯性定理: n 元二次型的规范型是唯一的.
2. 任意一个 n 元二次型的系数矩阵 A 是实数对称矩阵都合同与一个主对角线上只有 1, -1, 0 的对角矩阵.
3. 两个 n 元二次型等价 \Leftrightarrow 系数矩阵秩相等且正惯性指数相等 \Leftrightarrow 规范型相同 \Leftrightarrow 系数矩阵合同.

Chapter 28

正定矩阵和最小值

定义

对于任意非零向量 X , 我们有 $X^T A X > 0$, 我们称矩阵 A 为正定矩阵.

其实正定矩阵是在对称矩阵中某一些性质较好的矩阵, 满足特征值为正数或者主元都大于 0 或者顺序主子行列式都大于 0.

性质

1. 正定矩阵是对称矩阵.

2. 假设 A, B 都是正定矩阵, $(A+B)$ 也是正定矩阵, 因为 $X^T (A+B) X = X^T A X + X^T B X > 0$.

3. 矩阵 A 是 $m \times n$ 型矩阵, 矩阵 $A^T A$ 是正定矩阵. ($\text{rank}(A) = \min(m, n)$)

关于性质 3 的证明, 我们发现 $X^T A^T A X = (AX)^T A X \Rightarrow |AX|^2 \geq 0$, 当 AX 为零向量时成立, X 为非零向量, 只有当 A 的零空间只有零向量时, 矩阵 $A^T A$ 是正定矩阵.

我们接下来还会讲述一些正定矩阵的性质:

1. n 级实对称矩阵 A 当且仅当矩阵 A 所有的特征值全部大于 0 时是正定矩阵.

2. 实对称矩阵 A 正定的充要条件为 A 的所有顺序主子行列式全部大于 0.

Chapter 29

相似矩阵和若尔当型

相似矩阵

假设存在两个 $n \times n$ 的矩阵 A, B , 满足 $B = M^{-1}AM$, 我们就称 B 相似于 A . 有一个很好的例子就是矩阵对角化时: $A = SDS^{-1} \Rightarrow D = S^{-1}AS$, 矩阵 A 和 D 相似.

性质

1. 相似矩阵有相同的特征值.

关于性质 1 的证明, 我们假设 $A \sim B$, 我们有 $AX = \lambda X \Rightarrow AMM^{-1} = \lambda X \Rightarrow (M^{-1}AM)(M^{-1}X) = \lambda(M^{-1}X) \Rightarrow B(M^{-1}X) = \lambda(M^{-1}X)$, 所以 A, B 有相同的特征向量, 假设 x_i 是 A 的特征向量, 那么 $M^{-1}x_i$ 是 B 的特征向量.

注意: nI 只和自己相似

若尔当型

每一个方阵 A 都相似于一个若尔当型矩阵 J , $J = \begin{bmatrix} J_1 & & \\ & J_i & \\ & & J_j \end{bmatrix}$, blocks = $N(\text{vectors})$,

Chapter 30

奇异值分解

我们在特征值那一章中知道一些方阵可以进行对角化处理进而进行分解, 那么不是方阵的矩阵是否能进行分解呢?

其实答案是肯定的, 假设矩阵是 $m * n$ 型矩阵, 那么矩阵 A 可以分解为 $A = U \Sigma V^T$, 其中矩阵 $U(m * m)$ 是矩阵 A 列空间和左零空间的标准正交基的并; 矩阵 $V(n * n)$ 是矩阵 A 行空间和零空间的标准正交基的并; $\Sigma(m * n)$ 是除了主对角线上的元素外全为 0, 主对角线上的元素为奇异值.

我们换一个背景来理解, 假设矩阵行空间内有一组正交基 v_i , 列空间内有一组正交基 u_i , 这两组正交基很特殊, 满足:

$$\begin{cases} Av_1 = \sigma_1 u_1 \\ Av_2 = \sigma_2 u_2 \\ \dots \end{cases} \Rightarrow AV = U \Sigma, \Sigma = \text{diag}(\lambda_1, \lambda_2 \dots \lambda_n)$$

矩阵 $U = \begin{bmatrix} u_1 & u_2 & \dots & u_r & u_{r+1} & \dots & u_n \end{bmatrix}$, $V = \begin{bmatrix} v_1 & v_2 & \dots & v_r & v_{r+1} & \dots & v_n \end{bmatrix}$, $\text{rank}(A) = r$, 后面的是零空间和左零空间的基. 我们来看一下如何解决这个问题, 首先 $AV = U \Sigma \Rightarrow A = U \Sigma V^{-1} = U \Sigma V^T$, 我们根据实数对称矩阵的性质, 我们有以下的公式:

$$\begin{cases} AA^T = U \Sigma V^T V \Sigma^T U^T \\ A^T A = V \Sigma^T U^T U \Sigma V^T \end{cases} \Rightarrow \begin{cases} AA^T = U \Sigma \Sigma^T U^T \\ A^T A = V \Sigma^T \Sigma V^T \end{cases} \quad (30.1)$$

$\Sigma^T \Sigma = \Sigma \Sigma^T = \text{diag}(\lambda_1^2, \lambda_2^2 \dots \lambda_n^2)$, 我们知道对称矩阵 AA^T 和矩阵 $A^T A$

都可以对角化, 那么对于 (28.1) 两式中 $\sigma_i = \sqrt{\lambda_i}$, U, V 分别是两个对阵矩阵的特征向量矩阵.

我们来看几个例子:

1. $A = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$, 将矩阵 A 进行奇异值分解.

2. $A = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$, 将矩阵 A 进行奇异值分解.

Chapter 31

线性变换及对应矩阵

矩阵来源于线性变化,我们先来看几个例子: $T: R^n \rightarrow R^n$

1. 投影
2. 旋转
3. 平移不是一个线性变化

线性变换规则

$$\begin{cases} T(v+w) = T(v) + T(w) \\ T(cv) = cT(v) \end{cases} \Rightarrow T(cv + dw) = cT(v) + dT(w)$$

满足以上规则的变换都属于线性变换;理解线性变换的根本就是找出背后的矩阵.上面我们仅仅理解的是单个向量线性变换的结果,我们接下来看所有向量(线性空间)的线性变换情况.

假设空间中存在一组基 v_1, v_2, \dots, v_n , 空间中任意一个向量 k 可以写作为 $k = c_1v_1 + c_2v_2 + \dots + c_nv_n$, 我们将 (c_1, c_2, \dots, c_n) 成为向量 k 的坐标值, 存在唯一的表达式, 影响坐标值的只有向量空间的基的选取.

我们用一个矩阵 A 来表示一种线性变换: $T: R^n \rightarrow R^n$, 即将 R^n 中的向量映射到空间 R^n 中. 我们需要两组基, 第一组基用来确定映射前向量的坐标, 第二组基用来确定映射后的向量的坐标.

我们以投影为例, 假设在一个平面上的所有向量通过线性变换全部投影到平面上的一条直线上.

1. 我们选择平面的一组基为沿直线方向和垂直于直线方向的两个单位向量 v_1, v_2 , 直线的基为 v_1 , 我们需要将平面的一个任意一个向量 $V =$

$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ 通过线性变换 T 变为 $W = \begin{bmatrix} c_1 \\ 0 \end{bmatrix}$, 我们发现矩阵 $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ 满足 $AV = W$.

2. 我们选择标准正交向量作为输入空间和输出空间的一组基, 假设直线的倾斜角为 $\frac{\pi}{4}$, 我们用平面的标准正交基 $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, 我们在前面知道投影矩阵 $P = \frac{aa^T}{a^T a}$, 其中 a 是直线上的任意向量, 在这里不妨令 $a = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$, 对于平面上的向量, 在以标准正交基作为基底时, 通过线性变换可以使其变换为直线上的向量: $PV = W$.

接下来的工作就是如何找到线性变换的矩阵 A ?

假设输入空间为 R^n 的一组基为 v_1, v_2, \dots, v_n , 输出空间为 R^m 的一组基为 w_1, w_2, \dots, w_m , 假设将 R^n 中的基线性变换到 R^m 中, 我们有下面的公式:

$$\begin{cases} Av_1 = a_{11}w_1 + a_{12}w_2 + \dots + a_{1m}w_m \\ Av_2 = a_{21}w_1 + a_{22}w_2 + \dots + a_{2m}w_m \\ \dots \\ Av_n = a_{n1}w_1 + a_{n2}w_2 + \dots + a_{nm}w_m \end{cases} \quad (31.1)$$

我们知道 a_{ij} 是输出空间内向量的坐标, 针对输入空间的一组基, 其坐标

$$\text{为 } \begin{bmatrix} 1 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \dots \\ 1 \end{bmatrix}, \text{ 那么矩阵 } A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \text{ 对}$$

于输入向量空间中的基的线性变换都成立, 对于输入空间中任意一个向量 $x = c_1v_1 + c_2v_2 + \dots + c_nv_n$, 根据线性变换的规则:

$$T(x) = c_1T(v_1) + c_2T(v_2) + \dots + c_nT(v_n) = c_1Av_1 + c_2Av_2 + \dots + c_nAv_n = A(c_1v_1 + c_2v_2 + \dots + c_nv_n) = Ax$$

因此对于输入空间中任意向量矩阵 A 都满足, 我们找到了求矩阵 A 的方法.

特殊的线性变换幂函数求导:

假设 $f(x) = c_0 + c_1x + c_2x^2 + \dots + c_nx^n$, 我们该如何求 $f(x)$ 的导数?

我们不难发现其实这也是一个映射的问题, 将一个空间内的向量映射到另一个空间, 我们是要将 R^n 空间通过线性变换得到 R^{n-1} 空间的向量; 其中输入空间的一组基 $\{1, x, x^2, \dots, x^n\}$, 输出空间的一组基 $\{1, 2x, 3x^2, \dots, nx^{n-1}\}$,

我们直接能够得到输入向量的坐标 V , 我们能够找到 A , 使得 $AX = U, U$ 是线性变换后的向量.

举个简单的例子: $f(x) = 2 + x + 3x^2 + 4x^3 + 3x^4$, 那么 $\frac{df(x)}{dx} = AX, X =$

$$\begin{bmatrix} 2 \\ 1 \\ 3 \\ 4 \\ 3 \end{bmatrix}, \text{ 又因为 } \frac{df(x)}{dx} = \begin{bmatrix} 1 \\ 6 \\ 12 \\ 12 \end{bmatrix}, \text{ 所以 } A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}.$$

我们进一步推广, 当 $f(x) = c_0 + c_1x + c_2x^2 + \dots + c_nx^n$ 时, $A =$

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n-1 \end{bmatrix}, A \text{ 是 } (n-1) * n \text{ 型矩阵}.$$

Chapter 32

基变换和图像压缩

矩阵和线性变换的关系

1. 线性变换不一定需要坐标.
2. 矩阵是用坐标来表示线性变换.

图像压缩原理

图像是由一个个像素点构成的, 像素点中有表示 0 255 的数字, 灰白图像每个像素点只有一个分量, 而彩色像素点有三个分量. 比如我们有一张 $512 * 512$ 像素的照片, 我们可以得到一个向量 $x \in R^{512^2}$, 我们压缩图像的关键就是将相似灰度值的像素合在一起, 整体的样子没有太大变化, 但是清晰度降低. 我们需要找到一组很好的基, 来表示出图像的信息并且将少量的信息舍去. 在图像处理中, 常见的是将矩阵分为 $8 * 8$ 的矩阵, 采用 8 阶傅

里叶矩阵的列向量作为一组基, $F_8 = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & w & w^2 & \dots & w^7 \\ 1 & w^2 & w^4 & \dots & w^{14} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & w^7 & w^{14} & \dots & w^{49} \end{bmatrix}$, $w^8 = 1$,

我们通过基变换得到新的向量坐标 c' , 将向量中的小于某个标准的值都去掉 (失真), 这些部分对于图像还原后的影响较小. 我们能通过 $X_i = \sum c'_i V_i$ 将其还原, 不过精确性降低了.

除了傅里叶矩阵的列, 人们还发现了另一种很好的基: 小波基 (FWT) $P =$

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & -1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 & -1 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 1 & -1 & 0 & -1 & 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & -1 & 0 & 0 & 0 & -1 \end{bmatrix}, \text{我们发现矩阵 } P \text{ 的列向量是正}$$

交的, 这是一个很好的性质, 我们比较压缩前的和压缩后的得到 $X = PC \Rightarrow C = P^{-1}X \Rightarrow C = \sum P^T X$

基变换

简单来说就是已知两组基和两种线性变换; 这两种线性变换分别对应的矩阵的关系.

假设空间中有一组基 v_1, v_2, \dots, v_n 和这组基对应的线性变换 A ; 另外一组基 w_1, w_2, \dots, w_n 和这组基对应的线性变换 B , 满足 $A B \Rightarrow B = M^{-1} A M$.

第一组线性变换对于第一组的每一个基有:

$$\begin{cases} T(v_1) = a_{11}v_1 + a_{12}v_2 + \dots + a_{1n}v_n \\ T(v_2) = a_{21}v_1 + a_{22}v_2 + \dots + a_{2n}v_n \\ \dots \\ T(v_n) = a_{n1}v_1 + a_{n2}v_2 + \dots + a_{nn}v_n \end{cases} \Rightarrow A = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

, 对于第二组基, 我们同样可以得到:

$$\begin{cases} T(w_1) = c_{11}w_1 + c_{12}w_2 + \dots + c_{1n}w_n \\ T(w_2) = c_{21}w_1 + c_{22}w_2 + \dots + c_{2n}w_n \\ \dots \\ T(w_n) = c_{n1}w_1 + c_{n2}w_2 + \dots + c_{nn}w_n \end{cases} \Rightarrow B = \begin{bmatrix} c_{11} & c_{21} & \dots & c_{n1} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix}$$

我们进一步假设 v_1, v_2, \dots, v_n 是特征向量时, 我们有 $T(v_i) = \lambda_i v_i \Rightarrow A = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$

Chapter 33

左右逆和伪逆

m*n 型矩阵;rank=r

1.(1) 当矩阵可逆时: $AA^{-1} = A^{-1}A = I \Rightarrow m = n = r$

1.(2) 矩阵列满秩: $n = r, n < m; \dim(N(A)) = m - r$; 矩阵 $A^T A$ 是 $n * n$ 型对称可逆矩阵, $\text{rank}(A^T A) = n$. 我们将 $(A^T A)^{-1} A^T$ 称为矩阵 A 的左逆. 我们将左逆放在右边得到 $A(A^T A)^{-1} A^T \Rightarrow P$ 是矩阵列空间的投影矩阵.

1.(3) 矩阵行满秩: $m = r, m < n; N(A) = \{0\}$. 矩阵 AA^T 是 $m * m$ 型对称可逆矩阵, $\text{rank}(AA^T) = m$. 我们将 $A^T(AA^T)^{-1}$ 称为矩阵 A 的右逆. 我们将右逆放在左边得到 $A^T(AA^T)^{-1} A^T \Rightarrow P$ 是矩阵 A 行空间的投影矩阵

2. 矩阵满足 $r < m; r < n$, 我们根据 1.(2) 和 1.(3) 知道矩阵的 $C(A)$ 、 $R(A)$ 都是 r 维空间, 且两个空间中的向量一一对应, 对于 $\forall x \in R(A), Ay \in C(A); \forall y \in R(A), Ay \in C(A); Ax \neq Ay$. 既然两个空间一一对应, 那么我们也有 $\forall Ax \in C(A), y \in R(A); \forall Ay \in C(A), y \in R(A) \Rightarrow x = A^+(Ax)$, 我们称 A^+ 为矩阵 A 的伪逆.

伪逆

我们从奇异值分解开始: $A = U \Sigma V^T, \Sigma = \text{diag} \{\sigma_1, \sigma_2, \dots, \sigma_r, \dots, 0\}, \Sigma^+ \Sigma = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, (n * n, \text{rank}(\Sigma) = r)$.

首先 $\Sigma^+ = \text{diag} \left\{ \frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_r}, \dots, 0 \right\}$, 那么我们能够得出:

$$A^+ = (U \Sigma V^T)^+ \Rightarrow A^+ = V \Sigma^+ U^T$$

这是为了解决最小二乘法中 AA^T 不可逆的情况, 使用伪逆来解决这一问题.