

3.8 左右逆与伪逆

这讲内容我们就来系统的讲讲逆矩阵！

满秩逆矩阵

这个矩阵的数学表达是 $A^{-1}A = I = AA^{-1}$ 。首先这个矩阵一定是满秩的，不满秩不可逆！其次左乘和右乘相等，那么A一定是方阵。所以我们得到 $r = m = n$ 。所以A一定得是满秩的方阵！

这个也是我们常说的逆矩阵，如果不做特殊说明，我们说的逆矩阵都是这个情况！

左逆矩阵（左逆矩阵出现的情况必须满足矩阵列满秩！）

若 A 列满秩（full column rank），即 $r = n$ ，则 A 的列线性无关，其零空间仅含零向量。方程 $Ax = b$ 要么恰有一解，要么无解。

此时 $A^T A$ 是一个 $n \times n$ 的可逆对称矩阵，因此

$$(A^T A)^{-1} A^T A = I.$$

我们称

$$A_{\text{left}}^{-1} = (A^T A)^{-1} A^T$$

为 A 的一个左逆（可能还有其他左逆，但这是最常用的一种）。我们一般求一个矩阵的左逆就是套用这个公式！

注意： AA_{left}^{-1} 是一个 $m \times m$ 矩阵，仅当 $m = n$ 时才等于单位阵 I 。

矩形矩阵不可能有双侧逆，因为矩阵或其转置必有非零零空间。或者根据矩阵乘法也可以得出这时不可能的！

而且m是会小于n才会出现左逆，因为如果n大于m，这个时候秩最大只能等于m，这个时候无法列满秩，于是没有左逆出现了！

右逆矩阵(必须行满秩)

若 A 行满秩（full row rank），即 $r = m$ ，则 A^T 的零空间仅含零向量， A 的行线性无关。此时方程 $Ax = b$ 总有解； A 的零空间维数为 $n - m$ ，故若 $n > m$ 则有无穷多解。（这个不多解释，第一单元有讨论过！）

行满秩矩阵存在右逆 A_{right}^{-1} ，满足

$$AA_{\text{right}}^{-1} = I.$$

显然的是我们有：

$$AA^T(AA^T)^{-1} = I$$

最简洁的一个右逆是

$$A_{\text{right}}^{-1} = A^T(AA^T)^{-1}.$$

验证：

$$A[A^T(AA^T)^{-1}] = I.$$

大家来想想如果把左右逆矩阵调转位置会发生什么：

- 如果 A **列满秩**，则

$$AA_{\text{left}}^{-1} = A(A^T A)^{-1} A^T$$

是把 \mathbb{R}^m 投影到 A 的列空间的投影矩阵 P 。

- 类似地，若 A **行满秩**，则

$$A_{\text{right}}^{-1} A = A^T(AA^T)^{-1} A$$

是把 \mathbb{R}^n 投影到 A 的行空间的投影矩阵。

伪逆 (Pseudoinverse)

- 可逆矩阵 ($r = m = n$) 的零空间与左零空间均只有零向量。
- 列满秩矩阵 ($r = n$) 的零空间只有零向量。
- 行满秩矩阵 ($r = m$) 的左零空间只有零向量。

剩下需考虑的是 $r < n$ 且 $r < m$ 的情形。

非平凡零空间使得求逆变得困难。若存在非零 x 使 $Ax = 0$ ，则不可能有矩阵 A^{-1} 满足 $A^{-1}0 = x$ 。向量 Ax 总在 A 的列空间中。事实上，行空间 (r 维) 与列空间 (r 维) 之间的映射

$$x \mapsto Ax$$

是**一一对应**的。

证明：若 $x \neq y$ ，则 $Ax \neq Ay$

反设不成立，则存在 $x \neq y$ 且 x, y 均位于 A 的行空间，使得 $Ax = Ay$ 。

于是

$$A(x - y) = 0,$$

即 $x - y$ 属于 A 的零空间；又 $x - y$ 属于行空间（行空间对线性组合封闭）。

但零空间与行空间唯一的公共向量为零向量，故

$$x - y = 0,$$

与 $x \neq y$ 矛盾。

因此映射 $x \mapsto Ax$ 从行空间到列空间是可逆的，其逆称为**伪逆**（就是从 Ax 到 x ），在统计学中的线性回归里非常有用——即便矩阵未必列满秩 $r = n$ 。伪逆记作： A^+

其实伪逆每一个矩阵都有且是唯一的。当然正常的逆只有满秩方阵有，左逆和右逆都需要满足不同的条件。而伪逆就是为了在**当矩阵不可逆时，我们能找到的“最好、最接近逆矩阵”的东西**。

关于上面的证明其实有一个更加几何的视角去看待！对于一个矩阵而言，他会有列空间，行空间，零空间，左零空间，他们的关系我们之前讨论过，这里就不赘述！然而我们上面证明了，只要是一个矩阵，那么这个矩阵就充当着一个列空间向量与行空间向量之间的映射关系！当然这个逆是可以的，也正是我们追求的。但是如果是零空间与左零空间的映射，操作后得到 0，那么就无法逆回去了。这会造成信息的丢失！所以对于有逆的矩阵（逆，左逆，右逆都算），他们的零空间和左零空间都是只有零向量的！这就造成了这样的矩阵不存在不可逆的情况，但是其他矩阵就会存在！

那么伪逆 A^+ 满足：对所有位于行空间的 x ，有

$$x = A^+ Ax.$$

A^+ 的零空间等于 A 的左零空间。（就是说 A^+ 会把所有左零空间里的向量直接压成 0，而对其他向量则保留信息。）从计算上来说，就是让独立的列向量结果变成单位向量！

下面我们来看看怎么求伪逆吧：

求伪逆

基于 SVD 的构造

我们介绍 **SVD 奇异值分解** 方法。即有一 $m \times n$ 矩阵 A ，我们求解其伪逆时，首先将其分解：

$$A = U \Sigma V^T$$

其中 U, V^T 为正交矩阵； Σ 为一个对角矩阵，其中的元素为：

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r & & \\ & & & & \ddots & \end{bmatrix}_{(m \times n)}, \quad \sigma_i \text{ 为奇异值, } r \text{ 为秩}$$

Σ 的**伪逆**为 Σ^+ ：

$$\Sigma^+ = \begin{bmatrix} 1/\sigma_1 & & & \\ & 1/\sigma_2 & & \\ & & \ddots & \\ & & & 1/\sigma_r \\ & & & & \ddots \end{bmatrix}_{(n \times m)}, \quad r \text{ 为秩}$$

3.8 左右逆与伪逆

上面的求解应该不困难，就是要把有值的元素全部转化为1!

计算得到的 $\Sigma\Sigma^+$ 与 $\Sigma^+\Sigma$ 不同，一个是 $n \times n$ 矩阵，另一个是 $m \times m$ 的，即为 Σ 在 **行空间**、**列空间** 上的投影矩阵。这也表现了伪逆将我们代入两个很好的空间：行空间与列空间。

介绍完了 Σ 的伪逆求法，那么 A 本身的伪逆就好求了。 U, V^T 为正交矩阵，都可逆，直接求逆，得到：

$$A^+ = V\Sigma^+U^T$$

很明显，**SVD** 的特殊之处就在于将一切问题归结于对角矩阵上；在对角矩阵上，很多东西会变得明显，然计算简便许多！

讨论课

这节讨论课，在计算的范式上，没有习题课讲的好，所以大家参考习题课的题目足矣！而值得说的是在讨论课中说到了当 x 在不同空间中， A^+ 的作用是什么！首先，我们在上文说到过，我们把矩阵 A 看做是一个线性变换，而矩阵 A 有逆时，无论什么操作，逆矩阵都可以把这个变换撤销！但是如果不存在逆，我们就只能用伪逆，而伪逆他撤销变换是有范围的！如果 A 乘以 x 的结果在矩阵 A 的列空间中，那么伪逆可以为其撤销，但是如果 x 是在矩阵 A 的零空间中，那么这个就无法撤销！这个很值得大家理解！

习题课

问题1

求矩阵

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

的一个右逆。

解答：我们使用公式

$$A_{\text{right}}^{-1} = A^T(AA^T)^{-1}.$$

计算：

$$\bullet A^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

- $AA^T = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$
- $(AA^T)^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}$

于是

$$A_{\text{right}}^{-1} = A^T(AA^T)^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \\ 1/2 & 0 \end{bmatrix}.$$

我们可以快速验证： $AA_{\text{right}}^{-1} = I$ 。

问题2

矩阵

$$A = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$$

是否有左逆？右逆？伪逆？若答案为“是”，请给出相应的逆矩阵。

解答：

A 的第二行是第一行的倍数，因此矩阵的秩为 1，且 $\det A = 0$ 。由于 A 是方阵，行列式存在，我们可以使用性质 $\det(C)\det(D) = \det(CD)$ 来证明 A 不可能有左逆或右逆。（若 $AB = I$ ，则 $\det A \det B = \det I$ 推出 $0 = 1$ ，矛盾。）

我们可以求出其伪逆 $A^+ = V\Sigma^+U^T$ 。为此，先求 A 的奇异值分解 $U\Sigma V^T$ 。

在关于奇异值分解的讲座中已经算出了 A 的 SVD：

$$\underbrace{\begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}}_A = \underbrace{\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}}_U \underbrace{\begin{bmatrix} \sqrt{125} & 0 \\ 0 & 0 \end{bmatrix}}_\Sigma \underbrace{\begin{bmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{bmatrix}}_{V^T}.$$

于是

$$\Sigma^+ = \begin{bmatrix} 1/\sqrt{125} & 0 \\ 0 & 0 \end{bmatrix},$$

计算伪逆：

$$\begin{aligned} A^+ &= V\Sigma^+U^T \\ &= \begin{bmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{bmatrix} \begin{bmatrix} 1/\sqrt{125} & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \\ &= \frac{1}{125} \begin{bmatrix} 4 & 8 \\ 3 & 6 \end{bmatrix}. \end{aligned}$$

为验证结果，我们确认 A^+ 在 A 的行空间上逆转了 A 的作用（使用计算 SVD 时得到的基）。回忆：

- 若 $j \leq r$ ，则 $Av_j = \sigma_j u_j$ ；
- 若 $j > r$ ，则 $Av_j = 0$ 。

这里

- $u_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$,
- $A^+ u_1 = \frac{1}{\sqrt{125}} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \frac{1}{\sigma_1} v_1$ 。

同样可验证

- $A^+ u_2 = \frac{1}{125} \begin{bmatrix} 4 & 8 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ 。