Chapter32左逆，右逆，伪逆

Section1.引入

我们知道，在$m \times n$矩阵A满足$m = n = rank(A)$，也就是满秩方阵时，我们容易得到$A^{-1}A = I = AA^{-1}$，这个即是我们说到的逆。

我们同时也在Chapter08四个基本子空间，我们会知道:

- 列空间$C(A) \in \mathbb{R}^m$，$\dim C(A) = r$，左零空间$N\left(A^T\right) \in \mathbb{R}^m$，$\dim N\left(A^T\right) = m - r$，列空间与左零空间互为正交补；
- 行空间$C\left(A^T\right) \in \mathbb{R}^n$，$\dim C\left(A^T\right) = r$，零空间$N(A) \in \mathbb{R}^n$，$\dim N(A) = n - r$，行空间与零空间互为正交补。

因此我们自然而然会出现零空间与左零空间不全是0向量时候的情况，即$m = r < n$或者$n = r < m$或$r < m, r < m$时候的情形。

Section2.左逆

我们考虑到A列满秩时候的情形，那么我们会发现A^TA是满秩的，也就是会成立

$$\underbrace{\left(A^TA\right)^{-1}A^T}A = I$$

我们把大括号内的内容考虑为长方形矩阵A的左逆：

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

顺便复习一下最小二乘一讲，我们解$Ax = b$，A_{left}^{-1}被当做一个系数矩阵乘在b向量上，求得b向量投影在A的列空间之后的解$\hat{x} = \left(A^TA\right)^{-1}A^Tb$。如果我们强行给左逆左乘矩阵$A$，得到的矩阵就是投影矩阵$P = A\left(A^TA\right)^{-1}A^T$，来自$p = A\hat{x} = A\left(A^TA\right)^{-1}A^T$，它将右乘的向量$b$投影在矩阵$A$的列空间中。

再来观察AA^T矩阵，这是一个$m \times m$矩阵，秩为$rank(AA^T) = n < m$，也就是说AA^T是不可逆的，那么接下来我们看看右逆。

Section3.右逆

其实是类似的，我们知道

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

当然这是A是列满秩的时候才成立的

此时如果A是行满秩的时候，那么A^T就是列满秩了，那么

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

就会有

$$\left(AA^T\right)^{-1}AA^T=I\overset{\text{取转置}}{\rightarrow}A\,\underbrace{A^T(AA^T)^{-1}}=I$$

可以知道

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

同样的，如果我们强行给右逆右乘矩阵A，将得到另一个投影矩阵$P=A^T\left(AA^T\right)A$，与上一个投影矩阵不同的是，这个矩阵的A全部变为A^T了。所以这是一个能够将右乘的向量b投影在A的行空间中。

Section4.伪逆

前面我们提及了逆（方阵满秩），并讨论了左逆（矩阵列满秩）、右逆（矩阵行满秩），现在看一下第四种情况，$m\times n$矩阵A不满秩的情况。

现在任取一个向量x，乘上A后结果Ax一定落在矩阵A的列空间$C(A)$中。而根据维数，$x\in\mathbb{R}^n$，$Ax\in\mathbb{R}^m$，那么我们现在猜测，输入向量x全部来自矩阵的行空间，而输出向量Ax全部来自矩阵的列空间，并且是一一对应的关系，也就是$\mathbb{R}^n$的r维子空间到$\mathbb{R}^m$的r维子空间的映射。

而矩阵A现在有这些零空间存在，其作用是将某些向量变为零向量，这样$\mathbb{R}^n$空间的所有向量都包含在行空间与零空间中，所有向量都能由行空间的分量和零空间的分量构成，变换将零空间的分量消除。但如果我们只看行空间中的向量，那就全部变换到列空间中了。

那么，我们现在只看行空间与列空间，在行空间中任取两个向量x，$y\in C(A^T)$，则有$Ax\neq Ay$。所以从行空间到列空间，变换A是个不错的映射，如果限制在这两个空间上，A可以说“是个可逆矩阵”，那么它的逆就称作伪逆，而这个伪逆的作用就是将列空间的向量一一映射到行空间中。通常，伪逆记作A^+，因此$Ax=(Ax)$，$y=A^+(Ay)$。

现在我们来证明对于$x,y\in C\left(A^T\right)$，$x\neq y$，有$Ax,Ay\in C(A)$，$Ax\neq Ay$：

- 反证法，设$Ax = Ay$，则有$A(x-y)=0$，即向量$x-y \in N(A)$；
- 另一方面，向量$x, y \in C\left(A^T\right)$，所以两者之差$x-y$向量也在$C\left(A^T\right)$中，即$x-y \in C\left(A^T\right)$；
- 此时满足这两个结论要求的仅有一个向量，即零向量同时属于这两个正交的向量空间，从而得到$x=y$，与题设中的条件矛盾，得证。

伪逆在统计学中非常有用，以前我们做最小二乘需要矩阵列满秩这一条件，只有矩阵列满秩才能保证A^TA是可逆矩阵，而统计中经常出现重复测试，会导致列向量线性相关，在这种情况下A^TA就成了奇异矩阵，这时候就需要伪逆。

接下来我们介绍如何计算伪逆A^+：

其中一种方法是使用奇异值分解，$A=U\Sigma V^T$，其中的对角矩阵型为
$\Sigma=\left[\begin{array}{ccc|c}\sigma_1 & & & \\ & \ddots & & \\ & & \sigma_2 & \\ \hline & & & [0]\end{array}\right]$，对角线非零的部分来自$A^TA$，$AA^T$比较好的部分，剩下的来自左/零空间。

我们先来看一下Σ矩阵的伪逆是多少，这是一个$m \times n$矩阵，$rank(\Sigma)=r$，
$\Sigma^+=\left[\begin{array}{ccc|c}\frac{1}{\sigma_1} & & & \\ & \ddots & & \\ & & \frac{1}{\sigma_r} & \\ \hline & & & [0]\end{array}\right]$，伪逆与原矩阵有个小区别：这是一个$n \times m$矩阵。则有
$\Sigma\Sigma^+=\left[\begin{array}{ccc|c}1 & & & \\ & \ddots & & \\ & & 1 & \\ \hline & & & [0]\end{array}\right]_{m\times m}$，$\Sigma^+\Sigma=\left[\begin{array}{ccc|c}1 & & & \\ & \ddots & & \\ & & 1 & \\ \hline & & & [0]\end{array}\right]_{n\times n}$。

观察$\Sigma\Sigma^+$和$\Sigma^+\Sigma$不难发现，$\Sigma\Sigma^+$是将向量投影到列空间上的投影矩阵，而$\Sigma^+\Sigma$是将向量投影到行空间上的投影矩阵。我们不论是左乘还是右乘伪逆，得到的不是单位矩阵，而是投影矩阵，该投影将向量带入比较好的空间（行空间和列空间，而不是左/零空间）。

接下来就可以求A的伪逆：

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

注：
对于考研来说，由于时间原因我没有仔细去学习伪逆这一章，直接选用的mit-18.06-linalg-notes/docs/chapter34.md at master · apachecn/mit-18.06-linalg-notes (github.com)

如果想要深入理解，建议看MIT18.06 跟男神教授学线性代数 - 知乎 (zhihu.com)

这个也很不错MIT线性代数18.06-学习笔记 - 知乎 (zhihu.com)

以下是其中难得的latex代码：

接下来我们讨论最一般的情况： r<m,n，即矩阵的秩小于矩阵的行和列，此时零空间N(A) 和左零空间 N(A^{T}) 都不为零， \vec{x} 和 \vec{b} 都不可能被还原。但是，行空间 \vec{x_{r}} 和列空间 \vec{p} 之间是存在一一对应的，矩阵的逆和它的伪逆就是在这两个空间之间的变换和逆变换， A 的伪逆记为 A^{+} 。

再具体的说说我们对伪逆的期待。 \mathbb{R}^{n} 中任一矢量 \vec{x} 经过 A\vec{x} 变为列空间中的矢量 \vec{p} ，我们希望经过 A^{+} 逆变换为 A^{+}\vec{p}=A^{+}A\vec{x}=\vec{x*{r}} ， \vec{x} 并没有也不可能被恢复，乘 A 矩阵时已经抹杀了零空间的信息 \vec{x*{n}} ，我们最终希望得到的是 \vec{x} 在行空间的投影 \vec{x*{r}} 。另一方面，\mathbb{R}^{m} 中任一矢量 \vec{b} 经过 A^{+}\vec{b} 变为行空间中的矢量 \vec{x*{r}} ，再经过 A 变换为 A\vec{x_{r}}=AA^{+}\vec{b}=\vec{p} ， \vec{b} 也不可能被恢复，乘 A^{+} 的时候也已经抹杀了左零空间中的 \vec{e} ，我们最终得到的也只是 \vec{b} 在列空间的投影 \vec{p} 。换言之，我们希望 A^{+}A **是投影到** A **行空间的投影矩阵，**AA^{+} **是投影到** A **列空间的投影矩阵。** A 是 m\times n 矩阵，则 A^{+} 是 n\times m 矩阵。

一个找到伪逆的方法是通过 SVD : A=U\Sigma V^{T} ，我们选择 A 的伪逆为 A^{+}=V\Sigma^{+}U^{T} ，其中 \Sigma^{+}=\begin{bmatrix}\frac{1}{\sigma*{1}}&&\&\frac{1}{\sigma*{2}}&\&&...&\&&&\frac{1}{\sigma*{r}}&0&\0&0&..&&0\0&0&..&&0\end{bmatrix}，它是一个 n\times m 对角阵，正是因为 \Sigma 不可逆，我们才只能求伪逆。 \Sigma^{+} 正是 \Sigma 的伪逆， \Sigma^{+}\Sigma=\begin{bmatrix}I*{r\times r}&0\0&0\end{bmatrix}*{n\times n},\Sigma\Sigma^{+}=\begin{bmatrix}I*{r\times r}&0\0&0\end{bmatrix}_{m\times m} 。 A^{+} 的列空间和左零空间就是 A 的行空间和零空间， A^{+} 的行空间和零空间就是 A 的列空间和左零空间， A 和 A^{+} 的秩都是 r 。容易验证，**当矩阵可逆时** A^{+}=A^{-1}。

现在我们来验证：

1. **A^{+}A 是投影到** A **行空间的投影矩阵。** A^{+}A=V\Sigma^{+}U^{T}U\Sigma V^{T}=V\Sigma^{+}\Sigma V^{T} =\begin{bmatrix}|&|&&|&|&&|\v*{1}&v*{2}&.&v*{r}&\color{red}

{v{r+1}}&\color{red}.&\color{red}{v{n}}\|&|&&|&|&&|\end{bmatrix}\begin{bmatrix}I{r\times r}&0\0&0\end{bmatrix}\begin{bmatrix}|&|&&|&|&&|\v{1}&v{2}&.&v{r}&\color{red}{v{r+1}}&\color{red}.&\color{red}{v{n}}\|&|&&|&|&&|\end{bmatrix}^{T}= \begin{bmatrix}|&|&&|\v{1}&v{2}&.&v{r}\|&|&&|\end{bmatrix}\begin{bmatrix}|&|&&|\v{1}&v{2}&.&v_{r}\|&|&&|\end{bmatrix}^{T} ，这正是 A 行空间的投影矩阵。

2. AA^{+} **是投影到 A 列空间的投影矩阵。** AA^{+}=U\Sigma V^{T}V\Sigma^{+}U^{T}=U\Sigma \Sigma^{+} U^{T}=\begin{bmatrix}|&|&&|&|&&|\u{1}&u{2}&.&u{r}&\color{red}{u{r+1}}&\color{red}.&\color{red}{u{m}}\|&|&&|&|&&|\end{bmatrix}\begin{bmatrix}I{r\times r}&0\0&0\end{bmatrix}\begin{bmatrix}|&|&&|&|&&|\u{1}&u{2}&.&u{r}&\color{red}{u{r+1}}&\color{red}.&\color{red}{u_{m}}\|&|&&|&|&&|\end{bmatrix}^{T}

=\begin{bmatrix}|&|&&|\u{1}&u{2}&.&u{r}\|&|&&|\end{bmatrix}\begin{bmatrix}|&|&&|\u{1}&u{2}&.&u{r}\|&|&&|\end{bmatrix}^{T} ，这正是 A 列空间的投影矩阵。

最后，我们可以验证A^{+}=V\Sigma^{+}U^{T} 和上述 A^{-1}{left}，A^{-1}{right} 的表达式是一致的：

- 列满秩时， \Sigma=\begin{bmatrix}{\sigma{1}}&&\&{\sigma{2}}&\&&...&\&&&{\sigma{n}}\0&0&..&0\0&0&..&0\end{bmatrix}，\Sigma^{T}\Sigma=\begin{bmatrix}{\sigma{1}}&&&&0&0\&{\sigma{2}}&&&0&0\&&...&\&&&{\sigma{n}}&0&0\end{bmatrix}\begin{bmatrix}{\sigma{1}}&&\&{\sigma{2}}&\&&...&\&&&{\sigma{n}}\0&0&..&0\0&0&..&0\end{bmatrix} =\begin{bmatrix}{\sigma{1}^{2}}&&\&{\sigma{2}^{2}}&\&&...&\&&&{\sigma{n}^{2}}\end{bmatrix} 是个可逆矩阵，因此A^{-1}{left}=(A^{T}A)^{-1}A^{T}=((U\Sigma V^{T})^{T}U\Sigma V^{T})^{-1}(U\Sigma V^{T})^{T}=(V\Sigma^{T}\Sigma V^{T})^{-1}V\Sigma^{T}U^{T} =V(\Sigma^{T}\Sigma)^{-1}\Sigma^{T}U^{T}=V\begin{bmatrix}\frac{1}{\sigma{1}^{2}}&&\&\frac{1}{\sigma{2}^{2}}&\&&...&\&&&\frac{1}{\sigma{n}^{2}}\end{bmatrix}\begin{bmatrix}{\sigma{1}}&&&&0&0\&{\sigma{2}}&&&0&0\&&...&\&&&{\sigma_{n}}&0&0\end{bmatrix}U^{T}

=V\Sigma^{+}U^{T} 。

- 行满秩时， \Sigma=\begin{bmatrix}{\sigma{1}}&&&&0&0\&{\sigma{2}}&&&0&0\&&...&\&&&{\sigma{m}}&0&0\end{bmatrix} ，\Sigma\Sigma^{T}=\begin{bmatrix}{\sigma{1}}&&&&0&0\&{\sigma{2}}&&&0&0\&&...&\&&&{\sigma{m}}&0&0\end{bmatrix}\begin{bmatrix}{\sigma{1}}&&\&{\sigma{2}}&\&&...&\&&&

{\sigma{m}}\0&0&..&0\0&0&..&0\end{bmatrix}=\begin{bmatrix}{\sigma{1}^{2}}&&\&
{\sigma{2}^{2}}&\&&...&\&&&{\sigma{m}^{2}}\end{bmatrix} 也可逆，A^{-1}{right}=A^{T}
(AA^{T})^{-1}=V\Sigma^{T} U^{T}(U\Sigma\Sigma^{T}U^{T})^{-1}=V\Sigma^{T}
(\Sigma\Sigma^{T})^{-1}U^{T}= V\begin{bmatrix}{\sigma{1}}&&\&
{\sigma{2}}&\&&...&\&&&
{\sigma{m}}\0&0&..&0\0&0&..&0\end{bmatrix}\begin{bmatrix}\frac{1}
{\sigma{1}^{2}}&&\&\frac{1}{\sigma{2}^{2}}&\&&...&\&&&\frac{1}
{\sigma_{m}^{2}}\end{bmatrix}U^{T}=V\Sigma^{+}U^{T}

*最后唠叨一句关于“ A 与 A^{+} 四个基本空间的对应关系”。其实 A^{T} 与 A 也存在同样的对应关系：A^{T} 的列空间和左零空间就是 A 的行空间和零空间， A^{T} 的行空间和零空间就是 A 的列空间和左零空间。区别仅仅在于 \Sigma^{+} 和 \Sigma^{T}， A^{T} 包含的 \Sigma^{T} 和 A 中 \Sigma 的拉伸作用是不能互相抵消的， A^{+} 特地选择的 \Sigma^{+} 才可以抵消。

*最后再唠叨一句：微积分基本定理是说积分 T^{+}(f)=\int_{0}^{x}f(t)dt 是微分 T(f)=\frac{df}{dx} 的伪逆。