Lecture 33

一、矩阵的逆

之前在计算矩阵的逆都是针对方阵，并且矩阵是满秩，则矩阵 A 存在A^{-1}满足$AA^{-1}=\mathbf{I}$，这种情况下，左逆与右逆是相等的

二、左逆

当矩阵 A 不是方阵列满秩时，即 rank(A)=n<m，零空间只有零解，Ax=b 有一个解(b 在 A 的列空间)或者无解。在最小二乘法中，我们利用A^TA，来求解最优解。当矩阵 A 是列满秩的情况时间，A^TA是一个 n×n 的可逆矩阵，即满足$(A^TA)^{-1}A^TA=\mathbf{I}$，其中$(A^TA)^{-1}A^T$记为矩阵 A 的左逆：

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

用矩阵 A 左乘$A_{left}{}^{-1}$，得到$A(A^TA)^{-1}A^T$，可以看出这个是一个投影到矩阵 A 的列空间的投影矩阵。

三、右逆

当矩阵 A 不是方阵行满秩时，即 rank(A)=m<n，左零空间只有零解，Ax=b 将总是有解集，自由变量的个数是 n-r。

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

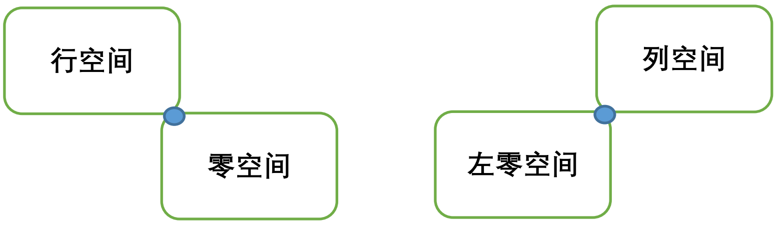
用矩阵 A 右乘$A_{left}{}^{-1}$，得到$A(A^TA)^{-1}A^T$，可以看出这个是一个投影到矩阵 A 的列空间的投影矩阵。当矩阵 A 是行满秩的情况时间，AA^T是一个 m×m 的可逆矩阵，即满足$AA^T(AA^T)^{-1}=\mathbf{I}$，其中$A^T(AA^T)^{-1}$记为矩阵 A 的右逆：

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

用矩阵 A 右乘$A_{right}{}^{-1}$，得到$A^T(AA^T)^{-1}A$，可以看出这个是一个投影到矩阵 A 的行空间的投影矩阵。

四、伪逆

回顾之前的四个子空间的关系，有如下的图：



给定任意一个向量 x，Ax 一定在矩阵 A 的列空间中，因为 Ax 的结果是矩阵的各列根据 x 的系数进行线性组合。根据矩阵乘法，x∈R^n，Ax∈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R^n的 r 维子空间到R^m的 r 维子空间的映射。可以认为矩阵 A 的变换将两个零空间全部消除了。

现在只看行空间和列空间，任取不相等的两个向量 x、y 属于 C(A)，则有 Ax≠Ay，从行空间到列空间，A 是一个很好的映射，如果只限制在这两个矩阵上，可以认为 A 是一个“可逆矩阵”，伪逆的作用就是从列空间映射到行空间，记做A^+。

下面证明对于不相等的两个向量 x、y 属于 C(A)，则有 Ax≠Ay。

反证法，设 Ax=Ay，有 A(x-y)=0，所以向量 x-y 属于零空间，又因为 x、y 属于行空间，则 x-y 属于行空间，x-y 不可能属于两个正交的空间，所以 x-y 只能是零向量，与 x、y 是两个不同的向量矛盾。

计算A^+

其中一种方法是通过奇异值分解的方法，$\mathbf{A}=\mathbf{U\Sigma V}^T$，其中对角阵

$$\mathbf{\Sigma}=\begin{bmatrix}\sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & [0]\end{bmatrix}$$

这是一个 m×n 的矩阵，rank(Σ)=r

$$\mathbf{\Sigma}^+=\begin{bmatrix}\frac{1}{\sigma_1} & & & \\ & \ddots & & \\ & & \frac{1}{\sigma_r} & \\ & & & [0]\end{bmatrix}$$

这是一个 n×m 的矩阵。则有：

$$\mathbf{\Sigma\Sigma}^+=\begin{bmatrix}1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & [0]\end{bmatrix}_{m\times m}$$

$$\mathbf{\Sigma}^+\mathbf{\Sigma}=\begin{bmatrix}1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & [0]\end{bmatrix}_{n\times n}$$

则$A^+=V\Sigma^+U^T$。