Chapter15正交矩阵与Gram-Schmidt正交化

Section1.正交矩阵的概念

Subsection1.简单的回顾

在标准正交向量组我们介绍了标准正交向量的概念，其实我们需要关注的就是**正交**和**标准**的概念，这些向量所组成的向量组我们可以写为如下的特性:

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

标准正交向量组又被称为标准正交基,显然,相互垂直的各列一定是线性无关的。我们关注标准正交向量的意义很明显，它们很规范，很便于操作。

Subsection2.正交矩阵

Subsection1.定义

标准正交矩阵Q,就是将标准正交向量组中的向量$q_1, q_2, \cdots, q_n$列到一个矩阵中去:

$$Q = [q_1, q_2, \cdots, q_n]$$

我们注意一个非常重要的性质:

$$Q^TQ = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I$$

一定要注意，是$Q^TQ = I$而不是$QQ^T = I$，原因参见下方概念辨析我们注意我们往往看待一个向量是看待为列向量。

特别的，当这种矩阵为方阵的时候，我们称之为正交矩阵。

注意概念辨析:

- 满足$Q^TQ = I$的是标准正交矩阵，只不过是简单的把标准正交向量组的列向量陈列在一个矩阵里面，**它不一定满足**$QQ^T = I$,例如:

$$Q=\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}，则Q^TQ=\begin{bmatrix}1&0\\0&1\end{bmatrix}=I，而QQ^T=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}\neq I$$

因为我们关注的是列向量的性质，即$q_1^Tq_1=1$，如果QQ^T则就不一定了。

- 正交矩阵是方阵的标准正交矩阵，因此它有性质$Q^TQ=I$之外，还有性质$QQ^T=I$.
原因我想是因为此时矩阵满足$rank(Q)=rank(C(A))=n$，因此行也是满秩的，那么转置过来成为列也是线性无关的。

Subsection2.性质

我们根据定义能够注意到正交矩阵有一些重要的性质:

- 拥有标准正交向量组的性质，即向量之间正交，向量的模为1
- 必须是方阵，定义决定
- 因为是方阵，说明A^T是A的逆矩阵
- 是对称阵

Subsection3.例子

- 回想我们在很久之前提到过的置换矩阵,P,当时也说明了置换矩阵P具有性质:$P^T=P^{-1}$,
而显然置换矩阵正是一个正交矩阵:$P=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}$,则 $Q^T=\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}$,易得 $Q^TQ=I$
。
- 另一个例子是我们上一讲介绍过的标准正交向量组: $\begin{bmatrix}\cos\theta\\ sin\theta\end{bmatrix}$, $\begin{bmatrix}-\sin\theta\\ cos\theta\end{bmatrix}$。写成矩阵形式即为 $Q=\begin{bmatrix}\cos\theta&-\sin\theta\\ \sin\theta&\cos\theta\end{bmatrix}$,显然该矩阵也有 $Q^TQ=I$。
- $\begin{bmatrix}1&1\\1&-1\end{bmatrix}$ 是正交矩阵吗?不是,因为虽然这个矩阵内各列正交,但列向量长度不为1,我们还需要再对其进行单位化(标准化),单位化以后得到 $Q=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$,这个矩阵是正交矩阵。
- 使用上一个例子中的矩阵 $Q=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$,令 $Q'=c\begin{bmatrix}Q&Q\\Q&-Q\end{bmatrix}$,取合适的$c$以使得向量长度为1也可以构造出一个正交矩阵:$Q=\frac{1}{2}\begin{bmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{bmatrix}$,这种构造方法以阿德玛(Adhemar)命名,阿德玛矩阵是一种只有1和-1的正交矩阵,我们现在知道的是这种构造方法在Q为2,4,16,64,...维矩阵时是有效的,但是没有人知道,究竟哪些维数的正交矩阵可以

由1和-1们构成,有些维度可以,但有些维度就不行,比如说5维的矩阵就不可能是阿德玛矩阵,这个比较简单,但有些大小没人能确切地说,能不能由1和-1构成。

Subsection4.用处

我们在考虑投影矩阵的时候，注意$A^TAx=A^Tb$，则$\hat{x}=(A^TA)^{-1}A^Tb$,则$A\hat{x}=A(A^TA)^{-1}A^Tb=\hat{b}$,那么投影矩阵就是$P=A(A^TA)^{-1}A^T$,此时如果$A$是标准正交矩阵，则$(A^TA)^{-1}=I$，那么$P=AA^T$，如果$A$是正交矩阵，那么$P=I$.

这也是很好理解的，如果A是一个满秩的方阵，那么A的列空间是n维的满空间，那么投影就是其本身，$P=I$是显然的了

但是我们注意到，如果如果A列不满秩，那么显然A^TA是不存在逆矩阵的，那么投影矩阵$P=A(A^TA)^{-1}A^T$不存在，因此我们要考虑一个方法，把矩阵中的列向量进行标准正交化，这样对于我们进行后面的处理大有脾益。

考虑我们最近讲过的最重要的方程,正规方程$A^TA\hat{x}=A^Tb$,现在变为了$Q^TQ\hat{x}=Q^Tb$,也就是$\hat{x}=Q^Tb$,分解开来看就是$\hat{x}_i=q_i^Tb$,这个式子在很多数学领域都有重要作用。当我们知道标准正交基,则在第i个基方向上的投影等于q_i^Tb。

Section2.Gram-Schmidt正交化

我们已经看到标准正交向量的性质特别好,但更多时候我们见到的是线性无关向量组,有没有一种方法能够将线性无关向量组转换为标准正交基呢?这也即今天要讲的第二个内容,Gram-Schmidt 正交化。

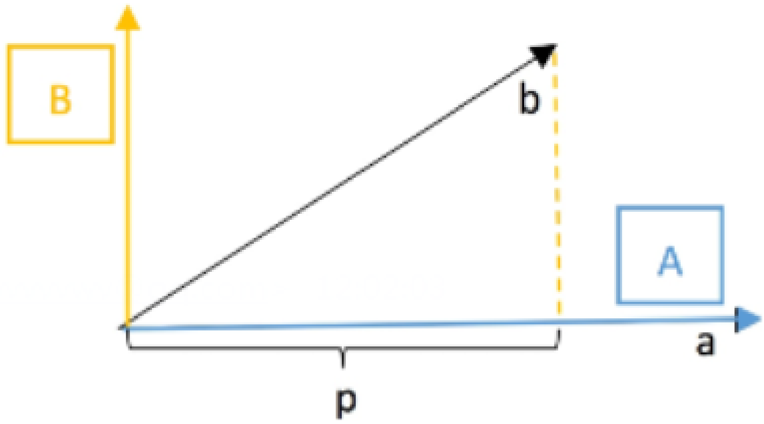
在介绍它之前,我们需要先说明,格拉姆-施密特正交化的缺点在于,由于要求的单位向量,我们不可避免地要除以向量的长度,而这个过程很容易产生根号,所以最终产生的标准正交向量组经常会带有根号。

Gram-Schmidt 正交化的过程如下:

$$\text{线性无关向量}a,b \rightarrow Graham\text{正交化向量}A,B \rightarrow Schmidt\text{标准化正交向量}q_1=\frac{A}{|A|},q_2=\frac{B}{|B|}$$

可以看到Schmidt也即单位化的过程是很简单的,正交化的关键在于找到A,B。

我们先从简单的情况开始,假设有两个线性无关的向量 a,b:



Subsection1.正交化

显然如果我们要考虑与a正交的向量，即我们之前考虑的误差向量e,而且我们有:
b在a上的投影写为$\frac{a^Tb}{a^Ta}b$

$$e = b - p = b - a(a^Ta)^{-1}a^Tb$$

注意到a, b均为向量,那么a^Ta与a^Tb均为标量，因此我们可以写为

$$e = b - \frac{a^Tb}{a^Ta}a$$

那么我们从a, b两个非正交的向量正交化为了两个正交的向量A, B，其中
$A = a, B = e = b - \frac{a^Tb}{a^Ta}a$

验证一下正交性:

$$A^TB = a^T(b - \frac{a^Tb}{a^Ta}a) = a^Tb - a^Tb = 0$$

因此A^T与B是正交的。

Subsection2.单位化

$$q_1 = \frac{A}{|A|}, q_2 = \frac{B}{|B|}$$

Subsection3.三维情况下的正交化与单位化

第三个矢量减去它在前两个矢量构成平面的投影，因此剩下的部分C 肯定也和 A, B 都正交。
这里一定要注意，第三个向量需要在第二个已经正交化之后的向量上投影，即B是已经正交化结束后的向量

$$C = c - \frac{c^T A}{A^T A} A - \frac{c^T B}{B^T B} B$$

接着做标准化就好。高维情况只需要按照同样的思路进行。

Subsection4.例子

对一下三个向量做施密特正交化:

$$a = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad b = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix} \quad c = \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix}$$

基于a来考虑正交向量，那么b在a上的投影为:

$$p_b = \frac{a^T b}{a^T a} a = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

则b与a正交的向量为:

$$b' = b - p_b = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

这里一定要注意，第三个向量需要在第二个已经正交化之后的向量上投影

同理，可以知道c在a和b'上的投影分别为:

$$p_a = \frac{a^T c}{a^T a} a = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}$$

$$p_b = \frac{b'^T c}{b'^T b'} b' = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

那么c与a, b'正交的向量为:

$$c' = c - p_a - p_b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

那么标准化之后我们可以知道三个标准正交向量为:

$$q_1 = \frac{a}{|a|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad q_2 = \frac{b'}{|b'|} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \quad q_3 = \frac{c'}{|c'|} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Section3.$A=QR$分解

我们曾经用矩阵的眼光审视消元法,故有 $A=LU$,这即是消元法的矩阵表示。

以同样的眼光来看待 Gram-Schmidt 正交化,故有 $A=QR$,这既是 Gram-Schmidt 正交化的矩阵表示。

设 A 有两个列向量:$\begin{bmatrix}a_1 & a_2\end{bmatrix}$,标准正交化后有 $\begin{bmatrix}a_1 & a_2\end{bmatrix}=\begin{bmatrix}q_1 & q_2\end{bmatrix}\begin{bmatrix}a_1^Tq_1 & a_1^Tq_2\\ a_2^Tq_1 & a_2^Tq_2\end{bmatrix}$, 而左下角的 $a_1^Tq_2$ 为 0。

$A=QR$ 的重点在于,R 是一个上三角矩阵,这是因为后来构造的向量总是正交于先前的向量。$A=QR$ 是用新基表示旧矢量,R 的列代表旧矢量在新基下的系数, R 是上三角矩阵因为每个旧矢量都和它后面的新基正交。