2.4 正交矩阵和 Gram-Schmidt 正交化

标准正交向量

给标准正交向量一个数学定义：
在一组正交向量中，设 q 是标准正交向量组中的任意向量，则

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

这很好地表现了标准正交向量组内各向量的性质。“标准”→ 长度为 1。

标准正交矩阵 Q

当我们把一组正交向量放到矩阵中得到：

$$Q = \begin{bmatrix} . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ q_1 & . & . & q_n \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \end{bmatrix}$$

这时，结合上面正交向量性质得到：$Q^TQ = I$。
特别的，当矩阵Q是方阵的时候，我们将这样的矩阵 Q 称为：正交矩阵。为什么我们要单独拿出来说，因为方阵有逆矩阵，而逆矩阵有一些特别的性质在。那就是：$Q^T = Q^{-1}$。证明：因为 $Q^TQ = I$！举一些例子：
$Q = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$，计算可以得到的是$Q^T = Q^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$。
然后我们需要注意的是要给矩阵单位化，$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$这个矩阵各列是正交的，但并不是正交矩阵，因为没有单位化，所以，正交矩阵不要忘了单位化向量。正确的正交矩阵是：$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$。由这个矩阵可以延伸出阿达马矩阵，像这样：$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$。单位化后是$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$。他还可以延申到更高维度，这里不做详细介绍。

标准正交矩阵的作用

上面介绍了标准正交矩阵 Q 的各种性质，很显然这是一种新的性质优良的矩阵，接下来主要介绍它的具体应用之一：投影矩阵。标准正交矩阵将会极大的简化上面三讲我们说到过的公式！

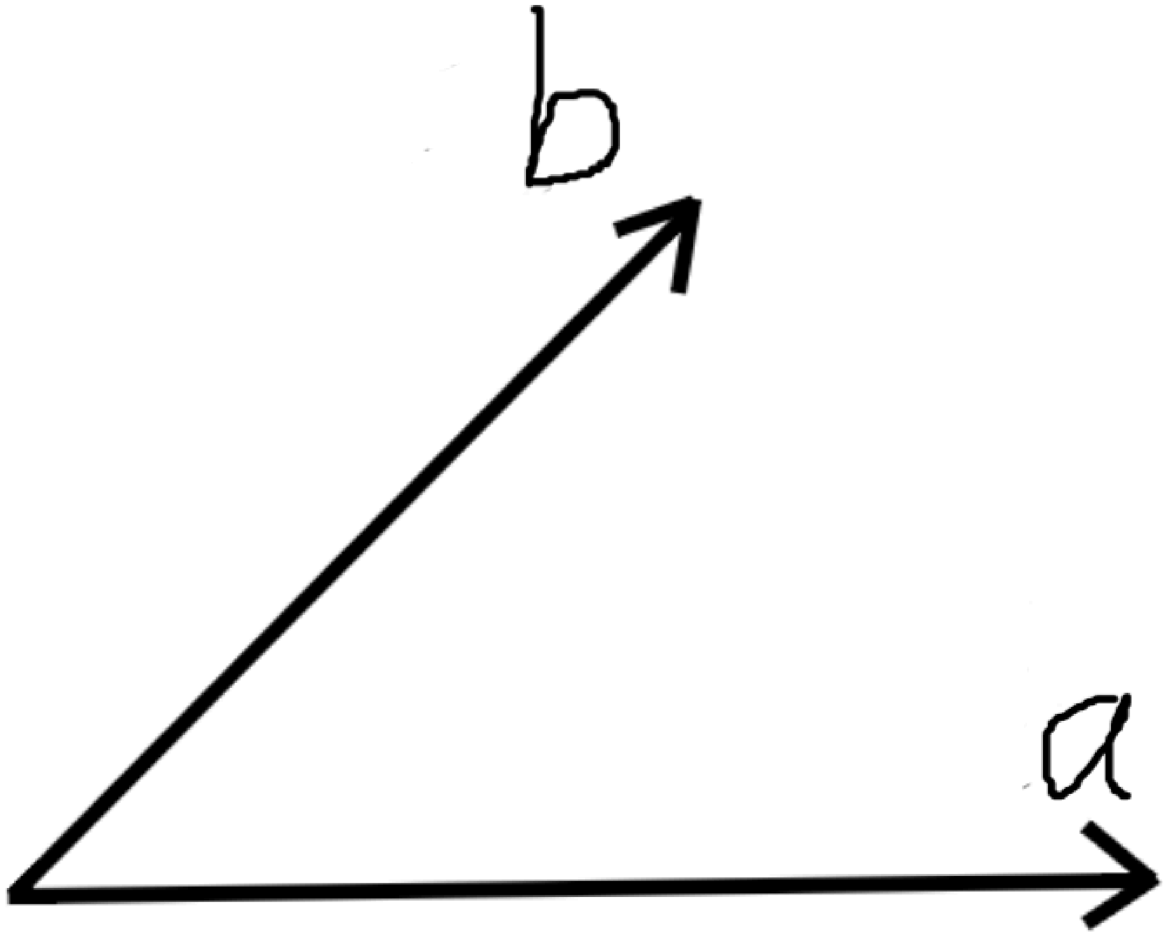
首先出场的是投影矩阵的公式：$A(A^TA)^{-1}A^T$，当我们把A换成Q的时候得到：$Q(Q^TQ)^{-1}Q^T$。然后我们来分析他会简便成什么样子：
（1）当Q是方阵的时候：
我们在上文中得知$Q^T=Q^{-1}$，那么得到：$Q(Q^TQ)^{-1}Q^T=QIQ^T=QQ^T$。而Q是方阵的时候，$QQ^T=I$。那么最后投影矩阵是单位向量！其实我们可以从空间角度去了解这件事，当Q是方阵且可逆，那么他的列空间是填满整个n维空间的。那么我们要投影的向量就一定会在这个列空间中，那么他本身就是他的投影！
（2）当Q不是方阵的时候：
我们在上文中得知$Q^T=Q^{-1}$，那么得到：$Q(Q^TQ)^{-1}Q^T=QIQ^T=QQ^T$。这样也简便了公式！

然后就是拟合方程，最小二乘法中的：$A^TA\hat{x}=A^Tb$。当我们把A换成Q的时候得到：
$Q^TQ\hat{x}=Q^Tb$
我们在上文中得知$Q^T=Q^{-1}$，那么得到$\hat{x}=Q^Tb$。这样化简之后，很明显$\hat{x}$的每个分量都是 Q 中对应列向量与 b 的点乘结果。即：$\hat{x}_i=q_i^Tb$。这个式子的意义就是，如果我们已知标准正交基，那么 b 在第 i 个基上的投影就是对应基向量q_i^Tb。

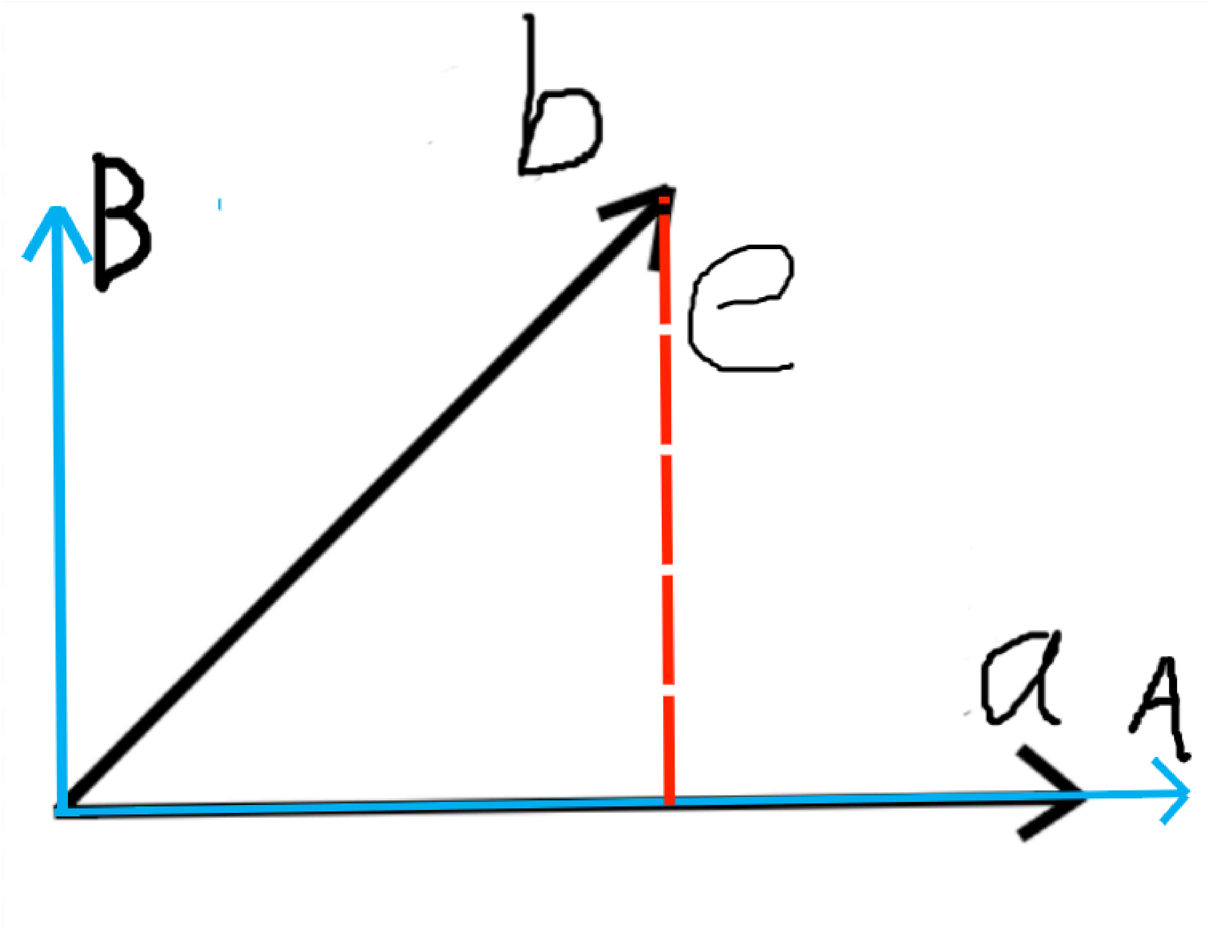
足以见得标准正交矩阵的优良性质！

Gram-Schmidt 正交化

当我们拥有一组线性无关的向量的时候，把他们集成为一个矩阵A！如何将A矩阵标准正交化。举个例子来看吧：
有两个线性无关的向量 a，b，他们并不正交。我们想从中得到标准正交向量q_1，q_2。
像这样：



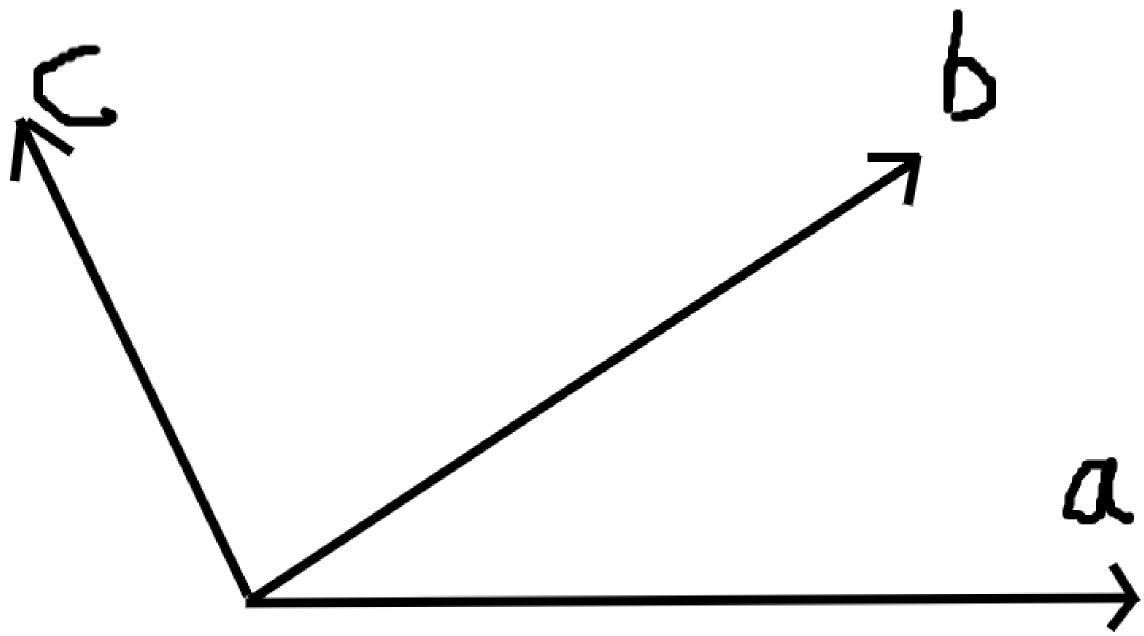
首先我们可以假设已知一组基 A,B 是正交的，其中 a=A。那么标准正交向量$q_1=\frac{A}{|A|},q_2=\frac{B}{|B|}$。
那么我们画个图来看看：



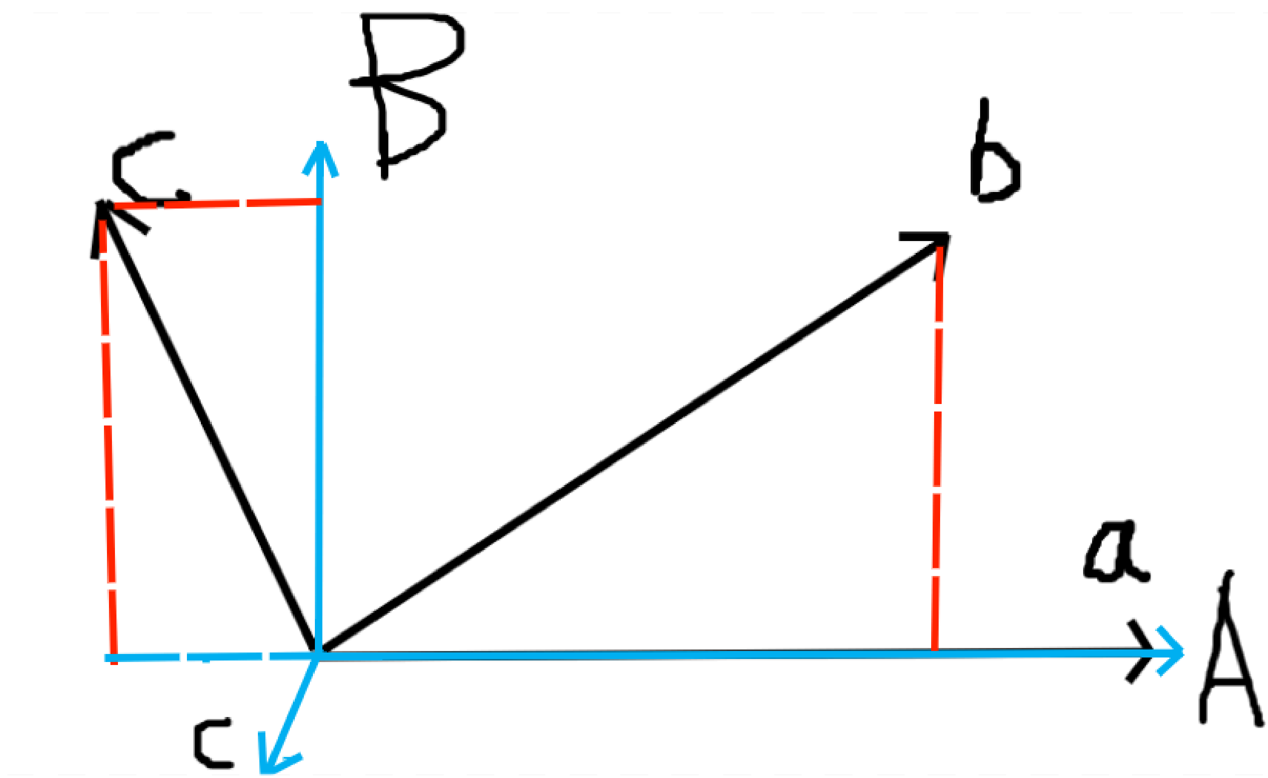
现在我们需要把向量b调到B的位置，也就是调到 e 的位置。怎么调，联想我们之前学习的投影，这个 B 即为 (b-p)！得到的是：$B=b-\frac{A^Tb}{A^TA}A$！带进去检验一下，$A^TB=A^T(b-\frac{A^Tb}{A^TA}A)=A^Tb-A^T\frac{A^Tb}{A^TA}A$，由于$A^TA$是一个常数，可以自由约分不受影响，所以$A^Tb-A^T\frac{A^Tb}{A^TA}A=A^Tb-A^Tb=0$，所以他们正交！原来矩阵是$[a|b]$，现在矩阵是$[a|B]$, 也就是$[A|B]$。

> tips：相信大家有疑惑，在之前学到的内容知道$p=\frac{aa^T}{a^Ta}b$, 也就是$\frac{AA^T}{A^TA}b$，这里怎么$p=\frac{A^Tb}{A^TA}A$，按理说线性代数的运算中矩阵位置是不可以随意调换的！我来解释一下。首先A^TA是一个标量，A^Tb也是标量（这是不难理解的）。所以在$\frac{AA^T}{A^TA}b$中只有A是矩阵，那么A^Tb倍的矩阵A和b倍的矩阵AA^T是一样的。所以$\frac{AA^T}{A^TA}b=A\frac{A^Tb}{A^TA}=\frac{A^Tb}{A^TA}A$。这样就解释通了！所以大家不必拘泥于形式！

同样的道理，推广到三维：
有三个线性无关变量a,b,c。如图：



然后我们有一组正交基A，B，C！其中 a=A。那么$q_1=\frac{A}{|A|},q_2=\frac{B}{|B|},q_3=\frac{C}{|C|}$。如图：



那么一样的：

A=a

B=$b-\frac{A^Tb}{A^TA}A$

C=$c-\frac{A^Tc}{A^TA}A-\frac{B^Tc}{B^TB}B$
得到三个正交的向量 A,B,C，再进行单位化即可。
再高维以此类推！

这个时候问一个问题：矩阵A在变成Q后，列空间改变了吗？观察两个矩阵的列空间，它们是相同的，也就是说我们的正交化过程都是在同一个空间中进行的，只是最后得到了一个更好的标准正交基而已。因为我们是在原先a,b列的基础上操作得来的！

从矩阵的角度来看，类似于 A 的 LU 分解，在 Gram-Schmidt 正交化中，A 可分解为 Q 与 R。其中 R 是上三角矩阵：A = QR！
为什么这么说呢，A=LU，其中A变成U，是对A的行进行操作得来的，L是记录行操作的矩阵！同样，Q是由A的列进行列操作得来的，那么R是记录列操作的矩阵！这么一看很像了吧！
而且矩阵R是上三角矩阵，原因如下：
A = QR，由于$Q^T=Q^{-1}$那么$Q^TA=Q^TQR\rightarrow Q^TA=Q^{-1}QR\rightarrow R=Q^TA$

那么A=$\begin{bmatrix}a & b & c & d & \dots\end{bmatrix}$，$Q=\begin{bmatrix}q_1 & q_2 & q_3 & q_4 & \dots\end{bmatrix}, Q^T=\begin{bmatrix}q_1^T \\ q_2^T \\ q_3^T \\ q_4^T \\ \dots\end{bmatrix}$

可得R=$\begin{bmatrix}q_1^T \\ q_2^T \\ q_3^T \\ q_4^T \\ \dots\end{bmatrix}\begin{bmatrix}a & b & c & d & \dots\end{bmatrix}=\begin{bmatrix}q_1^Ta & q_1^Tb & q_1^Tc & \dots \\ q_2^Ta & q_2^Tb & q_2^Tc & \dots \\ q_3^Ta & q_3^Tb & q_3^Tc & \dots \\ \dots\dots\dots\end{bmatrix}$，其中主对角线上的值为1，而对角线上的为不为0，对角线下面的为0（分析正交关系可得！）因为每一个q_i都正交于原来的矩阵中的$a_j(j\ \ i)$，比如q_3是由a_3减去在q_1,q_2方向的分量，自然是正交的！
如果上面内容太过于抽象，那么大家可以这么理解，R是记录A变成Q的列操作的！那么第一列自然是仅仅对第一列进行操作的，所以后面就都是0，后面的列以此类推！（讨论课我们有一个例子来解释这个）

到此我们终于结束对正交化的所有内容！

讨论课

有一个向量A有三个列向量a,b,c，A=$\begin{bmatrix}1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6\end{bmatrix}$，找到对应的矩阵Q矩阵，以及求出A=QR中的R！
解答：
显然$q_1=\frac{a}{1}=\begin{bmatrix}1 \\ 0 \\ 0\end{bmatrix}$，刚好长度是1，我们无需单位化。

而$q_2=b-\frac{q_1^Tb}{q_1^Tq_1}q_1=b-q_1^Tbq_1=b-2q_1=\begin{bmatrix}0\\0\\3\end{bmatrix}$。单位化后：$q_2=\begin{bmatrix}0\\0\\1\end{bmatrix}$。可得$q_2=3q_2$

$q_3=c-\frac{q_1^Tc}{q_1^Tq_1}q_1-\frac{q_2^Tc}{q_2^Tq_2}q_2=c-q_1^Tcq_1-q_2^Tcq_2=c-4q_1-6q_2=\begin{bmatrix}0\\5\\0\end{bmatrix}$,单位化后得到$q_3=\begin{bmatrix}0\\1\\0\end{bmatrix}$。可得$q_3=5q_3$

那么A=QR，则$\begin{bmatrix}1&2&4\\0&0&5\\0&3&6\end{bmatrix}=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1&2&4\\0&3&6\\0&0&5\end{bmatrix}$

结合上面的例子，a确实等于1倍的q_1，b等于2倍q_1加3倍的q_2，后面的类似！可以明白为什么R是上三角了吧！

习题课

问题一

标准正交向量自动线性无关。
用矩阵证明：证明若 $Qx=0$，则必有 $x=0$。由于 Q 可能是长方形矩阵，可以使用 Q^T 但不能使用 Q^{-1}。

解答

根据定义，矩阵 Q 的列是标准正交的，因此有 $Q^TQ=I$（Q 可以是长方形矩阵）。于是：

$$Qx=0\Rightarrow Q^TQx=Q^T0\Rightarrow Ix=0\Rightarrow x=0.$$

因此，Q 的零空间仅含零向量，其列向量线性无关。不存在任何非零线性组合能使 Q 的列向量之和为零向量。这说明**标准正交向量自动线性无关**。

问题 17.2（第4.4节习题18）

给定以下向量 a,b,c，用 **Gram-Schmidt正交化过程** 求出正交向量 A,B,C，并要求它们张成的空间与原向量组相同：

$$a=(1,-1,0,0),\quad b=(0,1,-1,0),\quad c=(0,0,1,-1).$$

进一步证明：集合 $\{A,B,C\}$ 和 $\{a,b,c\}$ 都是垂直于向量 $d=(1,1,1,1)$ 的空间的基底。

解答

对 a,b,c 应用 **Gram-Schmidt正交化过程**：

1. 令：

$$A = a = (1, -1, 0, 0).$$

2. 计算 B：

$$B = b - \frac{A^T b}{A^T A} A = (0, 1, -1, 0) - \frac{-1}{2}(1, -1, 0, 0) = \left(\frac{1}{2}, \frac{1}{2}, -1, 0\right).$$

3. 计算 C：

$$C = c - \frac{A^T c}{A^T A} A - \frac{B^T c}{B^T B} B = (0, 0, 1, -1) - \frac{0}{2} A - \frac{-1}{32}\left(\frac{1}{2}, \frac{1}{2}, -1, 0\right) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -1\right).$$

由问题一可知，集合 $\{A, B, C\}$ 中的向量线性无关，且每个向量都与 $(1, 1, 1, 1)$ 正交。垂直于 d 的向量空间是三维的（因为 $(1, 1, 1, 1)$ 的行空间是一维的，而行空间维数与零空间维数之和为4）。因此，$\{A, B, C\}$ 构成了垂直于 d 的空间的一组基底。

类似地，$\{a, b, c\}$ 也是垂直于 d 的空间的基底，因为这些向量线性无关、与 $(1, 1, 1, 1)$ 正交，且数量为3。