第十七讲：正交矩阵和Gram-Schmidt正交化法

标准正交矩阵

定义标准正交向量（orthonormal）：$q_i^T q_j = \begin{cases}0 & i \neq j\\ 1 & i = j\end{cases}$

我们将标准正交向量放入矩阵中，有$Q = [q_1 q_2 \cdots q_n]$。

上一讲我们研究了A^A的特性，现在来观察$Q^T Q = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} [q_1 q_2 \cdots q_n]$

根据标准正交向量的定义，计算$Q^T Q = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I$，我们也把$Q$成为标准正交矩阵（orthonormal matrix）。

特别的，当Q恰好是方阵时，由于正交性，易得Q是可逆的，又$Q^T Q = I$，所以$Q^T = Q^{-1}$。

- 举个置换矩阵的例子：$Q = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$，则$Q^T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$，易得$Q^T Q = I$。
- 使用上一讲的例子$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$，列向量长度为1，且列向量相互正交。

- 其他例子$Q=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$，列向量长度为1，且列向量相互正交。
- 使用上一个例子的矩阵，令$Q'=c\begin{bmatrix}Q&Q\\Q&-Q\end{bmatrix}$，取合适的$c$另列向量长度为1也可以构造标准正交矩阵：$Q=\frac{1}{2}\begin{bmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{bmatrix}$，这种构造方法以阿德玛（Adhemar）命名，对$2,4,16,64,\cdots$阶矩阵有效。
- 再来看一个例子，$Q=\frac{1}{3}\begin{bmatrix}1&-2&2\\2&-1&-2\\2&2&1\end{bmatrix}$，列向量长度为1，且列向量相互正交。

  格拉姆-施密特正交化法的缺点在于，由于要求得单位向量，所以我们总是除以向量的长度，这导致标准正交矩阵中总是带有根号，而上面几个例子很少有根号。

再来看标准正交化有什么好处，假设要做投影，将向量b投影在标准正交矩阵Q的列空间中，根据上一讲的公式得$P=Q(Q^TQ)^{-1}Q^T$，易得$P=QQ^T$。我们断言，当列向量为标准正交基时，QQ^T是投影矩阵。极端情况，假设矩阵是方阵，而其列向量是标准正交的，则其列空间就是整个向量空间，而投影整个空间的投影矩阵就是单位矩阵，此时$QQ^T=I$。可以验证一下投影矩阵的两个性质：$(QQ^T)^T=(Q^T)^TQ^T=QQ^T$，得证；$(QQ^T)^2=QQ^TQQ^T=Q(Q^TQ)Q^T=QQ^T$，得证。

我们计算的$A^TA\hat{x}=A^Tb$，现在变为$Q^TQ\hat{x}=Q^Tb$，也就是$\hat{x}=Q^Tb$，分解开来看就是$\underline{\hat{x}_i=q_i^Tb}$，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，则解向量第i个分量为基的第i个分量乘以b，在第i个基方向上的投影就等于q_i^Tb。

Gram-Schmidt正交化法

我们有两个线性无关的向量a,b，先把它们化为正交向量A,B，再将它们单位化，变为单位正交向量$q_1=\frac{A}{\|A\|},q_2=\frac{B}{\|B\|}$：

- 我们取定a向量的方向，$a=A$；
- 接下来将b投影在A的法方向上得到B，也就是求子空间投影一讲中，我们提到的误差向量$e=b-p$，即$B=b-\frac{A^Tb}{A^TA}A$。检验一下$A\bot B$，$A^TB=A^Tb-A^T\frac{A^Tb}{A^TA}A=A^Tb-\frac{A^TA}{A^TA}A^Tb=0$。（$\frac{A^Tb}{A^TA}A$就是$A\hat{x}=p$。）

如果我们有三个线性无关的向量a,b,c，则我们现需要求它们的正交向量A,B,C，再将它们单位化，变为单位正交向量$q_1=\frac{A}{\|A\|},q_2=\frac{B}{\|B\|},q_3=\frac{C}{\|C\|}$：

- 前两个向量我们已经得到了，我们现在需要求第三个向量同时正交于A, B；
- 我们依然沿用上面的方法，从c中减去其在A, B上的分量，得到正交与A, B的C：$C = c - \frac{A^Tc}{A^TA}A - \frac{B^Tc}{B^TB}B$。

现在我们试验一下推导出来的公式，$a = \begin{bmatrix}1\\1\\1\end{bmatrix}, b = \begin{bmatrix}1\\0\\2\end{bmatrix}$：

- 则$A = a = \begin{bmatrix}1\\1\\1\end{bmatrix}$；
- 根据公式有$B = a - hA$，h是比值$\frac{A^Tb}{A^TA} = \frac{3}{3}$，则$B = \begin{bmatrix}1\\1\\1\end{bmatrix} - \frac{3}{3}\begin{bmatrix}1\\0\\2\end{bmatrix} = \begin{bmatrix}0\\-1\\1\end{bmatrix}$。验证一下正交性有$A \cdot B = 0$。
- 单位化，$q_1 = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}, \quad q_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\2\end{bmatrix}$，则标准正交矩阵为$Q = \begin{bmatrix}\frac{1}{\sqrt{3}} & 0\\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}}\end{bmatrix}$，对比原来的矩阵$D = \begin{bmatrix}1 & 1\\1 & 0\\1 & 2\end{bmatrix}$，有$D, Q$的列空间是相同的，我们只是将原来的基标准正交化了。

我们曾经用矩阵的眼光审视消元法，有$A = LU$。同样的，我们也用矩阵表达标准正交化，$A = QR$。设矩阵A有两个列向量$[a_1 a_2]$，则标准正交化后有$[a_1 a_2] = [q_1 q_2]\begin{bmatrix}a_1^Tq_1 & a_2^Tq_1\\ a_1^Tq_2 & a_2^Tq_2\end{bmatrix}$，而左下角的$a_1^Tq_2$始终为0，因为Gram-Schmidt正交化总是使得$a_1 \perp q_2$，后来构造的向量总是正交于先前的向量。所以这个R矩阵是一个上三角矩阵。