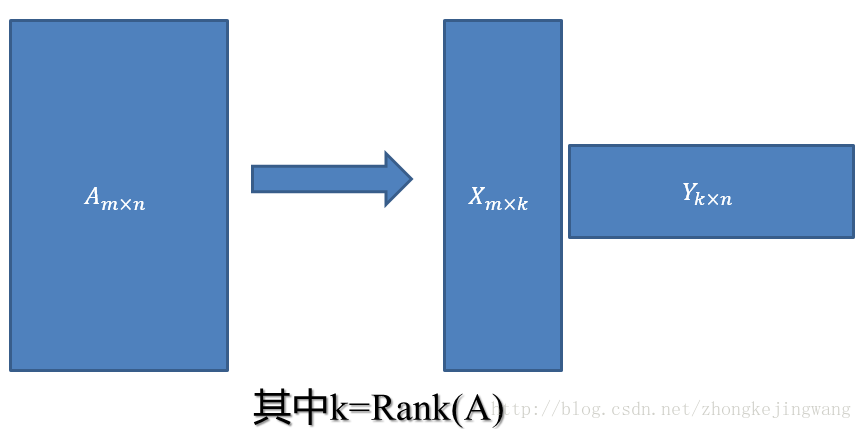
本周学习：SVD奇异值分解（其一）

首先，SVD可以用于分解压缩数据。

利用SVD实现，能够用足够小的数据集来表示原始数据集。这样实际上是去除了噪声和冗余信息。可以看成从噪声数据中提取相关特征。

SVD不仅是一个数学问题，在工程应用中的很多地方都有它的身影，比如PCA，掌握了SVD原理后对于理解PCA会变得简单。在推荐系统方面，SVD更是名声大噪，将它应用于推荐系统的是Netflix大奖的获得者Koren，可以在Google上找到他写的文章；用SVD可以很容易得到任意矩阵的满秩分解，用满秩分解可以对数据做压缩。可以用SVD来证明对任意M\*N的矩阵均存在如下分解：

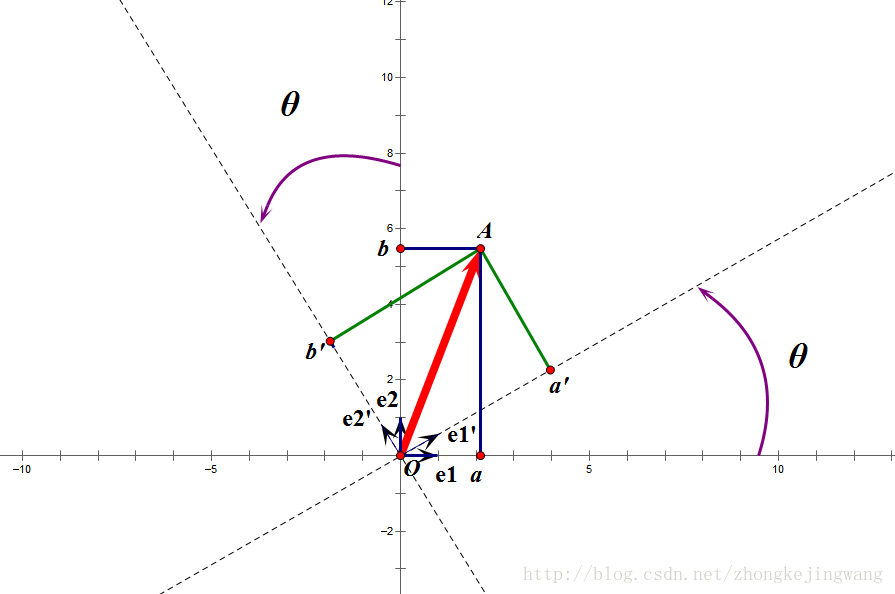


其中Rank（A）表示矩阵A的秩。

这个特点就可以运用在在数据降维压缩上。

正交矩阵

   正交矩阵是在欧几里得空间里的叫法，在酉空间里叫酉矩阵，一个正交矩阵对应的变换叫正交变换，这个变换的特点是不改变向量的尺寸和向量间的夹角（旋转是一个例子），那么它到底是个什么样的变换呢？看下面这张图



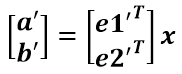
假设二维空间中的一个向量OA，它在标准坐标系也即e1、e2（标准正交基）表示的坐标是中表示为(a,b)'（用'表示转置），现在把它用另一组坐标e1'、e2'表示为(a',b')'，存在矩阵U使得(a',b')'=U(a,b)'，则U即为正交矩阵。从图中可以看到，正交变换只是将变换向量用另一组正交基表示，在这个过程中并没有对向量做拉伸，也不改变向量的空间位置，加入对两个向量同时做正交变换，那么变换前后这两个向量的夹角显然不会改变。上面的例子只是正交变换的一个方面，即旋转变换，可以把e1'、e2'坐标系看做是e1、e2坐标系经过旋转某个θ角度得到，怎么样得到该旋转矩阵U呢？如下

IMG_256

IMG_256

IMG_256

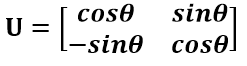
a'和b'实际上是x在e1'和e2'轴上的投影大小，所以直接做内积可得，可得



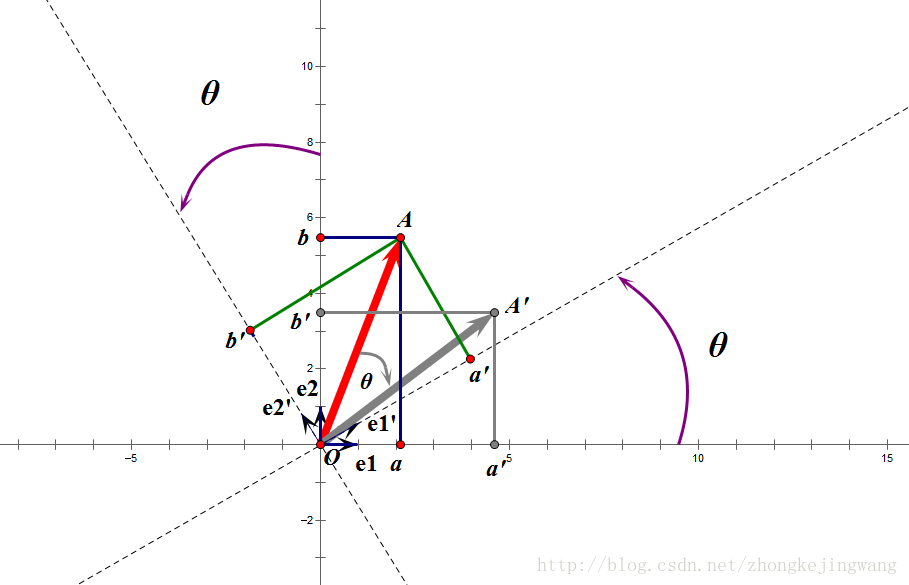
可以得到

IMG_256IMG_256

所以有



正交阵U行（列）向量之间都是单位正交向量。上面求得的是一个旋转矩阵，它对向量做旋转变换！也许你会有疑问：刚才不是说向量空间位置不变吗？怎么现在又说它被旋转了？对的，这两个并没有冲突，说空间位置不变是绝对的，但是坐标是相对的，加入你站在e1上看OA，随着e1旋转到e1'，看OA的位置就会改变。如下图：



如图，如果我选择了e1'、e2'作为新的标准坐标系，那么在新坐标系中OA（原标准坐标系的表示）就变成了OA'，这样看来就好像坐标系不动，把OA往顺时针方向旋转了θ角度，这个操作实现起来很简单：将变换后的向量坐标仍然表示在当前坐标系中。

旋转变换是正交变换的一个方面，这个挺有用的，比如在开发中需要实现某种旋转效果，直接可以用旋转变换实现。正交变换的另一个方面是反射变换，也即e1'的方向与图中方向相反，这个不再讨论。

总结：正交矩阵的行（列）向量都是两两正交的单位向量，正交矩阵对应的变换为正交变换，它有两种表现：旋转和反射。正交矩阵将标准正交基映射为标准正交基（即图中从e1、e2到e1'、e2'）

特征值分解——EVD

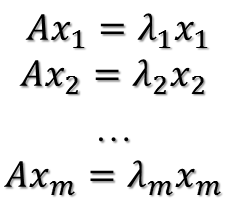
在讨论SVD之前先讨论矩阵的特征值分解（EVD），在这里，选择一种特殊的矩阵——对称阵（酉空间中叫hermite矩阵即厄米阵）。对称阵有一个很优美的性质：它总能相似对角化，对称阵不同特征值对应的特征向量两两正交。一个矩阵能相似对角化即说明其特征子空间即为其列空间，若不能对角化则其特征子空间为列空间的子空间。现在假设存在mxm的满秩对称矩阵A，它有m个不同的特征值，设特征值为

IMG_256

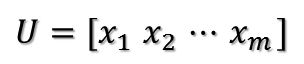
对应的单位特征向量为

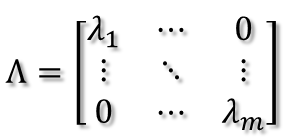
IMG_257

则有



进而  IMG_259





所以可得到A的特征值分解（由于对称阵特征向量两两正交，所以U为正交阵，正交阵的逆矩阵等于其转置）

IMG_262

这里假设A有m个不同的特征值，实际上，只要A是对称阵其均有如上分解。

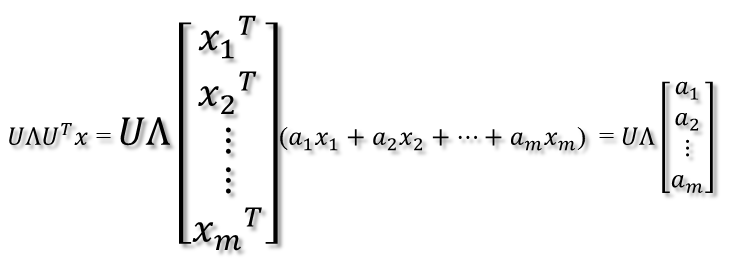
矩阵A分解了，相应的，其对应的映射也分解为三个映射。现在假设有x向量，用Ａ将其变换到Ａ的列空间中，那么首先由U'先对x做变换：

IMG_263

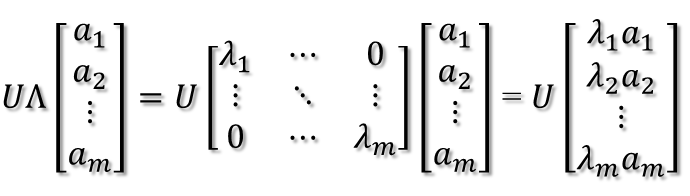
U是正交阵U'也是正交阵，所以U'对x的变换是正交变换，它将x用新的坐标系来表示，这个坐标系就是A的所有正交的特征向量构成的坐标系。比如将x用A的所有特征向量表示为：

IMG_264

则通过第一个变换就可以把x表示为[a1 a2 ... am]'：



紧接着，在新的坐标系表示下，由中间那个对角矩阵对新的向量坐标换，其结果就是将向量往各个轴方向拉伸或压缩：

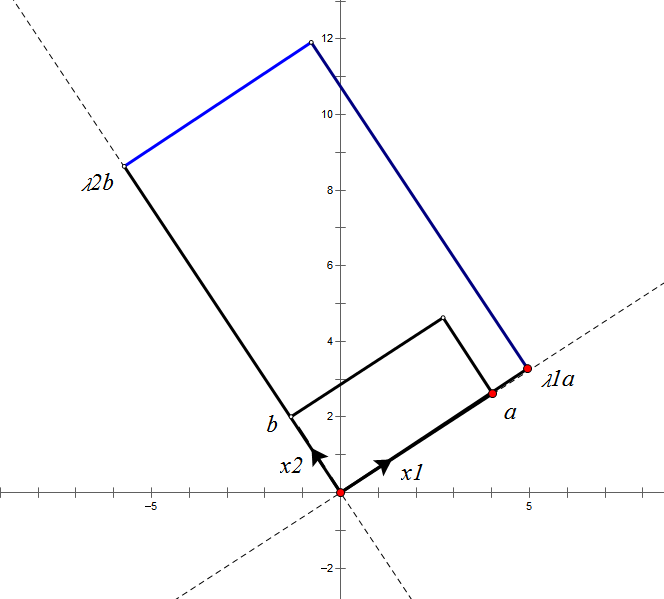


从上图可以看到，如果A不是满秩的话，那么就是说对角阵的对角线上元素存在0，这时候就会导致维度退化，这样就会使映射后的向量落入m维空间的子空间中。

最后一个变换就是U对拉伸或压缩后的向量做变换，由于U和U'是互为逆矩阵，所以U变换是U'变换的逆变换。

因此，从对称阵的分解对应的映射分解来分析一个矩阵的变换特点是非常直观的。假设对称阵特征值全为1那么显然它就是单位阵，如果对称阵的特征值有个别是0其他全是1，那么它就是一个正交投影矩阵，它将m维向量投影到它的列空间中。

根据对称阵A的特征向量，如果A是2\*2的，那么就可以在二维平面中找到这样一个矩形，是的这个矩形经过A变换后还是矩形：



这个矩形的选择就是让其边都落在A的特征向量方向上，如果选择其他矩形的话变换后的图形就不是矩形了！

参考资料：

[1].http://blog.csdn.net/zhongkejingwang/article/details/43053513

[2].机器学习实战第十四章-利用SVD简化数据