SVD

2015年7月27日 星期一　　下午10:34

第二次看，理解：

一、简介
是一种矩阵的变换，与矩阵的特征值分解类似。
矩阵的特征值分解是仅仅针对方阵的一种分解，分解成比较好理解、表示、做一些事情的形式。
奇异值分解针对所有的矩阵都可以做，分解成好理解的形式。

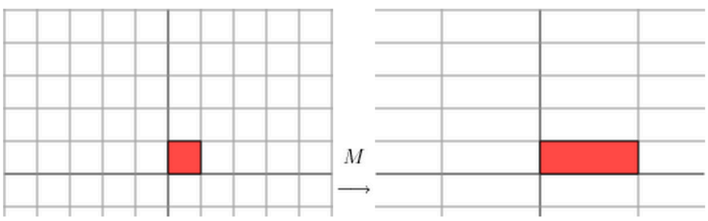
二、特征值分解/矩阵的特征分解
目的是将矩阵(仅适用于方阵)，用一种方式更形象、好理解的方式描述。
1、线性变换
首先，矩阵可以用来做线性变换，例如

$$M = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

从几何的角度，矩阵可以描述为一个变换：用矩阵乘法将平面上的点（x, y）变换成另外一个点（3x, y）：

$$\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ y \end{bmatrix}.$$

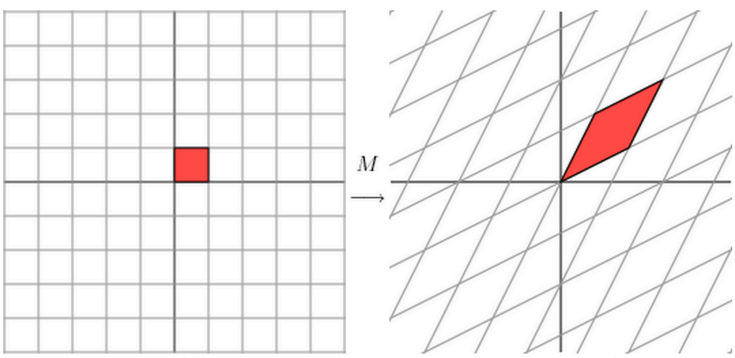
这种变换的效果如下：平面在水平方向被拉伸了3倍，在竖直方向无变化。



再如，

$$M = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

它会产生如下的效果



不过这张图貌似也并没有能够简洁、清晰的描述出上述矩阵变换的几何效果。然而，
我们能看出来，矩阵右乘一个向量可以做到，对这个向量线性变换，也就是在向量所在的空间进行拉伸等变换。

在向量所在的空间进行拉伸等变换。

还能看出，不同的矩阵对不同的向量有不同形式的拉伸，有的只在一个方向上拉伸，有的是多个方向上。

为什么会这样，这是和矩阵本身、向量本身有关

2、特征值分解

a）特征向量、特征值

任何一个方阵M，都存在一个正交向量组v，使得v里面的每个向量v_i，都满足：

$$Mv_i = \lambda_i v_i$$

其中，λ_i是个常量，如果方阵M是n维的，这个向量组就有n个这样的向量

正交向量组的含义是，这是一组向量，而不是一个向量，这一组向量两两之间是正交的。

正交我们可以理解为，就像坐标轴一样，是两个不同的方向，互不干扰。

一个严格的说法是，正交向量组中的所有向量是线性无关的(线性独立)

V_i叫做M的特征向量，λ_i叫做M的特征值

b）特征值分解/矩阵的特征分解

将$Mv_i = \lambda_i v_i$中的n个v_i(i=1,2,3....n)写在一起，写成在一个矩阵中就形成

$$A = Q\Sigma Q^{-1}$$

Q中就包含n个向量v_i(i=1,2,3....n)

这个就是特征值分解/矩阵的特征分解

这里Q是由M的n个特征向量组成的矩阵

c）特征值分解数学含义的理解

结合1）线性变换里面的内容考虑一下，特征值分解的数学含义，为什么这样分解，为什么一定存在特征向量和特征值？

不同的矩阵对不同的向量v_i有不同形式的拉伸，有的只在一个方向上拉伸，有的是多个方向上。

为什么会这样，这是和矩阵M本身、向量v_i本身有关，矩阵是对向量的一种线性变换，某个向量的方向比较巧的时候，这个线性变换做完以后只是在这个向量v_i的基础上做一个拉伸(不会更改方向，比如第一个例子。

我的理解是：这个矩阵所代表的n维空间上的线性变换(也可能不到n

维，但是肯定<=n维)，是作用与n个线性独立的方向上的(方向数有可能<=n，我们暂认为是n，不影响什么)，可以理解为2、3维空间的坐标轴

由于向量v_i的特殊性，他所指示的方向，正好在"坐标轴"上，这样这个向量v_i，做完矩阵M的线性变换之后，只会被做拉伸操作。

反过来，我们可以通过哪些向量满足“只被做了拉伸操作”来确定这个向量指示的方向是不是 矩阵M的n个独立的方向之一。

“只被做了拉伸操作”

也就是满足，经过M的变换之后，相当于在向量的基础上乘一个常数。

也就是满足$Mv_i = \lambda_i v_i$

“这个向量指示的方向是不是 矩阵M的n个独立的方向之一”

也就是这个向量是不是特征向量

如果是的话，我们就找出了这n个特征向量，这n个特征向量，指示了M变换的n维空间，特征值λ_i是每个方向上拉伸的长度

特征值越大，代表包含的信息越多，越是主要的方向，降维的时候，这个方向越会被保留

我的理解是，特征值分解之后的结果，是找出了一个方阵中(如果做线性变换的话)的方向，这叫做特征向量，也找出了在各个特征向量下拉伸的程度，这就是特征值。

d）应用

可以用来做降维等。

三、奇异值分解

1、定义

特征值分解只能对方阵做，所以提出奇异值分解，类似特征值分解

也是转化成三个矩阵的乘积，也是两个正交的矩阵中间夹着一个类似对角阵(但不是)的矩阵

$$A = U\Sigma V^T$$

假设A是一个N＊M的矩阵，那么得到的U是一个N＊N的方阵（里面的向量是正交的，U里面的向量称为左奇异向量），Σ是一个N＊M的矩阵（除了对角线的元素都是0，对角线上的元素称为奇异值），V’(V的转置)是一个N＊N的矩阵，里面的向量也是正交的，V里面的向量称为右奇异向量）

2、推导
借鉴特征值分解的方法，先A^T * A 转成一个方阵，然后先求A^T * A的特征向量和特征值

$$(A^T A)v_i = \lambda_i v_i$$

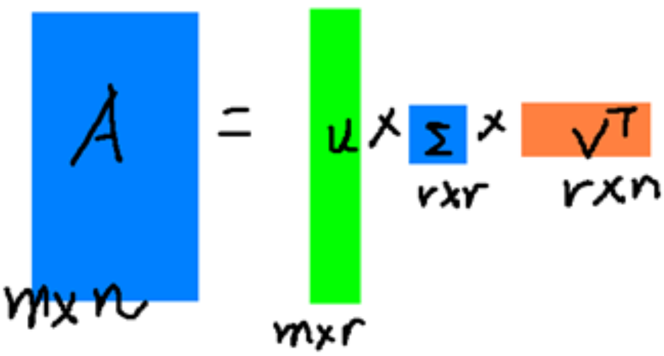
继续推导

$$\sigma_i = \sqrt{\lambda_i}$$

$$u_i = \frac{1}{\sigma_i} A v_i$$

3、降维
同特征值分解的道理，对λi(拉伸量的度量)排序，取top。
维度也降维r

转为：



四、参考资料：
不错的博客：http://www.cnblogs.com/LeftNotEasy/archive/2011/01/19/svd-and-applications.html
wiki：https://zh.wikipedia.org/wiki/%E5%A5%87%E5%BC%82%E5%80%BC%E5%88%86%E8%A7%A3
详细的英文博客：http://www.ams.org/samplings/feature-column/fcarc-svd

上面的英文博客的译文：http://www.flickering.cn/%E6%95%B0%E5%AD%A6%E4%B9%8B%E7%BE%8E/2015/01/%E5%A5%87%E5%BC%82%E5%80%BC%E5%88%86%E8%A7%A3%EF%BC%88we-recommend-a-singular-value-decomposition%EF%BC%89/

第一次看（Ng的课）：

一,简介

SVD是对一般的矩 阵m * n,进行分解的,可以将任意一个矩阵,分解成三个矩阵相乘的形式,

$$C = U\Sigma V^T$$

其中 U ∑ V 都有着一定的性质 ,他们的组合也有着一定的性质

Ng说是用的最广泛的一种分解,很常用的一种变换,SVD是LSA的数学基础.

二,适用范围

矩阵分解的概念

矩阵分解是仅仅针对方阵的,可以将方阵分解为三个矩阵相乘,详见 参考资料1 中的讲解

奇异值分解作为矩阵分解的补充,对任何矩阵都可以做

三, 形式

$M = U\Sigma V^*$

其 中 U 是 m×m 阶 酉矩阵 ；Σ是 半正定 m×n阶对角矩阵；而V*，即V的共轭转置，是n×n阶酉矩阵。

Σ对角线上的元素Σi,i 即为M的奇异值。 有大小顺序的,我们可以把奇异值排个序 选top K 什么的以达到降维的目的

注意:1. U 是方阵 本作者LSA的中的SVD写错了 2. ∑ 中 非零元素的个数可以自己控制 不受m和n的限制 (见wiki, Ng上例子)

四.性质

性质 1

$$M^*M = V\Sigma^*U^*U\Sigma V^* = V(\Sigma^*\Sigma)V^*$$

$$MM^* = U\Sigma V^*V\Sigma^*U^* = U(\Sigma\Sigma^*)U^*$$

这个性质在LSA中用到了

性质2

V 的列向量（右奇异向量）是 M^*M 的 特征向量 。

U 的列向量（左奇异向量）是 MM^* 的特征向量。

Σ 的非零对角元（非零奇异值）是 M^*M 或者 MM^* 的非零 特征值 的平方根。

五.实现

在openCV 上有实现 ,提供了python,MATLAB接口等

void cvSVD(CvArr* A, CvArr* W, CvArr* U=NULL, CvArr* V=NULL, int flags=0)

void cvSVD(CvArr* A, CvArr* W, CvArr* U=NULL, CvArr* V=NULL, int flags=0)
六,应用
1, 做PCA
Ng哥在公开课的第15讲 中说了 大概在18~20min的地方,SVD 可以用在PCA中,当PCA的维度过高时.
PCA中需要用到的协方差矩阵 $\sum = \sum \sum (x_i * x_j)$ 矩阵中对位相乘之后进行相加
用矩阵的形式就可以表示成 $\sum = X^T * X$
可以对X 进行SVD分解 $X = U * D * V$ 由于 SVD分解中U V 是正定矩阵 所以$X^T * X$的时候 U 就可以消掉了
$\sum = V^T * D^2 * V$ 这个计算维度就低了吧
(但为啥高维的时候SVD可以用 这个还没想明白)
2,LSA (latent semantic analyse)
潜在语义分析,详见另一篇笔记
七,参考资料:
1 http://blog.csdn.net/wangran51/article/details/7408414 SVD 不错
2 http://blog.csdn.net/wangran51/article/details/7408406 在LSA上应用
3 http://blog.sina.com.cn/s/blog_4b16455701016aau.html 奇异值与特征值关系 :除 了一般性的定义 文章里说了句: 奇异值是A*AH(A的 共轭转置)的特征值的平方根
4 http://baike.baidu.com/view/3068725.htm?fr=aladdin
5 http://zh.wikipedia.org/wiki/%E5%A5%87%E5%BC%82%E5%80%BC%E5%88%86%E8%A7%A3 比较详细 有些总结出来的性质不错 建议多看!