

捕获

359b033b5bb5c9ea8718f4add639b6003bf3b391

从直观上来看，协方差表示的是两个变量总体误差的期望。

如果两个变量的变化趋势一致，也就是说如果其中一个大于自身的期望值时另外一个也大于自身的期望值，那么两个变量之间的协方差就是正值；如果两个变量的变化趋势相反，即其中一个变量大于自身的期望值时另外一个却小于自身的期望值，那么两个变量之间的协方差就是负值。

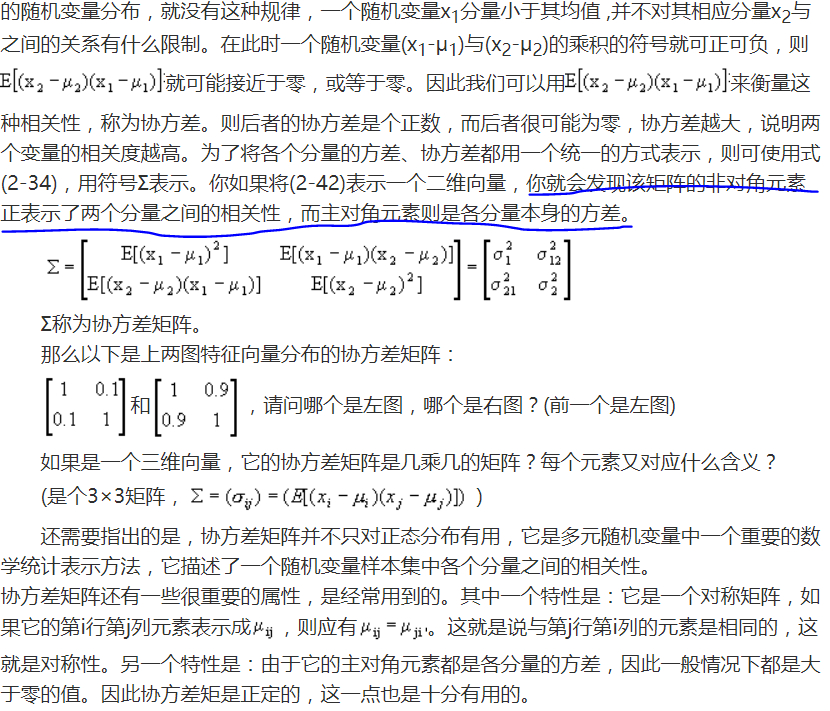
如果*X*与*Y*是统计独立的，那么二者之间的协方差就是0，因为两个独立的随机变量满足*E*[*XY*]=*E*[*X*]*E*[*Y*]。

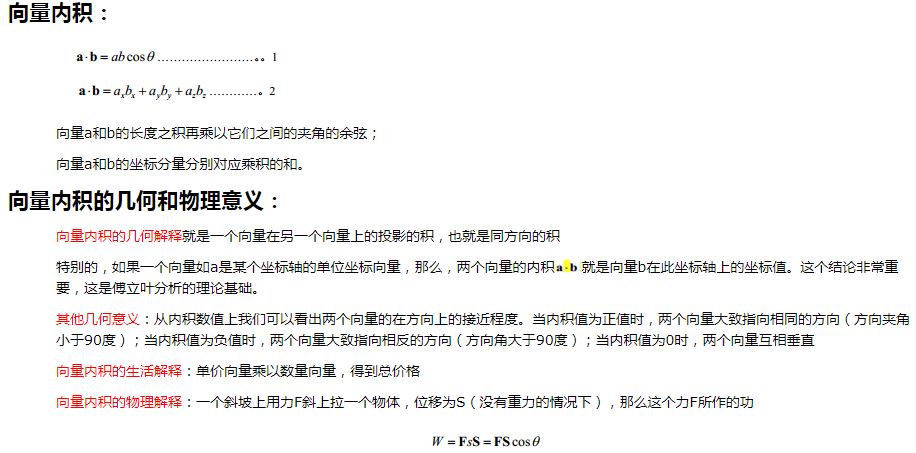
但是，反过来并不成立。即如果*X*与*Y*的协方差为0，二者并不一定是统计独立的。

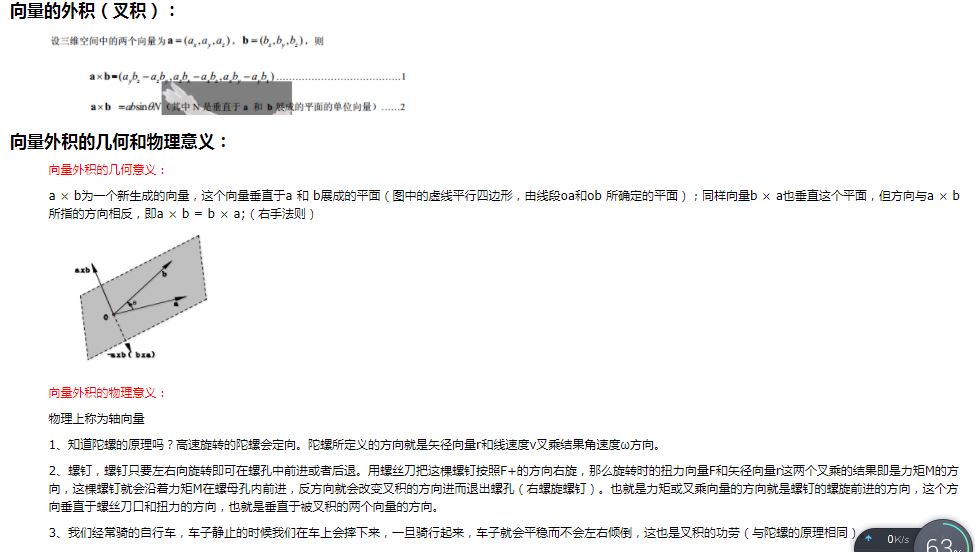
相关系数 v2-67c6163775717c78659fcbf8f92470a9_hd

翻译一下：就是用X、Y的协方差除以X的标准差和Y的标准差。是规范化后的协方差，消除了度量单位不一致的影响。

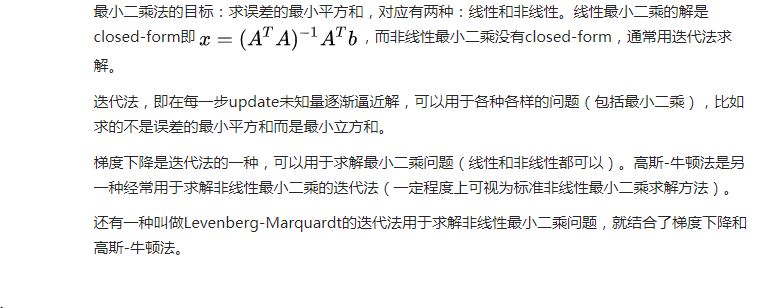
**多元正态分布**

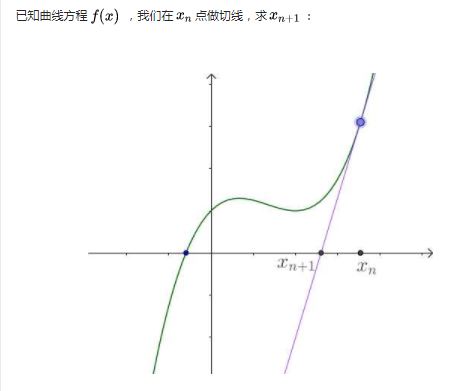


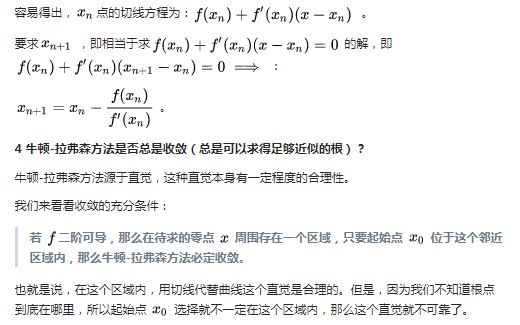




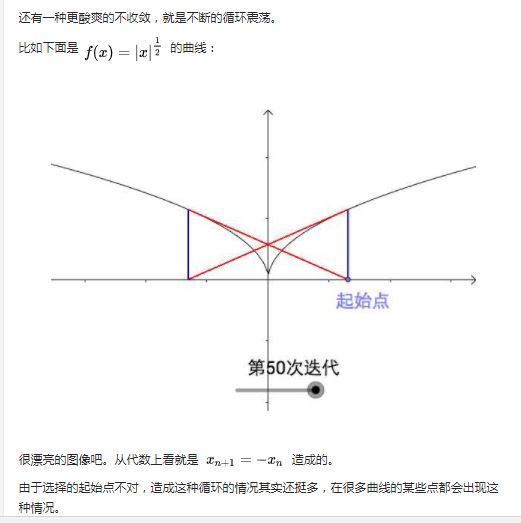
梯度法是一种迭代计算方法



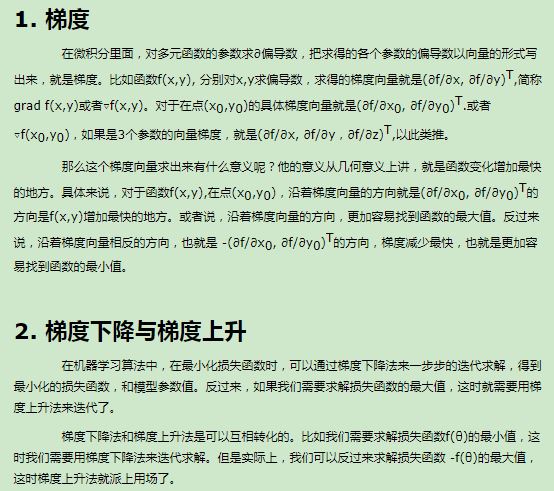




牛顿法有适用条件，不是所有的函数都可以用，有的用了也不能求解，比如下面的例子



梯度法和牛顿法都是迭代计算方法



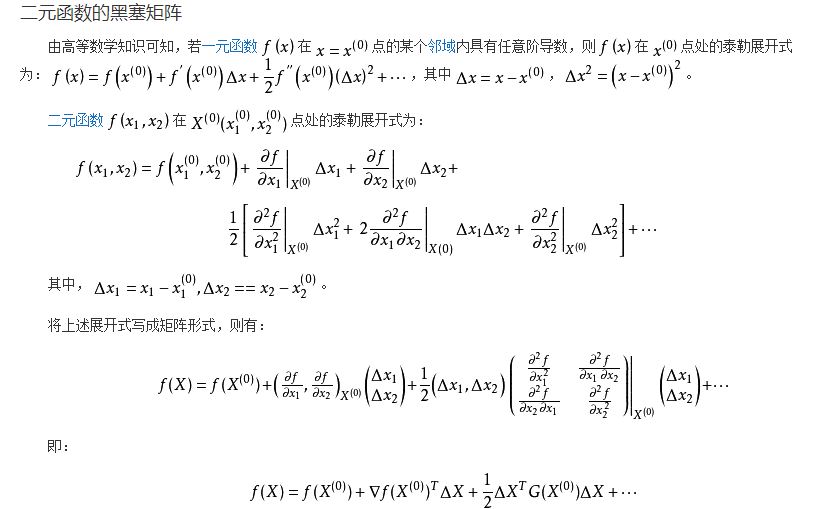
梯度下降的相关概念：1,步长：迭代的长度

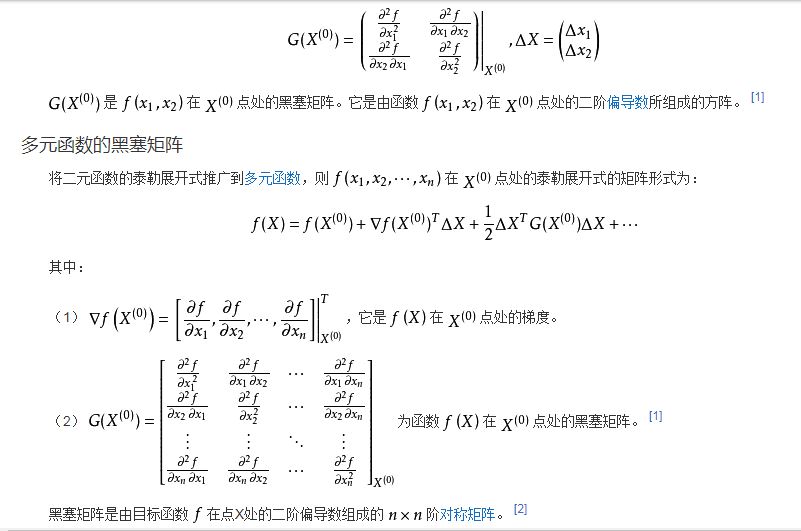
2,特征：样本的输入部分

3,假设函数：假设的模型

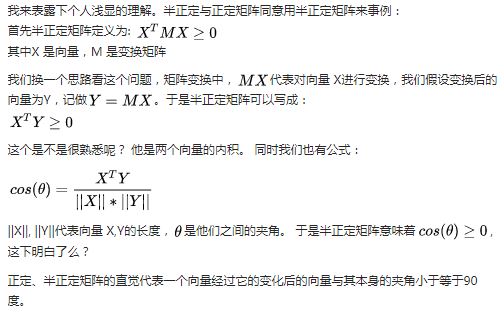
4,损失函数：计算假设函数和实际样本的拟合情况

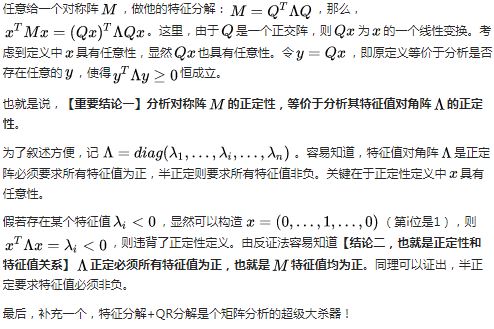
多元函数的黑塞矩阵

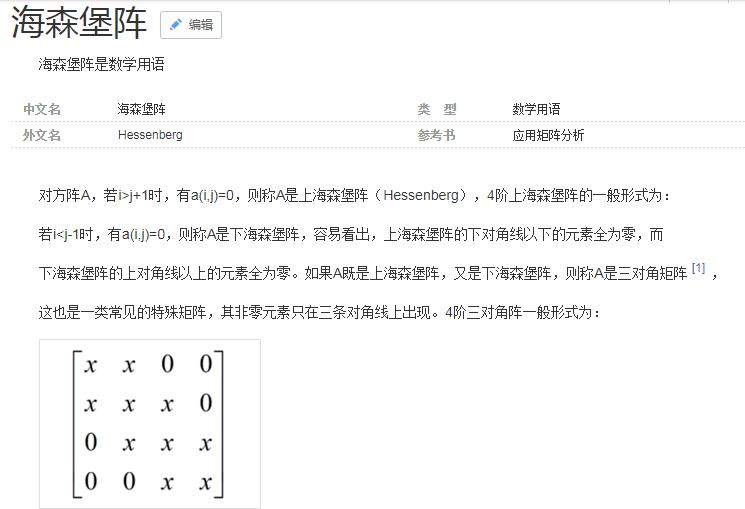




正定矩阵:





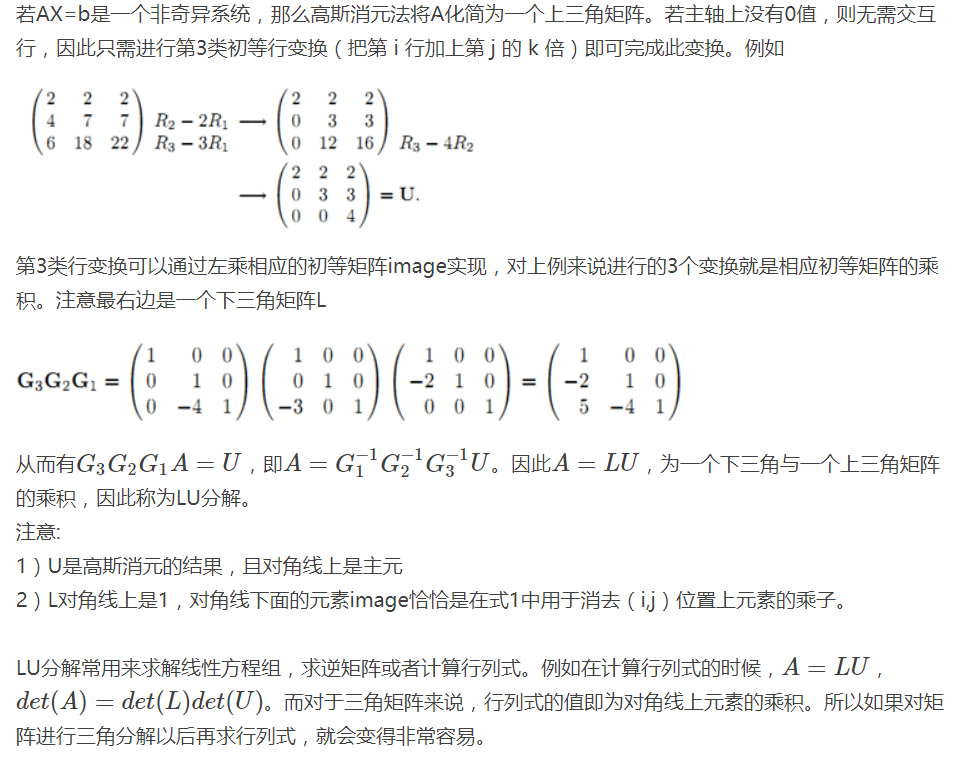


矩阵分解:

http://blog.csdn.net/bitcarmanlee/article/details/52662518

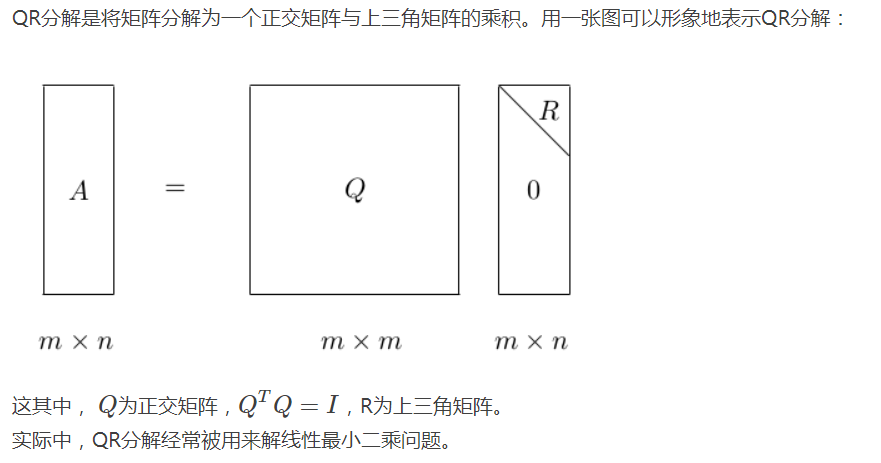
## 1.三角分解(LU分解) NXN矩阵

矩阵的LU分解是将一个矩阵分解为一个**下三角矩阵**与**上三角矩阵**的乘积。本质上，LU分解是高斯消元的一种表达方式。首先，对矩阵A通过初等行变换将其变为一个上三角矩阵。对于学习过线性代数的同学来说，这个过程应该很熟悉，线性代数考试中求行列式求逆一般都是通过这种方式来求解。然后，将原始矩阵A变为上三角矩阵的过程，对应的变换矩阵为一个下三角矩阵。这中间的过程，就是Doolittle algorithm(杜尔里特算法)。



在线性代数中已经证明，如果方阵A是非奇异的，即A的行列式不为0，LU分解总是存在的。

## 2.QR分解 MXN矩阵



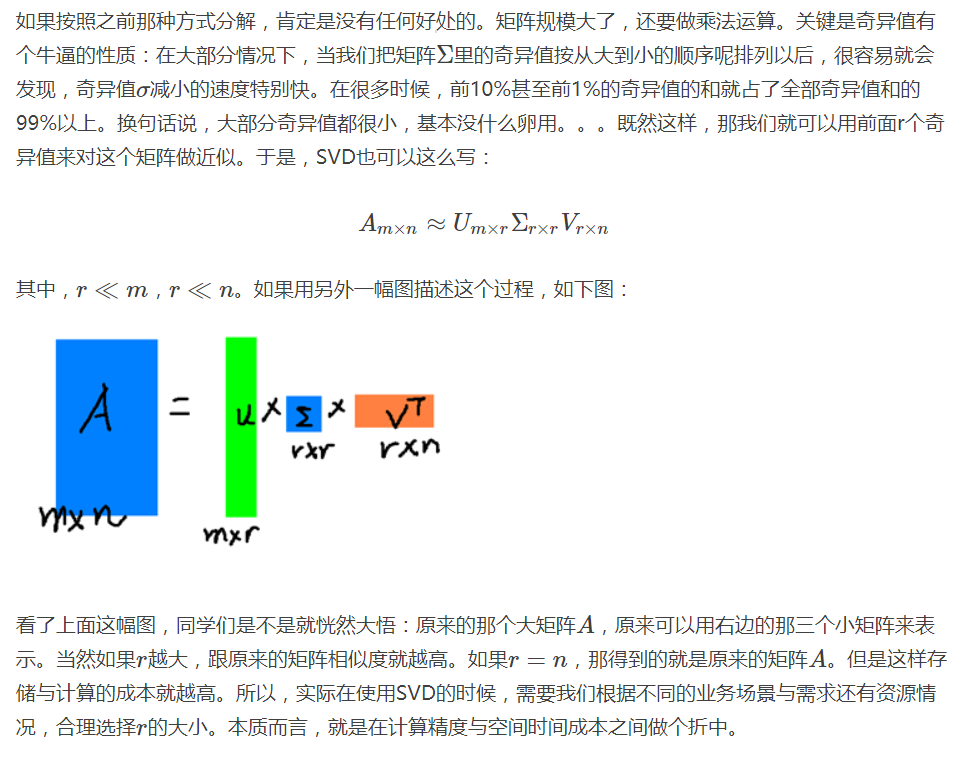
**SVD奇异值分解**



矩阵可以看做运动，包括旋转，拉伸，投影三个操作。

首先，矩阵可以认为是一种线性变换，而且这种线性变换的作用效果与基的选择有关。以Ax = b为例，x是m维向量，b是n维向量，m,n可以相等也可以不相等，表示矩阵可以将一个向量线性变换到另一个向量，这样一个线性变换的作用可以包含旋转、缩放和投影三种类型的效应。

奇异值分解正是对线性变换这三种效应的一个析构。A=U∑V，u和v是两组正交单位向量，是对角阵，表示奇异值，它表示我们找到了和这样两组基，A矩阵的作用是将一个向量从u这组正交基向量的空间旋转到v这组正交基向量空间，并对每个方向进行了一定的缩放，缩放因子就是各个奇异值。如果维度比大，则表示还进行了投影。可以说奇异值分解将一个矩阵原本混合在一起的三种作用效果，分解出来了。



矩阵可以看做运动，包括旋转，拉伸，投影三个操作。  
  
 首先，矩阵可以认为是一种线性变换，而且这种线性变换的作用效果与基的选择有关。以Ax = b为例，x是m维向量，  
b是n维向量，m,n可以相等也可以不相等，表示矩阵可以将一个向量线性变换到另一个向量，这样一个线性变换的作用  
可以包含旋转、缩放和投影三种类型的效应。  
 奇异值分解正是对线性变换这三种效应的一个析构。A=U∑V，u和v是两组正交单位向量，是对角阵，表示奇异值，它表示我们找到  
了和这样两组基，A矩阵的作用是将一个向量从u这组正交基向量的空间旋转到v这组正交基向量空间，并对每个方向进行了一定  
的缩放，缩放因子就是各个奇异值。如果维度比大，则表示还进行了投影。可以说奇异值分解将一个矩阵原本混合在一起的  
三种作用效果，分解出来了。