协方差的定义

http://hi.csdn.net/attachment/201103/23/0_13008475759rb0.gif

对于一般的分布，直接代入E(X)之类的就可以计算出来了，但真给你一个具体数值的分布，要计算协方差矩阵，根据这个公式来计算，还真不容易反应过来。网上值得参考的资料也不多，这里用一个例子说明协方差矩阵是怎么计算出来的吧。

记住，X、Y是一个列向量，它表示了每种情况下每个样本可能出现的数。比如给定

http://hi.csdn.net/attachment/201103/23/0_1300847734WwY3.gif

则X表示x轴可能出现的数，Y表示y轴可能出现的。注意这里是关键，给定了4个样本，每个样本都是二维的，所以只可能有X和Y两种维度。所以

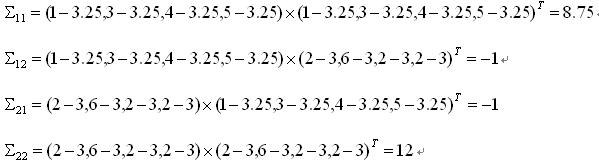
http://hi.csdn.net/attachment/201103/23/0_1300847802eQfp.gif

http://hi.csdn.net/attachment/201103/23/0_130084793593l9.gif

用中文来描述，就是：

协方差(i,j)=（第i列的所有元素-第i列的均值）\*（第j列的所有元素-第j列的均值）

这里只有X,Y两列，所以得到的协方差矩阵是2x2的矩阵，下面分别求出每一个元素：



       所以，按照定义，给定的4个二维样本的协方差矩阵为：

http://hi.csdn.net/attachment/201103/23/0_13008480227Mfz.gif

用matlab计算这个例子

z=[1,2;3,6;4,2;5,2]

cov(z)

ans =

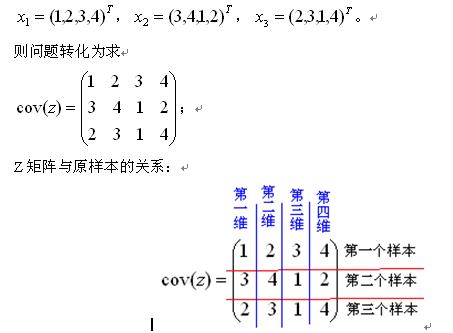
    2.9167   -0.3333

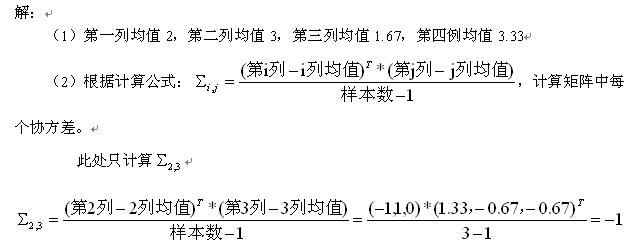
   -0.3333    4.0000

可以看出，matlab计算协方差过程中还将元素统一缩小了3倍。所以，协方差的matlab**计算公式**为：

**协方差(i,j)=（第i列所有元素-第i列均值）\*（第j列所有元素-第j列均值）/（样本数-1）**

       下面在给出一个4维3样本的实例，注意4维样本与符号X,Y就没有关系了，X,Y表示两维的，4维就直接套用计算公式，不用X,Y那么具有迷惑性的表达了。





        （3）与matlab计算验证

                     Z=[1 2 3 4;3 4 1 2;2 3 1 4]

                     cov(Z)

                     ans =

                          1.0000    1.0000   -1.0000   -1.0000

                          1.0000    1.0000   -1.0000   -1.0000

                         -1.0000   -1.0000    1.3333    0.6667

                          -1.0000   -1.0000    0.6667    1.3333

       可知该计算方法是正确的。我们还可以看出，协方差矩阵都是方阵，它的维度与样本维度有关（相等）。参考2中还给出了计算协方差矩阵的源代码，非常简洁易懂，在此感谢一下!