协方差和协方差矩阵

协方差的定义：

对于一般的分布，直接代入E(X)之类的就可以直接计算出来了，但如果真给一个具体的数值分布，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。

下面用一个例子来说明协方差矩阵是怎么计算出来的：

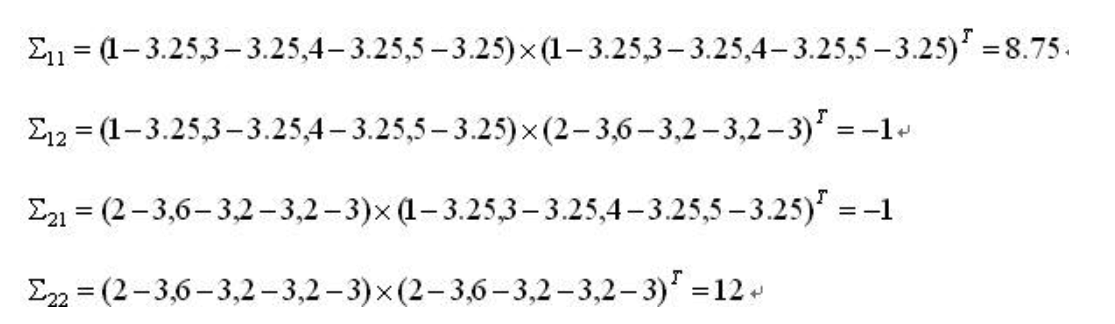
记住，X和Y是一个列向量，它表示了每种情况下每个样本可能出现的数，比如给定：

那么上面的Cov(X,Y)中的X则表示x轴可能出现的数，Y表示y轴可能出现的数。注意这里是关键，给定了4个样本，但每个样本都是二维的，所以只可能有X和Y两种维度。所以：

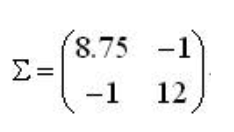
用中文来描述，就是：

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就是2\*2的矩阵，下面分别求出每一个元素：



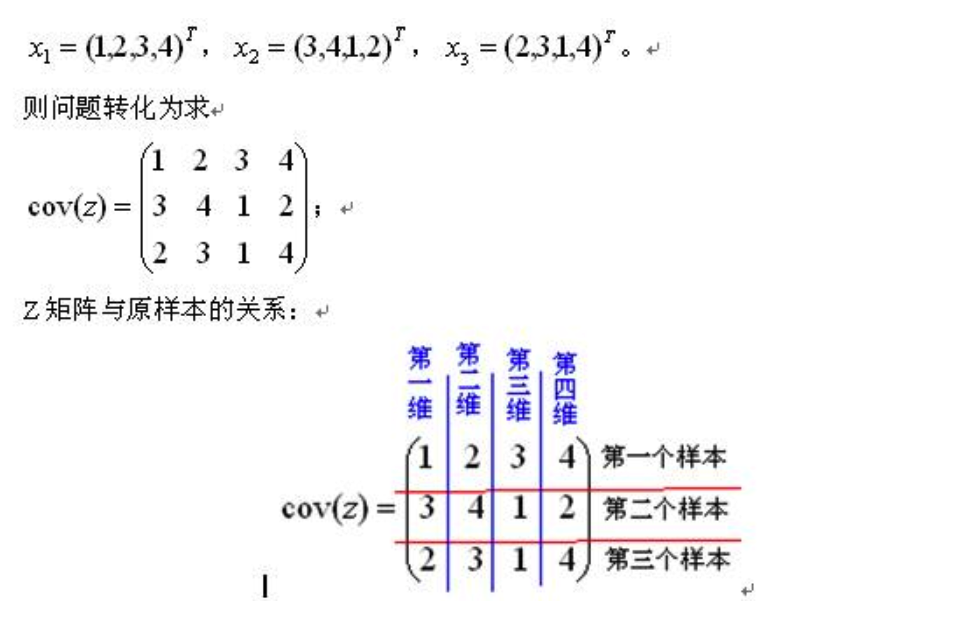
所以，按照定义，给定的4个二维样本的协方差矩阵为：



其中matlab协方差的计算公式中，要在分母除以（样本数-1）.

下面给出一个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。

如以下样本：



解法如下：

