如何求协方差矩阵（Covariance Matirx）

首先先来复习一下方差(variance)和协方差的公式(Covariance):

$\sigma_x^2 = \frac{1}{n-1} \sum_* i = 0^n (x_i - \bar{x})^2$

$\sigma(x,y) = \frac{1}{n-1} \sum_{i=0}^{n} (x_i - \bar{x})(y_i - \bar{y})$

制造一个矩阵，假设这个矩阵行(row)为属性，列为每个样本的具体数据，协方差矩阵是为了找属性和属性（也就是行）之间的关系:

$$\begin{bmatrix} X_{1,1} & X_{1,2} & X_{1,3} \\ X_{2,1} & X_{2,2} & X_{2,3} \end{bmatrix}$$

首先先求得每一行的平均数：$\bar{X}_1$ 和$\bar{X}_2$，接下来让每一行分别减去自己的平均数：

$$\begin{bmatrix} X_{1,1} - \bar{X}_1 & X_{1,2} - \bar{X}_1 & X_{1,3} - \bar{X}_1 \\ X_{2,1} - \bar{X}_2 & X_{2,2} - \bar{X}_2 & X_{2,3} - \bar{X}_2 \end{bmatrix}$$

然后求这个矩阵的转置：

$$\begin{bmatrix} X_{1,1} - \bar{X}_1 & X_{2,1} - \bar{X}_2 \\ X_{1,2} - \bar{X}_1 & X_{2,2} - \bar{X}_2 \\ X_{1,3} - \bar{X}_1 & X_{2,3} - \bar{X}_2 \end{bmatrix}$$

然后我们让两个矩阵相乘，我们先来观察一下，第一个的结果实际等于

$$\begin{aligned} C_{1,1} &= (X_{1,1} - \bar{X}_1)(X_{1,1} - \bar{X}_1) + (X_{1,2} - \bar{X}_1)(X_{1,2} - \bar{X}_1) \\ &+ (X_{1,3} - \bar{X}_1)(X_{1,3} - \bar{X}_1) \\ &= Variance(X_1) \times (N-1) \end{aligned}$$

第二个的结果为:

$$\begin{aligned} C_{1,2} &= (X_{1,1} - \bar{X}_1)(X_{2,1} - \bar{X}_2) + (X_{1,2} - \bar{X}_1)(X_{2,2} - \bar{X}_2) \\ &+ (X_{1,3} - \bar{X}_1)(X_{2,3} - \bar{X}_2) \\ &= Covariance(X_1) \times (N-1) \end{aligned}$$

因此协方差矩阵可以写为:

$$C = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(X_i - \bar{X})^T$$

得到的结果为:

$$\begin{bmatrix} variance(X_1) & Covariance(X_1, X_2) \\ Covariance(X_2, X_1) & variance(X_2) \end{bmatrix}$$