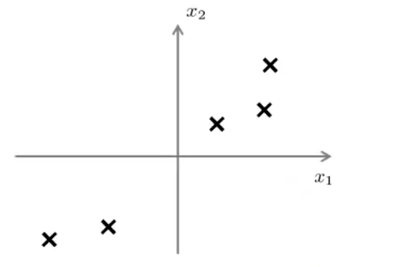
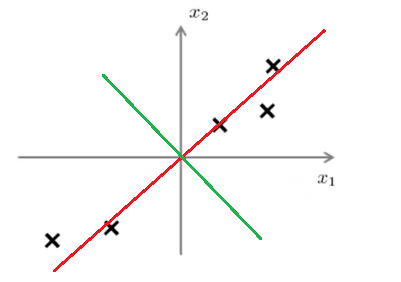
**PCA**

**1.二维数据表示**



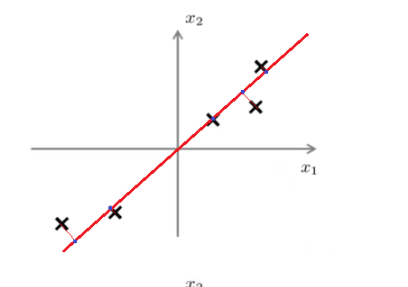
**2.降维（降维就会损失数据信息）**

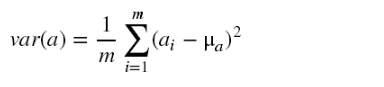
1.绿线和红线分别表示了数据投影在一维线上的情况。



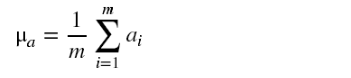
2.数据投影在红线上的分布比投影在绿线上的分布广。（尽可能地减少损失）

3.投影后的点（蓝色）





其中：μa



方便计算，所以所有的特征都减去该特征的均值，并依然用ai来表示，所以蓝点之间的方差可以记为：

