pca:

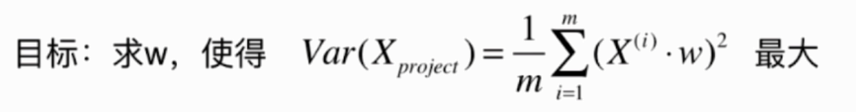
用于数据降维，将高维数据映射到低维度,具有减少计算量及降低噪音功能.

pca算法解析:

在降维的过程中,我们希望将数据映射到方差较大的维度上,方差越大,数据越离散,区分度越好.根据方差公式 方差等于 各样本数值 - 样本平均值 的平方和 除以样本数.



因为我们要求的是一个单位方向向量,所以其模等于 1 ,另外为了简单起见,化简过程中将数据的平均值设成0,对于现实中的数据,这相当于预处理时加上了一个demean操作.最终简其实等价于以下式子:



这个式子可以根据梯度上升法求得最大值.

求导得:

