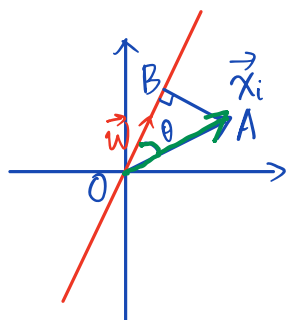


现有  $X: n \times p$ ,  $n$  是观测值的个数,  $p$  是 features 个数.  $X$  可以看成  $p$  维空间的一组点.

问题: 如何在尽可能保留信息的情况下, 把  $X$  的维度减少?

方案: 把  $X$  投影到  $q$  维空间上去, 使这  $n$  个点 to 投影点的距离最小.

从  $p$  维投影到过原点的一条直线上. (红线)



$\vec{w}$  是红线的 unit vector.

OB 的长度是  $\|\vec{x}_i\| \cdot \cos \theta = \|\vec{x}_i\| \frac{\vec{x}_i \cdot \vec{w}}{\|\vec{x}_i\| \|\vec{w}\|} = \vec{x}_i \cdot \vec{w}$

OB 的坐标是  $(\vec{x}_i \cdot \vec{w}) \vec{w}$

目标: 希望投影过后仍然尽可能保持  $\vec{x}_i$  的信息. 信息就是长度.

目标函数:  $\min_{\vec{w}} \frac{1}{n} \sum_{i=1}^n \|\vec{x}_i - (\vec{x}_i \cdot \vec{w}) \vec{w}\|^2$

$$\begin{aligned} \|\vec{x}_i - (\vec{x}_i \cdot \vec{w}) \vec{w}\|^2 &= \vec{x}_i \cdot \vec{x}_i - 2(\vec{x}_i \cdot \vec{w})^2 + (\vec{x}_i \cdot \vec{w})^2 \vec{w} \cdot \vec{w} \\ &= \vec{x}_i \cdot \vec{x}_i - (\vec{x}_i \cdot \vec{w})^2 \end{aligned}$$

$$MSE(\vec{w}) = \frac{1}{n} \sum_{i=1}^n \|\vec{x}_i\|^2 - \frac{1}{n} \sum_{i=1}^n (\vec{x}_i \cdot \vec{w})^2$$

最小化 MSE, 也即最大化  $\frac{1}{n} \sum_{i=1}^n (\vec{x}_i \cdot \vec{w})^2$   $E(x^2) = (E(x))^2 + \text{Var} x$ .

$$\frac{1}{n} \sum_{i=1}^n (\vec{x}_i \cdot \vec{w})^2 = \underbrace{\left( \frac{1}{n} \sum_{i=1}^n \vec{x}_i \cdot \vec{w} \right)^2}_{=0} + \boxed{\text{Var}(\vec{x}_i \cdot \vec{w})}$$

$= \left( \frac{1}{n} \sum_{i=1}^n \vec{x}_i \right) \cdot \vec{w}$

$\vec{x}_i$  已作标准化, 所以第一项为 0.

中心化,  $x$  均值为 0,  $\frac{1}{n} \sum_{i=1}^n x_i = 0, \forall x^j \in \{x^1, x^2, \dots, x^p\}$

若投影到  $q$  维空间

$\sum_{j=1}^q (\vec{x}_i \cdot \vec{w}_j) \vec{w}_j$ ,  $w_j$  是  $q$  维子空间的标准正交基  
 $q$  维空间中  $q$  个方向上的和, 与  $\vec{x}$  的距离  
 不希望太大

$$\|\vec{x}_i - \sum_{j=1}^q (\vec{x}_i \cdot \vec{w}_j) \vec{w}_j\|^2 = \vec{x}_i \cdot \vec{x}_i - 2 \sum_{j=1}^q (\vec{x}_i \cdot \vec{w}_j) \vec{x}_i \cdot \vec{w}_j + \left( \sum_{j=1}^q (\vec{x}_i \cdot \vec{w}_j) \vec{w}_j \right) \cdot \left( \sum_{j=1}^q (\vec{x}_i \cdot \vec{w}_j) \vec{w}_j \right)$$

$$\star = \left[ (\vec{x}_i \cdot \vec{w}_1) \vec{w}_1 + \dots + (\vec{x}_i \cdot \vec{w}_q) \vec{w}_q \right] \cdot \left[ (\vec{x}_i \cdot \vec{w}_1) \vec{w}_1 + \dots + (\vec{x}_i \cdot \vec{w}_q) \vec{w}_q \right]$$

$$= (\vec{x}_i \cdot \vec{w}_1)^2 \vec{w}_1 \cdot \vec{w}_1 + (\vec{x}_i \cdot \vec{w}_2)^2 \vec{w}_2 \cdot \vec{w}_2 + \dots + (\vec{x}_i \cdot \vec{w}_q)^2 \vec{w}_q \cdot \vec{w}_q$$

$$= \sum_{j=1}^q (\vec{x}_i \cdot \vec{w}_j)^2$$

$$\rightarrow = \vec{x}_i \cdot \vec{x}_i - \sum_{j=1}^q (\vec{x}_i \cdot \vec{w}_j)^2$$

最大化  $\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^q (\vec{x}_i \cdot \vec{w}_j)^2 = \sum_{j=1}^q \frac{1}{n} \sum_{i=1}^n (\vec{x}_i \cdot \vec{w}_j)^2$

$$= \frac{1}{n} \sum_{i=1}^n (\vec{x}_i \cdot \vec{w}_1)^2 + \frac{1}{n} \sum_{i=1}^n (\vec{x}_i \cdot \vec{w}_2)^2 + \dots + \frac{1}{n} \sum_{i=1}^n (\vec{x}_i \cdot \vec{w}_q)^2$$

$$= \frac{1}{n} (\sum \vec{w}_j)^T (\sum \vec{w}_j)$$

$$= \frac{1}{n} \vec{w}_j^T \sum \vec{x}^T \vec{x} \vec{w}_j$$

$$= \vec{w}_j^T \frac{\sum \vec{x}^T \vec{x}}{n} \vec{w}_j$$

$$= \vec{w}_j^T \text{Var}(\vec{x}) \vec{w}_j$$

约束条件:  $\vec{w}_j \cdot \vec{w}_j = 1$

Lagrange:  $\vec{w}_j^T \text{Var}(\vec{x}) \vec{w}_j - \lambda (\vec{w}_j \cdot \vec{w}_j - 1)$

$$\frac{\partial L}{\partial \lambda} : w^T w - 1 = 0$$

$$\frac{\partial L}{\partial w} : 2v\vec{w} - 2\lambda\vec{w} = 0$$

$$\Rightarrow \boxed{v\vec{w} = \lambda\vec{w}}$$

因此  $\vec{w}$  是  $v$  的特征向量,  $\lambda$  是特征值.