Derivation of ksums

有一堆向量S_w，其中一个是x_i，计算x_i与这堆向量的距离：

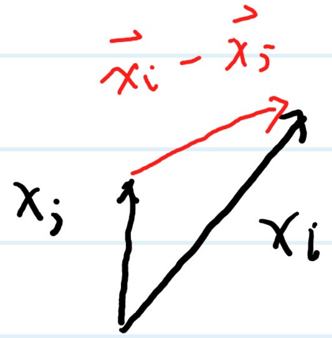
$$d(x_i, S_w) = \sum_{x_j \in S_w} \|x_i - x_j\|^2$$

简化：

$$d(x_i, S_w) = n_w \cdot x_i' \cdot x_i - 2 \cdot x_i' \cdot D_w + \sum_{x_j \in S_w} x_j' \cdot x_j$$

这一简化究竟是怎么来的？

① 从 $\vec{x_i} - \vec{x_j}$ 说起.



由余弦定理

$$(x_i - x_j)^2 = |x_i|^2 + |x_j|^2 - 2|x_i|\cdot|x_j|\cdot\cos\theta$$

$$\|x_i - x_j\|^2 \qquad\qquad \vec{x_i}\cdot\vec{x_j}$$

那么：

$$\sum \|\vec{x_i} - \vec{x_j}\|^2 = \sum x_i^2 + \sum x_j^2 - \sum 2\cdot|\vec{x_i}|\cdot|\vec{x_j}|$$

$$= n\cdot x_i^2 + \sum x_j^2 - 2\cdot|\vec{x_i}|\cdot\sum|\vec{x_j}|$$

向量化表示 $\Rightarrow$ $= n\cdot x_i^{\prime}\cdot x_i + \sum x_j^{\prime}\cdot x_j - 2\cdot x_i^{\prime}\cdot D_w$

$(D_w = \mathrm{sum}(x, 1))$

默认为列向量