

3. 解决对偶问题可使用 SMO 算法.

a. 选取 α_i, α_j

b. 固定 α_i, α_j 以外的参数, 求解式④获得更新后的 α_i, α_j .

c. 注意到只需选取的 α_i, α_j 中有一个不满足 KKT 条件, 目标 function 才会继续增大, 直观来看, KKT 条件违背的程度越大, 则变量更新后可导致的增大, 直观来看, KKT 条件违背的程度越大, 则变量更新后可导致的增大, 目标函数值增幅越大, 于是, SMO 先选取违背 KKT 条件程度最大的变量. 第二个变量应选取一个使目标函数值增长最快的变量, 但由于此两个变量所对应的目标 function 值增幅复杂度过高, 因此 SMO 采用一种启发式: 使选取的两变量所对应样本之间的间隔最大. 一种直观的解释是, 这样的两个变量有很大差别, 与对两个相似的数量进行更新相比, 对它们进行更新会带来目标函数值更大的变化.

d. 对于 b, 对任意支持向量 (\mathbf{x}_s, y_s) 有 $y_s f(\mathbf{x}_s) = 1$

$$y_s = \left(\sum_{i \in S} \alpha_i y_i \mathbf{x}_i^T \mathbf{x}_s + b \right) = 1.$$

$$b = \frac{1}{|S|} \sum_{i \in S} (1/y_i - \sum_{i \in S} \alpha_i y_i \mathbf{x}_i^T \mathbf{x}_s).$$

6.2









