

第七章

支持向量机

张志忠 华东师范大学

- 回顾——上节课的内容
 - 支持向量机的对偶形式
 - 支持向量机的软间隔算法
 - 核技巧
- 本节课的内容：如何高效求解SVM——SMO

例 7.1 数据与例 2.1 相同. 已知一个如图 7.4 所示的训练数据集, 其正例点是 $x_1 = (3,3)^T$, $x_2 = (4,3)^T$, 负例点是 $x_3 = (1,1)^T$, 试求最大间隔分离超平面.

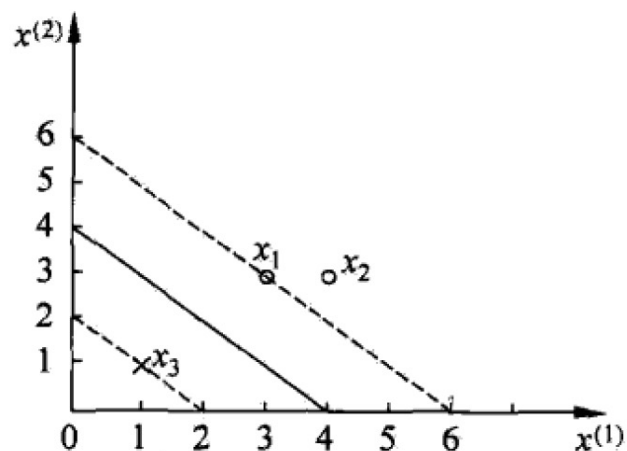


图 7.4 间隔最大分离超平面示例

原始形式:

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2} \|w\|^2 \\ \text{s.t.} \quad & y_i(w \cdot x_i + b) - 1 \geq 0, \quad i=1,2,\dots,N \end{aligned}$$



$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2} (w_1^2 + w_2^2) \\ \text{s.t.} \quad & 3w_1 + 3w_2 + b \geq 1 \\ & 4w_1 + 3w_2 + b \geq 1 \\ & -w_1 - w_2 - b \geq 1 \end{aligned}$$

例 7.1 数据与例 2.1 相同. 已知一个如图 7.4 所示的训练数据集, 其正例点是 $x_1 = (3,3)^T$, $x_2 = (4,3)^T$, 负例点是 $x_3 = (1,1)^T$, 试求最大间隔分离超平面.

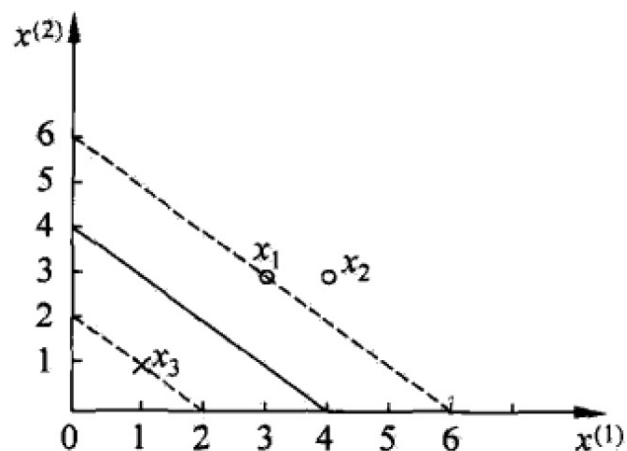


图 7.4 间隔最大分离超平面示例

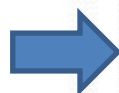
对偶形式:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^N \alpha_i y_i (w \cdot x_i + b) + \sum_{i=1}^N \alpha_i$$

$$\min_{\alpha} \quad \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^N \alpha_i$$

$$\text{s.t.} \quad \sum_{i=1}^N \alpha_i y_i = 0$$

$$\alpha_i \geq 0, \quad i=1,2,\dots,N$$



$$\min_{\alpha} \quad \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^N \alpha_i$$

$$= \frac{1}{2} (18\alpha_1^2 + 25\alpha_2^2 + 2\alpha_3^2 + 42\alpha_1\alpha_2 - 12\alpha_1\alpha_3 - 14\alpha_2\alpha_3) - \alpha_1 - \alpha_2 - \alpha_3$$

$$\text{s.t.} \quad \alpha_1 + \alpha_2 - \alpha_3 = 0$$

$$\alpha_i \geq 0, \quad i=1,2,3$$

原始问题

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2} \|w\|^2 \\ \text{s.t.} \quad & y_i(w \cdot x_i + b) - 1 \geq 0, \quad i=1,2,\dots,N \end{aligned}$$



对偶问题

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^N \alpha_i \\ \text{s.t.} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & \alpha_i \geq 0, \quad i=1,2,\dots,N \end{aligned}$$

拉格朗日函数

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^N \alpha_i y_i (w \cdot x_i + b) + \sum_{i=1}^N \alpha_i$$

求偏导

$$\nabla_w L(w, b, \alpha) = w - \sum_{i=1}^N \alpha_i y_i x_i = 0$$

$$\nabla_b L(w, b, \alpha) = \sum_{i=1}^N \alpha_i y_i = 0$$



$$w = \sum_{i=1}^N \alpha_i y_i x_i$$

$$\sum_{i=1}^N \alpha_i y_i = 0$$

$$\nabla_x L(x^*, \alpha^*, \beta^*) = 0$$

$$\nabla_{\alpha} L(x^*, \alpha^*, \beta^*) = 0$$

$$\nabla_{\beta} L(x^*, \alpha^*, \beta^*) = 0$$

$$\alpha_i^* c_i(x^*) = 0, \quad i=1,2,\dots,k$$

$$c_i(x^*) \leq 0, \quad i=1,2,\dots,k$$

$$\alpha_i^* \geq 0, \quad i=1,2,\dots,k$$

$$h_j(x^*) = 0, \quad j=1,2,\dots,l$$

KKT 条件

原始问题

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2} \|w\|^2 \\ \text{s.t.} \quad & y_i(w \cdot x_i + b) - 1 \geq 0, \quad i=1,2,\dots,N \end{aligned}$$



对偶问题

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^N \alpha_i \\ \text{s.t.} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & \alpha_i \geq 0, \quad i=1,2,\dots,N \end{aligned}$$

软间隔

松弛变量

$$\begin{aligned} \min_{w,b,\xi} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \\ \text{s.t.} \quad & y_i(w \cdot x_i + b) \geq 1 - \xi_i, \quad i=1,2,\dots,N \\ & \xi_i \geq 0, \quad i=1,2,\dots,N \end{aligned}$$

软间隔对偶形式

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^N \alpha_i \\ \text{s.t.} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & \underline{0 \leq \alpha_i \leq C}, \quad i=1,2,\dots,N \end{aligned}$$

常见的核函数:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^N \alpha_i \\ \text{s.t.} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & \alpha_i \geq 0, \quad i=1, 2, \dots, N \end{aligned}$$

多项式核: $K(x, z) = (x \cdot z + 1)^p$

高斯核: $K(x, z) = \exp\left(-\frac{\|x - z\|^2}{2\sigma^2}\right)$

维度转换函数 T

$$\begin{aligned} \vec{x}_i = (x_{i1}, x_{i2}) & \xrightarrow{T} T(\vec{x}_i) = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) \\ \vec{x}_j = (x_{j1}, x_{j2}) & \xrightarrow{T} T(\vec{x}_j) = (x_{j1}, x_{j2}, x_{j3}, \dots, x_{jn}) \end{aligned} \quad \xrightarrow{T(\vec{x}_i) \cdot T(\vec{x}_j) \text{ 点积}} \sum_{m=1}^n x_{im} x_{jm}$$

核函数 $K(\vec{x}_i, \vec{x}_j)$

$$T(\vec{x}_i) \cdot T(\vec{x}_j) = K(\vec{x}_i, \vec{x}_j)$$

直接计算
点积结果

四、序列最小最优化算法

☞ 序列最小最优化(sequential minimal optimization SMO)算法：1998年由Platt提出。

John C. Platt, "Using Analytic QP and Sparseness to Speed Training of Support Vector Machines" in *Advances in Neural Information Processing Systems 11*, M. S. Kearns, S. A. Solla, D. A. Cohn, eds (MIT Press, 1999), 557-63.

☞ 动机：

☞ 支持向量机的学习问题可以形式化为求解凸二次规划问题.这样的凸二次规划问题具有全局最优解，并且有许多最优化算法可以用于这一问题的求解；

☞ 但是当训练样本容量很大时，这些算法往往变得非常低效，以致无法使用.所以，如何**高效地实现支持向量机学习**就成为一个重要的问题。

序列最小最优化算法

☞ SMO (Sequential minimal optimization)

☞ 解如下凸二次规划的对偶问题

$$\min_{\alpha} \quad \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^N \alpha_i$$

$$\text{s.t.} \quad \sum_{i=1}^N \alpha_i y_i = 0$$

$$0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, N$$

☞ 注意：变量是拉格朗日乘子 α_i ，一个对应一个样本

SMO算法

⌘ 启发式算法，基本思路：

⌘ 如果所有变量的解都满足此最优化问题的KKT条件，那么得到解；

⌘ 否则，选择两个变量，固定其它变量，针对这两个变量构建一个二次规划问题，称为子问题，可通过解析方法求解，提高了计算速度。

⌘ SMO算法包括两个部分：

⌘ 求解两个变量二次规划的解析方法

⌘ 选择变量的启发式方法

两个变量二次规划的求解过程

选择两个变量，其它固定，SMO的 5 的子问题：

$$\min_{\alpha_1, \alpha_2} W(\alpha_1, \alpha_2) = \frac{1}{2} K_{11} \alpha_1^2 + \frac{1}{2} K_{22} \alpha_2^2 + y_1 y_2 K_{12} \alpha_1 \alpha_2$$

$$-(\alpha_1 + \alpha_2) + y_1 \alpha_1 \sum_{i=3}^N y_i \alpha_i K_{i1} + y_2 \alpha_2 \sum_{i=3}^N y_i \alpha_i K_{i2}$$

6

$$\text{s.t. } \alpha_1 y_1 + \alpha_2 y_2 = -\sum_{i=3}^N y_i \alpha_i = \zeta \rightarrow \alpha_1 = -y_1 \sum_{i=2}^N \alpha_i y_i \quad \text{只有一个自由变量}$$

$$0 \leq \alpha_i \leq C, \quad i = 1, 2$$

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^N \alpha_i$$

$$\text{s.t. } \sum_{i=1}^N \alpha_i y_i = 0$$

$$0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, N$$

两个变量二次规划的求解过程

选择两个变量，其它固定，SMO的 5 的子问题：

$$\begin{aligned} \min_{\alpha_1, \alpha_2} \quad & W(\alpha_1, \alpha_2) = \frac{1}{2} K_{11} \alpha_1^2 + \frac{1}{2} K_{22} \alpha_2^2 + y_1 y_2 K_{12} \alpha_1 \alpha_2 \\ & -(\alpha_1 + \alpha_2) + y_1 \alpha_1 \underbrace{\sum_{i=3}^N y_i \alpha_i K_{i1}}_{V_1} + y_2 \alpha_2 \underbrace{\sum_{i=3}^N y_i \alpha_i K_{i2}}_{V_2} \\ \text{s.t.} \quad & \alpha_1 y_1 + \alpha_2 y_2 = -\sum_{i=3}^N y_i \alpha_i = \zeta \\ & 0 \leq \alpha_i \leq C, \quad i=1, 2 \end{aligned}$$

$$\alpha_1 = -y_1 \sum_{i=2}^N \alpha_i y_i$$

只有一个自由变量

两个变量二次规划的求解过程

证明： 引进记号

$$v_i = \sum_{j=3}^N \alpha_j y_j K(x_i, x_j)$$

目标函数写成：

$$W(\alpha_1, \alpha_2) = \frac{1}{2} K_{11} \alpha_1^2 + \frac{1}{2} K_{22} \alpha_2^2 + y_1 y_2 K_{12} \alpha_1 \alpha_2 \\ - (\alpha_1 + \alpha_2) + y_1 v_1 \alpha_1 + y_2 v_2 \alpha_2$$

由 $\alpha_1 y_1 = \zeta - \alpha_2 y_2$ 及 $y_i^2 = 1$

$$\alpha_1 = (\zeta - y_2 \alpha_2) y_1$$



两个变量二次规划的求解过程

得到只是 α_2 的函数的目标函数

$$W(\alpha_2) = \frac{1}{2}K_{11}(\zeta - \alpha_2 y_2)^2 + \frac{1}{2}K_{22}\alpha_2^2 + y_2 K_{12}(\zeta - \alpha_2 y_2)\alpha_2 - (\zeta - \alpha_2 y_2)y_1 - \alpha_2 + v_1(\zeta - \alpha_2 y_2) + y_2 v_2 \alpha_2$$

对 α_2 求导

$$\frac{\partial W}{\partial \alpha_2} = K_{11}\alpha_2 + K_{22}\alpha_2 - 2K_{12}\alpha_2 - K_{11}\zeta y_2 + K_{12}\zeta y_2 + y_1 y_2 - 1 - v_1 y_2 + y_2 v_2$$

令其为0:

$$\zeta = \alpha_1^{\text{old}} y_1 + \alpha_2^{\text{old}} y_2$$

$$(K_{11} + K_{22} - 2K_{12})\alpha_2 = y_2(y_2 - y_1 + \zeta K_{11} - \zeta K_{12} + v_1 - v_2)$$

两个变量二次规划的求解过程

将 $\zeta = \alpha_1^{\text{old}} y_1 + \alpha_2^{\text{old}} y_2$ 代入:

$$g(x) = \sum_{i=1}^N \alpha_i y_i K(x_i, x) + b$$

$$E_i = g(x_i) - y_i = \left(\sum_{j=1}^N \alpha_j y_j K(x_j, x_i) + b \right) - y_i, \quad i=1,2$$

$$\begin{aligned} (K_{11} + K_{22} - 2K_{12})\alpha_2^{\text{new,unc}} &= y_2 ((K_{11} + K_{22} - 2K_{12})\alpha_2^{\text{old}} y_2 + y_2 - y_1 + g(x_1) - g(x_2)) \\ &= (K_{11} + K_{22} - 2K_{12})\alpha_2^{\text{old}} + y_2(E_1 - E_2) \end{aligned}$$

将 $\eta = K_{11} + K_{22} - 2K_{12}$ 代入:

$$\alpha_2^{\text{new,unc}} = \alpha_2^{\text{old}} + \frac{y_2(E_1 - E_2)}{\eta}$$

两个变量二次规划的求解过程

∞ 得到定理:

∞ 最优化问题 6 沿约束方向未经剪辑的解:

$$\alpha_2^{\text{new,unc}} = \alpha_2^{\text{old}} + \frac{y_2(E_1 - E_2)}{\eta}$$

$$\eta = K_{11} + K_{22} - 2K_{12} = \|\Phi(x_1) - \Phi(x_2)\|^2$$

∞ 剪辑后的解

$$\alpha_2^{\text{new}} = \begin{cases} H, & \alpha_2^{\text{new,unc}} > H \\ \alpha_2^{\text{new,unc}}, & L \leq \alpha_2^{\text{new,unc}} \leq H \\ L, & \alpha_2^{\text{new,unc}} < L \end{cases}$$

∞ 得到 α_1 的解

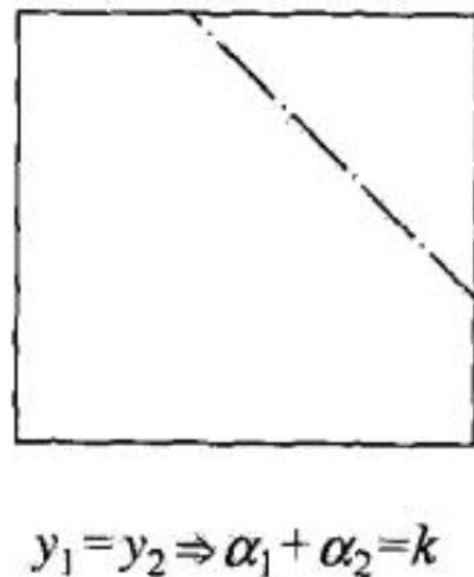
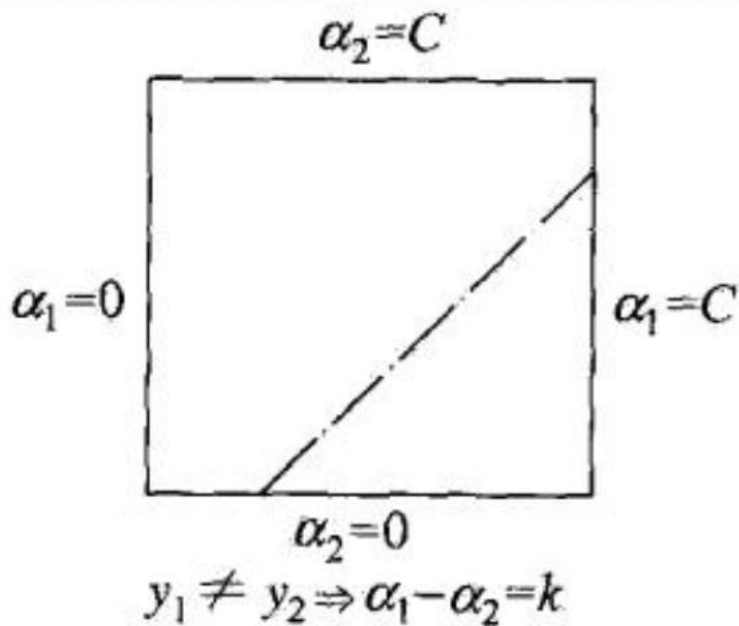
$$\alpha_1^{\text{new}} = \alpha_1^{\text{old}} + y_1 y_2 (\alpha_2^{\text{old}} - \alpha_2^{\text{new}})$$

两个变量二次规划的求解过程

两个变量，约束条件用二维空间中的图形表示

$$\alpha_1 y_1 + \alpha_2 y_2 = -\sum_{i=3}^N y_i \alpha_i = \zeta$$

$$0 \leq \alpha_i \leq C, \quad i=1,2$$



假设问题 6 的初始可行解为 $\alpha_1^{old}, \alpha_2^{old}$ ，最优解 $\alpha_1^{new}, \alpha_2^{new}$

设 α_2 未经剪辑时的最优解为 $\alpha_2^{new,unc}$

Step1:

定理 7.6 最优化问题 (7.101) ~ (7.103) 沿着约束方向未经剪辑时的解是

$$\alpha_2^{\text{new,unc}} = \alpha_2^{\text{old}} + \frac{y_2(E_1 - E_2)}{\eta} \quad (7.106)$$

其中,

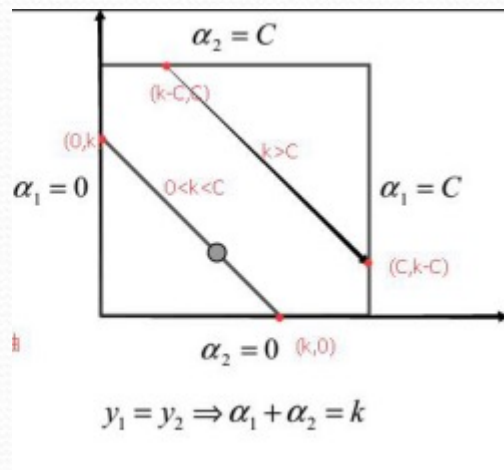
$$\eta = K_{11} + K_{22} - 2K_{12} = \|\Phi(x_1) - \Phi(x_2)\|^2 \quad (7.107)$$

Step2:

经剪辑后 α_2 的解是

$$\alpha_2^{\text{new}} = \begin{cases} H, & \alpha_2^{\text{new,unc}} > H \\ \alpha_2^{\text{new,unc}}, & L \leq \alpha_2^{\text{new,unc}} \leq H \\ L, & \alpha_2^{\text{new,unc}} < L \end{cases}$$

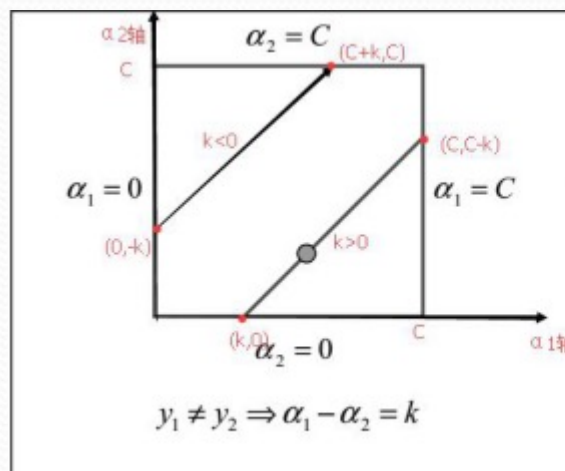
两个变量二次规划的求解过程



>2C 仍然是空集

- ① $\xi < 0$ 时，因为 $0 \leq \alpha_1 \leq C$ ，所以此时 $\alpha_1 + \alpha_2 = \xi$ 与方形区域一定没有任何交集，所以此时 α_2 的可行域为空集。而且也不会出现 $\xi < 0$ 的情况，因为 $0 \leq \alpha_1 \leq C$ 的限制，不可能 $\alpha_1 + \alpha_2$ 结果小于0
- ② $\xi = 0$ 时，此时 $\alpha_1 + \alpha_2 = 0$ ，此时与方形区域的交点就是 $(0, 0)$ ，那么可行域就是 $\alpha_2 = 0$
- ③ $0 < \xi < C$ 时，此时对应上图中右边的靠下的那种直线的情况，所以根据直线和方形区域的相交情况此时可以求出 α_2 的可行区间为 $[0, \xi]$ ，即 $[0, \alpha_1 + \alpha_2]$
- ④ 当 $C \leq \xi \leq 2C$ 时，可以求出此时对应上图右边情况靠上的那种直线，所以此时可以求出的可行区间为 $[\xi - C, C]$ ，即 $[\alpha_1 + \alpha_2 - C, C]$
(上面求可行区间就是看直线和方形区域的交点边界情况，确定的范围，因为既要在直线上，也要在方形区域内)

两个变量二次规划的求解过程



(3) 当 $y_1 = 1$ 且 $y_2 = -1$ 时, 此时的表达式是 $\alpha_1 - \alpha_2 = \xi$, 那么和之前一样此时我们可以根据 ξ 的不同取值来看 α_2 的可行域。

① $\xi > C$ 或者 $\xi < -C$ 时, 很明显此时直线与方形区域没有交点, 所以此时 α_2 可行域为空集

② $0 < \xi \leq C$ 时, 此时对应上面的左图中的靠下的那种直线的情况, 此时可以计算出 α_2 可行域为 $[0, C - \alpha_1 + \alpha_2]$

③ 当 $-C \leq \xi \leq 0$ 时, 此时对应左图中靠上的那种直线的情况, 此时可计算出 α_2 可行域为 $[\alpha_2 - \alpha_1, C]$

综上所述, α_2 的可行域为 $[\max(0, \alpha_1 - \alpha_2), \min(C, C - \alpha_1 + \alpha_2)]$

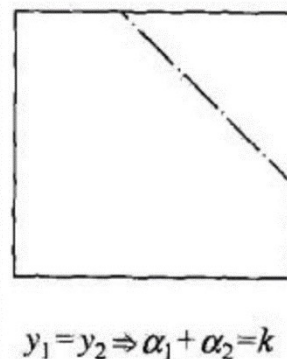
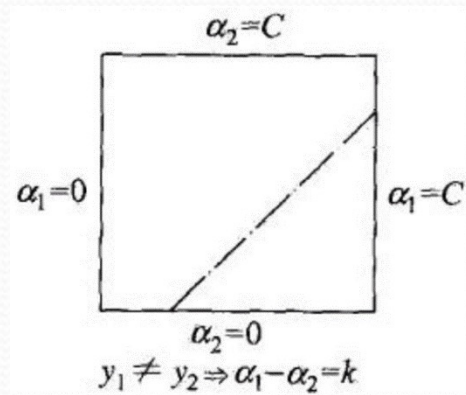
两个变量二次规划的求解过程

根据不等式条件 α_2^{new} 的取值范围:

$$L \leq \alpha_2^{\text{new}} \leq H$$

左图: $L = \max(0, \alpha_2^{\text{old}} - \alpha_1^{\text{old}})$ $H = \min(C, C + \alpha_2^{\text{old}} - \alpha_1^{\text{old}})$

右图: $L = \max(0, \alpha_2^{\text{old}} + \alpha_1^{\text{old}} - C)$ $H = \min(C, \alpha_2^{\text{old}} + \alpha_1^{\text{old}})$



变量的选择方法

- ⌘ SMO算法在每个子问题中选择两个变量优化，其中至少一个变量是违反KKT条件的
- ⌘ 1、第一个变量的选择：外循环
- ⌘ 违反KKT最严重的样本点，
- ⌘ 检验样本点是否满足KKT条件：

$$\alpha_i = 0 \Leftrightarrow y_i g(x_i) \geq 1$$

先检查 $\longrightarrow 0 < \alpha_i < C \Leftrightarrow y_i g(x_i) = 1$

$$\alpha_i = C \Leftrightarrow y_i g(x_i) \leq 1$$

$$g(x_i) = \sum_{j=1}^N \alpha_j y_j K(x_i, x_j) + b$$

变量的选择方法

2、第二个变量的检查：内循环，

选择的标准是希望能使目标函数有足够大的变化

即对应 $|E_1 - E_2|$ 最大，即 E_1, E_2 的符号相反，差异最大

如果内循环通过上述方法找到的点不能使目标函数有足够的下降

则：遍历间隔边界上的样本点，测试目标函数下降

如果下降不大，则遍历所有样本点

如果依然下降不大，则丢弃外循环点，重新选择

计算阈值b和Ei

3、每次完成两个变量的优化后，重新计算b，Ei

由KKT条件：
$$\sum_{i=1}^N \alpha_i y_i K_{i1} + b = y_1$$
 当 $0 < \alpha_1 < 0$ ，W和b用对偶变量代替

$$b_1^{\text{new}} = y_1 - \sum_{i=3}^N \alpha_i y_i K_{i1} - \alpha_1^{\text{new}} y_1 K_{11} - \alpha_2^{\text{new}} y_2 K_{21}$$

$$E_i = g(x_i) - y_i = \left(\sum_{j=1}^N \alpha_j y_j K(x_j, x_i) + b \right) - y_i, \quad i=1,2$$

$$E_1 = \sum_{i=3}^N \alpha_i y_i K_{i1} + \alpha_1^{\text{old}} y_1 K_{11} + \alpha_2^{\text{old}} y_2 K_{21} + b^{\text{old}} - y_1$$

计算阈值b和Ei

3、每次完成两个变量的优化后，重新计算b，Ei

由KKT条件：
$$\sum_{i=1}^N \alpha_i y_i K_{i1} + b = y_1$$

$$b_1^{\text{new}} = y_1 - \sum_{i=3}^N \alpha_i y_i K_{i1} - \alpha_1^{\text{new}} y_1 K_{11} - \alpha_2^{\text{new}} y_2 K_{21}$$

$$E_1 = \sum_{i=3}^N \alpha_i y_i K_{i1} + \alpha_1^{\text{old}} y_1 K_{11} + \alpha_2^{\text{old}} y_2 K_{21} + b^{\text{old}} - y_1$$

$$y_1 - \sum_{i=3}^N \alpha_i y_i K_{i1} = -E_1 + \alpha_1^{\text{old}} y_1 K_{11} + \alpha_2^{\text{old}} y_2 K_{21} + b^{\text{old}}$$

$$b_1^{\text{new}} = -E_1 - y_1 K_{11} (\alpha_1^{\text{new}} - \alpha_1^{\text{old}}) - y_2 K_{21} (\alpha_2^{\text{new}} - \alpha_2^{\text{old}}) + b^{\text{old}}$$

计算阈值 b 和 E_i

∞ 如果: $0 < \alpha_2^{\text{new}} < C$

$$b_2^{\text{new}} = -E_2 - y_1 K_{12}(\alpha_1^{\text{new}} - \alpha_1^{\text{old}}) - y_2 K_{22}(\alpha_2^{\text{new}} - \alpha_2^{\text{old}}) + b^{\text{old}}$$

$$E_i^{\text{new}} = \sum_S y_j \alpha_j K(x_i, x_j) + b^{\text{new}} - y_i$$

S 是所有支持向量 x_j 的集合

SMO算法

输入：训练数据集 $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$

$x_i \in \mathcal{X} = \mathbf{R}^n$ $y_i \in \mathcal{Y} = \{-1, +1\}$, $i = 1, 2, \dots, N$, 精度 ε

输出：近似解 α

(1) 取初值 $\alpha^{(0)} = 0$, 令 $k = 0$

(2) 选取优化变量 $\alpha_1^{(k)}, \alpha_2^{(k)}$, 解析求解两个变量的最优化问题
求得最优解 $\alpha_1^{(k+1)}, \alpha_2^{(k+1)}$, 更新 α 为 $\alpha^{(k+1)}$;

(3) 若在精度 ε 范围内满足停机条件

$$\begin{aligned} \sum_{i=1}^N \alpha_i y_i &= 0 \\ 0 \leq \alpha_i &\leq C, \quad i = 1, 2, \dots, N \end{aligned} \quad y_i \cdot g(x_i) = \begin{cases} \geq 1, & \{x_i \mid \alpha_i = 0\} \\ = 1, & \{x_i \mid 0 < \alpha_i < C\} \\ \leq 1, & \{x_i \mid \alpha_i = C\} \end{cases}$$

则转 (4); 否则令 $k = k + 1$, 转 (2);

$$g(x_i) = \sum_{j=1}^N \alpha_j y_j K(x_j, x_i) + b$$

(4) 取 $\hat{\alpha} = \alpha^{(k+1)}$

SMO算法

☞ LIBSVM: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>



Q & A