作业5

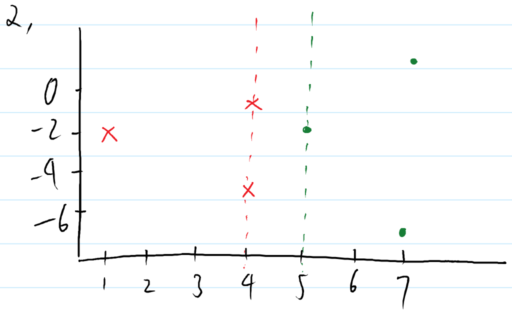
2017年4月18日 8:14

1, $\min_{b,w,\xi} \frac{1}{2}(w^Tw + C\sum_{n=1}^{N}\xi_n)$

s.t. $y_n(w^Tz_n+b) \ge 1-\xi_n$ and $\xi_n \ge 0$

共 d+1+N 个参数，2N 个约束条件 （C 视为超参数）

2,



$\phi_1(x)$	$\phi_2(x)$	y
1	-2	-1
4	-5	-1
4	-1	-1
5	-2	+1
7	-7	+1
7	1	+1
7	1	+1

易知该超平面为 $z_1 = 4.5$

3, 多项式核，应用 sklearn.svm.SVC 计算得

$\alpha = [0.64, 0.76, 0.88, 0.23, 0.28]$

$b = [-1.67]$

SV 下标 [1, 2, 3, 4, 5]

4 方程 $g_{svm} = sign[\sum_{SV} \alpha_n y_n K(x_n, x) + b]$

对应的超平面

$$-0.64(1+x_2)^2 - 0.76(1-x_2)^2 + 0.88(1-x_1)^2 + 0.23(1+2x_2)^2 + 0.28(1-2x_2^2) - 1.67$$

5, 显然不择，实际上，两者的 Z 空间就不一样

对 Q2 $(x_1, x_2) \longrightarrow (x_2^2-2x_1+3,\ x_1^2-2x_2-3)$

对 Q4 $(x_1, x_2) \longrightarrow (1, \sqrt{2}x_1, \sqrt{2}x_2, \sqrt{2}x_1x_2, x_1^2, x_2^2)$

6, $L(R, c, \lambda)$

$= R^2 + \sum_{n=1}^{N}\lambda_n(\|x_n-c\|^2 - R^2)$

7, KKT 条件

① $\|x_n - c\|^2 \le R^2$

② $\lambda_n \ge 0$

③ $\lambda_n(\|x_n-c\|^2 - R^2) = 0$

$\frac{\partial L}{\partial c_i} = 0 \Rightarrow 2\sum_{n=1}^{N}\lambda_n(x_n^{(i)} - c_i) = 0 \Rightarrow c_i\sum_{n=1}^{N}\lambda_n = \sum_{n=1}^{N}\lambda_n x_n^{(i)}$

故 $\sum_{n=1}^{N}\lambda_n \ne 0$ 时 $c = \frac{\sum_{n=1}^{N}\lambda_n x_n}{\sum_{n=1}^{N}\lambda_i}$

$\frac{\partial L}{\partial R} = 0 \Rightarrow 2R - 2R\sum_{n=1}^{N}\lambda_n = 0 \Rightarrow R(1-\sum_{n=1}^{N}\lambda_n) = 0$

8、 $R>0 \Rightarrow \sum_{n=1}^{N}\lambda_n = 1$ $\quad c = \sum_{n=1}^{N}\lambda_n x_n$

则 $L(R, c, \lambda) = \sum_{n=1}^{N}\lambda_n\|x_n - c\|^2 = \sum_{n=1}^{N}\lambda_n\|x_n - \sum_{m=1}^{N}\lambda_m x_m\|^2$ s.t. $\sum_{n=1}^{N}\lambda_n = 1$

9, $L = \sum_{n=1}^{N}\lambda_n\|x_n - \sum_{m=1}^{N}\lambda_m x_m\|^2 = \sum_{n=1}^{N}\lambda_n(x_n - \sum_{m=1}^{N}\lambda_m x_m)^T(x_n - \sum_{m=1}^{N}\lambda_m x_m))$

$= \sum_{n=1}^{N}\lambda_n x_n^Tx_n - 2\sum_{n=1}^{N}\sum_{m=1}^{N}\lambda_n\lambda_m x_m^Tx_n + \sum_{n=1}^{N}\lambda_n\sum_{m=1}^{N}\sum_{k=1}^{N}\lambda_m\lambda_k x_m^Tx_k$

$$= \sum_{n=1}^{N} \alpha_n x_n^T x_n - 2\sum_{n=1}^{N}\sum_{m=1}^{N} \alpha_n \alpha_m x_m^T x_n + \sum_{n=1}^{N} \alpha_n \sum_{m=1}^{N}\sum_{k=1}^{N} \alpha_m \alpha_k x_m^T x_k$$

$$= \sum_{n=1}^{N} \alpha_n k(x_n, x_n) - 2\sum_{n=1}^{N}\sum_{m=1}^{N} \alpha_n \alpha_m k(x_m, x_n) + \sum_{m=1}^{N}\sum_{k=1}^{N} \alpha_m \alpha_k \cdot \sum_{n=1}^{N} \alpha_n$$

$$= \sum_{n=1}^{N} \alpha_n k(x_n, x_n) - \sum_{n=1}^{N}\sum_{m=1}^{N} \alpha_n \alpha_m k(x_m, x_n)$$

10, 当 $\alpha_i > 0$ 时 $\|x_i - c\| - R^2 = 0$

$$R = \sqrt{(x_i - c)^T (x_i - c)} = \sqrt{x_i^T x_i - 2c^T x_i + c^T c}$$

$$= \sqrt{x_i^T x_i - 2\sum_{m=1}^{N} \alpha_m x_m^T x_i + \sum_{n=1}^{N}\sum_{m=1}^{N} \alpha_m \alpha_n x_n^T x_m}$$

$$= \sqrt{k(x_i, x_i) - 2\sum_{m=1}^{N} \alpha_m k(x_m, x_i) + \sum_{n=1}^{N}\sum_{m=1}^{N} \alpha_n \alpha_m k(x_n, x_m)}$$

11, $$\min_{w,b,\xi} \frac{1}{2} w^T w + C\sum_{n=1}^{N} \xi_n^2 = \min_{\tilde{w}, b} \frac{1}{2} \tilde{w}^T \tilde{w}$$

$$\Rightarrow \frac{1}{2}\tilde{w}^T \tilde{w} = \frac{1}{2} w^T w + C\sum_{n=1}^{N} \xi_n^2 = \frac{1}{2} w^T w + \frac{1}{2} (\sqrt{2C}\,\xi)^T (\sqrt{2C}\,\xi)$$

$\Rightarrow \tilde{w} = (w, \sqrt{2C}\,\xi)$ $\quad$ ξ 为 n 维向量

条件函数等价 $y_n(w^T x_n + b) + \xi_n = y_n(\tilde{w}^T \tilde{x}_n + b)$ $\quad$ 记 $\tilde{x}_n = (x_n, v) = (x_n, v_1, v_2 \cdots v_n)$

$$\Rightarrow y_n \sqrt{2C}\, \xi^T v = \xi_n$$

$$\Rightarrow \xi^T v = \frac{\xi_n}{y_n \sqrt{2C}}$$

$$\Rightarrow v_i = [[i = n]] \frac{1}{y_n \sqrt{2C}} \quad i \in [1, N]$$

(如何去掉 y_n?)

12, 考虑核函数对应的 对称半正定矩阵

$K1$: $[k_1(x_i, x_j)]_{m\times m}$ $\quad$ $K2$: $[k_2(x_i, x_j)]_{m\times m}$

a, $k_1 + k_2$ 对应矩阵 $[k_1(x_i, x_j) + k_2(x_i, x_j)]_{m\times m}$

$$= [k_1(x_i, x_j)]_{m\times m} + [k_2(x_i, x_j)]_{m\times m}$$

两个对称半正定矩阵的和 依然是对称半正定矩阵 √

b, 不妨设 $k_2 = 2k_1$ 则 $k_1 - k_2$ 对应矩阵 $-[k_1(x_i, x_j)]$，不是半正定的 ×

c, $k_1 \cdot k_2$ 对应矩阵 $[k_1(x_i, x_j) \cdot k_2(x_i, x_j)]_{m\times m}$

$$\sum_{i,j=1}^{m} c_i c_j k_1(x_i, x_j) \cdot k_2(x_i, x_j)$$

$$= \sum_{i,j=1}^{m} c_i c_j \phi_1(x_i) \phi_1(x_j) \phi_2(x_i) \phi_2(x_j)$$

$$= \left(\sum_i c_i \phi_1(x_i) \phi_2(x_i)\right) \left(\sum_j c_j \phi_1(x_j) \phi_2(x_j)\right)$$

$$= \left\|\sum c_i \phi_1(x_i) \phi_2(x_i)\right\|^2 \geq 0$$ 半正定 √

$= \left(\sum_i c_i \phi_1(x_i) \phi_2(x_i)\right) \left(\sum_j c_j \phi_1(x_j) \phi_2(x_j)\right)$

$= \left\| \sum_i c_i \phi_1(x_i) \phi_2(x_i) \right\|^2 \geq 0$ 半正定 ✓

d, 反例 $k_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ $k_2 = \begin{bmatrix} 4/5 & 1/5 \\ 1/5 & 4/5 \end{bmatrix}$ ×

13, a, $k_1 = \begin{bmatrix} 4/5 & 1/5 \\ 1/5 & 4/5 \end{bmatrix}$ $k = \begin{bmatrix} 1/25 & 16/25 \\ 16/25 & 1/25 \end{bmatrix}$ × 反例

b, 半正定矩阵乘常数 依然半正定 ✓

c, $k_1 = \begin{pmatrix} 4/5 & 1/5 \\ 1/5 & 4/5 \end{pmatrix}$ $k = \begin{bmatrix} e^{-4/5} & e^{-1/5} \\ e^{-1/5} & e^{-4/5} \end{bmatrix}$ × 反例

d, $\sum_{i,j=1}^{m} c_i c_j \frac{1}{1 - k_1(x_i x_j)}$

$= \sum_{i,j=1}^{m} c_i c_j \frac{k_1(x_i x_j)}{(1-k_1(x_i x_j)) k_1(x_i x_j)} > \sum_{i,j=1}^{m} c_i c_j k_1(x_i x_j) \geq 0$ ✓

14, $g_{svm}(x) = \mathrm{sign}\left(\sum_{SV} \alpha_n y_n k(x_n, x) + y_s - \sum_{SV} \alpha_n y_n k(x_n, x_s)\right)$

$k \to \tilde{k}$ $g_{svm}(x) = \mathrm{sign}\left(\sum_{SV} \alpha_n y_n p k(x_n, x) + y_s - \sum_{SV} \alpha_n y_n p k(x_n x_s)\right)$

为了保持结果一致, 令 $\alpha_n' = \alpha_n / p$

而 α 的求解 $\frac{1}{2} \min_{\alpha} \sum_{n=1}^{N} \sum_{m=1}^{N} \alpha_n \alpha_m y_n y_m z_n^T z_m - \sum_{n=1}^{N} \alpha_n$

s.t. $\sum_{n=0}^{N} \alpha_n y_n = 0$

$0 \leq \alpha_n \leq C$

则 α_n 与C成正比, 取 $\tilde{C} = C/p$

15 ~ 20 当见代码

15, $C \uparrow$, $\|w\| \uparrow$

16, $C \uparrow$ E_{in} 不变

17. $C \uparrow$ $\sum_{n=1}^{N} \alpha_n \uparrow$ 与C成正比

18. 对 free SV $d = \frac{1}{\|w\|}$ $\|w\|^2 = \alpha^T k \alpha$ (sklearn. SVC中如何求?)

($C \uparrow$ $\|w\| \uparrow$ $d \downarrow$)

19, $\gamma \uparrow$ E_{out} 先降后升 $\log_{10}\gamma = 1$ 时取得最小值

20, $\gamma \uparrow$ E_{val} 先降后升 $\log_{10}\gamma = 1$ 时取得最小值