作业6

2017年4月21日 9:34

1. $$\frac{\partial F}{\partial A} = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{1+e^{s}}e^{s}\frac{\partial s}{\partial A} = \frac{1}{N}\sum_{n=1}^{N}\frac{e^{s}}{1+e^{s}}(-y_n z_n) = \frac{1}{N}\sum_{n=1}^{N}-y_n z_n p_n$$

$$\frac{\partial F}{\partial B} = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{1+e^{s}}e^{s}\frac{\partial s}{\partial B} = \frac{1}{N}\sum_{n=1}^{N}\frac{e^{s}}{1+e^{s}}(-y_n) = \frac{1}{N}\sum_{n=1}^{N}-y_n p_n$$

$$\nabla F(A,B) = \frac{1}{N}\sum_{n=1}^{N}[-y_n p_n z_n, -y_n p_n]$$

2. $$\frac{\partial^2 F}{\partial A^2} = \frac{1}{N}\sum_{n=1}^{N}-y_n z_n\frac{\partial p_n}{\partial A} = \frac{1}{N}\sum_{n=1}^{N}-y_n z_n p_n(1-p_n)\cdot(-y_n z_n) = \frac{1}{N}\sum_{n=1}^{N}z_n^2 p_n(1-p_n)$$

$$\frac{\partial^2 F}{\partial A\partial B} = \frac{1}{N}\sum_{n=1}^{N}-y_n z_n\frac{\partial p_n}{\partial B} = \frac{1}{N}\sum_{n=1}^{N}-y_n z_n p_n(1-p_n)(-y_n) = \frac{1}{N}\sum_{n=1}^{N}z_n p_n(1-p_n)$$

$$\frac{\partial^2 F}{\partial B^2} = \frac{1}{N}\sum_{n=1}^{N}-y_n\frac{\partial p_n}{\partial B} = \frac{1}{N}\sum_{n=1}^{N}-y_n p_n(1-p_n)(-y_n) = \frac{1}{N}\sum_{n=1}^{N}p_n(1-p_n)$$

$$H(F) = \frac{1}{N}\sum_{n=1}^{N}\begin{bmatrix} z_n^2 p_n(1-p_n) & z_n p_n(1-p_n) \\ z_n p_n(1-p_n) & p_n(1-p_n)\end{bmatrix}$$

3. $$k(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2)$$

x_i, x_j 相异. $\|x_i - x_j\|^2 > 0$ $\gamma\to\infty$ 时, $k(x_i, x_j)\to 0$

$k\to 0$ $\beta = (\lambda I + K)^{-1}y \longrightarrow \frac{1}{\lambda}y$

4. P_1 P_2 分别代表对越界情况由 τ_j 的线性惩罚, 二次惩罚

$\Rightarrow$ P_2 的无约束形式 $\min\limits_{b,w}\frac{1}{2}w^T w + C\sum_{n=1}^{N}\max(0, |w^T z_n + b - y_n| - \varepsilon)^2$

严格证明?

5. $$F(b,\beta) = \frac{1}{2}\sum_{m=1}^{N}\sum_{n=1}^{N}\beta_n\beta_m K(x_n, x_m) + C\sum_{n=1}^{N}\max\left(0, \left|\sum_{m=1}^{N}\beta_m K(x_n, x_m) + b - y_n\right| - \varepsilon\right)^2$$

则 $$\frac{\partial F(b,\beta)}{\beta_m} = \sum_{n=1}^{N}\beta_n K(x_n, x_m) - 2C\sum_{n=1}^{N}[\![\,|s_n - y_n| > \varepsilon\,]\!]$$

6. $\sum_{m=1}^{M} g_t(\tilde{x}_m)\tilde{y}_m = \frac{1}{2}\sum_{m=1}^{M}\left[g_t^2(\tilde{x}_m) + \tilde{y}_m^2 - (g_t(\tilde{x}_m) - \tilde{y}_m)^2 \right]$

$$= \frac{M}{2}(s_t + e_0 - e_t)$$

7. 假设两个训练样本为 (x_1, x_1^2) (x_2, x_2^2)

$$L = (w_1x_1 + w_0 - x_1^2)^2 + (w_1x_2 + w_0 - x_2^2)^2$$

$$\begin{cases} \frac{\partial L}{\partial w_1} = 2(w_1x_1 + w_0 - x_1^2)x_1 + 2(w_1x_2 + w_0 - x_2^2)x_2 = 0 \\ \frac{\partial L}{\partial w_0} = 2(w_1x_1 + w_0 - x_1^2) + 2(w_1x_2 + w_0 - x_2^2) = 0 \end{cases}$$

$\Longrightarrow \begin{cases} w_1 = x_1 + x_2 \\ w_0 = -x_1x_2 \end{cases}$ 故 $g(x) = (x_1 + x_2)x - x_1x_2$

$$\bar{g}(x) = E[(x_1 + x_2)x - x_1x_2] = x - \frac{1}{4}$$

8. $u_n(y_n - w^Tx_n)^2 = (\sqrt{u_n}y_n - w^T\sqrt{u_n}x_n)^2$

故可取 $(\tilde{x}_n, \tilde{y}_n) = (\sqrt{u_n}x_n, \sqrt{u_n}y_n)$

9. $\epsilon_t = \frac{\sum_{n=1}^{N} u_n^{(t)} [[y_n \neq g_t(x_n)]]}{\sum_{n=1}^{N} u_n^t} = \frac{\sum_{n=1}^{N} \frac{1}{N} [[y_n \neq g_t(x_n)]]}{\sum_{n=1}^{N} \frac{1}{N}}$

$$= \frac{[[y_n \neq g_t(x_n)]]}{N} = 0.01$$

$$u_+^{(2)} / u_-^{(2)} = \diamond_t^{-1} / \diamond_t = \frac{\epsilon_t}{1-\epsilon_t} = \frac{1}{99}$$

10. 通过不同维度，不同 θ (取在两点之间)，不同方向，有决策树桩

$2d(R-L)$ 种

此外，尚有全正，全负两种预测

故总数为 $2d(R-L)+2$，本题中为 22 种

11. $K_{ds} = \sum_{j=1}^{|G|} \text{sign}(x_{ij} - \theta_j) \cdot \text{sign}(x'_{ij} - \theta_j)$

$|G| = 2d(R-L) + 2$ 包含了所有的情况

当且仅当 θ_j 位于 x_{ij} 和 x'_{ij} 之间时，符号不同，

考虑不同的 s，每维度 $2|x_i - x'_i|$ 种情况

考虑所有维度，符号不同，取到 -1 的次数为 $\sum_{i=1}^{d} 2|x_i - x'_i| = 2\|x - x'\|_1$

故 k_{ds} = 取+1的数 — 取-1的数 = 总的数 − 2 × 取-1的数

$$= 2d(k-1)+2 - 4\|x-x'\|_1$$

12-20 参见代码

12. $E_{in}(g_1) = 0.24$ $\quad\alpha_1 = 0.58$

13. $E_{in}(g_t)$ 跟随 t 没有什么规律，g_t 只是对上一轮的错误样本敏感，对降低总体 E_{in} 才敏需求

14. $E_{in}(G) = 0.0$

15. $U_2 = 0.854$ $\quad U_T = 0.005$

16. $\min \epsilon_t = 0.179$

17. $E_{out}(g_1) = 0.29$

18. $E_{out}(G) = 0.127$

19, 20.

γ	λ	E_{in}	E_{out}
32	0.001	0.0	0.45
	1	0.0	0.45
	1000	0.0	0.45
2	0.001	0.0	0.44
	1	0.0	0.44
	1000	0.0	0.44
0.125	0.001	0.0	0.46
	1	0.03	0.45
	1000	0.2925	0.39