作业8

2017年5月6日 8:22

1, ① 计算 $x_j^{(l)}$ $(A+1)B+(B+1)\cdot 1$

② 计算 $\delta_j^{(l)}$ $B\cdot 1$

③ 计算 $w_{ij}^{(l)}$ $(A+1)B+(B+1)\cdot 1$

总计算量 $2AB+5B+2$

2~3 $d^{(0)}+=10$ $d^{(L)}=1$

$\sum_{l=1}^{L-1}(d^{(l)}+1)=36$ $(d^{(l)}\geq 1)$

$T=\sum_{l=0}^{L-1}(d^{(l)}+1)\cdot d^{(l+1)}$

使用程序暴力求解， $d^{(1)}=21$ $d^{(2)}=13$ 时，有 $T_{max}=510$
$d^{(1)}=d^{(2)}=\cdots d^{(18)}=1$ 时，有 $T_{min}=46$

暂未找到更好的最优化方法

4, $err_n(w)=\|x_n-ww^Tx_n\|^2$

$=(x_n-ww^Tx_n)^T(x_n-ww^Tx_n)$

$=(x_n^T-x_n^Tww^T)(x_n-ww^Tx_n)$

$=x_n^Tx_n-2x_n^Tww^Tx_n+x_n^Tww^Tww^Tx_n$

记 $a=x_n^Tw=w^Tx_n$ 易知 $\frac{\partial a}{\partial w}=x_n$

$b=w^Tw$ $\frac{\partial b}{\partial w}=2w$

$\nabla_w err_n(w)=\nabla_w(-2a^2+a^2b)$

$=(-4a+2ab)\frac{\partial a}{\partial w}+a^2\frac{\partial b}{\partial w}$

$=-4w^Tx_nx_n+2w^Tx_nw^Twx_n+2(w^Tx_n)^2w$

5, $E_{in}(w)=\frac{1}{N}\sum_{n=1}^{N}\|x_n-ww^T(x_n+\varepsilon_n)\|^2$

$=\frac{1}{N}\sum_{n=1}^{N}(x_n-ww^T(x_n+\varepsilon_n))^T(x_n-ww^T(x_n+\varepsilon_n))$

$=\frac{1}{N}\sum_{n=1}^{N}(x_n-ww^Tx_n-ww^T\varepsilon_n)^T(x_n-ww^Tx_n-ww^T\varepsilon_n)$

$=\frac{1}{N}\sum_{n=1}^{N}(x_n-ww^Tx_n)^T(x_n-ww^Tx_n)-\varepsilon_n^Tww^T(x_n-ww^Tx_n)$

$-(x_n-ww^Tx_n)ww^T\varepsilon_n+\varepsilon_n^Tww^Tww^T\varepsilon_n$

$\varepsilon_n\sim N(0,1)$

$$-(x_n - ww^T x_n) ww^T \varepsilon_n + \varepsilon_n^T ww^T ww^T \varepsilon_n$$

$\varepsilon_n \sim N(0,1)$

$$\mathcal{E}(E_{in}(w)) = \frac{1}{N}\sum_{n=1}^{N} \|x_n - ww^T x_n\| + \mathcal{E}\left(\frac{1}{N}\sum_{n=1}^{N} \varepsilon_n^T ww^T ww^T \varepsilon_n\right)$$

$$\Omega(w) = \mathcal{E}\left(\frac{1}{N}\sum_{n=1}^{N} \varepsilon_n^T ww^T ww^T \varepsilon_n\right)$$

$$= \frac{1}{N}\sum_{n=1}^{N} \mathcal{E}(\varepsilon_n^T ww^T ww^T \varepsilon_n)$$

$$= w^T w \cdot \frac{1}{N}\sum_{n=1}^{N} \mathcal{E}(\varepsilon_n^T ww^T \varepsilon_n)$$

$$= w^T w \cdot \frac{1}{N}\sum_{n=1}^{N} \mathcal{E}\left(\sum_{i=1}^{d} \varepsilon_{ni}^2 w_i^2\right)$$

$$= w^T w \cdot \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{d} w_i^2 \mathcal{E}(\varepsilon_{ni}^2)$$

$$= w^T w \cdot \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{d} w_i^2 (\mathcal{E}^2(\varepsilon_{ni}) + var(\varepsilon_{ni}))$$

$$= w^T w \cdot \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{d} w_i^2 = w^T w \cdot \frac{1}{N}\sum_{n=1}^{N} w^T w = (w^T w)^2$$

6. $g_{LIN}(x) = sign(\|x - x_-\|^2 - \|x - x_+\|^2)$ （距离 x_+ 较小，为 +1）

$$= sign((x - x_-)^T(x - x_-) - (x - x_+)^T(x - x_+))$$

$$= sign(2(x_+ - x_-)^T x + \|x_-\|^2 - \|x_+\|^2)$$

故 $w = 2(w_+ - w_-)$ $\quad b = \|x_-\|^2 - \|x_+\|^2$

7. $\beta_+ \exp(-\|x - \mu_+\|^2) + \beta_- \exp(-\|x - \mu_-\|^2) \geq 0$

$$\Rightarrow \frac{\exp(-\|x - \mu_+\|^2)}{\exp(-\|x - \mu_-\|^2)} \geq -\frac{\beta_-}{\beta_+}$$

$$\Rightarrow \|x - \mu_-\|^2 - \|x - \mu_+\|^2 \geq \ln -\frac{\beta_-}{\beta_+}$$

$$\Rightarrow 2(\mu_+ - \mu_-)^T x + \|\mu_-\|^2 - \|\mu_+\|^2 - \ln(-\frac{\beta_-}{\beta_+}) \geq 0$$

故 $w = 2(\mu_+ - \mu_-)$

$b = \|\mu_-\|^2 - \|\mu_+\|^2 - \ln(-\frac{\beta_-}{\beta_+})$

8. $RBF(x, \mu) = [[x = \mu]]$

则 $z_n = [RBF(x_n, x_1), RBF(x_n, x_2) \cdots RBF(x_n, x_N)]$

$$= [0, 0 \cdots \underset{\uparrow_{n}}{1}, \cdots 0]$$

故 $Z = I$

$\beta = (Z^T Z)^{-1} Z^T y = y$ $\quad$ 即 $\beta_n = y_n$

9. 损失函数 $\sum_{m=1}^{M}\left(\sum_{(x_n,r_{nm})\in D_m}(r_{nm}-w_m^T v_n)^2\right)$

对 w_m 求偏导 $\sum_{(x_n,r_{nm})\in D_m} 2(r_{nm}-w_m^T v_n)(-v_n)$

注意到 $v_n=1$，令上式为0，则 $w_m=\frac{\sum_n r_{nm}}{\sum_n 1}$

故 w_m 为第 m 部电影评分的均值

10. $\max_m v_{N+1}^T w_m = \max_m \left(\frac{1}{N}\sum_{n=1}^{N} v_n\right)^T w_m$

$= \max_m \frac{1}{N}\sum_{n=1}^{N} v_n^T w_m = \max_m \frac{1}{N}\sum_{n=1}^{N} r_{nm}$

即为 m 部电影中，平均打分最高的

11～18 参见代码

12. $E_{in} = (0.0, 0.1, 0.16, 0.15, 0.14)$ 先升后降

14. $E_{out} = (0.394, 0.288, 0.316, 0.322, 0.303)$ 比较稳定

16. $E_{in} = (0.45, 0.45, 0.02, 0.0, 0.0)$ 下降

18. $E_{out} = (0.467, 0.448, 0.288, 0.396, 0.344)$ 先降后升

19～20 参见代码

20 $E_{in} = (2.676, 2.268, 1.897, 1.801, 1.662)$ 逐渐减小