Machine Learning

Assignment #2

Instructor: Beilun Wang   Name:Yue Yuanhao   (ID: 0

Problem 1

(1) $l(\theta) = logL(\theta) = mlog\frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^{m}(y^{(i)} - \theta^T X^{(i)})^2$

(2)$logL(\theta) = mlog\frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{2\sigma^2} \cdot ||y - \theta^T X||^2$

$l(\theta)\ has\ maximum\ value\ when\ ||y - \theta^T X||^2\ has\ minimum\ value$

(3) $suppose\ v\ is\ a\ vector\ ,\ M = X^T X + \lambda I\ ,\ v^T M v = v^T (X^T X + \lambda I) v$

$= v^T X^T X v + \lambda v^T I v = ||Xv||^2 + \lambda v^T v$

$v^T M v\ is\ positive\ ,\ so\ ,\ M\ is\ Positive\ Definite.$

(4) $J(\theta) = (y - X\theta)^T (y - X\theta) + \lambda \theta^T \theta$

$= \theta^T X^T X \theta - \theta^T X^T y - y^T X \theta + y^T y + \lambda \theta^T \theta$

$when\ \ \triangledown J(\theta) = 2X^T X \theta - 2X^T y + 2\lambda\theta = 0\ ,$

$(X^T X + \lambda I)\theta = X^T y$

(5)

Problem 2

(1) $f(y) - f(x) \le f'(x)(y - x) + \frac{\beta}{2}(y - x)^2$

$f(y) - f(x) - f'(x)(y - x)$

$= \int_0^1 f'(x + t(y - x))(y - x)dt - f'(x)(y - x)$

$= \int_0^1 (f'(x + t(y - x)) - f'(x))(y - x)dt$

$\le \int_0^1 \beta t (y - x)^2 dt$

$= \frac{\beta}{2}(y - x)^2$

(2) $f(x_{k+1}) - f(x_k) \le f'(x_k)(x_{k+1} - x_k) + \frac{\beta}{2}(x_{k+1} - x_k)^2$

$= -\eta (f'(x_k))^2 + \frac{\beta\eta^2}{2}(f'(x_k))^2$

$so\ ,\ f(x_{k+1}) \le f(x_k) - \eta(1 - \frac{\eta\beta}{2})(f'(x_k))^2$

(3) $from\ (2)\ and\ \eta = \frac{1}{\beta}\ we\ know\ ,$

$(f^{'}(x_k))^2 \leq 2\beta(f(x_k) - f(x_{k+1}))$

$\sum_{k=1}^{K}(f^{'}(x_k))^2 \leq 2\beta(f(x_1) - f(x_{k+1}))$

$the\ left\ series\ is\ convergent\ so\ ,\ \ \lim_{k->\infty}(f^{'}(x_k))^2 = 0\ \ thus,$

$\lim_{k->\infty} f^{'}(x_k) = 0$

$for\ \forall k \leq K : 2\beta(f(x_k) - f(x_{k+1})) \geq (f^{'}(x_k))^2 \geq 0$

$f(x_{k+1}) \leq f(x_k)\ so\ \lim_{K->\infty} f(x_k) = x^*$

(4) $f(y) - f(x) \geq f^{'}(x)(y-x) + \frac{1}{2\beta}(f^{'}(y) - f^{'}(x))^2$

$let\ z = y - \frac{1}{\beta}(f^{'}(y) - f^{'}(x))$

$f(x) - f(y) = f(x) - f(z) + f(z) - f(y)$

$\leq f^{'}(x)(x-z) + f^{'}(y)(z-y) + \frac{\beta}{2}(z-y)^2$

$= f^{'}(x)(x-y) + (f^{'}(x) - f^{'}(y))(y-z) + \frac{1}{2\beta}(f^{'}(x) - f^{'}(y))^2$

$= f^{'}(x)(x-y) - \frac{1}{2\beta}(f^{'}(x) - f^{'}(y))^2$

$so\ f(y) - f(x) \geq f^{'}(y-x) + \frac{1}{2\beta}(f^{'}(x) - f^{'}(y))^2$

Problem 3

(1) $K_{ij} = k(u_i, v_j) = f(u_i)f(v_j)$

$z^T K z = \sum_i \sum_j z_i K_{ij} z_j$

$= \sum_i \sum_j z_i f(u_i)^T f(v_j)$

$= \sum_k \sum_i \sum_j z_i f(u_i) f(v_i) z_j$

$= \sum_k (\sum z_i f(u^i) f(v^j))^2 \geq 0$

(2) $c_{ij} = k_1(i,j)\ \ d_{ij} = k_2(i,j)$

$e_{i,j} = c_{ij} d_{ij}$

$c_{ij} = a_i T a_j \quad d_{ij} = b_i^T b_j$

$u^T E u = \sum_{ij} u_i u_j c_{ij} d_{ij}$

$= \sum_{kl} (\sum_i u_i a_{ik} b_{il})^2 \geq 0$

(3)