Project part 2

Task 1.

$$\text{minimize} \quad \overbrace{\frac{1}{K}\sum_{k=1}^{K}\left[\log\left(1+\exp(s^T x_k - r)\right) - y_k(s^T x_k - r)\right]}^{f(s,r)}$$

let $x = [s^T \; r^T]^T$, $h_k = \begin{bmatrix} x_k \\ -1 \end{bmatrix}$

$$f(s,r) = f(x) = \frac{1}{K}\sum \log\left(1+\exp\left([x_k^T \; -1]^T x\right)\right) - \sum \frac{y_k}{K}[x_k^T \; -1]^T x$$

$$= \underbrace{\sum_k \overset{\geq 0}{\frac{1}{K}} \underbrace{\log\Big(\overbrace{1+\underbrace{\exp(\underbrace{h_k^T x_k}_{\text{affin map}})}_{\substack{\text{composition of}\\ \text{cvx (affine map)}\\ \Rightarrow \text{is convex}}}}^{\substack{\text{constant + cvx funct.}\\ \Rightarrow \text{is convex}}}\Big)}_{\substack{\text{composition of cvx and nondecreasing}\\ \text{nondecreasing} \circ \text{CVX}\\ \Rightarrow \text{is convex}}}}_{\substack{\text{sum of convex function weighted}\\ \text{by nonnegative scalars}\\ \Rightarrow \text{is convex}}} + \underbrace{\sum_k \overset{\geq 0}{\frac{y_k}{K}} \underbrace{(-h)^T x_k}_{\substack{\text{affine map}\\ \Rightarrow \text{is convex}}}}_{\substack{\text{sum of convex functions}\\ \text{weighted by nonnegative scalars}\\ \Rightarrow \text{is convex}}}$$

sum of 2 convex function is CVX
$\Rightarrow$ is CVX

$$p(x_k, y_k \mid s, r) = \begin{cases} \dfrac{\exp(s^T x_k - r)}{1+\exp(s^T x_k - r)}, & y_k = 1 \\[2ex] \dfrac{1}{1+\exp(s^T x_k - r)}, & y_k = 0 \end{cases}$$

$$\Rightarrow p(x_k y_k \mid s, r) = \frac{\exp\left(y_k(s^T x_k - r)\right)}{1+\exp(s^T x_k - r)}$$

$$\underset{(s,r)\in\mathbb{R}^2\times\mathbb{R}}{\text{maximize}} \; p(x_1 \cdots x_K, y_1 \cdots y_K \mid s, r) = \underset{(s,r)\in\mathbb{R}^2\times\mathbb{R}}{\text{maximize}} \prod_{k=1}^{K} p(x_k, y_k \mid s, r)$$

$$= \underset{(s,r)\in\mathbb{R}^2\times\mathbb{R}}{\text{maximize}} \prod_{k=1}^{K} \frac{\exp\left(y_k(s^T x_k - r)\right)}{1+\exp(s^T x_k - r)}$$

$$= \underset{(s,r)\in\mathbb{R}^2\times\mathbb{R}}{\text{maximize}} \frac{1}{K}\prod_{k=1}^{K} \log \frac{\exp\left(y_k(s^T x_k - r)\right)}{1+\exp(s^T x_k - r)}$$

$$= \underset{(s,r)\in\mathbb{R}^2\times\mathbb{R}}{\text{maximize}} \frac{1}{K}\sum_{k=1}^{K}\left[y_k(s^T x_k - r) - \log\left(1+\exp(s^T x_k - r)\right)\right]$$

$$= \underset{(s,r)\in\mathbb{R}^2\times\mathbb{R}}{\text{minimize}} \frac{1}{K}\sum_{k=1}^{K}\left[\log\left(1+\exp(s^T x_k - r)\right) - y_k(s^T x_k - r)\right]$$

$$f(x) = \frac{1}{K}\sum_{k=1}\left[\overbrace{\log(\underbrace{1+\exp(\overbrace{h^T x}^{l=l(x)})}_{U=f(x)})}^{g(U)} - y_k(h^T x)\right]$$

$$\nabla f \circ g(a) = \nabla f|_{x=g(a)} \circ \nabla g|_{x=a}$$

$$\frac{\partial f}{\partial x} = \frac{1}{K}\sum_{k=1}^{K}\left[\frac{\partial g}{\partial U}\Big|_{U=f(x)} \frac{\partial f}{\partial x}\Big|_{x} - y_k h\right]$$

$$= \frac{1}{K}\sum_{k=1}^{K}\left[\frac{1}{1+\exp(h^T x)}\left(\left(\frac{\partial f}{\partial l}\Big|_{l=l(x)}\right)^T \frac{\partial l}{\partial x}\Big|_x\right) - y_k h\right]$$

$$= \frac{1}{K}\sum_{k=1}^{K}\left[\frac{\exp(h^T x)}{1+\exp(h^T x)} h - y_k h\right]$$

$(\ln f)$

$\frac{2x}{x^2} \quad \frac{2}{x}$

$$= \frac{1}{K}\sum_{k=1}^{K}\left[\left(\frac{\exp(h^T x)}{1+\exp(h^T x)} - y_k\right)h\right]$$

$$f(x_1, x_2, x_3) = \frac{1}{K}\sum_{k=1}\left(\log(1+\exp(h_1 x_1 + h_2 x_2 + h_3 x_3)) - y_k(h_1 x_1 + h_2 x_2 + h_3 x_3)\right)$$

$$\frac{\partial f}{\partial x_1} = \frac{1}{K}\sum_{k=1}^{K} \frac{h_1 \exp(h_1 x_1 + h_2 x_2 + h_3 x_3)}{1+\exp(h_1 x_1 + h_2 x_2 + h_3 x_3)} - y_k h_1$$

$\vdots$

$$\frac{\partial f}{\partial x} = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{\exp(h^T x)}{1+\exp(h^T x)} - y_k\right)h$$

$$s_1 x_1 + s_2 x_2 = r$$

$$x_2 = \frac{1}{s_2}(r - s_1 x_1)$$

3.

$\phi : \mathbb{R} \longmapsto \mathbb{R}$ $\qquad p(x) = \sum_{k=1}^{K} \phi(a_k^T x)$

$p : \mathbb{R}^3 \longmapsto \mathbb{R}$ $\qquad u = f(x)$

a)

$$Dp(x) = \sum_{k=1}^{K} \left(D_u \phi \big|_{u = a_k^T x} D(a_k^T x) \right) = \sum \dot{\phi}(a_k^T x) a_k^T \qquad \nabla p(x) = D^T p(x) = \sum \dot{\phi}(a_k^T x) a_k$$

$$= [a_1 \cdots a_K] \begin{bmatrix} \dot{\phi}(a_1^T x) \\ \vdots \\ \dot{\phi}(a_K^T x) \end{bmatrix} = A v \quad , \quad A = [a_1, \dots, a_K]$$

$$v^T = [\dot{\phi}(a_1^T x) \cdots \dot{\phi}(a_K^T x)]$$

b) $h(x) : \mathbb{R}^m \longmapsto \mathbb{R} \qquad \phi : \mathbb{R} \longmapsto \mathbb{R}$

$$\frac{\partial^2}{\partial x^2} p(x) = [D^2 p(x)] = D(D^T p(x)) = D \sum \dot{\phi}(a_k^T x) a_k = \sum D(a_k^T u)\big|_{u = \dot{\phi}(a_k^T x)} D(\dot{\phi}(a_k^T x))^T$$

$$= \sum a_k D(\dot{\phi}(s))\big|_{s = a_k^T x} D(a_k^T x) = \sum a_k \ddot{\phi}(a_k^T x) a_k^T$$

$$= A \operatorname{diag}(\ddot{\phi}(a_1^T x), \dots, \ddot{\phi}(a_K^T x)) A^T$$

$$f(x) = \frac{1}{K} \sum_{i=1}^{K} \phi(h_k^T x_k) - \sum \frac{1}{K} y_k (h_k^T x)$$

$$\phi(t) = \log(1 + \exp(t))$$

$$\frac{\partial \phi}{\partial t} = \frac{\exp(t)}{1 + \exp(t)}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2}{\partial x^2} \left(\frac{1}{K} \sum_{i=1}^{K} \phi(h_k^T x_k) \right) - \frac{1}{K} \sum h^T$$

$$\frac{\partial^2 \phi}{\partial t^2} = \frac{\partial}{\partial t} \left(\frac{\exp(t)}{1 + \exp(t)} \right)$$

$$= \frac{1}{K} [$$

$$D^2 f(x) = [h_1 \cdots h_K] \operatorname{diag}\left(\frac{\partial^2 \phi}{\partial t^2}(h_1^T x), \dots, \frac{\partial^2 \phi}{\partial t^2}(h_K^T x) \right) \begin{bmatrix} h_1 \\ \vdots \\ h_K \end{bmatrix}$$

$$= \frac{\exp(t)(1 + \exp(t)) - \exp(t)\exp(t)}{(1 + \exp(t))^2}$$

u_1, u_2

$u = f(g_1, g_2, g$

Let $f : \mathbb{R}^m \longmapsto \mathbb{R}$

$$\frac{\partial^2 \phi}{\partial t^2} = \frac{\exp(t)}{(1 + \exp(t))^2}$$

$U = g : \mathbb{R}^m \longmapsto \mathbb{R}^m$

$(f \circ g)(x) : \mathbb{R}^m \longmapsto \mathbb{R}$

$$\boxed{[D(f \circ g)(x)] = D_x g|_x \left(D_u f|_{u = g(x)} \right)^T}$$

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f}{\partial u_1}\Big|_{u_1 = g_1(x)} \frac{\partial g_1}{\partial x_1} + \cdots + \frac{\partial f}{\partial u_m}\Big|_{u_m = g_N(x)} \frac{\partial g_m}{\partial x_1} \\ \vdots \qquad\qquad \vdots \\ \frac{\partial f}{\partial u_1}\Big|_{u_1 = g_1(x)} \frac{\partial g_1}{\partial x_m} + \cdots + \frac{\partial f}{\partial u_m}\Big|_{u_m = g_N(x)} \frac{\partial g_m}{\partial x_m} \end{bmatrix}_{m \times 1}$$

$$\frac{\partial}{\partial x} = \nabla_x g|_x \left(\nabla f|_{u = g(x)} \right)^T = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_1} \\ \vdots & & \\ \frac{\partial g_1}{\partial x_m} & \cdots & \frac{\partial g_m}{\partial x_m} \end{bmatrix}_{m \times m} \left[\frac{\partial f}{\partial u_1} \cdots \frac{\partial f}{\partial u_m} \right]^T_{m \times 1}$$

$g : \mathbb{R}^m \longmapsto \mathbb{R}^k$

$f : \mathbb{R}^k \longmapsto \mathbb{R}^m$

$(f \circ g)(x) : \mathbb{R}^m \longmapsto \mathbb{R}^m$

$$[D(f\circ g)]_{ij} = \frac{\partial}{\partial x_j}(f\circ g)_i$$

D = jacobiano

$$[Df]_{ij} = \frac{\partial f_i}{\partial x_j}$$

∇ = gradiente

$$[\nabla f]_i = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_m} \end{bmatrix}$$

$$D(f\circ g) = \begin{bmatrix} \left.\frac{\partial f_1}{\partial u_1}\right|_{u_1 = g_1(x)} \frac{\partial g_1}{\partial x_1} + \dots + \left.\frac{\partial f_1}{\partial u_k}\right|_{u_k = g_k(x)} \frac{\partial g_k}{\partial x_1} & \dots & \left.\frac{\partial f_1}{\partial u_1}\right|_{u_1 = g_1(x)} \frac{\partial g_1}{\partial x_m} + \dots + \left.\frac{\partial f_1}{\partial u_k}\right|_{u_k = g_k(x)} \frac{\partial g_k}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \left.\frac{\partial f_m}{\partial u_1}\right|_{u_1 = g_1(x)} \frac{\partial g_1}{\partial x_1} + \dots & & \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{k} \left.\frac{\partial f_1}{\partial u_i}\right|_{u_i = g_i(x)} \frac{\partial g_i}{\partial x_1} & \dots & \sum_{i=1}^{k} \left.\frac{\partial f_1}{\partial u_i}\right|_{u_i = g_i(x)} \frac{\partial g_i}{\partial x_m} \\ \vdots & & \\ \sum_{i=1}^{k} \left.\frac{\partial f_m}{\partial u_i}\right|_{u_i = g_i(x)} \frac{\partial g_i}{\partial x_1} & \dots & \sum_{i=1}^{k} \left.\frac{\partial f_m}{\partial u_i}\right|_{u_i = g_i(x)} \frac{\partial g_i}{\partial x_m} \end{bmatrix}$$

$$[D(f\circ g)]_{ij} = \sum_{s=1}^{k} \left.\frac{\partial f_i}{\partial u_s}\right|_{u_s = g_s(x)} \frac{\partial g_s}{\partial x_j}$$

$$[D f\circ g]_{ij} = [Df]_{is} [Dg]_{sj} \qquad D(f\circ g) = Df \, Dg$$

$$h^* h^T = \begin{bmatrix} x_k \\ -1 \end{bmatrix} [x_k^T \quad -1] = \begin{bmatrix} x_k x_k^T & -x_k \\ -x_k & 1 \end{bmatrix} \ddot{\phi}(a_k^T x_k)$$