Homework - Problem Set #1

Warm Up

1. (a) $f(x) = w f_1(x)$

To show:

$$w f_1(\lambda x + (1-\lambda) y) \leq \lambda w f_1(x) + (1-\lambda) w f_1(y) \quad \text{— (1)}$$

When f_1 is convex, we know that

$$f_1(\lambda x + (1-\lambda) y) \leq \lambda f_1(x) + (1-\lambda) f_1(y) \quad \text{— (2)}$$

Multiplying a constant w on both sides where $w \geq 0$ we get

$$w f_1(\lambda x + (1-\lambda) y) \leq w\left(\lambda f_1(x) + (1-\lambda) f_1(y)\right)$$

$$w f_1(\lambda x + (1-\lambda) y) \leq \lambda w f_1(x) + w(1-\lambda) f_1(y) \quad \text{— (3)}$$

$\Rightarrow$ (1) = (3)

Hence $w f_1(x)$ is convex when f_1 is convex and $w \geq 0$

(b) $f(x) = f_1(x) + f_2(x)$

When f_1 and f_2 are convex functions

$$f_1(\lambda x + (1-\lambda) y) \leq \lambda f_1(x) + (1-\lambda) f_1(y) \quad \text{— (1)}$$

$$f_2(\lambda x + (1-\lambda) y) \leq \lambda f_2(x) + (1-\lambda) f_2(y) \quad \text{— (2)}$$

To show, performing (1) + (2)

$$f_1(\lambda x + (1-\lambda) y) + f_2(\lambda x + (1-\lambda) y) \leq \lambda f_1(x) + (1-\lambda) f_1(y) + \lambda f_2(x) + (1-\lambda) f_2(y)$$

$$\Rightarrow f_1(\lambda x + (1-\lambda) y) + f_2(\lambda x + (1-\lambda) y) \leq \lambda\left(f_1(x) + f_2(x)\right) + (1-\lambda)\left(f_1(y) + f_2(y)\right)$$

$$\Rightarrow f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$$

Hence sum of convex functions is also a convex functions

1. (e) $f(x) = \max\{f_1(x), f_2(x)\}$

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \quad -①$$

Assuming $\max(f_1, f_2)$ yeilds a function f_i
Since f_i is convex

$$f_i(\lambda x + (1-\lambda)y) \leq \lambda f_i(x) + (1-\lambda) f_i(y) \quad -②$$

Since f_i is the max of f_1 and f_2.
② can be rewritten as

$$\max\{f_1(\lambda x + (1-\lambda)y), f_2(\lambda x + (1-\lambda)y))$$

$$\leq \lambda \max\{f_1(x), f_2(x)\} + (1-\lambda) \max\{f_1(y), f_2(y)\} \quad -③$$

① = ③

Hence maximum of 2 convex functions is also convex.

2. (a) $f(x) = \max\{x^2 - 2x, |x|\}$

at $x = 0$:

$f(0) = \max(0, 0) = 0$

Choosing subgradient for $x^2 - 2x$

$$\frac{\partial f}{\partial x} = 2x - 2$$

$$\left.\frac{\partial f}{\partial x}\right|_{x=0} = \underline{\underline{-2}}$$

at $x = -2$:

$f(-2) = \max(6, 2) = 6$

Subgradient for $x^2 - 2x$

$$\frac{\partial f}{\partial x} = 2x - 2$$

$$\left.\frac{\partial f}{\partial x}\right|_{x=-2} = 2(-2) - 2 = \underline{\underline{-6}}$$

(b) $g(x) = \max\{(x-1)^2, (x-2)^2\}$

at $x = 1.5$:

$g(1.5) = \max((0.5)^2, (0.5)^2)$

Choosing subgradient for $(x-1)^2$

$$\frac{\partial g}{\partial x} = 2(x-1) \quad \left.\frac{\partial g}{\partial x}\right|_{x=1.5} = 2(0.5) = \underline{\underline{1}}$$

at $x = 0$:

$g(0) = \max(1, 4)$

Subgradient for $(x-2)^2$

$$\frac{\partial f}{\partial x} = 2(x-2) \quad \left.\frac{\partial f}{\partial x}\right|_{x=0} = \underline{\underline{-4}}$$

Problem 3 :

1. Regression problem of least squares would be

Hypothesis space $= \exp(ax + b)$

$$f = \text{Loss function} = \frac{1}{M} \sum_m \left(\exp(ax^{(m)} + b) - y^{(m)} \right)^2$$

Optimal solution

$$f^* = \min_{a,b} \frac{1}{M} \sum_m \left(\exp(ax^{(m)} + b) - y^{(m)} \right)^2$$

2. Gradient descent uses gradients of the loss function wrt a, b to determine the direction of maximum descent which then can be used to update a and b.

$$a = a - \gamma_t \nabla f_a$$
$$b = b - \gamma_t \nabla f_b$$

where γ_t is the step size
∇f_a, ∇f_b are the corresponding gradients.

$$\nabla f_a = \frac{\partial f}{\partial a} = \frac{1}{M} \sum_m 2x \exp(ax^{(m)} + b) \left(\exp(ax^{(m)} + b) - y^{(m)} \right)$$

$$\nabla f_b = \frac{\partial f}{\partial b} = \frac{1}{M} \sum_m 2 \exp(ax^{(m)} + b) \left(\exp(ax^{(m)} + b) - y^{(m)} \right)$$

3. The optimization problem is convex. since $\exp(ax+b)$ is a convex function $(\exp(ax+b) - y)^2$ is also convex.