[Special Topics] Homework 2
Sudarshana Jagadeeshi
CS 4301.001

1-1.
Slaters condition states that the inequality constraints must be satisfied strictly.
(0.25, 0.25) accomplishes this as 2(0.25)+0.25 < 1, 0.25>0, and 0.25>0.

1-2

27) $L(x,y,\lambda) = (x^2+y^2) + \lambda_1(2x+y-1) + \lambda_2(-x) + \lambda_3(-y)$

$$\frac{\partial L}{\partial x} = 2x + 2\lambda_1 - \lambda_2 = 0 \qquad x = \frac{\lambda_2 - 2\lambda_1}{2}$$

$$\frac{\partial L}{\partial y} = 2y + \lambda_1 - \lambda_3 = 0 \qquad y = \frac{\lambda_3 - \lambda_1}{2}$$

Substituting,

$$\max_{\lambda} \left(\frac{\lambda_2 - 2\lambda_1}{2}\right)^2 + \left(\frac{\lambda_3 - \lambda_1}{2}\right)^2 + \lambda_1\left(\lambda_2 - 2\lambda_1 + \frac{\lambda_3 - \lambda_1}{2} - 1\right) + \lambda_2\left(\frac{2\lambda_1 - \lambda_2}{2}\right) + \lambda_3\left(\frac{\lambda_1 - \lambda_3}{2}\right)$$

$$\frac{\lambda_2^2 - 4\lambda_1\lambda_2 + 4\lambda_1^2}{4} + \frac{\lambda_3^2 - 2\lambda_1\lambda_3 + \lambda_1^2}{4}$$

$$+ \lambda_1\lambda_2 - 2\lambda_1^2 + \frac{\lambda_1\lambda_3 - \lambda_1^2}{2}$$

$$+ \frac{2\lambda_2\lambda_1 - \lambda_2^2}{2} + \frac{\lambda_1\lambda_3 - \lambda_3^2}{2}$$

$$= \lambda_2^2 - 4\lambda_1\lambda_2 + 4\lambda_1^2 + \lambda_3^2 - 2\lambda_1\lambda_3 + \lambda_1^2$$
$$+ 4\lambda_1\lambda_2 - 8\lambda_1^2 + 2\lambda_1\lambda_3 - 2\lambda_1^2$$
$$+ 4\lambda_2\lambda_1 - 2\lambda_2^2 + 2\lambda_1\lambda_3 - 2\lambda_3^2$$
$$= -\lambda_2^2 + 4\lambda_1\lambda_2 + 5\lambda_1^2 - \lambda_3^2 + 2\lambda_1\lambda_3$$

Find $= -5\lambda_1^2 - \lambda_2^2 - \lambda_3^2 + 4\lambda_1\lambda_2 + 2\lambda_1\lambda_3$

subject to $\left(\frac{\lambda_2 - 2\lambda_1}{2}\right)^2 + \left(\frac{\lambda_3 - \lambda_1}{2}\right)^2 = 1$

$$\min_{\lambda} -[-5\lambda_1^2 - \lambda_2^2 - \lambda_3^2 + 4\lambda_1\lambda_2 + 2\lambda_1\lambda_3]$$

subject to $\left(\frac{\lambda_2 - 2\lambda_1}{2}\right)^2 + \left(\frac{\lambda_3 - \lambda_1}{2}\right)^2 - 1 = 0$

$$\mathcal{L}(\lambda_1, \lambda_2, \lambda_3, v) = 5\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 4\lambda_1\lambda_2 - 2\lambda_1\lambda_3 + v\left[\left(\frac{\lambda_2 - 2\lambda_1}{2}\right)^2 + \left(\frac{\lambda_3 - \lambda_1}{2}\right)^2 - 1\right]$$

$$\frac{\partial L}{\partial \lambda_1} = 10\lambda - 4\lambda_2 - 2\lambda_3$$

$\min_{\lambda_1, \lambda_2, \lambda_3} \mathcal{L}$ is the dual

2-1

The gradient w.r.t to each lambda_i is:

(q-sum) * x(i)

https://colab.research.google.com/drive/13gxjrX9OcVPcd1ZwUCdDYuycRGSAJ8sO?usp=sharing

2-2

$$\mathcal{L}(\lambda, v, \mu) = \frac{1}{2} \| q - \sum_{m=1}^{M} \lambda_m x^{(m)} \|_2^2$$

eq ineq

$$+ v \left[\sum_{m=1}^{M} \lambda_m - 1 \right] + \sum_{m=1}^{M} \mu_m \lambda_m$$

$$\frac{\partial L}{\partial \lambda_k} = \frac{1}{2} \| q - x(k) \|_2^2 + v + \mu_k = 0$$

$$= \frac{1}{2} (-2 x_0 q_0$$

$$(q_0 - \lambda_0 x_0)^2 + (q_1 - \lambda_1 x_1)^2 + \cdots + (q_m - \lambda_m x_m)^2$$

$$\cancel{q_0^2} - 2 \cancel{\lambda_0} x_0 q_0 + 2 \lambda_0 x_0^2$$

$$-2 x_0 q_0 + 2 \lambda_0 x_0^2 + v + \mu_k = 0$$

$$\lambda_0 = \frac{2 x_0 q_0 - v - \mu_k}{2 x_0^2}$$

Dual:

$$\mathcal{L}(\lambda, v, \mu) = \frac{1}{2} \| q - \sum_{m=1}^{M} \frac{2 x_m q_m - v - \mu_m}{2 x_m^2} \|_2^2$$

$$+ v \left[\sum_{m=1}^{M} \frac{2 x_m q_m - v - \mu_m}{2 x_m^2} - 1 \right]$$

$$+ \sum_{m=1}^{M} \mu_m \frac{2 x_m q_m - v - \mu_m}{2 x_m^2}$$

3-1

Probably not. The function is not smooth, and there is no neighborhood around a point, so the derivative will not exist.

3-2-a

It ceases to be a subgradient when w takes sufficiently large values. It will no longer underestimate data point y(m).

3-2-b
Rewrite the constraint in order to eliminate the y. The problem then becomes:
maximize over w(argmin over points(y(m)- wT (x(m)-x)))
i.e. we must find the optimal point first and then optimize over w for each iteration.

https://colab.research.google.com/drive/1WIUCNled6KcgJk_OqXxrDkGdy_U4BZgZ?usp=sharing

3-2-c

$\tilde{w}$ $\dot{w}$

Construct the Dual.

$$\max_{w \in \mathbb{R}^n} \operatorname*{argmin}_{m \in \{1 \dots M\}} y^{(m)} - w^T(x^{(m)} - query)$$

$$w_0^2 + w_1^2 + w_2^2 \leq y^2$$

$$\text{subject to } \|w\|_2^2 \leq y^2$$

$$\mathcal{L}(w, \lambda) = y^{(m^*)} - w^T(x^{(m^*)} - x) + \lambda\left[\sum_{m=1}^{M} w_m^2 - \gamma^2\right]$$

$$\frac{\partial L}{\partial w_m} = -(x^{(m^*)} - x) + 2\lambda w_m = 0$$

$$w_m = \frac{(x^{(m^*)} - x)}{2\lambda} \quad \text{or} \quad w = \frac{x - query}{2\lambda}$$

$$\max_{\lambda} g(\lambda) = y^{(m^*)} - \frac{(x^{(m^*)} - query)}{2\lambda}(x^{(m)} - query)$$

$$+ \lambda\left[\sum_{m=1}^{M}\left\|\frac{(x^{(m^*)} - query)}{2\lambda}\right\|^2 - \gamma^2\right]$$

$$= y^{(m^*)} - \frac{\|x^{(m^*)} - query\|_2^2}{2\lambda} + \frac{\|x^{(m^*)} - query\|}{4\lambda}$$

$$- \lambda\gamma^2$$

3-2-d

https://colab.research.google.com/drive/1D0gUWAxJpealGwmBKLkZ1K9WR5ndePfq?usp=sharing