Solution

Friday, November 17, 2023 4:47 PM

Problem 1:

Solution:

(a) No. The inequality is equivalent to

$$x_3 - \sqrt{(x_1-y_1)^2+(x_2-y_2)^2+(x_3-y_3)^2} \leq 0.$$

However, the left-hand-side is a concave function. It is not hard to find a counter example.

(b) Yes.

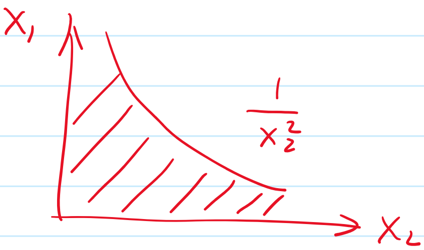
$$X > a \iff v^T(X - aI)v \geq 0$$
for all v

Thus, $\{X \mid X > a\}$

$$= \bigcap_v \{X \mid v^T(X-aI)v \geq 0\}$$

The set $\{X \mid v^T(X-aI)v \geq 0\}$ is convex since it is a halfspace.

(c) No. It looks like this



which is clearly not convex.

(d) Yes.

P_1

(d) Yes.

$$\frac{P_1}{P_2 + N} > 2 \iff 2P_2 - P_1 + N < 0$$

It is a convex set since $2P_2 - P_1 + N$ is a convex (actually linear) function.

(e) Yes.

KL distance is a convex function in both $\vec{p}$ & $\vec{q}$. This can be seen by the fact that $-p_i \log \frac{q_i}{p_i}$ is the perspective mapping of $-\log q_i$, which is convex.

Problem 2:

Solution: Let $z_1 \in A - B$. Then, there must exist $x_1 \in A$, $y_1 \in B$ such that $z_1 = x_1 - y_1$.

Similarly, let $z_2 \in A - B$, such that $z_2 = x_2 - y_2$, $x_2 \in A$ & $y_2 \in B$.

Consider the convex combination, $0 \le \theta \le 1$:

$$z = z_1 \cdot \theta + z_2 (1 - \theta).$$

$$= (x_1 - y_1)\theta + (x_2 - y_2)(1 - \theta)$$

$$= \left[x_1 \theta + x_2 (1-\theta)\right] - \left[y_1 \theta + y_2 (1 - \theta)\right]$$

Since A is convex, $x_1 \theta + x_2 (1-\theta) \in A$

Since B is convex, $y_1 \theta + x_2 (1 - \theta) \in B$

Thus, z must belong to $A - B$.

We can then conclude that $A - B$ is a convex set.

Problem 3:

(a) No. The first summation is equal to

$$\sum_{i: y_i = 1} w^T x_i - \underbrace{\log\left(1+e^{w^T x_i}\right)}_{\text{convex}},$$

which is concave in w. Similarly, the second term is also concave

(b) No. When y is large,

$$\log\left(\frac{y^2}{x}+1\right) \approx \log \frac{y^2}{x} = 2\log y - \log x$$

which is concave in y. Thus, the function cannot be convex in both x & y

(c) Yes. The function

$$\theta x - \log M_x(\theta)$$

is linear in x for every θ. Thus, $I(x)$ is convex in x since it is the pointwise maximum (over θ) of linear functions

(d) Yes. $\sqrt{xy} = x\sqrt{\frac{y}{x}}$ is concave since it is the perspective mapping of $\sqrt{y}$.

Hence, $-\sqrt{x(x+y)}$ is a convex function.

(e) Yes. This follows from the composition rules.

- e^y is convex and increasing

- x^2 is convex

Problem 4:

Solution:

(a) The Lagrangian is

$$L(\vec{y}, \lambda) = \sum_{t=1}^{T} g(x_t + y_t) + \lambda \left(C - \sum_{t=1}^{T} y_t \right)$$

$$= \sum_{t=1}^{T} \left[g(x_t + y_t) - \lambda y_t \right] + \lambda C$$

Thus, the KKT condition for optimality is

(i) $\sum_{t=1}^{T} y_t \geq C$, $y_t \geq 0$

(ii) $\lambda \in R$

(iii) y_t should minimize $g(x_t + y) - \lambda y$ over $y_t \geq 0$

(iv) $\lambda \left(\sum_{t=1}^{T} y_t - c \right) = 0$

(b) Consider any time-slot t such that $y_t > 0$.

The optimality condition for (iii) states that

$$g'(x_t + y_t) - \lambda = 0$$

Since $g''(z) > 0$ for all z, the value of z such that $g'(z) = \lambda$ is unique.

Therefore, the value of $x_t + y_t$ must be the same for all time-slots such that $y_t > 0$.

Problem 5:

Solution:

(a) The Lagrangian is

$$L(a_1, a_2, \lambda_1, \lambda_2) = a_1 x_1 + a_2 x_2 + \lambda_1 (2a_1 + a_2 - 4) + \lambda_2 (a_1 + 2a_2 - 5)$$

$$= a_1 (x_1 + 2\lambda_1 + \lambda_2) + a_2 (x_2 + \lambda_1 + 2\lambda_2) - 4\lambda_1 - 5\lambda_2$$

To minimize L over a_1, a_2, we get

To minimize L over a_1, a_2, we get

$$\min_{a_1,a_2} L(a_1,a_2,\lambda_1,\lambda_2) = \begin{cases} -4\lambda_1 - 5\lambda_2, & \text{if } \begin{array}{l} x_1 + 2\lambda_1 + \lambda_2 = 0 \\ x_2 + \lambda_1 + 2\lambda_2 = 0 \end{array} \\ -\infty, & \text{o/w.} \end{cases}$$

Therefore, the dual problem is

$$\max_{\lambda_1,\lambda_2} \quad -4\lambda_1 - 5\lambda_2 \qquad (7)$$

$$\text{sub to} \quad x_1 + 2\lambda_1 + \lambda_2 = 0 \qquad (8)$$

$$x_2 + \lambda_1 + 2\lambda_2 = 0$$

$$\lambda_1 \geq 0, \ \lambda_2 \geq 0$$

(b) By strong duality.

$$\min_{a \in P} a^T x = \max_{\lambda_1,\lambda_2 \in (8)} -4\lambda_1 - 5\lambda_2$$

Thus, the constraint that

$$\min_{a \in P} a^T x \geq 1$$

is equivalent to

$$\max_{\lambda_1,\lambda_2 \in (8)} -4\lambda_1 - 5\lambda_2 \geq 1$$

and is equivalent to

There exists $\lambda_1, \lambda_2 \in (8)$ such that $-4\lambda_1 - 5\lambda_2 \geq 1$

Therefore, (1) is equivalent to

$$\min \quad f_0(x)$$

$$\text{sub to} \quad -4\lambda_1 - 5\lambda_2 \geq 1$$

$$x_1 + 2\lambda_1 + \lambda_2 = 0$$

$$x_2 + \lambda_1 + 2\lambda_2 = 0$$

$$\lambda_1 \geq 0, \ \lambda_2 \geq 0$$

Problem 6:

Solution:

(a) The only non-convex constraint is (10).

Hence we can replace it by

(a) The only non-convex constraint is (10).

However, we can replace it by

$$r_l \le \log\left(1 + \frac{\delta_l P_l}{w_0}\right) \qquad (11)$$

because the optimal solution will still pick the largest r_l, which will make (11) tight with equality.

- Associate the Lagrange multiplier λ_l.

The Lagrangian is

$$L(\vec{x}, \vec{r}, \vec{p}, \vec{\lambda}) = -\sum_s U_s(x_s) + \sum_l P_l + \sum_l \lambda_l \left[\sum_s H_s^l x_s - r_l\right]$$

$$= -\sum_s \left[U_s(x_s) - x_s \cdot \sum_l H_s^l \lambda_l\right] + \sum_l \left[P_l - \lambda_l r_l\right]$$

- Therefore, to minimize L over $\vec{x}, \vec{r}, \vec{p}$ given $\vec{\lambda}$, each flow s can solve

$$x_s(t) = \max \; U_s(x_s) - x_s \sum_l H_s^l \lambda_l(t) \qquad (12)$$

Each link can solve

$$P_l(t) = \min \; P_l - \lambda_l r_l \qquad (13)$$

$$\text{sub to} \quad r_l \le \log\left(1 + \frac{\delta_l P_l}{w_0}\right)$$

$$\& \quad r_l(t) = \log\left(1 + \frac{\delta_l P_l(t)}{w_0}\right).$$

- To solve (12), we can set the derivative to zero (assuming that $x_s \ge 0$):

$$U_s'(x_s) = \sum_l H_s^l \lambda_l(t)$$

$$\Rightarrow \quad x_s(t) = \max\left\{0, U_s'^{-1}\left(\sum_l H_s^l \lambda_l(t)\right)\right\}$$

To solve (13), we note that it is equivalent to

To solve (13), we note that it is equivalent to

$$\min_{p_l} \quad p_l - \lambda_l \log\left(1 + \frac{\delta_l p_l}{w_0}\right)$$

Assuming $p_l > 0$, we can also set the derivative to zero:

$$1 - \lambda_l \frac{\frac{\delta_l}{w_0}}{1 + \frac{\delta_l p_l}{w_0}} = 0$$

$$\Rightarrow p_l(t) = \max\left\{0, \ \lambda_l - \frac{w_0}{\delta_l}\right\}.$$

- Finally, the dual algorithm can update λ_l by

$$\lambda_l(t+1) = \left\{\lambda_l(t) + \gamma\left[\sum_s H_s^l x_s(t) - r_l(t)\right]\right\}^+. \quad (14)$$

(b) Indeed, the above dual algorithm indeed represents a distributed & price-driven operation: Given λ_l

- each flow can solve x_s independently based on (12)
- each link can determine p_l independently based on (13).
- each link can update λ_l based on (14), without knowledge of $U_s(\cdot)$ & rate-power function.

Note: For problem (6), some students update x, p, λ all in the gradient direction:

$$x_s(t+1) = x_s(t) - \alpha \frac{\partial L}{\partial x_s}$$

$$p_l(t+1) = p_l(t) - \alpha \frac{\partial L}{\partial p_l}$$

$$\lambda_l(t+1) = \lambda_l(t) + \alpha \frac{\partial L}{\partial \lambda_l}.$$

However, note that such updates neither maximize nor minimize $L(x, p, \lambda)$. Thus, the convergence and optimality of this type of "primal-dual" algorithm is not obvious.

Instead, this problem asks for the "dual" algorithm, which is stated above. It is optimal since it maximizes the dual objective function.