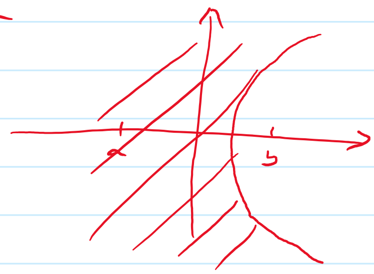
Solution

Wednesday, November 13, 2024 4:16 PM

Problem 1

(a) No. This set looks like



(b) Yes. The matrix $\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$

is positive-definite. Hence, the left-hand-side is a convex function of x.

(c) No. The left-hand-side is a log-sum-exp function of s, which is convex in s. A counterexample is easy to find

(d) Yes. $E[f(X)] = \sum_{i=1}^{10} p_i \cdot f(i)$, which is linear in $\vec{p}$. Hence, the set $\alpha \leq E[f(X)] \leq \beta$ is the intersection of two half-spaces.

(e) Yes. For each v, the set of X such that $v^T X v \leq 1$ is a half-space. The set in question is the intersection of these half-spaces, and hence is convex.

Problem 2.

Take (x_1, z_1) & $(x_2, z_2) \in S$. By the definition of S, there exists y_{11} & y_{12} such that

$$(x_1, y_{11}) \in S_1, \quad (x_1, y_{12}) \in S_2, \quad \& \quad z_1 = y_{11} + y_{12}.$$

Similarly, there exists y_{21} & y_{22} such that

$$(x_2, y_{21}) \in S_1, \quad (x_2, y_{22}) \in S_2, \quad \& \quad z_2 = y_{21} + y_{22}.$$

Now, for any $\theta \in (0,1)$, consider the convex combination

$$\theta(x_1, z_1) + (1-\theta)(x_2, z_2)$$

$$= [\theta x_1 + (1-\theta)x_2, \ \theta z_1 + (1-\theta)z_2].$$

Since S_1 is convex, we have

$$[\theta x_1 + (1-\theta)x_2, \ \theta y_{11} + (1-\theta)y_{21}] \in S_1.$$

Similarly,

$$[\theta x_1 + (1-\theta)x_2, \ \theta y_{12} + (1-\theta)y_{22}] \in S_2.$$

Noting that

$$\theta z_1 + (1-\theta)z_2 = \theta y_{11} + (1-\theta)y_{21} + \theta y_{12} + (1-\theta)y_{22},$$

it implies that

$$[\theta x_1 + (1-\theta)x_2, \ \theta z_1 + (1-\theta)z_2] \in S.$$

The result of the problem then follows.

Problem 3

(a) Yes. For u_1, suppose that $g(u_1)$ is attained by $x_1(u_1)$ & $x_2(u_2)$. In other words

$$g(u_1) = f(x_1(u_1), x_2(u_2)) \ \ \& \ \ x_1(u_1) + 2x_2(u_1) = u_1$$

Similarly, for u_2, we can write

$$g(u_2) = f(x_1(u_2), x_2(u_2)) \ \ \& \ \ x_1(u_2) + 2x_2(u_2) = u_2.$$

For any $\theta \in (0,1)$, consider $u = \theta u_1 + (1-\theta)u_2$. Let

$$x_1 = \theta x_1(u_1) + (1-\theta)x_1(u_2)$$
$$x_2 = \theta x_2(u_1) + (1-\theta)x_2(u_2).$$

Then

$$x_1 + 2x_2 \le \theta u_1 + (1-\theta)u_2 = u$$

In other words, (x_1, x_2) satisfies the constraint for $g(u)$. Hence

$$g(u) \le f(x_1, x_2)$$

$$g(u) \le f(x_1, x_2)$$

$$\le \theta f(x_1(u_1), x_2(u_1)) + (1-\theta) f(x_1(u_2), x_2(u_2))$$

(since f is convex)

$$= \theta g(u_1) + (1-\theta) g(u_2).$$

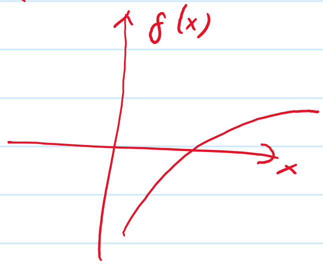
Thus, g is convex in u.

(b) Yes. Because u^2/v is convex, and the mapping $u = (x+y)$ & $v = y$ is affine.

(c) No. Consider the direction of x

$$g(x) \triangleq \frac{x}{x+2y} = 1 - \frac{2y}{x+2y}$$

is a concave function of x.



(d) Yes. $h(u) = \frac{1}{u}$ is convex & non-increasing on $u \ge 0$, & $g(x)$ is concave. Hence, $h(g(x))$ is convex

(e) Yes. For each θ, $\theta x - \log E[\exp(\theta X)]$ is linear in x. The pointwise supremum of linear func is convex.

Problem 4

(a) The optimization problem is

$$\min \quad \sum_{l=1}^{L} P_l$$

$$\text{sub to} \quad \sum_{j=1}^{J(s)} x_{sj} \ge A_s \quad \text{for all } s,$$

$$x_{sj} \ge 0 \quad \text{for all } s, j,$$

$$\sum_{s=1}^{N} \sum_{j=1}^{J(s)} H_{sj}^{l} x_{sj} \le r_l \quad \text{for all } l \qquad (1)$$

$$r_l \le W \log\left(1 + \frac{P_l}{n_0}\right), \quad P_l \ge 0$$

Since $\log(\cdot)$ is concave, and all other objectives &

Since $\log(\cdot)$ is concave, and all other objectives & constraints are linear, this is a convex problem. Note that the last constraint needs to be $\le$. Since we minimize P_L, the optimal solution will automatically make it $=$.

(b) Associate a Lagrange multiplier λ_l to constraint (1). The Lagrangian is

$$L(\vec{x}, \vec{r}, \vec{p}, \vec{\lambda}) = \sum_{l=1}^{L} p_l + \sum_{l} \lambda_l \left(\sum_{s=1}^{N} \sum_{j=1}^{J(s)} H_{sj}^{l} x_{sj} - r_l \right)$$

$$= \sum_{s=1}^{N} \sum_{j=1}^{J(s)} x_{sj} \left(\sum_{l=1}^{L} H_{sj}^{l} \lambda_l \right) + \sum_{l=1}^{L} (p_l - \lambda_l r_l)$$

Thus, to minimize $L(\cdot)$ over the remaining constraints, we can separately solve

- For each flow s,

$$\min \sum_{j=1}^{J(s)} x_{sj} \left(\sum_{l=1}^{L} H_{sj}^{l} \lambda_l \right)$$

$$\text{sub to} \quad \sum_{j=1}^{J(s)} x_{sj} \ge A_s, \quad x_{sj} \ge 0 \qquad (2)$$

- For each link l:

$$\min \quad p_l - \lambda_l r_l$$

$$\text{sub to} \quad r_l \le W \log\left(1 + \frac{p_l}{n_l}\right) \qquad (3)$$

(c) For (2), since the objective is linear, if there is any $x_{sj} > 0$ on a path whose $\left(\sum_{l=1}^{L} H_{sj}^{l} \lambda_l\right)$ is not the smallest, we can reduce the objective of (2) by moving it to the path with the smallest price. The conclusion of part (c) is therefore true.

(d) Given the current $\lambda_l(t)$, each flow s can solve (2) by assigning

$$x_{sj}(t) = A_s$$

on one of the paths j with the smallest total price.

Each link can solve (3), which is equivalent to

$$\min_{p_l \geq 0} \quad p_l - \lambda_l \cdot W \log\left(1 + \frac{p_l}{n_0}\right)$$

Setting the derivative to zero, we get

$$1 - \frac{\lambda_l \cdot W}{1 + \frac{p_l}{n_0}} \cdot \frac{1}{n_0} = 0$$

$$\Rightarrow \quad \lambda_l W = n_0 + p_l$$

$$p_l(t) = \begin{cases} \lambda_l(t) \cdot W - n_0 & \text{if } \lambda_l W - n_0 \geq 0 \\ 0 & \text{o/w} \end{cases}$$

$$\Rightarrow \quad r_l(t) = W \cdot \log\left(1 + \frac{(\lambda_l(t) \cdot W - n_0)^+}{n_0}\right).$$

Finally, the dual gradient algorithm for updating λ_l is

$$\lambda_l(t+1) = \left\{ \lambda_l(t) + \gamma \left[\sum_{S=1}^{N} \sum_{j=1}^{J(S)} x_{Sj}(t) H_{Sj}^l - r_l(t) \right] \right\}^+.$$

Problem 5

(a) The optimization problem is

$$\begin{aligned} \max \quad & \sum_{i=1}^{N} U_{i1} X_{i1} + U_{i2} X_{i2} \\ \text{sub to} \quad & \sum_{i=1}^{N} X_{i1} + 2 X_{i2} \leq W \\ & X_{i1}, X_{i2}, X_{i0} \in \{0, 1\} \\ & X_{i0} + X_{i1} + X_{i2} = 1 \end{aligned}$$

(b) The relaxed problem is

$$\begin{aligned} \max \quad & \sum_{i=1}^{N} U_{i1} X_{i1} + U_{i2} X_{i2} \\ \text{sub to} \quad & \sum_{i=1}^{N} X_{i1} + 2 X_{i2} \leq W \qquad (*) \\ & X_{i0}, X_{i1}, X_{i2} \geq 0 \\ & X_{i0} + X_{i1} + X_{i2} = 1 \end{aligned}$$

Associate a Lagrange multiplier λ to the constraint

Associate a Lagrange multiplier λ to the constraint $(*)$. The Lagrangian is

$$L(\vec{x},\lambda) = -\sum_{i=1}^{N}(u_{i1}x_{i1}+u_{i2}x_{i2}) + \lambda\left[\sum_{i=1}^{N}(x_{i1}+2x_{i2}) - W\right]$$
$$= -\sum_{i=1}^{N}\left[(u_{i1}-\lambda)x_{i1} + (u_{i2}-2\lambda)x_{i2}\right] - \lambda W.$$

Suppose that λ is the optimal dual variable. From the KKT condition, the optimal primal solution should minimize $L(\cdot,\cdot)$. In that case, $\vec{x}$ should minimize, for each i,

$$\max \quad (u_{i1}-\lambda)x_{i1} + (u_{i2}-2\lambda)x_{i2}$$
$$\text{sub to} \quad x_{i1}, x_{i2}, x_{i0} \geq 0$$
$$x_{i0}+x_{i1}+x_{i2} = 1$$

Since the objective is linear, the optimal solution should increase the x_{iw} value with the largest weight.
Thus, we can derive:

- If $u_{i1}-\lambda > u_{i2}-2\lambda$ & $u_{i1}-\lambda>0$, then
  $x_{i1}=1, \quad x_{i2}=x_{i0}=0$ (2)
- If $u_{i1}-\lambda < u_{i2}-2\lambda$ & $u_{i2}-2\lambda>0$, then
  $x_{i2}=1, \quad x_{i1}=x_{i0}=0$ (3)
- If $u_{i1}-\lambda = u_{i2}-2\lambda>0$, then
  $x_{i0}=0, \quad x_{i1}+x_{i2}=1$ (4)
- If $\max\{u_{i1}-\lambda, u_{i2}-2\lambda\}=0$, then
  x_{i0} can be any value ≥ 0 (5)
- If $\max\{u_{i1}-\lambda, u_{i2}-2\lambda\}<0$, then
  $x_{i0}=1, \quad x_{i1}=x_{i2}=0$ (6)

(c) Note that the KKT condition is
- $\vec{x}$ is primal feasible
- $\lambda \geq 0$
- $\vec{x}$ minimizes $L(\vec{x},\lambda)$ over $\vec{x}$
- $\lambda\left(\sum_{i=1}^{N} x_{i1}+2x_{i2}-W\right)=0.$

Suppose that λ is the optimal dual variable, and $\vec{x}$ is the optimal primal variable. Then $\vec{x}$ & λ satisfy the above KKT condition already.

the optimal primal variable. Then $\vec{x}$ & λ satisfy the above KKT condition already.

To find another optimal solution that is mostly 0 or 1, we just need to find another set of $\vec{x}'$ such that it has at most two non-0-and-1 users, and that also satisfies the structure in part (b), and

$$\sum_{i=1}^{N} x_{i1}' + 2x_{i2}' = \sum_{i=1}^{N} x_{i1} + 2x_{i2}. \qquad (**)$$

Clearly, for (2) (3), (6), we can simply take $x_{iw}' = x_{iw}$.

For those users that satisfy (4), let this set be J_1. Let

$$W_1 = \sum_{i \in J_1} x_{i1} + 2x_{i2}$$

denote their total bandwidth consumption. We will shift this W_1 among users in J_1. Specifically, since $x_{i1} + x_{i2} = 1$ for $i \in J_1$, we must have

$$|J_1| \le W_1 \le 2|J_1|.$$

We then set x_{iw}' for $i \in J_1$ as follows. All users $i \in J_1$ starts with $x_{i1}' = 1$ & $x_{i2}' = 0$.

Then, if the remaining bandwidth

$$W_1 - \sum_{i \in J_1} (x_{i1}' + 2x_{i2}') \ge 1,$$

we pick one user $i \in J_1$ with $x_{i1}' = 1$, and change it to $x_{i1}' = 0$ & $x_{i2}' = 1$. This will reduce the remaining bandwidth by 1. Continue doing so, until

$$0 \le \Delta \triangleq W_1 - \sum_{i \in J_1} (x_{i1}' + 2x_{i2}') < 1.$$

If $\Delta > 0$, then pick one more user $i \in J_1$ with $x_{i1}' = 0$, and change it to $x_{i2}' = \Delta$, $x_{i1}' = 1 - \Delta$.

(Note that such users i can always be found. Otherwise, we would have $\sum_{i \in J_1} x_{i1}' + 2x_{i2}' = 2|J_1|$, which is not possible when $\Delta > 0$.)

In this way, only one user $i \in J_1$ will have non-0-and-1 variables, but we get $(**)$.

We can perform a similar procedure for (5), resulting in another user i with non-0-or-1 variables.

In total, there are at most 2 users with non-0-or-1 variables.

(d) From the solution x_{iw}', we assign zero bandwidth to the (at most) 2 users with non-0-or-1 variables. This reduces the utility by at most

$$2 \cdot \max_{i=1,\dots,N,\ w=1,2} U_{iw}.$$

Let this utility by f_0', then

$$f_0' \geq f^* - 2 \max_{i,w} U_{iw}.$$

f_0 should only be better than f_0'. Hence

$$f_0 \geq f^* - 2 \max_{i,w} U_{iw}$$

Finally, since f^* is a relaxation of f_0,

$$f_0 \leq f^*.$$

The result then follows.

Problem 6

(a) The optimization problem is

$$\begin{aligned} \min \quad & r \\ \text{sub to} \quad & r_t \leq r \\ & \sum_{\tau=t}^{T} r_\tau \geq \sum_{\tau=t}^{T} a_\tau \quad \text{for all } t=1,\dots,T. \quad (*) \end{aligned}$$

This is a linear program, and hence convex.

(b) To show convexity of $p(\vec{a})$, take $\vec{a}^1$ & $\vec{a}^2$. By definition of $p(\vec{a}^1)$, there exists $\vec{r}_t^1$ such that

$$p(\vec{a}^1) = \max_t r_t^1 \quad \&$$

$$p(\vec{a}^1) = \max_t r_t^1 \quad \&$$
$$\sum_{\tau=t}^{T} r_\tau^1 \geq \sum_{\tau=t}^{T} a_\tau^1 \quad \text{for all } t=1,\cdots,T.$$

Similarly, for $\vec{a}^2$, there exists $\vec{r_t^2}$ such that

$$p(\vec{a}^2) = \max_t r_t^2 \quad \&$$
$$\sum_{\tau=t}^{T} r_t^2 \geq \sum_{\tau=t}^{T} a_\tau^2 \quad \text{for all } t=1,\cdots,T.$$

Now, for any $\theta \in (0,1)$, consider $\vec{a} = \theta\vec{a}^1 + (1-\theta)\vec{a}^2$.
If we simply take

$$r_t = \theta r_t^1 + (1-\theta) r_t^2,$$

we get

$$\sum_{\tau=t}^{T} r_\tau \geq \sum_{\tau=t}^{T} \{\theta a_\tau^1 + (1-\theta) a_\tau^2\} = \sum_{\tau=t}^{T} a_\tau$$

for all $t=1,\cdots,T$.

In other words, r_t satisfies the constraint for $p(\vec{a})$.
Thus, we have

$$p(\vec{a}) \leq \max_t r_t \leq \theta \max_t r_t^1 + (1-\theta)\max_t r_t^2$$

$$= \theta \cdot p(\vec{a}^1) + (1-\theta) p(\vec{a}^2).$$

Hence, $p(\vec{a})$ is convex in $\vec{a}$.

To show that $p(\vec{a})$ is non-decreasing in each element a_t,
consider $\vec{a}^1$ & $\vec{a}^2$ that differ only by $a_t^1 > a_t^2$.

For $p(\vec{a}^1)$, there exists $\vec{r}^1$ that satisfies (*).
Since $a_t^1 > a_t^2$, the same $\vec{r}^1$ also satisfies (*) for $\vec{a}^2$.
Thus $p(\vec{a}^2) \leq p(\vec{a}^1)$.

(c) For a fixed t, we use the following change of variables:

$$v = \frac{1}{\sum_{\tau=t}^{T} p(\vec{a}(\tau))} \qquad (**)$$

$$u_\tau = \frac{a_\tau}{\sum_{\tau=t}^{T} p(\vec{a}(\tau))}$$

$$u_\tau = \frac{a_\tau}{\sum_{\tau=t}^{T} p(\vec{a}(\tau))}$$

Then, our objective becomes $\sum_{\tau=t}^{T} u_\tau$.

v & u_τ must satisfy $\quad a_\tau = \frac{u_\tau}{v}$

$$\Rightarrow p(\vec{a}(\tau)) = p(a_1, \cdots, a_\tau) = p\left(\frac{u_1}{v}, \cdots, \frac{u_\tau}{v}\right)$$

$$\Rightarrow 1 = v\left(\sum_{\tau=t}^{T} p(\vec{a}(\tau))\right) = \sum_{\tau=t}^{T} v\, p\left(\frac{u_1}{v}, \cdots, \frac{u_\tau}{v}\right)$$

Thus, we can formulate the following optimization problem:

$$\max_{u_\tau \geq 0,\, v \geq 0} \sum_{\tau=t}^{T} u_\tau \qquad (*)$$

$$\text{Sub to } \sum_{\tau=t}^{T} v\, p\left(\frac{u_1}{v}, \cdots, \frac{u_\tau}{v}\right) \leq 1.$$

(Note that it is important to use $\leq$ in the constraint. Otherwise, an equality constraint will not be convex. We will justify below why $\leq$ has no loss of optimality.)

We now show that $(*)$ & (7) are equivalent. First, for any feasible solution to (7), we can use the mapping $(**)$ to get u_τ & v such that it is feasible for $(*)$ (with equality in the constraint). Thus,

$$\text{Opt of (7)} \leq \text{Opt of } (*).$$

Second, for the optimal solution to $(*)$, the constraint must be satisfied with equality. (Otherwise, we can increase u_τ to make the left-hand-side bigger.)

We can then use $a_\tau = \frac{u_\tau}{v}$, which is feasible for (7) and produces the same objective value because

$$\frac{\sum_{\tau=t}^{T} a_\tau}{\sum_{\tau=t}^{T} p(\vec{a}(\tau))} = \frac{\sum_{\tau=t}^{T} \frac{u_\tau}{v}}{\frac{1}{v}} = \sum_{\tau=t}^{T} u_\tau.$$

Hence, $\quad \text{Opt of (7)} \geq \text{Opt of } (*)$.

Putting it together, the two optimization problems are

Hence, opt of ([illegible]) [illegible] opt of ([illegible]).

Putting it together, the two optimization problems are equivalent.

Finally, (*) is convex because $v\,p\left(\frac{\vec{u}}{v}\right)$ is the perspective mapping of $p(\vec{u})$, which is convex.

Alternative solution to Problem 6(c):

① A common mistake is to convert (7) to the following

$$\min \quad \eta$$

$$\text{sub. to} \quad \eta \geq \frac{\sum_{\tau=t}^{T} a_t}{\sum_{\tau=t}^{T} p(\vec{a}(\tau))}$$

- However, this is not equivalent to (7). (7) looks for the maximum over all $\vec{a}$. The above constraint may only hold for one $\vec{a}$. Therefore, η may be much smaller than the solution to (7).

② Instead, this is an equivalent problem to (7)

$$\max \quad \eta$$

$$\text{sub to} \quad \eta \leq \frac{\sum_{\tau=t}^{T} a_\tau}{\sum_{\tau=t}^{T} p(\vec{a}(\tau))}$$

- However, the constraint is not convex.

③ Some students noticed that $p(\lambda\vec{a}) = \lambda p(\vec{a})$. Therefore, there is no loss of optimality forcing

$$\sum_{\tau=t}^{T} a_\tau = 1.$$

(7) is then equivalent to (the inverse of):

(7) is then equivalent to (the inverse of):

$$\min \sum_{\tau=t}^{T} p(\vec{a}(\tau))$$
$$\text{sub to} \sum_{\tau=t}^{T} a_\tau = 1$$

- This is a convex problem since $p(\cdot)$ is convex.

- However, this approach will not work if $p(\lambda\vec{a}) \neq \lambda p(\vec{a})$. My solution above does not require this additional assumption.