EE-588: Optimization for the Information and Data Sciences: Assignment #2

Due on September 27, 2023 at 23:59 pm

Professor Mahdi Soltanolkotabi

Jongmin Mun

2023-09-27 143947".png.bb"

Figure 1: points that vilolate convexity in Problem 3.2

Problem 3.2

Given level curves, there are two main strategy: looking at the sublevel set and looking at the line restriction.

1. The first picture. Let's denote the sublevel set as $S_\alpha : \{x \in dom f \mid f(x) \leq \alpha\}$. Looking at the inclusion relationship, it seems that the sublevel set is the boundary and interior of each curve, and the superlevel sets (which is sublevel set of $-f$) are the whole space with a hole in the middle.
   (a) [convex]: No. We can prove un-convexity by showing un-convexity of one line restriction. For each line restriction, we can prove un-convexity by finding one pair of points that breaks convexity. Note that the level curves have are drawn on uniform interval of y. Therefore points a, b, c in Figure 1, which represent a steep jump followed by a gradual jump, is a valid example for un-convexity.
   (b) [quasiconvex]: Yes, since the sublevel set seems convex.
   (c) [quasiconcave]: No, since the superlevel sets are not convex due to the hole in the middle.
   (d) [concave]: No. Because concave implies quasiconcave, un-quasiconcave implies un-concave.
2. The second picture. Sublevel sets are upper side of the curves, and superlevel sets are lower side of the curves.
   (a) [convex]: No, the sublevel sets are not convex.
   (b) [quasiconvex]: No, Not convex implies not quasiconvex.
   (c) [quasiconcave]: Yes, the superlevel sets seem convex.
   (d) [concave] Yes, all the line restriction seems like to be concave functions, but there might be violations that are not captures by level curves.

Problem 3.3

(from HW1) g is concave.
From high school math, we know graph of the inverse function of f. is $y = x$ symmetric transposition of the graph of f. By drawing a picture, we conjecture that the inverse function of increasing convex function is

concave. We prove using epigraph, since by drawing a simple picture we can conjecture this $y = x$ symmetric transposition will turn epigraph to hypograph and vice versa.

$$\begin{aligned}\textbf{hypp}g &= \{(y,t) : g(y) \le t\} \\ &= \{(y,t) : f(g(y)) \le f(t)\} \text{ (since } f \text{ is increasing)} \\ &= \{(y,t) : y \le f(t)\},\end{aligned}$$

which is just the epigraph of f whose coordinate order is reversed. Reversing the coordinate order is same as multiplying a matrix

$$\begin{pmatrix} 01 \\ 10 \end{pmatrix}, \tag{1}$$

to each point in the set (in fact, any permutation of coordinate is linear transformation) Since f is convex, the epigraph of f is convex. Therefoe, we have proved that the hypograph of g is convex. Thus g is concave.

Problem 3.15

(a) We can see numerator $\to 0$ and denominator $\to 0$. Therefore, by the lHopitals rule,

$$\lim_{\alpha \to 0} u_\alpha(x) = \lim_{\alpha \to 0} \frac{x^\alpha \log x}{1} = \log x.$$

(b) For fixed α, the first derivative is $x^{\alpha-1}$, so since $0 < \alpha < 1$ and $x > 0$, the first derivative is positive, so the function is increasing. The second derivative is $(\alpha - 1)x^{\alpha-2}$, which is always negative because of $\alpha - 1$. Therefore, the function is concave. Finally, $u_\alpha(1) = \frac{1^\alpha - 1}{\alpha} = \frac{1-1}{\alpha} = 0$.

Problem 3.16

(c) The domain $\mathbb{R}^2_{++}$ is an intersection of two open half-spaces, so it is convex.
Since the function is twice differentiable on the whole domain, we use the second order condition of convexity. The Hessian is

$$\frac{1}{x_1x_2}\begin{bmatrix} 2/(x_1^2) & 1/(x_1x_2) \\ 1/(x_1x_2) & 2/x_2^2 \end{bmatrix}.$$

Its determinant is $3/x_1^3x_2^3$, which is strictly positive for any $x \in \mathbb{R}^2_{++}$. Therefore, the Hessian is positive definite for all x in the domain.

Therefore, by the second order condition, f is convex, and not concave. Thus it is quasiconvex, and not quasiconcave.

(d) The domain $\mathbb{R}^2_{++}$ is an intersection of two open half-spaces, so it is convex.
The Hessian is

$$\begin{bmatrix} 0 & -1/x_2^2 \\ -1/x_2^2 & 2x_1/x_2^3 \end{bmatrix} = \frac{1}{x_2^2}\begin{bmatrix} 0 & -1 \\ -1 & 2x_1/x_2 \end{bmatrix} := \frac{1}{x^2}H.$$

Let us calculate the eigenvalues. Solving $\det(\lambda I - A) = 0$, we get $\lambda^2 = x_2/2x_1 > 0$. Therefore we have one positive and one negative eigenvalue. Thus The Hessian is not PSD nor PND.
Therefore, f is not convex, thus not quasiconvex. Also, f is not concave, thus not quasiconcave.

(e) The domain $\mathbb{R} \times \mathbb{R}_{++}$ is an open halfspace, therefore convex.
This function is quadratic-over-linear that we learned in the class as a convex function, but we can verify the convexity through second-order condition: the Hessian is

$$\begin{bmatrix} 2/x_2 & -2x_1/x_2^2 \\ -2x_1/x_2^2 & 2x_1^2/x_2^3 \end{bmatrix} = \frac{2}{x_2} \begin{bmatrix} 1 & -x_1/x_2 \\ -x_1/x_2 & x_1^2/x_2^2 \end{bmatrix} = \frac{2}{x_2} \begin{bmatrix} 1 \\ -x_1/x_2 \end{bmatrix} \begin{bmatrix} 1 & -x_1/x_2 \end{bmatrix},$$

then the quadratic form corresponding to this Hessian will be the $(1/x_2)$ multiplied by square of inner product between z and $(1, -x_1/x_2)$. Since $x_2 > 0$, the Hessian is PSD.
Therefore by the second order condition, f is convex, not concave. Thus it is quasiconvex, not quasiconcave.

(f) The domain $\mathbb{R}^2_{++}$ is an intersection of two open half-spaces, so it is convex.
We again rely on the second order condition. the Hessian is

$$\begin{aligned} \begin{bmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & (1-\alpha)(-\alpha)x_1^{\alpha}x_2^{-\alpha-1} \end{bmatrix} &= \alpha(1-\alpha)x_1^{\alpha}x_2^{1-\alpha} \begin{bmatrix} -1/x_1^2 & 1/x_1x_2 \\ 1/x_1x_2 & -1/x_1^2 \end{bmatrix} \\ &= -\alpha(1-\alpha)x_1^{\alpha}x_2^{1-\alpha} \begin{bmatrix} 1/x_1 \\ -1/x_2 \end{bmatrix} \begin{bmatrix} 1/x_1 & -1/x_2 \end{bmatrix}. \end{aligned}$$

This is the same form as (e), except for the sign of the scalar multiplier being flipped.
Therefore by the second-order condition, f is concave, not convex. Thus it is quasiconcave, not quasiconvex.

3.21

(a) $\|A^{(i)}x - b^{(i)}\|$ is norm (convex) of Affine transformation (convex), so convex. Pointwise maximim of convex functions is convex. Thus f is convex.

(b) Leaving only r element of a vector is a linear transformation (multiplying by an identity matrix, with jth diagonal changed to 0 if we want to elete jth coordniate). We consider $\binom{n}{r}$ all possible case of leaving only r element For each case, indexed by i, we build a function f_i whose pipeline is

leave 5 element (linear, thus convex) $\rightarrow$ calculate ℓ_1 norm (norm is convex).

Thus each function f_i is convex. Thus f, the pointwise maximum of all f_i, is convex.

3.22

(a) The outermost part is $-\log$, which is convex nonincreasing function. So it suffices to show its input (minus-log-sum-exp) is concave.
We can ignore affine mapping. Thus log-sum-exp part is convex, and therefore minus-log-sum-exp part is concave.

(b) Express the given function as

$$-\sqrt{u(v - x^\top x/u)}$$

Then the outermost function is $-\sqrt{x_1x_2}$, which is convex function and nonincreasing for each arguement. Thus it suffices to show its all two inputs are concave.

(i) u is projection, which is affine transformation, thus concave.

(ii) $v - x^\top x/u = v + (-x^\top x/u)$. We want to show each is concave; then the sum of the two is concave.
   - v is projection, thus concave.
   - $x^\top x/u$ is convex, thus $-x^\top x/u$ is concave.

(c) Express f as $-\log u - \log(v - x^\top x/u)$. This is convex becuase

(a) $-\log u$ is convex.

(b) $-\log(v - x^\top x/u)$ is convex, because

i. $-\log x$ is convex nonincreasing

ii. $v - x^\top x/u$ is concave, because

A. v is projection, so concave

B. $-x^\top x/u$ is minus perspective function, thus minus convex, thus concave

(d) Express f as

$$(t^p - \|x\|_p^p)^{1/p} = \left(t^p - t^{p-1}\frac{\|x\|_p^p}{t^{p-1}}\right)^{1/p} = \left(t^{p-1}\left(t - \frac{\|x\|_p^p}{t^{p-1}}\right)\right)^{1/p} = -t^{1-1/p}\left(t - \frac{\|x\|_p^p}{t^{p-1}}\right)^{1/p}.$$

$$= -\left(t - \frac{\|x\|_p^p}{t^{p-1}}\right)^{1/p} t^{1-1/p}.$$

The outermost function is $-x^{1/p}y^{1-1/p}$, whichi is convex and nonincreasing (since there is minus in front and powers of x and y are all positive). Therefore all we need to do is show that two inputs are concave.

(a) $t - \frac{\|x\|_p^p}{t^{p-1}}$ is concave, because

i. t is projection, so concave

ii. $\frac{\|x\|_p^p}{t^{p-1}}$ is convex, so with minus it's concave

(b) t is concave becuase it's projection.

(e) Express f as

$$-\log t^{p-1} - \log(t - \|x\|_p^p/t^{p-1}) = -(p-1)\log t - \log(t - \|x\|_p^p/t^{p-1})$$

This is convex, because

(a) $-(p-1)\log t$ is concave since $p > 1$

(b) $-\log(t - \|x\|_p^p/t^{p-1})$ is convex, because

i. $-\log$ is convex nonincreasing, and

ii. $t - \|x\|_p^p/t^{p-1}$ is convex, because

A. t is projection, thus concave

B. $-\|x\|_p^p/t^{p-1}$ is concave, because $\|x\|_p^p/t^{p-1}$ is convex.

3.49

(c) In order to show log-concavity, we must show that $\log f(x) = \sum_{i=1}^n \log x_i - \log \sum_{i=1}^n x_i$ is concave. The strategy is taking an abstract line restriction and proving convexity in one dimension.
For arbitrary starting point x and direction v in $\mathbb{R}^n$, define

$$\ell_{x,v}(t) := \sum_{i=1}^n \log(x_i + tv_i) - \log\sum_{i=1}^n (x_i + tv_i).$$

We show that this line restriction is convex through the second-order condtion. The derivative is

$$\ell'_{x,v}(t) = \sum_{i=1}^{n} \frac{v_i}{x_i + tv_i} - \frac{\mathbf{1}^\top v}{\mathbf{1}^\top x + t\mathbf{1}^\top v}$$

and the second derivative is

$$\ell''_{x,v}(t) = -\sum_{i=1}^{n} \frac{v_i^2}{(x_i + tv_i)^2} + \frac{(\mathbf{1}^\top v)^2}{(\mathbf{1}^\top x + t\mathbf{1}^\top v)^2}$$

We have to show that the $\ell''_{x,v}(t) \leq 0$ for all t, x, v. However, since x, v is arbitrary, we do not have to show for all t; it would be redundant since we consider all x and v. Therefore we only show

$$\ell''_{x,v}(0) = -\sum_{i=1}^{n} \frac{v_i^2}{v_i^2} + \frac{(\mathbf{1}^\top v)^2}{(\mathbf{1}^\top x)^2} \leq 0.$$

The inequality trivially holds if $\mathbf{1}^\top v = 0$. Thus let's assume $\mathbf{1}^\top v \neq 0$. Then we can see that scaling v has same effect for two terms (multiplying v by k results in multiplying k^2 to the whole). So, without loss of generality,we can scale v by $\frac{(\mathbf{1}^\top x)^2}{(\mathbf{1}^\top v)^2}$. Then the problem now amounts to showing

$$\sum_{i=1}^{n} \frac{v_i^2}{x_i^2} \geq 1.$$

for $x \succeq 0$ and all v.

d

4.11

(a) This is equivalent to

$$\begin{aligned} &\text{minimize } t \\ &\text{subject to } Ax - b \preceq t\mathbf{1} \\ &\qquad\qquad\quad Ax - b \geq -t\mathbf{1} \end{aligned}$$

If we fix x, the kth row of the first and second constraint block indicates $|a_k x - b_k| \leq t$. Thus the whole constraints imply $\|Ax - b\|_\infty = max_k |a_k^\top x - b_k| \leq t$. So for fixed x, the optimal value of LP is $\|Ax - b\|_\infty$, thus optimizing over t, x is equivalent to the original problem.

(b) This is equivalent to

$$\begin{aligned} &min\ \mathbf{1}^\top t \\ &s.t. Ax - b \succeq t \\ &\quad\ \ Ax - b \preceq -t \end{aligned}$$

If we fix x, the kth row of the first and second constraint block indicates $|a_k^\top x - b_k| \leq t_k$. Thus the whole constraints imply $\|Ax - b\|_\infty = max_k |a_k x - b_k| \leq t_k$. Since the objective function of LP is separatble, we achieve the optimim over t by choosing $t_k = |a_k^\top x - b_k|$. So for fixed x, the optimal value of LP is $\|Ax - b\|_1$, thus optimizing over x, t is equivalent to the original problem.

(c) This is equivalent to

$$\begin{aligned} &min\ \mathbf{1}^\top t \\ &s.t. - t \preceq Ax - b \preceq t \\ &- \mathbf{1} \leq x \leq \mathbf{1}. \end{aligned}$$

for $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$.

(d) This is equivalent to

$$\begin{aligned} &min\ \mathbf{1}^\top t \\ &s.t. - \mathbf{1} \preceq Ax - b \preceq \mathbf{1} \\ &- t \leq x \leq t. \end{aligned}$$

(e) This is equivalent to

$$\begin{aligned} &min\ \mathbf{1}^\top y + t \\ &s.t. - y \preceq Ax - b \preceq y \\ &- t\mathbf{1} \preceq x \preceq t\mathbf{1}. \end{aligned}$$