Problem Set 5

December 4, 2023 8:32 PM

Problem 5.1

December 4, 2023 8:32 PM

Problem 5.1 (Proving convexity-preserving operations)

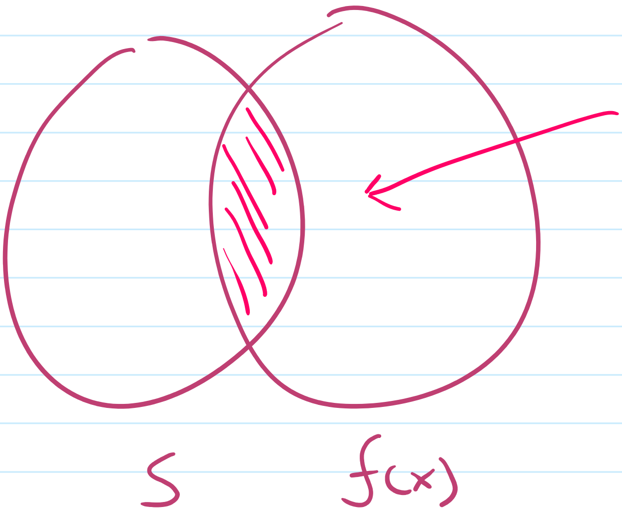
(a) Consider any affine function $f : \mathbb{R}^n \to \mathbb{R}^m$ and convex set $\mathcal{S} \subseteq \mathbb{R}^n$. Prove that the image of $\mathcal{S}$ under f, i.e., $f(\mathcal{S}) = \{f(x) | x \in \mathcal{S}\}$, is a convex set.

(b) Consider any affine function $f : \mathbb{R}^n \to \mathbb{R}^m$ and convex set $\mathcal{S} \subseteq \mathbb{R}^m$. Prove that the *inverse* (or *pre-*) image of $\mathcal{S}$ under f, i.e., $f^{-1}(\mathcal{S}) = \{x | f(x) \in \mathcal{S}\}$, is a convex set.

a)

- Affine sets are convex → from class
- S is an affine set

Image of S under f:

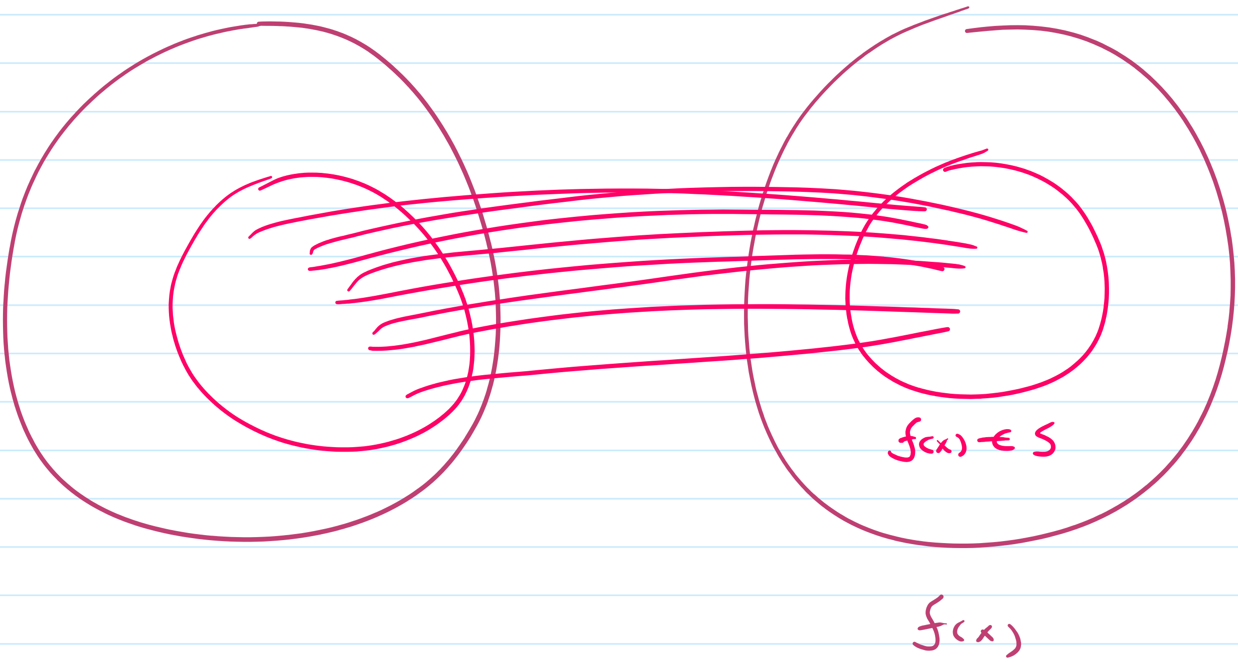


Intersection of convex sets is convex → from class

$\therefore$ Im (S, f) is convex

b) f is still an affine function, which is convex

S is a convex set



Since $f(x)$ is convex, and S (the input) is convex, the output of this process is also convex

Problem 5.2

December 4, 2023 8:33 PM

Problem 5.2 (Identifying convexity)
For each of the functions listed in parts (a)-(c) below identify whether the function is (i) convex, is (ii) quasi-convex, is (iii) concave, is (iv) quasi-concave.

(a) $f(x) = e^x - 1$ where $\mathrm{dom} f = \mathbb{R}$.

(b) $f(x_1, x_2) = x_1 x_2$ where $\mathrm{dom} f = \mathbb{R}^2_{++}$.

(c) $f(x) = 1/(x_1 x_2)$ where $\mathrm{dom} f = \mathbb{R}^2_{++}$.

Recall that a function f is "quasiconvex" if all its sublevel sets $\mathcal{S}_\alpha = \{x \in \mathrm{dom} f | f(x) \leq \alpha\}$ are convex sets, i.e., are convex sets for all $\alpha \in \mathbb{R}$. (Note the empty set is a convex set.) A function f is "concave" if the function $-f$ is convex. A function is "quasiconcave" if $-f$ is quasiconvex; equivalently, a function is quasiconcave if every "superlevel" set $\{x \in \mathrm{dom} f | f(x) \geq \alpha\}$ is a convex set.

If 2nd derivative is ≥ 0, then the function is convex

a) $f(x) = e^x - 1$ where $\mathrm{dom} f = \mathbb{R}$

$f'(x) = e^x \quad f''(x) = e^x \quad e^x \geq 0$

$\therefore f(x)$ is convex

b) $f(x_1, x_2) = x_1 x_2$

$f'(x_1, x_2) = \langle x_2, x_1 \rangle$

$f''(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \det\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) = (0 - 1) = -1$

$\therefore f(x_1, x_2)$ is concave

c) $f(x_1, x_2) = \dfrac{1}{x_1 x_2}$

$f'(x_1, x_2) = \left\langle \dfrac{-1}{x_2 x_1^2}, \dfrac{-1}{x_1 x_2^2} \right\rangle$

$$f''(x_1, x_2) = \begin{bmatrix} \frac{2}{x_2 x_1^3} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_1^3} \end{bmatrix}$$

$$\det(f''(x_1, x_2)) = \left(\frac{4}{x_1^4 x_2^4} - \frac{1}{x_1^4 x_2^4} \right)$$

$$= \frac{3}{x_1^4 x_2^4} > 0$$

$\therefore f(x_1, x_2)$ is convex

Problem 5.3

December 4, 2023 8:34 PM

Problem 5.3 (Formulating problems as LPs and QPs)
OptM Problem 9.1. Do the problem for objective functions f_1, f_2, f_4, and f_5, (i.e., skip objective function f_3.) Also, you can ignore the part about putting your "problem in standard form", just state your formulations using equality and inequality constraints.

Exercise 9.1 (Formulating problems as LPs or QPs) Formulate the problem

$$p_j^* \doteq \min_x f_j(x),$$

for different functions f_j, $j = 1, \ldots, 5$, with values given in Table 9.6, as QPs or LPs, or, if you cannot, explain why. In our formulations, we always use $x \in \mathbb{R}^n$ as the variable, and assume that $A \in \mathbb{R}^{m,n}$, $y \in \mathbb{R}^m$, and $k \in \{1, \ldots, m\}$ are given. If you obtain an LP or QP formulation, make sure to put the problem in standard form, stating precisely what the variables, objective, and constraints are. *Hint:* for the last one, see Example 9.10.

$$\begin{aligned}
f_1(x) &= \|Ax - y\|_\infty + \|x\|_1 \\
f_2(x) &= \|Ax - y\|_2^2 + \|x\|_1 \\
f_3(x) &= \|Ax - y\|_2^2 - \|x\|_1 \\
f_4(x) &= \|Ax - y\|_2^2 + \|x\|_1^2 \\
f_5(x) &= \sum_{i=1}^k |Ax - y|_{[i]} + \|x\|_2^2
\end{aligned}$$

Table 9.6 Table of the values of different functions f. $|z|_{[i]}$ denotes the element in a vector z that has the i-th largest magnitude.

f_1:

$$p_1^* = \min_x f_1(x)$$

$$= \min_x \|Ax - y\|_\infty + \|x\|_1$$

Problem 5.4 (Minimizing sum of logarithms)
OptM Problem 8.5. (Please use the Lagrangian method to deal with the constraint $\sum_i x_i = c$. For this problem, it is okay to ignore the constraints $x_i \geq 0$, because they would be automatically satisfied as long as x_i stays within the domain of ln() function.)

Exercise 8.5 (Minimizing a sum of logarithms) Consider the following problem:

$$p^* = \max_{x \in \mathbb{R}^n} \sum_{i=1}^n \alpha_i \ln x_i \quad \text{s.t.:} \quad x \geq 0, \quad \mathbf{1}^\top x = c,$$

where $c > 0$ and $\alpha_i > 0$, $i = 1, \ldots, n$. Problems of this form arise, for instance, in maximum-likelihood estimation of the transition probabilities of a discrete-time Markov chain. Determine in closed-form a minimizer, and show that the optimal objective value of this problem is

$$p^* = \alpha \ln(c/\alpha) + \sum_{i=1}^n \alpha_i \ln \alpha_i,$$

where $\alpha \doteq \sum_{i=1}^n \alpha_i$.

$$p^* = \max_{x \in \mathbb{R}^n} \sum_{i=1}^n \alpha_i \ln x_i \quad \text{s.t.} \quad x \geq 0, \quad 1^T x = c$$

equivalent to

$$p^* = \min_{x \in \mathbb{R}^n} \sum_{i=1}^n -\alpha_i \ln x_i = \min_{x \in \mathbb{R}^n} \sum_{i=1}^n \alpha_i \ln\left(\frac{1}{x_i}\right)$$

$$L(x, \mu) = \sum_{i=1}^n \alpha_i \ln \frac{1}{x_i} + \mu(1^T x_i - c)$$

$$= \sum_{i=1}^n \left(\alpha_i \ln \frac{1}{x_i} + \mu x_i\right) - \mu c$$

$$g(\mu) = \min_{x \geq 0} \sum_{i=1}^n \left(\alpha_i \ln \frac{1}{x_i} + \mu x_i\right) - \mu c$$

n

$$= -\mu c + \sum_{i=1}^{n} \min_{x \geq 0} \alpha_i \ln \frac{1}{x_i} + \mu x_i$$

$$\frac{\partial}{\partial x} \alpha_i \ln \frac{1}{x_i} + \mu x_i = 0$$

$$0 = \alpha_i x_i \cdot \frac{-1}{x_i^2} + \mu$$

$$0 = \frac{-\alpha_i}{x_i} + \mu$$

$$x_i = \frac{\alpha_i}{\mu}$$

$$= -\mu c + \sum_{i=1}^{n} \alpha_i \ln\left(\frac{\mu}{\alpha_i}\right) + \alpha_i$$

$$= -\mu c + \ln(\mu) \cdot \sum_{i=1}^{n} \alpha_i + \sum_{i=1}^{n} \alpha_i (1 - \ln(\alpha_i))$$

$$\frac{\partial g(\mu)}{\partial \mu} = 0 = -c + \frac{1}{\mu} \sum_{i=1}^{n} \alpha_i + 0$$

$$\longrightarrow \quad \mu = \frac{\sum_{i=1}^{n} \alpha_i}{c}$$

$$\therefore \longrightarrow \quad x_i = \frac{\alpha_i}{\mu} = \frac{c\alpha_i}{\sum^{n} \alpha_i}$$

$\therefore \longrightarrow x_i = \dfrac{\mu}{\sum_{i=1}^{n} \alpha_i}$

$$p^* = \min_x \sum_{i=1}^{n} \alpha_i \ln \frac{1}{x_i} = \sum_{i=1}^{n} \alpha_i \ln\left(\frac{\sum_{i=1}^{n} \alpha_i}{c\,\alpha_i} \right)$$