Boyd, Convex Optimization, Chapter 3

Qilin Ye

February 18, 2022

Problem 1

Suppose $f : \mathbb{R} \to \mathbb{R}$ is convex and $a < b$ are in the domain.

(a) Show that for all $x \in [a, b]$,

$$f(x) \leqslant \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b).$$

(b) Show that for all $x \in (a, b)$,

$$\frac{f(x) - f(a)}{x-a} \leqslant \frac{f(b) - f(a)}{b-a} \leqslant \frac{f(b) - f(x)}{b-x}.$$

(c) Suppose f is differentiable. Use (b) to show that

$$f'(a) \leqslant \frac{f(b) - f(a)}{b-a} \leqslant f'(b).$$

(d) Suppose f is twice differentiable. Use (c) to show that $f''(a) \geqslant 0$ and $f''(b) \geqslant 0$.

Proof. (a) Since $\dfrac{b-x}{b-a} + \dfrac{x-a}{b-a} = 1$, and $\dfrac{(b-x)a}{b-a} + \dfrac{(x-a)b}{b-a} = x$, this follows directly from definition of convexity.

(b) The first $\leqslant$ is established by subtracting $f(a)$ from both sides of (a):

$$f(x) - f(a) \leqslant \frac{a-x}{b-a} f(a) + \frac{x-a}{b-a} f(b) \implies \frac{f(x) - f(a)}{a-b} \leqslant \frac{f(b) - f(a)}{b-a}.$$

The other $\leqslant$ follows from subtracting $f(b)$ from both sides of (a).

(c) Let $h \searrow 0$. Taking the limits in (b), we have

$$f'(a) = \lim_{h \searrow 0} \frac{f(a+h) - f(a)}{h} \leqslant \lim_{h \searrow 0} \frac{f(b) - f(a)}{b-a} = \frac{f(b) - f(a)}{b-a} \leqslant \lim_{h \searrow 0} \frac{f(b) - f(b-h)}{h} = f'(b).$$

(d) (c) implies $(f'(b) - f'(a))/(b-a) \geqslant 0$. Taking $b \to a$ and $a \to b$ respectively gives $f''(a), f''(b) \geqslant 0$. □

Problem 3: Inverse of an Increasing Convex Function

Suppose $f : \mathbb{R} \to \mathbb{R}$ is increasing and convex on (a, b). Let g denote its inverse. What can you say about convexity or conveys of g?

Solution. Claim: g is concave. Let $u, v \in (f(a), f(b))$ be given. Our goal is to show that for all $\lambda \in (0,1)$,

$$f^{-1}(\lambda u + (1-\lambda)v) \geqslant \lambda f^{-1}(u) + (1-\lambda)f^{-1}(v).$$

Since

$$\begin{aligned} f(\lambda f^{-1}(u) + (1-\lambda)f^{-1}(v)) &\leqslant \lambda f(f^{-1}(u)) + (1-\lambda)ff^{-1}(v) && \text{(convexity of } f\text{)} \\ &= \lambda u + (1-\lambda)v && \text{(inverse)} \\ &= f(f^{-1}(\lambda u + (1-\lambda)v) && \text{(inverse)} \end{aligned}$$

and f is monotone increasing, we must have that the arguments

$$\lambda f^{-1}(u) + (1-\lambda)f^{-1}(v) \leqslant f^{-1}(\lambda u + (1-\lambda)v),$$

as claimed.

Problem 5: Runing Average

Suppose $f : \mathbb{R} \to \mathbb{R}$ is convex with domain containing $[0, \infty)$. Show that the *running average* F, defined by

$$F(x) := \frac{1}{x}\int_0^x f(t)\,\mathrm{d}t,$$

is convex. You can assume that f is differentiable.

Proof. Let $0 < a < b < \infty$ and let $\lambda \in (0,1)$. Using convexity of f and change of variables multiple times,

$$\begin{aligned} F(\lambda a + (1-\lambda)b) &= \frac{1}{\lambda a + (1-\lambda)b}\int_0^{\lambda a+(1-\lambda)b} f(t)\,\mathrm{d}t \\ &= \int_0^1 f(\lambda at + (1-\lambda)bt)\,\mathrm{d}t \\ &\leqslant \int_0^1 \lambda f(at) + (1-\lambda)f(bt)\,\mathrm{d}t \\ &= \lambda\int_0^1 f(at)\,\mathrm{d}t + (1-\lambda)\int_0^1 f(bt)\,\mathrm{d}t \\ &= \frac{\lambda}{a}\int_0^a f(t)\,\mathrm{d}t + \frac{1-\lambda}{b}\int_0^b f(t)\,\mathrm{d}t = \lambda F(a) + (1-\lambda)F(b). \end{aligned}$$

□

Problem 8: Second-Order Condition for Convexity

Prove that a twice differentiable function f is convex if and only if its domain is convex and $\nabla^2 f(x) \geq 0$ for all x in the domain.

Proof. For the case $n=1$,$\Rightarrow$ is proven in problem 1(d). Conversely, assume $f''(x)\geqslant 0$ for all x. Let $x<y$ be points in the domain and $\lambda\in(0,1)$. For convenience denote $z:=\lambda x+(1-\lambda)y$. Immediately we have $\lambda=(y-z)/(y-x)$ and $1-\lambda=(z-x)/(y-x)$. By the intermediate value theorem (IVT), there exists $\xi_1\in(x,z)$ and $\xi_2\in(z,y)$ such that

$$\frac{f(z)-f(x)}{z-x}=f'(\xi_1)\qquad\text{and}\ \frac{f(y)-f(z)}{y-z}=f'(\xi_2).$$

That $f''\geqslant 0$ implies $f'(\xi_2)\geqslant f'(\xi_1)$. Since

$$f(y)-f(x)=\frac{f(y)-f(z)}{y-z}(y-z)+\frac{f(z)-f(x)}{z-x}(z-x)=f'(\xi_2)(y-z)+f'(\xi_1)(z-x)\geqslant f'(\xi_1)(y-x),$$

we have

$$\begin{aligned}f(z)=f(x)+f'(\xi_1)(z-x)&\leqslant f(x)+\frac{f(y)-f(x)}{y-x}(z-x)\\&=\frac{y-z}{y-x}f(x)+\frac{z-x}{y-x}f(y)=\lambda f(x)+(1-\lambda)f(y),\end{aligned}$$

as claimed. To generalize the case $n=1$, note that a function is convex if and only if it is convex in each component. □

Problem 9: Second-Order Condition for Convexity on an Affine Set

Let $F\in\mathbb{R}^{n\times m}$, $\hat{x}\in\mathbb{R}^n$. The restriction of $f:\mathbb{R}^n\to\mathbb{R}$ to the affine set $\{Fz+\hat{x}:z\in\mathbb{R}^m\}$ is defined as functions $\tilde{f}:\mathbb{R}^m\to\mathbb{R}$ with

$$\tilde{f}(z)=f(Fz+\hat{x}),\qquad \text{domain of }\tilde{f}=\{z:Fz+\hat{x}\in\text{domain of }f\}.$$

(a) Show that $\tilde{f}$ is convex if and only if for all $z\in$ domain of $\tilde{f}$,

$$F^T\nabla^2 f(Fz+\hat{x})F\succeq 0.$$

(b) Suppose $A\in\mathbb{R}^{p\times n}$ is a matrix whose nullspace is equal to the range of F. Show that $\tilde{f}$ is convex if and only if for all z in the domain of $\tilde{f}$, there exists $\lambda\in\mathbb{R}$ such that

$$\nabla^2 f(Fz+\hat{x})+\lambda A^TA\succeq 0.$$

Proof. (a) The quantity given is precisely the Hessian of $\tilde{f}$.

(b) By (a), if $Ax=0$, then $x^TA^TAx=0$ and x is in the range of F, so$x^T\nabla^2 f(Fz+\hat{x})x\geqslant 0$. Therefore their sum $\geqslant 0$, which finishes the proof.

□

Problem 12

Suppose $f:\mathbb{R}^n\to\mathbb{R}$ is convex, $g:\mathbb{R}\to\mathbb{R}$ is concave, both functions are defined on all of $\mathbb{R}^n$, and $f\leqslant g$. Show that there exists an affine function h such that $g(x)\leqslant h(x)\leqslant f(x)$ for all x.

Proof. By assumption, the interior of the epigraph of f and the hypograph of g do not intersect and are both convex. Therefore there exists a hyperplane separating the two sets, and this hyperplane corresponds to the graph of our function of interest. □

Problem 14: Convex-Concave Functions and Saddle-Points

We say a function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is *convex-concave* if $f(x, z)$ is concave as a function of z and convex as a function of x. We also require the domain to have product form $A \times B$ where $A \subset \mathbb{R}^n, B \subset \mathbb{R}^m$ are convex.

(a) Give a second-order condition for a twice-differentiable function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ to be convex-concave in terms of its Hessian $\nabla^2 f(x, z)$.

(b) Suppose that $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is convex-concave and differentiable, with $\nabla f(\tilde{x}, \tilde{z}) = 0$. Show that the *saddle-point property* holds: for all x, z we have

$$f(\tilde{x}, z) \leqslant f(\tilde{x}, \tilde{z}) \leqslant f(x, \tilde{z})$$

and that this implies the *strong max-min property*

$$\sup_z \inf_x f(f, z) = \inf_x \sup_z f(x, z) = f(\tilde{x}, \tilde{z}).$$

(c) Now suppose that $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is differentiable but not necessarily convex-concave, but the saddle-point property holds at $\tilde{x}, \tilde{z}$:

$$f(\tilde{x}, z) \leqslant f(\tilde{x}, \tilde{z}) \leqslant f(x, \tilde{z}) \qquad \text{for all } x, z.$$

Show that $\nabla f(\tilde{x}, \tilde{z}) = 0$.

Proof. (a) $f(\cdot, z)$ for fixed z being convex implies $\nabla^2_{xx} f(x, z) \geq 0$ and $f(x, \cdot)$ for fixed x being concave implies $\nabla^2_{zz} f(x, z) \leq 0$.

(b) Since $\nabla f(\tilde{x}, \tilde{z}) = 0$ but $f(\cdot, \tilde{z})$ is convex, we know $f(\tilde{x}, \tilde{z})$ must attain the global minimum of $f(\cdot, \tilde{z})$. That is, $f(\tilde{x}, \tilde{z}) \leqslant f(x, \tilde{z})$ for all x. The other inequality follows analogously.

Taking limits and using continuity gives the strong max-min property.

(c) If the saddle-point property holds then $f(\tilde{x}, \tilde{z})$ minimizes $f(\cdot, \tilde{z})$ and $f(\tilde{x}, \tilde{z})$ also maximizes $f(\tilde{x}, \cdot)$. That is, $\nabla f_x = \nabla f_y = 0$ at $(\tilde{x}, \tilde{z})$. □

Problem 17

Suppose $p < 1, p \neq 0$. Show that

$$f(x) := \left(\sum_{i=1}^n x_i^p\right)^{1/p} \qquad \text{with domain } \mathbb{R}^n_{++}$$

is concave.

Proof. We want to show that for any x and any $v \in \mathbb{R}^n$,

$$v^T \nabla^2 f(x) v \leqslant 0.$$

We now compute the first-order partials:

$$\frac{\partial f(x)}{\partial x_i} = p x_i^{p-1} \cdot \frac{1}{p} \left(\sum_{i=1}^n x_i^p \right)^{1/p-1} = x_i^{p-1} \left(\sum_{i=1}^n x_i^p \right)^{(1-p)/p} = x_i^{p-1} f(x)^{1-p} = \left(\frac{f(x)}{x_i} \right)^{1-p}. \tag{1}$$

The second-order mixed-partials (i.e., for $i \neq j$) are

$$\begin{aligned} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} &= \frac{\partial}{\partial x_j} \left(\frac{f(x)}{x_i} \right)^{1-p} \\ &= \frac{1}{x_i^{1-p}} \cdot (1-p) f(x)^{-p} \left(\frac{f(x)}{x_j} \right)^{1-p} \\ &= \frac{1-p}{f(x)^p} \left(\frac{f(x)}{x_i x_j} \right)^{1-p} = \frac{1-p}{f(x)} \left(\frac{f(x)^2}{x_i x_j} \right)^{1-p} \end{aligned} \tag{1}$$

and the second-order unmixed partials (i.e., for $i = j$) are

$$\begin{aligned} \frac{\partial^2 f(x)}{\partial x_i^2} &= (1-p) \left(\frac{f(x)}{x_i} \right)^{-p} \left[\frac{x_i (f(x)/x_i)^{1-p} - f(x)}{x_i^2} \right] \\ &= (1-p) \left(\frac{f(x)}{x_i} \right)^{-p} \left[\frac{f(x)^{2-p}}{x_i^{-1-p}} - \frac{f(x)}{x_i^2} \right] \\ &= \frac{(1-p) f(x)^{2-2p}}{x_i^{-2+2p}} - \frac{(1-p) f(x)^{1-p}}{x_i^{-2+p}} \\ &= \frac{1-p}{f(x)} \left(\frac{f(x)}{x_i} \right)^{2(1-p)} - \frac{1-p}{x_i} \left(\frac{f(x)}{x_i} \right)^{1-p}. \end{aligned} \tag{2}$$

Summing over all i's, we have

$$\begin{aligned} v^T \nabla^2 f(x) v &= \sum_{i=1}^n \sum_{j=1}^n v_i v_j \frac{\partial^2 f(x)}{\partial x_j \partial x_j} \\ &= \sum_{i=1}^n v_i^2 \frac{1-p}{f(x)} \left(\frac{f(x)^{1-p}}{x_i^{1-p}} \right)^2 + 2 \sum_{i \neq j} v_i v_j \frac{1-p}{f(x)} \left(\frac{f(x)^{2(1-p)}}{x_i^{1-p} x_j^{1-p}} \right) - \sum_{i=1}^n v_i^2 \frac{1-p}{x_i} \left(\frac{f(x)}{x_i} \right)^{1-p} \\ &= \frac{1-p}{f(x)} \left(\sum_{i=1}^n \frac{v_i f(x)^{1-p}}{x_i^{1-p}} \right)^2 - \frac{1-p}{f(x)} \sum_{i=1}^n \frac{v_i^2 f(x)^{2-p}}{x_i^{2-p}} \\ &= \frac{1-p}{f(x)} \left(\left(\sum_{i=1}^n \frac{v_i f(x)^{1-p}}{x_i^{1-p}} \right)^2 - \sum_{i=1}^n \frac{v_i^2 f(x)^{2-p}}{x_i^{2-p}} \right). \end{aligned}$$

It remains to notice

$$f(x)^p = \sum_{i=1}^n x_i^p \implies \sum_{i=1}^n \frac{x_i^p}{f(x)^p} = 1 \implies \sum_{i=1}^n \left(\frac{f(x)}{x_i} \right)^{-p} = 1,$$

$$\frac{v_i f(x)^{1-p}}{x_i^{1-p}} = \left(\frac{f(x)}{x_i} \right)^{-p/2} \cdot v_i \left(\frac{f(x)}{x_i} \right)^{1-p/2},$$

and

$$\frac{v_i^2 f(x)^{2-p}}{x_i^{2-p}} = \left(v_i \left(\frac{f(x)}{x_i} \right)^{1-p/2} \right)^2.$$

Then, using Cauchy-Schwarz on

$$a_i := \left(\frac{f(x)}{x_i}\right)^{-p/2} \qquad \text{and} \qquad b_i := v_i\left(\frac{f(x)}{x_i}\right)^{1-p/2}$$

with respect to the standard Euclidean norm, we obtain

$$\left(\sum_{i=1}^n \frac{v_i f(x)^{1-p}}{x_i^{1-p}}\right)^2 \leqslant 1 \cdot \sum_{i=1}^n \frac{v_i^2 f(x)^{2-p}}{x_i^{2-p}},$$

so $v^T \nabla^2 f(x) v \leqslant 0$, which concludes the proof. □

Problem 21: Pointwise Maximum and Supremum

Show that the following functions $f : \mathbb{R}^n \to \mathbb{R}$ are convex.

(a) $f(x) = \max_{i=1,\dots,k} \|A^{(i)}x - b^{(i)}\|$ where $A^{(i)} \in \mathbb{R}^{m\times n}, b^{(i)} \in \mathbb{R}^m$, and $\|\cdot\|$ is a norm on $\mathbb{R}^m$.

(b) $f(x) := \sum_{i=1}^{r} |x|_{[i]}$ on $\mathbb{R}^n$, where $|x|_{[i]}$ denotes the i^{th} largest component of $|x|$.

Proof. (a) Each $\|A^{(i)}x - b^{(i)}\|$ is a translation of a norm and is therefore convex. Taking the max preserves convexity.

(b) For a given r, we have

$$f(x) = \sum_{i=1}^r |x|_{[i]} = \max_{\substack{j\in I\subset\{1,\dots,n\}\\ |I|=r}} \sum_{j=1}^r |x_j|$$

which is the maximum over a finite (in particular, n choose r) convex functions. It is therefore convex. □

Problem 22: Composition Rules

Show that the following functions are convex.

(a) $f(x) = -\log\left(-\log\left(\sum_{i=1}^n \exp(a_i^T x + b_i)\right)\right)$ on domain $\left\{x : \sum_{i=1}^m \exp(a_i^T x + b_i) < 1\right\}$. You may use the fact that $\log\left(\sum_{i=1}^n \exp(y_i)\right)$ is convex.

(b) $f(x,u,v) = -\sqrt{uv - x^T x}$ on $\{(x,u,v) : uv > x^T x \text{ and } u,v > 0\}$. Use the fact that $x^T x/u$ is convex for $u > 0$ and that $-\sqrt{x_1 x_2}$ is convex on $\mathbb{R}^2_{++}$.

(c) $f(x,u,v) = -\log(uv - x^T x)$ on the same domain as in (b).

(d) $f(x,t) = -(t^p - \|x\|_p^p)^{1/p}$ where $p > 1$ and domain of f is $\{(x,t) : t \geqslant \|x\|_p\}$. You can use the fact that $\|x\|_p^p/u^{p-1}$ is convex for $u > 0$ (see problem 23) and that $-x^{1/p}y^{1-1/p}$ is convex on $\mathbb{R}^2_+$.

(e) $f(x,t) = -\log(t^p - \|x\|_p^p)$ with same assumptions as in (d). You may use the fact that $\|x\|_p^p/u^{p-1}$ is convex for $u > 0$ (see problem 23 again).

Proof. (a) $a_i^T x + b_i$ is affine so composing it with the log-sum-exp function gives a convex function. Flipping the sign makes it concave, and composing it with $-\log$ again (convex and decreasing) makes the overall function convex.

(b) Note that

$$-\sqrt{uv - x^T x} = -\sqrt{u(v - x^T x/u)}$$

so that $v - x^T x/u$ is concave and $-\sqrt{uw}$ is convex and decreasing. Composing them gives the original function and shows it is convex.

(c) Since $uv - x^T x = u(v - x^T x/v)$ is concave and $-\log$ is convex and decreasing, the composition is convex in each component and therefore convex.

(d) Per the hint, we have

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

which is convex and deceasing with respect to either the argument $t - \|x\|_p^p/t^{p-1}$ or just t. Both are concave. Therefore the composition is convex.

(e) Since

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

where the first function is concave and so is the second (concave function composed with convex decreasing function), we see $f(x,t)$ is a sum of two convex functions and is therefore convex. □

Problem 25: Maximum Probability Distance between Distributions

Let $p, q \in \mathbb{R}^n$ represent two distributions on $\{1, \ldots, n\}$ so that $p, q \geq 0$ and $1^T p = 1^T q = 1$. We define the *maximum probability distamce*

$$d_{\text{mp}}(p,q) := \max\{|\mathbb{P}(p,C) - \mathbb{P}(q,C) : C \subset \{1,\ldots,n\}|\}$$

where $\mathbb{P}(p,C) := \sum_{i \in C} p_i$. Simplify the expression for $d_{\text{mp}}(p,q)$ using $\|\cdot\|_1$ and show that it is convex.

Solution. By assumption $\sum_{i=1}^n p_i = \sum_{i=1}^n q_i = 1$, so

$$\sum_{p_i > q_i} (p_i - q_i) + \sum_{p_i \leqslant q_i} (p_i - q_i) = 0 \implies \sum_{p_i > q_i} (p_i - q_i) = -\sum_{p_i \leqslant q_i} (p_i - q_i) \tag{1}$$

On the other hand,

$$\sum_{p_i > q_i} |p_i - q_i| + \sum_{p_i \leqslant q_i} |p_i - q_i| = \sum_{i=1}^n |p_i - q_i| = \|p - q\|_2, \tag{2}$$

and by using (1) and noticing that $\sum_{p_i \leqslant q_i} |p_i - q_i| = \sum_{p_i \leqslant q_i} -(p_i - q_i) = -\sum_{p_i \leqslant q_i} (p_i - q_i) = \sum_{p_i > q_i} (p_i - q_i)$, we have

$$\sum_{p_i > q_i} (p_i - q_i) = \frac{\|p - q\|_1}{2}.$$

From the definition of d_{mp}, it should be clear that this quantity is maximized if and only if $C \coloneqq \{i : p_i > q_i\}$, and if so, we have

$$d_{\mathrm{mp}}(p, q) = \sum_{p_i > q_i} (p_i - q_i) = \frac{\|p - q\|_1}{2},$$

clearly a convex function. □

Problem 30: Convex Hull or Envelop of a Function

The *convex hull* or *convex envelope* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$g(x) \coloneqq \inf\{t : (x, t) \in \operatorname{conv} \operatorname{epi} f\}.$$

Show that g is the largest convex underestimator of f.

Proof. By construction the epigraph of g is the convex hull of the epigraph of f. It follows from definition that g has a convex epigraph and is therefore convex. It again follows from definition that the epigraph of g is the minimal convex shape containing the epigraph of f, so if h is an convex underestimator of f, its epigraph must be a superset of the epigraph of g, i.e., $h \leqslant g$. □

Problem 31: Largest Homogeneous Underestimator

Let f be convex and define

$$g(x) \coloneqq \inf_{\alpha > 0} \frac{f(\alpha x)}{\alpha}.$$

(a) Show that g is homogeneous, i.e., $g(tx) = tg(x)$ for all $t \geqslant 0$.

(b) Show that g is the largest homogeneous underestimator of f.

(c) Show that g is convex.

Proof. (a) The claim is trivial for $t = 0$, and for $t > 0$,

$$g(tx) = \inf_{\alpha > 0} \frac{f(\alpha \cdot tx)}{\alpha} = t \inf_{\alpha > 0} \frac{f(\alpha \cdot tx)}{t\alpha} = tg(x).$$

(b) For any homogeneous underestimator h of f and any $\alpha > 0$,

$$h(x) = \frac{h(\alpha x)}{\alpha} \leqslant \frac{f(\alpha x)}{\alpha},$$

so taking the infimum gives $h(x) \leqslant \inf_{\alpha > 0} \frac{f(\alpha x)}{\alpha} = g(x)$.

(c) Since $g(x) = \inf_{\alpha > 0} \frac{f(\alpha x)}{\alpha} = \inf_{t^{-1} > 0} \frac{f(t^{-1} x)}{t^{-1}} = \inf_{t > 0} t f(x/t)$, we rewrote g as the infimum of a family of convex (perspective) functions, so it must be convex as well.

□

Problem 33: Direct Proof of the Perspective Theorem

Give a direct proof showing that $g(x,t) := tf(x/t)$ is convex if f is convex.

Proof. The domain of g is

$$\{(x,t) \subset \mathbb{R}^n \times \mathbb{R}_+ : x/t \in \text{domain of } f\}.$$

Given f is convex, dilating its domain by a factor of t preserves convexity; then, the Cartesian product with $\mathbb{R}^n \times \mathbb{R}_+$, a convex set, again preserves convexity.
Now let (x_1,t_1) and (x_2,t_2) be chosen from g's domain and let $\lambda \in (0,1)$. Then

$$\begin{aligned}
g(\lambda x_1 + (1-\lambda)x_2, \lambda t_1 + (1-\lambda)t_2) &= (\lambda t_1 + (1-\lambda)t_2) \cdot f\left(\frac{\lambda x_1 + (1-\lambda)x_2}{\lambda t_1 + (1-\lambda)t_2}\right) \\
&= (\lambda t_1 + (1-\lambda)t_2) f\left(\frac{\lambda t_1(x_1/t_1) + (1-\lambda)t_2(x_2/t_2)}{\lambda t_1 + (1-\lambda)t_2}\right) \\
&\leqslant (\lambda t_1 + (1-\lambda)t_2)\left[\frac{\lambda t_1}{\lambda t_1 + (1-\lambda)t_2} \cdot f(x_1/t_1) + \frac{(1-\lambda)t_2}{\lambda t_1 + (1-\lambda)t_2} \cdot f(x_2/t_2)\right] \\
&= \lambda t_1 f(x_1/t_1) + (1-\lambda)t_2 f(x_2/t_2) = \lambda g(x_1,t_1) + (1-\lambda)g(x_2,t_2),
\end{aligned}$$

where we used the convexity of f in the $\leqslant$. □

Problem 34: The Minkowski Function

The *Minkowski function* on a convex set C is defined as

$$M_C(x) := \inf\{t > 0 : t^{-1}x \in C\}.$$

(a) Give a geometric interpretation of how to find $M_C(x)$.

(b) Show that M_C is homogeneous, i.e., $M_C(\alpha x) = \alpha M_C(x)$ for $\alpha \geqslant 0$.

(c) What is its domain?

(d) Show that M_C is convex.

(e) Suppose C is closed[1] and symmetric with nonempty interior. Show that M_C induces a norm. What is the corresponding unit ball?

Solution. (a) Excluding the edge cases, we draw a line segment ℓ from the origin to x. Assuming the infimum exists (i.e., x is inside the domain), the line segment needs to intersect C. In the intersection $\ell \cap C$, there either exists a point p closest to x or there exists a sequence tending to a limit p, closer to x than anything in $\ell \cap C$. In either case, t^{-1} is ratio between $\|p\|$ and $\|x\|$. In other words, t is the reciprocal of the infimum of "scaling factors" transforming x into C.

(b) This directly follows from definition: for $\alpha > 0$,

$$M_C(\alpha x) = \inf\{t > 0 : t^{-1}\alpha x \in C\} = \alpha \inf\{t/\alpha > 0 : t^{-1}\alpha x \in C\} = \alpha M_C(x).$$

[1] I don't think being closed is sufficient. Maybe compact? Otherwise take $C := \mathbb{R}^n$, which is closed and convex, and $M_C(x) = 0$ for any x.

For $\alpha = 0$, $M_C(\alpha x) = M_C(0)$. Since 0 is in the domain only if $0 \in C$ (see below), we implicitly assume so. In this case $M_C(0) = 0$. On the other hand $\alpha M_C(x) = 0$, so homogeneity still holds.

(c) Its domain is $\{x : t^{-1}x \in C \text{ for some } t > 0\}$.

(d) We define the indicator function $I_C : \mathbb{R}^n \to \overline{\mathbb{R}}$ by

$$I_C(x) = \begin{cases} 0 & \text{if } x \in C \\ \infty & \text{otherwise} \end{cases}.$$

Then

$$M_C(x) = \inf\{t > 0 : t^{-1}x \in C\} = \inf_t(t + I_C(x/t)).$$

For fixed t, x/t is linear and I_C convex since C is convex. Hence $t + I_C(x/t)$ is convex, and taking infimum preserves the convexity.

(e) (*Here I assume in addition that C is bounded and so compact.*) Nondegeneracy is clear as $M_C(x)$ is nonnegative. If $x = 0$ then $M_C(x) = 0$ as shown above. Conversely, if $M_C(x) = 0$ but $x \neq 0$, then $nx \in C$ for all $n \in \mathbb{C}$ which implies C is unbounded.

Absolute homogeneity follows from homogeneity and symmetry of C (so that $M_C(-x) = M_C(x)$).

Finally, for subadditivity, we have

$$M_C(x + y) = 2M_C((x + y)/2) \leqslant M_C(x) + M_C(y)$$

where the $=$ is by homogeneity and the $\leqslant$ by convexity.

Problem 35: Support Function Calculus

Recall that the *support function* of a set $C \subset \mathbb{R}^n$ is defined as $S_C(y) := \sup\{y^T x : x \in C\}$. We showed that S_C is convex.

(a) Show that $S_B = S_{\mathbf{conv}B}$.

(b) Show that $S_{A+B} = S_A + S_B$.

(c) Show that $S_{A\cup B} = \max\{S_A, S_B\}$.

(d) Let B be closed and convex. Show that $A \subset B$ if and only if $S_A(y) \leqslant S_B(y)$ for all y.

Proof. (a) It is clear that $B \subset \mathbf{conv}B$ implies $S_B \leqslant S_{\mathbf{conv}B}$, so it remains to show that $<$ cannot happen. Suppose for contradiction that $S_B(y) < S_{\mathbf{conv}B}(y)$ for some y. Then there exist some $v \in \mathbf{conv}B$ such that $y^T v > S_B(y)$. That is,

$$y^T v > y^T u \text{ for all } u \in B. \tag{*}$$

By definition of convex hull, v is some convex combination of elements of B, i.e.,

$$v = \sum_{i=1}^{k} c_i u_i \qquad \text{where } u_i \in B, c_i \geqslant 0, \text{ and } \sum_{i=1}^{k} c_i = 1.$$

But then

$$y^T v = \sum_{i=1}^{k} c_i y^T u_i \overset{(*)}{<} \sum_{i=1}^{k} c_i y^T v = y^T v,$$

contradiction.

(b) $S_{A+B}(y) = \sup\{y^T(u+v) : u \in A, v \in B\} = \sup\{y^T u + y^T v\} = \sup\{y^T u\} + \sup\{y^T v\} = S_A(y) + S_B(y)$.

(c) $S_{A\cup B} = \sup\{y^T u : u \in A \cup B\} = \max\{\sup\{y^T u\}, \sup\{y^T v\}\} = \max\{S_A, S_B\}$.

(d) If $A \subset B$ then clearly $S_A \leqslant S_B$; it remains to show the converse.

If $A \not\subset B$ then there exists $x \in A$ but $x \notin B$. Since B is closed, $d(x, B) := \inf_{b\in B} d(x, b) > 0$. Hence there exists a separating hyperplane with $y^T x > y^T b$ for all $b \in B$. Then $S_A(y) > S_B(y)$, a contradiction. □

Problem 36: Conjugate Functions

Derive the conjugates of the following functions.

(a) Max: $f(x) := \max_{1\leqslant i\leqslant n} x_i$ on $\mathbb{R}^n$.

(b) Sum of largest elements: $f(x) := \sum_{i=1}^{r} x_{[i]}$ on $\mathbb{R}^n$.

(c) Piecewise linear: $f(x) := \max_{1\leqslant i\leqslant n}(a_i x + b_i)$ on $\mathbb{R}$, assuming $a_1 \leqslant \ldots \leqslant a_m$ and none of the functions $a_i x + b$ is redundant.

(d) Power: $f(x) := x^p$ with $p > 1$. Repeat for $p < 0$.

(e) Geometric mean: $f(x) := -\left(\prod_{i=1}^{n} x_i\right)^{1/n}$ on $\mathbb{R}^n_{++}$.

(f) Negative generalized logarithm for second-order cone: $f(x,t) := -\log(t^2 - x^T x)$ on $\{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \|x\|_2 < t\}$.

Solution. For convenience I first write the deifnition of a conjugate:

$$f^*(y) := \sup_{x\in \mathbf{dom} f} (y^T x - f(x)).$$