MATH 625 Homework 2

Qilin Ye

June 20, 2022

Problem 1

Let $f : \mathbb{R} \to \mathbb{R}$ be strictly convex for $x > 0$. Show that $g(x) := xf(1/x)$ is also strictly convex.

Proof. We use the following characterization of convexity:

> $f : \mathbb{R} \to \mathbb{R}$ is convex if and only if for each $x_0 \in \mathbb{R}$, the secant slope function $x \mapsto (f(x) - f(x_0))/(x - x_0)$ is increasing on $\mathbb{R} \backslash \{x_0\}$.

Now pick $x_0 > 0$. For $x \in \mathbb{R}_{>0} \backslash \{x_0\}$, we have

$$\begin{aligned} \frac{g(x) - g(x_0)}{x - x_0} &= \frac{xf(1/x) - x_0 f(1/x_0)}{x - x_0} \\ &= \frac{xf(1/x) + xf(1/x_0) - xf(1/x_0) - x_0 f(1/x_0)}{x - x_0} \\ &= f(1/x_0) + \frac{xf(1/x) - xf(1/x_0)}{x - x_0} \\ &= f(1/x_0) - \frac{1}{x_0} \frac{f(1/x) - f(1/x_0)}{1/x - 1/x_0}. \end{aligned} \tag{*}$$

Using the convexity characterization of f, we see that as x increases, $1/x$ decreases, so (*) increases as x increases. Since x_0 is arbitrary, we have shown that g is convex. □

Problem 2

Let C be nonempty convex in $\mathbb{R}^{n+1}$ and let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be defined by

$$f(x) := \inf\{w : (x, w) \in C\}.$$

Show that f is convex.

Proof. We show that the epigraph of f is convex. Let (x, w) and (y, v) be two points in the epigraph of f and let $\lambda \in (0, 1)$. By definition, there exist sequences $(x, w_n)_{n \geqslant 1}$ and $(y, v_n)_{n \geqslant 1}$ in C converging to (x, w) and (y, v) respectively. WLOG assume for each n, $w_n \leqslant w + 1/n$ and $v_n \leqslant v + 1/n$.
For each n, by convexity we have

$$\lambda(x, w_n) + (1 - \lambda)(y, v_n) = (\lambda x + (1 - \lambda)y, \lambda w_n + (1 - \lambda)v_n) \in C,$$

so by definition

$$f(\lambda x + (1-\lambda)y) \leqslant \lambda w_n + (1-\lambda)v_n \leqslant \lambda w + (1-\lambda)v + 1/n.$$

Letting $n \to \infty$ we see $f(\lambda x + (1-\lambda)y) \leqslant \lambda w + (1-\lambda)v$, i.e.,

$$\lambda(x, w) + (1-\lambda)(y, v) \in \operatorname{epi}(f).$$

This completes the proof. The domain is all of $\mathbb{R}^n$ since $\inf \varnothing = \infty$ which is by assumption part of the codomain.

□

Problem 3: Boyd 3.4

Show that a continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if for every line segment, its average value on the segment is less than or equal to the average of its values at the endpoints. That is, for $x, y \in \mathbb{R}^n$,

$$\int_0^1 f(x + \lambda(y-x))\,\mathrm{d}\lambda \leqslant \frac{f(x)+f(y)}{2}.$$

Proof. If f is convex, then $f(x + \lambda(y-x)) = f(\lambda y + (1-\lambda)x) \leqslant \lambda f(y) + (1-\lambda)f(x)$, so

$$\int_0^1 f(x + \lambda(y-x))\,\mathrm{d}\lambda \leqslant \int_0^1 \lambda f(y)\,\mathrm{d}\lambda + \int_0^1 (1-\lambda)f(x)\,\mathrm{d}\lambda = \frac{f(x)+f(y)}{2}.$$

Conversely, suppose there exist x_0, y_0 and $\lambda_0 \in (0,1)$ such that $f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda)y$. Let $a \in [0, \lambda_0)$ be the largest solution to $f(\lambda x + (1-\lambda)y) = \lambda(f(x) + (1-\lambda)y$ and let $b \in (\lambda_0, 1]$ the smallest solution to the same equation. Let $x := ax_0 + (1-a)y_0$ and $y := bx_0 + (1-b)y_0$. Then

$$f(\lambda x_0 + (1-\lambda)y_0) > \lambda f(x_0) + (1-\lambda)f(y_0) \qquad \text{for } \lambda \in (a, b)$$

which implies

$$f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda)f(y) \qquad \text{for } \lambda \in (0, 1).$$

Integrating then gives

$$\int_0^1 f(y + \lambda(x-y))\,\mathrm{d}\lambda > \frac{f(x)+f(y)}{2}.$$

□

Problem 4: Boyd 3.11

A function $\psi : \mathbb{R}^n \to \mathbb{R}$ is monotone if for all x, y,

$$(\psi(x) - \psi(y))^T(x-y) \geqslant 0.$$

Suppose f is differentiable and convex. Show ∇f is monotone. Is the converse true, i.e., is every monotone mapping the gradient of a convex function?

Proof. Using definition we have $f(x) \geqslant f(y) + \nabla f(y)^T(x-y)$ and $f(y) \geqslant f(x) + \nabla f(x)^T(y-x)$. Combining these two gives the desired inequality for monotonicity. The converse is not true. Let A be any matrix of form $\begin{bmatrix} a & 0 \\ b & a \end{bmatrix}$ with $0 < b < a$. Then this matrix is positive definite. Let $\psi(x)$ be defined by $x \mapsto Ax$. Then by positive definiteness

ψ is monotone. However, ψ cannot have a primitive, for otherwise its mixed partials must equal, whereas here we have $0 \neq b$. □

Problem 5

A function $f : \mathbb{R} \to \mathbb{R}$ is **α-strongly convex**, where $\alpha > 0$, if

$$f_\alpha(x) := f(x) - \alpha x^2$$

is convex.

(1) Assuming f is twice differentiable, give a characterization of α-strongly convex functions in terms of f''.

(2) Given an example of a strictly convex function which is not strongly convex for any $\alpha > 0$.

Solution. (1) If f is twice differentiable, so is f_α with $f''_\alpha(x) = f''(x) - 2\alpha$. Therefore in this case f is α-strongly convex if and only if $f''(x) \geqslant 2\alpha$ for all x.

(2) Intuitively we want a function that grows sub-quadraticly. One such example is e^x *on* $\mathbb{R}$. It is not strongly convex on $(-\infty, 0)$.

Problem 6

Let f be convex in some interval $I \subset \mathbb{R}$. Let $x, y, z \in I$ with $x < y < z$. Show that

$$\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ f(x) & f(y) & f(z) \end{vmatrix} \geqslant 0$$

and give a geometric interpretation of this result.

Proof. We compute by brute force.

$$\begin{aligned} \det &= yf(z) - zf(x) - xf(z) + zf(x) + xf(y) - yf(x) \\ &= (z-y)f(x) + (y-x)f(z) + (x-z)f(y). \end{aligned}$$

Since $x < z$, it suffices to show that

$$(z-x)f(y) \leqslant (z-y)f(x) + (y-x)f(z).$$

Using convexity, we have

$$f\Bigg(\underbrace{\frac{z-y}{z-x} \cdot x + \frac{y-x}{z-x} \cdot z}_{=y}\Bigg) \leqslant \frac{z-y}{z-x} f(x) + \frac{y-x}{z-x} f(z).$$

Now multiply both sides by $z - x$ and we are done.

For a geometric implication, note that the determinant is equal to the signed area of the triangle with vertices $(x, f(x)), (y, f(y))$, and $(z, f(z))$. This should be straightforward by the shoelace theorem. Given that $x < y < z$, y is on or below the segment passing through x and z, so the signed area is positive, as claimed. □

Problem 7

Let $f : \mathbb{R}^n \to \mathbb{R}$ and let f^* be its conjugate. Define $g(x) := f(Ax + b)$ where A is an $n \times n$ matrix and $b \in \mathbb{R}^n$. Prove that

$$g^*(z) = f^*(A^{-T}z) - b^T A^{-T} z$$

for $z \in \operatorname{dom} g^* = A^T \operatorname{dom} f^*$.

Proof. By definition,

$$\begin{aligned} g^*(z) = \sup_{x \in \operatorname{dom} g} \{z^T x - g(x)\} &= \sup_{Ax+b \in \operatorname{dom} f} \{z^T x - f(Ax+b)\} \\ &= \sup_{Ax+b \in \operatorname{dom} f} \{z^T A^{-1}(Ax + b - b) - f(Ax+b)\} \\ &= \sup_{Ax+b \in \operatorname{dom} f} \{z^T A^{-1}(Ax + b) - f(Ax+b) - z^T A^{-1} b\} \\ &= \sup_{Ax+b \in dom f} \{(A^{-T}z)^T (Ax + b) - f(Ax+b) - z^T A^{-1} b\} \\ &= \sup_{Ax+b \in \operatorname{dom} f} \{(A^{-T}z)^T (Ax + b) - f(Ax+b)\} - z^T A^{-1} b \\ &= f^*(A^{-T}z) - b^T A^{-T} z \end{aligned}$$

since $b^T A^{-T} z$ is just a number and transposition has no effect. □

Problem 8: Boyd 3.34

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

[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.

$\ell \cap C$, 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 $\alpha > 0$,

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

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 9: Boyd 3.38

(Young's inequality.) Let $f : \mathbb{R} \to \mathbb{R}$ be increasing with $f(0) = 0$ and let g be its inverse. Define F and G as

$$F(x) := \int_0^x f(a)\,\mathrm{d}a \qquad \text{and} \qquad G(y) := \int_0^y g(a)\,\mathrm{d}a.$$

Show that F and G are conjugates. Give a simple graphical interpretation of Young's inequality

$$xy \leqslant F(x) + G(y).$$

Proof. We first draw the graph of f on $\mathbb{R}^2$. Given x_0 and y_0, $F(x)$ denotes the area bounded by the x-axis, the vertical line $x = x_0$, and the graph of f. Similarly, $G(y)$ denotes the area bounded by the y-axis, the horizontal line $y = y_0$, and the graph of x. Then it is clear that $xy \leqslant F(x) + G(y)$ and equality happens if and only if $y = f(x)$.

From this we recover the conjugate equations $G(y) = \sup_x(xy - F(x))$ and $F(x) = \sup_y(xy - G(y))$. □

Problem 10: Boyd 3.46

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is quasilinear and continuous with domain $\mathbb{R}^n$. Show that it can be expressed as $f(x) = g(a^T x)$ where $g : \mathbb{R} \to \mathbb{R}$ is monotone and $x \in \mathbb{R}^n$.

Proof. Since f is quasilinear, its sublevel and superlevel sets are both convex. Since for each fixed level, these two sets make up the whole space, they have to be hyperplanes. That is, for $x_0 \in \mathbb{R}$, there exists $a(x_0) \in \mathbb{R}^n$ and $b(x_0) \in \mathbb{R}$ such that

$$\{x : f(x) \leqslant x_0\} = \{x : a(x_0)^T x \leqslant b(x_0)\}.$$

Furthermore, we claim that $a(x_0)$ is independent of x_0: this is because the sublevel sets are nested, meaning that the hyperplanes must all have the same orthogonal vector, which we now call a. Also, $b(\cdot)$ is an increasing (nondecreasing) function. We define g to be the generalized inverse of b, namely

$$g(t) := \sup\{x : b(x) \leqslant t\}.$$

Then $f(x) = g(a^T x)$, as claimed. □