CONVEX FUNCTIONS

MAT157, FALL 2020. YAEL KARSHON.

These notes supplement Spivak's appendix to Chapter 11. Please let me know if you find mistakes or typos.

Please see the lecture notes from class for diagrams that accompany this handout.

Intervals

For any two real numbers a and b such that $a < b$, the interval $[a, b]$ is given by

$$[a, b] = \{(1 - t)a + tb \mid 0 \leq t \leq 1\}.$$

Given a number x in this interval, the corresponding coefficients are given by

$$t = \frac{x - a}{b - a} \quad \text{and} \quad 1 - t = \frac{b - x}{b - a}.$$

For any two distinct points p and q in $\mathbb{R}^2$, the segment (also called "straight line segment" or "interval") from p to q is

$$[p, q] := \{(1 - t)p + tq \mid 0 \leq t \leq 1\}.$$

Convex functions

Let $f \colon I \to \mathbb{R}$ be a real valued function on an interval I.

Definition. The function f is **weakly convex** if, for every $a, b \in I$ and every $t \in (0, 1)$,

$$f((1 - t)a + tb) \leq (1 - t)f(a) + tf(b).$$

The function f is **strictly convex** if, for every $a, b \in I$ and every $t \in (0, 1)$,

$$f((1 - t)a + tb) < (1 - t)f(a) + tf(b).$$

Terminology: Spivak's "convex" means strictly convex. Many other author's "convex" means weakly convex.

Example: the function $f(x) = 1/x$ is convex on $(0, \infty)$ and concave on $(-\infty, 0)$.

Geometric meaning: f is weakly convex iff for each $a, b \in I$ the interval from $(a, f(a))$ to $(b, f(b))$ is weakly above the graph of f. Similarly, f is strictly convex iff for each $a, b \in I$ the interval from $(a, f(a))$ to $(b, f(b))$ is strictly above the graph of f, except at its endpoints.

Weakly concave and **strictly concave** are defined similarly, with the opposite inequalities.

Relation with convex sets (optional)

Definition. A *subset* A of $\mathbb{R}^2$ is **convex** if for each two points p and q, if p and q are both in A, then the segment from p to q is contained A.

Exercise: $f\colon I \to \mathbb{R}$ is weakly convex iff $\{(x,y) \mid x \in I,\ y \geq f(x)\}$ is a convex subset of $\mathbb{R}^2$.

Criteria for convexity: increasing slopes

Lemma. *$f\colon I \to \mathbb{R}$ is weakly convex if and only if for any $a < b < c$ in I*

$$\frac{f(b)-f(a)}{b-a} \leq \frac{f(c)-f(b)}{c-b}.$$

Proof. By algebraic manipulations, this inequality is equivalent to

$$f(b) \leq \frac{c-b}{c-a} f(a) + \frac{b-a}{c-a} f(c).$$

This is equivalent to the inequality in the definition of "weakly convex", because the coefficients $1-t$ and t in the expression $b = (1-t)a + tc$ are exactly $1-t = \frac{c-b}{c-a}$ and $t = \frac{b-a}{c-a}$. ☐

Criteria for convexity: increasing first derivatives

Theorem. *Let $f\colon I \to \mathbb{R}$ be a differentiable function. Then:*
$f\colon I \to \mathbb{R}$ is weakly convex if and only if $f'\colon I \to \mathbb{R}$ is weakly increasing.
$f\colon I \to \mathbb{R}$ is strictly convex if and only if $f'\colon I \to \mathbb{R}$ is strictly increasing.

Proof for "weakly". Suppose that f is weakly convex. Take any $a < b$ in I. By the above lemma, for any u, v such that $a < u < v < b$,

$$\frac{f(u)-f(a)}{u-a} \leq \frac{f(v)-f(u)}{v-u} \leq \frac{f(b)-f(v)}{b-v}.$$

In particular,

$$\frac{f(u)-f(a)}{u-a} \leq \frac{f(b)-f(v)}{b-v}.$$

Fixing a, v, b and taking the limit of the left hand side as u approaches a from the right,

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

Fixing a, b and taking the limit of the right hand side as v approaches b from the left,

$$f'(a) \leq f'(b).$$

Since $a < b$ in I were arbitrary, this shows that f' is weakly increasing.

Now suppose that f' is weakly increasing. Take any $a < b < c$ in I. By the intermediate value theorem for f on $[a,b]$ and on $[b,c]$, there exist a point $\overline{x}_1$ between a and b and a point $\overline{x}_2$ between b and c such that

$$\frac{f(b)-f(a)}{b-a} = f'(\overline{x}_1) \quad \text{and} \quad \frac{f(c)-f(b)}{c-b} = f'(\overline{x}_2).$$

Because f' is weakly increasing, $f'(\overline{x}_1) \leq f'(\overline{x}_2)$. So

$$\frac{f(b)-f(a)}{b-a} \leq \frac{f(c)-f(b)}{c-b}.$$

Because $a < b < c$ in I were arbitrary, by the above lemma, the function f is convex. ☐

Exercise: prove the "strictly" part of the theorem. (The direction $\Longrightarrow$ requires some care.)