Appendix C

Convexity

SECTION 1 defines convex sets and functions.
SECTION 2 shows that convex functions defined on subintervals of the real line have left- and right-hand derivatives everywhere.
SECTION 3 shows that convex functions on the real line can be recovered as integrals of their one-sided derivatives.
SECTION 4 shows that convex subsets of Euclidean spaces have nonempty relative interiors.
SECTION 5 derives various facts about separation of convex sets by linear functions.

1. Convex sets and functions

A subset C of a vector space is said to be convex if it contains all the line segments joining pairs of its points, that is,

$$\alpha x_1 + (1-\alpha)x_2 \in C \qquad \text{for all } x_1, x_2 \in C \text{ and all } 0 < \alpha < 1.$$

A real-valued function f defined on a convex subset C (of a vector space $\mathcal{V}$) is said to be convex if

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2) \qquad \text{for all } x_1, x_2 \in C \text{ and } 0 < \alpha < 1.$$

Equivalently, the ***epigraph*** of the function,

$$\text{epi}(f) := \{(x,t) \in C \times \mathbb{R} : t \ge f(x)\},$$

is a convex subset of $C \times \mathbb{R}$. Some authors (such as Rockafellar 1970) define $f(x)$ to equal $+\infty$ for $x \in \mathcal{V}\backslash C$, so that the function is convex on the whole of $\mathcal{V}$, and epi(f) is a convex subset of $\mathcal{V} \times \mathbb{R}$.

This Appendix will establish several facts about convex functions and sets, mostly for Euclidean spaces. In particular, the facts include the following results as special cases.

(i) For a convex function f defined at least on an open interval of the real line (possibly the whole real line), there exists a countable collection of linear functions for which $f(x) = \sup_{i \in \mathbb{N}} (\alpha_i + \beta_i x)$ on that interval.

(ii) If a real-valued function f has an increasing, real-valued right-hand derivative at each point of an open interval, then f is convex on that interval. In particular, if f is twice differentiable, with $f'' \ge 0$, then f is convex.

(iii) If a convex function f on a convex subset $C \subseteq \mathbb{R}^n$ has a local minimum at a point x_0, that is, if $f(x) \geq f(x_0)$ for all x in a neighborhood of x_0, then $f(w) \geq f(x_0)$ for all w in C.

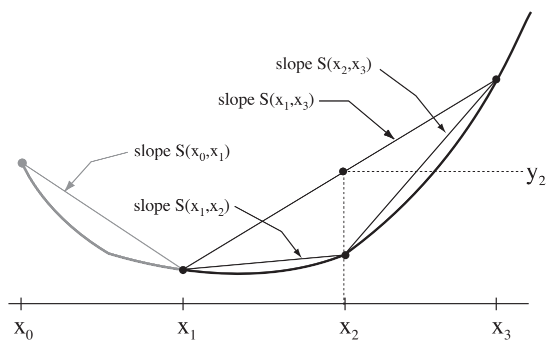
(iv) If C_1 and C_2 are disjoint convex subsets of $\mathbb{R}^n$ then there exists a nonzero ℓ in $\mathbb{R}^n$ for which $\sup_{x\in C_1} x \cdot \ell \leq \inf_{x\in C_2} x \cdot \ell$. That is, the linear functional $x \mapsto x \cdot \ell$ ***separates*** the two convex sets.

2. One-sided derivatives

Let f be a convex function, defined and real-valued at least on an interval J of the real line.

Consider any three points $x_1 < x_2 < x_3$, all in J. (For the moment, ignore the point x_0 shown in the picture.) Write α for $(x_2 - x_1)/(x_3 - x_1)$, so that $x_2 = \alpha x_3 + (1-\alpha)x_1$. By convexity, $y_2 := \alpha f(x_3) + (1-\alpha) f(x_1) \geq f(x_2)$. Write $S(x_i, x_j)$ for $\big(f(x_j) - f(x_i)\big)/(x_j - x_i)$, the slope of the chord joining the points $(x_i, f(x_i))$ and $(x_j, f(x_j))$. Then

$$
\begin{aligned}
S(x_2, x_3) &= \frac{f(x_3) - f(x_2)}{x_3 - x_2} \\
&\geq \frac{f(x_3) - y_2}{x_3 - x_2} = S(x_1, x_3) = \frac{y_2 - f(x_1)}{x_2 - x_1} \\
&\geq \frac{f(x_2) - f(x_1)}{x_2 - x_1} = S(x_1, x_2).
\end{aligned}
$$



From the second inequality it follows that $S(x_1, x)$ decreases as x decreases to x_1. That is, f has right-hand derivative $D_+(x_1)$ at x_1, if there are points of J that are larger than x_1. The limit might equal $-\infty$, as in the case of the function $f(x) = -\sqrt{x}$ defined on $\mathbb{R}^+$, with $x_1 = 0$. However, if there is at least one point x_0 of J for which $x_0 < x_1$ then the limit $D_+(x_1)$ must be finite: Replacing $\{x_1, x_2, x_3\}$ in the argument just made by $\{x_0, x_1, x_2\}$, we have $S(x_0, x_1) \leq S(x_1, x_2)$, implying that $-\infty < S(x_0, x_1) \leq D_+(x_1)$.

The inequality $S(x_1, x) \leq S(x_1, x_2) \leq S(x_2, x')$ if $x_1 < x < x_2 < x'$, leads to the conclusion that D_+ is an increasing function. Moreover, it is continuous from the right, because

$$\begin{aligned} D_+(x_2) \le S(x_2, x_3) &\to S(x_1, x_3) \qquad \text{as } x_2 \downarrow x_1, \text{ for fixed } x_3 \\ &\to D_+(x_1) \qquad \text{as } x_3 \downarrow x_1. \end{aligned}$$

Analogous arguments show that $S(x_0, x_1)$ increases to a limit $D_-(x_1)$ as x_0 increases to x_1. That is, f has left-hand derivative $D_1(x_1)$ at x_1, if there are points of J that are smaller than x_1.

If x_1 is an interior point of J then both left-hand and right-hand derivatives exist, and $D_-(x_1) \le D_+(x_1)$. The inequality may be strict, as in the case where $f(x) = |x|$ with $x_1 = 0$. The left-hand derivative has properties analogous to those of the right-hand derivative. The following Theorem summarizes.

<1> **Theorem.** *Let f be a convex, real-valued function defined (at least) on a bounded interval $[a, b]$ of the real line. The following properties hold.*

(i) The right-hand derivative $D_+(x)$ exists,

$$\frac{f(y) - f(x)}{y - x} \downarrow D_+(x) \qquad \text{as } y \downarrow x,$$

for each x in $[a, b)$. The function $D_+(x)$ is increasing and right-continuous on $[a, b)$. It is finite for $a < x < b$, but $D_+(a)$ might possibly equal $-\infty$.

(ii) The left-hand derivative $D_-(x)$ exists,

$$\frac{f(x) - f(z)}{x - z} \uparrow D_-(x) \qquad \text{as } z \uparrow x,$$

for each x in $(a, b]$. The function $D_-(x)$ is increasing and left-continuous function on $(a, b]$. It is finite for $a < x < b$, but $D_-(b)$ might possibly equal $+\infty$.

(iii) For $a \le x < y \le b$,

$$D_+(x) \le \frac{f(y) - f(x)}{y - x} \le D_-(y).$$

(iv) $D_-(x) \le D_+(x)$ for each x in (a, b), and

$$f(w) \ge f(x) + c(w - x) \qquad \text{for all } w \text{ in } [a, b],$$

for each real c with $D_-(x) \le c \le D_+(x)$.

Proof. Only the second part of assertion (iv) remains to be proved. For $w > x$ use

$$\frac{f(w) - f(x)}{w - x} = S(x, w) \ge D_+(x) \ge c;$$

for $w < x$ use

$$\frac{f(x) - f(w)}{x - w} = S(w, x) \le D_-(x) \le c,$$

□ where $S(\cdot, \cdot)$ denotes the slope function, as above.

<2> **Corollary.** *If a convex function f on a convex subset $C \subseteq \mathbb{R}^n$ has a local minimum at a point x_0, that is, if $f(x) \ge f(x_0)$ for all x in a neighborhood of x_0, then $f(w) \ge f(x_0)$ for all w in C.*

Proof. Consider first the case $n = 1$. Suppose $w \in C$ with $w > x_0$. The right-hand derivative $D_+(x_0) = \lim_{y \downarrow x_0} (f(y) - f(x_0))/(y - x_0)$ must be nonnegative, because $f(y) \geq f(x_0)$ for y near x_0. Assertion (iv) of the Theorem then gives

$$f(w) \geq f(x_0) + (w - x_0)D_+(x_0) \geq f(x_0).$$

The argument for $w < x_0$ is similar.

For general $\mathbb{R}^n$, apply the result for $\mathbb{R}$ along each straight line through x_0. □

Existence of finite left-hand and right-hand derivatives ensures that f is continuous at each point of the open interval (a, b). It might not be continuous at the endpoints, as shown by the example

$$f(x) = \begin{cases} -\sqrt{x} & \text{for } x > 0 \\ 1 & \text{for } x = 0. \end{cases}$$

Of course, we could recover continuity by redefining $f(0)$ to equal 0, the value of the limit $f(0+) := \lim_{w \downarrow 0} f(w)$.

<3> **Corollary.** *Let f be a convex, real-valued function on an interval $[a, b]$. There exists a countable collection of linear functions $d_i + c_i w$, for which the convex function $\psi(w) := \sup_{i \in \mathbb{N}} (d_i + c_i w)$ is everywhere $\leq f(w)$, with equality except possibly at the endpoints $w = a$ or $w = b$, where $\psi(a) = f(a+)$ and $\psi(b) = f(b-)$.*

Proof. Let $\mathcal{X}_0 := \{x_i : i \in \mathbb{N}\}$ be a countable dense subset of (a, b). Define $c_i := D_+(x_i)$ and $d_i := f(x_i) - c_i x_i$. By assertion (iv) of the Theorem, $f(w) \geq d_i + c_i w$ for $a \leq w \leq b$ for each i, and hence $f(w) \geq \psi(w)$.

If $a < w < b$ then (iv) also implies that $f(x_i) \geq f(w) + (x_i - w)D_+(w)$, and hence

$$\psi(w) \geq f(x_i) + c_i(w - x_i) \geq f(w) - (x_i - w)\left(D_+(x_i) - D_+(w)\right) \qquad \text{for all } x_i.$$

Let x_i decrease to w (through $\mathcal{X}_0$) to conclude, via right-continuity of D_+ at w, that $\psi(w) \geq f(w)$.

If $D_+(a) > -\infty$ then f is continuous at a, and

$$f(a) \geq \psi(a) \geq \limsup_{x_i \downarrow a} \left(f(x_i) + (a - x_i)c_i\right) = f(a+) = f(a).$$

If $D_+(a) = -\infty$ then f must be decreasing in some neighborhood $\mathcal{N}$ of a, with $c_i < 0$ when $x_i \in \mathcal{N}$, and

$$\psi(a) \geq \sup_{x_i \in \mathcal{N}} \left(f(x_i) + (a - x_i)c_i\right) \geq \sup_{x_i \in \mathcal{N}} f(x_i) = f(a+).$$

If $\psi(a)$ were strictly greater than $f(a+)$, the open set

$$\{w : \psi(w) > f(a+)\} = \cup_i \{w : d_i + c_i w > f(a+)\}$$

would contain a neighborhood of a, which would imply existence of points w in $\mathcal{N} \backslash \{a\}$ for which $\psi(w) > f(a+) \geq f(w)$, contradicting the inequality $\psi(w) \leq f(w)$. A similar argument works at the other endpoint. □

3. Integral representations

Convex functions on the real line are expressible as integrals of one-sided derivatives.