Convex functions of one real variable

In what follows, $I \subset \mathbb{R}$ will be a nondegenerate interval, and $f\colon I \to \mathbb{R}$ a (finite) convex function.

For $x, y \in I$, $x \neq y$, we denote by $Q(x,y)$ the correspondent difference quotient:

$$Q(x,y) = \frac{f(y) - f(x)}{y - x}\,.$$

Lemma 0.1. *Let I, f be as above, and $x, y, z \in I$. Then*

$$x < y < z \quad \Longrightarrow \quad Q(x,y) \le Q(x,z) \le Q(y,z)\,.$$

In other words, the difference quotient $Q(\cdot,\cdot)$ is nondecreasing in each of the two variables.

Proof. Since f is convex, the point $P_2 := (y, f(y))$ lies below (or on) the segment with endpoints $P_1 := (x, f(x))$ and $P_3 := (z, f(z))$, we have the following inequalities for slopes

$$\text{slope}\left(\overleftrightarrow{P_1P_2}\right) \le \text{slope}\left(\overleftrightarrow{P_1P_3}\right) \le \text{slope}\left(\overleftrightarrow{P_2P_3}\right)\,,$$

which are equivalent to the required inequalities for difference quotients.

For completeness, let us give a formal proof, too. Write y as a convex combination $y = \frac{z-y}{z-x}x + \frac{y-x}{z-x}z$. Then, by convexity,

$$f(y) \le \tfrac{z-y}{z-x} f(x) + \tfrac{y-x}{z-x} f(z)\,. \tag{1}$$

Substracting $f(x)$ from both sides of (1), we get

$$f(y) - f(x) \le \tfrac{y-x}{z-x}\big[f(z) - f(x)\big]$$

which gives the first inequality. Substracting $\frac{y-x}{z-x}f(z) + \frac{z-y}{z-x}f(y)$ from both sides of (1), we get

$$\tfrac{y-x}{z-x}\,[f(y) - f(z)] \le \tfrac{z-y}{z-x}\,[f(x) - f(y)]$$

which gives the second inequality. $\square$

Theorem 0.2 (behaviour at interior points)**.** *Let $I \subset R$ be an open interval, $f\colon I \to \mathbb{R}$ a convex function.*

(a) *For each $x \in I$, both one-sided derivatives $f'_+(x), f'_-(x)$ exist and are finite.*
(b) *f is Lipschitz on each compact interval $[a,b] \subset I$. (In particular, f is continuous on I.)*
(c) *If $x, y \in I$, $x < y$, then $f'_-(x) \le f'_+(x) \le f'_-(y) \le f'_+(y)$. (In particular, f'_+ and f'_- are nondecreasing on I.)*
(d) *f'_+ is right-continuous, f'_- is left continuous.*
(e) *The set of nondifferentiability points of f is at most countable.*

Proof. (a) Let $x \in I$. Fix $\delta > 0$ such that $x \pm \delta \in I$. Then, by Lemma 0.1, we have $Q(x-\delta, x) \le Q(x, x+h)$ whenever $h \in (0,\delta)$. Since the function $h \mapsto Q(x, x+h)$ is nondecreasing on $(0,\delta)$, it admits a limit for $h \to 0^+$, and

$$Q(x-\delta, x) \le f'_+(x) = \lim_{h\to 0+} Q(x, x+h) = \inf_{0<h<\delta} Q(x, x+h)\,.$$

Hence $f'_+(x)$ is finite. The proof for $f'_-(x)$ is analogous.

(b) Let $x, y \in [a,b]$, $x < y$. Let $\delta > 0$ be such that $a + \delta < y$ and $b - \delta > x$. For each $h \in (0, \delta)$, we have by Lemma 0.1

$$Q(a, a+h) \le Q(x,y) \le Q(b-h, b).$$

Passing to limits for $h \to 0^+$, we get

$$f'_+(a) \le Q(x,y) \le f'_-(b).$$

Consequently, $|f(y) - f(x)| \le L|y-x|$ where $L = \max\bigl\{|f'_+(a)|, |f'_-(b)|\bigr\}$.
(c) Fix $\delta > 0$ such that $x - \delta \in I$, $y + \delta \in I$ and $x + \delta < y - \delta$. By Lemma 0.1,

$$Q(x-h, x) \le Q(x, x+h) \le Q(y-h, x) \le Q(y, y+h) \qquad (0 < h < \delta).$$

Pass to limits for $h \to 0^+$.
(d) Fix $x_0 \in I$. Let $y \in I$ be such that $y > x_0$. Lemma 0.1 implies that, for each $x \in (x_0, y)$,

$$f'_+(x) \le Q(x,y).$$

Passing to limit for $x \to (x_0)^+$, we get

$$\lim_{x \to (x_0)+} f'_+(x) \le Q(x_0, y)$$

(the limit on the left-hand side exists since f'_+ is nondecreasing). Passing to limit for $y \to (x_0)^+$, we obtain

$$\lim_{x \to (x_0)+} f'_+(x) \le f'_+(x_0)\,.$$

The other inequality follows immediately from (c).
(e) Let N_f be the set of nondifferentiability points of f. Then

$$N_f = \{x \in I : f'_-(x) < f'_+(x)\}\,.$$

Observe that (c) implies that the (nonempty) open intervals $J_x := \bigl(f'_-(x), f'_+(x)\bigr)$ $(x \in N_f)$ are pairwise disjoint. Since each of them contains a rational number, they are at most countably many. ☐

Corollary 0.3. *Let I, f be as in the previous theorem. Then, for each $x \in I$,*

$$f'_+(x) = \lim_{y \to x+} f'_-(y)\,, \quad f'_-(x) = \lim_{y \to x-} f'_+(y)\,.$$

Proof. The first formula follows from Theorem 0.2(c,d), since for $y > x$ we have $f'_+(x) \le f'_-(y) \le f'_+(y)$. The second formula is analogous. ☐

Corollary 0.4. *Let I, f be as in Theorem 0.2. Then, for each $a, b \in I$,*

$$f(b) - f(a) = (\mathrm{L})\!\int_a^b f'(x)\,dx = (\mathrm{R})\!\int_a^b f'_+(x)\,dx = (\mathrm{R})\!\int_a^b f'_-(x)\,dx$$

(the letters "L, R" indicate that the integral is intended in the Lebesgue or Riemann sense).

Proof. Let $a < b$. By Theorem 0.2(b), f is Lipschitz on $[a, b]$. Since Lipschitz functions are absolutely continuous, the first equality follows from a well known result of Real Analysis. The other two equalities are clear for Lebesgue integrals, since $f'_\pm(x) = f'(x)$ except a countable set N_f. Observe that the last two integrals exist also in the Riemann sense, since the one-sided derivatives $f'_\pm$ are bounded and monotone on $[a, b]$. □

Proposition 0.5 (behaviour at endpoints)**.** *Let $I = [a, b]$, $f : I \to \mathbb{R}$ a convex function.*

(a) *f is upper-semicontinuous, but not necessarily continuous, at a (from the right) and at b (from the left).*
(b) *f is bounded on $[a, b]$.*
(c) *$f'_+(a)$ exists in $[-\infty, +\infty)$, and $f'_-(b)$ exists in $(-\infty, +\infty]$.*
(d) *If f is right-continuous at a, then $f'_+(a) = \lim_{x\to a+} f'_+(x) = \lim_{x\to a+} f'_-(x)$. If f is left-continuous at b, then $f'_-(b) = \lim_{x\to b-} f'_-(x) = \lim_{x\to b-} f'_+(x)$.*
(e) *If $f'_+(a) > -\infty$, then f is Lipschitz on $[a, c]$ whenever $a < c < b$. (In particular, f is right-continuous at a.)*
If $f'_-(b) < +\infty$, then f is Lipschitz on $[c, b]$ whenever $a < c < b$. (In particular, f is left-continuous at b.)

Proof. (a) We have to prove that $f(a) \geq \limsup_{x\to a+} f(x)$. Fix an arbitrary $c \in (a, b)$ and observe that

$$\limsup_{x\to a+} f(x) = \limsup_{t\to 0+} f\big((1-t)a + tc\big) \leq \limsup_{t\to 0+}\big[(1-t)f(a) + tf(c)\big] = f(a).$$

To see that f need not be continuous at a, consider the function $f(a) = 1$, $f(x) = 0$ for $x \in (a, b)$.

(b) For each $x \in (a, b)$, $f'_+(a) \leq Q(a, x)$; thus $f(x) \geq f(a) + f'_+(a)(x - a) \geq f(a) - |f'_+(a)|(b - a)$. Consequently, f is lower bounded on $[a, b]$. If $x = (1-t)a + tb$ with $t \in [0, 1]$, then $f(x) \leq (1-t)f(a) + tf(b) \leq \max\{f(a), f(b)\}$.

(c) can be proved in the same way as Theorem 0.2(a); (d) can be proved in the same way as Theorem 0.2(d) and Corollary 0.3; (e) can be proved in the same way as Theorem 0.2(b). □

Exercise 0.6. Let $f : [a, b] \to \mathbb{R}$ be a continuous convex function. Then $f(b) - f(a) = (\mathrm{L})\int_a^b f'(x)\,dx = (\mathrm{R})\int_a^b f'_+(x)\,dx = (\mathrm{R})\int_a^b f'_-(x)\,dx$.

Subdifferential. Let $I \subset \mathbb{R}$ be an open interval, and $f : I \to \mathbb{R}$ a convex function. Given $x_0 \in I$, each non-vertical line passing through the point $(x_0, f(x_0))$ is of the form $y = f(x_0) + m(x - x_0)$. Thus the following definition defines $\partial f(x_0)$ as the set of all angular coefficients m for which the corresponding line lies under (in the weak sense) the (epi)graph of f. By 2^E we denote the set of all subsets of a set E.

Definition 0.7. Let I, f, x_0 be as above. The *subdifferential* of f at x_0 is the set

$$\partial f(x_0) = \{m \in \mathbb{R} : \ f(x) \geq f(x_0) + m(x - x_0) \text{ whenever } x \in I\}.$$

Any $m \in \partial f(x_0)$ is called a *subgradient* of f at x_0.
The multivalued mapping $\partial f \colon I \to 2^{\mathbb{R}}$, $x \mapsto \partial f(x)$, is called the *subdifferential mapping* (or just the *subdifferential*) of the function f.

Proposition 0.8. *Let $I \subset \mathbb{R}$ be an open interval, $f \colon I \to \mathbb{R}$ a convex function, $x_0 \in I$. Then*

$$\partial f(x_0) = \left[f'_-(x_0), f'_+(x_0)\right].$$

Proof. The formula is geometrically clear. However, we give also a formal proof. Let $m \in \partial f(x_0)$. Let $\delta > 0$ be such that $x_0 \pm \delta \in I$. An elementary calculation shows that, for any $h \in (0,\delta)$, we have $Q(x_0 - h, x_0) \le m \le Q(x_0, x_0 + h)$. Passing to limits for $h \to 0^+$, we get $f'_-(x_0) \le m \le f'_+(x_0)$.

Now, let $m \in \left[f'_-(x_0), f'_+(x_0)\right]$. Fix $x \in I$. If $x > x_0$, we have $Q(x_0, x) \ge f'_+(x_0) \ge m$, and this immediately implies $f(x) \ge m(x - x_0)$. If $x < x_0$, we have $Q(x, x_0) \le f'_-(x_0) \le m$, which implies $f(x) \ge m(x - x_0)$ again. Thus $m \in \partial f(x_0)$. □

We say that a function $\varphi \colon I \to \mathbb{R}$ is a *single-valued selection* of ∂f if $\varphi(x) \in \partial f(x)$ for each $x \in I$.

Corollary 0.9. *Let I, f, x_0 be as in Proposition 0.8. Then the following assertions are equivalent:*

(i) *f is differentiable at x_0;*
(ii) *$\partial f(x_0)$ is a singleton;*
(iii) *∂f admits a single-valued selection which is continuous at x_0.*
(iii) *each single-valued selection of ∂f is continuous at x_0.*

Proof. Exercise, based on Proposition 0.8, Theorem 0.2(d), and Corollary 0.3. □

Second-order differentiability. We would like to define second-order differentiability for convex functions which are not necessarily everywhere differentiable. There are several possibilities to do that. The following theorem shows that they are all equivalent.

Theorem 0.10. *Let $I \subset \mathbb{R}$ be an open interval, $f \colon I \to \mathbb{R}$ a convex function, $x_0 \in I$, $\Delta \in \mathbb{R}$. Then the following assertions are equivalent.*

(i) *f is differentiable at x_0 and, denoting by D_1 the set of differentiability points of f in I, one has*

$$\lim_{\substack{x \to x_0 \\ x \in D_1}} \frac{f'(x) - f'(x_0)}{x - x_0} = \Delta.$$

(ii) *f'_+ is differentiable at x_0 with $(f'_+)'(x_0) = \Delta$.*
(iii) *f'_- is differentiable at x_0 with $(f'_-)'(x_0) = \Delta$.*
(iv) *Some single-valued selection φ of ∂f is differentiable at x_0 with $\varphi'(x_0) = \Delta$.*
(v) *Each single-valued selection φ of ∂f is differentiable at x_0 with $\varphi'(x_0) = \Delta$.*
(vi) *f is differentiable at x_0 and*

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \Delta \tfrac{h^2}{2} + o(h^2) \qquad (\text{as } h \to 0).$$

Moreover, $\Delta \geq 0$ in this case.

Proof.
$(i) \Rightarrow (ii)$. For $x_0 + h \in I$, that is $h \in I \setminus x_0$, define

$$\omega(h) = f'_+(x_0 + h) - f'(x_0) - \Delta h\,. \tag{2}$$

We have that $\lim_{h\to 0,\, h\in D_1 - x_0} \frac{\omega(h)}{h} = 0$ by (1). We want to prove that $\omega(h) = o(h)$ as $h \to 0$. For each $0 \neq h \in D_1 - x_0$ choose $s_h \in (h, h + h^2) \cap (D_1 - x_0)$ and $s'_h \in (h - h^2, h) \cap (D_1 - x_0)$. Notice that $f'_+(s'_h) \leq f'_+(h) \leq f'_+(s_h)$. Consequently, for $I - x_0 \ni h > 0$ we have

$$\frac{\omega(h)}{h} \leq \frac{\omega(s_h) + \Delta(s_h - h)}{h} = \frac{\omega(s_h)}{h} + \Delta\left(\frac{s_h}{h} - 1\right) \to 0 \quad \text{as } h \to 0^+,$$

and also

$$\frac{\omega(h)}{h} \geq \frac{\omega(s'_h) + \Delta(s'_h - h)}{h} = \frac{\omega(s'_h)}{h} + \Delta\left(\frac{s'_h}{h} - 1\right) \to 0 \quad \text{as } h \to 0^+.$$

The same with inverse inequalities holds for $I - x_0 \ni h < 0$. This proves (ii). Analogously we obtain the implication $(i) \Rightarrow (iii)$.
$(ii) \Rightarrow (vi)$. f'_+ is continuous at x_0, hence

$$f'_-(x_0) = \lim_{x\to x_0^-} f'_+(x) = \lim_{x\to x_0^+} f'_+(x) = f'_+(x_0)$$

and f is differentiable at x_0. Now, let ω be the function from (2). Then $\omega(0) = 0$, $\omega(h) = o(h)$ as $h \to 0$ (in particular, ω is continuous at 0), and ω is Riemann-integrable on each compact interval contained in $I - x_0$. Using Corollary 0.4, we get (for $h \in I - x_0$)

$$f(x_0 + h) - f(x_0) = \int_{x_0}^{x_0+h} f'(x)\,dx = f'(x_0)h + \Delta\frac{h^2}{2} + \int_0^h \omega(t)\,dt\,.$$

Notice that the last integral is $o(h^2)$ by the De L'Hopital rule. In the same way, $(iii) \Rightarrow (vi)$ holds.
$(vi) \Rightarrow (v)$. By substracting an affine function, we can suppose that $f(x_0) = f'(x_0) = 0$. In this case, $f(x_0 + h) = \Delta\frac{h^2}{2} + o(h^2)$ for $h \to 0$; and $\varphi(x_0) = 0$. Thus, for each $\varepsilon \in (0, 1)$,

$$\begin{aligned}
\limsup_{h\to 0+} \frac{\varphi(x_0 + h) - \varphi(x_0)}{h} &\leq \limsup_{h\to 0+} \frac{f'_+(x_0 + h)}{h} \leq \limsup_{h\to 0+} \frac{Q(x_0 + h, x_0 + h + \varepsilon h)}{h} \\
&= \limsup_{h\to 0+} \frac{1}{\varepsilon h^2}\left[f(x_0 + h + \varepsilon h) - f(x_0 + h)\right] \\
&= \limsup_{h\to 0+} \frac{1}{\varepsilon h^2}\left[\frac{\Delta\, h^2\big((1+\varepsilon)^2 - 1\big)}{2} + o(h^2)\right] \\
&= \limsup_{h\to 0+} \left[\frac{\Delta(2+\varepsilon)}{2} + o(1)\right] = \Delta\left(1 + \frac{\varepsilon}{2}\right).
\end{aligned}$$

In a similar way, we obtain

$$\liminf_{h\to 0+} \frac{\varphi(x_0+h)-\varphi(x_0)}{h} \geq \liminf_{h\to 0+} \frac{f'_-(x_0+h)}{h} \geq \liminf_{h\to 0+} \frac{Q(x_0+h-\varepsilon h, x_0+h)}{h}$$
$$= \Delta\left(1-\frac{\varepsilon}{2}\right).$$

By arbitrariety of ε, $\varphi'_+(x_0)=\Delta$. In a completely analogous way, one obtains that $\varphi'_-(x_0)=\Delta$.
The implication $(v)\Rightarrow(iv)$ is obvious. Moreover, $(iv)\Rightarrow(i)$ is immediate since $\varphi(x)=f'(x)$ for each $x\in D_1$.
Finally, $\Delta\geq 0$ since it is the derivative at x_0 of the nondecreasing function f'_+. □

Observation 0.11. *Let $I\subset\mathbb{R}$ be an open interval, $f\colon I\to\mathbb{R}$ a differentiable (not necessarily convex) function, $x_0\in I$. If f is twice differentiable at x_0, then the condition (vi) of Theorem 0.10 holds (with $\Delta=f''(x_0)$). However, the vice-versa is not necessarily true (consider the function $f(x)=x^3\sin(1/x)$, $f(0)=0$, at $x_0=0$). On the other hand, Theorem 0.10 implies that the vice-versa holds for convex functions.*

Definition 0.12. Let $I\subset\mathbb{R}$ be an open interval. A convex function f on I is said to be *twice differentiable* at a point $x_0\in I$ if the equivalent conditions (i)-(vi) in Theorem 0.10 are satisfied.

As a corollary, we obtain the following important result.

Theorem 0.13. *Every convex function on an open interval $I\subset\mathbb{R}$ is twice differentiable almost everywhere (with respect to the Lebesgue measure).*

Proof. The nondecreasing function f'_+ is differentiable almost everywhere by a well known Lebesgue's theorem in Real Analysis. □