Chapter 3

Regularity of convex functions

3.1 Convexity and continuity

Proposition 3.1.15 (Continuity of convex functions) *Let $f : E \to \mathbb{R} \cup \{+\infty\}$ be a convex function and $x_0 \in int(dom(f))$. If f is upper bounded in the neighborhood of x_0, it is continuous, and also Lipschitz, in the neighborhood of x_0.*

Proof. i) We first prove the continuity of f at x_0. Let $r_0 > 0$ and $M \in \mathbb{R}$ such that $f(x) \leq M$ for all $x \in \bar{B}(x_0, r_0)$.

For $x \in \bar{B}(x_0, r_0)$ define y such that $y - x_0 = r_0 \frac{x - x_0}{|x - x_0|}$ and set $\lambda = \frac{|x - x_0|}{r_0}$, then $y \in \bar{B}(x_0, r_0)$ and $\lambda \in [0, 1]$. By convexity we have

$$\begin{aligned} f(x) - f(x_0) &= f(\lambda y + (1-\lambda)x_0) - f(x_0) \\ &\leq \lambda f(y) + (1-\lambda) f(x_0) - f(x_0) \\ &\leq \lambda(f(y) - f(x_0)) \\ &\leq \lambda\big(M - f(x_0)\big) = \frac{|x - x_0|}{r_0}(M - f(x_0)) \,. \end{aligned} \tag{3.1.1}$$

Define $z = x_0 - (x - x_0)$. Then, $z \in B(x_0, r_0)$ and $x_0 = (x+z)/2$. We then have

$$f(x_0) = f\big(\frac{x+z}{2}\big) \leq \frac{f(x) + f(z)}{2}$$

and

$$f(z) - f(x_0) \quad \geq \quad -(f(x) - f(x_0)) \ .$$

Applying (3.1.1) to z we get

$$\begin{aligned} \frac{|x - x_0|}{r_0}(M - f(x_0)) \quad &= \quad \frac{|z - x_0|}{r_0}(M - f(x_0)) \\ &\geq \quad f(z) - f(x_0) \quad \geq \quad -(f(x) - f(x_0)) \end{aligned}$$

and

$$|f(x) - f(x_0)| \quad \leq \quad \frac{|x - x_0|}{r_0}(M - f(x_0))$$

for all $x \in \bar{B}(x_0, r_0)$, which gives the continuity of f at x_0.

ii) We now prove that f is Lipschitz continuous on $B(x_0, r)$, for $r < r_0$. For $x \in B(x_0, r)$, we have $\bar{B}(x, r_0 - r) \subset \bar{B}(x_0, r_0)$. Therefore f is upper bounded by M on $\bar{B}(x, r_0 - r)$. We then get

$$|f(y) - f(x)| \quad \leq \quad \frac{|y - x|}{r_0 - r}(M - f(x))$$

for all $y \in B(x, r_0 - r)$. Since $x \in B(x_0, r_0)$, we have $f(x) \geq f(x_0) - (M - f(x_0)) = 2f(x_0) - M$. Therefore

$$|f(y) - f(x)| \quad \leq \quad 2\frac{|y - x|}{r_0 - r}(M - f(x_0))$$

for all $y \in \bar{B}(x, r_0 - r)$. For $x, y \in B(x_0, r)$, we divide $[x, y]$ into consecutive sets $[x_i, x_{i+1}]$, $i = 0, \ldots, N$ with length strictly smaller than $r_0 - r$. Since $[x, y] \subset B(x_0, r)$ we get

$$\begin{aligned} |f(y) - f(x)| \quad &\leq \quad \sum_{i=1}^{N-1} |f(x_{i+1}) - f(x_i)| \\ &\leq \quad \sum_{i=1}^{N-1} L|x_{i+1} - x_i| \quad = \quad L|x - y| \end{aligned}$$

with $L = 2\frac{M - f(x_0)}{r_0 - r}$. $\square$

Definition 3.1.13 (Locally Lipschitz functions) *A function f is said to be locally Lipschitz on $int(dom(f))$ if it is Lipschitz in the neighborhood of all points of $int(dom(f))$.*

Corollary 3.1.5 *Let $f: E \to \mathbb{R} \cup \{+\infty\}$ be a convex function. The following assertions are equivalent.*

i) there exists $x_0 \in dom(f)$ such that f is upper bounded at the neighborhood of x_0.

ii) There exists $x_0 \in dom(f)$ such that f is continuous at x_0.

iii) $int(dom(f)) \neq \emptyset$ and f is continuous on $dom(f)$.

iv) $int(dom(f)) \neq \emptyset$ and f is locally Lipschitz on $dom(f)$.

Proof. We obviously have $iv) \Rightarrow iii) \Rightarrow ii) \Rightarrow i)$.

We now prove $i) \Rightarrow iv)$. Suppose f upper bounded on $\bar{B}(x_0, r_0)$. Then, $x_0 \in int(dom(f)) \neq \emptyset$. Let $y_0 \in int(dom(f))$. There exists $z_0 \in int(dom(f))$ such that $y_0 \in [x_0, z_0)$. Let h be the homothety centered at z_0 with coefficient λ such that $h(x_0) = y_0$. We then have $h(x) = z_0 + \lambda(x - z_0) = \lambda x + (1-\lambda)z_0$, with $0 < \lambda \leq 1$ and $h(\bar{B}(x_0, r_0)) = \bar{B}(y_0, \lambda r_0)$. By convexity of f we have

$$\begin{aligned} f(h(x)) &= f(\lambda x + (1-\lambda)z_0) \leq \lambda f(x) + (1-\lambda)f(z_0) \\ &\leq \max(M, f(z_0)) \end{aligned}$$

for all $x \in \bar{B}(x_0, r_0)$. This implies that f is upper bounded on $\bar{B}(y_0, \lambda r_0)$, and therefore Lipschitz continuous on $\bar{B}(y_0, \lambda r)$, for all $r < r_0$, by the previous proposition. $\square$

Corollary 3.1.6 *Let $f: E \to \mathbb{R} \cup \{+\infty\}$ be a convex function. Then, f is continuous and even locally Lipschitz on $int(dom(f))$.*

Proof. Suppose that $int(dom(f)) \neq \emptyset$ and w.l.o.g. $0 \in int(dom(f))$. Let $r_0 > 0$ such that $\bar{B}(x_0, r_0) \subset dom(f)$. We can then find $x_1, \ldots, x_n \in \partial B(0, r_0)$ such that $(x_1, \ldots, x_n)$ is a basis of E (recall that E is finite dimensional). Then f is bounded on $conv(x_1, -x_1, \ldots, x_n, -x_n)$. Since $(x_1, \ldots, x_n)$ is a basis of E we have $int(conv(x_1, -x_1, \dots, x_n, -x_n)) \neq \emptyset$. From the previous proposition f is Lipschitz continuous on $int(conv(x_1, -x_1, \dots, x_n, -x_n))$ which contains 0. $\square$

Corollary 3.1.7 *Let $f : E \to \mathbb{R} \cup \{+\infty\}$ be a convex function. Then, f is continuous at a point of its domain if and only if $int(epi(f)) \neq \emptyset$.*

Proof. The continuity assumption implies that $int(epi(f)) \neq \emptyset$. Indeed, if $f \leq M$ on $B(y_0, r)$ then $B(y_0, r) \times (M, +\infty) \subset epi(f)$. The reverse implication is a consequence of Corollary 3.1.5. $\square$

3.2 Convexity and differentiability

Definition 3.2.14 (Directional derivative) *Let $f : E \to \mathbb{R} \cup \{+\infty\}$, $x_0 \in dom(f)$ and $h \in E \setminus \{0\}$. The function f is said to be right-differentiable at x_0 in the direction h if*

$$\lim_{t \to 0+} \frac{f(x_0 + th) - f(x_0)}{t}$$

exists in $\mathbb{R} \cup \{+\infty\}$. When it exists, this limit is denoted by $f'_d(x_0, h)$ and is called directional derivative of f in the direction h at x_0.

Proposition 3.2.16 *A convex function $f : E \to \mathbb{R} \cup \{+\infty\}$ is right-differentiable at any $x_0 \in dom(f)$ in any direction h. Moreover, we have the inequality*

$$f(x) - f(x_0) \quad \geq \quad f'_d(x_0, x - x_0) \tag{3.2.2}$$

for all $x \in E$.

This result is a direct consequence of the following lemma.

Lemma 3.2.2 *Let $f : E \to \mathbb{R} \cup \{+\infty\}$ be a convex function, $x_0 \in dom(f)$ and $h \in E \setminus \{0\}$. The function $\Delta_{x_0,h} : (0, +\infty) \to \mathbb{R} \cup \{+\infty\}$ defined by*

$$\Delta_{x_0,h}(t) \quad = \quad \frac{f(x_0 + th) - f(x_0)}{t}\,, \quad t \in (0, +\infty)$$

is nondecreasing.

Proof. Define $\phi(t) = f(x_0 + th) - f(x_0)$. ϕ is a convex function from $[0, +\infty)$ to $\mathbb{R} \cup \{+\infty\}$, with $\phi(0) = 0$. Therefore

$$\phi(\alpha t) \;=\; \phi(\alpha t + (1-\alpha).0) \;\leq\; \alpha\phi(t) + (1-\alpha)\phi(0) \;=\; \alpha\phi(t)$$

for all $\alpha, t \in [0, +\infty)$. Setting $s = \alpha t$, we get $\phi(s) \leq (s/t)\phi(t)$ and $\phi(s)/s \leq \phi(t)/t$ for $s > 0$.

This means $\Delta_{x_0,h}(s) \leq \Delta_{x_0,h}(t)$. Therefore $\Delta_{x_0,h}$ is a nondecreasing function from $(0, +\infty)$ to $\mathbb{R} \cup \{+\infty\}$. □

Corollary 3.2.8 (Minimum of a convex function) *A convex proper function f reaches its minimum at x_0 if and only if all its directional derivatives at x_0 are nonnegatives.*

Proof. Let x_0 be a minimum point of f. Then we have

$$f(x_0 + th) - f(x_0) \;\geq\; 0$$

for any $t > 0$ and any $h \in E$. Sending t to $0+$ we get $f'_d(x_0, h) \geq 0$ for any $h \in E$.

Conversely, let x_0 with nonnegative directional derivatives. Using (3.2.2), we get that x_0 is a minimum point of f. □

We turn to differentiable convex functions.

Proposition 3.2.17 *Let $f : E \to \mathbb{R} \cup \{+\infty\}$ be a function differentiable on $dom(f)$ that is suppose to be convex. Then f is convex if and only if*

$$f(y) - f(x) \;\geq\; \langle \nabla f(x), y - x \rangle$$

for all $x \in dom(f)$ and $y \in E$.

Proof. If f is convex, the inequality follows from Proposition 3.2.16.

Conversely, suppose the inequality holds. Let $x, y \in dom(f)$, $\lambda \in [0, 1]$ and $z = \lambda x + (1-\lambda)y$. We then have

$$\begin{aligned} f(x) - f(z) \;&\geq\; \langle \nabla f(z), x - z \rangle \;, \\ f(y) - f(z) \;&\geq\; \langle \nabla f(z), y - z \rangle \;. \end{aligned}$$

Hence

$$\lambda f(x) + (1-\lambda)f(y) - f(z) \geq \langle \nabla f(z), \lambda(x-z) + (1-\lambda)(y-z)\rangle = 0$$

and f is convex. □

We provide a second convexity criterium based on the gradient.

Proposition 3.2.18 *Let $f: E \to \mathbb{R}\cup\{+\infty\}$ be a function differentiable on $dom(f)$ that is suppose to be convex. Then f is convex if and only if*

$$\langle \nabla f(y) - \nabla f(x), y - x\rangle \quad \geq \quad 0$$

for all $x, y \in dom(f)$.

Proof. i) Suppose that f is convex. let $x, y \in dom(f)$. From Proposition 3.2.16 we have

$$\begin{aligned} f(y) - f(x) &\geq \langle \nabla f(x), y - x\rangle , \\ f(x) - f(y) &\geq \langle \nabla f(y), x - y\rangle . \end{aligned}$$

Summing these two inequalities gives the result.

ii) Conversely, supposons that ∇f satisfies the inequality. Let $x, y \in dom(f)$ and consider the map $\phi : [0,1] \to \mathbb{R}$, $\lambda \mapsto \phi(\lambda) = f(\lambda x + (1-\lambda)y) = f(y + \lambda(x-y))$. ϕ is differentiable with derivative given by

$$\phi'(\lambda) \quad = \quad \langle \nabla f(y + \lambda(x-y)), x - y\rangle$$

For $\lambda_1 < \lambda_2$ we have

$$\phi'(\lambda_1) - \phi'(\lambda_2) \quad = \quad \langle \nabla f(y + \lambda_1(x-y)) - \nabla f(y + \lambda_2(x-y)), x - y\rangle \quad \leq \quad 0$$

since the difference between the arguments of ∇f is $(\lambda_1 - \lambda_2)(x-y)$. We then deduce that ϕ has a nondecreasing derivative and is therefore convex. In particular

$$\phi(\lambda) \quad \leq \quad \lambda\phi(1) + (1-\lambda)\phi(0) \quad = \quad \lambda f(x) + (1-\lambda)f(y)$$

which is the convexity property for f. □

We provide a more precise result on the lower bound in the case of a Lipschitz continuous gradient.

Proposition 3.2.19 (Coercivity of the gradient) *Let f be convex and differentiable function from E to $\mathbb{R}$. We suppose that ∇f L-Lipschitz continuous on E. Then*

$$\langle \nabla f(x) - \nabla f(y), x - y\rangle \quad \geq \quad \frac{1}{L}|\nabla f(x) - \nabla f(y)|^2$$

for $x, y \in E$.

Proof. Fix $y, z \in E$ and g a function from E to $\mathbb{R}$ satisfying the same assumptions as f. Applying the first order Taylor formula to the function $t \in [0,1] \mapsto f(ty + (1-t)z)$ and using L-Lipschitz continuity of ∇f we have

$$g(z) \quad \leq \quad g(y) + \langle \nabla g(y), z - y\rangle + \frac{L}{2}|z - y|^2 \ .$$

Taking $y = x$ and $z = x - \frac{1}{L}\nabla g(x)$ we get

$$\frac{1}{2L}|\nabla g(x)|^2 \quad \leq \quad g(x) - g(x - \frac{1}{L}\nabla f(x)) \ .$$

Suppose now that g is lower bounded. We then have

$$\frac{1}{2L}|\nabla g(x)|^2 \quad \leq \quad g(x) - M \tag{3.2.3}$$

for all $x \in E$, where M is a lower bound of g. We next define the functions h_1 and h_2 by

$$\begin{aligned} h_1(u) &= f(u) - \langle \nabla f(x), u\rangle \ , \\ h_2(u) &= f(u) - \langle \nabla f(y), u\rangle \ , \end{aligned}$$

for $u \in E$. These two fonctions are convex and lower bounded respectively by $h_1(x)$ and $h_2(y)$ from Proposition 3.2.17. Applying (3.2.3) to h_1 and h_2 we get

$$h_1(x) \ \leq \ h_1(y) - \frac{1}{2L}|\nabla h_1(x)|^2 \quad \text{and} \quad h_2(y) \ \leq \ h_2(x) - \frac{1}{2L}|\nabla h_2(y)|^2 \ .$$

Summing these two inequalities gives the result. □

We end this section with a second order property for convex functions.

Proposition 3.2.20 *Let $f : E \to \mathbb{R} \cup \{+\infty\}$ be a function differentiable on dom(f). Suppose that the function $t \in \mathbb{R}_+ \mapsto \nabla f(x+th)$ is differentiable and*

$$\left\langle \frac{\nabla f(x+th) - \nabla f(x)}{t}, h \right\rangle \xrightarrow[t \to 0+]{} Q(x,h)$$

with $Q(x,h) \geq 0$ for all $h \in E$ and all $x \in$ dom(f). Then f is convex.

Proof. From the mean value Theorem, there exists $\theta \in [0,1]$ such that

$$\begin{aligned} \langle \nabla f(y) - \nabla f(x), y - x \rangle &= \frac{d}{dt} \langle \nabla f(y + t(x-y)), x - y \rangle |_{t=\theta} \\ &= Q(\theta x + (1-\theta) y, x - y) \quad \geq \quad 0 \end{aligned}$$

and f is convex from the previous proposition. $\square$

We deduce from this last result that for a function $f : U \to \mathbb{R}$ that is C^2 on the open convex subset U of E, such that

$$\nabla^2 f(x) \quad \geq \quad 0$$

for all $x \in U$, then f is convex on U.