Appendix A
Convex functions and differentiability

We review some basic concepts and properties from the theory of differentiable functions (See, e.g., [9]) and from Convex Analysis (See, e.g., [16] and [88]).

A.1 Maxima of parameterized families

In this section $X$ is a topological space, $D$ is a compact topological space, and $\Phi : X \times D \to \mathbb{R}$ is a continuous function. The function $\Phi$ may also be considered as a family of functions $\Phi_p : X \to \mathbb{R}$, parameterized by $p \in D$, where $\Phi_p(x) = \Phi(x, p)$. We will also use the shorthand notation $\Phi(\cdot, p)$ for $\Phi_p$. Furthermore, we consider the maximum-function $\varphi : X \to \mathbb{R}$, defined by

$$\varphi(x) = \max_{p \in D} \Phi(x, p).$$

Lemma 63. The function $\varphi$ is continuous.

Proof. Let $x \in X$, and let $\varepsilon > 0$. We shall prove that there is a neighborhood $X_0$ of $x$ in $X$ such that $|\varphi(\xi) - \varphi(x)| \leq \varepsilon$ for $\xi \in X_0$. For $p \in D$ there is a neighborhood $X_p$ of $x$ in $X$ and a neighborhood $D_p$ of $p$ in $D$ such that

$$|\Phi(\xi, \eta) - \varphi(x)| \leq \varepsilon,$$

for $(\xi, \eta) \in X_p \times D_p$. Since $D$ is compact, the cover $\{D_p \mid p \in D\}$ of $D$ contains a finite sub cover. In other words, there are $p_1, \ldots, p_n \in D$ such that $D = D_{p_1} \cup \cdots \cup D_{p_n}$. Let $X_0 = X_{p_1} \cap \cdots \cap X_{p_n}$, then $X_0$ is a neighborhood of $x$ in $X$ such that

$$|\Phi(\xi, \eta) - \varphi(x)| \leq \varepsilon,$$

for $\xi \in X_0$ and all $\eta \in D$. It follows that $|\varphi(\xi) - \varphi(x)| \leq \varepsilon$ for $\xi \in X_0$. In other words, $\varphi$ is continuous in $x$. $\square$

In general, for given $x \in X$ the value of $p$ for which $\Phi(x, p)$ attains its maximum is not unique. However, if it is unique it depends continuously on $x$. 
Appendix A. Convex functions and differentiability

Lemma 64. If for each $x$ in $X$ there is a unique value $\lambda(x) \in D$ such that

$$\varphi(x) = \Phi(x, \lambda(x)),$$

then the function $\lambda : X \to D$ is continuous.

Proof. Let $x \in X$, and consider a sequence $\{x_n\}$ of points in $X$, converging to $x$. We shall prove that the sequence $\{\lambda(x_n)\}$ has $\lambda(x)$ as its unique limit point, which allows us to conclude that $\lambda$ is continuous at $x$.

Let $\eta \in D$ be a limit point of $\{\lambda(x_n)\}$. Such a limit point exists, since $D$ is compact. Passing to subsequences if necessary we may assume that the sequence converges to $\eta$. Since $\Phi$ is continuous by assumption, and $\varphi$ is continuous according to Lemma 63, taking limits in

$$\varphi(x_n) = \Phi(x_n, \lambda(x_n))$$

we obtain the identity

$$\varphi(x) = \Phi(x, \eta).$$

Since the point at which $\Phi(x, \cdot)$ attains its maximum is unique, we conclude that $\eta = \lambda(x)$. In other words, $\lambda(x)$ is the unique limit point of the sequence $\{\lambda(x_n)\}$. \qed

A.2 Gateaux- and Fréchet-differentiability

Consider a function $f : \mathbb{R}^d \to \mathbb{R}$. The one-sided directional derivative of $f$ at $x \in \mathbb{R}^d$ in the direction $v$, $v \in \mathbb{R}^d$, is

$$f'(x; v) = \lim_{h \to 0} \frac{f(x + hv) - f(x)}{h},$$

provided this limit exists.

The function $f$ is called Gateaux differentiable at $x$ if the two-sided directional derivative in the direction $v$

$$\lim_{h \to 0} \frac{f(x + hv) - f(x)}{h}$$

exists for all $v \in \mathbb{R}^d$. In this case $f'(x; v) = -f'(x; -v)$. The function $f$ is called Fréchet differentiable at $x \in \mathbb{R}^d$ if there is a linear function $T_x : \mathbb{R}^d \to \mathbb{R}$ such that

$$f(x + v) = f(x) + T_x(v) + \|v\|E(v),$$

for all $v$ is some neighborhood of $x$, where $E : \mathbb{R}^d \to \mathbb{R}$ is a function such that $E(v) \to 0$ if $\|v\| \to 0$. The linear transformation $T_x$ is called the total derivative of $f$ at $x$, and is denoted by $f'(x)$. If $f$ is Fréchet-differentiable at $x$, then $f$ is Gateaux-differentiable at $x$ in all directions $v$, and $f'(x; v) = f'(x)(v)$ for all $v \in \mathbb{R}^d$. Conversely, if $f$ is Gateaux-differentiable at $x$, and all its partial derivatives are continuous at $x$, then $f$ is Fréchet-differentiable at $x$. 

A.3. Convex functions

Note that this last property does not necessarily hold if the partial derivatives are not continuous. To see this, consider the function $f: \mathbb{R}^2 \to \mathbb{R}$, defined by

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Then $f$ is Gateaux-differentiable at $x_0 = (0,0)$. In fact, if $v = (a, b)$ then $f'(x_0; v) = \frac{b^2}{a}$, if $a \neq 0$, and $f'(x_0; v) = 0$, if $a = 0$. On the other hand $f$ is not even continuous at $(0,0)$, since $f(y^2, y) = \frac{1}{2}$, if $y \neq 0$, and $f(0,0) = 0$. In particular, $f$ is not Fréchet-differentiable at $(0,0)$.

A.3 Convex functions

We continue the discussion of Appendix A.1, but now we assume that $X = \mathbb{R}^d$ and $D$ is a compact subset of $\mathbb{R}^n$. First we recall some results and terminology from Convex Analysis.

A set $A$ is convex if for any two points $x, x' \in A$ the line segment $xx'$ lies in $A$ and $A$ is called strictly convex if the open line segment $(x, x')$ lies in the interior of $A$ for any two points $x, x' \in A$.

The set of convex functions on a compact subset of a Euclidean space is closed.

**Lemma 65.** If a sequence of convex continuous functions defined on a compact subset $D$ of $\mathbb{R}^d$ is convergent, then the limit function is convex on $D$.

The proof is straightforward. Note that the result is not true if we restrict to strictly convex functions.

The *epigraph* of a function $f: X \to \mathbb{R}$ is the set of points in $X \times \mathbb{R}$ above the graph of $f$:

$$\text{epi}(f) = \{(x, z) | x \in X, z \in \mathbb{R} \text{ and } f(x) \leq z\}.$$ 

A function is (strictly) convex if and only if its epigraph is a (strictly) convex set.

The following criterion for strict convexity is useful in the context of piecewise smooth functions.

**Lemma 66.** Let $f$ be a convex continuous function defined on a closed interval $I$ of the real line. If $f$ is strictly convex on each connected component of the complement of a finite set of points in $I$, then $f$ is strictly convex on $I$.

**Proof.** Let $S$ be the finite set of points such that $f$ is strictly convex on each connected component of $I \setminus S$. Consider a point $\xi \in S$, and a neighborhood $U$ of $\xi$, in $I$ that contains no other points of $S$. It is sufficient to prove that $f$ is strictly convex on $U$, since a function is strictly convex on $I$ if every point of $I$ has a neighborhood on which $f$ is strictly convex.

Let $x_0, x_1 \in U$ with $x_0 < x_1$, and let $x_t = (1-t)x_0 + tx_1$ for $0 < t < 1$. Since $f$ is convex on $[x_0, x_1]$, we know that

$$f(x_t) \leq (1-t)f(x_0) + tf(x_1),$$

which is the desired inequality.

Note that this last property does not necessarily hold if the partial derivatives are not continuous. To see this, consider the function $f: \mathbb{R}^2 \to \mathbb{R}$, defined by

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Then $f$ is Gateaux-differentiable at $x_0 = (0,0)$. In fact, if $v = (a, b)$ then $f'(x_0; v) = \frac{b^2}{a}$, if $a \neq 0$, and $f'(x_0; v) = 0$, if $a = 0$. On the other hand $f$ is not even continuous at $(0,0)$, since $f(y^2, y) = \frac{1}{2}$, if $y \neq 0$, and $f(0,0) = 0$. In particular, $f$ is not Fréchet-differentiable at $(0,0)$.

A.3 Convex functions

We continue the discussion of Appendix A.1, but now we assume that $X = \mathbb{R}^d$ and $D$ is a compact subset of $\mathbb{R}^n$. First we recall some results and terminology from Convex Analysis.

A set $A$ is convex if for any two points $x, x' \in A$ the line segment $xx'$ lies in $A$ and $A$ is called strictly convex if the open line segment $(x, x')$ lies in the interior of $A$ for any two points $x, x' \in A$.

The set of convex functions on a compact subset of a Euclidean space is closed.

**Lemma 65.** If a sequence of convex continuous functions defined on a compact subset $D$ of $\mathbb{R}^d$ is convergent, then the limit function is convex on $D$.

The proof is straightforward. Note that the result is not true if we restrict to strictly convex functions.

The *epigraph* of a function $f: X \to \mathbb{R}$ is the set of points in $X \times \mathbb{R}$ above the graph of $f$:

$$\text{epi}(f) = \{(x, z) | x \in X, z \in \mathbb{R} \text{ and } f(x) \leq z\}.$$ 

A function is (strictly) convex if and only if its epigraph is a (strictly) convex set.

The following criterion for strict convexity is useful in the context of piecewise smooth functions.

**Lemma 66.** Let $f$ be a convex continuous function defined on a closed interval $I$ of the real line. If $f$ is strictly convex on each connected component of the complement of a finite set of points in $I$, then $f$ is strictly convex on $I$.

**Proof.** Let $S$ be the finite set of points such that $f$ is strictly convex on each connected component of $I \setminus S$. Consider a point $\xi \in S$, and a neighborhood $U$ of $\xi$, in $I$ that contains no other points of $S$. It is sufficient to prove that $f$ is strictly convex on $U$, since a function is strictly convex on $I$ if every point of $I$ has a neighborhood on which $f$ is strictly convex.

Let $x_0, x_1 \in U$ with $x_0 < x_1$, and let $x_t = (1-t)x_0 + tx_1$ for $0 < t < 1$. Since $f$ is convex on $[x_0, x_1]$, we know that

$$f(x_t) \leq (1-t)f(x_0) + tf(x_1),$$

which is the desired inequality.
for $0 < t < 1$. If there is a $t$, with $0 < t < 1$, for which equality holds, then equality holds for all $t$. In that case $f$ is an affine function on $[x_0, x_1]$. This contradicts the fact that $f$ is strictly convex on $(x_0, \xi)$ and on $(\xi, x_1)$. Hence the inequality is strict for all $t$ with $0 < t < 1$. In other words, $f$ is strictly convex on $U$. □

**Lemma 67.** If $f$ is $C^2$ on a convex domain $D$, then $f$ is (strictly) convex iff the Hessian of $f$ is nonnegative definite (positive definite).

*Proof.* If the Hessian of $f$ is nonnegative definite then the function is locally strictly convex. A function that is everywhere locally strictly convex, is strictly convex. □

We present a simple situation in which the maximum-function $\varphi$, introduced in Section A.1, is convex.

**Lemma 68.** If $\Phi(\cdot, p) : \mathbb{R}^d \to \mathbb{R}$ is convex, for all $p \in D$, then the function $\varphi : \mathbb{R}^d \to \mathbb{R}$ is convex.

*Proof.* Note that $(x, z) \in \text{epi}(\varphi)$ if and only if $\varphi(x) \leq z$, i.e., if and only if $\Phi(x, p) \leq z$ for all $p \in D$. Therefore:

$$\text{epi}(\varphi) = \cap_{p \in D} \text{epi}(\Phi(\cdot, p)).$$

The right hand side is an intersection of convex sets, so it is a convex set. □

**Lemma 69.** A convex function on $\mathbb{R}^d$ has a (one sided) directional derivative at all points in all directions. Furthermore,

$$-f'(x; -v) \leq f'(x; v),$$

for all $x, v \in \mathbb{R}^d$. 

*Proof.* Consider a monotonically decreasing sequence $h_n$ of positive numbers tending to 0. Since $f$ is convex, the sequence

$$\frac{f(x + h_n v) - f(x)}{h_n}$$

is monotonically decreasing, so it has a limit. In other words, $f'(x; v)$ exists for all $x, v \in \mathbb{R}^d$. Now let $h_-$ and $h_+$ be arbitrary positive numbers. Convexity of $f$ implies that

$$\frac{f(x) - f(x - h_- v)}{h_-} \leq \frac{f(x + h_+ v) - f(x)}{h_+}.$$

Taking limits for $h_- \downarrow 0$ and $h_+ \downarrow 0$ yields $-f'(x; -v) \leq f'(x; v)$. □