13.1 Talking convex analysis

In convex analysis, convex and concave functions are defined everywhere on a vector space $X$, and are allowed to assume the extended values $\infty$ (sometimes denoted $+\infty$) and $-\infty$.

(Actually, convex analysts talk mostly about convex functions, and only occasionally about concave functions.) Recall that $\mathbb{R}^\sharp$ denotes the extended real numbers, $\mathbb{R}^\sharp = \mathbb{R} \cup \{\infty, -\infty\}$. Given an extended real-valued function $f : X \to \mathbb{R}^\sharp$, recall that the hypograph of $f$ is the subset of $X \times \mathbb{R}$ defined by

$$\text{hypo } f = \{(x, \alpha) \in X \times \mathbb{R} : \alpha \leq f(x)\}.$$  

The epigraph is defined by reversing the inequality

$$\text{epi } f = \{(x, \alpha) \in X \times \mathbb{R} : \alpha \geq f(x)\}.$$  

Note well that the hypograph or epigraph of $f$ is a subset of $X \times \mathbb{R}$, not of $X \times \mathbb{R}^\sharp$. That is, each $\alpha$ in the definition of hypograph or epigraph is a real number, not an infinity. For example, the epigraph of the constant function $\infty$ is the empty set.

To a convex analyst, an extended real-valued function is convex if its epigraph is a convex subset of $X \times \mathbb{R}$. Given a convex function $f : X \to \mathbb{R}^\sharp$, its effective domain is

$$\text{dom } f = \{x \in X : f(x) < \infty\}.$$  

We can extend a conventional real-valued convex function $f$ defined on a subset $C$ of a vector space $X$ to an extended real-valued function $\tilde{f}$ defined on all of $X$ by setting

$$\tilde{f}(x) = \begin{cases} f(x) & x \in C \\ \infty & x \notin C. \end{cases}$$  

Note that $\tilde{f}$ is convex in the convex analyst’s sense if and only if $f$ is convex in the conventional sense, in which case we also have that $\text{dom } \tilde{f} = C$.

We can similarly extend conventionally defined concave functions to $X$ by setting them equal to $-\infty$ where conventionally undefined.
Proper functions

In the language of convex analysis, a convex function is proper if its effective domain is nonempty and its epigraph contains no vertical lines. A concave function is proper if its effective domain is nonempty and its hypograph contains no vertical lines. (A vertical line in $X \times \mathbb{R}$ is a set of the form $\{x\} \times \mathbb{R}$ for some $x \in X$.)

That is, a convex $f$ is proper if $f(x) < \infty$ for at least one $x$ and $f(x) > -\infty$ for every $x$. Every proper convex function is gotten by taking a finite-valued convex function defined on some nonempty convex set and extending it to all of $X$ as above. Clearly, a convex function $f$ is proper if and only if $-f$ is a proper concave function. Thus:

> Every convex function in the conventional sense is a proper convex function in the sense of convex analysis. Likewise for concave functions. For a proper function, its effective domain is the set of points where it is finite.

As an example of a nontrivial improper convex function, consider this one taken from Rockafellar [4, p. 24].

13.1.1 Example (A nontrivial improper convex function) The function $f: \mathbb{R} \to \mathbb{R}^\sharp$ defined by

$$
\begin{cases}
-\infty & |x| < 1 \\
0 & |x| = 1 \\
\infty & |x| > 1
\end{cases}
$$

is an improper convex function that is not constant.

Some authors, for example, Aubin [1] or Hiriart-Urruty and Lemaréchal [3], do not permit convex functions to assume the value $-\infty$, so for them, properness is equivalent to nonemptyness of the effective domain.

Indicator functions

Another deviation from conventional terminology is the convex analysts’ definition of the indicator function. The indicator function of the set $C$, denoted $\delta(\cdot \mid C)$, is defined by

$$
\delta(x \mid C) = \begin{cases}
0 & x \in C \\
\infty & x \notin C.
\end{cases}
$$

The indicator of $C$ is a convex function if and only if $C$ is a convex set. It is proper if and only if $C$ is nonempty. This is not to be confused with the probabilists’ indicator function $1_C$ defined by

$$
1_C(x) = \begin{cases}
1 & x \in C \\
0 & x \notin C.
\end{cases}
$$
We now turn your attention to a few simple facts.

- Affine functions are both concave and convex.
- If $f$ is continuous on $\text{dom } f$ and $\text{dom } f$ is closed in $X$, then $f$ is lower semicontinuous as an extended real-valued function on $X$.
- Note well that the epigraph of an extended-real valued function is a subset of $X \times \mathbb{R}$, not a subset of $X \times \mathbb{R}^\#$. As a result, for a convex function $f$,
  $$x \in \text{dom } f \iff (x, f(x)) \in \text{epi } f.$$  
In other words, the effective domain of $f$ is the projection on $X$ of its epigraph. For concave functions replace epigraph by hypograph.
- The effective domain of a convex or concave function is a convex set.
- The constant function $f = -\infty$ is convex (its epigraph is $X \times \mathbb{R}$), but not proper, and the constant function $g = \infty$ is also convex (its epigraph is the empty set, which is convex), but not proper. These functions are also concave.
- If a convex function is proper, then its epigraph is a nonempty proper subset of $X \times \mathbb{R}$. If a concave function is proper, then its hypograph is a nonempty proper subset of $X \times \mathbb{R}$.
- Let $f$ be an extended real-valued function on a tvs $X$. If $f$ finite at $x$ and continuous at $x$, then in fact $x$ belongs to the interior of the effective domain of $f$. Why?
- A convex function need not be finite at all points of continuity. The proper convex function $f$ on $\mathbb{R}$ defined by $f(x) = 1/x$ for $x > 0$, and $f(x) = \infty$ for $x \leq 0$ is continuous everywhere, even at zero.

13.1.2 Exercise (Rockafellar [4, Lemma 7.3, p. 54]) Prove the following.
For any concave function $f$ on $\mathbb{R}^m$,
$$\text{ri hypo } f = \{(x, \alpha) \in \mathbb{R}^m \times \mathbb{R} : x \in \text{ri dom } f, \alpha < f(x)\}.$$  
For a convex function $f$ the corresponding result is
$$\text{ri epi } f = \{(x, \alpha) \in \mathbb{R}^m \times \mathbb{R} : x \in \text{ri dom } f, \alpha > f(x)\}.$$  
$\square$
13.2 Hyperplanes in $X \times \mathbb{R}$ and affine functions on $X$

I will refer to a typical element in $X \times \mathbb{R}$ as a point $(x, \alpha)$ where $x \in X$ and $\alpha \in \mathbb{R}$. I may call $x$ the “vector component” and $\alpha$ the “real component,” even when $X = \mathbb{R}$. A hyperplane in $X \times \mathbb{R}$ is defined in terms of its “normal vector” $(p, \lambda)$, which belongs to the dual space $(X \times \mathbb{R})^* \equiv X^* \times \mathbb{R}$. (That is, every (continuous) linear functional $\ell$ on $X \times \mathbb{R}$ is of the form $\ell(x, \alpha) = p(x) + \lambda \alpha$, where $p$ is a (continuous) linear functional on $X$. When $X$ is a Euclidean space, then $p(x) = p \cdot x$ and $(p, \lambda)$ is indeed the normal vector to the hyperplane.) If the real component $\lambda = 0$, we say the hyperplane is vertical. If the hyperplane is not vertical, by homogeneity we can arrange for $\lambda$ to be $-1$ (you will see why in just a moment). Here is an obvious fact about vertical hyperplanes that is just begging for a name:

13.2.1 Verticality  If $H$ is a vertical hyperplane in $X \times \mathbb{R}$ and if $(x, \alpha) \in H$, then for every $\beta \in \mathbb{R}$, we have $(x, \beta) \in H$. Consequently a vertical hyperplane can never properly separate $(x, \alpha)$ and $(x, \beta)$.

13.2.2 Proposition  Non-vertical hyperplanes in $X \times \mathbb{R}$ are precisely the graphs of affine functions on $X$. That is,

$$\text{gr}(x \mapsto p(x) - \beta) \text{ is the non-vertical hyperplane } \{ (x, \alpha) \in X \times \mathbb{R} : (p, -1) \cdot (x, \alpha) = \beta \}. $$

And the non-vertical hyperplane

$$\{ (x, \alpha) \in X \times \mathbb{R} : (p, \lambda) \cdot (x, \alpha) = \beta \} \text{ is gr}(x \mapsto (-1/\lambda)p(x) + \beta/\lambda). $$

See Figure 13.2.1.

13.3 Lower semicontinuous convex functions

Recall that an extended real-valued function on a topological space $X$ is lower semicontinuous if and only if its epigraph is closed (Theorem 13.4.3 in the Appendix). Recall that a locally convex space is a tvs where each neighborhood of a point includes a convex neighborhood of the point. It is for these spaces that the strong separating hyperplane theorem for continuous linear functionals (topologically closed half-spaces) holds.

Taking a page from Barry Simon’s book [5], let’s make the following definition.

13.3.1 Definition (Regular convex functions)  A regular convex function on a topological vector space is a lower semicontinuous proper convex function.

A regular concave function on a topological vector space is an upper semicontinuous proper concave function.
Figure 13.2.1. The graph of $g: y \mapsto p \cdot y - \beta$ is the hyperplane $H = \{(x, \alpha) \in X \times \mathbb{R} : (p, -1) \cdot (x, \alpha) = \beta\}$. 
It is the class of regular convex functions that are most useful and the nicest to work with.

13.3.2 Lemma Let $X$ be a locally convex Hausdorff space (such as $\mathbb{R}^m$), and let $f : X \to \mathbb{R}^\sharp$ be a regular convex function. If $x$ belongs to the effective domain of $f$ and $\alpha \in \mathbb{R}$ satisfies $\alpha < f(x)$, then there exists a continuous affine function $g$ satisfying

$$g(x) = \alpha \quad \text{and} \quad g \ll f,$$

where $g \ll f$ means $g(y) < f(y)$ for all $y \in X$.

Proof: First note that the epigraph of $f$ is a nonempty closed convex subset of $X \times \mathbb{R}$, and by hypothesis $(x, \alpha)$ does not belong to $\text{epi} \ f$. Thus by the Strong Separating Hyperplane Theorem 8.3.2 there is a closed hyperplane $H$ that strongly separates $(x, \alpha)$ from $\text{epi} \ f$. The hyperplane $H$ cannot be vertical, for by Verticality 13.2.1 a vertical hyperplane cannot properly separate $(x, \alpha)$ from $(x, f(x))$. Therefore $H$ is the graph of a continuous affine function $g$ (Proposition 13.2.2). Moreover, $g \ll f$, for if $g(y) \geq f(y)$, then $(y, g(y)) \in \text{gr} \ g \cap \text{epi} \ f$, which is ruled out by strong separation. See Figure 13.3.1.

\[\text{Figure 13.3.1. The linear function } (p, \lambda) \text{ (or, equivalently, the non-vertical hyperplane } H = \text{gr} \ g) \text{ strongly separates } (x, \alpha) \text{ from } \text{epi} \ f.\]

13.3.3 Theorem Let $X$ be a locally convex Hausdorff space, and let $f : X \to \mathbb{R}^\sharp$ be a regular convex function. Then for each $x$ we have

$$f(x) = \sup \{g(x) : g \ll f \text{ and } g \text{ is affine and continuous}\}.$$
Proof: Fix \( x \) and let \( \alpha \in \mathbb{R} \) satisfy \( \alpha < f(x) \). (Since \( f \) is proper, we cannot have \( f(x) = -\infty \), so such a finite real number \( \alpha \) exists.) We need to show that there is a continuous affine function \( g \) with \( g \ll f \) and \( g(x) \geq \alpha \).

There are two cases to consider. The first is that \( x \) belongs to the effective domain of \( f \). This is covered by Lemma 13.3.2 directly.

In case \( x \) is not in the effective domain, we may still proceed as in the proof of Lemma 13.3.2 to show that there exists a hyperplane \( H \) defined by a nonzero continuous linear functional \( (p, \lambda) \) that strongly separates \((x, \alpha)\) from \( \text{epi} f \). But we cannot use the previous argument to conclude that the hyperplane is non-vertical! So suppose that \( \lambda = 0 \). Then strong separation can be written as

\[ p(y) < p(x) - \varepsilon \text{ for every } y \in \text{dom } f \]

for some \( \varepsilon > 0 \). See Figure 13.3.2. Define the affine function \( h \) by

\[ h(z) = p(z) - p(x) + \varepsilon/2 \]

and observe that

\[ h(x) = \varepsilon/2 > 0 \quad \text{and for } y \in \text{dom } f, \quad h(y) < -\varepsilon/2 < 0. \]

Next pick some \( \bar{y} \in \text{dom } f \), and use Lemma 13.3.2 to find an affine function \( \bar{g} \) satisfying

\[ \bar{g} \ll f \]

(and \( \bar{g}(\bar{y}) = f(\bar{y}) - 1 \), which is irrelevant for our purpose). Now consider the affine functions \( g_\gamma \) of the form

\[ g_\gamma(z) = \gamma h(z) + \bar{g}(z), \quad \gamma > 0. \]

For \( y \in \text{dom } f \) we have \( h(y) < 0 \) so \( g_\gamma(y) < \bar{g}(y) < f(y) \). For \( y \notin \text{dom } f \), we have \( f(y) = \infty \). Thus for any \( \gamma > 0 \) and any \( y \), we have

\[ g_\gamma(y) < f(y). \]

But \( h(x) > 0 \), so for \( \gamma \) large enough,

\[ g_\gamma(x) > \alpha, \]

as desired.

A remark is in order. We know that the epigraph of a regular convex function is a proper closed convex subset of \( X \times \mathbb{R} \). Therefore it is the intersection of all the closed half-spaces that include it. The theorem refines this to the intersection of all the closed non-vertical half spaces that include it.
Figure 13.3.2. The proof of Theorem 13.3.3: Here $x \notin \text{dom } f$, but $\bar{y} \in \text{dom } f$. The linear function $(p, 0)$ (or the vertical hyperplane $H$) strongly separates $(x, \alpha)$ from $\text{epi } f$ by $\varepsilon$. (That is, $p(y) < p(x) - \varepsilon$ for all $y \in \text{dom } f$.) The affine function $\bar{g}$ satisfies $\bar{g} \ll f$, and the affine function $h: y \mapsto p(y) - p(x) + \varepsilon/2$ satisfies $h(x) = \varepsilon/2 > 0$ and $h(y) < -\varepsilon/2 < 0$ for $y \in \text{dom } f$. The affine function $g_\gamma = \gamma h + \bar{g}$ satisfies $g_\gamma \ll f$ for $\gamma > 0$, and for $\gamma$ large enough $g_\gamma(x) > \alpha$. 

13.4 Appendix: Semicontinuous functions

The real-valued function \( f : X \to \mathbb{R} \) is upper semicontinuous on \( X \) if for each \( \alpha \in \mathbb{R} \), the superlevel set \( \{ f \geq \alpha \} \) is closed, or equivalently, the strict sublevel set \( \{ f < \alpha \} \) is open. It is lower semicontinuous if every sublevel set \( \{ f \leq \alpha \} \) is closed, or equivalently, the strict superlevel set \( \{ f > \alpha \} \) is open.

The extended real valued function \( f \) is upper semicontinuous at the point \( x \) if \( f(x) < \infty \) and

\[
(\forall \varepsilon > 0) \ (\exists \delta > 0) \ [d(y, x) < \delta \implies f(y) < f(x) + \varepsilon].
\]

Similarly, \( f \) is lower semicontinuous at the point \( x \) if \( f(x) > -\infty \) and

\[
(\forall \varepsilon > 0) \ (\exists \delta > 0) \ [d(y, x) < \delta \implies f(y) > f(x) - \varepsilon].
\]

Equivalently, \( f \) is upper semicontinuous at \( x \) if \( f(x) < \infty \) and

\[
f(x) \geq \limsup_{y \to x} f(y) = \inf_{\varepsilon > 0} \sup_{0 < d(y, x) < \varepsilon} f(y).
\]

Similarly, \( f \) is lower semicontinuous at \( x \) if \( f(x) > -\infty \) and

\[
f(x) \leq \liminf_{y \to x} f(y) = \sup_{\varepsilon > 0} \inf_{0 < d(y, x) < \varepsilon} f(y).
\]

Note that \( f \) is upper semicontinuous if and only if \(-f\) is lower semicontinuous.

13.4.1 Lemma A real valued function \( f : X \to \mathbb{R} \) is upper semicontinuous on \( X \) if and only if it is upper semicontinuous at each point of \( X \). It is lower semicontinuous on \( X \) if and only if it is lower semicontinuous at each point of \( X \).

Proof: I’ll prove the result for upper semicontinuity. Assume that \( f \) is upper semicontinuous on \( X \). For any real number \( \alpha \), if \( f(x) < \beta < \alpha \), then \( \{ y \in X : f(y) < \beta \} \) is an open neighborhood of \( x \). Thus for \( \varepsilon > 0 \) small enough \( d(y, x) < \varepsilon \) implies \( f(y) < \beta \). Therefore \( \limsup_{y \to x} f(y) \leq \beta < \alpha \). Setting \( \alpha = \limsup_{y \to x} f(y) \), we see that it cannot be the case that \( f(x) < \limsup_{y \to x} f(y) \), for then \( f(x) < \alpha = \limsup_{y \to x} f(y) < \alpha \), a contradiction. That is, \( f \) is upper semicontinuous at \( x \).

For the converse, assume that \( f \) is upper semicontinuous at each \( x \). Fix a real number \( \alpha \), and let \( f(x) < \alpha \). Since \( f(x) \geq \limsup_{y \to x} f(y) \), there is \( \varepsilon > 0 \) small enough so that \( \sup_{0 < d(y, x) < \varepsilon} f(y) < \alpha \), but this implies \( \{ x \in X : f(x) < \alpha \} \) is open, so \( f \) is upper semicontinuous on \( X \).

13.4.2 Corollary A real-valued function is continuous if and only if it is both upper and lower semicontinuous.

13.4.3 Theorem An extended real-valued function \( f \) is upper semicontinuous on \( X \) if and only if its hypograph is closed. It is lower semicontinuous on \( X \) if and only if its epigraph is closed.
Proof: Assume \( f \) is upper semicontinuous, and let \((x_n, \alpha_n)\) be a sequence in its hypograph, that is, \( f(x_n) \geq \alpha_n \) for all \( n \). Therefore \( \limsup_n f(x_n) \geq \limsup_n \alpha_n \). If \((x_n, \alpha_n) \to (x, \alpha)\), since \( f \) is upper semicontinuous at \( x \), we have \( \alpha = \lim_n \alpha_n \leq \limsup_n f(x_n) \leq f(x) \). Thus \( \alpha \leq f(x) \), or \((x, \alpha)\) belong to the hypograph of \( f \). Therefore the hypograph is closed.

Assume now that the hypograph is closed. Pick \( x \) and let \( \alpha = \limsup_{y \to x} f(y) \). Then there is a sequence \( x_n \to x \) with \( f(x_n) \uparrow \alpha \). Since \((x_n, f(x_n))\) belongs to the hypograph for each \( n \), so does its limit \((x, \alpha)\). That is, \( \limsup_{y \to x} f(y) = \alpha \leq f(x) \), so \( f \) is upper semicontinuous at \( x \).

13.4.4 Exercise Prove that if both the epigraph and hypograph of a function are closed, then the graph is closed. Give an example to show that the converse is not true. □

13.4.5 Proposition The infimum of a family of upper semicontinuous functions is upper semicontinuous. The supremum of a family of lower semicontinuous functions is lower semicontinuous.

Proof: Let \( \{f_\nu\}_\nu \) be a family of upper semicontinuous functions, and let \( f_x = \inf_\nu f_\nu(x) \). Then \( \{f \geq \alpha\} = \bigcap_\nu \{f_\nu \geq \alpha\} \), which is closed. Lower semicontinuity is dealt with mutatis mutandis. □

13.4.6 Definition Given an extended real-valued function \( f \) on the metric space \( X \), we define the upper envelope \( \overline{f} \) of \( f \) by

\[
\overline{f}(x) = \max\{f(x), \limsup_{y \to x} f(y)\} = \inf_{\varepsilon > 0} \sup_{d(y, x) < \varepsilon} f(y),
\]

and the lower envelope \( \underline{f} \) of \( f \) by

\[
\underline{f}(x) = \min\{f(y), \liminf_{y \to x} f(y)\} = \sup_{\varepsilon > 0} \inf_{d(y, x) < \varepsilon} f(y).
\]

Clearly if \( f \) is upper semicontinuous at \( x \), then \( f(x) = \overline{f}(x) \), and if \( f \) is lower semicontinuous at \( x \), then \( f(x) = \underline{f}(x) \). Consequently, \( f \) is upper semicontinuous if and only \( f = \overline{f} \), and \( f \) is lower semicontinuous if and only \( f = \underline{f} \).

We say that the real-valued function \( g \) dominates the real-valued function \( f \) on \( X \) if for every \( x \in X \) we have \( g(x) \geq f(x) \).

13.4.7 Theorem The upper envelope \( \overline{f} \) is the smallest upper semicontinuous function that dominates \( f \) and the lower envelope \( \underline{f} \) is the greatest lower semicontinuous function that \( f \) dominates.

Moreover,

\[
\operatorname{hypo} \overline{f} = \overline{\operatorname{hypo} f},
\]

and

\[
\operatorname{epi} \underline{f} = \overline{\operatorname{epi} f}.
\]
Proof: Clearly, $\overline{f}$ dominates $f$ and $f$ dominates $\overline{f}$.

Now suppose $g$ is upper semicontinuous and dominates $f$. Then for any $x$, we have $g(x) \geq \lim \sup_{y \to x} g(y) \geq \lim \sup_{y \to x} f(y)$, so $g(x) \geq \overline{f}(x)$. That is, $g$ dominates $\overline{f}$.

Similarly if $g$ is lower semicontinuous and $f$ dominates $g$, then $\overline{f}$ dominates $g$.

It remains to show that $\overline{f}$ is upper semicontinuous. It suffices to prove that the hypograph of $\overline{f}$ is closed. We prove the stronger result that $\text{hypo} \overline{f} = \overline{\text{hypo} f}$.

Let $(x_n, \alpha_n)$ be a sequence in the hypograph of $f$, and assume it converges to a point $(x, \alpha)$. Since $\alpha_n \leq f(x_n)$, we must have $\alpha \leq \lim \sup_{y \to x} f(y)$, so $\alpha \leq \overline{f}$. That is, $\overline{\text{hypo} f} \subset \text{hypo} \overline{f}$. For the opposite inclusion, suppose by way of contradiction that $(x, \alpha)$ belongs to the hypograph of $\overline{f}$, but not to $\overline{\text{hypo} f}$. Then there is a neighborhood $B_\varepsilon(x) \times B_\varepsilon(\alpha)$ disjoint from $\text{hypo} \overline{f}$. In particular, if $d(y, x) < \varepsilon$, then $f(y) < \alpha \leq \overline{f}(x)$, which implies $\overline{f}(x) > \max\{f(x), \lim \sup_{-y \to x} f(y)\}$, a contradiction. Therefore $\overline{\text{hypo} f} \supset \text{hypo} \overline{f}$.

The case of $\underline{f}$ is similar.

13.5 ★ Appendix: Closed functions revisited

Rockafellar [4, p 52, pp. 307–308] makes the following definition. Recall from Definition 13.4.6 that the upper envelope of $f$ is defined by

$$\overline{f}(x) = \inf_{\varepsilon > 0} \sup_{d(y, x) < \varepsilon} f(y),$$

and that the upper envelope is real-valued if $f$ is locally bounded, and is upper semicontinuous.

13.5.1 Definition The closure $\text{cl} f$ of a convex function $f$ on $\mathbb{R}^m$ is defined by

1. $\text{cl} f(x) = -\infty$ for all $x \in \mathbb{R}^m$ if $f(y) = -\infty$ for some $y$.
2. $\text{cl} f(x) = +\infty$ for all $x \in \mathbb{R}^m$ if $f(x) = +\infty$ for all $x$.
3. $\text{cl} f$ is the lower envelope of $f$ if $f$ is a proper convex function.

The closure $\text{cl} f$ of a concave function $f$ on $\mathbb{R}^m$ is defined by

1’. $\text{cl} f(x) = +\infty$ for all $x \in \mathbb{R}^m$ if $f(y) = +\infty$ for some $y$.
2’. $\text{cl} f(x) = -\infty$ for all $x \in \mathbb{R}^m$ if $f(x) = -\infty$ for all $x$.
3’. $\text{cl} f$ is the upper envelope of $f$ if $f$ is a proper concave function.

13.5.2 Definition (Closed functions à la Rockafellar) A convex (or concave) function is closed if and only if $f = \text{cl} f$. 
13.5.3 Proposition If \( f : \mathbb{R}^m \to \mathbb{R}^\# \) is convex, then \( \text{cl} \, f \) is convex. If \( f \) is concave, then \( \text{cl} \, f \) is concave.

**Proof:** I shall just prove the concave case. If \( f \) is concave and does not assume the value \(+\infty\), Theorem 13.4.7 asserts that the hypograph of the closure of \( f \) is the closure of the hypograph of \( f \), which is convex. If \( f \) does assume the value \(+\infty\), then \( \text{cl} \, f \) is identically \(+\infty\), so its hypograph is \( \mathbb{R}^m \times \mathbb{R} \), which is convex. Either way, the hypograph of \( \text{cl} \, f \) is convex. \( \square \)

The next result is that my Definition 21.1.3 and Rockafellar’s Definition 13.5.1 agree.

13.5.4 Theorem Let \( f : \mathbb{R}^m \to \mathbb{R}^\# \) be concave. Then for every \( x \in \mathbb{R}^m \),

\[
\text{cl} \, f(x) = \inf \{ h(x) : h \geq f \text{ and } h \text{ is affine and continuous} \} = \hat{f}.
\]

If \( f : \mathbb{R}^m \to \mathbb{R}^\# \) is convex, then for every \( x \in \mathbb{R}^m \),

\[
\text{cl} \, f(x) = \sup \{ h(x) : f \geq h \text{ and } h \text{ is affine and continuous} \} = \check{f}.
\]

**Proof:** I shall prove the concave case. There are three subcases. If \( f \) is improper and assumes the value \(+\infty\), then by definition \( \text{cl} \, f \) is the constant function \(+\infty\). In this case, no affine function, which is (finite) real-valued, dominates \( f \) so the infimum is over the empty set, and thus \(+\infty\). The second subcase is that \( f \) is the improper constant function \(-\infty\). In this case every affine function dominates \( f \), so the infimum is \(-\infty\).

So assume we are in the third subcase, namely that \( f \) is proper. That is, \( f(x) < \infty \) for all \( x \in \mathbb{R}^m \), and \( \text{dom} \, f \) is nonempty. Then by definition \( \text{cl} \, f \) is the upper envelope of \( f \). That is,

\[
\text{cl} \, f(x) = \inf_{\varepsilon > 0} \sup_{d(y,x)<\varepsilon} f(y).
\]

Define \( g(x) = \inf \{ h(x) : h \geq f \text{ and } h \text{ is affine and continuous} \} \). If \( h \) is affine, continuous, and dominates \( f \), then by Theorem 13.4.7, \( h \) dominates \( \text{cl} \, f \), so \( g \) dominates \( \text{cl} \, f \).

We now show that \( \text{cl} \, f \geq g \). It suffices to show that for any \((x, \alpha)\) with \( \alpha > \text{cl} \, f(x) \), there is an affine function \( h \) dominating \( f \) with \( h(x) \leq \alpha \). Now \( \alpha > \text{cl} \, f(x) = \lim \sup_{y \to x} f(y) \) implies that \((x, \alpha)\) does not belong to the closure of the hypograph of \( f \).

There are two cases to consider. The simpler case is that \( x \) belongs to \( \text{dom} \, f \). So assume now that \( x \in \text{dom} \, f \). Since \( f \) is concave, its hypograph and the closure thereof are convex, and since \( f \) is proper, its hypograph is nonempty. So by Corollary 8.3.2 there is a nonzero \((p, \lambda) \in \mathbb{R}^m \times \mathbb{R} \) strongly separating \((x, \alpha)\) from the closure of the hypograph of \( f \). In particular, for each \( y \in \text{dom} \, f \) the point
\((y, f(y))\) belongs to the hypograph of \(f\). Thus strong separation implies that for some \(\varepsilon > 0\), for any \(y \in \text{dom}\ f\),

\[
p \cdot x + \lambda \alpha > p \cdot y + \lambda f(y) + \varepsilon.
\]

(1)

The same argument as that in the proof of Lemma 13.3.2 shows that \(\lambda \geq 0\). Moreover, taking \(y = x\) (since \(x \in \text{dom}\ f\)) shows that \(\lambda \neq 0\). So dividing by \(\lambda\) gives

\[
(1/\lambda)p \cdot (x - y) + \alpha > f(y) + (\varepsilon/\lambda)
\]

for all \(y \in \text{dom}\ f\). Define

\[
h(y) = (1/\lambda)p \cdot (x - y) + \alpha.
\]

Then \(h\) is a continuous affine function satisfying

\[
h(y) > f(y) + \eta \quad \text{for all} \ y \in \text{dom}\ f;
\]

where \(\eta = (\varepsilon/\lambda) > 0\) and \(h(x) = \alpha\), as desired.

The case where \((x, \alpha)\) satisfies \(\alpha > \text{cl}\ f(x)\), but \(x \notin \text{dom} f\) is more subtle. The reason the above argument does not work is that the hyperplane may be vertical (\(\lambda = 0\)), and hence not the graph of any affine function. So assume that \(\lambda = 0\). Then (1) becomes

\[
p \cdot x > p \cdot y + \varepsilon
\]

for all \(y \in \text{dom}\ f\). Define the continuous affine function \(g\) by

\[
g(y) = p \cdot (x - y) - \varepsilon/2,
\]

and note that \(g(x) < 0\), and \(g(y) > 0\) for all \(y \in \text{dom}\ f\).

But we still have (from the above argument) a continuous affine function \(h\) satisfying

\[
h(y) > f(y) \quad \text{for all} \ y \in \text{dom}\ f.
\]

Now for any \(\gamma > 0\), we have

\[
\gamma g(y) + h(y) > f(y) \quad \text{for all} \ y \in \text{dom}\ f,
\]

and for \(y \notin \text{dom} f\), \(f(y) = -\infty\), so the inequality holds for all \(y\) in \(\mathbb{R}^m\). But since \(g(x) < 0\), for \(\gamma\) large enough, \(\gamma g(x) + h(x) < \alpha\), so this is the affine function we wanted.

I think that covers all the bases (and cases).

The case of a convex function is dealt with by replacing the epigraph with the hypograph and reversing inequalities.
References


