Minimization of convex functions

In this chapter, we shall apply weak topologies to minimization of convex functions. The main general result in which convexity plays an important role is Theorem 0.5(c).

Let us recall that, given a Hausdorff topological space T, a function $f\colon T \to (-\infty, +\infty]$ is lower semicontinuous (l.s.c.) if, for each $t_0 \in T$ and each $\alpha < f(t_0)$ there exists a neighborhood U of t_0 such that $\alpha < f(t)$ for each $t \in U$. It is an easy exercise to show that *the following assertions are equivalent:*

- *f is l.s.c.;*
- *for each $\alpha \in \mathbb{R}$, the set $\{f \le \alpha\}$ is closed in T;*
- *the epigraph* $\mathrm{epi}(f)$ *is closed in $T \times \mathbb{R}$.*

Moreover, a pointwise supremum of any family of l.s.c. functions is again a l.s.c. function (this follows easily from the above epigraph ciriterion), and a sum of finitely many l.s.c. functions is again a l.s.c. function (this follows easily from the definition, using the fact that if $t < f(x_0) + g(x_0)$, there exist $t_1 < f(x_0)$ and $t_2 < g(x_0)$ such that $t = t_1 + t_2$).

The following theorem is the “lower part” of the well-known Weierstrass theorem.

Theorem 0.1. *Let $f\colon T \to (-\infty, +\infty]$ be a l.s.c. function on a compact Hausdorff topological space T. Then T attains its infimum over T.*

Proof. Let $\alpha_0 = \inf f(T)$. If $\alpha_0 = +\infty$, then $f \equiv +\infty$ and the assertion trivially holds. Let $\alpha_0 < +\infty$. Then, for each real $\alpha > \alpha_0$, the set $\{f \le \alpha\}$ is closed and nonempty. Obviously, any finite collection of such sets has a nonempty intersection. By compactness, also the set $\bigcap_{\alpha > \alpha_0} \{f \le \alpha\} = \{f \le \alpha_0\} = f^{-1}(\alpha_0)$ is nonempty. (In particular, this implies that α_0 is finite.) □

Theorem 0.2. *Let X be a normed space, and $f\colon X \to (-\infty, +\infty]$ a l.s.c. convex function. Then f is w-l.s.c. (In particular, the norm of X is w-l.s.c.)*

Proof. The sets $\{f \le \alpha\}$ are convex and closed, and hence w-closed. □

Corollary 0.3. *Let X be a normed space, $f\colon X \to (-\infty, +\infty]$ a l.s.c. convex function. Then f has a minimum on each nonempty w-compact subset of X. (In particular, if X is a reflexive Banach space, then f has a minimum on each bounded closed convex subset of X.)*

Lemma 0.4. *Let X be a normed space. Then the dual norm $\|\cdot\|_{X^*}$ is w^*-l.s.c.*

Proof. The sets $\{\|\cdot\|_{X^*} \le \alpha\}$ are closed balls, and hence w^*-closed. □

Now, we are ready to prove a general theorem about existence of global minima.

Theorem 0.5. *Let X be a Banach space, $f\colon X \to (-\infty, +\infty]$ a function. Assume that f is* coercive*, that is,*

$$\lim_{\|x\| \to +\infty} f(x) = +\infty\,.$$

Suppose that at least one of the following conditions is satisfied:

(a) *$X = Z^*$ for some normed space Z, and f is w^*-l.s.c. (that is, l.s.c. in the $\sigma(Z^*, Z)$-topology).*

(b) *X is a reflexive and f is w-l.s.c.;*
(c) *X is reflexive, and f is convex and (norm-)l.s.c.;*

Then f attains a global minimum.

Proof. Let $\alpha_0 = \inf f(X)$ and $\alpha \in (\alpha_0, +\infty)$. By coercivity, there exists $r > 0$ such that $f(x) > \alpha$ whenever $\|x\| > r$. It follows that $\alpha_0 = \inf f(rB_X)$. Now, the case (a) follows from Theorem 0.1 since the ball $rB_X = rB_{Z^*}$ is w^*-compact. The case (b) follows from (a) since we can take $Z = X^*$, and, in this case, Z is reflexive and $w_X^* = \sigma(Z^*, Z) = \sigma(Z^*, Z^{**}) = \sigma(X, X^*) = w_X$. Finally, (c) follows from (b) by Theorem 0.2. □

Let us state two applications in Abstract Approximation Theory.

Existence of nearest points. Let M be a metric space, $E \subset M$ a set and $x \in M$ a point. We say that $y \in M$ is a *nearest point* to x in M if $d(x, y) = d(x, M)$. The set E is called *proximinal* if each point of M has at least one nearest point in M.

Theorem 0.6.
(a) *Every nonempty w^*-closed set in a dual of a normed space is proximinal.*
(b) *Every nonempty w-closed set in a reflexive Banach space is proximinal.*
(c) *Every nonempty closed convex set in a reflexive Banach space is proximinal.*

Proof. As in the proof of Theorem 0.5, it suffices to prove the case (a). Let $E \subset X^*$ be a nonempty w^*-closed set, and $x_0^* \in X^* \setminus E$. The function $x^* \mapsto \|x^* - x_0^*\|$ is w^*-l.s.c. on X^* (Lemma 0.4) and coercive. Then also the function

$$f : X^* \to (-\infty, +\infty], \quad f(x^*) = \begin{cases} \|x^* - x_0^*\| & \text{if } x^* \in E \\ +\infty & \text{otherwise} \end{cases},$$

is w^*-l.s.c. and coercive. By Theorem 0.5, f attains a global minimum. Any such point of minimum is a nearest point to x_0^* in E. □

Using the James theorem, we can show that the property (c) in Theorem 0.6 characterizes reflexive spaces among all Banach spaces.

Theorem 0.7. *For a Banach space X, the following assertions are equivalent.*
(i) *X is reflexive.*
(ii) *Each nonempty closed convex set $C \subset X$ is proximinal.*
(iii) *Each closed hyperplane $H \subset X$ is proximinal.*

Proof. The implication $(i) \Rightarrow (ii)$ follows from Theorem 0.6, and $(ii) \Rightarrow (iii)$ is trivial. It remains to show that (iii) implies (i).

Assume (iii). Let $x^* \in X^*$ be such that $\|x^*\| = 1$, and consider the closed hyperplane $H = \{y \in X : x^*(y) = 1\}$. We claim that $d(0, H) = 1$. Indeed, $\sup x^*(rB_X) = r\|x^*\| = r$ for each $r > 0$. It follows easily that $rB_X \cap H = \emptyset$ whenever $0 < r < 1$, and $rB_X \cap H \neq \emptyset$ whenever $r > 1$; and this gives our claim. Since H is proximinal, there exists $y \in X$ such that $\|y\| = 1$ and $x^*(y) = 1$, that is, x^* attains its norm. It follows that each element of X^* attains its supremum over B_X. By the James theorem, B_X is w-compact, and hence X is reflexive. □

A normed space X is called *strictly convex* if its unit sphere $S_X = \{x \in X : \|x\| = 1\}$ contains no nontrivial line segment. We omit the simple proof of the following theorem about uniqueness of nearest points.

Theorem 0.8. *For a normed space X, the following assertions are equivalent.*

(i) *X is strictly convex.*
(ii) *For each convex set $C \subset X$, every point of X has at most one nearest point in C.*
(iii) *For each line $L \subset X$, every point of X has exactly one nearest point in L.*

A *Chebyshev set* is a set $E \subset X$, such that each point of X has a unique nearest point in E.

Corollary 0.9. *For a Banach space X, the following are equivalent:*

(i) *X is reflexive and strictly convex;*
(ii) *each nonempty closed convex set $C \subset X$ is a Chebyshev set.*

Lower semicontinuity of the distance function in weak topologies. For this application of We shall need the following general lemma.

Lemma 0.10. *Let (X, τ) be a t.v.s., $A, B \subset X$ two sets. If A is τ-closed and B is τ-compact, then $A + B$ is τ-closed.*

Proof. Fix an arbitrary $x^* \in X^* \setminus (A+B)$. This means that $(x^* - A) \cap B = \emptyset$. Since $x^* - A$ is τ-closed and B is τ-compact, there exists $V \in \mathcal{U}_\tau(0)$ such that $(x^* - A) \cap (B+V) = \emptyset$ (see a lemma before the Hahn-Banach Strong Separation Theorem). This easily implies that $(x^* - V) \cap (A + B) = \emptyset$. Since $x^* - V$ is a τ-neighborhood of x^*, and x^* was an arbitrary element of the complement of $A + B$, this shows that this complement is τ-open. $\square$

Theorem 0.11. *In a dual Banach space, the distance function of every w^*-closed set is w^*-l.s.c. (In particular, the distance function of any w-closed subset of a reflexive Banach space is w-l.s.c.)*

Proof. Let $E \subset X^*$ be a nonempty w^*-closed set. Since E is proximinal (Theorem 0.6), we have (for $r \geq 0$)

$$\{x^* : d_E(x^*) \leq r\} = \{x^* : \ \|x^* - e^*\| \leq r \text{ for some } e^* \in E\} = E + rB_{X^*},.$$

By previous lemma, all sublevel sets $\{d_E \leq r\}$ are w^*-closed. Thus d_E is w^*-l.s.c. $\square$

Remark 0.12. In the proof of Theorem 0.11, the equality $\{d_E \leq r\} = E + rB_{X^*}$ holds thanks to proximinality of E. For a general set E, we can only say that $\{d_E \leq r\} = \bigcap_{s>r}(E + sB_{X^*})$.

Existence of centers and medians. Let X be a normed space, and $A \subset X$ a nonempty bounded set. We are looking for a point $x_0 \in X$ that would somehow approximate the set A. Of course, we should decide what "approximates" means for us.

One possibility is to consider points, for which is minimal the "worst error", that is, minimizers (over X) of the function

$$\varphi_{A,\infty}(x) = \sup_{a \in A} \|x - a\| .$$

The points of X that minimize $\varphi_{A,\infty}$ are called *Chebyshev centers* of the set A.

This is not the unique possibility. For instance, if the set A is finite, that is $A = \{a_1, \ldots, a_n\}$, we can try to minimize the sum of the distances

$$\varphi_{A,1}(x) = \sum_{i=1}^{n} \|x - a_i\| ,$$

or the sum of their squares

$$\varphi_{A,2}(x) = \sum_{i=1}^{n} \|x - a_i\|^2 .$$

Any minimizer of $\varphi_{A,1}$ is called a *median* for A, and any minimizer of $\varphi_{A,2}$ is called a *square median* for A.

Since the functions of the type $x \mapsto \|x - a\|^p$ (with $a \in X$, $p \geq 1$) are convex, also the functions $\varphi_{A,k}$ ($k \in \{1, 2, \infty\}$) are convex. Moreover, they are also continuous since they are bounded an bounded sets.

Observe that, if X is a dual Banach space (that is $X = Z^*$ for some normed space Z), the functions $x \mapsto \|x - a\|^p$ are also w^*-l.s.c., and hence the same holds for the functions $\varphi_{A,k}$ ($k \in \{1, 2, \infty\}$). These latter functions are clearly also coercive. Thus we can apply Theorem 0.5 to obtain the following result.

Theorem 0.13. *In a dual Banach space (in particular, in a reflexive Banach space), every bounded set has a Chebyshev center, and every finite set has a median and a square median.*

Remark 0.14. It is known that even a three-point set can have no Chebyshev center, no median and no square median. Moreover, even if one of the three types of centers exists, it is not necessarily unique.