Chapter IV
Convex Analysis

March 11, 2023

1 Continuity of Convex Functions

4.1.6 Polar sets and strict separation

Fix a nonempty set C in $\mathbb{E}$.

- For points x in $\operatorname{int}(C)$ and ϕ in C°, prove $\langle \phi, x \rangle < 1$.
- Assume further that C is a convex set. Prove γ_C is sublinear.
- Assume in addition that $0 \in \operatorname{core}(C)$. Deduce
$$\operatorname{cl}(C) = \{x : \gamma_C(x) \leq 1\}$$
- Finally, suppose in addition that $D \subseteq \mathbb{E}$ is a convex set disjoint from the interior of C. By considering the Fenchel problem $\inf\{\delta_D + \gamma_C\}$, prove there is a closed halfspace containing D but disjoint from the interior of C.

Proof:

- Note that $\langle \phi, z \rangle \leq 1$ for all $z \in C$. Now since $x \in C^\circ$, we have $x + \epsilon d \in C$ for all $||d|| = 1$, for some $\epsilon > 0$. Thus, $\langle \phi, x \rangle + \langle \phi, \epsilon \frac{\phi}{||\phi||} \rangle \leq 1$. Hence, $\langle \phi, x \rangle < 1$.
- Note that $\gamma(\mu c) = \inf\{\lambda : \mu x \in \lambda C\} = \inf\{\lambda\mu : \mu x \in \lambda\mu C\} = \mu \inf\{\lambda : x \in \lambda C\}$, for $\mu > 0$. Thus, γ_C is homogeneous. Now notice
$$\{\lambda_1 : x \in \lambda_1 C\} + \{\lambda_2 : y \in \lambda_2 C\} \subseteq \{\lambda : x + y \in \lambda C\},$$
as C is convex we have $\lambda_1 C + \lambda_2 C = (\lambda_1 + \lambda_2)C$, and thus
$$\inf\{\lambda : x+y \in \lambda C\} \leq \inf\{\lambda_1 : x \in \lambda_1 C\} + \inf\{\lambda_2 : x \in \lambda_2 C\} \Rightarrow \gamma_C(x+y) \leq \gamma_C(x) + \gamma_C(y).$$
- Note that for $x \in C$, we have $\gamma_C(x) \leq 1$ and thus $C \subseteq \{x : \gamma_C(x) \leq 1\}$. Since the latter is closed, as γ is everywhere finite continuous, we have $\operatorname{cl}(C) \subseteq \{x : \gamma_C(x) \leq 1\}$. Now let $x \in \{x : \gamma_C(x) \leq 1\}$. Then since $0 \in C$, $\lambda_1 C \subseteq \lambda_2 C$ for all $\lambda_1 \leq \lambda_2$. Thus, $x \in (1+\epsilon)C$ for all $\epsilon > 0$. Hence, $x \in \operatorname{cl}(C)$.
- Note that
$$\inf_{x \in \mathbb{E}}\{\delta_D(x) + \gamma_C(x)\} \geq \sup_{\phi \in Y}\{-\delta_D^*(\phi) - \gamma_C^*(-\phi)\}.$$
Note that $\operatorname{dom}\gamma_C - \operatorname{dom}\delta_D = \mathbb{E} - D = \mathbb{E}$ hence CQ holds. However, note that if $\inf_{x \in \mathbb{E}}\{\delta_D(x) + \gamma_C(x)\} < 1$ then there exists $x \in D$ such that
$$x \in \lambda C \subseteq C$$
for some $\lambda < 1$. But, $\lambda C \subseteq \operatorname{int}(C)$ and hence $x \in D \cap \operatorname{int}(C)$ which is a contradiction. Thus, $\inf_{x \in \mathbb{E}}\{\delta_D(x) + \gamma_C(x)\} \geq 1$.

  But, this contradicts the fact that $D \cap \operatorname{int}(C) = \emptyset$. Thus, $\inf_{x \in \mathbb{E}}\{\delta_D(x) + \gamma_C(x)\} > 0$. Hence, there exists $\phi^* \in Y$ such that $-\delta_D^*(\phi^*) - \gamma_C^*(-\phi^*) > 0$ or $\delta_D^*(\phi^*) + \gamma_C^*(-\phi^*) < 0$. Note that $\gamma_C^* = \delta_{C^\circ}$. Thus, $-\phi^* \in C^\circ$ and also $\langle \phi^*, y \rangle \leq -1$ for all $y \in D$. Thus,
$$\langle \phi^*, y \rangle \leq -1 < \langle \phi^*, x \rangle \; \forall x \in C^\circ, y \in D.$$

4.1.7. Polar calculus

Suppose C and D are subsets of $\mathbb{E}$.

- Prove $(C \cup D)^\circ = C^\circ \cap D^\circ$.
- If C and D are convex, prove

$$\operatorname{conv}(C \cup D) = \cup_{\lambda\in[0,1]}(\lambda C + (1-\lambda)D).$$

- If C is a convex cone and the convex set D contains 0, prove

$$C + D \subseteq cl\, \operatorname{conv}(C \cup D).$$

  Now suppose the closed convex sets K and H of $\mathbb{E}$ both contain 0.

- Prove $(K \cap H)^\circ = cl\, \operatorname{conv}(K^\circ \cup H^\circ)$.

Proof:

- Note that

$$\phi \in (C\cup D)^\circ \iff \langle \phi, x\rangle \le 1\ \forall x \in C\cup D \iff \langle \phi, x\rangle \le 1\ \forall x \in C\ \&\ \langle \phi, y\rangle \le 1\ \forall y \in D.$$

  Thus, $x \in (C \cup D)^\circ$ if and only if $x \in C^\circ \cap D^\circ$.

- For $\lambda \in [0,1]$, let $X_\lambda = \lambda C + (1-\lambda)D$. Now since $\operatorname{conv}(C \cup D)$ is convex and $C, D \subseteq \operatorname{conv}(C\cup D)$, $X_\lambda \subseteq \operatorname{conv}(C\cup D)$ for all $\lambda \in [0,1]$. So, $\cup_{\lambda\in[0,1]}X_\lambda \subseteq \operatorname{conv}(C\cup D)$.

  Conversely, since, $X_1 = C$ and $X_0 = D$. So, $C \cup D \subseteq \cup_{\lambda\in[0,1]}X_\lambda$. Thus, in order to prove, $\operatorname{conv}(C \cup D) \subseteq \cup_{\lambda\in[0,1]}X_\lambda$, we just need to show that $\cup_{\lambda\in[0,1]}X_\lambda$ is convex. Let $\lambda_1, \lambda_2 \in [0,1]$ and

$$\lambda_1 c_1 + (1-\lambda_1)d_1 \in X_{\lambda_1}\ \&\ \lambda_2 c_2 + (1-\lambda_2)d_2 \in X_{\lambda_2},$$

  in which $c_1, c_2 \in C$ and $d_1, d_2 \in D$. Now we need to show for any $\mu \in [0,1]$ we have

$$\mu(\lambda_1 c_1 + (1-\lambda_1)d_1) + (1-\mu)(\lambda_2 c_2 + (1-\lambda_2)d_2) \in \cup_{\lambda\in[0,1]}X_\lambda.$$

  However, $\mu(\lambda_1 c_1 + (1-\lambda_1)d_1) + (1-\mu)(\lambda_2 c_2 + (1-\lambda_2)d_2)$ equals to

$$(\mu\lambda_1 c_1 + (1-\mu)\lambda_2 c_2) + (\mu(1-\lambda_1)d_1 + (1-\mu)(1-\lambda_2))d_2,$$

  and if $t = \mu\lambda_1 + (1-\mu)\lambda_2$, then the above equals to, notice $1-t = \mu(1-\lambda_1) + (1-\mu)(1-\lambda_2)$,

$$t(\frac{\mu\lambda_1}{t}c_1 + \frac{(1-\mu)\lambda_2}{t}c_2) + (1-t)(\frac{\mu(1-\lambda_1)}{1-t}d_1 + \frac{(1-\mu)(1-\lambda_2)}{1-t}d_2).$$

  Note that $0 \le t, 1-t$ and thus $t \in [0,1]$.

- Note that C is a cone and thus $\lambda C = C$ for all $\lambda \in (0,1]$. So

$$\cup_{\lambda\in[0,1]}(\lambda C + (1-\lambda)D) = D \cup \left(\cup_{\lambda\in(0,1]}(C + (1-\lambda)D)\right).$$

  Now since, $0 \in D$ we should have $(1-\lambda)D \subseteq D$ for all $\lambda \in [0,1]$. Thus,

$$\cup_{\lambda\in(0,1]}(C + (1-\lambda)D) = C + D.$$

  So,

$$\operatorname{conv}(C \cup D) = \cup_{\lambda\in[0,1]}(\lambda C + (1-\lambda)D) = (C+D) \cup D.$$

  So, the proof is complete.

- Note that since K, H are closed convex set we obtain

$$K = K^{\circ\circ}, H = H^{\circ\circ}.$$

  Thus, due to part 1,

$$(K \cap H)^\circ = (K^\circ \cup H^\circ)^\circ$$

4.1.13. Existence of extreme points

Prove any nonempty compact convex set $C \subseteq \mathbb{E}$ has an extreme point, without using Minkowski's theorem, by considering the furthest point in C from the origin.

Proof: Since C is compact closed, there exists $\bar{x} \in C$ such that $||\bar{x}|| = \sup_{a\in C} ||a||$. We prove $\bar{x}$ is an extreme point. Let $\bar{x} = \lambda a + (1-\lambda)b$ for some $a, b \in C$. Then $||\bar{x}|| \leq \lambda||a|| + (1-\lambda)||b|| \leq \lambda||\bar{x}|| + (1-\lambda)||\bar{x}|| = ||\bar{x}||$. Thus $||a|| = ||b|| = ||\bar{x}||$ and hence $a, b, \bar{x}$ are collinear and since they have the same norm we conclude that $a = b = \bar{x}$.

Note that fixing any point $c \in \mathbb{E}$, there exists $\bar{x} \in C$ such that $||\bar{x} - c|| = \sup_{a\in C} ||a - c||$. Thus, $\bar{x}$ is also an extreme point.

Remark: All the extreme points can be obtained this way.

4.1.14.

Given a supporting hyperplane H of a convex set $C \subseteq \mathbb{E}$, any extreme points of $C \cap H$ is also an extreme point of C.

Proof: Let $\bar{x}$ be an extreme point of $C \cap H$ in which

$$H = \{x \in \mathbb{E} : \langle a, x - \bar{x}\rangle = 0\},$$

and for all $x \in C$, $\langle a, x\rangle \geq \langle a, \bar{x}\rangle$. Now let $\bar{x} = \lambda x_1 + (1-\lambda)x_2$ for some $x_1, x_2 \in C$. Then

$$\langle a, x_1\rangle, \langle a, x_2\rangle \geq \langle a, \bar{x}\rangle \;\&\; \langle a, \bar{x}\rangle = \lambda\langle a, x_1\rangle + (1-\lambda)\langle a, x_2\rangle.$$

Thus, $\langle a, x_1\rangle = \langle a, x_2\rangle = \langle a, \bar{x}\rangle$, and hence $x_1, x_2 \in H$. However, $\bar{x}$ is an extreme point of $C \cap H$. Thus, $\bar{x} = x_1 = x_2$. This completes the proof.

4.1.15.

For any compact convex set $C \subseteq \mathbb{E}$, prove $C = \operatorname{conv}(\operatorname{bd} C)$.

Proof: Clearly, $\operatorname{conv}(\partial C) \subseteq C$. Now let $\bar{x} \in C \setminus \operatorname{conv}(\partial C)$. $\partial C = C \setminus \operatorname{int}(C)$ and thus ∂C is closed. Also $\partial C \subseteq C$ and thus it is compact. We know that the convex hull of a closed set is closed, we realize that $\operatorname{conv}(\partial C)$ is closed ,convex and bounded.

Now consider the following general case:

Question: Let $C_1 \subsetneq C_2$ be two compact convex sets in $\mathbb{E}$. We know there exists $\bar{x} \in C_2$ such that $d(\bar{x}, C_1) = \sup_{a \in C_2} d(a, C_1)$ as C_2 is compact. Is it true that $\bar{x}$ is an extreme point of C_2?

Answer: Suppose $\bar{x} = \lambda x_1 + (1-\lambda)x_2$, with $x_1, x_2 \in C_2$. Now, since C_1 is compact, there exists $c \in C_1$ such that $||\bar{x} - c|| = \inf_{a \in C_1} ||x - a||$. Thus, due to the way we picked $\bar{x}$, we have

$$||\bar{x} - a|| \geq ||x_1 - a||, ||\bar{x} - a|| \geq ||x_2 - a||.$$

But,

$$||x-a|| = ||\lambda(x_1-a)+(1-\lambda)(x_2-a)|| \leq \lambda||x_1-a||+(1-\lambda)||x_2-a|| \leq \lambda||x-a||+(1-\lambda)||x-a|| = ||x-a||.$$

Thus, $||x - a|| = ||x_1 - a|| = ||x_2 - a||$. Since in the triangle inequality above, equality holds, $x - a, x_1 - a, x_2 - a$ are collinear, and since $||x - a|| = ||x_1 - a|| = ||x_2 - a||$, we have $x_1 = x_2 = x$. Hence, x is an extreme point of C_2.

Now, back to the main question, since $\operatorname{conv}(\partial C) \subseteq C$, if $\operatorname{conv}(\partial C) \neq C$, then the above discussion gives us an extreme point $\bar{x}$ of C lying outside of $\operatorname{conv}(\partial C)$. However, we know that no point in $\operatorname{int}(C)$ can be an extreme point. Thus $\bar{x} \in \partial(C)$ and this is a contradiction.

4.1.16. A converse of Minkowski's theorem

Suppose D is a subset of a compact convex set $C \subseteq \mathbb{E}$ satisfying $\operatorname{cl}(\operatorname{conv}(D)) = C$. Prove $\operatorname{ext} C \subseteq \operatorname{cl} D$.

Proof: Since $\operatorname{conv}(\operatorname{cl} D) = \operatorname{cl}(\operatorname{conv}(D))$, we need to prove for $D \subseteq \mathbb{E}$ closed with $C = \operatorname{conv}(D)$, we should have $\operatorname{ext}(C) \subseteq D$. Suppose that $\operatorname{aff}(C) = \mathbb{E}$. Then, let $\bar{x} \in \operatorname{ext}(C)$. Then $\bar{x} = \sum_{i=1}^{m} \lambda_i x_i$ for some $\lambda_i \geq 0$ with $\sum_{i=1}^{m} \lambda_i = 1$ and some $x_i \in D$. Since, $s \in \operatorname{bd}(C)$, there exists a supporting hyperplane for C at $\bar{x}$. Clearly, $x_i \in H$ and thus $x \in \operatorname{conv}(D \cap H)$. x must be an extreme point of $\operatorname{conv}(D \cap H)$ and thus due to induction, $x \in D \cap H$. Note that $\dim(D \cap H) < \dim(D)$. This completes the proof.

4.1.17 Extreme points

Consider a compact convex set $C \subseteq \mathbb{E}$.

- If $\dim \mathbb{E} \leq 2$, prove the set $\operatorname{ext}(C)$ is closed.
- If $\mathbb{E}$ is $\mathbb{R}^3$ and C is the convex hull of the set

$$\{(x, y, 0) : x^2 + y^2 = 1\} \cup \{(1, 0, 1), (1, 0, -1)\},$$

prove $\operatorname{ext}(C)$ is not closed.

Proof:

- Suppose that $\mathrm{aff}(C) = \mathbb{E}$. If $\dim \mathbb{E} = 0$ then $\mathrm{ext}(C) = C = \mathbb{E} = \{0\}$. If $\dim \mathbb{E} = 1$, then C is closed and compact, $C = [a, b]$ for some $a, b \in \mathbb{R}$. In this case, $\mathrm{ext}(C) = \{a, b\}$.

  Now suppose $\dim(E) = 2$. Let $x_i \in \mathrm{ext}(C) \to x$ for some $x \in \mathrm{bd}(C)$. Then there exists a hyperplane
  $$H := \{\phi : \langle a, \phi\rangle = \beta\},$$
  such that $x \in H$ and $\langle a, y\rangle \geq \beta$ for all $y \in C$. If $x \in \mathrm{ext}(C \cap H)$ then $x \in \mathrm{ext}(C)$. So suppose that $x \notin \mathrm{ext}(C \cap H)$. Then if $||b|| = 1$ such that $b \in H$ and $\langle b, a\rangle = 0$, then for some $\epsilon > 0$, $x + tb \in C$ for all $t \in [-\epsilon, +\epsilon]$. Now let $\bar{x} \in C\backslash H$. Then let $D = \mathrm{conv}\{x + \epsilon b, x - \epsilon b, \bar{x}\}$. Now there exists $\delta > 0$ such that $B_\delta(x) \cap C \subseteq D$. So since, $x_i \to x$, for some n, $x_n \in D$. So, x_n can't be an extreme point as it is a convex combination of $x + \epsilon, x - \epsilon, \bar{x}$.

- We show that $P = (1, 0, 0) \in \mathrm{cl}(\mathrm{ext}(C))$ but at the same time $(1, 0, 0)$ is not an extreme point. In fact, P is not an extreme point is clear: $P = \frac{1}{2}((1, 0, 1) + (1, 0, -1))$ and thus P is not extreme point. Now we show $Q = (x, y, 0) \in \mathrm{ext}(C)$ for all $Q \neq P$. Let
  $$H_Q = \{z \in \mathbb{R}^3 : \langle z, Q\rangle = 1\}.$$
  Then since $x^2 + y^2 = 1$ and $Q \neq P$, $x < 1$ and hence
  $$\langle (1, 0, -1), Q\rangle = \langle (1, 0, 1), Q\rangle = x < 1 \Rightarrow (1, 0, -1), (1, 0, 1) \in \{z \in \mathbb{R}^3 : \langle z, Q\rangle < 1\}.$$
  However, let $(x', y', 0) \in C$ with $(x', y') \neq (x, y)$. Then due to Cauchy-Schwartz inequality we have
  $$\langle (x', y'), (x, y)\rangle < \sqrt{||(x', y')||||(x, y)||} = 1 \Rightarrow \langle (x', y', 0), Q\rangle < 1.$$
  Thus, C which is the convex hull of $(x', y', 0)$ with $x'^2 + y'^2 = 1$ and $(1, 0, 1), (1, 0, -1)$ lie inside $\{z \in \mathbb{R}^3 : \langle z, Q\rangle \leq 1\}$ and P is the only point of $C \cap H$. Hence, P is a vertex.

  Now let $P_n = (\sqrt{1 - \frac{1}{n}}, \sqrt{\frac{1}{n}}, 0)$ for $n \in \mathbb{N}$. Then $\lim_n P_n = P$ and thus $P \in \mathrm{cl}(\mathrm{ext}(C))$.

4.1.21 Essential smoothness

For any convex function f and any point $x \in \mathrm{bd}\,\mathrm{dom}\, f$, prove $\partial f(x)$ is either empty or unbounded. Deduce that a function is essentially smooth if and only if its subdifferential is always singleton or empty.

Proof: Let $C = \mathrm{dom}\, f$. Note that according to problem 4.1.20, $N_C(\bar{x}) = \{0\}$ implies $\bar{x} \in \mathrm{ri}C$. Thus, there exists $s \in N_C(\bar{x})$ and if $\phi \in \partial f(\bar{x})$ then
$$\langle \phi + s, y - \bar{x}\rangle + f(\bar{x}) \leq f(y)\ \forall y \in \mathbb{E}.$$
Note that the above holds for $y \notin C$ obviously, and if $y \in C$, then $\langle s, y - \bar{x}\rangle \leq 0$ and the result follows from the fact that $\phi \in \partial f(\bar{x})$. Thus, $\phi + ts \in \partial f(\bar{x})$ for all $t \in \mathbb{R}_+$.

---

2 Fenchel Biconjugation

4.2.12 Compact bases for cones

Consider a closed convex cone K. Using Moreau-Rockafellar theorem, show that a point x lies in $\mathrm{int}(K)$ if and only if the set $\{\phi \in K^- : \langle \phi, x\rangle \geq -1\}$ is bounded. If the set $\{\phi \in K^- : \langle \phi, x\rangle = -1\}$ is nonempty and bounded, prove $x \in \mathrm{int}(K)$.

Proof: First, suppose $x \in \mathrm{int}(K)$. Then there exists $\epsilon > 0$ such that $x + \epsilon d \in K$ for all $d \in \mathbb{E}$ with $||d|| = 1$. So, if $\langle \phi, x\rangle \geq -1$ for some $0 \neq \phi \in K^-$, then since $x + \epsilon\phi/||\phi|| \in K$, then $\langle \phi, x + \epsilon\phi/||\phi||\rangle \leq 0$. Thus, $\epsilon||\phi|| - 1 \leq 0$ and hence $||\phi|| \leq 1/\epsilon$.

Conversely, suppose $\{\phi \in K^- : \langle \phi, x\rangle \geq -1\}$ is bounded. Thus first note that if $\langle \phi, x\rangle = 0$ for some $\phi \in K^-$ then $\phi = 0$. Now, let $f(.) = \langle ., x\rangle + \delta_{K^-}(.)$. Then clearly,

$$f^*(\psi) = \sup_{\phi\in K^-} \langle \psi, \phi\rangle - \langle x, \phi\rangle = \sup_{\phi\in K^-} \langle \psi - x, \phi\rangle.$$

Note that $f^*(0) = 0$ and thus f^* is bounded about 0. Thus, for all $\psi_i \to x$, $\langle \psi_i - x, \phi\rangle \leq 0$ for all $\phi \in K^-$. So, $\langle \psi_i, \phi\rangle \leq \langle x, \phi\rangle \leq 0$ for all $\phi \in K^-$. Thus, $\psi_i \in K$, and hence x lies inside the interior of K.

Now suppose the set $\{\phi \in K^- : \langle \phi, x\rangle = -1\}$ is nonempty and bounded. Then let $0 \neq \phi \in K^-$ such that $\langle \phi, x\rangle \geq -1$. Then since $\langle \phi, x\rangle \neq 0$, then $\langle \phi/|\langle \phi, x\rangle|, x\rangle = -1$. So, $||\phi||/|\langle \phi, x\rangle| \leq M$. Thus, $||\phi|| \leq M|\langle \phi, x\rangle| \leq M$.

4.2.13

For any function $h : \mathbb{E} \to [-\infty, +\infty]$, prove the set $\mathrm{cl}(\mathrm{epi}\, h)$ is the epigraph of some function.

Proof: Let

$$f(x) := \inf\{y : (x, y) \in \mathrm{cl}(\mathrm{epi}\, h)\}.$$

Then, note that if $(x, y_i) \in \mathrm{cl}(\mathrm{epi}\, h) \to (x, f(x))$, then $(x, f(x)) \in \mathrm{cl}(\mathrm{epi}\, h)$ as the latter is closed. Now, note that for $r > 0$ and some $(x, y) \in \mathrm{cl}(\mathrm{epi}\, h)$, since there exists $(x_i, y_i) \in \mathrm{epi}\, h$ such that $(x_i, y_i) \to (x, y)$. Then since $(x_i, y_i + r) \in \mathrm{epi}\, h$ and $(x_i, y_i + r) \to (x, y + r)$ we realize that $(x, y + r) \in \mathrm{cl}(\mathrm{epi}\, h))$. Thus, $\mathrm{cl}(\mathrm{epi}\, h) = \mathrm{epi}\, f$.

4.2.14 Lower semicontinuity and closure

For any convex function $h : \mathbb{E} \to [-\infty, +\infty]$ and any point x^0 in $\mathbb{E}$, prove

$$(\mathrm{cl}\, h)(x^0) = \lim_{\delta\downarrow 0} \inf_{||x-x^0||\leq\delta} h(x).$$

Deduce

Proposition 4.2.7 If a function $f : \mathbb{E} \to [-\infty, +\infty]$ is convex then it is lower semi-continuous at a point x where it is finite if and only if $f(x) = \mathrm{cl}\, f(x)$. In this case, f is proper.

Proof:

4.2.15

For any point x in $\mathbb{E}$ and any function $h : \mathbb{E} \to (-\infty, +\infty]$ with a sub-gradient at x, prove h is lower semicontinuous at x.

Proof: Note that
$$(\mathrm{cl}\, f)(x) = \lim_{\delta \downarrow 0} \inf_{||y-x|| \leq \delta} f(y) \leq f(x).$$
However, suppose $s \in \partial f(x)$. Then
$$\langle s, y-x \rangle + f(x) \leq f(y)\ \forall y \in \mathbb{E},$$
and thus
$$\inf_{||y-x|| \leq \delta} f(y) \geq f(x) - \delta ||s|| \Rightarrow (\mathrm{cl}\, f)(x) \geq f(x).$$
Thus, $\mathrm{cl}\, f(x) = f(x)$ and due to Proposition 4.2.7, since f is finite at x as otherwise $f \equiv +\infty$ and in that case obviously f is lower semicontinuous everywhere, we have f is lowersemicontinuous at x.

4.2.16. Von Neumann's minmax theorem

Suppose Y is a Euclidean space. Suppose that the sets $C \subseteq \mathbb{E}$ and $D \subseteq Y$ are nonempty and convex with D closed and that the map $A : \mathbb{E} \to Y$ is linear.

Proof:

- By considering the Fenchel problem
$$\inf_{x \in \mathbb{E}} \{\delta_C(x) + \delta_D^*(Ax)\}$$
prove
$$\inf_{x \in \mathbb{E}} \sup_{y \in D} \langle y, Ax \rangle = \max_{y \in D} \inf_{x \in C} \langle y, Ax \rangle$$
(where the max is attained if finite), under the assumption
$$0 \in \mathrm{core}(\mathrm{dom}\, \delta_D^* - AC).$$

- Prove property above holds in either of the two cases
  1. D is bounded, or
  2. A is surjective and 0 lies in $\mathrm{int}\, C$.

- Suppose both C and D are compact. Prove
$$\min_{x \in C} \max_{y \in D} \langle y, Ax \rangle = \max_{y \in D} \min_{x \in C} \langle A^* y, x \rangle.$$

Proof:

- Note that under the assumption $0 \in \mathrm{core}(\mathrm{dom}\, \delta_D^* - AC)$.
$$\inf_{x \in \mathbb{E}} \{\delta_C(x) + \delta_D^*(Ax)\} = \sup_{y \in Y} \{-\delta_C^*(A^* y) - \delta_D(-y)\},$$
since D is closed, δ_D is a closed convex function. Also,
$$\inf_{x \in \mathbb{E}} \{\delta_C(x) + \delta_D^*(Ax)\} = \inf_{x \in C} \{\delta_D^*(Ax)\} = \inf_{x \in C} \sup_{y \in D} \langle y, Ax \rangle = \inf_{x \in C} \max_{y \in D} \langle y, Ax \rangle.$$

Also,

$$\sup_{y\in Y}\{-\delta_C^*(A^*y)-\delta_D(-y)\} = \sup_{y\in -D}\{-\delta_C^*(A^*y)\} = \sup_{y\in -D}\{-\sup_{x\in C}\langle x, A^*y\rangle\} = \sup_{y\in -D}\inf_{x\in C}\langle x, -A^*y\rangle,$$

which equals to $\max_{y\in D}\inf_{x\in C}\langle x, A^*y\rangle$.

- Note that
$$\delta_D^*(x) = \sup_{y\in D}\langle x, y\rangle \Rightarrow \operatorname{dom}\delta_D^* = \mathbb{E}.$$
Also, if A is surjective and 0 lies in interior of C, then since $0 \in \operatorname{dom}\delta_D^*$, then due problem 4.1.9 we are done.

- Clear!

4.2.8. Closed subdifferential
If a function $h : \mathbb{E} \to (\infty, +\infty]$ is closed, prove the multifunction ∂h is closed: that is,

$$\phi_r \in \partial h(x_r), x_r \to x, \phi_r \to \phi \Rightarrow \phi \in \partial h(x).$$

Deduce that if h is essentially smooth and a sequence of points x_r in $\operatorname{int}(\operatorname{dom} h)$ approaches a point in $\partial(\operatorname{dom} h)$ then $||\nabla h(x_r)|| \to \infty$.

Proof: Let $y \in \mathbb{E}$ then

$$\langle \phi_r, y - x_r\rangle + h(x_r) \leq h(y).$$

But, since h is lowersemicontinuous at x, we have

$$h(x) \leq \liminf_{r\to+\infty} h(x_r) \leq h(y) - \langle \phi, y - x\rangle.$$

Hence, $\phi \in \partial h(x)$.

Note that if $\partial h(x) \neq \emptyset$ then h has Gâteaux differential at x and hence $x \in \operatorname{int}\operatorname{dom}(h)$. So $\partial h(x) = \emptyset$ and thus if $||\nabla h(x_r)|| \leq C$ for some C, w can assume that $\nabla h(x_r) \to \phi$ for some $\phi \in \mathbb{E}$ and thus $\phi \in \partial h(x)$ which is a contradiction.

4.2.9. Support functions
Prove that if the set $C \subseteq \mathbb{E}$ is nonempty then δ_C^* is a closed sublinear function and $\delta_C^{**} = \delta_{\operatorname{cl}\operatorname{conv} C}$. Prove that if C is also bounded then δ_C^* is everywhere finite.

- Prove that any sets $C, D \subseteq \mathbb{E}$ satisfy
$$\delta_{C+D}^* = \delta_C^* + \delta_D^* \text{ and}$$
$$\delta_{\operatorname{conv}(C\cup D)}^* = \max(\delta_C^*, \delta_D^*).$$

- Suppose the function $h : \mathbb{E} \to (-\infty, +\infty]$ is positively homogeneous, and define a closed convex set
$$C = \{\phi \in \mathbb{E} : \langle \phi, d\rangle \leq h(d)\ \forall d\}.$$
Prove $h^* = \delta_C$. Prove that if h is in fact sublinear and everywhere finite then C is nonempty and compact.

Proof:

- Let $c \in \mathbb{R}$, then
$$\delta_C^*(x) \le c \iff \langle x, y\rangle - \delta_C(y) \le c \; \forall y \in C \iff x \in \cap_{y\in C} H_y,$$
where $H_y = \{\phi \in \mathbb{E} : \langle \phi, y\rangle \le c\}$. Thus δ_C^* is closed.
Also, let $\lambda \in \mathbb{R}_+$, then
$$\delta_C^*(\lambda x) = \sup_{y\in\mathbb{E}} \langle \lambda x, y\rangle - \delta_C(y) = \lambda \sup_{y\in\mathbb{E}} \langle x, y\rangle - \delta_C(y) = \lambda \delta_C^*(x).$$
Now let $x, y \in \mathbb{E}$, then
$$\langle x + y, z\rangle - \delta_C(z) = \langle x + y, z\rangle - 2\delta_C(z).$$
On the other hand,
$$\delta_C^*(x) + \delta_C(z) \ge \langle x, z\rangle \; , \; \delta_C^*(y) + \delta_C(z) \ge \langle y, z\rangle.$$
Thus, it can be derived from the above two statements
$$\langle x + y, z\rangle - \delta_C(z) \le \delta_C^*(x) + \delta_C^*(y).$$
Now, let $\bar{x} \in \operatorname{cl}\operatorname{conv}(C)$, then we want to show that
$$\delta_C^{**}(\bar{x}) = 0.$$
Note that $\delta_C^*(0) = 0$ as C is nonempty, thus $\delta_C^{**}(\bar{x}) \ge 0$. Now we want to show for all $y \in \mathbb{E}$, $\langle \bar{x}, y\rangle - \delta_C^*(y) \le 0$, or $\langle \bar{x}, y\rangle \le \delta_C^*(y)$. So, we can suppose that $x \in \operatorname{conv}(C)$ and hence $x = \sum_i \lambda_i x_i$ for some $x_i \in C$. Now note that $\langle x_i, y\rangle \le \delta_C^*(y)$ and hence
$$\lambda_i \langle x_i, y\rangle \le \lambda_i \delta_C^*(y) \Rightarrow \langle \bar{x}, y\rangle \le \delta_C^*(y),$$
for all $y \in C$. Thus $\delta_C^{**}(\bar{x}) \le 0$.
Now let $\bar{x} \in C \setminus \operatorname{cl}\operatorname{conv}(C)$. So, there exists $\phi \in \mathbb{E}$ such that $\langle \bar{x}, \phi\rangle > b \ge \langle y, \phi\rangle$ for all $y \in C$. Thus $\delta_C^*(y) \le b$
$$\delta_C^{**}(\bar{x}) \ge \langle \bar{x}, \phi\rangle - \delta_C^*(\phi) \ge \langle \bar{x}, \phi\rangle - b > 0.$$
Thus $\delta_C^{**}(\bar{x}) \ge \langle \bar{x}, \lambda\phi\rangle - \delta_C^*(\lambda\phi) = \lambda(\langle \bar{x}, \phi\rangle - \delta_C^*(\phi)) \to +\infty$. Thus, $\delta_C^{**}(\bar{x}) = +\infty$. Now, suppose $C \subseteq rB$ for some $r > 0$. Then $\delta_C^*(\bar{x}) = \sup_{y\in C} \langle x, y\rangle \le \sup_{y\in rB} \langle x, y\rangle \le r||x|| < +\infty$.

- Very easy! Omitted.

- C is trivially closed and convex, and if $\phi \notin C$, then $h^*(\phi) > 0$ by definition of C. However, if $\langle \phi, d\rangle - h(d) > 0$ for some $d \in \mathbb{E}$, then $\langle \phi, td\rangle - h(td) \to +\infty$ as $t \to +\infty$. Thus, $h^*(\phi) = \infty$. Now, let $\phi \in C$, then $\langle \phi, d\rangle - h(d) \le 0$ by definition and hence $h^*(\phi) \le 0$. However, $\langle \phi, 0\rangle - h(0) = 0$. Thus, $h^*(\phi) = 0$. Therefore, we have proved that $h^* = \delta_C$. Now, since h is sublinear and everywhere finite $h(d) \le M$, for some $M > 0$, for all $||d|| = 1$. Thus, for $\phi \in C$, we have $\langle \phi, d\rangle \le M$ for all $||d|| = 1$ and hence $||\phi|| \le M$. Now, if $C = \emptyset$, then $h^* \equiv +\infty$ and thus since h is closed and convex we have $h = h^{**}$, we obtain $h \equiv -\infty$ which is a contradiction.

4.2.21 cofiniteness

Consider a function $h : \mathbb{E} \to (\infty, +\infty]$ and the following properties:

1. $h(.) - \langle \phi, . \rangle$ has bounded level sets for all ϕ in $\mathbb{E}$.
2. $\lim_{||x|| \to \infty} ||x||^{-1} h(x) = +\infty$.
3. h^* is everywhere finite.

Complete the following steps.

- Prove properties 1 and 2 are equivalent.
- If h is closed, convex and proper, use and Moreau-Rockafellar theorem to prove properties 1 and 3 are equivalent.

Proof:

- Suppose 1 holds. Then if $\lim_{||x|| \to \infty} ||x||^{-1} h(x) = +\infty$ does not hold true then there exists x_i with $||x_i|| \to +\infty$ such that $||x_i||^{-1} h(x_i) \leq C$ for some constant $C > 0$. Thus, since $||x_i|| \to +\infty$ and hence

  $$h(x_i) - \langle \phi, x_i \rangle \to +\infty,$$

  for all $\phi \in \mathbb{E}$. However, without loss of generality, suppose that $||x_i||^{-1} x_i \to v$ for some $v \in \mathbb{E}$. Then, let $\phi = Cv$. We have

  $$||x_i||^{-1} \langle \phi, x_i \rangle \uparrow C.$$

  However, $||x_i||^{-1}(h(x_i) - \langle \phi, x_i \rangle) \leq C - C \leq 0$ and hence $h(x_i) - \langle \phi, x_i \rangle \to +\infty$ cannot hold true.

  Conversely, suppose that 2 holds, and $\phi \in \mathbb{E}$. Then $\lim_{||x|| \to \infty} ||x||^{-1}(h(x) - \langle \phi, x \rangle) = +\infty$. and thus without loss of generality we can assume that $\phi = 0$. Now suppose $h(x) \leq M$ is not bounded for some $M \in \mathbb{R}$ and hence there exists $||x_i|| \to +\infty$ such that $h(x_i) \leq M$. So $\lim_{i \to \infty} ||x_i||^{-1} h(x_i) = 0$, which is a contradiction.

- Let $h_\phi(.) := h(.) - \langle \phi, . \rangle$. Then $h^*_\phi(\psi) = h^*(\phi + \psi)$. Thus h^* is continuous at ϕ if and only if h^*_ϕ is continuous at 0. But, we know h^*_ϕ has bounded level set if and only if h^*_ϕ is continuous at zero. Hence, h^* is finite everywhere.

4.2.22 Computing closures

- Prove that any closed convex function $g : \mathbb{R} \to (\infty, +\infty]$ is continuous on its domain.
- Consider a convex function $f : \mathbb{E} \to (\infty, +\infty]$. For any point $x \in \mathbb{E}$ and any $y \in \operatorname{int}(\operatorname{dom} f)$, prove

  $$f^{**}(x) = \lim_{t \uparrow 1} f(y + t(x - y)).$$

Proof:

- Without loss of generality assume that $\mathrm{aff}(\mathrm{dom}\, f) = \mathbb{E}$. Note that $f = f^{**}$ is continuous at 0 if and only if f^* has bounded level sets.

  Suppose this does not hold. Then there exists $x_i \in \mathbb{E}$ such that $||x_i|| \to +\infty$ and

  $$f^*(x_i) \leq M \text{ for some } M > 0.$$

  So, for every $y \in \mathbb{E}$ and every $i = 1, 2, \cdots$ we have

  $$\langle x_i, y\rangle \leq f(y) + M \Rightarrow \langle x, y\rangle \leq f(y) + M \ \forall x \in \mathrm{conv}\{x_1, x_2, \cdots\}.$$

  But since $||x_i|| \to +\infty$, $C = \mathrm{conv}\{x_1, x_2, \cdots\}$ is unbounded and thus $d \in 0^+(C)$ for some $d \neq 0$. Thus, for all $t \in \mathbb{R}_+$,

  $$\langle x_1 + td, y\rangle \leq f(y) + M \Rightarrow \langle d, y\rangle \leq 0 \ \forall y \in \mathrm{ri}(\mathrm{dom}\, f).$$

  But, then $d = 0$ as $\mathrm{aff}(\mathrm{dom}\, f) = \mathbb{E}$. This contradiction completes the proof.

4.2.24 Fisher information function

Let $f : \mathbb{R} \to (\infty, +\infty]$ be a given function, and define a function $g : \mathbb{R}^2 \to (\infty, +\infty]$ by

$$g(x, y) = \begin{cases} yf(\frac{x}{y}) & \text{if } y > 0 \\ +\infty & \text{otherwise.} \end{cases}$$

- Prove g is convex if and only if f is convex.

Proof:

- Suppose f is convex with $\mathrm{epi}\, f = C \subseteq \mathbb{R}^2$. Then C is convex. Now $((x, y), r) \in \mathrm{epi}\, g$ for some $y > 0$ if and only if

  $$r \geq yf(\frac{x}{y}) \iff (\frac{x}{y}, \frac{r}{y}) \in C \iff (x, r) \in yC.$$

  Now for some $\lambda \in (0, 1)$ and $((x_1, y_1), r_1), ((x_2, y_2), r_2) \in \mathrm{epi}\, g$ we have $\lambda((x_1, y_1), r_1) + (1 - \lambda)((x_2, y_2), r_2) \in \mathrm{epi}\, g$ if and only if

  $$(\lambda x_1 + (1 - \lambda)x_2, \lambda r_1 + (1 - \lambda)r_2) \in (\lambda y_1 + (1 - \lambda)y_2)C,$$

  which holds true as C is convex and $(x_1, r_1) \in y_1C$ and $(x_2, r_2) \in y_2C$.

  Now conversely suppose that $\mathrm{epi}\, g$ is convex, then $(x, r) \in C$ if and only if $((x, 1), r) \in \mathrm{epi}\, g$. Thus, C is the image of the projection of $\mathrm{epi}\, g \cap \{y = 1\}$ onto $\mathbb{R}^2$. Hence, C is convex.

---

3 Lagrangian Duality

4.3.1 Weak duality

Prove that the primal and dual values p and d defined by equations

$$p = \inf_{x\in\mathbb{E}} \sup_{\lambda\in\mathbb{R}^n_+} L(x;\lambda), d = \sup_{\lambda\in\mathbb{R}^n_+} \inf_{x\in\mathbb{E}} L(x;\lambda),$$

satisfies $d \leq p$.

Proof: We only need to show $\inf_{x\in\mathbb{E}} L(x;\tilde{\lambda}) \leq \sup_{\lambda\in\mathbb{R}^n_+} L(\tilde{x};\lambda)$ for fixed $\tilde{\lambda} \in \mathbb{R}^n_+$ and $\tilde{x} \in \mathbb{E}$. But, $\inf_{x\in\mathbb{E}} L(x;\tilde{\lambda}) \leq L(\tilde{x};\tilde{\lambda}) \leq \sup_{\lambda\in\mathbb{R}^n_+} L(\tilde{x};\lambda)$.

4.3.2

Calculate the Lagrangian dual of the problem:

$$\inf_{x\in\mathbb{R}^n_{++}} \{\sum_{i=1}^n \frac{c_i}{x_i} : \sum_{i=1}^n a_i x_i \leq b\},$$

where $a_1, c_1, \cdots, a_n, c_n, b \in \mathbb{R}_{++}$.

Proof: Define $\Phi(\lambda) = \inf_{x\in\mathbb{R}^n_{++}} L(x;\lambda)$. Fix $\tilde{\lambda} \in \mathbb{R}_{++}$. Then

$$L(x;\tilde{\lambda}) = \sum_{i=1}^n \frac{c_i}{x_i} + \tilde{\lambda}(\sum_{i=1}^n a_i x_i - b) = -\tilde{\lambda}b + \sum_{i=1}^n \frac{c_i}{x_i} + \tilde{\lambda}a_i x_i \geq -\tilde{\lambda}b + 2\sum_{i=1}^n \sqrt{\tilde{\lambda}c_i a_i}.$$

Note that equality happens if $x_i = \frac{c_i}{\tilde{\lambda}a_i}$ and $x_i = +\infty$ if $\tilde{\lambda} = 0$. Thus, $\Phi(\tilde{\lambda}) = -\tilde{\lambda}b + 2\sum_{i=1}^n \sqrt{\tilde{\lambda}c_i a_i}$. Note that $\Phi(\lambda^2)$ is a concave function in λ. Thus,

$$\Phi(\lambda^2) = -\lambda^2 b + 2\lambda\sum_{i=1}^n \sqrt{c_i a_i} \Rightarrow \sup_{\lambda\in\mathbb{R}^n_+} \Phi(\lambda^2) = \sup_{\lambda\in\mathbb{R}^n_+} -\lambda^2 b + 2\lambda\sum_{i=1}^n \sqrt{c_i a_i}.$$

However, the supremum happens at $\lambda^* = \frac{\sum\sqrt{c_i a_i}}{b}$. Thus $d = \frac{(\sum\sqrt{c_i a_i})^2}{b}$.

4.3.3 (Slater and compactness)

Prove the Slater condition holds for problem

$$\inf\{f(x) : g(x) \leq 0, x \in \mathbb{E}\},$$

if and only if there exists $\hat{x} \in \mathbb{E}$ for which the level sets

$$\{\lambda \in \mathbb{R}^n_+ : -L(\hat{x};\lambda) \leq \alpha\},$$

is compact for all $\alpha \in \mathbb{R}$.

Proof: Suppose there exists a Slater point, then $-\lambda^T g(\hat{x}) \leq \alpha$ has compact level sets for all $\alpha \in \mathbb{R}$. In fact, for each $i = 1, \cdots, m$, we have $-\lambda_i g_i(\hat{x}) \leq \alpha$. Thus, $\lambda_i \leq \frac{\alpha}{-g_i(\hat{x})}$. Thus, λ is bounded above.

Conversely, suppose $\{\lambda \in \mathbb{R}^n_+ : -L(\hat{x};\lambda) \le \alpha\}$, is compact for all $\alpha \in \mathbb{R}$. Then, if some i, $g_i(\hat{x}) \ge 0$, then if $\mu \in \{\lambda \in \mathbb{R}^n_+ : -L(\hat{x};\lambda) \le \alpha\}$, then so is $\mu + te_i$ for all $t \ge 0$. Thus, $\{\lambda \in \mathbb{R}^n_+ : -L(\hat{x};\lambda) \le \alpha\}$ is empty for all real α, which is a contradiction.

4.3.4 (Examples of duals)
Calculate the Lagrangian duals for the following problem:

- The *linear program*

$$\inf_{x\in\mathbb{R}^n}\{\langle c,x\rangle : \langle a^i,x\rangle \le b_i \text{ for } i=1,\cdots,m\}.$$

Proof: Let $A=[a^1|\cdots|a^m]$, then $\langle a^i,x\rangle \le b_i$ translates into $A^Tx \le b$. We have

$$\langle c,x\rangle + \lambda^T(A^Tx-b) = \langle c+A\lambda, x\rangle - \lambda^T b.$$

Thus the dual problem is as follows:

$$\sup_{A\lambda+c=0,\lambda\in\mathbb{R}^m_+} -\lambda^T b.$$

- Another linear program

$$\inf_{x\in\mathbb{R}^n}\{\langle c,x\rangle + \delta_{\mathbb{R}^n_+(x)} : \langle a^i,x\rangle \le b_i \text{ for } i=1,\cdots,m\}.$$

Proof: Again

$$\langle c,x\rangle + \lambda^T(A^Tx-b) = \langle c+A\lambda, x\rangle - \lambda^T b.$$

Thus, the dual problem is a follows:

$$\sup_{A\lambda+c\ge 0,\lambda\in\mathbb{R}^m_+} -\lambda^T b.$$

- The *quadratic program* for some $C \in \mathbb{S}^n_{++}$

$$\inf_{x\in\mathbb{R}^n}\{\frac{1}{2}(x^TCx) : \langle a^i,x\rangle \le b_i \text{ for } i=1,\cdots,m\}.$$

Proof: We have

$$\frac{1}{2}(x^TCx) + \lambda^T(A^Tx-b) \text{ is strictly convex w.r.t } x,$$

thus the dual function equals to (for $Cx^* + A\lambda = 0$)

$$\inf_{x\in\mathbb{E}}\Phi(\lambda) = -\frac{1}{2}x^{*T}A\lambda + \lambda^TA^Tx^* - \lambda^Tb = \frac{1}{2}x^{*T}A\lambda - \lambda^Tb = -[\frac{1}{2}(C^{-1}A\lambda)^TA\lambda + \lambda^Tb],$$

Hence,

$$\sup_{\lambda\ge 0} -[\frac{1}{2}(C^{-1}A\lambda)^TA\lambda + \lambda^Tb] = -\inf_{\lambda\ge 0}\frac{1}{2}(\lambda^TA^TC^{-T}A\lambda) + \lambda^Tb = \frac{1}{2}b^TA^{-1}CA^{-T}b.$$

- The *separable* problem

$$\inf_{x\in\mathbb{R}^n}\{\sum_{j=1}^{n} p(x_j) : \langle a^i, x\rangle \le b_i \text{ for } i = 1, \cdots, m\}$$

for a given function $p : \mathbb{R} \to (\infty, +\infty]$.

Proof: $\Phi(\lambda) = \inf_{x\in\mathbb{E}} \sum_{j=1}^{n}(p(x_j) + \lambda_j(\langle a^j, x\rangle - b_j))$

4.3.7
Given a matrix C in $\mathbb{S}^n_{++}$, calculate

$$\inf_{X\in\mathbb{S}^n_{++}} \{\mathrm{Tr}(CX) : -\log\det(X) \le 0\}$$

by the Lagrangian duality.
Proof: Suppose that $b < n$ and consider the problem

$$\inf_{X\in\mathbb{S}^n_{++}} \{\mathrm{Tr}(CX) : -\log\det(X) \le b\}$$

Let $\lambda \in \mathbb{R}_+$. Then the Lagrangian equals to

$$\mathrm{Tr}(CX) - \lambda\log\det(X) - \lambda b \Rightarrow \Phi(\lambda) = -\lambda b + \inf_{X\in\mathbb{S}^n_{++}} \mathrm{Tr}(CX) - \lambda\log\det(X).$$

Note that $\nabla^2(\mathrm{Tr}(CX)-\lambda\log\det(X)) = \lambda X^{-2} \succeq 0$. Thus, since $\nabla(\mathrm{Tr}(CX)-\lambda\log\det(X)) = C - \lambda X^{-1}$, we have $\Phi(\lambda) = -\lambda b + n\lambda + \lambda\log\det(C) - n\lambda\log\lambda$ for $\lambda > 0$ and $\Phi(0) = 0$.
Now note that $(n\lambda + \lambda\log\det(C) - n\lambda\log\lambda)'' = -n/\lambda < 0$ for $\lambda \in \mathbb{R}_{++}$. Now notice $(-\lambda b + n\lambda + \lambda\log\det(C) - n\lambda\log\lambda)' = -b + \log\det(C) - n\log\lambda$, thus, $\lambda^* = \sqrt[1/n]{e^{-b}\det(C)}$. Notice that $\sup_{\lambda\ge 0}\Phi(\lambda) = \Phi(\lambda^*) = n\lambda^* - \lambda^* b > 0$.

4.3.8. Mixed constraints
Explain why an appropriate dual for the problem

$$\inf\{f(x) : g(x) \le 0, h(x) = 0\}$$

for a function $h : \mathrm{dom}\, f \to \mathbb{R}^k$ is

$$\sup_{\lambda\in\mathbb{R}^m_+,\mu\in\mathbb{R}^k} \inf_{x\in\mathrm{dom}\, f}\{f(x) + \lambda^T g(x) + \mu^T h(x)\}.$$

Proof: Come on!

4.3.9. Fenchel and Lagrangian duality
Let Y be a Euclidean space. By suitably rewriting the problem Fenchel problem

$$\inf_{x\in\mathbb{E}}\{f(x) + g(Ax)\}$$

for given function $f : \mathbb{E} \to (\infty, +\infty]$, $g : Y \to (\infty, +\infty]$ and linear map $A : \mathbb{E} \to Y$, interpret the dual Fenchel problem

$$\sup_{\phi \in Y} \{-f^*(A^*\phi) - g^*(-\phi)\}$$

as a Lagrangian dual problem.

Proof: Note that

$$\inf_{x \in \mathbb{E}} \{f(x) + g(Ax)\} = \inf_{(x,y) \in \mathbb{E}^2} \{f(x) + g(y) : Ax = y\}.$$

Thus, $L(x, y, \phi) = f(x) + g(y) + \langle \phi, Ax - y \rangle$ with $L : \mathbb{E}^2 \times \mathbb{R} \to [-\infty, +\infty]$. Then

$$\Phi(\phi) = \inf_{(x,y)} f(x) + g(y) + \langle \phi, Ax - y \rangle = -\sup_{x} [\langle -A^*\phi, x \rangle - f(x)] - \sup_{y \in \mathbb{E}} [\langle \phi, y \rangle - g(y)]$$

which equals to $-f^*(-A^*\phi) - g^*(\phi)$. This completes the proof.