The Hahn–Banach–Lagrange theorem

by

Stephen Simons

<simons@math.ucsb.edu>.

Abstract

We discuss the Hahn–Banach–Lagrange theorem, a generalized form of the Hahn–Banach theorem. As applications, we derive various results on the existence of linear functionals in functional analysis, on the existence of Lagrange multipliers for convex optimization problems, with an explicit sharp lower bound on the norm of the solutions (multipliers), on finite families of convex functions (leading rapidly to a minimax theorem), on the existence of subgradients of convex functions, and on the Fenchel conjugate of a convex function. We give a complete proof of Rockafellar's version of the Fenchel duality theorem, and an explicit sharp lower bound for the norm of the solutions of the Fenchel duality theorem in terms of elementary geometric concepts.

Downloads

You can download files containing these slides and a related paper from
<www.math.ucsb.edu/~simons/WCOM.html>.

Sublinear functionals

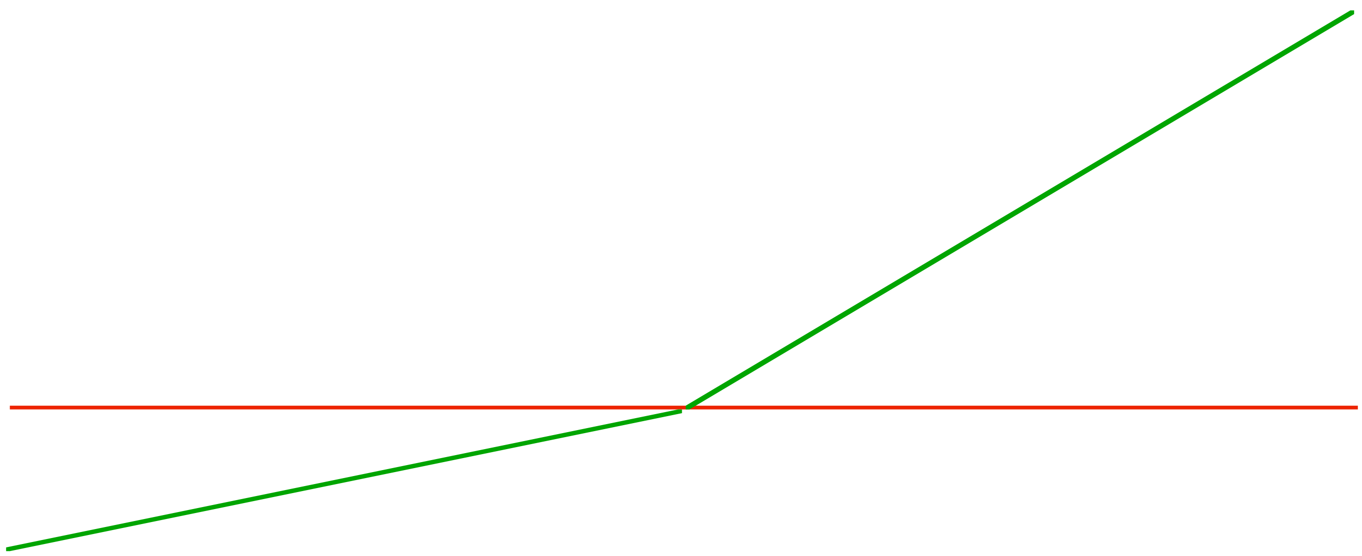
Let E be a nonzero real vector space[†]. A **sublinear functional** on E is a map $P\colon\ E \to \mathbb{R}$ such that

$$P \text{ is subadditive:} \qquad x, y \in E \quad \Longrightarrow \quad P(x+y) \le P(x) + P(y)$$

and

$$P \text{ is positively homogeneous:} \qquad x \in E \text{ and } \lambda > 0 \quad \Longrightarrow \quad P(\lambda x) = \lambda P(x).$$

Sublinear functional



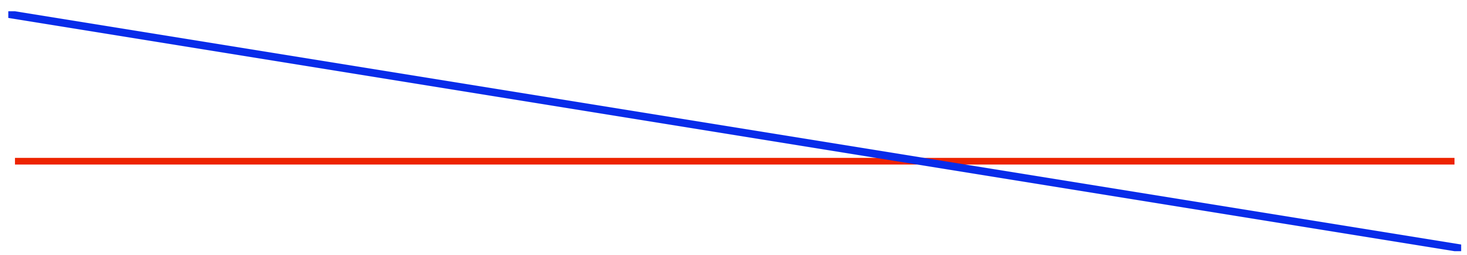
- Norms and linear functionals are sublinear.

Affine functions

Let D be a nonempty convex subset of a vector space, E be a vector space and $a\colon D \to E$. a is **affine** if

$$x, y \in D \text{ and } \lambda \in]0,1[\quad \Longrightarrow \quad a\big(\lambda x + (1-\lambda)y\big) = \lambda a(x) + (1-\lambda)a(y).$$

Affine function



- Note that an affine function can map into a vector space.

Convex functions

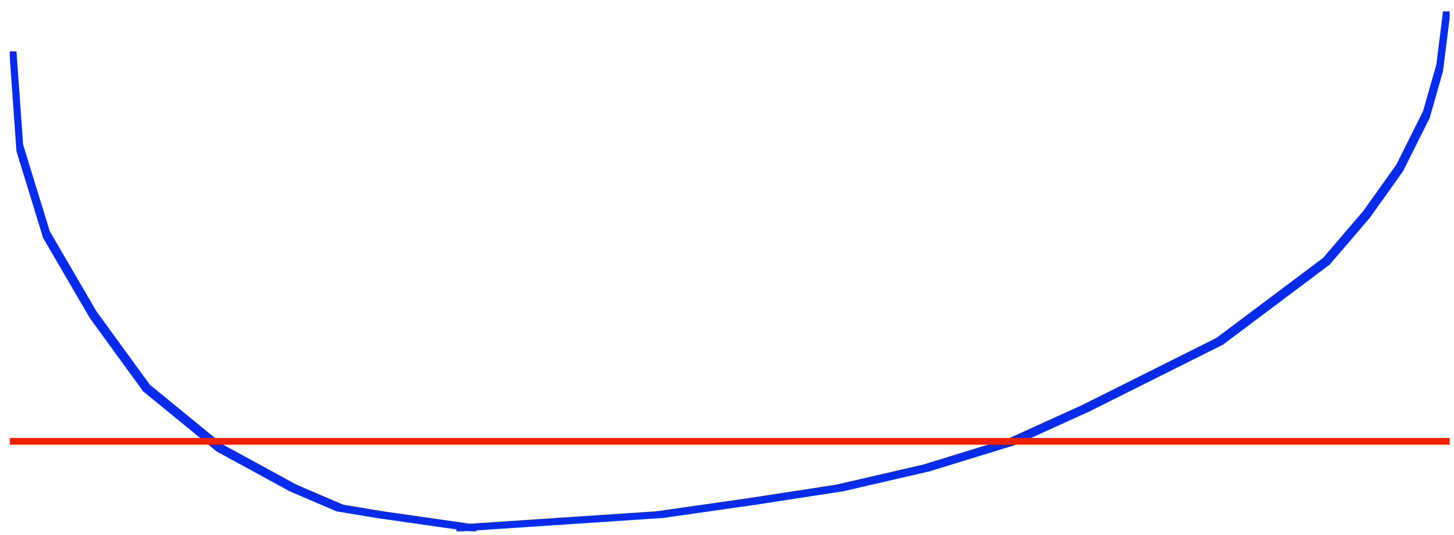
Let C be a nonempty convex subset of a vector space, and $f\colon C \to]-\infty, \infty]$. f is **convex** if

$$x, y \in C \text{ and } \lambda \in]0,1[\quad \Longrightarrow \quad f\big(\lambda x + (1-\lambda)y\big) \le \lambda f(x) + (1-\lambda)f(y),$$

provided $\infty + \infty := \infty$, and $\lambda \times \infty := \infty$ for $\lambda > 0$. f is proper if

$$\exists\, x \in C \text{ such that } f(x) \in \mathbb{R}.$$

Convex function



- Sublinear functionals are convex.

Sublinear functionals

Let E be a nonzero real vector space[†]. A **sublinear functional** on E is a map $P\colon\ E \to \mathbb{R}$ such that

$$P \text{ is subadditive:} \qquad x, y \in E \quad\Longrightarrow\quad P(x+y) \le P(x) + P(y)$$

and

$$P \text{ is positively homogeneous:} \qquad x \in E \text{ and } \lambda > 0 \quad\Longrightarrow\quad P(\lambda x) = \lambda P(x).$$

The Hahn-Banach theorem

Let P be a sublinear functional on E. Then $\exists$ a linear functional L on E such that[†]

$$L \le P \text{ on } E.$$

The Mazur-Orlicz theorem

Let P be a sublinear functional on E and D be a nonempty convex subset of E. Then $\exists$ a linear functional L on E such that

$$L \le P \text{ on } E \quad \text{and} \quad \inf_D L = \inf_D P.$$

The Mazur-Orlicz theorem

Let P be a sublinear functional on E and D be a nonempty convex subset of E. Then $\exists$ a linear functional L on E such that

$$L \le P \text{ on } E \quad \text{and} \quad \inf_D L = \inf_D P.$$

Proof Let $\beta := \inf_D P$. If $\beta = -\infty$, the result is immediate from the sublinear version of the Hahn-Banach theorem (take any linear functional L on E such that $L \le P$ on E). So we can suppose that $\beta \in \mathbb{R}$. Define $Q\colon\ E \to [-\infty, \infty[$ by

$$Q(x) := \inf_{d \in D,\ \lambda > 0}\big[P\big(x + \lambda d\big) - \lambda\beta\big].$$

Then $Q\colon\ E \to \mathbb{R}$ and Q is sublinear. If $x \in E$ and $d \in D$ then, $\forall\ \lambda > 0$,

$$Q(x) \le P(x + \lambda d) - \lambda\beta \le P(x) + P(\lambda d) - \lambda\beta \le P(x) + \lambda P(d) - \lambda\beta.$$

If we let $\lambda \to 0$, we get $Q(x) \le P(x)$, i.e., $Q \le P$ on E. We now apply the sublinear version of the Hahn-Banach theorem to Q, and obtain a linear functional L on E such that $L \le Q$ on E. Obviously, $L \le P$ on E. If $d \in D$ then

$$-L(d) = L(-d) \le Q(-d) \le P(-d + 1d) - 1\beta = -\beta.$$

Thus $L(d) \ge \beta = \inf_D P$. Consequently, $\inf_D L \ge \beta$. On the other hand, since $L \le P$ on E, $\inf_D L \le \inf_D P$. Combining these inequalities:

$$\inf_D L = \inf_D P. \qquad \blacksquare$$

• The technique used above is called the technique of the "auxiliary sublinear functional".

The Mazur-Orlicz theorem

Let P be a sublinear functional on E and D be a nonempty convex subset of E. Then $\exists$ a linear functional L on E such that

$$L \leq P \text{ on } E \quad \text{and} \quad \inf_D L = \inf_D P.$$

• If E is a normed space, E^* stands for the norm–dual of E.

A separation theorem ("bipolar theorem")

Let C be a nonempty convex subset of a normed space E and $x \in E \setminus \overline{C}$. Then $\exists$ $z^* \in E^*$ such that

$$\sup_C z^* < \langle x, z^* \rangle.$$

Proof We know that $\delta := \inf_{y \in C} \|x - y\| > 0$. Let $P := \|\cdot\|$ and $D := x - C$, so that

$$\inf_D P = \delta.$$

From the MOt, there exists a linear functional L on E such that

$$L \leq \|\cdot\| \text{ on } E \quad \text{and} \quad \inf_{y \in C} L(x - y) = \delta.$$

Since L is linear, this implies that

$$L(x) = \sup_{y \in C} L(y) + \delta > \sup_{y \in C} L(y).$$

Since $L \leq \|\cdot\|$ on E, $L \in E^*$ (in fact, $\|L\| \leq 1$). ▮

The Mazur-Orlicz theorem

Let P be a sublinear functional on E and D be a nonempty convex subset of E. Then $\exists$ a linear functional L on E such that

$$L \le P \text{ on } E \quad \text{and} \quad \inf_D L = \inf_D P.$$

We will prove:

The Hahn–Banach–Lagrange theorem

Let P be a sublinear functional on E. Let C be a nonempty convex subset of a (possibly different) vector space, $k\colon\ C \to\,]-\infty, \infty]$ be proper and convex and $j\colon\ C \to E$ be P–convex. Then $\exists$ a linear functional L on E such that

$$L \le P \text{ on } E \quad \text{and} \quad \inf_C \big[L \circ j + k\big] = \inf_C \big[P \circ j + k\big].$$

- "j is P–convex" means that

$x_1, x_2 \in C,\ \mu_1, \mu_2 > 0$ and $\mu_1 + \mu_2 = 1 \quad \Longrightarrow \quad j(\mu_1 x_1 + \mu_2 x_2) \le_P \mu_1 j(x_1) + \mu_2 j(x_2)$,

where the ordering "$\le_P$" on E is defined by

$$y \le_P z \iff P(y - z) \le 0.$$

The Hahn–Banach–Lagrange theorem

Let P be a sublinear functional on E. Let C be a nonempty convex subset of a (possibly different) vector space, $k\colon C \to]-\infty, \infty]$ be proper and convex and $j\colon C \to E$ be P–convex. Then $\exists$ a linear functional L on E such that

$$L \le P \text{ on } E \quad \text{and} \quad \inf_C \big[L \circ j + k\big] = \inf_C \big[P \circ j + k\big].$$

Picture :

$$\begin{array}{ccc} & C & \\ \swarrow j & & k \searrow \\ E & \xrightarrow{P,\,L} &]-\infty, \infty]. \end{array}$$

Proof This follows from the MOt with E replaced by $E \times \mathbb{R}$, the sublinear functional defined on $E \times \mathbb{R}$ by $(y, \lambda) \mapsto P(y) + \lambda$, and the convex set D defined by

$$D := \bigcup_{x \in C} \big\{(y, \lambda) \in E \times \mathbb{R}\colon\ P\big(j(x) - y\big) \le 0,\ k(x) \le \lambda\big\}$$

$$= \bigcup_{x \in C} \big\{(y, \lambda) \in E \times \mathbb{R}\colon\ j(x) \le_P y,\ k(x) \le \lambda\big\}. \qquad \blacksquare$$

• Affine functions are P–convex, and this result is frequently (but not always) applied with j affine.

The Hahn–Banach–Lagrange theorem

Let P be a sublinear functional on E. Let C be a nonempty convex subset of a (possibly different) vector space, $k\colon C \to \left]-\infty, \infty\right]$ be proper and convex and $j\colon C \to E$ be P–convex. Then $\exists$ a linear functional L on E such that

$$L \le P \text{ on } E \quad \text{and} \quad \inf_C \left[L \circ j + k\right] = \inf_C \left[P \circ j + k\right].$$

Sandwich theorem

Let P be a sublinear functional on E and $k\colon E \to \left]-\infty, \infty\right]$ be proper and convex and $-k \le P$ on E. Then $\exists$ a linear functional L on E such that

$$-k \le L \le P \text{ on } E.$$

Proof We know that $\inf_E \left[P + k\right] \ge 0$. Let $C := E$ and $j(x) := x$. From the HBLt, there exists a linear functional L on E such that

$$L \le P \text{ on } E \quad \text{and} \quad \inf_E \left[L + k\right] \ge 0.$$

But this last inequality implies that $-k \le L$ on E. ∎

The Hahn–Banach–Lagrange theorem

Let P be a sublinear functional on E. Let C be a nonempty convex subset of a (possibly different) vector space, $k\colon C \to \,]-\infty, \infty]$ be proper and convex and $j\colon C \to E$ be P–convex. Then $\exists$ a linear functional L on E such that

$$L \le P \text{ on } E \quad \text{and} \quad \inf_C \big[L \circ j + k\big] = \inf_C \big[P \circ j + k\big].$$

The extension form of the Hahn-Banach theorem

Let E be a normed space, F be a subspace of E and $y^* \in F^*$. Then $\exists\, x^* \in E^*$ such that

$$x^*|_F = y^* \quad \text{and} \quad \|x^*\|_E \le \|y^*\|_F.$$

Proof It follows from the definition of $\|y^*\|_F$ that, for all $y \in F$, $\langle y, y^*\rangle \le \|y\|\|y^*\|_F$. Let $P := \|y^*\|_F\|\cdot\|$ and $k(y) := -\langle y, y^*\rangle$. So we have $\inf_F \big[P + k\big] \ge 0$. Let $C := F$ and $j(y) := y$. From the HBLt, there exists a linear functional L on E such that

$$x \in E \quad \Longrightarrow \quad L(x) \le P(x) = \|y^*\|_F\|x\| \qquad (☾)$$

and

$$y \in F \quad \Longrightarrow \quad L(y) - \langle y, y^*\rangle \ge 0.$$

Since $L - y^*$ is linear on the subpace F, it follows from this that

$$y \in F \quad \Longrightarrow \quad L(y) - \langle y, y^*\rangle = 0,$$

in other words, $L|_F = y^*$. Furthermore, (☾) implies that $L \in E^*$ and $\|L\|_E \le \|y^*\|_F$. ▌

The Hahn–Banach–Lagrange theorem

Let P be a sublinear functional on E. Let C be a nonempty convex subset of a (possibly different) vector space, $k\colon C \to]-\infty, \infty]$ be proper and convex and $j\colon C \to E$ be P–convex. Then $\exists$ a linear functional L on E such that

$$L \le P \text{ on } E \quad \text{and} \quad \inf_C \big[L \circ j + k\big] = \inf_C \big[P \circ j + k\big].$$

Lemma on m convex functions

Let C be a nonempty convex subset of a vector space and $f_1, \ldots, f_m$ be convex real functions on C. Then $\exists\ \lambda_1, \ldots, \lambda_m \ge 0$ such that

$$\lambda_1 + \cdots + \lambda_m = 1 \quad \text{and} \quad \inf_C[f_1 \vee \cdots \vee f_m] = \inf_C[\lambda_1 f_1 + \cdots + \lambda_m f_m].$$

Proof This follows from the HBLt with[†] $E := \mathbb{R}^m$, $k := 0$, and P and j defined by

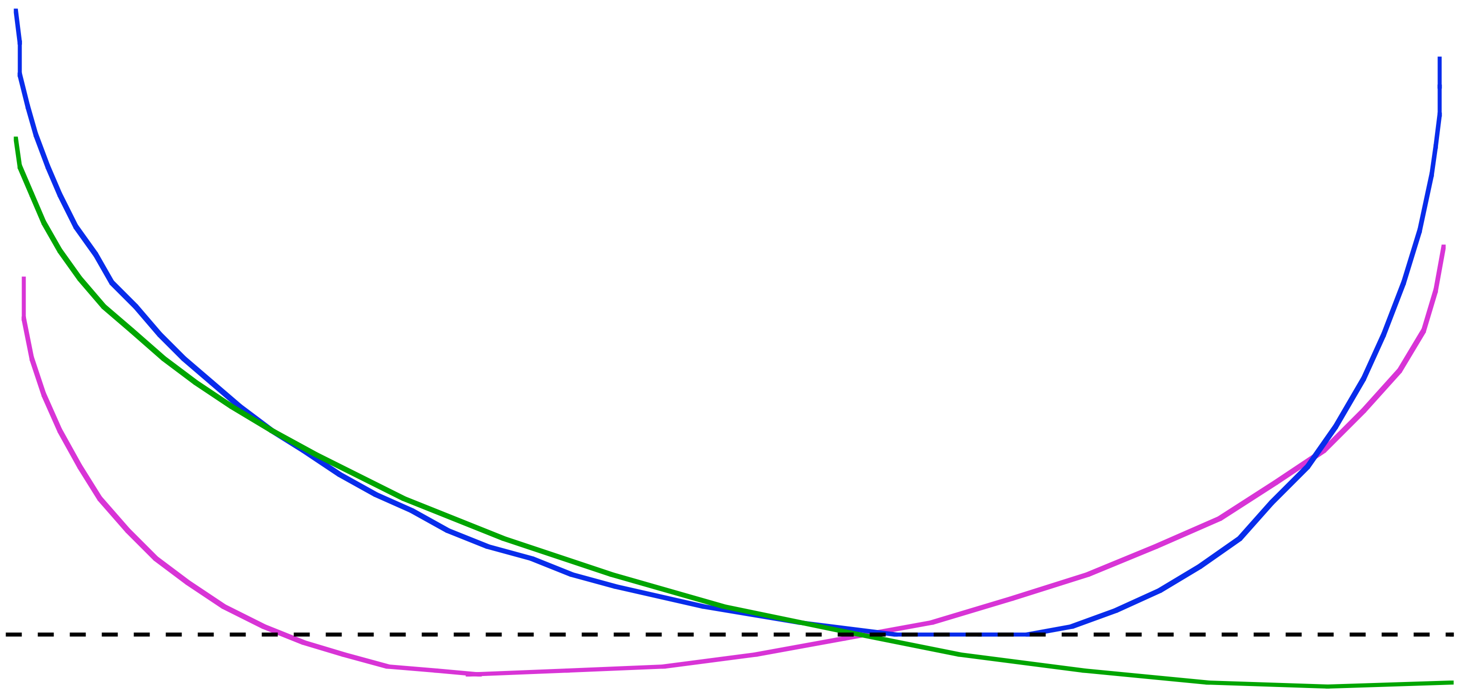
$$P(\mu_1, \ldots, \mu_m) := \mu_1 \vee \cdots \vee \mu_m \quad \text{and} \quad j(c) := \big(f_1(c), \ldots, f_m(c)\big).$$

▌

Lemma on m convex functions

Let C be a nonempty convex subset of a vector space and $f_1, \ldots, f_m$ *be convex real functions on C. Then* $\exists\ \lambda_1, \ldots, \lambda_m \ge 0$ *such that*

$$\lambda_1 + \cdots + \lambda_m = 1 \quad \textit{and} \quad \inf_C[f_1 \vee \cdots \vee f_m] = \inf_C[\lambda_1 f_1 + \cdots + \lambda_m f_m].$$



Let A, B be nonempty sets, and $h\colon A \times B \to \mathbb{R}$.

- It is easily seen that

$$\sup_{a \in A} \inf_{b \in B} h(a,b) \le \inf_{b \in B} \sup_{a \in A} h(a,b).$$

- This inequality can be strict, take for instance $A = B = \{0,1\}$ and $h(a,b) = 0$ if $a \neq b$ and $h(a,b) = 1$ if $a = b$.

The minimax theorem

Let A be a nonempty convex subset of a vector space, B be a nonempty convex subset of a vector space and B also be a compact space. Let $h\colon A \times B \to \mathbb{R}$ be concave on A, and convex and lower semicontinuous on B. Then

$$\sup_{a \in A} \min_{b \in B} h(a,b) = \min_{b \in B} \sup_{a \in A} h(a,b).$$

- h is "concave on A" means that

$$\forall b \in B, \quad -h(\cdot, b) \text{ is convex on } A.$$

- h is "convex and lower semicontinuous on B" mean that

$$\forall a \in A, \quad h(a, \cdot) \text{ is convex and lower semicontinuous on } B.$$

- Note that the set A has no topological structure.

- We can write "min" instead of "inf" because h is lower semicontinuous on B, and B is compact.

The minimax theorem

Let A be a nonempty convex subset of a vector space, B be a nonempty convex subset of a vector space and B also be a compact space. Let $h\colon\ A \times B \to \mathbb{R}$ be concave on A, and convex and lower semicontinuous on B. Then

$$\sup_{a\in A}\min_{b\in B} h(a,b) = \min_{b\in B}\sup_{a\in A} h(a,b).$$

Proof Let $\beta := \sup_{a\in A}\min_{b\in B} h(a,b)$. If we had $\beta < \min_{b\in B}\sup_{a\in A} h(a,b)$ then

$$\bigcup_{a\in A}\big\{b \in B\colon\ h(a,b) > \beta\big\} = B.$$

Since h is lower semicontinuous on B, the sets $\big\{b \in B\colon\ h(a,b) > \beta\big\}$ are open and B is compact, there would exist $a_1, \ldots, a_m \in A$ such that

$$\big\{b \in B\colon\ h(a_1,b) > \beta\big\} \cup \cdots \cup \big\{b \in B\colon\ h(a_m,b) > \beta\big\} = B$$

and so $\min_{b\in B}[h(a_1,b) \vee \cdots \vee h(a_m,b)] > \beta$. From the Lemma on m convex functions with $f_i := h(a_i,\cdot)$, there would exist $\lambda_1, \ldots, \lambda_m \geq 0$ such that $\lambda_1 + \cdots + \lambda_m = 1$ and

$$\min_{b\in B}[\lambda_1 h(a_1,b) + \cdots + \lambda_m h(a_m,b)] > \beta.$$

Since h is concave on A, it would follow from this that

$$\min_{b\in B} h(\lambda_1 a_1 + \cdots + \lambda_m a_m, b) > \beta,$$

which would contradict the definition of β. So

$$\beta \geq \min_{b\in B}\sup_{a\in A} h(a,b). \qquad \blacksquare$$

Finite dimensional Lagrange multipliers for a constrained infimum

- C is a nonempty convex subset of a vector space.
- $k\colon C \to]-\infty, \infty]$ is convex.
- $f_1, \ldots, f_m$ are convex real functions on C.
- $\inf\big\{k(x)\colon x \in C,\ f_1(x) \le 0, \ldots, f_m(x) \le 0\big\} = \mu_0 \in \mathbb{R}$.

When can we assert that

$\exists\, \lambda_1, \ldots, \lambda_m \ge 0$ such that $\inf\big\{\lambda_1 f_1(x) + \cdots + \lambda_m f_m(x) + k(x)\colon x \in C\big\} = \mu_0$? (👓)

A vector $(\lambda_1, \ldots, \lambda_m)$ satisfying (👓) is known as a Lagrange multiplier.
Now let $E := \mathbb{R}^m$ and $\preceq$ be the usual ordering on E. Define $j\colon C \to E$ and $z^* \in E^*$ by

$$j(x) := \big(f_1(x), \ldots, f_m(x)\big) \quad \text{and} \quad z^* = (\lambda_1, \ldots, \lambda_m).$$

j is $\preceq$–convex and z^* is $\preceq$–positive in the obvious senses.
So we can reformulate our hypothesis for the constrained infimum as:

- $\inf\big\{k(x)\colon x \in C, j(x) \preceq 0\big\} = \mu_0 \in \mathbb{R}$

and our question as:

when can we assert that

$\exists$ $\preceq$–positive $z^* \in E^*$ such that $\inf\big\{\langle j(x), z^*\rangle + k(x)\colon x \in C\big\} = \mu_0$? (👓)

— The Hahn–Banach–Lagrange theorem —

- E is a normed space and $\preceq$ is a vector ordering on E.
- C is a nonempty convex subset of a vector space.
- $k\colon C \to]-\infty, \infty]$ is convex and $j\colon C \to E$ is $\preceq$–convex.
- $\inf\big\{k(x)\colon x \in C,\ j(x) \preceq 0\big\} = \mu_0 \in \mathbb{R}$.
- Let $N := \{y \in E\colon\ y \preceq 0\}$ and $A := \big\{x \in C\colon\ k(x) < \mu_0\big\} \neq \emptyset$.

Lagrange multipliers for constrained convex problems

When can we assert that

$$\exists\ \preceq\text{–positive } z^* \in E^* \quad \text{such that} \quad \inf\big\{\langle j(x), z^*\rangle + k(x)\colon\ x \in C\big\} = \mu_0? \qquad (👓)$$

Classical Slater condition result: Let $B := \big\{x \in C\colon\ j(x) \in \operatorname{int} N\big\} \neq \emptyset$ then (👓).

Necessary condition with a bound on the norm

Suppose that $B \neq \emptyset$. *Then* (👓) *with*

$$\|z^*\| \le \inf_{v \in B} \frac{k(v) - \mu_0}{\operatorname{dist}\big(j(v), E \setminus N\big)}.$$

Necessary and sufficient condition with sharp bound on the norm

$$(👓) \iff \sup_{x \in A} \frac{\mu_0 - k(x)}{\operatorname{dist}\big(j(x), N\big)} < \infty.$$

Further,

$$\sup_{x \in A} \frac{\mu_0 - k(x)}{\operatorname{dist}\big(j(x), N\big)} = \min\big\{\|z^*\|\colon\ z^* \text{ satisfies } (👓)\big\}.$$

- E is a normed space and $\preceq$ is a vector ordering on E.
- C is a nonempty convex subset of a vector space.
- $k\colon C \to]-\infty, \infty]$ is convex and $j\colon C \to E$ is $\preceq$–convex.
- $\inf\{k(x)\colon x \in C,\ j(x) \preceq 0\} = \mu_0 \in \mathbb{R}$.
- Let $N := \{y \in E\colon y \preceq 0\}$ and $A := \{x \in C\colon k(x) < \mu_0\} \neq \emptyset$.

A property of Lagrange multipliers

Let z^* be a Lagrange multiplier, that is to say z^* is a $\preceq$–positive element of E^* and

$$\inf\{\langle j(x), z^*\rangle + k(x)\colon x \in C\} = \mu_0$$

Then

$$\sup_{x \in A} \frac{\mu_0 - k(x)}{\operatorname{dist}\big(j(x), N\big)} \le \|z^*\|.$$

Proof. Let x and y be arbitrary elements of A and N, respectively. Since z^* is $\preceq$–positive and $y \preceq 0$, using the definition of μ_0,

$$\|j(x) - y\|\,\|z^*\| \ge \langle j(x) - y, z^*\rangle = \langle j(x), z^*\rangle - \langle y, z^*\rangle \ge \langle j(x), z^*\rangle \ge \mu_0 - k(x) > 0.$$

Taking the infimum over $y \in N$,

$$\operatorname{dist}\big(j(x), N\big)\|z^*\| \ge \mu_0 - k(x) > 0.$$

Divide by $\operatorname{dist}\big(j(x), N\big)$ and take the supremum over $x \in A$. ∎

- E is a normed space and $\preceq$ is a vector ordering on E.
- C is a nonempty convex subset of a vector space.
- $k\colon C \to]-\infty, \infty]$ is convex and $j\colon C \to E$ is $\preceq$–convex.
- $\inf\big\{k(x)\colon x \in C,\ j(x) \preceq 0\big\} = \mu_0 \in \mathbb{R}$.
- Let $N := \{y \in E\colon y \preceq 0\}$ and $A := \big\{x \in C\colon k(x) < \mu_0\big\} \neq \emptyset$.

Lemma

Let $M > 0$, and define $P\colon E \to [0, \infty[$ by

$$P(y) := M\operatorname{dist}(y, N) = M \inf_{u \in N} \|y - u\| \qquad (y \in E).$$

Then P is sublinear, $P = 0$ on N and j is P–convex.

Proof. Let $y_1, y_2 \in E$ and $u_1, u_2 \in N$ be arbitrary. Then, since $u_1 + u_2 \in N$,

$$M\|y_1 - u_1\| + M\|y_2 - u_2\| \geq M\big\|y_1 + y_2 - (u_1 + u_2)\big\| \geq P(y_1 + y_2).$$

Take the infimum over u_1 and u_2, $P(y_1) + P(y_2) \geq P(y_1 + y_2)$. Thus P is subadditive. It is easy to see that P is positively homogeneous, so P is sublinear. It is obvious that $P = 0$ on N. Let $y, z \in E$. Then

$$y \preceq z \iff y - z \in N \implies P(y - z) = 0 \implies y \leq_P z.$$

Thus the $\preceq$–convexity of j implies that j is P–convex. ∎

— The Hahn–Banach–Lagrange theorem —

- E is a normed space and $\preceq$ is a vector ordering on E.
- C is a nonempty convex subset of a vector space.
- $k\colon C \to]-\infty,\infty]$ is convex and $j\colon C \to E$ is $\preceq$–convex.
- $\inf\big\{k(x)\colon\ x \in C,\ j(x) \preceq 0\big\} = \mu_0 \in \mathbb{R}$.
- Let $N := \{y \in E\colon\ y \preceq 0\}$ and $A := \big\{x \in C\colon\ k(x) < \mu_0\big\} \neq \emptyset$.

Lemma

Let $M > 0$, and define $P\colon\ E \to [0,\infty[$ by

$$P(y) := M\operatorname{dist}(y,N) = M\inf_{u\in N}\|y-u\| \qquad (y \in E).$$

Then P is sublinear, $P = 0$ on N and j is P–convex.

The Hahn–Banach–Lagrange theorem

Let P be a sublinear functional on E. Let C be a nonempty convex subset of a (possibly different) vector space, $k\colon\ C \to]-\infty,\infty]$ be proper and convex and $j\colon\ C \to E$ be P–convex. Then $\exists$ a linear functional L on E such that

$$L \le P \text{ on } E \quad \text{and} \quad \inf\nolimits_C \big[L\circ j + k\big] = \inf\nolimits_C \big[P\circ j + k\big].$$

— The Hahn–Banach–Lagrange theorem —

- E is a normed space and $\preceq$ is a vector ordering on E.
- C is a nonempty convex subset of a vector space.
- $k\colon C \to]-\infty,\infty]$ is convex and $j\colon C \to E$ is $\preceq$–convex.
- $\inf\big\{k(x)\colon x \in C,\ j(x) \preceq 0\big\} = \mu_0 \in \mathbb{R}$.
- Let $N := \{y \in E\colon y \preceq 0\}$ and $A := \big\{x \in C\colon k(x) < \mu_0\big\} \neq \emptyset$.

The existence of Lagrange multipliers

Suppose that $0 < M := \sup_{x\in A} \dfrac{\mu_0 - k(x)}{\operatorname{dist}\big(j(x), N\big)} < \infty.$ *Then* $\exists$ *a Lagrange multiplier* z_0^* *such that* $\|z_0^*\| \le M.$

Proof. Define $P\colon E \to [0,\infty[$ by $P(y) := M\operatorname{dist}(y,N)$ $(y \in E)$. Then, for all $x \in A$, $P\big(j(x)\big) = M\operatorname{dist}\big(j(x),N\big) \ge \mu_0 - k(x)$, thus $\inf_A[P\circ j + k] \ge \mu_0$. Since $k \ge \mu_0$ on $C \setminus A$ and $P \ge 0$ on E, in fact

$$\inf_C[P\circ j + k] \ge \mu_0.$$

From the Lemma, P is sublinear and j is P–convex, so the HBLt gives a linear functional L on E such that $L \le P$ on E and $\inf_C[L\circ j + k] = \inf_C[P\circ j + k]$. Thus

$$\inf_C[L\circ j + k] \ge \mu_0.$$

Since $P \le M\|\cdot\|$ on E and $P = 0$ on N, $L \in E^*$, $\|L\| \le M$ and $L \le 0$ on N. So

$$\mu_0 = \inf_{j^{-1}N} k \ge \inf_{j^{-1}N}[L\circ j + k] \ge \inf_C[L\circ j + k].$$

Combining these: $\inf_C[L\circ j + k] = \mu_0$. Take $z_0^* = L$. ∎

A property of Lagrange multipliers

Let z^* *be a Lagrange multiplier, that is to say* z^* *is a* $\preceq$*–positive element of* E^* *and*

$$\inf\big\{\langle j(x), z^*\rangle + k(x)\colon\ x \in C\big\} = \mu_0$$

Then $$\sup_{x\in A} \frac{\mu_0 - k(x)}{\operatorname{dist}\big(j(x), N\big)} \le \|z^*\|.$$

The existence of Lagrange multipliers

Suppose that $0 < M := \sup_{x\in A} \dfrac{\mu_0 - k(x)}{\operatorname{dist}\big(j(x), N\big)} < \infty.$ *Then* $\exists$ *a Lagrange multiplier* z_0^* *such that* $\|z_0^*\| \le M$.

Combine:

Lagrange multipliers for constrained convex problems

$$\exists \preceq\text{–positive } z^* \in E^* \text{ such that } \inf\big\{\langle j(x), z^*\rangle + k(x)\colon\ x \in C\big\} = \mu_0 \qquad (👓)$$

$$\Updownarrow$$

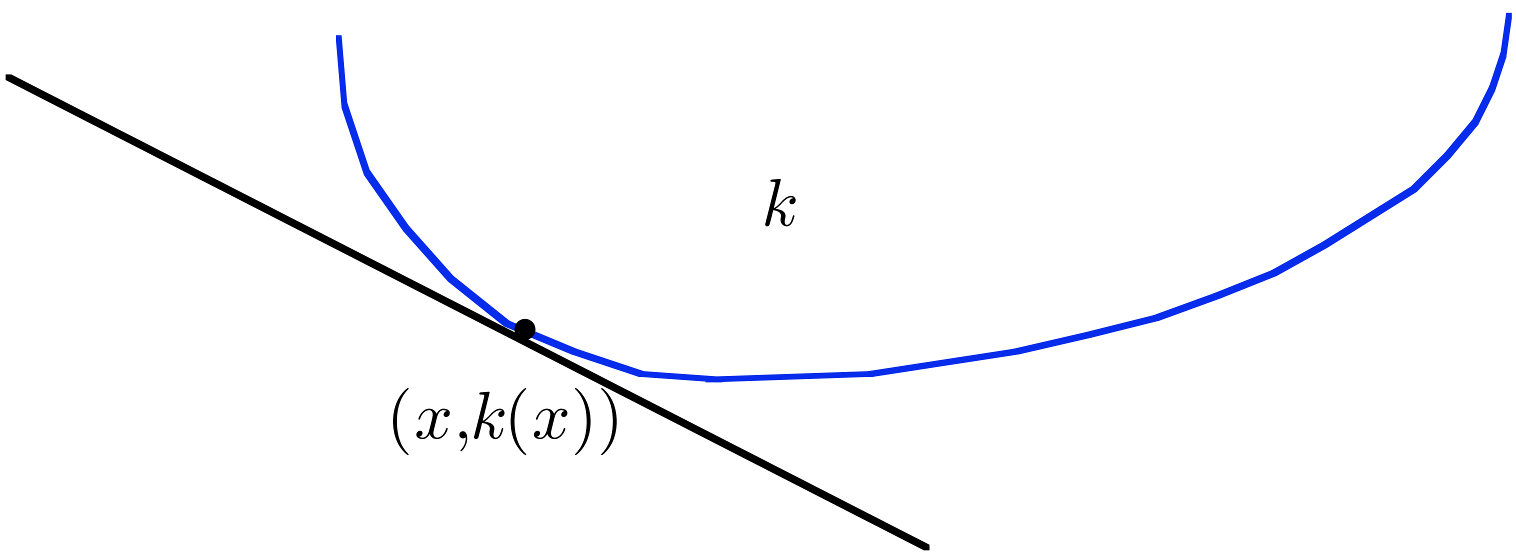
$$\sup_{x\in A} \frac{\mu_0 - k(x)}{\operatorname{dist}\big(j(x), N\big)} < \infty.$$

Further, $$\sup_{x\in A} \frac{\mu_0 - k(x)}{\operatorname{dist}\big(j(x), N\big)} = \min\big\{\|z^*\|\colon\ z^* \text{ satisfies } (👓)\big\}.$$

On the existence of subgradients

Let E be a normed space, $k\colon E \to]-\infty, \infty]$ be convex, $x \in E$ and $k(x) \in \mathbb{R}$. Does there exist $x^* \in E^*$ such that

$$y \in E \quad \Longrightarrow \quad k(x) + \langle y - x, x^* \rangle \le k(y)?$$



The Hahn–Banach–Lagrange theorem

Let P be a sublinear functional on E. Let C be a nonempty convex subset of a (possibly different) vector space, $k\colon C \to \,]-\infty,\infty]$ be proper and convex and $j\colon C \to E$ be P–convex. Then $\exists$ a linear functional L on E such that

$$L \le P \text{ on } E \quad \text{and} \quad \inf_C \big[L \circ j + k\big] = \inf_C \big[P \circ j + k\big].$$

On the existence of subgradients

Let E be a normed space, $k\colon E \to \,]-\infty,\infty]$ be convex, $x \in E$ and $k(x) \in \mathbb{R}$. Does there exist $x^* \in E^*$ such that $y \in E \implies k(x) + \langle y - x, x^*\rangle \le k(y)$?

$$\iff$$

Do there exist $M \ge 0$ and a linear functional L on E such that $L \le M\|\cdot\|$ on E and

$$y \in E \implies k(y) + L(x-y) \ge k(x)?$$

From the HBLt with $P := M\|\cdot\|$, $C := E$ and $j(y) := x - y$, this $\iff$

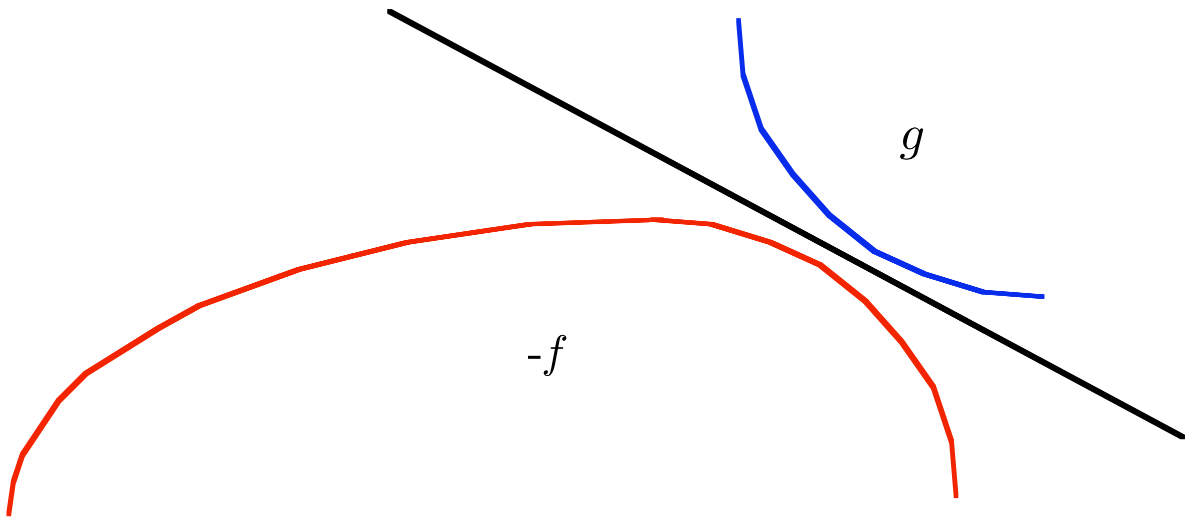
Does there exist $M \ge 0$ such that $y \in E \implies k(y) + M\|x-y\| \ge k(x)$?

Thus we have transformed the original problem on the existence of continuous linear functionals into the (much simpler) problem of finding a real constant M. This is an example of the "discovery method".

Separating a convex and a concave function

Let E be a normed space and $f, g\colon\ E \to]-\infty, \infty]$ be proper and convex. Do there exist $z^* \in E^*$ and $\beta \in \mathbb{R}$ such that

$$-f \le z^* + \beta \le g \quad \text{on} \quad E?$$



Using the same technique as before, with $C := E \times E$, $j(x, y) := x - y$ and $k(x, y) := f(x) + g(y)$, the above problem reduces to:

Does there exist $M \ge 0$ such that

$$\forall\ x,\ y \in E, \quad f(x) + g(y) + M\|x - y\| \ge 0?$$

Separating a convex and a concave function

Let E be a normed space and $f, g\colon\ E \to]-\infty, \infty]$ be proper and convex. Do there exist $z^* \in E^*$ and $\beta \in \mathbb{R}$ such that

$$-f \le z^* + \beta \le g \quad \text{on} \quad E? \qquad (\ )$$

• The **Fenchel conjugate**, f^*, of f is defined by $f^*(x^*) := \sup_E (x^* - f)$. f^* is also known as the **Legendre transform** of f.

•
$$(\ ) \iff -z^* - f \le \beta \text{ on } E \text{ and } z^* - g \le -\beta \text{ on } E$$
$$\iff f^*(-z^*) \le \beta \text{ and } g^*(z^*) \le -\beta,$$

• So our question $\iff$ is it true that

$$\exists\ z^* \in E^* \quad \text{such that} \quad f^*(-z^*) + g^*(z^*) \le 0? \qquad (\ )$$

When ( ) holds, we say that the Fenchel duality theorem is true.

• Rockafellar and Attouch–Brezis have given sufficient conditions for the Fenchel duality theorem to be true. The condition ( ) on the previous slide is both necessary and sufficient.

• We will use ( ) to prove Rockafellar's version in the following form: that ( ) is true if $f + g \ge 0$ on E and g is bounded above in a neighborhood of a point $z \in E$ at which $f(z) \in \mathbb{R}$.

- Let E be a normed space and $f, g\colon\ E \to \,]-\infty, \infty]$. If $w \in E$, let
$$(f \ominus g)(w) := \inf\nolimits_{z\in E}\big[f(z) + g(z-w)\big].$$

The $\ominus$ lemma

Let E be a normed space, $f, g\colon\ E \to \,]-\infty, \infty]$ be convex and $f + g \ge 0$ on E. Suppose that $\varepsilon > 0$, $N \ge 0$ and $\|w\| < \varepsilon \implies (f \ominus g)(w) < N$. Then $\exists\ M \ge 0$ such that
$$\forall\ x,\ y \in E, \quad f(x) + g(y) + M\|x - y\| \ge 0. \qquad (🌵)$$

Proof. Let $M = N/\varepsilon \ge 0$. Let $x, y \in E$. Let $\lambda > \|x - y\| \ge 0$ and $w := \varepsilon(y - x)/\lambda \in F$. Since $\|w\| < \varepsilon$, $(f \ominus g)(w) < N$ and so $\exists\ z \in E$ such that $f(z) + g(z - w) < N$. Now $\varepsilon x + \lambda z = \varepsilon y + \lambda(z - w)$, and so
$$\frac{\lambda z + \varepsilon x}{\lambda + \varepsilon} = \frac{\lambda(z - w) + \varepsilon y}{\lambda + \varepsilon}.$$
Thus
$$0 \le f\Big(\frac{\lambda z + \varepsilon x}{\lambda + \varepsilon}\Big) + g\Big(\frac{\lambda(z - w) + \varepsilon y}{\lambda + \varepsilon}\Big) \le \frac{\lambda f(z) + \varepsilon f(x) + \lambda g(z - w) + \varepsilon g(y)}{\lambda + \varepsilon}.$$
Consequently,
$$\lambda f(z) + \varepsilon f(x) + \lambda g(z - w) + \varepsilon g(y) \ge 0,$$
from which $\varepsilon\big[f(x) + g(y)\big] + \lambda N \ge 0$, and so
$$f(x) + g(y) + \lambda M \ge 0.$$
(🌵) now follows by letting $\lambda \to \|x - y\|$. ∎

- Let E be a normed space and $f, g\colon\ E \to]-\infty, \infty]$. If $w \in E$, let
$$(f \ominus g)(w) := \inf_{z \in E} \big[f(z) + g(z - w)\big].$$

The $\ominus$ lemma

Let E be a normed space, $f, g\colon\ E \to]-\infty, \infty]$ be convex and $f + g \geq 0$ on E. Suppose that $\varepsilon > 0$, $N \geq 0$ and $\|w\| < \varepsilon \implies (f \ominus g)(w) < N$. Then $\exists\ M \geq 0$ such that
$$\forall\ x,\ y \in E, \quad f(x) + g(y) + M\|x - y\| \geq 0. \qquad (🌵)$$

Rockafellar's version of the Fenchel duality theorem

Let E be a normed space, $f, g\colon\ E \to]-\infty, \infty]$ be convex, $f + g \geq 0$ on E, $z \in E$ be such that $f(z) \in \mathbb{R}$, and $\varepsilon > 0$ and $K \in \mathbb{R}$ be such that
$$\|w\| < \varepsilon \implies g(z - w) < K.$$
Then
$$\exists\ z^* \in E^* \quad \text{such that} \quad f^*(-z^*) + g^*(z^*) \leq 0. \qquad (🐕)$$

Proof. Let
$$N := f(z) + K > f(z) + g(z) = (f + g)(z) \geq 0.$$
If $w \in E$ and $\|w\| < \varepsilon$ then
$$(f \ominus g)(w) \leq f(z) + g(z - w) < N,$$
and so the $\ominus$ lemma gives (🌵). But we know already that (🌵) implies (🐕). ∎

A sharp version of the Fenchel duality theorem

Let E be a normed space, $f, g\colon E \to\,]-\infty, \infty]$ be convex, $f + g \geq 0$ on E and

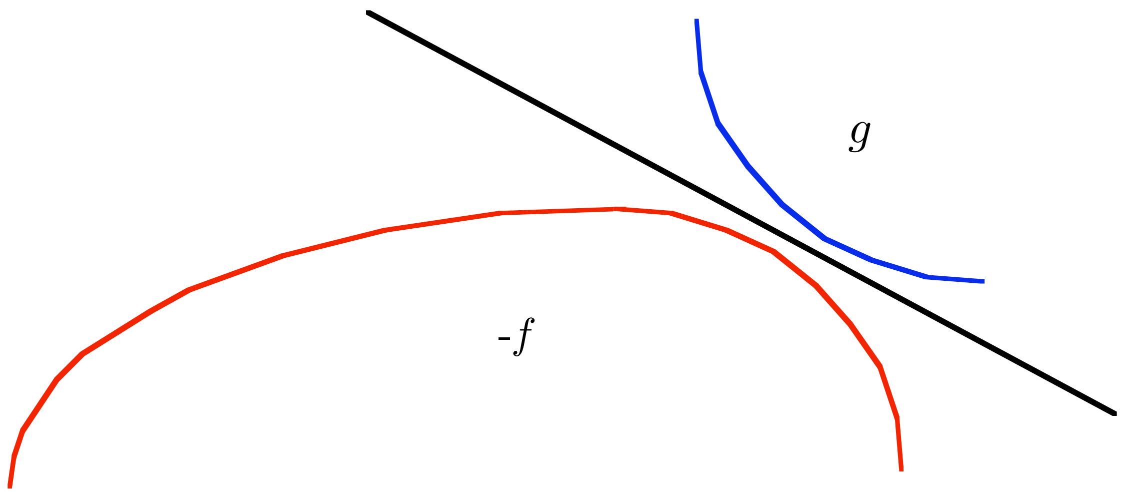
$$\sup_{x,\, y \in E,\ x \neq y} \frac{-f(x) - g(y)}{\|x - y\|} < \infty.$$

Then

$$\exists\ z^* \in E^* \quad \text{such that} \quad f^*(-z^*) + g^*(z^*) \leq 0 \qquad (🐕)$$

and

$$\min\left\{\|z^*\|\colon\ z^* \in E^*,\ z^* \text{ satisfies } (🐕)\right\} = \sup_{x,\, y \in E,\ x \neq y} \frac{-f(x) - g(y)}{\|x - y\|}.$$



Note that the quotient above is the slope of the line–segment going from $\big(y, g(y)\big)$ to $\big(x, -f(x)\big)$.