Subdifferential Calculus and Duality Theory

Subdifferential Calculus

For all $x \in \operatorname{dom} \partial f$,

$$\begin{aligned} \partial f(x) &= \{ v \mid f(x) + \langle v, y - x \rangle \leq f(y) \ \forall\, y \in \mathbf{E} \} \\ &= \{ v \mid f(x) + f^*(v) \geq \langle v, x \rangle \} \\ &= \operatorname*{argmax}_v [\langle v, x \rangle - f^*(v)] \ . \end{aligned}$$

The subdifferential calculus is more subtle than differential calculus due to issues with domains of functions under various operations that preserve convexity.

For example, the sum rule for the subdifferential may fail:

$$\partial (f_1 + f_2)(x) \neq \partial f_1(x) + \partial f_2(x).$$

Subdifferential Calculus

For all $x \in \operatorname{dom} \partial f$,

$$\begin{aligned} \partial f(x) &= \{ v \mid f(x) + \langle v, y - x \rangle \le f(y) \ \forall\, y \in \mathbf{E} \} \\ &= \{ v \mid f(x) + f^*(v) \ge \langle v, x \rangle \} \\ &= \operatorname*{argmax}_v [\langle v, x \rangle - f^*(v)] \ . \end{aligned}$$

The subdifferential calculus is more subtle than differential calculus due to issues with domains of functions under various operations that preserve convexity.

For example, the sum rule for the subdifferential may fail:

$$\partial(f_1 + f_2)(x) \ne \partial f_1(x) + \partial f_2(x).$$

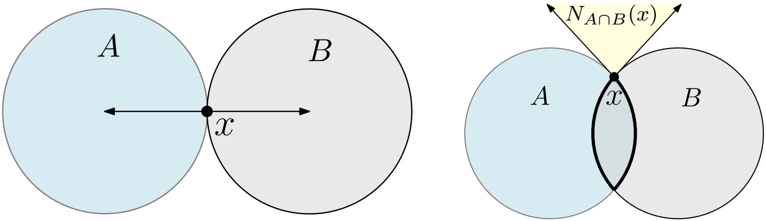
For $A := \operatorname{cl} B_1(-1, 0)$, $B := \operatorname{cl}(1, 0)$,

$$\partial\delta_{A\cap B}(0,0) = \partial(\delta_A + \delta_B)(0,0) \ne \partial\delta_A(0,0) + \partial\delta_B(0,0).$$

But for $A := \operatorname{cl} B_1(-3/4, 0)$, $B := \operatorname{cl}(3/4, 0)$,

$$\partial\delta_{A\cap B}(0,0) = \partial(\delta_A + \delta_B)(0,0) = \partial\delta_A(0,0) + \partial\delta_B(0,0).$$

Subdifferential Calculus



(a) Sum rule fails

(b) Sum rule holds

Figure: Normal cone to an intersection.

Parametric Optimization Problems

Consider a convex function $F\colon \mathbf{E} \times \mathbf{Y} \to \overline{\mathbf{R}}$ and the parametric optimization problem:

$$p(y) := \inf_x \; F(x, y).$$

Think of y as a perturbation parameter and the problem corresponding to $p(0)$ as the original "primal" problem. The assignment $y \mapsto p(y)$ is called the *value function*.

Recall that p as the infimal projection of F along the x component.

We study the variational behavior of $p(y)$ near $y = 0$. In particular, we compute $\partial p(0)$ and examine when $0 \in \operatorname{dom} \partial p$.

In conjunction with the "primal" function p, we define a corresponding "dual" function

$$q(x) := \sup\nolimits_y \; -F^{\star}(x, y).$$

We call $q(0)$ the parametric dual to $p(0)$.

$p^\star$, $p^{**}(0)$, and Weak Duality

$$
\begin{aligned}
p^*(v) &= \sup_y [\langle v, y\rangle - p(y)] \\
&= \sup_y [\langle v, y\rangle - \inf_x F(x,y)] \\
&= \sup_{(x,y)} [\langle (0,v), (x,y)\rangle - F(x,y)] \\
&= F^*(0,v).
\end{aligned}
$$

Therefore,

$$
p^{**}(0) = \sup_v [\langle v, 0\rangle - p^*(v)] = \sup_v -F^*(0,v) = q(0),
$$

so that

$$
p(0) \geq p^{**}(0) = q(0).
$$

Parametric Optimization and Duality

Theorem: Suppose that $F\colon \mathbf{E} \times \mathbf{Y} \to \overline{\mathbf{R}}$ is proper, closed, and convex. Then the following are true.

(1) **(Weak duality)** The inequality $p(0) \geq q(0)$ always holds.

(2) **(Subdifferential)**If $p(0)$ is finite, then

$$\partial p(0) \subset \operatorname*{argmax}_{y} \; -F^{\star}(0, y).$$

If, in addition, the inclusion $0 \in \operatorname{ri}(\operatorname{dom} p)$ holds, then equality holds.

(3) **(Strong duality)** If the subdifferential $\partial p(0)$ is nonempty, then $p(0) = q(0)$ and the supremum $q(0)$ is attained.

Example: Linear Programming Duality

Consider the linear program $\min\{\langle b, x\rangle \mid Ax \geq c\}$ and define

$$F(x,y) := \langle b, x\rangle + \delta(y + c - Ax|\mathbf{R}^n{}_-).$$

Then, $p(0) = \min\{\langle b, x\rangle \mid Ax \geq c\}$ and

$$F^*(u,v) = \sup_{x,y}[\langle (u,v),(x,y)\rangle - \langle b, x\rangle - \delta(y + c - Ax|\mathbf{R}^m{}_-)]$$

(use the substitution $w := y + c - Ax$ so $y = w - c + Ax$)

$$= \sup_{x,w}[\langle (u,v),(x, w + Ax - b)\rangle - \langle c, x\rangle - \delta(w|\mathbf{R}^n{}_-)]$$

$$= \sup_{x,w}[-\langle v, c\rangle + (\langle u - b + A^T v, x\rangle - \delta_{\mathbf{R}^n}(x)) + (\langle v, w\rangle - \delta(w|\mathbf{R}^m_-)]$$

$$= -\langle v, c\rangle + \delta^*_{\mathbf{R}^n}(u - b + A^T v) + \delta^*_{\mathbf{R}^m_-}(v)$$

$$= -\langle v, c\rangle + \delta_{\{0\}}(u - b + A^T v) + \delta_{\mathbf{R}^m{}_+}(v) \ ,$$

giving the dual

$$q(0) \ = \ \sup_v -F^*(x,y) \ = \ \sup\left\{\langle c, v\rangle \,\middle|\, 0 \leq v,\ A^T v = b\right\} \ .$$

Parametric Optimization and Duality

Theorem: Suppose that $F\colon \mathbf{E} \times \mathbf{Y} \to \overline{\mathbf{R}}$ is proper, closed, and convex. Then the following are true.

(1) **(Weak duality)** The inequality $p(0) \geq q(0)$ always holds.

(2) **(Subdifferential)**If $p(0)$ is finite, then

$$\partial p(0) \subset \operatorname*{argmax}_{y} \; -F^{\star}(0, y).$$

If, in addition, the inclusion $0 \in \operatorname{ri}(\operatorname{dom} p)$ holds, then equality holds.

(3) **(Strong duality)** If the subdifferential $\partial p(0)$ is nonempty, then $p(0) = q(0)$ and the supremum $q(0)$ is attained.

Parametric Optimization and Duality

Proof (2): Since p is proper, $v \in \partial p(0)$ iff $p(0) + p^*(v) = \langle 0, v \rangle = 0$. Hence,

$$q(0) = \sup_w -F^*(0, w) = p^{**}(0) \leq p(0) = -p^*(v) = -F^*(0, v),$$

so that $\partial p(0) \subset \operatorname{argmax}_y \; - F^*(0, y)$.

If $0 \in \operatorname{ri} \operatorname{dom} p$, then

$$p(0) = \operatorname{cl} p(0) = p^{**}(0),$$

and we have equality throughout the above inequality.

Proof (3): If $v \in \partial p(0)$, then

$$p(0) = \operatorname{cl} p(0) = -F^*(0, v) = \sup_w -F^*(0, w) = q(0) \; .$$

Fenchel-Rockafellar Duality

$$(P) \qquad \inf_{x\in\mathbf{E}} \; h(\mathcal{A}x) + g(x)$$

We compute the dual to (P) using the inequality

$$f(x) \geq f^{**}(x) \geq \langle y, x\rangle - f^*(x) \quad \forall\, y \in \mathbf{E}.$$

This yields

$$\begin{aligned} \mathrm{val}(P) &= \inf_x h^{\star\star}(\mathcal{A}x) + g(x), \\ &\geq \inf_x \; \langle \bar{y}, \mathcal{A}x\rangle - h^\star(y) + g(x) && \forall\, y \in \mathbf{E} \\ &= -h^\star(y) - \sup_{x\in\mathbf{E}} \; \{\langle -\mathcal{A}^* y, x\rangle - g(x)\} \\ &= -h^\star(y) - g^\star(-\mathcal{A}^* y) && \forall\, y \in \mathbf{E}. \end{aligned}$$

Giving the dual

$$(D) \qquad \sup_y -h^\star(y) - g^\star(-\mathcal{A}^* y) \; ,$$

with $\mathrm{val}(P) \geq \mathrm{val}(D)$.

Examples: Fenchel-Rockafellar Duality

Primal (P)	**Dual** (D)
$\min_x \ \frac{1}{2}\|Ax-b\|_2^2 + \|x\|_1$	$\max_y \ \left\{-\frac{1}{2}\|y\|^2 - \langle b, y\rangle : \|A^T y\|_\infty \le 1\right\}$
$\min_{x:\, \|x\|_q \le 1} \ \|Ax-b\|_p$	$\max_{y:\|y\|_{\bar{p}} \le 1} \ -\|A^T y\|_{\bar{q}} - \langle b, y\rangle$
$\min_x \ \{\langle c, x\rangle : \mathcal{A}x = b, x \in K\}$	$\max_y \ \{\langle b, y\rangle : \mathcal{A}^* y - c \in K^\circ\}$
$\min_x \ \left\{\frac{1}{2}\langle Qx, x\rangle + \langle c, x\rangle : Ax \ge b\right\}$	$\max_{y \ge 0} -\frac{1}{2}\langle Q^{-1}(c - A^T y), c - A^T y\rangle + \langle b, y\rangle$

Table: Fenchel-Rockafellar dual pairs. The parameters are: K is a convex cone, $Q \succ 0$, and $p, \bar{p}, q, \bar{q} \in [1, \infty]$ satisfy $p^{-1} + \bar{p}^{-1} = q^{-1} + \bar{q}^{-1} = 1$.

Fenchel-Rockafellar Duality

We now establish strong duality (P) and (D). Define

$$p(y) = \inf_x F(x, y) := h(\mathcal{A}x + y) + g(x)$$

so that the primal problem is $p(0)$.

For the dual, observe that

$$F^*(u, v) = \sup_{x,y}[\langle (u, v), (x, y)\rangle - h(\mathcal{A}x + y) + g(x)]$$

(use the substitution $w := \mathcal{A}x + y$ so that $y = w - \mathcal{A}x$

$$= \sup_{x,w}[\langle u, x\rangle + \langle v, w - \mathcal{A}x\rangle - h(w) - g(x)]$$

$$= \sup_{x,w}[(\langle u - A^T v, x\rangle - g(x)) + (\langle v, w\rangle - h(w))]$$

$$= g^*(u - A^T v) + h^*(v)$$

giving the dual

$$q(0) = \sup_y -F^*(0, y) \ = \ \sup_y[-h^*(y) - g^*(-A^T y)].$$

Fenchel-Rockafellar Duality: Strong Duality

Strong duality follows from the condition $0 \in \operatorname{ri}\operatorname{dom} p$, where

$$p(y) = \inf_x F(x,y) := h(\mathcal{A}x + y) + g(x) \ .$$

Note that $y \in \operatorname{dom} p$ iff $\exists\, x \in \operatorname{dom} g$ such that $\mathcal{A}x + y \in \operatorname{dom} h$, or equivalently, $x \in \operatorname{dom} g$ and $y \in \operatorname{dom} h - \mathcal{A}x$. In other words,

$$\operatorname{dom} p = \operatorname{dom} h \ - \ \mathcal{A}\operatorname{dom} g.$$

Therefore,

$$\operatorname{ri}\operatorname{dom} p = \operatorname{ri}(\operatorname{dom} h \ - \ \mathcal{A}\operatorname{dom} g) = \operatorname{ri}\operatorname{dom} h \ - \ \mathcal{A}\operatorname{ri}\operatorname{dom} g.$$

Consequently, $0 \in \operatorname{ri}\operatorname{dom} p$ if and only

$$0 \in \operatorname{ri}\operatorname{dom} h \ - \ \mathcal{A}\operatorname{ri}\operatorname{dom} g,$$

or equivalently,

$$\exists\, x \in \operatorname{ri}\operatorname{dom} g \ \text{ such that } \ \mathcal{A}x \in \operatorname{ri}\operatorname{dom} h\,.$$

Fenchel-Rockafellar Duality: Strong Duality

Theorem: Consider the problems:

$$
\begin{aligned}
(P) \qquad & \min_x \; h(\mathcal{A}x) + g(x) \\
(D) \qquad & \max_y \; -g^\star(-\mathcal{A}^* y) - h^\star(y).
\end{aligned}
$$

where $g\colon \mathbf{E} \to \overline{\mathbf{R}}$ and $h\colon \mathbf{Y} \to \overline{\mathbf{R}}$ are proper, closed convex functions, and $\mathcal{A}\colon \mathbf{E} \to \mathbf{Y}$ is a linear map. If the regularity condition

$$0 \in \operatorname{ri}(\operatorname{dom} h) - \mathcal{A}(\operatorname{ri}\operatorname{dom} g) \tag{1}$$

holds, then the primal and dual optimal values are equal and the dual optimal value is attained, if finite.

Primal-dual optimality conditions

$$
\begin{array}{ll}
(P) & \min_{x} \; h(\mathcal{A}x) + g(x) \\
(D) & \max_{y} \; -g^{\star}(-\mathcal{A}^{*}y) - h^{\star}(y).
\end{array}
$$

Note that the direct optimality conditions for F-R primal-dual pair are

$$
\left\{
\begin{array}{l}
0 \in \mathcal{A}^{*}\partial h(\mathcal{A}x) + \partial g(x) \\
0 \in -\mathcal{A}\partial g^{\star}(-\mathcal{A}^{\star}y) + \partial h^{\star}(y)
\end{array}
\right\}.
$$

Two disadvantages of this representation are

(1) The variables x and y appear unrelated, even though they are closely related.

(2) The fact that the subdifferentials ∂h and ∂g are evaluated at points in the image of $\mathcal{A}$ and $\mathcal{A}^{*}$, respectively, is inconvenient for computation.

Primal-Dual optimality conditions

Theorem: Consider the Fenchel-Rockafellar duality framework:

$$\begin{aligned} (P) \qquad & \min_x \; h(\mathcal{A}x) + g(x) \\ (D) \qquad & \max_y \; -g^\star(-\mathcal{A}^* y) - h^\star(y). \end{aligned}$$

where $g\colon \mathbf{E} \to \overline{\mathbf{R}}$ and $h\colon \mathbf{Y} \to \overline{\mathbf{R}}$ are proper, closed, convex functions, and $\mathcal{A} \in L[\mathbf{E}, \mathbf{Y}]$. Suppose that the optimal values of (P) and (D) are equal, as is implied, for example, by either of the two regularity conditions:

$$\begin{aligned} 0 &\in \mathrm{ri}\,(\mathrm{dom}\, h) - \mathcal{A}(\mathrm{ri}\,\mathrm{dom}\, g) \\ 0 &\in \mathrm{ri}\,(\mathrm{dom}\, g^\star) + \mathcal{A}^*(\mathrm{ri}\,\mathrm{dom}\, h^\star). \end{aligned}$$

Then x is the minimizer of (P) and y is the maximizer of (D) if and only if

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \in \begin{bmatrix} 0 & \mathcal{A}^* \\ -\mathcal{A} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \partial g(x) \times \partial h^\star(y).$$

Primal-Dual Optimality Conditions: Proof

Since the primal and dual optimal values are equal, we deduce that x is a minimizer of (P) and y is a maximizer of (D) if and only if equality holds:

$$0 = (h(\mathcal{A}x) + g(x)) + (g^\star(-\mathcal{A}^* y) + h^\star(y)) \,. \tag{2}$$

The F-Y ineq. guarantees

$$h(\mathcal{A}x) + h^\star(y) \geq \langle \mathcal{A}x, y\rangle \quad \text{and} \quad g^\star(-\mathcal{A}^* y) + g(x) \geq \langle -\mathcal{A}^* y, x\rangle . \tag{3}$$

Adding the two inequalities in (3), we see that the right side of (2) is always lower-bounded by zero. We therefore deduce that (2) holds if and only if the inequalities (3) hold as equalities. This happens precisely when the inclusions, $\mathcal{A}x \in \partial h^\star(y)$ and $-\mathcal{A}^* y \in \partial g(x)$, hold. Again, by the F-Y ineq., these two inclusions are exactly the system (3) which implies (2).

Subdifferential Calculus

Theorem: Let $g\colon \mathbf{E} \to \overline{\mathbf{R}}$ and $h\colon \mathbf{Y} \to \overline{\mathbf{R}}$ be proper, closed convex functions and $\mathcal{A}\colon \mathbf{E} \to \mathbf{Y}$ a linear map. Then for any point $x \in \mathbf{E}$,

$$\partial g(x) + \mathcal{A}^* \partial h(\mathcal{A}x) \subset \partial(g + h \circ \mathcal{A})(x) \ .$$

Moreover, equality holds if

$$0 \in \operatorname{ri}(\operatorname{dom} h) - \mathcal{A}(\operatorname{ri}\operatorname{dom} g).$$

Subdifferential Calculus

Proof: If $v \in \partial g(x)$ and $w \in \partial h(\mathcal{A}x)$, then

$$\left.\begin{array}{r} g(x) + \langle v, y - x\rangle \leq g(y) \\ h(\mathcal{A}x) + \langle \mathcal{A}^* w, y - x\rangle \leq h(\mathcal{A}y) \end{array}\right\} \quad \forall\, y \in \mathbf{E}\ .$$

Adding these two inequalities yields

$$g(x) + h(\mathcal{A}x) + \langle v + \mathcal{A}^* w, y - x\rangle \leq g(y) + h(\mathcal{A}y) \quad \forall\, y \in \mathbf{E}.$$

Hence,

$$\partial g(x) + \mathcal{A}^* \partial h(\mathcal{A}x) \subset \partial (g + h \circ \mathcal{A})(x).$$

For the reverse inclusion we set $f(x) := g(x) + h(\mathcal{A}x)$.

Subdifferential Calculus

Now assume $0 \in \mathrm{ri}\,(\mathrm{dom}\, h) - \mathcal{A}(\mathrm{ri}\,\mathrm{dom}\, g)$ and let $v \in \partial f(x)$.
WLOG $v = 0$, else replace g by $g - \langle v, \cdot \rangle$.

Then $x \in \mathrm{argmin}\, f$. The F-R Duality Theorem guarentees that $f(x) = \max_y -g^*(-\mathcal{A}^* y) - h^*(y)$ and $S := \mathrm{argmax}_y[-g^*(-\mathcal{A}^* y) - h^*(y)] \neq \emptyset$. Then, for any $y \in S$,

$$\begin{aligned} 0 &= (g(x) + h(\mathcal{A}x)) + (g^\star(-\mathcal{A}^* y) + h^\star(y)) \\ &= (g(x) + g^\star(-\mathcal{A}^* y)) + (h(\mathcal{A}x) + h^\star(y)) \\ &\geq \langle x, -\mathcal{A}^* y \rangle + \langle \mathcal{A}x, y \rangle = 0, \end{aligned}$$

where the final inequality follows from the F-Y Ineq. Hence equality holds throughout and, again by F-Y Ineq.

$$g(x) + g^\star(-\mathcal{A}^* y) = \langle x, -\mathcal{A}^* y \rangle \quad \text{and} \quad h(\mathcal{A}x) + h^\star(y) = \langle \mathcal{A}x, y \rangle.$$

Hence

$$0 = -\mathcal{A}^* y + \mathcal{A}^* y \in \partial g(x) + \mathcal{A}^* \partial h(\mathcal{A}x),$$

as claimed.

$N_{A\cap B}(x) = N_A(x) + N_B(x)$

Corollary: Let A and B be close convex sets in $\mathbf{E}$ such that $\operatorname{ri} A \cap \operatorname{ri} B \neq \emptyset$. Then, for all $x \in A \cap B$,

$$N_{A\cap B}(x) = N_A(x) + N_B(x).$$

Proof: The theorem tells us that

$$\begin{aligned}\partial\delta_{A\cap B}(x) &= \partial(\delta_A + \delta_B)(x)\\ &= \partial\delta_A(x) + \partial\delta_B(x)\\ &= N_A(x) + N_B(x)\ ,\end{aligned}$$

since $\operatorname{ri}\operatorname{dom}\delta_A = \operatorname{ri} A$ and $\operatorname{ri}\operatorname{dom}\delta_B = \operatorname{ri} B$.

$f(x) := \max\{f_i(x) \mid i = 1, \ldots, k\}$

Theorem: Let $f_i : \mathbf{E} \to \mathbf{R}$ be closed proper cvx, $i = 1, 2, \ldots, k$, and define $f(x) := \max\{f_i(x) \mid i = 1, \ldots, k\}$. Then

$$\partial f(x) = \mathrm{conv}\left(\bigcup_{i \in I(x)} \partial f_i(x)\right),$$

where $I(x) := \{i \mid f_i(x) = f(x)\}$.

Proof:

$$\begin{aligned}
v \in \partial f(x) &\iff (v, -1) \in N_{\mathrm{epi}\, f}(x, f(x)) \\
&\iff (v, -1) \in N((x, f(x)) \mid \bigcap_{i=1}^{k} \mathrm{epi}\, f_i) \\
&\iff (v, -1) \in \sum_{i=1}^{k} N_{\mathrm{epi}\, f_i}(x, f(x)) \\
&\iff \exists\, (w_i, \nu_i) \in N_{\mathrm{epi}\, f_i}(x, f(x))\, (i = 1, \ldots, k) \text{ s.t. } (v, -1) = \sum_{i=1}^{k} (w_i, \nu_i)
\end{aligned}$$

$f(x) := \max\{f_i(x) \mid i = 1, \dots, k\}$

$$\iff \begin{cases} \exists\, (w_i, \nu_i) \in N_{\mathrm{epi}\, f_i}(x, f(x)),\ \nu_i < 0\, (i \in I(x)) \\ \text{s.t. } (v, -1) = \displaystyle\sum_{i \in I(x)}^{k} (w_i, \nu_i), \end{cases}$$

where the final equivalence comes from the fact that if $f_i(x) < f(x)$ then $(x, f(x)) \in \operatorname{intr\,epi} f_i$ so $N_{\mathrm{epi}\, f_i}(x, f(x)) = \{(0,0)\}$.

Set $\lambda_i = -\nu_i$, $\lambda_i v_i = w_i$ $(i \in I(x))$. Then

$$v \in \partial f(x) \iff (v, -1) = \sum_{i \in I(x)} \lambda_i (v_i, -1)$$

with $\sum_{i \in I(x)} \lambda_i = 1$ and $0 \le \lambda_i$ $(i \in I(x))$. Therefore,

$$\partial f(x) = \operatorname{conv}\left(\bigcup_{i \in I(x)} \partial f_i(x) \right).$$

$N_{\mathcal{A}^{-1}Q}(\bar{x}) = A^* N_Q(\mathcal{A}\bar{x})$

Corollary: Let $\mathcal{A} \in L[\mathbf{E}, \mathbf{Y}]$ and $Q \subset \mathbf{Y}$ be such that Q is closed cvx and $\mathrm{Im}(\mathcal{A}) \cap \mathrm{ri}\, Q \neq \emptyset$. Then, for all $x \in \Omega$, $N_{\mathcal{A}^{-1}Q}(x) = \mathcal{A}^* N_Q(\mathcal{A}x)$.

Proof: In the subdifferential calculus theorem take $h = \delta_Q$ and $g \equiv 0 = \delta_{\mathbf{E}}$.

Then, by hypothesis, $0 \in \mathrm{ri}\,\mathrm{dom}\, h - \mathcal{A}\,\mathrm{ri}\,\mathrm{dom}\, g = \mathrm{ri}\, Q - \mathrm{Im}(A)$.

Since $\delta_{\mathcal{A}^{-1}Q}(x) = \delta_Q(\mathcal{A}x)$, for all $x \in A^{-1}Q$,

$$\begin{aligned} N_{\mathcal{A}^{-1}Q}(x) &= \partial\delta_{\mathcal{A}^{-1}Q}(x) \\ &= \partial(\delta_Q \circ \mathcal{A})(x) \\ &= \mathcal{A}^* N_Q(\mathcal{A}x)\,. \end{aligned}$$

Level Sets

Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be closed proper cvx, and consider the lower level sets

$$\text{lev}_f(r) := \{x \,|\, f(x) \le r\}.$$

If we let M_r be the affine set $M_r := \mathbf{E} \times \{r\}$ and P be the projection $P(x, \mu) := x$, then

$$\text{lev}_f(r) = P(M_r \cap \text{epi}\, f) = P(\{(x, \mu) \,|\, \mu = r,\ (x, \mu) \in \text{epi}\, f\}).$$

Hence, if $r > \inf f$,

$$\begin{aligned} \text{ri}\,\text{lev}_f(r) &= \text{ri}\, P(M_r \cap \text{epi}\, f) = P\,\text{ri}\,(M_r \cap \text{epi}\, f) = P(M_r \cap \text{ri}\,\text{epi}\, f) \\ &= \{x \in \text{ri}\,\text{dom}\, f \,|\, f(x) < r\}, \end{aligned}$$

and

$$\text{cl}\,\text{lev}_f(r) = \{x \,|\, \text{cl}\, f(x) \le r\}.$$

Moreover, all these sets have the same closure.

Level Sets: Tangent and Normal Cones

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be closed proper cvx and $\bar{x} \in \operatorname{dom} \partial f$ be such that $f(\bar{x}) > \inf f$. Then

$$T(\bar{x} | \operatorname{lev}_f(f(\bar{x}))) = \{ d \mid f'(\bar{x}; d) \leq 0 \} \quad \text{and} \quad N(\bar{x} | \operatorname{lev}_f(f(\bar{x}))) = \mathbf{R}_+ \partial f(\bar{x}) \, .$$

Level Sets: Tangent and Normal Cones

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be closed proper cvx and $\bar{x} \in \operatorname{dom} \partial f$ be such that $f(\bar{x}) > \inf f$. Then

$$T(\bar{x} | \operatorname{lev}_f(f(\bar{x}))) = \{ d \mid f'(\bar{x}; d) \leq 0 \} \quad \text{and} \quad N(\bar{x} | \operatorname{lev}_f(f(\bar{x}))) = \mathbf{R}_+ \partial f(\bar{x}) \,.$$

Proof: Since $\bar{x} \in \operatorname{dom} \partial f$, $f'(\bar{x}; \cdot) = \delta^*(x | \partial f(\bar{x}))$ is closed proper cvx. Since $f(\bar{x}) > \inf f$, $\operatorname{lev}_f(f(\bar{x})) = \operatorname{cl} \{ x \in \operatorname{ri} \operatorname{dom} f \mid f(x) < f(\bar{x}) \}$ so

$$\begin{aligned} T(\bar{x} | \operatorname{lev}_f(f(\bar{x}))) &= \operatorname{cl} \{ \lambda(x - \bar{x}) \mid f(x) < f(\bar{x}), \ \lambda \geq 0 \} \\ &= \operatorname{cl} \{ d \mid \exists\, t > 0 \text{ s.t. } f(x + td) - f(\bar{x}) < 0 \} \\ &= \operatorname{cl} \{ d \mid f'(x; d) < 0 \} \\ &= \{ d \mid f'(x; d) \leq 0 \}. \end{aligned}$$

In addition,

$$\begin{aligned} N(\bar{x} | \operatorname{lev}_f(f(\bar{x}))) &= T(\bar{x} | \operatorname{lev}_f(f(\bar{x})))^\circ \\ &= \{ d \mid \delta^*(d | \partial f(\bar{x})) \leq 0 \}^\circ \\ &= \{ d \mid \langle v, d \rangle \leq 0 \ \forall\, v \in \partial f(\bar{x}) \}^\circ \\ &= \{ d \mid \langle v, d \rangle \leq 0 \ \forall\, v \in \mathbf{R}_+ \partial f(\bar{x}) \}^\circ \\ &= ((\mathbf{R}_+ \partial f(\bar{x}))^\circ)^\circ \\ &= \mathbf{R}_+ \partial f(\bar{x}) \qquad \text{(since } 0 \notin \partial f(\bar{x})\text{)} \end{aligned}$$

Normal Cones to Constraint Regions

Theorem: Let $f_i : \mathbf{E} \to \mathbf{R}$ $(i = 1, \ldots, k)$, $\mathcal{A} \in L[\mathbf{E}, \mathbf{Y}]$, and $Q \subset \mathbf{Y}$. Define $F : \mathbf{E} \to \mathbf{R}^k$ to have component functions f_i, $K := \mathbf{R}^k_- \times Q$, and $\Omega := \{x \,|\, (F(x), \mathcal{A}x) \in K\}$. If there exists $\hat{x} \in \mathbf{E}$ such that

$$f_i(\hat{x}) < 0 \;(i = 1, \ldots, k) \quad \text{and} \quad \mathcal{A}\hat{x} \in \operatorname{ri} Q,$$

then, for every $\bar{x} \in \Omega$,

$$N_\Omega(\bar{x}) = \textstyle\sum_{i \in I(\bar{x})} \mathbf{R}_+ \partial f_i(\bar{x}) \;+\; A^* N_Q(\mathcal{A}\bar{x}),$$

where $I(\bar{x}) = \{i \,|\, f_i(x) = 0\}$.

Proof: Let $h = \delta_Q$ and $g := \sum_{i=1}^k \delta_{\mathrm{lev}_{f_i}(0)}$. The hypotheses imply that $f := h \circ \mathcal{A} + g$ satisfied the regularity conditions of our theorem on the subgradient calculus, hence

$$\begin{aligned} \partial f(\bar{x}) &= \mathcal{A}^* \partial h(\mathcal{A}\bar{x}) + \partial g(\bar{x}) \\ &= \mathcal{A}^* N_Q(\mathcal{A}\bar{x}) + \sum_{i=1}^k N(\bar{x} | \operatorname{lev}_{f_i}(0)) \\ &= \mathcal{A}^* N_Q(\mathcal{A}\bar{x}) + \sum_{i \in I(\bar{x})} \mathbf{R}_+ \partial f_i(\bar{x}) \,. \end{aligned}$$

Normal Cones to Constraint Regions

Theorem: Let $f_i : \mathbf{E} \to \mathbf{R}$ $(i = 1, \ldots, k)$, $\mathcal{A} \in L[\mathbf{E}, \mathbf{Y}]$, and $Q \subset \mathbf{Y}$. Define $F : \mathbf{E} \to \mathbf{R}^k$ to have component functions f_i, $K := \mathbf{R}^k_- \times Q$, and $\Omega := \{x \,|\, (F(x), \mathcal{A}x) \in K\}$. If there exists $\hat{x} \in \mathbf{E}$ such that

$$f_i(\hat{x}) < 0 \ (i = 1, \ldots, k) \quad \text{and} \quad \mathcal{A}\hat{x} \in \operatorname{ri} Q,$$

then, for every $\bar{x} \in \Omega$,

$$N_\Omega(\bar{x}) = \textstyle\sum_{i \in I(\bar{x})} \mathbf{R}_+ \partial f_i(\bar{x}) \ + \ A^* N_Q(\mathcal{A}\bar{x}),$$

where $I(\bar{x}) = \{i \,|\, f_i(x) = 0\}$.

Proof: Let $h = \delta_Q$ and $g := \sum_{i=1}^k \delta_{\operatorname{lev}_{f_i}(0)}$. The hypotheses imply that $f := h \circ \mathcal{A} + g$ satisfied the regularity conditions of our theorem on the subgradient calculus, hence

$$\begin{aligned}
\partial f(\bar{x}) &= \mathcal{A}^* \partial h(\mathcal{A}\bar{x}) + \partial g(\bar{x}) \\
&= \mathcal{A}^* N_Q(\mathcal{A}\bar{x}) + \sum_{i=1}^k N(\bar{x} | \operatorname{lev}_{f_i}(0)) \\
&= \mathcal{A}^* N_Q(\mathcal{A}\bar{x}) + \sum_{i \in I(\bar{x})} \mathbf{R}_+ \partial f_i(\bar{x}) \,.
\end{aligned}$$

Slater Constraint Qualification:
$\exists\, \hat{x}$ s.t. $f_i(\hat{x}) < 0$ $(i = 1, \ldots, k)$, $\mathcal{A}\hat{x} \in \operatorname{ri} Q$.

Constrained Convex Optimization

Theorem: Let $f_i : \mathbf{E} \to \mathbf{R}$ $(i = 1, \ldots, k)$, $\mathcal{A} \in L[\mathbf{E}, \mathbf{Y}]$, and $Q \subset \mathbf{Y}$ satisfy the conditions of the previous result including the Slater CQ. In addition, let $f_0 : \mathbf{E} \to \overline{\mathbf{R}}$ be closed proper convex and $\exists\, \hat{x} \in \operatorname{ri} \operatorname{dom} f_0$ s.t. $f_i(\hat{x}) < 0$ $(i = 1, \ldots, k)$, $\mathcal{A}\hat{x} \in \operatorname{ri} Q$. Then $\bar{x}$ solves $\min_{x \in \Omega} f$ if and only if there exist multipliers $y_i \geq 0$ $(i \in I(\bar{x}))$ such that

$$0 \in \partial f_0(\bar{x}) + \sum_{i \in I(\bar{x})} y_i \partial f_i(\bar{x}) \; + \; \mathcal{A}^* N_Q(\mathcal{A}\bar{x}),$$

where $I(\bar{x}) = \{ i \,|\, f_i(x) = 0 \}$.

Proof: The hypotheses imply that the function $f = f_0 + \sum_{i=1}^{k} \delta_{\operatorname{lev}_{f_i}(0)} + \delta_Q \circ \mathcal{A}$ is closed proper cvx with $\hat{x} \in \operatorname{ri} \operatorname{dom} f$. Hence $\bar{x}$ solves $\min_{x \in \Omega} f_0$ if and only if $0 \in \partial f(\bar{x})$. The previous results show that the inclusion $0 \in \partial f_0(\bar{x})$ is equivalent to the statement given in the theorem.

Lagrangian Duality

Consider the constrained optimization problem

$$\begin{array}{lll} (P) & \text{minimize}_x & f_0(x) \\ & \text{subject to} & f_i(x) \leq 0 \ (i = 1, \ldots, k), \\ & & f_i(x) = 0 \ (i = k+1, \ldots, m), \end{array}$$

where $f_0 : \mathbf{E} \to \overline{\mathbf{R}}$ and $f_i : \mathbf{E} \to \mathbf{R}$ $(i = 1, \ldots, k)$ are closed proper cvx and $f_i : \mathbf{E} \to \mathbf{R}$ $(i = k+1, \ldots, m)$ are affine.

We define the *Lagrangian* for (P) to be the mapping $L : \mathbf{E} \times \mathbf{R}^k \to \overline{\mathbf{R}}$ given by

$$L(x, y) := f_0(x) + \langle y, F(x) \rangle - \delta_K^*(y),$$

where $K := \mathbf{R}_-^k \times \{0\}^{m-k}$. Since K is a closed convex cone, we have $\delta_K^* = \delta_{K^\circ}$ where $K^\circ = \mathbf{R}_+^k \times \mathbf{R}^{m-k}$.

Lagrangian Duality

Consider the constrained optimization problem

$$
\begin{array}{lll}
(P) & \text{minimize}_x & f_0(x) \\
 & \text{subject to} & f_i(x) \le 0 \ (i = 1, \ldots, k), \\
 & & f_i(x) = 0 \ (i = k+1, \ldots, m),
\end{array}
$$

where $f_0 : \mathbf{E} \to \overline{\mathbf{R}}$ and $f_i : \mathbf{E} \to \mathbf{R}$ $(i = 1, \ldots, k)$ are closed proper cvx and $f_i : \mathbf{E} \to \mathbf{R}$ $(i = k+1, \ldots, m)$ are affine.

We define the *Lagrangian* for (P) to be the mapping $L : \mathbf{E} \times \mathbf{R}^k \to \overline{\mathbf{R}}$ given by

$$L(x, y) := f_0(x) + \langle y, F(x) \rangle - \delta_K^*(y),$$

where $K := \mathbf{R}_-^k \times \{0\}^{m-k}$. Since K is a closed convex cone, we have $\delta_K^* = \delta_{K^\circ}$ where $K^\circ = \mathbf{R}_+^k \times \mathbf{R}^{m-k}$.

For every $x \in \mathbf{E}$,

$$\sup_y L(x, y) \;=\; f_0(x) + \delta_K(F(x)),$$

hence the problem (P) can be written as

$$(P) \qquad \inf_x \sup_y L(x, y).$$

Lagrangian Duality

The Lagrangian dual to the problem

$$(P) \qquad \min_{x\in\Omega} f_0(x) \;=\; \inf_x \sup_y L(x,y)$$

is the problem

$$(D) \qquad \sup_y \Phi(y) \;=\; \sup_y \inf_x L(x,y),$$

where dual objective function Φ is given by

$$\Phi(y) := \inf_x L(x,y) \;=\; \inf_x f_0(x) + \langle y, F(x)\rangle - \delta^*_K(y).$$

We may write the dual as

$$\sup_{y\in K^\circ} \inf_x [f_0(x) + \langle y, F(x)\rangle] \;.$$

The weak duality inequality is

$$\mathrm{Val}(P) = \inf_x \sup_y L(x,y) \;\geq\; \sup_y \inf_x L(x,y) = \mathrm{Val}(D) \;.$$

Strong Duality Theorem

$$\begin{array}{lccc} (P) & \min_{x\in\Omega} f_0(x) & = & \inf_x \sup_y L(x,y) \\ (D) & \sup_{y\in K^\circ} \Phi(y) & = & \sup_y \inf_x L(x,y) \end{array}$$

$$\Omega = \{x \,|\, f_i(x) \leq 0 \ (i = 1,\ldots,k),\ f_i(x) = 0 \ (i = k+1,\ldots m)\} = \{x \,|\, F(x) \in K\}$$

Theorem: Consider the problem (P) as defined above. If

$$\exists\, x \in \mathrm{ri}\,\mathrm{dom}\, f_0 \ \text{ s.t. } \ f_i(x) < 0 \, (i = 1,\ldots,k) \ \text{ and } \ f_i(x) = 0 \, (i = k{+}1,\ldots m),$$

then the primal and dual optimal values are equal and the dual optimal value is attained.

Strong Duality Theorem: Proof

The proof follows from the perturbation framework given by

$$\mathbf{F}(x,z) := f_0(x) + \delta_K(F(x)+z).$$

The strong duality assumptions are satisfied as the Slater condition holds. To see the that we have computed the dual correctly, observe that

$$\begin{aligned}
\mathbf{F}^*(0,y) &= \sup_{(x,z)} [\langle (0,y),(x,z)\rangle - f_0(x) - \delta_K(F(x)+z)] \\
&= \sup_{(x,w)} \langle y, w - F(x)\rangle - f_0(x) - \delta_K(w) \\
&= \sup_{w} [\langle y, w\rangle - \delta_K(w)] \ - \ \inf_x [f_0(x) + \langle y, F(x)\rangle] \\
&= \delta_{K^\circ}(y) \ - \ \inf_x [f_0(x) + \langle y, F(x)\rangle].
\end{aligned}$$

So the dual is

$$\sup_y -\mathbf{F}^*(0,y) = \sup_{y\in K^\circ} \inf_x [f_0(x) + \langle y, F(x)\rangle] = \sup_{y\in K^\circ} \Phi(y).$$

Lagrangian Duality: Quadratic Programming

$$
(QP) \qquad \begin{array}{ll} \text{minimize} & \frac{1}{2}x^T Q x + c^T x \\ \text{subject to} & Ax \leq b \end{array}
$$

where $Q \in \mathbb{S}^n_+, \; c \in \mathbf{R}^n, \; A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$.

Lagrangian Duality: Quadratic Programming

$$(QP) \qquad \begin{array}{ll} \text{minimize} & \frac{1}{2}x^T Q x + c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

where $Q \in \mathbb{S}^n_+,\ c \in \mathbf{R}^n,\ A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$.

Let Q have Cholesky factorization $Q = LL^T$ where $L \in \mathbf{R}^{n \times k}$ with k the rank of Q. Then rewrite (QP) as

$$\widehat{(QP)} \qquad \begin{array}{ll} \text{minimize} & \frac{1}{2}\|z\|^2 + c^T x \\ \text{subject to} & Ax \leq b,\ [L^T, -I]\begin{pmatrix} x \\ z \end{pmatrix} = 0. \end{array}$$

Lagrangian Duality: Quadratic Programming

In this case, $K = \mathbf{R}^m_- \times \{0\}^k$, and

$$f_0(x,z) = \tfrac{1}{2}\|z\|^2 + c^T x \quad \text{and} \quad F(x,z) = \begin{bmatrix} A & 0 \\ L^T & -I \end{bmatrix} \begin{pmatrix} x \\ z \end{pmatrix} - \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

Given $(u,v) \in K^\circ = \mathbf{R}^m_+ \times \mathbf{R}^k$ the dual objective is

$$\begin{aligned} \Phi(u,v) &= \inf_{(x,z)} \frac{1}{2}\|z\|^2 + c^T x + \langle (u,v), F(x,z) \rangle \\ &= \inf_{(x,z)} \frac{1}{2}\|z\|^2 + c^T x + \langle u, Ax - b \rangle + \langle v, L^T x - z \rangle. \end{aligned}$$

Lagrangian Duality: Quadratic Programming

In this case, $K = \mathbf{R}^m_- \times \{0\}^k$, and

$$f_0(x,z) = \frac{1}{2}\|z\|^2 + c^T x \quad \text{and} \quad F(x,z) = \begin{bmatrix} A & 0 \\ L^T & -I \end{bmatrix} \begin{pmatrix} x \\ z \end{pmatrix} - \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

Given $(u,v) \in K^\circ = \mathbf{R}^m_+ \times \mathbf{R}^k$ the dual objective is

$$\begin{aligned} \Phi(u,v) &= \inf_{(x,z)} \frac{1}{2}\|z\|^2 + c^T x + \langle (u,v), F(x,z) \rangle \\ &= \inf_{(x,z)} \frac{1}{2}\|z\|^2 + c^T x + \langle u, Ax - b \rangle + \langle v, L^T x - z \rangle. \end{aligned}$$

This optimization problem can be solved by solving the equations

$$\begin{aligned} 0 &= c + A^T u + Lv \\ 0 &= z - v \end{aligned}$$

Lagrangian Duality: Quadratic Programming

$$0 = c + A^T u + Lv$$
$$0 = z - v$$

Plugging this information into

$\Phi(u, v) = \inf_{(x,z)} \frac{1}{2} \|z\|^2 + c^T x + \langle u, Ax - b\rangle + \langle v, L^T x - z\rangle$

we find that

$\Phi(u, v) = \frac{1}{2} \|v\|^2 - \langle u, b\rangle - \|v\|^2 + \delta_{\{0\}}(c + A^T u + Lv)$.

Hence the dual problem becomes

$$\sup_{(u,v)} -[\frac{1}{2} \|v\|^2 + \langle u, b\rangle] \;\text{ s.t. }\; c + A^T u + Lv = 0,\; 0 \le u.$$

Lagrangian Duality: Quadratic Programming

Since $\ker L = \{0\}$, $(L^T L)^{-1}$ exists, so we can multiply $c + A^T u + Lv = 0$ by L^T to find $v = -(L^T L)^{-1} L^T (c + A^T u)$ allowing us to remove v from the dual and obtain the dual

$$\sup_{0 \leq u} -\left[(c + A^T u)^T L (L^T L)^{-2} L^T (c + A^T u) + \langle u, b \rangle\right] .$$

Lagrangian Duality: Quadratic Programming

Since $\ker L = \{0\}$, $(L^T L)^{-1}$ exists, so we can multiply $c + A^T u + Lv = 0$ by L^T to find $v = -(L^T L)^{-1} L^T (c + A^T u)$ allowing us to remove v from the dual and obtain the dual

$$\sup_{0 \le u} -\left[(c + A^T u)^T L (L^T L)^{-2} L^T (c + A^T u) + \langle u, b \rangle\right] .$$

Primal Solution Recovery: Suppose instead we use the compact singular value decomposition of $Q = UDU^T$, where D is the diagonal matrix of the k nonzero singular values of Q and $U^T U = I_k$.

In this case that $L = UD^{1/2}$, where $D^{1/2}$ is the diagonal matrix of the square roots of the singular values. If u solves the dual, then the optimal x satisfies

$$\begin{aligned} D^{1/2} U^T x = L^T x = z = v \\ &= -(L^T L)^{-1} L^T (c + A^T u) \\ &= -(D^{1/2} U^T U D^{1/2})^{-1} D^{1/2} U^T (c + A^T u) \\ &= -D^{-1/2} U^T (c + A^T u). \end{aligned}$$

So $U^T x = -D^{-1} U^T (c + A^T u)$.

Horizon Cones

Given $S \subset \mathbf{E}$, we define the *horizon cone* of S to be

$$S^\infty := \begin{cases} \left\{ d \,\middle|\, \exists \{x_k\} \subset S,\, t_k \downarrow 0 \text{ s.t. } \left\|x^k\right\| \uparrow \infty \text{ and } t_k x_k \to d \right\} & , \; S \neq \emptyset, \\ \{0\} & , \; S = \emptyset. \end{cases}$$

Clearly, S^∞ is always a closed nonempty cone, and S is bounded iff $S^\infty = \{0\}$.

Horizon Cones

Given $S \subset \mathbf{E}$, we define the *horizon cone* of S to be

$$S^\infty := \begin{cases} \left\{ d \,\middle|\, \exists \{x_k\} \subset S,\ t_k \downarrow 0 \text{ s.t. } \left\| x^k \right\| \uparrow \infty \text{ and } t_k x_k \to d \right\} & ,\ S \neq \emptyset, \\ \{0\} & ,\ S = \emptyset. \end{cases}$$

Clearly, S^∞ is always a closed nonempty cone, and S is bounded iff $S^\infty = \{0\}$.

Lemma: If $Q \subset \mathbf{E}$ is convex, then Q^∞ is a convex cone and

$$Q^\infty = \{ d \,|\, x + \lambda d \in Q \ \forall\, x \in Q,\ \lambda \geq 0 \}.$$

Closedness of the linear image of sets

Theorem: Let $C \subset \mathbf{E}$ be closed and $\mathcal{A} \in L[\mathbf{E}, \mathbf{Y}]$. If $\ker \mathcal{A} \cap C^\infty = \{0\}$, then $\mathcal{A}C$ is closed and $(\mathcal{A}C)^\infty = \mathcal{A}C^\infty$ although, in general, we only have $\mathcal{A}C^\infty \subset (\mathcal{A}C)^\infty$.

Closedness of the linear image of sets

Theorem: Let $C \subset \mathbf{E}$ be closed and $\mathcal{A} \in L[\mathbf{E}, \mathbf{Y}]$. If $\ker \mathcal{A} \cap C^\infty = \{0\}$, then $\mathcal{A}C$ is closed and $(\mathcal{A}C)^\infty = \mathcal{A}C^\infty$ although, in general, we only have $\mathcal{A}C^\infty \subset (\mathcal{A}C)^\infty$.

Proof: First suppose that $\{x_k\} \subset C$ is unbounded. We show that $\ker \mathcal{A} \cap C^\infty = \{0\}$ implies that $\{\mathcal{A}x_k\}$ is also unbounded. Indeed, WLOG $x_k / \|x_k\| \to d \in C^\infty$ with $\|d\| = 1$. If $\{\mathcal{A}x_k\}$ is bounded, then $0 = \mathcal{A}(x_k / \|x_k\|) = \mathcal{A}d$ giving the contradiction $d \in \ker \mathcal{A} \cap C^\infty$.
Next let $y \in \operatorname{cl} \mathcal{A}C$ so $\exists \{x_k\} \subset C$ s.t. $y_k = \mathcal{A}x_k \to y$. By what we have just shown, $\{x_k\}$ must be bounded so WLOG $\exists, x \in C$ such that $x_k \to x$ so $y = \mathcal{A}x \in \mathcal{A}C$.

Closedness of the linear image of sets

Theorem: Let $C \subset \mathbf{E}$ be closed and $\mathcal{A} \in L[\mathbf{E}, \mathbf{Y}]$. If $\ker \mathcal{A} \cap C^\infty = \{0\}$, then $\mathcal{A}C$ is closed and $(\mathcal{A}C)^\infty = \mathcal{A}C^\infty$ although, in general, we only have $\mathcal{A}C^\infty \subset (\mathcal{A}C)^\infty$.

Proof: First suppose that $\{x_k\} \subset C$ is unbounded. We show that $\ker \mathcal{A} \cap C^\infty = \{0\}$ implies that $\{\mathcal{A}x_k\}$ is also unbounded. Indeed, WLOG $x_k / \|x_k\| \to d \in C^\infty$ with $\|d\| = 1$. If $\{\mathcal{A}x_k\}$ is bounded, then $0 = \mathcal{A}(x_k / \|x_k\|) = \mathcal{A}d$ giving the contradiction $d \in \ker \mathcal{A} \cap C^\infty$.
Next let $y \in \operatorname{cl} \mathcal{A}C$ so $\exists \{x_k\} \subset C$ s.t. $y_k = \mathcal{A}x_k \to y$. By what we have just shown, $\{x_k\}$ must be bounded so WLOG $\exists, x \in C$ such that $x_k \to x$ so $y = \mathcal{A}x \in \mathcal{A}C$.
Clearly, $\mathcal{A}C^\infty \subset (\mathcal{A}C)^\infty$. The reverse inclusion uses $\ker \mathcal{A} \cap C^\infty = \{0\}$. Let $y \in (\mathcal{A}C)^\infty$,i.e., $\exists \{x_k\} \subset C,\ t_k \downarrow 0$ s.t. $t_k \mathcal{A}x_k \to y$. We have shown that this implies that $\{t_k x_k\}$ is bounded, so WLOG $t_k x_k \to d \in C^\infty$, i.e., $y \in \mathcal{A}C^\infty$.

Closure of the sum of sets

Corollary: Let $C_i \subset \mathbf{E}$ $(i = 1, \ldots, k)$ be closed. If

$$[0 = \sum_{i=1}^{k} d_i,\ d_i \in C_i^\infty\ i = 1, \ldots, k] \implies [d_i = 0\ i = 1, \ldots, k],$$

then $\sum_{i=1}^{k} C_i$ is closed.

Closure of the sum of sets

Corollary: Let $C_i \subset \mathbf{E}$ $(i = 1, \ldots, k)$ be closed. If

$$[0 = \sum_{i=1}^{k} d_i,\ d_i \in C_i^\infty\ i = 1, \ldots, k] \implies [d_i = 0\ i = 1, \ldots, k],$$

then $\sum_{i=1}^{k} C_i$ is closed.

Proof: Let $\mathbf{X} := \prod_{i=1}^{k} \mathbf{E}$, $C = \prod_{i=1}^{k} C_i$, and $\mathcal{A} \in L[\mathbf{X}, \mathbf{E}]$ be given by $\mathcal{A}(x_1, \ldots, x_k) = \sum_{i=1}^{k} x_i$. Then
$\ker \mathcal{A} = \left\{ (d_1, \ldots, d_k) \,\middle|\, 0 = \sum_{i=1}^{k} d_i \right\}$ and $C^\infty = \prod_{i=1}^{k} C_i^\infty$.
Hence, the result follows from the theorem.

Horizon Functions

Given $f : \mathbf{E} \to \overline{\mathbf{R}}$, we define the *horizon function* for f to be the function $f^\infty : \mathbf{E} \to \overline{\mathbf{R}}$ satisfying $\operatorname{epi} f^\infty = (\operatorname{epi} f)^\infty$. An exceptional case occurs when $f \equiv +\infty$, in this case $\operatorname{epi} f = \emptyset$ so $(\operatorname{epi} f)^\infty = \{(0,0\}$ which is not the epigraph of a function.

Lemma: The proper lsc function $f : \mathbf{E} \to \overline{\mathbf{R}}$ is coercive iff $f^\infty(d) > 0 \; \forall \, d \in \mathbf{E}$. If f is cvx, the requirement that f be proper lsc can be omitted.

Horizon Functions and the Perspective map f^π

Theorem: For any $f : \mathbf{E} \to \overline{\mathbf{R}}$ with $f \not\equiv +\infty$, f^∞ is positively homogeneous with

$$(\star) \qquad f^\infty(d) = \liminf_{\substack{u \to d \\ \lambda \downarrow 0}} f^\pi(u, \lambda),$$

where f^π is the perspective map for f, i.e., $f^\pi(u, \lambda) = \lambda f(u/\lambda)$ for $\lambda > 0$.

If f is convex, then f^∞ is sublinear, and if f is closed proper cvx, then, for every $\bar{x} \in \operatorname{dom} f$,

$$(\star\star) \qquad f^\infty(d) = \lim_{\tau \uparrow \infty} \frac{f(\bar{x} + \tau d) - f(\bar{x})}{\tau} = \sup_{\tau > 0} \frac{f(\bar{x} + \tau d) - f(\bar{x})}{\tau}.$$

Horizon Functions and the Perspective map f^π

Proof: By definition, f^∞ is lsc and pos. homog. . Since

$f^\infty(d) = \inf\{\mu \,|\, \lambda_k \downarrow 0,\ \lambda_k(x_k, \mu_k) \to (d, \mu),\ (x_k, \mu_k) \in (\text{epi}\, f)\, \forall\, k\}$,

$f^\infty(d)$ is the inf of the values μ for which

$$\exists\, \lambda_k \downarrow 0,\, u_k \to d \text{ s.t. } \lambda_k f(u_k/\lambda_k) \to \mu$$

giving $(\star)$.

In the cvx case, we have shown that $\frac{f(\bar{x}+\tau d)-f(\bar{x})}{\tau}$ is a nondecreasing function of $\tau > 0$ for all $x \in \text{dom}\, f$. Hence, the supremum in $(\star\star)$ exists so that $(\star\star)$ follows from $(\star)$.

$f^\infty = \delta^*_{\mathrm{dom}\, f^*}$

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be closed proper cvx. Then

$$f^\infty = \delta^*_{\mathrm{dom}\, f^*} \quad \text{and} \quad (f^*)^\infty = \delta^*_{\mathrm{dom}\, f} \ .$$

Proof: Since $f^{**} = f$, we need only show $(f^*)^\infty = \delta^*_{\mathrm{dom}\, f}$.
Given $v \in \mathrm{dom}\, f^*$, $d \in \mathbf{E}$ and $\tau > 0$,

$$f^*(v + \tau d) = \sup_{x \in \mathrm{dom}\, f} [\langle v + \tau d, x\rangle - f(x)]$$

$$\leq \sup_{x \in \mathrm{dom}\, f} [\langle v, x\rangle - f(x)] + \tau \sup_{x \in \mathrm{dom}\, f} \langle v, d\rangle \quad = f^*(v) + \tau \delta^*_{\mathrm{dom}\, f}(d).$$

Hence,

$$(f^*)^\infty(d) = \sup_{\tau > 0} [\tfrac{f^*(v+\tau d) - f^*(v)}{\tau}] \leq \delta^*_{\mathrm{dom}\, f}(d).$$

On the other hand, if $(f^*)^\infty(d) \leq \beta$, then, for all $v \in \mathrm{dom}\, f^*$, $f^*(v + \tau d) \leq f^*(v) + \tau\beta \ \forall\, \tau > 0$. Hence, $\forall\, x \in \mathbf{E}$,

$$f(x) \geq \langle v + \tau d, x\rangle - f^*(v + \tau d)$$

$$\geq \langle v, x\rangle - f^*(v) + \tau(\langle d, x\rangle - \beta) \quad \forall\, \tau > 0\,.$$

Hence, for all $x \in \mathrm{dom}\, f$, $\langle d, x\rangle \leq \beta$ so that $\delta^*_{\mathrm{dom}\, f}(d) \leq \beta$ giving $\delta^*_{\mathrm{dom}\, f}(d) \leq (f^*)^\infty(d)$.

The horizon cone of f

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be closed proper cvx. Then there is a nonempty cvx cone $K \subset \mathbf{E}$ such that $K = (\mathrm{lev}_f(\alpha))^\infty$ for all $\alpha \geq \inf f$.

Proof: Let $\inf f \leq \alpha_1 \leq \alpha_2$. Cleary, $\mathrm{lev}_f(\alpha_1)^\infty \subset \mathrm{lev}_f(\alpha_2)^\infty$, so we show the reverse inclusion.
Let $d \in \mathrm{lev}_f(\alpha_2)^\infty$, $\lambda \in (0,1)$, $t > 0$, $x_i \in \mathrm{lev}_f(\alpha_i)$ $i = 1,2$ and set $\mu := \frac{\lambda}{(1-\lambda)}t$. Then, $f(\lambda x_1 + (1-\lambda)(x_2 + \mu d)) \leq \lambda\alpha_1 + (1-\lambda)\alpha_2$ and so

$$\begin{aligned} f(\lambda(x_1 + td) + (1-\lambda)x_2) &= f(\lambda x_1 + (1-\lambda)(x_2 + \frac{\lambda}{(1-\lambda)}td)) \\ &= f(\lambda x_1 + (1-\lambda)(x_2 + \mu d)) \\ &\leq \lambda\alpha_1 + (1-\lambda)\alpha_2. \end{aligned}$$

Since f is lsc, we can take the limit as $\lambda \uparrow 1$ to obtain $f(x_1 + td) \leq \alpha_1$. Since $t > 0$ was arbitrarily chosen, we obtain the result.

The horizon cone of f

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be closed proper cvx. Then there is a nonempty cvx cone $K \subset \mathbf{E}$ such that $K = (\mathrm{lev}_f(\alpha))^\infty$ for all $\alpha \geq \inf f$.

Proof: Let $\inf f \leq \alpha_1 \leq \alpha_2$. Cleary, $\mathrm{lev}_f(\alpha_1)^\infty \subset \mathrm{lev}_f(\alpha_2)^\infty$, so we show the reverse inclusion.
Let $d \in \mathrm{lev}_f(\alpha_2)^\infty$, $\lambda \in (0,1)$, $t > 0$, $x_i \in \mathrm{lev}_f(\alpha_i)$ $i = 1, 2$ and set $\mu := \frac{\lambda}{(1-\lambda)} t$. Then, $f(\lambda x_1 + (1-\lambda)(x_2 + \mu d)) \leq \lambda\alpha_1 + (1-\lambda)\alpha_2$ and so

$$\begin{aligned} f(\lambda(x_1 + td) + (1-\lambda)x_2) &= f(\lambda x_1 + (1-\lambda)(x_2 + \frac{\lambda}{(1-\lambda)} td)) \\ &= f(\lambda x_1 + (1-\lambda)(x_2 + \mu d)) \\ &\leq \lambda\alpha_1 + (1-\lambda)\alpha_2. \end{aligned}$$

Since f is lsc, we can take the limit as $\lambda \uparrow 1$ to obtain $f(x_1 + td) \leq \alpha_1$. Since $t > 0$ was arbitrarily chosen, we obtain the result.

We call $K := \mathrm{hzn}\, f$ the *horizon cone* of f.

$\operatorname{hzn}(f^*) = (\mathbf{R}_+ \operatorname{dom} f)^\circ$

Lemma: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be proper cvx. Then $\operatorname{hzn} f^* = (\mathbf{R}_+ \operatorname{dom} f)^\circ$ and $(\operatorname{hzn} f^*)^\circ = \operatorname{cl} \mathbf{R}_+ \operatorname{dom} f$.

Proof: Let $\alpha \geq \inf f^*$ and $\overline{w} \in \operatorname{dom} f^*$. Then

$$\begin{aligned}
\operatorname{hzn} f^* &= \{ w \,|\, f^*(w) \leq \alpha \}^\infty \\
&= \{ v \,|\, f^*(\overline{w} + tv) \leq \alpha \ \forall\, t > 0 \} \\
&= \{ v \,|\, \langle \overline{w} + tv, x \rangle - f(x) \leq \alpha \ \forall\, x \in \operatorname{dom} f,\ t > 0 \} \\
&= \left\{ v \,\middle|\, \langle v, x \rangle \leq t^{-1}(\alpha + f(x) - \langle \overline{w}, x \rangle) \ \forall\, x \in \operatorname{dom} f,\ t > 0 \right\} \\
&= \{ v \,|\, \langle v, x \rangle \leq 0 \ \forall\, x \in \operatorname{dom} f \} \\
&= \{ v \,|\, \langle v, x \rangle \leq 0 \ \forall\, x \in \mathbf{R}_+ \operatorname{dom} f \} \\
&= (\mathbf{R}_+ \operatorname{dom} f)^\circ.
\end{aligned}$$

Convexity of Compositions

Theorem: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be closed proper cvx and $G : \mathbf{Y} \to \mathbf{E}$ be concave with respect to hzn f, i.e.,

$$G(\lambda y_1 + (1-\lambda) y_2) - [\lambda G(y_1) + (1-\lambda) G(y_2)] \in \operatorname{hzn} f \;\; \forall\, y_1, y_2 \in \mathbf{Y}\; \lambda \in [0,1].$$

Then f is non-increasing relative to hzn f, i.e.,

$$f(x+w) \leq f(x) \quad \text{whenever } w \in \operatorname{hzn} f,$$

and $f \circ G$ is convex on $\operatorname{dom} f \circ G = \{ y \,|\, G(y) \in \operatorname{dom} f \}$.

Convexity of Compositions

Proof: Let $x \in \mathbf{E}$ be such that $w \in \operatorname{hzn} f = (\mathbf{R}_+ \operatorname{dom} f^*)^\circ$. Since $f = f^{**}$,

$$\begin{aligned} f(x+w) &= \sup_{v \in \operatorname{dom} f^*} \langle x+w, v\rangle - f^*(v) \\ &= \sup_{v \in \operatorname{dom} f^*} \langle x, v\rangle - f^*(v) + \langle w, v\rangle \\ &\le \sup_{v \in \operatorname{dom} f^*} [\langle x, v\rangle - f^*(v)] \qquad (\text{since } \langle w, v\rangle \le 0) \\ &= f(x), \end{aligned}$$

so f is non-increasing relative to $\operatorname{hzn} f$. Hence, for $y_1, y_2 \in \operatorname{dom} f \circ G$, $\lambda \in [0,1]$, $x = \lambda G(y_1) + (1-\lambda) G(y_2)$ and $w = G(\lambda y_1 + (1-\lambda) y_2) - (\lambda G(y_1) + (1-\lambda) G(y_2))$,

$$\begin{aligned} (f \circ G)((1-\lambda) y_1 + \lambda y_2) &= f(x+w) \\ &\le f(x) \\ &= (1-\lambda)(f\circ G)(y_1) + \lambda (f \circ G)(y_2). \end{aligned}$$

Closedness of the linear image of epigraphs

Lemma: Let $f : \mathbf{E} \to \overline{\mathbf{R}}$ be proper cvx and $L \subset \mathbf{E}$ a subspace. Then

$$L \cap \mathrm{ri}\,\mathrm{dom}\, f \neq \emptyset \iff L^\perp \cap \mathrm{lev}_{\delta^*_{dom\, f}}(0) = \{0\}$$
$$\iff L^\perp \cap \mathrm{lev}_{(f^*)^\infty}(0) = \{0\}.$$

Proof: We prove the equivalence of the negation. Observe that $L \cap \mathrm{ri}\,\mathrm{dom}\, f = \emptyset$ iff $0 \notin (\mathrm{ri}\,\mathrm{dom}\, f) - L = \mathrm{ri}\,(\mathrm{dom}\, f - L)$. The separation theorem tells us that

$$\begin{aligned} 0 \notin \mathrm{ri}\,(\mathrm{dom}\, f - L) &\iff \exists\, v \text{ s.t. } \langle v, x - w\rangle < 0\ \forall\, x \in \mathrm{ri}\,\mathrm{dom}\, f,\ w \in L \\ &\iff \exists\, v \text{ s.t. } \langle v, x\rangle < \langle v, w\rangle\ \forall\, x \in \mathrm{ri}\,\mathrm{dom}\, f,\ w \in L \\ &\iff \exists\, v \in L^\perp \text{ s.t. } \langle v, x\rangle < 0\ \forall\, x \in \mathrm{ri}\,\mathrm{dom}\, f \\ &\iff \exists\, v \in L^\perp, 0 \neq v \text{ s.t. } \delta^*_{\mathrm{dom}\, f}(v) \leq 0. \end{aligned}$$