CALIFORNIA INSTITUTE OF TECHNOLOGY
Division of the Humanities and Social Sciences

# Convex analysis and profit/cost/support functions

KC Border

October 2004
Revised January 2009
v. 2012.02.10::11.34

Let $A$ be a subset of $\boldsymbol{R}^{\mathrm{m}}$. Convex analysts may give one of two definitions for the **support function** of $A$ as either an infimum or a supremum. Recall that the **supremum** of a set of real numbers is its least upper bound and the **infimum** is its greatest lower bound. If $A$ has no upper bound, then by convention $\sup A = \infty$ and if $A$ has no lower bound, then $\inf A = -\infty$. For the empty set, $\sup A = -\infty$ and $\inf A = \infty$; otherwise $\inf A \leqslant \sup A$. (This makes a kind of sense: Every real number $\lambda$ is an upper bound for the empty set, since there is no member of the empty set that is greater than $\lambda$. Thus the least upper bound must be $-\infty$. Similarly, every real number is also lower bound, so the infimum is $\infty$.) Thus support functions (as infima or suprema) may assume the values $\infty$ and $-\infty$.

By convention, $0 \cdot \infty = 0$; if $\lambda > 0$ is a real number, then $\lambda \cdot \infty = \infty$ and $\lambda \cdot (-\infty) = -\infty$; and if $\lambda < 0$ is a real number, then $\lambda \cdot \infty = -\infty$ and $\lambda \cdot (-\infty) = \infty$. These conventions are used to simplify statements involving positive homogeneity.

Rather than choose one definition, I shall give the two definitions different names based on their economic interpretation.

### Profit maximization

The **profit function** $\pi_A$ of $A$ is defined by

$$\pi_A(p) = \sup_{y \in A} p \cdot y.$$

Clearly,

$$\pi_A(p) = -c_A(-p).$$

**Proposition** *$\pi_A$ is convex, lower semicontinuous, and positively homogeneous of degree 1.*

Positive homogeneity of $\pi_A$ is obvious given the conventions on multiplication of infinities. To see that it is convex, let $g_x$ be the linear (hence convex) function defined by $g_x(p) = x \cdot p$. Then $\pi_A(p) = \sup_{x \in A} g_x(p)$. Since the pointwise supremum of a family of convex functions is convex, $\pi_A$ is convex. Also each $g_x$ is continuous, hence lower semicontinuous, and the supremum of a family of lower semicontinuous functions is lower semicontinuous. See my notes on maximization.

### Cost minimization

The **cost function** $c_A$ of $A$ is defined by

$$c_A(p) = \inf_{y \in A} p \cdot y.$$

Clearly

$$c_A(p) = -\pi_A(-p).$$

**Proposition** *$c_A$ is concave, upper semicontinuous, and positively homogeneous of degree 1.*

Positive homogeneity of $c_A$ is obvious given the conventions on multiplication of infinities. To see that it is concave, let $g_x$ be the linear (hence concave) function defined by $g_x(p) = x \cdot p$. Then $c_A(p) = \inf_{x \in A} g_x(p)$. Since the pointwise infimum of a family of concave functions is concave, $c_A$ is concave. Also each $g_x$ is continuous, hence upper semicontinuous, and the infimum of a family of upper semicontinuous functions is upper semicontinuous. See my notes on maximization.

**Proposition** *The set*

$$\{p \in \boldsymbol{R}^{\mathrm{m}} : \pi_A(p) < \infty\}$$

*is a closed convex cone, called the* ***effective domain*** *of* $\pi_A$*, and denoted* $\operatorname{dom} \pi_A$.

**Proposition** *The set*

$$\{p \in \boldsymbol{R}^{\mathrm{m}} : c_A(p) > -\infty\}$$

*is a closed convex cone, called the* ***effective domain*** *of* $c_A$*, and denoted* $\operatorname{dom} c_A$.

The effective domain will always include the point 0 provided $A$ is nonempty. By convention $\pi_\varnothing(p) = -\infty$ for all $p$, and we say that $\pi_A$ is **improper**. If $A = \boldsymbol{R}^{\mathrm{m}}$, then 0 is the only point in the effective domain of $\pi_A$.

The effective domain will always include the point 0 provided $A$ is nonempty. By convention $c_\varnothing(p) = \infty$ for all $p$, and we say that $c_\varnothing$ is **improper**. If $A = \boldsymbol{R}^{\mathrm{m}}$, then 0 is the only point in the effective domain of $c_A$.

It is easy to see that the effective domain $\operatorname{dom} \pi_A$ of $\pi_A$ is a cone, that is, if $p \in \operatorname{dom} \pi_A$, then $\lambda p \in \operatorname{dom} \pi_A$ for every $\lambda \geqslant 0$. (Note that $\{0\}$ is a (degenerate) cone.)

It is easy to see that the effective domain $\operatorname{dom} c_A$ of $c_A$ is a cone, that is, if $p \in \operatorname{dom} c_A$, then $\lambda p \in \operatorname{dom} c_A$ for every $\lambda \geqslant 0$. (Note that $\{0\}$ is a (degenerate) cone.)

It is also straightforward to show that $\operatorname{dom} \pi_A$ is convex. For if $\pi_A(p) < \infty$ and $\pi_A(q) < \infty$, for $0 \leqslant \lambda \leqslant 1$, by convexity of $\pi_A$, we have

$$\begin{aligned} \pi_A\big(\lambda x + (1-\lambda) y\big) &\leqslant \lambda \pi_A(p) + (1-\lambda)\pi_A(q) \\ &< \infty. \end{aligned}$$

The closedness of $\operatorname{dom} \pi_A$ is more difficult.

It is also straightforward to show that $\operatorname{dom} c_A$ is convex. For if $c_A(p) > -\infty$ and $c_A(q) > -\infty$, for $0 \leqslant \lambda \leqslant 1$, by concavity of $c_A$, we have

$$\begin{aligned} c_A\big(\lambda x + (1-\lambda) y\big) &\geqslant \lambda c_A(p) + (1-\lambda) c_A(q) \\ &> -\infty. \end{aligned}$$

The closedness of $\operatorname{dom} c_A$ is more difficult.

## Recoverability

**Separating Hyperplane Theorem** *If $A$ is a closed convex set, and $x$ does not belong to $A$, then there is a nonzero $p$ satisfying*

$$p \cdot x > \pi_A(p).$$

For a proof see my notes. Note that given our conventions, $A$ may be empty. From this theorem we easily get the next proposition.

**Separating Hyperplane Theorem** *If $A$ is a closed convex set, and $x$ does not belong to $A$, then there is a nonzero $p$ satisfying*

$$p \cdot x < c_A(p).$$

For a proof see my notes. Note that given our conventions, $A$ may be empty. From this theorem we easily get the next proposition.

**Proposition** *The closed convex hull $\overline{\mathrm{co}}\, A$ of satisfies*

$$\overline{\mathrm{co}}\, A = \big\{y \in \boldsymbol{R}^{\mathrm{m}} : (\forall p \in \boldsymbol{R}^{\mathrm{m}})\ [p \cdot y \leqslant \pi_A(p)]\big\}.$$

**Proposition** *The closed convex hull $\overline{\mathrm{co}}\, A$ of satisfies*

$$\overline{\mathrm{co}}\, A = \big\{y \in \boldsymbol{R}^{\mathrm{m}} : (\forall p \in \boldsymbol{R}^{\mathrm{m}})\ [p \cdot y \geqslant c_A(p)]\big\}.$$

Now let $f$ be a continuous real-valued function defined on a closed convex cone $D$. We can extend $f$ to all of $\boldsymbol{R}^{\mathrm{m}}$ by setting it to $\infty$ outside of $D$ if $f$ is convex or $-\infty$ if $f$ is concave.

**Proposition** *If $f$ is convex and positively homogeneous of degree 1, define*

$$A = \big\{y \in \boldsymbol{R}^{\mathrm{m}} : (\forall p \in \boldsymbol{R}^{\mathrm{m}})\ [p \cdot y \leqslant f(p)]\big\}.$$

*Then $A$ is closed and convex and*

$$f = \pi_A.$$

**Proposition** *If $f$ is concave and positively homogeneous of degree 1, define*

$$A = \big\{y \in \boldsymbol{R}^{\mathrm{m}} : (\forall p \in \boldsymbol{R}^{\mathrm{m}})\ [p \cdot y \geqslant f(p)]\big\}.$$

*Then $A$ is closed and convex and*

$$f = c_A.$$

## Extremizers are subgradients

**Proposition** *If $\tilde{y}(p)$ maximizes $p$ over $A$, that is, if $\tilde{y}(p)$ belongs to $A$ and $p \cdot \tilde{y}(p) \geqslant p \cdot y$ for all $y \in A$, then $\tilde{y}(p)$ is a subgradient of $\pi_A$ at $p$. That is,*

$$\pi_A(p) + \tilde{y}(p) \cdot (q - p) \leqslant \pi_A(q) \qquad (*)$$

*for all $q \in \boldsymbol{R}^{\mathrm{m}}$.*

To see this, note that for any $q \in \boldsymbol{R}^{\mathrm{m}}$, by definition we have

$$q \cdot \tilde{y}(p) \leqslant \pi_A(q).$$

Now add $\pi_A(p) - p \cdot \tilde{y}(p) = 0$ to the left hand side to get the subgradient inequality.

Note that $\pi_A(p)$ may be finite for a closed convex set $A$, and yet there may be no maximizer. For instance, let

$$A = \{(x, y) \in \boldsymbol{R}^2 : x < 0,\ y < 0,\ xy \geqslant 1\}.$$

Then for $p = (1, 0)$, we have $\pi_A(p) = 0$ as $(1, 0) \cdot (-1/n, -n) = -1/n$, but $(1, 0) \cdot (x, y) = x < 0$ for each $(x, y) \in A$. Thus there is no maximizer in $A$.

It turns out that if there is no maximizer of $p$, then $\pi_A$ has no subgradient at $p$. In fact, the following is true, but I won't present the proof, which relies on the Separating Hyperplane Theorem. (See my notes for a proof.)

**Theorem** *If $A$ is closed and convex, then $x$ is a subgradient of $\pi_A$ at $p$ if and only if $x \in A$ and $x$ maximizes $p$ over $A$.*

**Proposition** *If $\hat{y}(p)$ minimizes $p$ over $A$, that is, if $\hat{y}(p)$ belongs to $A$ and $p \cdot \hat{y}(p) \leqslant p \cdot y$ for all $y \in A$, then $\hat{y}(p)$ is a supergradient of $c_A$ at $p$. That is,*

$$c_A(p) + \hat{y}(p) \cdot (q - p) \geqslant c_A(q) \qquad (*)$$

*for all $q \in \boldsymbol{R}^{\mathrm{m}}$.*

To see this, note that for any $q \in \boldsymbol{R}^{\mathrm{m}}$, by definition we have

$$q \cdot \hat{y}(p) \geqslant c_A(q).$$

Now add $c_A(p) - p \cdot \hat{y}(p) = 0$ to the left hand side to get the supergradient inequality.

Note that $x_A(p)$ may be finite for a closed convex set $A$, and yet there may be no minimizer. For instance, let

$$A = \{(x, y) \in \boldsymbol{R}^2 : x > 0,\ y > 0,\ xy \geqslant 1\}.$$

Then for $p = (1, 0)$, we have $\pi_A(p) = 0$ as $(1, 0) \cdot (1/n, n) = 1/n$, but $(1, 0) \cdot (x, y) = x > 0$ for each $(x, y) \in A$. Thus there is no minimizer in $A$.

It turns out that if there is no minimizer of $p$, then $c_A$ has no supergradient at $p$. In fact, the following is true, but I won't present the proof, which relies on the Separating Hyperplane Theorem. (See my notes for a proof.)

**Theorem** *If $A$ is closed and convex, then $x$ is a supergradient of $c_A$ at $p$ if and only if $x \in A$ and $x$ minimizes $p$ over $A$.*

## Comparative statics

**Proposition** *Consequently, if $A$ is closed and convex, and $\tilde{y}(p)$ is the unique maximizer of $p$ over $A$, then $\pi_A$ is differentiable at $p$ and*

$$\tilde{y}(p) = \pi_A'(p). \qquad (**)$$

**Proposition** *Consequently, if $A$ is closed and convex, and $\hat{y}(p)$ is the unique minimizer of $p$ over $A$, then $c_A$ is differentiable at $p$ and*

$$\hat{y}(p) = c_A'(p). \qquad (**)$$

To see that differentiability of $\pi_A$ implies the profit maximizer is unique, consider $q$ of the form $p \pm \lambda e^i$, where $e^i$ is the $i^{\text{th}}$ unit coordinate vector, and $\lambda > 0$.

The subgradient inequality for $q = p + \lambda e^i$ is

$$\tilde{y}(p) \cdot \lambda e^i \leqslant \pi_A(p + \lambda e^i) - \pi_A(p)$$

and for $q = p - \lambda e^i$ is

$$-\tilde{y}(p) \cdot \lambda e^i \leqslant \pi_A(p - \lambda e^i) - \pi_A(p).$$

Dividing these by $\lambda$ and $-\lambda$ respectively yields

$$y_i^*(p) \leqslant \frac{\pi_A(p + \lambda e^i) - \pi_A(p)}{\lambda}$$
$$y_i^*(p) \geqslant \frac{\pi_A(p - \lambda e^i) - \pi_A(p)}{\lambda}.$$

so

$$\tfrac{\pi_A(p-\lambda e^i)-\pi_A(p)}{\lambda} \leqslant y_i^*(p) \leqslant \tfrac{\pi_A(p+\lambda e^i)-\pi_A(p)}{\lambda}.$$

Letting $\lambda \downarrow 0$ yields $\tilde{y}_i(p) = D_i \pi_A(p)$.

**Proposition** *Thus if $\pi_A$ is twice differentiable at $p$, that is, if the maximizer $\tilde{y}(p)$ is differentiable with respect to $p$, then the $i^{\text{th}}$ component satisfies*

$$D_j \tilde{y}_i(p) = D_{ij} \pi_A(p). \tag{***}$$

Consequently, the matrix

$$\big[D_j \tilde{y}_i(p)\big]$$

is positive semidefinite.

In particular,

$$D_i \tilde{y}_i \geqslant 0.$$

Even without twice differentiability, from the subgradient inequality, we have

$$\pi_A(p) + \tilde{y}(p) \cdot (q - p) \leqslant \pi_A(q)$$
$$\pi_A(q) + \tilde{y}(q) \cdot (p - q) \leqslant \pi_A(p)$$

so adding the two inequalities, we get

$$\big(\tilde{y}(p) - \tilde{y}(q)\big) \cdot (p - q) \geqslant 0.$$

To see that differentiability of $c_A$ implies the cost minimizer is unique, consider $q$ of the form $p \pm \lambda e^i$, where $e^i$ is the $i^{\text{th}}$ unit coordinate vector, and $\lambda > 0$.

The supergradient inequality for $q = p + \lambda e^i$ is

$$\hat{y}(p) \cdot \lambda e^i \geqslant c_A(p + \lambda e^i) - c_A(p)$$

and for $q = p - \lambda e^i$ is

$$-\hat{y}(p) \cdot \lambda e^i \geqslant c_A(p - \lambda e^i) - c_A(p).$$

Dividing these by $\lambda$ and $-\lambda$ respectively yields

$$y_i^*(p) \geqslant \frac{c_A(p + \lambda e^i) - c_A(p)}{\lambda}$$
$$y_i^*(p) \leqslant \frac{c_A(p - \lambda e^i) - c_A(p)}{\lambda}.$$

so

$$\tfrac{c_A(p+\lambda e^i)-c_A(p)}{\lambda} \leqslant y_i^*(p) \leqslant \tfrac{c_A(p-\lambda e^i)-c_A(p)}{\lambda}.$$

Letting $\lambda \downarrow 0$ yields $\hat{y}_i(p) = D_i c_A(p)$.

**Proposition** *Thus if $c_A$ is twice differentiable at $p$, that is, if the minimizer $\hat{y}(p)$ is differentiable with respect to $p$, then the $i^{\text{th}}$ component satisfies*

$$D_j \hat{y}_i(p) = D_{ij} c_A(p). \tag{***}$$

Consequently, the matrix

$$\big[D_j \hat{y}_i(p)\big]$$

is negative semidefinite.

In particular,

$$D_i \hat{y}_i \leqslant 0.$$

Even without twice differentiability, from the supergradient inequality, we have

$$c_A(p) + \hat{y}(p) \cdot (q - p) \geqslant c_A(q)$$
$$c_A(q) + \hat{y}(q) \cdot (p - q) \geqslant c_A(p)$$

so adding the two inequalities, we get

$$\big(\hat{y}(p) - \hat{y}(q)\big) \cdot (p - q) \leqslant 0.$$

**Proposition** *Thus if $q$ differs from $p$ only in its $i^{\text{th}}$ component, say $q_i = p_i + \triangle p_i$, then we have*

$$\triangle \tilde{y}_i \triangle p_i \geqslant 0.$$

*Dividing by the positive quantity $(\triangle p_i)^2$ does not change this inequality, so*

$$\frac{\triangle \tilde{y}_i}{\triangle p_i} \geqslant 0.$$

**Proposition** *Thus if $q$ differs from $p$ only in its $i^{\text{th}}$ component, say $q_i = p_i + \triangle p_i$, then we have*

$$\triangle \hat{y}_i \triangle p_i \leqslant 0.$$

*Dividing by the positive quantity $(\triangle p_i)^2$ does not change this inequality, so*

$$\frac{\triangle \hat{y}_i}{\triangle p_i} \leqslant 0.$$

### Cyclical monotonicity and empirical restrictions

A real function $g \colon X \subset \boldsymbol{R} \to \boldsymbol{R}$ is **increasing** if

$$x \geqslant y \implies g(x) \geqslant g(y).$$

That is, $g(x) - g(y)$ and $x - y$ have the same sign. An equivalent way to say this is

$$\big(g(x) - g(y)\big)\big(x - y\big) \geqslant 0 \qquad \text{for all } x, y,$$

which can be rewritten as

$$g(x)(y - x) + g(y)(x - y) \leqslant 0 \qquad \text{for all } x, y.$$

We can generalize this to a function $g$ from $\boldsymbol{R}^{\mathrm{m}}$ into $\boldsymbol{R}^{\mathrm{m}}$ like this:

**Definition** *A function $g \colon X \subset \boldsymbol{R}^{\mathrm{m}} \to \boldsymbol{R}^{\mathrm{m}}$ is* ***monotone (increasing)*** *if*

$$g(x) \cdot (y - x) + g(y) \cdot (x - y) \leqslant 0$$

*for all $x, y \in X$.*

We have already seen that the (sub)gradient of $\pi_A$ is monotone (increasing).

More is true.

**Definition** *A mapping $g \colon X \subset \boldsymbol{R}^{\mathrm{m}} \to \boldsymbol{R}^{\mathrm{m}}$ is* ***cyclically monotone (increasing)*** *if for every cycle $x_0, x_1, \ldots, x_n, x_{n+1} = x_0$ in $X$, we have*

$$g(x_0) \cdot (x_1 - x_0) + \cdots + g(x_n) \cdot (x_{n+1} - x_n) \leqslant 0.$$

**Proposition** *If $f \colon X \subset \boldsymbol{R}^{\mathrm{m}} \to \boldsymbol{R}$ is convex, and $g \colon X \subset \boldsymbol{R}^{\mathrm{m}} \to \boldsymbol{R}^{\mathrm{m}}$ is a selection from the subdifferential of $f$, that is, if $g(x)$ is a subgradient of $f$ at $x$ for every $x$, then $g$ is cyclically monotone (increasing).*

A real function $g \colon X \subset \boldsymbol{R} \to \boldsymbol{R}$ is **decreasing** if

$$x \geqslant y \implies g(x) \leqslant g(y).$$

That is, $g(x) - g(y)$ and $x - y$ have the opposite sign. An equivalent way to say this is

$$\big(g(x) - g(y)\big)\big(x - y\big) \leqslant 0 \qquad \text{for all } x, y,$$

which can be rewritten as

$$g(x)(y - x) + g(y)(x - y) \geqslant 0 \qquad \text{for all } x, y.$$

We can generalize this to a function $g$ from $\boldsymbol{R}^{\mathrm{m}}$ into $\boldsymbol{R}^{\mathrm{m}}$ like this:

**Definition** *A function $g \colon X \subset \boldsymbol{R}^{\mathrm{m}} \to \boldsymbol{R}^{\mathrm{m}}$ is* ***monotone (decreasing)*** *if*

$$g(x) \cdot (y - x) + g(y) \cdot (x - y) \geqslant 0$$

*for all $x, y \in X$.*

We have already seen that the (super)gradient of $c_A$ is monotone (decreasing).

More is true.

**Definition** *A mapping $g \colon X \subset \boldsymbol{R}^{\mathrm{m}} \to \boldsymbol{R}^{\mathrm{m}}$ is* ***cyclically monotone (decreasing)*** *if for every cycle $x_0, x_1, \ldots, x_n, x_{n+1} = x_0$ in $X$, we have*

$$g(x_0) \cdot (x_1 - x_0) + \cdots + g(x_n) \cdot (x_{n+1} - x_n) \geqslant 0.$$

**Proposition** *If $f \colon X \subset \boldsymbol{R}^{\mathrm{m}} \to \boldsymbol{R}$ is concave, and $g \colon X \subset \boldsymbol{R}^{\mathrm{m}} \to \boldsymbol{R}^{\mathrm{m}}$ is a selection from the superdifferential of $f$, that is, if $g(x)$ is a supergradient of $f$ at $x$ for every $x$, then $g$ is cyclically monotone (decreasing).*

*Proof*: Let $x_0, x_1, \ldots, x_n, x_{n+1} = x_0$ be a cycle in $X$. From the subgradient inequality at $x_i$, we have

$$f(x_i) + g(x_i) \cdot (x_{i+1} - x_i) \leqslant f(x_{i+1})$$

or

$$g(x_i) \cdot (x_{i+1} - x_i) \leqslant f(x_{i+1}) - f(x_i)$$

for each $i = 0, \ldots, n$. Summing gives

$$\sum_{i=0}^{n} g(x_i) \cdot (x_{i+1} - x_i) \leqslant 0,$$

where the right-hand side takes into account $f(x_{n+1}) = f(x_0)$. ∎

**Corollary** *The profit maximizing points correspondence $\hat{y}$ is cyclically monotonic (increasing).*

Remarkably the converse is true.

**Theorem (Rockafellar)** *Let $X$ be a convex set in $\boldsymbol{R}^{\mathrm{m}}$ and let $\varphi \colon X \twoheadrightarrow \boldsymbol{R}^{\mathrm{m}}$ be a cyclically monotone (increasing) correspondence (that is, if every selection from $\varphi$ is a cyclically monotone (increasing) function). Then there is a lower semicontinuous convex function $f \colon X \to \boldsymbol{R}$ such that*

$$\varphi(x) \subset \partial f(x)$$

*for every $x$.*

*Proof*: Let $x_0, x_1, \ldots, x_n, x_{n+1} = x_0$ be a cycle in $X$. From the supergradient inequality at $x_i$, we have

$$f(x_i) + g(x_i) \cdot (x_{i+1} - x_i) \geqslant f(x_{i+1})$$

or

$$g(x_i) \cdot (x_{i+1} - x_i) \geqslant f(x_{i+1}) - f(x_i)$$

for each $i = 0, \ldots, n$. Summing gives

$$\sum_{i=0}^{n} g(x_i) \cdot (x_{i+1} - x_i) \geqslant 0,$$

where the right-hand side takes into account $f(x_{n+1}) = f(x_0)$. ∎

**Corollary** *The cost minimizing points correspondence $\hat{y}$ is cyclically monotonic (decreasing).*

Remarkably the converse is true.

**Theorem (Rockafellar)** *Let $X$ be a convex set in $\boldsymbol{R}^{\mathrm{m}}$ and let $\varphi \colon X \twoheadrightarrow \boldsymbol{R}^{\mathrm{m}}$ be a cyclically monotone (decreasing) correspondence (that is, if every selection from $\varphi$ is a cyclically monotone (decreasing) function). Then there is an upper semicontinuous concave function $f \colon X \to \boldsymbol{R}$ such that*

$$\varphi(x) \subset \partial f(x)$$

*for every $x$.*