On Information Functions and Their Polars[1]

F. PUKELSHEIM[2]

Communicated by L. Cesari

Abstract. Let $\mathscr{P}$ be a closed convex cone. Information functions, i.e., nonnegative functions on $\mathscr{P}$ which are positively homogeneous and concave, are shown to be in a one-to-one correspondence with certain convex subsets of $\mathscr{P}$. Information functions are always isotone with respect to the vector ordering induced by $\mathscr{P}$, and this order-preserving property distinguishes them from their convex analogues, gauge functions. A polarity concept for information functions is proposed which slightly deviates from the well-known polarity correspondence for gauge functions. Finally, those functions are characterized which differ from information functions only by some nondecreasing concave transformation.

1. Introduction and Summary

Information functions, as defined below, closely resemble their convex analogues, gauge functions, except that the effective domain of information functions is a closed convex cone $\mathscr{P}$, only, and except that information functions are always isotone with respect to the vector ordering $\leqq$ induced by $\mathscr{P}$. Hence, information functions may be of interest whenever a problem formulation implies a cone ordering.

The present paper originates from one such example in the statistical analysis of experimental design (Ref. 1); as a motivation, we outline briefly this background. Suppose that we wish to design an experiment in order to estimate, test, or otherwise investigate s unknown real parameters. Under suitable assumptions on the statistical model, the performance of a design

[1] The author would like to thank Professors L. Cesari and D. W. Jorgenson for drawing his attention to Refs. 7 and 8.

[2] Professor, Lehrstuhl für Stochastik und ihre Anwendungen, Universität Augsburg, Augsburg, Federal Republic of Germany.

ξ is reflected by an $s \times s$ information matrix $(n/\sigma^2)J(\xi)$; i.e., information is directly proportional to the size n of the sample, inversely proportional to the error variance σ^2, and otherwise given by a nonnegative definite matrix $J(\xi)$. The objective is to maximize a real functional j of $(n/\sigma^2)J(\xi)$. In order that j conforms with the notion of information (and rightly j is called an information function), we demand that j be (a) positive, i.e., a positive (or nonnegative) definite information matrix is mapped into a positive (or nonnegative) information number; (b) positively homogeneous, whence the common scalar factor n/σ^2 may be neglected; and (c) concave, so that information cannot be increased by interpolation. Convex analysis then allows one to characterize and study those designs which have maximal j-information for the s parameters under investigation.

Moreover, every information function is isotone with respect to the Loewner ordering $\leqq$, which is induced by the closed convex cone $\mathscr{P}$ of nonnegative definite matrices. A characterization of designs whose information matrix is maximal in the Loewner ordering is easily deduced from those results which pertain to maximizing j-information; admissible information matrices necessarily have maximal j-information, for certain functions j and for certain parameters; see Corollaries 5.2 and 8.4 in Ref. 1. These results point to a wider applicability of information functions beyond the mere statistical problem, inasmuch as determination of admissible (sometimes called efficient or Pareto-optimal) solutions plays a central role in the theory of vector optimization. Of course, existence and characterizations of admissible points may also be established by alternative methods; see Refs. 2 and 3. Yet, it appears natural to investigate a given vector ordering by means of a suitable family of scalar functionals (compare the approaches taken in Refs. 4 and 5), and the monotonicity behavior of information functions makes them prime candidates for this purpose.

Therefore, we have collected below what seems to us the basic properties of information functions. In Section 2, it is shown that information functions are in a one-to-one correspondence with those closed convex subsets of $\mathscr{P}$ which do not contain 0 and which recede in all directions of $\mathscr{P}$. The polarity concept for information functions leads to information functions on the dual cone $\mathscr{P}^d$, as detailed in Section 3. Information-like functions are discussed in the concluding Section 4. Although there is no counterpart for the cone $\mathscr{P}$ in the theory of gauge functions, our presentation will emphasize the common aspects of the convex and the concave case, in that exposition and terminology closely follow Section 15 of Rockafellar (Ref. 6) on polars of convex functions. Thus, our frame of reference is the Euclidean n-space; then, application to matrix space and generalization to paired spaces are straightforward.

Partial results when $\mathcal{P} = \mathbb{R}^n_+$ is the nonnegative quadrant have been obtained by Rockafellar (Ref. 7) and McFadden (Ref. 8). Rockafellar introduces monotone concave gauges and proves (Theorem 3.4) a polarity correspondence similar to our Theorem 3.1, while McFadden defines distance functions and establishes (Lemma 7.4) a function-set correspondence similar to our Theorem 2.1.

Finally, let us recall some notation. We write $\{j \geqq 1\}$ for the set $\{x \in \mathbb{R}^n \mid j(x) \geqq 1\}$. A convex set $C \subset \mathbb{R}^n$ is said to recede in the direction of $y \in \mathbb{R}^n$ if

$$x + \lambda y \in C, \qquad \text{for every } \lambda \geqslant 0 \text{ and } x \in C.$$

These directions form the recession cone 0^+C of C. A concave function is proper if it is not identically $-\infty$ and never attains the value $+\infty$; a proper concave function is said to be closed if it is upper semicontinuous. The polarity concepts for gauge and information functions are directed toward 0 and ∞, respectively; in Section 3, our notation seeks to express this orientation.

2. Information Functions

Let $\mathcal{P}$ be a closed convex cone in the Euclidean n-space $\mathbb{R}^n$. A real function j defined on $\mathcal{P}$ will be called an *information function on* $\mathcal{P}$, if j is

(a) nonnegative on $\mathcal{P}$ and positive on its relative interior, ri $\mathcal{P}$;
(b) positively homogeneous, $j(\lambda x) = \lambda j(x), \forall \lambda > 0, \forall x \in \mathcal{P}$;
(c) superadditive, $j(x + y) \geqq j(x) + j(y), \forall x, y \in \mathcal{P}$.

Outside $\mathcal{P}$, the value of $j(x)$ is taken to be $-\infty$; thus, j is a concave function on $\mathbb{R}^n$ with effective domain $\mathcal{P}$.

Let $x \leqq y$ denote the vector ordering induced by $\mathcal{P}$, i.e.,

$$x \leqq y \Leftrightarrow y - x \in \mathcal{P}.$$

As a consequence of the definition, every information function j is concave, isotone, and satisfies

$$j(0) = 0;$$

if j is strictly concave, then it is strictly isotone, i.e.,

$$x \leqq y \text{ and } x \neq y \text{ imply } j(x) < j(y).$$

The average of an arbitrary collection of information functions and the minimum of a finite collection again are information functions. Even the infimum of an arbitrary collection is an information function, provided it is positive on ri $\mathcal{P}$.

The convex analogue of information functions are gauge functions, i.e., real functions k on $\mathbb{R}^n$ which are nonnegative, positively homogeneous, and subadditive. Just as gauge functions k are in a one-to-one correspondence with their unit ball $\{k \leqq 1\}$, the same is true for information functions j and their *unit set* $\{j \geqq 1\}$. If j is a closed information function, then $\{j \geqq 1\}$ is a closed convex subset of $\mathcal{P}$ which does not contain 0 and which recedes in all directions of $\mathcal{P}$.

Conversely, define the function $\eta(\cdot \mid C)$ for a subset C of $\mathbb{R}^n$ by

$$\eta(x \mid C) = \begin{cases} \sup\{\mu \geqq 0 \mid \mu = 0 \text{ or } x \in \mu C\}, & \text{for } x \in \mathcal{P}, \\ -\infty, & \text{for } x \notin \mathcal{P}. \end{cases}$$

Under suitable conditions on C, this is an information function.

Lemma 2.1. If C is a closed convex subset of $\mathcal{P}$, which does not contain 0 and which recedes in all directions of $\mathcal{P}$, then $\eta(\cdot \mid C)$ is a closed information function on $\mathcal{P}$.

Proof. We write

$$\eta(x) = \eta(x \mid C).$$

For $x \in \mathcal{P}$, then

$$\eta(x) \geqq 0,$$

by definition, and

$$\eta(x) < \infty.$$

For, since 0 is not contained in C, there exists some $\varepsilon > 0$ such that εB lies outside C, where B denotes the closed Euclidean unit ball. If

$$\mu \geqq \|x\| / \varepsilon,$$

then x/μ lies in εB and not in C; hence,

$$\eta(x) \leqq \|x\| / \varepsilon.$$

In order to show that

$$\eta(x) > 0, \qquad \text{for } x \in \text{ri } \mathcal{P},$$

there is no loss of generality to assume that $\mathcal{P}$ actually has nonempty interior. But

$$\eta(x) = 0 \quad \text{and} \quad x \in \text{int } \mathcal{P}$$

lead to a contradiction. From

$$\eta(x)=0,$$

it follows that the half-ray

$$R=\{\alpha x \mid \alpha>0\}$$

does not meet C; hence, there exists a hyperplane H which properly separates C and R. Let $H+$ be the open half-space determined by H which does not contain C. If, if addition,

$$x \in \operatorname{int} \mathscr{P},$$

then H meets int $\mathscr{P}$, and $\mathscr{P} \cap H+$ cannot be empty. But, for every $y \in \mathscr{P} \cap H+$, one has $x+y \in H+$, by construction, and $x+y \in C$, by the assumption that y is a direction of recession of C; this contradicts the choice of $H+$.

Homogeneity of η is immediate. Superadditivity is obvious if

$$x \notin \mathscr{P} \quad \text{or} \quad y \notin \mathscr{P},$$

or if

$$\eta(x)=\eta(y)=0.$$

If

$$\eta(x)>0 \quad \text{and} \quad \eta(y)=0,$$

then

$$\alpha>0 \quad \text{and} \quad x \in \alpha C$$

imply

$$x+y=\alpha[x/\alpha+y/\alpha] \in \alpha C,$$

since y/α is a direction of recession of C. If

$$\eta(x)>0 \quad \text{and} \quad \eta(y)>0,$$

then

$$\alpha, \beta>0 \quad \text{and} \quad x \in \alpha C, \qquad y \in \beta C$$

imply

$$x+y=(\alpha+\beta)\left[\frac{\alpha}{\alpha+\beta} \frac{x}{\alpha}+\frac{\beta}{\alpha+\beta} \frac{x}{\beta}\right] \in(\alpha+\beta) C.$$

In any case,

$$\eta(x+y) \geqq \eta(x)+\eta(y).$$

Finally, closedness of η is equivalent to closedness of the sets $\{\eta \geqq \alpha\}$, for all $\alpha \in \mathbb{R}$. Certainly,

$$\{\eta \geqq \alpha\}=\mathscr{P}, \qquad \text{if } \alpha \leqq 0.$$

If $\alpha > 0$, then

$$\{\eta \geqq \alpha\} = \alpha C;$$

for, if

$$\mu \geqq \alpha > 0 \quad \text{and} \quad x \in \mu C,$$

then

$$x = \alpha[x/\mu + (1/\alpha - 1/\mu)x] \in \alpha C, \qquad \mu C \subset \alpha C,$$

and therefore

$$\{\eta \geqq \alpha\} \subset \alpha C,$$

the converse inclusion following from the definitions. □

The condition that C recedes in all directions of $\mathscr{P}$ may also be expressed solely in terms of C. Indeed, the proof of Lemma 2.1 shows that

$$\text{ri}\,\mathscr{P} \subset \{\eta > 0\} \subset \bigcup_{\alpha>0} \alpha C \subset \mathscr{P};$$

and, taking closures, this yields

$$\mathscr{P} = \text{cl} \bigcup_{\alpha>0} \alpha C.$$

Hence, for closed convex sets C which do not contain 0 and whose recession cone 0^+C coincides with $\text{cl} \bigcup_{\alpha>0} \alpha C$, Lemma 2.1 applies with $\mathscr{P} = 0^+C$. Then, notice that $\mathscr{P}$ is the Kuratowski limit of αC as $\alpha \downarrow 0$; see Salinetti and Wets (Ref. 9, p. 19). Summarizing, we have the following theorem.

Theorem 2.1. The relations

$$j(x) = \eta(x \mid C), \qquad C = \{x \mid j(x) \geqq 1\},$$

define a one-to-one correspondence between the closed information functions j on $\mathscr{P}$ and the closed convex subsets C of $\mathscr{P}$, which do not contain 0 and which recede in all directions of $\mathscr{P}$.

Proof. Given C, Lemma 2.1 showed that

$$C = \{\eta \geqq 1\}.$$

Given j,

$$j(x) \leqq \eta(x)$$

follows from

$$x \in j(x)\{j \geqq 1\},$$

while

$$j(x) \geqq \eta(x)$$

is implied by

$$j(x) \geqq \mu,$$

whenever $x \in \mu C$. □

3. Polar Information Functions

The dual cone of $\mathscr{P}$ is

$$\mathscr{P}^d = \{u \in \mathbb{R}^n \mid \langle x, u\rangle \geqq 0, \forall x \in \mathscr{P}\},$$

where $\langle x, u\rangle$ denotes the standard inner product of $\mathbb{R}^n$. In fact, the mapping

$$j_u(x) = \langle x, u\rangle$$

itself is an information function on $\mathscr{P}$, provided

$$u \in \mathscr{P}^d \quad \text{and} \quad u \notin \mathscr{P}^\perp.$$

Moreover, if $u \in \operatorname{ri} \mathscr{P}^d$ and $\mathscr{P}$ contains no lines, then j_u is strictly isotone (but fails to be strictly concave).

The *polar set* C^∞ of a convex set C not containing 0 and the *polar function* j^∞ of an information function j on $\mathscr{P}$ are defined by

$$C^\infty = \{u \in \mathbb{R}^n \mid \langle x, u\rangle \geqq 1, \forall x \in C\},$$
$$j^\infty(u) = \inf\{\langle x, u\rangle \mid x \in \{j \geqq 1\}\}.$$

These definitions and the equality

$$\operatorname{cl}\{j \geqq 1\} = \{\operatorname{cl} j \geqq 1\}$$

entail

$$C^\infty = (\operatorname{cl} C)^\infty = (\operatorname{ri} C)^\infty,$$

$$j^\infty = (\operatorname{cl} j)^\infty = \begin{cases} \inf\{\langle x/j(x), \cdot\rangle \mid x \in \operatorname{ri} \mathscr{P}\}, & \text{on } \mathscr{P}^d, \\ -\infty, & \text{otherwise.} \end{cases}$$

The last alternative follows from

$$(\operatorname{cl} j)^\infty(u) = \inf\{\langle x, u\rangle \mid x \in \{\operatorname{cl} j = 1\}\} = \inf\{\langle y/(\operatorname{cl} j)(y), u\rangle \mid y \in \{\operatorname{cl} j > 0\}\},$$
$$\text{for } u \in \mathscr{P}^d,$$

upon observing that

$$u \in \mathcal{P}^d \quad \text{and} \quad y \in \{\mathrm{cl}\, j > 0\}$$

imply

$$\langle y/(\mathrm{cl}\, j)(y), u\rangle \in \mathrm{cl}\{\langle x/j(x), u\rangle \mid x \in \mathrm{ri}\, \mathcal{P}\}.$$

The relations between polars of sets and functions are canonical.

Lemma 3.1. If C is a convex subset of $\mathcal{P}$, which does not contain 0 and which recedes in all directions of $\mathcal{P}$, then C^∞ is a closed convex subset of $\mathcal{P}^d$, which does not contain 0 and which recedes in all directions of $\mathcal{P}^d$; and if j is an information function on $\mathcal{P}$, then j^∞ is a closed information function on $\mathcal{P}^d$. Moreover,

$$(C^\infty)^\infty = \mathrm{cl}\, C \quad \text{and} \quad (j^\infty)^\infty = \mathrm{cl}\, j.$$

Proof. The set C^∞ is closed and convex. If $u \in C^\infty$, then

$$\langle x + \lambda y, u\rangle \geqq 1, \qquad \text{for all } x \in C,\, y \in \mathcal{P},\, \lambda > 0;$$

hence, $u \in \mathcal{P}^d$. Certainly, $0 \notin C$. If $v \in 0^+(C^\infty)$, then

$$\langle x, u + \lambda v\rangle \geqq 1, \qquad \text{for all } x \in C,\, u \in C^\infty,\, \lambda > 0,$$

and

$$\langle x, v\rangle \geqq 0, \qquad \text{for all } x \in C;$$

this again implies $v \in \mathcal{P}^d$. The converse inclusion $\mathcal{P}^d \subset 0^+(C^\infty)$ is immediate. Now, if j is an information function and

$$C = \{j \geqq 1\},$$

the assertion then follows from Lemma 2.1 by showing that

$$j^\infty = \eta(\cdot \mid C^\infty).$$

But, for $\mu > 0$, one has

$$u \in \mu C^\infty \Leftrightarrow \mu \leqq j^\infty(u),$$

and this implies

$$\eta(u \mid C^\infty) = j^\infty(u), \qquad \text{if } \eta(u \mid C^\infty) > 0,$$
$$j^\infty(u) = 0, \qquad \text{if } \eta(u \mid C^\infty) = 0.$$

For studying double polars, assume C to be closed. Since $C \subset C^{\infty\infty}$ is immediate, fix $x_0 \notin C$. If $x_0 \notin \mathcal{P}$, then $x_0 \notin C^{\infty\infty}$, since

$$C^{\infty\infty} \subset \mathcal{P}^{dd} = \mathcal{P}.$$

If $x_0 \in \mathscr{P}$, choose some vector $u \in \mathbb{R}^n$ such that

$$\langle x_0, u\rangle < \inf\{\langle x, u\rangle \mid x \in C\} = a,$$

say. Then,

$$\langle x + \lambda y, u\rangle > \langle x_0, u\rangle, \qquad \text{for all } x \in C,\ y \in \mathscr{P},\ \lambda > 0,$$

and therefore,

$$u \in \mathscr{P}^d \quad \text{and} \quad \langle x_0, u\rangle \geqq 0.$$

Now,

$$\langle x, u/a\rangle \geqq 1 > \langle x_0, u/a\rangle, \qquad \text{for all } x \in C,$$

and this establishes $u/a \in C^\infty$ as well as $x_0 \notin C^{\infty\infty}$. This also yields

$$\{j^{\infty\infty} \geqq 1\} = \{\mathrm{cl}\, j \geqq 1\},$$

and Theorem 2.1 gives

$$j^{\infty\infty} = \mathrm{cl}\, j.$$

Thus, the proof is complete. □

In summary, we have the following analogy with the polarity correspondence of gauge functions. However, notice that, as in Theorem 2.1, the cones $\mathscr{P}$ and $\mathscr{P}^d$ are additional ingredients which do not have a counterpart in the theory of gauge functions.

Theorem 3.1. The polarity operation $j \to j^\infty$ induces a one-to-one symmetric correspondence in the classes of all closed information functions on $\mathscr{P}$ and $\mathscr{P}^d$. Two closed convex subsets of $\mathscr{P}$ and $\mathscr{P}^d$, which do not contain 0 and which recede in all directions of $\mathscr{P}$ and $\mathscr{P}^d$, respectively, are polar to each other if and only if their information functions are polar to each other.

The relation with the support function $\delta^*(\cdot \mid C)$ of C is also straightforward. If C is a closed convex subset of $\mathscr{P}$ which does not contain 0 and which recedes in all directions of $\mathscr{P}$, the function $\eta(\cdot \mid C)$ and the function h given by

$$h(u) = -\delta^*(-u \mid C)$$

are information functions polar to each other.

Examples of information functions on the nonnegative quadrant $\mathbb{R}^n_+$ are the generalized means of order $p \in [-\infty, +1]$, defined for $x \in \mathrm{ri}\, \mathbb{R}^n_+$ by

$$j_p(x) = (n^{-1}\Sigma x_i^p)^{1/p}, \qquad p \neq -\infty, p \neq 0,$$

$$j_0(0) = (x_1, \ldots, x_n)^{1/n},$$

$$j_{-\infty}(x) = \min\{x_1, \ldots, x_n\},$$

and extended to all of $\mathbb{R}^n_+$ by semicontinuity. The polar of j_p is nj_q, by Hölder's inequality, where $p, q \in [-\infty, +1]$ must satisfy

$$p+q=pq.$$

Notice that, in the j_p-family, the geometric mean j_0 is the only member which, up to positive proportionality, is self-polar.

The j_p-means serve as prime examples in experimental design theory, where they are interpreted as information functions (cf. Section 2 in Ref. 1), and in econometric production theory, where they serve as production functions (cf. Section 1.9 in Ref. 8).

4. Information-Like Functions

It is of natural interest to determine those functions which, except for a monotone transformation, coincide with some information function. A concave function g on $\mathbb{R}^n$ will be called *information-like on* $\mathcal{P}$ if g is proper, dom $g \subset \mathcal{P}$,

$$g(0)=\inf\{g(x) \mid x \in \mathcal{P}\}<g(x), \qquad \text{for some } x \in \mathcal{P},$$

and, for $\alpha \in (g(0), \sup g)$, the sets $\{g \geqq \alpha\}$ are all positively proportional and have recession cone $\mathcal{P}$. Since any such function g is nonconstant, it is necessarily proper, provided it is closed.

Theorem 4.1. A closed concave function g is information-like on $\mathcal{P}$ if and only if it can be expressed in the form

$$g=h \circ j,$$

where j is a closed information function on $\mathcal{P}$ and h is a nondecreasing closed concave function on $[0, +\infty)$ such that

$$h(0)<h(\zeta), \qquad \text{for some } \zeta>0,$$

while

$$h(-\infty)=-\infty.$$

Proof. For the direct part, assume g to be given. Let I be the open interval $(g(0), \sup g)$, and fix

$$C=\{g \geqq \alpha_1\}, \qquad \text{for some } \alpha_1 \in I.$$

Then, C is closed and convex, it is contained in $\mathcal{P}$, but does not contain 0, it recedes in all directions of $\mathcal{P}$; and, for $\alpha \in I$, there exists some $\zeta>0$ such that

$$\{g \geqq \alpha\} = \zeta C.$$

Hence, choose

$$j = \eta(\cdot \mid C);$$

and, for $\zeta \geqq 0$, define h by

$$h(\zeta) = g(0),$$

if ζC is not contained in any of the sets $\{g \geqq \alpha\}$ for $\alpha \in I$, and define h by

$$h(\zeta) = \sup\{\alpha \in I \mid \zeta C \subset \{g \geqq \alpha\}\},$$

otherwise. Now, h is nondecreasing; for, if $0 < \zeta < \xi$, then

$$\xi C = \{\zeta x + \zeta(\xi/\zeta - 1)x \mid x \in C\} \subset \zeta C,$$

and hence,

$$h(0) = g(0) \leqq h(\zeta) \leqq h(\xi).$$

Furthermore, for $x \in \mathscr{P}$ with $g(x) > g(0)$, one has

$$\begin{aligned} g(x) &= \sup\{\alpha \in \mathbb{R} \mid x \in \{g \geqq \alpha\}\} \\ &= \sup\{\alpha \in I \mid \exists \zeta > 0 \colon x \in \zeta C \text{ and } \zeta C \subset \{g \geqq \alpha\}\} \\ &= \sup\{h(\zeta) \mid \zeta > 0 \text{ and } x \in \zeta C\} \\ &= h \circ j(x). \end{aligned}$$

Otherwise,

$$x \in \mathscr{P} \quad \text{and} \quad g(x) = g(0),$$

or

$$x \notin \mathscr{P}.$$

In both cases, the equality

$$g(x) = h \circ j(x)$$

follows from the definitions. Any $x_1 \in C$ with $j(x_1) = 1$ permits the alternative representation of h as

$$h(\zeta) = g(\zeta x_1),$$

showing that h has all the required properties.

For the converse part, assume h and j to be given. Then,

$$g = h \circ j$$

is a closed proper concave function with

$$\operatorname{dom} g \subset \mathcal{P}$$

and

$$g(0) = h(0) = \inf\{g(x) \mid x \in \mathcal{P}\} < \sup g = \sup h.$$

Fix

$$\alpha < \sup g,$$

and introduce the pseudo-inverse

$$h^{-1}(\alpha) = \inf\{\zeta \geqq 0 \mid h(\zeta) \geqq \alpha\}.$$

Then,

$$h^{-1}(\alpha) \geqq 0,$$

with equality if and only if

$$\alpha \leqq h(0) = g(0).$$

With

$$C = \{j \geqq 1\},$$

this yields

$$\{g \geqq \alpha\} = \{j \geqq h^{-1}(\alpha)\} = h^{-1}(\alpha)C, \qquad \text{if } \alpha > g(0). \qquad \square$$

The concave conjugate of information-like functions again is information-like. Indeed, the conjugate of a function

$$g = h \circ j$$

as in Theorem 5.1 turns out to be

$$g^* = h^- \circ j^\infty,$$

where the monotone conjugate h^- of h is given by

$$h^-(\xi) = \inf\{\zeta\xi - h(\zeta) \mid \zeta \geqq 0\};$$

see page 111 in Rockafellar (Ref. 6). Moreover, applying Corollary 12.2.2 of Rockafellar (Ref. 6) and using

$$\operatorname{ri} \mathcal{P} \subset \bigcup_{\zeta > 0} \zeta D,$$

with

$$D = \{\operatorname{cl} j = 1\},$$

the following argument, for $u \in \mathscr{P}^d$, does not require j to be closed:

$$\begin{aligned} g^*(u) &= \inf_{x \in \mathscr{P}} \langle x, u\rangle - h \circ j(x) \\ &= \inf_{x \in \mathrm{ri}\, \mathscr{P}} \langle x, u\rangle - h \circ (\mathrm{cl}\, j)(x) \\ &= \inf_{\zeta > 0} \zeta \inf_{x \in D} \langle x, u\rangle - h(\zeta) \\ &= h^- \circ j^\infty(u). \end{aligned}$$

An instructive example is given by the transformation

$$h(\zeta) = \zeta,$$

with monotone conjugate $h^-(\xi)$ equal to $-\infty$ or 0 according as $\xi < 1$ or $\xi \geqq 1$. The transformation

$$h(\zeta) = \log \zeta$$

is employed in the proof of Theorem 4 in Pukelsheim (Ref. 1); its monotone conjugate is

$$\log^- \xi = 1 + \log \xi.$$

Further examples are listed in Section 7 of Bellman and Karush (Ref. 10).

References

1. PUKELSHEIM, F., *On Linear Regression Designs Which Maximize Information*, Journal of Statistical Planning and Inference, Vol. 4, pp. 339–364, 1980.
2. CORLEY, H. W., *An Existence Result for Maximizations with Respect to Cones*, Journal of Optimization Theory and Applications, Vol. 31, pp. 277–281, 1980.
3. TANINO, T., and SAWARAGI, Y., *Conjugate Maps and Duality in Multiobjective Optimization*, Journal of Optimization Theory and Applications, Vol. 31, pp. 473–499, 1980.
4. BITRAN, G. R., and MAGNANTI, T. L., *The Structure of Admissible Points with Respect to Cone Dominance*, Journal of Optimization Theory and Applications, Vol. 29, pp. 573–614, 1979.
5. GEARHART, W. B., *Compromise Solutions and Estimation of the Noninferior Set*, Journal of Optimization Theory and Applications, Vol. 28, pp. 29–47, 1979.
6. ROCKAFELLAR, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
7. ROCKAFELLAR, R. T., *Monotone Processes of Convex and Concave Type*, Memoirs of the American Mathematical Society, Vol. 77, p. 74, 1967.
8. MCFADDEN, D., *Cost, Revenue, and Profit Functions. Production Economics: A Dual Approach to Theory and Applications, Vol. 1, The Theory of Production*, Edited by M. Fuss and D. McFadden, North-Holland, Amsterdam, Holland, pp. 1–109, 1978.

9. SALINETTI, G., and WETS, J.-B., *On the Convergence of Sequences of Convex Sets in Finite Dimensions,* SIAM Review, Vol. 21, pp. 18–33, 1979.
10. BELLMAN, R., and KARUSH, W., *Mathematical Programming and the Maximum Transform,* SIAM Journal on Applied Mathematics, Vol. 10, pp. 550–567, 1962.