Mathematical Programming 18 (1980) 215–227.
North-Holland Publishing Company

A GENERALIZATION OF THE FRANK–WOLFE THEOREM*

André F. PEROLD**

Department of Operations Research, Stanford University, Stanford, CA, U.S.A.

Received 21 November 1977
Revised manuscript received 15 August 1979

The Frank–Wolfe theorem states that a quadratic function, bounded below on a nonempty polyhedral convex set, attains its infimum there. This paper gives sufficient conditions under which a function either attains its infimum on a nonempty polyhedral convex set or is unbounded below on some halfline of that set. Quadratic functions are shown to satisfy these sufficient conditions.

Key words: Frank–Wolfe Theorem, Quadratic Programming, Norm Coercive Functions, Polyhedral Sets.

1. Introduction

The existence theorem for quadratic programming states that a quadratic function Q bounded below on a nonempty polyhedral convex set C attains its infimum there. This result was first proved in 1956 by Frank and Wolfe [5]. Alternative proofs have since been given by Collatz and Wetterling [2] (for the case when Q is convex), Eaves [4] and Blum and Oettli [1].

Eaves [4] also improved on this result by showing that if Q does not attain its infimum on C, then Q must be unbounded below on some halfline contained in C. This was first claimed, but not proved, by Dennis in 1959 [3].

In [5], the example $Q(x_1,x_2) = x_1^2 + (1 - x_1x_2)^2$ is given to show that these results do not hold in general for higher order polynomials. (Q here does not attain its infimum, zero, in the plane.) This leads one to ask: what is so special about a quadratic; and also perhaps, what is so special about polyhedral convex sets. Unfortunately, the proofs in [1], [2], [4] and [5] are specifically tailored to the quadratic and polyhedral case, and shed little light on the answers to these questions.

In this paper we shall deal with the first question. Let $\mathcal{F}$ be the class of all continuous functions $f:\mathbf{R}^n \to \mathbf{R}^1$ (some n) such that for any polyhedral convex set $C \subseteq \mathbf{R}^n$, f either attains its infimum or C or is unbounded below on some

* Research and reproduction of this report were partially supported by the National Science Foundation Grant MCS76–81259; and the Office of Naval Research Contract N00014–75–C–0267.

** Current address: Graduate School of Business Administration, Harvard University, Boston, MA. 02163, U.S.A.

halfline contained in C. We shall define a class of functions $\mathscr{G}$ which, in particular, contains all quadratic functions, and shall show that $\mathscr{G}$ is contained (strictly) in $\mathscr{F}$.

2. Notation and definitions

Let $\boldsymbol{n}$ denote $\{1,2,\dots,n\}$. For $x\in\mathbf{R}^n$, $\alpha\subseteq\boldsymbol{n}$, let $x_\alpha\in\mathbf{R}^k$ denote

$$(x_{\alpha_1},\dots,x_{\alpha_k})^{\mathrm{T}}\quad\text{where } \alpha=\{\alpha_1,\dots,\alpha_k\},\quad \alpha_1<\cdots<\alpha_K.$$

For $A\in\mathbf{R}^{m\times n}$, $\alpha\subseteq\boldsymbol{m}$, $\beta\subseteq\boldsymbol{n}$, let $A_{\alpha\cdot}$ denote the submatrix of rows of A indexed by α; let $A_{\cdot\beta}$ denote the submatrix of columns of A indexed by β; let $A_{\alpha\beta}$ denote $(A_{\alpha\cdot})_{\cdot\beta}$.

A function $f:\mathbf{R}^n\to\mathbf{R}^1$ is said to be *norm-coercive* if $\lim_{k\to\infty}|f(x_k)|=\infty$ for every sequence $\{x_k\}\subseteq\mathbf{R}^n$ such that $\lim_{k\to\infty}\|x_k\|=\infty$.

3. The class $\mathscr{G}$ and some examples

We define $\mathscr{G}$ by induction on the number of variables.

Definition 3.1. Let $f:\mathbf{R}^n\to\mathbf{R}^1$ be continuous. Then $f\in\mathscr{G}$ if there exists a nonsingular matrix Q, and a partition of $\boldsymbol{n}$, $\boldsymbol{n}=\alpha\cup\beta$, such that $g:\mathbf{R}^n\to\mathbf{R}$ defined by $g(x)=f(Qx)$ satisfies

(i) if $\beta\neq\emptyset$, then $g(x_\alpha,\cdot)$ is a concave function;

(ii) if $\beta=\emptyset$, then g is norm-coercive;

(iii) if $\alpha\neq\emptyset$ and $\beta\neq\emptyset$, then for all A and b of appropriate dimensions, the function $h(\cdot)$, defined by $h(x_\alpha)=g(x_\alpha,Ax_\alpha+b)$, is a member of $\mathscr{G}$.

Note that for $n=1$, condition (iii) is trivially satisfied since either α or β must be empty. Hence a function of one variable, f, is in $\mathscr{G}$ if f is concave or if f is norm-coercive and continuous.

Note also that the induction step in the definition of $\mathscr{G}$ is in condition (iii).

Examples 3.2. (i) All continuous norm-coercive functions are in $\mathscr{G}$. Set $Q=1$, $\alpha=\boldsymbol{n}$, $\beta=\emptyset$ in Definition 3.1.

(ii) All concave functions are in $\mathscr{G}$. Set $Q=I$, $\alpha=\emptyset$, $\beta=\boldsymbol{n}$ in Definition 3.1.

(iii) All quadratic functions are in $\mathscr{G}$.

We shall prove this by induction on n, the number of variables. Let

$$f(x)=c^{\mathrm{T}}x+\tfrac{1}{2}x^{\mathrm{T}}Dx,\quad x\in\mathbf{R}^n.$$

If $n=1$, then f is either concave or $f(x)\to\infty$ as $|x|\to\infty$, and we are done.

Suppose the result is true for quadratic functions of $n-1$ variables or fewer. Without loss of generality we may assume that D is symmetric. Hence there is an orthogonal matrix Q such that

$$Q^{\mathrm{T}}DQ = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$$

where the λ_i are the real eigenvalues of D.

Set $d = Q^{\mathrm{T}}c$, and $g(x) = f(Qx)$. Then

$$g(x) = \sum_{\lambda_i>0} (\lambda_i x_i^2 + d_i x_i) + \sum_{\lambda_i \le 0} (\lambda_i x_i^2 + d_i x_i).$$

Let $\alpha = \{i : \lambda_i > 0\}$, $\beta = \{i : \lambda_i \le 0\}$. Then α and β partition $\boldsymbol{n}$, and $g(x)$ clearly satisfies conditions (i) and (ii) of Definition 3.1.

Also, since for any A and $b, g(x_\alpha, Ax_\alpha + b)$ is a quadratic function in fewer variables, the induction hypothesis applies and condition (iii) of Definition 3.1 is established.

(iv) Any function of the form

$$f(u,v) = (u^{\mathrm{T}}Bu)^k - (v^{\mathrm{T}}Dv)^m$$

where B is positive definite, D is positive semi-definite and $k > m \ge 0$, is in $\mathscr{G}$.

Set $Q = I$, and identify x_α with u and x_β with v. Since B is positive definite

$$x_\alpha{}^{\mathrm{T}}Bx_\alpha \ge \xi \|x_\alpha\|^2 \quad \forall x_\alpha \tag{1}$$

where $\xi > 0$ is the smallest eigenvalue of B.

Since D is positive semi-definite, and the function t^m is nondecreasing in $t \ge 0, (x_\beta^{\mathrm{T}}Dx_\beta)^m$ is convex in x_β.

Hence conditions (i) and (ii) of Definition 3.1 are satisfied.

Further, since D is positive semi-definite,

$$0 \le x_\beta^{\mathrm{T}}Dx_\beta \le \eta \|x_\beta\|^2 \quad \forall x_\beta$$

where $\eta \ge 0$ is the largest eigenvalue of D. Thus for any A and b

$$\begin{aligned} 0 &\le (Ax_\alpha + b)^{\mathrm{T}}D(Ax_\alpha + b) \\ &\le \eta \|Ax_\alpha + b\|^2 \\ &\le \eta(\mu\|x_\alpha\| + \|b\|)^2 \end{aligned} \tag{2}$$

where μ is the largest row norm of A.

Combining (1) and (2) we obtain

$$f(x_\alpha, Ax_\alpha + b) \ge \xi^k \|x_\alpha\|^{2k} - \eta^m(\mu\|x_\alpha\| + \|b\|)^{2m}.$$

The right-hand side of this inequality is a polynomial in $\|x_\alpha\|$ with leading coefficient $\xi^k > 0$. Hence $f(x_\alpha, Ax_\alpha + b) \to \infty$ as $\|x_\alpha\| \to \infty$. This establishes condition (iii) of Definition 3.1 by an application of example (i).

Remark. Example (iv) shows that the class $\mathscr{G}$ is indeed interesting, that is, $\mathscr{G}$ consists of functions to which the Frank–Wolfe theorem as it stands is not applicable.

4. Preliminary results

Let C be a nonempty polyhedral convex set of the form

$$C = \{x \in \mathbf{R}^n : Ax \leq b\}$$

where $A \in \mathbf{R}^{mxn}$.

Definition 4.1. Let $x \in C$, and let $\gamma = \{i : A_{i\cdot}x = b_i\}$. Then x is called a *pseudo-extreme point* of C if x uses linearly independent columns of $A_{\gamma\cdot}$, i.e. if $\delta = \{j : x_j \neq 0\}$, then $A_{\gamma\delta}$ has full column rank.

For convenience, when A is the zero matrix, define the origin to be the pseudo-extreme point of C.

One can easily show that C always has pseudo-extreme points. Furthermore, it follows by Lemma 4.3 (stated below) that the extreme points of C, when they exist, are pseudo-extreme points of C.

Geometrically, the pseudo-extreme points of C are the extreme points of all the sections of C at $x_\alpha = 0$ where α ranges over all subsets (including the empty set) of $\boldsymbol{n}$.

In the following examples, the pseudo-extreme points are marked with asteriscs. The set C is the shaded area.

(i) $$C = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\};$$

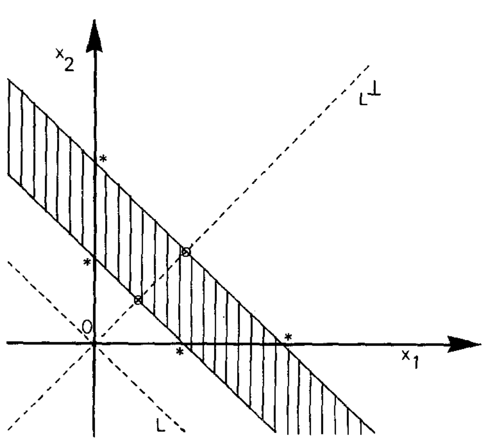

Fig. 1.

(ii) $$C=\left\{\begin{pmatrix}x_1\\x_2\end{pmatrix}:\begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}\leq\begin{pmatrix}1\\1\end{pmatrix}\right\}.$$

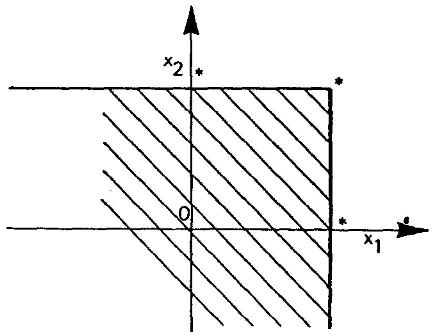

Fig. 2.

The following theorem shows that C has a representation in terms of its pseudo-extreme points.

Theorem 4.2. *C has a representation*

$$C=\{s+\mu t: s\in S, t\in T, \mu\geq 0\}$$

where S is the convex hull of the pseudo-extreme points of C, and T is the intersection of a polyhedral convex cone with the unit sphere.

To prove this result, we require 3 lemmas.

Lemma 4.3. *Let $x\in C$ and let $\gamma=\{i: A_{i\cdot}x=b_i\}$. Then x is an extreme point of C iff the rank of $A_{\gamma\cdot}$ is n.*

Lemma 4.4. *If C has extreme points, then C has the representation*

$$C=\{s+\mu t: s\in S, t\in T, \mu\geq 0\}$$

where S is the convex hull of the extreme points of C and T is the intersection of a polyhedral convex cone with the unit sphere.

The proofs of these two lemmas can be found in Goldman [6].

In the following, let x^+ and x^- denote the positive and negative parts of x, respectively.

Lemma 4.5. *Let*

$$\bar{C}=\left\{\begin{pmatrix}u\\v\end{pmatrix}\geq 0: Au-Av\leq b\right\}.$$

Then x is a pseudo-extreme point of C if and only if $\binom{x^+}{x^-}$ *is an extreme point of* $\bar{C}$.

Proof. Let x be a pseudo-extreme point of C. Let $\gamma = \{i : A_{i\cdot}x = b_i\}$. Then since x uses linearly independent columns of $A_{\gamma\cdot}$, $\binom{x^+}{x^-}$ uses linearly independent columns, say $\delta \subseteq \mathbf{2n}$, of the partitioned matrix $(A_{\gamma\cdot}, -A_{\gamma\cdot})$. Let $B \doteq (A, -A)$ and let I be the identity matrix of dimension $2n$. Then $\bar{C}$ can be written as

$$\bar{C} = \left\{\binom{u}{v} : \bar{A}\binom{u}{v} \le \binom{b}{0}\right\}$$

where

$$\bar{A} = \binom{B}{I}.$$

Thus the row submatrix of $\bar{A}$ corresponding to the constraints which $\binom{x^+}{x^-}$ satisfies with equality is

$$\binom{B_{\gamma\cdot}}{I_{\bar{\delta}\cdot}}$$

where $\bar{\delta}$ denotes $\mathbf{2n} \sim \delta$.

Rewrite this matrix as

$$\begin{pmatrix} B_{\gamma\delta} & B_{\gamma\bar{\delta}} \\ 0 & I_{\bar{\delta}\bar{\delta}} \end{pmatrix}.$$

Since $B_{\gamma\delta}$ and $I_{\bar{\delta}\bar{\delta}}$ both have full column rank it follows that

$$\operatorname{rank}\binom{B_{\gamma\cdot}}{I_{\bar{\delta}\cdot}} = 2n.$$

By Lemma 4.3 $\binom{x^+}{x^-}$ is an extreme point of $\bar{C}$, and the first implication is established.

The converse follows easily along similar lines.

Proof of Theorem 4.2. Let $\bar{C}$ be as in the statement of Lemma 4.5. Since C has pseudo-extreme points, it follows by Lemma 4.5 that $\bar{C}$ has extreme points.

Thus by Lemma 4.4 $\bar{C}$ has a representation

$$\bar{C} = \left\{\binom{s'}{s''} + \mu\binom{t'}{t''} : \binom{s'}{s''} \in \bar{S}, \binom{t'}{t''} \in \bar{T}, \mu \ge 0\right\}$$

where $\bar{S}$ is the convex hull of the extreme points of $\bar{C}$, and $\bar{T}$ is the intersection of a polyhedral convex cone with the unit sphere.

Now C can be written as

$$C = \left\{u - v : \begin{pmatrix} u \\ v \end{pmatrix} \in \bar{C}\right\}.$$

Hence

$$C = \left\{(s' - s'') + \mu(t' - t'') : \begin{pmatrix} s' \\ s'' \end{pmatrix} \in \bar{S}, \begin{pmatrix} t' \\ t'' \end{pmatrix} \in \bar{T}, \mu \geq 0\right\}.$$

Set

$$S = \left\{s' - s'' : \begin{pmatrix} s' \\ s'' \end{pmatrix} \in \bar{S}\right\}.$$

Then by Lemma 4.5 S is the convex hull of the pseudo-extreme points of C. Set

$$T = \left\{\frac{t' - t''}{\|t' - t''\|} : \begin{pmatrix} t' \\ t'' \end{pmatrix} \in \bar{T}, t' - t'' \neq 0\right\}.$$

Then

$$C = \{s + \mu t : s \in S, t \in T, \mu \geq 0\}$$

as required.

Theorem 4.6. *Let $g : C \to \mathbf{R}$ be concave. Then either g attains its infimum at a pseudo-extreme point of C or g is unbounded below on some halfline of C.*

In 1961 Hirsch and Hoffman [7] proved a similar theorem using a different representation of C. They decomposed C as

$$C = M \oplus L$$

where L is a linear subspace and M is the $L^{\perp}$-section of C, and showed that a concave function bounded below on C attains its infimum at an extreme point of M. In Fig. 1 the extreme points of M are circled. Note that they are not the same as the pseudo-extreme points of C.

Proof of Theorem 4.6. Let $\{p_1, \dots, p_k\}$ be the set of pseudo-extreme points of C. Let

$$g(p_m) = \min_{1 \leq i \leq k} g(p_i).$$

Let S and T be as in Theorem 4.2. For any $s \in S$ $\exists \lambda_i \geq 0, i = 1, \dots, k, \sum_{i=1}^{k} \lambda_i = 1$ such that

$$s = \sum_{i=1}^{k} \lambda_i p_i.$$

By the concavity of g,

$$g(s) = g\left(\sum_{i=1}^{k} \lambda_i p_i\right) \geq \sum_{i=1}^{k} \lambda_i g(p_i).$$

Hence

$$g(s) \geq g(p_m) \ \forall s \in S. \tag{1}$$

Now suppose that g does not attain its infimum at a pseudo-extreme point of C. Then by (1) $\exists x \in C \setminus S$ with

$$g(x) < g(p_m).$$

By Theorem 4.2 x can be written as

$$x = s + \mu t$$

for some $s \in S, t \in T, \mu \geq 0$.

Since $x \notin S$, $\mu > 0$.

Further, the half-line

$$H = \{s + \xi t : \xi \geq \mu\}$$

is contained in C.

Now $\forall \xi \geq \mu$

$$s + \mu t = \left(1 - \frac{\mu}{\xi}\right)s + \left(\frac{\mu}{\xi}\right)(s + \xi t).$$

Since g is concave

$$g(s + \mu t) \geq \left(1 - \frac{\mu}{\xi}\right)g(s) + \left(\frac{\mu}{\xi}\right)g(s + \xi t).$$

Upon rearrangement we get

$$g(s + \xi t) \leq g(s) - \left(\frac{g(s) - g(s + \mu t)}{\mu}\right)\xi.$$

Since

$$g(s) - g(s + \mu t) \geq g(pm) - g(x) > 0$$

it follows that

$$g(s + \xi t) \to -\infty \quad \text{as } \xi \to \infty$$

i.e. g is unbounded below on H.

The following result is of interest in its own right.

Theorem 4.7. *Let $g : \mathbf{R}^n \to \mathbf{R}^1$, $n \geq 2$, be continuous and norm-coercive. Then g is either bounded above or bounded below.*

The proof of this result uses the concept of path connectedness. A set $D \subset \mathbf{R}^n$ is said to be *path connected* if for any x, $y \in D$ there is a continuous mapping

$p:[0,1]\to D$ such that $p(0)=x$ and $p(1)=y$.

Let K_r denote the closed ball of radius r, and $\bar{K}_r$ its complement. For $n\geq 2$ $\bar{K}_r$ is path connected.

Proof of Theorem 4.7. By assumption on g, there exists $r>0$ such that

$$x\in\bar{K}_r \Rightarrow |g(x)|\geq 1. \tag{1}$$

We shall show that either $g(x)\geq 1 \forall x\in\bar{K}_r$ or $g(x)\leq -1 \forall x\in\bar{K}_r$. If not $\exists x,y\in\bar{K}_r \ni g(x)\geq 1$ and $g(y)\leq -1$. Since $\bar{K}_r$ is path connected there is a continuous mapping $p:[0,1]\to\bar{K}_r \ni p(0)=x$, $p(1)=y$. Since g is continuous, the mapping $h:[0,1]\to\mathbf{R}$ defined by

$$h(t)=g(p(t))$$

is continuous. Moreover $h(0)\geq 1$ and $h(1)\leq -1$. By the intermediate value theorem there exists $t_0\in[0,1]$ such that $h(t_0)=0$, i.e.

$$g(p(t_0))=0.$$

Since $p(t_0)\in\bar{K}_r$ this contradicts (1).

Hence g is either bounded above or below on $\bar{K}_r$. Since K_r is compact and g is continuous, g is bounded on K_r.

The following theorem is obtained as a consequence of the above result.

Theorem 4.8. *Let $g:\mathbf{R}^n\to\mathbf{R}$ be continuous and norm-coercive. Let P be any non-empty closed convex set. Then g either attains its infimum on P or is unbounded below on some half-line contained in P.*

Proof. If g is bounded below on P, then

$$\|x\|\to\infty,\ x\in P\Rightarrow g(x)\to\infty.$$

Thus for k sufficiently large the set

$$R=\{x:g(x)\leq k\}\cap P$$

is nonempty and compact.

Since g is continuous g attains its infimum on R. The infimum of g on R is equal to the infimum of g on P and we are done.

Suppose g is unbounded below on P. Then since g is continuous, there is a sequence $\{x_n\}\subseteq P\ni g(x_n)\to -\infty$ and $\|x_n\|\to\infty$.

If $n=1$, one of the half-lines $P\cap[-\infty,0]$, $P\cap[0,\infty)$ will do.

If $n\geq 2$, it follows by Theorem 4.7 that

$$\|x\|\to\infty,\ x\in P\Rightarrow g(x)\to -\infty.$$

Since any unbounded convex set contains half-lines, any half-line in P will do.

Lemma 4.9. *Let $P(b)=\{x: Ax\le b\}$. Then there is a finite collection $\{(\gamma_i,\delta_i): i=1,\dots,l\}$ where $\gamma_i\subseteq \boldsymbol{m}$, $\delta_i\subseteq \boldsymbol{n}$ such that $\forall b$, x is a pseudo-extreme point of $P(b)\Rightarrow$ there exists $i\in \boldsymbol{l}$ such that*

(i) $A_{\gamma_i\delta_i}x_{\delta_i}=b_{\gamma_i}$;
(ii) $A^{-1}_{\gamma_i\delta_i}$ *exists*;
(iii) $x_{\bar\delta_i}=0$ *where* $\bar\delta_i=\boldsymbol{n}\sim\delta_i$.

The proof is immediate from Definition 4.1 and will be omitted.

5. Proof of the main result: $\mathcal{G}\subset\mathcal{F}$

The proof is by induction on the number of variables. Let C be any non-empty polyhedral convex set defined by

$$C=\{x\in\mathbf{R}^n: Ax\le b\}$$

where $A\in\mathbf{R}^{m\times n}$.

For $n=1$, if $f\in\mathcal{G}$, then f is either concave or $|f(x)|\to\infty$ as $|x|\to\infty$. In the former case, Theorem 4.6 applies. In the latter case, Theorem 4.8 applies.

Suppose the result holds for all $k<n$. Let $f\in\mathcal{G}$, $f:\mathbf{R}^n\to\mathbf{R}^1$. Let Q, α and β be as in Definition 3.1. Since Q is nonsingular, it suffices to prove the result for g defined by

$$g(x)=f(Qx)$$

on the set

$$\bar C=\{x:\bar A x\le b\}$$

where $\bar A=AQ$.

If either $\alpha=\emptyset$ or $\beta=\emptyset$, then g is either concave or norm-coercive and we can use Theorem 4.6 or 4.8.

So assume $\alpha\ne\emptyset$, $\beta\ne\emptyset$. Let $\{x^k\}=\{(x^k_\alpha,x^k_\beta)\}$ be a minimizing sequence for g in the sense that $g(x^k)\searrow-\infty$ if g is unbounded below on $\bar C$, or else

$$g(x^k)\searrow\inf\{g(x): x\in\bar C\}=\rho>-\infty\quad\text{(say).}$$

Now for each k hold x^k_α fixed and solve

$$\begin{aligned}&\min && g(x^k_\alpha,u_\beta),\\ &\text{subject to} && u_\beta\in\bar C_k\overset{\Delta}{=}\{y:\bar A_{\cdot\alpha}x^k_\alpha+\bar A_{\cdot\beta}y\le b\}.\end{aligned}$$

Since $g(x^k_\alpha,\cdot)$ is concave, we can apply Theorem 4.6:

If $g(x^k_\alpha,\cdot)$ is unbounded below on $\bar C_k$, then there is a half line $\{s+\mu t:\mu\ge 0\}\subseteq\bar C_k$

such that

$$g(x_\alpha^k, s+\mu t) \to -\infty \quad \text{as } \mu \to \infty$$

i.e. g is unbounded below on the half-line

$$\left\{\begin{pmatrix} x_\alpha^k \\ s \end{pmatrix} + \mu \begin{pmatrix} 0 \\ t \end{pmatrix} : \mu \geq 0\right\} \subseteq \bar{C}$$

and the result is proved.

Otherwise, $\forall k$, $g(x_\alpha^k, \cdot)$ attains its infimum at a pseudo-extreme point $u_\beta^k \in \bar{C}_k$. By Lemma 4.9 there exists (γ_i, δ_i) such that

$$\begin{aligned} &\bar{A}_{\gamma_i\delta_i} u_{\delta_i}^k = b_{\gamma_i} - \bar{A}_{\gamma_i\alpha} x_\alpha^k; \\ &\bar{A}_{\gamma_i\delta_i}^{-1} \quad \textit{exists}; \\ &u_{\bar{\delta}_i}^k = 0 \quad \text{where } \bar{\delta}_i = \beta \sim \delta_i. \end{aligned} \tag{1}$$

Since the pair (x_α^k, u_β^k) is feasible

$$\bar{A}_{\bar{\delta}_i\beta} u_\beta^k \leq b_{\bar{\gamma}_i} - \bar{A}_{\bar{\gamma}_i\alpha} x_\alpha^k \tag{2}$$

where $\bar{\gamma}_i = m \sim \gamma_i$.

Since there are only finitely many such (γ_i, δ_i) there is an infinite subsequence K such that some $(\gamma_i, \delta_i) = (\gamma, \delta)$ (say) is repeated for all $k \in K$.

Using (1) to eliminate u_β^k from (2) we get, $\forall k \in K$

$$\begin{aligned} &u_\beta^k = d - Dx_\alpha^k, \\ &Bx_\alpha^k \leq c \end{aligned}$$

where

$$\begin{aligned} &d = \begin{pmatrix} d_\delta \\ d_{\bar{\delta}} \end{pmatrix} = \begin{pmatrix} \bar{A}_{\gamma\delta}^{-1} b_\gamma \\ 0 \end{pmatrix}, \\ &D = \begin{pmatrix} D_{\delta\cdot} \\ D_{\bar{\delta}\cdot} \end{pmatrix} = \begin{pmatrix} \bar{A}_{\gamma\delta}^{-1} \bar{A}_{\gamma\alpha} \\ 0 \end{pmatrix}, \\ &c = b_{\bar{\gamma}} - \bar{A}_{\bar{\gamma}\beta} d, \\ &B = \bar{A}_{\bar{\gamma}\alpha} - \bar{A}_{\bar{\gamma}\beta} D. \end{aligned}$$

Now since $\{(x_\alpha^k, x_\beta^k)\}$ is a minimizing sequence for $g(\cdot)$ on $\bar{C}$, it follows that $\{(x_\alpha^k, u_\beta^k)\}$ is a minimizing sequence for $g(\cdot)$ on $\bar{C}$ since by definition of u_β^k,

$$g(x_\alpha^k, u_\beta^k) \leq g(x_\alpha^k, x_\beta^k).$$

But $(x_\alpha^k, u_\beta^k) = (x_\alpha^k, d - Dx_\alpha^k)\ \forall k \in K$. Hence $\{x_\alpha^k\}_{k \in K}$ is a minimizing sequence for $h(\cdot)$ on $\bar{\bar{C}}$ where

$$h(z) \stackrel{\Delta}{=} g(z, d - Dz),$$

$$\bar{\bar{C}} = \{z : Bz \leq c\}.$$

We can now apply the induction hypothesis to $h(\cdot)$ on $\bar{\bar{C}}$ using condition (iii) of Definition 3.1 as follows:

If g is bounded below on $\bar{C}$, then

$$g(x_\alpha^k, x_\beta^k) \searrow \rho \quad \text{as } k \to \infty$$

so that

$$g(x_\alpha^k, d - Dx_\alpha^k) = h(x_\alpha^k) \searrow \rho \quad \text{as } k \to \infty, k \in K.$$

Hence there exists $\bar{x}_\alpha \in \bar{\bar{C}}$ such that $h(\bar{x}_\alpha) = \rho$ i.e. $g(\bar{x}_\alpha, d - \bar{D}x_\alpha) = \rho$, and we are done.

Otherwise $g(x_\alpha^k, x_\beta^k) \searrow -\infty$ so that $h(z_\alpha^k) \searrow -\infty, k \in K$, i.e. $h(\cdot)$ is unbounded below on $\bar{\bar{C}}$. Hence there exists a half-line $\{s + \mu t: \mu \geq 0\} \subseteq \bar{\bar{C}}$ such that

$$h(s + \mu t) \searrow -\infty \quad \text{as } \mu \to \infty$$

i.e. $g(\cdot)$ is unbounded below on the half-line

$$\left\{\begin{pmatrix} s \\ d - Ds \end{pmatrix} + \mu \begin{pmatrix} t \\ -Dt \end{pmatrix} : \mu \geq 0\right\}$$

in $\bar{C}$. This completes the proof.

6. Concluding remarks

We might ask questions about the structures of $\mathcal{G}$ and $\mathcal{F}$.

(i) Neither is closed under addition: Let $f(x) = x^2 + e^{-x}$ and $g(x) = -x^2$. Then both f and $g \in \mathcal{G}$ while $f + g \notin \mathcal{F}$.

(ii) $\mathcal{F}$ is closed under arbitrary affine transformations of the variables: Let $f \in \mathcal{F}$, $f: \mathbf{R}^n \to \mathbf{R}^1$. Let $B \in \mathbf{R}^{n \times k}$ (any k) and $d \in \mathbf{R}^n$ be arbitrary. We must show that $f(Bx + d)$ has the required properties on any $C = \{x : Ax \leq b\} \subseteq \mathbf{R}^k$.

Let $E = \{(x,y): Ax \leq b, y = Bx + d\}$. Since E is a polyhedral convex set, so is its projection onto y space, E_Y (say).

If $\inf\{f(Bx + d): x \in C\} = \rho > -\infty$, then $\inf\{f(y): y \in E_Y\} = \rho$. Since $f \in \mathcal{F}$, there exists $\bar{y} \in E_Y$ such that $f(\bar{y}) = \rho$. By definition of E_Y there exists $\bar{x} \in C$ such that $\bar{y} = B\bar{x} + d$. Hence $f(Bx + d) = \rho$ and we are done.

If $f(Bx + d)$ is unbounded below on C, then $f(y)$ is unbounded below on E_Y. Hence, there is a halfline $H = \{s + \mu t: \mu \geq 0\} \subseteq E_Y$ such that $f(s + \mu t) \to -\infty$ as $\mu \to \infty$. By using a theorem of the alternative [8] or otherwise, it can be shown that there is a halffline $G = \{\bar{s} + \mu \bar{t}: \mu \geq 0\} \subseteq C$ such that

$$s = B\bar{s} + d,$$
$$t = B\bar{t}.$$

Hence $f(Bx + d) \to -\infty$ on G as required.

(iii) It can be easily shown that $\mathcal{G}$ is closed under affine transformations of the variables $y = Bx + d$ where B has full row rank. The result for general B is not known at present.

(iv) The result of this paper can be slightly generalized by replacing in Definition 3.1 the class of norm-coercive functions by an arbitrary subclass of $\mathscr{F}$. That this new $\mathscr{G}$ will still be in $\mathscr{F}$ follows immediately from the proof in Section 5 where the only property required of norm-coercive functions is that they be in $\mathscr{F}$.

Acknowledgement

The author wishes to express his appreciation to Professors R.W. Cottle and V. Chvátal for their helpful suggestions during the preparation of this manuscript.

References

[1] E. Blum and W. Oettli, "Direct proof of the existence theorem in quadratic programming", *Operations Research* 20 (1972) 165–167.

[2] L. Collatz and W. Wetterling, *Optimization problems* (Springer, Berlin, 1975).

[3] J.B. Dennis, *Mathematical programming and electrical networks* (Wiley and Technical Press of MIT, New York, 1959).

[4] B.C. Eaves, "On quadratic programming", *Management Science* 17 (11) (1971) 698–711.

[5] M. Frank and P. Wolfe, "An algorithm for quadratic programming", *Naval Research Logistics Quarterly* 3 (1956) 95–110.

[6] A.J. Goldman, "Resolution and separation theorems for polyhedral convex sets", in: H.W. Kuhn and A.W. Tucker, eds., *Linear inequalities and related systems* (1956) pp. 41–51.

[7] W.M. Hirsch and A.J. Hoffman, "Extreme varieties, concave functions and the fixed charge problem", *Communications on Pure and Applied Mathematics* XIV (1961).

[8] J. Stoer and C. Witzgall, *Convexity and optimization in finite dimensions I* (Springer, New York, 1970).