JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 25, No. 3, JULY 1978

TECHNICAL NOTE

Calculating the Cone of Directions of Constancy[1]

H. WOLKOWICZ[2]

Communicated by A. V. Fiacco

Abstract. This note presents an algorithm that finds the cone of directions of constancy of a differentiable, faithfully convex function.

Key Words. Cone of directions of constancy, faithfully convex functions, gradient.

1. Introduction

For a function $f: R^n \to R$, the *cone of directions of constancy* of f at $x \in R^n$ is defined as

$$D_f(x) = \{d \in R^n: \exists \bar{\alpha} > 0 \ni f(x + \alpha d) = f(x) \text{ for all } 0 < \alpha < \bar{\alpha}\}.$$

If f is a differentiable convex function, then $D_f(x)$ is a convex cone (e.g., Ref. 1).

The cone of directions of constancy has been recently used in various characterizations of optimality (e.g., Refs. 1–4) and numerical algorithms (e.g., Ref. 5). This cone is of particular importance when f belongs to the class of *faithfully convex functions,* i.e., convex functions which are not affine along any line segment, unless they are affine along the entire line extending the segment (e.g., Ref. 6). In this case, this cone is a subspace independent of the choice of x (e.g., Ref. 4). The class of faithfully convex functions is large and it includes all analytic convex functions as well as all strictly convex functions. Note that, in the case of faithfully convex functions, one can, by using the cone of directions of constancy, produce dual programs which are linearly constrained (e.g., Ref. 6).

[1] This work was supported by the National Research Council of Canada. The author is indebted to Professor S. Zlobec for suggesting the topic and for his guidance.

[2] Graduate Student and Lecturer, McGill University, Department of Mathematics, Montreal, Quebec, Canada.

2. Algorithm

Suppose that $f: R^n \to R$ is a differentiable faithfully convex function. In this section, we formulate an algorithm that finds $D_f \cap \mathcal{R}(A_0)$, where D_f is the cone of directions of constancy, A_0 is any specified $n \times p$ matrix, and $\mathcal{R}(A_0)$ denotes the range space of A_0. Calculation of the intersection $D_f \cap \mathcal{R}(A_0)$ is useful in the situation when the intersection of two or more cones of directions of constancy is needed (e.g., Ref. 4). If $A_0 = I$, the identity matrix, then the algorithm calculates the cone of directions of constancy of f.

The algorithm is based on the fact that D_f lies in the orthogonal complement of $\nabla f(x)$, the gradient of f at x. By repeatedly considering the restriction of f to this orthogonal complement, we calculate D_f.

First we need a useful observation which is given without proof.

Lemma 2.1. Suppose that $0 \neq d \in R^k$ and i_0 is the smallest positive integer such that the i_0th component of d is nonzero, i.e., $d_{i_0} \neq 0$. Let

$$A = \left[\begin{array}{c|c} I_{(i_0-1)\times(i_0-1)} & 0 \\ \hline 0 & \begin{array}{c} d_{i_0+1}/d_{i_0} \ \dots \ d_k/d_{i_0} \\ \hline -I_{(k-i_0)\times(k-i_0)} \end{array} \end{array}\right].$$

Then,

$$\mathcal{R}(A) = \mathcal{N}(d),$$

where $\mathcal{N}(d)$ denotes the null space of d.

Let $E_k = \{e_i: i = 1, \dots, k\}$ denote the set of unit vectors in R^k and $A_0 \in R^{n\times p}$ be given.

Algorithm

Initialization. Set $P_0 = A_0$ and $i = 1$.

ith step, $1 \leq i \leq p$. Find a point x in the set of $p - i + 2$ vectors $\{0\} \cup E_{p-i+1}$ such that

$$\nabla f(P_{i-1}x)P_{i-1} \neq 0. \tag{1}$$

Case (i). If such an x exists and $i < p$, then, using Lemma 2.1, determine

$$A_i \in R^{(p-i+1)\times(p-i)}$$

such that

$$\mathscr{R}(A_i) = \mathscr{N}(\nabla f(P_{i-1}x)P_{i-1}). \tag{2}$$

Set

$$P_i = P_{i-1}A_i,$$

and proceed to step $i+1$.

Case (*ii*). If such an x exists but $i = p$, then stop.

Conclusion. $D_f \cap \mathscr{R}(A_0) = \{0\}$.

Case (*iii*). If such an x does not exist, then stop.

Conclusion. $D_f \cap \mathscr{R}(A_0) = \mathscr{R}(P_{i-1})$.

Theorem 2.1. Suppose that $f: R^n \to R$ is a faithfully convex function and A_0 is some given $n \times p$ matrix. Then, the above algorithm finds $D_f \cap \mathscr{R}(A_0)$ in at most $p - s + 1$ steps, where

$$s = \dim(D_f \cap \mathscr{R}(A_0)).$$

Proof. Let x^i denote the point x which satisfies (1) at the ith step; and, for $i \geq 0$, let $f_i = f \circ P_i$ denote the composite function formed by applying first P_i and then f. By the linearity of P_i, f_i is a faithfully convex function and so D_{f_i} is a fixed subspace of R^{p-i}. Furthermore,

$$\nabla f_i(x) = \nabla f(P_i x)P_i.$$

Now, suppose that Case (i) has occurred, i.e.,

$$x^i \in \{0\} \cup E_{p-i+1},$$

$$\nabla f_{i-1}(x^i) = \nabla f(P_{i-1}x^i)P_{i-1} \neq 0,$$

and $i < p$. Let us show that

$$D_f \cap \mathscr{R}(A_0) = P_i D_{f_i}. \tag{3}$$

First, let us show that

$$D_f \cap \mathscr{R}(A_0) = A_0 D_{f_0}. \tag{4}$$

Suppose that $d \in D_{f_0}$. This means that

$$f_0(\alpha d) = f_0(0)$$

for all $\alpha \in R$. By definition of f_0 and the linearity of A_0, this gives

$$f(\alpha A_0 d) = f(0)$$

for all $\alpha \in R$, i.e., $A_0 d \in D_f$. Furthermore, since $A_0 d \in \mathscr{R}(A_0)$,

$$A_0 d \in D_f \cap \mathscr{R}(A_0).$$

Conversely, suppose that $d \in D_f \cap \mathscr{R}(A_0)$. Then, there exists a $\bar{d} \in R^p$ such that

$$d = A_0 \bar{d} \qquad \text{and} \qquad f(\alpha A_0 \bar{d}) = f(0)$$

for all $\alpha \in R$. Again, by definition of f_0 and the linearity of A_0, we get that

$$f_0(\alpha \bar{d}) = f_0(0)$$

for all $\alpha \in R$, i.e., $\bar{d} \in D_{f_0}$, where $d = A_0 \bar{d}$. This proves (4).

Next, let us show that

$$D_{f_{i-1}} = A_i D_{f_i}, \qquad i \geq 1. \tag{5}$$

Suppose that $d \in D_{f_i}$. This means that

$$f_i(\alpha d) = f_i(0)$$

for all $\alpha \in R$. Since

$$f_i = f_{i-1} \circ A_i,$$

we get that

$$f_{i-1}(\alpha A_i d) = f_{i-1}(0)$$

for all $\alpha \in R$; i.e.,

$$A_i d \in D_{f_{i-1}}.$$

Conversely, suppose that $d \in D_{f_{i-1}}$, i.e.,

$$f_{i-1}(\alpha d) = f_{i-1}(0)$$

for all $\alpha \in R$. But

$$D_{f_{i-1}} \subset \mathcal{N}(\nabla f_{i-1}(x^i))$$

and

$$\mathcal{N}(\nabla f_{i-1}(x^i)) = \mathcal{N}(\nabla f(P_{i-1} x^i) P_{i-1}) = \mathscr{R}(A_i),$$

by (2). Therefore, there exists a $\bar{d} \in R^{p-i}$ such that $d = A_i \bar{d}$. So,

$$f_i(\alpha \bar{d}) = f_{i-1}(\alpha A_i \bar{d}) = f_{i-1}(\alpha d) = f_{i-1}(0) = f_i(0)$$

for all $\alpha \in R$, i.e., $\bar{d} \in D_{f_i}$ and $d = A_i \bar{d}$. This proves (5).

By repeated substitution of (5) into (4), one gets that

$$D_f \cap \mathscr{R}(A_0) = A_0 D_{f_0} = A_0 A_1 D_{f_1} = \cdots = P_i D_{f_i},$$

which proves (3).

Now, suppose that Case (ii) has occurred, i.e.,

$$x^i \in \{0\} \cup E_{p-i+1}, \qquad \nabla f_{i-1}(x^i) \neq 0,$$

but $i = p$. Since f_{p-1}: $R \to R$ is faithfully convex, we get that

$$D_{f_{p-1}} = \{0\}.$$

But, by (3), the $(p-1)$th step implies that

$$D_f \cap \mathscr{R}(A_0) = P_{p-1} D_{f_{p-1}}.$$

Substituting for $D_{f_{p-1}}$ yields the desired result that

$$D_f \cap \mathscr{R}(A_0) = \{0\}.$$

Finally, suppose that Case (iii) has occurred, i.e.,

$$\nabla f_{i-1}(y) = 0 \qquad \text{for all} \qquad y \in \{0\} \cup E_{p-i+1}.$$

Then, by the convexity of f_{i-1}, the complete set E_{p-i+1} lies in $D_{f_{i-1}}$. But $D_{f_{i-1}}$ is a subspace of R^{p-i+1}, and so we conclude that

$$D_{f_{i-1}} = R^{p-i+1}.$$

Substituting into (3) yields

$$D_f \cap \mathscr{R}(A_0) = \mathscr{R}(P_{i-1}). \qquad \square$$

The algorithm will be illustrated by two examples.

Example 2.1. Consider the function

$$f(x) = -(4 + (x_1 + x_2)^2)^{1/2} + x_1 + x_2 + x_3^2.$$

This function is convex and analytic, and so is faithfully convex. Let us determine its cone of directions of constancy D_f.

Initialization. Set

$$P_0 = A_0 = I_{3\times 3}, \qquad i = 1.$$

Step 1. Since

$$\nabla f(x) = (1 - (x_1 + x_2)/(4 + (x_1 + x_2)^2)^{1/2},\ 1 - (x_1 + x_2)/(4 + (x_1 + x_2)^2)^{1/2},\ 2x_3),$$

we see that $0 \in \{0\} \cup E_3$ and

$$\nabla f(P_0 0) P_0 = \nabla f(0) = (1, 1, 0) \neq 0.$$

Furthermore, since $i=1<p=3$, we are in Case (i). Using Lemma 2.1, we find that

$$P_1 = A_0A_1 = A_1 = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Step 2. For $x \in R^2$,

$$\nabla f(P_1x)P_1 = \nabla f\begin{pmatrix} x_1 \\ -x_1 \\ -x_2 \end{pmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} = (0, 2x_2).$$

Therefore,

$$\nabla f(P_1e_2)P_1 = (0, 2) \neq 0,$$

where $e_2 \in E_2$. Furthermore, since $i=2<p$, we are in Case (i) again, and so we find that

$$A_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad P_2 = P_1A_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Step 3. The finite point set $\{0\} \cup E_{p-i+1}$ is $\{0, 1\}$ and

$$\nabla f(P_2 0)P_2 = \nabla f(P_2 1)P_2 = (1, 1, 0)\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = 0.$$

Therefore, we are in Case (iii) and

$$D_f = \mathcal{R}(P_2) = \left\{ \begin{bmatrix} d \\ -d \\ 0 \end{bmatrix} \in R^3 : d \in R \right\}.$$

Example 2.2. Now consider the faithfully convex function

$$g(x) = -x_1 - x_2 + x_3^2,$$

and suppose that we wish to find

$$D_g \cap D_f = D_g \cap \mathcal{R}\left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\right),$$

where f is the function in the previous example.

Initialization. Set

$$P_0 = A_0 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \qquad i = 1.$$

Step 1. Since

$$p = 1, \qquad \nabla g(x) = (-1, -1, 2x_3),$$

we see that $\{0\} \cup E_1 = \{0, 1\}$ and that

$$\nabla g(P_0 0)P_0 = \nabla g(P_0 1)P_0 = 0.$$

Therefore, we are in Case (iii) and

$$D_f \cap D_g = \mathscr{R}(P_0) = \left\{ \begin{pmatrix} d \\ -d \\ 0 \end{pmatrix} \in R^3 : d \in R \right\}.$$

The author has implemented the above algorithm on the IBM/370 at McGill University.

References

1. BEN-ISRAEL, A., BEN-TAL, A., and ZLOBEC, S., *Optimality Conditions in Convex Programming,* Proceedings of the 9th International Symposium on Mathematical Programming, Budapest, Hungary, 1976.
2. ABRAMS, R. A., and KERZNER, L., *A Simplified Test for Optimality,* Journal of Optimization Theory and Applications (to appear).
3. BEN-ISRAEL, A., BEN-TAL, A., and ZLOBEC, S., *Characterization of Optimality in Convex Programming without a Constraint Qualification,* Journal of Optimization Theory and Applications, Vol. 20, pp. 417–437, 1976.
4. BEN-ISRAEL, A., and BEN-TAL, A., *On a Characterization of Optimality in Convex Programming,* Mathematical Programming, Vol. 11, pp. 81–88, 1976.
5. BEN-TAL, A., and ZLOBEC, S., *A New Class of Feasible Direction Methods,* The University of Texas, Austin, Texas, Center for Cybernetics Studies, Report No. CCS-216, 1977.
6. ROCKAFELLAR, R. T., *Some Convex Programs Whose Duals are Linearly Constrained,* Nonlinear Programming, Edited by J. B. Rosen, O. L. Mangasarian, and K. Ritter, pp. 293–322, Academic Press, New York, New York, 1970.