REFERENCES

1. T. M. Apostol, *Linear Algebra: A First Course with Applications to Differential Equations*, Wiley, New York, 1997.
2. S. N. Elaydi, *An Introduction to Difference Equations*, 2nd ed., Springer-Verlag, New York, 1999.
3. S. N. Elaydi and W. A. Harris, Jr., On the computation of A^n, *SIAM Rev.* **40** (1998) 965–971.
4. I. E. Leonard, The matrix exponential, *SIAM Rev.* **38** (1996) 507–512.
5. E. J. Putzer, Avoiding the Jordan canonical form in the discussion of linear systems with constant coefficients, this MONTHLY **73** (1966) 2–7.

Department of Basic Sciences, University of Sharjah, P.O.Box 27272, Sharjah, United Arab Emirates.
rabusaris@sharjah.ac.ae

Department of Electrical and Electronics Engineering, University of Sharjah, P.O.Box 27272, Sharjah, United Arab Emirates.
wajdi@sharjah.ac.ae

Exact Solutions for Minimax Optimization Problems

Francesc Comellas and J. Luis A. Yebra

1. INTRODUCTION. Functions defined as pointwise maxima or minima of a finite set of smooth functions need not be differentiable, thus disallowing the characterization of their extrema by the standard methods of elementary calculus. However, the minimization of functions such as $f = \max\{f_1, f_2, \ldots, f_m\}$ do often surface in applications, so methods for sorting out their critical points may be of interest. In fact, an appropriate characterization can be used to find effectively exact solutions to problems of this type that are generally studied in the field of nonsmooth optimization and solved approximatively by various algorithmic methods (see [**1**], [**3**], [**5**], and [**6**]). A concrete application by the authors can be found in [**2**].

To give a feel for the problem, we consider first a one variable example in which we look for the minima of the function

$$f(x) = \max\left\{f_1(x),\ f_2(x)\right\} = \max\left\{x^2,\ \frac{8+4x-x^2}{3}\right\}$$

whose graph is shown in Figure 1. The function has two critical points at which $f'(x)$ does not exist: $x^* = -1$ and $x^* = 2$. It attains its minimum value at the former, while at the latter it does not have even a local minimum. This may be inferred from a known version of the first derivative test that assumes the existence of the left- and right-hand derivatives of f at x^* and states: $f'(x^{*-}) \leq 0$ and $f'(x^{*+}) \geq 0$ are necessary conditions for f to have a minimum at x^*, while $f'(x^{*-}) < 0$ and $f'(x^{*+}) > 0$ are sufficient conditions for it. In the example f satisfies the necessary and sufficient conditions at -1 but only the necessary conditions at 2. Of course, reversing all the inequalities characterizes a maximum. Now a simple and common characterization of an extremum (either a maximum or a minimum) comes out: namely, the existence of

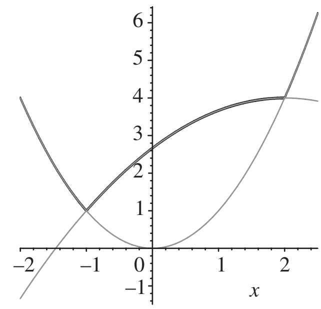

Figure 1. $f(x) = \max\{x^2,\ (8 + 4x - x^2)/3\}$.

constants α and β such that $\alpha + \beta > 0$ and

$$\alpha f'(x^{*-}) + \beta f'(x^{*+}) = \alpha f_1'(x^*) + \beta f_2'(x^*) = 0$$

is a sufficient condition for an extremum when both constants are required to be positive, whereas it is a necessary condition only when one of them is allowed to vanish. Moreover, this characterization can be carried over to the case of several variables.

2. CHARACTERIZATION OF THE MINIMA. Consider now the minimization of a real function f defined in some open set in R^n as the maximum of m continously differentiable functions of $\mathbf{x} = (x_1, x_2, \dots, x_n)$:

$$f = \max\{f_1, f_2, \dots, f_m\}. \tag{1}$$

Besides regular critical points $\mathbf{x}^*$ where some function f_i is larger than the others, hence characterized by $\nabla f_i(\mathbf{x}^*) = 0$, we have to study the critical points $\mathbf{x}^*$ where the values of several functions coincide and their common value is larger than those of the remaining functions. The following characterization can be found in [**3**, chap. 3, sec. 2] and [**6**, sec. 2.1] in a somewhat more general framework.

Theorem 1. *Let* $f = \max\{f_1, f_2, \dots, f_m\}$, *where the functions* f_i *are continously differentiable in some open set in* R^n, *and let* $\mathbf{x}^*$ *be a critical point of* f. *Suppose that the labeling of the functions is such that*

$$f_1(\mathbf{x}^*) = f_2(\mathbf{x}^*) = \cdots = f_{k+1}(\mathbf{x}^*) > f_{k+2}(\mathbf{x}^*) \geq \cdots \geq f_m(\mathbf{x}^*) \tag{2}$$

for some k. *Then a necessary condition for* f *to have a local minimum at* $\mathbf{x}^*$ *is that the* $k + 1$ *gradients* $\nabla f_1(\mathbf{x}^*), \nabla f_2(\mathbf{x}^*), \dots, \nabla f_{k+1}(\mathbf{x}^*)$ *have a vanishing nontrivial linear combination with nonnegative coefficients:*

$$\sum_{i=1}^{k+1} c_i \nabla f_i(\mathbf{x}^*) = \mathbf{0}, \qquad c_i \geq 0, \qquad \sum_{i=1}^{k+1} c_i > 0. \tag{3}$$

Moreover, when $k = n$ *and the* $n + 1$ *gradients* $\nabla f_i(\mathbf{x}^*)$ $(1 \leq i \leq n + 1)$ *span* R^n, *a sufficient condition for* f *to have a strict minimum at* $\mathbf{x}^*$ *is that (3) hold with positive coefficients* c_i:

$$\sum_{i=1}^{n+1} c_i \nabla f_i(\mathbf{x}^*) = \mathbf{0}, \qquad c_i > 0. \tag{4}$$

Characterizations (3) and (4) still hold when looking for the maxima of the minimum of a (finite) set of functions.

3. EXACT SOLUTIONS. Any point x^* at which (3) holds is called a *stationary point* of f. While the determination of global minima is often accomplished via different iterative algorithmic methods (see, for instance, [**6**]), in many cases stationary points can be determined exactly through an appropriate use of condition (3). Indeed, when $k < n$ the linear dependence expressed by (3) provides $n - k$ equations that, together with the k equations $f_1(\mathbf{x}^*) = f_2(\mathbf{x}^*) = \cdots = f_{k+1}(\mathbf{x}^*)$, lead to the critical points among which stationary points (and thus extrema) are to be found, as shown in the following examples.

Example 1. We look here for the point closest to three given points in the plane (i.e., at minimum distance from the furthest one). In the preceding formulation, we must find the minimum of the maximum of three functions: $f = \max\{f_1, f_2, f_3\}$, where f_1, f_2, and f_3 are the (squared) distances from a point (x, y) to the given points.

1. Consider first the points $(-4, 0)$, $(4, 0)$, and $(0, -2)$. Then

$$f_1(x, y) = (x+4)^2 + y^2, \quad f_2(x, y) = (x-4)^2 + y^2, \quad f_3(x, y) = x^2 + (y+2)^2.$$

We have $f_1(x, y) = f_2(x, y) = f_3(x, y) = 25$ at $(x^*, y^*) = (0, 3)$. At this point

$$5\nabla f_1(0, 3) + 5\nabla f_2(0, 3) - 6\nabla f_3(0, 3) = 5(8, 6) + 5(-8, 6) - 6(0, 10) = (0, 0).$$

Since the necessary condition (3) does not hold, no minimum can be attained at this point.

Considering now $f_1(x, y) = f_2(x, y)$, we obtain $x = 0$, and for points on this line $\nabla f_1(0, y) = (8, 2y)$ and $\nabla f_2(0, y) = (-8, 2y)$, which are linearly dependent only when $y = 0$. Now, at $(x^*, y^*) = (0, 0)$ we have $f_1(0, 0) = f_2(0, 0) = 16 > f_3(0, 0) = 4$ and also

$$\nabla f_1(0, 0) + \nabla f_2(0, 0) = (8, 0) + (-8, 0) = (0, 0),$$

satisfying the necessary condition (3). Direct verification along the line $x = 0$, where $f(0, y) = 16 + y^2$, shows that the function does attain its minimum value at this point.

2. Now let the points be $(0, 0)$, $(2, 0)$, and $(0, 2)$, so that

$$f_1(x, y) = x^2 + y^2, \quad f_2(x, y) = (x-2)^2 + y^2, \quad f_3(x, y) = x^2 + (y-2)^2.$$

These functions satisfy $f_1(x, y) = f_2(x, y) = f_3(x, y) = 2$ at $(x^*, y^*) = (1, 1)$, at which point

$$0\nabla f_1(1, 1) + \nabla f_2(1, 1) + \nabla f_3(1, 1) = 0(2, 2) + (-2, 2) + (2, -2) = (0, 0).$$

Thus at $(1, 1)$ the necessary condition (3) is satisfied, but not the sufficient condition (4). However, direct verification ($f(1 + \epsilon_1, 1 + \epsilon_2) - f(1, 1) \geq \epsilon_1^2 + \epsilon_2^2$ near $(1, 1)$) shows that the function does attain its minimum value at this point.

3. Finally, consider the points $(-4, 0)$, $(4, 0)$, and $(0, 8)$. In this instance

$$f_1(x, y) = (x+4)^2 + y^2, \quad f_2(x, y) = (x-4)^2 + y^2, \quad f_3(x, y) = x^2 + (y-8)^2.$$

Now $f_1(x, y) = f_2(x, y) = f_3(x, y) = 25$ at $(x^*, y^*) = (0, 3)$, where

$$5\nabla f_1(0, 3) + 5\nabla f_2(0, 3) + 6\nabla f_3(0, 3) = 5(8, 6) + 5(-8, 6) + 6(0, -10) = (0, 0).$$

At $(0, 3)$ both the necessary condition (3) and the sufficient condition (4) are satisfied, so this is the point we were after (see Figure 2a).

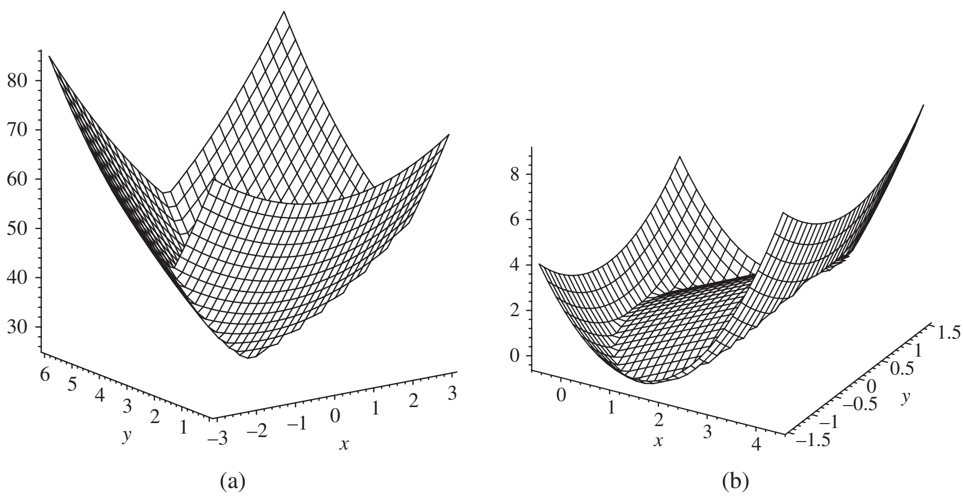

Figure 2. (a) Example 1, case 3; (b) Example 2.

Of course, this is an easy example because each function f_i (hence f) is convex, so the minimum is always attained: at an inner point (the triangle's circumcenter) when the three points are the vertices of an acute triangle, and at the midpoint of the longest side when the points are the vertices of an obtuse or right triangle. The next example describes quite a different situation.

Example 2. Consider

$$f(x, y) = \max\{f_1(x, y), f_2(x, y)\} = \max\{x - y^2, -3x + x^2 + y^2\}$$

(see Figure 2b). We have $f_1 = f_2$ on the ellipse $E : x^2 - 4x + 2y^2 = 0$, with $f_1 > f_2$ inside and $f_1 < f_2$ outside E. Inside the ellipse $\nabla f_1(x, y) = (1, -2y) \neq (0, 0)$, while outside it $\nabla f_2(x, y) = (-3 + 2x, 2y) \neq (0, 0)$ as well. Next, for points on the ellipse ∇f_1 and ∇f_2 can be linearly dependent only under the following conditions:

(a) if $y \neq 0$, then $x = 1$, which yields the two points $A = (1, \sqrt{3/2})$ and $B = (1, -\sqrt{3/2})$;

(b) if $y = 0$, then $x^2 - 4x = 0$, giving two new points, $C = (0, 0)$ and $D = (4, 0)$.

At A and B, $\nabla f_1 + \nabla f_2 = 0$. Parameterizing the ellipse E in the standard way by $\mathbf{g} : [-\pi, \pi] \to R^2$ we obtain

$$\mathbf{g}(s) = \big(x(s), y(s)\big) = (2 - 2\cos s,\ \sqrt{2}\sin s),$$

with $A = \mathbf{g}(\pi/3)$ and $B = \mathbf{g}(-\pi/3)$. Set

$$\phi(s) = f\big[\mathbf{g}(s)\big] = 2\,(\cos^2 s - \cos s).$$

Then $\phi'(\pm\pi/3) = 0$ and $\phi''(\pm\pi/3) = 3 > 0$, revealing that f attains its minimum value $\phi(\pm\pi/3) = -1/2$ at these points. At C, we get $3\nabla f_1 + \nabla f_2 = 0$. Since $C = \mathbf{g}(0)$ and $\phi'(0) = 0$ but $\phi''(0) = -2 < 0$, f does not attain a minimum at C. Lastly, at the point D, $\nabla f_1 = (1, 0)$ and $\nabla f_2 = (5, 0)$, so the necessary condition (3) does not hold at D.

Finally, we consider one of the standard test examples used to compare different algorithms in nonsmooth optimization (see [**5**, p. 139] or [**4**]). Most problems in these tests admit analogous treatment.

Example 3. We look for the minimum of

$$\begin{aligned} f(x, y) &= \max\left\{f_1(x, y),\ f_2(x, y),\ f_3(x, y)\right\} \\ &= \max\left\{x^2 + y^4,\ (x-2)^2 + (y-2)^2,\ 2e^{-x+y}\right\}. \end{aligned}$$

All the algorithms find approximate solutions near $(x', y') = (1.14, 0.9)$, for which

$$f_1(x', y') = 1.9557, \quad f_2(x', y') = 1.9496, \quad f_3(x', y') = 1.5733.$$

In our context we have to study the behavior of f at a critical point (x^*, y^*) for which $f_1(x^*, y^*) = f_2(x^*, y^*) > f_3(x^*, y^*)$. Therefore (x^*, y^*) must be a solution of

$$x^2 + y^4 = (x-2)^2 + (y-2)^2, \quad x(y-2) = 2y^3(x-2),$$

where the second equation comes from the linear dependence of the gradients $\nabla f_1(x, y) = (2x, 4y^3)$ and $\nabla f_2(x, y) = 2(x-2, y-2)$. Solving these equations, we find the "exact" optimal solution:

$$\begin{aligned} x^* &= 1.139037651992656, \\ y^* &= 0.899559938395897, \\ f(x^*, y^*) &= 1.952224493870659. \end{aligned}$$

ACKNOWLEDGMENTS. This research was supported by the MCyT Grant TIC-2002-00155, Spain.

REFERENCES

1. F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolensky, *Nonsmooth Analysis and Control Theory*, Springer-Verlag, New York, 1998.
2. F. Comellas and J. L. A. Yebra, New lower bounds for Heilbronn numbers, *Elect. J. Combinatorics* **9** (2002) R6.
3. V. F. Dem'yanov and V. N. Malozenov, *Introduction to Minimax*, John Wiley & Sons, New York, 1974; reprinted by Dover, New York, 1990.
4. L. Luksan and J. Vlcek, *Test Problems for Nonsmooth Unconstrained and Linearly Constrained Optimization*, ICS Technical Report, no. 798, 2000; available at http://www.cs.cas.cz/~luksan/test.html
5. M. M. Mäkelä and P. Neittaanmäki, *Nonsmooth Optimization*, World Scientific, Singapore, 1992.
6. E. Polak, *Optimization*, Springer-Verlag, New York, 1997.

Departament de Matemàtica Aplicada IV, E. P. S. C., Universitat Politècnica de Catalunya, Av. del Canal Olímpic, s.n., 08860 Castelldefels, Catalonia, Spain
{*comellas,yebra*}*@mat.upc.es*