Global Minimum Point of a Convex Function

J. M. Soriano

Departamento de Análisis Matemático
Facultad de Matemáticas
Universidad de Sevilla
Apart. 1160
41080 Sevilla, Spain

ABSTRACT

In this paper we prove the existence of an unique global minimum point of a convex function under some smoothness conditions. Our proof permits us to calculate numerically such a minimum point utilizing a constructive homotopy method.

1. INTRODUCTION

With the invention of high-speed computers, large-scale problems from such diverse fields as economics, agriculture, military planning, and flows in networks became at least potentially solvable, a lot of them being extremum problems.

The great importance of extremum problems in applied mathematics leads us to the general study of the extremum of functions from $\mathbf{R}^n$ to $\mathbf{R}$. It is not easy to know the extremum points, for differentiable functions because it is not always possible to solve the equation $\nabla f(x) = 0$ to calculate critical points. Convex functions have a particularly simple extremal structure [2], and there exist algorithms to calculate extremum points, supposing its existence. However, it is not easy to prove the existence of extremum even in the case of convex differentiable functions [2, 3]. Therefore, it is very important to give sufficient conditions to guarantee this existence.

2. PROOF OF A UNIQUE GLOBAL MINIMUM POINT

Given a strictly convex function f from $\mathbf{R}^n$ to $\mathbf{R}$, we prove the existence of a unique global minimum point for f if the following condition is verified:

655 Avenue of the Americas, New York, NY 10010 0096-3003/93/$6.00

$$\lim_{X \to \infty} \frac{\delta f(X)/\delta X_i}{X_i} > 0 \quad \text{for any value of } i \in 1, \ldots, n. \tag{1}$$

This proof is founded on the continuation method and the method can serve to determine that point numerically as we have shown in [5, 6]; see also [1], and [7].

THEOREM 1. *Let $f: \mathbf{R}^n \to \mathbf{R}$, $f \in C^3(\mathbf{R}^n)$ be a strictly convex function verifying* (1). *Then there exists an unique minimum point for f.*

PROOF. We have $\nabla f: \mathbf{R}^n \to \mathbf{R}^n$, with $\nabla f \in C^2(\mathbf{R}^n)$. We construct the function

$$H: \mathbf{R}^n[0,1] \to \mathbf{R}^n, \qquad H(X,t) = (1-t)X + t\,\nabla f(X).$$

1) Let us first prove that zero is a regular value for H. Since f is strictly convex, zero is a regular value for ∇f, and hence

$$\operatorname{rank}\left(\frac{\delta^2 f(X)}{\delta X_i \delta X_j}\right)_{1 \leqslant i \leqslant j \leqslant n} = n.$$

Moreover, the former matrix is the matrix of a symmetric definite positive bilinear transformation.

Let us now consider the matrix

$$H_X(X,t) = \begin{pmatrix} (1-t) + t\dfrac{\delta^2 f(X)}{\delta X_1^2} & \cdots & t\dfrac{\delta^2 f(X)}{\delta X_n \delta X_1} \\ \cdots & \cdots & \cdots \\ t\dfrac{\delta^2 f(X)}{\delta X_1 \delta X_n} & \cdots & (1-t) + t\dfrac{\delta^2 f(X)}{\delta X_n^2} \end{pmatrix}$$

and the symmetric bilinear transformation

$$hH_X(X,t)h^T = t\left[h\left(\frac{\delta^2 f(X)}{\delta X_i \delta X_j}\right)_{1\leqslant i\leqslant j\leqslant n} h^T\right] + (1-t)hh^T$$

with $h \in \mathbf{R}^n$. If $h \neq 0$ and $t \in (0,1]$, both summands are greater than zero, whereby

$$hH_X(X,t)h^T > 0,$$

if $t \in [0,1]$. Therefore, $H_X(X,t)$ is the matrix of a symmetric definite positive bilinear transformation, and so Sylvester's theorem implies that $\det H_X(X,t) > 0$. Thus, zero is a regular value for H and for $H|\delta(R^n[0,1])$. Moreover, $H \in C^2$ as composition of C^2-functions.

2) Let's now prove that $H^{-1}(0)$ includes an arc passing through the point $X = 0, t = 0$. We have for every point

$$Y^0 = (X^0, t^0) \in H^{-1}(0), \qquad \det H_X(X^0, t^0) \neq 0$$

and so, the Implicit Function Theorem implies the existence of a neighborhood N of t^0 and exactly one function $g \in C^2$, such that $g(t^0) = X^0$, $H(g(t), t) = 0$, $\forall t \in N$. Therefore, $H^{-1}(0)$ consists of arcs and only arcs, and as $(0,0) \in H^{-1}(0)$, there exists an arc C of $H^{-1}(0)$ passing through $(0,0)$.

3) Let's see that, in the analytical continuation of C, the coordinate t is strictly monotonous as a function of the arc length s. Let us parameterize C with respect to s,

$$Y = (Y_1, \ldots, Y_{n+1}) = (X, t) = Y(s).$$

When s grows, $Y(s)$ describes C and we have in a 0-neighbourhood

$$H(Y(s)) = 0. \tag{2}$$

Differentiating, we obtain

$$\sum_{i=1}^{n+1} \frac{\delta H(Y(s))}{\delta Y_i} \frac{dY_i}{ds} = 0,$$

or equivalently

$$H'(Y)\left(\frac{dY}{ds}\right)^T = 0. \tag{3}$$

Let's consider the linear system (3), clearly indeterminated. Since

$$\det H_X(X,t) \neq 0,$$

as solution of (3) is

$$\frac{dY_i}{ds} = (-1)^i \det H'_{-i}(Y), \quad i = 1,\dots,n+1, \tag{4}$$

where $H'_{-i}(Y)$ is the result of suppressing the ith column in $H'(Y)$. The initial value problem formed by the system (4) and an initial value

$$(s^*, Y^0), Y^0 = Y(s^*) \in C$$

has a unique solution of class two (Picard-Lindelof theorem) $Y^*(s)$ defined on $D = [s^*, s^* + k]$ ($k \in \mathbf{R}^+$) that verifies (1) and $Y^* \equiv C$ on $t(D) \cap N$. We define

$$u\colon D \subseteq \mathbf{R} \to \mathbf{R}^{n+1} \quad \text{by } u(s) = H(Y^*(s)).$$

Differentiating,

$$\frac{du(s)}{ds} = \sum_{i=1}^{n+1} \frac{\delta H(Y^*(s))}{\delta Y_i} \frac{dY_i^*(s)}{ds} = 0 \Rightarrow u(s) = \text{constant},$$

but

$$u(s^*) = H(Y^*(s^*)) = H(Y^0) = 0 \Rightarrow H(Y^*(s)) = 0.$$

For any continuation of $Y^*(s)$, the t-coordinate verifies $dt/ds = (-1)^{n+1}\det H_X(X,t)$, and $\det H_X(X,t) > 0$. Therefore, $t(s)$ is strictly monotonous.

4) Let us now show that Y^* can only be continued by bounded values of $\|X\|$. This is a consequence of the condition (1) of the theorem because the equation

$$t\,\nabla f(X) + (1-t)X = 0$$

leads us to

$$\frac{\delta f(x)/\delta X_i}{X_i} = \frac{1-t}{-t}, \qquad i = 1, \dots, n, \tag{5}$$

and $t \in (0, 1]$, $(1 - t/-t) \leqslant 0$ (or $-\infty$ when $t \to 0$). But (5) is absurd for a sufficiently great $\|X\|$ since it implies

$$\lim_{\|X\| \to \infty} \frac{\delta f(X)/X_i}{X_i} \leqslant 0$$

against the hypothesis.

5) The right extreme point of the maximal continuation of Y^* belongs to the hyperplane $t = 1$. That follows from det $H'_{-i}(Y)$ $(i = 1, \dots, n)$ being continuously differentiable, and the system (4) autonomous; so it is possible to continue Y^* to the boundary of $\mathbf{R}^n[0, 1]$ with $s \in [0, +\infty)$ [4]. Clearly, that right extreme point T^+ of the maximal prolongation of Y^* cannot coincide with the initial value $t = 0$ by 3), and the trajectory of this prolongation is only defined for bounded values of X by 4). Therefore,

$$T^+ = (A, 1) \in \mathbf{R}^n\{1\}.$$

6) Finally, let's note that A is a minimum local point for $f(X)$ since

$$\lim_{(X,t) \to (A,1)} H(X, t) = 0$$

implies that $\nabla f(A) = 0$ due to the continuity of H.
Theorem B [2, p. 124] implies that this minimum is global and A is the unique minimum point for f. ■

We don't develop here the numerical aspect, but is has been developed in similar conditions in other papers of ours [5, 6].

REFERENCES

1 C. B. Garcia and W. I. Zangwill, *Pathways to Solution, Fixed Points and Equilibria*, Prentice-Hall Series in Computational Mathematics. Prentice-Hall, London, 1981.

2 A. W. Roberts and D. E. Varberg, Convex functions. *Pure Appl. Math.* 57 (1973).
3 R. T. Rockafellar, *Convex Analysis*, Princeton Mathematical Series 28 Princeton University Press, Princeton, New Jersey, 1970.
4 I. G. Petrovski, *Ordinary Differential Equations*, Dover Publications, New York, 1966.
5 J. M. Soriano, Sobre las existencia y el cálculo de ceros de funciones regulares, *Rev. Real Acad. Cienc. Exact. Fis. Natur. Madrid* LXXXII(3-4):523–531 (1988).
6 J. M. Soriano, A special type of triangulation in numerical nonlinear analysis, *Collect. Math.* 41(1):45–58 (1990).
7 E. Zeidler, *Nonlinear Functional Analysis and its Applications*, Springer-Verlag, New York, 1985.