R. P. Brent

On the Davidenko-Branin Method for Solving Simultaneous Nonlinear Equations

Abstract: It has been conjectured that the Davidenko-Branin method for solving simultaneous nonlinear equations is globally convergent, provided that the surfaces on which each equation vanishes are homeomorphic to hyperplanes. We give an example to show that this conjecture is false. A more complicated example shows that the method may fail to converge to a zero of the gradient of a scalar function, so the associated method for function minimization is not globally convergent.

Introduction

In an attempt to increase the domain of convergence of iterative methods for solving systems of nonlinear equations, various techniques involving the solution of an associated system of differential equations have been proposed [1–6]. Particularly promising is Branin's modification [1,2] of a method proposed by Davidenko [3,4]. To solve the system

$$f(x) = 0 \tag{1}$$

of n nonlinear equations in n unknowns, a parameter t is introduced, and the system of differential equations

$$\frac{d}{dt} f[x(t)] = S[x(t)]f[x(t)] \tag{2}$$

is integrated numerically, both forward and backward in t, from some starting point $x(0)$. Here $S(x)$ is the sign of the determinant of the Jacobian matrix

$$J(x) = (\partial f_i/\partial x_j) . \tag{3}$$

(We assume that each f_i is continuously differentiable and that a continuous trajectory $x(t)$ exists such that (2) holds except on a set having no limit points.) At points where $J(x)$ is nonsingular, (2) may be written as

$$dx/dt = S(x)J^{-1}(x)f(x) , \tag{4}$$

which shows the connection with Newton's method.

It is important to observe that, with the assumptions given above,

$$f[x(t)] = c(t)f[x(0)] , \tag{5}$$

where $c(t)$ is a positive continuous scalar function of t. In fact, if $J[x(t)]$ is nonsingular for $t \in [t_0, t_1)$, then

$$c(t) = c(t_0)\exp\{S[x(t_0)](t - t_0)\} \tag{6}$$

on $[t_0, t_1]$.

The idea of the Davidenko-Branin method is that, as t tends to $+\infty$ or $-\infty$, $c(t)$ may tend to zero and $x(t)$ may tend to a zero of f. For this to be true if $n \geq 2$, a condition that might be imposed on f is that the "Davidenko surfaces"

$$S_i = \{x \in R^n | f_i(x) = 0\} \tag{7}$$

are homeomorphic to hyperplanes. The reason for imposing this condition is that if all components of $f[x(0)]$ are nonzero, all components of $f[x(t)]$ are nonzero for all t [from Eq. (5)], so the trajectory $x(t)$ never crosses any of the S_i. Thus, if some S_i includes the surface of an n-sphere B containing $x(0)$, then $x(t)$ can never leave B, and a zero outside B can never be approached by $x(t)$.

It has been conjectured [1,2] that the method is globally convergent if the Davidenko surfaces S_i are homeomorphic to hyperplanes. In the next section we give a simple example to show that this conjecture is false: for certain starting values the method fails to converge, even though $f(x)$ has a unique, simple zero. The example is for $n = 2$, but similar examples may easily be constructed for $n > 2$. In the final section we give an example in which f is the gradient of a scalar function. This example is of interest because it shows that the Davidenko-Branin method, when applied to the gradient of a function, may fail to find a local (much less a global) minimum or maximum.

Counterexample

Let $f(x)$ be the vector function of two variables defined by

$$f_1(x_1,x_2) = 4(x_1 + x_2) \tag{8}$$

and

$$f_2(x_1,x_2) = 4(x_1 + x_2) + (x_1 - x_2)[(x_1 - 2)^2 + x_2^2 - 1], \tag{9}$$

and let C be the circular disc described by $(x_1 - 2)^2 + x_2^2 \leq 1$.

• *Theorem 1*

The curves $S_i = \{x \in R^2 | f_i(x) = 0\}$, $i = 1, 2$, are homeomorphic to straight lines and intersect only at $x = 0$. If $x(t)$ is any continuous trajectory satisfying Eq. (5) [for any scalar function $c(t)$], and $x(0) \in C$, then $x(t) \in C$ for all t.

The proof follows from some simple lemmas.

• *Lemma 1*

$f(x) = 0$ has a unique solution $x = 0$ in R^2.

Proof

If $x = (x_1,x_2)$ is a solution, then (8) and (9) give $x_2 = -x_1$ and $x_1[2(x_1 - 1)^2 + 1] = 0$, and the unique real solution is thus $x_1 = x_2 = 0$.

• *Lemma 2*

For all $x \in R^2$, $\partial f_2/\partial x_1 \geq 1$.

Proof

From (9),

$$\partial f_2/\partial x_1 = (x_1 - x_2 - 2)^2 + 2(x_1 - 1)^2 + 1 \geq 1. \tag{10}$$

• *Lemma 3*

The curve $S_2 = \{x \in R^2 | f_2(x) = 0\}$ is homeomorphic to a straight line.

Proof

Lemma 2 shows that for each $x_2 \in R$, there is precisely one value of x_1, say $x_1 = g(x_2)$, such that $f_2(x_1,x_2) = 0$. Also, it is easy to see that $g(x_2)$ is continuously differentiable. Hence $S_2 = \{(g(x_2),x_2) | x_2 \in R\}$ is homeomorphic to a straight line.

• *Lemma 4*

If $x_1 \geq |x_2|$, then

$$f_2(x_1,x_2) \geq x_1. \tag{11}$$

Proof

By Lemma 2, it is sufficient to prove (11) for $x_1 = |x_2|$. If $x_1 = x_2 \geq 0$, then, from (9), $f_2(x_1,x_2) = 8x_1 \geq x_1$. On the other hand, if $x_1 = -x_2 > 0$, then

$$f_2(x_1,x_2) = 2x_1(2(x_1 - 1)^2 + 1) \geq x_1. \tag{12}$$

• *Lemma 5*

Let S be the annulus $1 < (x_1 - 2)^2 + x_2^2 < 2$. Then $f_2(x) > f_1(x) > 0$ for $x \in S$, and $f_1(x) \geq f_2(x) > 0$ for $x \in C$.

Proof

This proof follows easily from (8), (9) and Lemma 4.

Theorem 1 now follows Lemma 5, using Eq. (5) and the continuity of $x(t)$: The trajectory $x(t)$ can never cross S and so must remain in C. In particular, $x(t)$ can never approach zero.

Global minimization

Since it has been proposed in Ref. 7 that the Davidenko-Branin method should be used to find the global minimum and/or maximum of a function of several variables, the following theorem is of interest. It shows that the method may fail to converge to either a minimum or a maximum.

• *Theorem 2*

Let

$$F(x) = (x_1 + 10)^2 + (x_2 + 10)^2 + \exp[-(x_1^2 + x_2^2)],$$

$$f(x) = (\partial F/\partial x_i),$$

for suitable $\alpha > 1$

and let $x(t)$ be a continuous trajectory satisfying Eq. (5) [for some scalar function $c(t)$] with $x(0) = \frac{1}{3}(1, -1)$. Then the curves $S_i = \{x \in R^2 | f_i(x) = 0\}$ are homeomorphic to straight lines and the vector equation $f(x) = 0$ has a unique solution $x = y$ [near $(-10, -10)$], but $x(t)$ does not approach y. In fact, $x(t)$ lies in a component T_1 of the set $T = \{x \in R^2 | f_2(x) > f_1(x) > 0\}$, and y is not in the closure of T_1.

Remarks

The proof of Theorem 2 is similar to that of Theorem 1, and the details are omitted. The idea is that the term $\exp[-(x_1^2 + x_2^2)]$ introduces a slight bump in the graph of $F(x)$. The bump is not sufficient to introduce another stationary point or to appreciably perturb the curves S_i, but it is sufficient to introduce the component T_1 of T, and Eq. (5) shows that $x(t) \in T_1$ for all t.

In practice this kind of breakdown of the method seems quite likely to occur when the objective function F is not convex. If F is twice continuously differentiable, strictly convex, and has bounded level sets, then the method is probably globally convergent, but so are many other methods.

Acknowledgment

I express my gratitude to Philip Wolfe for several helpful suggestions and discussions.

References

1. F. H. Branin, "Solution of Nonlinear DC Network Problems via Differential Equations," *Memoirs IEEE Conference on Systems, Networks and Computers*, Oaxtepec, Mexico, 1971.
2. F. H. Branin, "A Globally Convergent Method for Finding Multiple Solutions of Nonlinear Simultaneous Equations," *Conference on Numerical Methods for Nonlinear Optimization*, Dundee, Scotland, 1971. (An expanded version of this paper will appear in this Journal).
3. D. F. Davidenko, "On a New Method of Numerical Solution of Systems of Nonlinear Equations" (in Russian), *Dokl. Akad. Nauk SSSR* **88,** 601–602 (1953).
4. D. F. Davidenko, "On the Approximate Solution of Systems of Nonlinear Equations" (in Russian), *Ukr. Mat. Z.* **5,** 196–206 (1953).
5. W. Kizner, "A Numerical Method for Finding Solutions of Nonlinear Equations," *J. SIAM* **12,** 424–428 (1964).
6. J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970, Sec. 7.5.
7. F. H. Branin and S. K. Hoo, "A Method for Finding the Global Maximum and Global Minimum of a Function of *n* Variables," *Conference on Numerical Methods for Nonlinear Optimization*, Dundee, Scotland, 1971.

Received October 8, 1971

The author is located at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.