Ann. Inst. Statist. Math.
37 (1985), Part B, 189-203

GLOBAL ANALYSIS OF CONTINUOUS ANALOGUES OF THE LEVENBERG-MARQUARDT AND NEWTON-RAPHSON METHODS FOR SOLVING NONLINEAR EQUATIONS

KUNIO TANABE

(Received Mar. 31, 1984)

Summary

Global analyses are given to continuous analogues of the Levenberg-Marquardt method $dx/dt=-(J^t(x)J(x)+\delta I)^{-1}J^t(x)g(x)$, and the Newton-Raphson-Ben-Israel method $dx/dt=-J^+(x)g(x)$, for solving an over- and under-determined system $g(x)=0$ of nonlinear equations. The characteristics of both methods are compared. Errors in some literature which dealt with related continuous analogue methods are pointed out.

1. Introduction

The problems of finding a solution of nonlinear equations and a maximum of a nonlinearly constrained function arise in diverse fields of mathematical sciences and numerous methods have been proposed for solving the problems in literature. Recently much attention has been paid to continuous analogues of discrete iterative methods, which are based on solving the related systems of ordinary differential equations [3]-[8], [10]-[27]. As was well demonstrated in Branin [7], Meyer [15] and Tanabe [23], [24], continuous analogues not only have larger region of convergence than the original discrete methods but also their analysis provides us qualitative information on the behavior of the original methods. Further, an important advantage of the differential equation approach is that it facilitates global analysis of the behavior of the methods, which does not seem to be widely recognized. There have been several papers on global analysis of continuous analogue methods [7], [8], [12], [15], [17], [18]-[27]. The analyses given in Boggs [3], [4], Boggs and Dennis [5] are essentially local ones. The results given in Meyer [15] are not completely global because he considered

Key words and phrases: Stability of dynamical systems, continuous Levenberg-Marquardt method, continuous Newton-Raphson method, underdetermined simultaneous nonlinear equations, nonlinear least squares, generalized inverse.

only the cases where solutions of related differential equations exist for the infinite interval. Branin [7] showed, by numerical experiments, the global behavior of the continuous analogue of Newton-Raphson method. In spite of lack of mathematical rigor, his paper is thought-provoking and an excellent source of information on the method. A theoretical analysis of the method in the case where a system of nonlinear equations which is possibly underdetermined was given in Tanabe [22]-[25]. An analysis was also given to the continuous analogue of the gradient projection method for solving nonlinear constrained optimization problems in [18]-[20]. In this connection, an analysis was given to a continuous analogue of the gradient projection method with enforced feasibility in [22], [24], [25]. In the analysis of global behaviors of these methods, a differential geometric method played an important role.

In Section 2, mathematical errors which were made in some literature treating continuous analogue methods are pointed out. In Section 3, a global analysis is given to a continuous analogue of the Levenberg-Marquardt method in both cases where systems of nonlinear equations are overdetermined and underdetermined. In Section 4, a global analysis is given to a continuous analogue of the Newton-Raphson method in the case where a system is overdetermined. The characteristic of the continuous Levenberg-Marquardt method is shown by comparing both methods.

2. Common errors

The problem discussed in this paper is to solve a system of nonlinear equations,

$$g(x)=0 \in R^m , \tag{1}$$

where g is a twice continuously differentiable mapping from n-dimensional Euclidian space R^n to m-dimensional space R^m.

Gay [13] proposed the method,

$$dx/dt=-\hat{J}_g^+(x)g(x) , \tag{2}$$

for solving (1) in the case where $m=n$, where $\hat{J}_g^+(x)$ is the Levenberg-Marquardt type modification of the inverse of the Jacobian matrix $J_g(x)$ of g. He assumed the following condition to obtain his result,

$$g^t(x)(J_g(x)\hat{J}_g^+(x))g(x)\geqq\theta\|g(x)\|^2 \qquad \text{for all } x\in R^n , \tag{3}$$

which is rather a restrictive condition since $J_g(x)\hat{J}_g^+(x)$ may be a rank deficient matrix. He claimed in Theorem (23) of [13] that for each $x^0 \in R^n$, there exists a solution $x(t)$ of (2) with $x(t^0)=x^0$ for the interval,

(4) $$t^0 \leqq t < \infty ,$$

and that such a solution has an asymptote x^*

(5) $$x^* = \lim_{t\to\infty} x(t)$$

with $g(x^*)=0$ and

(6) $$\|x(t)-x^*\| \leqq (\|g(x^0)\|/(\theta\varepsilon)) \exp(-\theta t) .$$

Unfortunately this result is unfounded. His analysis was based on Theorem (5) of [13] which he erroneously cited from Coddington and Levinson [9] claiming that if $f: R^n \mapsto R^n$ is continuous, then for each $x^0 \in R^n$ and $t^0 \in R$ there exists a continuously differentiable function x: $R \mapsto R^n$ such that $x(t^0)=x^0$ and $dx(t)/dt = f(x(t))$ for

(7) $$-\infty < t < \infty .$$

This is incorrect and (7) should have read,

(8) $$t^0 - M^- < t < t^0 + M^+$$

where M^- and M^+ are positive numbers or infinity. A counter example against his claim in Theorem (5) of [13] is given by

(9) $$dx/dt = x^2 \in R .$$

If the initial value x^0 is not zero, the solution of (9) blows up in finite time. For example, it blows up at $t=1$ when $x^0=1$. The present author includes this simple example, since in many literature [3]–[7], [11], [13], [27] which dealt with continuous analogue methods, the authors assumed a priori and sometimes erroneously that their solutions exist for the intervals (4) or (7). Although the errors are subtle and most of them could be rectified by providing suitable additional assumptions, which lose the generality of the propositions however, these results are incorrect or misleading since the assumptions (4) and (7) almost imply that the solution $x(t)$ of the autonomous system converges as t tends to infinity. In particular, the proof of Theorem (23) of [13] is invalid because it depends on the erroneous assumption (4). An alternative analysis of a related method is given in the next section.

The following theorem, which should replace Theorem (5) of [13], will frequently be used in the analysis given in Sections 3 and 4.

THEOREM 0. *If a mapping $f(t, x): R^{n+1} \mapsto R^n$ is continuous and satisfies the local Lipschitz condition on an open set $D \subset R^{n+1}$, there exists a unique solution $x(t, x^0)$, $0 \leqq t < M$, of*

(10) $$dx/dt = f(t, x) , \qquad t \in R \text{ and } x \in R^n ,$$

with $x(0, x^0)=x^0$, $(0, x^0)\in D$ *and the integration curve of the solution can be prolonged until it reaches the boundary* ∂D *of* D.

In this paper we consider only solutions with maximum interval of existence in the positive direction for given initial values.

3. Continuous analogue of Levenberg-Marquardt method

We consider a continuous analogue of the Levenberg-Marquardt method,

$$dx/dt=-(J_g^t(x)J_g(x)+\delta I_n)^{-1}J_g^t(x)g(x) \tag{11-1}$$

$$=-J_g^t(x)(J_g(x)J_g^t(x)+\delta I_m)^{-1}g(x)=\Phi(x) \tag{11-2}$$

where $\delta>0$, I_m and I_n are identity matrices of respective dimension m and n. Note that the expressions (11-1 and 2) are equivalent.

THEOREM 3.1. *For each* $x^0\in R^n$, *there exists a solution* $x(t, x^0)$, $0\leq t<M$, *with* $x(0, x^0)=x^0$, *of* (11). *As* t *tends to* M, *its trajectory will either* (i) *converge to a solution* $x^*\in V_g\cap S_1^c$, *in which case* $M=\infty$, *or* (ii) *approach the set* $E-(V_g\cap S_1^c)\subset S_2\cup V_g$, *of certain equilibrium points of the system* (11), *or* (iii) *diverge, where* V_g, S_1, S_2 *and* E *are possibly empty sets defined respectively by*

$$V_g=\{x\in R^n:\ g(x)=0\}\ , \tag{12}$$

$$S_1=\{x\in R^n:\ \operatorname{rank} J_g(x)<n\}\ , \tag{13}$$

$$S_2=\{x\in R^n:\ \operatorname{rank} J_g(x)<m\}\ , \tag{14}$$

$$E=\{x\in R^n:\ J_g^t(x)g(x)=0\}\subset V_g\cup S_2\ ,$$

and S^c *is the complement of* S *in* R^n.

Note that if V_g is empty then the set E contains the possibly empty set L_g of all the least squares solutions of the system (1). Before proving the theorem we prepare a few lemmas.

LEMMA 3.2. *If* $\operatorname{rank} J_g(x^*)=n$ *for a solution* $x^*\in V_g$, *then* x^* *is an asymptotically stable point of the system* (11).

PROOF. By the continuity of $J_g(x)$, there exists a neighbourhood U of x^* such that x^* is the only solution of (1) in U and

$$\operatorname{rank} J_g(x)=n \qquad \text{for } x\in U\,.$$

Let us consider the function,

$$V(x)=\|x-x^*\|^2/2$$

defined in U. Then

$$\begin{aligned} dV/dt &= \langle dx/dt,\ x-x^* \rangle \\ &= \langle \Phi(x),\ x-x^* \rangle \\ &= -\langle (J_g(x)J_g^t(x)+\delta I_m)^{-1}(g(x)-g(x^*)),\ J_g(x)(x-x^*) \rangle . \end{aligned}$$

Since $g(x)-g(x^*)=J_g(x)(x-x^*)+o(\|x-x^*\|)$, we have

$$\begin{aligned} dV/dt = &-\langle (J_g(x)J_g^t(x)+\delta I_m)^{-1}J_g(x)(x-x^*),\ J_g(x)(x-x^*) \rangle \\ &+o(\|x-x^*\|^2) . \end{aligned}$$

Hence there exists an open ball,

$$B=\{x \in R^n:\ \|x-x^*\|<\varepsilon\} \subset U ,$$

such that $dV/dt<0$ for $x \in B-\{x^*\}$. This implies that the function V is a strict Liapunov function for x^*, hence we have the desired result.

LEMMA 3.3. *If a solution* $x(t, x^0)$, $0 \leqq t < M$, *of* (11) *is bounded then* $M=\infty$ *and the positive limit set* Γ *of the solution is contained in the set* E *of equilibrium points. Note that* $E \supset L_g \supset V_g$.

PROOF. By Theorem 0 it is obvious that $M=\infty$. Let us consider the function,

$$L(x)=\|g(x)\|^2/2=\left(\sum_{i=1}^{m} g_i^2(x)\right)\Big/2 .$$

Then

$$dL/dt=-\langle J_g^t(x)g(x),\ (J_g^t(x)J_g(x)+\delta I_n)^{-1}J_g^t(x)g(x)\rangle \leqq 0 \quad \text{for } x=x(t, x^0),\ 0 \leqq t < \infty \tag{15}$$

and the equality holds if and only if $x^0 \in E$, because if the equality holds at $t=T>0$, then $x(T, x^0) \in E$ and

$$x(t) \equiv x(T, x^0)$$

is obviously a solution of (11), hence by the uniqueness of the solution we have

$$x(t)=x(t, x^0)=x^0 \in E .$$

Thus we have either that $x(t, x^0) \equiv x^0 \in E$ or that $L(x(t, x^0))$ is a strictly monotone decreasing function of t. Therefore there is no periodic solution of (11). In both cases, there exists the limit

$$\lim_{t \to \infty} L(x(t, x^0))=k \geqq 0 . \tag{16}$$

Hence by the continuity of $L(x)$ we have

$$L(\Gamma)=k\ . \tag{17}$$

Suppose Γ is not contained in E, then there exists a point $x'\in\Gamma\cap E^c$ and there exists a solution $x(t, x')$, $0\leqq t<M'$, with $x(0, x')=x'$. Because the positive limit set Γ is an invariant set of the system (11), the trajectory of $x(t, x')$ is contained in Γ. However, $L(x(t, x'))$ is a strictly monotone decreasing function of $t\in[0, M')$, since $x'\notin E$. This contradicts the fact that $L(x)$ is constant k on Γ. Thus we have $\Gamma\subset E$.

Combining these two lemmas and Theorem 0, we obtain Theorem 3.1.

COROLLARY 3.4. *If the connected component $C(x^0)$ of the level set,*

$$\{x\in R^n:\ \|g(x)\|\leqq\|g(x^0)\|\}\ ,$$

that contains x^0, is bounded, the conclusion of Lemma 3.3 is valid. Further if $E\cap C(x^0)$ consists of isolated points, the solution converges to a point in E as t tends to infinity. Hence in the case where $E=L_g$, it converges to a least squares solution of (1).

THEOREM 3.5. *If a solution $x(t, x^0)$, $0\leqq t<\infty$, of* (11) *with $x(0, x^0)=x^0$, converges, as t tends to infinity, to a solution $x^*\in V_g$ such that* rank $J_g(x^*)=m$, *then there exist positive numbers k and θ such that*

$$\|x(t, x^0)-x^*\|\leqq k\|g(x^0)\|\exp(-\theta t)\ ,\qquad 0\leqq t<\infty\ , \tag{18}$$

PROOF. There exists a neighbourhood U of x^* such that the topological closure $\bar{U}$ of U is compact and rank $J_g(x)=m$ for all $x\in\bar{U}$. Considering the function $L(x)$ defined in the proof of Lemma 3.3, we have

$$dL/dt=-\langle g(x),\ J_g(x)\hat{J}_g^+(x)g(x)\rangle\ ,$$

where $\hat{J}_g^+(x)=(J_g^t(x)J_g(x)+\delta I_n)^{-1}J_g^t(x)$. Since $J_g(x)\hat{J}_g^+(x)$ is non-singular and continuous on the compact set $\bar{U}$, there exists a positive number θ_1 such that

$$\langle g(x),\ J_g(x)\hat{J}_g^+(x)g(x)\rangle\geqq\theta_1\|g(x)\|^2\qquad \text{for } x\in\bar{U}\ . \tag{19}$$

On the other hand, there exists T such that $x(t, x^0)\in U$ for $t\geqq T$. Let us consider the function defined on V_g^c,

$$q(x)=\langle g(x),\ J_g(x)\hat{J}_g^+(x)g(x)\rangle/\|g(x)\|^2\geqq 0\ ,$$

which does not vanish on the trajectory of a solution with an initial point in E^c, by the proof of Lemma 3.3. Since $q(x(t, x^0))$ is a continuous function of t on the closed interval $[0, T]$, there exist the minimum

θ_2 of $q(x(t, x^0))$ on the interval. Let

$$\theta=\min(\theta_1, \theta_2) ,$$

and let the topological closure of the trajectory of the solution be denoted by $\Lambda(x^0)$, i.e.,

$$\Lambda(x^0)=(\bigcup_{0\leq t<\infty}\{x(t, x^0)\})\cup\{x^*\} .$$

Then we have

$$\langle g(x), J_g(x)\hat{J}_g^+(x)g(x)\rangle\geqq\theta\|g(x)\|^2 \qquad \text{for } x\in\Lambda(x^0) .$$

Thus we have

$$dL(x(t, x^0))/dt\leqq-2\theta L(x(t, x^0)) , \qquad 0\leqq t<\infty .$$

Therefore

$$\|g(x(t, x^0))\|^2\leqq\|g(x^0)\|^2\exp(-2\theta t) , \qquad 0\leqq t<\infty .$$

Since $\Lambda(x^0)$ is a compact set and $\hat{J}_g^+(x)$ is continuous on $\Lambda(x^0)$, there exists a positive number k such that

$$\|\hat{J}_g^+(x)\|\leqq k<\infty .$$

Thus we have

$$\begin{aligned}\|x(t, x^0)-x^*\|&=\left\|\int_t^\infty(dx/dt)dt\right\|\\&\leqq\int_t^\infty\|dx/dt\|\,dt\\&=\int_t^\infty\|\hat{J}_g^+(x)g(x)\|\,dt\\&\leqq k\int_t^\infty\|g(x(t, x^0))\|\,dt\\&\leqq k\|g(x^0)\|\int_t^\infty\exp(-\theta t)\\&=k\|g(x^0)\|\exp(-\theta t) ,\end{aligned}$$

which completes the proof. Finally note that if $x^0\notin E$ then $\|g(x(t, x^0))\|^2$ is a strictly monotone decreasing function of $t\in[0, M)$ in all cases.

4. Continuous analogue of Newton-Raphson method

In this section we consider a continuous analogue of the Newton-Raphson-Ben Israel method,

$$dx/dt=-J_g^+(x)g(x)=\Psi(x) , \tag{20}$$

where $J_g^+(x)$ is the Moore-Penrose inverse of $J_g(x)$, to clarify the characteristics that distinguish the Levenberg-Marquardt modification (11) from the Newton-Raphson method. The autonomous system (20) was analyzed in Tanabe [22]-[24] in the case where $m \leqq n$ and the following local and global theorems were obtained without assuming such a restrictive assumption as (3). For proofs and a more detailed treatment see [22]-[24].

THEOREM 4.1. *Let us assume that $g(x)$ is a twice continuously differentiable mapping from R^n to R^m, where $m \leqq n$. If* $\operatorname{rank} J_g(x^*)=m$ *for a solution $x^* \in V_g$ then there exists a neighbourhood U^* of x^* such that for each $x^0 \in U^*$ there exists a solution $x(t, x^0)$, $0 \leqq t < \infty$, of* (20) *with $x(0, x^0)=x^0$, and as t tends to infinity it always converges to a point in V_g which may be different from x^* in the case where $m<n$.*

THEOREM 4.2. *Under the same assumption as the previous theorem, for a given $x^0 \in S_2^c$, there exists a solution $x(t, x^0)$, $0 \leqq t < M$, of* (20) *with $x(0, x^0)=x^0$. As t tends to M, its trajectory will either* (i) *converge to a solution $x^* \in V_g \cap S_2^c$, in which case we have $M=\infty$ and*

$$\|x(t, x^0)-x^*\| \leqq k\|g(x^0)\| \exp(-t)\,, \qquad 0 \leqq t < \infty\,, \tag{21}$$

for some positve number k, or (ii) *approach the set,*

$$S_2=\{x \in R^n:\ \operatorname{rank} J_g(x)<m\}\,,$$

of singular points of (20), *or* (iii) *diverge.*

Note that the Liapunov methods are of no use when $m<n$ because then the rank of the Jacobian matrix (the first derivative) of the right-hand-side, $\Psi(x)$ of (20) is always less than n at a solution of (1). The proofs of the theorems depended heavily on a differential geometric treatment [22]-[24]. Note also that the inequality (21) is essentially the same as (6) but it is obtained without assuming such a condition as (3).

The proof of the inequality (21) depends also on a distinctive feature of the continuous analogue of the Newton-Raphson method, that a solution $x(t, x^0)$, $0 \leqq t < M$, of (20) with $x(0, x^0)=x^0$ satisfies the 'first integral' of (20),

$$g(x(t, x^0))=\exp(-t)g(x^0) \qquad \text{for } 0 \leqq t < M\,, \tag{23}$$

which does not hold in the case of the continuous analogue of the Levenberg-Marquardt method (11). Let the topological closure of the trajectory of the solution be denoted by $\Lambda(x^0)$, i.e.,

$$\Lambda(x^0)=\Big(\bigcup_{0 \leqq t < \infty}\{x(t, x^0)\}\Big) \cup \{x^*\}$$

where

$$x^* = \lim_{t\to\infty} x(t, x^0) .$$

Since $\Lambda(x^0)$ is a compact set and $J_g^+(x)$ is continuous on $\Lambda(x^0)$, there exists a positive number k such that

$$\|J_g^+(x)\| \leqq k \qquad \text{for } x \in \Lambda(x^0) .$$

Thus we have

$$\begin{aligned} \|x(t, x^0) - x^*\| &= \left\| \int_t^\infty (dx/dt)dt \right\| \\ &\leqq \int_t^\infty \|dx/dt\| \, dt \\ &= \int_t^\infty \|J_g^+(x)g(x)\| \, dt \\ &\leqq k \int_t^\infty \|g(x(t, x^0))\| \, dt \\ &= k \|g(x^0)\| \int_t^\infty \exp(-t) dt \\ &= k \|g(x^0)\| \exp(-t) . \end{aligned}$$

Incidentally, if V_g is empty in Theorem 4.2, the set S_2 of singular points contains the set L_g of all the least squares solutions of the system (1) if they exist.

Note also that for any solution $x(t, x^0)$, $0 \leqq t < M$, of (20) with $x(0, x^0) = x^0$, we have, by the equality (23),

$$\text{(24)} \qquad \operatorname{Sign}(g_i(x(t, x^0))) = \operatorname{Sign}(g_i(x^0)) \qquad \text{for } 0 \leqq t < M ,$$

where $\operatorname{Sign}(x) = 1$(if $x > 0$), $= 0$(if $x = 0$), $= -1$(if $x < 0$), and

$$g(x) \equiv (g_1(x), g_2(x), \cdots, g_m(x))^t .$$

This implies that the trajectory of a solution of (20) never crosses any of the hyperplanes V_i defined by

$$V_i = \{x \in R^n : \; g_i(x) = 0\} .$$

COROLLARY 4.3. *Under the same assumption as the previous theorem, if the connected component $C^*(x^0)$ of the level set,*

$$\{x \in R^n : \; 0 \leqq \operatorname{Sign}(g_i(x^0)) g_i(x) \leqq \operatorname{Sign}(g_i(x^0)) g_i(x^0) \text{ for all } i\}$$

that contains x^0, is bounded and does not contain a point in S_2, then $M = \infty$ and the trajectory of the solution $x(t, x^0)$, $0 \leqq t < \infty$, of (20) with $x(0, x^0) = x^0$, converges as t tends to infinity, to a solution $x^* \in V_g \cap S_2^c$ of (1) and the inequality (21) holds.

In the case where $m>n$ however, solutions of the autonomous system (20) behave quite differently in general. The following theorem describes its behavior in this case.

THEOREM 4.4. *For each* $x^0 \in S_1^c$, *there exists a solution* $x(t, x^0)$, $0 \leqq t < M$, *with* $x(0, x^0) = x^0$. *As* t *tends to* M, *its trajectory will either* (i) *converge to a solution* $x^* \in V_g \cap S_1^c$, *in which case* $M=\infty$, *or* (ii) *approach the set* E, *or* (iii) *approach the set* S_1, *or* (iv) *diverge, where* V_g, S_1 *and* E *are defined in Theorem* 3.1. *In all cases* $\|g(x(t, x^0))\|^2$ *is a strictly monotone decreasing function of* $t \in [0, M)$.

Before proving the theorem we prepare a few lemmas.

LEMMA 4.5. *If* $\operatorname{rank} J_g(x^*) = n$ *for a solution* $x^* \in V_g$, *then* x^* *is an asymptotically stable point of the system* (20).

PROOF. The Jacobian matrix $J_\Psi(x)$ of the right-hand-side $\Psi(x)$ of (20) is given by

$$J_\Psi(x) = -[\partial J_g^+(x)/\partial x]g(x) - J_g^+(x)J_g(x) ,$$

where $[\partial J_g^+(x)/\partial x]g(x) \equiv [(\partial J_g^+(x)/\partial x_1)g(x) : \cdots : (\partial J_g^+(x)/\partial x_n)g(x)]$. Hence

$$J_\Psi(x^*) = -J_g^+(x^*)J_g(x^*) = -I_n .$$

Thus by Liapunov's first method, we have the desired result.

LEMMA 4.6. *If a solution* $x(t, x^0)$, $0 \leqq t < M$, *of* (20) *with* $x(0, x^0) = x^0$, *is bounded, then as* t *tends to* M, *its trajectory will either* (i) *approach the set* S_1, *or* (ii) *approach the set* E *in which case* $M=\infty$ *and its positive limit set* Γ *is contained in the set* E. *Note that* $E \supset L_g \supset V_g$.

PROOF. If its trajectory does not approach the set S_1, then by Theorem 0, we have $M=\infty$. Let us consider the function

$$L(x) = \|g(x)\|^2/2$$

then

$$dL/dt = -\langle g(x), J_g(x)J_g^+(x)g(x)\rangle = -\|J_g(x)J_g^+(x)g(x)\|^2 \leqq 0 . \tag{25}$$

Hence we have

$$dL(x(t, x^0))/dt \leqq 0 , \qquad \text{for } 0 \leqq t < \infty ,$$

and the equality holds if and only if $x^0 \in E$, because if the equality holds at $t=T$, then at the point $x' = x(T, x^0)$ we have

$$J_g(x')J_g^+(x')g(x') = 0 ,$$

which implies

$$J_g^+(x')g(x')=0 \qquad \text{and} \qquad J_g^t(x')g(x')=0 .$$

Hence

$$x(t)\equiv x(T, x^0)\in E$$

is obviously a solution of (20). The rest of the proof is exactly the same as the latter part of the proof of Lemma 3.3, so it is omitted here.

Combining these two lemmas and Theorem 0, we obtain Theorem 4.4.

COROLLARY 4.7. *If the connected component $C(x^0)$ of the level set,*

$$\{x\in R^n:\ \|g(x)\|\leqq\|g(x^0)\|\} ,$$

that contains x^0, is bounded then the conclusion of Lemma 4.6 is valid. Further, if $S_1\cap C(x_0)=\phi$ and $E\cap C(x^0)$ consists of isolated points, then the solution converges to a point in E as t tend to infinity. Hence in the case where $E=L_g$, it converges to a least squares solution of the system (1).

In general, the equality (23) does not hold in the case where $m>n$. However, if the very restrictive condition,

$$\text{(26)} \qquad \operatorname{Im} J_g(x)\ni g(x) , \qquad \text{for any } x\in D ,$$

is satisfied, then (23) hold for a solution of (20) which starts from a point in D, where $\operatorname{Im} J_g(x)$ is the range space of $J_g(x)$ and D is a domain in R^n. Note that the condition (26) is similar to (3).

Now we compare the two methods (11) and (20). We consider the case where $m\leqq n$. The trajectories of the solutions of the systems (11) and (20) approach E and $S_2\cup V_g$ respectively if they don't diverge. We have

$$E=(V_g\cap S_2^c)\cup(E-(V_g\cap S_2^c))\supset L_g\supset V_g .$$

Since

$$E-(V_g\cap S_2^c)\subset S_2 ,$$

the method (11) seems to have a better chance to converge to a solution $x^*\in V_g$ than the method (20). In conclusion, the Levenberg-Marquardt method (11) resolves the singularity of the Newton-Raphson method at the cost of the equality (23). To see this in more detail, we consider the following example,

$$g(x)=\begin{pmatrix}g_1\\g_2\end{pmatrix}=\begin{pmatrix}x_1^2-x_2\\x_2^2-x_1\end{pmatrix}=0\in R^2\,, \tag{27}$$

which has the solution set,

$$V_g=\{(0,0),(1,1)\}\ .$$

The continuous analogue of the Newton-Raphson method for (27) is given by

$$\frac{d}{dt}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}2x_1^2x_2-x_2^2-x_1\\2x_2^2x_1-x_1^2-x_2\end{pmatrix}\Big/(1-4x_1x_2)\,, \tag{28}$$

which has the singular set,

$$S_2=\{(x_1,x_2):\ 1-4x_1x_2=0\}\ .$$

Trajectories of the solutions are shown in Fig. 1, which was given in [22], [24]. The domains of attraction of the solutions (0, 0) and (1, 1) are respectively

$$\{(x_1,x_2):\ 1-4x_1x_2>0 \text{ and } x_1+x_2+1>0\}\ ,$$

and

$$\{(x_1,x_2):\ 1-4x_1x_2<0 \text{ and } x_1>0\}\ .$$

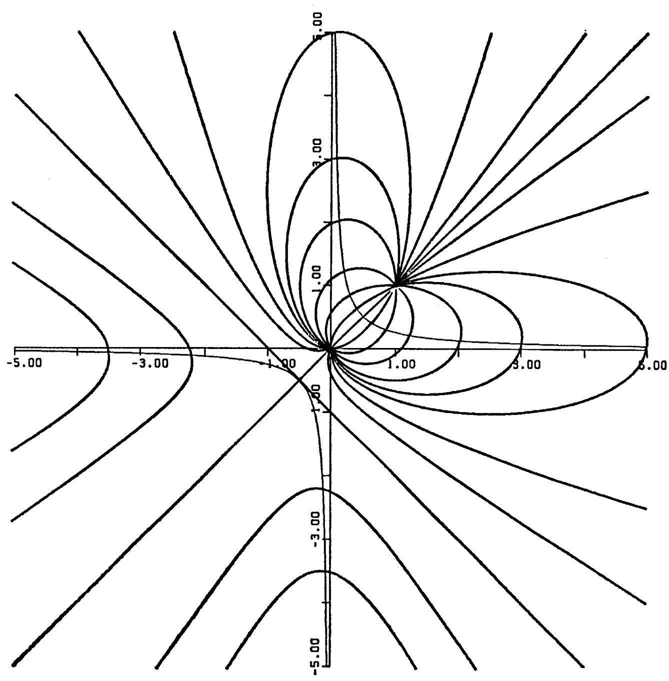

Fig. 1. Trajectories of the continuous Newton-Raphson method.

However, any trajectory emanating from a point in the region,

$$\{(x_1, x_2):\ x_1+x_2+1\leqq 0\}$$

converges to a singular point which belongs to the branch,

$$\{(x_1, x_2):\ 1-4x_1x_2=0, \text{ and } x_1<0\}\ ,$$

of the set S_2 in finite time. Note also that by (23) any solution of (20) is contained in the set $C^*(x^0)$ defined in Corollary 4.3 and that $C^*(x^0)$ for this example is always a bounded set for a given x^0.

The continuous analogue of the Levenberg-Marquardt method for (27) is given by,

$$\text{(29)}\qquad \frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -\begin{pmatrix} 4x_1^2+1+\delta & -2(x_1+x_2) \\ -2(x_1+x_2) & 4x_2^2+1+\delta \end{pmatrix}^{-1}\begin{pmatrix} 2x_1 & -1 \\ -1 & 2x_2 \end{pmatrix}\begin{pmatrix} x_1^2-x_2 \\ x_2^2-x_1 \end{pmatrix}$$

which has the set E of equilibrium points,

$$E=\{(0, 0), (1, 1), (1/2, 1/2)\}\ .$$

Corollary 3.4 applies in this example. Note that $E-V_g=\{(1/2, 1/2)\}\subset S_2$ is smaller than the singular set S_2. Further, the equilibrium point (1/2, 1/2) is an unstable point of the system (29). Hence there is little danger of a numerical solution of (29) converging to a point which is

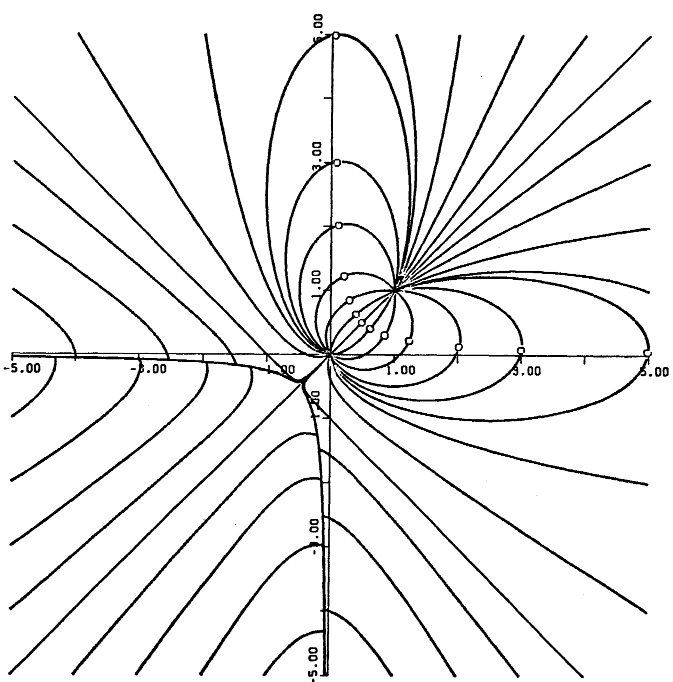

Fig. 2.1. Trajectories of the continuous Levenberg-Marquardt method ($\delta=0.01$).

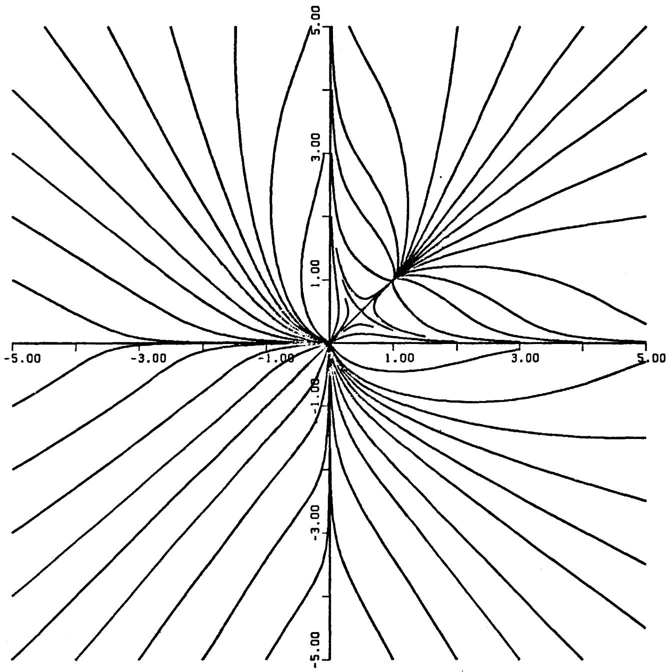

Fig. 2.2. Trajectories of the continuous Levenberg-Marquardt method ($\delta=1.0$).

not the solution of the system (27). See also Figs. 2.1 and 2.2 which show trajectories of the solutions of (29). Thus, the Levenberg-Marquardt method (29) is preferable to the Newton-Raphson method (28) as far as convergence region is concerned. Note also that when δ is smaller the trajectories of (29) are *locally* more similar to those of (28), which is expected because the system (28) is the limit of (29), in the sense,

$$J_g^+(x)=\lim_{\delta\to+0}(J_g^t(x)J_g(x)+\delta I_n)^{-1}J_g^t(x) .$$

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Ben Israel, A. (1966). A Newton-Raphson method for the solution of systems of equations, *J. Math. Anal. Appl.*, **15**, 243-252.

[2] Bhatia, N. P. and Szegö, G. P. (1967). *Dynamical Systems: Stability Theory and Applications*, Springer Verlag, New York.

[3] Boggs, P. T. (1971). The solution of nonlinear systems of equations by A-stable integration techniques, *SIAM J. Numer. Anal.*, **8**, 767-785.

[4] Boggs, P. T. (1976). The convergence of the Ben-Israel iteration for nonlinear least squares problems, *Math. Comp.*, **30**, 512-522.

[5] Boggs, P. T. and Dennis, J. E., Jr. (1976). A stability analysis for perturbed nonlinear iterative methods, *Math. Comp.*, **30**, 199-215.

[6] Boggs, P. T. (1977). An algorithm based on singular perturbation theory for ill-conditioned minimization problem, *SIAM Numer. Anal.*, **14**, 830-843.

[7] Branin, F. H., Jr. (1972). Widely convergent method for finding multiple solutions of simultaneous nonlinear equations, *IBM J. Res. Develop.*, Sep., 504-522.

[8] Brent, R. P. (1972). On the Davidenko-Branin method for solving simultaneous nonlinear equations, *IBM J. Res. Develop.*, July, 434-436.

[9] Coddington, E. A. and Levinson, J. E. (1955). *Theory of Ordinary Differential Equations*, McGraw-Hill, New York.

[10] Dekker, L. (1973). Hybrid computation in mathematical programming, *Mathematical Programming*, **4**, 21-43.

[11] Dennis, J. E., Jr. (1977). Nonlinear least squares and equations, (ed. D. A. H. Jacobs) *The State of Art in Numerical Analysis*, Academic Press, London, 269-312.

[12] Evtushenko, Y. G. and Zhadan, V. G. (1975). Application of the method of Liapunov functions to the study of the convergence of numerical methods, *Zh. vy chisl. Mat. mat. Fiz.*, **15**, 101-112.

[13] Gay, D. (1977). Modifying singular values: Existence of solutions to systems of nonlinear equations having a possibly singular Jacobian matrix, *Math. Comp.*, **31**, 962-973.

[14] Kizner, W. (1964). A numerical method for finding solutions of nonlinear equations, *SIAM J. Appl. Math.*, **12**, 424-428.

[15] Meyer, G. H. (1968). On solving nonlinear equations with a one-parameter operator imbedding, *SIAM J. Numer. Anal.*, **5**, 739-752.

[16] Ramsay, J. O. (1970). A family of gradient methods for optimization, *The Computer J.*, **13**, 413-417.

[17] Shinohara, Y. (1969). A geometric method of numerical solution of nonlinear equations and error estimation by Urabe's proposition, *Publ. RIMS Kyoto Univ.*, **5**, 1-9.

[18] Tanabe, K. (1969). Algorithm for the constrained maximization technique for nonlinear programming, *Proc. 2nd Hawaii International Conf. on System Science*, 487-490.

[19] Tanabe, K. (1969). An algorithm for the constrained maximization in nonlinear programming, *Research Memorandum*, no. **31**, Institute of Statistical Mathematics.

[20] Tanabe, K. (1974). An algorithm for the constrained maximization in nonlinear programming, *J. Operations Research Society of Japan*, **17**, 184-201.

[21] Tanabe, K. (1976). Analogue-type nonlinear programming (in Japanese), *Mathematical Science*, **157**, 45-51. (Presented at the Japan-France Seminar on Functional Analysis and Numerical Analysis, Tokyo, 1976.)

[22] Tanabe, K. (1977). A geometric method in nonlinear programming, *Stanford University Computer Science Report*, STAN-CS-77-643. (Presented at the Gatlinburg VII Conference, Asilomar, 1977.)

[23] Tanabe, K. (1979). Continuous Newton-Raphson method for solving an underdetermined system of nonlinear equations, *Nonlinear Analysis*, **3**, 495-503.

[24] Tanabe, K. (1980). A geometric method in nonlinear programming, *J. Optimization Theory and Applications*, **30**, 181-210.

[25] Tanabe, K. (1982). Differential geometric approach to extended GRG method with enforced feasibility in nonlinear programming; Global analysis, *Recent Applications of Generalized Inverses* (ed. S. L. Campbell), Pitman, Boston.

[26] Tanabe, K. (1978). Global analysis of continuous analogues of the Levenberg-Marquardt and Newton-Raphson methods for solving nonlinear equations, *Brookhaven National Laboratory Report*, BNL-24022 AMD-789.

[27] Yamashita, H. (1980). A differential equation approach to nonlinear programming, *Mathematical Programming*, **18**, 155-168.