SOME GLOBAL CONVERGENCE PROPERTIES OF THE LEVENBERG-MARQUARDT METHODS WITH LINE SEARCH[†]

SHOU-QIANG DU

ABSTRACT. In this paper, we consider two kinds of the Levenberg-Marquardt method for solve a system of nonlinear equations. We use line search on every iteration to guarantee that the Levenberg-Marquardt methods are globally convergent. Under mild conditions, we prove that while the descent condition can be satisfied at the iteration of the Levenberg-Marquardt method, the global convergence of the method can be established.

AMS Mathematics Subject Classification : 65K05, 90C30.
Key words and phrases : Global convergence, Levenberg-Marquardt method, line search, nonlinear equations.

1. Introduction

In this paper, the Levenberg-Marquardt method is designed to solve the following nonlinear equations

$$F(x) = 0, \tag{1.1}$$

where $F : R^n \to R^m$ is continuously differentiable and the Jacobian $J(x_k)$ is Lipschitz continuous on some neighbor of $x^* \in X^*$(We assume that the solution set of (1.1) is nonempty and denoted X^*). Nonlinear equations arise in various applications [1-4], such as some monotone variational inequality problems can be converted into (1.1) and et al. The iterative methods for solving (1.1) have been widely studied by many authors (see [1-7]). Due to its simplicity and global convergence, the Levenberg-Marquardt method has played a special role for solving the above nonlinear equations (see [5-8]). The classical Levenberg-Marquardt method for (1.1) computes d_k by

$$d_k = -(J(x_k)^T J(x_k) + \mu_k I)^{-1} J(x_k)^T F(x_k), \tag{1.2}$$

where $\mu_k \geq 0$ is a parameter being updated from iteration to iteration, such as $\mu_k = \|F(x_k)\|^\delta, \delta \in [1,2]$ and $\mu_k = \|F(x_k)\|^2$, in [7,8].

In this paper, we only consider the globalization of the Levenberg-Marquardt method. Similar to unconstrained optimization problems, line search conditions are used to guarantee the global convergence. A line search algorithms computes a search direction at the k-th iteration and computes the next iterate by $x_{k+1} = x_k + \alpha_k d_k$, where α_k is computed by different line search techniques, such as Wolfe line search, Armijo line search and some Wolfe-type line search, Armijo-type line search, et al (see [9-12]). Considering that the line search plays an important role in the proof of the global convergence of the Levenberg-Marquardt method, we give an other point of view to guarantee the global convergence of the Levenberg-Marquardt method.

The paper is organized as follows. In Section 2, we give some kinds of line search conditions for ensuring the global convergence of the Levenberg-Marquardt method. In Section 3, the Levenberg-Marquardt methods and their convergence results are given. Finally, we give some final remarks.

Notation. In the following paper, a quantity with a subscript k denotes that quantity evaluated at x_k, the vector norm is the l_2 norm.

2. Preliminaries

In this section, we give two kinds of inexact line search(see [9-10]), which can be used in the following Levenberg-Marquardt methods. The inexact line search conditions are based on the reduction of the following merit function

$$\Psi(x) = \frac{1}{2}\|F(x)\|^2.$$

Line search I
The inexact Wolfe type line search requires α_k to satisfy

$$\|F(x_k)\|^2 - \|F(x_k + \alpha_k d_k)\|^2 \geq \rho \alpha_k^2 \|d_k\|^2, \tag{2.1}$$

$$F(x_k + \alpha_k d_k)^T J(x_k + \alpha_k d_k) d_k \geq -2\sigma\alpha_k\|d_k\|^2, \tag{2.2}$$

where $0 < \rho < \sigma < 1$.

Line search II
The inexact Wolfe type line search requires α_k to satisfy

$$\|F(x_k + \alpha_k d_k)\|^2 - \|F(x_k)\|^2 \leq \max\{\delta\alpha_k F(x_k)^T J_k d_k, -\rho\alpha_k^2\|d_k\|^2\}, \tag{2.3}$$

$$F(x_k + \alpha_k d_k)^T J(x_k + \alpha_k d_k) d_k \geq \max\{\delta F(x_k)^T J_k d_k, -2\sigma\alpha_k\|d_k\|^2\}, \tag{2.4}$$

where $0 < \delta < \sigma < 1, 0 < \rho < 1$.

Assumption 2.1 Assume that the Jacobian $J(x)$ is Lipschitz continuous on some neighborhood of $x^* \in X^*$, i.e.

$$\|J(y) - J(x)\| \leq L\|y - x\|,$$

where $x, y \in N(x^*)$, L is a positive constant.

Normally, the search direction d_k is chosen when it is a descent direction unless $\nabla\Psi(x_k) = 0$. So we can see $d_k^T\nabla\Psi < 0$, if $\nabla\Psi \neq 0$. Based on the above analysis, we get the following Lemmas.

Lemma 2.1. *Suppose that Assumption 2.1 holds and that $d_k^T\nabla\Psi < 0$, the formula (2.1) and (2.2) imply that*

$$\|F(x_{k+1})\|^2 \leq \|F(x_k)\|^2 - \rho(\frac{1}{2\sigma + L})^2 \frac{(F_k^T J_k d_k)^2}{\|d_k\|^2}. \tag{2.5}$$

Proof. By the conditions of the lemma, we get

$$(F_{k-1}^T J_{k-1} - F_k^T J_k)^T d_k \geq -F_k^T J_k d_k - 2\sigma\alpha_k\|d_k\|^2,$$
$$(F_{k-1}^T J_{k-1} - F_k^T J_k)^T d_k \leq L\alpha_k\|d_k\|^2,$$

so

$$(2\sigma + L)\alpha_k\|d_k\|^2 \geq -F_k^T J_k d_k,$$
$$(\frac{1}{2\sigma + L})^2 \frac{(F_k^T J_k d_k)^2}{\|d_k\|^2} \leq \alpha_k^2\|d_k\|^2,$$

and

$$\alpha_k^2\|d_k\|^2 \leq \frac{1}{\rho}(\|F(x_k)\|^2 - \|F(x_k + \alpha_k d_k)\|^2).$$

From the above inequations, we get (2.5). $\qquad\square$

Corollary 2.2. *Suppose that the assumptions of Lemma 2.1 holds, the formula (2.3) and (2.4) also imply that (2.5) holds.*

3. Levenberg-Marquardt methods and their convergence results

In this section, we describe two Levenberg-Marquardt methods for (1.1). The global convergence of the two methods are also given.

Levenberg-Marquardt Method(I).
Step 1. Given $x_0 \in R^n$, $\delta \in [1,2], \eta \in (0,1), k := 0$.
Step 2. If $\|\nabla\Psi\| = 0$ then stop. Otherwise set $\mu_k = \|F_k\|^\delta$, compute d_k by (1.2).
Step 3. If d_k satisfies $\|F(x_k + d_k)\| \leq \eta\|F(x_k)\|$, set $x_{k+1} = x_k + d_k$, otherwise compute x_{k+1} by $x_{k+1} = x_k + \alpha_k d_k$, where α_k is computed by Line Search I(see (2.1) and (2.2)). Let $k := k+1$, and go to Step 2.

Levenberg-Marquardt Method(II).
Step 1. Given $x_0 \in R^n$, $\delta \in [1,2], \eta \in (0,1), k := 0$.
Step 2. If $\|\nabla\Psi\| = 0$ then stop. Otherwise set $\mu_k = \|F_k\|^\delta$, compute d_k by (1.2).
Step 3. If d_k satisfies $\|F(x_k + d_k)\| \leq \eta\|F(x_k)\|$, set $x_{k+1} = x_k + d_k$, otherwise compute x_{k+1} by $x_{k+1} = x_k + \alpha_k d_k$, where α_k is computed by Line Search II(see (2.3) and (2.4)). Let $k := k+1$, and go to Step 2.

Based upon the above analysis, similar to [7], we also can get the following global convergence results.

Theorem 3.1. *Suppose that Assumption 2.1 holds and $\{x_k\}$ is a sequence generated by the above Levenberg-Marquardt Method(I). Then any accumulation point x^* of $\{x_k\}$ is a stationary points of Ψ.*

Proof. Since $\nabla\Psi(x_k) \neq 0$ implies $d_k \neq 0$, we have

$$\nabla\Psi(x_k)^T d_k = -((J(x_k)^T J(x_k) + \mu_k I)d_k)^T d_k < 0.$$

Then follows from $\|F(x_k + d_k)\| \leq \eta\|F(x_k)\|$ and (2.1), (2.2), we know that $\{\Psi(x_k)\}$ is monotonically decreasing. If $\|F(x_k)\| \to 0$, any accumulation point of $\{x_k\}$ is a solution of Ψ. Otherwise, if $\|F(x_k)\| \to \zeta > 0$, so $\|F_{k+1}\| \leq \eta\|F_k\|$ holds for finitely many times. We can see that (2.5) is satisfied only for all large k. By (1.2) and (2.5), we know

$$(F_k^T J_k d_k)^2 = (d_k^T (J_k^T J_k + \mu_k I)d_k)^2 \geq \zeta^{2\delta}\|d_k\|^4. \tag{3.1}$$

From (2.5), (3.1), we can see

$$\lim_{k\to\infty} \|d_k\| = 0.$$

The above limit and (1.2) imply that $J(x^*)F(x^*) = 0$. So we complete the proof. $\qquad\square$

By Corollary 2.2, we can easily get the following global convergence of Levenberg-Marquardt Method(II).

Theorem 3.2. *Suppose that Assumption 2.1 holds and $\{x_k\}$ is a sequence generated by the above Levenberg-Marquardt Method(II). Then any accumulation point x^* of $\{x_k\}$ is a stationary points of Ψ.*

Proof. By Assumption 2.1, (2.3) and (2.4), we can get (3.1). Since it is a straightforward modification of the proof of Theorem 3.1, we omit the proof. $\qquad\square$

4. Final remarks

In the above section, we give some global convergence properties of Levenberg-Marquardt method under inexact line search for (1.1). From the global convergence analysis of Levenberg-Marquardt Method(I) and Levenberg-Marquardt Method(II), we can see that (2.5) is used to prove the global convergence of Levenberg-Marquardt methods. So we can get the global convergence results of Levenberg-Marquardt method for any line search, which can get (2.5) under Assumption 2.1, Such as we can also get (2.5) by the following strong Wolfe line search conditions and generalized wolfe line search conditions and exact line search [12].

Line search III
The strong Wolfe line search requires α_k to satisfy

$$\|F(x_k + \alpha_k d_k)\|^2 \leq \|F(x_k)\|^2 + \beta_1 \alpha_k F_k^T J_k d_k, \tag{4.1}$$

$$|F(x_k + \alpha_k d_k)^T J(x_k + \alpha_k d_k)d_k| \leq -\beta_2 F_k^T J_k d_k, \tag{4.2}$$

where $0 < \beta_1 \leq \beta_2 < 1$.

Line search IV
The generalized Wolfe line search requires α_k to satisfy (4.1) and

$$\beta_3 F_k^T J_k d_k \leq F(x_k + \alpha_k d_k)^T J(x_k + \alpha_k d_k)d_k \leq -\beta_4 F_k^T J_k d_k, \qquad (4.3)$$

where $0 < \beta_1 \leq \beta_3 < 1$, $\beta_4 \geq 0$.

Line search V
The exact line search requires α_k to satisfy

$$\|F(x_k + \alpha_k d_k)\|^2 = \min_{\alpha \geq 0} \|F(x_k + \alpha d_k)\|^2. \qquad (4.4)$$

Remark 4.1. Suppose that Assumption 2.1 holds and that $d_k^T \nabla \Psi < 0$. Consider the formula (4.1) and (4.2) or (4.1) and (4.3) or (4.4), then we can get (2.5).

Remark 4.2. If we replace Line search I in Levenberg-Marquardt Method(I) by Line search III or Line search IV or Line search V, we also can get the global convergence result of Levenberg-Marquardt Method(I).

We also can use the above results to solve the linear complementarity problem(LCP)(see [13-17]), which is to find $x \in R^n$ such that

$$x \geq 0, Mx + q \geq 0, x^T(Mx + q) = 0,$$

where M is an $n \times n$ matrix and q is a n-dimension vector. Base on the F-B function

$$\phi(a, b) = \sqrt{a^2 + b^2} - a - b$$

in [13], we consider the following reformulation of the linear complementarity problem(LCP)

$$F(x) = \begin{pmatrix} \phi(x_1, (Mx+q)_1) \\ \vdots \\ \phi(x_n, (Mx+q)_n) \end{pmatrix} = 0.$$

Choosing $J_k \in \partial F(x_k)$, we can compute d_k by the following linear system

$$(J_k^T J_k + \mu_k I)d = -\nabla\Psi(x_k).$$

Since $\Psi(x) = \frac{1}{2}\|F(x)\|^2$ is differentiable, when M is a P_0 matrix, any stationary point of Ψ is a solution of the linear complementarity problem(LCP). Using the analysis and results obtained in the above, we can get the following result.

Remark 4.3. Suppose that M is a P_0 matrix, then any accumulation point of the sequence generated by the above Levenberg-Marquardt methods is a solution of the linear complementarity problem(LCP).

References

1. A.N.Iusem and M.V. Solodov, *Newton-type methods with generalized distances for constrained optimization*, Optimization **41** (1997), 257-278.

2. Y.B.Zhao and D. Li, *Monotonicity of fixed point and normal mappings associated with variational inequality and its application*, SIAM Journal on Optimization **11** (2001), 962-973.

3. C. Wang and Y. Wang and C. Xu, *A projection method for a system of nonlinear monotone equations with convex constraints*, Mathematical Methods of Operations Research **66** (2007), 36-46.

4. W.J.Zhou and D.Li, *Limited memory BFGS method for nonlinear monotone equations*, Journal of Computational Mathematics **25** (2007),89-96.

5. K. Levenberg, *A method for the solution of certain nonlinear problems in least squares*, Quart. Appl. Math. **2** (1944), 164-166.

6. D.W. Marquardt, *An algorithm for least-squares estimation of nonlinear inequalities*, SIAM J. Appl. Math. **11** (1963), 431-441.

7. J.Y.Fan and Y.X.Yuan , *On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption*, Computing **74** (2005), 23-39.

8. N.Yamashita and M. Fukushima, *On the rate of convergence of the Levenberg-Marquardt method*, Computing **15** (2001), 237-249.

9. C.Y.Wang,Y.Y.Chen and S.Q.Du, *Further insight into the Shamanskii modification of Newton method*, Appl. Math. Comput.**180** (2006), 46-52.

10. X. J.Zhang and A. N.Xu, *Global convergence properties of a new class of nonlinear conjugate gradient methods*, Guangxi Science **12** (2005), 282-283.

11. Y.H. Dai, *Conjugate gradient methods with Armijo-type line search*, Acta Mathematicae Applicatae Sinica (2002), 123-130.

12. Y.H. Dai and Y.X. Yuan, *Nonlinear conjugate gradient methods*, Shanghai Scientific and Technical Publishers, Shanghai, 2000.

13. A. Fischer, *Solution of monotone complementarity problems with locally lipschitzian functions*, Mathematical Programming **76** (1997), 513-532.

14. M.C.Ferris, O.L.Mangasarian and J.S. Pang, *Complementarity: Applications, Algorithms and Extensions*, Kluwer Academic, Dordrecht, 2001.

15. F.Facchinei and J.S. Pang, *Finite-dimensional variational inequalities and complementarity problem, Volume I-II*, Springer, New York, 2003.

16. M.Dehghan and M.Hajarian, *Convergence of SSOR methods for linear complementarity problems*, Operations Research Letters **37** (2009), 219-223.

17. Y.Elfoutayeni and M.Khaladi, *Using vector divisions in solving the linear complementarity problem*, Journal of Computational and Applied Mathematics **236** (2012), 1919-1925.

Shou-qiang Du PhD, Vice Professor. Since 2003 he has been at College of Mathematics, Qingdao University. His research interests include numerical optimization and nonsmooth optimization.

College of Mathematics, Qingdao University, Qingdao, 266071, China.
e-mail: dsq8933@163.com; sqdu@qdu.edu.cn