Numerische Mathematik manuscript No.
(will be inserted by the editor)

On the convergence of a regularizing Levenberg-Marquardt scheme for nonlinear ill-posed problems

Marlis Hochbruck · Michael Hönig

Received: July 21, 2009/ Accepted: date

Abstract In this note we study the convergence of the Levenberg-Marquardt regularization scheme for nonlinear ill-posed problems. We consider the case that the initial error satisfies a source condition. Our main result shows that if the regularization parameter does not grow too fast (not faster than a geometric sequence), then the scheme converges with optimal convergence rates. Our analysis is based on our recent work on the convergence of the exponential Euler regularization scheme [3].

Keywords Nonlinear ill-posed problems · Levenberg-Marquardt method · optimal convergence rates

Mathematics Subject Classification (2000) 65J15 · 65J20

1 Introduction

In this note we study the convergence rates of the Levenberg-Marquardt method for solving the nonlinear ill-posed problem

$$F(x) = y. \tag{1}$$

Here $F : \mathcal{D}(F) \subset X \to Y$ is a nonlinear differentiable operator between the Hilbert spaces X and Y, whose Fréchet derivative $F'(u)$ is locally uniformly bounded. We always assume that (1) has a solution $x_* \in \mathcal{D}(F)$ but we do not assume that this solution is unique. We are interested in the case that only perturbed data $y^\delta \approx y$ satisfying

$$\left\| y^\delta - y \right\| \le \delta, \tag{2}$$

is available. Throughout the paper, the norm in both Hilbert spaces X and Y is denoted by $\|\cdot\|$, the corresponding inner product by $\langle \cdot\,, \cdot \rangle$.

It has been shown by Hanke [1] that the Levenberg-Marquardt method

$$u_{n+1} = u_n + h_n\big(I + h_n J_n\big)^{-1} F'(u_n)^* \big(y^\delta - F(u_n)\big), \qquad n = 0, 1, 2, \ldots \tag{3}$$

M. Hochbruck, M. Hönig
Mathematisches Institut, Heinrich-Heine Universität Düsseldorf, D-40225 Düsseldorf, Germany

with

$$J(u) = F'(u)^* F'(u), \qquad J_n = J(u_n) \tag{4}$$

converges to a solution of the unperturbed problem (1) in the limit $\delta \to 0$ if the regularization parameter is chosen appropriately and if the iteration is stopped as soon as the standard discrepancy principle

$$\left\| \Delta F_{n_*}^\delta \right\| \le \tau\delta < \left\| \Delta F_n^\delta \right\| \qquad \text{for all } n < n_*, \tag{5}$$

is satisfied for some parameter $\tau > 1$. Hanke [1] suggests to select h_n such that the following discrepancy principle

$$\left\| \Delta F_n^\delta - F'(u_n)(u_{n+1} - u_n) \right\| = \mu \left\| \Delta F_n^\delta \right\|, \quad \mu < 1, \tag{6}$$

is satisfied. Here

$$\Delta F_n^\delta = y^\delta - F(u_n)$$

denotes the residual of the perturbed problem.

Rieder [7,8] managed to prove nearly optimal convergence rates for yet different adaptively chosen step sizes. Only recently, Jin [4] proved optimal convergence rates for an a priori chosen geometric step size sequence.

The aim of this note is to show that if the initial error satisfies a source condition, then the method converges with optimal rate for quite general step size sequences including the geometric sequence studied in [4]. Our analysis is based on our recent work [3], where we proved an analogous result for the exponential Euler regularization.

2 Preliminaries

In order to verify optimal convergence rates, certain assumptions have to be imposed. Let x_+ be the solution of minimal distance to x_0. The following assumptions ensure, that this solution is unique, see [6, Proposition 2.1]. Our main assumption is that the initial error satisfies a source condition.

Assumption 1 There exists $w \in X$ and constants $\gamma \in (0, 1/2]$ and $\rho \ge 0$ such that

$$e_0 = x_0 - x_+ = J(x_+)^\gamma w, \qquad \|w\| \le \rho.$$

Moreover, we have to assume relations between the Fréchet derivatives evaluated at two different points in $B_r(x_+)$.

Assumption 2 For all $x, \widetilde{x} \in B_r(x_+)$ there exist linear bounded operators $R(x, \widetilde{x}) : Y \to Y$ and a constant $C_R \ge 0$ such that

1. $F'(x) = R(x, \widetilde{x}) F'(\widetilde{x})$
2. $\|R(x, \widetilde{x}) - I\| \le C_R \|x - \widetilde{x}\|$.

Both assumptions are standard assumptions arising in the literature, see, e.g., [4–7]. Note, that for $C_R r < 1/2$ Assumption 2 implies the so-called tangential cone condition

$$\left\| F(\widetilde{x}) - F(x) - F'(x)(\widetilde{x} - x) \right\| \le \eta \left\| F(x) - F(\widetilde{x}) \right\|, \qquad x, \widetilde{x} \in B_r(x_+). \tag{7}$$

with $\eta = C_r r/(1 - C_r r) < 1$, see, e.g., [7]. Moreover it is possiple to slightly weaken Assumption 2 by fixing $\widetilde{x} = x_+$. This results in a slightly larger constant of $3/2C_R$ in (15) below, cf. equation (3.4) in [2].

To simplify the presentation we further assume without loss of generality that the problem is appropriately scaled, i.e.,

$$\|F'(x)\| \le 1, \qquad x \in B_r(x_+). \tag{8}$$

3 Convergence rates

The aim of this section is to show that the Levenberg-Marquardt regularization in fact converges with optimal rates. Our results are valid under weak restrictions on the step sizes, namely we assume that there exist constants c_0, c_h such that

$$h_0 \le c_0, \qquad 0 < h_j \le c_h t_j, \qquad j \ge 1, \tag{9}$$

where

$$t_0 = 0, \qquad t_{j+1} = t_j + h_j, \qquad j = 0, 1, 2, \dots. \tag{10}$$

Note that this step size restriction allows to choose $(h_j)_{j\ge 0}$ as a geometric sequence and thus our result generalizes the recent result [4].

Theorem 1 *Let Assumptions* 1 *and* 2 *hold and assume that the step sizes h_j satisfy* (9) *for all $j \le n_*$ and that $t_j \to \infty$ for $j \to \infty$. Here, the stopping index n_* is defined by* (5)*, with τ satisfying*

$$\tau > \frac{2-\eta}{1-\eta}. \tag{11}$$

Then for ρ sufficiently small, the iterates u_n stay in $B_r(x_+)$ for $n = 0, 1, \dots, n_*$ and the iteration stops after $n_* < \infty$ steps. Moreover, there exists a constant $C = C(\tau, \eta, C_R, c_0, c_h, \gamma, r) > 0$ such that

$$\|u_{n_*} - x_+\| \le C\rho^{1/(2\gamma+1)}\delta^{2\gamma/(2\gamma+1)}.$$

The proof of this theorem is postponed to the end of this note.

Remark. The assumption $t_j \to \infty$ for $j \to \infty$ is satisfied if the step sizes are bounded away from zero or if they do not decay faster than $1/j$, for instance.

Our analysis uses the discrete variation-of-constants formula (Theorem 2), which is derived from the following suitably written error recursion. Throughout the paper we denote the operators by

$$\begin{aligned} A_+ &= F'(x_+), & A_n &= F'(u_n),\\ J_+ &= A_+^* A_+, & J_n &= A_n^* A_n,\\ K_+ &= A_+ A_+^*, & K_n &= A_n A_n^*, \end{aligned}$$

and the corresponding operator functions by

$$\begin{aligned} \Phi_{n,+} &= (I + h_n J_+)^{-1}, & \Phi_n &= (I + h_n J_n)^{-1},\\ \widetilde{\Phi}_{n,+} &= (I + h_n K_+)^{-1}, & \widetilde{\Phi}_n &= (I + h_n K_n)^{-1}. \end{aligned}$$

Lemma 1 *Let Assumption* 2 *hold. Then the error*

$$e_n = u_n - x_+$$

of the Levenberg-Marquardt recursion (3) *satisfies*

$$e_{n+1} = \Phi_{n,+} e_n + h_n A_+^* \widetilde{\Phi}_{n,+} (r_n + y^\delta - y) \tag{12}$$

where, for $R_n = R(u_n, x_+)$ *and* $\widetilde{R}_n = R(x_+, u_n)$,

$$r_n = F(x_+) - F(u_n) + A_+ e_n + \left(R_n^* - I + (\widetilde{R}_n - R_n^*) \widetilde{\Phi}_n h_n K_n \right) \Delta F_n^\delta .$$

If in addition the the stopping index n_* *is defined by* (5), *then there is a constant* $C_1 = C_1(\tau, \eta, C_R, c_0, c_h, \gamma, r)$ *such that for* $n < n_*$ *we have*

$$\|r_n\| \leq C_1 \|e_n\| \|A_+ e_n\| . \tag{13}$$

Proof By (3), the following error recursion holds

$$\begin{aligned} e_{n+1} &= \Phi_{n,+} e_n + h_n A_+^* \widetilde{\Phi}_{n,+} (F(x_+) - F(u_n) + A_+ e_n) \\ &\quad + h_n \Phi_{n,+} A_+^* \left[(R_n^* - I) + h_n [(\widetilde{R}_n - I) - (R_n^* - I)] \widetilde{\Phi}_n K_n \right] \Delta F_n^\delta \\ &\quad + h_n \Phi_{n,+} A_+^* (y^\delta - y) \\ &= \Phi_{n,+} e_n + h_n A_+^* \widetilde{\Phi}_{n,+} \Big\{ F(x_+) - F(u_n) + A_+ e_n + y^\delta - y \\ &\qquad + \left[(R_n^* - I) + h_n [(\widetilde{R}_n - I) - (R_n^* - I)] \widetilde{\Phi}_n K_n \right] \Delta F_n^\delta \Big\} . \end{aligned}$$

This proves the error recursion.

It was shown in [3, Lemma 4.3], that if the stopping index n_* is defined by (5), then we have

$$\left\| \Delta F_n^\delta \right\| \leq \frac{\tau}{(\tau - 1)(1 - \eta)} \|A_+ e_n\| , \qquad n < n_* . \tag{14}$$

Moreover, equation (3.4) in [2] (for a slightly weaker form of Assumption 1) or [9, Proposition 4] yield

$$\|F(x_+) - F(u_n) + A_+ e_n\| \leq \frac{1}{2} C_R \|e_n\| \|A_+ e_n\| . \tag{15}$$

Defining

$$C_1 = C_R \left(\frac{1}{2} + 3 \frac{\tau}{(\tau - 1)(1 - \eta)} \right)$$

gives the bound (13). □

Next we prove that the error norms $\|e_n\|$ and $\|A_+ e_n\|$ decay with a rate proportional to $(1 + t_n)^\gamma$ and $(1 + t_n)^{\gamma + 1/2}$, respectively.

Theorem 2 *Let the assumptions of Theorem* 1 *hold. Then for* ρ *sufficiently small there is a constant* $C_* = C_*(\tau, \eta, C_R, c_0, c_h, \gamma, r)$ *such that for* $n \leq n_*$

$$\begin{aligned} \|e_n\| &\leq C_* \frac{\rho}{(1 + t_n)^\gamma} , \\ \|A_+ e_n\| &\leq C_* \frac{\rho}{(1 + t_n)^{\gamma + 1/2}} . \end{aligned}$$

Proof For an arbitrary $n \in \mathbb{N}$ the error recursion (12) leads to the following discrete variation-of-constants formulas

$$\begin{aligned} e_n &= \prod_{j=0}^{n-1} \Phi_{j,+} e_0 + \sum_{j=0}^{n-1} h_j \prod_{k=j+1}^{n-1} \Phi_{k,+} A_+^* \widetilde{\Phi}_{j,+} (r_j + y^\delta - y) \\ &= \prod_{j=0}^{n-1} \Phi_{j,+} e_0 + \sum_{j=0}^{n-1} h_j A_+^* \prod_{k=j}^{n-1} \widetilde{\Phi}_{k,+} (r_j + y^\delta - y). \end{aligned} \tag{16}$$

Moreover, we have

$$A_+ e_n = A_+ \prod_{j=0}^{n-1} \Phi_{j,+} e_0 + \sum_{j=0}^{n-1} h_j K_+ \prod_{k=j}^{n-1} \widetilde{\Phi}_{k,+} (r_j + y^\delta - y). \tag{17}$$

By Lemma 2 below, the sum multiplying $y^\delta - y$ in (16) can be bounded by

$$\left\| \sum_{j=0}^{n-1} h_j A_+^* \prod_{k=j}^{n-1} \widetilde{\Phi}_{k,+} \right\| \leq \sum_{j=0}^{n-1} h_j (t_n - t_j)^{-1/2} \leq \int_0^{t_n} \frac{1}{\sqrt{t_n - x}} \,\mathrm{d}x = 2\sqrt{t_n}$$

while the corresponding sum in (17) can be bounded by one by using the identity

$$\sum_{j=0}^{n-1} h_j K_+ \prod_{k=j}^{n-1} \widetilde{\Phi}_{k,+} = I - \prod_{j=0}^{n-1} \widetilde{\Phi}_{j,+}. \tag{18}$$

Thus, by Assumption 1, (13) and Lemma 2 we have

$$\|e_n\| \leq \frac{\rho}{(1+t_n)^\gamma} + 2\sqrt{t_n}\,\delta + C_1 \sum_{j=0}^{n-1} h_j \frac{1}{\sqrt{1 + t_n - t_j}} \|e_j\| \, \|A_+ e_j\|$$

and

$$\|A_+ e_n\| \leq \frac{\rho}{(1+t_n)^{\gamma+1/2}} + \delta + C_1 \sum_{j=0}^{n-1} h_j \frac{1}{1 + t_n - t_j} \|e_j\| \, \|A_+ e_j\| \,.$$

Following the proof of Theorem 4.8 in [3], we proceed by induction for $n = 0, 1, \ldots, n_*$. By Assumption 1, the statement is true for $n = 0$ if $C_* \geq 1$. Assuming that the bounds hold for all indices up to $n - 1$, we obtain

$$\|e_n\| \leq \frac{\rho}{(1+t_n)^\gamma} + 2\sqrt{t_n}\,\delta + C_*^2 \rho^2 C_1 S_n \left(\tfrac{1}{2}, 2\gamma + \tfrac{1}{2}\right)$$

and

$$\|A_+ e_n\| \leq \frac{\rho}{(1+t_n)^{\gamma+1/2}} + \delta + C_*^2 \rho^2 C_1 S_n \left(1, 2\gamma + \tfrac{1}{2}\right),$$

where

$$S_n(\alpha, \beta) = \sum_{j=0}^{n-1} \frac{h_j}{(1 + t_n - t_j)^\alpha (1 + t_j)^\beta}. \tag{19}$$

It was shown in Lemma 4.11 in [3], that the discrete sums can be bounded by

$$S_n \left(\alpha, 2\gamma + \tfrac{1}{2}\right) \leq C_2 \frac{1}{(1+t_n)^{\alpha+\gamma-1/2}}, \tag{20}$$

provided that the step sizes satisfy (9). This leads to

$$\|e_n\| \le \frac{\rho}{(1+t_n)^\gamma}\Big(1 + C_*^2\rho C_1 C_2\Big) + 2\sqrt{t_n}\,\delta, \tag{21}$$

$$\|A_+ e_n\| \le \frac{\rho}{(1+t_n)^{\gamma+1/2}}\,(1 + C_*^2\rho C_1 C_2) + \delta. \tag{22}$$

By induction hypothesis and by applying (5) and (14) we get

$$\begin{aligned}\delta &\le \frac{1}{(\tau-1)(1-\eta)}\,\|A_+ e_{n-1}\| \\ &\le \frac{1}{(\tau-1)(1-\eta)}\left(\frac{\rho}{(1+t_{n-1})^{\gamma+1/2}}\Big(1 + C_*^2\rho C_1 C_2\Big) + \delta\right).\end{aligned}$$

Using (9), we have

$$\frac{1}{1+t_{n-1}} \le \frac{1+c_h}{1+t_n}, \qquad n = 1, 2, \ldots$$

so that

$$\delta \le C_3 \frac{\rho}{(1+t_n)^{\gamma+1/2}} \tag{23}$$

with

$$C_3 = \frac{1+c_h}{(\tau-1)(1-\eta)-1}\Big(1 + C_*^2\rho C_1 C_2\Big)$$

holds. Inserting this relation into (21) shows

$$\|e_n\| \le \frac{\rho}{(1+t_n)^\gamma}\Big(1 + C_*^2\rho C_1 C_2 + 2C_3\Big),$$

$$\|A_+ e_n\| \le \frac{\rho}{(1+t_n)^{\gamma+1/2}}\Big(1 + C_*^2\rho C_1 C_2 + C_3\Big).$$

This yields the desired result, as long as

$$1 + C_*^2\rho C_1 C_2 + 2C_3 \le C_*,$$

holds, which can be achieved for ρ sufficiently small. □

In the previous proof, we have used the following estimate.

Lemma 2 *For $0 \le \alpha \le 1$ we have*

$$\left\| K_+^\alpha \prod_{k=j}^{n-1} \widetilde{\Phi}_{k,+} \right\| \le \min\{(t_n - t_j)^{-\alpha}, (1 + t_n - t_j)^{-\alpha}\}.$$

Proof The inequality

$$\prod_{k=j}^{n-1}(1 + h_k\lambda) \ge 1 + \lambda\sum_{k=j}^{n-1} h_k = 1 + \lambda(t_n - t_j)$$

shows that

$$\lambda^\alpha \prod_{k=j}^{n-1}(1 + h_k\lambda)^{-\alpha} \le \left(\frac{\lambda}{1 + \lambda(t_n - t_j)}\right)^\alpha.$$

For $x \in [0,1]$ the function $x/\big(1 + x(t_n - t_j)\big)$ attains its maximum at $x = 1$. This proves the second bound.

The first part of the bound was also used in [6, p. 109] or [4, Lemma 2]. □

Remark. If the maximum possible step sizes $h_j = c_h t_j$, $j = 1, \dots, n_* - 1$ are chosen, then (14), Theorem 2, and (5) show that there is a constant c such that the stopping index satisfies $n_* \le c\,|\log \delta|$.

It remains to prove our main theorem.

Proof (of Theorem 1) By Theorem 2, the iterates u_n stay in $B_{C_*\rho}(x_+)$ for all $n = 0, 1, \dots, n_*$. Moreover, using (14) and $t_n \to \infty$, the bound of $\|A_+ e_n\|$ also shows that the stopping index n_* is finite.

In order to prove the convergence rate, we write (16) for $n = n_*$ in the form

$$e_{n_*} = J_+^\gamma v_* + \sum_{j=0}^{n-1} h_j A_+^* \prod_{k=j}^{n_*-1} \widetilde{\Phi}_{k,+}(y^\delta - y),$$

where

$$v_* = \prod_{k=0}^{n_*-1} \Phi_{k,+} w + \sum_{j=0}^{n_*-1} h_j \prod_{k=j}^{n_*-1} \Phi_{k,+} J_+^{-\gamma} A_+^* r_j .$$

Note that v_* is well defined since

$$J_+^{-\gamma} A_+^* : \mathcal{N}(A_+^*)^\perp \to X$$

is a bounded operator for $\gamma \le \frac{1}{2}$.

Using (13), Theorem 2, Lemma 2, (19), and (20) we obtain

$$\|v_*\| \le \rho + C_1 C_*^2 \rho^2 S_n \left(\frac{1}{2} - \gamma, 2\gamma + \frac{1}{2}\right) \le C_4 \rho .$$

Moreover, the telecopic identity (18) and (7) imply

$$\left\|A_+ J_+^\gamma v_*\right\| \le \|A_+ e_{n_*}\| + \delta \le (1+\eta)\left(\left\|\Delta F_{n_*}^\delta\right\| + \delta\right) + \delta \le C_5 \delta .$$

with $C_5 = (1+\eta)(1+\tau) + 1$. The desired bounds follow as in [3]. □

4 Concluding Remarks

In this paper we proved that the Levenberg-Marquardt regularization method converges with optimal rates under suitable assumptions. If the step sizes are chosen according to the discrepancy principle (6) proposed by Hanke [1], then it was shown in [1] that the method converges without requiring a source condition. If the source condition (cf. Assumption 1) is satisfied, then Theorem 1 shows that the rate of convergence is optimal, if the step sizes chosen by (6) do not grow faster than (9). Note that (9) is satisfied if $h_{j+1}/h_j \le$ const, $j = 0, 1, \dots$, so that this result appears to be relevant for practical applications. However, if (9) fails to be true, then Theorems 1 guarantees that one can switch to any step size sequence satisfying (9) and being bounded away from zero and still gets optimal convergence rates.

References

1. Hanke, M.: A regularizing Levenberg–Marquardt scheme, with applications to inverse groundwater filtration problems. Inverse Problems **13**(1), 79–95 (1997)
2. Hanke, M., Neubauer, A., Scherzer, O.: A convergence analysis of the Landweber iteration for nonlinear ill-posed problems. Numerische Mathematik **72**(1), 21–37 (1995)
3. Hochbruck, M., Hönig, M., Ostermann, A.: A convergence analysis of the exponential Euler iteration for nonlinear ill-posed problems. Inverse Problems (to appear)
4. Jin, Q.: On a regularized Levenberg-Marquardt method for solving nonlinear inverse problems. Tech. rep., University of Texas at Austin (2009)
5. Jin, Q., Tautenhahn, U.: On the discrepancy principles for some Newton type methods for solving nonlinear inverse problems. Numerische Mathematik **111**(4), 509–558 (2009)
6. Kaltenbacher, B., Neubauer, A., Scherzer, O.: Iterative Regularization Methods for Nonlinear Ill-Posed Problems. De Gruyter, Berlin, New York (2008)
7. Rieder, A.: On the regularization of nonlinear ill-posed problems via inexact Newton iterations. Inverse Problems **15**(1), 309–327 (1999)
8. Rieder, A.: On convergence rates of inexact Newton regularizations. Numerische Mathematik **88**(2), 347–365 (2001)
9. Tautenhahn, U.: On the asymptotical regularization of nonlinear ill-posed problems. Inverse Problems **10**(6), 1405–1418 (1994)