EESTI NSV TEADUSTE AKADEEMIA TOIMETISED.
FÜÜSIKA * MATEMAATIKA
ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ФИЗИКА * МАТЕМАТИКА
PROCEEDINGS OF THE ACADEMY OF SCIENCES OF THE ESTONIAN SSR.
PHYSICS * MATHEMATICS

1989, 38, 2

УДК 517.948 : 513.88

O. VAARMANN

SOLUTION OF NONLINEAR LEAST SQUARES PROBLEMS BY LEVENBERG—MARQUARDT TYPE METHODS

(Presented by G. Vainikko)

This paper is concerned with minimization of a functional

$$\varphi(x) = \|F(x)\|^2, \tag{1}$$

where F is a Frechet differentiable operator which maps one Hilbert space, H_1, into another, H_2, and its derivative has the bounded pseudo-inverse $[F'(x)]^+$. A necessary condition for x^* to be a local minimum point of φ is

$$[F'(x)]^* F(x) = 0. \tag{2}$$

To solve the equation (2) we consider methods in the form

$$x_{k+1} = x_k - \varepsilon_k A_k F(x_k), \quad k = 0, 1, \ldots, \tag{3}$$

where ε_k is a relaxation (damping) parameter ($0 < \varepsilon_k \leqslant 1$), A_k is a linear operator from H_2 into H_1 approximating the operator $U^{-1}(F'(x)U^{-1})^+$, and U is an arbitrary nonsingular linear operator from H_1 into itself. In the case $\varepsilon_k = 1$ and $A_k = U^{-1}(F'(x_k)U^{-1})^+$ one gets an extension of Ben-Israel method which coincides with the usual Ben-Israel method if $U = I$, where I denotes identity mapping. In keeping with the terminology of [1] the operator $U^{-1}(F'(x)U^{-1})^+$ can be called the IU-weighted pseudoinverse of $F'(x)$ and, in general case U is allowed [1,2,3] to be a singular operator. Note that the introduction of the nonsingular operator U into the computational scheme (3) is not principal but it is due to the convenience of the treatment of some problems in numerical practice [1,3]. The solution of a problem depends on the spaces considered. Via a change in the topology, the initial problem can be converted into one with $U = I$. The philosophy underlying many methods of regularization lies in deciding in what sense one should understand an approximate solution of an ill-posed problem.

The main aim of this paper is to study the convergence properties of various methods based on solving the normal equation within a single framework of the notion of the IU-weighted pseudoinverse. Introduction of the damping parameter ε is a common strategy for enlarging the domain of convergence and sometimes also for converging acceleration in the sense of the required bulk of iterations or run time in seconds on the computer [4,5].

The family of methods (3) involves in particular cases underrelaxed or damped Levenberg—Marquardt type methods as follows

$$x_{k+1} = x_k - \varepsilon_k D_k [F'(x_k)]^* F(x_k), \quad k = 0, 1, \ldots,$$

where D_k is an approximation to $M^{-1}(x_k, \alpha_k) := \{[F'(x_k)]^* F'(x_k) + \alpha_k H\}^{-1}$ with $\alpha_k > 0$ and $A_k := D_k [F'(x_k)]^*$. It is assumed that H is represented in the form $H = U^* U$. In order to guarantee the global convergence for

Levenberg—Marquardt (regularized Gauss—Newton) type methods, the regularization parameter α used in [4] is chosen from the considerations that operators $M(\cdot,\alpha)$ appear to be positive definite ones. Here for obtaining the convergence for methods (3) from a poor starting point x_0 operators A_k need not necessarily be positive definite.

Great values of parameter α can lower the convergence. On the other hand, tending the parameter α to zero can cause a numerical instability. This is the reason why we are interested in the variants of Levenberg—Marquardt type methods with successive approximation of the inverse (pseudoinverse) operator using only linear operators multiplications because the formula (17) appears to be self-correcting one with $\alpha>0$. Moreover, methods (3) based on formula (17) do not break down even with $\alpha=0$, provided operators $F'(x)$ and $F'(y)$ are in the acute case for all x and y belonging to a region under discussion [6].

1. Since there exists the bounded pseudoinverse, then the operator $F'(x)$ has the closed range $R(x)=R(F'(x))$.

Let

$$\Phi(x)=[F'(x)U^{-1}]^{+}, \quad \Psi(x)=\Phi^{+}(x)=F'(x)U^{-1}, \tag{4}$$

and let $P_{R(x)}=P_{R(\Psi(x))}$ $(P_k=P_{R(\Psi(x_k))})$ denote the orthogonal projector of H_2 onto $R(\Psi(x))$. Let further C, $\bar{C}$, $\bar{L}_0$, L_1, N, N_0, N'_1, N_2 and L denote some positive constants. Supposing that $\bar{A}_k$ is an approximation to $[F'(x_k)U^{-1}]^{+}$, i. e. $\bar{A}_k=UA_k$ and the following relations hold

$$\bar{A}_k=\bar{A}_kP_k, \quad \|\bar{A}_k\|\leqslant\bar{\lambda}_k\leqslant\bar{\lambda}<\infty, \quad \|A_k\|\leqslant\lambda_k\leqslant\lambda<\infty; \tag{5}$$

using the results ([7] p. 195—197), it is not difficult to prove the following assertion.

Lemma 1. *Let $F'(x)$ be Lipschitz continuos and $\Phi(x)$ be uniformly bounded on some set S, i. e.*

$$\|F'(x)-F'(y)\|\leqslant L_1\|x-y\|, \quad \forall x, y\in S, \tag{6}$$

$$\|\Phi(x)\|\leqslant\bar{C}, \quad \forall x\in S, \tag{7}$$

then

$$\|P_{R(x)}-P_{R(y)}\|\leqslant\bar{L}_0\|x-y\|, \quad \forall x, y\in S, \tag{8}$$

where $\bar{L}_0=\bar{C}L_1\|U^{-1}\|$.

Proof of lemma 1 is similar to that in [5].

Remark 1. From Lemma 1 it follows that there exist scalars N' and N such that

$$\|(P_{R(y)}-P_{R(y)}P_{R(x)})F(x)\|=\|(P_{R(y)}-P_{R(x)})(I-P_{R(x)})F(x)\|\leqslant \leqslant N'\|x-y\|, \quad x, y\in S, \tag{9}$$

and

$$\|(P_{k+1}-P_{k+1}P_k)F(x_k)\|\leqslant\varepsilon_kN\|P_kF(x_k)\|, \quad x_k, x_{k+1}\in S, \tag{10}$$

where $\sup\limits_{x\in S}\|(I-P_{R(x)})F(x)\|<N''$, $N'=N''\bar{L}_0$, and $N=N'\lambda$. It can be readily shown that if F satisfies Lipschitz condition with some constant L_1 then Ψ satisfies Lipschitz condition with the constant $\bar{L}_1=L_1\|U^{-1}\|$.

2. Let $\bar{A}_k=UA_k$. Introducing a sequence $\{\bar{\gamma}_k\}$ defined as

$$\|P_k-\Psi(x_k)\bar{A}_k\|\leqslant\bar{\gamma}_k, \quad k=0, 1, \ldots, \tag{11}$$

one gets

Theorem 1. *Let $x_0 \in H_1$, $S=\{x \in H_1: \|x-x_0\| \leqslant \varrho\}$, and the following conditions be valid in S:*

1° *operator F has Frechet-derivative F′:*

2° *derivative F′ satisfies Lipschitz condition*
$\|F'(x)-F'(y)\| \leqslant L_1\|x-y\|$;

3° $\delta=\delta_0<1$ (*below* δ_0 *is defined differently in the cases A and B*);

4° $0<\varepsilon_0 \leqslant \varepsilon_{k-1} \leqslant \varepsilon_k = \min\{1, \varepsilon_{k-1}\delta^{-1/2}\}$.

A. *If there exists a constant* $\bar{C}$ *such that* $\|\Phi(x)\| \leqslant \bar{C}$, $\bar{\gamma}_k \leqslant \bar{\gamma}_0$ *and* $r_1 = \lambda\|P_0F(x_0)\|/(1-\delta) \leqslant \varrho$, *then the sequence* $\{x_k\}$ *generated by the method* (3) *has a limit* x^* *which appears to be a solution of the equation* $[F'(x)]^*F(x)=0$ *with* $\|x^*-x_0\| \leqslant r_1$, $\delta_k \leqslant \delta$, *and* $\|x_k-x^*\| \leqslant r_1\delta^k$, *where* $\delta=\delta_0=1-\varepsilon_0+\varepsilon_0(\bar{\gamma}_0+N)+\frac{1}{2}\varepsilon_0^2\lambda^2L_1\|P_0F(x_0)\|$.

B. *If operator* $P_{R(x)}$ *is independent of x and* $\bar{\gamma} \geqslant \bar{\gamma}_0 \geqslant \ldots \geqslant \bar{\gamma}_k \geqslant \ldots \geqslant 0$ ($\bar{\gamma}_k \to 0$ *as* $k \to \infty$), *then the sequence* $\{x_k\}$ *converges superlinearly with* $\|x^*-x_0\| \leqslant r_1$, *and*

$$\|x_k-x^*\| \leqslant r_1 \prod_{i=0}^{k-1} \delta_i,$$

where $\delta_i = 1-\varepsilon_i+\varepsilon_i\bar{\gamma}_i+\frac{1}{2}\varepsilon_i^2\lambda^2L_1\|P_iF(x_i)\| \to 0$ *as* $i \to \infty$.

Proof. Denote $\varphi_k(\varepsilon)=\varepsilon\psi_k+1-\varepsilon$, where ψ_k are some real numbers. If $\psi_k \leqslant \psi < 1$ for all $k=0, 1, \ldots$, then $\frac{d}{d\varepsilon}\varphi_k(\varepsilon)=\psi_k-1 \leqslant \psi-1<0$ and hence $\varphi_k(\varepsilon)$ is a decreasing function in ε.

Since $\|P_k-\varepsilon\Psi(x_k)\bar{A}_k\| \leqslant \varepsilon\|P_k-\Psi(x_k)\bar{A}_k\|+1-\varepsilon \leqslant \varepsilon\bar{\gamma}+1-\varepsilon$, in the capacity of $\bar{\gamma}_k(\varepsilon)$, one can take $\bar{\gamma}_k(\varepsilon) \stackrel{\triangle}{=} \varepsilon\bar{\gamma}_k+1-\varepsilon$.

According to Taylor expansion one obtains

$$P_{k+1}F(x_{k+1}) = (P_{k+1}-P_{k+1}P_k)F(x_k)+P_{k+1}\{(P_k-\varepsilon_k\Psi(x_k)\bar{A}_k)F(x_k)+ \\ +\varepsilon_k\int_0^1[F'(x_k)-F'(x_k+t(x_{k+1}-x_k))]A_kF(x_k)\,dt.\}. \quad (12)$$

A. On the basis of the remark 1 one gets $\|P_{k+1}F(x_{k+1})\| \leqslant \delta_k\|P_kF(x_k)\|$ where $\delta_k=\varepsilon_kN+\bar{\gamma}_k(\varepsilon)+\frac{1}{2}\varepsilon_k^2\lambda^2L_1\|P_kF(x_k)\|$. Supposing $\delta_k \leqslant \delta_0$, it will be shown that the relation $\delta_{k+1} \leqslant \delta_0$ holds as well.

Indeed in virtue of the preposition 4° $\varepsilon_{k+1}^2\delta_k \leqslant \varepsilon_{k+1}^2\delta \leqslant \varepsilon_k^2$ and $\|P_{k+1}F(x_{k+1})\| \leqslant \delta_k\|P_kF(x_k)\|$ the following relations $\varepsilon_{k+1}^2\lambda^2L_1 \times \|P_{k+1}F(x_{k+1})\| \leqslant \varepsilon_{k+1}^2\lambda^2L_1\delta_k\|P_kF(x_k)\| \leqslant \varepsilon_k^2\lambda^2L_1\|P_kF(x_k)\|$ are valid. Setting now $\psi_k=\bar{\gamma}_k+N$ from the proposition 3° and $\bar{\gamma}_k \leqslant \bar{\gamma}_0$, it follows that $\psi_k \leqslant \psi_0 < 1$.

Therefore $\varphi_{k+1}(\varepsilon_{k+1}) \leqslant \varphi_{k+1}(\varepsilon_0) \leqslant \varphi_0(\varepsilon_0)$ and

$$\delta_{k+1}=\varphi_{k+1}(\varepsilon_{k+1})+\frac{1}{2}\varepsilon_{k+1}^2\lambda^2L_1\|P_{k+1}F(x_{k+1})\| \leqslant \ldots \\ \ldots \leqslant \varphi_0(\varepsilon_0)+\frac{1}{2}\varepsilon_0^2\lambda^2L_1\|P_0F(x_0)\|=\delta_0.$$

Further completion of the proof is similar to that in section B.

B. If $P_{R(x)}$ is independent of x, then $\delta_k=\bar{\gamma}_k(\varepsilon_k)+\frac{1}{2}\varepsilon_k^2\lambda^2L_1\|P_kF(x_k)\|$ and

$$\bar{\gamma}_{k+1}(\varepsilon_{k+1})=\varphi_{k+1}(\varepsilon_{k+1}) \leqslant \varphi_{k+1}(\varepsilon_k) \leqslant \varphi_k(\varepsilon_k)=\bar{\gamma}_k(\varepsilon_k).$$

If $\varepsilon_{k+1}^2\delta_k \leqslant \varepsilon_{k+1}^2\delta \leqslant \varepsilon_k^2$ or $\varepsilon_{k+1}=\varepsilon_k$, then one concludes by the induction that $\delta_{k+1}\leqslant\delta_k\leqslant\ldots\leqslant\delta_0<1$. The latter relations with the assumption $\bar{\gamma}_k\to 0$, while $k\to\infty$ implies that $\delta_k\to 0$. It is not difficult to see that (cf. [8])

$$\|x_{k+1}-x_k\|\leqslant\lambda_k\|P_kF(x_k)\|\leqslant\lambda_k\|P_0F(x_0)\|\prod_{i=0}^{k-1}\delta_i\leqslant\lambda\|P_0F(x_0)\|\delta^k,$$

$$\|x_m-x_k\|\leqslant r_1(\delta^k-\delta^m),\qquad (m\geqslant k),$$

$$\|x_k-x_0\|\leqslant r_1\leqslant\varrho,$$

$$x^*=\lim_{k\to\infty}x_k,\qquad \lim_{k\to\infty}\|P_kF(x_k)\|=0,\qquad \|x^*-x_k\|\leqslant r_1\delta^k.$$

Observing that

$$\|[F'(x)]^*F(x)\|\leqslant\|H\|\,\|H^{-1}[F'(x)]^*F(x)\|\leqslant\|H\|\,\|U^{-1}(F'(x)U^{-1})^*F(x)\|\leqslant$$
$$\leqslant\|H\|\,\|U^{-1}(F'(x)U^{-1})^*\Psi(x)\Phi(x)F(x)\|\leqslant$$
$$\leqslant\|H\|\,\|U^{-1}(F'(x)U^{-1})^*\|\,\|P_{R(x)}F(x)\|$$

and under the assumption of Theorem 1 the operator $P_{R(x)}F(x)$ is continuous, it is evident that $[F'(x^*)]^*F(x^*)=0$.

Remark 2. The rate of change for the parameter ε is controlled by the condition 4° of Theorem 1 which guarantees the validity of $\delta_k\leqslant\delta_0$ or $\delta_k\leqslant\delta_{k-1}$. It is obvious that the condition 4° can be replaced by a less severe restriction $0<\varepsilon_0\leqslant\varepsilon_{k-1}\leqslant\varepsilon_k=\min\{1,\ \varepsilon_{k-1}\delta_{k-1}^{-1/2}\}$.

The condition $\delta=\delta_0<1$ can be rewritten as

$$\varepsilon_0\left(\psi_0+\frac{1}{2}\varepsilon_0\lambda^2L_1\|P_0F(x_0)\|\right)\leqslant\varepsilon_0,$$

where $\psi_0=\bar{\gamma}_0+N$ and $\psi_0=\bar{\gamma}_0$ in the cases A and B, respectively. Thus if $\psi_0<1$, then, clearly, by decreasing ε_0 one can accomplish that $\psi_0+\frac{1}{2}\varepsilon_0\lambda^2L_1\|P_0F(x_0)\|<1$.

It means that the condition $\delta_0<1$ can be guaranteed for all $x\in S$.

3. Now consider the case where in extension of Ben-Israel method instead of the derivative $F'(x)$ a finite-difference approximation $\Delta F(x,h)$ $(h\in H_1)$ is used, i. e. $A_k=U^{-1}[\Delta F(x_k,h_k)U^{-1}]^+$. Let C_1, $\bar{C}_1$, C_2 and C_3 be some constants. Assuming $R(F'(x))=R(\Delta F(x,h))$, $R([F'(x)]^*)=$ $=R([\Delta F(x,h)]^*)$ ([9] Section II. 25, [10]) and $\|\bar{A}_k\|=\|[\Delta F(x_k,h_k)\times$ $\times U^{-1}]^+\leqslant\bar{C}_1$ for all $x\in S$, then

$$[\Delta F(x,h)]^+=[\Delta F(x,h)]^+P_{R(F'(x))}=[\Delta F(x,h)]^+\Delta F(x,h)[\Delta F(x,h)]^+,$$

$$\|P_k-\psi(x_k)\bar{A}_k\|=\|(\Delta F(x_k,h_k)-F'(x_k))U^{-1}\bar{A}_k\|,$$

$$\|A_k\|=\|U^{-1}\bar{A}_k\|\leqslant\|U^{-1}\|\bar{C}_1=\lambda.$$

If, in addition, the $\Delta F(x,h)$ is a «consistent approximation» to $F'(x)$ ([9] p. 344—352), then, from the assumption 2° of Theorem 1, it follows that

$$\|\Delta F(x,h)-\Delta F(y,h)\|\leqslant L_\Delta\|x-y\|,\tag{13}$$

where L_Δ is a constant (scalar).

Remark 3. If in the capacity of $\Delta F(x_k,h_k)$ to put a first order divided difference $F(y_k;x_k)$ where the basic elements x_k and y_k satisfy the condition $\|y_k-x_k\|\leqslant C_2\|P_kF(x_k)\|$ and symmetric second order divided difference $F(y_k;x_k;x_k)$ is bounded with a constant say K', then one can take $\bar{\gamma}_k=C_3\|P_kF(x_k)\|$.

Indeed,

$$\|F(y_k; x_k) - F'(x_k)\| = \|F(y_k; x_k) - F(x_k; x_k)\| \leqslant \|F(y_k; x_k; x_k)\| \, \|y_k - x_k\|,$$

$$\|P_k - \Psi(x_k)\bar{A}_k\| \leqslant C_3 \|P_k F(x_k)\| \quad \text{with} \quad K' = L_\Delta \quad \text{and} \quad C_3 = K' C_2 \lambda.$$

4. Let $B(x) = [F'(x)]^* F'(x)$, $B_k = B(x_k)$, $M(x, \alpha) = B(x) + \alpha H$, $M_k = M(x_k, \alpha_k)$, $L(x, \alpha) = M^{-1}(x, \alpha)[F'(x)]^*$, where H is a symmetric positive definite operator from H_1 into itself and α is a positive-valued scalar parameter.

For solving (2) one might use Levenberg—Marquardt-type methods

$$x_{k+1} = x_k - \varepsilon_k D_k [F'(x_k)]^* F(x_k), \tag{14}$$

where D_k is an approximation to M_k^{-1} and $A_k = D_k [F'(x_k)]^*$.

Let

$$\omega_k = \mu_k / \alpha_k, \qquad V = \|U\| \, \|H^{-1}\| \tag{15}$$

and μ_k and K be some scalars satisfying

$$\|I - M_k D_k\| \leqslant \mu_k < 1, \qquad \|F'(x)\| \leqslant K. \tag{16}$$

Theorem 2. *Let $x_0 \in H_1$, $S = \{x \in H_1 : \|x - x_0\| \leqslant \varrho\}$, and let conditions $1° - 4°$ be fulfilled in S.*

A. *If $\|\Phi(x)\| \leqslant \bar{C}$, $\alpha_k \leqslant \alpha_0$, $\omega_k \leqslant \omega_0$ $(\alpha_0, \omega_0 < \infty)$ and $r_1 = \lambda \|P_0 F(x_0)\| / (1 - \delta) \leqslant \varrho$, where $\delta = \delta_0 = 1 - \varepsilon_0 + \varepsilon_0 [N + K\|U^{-1}\| (\alpha_0 \bar{C}^3 + \omega_0 K \|H^{-1}\|)] + \frac{1}{2} \varepsilon_0^2 \lambda^2 L_1 \|P_0 F(x_0)\|$, then the sequence $\{x_k\}$ generated by the method (14) has a limit x^* which appears to be a solution of the equation (2) with $\|x^* - x_0\| \leqslant r_1$ and*

$$\|x_k - x^*\| \leqslant r_1 \delta^k.$$

B. *If $P_{R(x)}$ is independent of x, $n\xi^k \leqslant \alpha_k \leqslant m\xi^k$ $(0 < m, n < \infty)$, $0 < \xi < 1$, $\omega_k \leqslant s l^k$, $s > 0$, $0 < l < 1$, then the sequence (14) converges superlinearly with $\|x^* - x_0\| \leqslant r_1$ and*

$$\|x_k - x^*\| \leqslant r_1 \prod_{i=0}^{k-1} \delta_i,$$

where $\delta_i = 1 - \varepsilon_i + \varepsilon_i K \|U^{-1}\| (m \bar{C}^3 \xi^i + s l^i K \|U\| \, \|H^{-1}\|) + \frac{1}{2} \varepsilon_i^2 \lambda_i^2 L_1 \times \|P_i F(x_j)\| \to 0$ *as* $i \to \infty$.

Proof. Since

$$\|(B_k + \alpha_k H)^{-1}\| = \|(H^{-1} B_k + \alpha_k I)^{-1} H^{-1}\| \leqslant \alpha_k^{-1} \|H^{-1}\|,$$

$$\|L(x_k, \alpha_k)\| = \|M_k^{-1} [F'(x_k)]^*\| = \|(H^{-1} B_k + \alpha_k I)^{-1} [F'(x_k)]^*\| =$$

$$= \|U^{-1} (U^{*-1} B_k U^{-1} + \alpha_k I)^{-1} U^{*-1} [F'(x_k)]^*\| \leqslant \|U^{-1}\| \, \|[F'(x_k) U^{-1}]^+\|,$$

then

$$\|\Phi(x_k) - UL(x_k, \alpha_k)\| = \|\Phi(x_k) - (\Psi^*(x_k) \Psi(x_k) + \alpha_k I)^{-1} \Psi^*(x_k)\| =$$

$$= \|(\Psi^*(x_k) \Psi(x_k) + \alpha_k I)^{-1} \alpha_k \Phi(x_k)\| =$$

$$= \|\alpha_k (\Psi^*(x_k) \Psi(x_k) + \alpha_k I) \Psi^*(x_k) \Psi^{*+}(x_k) \Phi(x_k)\| \leqslant$$

$$\leqslant \alpha_k \|UL(x_k, \alpha_k)\| \, \|\bar{C}\|^2 < \alpha_k \bar{C}^3,$$

$$\|L(x_k, \alpha_k) - A_k\| = \|(B_k + \alpha_k H)^{-1} (I - M_k D_k) [F'(x_k)]^*\| \leqslant \omega_k K \|H^{-1}\|,$$

$$\|\bar{A}_k\| = \|\Phi(x_k) + U(L(x_k, \alpha_k) - L(x_k, \alpha_k)) + \bar{A}_k - \Phi(x_k)\| \leqslant$$

$$\leqslant \bar{C} + \|U(L(x_k, \alpha_k) - A_k)\| + \|UL(x_k, \alpha_k) - \Phi(x_k)\| \leqslant \bar{C} + \omega_k K V + \alpha_k \bar{C}^3,$$

$$\|P_k - \Psi(x_k) \bar{A}_k\| = \|\Psi(x_k) [\Phi(x_k) - \bar{A}_k]\| = \|\Psi(x_k) [\Phi(x_k) -$$

$$- U(L(x_k, \alpha_k) - L(x_k, \alpha_k)) - \bar{A}_k\| \leqslant K \|U^{-1}\| (\alpha_k \bar{C}^3 + \omega_k K V).$$

In the case of A and B, one can take, respectively,

$$\bar{\lambda}=\bar{C}+\alpha_0\bar{C}^3+\omega_0 KV,$$

$$\bar{\lambda}=\bar{C}+m\bar{C}^3+sKV,$$

and

$$\bar{\gamma}_k=\alpha_k\bar{C}^3K\|U^{-1}\|+\omega_k K^2V|U^{-1}\|,$$

$$\bar{\gamma}_k=m\xi^k\bar{C}^3K\|U^{-1}\|+sl^kK^2V\|U^{-1}\|.$$

Existence of scalar K such that $\|F'(x)\|\leqslant K$ follows from (7).

Since $\|I-M_kD_k\|\leqslant\mu_k<1$ and $\lim\limits_{\alpha_k\to 0} M_k^{-1}[F'(x_k)]^*=U^{-1}[F'(x_k)U^{-1}]^+$, then $A_k=D_k[F'(x_k)]^*$ approximates the $U^{-1}\Phi(x_k)$ and one can apply Theorem 1.

R e m a r k 4. Regularizing properties of method (14) provided $D_k=M_k^{-1}$ and its connections with the regularization method by A. N. Tikhonov are discussed in [11, 12].

R e m a r k 5. Methods of type (14) can also be called methods «of two parametric damping» because ε as well as α limit the change $x_{k+1}-x_k$. It is well known that under certain conditions Levenberg—Marquardt method is globally convergent [4, 13]. There exists sufficiently large α_k such that ε_k can be chosen as $\varepsilon_k=1$ for all k. Computational difficulties arise with these methods when the operator $F(x)$ is badly nonlinear. Moreover, introducing the damping parameter ε enables to sparse the amount of the arithmetic per iteration step, avoiding the need for solving repeatedly one linear equation or inverting one linear operator at an iteration step ([4] p. 258).

5. Consider now variants Levenberg—Marquardt-type methods where D_k is generated recurrently by the following prescription

$$D_{k+1}=D_k\sum_{i=0}^{q-1}(I-M_{k+1}D_k)^i,\quad q\geqslant 2,\quad k=0,\ 1,\ \ldots . \tag{17}$$

In addition to the quantities μ_k, $\bar{\gamma}_k$, λ_k, $\bar{\lambda}_k$ and ω_k given by (5), (11), (15), (16) let us define σ_k as follows

$$\|I-M_{k+1}D_k\|\leqslant\sigma_k,\quad k=0,\ 1,\ \ldots, \tag{18}$$

and put

$$\mu_0=\max\{\|I-M_0D_0\|,\ \eta^q\},\quad \max\{\eta_0,\sigma_0\}\leqslant\eta,\quad \max\{\sigma_i,\xi\}\leqslant\zeta_i, \tag{19}$$

$$\eta=(1+\sigma_i+\ldots+\sigma_i^{q-1})\zeta_i,\quad \bar{\lambda}=\bar{\lambda}_0=\bar{C}+\alpha_0\bar{C}^3+\Omega K\|U\|\,\|H^{-1}\|, \tag{20}$$

$$\Omega=\max\{\mu_0/\alpha_0,\ h^q/n\xi\},\quad h=\max\{\eta^{q-1}+\Lambda,\ \mu_0+\Lambda\}, \tag{21}$$

$$\Lambda=(2\varepsilon_0KL_1\lambda\|P_0F(x_0)\|+(m-n\xi)\|H\|)\|D_0\|. \tag{22}$$

T h e o r e m 3. *Let* $x_0\in H_1$, $S=\{x\in H_1:\|x-x_0\|\leqslant\varrho\}$, *and let, in addition to the conditions* $1°-4°$ *of Theorem* 1, *the following conditions be fulfilled:*

1° $\eta^q\leqslant\xi$, $h^q\leqslant\eta$ *and* $\eta, h, \xi<1$;

2° $n\xi^k\leqslant\alpha_k\leqslant m\xi^k$ $(0<m,\ n<\infty)$.

A. *If the condition* (7) *holds,* $r_1=\lambda_0\|P_0F(x_0)\|/(1-\delta)\leqslant\varrho$, *where* $\delta=\delta_0=1-\varepsilon_0+\varepsilon_0[N+K\|U^{-1}\|(m\bar{C}^3+\Omega K\|U\|\,\|H^{-1}\|)]+\frac{1}{2}\varepsilon_0^2\lambda^2L_1\|P_0F(x_0)\|,$

then the sequence $\{x_k\}$ *generated by* (14), (17) *has a limit* x^* *which appears to be solution of the equation* (2) *with* $\|x^*-x_0\|\leqslant r_1$ *and*

$$\|x_k-x^*\|\leqslant r_1\delta^k.$$

B. *If* $P_{R(x)}$ *is independent of* x, *then the sequence* $\{x_k\}$ *converges superlinearly with* $\|x^*-x_0\|\leqslant r_1$ *and*

$$\|x_k-x^*\|\leqslant r_1\prod_{i=0}^{k-1}\delta_i,$$

where

$$\delta_i=1-\varepsilon_i+\varepsilon_i K\|U^{-1}\|[m\bar{C}^3\xi^i+\Omega K\|U\|\,\|H^{-1}\|(\eta^q/\xi)^i]+$$
$$+\frac{1}{2}\varepsilon_i^2\lambda_i^2 L\|P_iF(x_i)\|\to 0 \quad as \quad i\to\infty.$$

Proof. From (17), (18) it follows

$$\|I-M_kD_k\|\leqslant\|I-M_kD_{k-1}\|^q,\qquad \|D_k\|\leqslant\|D_0\|\prod_{i=0}^{k-1}\frac{1-\sigma_i^q}{1-\sigma_i},$$

and from the definition of B_k and M_k one concludes that

$$\|(M_{k-1}-M_k)D_{k-1}\|\leqslant\|\{[F'(x_k)]^*(F'(x_{k-1})-F'(x_k))+$$
$$+([F'(x_{k-1})]^*-[F'(x_k)]^*)F'(x_{k-1})+(\alpha_{k-1}-\alpha_k)H\}D_{k-1}\|\leqslant$$
$$\leqslant(2\varepsilon_{k-1}\lambda_{k-1}KL_1\|P_{k-1}F(x_{k-1})\|+(m-n\xi)\|H\|)\|D_0\|\prod_{i=0}^{k-2}\frac{(1-\sigma_i^q)}{1-\sigma_i},$$
$$\|I-M_kD_{k-1}\|\leqslant\|I-M_{k-1}D_{k-1}\|+\|(M_{k-1}-M_k)D_{k-1}\|\leqslant\sigma_{k-2}^q+\Lambda\eta^{k-1}.$$

According to (12), one gets

$$\|P_{k+1}F(x_{k+1})\|\leqslant\delta_k\|P_kF(x_k)\|,$$

where

$$\delta_k=1-\varepsilon_k+\varepsilon_k[N+K\|U^{-1}\|(m\bar{C}^3\xi^k+\Omega KV)(\eta^q/\xi)^k]+\frac{1}{2}\varepsilon_k^2\lambda_k^2L_1\|P_kF(x_k)\|$$

and

$$\delta_k=1-\varepsilon_k+\varepsilon_kK\|U^{-1}\|(m\bar{C}^3\xi^k+\Omega KV)(\eta^q/\xi)^k+\frac{1}{2}\varepsilon_k^2\lambda_k^2L_1\|P_kF(x_k)\|$$

in the cases of A and B, respectively.

Setting

$$\bar{\lambda}=\bar{C}+\alpha_0\bar{C}+\Omega KV,$$
$$\bar{\gamma}_k=m\xi^k\bar{C}^3K\|U^{-1}\|+sl^kK^2V\|U^{-1}\|,$$

and making use of (21)—(24), one has

$$\mu_1=\sigma_0^q\leqslant\mu_0,\quad \sigma_0\leqslant\mu+\Lambda\leqslant h,\quad \omega_1=\mu_1/\alpha_1\leqslant h^q/n\xi\leqslant\Omega.$$

In the remainder the proof is similar to proofs corresponding to Theorems in [5, 8].

Remark 6. In a similar way as in ([14] Theorem 2) it can be shown that in the case $\varepsilon_k=1$ and $P_{R(x)}$ being independent of x the sequence $\{x_k\}$ generated by the formulas (14) and (17) converges quadratically to a solution of (2).

The author would like to thank Prof. G. Vainikko for very helpful discussions.

REFERENCES

1. *Eldén, L.* A note on weighted pseudoinverses with application to the regularization of Fredholm integral equations of the first kind. Report LiTH — MAT — R — 1975 — 11, Department of Mathematics, Linköping University (Sweden), 1975.
2. *Diaconu, A.* // Mathematicǎ (RSR), 1980, **22**, № 2, 247—261.
3. *Locker J., Prenter, P. M.* // SIAM J. Numer. Anal., 1980, **17**, 247—267.
4. *Schwetlick, H.* Numerische Lösung nichtlinear Gleichungen. Berlin, VEB Deutscher Verlag der Wissenschaften, 1979.
5. *Ваарманн О., Ломп М.* // Изв. АН ЭССР. Физ. Матем., 1982, **31**, № 4, 410—417.
6. *Vaarmann, O.* // Proc. Acad. Sci. ESSR. Phys. Math., 1978, **27**, № 3, 251—258.
7. *Красносельский М. А., Вайникко Г. М., Забрейко П. П., Рутицкий Я. Б., Стеценко В. Я.* Приближенное решение операторных уравнений. М., Наука, 1969.
8. *Ваарманн О.* // Изв. АН ЭССР. Физ. Матем., 1970, **19**, № 3, 265—274.
9. *Ortega, J. M., Rheinboldt, W. C.* Iterative Solution of Nonlinear Equations in Several Variables. New York; London, Academic Press, 1970.
10. *Boggs, P. T., Dennis, J. E.* jr. // Math. Comput., 1976, **30**, 199—215.
11. *Оганесян С. М., Старостенко В. И.* // Изв. АН СССР. Физ. Земли, 1978, № 1, 63—75.
12. *Старостенко В. И., Оганесян С. М.* // Методы решения некорректных задач и приложения (ред. А. Н. Тихонов). Нарва-Йыэсуу; Москва, 1982, 119—128.
13. *Marquardt, D. W.* // SIAM J. Appl. Math., 1963, **11**, 431—441.
14. *Vaarmann, O.* // Proc. Acad. Sci. ESSR. Phys. Math., 1980, **29**, № 3, 233—240.

Academy of Sciences of the Estonian SSR,
Institute of Cybernetics

Received
March 31, 1988

O. VAARMANN

MITTELINEAARSETE VÄHIMRUUTUDE ÜLESANNETE LAHENDAMINE LEVENBERG—MARQUARDTI TÜÜPI MEETODITEGA

Mittelineaarse operaatorvõrrandi üldistatud lahendi arvutamiseks on välja töötatud üks üldistatud pseudopöördoperaatori aproksimeerimisel baseeruv iteratsioonimeetodite pere, mis erijuhul sisaldab Gauss—Newtoni, Ben—Israeli, Levenberg—Marquardti meetodid, aga ka üldistatud pseudopöördoperaatori järkjärgulisel aproksimeerimisel põhinevad meetodid. Iteratsioonimeetodite koonduvusomaduste parandamiseks, sealhulgas koonduvuspiirkonna laiendamiseks, on kasutusele võetud iteratsioonisammu pikkust reguleerivad nn. relaksatsiooniparameetrid. On tõestatud koonduvusteoreemid selle iteratsioonimeetodite pere jaoks.

О. ВААРМАНН

РЕШЕНИЕ НЕЛИНЕЙНЫХ ЗАДАЧ НАИМЕНЬШИХ КВАДРАТОВ МЕТОДАМИ ТИПА ЛЕВЕНБЕРГА—МАРКВАРДТА

Понятие псевдообратного оператора позволяет укладывать в единые рамки различные итерационные методы. В случае гильбертова пространства рассматриваются методы решения задач определения обобщенного решения нелинейного операторного уравнения в терминах обобщенной псевдоинверсии. Доказаны теоремы о сходимости для одного семейства итерационных методов, включающих в себя как частные случаи методы Гаусса—Ньютона, Бен-Израэля, Левенберга—Марквардта, а также методы с последовательной аппроксимацией обобщенной псевдоинверсии. Установлена область допустимых значений релаксационного (шагового) параметра, позволяющего ослаблять ограничения на выбор начального приближения.