NOVI SAD J. MATH.
VOL. 29, NO. 1, 1999, 169-175

ON AN ITERATIVE METHOD WITH DEFLATION FOR SYSTEMS OF EQUATIONS

Dragoslav Herceg[1], **Nataša Krejić**[1], **Helena Maličić**[1]

Abstract. In [2] the recursive projection method based on deflation for solving nonlinear parameter dependent problems was presented. In this paper we introduce some new variants of this iterative method and prove their convergence.

AMS *Mathematics Subject Classification (1991):* 65B99, 65M12

Key words and Phrases: stabilization procedures, projection methods

1. Introduction

In this paper we consider solving of parameter dependent nonlinear systems of the form

(1) $$x = F(x, \lambda), \qquad F : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n.$$

One of the usual techniques for solving (1) is using "fixed-point" iterations

(2) $$x^{k+1} = F\left(x^k, \lambda\right), \qquad x^k = x^k(\lambda),$$

where we assume F is smooth and that the sequence $\left\{x^k(\lambda)\right\}$ converges to the solution $x^*(\lambda)$ for which it holds

$$x^*(\lambda) = F(x^*(\lambda), \lambda), \qquad \lambda \in [\lambda_a, \lambda_b].$$

Let Γ be a smooth arc of solutions $x^*(\lambda)$ parameterized by $\lambda \in [\lambda_a, \lambda_b]$. Varying the parameter λ, the convergence rate may change or the procedure may even diverge. In [2] the authors presented the recursive projection (RP) method as stabilization procedure for correcting this situation.

With $F_{x^*} = F_x(x^*(\lambda), \lambda)$ we shall denote Jacobian matrix for the function F and fixed value of the parameter λ, and let

$$\{\mu_k : \; k = 1, \ldots, n\}$$

be the eigenvalues of F_{x^*}. From the Ostrowski theorem [1], it is known that the iterations (2) converge locally in a neighborhood of the solution when all

[1]University of Novi Sad, Faculty of Science, Trg Dositeja Obradovića 4, 21000 Novi Sad, Yugoslavia, e-mail: {hercegd, natasa, helena}@unsim.ns.ac.yu

eigenvalues of the Jacobian matrix F_{x^*} belong to the unit disc $\{|z| < 1\}$. In that case

$$\rho(F_{x^*}) < 1,$$

where $\rho(A)$ represents spectral radius of a matrix A. More interesting case that we shall consider will be when for some $\delta > 0$ we have

$$|\mu_1| \geq \cdots \geq |\mu_m| > 1 - \delta \geq |\mu_{m+1}| \geq \cdots \geq |\mu_n|.$$

The basic idea of this method is to write space $\mathbb{R}^n$ as a direct sum of the span of the unstable eigenspace $\mathbb{P}$ and its orthogonal complement $\mathbb{Q}$. On the subspace $\mathbb{P}$ standard Newton's method will be performed and on the subspace $\mathbb{Q}$ "fixed-point" iterations will be used. It will be shown that the latter iterations do indeed locally converge even if the iterations (2) diverge. Here we shall also present some new variants of this method based on deflation for which convergence will be proved.

2. The recursive projection method

Using notation from [2], let $\mathbb{P}$ be the maximal invariant subspace of F_{x^*} belonging to $\{\mu_k : k = 1, 2, \ldots, m\}$, where

$$|\mu_1| \geq \cdots \geq |\mu_m| > 1 - \delta,$$

and let $\mathbb{Q}$ be the orthogonal complement of $\mathbb{P}$. If P and Q are orthogonal projections of $\mathbb{R}^n$ onto these subspaces, then every vector $x \in \mathbb{R}^n$ can be uniquely represented as

$$x = p + q, \qquad p = Px \in \mathbb{P}, \qquad q = Qx = (I - P)x \in \mathbb{Q}.$$

Now, applying this decomposition to the system (1), a new system is obtained

$$p = f(p, q, \lambda),$$

$$q = g(p, q, \lambda),$$

where

$$f(p, q, \lambda) = PF(p + q, \lambda),$$

$$g(p, q, \lambda) = QF(p + q, \lambda).$$

As mentioned before, using Newton's method on subspace $\mathbb{P}$ and "fixed-point" iterations on $\mathbb{Q}$, the recursive projection method is finally of the form

$$p^0 = Px^0(\lambda), \qquad q^0 = Qx^0(\lambda),$$

$$p^{k+1} = h(p^k, q^k, \lambda), \qquad k = 0, 1, \ldots, N - 1,$$

(3)

$$q^{k+1} = g(p^k, q^k, \lambda), \qquad k = 0, 1, \ldots, N - 1,$$

$$x^*(\lambda) = p^N + q^N = p^*(\lambda) + q^*(\lambda),$$

where

$$h(p,q,\lambda) = p + (I - f_p(p,q,\lambda))^{-1}(f(p,q,\lambda) - p),$$

$$f_p(p,q,\lambda) = PF_x(x,\lambda)P.$$

For the convenience, we wrote p^k and q^k instead of $p^k(\lambda)$ and $q^k(\lambda)$. The number N represents the first iteration which satisfies the termination criterion

$$\left\|x^N(\lambda) - F\left(x^N(\lambda),\lambda\right)\right\|_2 < \varepsilon, \qquad \varepsilon > 0.$$

Given two points $x_{-1} \equiv x^*(\lambda_{-1})$, $x_0 \equiv x^*(\lambda_0)$, $x_{-1}, x_0 \in \Gamma$, where $\lambda_{-1}, \lambda_0 \in [\lambda_a, \lambda_b]$, $\lambda_{-1} < \lambda_0$, a starting value $x^0(\lambda)$ is determined by the secant method

$$\lambda = \lambda_0 + \delta_\lambda,$$

$$x^0(\lambda) = x_0 + \frac{\delta_\lambda}{\lambda_0 - \lambda_{-1}}(x_0 - x_{-1}),$$

with the step size δ_λ.

It is important to mention that in (3), the vectors $x^k(\lambda)$ are

$$x^k(\lambda) = p^k(\lambda) + q^k(\lambda)$$

and they are different from the vectors formed by (2).

From the next lemma, the recursion

$$q^{k+1} = g\left(p^k, q^k, \lambda\right)$$

is locally convergent on $\mathbb{Q}$ in some neighborhood of (p^*, q^*).

Lemma 1. [2] *All the eigenvalues of*

$$g_{q^*} \equiv g_q(p^*, q^*, \lambda) = QF_{x^*}Q$$

lie in the disc $\{|z| < 1\}$.

Lemma 2. [2] *For the matrix* $g_{p^*} \equiv g_p(p^*, q^*, \lambda) = QF_{x^*}P$, *it holds*

$$g_{p^*} = 0.$$

For the method (3), if $f(p,q,\lambda)$ is smooth enough and $I - f_p(p^*, q^*, \lambda)$ is regular, the Jacobian J_J is

$$J_J = \begin{bmatrix} h_{p^*} & h_{q^*} \\ g_{p^*} & g_{q^*} \end{bmatrix} = \begin{bmatrix} 0 & (I - f_{p^*})^{-1} f_{q^*} \\ 0 & g_{q^*} \end{bmatrix},$$

since

$$h_{p^*} = I + (I - f_{p^*})^{-1}(f_{p^*} - I) = 0. \tag{4}$$

Finally, from Lemma 1, we have

$$\rho(J_J) = \rho(g_{q^*}) \le 1 - \delta < 1$$

and the convergence of the method (3) follows from the next theorem.

Theorem 1. [2] *For some* $\varepsilon_1 = \varepsilon_1(\lambda) > 0$, *let* $F(x, \lambda)$ *satisfy on the set* $\mathcal{B}_{\varepsilon_1}(x^*) = \{x : \|x - x^*\| < \varepsilon_1\}$:

a) $F(x, \lambda)$ *is twice differentiabile,*

b) $1 \notin \sigma(f_{p^*})$.

Then the iterations (3) *converge for all initial values* $x^0 \in \mathcal{B}_\varepsilon(x^*)$ *for some* $\varepsilon < \varepsilon_1$.

For implementation of the recursive projection method it is necessary to determine the projectors P and Q onto the subspaces $\mathbb{P}$ and $\mathbb{Q}$. This will be omitted here, but the detailed descripiton can be found in [2]. Let us just mention that the key role has an orthonormal basis Z for $\mathbb{P}$, since $P = ZZ^\top$, $Q = I - ZZ^\top$.

3. New variants of the RP method

The iterations (3) represent coupled iterations of Jacobi type, since the next iteration is formed using only the previous one. Introducing

$$\begin{aligned} p^{k+1} &= h(p^k, q^k, \lambda), \\ q^{k+1} &= g(p^{k+1}, q^k, \lambda), \end{aligned} \tag{5}$$

and

$$\begin{aligned} q^{k+1} &= g(p^k, q^k, \lambda), \\ p^{k+1} &= h(p^k, q^{k+1}, \lambda), \end{aligned} \tag{6}$$

a Gauss-Seidel (GS) and backward Gauss-Seidel (BGS) type of iterations is obtained. Similary,

$$\begin{aligned} p^{k+1} &= h(p^k, q^k, \lambda), \\ q^{k+1} &= g\left((1-\omega)p^k + \omega p^{k+1}, q^k, \lambda\right), \end{aligned} \tag{7}$$

and

$$\begin{aligned} q^{k+1} &= g(p^k, q^k, \lambda), \\ p^{k+1} &= h\left(p^k, (1-\omega)q^k + \omega q^{k+1}, \lambda\right), \end{aligned} \tag{8}$$

represent succesive overrelaxation (SOR) and backward SOR type of iterations respectively with the relaxation parameter ω. Of course, if in (7) or in (8) with $\omega = 0$ the coupled Jacobi iterations are obtained, and for $\omega = 1$ we get GS and BGS couplings, respectively. Let J_G, J_{BG}, J_S and J_{BS} be the corresponding Jacobians. Then

$$J_G = \begin{bmatrix} h_{p^*} & h_{q^*} \\ g_{p^*}h_{p^*} & g_{p^*}h_{q^*} + g_{q^*} \end{bmatrix}, \qquad J_{BG} = \begin{bmatrix} g_{q^*} & g_{p^*} \\ h_{q^*}g_{q^*} & h_{q^*}g_{p^*} + h_{p^*} \end{bmatrix};$$

$$J_S = \begin{bmatrix} h_{p^*} & h_{q^*} \\ g_{p^*}\left((1-\omega)I + \omega h_{p^*}\right) & \omega g_{p^*} h_{q^*} + g_{q^*} \end{bmatrix},$$

$$J_{BS} = \begin{bmatrix} g_{q^*} & g_{p^*} \\ h_{q^*}\left((1-\omega)I + \omega g_{q^*}\right) & \omega h_{q^*} g_{p^*} + h_{p^*} \end{bmatrix}.$$

Lemma 3.

$$\rho(J_G) = \rho(J_{BG}) = \rho(J_S) = \rho(J_{BS}) = \rho(g_{q^*}).$$

Proof. Using Lemma 2 and (4), the Jacobians are

$$J_G = \begin{bmatrix} 0 & h_{q^*} \\ 0 & g_{q^*} \end{bmatrix}, \qquad J_{BG} = \begin{bmatrix} g_{q^*} & 0 \\ h_{q^*} g_{q^*} & 0 \end{bmatrix},$$

$$J_S = \begin{bmatrix} 0 & h_{q^*} \\ 0 & g_{q^*} \end{bmatrix}, \qquad J_{BS} = \begin{bmatrix} g_{q^*} & 0 \\ h_{q^*}\left((1-\omega)I + \omega g_{q^*}\right) & 0 \end{bmatrix},$$

and the stated follows directly. □

The following theorem proves local convergence of the new variants (5), (6), (7) and (8) of the recursive projection method.

Theorem 2. *For some* $\varepsilon_1 = \varepsilon_1(\lambda) > 0$, *let* $F(x, \lambda)$ *satisfy on the set* $\mathcal{B}_{\varepsilon_1}(x^*)$:

a) $F(x, \lambda)$ *is twice differentiabile,*

b) $1 \notin \sigma(f_{p^*})$.

Then the iterations (5), (6), (7) *and* (8) *converge for all initial values* $x^0 \in \mathcal{B}_\varepsilon(x^*)$ *for some* $\varepsilon < \varepsilon_1$.

Proof. Let $\{v^k\}$ be the sequence from $\mathbb{R}^{2n}$ defined with

$$v^k = \begin{bmatrix} p^k \\ q^k \end{bmatrix}, \qquad k = 0, 1, \dots,$$

where $\{p^k\}$ and $\{q^k\}$ are the iterates from the methods (5), (6), (7) and (8). If we show that for some norm $\|\cdot\|$ there is a constant $C < 1$ for which

$$\left\|v^{k+1} - v^*\right\| \leqslant C \left\|v^k - v^*\right\|, \qquad k = 0, 1, \dots, \tag{9}$$

where

$$v^* = \begin{bmatrix} p^* \\ q^* \end{bmatrix},$$

then the convergence of the sequence $\{x^k\}$ will follow directly from the convergence of the sequences $\{p^k\}$ and $\{q^k\}$. Firstly, using induction it can be easily shown

$$x^k = p^k + q^k \in \mathcal{B}_\varepsilon(x^*), \qquad \varepsilon < \varepsilon_1,$$

which ensures that v^{k+1} is well defined. In general, a Taylor expansion about v^* yields

$$v^{k+1} - v^* = J(p^*, q^*)(v^k - v^*) + \mathcal{O}\left(\left\|v^k - v^*\right\|^2\right),$$

where $J(p^*, q^*)$ represents the Jacobian of the new methods. For the matrix $J = J(p^*, q^*)$ and for any $\eta > 0$ there exists a norm $\|\cdot\|_{J,\eta}$ such that

$$\|J(p^*, q^*)\|_{J,\eta} \leqslant \rho(J(p^*, q^*)) + \eta.$$

Now using Lemmas 1 and 4, we have

$$\begin{aligned}\left\|v^{k+1} - v^*\right\|_{J,\eta} &\leq (1 - \delta + \eta)\left\|v^k - v^*\right\|_{J,\eta} + \mathcal{O}\left(\left\|v^k - v^*\right\|_{J,\eta}^2\right)\\ &\leq \left(1 - \delta + \eta + \mathcal{O}\left(\left\|v^k - v^*\right\|_{J,\eta}\right)\right)\left\|v^k - v^*\right\|_{J,\eta}.\end{aligned}$$

Specially, for $k = 0$

$$\left\|v^1 - v^*\right\|_{J,\eta} \leq \left(1 - \delta + \eta + \mathcal{O}\left(\left\|v^0 - v^*\right\|_{J,\eta}\right)\right)\left\|v^0 - v^*\right\|_{J,\eta}.$$

From the condition $x^0 \in \mathcal{B}_\varepsilon(x^*)$, $\varepsilon < \varepsilon_1$, it holds

$$\begin{aligned}\left\|v^0 - v^*\right\|_{J,\eta} &= \left\|\begin{bmatrix} p^0 - p^* \\ q^0 - q^* \end{bmatrix}\right\|_{J,\eta} = \left\|\begin{bmatrix} P(x^0 - x^*) \\ Q(x^0 - x^*) \end{bmatrix}\right\|_{J,\eta}\\ &\leq \left\|\begin{bmatrix} ZZ^\top \\ I - ZZ^\top \end{bmatrix}\right\|_{J,\eta} \cdot \left\|x^0 - x^*\right\|_{J,\eta}\\ &\leq M\varepsilon,\end{aligned}$$

where $M > 0$. Finally

$$\left\|v^1 - v^*\right\|_{J,\eta} \leq (1 - \delta + \eta + \mathcal{O}(\varepsilon))\left\|v^0 - v^*\right\|_{J,\eta},$$

and the constant C can be chosen as

$$C = 1 - \delta + \eta + \mathcal{O}(\varepsilon),$$

because for the small enough values η and ε it holds $C < 1$. The convergence of the sequence $\{v^k\}$ follows from the inductive hypothesis

$$\left\|v^k - v^*\right\|_{J,\eta} \leq C\left\|v^{k-1} - v^*\right\|_{J,\eta}$$

and

$$\left\| v^k - v^* \right\|_{J,\eta} \leq C^k \left\| v^0 - v^* \right\|_{J,\eta} .$$

The solution of the problem (1) is the vector $x^* = p^* + q^*$, where p^* and q^* are the limits of the convergent sequences $\{p^k\}$ and $\{q^k\}$. □

References

[1] Ortega, J.M., Rheinboldt, W.C., Iterative Solution of Nonlinear Equations in Several Variables, First ed., New York London, Academic Press, 1970.

[2] Shroff, G., Keller, H.B., Stabilization of unstable procedures: The recursive projection method, SIAM J. Numer. Anal., 30 (1993), 1099-1120.

[3] Varga, R.S., Matrix Iterative Analysis, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962.

Received by the editors December 3, 1998.