A note on defect correction processes with an approximate inverse of deficient rank

P. W. Hemker *

ABSTRACT

In view of the analysis of multiple grid methods, in this note we consider defect correction processes of deficient rank. Both for the error and for the residual, the convergence of the defect iterative process is studied in terms of the range and the kernel of the approximate inverse. Since the coarse grid correction in the multiple grid algorithm can be seen as a step in such an iterative process, the present study can be used in the convergence analysis of these algorithms. It shows that pre-relaxation is advantageous for reducing the error, whereas post-relaxation is better for reducing the residual.

1. THE DEFECT CORRECTION PROCESS

In order to solve the operator equation

$$Fx = y, \tag{1}$$

$F : B_1 \to B_2$, B_1, B_2 Banach spaces, we consider the defect correction iterative process

$$\begin{cases} x_0 = \tilde{G}y, \quad y \in B_2, \\ x_{i+1} = x_i - \tilde{G}Fx_i + \tilde{G}y. \end{cases} \tag{2}$$

The process is determined by the operator $\tilde{G} : B_2 \to B_1$, which is called the *approximate inverse* of F.
In this paper we consider only linear operators F and $\tilde{G}$.
We notice that the process (2) converges to the solution x^* of (1) if $\tilde{G}$ is injective and

$$\|I - \tilde{G}F\|_{B_1 \to B_1} < 1.$$

The value $e_i = x_i - x^*$ is called the *error* of x_i; and the operator

$$M = I - \tilde{G}F$$

the transistion matrix, or is the *amplification operator of the error,* since

$$e_{i+1} = Me_i.$$

We notice also that, due to the linearity of $\tilde{G}$, the process (2) is equivalent with

$$\begin{cases} \ell_0 = y, \\ \ell_{i+1} = \ell_i - F\tilde{G}\ell_i + y, \end{cases} \tag{3}$$

when x_i is identified with

$$x_i = \tilde{G}\ell_i.$$

The process (3) converges to the solution of (1) if

$$\|I - F\tilde{G}\|_{B_2 \to B_2} < 1;$$

the value $r_i = y - Fx_i$ is called the *residual* of x_i and the operator

$$\hat{M} = I - F\tilde{G}$$

is the *amplification operator of the residual* since $r_{i+1} = \hat{M}r_i$.
In particular we shall here consider the processes (2) and (3) where F and $\tilde{G}$ are operators $\mathbb{R}^n \to \mathbb{R}^n$, where F is a full rank matrix, such that the original problem (1) has a unique solution, and $\tilde{G}$ is of deficient rank, i.e. $\tilde{G}$ is neither injective nor surjective.
Because rank $(\tilde{G}) = k < n$, we know that $N = \text{Range}(\tilde{G})$ is a k-dimensional subspace of $\mathbb{R}^n$ and $Z = \text{Kernel}(\tilde{G})$ is a (n-k)-dimensional subspace of $\mathbb{R}^n$.
In order to define orthonormal bases in N and Z, we can decompose the n*n matrix G into its singular value decomposition [4] :

$$\tilde{G} = U \Sigma V^T,$$

where U, Σ and V are n*n matrices, U and V are orthonormal and Σ is a nonnegative diagonal matrix. Except for the ordering of the elements of Σ, this decomposition is uniquely determined. The diagonal elements of Σ are the singular values and normally they are ordered such that

$$\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_n \geqslant 0.$$

Because rank $(\tilde{G}) = k$, we know that $\sigma_1, \sigma_2, \ldots, \sigma_k$ are non-zero and $\sigma_j = 0$, $j = k+1, \ldots, n$.

* P. W. Hemker, Stichting Mathematisch Centrum, Postbus 4079, 1009 AB Amsterdam, The Netherlands.

Since $\tilde{G}$ is not surjective, possibly $x^* \notin N$; however, from (2) we see that all approximate solutions $\{x_i\}$ are in N. Hence, if $\{x_i\}_{i=0,1,\ldots}$ attains a stationary point $\hat{x}$, $\hat{x}$ is not necessarily the solution of (1). However, we know

$$\tilde{G}(y - F\hat{x}) = 0,$$

i.e. the residual $\hat{r} = y - F\hat{x} \in Z$. Thus with $\Delta = V_1 V_1^T$ the projection $\mathbb{R}^n \to Z^\perp$, instead of the sequence $\{\ell_i\}$ in (3) we may consider the sequence $\{\lambda_i\}$, with $\lambda_i = \Delta\ell_i$:

$$\begin{cases} \lambda_0 = \Delta y, \\ \lambda_{i+1} = \lambda_i - \Delta F\tilde{G}\lambda_i + \Delta y, \end{cases}$$

which has a unique stationary point $\hat{\lambda}$, satisfying

$$\Delta F\tilde{G}\hat{\lambda} = \Delta y.$$

Clearly, $N = \text{Span}(U_1)$, where U_1 are the first k column vectors of U and $N^\perp = \text{Span}(U_2)$, the last n-k columns of U. Analogously, $Z = \text{Span}(V_2)$ and $Z^\perp = \text{Span}(V_1)$.

From the singular value decomposition we easily see that for an arbitrary $P : \mathbb{R}^k \to \mathbb{R}^n$ and $R : \mathbb{R}^n \to \mathbb{R}^k$, with range (P) = N and Kernel (R) = Z, we may write

$$\tilde{G} = PSR,$$

where $S : \mathbb{R}^k \to \mathbb{R}^k$ is the nonsingular k*k matrix for which

$$S^{-1} = (RV_1)\,\text{diag}\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots, \frac{1}{\sigma_k}\right)(U_1^T P).$$

The operators P and R are called *prolongation* and *restriction* respectively.

Because P and R are full rank matrices, rank (P) = rank (R) = k, P has the left-inverse $\hat{R} = (U_1^T P)^{-1} U_1^T$ and R has the right-inverse $\hat{P} = V_1 (RV_1)^{-1}$. Moreover, we know that

$$P\hat{R} = P(U_1^T P)^{-1} U_1^T : \mathbb{R}^n \to N,$$

and

$$\hat{P}R = V_1 (RV_1)^{-1} R : \mathbb{R}^n \to Z^\perp,$$

are projection operators.

Now we can consider what happens to the error to the solution or to the residual after one iteration step of the defect correction process.

1.1. The error in the solution

To study the effect on the error of the solution, we consider (2), of which the transition matrix is

$$M = I - \tilde{G}F = I - PSRF.$$

We decompose the error e into two parts : $e = e_s + e_u$, with $e_s \in N$ and $e_u \in N^\perp$. Analogously, we write

$$Me = (Me)_s + (Me)_u,$$

with $(Me)_s \in N$ and $(Me)_u \in N^\perp$.

From the relation

$$Me_s = MP\hat{R}e_s = (P\hat{R} - PSRFP\hat{R})e_s = P(I - SRFP)\hat{R}e_s,$$

we see that $Me_s \in N$. Moreover, we notice that in the special case when $S^{-1} = RFP$, we have $Me_s = 0$. In the general case, with $S^{-1} = RFP + E$ we have

$$Me_s = PSE\hat{R}e_s = \tilde{G}\hat{P}E\hat{R}e_s.$$

The contribution from e_u to Me is given by

$$Me_u = e_u - \tilde{G}Fe_u,$$

with $\tilde{G}Fe_u \in N$ and $e_u \in N^\perp$.

We conclude that

$$\begin{cases} (Me)_s = \tilde{G}\hat{P}E\hat{R}e_s - \tilde{G}Fe_u, \\ (Me)_u = e_u. \end{cases} \tag{4}$$

1.2. The residual

For the residual, the transition matrix is

$$\hat{M} = I - F\tilde{G} = I - FPSR.$$

Now we decompose the residual r into two parts $r = r_s + r_u$, with $r_u \in Z$ and $r_s \in Z^\perp$. Analogously we write

$$\hat{M}r = (\hat{M}r)_s + (\hat{M}r)_u.$$

Again, a simple computation shows

$$\begin{cases} (\hat{M}r)_s = \hat{P}E\hat{R}\tilde{G}r_s, \\ (\hat{M}r)_u = -(I - \hat{P}R)F\tilde{G}r_s + r_u. \end{cases} \tag{5}$$

1.3. Summary

Summarizing the effect of one iteration step in a defect correction process with an approximate inverse of deficient rank, we find the following transitions in a single iteration step of the form (2).

For the error in the solution :

$$\begin{array}{lll} \text{Smooth components} & = \text{Range}(P) = N & \xrightarrow{\tilde{G}\hat{P}E\hat{R}} N \\ & & \quad \searrow^{\tilde{G}F} \\ \text{Unsmooth components} & = \text{Kernel}(\hat{R}) = N^\perp & \xrightarrow[I]{} N^\perp \end{array}$$

For the residual :

$$\begin{array}{lll} \text{Smooth components} & = \text{Range}(\hat{P}) = Z^\perp & \xrightarrow{\hat{P}E\hat{R}\tilde{G}} Z^\perp \\ & & \quad \searrow^{(\hat{P}R - I)F\tilde{G}} \\ \text{Unsmooth components} & = \text{Kernel}(R) = Z & \xrightarrow[I]{} Z. \end{array}$$

We note that in the special case when $R = P^T$ we have

$$N = \text{Range}(P) = \text{Span}(U_1) = \text{Span}(V_1) = Z^\perp,$$

$$Z = \text{Kernel}(R) = \text{Span}(U_2) = \text{Span}(V_2) = N^\perp.$$

In this case the subspace of the smooth (resp. unsmooth) components of the residual is the same as the

subspace of the smooth (unsmooth) components in the error.

2. APPLICATION TO THE MULTIGRID METHOD

We want to apply the above results for the explanation of some phenomena in multiple grid methods. The multiple grid method (see [1,2,3]) is an iterative method for the solution of a linear system

$$A_h x_h = f_h,$$

arising from the discretization of an elliptic partial differential equation. The multi-grid method is an iterative method which consists of relaxation steps and coarse-grid correction (CGC) steps. It is well known that relaxation steps are efficient for the reduction of *non-smooth* components in the error or in the residual. In the multi-grid method the CGC is used to reduce the *smooth* components efficiently.
The CGC-step can be written in the form (2), with

$$F = A_h$$

and

$$\tilde{G} = PSR = P_{hH} A_H^{-1} R_{Hh},$$

where A_h is the fine-grid discretization of the continuous operator, A_H is its coarse grid discretization; R_{Hh} is the restriction of a fine-grid function to a coarse grid and P_{hH} is the prolongation (interpolation) of a coarse grid function to a fine grid.
For A_H any convergent discretization of the continuous problem can be used. However, one particularly efficient choice is

$$A_H = R_{Hh} A_h P_{hH}, \tag{6}$$

the *Galerkin approximation* of A_h on the coarse grid.
This choice corresponds with $S^{-1} = RFP$, i.e. $E = 0$ in the discussion in the previous section.
The treatment in section 1 holds for arbitrary P and R. In the context of a multi-grid method we choose $P = P_{hH}$ and $R = R_{Hh}$. This implies that the components in N are those grid functions in the fine grid that can be obtained by prolongation from a coarse grid function; therefore they are the *smooth components of the error*. Those in $N^{\perp}$ are the *unsmooth components of the error.*
Similarly, we find in the right-hand-side space B_2 that the components in Z are those grid function on the fine grid that vanish by restriction to the coarse grid and therefore they are the *unsmooth components of the residual*; those in $Z^{\perp}$ are the *smooth components of the residual.*
Application of the results in (4) and (5) to the CGC of the multigrid method shows that

(1) In the case of the Galerkin approximation (6), smooth errors in the solution (and smooth residuals) do not give rise to new smooth errors (resp. residuals) after a CGC step. If A_H is not the Galerkin approximation, the transfer from smooth to smooth components is proportional to the deviation of the Galerkin approximation, i.e.
$E = A_H - R_{Hh} A_h P_{hH}$.
(2) Unsmooth components in the error or in the residual before a CGC-step give rise to the same unsmooth components after the CGC-step.
(3) Since smooth components in the error don't induce unsmooth components in the error, but unsmooth components induce smooth components, reduction of the unsmooth component before a CGC is more useful than after a CGC. Hence, to obtain a small error in a multi-grid cycle pre-relaxation is preferred.
(4) Since unsmooth components in the residual don't induce smooth components, but smooth components do induce unsmooth components in the residual, post-relaxation is preferred in a multi-grid cycle to obtain a small residual.

REFERENCES

1. BRANDT A. : 'Multi-level adaptive solutions to boundary value problems'. Math. Comp. 31 (1977) 333-390.
2. HACKBUSCH W. : 'On the multi-grid method applied to difference equations'. Computing 20 (1978) 291-306.
3. HEMKER P. W. : 'Introduction to multigrid methods'. Nw. Arch. Wisk. 29 (1981) 71-101.
4. LAWSON C. L. & HANSON R. J. : *Solving least squares problems,* Prentice-Hall Inc., N. J., 1974.
5. STETTER H. J. : 'The defect correction principle and discretization methods', Num.,Math. 29.(1978) 425-443.