Rüdiger Weiss

# Transformations of and Relations between Iterative Linear Solvers

*A general overview on the tranformations of iterative linear solvers by smoothing is given. Relations between different methods and convergence estimates are shown. In particular the same bounds are given for the residuals of Schönauer's smoothing applied to biconjugate gradients as for Freund's and Nachtigal's QMR method.*

## 1. Background

Smoothing of iterative methods has attracted considerable attention recently; see for example [1,4,6,8,9]. Let $x_k$ be an approximation resulting from an iterative method for the solution of the linear system $Ax = b$ and let $r_k = Ax_k - b$ be the residual. The sequence $s_0 = r_0$ , $z_0 = x_0$ ,

$$s_k = s_{k-1} + \gamma_k(r_k - s_{k-1}),\ z_k = z_{k-1} + \gamma_k(x_k - z_{k-1}) \tag{1}$$

is a corresponding *smoothed method* delivering approximations $z_k$ and residuals $s_k$. The smoothed sequence is a mean of the actual element of the original sequence and the previous element of the smoothed sequence. Different smoothing techniques vary by the determination of the coefficients $\gamma_k$.

In [8] smoothing was investigated for basic methods with orthogonal residuals. In general, smoothing transforms methods. For example ORTHORES is transformed to GMRES and CGNE to CGNR by smoothing. The further main results in [8] are briefly summarized. Let $Z$ be a positive definite matrix and $||y||_Z = \sqrt{y^T Zy}$ for any vector $y$. If $r_k$ is $Z$-orthogonal to $r_i$ for $i = 0, \ldots, k-1$, i. e. $r_k^T Z r_i = 0$, then the following equations are equivalent

$$||s_k||_Z = min_{\gamma_k} ||s_{k-1} + \gamma_k(r_k - s_{k-1})||_Z \ , \tag{2}$$

$$||s_k||_Z = min_{\gamma_1,\ldots,\gamma_k} ||s_{k-1} + \sum_{i=1}^{k} \gamma_i(r_i - s_{i-1})||_Z \ , \tag{3}$$

$$s_k = \frac{1}{\sum_{i=0}^{k} \frac{1}{||r_i||_Z^2}} \sum_{i=0}^{k} \frac{1}{||r_i||_Z^2} r_i \ , \tag{4}$$

$$\frac{1}{||s_k||_Z^2} = \sum_{i=0}^{k} \frac{1}{||r_i||_Z^2} \ . \tag{5}$$

The meaning of (3) is that the smoothing technique is equivalent to a k-dimensional minimization, though only a 1-dimensional minimization is performed in each step resulting from (2). By (5) we also get the norm of the unique minimizer:

$$||s_k||_Z \quad = \quad \frac{1}{\sqrt{\sum_{i=0}^{k} \frac{1}{||r_i||_Z^2}}} \ . \tag{6}$$

If the residuals of the original sequence are in a Krylov space, i. e. $r_k = \sum_{i=1}^{k} \alpha_i A^i r_0 + r_0$, then

$$||s_k||_Z \quad = \quad \min_{\eta_1,\ldots,\eta_k} || \sum_{i=1}^{k} \eta_i A^i r_0 + r_0 ||_Z \ . \tag{7}$$

## 2. Smoothing applied to non-orthogonal residuals

We will consider only the case $Z = I$ because it is the most important from a practical viewpoint and the extension to arbitrary matrices $Z$ is quite simple. If the $r_i$ are non-orthogonal, then the equations (2) – (5) are not equivalent and smoothing can be applied by using (2) or (4) for determining $\gamma_k$.

If we choose (2), then the technique is called minimal residual (MR) smoothing. This is the original algorithm of Schönauer [6]. Let us denote by $s^{MR}$ the residual of such a method. We get

$$\gamma_k^{MR} = -\frac{\left(s_{k-1}^{MR}\right)^T (r_k - s_{k-1}^{MR})}{||r_k - s_{k-1}^{MR}||^2} \ . \tag{8}$$

The residual is a function of the iteration index, with a non-increasing norm, i. e.

$$||s_k^{MR}|| \leq ||s_{k-1}^{MR}|| \ , \ and \ ||s_k^{MR}|| \leq ||r_k|| \ . \tag{9}$$

This sort of smoothing is locally optimal, because the residual is minimized in each step. However, very little has been known about the global behavior. We can prove the following estimate for the convergence that is the same as for QMR.

Theorem

$$||s_k^{MR}|| \leq \sqrt{k+1}\ \frac{1}{\sqrt{\sum_{i=0}^{k}\frac{1}{||r_i||^2}}}\ . \tag{10}$$

If the original sequence is generated by biconjugate gradients, i. e. $x_k = x_0 + V_k\bar{z}$, $r_k = r_0 + AV_k\bar{z}$, with $\bar{z} \in I\!R^k$ and $V_k$ is the matrix matrix resulting from the non-symmetric Lanczos process with $AV_k = V_{k+1}H_k$, where $H_k \in I\!R^{(k+1)\times k}$ is the tridiagonal matrix or a block tridiagonal matrix, then

$$||s_k^{MR}|| \leq \sqrt{k+1}\ \min_{z\in I\!R^k} ||\ ||r_0||e_1 - H_k z||\ , \tag{11}$$

where $e_1$ is the first unit vector.

Proof. By (9)

$$||s_k^{MR}|| \leq \min_{i=0,\ldots,k} ||r_i|| = \frac{1}{\sum_{i=0}^{k}\frac{1}{||r_i||^2}} \sum_{i=0}^{k} \frac{1}{||r_i||^2} \min_{i=0,\ldots,k} ||r_i|| \leq \frac{1}{\sum_{i=0}^{k}\frac{1}{||r_i||^2}} \sum_{i=0}^{k} \frac{1}{||r_i||^2}||r_i||$$

$$= \frac{1}{\sum_{i=0}^{k}\frac{1}{||r_i||^2}} \sum_{i=0}^{k} \frac{1}{||r_i||} \leq \frac{1}{\sum_{i=0}^{k}\frac{1}{||r_i||^2}}\sqrt{k+1}\sqrt{\sum_{i=0}^{k}\frac{1}{||r_i||^2}} = \sqrt{k+1}\ \frac{1}{\sqrt{\sum_{i=0}^{k}\frac{1}{||r_i||^2}}}\ .$$

The equivalence of the right-hand sides of (10) and (11) for the biconjugate gradients is proven in [9].

Note that also the minimum of the original residuals fulfils (10) and (11) resulting from the proof of the Theorem. Thus this minimum would be a serious smoother, but worse than minimal residual smoothing.

The idea of quasi-minimal residual smoothing [9] is to satisfy (4). This can be realized as follows. Choose $\tau_0^2 = ||r_0||^2$ and calculate for $k \geq 1$ $\tau_k^2 = \frac{1}{\frac{1}{\tau_{k-1}^2}+\frac{1}{||r_k||^2}}$, $\gamma_k^{QMR} = \frac{\tau_k^2}{||r_k||^2}$. It is obvious that

$$0 \leq \gamma_k^{QMR} \leq 1\ . \tag{12}$$

Zhou and Walker [9] show that quasi-minimal residual smoothing transforms biconjugate gradients [5] into QMR [3] and an auxiliary sequence based on CGS [7] into TFQMR [2] For the residual $s_k^{QMR}$ resulting from the biconjugate gradients and CGS as basic methods holds [9]

$$||s_k^{QMR}|| \leq \sqrt{k+1}\ \frac{1}{\sqrt{\sum_{i=0}^{k}\frac{1}{||r_i||^2}}}\ . \tag{13}$$

From the original paper on QMR [3] follows:

$$||s_k^{QMR}|| \leq \sqrt{k+1}\ \min_{z\in I\!R^k} ||\ ||r_0||e_1 - H_k z||. \tag{14}$$

Note the equivalence of (13) and (14) with the Theorem. As a result the global behavior of quasi-minimal residual smoothing is known, but still very little is known theoretically about the local behavior, except heuristics.

### 3. References


1 C. Brezinski and M. Redivo-Zaglia. Hybrid procedures for solving linear systems. *Numer. Math.*, 67:1–19, 1994.

2 R.W. Freund. A transpose-free quasi-minimal residual algorithm for non-Hermitian linear systems. *SIAM J. Sci. Comput.*, 14(2):470–482, 1993.

3 R.W. Freund and N.M. Nachtigal. QMR: a quasi-minimal residual method for non-Hermitian linear systems. *Numer. Math.*, 60:315–339, 1991.

4 M.H. Gutknecht. Changing the norm in conjugate gradient type algorithms. *SIAM J. Numer. Anal.*, 30(1):40–56, 1993.

5 C. Lanczos. Solution of systems of linear equations by minimized iterations. *J. Res. Nat. Bur. Standards*, 49:33–53, 1952.

6 W. Schönauer. *Scientific Computing on Vector Computers.* North-Holland, Amsterdam, New York, Oxford, Tokyo, 1987.

7 P. Sonneveld. CGS, a fast Lanczos-type solver for nonsymmetric linear systems. *SIAM J. Sci. Stat. Comp.*, 10(1):36–52, 1989.

8 R. Weiss. Convergence behavior of generalized conjugate gradient methods. Internal report 43/90, University of Karlsruhe, Computing Center, 1990.

9 L. Zhou and H.F. Walker. Residual smoothing techniques for iterative methods. *SIAM J. Sci. Comput.*, 15(2):297–312, 1994.



*Addresses:* Dr. Rüdiger Weiss, Rechenzentrum Universität Karlsruhe, Postfach 6980, D-76128 Karlsruhe, Germany.