Duality between Solvers for Ordinary Differential Equations and Linear Systems

Rüdiger Weiss
Numerikforschung für Supercomputer, Rechenzentrum Universität Karlsruhe
Postfach 6980, 76128 Karlsruhe, Germany
e-mail: weiss@rz.uni-karlsruhe.de

Keywords: linear systems, ordinary differential equations, iterative solvers

ABSTRACT

The successive approximations of an iterative linear solver can be considered as time-dependent variables of a dynamic system. In this aspect the iteration index plays the role of a virtual time. In this article finite difference approximations of an ordinary differential equation are identified with linear solver iterates. The solution of the linear system is achieved when time tends to infinity. This duality between ordinary differential equations and iterative solution methods for linear systems is investigated.

By the duality large time steps correspond to a fast convergence of the linear solver while small time steps lead to stagnation. It is well known that explicit finite difference approximations are limited in the time step. Implicit or semi-implicit finite difference approximations allow larger time steps. In particular, mixed implicit-explicit finite difference approximations are considered that lead to new iterative linear solvers.

SOLVERS FOR ORDINARY DIFFERENTIAL EQUATIONS

Let us consider the initial value problem for the system of ordinary differential equations

$$\frac{dx}{dt} = Ax - b, \qquad x(t_0) = x_0, \tag{1}$$

where $x, b \in I\!R^n$ and $A \in I\!R^{n\times n}$ is a real, square and non singular matrix of dimension n. The solution of (1) is formally given by

$$x(t) = \mathrm{e}^{A(t-t_0)}x_0 + \left(I - \mathrm{e}^{A(t-t_0)}\right)A^{-1}b. \tag{2}$$

Clearly, if A is similar to a diagonal matrix and the real parts of all eigenvalues are negative (positive), then for $t \to \infty$ $(-\infty)$ the solution converges to the limit

$$\bar{x} = \lim_{t\to\pm\infty} x(t) = A^{-1}b. \tag{3}$$

Thus, the solution of (1) converges to the solution of the linear system

$$A\bar{x} = b. \tag{4}$$

Let x_k be an approximation for $x(t_k)$ and the residual $r_{k-1} = Ax_{k-1} - b$. An explicit finite difference approximation for the differential equation (1) is

$$\sum_{i=1}^{\sigma_k} \gamma_{i,k}\frac{x_k - x_{k-i}}{t_k - t_{k-i}} = Ax_{k-1} - b = r_{k-1} \tag{5}$$

with

$$\sum_{i=1}^{\sigma_k} \gamma_{i,k} = 1. \tag{6}$$

The approximation x_k for the solution of (1) at the time t_k is obtained from (5) by

$$x_k \;=\; \alpha_1 r_{k-1} + \sum_{i=2}^{\sigma_k} \alpha_i (x_{k-i} - x_{k-1}) + x_{k-1} \tag{7}$$

with

$$\alpha_1 = \left(\sum_{i=1}^{\sigma_k} \gamma_{i,k} \frac{1}{t_k - t_{k-i}} \right)^{-1}, \qquad \alpha_j = \alpha_1 \gamma_{j,k} \frac{1}{t_k - t_{k-j}} \tag{8}$$

for $j = 2, \ldots, \sigma_k$.

SOLVERS FOR LINEAR SYSTEMS

Krylov subspace methods have been quite successfully applied for the solution of large linear systems. Equation (7) describes precisely the recursion for such a method like GMRES, BCG or QMR. Thus, we can consider these iterative linear solvers as explicit finite difference approximations for (1). The solution of the linear system (4) is obtained for $t = \pm\infty$.

Small time steps lead to a slow convergence while larger time steps lead to a fast convergence of the methods. If these considerations hold, then different linear solvers achieve solutions of the same quality for the same virtual time. Let us reconstruct the virtual time from the iteration recursion. From the second equation in (8) we obtain

$$\gamma_{j,k} = \frac{\alpha_j}{\alpha_1} (t_k - t_{k-j}) \tag{9}$$

for $j = 2, \ldots, \sigma_k$. From the first equation in (8) we obtain

$$\begin{aligned} \alpha_1^{-1} &= \gamma_{1,k} \frac{1}{t_k - t_{k-1}} + \sum_{i=2}^{\sigma_k} \gamma_{i,k} \frac{1}{t_k - t_{k-i}} \\ &\overset{(6)}{=} \frac{1}{t_k - t_{k-1}} + \sum_{i=2}^{\sigma_k} \gamma_{i,k} \left(\frac{1}{t_k - t_{k-i}} - \frac{1}{t_k - t_{k-1}} \right). \\ &\overset{(9)}{=} \frac{1}{t_k - t_{k-1}} + \sum_{i=2}^{\sigma_k} \frac{\alpha_i}{\alpha_1} \frac{t_{k-i} - t_{k-1}}{t_k - t_{k-1}} \end{aligned}$$

Thus, if approximations of the linear system are given according to (7), then the virtual time of the differential equation can be calculated step by step from

$$t_k = \alpha_1 + \sum_{i=2}^{\sigma_k} \alpha_i (t_{k-i} - t_{k-1}) + t_{k-1}. \tag{10}$$

Note the similarity with (7). If there would be no discretization error for the finite difference approximation of the partial differential equation, then any associated iterative linear solver should deliver the same solution for the same virtual time. As there is a discretization error this cannot be expected. However, the approximations of different linear solvers are expected to be closer if they are considered as function of the virtual time than as functions of the number of iterations.

Let us consider a rough model of the 3-dimensional Navier-Stokes equations. Let be $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)^T$, $\Delta = \nabla^T \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. We will solve the following partial differential equation (PDE) for the velocity $v = (v_1, v_2, v_3)^T$,

$$\Delta v + v + R(v^T \nabla) v = h \; , \tag{11}$$

with Dirichlet boundary conditions on the unit cube. The parameter R simulates a Reynolds number. The right-hand side $h = (h_1, h_2, h_3)^T$ is determined so that equation (11) has as exact solution

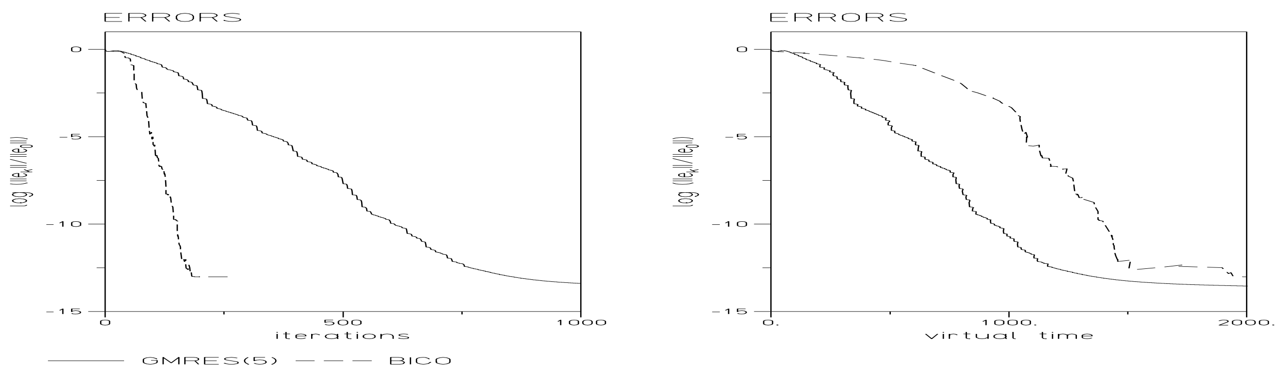

Figure 1: Relative errors for the sample problem R=100

trigonometric functions. The linear system arises from a finite difference discretization with consistency order 4 and from the linearization in the first Newton step. It is generated by the FIDISOL program package [1]. The calculations have been performed on a $20 \times 20 \times 20$ grid. The dimension of the linear system is $n = 24000$.
We apply the restarted GMRES(5) method and the residual smoothed biconjugate gradients BICO (see e. g. [2]) to this linear system with the Reynolds number $R = 100$. In figure 1 the convergence history is depicted. In the left part the relative norm of the errors is plotted over the iteration, in the right part it is plotted over the virtual time. The curves are quite different for both methods if we consider them in dependence of the iterations. If we consider the dependence of the virtual time the curves are closer. We also see that approximately 1000 iterations correspond to a virtual time of 2000 for GMRES(5). Thus, we have a relatively large average time step of 2 producing large discretization errors. Therefore, the two curves are still quite different in dependence of the virtual time. This test confirms the dual viewpoint for iterative linear solvers because the curves of the time plot are closer, but it also shows its limits because the curves are not close enough.

EXTRAPOLATION

Let us assume that k iterations of a generalized CG method have been performed according to equation (7). It is close at hand to use the duality to ordinary differential equations to try to extrapolate $\bar{x}$. Substituting $x(t_k)$ by the approximation x_k in equation (2) leads to

$$x_k = -\left(I - \mathrm{e}^{A(t_k - t_0)}\right) x_0 + \left(I - \mathrm{e}^{A(t_k - t_0)}\right) \bar{x} + x_0. \tag{12}$$

Thus,

$$\bar{x} = \left(I - \mathrm{e}^{A(t_k - t_0)}\right)^{-1} (x_k - x_0) + x_0. \tag{13}$$

This expression cannot be evaluated directly because a matrix exponential has to be formed and a matrix has to be inverted. If A is similar to a diagonal matrix so that $A = C^{-1}JC$, where $J = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix consisting of the eigenvalues λ_i of A, then

$$I - \mathrm{e}^{A(t_k - t_0)} = C^{-1} \left(I - \mathrm{e}^{J(t_k - t_0)}\right) C. \tag{14}$$

Let

$$D_k = \left(I - \mathrm{e}^{J(t_k - t_0)}\right)^{-1} = \mathrm{diag}\left(\frac{1}{1 - \mathrm{e}^{\lambda_1(t_k - t_0)}}, \ldots, \frac{1}{1 - \mathrm{e}^{\lambda_n(t_k - t_0)}}\right), \tag{15}$$

then

$$\bar{x} = C^{-1} D_k C (x_k - x_0) + x_0 \tag{16}$$

and

$$C\bar{x} = D_k C(x_k - x_0) + Cx_0. \tag{17}$$

Building the difference of equation (17) for k and $k-1$ gives

$$D_k C(x_k - x_0) = D_{k-1} C(x_{k-1} - x_0). \tag{18}$$

As D_k and D_{k-1} are diagonal matrices equation (18) gives a relation for the entries of the matrices D_k and D_{k-1}. Let $[C(x_k - x_0)]_i$ be the i-th component of $C(x_k - x_0)$, then

$$\frac{1}{1-\mathrm{e}^{\lambda_i(t_k-t_0)}}[C(x_k - x_0)]_i = \frac{1}{1-\mathrm{e}^{\lambda_i(t_{k-1}-t_0)}}[C(x_{k-1} - x_0)]_i. \tag{19}$$

x_k, x_{k-1} and x_0 are known, t_k, t_{k-1} and t_0 can be calculated from (8). If C is known, then the λ_i can be calculated from (19) and D_k can be determined by (15). If C^{-1} is known, then $\bar{x}$ can be extrapolated according to (16).
But C and C^{-1} are unknown so that this approach is not possible. However, the analysis shows that there is an asymptotic exponential behavior for CG-type methods. The different exponents are coupled by the eigenvector matrices C and C^{-1}.
The eigenvector matrices C and C^{-1} can be approximated by the Lanczos process and the above mentioned approach can be performed. However, the extrapolation vector $\bar{x}$ is then obtained from (16), where C^{-1} is substituted by the Lanczos approximation. Therefore, $\bar{x}$ is in a (by x_0 shifted) space spanned by the Lanczos vectors. This means, $\bar{x}$ is in the same space as generated by BICO or QMR. We do not believe that the convergence of these methods can be accelerated.

IMPLICIT-EXPLICIT APPROXIMATIONS

Using the duality of iterative linear solvers and solution methods for ordinary differential equations mixed implicit-explicit finite difference approximations are investigated. In general implicit methods allow larger time steps than explicit methods.
Let us consider mixed implicit-explicit finite difference approximations for (1) of the form

$$\sum_{i=1}^{\sigma_k} \gamma_{i,k} \frac{x_k - x_{k-i}}{t_k - t_{k-i}} \quad = \quad \delta(Ax_k - b) + (1-\delta) r_{k-1} \tag{20}$$

Setting again α_j as in (8) for $j = 1, \ldots, \sigma_k$ we obtain the implicit formula

$$x_k \quad = \quad \delta\alpha_1(Ax_k - b) + (1-\delta)\alpha_1 r_{k-1} + \sum_{i=2}^{\sigma_k} \alpha_i (x_{k-i} - x_{k-1}) + x_{k-1}. \tag{21}$$

Introducing

$$\eta = \frac{-\delta\alpha_1}{1-\delta\alpha_1}, \tag{22}$$

$$B_\eta = (1-\eta)I + \eta A \tag{23}$$

and

$$\beta_i = \frac{\alpha_i}{1-\delta\alpha_1} \stackrel{(22)}{=} \alpha_i(1-\eta), \tag{24}$$

(21) can be written as

$$B_\eta x_k = \beta_1 r_{k-1} + \sum_{i=2}^{\sigma_k} \beta_i (x_{k-i} - x_{k-1}) + B_\eta x_{k-1}. \tag{25}$$

Solving (25) for x_k leads to

$$x_k = \beta_1 B_\eta^{-1} r_{k-1} + \sum_{i=2}^{\sigma_k} \beta_i B_\eta^{-1}(x_{k-i} - x_{k-1}) + x_{k-1} \tag{26}$$

If $\eta = 0$, then $B_0 = I$ and we have the explicit scheme (7). If $\eta = 1$, then $B_1 = A$ and we have the fully implicit scheme. For $\eta = 1$ the solution $\bar{x} = A^{-1}b$ of (4) is obtained in the first step, if the β_i are chosen such that any residual norm is minimized:

$$\begin{aligned} \|r_1\|_{any} &= \|Ax_1 - b\|_{any} \overset{(26)}{=} \|A(\beta_1 B_1^{-1} r_0 + x_0) - b\|_{any} \\ &\overset{B_1=A}{=} \|\beta_1(Ax_0 - b) + Ax_0 - b\|_{any} \overset{\beta_1=-1}{=} 0. \end{aligned}$$

The approximation can be calculated by

$$x_k = \sum_{i=1}^{\sigma_k} \beta_i y_{k-i} + x_{k-1}, \tag{27}$$

where in each iteration step the two linear systems

$$\begin{aligned} B_\eta y_{k-1} &= r_{k-1} \\ B_\eta y_{k-2} &= x_{k-2} - x_{k-1} \end{aligned} \tag{28}$$

have to be solved. If η is not too far from 0, then the matrix B_η is close to the unit matrix so that iterative methods applied to (28) converge very fast. If η is close to 1, then the solution of (28) by iterative methods is a similar problem as the solution of (4) but only a few outer iterations have to be performed.
The coefficients β_i are determined so that the residuals are minimized

$$\|r_k\| = \min_{\beta_1,\ldots,\beta_{\sigma_k}} \left\| \sum_{i=1}^{\sigma_k} \beta_i A y_{k-i} + r_{k-1} \right\|. \tag{29}$$

NUMERICAL TESTS

Throughout the numerical tests we applied as inner iteration method a smoothed ORTHORES method using 20 preceding aproximations. The inner method is stopped after 100 iterations. For the outer iteration we used $\sigma_k = \min(k, 5)$ and restarted every 20 iterations. Please note that by this choice of the inner solver the Krylov space is the same as for explicit approximations. We believe that an essential improvement will be achieved if the Krylov space is abandoned by some other inner solver. This work is still ahead and beyond the scope of this article.
The dependence of η is depicted in figure 2 for a Reynolds number $R = 500$. For $\eta = 0$ (corresponding to a truncated GMRES(5)) the method does not converge. The number of iterations decreases if η increases. For $\eta = 1$ the best result is obtained. Note that we do not get the true solution for $\eta = 1$ in the first step because we do not solve the equation exactly. On the contrary, for $R = 500$ we obtain after 100 inner iterations only a relative reduction of the residual norm by a factor of 0.98 in average. Thus indeed a non-converging explicit method ($\eta = 0$) becomes convergent if the implicit part increases.
The here discussed methods are similar to preconditioned methods, but different. If we would precondition the truncated GMRES(5) method by B_η, the we get the approximation by a formula similar to (26)

$$x_k = \beta_1 B_\eta^{-1} r_{k-1} + \sum_{i=2}^{\sigma_k} \beta_i (x_{k-i} - x_{k-1}) + x_{k-1}. \tag{30}$$

Again the equation $B_\eta y_{k-1} = r_{k-1}$ is solved by a smoothed ORTHORES method using 20 preceding aproximations and stopped after 100 iterations. In figure 3 the truncated GMRES(5), the here proposed method with $\eta = 1$ and the preconditioned GMRES(5) according to (30) are plotted for

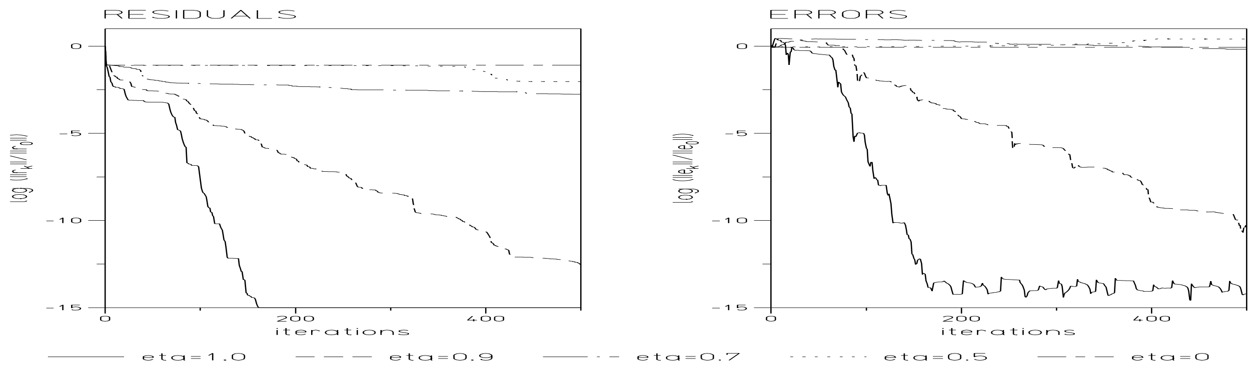

Figure 2: Relative residuals and errors for the sample problem R=500

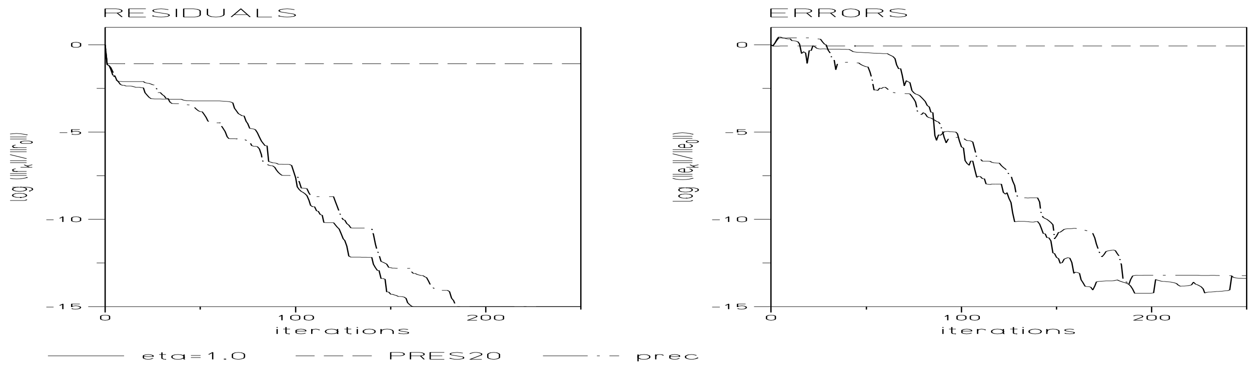

Figure 3: Relative residuals and errors for the sample problem R=500

the sample problem with $R = 500$. While the basic method does not converge both other methods converge similar. For preconditioning one additional linear system has to be approximated in each iteration step while for the implicit method two systems have to be solved. Therefore, preconditioning is cheaper. For $R = 600$ neither the basic method nor the preconditioned system converge. But the implicit method still delivers a solution. The situation remains the same until a Reynolds number of $R = 1000$ where all three investigated methods fail. Thus this example shows that implicit methods deliver comparable results as preconditioning but are more robust.

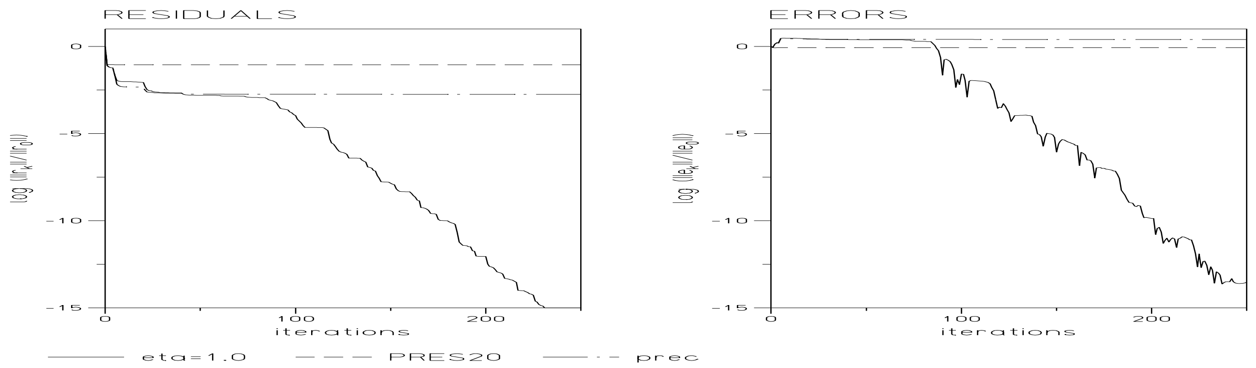

Figure 4: Relative residuals and errors for the sample problem R=600

REFERENCES

[1] W. Schönauer, E. Schnepf, and H. Müller. The FIDISOL program package. Internal report 27/85, University of Karlsruhe, Computing Center, Postfach 6980, 76128 Karlsruhe, Germany, 1985.

[2] R. Weiss. *Parameter-Free Iterative Linear Solvers*. Mathematical Research, vol. 97. Akademie Verlag, Berlin, 1996.