High-performance Deflated Conjugate Gradient Method

J. Kruzík 1
D. Horák 2

The talk deals with the choice of the deflation space for the deflated conjugate gradients (DCG) as well as with the DCG coarse problem solution employing a multilevel approach. A PETSc-based implementation of the method will demonstrate the effectiveness of the method on various large-scale benchmarks.

The conjugate gradient (CG) algorithm is often the method of choice for the solution of large symmetric positive definite linear systems of the form

$$Ax = b.$$  

In order to accelerate the convergence of CG we often need a suitable preconditioner. However, there also exists a complementary approach to the preconditioning known as deflation. The deflation utilizes a deflation space that should represent slowly converging components of the solution.

The deflated conjugate gradient (DCG) method [1], introduced in [2, 3, 4], works by splitting the solution of the linear system into two parts. The first part represents the solution in the deflation space and is directly obtained. The second one is computed by CG iterations that operate only on the $A$-conjugate complement of the deflation space.

Given a full rank deflation matrix $W$ whose columns span the deflation space, we can create a projection on the $A$-conjugate complement of $W$

$$P = I - W (W^T AW)^{-1} W^T A.$$  

This projection is then used inside CG iterations to keep the solution (and the descend directions) in the $A$-conjugate complement of the deflation space.

It can be shown [5] that DCG act as CG "preconditioned" by the projector $P$ as the convergence is governed by the spectrum of $PA$ operator. Moreover, $PA$ operator with a good choice of the deflation space has some of its eigenvalues shifted to zero (deflated).

A good choice of deflation space is crucial for making DCG converge quickly. In practice, there were two main deflation spaces.

The first one uses eigenvectors of $A$ as the deflation space. The associated eigenvalues of the eigenvectors belonging to the deflation space are shifted to zero in the spectrum of the DCG operator $PA$. Particularly, eigenvectors belonging to the smallest eigenvalues are used as they

---

1 Institute of Geonics of the Czech Academy of Sciences & VSB-Technical University of Ostrava
2 Institute of Geonics of the Czech Academy of Sciences & VSB-Technical University of Ostrava
slow down the convergence of CG the most. In our experiments, this approach works very well. The problem is how to obtain the eigenvectors.

The second approach is subdomains aggregation. Given a decomposition of the computational domain, each subdomain contributes a single vector into the deflation space. This vector contains ones on the indices of unknowns belonging to the subdomain and zeros otherwise. Such space often approximates a similar space as in the eigenvector approach. We can use, e.g., METIS to obtain the domain decomposition. However, assuming a single computational core owns the whole subdomain then, to utilize the cores appropriately, the subdomains have to be fairly large making the deflation space too coarse to be effective.

A new approach based on wavelet compression was suggested in [6]. The basic idea is that given the wavelet scaling coefficients \(h_1, \ldots, h_k\) we create a projection onto the scaling subspace

\[
H_{1,n} = \begin{pmatrix}
    h_1 & h_2 & h_3 & \cdots & 0 & \cdots & 0 & 0 \\
    0 & 0 & h_1 & h_2 & \cdots & 0 & 0 \\
    \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
    h_{k-1} & h_k & 0 & 0 & 0 & \cdots & h_{k-3} & h_{k-2} \\
\end{pmatrix} \in \mathbb{R}^{\frac{n}{2} \times n}.
\]

Then \(H_{1,n}AH_{1,n}^T\) contains trends of \(A\). Moreover, we can repeat this compression process to use up to \(m\) levels of the compression

\[
H_{1,n/2^{m-1}} \cdots H_{1,n/2}H_{1,n}AH_{1,n}^T \cdots H_{1,n/2}H_{1,n}AH_{1,n}^T\cdots H_{1,n/2} = H_{m,n}AH_{m,n}^T
\]

Since \(H_{m,n}\) cuts off the high frequencies, we can set \(W = H_{m,n}^T\).

The suggested wavelet compression is also used in the algebraic multigrid [7]. Therefore, using the prolongation matrices from multigrid in place of the deflation matrix might work as well. Moreover, the prolongation operators can be chained, as in the wavelet-based deflation, without the use of any smoothers between multigrid levels.

While large deflation spaces can be highly effective in decreasing the number of iterations, they also make the solution of the inverse (coarse problem) in the projector \(P\) difficult. To alleviate this problem, we turn to the multilevel deflation [8], where DGG is recursively used to solve the coarse problem on each level until the coarse problem is sufficiently small to be quickly solved by a direct solver.

In order to evaluate the aforementioned deflation spaces, an efficient, parallel implementation of DCG was created. It is written as a solver for linear systems in PETSc [9] (KSP). Currently, it is part of the PETSc-based, open-source PERMON library [10].

The benchmarks used in the numerical experiments include, e.g., matrices from SuiteSparse Matrix collection, 3D linear elasticity multi-material cantilever beam and 2D Laplace discretized by boundary element method on an L-shaped domain. Some of the results with appropriate discussion can be found in [5].
References


