Multi-shift conjugate gradient algorithm

Martin Lüscher January 2012

1. Introduction

Let A be a positive hermitian operator acting on the states x of a linear space. For any shift $s \geq 0$, the equation

$$(A + s)x = b \tag{1.1}$$

can be solved iteratively using the conjugate gradient (CG) algorithm (see ref. [1], for example). With little further work, it is in fact possible to solve the equation for several shifts s at once [2]. In this note the multi-shift algorithm is briefly derived starting from the standard formulation of the CG algorithm.

2. Basic CG algorithm

For a given source b, the basic CG algorithm solves the linear system

$$Ax = b \tag{2.1}$$

iteratively by generating a sequence

$$x_k, p_k, \quad k = 1, 2, 3, \ldots \tag{2.2}$$

of approximate solutions x_k and search directions p_k recursively. In the course of the recursion, the residues

$$r_k = b - Ax_k \tag{2.3}$$

are also computed.

The CG recursion starts from the initial values

$$x_1 = 0, \quad r_1 = p_1 = b. \tag{2.4}$$

Once x_k, r_k and p_k have been computed, the next approximate solution,

$$x_{k+1} = x_k + \alpha_k p_k, \tag{2.5}$$

is obtained by calculating Ap_k and the coefficient

$$\alpha_k = \frac{(r_k, r_k)}{(p_k, Ap_k)}, \tag{2.6}$$

where the bracket $(\cdot\,,\cdot)$ denotes the scalar product in the space of states.

The residue of the new solution,

$$r_{k+1} = r_k - \alpha_k A p_k, \tag{2.7}$$

may then be calculated easily since Ap_k is already known. After that the coefficient

$$\beta_k = \frac{(r_{k+1}, r_{k+1})}{(r_k, r_k)} \tag{2.8}$$

and the next search direction

$$p_{k+1} = r_{k+1} + \beta_k p_k \tag{2.9}$$

can be computed, thus completing the recursion.

3. Shifted system

The CG algorithm can also be used to solve the shifted system

$$(A + s)\hat{x} = b \tag{3.1}$$

for given source b and shift $s \geq 0$. Explicitly, the appoximate solutions $\hat{x}_k$, associated residues $\hat{r}_k$ and search directions $\hat{p}_k$ are, in this case, generated through the recursion

$$\hat{x}_{k+1} = \hat{x}_k + \hat{\alpha}_k \hat{p}_k, \qquad \hat{\alpha}_k = \frac{(\hat{r}_k, \hat{r}_k)}{(\hat{p}_k, (A+s)\hat{p}_k)}, \tag{3.2}$$

$$\hat{r}_{k+1} = \hat{r}_k - \hat{\alpha}_k (A+s)\hat{p}_k, \tag{3.3}$$

$$\hat{p}_{k+1} = \hat{r}_{k+1} + \hat{\beta}_k \hat{p}_k, \qquad \hat{\beta}_k = \frac{(\hat{r}_{k+1}, \hat{r}_{k+1})}{(\hat{r}_k, \hat{r}_k)}. \tag{3.4}$$

The initial values

$$\hat{x}_1 = 0, \quad \hat{r}_1 = \hat{p}_1 = b, \tag{3.5}$$

are the same as before.

Independently of s, the residues $\hat{r}_1, \ldots, \hat{r}_n$ span the same linear subspace of states. The CG algorithm moreover has the feature of producing orthogonal residues, viz.

$$(\hat{r}_k, \hat{r}_l) = 0 \quad \text{for all} \quad k \neq l \tag{3.6}$$

(for a proof of this property, see ref. [1] for example). In particular, $\hat{r}_k$ is orthogonal to the space spanned by $\hat{r}_1, \ldots, \hat{r}_{k-1}$. Since this space does not depend on s, the relation

$$\hat{r}_k = \hat{\gamma}_k r_k \tag{3.7}$$

must hold for some complex (s-dependent) constants $\hat{\gamma}_k$.

It is possible to compute the coefficients $\hat{\gamma}_k$, $\hat{\alpha}_k$ and $\hat{\beta}_k$ recursively together with α_k and β_k. A little algebra shows that

$$\hat{\gamma}_1 = 1, \qquad \hat{\gamma}_2 = \frac{1}{1 + s\alpha_1}, \qquad \hat{\alpha}_1 = \hat{\gamma}_2 \alpha_1, \qquad \hat{\beta}_1 = \hat{\gamma}_2^2 \beta_1. \tag{3.8}$$

The recursion for the higher-order coefficients may be derived by noting that

$$r_{k+1} = -\alpha_k A r_k + (1 + \omega_k) r_k - \omega_k r_{k-1}, \qquad \omega_k = \frac{\alpha_k \beta_{k-1}}{\alpha_{k-1}}, \tag{3.9}$$

$$\hat{r}_{k+1} = -\hat{\alpha}_k A \hat{r}_k + (1 + \hat{\omega}_k - s\hat{\alpha}_k)\hat{r}_k - \hat{\omega}_k \hat{r}_{k-1}, \qquad \hat{\omega}_k = \frac{\hat{\alpha}_k \hat{\beta}_{k-1}}{\hat{\alpha}_{k-1}}, \tag{3.10}$$

for all $k \geq 2$. After substitution of eq. (3.7) and division of eq. (3.10) by $\hat{\gamma}_{k+1}$, the coefficients of Ar_k, r_k and r_{k-1} in these two equations must match. The relations

$$\hat{\alpha}_k = \hat{\rho}_k \alpha_k, \tag{3.11}$$

$$\hat{\omega}_k = -1 + \hat{\rho}_k(1 + \omega_k + s\alpha_k) \tag{3.12}$$

$$\hat{\omega}_k = \hat{\rho}_k \hat{\rho}_{k-1} \omega_k \tag{3.13}$$

are thus obtained, where the abbreviation

$$\hat{\rho}_k = \frac{\hat{\gamma}_{k+1}}{\hat{\gamma}_k}. \tag{3.14}$$

was used.

Recalling the definition of ω_k and $\hat{\omega}_k$, the combination of eqs. (3.11) and (3.13) leads to

$$\hat{\beta}_k = \hat{\rho}_k^2 \beta_k, \tag{3.15}$$

a relation that could also be directly obtained from the definitions (2.8) and (3.4) of β_k and $\hat{\beta}_k$. The combination of (3.12) and (3.13) moreover shows that

$$\hat{\rho}_k = \{1 + s\alpha_k + (1 - \hat{\rho}_{k-1})\omega_k\}^{-1}, \tag{3.16}$$

which allows the ratios $\hat{\rho}_k$ to be computed recursively starting from $\hat{\rho}_1 = (1+s\alpha_1)^{-1}$.

4. Multi-shift algorithm

In this section, an implementation of the multi-shift CG algorithm is desribed that solves two equations,

$$Ax = b, \tag{4.1}$$

$$(A + s)x = b, \tag{4.2}$$

simultaneously. Further systems with different shifts s can be easily included in this algorithm by dublicating the steps required for the solution of the second equation.

The algorithm updates the states x_k, r_k, p_k and $\hat{x}_k, \hat{p}_k$ recursively, starting from

$$x_1 = \hat{x}_1 = 0, \quad r_1 = p_1 = \hat{p}_1 = b. \tag{4.3}$$

In the course of the recursion, the coefficients $\alpha_{k-1}, \beta_{k-1}$ and $\hat{\gamma}_k, \hat{\rho}_{k-1}$ are updated together with the states. Their values at the beginning of the recursion,

$$\beta_0 = 0, \quad \alpha_0 = \hat{\gamma}_1 = \hat{\rho}_0 = 1, \tag{4.4}$$

are chosen so that the update rules specified below will give the correct sequence of coefficients.

The step that leads from $x_k, \ldots, \hat{p}_k$ to $x_{k+1}, \ldots, \hat{p}_{k+1}$ first requires the computation of Ap_k and of the coefficients

$$\alpha_k = \frac{(r_k, r_k)}{(p_k, Ap_k)}, \tag{4.5}$$

$$\omega_k = \frac{\alpha_k \beta_{k-1}}{\alpha_{k-1}}, \tag{4.6}$$

$$\hat{\rho}_k = \frac{1}{1 + s\alpha_k + (1 - \hat{\rho}_{k-1})\omega_k}, \tag{4.7}$$

$$\hat{\alpha}_k = \hat{\rho}_k \alpha_k. \tag{4.8}$$

One can then compute the linear combinations

$$x_{k+1} = x_k + \alpha_k p_k, \tag{4.9}$$

$$\hat{x}_{k+1} = \hat{x}_k + \hat{\alpha}_k \hat{p}_k, \tag{4.10}$$

$$r_{k+1} = r_k - \alpha_k A p_k. \tag{4.11}$$

At this point the coefficients

$$\beta_k = \frac{(r_{k+1}, r_{k+1})}{(r_k, r_k)}, \tag{4.12}$$

$$\hat{\beta}_k = \hat{\rho}_k^2 \beta_k, \tag{4.13}$$

$$\hat{\gamma}_{k+1} = \hat{\rho}_k \hat{\gamma}_k, \tag{4.14}$$

can be calculated and subsequently the linear combinations

$$p_{k+1} = r_{k+1} + \beta_k p_k, \tag{4.15}$$

$$\hat{p}_{k+1} = \hat{\gamma}_{k+1} r_{k+1} + \hat{\beta}_k \hat{p}_k, \tag{4.16}$$

thus completing the step that leads from $x_k, \ldots, \hat{p}_k$ to $x_{k+1}, \ldots, \hat{p}_{k+1}$.

When the residue

$$\|\hat{r}_k\| = \hat{\gamma}_k \|r_k\| \tag{4.17}$$

reaches the required tolerance, the recursion can be restricted to the basic system until $\|r_k\|$ satisfies the convergence criterion for the latter. Note that $\hat{\gamma}_k < 1$ if s is positive. The convergence rates are in fact determined by the condition number of $A + s$ and A, respectively, and are thus widely different if s much larger than the lowest eigenvalue of A.

References

[1] G. H. Golub, C. F. van Loan, *Matrix computations*, 2nd ed. (The Johns Hopkins University Press, Baltimore, 1989)

[2] B. Jegerlehner, *Krylov space solvers for shifted linear systems*, preprint IUHET-353 (1996) [arXiv: hep-lat/9612014]