

Mathematical Tripos Part II: Michaelmas Term 2023

Numerical Analysis – Lecture 19

Conjugate gradient method The conjugate gradient method is the method of conjugate directions (Theorem 4.22 from previous lecture) where the directions $\mathbf{d}^{(i)}$ are chosen so that they A -orthogonalize the residuals, i.e., the $\mathbf{d}^{(i)}$ satisfy

$$\text{span}(\mathbf{d}^{(0)}, \dots, \mathbf{d}^{(k-1)}) = \text{span}(\mathbf{r}^{(0)}, \dots, \mathbf{r}^{(k-1)}) \quad (4.8)$$

for every iteration k , in addition to being pairwise A -orthogonal. This can be achieved by applying the Gram-Schmidt procedure to the residuals $\{\mathbf{r}^{(i)}\}$. The conjugate gradient method can thus be expressed using the following iterates. Starting from $\mathbf{x}^{(0)} \in \mathbb{R}^n$, and letting $\mathbf{r}^{(0)} = \mathbf{b} - A\mathbf{x}^{(0)}$, iterate, for $k \geq 0$:

$$\begin{cases} \mathbf{d}^{(k)} = \mathbf{r}^{(k)} - \sum_{i < k} \frac{\langle \mathbf{r}^{(k)}, \mathbf{d}^{(i)} \rangle_A}{\langle \mathbf{d}^{(i)}, \mathbf{d}^{(i)} \rangle_A} \mathbf{d}^{(i)} \\ \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \frac{\langle \mathbf{r}^{(k)}, \mathbf{d}^{(k)} \rangle}{\langle \mathbf{d}^{(k)}, A\mathbf{d}^{(k)} \rangle} \mathbf{d}^{(k)}. \end{cases} \quad (4.9)$$

As written, the iterates above are not attractive from a computational point of view, because they require computing at least $k - 1$ inner products at each step k . However it turns out that the equation above defining $\mathbf{d}^{(k)}$ can be simplified dramatically and the terms $i < k - 1$ in the summation above happen to be zero! This is one of the key points of the CG method. To prove this, we first introduce the following important definition.

Definition 4.26 Let $\mathbf{v} \in \mathbb{R}^n$. The m 'th Krylov subspace of A with respect to \mathbf{v} is

$$K_m(A, \mathbf{v}) = \text{span}\{A^i \mathbf{v}\}_{i=0}^{m-1}.$$

The following theorem collects important observations about the iterates (4.9).

Theorem 4.27 (Properties of CGM) For every $m \geq 0$ such that $\mathbf{r}^{(m)} \neq 0$, the conjugate gradient method has the following properties.

- (1) The linear space spanned by the residuals $\{\mathbf{r}^{(i)}\}_{i=0}^m$ is the same as the linear space spanned by the conjugate directions $\{\mathbf{d}^{(i)}\}_{i=0}^m$ and it coincides with the Krylov subspace $K_{m+1}(A, \mathbf{r}^{(0)})$:

$$\{\mathbf{r}^{(i)}\}_{i=0}^m = \{\mathbf{d}^{(i)}\}_{i=0}^m = K_{m+1}(A, \mathbf{r}^{(0)}).$$

- (2) The residuals satisfy the orthogonality conditions: $\langle \mathbf{r}^{(m)}, \mathbf{r}^{(i)} \rangle = \langle \mathbf{r}^{(m)}, \mathbf{d}^{(i)} \rangle = 0$ for $i < m$.

- (3) The directions are conjugate (A -orthogonal): $\langle \mathbf{d}^{(m)}, \mathbf{d}^{(i)} \rangle_A = \langle \mathbf{d}^{(m)}, A\mathbf{d}^{(i)} \rangle = 0$ for $i < m$.

Proof. (1) For notational convenience, we let $K_m = K_m(A, \mathbf{r}^{(0)})$.

By the properties of the Gram-Schmidt procedure we know $\text{span}\{\mathbf{d}^{(i)}\}_{i=0}^m = \text{span}\{\mathbf{r}^{(i)}\}_{i=0}^m$. Furthermore, from Theorem 4.22 of the previous lecture, we know that $\{\mathbf{r}^{(0)}, \dots, \mathbf{r}^{(m)}\}$ is an orthogonal basis of this subspace. To show that this subspace is the Krylov subspace K_{m+1} , we proceed by induction. The case $m = 0$ is trivial. For $m \geq 1$, note that since the iterates $\mathbf{x}^{(k)}$ satisfy $\mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} + \alpha_m \mathbf{d}^{(m)}$, we get $\mathbf{r}^{(m+1)} = \mathbf{r}^{(m)} - \alpha_m A\mathbf{d}^{(m)}$. By the induction hypothesis, $\mathbf{r}^{(m)}, \mathbf{d}^{(m)} \in K_{m+1}$, and so $A\mathbf{d}^{(m)} \in K_{m+2}$. Thus $\mathbf{r}^{(m+1)} \in K_{m+2}$ as desired. This shows the inclusion $\text{span}\{\mathbf{r}^{(i)}\}_{i=0}^{m+1} \subseteq K_{m+2}$; and one can show that we have equality by dimension counting. Indeed since $\mathbf{r}^{(m+1)} \neq 0$ (by assumption), we know that $\mathbf{r}^{(m+1)}$ is orthogonal to $\mathbf{r}^{(i)}$ for $i \leq m$, and so we have

$$\dim \text{span}\{\mathbf{r}^{(i)}\}_{i=0}^{m+1} = 1 + \dim \text{span}\{\mathbf{r}^{(i)}\}_{i=0}^m = 1 + \dim K_{m+1}(A, \mathbf{r}^{(0)}) \geq \dim K_{m+2}.$$

Combining this with $\text{span}\{\mathbf{r}^{(i)}\}_{i=0}^{m+1} \subseteq K_{m+2}$ we get equality of the subspaces.

(2) From Theorem 4.22 of the previous lectures, we know that $\langle \mathbf{r}^{(m)}, \mathbf{d}^{(i)} \rangle = 0$ for all $i \leq m-1$ which is exactly what we want.

(3) This is by definition, since the $\mathbf{d}^{(i)}$ are obtained by A -orthogonalizing the residual vectors. \square

The properties above allow us to show that $\langle \mathbf{r}^{(k)}, \mathbf{d}^{(i)} \rangle_A = 0$ for $i \leq k-2$. Indeed, if we take $i \leq k-2$, then $\mathbf{d}^{(i)} \in K_{k-1}(A, \mathbf{r}^{(0)})$, which implies $A\mathbf{d}^{(i)} \in K_k(A, \mathbf{r}^{(0)})$ and so $\langle \mathbf{r}^{(k)}, \mathbf{d}^{(i)} \rangle_A = \langle \mathbf{r}^{(k)}, A\mathbf{d}^{(i)} \rangle = 0$.

As a result, the algorithm (4.9) simplifies and reduces to the following: Set $\mathbf{d}^{(0)} = \mathbf{r}^{(0)} = \mathbf{b} - A\mathbf{x}^{(0)}$ and iterate, for $k \geq 0$:

$$\begin{cases} \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)} & \alpha_k = \frac{\langle \mathbf{r}^{(k)}, \mathbf{d}^{(k)} \rangle}{\langle \mathbf{d}^{(k)}, A\mathbf{d}^{(k)} \rangle} \\ \mathbf{d}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k \mathbf{d}^{(k)} & \beta_k = -\frac{\langle \mathbf{r}^{(k+1)}, A\mathbf{d}^{(k)} \rangle}{\langle \mathbf{d}^{(k)}, A\mathbf{d}^{(k)} \rangle} \end{cases} \quad (4.10)$$

where $\mathbf{r}^{(k)}$ stands for $\mathbf{b} - A\mathbf{x}^{(k)}$.

We can further simplify the expressions for α_k and β_k using the properties stated in Theorem 4.27. Indeed, using the second equation in (4.10), and the fact that $\mathbf{r}^{(k)} \perp \mathbf{d}^{(k-1)}$, we have

$$\langle \mathbf{r}^{(k)}, \mathbf{d}^{(k)} \rangle = \langle \mathbf{r}^{(k)}, \mathbf{r}^{(k)} \rangle = \|\mathbf{r}^{(k)}\|_2^2 \quad (4.11)$$

which shows that

$$\alpha_k = \frac{\|\mathbf{r}^{(k)}\|_2^2}{\langle \mathbf{d}^{(k)}, A\mathbf{d}^{(k)} \rangle} > 0.$$

Also, we can write:

$$\beta_k = -\frac{\langle \mathbf{r}^{(k+1)}, A\mathbf{d}^{(k)} \rangle}{\langle \mathbf{d}^{(k)}, A\mathbf{d}^{(k)} \rangle} \stackrel{(a)}{=} -\frac{\langle \mathbf{r}^{(k+1)}, \mathbf{r}^{(k+1)} - \mathbf{r}^{(k)} \rangle}{\langle \mathbf{d}^{(k)}, \mathbf{r}^{(k+1)} - \mathbf{r}^{(k)} \rangle} \stackrel{(b)}{=} \frac{\|\mathbf{r}^{(k+1)}\|^2}{\langle \mathbf{d}^{(k)}, \mathbf{r}^{(k)} \rangle} \stackrel{(c)}{=} \frac{\|\mathbf{r}^{(k+1)}\|^2}{\|\mathbf{r}^{(k)}\|^2} > 0.$$

where we used in (a) the fact that $A\mathbf{d}^{(k)}$ is a multiple of $\mathbf{r}^{(k+1)} - \mathbf{r}^{(k)}$, and in (b) orthogonality of $\mathbf{r}^{(k+1)}$ to both $\mathbf{r}^{(k)}$, $\mathbf{d}^{(k)}$ (Theorem 4.27(2)), and in (c) we used (4.11).

The complete conjugate gradient method can thus be written as follows:

Algorithm 4.28 (Standard form of the conjugate gradient method) –

- (1) Set $k = 0$, $\mathbf{x}^{(0)} = 0$, $\mathbf{r}^{(0)} = \mathbf{b}$, and $\mathbf{d}^{(0)} = \mathbf{r}^{(0)}$;
- (2) Calculate the matrix-vector product $\mathbf{v}^{(k)} = A\mathbf{d}^{(k)}$ and $\alpha_k = \|\mathbf{r}^{(k)}\|^2 / \langle \mathbf{d}^{(k)}, \mathbf{v}^{(k)} \rangle > 0$;
- (3) Apply the formulae $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$ and $\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k \mathbf{v}^{(k)}$;
- (4) Stop if $\|\mathbf{r}^{(k+1)}\|$ is acceptably small;
- (5) Set $\mathbf{d}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k \mathbf{d}^{(k)}$, where $\beta_k = \|\mathbf{r}^{(k+1)}\|^2 / \|\mathbf{r}^{(k)}\|^2 > 0$;
- (6) Increase $k \rightarrow k+1$ and go back to (2).

The total work is dominated by the number of iterations, multiplied by the time it takes to compute $\mathbf{v}^{(k)} = A\mathbf{d}^{(k)}$. Thus the conjugate gradient algorithm is highly suitable when most of the elements of A are zero, i.e. when A is *sparse*.

Finite termination We have already seen that the method of conjugate directions (Theorem 4.22 in previous lecture) terminates after at most n steps. We restate this result in the special case of the conjugate gradient method.

Corollary 4.29 (A termination property) *If the conjugate gradient method is applied in exact arithmetic, then, for any $\mathbf{x}^{(0)} \in \mathbb{R}^n$, termination occurs after at most n iterations. More precisely, termination occurs after at most s iterations, where $s = \dim \text{span}\{A^i \mathbf{r}_0\}_{i=0}^{n-1}$ (which can be smaller than n).*

Proof. Assertion (2) of Theorem 4.27 states that residuals $(\mathbf{r}^{(k)})_{k \geq 0}$ form a sequence of mutually orthogonal vectors in \mathbb{R}^n , therefore at most n of them can be nonzero. Since they also belong to the space $\text{span}\{A^i \mathbf{r}_0\}_{i=0}^{n-1}$, their number is bounded by the dimension of that space. \square

We can bound the dimension of the Krylov subspace $\text{span}\{A^i \mathbf{r}_0\}_{i=0}^{n-1}$ using the number of distinct eigenvalues of A .

Theorem 4.30 (Number of iterations in CGM) *Let s be the number of distinct eigenvalues of A . Then, for any \mathbf{v} ,*

$$\dim K_m(A, \mathbf{v}) \leq s \quad \forall m. \quad (4.12)$$

Hence the number of iterations of CGM for solving $A\mathbf{x} = \mathbf{b}$ is bounded by the number of distinct eigenvalues of A .

Proof. Inequality (4.12) is true for any diagonalizable matrix A . Indeed if (\mathbf{u}_i) are the n linearly independent eigenvectors of A , one can expand $\mathbf{v} = \sum_{i=1}^n a_i \mathbf{u}_i$, and then group together eigenvectors with the same eigenvalues: for each λ_ν we set $\mathbf{w}_\nu = \sum_{k=1}^{m_\nu} a_{i_k} \mathbf{u}_{i_k}$ if $A\mathbf{u}_{i_k} = \lambda_\nu \mathbf{u}_{i_k}$. Then

$$\mathbf{v} = \sum_{\nu=1}^s c_\nu \mathbf{w}_\nu, \quad c_\nu \in \{0, 1\},$$

hence $A^i \mathbf{v} = \sum_{\nu=1}^s c_\nu \lambda_\nu^i \mathbf{w}_\nu$, thus for any m we get $K_m(A, \mathbf{v}) \subseteq \text{span}\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_s\}$, and that proves (4.12). By Corollary 4.29, the number of iteration in CGM is bounded by $\dim K_m(A, \mathbf{r}^{(0)})$, hence the final conclusion. \square

Remark 4.31 Theorem 4.30 shows that, unlike other iterative schemes, the conjugate gradient method is both iterative and direct: each iteration produces a reasonable approximation to the exact solution, and the exact solution itself will be recovered after n iterations at most.