CME 302: NUMERICAL LINEAR ALGEBRA
FALL 2005/06
LECTURE 19

GENE H. GOLUB

1. Conjugate Gradient Method

Many iterative methods for solving $A\mathbf{x} = \mathbf{b}$ have the form

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k-1)} + \omega_{k+1}(\alpha_k \mathbf{z}^{(k)} - \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}) \tag{1.1}$$

where

$$M\mathbf{z}^{(k)} = \mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)} \tag{1.2}$$

for some M. In particular, if $\omega_k \equiv 1$ and $\alpha_k \equiv 1$ then this reduces to

$$\mathbf{x}^{(k+1)} = M^{-1}(\mathbf{b} - A\mathbf{x}^{(k)}) - \mathbf{x}^{(k)}$$

or

$$M\mathbf{x}^{(k+1)} = \mathbf{b} - (A - M)\mathbf{x}^{(k)} = N\mathbf{x}^{(k)} + \mathbf{b}$$

where $A = M - N$. Our goal is to choose the parameters α_k and ω_k so that $\|P_k(M^{-1}A)\mathbf{e}^{(0)}\|$ is minimized, where $\mathbf{e}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$ and $\mathbf{e}^{(k)} = P_k(M^{-1}A)\mathbf{e}^{(0)}$.

Suppose that we can impose the condition that

$$(\mathbf{z}^{(k)}, M\mathbf{z}^{(k)}) = \delta_{jk}$$

where both M and A are $n \times n$ and required to be symmetric positive definite. If this is possible, then it follows that $\mathbf{z}^{(n+1)} = \mathbf{0}$, and therefore $\mathbf{r}^{(n+1)} = \mathbf{0}$, implying convergence in n iterations.

It follows from (1.1) that

$$\mathbf{b} - A\mathbf{x}^{(k+1)} = \mathbf{b} - A\mathbf{x}^{(k-1)} - \omega_{k+1}(\alpha_k A\mathbf{z}^{(k)} + A\mathbf{x}^{(k)} - \mathbf{b} + \mathbf{b} - A\mathbf{y}^{(k-1)})$$

which simplifies to

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k-1)} - \omega_{k+1}(\alpha_k A\mathbf{z}^{(k)} - \mathbf{r}^{(k)} + \mathbf{r}^{(k-1)}).$$

From (1.2), we obtain

$$M\mathbf{z}^{(k+1)} = M\mathbf{z}^{(k-1)} - \omega_{k+1}(\alpha_k A\mathbf{z}^{(k)} - M\mathbf{z}^{(k)} + M\mathbf{z}^{(k-1)}).$$

We use the induction hypothesis

$$(\mathbf{z}^{(p)}, M\mathbf{z}^{(q)}) = 0, \quad p \neq q, \quad p = 1, 2, \dots, k.$$

Then

$$(\mathbf{z}^{(k)}, M\mathbf{z}^{(k+1)}) = (\mathbf{z}^{(k)}, M\mathbf{z}^{(k-1)}) - \omega_{k+1}[(\alpha_k \mathbf{z}^{(k)}, A\mathbf{z}^{(k)}) - (\mathbf{z}^{(k)}, M\mathbf{z}^{(k)}) + (\mathbf{z}^{(k)}, M\mathbf{z}^{(k-1)})]$$

which yields

$$\alpha_k = \frac{(\mathbf{z}^{(k)}, M\mathbf{z}^{(k)})}{(\mathbf{z}^{(k)}, A\mathbf{z}^{(k)})}.$$

Similarly,

$$(\mathbf{z}^{(k-1)}, M\mathbf{z}^{(k+1)}) = (\mathbf{z}^{(k-1)}, M\mathbf{z}^{(k-1)}) - \omega_{k+1}[(\alpha_k \mathbf{z}^{(k-1)}, A\mathbf{z}^{(k)}) - (\mathbf{z}^{(k-1)}, M\mathbf{z}^{(k)}) + (\mathbf{z}^{(k-1)}, M\mathbf{z}^{(k-1)})]$$

Date: December 7, 2005, version 1.0.
Notes originally due to James Lambers. Minor editing by Lek-Heng Lim.

which yields

$$\omega_{k+1} = \frac{(\mathbf{z}^{(k-1)}, M\mathbf{z}^{(k-1)})}{\alpha_k(\mathbf{z}^{(k-1)}, A\mathbf{z}^{(k)}) + (\mathbf{z}^{(k-1)}, M\mathbf{z}^{(k-1)})}.$$

We can simplify this expression for ω_{k+1} by noting that by symmetry,

$$(\mathbf{z}^{(k-1)}, A\mathbf{z}^{(k)}) = (\mathbf{z}^{(k)}, A\mathbf{z}^{(k-1)})$$

and therefore

$$\begin{aligned}(\mathbf{z}^{(k)}, M\mathbf{z}^{(k)}) &= (\mathbf{z}^{(k)}, M\mathbf{z}^{(k-2)}) + \omega_k(\alpha_{k-1}(\mathbf{z}^{(k)}, A\mathbf{z}^{(k-1)}) - (\mathbf{z}^{(k)}, M\mathbf{z}^{(k-1)}) + (\mathbf{z}^{(k)}, M\mathbf{z}^{(k-2)}))\\ &= \omega_k(\alpha_{k-1}(\mathbf{z}^{(k)}, A\mathbf{z}^{(k-1)}))\end{aligned}$$

which yields

$$\omega_{k+1} = \frac{(\mathbf{z}^{(k-1)}, M\mathbf{z}^{(k-1)})}{-\frac{\alpha_k}{\alpha_{k+1}}\frac{1}{\omega_k}(\mathbf{z}^{(k)}, M\mathbf{z}^{(k)}) + (\mathbf{z}^{(k-1)}, M\mathbf{z}^{(k-1)})}$$

or

$$\omega_{k+1} = \left[1 - \frac{\alpha_k}{\alpha_{k-1}}\frac{1}{\omega_k}\frac{(\mathbf{z}^{(k)}, M\mathbf{z}^{(k)})}{(\mathbf{z}^{(k-1)}, M\mathbf{z}^{(k-1)})}\right]^{-1}.$$

We have shown that

$$(\mathbf{z}^{(k)}, M\mathbf{z}^{(k+1)}) = (\mathbf{z}^{(k-1)}, M\mathbf{z}^{(k+1)}) = 0.$$

It can easily be shown that

$$(\mathbf{z}^{(\ell)}, M\mathbf{z}^{(k+1)}) = 0, \quad \ell < k-1.$$

We now state the *classical conjugate gradient* algorithm:

$\mathbf{x}^{(0)}$ given
Solve $M\mathbf{z}^{(0)} = \mathbf{r}^{(0)}$
$\mathbf{p}^{(0)} = \mathbf{z}^{(0)}$
for $k = 0, \ldots$
  $\alpha_k = \frac{(\mathbf{z}^{(k)}, M\mathbf{z}^{(k)})}{(\mathbf{p}^{(k)}, A\mathbf{p}^{(k)})}$
  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{p}^{(k)}$
  $\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} + \alpha_k A\mathbf{p}^{(k)}$
  Test for convergence
  Solve $M\mathbf{z}^{(k+1)} = \mathbf{r}^{(k+1)}$
  $\beta_{k+1} = \frac{(\mathbf{z}^{(k+1)}, M\mathbf{z}^{(k+1)})}{(\mathbf{z}^{(k)}, M\mathbf{z}^{(k)})}$
  $\mathbf{p}^{(k+1)} = \mathbf{z}^{(k+1)} + \beta_{k+1}\mathbf{p}^{(k)}$
end

It can be shown that

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(0)} + P_k(K)\mathbf{z}^{(0)}$$

where $K = M^{-1}A$. Furthermore, amongst all methods which generate a polynomial for a given $\mathbf{x}^{(0)}$, the conjugate gradient method minimizes the quantity

$$\epsilon^{k+1} = \mathbf{e}^{(k+1)\top} A\mathbf{e}^{(k+1)}.$$

Most notable of all is that if A has p distinct eigenvalues, then the conjugate gradient method converges in p steps. This is particularly useful in domain decomposition, where the interface between two subdomains consists of only a small number of points.

Department of Computer Science, Gates Building 2B, Room 280, Stanford, CA 94305-9025
E-mail address: golub@stanford.edu