CME 302: NUMERICAL LINEAR ALGEBRA
FALL 2005/06
LECTURE 17

GENE H. GOLUB

1. Chebyshev Iteration

Consider the iterative method

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_{k+1}\mathbf{r}^{(k)}$$

for solving $A\mathbf{x} = \mathbf{b}$. If we define $\mathbf{e}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$, then

$$\mathbf{e}^{(k)} = P_k(A)\mathbf{e}^{(0)}$$

where

$$P_k(A) = (I - \alpha_k A)(I - \alpha_{k-1}A)\cdots(I - \alpha_1 A).$$

Therefore

$$\frac{\|\mathbf{e}^{(k)}\|_2}{\|\mathbf{e}^{(0)}\|_2} \leq \|P_k(A)\|_2 \leq \max_{1\leq i\leq N} |P_k(\lambda_i)| \leq \max_{a\leq\lambda_i\leq b} |P_k(\lambda)|$$

where $P_k(0) = I$ and $b = \lambda_1 \geq \cdots \geq \lambda_N = a$ are the eigenvalues of A.

Recall that a good choice for the polynomial P_k arises from the Chebyshev polynomials

$$C_k(\cos\theta) = \cos k\theta, \quad \theta = \cos^{-1} x.$$

If we fix k, then we have

$$\alpha_j^{(k)} = \left[\frac{b+a}{2} - \left(\frac{b-a}{2}\right)\cos\frac{(2j+1)\pi}{2k}\right]^{-1}, \quad j = 0,\ldots,k-1.$$

Note that

$$\alpha_0^{(1)} = \frac{2}{b+a},$$

which is the same optimal parameter obtained using a different analysis.

Therefore, we can select k and then use the parameters $\alpha_0^{(k)},\ldots,\alpha_{k-1}^{(k)}$. If $\|\mathbf{r}^{(k)}\|/\|\mathbf{r}^{(0)}\| \leq \epsilon$, we can stop; otherwise, we simply recycle these parameters. The process should not be stopped before the full cycle, because a partial polynomial may not be small on the interval $[a,b]$. Also, using the parameters in an arbitrary order may lead to numerical instabilities even though mathematically the order does not matter. For a long time, the determination of a suitable ordering was an open problem, but it has now been solved. It has been shown that when solving Laplace's equation using 128 parameters, a simple left-to-right ordering results in $\|\mathbf{e}^{(128)}\| \approx 10^{35}$, while the optimal ordering yields $\|\mathbf{e}^{(128)}\| \approx 10^{-7}$.

In the absence of roundoff error, using Chebyshev polynomials yields

$$\frac{\|\mathbf{e}^{(k)}\|_2}{\|\mathbf{e}^{(0)}\|_2} \leq \frac{2}{\left(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\right)^k + \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k} \approx \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k$$

Date: November 30, 2005, version 1.0.

Notes originally due to James Lambers. Minor editing by Lek-Heng Lim.

whereas, with steepest descent,

$$\frac{\|\mathbf{e}^{(k)}\|_2}{\|\mathbf{e}^{(0)}\|_2} \approx \left(\frac{\kappa-1}{\kappa+1}\right)^k.$$

2. Convergence Acceleration

Consider the iteration

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

where $A = M - N$ is symmetric positive definite. This iteration can be rewritten as

$$\mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \mathbf{c}$$

where $B = M^{-1}N$ and $\mathbf{c} = M^{-1}\mathbf{b}$. Therefore $\mathbf{e}^{(k+1)} = B\mathbf{e}^{(k)}$ where $\mathbf{e}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$. In an attempt to accelerate convergence, we define

$$\mathbf{y}^{(k)} = \sum_{\ell=0}^{k} a_{k\ell} x^{(\ell)}, \quad \sum_{\ell=0}^{k} a_{k\ell} = 1.$$

Then

$$\mathbf{x} - \mathbf{y}^{(k)} = \sum_{\ell=0}^{k} a_{k\ell}(\mathbf{x} - \mathbf{x}^{(\ell)}) = \sum_{\ell=0}^{k} a_{k\ell} B^\ell \mathbf{e}^{(0)}$$

which yields

$$\hat{\mathbf{e}}^{(k)} = P_k(B)\mathbf{e}^{(0)}$$

where $\hat{\mathbf{e}}^{(k)} = \mathbf{x} - \mathbf{y}^{(k)}$ and

$$P_k(\lambda) = \sum_{\ell=0}^{k} a_{k\ell}\lambda^\ell, \quad P_k(1) = 1.$$

It follows that

$$\frac{\|\hat{\mathbf{e}}^{(k)}\|_2}{\|\hat{\mathbf{e}}^{(0)}\|_2} \leq \|P_k(B)\|_2.$$

If B is symmetric, then we can write $B = Q\Lambda Q^T$ and obtain

$$\|P_k(B)\|_2 = \|P_k(\Lambda)\|_2 = \max_{\lambda=\lambda_i} |P_k(\lambda)| \leq \max_{a\leq\lambda\leq b} |P_k(\lambda)|.$$

Recall that the Chebyshev polynomials $C_k(x)$ satisfy the three-term recurrence relation

$$C_{k+1}(x) = 2xC_k(x) - C_{k-1}(x).$$

If we let $B = I - \alpha A$ where $\alpha = 2/(a+b)$, then B is symmetric and we can use the iteration

$$\mathbf{y}^{(\ell+1)} = \omega_{\ell+1}(B\mathbf{y}^{(\ell)} + \mathbf{c} - \mathbf{y}^{(\ell-1)}) + \mathbf{y}^{(\ell-1)}$$

with initial vectors

$$\mathbf{y}^{(0)} = \mathbf{x}^{(0)}, \quad \mathbf{y}^{(1)} = B\mathbf{y}^{(0)} + \mathbf{c}.$$

The parameters $\omega_{\ell+1}$ are defined by

$$\omega_{\ell+1} = \left(1 - \frac{\rho^2\omega_\ell}{4}\right)^{-1}, \quad \ell \geq 1, \quad \rho = \frac{b-a}{b+a}.$$

It follows that

$$\omega_2 \geq \omega_3 \geq \cdots \geq \omega^* > 1$$

where

$$\omega_* = \lim_{\ell\to\infty} \omega_\ell.$$

What is the limit ω^*? This limit satisfies

$$\omega^* = \left(1 - \frac{\rho^2\omega^*}{4}\right)^{-1}$$

which is a quadratic equation with solutions

$$\omega^* = \frac{1 \pm \sqrt{1-\rho^2}}{\rho^2/2}.$$

Choosing the plus sign, we have

$$1 < \omega_* = \frac{2}{1+\sqrt{1-\rho^2}} < 2.$$

Recall that for solving Poisson's equation, $\rho = 1 - ch^2 + O(h^4)$ for the Jacobi method, while $\rho = 1 - c'h + O(h^2)$ for the Chebyshev method.

3. Convergence for Positive Definite Systems

Let $A = M - N$, where $A = A^*$ and M is invertible, and define $Q = M + M^* - A$. If Q and A are both positive definite, then $\rho(M^{-1}N) < 1$. To prove this, we define $B = M^{-1}N = I - M^{-1}A$. It follows that if $B\mathbf{u} = \lambda\mathbf{u}$, then

$$A\mathbf{u} = (1-\lambda)M\mathbf{u},$$

where $\lambda \neq 1$ since A is invertible. Taking the inner product of both sides with $\mathbf{u}$ yields

$$\mathbf{u}^*A\mathbf{u} = (1-\lambda)\mathbf{u}^*M\mathbf{u},$$

but since A is symmetric positive definite, we also have

$$\mathbf{u}^*A\mathbf{u} = (1-\bar{\lambda})\mathbf{u}^*M^*\mathbf{u}.$$

Adding these relaions yields

$$\begin{aligned}
\mathbf{u}^*(M+M^*)\mathbf{u} &= \left(\frac{1}{1-\lambda} + \frac{1}{1-\bar{\lambda}}\right)\mathbf{u}^*A\mathbf{u} \\
&= 2\,\mathrm{Re}\left(\frac{1}{1-\lambda}\right)\mathbf{u}^*A\mathbf{u}
\end{aligned}$$

which can be rewritten as

$$\frac{\mathbf{u}^*(Q+A)\mathbf{u}}{\mathbf{u}^*A\mathbf{u}} = 1 + \frac{\mathbf{u}^*Q\mathbf{u}}{\mathbf{u}^*A\mathbf{u}} = 2\,\mathrm{Re}\left(\frac{1}{1-\lambda}\right).$$

Since both Q and A are positive definite, we must have

$$2\,\mathrm{Re}\left(\frac{1}{1-\lambda}\right) > 1.$$

If we write $\lambda = \alpha + i\beta$, then it follows that

$$\frac{2(1-\alpha)}{(1-\alpha)^2+\beta^2} > 1$$

which yields $\alpha^2 + \beta^2 = |\lambda|^2 < 1$.

4. Successive Overrelaxation with Positive Definite Systems

Let $A = D + L + U$ be positive definite with $D = I$. Then the iteration matrix for SOR is

$$\mathcal{L}_\omega = \left(\frac{1}{\omega}I + L\right)^{-1}\left(\left(\frac{1}{\omega}-1\right)I - U\right).$$

Then $Q = M + M^* - A$ is

$$Q = \left(\frac{1}{\omega}I + L\right) + \left(\frac{1}{\omega}I + U\right) - (I+L+U) = \left(\frac{2}{\omega}-1\right)I.$$

For convergence, we want Q to be positive definite, so we must have $2/\omega - 1 > 0$ or $0 < \omega < 2$. It follows that SOR will converge for all $0 < \omega < 2$ when A is positive definite.

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