AMSC 600 /CMSC 760 Advanced Linear Numerical Analysis
Fall 2007
Chebyshev Semi-Iterative Methods
Dianne P. O'Leary

Notes on Chebyshev Semi-Iterative Methods

In this set of notes we consider a non-stationary iteration for solving a linear system of equations.

The idea is built upon (any) SIM

$$\mathbf{x}^{(k+1)} = \mathbf{G}\mathbf{x}^{(k)} + \mathbf{d},$$

that converges to a fixed point $\mathbf{x}^*$.

Note: The notation for the constant vector in this set of notes is $\mathbf{d}$ instead of $\mathbf{c}$, because Chebyshev polynomials are almost always written using the letter c.

The Main Idea

- Suppose that we have a basic stationary iterative method

$$\mathbf{x}^{(k+1)} = \mathbf{G}\mathbf{x}^{(k)} + \mathbf{d},$$

  that converges to a fixed point $\mathbf{x}^*$.

- Consider the accelerated sequence

$$\bar{\mathbf{x}}^{(k)} = \sum_{j=0}^{k} \nu_j(k)\, \mathbf{x}^{(j)}$$

  where the $\nu_j(k)$ are scalar parameters to be determined.

- The demand that $\mathbf{x}^*$ remain a fixed point of the iteration adds the constraint

$$\sum_{j=0}^{k} \nu_j(k) = 1.$$

- We want to determine the parameters $\nu_j(k)$ to accelerate convergence.

Measuring convergence

Let

$$\bar{\mathbf{e}}^{(k)} = \bar{\mathbf{x}}^{(k)} - \mathbf{x}^*, \;\; \mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^*.$$

and let

$$p_k(z) = \sum_{j=0}^{k} \nu_j(k)\, z^j$$

be a polynomial of degree k. Our constraint on the coefficients $\nu_j(k)$ means that $p_k(1) = 1$.

(Notice that the superscripts on $\mathbf{x}$ and $\mathbf{e}$ denote iteration numbers, while those on $\mathbf{z}$ denote exponentiation.)

Now,

$$\begin{aligned}
\bar{\mathbf{e}}^{(k)} &= \sum_{j=0}^{k} \nu_j(k)\mathbf{x}^{(j)} - \mathbf{x}^* \\
&= \sum_{j=0}^{k} \nu_j(k)(\mathbf{x}^{(j)} - \mathbf{x}^*) \\
&= \sum_{j=0}^{k} \nu_j(k)\mathbf{G}^j\mathbf{e}^{(0)} \\
&= p_k(\mathbf{G})\mathbf{e}^{(0)}
\end{aligned}$$

Therefore, our problem becomes this:

> Given some information about the eigenvalues of $\mathbf{G}$, find coefficients of p_k, with $p_k(1) = 1$, so that $p_k(\mathbf{G})\mathbf{e}^{(0)}$ is small for every choice of $\mathbf{e}^{(0)}$.

Digression: The Chebyshev Polynomials

The Chebyshev polynomials are defined by

$$\begin{aligned}
c_0(z) &= 1, \\
c_1(z) &= z, \\
c_{m+1}(z) &= 2zc_m(z) - c_{m-1}(z), \;\; m \geq 1\,.
\end{aligned}$$

Properties

1. For $-1 < z < 1$, $c_m(z) = \cos m\theta$, where $\cos\theta = z$.

   Proof: True for $m = 0, 1$.
   Recall that $\cos(m+1)\theta = 2\cos\theta\cos m\theta - cos(m-1)\theta$ for $m \geq 1$ and use the definitions above. $\square$

2. $\max_{-1\leq z\leq 1} |c_m(z)| = 1$ for all $m \geq 0$, because of the properties of $\cos$.

3. Again, because of the properties of $\cos$, $|c_m(z)|$ has $m+1$ maximas in $[-1, 1]$ for $m > 0$. These occur when $\cos m\theta = \pm 1$, or equivalently for $\theta_k = \pi k/m$ or $z_k = \cos\pi k/m$, $k = 0, 1, \ldots, m$.

4. The $m+1$ maximas and minimums of $c_m(z)$ alternate in sign and thus by continuity we have a root between each pair. This gives m roots, and since c_m is a polynomial of degree m, this is all of them.

5. Given two numbers s, t, with $s \notin [-1, 1]$, let $\gamma = t/c_m(s)$. Then γc_m is the polynomial that solves the problem

   $$\min_{p_m(s)=t} \max_{-1\leq z\leq 1} |p_m(z)|$$

   over all polynomials of degree m.

   Proof:

   (a) Note that γc_m has the correct degree and equals t at s.

   (b) Assume that $p^* \neq \gamma c_m$ solves the problem with a smaller maximum value. Let $r = \gamma c_m - p^*$. Then r is also a polynomial of degree at most m, and the $m+1$ values $r(z_k)$, $k = 0, \ldots, m$ alternate in sign. Therefore, r has m roots in $[-1, 1]$. But it also has a root at s since $r(s) = 0$. Therefore r must be the zero polynomial, a contradiction.

   (c) Uniqueness follows from a similar argument. $\square$

Semi-Iteration

We give several solutions to the problem

> Given some information about the eigenvalues of $\mathbf{G}$, find coefficients of p_k, with $p_k(1) = 1$, so that $p_k(\mathbf{G})\mathbf{e}^{(0)}$ is small for every choice of $\mathbf{e}^{(0)}$.

Case 1

(impractical, but we'll see later that conjugate gradients accomplishes this in a different way)

Let $\mathbf{G}$ have eigenvalues $\lambda_1, \ldots, \lambda_n$. Then its characteristic equation is $p_n^*(\lambda) = \det(\mathbf{G} - \lambda \mathbf{I}) = 0$ and $p_n^*(\mathbf{G}) = \mathbf{0}$.

Therefore, we could take n steps of any iterative method, take $p_n = p_n^*$, and have

$$\bar{\mathbf{e}}^n = p_n^*(\mathbf{G})\mathbf{e}^0 = \mathbf{0}.$$

Case 2

Suppose $\mathbf{G}$ is symmetric and $-1 < a \leq \lambda(\mathbf{G}) \leq b < 1$. Then

$$\begin{aligned} ||\bar{\mathbf{e}}^k||_2 &= ||p_k(\mathbf{G})\mathbf{e}^{(0)}||_2 \\ &\leq ||p_k(\mathbf{G})||_2 \, ||\mathbf{e}^{(0)}||_2 \\ &= \max_{\lambda(\mathbf{G})} |p_k(\lambda)| \, ||\mathbf{e}^{(0)}||_2. \end{aligned}$$

One polynomial that makes this last expression small is the one that solves

$$\min_{p(1)=1} \max_{a \leq \lambda \leq b} |p(\lambda)|$$

over all polynomials of degree at most m. The solution is a scaled and shifted Chebyshev polynomial:

$$p_k(\lambda) = \frac{c_k\left(\frac{2\lambda - (b+a)}{b-a}\right)}{c_k\left(\frac{2-(b+a)}{b-a}\right)}.$$

Case 3

Suppose $\mathbf{G}$ is symmetric and $-1 < -b \leq \lambda(\mathbf{G}) \leq b < 1$. (This is a special case of 2.) Then the solution polynomial is

$$p_k(\lambda) = \frac{c_k\left(\frac{2\lambda}{2b}\right)}{c_k\left(\frac{2}{2b}\right)}$$

Next we derive the iteration: we find a formula for $\bar{\mathbf{x}}^{(k)}$.

We have

$$\begin{aligned} c_{k+1}(z) &= 2z c_k(z) - c_{k-1}(z), \\ \bar{\mathbf{e}}^{(k)} &= p_k(\mathbf{G})\mathbf{e}^{(0)}, \\ c_k(\lambda/b) &= c_k(1/b)p_k(\lambda). \end{aligned}$$

Therefore,

$$c_{k+1}(1/b)p_{k+1}(\lambda) = \frac{2\lambda}{b} c_k(1/b)p_k(\lambda) - c_{k-1}(1/b)p_{k-1}(\lambda).$$

Multiply this by $\mathbf{e}^{(0)}$ and evaluate the polynomials at $\mathbf{G}$, giving

$$c_{k+1}(1/b)\bar{\mathbf{e}}^{(k+1)} = \frac{2}{b}\mathbf{G}c_k(1/b)\bar{\mathbf{e}}^{(k)} - c_{k-1}(1/b)\bar{\mathbf{e}}^{(k-1)}.$$

Now we use the definition $\bar{\mathbf{e}}^{(k)} = \bar{\mathbf{x}}^{(k)} - \mathbf{x}^*$, getting

$$c_{k+1}(1/b)(\bar{\mathbf{x}}^{(k+1)} - \mathbf{x}^*) = \frac{2c_k(1/b)}{b}\mathbf{G}(\bar{\mathbf{x}}^{(k)} - \mathbf{x}^*) - c_{k-1}(1/b)(\bar{\mathbf{x}}^{(k-1)} - \mathbf{x}^*),$$

and therefore

$$\begin{aligned} c_{k+1}(1/b)\bar{\mathbf{x}}^{(k+1)} &= \frac{2c_k(1/b)}{b}\mathbf{G}\bar{\mathbf{x}}^{(k)} - c_{k-1}(1/b)\bar{\mathbf{x}}^{(k-1)} \\ &\quad + [c_{k+1}(1/b) - \frac{2c_k(1/b)}{b}\mathbf{G} + c_{k-1}(1/b)]\mathbf{x}^*. \end{aligned}$$

Now, since $c_{k+1}(1/b) - \frac{2c_k(1/b)}{b} + c_{k-1}(1/b) = 0$, the red piece of this expression becomes

$$[c_{k+1}(1/b) - \frac{2c_k(1/b)}{b}\mathbf{G} + c_{k-1}(1/b)]\mathbf{x}^* = \frac{2c_k(1/b)}{b}(\mathbf{I} - \mathbf{G})\mathbf{x}^* = \frac{2c_k(1/b)}{b}\mathbf{d},$$

so

$$\begin{aligned} \bar{\mathbf{x}}^{(k+1)} &= \frac{\frac{2}{b}c_k(1/b)(\mathbf{G}\bar{\mathbf{x}}^{(k)} + \mathbf{d}) - c_{k-1}(1/b)\bar{\mathbf{x}}^{(k-1)}}{c_{k+1}(1/b)} \\ &= w_{k+1}(\mathbf{G}\bar{\mathbf{x}}^{(k)} + \mathbf{d} - \bar{\mathbf{x}}^{(k-1)}) + \bar{\mathbf{x}}^{(k-1)} \end{aligned}$$

where

$$w_{k+1} = \frac{2c_k(1/b)}{bc_{k+1}(1/b)} = 1 + \frac{c_{k-1}(1/b)}{c_{k+1}(1/b)}$$

for $k > 1$, with $w_1 = 1$.

Notes:

1. The $\mathbf{x}^{(k)}$ sequence need not be computed at all!
2. Two previous iterates must be saved, and the correct starting condition is $\bar{\mathbf{x}}^{(0)}$ arbitrary and $\bar{\mathbf{x}}^{(-1)} = \mathbf{0}$.
3. We have the relations

   $$\max_{-b\le\lambda\le b} |p_k(\lambda)| = \max_{-b\le\lambda\le b} \left|\frac{c_k(\lambda/b)}{c_k(1/b)}\right| = \frac{1}{c_k(1/b)},$$

   and further algebra gives

   $$\frac{1}{c_k(1/b)} \le \frac{2(w_b - 1)^{k/2}}{1 + (w_b - 1)^k}$$

   where

   $$w_k = \frac{2}{1 + \sqrt{1 - \rho^2(G)}}$$

   where ρ denotes the spectral radius. This gives a bound on the rate of convergence.

Case 4

Suppose $\mathbf{G}$ is nonsymmetric.

We have just constructed an iteration using a min-max problem over an interval known to contain the eigenvalues.

In the nonsymmetric case, the min-max problem is over an ellipse. The solution was constructed by Tom Manteuffel, *Numerische Mathematik* 31 (1978) 183-208.