Mathematical Tripos Part IB: Easter Term 2002

Numerical Analysis – Lecture 6

6 Iterative methods for linear systems (cont.)

6.1 Sufficient condition for convergence

The next statement provides a condition which guarantees convergence of an iterative method.

Theorem 6.1 *Let $H \in \mathbb{R}^{n\times n}$ and let $\|SHS^{-1}\| < 1$ for some induced matrix norm $\|\cdot\|$ and for some nonsingular S (e.g., $\|H\|_\infty < 1$). Then the method*

$$x^{k+1} = Hx^k + v, \qquad x^0, v \in \mathbb{R}^n,$$

converges.

Proof. Set $\widetilde{H} := SHS^{-1}$ and let $\alpha := \|S\|\,\|S^{-1}\|$. Then

$$\|e^k\| = \|H^k e^0\| = \|S^{-1}\widetilde{H}^k S e^0\| \le \|S^{-1}\|\,\|\widetilde{H}^k\|\,\|S\|\,\|e^0\| \le \alpha\|\widetilde{H}\|^k\,\|e^0\| \le \alpha\gamma^k\|e^0\| \to 0. \qquad \blacksquare$$

As an application of this result, consider the following partially familiar class of matrices.

Definition 6.2 A matrix A is called strictly diagonally dominant by rows or by columns, if

$$|a_{ii}| > \textstyle\sum_{j\ne i} |a_{ij}|, \quad i = \overline{1,n}\,, \qquad \text{or} \qquad |a_{ii}| > \textstyle\sum_{i\ne j} |a_{ij}|, \quad j = \overline{1,n}\,,$$

respectively.

Theorem 6.3 *Let A be strictly diagonally dominant by rows. Then both Jacobi and Gauss–Seidel methods converge.*

Proof. We give a proof only for Gauss-Seidel. The Jacobi method is treated similarly (and simpler). We set $e^{(k)} = x^{(k)} - x^*$, so that

$$e_i^{(k+1)} = -\textstyle\sum_{j<i} \frac{a_{ij}}{a_{ii}}\, e_j^{(k+1)} - \sum_{j>i} \frac{a_{ij}}{a_{ii}}\, e_j^{(k)}\,,$$

and notice that the strict diagonal dominance by rows implies

$$\textstyle\sum_{j\ne i} \left|\frac{a_{ij}}{a_{ii}}\right| \le \gamma < 1.$$

For the 1-st component of $e^{(k+1)}$ we have the inequality

$$|e_1^{(k+1)}| = \left|\textstyle\sum_{j\ne 1} \frac{a_{1j}}{a_{11}}\, e_j^{(k)}\right| \le \max_j |e_j^{(k)}| \textstyle\sum_{j\ne 1} \left|\frac{a_{1j}}{a_{11}}\right| \le \gamma\|e^{(k)}\|_\infty\,.$$

Assuming that the similar estimate is valid for the first $(i-1)$ components of $e^{(k+1)}$, i.e.,

$$|e_j^{(k+1)}| \le \gamma\,\|e^{(k)}\|_\infty, \qquad j = \overline{1, i-1},$$

we derive the same estimate for the i-th component

$$|e_i^{(k+1)}| = \left|\textstyle\sum_{j<i} \frac{a_{ij}}{a_{ii}}\, e_j^{(k+1)} + \sum_{j>i} \frac{a_{ij}}{a_{ii}}\, e_j^{(k)}\right| \le \|e^{(k)}\|_\infty \textstyle\sum_{j\ne i} \left|\frac{a_{ij}}{a_{ii}}\right| \le \gamma\|e^{(k)}\|_\infty.$$

Thus, $\|e^{(k+1)}\|_\infty := \max_i |e_i^{(k+1)}| \le \gamma\|e^{(k)}\|_\infty$, hence $\|e^{(k)}\|_\infty \le \gamma^k\|e^0\|_\infty \to 0$. $\blacksquare$

Remark 6.4 We did not use Theorem 6.1 explicitly, however we proved that, for a strictly diagonally dominant A, the Gauss-Seidel method implies $\|e^{k+1}\|_\infty \le \gamma\|e^k\|_\infty$, and this is true for any e^k. Since $e^{k+1} = H_{\rm GS}e^k$, where $H_{\rm GS}$ is the iteration matrix, this means the following: for a strictly diagonally dominant A,

$$\|H_{\rm GS}z\|_\infty \le \gamma\|z\|_\infty \quad \forall z \in \mathbb{R}^n \quad \Rightarrow \quad \|H_{\rm GS}\|_\infty \le \gamma < 1,$$

i.e., the condition of Theorem 6.1 is fulfilled.

6.2 Necessary and sufficient conditions for convergence

Theorem 6.5 *Let $H \in \mathbb{R}^{n\times n}$. Then $\lim\limits_{k\to\infty} H^k z = 0$ for all $z \in \mathbb{R}^n$ if and only if $\rho(H) < 1$.*

Proof. 1) We commence with the case $\rho(H) \geq 1$ and wish to demonstrate that, for some *real* z, the vector $H^k z$ need not tend to 0. Let λ be an eigenvalue of (the real) H, real or complex, such that $|\lambda| = \rho(H) \geq 1$, and let w be a corresponding eigenvector, i.e., $Hw = \lambda w$. Then $H^k w = \lambda^k w$, and

$$\|H^k w\|_\infty = |\lambda|^k \|w\|_\infty \geq \|w\|_\infty =: \gamma > 0. \tag{6.1}$$

If w is real, we choose $z = w$, hence $H^k z \geq \gamma$, and this cannot tend to zero.

If w is complex, then $w = u + \mathrm{i}v$ with some real vectore u, v. But then at least one of the sequences $(H^k u), (H^k v)$ does not tend to zero. For if both do, then also $H^k w = H^k u + \mathrm{i} H^k v \to 0$, and this contradicts (6.1). This completes the proof of the 'only if' part of the theorem.

2) Now we turn to the 'if' case, i.e., $\rho(H) < 1$, where we assume for simplicity that H possesses n linearly independent eigenvectors (w_j) such that $Hw_j = \lambda_j w_j$. Linear independence means that every $z \in \mathbb{R}^n$ can be expressed as a linear combination of the eigenvectors, i.e., there exist $(c_j) \in \mathbb{C}$ such that $z = \sum_{j=1}^n c_j w_j$. Thus,

$$H^k z = \textstyle\sum_{j=1}^n c_j \lambda_j^k w_j, \qquad k = 0, 1, \ldots$$

and since $|\lambda_j| \leq \rho(H) < 1$ we have $\lim_{k\to\infty} H^k z = 0$, as required. ∎

Theorem 6.6 *Suppose that x^*, a solution of $Ax = b$, satisfies $x^* = Hx^* + v$ and we are given the iterative scheme*

$$x^{k+1} = Hx^k + v \tag{6.2}$$

Then $x^k \to x^*$ for any choice of x^0 if an only if $\rho(H) < 1$.

Proof. Set $e^k := x^k - x^*$. Since $x^* = Hx^* + v$, substracting this equation from (6.2) we obtain $e^{k+1} = He^k$, hence $e^k = H^k e^0$. Now, by Theorem 6.5, $e^k \to 0$ for any e^0 iff $\rho(H) < 1$. ∎

Corollary 6.7 *Let A be positive definite. Then the method of simple iteration*

$$x^{k+1} = x^k - \tau(Ax^k - b)$$

converges if and only if $\tau \in (0, \frac{2}{\rho(A)})$.

Proof. Let $\{\lambda_i\}$ be the spectrum of A. Then the spectrum of the iteration matrix H_τ is

$$\sigma(H_\tau) = \sigma(I - \tau A) = \{1 - \tau\lambda_i\}$$

whence the spectral radius of H_τ has the value

$$\rho(H_\tau) = \max\left(|1 - \tau\lambda_{\min}|, |1 - \tau\lambda_{\max}|\right).$$

Since for a positive definite A we have $0 < \lambda_{\min} \leq \lambda_i \leq \lambda_{\max} = \rho(A)$, it follows that

$$\rho(H_\tau) < 1 \quad \Leftrightarrow \quad -1 < 1 - \tau\lambda_i < 1 \quad \Leftrightarrow \quad \tau \in (0, \tfrac{2}{\rho(A)}).$$ ∎

Remark 6.8 The sufficient condition of Theorem 6.1 is at first sight stronger than that of Theorem 6.5 because, for any nonsingular S and for any induced matrix norm, we have

$$\rho(H) = \rho(SHS^{-1}) \leq \|SHS^{-1}\|.$$

(The first equality holds for the entire spectra, the second one is due to Lemma 5.6.) However, they are equivalent: for any matrix H, and for any $\varepsilon > 0$ there exists an induced matrix norm $\|\cdot\|_*$ such that $\|H\|_* \leq \rho(H) + \varepsilon$ (however, one cannot take $\varepsilon = 0$).

6.3 Extras

Theorem 6.9 *Let A be strictly diagonally dominant by columns. Then both Jacobi and Gauss–Seidel methods converge.*

Proof. Here is a proof of the simpler Jacobi case. The iteration matrix is $H_{\rm J} := D^{-1}(L_0 + U_0)$, so with $S = D$ we obtain

$$\widetilde{H}_{\rm J} := S H_{\rm J} S^{-1} = (L_0 + U_0) D^{-1} = \begin{bmatrix} 0 & \frac{a_{12}}{a_{22}} & \cdots & \frac{a_{1n}}{a_{nn}} \\ \frac{a_{21}}{a_{11}} & 0 & \cdots & \frac{a_{2n}}{a_{nn}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{a_{11}} & \frac{a_{n2}}{a_{22}} & \cdots & 0 \end{bmatrix}.$$

Now, if we take the matrix ℓ_1-norm, the diagonal dominance of A by columns implies

$$\|\widetilde{H}_{\rm J}\|_1 := \max_j \textstyle\sum_i |\widetilde{h}_{ij}| = \max_j \sum_{i \neq j} |\frac{a_{ij}}{a_{jj}}| \le \gamma < 1,$$

and application of Theorem 6.1 yields the result.

For the Gauss-Seidel case, the proof again follows with $S = (L_0 + D)$, and the ℓ_1-norm. Hint: transpose the iteration matrix and use the equality $\|A\|_1 = \|A^T\|_\infty$ together with some consequences of Theorem 6.3. ∎

Remark 6.10 The complete proof of the 'if' case of Theorem 6.5 exploits the so-called Jordan normal form of the matrix H, namely $H = SJS^{-1}$, where J is a block diagonal matrix consisting of the Jordan blocks

$$J = \begin{bmatrix} \boxed{J_1} & & & \\ & \boxed{J_2} & & \\ & & \ddots & \\ & & & \boxed{J_r} \end{bmatrix}, \qquad J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}, \qquad J_i \in \mathbb{R}^{n_i \times n_i}, \qquad \textstyle\sum_i n_i = n.$$

To prove that $J_i^k \to 0$ if $|\lambda_i| < 1$, one can split $J_i = \lambda_i I + P$, notice that $P^\ell = 0$ for $\ell \ge n_i$, and evaluate the terms of expansion $(\lambda_i I + P)^k = \sum_{\ell=0}^{n_i - 1} \binom{k}{\ell} \lambda_i^{k-\ell} P^\ell$.

Remark 6.11 In the method of simple iteration, if we know the bound not only for $\rho(A) = \lambda_{\max} =: M$ but also for $\lambda_{\min} =: m$, then the choice

$$\tau_{\rm opt} := \frac{2}{M + m}$$

is optimal in the sense that it provides the best rate of convergence (of the method of simple iteration) on the class of SPD-matrices A with $\sigma(A) \in [m, M]$, namely

$$\|e^k\|_2 \le \gamma_{\rm opt}^k \|e^0\|_2, \qquad \gamma_{\rm opt} = \frac{M - m}{M + m} < 1\,.$$