Dr A. Hansen

Mathematical Tripos Part II: Michaelmas Term 2020

Numerical Analysis – Lecture 17

Definition 4.10 (Strictly diagonally dominant matrices) A matrix A is called strictly diagonally dominant by rows (resp. by columns) if

$$|a_{ii}| > \sum_{j\neq i} |a_{ij}|, \quad i = 1..n \qquad (\text{resp.} \quad |a_{jj}| > \sum_{i\neq j} |a_{ij}|, \quad j = 1..n\,).$$

From Gershgorin theorem, it follows that strictly diagonally dominant matrices are nonsingular.

Theorem 4.11 *If A is strictly diagonally dominant, then both the Jacobi and the Gauss-Seidel methods converge.*

Proof. For the Gauss-Seidel method, the eigenvalues of the iteration matrix $H_{\rm GS} = -(L_0+D)^{-1}U_0$ satisfy the equation

$$\det[H_{\rm GS} - \lambda I] = \det[-(L_0 + D)^{-1}U_0 - \lambda I] = 0 \quad \Rightarrow \quad \det[A_\lambda] := \det[U_0 + \lambda D + \lambda L_0] = 0$$

It is easy to see that if $A = L_0 + D + U_0$ is strictly diagonally dominant, then for $|\lambda| \geq 1$ the matrix $A_\lambda = \lambda L_0 + \lambda D + U_0$ is strictly diagonally dominant too, hence it is nonsingular, and therefore the equality $\det[A_\lambda] = 0$ is impossible. Thus $|\lambda| < 1$, hence convergence. The proof for the Jacobi method is the same. □

Theorem 4.12 (The Householder–John theorem) *If A and B are real matrices such that both A and $A-B-B^T$ are symmetric positive definite, then the spectral radius of $H = -(A - B)^{-1}B$ is strictly less than one.*

Proof. Let λ be an eigenvalue of H, so $H\boldsymbol{w} = \lambda\boldsymbol{w}$ holds, where $\boldsymbol{w} \neq \boldsymbol{0}$ is an eigenvector. (Note that both λ and $\boldsymbol{w}$ may have nonzero imaginary parts when H is not symmetric, e.g. in the Gauss–Seidel method.) The definition of H provides equality $-B\boldsymbol{w} = \lambda(A-B)\boldsymbol{w}$, and we note that $\lambda \neq 1$ since otherwise A would be singular (which it is not). Thus, we deduce

$$\overline{\boldsymbol{w}}^T B\boldsymbol{w} = \frac{\lambda}{\lambda - 1}\overline{\boldsymbol{w}}^T A\boldsymbol{w}, \tag{4.3}$$

where the bar means complex conjugation. Moreover, writing $\boldsymbol{w} = \boldsymbol{u}+i\boldsymbol{v}$, where $\boldsymbol{u}$ and $\boldsymbol{v}$ are real, we find (for $C = C^T$) the identity $\overline{\boldsymbol{w}}^T C\boldsymbol{w} = \boldsymbol{u}^T C\boldsymbol{u} + \boldsymbol{v}^T C\boldsymbol{v}$, so symmetric positive definiteness in the assumption implies $\overline{\boldsymbol{w}}^T A\boldsymbol{w} > 0$ and $\overline{\boldsymbol{w}}^T(A - B - B^T)\boldsymbol{w} > 0$. In the latter inequality, we use relation (4.3) and its conjugate transpose to obtain

$$0 < \overline{\boldsymbol{w}}^T A\boldsymbol{w} - \overline{\boldsymbol{w}}^T B\boldsymbol{w} - \overline{\boldsymbol{w}}^T B^T\boldsymbol{w} = \left(1 - \frac{\lambda}{\lambda - 1} - \frac{\overline{\lambda}}{\overline{\lambda} - 1}\right)\overline{\boldsymbol{w}}^T A\boldsymbol{w} = \frac{1 - |\lambda|^2}{|\lambda - 1|^2}\,\overline{\boldsymbol{w}}^T A\boldsymbol{w}.$$

Now $\lambda \neq 1$ implies $|\lambda - 1|^2 > 0$. Hence, recalling that $\overline{\boldsymbol{w}}^T A\boldsymbol{w} > 0$, we see that $1 - |\lambda|^2$ is positive. Therefore $|\lambda| < 1$ occurs for every eigenvalue of H as required. □

Corollary 4.13 1) *If A is symmetric positive definite, then the Gauss-Seidel method converges.*
2) *If both A and $2D-A$ are symmetric positive definite, then the Jacobi method converges.*

Proof. 1) For the Gauss-Seidel method, B is the superdiagonal part of symmetric A, hence $A-B-B^T$ is equal to D, the diagonal part of A, and if A is positive definite, then D is positive definite too (this is the first part of the Exercise 23 from Example Sheets).

2) For the Jacobi method, we have $B = A-D$, and if A is symmetric, then $A-B-B^T = 2D-A$. (The latter matrix is the same as A except that the signs of the off-diagonal elements are reversed.) □

Example 4.14 (Poisson's equation on a square) As we have seen in the previous sections linear systems $A\boldsymbol{x} = \boldsymbol{b}$, where A is a real symmetric positive (negative) definite matrix, frequently occur in numerical methods for solving elliptic partial differential equations. A typical example we already encountered is Poisson's equation on a square where the *five-point formula* approximation yields an $n \times n$ system of linear equations with $n = m^2$ unknowns $u_{p,q}$:

$$u_{p-1,q} + u_{p+1,q} + u_{p,q-1} + u_{p,q+1} - 4u_{p,q} = h^2 f(ph, qh) \tag{4.4}$$

(Note that when p or q is equal to 1 or m, then the values $u_{0,q}$, $u_{p,0}$ or $u_{p,m+1}$, $u_{m+1,q}$ are known boundary values and they should be moved to the right-hand side, thus leaving fewer unknowns on the left.)

For any ordering of the grid points (ph, qh) we have shown in Lemma 1.11 that the matrix A of this linear system is symmetric and negative definite.

Corollary 4.15 *For linear system (4.4), for any ordering of the grid, both Jacobi and Gauss-Seidel methods converge.*

Proof. By Lemma 1.11, A is symmetric and negative definite, hence convergence of Gauss-Seidel. To prove convergence of the Jacobi method, we need negative definiteness of the matrix $2D - A$, and that follows by the same arguments as in Lemma 1.11: recall that the proof operates with the modulus of the off-diagonal elements and does not depend on their sign. □

Method 4.16 (Relaxation) It is often possible to improve the efficiency of the splitting method by *relaxation*. Specifically, instead of letting $(A - B)\boldsymbol{x}^{(k+1)} = -B\boldsymbol{x}^{(k)} + \boldsymbol{b}$, we let

$$(A - B)\widehat{\boldsymbol{x}}^{(k+1)} = -B\boldsymbol{x}^{(k)} + \boldsymbol{b}, \quad \text{and then} \quad \boldsymbol{x}^{(k+1)} = \omega\widehat{\boldsymbol{x}}^{(k+1)} + (1-\omega)\boldsymbol{x}^{(k)} \qquad k = 0, 1, \ldots,$$

where ω is a real constant called the *relaxation parameter*. (Note that $\omega = 1$ corresponds to the standard "unrelaxed" iteration.) Good choice of ω leads to a smaller spectral radius of the iteration matrix (compared with the "unrelaxed" method), and the smaller the spectral radius, the faster the iteration converges. To this end, let us express the relaxation iteration matrix H_ω in terms of $H = -(A - B)^{-1}B$. We have

$$\widehat{\boldsymbol{x}}^{(k+1)} = H\boldsymbol{x}^{(k)} + \boldsymbol{v} \quad \Rightarrow \quad \boldsymbol{x}^{(k+1)} = \omega\widehat{\boldsymbol{x}}^{(k+1)} + (1-\omega)\boldsymbol{x}^{(k)} = \omega H\boldsymbol{x}^{(k)} + (1-\omega)\boldsymbol{x}^{(k)} + \omega\boldsymbol{v}$$

hence

$$H_\omega = \omega H + (1-\omega)I.$$

It follows that the spectra of H_ω and H are related by the rule $\lambda_\omega = \omega\lambda + (1-\omega)$, therefore one may try to choose $\omega \in \mathbb{R}$ to minimize

$$\rho(H_\omega) = \max\{|\omega\lambda + (1-\omega)| : \lambda \in \sigma(H)\}.$$

In general, $\sigma(H)$ is unknown, but often we have some information about it which can be utilized to find a "good" (rather than "best") value of ω. For example, suppose that it is known that $\sigma(H)$ is real and resides in the interval $[\alpha, \beta]$ where $-1 < \alpha < \beta < 1$. In that case we seek ω to minimize

$$\max\{|\omega\lambda + (1-\omega)| : \lambda \in [\alpha, \beta]\}.$$

It is readily seen that, for a fixed $\lambda < 1$, the function $f(\omega) = \omega\lambda + (1-\omega)$ is decreasing, therefore, as ω increases (decreases) from 1 the spectrum of H_ω moves to the left (to the right) of the spectrum of H. It is clear that the optimal location of the spectrum $\sigma(H_\omega)$ (or of the interval $[\alpha_\omega, \beta_\omega]$ that contains $\sigma(H_\omega)$) is the one which is centralized around the origin:

$$-[\omega\alpha + (1-\omega)] = \omega\beta + (1-\omega) \quad \Rightarrow \quad \omega_{\text{opt}} = \tfrac{2}{2-(\alpha+\beta)}, \quad -\alpha_{\omega_{\text{opt}}} = \beta_{\omega_{\text{opt}}} = \tfrac{\beta-\alpha}{2-(\alpha+\beta)}.$$