Dr C.-B. Schönlieb

Mathematical Tripos Part II: Michaelmas Term 2015

Numerical Analysis – Lecture 16

4 Iterative methods for linear algebraic systems

Technique 4.1 (Splitting) For $n \geq 0$ let $A \in \mathbb{R}^{n\times n}$, $b, x \in \mathbb{R}^n$. One way of solving a linear system $A\boldsymbol{x} = \boldsymbol{b}$ is to write it in the form

$$(A - B)\,\boldsymbol{x} = -B\boldsymbol{x} + \boldsymbol{b},$$

where the matrix B is chosen in such a way that it is relatively easy to solve the system $(A-B)\,\boldsymbol{x} = \boldsymbol{y}$ for any given $\boldsymbol{y}$ (this also requires $A - B$ to be nonsingular). Then the iteration commences with an estimate $\boldsymbol{x}^{(0)}$ of the required solution, and generates the sequence $(\boldsymbol{x}^{(k)})$ by solving

$$(A - B)\,\boldsymbol{x}^{(k+1)} = -B\boldsymbol{x}^{(k)} + \boldsymbol{b}, \qquad k = 0, 1, 2, \ldots \tag{4.1}$$

If the sequence converges to a limit, $\boldsymbol{x}^{(k)} \to \boldsymbol{x}^*$, say, then the limit has the property $(A - B)\boldsymbol{x}^* = -B\boldsymbol{x}^* + \boldsymbol{b}$. Therefore $\boldsymbol{x}^*$ is a solution of $A\boldsymbol{x} = \boldsymbol{b}$ as required.

Discussion 4.2 (Conditions for convergence) Suppose that $A - B$ is nonsingular and let $H := -(A - B)^{-1}B$ the *iteration matrix* and $\mathbf{c} := (A - B)^{-1}\mathbf{b}$, hence $\mathbf{x}^{(k+1)} = H\mathbf{x}^{(k)} + \mathbf{c}$. Define $\boldsymbol{x}^*$ by $A\boldsymbol{x}^* = \boldsymbol{b}$. Thus, since $(A - B)\mathbf{x}^* = -B\mathbf{x}^* + \mathbf{b}$, we have $\mathbf{x}^* = H\mathbf{x}^* + \mathbf{c}$. Then, denoting by $\boldsymbol{e}^{(k)} := \boldsymbol{x}^{(k)} - \boldsymbol{x}^*$ the *error* of the k-th iteration and subtracting equations, we have

$$\boldsymbol{e}^{(k+1)} = H\boldsymbol{e}^{(k)} = \cdots = H^{k+1}\boldsymbol{e}^{(0)}\,.$$

In other words, if $H^k \overset{k\to\infty}{\longrightarrow} O$ then the iteration converges for all $\mathbf{e}^{(0)}$.

Lemma 4.3 *Suppose that H has n linearly independent eigenvectors and that the moduli of all its eigenvalues are less than one. Then $\mathbf{e}^{(k)} \overset{k\to\infty}{\longrightarrow} \mathbf{0}$.*

Proof. Let $H\mathbf{v}_j = \lambda_j\mathbf{v}_j$, $j = 1, \ldots, n$, where $\|\mathbf{v}_j\| = 1$ (the Euclidean length). We can expand every vector in R^n as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Thus, $\mathbf{e}^{(0)} = \sum_{j=1}^n \theta_j\mathbf{v}_j$, say. But $H^k\mathbf{v}_j = \lambda_j^k\mathbf{v}_j$, hence $\mathbf{e}^{(k)} = \sum_{j=1}^n \theta_j\lambda_j^k\mathbf{v}_j$. Consequently,

$$\|\mathbf{e}^{(k)}\| \leq \sum_{j=1}^n \left\|\theta_j\lambda_j^k\mathbf{v}_j\right\| = \sum_{j=1}^n |\theta_j\lambda_j^k| \leq \left(\sum_{j=1}^n |\theta_j|\right) \max_{j=1,\ldots,n} |\lambda_j^k|.$$

Since $|\lambda_j| < 1$, $j = 1, \ldots, n$, it follows that $\mathbf{e}^{(k)} \to \mathbf{0}$. □

Lemma 4.4 *If the modulus of an eigenvalue of H is ≥ 1 then there exists a choice of $\mathbf{x}^{(0)}$ s.t. the iteration fails to converge.*

Proof. We use the notation from the last proof and assume that $|\lambda_1| > 1$. If λ_1 is complex then $H\bar{\mathbf{v}}_1 = \bar{\lambda}_1\bar{\mathbf{v}}_1$. We choose $\mathbf{e}^{(0)} = \frac{1}{2}(\mathbf{v}_1 + \bar{\mathbf{v}}_1)$ (clearly, if λ_j is real then $\mathbf{e}^{(0)} = \mathbf{v}_1$). Therefore $\mathbf{e}^{(k)} = \frac{1}{2}(\lambda_1^k\mathbf{v}_1 + \bar{\lambda}_1^k\bar{\mathbf{v}}_1)$. Except for the case $\lambda_1 = 1$ (when $\mathbf{e}^{(k)} = \mathbf{v}_1$) this precludes convergence. Finally, in the case $\lambda_1 = 1$[1] we choose $\mathbf{x}^{(0)} = \mathbf{c} = \mathbf{v}_1$. It follows by induction that $\mathbf{x}^{(k)} = (k+1)\mathbf{v}_1$ – again, no convergence. □

Theorem 4.5 *The iteration $\mathbf{x}^{(k+1)} = H\mathbf{x}^{(k)} + \mathbf{c}$, $k \in Z^+$, converges to $\mathbf{x}^*$ for every choice of $\mathbf{x}^{(0)}$ and $\mathbf{c}$ if and only if all the moduli of eigenvalues of H are less than 1.*

[1] This may happen only if A is singular – verify!

Proof. If there are n independent eigenvalues, the theorem follows from the last two lemmas. Otherwise we can use the Jordan canonical form, but we leave out this analysis. □

Remark 4.6 *In other words Theorem 4.5 states that convergence $\boldsymbol{e}^{(k)} \to 0$ is achieved for any choice of $\boldsymbol{x}^{(0)}$ if and only if H has the property $\rho(H) < 1$. Here $\rho(H)$ is the spectral radius of H, which means the largest modulus of an eigenvalue of H. (Some of the eigenvalues may have nonzero imaginary parts.)*

Example 4.7 *An example of the situation mentioned above is when*

$$H = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

In that case

$$H^k = \begin{bmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{bmatrix},$$

therefore $H^k \to O$ iff $|\lambda| < 1$, as claimed in the theorem.

Note: *There is more to iteration than just 'convergence'! For example, we achieve convergence with*

$$H = \begin{bmatrix} 0.99 & 10^6 & 10^{12} \\ 0 & 0.99 & 10^{20} \\ 0 & 0 & 0.99 \end{bmatrix},$$

but it will take quite a long time...

Method 4.8 (Jacobi and Gauss–Seidel) Both of these methods are versions of splitting which can be applied to any A with nonzero diagonal elements. We write A as the sum of three matrices $L_0 + D + U_0$: subdiagonal (strictly lower-triangular), diagonal and superdiagonal (strictly upper-triangular) portions of A, respectively.

1) *Jacobi method*. We set $A - B = D$, the diagonal part of A, and we obtain the next iteration by solving the diagonal system $D\boldsymbol{x}^{(k+1)} = -(L_0 + U_0)\boldsymbol{x}^{(k)} + \boldsymbol{b}$.

2) *Gauss–Seidel method*. We take $A - B = L_0 + D = L$, the lower-triangular part of A, and we generate the sequence $(\boldsymbol{x}^{(k)})$ by solving the triangular system $(L_0 + D)\,\boldsymbol{x}^{(k+1)} = -U_0\boldsymbol{x}^{(k)} + \boldsymbol{b}$. There is no need to invert $(L_0 + D)$, we calculate the components of $\boldsymbol{x}^{(k+1)}$ in sequence by forward substitution:

$$a_{ii}x_i^{(k+1)} = -\sum_{j<i} a_{ij}x_j^{(k+1)} - \sum_{j>i} a_{ij}x_j^{(k)} + b_i, \qquad i = 1..n.$$

As we mentioned above, the sequence $\boldsymbol{x}^{(k)}$ converges to solution of $A\boldsymbol{x} = \boldsymbol{b}$ if the spectral radius of the iteration matrix, $H_{\mathrm{J}} = -D^{-1}(L_0 + U_0)$ or $H_{\mathrm{GS}} = -(L_0 + D)^{-1}U_0$, respectively, is less than one. Our next goal is to prove that this is the case for two important classes of matrices A: a) diagonally dominant and b) positive definite matrices. We start with recalling the simple, but very useful Gershgorin theorem.

Revision 4.9 (Gershgorin theorem) *All eigenvalues of an $n \times n$ matrix A are contained in the union of the Gershgorin discs in the complex plane:*

$$\sigma(A) \subset \cup_{i=1}^n \Gamma_i\,, \qquad \Gamma_i := \{z \in \mathbb{C} : |z - a_{ii}| \le r_i\}, \qquad r_i := \sum_{j \ne i} |a_{ij}|\,.$$