NLA 2021-2022
Iterative methods for linear equation solving

Martin Sombra

xx xxx 2021

Basic iterative methods

Let A be a nonsingular $n \times n$ matrix and b an n-vector

Given an *initial n-vector* x_0, these methods generate a sequence of other n-vectors

$$(x_l)_{l \geqslant 0}$$

hopefully converging to the solution $x = A^{-1}\,b$ and where for each $l \geqslant 0$, the vector x_{l+1} is easy to compute from x_l

Splittings

A *splitting* of A is a decomposition

$$A = M - K$$

with M nonsingular

$\rightsquigarrow$ iterative method:

$$Ax = b \iff Mx = Kx + b \iff x = M^{-1}Kx + M^{-1}b$$

We then set

$$x_{l+1} = R\,x_l + c \qquad \text{for } l \geqslant 0 \tag{1}$$

A convergence criterion

The *spectral radius*

$$\rho(R)$$

is the maximum absolute value of the eigenvalues of R

The iteration in (1) converges for every choice of initial vector x_0 if and only if

$$\rho(R) < 1$$

A convergence criterion (cont.)

Indeed, for every $\varepsilon > 0$ there is a vector norm $\|\cdot\|$ such that the associated operator norm verifies that

$$\|R\| \leqslant \rho(R) + \varepsilon$$

Hence if $\rho(R) < 1$ then we can choose ε sufficiently small so that the associated vector norm $\|\cdot\|$ verifies that $\|R\| < 1$

Then

$$\begin{array}{rrcl} & x & = & R\,x + c \\ - & x_{l+1} & = & R\,x_l + c \\ \hline & \multicolumn{3}{c}{x - x_{l+1} = R\,(x - x_l)} \end{array}$$

and so in this case

$$\|x - x_{l+1}\| = \|R\,(x - x_l)\| \leqslant \|R\|\,\|x - x_l\| \leqslant \|R\|^{l+1}\,\|x - x_0\| \longrightarrow 0$$

when $l \to +\infty$

A convergence criterion (cont.)

Conversely, if $\rho(R) \geqslant 1$ then there is $x_0 \neq x$ such that $x - x_0$ is an eigenvector for an eigenvalue λ of absolute value $\geqslant 1$

Hence

$$x - x_{l+1} = R^{l+1}(x - x_0) = \lambda^{l+1}(x - x_{l+1})$$

which does not approach 0 when $k \to +\infty$

The speed of convergence

When $\rho(R) < 1$ the approximation of the l-th iteration is

$$-\log_b \|x - x_l\| \geqslant -l \log_b \rho(R) - \log_b \|x - x_0\|$$

$\rightsquigarrow$ the increase in precision at each step is *linear* with rate

$$-\log_b \rho(R)$$

Our goal is to fing a splitting $A = M - K$ verifying the conditions

- $R = M^{-1}K$ and $c = M^{-1}b$ are easy to compute
- $\rho(R)$ is small

The classical iterative methods

Suppose that no diagonal entry of A is zero and set

$$A = D - \widetilde{L} - \widetilde{U} = D\left(\mathbb{1}_n - L - U\right)$$

with

- D is diagonal
- $\widetilde{L}$ and L strictly lower triangular
- $\widetilde{U}$ and U are strictly upper triangular

Jacobi's method

Jabobi's method can be interpreted as going successively through the equations, changing at the j-th step of the $(l+1)$th iteration the variable x_j so that the j-th equation is satisfied:

$$\begin{aligned}&\text{for } j = 1, \ldots, n\\ &\qquad x_{l+1,j} \leftarrow \tfrac{1}{a_{j,j}} \left(b_j - \textstyle\sum_{k \neq j} a_{j,k}\, x_{l,k}\right)\end{aligned}$$

Hence for all j we have that

$$a_{j,1}\, x_{l,1} + \cdots + a_{j,j-1}\, x_{l,j-1} + a_{j,j}\, x_{l+1,j} + a_{l,j+1}\, x_{l,j+1} + \cdots + a_{l,n}\, x_{l,n} = b_j$$

In matrix notation, $x_{l+1} = R_{\mathrm{J}}\, x_l + c_{\mathrm{J}}$ with

$$R_{\mathrm{J}} = D^{-1}(\widetilde{L} + \widetilde{U}) = L + U \quad \text{and} \quad c_{\mathrm{J}} = D^{-1} b$$

Gauss-Seidel method

The *Gauss-Seidel method* takes advantage of the previously computed coordinates $x_{l+1,k}$, $k = 1, \ldots, j-1$:

$$\text{for } j = 1, \ldots, n$$

$$x_{l+1,j} \leftarrow \frac{1}{a_{j,j}} \Big(b_j - \underbrace{\sum_{k<j} a_{j,k}\, x_{l+1,k}}_{\text{updated } x\text{'s}} - \underbrace{\sum_{k>j} a_{j,k}\, x_{l,k}}_{\text{older } x\text{'s}} \Big)$$

Hence for all j

$$a_{j,1}\, x_{l+1,1} + \cdots + a_{j,j-1}\, x_{l+1,j-1} + a_{j,j}\, x_{l+1,j} + a_{l+1,j+1}\, x_{l,j+1} + \cdots + a_{l,n}\, x_{l,n} = b_j$$

In matrix notation, $x_{l+1} = R_{\mathrm{GS}}\, x_l + c_{\mathrm{GS}}$ with

$$R_{\mathrm{GS}} = (D - \widetilde{L})^{-1} \widetilde{U} = (\mathbb{1}_n - L)^{-1} U$$

$$c_{\mathrm{GS}} = (D - \widetilde{L})^{-1} b = (\mathbb{1}_n - L)^{-1} D^{-1} b$$

SOR

Successive overrelaxation for a parameter $\omega \in \mathbb{R}$ is a weighted average of the vectors x_{l+1} and x_l from the Gauss-Seidel method:

$$x_{l+1}^{\mathrm{SOR}(\omega)} = (1-\omega)\, x_l^{\mathrm{GS}} + \omega x_{l+1}^{\mathrm{GS}}$$

for $j = 1, \ldots, n$

$$x_{l+1,j} \leftarrow (1-\omega) x_{l,j} + \tfrac{\omega}{a_{j,j}} \left(b_j - \sum_{k<j} a_{j,k}\, x_{l+1,k} - \sum_{k>j} a_{j,k}\, x_{l,k} \right)$$

In matrix notation, $x_{l+1} = R_{\mathrm{SOR}(\omega)}\, x_l + c_{\mathrm{SOR}(\omega)}\, b$ with

$$\begin{aligned} R_{\mathrm{SOR}(\omega)} &= (D - \omega\, \widetilde{L})^{-1} ((1-\omega)\, D + \omega\, \widetilde{U}) \\ &= (\mathbb{1}_n - \omega\, L)^{-1} ((1-\omega)\, \mathbb{1}_n + \omega\, U) \\ c_{\mathrm{SOR}(\omega)} &= \omega\, (D - \omega\, \widetilde{L}) = \omega\, (\mathbb{1}_n - \omega\, L)\, D^{-1} \end{aligned}$$

- $\omega = 1$: Gauss-Seidel method
- $\omega > 1$: *overrelaxation*
- $\omega < 1$: *underrelaxation*

Example

Consider the 2×2 matrix and the 2-vector

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

The solution of $A\,x = b$ is the 2-vector

$$x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

The decompositions $A = D - \widetilde{L} - \widetilde{U} = D\,(\mathbb{1}_2 - L - U)$ are

$$\begin{aligned} A &= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ \frac{-1}{2} & 0 \end{bmatrix} - \begin{bmatrix} 0 & \frac{-1}{2} \\ 0 & 0 \end{bmatrix} \right). \end{aligned}$$

Example (cont.)

Hence

$$R_{\mathrm{J}} = L + U = \begin{bmatrix} 0 & \frac{-1}{2} \\ \frac{-1}{2} & 0 \end{bmatrix} \quad \text{and} \quad c_{\mathrm{J}} = D^{-1} b = \begin{bmatrix} \frac{1}{2} \\ \frac{-1}{2} \end{bmatrix}$$

and so the corresponding Jacobi iteration writes as

$$\begin{bmatrix} x_{l+1,1} \\ x_{l+1,2} \end{bmatrix} = \begin{bmatrix} 0 & \frac{-1}{2} \\ \frac{-1}{2} & 0 \end{bmatrix} \begin{bmatrix} x_{l,1} \\ x_{l,2} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ \frac{-1}{2} \end{bmatrix} = \begin{bmatrix} \frac{-x_{l,2}}{2} + \frac{1}{2} \\ \frac{x_{l,1}}{2} - \frac{1}{2} \end{bmatrix}$$

The characteristic polynomial of the Jacobi matrix is

$$\chi_{R_{\mathrm{J}}} = \det(R_{\mathrm{J}} - t\, \mathbb{1}_2) = \det \begin{bmatrix} -t & \frac{-1}{2} \\ \frac{-1}{2} & -t \end{bmatrix} = t^2 - \frac{1}{4}$$

Hence $\lambda(R_{\mathrm{J}}) = \{t \mid \chi_{R_{\mathrm{J}}} = 0\} = \left\{ \pm \frac{1}{2} \right\}$ and so the spectral radius is

$$\rho(R_{\mathrm{J}}) = \frac{1}{2}$$

$\rightsquigarrow$ the method converges to x for every choice of initial vector x_0

Example (cont.)

We also have that

$$R_{\mathrm{GS}} = (\mathbb{1}_2 - L)^{-1} U = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & \frac{-1}{2} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{-1}{2} \\ 0 & \frac{1}{4} \end{bmatrix}$$

$$c_{\mathrm{GS}} = (D - \widetilde{L})^{-1}\, b = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{-1}{4} & \frac{1}{2} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{-3}{4} \end{bmatrix}$$

and so the Gauss-Seidel iteration writes as

$$\begin{bmatrix} x_{l+1,1} \\ x_{l+1,2} \end{bmatrix} = \begin{bmatrix} 0 & \frac{-1}{2} \\ 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} x_{l,1} \\ x_{l,2} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ \frac{-3}{4} \end{bmatrix} = \begin{bmatrix} \frac{-x_{l,2}}{2} + \frac{1}{2} \\ \frac{x_{l,2}}{4} - \frac{3}{4} \end{bmatrix}$$

Example (cont.)

We have that

$$\chi_{R_{\mathrm{GS}}} = \det(R_{\mathrm{GS}} - t\,\mathbb{1}_2) = \det \begin{bmatrix} -t & \frac{1}{2} \\ 0 & -t + \frac{1}{4} \end{bmatrix} = t^2 - \frac{1}{4}\,t$$

Hence $\lambda(R_{\mathrm{GS}}) = \{0, \frac{1}{4}\}$ and so $\rho(R_{\mathrm{GS}}) = \frac{1}{4}$

In this example

$$\rho(R_{\mathrm{GS}}) = \rho(R_{\mathrm{J}})^2$$

and so the Gauss-Seidel method converges with the *double of the speed* of the Jacobi method

Example (cont.)

For $\omega \in \mathbb{R}$

$$R_{\mathrm{SOR}(\omega)} = (\mathbb{1}_2 - \omega\, L)^{-1}((1-\omega)\,\mathbb{1}_2 + \omega\, U)$$
$$= \begin{bmatrix} 1 & 0 \\ \omega/2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1-\omega & -\omega/2 \\ 0 & 1-\omega \end{bmatrix} = \begin{bmatrix} 1-\omega & -\frac{\omega}{2} \\ \frac{\omega^2}{2} - \frac{\omega}{2} & \frac{\omega^2}{4} - \omega + 1 \end{bmatrix}$$

and

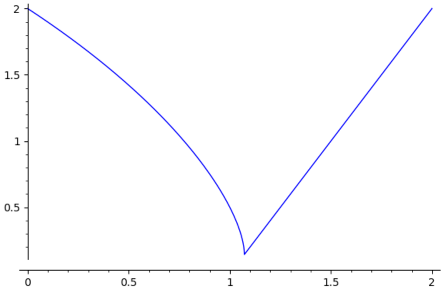
$$c_{\mathrm{SOR}(\omega)} = \omega\,(D - \omega\,\widetilde{L})^{-1} b = \omega \begin{bmatrix} 2 & 0 \\ \frac{\omega}{2} & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{\omega}{2} \\ \frac{-\omega^2}{4} - \frac{\omega}{2} \end{bmatrix}$$

Hence

$$\chi_{R_{\mathrm{SOR}(\omega)}} = \det(R_{\mathrm{SOR}(\omega)} - t\,\mathbb{1}_2) = \det \begin{bmatrix} 1-\omega-t & -\frac{\omega}{2} \\ \frac{\omega^2}{2} - \frac{\omega}{2} & \frac{\omega^2}{4} - \omega + 1 - t \end{bmatrix}$$
$$= t^2 + \left(\frac{-\omega^2}{4} + 2\,\omega - 2\right) t + (\omega^2 - 2\,\omega + 1).$$

Example (cont.)

The spectral radius $R_{\mathrm{SOR}(\omega)}$ for $\omega \in [0,2]$:



The optimal value is $\omega_{\mathrm{opt}} = 1.0717$ and we have that

$$\rho(R_{\mathrm{SOR}(1.0717)}) = 0.1535$$

Convergence of the basic iterative schemes

Consider a splitting $A = M - K$ with M nonsingular and the associated iterative scheme

$$x_{l+1} \leftarrow R\, x_l \qquad \text{with } R = M^{-1}K \text{ and } c = M^{-1}b$$

If it converges, its limit $x_\infty = \lim_{l\to\infty} x_l$ satisfies

$$x_\infty = R\, x_\infty + c = M^{-1}K\, x_\infty + M^{-1}b$$

which implies that x_∞ is the (unique) solution of the equation $A\, x = b$

Convergence of the basic iterative schemes (cont.)

The convergence of the iteration for an arbitrary initial vector x_0 is equivalent to the condition that

$$\rho(R) < 1$$

In this case, the rate of convergence of the method is linear with coefficient

$$-\log_b \rho(R)$$

In practice, the spectral radius is difficult to compute: need sufficient conditions that are easy to apply!

Example

The *Richardson iteration* for a parameter $\omega \in \mathbb{R}$ is

$$x_{l+1} \leftarrow R_\omega \, x_l + \omega \, b$$

with $R_\omega = \mathbb{1}_n - \omega \, A$. It corresponds to the splitting

$$A = \omega^{-1} \, \mathbb{1}_n - (\omega^{-1} \, \mathbb{1}_n - A)$$

Suppose that the eigenvalues of A are real and ordered as

$$\lambda_1 \geqslant \cdots \geqslant \lambda_n$$

Then the eigenvalues of R_ω are also real and satisfy that

$$\mu_1 = 1 - \omega \, \lambda_1 \leqslant \cdots \leqslant \mu_n = 1 - \omega \, \lambda_n$$

If $\lambda_n \leqslant 0$ then $\mu_n \geqslant 1$ and so $\rho(R_\omega) \geqslant 1$ and the iteration does not converge

Example (cont.)

Else suppose that $\lambda_n \geqslant 0$. Then the iteration converges if and only if

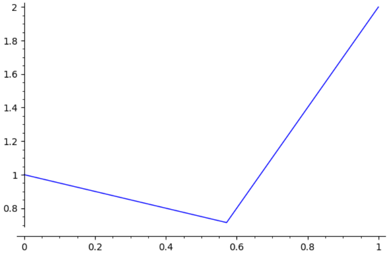
$$-1 \leqslant 1 - \omega\,\lambda_1 \quad \text{and} \quad 1 - \omega\,\lambda_n \leqslant 1$$

that is, if and only if

$$0 < \omega < \frac{2}{\lambda_1}$$

The spectral radius is the piecewise affine function

$$\rho(R_\omega) = \max(|1 - \omega\,\lambda_n|, |1 - \omega\,\lambda_1|)$$



Example (cont.)

The optimal value is given by $-1 + \lambda_1\,\omega_{\rm opt} = 1 - \lambda_n\omega_{\rm opt}$ and so

$$\omega_{\rm opt} = \frac{2}{\lambda_1 + \lambda_n}$$

The corresponding spectral radius is

$$\rho(\omega_{\rm opt}) = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}$$

- If A has both very large and very small eigenvalues, then the iteration will be slow
- The determination of the optimal parameter $\omega_{\rm opt}$ requires knowledge of these eigenvalues

Diagonally dominant matrices

The matrix A is *diagonally dominant* if for all j

$$|a_{j,j}| > \sum_{i \neq j} |a_{i,j}|$$

If A is diagonally dominant then both the Jacobi and the Gauss-Seidel iterations will converge for every choice of initial vector x_0

Diagonally dominant matrices (cont.)

For instance, let

$$R_{\mathrm{J}} = D^{-1}(\widetilde{L} + \widetilde{U})$$

be the iteration matrix of Jacobi's method

For an eigenvalue $\lambda \in \lambda(R_{\mathrm{J}})$ let x be a corresponding eigenvector and m the index of its largest component, normalized so that

$$x_m = 1 \quad \text{and} \quad |x_j| \leqslant 1 \text{ for all } j$$

Then from $R_{\mathrm{J}}\, x = \lambda\, x$ we deduce that

$$\lambda\, x_m = -\sum_{j \neq m} \frac{a_{m,j}}{a_{m,m}}\, x_j$$

and so

$$|\lambda| \leqslant \sum_{j \neq m} \frac{|a_{m,j}|}{|a_{m,m}|}\, |x_j| < 1$$

which proves the result for the Jacobi iteration

Convergence of SOR

For successive overrelaxation, the condition $0 < \omega < 2$ is necessary for the convergence of the method

Indeed

$$R_{\mathrm{SOR}(\omega)} = (\mathbb{1}_n - \omega\, L)^{-1}((1-\omega)\,\mathbb{1}_n + \omega\, U)$$

and so $\chi_{R_{\mathrm{SOR}(\omega)}}(0) = \det(R_{\mathrm{SOR}(\omega)}) = (1-\omega)^n$

On the other hand

$$|\chi_{R_{\mathrm{SOR}(\omega)}}(0)| = \prod_{\lambda} |\lambda| \geqslant \rho(R_{\mathrm{SOR}(\omega)})^n$$

the product being over the eigenvalues of $R_{\mathrm{SOR}(\omega)}$, and so

$$\rho(R_{\mathrm{SOR}(\omega)}) \geqslant |1-\omega|$$

Convergence of SOR

When A is symmetric and positive definite,

$$\mathrm{SOR}(\omega)$$

converges for every $0 < \omega < 2$

In particular, in this situation the Gauss-Seidel method ($\omega = 1$) also converges