

MTH 451-551/Peszynska, Nov. 2017
Stationary iterative methods for solving $Ax = b$

Consider a nonsingular square matrix $A \in \mathbb{R}^{n \times n}$, a right-hand side vector $b \in \mathbb{R}^n$, and a solution to

$$Ax = b. \tag{1}$$

If A is sparse and has some additional properties (see below), then one can solve (1) by iteration, with fewer flops than a direct solver would require.

An iteration proceeds from an initial guess for the solution $x^{(0)}$, as follows

$$x^{(0)} \mapsto x^{(1)} \mapsto \dots x^{(k-1)} \mapsto x^{(k)} \mapsto \dots, \tag{2}$$

until it satisfies some stopping criteria, say for some $k = k_{STOP}$. When the iteration stops, the iterate $x^{(k_{STOP})}$ should be close to the true solution x of (1).

There are two classes of iterative methods: stationary and non-stationary.

A stationary iterative method for solving has a form: given an initial guess $x^{(0)}$, for $k = 1, 2, \dots$ until convergence, find

$$x^{(k)} = Gx^{(k-1)} + c \tag{3}$$

where G is defined depending on A , is different for every iterative method, but does not change between iterations.

The key is to rearrange (1) in the form

$$x = Gx + c, \tag{4}$$

that would ensure convergence, and be efficient.

1 General information

A method is *consistent* if (3) is satisfied by the solution x to (1).

The *error* equation is derived by subtracting (3) from (4) to find that the error in k 'th iteration $e_k := x - x^{(k)}$ satisfies

$$x - x^{(k)} = G(x - x^{(k-1)}) \tag{5}$$

From this it follows that

$$x - x^{(k)} = G^k(x - x^{(0)}) \tag{6}$$

or that

$$\|x - x^{(k)}\| \leq \|G\|^k \|x - x^{(0)}\| \tag{7}$$

where $\|\cdot\|$ is a given vector norm and where $\|G\|$ is evaluated by the associated subordinate norm.

We find then that for the iteration to converge, that is, for $x^{(k)} \rightarrow x$ as $k \rightarrow \infty$ we have to have that $\rho(G) < 1$ or (a stronger condition) that $\|G\| < 1$.

Next, we find out how to estimate the current error without knowing x . We use (7) and triangle inequality to derive

$$\|x - x^{(k)}\| \leq \|G\| \left(\|x - x^{(k)}\| + \|x^{(k)} - x^{(k-1)}\| \right)$$

from which (and if $\|G\| < 1$), it follows that

$$\|x - x^{(k)}\| \leq \frac{\|G\|}{1 - \|G\|} \|x^{(k)} - x^{(k-1)}\|$$

This *a-posteriori* estimate allows us to stop the iteration when $\|x^{(k)} - x^{(k-1)}\| < \textit{tolerance}$. Note that we can easily compute $\|x^{(k)} - x^{(k-1)}\|$ without knowing x .

Finally, we say that the convergence of stationary iterative methods is linear (order $\alpha = 1$) with rate $C = \rho(G)$. Here we refer to the definition of rate/order of convergence of iterative methods: method converges with order *alpha* and rate C if we can bound

$$\|e^{(k)}\| \leq C \|e^{(k-1)}\|^\alpha.$$

2 Splitting methods

Most of the splitting methods rely on an inexpensive splitting of matrix A . The ones below use the following splitting $A = D - L - U$ where $D, -L, -U$ are build from, respectively, the diagonal, the subdiagonal, and the superdiagonal part of A .

2.1 Jacobi method

Jacobi method is based on the splitting $Dx^{(k)} = b + (L + U)x^{(k-1)}$.

Assuming D is invertible, we derive (3) with $G_{JAC} = D^{-1}(L + U)$.

One can prove (from Gerschgorin's theorem) that if A is strictly diagonally dominant, then Jacobi method converges. There are ways to weaken this assumption by assuming for example only weak diagonal dominance plus some other conditions.

Jacobi method has a very natural parallel implementation.

2.2 Gauss-Seidel method

Gauss-Seidel uses the splitting $(D - L)x^{(k)} = b + Ux^{(k-1)}$. It is similar to Jacobi's method except that, when processing a row, it uses the new information available in rows above.

One can prove that Gauss-Seidel method converges based on the same conditions as Jacobi, and that $\rho(G_{GS}) = \frac{1}{2}\rho(G_{JAC})$ for the same matrix A . This means that if it converges, it converges twice as fast as Jacobi.

However, it is not well suited for parallel implementation.

2.3 SOR=Successive OverRelaxation

SOR introduces a relaxation and optimally damps most frequencies of the error; it takes advantage of those slowly decaying and those fast decaying via weighing with a parameter ω as follows

$$x^{(k)} = (1 - \omega)x^{(k-1)} + \omega x^{(k),GS} \quad (8)$$

where $x^{(k),GS}$ would be obtained by a Gauss-Seidel step from $x^{(k-1)}$.

The iteration is defined via

$$(D - \omega L)x^{(k)} = (\omega U + (1 - \omega)D)x^{(k-1)} + \omega b \quad (9)$$

One can prove (Rellich/Ostrowski) that for convergence of SOR, it is necessary that $0 < \omega < 2$. SOR converges for spd matrices and strictly diagonal matrices.

One can find an optimal value of SOR for which the $\rho(G_{SOR,\omega})$, based on the knowledge of spectral radius $\rho(G_{JAC})$. In fact, we have

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \rho(G_{JAC})^2}}.$$

Unfortunately, in practice it is not easy to determine an optimal ω .

2.4 Richardson' iteration

Here we write $x^{(k)} = x^{(k-1)} + \alpha(b - Ax^{(k-1)})$ and choose α so that the iteration matrix $G_\alpha = I - \alpha A$ satisfies $\rho(G_\alpha) < 1$.

Assume A is spd. Then one can show that for $0 < \alpha < \frac{2}{\lambda_{max}(A)}$ we have $\rho(G_\alpha) < 1$.

One can also find an optimal $\alpha = \frac{2}{\lambda_{min}(A) + \lambda_{max}(A)}$: one for which the $\rho(G_\alpha)$ is the smallest.

3 Summary

Albeit stationary iterative methods converge, in general, much slower than the modern Krylov subspace methods (such as GMRES and CG), their importance is in that they are easy to implement and can be extended to simple nonlinear iterations. Most importantly, they are frequently used as preconditioners or as block preconditioners in Krylov subspace methods, and are used as smoothers in the Multigrid method.