Chapter 7

Iterative methods for linear systems

In this chapter we revisit the problem of solving linear systems of equations, but now in the context of large sparse systems for which direct methods are too expensive, in memory and execution time.

We introduce instead iterative methods, for which matrix sparsity is exploited to develop fast algorithms with a low memory footprint.

7.1 Stationary iterative methods

Iterative methods

For a given nonsingular matrix $A \in \mathbb{R}^{n\times n}$ and vector $b \in \mathbb{R}^n$, we consider the problem of finding a vector $x \in \mathbb{R}^n$, such that

$$Ax = b, \tag{7.1}$$

where the size n of the system is large, and the matrix A is *sparse* with the number of nonzero elements being $O(n)$ and not $O(n^2)$.

We do not seek to construct the exact solution $x = A^{-1}b$, but instead we will develop *iterative methods* based on algorithms that generate a sequence of approximations $\{x^{(k)}\}_{k\geq 0}$ that converges towards x, with

$$e^{(k)} = x - x^{(k)}, \tag{7.2}$$

the *error* at iteration k.

The error is not directly computable since the exact solution is unknown, but the error can be expressed in terms of the *residual* $r^{(k)} = b - Ax^{(k)}$, as

$$r^{(k)} = b - Ax^{(k)} = Ax - Ax^{(k)} = Ae^{(k)}, \tag{7.3}$$

so that for $\|\cdot\| = \|\cdot\|_2$, we have that

$$\|e^{(k)}\| = \|A^{-1}r^{(k)}\| \leq \|A^{-1}\|\|r^{(k)}\|, \tag{7.4}$$

and similarly

$$\|r^{(0)}\| = \|Ae^{(0)}\| \leq \|A\|\|e^{(0)}\|. \tag{7.5}$$

The *condition number* of A relative to the norm $\|\cdot\|$ is defined as

$$\kappa(A) = \|A\|\|A^{-1}\|, \tag{7.6}$$

which together with (7.4) and (7.5) provides an estimate of the relative error in terms of the relative residual.

Theorem 10 (Error estimate)**.** *For $\{x^{(k)}\}_{k\geq 0}$ a sequence of approximate solutions to the linear system of equations $Ax = b$, the relative error can be estimated as*

$$\frac{\|e^{(k)}\|}{\|e^{(0)}\|} \leq \kappa(A)\frac{\|r^{(k)}\|}{\|r^{(0)}\|}. \tag{7.7}$$

The error estimate (7.7) may be used as a stopping criterion for when to terminate an iterative algorithm,

$$\frac{\|r^{(k)}\|}{\|r^{(0)}\|} < TOL, \tag{7.8}$$

with $TOL > 0$ the chosen tolerance.

Although, to use the relative error with respect to the initial approximation can be problematic, since the choice of $x^{(0)}$ may be completely arbitrary and not of significance for the problem at hand. Instead it is more suitable to formulate a stopping criterion based on the following condition,

$$\frac{\|r^{(k)}\|}{\|b\|} < TOL, \tag{7.9}$$

corresponding to $x^{(0)} = 0$.

Stationary iterative methods

Stationary iterative methods are formulated as a linear *fixed point iteration* of the form

$$x^{(k+1)} = Mx^{(k)} + c, \tag{7.10}$$

with $M \in \mathbb{R}^{n\times n}$ the *iteration matrix*, $\{x^{(k)}\}_{k\geq 0} \subset \mathbb{R}^n$ a sequence of approximations, and $c \in \mathbb{R}^n$ a vector. If $\|M\| < 1$, the fixed point iteration (7.10)

converges to $x = (I - M)^{-1}c$, which follows from the Banach fixed point theorem. Further, we can show that an equivalent condition for convergence is that the *spectral radius* $\rho(M) < 1$, with

$$\rho(A) = \max_{\lambda \in \Lambda(A)} |\lambda|. \tag{7.11}$$

The linear system $Ax = b$ can be formulated as a fixed point iteration through the *Richardson iteration*

$$x^{(k+1)} = (I - A)x^{(k)} + b, \tag{7.12}$$

with an iteration matrix $M = I - A$, which will converge if $\|I - A\| < 1$, or $\rho(A) < 1$.

Preconditioning

To improve convergence of Richardson iteration we can *precondition* the system $Ax = b$ by multiplication of both sides of the equation by a matrix B, so that we get the new system

$$BAx = Bb, \tag{7.13}$$

for which Richardson iteration will converge if $\|I - BA\| < 1$, or equivalently $\rho(BA) < 1$, and we then refer to B as an *approximate inverse* of A. The preconditioned Richardson iteration takes the form

$$x^{(k+1)} = (I - BA)x^{(k)} + Bb, \tag{7.14}$$

and the preconditioned residual $Bb - BAx^{(k)}$ is used as basis for a stopping criterion.

Iterative methods based on matrix splitting

An alternative to Richardson iteration is *matrix splitting*, where stationary iterative methods are formulated based on splitting the matrix into a sum

$$A = A_1 + A_2, \tag{7.15}$$

where A_1 is chosen as a nonsingular matrix easy to invert, such as a diagonal matrix D, a (strict) lower triangular matrix L or (strict) upper triangular matrix U, where L and U have zeros on the diagonal.

Jacobi iteration

Jacobi iteration is based on the splitting

$$A_1 = D, \quad A_2 = L + U, \tag{7.16}$$

which gives the iteration matrixs $M_J = -D^{-1}(L + U)$, or in terms of the elements of $A = (a_{ij})$,

$$x_i^{(k+1)} = a_{ii}^{-1}(b - \sum_{j \neq i} a_{ij} x_j^{(k)}), \tag{7.17}$$

where the diagonal matrix D is trivial to invert. To use Jacobi iteration as a preconditioner, we choose $B = D^{-1}$.

Gauss-Seidel iteration

Gauss-Seidel iteration is based on the splitting

$$A_1 = D + L, \quad A_2 = U, \tag{7.18}$$

which gives the iteration matrixs $M_{GS} = -(D + L)^{-1}U$, or

$$x_i^{(k+1)} = a_{ii}^{-1}(b - \sum_{j<i} a_{ij} x_j^{(k+1)} - \sum_{j>i} a_{ij} x_j^{(k)}), \tag{7.19}$$

where the matrix $D + L$ is inverted by forward substitution.

7.2 Krylov methods

Krylov subspace

A *Krylov method* is an iterative method for the solution of the system $Ax = b$ based on, for each iteration, finding an approximation $x^{(k)} \approx x = A^{-1}b$ in a *Krylov subspace* $\mathcal{K}_k$, spanned by the vectors $b, Ab, ..., A^{k-1}b$, that is

$$\mathcal{K}_k =< b, Ab, ..., A^{k-1}b > . \tag{7.20}$$

The basis for Krylov methods is that, by the *Cayley-Hamilton theorem*, the inverse of a matrix A^{-1} can be expressed as a linear combination of its powers A^k.

GMRES

The idea of *GMRES* (generalized minimal residuals) is that, at each step k of the iteration, find the vector $x^{(k)} \in \mathcal{K}_k$ that minimizes the norm of the residual $r^{(k)} = b - Ax^{(k)}$, which corresponds to the least squares problem

$$\min_{x^{(k)} \in \mathcal{K}_k} \|b - Ax^{(k)}\|. \tag{7.21}$$

But instead of expressing the approximation $x^{(k)}$ as a linear combination of the Krylov vectors $b, Ab, ..., A^{k-1}b$, which leads to an unstable algorithm, we construct an orthonormal basis $\{q_j\}_{j=1}^{k}$ for $\mathcal{K}_k$, such that

$$\mathcal{K}_k =< q_1, q_2, ..., q_k >, \tag{7.22}$$

with Q_k the $n \times k$ matrix with the basis vectors q_j as columns.

Thus we can express the approximation as $x^{(k)} = Q_k y$, with $y \in \mathbb{R}^k$ a vector with the coordinates of $x^{(k)}$, so that the least squares problem take the form

$$\min_{y \in \mathbb{R}^k} \|b - AQ_k y\|. \tag{7.23}$$

The *Arnoldi iteration* constructs a partial similarity transformation of A into an *Hessenberg matrix* $\tilde{H}_k \in \mathbb{R}^{n+1 \times n}$,

$$AQ_k = Q_{k+1}\tilde{H}_k, \tag{7.24}$$

that is

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \left[\begin{array}{c|c|c} & & \\ q_1 & \cdots & q_n \\ & & \end{array}\right] = \left[\begin{array}{c|c|c|c} & & & \\ q_1 & \cdots & q_{n+1} & \\ & & & \end{array}\right] \begin{bmatrix} h_{11} & \cdots & h_{1n} \\ h_{21} & \cdots & \\ & \ddots & \vdots \\ & & h_{nn} \end{bmatrix},$$

and multiplication of (7.23) by Q_{n+1}^T does not change the norm, so that the least squares problem takes the form,

$$\min_{y \in \mathbb{R}^k} \|Q_{n+1}^T b - \tilde{H}_k y\|, \tag{7.25}$$

where we note that by construction $Q_{n+1}^T b = \|b\| e_1$, with $e_1 = (1, 0, ...)^T$, that gives

$$\min_{y \in \mathbb{R}^k} \|\|b\| e_1 - \tilde{H}_k y\|, \tag{7.26}$$

which is the least squares problem we solve for y at each iteration k, to get $x^{(k)} = Q_k y$.

Algorithm 7: GMRES

$q_1 = b/\|b\|$
for $k = 1, 2, 3$ **do**
  Arnoldi algorithm to get $Q_k, \tilde{H}_k$
  $\min_{y \in \mathbb{R}^k} \|\|b\| e_1 - \tilde{H}_k y\|$
  $x^{(k)} = Q_k y$
end