Solving Linear Systems: Explanation

June 10, 2025

1 Solving Linear Systems $Ax = b$

The given task is to solve the linear system of equations $Ax = b$, where $A \in \mathbb{R}^{n \times n}$ is a square matrix and $b \in \mathbb{R}^n$ is a vector.

Consider the general iterative method:

$$x_{k+1} = x_k - M^{-1}(Ax_k - b) \tag{1}$$

where $M \in \mathbb{R}^{n \times n}$ is an arbitrary non-singular (regular) matrix.

1.1 Explanation of the Iterative Method's Convergence

To analyze the convergence of this method, let's define the error at step k as $e_k = x_k - x^*$, where x^* is the exact solution to the system $Ax = b$. We can rewrite the iterative method (1) as:

$$x_{k+1} = x_k - M^{-1}Ax_k + M^{-1}b$$
$$x_{k+1} = (I - M^{-1}A)x_k + M^{-1}b$$

Let $G = I - M^{-1}A$. This matrix G is called the *iteration matrix*. Then: $x_{k+1} = Gx_k + M^{-1}b$.

Since x^* is the exact solution, it also satisfies this same formula: $x^* = Gx^* + M^{-1}b$. Subtracting this from the previous equation: $x_{k+1} - x^* = Gx_k - Gx^*$ $x_{k+1} - x^* = G(x_k - x^*)$ Substituting the definition of the error, we get:

$$e_{k+1} = Ge_k \tag{2}$$

This equation shows how the error changes from one iteration to the next. Recursively applying it, we get: $e_k = G^k e_0$, where $e_0 = x_0 - x^*$ is the initial error.

Convergence Condition via Norms: For the method to converge, we require that the error $\|e_k\|$ approaches zero as the number of iterations k tends to infinity. Taking the norm of both sides of the error equation (2): $\|e_{k+1}\| = \|Ge_k\|$

We know that for any matrix G and vector v, $\|Gv\| \leq \|G\|\|v\|$, where $\|G\|$ is a *matrix norm* (which measures how much a matrix can "stretch" a vector). Therefore, $\|e_{k+1}\| \leq \|G\|\|e_k\|$.

If $\|G\| < 1$ (for any consistent matrix norm), it means that at each step, the "size" of the error decreases by at least a factor of $\|G\|$. This leads to: $\|e_k\| \leq \|G\|^k \|e_0\|$

Since $\|G\| < 1$, $\|G\|^k$ approaches zero as $k \to \infty$. Thus, $\|e_k\| \to 0$, which means that the sequence x_k converges to the exact solution x^*.

The Sharpest Convergence Condition (Spectral Radius): While the condition $\|G\| < 1$ for any matrix norm is sufficient for convergence, there is a most precise (necessary and sufficient) condition. It is related to the *eigenvalues* of the matrix G. The eigenvalues λ of a matrix G are scalars for which there exists a non-zero vector v (eigenvector) such that $Gv = \lambda v$. An eigenvalue λ indicates by how much the eigenvector v is "stretched" or "compressed" by the action of matrix G.

The *spectral radius* of matrix G, denoted as $\rho(G)$, is the largest modulus (absolute value) among all its eigenvalues:

$$\rho(G) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } G\}$$

This explains why the "largest stretching factor" (which is the spectral radius) for the iteration matrix G must be less than one. If it is less than 1, then with each iteration, the error is guaranteed to decrease, and the method will converge to the exact solution.

1.2 Connection to Gradient Descent

Solving the linear system $Ax = b$ (provided that the matrix A is symmetric and positive definite) is equivalent to finding the minimum of the quadratic function:

$$f(x) = \frac{1}{2}x^T A x - b^T x$$

The gradient of this function is computed as:

$$\nabla f(x) = Ax - b$$

The general formula for the gradient descent method is:

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) \tag{3}$$

where $\alpha_k > 0$ is the step size at the k-th iteration.

Substituting the expression for the gradient into (3), we get:

$$x_{k+1} = x_k - \alpha_k(Ax_k - b)$$

Let's compare this formula with the original iterative method (1):

$$x_{k+1} = x_k - M^{-1}(Ax_k - b)$$

We can see that the gradient descent method is a special case of the iterative method (1) if the matrix M^{-1} is a scalar multiple of the identity matrix, i.e., $M^{-1} = \alpha_k I$, where I is the identity matrix, and α_k is the step size. In this case, $M = \frac{1}{\alpha_k}I$.

In a more general case, when M^{-1} is a non-scalar matrix, the iterative method (1) corresponds to *preconditioned gradient descent*. The matrix M^{-1} in this context acts as a *preconditioner*, which reshapes the geometry of the space or scales the gradient direction to accelerate the algorithm's convergence. The choice of matrix M (such as A_D for the Jacobi method or $A_D + A_L$ for the Gauss-Seidel method) essentially defines the preconditioning strategy for solving the linear system.