Solution of Linear Equations

Many computational problems involve the solution of systems of linear equations. A system of linear equations is typically expressed as

$$Ax = b,$$  \hspace{1cm} (1)

where $A$ is $m \times n$ coefficient matrix, $b$ is $m \times p$ right-hand sides, $x$ is the unknowns. When $b$ is the identity matrix, the unknown is the inverse matrix if it exits. Solution methods for (1) are important building blocks for computational sciences.

Computational methods for solving a system of linear equations may be categorized as follows

\begin{itemize}
  \item direct methods,
  \item iterative methods, and
  \item hybrid methods
\end{itemize}

In solving a system of linear equations, we are concerned with the following issues

\begin{itemize}
  \item the existence of a solution,
  \item the uniqueness of solutions,
  \item the sensitivity of a solution to perturbations,
  \item the stability of a computational method,
  \item the computational complexity of a computational method.
\end{itemize}

The first three issues are determined by the system itself and the last two issues are about the evaluation of a solution method, especially, when compared to other methods.
Review:

- There exists a solution iff \( b \in \text{span}(A) \).
- The solution, when exists, is unique iff \( \text{null}(A) = \{0\} \).
- Let \( k = \text{rank}(A) \). Then, \( k = \dim(\text{span}(A)) = \dim(\text{span}(A^T)) \).

**Direct Methods**

**Special systems**

We consider first special systems and then reduce a general system into special systems. Assume that \( A \) is non-singular \((k = m = n)\) and has at least one of the following properties.

1. \( A \) is diagonal, i.e., the system is decoupled.
   Solution: solve decoupled equations simultaneously (elementwise inversion).

2. \( A \) is triangular, i.e., the unknowns in the system are incrementally coupled.
   Solution: find the coupling ordering and do forward or backward substitution

3. \( A \) is unitary or orthogonal,
   in particular, permutation matrices are orthogonal,
   Solution: use the inverse \( A^H \).

4. \( A \) is of rank-1 update to the identity matrix

\[
A = I - xy^H, \quad y^Hx \neq 1.
\]

( In this case, \( A^{-1} = I - \beta xy^H \) with \( \beta = 1/(y^Hx - 1) \). )

In these special cases, either the inverse can be obtained easily, or the solution \( x \) can be obtained without computing the inverse.
Reduction to special systems via matrix factorization

For the general case the matrix is factored into special matrices, and the solution is obtained by successively solving the decomposed special systems. The process can be described as follows. Let

$$A = A_1 A_2 \cdots A_k$$

be a factorization of \( A \), the factors have special structures. Then the solution \( x \), if exists, can be obtained by solving the factored, special systems one by one

$$A_j y_j = y_{j-1}, \quad j = 1 : k, \quad y_0 = b, x = y_k.$$ 

We introduce some basic matrix factorizations. In many cases it is known that the matrix in question has linearly independent columns and the fact is taken into account to simplify the factorizations and reduce computation complexity.

- For \( A \) with linearly independent columns (\( \text{rank}(A) = n \leq m \)).

  1. **LU factorization with partial pivoting.** There exists a permutation matrix \( P \), a lower unit triangular matrix \( L \) and an upper triangular matrix \( U \) such that

  $$PA = LU, \quad U = \begin{bmatrix} U_{11} \\ 0 \end{bmatrix},$$

  where \( U_{11} \) is of order \( n \) and nonsingular.

  2. **QR factorization.** There exists a unitary matrix \( Q \), and an upper triangular matrix \( R \) such that

  $$A = QR, \quad R = \begin{bmatrix} R_{11} \\ 0 \end{bmatrix},$$

  where \( R_{11} \) is of order \( n \) and nonsingular.

- For arbitrary matrix,

  1. **LU factorization with total pivoting.** There exist permutation matrices \( P_r \) and \( P_c \), a lower unit triangular matrix \( L \) and an upper triangular matrix \( U \) such that

  $$P_r A P_c = LU, \quad U = \begin{bmatrix} U_{11} & U_{12} \\ 0 & 0 \end{bmatrix},$$
where $U_{11}$ is of order $k = \text{rank}(A)$ and nonsingular. Such a factorization reveals the matrix rank.

2. **QR factorization with column pivoting.** There exists a permutation matrix $P$, a unitary matrix $Q$, and an upper triangular matrix $R$ such that

$$AP = QR, \quad R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix},$$

where $R_{11}$ is of order $k = \text{rank}(A)$ and nonsingular.

In the above, we use factorization to indicate that the factors can be obtained with finite many arithmetic operations. In fact the proofs for such factorizations are often computationally constructive. We introduce the following decompositions, which may involve iterative procedures in numerical computation. These decompositions are very useful for matrix analysis as well as in matrix computations.

- **Eigenvalue decomposition (EVD).**
  For any square matrix $A$, there exists a nonsingular matrix $X$ and an upper bi-diagonal matrix $J$ such that

$$A = X J X^{-1}.$$  

It is called the eigenvalue decomposition of $A$. The diagonal elements of $J$ are eigenvalues of $A$. When $J$ is diagonal, $A$ is said to be diagonalizable, or decoupled, by similarity transformation and have a non-defect eigen-system. In such case, a function of $A$, $f(A)$, may be studied in the spectral domain. In particular, if $p$ is a polynomial, then $p(A) = X p(J) X^{-1}$.

We mention two special cases. A circulant matrix is diagonalizable by the DFT matrix. A Hermitian (or symmetric) matrix is diagonalizable by a unitary similarity transformation.

- **Singular value decomposition (SVD).**
  For any matrix, there exist unitary matrices $U$ and $V$, and a non-negative diagonal matrix $\Sigma = \text{diag}(\sigma_i)$ such that

$$A = U \Sigma V^H.$$  

The diagonal elements of $\Sigma$ are the singular values of $A$. By convention, the singular values $\sigma_i$ are non-decreasingly ordered, $\sigma_i \geq \sigma_{i+1}$. The $j$th columns of $U$ and $V$ are, respectively, the left and right singular vectors corresponding to the $j$-th singular value $\sigma_j$.  


There are other factorizations and decompositions, many are variations of the basic ones introduced above. The EVD is often used in convergence analysis of iterative methods for the solution of linear equations. The SVD is often employed to study the sensitivity of the solution to perturbation, or the stability of an algorithm. It can be also used in the analysis of LS solution.

We mention the particular relationship between Hermitian EVD and SVD.

- A SVD is related to the EVDs of two symmetric positive definite matrices
  \[ A^	ext{H}A = V\Sigma^2V^\text{H}, \quad \text{and} \quad AA^\text{H} = U\Sigma^2U^\text{H}, \]

- A SVD is related to the EVD of a symmetric, indefinite matrix
  \[
  \begin{bmatrix}
  0 & A \\
  A^\text{H} & 0 
  \end{bmatrix}
  = 1/2 \begin{bmatrix}
  U & U \\
  V & -V 
  \end{bmatrix} \begin{bmatrix}
  \Sigma & 0 \\
  0 & -\Sigma 
  \end{bmatrix} \begin{bmatrix}
  U & U \\
  V & -V 
  \end{bmatrix}^\text{H}.
  \]

**Stability analysis**

There are two aspects in stability analysis for the solution of a system of linear equations. We ask first whether or not the solution changes dramatically with a small perturbation in the data \((A, b)\). Perturbations may come from many sources, including those in the data. Perturbations may be also introduced by a numerical method, especially with finite-precision computation. We therefore evaluate a method based on an estimation how much perturbation it introduces potentially. In short, we are concerned in the first aspect with the sensitivity of the solution to any perturbations, and we are concerned in the second aspect how much particular perturbation a method may introduce potentially.

We illustrate the analysis with the case that \(A\) is nonsingular and we use 2-norm for the metric on the change. We consider the effects of perturbations in the matrix \(A\) and in the right hand side separately, and leave their joint effect as an exercise problem.

1. Sensitivity to perturbation in the right hand side \(b\)

   \[ A(x + \Delta x) = b + \Delta b. \]

   First we have \(\Delta x = A^{-1}\Delta\). Then, we can verify the following bound on \(\Delta x\), via the SVD of \(A\),

   \[
   \frac{\|\Delta x\|_2}{\|x\|_2} \leq \kappa_2(A) \frac{\|\Delta b\|_2}{\|b\|_2}.
   \]
where \( \kappa_2 = \sigma_{\text{max}} / \sigma_{\text{min}} = \| A \|_2 || A^{-1} ||_2 \) is called the condition number of \( A \) in 2-norm, which indicates the sensitivity of the solution to the perturbation. Via the SVD, it is not hard to find the perturbations in \( b \) that maximize the deduced perturbation in \( x \) and achieve the perturbation bound. A large condition number is therefore undesirable, unless we have control over the perturbation in \( b \).

2. Sensitivity to perturbation in the matrix \( A \)

\[
(A + \Delta A)(x + \Delta x) = b.
\]

If \( \| A^{-1} \Delta A \|_2 < 1/k \), for \( k > 1 \), then

\[
\frac{\| \Delta x \|_2}{\| x \|_2} \leq \frac{\| A^{-1} \Delta A \|_2}{1 - \| A^{-1} \Delta A \|_2} < \frac{k}{k - 1} \kappa_2(A) \frac{\| \Delta A \|_2}{\| A \|_2}.
\]

Thus, \( \kappa_2 \) also indicates the sensitivity of the solution to the perturbation in the matrix.

3. Residual bound.

In practical computation, the residual, \( r = b - A\bar{x} \), can be for a computed solution \( \bar{x} \). In the ideal case, the residual is zero. Otherwise, we would like to estimate \( \Delta x \), the distance of the computed solution from the true solution, from the residual.

The residual may be seen as a perturbation in the right hand side, \( A\bar{x} = b - r \). Thus,

\[
\frac{\| \Delta x \|_2}{\| x \|_2} \leq \kappa_2(A) \frac{\| r \|_2}{\| b \|_2}.
\]

It states that a small residual does not indicate a small error in the solution unless the matrix is well-conditioned, i.e., the matrix has a modest condition number for the problem of interest. The condition number of any matrix is bounded from below by 1, which is reached by unitary matrices.

The residual may be also seen as a perturbation in the matrix

\[
(A + \Delta A)\bar{x} = b, \quad \Delta A = ry^H / (y^H \bar{x}), \quad y^H \bar{x} \neq 0,
\]

i.e., the perturbation in \( A \) is of rank-1 and \( y \) can be any vector not orthogonal to \( \bar{x} \). A special choice of \( y \) is \( \bar{x} \).

In practice, the residual is computed in higher precision.
4. Backward analysis of numerical methods.

An approach to evaluating a numerical method is the backward analysis. In backward analysis, one proves that a numerical method renders an exact solution solution to the system with a perturbed matrix $A + \Delta A$ and a perturbed right hand $b + \Delta b$. If the perturbations associated with the method are large, or prohibitively large, the method is considered backward unstable. In practice, the degree of perturbation by a numerical method is to be compared against the perturbation or uncertainty in the data.

**Algorithm complexity**

The direct methods, in general, are of $O(n^3)$ in arithmetic complexity and $O(n^2)$ in memory space requirement for a system with an $n \times n$ matrix, where the decompositions are computed according to a fixed machine precision. Systems of small size can be solved quickly by these methods on modern computers. Direct methods become too expensive to use directly for large systems. Many systems are sparse and irregularly structured, iterative methods are used instead to exploit the sparsity. An iterative method can exploit many kinds of system structures, while direct methods preserve only a few special ones. For large and dense matrices, compressed representations are sought first and exploited in iterative methods. It is common to use direct methods for smaller systems induced by an iterative method for a large system.