Gaussian Elimination and Matrices

Notes on Linear Algebra

Chia-Ping Chen

Department of Computer Science and Engineering

National Sun Yat-Sen University

Kaohsiung, Taiwan ROC

Course Syllabus

- Vectors
- Matrices
- Linear equations
- Vector spaces
- Orthogonality
- Determinants
- Eigenvalues and eigenvectors
- Positive definite matrices
- Computation with matrices
- Matrix decomposition

Introduction

- A basic problem of linear algebra is to find the values of unknowns that satisfy a given system of linear equations.
- Suppose there are n unknowns and n equations. What methods do you know to solve such a problem?
  - method of elimination (Gaussian elimination)
  - method by determinant (Cramer's rule)

The Geometry of Linear Equations

There are two ways to look at the following system of linear equations.

$$\begin{cases} 2x - y = 1 \\ x + y = 5. \end{cases}$$

- (row picture) Each equation represents a line and we are looking for the intersection point(s) of these lines.
- (column picture) The set of coefficients of a variable $(2, 1$ for $x)$ represents a vector. We are looking for the combination of these vectors that equals the right-hand side.

Row Picture

- The row picture is not unfamiliar.
  - for $n = 2$, a linear equation represents a line (1-d) and two lines intersect at a point
  - for $n = 3$, a linear equation represents a plane (2-d) and three planes intersect at a point
  - for $n > 3$, a linear equation represents a "plane" of dimension $n - 1$, and we need n such planes to intersect at a point
- The point of intersection is the solution.

Column Picture

- We form a *column vector* by putting all coefficients for the unknown x in a column. Similarly for other unknowns.
- With these column vectors, the system of equations can be re-written as a single vector equation.

$$x \begin{bmatrix} 2 \\ 1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

- This meaning of this equation can be explained in the column picture.

Vector Operations

- Addition

$$\begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} a+c \\ b+d \end{bmatrix}$$

- Multiplication by a scalar

$$r \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} ra \\ rb \end{bmatrix}$$

Linear Combination

- A *linear combination* of a set of vectors is defined by

$$\lambda_1 v_1 + \cdots + \lambda_n v_n,$$

  where λ_is are scalars and v_is are the vectors.
- In column picture, we are looking for a linear combination of the n column vectors to equal the right-hand side.
- An intersection point of row-picture planes must provide the linear combination coefficients for the column vectors!

Singular Cases

- Ordinarily, n linear equations with n unknowns has a unique solution. But there are also singular cases.
  - inconsistent equations, no solution
  - infinitely many solutions
- Geometrically, in the row picture, no solution means the planes do not intersect at any point, and infinitely many solutions mean that the planes intersect at a line or a plane.
- In the column picture, singular cases means there are either no or more than 1 linear combinations for the right-side. In this case, one of the column vectors is a linear combination of the others.

Gaussian Elimination

- Example

$$
\begin{aligned}
2u + \quad v + \quad w &= 5 \\
4u - \ 6v \qquad\qquad &= -2 \\
-2u + \ 7v + \ 2w &= 9.
\end{aligned}
$$

- We want to first eliminate unknown u from the second and third equations and then v from the third equation.
- Let's work out the details.

Pivots and Multipliers

- The number 2, the coefficient of the first unknown in the first equation, is called a *pivot*. By definition, a pivot cannot be zero.
- The factors $2, -1$ to multiply the first equation to eliminate the first unknown in the other equations, are called the *multipliers*.
- The second pivot is -8, and the corresponding multiplier is -1.
- The third pivot is 1.

Back Substitution

- After elimination, the system of equations becomes

$$\begin{aligned} 2u + \quad v + \quad w &= 5 \\ -8v - \ 2w &= -12 \\ w &= 2. \end{aligned}$$

- Note that it is in a "triangular" form. It is easily solve by *back substitution*.
- The solution is $w = 2, v = 1, u = 1$.

Matrix Representation

- To represent a system of linear equations, we can simply write down the coefficients and right-hand side. (The $''+'','' =''$ and unknowns are implicit).
- The new representation is a matrix.
- The Gaussian elimination of the above system of linear equations is represented by

$$\begin{bmatrix} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Breakdown of Elimination

- When does the elimination process break down?
  - Something must go wrong in singular cases.
  - Something may go wrong in non-singular cases.
- If a zero shows up in a pivot position during the elimination process, then it has to stop because the multipliers cannot be found.
  - In non-singular cases, such problems can be cured by re-arranging the equations.
  - In singular cases, a zero always shows up at some pivot position no matter how.

Elimination Cost

- How many arithmetic operations does elimination require for n equations in n unknowns?
- Let's count each division or multiply-subtract one operation. For a row under a pivot,
  - the leading entry is divided by the pivot to find the multiplier (1 op)
  - the remaining entries are multiplied and subtracted (0 to $n-1$ ops)
- The total number of operations is

$$\sum_{k=1}^{n}(n-k+1)(n-k) = \sum_{k'=1}^{n}(k'-1)k' = O(n^3).$$

Back Substitution Cost

- The last equation requires one operation (division by pivot).
- The second to last requires two operations (one multiplication followed by subtraction, and one division).
- So the total number of operations is

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

- The computational cost is mostly spent on elimination (n^3), not back substitution (n^2).

Matrix Notation

- There are three different types of quantities in a system of linear equations.
  - the unknowns u, v, w.
  - the right-hand side $5, -2, 9$
  - the coefficients
- We can put them in matrices and represent the system by

$$Ax = b,$$

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}, \quad x = \begin{bmatrix} u \\ v \\ w \end{bmatrix}, \quad b = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}$$

Multiplication of Matrix and Vector

- We want to define the multiplication of a matrix and a vector in such a way that $Ax = b$ reproduces the original system of equations.
- Specifically, we want

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} \Rightarrow \begin{aligned} 2u + \quad v + \quad w &= 5 \\ 4u - \quad 6v \qquad\quad &= -2 \\ -2u + \quad 7v + \quad 2w &= 9. \end{aligned}$$

Inner Product

- For the first component, we want

$$\begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 2u + v + w \end{bmatrix}.$$

  So we define the first component of the product Ax to be the sum of component-wise multiplication of x and the first row of A.

- The sum of component-wise multiplication of two vectors is also known as *inner product*.

Linear Combination of Columns

- Ax is a linear combination of the columns of A with the components of x as the coefficients.

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 2u + 1v + 1w \\ 4u + (-6)v + 0w \\ -2u + 7v + 2w \end{bmatrix}$$

$$= u \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} + v \begin{bmatrix} 1 \\ -6 \\ 7 \end{bmatrix} + w \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

Linear Combination of Rows

- The inner product view for Ax has another interpretation: the ith component y_i is a linear combination of the components of x, with the entries in row i of A as coefficients.

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 2u + 1v + 1w \\ 4u + (-6)v + 0w \\ -2u + 7v + 2w \end{bmatrix}$$

Matrix Multiplication

- We are now ready to define the multiplication of two matrices A and B. Let $\mathbf{b}_j$ be column j of B and $\mathbf{a}_i$ be row i of A.
  - Column j of AB is a linear combination of columns of A using $\mathbf{b}_j$ as coefficients.
  - Row i of AB is a linear combination of rows of B using $\mathbf{a}_i$ as coefficients.
  - The ijth entry is the inner product of $\mathbf{a}_i$ and $\mathbf{b}_j$.
- Note that for AB to be defined, the number of columns in A must be the same as the number of rows in B.

Elementary Matrix

- Define an elementary matrix $E_{ij}(l)$ which has 1's on the diagonal, and $-l$ in the intersection of row i and column j, and 0's elsewhere.
- When a matrix is multiplied by $E_{ij}(l)$ from the left, the new row i is the old row i subtract l times row j.
- That's exactly the row operation: a row operation is equivalent to a multiplication of an elementary matrix from the left.
- Gaussian elimination consists of a sequence of row operations. Thus it is equivalent to a product of elementary matrices.

Example

- In our example, $Ax = b \Rightarrow Ux = c$, or

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -12 \\ 2 \end{bmatrix}$$

- The multipliers are $l_{21} = 2, l_{31} = -1, l_{32} = -1$. So

$$E_{32}(-1)E_{31}(-1)E_{21}(2)A = U.$$

It's a good exercise to work out the details!

Triangular Matrices

- The matrix U is an upper-triangular matrix: all non-zero entries are on or above the diagonal.
- An elementary matrix E_{ij} is a lower-triangular matrix. All non-zero entries are on or below the diagonal.
- To go back to A from U, we need to find the (inverse) matrices that “undo” the effects of elementary matrices.
  - For an elementary matrix, simply reverse the sign of l, i.e., use $E_{ij}(-l)$.
  - For a product of elementary matrices, apply the inverse matrices in the reversed order.

LU Decomposition

- If no row exchanges are required, the coefficient matrix for a non-singular system of linear equations can be written as

$$A = LU,$$

where L is a lower-triangular matrix and U is an upper-triangular matrix.

- Moreover, U is the matrix after elimination and L has the multipliers l_{ij} as the (i, j) entry.

Example

- In our example,

$$E_{32}(-1)E_{31}(-1)E_{21}(2)A = U \Rightarrow A = E_{21}(-2)E_{31}(1)E_{32}(1)U.$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

Proof

- Let L and U be as defined. Apply the elimination to

$$\begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} \text{row 1 of } U \\ \text{row 2 of } U \\ \text{row 3 of } U \end{bmatrix}.$$

If the result is U, then $LU = A$ since applying the elimination steps (row operations) to A also yields U.

- It does yield U since the matrix on the left is the identity matrix when elimination is applied.

LU and Elimination

- We can solve $Ax = b$ via LU decomposition

$$Ax = b \Rightarrow LUx = b \Rightarrow Lc = b \text{ and } Ux = c.$$

- Note that triangular systems can be solved quickly (via back or forward substitutions).
- The above is exactly what the Gaussian elimination method does.

LDU Decomposition

- In LU decomposition, L is lower-triangular with 1's on the diagonal, while U is upper-diagonal with the pivots on the diagonal.
- We can make it more "symmetric" by requiring the diagonal entries of U to be 1. This is the LDU decomposition: For a nonsingular A,

$$A = LDU, \text{ where } \begin{cases} L \text{ is unit lower-triangular} \\ U \text{ is unit upper-triangular} \\ D \text{ is diagonal} \end{cases}$$

- The LDU decomposition of a matrix is unique.

Row Exchanges

- In the process of Gaussian elimination, we may encounter a zero in a pivot position at some point.
- In this case, we look for a non-zero entry under the position in the same column.
  - If none can be found, the system of equations is singular and there is no unique solution.
  - If one is found, we exchange that row with current row and proceed.
- If an elimination can be completed with row exchanges, then those exchanges can be done ahead of time without changing solution.

Permutation Matrices

- A permutation π is a re-ordering of $(1, 2, \ldots, n)$.
- Every permutation can be resulted from a sequence of swappings of pairs of elements.
- A permutation π is *even* if the number of swappings to go from $(1, 2, \ldots, n)$ to π is even, and *odd* otherwise.
- A permutation matrix is a matrix whose rows are a permutation of the rows of the identity matrix.
- An exchange of row i and row j is equivalent to the multiplication from left by the permutation matrix P_{ij}, resulting from exchanging row i and row j of the identity matrix.

Inverses

- The inverse of an n by n matrix A is another n by n matrix, written as A^{-1}, such that

$$AA^{-1} = I = A^{-1}A.$$

- A^{-1} may not exist. When it does, it is unique and A is said to be invertible.
- If A is invertible, then

$$Ax = b \Leftrightarrow x = A^{-1}b.$$

Properties of Inverses

- The inverse of AB is $B^{-1}A^{-1}$, since

$$ABB^{-1}A^{-1} = I = B^{-1}A^{-1}AB.$$

- The inverse of A^{-1} is A itself, since

$$A^{-1}A = I = AA^{-1}.$$

Calculation of A^{-1}

- Since $AA^{-1} = I$, A^{-1} consists of n vectors that

$$A \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} = \begin{bmatrix} e_1 & \dots & e_n \end{bmatrix}.$$

- Here we have n systems of linear equations with the same coefficient matrix A and different right sides.
- It is possible to solve all systems of equations simultaneously. The Gauss-Jordan method does that.

Gauss-Jordan Method

- The Gauss-Jordan method begins with $\begin{bmatrix} A & \mid & I \end{bmatrix}$ and ends with $\begin{bmatrix} I & \mid & A^{-1} \end{bmatrix}$.
- Starting with elimination that makes the first n columns upper-triangular

$$\begin{bmatrix} A & \mid & I \end{bmatrix} \xrightarrow{L^{-1}} \begin{bmatrix} U & \mid & L^{-1} \end{bmatrix},$$

it continues by subtracting multiples of a row from the rows above to make the first n columns to become I,

$$\begin{bmatrix} U & \mid & L^{-1} \end{bmatrix} \xrightarrow{U^{-1}} \begin{bmatrix} I & \mid & A^{-1} \end{bmatrix}.$$

Example

- Work out the details for the case

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}.$$

Invertible and Non-singular

- If a matrix A has a full set of pivots (by definition A is non-singular), then it is invertible.
- Proof
  - right inverse: the columns of A^{-1} can be solved one by one for $AA^{-1} = I$.
  - We must also show that A^{-1} is also a left inverse. From Gauss-Jordan, A has a left inverse which is a product of three types of matrices: elementary, permutation and diagonal.
  - For A, a left inverse B must equal to a right inverse C, since $B = B(AC) = (BA)C = C$.

Matrix Transposes

- The *transpose* of A, denoted by A^T, is defined by

$$(A^T)_{ij} = A_{ji}.$$

- Row i of A is column i of A^T. Likewise, column j of A is row j of A^T.
- For transposes,

$$(A^T)^T = A$$
$$(AB)^T = B^T A^T$$
$$(A^{-1})^T = (A^T)^{-1}$$

Symmetric Matrices

- A matrix A is said to be *symmetric* if it equals its transpose
$$A^T = A.$$
- Symmetric matrices appear often in statistics.
- They also appear in positive definite matrices.
- The eigenvalues of a symmetric matrix are all real.

Theorem

- If A is symmetric and if it can be decomposed to $A = LDU$ without row exchanges, then $U = L^T$. That is, the factorization becomes $A = LDL^T$.
- To prove, note

$$A = LDU \Rightarrow A^T = U^T D^T L^T = L' D U'$$
$$\Rightarrow L' D U' = A^T = A = LDU$$

From the uniqueness of LDU decomposition, we have $L = L' = U^T$.