Linear Algebra

Chapter 2: Systems of Linear Equations

University of Seoul
School of Computer Science
Minho Kim

Table of contents

Outline

Triviality: "Three Roads"

$$\begin{array}{rcrcr} 2x & + & y & = & 8 \\ x & - & 3y & = & -3 \end{array}$$

1. Geometric meaning:
   "Find the position vector that is the **intersection of two lines** with equations $2x + y = 8$ and $x - 3y = -3$."
   (Problem 1)
2. Linear combination:
   "Let $\boldsymbol{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} 8 \\ -3 \end{bmatrix}$. Fine the coefficients x and y of the linear combination of $\boldsymbol{u}$ and $\boldsymbol{v}$ such that $x\boldsymbol{u} + y\boldsymbol{v} = \boldsymbol{w}$." (Problems 2∼4)
3. Numerical view:
   "How can we find the solution?" (Problems 5∼6)

Outline

Linear Equations

Definition

A **linear equation** in the n variables $x_1, x_2, \cdots, x_n$ is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where the **coefficients** $a_1, a_2, \cdots, a_n$ and the **constant term** b are constants.

- Examples of linear equations

$$3x - 4y = -1 \quad x_1 + 5x_2 = 3 - x_3 + 2x_4 \quad \sqrt{2}x + \frac{\pi}{4}y - \left(\sin\frac{\pi}{5}\right)z = 1$$

- Examples of nonlinear equations

$$xy + 2z = 1 \quad x_1^2 - x_2^3 = 3 \quad \sqrt{2}x + \frac{\pi}{4}y - \sin\left(\frac{\pi}{5}z\right) = 1$$

Systems of Linear Equations

- **System of linear equations**: finite set of linear equations, each with the same variables
- **Solution** (of a system of linear equations): a vector that is *simultaneously* a solution of each equation in the system
- **Solution set** (of a system of linear equations): set of *all* solutions of the system
- Three cases
  1. a unique solution (a **consistent** system)
  2. infinitely many solutions (a **consistent** system)
  3. no solutions (an **inconsistent** system)
- **Equivalent** linear systems: different linear systems having the same solution sets.

Solving a System of Linear Equations

A linear system with **triangular pattern** can be easily solved by applying **back substitution**. (Example 2.5)

- Is the solution of the system in triangular form also the solution of the original one? Why? $\rightarrow$ They are *equivalent*.
- How can we transform a linear system into an equivalent triangular linear system?
  $\rightarrow$ Example 2.6

Numerical Errors

- Example (p.66):

$$\begin{array}{ccccc} x & + & y & = & 0 \\ x & + & \frac{801}{800}y & = & 1 \end{array}$$

- Due to the **roundoff errors** introduced by computers
- **Ill-conditioned** system: extremely sensitive to roundoff errors
- Geometric view?

Outline

Matrices Related to Linear Systems

For the system (of linear equations)

$$\begin{array}{ccccccc} a_1x & + & b_1y & + & c_1z & = d_1 \\ a_2x & + & b_2y & + & c_2z & = d_2 \\ a_3x & + & b_3y & + & c_3z & = d_3 \end{array}$$

the **coefficient matrix** is

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

and the **augmented matrix** is

$$\left[\begin{array}{ccc|c} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{array}\right]$$

Echelon Form

Can we always reduce any matrix to triangular form?
$\rightarrow$ echelon form

Definition

A matrix is in **row echelon form** if it satisfies the following properties:

1. Any rows consisting entirely of zeros are at the bottom
2. In each nonzero row, the first nonzero entry (called the **leading entry** is in a column to the left of any leading entries below it.

- In any column containing a leading entry, all entries below the leading entry are zeros.
- What makes the row echelon form good?
- Is the row echelon form unique for a given matrix? $\rightarrow$ No. (p.71) Example?

Elementary Row Operations

Allowable operations that can be performed on a system of linear equations to transform it into an equivalent system.

Definition

The following **elementary row operations** can be performed on a matrix:

1. Intercahange two rows.
   $R_i \leftrightarrow R_j$
2. Multiply a row by a nonzero constant.
   kR_i
3. Add a multiple of a row to another row.
   $R_i + kR_j$: Add k times row j to row i and replace row i with the result.

- Why are they allowable? Which operations are NOT allowable?
- **Row reduction**: The process of applying elementary row operations to bring a matrix into row echelon form.
- **Pivot**: The entry chosen to become a leading entry

Elementary Row Operations (cont'd)

- Row reduction is reversible $\rightarrow$ How?

Definition

Matrices A and B are **row equivalent** if there is a sequence of elementary row operations that converts A into B.

Theorem 2.1

Matrices A and B are row equivalent iff they can be reduced to the same row echelon form.

Gaussian Elimination

- A method to solve a system of linear equations

1. Write the augmented matrix of the system of linear equations.
2. Use elementary row operations to reduce the augmented matrix to row echelon form.
   (a) Locate the leftmost column that is not all zeros.
   (b) Create a leading entry at the top of this column. (Making it 1 makes your life easier.)
   (c) Use the leading entry to create zeros below it.
   (d) Cover up (Hide) the row containing the leading entry, and go back to step (a) to repeat the procedure on the remaining submatrix. Stop when the entire matrix is in row echelon form.
3. Using back substitution, solve the equivalent system that corresponds to the row-reduced matrix.

Gaussian Elimination (cont'd)

- What if there are more than one ways to assign values in the final back substitution? (Example 2.11)
  $\rightarrow$ Solution in vector form writing the **leading variables** in terms of **free variables.**
- In a consistent system,
  - Leading variables: variables corresponding to the leading entries $\rightarrow$ How many leading variables do we have?
  - Free variables: variables that are not leading variables
- Given a matrix, the number of nonzero rows is the same in *all* row echelon forms. $\rightarrow$ **rank**

Rank

Definition

The **rank** of a matrix is the number of nonzero rows in its row echelon form.

- # of nonzero rows = # of leading variables
- The "rank of a matrix A" is denoted by $\mathrm{rank}(A)$.

Theorem 2.2: The rank theorem

Let A be the coefficient matrix of a system of linear equations with n variables. If the system is consistent, then

$$\text{number of free variables} = n - \mathrm{rank}(A)$$

Reduced Row Echelon Form

Definition

A matrix is in **reduced row echelon form** if it satisfies the following properties:

1. It is in row echelon form.
2. The leading entry in each nonzero row is a 1 (called a **leading** 1).
3. Each column containing a leading 1 has zeros everywhere else.

- Unique! cf) Row echelon form is not unique. $\rightarrow$ Proof? Not easy!

Example

$$\begin{bmatrix} 1 & 2 & 0 & 0 & -3 & 1 & 0 \\ 0 & 0 & 1 & 0 & 4 & -1 & 0 \\ 0 & 0 & 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Gauss-Jordan Elimination

- Simplifies the back substitution step of Gauss elimination.

Steps

1. Write the augmented matrix of the system of linear equations.
2. Use elementary row operations to reduce the augmented matrix to **reduced row echelon form**.
3. If the resulting system is consistent, solve for the leading variables in terms of any remaining free variables.

Linear Systems and Geometry

- Plane-plane intersection (Example 2.14)
- Line-line intersection (Example 2.15)

Homogeneous Systems

Definition

A system of linear equations is called **homogeneous** if the constant term in each equation is zero.

- Always have at least one solution $\rightarrow$ What is it? $\rightarrow$ *trivial* solution
- When does a homogeneous linear system have infinitely many solutions?

Theorem 2.3

If $[A|\mathbf{0}]$ is a homogeneous system of m linear equations with n variables, where $m < n$, then the system has infinitely many solutions.

Outline

Linear Systems and Linear Combinations

"Does a linear system have a solution?"
$\Leftrightarrow$ "Is the vector $\boldsymbol{w}$ a linear combination of the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$?"
Example 2.18:
Does the following linear system have a solution?

$$\begin{array}{rcrcl} x & - & y & = & 1 \\ & & y & = & 2 \\ 3x & - & 3y & = & 3 \end{array} \rightarrow \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$\Leftrightarrow$ Is the vector $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ a linear combination of the vectors

$\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix}$? $\rightarrow \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} x + \begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix} y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Spanning Sets of Vectors

Theorem 2.4

A system of linear equations with augmented matrix $[A|\boldsymbol{b}]$ is consistent iff $\boldsymbol{b}$ is a linear combination of the columns of A.

Definition

If $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k\}$ is a set of vectors in $\mathbb{R}^n$, then the set of all linear combinations of $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k$ is called the **span** of $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k$ and is denoted by $\mathrm{span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k)$ or $\mathrm{span}(S)$. If $\mathrm{span}(S) = \mathbb{R}^n$, then S is called a **spanning set** for $\mathbb{R}^n$.

- How big is $\mathrm{span}(S)$ if $S \neq \emptyset$?
- $\mathrm{span}(S) = \mathbb{R}^n$
  $\Leftrightarrow$ Any vector in $\mathbb{R}^n$ can be written as a linear combination of the vectors in S. (Example 2.19)
- What do the vectors in S span if $\mathrm{span}(S) \neq \mathbb{R}^n$? (Example 2.21)
  $\rightarrow$ How to find the general equation of a plane? (three methods)

Linear Independence

Given the vectors $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$, can any vector be wrritten as a linear combination of others?

Definition

A set of vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k$ is **linearly dependent** if there are scalars $c_1, c_2, \cdots, c_k$, *at least one* of which is not zero, such that

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k = \boldsymbol{0}.$$

A set of vectors that is *not* linearly dependent is called **linearly independent.**

Theorem 2.5

Vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$ in $\mathbb{R}^n$ are linearly dependent *iff* at least one of the vectors can be expressed as a linear combination of the others.

- What if one of the vectors is $\boldsymbol{0}$? (Example 2.22)

Checking Linear Independence

- Example 2.23

Theorem 2.6

Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$ be (column) vectors in $\mathbb{R}^n$ and let A be the $n \times m$ matrix $[\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ \cdots \ \boldsymbol{v}_m]$ with these vectors as its columns. Then $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$ are linearly dependent *iff* the homogeneous linear system with augmented matrix $[A|\mathbf{0}]$ has a nontrivial solution.

Checking Linear Independence (cont'd)

- Example 2.25 $\rightarrow$ Performing elementary row operations $=$ Constructing linear combination of (row) vectors

Theorem 2.7

Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$ be (row) vectors in $\mathbb{R}^n$ and let A be the $m \times n$ matrix $\begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_m \end{bmatrix}$ with these vectors as its rows. Then $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$ are linearly dependent *iff* $\text{rank}(A) < m$.

Theorem 2.8

Any set of m vectors in $\mathbb{R}^n$ are linearly dependent *if* $m > n$.

Outline

Applications

1. Allocation of resources – to allocate limited resources subject to a set of constraints
2. Balanced chemical equations – relative number of reactants and products in the reaction keeping the number of atoms $\rightarrow$ homogeneous linear system
3. Network analysis – “conservation of flow”: At each node, the flow in equals the flow out.
4. Electrical networks – specialized type of network
5. Finite linear games – finite number of *states*
6. Global positioning system (GPS) – to determine geographical locations from the satellite data

Outline

Iterative Method

- Usually *faster* and *more accurate* than the direct methods
- Can be stopped when the approximate solution is sufficiently accurate
- Two methods:
  1. Jacobi's method
  2. Gauss-Seidel method

Jacobi's Method

$$\begin{array}{rcrcr} 7x_1 & - & x_2 & = & 5 \\ 3x_1 & - & 5x_2 & = & -7 \end{array}$$

1. Solve the 1st eq. for x_1 and the 2nd eq. for x_2:

$$x_1 = \frac{5 + x_2}{7} \quad \text{and} \quad x_2 = \frac{7 + 3x_1}{5}$$

2. Assign **initial approximation** values, e.g., $x_1 = 0, x_2 = 0$.

$$x_1 = 5/7 \approx 0.714 \quad \text{and} \quad x_2 = 7/5 \approx 1.400$$

3. Substitute the new x_1 and x_2 into those in step 1 and repeat.
4. The solution **converges** to the exact solution $x_1 = 1, x_2 = 2$!

Gauss-Seidel Method

- Modification of Jacobi's method
- Use each value *as soon as we can*. $\to$ converges faster
- Different zigzag pattern
- Nice geometric interpretation in two variables

1. Solve the 1st eq. for x_1 and assign the initial approximation, of x_2, e.g., $x_2 = 0$:

$$x_1 = \frac{5+0}{7} = \frac{5}{7} \approx 0.714$$

2. Solve the 2nd eq. for x_2 and assign the value for x_1 just computed.

$$x_2 = \frac{7 + 3 \cdot (5/7)}{5} \approx 1.829$$

3. Repeat.

Generalization

How can we generalize each method to the linear systems of n variables?

Questions

- Do these methods always converge? (Example 2.36)
  $\rightarrow$ **divergence**
- If not, *when* do they converge?
  $\rightarrow$ Chapter 7

Gaussian Elimination? Iterative Methods?

- Gaussian elimination is sensitive to roundoff errors.
- Using Gaussian elimination, we cannot improve on a solution once we found it.