DM559

Linear and Integer Programming

Lecture 1
Introduction

Marco Chiarandini

Department of Mathematics & Computer Science
University of Southern Denmark

Outline

Outline

1. Course Organization

2. Preliminaries

Aims of the course

Learn about:

- both the theory and the practice of Linear Algebra
- one of the most important applications of Linear Algebra:
  - Mathematical optimization: linear programming
  - Discrete optimization: integer programming

$\rightsquigarrow$ Practical experience with computer software

$\rightsquigarrow$ The solution to really many real-life problems will be within short reach

Contents of the Course (1/2)

(see Syllabus)

Linear Algebra: manipulation of matrices and vectors with some theoretical background

Linear Algebra

1 Matrices and vectors - Matrix algebra, Geometric insight
2 Systems of Linear Equations - Gaussian elimination
3 Matrix inversion and determinants
4 Vector spaces
5 Linear Transformations - Matrix representation
6 Diagonalization - Eigenvalues and Eigenvectors
(7 Orthogonality or Numerical Methods)

Contents of the Course (2/2)

(see Syllabus)

Linear Programming

1 Introduction - Linear Programming, Notation
2 Linear Programming, Simplex Method
3 Exception Handling
4 Duality Theory
5 Sensitivity
6 Revised Simplex Method

Integer Linear Programming

7 Modeling Examples, Good Formulations, Relaxations
8 Well Solved Problems
9 Network Optimization Models (Max Flow, Min cost flow, matching)
10 Cutting Planes & Branch and Bound
11 More on Modeling

Practical Information

Teacher: Marco Chiarandini (http://www.imada.sdu.dk/~marco)
Instructor: Jens Østergaard

Schedule:

- Introductory classes: 44 hours (22 classes)
- Training classes: 50 hours
  - Exercises: 42 hours
  - Laboratory: 8 hours

Alternative views of the schedule:

- http://mitsdu.sdu.dk/skema, SDU Mobile
- Official course description (læserplaner)
- http://www.imada.sdu.dk/~marco/DM559

Communication Means

- BlackBoard (BB) $\Leftrightarrow$ Main Web Page (WWW) (link http://www.imada.sdu.dk/~marco/DM559)
- Announcements in BlackBoard
- Discussion Board in (BB) - allowed anonymous posting and rating
- Write to Marco (marco@imada.sdu.dk) and to instructor
- Ask peers
- You are welcome to visit me in my office in working hours (8-16)

$\leadsto$ It is good to ask questions!!

$\leadsto$ Let me know when you think we should do things differently! Things can be changed.

Books

Linear Algebra Part:

- AR Howard Anton and Chris Rorres. Elementary Linear Algebra. 11th Edition. 2014. Wiley

Other books:

- AH Martin Anthony and Michele Harvey, Linear Algebra, Concepts and Methods. 2012. Cambridge
- Le Steven J. Leon, Linear Algebra with Applications, 8th edition, Prentice Hall (2010).
- FSV Computing with Python: An introduction to Python for science and engineering Claus Führer, Jan Erik Solem, Olivier Verdier

Books

Linear and Integer Programming Part:

LN Lecture Notes

MG J. Matousek and B. Gartner. Understanding and Using Linear Programming. Springer Berlin Heidelberg, 2007

Wo L.A. Wolsey. Integer programming. John Wiley & Sons, New York, USA, 1998

Other books:

HL Frederick S Hillier and Gerald J Lieberman, Introduction to Operations Research, 9th edition, 2010

- ...

Course Material

Main Web Page (WWW) is the main reference for list of contents (ie[1], syllabus, pensum).

It contains:

- slides
- list of topics and references
- exercises
- links
- software

[1] ie = id est, eg = exempli gratia, wrt = with respect to

Assessment

- 7.5 ECTS
- (Four) Obligatory Assignments, pass/fail, evaluation by teacher
  both theoretical and practical (programming) exercises
- 4 hour written exam, 7-grade scale, external censor
  (theory part)
  similar to exercises in class and past exams
  in June
- (language: Danish and/or English)

Obligatory Assignments

- Small practical tasks must be passed to attend the written exam
- Individual work
- They require the use of Python + a MILP Solver (2nd part) See Tools from Public Web Page

Training Sessions

- Prepare them in advance to get out the most
- Best carried out in small groups
- Exam rehearsal (in June?)

Some Numbers

DM559

Obligatory Assignments

In Blackboard	60
Registered	43
Submitted	24
Passed	19

Final exam:

Registered	19
Showed up	14
Passed	6

Reexam:

Registered	12
Showed up	6
Passed	4

DM545

Obligatory Assignments

In Blackboard	60
Registered	27
Submitted	19
Passed	18

Final exam:

Registered	29
Showed up	21
Passed	16

Reexam:

Registered	12
Showed up	4
Passed	4

Who is here?

65 registered in BlackBoard...

- Computer Science
  (2nd year, 4th semester)
- Applied Mathematics?
- Math-economy?
- Returning CS

Prerequisites

- Goodwill
- Algorithms and Data Structures,
  Programming

Coding

Coding smal examples can help you to understand math!
Beside:

- listening to lectures
- watching an instructor work through a derivation
- working through numerical examples by hand

you can learn by doing by interacting with Python.

- jupyter.org/ipython (= interactive python)
- Python 2.7 or 3.5?

Outline

1. Course Organization

2. Preliminaries

Overview

- Notation
- Martices and vectors:
  - matrix arithmetic operations (addition, subtraction, and multiplication)
  - scalar multiplication and transposition.

Sets

- A set is a collection of objects. eg.: $A = \{1, 2, 3\}$
- $A = \{n \mid n \text{ is a whole number and } 1 \leq n \leq 3\}$
  ('$|$' reads 'such that')
- $B = \{x \mid x \text{ is a reader of this book}\}$
- $x \in A$
  x belongs to A
- set of no members: empty set, denoted $\emptyset$
- if a set S is a (proper) subset of a set T, we write ($S \subset T$) $T \subseteq S$
  $\{1, 2, 5\} \subset \{1, 2, 4, 5, 6, 30\}$
- for two sets A and B, the union $A \cup B$ is $\{x \mid x \in A \text{ or } x \in B\}$
- for two sets A and B, the intersection $A \cap B$ is $\{x \mid x \in A \text{ and } x \in B\}$
  $\{1, 2, 3, 5\}$ and $B = \{2, 4, 5, 7\}$, then $A \cap B = \{2, 5\}$

Numbers

- set of real numbers: $\mathbb{R}$
- set of natural numbers: $\mathbb{N} = \{1, 2, 3, 4, ...\}$ (positive integers)
- set of all integers: $\mathbb{Z} = \{..., -3, -2, -1, 0, 1, 2, 3, ...\}$
- set of rational numbers: $\mathbb{Q} = \{p/q \mid p, q \in \mathbb{Z}, q \neq 0\}$
- set of complex numbers: $\mathbb{C}$
- absolute value (non-negative):

$$|a| = \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a \leq 0 \end{cases}$$

  $|a + b| \leq |a| + |b|$, $a, b \in \mathbb{R}$

- the set $\mathbb{R}^2$ is the set of ordered pairs (x, y) of real numbers (eg, coordinates of a point wrt a pair of axes, the Cartesian plane)

Matrices and Vectors

- A matrix is a rectangular array of numbers or symbols. It can be written as

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

- An $n \times 1$ matrix is a column vector, or simply a vector:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

- the set $\mathbb{R}^n$ is the set of vectors $[x_1, x_2, \ldots, x_n]^T$ of real numbers (eg, coordinates of a point wrt an n-dimensional space, the Euclidean Space)

Basic Algebra

Elementary Algebra: the study of symbols and the rules for manipulating symbols. It differs from arithmetic in the use of abstractions, such as using letters to stand for numbers that are either unknown or allowed to take on many values

- collecting up terms: eg. $2a + 3b - a + 5b = a + 8b$
- multiplication of variables: eg:

$$a(-b) - 3ab + (-2a)(-4b) = -ab - 3ab + 8ab = 4ab$$

- expansion of bracketed terms: eg:

$$\begin{aligned} -(a-2b) &= -a + 2b, \\ (2x-3y)(x+4y) &= 2x^2 - 3xy + 8xy - 12y^2 \\ &= 2x^2 + 5xy - 12y^2 \end{aligned}$$

- $a^r a^s = a^{r+s}, \quad (a^r)^s = a^{rs}, \quad a^{-n} = 1/a^n,$
  $a^{1/n} = x \iff x^n = a, \quad a^{m/n} = (a^{1/n})^m$

Variables

- In Mathematics and Statistics, a variable is an alphabetic character representing a value, which is unknown. They are used in symbolic calculations.
  Commonly given one-character names.

- in contrast, a constant or given or scalar is a known real number

- In Computer Science, a variable is a storage location paired with an associated identifier, which contains a value, which may be known or unknown.
  Commonly given long, explanatory names.

Functions

- a function f on a set $\mathcal{X}$ into a set $\mathcal{Y}$ is a rule that assigns a unique element $f(x)$ in S to each element x in $\mathcal{X}$.

$$y = f(x)$$

y dependent variable

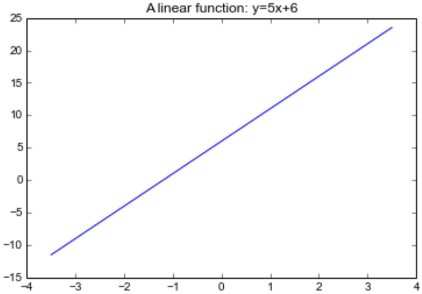
x independent variable

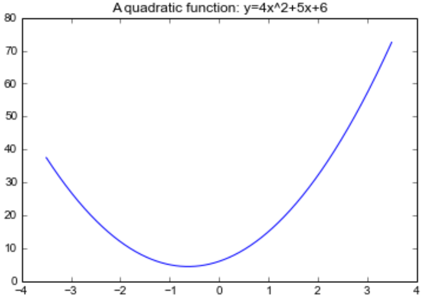
- a linear function has only sums and scalar multiplications, that is, for variable $x \in \mathbb{R}$ and scalars $a, b \in \mathbb{R}$:

$$f(x) := ax + b$$

Graphs of Functions

The graph of a function f consists of those points in the Cartesian plane whose coordinates (x, y) are pairs of input-output values for f.





Equations

- for a linear equation: $ax + b = 0$ with $a, b \in \mathbb{R}$
  the solution is a real number x for which the equation is true

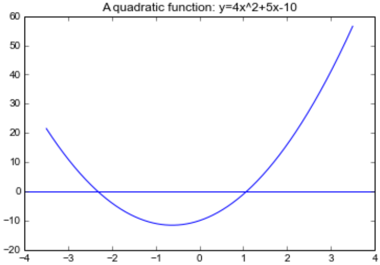
- Quadratic equation

$$ax^2 + bx + c = 0, \qquad a \neq 0.$$

- closed form or analytical solution:

$$x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

$$x_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$



- Quadratic equation

$$ax^2 + bx + c = 0, \qquad a \neq 0.$$

- quadratic formula:

$$x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \qquad x_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

  the term $b^2 - 4ac$ is called discriminant

- Solutions from discriminant:
  - if $b^2 - 4ac > 0 \implies$ two real solutions
  - if $b^2 - 4ac = 0 \implies$ exactly one solution: $x = -b/(2a)$
  - if $b^2 - 4ac < 0 \implies$ no real solution but complex solutions

- Can be solved also by factorization, eg:

$$x^2 - 6x + 5 = (x - 1)(x - 5) = 0$$

  then either $x - 1 = 0$ or $x - 5 = 0$.
  or by completing the square $\implies a(x + d)^2 + e = 0$, eg:
  i) $x^2 + 6x + 9 = 0$, and ii) $x^2 - 2x + 3 = 0$

Polynomial Equations

- A polynomial of degree n in x is an expression of the form:

$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

where the a_i are real constants, $a_n \neq 0$, and x is a real variable.

- $P_n(x) = 0$ has at most n solutions, eg:

$$x^3 - 7x + 6 = (x-1)(x-2)(x+3) = 0,$$

which are called roots or zeros of $P_n(x)$

- No general (closed) formula

- If α is a solution then $(x - \alpha)$ must be a factor for $P_n(x)$
We find α by trial and error and then set $(x-\alpha)Q(x)$ where $Q(x)$ is a polynomial of degree $n-1$

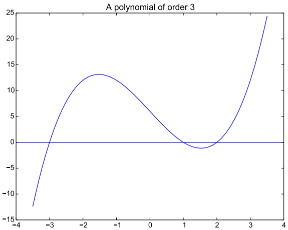
- Eg, $x^3 - 7x + 6$

In Python

```
import matplotlib.pyplot as plt
import numpy as np
a=[1,0,-7,6]
P=np.poly1d(a)
print(P)
```

```
   3
1 x - 7 x + 6
```

```
x = np.linspace(-3.5, 3.5, 500)
plt.plot(x, P(x), '-')
plt.axhline(y=0)
plt.title('A polynomial of order 3');
```



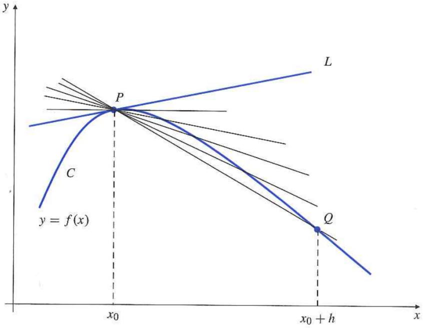
Differentiation

A line L through a point $(x_0, f(x_0))$ of f can be described by:

$$y = m(x - x_0) + f(x_0)$$

The derivative is the slope of the line that is tangent to the curve:

$$y = f'(x_0)(x - x_0) + f(x_0)$$



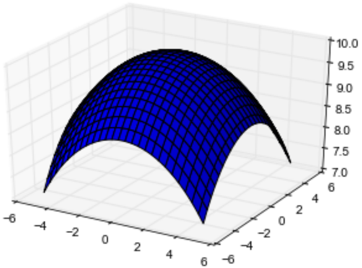
Functions of Several Variables

- A function f of n real variables is a rule that assigns a unique real number $f(x_1, x_2, \ldots, x_n)$ to each point $(x_1, x_2, \ldots, x_n)$

Example in $\mathbb{R}^2$:

$$f(x, y) = \sqrt{10^2 - x^2 - y^2}$$

$$x^2 + y^2 + z^2 = 10$$



Partial Derivatives

- The first partial derivative of the function $f(x, y)$ with respect to the variables x and y are:

$$f_1(x, y) = \lim_{h \to 0} \frac{f(x+h, y) - f(x, y)}{h} = \frac{\partial}{\partial x} f(x, y)$$

$$f_2(x, y) = \lim_{k \to 0} \frac{f(x, y+k) - f(x, y)}{k} = \frac{\partial}{\partial y} f(x, y)$$

- Their value in a point (x_0, y_0) is given by:

$$f_1(x_0, y_0) = \left(\frac{\partial}{\partial x} f(x, y) \right) \Bigg|_{(x_0, y_0)}$$

$$f_2(x_0, y_0) = \left(\frac{\partial}{\partial y} f(x, y) \right) \Bigg|_{(x_0, y_0)}$$

Trigonometry

- sine and cosine functions, $\sin\theta$ and $\cos\theta$, geometrical meaning
- angles measured in radiants rather than degrees ($\pi = 3.141..., \pi = 180$)
- $\cos x = \sin(x + \pi/2)$
- $\sin^2\theta + \cos^2\theta = 1$
- $\sin(\theta + \phi) = \sin\theta\cos\phi + \cos\theta\sin\phi$
- $\cos(\theta + \phi) = \cos\theta\cos\phi - \sin\theta\sin\phi$

DM559

Linear and Integer Programming

Lecture 2
Matrices and Vectors

Marco Chiarandini

Department of Mathematics & Computer Science
University of Southern Denmark

Outline

Outline

Matrices and Vectors

Definition (Matrix)

A matrix is a rectangular array of numbers or symbols. It can be written as

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

- We denote this array by a single letter A or by (a_{ij}) and
- we say that A has m rows and n columns, or that it is an $m \times n$ matrix.
- The size of A is $m \times n$.
- The number a_{ij} is called the (i, j) entry or scalar.

- A square matrix is an $n \times n$ matrix.
- The diagonal of a square matrix is the list of entries $a_{11}, a_{22}, \ldots, a_{nn}$
- The diagonal matrix is a matrix $n \times n$ with $a_{ij} = 0$ if $i \neq j$(ie, a square matrix with all the entries which are not on the diagonal equal to 0):

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{mn} \end{bmatrix}$$

Definition (Equality)

Two matrices are equal if they have the same size and if corresponding entries are equal. That is, if $A = (a_{ij})$ and $B = (b_{ij})$ are both $m \times n$ matrices, then:

$$A = B \iff a_{ij} = b_{ij} \qquad 1 \leq i \leq m, 1 \leq j \leq n$$

Matrix Addition and Multiplication

Definition (Addition)

If $A = (a_{ij})$ and $B = (b_{ij})$ are both $m \times n$ matrices, then

$$A + B = (a_{ij} + b_{ij}) \qquad 1 \leq i \leq m, 1 \leq j \leq n$$

Definition (Scalar Multiplication)

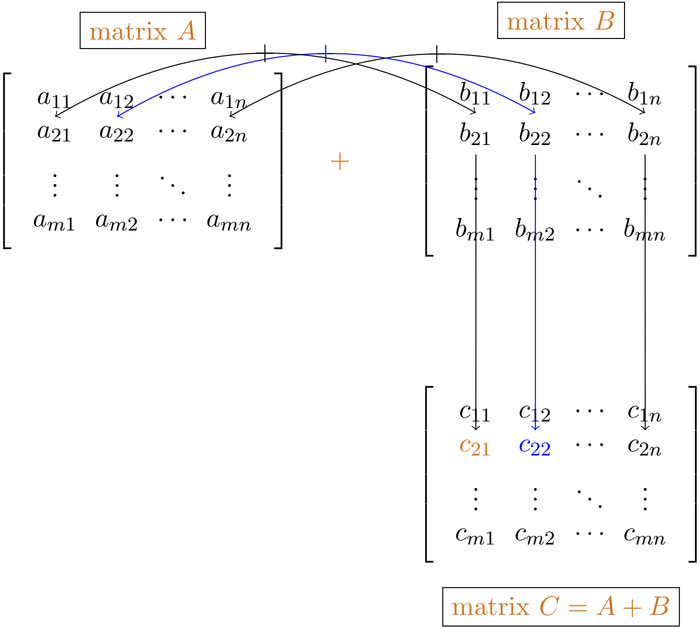
If $A = (a_{ij})$ is an $m \times n$ matrix and $\lambda \in \mathbb{R}$, then

$$\lambda A = (\lambda a_{ij}) \qquad 1 \leq i \leq m, 1 \leq j \leq n$$

Eg:

$$A + B = \begin{bmatrix} 3 & 1 & 2 \\ 0 & 5 & -2 \end{bmatrix} + \begin{bmatrix} -1 & 1 & 4 \\ 2 & -3 & 1 \end{bmatrix} = ?$$

$$-2A = ?$$


matrix A
matrix B
$+$
matrix $C = A + B$

Matrix Multiplication

Two matrices can be multiplied together, depending on the size of the matrices

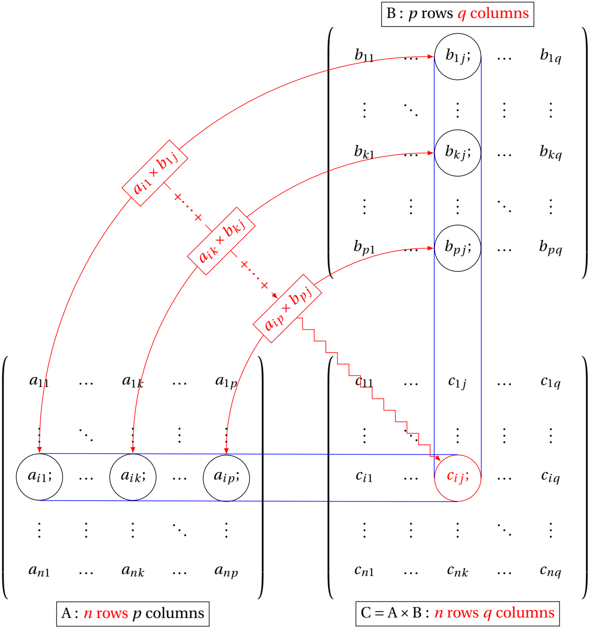
Definition (Matrix Multiplication)

If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then the product is the matrix $AB = C = (c_{ij})$ with

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

$$\begin{bmatrix} \\ \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \\ \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix}$$

What is the size of C?


B : p rows q columns
A : n rows p columns
C = A × B : n rows q columns
a_{i1} × b_{1j}
a_{ik} × b_{kj}
a_{ip} × b_{pj}
+ ... +
+ ... +

$$AB = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 1 & 2 & 4 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 1 & 1 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 5 & 3 \\ 1 & 14 \\ 9 & -1 \end{bmatrix}$$

$$(2)(3) + (0)(1) + (1)(-1) = 5$$

The motivation behind this definition will become clear later. It is exactly what is needed in our study of linear algebra

- $AB \neq BA$ in general, ie, not commutative
  try with the example of previous slide...

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Matrix Algebra

Matrices are useful because they provide compact notation and we can perform algebra with them

Bear in mind to use only operations that are defined. In the following rules, the sizes are dictated by the operations being defined.

- commutative $A + B = B + A$. Proof?
- associative:
  - $(A + B) + C = A + (B + C)$
  - $\lambda(AB) = (\lambda A)B = A(\lambda B)$
  - $(AB)C = A(BC)$

Size?

- distributive:
  - $A(B + C) = AB + AC$
  - $(B + C)A = BA + CA$
  - $\lambda(A + B) = \lambda A + \lambda B$

Why both first two rules?

Definition (Zero Matrix)

A zero matrix, denoted 0, is an $m \times n$ matrix with all entries zero:

$$\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

- additive identity: $A - A = 0$
  - $A + 0 = A$
  - $A - A = 0$
  - $0A = 0, A0 = 0$

Definition (Identity Matrix)

The $n \times n$ identity matrix, denoted I_n or I is the diagonal matrix with $a_{ii} = 1$:
zero:

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

- multiplicative identity (like 1 does for scalars)
  - $AI = A$ and $IA = A$

A of size $m \times n$.
What size is I?

Exercise: $3A + 2B = 2(B - A + C)$

Matrix Inverse

- If $AB = AC$ can we conclude that $B = C$?

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & -1 \\ 3 & 5 \end{bmatrix}, \qquad C = \begin{bmatrix} 8 & 0 \\ -4 & 4 \end{bmatrix}$$

$$AB = AC = \begin{bmatrix} 0 & 0 \\ 4 & 4 \end{bmatrix}$$

- $A + 5B = A + 5C \implies B = C$
  addition and scalar multiplication have inverses ($-A$ and $1/c$)

- Is there a multiplicative inverse?

Definition (Inverse Matrix)

The $n \times n$ matrix A is invertible if there is a matrix B such that

$$AB = BA = I$$

where I is the $n \times n$ identity matrix. The matrix B is called the inverse of A and is denoted by A^{-1}.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}$$

Theorem

If A is an $n \times n$ invertible matrix, then the matrix A^{-1} is unique.

Proof: Assume A has two inverses B, C so $AB = BA = I$ and $AC = CA = I$. Consider the product CAB:

$$CAB = C(AB) = CI = C \qquad \text{associativity} + AB = I$$
$$CAB = (CA)B = IB = B \qquad \text{associativity} + CA = I$$

- If a matrix has an inverse we say that it is invertible or non-singular
  If a matrix has no inverse we say that it is non-invertible or singular
  Eg:

$$\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- If

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad ad - bc \neq 0$$

  then A has the inverse

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \qquad ad - bc \neq 0$$

- The scalar $ad - bc$ is called determinant of A and denoted $|A|$.

Matrix Inverse

Back to the question:

- If $AB = AC$ can we conclude that $B = C$?
  If A is invertible then the answer is yes:

$$A^{-1}AB = A^{-1}AC \implies IB = IC \implies B = C$$

- But $AB = CA$ then we cannot conclude that $B = C$.

Properties of the Inverse

Let A be invertible $\implies$ A^{-1} exists

- $(A^{-1})^{-1} = A$

- $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$
  the inverse of the matrix (λA) is a matrix C that satisfies $(\lambda A)C = C(\lambda A) = I$. Using matrix algebra:

$$(\lambda A)\left(\frac{1}{\lambda}A^{-1}\right) = \lambda\frac{1}{\lambda}AA^{-1} = I \text{ and } \left(\frac{1}{\lambda}A^{-1}\right)(\lambda A) = \frac{1}{\lambda}\lambda A^{-1}A = I$$

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

Powers of a matrix

For A an $n \times n$ matrix and $r \in \mathbb{N}$

$$A^r = \underbrace{AA \dots A}_{r \text{ times}}$$

For the associativity of matrix multiplication:

- $(A^r)^{-1} = (A^{-1})^r$
- $A^r A^s = A^{r+s}$
- $(A^r)^s = A^{rs}$

Transpose Matrix

Definition (Transpose)

The transpose of an $m \times n$ matrix A is the $n \times m$ matrix B defined by

$$b_{ij} = a_{ji} \qquad \text{for } i = 1, \ldots, n \text{ and } j = 1, \ldots, m$$

It is denoted A^T

$$A = (a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad A^T = (a_{ji}) = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nm} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

Note that if D is a diagonal matrix: $D^T = D$

Properties of the transpose

- $(A^T)^T = A$
- $(\lambda A)^T = \lambda A^T$
- $(A + B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$ (consider first which matrix sizes make sense in the multiplication, then rewrite the terms)
- if A is invertible, $(A^T)^{-1} = (A^{-1})^T$

$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$$

$$(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$$

Symmetric Matrix

Definition (Symmetric Matrix)

A matrix A is symmetric if it is equal to its transpose, $A = A^T$.
(only square matrices can be symmetric)

Vectors

- An $n \times 1$ matrix is a column vector, or simply a vector:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

The numbers $v_1, v_2, \ldots$ are known as the components (or entries) of v.

- A row vector is a $1 \times n$ matrix
- We write vectors in lower boldcase type (writing by hand we can either underline them or add an arrow over $\mathbf{v}$).
- Addition and scalar multiplication are defined for vectors as for $n \times 1$ matrices:

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{bmatrix} \qquad \lambda\mathbf{v} = \begin{bmatrix} \lambda v \\ \lambda v \\ \vdots \\ \lambda v \end{bmatrix}$$

- For a fixed n, the set of vectors together with the operations of addition and multiplication form the set $\mathbb{R}^n$, usually called Euclidean space

- For vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$ and scalars $\alpha_1, \alpha_2, \ldots, \alpha_k$ in $\mathbb{R}$, the vector

  $$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_k \mathbf{v}_k$$

  is known as linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$
- A zero vector is denoted by $\mathbf{0}$;
  $\mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{v}$;
  $0\mathbf{v} = \mathbf{0}$

- The matrix product of $\mathbf{v}$ and $\mathbf{w}$ cannot be calculated
- The matrix product of $\mathbf{v}^T\mathbf{w}$ gives an 1×1 matrix
- The matrix product of $\mathbf{v}\mathbf{w}^T$ gives an $n \times n$ matrix

Inner product of two vectors

Definition (Inner product)

Given

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

the inner product denoted $\langle \mathbf{v}, \mathbf{w} \rangle$, is the real number given by

$$\langle \mathbf{v}, \mathbf{w} \rangle = \left\langle \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \right\rangle = v_1 w_1 + v_2 w_2 + \ldots + v_n w_n = \mathbf{v}^T \mathbf{w}$$

It is also called scalar product or dot product (and written $\mathbf{v} \cdot \mathbf{w}$).

$$\mathbf{v}^T\mathbf{w} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = v_1 w_1 + v_2 w_2 + \ldots + v_n w_n =$$

Theorem

The inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n, \qquad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

satisfies the following properties for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ *and for all* $\alpha \in \mathbb{R}$*:*

- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
- $\alpha \langle \mathbf{x}, \mathbf{y} \rangle = \langle \alpha\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \alpha\mathbf{y} \rangle$
- $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
- $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ *and* $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ *if and only if* $\mathbf{x} = \mathbf{0}$

Note: vectors from different Euclidean spaces live in different 'worlds'

Vectors and Matrices

Theorem

Let A be an $m \times n$ matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

and denote the columns of A by the column vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, so that

$$\mathbf{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}, \qquad i = 1, \ldots, n.$$

Then if $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ is any vector in $\mathbb{R}^n$

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n$$

(vector $A\mathbf{x}$ in $\mathbb{R}^m$ as a linear combination of the column vectors of A)

Resume

1. Matrices and Vectors

DM559

Linear and Integer Programming

Lecture 3
Matrices and Vectors: Geometric Insight

Marco Chiarandini

Department of Mathematics & Computer Science
University of Southern Denmark

Outline

Outline

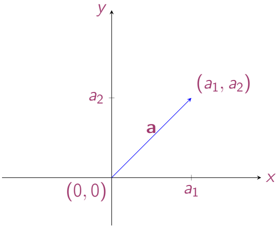
1. Geometric Insight

2. Linear Systems

Geometric Insight

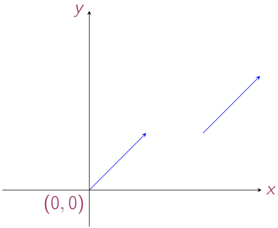
- Set $\mathbb{R}$ can be represented by real-number line. Set $\mathbb{R}^2$ of real number pairs (a_1, a_2) can be represented by the Cartesian plane.

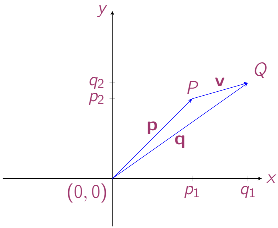
- To a point in the plane $A = (a_1, a_2)$ it is associated a position vector $\mathbf{a} = (a_1, a_2)^T$, representing the displacement from the origin $(0, 0)$. ◇



- Two displacement vectors of same length and direction are considered to be equal even if they do not both start from the origin

- If object displaced from O to P by displacement $\mathbf{p}$ and from P to Q by displacement $\mathbf{v}$, then the total displacement satisfies $\mathbf{q} = \mathbf{p} + \mathbf{v} = \mathbf{v} + \mathbf{q}$





- $\mathbf{v} = \mathbf{q} - \mathbf{p}$, think of $\mathbf{v}$ as the vector that is added to $\mathbf{p}$ to obtain $\mathbf{q}$.

- the length of a vector $\mathbf{a} = (a_1, a_2)^T$ is denoted by $||\mathbf{a}||$ and from Pythagoras

$$||\mathbf{a}|| = \sqrt{a_1^2 + a_2^2} = \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle}$$

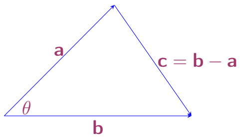
- the direction is given by the components of the vector
- the unit vector can be derived by normalizing it, that is:

$$\mathbf{u} = \frac{1}{||\mathbf{v}||}\mathbf{v}$$

Theorem (Inner Product)

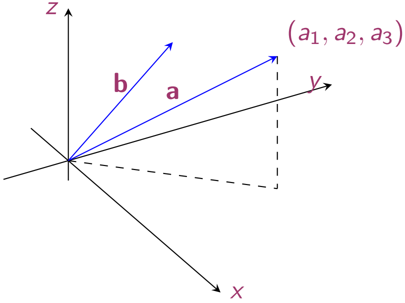
Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ and let θ denote the angle between them. Then (from the law of cosines),

$$\langle \mathbf{a}, \mathbf{b} \rangle = ||\mathbf{a}|| \, ||\mathbf{b}|| \cos\theta$$



Two vectors $\mathbf{a}$ and $\mathbf{b}$ are orthogonal (or normal or perpendicular) if and only if $\langle \mathbf{a}, \mathbf{b} \rangle = 0$.

Vectors in $\mathbb{R}^3$



$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \qquad \|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

$$\langle \mathbf{a}, \mathbf{b} \rangle = \|\mathbf{a}\| \, \|\mathbf{b}\| \cos\theta$$

Lines in $\mathbb{R}^2$

- Cartesian equation of a line: $y = ax + b$

- another way is by giving position vectors.
  We can let $x = t$ where t is any real number. Then
  $y = ax + b = at + b$. Hence the position vector $\mathbf{x} = (x, y)^T$

  $$\mathbf{x} = \begin{bmatrix} t \\ at + b \end{bmatrix} = t \begin{bmatrix} 1 \\ a \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix} = t\mathbf{v} + (0, b)^T, \qquad t \in \mathbb{R}$$

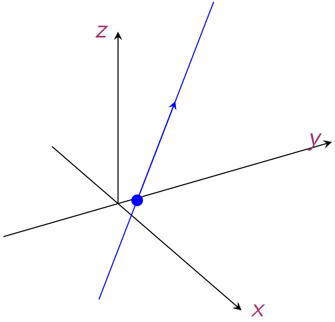
- To derive the Cartesian equation: locate one particular point on the line, eg, the y intercept. Then the position vector of any point on the line is a sum of two displacements, first going to the point and then along the direction of the line. Try with $P = (-1, 1)$ and $Q = (3, 2)$

- In general, any line in $\mathbb{R}^2$ is given by a vector equation with one parameter of the form

  $$\mathbf{x} = \mathbf{p} + t\mathbf{v}$$

  where $\mathbf{x}$ is the position vector, $\mathbf{p}$ is any particular point and $\mathbf{v}$ is the direction of the line

Lines in $\mathbb{R}^3$



$$\mathbf{x} = \mathbf{p} + t\mathbf{v}$$

$$\mathbf{x} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} 3 \\ 7 \\ 2 \end{bmatrix} + s \begin{bmatrix} -3 \\ -6 \\ 3 \end{bmatrix}, \quad s, t \in \mathbb{R}$$

Are these lines intersecting?
What is the Cartesian equation of the first?

In $\mathbb{R}^2$, two lines are:

- parallel
- intersecting in a unique point

In $\mathbb{R}^3$, two lines are:

- parallel
- intersecting in a unique point
- skew (lay on two parallel planes)

What about these lines? Do they intersect? Are they coplanar?

$$L_1 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$L_2 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ 1 \end{bmatrix} + t \begin{bmatrix} -2 \\ 1 \\ 7 \end{bmatrix}$$

Planes in $\mathbb{R}^3$

Vector parametric equation:

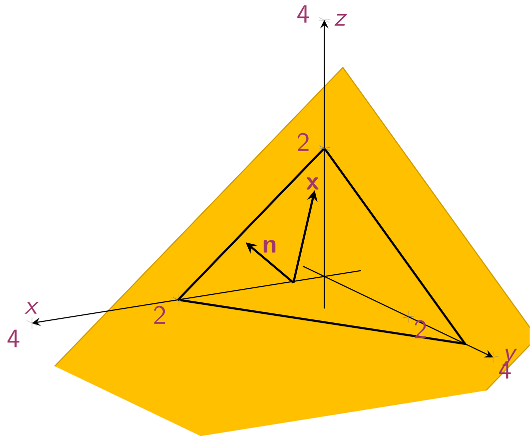
- The position of vectors of points on a plane is described by:

  $$\mathbf{x} = \mathbf{p} + s\mathbf{v} + t\mathbf{w}, \quad s, t \in \mathbb{R}$$

  provided $\mathbf{v}$ and $\mathbf{w}$ are non-zero and not parallel.
  ($\mathbf{p}$ position vector, $\mathbf{v}$ and $\mathbf{w}$ displacement vectors).

- How is the plane through the origin? What if $\mathbf{v}$ and $\mathbf{w}$ are parallel?

- Two intersecting lines determine a plane. What is its description?


4
z
2
x
n
x
2
4
2
y
4

Alternative Description of Planes

Cartesian equation:

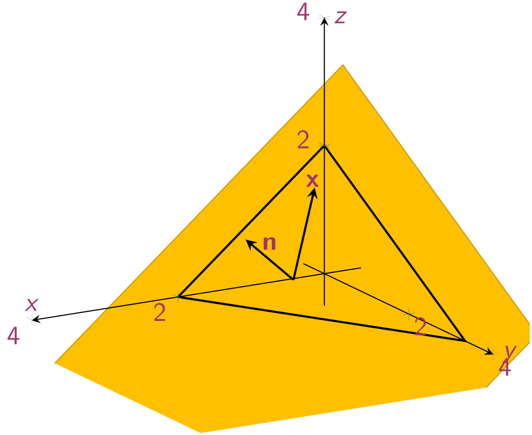
- Let $\mathbf{n}$ be a given vector in $\mathbb{R}^3$. All positions represented by postion vectors $\mathbf{x}$ that are orthogonal to $\mathbf{n}$ describe a plane through the origin. ($\mathbf{n}$ is called a normal vector to the plane)

- Vectors $\mathbf{n}$ and $\mathbf{x}$ are orthogonal iff

$$\langle \mathbf{n}, \mathbf{x} \rangle = 0,$$

hence this equation describes a plane.
If $\mathbf{n} = (a, b, c)^T$ and $\mathbf{x} = (x, y, z)^T$, then the equation becomes:

$$ax + by + cz = 0$$


4
z
2
x
n
x
2
4
2
y
4

- For a point P on the plane with position vector $\mathbf{p}$ and a position vector $\mathbf{x}$ of any other point on the plane, the displacement vector $\mathbf{x} - \mathbf{p}$ lies on the plane and $\mathbf{n} \perp \mathbf{x} - \mathbf{p}$

- Conversely, if the position vector $\mathbf{x}$ of a point is such that

$$\langle \mathbf{n}, \mathbf{x} - \mathbf{p} \rangle = 0$$

  then the point represented by $\mathbf{x}$ lies on the plane.

- hence, $\langle \mathbf{n}, \mathbf{x} \rangle = \langle \mathbf{n}, \mathbf{p} \rangle = d$ and the equation becomes:

$$ax + by + cz = d$$

Eg.: $2x - 3y - 5z = 2$ has $\mathbf{n} = (2, -3, -5)^T$ and passes through $(0, 0, e)$

Vector parametric equation $\iff$ Cartesian equation

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = s \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + t \begin{bmatrix} 2 \\ 1 \\ 7 \end{bmatrix} = s\mathbf{v} + t\mathbf{w}, \quad s, t \in \mathbb{R}$$

$$3x - y + z = 0, \qquad \mathbf{n} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$\langle \mathbf{n}, \mathbf{v} \rangle = 0, \langle \mathbf{n}, \mathbf{w} \rangle = 0 \text{ and } \langle \mathbf{n}, s\mathbf{v} + t\mathbf{w} \rangle = 0 \text{ for } s, t \in \mathbb{R}$$

What will change if the plane does not pass through the origin?

Are the two following planes parallel?

$$x + 2y - 3x = 0 \text{ and } -2x - 4y + 6z = 4$$

and these?

$$x + 2y - 3x = 0 \text{ and } x - 2y + 5z = 4$$

Lines and Hyperplanes in $\mathbb{R}^n$

- Point in $\mathbb{R}^n$: $\mathbf{a} = (a_1, a_2, \ldots, a_n)^T$
- Length of a vector $\mathbf{x} = (x_, x_2, \ldots, x_n)^T$

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

- The vectors in $\mathbb{R}^n$ are orthogonal iff

$$\langle \mathbf{v}, \mathbf{w} \rangle = 0.$$

- Line:

$$\mathbf{x} = \mathbf{p} + t\mathbf{v}, \quad t \in \mathbb{R}$$

How many Cartesian equations?

- The set of points $(x_1, x_2, \ldots, x_n)$ that satisfy a Cartesian equation

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = d$$

is called hyperplane. ($\langle \mathbf{n}, \mathbf{x} - \mathbf{p} \rangle = 0$.) What is the vector equation?

Outline

1. Geometric Insight

2. Linear Systems

Systems of Linear Equations

Definition (System of linear equations, aka linear system)

A system of m linear equations in n unknowns $x_1, x_2, \ldots, x_n$ is a set of m equations of the form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ & & & & \vdots & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

The numbers a_{ij} are known as the coefficients of the system.

We say that $s_1, s_2, \ldots, s_n$ is a solution of the system if all m equations hold true when

$$x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$$

Examples

$$\begin{array}{rcrcrcrcrcl} x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & = & 3 \\ 2x_1 & + & x_2 & + & x_3 & + & x_4 & + & 2x_5 & = & 4 \\ x_1 & - & x_2 & - & x_3 & + & x_4 & + & x_5 & = & 5 \\ x_1 & & & & & + & x_4 & + & x_5 & = & 4 \end{array}$$

has solution

$$x_1 = -1, x_2 = -2, x_3 = 1, x_4 = 3, x_5 = 2.$$

Is it the only one?

$$\begin{array}{rcrcrcrcrcl} x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & = & 3 \\ 2x_1 & + & x_2 & + & x_3 & + & x_4 & + & 2x_5 & = & 4 \\ x_1 & - & x_2 & - & x_3 & + & x_4 & + & x_5 & = & 5 \\ x_1 & & & & & + & x_4 & + & x_5 & = & 6 \end{array}$$

has no solutions

Definition (Coefficient Matrix)

The matrix $A = (a_{ij})$, whose (i, j) entry is the coefficient a_{ij} of the system of linear equations is called the coefficient matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Let $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ then

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}^{m \times n} \overbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}^{n \times 1} = \overbrace{\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}}^{n \times 1}$$

hence, the linear system can be written also as $A\mathbf{x} = \mathbf{b}$

Row operations

How do we find solutions?

$$\begin{array}{l|rcrcrcr}
\texttt{R1:} & x_1 & + & x_2 & + & x_3 & = & 3 \\
\texttt{R2:} & 2x_1 & + & x_2 & + & x_3 & = & 4 \\
\texttt{R3:} & x_1 & - & x_2 & + & 2x_3 & = & 5
\end{array}$$

Eliminate one of the variables from two of the equations

$$\begin{array}{l|rcrcrcr}
\texttt{R1'=R1:} & x_1 & + & x_2 & + & x_3 & = & 3 \\
\texttt{R2'=R2-2*R1:} & & - & x_2 & - & x_3 & = & -2 \\
\texttt{R3'=R3:} & x_1 & - & x_2 & + & 2x_3 & = & 5
\end{array}$$

$$\begin{array}{l|rcrcrcr}
\texttt{R1'=R1:} & x_1 & + & x_2 & + & x_3 & = & 3 \\
\texttt{R2'=R2:} & & - & x_2 & - & x_3 & = & -2 \\
\texttt{R3'=R3-R1:} & & - & 2x_2 & + & x_3 & = & 2
\end{array}$$

We can now eliminate one of the variables in the last two equations to obtain the solution

Row operations that do not alter solutions:

O1: multiply both sides of an equation by a non-zero constant

O2: interchange two equations

O3: add a multiple of one equation to another

These operations only act on the coefficients of the system
For a system $A\mathbf{x} = \mathbf{b}$:

$$\left[\, A \,\middle|\, \mathbf{b} \,\right] = \begin{bmatrix} 1 & 1 & 1 & 3 \\ 2 & 1 & 1 & 4 \\ 1 & -1 & 2 & 5 \end{bmatrix}$$

Augmented Matrix

Definition (Augmented Matrix and Elementary row operations)

For a system of linear equations $A\mathbf{x} = \mathbf{b}$ with

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

the augmented matrix of the system and the row operations are:

$$\left[\, A \,\middle|\, \mathbf{b} \,\right] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

RO1: multiply a row by a non-zero constant

RO2: interchange two rows

RO3: add a multiple of one row to another

DM559

Linear and Integer Programming

Lecture 4

Systems of Linear Equations

Marco Chiarandini

Department of Mathematics & Computer Science
University of Southern Denmark

Outline

1. Solving Linear Systems

Outline

Problem Statement

Given the system of linear equations:

$$\begin{array}{l|rcrcrcr}
\texttt{R1:} & x_1 & + & x_2 & + & x_3 & = & 3 \\
\texttt{R2:} & 2x_1 & + & x_2 & + & x_3 & = & 4 \\
\texttt{R3:} & x_1 & - & x_2 & + & 2x_3 & = & 5
\end{array}$$

Find whether it has any solution and in case characterize the solutions.

Augmented Matrix

Definition (Augmented Matrix and Elementary row operations)

For a system of linear equations $A\mathbf{x} = \mathbf{b}$ with

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

the augmented matrix of the system and the row operations are:

$$\left[\, A \,\middle|\, \mathbf{b} \,\right] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

RO1: multiply a row by a non-zero constant

RO2: interchange two rows

RO3: add a multiple of one row to another

They modify the linear system into an equivalent system (same solutions)

Gaussian Elimination: Example

Let's consider the system $A\mathbf{x} = \mathbf{b}$ with:

$$\left[\, A \,\middle|\, \mathbf{b} \,\right] = \begin{bmatrix} 1 & 1 & 1 & 3 \\ 2 & 1 & 1 & 4 \\ 1 & -1 & 2 & 5 \end{bmatrix}$$

1. Left most column that is not all zeros (it is column 1)
2. A non-zero entry at the top of this column (it is the one on the top)
3. Make the entry 1 (it is already)

$$\begin{bmatrix} 1 & 1 & 1 & 3 \\ 2 & 1 & 1 & 4 \\ 1 & -1 & 2 & 5 \end{bmatrix}$$

4. make all entries below the leading one zero:

$$\begin{array}{l} \texttt{R1'=R1} \\ \texttt{R2'=R2-2R1} \\ \texttt{R3'=R3-R1} \end{array} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & -1 & -1 & -2 \\ 0 & -2 & 1 & 2 \end{bmatrix}$$

Example, cntd. Row Echelon Form

5. Cover up the top row and apply steps (1) and (4) again
1. Left most column that is not all zeros is column 2
2. Non-zero entry at the top of the column
3. Make this entry the leading 1 by elementary row operations RO1 or RO2.
4. Make all entries below the leading 1 zero by RO3

$$\left[\begin{array}{cccc} \cancel{1} & \cancel{1} & \cancel{1} & \cancel{3} \\ 0 & -1 & -1 & -2 \\ 0 & -2 & 1 & 2 \end{array}\right] \rightarrow \left[\begin{array}{cccc} 1 & 1 & 1 & 3 \\ 0 & -1 & -1 & -2 \\ 0 & 0 & 3 & 6 \end{array}\right] \equiv \begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & = & 3 \\ & & x_2 & + & x_3 & = & 2 \\ & & & & x_3 & = & 2 \end{array}$$

Definition (Row echelon form)

A matrix is said to be in row echelon form (or echelon form) if it has the following three properties:

1. the first nonzero entry in each nonzero row is 1
2. a leading 1 in a lower row is further to the right
3. zero rows are at the bottom of the matrix

Back substitution

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & = & 3 \\ & & x_2 & + & x_3 & = & 2 \\ & & & & x_3 & = & 2 \end{array}$$

From the row echelon form we solve the system by back substitution:

- from the last equation: set $x_3 = 2$
- substitute x_3 in the second equation $\rightsquigarrow x_2$
- substitute x_2 and x_3 in the first equation $\rightsquigarrow x_1$

Reduced Row Echelon Form

In the augmented matrix representation:

6. Begin with the last row and add suitable multiples to each row above to get zero above the leading 1.

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

Definition (Reduced row echelon form)

A matrix is said to be in reduced (row) echelon form if it has the following properties:

1. The matrix is in row echelon form
2. Every column with a leading 1 has zeros elsewhere

From a Reduced Row Echelon Form (RREF) we can read the solution:

$$\left[\,A\,\middle|\,\mathbf{b}\,\right] = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

The system has a unique solution.
Is it a correct solution? Let's check:

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & = & 3 \\ 2x_1 & + & x_2 & + & x_3 & = & 4 \\ x_1 & - & x_2 & + & 2x_3 & = & 5 \end{array} \quad \rightsquigarrow \quad \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$$

Gaussian Elimination: Algorithm

Gaussian Elimination algorithm for solving a linear system:
(puts the augmented matrix in a form from which the solution can be read)

1. Find left most column that is not all zeros
2. Get a non-zero entry at the top of this column (pivot element)
3. Make this entry 1 by elementary row operations RO1 or RO2. This entry is called leading one
4. Add suitable multiples of the top row to rows below so that all entries below the leading one become zero
5. Cover up the top row and apply steps (1) and (4) again
   The matrix left is in (row) echelon form
6. Back substitution

Gauss-Jordan Reduction

Gauss Jordan Reduction algorithm for solving a linear system:
(puts the augmented matrix in a form from which the solution can be read)

1. Find left most column that is not all zeros
2. Get a non-zero entry at the top of this column (pivot element)
3. Make this entry 1 by elementary row operations RO1 or RO2. This entry is called leading one
4. Add suitable multiples of the top row to rows below so that all entries below the leading one become zero
5. Cover up the top row and apply steps (1) and (4) again
   The matrix left is in (row) echelon form
6. Begin with the last row and add suitable multiples to each row above to get zero above the leading 1.
   The matrix left is in reduced (row) echelon form

Will there always be exactly one solution?

$$\begin{array}{l|rcrcl} \texttt{R1:} & & & 2x_3 & = & 3 \\ \texttt{R2:} & 2x_2 & + & 3x_3 & = & 4 \\ \texttt{R3:} & & & x_3 & = & 5 \end{array} \quad \to \quad \left[\, A \,\middle|\, \mathbf{b} \,\right] = \begin{bmatrix} 0 & 0 & 2 & 3 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 2 & 3 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 1 & 5 \end{bmatrix} \to \begin{array}{c} \texttt{R2} \\ \texttt{R1} \\ \end{array} \begin{bmatrix} 0 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 1 & 5 \end{bmatrix} \to \begin{array}{c} \texttt{R1/2} \\ \\ \end{array} \begin{bmatrix} 0 & 1 & \frac{3}{2} & 2 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 1 & 5 \end{bmatrix} \to$$

$$\to \begin{array}{c} \\ \texttt{R3} \\ \texttt{R2} \end{array} \begin{bmatrix} 0 & 1 & \frac{3}{2} & 2 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 2 & 3 \end{bmatrix} \to \begin{array}{c} \\ \\ \texttt{R3-2R2} \end{array} \begin{bmatrix} 0 & 1 & \frac{3}{2} & 2 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & -7 \end{bmatrix} \to$$

$$\to \begin{array}{c} \\ \\ \texttt{-R3/7} \end{array} \begin{bmatrix} 0 & 1 & \frac{3}{2} & 2 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & \frac{3}{2} \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}$$

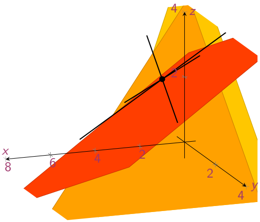
No Solution!

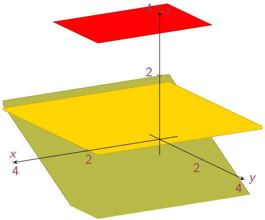
Definition (Consistent)

A system of linear equations is said to be consistent if it has at least one solution. It is inconsistent if there are no solutions.

$$\begin{cases} x_1 + x_2 + x_3 = 3 \\ 2x_1 + x_2 + x_3 = 4 \\ x_1 - x_2 + 2x_3 = 5 \end{cases} \qquad \begin{cases} 2x_3 = 3 \\ 2x_2 + 3x_3 = 4 \\ x_3 = 5 \end{cases}$$

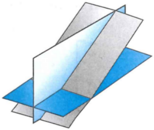
$$\left[\, A \,\middle|\, \mathbf{b} \,\right] = \begin{bmatrix} 1 & 1 & 1 & 3 \\ 2 & 1 & 1 & 4 \\ 1 & -1 & 2 & 5 \end{bmatrix} \qquad \left[\, A \,\middle|\, \mathbf{b} \,\right] = \begin{bmatrix} 0 & 0 & 2 & 3 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$



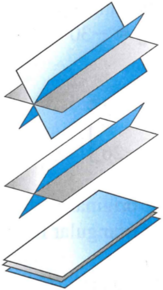


Geometric Interpretation

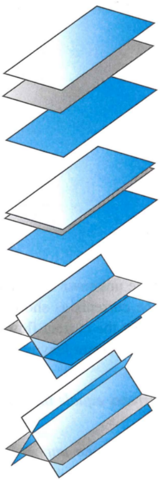
Three equations in three unknowns interpreted as planes in space



Unique solution



Infinitely many solutions



No solution

Definition (Overdetermined)

A linear system is said to be over-determined if there are more equations than unknowns. Over-determined systems are usually (but not always) inconsistent.

Definition (Underdetermined)

A linear system of m equations and n unknowns is said to be under-determined if there are fewer equations than unknowns ($m < n$). They have usually infinitely many solutions (never just one).

Linear systems with free variables

$$\begin{array}{rcrcrcrcrcr} x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & = & 3 \\ 2x_1 & + & x_2 & + & x_3 & + & x_4 & + & 2x_5 & = & 4 \\ x_1 & - & x_2 & - & x_3 & + & x_4 & + & x_5 & = & 5 \\ x_1 & & & & & + & x_4 & + & x_5 & = & 4 \end{array}$$

$$\left[\, A \,\middle|\, \mathbf{b} \,\right] = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 2 & 1 & 1 & 1 & 2 & 4 \\ 1 & -1 & -1 & 1 & 1 & 5 \\ 1 & 0 & 0 & 1 & 1 & 4 \end{bmatrix}$$

$$\rightarrow \quad \begin{array}{r} \\ \texttt{R2-2R1} \\ \texttt{R3-R1} \\ \texttt{R4-R1} \end{array} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & -1 & -1 & -1 & 0 & -2 \\ 0 & -2 & -2 & 0 & 0 & 2 \\ 0 & -1 & -1 & 0 & 0 & 1 \end{bmatrix}$$

$$\rightarrow \quad \begin{array}{r} \\ \texttt{(-1)R2} \\ \\ \\ \end{array} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & -2 & -2 & 0 & 0 & 2 \\ 0 & -1 & -1 & 0 & 0 & 1 \end{bmatrix}$$

$$\rightarrow \quad \begin{array}{r} \\ \\ \texttt{R3+2R2} \\ \texttt{R4+R2} \end{array} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 2 & 0 & 6 \\ 0 & 0 & 0 & 1 & 0 & 3 \end{bmatrix}$$

$$\rightarrow \quad \begin{array}{r} \\ \\ \texttt{(1/2)R3} \\ \\ \end{array} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 3 \end{bmatrix}$$

$$\rightarrow \quad \begin{array}{r} \\ \\ \\ \texttt{R4-R3} \end{array} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{Row echelon form}$$

$$
\to \quad \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

$$
\to \quad \begin{array}{l} \texttt{R1-R3} \\ \texttt{R2-R3} \\ \\ \\ \end{array} \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

$$
\to \quad \begin{array}{l} \texttt{R1-R2} \\ \\ \\ \\ \end{array} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

$$
\begin{array}{ccccccccccc}
x_1 & + & 0 & + & 0 & + & 0 & + & x_5 & = & 1 \\
 & + & x_2 & + & x_3 & + & 0 & + & 0 & = & -1 \\
 & & & & & + & x_4 & + & 0 & = & 3
\end{array}
$$

$$
\begin{array}{ccccccccccc}
x_1 & + & 0 & + & 0 & + & 0 & + & x_5 & = & 1 \\
 & + & x_2 & + & x_3 & + & 0 & + & 0 & = & -1 \\
 & & & & & + & x_4 & + & 0 & = & 3
\end{array}
$$

Definition (Leading variables)

The variables corresponding with leading ones in the reduced row echelon form of an augmented matrix are called leading variables. The other variables are called non-leading variables

- x_1, x_2 and x_4 are leading variables.
- x_3, x_5 are non-leading variables.
- we assign x_3, x_5 the arbitrary values $s, t \in \mathbb{R}$ and solve for the leading variables.
- there are infinitely many solutions, represented by the general solution:

$$
\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1-t \\ -1-s \\ s \\ 3 \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 3 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}
$$

Solution Sets

Theorem

A system of linear equations either has no solutions, a unique solution or infinitely many solutions.

Proof.

Let's assume the system $A\mathbf{x} = \mathbf{b}$ has two solutions $\mathbf{p}$ and $\mathbf{q}$. Then all points on the line connecting these two points are also solutions and so there are infinitely many solutions.

$A\mathbf{p} = \mathbf{b} \qquad \mathrm{A}\mathbf{q}\mathbf{=}\mathbf{b} \qquad \mathbf{p} \neq \mathbf{q}$

$\mathbf{v} = \mathbf{p} + t(\mathbf{q} - \mathbf{p}), t \in \mathbb{R}$

$A\mathbf{v} = A(\mathbf{p} + t(\mathbf{q} - \mathbf{p})) = A\mathbf{p} + t(A\mathbf{q} - A\mathbf{p}) = \mathbf{b} + t(\mathbf{b} - \mathbf{b}) = \mathbf{b}$

Homogeneous systems

Definition (Homogenous system)

An homogeneous system of linear equations is a linear system of the form $A\mathbf{x} = \mathbf{0}$.

- A homogeneous system $A\mathbf{x} = \mathbf{0}$ is always consistent $A\mathbf{0} = \mathbf{0}$.
- If $A\mathbf{x} = \mathbf{0}$ has a unique solution, then it must be the trivial solution $\mathbf{x} = \mathbf{0}$.

In the augmented matrix the last column stays always zero $\rightsquigarrow$ we can omit it.

Example

$$\begin{array}{ccccccccc} x & + & y & + & 3z & + & w & = & 0 \\ x & - & y & + & z & + & w & = & 0 \\ & & y & + & 2z & + & 2w & = & 0 \end{array} \qquad A = \begin{bmatrix} 1 & 1 & 3 & 1 \\ 1 & -1 & 1 & 1 \\ 0 & 1 & 2 & 2 \end{bmatrix}$$

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

$$\to \begin{bmatrix} 1 & 1 & 3 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & -5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = t \begin{bmatrix} 3 \\ 2 \\ -2 \\ 1 \end{bmatrix}, t \in \mathbb{R}$$

Theorem

If A is an $m \times n$ matrix with $m < n$, then $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions.

Proof.

- The system is always consistent since homogeneous.
- Matrix A brought in reduced echelon form contains at most m leading ones (variables).
- $n - m \geq 1$ non-leading variables

How about $A\mathbf{x} = \mathbf{b}$ with A $m \times n$ and $m < n$?
If the system is consistent, then there are infinitely many solutions.

Example

$$\begin{array}{rcrcrcrcr} x & + & y & + & 3z & + & w & = & 2 \\ x & - & y & + & z & + & w & = & 4 \\ & & y & + & 2z & + & 2w & = & 0 \end{array}
\qquad [A|\mathbf{b}] = \begin{bmatrix} 1 & 1 & 3 & 1 & 2 \\ 1 & -1 & 1 & 1 & 4 \\ 0 & 1 & 2 & 2 & 0 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & 0 & -3 & 1 \\ 0 & 1 & 0 & -2 & -2 \\ 0 & 0 & 1 & 2 & 1 \end{bmatrix}
\qquad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ 2 \\ -2 \\ 1 \end{bmatrix}, t \in \mathbb{R}$$

$$A\mathbf{x} = \mathbf{0}$$

$$RREF(A)$$

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

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = t \begin{bmatrix} 3 \\ 2 \\ -2 \\ 1 \end{bmatrix}, t \in \mathbb{R}$$

$$A\mathbf{x} = \mathbf{b}$$

$$RREF([A|\mathbf{b}])$$

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

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ 2 \\ -2 \\ 1 \end{bmatrix}, t \in \mathbb{R}$$

Definition (Associated homogenous system)

Given a system of linear equations, $A\mathbf{x} = \mathbf{b}$, the linear system $A\mathbf{x} = \mathbf{0}$ is called the associated homogeneous system

Eg:

$$RREF(A) = \begin{bmatrix} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

How can you tell from here that $A\mathbf{x} = \mathbf{b}$ is consistent with infinitely many solutions?

Definition (Null space)

For an $m \times n$ matrix A, the null space of A is the subset of $\mathbb{R}^n$ given by

$$N(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

where $\mathbf{0} = (0, 0, \ldots, 0)^T$ is the zero vector of $\mathbb{R}^n$

Theorem (Principle of Linearity)

Suppose that A is an $m \times n$ matrix, that $\mathbf{b} \in \mathbb{R}^m$ and that the system $A\mathbf{x} = \mathbf{b}$ is consistent. Suppose that $\mathbf{p}$ is any solution of $A\mathbf{x} = \mathbf{b}$. Then the set of all solutions of $A\mathbf{x} = \mathbf{b}$ consists precisely of the vectors $\mathbf{p} + \mathbf{z}$ for $\mathbf{z} \in N(A)$; ie,

$$\{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}\} = \{\mathbf{p} + \mathbf{z} \mid \mathbf{z} \in N(A)\}.$$

Proof: We show that

1. $\mathbf{p} + \mathbf{z}$ is a solution for any $\mathbf{z}$ in the null space of A $(\{\mathbf{p} + \mathbf{z} \mid \mathbf{z} \in N(A)\} \subseteq \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}\})$
2. all solutions, $\mathbf{x}$, of $A\mathbf{x} = \mathbf{b}$ are of the form $\mathbf{p} + \mathbf{z}$ for some $\mathbf{z} \in N(A)$ $(\{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}\} \subseteq \{\mathbf{p} + \mathbf{z} \mid \mathbf{z} \in N(A)\})$

1. $A(\mathbf{p} + \mathbf{z}) = A\mathbf{p} + A\mathbf{z} = \mathbf{b} + \mathbf{0} = \mathbf{b}$ so $\mathbf{p} + \mathbf{z} \in \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}\}$

2. Let x be a solution. Because $\mathbf{p}$ is also we have $A\mathbf{p} = \mathbf{b}$ and$A(\mathbf{x} - \mathbf{p}) = A\mathbf{x} - A\mathbf{p} = \mathbf{b} - \mathbf{b} = \mathbf{0}$ so $\mathbf{z} = \mathbf{x} - \mathbf{p}$ is a solution of $A\mathbf{z} = \mathbf{0}$ and $\mathbf{x} = \mathbf{p} + \mathbf{z}$

(Check validity of the theorem on the previous examples.)

Summary

- If $A\mathbf{x} = \mathbf{b}$ is consistent, the solutions are of the form:

$$\{\text{solutions of } A\mathbf{x} = \mathbf{b}\} = \mathbf{p} + \{\text{solutions of } A\mathbf{x} = \mathbf{0}\}$$

  - if $A\mathbf{x} = \mathbf{b}$ has a unique solution, then $A\mathbf{x} = \mathbf{0}$ has only the trivial solution
  - if $A\mathbf{x} = \mathbf{b}$ has a infinitely many solutions, then $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions

- $A\mathbf{x} = \mathbf{b}$ may be inconsistent, but $A\mathbf{x} = \mathbf{0}$ is always consistent.

DM559

Linear and Integer Programming

Lecture 5
Matrix Inverse and Determinants

Marco Chiarandini

Department of Mathematics & Computer Science
University of Southern Denmark

Outline

Outline

1. Elementary Matrices

2. Matrix Inverse

3. Determinants

4. Matrix Inverse and Cramer's rule

Row Operations Revisited

Let's examine the process of applying the elementary row operations:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \overrightarrow{\mathbf{a}}_1 \\ \overrightarrow{\mathbf{a}}_2 \\ \vdots \\ \overrightarrow{\mathbf{a}}_m \end{bmatrix}$$

($\overrightarrow{\mathbf{a}}_i$ row ith of matrix A)
Then the three operations can be described as:

$$\begin{bmatrix} \overrightarrow{\mathbf{a}}_1 \\ \lambda\overrightarrow{\mathbf{a}}_2 \\ \vdots \\ \overrightarrow{\mathbf{a}}_m \end{bmatrix} \qquad \begin{bmatrix} \overrightarrow{\mathbf{a}}_2 \\ \overrightarrow{\mathbf{a}}_1 \\ \vdots \\ \overrightarrow{\mathbf{a}}_m \end{bmatrix} \qquad \begin{bmatrix} \overrightarrow{\mathbf{a}}_1 \\ \overrightarrow{\mathbf{a}}_2 + \lambda\overrightarrow{\mathbf{a}}_1 \\ \vdots \\ \overrightarrow{\mathbf{a}}_m \end{bmatrix}$$

For any $n \times n$ matrices A and B:

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{a}}_1 B \\ \vec{\mathbf{a}}_2 B \\ \vdots \\ \vec{\mathbf{a}}_n B \end{bmatrix}$$

$$\begin{bmatrix} \vec{\mathbf{a}}_1 B \\ \vec{\mathbf{a}}_2 B + \lambda \vec{\mathbf{a}}_1 B \\ \vdots \\ \vec{\mathbf{a}}_n B \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{a}}_1 B \\ (\vec{\mathbf{a}}_2 + \lambda \vec{\mathbf{a}}_1) B \\ \vdots \\ \vec{\mathbf{a}}_n B \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{a}}_1 \\ \vec{\mathbf{a}}_2 + \lambda \vec{\mathbf{a}}_1 \\ \vdots \\ \vec{\mathbf{a}}_n \end{bmatrix} B$$

(matrix obtained by a row operation on AB)

$=$ (matrix obtained by a row operation on A)B

(matrix obtained by a row operation on B)

$=$ (matrix obtained by a row operation on I)B

Elementary matrix

Definition (Elementary matrix)

An elementary matrix, E, is an $n \times n$ matrix obtained by doing exactly one row operation on the $n \times n$ identity matrix, I.

Example:

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

$$B = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 6 \\ -1 & 0 & 1 \end{bmatrix} \xrightarrow{ii-i} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{ii-i} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = E_1$$

$$E_1 B = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 6 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Outline

1. Elementary Matrices
2. Matrix Inverse
3. Determinants
4. Matrix Inverse and Cramer's rule

Matrix Inverse

The three elementary row operations are trivially invertible.

Theorem

Any elementary matrix is invertible, and the inverse is also an elementary matrix

$$E_1 B = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 6 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$E_1^{-1}(E_1 B) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 6 \\ -1 & 0 & 1 \end{bmatrix}$$

Row equivalence

To be an equivalence relation a relation must satisfy three properties:

- reflexive: $A \sim A$
- symmetric: $A \sim B \implies B \sim A$
- transitive: $A \sim B$ and $B \sim C \implies A \sim C$

Definition (Row equivalence)

If two matrices A and B are $m \times n$ matrices, we say that A is row equivalent to B if and only if there is a sequence of elementary row operations to transform A to B.

Theorem

Every matrix is row equivalent to a matrix in reduced row echelon form

Invertible Matrices

Theorem

If A is an $n \times n$ matrix, then the following statements are equivalent:

1. A^{-1} *exists*
2. $A\mathbf{x} = \mathbf{b}$ *has a unique solution for any* $\mathbf{b} \in \mathbb{R}^n$
3. $A\mathbf{x} = \mathbf{0}$ *only has the trivial solution,* $\mathbf{x} = \mathbf{0}$
4. *The reduced row echelon form of A is I.*

Proof: $(1) \implies (2) \implies (3) \implies (4) \implies (1)$.

- $(1) \implies (2)$

  $$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \implies I\mathbf{x} = A^{-1}\mathbf{b} \implies \mathbf{x} = A^{-1}\mathbf{b}$$

  hence $\mathbf{x} = A^{-1}\mathbf{b}$ is the only possible solution and it is a solution indeed: $A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}, \quad \forall \mathbf{b}$

- $(2) \implies (3)$
  If $A\mathbf{x} = \mathbf{b}$ has a unique solution for all $\mathbf{b} \in \mathbb{R}^n$, then this is true for $\mathbf{b} = 0$. The unique solution of $A\mathbf{x} = \mathbf{0}$ must be the trivial solution, $\mathbf{x} = \mathbf{0}$

- (3) $\implies$ (4)
  then in the reduced row echelon form of A there are no non-leading (free) variables and there is a leading one in every column hence also a leading one in every row (because A is square and in RREF) hence it can only be the identity matrix

- (4) $\implies$ (1)
  $\exists$ sequence of row operations and elementary matrices $E_1, \ldots, E_r$ that reduce A to I ie,

  $$E_r E_{r-1} \cdots E_1 A = I$$

  Each elementary matrix has an inverse hence multiplying repeatedly on the left by E_r^{-1}, E_{r-1}^{-1}:

  $$A = E_1^{-1} \cdots E_{r-1}^{-1} E_r^{-1} I$$

  hence, A is a product of invertible matrices hence invertible.
  (Recall that $(AB)^{-1} = B^{-1}A^{-1}$)

Matrix Inverse via Row Operations

We saw that:

$$A = E_1^{-1} \cdots E_{r-1}^{-1} E_r^{-1} I$$

taking the inverse of both sides:

$$A^{-1} = (E_1^{-1} \cdots E_{r-1}^{-1} E_r^{-1})^{-1} = E_r \cdots E_1 = E_r \cdots E_1 I$$

Hence:

$$\text{if} \quad E_r E_{r-1} E \cdots E_1 A = I \quad \text{then} \quad A^{-1} = E_r E_{r-1} \cdots E_1 I$$

Method:

- Construct $[A \mid I]$
- Use row operations to reduce this to $[I \mid B]$
- If this is not possible then the matrix is not invertible
- If it is possible then $B = A^{-1}$

Example

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 6 \\ -1 & 0 & 1 \end{bmatrix} \to [A \mid I] = \left[\begin{array}{ccc|ccc} 1 & 2 & 4 & 1 & 0 & 0 \\ 1 & 3 & 6 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[iii+i]{ii-i} \left[\begin{array}{ccc|ccc} 1 & 2 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 2 & 5 & 1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{iii-2ii} \left[\begin{array}{ccc|ccc} 1 & 2 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1 & 3 & -2 & 1 \end{array}\right] \xrightarrow[ii-2iii]{i-4iii} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -11 & 8 & -4 \\ 0 & 1 & 0 & -7 & 5 & -2 \\ 0 & 0 & 1 & 3 & -2 & 1 \end{array}\right]$$

$$\xrightarrow{i-2ii} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -2 & 0 \\ 0 & 1 & 0 & -7 & 5 & -2 \\ 0 & 0 & 1 & 3 & -2 & 1 \end{array}\right]$$

$$A^{-1} = \begin{bmatrix} 3 & -2 & 0 \\ -7 & 5 & -2 \\ 3 & -2 & 1 \end{bmatrix}$$

Verify by checking $AA^{-1} = I$ and $A^{-1}A = I$.
What would happen if the matrix is not invertible?

Verifying an Inverse

Theorem

If A and B are $n \times n$ matrices and $AB = I$, then A and B are each invertible matrices, and $A = B^{-1}$ and $B = A^{-1}$.

Proof: show that $B\mathbf{x} = \mathbf{0}$ has unique solution $\mathbf{x} = 0$, then B is invertible.

$$B\mathbf{x} = \mathbf{0} \implies A(B\mathbf{x}) = A\mathbf{0} \implies (AB)\mathbf{x} = \mathbf{0} \overset{AB=I}{\implies} I\mathbf{x} = \mathbf{0} \implies \mathbf{x} = \mathbf{0}$$

So B^{-1} exists for the previous theorem. Hence:

$$AB = I \implies (AB)B^{-1} = IB^{-1} \implies A(BB^{-1}) = B^{-1} \implies A = B^{-1}$$

So A is the inverse of B, and therefore also invertible and

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

Outline

1. Elementary Matrices

2. Matrix Inverse

3. Determinants

4. Matrix Inverse and Cramer's rule

Determinants

- The determinant of a matrix A is a particular number associated with A, written $|A|$ or $\det(A)$, that tells whether the matrix A is invertible.

- For the 2×2 case:

$$[A \mid I] = \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right] \xrightarrow{(1/a)R_1} \left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ c & d & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_2 - cR_1} \left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & d - cb/a & -c/a & 1 \end{array}\right] \xrightarrow{aR_2} \left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & (ad - bc) & -c & a \end{array}\right]$$

Hence A^{-1} exists if and only if $ad - bc \neq 0$.

- hence, for a 2×2 matrix the determinant is

$$\left| \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

- The extension to $n \times n$ matrices is done recursively

Definition (Minor)

For an $n \times n$ matrix the (i, j) minor of A, denoted by M_{ij}, is the determinant of the $(n-1) \times (n-1)$ matrix obtained by removing the ith row and the jth column of A.

Definition (Cofactor)

The (i, j) cofactor of a matrix A is

$$C_{ij} = (-1)^{i+j} M_{ij}$$

Definition (Cofactor Expansion of $|A|$ by row one)

The determinant of an $n \times n$ matrix is given by

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$$

Example

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

$$\begin{aligned} |A| &= 1C_{11} + 2C_{12} + 3C_{13} \\ &= 1\begin{vmatrix} 1 & 1 \\ 3 & 0 \end{vmatrix} - 2\begin{vmatrix} 4 & 1 \\ -1 & 0 \end{vmatrix} + 3\begin{vmatrix} 4 & 1 \\ -1 & 3 \end{vmatrix} \\ &= 1(-3) - 2(1) + 3(13) = 34 \end{aligned}$$

Theorem

If A is an $n \times n$ matrix, then the determinant of A can be computed by multiplying the entries of any row (or column) by their cofactors and summing the resulting products:

$$\begin{aligned} |A| =& a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \\ & \quad \textit{(cofactor expansion by row } i\textit{)} \\ |A| =& a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \\ & \quad \textit{(cofactor expansion by column } j\textit{)} \end{aligned}$$

A mnemonic rule for the 3×3 matrix determinant: the rule of Sarrus

$$\begin{array}{ccc:cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

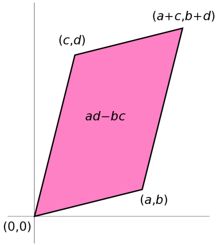
$$\begin{aligned} |A| = & + a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \end{aligned}$$

Verify the rule:

- from the conditions of existence of an inverse
- as a consequence of the general recursive rule for the determinants

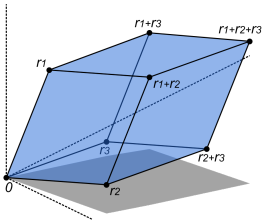
Geometric interpretation

2×2



The area of the parallelogram is the absolute value of the determinant of the matrix formed by the vectors representing the parallelogram's sides.

3×3



The volume of this parallelepiped is the absolute value of the determinant of the matrix formed by the rows constructed from the vectors $r1$, $r2$, and $r3$.

Properties of Determinants

Let A be an $n \times n$ matrix, then it follows from the previous theorem:

1. $|A^T| = |A|$
2. If a row of A consists entirely of zeros, then $|A| = 0$.
3. If A contains two rows which are equal, then $|A| = 0$.

$$|A| = \begin{vmatrix} a & b \\ a & b \end{vmatrix} = ab - ab = 0$$

$$|A| = \begin{vmatrix} a & b & c \\ d & e & f \\ a & b & c \end{vmatrix} = -d \begin{vmatrix} b & c \\ b & c \end{vmatrix} + e \begin{vmatrix} a & c \\ a & c \end{vmatrix} - f \begin{vmatrix} a & b \\ a & b \end{vmatrix} = 0 + 0 + 0$$

For 1. we can substitute row with column in 2., 3., 4.

4. If the cofactors of one row are multiplied by the entries of a different row and added, then the result is 0. That is, if $i \neq j$, then $a_{j1}C_{i1} + a_{j2}C_{i2} + \cdots + a_{jn}C_{in} = 0$.

$$A = \begin{array}{cc} \begin{bmatrix} \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix} & \begin{array}{c} \\ i\text{th} \\ \\ \\ \\ \end{array} \end{array} \qquad |A| = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

$$B = \begin{array}{cc} \begin{bmatrix} \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix} & \begin{array}{c} \\ i\text{th} \\ \\ \\ \\ \end{array} \end{array} \qquad |B| = a_{j1}C_{i1} + a_{j2}C_{i2} + \cdots + a_{jn}C_{in} = 0$$

5. If $A = (a_{ij})$ and if each entry of one of the rows, say row i, can be expressed as a sum of two numbers, $a_{ij} = b_{ij} + c_{ij}$ for $i \leq j \leq n$, then $|A| = |B| + |C|$, where B is the matrix A with row i replaced by $b_{i1}, b_{i2}, \cdots, b_{in}$ and C is the matrix A with row i replaced by $c_{i1}, c_{i2}, \cdots, c_{in}$.

$$|A| = \begin{vmatrix} a & b & c \\ d+p & e+q & f+r \\ g & h & i \end{vmatrix} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} + \begin{vmatrix} a & b & c \\ p & q & r \\ g & h & i \end{vmatrix} = |B| + |C|$$

Triangular Matrices

Definition (Triangular Matrices)

An $n \times n$ matrix is said to be upper triangular if $a_{ij} = 0$ for $i > j$ and lower triangular if $a_{ij} = 0$ for $i < j$. Also A is said to be triangular if it is either upper triangular or lower triangular.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \qquad \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Definition (Diagonal Matrices)

An $n \times n$ matrix is diagonal if $a_{ij} = 0$ whenever $i \neq j$.

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Determinant using row operations

- Which row or column would you choose for the cofactor expansion in this case:

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} =? = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}$$

- if A is upper/lower triangular or diagonal, then $|A| = a_{11}a_{22}\cdots a_{nn}$

- Idea: a square matrix in REF is upper triangular. What is the effect of row operations on the determinant?

RO1 multiply a row by a non-zero constant

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \quad |B| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$|B| = \alpha a_{i1} C_{i1} + \alpha a_{i2} C_{i2} + \cdots + \alpha a_{in} C_{in} = \alpha |A|$$

$\rightsquigarrow$ $|A|$ changes to $\alpha|A|$

RO2 interchange two rows

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb \qquad |B| = \begin{vmatrix} c & d \\ a & b \end{vmatrix} = cb - ad \implies |B| = -|A|$$

$$|A| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \qquad |B| = \begin{vmatrix} g & h & i \\ d & e & f \\ a & b & c \end{vmatrix} \implies |B| = -|A|$$

$\rightsquigarrow$ $|A|$ changes to $-|A|$ (by induction)

RO3 add a multiple of one row to another

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \quad |B| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} + 4a_{11} & a_{22} + 4a_{12} & \cdots & a_{2n} + 4a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$\begin{aligned} |B| =& (a_{j1} + \lambda a_{i1})C_{j1} + (a_{j2} + \lambda a_{i2})C_{j2} + \cdots + (a_{jn} + \lambda a_{in})C_{jn} \\ =& a_{j1}C_{j1} + a_{j2}C_{j2} + \cdots + a_{jn}C_{jn} + \lambda(a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn}) \\ =& |A| + 0 \end{aligned}$$

$\rightsquigarrow$ there is no change in $|A|$

Example

$$|A| = \begin{vmatrix} 1 & 2 & -1 & 4 \\ -1 & 3 & 0 & 2 \\ 2 & 1 & 1 & 2 \\ 1 & 4 & 1 & 3 \end{vmatrix} \overset{RO3s}{=} \begin{vmatrix} 1 & 2 & -1 & 4 \\ 0 & 5 & -1 & 6 \\ 0 & -3 & 3 & -6 \\ 0 & 2 & 2 & -1 \end{vmatrix} \overset{\alpha R_3}{=} -3 \begin{vmatrix} 1 & 2 & -1 & 4 \\ 0 & 5 & -1 & 6 \\ 0 & 1 & -1 & 2 \\ 0 & 2 & 2 & -1 \end{vmatrix}$$

$$\overset{RO2}{=} 3 \begin{vmatrix} 1 & 2 & -1 & 4 \\ 0 & 1 & -1 & 2 \\ 0 & 5 & -1 & 6 \\ 0 & 2 & 2 & -1 \end{vmatrix} \overset{RO3s}{=} 3 \begin{vmatrix} 1 & 2 & -1 & 4 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 4 & -4 \\ 0 & 0 & 4 & -5 \end{vmatrix} \overset{RO3s}{=} 3 \begin{vmatrix} 1 & 2 & -1 & 4 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 4 & -4 \\ 0 & 0 & 4 & -5 \end{vmatrix}$$

$$\overset{RO3s}{=} 3 \begin{vmatrix} 1 & 2 & -1 & 4 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 4 & -4 \\ 0 & 0 & 0 & -1 \end{vmatrix} = 3(1 \times 1 \times 4 \times (-1)) = -12$$

Determinant of a Product

Theorem

If A and B are $n \times n$ matrices, then $|AB| = |A||B|$

Proof:

- Let E_1 be an elementary matrix that multiplies a row by a non-zero constant k
- $|E_1| = |E_1 I| = k|I| = k$ and $|E_1 B| = k|B| = |E_1||B|$
- similarly: $|E_2 B| = -|B| = |E_2||B|$ and $|E_3 B| = |B| = |E_3||B|$
- by row equivalence we have

  $$A = E_r E_{r-1} \cdots E_1 R$$

  where R is in RREF. Since A is square, R is either I or has a row of zeros.
- $|A| = |E_r E_{r-1} \cdots E_1 R| = |E_r||E_{r-1}| \cdots |E_1||R|$ and $|E_i| \neq 0$
- If $R = I$:

  $$\begin{aligned} |AB| &= |(E_r E_{r-1} \cdots E_1 I)B| = |E_r E_{r-1} \cdots E_1 B| \\ &= |E_r||E_{r-1}| \cdots |E_1||B| = |E_r E_{r-1} \cdots E_1||B| = |A||B| \end{aligned}$$

- If $R \neq I$ then $|AB| = |E_r \ldots E_1 RB| = |E_r| \ldots |E_1||RB|$ and $|AB| = 0$

Matrix Inverse using Cofactors

Theorem

If A is an $n \times n$ matrix, then A is invertible if and only if $|A| \neq 0$.

Proof:

- (implied already by the first theorem of today: by (4) either R is I or it has a row of zeros.)

$\Rightarrow$ If A is invertible then $|AA^{-1}| = |A||A^{-1}| = |I|$. Hence $|A| \neq 0$. We get also that: and

$$|A^{-1}| = \frac{1}{|A|}$$

$\Leftarrow$ if $|A| \neq 0$ then A is invertible: we show this by construction:

Definition (Adjoint)

If A is an $n \times n$ matrix, the matrix of cofactors of A is the matrix whose (i, j) entry is C_{ij}, the (i, j) cofactor of A.
The adjoint or (adjugate) of A is the transpose of the matrix of cofactors, ie:

$$\mathrm{adj(A)} = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{bmatrix}^T = \begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix}$$

- $$A\,\mathrm{adj(A)} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix}$$

- entry $(1,1)$ is $a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$, ie, cofactor by row 1
  entry $(1,2)$ is $a_{11}C_{21} + a_{12}C_{22} + \cdots + a_{1n}C_{2n}$, ie, entries of row 1 multiplied by cofactors of row 2

$$A\,\mathrm{adj(A)} = \begin{bmatrix} |A| & 0 & \dots & 0 \\ 0 & |A| & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |A| \end{bmatrix} = |A|I$$

- Since $|A| \neq 0$ we can divide:

$$A\left(\frac{1}{|A|}\mathrm{adj}(A)\right) = I \qquad A^{-1} = \frac{1}{|A|}\mathrm{adj}(A)$$

$\square$

Outline

1. Elementary Matrices

2. Matrix Inverse

3. Determinants

4. Matrix Inverse and Cramer's rule

Matrix Inverse using Cofactors

Example

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

What is A^{-1}?

- $|A| = 1(2-1) - 2(-1-4) + 3(-1-8) = -16 \neq 0 \implies$ invertible
- Matrix of cofactors

$$\begin{bmatrix} +M_{11} & -M_{12} & +M_{13} & -M_{14} & \cdots \\ -M_{21} & +M_{22} & -M_{23} & +M_{24} & \cdots \\ +M_{31} & -M_{32} & +M_{33} & -M_{34} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \to \begin{bmatrix} 1 & 5 & -9 \\ 1 & -11 & 7 \\ -4 & 4 & 4 \end{bmatrix}$$

- $$A^{-1} = \frac{1}{|A|}\operatorname{adj}(A) = -\frac{1}{16}\begin{bmatrix} 1 & 5 & -9 \\ 1 & -11 & 7 \\ -4 & 4 & 4 \end{bmatrix}^T = -\frac{1}{16}\begin{bmatrix} 1 & 1 & -4 \\ 5 & -11 & 4 \\ -9 & 7 & 4 \end{bmatrix}$$

Matrix Inverse using Cofactors

Example (cntd)

- Verify $AA^{-1} = I$:

$$-\frac{1}{16}\begin{bmatrix} 1 & 2 & 3 \\ -1 & 2 & 1 \\ 4 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & -4 \\ 5 & -11 & 4 \\ -9 & 7 & 4 \end{bmatrix} = -\frac{1}{16}\begin{bmatrix} -16 & 0 & 0 \\ 0 & -16 & 0 \\ 0 & 0 & -16 \end{bmatrix} = I$$

Cramer's rule

Theorem (Cramer's rule)

If A is $n \times n$, $|A| \neq 0$, and $\mathbf{b} \in \mathbb{R}^n$, then the solution $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ of the linear system $A\mathbf{x} = \mathbf{b}$ is given by

$$x_i = \frac{|A_i|}{|A|},$$

where A_i is the matrix obtained from A by replacing the ith column with the vector $\mathbf{b}$.

Proof: Since $|A| \neq 0$, A^{-1} exists and we can solve for $\mathbf{x}$ by multiplying $A\mathbf{x} = \mathbf{b}$ on the left by A^{-1}. The $\mathbf{x} = A^{-1}\mathbf{b}$:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{|A|} \begin{bmatrix} C_{11} & C_{21} & \ldots & C_{n1} \\ C_{12} & C_{22} & \ldots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \ldots & C_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$\implies x_i = \frac{1}{|A|}(b_1 C_{1i} + b_2 C_{2i} + \cdots + b_n C_{ni})$, ie, cofactor expansion of column i of A with column i replaced by $\mathbf{b}$, ie, $|A_i|$ $\square$

Matrix Inverse using Cofactors

Example

Use Cramer's rule to solve:

$$\begin{array}{rcrcrcr} x & + & 2y & + & 3z & = & 7 \\ -x & + & 2y & + & z & = & -3 \\ 4x & + & y & + & z & = & 5 \end{array}$$

- In matrix form:

$$\begin{bmatrix} 1 & 2 & 3 \\ -1 & 2 & 1 \\ 4 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7 \\ -3 \\ 5 \end{bmatrix}$$

- $|A| = -16 \neq 0$
- $$x = \frac{\begin{vmatrix} 7 & 2 & 3 \\ -3 & 2 & 1 \\ 5 & 1 & 1 \end{vmatrix}}{|A|} = 1, \quad y = \frac{\begin{vmatrix} 1 & 7 & 3 \\ -1 & -3 & 1 \\ 4 & 5 & 1 \end{vmatrix}}{|A|} = -3, \quad z = \frac{\begin{vmatrix} 1 & 2 & 7 \\ -1 & 2 & -3 \\ 4 & 1 & 5 \end{vmatrix}}{|A|} = 4$$

Summary (1/2)

- There are three methods to solve $A\mathbf{x} = \mathbf{b}$ if A is $n \times n$ and $|A| \neq 0$:
  1. Gaussian elimination
  2. Matrix solution: find A^{-1}, then calculate $\mathbf{x} = A^{-1}\mathbf{b}$
  3. Cramer's rule
- There is one method to solve $A\mathbf{x} = \mathbf{b}$ if A is $m \times n$ and $m \neq n$ or if $|A| = 0$:
  1. Gaussian elimination
- There are two methods to find A^{-1}:
  1. using cofactors for the adjoint matrix
  2. by row reduction of $[A \mid I]$ to $[I \mid A^{-1}]$

Summary (2/2)

- If A is an $n \times n$ matrix, then the following statements are equivalent:
  1. A is invertible
  2. $A\mathbf{x} = \mathbf{b}$ has a unique solution for any $\mathbf{b} \in \mathbb{R}$
  3. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, $\mathbf{x} = \mathbf{0}$
  4. the reduced row echelon form of A is I.
  5. $|A| \neq 0$

- Solving $A\mathbf{x} = \mathbf{b}$ in practice and at the computer:
  - via LU factorization (much quicker if one has to solve several systems with the same matrix A but different vectors b)
  - if A is symmetric positive definite matrix then Cholesky decomposition (twice as fast)
  - if A is large or sparse then iterative methods

DM559

Linear and Integer Programming

Lecture 6
Rank and Range
Vector Spaces

Marco Chiarandini

Department of Mathematics & Computer Science
University of Southern Denmark

Outline

1. Rank

2. Range

3. Vector Spaces

Outline

Rank

- Synthesis of what we have seen so far under the light of two new concepts: rank and range of a matrix
- We saw that:
  every matrix is row-equivalent to a matrix in reduced row echelon form.

Definition (Rank of Matrix)

The rank of a matrix A, $\text{rank}(A)$, is

- the number of non-zero rows, or equivalently
- the number of leading ones

in a row echelon matrix obtained from A by elementary row operations.

$\rightsquigarrow$ For an $m \times n$ matrix A,

$$\text{rank}\, A \leq \min\{m, n\},$$

where $\min\{m, n\}$ denotes the smaller of the two integers m and n.

Example

$$M = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 3 & 0 & 5 \\ 3 & 5 & 1 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 3 & 0 & 5 \\ 3 & 5 & 1 & 6 \end{bmatrix} \xrightarrow[\;]{\substack{R_2'=R_2-2R_1 \\ R_3'=R_3-3R_1}} \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & -1 & -2 & 3 \\ 0 & -1 & -2 & 3 \end{bmatrix} \xrightarrow[\;]{\substack{R_2'=-R_2 \\ R_3'=R_3-R_2}} \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$\rightsquigarrow \operatorname{rank}(M) = 2$

Extension of the main theorem

Theorem

If A is an $n \times n$ matrix, then the following statements are equivalent:

1. *A is invertible*
2. *$A\mathbf{x} = \mathbf{b}$ has a unique solution for any $\mathbf{b} \in \mathbb{R}$*
3. *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution, $\mathbf{x} = \mathbf{0}$*
4. *the reduced row echelon form of A is I.*
5. *$|A| \neq 0$*
6. *the rank of A is n*

Rank and Systems of Linear Equations

$$\begin{array}{rcrcrcr} x & + & 2y & + & z & = & 1 \\ 2x & + & 3y & & & = & 5 \\ 3x & + & 5y & + & z & = & 4 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 3 & 0 & 5 \\ 3 & 5 & 1 & 4 \end{array}\right] \xrightarrow[R_3'=R_3-3R_1]{R_2'=R_2-2R_1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & -1 & -2 & 3 \\ 0 & -1 & -2 & 1 \end{array}\right] \xrightarrow[R_3'=R_3-R_2]{R_2'=-R_2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & -2 \end{array}\right]$$

$$\begin{array}{rcrcrcr} x & + & 2y & + & z & = & 1 \\ & & x & + & 2z & = & -3 \\ 0x & + & 0y & + & 0z & = & -2 \end{array}$$

It is inconsistent!

The last row is of the type $0 = a, a \neq 0$, that is, the augmenting matrix has a leading one in the last column

$\text{rank}(A) = 2 \neq \text{rank}(A \mid \mathbf{b}) = 3$

1. A system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if the rank of the augmented matrix is precisely the same as the rank of the matrix A.

2. If an $m \times n$ matrix A has rank m, the system of linear equations, $A\mathbf{x} = \mathbf{b}$, will be consistent for all $\mathbf{b} \in \mathbb{R}^n$

- Since A has rank m then there is a leading one in every row. Hence $[A \mid \mathbf{b}]$ cannot have a row $[0, 0, \ldots, 0, 1] \implies \operatorname{rank} A \not< \operatorname{rank}(A \mid \mathbf{b})$
- $[A \mid \mathbf{b}]$ has also m rows $\implies \operatorname{rank}(A) \not> \operatorname{rank}(A \mid \mathbf{b})$
- Hence, $\operatorname{rank}(A) = \operatorname{rank}(A \mid \mathbf{b})$

Example

$$B = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 3 & 0 & 5 \\ 3 & 5 & 1 & 4 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & -3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \operatorname{rank}(B) = 3$$

Any system $B\mathbf{x} = \mathbf{d}$ in 4 unknowns and 3 equalities with $\mathbf{d} \in \mathbb{R}^3$ is consistent.

Since $\operatorname{rank}(A)$ is smaller than the number of variables, then there is a non-leading variable. Hence infinitely many solutions!

Example

$$[A|\mathbf{b}] = \begin{bmatrix} 1 & 3 & -2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 3 & 0 & 4 & 0 & -28 \\ 0 & 0 & 1 & 2 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\operatorname{rank}([A|\mathbf{b}]) = 3 < 5 = n$

$$\begin{array}{ccccccccc} x_1 & + & 3x_2 & & & + & 4x_4 & & = & -28 \\ & & & & x_3 & + & 2x_4 & & = & -14 \\ & & & & & & & x_5 & = & 5 \end{array}$$

x_1, x_3, x_5 are leading variables; x_2, x_4 are non-leading variables (set them to $s, t \in \mathbb{R}$)

$$\begin{aligned} x_1 &= -28 - 3s - 4t \\ x_2 &= s \\ x_3 &= -14 - 2t \\ x_4 &= t \\ x_5 &= 5 \end{aligned} \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -28 \\ 0 \\ -14 \\ 0 \\ 5 \end{bmatrix} + \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} s + \begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} t$$

Summary

Let $A\mathbf{x} = \mathbf{b}$ be a general linear system in n variables and m equations:

- If $\text{rank}(A) = r < m$ and $\text{rank}(A \mid b) = r + 1$ then the system is inconsistent. (the row echelon form of the augmented matrix has a row $[0\ 0\ \ldots\ 0\ 1]$)

- If $\text{rank}(A) = r = \text{rank}(A \mid \mathbf{b})$ then the system is consistent and there are $n - r$ free variables;
  if $r < n$ there are infinitely many solutions, if $r = n$ there are no free variables and the solution is unique

Let $A\mathbf{x} = \mathbf{0}$ be an homogeneous system in n variables and m equations, $\text{rank}(A) = r$ (always consistent):

- if $r < n$ there are infinitely many solutions, if $r = n$ there are no free variables and the solution is unique, $\mathbf{x} = \mathbf{0}$.

General solutions in vector notation

Example

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_2 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -28 \\ 0 \\ -14 \\ 0 \\ 5 \end{bmatrix} + \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} s + \begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} t, \quad \forall s, t \in \mathbb{R}$$

For $A\mathbf{x} = \mathbf{b}$:

$$\mathbf{x} = \mathbf{p} + \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_{n-r} \mathbf{v}_{n-r}, \quad \forall \alpha_i \in \mathbb{R}, i = 1, \ldots, n-r$$

Note:

- if $\alpha_i = 0, \forall i = 1, \ldots, n-r$ then $A\mathbf{p} = \mathbf{b}$, ie, $\mathbf{p}$ is a particular solution
- if $\alpha_1 = 1$ and $\alpha_i = 0, \forall i = 2, \ldots, n-r$ then

$$A(\mathbf{p} + \mathbf{v}_1) = \mathbf{b} \;\rightarrow\; A\mathbf{p} + A\mathbf{v}_1 = \mathbf{b} \;\xrightarrow{A\mathbf{p}=\mathbf{b}}\; A\mathbf{v}_1 = 0$$

Thus (recall that $\mathbf{x} = \mathbf{p} + \mathbf{z}$, $\mathbf{z} \in N(A)$):

- If A is an $m \times n$ matrix of rank r, the general solutions of $A\mathbf{x} = \mathbf{b}$ is the sum of:
    - a particular solution $\mathbf{p}$ of the system $A\mathbf{x} = \mathbf{b}$ and
    - a linear combination $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_{n-r}\mathbf{v}_{n-r}$ of solutions $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{n-r}$ of the homogeneous system $A\mathbf{x} = \mathbf{0}$
- If A has rank n, then $A\mathbf{x} = \mathbf{0}$ only has the solution $\mathbf{x} = \mathbf{0}$ and so $A\mathbf{x} = \mathbf{b}$ has a unique solution: $\mathbf{p}$

Outline

Range

Definition (Range of a matrix)

Let A be an $m \times n$ matrix, the range of A, denoted by $R(A)$, is the subset of $\mathbb{R}^m$ given by

$$R(A) = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}$$

That is, the range is the set of all vectors $\mathbf{y} \in \mathbb{R}^m$ of the form $\mathbf{y} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$, or
all $\mathbf{y} \in \mathbb{R}^m$ for which the system $A\mathbf{x} = \mathbf{y}$ is consistent.

Recall, if $\mathbf{x} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$ is any vector in $\mathbb{R}^n$ and

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \mathbf{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}, \qquad i = 1, \ldots, n.$$

Then $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ and

$$A\mathbf{x} = \alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \ldots + \alpha_n \mathbf{a}_n$$

that is, vector $A\mathbf{x}$ in $\mathbb{R}^n$ as a linear combination of the column vectors of A
Proof?

Hence $R(A)$ is the set of all linear combinations of the columns of A.
$\rightsquigarrow$ the range is also called the column space of A:

$$R(A) = \{\alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \ldots + \alpha_n \mathbf{a}_n \mid \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}\}$$

Thus, $A\mathbf{x} = \mathbf{b}$ is consistent iff $\mathbf{b}$ is in the range of A, ie, a linear combination of the columns of A

Example

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 2 & 1 \end{bmatrix}$$

Then, for $\mathbf{x} = [\alpha_1, \alpha_2]^T$

$$A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \alpha_1 + 2\alpha_2 \\ -\alpha_1 + 3\alpha_2 \\ 2\alpha_1 + \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \alpha_1 + \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \alpha_2$$

so

$$R(A) = \left\{ \begin{bmatrix} \alpha_1 + 2\alpha_2 \\ -\alpha_1 + 3\alpha_2 \\ 2\alpha_1 + \alpha_2 \end{bmatrix} \middle| \; \alpha_1, \alpha_2 \in \mathbb{R} \right\}$$

Example

$$\begin{cases} \quad x + 2y = 0 \\ -x + 3y = -5 \\ 2x + y = 3 \end{cases}$$

$$\begin{cases} \quad x + 2y = 1 \\ -x + 3y = -5 \\ 2x + y = 2 \end{cases}$$

$$A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -5 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ -5 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = 2\mathbf{a}_1 - \mathbf{a}_2$$

$$A\mathbf{x} = \mathbf{0}$$

has only the trivial solution $\mathbf{x} = \mathbf{0}$. (Why?) Only way:

$$0 \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = 0\mathbf{a}_1 + 0\mathbf{a}_2 = \mathbf{0}$$

Hence no way to express $[1, -5, 2]$ as linear expression of the two columns of A.

Outline

Premise

- We move to a higher level of abstraction
- A vector space is a set with an addition and scalar multiplication that behave appropriately, that is, like $\mathbb{R}^n$
- Imagine a vector space as a class of a generic type (template) in object oriented programming, equipped with two operations.

Vector Spaces

Definition (Vector Space)

A (real) vector space V is a non-empty set equipped with an addition and a scalar multiplication operation such that for all $\alpha, \beta \in \mathbb{R}$ and all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$:

1. $\mathbf{u} + \mathbf{v} \in V$ (closure under addition)
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (commutative law for addition)
3. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ (associative law for addition)
4. there is a single member $\mathbf{0}$ of V, called the zero vector, such that for all $\mathbf{v} \in V, \mathbf{v} + \mathbf{0} = \mathbf{v}$
5. for every $\mathbf{v} \in V$ there is an element $\mathbf{w} \in V$, written $-\mathbf{v}$, called the negative of $\mathbf{v}$, such that $\mathbf{v} + \mathbf{w} = \mathbf{0}$
6. $\alpha\mathbf{v} \in V$ (closure under scalar multiplication)
7. $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$ (distributive law)
8. $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$ (distributive law)
9. $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$ (associative law for vector multiplication)
10. $1\mathbf{v} = \mathbf{v}$

Examples

- set $\mathbb{R}^n$

- but the set of objects for which the vector space defined is valid are more than the vectors in $\mathbb{R}^n$.

- set of all functions $F : \mathbb{R} \to \mathbb{R}$.
  We can define an addition $f + g$:

  $$(f + g)(x) = f(x) + g(x)$$

  and a scalar multiplication αf:

  $$(\alpha f)(x) = \alpha f(x)$$

- Example: $x + x^2$ and $2x$. They can represent the result of the two operations.

- What is $-f$? and the zero vector?

The axioms given are minimum number needed.
Other properties can be derived:
For example:

$$(-1)\mathbf{x} = -\mathbf{x}$$

$$\mathbf{0} = 0\mathbf{x} = (1 + (-1))\mathbf{x} = 1\mathbf{x} + (-1)\mathbf{x} = \mathbf{x} + (-1)\mathbf{x}$$

Adding $-\mathbf{x}$ on both sides:

$$-\mathbf{x} = -\mathbf{x} - \mathbf{0} = -\mathbf{x} + \mathbf{x} + (-1)\mathbf{x} = (-1)\mathbf{x}$$

which proves that $-\mathbf{x} = (-1)\mathbf{x}$.

Try the same with $-f$.

Examples

- $V = \{\mathbf{0}\}$

- the set of $m \times n$ all matrices

- the set of all infinite sequences of real numbers, $\mathbf{y} = \{y_1, y_2, \ldots, y_n, \ldots, \}, y_i \in \mathbb{R}$. ($\mathbf{y} = \{y_n\}$, $n \geq 1$)

  addition of $\mathbf{y} = \{y_1, y_2, \ldots, y_n, \ldots, \}$ and $\mathbf{z} = \{z_1, z_2, \ldots, z_n, \ldots, \}$ then:

  $$\mathbf{y} + \mathbf{z} = \{y_1 + z_1, y_2 + z_2, \ldots, y_n + z_n, \ldots, \}$$

  multiplication by a scalar $\alpha \in \mathbb{R}$:

  $$\alpha\mathbf{y} = \{\alpha y_1, \alpha y_2, \ldots, \alpha y_n, \ldots, \}$$

- set of all vectors in $\mathbb{R}^3$ with the third entry equal to 0 (verify closure):

  $$W = \left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \middle| \; x, y \in \mathbb{R} \right\}$$

Linear Combinations

Definition (Linear Combination)

For vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in a vector space V, the vector

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_k\mathbf{v}_k$$

is called a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$.
The scalars α_i are called coefficients.

- To find the coefficients that given a set of vertices express by linear combination a given vector, we solve a system of linear equations.
- If F is the vector space of functions from $\mathbb{R}$ to $\mathbb{R}$ then the function $f : x \mapsto 2x^2 + 3x + 4$ can be expressed as a linear combination of:

  $$f = 2g + 3h + 4k$$

  where $g : x \mapsto x^2$, $h : x \mapsto x$, $k : x \mapsto 1$
- Given two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$, is it possible to represent any point in the Cartesian plane?

Subspaces

Definition (Subspace)

A subspace W of a vector space V is a non-empty subset of V that is itself a vector space under the same operations of addition and scalar multiplication as V.

Theorem

Let V be a vector space. Then a non-empty subset W of V is a subspace if and only if both the following hold:

- *for all $\mathbf{u}, \mathbf{v} \in W$, $\mathbf{u} + \mathbf{v} \in W$*
  (W is closed under addition)
- *for all $\mathbf{v} \in W$ and $\alpha \in \mathbb{R}$, $\alpha\mathbf{v} \in W$*
  (W is closed under scalar multiplication)

ie, all other axioms can be derived to hold true

Example

- The set of all vectors in $\mathbb{R}^3$ with the third entry equal to 0.
- The set $\{\mathbf{0}\}$ is not empty, it is a subspace since $\mathbf{0}+\mathbf{0}=\mathbf{0}$ and $\alpha\mathbf{0}=\mathbf{0}$ for any $\alpha \in \mathbb{R}$.

Example

In $\mathbb{R}^2$, the lines $y = 2x$ and $y = 2x+1$ can be defined as the sets of vectors:

$$S = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \;\middle|\; y = 2x, x \in \mathbb{R} \right\} \qquad U = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \;\middle|\; y = 2x+1, x \in \mathbb{R} \right\}$$

$$S = \{\mathbf{x} \mid \mathbf{x} = t\mathbf{v}, t \in \mathbb{R}\} \qquad U = \{\mathbf{x} \mid \mathbf{x} = \mathbf{p} + t\mathbf{v}, t \in \mathbb{R}\}$$

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ \mathbf{p} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Example (cntd)

1. The set S is non-empty, since $\mathbf{0} = 0\mathbf{v} \in S$.
2. closure under addition:

$$\mathbf{u} = s \begin{bmatrix} 1 \\ 2 \end{bmatrix} \in S, \quad \mathbf{w} = t \begin{bmatrix} 1 \\ 2 \end{bmatrix} \in S, \qquad \text{for some } s, t \in \mathbb{R}$$

$\mathbf{u} + \mathbf{w} = s\mathbf{v} + t\mathbf{v} = (s + t)\mathbf{v} \in S$ since $s + t \in \mathbb{R}$

3. closure under scalar multiplication:

$$\mathbf{u} = s \begin{bmatrix} 1 \\ 2 \end{bmatrix} \in S \qquad \text{for some } s \in \mathbb{R}, \qquad \alpha \in \mathbb{R}$$

$\alpha\mathbf{u} = \alpha(s(\mathbf{v})) = (\alpha s)\mathbf{v} \in S$ since $\alpha s \in \mathbb{R}$

Note that:

- $\mathbf{u}, \mathbf{w}$ and $\alpha \in \mathbb{R}$ must be arbitrary

Example (cntd)

1. $\mathbf{0} \notin U$
2. U is not closed under addition:

$$\begin{bmatrix}0\\1\end{bmatrix} \in U, \begin{bmatrix}1\\3\end{bmatrix} \in U \quad \text{but} \quad \begin{bmatrix}0\\1\end{bmatrix} + \begin{bmatrix}1\\3\end{bmatrix} = \begin{bmatrix}1\\4\end{bmatrix} \notin U$$

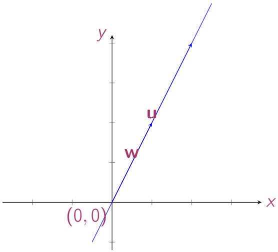
3. U is not closed under scalar multiplication

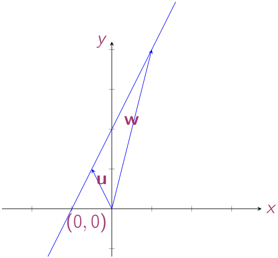
$$\begin{bmatrix}0\\1\end{bmatrix} \in U, 2 \in \mathbb{R} \quad \text{but} \quad 2\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\2\end{bmatrix} \notin U$$

Note that:

- proving just one of the above couterexamples is enough to show that U is not a subspace
- it is sufficient to make them fail for particular choices
- a good place to start is checking whether $\mathbf{0} \in S$. If not then S is not a subspace

Geometric interpretation:





$\rightsquigarrow$ The line $y = 2x + 1$ is an affine subset, a „translation" of a subspace

Theorem

A non-empty subset W of a vector space is a subspace if and only if for all $\mathbf{u}, \mathbf{v} \in W$ and all $\alpha, \beta \in \mathbb{R}$, we have $\alpha\mathbf{u} + \beta\mathbf{v} \in W$.
That is, W is closed under linear combination.

Summary

- Rank of a matrix and relation to number of solutions of a linear system
- General solutions of a linear system in vector notation
- Range, set of linear combinations of the columns of a matrix
- Vector spaces: properties
- Linear combination
- Subspaces: non-empty + closed under linear combination

DM559

Linear and Integer Programming

Lecture 7
Vector Spaces (cntd)
Linear Independence, Bases and Dimension

Marco Chiarandini

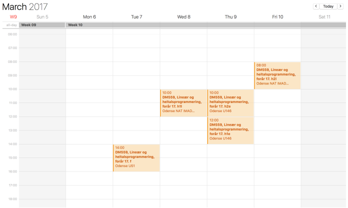
Department of Mathematics & Computer Science
University of Southern Denmark

Outlook

Week 9:

Section H1:	Section H2:
Today, Introductory class	Today, Introductory class
Tomorrow, 14-16, Applications	Tomorrow, 12-14, Introductory class
Thursday, 12-14, Introductory class	**Tomorrow, 14-16, Applications**

Outlook



Week 10:

Section H1:

Tuesday, 14-16, Exercise class
Wednesday, 10-12, Laboratory class
Thursday, 12-14, Exercise class

Section H2:

Tuesday, 14-16, Exercise class
Thursday, 10-12, Exercise class
Friday, 08-10, Laboratory class

a) Join? H1, Thursday, 12-14 $\iff$ H2, Thursday, 10-12
   1. Move H1 from Thursday, 12-14 to Thursday, 10-12?
   2. Move H2 from Thursday, 10-12 to Thursday, 12-14? ✔

b) Join? H1, Wednesday, 10-12 $\iff$ H2, Friday, 08-10
   1. Move H1 from Wednesday, 10-12 to Friday, 08-10?
   2. Move H2 from Friday, 08-10 to Wednesday, 10-12? ✔

Resume

- Matrix Calculus
- Geometric Insight
- Systems of Linear Equations. Gaussian Elimination.
- Elementary Matrices, Determinants, Matrix Inverse.

Last Time:

- Rank (number of leading ones in REF)
  Relationship with linear systems
- (Numerical methods, LU + iterative)
- Range of a matrix
- Vector Spaces: Definition, Examples. Linear combination.
- Subspaces

Outline

Outline

1. Vector Spaces (cntd)

2. Linear independence

3. Bases

4. Dimension

Linear Span

- If $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_k\mathbf{v}_k$ and $\mathbf{w} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \ldots + \beta_k\mathbf{v}_k$, then $\mathbf{v} + \mathbf{w}$ and $s\mathbf{v}, s \in \mathbb{R}$ are also linear combinations of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$.
- The set of all linear combinations of a given set of vectors of a **vector space** V forms a **subspace**:

Definition (Linear span)

Let V be a **vector space** and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in V$. The linear span of $X = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is the set of all linear combinations of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, denoted by $\mathrm{Lin}(X)$, that is:

$$\mathrm{Lin}(\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}) = \{\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_k\mathbf{v}_k \mid \alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}\}$$

Theorem

*If $X = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a set of vectors of a vector space V, then $\mathrm{Lin}(X)$ is a **subspace** of V and is also called the **subspace** spanned by X.*
*It is the smallest **subspace** containing the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$.*

Example

- $\text{Lin}(\{\mathbf{v}\}) = \{\alpha\mathbf{v} \mid \alpha \in \mathbb{R}\}$ defines a line in $\mathbb{R}^n$.
- Recall that a plane in $\mathbb{R}^3$ has two equivalent representations:

$$ax + by + cz = d \qquad \text{and} \qquad \mathbf{x} = \mathbf{p} + s\mathbf{v} + t\mathbf{w}, \quad s, t \in \mathbb{R}$$

  where $\mathbf{v}$ and $\mathbf{w}$ are non parallel.

  - If $d = 0$ and $\mathbf{p} = \mathbf{0}$, then

$$\{\mathbf{x} \mid \mathbf{x} = s\mathbf{v} + t\mathbf{w}, s, t, \in \mathbb{R}\} = \text{Lin}(\{\mathbf{v}, \mathbf{w}\})$$

    and hence a **subspace** of $\mathbb{R}^n$.

  - If $d \neq 0$, then the plane is not a **subspace**. It is an affine subset, a translation of a **subspace**.

    (recall that one can also show directly that a subset is a **subspace** or not)

Outline

Linear Independence

Definition (Linear Independence)

Let V be a **vector space** and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in V$. Then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent (or form a linearly independent set) if and only if the vector equation

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_k \mathbf{v}_k = \mathbf{0}$$

has the unique solution

$$\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$$

Definition (Linear Dependence)

Let V be a **vector space** and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in V$. Then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly dependent (or form a linearly dependent set) if and only if there are real numbers $\alpha_1, \alpha_2, \cdots, \alpha_k$, not all zero, such that

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_k \mathbf{v}_k = \mathbf{0}$$

Example

In $\mathbb{R}^2$, the vectors

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{w} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

are linearly independent. Indeed:

$$\alpha \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \Longrightarrow \qquad \left\{ \begin{array}{rcrcl} \alpha & + & \beta & = & 0 \\ 2\alpha & - & \beta & = & 0 \end{array} \right.$$

The homogeneous linear system has only the trivial solution, $\alpha = 0, \beta = 0$, so linear independence.

Example

In $\mathbb{R}^3$, the following vectors are linearly dependent:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}, \qquad \mathbf{v}_3 = \begin{bmatrix} 4 \\ 5 \\ 11 \end{bmatrix}$$

Indeed: $2\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0}$

Theorem

The set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \subseteq V$ is linearly dependent if and only if at least one vector $\mathbf{v}_i$ is a linear combination of the other vectors.

Proof

$\Longrightarrow$

If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ are linearly dependent then

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k = \mathbf{0}$$

has a solution with some $\alpha_i \neq 0$, then:

$$\mathbf{v}_i = -\frac{\alpha_1}{\alpha_i}\mathbf{v}_1 - \frac{\alpha_2}{\alpha_i}\mathbf{v}_2 - \cdots - \frac{\alpha_{i-1}}{\alpha_i}\mathbf{v}_{i-1} - \frac{\alpha_{i+1}}{\alpha_i}\mathbf{v}_{i+1} + \cdots - \frac{\alpha_k}{\alpha_i}\mathbf{v}_k$$

which is a linear combination of the other vectors

$\Longleftarrow$

If $\mathbf{v}_i$ is a lin combination of the other vectors, eg,

$$\mathbf{v}_i = \beta_1\mathbf{v}_1 + \cdots + \beta_{i-1}\mathbf{v}_{i-1} + \beta_{i+1}\mathbf{v}_{i+1} + \cdots + \beta_k\mathbf{v}_k$$

then

$$\beta_1\mathbf{v}_1 + \cdots + \beta_{i-1}\mathbf{v}_{i-1} - \mathbf{v}_i + \beta_{i+1}\mathbf{v}_{i+1} + \cdots + \beta_k\mathbf{v}_k = \mathbf{0} \qquad \square$$

Corollary

Two vectors are linearly dependent if and only if at least one vector is a scalar multiple of the other.

Example

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}$$

are linearly independent

Theorem

In a **vector space** V*, a non-empty set of vectors that contains the zero vector is linearly dependent.*

Proof:

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \subset V$$

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{0}\}$$

$$0\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 0\mathbf{v}_k + a\mathbf{0} = \mathbf{0}, \qquad a \neq 0$$

Uniqueness of linear combinations

Theorem

If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent vectors in V and if

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_k\mathbf{v}_k = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \ldots + b_k\mathbf{v}_k$$

then

$$a_1 = b_1, \quad a_2 = b_2, \quad \ldots \quad a_k = b_k.$$

- If a vector $\mathbf{x}$ can be expressed as a linear combination of linearly independent vectors, then this can be done in only one way

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k$$

Testing for Linear Independence in $\mathbb{R}^n$

For k vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k$$

is equivalent to

$$A\mathbf{x}$$

where A is the $n \times k$ matrix whose columns are the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ and $\mathbf{x} = [\alpha_1, \alpha_2, \ldots, \alpha_k]^T$:

Theorem

The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$ are linearly dependent if and only if the linear system $A\mathbf{x} = \mathbf{0}$, where A is the matrix $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_k]$, has a solution other than $\mathbf{x} = \mathbf{0}$.
Equivalently, the vectors are linearly independent precisely when the only solution to the system is $\mathbf{x} = \mathbf{0}$.

If vectors are linearly dependent, then any solution $\mathbf{x} \neq \mathbf{0}$, $\mathbf{x} = [\alpha_1, \alpha_2 \ldots, \alpha_k]^T$ of $A\mathbf{x} = \mathbf{0}$ gives a non-trivial linear combination $A\mathbf{x} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_k\mathbf{v}_k = \mathbf{0}$

Example

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \mathbf{v}_3 = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$$

are linearly dependent.
We solve $A\mathbf{x} = \mathbf{0}$

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & -1 & -5 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \end{bmatrix}$$

The general solution is

$$\mathbf{v} = \begin{bmatrix} t \\ -3t \\ t \end{bmatrix}$$

and $A\mathbf{x} = t\mathbf{v}_1 - 3t\mathbf{v}_2 + t\mathbf{v}_3 = \mathbf{0}$
Hence, for $t = 1$ we have: $1\begin{bmatrix} 1 \\ 2 \end{bmatrix} - 3\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ -5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Recall that $A\mathbf{x} = \mathbf{0}$ has precisely one solution $\mathbf{x} = \mathbf{0}$ iff the $n \times k$ matrix is row equiv. to a row echelon matrix with k leading ones, ie, iff $\text{rank}(A) = k$

Theorem

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$. The set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly independent iff the $n \times k$ matrix $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_k]$ has rank k.

Theorem

The maximum size of a linearly independent set of vectors in $\mathbb{R}^n$ is n.

- $\text{rank}(A) \leq \min\{n, k\}$, hence $\text{rank}(A) \leq n \Rightarrow$ when lin. indep. $k \leq n$.
- we exhibit an example that has exactly n independent vectors in $\mathbb{R}^n$ (there are infinite examples):

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \qquad \ldots, \qquad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

This is known as the standard basis of $\mathbb{R}^n$.

Example

$$L_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 9 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 9 \\ 1 \end{bmatrix} \right\} \quad \text{lin. dep.} \quad \text{since } 5 > n = 4$$

$$L_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 9 \\ 2 \end{bmatrix} \right\} \quad \text{lin. indep.}$$

$$L_3 = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 9 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \\ 1 \end{bmatrix} \right\} \quad \text{lin. dep. since } \operatorname{rank}(A) = 2$$

$$L_4 = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 9 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\} \quad \text{lin. dep. since } L_3 \subseteq L_4$$

Linear Independence and Span in $\mathbb{R}^n$

Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a set of vectors in $\mathbb{R}^n$.
What are the conditions for S to span $\mathbb{R}^n$ and be linearly independent?

Let A be the $n \times k$ matrix whose columns are the vectors from S.

- S spans $\mathbb{R}^n$ if for any $v \in \mathbb{R}^n$ the linear system $A\mathbf{x} = \mathbf{v}$ is consistent. This happens when $\operatorname{rank}(A) = n$, hence $k \geq n$
- S is linearly independent iff the linear system $A\mathbf{x} = \mathbf{0}$ has a unique solution. This happens when $\operatorname{rank}(A) = k$, Hence $k \leq n$

Hence, to span $\mathbb{R}^n$ and to be linearly independent, the set S must have exactly n vectors and the square matrix A must have $\det(A) \neq 0$

Example

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix} \qquad |A| = \begin{vmatrix} 1 & 2 & 4 \\ 2 & 1 & 5 \\ 3 & 5 & 1 \end{vmatrix} = 30 \neq 0$$

Outline

Bases

Definition (Basis)

Let V be a **vector space**. Then the subset $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of V is said to be a basis for V if:

1. B is a linearly independent set of vectors, and
2. B spans V; that is, $V = \mathrm{Lin}(B)$

Theorem

*If V is a **vector space**, then a smallest spanning set is a basis of V.*

Theorem

$B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis of V if and only if any $\mathbf{v} \in V$ is a unique linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$

Example

$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$.
the vectors are linearly independent and for any $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n$,
$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$, ie,

$$\mathbf{x} = x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \ldots + x_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Example

The set below is a basis of $\mathbb{R}^2$:

$$S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

- any vector $\mathbf{x} \in \mathbb{R}^2$ can be written as a linear combination of vectors in S.
- any vector $\mathbf{b}$ is a linear combination of the two vectors in S
  $\rightsquigarrow A\mathbf{x} = \mathbf{b}$ is consistent for any $\mathbf{b}$.
- S spans $\mathbb{R}^2$ and is linearly independent

Example

Find a basis of the **subspace** of $\mathbb{R}^3$ given by

$$W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \;\middle|\; x + y - 3z = 0 \right\}.$$

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ -x + 3z \\ z \end{bmatrix} = x \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} = x\mathbf{v} + z\mathbf{w}, \quad \forall x, z \in \mathbb{R}$$

The set $\{\mathbf{v}, \mathbf{w}\}$ spans W. The set is also independent:

$$\alpha\mathbf{v} + \beta\mathbf{w} = \mathbf{0} \implies \alpha = 0, \beta = 0$$

Coordinates

Definition (Coordinates)

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis of a vector space V, then any vector $\mathbf{v} \in V$ can be expressed uniquely as $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n$ then the real numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the coordinates of $\mathbf{v}$ with respect to the basis S. We use the notation

$$[\mathbf{v}]_S = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}_S$$

to denote the coordinate vector of $\mathbf{v}$ in the basis S.

Example

Consider the two basis of $\mathbb{R}^2$:

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \qquad S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

$$[\mathbf{v}]_B = \begin{bmatrix} 2 \\ -5 \end{bmatrix}_B \qquad [\mathbf{v}]_S = \begin{bmatrix} -1 \\ 3 \end{bmatrix}_S$$

In the standard basis the coordinates of $\mathbf{v}$ are precisely the components of the vector $\mathbf{v}$.
In the basis S, they are such that

$$\mathbf{v} = -1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$$

Extension of the main theorem

Theorem

If A is an $n \times n$ matrix, then the following statements are equivalent:

1. *A is invertible*
2. *$A\mathbf{x} = \mathbf{b}$ has a unique solution for any $\mathbf{b} \in \mathbb{R}$*
3. *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution, $\mathbf{x} = \mathbf{0}$*
4. *the reduced row echelon form of A is I.*
5. *$|A| \neq 0$*
6. *The rank of A is n*
7. *The column vectors of A are a basis of $\mathbb{R}^n$*
8. *The rows of A (written as vectors) are a basis of $\mathbb{R}^n$*

(The last statement derives from $|A^T| = |A|$.)
Hence, simply calculating the determinant can inform on all the above facts.

Example

$$\mathbf{v}_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix}2\\1\\5\end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix}4\\5\\11\end{bmatrix}$$

This set is linearly dependent since $\mathbf{v}_3 = 2\mathbf{v}_1 + \mathbf{v}_2$
so $\mathbf{v}_3 \in \text{Lin}(\{\mathbf{v}_1, \mathbf{v}_2\})$ and $\text{Lin}(\{\mathbf{v}_1, \mathbf{v}_2\}) = \text{Lin}(\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\})$.
The linear span of $\{\mathbf{v}_1, \mathbf{v}_2\}$ in $\mathbb{R}^3$ is a plane:

$$\mathbf{x} = \begin{bmatrix}x\\y\\z\end{bmatrix} = s\mathbf{v}_1 + t\mathbf{v}_2 = s\begin{bmatrix}1\\2\\3\end{bmatrix} + t\begin{bmatrix}2\\1\\5\end{bmatrix}$$

The vector $\mathbf{x}$ belongs to the **subspace** iff it can be expressed as a linear combination of $\mathbf{v}_1, \mathbf{v}_2$, that is, if $\mathbf{v}_1, \mathbf{v}_2, \mathbf{x}$ are linearly dependent or:

$$|A| = \begin{vmatrix}1 & 2 & x\\2 & 1 & y\\3 & 5 & z\end{vmatrix} = 0 \qquad \Longrightarrow \qquad |A| = 7x + y - 3z = 0$$

Outline

1. Vector Spaces (cntd)
2. Linear independence
3. Bases
4. Dimension

Dimension

Theorem

Let V be a **vector space** *with a basis*

$$B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$$

of n vectors. Then any set of $n+1$ vectors is linearly dependent.

Proof:
Omitted (choose an arbitrary set of $n+1$ vectors in V and show that since any of them is spanned by the basis then the set must be linearly dependent.)

It follows that:

Theorem

Let a **vector space** V *have a finite basis consisting of* r *vectors. Then any basis of* V *consists of exactly* r *vectors.*

Definition (Dimension)

The number of k vectors in a finite basis of a **vector space** V is the dimension of V and is denoted by $\dim(V)$.
The **vector space** $V = \{\mathbf{0}\}$ is defined to have dimension 0.

- a plane in $\mathbb{R}^2$ is a two-dimensional **subspace**
- a line in $\mathbb{R}^n$ is a one-dimensional **subspace**
- a hyperplane in $\mathbb{R}^n$ is an $(n-1)$-dimensional **subspace** of $\mathbb{R}^n$
- the **vector space** F of real functions is an infinite-dimensional **vector space**
- the **vector space** of real-valued sequences is an infinite-dimensional **vector space**.

Dimension and bases of Subspaces

Example

The plane W in $\mathbb{R}^3$

$$W = \{\mathbf{x} \mid x + y - 3z = 0\}$$

has a basis consisting of the vectors $\mathbf{v}_1 = [1, 2, 1]^T$ and $\mathbf{v}_2 = [3, 0, 1]^T$.

Let $\mathbf{v}_3$ be any vector $\notin W$, eg, $\mathbf{v}_3 = [1, 0, 0]^T$. Then the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of $\mathbb{R}^3$.

Basis of a Linear Space

If we are given k vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$, how can we find a basis for $\mathrm{Lin}(\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\})$?

We can:

- create an $n \times k$ matrix (vectors as columns) and find a basis for the column space by putting the matrix in reduced row echelon form

Summary

- Linear dependence and independence
- Determine linear dependency of a set of vectors, ie, find non-trivial lin. combination that equal zero
- Basis
- Find a basis for a linear space
- Dimension (finite, infinite)

DM559

Linear and Integer Programming

Lecture 8
Linear Transformations

Marco Chiarandini

Department of Mathematics & Computer Science
University of Southern Denmark

Outline

Resume

- Linear dependence and independence
- Determine linear dependency of a set of vectors, ie, find non-trivial lin. combination that equal zero
- Basis
- Find a basis for a linear space
- Dimension (finite, infinite)

Outline

1. Linear Transformations

2. Coordinate Change

3. More on Change of Basis

Linear Transformations

Definition (Linear Transformation)

Let V and W be two vector spaces. A function $T : V \to W$ is linear if for all $\mathbf{u}, \mathbf{v} \in V$ and all $\alpha \in \mathbb{R}$:

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
2. $T(\alpha\mathbf{u}) = \alpha T(\mathbf{u})$

A linear transformation is a linear function between two vector spaces

- If $V = W$ also known as linear operator
- Equivalent condition: $T(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha T(\mathbf{u}) + \beta T(\mathbf{v})$
- for all $\mathbf{0} \in V$, $T(\mathbf{0}) = \mathbf{0}$

Example (Linear Transformations)

- vector space $V = \mathbb{R}$, $F_1(x) = px$ for any $p \in \mathbb{R}$

$$\begin{aligned}\forall x, y \in \mathbb{R}, \alpha, \beta \in \mathbb{R} : F_1(\alpha x + \beta y) &= p(\alpha x + \beta y) = \alpha(px) + \beta(px) \\ &= \alpha F_1(x) + \beta F_1(y)\end{aligned}$$

- vector space $V = \mathbb{R}$, $F_2(x) = px + q$ for any $p, q \in \mathbb{R}$ or $F_3(x) = x^2$ are not linear transformations

$$T(x + y) \neq T(x) + T(y) \quad \text{for some } x, y \in \mathbb{R}$$

- vector spaces $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, $m \times n$ matrix A, $T(\mathbf{x}) = A\mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^n$

$$\begin{aligned}&T(\mathbf{u} + \mathbf{v}) = A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = T(\mathbf{u}) + T(\mathbf{v}) \\ &T(\alpha\mathbf{u}) = A(\alpha\mathbf{u}) = \alpha A\mathbf{u} = \alpha T(\mathbf{u})\end{aligned}$$

Example (Linear Transformations)

- vector spaces $V = \mathbb{R}^n$, $W : f : \mathbb{R} \to \mathbb{R}$. $T : \mathbb{R}^n \to W$:

$$T(\mathbf{u}) = T\left(\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}\right) = p_{u_1, u_2, \ldots, u_n} = p_{\mathbf{u}}$$

$$p_{u_1, u_2, \ldots, u_n} = u_1 x^1 + u_2 x^2 + u_3 x^3 + \cdots + u_n x^n$$

$$\begin{aligned} p_{\mathbf{u}+\mathbf{v}}(x) &= \cdots = (p_{\mathbf{u}} + p_{\mathbf{v}})(x) \\ p_{\alpha\mathbf{u}}(\mathbf{x}) &= \cdots = \alpha p_u(x) \end{aligned}$$

Linear Transformations and Matrices

- any $m \times n$ matrix A defines a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m \rightsquigarrow T_A$
- for every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ there is a matrix A such that $T(\mathbf{v}) = A\mathbf{v} \rightsquigarrow A_T$

Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ denote the standard basis of $\mathbb{R}^n$ and let A be the matrix whose columns are the vectors $T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)$: that is,

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \ldots & T(\mathbf{e}_n) \end{bmatrix}$$

Then, for every $\mathbf{x} \in \mathbb{R}^n$, $T(\mathbf{x}) = A\mathbf{x}$.

Proof: write any vector $\mathbf{x} \in \mathbb{R}^n$ as lin. comb. of standard basis and then make the image of it.

Example

$T : \mathbb{R}^3 \to \mathbb{R}^3$

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y+z \\ x-y \\ x+2y-3z \end{bmatrix}$$

- The image of $\mathbf{u} = [1,2,3]^T$ can be found by substitution: $T(\mathbf{u}) = [6,-1,-4]^T$.

- to find A_T:

$$T(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad T(\mathbf{e}_2) = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \quad T(\mathbf{e}_3) = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$$

$$A = [T(\mathbf{e}_1)\ T(\mathbf{e}_2)\ T(\mathbf{e}_n)] = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 2 & -3 \end{bmatrix}$$

$T(\mathbf{u}) = A\mathbf{u} = [6,-1,-4]^T$.

Linear Transformation in $\mathbb{R}^2$

- We can visualize them!

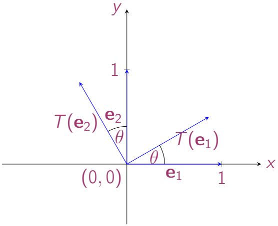
- Reflection in the x axis:

$$T : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x \\ -y \end{bmatrix} \qquad A_T = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

- Stretching the plane away from the origin

$$T(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- Rotation anticlockwise by an angle θ



we search the images of the standard basis vector $\mathbf{e}_1, \mathbf{e}_2$

$$T(\mathbf{e}_1) = \begin{bmatrix} a \\ c \end{bmatrix}, \quad T(\mathbf{e}_2) = \begin{bmatrix} b \\ d \end{bmatrix}$$

they will be orthogonal and with length 1.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

For $\pi/4$:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Identity and Zero Linear Transformations

- For $T : V \to V$ the linear transformation such that $T(\mathbf{v}) = \mathbf{v}$ is called the identity.

- if $V = \mathbb{R}^n$, the matrix $A_T = I$ (of size $n \times n$)

- For $T : V \to W$ the linear transformation such that $T(\mathbf{v}) = \mathbf{0}$ is called the zero transformation.

- If $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$, the matrix A_T is an $m \times n$ matrix of zeros.

Composition of Linear Transformations

- Let $T : V \to W$ and $S : W \to U$ be linear transformations.
  The composition of ST is again a linear transformation given by:

  $$ST(\mathbf{v}) = S(T(\mathbf{v})) = S(\mathbf{w}) = \mathbf{u}$$

  where $\mathbf{w} = T(\mathbf{v})$

- ST means do T and then do S: $V \xrightarrow{T} W \xrightarrow{S} U$

- if $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^p$ in terms of matrices:

  $$ST(\mathbf{v}) = S(T(\mathbf{v})) = S(A_T\mathbf{v}) = A_S A_T \mathbf{v}$$

  note that composition is not commutative

Combinations of Linear Transformations

- If $S, T : V \to W$ are linear transformations between the same vector spaces, then $S + T$ and αS, $\alpha \in \mathbb{R}$ are linear transformations.

- hence also $\alpha S + \beta T$, $\alpha, \beta \in \mathbb{R}$ is

Inverse Linear Transformations

- If V and W are finite-dimensional vector spaces of the same dimension, then the inverse of a lin. transf. $T : V \to W$ is the lin. transf such that

  $$T^{-1}(T(v)) = \mathbf{v}$$

- In $\mathbb{R}^n$ if T^{-1} exists, then its matrix satisfies:

  $$T^{-1}(T(v)) = A_{T^{-1}}A_T\mathbf{v} = I\mathbf{v}$$

  that is, T^{-1} exists iff $(A_T)^{-1}$ exists and $A_{T^{-1}} = (A_T)^{-1}$
  (recall that if $BA = I$ then $B = A^{-1}$)

- In $\mathbb{R}^2$ for rotations:

  $$A_{T^{-1}} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Example

Is there an inverse to $T : \mathbb{R}^3 \to \mathbb{R}^3$

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y+z \\ x-y \\ x+2y-3z \end{bmatrix}$$

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

Since $\det(A) = 9$ then the matrix is invertible, and T^{-1} is given by the matrix:

$$A^{-1} = \frac{1}{9}\begin{bmatrix} 3 & 5 & 1 \\ 3 & -4 & 1 \\ 3 & -1 & -2 \end{bmatrix} \qquad T^{-1}\left(\begin{bmatrix} u \\ v \\ w \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{3}u + \frac{5}{9}v + \frac{1}{9}w \\ \frac{1}{3}u - \frac{4}{9}v + \frac{1}{9}w \\ \frac{1}{3}u + \frac{1}{9}v - \frac{2}{9}w \end{bmatrix}$$

Outline

1. Linear Transformations

2. Coordinate Change

3. More on Change of Basis

Coordinates

Recall:

Definition (Coordinates)

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a basis of a vector space V, then

- any vector $\mathbf{v} \in V$ can be expressed uniquely as $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n$
- and the real numbers $\alpha_1, \alpha_2, \dots, \alpha_n$ are the coordinates of $\mathbf{v}$ wrt the basis S.

To denote the coordinate vector of $\mathbf{v}$ in the basis S we use the notation

$$[\mathbf{v}]_S = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}_S$$

- In the standard basis the coordinates of $\mathbf{v}$ are precisely the components of the vector $\mathbf{v}$: $\mathbf{v} = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + \cdots + v_n \mathbf{e}_n$
- How to find coordinates of a vector $\mathbf{v}$ wrt another basis?

Transition from Standard to Basis B

Definition (Transition Matrix)

Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis of $\mathbb{R}^n$. The coordinates of a vector $\mathbf{x}$ wrt B, $\mathbf{a} = [a_1, a_2, \ldots, a_n]^T = [\mathbf{x}]_B$, are found by solving the linear system:

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n = \mathbf{x} \qquad \text{that is} \qquad \mathbf{x} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \mathbf{v}_n][\mathbf{x}]_B$$

We call P the matrix whose columns are the basis vectors:

$$P = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \mathbf{v}_n]$$

Then for any vector $\mathbf{x} \in \mathbb{R}^n$

$$\mathbf{x} = P[\mathbf{x}]_B \qquad \text{transition matrix from } B \text{ coords to standard coords}$$

moreover P is invertible (columns are a basis):

$$[\mathbf{x}]_B = P^{-1}\mathbf{x} \qquad \text{transition matrix from standard coords to } B \text{ coords}$$

Example

$$B = \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \right\} \qquad [\mathbf{v}]_B = \begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix}$$

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

$\det(P) = 4 \neq 0$ so B is a basis of $\mathbb{R}^3$
We derive the standard coordinates of $\mathbf{v}$:

$$\mathbf{v} = 4 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} - 5 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -9 \\ -3 \\ -5 \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 2 \\ -1 & 4 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix}_B = \begin{bmatrix} -9 \\ -3 \\ -5 \end{bmatrix}$$

Example (cntd)

$$B = \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \right\}, \qquad [\mathbf{x}] = \begin{bmatrix} 5 \\ 7 \\ -3 \end{bmatrix}$$

We derive the B coordinates of vector $\mathbf{x}$:

$$\begin{bmatrix} 5 \\ 7 \\ -3 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + a_2 \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} + a_3 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

either we solve $P\mathbf{a} = \mathbf{x}$ in $\mathbf{a}$ by Gaussian elimination or we find the inverse P^{-1}:

$$[\mathbf{x}]_B = P^{-1}\mathbf{x} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}_B \qquad \text{check the calculation}$$

What are the B coordinates of the basis vector? ($[1, 0, 0], [0, 1, 0], [0, 0, 1]$)

Change of Basis

Since $T(\mathbf{x}) = P\mathbf{x}$ then $T(\mathbf{e}_i) = \mathbf{v}_i$, ie, T maps standard basis vector to new basis vectors

Example

Rotate basis in $\mathbb{R}^2$ by $\pi/4$ anticlockwise, find coordinates of a vector wrt the new basis.

$$A_T = \begin{bmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Since the matrix A_T rotates $\{\mathbf{e}_1, \mathbf{e}_2\}$, then $A_T = P$ and its columns tell us the coordinates of the new basis and $\mathbf{v} = P[\mathbf{v}]_B$ and $[\mathbf{v}]_B = P^{-1}\mathbf{v}$. The inverse is a rotation clockwise:

$$P^{-1} = \begin{bmatrix} \cos(-\frac{\pi}{4}) & -\sin(-\frac{\pi}{4}) \\ \sin(-\frac{\pi}{4}) & \cos(-\frac{\pi}{4}) \end{bmatrix} = \begin{bmatrix} \cos(\frac{\pi}{4}) & \sin(\frac{\pi}{4}) \\ -\sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Example (cntd)

Find the new coordinates of a vector $\mathbf{x} = [1, 1]^T$

$$[\mathbf{x}]_B = P^{-1}\mathbf{x} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix}$$

Change of basis from B to B'

Given a basis B of $\mathbb{R}^n$ with transition matrix P_B,
and another basis B' with transition matrix $P_{B'}$,
how do we change from coords in the basis B to coords in the basis B'?

coordinates in B $\xrightarrow{\mathbf{v}=P_B[\mathbf{v}]_B}$ standard coordinates $\xrightarrow{[\mathbf{v}]_{B'}=P_{B'}^{-1}\mathbf{v}}$ coordinates in B'

$$[\mathbf{v}]_{B'} = P_{B'}^{-1}P_B[\mathbf{v}]_B$$

$$M = P_{B'}^{-1}P_B = P_{B'}^{-1}[\mathbf{v}_1\ \mathbf{v}_2\ \dots\ \mathbf{v}_n] \overset{ex11sh2}{=} [P_{B'}^{-1}\mathbf{v}_1\ P_{B'}^{-1}\mathbf{v}_2\ \dots\ P_{B'}^{-1}\mathbf{v}_n]$$

Theorem

If B and B' are two bases of $\mathbb{R}^n$, with

$$B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$$

then the transition matrix from B coordinates to B' coordinates is given by

$$M = \left[[\mathbf{v}_1]_{B'}\ \ [\mathbf{v}_2]_{B'}\ \ \cdots\ \ [\mathbf{v}_n]_{B'}\right]$$

(the columns of M are the B' coordinates of the basis B)

Example

$$B = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\} \qquad B' = \left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \end{bmatrix} \right\}$$

are basis of $\mathbb{R}^2$, indeed the corresponding transition matrices from standard basis:

$$P = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \qquad Q = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$$

have $\det(P) = 3$, $\det(Q) = 1$. Hence, lin. indep. vectors.
We are given

$$[\mathbf{x}]_B = \begin{bmatrix} 4 \\ -1 \end{bmatrix}_B$$

find its coordinates in B'.

Example (cntd)

1. find first the standard coordinates of $\mathbf{x}$

$$\mathbf{x} = 4\begin{bmatrix}1\\2\end{bmatrix} - \begin{bmatrix}-1\\1\end{bmatrix} = \begin{bmatrix}1 & -1\\2 & 1\end{bmatrix}\begin{bmatrix}4\\-1\end{bmatrix} = \begin{bmatrix}5\\7\end{bmatrix}$$

   and then find B' coordinates:

$$[\mathbf{x}]_S = Q^{-1}\mathbf{x} = \begin{bmatrix}2 & -5\\-1 & 3\end{bmatrix}\begin{bmatrix}5\\7\end{bmatrix} = \begin{bmatrix}-25\\16\end{bmatrix}_S$$

2. use transition matrix M from B to B' coordinates:
   $\mathbf{v} = P[\mathbf{v}]_B$ and $\mathbf{v} = Q[\mathbf{v}]_{B'}$ $\rightsquigarrow$ $[\mathbf{v}]_{B'} = Q^{-1}P[\mathbf{v}]_B$:

$$M = Q^{-1}P = \begin{bmatrix}2 & -5\\-1 & 3\end{bmatrix}\begin{bmatrix}1 & -1\\2 & 1\end{bmatrix} = \begin{bmatrix}-8 & -7\\5 & 4\end{bmatrix}$$

$$[\mathbf{x}]_{B'} = \begin{bmatrix}-8 & -7\\5 & 4\end{bmatrix}\begin{bmatrix}4\\-1\end{bmatrix} = \begin{bmatrix}-25\\16\end{bmatrix}_{B'}$$

Outline

1. Linear Transformations

2. Coordinate Change

3. More on Change of Basis

Change of Basis for a Lin. Transf.

We saw how to find A for a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ using standard basis in both $\mathbb{R}^n$ and $\mathbb{R}^m$. Now: is there a matrix that represents T wrt two arbitrary bases B and B'?

Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ and $B' = \{\mathbf{v}'_1, \mathbf{v}'_2, \dots, \mathbf{v}'_n\}$ be bases of $\mathbb{R}^n$ and $\mathbb{R}^m$. Then for all $\mathbf{x} \in \mathbb{R}^n$, $\qquad [T(\mathbf{x})]_{B'} = M[\mathbf{x}]_B$
where $M = A_{[B,B']}$ is the $m \times n$ matrix with the ith column equal to $[T(\mathbf{v}_i)]_{B'}$, the coordinate vector of $T(\mathbf{v}_i)$ wrt the basis B'.

Proof:

$$\begin{array}{cl}
\text{change } B \text{ to standard} & \mathbf{x} = P_B^{n\times n}[\mathbf{x}]_B \quad \forall \mathbf{x} \in \mathbb{R}^n \\
\downarrow & \\
\text{perform linear transformation} & T(\mathbf{x}) = A\mathbf{x} = AP_B^{n\times n}[\mathbf{x}]_B \\
\text{in standard coordinates} & \\
\downarrow & \\
\text{change to basis } B' & [\mathbf{u}]_{B'} = (P_{B'}^{m\times m})^{-1}\mathbf{u} \quad \forall \mathbf{u} \in \mathbb{R}^m \\
 & [T(\mathbf{x})]_{B'} = (P_{B'}^{m\times m})^{-1}AP_B^{n\times n}[\mathbf{x}]_B \\
 & M = (P_{B'}^{m\times m})^{-1}AP_B^{n\times n}
\end{array}$$

How is M done?

- $P_B = [\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_n]$
- $AP_B = A[\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_n] = [A\mathbf{v}_1 \ A\mathbf{v}_2 \ \ldots \ A\mathbf{v}_n]$
- $A\mathbf{v}_i = T(\mathbf{v}_i)$: $AP_B = [T(\mathbf{v}_1) \ T(\mathbf{v}_2) \ \ldots \ T(\mathbf{v}_n)]$
- $M = P_{B'}^{-1}AP_B = [P_{B'}^{-1}T(\mathbf{v}_1) \ P_{B'}^{-1}T(\mathbf{v}_2) \ \ldots \ P_{B'}^{-1}T(\mathbf{v}_n)]$
- $M = [[T(\mathbf{v}_1)]_{B'} \ [T(\mathbf{v}_2)]_{B'} \ \ldots \ [T(\mathbf{v}_n)]_{B'}]$

Hence, if we change the basis from the standard basis of $\mathbb{R}^n$ and $\mathbb{R}^m$ the matrix representation of T changes

Similarity

Particular case $m = n$:

Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a basis $\mathbb{R}^n$. Let A be the matrix corresponding to T in standard coordinates: $T(\mathbf{x}) = A\mathbf{x}$. Let

$$P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}$$

be the matrix whose columns are the vectors of B. Then for all $\mathbf{x} \in \mathbb{R}^n$,

$$[T(\mathbf{x})]_B = P^{-1}AP[\mathbf{x}]_B$$

Or, the matrix $A_{[B,B]} = P^{-1}AP$ performs the same linear transformation as the matrix A but expressed it in terms of the basis B.

Similarity

Definition

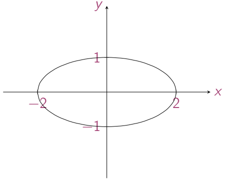
A square matrix C is similar (represent the same linear transformation) to the matrix A if there is an invertible matrix P such that

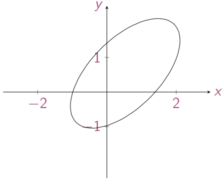
$$C = P^{-1}AP.$$

Similarity defines an equivalence relation:

- (reflexive) a matrix A is similar to itself
- (symmetric) if C is similar to A, then A is similar to C
  $C = P^{-1}AP, \quad A = Q^{-1}CQ, \quad Q = P^{-1}$
- (transitive) if D is similar to C, and C to A, then D is similar to A

Example





- $x^2 + y^2 = 1$ circle in standard form
- $x^2 + 4y^2 = 4$ ellipse in standard form
- $5x^2 + 5y^2 - 6xy = 2$??? Try rotating $\pi/4$ anticlockwise

$$A_T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = P$$

$$\mathbf{v} = P[\mathbf{v}]_B \iff \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$$

$$X^2 + 4Y^2 = 1$$

Example

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$:

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + 3y \\ -x + 5y \end{bmatrix}$$

What is its effect on the xy-plane?
Let's change the basis to

$$B = \{\mathbf{v}_1, \mathbf{v}_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}$$

Find the matrix of T in this basis:

- $C = P^{-1}AP$, A matrix of T in standard basis, P is transition matrix from B to standard

$$C = P^{-1}AP = \frac{1}{2}\begin{bmatrix} -1 & 3 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

Example (cntd)

- the B coordinates of the B basis vectors are

$$[\mathbf{v}_1]_B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_B, \quad [\mathbf{v}_2]_B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}_B$$

- so in B coordinates T is a stretch in the direction $\mathbf{v}_1$ by 4 and in dir. $\mathbf{v}_2$ by 2:

$$[T(\mathbf{v}_1)]_B = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}_B = \begin{bmatrix} 4 \\ 0 \end{bmatrix}_B = 4[\mathbf{v}_1]_B$$

- The effect of T is however the same no matter what basis, only the matrices change! So also in the standard coordinates we must have:

$$A\mathbf{v}_1 = 4\mathbf{v}_1 \qquad A\mathbf{v}_2 = 2\mathbf{v}_2$$

Summary

- Linear transformations and proofs that a given mapping is linear
- two-way relationship between matrices and linear transformations
- change from standard (S) to arbitrary basis (B)
- change of basis between two arbitrary basis (from B to B')
- Matrix representation of a transformation with respect to two arbitrary basis
- Similarity of square matrices