Basic Mathematics , Fall 2020

Karen Keryan
ASDS, YSU

September 16, 2020

WELCOME TO

Basic Mathematics!

Course Details

- Course: **Basic Mathematics**, Credits:**6**
- Lectures: **Tuesday** & **Wednesday 18:30-19:50**
- Instructor: Karen Keryan
- Grading Policy:

  Total=0.1∗(HW)+0.25∗(M1+M2)+0.4∗ F

Course Format

- Lectures:
  Tuesdays& Wednesday, 18:30-19:50 (Sep & Oct)
  Tuesday, 18:30-19:50 (Nov & Dec)
- Homeworks: weekly (except the exam weeks)
- No late Homework! (Except for some exceptional cases)

Course Content

- Linear Algebra
- Calculus
- Probability theory

Textbooks

- Mathematics for Machine Learning by Marc Peter Deisenroth, A Aldo Faisal, and Cheng Soon Ong. https://mml-book.github.io/

Additional textbooks:

- Linear Algebra: A Modern Introduction, 4th Ed, by David Poole
- Calculus, 7th Ed, by James Stewart
- A First Course in Probability, 8th Ed, by Sheldon Ross

Find other interesting stuff on the syllabus!

Definition

With $m, n \in \mathbb{N}$ a real-valued (m, n) matrix A is matrix an $m \cdot n$-tuple of elements a_{ij} , $i = 1, \ldots, m,\ \ j = 1, \ldots, n,$ which is ordered in m rows and n columns:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, a_{ij} \in \mathbb{R}.$$

$(1, n)$*-matrices are called* **rows,** $(m, 1)$*-matrices are called* **columns**. *These special matrices are also called* **row/column vectors**.

$\mathbb{R}^{m \times n}$ is the set of all real-valued (m, n)-matrices. $A \in R^{m \times n}$ can be equivalently represented as $a \in \mathbb{R}^{mn}$ by stacking all n columns of the matrix into a long vector.

Matrix Multiplication

Definition

If A is an $m \times n$ matrix and B is an $n \times r$ matrix, then the product $C = AB$ is an $m \times r$ matrix. The (i, j) entry of the product is computed as follows:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Remark

1. The number of columns of A must be the same as the number of rows of B.
2. The (i, j) entry of the matrix AB is the dot product of the ith row of A and the jth column of B :

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

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \ldots & b_{1j} & \ldots & b_{1r} \\ b_{21} & \ldots & b_{2j} & \ldots & b_{2r} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \ldots & b_{nj} & \ldots & b_{nr} \end{bmatrix}$$

Example

Compute AB if

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

Check that $AB \neq BA$.

Definition

In $R^{n \times n}$, we define the **identity matrix**

$$I_n = \begin{bmatrix} 1 & 0 & 0 \dots & 0 \\ 0 & 1 & 0 \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 \dots & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}$$

as the $n \times n$-matrix containing 1 on the diagonal and 0 everywhere else.
With this, $A \cdot I_n = A = I_n \cdot A$ for all $A \in \mathbb{R}^{n \times n}$.

Properties of matrices

- Associativity:
  $\forall A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{p \times q} : (AB)C = A(BC)$
- Distributivity:
  $$\forall A, B \in \mathbb{R}^{m \times n},\ C, D \in \mathbb{R}^{n \times p} : \quad \begin{aligned} (A+B)C &= AC + BC \\ A(C+D) &= AC + AD \end{aligned}$$
- Neutral element (identity element):
  $\forall A \in \mathbb{R}^{m \times n} : \quad I_m A = A I_n = A.$
  Note that $I_m \neq I_n$ for $m \neq n$.

Definition

For a square matrix $A \in \mathbb{R}^{n\times n}$ a matrix $B \in \mathbb{R}^{n\times n}$ with $AB = I_n = BA$ is called **inverse** *and denoted by A^{-1}.*

Not every matrix A has an inverse A^{-1}. If this inverse does exist, A is called **regular/invertible/non-singular**, otherwise **singular/non-invertible.**

Existence of the Inverse of a 2×2-Matrix

Consider a matrix

$$A := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Then for $B = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$ we get

$$AB = \begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} = (a_{11}a_{22} - a_{12}a_{21}) I_2.$$

Therefore

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix},$$

if $a_{11}a_{22} - a_{12}a_{21} \neq 0$.

Example

The matrices

$$A = \begin{bmatrix} 4 & 5 & 4 \\ 7 & 7 & 6 \\ 2 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 2 \\ 5 & -4 & 4 \\ -7 & 6 & -7 \end{bmatrix}$$

are inverse to each other since $AB = I = BA$.

Definition

For $A \in \mathbb{R}^{m \times n}$ the matrix $B \in \mathbb{R}^{n \times m}$ with $b_{ij} = a_{ji}$ is called the transpose of A. We write $B = A^T$.

A^T is achieved by any one of the following equivalent actions:

- reflect A over its main diagonal (which runs from top-left to bottom-right) to obtain A^T,
- write the rows of A as the columns of A^T,
- write the columns of A as the rows of A^T.

Properties of inverses and transposes

- $AA^{-1} = I = A^{-1}A$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^T)^T = A$
- $(A+B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$
- If A is invertible, then so is A^T and $(A^{-1})^T = (A^T)^{-1}$

Definition

A square matrix A is symmetric if $A = A^T$.

Example

Prove that for any square matrix A the matrices $A + A^T$ and $A \cdot A^T$ are symmetric.

Example

Prove that for any symmetric matrices A, B the matrix $A + B$ is symmetric.

Remark

The product of two symmetric matrices is not symmetric in general, e.g. the matrices $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, *but the matrix* $AB = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ *is not symmetric.*

Multiplication by a Scalar

If A is an $m \times n$ matrix and λ is a scalar, then the scalar multiple λA is the $m \times n$ matrix

$$\lambda A = \lambda[a_{ij}] = [\lambda a_{ij}].$$

- Distributivity:
  $(\lambda_1 + \lambda_2)C = \lambda_1 C + \lambda_2 C,\ C \in \mathbb{R}^{m\times n}$
  $\lambda(B + C) = \lambda B + \lambda C,\ B, C \in \mathbb{R}^{m\times n}$
- Associativity:
  $(\lambda_1\lambda_2)C = \lambda_1(\lambda_2 C),\ C \in \mathbb{R}^{m\times n}$
  $\lambda(BC) = (\lambda B)C = B(\lambda C) = (BC)\lambda, \quad B \in \mathbb{R}^{m\times n},\ C \in \mathbb{R}^{n\times k}$
- $(\lambda C)^T = \lambda C^T$

Example

Check that

$$(\lambda_1 + \lambda_2)C = \lambda_1 C + \lambda_2 C$$

for any $\lambda_1, \lambda_2 \in R$ *and* $C := \begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix}$.

Compact Representations of Systems of Linear Equations

The system of linear equations

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

can be written as vector equation

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

and can be compactly written in matrix notation as follows

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

Note that Ax is a linear combination of the columns of A:

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

Let us introduce the general form of a system of linear equations,

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

where $a_{ij} \in \mathbb{R}$ and $b_i \in \mathbb{R}$ are known constants and x_j are unknowns, $i = 1, \ldots, m,\ j = 1, \ldots, n.$
It can be written in matrix form $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \\ a_{m1} & \ldots & a_{mn} \end{bmatrix},\ \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

Particular and General Solution

Consider the following system of linear equations (SLE):

$$\begin{bmatrix} 1 & 0 & 7 & 2 \\ 0 & 1 & 3 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$$

or, equivalently $\sum_{i=1}^{4} x_i \mathbf{c}_i = \mathbf{b}$, where $\mathbf{c}_i$ is the ith column of the matrix and $\mathbf{b}$ is the right-hand-side.
$[5, 8, 0, 0]^T$ is a solution, since $\mathbf{b} = 5\mathbf{c}_1 + 8\mathbf{c}_2 + 0\mathbf{c}_3 + 0\mathbf{c}_4$.
This solution is called a **particular solution** or **special solution**.

To capture all the other solutions, we express the third column using the first two columns

$$\begin{bmatrix}7\\3\end{bmatrix} = 7\begin{bmatrix}1\\0\end{bmatrix} + 3\begin{bmatrix}0\\1\end{bmatrix}, \text{equivalently } \mathbf{0} = 7\mathbf{c}_1 + 3\mathbf{c}_2 - 1\mathbf{c}_3 + 0\mathbf{c}_4$$

So $(x_1, x_2, x_3, x_4) = (7, 3, -1, 0)$ produces the $\mathbf{0}$ vector as well as any scaling of it by λ_1, i.e. $\lambda_1(7, 3, -1, 0)$.

$$\begin{bmatrix}1 & 0 & 7 & 2\\0 & 1 & 3 & -4\end{bmatrix}\left(\lambda_1\begin{bmatrix}7\\3\\-1\\0\end{bmatrix}\right) = \lambda_1(7\mathbf{c}_1 + 3\mathbf{c}_2 - 1\mathbf{c}_3 + 0\mathbf{c}_4) = \mathbf{0}$$

Similarly,

$$\begin{bmatrix}1 & 0 & 7 & 2\\0 & 1 & 3 & -4\end{bmatrix}\left(\lambda_2\begin{bmatrix}2\\-4\\0\\-1\end{bmatrix}\right) = \lambda_1(2\mathbf{c}_1 - 4\mathbf{c}_2 + 0\mathbf{c}_3 - 1\mathbf{c}_4) = \mathbf{0}$$

All solutions of the SLE, which is called the **general solution**, is the set

$$\left\{ \mathbf{x} \in \mathbb{R}^4 ; \mathbf{x} = \begin{bmatrix} 5 \\ 8 \\ 0 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 7 \\ 3 \\ -1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2 \\ -4 \\ 0 \\ -1 \end{bmatrix}, \lambda_1, \lambda_2 \in \mathbb{R} \right\}.$$

The algorithm.

1. Find a particular solution to $A\mathbf{x} = \mathbf{b}$.
2. Find all solutions to $A\mathbf{x} = \mathbf{0}$.
3. Combine the solutions from 1. and 2. to the general solution.

Definition

Given a SLE

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases}.$$

$$A := \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$ *is called the* **coefficient matrix**

and $\left[\ A \mid \mathbf{b}\ \right]$ *is called the* **augmented matrix**,

where $\mathbf{b} = \left[b_1, b_2, \cdot\cdot, b_n\right]^T$

Definition

A matrix is in **row echelon form (REF)** *if it satisfies the following properties:*
1. Any rows consisting entirely of zeros are at the bottom.
2. In each nonzero row, the first nonzero entry (called the **leading entry** *or* **pivot***) is in a column to the left of any leading entries below it.*

Example

The following matrices are in row echelon form:

$$\begin{bmatrix} 1 & 3 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & -2 & 7 \\ 0 & 4 & 1 \\ 0 & 0 & 3 \end{bmatrix} \quad \begin{bmatrix} 3 & 1 & -2 & 7 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

Assuming that each of these matrices is an augmented matrix, write out the corresponding SLE and solve them.

Elementary Row Operations

Definition

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

1. Interchange two rows.
2. Multiply a row by a nonzero constant.
3. Add a multiple of a row to another row.

Shorthand notation

1. $R_i \leftrightarrow R_j$ means interchange rows i and j.
2. kR_i means multiply row i by k.
3. $R_i + kR_j$ means add k times row j to row i (and replace row i with the result).

Example

Reduce the following matrix to echelon form:

$$\begin{bmatrix} 1 & -1 & 2 & 4 \\ 3 & -1 & 8 & 15 \\ -2 & 0 & -4 & -2 \end{bmatrix}.$$

Definition

The variables corresponding to the pivots in the row-echelon form are called **basic variables** *or* **leading variables**, *the other variables are* **free variables**.

Example

Let the REF of the augmented matrix be

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

The leading entries in row echelon form are x_1, x_3, *and the free variable is* x_2.