Mathematics for Computer Science
Linear Algebra

Lecture 1: Matrices

Andrei Krokhin

October 6, 2020

Communication

- All lectures are on Zoom, recordings will appear on DUO.
- During lectures, I will:
  - monitor the chat during lectures - if you have a question, type it or just say "question" in the chat.
  - use (anonymous) polls to check your understanding
  - sometimes ask you to generate ideas (e.g. for a proof). I plan to use breakout rooms on Zoom for this.
- Outside lectures, I'll be happy to
  - answer questions by email
  - meet (online) by appointment or during my office hours (TBA)
  - use MCS Chat forum on DUO

Formalities

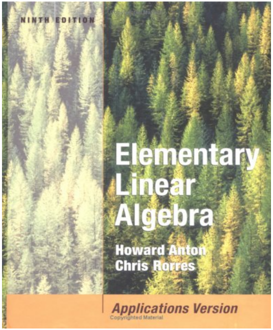
- Lectures: one LA lecture per week, teaching weeks 1-22
  - new material – week 1-20 (terms 1 and 2)
  - revision lectures – weeks 21-22 (term 3)
- Practicals: starting week 2, every second week (alternating with Calculus)

Assessment:

- Exam: 66% of the module, in May/June next year (term 3)
- Summative coursework: 34%, combined assignment for Calc and LA
  - Hand-out in November, hand-in in December (last week of term 1)

Textbook

"Elementary linear algebra" by Howard Anton and Chris Rorres



- Many versions/editions of this book in the library/colleges, all are good for us. You don't have to buy it.
- We'll follow it, but not closely, and lecture slides should be sufficient.
- You can use it to get additional explanations and examples.

Plan for this term

Linear algebra is the study of matrices and vectors

- Matrices, determinants and linear systems (5 lectures)
- Vector spaces (5 lectures)

Where is Linear Algebra applied in Computer Science?

- Network Science (e.g. ranking results from an Internet search engine)
- Computer Graphics / Computer Vision / Virtual Reality
- Machine Learning / Data Science
- Quantum Computing
- Codes and Cryptography
- Parallel Scientific Computing
- ...

- We teach a module on each of the above, in Levels 2-4
- Many more: > 200 pages of various applications in the textbook

Watch the popular video "The applications of matrices: What I wish my teachers told me way earlier": `https://youtu.be/rowWM-MijXU`.

Contents for today's lecture

- Matrix operations: addition and multiplication;
- Matrix operations: transpose and inverse;
- Special matrices: symmetric, diagonal, and triangular
- Examples and exercises.

Matrices

Definition

A matrix is a rectangular array of numbers. The entry in row i and column j is denoted by a_{ij}, b_{ij}, etc.

$$\begin{pmatrix} 1 & 2 \\ 3 & 0 \\ -1 & 4 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 0 & -3 \end{pmatrix}, \begin{pmatrix} \pi & \sqrt{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \end{pmatrix}$$

- A matrix with m rows and n columns is said to have size $m \times n$.
- A general $m \times n$ matrix can be written as

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

- A matrix of size $n \times n$ is called a square matrix of order n.

Matrix operations

Let $A = (a_{ij})$ and $B = (b_{ij})$ be $m \times n$ matrices.

- The sum $A + B$ is defined as the $m \times n$ matrix $C = (c_{ij})$ such that $c_{ij} = a_{ij} + b_{ij}$.
- The difference $A - B$ is defined similarly.
- If α is a number (scalar) then the product (of a matrix by a scalar) αA is the $m \times n$ matrix $C = (c_{ij})$ such that $c_{ij} = \alpha \cdot a_{ij}$.

Example: Let

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 3 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 2 & 7 \\ -1 & 3 & -5 \end{pmatrix}$$

Then

$$2A - B = \begin{pmatrix} 4 & 6 & 8 \\ 2 & 6 & 2 \end{pmatrix} - \begin{pmatrix} 0 & 2 & 7 \\ -1 & 3 & -5 \end{pmatrix} = \begin{pmatrix} 4 & 4 & 1 \\ 3 & 3 & 7 \end{pmatrix}$$

Matrix multiplication

- If A is an $m \times r$ matrix and B an $r' \times n$ matrix then the product matrix AB is defined only if $r = r'$.
- If $r = r'$ then the product AB is the $m \times n$ matrix $C = (c_{ij})$ such that

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ir}b_{rj}.$$

Example:

$$\begin{pmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{pmatrix} \begin{pmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{pmatrix} = \begin{pmatrix} 12 & 27 & 30 & y \\ 8 & x & 26 & 12 \end{pmatrix}$$

$$(2 \cdot 4) + (6 \cdot 3) + (0 \cdot 5) = 26$$

Exercise: Find x and y.

- $x = 2 \cdot 1 + 6 \cdot (-1) + 0 \cdot 7 = -4$
- $y = 1 \cdot 3 + 2 \cdot 1 + 4 \cdot 2 = 13$

Computing matrix product

- How many additions and multiplications (of numbers) does one need to perform to multiply two $n \times n$ matrices?
- Computing by definition takes ?? multiplications and ?? additions.
- In practice, one needs to multiply (big) matrices very often – hence, more efficient matrix multiplication algorithms are desirable.
- Often, ω denotes the matrix multiplication exponent: smallest number such that a mm algorithm using (the order of) n^ω operations exists.
- The exact value of ω is unknown! State-of-the-art: $2 \leq \omega \leq 2.3726$.
- In L2, you will (probably) learn some good algorithms for matrix multiplication, e.g. Strassen's algorithm using the order of $n^{\log_2 7}$ operations ($\log_2 7 \approx 2.81$).

Properties of matrix arithmetic

Theorem

Assuming that the sizes of the matrices are such that the operations can be performed, the following rules are valid:

1. $A + B = B + A$
2. $A + (B + C) = (A + B) + C$
3. $A(BC) = (AB)C$
4. $A(B \pm C) = AB \pm AC$
5. $(B \pm C)A = BA \pm CA$
6. $\alpha(B \pm C) = \alpha B \pm \alpha C$
7. $(\alpha \pm \beta)A = \alpha A \pm \beta A$
8. $\alpha(\beta A) = (\alpha\beta)A$
9. $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

The proofs (omitted) for all items easily follow from definitions.

More properties of matrix arithmetic

Some properties of matrix arithmetic differ from number arithmetic:

In general, even for square matrices of the same size, it is possible that

- $AB \neq BA$ (in words: matrix multiplication is not commutative);
- $AB = O$, but $A \neq O$ and $B \neq O$, where O is the (all-)zero matrix of appropriate size;
- $AC = BC$ for some $C \neq O$, but $A \neq B$.

Example: use

$$\begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 3 & 7 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 7 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 17 \\ 0 & 0 \end{pmatrix}$$

Matrix transpose

For an $m \times n$ matrix A, the transpose of A is the $n \times m$ matrix A^T such that the i-th row of A is the i-th column of A^T.

Example:

$$\text{If } A = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 3 & 1 \end{pmatrix} \text{ then } A^T = \begin{pmatrix} 2 & 1 \\ 3 & 3 \\ 4 & 1 \end{pmatrix}$$

Theorem

If the sizes of the matrices are such that the operations can be performed then the following rules are valid:

1. $(A^T)^T = A$
2. $(A \pm B)^T = A^T \pm B^T$
3. $(\alpha A)^T = \alpha A^T$
4. $(AB)^T = B^T A^T$.

Symmetric matrices

Definition

A square matrix $A = (a_{ij})$ is called symmetric if $A = A^T$. In other words, if we have $a_{ij} = a_{ji}$ for all i, j.

The symmetry is about the main diagonal (i.e. the line $a_{11}, a_{22}, \ldots$)

Examples: the first matrix below is symmetric, while the second is not

$$\begin{pmatrix} -2 & 2 & 0 \\ 2 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 2 & 0 \\ 2 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Easy to see: If A and B are symmetric then so are αA and $A + B$.

Question: If A and B are symmetric, when is AB symmetric?

The identity matrix and matrix inverse

- A square matrix (a_{ij}) such that $a_{ii} = 1$ and $a_{ij} = 0$ if $i \neq j$ is called the identity matrix, denoted I_n.

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

- It is easy to check that, for any $m \times n$ matrix A, $AI_n = A = I_m A$.

- If A is a square matrix of order n and if a matrix B of the same size can be found such that $AB = BA = I_n$ then A is said to be invertible (or non-singular), and B is called an inverse of A.
- In this case A and B are inverse of each other.
- If no such B can be found then A is called singular.

Properties of matrix inverse

- If B and C are both inverses of A then $B = C$. Indeed, we have

$$B = BI = B(AC) = (BA)C = IC = C$$

  So, we can speak of the inverse of A, usually denoted by A^{-1}.

- If A and B are invertible matrices of the same size then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$. Indeed,

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$$

- If A is invertible then A^T is also invertible and $(A^T)^{-1} = (A^{-1})^T$. Indeed,

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

  It follows that if, in addition, A is symmetric then so is A^{-1}.

Diagonal matrices

Diagonal matrices will play a special role later, they are of the form

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 & 0 \\ 0 & d_2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & d_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & d_n \end{pmatrix}$$

Question: Let D be a diagonal matrix. When does D^{-1} exists?

Question: If D is diagonal and DA exists, can you describe DA?

Triangular matrices

Definition

A square matrix $A = (a_{ij})$ is called

- lower triangular if $a_{ij} = 0$ for all $i < j$ (above main diagonal), and
- upper triangular if $a_{ij} = 0$ for all $i > j$ (below main diagonal)

Examples: L is lower triangular and U is upper triangular.

$$L = \begin{pmatrix} -2 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Such matrices have important algorithmic applications (as we'll see soon)

Question: what matrices are both symmetric and (lower) triangular?

Basic properties of triangular matrices

Theorem

Let A and B be lower triangular matrices of the same size. Then

1. $(A)^T$ *is upper triangular,*
2. $A \pm B$ *is lower triangular,*
3. αA *is lower triangular,*
4. AB *is lower triangular.*

Proof. (1-3) are obvious. To see (4), let $A = (a_{ij}), B = (b_{ij})$, and let $AB = C = (c_{ij})$. By assumption, we have $a_{ij} = b_{ij} = 0$ for all $i < j$.

Need to show that $c_{ij} = 0$ if $i < j$. Fix any such pair i, j. We have

$$c_{ij} = \underbrace{a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots a_{i(j-1)}b_{(j-1)j}}_{=0 \text{ because each } b \text{ here is } 0} + \underbrace{a_{ij}b_{jj} + \cdots + a_{in}b_{nj}}_{=0 \text{ because each } a \text{ here is } 0} = 0.$$

$\square$

What we learnt today

1. What matrices are and how operations on matrices work:
   - addition, multiplication by a scalar
   - multiplication - behaves not like multiplication of numbers
   - transpose and inverse
2. Important special types of matrices:
   - zero and identity
   - symmetric
   - diagonal
   - lower and upper triangular
3. Open problem: the matrix multiplication exponent.