Linear Algebra
Week 2

G-07

03 X 2024

1 Matrices and Linear Combinations

Matrices are a set of vectors. But they are more than that. They are linear systems of equations and also functions. Matrices are images, neural network models, backbone of AI and computation.

1.1 Definitions

Matrices are rectangular array of (real) numbers. Here is an $m \times n$ matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}, A \in \mathbb{R}^{m \times n}$$

There are other notations which make writing and understanding the matrices relatively easier.

$$A = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_n \\ | & | & & | \end{bmatrix}$$

(a) Column Notation

$$\begin{bmatrix} — & u_1 & — \\ — & u_2 & — \\ & \vdots & \\ — & u_m & — \end{bmatrix}$$

(b) Row Notation

Question 2.1 What are the dimensions of the vectors $v_1, \dots, v_n$ and the vecotrs $u_1, \dots, u_m$?

1.2 Matrix Addition and Scalar Multiplication

Matrix addition and scalar multiplication are defined element-wise. This is why, in order to add two matrices A and B **their dimensions must match!**

Matrix Addition: $A + B$

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

$$= \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{nm1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

Scalar Multiplication: $\lambda \cdot A$

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

Just like vectors. Coincidence? No. The truth is, vectors are $1 \times n$ (row vector) or $n \times 1$ (column vector) matrices. So vector addition and scalar multiplication with vectors are just special cases of matrix addition and scalar multiplication.

1.3 Matrix Shapes

The names that we came up with to describe the matrices might sound funny but they tell us a lot about the matrix. For now just be aware of the intuitive names:

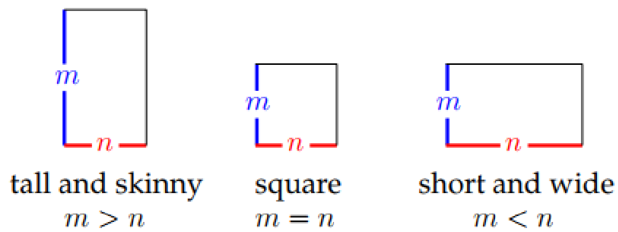

Figure 2.1. from the lecture notes

1.3.1 Square Matrices

There are special square matrices. We use them at various tasks as you will see later. Be aware of the fact that the definition of these matrices are made on **square matrices** in the lecture notes. This implies that a non square matrix *can not* belong to one of these categories.

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

identity matrix — diagonal matrix — upper triangular matrix — lower triangular matrix — symmetric matrix

Examples from lecture notes p.40

Question 2.2 Try to think about the formal definitions of these special square matrices. Can you write them down without the help of the lecture notes?

A square matrix can be attributed with more than one of the features listed above. For example a diagonal matrix is an upper triangular and also a lower triangular matrix, it is symmetric as well. *E.g.* a zero matrix has all four features together.

1.4 Matrix Vector Multiplication

Matrix vector multiplication is nothing else than the linear combination of the columns of the matrix.

$$2\begin{bmatrix}1\\3\\4\end{bmatrix}+3\begin{bmatrix}2\\9\\7\end{bmatrix}=\begin{bmatrix}1&2\\3&9\\4&7\end{bmatrix}\begin{bmatrix}2\\3\end{bmatrix}$$

In general:

$$A=\begin{bmatrix}| & | & & |\\ v_1 & v_2 & \dots & v_n\\ | & | & & |\end{bmatrix}\in\mathbb{R}^{m\times n}, \mathbf{x}=\begin{bmatrix}x_1\\x_2\\\dots\\x_n\end{bmatrix}\in\mathbb{R}^n$$

Their multiplication is the linear combination of the column vectors of the matrix A:

$$A\mathbf{x}:=\sum_{j=1}^{n}x_jv_j\in\mathbb{R}^m$$

In column notation: with $x_i\in\mathbb{R}$ (scalar) and $\mathbf{v}_i\in\mathbb{R}^m$:

$$\begin{bmatrix}| & | & & |\\ x_1\mathbf{v}_1 & x_2\mathbf{v}_2 & \dots & x_n\mathbf{v}_n\\ | & | & & |\end{bmatrix}$$

Question 2.3 Write $A\mathbf{x}$ in row notation.

1.5 Column Space

Column space of a matrix is the span of its column vectors. Formally for $A \in \mathbb{R}^{m \times n}$ the column space $\mathbf{C}(A)$ is:

$$\mathbf{C}(A) := \{A\mathbf{x} : x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$$

Zero is always in the column space of a matrix. $\mathbf{0} \in \mathbf{C}(A)$. Do not forget that the $\mathbf{0}$ is an abuse of notation here and $\mathbf{0} \in \mathbb{R}^m$ just as any vector in $\mathbf{C}(A)$.

1.6 Rank

As simple as it seems, the rank of a matrix is one of the most important features of a matrix.

The rank of a matrix is the number of its independent columns.

A column of the matrix $\mathbf{v}_j$ is independent if it can not be written as the linear combination of the previous columns.
An important observation is that a matrix $A \in \mathbb{R}^{m \times n}$ (a matrix that has n columns) can only have

$$0 \leq \mathbf{rank}(A) \leq n$$

1.7 The Transpose

There are many ways to think about the transpose of a matrix:

- You get the transpose of a matrix by mirroring the entries along the diagonal.
- The entry in a_{ij} of a matrix A is swapped with a_{ji}.
- Transposing a matrix interchanges columns with rows.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 9 \\ 4 & 7 \end{bmatrix} \qquad A^\top = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 9 & 7 \end{bmatrix}$$

As the vectors are also matrices, we also have transposes of vectors. With this fact, it is easier to think about the transpose of a matrix in terms of column and vector notations.

$$A = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} \quad \Leftrightarrow \quad A^\top = \begin{bmatrix} — & \mathbf{v}_1^\top & — \\ — & \mathbf{v}_2^\top & — \\ & \vdots & \\ — & \mathbf{v}_n^\top & — \end{bmatrix},$$

$$A = \begin{bmatrix} — & \mathbf{u}_1 & — \\ — & \mathbf{u}_2 & — \\ & \vdots & \\ — & \mathbf{u}_m & — \end{bmatrix} \quad \Leftrightarrow \quad A^\top = \begin{bmatrix} | & | & & | \\ \mathbf{u}_1^\top & \mathbf{u}_2^\top & \cdots & \mathbf{u}_m^\top \\ | & | & & | \end{bmatrix}.$$

The transpose in column and row notations, form lecture notes p.46

Important Notes on the Transpose:

Mirroring tow times gives us the initial matrix: $(A^\top)^\top = A$.
A square matrix is symmetric *if and only if* $A = A^\top$. (use this for proofs/calculations)

1.8 Row Space

Row space of a matrix is the span of its rows. Since the rows of A are the columns of $A^\top$ we can define the row space of a as follows:

$$\mathbf{R}(A) := \mathbf{C}(A^\top)$$

Similarly we define the *row rank* of a matrix as the number of its independent rows, or equivalently the number of the independent columns of its transpose.

Note that for all matrices the following relation holds:

$$\boxed{\text{row rank} = \text{column rank}}$$

Which we are going to proof later.

2 Matrix Multiplication

This operation is one of the most important operations of today's world. People are excited about multiplying matrices and have been doing research on how to multiply them faster. Matrix multiplication can be seen as several linear combinations of the columns of a matrix.

2.1 How to multiply matrices?

There are many ways to think about how to multiply two matrices with each other. Here are a few of them illustrated, from the lecture notes:

* * *

Definition 2.16 (Matrix multiplication). *Let A be an $a \times n$ matrix and*

$$B = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_b \\ | & | & & | \end{bmatrix}$$

an $n \times b$ matrix. The $a \times b$ matrix

$$AB := \begin{bmatrix} | & | & & | \\ A\mathbf{x}_1 & A\mathbf{x}_2 & \cdots & A\mathbf{x}_b \\ | & | & & | \end{bmatrix}$$

is the product *of A and B.*

Observation 2.17. *Let*

$$A = \begin{bmatrix} — & \mathbf{u}_1 & — \\ — & \mathbf{u}_2 & — \\ & \vdots & \\ — & \mathbf{u}_a & — \end{bmatrix} \in \mathbb{R}^{a \times n}, \quad B = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_b \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{n \times b}.$$

Then

$$AB = \underbrace{\begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{x}_1 & \mathbf{u}_1 \cdot \mathbf{x}_2 & \cdots & \mathbf{u}_1 \cdot \mathbf{x}_b \\ \mathbf{u}_2 \cdot \mathbf{x}_1 & \mathbf{u}_2 \cdot \mathbf{x}_2 & \cdots & \mathbf{u}_2 \cdot \mathbf{x}_b \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_a \cdot \mathbf{x}_1 & \mathbf{u}_a \cdot \mathbf{x}_2 & \cdots & \mathbf{u}_a \cdot \mathbf{x}_b \end{bmatrix}}_{ab \text{ scalar products}} = [\mathbf{u}_i \cdot \mathbf{x}_j]_{i=1,j=1}^{a \quad b} \in \mathbb{R}^{a \times b}.$$

And also

$$\underbrace{\begin{bmatrix} — & \mathbf{u}_1 B & — \\ — & \mathbf{u}_2 B & — \\ & \vdots & \\ — & \mathbf{u}_m B & — \end{bmatrix}}_{AB\text{, row notation}} = \underbrace{\begin{bmatrix} — & \mathbf{u}_1 & — \\ — & \mathbf{u}_2 & — \\ & \vdots & \\ — & \mathbf{u}_m & — \end{bmatrix}}_{A\text{, row notation}} \underbrace{\begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ | & | & & | \end{bmatrix}}_{B\text{, column notation}} = \underbrace{\begin{bmatrix} | & | & & | \\ A\mathbf{x}_1 & A\mathbf{x}_2 & \cdots & A\mathbf{x}_n \\ | & | & & | \end{bmatrix}}_{AB\text{, column notation}}$$

* * *

Important Notes on Matrix Multiplication

- The product of matrices A and B are only then defined when the dimensions of the two matrices match. AB is defined when A has so many columns as b has rows. In general $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times k}$
- Matrix multiplication is NOT commutative. $AB \neq BA$ in general. Sometimes one of the products might not even be defined if the dimensions do not match.
- *Matrix Multiplication and Transposition:* $(AB)^\top = B^\top A^\top$
- For the identity matrix $I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$ we have: $IA = AI = A$ for all matrices A (and therefore also for vectors).

2.2 Distributivity and Associativity

The matrix multiplication is distributive and associative. From the lecture notes:

Lemma 2.22. *Let A, B, C be three matrices such that all sums and products in the following are defined. Then*

(i) $A(B+C) = AB + AC$ *and* $(B+C)D = BD + CD$ (distributivity);

(ii) $(AB)C = A(BC)$ (associativity).

3 Hints

1. In-Class
2. Lemma 2.21 can help. Besides, think about what happens when $\mathbf{x}$ is orthogonal to $\mathbf{w}$?
3. a) direct calculation, b) A and B are commuting. Make them commute in the product of k AB's, c) use b), d) Use distributivity of matrix multiplication, e) Use induction
4. First, prove that if a vector $\mathbf{v}$ is orthogonal to both $\mathbf{x}$ and $\mathbf{y}$ then it is also orthogonal to the vector $(\lambda\mathbf{x} + \mu\mathbf{y})$.
5. No hints.
6. No hints.
7. Be careful with the dimensions.

mkilic

♠