Linear Algebra
Week 2

G-17

2 X 2025

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

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: $\lambda \cdot A$

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

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:

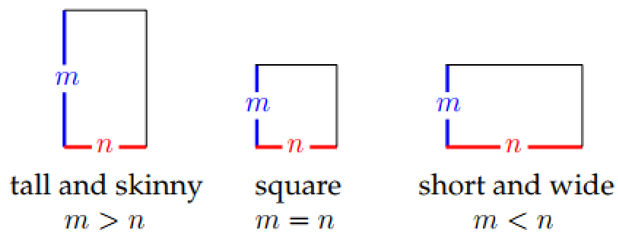


Figure 2.1. from the lecture notes

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.

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 4 & 7 \\ 0 & 0 & 5 \end{bmatrix}$	$\begin{bmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 7 & 5 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & 7 \\ 0 & 7 & 5 \end{bmatrix}$
identity matrix	diagonal matrix	upper triangular matrix	lower triangular matrix	symmetric matrix

Examples from lecture notes p.48

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:

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

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

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

Observation 2.5 Let A be an $m \times n$ matrix.

(i) A vector $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of the columns of A if and only if there is a vector $\mathbf{x} \in \mathbb{R}^n$ (of suitable scalars) such that $A\mathbf{x} = \mathbf{b}$.

(ii) The columns of A are linearly independent if and only if $\mathbf{x} = 0$ is the only vector such that $A\mathbf{x} = 0$.

You can also view the matrix vector multiplication as the scalar products of the rows of the matrix with the given vector.

Observation 2.8 Matrix-vector multiplication with A in row notation Let

$$A = \begin{bmatrix} - & \mathbf{u}_1^\top & - \\ - & \mathbf{u}_2^\top & - \\ & \vdots & \\ - & \mathbf{u}_m^\top & - \end{bmatrix} \in \mathbb{R}^{m \times n}, \qquad \mathbf{x} \in \mathbb{R}^n. \text{ Then } A\mathbf{x} = \underbrace{\begin{pmatrix} \mathbf{u}_1^\top x \\ \mathbf{u}_2^\top x \\ \vdots \\ \mathbf{u}_m^\top x \end{pmatrix}}_{m \text{ scalar products}}.$$

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.

$$\boxed{\text{The rank of a matrix is the number of its independent columns.}}$$

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

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

Note: $rank(A) = 0$ only if $A = 0$, the zero matrix.

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.

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

Remember 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.54

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 Nullspace

Nullspace is one of the fundamental spaces associated with a matrix. It consists of all the vectors that result in the zero vector $\mathbf{0}$ when multiplied with a matrix A. By this definition, $\mathbf{0}$ vector is always in the Nullspace of any matrix. Formally:

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

An important observation is that if an $m \times n$ matrix has $\mathbf{N}(A) = \{\mathbf{0}\}$, *ie* $\mathbf{0}$ is the only vector in the Nullspace, then the columns of the matrix are linearly independent according to *Observation 2.5*.

3 Linear Transformations

Linear Transformations are one of the fundamental concepts of linear algebra. Under the hood matrices are tools to represent linear transformations. At this point we should consider matrices as functions.

3.1 Matrices as Functions

Matrices are functions! They are defined as follows in the lecture notes:

Definition 2.18 (Matrix transformation) *Let A be an $m \times n$ matrix. $T_A : \mathbb{R}^n \mapsto \mathbb{R}^m$ is the function where $\boldsymbol{x} \in \mathbb{R}^n$ and $A\boldsymbol{x} \in \mathbb{R}^m$ defined by:*

$$T_A(\mathbf{x}) = A\mathbf{x}$$

This shows that every matrix defines a linear transformation.

There is an important set of observations about matrices as functions, that show that these transformations are linear:
Observation Let A be an $m \times n$ matrix, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. Then

1. $A(\mathbf{x} + \mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y})$ and
2. $A(\lambda\mathbf{x}) = \lambda A(\mathbf{x})$.

By combining (i) and (ii), we also get

$$T_A(\lambda x + \mu y) = \lambda T_A(x) + \mu T_A(y).$$

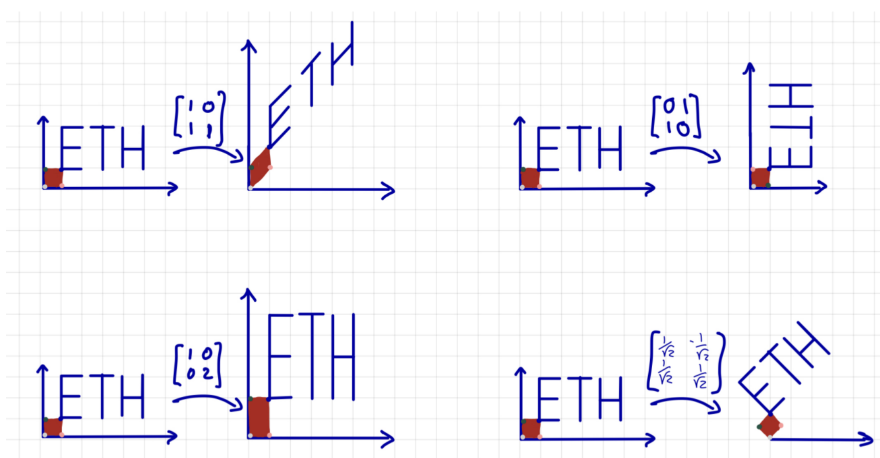
More formally:

Lemma 2.19 Let A be an $m \times n$ matrix, $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ and $\lambda_1, \lambda_2 \in \mathbb{R}$. Then

$$A(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2) = \lambda_1 A\mathbf{x}_1 + \lambda_2 A\mathbf{x}_2.$$

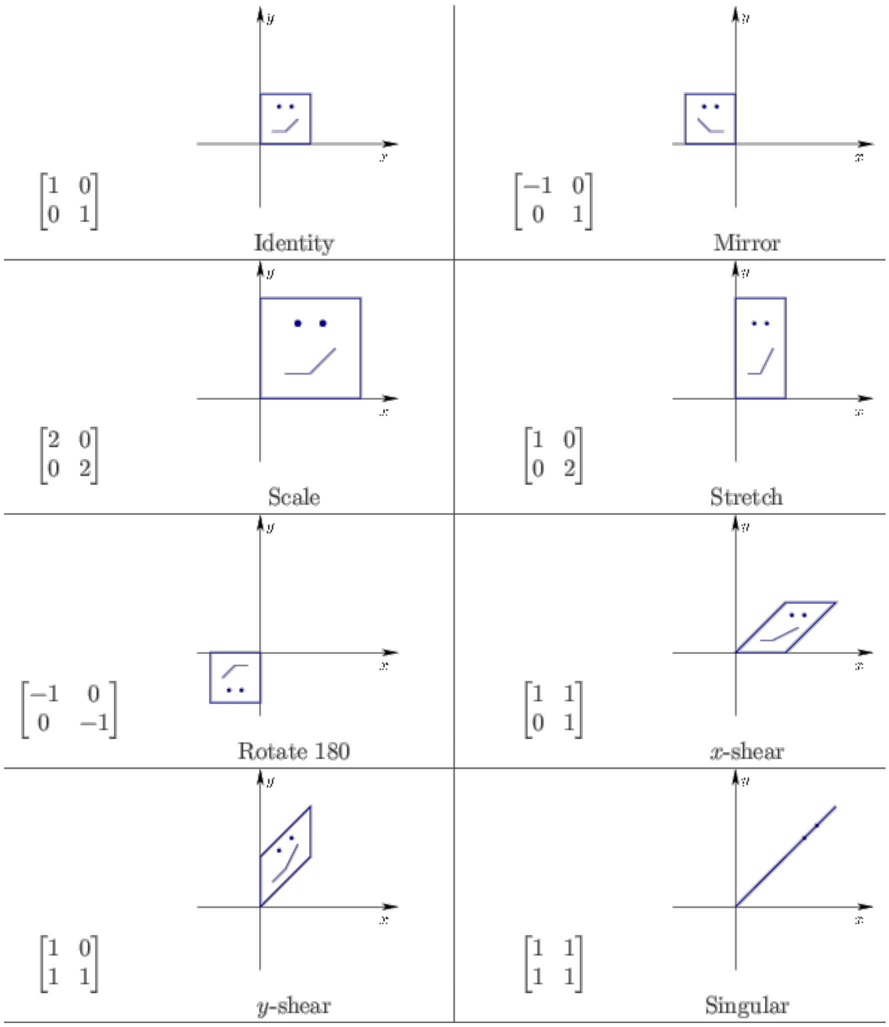
To understand in what ways a matrix transforms a vector we are specifically interested in the standard unit vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. To intuitively sum up why, it is because we can express everything in a space $\mathbb{R}^n$ using the set of standard unit vectors. If we know what A does to this set of vectors, we can generalize the effect of A to all vectors in $\mathbb{R}^n$.

This is demonstrated below. If you know what happens to the unit square (little red guy), then you can generalize that effect to any other shape or *e.g.* to any set of vectors.



Example from HS23

Certain operations have special names that intuitively describe the operation a matrix does on a given object.



Special Linear Transformations

4 Hints

1. In-Class
2. Section 2.1.4 from the lecture notes is helpful.
3. a) You can try to visualize this on unit vectors in $\mathbb{R}^3$. You can also first consider the matrix
$$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$
4. No hints.
5. No hints.
6. d) Write the matrix $A \in \mathbb{R}^{3\times 3}$ in a generic way. Remember that if you can find one vector other than $\mathbf{0}$ with $A\mathbf{x} = \mathbf{0}$, then you can prove $\mathbf{rank}(A) \leq 2$.
7. No hints.

mkilic

♠