# LINEAR ALGEBRA: DEFINITIONS AND FACTS

SPRING, 2026

## Week 1

**Definitions:**

(1) A **matrix** is a rectangular array of numbers made up of **rows** which are horizontal lists and **columns** which are vertical lists. The individual numbers in the matrix are called **entries**. The notation $a_{ij}$ is used to indicate the entry which is in the $i$th row and the $j$th column.
(2) The **size** of a matrix $A$ is $n \times m$ where $n$ is the number of rows in $A$ and $m$ is the number of columns. A matrix is called **square** if the number of rows is equal to the number of columns.
(3) A square matrix with entries $a_{ij}$ is called **diagonal** $a_{ij} = 0$ unless $i = j$. A square matrix is called **upper (or lower) triangular** if all entires on or below (above) the diagonal entries are equal to zero. If all the entries of a matrix are zero, it is called a **zero matrix**.
(4) A matrix with a single row is called a **row vector**.
(5) A matrix with a single column is called a **column vector** or simply **vector**.
(6) The set of all column vectors with $n$ components is denoted by $\mathbb{R}^n$. We will refer to $\mathbb{R}^n$ as a **vector space**.
(7) A **component** is an entry in a row or column vector.
(8) The $n \times n$ matrix whose diagonal entries are equal to 1 and whose off-diagonal entries are all zero is called the **n by n identity matrix** and is denoted by $I_n$.
(9) A matrix is in **reduced row-echelon form (rref)** if it satisfies the following conditions.
    (a) If a row has nonzero entries, then the first nonzero entry is a 1. Such a 1 is called a **leading 1** or **pivot**.
    (b) If a column contains a leading 1, then all other entries in the column are 0.
    (c) Each leading 1 is to the *right* of any leading 1s in the row above.
    (d) If there exists a row of all 0s, it appears as the last row of the matrix. (This condition is implied by the others).
(10) The following three operations on rows are called **elementary row operations**.
    (a) Divide a row by a nonzero scalar (number).
    (b) Subtract a multiple of a row from another row.
    (c) Swap two rows.
(11) A **linear equation** is an equation of the form

$$y = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

where each $a_i$ is a number, and where $x_i, y$ are variables.
(12) A **system of linear equations** is a collection of 2 or more linear equations.
(13) A system of linear equations is called **consistent** if it has *at least one* solution. A system is called **inconsistent** if it has *no solutions.*
(14) Given a system of $n$ equations in $m$ variables, that is,

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m \\ \vdots &= \vdots + \vdots + \cdots + \vdots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m, \end{aligned}$$

the matrix

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}$$

is called the **coefficient matrix** for the system. The matrix

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} & | & y_1 \\ a_{21} & a_{22} & \dots & a_{2m} & | & y_2 \\ & \vdots & \vdots & \ddots & | & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & | & y_n \end{bmatrix}$$

is called the **augmented matrix** for the system and is often denoted by $[A|\vec{y}]$.

(15) Note that the coefficient matrix of a system of linear equations in $m$ variables has $m$ columns, one for each variable. Let $A$ be the coefficient matrix of such a system. The variable $x_i$ is called **a free variable** if the $i$th column in $\mathrm{rref}(A)$ *does not contain a leading 1.*

(16) **Gauss–Jordan elimination** is an algorithm that transforms a matrix into reduced row echelon form through a finite sequence of elementary row operations. The algorithm is as follows.
  (a) Form the augmented matrix of the system of linear equations.
  (b) Beginning with the leftmost column, locate a nonzero entry and interchange rows if necessary to place it in the highest available row; this entry is the pivot.
  (c) Scale the pivot row so that the pivot entry equals 1.
  (d) Use row replacement operations to make all other entries in the pivot column equal to 0, both above and below the pivot.
  (e) Move to the next column to the right and the next row down, and repeat Steps 2–4 until no further pivots can be found.
  (f) Continue until the matrix is in reduced row echelon form, where each pivot column contains a single 1 and zeros elsewhere.

**Facts:**

(1) Linear equations represent lines, planes or their higher-dimensional analogues.
(2) A solution to a system of linear equations (if it exists) represents a point in the intersection of the lines or planes described by each equation in the system.
(3) The reduced row echelon form of a matrix does *not* depend on the sequence of elementary row operations used.

## Week 2

**Definitions:**

(1) The **rank** of a matrix $A$ is the number of leading 1s in its reduced row-echelon form (which is uniquely determined by $A$).
(2) Suppose that $\vec{v}$ and $\vec{w}$ are two vectors with the same number of components. Suppose that the components of $\vec{v}$ are $v_1, \dots, v_n$ and the components of $\vec{w}$ are $w_1, \dots, w_n$. The **dot product** of $\vec{v}$ and $\vec{w}$ is defined to be the scalar

$$\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 + \dots + v_n w_n.$$

**Note:** this operation is defined even when $\vec{v}$ and $\vec{w}$ are vectors of different types. That is either can be either a column of row vector.

(3) If $A$ is an $n \times m$ matrix with row vectors $\vec{w}_1, \dots, \vec{w}_n$, and $\vec{x}$ is a vector in $\mathbb{R}^n$, then the **product of the matrix $A$ with the column vector $\vec{x}$** is defined to be

$$A\vec{x} = \begin{bmatrix} -\vec{w}_1- \\ \vdots \\ -\vec{w}_n- \end{bmatrix} \vec{x} = \begin{bmatrix} \vec{w}_1 \cdot \vec{x} \\ \vdots \\ \vec{w}_n \cdot \vec{x} \end{bmatrix}.$$

(4) A vector $\vec{b} \in \mathbb{R}^n$ is called a **linear combination** of the vectors $\vec{v}_1, \ldots, \vec{v}_m$ in $\mathbb{R}^n$ if there exist scalars $x_1, \ldots, x_n$ such that

$$\vec{b} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n$$

**Facts:**

(1) If a linear system is consistent, then it has either
- Infinitely many solutions (in this case, there is at least one free variable) or
- exactly one solution (in this case, all the variables are leading (correspond to pivot rows).

(2) Suppose that $A$ is the an $n \times m$ matrix representing a system of $n$ linear equations in $m$ variables. Then
- $\text{rank}(A) \leqslant n$ and $\text{rank}(A) \leqslant m$.
- If $\text{rank}(A) = n$ then the system is consistent.
- If $\text{rank}(A) = m$, then the system has at most one solution.
- If $\text{rank}(A) < m$, then the system has either infinitely many solutions, or none.

(3) A linear system with fewer equations than unknowns has either no solutions or infinitely many solutions.

(4) A linear system of $n$ equations in $n$ variables has a unique solution iff the rank of its coefficient matrix $A$ is $n$. In this case, $\text{rref}(A) = I_n$.

(5) If the column vectors of an $n \times m$ matrix $A$ are $\vec{v}_1, \ldots, \vec{v}_m$ and $\vec{x}$ is a vector in $\mathbb{R}^m$ with components $x_1, \ldots, x_m$, then

$$A\vec{x} = \begin{bmatrix} | & & | \\ \vec{v}_1 & \ldots & \vec{v}_m \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = x\vec{v}_1 + \cdots + x_m\vec{v}_m.$$

(6) If $A$ is an $n \times m$ matrix, $\vec{x}, \vec{y} \in \mathbb{R}^m$, and $k \in \mathbb{R}$, then
- $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$.
- $A(k\vec{x}) = k\,(A\vec{x})$.

## Week 3

**Definitions**

(1) Consider two sets $X$ and $Y$. A **function** $T$ from $X$ to $Y$ is a rule that associates with each element $x \in X$ a unique element $y \in Y$. The set $X$ is called the **domain** or **input** of the function $T$. The set $Y$ is called the **target space, codomain,** or **output** of $T$.

(2) A function $T : \mathbb{R}^m \to \mathbb{R}^n$ is called a **linear transformation** if there exists an $n \times m$ matrix $A$ such that

$$T(\vec{x}) = A\vec{x}$$

for all $\vec{x}$ in the vector space $\mathbb{R}^m$.

(3) **Matrix Multiplication:**
  (a) Let $B$ be an $n \times p$ matrix and $A$ a $q \times m$ matrix. The **product** $BA$ is defined if and only of $p = q$.
  (b) If $B$ is an $n \times p$ matrix and $A$ is a $p \times m$ matrix, then the product $BA$ is defined as the matrix of the linear transformation $T(\vec{x}) = B\,(A\vec{x})$.
  (c) Suppose that $A$ and $B$ are two matrices for which $AB$ and $BA$ are both defined. If $AB = BA$ then we say that $A$ **commutes** with $B$.

(4) A function $T : X \to Y$ is called **invertible** if the equation $T(x) = y$ has a unique solution $x \in X$ for each $y \in Y$. In this case, the **inverse** $T^{-1} : Y \to X$ is defined by

$$T^{-1}(y) = \text{the unique } x \in X \text{ such that } T(x) = y.$$

**Facts**

(1) Consider a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$. Then the matrix of $T$ is
$$A = \begin{bmatrix} | & | & & | \\ T(\vec{e}_1) & T(\vec{e}_2) & \dots & T(\vec{e}_m) \\ | & | & & | \end{bmatrix}$$
where $\vec{e}_i$ is the vector in $\mathbb{R}^m$ whose components are all zero except for the $i$th component which is equal to 1.

(2) A transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ is linear iff
- $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$ for all vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^m$, and
- $T(k\vec{v}) = kT(\vec{v})$ for all $\vec{v} \in \mathbb{R}^m$ and $k \in \mathbb{R}$.

(3) In general, it is *not the case* that $AB = BA$. That is, matrix multiplication is **non-commutative**.

(4) For any $n \times m$ matrix $A$,
$$AI_m = I_n A = A.$$

(5) Matrix multiplication satisfies the following properties
  (a) $(AB)C = A(BC)$ (associativity)
  (b) $A(C + D) = AC + AD$ and $(A + B)C = AC + BC$ (distributivity)
  (c) $(kA)B = A(kB) = k(AB)$ (commutativity of scalar multiplication).

(6) Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ which rotates a vector $\vec{v}$ *counterclockwise* by an angle $\theta$. The matrix which represents this rotation is
$$R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$
To rotate in the *clockwise* direction, multiply the vector by
$$R_{-\theta} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}.$$

(7) Suppose a line $L$ passes through the origin and makes an angle $\theta$ with the $x$-axis. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which reflects a vector $\vec{v}$ over the line $L$. The matrix which represents this reflection is
$$B_\theta = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}.$$

## Week 4

### Definitions

(1) A function $T : X \to Y$ is called **invertible** if the equation $T(x) = y$ has a unique solution $x \in X$ for each $y \in Y$. In this case, the **inverse** $T^{-1} : Y \to X$ is defined by
$$T^{-1}(y) = \text{the unique } x \in X \text{ such that } T(x) = y.$$

(2) A square matrix $A$ is called **invertible** if the linear transformation $T$ defined by $T(\vec{x}) = A\vec{x}$ is invertible.

(3) Suppose that $A$ is an invertible square matrix. Then the **inverse matrix of** $A$ denoted by $A^{-1}$ is the unique square matrix which satisfies
$$AA^{-1} = A^{-1}A = I_n.$$
For $T(\vec{x}) = A\vec{x}$, we also have $T^{-1}(\vec{x}) = A^{-1}\vec{x}$.

(4) Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The quantity $ad - bc$ is called the **determinant** of the matrix $A$.

(5) The **image** of a function $f : X \to Y$ is
$$\text{image}(f) = \{b \in Y \mid b = f(x) \text{ for some } x \in X\}.$$
Similarly, the **image of an n by m matrix** $A$ is
$$\text{image}(A) = \{\vec{y} \in \mathbb{R}^n \mid \vec{y} = A\vec{x} \text{ for some } \vec{x} \in \mathbb{R}^m\}.$$

(6) Consider a set of vectors $\{v_1, \ldots, v_m\}$ in $\mathbb{R}^n$. The set of *all* linear combinations $c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$ is called the **span** of this set. We write

$$\text{span}\,\{v_1, \ldots, v_m\} = \left\{ \sum_{i=1}^{m} c_i\vec{v}_i \mid c_i \in \mathbb{R}, \vec{v}_i \in \{\vec{v}_1, \ldots, \vec{v}_m\} \right\}.$$

(7) The **kernel** or **nullspace** of a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ is

$$\ker(T) = \left\{ \vec{x} \in \mathbb{R}^m \mid T(\vec{x}) = \vec{0} \in \mathbb{R}^n \right\}.$$

Similarly, the **kernel** of an $n \times m$ matrix $A$ is

$$\ker(A) = \left\{ \vec{x} \in \mathbb{R}^m \mid A\vec{x} = \vec{0} \in \mathbb{R}^n \right\}.$$

**Facts**

(1) Let $A$ be an $n \times n$ matrix. TFAE (The following are equivalent).
- $A$ is invertible
- $\text{rref}(A) = I_n$
- $\text{rank}(A) = n$.
- $A\vec{x} = 0$ has exactly one solution (which turns out to be $\vec{x} = \vec{0}$).
- $A\vec{x} = \vec{b}$ has exactly one solution (which turns out to be $A^{-1}\vec{b}$).

(2) Let $A$ be an $n \times n$ matrix that is *not* invertible. Then $A\vec{x} = \vec{b}$ has infinitely many solutions or none.

(3) To find the *inverse* of an $n \times n$ matrix, compute $\text{rref}[A|I_n]$.
- If $\text{rref}\,[A|I_n]$ is of the form $[I_n|B]$ then $A$ is invertible and $A^{-1} = B$.
- If $\text{rref}\,[A|I_n]$ is of any other form, $A$ is not invertible.

(4) If $A$ and $B$ are invertible $n \times n$ matrices then $BA$ is invertible as well and

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

(5) If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible, then

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

(6) The image of a linear transformation $T(\vec{x}) = A\vec{x}$ is the span of the column vectors of $A$.

(7) Suppose that $A$ is an $n \times m$ matrix. Then

$$\ker A = \left\{\vec{0}\right\} \iff \text{rank}(A) = m$$

(8) Suppose that $A$ is $n \times n$. Then

$$\ker A = \left\{\vec{0}\right\} \iff A \text{ is invertible}$$

## Week 5

**Definitions**

(1) A set $W \subseteq \mathbb{R}^n$ is said to be **closed under addition** if

$$\vec{w}_1, \vec{w}_2 \in W \quad \Longrightarrow \quad \vec{w}_1 + \vec{w}_2 \in W.$$

(2) A set $W \subseteq \mathbb{R}^n$ is said to be **closed under scalar multiplication** if

$$\vec{w} \in W \quad \Longrightarrow \quad k\vec{w} \in W \;\; \forall\, k \in \mathbb{R}.$$

(3) A subset $W$ of $\mathbb{R}^n$ is called a **linear subspace** if
   (a) $\vec{0} \in W$.
   (b) $W$ is closed under addition
   (c) $W$ is closed under scalar multiplication.

(4) Consider the set of vectors $\{\vec{v}_1, \ldots, \vec{v}_m\}$ in $\mathbb{R}^n$.
   (a) $\vec{v}_i$ in the list $\vec{v}_1, \ldots, \vec{v}_k$ is called **redundant** if $\vec{v}_i$ is a linear combination of the preceding vectors $\vec{v}_1, \ldots, \vec{v}_{i-1}$.

   (b) the vectors $\vec{v}_1, \ldots, \vec{v}_m$ are called **linearly independent** if none of them is redundant.
(5) The set of vectors $\{\vec{v}_1, \ldots, \vec{v}_m\}$ form a **basis** of a subspace $V$ of $\mathbb{R}^n$ if
   - $V = \operatorname{span}\{\vec{v}_1, \ldots, \vec{v}_m\}$ and
   - $\{\vec{v}_1, \ldots, \vec{v}_m\}$ are linearly independent.
(6) Consider the set of vectors $\{\vec{v}_1, \ldots, \vec{v}_m\}$. An equation of the form

$$c_1\vec{v}_1 + \cdots + c_m\vec{v}_m = \vec{0}$$

where $c_i \in \mathbb{R}$ is called a **relation** among the vectors $\vec{v}_1, \ldots, \vec{v}_m$. A relation is called **trivial** if $c_1 = \cdots = c_m = 0$ and **non-trivial** if it is not trivial.
(7) The number of elements in a basis of a linear subspace $V \subseteq \mathbb{R}^n$ is called the **dimension** of $V$.
(8) (Optional) Suppose that $W$ is an **affine space**. That is, each element $\vec{w} \in W$ is of the form $\vec{w} = \vec{x}_0 + \vec{v}$ for some fixed $\vec{x}_0$ and some $\vec{v}$ in a vector space $V$. Then the **dimension** of $W$ is just the dimension of $V$.

**Facts**

(1) The image of a linear transformation $T(\vec{x}) = A\vec{x}$ is the span of the column vectors of $A$.
(2) The image of a linear transformation $T : R^m \to \mathbb{R}^n$ has the following properties
   (a) $\vec{0} \in \operatorname{image} T$
   (b) $\operatorname{image} T$ is *closed under addition.*
   (c) $\operatorname{image} T$ is *closed under scalar multiplication.*
   Thus $\operatorname{image} T$ is a *subspace* of $\mathbb{R}^n$.
(3) Consider a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$.
   (a) $\vec{0} \in \ker T$.
   (b) $\ker T$ is closed under addition
   (c) $\ker T$ is closed under scalar multiplication.
   Thus $\ker T$ is a subspace of $\mathbb{R}^m$.
(4) All bases of a subspace $V \subseteq \mathbb{R}^n$ have the same number of elements.
(5) Consider an $m$-dimensional subspace $V$ of $\mathbb{R}^n$.
   (a) One can find *at most* $m$ linearly independent vectors in $V$.
   (b) We need *at least* $m$ vectors to span $V$.
   (c) If $m$ vectors in $V$ are linearly independent, then they form a basis of $V$.
   (d) If $m$ vectors in $V$ span $V$ then they form a basis for $V$.
(6) For any matrix $A$,

$$\dim \operatorname{image}(A) = \operatorname{rank}(A)$$

(7) **Rank–Nullity Theorem:** For any $n \times m$ matrix $A$,

$$\dim \ker A + \dim \operatorname{image} A = m.$$

In other words,

$$\text{nullity of } A + \operatorname{rank} A = m.$$

(8) The vectors $\{\vec{v}_1, \ldots, \vec{v}_m\}$ form a basis of $\mathbb{R}^n$ iff the matrix

$$\begin{bmatrix} | & & | \\ \vec{v}_1 & \ldots & \vec{v}_n \\ | & & | \end{bmatrix}$$

is invertible.

## Week 6

**Definitions**

(1) Consider a basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_m\}$ of a subspace $V \subseteq \mathbb{R}^n$. Then any vector $\vec{x} \in V$ can be written uniquely as

$$\vec{x} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m.$$

The scalars $c_1, c_2, \dots, c_m$ are called the **$\mathcal{B}$-coordinates** of $\vec{x}$ and the vector

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

is called the **$\mathcal{B}$-coordinate vector of** $\vec{x}$ and is denoted by $[\vec{x}]_{\mathcal{B}}$.

(2) Consider a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ and a basis $\mathcal{B}$ of $\mathbb{R}^n$. The $n \times n$ matrix $B$ satisfying

$$[T(\vec{x})]_{\mathcal{B}} = B\,[\vec{x}]_{\mathcal{B}}$$

for all $\vec{x} \in \mathbb{R}^n$ is called the **$\mathcal{B}$-matrix** of $T$.

(3) Let $V$ be an $m$-dimensional subspace of $\mathbb{R}^n$, and let $\mathcal{B} = \{b_1, \dots, b_m\}$ and $\mathcal{C} = \{c_1, \dots, c_n\}$ be bases of $V$. The *change of basis matrix from $\mathcal{C}$ to $\mathcal{B}$*, denoted $S_{\mathcal{B}\leftarrow\mathcal{C}}$, is the unique $n \times n$ matrix such that

$$[\vec{v}]_{\mathcal{B}} \;=\; S_{\mathcal{B}\leftarrow\mathcal{C}}\,[\vec{v}]_{\mathcal{C}} \quad \text{for all } \vec{v} \in V.$$

(4) Two matrices $A, B$ are said to be **similar** if there exists an invertible matrix $M$ such that

$$A = MBM^{-1}.$$

(5) Suppose that $W \subseteq \mathbb{R}^n$ is a subspace. It's *orthogonal complement* is the set $W^\perp$ defined by

$$W^\perp = \{\vec{v} \in \mathbb{R}^n \mid \vec{v} \cdot \vec{w} = 0 \;\forall\, \vec{w} \in W\}$$

**Facts**

(1) Let $W \subseteq \mathbb{R}^n$ be a subspace. Then $\dim W^\perp = n - \dim W$.
(2) Let $A$ be an $m \times n$ matrix. Then $(\ker A)^\perp = \operatorname{Im}\left(A^\top\right)$.
(3) Let $V$ be a linear subspace with two given bases $\mathcal{B}$ and $\mathcal{B}'$. Consider a linear transformation $T : V \to V$ and let $A, B$ be the $\mathcal{B}$- and $\mathcal{B}'$-matrix of $T$, respectively. Let $S$ be the change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$. Then $A$ is similar to $B$ and

$$AS = SB \iff A = SBS^{-1} \iff B = S^{-1}AS.$$

.

## Weeks 9 and 10

**Definitions**

(1) A **linear space** (or **vector space**) $V$ is a set endowed with a rule for addition ($f, g \in V \Rightarrow f + g \in V$) and a rule for scalar multiplication ($f \in V \Rightarrow kf \in V \;\forall\, k \in \mathbb{R}$) such that (for all $f, g, h \in V$ and $c, k \in \mathbb{R}$) these operations satisfy the following 8 rules:
  (a) $(f + g) + h = f + (g + h)$
  (b) $f + g = g + f$
  (c) $\exists!$ **neutral element** $n \in V$ such that $f + n = f \;\forall\, f \in V$. This element is denoted by 0.
  (d) For each $f \in V, \exists! g \in V \mid f + g = 0$. This $g$ is denoted by $-f$.
  (e) $k(f + g) = kf + kg$.
  (f) $(c + k)f = cf + kf$
  (g) $c(kf) = (ck)f$
  (h) $1f = f$.
(2) A subset $W$ of a linear space $V$ is called a **linear subspace** if
  (a) $\vec{0} \in W$.
  (b) $W$ is closed under addition
  (c) $W$ is closed under scalar multiplication.
(3) A function $T : V \to W$ between linear spaces is called a **linear transformation** if
  - $T(f + g) = T(f) + T(g)$ for all $f, g \in V$.
  - $T(kf) = kT(f)$ for all $f \in V$ and all $k \in \mathbb{R}$.

(4) An invertible linear transformation is called an **isomorphism**.
(5) Two matrices $A, B$ are said to be **similar** if there exists an invertible matrix $M$ such that
$$A = MBM^{-1} \qquad B = M^{-1}AM.$$
(6) Consider two bases $\mathscr{B}, \mathscr{B}'$ of an $n$-dimensional linear space $V$. Consider the linear transformation $L_{\mathscr{B}} \circ L_{\mathscr{B}'}^{-1}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$. Let $S\vec{x} = \left(L_{\mathscr{B}'} \circ L_{\mathscr{B}'}^{-1}\right)(\vec{x})$. Then
$$[f]_{\mathscr{B}} = S\,[f]_{\mathscr{B}'}, \qquad \text{for all} f \in V.$$
This matrix $S$ is called the **change of basis matrix**.
(7) A set of vectors $\{\vec{v}_1, \dots, \vec{v}_k\}$ is called **orthonormal** if
- $\vec{v}_i \cdot \vec{v}_j = 0$ for all $i \neq j$ (any two vectors in the set are orthogonal) and
- $|\vec{v}_i| = 1$ for all $i$ (every vector in the set has length 1).

**Facts**

(1) Let $V$ be a linear subspace with two given bases $\mathscr{B}$ and $\mathscr{B}'$. Consider a linear transformation $T : V \to V$ and let $A, B$ be the $\mathscr{B}$- and $\mathscr{B}'$-matrix of $T$, respectively. Let $S$ be the change of basis matrix from $\mathscr{B}$ to $\mathscr{B}'$. Then $A$ is similar to $B$ and
$$AS = SB \iff A = SBS^{-1} \iff B = S^{-1}AS.$$
.
(2) Suppose that $V$ and $V'$ are two vector spaces with the same (finite) dimension. Then there exists an isomorphism $T : V \to V'$. The converse also holds. That is, if there is an isomorphism $T : V \to V'$ then $\dim V = \dim V'$
(3) Suppose that $W$ is a subspace of $\mathbb{R}^n$. Let $\{\vec{w}_1, \dots, \vec{w}_k\}$ be an *orthogonal* basis for $W$. That is, $\vec{w}_i \cdot \vec{w}_j = 0$ for all $i \neq j$. Let $\vec{x} \in \mathbb{R}^n$ be any vector. Then the *orthogonal projection* of $\vec{x}$ onto $W$ is
$$\vec{x}_W = \sum_{i=1}^{k} \frac{\vec{x} \cdot \vec{w}_i}{\vec{w}_i \cdot \vec{w}_i} \vec{w}_i.$$
**Warning!** This formula is *only* valid if the basis for $W$ is *orthogonal.*
(4) If two linear spaces $V$ and $W$ have the same (finite) dimension, then they are isomorphic. That is, there exists an isomorphism $T : V \to W$. Conversely, if $T : V \to W$ is an isomorphism, then $V$ and $W$ have the same dimension.
(5) Consider a linear transformation $T : V \to V$. Let $B$ be a the matrix of $T$ with respect to a basis $\mathscr{B} = (f_1, \dots, f_n)$ of $V$. Then the $i$th column of $B$ is given by
$$[T(f_i)]_{\mathscr{B}}.$$

## Week 11

**Definitions**

(1) A set of vectors $\{u_1, \dots, u_m\}$ in $\mathbb{R}^n$ is called an **orthonormal basis** of a subspace $V$ if
- $u_i \cdot u_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$
- $V = \operatorname{span}\{u_1, \dots, u_m\}$.

(2) Let $V$ be a subspace of $\mathbb{R}^n$ with an orthonormal basis $\{u_1, u_2, \dots, u_m\}$. The **orthogonal projection** of a vector $x \in \mathbb{R}^n$ onto $V$ is defined by
$$\operatorname{proj}_V(x) = (x \cdot u_1)u_1 + (x \cdot u_2)u_2 + \cdots + (x \cdot u_m)u_m = \sum_{i=1}^{m} (x \cdot u_i)u_i.$$
The difference $x - \operatorname{proj}_V(x)$ lies in $V^\perp$, the orthogonal complement of $V$.
(3) Given a basis $\{v_1, \dots, v_m\}$ of a subspace $V \subseteq \mathbb{R}^n$, the vectors
$$u_1 = \frac{v_1}{\|v_1\|}, \qquad u_k = \frac{v_k - \sum_{j=1}^{k-1}(v_k \cdot u_j)u_j}{\|v_k - \sum_{j=1}^{k-1}(v_k \cdot u_j)u_j\|} \quad (k = 2, \dots, m)$$

form an orthonormal basis of $V$; this procedure is called the **Gram–Schmidt process**.

(4) If the columns of a matrix $M$ form a basis $\{v_1, \ldots, v_m\}$ of a subspace $V$ and $\{u_1, \ldots, u_m\}$ is the orthonormal basis obtained by the Gram–Schmidt process, then

$$M = QR,$$

where $Q$ is the $n \times m$ matrix whose columns are $u_i$ (orthonormal) and $R$ is the $m \times m$ upper-triangular matrix with positive diagonal entries giving the coordinates of $v_i$ in the $u$-basis. This decomposition is called the **QR factorization** of $M$.

(5) For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, the **length** of $x$ is

$$\|x\| \;=\; \sqrt{x \cdot x} \;=\; \sqrt{x_1^2 + \cdots + x_n^2}.$$

**Facts**

(1) Every vector $x \in \mathbb{R}^n$ can be uniquely written as

$$x = \mathrm{proj}_V(x) + w,$$

where $\mathrm{proj}_V(x) \in V$ and $w \in V^\perp$. The vector $\mathrm{proj}_V(x)$ is the closest point in $V$ to $x$.

(2) For any subspace $V \subseteq \mathbb{R}^n$ and any $x \in \mathbb{R}^n$,

$$\|x - \mathrm{proj}_V(x)\| \leqslant \|x - v\| \quad \text{for all } v \in V.$$

Equality holds only when $v = \mathrm{proj}_V(x)$.

(3) If the columns of $Q$ form an orthonormal basis for $\mathbb{R}^n$, then $Q^{-1} = Q^\top$.

(4) If the columns of $Q$ form an orthonormal basis of $V$, then

$$\mathrm{proj}_V(x) = QQ^T x, \qquad \text{and} \qquad \mathrm{proj}_{V^\perp}(x) = (I - QQ^T)x.$$

(5) The Gram–Schmidt process applied to any linearly independent set $\{v_1, \ldots, v_m\}$ produces an orthonormal set $\{u_1, \ldots, u_m\}$ with

$$\mathrm{span}\{v_1, \ldots, v_k\} = \mathrm{span}\{u_1, \ldots, u_k\}, \quad \forall\, k.$$

(6) If $A$ is an $n \times m$ matrix with linearly independent columns, then there exist matrices

$$Q = [\, u_1 \; u_2 \; \cdots \; u_m \,], \qquad R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ 0 & r_{22} & \cdots & r_{2m} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & r_{mm} \end{bmatrix},$$

such that

$$A = QR, \quad \text{where } Q^T Q = I_m \text{ and } r_{ii} > 0.$$

This decomposition is unique if the diagonal entries of $R$ are positive.

(7) In the above decomposition, for $i \leqslant j$, $r_{ij} = \vec{u}_i \cdot \vec{a}_j$ where $\vec{u}_i$ is a column vector of $Q$ and $\vec{a}_j$ is a column vector of $A$. For $i > j$, $r_{ij} = 0$.

(8) If $Q$ has orthonormal columns, then for all $x \in \mathbb{R}^m$,

$$\|Qx\| = \|x\|, \qquad (Qx) \cdot (Qy) = x \cdot y.$$

Thus, $Q$ preserves both lengths and angles.

## Week 12

**Definitions**

(1) The **determinant of a** $3 \times 3$ **matrix** $A = [\vec{u}\; \vec{v}\; \vec{w}]$, in terms of its columns, is

$$\det A = \vec{u} \cdot (\vec{v} \times \vec{w})\,.$$

In calculus, this is called the **scalar triple product**.

(2) A **pattern** in an $n \times n$ matrix $A$ is a choice of $n$ entries of the matrix so that there is one chosen entry in each row and one chosen entry in each column of $A$. Two entries in a pattern are sait to be **inverted** if one of them is located to the right and above the other in the matrix. The **signature** or **sign** of a pattern is defined as

$$\text{sgn } P = (-1)^{\text{number of inversions in } P}.$$

The **determinant** of an $n \times n$ matrix $A$ is defined as

$$\det A = \sum_P (\text{sgn } P)\,(\text{prod } P)\,.$$

Here, the sum is taken over *all possible patterns in* $A$ and $\text{prod}\, P$ is defined to be the product of all entries in $P$.

(3) Consider a linear transformation $T$ from $V$ to $V$, where $V$ is a finite-dimensional linear space. If $\mathscr{B}$ is a basis of $V$ and $A$ is the $\mathscr{B}$-matrix of $T$, then we define

$$\det T = \det A.$$

(4) Consider an invertible $n \times n$ matrix $A$. The **classical adjoint** $\text{adj}(A)$ is the $n \times n$ matrix whose $ij$th entry is $(-1)^{i+j} \det(A_{ji})$.

(5) A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is called **orthogonal** if

$$\|T(\vec{v})\| = \|\vec{v}\|.$$

If $T$ is an orthogonal transformation and $T(\vec{v}) = Q\vec{v}$, then $Q$ is called an **orthogonal matrix**.

**Facts**

(1) The determinant of a triangular matrix is the product of the diagonal entries of the matrix.

(2) $\det\left(A^\top\right) = \det A$

(3) If a matrix $B$ is obtained from a (square) matrix $A$ by...

(a) dividing a row of $A$ by a scalar $k$, then

$$\det B = (1/k)\det A$$

(b) a row swap, then

$$\det B = -\det A.$$

(c) adding a multiple of a row of $A$ to another row, then

$$\det B = \det A.$$

Warning! This one applies to row operations of the form

$$kR_i + R_j \to R_j.$$

Note that the "replaced row", $R_j$, does not get multiplied by a scalar.

(4) A square matrix is invertible if and only if $\det A \neq 0$.

(5) $\det(AB) = (\det A)\,(\det B)$.

(6) If a matrix $A$ is similar to $B$, then $\det A = \det B$.

(7) If $A$ is invertible, then

$$\det\left(A^{-1}\right) = \frac{1}{\det A} = (\det A)^{-1}.$$

(8) The determinant of an orthogonal matrix is either 1 or $-1$.

(9) If $A$ is an $n \times n$ matrix with columns $\vec{v}_1, \ldots \vec{v}_n$, then

$$|\det A| = \|\vec{v}_1\|\left\|\vec{v}_2^\perp\right\| \ldots \left\|\vec{v}_n^\perp\right\|$$

where $\vec{v}_k^\perp$ is the component of $\vec{v}_k$ perpendicular to $\text{span}\left\{\vec{v}_{1,\ldots,\vec{v}_{k-1}}\right\}$.

(10) Consider the determinant tells you how a volume or area changes under a linear transformation.

(11) $A^{-1} = \frac{1}{\det A}\text{adj}(A)$.

(12) The following result is called **Cramer's rule**. Consider the equation $A\vec{x} = \vec{b}$ with $\det A \neq 0$. The $i$th coordinate of the solution $\vec{x}$ is given by
$$x_i = \frac{\det A_i}{\det A}$$
where $A_i$ is $A$ with the $i$th column replaced by $\vec{b}$.

(13) The following are equivalent
- $Q$ is orthogonal
- The columns of $Q$ form an orthonormal basis for $\mathbb{R}^n$.
- $Q\vec{v} \cdot Q\vec{u} = \vec{v} \cdot \vec{u}$.
- $Q^{-1} = Q^\top$
- $\|Q\vec{v}\| = \|\vec{v}\|$.

## Week 13

### Definitions

(1) Suppose that $A$ is an $n \times n$ matrix. A *nonzero* vector $\vec{v} \in \mathbb{R}^n$ is called an **eigenvector** of $A$ if there exists a scalar $\lambda$ such that
$$A\vec{v} = \lambda\vec{v}.$$
The scalar $\lambda$ is called the **eigenvalue** for the eigenvector $\vec{v}$.

(2) A **discrete dynamical system** is a system whose **state** $x$ is recorded at successive time steps $t = 0, 1, 2, \ldots$ and whose next state is determined from the current one. If the **state vector** is
$$x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix},$$
and there is an $n \times n$ matrix $A$ such that
$$x(t+1) = Ax(t)$$
then the system is called a **discrete linear dynamical system**.

(3) For an $n \times n$ matrix $A$, the equation $\det(A - \lambda I_n) = 0$ is called the **characteristic equation** of the matrix $A$.

(4) The sum of the diagonal entries of a square matrix $A$ is called the **trace** of $A$ and is denoted by $\mathrm{tr}A$.

(5) We say that an eigenvalue $\lambda_0$ of a square matrix $A$ has **algebraic multiplicity** $k$ if $\lambda_0$ is the root of multiplicity $k$ of the characteristic polynomial $f_A(\lambda)$.

### Facts

(1) If $\vec{v}$ is an eigenvector of a matrix $A$ then $\vec{v}$ is an eigenvector of matrices $A^2, A^3, A^4, \ldots$ as well. Furthermore, if $\lambda$ is the eigenvalue for $\vec{v}$, then
$$A^t\vec{v} = \lambda^t\vec{v}$$
for all positive integers $t$.

(2) If $\lambda$ is an eigenvalue of an orthogonal matrix, then $\lambda = \pm 1$.

(3) A square matrix $A$ is invertible if and only if 0 fails to be an eigenvalue of $A$.

(4) Consider an $n \times n$ matrix $A$ and $\lambda \in \mathbb{R}$. Then $\lambda$ is an eigenvalue of $A$ if and only if
$$\det(A - \lambda I_n) = 0.$$

(5) The eigenvalues of a triangular matrix are its diagonal entries.

(6) If $A$ is an $n \times n$ matrix, then
$$\det(A - \lambda I_n) = (-1)^n \lambda^n + (-1)^{n-1} (\mathrm{tr}A) \lambda^{n-1} + \cdots + \det A.$$
This polynomial is called the **characteristic polynomial of the matrix** $A$. Note that its roots are eigenvalues.

(7) An $n \times n$ matrix has *at most* $n$ real eigenvalues, counted with algebraic multiplicity.

(8) If an $n \times n$ matrix $A$ has the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, listed with their algebraic multiplicities, then

$$\det A = \prod_{i=1}^{n} \lambda_i$$

and

$$\mathrm{tr} A = \sum_{i=1}^{n} \lambda_i .$$