

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

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1m} & y_1 \\ a_{21} & a_{22} & \dots & a_{2m} & y_2 \\ & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & y_n \end{array} \right]$$

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 $\text{rref}(A)$ does not contain a leading 1.

- (16) “**TFAE**” is short for *The following are equivalent.*

- (17) Suppose that $\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$ are two vectors with the same number of components. The **dot product** of \vec{v} and \vec{w} is defined to be the scalar

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

- (18) 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}^m , then the **product of the matrix A with the vector \vec{x}** is defined to be

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

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

$$\vec{b} = x_1 \vec{v}_1 + \dots + x_m \vec{v}_m$$

1.2. 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
 - $\text{rank}(A) \leq n$ and $\text{rank}(A) \leq m$.
 - 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 $\vec{v}_1, \dots, \vec{v}_m$ and \vec{x} is a vector in \mathbb{R}^m with components x_1, \dots, x_m , then

$$A\vec{x} = \left[\begin{array}{c|ccc|c} \vdots & \vdots & \dots & \vdots & \vdots \\ \hline \vec{v}_1 & \dots & \vec{v}_m & \vdots & \\ \hline \vdots & & & \vdots & \end{array} \right] \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = x_1 \vec{v}_1 + \dots + x_m \vec{v}_m.$$

2. LINEAR TRANSFORMATIONS

2.1. 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 $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ 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 if $p = q$.
- (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 $T : X \rightarrow Y$ 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** $T^{-1} : Y \rightarrow X$ is defined by

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

- (5) A square matrix A is called **invertible** if the linear transformation T defined by $T(\vec{x}) = A\vec{x}$ is invertible.
- (6) 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 $T^{-1}(\vec{x}) = A^{-1}\vec{x}$.

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

2.2. **Facts.**

- (1) Consider a linear transformation $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$. 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 $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ 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) Let A be an $n \times n$ matrix. TFAE (The following are equivalent).
- A is invertible
 - $\text{rref}(A) = I_n$
 - $\text{rank}(A) = n$.
 - $A\vec{x} = \vec{0}$ 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 $A^{-1}\vec{b}$).
- (7) Let A be an $n \times n$ matrix that is *not* invertible. Then $A\vec{x} = \vec{b}$ has infinitely many solutions or none.
- (8) 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.

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

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

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

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

3. SUBSPACES OF \mathbb{R}^n AND THEIR DIMENSIONS

3.1. Definitions.

(1) The **image** of a function $f : X \rightarrow Y$ 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

$$\text{image}(A) = \{\vec{y} \in \mathbb{R}^n \mid \vec{y} = A\vec{x} \text{ for some } \vec{x} \in \mathbb{R}^m\}.$$

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

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

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

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

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

$$\vec{w} \in W \implies k\vec{w} \in W \quad \forall k \in \mathbb{R}.$$

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

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

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

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

(6) 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.

(7) Consider the set of vectors $\{\vec{v}_1, \dots, \vec{v}_m\}$ in \mathbb{R}^n .

- (a) \vec{v}_i in the list $\vec{v}_1, \dots, \vec{v}_k$ is called **redundant** if \vec{v}_i is a linear combination of the preceding vectors $\vec{v}_1, \dots, \vec{v}_{i-1}$.
- (b) the vectors $\vec{v}_1, \dots, \vec{v}_m$ are called **linearly independent** if none of them is redundant.

(8) The set of vectors $\{\vec{v}_1, \dots, \vec{v}_m\}$ form a **basis** of a subspace V of \mathbb{R}^n if

- $V = \text{span}\{\vec{v}_1, \dots, \vec{v}_m\}$ and
- $\{\vec{v}_1, \dots, \vec{v}_m\}$ are linearly independent.

(9) Consider the set of vectors $\{\vec{v}_1, \dots, \vec{v}_m\}$. An equation of the form

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

where $c_i \in \mathbb{R}$ is called a **relation** among the vectors $\vec{v}_1, \dots, \vec{v}_m$. A relation is called **trivial** if $c_1 = \dots = c_m = 0$ and **non-trivial** if it is not trivial.

(10) The number of elements in a basis of a linear subspace $V \subseteq \mathbb{R}^n$ is called the **dimension** of V . 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 .

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

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

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 $[\vec{x}]_{\mathcal{B}}$.

- (12) Consider a linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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 .

- (13) 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

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

3.2. 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 $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ has the following properties
 (a) $\vec{0} \in \text{image } T$
 (b) $\text{image } T$ is *closed under addition*.
 (c) $\text{image } T$ is *closed under scalar multiplication*.

Thus $\text{image } T$ is a *subspace* of \mathbb{R}^n .

- (3) Consider a linear transformation $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$.
 (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) Suppose that A is an $n \times m$ matrix. Then

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

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

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

- (6) The vectors in the kernel of an $n \times m$ matrix A correspond to the nontrivial linear relations among the column vectors of A . That is,

$$A\vec{x} = \vec{0} \implies x_1\vec{v}_1 + \dots + x_m\vec{v}_m = \vec{0}.$$

- (7) All bases of a subspace $V \subseteq \mathbb{R}^n$ have the same number of elements.
 (8) 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 .

- (9) For any matrix A ,

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

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

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

In other words,

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

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

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

is invertible.

4. LINEAR SPACES

4.1. 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:
- $(f + g) + h = f + (g + h)$
 - $f + g = g + f$
 - $\exists!$ **neutral element** $n \in V$ such that $f + n = f \forall f \in V$. This element is denoted by 0 .
 - For each $f \in V, \exists!g \in V \mid f + g = 0$. This g is denoted by $-f$.
 - $k(f + g) = kf + kg$.
 - $(c + k)f = cf + kf$
 - $c(kf) = (ck)f$
 - $1f = f$.
- (2) A subset W of a linear space V is called a **linear subspace** if
- $\vec{0} \in W$.
 - W is closed under addition
 - W is closed under scalar multiplication.
- (3) A function $T : V \rightarrow W$ 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

$$A = MBM^{-1} \quad B = M^{-1}AM.$$

- (6) Consider two bases $\mathcal{B}, \mathcal{B}'$ of an n -dimensional linear space V . Consider the linear transformation $L_{\mathcal{B}} \circ L_{\mathcal{B}'}^{-1}$ from \mathbb{R}^n to \mathbb{R}^n . Let $S\vec{x} = (L_{\mathcal{B}} \circ L_{\mathcal{B}'}^{-1})(\vec{x})$. Then

$$[f]_{\mathcal{B}} = S[f]_{\mathcal{B}'}, \quad \text{for all } f \in V.$$

This matrix S is called the **change of basis matrix**.

4.2. Facts.

- (1) Let V be a linear subspace with two given bases \mathcal{B} and \mathcal{B}' . Consider a linear transformation $T : V \rightarrow V$ 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

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

- (2) Suppose that V and V' are two vector spaces with the same (finite) dimension. Then there exists an isomorphism $T : V \rightarrow V'$. The converse also holds. That is, if there is an isomorphism $T : V \rightarrow V'$ then $\dim V = \dim V'$

5. ORTHOGONALITY AND LEAST SQUARES

5.1. 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
- $u_i \cdot u_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$
 - $V = \text{span}\{u_1, \dots, u_m\}$.

- (2) For a subspace V of \mathbb{R}^n , the **orthogonal complement** of V is the set

$$V^\perp = \{x \in \mathbb{R}^n \mid v \cdot x = 0 \text{ for all } v \in V\}.$$

- (3) 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

$$\text{proj}_V(x) = (x \cdot u_1)u_1 + (x \cdot u_2)u_2 + \dots + (x \cdot u_m)u_m = \sum_{i=1}^m (x \cdot u_i)u_i.$$

The difference $x - \text{proj}_V(x)$ lies in V^\perp , the orthogonal complement of V .

- (4) Given a basis $\{v_1, \dots, v_m\}$ of a subspace $V \subseteq \mathbb{R}^n$, the vectors

$$u_1 = \frac{v_1}{\|v_1\|}, \quad u_k = \frac{v_k - \sum_{j=1}^{k-1} (v_k \cdot u_j)u_j}{\|v_k - \sum_{j=1}^{k-1} (v_k \cdot u_j)u_j\|} \quad (k = 2, \dots, m)$$

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

- (5) If the columns of a matrix M form a basis $\{v_1, \dots, v_m\}$ of a subspace V and $\{u_1, \dots, u_m\}$ 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 .

- (6) If Q is an $m \times n$ matrix with entries Q_{ij} , its **transpose** Q^T is the $n \times m$ matrix obtained by interchanging rows and columns:

$$(Q^T)_{ij} = Q_{ji}.$$

Equivalently, for all vectors $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$,

$$(Qx) \cdot y = x \cdot (Q^T y).$$

(In particular, if the columns of Q are orthonormal, then $Q^T Q = I$.)

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

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

5.2. Facts.

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

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

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

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

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

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

- (3) $(V^\perp)^\perp = V$.

- (4) $\dim(V) + \dim(V^\perp) = n$.

- (5) If the columns of Q form an orthonormal basis of V , then

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

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

$$\text{span}\{v_1, \dots, v_k\} = \text{span}\{u_1, \dots, u_k\}, \quad \forall k.$$

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

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

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.

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

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

Thus, Q preserves both lengths and angles.