6. Vector spaces

- definition
- linear independence
- basis and dimension
- coordinate and change of basis
- range space and null space
- rank and nullity

Vector space

a vector space or linear space (over **R**) consists of

- a set $\mathcal{V}$
- a vector sum $+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$
- a scalar multiplication : $\mathbf{R} \times \mathcal{V} \to \mathcal{V}$
- a distinguished element $0 \in \mathcal{V}$

which satisfy a list of properties

- $x + y \in \mathcal{V} \quad \forall x, y \in \mathcal{V}$ (closed under addition)
- $x + y = y + x$, $\forall x, y \in \mathcal{V}$ (+ is commutative)
- $(x + y) + z = x + (y + z)$, $\forall x, y, z \in \mathcal{V}$ (+ is associative)
- $0 + x = x$, $\forall x \in \mathcal{V}$ (0 is additive identity)
- $\forall x \in \mathcal{V}\ \exists (-x) \in \mathcal{V}$ s.t. $x + (-x) = 0$ (existence of additive inverse)
- $\alpha x \in \mathcal{V}$ for any $\alpha \in \mathbf{R}$ (closed under scalar multiplication)
- $(\alpha\beta)x = \alpha(\beta x)$, $\forall \alpha, \beta \in \mathbf{R}\ \forall x \in \mathcal{V}$ (scalar multiplication is associative)
- $\alpha(x + y) = \alpha x + \alpha y$, $\forall \alpha \in \mathbf{R}\ \forall x, y \in \mathcal{V}$ (right distributive rule)
- $(\alpha + \beta)x = \alpha x + \alpha y$, $\forall \alpha, \beta \in \mathbf{R}\ \ \forall x \in \mathcal{V}$ (left distributive rule)
- $\mathbf{1}x = x$, $\forall x \in \mathcal{V}$ (1 is multiplicative identity)

notation

- $(\mathcal{V}, \mathbf{R})$ denotes a vector space $\mathcal{V}$ over **R**
- an element in $\mathcal{V}$ is called a **vector**

Theorem: let u be a vector in $\mathcal{V}$ and k a scalar; then

- $0u = 0$ (multiplication with zero gives the zero vector)
- $k0 = 0$ (multiplication with the zero vector gives the zero vector)
- $(-1)u = -u$ (multiplication with -1 gives the additive inverse)
- if $ku = 0$, then $k = 0$ or $u = 0$

roughly speaking, a vector space must satisfy the following operations

1. **vector addition**

$$x, y \in \mathcal{V} \quad \Rightarrow \quad x + y \in \mathcal{V}$$

2. **scalar multiplication**

$$\text{for any } \alpha \in \mathbf{R}, \;\; x \in \mathcal{V} \quad \Rightarrow \quad \alpha x \in \mathcal{V}$$

the second condition implies that a vector space contains the **zero vector**

$$0 \in \mathcal{V}$$

in otherwords, if $\mathcal{V}$ is a vector space then $0 \in \mathcal{V}$

(but the converse is *not true*)

examples: the following sets are vector spaces (over $\mathbf{R}$)

- $\mathbf{R}^n$
- $\{0\}$
- $\mathbf{R}^{m\times n}$
- $\mathbf{C}^{m\times n}$: set of $m\times n$-complex matrices
- $\mathbf{P}_n$: set of polynomials of degree $\leq n$

$$\mathbf{P}_n = \{p(t) \mid p(t) = a_0 + a_1 t + \cdots + a_n t^n\}$$

- $\mathbf{S}^n$: set of symmetric matrices of size n
- $C(-\infty,\infty)$: set of real-valued continuous functions on $(-\infty,\infty)$
- $C^n(-\infty,\infty)$: set of real-valued functions with continuous nth derivatives on $(-\infty,\infty)$

✎ check whether any of the following sets is a vector space (over $\mathbf{R}$)

- $\{0, 1, 2, 3, \ldots\}$

- $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$

- $\left\{ x \in \mathbf{R}^2 \mid x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}, \quad x_1 \in \mathbf{R} \right\}$

- $\left\{ p(x) \in \mathbf{P}_2 \mid p(x) = a_1 x + a_2 x^2 \quad \text{for some } a_1, a_2 \in \mathbf{R} \right\}$

Subspace

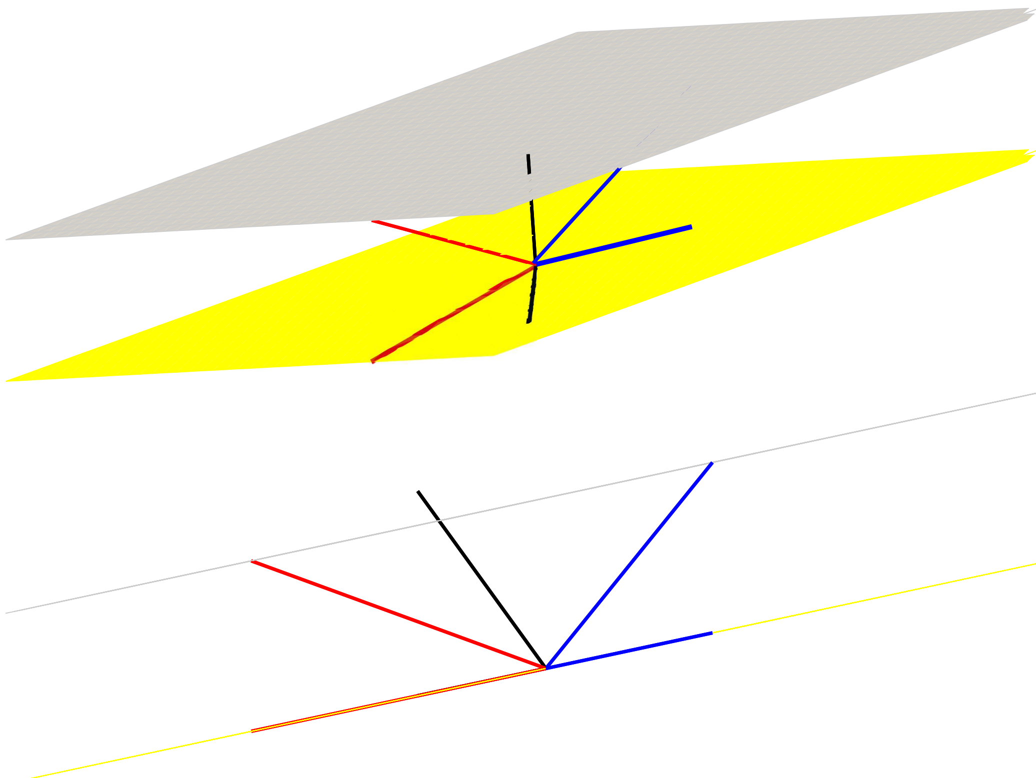
- a **subspace** of a vector space is a *subset* of a vector space which is itself a vector space
- a subspace is closed under vector addition and scalar multiplication

examples:

- $\{0\}$ is a subspace of $\mathbf{R}^n$
- $\mathbf{R}^{m\times n}$ is a subspace of $\mathbf{C}^{m\times n}$
- $\left\{x \in \mathbf{R}^2 \mid x_1 = 0\right\}$ is a subspace of $\mathbf{R}^2$
- $\left\{x \in \mathbf{R}^2 \mid x_2 = 1\right\}$ is not a subspace of $\mathbf{R}^2$
- $\left\{\begin{bmatrix} 1 & 4 \\ -3 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\right\}$ is not a subspace of $\mathbf{R}^{2\times 2}$
- the solution set $\{x \in \mathbf{R}^n \mid Ax = b\}$ for $b \neq 0$ is a not subspace of $\mathbf{R}^n$

examples: two hyperplanes; one is a subspace but the other one is not

$$2x_1 - 3x_2 + x_3 = 0 \quad \text{(yellow)}, \qquad 2x_1 - 3x_2 + x_3 = 20 \quad \text{(grey)}$$



black = **red** + **blue**

$x = (-3, -2, 0)$ & $y = (1, -1, -5)$ are on the yellow plane, and so is $x + y$

$x = (-3, -2, 20)$ & $y = (1, -1, 15)$ are on the grey plane, but $x + y$ is not

Linear Independence

Definition: a set of vectors $\{v_1, v_2, \ldots, v_n\}$ is **linearly independent** if

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0 \Longrightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$$

equivalent conditions:

- coefficients of $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ are uniquely determined, i.e.,

  $$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$$

  implies $\alpha_k = \beta_k$ for $k = 1, 2, \ldots, n$

- no vector v_i can be expressed as a linear combination of the other vectors

examples:

- $\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}3\\1\\0\end{bmatrix}$ are independent

- $\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}3\\1\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\1\end{bmatrix}$ are independent

- $\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}3\\1\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\1\end{bmatrix}, \begin{bmatrix}4\\2\\0\end{bmatrix}$ are not independent

- $\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}3\\1\\0\end{bmatrix}, \begin{bmatrix}2\\-1\\-1\end{bmatrix}$ are not independent

Linear span

Definition: the linear span of a set of vectors

$$\{v_1, v_2, \ldots, v_n\}$$

is the set of all linear combinations of $v_1, \ldots, v_n$

$$\text{span}\{v_1, v_2, \ldots, v_n\} = \{a_1 v_1 + a_2 v_2 + \cdots + a_n v_n \mid a_1, \ldots, a_n \in \mathbf{R}\}$$

example:

$\text{span}\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ is the set of 2×2 symmetric matrices

Fact: if $v_1, \ldots, v_n$ are vectors in $\mathcal{V}$, $\text{span}\{v_1, \ldots, v_n\}$ is a subspace of $\mathcal{V}$

Basis and dimension

Definition: set of vectors $\{v_1, v_2, \cdots, v_n\}$ is a **basis** for a vector space $\mathcal{V}$ if

- $\{v_1, v_2, \ldots, v_n\}$ is linearly independent
- $\mathcal{V} = \text{span}\ \{v_1, v_2, \ldots, v_n\}$

equivalent condition: *every* $v \in \mathcal{V}$ can be *uniquely* expressed as

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

Definition: the **dimension** of $\mathcal{V}$, denoted $\dim(\mathcal{V})$, is the number of vectors in a basis for $\mathcal{V}$

Theorem: the number of vectors in *any* basis for $\mathcal{V}$ is the same

(we assign $\dim\{0\} = 0$)

examples:

- $\{e_1, e_2, e_3\}$ is a standard basis for $\mathbf{R}^3$ $\qquad (\dim \mathbf{R}^3 = 3)$
- $\left\{ \begin{bmatrix} -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right\}$ is a basis for $\mathbf{R}^2$ $\qquad (\dim \mathbf{R}^2 = 2)$
- $\{1, t, t^2\}$ is a basis for $\mathbf{P}_2$ $\qquad (\dim \mathbf{P}_2 = 3)$
- $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ is a basis for $\mathbf{R}^{2\times 2}$ $(\dim \mathbf{R}^{2\times 2} = 4)$
- $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ cannot be a basis for $\mathbf{R}^3$ why ?
- $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \end{bmatrix} \right\}$ cannot be a basis for $\mathbf{R}^2$ why ?

Coordinates

let $S = \{v_1, v_2, \dots, v_n\}$ be a basis for a vector space $\mathcal{V}$

suppose a vector $v \in \mathcal{V}$ can be written as

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$$

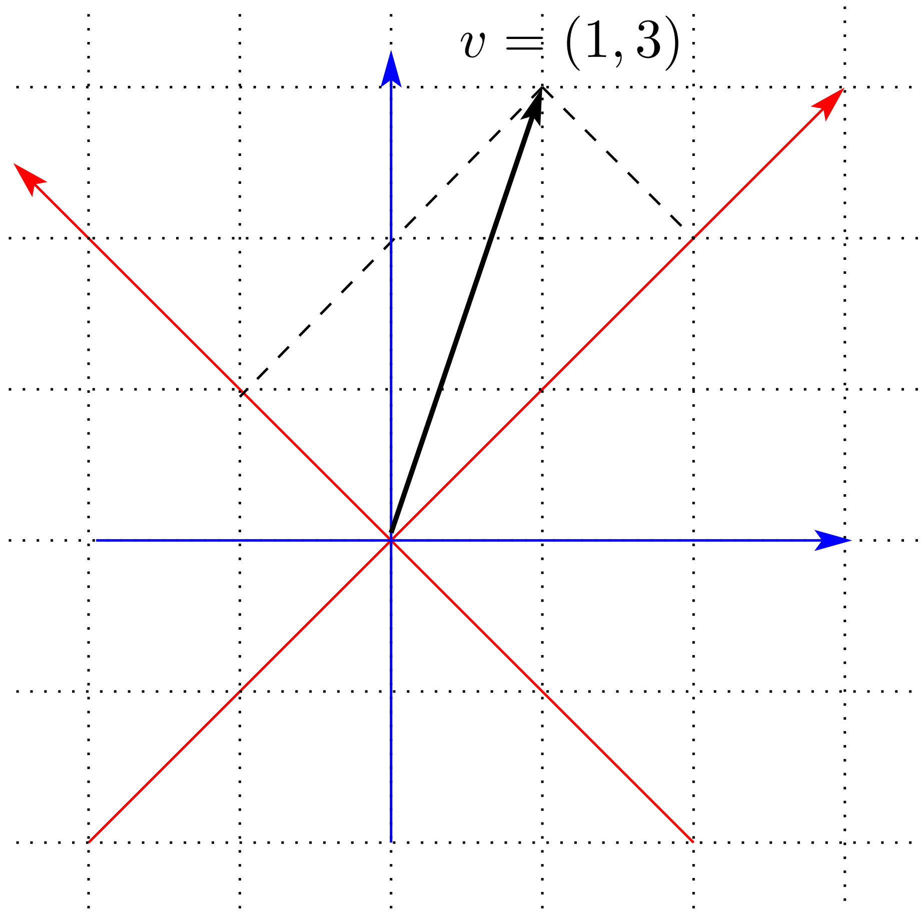
Definition: the coordinate vector of v relative to the basis S is

$$[v]_S = (a_1, a_2, \dots, a_n)$$

- linear independence of vectors in S ensures that a_k's are *uniquely* determined by S and v
- changing the basis yields a different coordinate vector

Geometrical interpretation

new coordinate in a new reference axis



$$v = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

examples:

- $S = \{e_1, e_2, e_3\}$, $v = (-2, 4, 1)$

$$v = -2e_1 + 4e_2 + 1e_3, \quad [v]_S = (-2, 4, 1)$$

- $S = \{(-1, 2, 0), (3, 0, 0), (-2, 1, 1)\}$, $v = (-2, 4, 1)$

$$v = \begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix} = \frac{3}{2}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + 1\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \quad [v]_S = (3/2, 1/2, 1)$$

- $S = \{1, t, t^2\}$, $v(t) = -3 + 2t + 4t^2$

$$v(t) = -3 \cdot 1 + 2 \cdot t + 4 \cdot t^2, \quad [v]_S = (-3, 2, 4)$$

- $S = \{1, t - 1, t^2 + t\}$, $v(t) = -3 + 2t + 4t^2$

$$v(t) = -5 \cdot 1 - 2 \cdot (t - 1) + 4 \cdot (t^2 + t), \quad [v]_S = (-5, -2, 4)$$

Change of basis

let $U = \{u_1, \ldots, u_n\}$ and $W = \{w_1, \ldots, w_n\}$ be bases for a vector space $\mathcal{V}$

a vector $v \in \mathcal{V}$ has the coordinates relative to these bases as

$$[v]_U = (a_1, a_2, \ldots, a_n), \quad [v]_W = (b_1, b_2, \ldots, b_n)$$

suppose the coordinate vectors of w_k relative to U is

$$[w_k]_U = (c_{1k}, c_{2k}, \ldots, c_{nk})$$

or in the matrix form as

$$\begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

the coordinate vectors of v relative to U and W are related by

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \triangleq P \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

- we obtain $[v]_U$ by multiplying $[v]_W$ with P
- P is called the **transition** matrix from W to U
- the columns of P are the coordinate vectors of the basis vectors in W relative to U

Theorem ✌

P is invertible and P^{-1} is the transition matrix from U to W

example: find $[v]_U$, given

$$U = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}, \quad W = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}, \quad [v]_W = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

first, find the coordinate vectors of the basis vectors in W relative to U

$$\begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

from which we obtain the transition matrix

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 3 & 1 \\ -1 & -1 \end{bmatrix}$$

and $[v]_U$ is given by

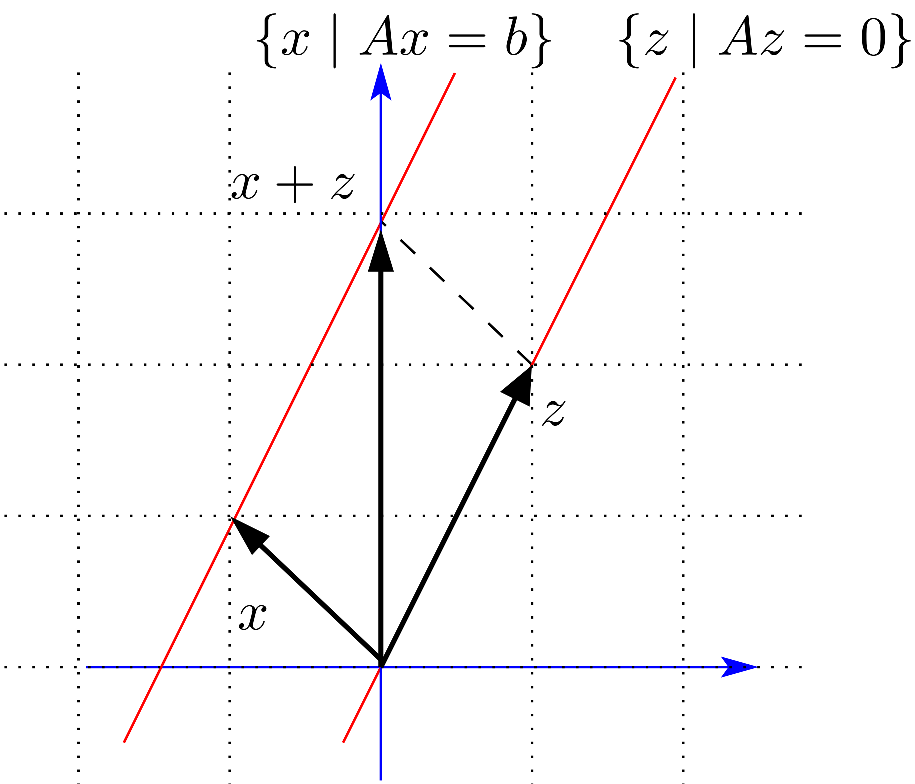
$$[v]_U = \frac{1}{2} \begin{bmatrix} 3 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -5 \\ 1 \end{bmatrix}$$

Nullspace

the **nullspace** of an $m \times n$ matrix is defined as

$$\mathcal{N}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\}$$

- the set of all vectors that are mapped to zero by $f(x) = Ax$
- the set of all vectors that are orthogonal to the rows of A
- if $Ax = b$ then $A(x + z) = b$ for all $z \in \mathcal{N}(A)$
- also known as **kernel** of A
- $\mathcal{N}(A)$ is a subspace of $\mathbf{R}^n$ ✎



$$A = \begin{bmatrix} 2 & -1 \\ -4 & 2 \\ -6 & 3 \end{bmatrix}, \quad b = \begin{bmatrix} -3 \\ 6 \\ 9 \end{bmatrix}$$

- $\mathcal{N}(A) = \{x \mid 2x_1 - x_2 = 0\}$
- the solution set of $Ax = b$ is $\{x \mid 2x_1 - x_2 = -3\}$
- the solution set of $Ax = b$ is the translation of $\mathcal{N}(A)$

Zero nullspace matrix

- A has a zero nullspace if $\mathcal{N}(A) = \{0\}$
- if A has a zero nullspace and $Ax = b$ is solvable, the solution is unique
- columns of A are independent

✌ **equivalent conditions:** $A \in \mathbf{R}^{n \times n}$

- A has a zero nullspace
- A is invertible or nonsingular
- columns of A are a basis for $\mathbf{R}^n$

Range space

the **range** of an $m \times n$ matrix A is defined as

$$\mathcal{R}(A) = \{y \in \mathbf{R}^m \mid y = Ax \text{ for some } x \in \mathbf{R}^n\}$$

- the set of all m-vectors that can be expressed as Ax
- the set of all linear combinations of the columns of $A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$

$$\mathcal{R}(A) = \{y \mid y = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n, \quad x \in \mathbf{R}^n\}$$

- the set of all vectors b for which $Ax = b$ is solvable
- also known as the **column space** of A
- $\mathcal{R}(A)$ is a subspace of $\mathbf{R}^m$ ✎

Full range matrices

A has a full range if $\mathcal{R}(A) = \mathbf{R}^m$

✌ **equivalent conditions:**

- A has a full range
- columns of A span $\mathbf{R}^m$
- $Ax = b$ is solvable for *every* b
- $\mathcal{N}(A^T) = \{0\}$

Bases for $\mathcal{R}(A)$ and $\mathcal{N}(A)$

A and B are row equivalent matrices, $i.e.$,

$$B = E_k \cdots E_2 E_1 A$$

Facts ✌

- elementary row operations *do not alter* $\mathcal{N}(A)$

$$\mathcal{N}(B) = \mathcal{N}(A)$$

- columns of B are independent if and only if columns of A are
- a given set of column vectors of A forms a basis for $\mathcal{R}(A)$ if and only if the corresponding column vectors of B form a basis for $\mathcal{R}(B)$

example: given a matrix A and its row echelon form B:

$$A = \begin{bmatrix} -1 & 2 & 4 & 1 \\ 0 & 1 & 2 & 1 \\ 2 & 3 & 6 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

basis for $\mathcal{N}(A)$: from $\{x \mid Ax = 0\} = \{x \mid Bx = 0\}$, we read

$$x_1 + x_4 = 0, \quad x_2 + 2x_3 + x_4 = 0$$

define x_3 and x_4 as free variables, any $x \in \mathcal{N}(A)$ can be written as

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -x_4 \\ -2x_3 - x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

(a linear combination of $(0, -2, 1, 0)$ and $(-1, -1, 0, 1)$

hence, a basis for $\mathcal{N}(A)$ is $\left\{ \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ and $\dim \mathcal{N}(A) = 2$

basis for $\mathcal{R}(A)$: pick a set of the independent column vectors in B (here pick the 1st and the 2nd columns)

the corresponding columns in A form a basis for $\mathcal{R}(A)$:

$$\left\{ \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \right\}$$

$\dim \mathcal{R}(A) = 2$

✌ **conclusion:** if R is the row reduced echelon form of A

- the pivot column vectors of R form a basis for the range space of R
- the column vectors of A *corresponding* to the pivot columns of R form a basis for the range space of A
- $\dim \mathcal{R}(A)$ is the number of leading 1's in R
- $\dim \mathcal{N}(A)$ is the number of free variables in solving $Rx = 0$

Rank and Nullity

rank of a matrix $A \in \mathbf{R}^{m \times n}$ is defined as

$$\mathbf{rank}(A) = \dim \mathcal{R}(A)$$

nullity of a matrix $A \in \mathbf{R}^{m \times n}$ is

$$\mathbf{nullity}(A) = \dim \mathcal{N}(A)$$

Facts ✌

- $\mathbf{rank}(A)$ is maximum number of independent columns (or rows) of A

$$\mathbf{rank}(A) \leq \min(m, n)$$

- $\mathbf{rank}(A) = \mathbf{rank}(A^T)$

Full rank matrices

for $A \in \mathbf{R}^{m \times n}$ we always have $\mathbf{rank}(A) \leq \min(m, n)$

we say A is **full rank** if $\mathbf{rank}(A) = \min(m, n)$

- for **square** matrices, full rank means nonsingular (invertible)
- for **skinny** matrices ($m \geq n$), full rank means columns are independent
- for **fat** matrices ($m \leq n$), full rank means rows are independent

Rank-Nullity Theorem

for any $A \in \mathbf{R}^{m \times n}$,

$$\mathbf{rank}(A) + \dim \mathcal{N}(A) = n$$

Proof:

- a homogeneous linear system $Ax = 0$ has n variables
- these variables fall into two categories
  - leading variables
  - free variables
- $\#$ of leading variables $=$ $\#$ of leading 1's in reduced echelon form of A $= \mathbf{rank}(A)$
- $\#$ of free variables $=$ nullity of A

MATLAB commands

- `rref(A)` produces the reduced row echelon form of A

```
>> A = [-1 2 4 1;0 1 2 1;2 3 6 5]
A =
    -1     2     4     1
     0     1     2     1
     2     3     6     5

>> rref(A)

ans =

     1     0     0     1
     0     1     2     1
     0     0     0     0
```

- `rank(A)` provides an estimate of the rank of A

- `null(A)` gives normalized vectors in a basis for $\mathcal{N}(A)$

```
>> A

A =

     1    -3     2
     2    -6     4
     3    -9     6

>> U = null(A)

U =

   -0.8729   -0.4082
   -0.4364    0.4082
   -0.2182    0.8165
```

(and we can verify that $AU = 0$)

References

Chapter 4 in

H. Anton, *Elementary Linear Algebra*, 10th edition, Wiley, 2010

Lecture note on

Linear algebra review, EE263, S. Boyd, Stanford University

Lecture note on

Theory of linear equations, EE103, L. Vandenberghe, UCLA