Vector Space

Chia-Ping Chen

Professor
Department of Computer Science and Engineering
National Sun Yat-sen University

Linear Algebra

Outline

- Under-determined system of linear equations
- Vector space
- The fundamental subspaces of a matrix
- Linear transformation

NOTATION

- $\boldsymbol{Ax} = \boldsymbol{b}$: under-determined system of linear equations
- $\boldsymbol{U}$: echelon matrix
- $\mathbb{V}, \mathbb{S}$: vector space or subspace
- $\mathcal{B}, \mathcal{B}'$: basis
- $\boldsymbol{T} : \mathbb{D} \to \mathbb{R}$: linear transform from domain $\mathbb{D}$ to range $\mathbb{R}$
- $[\boldsymbol{x}_{\mathcal{B}}]$: column representation of vector $\boldsymbol{x}$ using basis $\mathcal{B}$
- $[\boldsymbol{T}_{\mathcal{B}\mathcal{B}'}]$: matrix representation of $\boldsymbol{T} : \mathbb{D} \to \mathbb{R}$ using basis $\mathcal{B}$ for $\mathbb{D}$ and basis $\mathcal{B}'$ for $\mathbb{R}$

Under-determined System of Linear Equations

UNDER-DETERMINED SYSTEM

Consider a system of linear equations with m equations and n unknowns, and $m < n$.

- It is called an **under-determined** system
- It can be represented by $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$
- $\boldsymbol{A}$ is of order $m \times n$, $\boldsymbol{x}$ is $n \times 1$, and $\boldsymbol{b}$ is $m \times 1$

Consider

$$\begin{cases} u + 3v + 3w = 1 \\ 2u + 6v + 9w = 5 \end{cases}$$

- It is under-determined with $m = 2$ and $n = 3$
- It can be represented by $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ where

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 3 \\ 2 & 6 & 9 \end{bmatrix}, \ \boldsymbol{x} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}, \ \boldsymbol{b} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

From under-determined system to square system

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be an under-determined system of linear equations with n unknowns and m equations. It can be converted to square system by moving $n-m$ unknowns to right side.

Consider

$$\begin{cases} u + 3v + 3w = 1 \\ 2u + 6v + 9w = 5 \end{cases}$$

Moving w to the right side, we get

$$\begin{cases} u + 3v = 1 - 3w \\ 2u + 6v = 5 - 9w \end{cases}$$

which can be seen as a square system with 2 unknowns.

Which unknowns to move

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be an under-determined system of linear equations. We convert it to a square system by moving unknowns.

- Moving the right unknowns makes it non-singular
- Moving the wrong unknowns makes it singular

Consider

$$\begin{cases} u + 3v + 3w = 1 \\ 2u + 6v + 9w = 5 \end{cases}$$

- Moving v to the right side makes it non-singular

$$\begin{cases} u + 3w = 1 - 3v \\ 2u + 9w = 5 - 6v \end{cases}$$

- Moving w to the right side makes it singular

$$\begin{cases} u + 3v = 1 - 3w \\ 2u + 6v = 5 - 9w \end{cases}$$

Theorem (solving an under-determined system)

Let $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ be an under-determined system of linear equations. Exactly one of the following cases is true.

1. *No solution*
2. *Infinite solutions*

3 steps to solve an under-determined system

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be an under-determined system of linear equations.

1. Solve (the **homogeneous equation**) $\boldsymbol{Ax} = \boldsymbol{0}$

$$\mathbb{H} = \{\boldsymbol{x}_n \,|\, \boldsymbol{Ax}_n = \boldsymbol{0}\}$$

2. Find (a **particular solution**) $\boldsymbol{x}_p$ such that

$$\boldsymbol{Ax}_p = \boldsymbol{b}$$

3. A **general solution** is

$$\boldsymbol{x}_g = \boldsymbol{x}_p + \boldsymbol{x}_n$$

Example (solve an under-determined system)

$$\mathcal{P}:\begin{cases} u + 3v + 3w + 2y = 1 \\ 2u + 6v + 9w + 7y = 5 \\ -u - 3v + 3w + 4y = 5 \end{cases}$$

Replace right side by $\mathbf{0}$ and solve the homogeneous equation.

$$\mathcal{P} \xrightarrow{\boldsymbol{b}\leftarrow\mathbf{0}} \begin{cases} u + 3v + 3w + 2y = 0 \\ 2u + 6v + 9w + 7y = 0 \\ -u - 3v + 3w + 4y = 0 \end{cases}$$

$$\xrightarrow{\text{elimination}} \begin{cases} u + 3v + 3w + 2y = 0 \\ 3w + 3y = 0 \\ 6w + 6y = 0 \end{cases}$$

$$\xrightarrow{\text{elimination}} \begin{cases} u + 3v + 3w + 2y = 0 \\ 3w + 3y = 0 \end{cases}$$

So

$$w = -y, \quad u = -3v + y$$

- Variables u and w stays on the left
- Variables v and y are moved to the right
- Values of u and w are determined by values of v and y

- A solution can be represented by a vector

$$\boldsymbol{x}_n = \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} -3v + y \\ v \\ -y \\ y \end{bmatrix} = v \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$
$$= v\boldsymbol{x}_1 + y\boldsymbol{x}_2$$

- It is a linear combination of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$

Restore the right side $\boldsymbol{b}$ and find a particular solution $\boldsymbol{x}_p$.

$$\mathcal{P}: \left\{\begin{array}{rcrcrcrcr} u & + & 3v & + & 3w & + & 2y & = & 1 \\ 2u & + & 6v & + & 9w & + & 7y & = & 5 \\ -u & - & 3v & + & 3w & + & 4y & = & 5 \end{array}\right.$$

- Letting $v = y = 0$, we have

$$\left\{\begin{array}{rcrcr} u & + & 3w & = & 1 \\ 2u & + & 9w & = & 5 \\ -u & + & 3w & = & 5 \end{array}\right.$$

  so $w = 1$ and $u = -2$.
- This particular solution can be represented by a vector

$$\boldsymbol{x}_p = \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

General Solution

Let $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ be an under-determined system of linear equations. The sum of a homogeneous solution and a particular solution is a solution.

Consider $\boldsymbol{x}_n + \boldsymbol{x}_p$. It is a solution of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ since

$$\boldsymbol{A}(\boldsymbol{x}_n + \boldsymbol{x}_p) = \boldsymbol{A}\boldsymbol{x}_n + \boldsymbol{A}\boldsymbol{x}_p = \boldsymbol{b}$$

In the current example

$$\boldsymbol{x}_n + \boldsymbol{x}_p = v \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

SOLVING THE HOMOGENEOUS EQUATION VIA MATRIX

- The homogeneous equation was solved by elimination, i.e. a sequence of elimination steps
- Elimination step is equivalent to row operation on the coefficient matrix
- In particular

$$\underbrace{\begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix}}_{\boldsymbol{A}} \rightarrow \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 6 & 6 \end{bmatrix} \rightarrow \underbrace{\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{\boldsymbol{U}}$$

Echelon Matrix (Row Echelon Form)

Elimination converts $\boldsymbol{Ax} = \mathbf{0}$ to $\boldsymbol{Ux} = \mathbf{0}$ where $\boldsymbol{U}$ is an **echelon matrix**.

- In each non-zero row of $\boldsymbol{U}$, the first non-zero element is a **pivot**
- Pivots descend to the right
- Using pivots as anchors, we can draw a zigzag line on $\boldsymbol{U}$ such that the elements below the line are 0
- An echelon matrix can be converted to a **reduced echelon matrix** (a.k.a. reduced form) where every pivot is 1

$$U = \begin{bmatrix} \bullet & * & * & * & * & * & * & * \\ 0 & \bullet & * & * & * & * & * & * \\ 0 & 0 & 0 & \bullet & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bullet \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad R = \begin{bmatrix} \mathbf{1} & \mathbf{0} & * & \mathbf{0} & * & * & * & \mathbf{0} \\ 0 & \mathbf{1} & * & \mathbf{0} & * & * & * & \mathbf{0} \\ 0 & 0 & 0 & \mathbf{1} & * & * & * & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Figure 2.3: The entries of a 5 by 8 echelon matrix U and its reduced form R.

Pivot variables and free variables

Suppose elimination converts $\boldsymbol{Ax} = \mathbf{0}$ to $\boldsymbol{Ux} = \mathbf{0}$ where $\boldsymbol{U}$ is an echelon matrix.

- Pivot positions correspond to **pivot variables**
- The other variables are **free variables**

Consider the system

$$\begin{cases} u + 3v + 3w + 2y = 1 \\ 2u + 6v + 9w + 7y = 5 \\ -u - 3v + 3w + 4y = 5 \end{cases}$$

- u and w are pivot variables
- v and y are free variables

HOMOGENEOUS/PARTICULAR/GENERAL SOLUTIONS

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be an under-determined system with m equations and n unknowns. Suppose elimination converts $\boldsymbol{Ax} = \boldsymbol{0}$ to $\boldsymbol{Ux} = \boldsymbol{0}$ where $\boldsymbol{U}$ is an echelon matrix. Let r be the number of pivots in $\boldsymbol{U}$.

- The number of pivot variables is r
- The number of free variables is $n - r$
- We can find $n - r$ homogeneous solutions by setting one free variable to 1 and the other free variables to 0
- If the system is solvable, we can find a particular solution by setting free variables to 0
- The sum of a homogeneous solution and a particular solution is a general solution

For $\mathcal{P}$, we have homogeneous solution

$$\boldsymbol{x}_n = \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} -3v + y \\ v \\ -y \\ y \end{bmatrix} = v \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

and a particular solution

$$\boldsymbol{x}_p = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

The general solution is

$$\boldsymbol{x}_g = \boldsymbol{x}_p + \boldsymbol{x}_n = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + v \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

$$\begin{cases} y + z = 2 \\ 2y + 2z = 4 \end{cases} \Rightarrow \boldsymbol{x}_n = c\begin{bmatrix}-1\\1\end{bmatrix}, \boldsymbol{x}_p = \begin{bmatrix}1\\1\end{bmatrix}$$

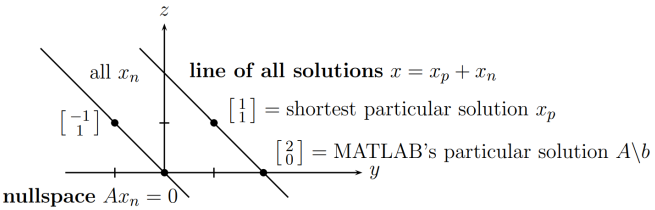


Figure 2.2: The parallel lines of solutions to $Ax_n = 0$ and $\left[\begin{smallmatrix}1&1\\2&2\end{smallmatrix}\right]\left[\begin{smallmatrix}y\\z\end{smallmatrix}\right] = \left[\begin{smallmatrix}2\\4\end{smallmatrix}\right]$.

Example (exercise)

$$\begin{cases} 1x_1 + 2x_2 + 3x_3 + 5x_4 = 0 \\ 2x_1 + 4x_2 + 8x_3 + 12x_4 = 6 \\ 3x_1 + 6x_2 + 7x_3 + 13x_4 = -6 \end{cases}$$

Vector Space

Definition (vector space)

Let $\mathbb{V}$ be a set of vectors. $\mathbb{V}$ is a **space** if

- **addition** and **scalar multiplication** are defined for $\mathbb{V}$
- $\mathbb{V}$ is **closed** under addition and scalar multiplication

The following rules hold for addition and scalar multiplication.

1. $\exists\, \mathbf{0} \in \mathbb{V}$ such that $\forall\, \boldsymbol{x} \in \mathbb{V}$ we have $\boldsymbol{x} + \mathbf{0} = \boldsymbol{x}$
2. $\forall\, \boldsymbol{x} \in \mathbb{V}$, $\exists\, \boldsymbol{y} \in \mathbb{V}$ such that $\boldsymbol{x} + \boldsymbol{y} = \mathbf{0}$
3. $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{V}$, we have $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$
4. $\forall\, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{V}$, we have $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$
5. $\forall\, \boldsymbol{x} \in \mathbb{V}$, we have $1\boldsymbol{x} = \boldsymbol{x}$
6. $\forall\, \boldsymbol{x} \in \mathbb{V}$, we have $c_1(c_2\boldsymbol{x}) = (c_1c_2)\boldsymbol{x}$ for any c_1, c_2
7. $\forall\, \boldsymbol{x} \in \mathbb{V}$, we have $(c_1 + c_2)\boldsymbol{x} = c_1\boldsymbol{x} + c_2\boldsymbol{x}$ for any c_1, c_2
8. $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{V}$, we have $c(\boldsymbol{x} + \boldsymbol{y}) = c\boldsymbol{x} + c\boldsymbol{y}$ for any c

Example (vector space)

- $\mathbb{R}^1$
- $\mathbb{R}^2$
- $\mathbb{R}^3$
- $\mathbb{R}^n$
- $\mathbb{M}_{3\times 2}$: the set of matrices of order 3×2
- $\mathbb{F}_{[a,b]}$: the set of functions defined over $[a, b]$

Definition (vector subspace)

Let $\mathbb{S}$ be a set of vectors. $\mathbb{S}$ is a **subspace** if

1. $\mathbb{S} \subset \mathbb{V}$ where $\mathbb{V}$ is a space
2. $\mathbb{S}$ is a space

Example (vector subspace)

1. $\{[0,0,0]\}$: subspace of $\mathbb{R}^3$
2. z-axis: subspace of $\mathbb{R}^3$
3. xy-plane: subspace of $\mathbb{R}^3$
4. $\mathbb{S}_{6\times 6}$ (6×6 symmetric matrices): subspace of $\mathbb{M}_{6\times 6}$
5. $\mathbb{L}_{5\times 5}$ (5×5 lower-triangular matrices): subspace of $\mathbb{M}_{5\times 5}$

Definition (linear combination)

Let $\mathbb{V}$ be a space and $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ be vectors of $\mathbb{V}$. The **linear combination** of $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ is

$$\sum_{i=1}^{n} c_i \boldsymbol{v}_i = c_1 \boldsymbol{v}_1 + \cdots + c_n \boldsymbol{v}_n$$

where $c_1, \cdots, c_n$ are scalars called **combination coefficients**.

The linear combination $\sum_{i=1}^{n} c_i \boldsymbol{v}_i$ is the ending point of a walk in space $\mathbb{V}$ with segments $c_i \boldsymbol{v}_i$'s starting from the origin.

Definition (span)

Let $\mathbb{V}$ be a space and $\mathcal{V} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be a vector set in $\mathbb{V}$. The **span** of $\mathcal{V}$ is

$$\textbf{span}(\mathcal{V}) = \{\boldsymbol{v} \,|\, \boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n\}$$

Let $\mathbb{B} = \textbf{span}(\mathcal{V})$.

- $\mathbb{B}$ is a subspace of $\mathbb{V}$
- $\mathbb{B}$ is the set of points reachable from the origin moving only in the directions of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$
- We say "$\mathcal{V}$ spans $\mathbb{B}$" or "$\mathcal{V}$ is a **spanning set** of $\mathbb{B}$"

Trivial Linear Combination

Let $\sum_{i=1}^{n} c_i \boldsymbol{v}_i$ be a linear combination of $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$.

- It is **trivial** if $c_i = 0$ for all i
- It is **non-trivial** if there exists $c_i \neq 0$

- A trivial linear combination is always $\mathbf{0}$
- A non-trivial linear combination may be $\mathbf{0}$

Definition (linear independence)

Let $\mathcal{V} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be a set of vectors.

- $\mathcal{V}$ is **linearly independent** if every non-trivial linear combination of $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ is a non-zero vector
- Otherwise, $\mathcal{V}$ is **linearly dependent**
- That is, $\mathcal{V}$ is linearly dependent if there exists $c_i \neq 0$ such that

$$\sum_{i=1}^{n} c_i \boldsymbol{v}_i = \mathbf{0}$$

Example (linear dependence)

Convert $\boldsymbol{A}$ to $\boldsymbol{U}$ by row operations

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \boldsymbol{U}$$

- $\{\boldsymbol{u}_{1:}, \boldsymbol{u}_{2:}\}$ is linearly independent
- $\{\boldsymbol{a}_{1:}, \boldsymbol{a}_{2:}\}$ is linearly independent
- $\{\boldsymbol{u}_{1}, \boldsymbol{u}_{3}\}$ is linearly independent
- $\{\boldsymbol{a}_{1}, \boldsymbol{a}_{3}\}$ is linearly independent

Definition (dependent vector)

Let $\mathcal{V} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be a set of vectors. If $\boldsymbol{v}_i$ is a linear combination of $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{i-1}$, it is a **dependent vector** of $\mathcal{V}$.

- By definition, $\boldsymbol{v}_i$ is a dependent vector of $\mathcal{V}$ if

$$\boldsymbol{v}_i = \sum_{j=1}^{i-1} c_j \boldsymbol{v}_j$$

- $\mathbf{0}$ is a dependent vector since

$$\mathbf{0} = \sum_{j=1}^{i-1} 0\, \boldsymbol{v}_j$$

Linear Dependence and Dependent Vector

A set of vectors $\mathcal{V} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is linearly dependent if and only if there exists a dependent vector in $\mathcal{V}$.

- Suppose $\boldsymbol{v}_i$ is a dependent vector so $\boldsymbol{v}_i = \sum_{j=1}^{i-1} c_j \boldsymbol{v}_j$. Then the linear combination $\boldsymbol{v}_i - \sum_{j=1}^{i-1} c_j \boldsymbol{v}_j$ is $\mathbf{0}$ and it is non-trivial (since $c_i = 1 \neq 0$). Hence $\mathcal{V}$ is linearly dependent.
- Suppose $\mathcal{V}$ is linearly dependent so there exists non-trivial linear combination $\sum_{j=1}^{n} c'_j \boldsymbol{v}_j$ that is $\mathbf{0}$. Let i be the largest integer with $c'_i \neq 0$. Then

$$\sum_{j=1}^{i} c'_j \boldsymbol{v}_j = \mathbf{0} \Rightarrow c'_i \boldsymbol{v}_i = -\sum_{j=1}^{i-1} c'_j \boldsymbol{v}_j \Rightarrow \boldsymbol{v}_i = \sum_{j=1}^{i-1} \left(\frac{-c'_j}{c'_i} \right) \boldsymbol{v}_j$$

Hence $\boldsymbol{v}_i$ is a dependent vector.

Theorem

Let $\mathcal{V} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ *be a linearly dependent set.*

- *We can move dependent vectors out of* $\mathcal{V}$ *until it is linearly independent*
- *Moving a dependent vector out of* $\mathcal{V}$ *does not change the space it spans*

Proof.

- Since $\mathcal{V}$ is linearly dependent, a dependent vector exists and we move it out of $\mathcal{V}$. Continue until a dependent vector cannot be found. The remaining set is linearly independent.
- Let $\boldsymbol{v}_i$ be a dependent vector so

$$\boldsymbol{v}_i = \sum_{j=1}^{i-1} c_j \boldsymbol{v}_j$$

Note linear combination of $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_i$ can be written as linear combination of $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{i-1}$. It follows that any linear combination of $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ can be written as a linear combination of $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{i-1}, \boldsymbol{v}_{i+1}, \cdots, \boldsymbol{v}_n$. Hence

$$\textbf{span}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{i-1}, \boldsymbol{v}_{i+1}, \cdots, \boldsymbol{v}_n) = \textbf{span}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n)$$

$\square$

Definition (basis)

Let $\mathbb{S}$ be a space. A set of vectors $\mathcal{B}$ is a **basis** of $\mathbb{S}$ if

- $\mathcal{B}$ is a spanning set of $\mathbb{S}$
- $\mathcal{B}$ is linearly independent

constructing basis from spanning set

Let $\mathcal{V} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be a spanning set of $\mathbb{S}$.

- If $\mathcal{V}$ is linearly independent, $\mathcal{V}$ is a basis of $\mathbb{S}$.
- If $\mathcal{V}$ is linearly dependent, remove dependent vectors until the remaining set

$$\mathcal{V}' = \{\boldsymbol{v}'_1, \cdots, \boldsymbol{v}'_m\} \subset \mathcal{V}$$

is linearly independent. Then $\mathcal{V}'$ is a basis of $\mathbb{S}$.

Example (spanning set vs. basis)

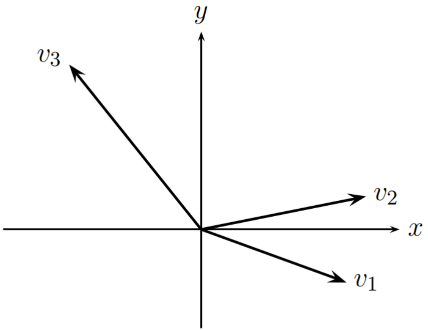


Figure 2.4: A spanning set v_1, v_2, v_3. Bases v_1, v_2 and v_1, v_3 and v_2, v_3.

Lemma (independent set and spanning set)

Let $\mathbb{S}$ be a space, $\mathcal{V} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$ be a linearly independent set and $\mathcal{U} = \{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_l\}$ be a spanning set of $\mathbb{S}$. Then $k \leq l$.

Proof by contradiction. Suppose $k > l$. Since $\mathcal{U}$ spans $\mathbb{S}$

$$\boldsymbol{v}_j = \sum_{i=1}^{l} a_{ij} \boldsymbol{u}_i, \ j = 1, \ldots, k$$

Define $\boldsymbol{A} = \{a_{ij}\}_{l \times k}$ and consider $\boldsymbol{Ax} = \boldsymbol{0}$. Since it is an under-determined system, $\exists\, \boldsymbol{c} \neq \boldsymbol{0}$ such that $\boldsymbol{Ac} = \boldsymbol{0}$. Then

$$\sum_{j=1}^{k} c_j \boldsymbol{v}_j = \sum_{j=1}^{k} c_j \sum_{i=1}^{l} a_{ij} \boldsymbol{u}_i = \sum_{i=1}^{l} \left(\sum_{j=1}^{k} a_{ij} c_j \right) \boldsymbol{u}_i = \boldsymbol{0}$$

This contradicts the assumption that $\mathcal{V}$ is linearly independent.

Theorem (size of a basis of a space)

Let $\mathbb{S}$ be a space. Every basis of $\mathbb{S}$ has the same number of vectors (a.k.a. cardinality).

Proof. Let $\mathcal{U} = \{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m\}$ and $\mathcal{V} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be bases of $\mathbb{S}$. Since $\mathcal{U}$ is a linearly independent set and $\mathcal{V}$ is a spanning set, we have

$$m \leq n$$

Since $\mathcal{V}$ is a linearly independent set and $\mathcal{U}$ is a spanning set, we also have

$$n \leq m$$

Hence

$$m = n$$

Definition (dimension)

The **dimension** of a space is the number of vectors in a basis of the space. Let $\mathbb{S}$ be a space and $\mathcal{V} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be a basis of $\mathbb{S}$.

- The dimension of $\mathbb{S}$ is n
- This is denoted by $\mathbf{dim}\,\mathbb{S} = n$

Suppose $\mathbf{dim}\,\mathbb{S} = n$.

- A linearly independent set of $\mathbb{S}$ has at most n vectors
- A spanning set of $\mathbb{S}$ has at least n vectors

Fundamental Subspaces of a Matrix

Definition (column space and row space)

Let $\boldsymbol{A}$ be a matrix.

- The **column space** of $\boldsymbol{A}$ is the space spanned by the column vectors of $\boldsymbol{A}$
- The **row space** is the space spanned by the row vectors

Let $\boldsymbol{A}$ be a matrix of order $m \times n$. We have

$$\mathbb{C}(\boldsymbol{A}) = \textbf{span}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n)$$

$$\mathbb{C}\left(\boldsymbol{A}^T\right) = \textbf{span}\left(\boldsymbol{a}_{1:}^T, \cdots, \boldsymbol{a}_{m:}^T\right)$$

Note

$$\mathbb{C}(\boldsymbol{A}) \subset \mathbb{R}^m,\ \mathbb{C}\left(\boldsymbol{A}^T\right) \subset \mathbb{R}^n$$

Example (column space)

$$A = \begin{bmatrix} 1 & 0 \\ 5 & 4 \\ 2 & 4 \end{bmatrix}$$

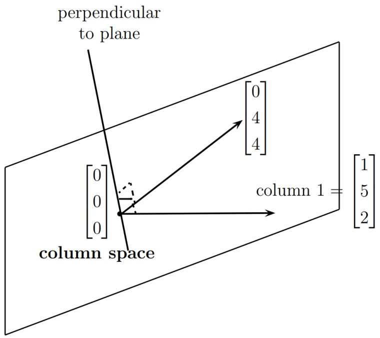


Figure 2.1: The column space $C(A)$, a plane in three-dimensional space.

Definition (nullspace and left nullspace)

Let $\boldsymbol{A}$ be a matrix.

- The **nullspace** of $\boldsymbol{A}$ is defined by

$$\mathbb{N}(\boldsymbol{A}) = \left\{ \boldsymbol{x} \,\middle|\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \right\}$$

- The **left nullspace** of $\boldsymbol{A}$ is defined by

$$\mathbb{N}\left(\boldsymbol{A}^T\right) = \left\{ \boldsymbol{y} \,\middle|\, \boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0} \right\}$$

Let $\boldsymbol{A}$ be a matrix of order $m \times n$. We have

$$\mathbb{N}(\boldsymbol{A}) \subset \mathbb{R}^n, \; \mathbb{N}(\boldsymbol{A}^T) \subset \mathbb{R}^m$$

$\mathbb{N}\left(\boldsymbol{A}^T\right)$ is called the left nullspace because

$$\boldsymbol{y} \in \mathbb{N}\left(\boldsymbol{A}^T\right) \;\Rightarrow\; \boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0} \;\Rightarrow\; \boldsymbol{y}^T\boldsymbol{A} = \boldsymbol{0}$$

Example (fundamental subspaces of a matrix)

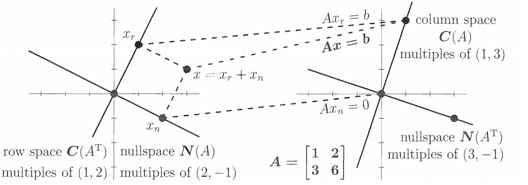


Figure 2.5: The four fundamental subspaces (lines) for the singular matrix A.

$$\mathbb{C}(\boldsymbol{A}) = \left\{ \boldsymbol{y} \middle| \boldsymbol{y} = c_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 6 \end{bmatrix} \right\} = \left\{ \boldsymbol{y} \middle| \boldsymbol{y} = c \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$$

$$\mathbb{C}\left(\boldsymbol{A}^T\right) = \left\{ \boldsymbol{x} \middle| \boldsymbol{x} = c_1 \begin{bmatrix} 1 & 2 \end{bmatrix}^T + c_2 \begin{bmatrix} 3 & 6 \end{bmatrix}^T \right\} = \left\{ \boldsymbol{x} \middle| \boldsymbol{x} = c \begin{bmatrix} 1 & 2 \end{bmatrix}^T \right\}$$

$$\mathbb{N}(\boldsymbol{A}) = \left\{ \boldsymbol{x} \middle| \boldsymbol{x} = c \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$$

$$\mathbb{N}\left(\boldsymbol{A}^T\right) = \left\{ \boldsymbol{y} \middle| \boldsymbol{y} = c \begin{bmatrix} -3 & 1 \end{bmatrix}^T \right\}$$

Definition (rank)

The **rank** of $\boldsymbol{A}$ is the number of pivots in the echelon matrix converted from $\boldsymbol{A}$.

- For example

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 6 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\textbf{rank}(\boldsymbol{A}) = 2$$

- Let $\boldsymbol{A}$ be a matrix of order $m \times n$. Then

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

Bases of Fundamental Subspaces

Let $\boldsymbol{A}$ be a matrix of order $m \times n$ with rank r. Let $\boldsymbol{U}$ be the echelon matrix converted from $\boldsymbol{A}$ so $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$.

- The r pivot columns of $\boldsymbol{A}$ constitute a basis of $\mathbb{C}(\boldsymbol{A})$
- The r pivot rows of $\boldsymbol{U}$ constitute a basis of $\mathbb{C}\left(\boldsymbol{A}^T\right)$
- The $(n-r)$ independent solutions of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ constitute a basis of $\mathbb{N}(\boldsymbol{A})$
- The last $(m-r)$ rows in $\boldsymbol{L}^{-1}$ where $\boldsymbol{U} = \boldsymbol{L}^{-1}\boldsymbol{A}$ constitute a basis of $\mathbb{N}\left(\boldsymbol{A}^T\right)$

Example (bases of fundamental subspaces)

Find the bases of the fundamental subspaces of

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix}$$

The echelon matrix $\boldsymbol{U}$ converted from $\boldsymbol{A}$ is

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \boldsymbol{U}$$

Note $\boldsymbol{U} = \boldsymbol{L}^{-1}\boldsymbol{A}$ where

$$\boldsymbol{L}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -2 & 1 \end{bmatrix}$$

- Row 1 and row 2 are the pivot rows

$$\mathbb{C}\left(\boldsymbol{A}^T\right): \left\{\begin{bmatrix}1 & 3 & 3 & 2\end{bmatrix}^T, \begin{bmatrix}0 & 0 & 3 & 3\end{bmatrix}^T\right\}$$

- Column 1 and column 3 are the pivot columns

$$\mathbb{C}(\boldsymbol{A}): \left\{\begin{bmatrix}1\\2\\-1\end{bmatrix}, \begin{bmatrix}3\\9\\3\end{bmatrix}\right\}$$

- For the left nullspace

$$\mathbb{N}\left(\boldsymbol{A}^T\right): \left\{\begin{bmatrix}5 & -2 & 1\end{bmatrix}^T\right\}$$

- For the nullspace

$$\mathbb{N}(\boldsymbol{A}): \left\{\begin{bmatrix}-3\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\-1\\1\end{bmatrix}\right\}$$

Theorem (fundamental theorem part i)

Let $\boldsymbol{A}$ be a matrix of order $m \times n$ and rank r.

- *The column space $\mathbb{C}(\boldsymbol{A})$ is of dimension r*
- *The row space $\mathbb{C}(\boldsymbol{A}^T)$ is of dimension r*
- *The nullspace $\mathbb{N}(\boldsymbol{A})$ is of dimension $n - r$*
- *The left nullspace $\mathbb{N}(\boldsymbol{A}^T)$ is of dimension $m - r$*

Corollary (row space and column space)

The row space and the column space of a matrix always have the same dimension.

Theorem (existence and uniqueness of solution)

Let $\boldsymbol{A}$ be of order $m \times n$ and rank r. Consider $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

- *If $\boldsymbol{b} \in \mathbb{C}(\boldsymbol{A})$, $\boldsymbol{b}$ is a linear combination of the column vectors of $\boldsymbol{A}$, so a solution exists*
- *If $r = m$, $\mathbb{C}(\boldsymbol{A}) = \mathbb{R}^m$, so solution exists for every $\boldsymbol{b}$*
- *If $r = n$, the column vectors are linearly dependent, so there is at most one solution for every $\boldsymbol{b}$*

Linear Transformation

Transformation and Linear Transformation

Let $\mathbb{D}$ (domain) and $\mathbb{R}$ (range) be vector spaces.

- A **transformation** from $\mathbb{D}$ to $\mathbb{R}$ maps a vector in $\mathbb{D}$ to a vector in $\mathbb{R}$. This is denoted by

$$\boldsymbol{T} : \mathbb{D} \mapsto \mathbb{R}$$

- Let $\boldsymbol{T} : \mathbb{D} \mapsto \mathbb{R}$ be a transformation. $\boldsymbol{T}$ is **linear** if

$$\boldsymbol{T}(c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2) = c_1\boldsymbol{T}(\boldsymbol{x}_1) + c_2\boldsymbol{T}(\boldsymbol{x}_2)$$

  for any scalars c_1, c_2 and $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{D}$.

Theorem (matrix and linear transformation)

A **linear transformation** *can be represented by* **matrix**.

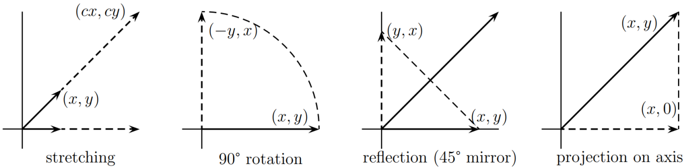


Figure 2.9: Transformations of the plane by four matrices.

$$\begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Representation of a vector by a column

Let $\mathbb{V}$ be a space of dimension n. Through a basis, a vector of $\mathbb{V}$ can be represented by a column of size n.

Let $\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be a basis of $\mathbb{V}$. For any $\boldsymbol{x} \in \mathbb{V}$, there exist $x_1, \cdots, x_n$ (to be shown to be unique) such that

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{v}_i$$

Thus $\boldsymbol{x}$ can be represented by a column of size n

$$[\boldsymbol{x}_{\mathcal{B}}] = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \Leftrightarrow \boldsymbol{x} = x_1\boldsymbol{v}_1 + \cdots + x_n\boldsymbol{v}_n$$

Theorem (uniqueness of representation)

Given basis, the representation of a vector is unique.

Let $\mathbb{V}$ be a space of dimension n and $\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be a basis of $\mathbb{V}$. For any $\boldsymbol{x} \in \mathbb{V}$, suppose

$$\boldsymbol{x} = a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n = b_1\boldsymbol{v}_1 + \cdots + b_n\boldsymbol{v}_n$$

Then

$$\boldsymbol{x} - \boldsymbol{x} = \sum_{i=1}^{n} a_i\boldsymbol{v}_i - \sum_{i=1}^{n} b_i\boldsymbol{v}_i = \sum_{i=1}^{n}(a_i - b_i)\boldsymbol{v}_i = \mathbf{0}$$

Since $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is linearly independent, the linear combination $\sum_{i=1}^{n}(a_i - b_i)\boldsymbol{v}_i$ must be trivial. Hence

$$a_i = b_i, \quad \forall i$$

CHARACTERIZATION OF LINEAR TRANSFORMATION

Let $\boldsymbol{T} : \mathbb{D} \mapsto \mathbb{R}$ be a linear transformation. $\boldsymbol{T}$ can be specified by the transformation by $\boldsymbol{T}$ for the vectors in a basis of $\mathbb{D}$.

Let $\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be a basis of $\mathbb{D}$.

- Let $\boldsymbol{T}(\boldsymbol{v}_j)$ be the transformation by $\boldsymbol{T}$ for $\boldsymbol{v}_j$
- For any $\boldsymbol{x} \in \mathbb{D}$, we have $\boldsymbol{x} = \sum_j x_j \boldsymbol{v}_j$
- By the linearity of $\boldsymbol{T}$, the transformation by $\boldsymbol{T}$ for $\boldsymbol{x}$ is

$$\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{T}\left(\sum_j x_j \boldsymbol{v}_j\right) = \sum_j x_j \boldsymbol{T}(\boldsymbol{v}_j)$$

REPRESENTATION OF LINEAR TRANSFORM BY MATRIX

Let $\boldsymbol{T} : \mathbb{D} \mapsto \mathbb{R}$ be linear, $\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be a basis of $\mathbb{D}$, and $\mathcal{B}' = \{\boldsymbol{v}'_1, \cdots, \boldsymbol{v}'_m\}$ be a basis of $\mathbb{R}$. Suppose

$$\boldsymbol{T}(\boldsymbol{v}_j) = \sum_{i=1}^{m} a_{ij}\boldsymbol{v}'_i, \quad j = 1, \cdots, n$$

- Coefficients a_{ij} specify the transformation by $\boldsymbol{T}$ for the basis vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$
- Define matrix

$$[\boldsymbol{T}_{\mathcal{B}\mathcal{B}'}] = \{a_{ij}\}$$

  The size is $m \times n$, where column j is decided by $\boldsymbol{T}(\boldsymbol{v}_j)$

The matrix $[\boldsymbol{T}_{\mathcal{B}\mathcal{B}'}]$ completely specifies $\boldsymbol{T} : \mathbb{D} \mapsto \mathbb{R}$ through basis $\mathcal{B}$ for $\mathbb{D}$ and basis $\mathcal{B}'$ for $\mathbb{R}$.

Linear Transform as Matrix Multiplication

Let $\boldsymbol{T} : \mathbb{D} \mapsto \mathbb{R}$ be linear, $\boldsymbol{x} \in \mathbb{D}$, and $\boldsymbol{y} \in \mathbb{R}$. Let $\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ be a basis of $\mathbb{D}$, and $\mathcal{B}' = \{\boldsymbol{v}'_1, \cdots, \boldsymbol{v}'_m\}$ be a basis of $\mathbb{R}$. We have

$$\boldsymbol{y} = \boldsymbol{T}(\boldsymbol{x}) \;\Leftrightarrow\; [\boldsymbol{y}_{\mathcal{B}'}] = [\boldsymbol{T}_{\mathcal{B}\mathcal{B}'}]\,[\boldsymbol{x}_{\mathcal{B}}]$$

Let $\boldsymbol{x} = \sum_{j=1}^{n} x_j \boldsymbol{v}_j$ and $\boldsymbol{y} = \sum_{i=1}^{m} y_i \boldsymbol{v}'_i$.

$$\boldsymbol{T}(\boldsymbol{x}) = \sum_{j=1}^{n} x_j \boldsymbol{T}(\boldsymbol{v}_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} \boldsymbol{v}'_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) \boldsymbol{v}'_i = \boldsymbol{y}$$

$$\Rightarrow\; y_i = \sum_j a_{ij} x_j,\; i = 1, \cdots, m$$

$$\Rightarrow\; [\boldsymbol{y}_{\mathcal{B}'}] = [\boldsymbol{T}_{\mathcal{B}\mathcal{B}'}]\,[\boldsymbol{x}_{\mathcal{B}}]$$

Example (differentiation: derivation of matrix)

Let $\boldsymbol{D} : \mathcal{P}_3 \mapsto \mathcal{P}_2$ be differentiation on polynomials. A basis of $\mathcal{P}_3$ is $\mathcal{B} = \{\boldsymbol{v}_1 = 1,\ \boldsymbol{v}_2 = t,\ \boldsymbol{v}_3 = t^2,\ \boldsymbol{v}_4 = t^3\}$, and a basis of $\mathcal{P}_2$ is $\mathcal{B}' = \{\boldsymbol{v}'_1 = 1,\ \boldsymbol{v}'_2 = t,\ \boldsymbol{v}'_3 = t^2\}$. We have

$$\begin{aligned}
\dot{\boldsymbol{v}}_1 &= 0 = 0\boldsymbol{v}'_1 + 0\boldsymbol{v}'_2 + 0\boldsymbol{v}'_3 \\
\dot{\boldsymbol{v}}_2 &= 1 = 1\boldsymbol{v}'_1 + 0\boldsymbol{v}'_2 + 0\boldsymbol{v}'_3 \\
\dot{\boldsymbol{v}}_3 &= 2t = 0\boldsymbol{v}'_1 + 2\boldsymbol{v}'_2 + 0\boldsymbol{v}'_3 \\
\dot{\boldsymbol{v}}_4 &= 3t^2 = 0\boldsymbol{v}'_1 + 0\boldsymbol{v}'_2 + 3\boldsymbol{v}'_3
\end{aligned}$$

The coefficients go to the columns of a matrix

$$[\boldsymbol{D}_{\mathcal{B}\mathcal{B}'}] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Example (alternative derivation of matrix)

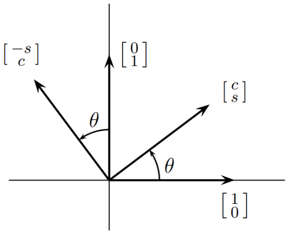
Suppose $\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + x_3\boldsymbol{v}_3 + x_4\boldsymbol{v}_4$, and $\boldsymbol{D}(\boldsymbol{x}) = \boldsymbol{y}$.

$$\begin{aligned}
\boldsymbol{y} = \dot{\boldsymbol{x}} &= x_1\dot{\boldsymbol{v}}_1 + x_2\dot{\boldsymbol{v}}_2 + x_3\dot{\boldsymbol{v}}_3 + x_4\dot{\boldsymbol{v}}_4 \\
&= x_1(0\boldsymbol{v}_1' + 0\boldsymbol{v}_2' + 0\boldsymbol{v}_3') + x_2(1\boldsymbol{v}_1' + 0\boldsymbol{v}_2' + 0\boldsymbol{v}_3') \\
&\quad + x_3(0\boldsymbol{v}_1' + 2\boldsymbol{v}_2' + 0\boldsymbol{v}_3') + x_4(0\boldsymbol{v}_1' + 0\boldsymbol{v}_2' + 3\boldsymbol{v}_3') \\
&= y_1\boldsymbol{v}_1' + y_2\boldsymbol{v}_2' + y_3\boldsymbol{v}_3'
\end{aligned}$$

$$\Rightarrow \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$\Rightarrow [\boldsymbol{D}_{\mathcal{B}\mathcal{B}'}] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Example (rotation and projection)



$$R = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

$$P = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}$$

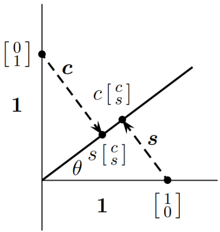


Figure 2.10: Rotation through θ (left). Projection onto the θ-line (right).

Example (rotation: derivation of matrix)

Let $\boldsymbol{R}$ be the **counter-clockwise rotation** by θ in $\mathbb{R}^2$. Let $\mathcal{B} = \mathcal{B}' = \left\{ \boldsymbol{e}_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}^T, \boldsymbol{e}_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}^T \right\}$. Rotation of the basis vectors leads to

$$\begin{aligned} \boldsymbol{R}(\boldsymbol{e}_1) &= \cos\theta\, \boldsymbol{e}_1 + \sin\theta\, \boldsymbol{e}_2 \\ \boldsymbol{R}(\boldsymbol{e}_2) &= \cos\left(\frac{\pi}{2} + \theta\right) \boldsymbol{e}_1 + \sin\left(\frac{\pi}{2} + \theta\right) \boldsymbol{e}_2 \\ &= -\sin\theta\, \boldsymbol{e}_1 + \cos\theta\, \boldsymbol{e}_2 \end{aligned}$$

Hence

$$[\boldsymbol{R}_{\mathcal{B}\mathcal{B}'}] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Example (alternative derivation of matrix)

Suppose $\boldsymbol{R}$ rotates $\boldsymbol{x} = [x_1 \;\; x_2]^T$ to $\boldsymbol{y} = [y_1 \;\; y_2]^T$. Let the angle between $\boldsymbol{x}$ and the horizontal axis be ϕ. Then the angle between $\boldsymbol{y}$ and the horizontal axis is $(\phi + \theta)$.

$$\begin{aligned}
x_1 &= |\boldsymbol{x}| \cos\phi, \;\; x_2 = |\boldsymbol{x}| \sin\phi, \;\; |\boldsymbol{y}| = |\boldsymbol{x}| \\
y_1 &= |\boldsymbol{y}| \cos(\theta + \phi) = |\boldsymbol{x}|(\cos\theta\cos\phi - \sin\theta\sin\phi) \\
&= \cos\theta\, x_1 - \sin\theta\, x_2 \\
y_2 &= |\boldsymbol{y}| \sin(\theta + \phi) = |\boldsymbol{x}|(\sin\theta\cos\phi + \cos\theta\sin\phi) \\
&= \sin\theta\, x_1 + \cos\theta\, x_2
\end{aligned}$$

$$\Rightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\Rightarrow [\boldsymbol{R}_{\mathcal{B}\mathcal{B}'}] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Example (projection: derivation of matrix)

Let $\boldsymbol{P}$ be the **projection** to line L, which is at angle θ to the horizontal axis. Projection of the basis vectors leads to

$$\boldsymbol{P}(\boldsymbol{e}_1) = \cos\theta(\cos\theta\,\boldsymbol{e}_1 + \sin\theta\,\boldsymbol{e}_2) = \cos^2\theta\,\boldsymbol{e}_1 + \cos\theta\sin\theta\,\boldsymbol{e}_2$$

$$\boldsymbol{P}(\boldsymbol{e}_2) = \sin\theta(\cos\theta\,\boldsymbol{e}_1 + \sin\theta\,\boldsymbol{e}_2) = \sin\theta\cos\theta\,\boldsymbol{e}_1 + \sin^2\theta\,\boldsymbol{e}_2$$

Hence

$$[\boldsymbol{P}_{\mathcal{B}\mathcal{B}'}] = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}$$

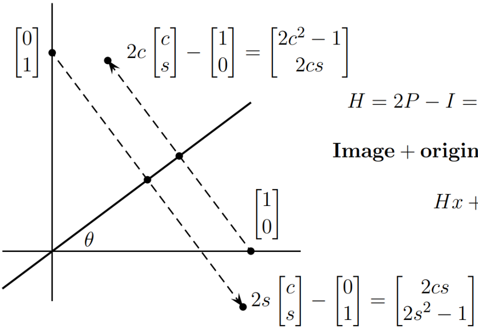
Example (alternative derivation of matrix)

Suppose $\boldsymbol{P}$ projects $\boldsymbol{x} = [x_1 \ \ x_2]^T$ to $\boldsymbol{y} = [y_1 \ \ y_2]^T$. Let the angle between $\boldsymbol{x}$ and the horizontal axis be ϕ. Then the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$ is $(\theta - \phi)$, and the angle between $\boldsymbol{y}$ and horizontal axis is θ.

$$\begin{aligned}
x_1 &= |\boldsymbol{x}|\cos\phi,\ \ x_2 = |\boldsymbol{x}|\sin\phi,\ \ |\boldsymbol{y}| = |\boldsymbol{x}|\cos(\theta-\phi) \\
y_1 &= |\boldsymbol{y}|\cos\theta = |\boldsymbol{x}|\cos(\theta-\phi)\cos\theta \\
&= |\boldsymbol{x}|(\cos\theta\cos\phi + \sin\theta\sin\phi)\cos\theta \\
&= \cos^2\theta\, x_1 + \sin\theta\cos\theta\, x_2 \\
y_2 &= |\boldsymbol{y}|\sin\theta = |\boldsymbol{x}|\cos(\theta-\phi)\sin\theta \\
&= |\boldsymbol{x}|(\cos\theta\cos\phi + \sin\theta\sin\phi)\sin\theta \\
&= \cos\theta\sin\theta\, x_1 + \sin^2\theta\, x_2
\end{aligned}$$

$$\Rightarrow \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Example (reflection)



$$H = 2P - I = \begin{bmatrix}2c^2-1 & 2cs\\ 2cs & 2s^2-1\end{bmatrix}$$

Image + original = 2 × projection

$$Hx + x = 2Px$$

Figure 2.11: Reflection through the θ-line: the geometry and the matrix.

Example (reflection: derivation of matrix)

Let $\boldsymbol{H}$ be the **reflection** with respect to line L, which is at angle θ to the horizontal axis. Reflection of the basis vectors leads to

$$\begin{aligned} \boldsymbol{H}(\boldsymbol{e}_1) &= \cos 2\theta\, \boldsymbol{e}_1 + \sin 2\theta\, \boldsymbol{e}_2 \\ \boldsymbol{H}(\boldsymbol{e}_2) &= \cos\left(2\theta - \frac{\pi}{2}\right) \boldsymbol{e}_1 + \sin\left(2\theta - \frac{\pi}{2}\right) \boldsymbol{e}_2 \\ &= \sin 2\theta\, \boldsymbol{e}_1 - \cos 2\theta\, \boldsymbol{e}_2 \end{aligned}$$

Hence

$$[\boldsymbol{H}_{\mathcal{B}\mathcal{B}'}] = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$$

Example (alternative derivation of matrix)

Suppose $\boldsymbol{H}$ reflects $\boldsymbol{x} = [x_1 \;\; x_2]^T$ to $\boldsymbol{y} = [y_1 \;\; y_2]^T$. Let the angle between $\boldsymbol{x}$ and the horizontal axis be ϕ. Then the angle between $\boldsymbol{y}$ and horizontal axis is $(2\theta - \phi)$.

$$
\begin{aligned}
x_1 &= |\boldsymbol{x}| \cos\phi, \;\; x_2 = |\boldsymbol{x}| \sin\phi, \;\; |\boldsymbol{y}| = |\boldsymbol{x}| \\
y_1 &= |\boldsymbol{y}| \cos(2\theta - \phi) = |\boldsymbol{x}| \cos 2\theta \cos\phi + |\boldsymbol{x}| \sin 2\theta \sin\phi \\
&= \cos 2\theta \, x_1 + \sin 2\theta \, x_2 \\
y_2 &= |\boldsymbol{y}| \sin(2\theta - \phi) = |\boldsymbol{x}| \sin 2\theta \cos\phi - |\boldsymbol{x}| \cos 2\theta \sin\phi \\
&= \sin 2\theta \, x_1 - \cos 2\theta \, x_2
\end{aligned}
$$

$$
\Rightarrow \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
$$

Summary of Linear Transformation

A linear transformation $\boldsymbol{T} : \mathbb{D} \mapsto \mathbb{R}$ from $\mathbb{D}$ of dimension n to $\mathbb{R}$ of dimension m is completely represented by a matrix of order $m \times n$. Such a matrix is constructed as follows.

- Find a basis of $\mathbb{D}$ and a basis of $\mathbb{R}$
- Apply $\boldsymbol{T}$ to a basis vector of $\mathbb{D}$ and express the result as a linear combination of the basis vectors of $\mathbb{R}$
- Put the coefficients in a column of a matrix
- Repeat until every basis vector of $\mathbb{D}$ has been processed