

Lecture 2: Linear Algebra

Yi, Yung (이웅)

Mathematics for Machine Learning
<https://yung-web.github.io/home/courses/mathml.html>
 KAIST EE

April 2, 2021

- (1) Systems of Linear Equations
- (2) Matrices
- (3) Solving Systems of Linear Equations
- (4) Vector Spaces
- (5) Linear Independence
- (6) Basis and Rank
- (7) Linear Mappings
- (8) Affine Spaces

April 2, 2021 1 / 56

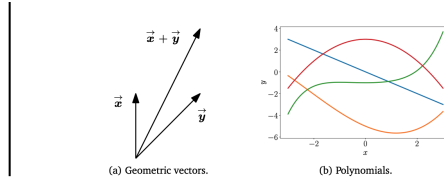
April 2, 2021 2 / 56

- (1) Systems of Linear Equations
- (2) Matrices
- (3) Solving Systems of Linear Equations
- (4) Vector Spaces
- (5) Linear Independence
- (6) Basis and Rank
- (7) Linear Mappings
- (8) Affine Spaces

- Scalars: $a, b, c, \alpha, \beta, \gamma$
- Vectors: $\mathbf{x}, \mathbf{y}, \mathbf{z}$
- Matrices: $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$
- Sets: $\mathcal{A}, \mathcal{B}, \mathcal{C}$
- (Ordered) tuple: $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$
- Matrix of column vectors: $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$ or $\mathbf{B} = (\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3)$
- Set of vectors: $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$
- $\mathbb{R}, \mathbb{C}, \mathbb{Z}, \mathbb{N}, \mathbb{R}^n$, etc

- Algebra: a set of objects and a set of rules or operations to manipulate those objects
- Linear algebra
 - Object: vectors \mathbf{v}
 - Operations: their additions ($\mathbf{v} + \mathbf{w}$) and scalar multiplication ($k\mathbf{v}$)
- Examples

- Geometric vectors
 - High school physics
- Polynomials
- Audio signals
- Elements of \mathbb{R}^n



L2(1)

April 2, 2021 5 / 56

- For unknown variables $(x_1, \dots, x_n) \in \mathbb{R}^n$,

$$a_{11}x_1 + \dots + a_{1n}x_n = b_1$$

$$\vdots$$

$$a_{m1}x_1 + \dots + a_{mn}x_n = b_m$$

- Three cases of solutions

- No solution

$$x_1 + x_2 + x_3 = 3$$

$$x_1 - x_2 + 2x_3 = 2$$

$$2x_1 + 3x_3 = 1$$

- Unique solution

$$x_1 + x_2 + x_3 = 3$$

$$x_1 - x_2 + 2x_3 = 2$$

$$x_2 + 3x_3 = 1$$

- Infinitely many solutions

$$x_1 + x_2 + x_3 = 3$$

$$x_1 - x_2 + 2x_3 = 2$$

$$2x_1 + 3x_3 = 5$$

- Question: Under what conditions, one of the above three cases occur.

L2(1)

April 2, 2021 6 / 56

- A collection of linear equations

$$a_{11}x_1 + \dots + a_{1n}x_n = b_1$$

$$\vdots$$

$$a_{m1}x_1 + \dots + a_{mn}x_n = b_m$$

- Matrix representations:

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \dots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \iff \underbrace{\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}}_{\mathbf{b}}$$

- Understanding \mathbf{A} is the key to answering various questions about this linear system $\mathbf{Ax} = \mathbf{b}$.

L2(1)

April 2, 2021 7 / 56

(1) Systems of Linear Equations

(2) **Matrices**

(3) Solving Systems of Linear Equations

(4) Vector Spaces

(5) Linear Independence

(6) Basis and Rank

(7) Linear Mappings

(8) Affine Spaces

L2(2)

April 2, 2021 8 / 56

- For two matrices $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{m \times n}$,

$$\mathbf{A} + \mathbf{B} := \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n}$$

- For two matrices $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{n \times k}$, the elements c_{ij} of the product $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{m \times k}$ is:

$$c_{ij} = \sum_{l=1}^n a_{il} b_{lj}, \quad i = 1, \dots, m, \quad j = 1, \dots, k.$$

- Example.** $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$, compute \mathbf{AB} and \mathbf{BA} .

L2(2)

April 2, 2021 9 / 56

- A square matrix¹ \mathbf{I}_n with $I_{ii} = 1$ and $I_{ij} = 0$ for $i \neq j$, where n is the number of rows and columns. For example,

$$\mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- Associativity:** For $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times p}$, $\mathbf{C} \in \mathbb{R}^{p \times q}$, $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- Distributivity:** For $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$, and $\mathbf{C}, \mathbf{D} \in \mathbb{R}^{n \times p}$,
(i) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ and (ii) $\mathbf{A}(\mathbf{C} + \mathbf{D}) = \mathbf{AC} + \mathbf{AD}$
- Multiplication with the identity matrix:** For $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{I}_m \mathbf{A} = \mathbf{A} \mathbf{I}_n = \mathbf{A}$

¹# of rows = # of cols
L2(2)

April 2, 2021 10 / 56

- For a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, \mathbf{B} is the **inverse** of \mathbf{A} , denoted by \mathbf{A}^{-1} , if

$$\mathbf{AB} = \mathbf{I}_n = \mathbf{BA}.$$

- Called **regular/invertible/nonsingular**, if it exists.
- If it exists, it is unique.
- (Q).** How to compute?
- For 2×2 matrix,

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

- For a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, the $\mathbf{B} \in \mathbb{R}^{n \times m}$ with $b_{ij} = a_{ji}$ is the **transpose** of \mathbf{A} , which we denote by \mathbf{A}^T .

$$\text{Example. For } \mathbf{A} = \begin{pmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{pmatrix},$$

$$\mathbf{A}^T = \begin{pmatrix} 0 & 1 & 0 \\ 2 & -1 & 1 \end{pmatrix}$$

- If $\mathbf{A} = \mathbf{A}^T$, \mathbf{A} is called **symmetric**.

L2(2)

April 2, 2021 11 / 56

- $\mathbf{AA}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A}$
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
- $(\mathbf{A} + \mathbf{B})^{-1} \neq \mathbf{A}^{-1} + \mathbf{B}^{-1}$
- $(\mathbf{A}^T)^T = \mathbf{A}$
- $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
- $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$
- If \mathbf{A} is invertible, so is \mathbf{A}^T .

L2(2)

April 2, 2021 12 / 56

- Multiplication by a scalar $\lambda \in \mathbb{R}$ to $\mathbf{A} \in \mathbb{R}^{m \times n}$

- **Example.** For $\mathbf{A} = \begin{pmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$, $3 \times \mathbf{A} = \begin{pmatrix} 0 & 6 \\ 3 & -3 \\ 0 & 3 \end{pmatrix}$

- **Associativity**

- $(\lambda\psi)\mathbf{C} = \lambda(\psi\mathbf{C})$
- $\lambda(\mathbf{BC}) = (\lambda\mathbf{B})\mathbf{C} = \mathbf{B}(\lambda\mathbf{C}) = (\mathbf{BC})\lambda$
- $(\lambda\mathbf{C})^T = \mathbf{C}^T\lambda^T = \mathbf{C}^T\lambda = \lambda\mathbf{C}^T$

- **Distributivity**

- $(\lambda + \psi)\mathbf{C} = \lambda\mathbf{C} + \psi\mathbf{C}$
- $\lambda(\mathbf{B} + \mathbf{C}) = \lambda\mathbf{B} + \lambda\mathbf{C}$

L2(2)

April 2, 2021 13 / 56

(1) Systems of Linear Equations

(2) Matrices

(3) Solving Systems of Linear Equations

(4) Vector Spaces

(5) Linear Independence

(6) Basis and Rank

(7) Linear Mappings

(8) Affine Spaces

L2(3)

April 2, 2021 14 / 56

Example

$$\begin{array}{rcl} -3x & + 2z & = -1 \\ x - 2y + 2z & = & -5/3 \\ -x - 4y + 6z & = & -13/3 \end{array}$$

- ρ_i : i -th equation
- Express the equation as its **augmented matrix**.

$$\begin{pmatrix} -3 & 0 & 2 & | & -1 \\ 1 & -2 & 2 & | & -5/3 \\ -1 & -4 & 6 & | & -13/3 \end{pmatrix} \xrightarrow{\begin{array}{l} (1/3)\rho_1 + \rho_2 \\ -(1/3)\rho_1 + \rho_3 \end{array}} \begin{pmatrix} -3 & 0 & 2 & | & -1 \\ 0 & -2 & 8/3 & | & -2 \\ 0 & -4 & 16/3 & | & -4 \end{pmatrix} \xrightarrow{-2\rho_2 + \rho_3} \begin{pmatrix} -3 & 0 & 2 & | & -1 \\ 0 & -2 & 8/3 & | & -2 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}$$

The two nonzero rows give $-3x + 2z = -1$ and $-2y + (8/3)z = -2$.

L2(3)

April 2, 2021 15 / 56

- Parametrizing $-3x + 2z = -1$ and $-2y + (8/3)z = -2$ gives:

$$x = (1/3) + (2/3)z$$

$$y = 1 + (4/3)z$$

$$z = z$$

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 2/3 \\ 4/3 \\ 1 \end{pmatrix} z \mid z \in \mathbb{R} \right\}$$

This helps us understand the set of solutions, e.g., each value of z gives a different solution.

	z	0	1	2	$-1/2$
solution	$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$	$\begin{pmatrix} 1/3 \\ 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 7/3 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 5/3 \\ 11/3 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1/3 \\ -1/2 \end{pmatrix}$

L2(3)

April 2, 2021 16 / 56

- The system $\begin{cases} x + 2y - z = 2 \\ 2x - y - 2z + w = 5 \end{cases}$ reduces in this way.

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 2 \\ 2 & -1 & -2 & 1 & 5 \end{array}\right) \xrightarrow{-2\rho_1+\rho_2} \left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 2 \\ 0 & -5 & 0 & 1 & 1 \end{array}\right)$$

- It has solutions of this form.

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 12/5 \\ -1/5 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} z + \begin{pmatrix} -2/5 \\ 1/5 \\ 0 \\ 1 \end{pmatrix} w \quad \text{for } z, w \in \mathbb{R}$$

- Note that taking $z = w = 0$ shows that the first vector is a **particular solution** of the system.

L2(3)

April 2, 2021 17 / 56

- General approach

- Find a particular solution to $\mathbf{Ax} = \mathbf{b}$
- Find all solutions to the homogeneous equation $\mathbf{Ax} = \mathbf{0}$
 - 0 is a trivial solution
- Combine the solutions from steps 1. and 2. to the general solution

- Questions: A formal algorithm that performs the above?

- Gauss-Jordan method: convert into a "beautiful" form (formally **reduced row-echelon** form)
- Elementary transformations: (i) row swapping (ii) multiply by a constant (iii) row addition

- Such a form allows an algorithmic way of solving linear equations

L2(3)

April 2, 2021 18 / 56

Example: Unique Solution

- Start as usual by getting echelon form.

$$\begin{array}{rcl} x + y - z = 2 & & x + y - z = 2 \\ 2x - y = -1 & \xrightarrow{-2\rho_1+\rho_2} & -3y + 2z = -5 \\ x - 2y + 2z = -1 & \xrightarrow{-1\rho_1+\rho_3} & -3y + 3z = -3 \end{array} \xrightarrow{-1\rho_2+\rho_3} \begin{array}{rcl} x + y - z = 2 & & x + y - z = 2 \\ -3y + 2z = -5 & & -3y + 2z = -5 \\ z = 2 & & z = 2 \end{array}$$

- Make all the leading entries one.

$$\begin{array}{rcl} x + y - z = 2 & & x + y - z = 2 \\ \xrightarrow{(-1/3)\rho_2} y - (2/3)z = 5/3 & & y - (2/3)z = 5/3 \\ z = 2 & & z = 2 \end{array}$$

- Finish by using the leading entries to eliminate upwards, until we can read off the solution.

$$\begin{array}{rcl} x + y - z = 2 & & x + y = 4 \\ y - (2/3)z = 5/3 & \xrightarrow{\rho_3+\rho_2} & y = 3 \\ z = 2 & \xrightarrow{(2/3)\rho_3+\rho_2} & z = 2 \end{array} \xrightarrow{-\rho_2+\rho_1} \begin{array}{rcl} x & = & 1 \\ y & = & 3 \\ z & = & 2 \end{array}$$

L2(3)

April 2, 2021 19 / 56

Example: Infinite Number of Solutions

$$\begin{array}{rcl} x - y - 2w = 2 & & x - y - 2w = 2 \\ x + y + 3z + w = 1 & & x + y + 3z + w = 1 \\ -y + z - w = 0 & & -y + z - w = 0 \end{array}$$

- Start by getting echelon form and turn the leading entries to 1's.

$$\begin{array}{rcl} \xrightarrow{-1\rho_1+\rho_2} \left(\begin{array}{cccc|c} 1 & -1 & 0 & -2 & 2 \\ 0 & 2 & 3 & 3 & -1 \\ 0 & -1 & 1 & -1 & 0 \end{array}\right) \\ \xrightarrow{(1/2)\rho_2+\rho_3} \left(\begin{array}{cccc|c} 1 & -1 & 0 & -2 & 2 \\ 0 & 2 & 3 & 3 & -1 \\ 0 & 0 & 5/2 & 1/2 & -1/2 \end{array}\right) \\ \xrightarrow{(1/2)\rho_2} \left(\begin{array}{cccc|c} 1 & -1 & 0 & -2 & 2 \\ 0 & 1 & 3/2 & 3/2 & -1/2 \\ 0 & 0 & 1 & 1/5 & -1/5 \end{array}\right) \end{array}$$

- Eliminate upwards.

$$\begin{array}{rcl} \xrightarrow{-(3/2)\rho_3+\rho_2} \left(\begin{array}{cccc|c} 1 & -1 & 0 & -2 & 2 \\ 0 & 1 & 0 & 6/5 & -1/5 \\ 0 & 0 & 1 & 1/5 & -1/5 \end{array}\right) \\ \xrightarrow{\rho_2+\rho_1} \left(\begin{array}{cccc|c} 1 & 0 & 0 & -4/5 & 9/5 \\ 0 & 1 & 0 & 6/5 & -1/5 \\ 0 & 0 & 1 & 1/5 & -1/5 \end{array}\right) \end{array}$$

- The parameterized solution set is:

$$\left\{ \begin{pmatrix} 9/5 \\ -1/5 \\ -1/5 \\ 0 \end{pmatrix} + \begin{pmatrix} 4/5 \\ -6/5 \\ -1/5 \\ 1 \end{pmatrix} w \mid w \in \mathbb{R} \right\}$$

L2(3)

April 2, 2021 20 / 56

		number of solutions of the homogeneous system	
		one	infinitely many
particular solution exists?	yes	unique solution	infinitely many solutions
	no	no solutions	no solutions

L2(3)

April 2, 2021 21 / 56

1. Pseudo-inverse

$$\mathbf{Ax} = \mathbf{b} \iff \mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b} \iff \mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$$

- $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$: *Moore-Penrose pseudo-inverse*
- many computations: matrix product, inverse, etc

2. Gaussian elimination

- intuitive and constructive way
- cubic complexity (in terms of # of simultaneous equations)

3. Iterative methods

- practical ways to solve indirectly
- (a) stationary iterative methods: Richardson method, Jacobi method, Gauss-Seidel method, successive over-relaxation method
- (b) Krylov subspace methods: conjugate gradients, generalized minimal residual, biconjugate gradients

L2(3)

April 2, 2021 22 / 56

- (1) Systems of Linear Equations
- (2) Matrices
- (3) Solving Systems of Linear Equations
- (4) **Vector Spaces**
- (5) Linear Independence
- (6) Basis and Rank
- (7) Linear Mappings
- (8) Affine Spaces

L2(4)

April 2, 2021 23 / 56

- A set \mathcal{G} and an operation $\otimes : \mathcal{G} \times \mathcal{G} \mapsto \mathcal{G}$. $G := (\mathcal{G}, \otimes)$ is called a **group**, if:
 1. **Closure**. $\forall x, y \in \mathcal{G}, x \otimes y \in \mathcal{G}$
 2. **Associativity**. $\forall x, y, z \in \mathcal{G}, (x \otimes y) \otimes z = x \otimes (y \otimes z)$
 3. **Neutral element**. $\exists e \in \mathcal{G}, \forall x \in \mathcal{G}, x \otimes e = x$ and $e \otimes x = x$
 4. **Inverse element**. $\forall x \in \mathcal{G}, \exists y \in \mathcal{G}, x \otimes y = e$ and $y \otimes x = e$. We often use $x^{-1} = y$.
- $G = (\mathcal{G}, \otimes)$ is an **Abelian group**, if the following is additionally met:
 - **Communicativity**. $\forall x, y \in \mathcal{G}, x \otimes y = y \otimes x$

L2(4)

April 2, 2021 24 / 56

- $(\mathbb{Z}, +)$ is an Abelian group
- $(\mathbb{N} \cup \{0\}, +)$ is not a group (because inverses are missing)
- (\mathbb{Z}, \cdot) is not a group
- (\mathbb{R}, \cdot) is not a group (because of no inverse for 0)
- $(\mathbb{R}^n, +)$, $(\mathbb{Z}^n, +)$ are Abelian, if $+$ is defined componentwise
- $(\mathbb{R}^{m \times n}, +)$ is Abelian (with componentwise $+$)
- $(\mathbb{R}^{n \times n}, \cdot)$
 - Closure and associativity follow directly
 - Neutral element: I_n
 - The inverse \mathbf{A}^{-1} may exist or not. So, generally, it is not a group. However, the set of invertible matrices in $\mathbb{R}^{n \times n}$ with matrix multiplication is a group, called **general linear group**.

L2(4)

April 2, 2021 25 / 56

Definition. A real-valued vector space $V = (\mathcal{V}, +, \cdot)$ is a set \mathcal{V} with two operations

$+: \mathcal{V} \times \mathcal{V} \mapsto \mathcal{V}$ (vector addition)

$\cdot: \mathbb{R} \times \mathcal{V} \mapsto \mathcal{V}$ (scalar multiplication),

where

1. $(\mathcal{V}, +)$ is an Abelian group

2. **Distributivity.**

$$\circ \forall \lambda \in \mathbb{R}, \mathbf{x}, \mathbf{y} \in \mathcal{V}, \lambda \cdot (\mathbf{x} + \mathbf{y}) = \lambda \cdot \mathbf{x} + \lambda \mathbf{y}$$

$$\circ \forall \lambda, \psi \in \mathbb{R}, \mathbf{x} \in \mathcal{V}, (\lambda + \psi) \cdot \mathbf{x} = \lambda \cdot \mathbf{x} + \psi \cdot \mathbf{x}$$

3. **Associativity.** $\forall \lambda, \psi \in \mathbb{R}, \mathbf{x} \in \mathcal{V}, \lambda \cdot (\psi \cdot \mathbf{x}) = (\lambda\psi) \cdot \mathbf{x}$

4. **Neutral element.** $\forall \mathbf{x} \in \mathcal{V}, 1 \cdot \mathbf{x} = \mathbf{x}$

L2(4)

April 2, 2021 26 / 56

- $\mathcal{V} = \mathbb{R}^n$ with
 - Vector addition: $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_n + y_n)$
 - Scalar multiplication: $\lambda \mathbf{x} = (\lambda x_1, \dots, \lambda x_n)$

- $\mathcal{V} = \mathbb{R}^{m \times n}$ with

$$\circ \text{Vector addition: } \mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

$$\circ \text{Scalar multiplication: } \lambda \mathbf{A} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}$$

L2(4)

April 2, 2021 27 / 56

Definition. Consider a vector space $V = (\mathcal{V}, +, \cdot)$ and $\mathcal{U} \subset \mathcal{V}$. Then, $U = (\mathcal{U}, +, \cdot)$ is called **vector subspace** (simply linear subspace or subspace) of V if U is a vector space with two operations ' $+$ ' and ' \cdot ' restricted to $\mathcal{U} \times \mathcal{U}$ and $\mathbb{R} \times \mathcal{U}$.

Examples

- For every vector space V , V and $\{0\}$ are the trivial subspaces.
- The solution set of $\mathbf{Ax} = 0$ is the subspace of \mathbb{R}^n .
- The solution of $\mathbf{Ax} = \mathbf{b}$ ($\mathbf{b} \neq 0$) is not a subspace of \mathbb{R}^n .
- The intersection of arbitrarily many subspaces is a subspace itself.

L2(4)

April 2, 2021 28 / 56

- (5) Systems of Linear Equations
- (5) Matrices
- (5) Solving Systems of Linear Equations
- (5) Vector Spaces
- (5) **Linear Independence**
- (5) Basis and Rank
- (5) Linear Mappings
- (5) Affine Spaces

L2(5)

April 2, 2021 29 / 56

- **Definition.** For a vector space V and vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$, every $\mathbf{v} \in V$ of the form $\mathbf{v} = \lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k$ with $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ is a **linear combination** of the vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$.
- **Definition.** If there is a non-trivial linear combination such that $0 = \sum_{i=1}^k \lambda_i \mathbf{x}_i$ with at least one $\lambda_i \neq 0$, the vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ are **linearly dependent**. If only the trivial solution exists, i.e., $\lambda_1 = \dots = \lambda_k = 0$, $\mathbf{x}_1, \dots, \mathbf{x}_n$ are **linearly independent**.
- **Meaning.** A set of linearly independent vectors consists of vectors that have no redundancy.
- **Useful fact.** The vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ are linearly dependent, iff (at least) one of them is a linear combination of the others.
 - $x - 2y = 2$ and $2x - 4y = 4$ are linearly dependent.

L2(5)

April 2, 2021 30 / 56

- Gauss elimination to get the row echelon form
- All column vectors are linearly independent iff all columns are pivot columns (why?).
- **Example.**

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \\ -3 \\ 4 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} -1 \\ -2 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & -2 \\ -3 & 0 & 1 \\ 4 & 2 & 1 \end{pmatrix} \rightsquigarrow \dots \rightsquigarrow \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

- Every column is a pivot column. Thus, $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are linearly independent.

L2(5)

April 2, 2021 31 / 56

- Vector space V with k linearly independent vectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k$
- m linear combinations $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. (Q) Are they linearly independent?

$$\begin{array}{l} \mathbf{x}_1 = \lambda_{11} \mathbf{b}_1 + \lambda_{21} \mathbf{b}_2 + \dots + \lambda_{k1} \mathbf{b}_k \\ \vdots \\ \mathbf{x}_m = \lambda_{1m} \mathbf{b}_1 + \lambda_{2m} \mathbf{b}_2 + \dots + \lambda_{km} \mathbf{b}_k \end{array} \quad \left| \quad \mathbf{x}_j = \overbrace{(\mathbf{b}_1, \dots, \mathbf{b}_k)}^{\mathbf{B}} \overbrace{\begin{pmatrix} \lambda_{1j} \\ \vdots \\ \lambda_{kj} \end{pmatrix}}^{\lambda_j}, \quad \mathbf{x}_j = \mathbf{B} \lambda_j$$

- $\sum_{j=1}^m \psi_j \mathbf{x}_j = \sum_{j=1}^m \psi_j \mathbf{B} \lambda_j = \mathbf{B} \sum_{j=1}^m \psi_j \lambda_j$
- $\{\mathbf{x}\}$ linearly independent $\iff \{\lambda\}$ linearly independent

L2(5)

April 2, 2021 32 / 56

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{b}_1 - 2\mathbf{b}_2 + \mathbf{b}_3 - \mathbf{b}_4 \\ \mathbf{x}_2 &= -4\mathbf{b}_1 - 2\mathbf{b}_2 + 4\mathbf{b}_4 \\ \mathbf{x}_3 &= 2\mathbf{b}_1 + 3\mathbf{b}_2 - \mathbf{b}_3 - 3\mathbf{b}_4 \\ \mathbf{x}_4 &= 17\mathbf{b}_1 - 10\mathbf{b}_2 + 11\mathbf{b}_3 + \mathbf{b}_4 \end{aligned}$$

$$\mathbf{A} = (\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3 \ \mathbf{x}_4) = \begin{pmatrix} 1 & -4 & 2 & 17 \\ -2 & -2 & 3 & -10 \\ 1 & 0 & -1 & 11 \\ -1 & -4 & -3 & 1 \end{pmatrix} \rightsquigarrow \dots \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & -15 \\ 0 & 0 & 1 & -18 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- The last column is not a pivot column. Thus, $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ are linearly dependent.

- (1) Systems of Linear Equations
- (2) Matrices
- (3) Solving Systems of Linear Equations
- (4) Vector Spaces
- (5) Linear Independence
- (6) **Basis and Rank**
- (7) Linear Mappings
- (8) Affine Spaces

L2(5)

April 2, 2021 33 / 56

L2(6)

April 2, 2021 34 / 56

- Definition.** A vector space $V = (\mathcal{V}, +, \cdot)$ and a set of vectors $\mathcal{A} = \{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subset \mathcal{V}$.
 - If every $\mathbf{v} \in \mathcal{V}$ can be expressed as a linear combination of $\mathbf{x}_1, \dots, \mathbf{x}_k$, \mathcal{A} is called a **generating set** of V .
 - The set of all linear combinations of \mathcal{A} is called the **span** of \mathcal{A} .
 - If \mathcal{A} spans the vector space V , we use $V = \text{span}[\mathcal{A}]$ or $V = \text{span}[\mathbf{x}_1, \dots, \mathbf{x}_k]$
- Definition.** The minimal generating set \mathcal{B} of V is called **basis** of V . We call each element of \mathcal{B} **basis vector**. The number of basis vectors is called **dimension** of V .
- Properties**
 - \mathcal{B} is a maximally² linearly independent set of vectors in V .
 - Every vector $\mathbf{x} \in V$ is a linear combination of \mathcal{B} , which is unique.

²Adding any other vector to this set will make it linearly dependent.

L2(6)

April 2, 2021 35 / 56

L2(6)

April 2, 2021 36 / 56

- Different bases \mathbb{R}^3

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\},$$

$$\mathcal{B}_3 = \left\{ \begin{pmatrix} 0.5 \\ 0.8 \\ 0.4 \end{pmatrix}, \begin{pmatrix} 1.8 \\ 0.3 \\ 0.3 \end{pmatrix}, \begin{pmatrix} -2.2 \\ -1.3 \\ 3.5 \end{pmatrix} \right\}$$

- Linearly independent, but not maximal. Thus, not a basis.

$$\mathcal{A} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ -4 \end{pmatrix} \right\}$$

- Want to find a basis of a subspace $U = \text{span}[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]$
 1. Construct a matrix $\mathbf{A} = (\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_m)$
 2. Find the row-echelon form of \mathbf{A} .
 3. Collect the pivot columns.
- Logic: Collect \mathbf{x}_i so that we have only trivial solution. Pivot columns tell us which set of vectors is linearly independent.
- See example 2.17 (pp. 35)

- **Definition.** The **rank** of $\mathbf{A} \in \mathbb{R}^{m \times n}$ denoted by $\text{rk}(\mathbf{A})$ is # of linearly independent columns
 - Same as the number of linearly independent rows

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{pmatrix} \rightsquigarrow \cdots \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus, $\text{rk}(\mathbf{A}) = 2$.

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

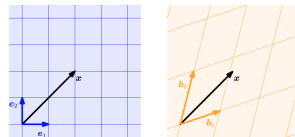
- The **columns** (resp. **rows**) of \mathbf{A} span a subspace U (resp. W) with $\dim(U) = \text{rk}(\mathbf{A})$ (resp. $\dim(W) = \text{rk}(\mathbf{A})$), and a basis of U (resp. W) can be found by Gauss elimination of \mathbf{A} (resp. \mathbf{A}^T).
- For all $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\text{rk}(\mathbf{A}) = n$, iff \mathbf{A} is regular (invertible).
- The linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ is solvable, iff $\text{rk}(\mathbf{A}) = \text{rk}(\mathbf{A}|\mathbf{b})$.
- For $\mathbf{A} \in \mathbb{R}^{m \times n}$, the subspace of solutions for $\mathbf{A}\mathbf{x} = \mathbf{0}$ possesses dimension $n - \text{rk}(\mathbf{A})$.
- $\mathbf{A} \in \mathbb{R}^{m \times n}$ has **full rank** if its rank equals the largest possible rank for a matrix of the same dimensions. The rank of the full-rank matrix \mathbf{A} is $\min(\# \text{ of cols}, \# \text{ of rows})$.

- (1) Systems of Linear Equations
- (2) Matrices
- (3) Solving Systems of Linear Equations
- (4) Vector Spaces
- (5) Linear Independence
- (6) Basis and Rank
- (7) **Linear Mappings**
- (8) Affine Spaces

- Interest: A mapping that preserves the structure of the vector space
- **Definition.** For vector spaces V, W , a mapping $\Phi : V \mapsto W$ is called a **linear mapping** (or homomorphism/linear transformation), if, for all $\mathbf{x}, \mathbf{y} \in V$ and all $\lambda \in \mathbb{R}$,
 - $\Phi(\mathbf{x} + \mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y})$
 - $\Phi(\lambda \mathbf{x}) = \lambda \Phi(\mathbf{x})$
- **Definition.** A mapping $\Phi : \mathcal{V} \mapsto \mathcal{W}$ is called
 - **Injective** (단사), if $\forall \mathbf{x}, \mathbf{y} \in \mathcal{V}, \Phi(\mathbf{x}) = \Phi(\mathbf{y}) \implies \mathbf{x} = \mathbf{y}$
 - **Surjective** (전사), if $\Phi(\mathcal{V}) = \mathcal{W}$
 - **Bijective** (전단사), if it is injective and surjective.

- For bijective mapping, there exists an inverse mapping Φ^{-1} .
- **Isomorphism** if Ψ is linear and bijective.
- **Theorem.** Vector spaces V and W are isomorphic, iff $\dim(V) = \dim(W)$.
 - Vector spaces of the same dimension are kind of the same thing.
- Other properties
 - For two linear mappings Φ and Ψ , $\Phi \circ \Psi$ is also a linear mapping.
 - If Φ is an isomorphism, so is Φ^{-1} .
 - For two linear mappings Φ and Ψ , $\Phi + \Psi$ and $\lambda \Psi$ for $\lambda \in \mathbb{R}$ are linear.

- A basis defines a coordinate system.



- Consider an ordered basis $B = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ of vector space V . Then, for any $\mathbf{x} \in V$, there exists a unique linear combination

$$\mathbf{x} = \alpha_1 \mathbf{b}_1 + \dots + \alpha_n \mathbf{b}_n.$$

- We call $\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ the coordinate of \mathbf{x} with respect to $B = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$.

- Consider a vector space V and two coordinate systems defined by $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ and $B' = (\mathbf{b}'_1, \dots, \mathbf{b}'_n)$.
- **Question.** For $(x_1, \dots, x_n)_B \rightarrow (y_1, \dots, y_n)_{B'}$, what is $(y_1, \dots, y_n)_{B'}$?
- **Theorem.**
$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (\mathbf{b}'_1 \dots \mathbf{b}'_n)^{-1} (\mathbf{b}_1 \dots \mathbf{b}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$
- Regard $\mathbf{A}_\Phi = (\mathbf{b}'_1 \dots \mathbf{b}'_n)^{-1} (\mathbf{b}_1 \dots \mathbf{b}_n)$ as a linear map

- $B = ((1, 0), (0, 1))$ and $B' = ((2, 1), (1, 2))$

- $(4, 2)_B \rightarrow (x, y)_{B'}$?

- Using $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (\mathbf{b}'_1 \ \dots \ \mathbf{b}'_n)^{-1} (\mathbf{b}_1 \ \dots \ \mathbf{b}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

L2(7)

April 2, 2021 45 / 56

- Two vector spaces

- V with basis $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ and W with basis $C = (\mathbf{c}_1, \dots, \mathbf{c}_m)$

- What is the coordinate in C -system for each basis \mathbf{b}_j ? For $j = 1, \dots, n$,

$$\mathbf{b}_j = \alpha_{1j}\mathbf{c}_1 + \dots + \alpha_{mj}\mathbf{c}_m \iff \mathbf{b}_j = (\mathbf{c}_1 \ \dots \ \mathbf{c}_m) \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}$$

$$\implies (\mathbf{b}_1 \ \dots \ \mathbf{b}_n) = (\mathbf{c}_1 \ \dots \ \mathbf{c}_m) \overbrace{\begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}}^{\mathbf{A}_\Phi}$$

- $\hat{x} = \mathbf{A}_\Phi \hat{y}$, where \hat{x} is the vector w.r.t B and \hat{y} is the vector w.r.t. C

L2(7)

April 2, 2021 46 / 56

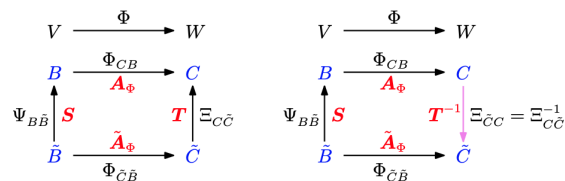
Basis Change: General Case

- For linear mapping $\Phi : V \mapsto W$, consider bases B, B' of V and C, C' of W
 $B = (\mathbf{b}_1 \ \dots \ \mathbf{b}_n)$, $B' = (\mathbf{b}'_1 \ \dots \ \mathbf{b}'_n)$ $C = (\mathbf{c}_1 \ \dots \ \mathbf{c}_m)$, $C' = (\mathbf{c}'_1 \ \dots \ \mathbf{c}'_m)$.

- (inter) transformation matrices \mathbf{A}_Φ from B to C and \mathbf{A}'_Φ from B' to C'

- (intra) transformation matrices S from B' to B and T from C' to C

- **Theorem.** $\mathbf{A}'_\Phi = T^{-1} \mathbf{A}_\Phi S$



L2(7)

April 2, 2021 47 / 56

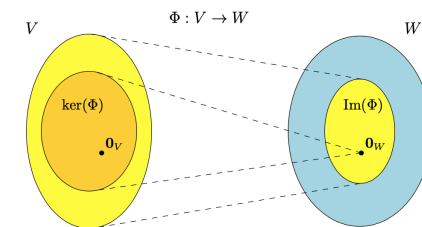
Image and Kernel

- Consider a linear mapping $\Phi : V \mapsto W$. The **kernel** (or **null space**) is the set of vectors in V that maps to $0 \in W$ (i.e., neutral element).

Definition. $\ker(\Phi) := \Phi^{-1}(0_W) = \{\mathbf{v} \in V : \Phi(\mathbf{v}) = 0_W\}$

- **Image/range:** set of vectors $w \in W$ that can be reached by Φ from any vector in V

- V : **domain**, W : **codomain**



L2(7)

April 2, 2021 48 / 56

- $0_V \in \ker(\Phi)$ (because $\Phi(0_V) = 0_W$)
- Both $\text{Im}(\Phi)$ and $\ker(\Phi)$ are subspaces of W and V , respectively.
- Φ is one-to-one (injective) $\iff \ker(\Phi) = \{0\}$ (i.e., only 0 is mapped to 0)
- Since Φ is a linear mapping, there exists $\mathbf{A} \in \mathbb{R}^{m \times n}$ such that $\Phi : \mathbf{x} \mapsto \mathbf{Ax}$. Then, $\text{Im}(\Phi) = \text{column space of } \mathbf{A}$ which is the span of column vectors of \mathbf{A} .
- $\text{rk}(\mathbf{A}) = \dim(\text{Im}(\Phi))$
- $\ker(\Phi)$ is the solution set of the homogeneous system of linear equations $\mathbf{Ax} = 0$

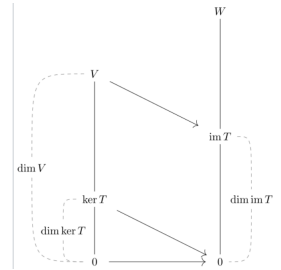
Theorem.

$$\dim(\ker(\Phi)) + \dim(\text{Im}(\Phi)) = \dim(V)$$

- If $\dim(\text{Im}(\Phi)) < \dim(V)$, the kernel contains more than just 0.
- If $\dim(\text{Im}(\Phi)) < \dim(V)$, $\mathbf{A}_\Phi \mathbf{x} = 0$ has infinitely many solutions.
- If $\dim(V) = \dim(W)$ (e.g., $V = W = \mathbb{R}^n$), the followings are equivalent: Φ is
 - (1) injective, (2) surjective, (3) bijective,
 - In this case, Φ defines $\mathbf{y} = \mathbf{Ax}$, where \mathbf{A} is regular.
- **Simplified version.** For $\mathbf{A} \in \mathbb{R}^{m \times n}$,

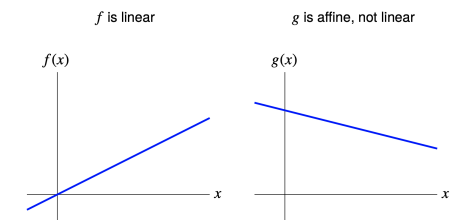
$$\text{rk}(\mathbf{A}) + \text{nullity}(\mathbf{A}) = n$$

²Nullity: the dimension of null space (kernel)



- (1) Systems of Linear Equations
- (2) Matrices
- (3) Solving Systems of Linear Equations
- (4) Vector Spaces
- (5) Linear Independence
- (6) Basis and Rank
- (7) Linear Mappings
- (8) **Affine Spaces**

- **linear function:** $f(x) = ax$
- **affine function:** $f(x) = ax + b$
- sometimes (ignorant) people refer to affine functions as linear



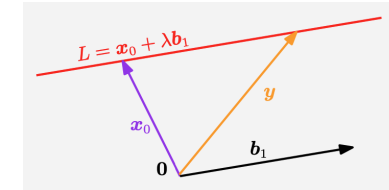
- Spaces that are offset from the origin. Not a vector space.
- **Definition.** Consider a vector space V , $\mathbf{x}_0 \in V$ and a subspace $U \subset V$. Then, the subset $L = \mathbf{x}_0 + U := \{\mathbf{x}_0 + \mathbf{u} : \mathbf{u} \in U\}$ is called **affine subspace** or **linear manifold** of V .
- U is called **direction** or **direction space**, and \mathbf{x}_0 is **support** point.
- An affine subspace is not a vector subspace of V for $\mathbf{x}_0 \notin U$.
- **Parametric equation.** A k -dimensional affine space $L = \mathbf{x}_0 + U$. If $(\mathbf{b}_1, \dots, \mathbf{b}_k)$ is an ordered basis of U , any element $\mathbf{x} \in L$ can be uniquely described as

$$\mathbf{x} = \mathbf{x}_0 + \lambda_1 \mathbf{b}_1 + \dots + \lambda_k \mathbf{b}_k, \quad \lambda_1, \dots, \lambda_k \in \mathbb{R}$$

L2(8)

April 2, 2021 53 / 56

- In \mathbb{R}^2 , one-dimensional affine subspace: **line**. $\mathbf{y} = \mathbf{x}_0 + \lambda \mathbf{b}_1$. $U = \text{span}[\mathbf{b}_1]$
- In \mathbb{R}^3 , two-dimensional affine subspace: **plane**. $\mathbf{y} = \mathbf{x}_0 + \lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2$. $U = \text{span}[\mathbf{b}_1, \mathbf{b}_2]$
- In \mathbb{R}^n , $(n-1)$ -dimensional affine subspace: **hyperplane**. $\mathbf{y} = \mathbf{x}_0 + \sum_{k=1}^{n-1} \lambda_k \mathbf{b}_k$.
 $U = \text{span}[\mathbf{b}_1, \dots, \mathbf{b}_n]$



- For a linear mapping $\Phi : V \mapsto W$ and a vector $\mathbf{a} \in W$, the mapping $\phi : V \mapsto W$ with $\phi(\mathbf{x}) = \mathbf{a} + \Phi(\mathbf{x})$ is an **affine mapping** from V to W . The vector \mathbf{a} is called the **translation vector**.

L2(8)

April 2, 2021 54 / 56

Questions?

1)

L2(8)

April 2, 2021 55 / 56

L2(8)

April 2, 2021 56 / 56