

Lecture 2: Linear Algebra

Yi, Yung (이용)

Mathematics for Machine Learning

 $\label{limits} \begin{tabular}{ll} $https://yung-web.github.io/home/courses/mathml.html \\ & KAIST \ EE \end{tabular}$

April 7, 2021

April 7, 2021 1 / 56

Basic Notations



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

Roadmap



- (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 7, 2021 3 / 56

Roadmap



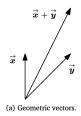
- (1) Systems of Linear Equations
- (2) Matrices
- (3) Solving Systems of Linear Equations
- (4) Vector Spaces
- (5) Linear Independence
- (6) Basis and Rank
- (7) Linear Mappings
- (8) Affine Spaces

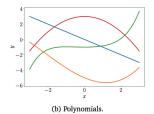
L2(1) April 7, 2021 4 / 56

Linear Algebra



- Algebra: a set of objects and a set of rules or operations to manipulate those objects
- Linear algebra
 - Object: vectors 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 7, 2021 5 / 56

System of Linear Equations



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

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

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

- 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)

Matrix Representation



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:

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

• Understanding \boldsymbol{A} is the key to answering various questions about this linear system $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$.

L2(1) April 7, 2021 7 / 56

Roadmap



- (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)

Matrix: Addition and Multiplication



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

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

• 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 $\mathbf{C} = \mathbf{A}\mathbf{B} \in \mathbb{R}^{m \times k}$ is:

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

• 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 7, 2021 9 / 56

Identity Matrix and Matrix Properties



• A square matrix¹ 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,

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- Associativity: For $\pmb{A} \in \mathbb{R}^{m \times n}$, $\pmb{B} \in \mathbb{R}^{n \times p}$, $\pmb{C} \in \mathbb{R}^{p \times q}$, $\pmb{(AB)C} = \pmb{A(BC)}$
- Distributivity: For $A, B \in \mathbb{R}^{m \times n}$, and $C, D \in \mathbb{R}^{n \times p}$, (i) (A + B)C = AC + BC and (ii) A(C + D) = AC + 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 $^{\text{L2(2)}}$

Inverse and Transpose



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

$$AB = I_n = BA$$
.

- Called regular/invertible/nonsingular, if it exists.
- If it exists, it is unique.
- 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 $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times m}$ with $b_{ij} = a_{ji}$ is the transpose of \mathbf{A} , which we denote by \mathbf{A}^{T} .
- Example. For $\mathbf{A} = \begin{pmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$,

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

• If $\mathbf{A} = \mathbf{A}^{\mathsf{T}}$, \mathbf{A} is called symmetric.

L2(2) April 7, 2021 11 / 56

Inverse and Transpose: More Properties



•
$$AA^{-1} = I = A^{-1}A$$

•
$$(AB)^{-1} = B^{-1}A^{-1}$$

•
$$(A + B)^{-1} \neq A^{-1} + B^{-1}$$

$$\bullet (\mathbf{A}^{\mathsf{T}})^{\mathsf{T}} = \mathbf{A}$$

$$\bullet \ (\mathbf{A} + \mathbf{B})^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}} + \mathbf{B}^{\mathsf{T}}$$

•
$$(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$$

• If \boldsymbol{A} is invertible, so is $\boldsymbol{A}^{\mathsf{T}}$.

Scalar Multiplication



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

• 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

$$\circ (\lambda \psi) \mathbf{C} = \lambda (\psi \mathbf{C})$$

•
$$\lambda(BC) = (\lambda B)C = B(\lambda C) = (BC)\lambda$$

$$(\lambda \mathbf{C})^{\mathsf{T}} = \mathbf{C}^{\mathsf{T}} \lambda^{\mathsf{T}} = \mathbf{C}^{\mathsf{T}} \lambda = \lambda \mathbf{C}^{\mathsf{T}}$$

Distributivity

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

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

L2(2) April 7, 2021 13 / 56

Roadmap



- (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

Example



$$-3x + 2z = -1$$

$$x - 2y + 2z = -5/3$$

$$-x - 4y + 6z = -13/3$$

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

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

$$\xrightarrow{-2\rho_2+\rho_3}
\begin{pmatrix}
-3 & 0 & 2 & | & -1 \\
0 & -2 & 8/3 & | & -2 \\
0 & 0 & 0 & | & 0
\end{pmatrix}$$

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

KAIST EE

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

$$\begin{aligned}
 x &= (1/3) + (2/3)z \\
 y &= 1 + (4/3)z \\
 z &= z
 \end{aligned}
 \qquad
 \begin{cases}
 (x) \\
 (y) \\
 z) = (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3) \\
 (1/3)$$

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

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

L2(3) April 7, 2021 16 / 56

¹Examples from this slide to the next several slides come from Jim Hefferson's Linear Algebra book.

L2(3)

April 7, 2021 15 / 56

Form of solution sets



$$\begin{pmatrix} 1 & 2 & -1 & 0 & 2 \\ 2 & -1 & -2 & 1 & 5 \end{pmatrix} \quad \stackrel{-2\rho_1+\rho_2}{\longrightarrow} \quad \begin{pmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & -5 & 0 & 1 & 1 \end{pmatrix}$$

• It has solutions of this form.

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

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

L2(3) April 7, 2021 17 / 56

General = Particular + Homogeneous



- General approach
 - 1. Find a particular solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$
 - 2. Find all solutions to the homogeneous equation Ax = 0
 - ▶ 0 is a trivial solution
 - 3. 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 7, 2021 18 / 56

Example: Unique Solution



• Start as usual by getting echelon form.

Make all the leading entries one.

$$(-1/3)\rho_2 \xrightarrow{(-1/3)\rho_2} x + y - z = 2$$

$$y - (2/3)z = 5/3$$

$$z = 2$$

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

L2(3) April 7, 2021 19 / 56

Example: Infinite Number of Solutions

KAIST EE

$$x-y - 2w = 2$$

$$x+y+3z+ w = 1$$

$$-y+z-w=0$$

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

• Eliminate upwards.

• 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 7, 2021 20 / 56

Cases of Solution Sets



number of solutions of the homogeneous system

particular solution exists?

one infinitely many unique solution solutions solutions

no no no solutions solutions

L2(3) April 7, 2021 21 / 56

Algorithms for Solving System of Linear Equations



1. Pseudo-inverse

$$\mathbf{A}\mathbf{x} = \mathbf{b} \Longleftrightarrow \mathbf{A}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \mathbf{A}^{\mathsf{T}}\mathbf{b} \Longleftrightarrow \mathbf{x} = (\mathbf{A}^{\mathsf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{T}}\mathbf{b}$$

$$\circ (\mathbf{A}^{\mathsf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{T}} : Moore-Penrose pseudo-inverse$$

- o 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, Gaus-Seidel method, successive over-relaxation method
 - (b) Krylov subspace methods: conjugate gradients, generalized minimal residual, biconjugate gradients

L2(3) April 7, 2021 22 / 56

Roadmap



- (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 7, 2021 23 / 56

Group



- 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. $\exists e \in \mathcal{G}, \ \forall x \in \mathcal{G}, \ x \otimes e = x \ \text{and} \ e \otimes x = x$
 - 4. Inverse element. $\forall x \in \mathcal{G}, \exists y \in \mathcal{G}, x \otimes y = e \text{ and } y \otimes x = e.$ 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)

Examples



- $(\mathbb{Z},+)$ 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 7, 2021 25 / 56

Vector Spaces



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

- (a) $+: \mathcal{V} \times \mathcal{V} \mapsto \mathcal{V}$ (vector addition)
- (b) $\cdot: \mathbb{R} \times \mathcal{V} \mapsto \mathcal{V}$ (scalar multiplication),

where

- 1. $(\mathcal{V},+)$ is an Abelian group
- 2. Distributivity.
 - $\circ \ \forall \lambda \in \mathbb{R}, \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}, \ \lambda \cdot (\boldsymbol{x} + \boldsymbol{y}) = \lambda \cdot \boldsymbol{x} + \lambda \boldsymbol{y}$
 - $\circ \ \forall \lambda, \psi \in \mathbb{R}, \mathbf{x} \in \mathcal{V}, (\lambda + \psi) \cdot \mathbf{x} = \lambda \cdot \mathbf{x} + \psi \cdot \mathbf{x}$
- 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. $\forall x \in \mathcal{V}, 1 \cdot x = x$

L2(4) April 7, 2021 26 / 56

Example



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

$$\mathcal{V} = \mathbb{R}^{m \times n}$$
 with \circ 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 Scalar multiplication: $\lambda \mathbf{A} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}$

• 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 7, 2021 27 / 56

Vector Subspaces



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 '.' 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 $\mathbf{A}\mathbf{x} = 0$ is the subspace of \mathbb{R}^n .
- The solution of $\mathbf{A}\mathbf{x} = \mathbf{b} \ (\mathbf{b} \neq 0)$ is not a subspace of \mathbb{R}^n .
- The intersection of arbitrarily many subspaces is a subspace itself.

L2(4) April 7, 2021 28 / 56

Roadmap



- (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 7, 2021 29 / 56

Linear Independence



- Definition. For a vector space V and vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in V$, every $\mathbf{v} \in V$ of the form $\mathbf{v} = \lambda_1 \mathbf{x}_1 + \cdots + \lambda_k \mathbf{x}_k$ with $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ is a linear combination of the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in V$.
- 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 $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly dependent. If only the trivial solution exists, i.e., $\lambda_1 = \ldots = \lambda_k = 0, \mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent.
- Meaning. A set of linearly independent vectors consists of vectors that have no redundancy.
- Useful fact. The vectors $\{x_1, \dots, x_n\}$ 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 7, 2021 30 / 56

Checking Linear Independence



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

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

$$\begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & -2 \\ -3 & 0 & 1 \\ 4 & 2 & 1 \end{pmatrix} \quad \stackrel{\sim}{\sim} \cdots \sim \quad \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

• Every column is a pivot column. Thus, x_1 , x_2 , x_3 are linearly independent.

L2(5) April 7, 2021 31 / 56

Linear Combinations of Linearly Independent Vectors



- Vector space V with k linearly independent vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_k$
- m linear combinations x_1, x_2, \dots, x_m . (Q) Are they linearly independent?

$$\mathbf{x}_1 = \lambda_{11}\mathbf{b}_1 + \lambda_{21}\mathbf{b}_2 + \dots + \lambda_{k1}\mathbf{b}_k$$

$$\vdots$$

$$\mathbf{x}_m = \lambda_{1m}\mathbf{b}_1 + \lambda_{2m}\mathbf{b}_2 + \dots + \lambda_{km}\mathbf{b}_k$$

$$oldsymbol{x}_j = \overbrace{\left(oldsymbol{b}_1, \; \cdots, \; oldsymbol{b}_k
ight)}^{oldsymbol{\mathcal{B}}} \overbrace{\left(egin{array}{c} \lambda_{1j} \ dots \ \lambda_{kj} \end{array}
ight)}^{oldsymbol{\lambda}_j}, \quad oldsymbol{x}_j = oldsymbol{B} oldsymbol{\lambda}_j$$

•
$$\sum_{j=1}^{m} \psi_j \mathbf{x}_j = \sum_{j=1}^{m} \psi_j \mathbf{B} \lambda_j = \mathbf{B} \sum_{j=1}^{m} \psi_j \lambda_j$$

• $\{m{x}\}$ linearly independent $\Longleftrightarrow \{m{\lambda}\}$ linearly independent

L2(5) April 7, 2021 32 / 56

Example



$$x_1 = b_1 - 2b_2 + b_3 - b_4$$

 $x_2 = -4b_1 - 2b_2 + 4b_4$
 $x_3 = 2b_1 + 3b_2 - b_3 - 3b_4$
 $x_4 = 17b_1 - 10b_2 + 11b_3 + b_4$

$$\mathbf{A} = \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{pmatrix} = \begin{pmatrix} 1 & -4 & 2 & 17 \\ -2 & -2 & 3 & -10 \\ 1 & 0 & -1 & 11 \\ -1 & -4 & -3 & 1 \end{pmatrix} \rightsquigarrow \cdots \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & -15 \\ 0 & 0 & 1 & -18 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

• The last column is not a pivot column. Thus, x_1, x_2, x_3, x_3 are linearly dependent.

L2(5) April 7, 2021 33 / 56

Roadmap



- (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(6) April 7, 2021 34 / 56

Generating Set and Basis



- Definition. A vector space $V = (\mathcal{V}, +, \cdot)$ and a set of vectors $\mathcal{A} = \{x_1, \dots, x_k\} \subset \mathcal{V}$.
 - If every $v \in \mathcal{V}$ 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 A is called the span of 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
 - \circ ${\cal B}$ is a maximally 2 linearly independent set of vectors in V.
 - Every vector $x \in V$ is a linear combination of \mathcal{B} , which is unique.

April 7, 2021 35 / 56

Examples



• Different bases \mathbb{R}^3

$$\mathcal{B}_{1} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \mathcal{B}_{2} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\},$$

$$\begin{pmatrix} 0.5 \\ 1.8 \\ -2.2 \\ \end{pmatrix}$$

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

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

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

L2(6) April 7, 2021 36 / 56

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

Determining a Basis



- 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 **A**.
 - 3. Collect the pivot columns.
- Logic: Collect 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)

L2(6) April 7, 2021 37 / 56

Rank (1)



- Definition. The rank of $\mathbf{A} \in \mathbb{R}^{m \times n}$ denoted by $\operatorname{rk}(\mathbf{A})$ 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, $rk(\mathbf{A}) = 2$.

• $rk(\mathbf{A}) = rk(\mathbf{A}^T)$

L2(6) April 7, 2021 38 / 56

Rank (2)



- The columns (resp. rows) of \mathbf{A} span a subspace U (resp. W) with $\dim(U) = \operatorname{rk}(\mathbf{A})$ (resp. $\dim(W) = \operatorname{rk}(\mathbf{A})$), and a basis of U (resp. W) can be found by Gauss elimination of \mathbf{A} (resp. \mathbf{A}^{T}).
- For all $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\operatorname{rk}(\mathbf{A}) = n$, iff \mathbf{A} is regular (invertible).
- The linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ is solvable, iff $\mathrm{rk}(\mathbf{A}) = \mathrm{rk}(\mathbf{A}|\mathbf{b})$.
- For $\mathbf{A} \in \mathbb{R}^{m \times n}$, the subspace of solutions for $\mathbf{A}\mathbf{x} = 0$ 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(# of cols, # of rows).

L2(6) April 7, 2021 39 / 56

Roadmap



- (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(7) April 7, 2021 40 / 56

Linear Mapping (1)



- 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 $x, y \in V$ and all $\lambda \in \mathbb{R}$,
- Definition. A mapping $\Phi: \mathcal{V} \mapsto \mathcal{W}$ is called
 - \circ Injective (단사), if $\forall x, y \in \mathcal{V}$, $\Phi(x) = \Phi(y) \implies x = y$
 - \circ Surjective (전사), if $\Phi(\mathcal{V}) = \mathcal{W}$
 - Bijective (전단사), if it is injenctive and surjective.

L2(7) April 7, 2021 41 / 56

Linear Mapping (2)



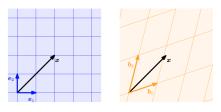
- For bjective mapping, there exists an inverse mapping Φ^{-1} .
- Isomorphism if Ψ is linear and bijective.
- Theorem. Vector spaces V and W are isomorphic, iff $\dim(V) = \dim(W)$.
 - Vector spaces of the same dimension are kind of the same thing.
- Other properties
 - $\circ~$ For two linear mappings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

L2(7) April 7, 2021 42 / 56

Coordinates



 A basis defines a coordinate system.



- Consider an ordered basis $B = (\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n)$ of vector space V. Then, for any $\boldsymbol{x} \in V$, there exists a unique linear combination $\boldsymbol{x} = \alpha_1 \boldsymbol{b}_1 + \dots + \alpha_n \boldsymbol{b}_n$.
- We call $\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ the coordinate of ${\pmb x}$ with respect to $B = ({\pmb b}_1, {\pmb b}_2, \dots, {\pmb b}_n).$
- Basis change ⇒ Coordinate change

L2(7) April 7, 2021 43 / 56

Basis Change



- Consider a vector space V and two coordinate systems defined by $B = (\boldsymbol{b}_1, \dots, \boldsymbol{b}_n)$ and $B' = (\boldsymbol{b}_1', \dots, \boldsymbol{b}_n')$.
- Question. For $(x_1,\ldots,x_n)_B \to (y_1,\ldots,y_n)_{B'}$, what is $(y_1,\ldots,y_n)_{B'}$?

• Theorem.
$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_1' & \dots & \boldsymbol{b}_n' \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{b}_1 & \dots & \boldsymbol{b}_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

• Regard $m{A}_{\Phi} = m{b}_1' \ \dots \ m{b}_n'^{-1} m{b}_1 \ \dots \ m{b}_n$ as a linear map

L2(7) April 7, 2021 44 / 56

Example



- B = ((1,0),(0,1) and B' = ((2,1),(1,2))
- $(4,2)_B \to (x,y)_{B'}$?

• Using
$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \mathbf{b}'_1 & \dots & \mathbf{b}'_n \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$
, $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$

L2(7) April 7, 2021 45 / 56

Transformation Matrix



- Two vector spaces
 - V with basis $B = (\boldsymbol{b}_1, \dots, \boldsymbol{b}_n)$ and W with basis $C = (\boldsymbol{c}_1, \dots, \boldsymbol{c}_m)$
- What is the coordinate in *C*-system for each basis b_i ? For $j=1,\ldots,n$,

$$\mathbf{b}_{j} = \alpha_{1j}\mathbf{c}_{1} + \dots + \alpha_{mj}\mathbf{c}_{m} \iff \mathbf{b}_{j} = (\mathbf{c}_{1} \cdots \mathbf{c}_{m}) \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}$$

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

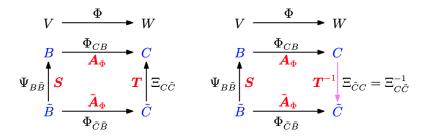
• $\hat{x} = \mathbf{A}_{\Phi}\hat{y}$, where \hat{x} is the vector w.r.t B and \hat{y} is the vector w.r.t. C

L2(7) April 7, 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 $B = (\mathbf{b}_1 \cdots \mathbf{b}_n), B' = (\mathbf{b}_1' \cdots \mathbf{b}_n') \quad C = (\mathbf{c}_1 \cdots \mathbf{c}_m), C' = (\mathbf{c}_1' \cdots \mathbf{c}_m').$
- (inter) transformation matrices ${m A}_\Phi$ from B to C and ${m A}'_\Phi$ from B' to C'
- (intra) transformation matrices S from B' to B and T from C' to C
- Theorem. ${m A}_{\Phi}' = T^{-1}{m A}_{\Phi}S$



L2(7) April 7, 2021 47 / 56

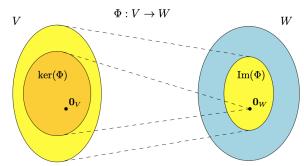
Image and Kernel



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

- 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 7, 2021 48 / 56

Image and Kernel: Properties



- $0_V \in \ker(\Phi)$ (because $\Phi(0_V) = 0_W$)
- Both $Im(\Phi)$ 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 $\Phi : \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$. Then, $Im(\Phi) = column$ space of **A** which is the span of column vectors of **A**.
- $\operatorname{rk}(\mathbf{A}) = \dim(\operatorname{Im}(\Phi))$
- $\ker(\Phi)$ is the solution set of the homogeneous system of linear equations $\mathbf{A}\mathbf{x}=0$

L2(7) April 7, 2021 49 / 56

Rank-Nullity Theorem

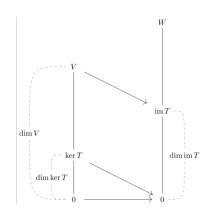


Theorem.

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

- If $\dim(\operatorname{Im}(\Phi)) < \dim(V)$, the kernel contains more than just 0.
- If $\dim(\operatorname{Im}(\Phi)) < \dim(V)$, $\mathbf{A}_{\Phi}\mathbf{x} = 0$ 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 y = Ax, where A is regular.
- Simplified version. For $\mathbf{A} \in \mathbb{R}^{m \times n}$,

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



²Nullity: the dimension of null space (kernel)

Roadmap



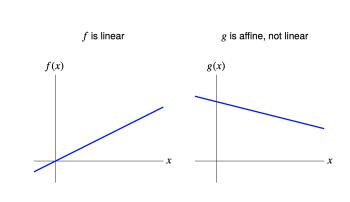
- (1) Systems of Linear Equations
- (2) Matrices
- (3) Solving Systems of Linear Equations
- (4) Vector Spaces
- (5) Linear Independence
- (6) Basis and Rank
- (7) Linear Mappings
- (8) Affine Spaces

L2(8) April 7, 2021 51 / 56

Linear vs. Affine Function



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



L2(8) April 7, 2021 52 / 56

Affine Subspace



- 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 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

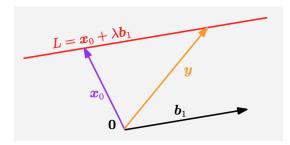
$$\mathbf{x} = \mathbf{x}_0 + \lambda_1 \mathbf{b}_1 + \cdots + \lambda_k \mathbf{b}_k, \quad \lambda_1, \dots, \lambda_k \in \mathbb{R}$$

L2(8) April 7, 2021 53 / 56

Example



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



• 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 7, 2021 54 / 56



Questions?

L2(8)

April 7, 2021 55 / 56

Review Questions



1)

L2(8) April 7, 2021 56 / 56