UPM
UNIVERSITI PUTRA MALAYSIA
BERILMU BERBAKTI

UNIVERSITI PUTRA MALAYSIA

AGRICULTURE • INNOVATION • LIFE

CHAPTER 1
ALGEBRAIC EXPRESSION

SKM 3002

MATHEMATICS FOR MULTIMEDIA COMPUTING

2022/2023 SEMESTER 2

Course Outline

- **1 - Chapter 1: Algebraic Expressions**
- 2 - Chapter 2: Linear Transformation
- 3 - Chapter 3: Analytic Geometry
- 4 - Chapter 4: Matrix Analysis
- 5 - Chapter 5: Vector Calculus
- 6 - Chapter 6: Probability
- 7 - Chapter 7: Distribution
- 8 -Chapter 8: Optimization Using Gradient Descent
- 9 - Chapter 9: Convex Optimization

Contents

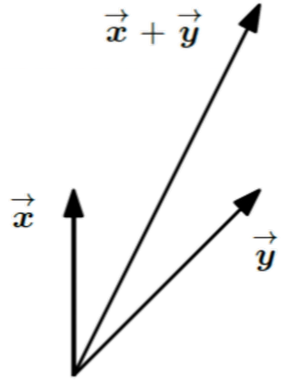
- Introduction
- System and Solving Systems of Linear Equations
- Linear inequality
- Matrices and vector spaces
- Linear mappings
- Basis, Rank and Affine Spaces
- Case Study: Manipulate digital audio signals

Introduction

- When formalizing intuitive concepts, a common approach is to construct a set of objects (symbols) and a set of rules to manipulate these objects.
- This is known as an algebra.
- Linear algebra is the study of vectors and certain rules to manipulate vectors.
- The vectors , which are usually denoted by a small arrow above the letter, e.g., $\vec{x}$ and $\vec{y}$
- In general, vectors are special objects that can be added together and multiplied by scalars to produce another object of the same kind.
- From an abstract mathematical viewpoint, any object that satisfies these two properties can be considered a vector.

Introduction cont…

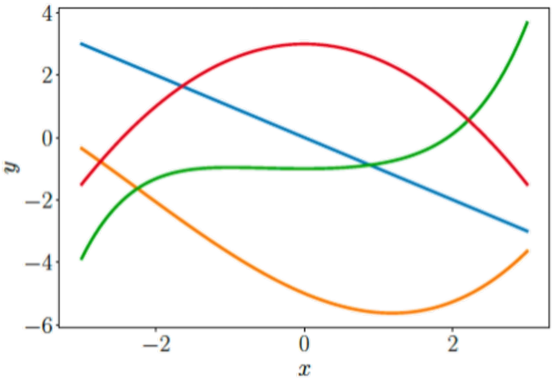
Geometric vectors. This example of a vector may be familiar from high school mathematics and physics. Geometric vectors – see Figure 2.1(a) – are directed segments, which can be drawn (at least in two dimensions). Two geometric vectors $\vec{x}$, $\vec{y}$ can be added, such that $\vec{x} + \vec{y} = \vec{z}$ is another geometric vector. Furthermore, multiplication by a scalar $\lambda \vec{x}$, $\lambda \in \mathbb{R}$, is also a geometric vector. In fact, it is the original vector scaled by λ. Therefore, geometric vectors are instances of the vector concepts introduced previously. Interpreting vectors as geometric vectors enables us to use our intuitions about direction and magnitude to reason about mathematical operations.



(a) Geometric vectors.

Introduction cont.

2. Polynomials are also vectors; see Figure 2.1(b): Two polynomials can be added together, which results in another polynomial; and they can be multiplied by a scalar $\lambda \in \mathbb{R}$, and the result is a polynomial as well. Therefore, polynomials are (rather unusual) instances of vectors. Note that polynomials are very different from geometric vectors. While geometric vectors are concrete "drawings", polynomials are abstract concepts. However, they are both vectors in the sense previously described.



(b) Polynomials.

Introduction cont.

3. Audio signals are vectors. Audio signals are represented as a series of numbers. We can add audio signals together, and their sum is a new audio signal. If we scale an audio signal, we also obtain an audio signal. Therefore, audio signals are a type of vector, too.
4. Elements of $\mathbb{R}^n$ (tuples of n real numbers) are vectors. $\mathbb{R}^n$ is more abstract than polynomials, and it is the concept we focus on in this book. For instance,

$$\boldsymbol{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^3 \tag{2.1}$$

is an example of a triplet of numbers. Adding two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ component-wise results in another vector: $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{c} \in \mathbb{R}^n$. Moreover, multiplying $\boldsymbol{a} \in \mathbb{R}^n$ by $\lambda \in \mathbb{R}$ results in a scaled vector $\lambda\boldsymbol{a} \in \mathbb{R}^n$. Considering vectors as elements of $\mathbb{R}^n$ has an additional benefit that it loosely corresponds to arrays of real numbers on a computer. Many programming languages support array operations, which allow for convenient implementation of algorithms that involve vector operations.

Introduction cont.

- Linear algebra focuses on the similarities between these vector concepts.
- We can add them together and multiply them by scalars. We will largely focus on vectors in $\mathbb{R}^n$ since most algorithms in linear algebra are formulated in $\mathbb{R}^n$

System and Solving Systems of Linear Equations

- Systems of linear equations play a central part of linear algebra.
- Many problems can be formulated as systems of linear equations, and linear algebra gives us the tools for solving them.

Example:

A company produces products $N_1, \ldots, N_n$ for which resources $R_1, \ldots, R_m$ are required. To produce a unit of product N_j, a_{ij} units of resource R_i are needed, where $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

The objective is to find an optimal production plan, i.e., a plan of how many units x_j of product N_j should be produced if a total of b_i units of resource R_i are available and (ideally) no resources are left over.

If we produce $x_1, \ldots, x_n$ units of the corresponding products, we need

a total of

$$a_{i1}x_1 + \cdots + a_{in}x_n \tag{2.2}$$

many units of resource R_i. An optimal production plan $(x_1, \ldots, x_n) \in \mathbb{R}^n$, therefore, has to satisfy the following system of equations:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}, \tag{2.3}$$

where $a_{ij} \in \mathbb{R}$ and $b_i \in \mathbb{R}$.

System and Solving Systems of Linear Equations cont.

Equation (2.3) is the general form of a system of linear equations, and x_1, x_2, …., x_n are the unknowns of this system.
Every n-tuple $(x_1, x_2, …., x_n) \epsilon R^n$ that satisfies (2.3) is a *solution* of the linear equation system.

System and Solving Systems of Linear Equations cont. Example:

The system of linear equations

$$\begin{array}{rcrcrcrl} x_1 & + & x_2 & + & x_3 & = & 3 & (1) \\ x_1 & - & x_2 & + & 2x_3 & = & 2 & (2) \\ 2x_1 & & & + & 3x_3 & = & 1 & (3) \end{array} \qquad (2.4)$$

has *no solution:* Adding the first two equations yields $2x_1 + 3x_3 = 5$, which contradicts the third equation (3).

Let us have a look at the system of linear equations

$$\begin{array}{rcrcrcrl} x_1 & + & x_2 & + & x_3 & = & 3 & (1) \\ x_1 & - & x_2 & + & 2x_3 & = & 2 & (2) \\ & & x_2 & + & x_3 & = & 2 & (3) \end{array} . \qquad (2.5)$$

From the first and third equation, it follows that $x_1 = 1$. From (1)+(2), we get $2x_1 + 3x_3 = 5$, i.e., $x_3 = 1$. From (3), we then get that $x_2 = 1$. Therefore, $(1, 1, 1)$ is the only possible and *unique solution* (verify that $(1, 1, 1)$ is a solution by plugging in).

As a third example, we consider

$$\begin{array}{rcrcrcrl} x_1 & + & x_2 & + & x_3 & = & 3 & (1) \\ x_1 & - & x_2 & + & 2x_3 & = & 2 & (2) \\ 2x_1 & & & + & 3x_3 & = & 5 & (3) \end{array} . \qquad (2.6)$$

Since (1)+(2)=(3), we can omit the third equation (redundancy). From (1) and (2), we get $2x_1 = 5 - 3x_3$ and $2x_2 = 1 + x_3$. We define $x_3 = a \in \mathbb{R}$ as a free variable, such that any triplet

$$\left(\frac{5}{2} - \frac{3}{2}a, \frac{1}{2} + \frac{1}{2}a, a\right), \quad a \in \mathbb{R} \qquad (2.7)$$

is a solution of the system of linear equations, i.e., we obtain a solution set that contains *infinitely many* solutions.

System and Solving Systems of Linear Equations cont.

- In a system of linear equations with two variables x_1 ,x_2, each linear equation defines a line on the $x_1 x_2$ -plane.
- Since a solution to a system of linear equations must satisfy all equations simultaneously, the solution set is the intersection of these lines.
- This intersection set can be a line (if the linear equations describe the same line), a point, or empty (when the lines are parallel).
- An illustration is given in Figure 2.3 for the system

$$\begin{aligned} 4x_1 + 4x_2 &= 5 \\ 2x_1 - 4x_2 &= 1 \end{aligned} \qquad (2.8)$$

where the solution space is the point (x_1 , x_2) = (1; 1/4). Similarly, for three variables, each linear equation determines a plane in three-dimensional space. When we intersect these planes, i.e., satisfy all linear equations at the same time, we can obtain a solution set that is a plane, a line, a point or empty (when the planes have no common intersection).

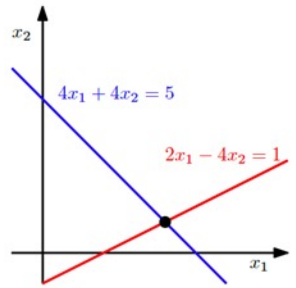
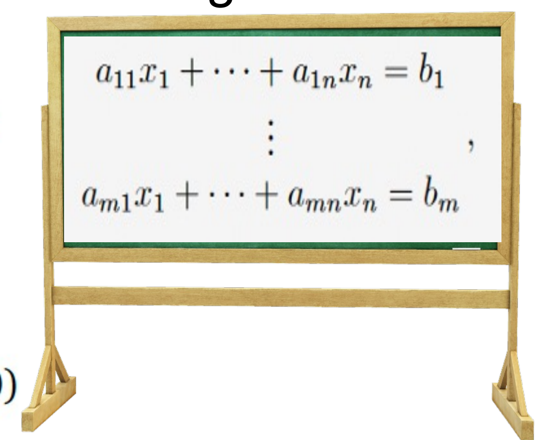


Figure 2.3 The solution space of a system of two linear equations with two variables can be geometrically interpreted as the intersection of two lines. Every linear equation represents a line.

System and Solving Systems of Linear Equations cont.

- For a systematic approach to solving systems of linear equations, we will introduce a useful compact notation. We collect the coefficients a_{ij} into vectors and collect the vectors into matrices. In other words, we write the system from (2.3) in the following form:

$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{2.9}$$



$$\iff \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} . \tag{2.10}$$

- In the following, we will closely examine these matrices and define computation rules.
- We will return to solving linear equations in Section 2.3.

System and Solving Systems of Linear Equations cont.

- Matrices play a central role in linear algebra. They can be used to compactly represent systems of linear equations, but they also represent linear functions (linear

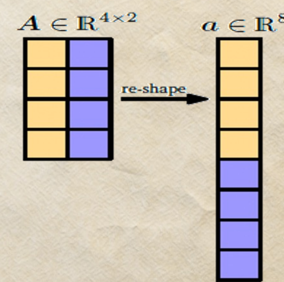
Definition 2.1 (Matrix). With $m, n \in \mathbb{N}$ a real-valued (m, n) *matrix* $\boldsymbol{A}$ is an $m \cdot n$-tuple of elements a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, which is ordered according to a rectangular scheme consisting of m rows and n columns:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad a_{ij} \in \mathbb{R}. \tag{2.11}$$

By convention $(1, n)$-matrices are called *rows* and $(m, 1)$-matrices are called *columns*. These special matrices are also called *row/column vectors*.

$\mathbb{R}^{m \times n}$ is the set of all real-valued (m, n)-matrices. $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ can be equivalently represented as $\boldsymbol{a} \in \mathbb{R}^{mn}$ by stacking all n columns of the matrix into a long vector; see Figure 2.4.

Figure 2.4 By stacking its columns, a matrix $\boldsymbol{A}$ can be represented as a long vector $\boldsymbol{a}$.



Note the size of the matrices.

Matrices

The sum of two matrices $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{m\times n}$ is defined as the element-wise sum, i.e.,

$$\boldsymbol{A}+\boldsymbol{B} := \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}. \tag{2.12}$$

For matrices $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times k}$, the elements c_{ij} of the product $\boldsymbol{C} = \boldsymbol{AB} \in \mathbb{R}^{m\times k}$ are computed as

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, k. \tag{2.13}$$

- This means, to compute element cij we multiply the elements of the ith row of A with the jth column of B and sum them up.

Matrices cont.

- Matrices can only be multiplied if their "neighboring" dimensions match.
- For instance, an $n \times k$ -matrix A can be multiplied with a $k \times m$ matrix B, but only from the left side:

$$\underbrace{A}_{n\times k}\underbrace{B}_{k\times m} = \underbrace{C}_{n\times m} \tag{2.14}$$

- The product BA is not defined if m $\neq$ n since the neighboring dimensions do not match.

Matrices cont.

- Matrix multiplication is not defined as an element-wise operation on matrix elements, i.e., $c_{ij} \neq a_{ij}b_{ij}$ (even if the size of A, B was chosen appropriately).
- This kind of element-wise multiplication often appears in programming languages when we multiply (multi-dimensional) arrays with each other and is called a ***Hadamard product.***

Example 2.3

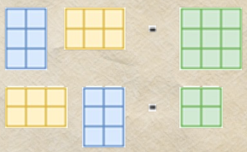
For $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 3}$, $\boldsymbol{B} = \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 2}$, we obtain

$$\boldsymbol{AB} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 2 & 5 \end{bmatrix} \in \mathbb{R}^{2\times 2}, \tag{2.15}$$

$$\boldsymbol{BA} = \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 4 & 2 \\ -2 & 0 & 2 \\ 3 & 2 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 3}. \tag{2.16}$$

- multiplication is not commutative, i.e., $AB \neq BA$; see also Figure 2.5 for an illustration.

Figure 2.5 Even if both matrix multiplications AB and BA are defined, the dimensions of the results can be different.



identity matrix

Matrices cont.

Definition 2.2 (Identity Matrix). In $\mathbb{R}^{n\times n}$, we define the *identity matrix*

$$I_n := \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n\times n} \tag{2.17}$$

as the $n \times n$-matrix containing 1 on the diagonal and 0 everywhere else.

Now that we defined matrix multiplication, matrix addition and the identity matrix, let us have a look at some properties of matrices:

- *Associativity:*

$$\forall \boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{B} \in \mathbb{R}^{n\times p}, \boldsymbol{C} \in \mathbb{R}^{p\times q} : (\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC}) \tag{2.18}$$

- *Distributivity:*

$$\forall \boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m\times n}, \boldsymbol{C}, \boldsymbol{D} \in \mathbb{R}^{n\times p} : (\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC} \tag{2.19a}$$

$$\boldsymbol{A}(\boldsymbol{C}+\boldsymbol{D}) = \boldsymbol{AC} + \boldsymbol{AD} \tag{2.19b}$$

- Multiplication with the identity matrix:

$$\forall \boldsymbol{A} \in \mathbb{R}^{m\times n} : \boldsymbol{I}_m \boldsymbol{A} = \boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{A} \tag{2.20}$$

Note that $\boldsymbol{I}_m \neq \boldsymbol{I}_n$ for $m \neq n$.

Matrices cont.

Definition 2.3 (Inverse). Consider a square matrix $A \in \mathbb{R}^{n \times n}$. Let matrix $B \in \mathbb{R}^{n \times n}$ have the property that $AB = I_n = BA$. B is called the *inverse* of A and denoted by A^{-1}.

- Unfortunately, not every matrix A possesses an inverse A^{-1}
- If this inverse does exist, A is called regular / invertible / non-singular, otherwise singular / noninvertible.
- When the matrix inverse exists, it is unique.

Matrices cont.

Remark (Existence of the Inverse of a 2×2-matrix). Consider a matrix

$$\boldsymbol{A} := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}. \tag{2.21}$$

If we multiply $\boldsymbol{A}$ with

$$\boldsymbol{A}' := \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \tag{2.22}$$

we obtain

$$\boldsymbol{A}\boldsymbol{A}' = \begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} = (a_{11}a_{22} - a_{12}a_{21})\boldsymbol{I}. \tag{2.23}$$

Therefore,

$$\boldsymbol{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \tag{2.24}$$

$if\ and\ only\ if\ a_{11}a_{22} - a_{12}a_{21} \neq 0.$

The $a_{11}a_{22} - a_{12}a_{21} \neq 0$ is the determinant of a $2\times 2-$matrix.

The determinant can be used to check whether a matrix is invertible.

Example 2.4 (Inverse Matrix)

The matrices

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 4 & 5 \\ 6 & 7 & 7 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} -7 & -7 & 6 \\ 2 & 1 & -1 \\ 4 & 5 & -4 \end{bmatrix} \tag{2.25}$$

are inverse to each other since $\boldsymbol{AB} = \boldsymbol{I} = \boldsymbol{BA}$.

Matrices cont.

Definition 2.4 (Transpose). For $A \in \mathbb{R}^{m\times n}$ the matrix $B \in \mathbb{R}^{n\times m}$ with $b_{ij} = a_{ji}$ is called the *transpose* of A. We write $B = A^\top$.

In general, $A^\top$ can be obtained by writing the columns of A as the rows of $A^\top$. The following are important properties of inverses and transposes:

$$AA^{-1} = I = A^{-1}A \tag{2.26}$$

$$(AB)^{-1} = B^{-1}A^{-1} \tag{2.27}$$

$$(A+B)^{-1} \neq A^{-1} + B^{-1} \tag{2.28}$$

$$(A^\top)^\top = A \tag{2.29}$$

$$(A+B)^\top = A^\top + B^\top \tag{2.30}$$

$$(AB)^\top = B^\top A^\top \tag{2.31}$$

Definition 2.5 (Symmetric Matrix). A matrix $A \in \mathbb{R}^{n\times n}$ is *symmetric* if $A = A^\top$.

Note that only (n, n)-matrices can be symmetric. Generally, we call (n, n)-matrices also *square matrices* because they possess the same number of rows and columns. Moreover, if A is invertible, then so is $A^\top$, and $(A^{-1})^\top = (A^\top)^{-1} =: A^{-\top}$.

Remark (Sum and Product of Symmetric Matrices). The sum of symmetric matrices $A, B \in \mathbb{R}^{n\times n}$ is always symmetric. However, although their product is always defined, it is generally not symmetric:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}. \tag{2.32}$$

Matrices cont.

Let us look at what happens to matrices when they are multiplied by a scalar $\lambda \in \mathbb{R}$. Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\lambda \in \mathbb{R}$. Then $\lambda\boldsymbol{A} = \boldsymbol{K}$, $K_{ij} = \lambda\, a_{ij}$. Practically, λ scales each element of $\boldsymbol{A}$. For $\lambda, \psi \in \mathbb{R}$, the following holds:

- *Associativity:*
  $(\lambda\psi)C = \lambda(\psi C), \quad C \in \mathbb{R}^{m\times n}$
- $\lambda(BC) = (\lambda B)C = B(\lambda C) = (BC)\lambda, \quad B \in \mathbb{R}^{m\times n}, C \in \mathbb{R}^{n\times k}$.
  Note that this allows us to move scalar values around.
- $(\lambda C)^\top = C^\top \lambda^\top = C^\top \lambda = \lambda C^\top$ since $\lambda = \lambda^\top$ for all $\lambda \in \mathbb{R}$.
- *Distributivity:*
  $(\lambda + \psi)C = \lambda C + \psi C, \quad C \in \mathbb{R}^{m\times n}$
  $\lambda(B + C) = \lambda B + \lambda C, \quad B, C \in \mathbb{R}^{m\times n}$

Example 2.5 (Distributivity)
If we define

$$C := \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \tag{2.33}$$

then for any $\lambda, \psi \in \mathbb{R}$ we obtain

$$(\lambda+\psi)C = \begin{bmatrix} (\lambda+\psi)1 & (\lambda+\psi)2 \\ (\lambda+\psi)3 & (\lambda+\psi)4 \end{bmatrix} = \begin{bmatrix} \lambda+\psi & 2\lambda+2\psi \\ 3\lambda+3\psi & 4\lambda+4\psi \end{bmatrix} \tag{2.34a}$$

$$= \begin{bmatrix} \lambda & 2\lambda \\ 3\lambda & 4\lambda \end{bmatrix} + \begin{bmatrix} \psi & 2\psi \\ 3\psi & 4\psi \end{bmatrix} = \lambda C + \psi C. \tag{2.34b}$$

Matrices cont…

If we consider the system of linear equations

$$\begin{aligned} 2x_1 + 3x_2 + 5x_3 &= 1 \\ 4x_1 - 2x_2 - 7x_3 &= 8 \\ 9x_1 + 5x_2 - 3x_3 &= 2 \end{aligned} \tag{2.35}$$

and use the rules for matrix multiplication, we can write this equation system in a more compact form as

$$\begin{bmatrix} 2 & 3 & 5 \\ 4 & -2 & -7 \\ 9 & 5 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \\ 2 \end{bmatrix} . \tag{2.36}$$

Note that x_1 scales the first column, x_2 the second one, and x_3 the third one.

- Generally, a system of linear equations can be compactly represented in their matrix form as Ax =b
- The product Ax is a (linear) combination of the columns of A.

System and Solving Systems of Linear Equations – Matrices

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}, \tag{2.3}$$

where $a_{ij} \in \mathbb{R}$ and $b_i \in \mathbb{R}$ are known constants and x_j are unknowns, $i = 1, \ldots, m$, $j = 1, \ldots, n$. Thus far, we saw that matrices can be used as a compact way of formulating systems of linear equations so that we can write $Ax = b$, see (2.10). Moreover, we defined basic matrix operations, such as addition and multiplication of matrices. In the following, we will focus on solving systems of linear equations and provide an algorithm for finding the inverse of a matrix.

Recalled: The general form of an equation system

System and Solving Systems of Linear Equations – Matrices

Consider the system of equations:

$$\begin{bmatrix} \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & \begin{matrix} 8 & -4 \\ 2 & 12 \end{matrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 42 \\ 8 \end{bmatrix}. \qquad (2.38)$$

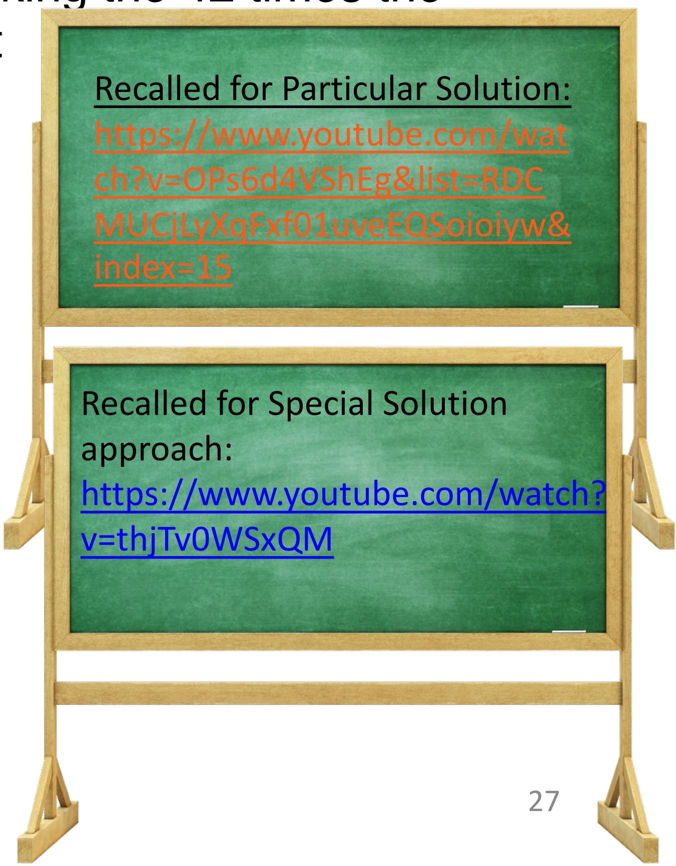
- The system has two equations and four unknowns. Therefore, in general we would expect infinitely many solutions.
- This system of equations is in a particularly easy form, where the first two columns consist of a 1 and a 0

System and Solving Systems of Linear Equations – Matrices

- Remember that we want to find scalars x_1, x_2, x_3, x_4 such that $\sum_{i=1}^{4} x_i c_i = b$, where we define c_i to be the ith column of the matrix and b the right-hand-side of (2.38).
- A solution to (2.38) can be found immediately by taking the 42 times the first column and 8 times the second column so that

$$\boldsymbol{b} = \begin{bmatrix} 42 \\ 8 \end{bmatrix} = 42 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 8 \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \qquad (2.39)$$

- Therefore, a solution is $[42,8,0,0]^T$.
- This solution is called a **particular solution** or **special solution**
- However, this is not the only solution of this system of linear equations.
- Therefore, To capture all the other solutions, we need to be creative in generating 0 in a non-trivial way using the columns of the matrix -> Adding 0 to our special solution does not change the special solution.



Matrices

System and Solving Systems of Linear Equations

$$\begin{bmatrix}1 & 0 & 8 & -4\\ 0 & 1 & 2 & 12\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix} = \begin{bmatrix}42\\ 8\end{bmatrix}. \quad (2.38)$$

Let say x3 =1, x4 =0

x1 = -8

X2 = -2

X3 = 1

X4 =0

- To do so, we express the third column using two columns

$$\begin{bmatrix}8\\ 2\end{bmatrix} = 8\begin{bmatrix}1\\ 0\end{bmatrix} + 2\begin{bmatrix}0\\ 1\end{bmatrix} \quad (2.40)$$

so that $\mathbf{0} = 8\boldsymbol{c}_1 + 2\boldsymbol{c}_2 - 1\boldsymbol{c}_3 + 0\boldsymbol{c}_4$ and $(x_1, x_2, x_3, x_4) = (8, 2, -1, 0)$. In fact, any scaling of this solution by $\lambda_1 \in \mathbb{R}$ produces the $\mathbf{0}$ vector, i.e.,

$$\begin{bmatrix}1 & 0 & 8 & -4\\ 0 & 1 & 2 & 12\end{bmatrix}\left(\lambda_1\begin{bmatrix}8\\ 2\\ -1\\ 0\end{bmatrix}\right) = \lambda_1(8\boldsymbol{c}_1 + 2\boldsymbol{c}_2 - \boldsymbol{c}_3) = \mathbf{0}\,. \quad (2.41)$$

Following the same line of reasoning, we express the fourth column of the matrix in (2.38) using the first two columns and generate another set of non-trivial versions of $\mathbf{0}$ as

$$\begin{bmatrix}1 & 0 & 8 & -4\\ 0 & 1 & 2 & 12\end{bmatrix}\left(\lambda_2\begin{bmatrix}-4\\ 12\\ 0\\ -1\end{bmatrix}\right) = \lambda_2(-4\boldsymbol{c}_1 + 12\boldsymbol{c}_2 - \boldsymbol{c}_4) = \mathbf{0} \quad (2.42)$$

for any $\lambda_2 \in \mathbb{R}$. Putting everything together, we obtain all solutions of the equation system in (2.38), which is called the *general solution*, as the set

$$\left\{\boldsymbol{x} \in \mathbb{R}^4 : \boldsymbol{x} = \begin{bmatrix}42\\ 8\\ 0\\ 0\end{bmatrix} + \lambda_1\begin{bmatrix}8\\ 2\\ -1\\ 0\end{bmatrix} + \lambda_2\begin{bmatrix}-4\\ 12\\ 0\\ -1\end{bmatrix}, \lambda_1, \lambda_2 \in \mathbb{R}\right\}. \quad (2.43)$$

Recalled for Particular Solution:
https://www.youtube.com/watch?v=OPs6d4VShEg&list=RDCMUCjLyXqFxf01uveEQSoioiyw&index=15

Recalled for Special Solution approach: https://www.youtube.com/watch?v=thjTv0WSxQM

Solving Systems of Linear Equations - remarks

Remark. The general approach we followed consisted of the following three steps:

1. Find a particular solution to $\boldsymbol{Ax} = \boldsymbol{b}$.
2. Find all solutions to $\boldsymbol{Ax} = \boldsymbol{0}$.
3. Combine the solutions from steps 1. and 2. to the general solution.

- The system of linear equations in the preceding example was easy to solve because the matrix in (2.38) has this particularly convenient form, which allowed us to find the particular and the general solution by inspection.
- However, general equation systems are not of this simple form.

Solving Systems of Linear Equations - remarks

- Fortunately, there exists a constructive algorithmic way of transforming any system of linear equations into this particularly simple form: **Gaussian elimination**.
- Key to Gaussian elimination are elementary transformations of systems of linear equations, which transform the equation system into a simple form.
- Then, we can apply the three steps to the simple form that we just discussed in the context of the example in (2.38).

Solving Systems of Linear Equations – Elementary Transformation

- Key to solving a system of linear equations are elementary transformations transformations that keep the solution set the same, but that transform the equation system into a simpler form:
  - – Exchange of two equations (rows in the matrix representing the system of equations)
  - – Multiplication of an equation (row) with a constant $\lambda \in \mathbb{R}\backslash\{0\}$
  - – Addition of two equations (rows)

Elementary Transformation

System and Solving Systems of Linear Equations

Example 2.6
For $a \in \mathbb{R}$, we seek all solutions of the following system of equations:

$$\begin{array}{rcrcrcrcrcr} -2x_1 & + & 4x_2 & - & 2x_3 & - & x_4 & + & 4x_5 & = & -3 \\ 4x_1 & - & 8x_2 & + & 3x_3 & - & 3x_4 & + & x_5 & = & 2 \\ x_1 & - & 2x_2 & + & x_3 & - & x_4 & + & x_5 & = & 0 \\ x_1 & - & 2x_2 & & & - & 3x_4 & + & 4x_5 & = & a \end{array} . \tag{2.44}$$

We start by converting this system of equations into the compact matrix notation $\boldsymbol{Ax} = \boldsymbol{b}$. We no longer mention the variables $\boldsymbol{x}$ explicitly and build the *augmented matrix* (in the form $[\boldsymbol{A} \,|\, \boldsymbol{b}]$)

augmented matrix

$$\begin{array}{cl} \left[\begin{array}{rrrrr|r} -2 & 4 & -2 & -1 & 4 & -3 \\ 4 & -8 & 3 & -3 & 1 & 2 \\ 1 & -2 & 1 & -1 & 1 & 0 \\ 1 & -2 & 0 & -3 & 4 & a \end{array}\right] & \begin{array}{l} \text{Swap with } R_3 \\ \\ \text{Swap with } R_1 \\ \\ \end{array} \end{array}$$

where we used the vertical line to separate the left-hand side from the right-hand side in (2.44). We use $\rightsquigarrow$ to indicate a transformation of the augmented matrix using elementary transformations.

The augmented matrix $[\boldsymbol{A} \,|\, \boldsymbol{b}]$ compactly represents the system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$.

Swapping Rows 1 and 3 leads to

$$\begin{array}{cl} \left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 4 & -8 & 3 & -3 & 1 & 2 \\ -2 & 4 & -2 & -1 & 4 & -3 \\ 1 & -2 & 0 & -3 & 4 & a \end{array}\right] & \begin{array}{l} \\ -4R_1 \\ +2R_1 \\ -R_1 \end{array} \end{array}$$

When we now apply the indicated transformations (e.g., subtract Row 1 four times from Row 2), we obtain

$$\begin{array}{ccl} & \left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & -1 & -2 & 3 & a \end{array}\right] & \begin{array}{l} \\ \\ \\ -R_2 - R_3 \end{array} \\ \rightsquigarrow & \left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right] & \begin{array}{l} \\ \cdot(-1) \\ \cdot(-\frac{1}{3}) \\ \\ \end{array} \\ \rightsquigarrow & \left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right] & \end{array}$$

Solving Systems of Linear Equations – Elementary Transformation

row-echelon form

This (augmented) matrix is in a convenient form, the *row-echelon form* (REF). Reverting this compact notation back into the explicit notation with the variables we seek, we obtain

$$\begin{array}{rcrcrcrcrcr} x_1 & - & 2x_2 & + & x_3 & - & x_4 & + & x_5 & = & 0 \\ & & & & x_3 & - & x_4 & + & 3x_5 & = & -2 \\ & & & & & & x_4 & - & 2x_5 & = & 1 \\ & & & & & & & & 0 & = & a+1 \end{array} \quad . \tag{2.45}$$

particular solution

Only for $a = -1$ this system can be solved. A *particular solution* is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} . \tag{2.46}$$

general solution

The *general solution*, which captures the set of all possible solutions, is

$$\left\{ \boldsymbol{x} \in \mathbb{R}^5 : \boldsymbol{x} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix} , \quad \lambda_1, \lambda_2 \in \mathbb{R} \right\} . \tag{2.47}$$

Solving Systems of Linear Equations – Elementary Transformation

Definition 2.6 (Row-Echelon Form). A matrix is in *row-echelon form* if

- All rows that contain only zeros are at the bottom of the matrix; correspondingly, all rows that contain at least one nonzero element are on top of rows that contain only zeros.
- Looking at nonzero rows only, the first nonzero number from the left (also called the *pivot* or the *leading coefficient*) is always strictly to the right of the pivot of the row above it.

Solving Systems of Linear Equations – Elementary Transformation

Remark (Basic and Free Variables). The variables corresponding to the pivots in the row-echelon form are called *basic variables* and the other variables are *free variables*. For example, in (2.45), x_1, x_3, x_4 are basic variables, whereas x_2, x_5 are free variables. ◊

$$
\begin{array}{ccccccccccc}
x_1 & - & 2x_2 & + & x_3 & - & x_4 & + & x_5 & = & 0 \\
 & & & & x_3 & - & x_4 & + & 3x_5 & = & -2 \\
 & & & & & & x_4 & - & 2x_5 & = & 1 \\
 & & & & & & & & 0 & = & a+1
\end{array} . \qquad (2.45)
$$

Solving Systems of Linear Equations – Elementary Transformation

Remark (Obtaining a Particular Solution). The row-echelon form makes our lives easier when we need to determine a particular solution. To do this, we express the right-hand side of the equation system using the pivot columns, such that $\boldsymbol{b} = \sum_{i=1}^{P} \lambda_i \boldsymbol{p}_i$, where $\boldsymbol{p}_i,\ i = 1, \ldots, P$, are the pivot columns. The λ_i are determined easiest if we start with the rightmost pivot column and work our way to the left.

In the previous example, we would try to find $\lambda_1, \lambda_2, \lambda_3$ so that

$$\lambda_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} . \tag{2.48}$$

From here, we find relatively directly that $\lambda_3 = 1, \lambda_2 = -1, \lambda_1 = 2$. When we put everything together, we must not forget the non-pivot columns for which we set the coefficients implicitly to 0. Therefore, we get the particular solution $\boldsymbol{x} = [2, 0, -1, 1, 0]^\top$. ◇

Solving Systems of Linear Equations – Elementary Transformation

Remark (Gaussian Elimination). *Gaussian elimination* is an algorithm that performs elementary transformations to bring a system of linear equations into reduced row-echelon form. ◇

Example 2.7 (Reduced Row Echelon Form)
Verify that the following matrix is in reduced row-echelon form (the pivots are in **bold**):

$$\boldsymbol{A} = \begin{bmatrix} \mathbf{1} & 3 & 0 & 0 & 3 \\ 0 & 0 & \mathbf{1} & 0 & 9 \\ 0 & 0 & 0 & \mathbf{1} & -4 \end{bmatrix}. \tag{2.49}$$

The key idea for finding the solutions of $\boldsymbol{Ax} = \mathbf{0}$ is to look at the *non-pivot columns*, which we will need to express as a (linear) combination of the pivot columns. The reduced row echelon form makes this relatively straightforward, and we express the non-pivot columns in terms of sums and multiples of the pivot columns that are on their left: The second column is 3 times the first column (we can ignore the pivot columns on the right of the second column). Therefore, to obtain $\mathbf{0}$, we need to subtract

the second column from three times the first column. Now, we look at the fifth column, which is our second non-pivot column. The fifth column can be expressed as 3 times the first pivot column, 9 times the second pivot column, and -4 times the third pivot column. We need to keep track of the indices of the pivot columns and translate this into 3 times the first column, 0 times the second column (which is a non-pivot column), 9 times the third column (which is our second pivot column), and -4 times the fourth column (which is the third pivot column). Then we need to subtract the fifth column to obtain $\mathbf{0}$. In the end, we are still solving a homogeneous equation system.

To summarize, all solutions of $\boldsymbol{Ax} = \mathbf{0}, \boldsymbol{x} \in \mathbb{R}^5$ are given by

$$\left\{ \boldsymbol{x} \in \mathbb{R}^5 : \boldsymbol{x} = \lambda_1 \begin{bmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{bmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \right\}. \tag{2.50}$$

Solving Systems of Linear Equations – The Minus-1 Trick

In the following, we introduce a practical trick for reading out the solutions $\boldsymbol{x}$ of a homogeneous system of linear equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, where $\boldsymbol{A} \in \mathbb{R}^{k\times n}, \boldsymbol{x} \in \mathbb{R}^n$.

To start, we assume that $\boldsymbol{A}$ is in reduced row-echelon form without any rows that just contain zeros, i.e.,

$$\boldsymbol{A} = \begin{bmatrix} 0 & \cdots & 0 & \mathbf{1} & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & * \\ \vdots & & \vdots & 0 & 0 & \cdots & 0 & \mathbf{1} & * & \cdots & * & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & 0 & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \mathbf{1} & * & \cdots & * \end{bmatrix}, \tag{2.51}$$

where $*$ can be an arbitrary real number, with the constraints that the first nonzero entry per row must be 1 and all other entries in the corresponding column must be 0. The columns $j_1, \ldots, j_k$ with the pivots (marked in **bold**) are the standard unit vectors $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k \in \mathbb{R}^k$. We extend this matrix to an $n \times n$-matrix $\tilde{\boldsymbol{A}}$ by adding $n - k$ rows of the form

$$\begin{bmatrix} 0 & \cdots & 0 & -1 & 0 & \cdots & 0 \end{bmatrix} \tag{2.52}$$

so that the diagonal of the augmented matrix $\tilde{\boldsymbol{A}}$ contains either 1 or -1. Then, the columns of $\tilde{\boldsymbol{A}}$ that contain the -1 as pivots are solutions of the homogeneous equation system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$. To be more precise, these columns form a basis (Section 2.6.1) of the solution space of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, which we will later call the *kernel* or *null space* (see Section 2.7.3).

Solving Systems of Linear Equations – The Minus-1 Trick

Example 2.8 (Minus-1 Trick)
Let us revisit the matrix in (2.49), which is already in REF:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix} . \tag{2.53}$$

We now augment this matrix to a 5×5 matrix by adding rows of the form (2.52) at the places where the pivots on the diagonal are missing and obtain

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} . \tag{2.54}$$

From this form, we can immediately read out the solutions of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ by taking the columns of $\tilde{\boldsymbol{A}}$, which contain -1 on the diagonal:

$$\left\{ \boldsymbol{x} \in \mathbb{R}^5 : \boldsymbol{x} = \lambda_1 \begin{bmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{bmatrix} , \quad \lambda_1, \lambda_2 \in \mathbb{R} \right\} , \tag{2.55}$$

which is identical to the solution in (2.50) that we obtained by "insight".

Class activity

- Use this link to know what simple solution that required system linear equation
- [New chat (openai.com)]
- We will discuss in the class

UNIVERSITI PUTRA MALAYSIA

BERILMU BERBAKTI

TERIMA KASIH/*THANK YOU*

www.upm.edu.my

BERILMU BERBAKTI

WITH KNOWLEDGE WE SERVE