8. Linear maps and linear operators

Instructor: Prof. Christopher L. Nehaniv, Ph.D.

slides by

Douglas Wilhelm Harder, LEL, M.Math.
dwharder@uwaterloo.ca
dwharder@gmail.com
Introduction

• In this topic, we will
  – Define maps and operators on vector spaces
  – Give some background
  – Define the image and range of maps
  – Give some examples
In engineering, any signal, be it continuous-time or discrete-time, may be represented as a vector:
- Voltages
- Sound
- Temperatures
- Concentrations of chemicals
- Global positioning system coordinates

Engineers must extract information from such signals.
Background

• We may be calculating, for example:
  – Average values of a signal or vector
  – Reducing noise within a signal
  – Increasing the resolution
    • Displaying an analog TV signal on a 4K screen

• In all cases, the output is another vector
  – In the first, the output is $\mathbb{R}^1$
  – In the second, it is the same vector space
  – In the third, it is a higher-dimensional vector space
Maps between vector spaces

• Given two vector spaces $\mathcal{U}$ and $\mathcal{V}$,
  a map $A$ is a function that takes a vector in one vector space and maps it to a vector in another vector space.

\[
\begin{align*}
\mathcal{U} \xrightarrow{A} \mathcal{V}
\end{align*}
\]

• We will write this as $A: \mathcal{U} \rightarrow \mathcal{V}$
  – Here, $A$ must map every vector in $\mathcal{U}$ to some vector in $\mathcal{V}$
  – We will say $\mathbf{u} \in \mathcal{U}$ is mapped to the vector $A\mathbf{u} \in \mathcal{V}$
  – We may sometimes write $A(\mathbf{u})$ to avoid confusion,
    • For example $A(\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2)$
Definitions

• Given a map \( A: U \rightarrow V \),
  we will refer to \( U \) as the domain
  we will refer to \( V \) as the codomain

• If \( A: U \rightarrow U \), that is, the domain and codomain are equal,
  we will refer to \( A \) as an operator

• If \( u \in U \), then \( Au \in V \) is the image of \( u \) under the map \( A \)

• If \( v \in V \), and \( u \in U \) such that \( v = Au \),
  we will say \( u \) is a pre-image of \( v \) under the map \( A \)

• If \( S \subseteq U \), then \( AS \) is the collection of all \( Au \) where \( u \in S \)
  – \( AS \) is the image of \( S \) under the map \( A \)

• Let \( A \mathcal{U} \) be the collection of all images \( Au \) for every \( u \in U \)
  – This is said to be the range of \( A \)
Example

- Here are examples of two operators, each mapping $\mathbb{R}^3$ to $\mathbb{R}^3$
  - The first maps to cylindrical coordinates
  - The second to spherical coordinates

\[
\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} \sqrt{x^2 + y^2} \\ \tan^{-1} \left( \frac{y}{x} \right) \\ z \end{pmatrix}
\]

\[
\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} \sqrt{x^2 + y^2 + z^2} \\ \tan^{-1} \left( \frac{y}{x} \right) \\ \cos^{-1} \left( \frac{z}{\sqrt{x^2 + y^2 + z^2}} \right) \end{pmatrix}
\]
Example

- Here are three other operators
  - The first maps the vector to the average value
  - The second subtracts off the average of the entries
  - The third averages each point with its neighbors

\[
\begin{align*}
\begin{pmatrix} x \\ y \\ z \end{pmatrix} & \mapsto \left( \frac{x + y + z}{3} \right) \\
\begin{pmatrix} x \\ y \\ z \end{pmatrix} & \mapsto \left( \begin{array}{c} x - \frac{x + y + z}{3} \\ y - \frac{x + y + z}{3} \\ z - \frac{x + y + z}{3} \end{array} \right) \\
\begin{pmatrix} x \\ y \\ z \end{pmatrix} & \mapsto \left( \begin{array}{c} \frac{x + y}{2} \\ \frac{x + y + z}{3} \\ \frac{y + z}{2} \end{array} \right)
\end{align*}
\]
Summary

• Following this topic, you now
  – Understand the of a map or operator
    • A map takes vectors in the domain
      and maps them to vectors in the codomain
  – Know the terms image and range
    • We can take the image of a vector, a subset of the domain,
      or the entire domain
  – Have seen some examples
8.1 The superposition principle

Instructor: Prof. Christopher L. Nehaniv, Ph.D.

slides by Douglas Wilhelm Harder, LEL, M.Math.
dwharder@uwaterloo.ca
dwharder@gmail.com
Introduction

- In this topic, we will
  - Describe the superposition principle
  - Give some examples
  - Examine the need for functions that preserve vector space operations
The superposition principle is an observation from physics that there are many systems where the net response at a given place and time caused by two or more stimuli is the sum of the response that would have been caused by each stimulus individually.
The superposition principle

• This image shows the superposition of waves:

  – One specific single wave, a *soliton*, does not obey the superposition principle when two such waves meet

Photograph by Flickr user Spiralz
Examples

- Here we see middle C combined with noise

- One goal of sound engineering is to remove noise
Examples

- The goal of noise-cancelling headphones is to:
  - Estimate the noise vector $\mathbf{u}$
  - Introduce $-\mathbf{u}$ to the sound generated by the headphones
- This cancels out the outside noise
Examples

- The gravitational and electromagnetic forces also obey the superposition principle

\[ F_{\text{gravitational}} = m_0 a_k = \sum_{k=1}^{n} G \frac{m_0 m_k}{\|x_0 - x_k\|_2^3} (x_0 - x_k) \]

\[ = \left( Gm_0 \sum_{k=1}^{n} \frac{m_k}{\|x_0 - x_k\|_2^3} \right) x_0 - Gm_0 \sum_{k=1}^{n} \frac{m_k x_k}{\|x_0 - x_k\|_2^3} \]

\[ F_{\text{electric}} = m_0 a_k = \sum_{k=1}^{n} k_e \frac{q_0 q_k}{\|x_0 - x_k\|_2^3} (x_0 - x_k) \]

\[ = \left( k_e q_0 \sum_{k=1}^{n} \frac{q_k}{\|x_0 - x_k\|_2^3} \right) x_0 - k_e q_0 \sum_{k=1}^{n} \frac{q_k x_k}{\|x_0 - x_k\|_2^3} \]
Examples

• Also, alternating current (AC) circuits with resistors, inductors and capacitors display a vector-like behavior:
  – If you double a current, the voltage response double
    • Scalar multiplication
  – If you add two currents, the voltage response is the sum of the individual voltage response
    • Vector addition
Consequence

• Engineers prefer to work with mappings that preserve vector space operations
  – The response to a sum of inputs, should be the sum of the responses to the individual inputs
  – The response to an attenuated or amplified input, should be the response similarly attenuated or amplified

• For example, if you develop a sound system that
  – Nearly eliminates white noise
  – Leaves the volumes of human voices essentially unchanged then when you mix white noise with human voices, the noise should still be eliminated while the human voices remain
Looking forward

• Consequently, we will look at linear maps and linear operators:
  \[ f(\alpha u + \beta v) = \alpha f(u) + \beta f(v) \]

• You have already seen one linear operator
  the derivative:
  \[ \frac{d}{dx}(\alpha f(x) + \beta g(x)) = \alpha \frac{d}{dx} f(x) + \beta \frac{d}{dx} g(x) \]
Summary

• Following this topic, you now
  – Know that the superposition principle comes from natural phenomena
  – Understand these phenomena can be described through vector space operations
  – Can see the potential need for functions that preserve vector space operations
    • The response to a linear combination of inputs, is that same linear combination of the individual responses
8.2 Definition of linear maps and linear operators

Instructor: Prof. Chrystopher L. Nehaniv, Ph.D.

slides by Douglas Wilhelm Harder, LEL, M.Math.
dwharder@uwaterloo.ca
dwharder@gmail.com
Introduction

• In this topic, we will
  – Define a linear map
  – Understand what it implies
  – Describe a test for linearity
  – Look at some examples
  – See that all finite-dimensional linear maps can be represented by matrices
  – Learn that the image of a linear map can be calculated by performing a matrix-vector multiplication
Background

• The two operations we can perform on vectors is:
  – Scalar multiplication: $\alpha \mathbf{u}$
  – **Vector** addition: $\mathbf{u} + \mathbf{v}$

Definition:

A map $A: \mathcal{U} \rightarrow \mathcal{V}$ is said to be a *linear map* if both

$A(\alpha \mathbf{u}) = \alpha A\mathbf{u}$

$A(\mathbf{u}_1 + \mathbf{u}_2) = A\mathbf{u}_1 + A\mathbf{u}_2$

– The response to a linear combination of inputs is equal to that same linear combination of the responses to the individual inputs

$A(\alpha \mathbf{u}_1 + \beta \mathbf{u}_2) = \alpha A\mathbf{u}_1 + \beta A\mathbf{u}_2$
The superposition principle

- Visually, we can approach this in two ways:
Results

Theorem

If $A: \mathcal{U} \to \mathcal{V}$ is linear, then $A0_{\mathcal{U}} = 0_{\mathcal{V}}$.

- That is, the zero vector in $\mathcal{U}$ is mapped to the zero vector in $\mathcal{V}$

Proof:

Let $u$ be any vector in $\mathcal{U}$

Recall that $0_{\mathcal{U}} = 0$

Thus, $A0_{\mathcal{U}} = A(0u) = 0Au$

but $Au$ is a vector in $\mathcal{V}$

$= 0_{\mathcal{V}} \blacksquare$
Theorem

If $A: \mathcal{U} \rightarrow \mathcal{V}$ is linear, then $A$ maps lines onto lines or a single point.

Proof:

A line is of the form $\mathbf{u}_1 + \alpha \mathbf{u}_2$ in $\mathcal{U}$

Thus, $A(\mathbf{u}_1 + \alpha \mathbf{u}_2) = A\mathbf{u}_1 + \alpha A\mathbf{u}_2$

If $A\mathbf{u}_2 = \mathbf{0}_\mathcal{V}$, then $A(\mathbf{u}_1 + \alpha \mathbf{u}_2) = A\mathbf{u}_1$

Otherwise, $A\mathbf{u}_2 \neq \mathbf{0}_\mathcal{V}$, in which case, the result is a line in $\mathcal{V}$.
Results

Theorem

If $A: \mathcal{U} \rightarrow \mathcal{V}$ is linear and $A\mathbf{u}_1 = A\mathbf{u}_2 = \mathbf{v} \in \mathcal{V}$, then $A(\alpha \mathbf{u}_1 + (1 - \alpha)\mathbf{u}_2) = \mathbf{v}$ for all $\alpha \in \mathbb{F}$.

Proof:

$$A(\alpha \mathbf{u}_1 + (1 - \alpha)\mathbf{u}_2) = \alpha A\mathbf{u}_1 + (1 - \alpha)A\mathbf{u}_2$$

$$= \alpha \mathbf{v} + (1 - \alpha)\mathbf{v}$$

$$= (\alpha + 1 - \alpha)\mathbf{v} = 1 \cdot \mathbf{v} = \mathbf{v} \blacksquare$$

- Note, this says: if there are two solutions, to $A\mathbf{u} = \mathbf{v}$, then there are infinitely many solutions
Results

Theorem
If $A: \mathcal{U} \rightarrow \mathcal{V}$ is linear and $A\mathbf{u}_1 = A\mathbf{u}_2 = \mathbf{v} \in \mathcal{V}$, then $A(\mathbf{u}_1 - \mathbf{u}_2) = \mathbf{0}_\mathcal{V}$.

Proof:

\[
A(\mathbf{u}_1 - \mathbf{u}_2) = A\mathbf{u}_1 - A\mathbf{u}_2 \\
= \mathbf{v} - \mathbf{v} \\
= \mathbf{0}_\mathcal{V} \blacksquare
\]
Background

Definition:
A linear map $A: \mathcal{U} \to \mathcal{U}$ is said to be a linear operator if the domain and codomain are the same vector space.

Definition
Given two linear maps, $A, B: \mathcal{U} \to \mathcal{V}$, we will say that $A = B$ if and only if $Au = Bu$ for all $u \in \mathcal{U}$. 
Proving linearity

• You can show a map \( A: \mathcal{U} \rightarrow \mathcal{V} \) is linear by either showing:
  1. \( A(\alpha \mathbf{u}) = \alpha A\mathbf{u} \) for all \( \mathbf{u} \in \mathcal{U} \) and all \( \alpha \in F \)
  2. \( A(\mathbf{u}_1 + \mathbf{u}_2) = A\mathbf{u}_1 + A\mathbf{u}_2 \) for all \( \mathbf{u}_1, \mathbf{u}_2 \in \mathcal{U} \)
     – You can show one or the other, Requirement 1 is true if and only if Requirement 2 is true.

• To show that \( A: \mathcal{U} \rightarrow \mathcal{V} \) is not linear, it is only necessary to show one of:
  1. \( A(0_\mathcal{U}) \neq 0_\mathcal{V} \)
  2. \( A(\alpha \mathbf{u}) \neq \alpha A(\mathbf{u}) \) for one \( \mathbf{u} \in \mathcal{U} \) and one \( \alpha \in F \)
  3. \( A(\mathbf{u}_1 + \mathbf{u}_2) \neq A\mathbf{u}_1 + A\mathbf{u}_2 \) for two specific \( \mathbf{u}_1, \mathbf{u}_2 \in \mathcal{U} \)
Examples

• Let $A: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ be defined as $Au = A \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = u_1 \begin{pmatrix} 0.5 \\ 1.2 \\ 0.9 \end{pmatrix} + u_2 \begin{pmatrix} 0.8 \\ 2.3 \\ 1.7 \end{pmatrix}$

• Now, $A(\alpha u_1 + \beta u_2) = (\alpha u_{1,1} + \beta u_{2,1}) \begin{pmatrix} 0.5 \\ 1.2 \\ 0.9 \end{pmatrix} + (\alpha u_{1,2} + \beta u_{2,2}) \begin{pmatrix} 0.8 \\ 2.3 \\ 1.7 \end{pmatrix}$

$$= \alpha \begin{pmatrix} 0.5 \\ 1.2 \\ 0.9 \end{pmatrix} + \beta \begin{pmatrix} 0.5 \\ 1.2 \\ 0.9 \end{pmatrix} + \alpha \begin{pmatrix} 0.8 \\ 2.3 \\ 1.7 \end{pmatrix} + \beta \begin{pmatrix} 0.8 \\ 2.3 \\ 1.7 \end{pmatrix}$$

$$= \alpha \begin{pmatrix} u_{1,1} \\ u_{1,2} \\ u_{2,1} \\ u_{2,2} \end{pmatrix} + \beta \begin{pmatrix} u_{1,1} \\ u_{1,2} \\ u_{2,1} \\ u_{2,2} \end{pmatrix}$$

$$= \alpha Au_1 + \beta Au_2$$
Examples

- Let $B: \mathbb{R}^4 \rightarrow \mathbb{R}^2$ be defined as

$$Bu = B \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = u_1 \begin{pmatrix} 3.7 \\ 9.2 \end{pmatrix} + u_2 \begin{pmatrix} 4.8 \\ 6.0 \end{pmatrix} + u_3 \begin{pmatrix} 0.5 \\ 3.4 \end{pmatrix} + u_4 \begin{pmatrix} 0.7 \\ 2.6 \end{pmatrix} + \begin{pmatrix} 8.9 \\ 1.2 \end{pmatrix}$$

- Well, $B0_4 = B \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0 \cdot \begin{pmatrix} 3.7 \\ 9.2 \end{pmatrix} + 0 \cdot \begin{pmatrix} 4.8 \\ 6.0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0.5 \\ 3.4 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0.7 \\ 2.6 \end{pmatrix} + \begin{pmatrix} 8.9 \\ 1.2 \end{pmatrix} = \begin{pmatrix} 8.9 \\ 1.2 \end{pmatrix}$

and $\begin{pmatrix} 8.9 \\ 1.2 \end{pmatrix} \neq 0_2$, so this $B$ is not linear
Examples

• Let $C: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be defined as $Cu = C \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} u_1 - u_2 \\ u_2 - u_3 \end{pmatrix}$

• Now, $Cu = C \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = u_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix}$,

so using the same approach as for $A$, we can show that $C$, too, is linear
Examples

Let $D: \mathbb{R}^4 \to \mathbb{R}^4$ be defined as $Du = D\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 3u_1 + 7u_2 - 9u_4 \\ 2u_2 - 4u_3 - 3u_4 \\ 6u_1 + 5u_3 + 4u_4 \\ 9u_1 - 6u_2 - 7u_3 \end{pmatrix}$.

Now, $Du = \begin{pmatrix} 3u_1 + 7u_2 - 9u_4 \\ 2u_2 - 4u_3 - 3u_4 \\ 6u_1 + 5u_3 + 4u_4 \\ 9u_1 - 6u_2 - 7u_3 \end{pmatrix} = u_1 \begin{pmatrix} 3 \\ 0 \\ 6 \\ 9 \end{pmatrix} + u_2 \begin{pmatrix} 7 \\ 2 \\ 0 \\ -6 \end{pmatrix} + u_3 \begin{pmatrix} 0 \\ -4 \\ 5 \\ -7 \end{pmatrix} + u_4 \begin{pmatrix} -9 \\ -3 \\ 4 \end{pmatrix}$.

so using the same approach as for $A$ and $C$, we can show that $D$, too, is linear.
Examples

• Let \( E : \mathbb{R}^2 \to \mathbb{R}^2 \) be defined as \( E \mathbf{u} = E \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \cos(u_1) - 1 \\ \sin(u_2) \end{pmatrix} \)

• Now, \( E \mathbf{0}_2 = E \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(0) - 1 \\ \sin(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{0}_2 \), so we can’t do this easily...

• However, \( E \begin{pmatrix} 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{pmatrix} = E \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} \cos(2) - 1 \\ \sin(2) \end{pmatrix} \) and \( 2 \cdot E \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \cos(1) - 2 \\ 2 \sin(1) \end{pmatrix} \)

  – But \( 1.6829 \approx 2 \sin(1) \neq \sin(2) \approx 0.90930 \)

  – Thus, \( E \) is not linear
Examples

- Let $F: \mathbb{R}^1 \to \mathbb{R}^1$ be defined as $F(u) = F(u_1) = (2u_1 + 1)$
- Now, $F0_1 = (1) \neq 0_1$, so $F$ is not linear
Examples

• Let $G: \mathbb{R}^2 \to \mathbb{R}^2$ be defined as $Gu = G \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} u_1^3 \\ u_2^3 \end{pmatrix}$

- Now, $G \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, yet $G \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $G \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ 8 \end{pmatrix}$,

  but $2G \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \neq \begin{pmatrix} 8 \\ 8 \end{pmatrix} = G \begin{pmatrix} 2 \cdot 1 \\ 2 \cdot 1 \end{pmatrix}$

- Thus, $G$ is not linear
A general observation...

- For finite-dimensional vector spaces, it seems that a mapping $A: \mathbb{F}^n \rightarrow \mathbb{F}^m$ is linear if and only if $A$ is defined as the entries of $u$ being the coefficients of a linear combination of $n$ vectors in $\mathbb{F}^m$

- That is, there exist $n$ vectors $v_1, \ldots, v_n$ such that
  \[ A u = u_1 v_1 + \cdots + u_n v_n \]
A general observation...

• Can we go the other way?
  Given a linear mapping $A : \mathbb{F}^n \to \mathbb{F}^m$, does it necessarily define a linear combination of $n$ vectors in $\mathbb{F}^m$ with coefficients from the vector $\mathbf{u}$ from $\mathbb{F}^n$

• That is, if $A$ is linear, do there exist $n$ vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ such that
  $$A\mathbf{u} = u_1 \mathbf{v}_1 + \cdots + u_n \mathbf{v}_n$$
A general observation...

- Solution:
  - In the domain $\mathbb{F}^n$,
    
    let $\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_n$ be the canonical basis, so $\hat{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, etc.

  - Define $v_k = A\hat{e}_k$, so given any vector $u \in \mathbb{F}^n$, we see
    
    $A u = A (u_1 \hat{e}_1 + u_2 \hat{e}_2 + \cdots + u_n \hat{e}_n)$
    
    $= u_1 A\hat{e}_1 + u_2 A\hat{e}_2 + \cdots + u_n A\hat{e}_n$
    
    $= u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$
Matrix representation of a linear map

- Now, given a linear map $A: \mathbb{F}^n \rightarrow \mathbb{F}^m$, we will represent $A$ as the $m \times n$ matrix

\[
A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}
\]

- Each column $a_k \in \mathbb{F}^m$ and we will describe $Au$ as the matrix-vector multiplication

\[
Au = \left( a_1 \ a_2 \ \cdots \ a_n \right) u = u_1a_1 + u_2a_2 + \cdots + u_na_n
\]
Matrix representation of a linear map

- It is also possible to think of matrix-vector multiplication as follow:

\[
A\mathbf{u} = \begin{pmatrix}
    a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & \cdots & a_{1,n} \\
    a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & \cdots & a_{2,n} \\
    a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & \cdots & a_{3,n} \\
    \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
    a_{m,1} & a_{m,2} & a_{m,3} & a_{m,4} & \cdots & a_{m,n}
\end{pmatrix}
\begin{pmatrix}
    \mathbf{u}_1 \\
    \mathbf{u}_2 \\
    \mathbf{u}_3 \\
    \vdots \\
    \mathbf{u}_n
\end{pmatrix}
= \begin{pmatrix}
    \mathbf{v}_1 \\
    \mathbf{v}_2 \\
    \mathbf{v}_3 \\
    \vdots \\
    \mathbf{v}_m
\end{pmatrix}
\]

- For each row of the matrix:
  
  \[a_{1,1}u_1 + a_{1,2}u_2 + a_{1,3}u_3 + a_{1,4}u_4 + \cdots + a_{1,n}u_n = v_1\]
  
  \[a_{2,1}u_1 + a_{2,2}u_2 + a_{2,3}u_3 + a_{2,4}u_4 + \cdots + a_{2,n}u_n = v_2\]
  
  \[a_{3,1}u_1 + a_{3,2}u_2 + a_{3,3}u_3 + a_{3,4}u_4 + \cdots + a_{3,n}u_n = v_3\]
  
  \[a_{m,1}u_1 + a_{m,2}u_2 + a_{m,3}u_3 + a_{m,4}u_4 + \cdots + a_{m,n}u_n = v_m\]
Examples

- Recall \(A: \mathbb{R}^2 \rightarrow \mathbb{R}^3\) was \(Au = A\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = u_1 \begin{pmatrix} 0.5 \\ 1.2 \\ 0.9 \end{pmatrix} + u_2 \begin{pmatrix} 0.8 \\ 2.3 \\ 1.7 \end{pmatrix}\)

  - The matrix representation is \(A = \begin{pmatrix} 0.5 & 0.8 \\ 1.2 & 2.3 \\ 0.9 & 1.7 \end{pmatrix}\)

  - For example, if \(u = \begin{pmatrix} 0.3 \\ -0.1 \end{pmatrix}\) then

\[
Au = \begin{pmatrix} 0.5 & 0.8 \\ 1.2 & 2.3 \\ 0.9 & 1.7 \end{pmatrix} \begin{pmatrix} 0.3 \\ -0.1 \end{pmatrix} = \begin{pmatrix} 0.07 \\ 0.13 \\ 0.10 \end{pmatrix}
\]
Examples

- Recall $C : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ was $Cu = C \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} u_1 - u_2 \\ u_2 - u_3 \end{pmatrix}$

  - The matrix representation is $C = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$

  - For example, if $u = \begin{pmatrix} 0.5 \\ 0.2 \\ -0.7 \end{pmatrix}$ then

    $$Cu = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0.5 \\ 0.2 \\ -0.7 \end{pmatrix} = \begin{pmatrix} 0.3 \\ 0.9 \end{pmatrix}$$
Examples

• Recall $D: \mathbb{R}^4 \to \mathbb{R}^4$ was $Du = D\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 3u_1 + 7u_2 - 9u_4 \\ 2u_2 - 4u_3 - 3u_4 \\ 6u_1 + 5u_3 + 4u_4 \\ 9u_1 - 6u_2 - 7u_3 \end{pmatrix}$

  \begin{pmatrix}
  3 & 7 & 0 & -9 \\
  0 & 2 & -4 & -3 \\
  6 & 0 & 5 & 4 \\
  9 & -6 & -7 & 0
  \end{pmatrix}

  – The matrix representation is $D = \begin{pmatrix} 0.2 \\ -0.3 \\ 0.1 \\ 0.4 \end{pmatrix}$

  – If $u = \begin{pmatrix} 0.2 \\ -0.3 \\ 0.1 \\ 0.4 \end{pmatrix}$ then $Du = \begin{pmatrix} 3 & 7 & 0 & -9 \\ 0 & 2 & -4 & -3 \\ 6 & 0 & 5 & 4 \\ 9 & -6 & -7 & 0 \end{pmatrix} \begin{pmatrix} 0.2 \\ -0.3 \\ 0.1 \\ 0.4 \end{pmatrix} = \begin{pmatrix} -5.1 \\ -2.2 \\ 3.3 \\ 2.9 \end{pmatrix}$
Summary

• Following this topic, you now
  – Understand the definition of linearity
  – Know the test for linearity
  – Know also how to show a mapping is not linear
  – Know that the zero vector is always mapped to the zero vector by a linear mapping
  – Have observed a number of examples
  – Know that a finite-dimensional linear map can always be represented by a matrix
  – Understand matrix-vector multiplication
8.4 Special linear maps and linear operators

Instructor: Prof. Chrystopher L. Nehaniv, Ph.D.

slides by Douglas Wilhelm Harder, LEL, M.Math.
dwharder@uwaterloo.ca
dwharder@gmail.com
Introduction

- In this topic, we will
  - Define a number of special linear maps and operators
    - The zero map
    - The identity operator
    - Diagonal maps
    - Super-diagonal maps
    - Sub-diagonal maps
    - Tri-diagonal maps
  - Also introduce the delay and advance operators for discrete-time signals
The zero map and operator

• The zero map $O: \mathcal{U} \rightarrow \mathcal{V}$ maps every vector in $\mathcal{U}$ onto the zero vector $0_{\mathcal{V}}$ in $\mathcal{V}$

Theorem

The zero map is linear.

Proof:

\[ O(\alpha \mathbf{u}_1 + \beta \mathbf{u}_2) = 0_{\mathcal{V}} \]
\[ \alpha O\mathbf{u}_1 + \beta O\mathbf{u}_2 = \alpha 0_{\mathcal{V}} + \beta 0_{\mathcal{V}} = 0_{\mathcal{V}} \]
Since these are equal, this map is linear. □

• The zero map $O: \mathbb{F}^n \rightarrow \mathbb{F}^m$ is represented by an $m \times n$ matrix of all zeros, and is often denoted $O_{m,n}$
The identity operator

- The identity operator $I: \mathcal{U} \rightarrow \mathcal{U}$ maps every vector in $\mathcal{U}$ onto itself.

Theorem

The identity operator is linear.

Proof:

$I(\alpha \mathbf{u}_1 + \beta \mathbf{u}_2) = \alpha \mathbf{u}_1 + \beta \mathbf{u}_2$

$\alpha I\mathbf{u}_1 + \beta I\mathbf{u}_2 = \alpha \mathbf{u}_1 + \beta \mathbf{u}_2$

Since these are equal, this operator is linear.
The identity operator

- The identity operator $I: \mathbb{F}^n \rightarrow \mathbb{F}^n$ is represented by an $n \times n$ matrix of all zeros except for all ones on the diagonal, and is denoted $I_n$
- For example,

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

- Note that the columns of the identity operator are the $n$ canonical basis vectors of $\mathbb{F}^n$
Diagonal maps and operators

- Any $m \times n$ matrix with zeros everywhere except perhaps on the diagonal called a diagonal map or diagonal operator
  - Generally, we only use diagonal operators

\[
\begin{pmatrix}
\lambda_1 & 0 & 0 & 0 \\
0 & \lambda_2 & 0 & 0 \\
0 & 0 & \lambda_3 & 0 \\
0 & 0 & 0 & \lambda_4
\end{pmatrix}
\begin{pmatrix}
\lambda u_1 \\
\lambda_2 u_2 \\
\lambda_3 u_3 \\
\lambda_4 u_4
\end{pmatrix}
\]
Super-diagonal maps and operators

• Any $m \times n$ matrix with zeros everywhere except perhaps on the super-diagonal is called a super-diagonal map or super-diagonal operator

\[
S = \begin{pmatrix}
0 & \lambda_1 & 0 & 0 \\
0 & 0 & \lambda_2 & 0 \\
0 & 0 & 0 & \lambda_3 \\
0 & 0 & 0 & 0
\end{pmatrix}, \text{ then } Su = \begin{pmatrix}
\lambda_1 u_2 \\
\lambda_2 u_3 \\
\lambda_3 u_4 \\
0
\end{pmatrix}
\]

• Note that if all the super-diagonal entries are one, this has an effect of shifting the entries of $u$ up by one
Sub-diagonal maps and operators

- Any $m \times n$ matrix with zeros everywhere except perhaps on the sub-diagonal is called a sub-diagonal map or sub-diagonal operator.

$$S = \begin{pmatrix}
0 & 0 & 0 & 0 \\
\lambda_1 & 0 & 0 & 0 \\
0 & \lambda_2 & 0 & 0 \\
0 & 0 & \lambda_3 & 0 \\
\end{pmatrix}$$

If $S = \begin{pmatrix}
0 & 0 & 0 & 0 \\
\lambda_1 & 0 & 0 & 0 \\
0 & \lambda_2 & 0 & 0 \\
0 & 0 & \lambda_3 & 0 \\
\end{pmatrix}$, then $S\mathbf{u} = \begin{pmatrix}
0 \\
\lambda_1 u_1 \\
\lambda_2 u_2 \\
\lambda_3 u_3 \\
\end{pmatrix}$

- Note that if all the sub-diagonal entries are one, this has an effect of shifting the entries of $\mathbf{u}$ down by one.
Tri-diagonal maps and operators

- Any $m \times n$ matrix with zeros everywhere except perhaps on the super-diagonal, the diagonal and the sub-diagonal is called a tri-diagonal map or tri-diagonal operator.

- If $S = \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & 0 & 0 & 0 \\ \lambda_{2,1} & \lambda_{2,2} & \lambda_{2,3} & 0 & 0 \\ 0 & \lambda_{3,2} & \lambda_{3,3} & \lambda_{3,4} & 0 \\ 0 & 0 & \lambda_{4,3} & \lambda_{4,4} & \lambda_{4,5} \\ 0 & 0 & 0 & \lambda_{5,4} & \lambda_{5,5} \end{pmatrix}$, then $Su = \begin{pmatrix} \lambda_{1,1}u_1 + \lambda_{1,2}u_2 \\ \lambda_{2,1}u_1 + \lambda_{2,2}u_2 + \lambda_{2,3}u_3 \\ \lambda_{3,2}u_2 + \lambda_{3,3}u_3 + \lambda_{3,4}u_4 \\ \lambda_{4,3}u_4 + \lambda_{4,4}u_4 + \lambda_{4,5}u_5 \\ \lambda_{5,4}u_4 + \lambda_{5,5}u_5 \end{pmatrix}$.
Operators on discrete-time signals

- For the vector space of discrete-time signals, two common operators are the delay and advance operators.

- If $x = (x[0], x[1], x[2], x[3], \ldots)$, then the action of the delay operator $D$ is defined as
  $$Dx = (0, x[0], x[1], x[2], \ldots)$$

  and the action of the advance operator $E$ is defined as
  $$Ex = (x[1], x[2], x[3], x[4], \ldots)$$
Summary

• Following this topic, you now
  – Are aware of some of the more common linear maps and operators
  – Know the delay and advance operators for discrete-time signals
8.5 The range of a linear map

Instructor: Prof. Chrystopher L. Nehaniv, Ph.D.

slides by Douglas Wilhelm Harder, LEL, M.Math.
dwharder@uwaterloo.ca
dwharder@gmail.com
Introduction

• In this topic, we will
  – Define the range of a linear map
  – Consider some consequences
  – Look at finite-dimensional maps
  – Consider the advance and delay operators
Describing polynomials

• Recall that a polynomial of degree $n$ is a collection of $n + 1$ coefficients
  – What is the significance of these coefficients?
  – Can changing one number completely change the properties of that polynomial?
    • Short answer, yes...
  – What you did in secondary school is determine the important characteristics of polynomials:
    • The roots of a polynomial
    • The coefficient of the leading term
Describing linear maps

• An $m \times n$ matrix is a collection of $mn$ entries
  – What is the significance of this?
  – Can changing one number completely change the properties of that matrix?
    • Short answer, yes...
  – What we want to do is describe the useful characteristics of linear maps
    • These characteristics help understand what a linear map does, and how to design linear maps
  – The first two such features are to describe the range and null space of that linear map
    • We will start with the range
The zero map and operator

• Given a linear map $A: \mathcal{U} \rightarrow \mathcal{V}$, the range is the collection of all vectors in $\mathcal{V}$ such that at least one vector in $\mathcal{U}$ is mapped to that vector.

• That is, $\text{range}(A)$ equals all vectors $Au$ where $u$ is a vector in $\mathcal{U}$
  – You will note that $\text{range}(A) \subseteq \mathcal{V}$
  – Also, $\text{range}(A)$ always has at least one element: $0_\mathcal{V}$
The range of a linear map

The range is a subspace

**Theorem**

If $A: U \to V$ is a linear map, the range of $A$ is a subspace of $V$.

**Proof:**

If $v_1, v_2 \in \text{range}(A)$, then we must show $\alpha v_1 + \beta v_2 \in \text{range}(A)$

If $v_1, v_2 \in \text{range}(A)$, then there must be $u_1, u_2 \in U$

such that $Au_1 = v_1$ and $Au_2 = v_2$

Now, because $U$ is a vector space, $\alpha u_1 + \beta u_2 \in U$

But because $A$ is linear, $A(\alpha u_1 + \beta u_2) = \alpha Au_1 + \beta Au_2$

$= \alpha v_1 + \beta v_2$

So since $\alpha u_1 + \beta u_2 \in U$, it follows

$A(\alpha u_1 + \beta u_2) = \alpha v_1 + \beta v_2 \in \text{range}(A)$

Thus, the range is closed under vector addition and scalar multiplication, so it is a subspace. □
The image of a subspace is a subspace

Theorem

If $S$ is a subspace of $\mathcal{U}$ and $A: \mathcal{U} \rightarrow \mathcal{V}$ is a linear map, the image of $S$, $AS$, is a subspace of $\mathcal{V}$.

Proof:

The proof is similar to the previous proof.
The range of the zero map

Theorem

If $A: U \rightarrow V$ is a linear map,

$$\text{range}(A) = \{0_V\} \text{ if and only if } A = O$$

Proof

If $\text{range}(A) = \{0_V\}$, then $Au = 0_V$ for all vectors $u$

However, the zero map $O$ is that map such that

$$u = 0_V \text{ for all vectors } u$$

Two maps are equal if $Au = 0u$ for all $u$,

so the range is just the zero vector if and only if the map is the zero map. ■
Definition: onto

Definition

If $A: \mathcal{U} \rightarrow \mathcal{V}$ and $\text{range}(A) = \mathcal{V}$, we say that $A$ is onto.

It is also said that $A$ is surjective.

– Recall that

• A surcharge is a charge on top of what you’re already paying for a product
• A surplus is over and above what you need
• Thus, a map is surjective if every point in the domain is mapped onto
Definition: onto

• What is “onto”?
  – Any odd degree polynomial with real coefficients of a real variable maps onto the reals. (It takes arbitrarily large values with opposite signs as the argument as to plus and to minus infinite, so must take every value in between.)
  – No even degree polynomial with real coefficients of a real variable maps onto the reals
    • Such a polynomial always has either a global minimum or a global maximum
      – For $t > 0$, $\ln(t)$ maps onto the real numbers
      – The range of $e^t$ for real $t$ is the open interval $(0, \infty)$, so therefore it is not onto
      – On the other hand, $e^t \sin(t)$ is onto the reals
    • Challenge: prove that every non-constant polynomial with complex coefficients maps onto $\mathbb{C}$
Range of finite-dimensional linear mappings

• If $A: \mathbb{F}^n \rightarrow \mathbb{F}^m$, the range of $A$ equals the span of the column vectors
  – That is, if $A = (a_1 \cdots a_n)$, then $\text{range}(A) = \text{span}\{a_1, \ldots, a_n\}$
  – You can use Gaussian elimination to find a basis for the range

• For example,

$$
A = \begin{pmatrix}
1 & 2 & -1 & -3 & -2 \\
3 & 8 & 2 & -11 & -2 \\
4 & 14 & 11 & -18 & 7 \\
-2 & -8 & -8 & 10 & -10
\end{pmatrix} \sim \begin{pmatrix}
1 & 2 & -1 & -3 & -2 \\
0 & 2 & 5 & -2 & 4 \\
0 & 0 & 0 & 0 & 3 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

– Thus, a basis is Columns 1, 2 and 5 of the matrix $A$
– The range is a 3-dimensional subspace of $\mathbb{R}^4$
Range of finite-dimensional linear mappings

- If $A: \mathbb{F}^n \rightarrow \mathbb{F}^m$,
  - The dimension of range($A$) is rank($A$)
  - $A$ is onto if rank($A$) = $m$
Operators on discrete-time signals

- Recall the delay and advance operators on discrete-time signals:
- If \( x = (x[0], x[1], x[2], x[3], \ldots) \),
  \( Dx = (0, x[0], x[1], x[2], \ldots) \) and the range is all signals \( x \) with \( x[0] = 0 \)
  \( Ex = (x[1], x[2], x[3], x[4], \ldots) \) and the range is all signals
Summary

- Following this topic, you now
  - Know the definition of the range
  - Know that the range of a linear map is always a subspace
  - Have become familiar with some additional results
  - Understand the definition of a map being onto
  - Know how to determine the range of a finite-dimensional linear map
8.5 The null space of a linear map

Instructor: Prof. Chrystopher L. Nehaniv, Ph.D.

slides by Douglas Wilhelm Harder, LEL, M.Math.
dwharder@uwaterloo.ca
dwharder@gmail.com
Introduction

• In this topic, we will
  – Define the null space of a linear map
  – Consider some consequences
  – Look at finite-dimensional maps
  – See the relationship between the range and the null space
  – Consider the advance and delay operators
Describing linear maps

• Recall that we want to describe the useful characteristics of a linear map
  – This is not information we can get by looking at, for example, the individual numbers in a given matrix
  – The first two such features are to describe the range and null space of that linear map
    • The last topic covered the range, now we will describe the null space
The null space

• Given a linear map \( A: \mathcal{U} \rightarrow \mathcal{V} \), the null space is the collection of all vectors in \( \mathcal{U} \) that map onto \( 0_\mathcal{V} \).

• That is, \( \text{null}(A) \) equals all vectors \( u \) where \( Au = 0_\mathcal{V} \)
  – You will note that \( \text{null}(A) \subseteq \mathcal{U} \)
  – Also, \( \text{null}(A) \) always has at least one element: \( 0_\mathcal{U} \)
The null space of a linear map

The null space is a subspace

Theorem

If \( A: \mathcal{U} \rightarrow \mathcal{V} \) is a linear map,
the null space of \( A \) is a subspace of \( \mathcal{U} \).

Proof:

If \( u_1, u_2 \in \text{null}(A) \), then we must show \( \alpha u_1 + \beta u_2 \in \text{null}(A) \).

But because \( A \) is linear, \( A(\alpha u_1 + \beta u_2) = \alpha Au_1 + \beta Au_2 = \alpha 0 + \beta 0 = 0 \).

Thus, \( \alpha u_1 + \beta u_2 \in \text{null}(A) \).

Thus, the null space is closed under vector addition and scalar multiplication, so it is a subspace. \( \blacksquare \)
Solutions to $Au = v$

**Theorem**

If $A: \mathcal{U} \rightarrow \mathcal{V}$ is a linear map, and $Au = v$ and $u_0 \in \text{null}(A)$, then $A(u + \alpha u_0) = v$ for all $\alpha \in F$

**Proof:**

Because $A$ is linear, we have $A(u + \alpha u_0) = Au + \alpha \cdot Au_0$

$= v + \alpha \cdot 0$

$= v + 0 = v$
The difference of two solutions to $Au = v$

Theorem

If $A: U \rightarrow V$ is a linear map, and $Au_1 = Au_2 = v$, then $u_1 - u_2 \in \text{null}(A)$.

Proof:

Because $A$ is linear, we have $A(u_1 - u_2) = Au_1 - Au_2$

$= v - v$

$= 0 \checkmark$
Definition: one-to-one

Definition

If $A: \mathcal{U} \rightarrow \mathcal{V}$ and if for each $u \in \mathcal{U}$, the value $Au$ is unique, we will say that $A$ is one-to-one.

That is, each vector $v \in \text{range}(A)$ has a unique pre-image.

It is also said that $A$ is injective.
Definition: one-to-one

• What is “one-to-one”?
  – A function $y(t)$ is one-to-one if different $t$ values evaluate to different $y$ values
  – Any linear polynomial $at + b$ with a non-zero slope is one-to-one
  – The exponential function $e^t$ is one-to-one
  – Any differentiable function $y(t)$ such that $\frac{d}{dt} y(t) > 0$ for most $t$
    and $\frac{d}{dt} y(t) = 0$ for only isolated points is one-to-one
  – No even degree polynomial with real coefficients and of a real variable is one-to-one (it will take arbitrary large values with the same sign as argument goes both to plus infinity and to minus infinity, so by continuity will take the same values on both sides).
  – For $t > 0$, $\ln(t)$ and $\sqrt{t}$ are one-to-one

• Challenge: prove that every polynomial of degree greater than one with complex coefficients is never one-to-one
When is a linear map one-to-one

Theorem
If \( A: U \rightarrow V \) is a linear map, then \( A \) is one-to-one if and only if the null space is just the zero vector; i.e., \( \text{null}(A) = \{0_U\} \).

Proof:
If \( A \) is one-to-one, then only one vector in \( U \) can map onto \( 0_V \), and since \( A0_U = 0_V \), it follows \( \text{null}(A) = \{0_U\} \).

If \( \text{null}(A) = \{0_U\} \), suppose \( Au_1 = Au_2 \).
From a previous theorem, we have \( u_1 - u_2 \in \text{null}(A) \).
But therefore, \( u_1 - u_2 = 0_U \), so \( u_1 = u_2 \).
Therefore, \( A \) is one-to-one. □
Range of finite-dimensional linear mappings

- If $A: \mathbb{F}^n \to \mathbb{F}^m$, the null space is found by solving the homogenous system of linear equations.
- For example,

$$A = \begin{pmatrix} 1 & 2 & -1 & -3 & -2 & 0 \\ 3 & 8 & 2 & -11 & -2 & 0 \\ 4 & 14 & 11 & -18 & 7 & 0 \\ -2 & -8 & -8 & 10 & -10 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 & -3 & -2 & 0 \\ 0 & 2 & 5 & -2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- We see that $3x_5 = 0$, so $x_5 = 0$
- Also, $x_3$ and $x_4$ are free variables
- Thus, $2x_2 + 5x_3 - 2x_4 = 0$, so $x_2 = -2.5x_3 + x_4$
- Thus, $x_1 + 2x_2 - x_3 - 3x_4 = 0$, so $x_1 = 6x_3 + x_4$
The range and the null space

Theorem

If \( A: \mathbb{F}^n \rightarrow \mathbb{F}^m \) is a linear map, the dimension of the null space plus the dimension of the range equals \( n \).

That is, \( \dim(\text{null}(A)) + \dim(\text{range}(A)) = n \)

Proof:

Convert the matrix to row-echelon form with operations:

– Each column associated with a leading non-zero coefficient contributes one vector to the basis of the range, thus the dimension of the range is the number of columns with leading non-zero entries.

– Each column associated with a free variable contributes one vector to the basis of the null space, thus, the dimension of the null space is the number of columns without leading non-zero entries.

There are \( n \) columns, so these two add up to \( n \). \( \blacksquare \)
When a matrix can never be onto

Theorem

If \( A: \mathbb{F}^n \rightarrow \mathbb{F}^m \) is a linear map, and \( n < m \),
then the dimension of the range must be less than \( m \).

Proof:

The maximum \( \text{rank}(A) = n \),
so the maximum dimension of \( \text{range}(A) \) is \( n \)
Thus \( \dim(\text{range}(A)) \leq n < m \). ■

Thus, if \( n < m \), the linear map is never onto.
When a matrix can never be one-to-one

Theorem

If \( A : \mathbb{F}^n \rightarrow \mathbb{F}^m \) is a linear map, and \( n > m \), then the dimension of the null space must be greater than or equal to \( n - m \).

Proof:

The maximum \( \text{rank}(A) = m \), so the maximum dimension of \( \text{range}(A) \) is \( m \):

\[
\dim(\text{range}(A)) \leq m
\]

\[
\dim(\text{range}(A)) \geq -m
\]

But \( \dim(\text{range}(A)) + \dim(\text{null}(A)) = n \)

Adding these last two, we get \( \dim(\text{null}(A)) \geq n - m \).

Thus, if \( n > m \), the linear map is never one-to-one.
When can a matrix be one-to-one and onto?

Theorem

If \( A: \mathbb{F}^n \to \mathbb{F}^m \) is a linear map, then for \( A \) to be one-to-one and onto, \( n = m \)

Proof:

For a linear map to be onto, \( n \geq m \).
For a linear map to be one-to-one, \( n \leq m \).
Consequently, for such a matrix to be one-to-one and onto, it follows that \( n = m \). □

Important: Not all linear operators \( A: \mathbb{F}^n \to \mathbb{F}^n \) are one-to-one and onto, but for a linear map to be one-to-one and onto, it must be a linear operator.
When is a matrix is one-to-one and onto?

Theorem

If \( A : \mathbb{F}^n \to \mathbb{F}^n \) is a linear operator, then \( A \) is either both one-to-one and onto, or neither.

Proof:

If \( A \) is onto, then \( \dim(\text{range}(A)) = n \).

Thus, as \( \dim(\text{range}(A)) + \dim(\text{null}(A)) = n \)

it follows that \( \dim(\text{null}(A)) = 0 \), so \( A \) is one-to-one.

If \( A \) is not onto, then \( \dim(\text{range}(A)) < n \).

Thus, as \( \dim(\text{range}(A)) + \dim(\text{null}(A)) = n \)

adding \( -\dim(\text{range}(A)) > -n \)

Thus, \( \dim(\text{null}(A)) > 0 \), so \( A \) is not one-to-one.
Applications of these theorems

• Given

\[
A = \begin{pmatrix}
2.3 & -1.2 & -0.5 & 2.5 & 3.1 & -1.3 \\
2.1 & 4.5 & 3.2 & -1.3 & -2.4 & 0.9 \\
4.1 & 1.3 & 1.2 & -1.8 & 1.3 & 0.4 \\
-1.3 & -1.5 & -1.8 & 1.0 & -1.3 & -2.4
\end{pmatrix}
\]

• We note \(A: \mathbb{R}^6 \to \mathbb{R}^4\)
  – Because \(6 > 4\), this linear map cannot be one-to-one
    • The minimum dimension of the null space is two
  – If \(\text{rank}(A) = 4\), this matrix is onto,
    but if \(\text{rank}(A) < 4\), this matrix is not onto
Applications of these theorems

- Given

\[
B = \begin{pmatrix}
4.2 & -4.5 & -2.3 \\
2.7 & 1.3 & 0.9 \\
-3.2 & 1.0 & 3.4 \\
5.9 & -3.2 & -0.3 \\
0.8 & -1.3 & -2.7 \\
-0.8 & 0.3 & 0.7 \\
1.3 & 3.9 & -2.8
\end{pmatrix}
\]

- We note \( B: \mathbb{R}^3 \to \mathbb{R}^7 \)
  - Because 3 < 7, this linear map cannot be onto
    - The maximum dimension of the range is three
  - If \( \text{rank}(B) = 3 \), this matrix is one-to-one,
    but if \( \text{rank}(B) < 3 \), this matrix is not one-to-one
Applications of these theorems

• Given

\[
C = \begin{pmatrix}
4.8 & -4.2 & -4.4 & 18.6 \\
8.0 & 8.0 & -9.0 & 11.0 \\
5.6 & 2.9 & -3.0 & 10.8 \\
4.0 & -3.2 & 2.3 & 16.1 \\
\end{pmatrix}
\sim
\begin{pmatrix}
8 & 8 & -9 & 11 \\
0 & -9 & 1 & 12 \\
0 & 0 & 6 & 1 \\
0 & 0 & 0 & -1 \\
\end{pmatrix}
\]

• We note \(C: \mathbb{R}^4 \rightarrow \mathbb{R}^4\)
  – Because, the domain and co-domain are equal, either \(C\) is both one-to-one and onto, or it is neither
  – If \(\text{rank}(C) = 4\), this matrix is one-to-one and onto, but if \(\text{rank}(C) < 4\), this matrix is neither one-to-one nor onto
  – We note that this matrix is both one-to-one and onto
Applications of these theorems

- Given

\[
D = \begin{pmatrix}
  7.2 & -9.0 & -2.4 & -20.8 \\
  -8.0 & 10.0 & 2.0 & 11.0 \\
  -7.2 & 9.0 & -4.2 & 0.9 \\
  7.2 & -9.0 & -5.4 & -21.3
\end{pmatrix} \sim \begin{pmatrix}
  -8 & 10 & 2 & 11 \\
  0 & 0 & -6 & -9 \\
  0 & 0 & 0 & -10 \\
  0 & 0 & 0 & 0
\end{pmatrix}
\]

- We note \( D: \mathbb{R}^4 \rightarrow \mathbb{R}^4 \)
  - Because, the domain and co-domain are equal, either \( D \) is both one-to-one and onto, or it is neither
  - We can determine this by finding the rank
  - We note that this matrix is neither onto nor one-to-one
Operators on discrete-time signals

- Recall the delay and advance operators on discrete-time signals:
  - If \( x = (x[0], x[1], x[2], x[3], \ldots) \)
    \[ Dx = (0, x[0], x[1], x[2], \ldots) \]
    and the null space is just the zero signal
    \[ Ex = (x[1], x[2], x[3], x[4], \ldots) \]
    and the null space is all signals with zero in all entries except for the first

- The rule we just saw, that a linear operator is either one-to-one and onto or neither, only applies to finite-dimensional vector spaces
  - \( D \) and \( E \) are both linear operators
    on the vector space of discrete-time signals
  - \( D \) is one-to-one but not onto
  - \( E \) is onto but not one-to-one
Review

• Given a linear map $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$,
  by using Gaussian elimination and backward substitution
    – We can determine the range and find a basis for it
    – We can determine if it is onto
    – We can determine the null space and find a basis for it
    – We can determine if it is one-to-one

• These are all useful properties to understand about a given linear map or operator
Summary

• Following this topic, you now
  – Know the definition of the null space
  – Know that the null space of a linear map is always a subspace
  – Have become familiar with some additional results
  – Understand the relationship between matrices, the rank and the dimensions of the range and null space
  – Know how to determine a number of properties about a matrix
References

8.8 The vector space of linear maps

Instructor: Prof. Chrystopher L. Nehaniv, Ph.D.

slides by Douglas Wilhelm Harder, LEL, M.Math.
dwharder@uwaterloo.ca
dwharder@gmail.com
Introduction

• In this topic, we will
  – Define scalar multiplication and addition of linear maps
  – Consider the consequences for matrices
  – Prove that the vector space of all such linear maps is itself a vector space
    • The vectors in this vector space are the linear maps
  – Consider a definition for the norm of such a vector
Multiplying a linear map by a scalar

- In order to claim that linear maps are a vector space, we need a definition of:
  - Scalar multiplication of a linear map
  - The sum of two linear maps

- We will see how these definitions apply to matrices
- We will then prove that using these definitions, the collection of all linear maps $A: U \rightarrow V$ forms a vector space
  - Important: each pair of vector spaces $U$ and $V$ defines a different vector space of linear maps
Given a linear map $A: \mathcal{U} \rightarrow \mathcal{V}$, and a scalar $\alpha \in \mathbb{F}$, we will define the linear map $(\alpha A): \mathcal{U} \rightarrow \mathcal{V}$ where 

$$(\alpha A)u = \alpha (Au)$$

That is, the map $\alpha A$ applied to $u$ is has the same effect of

- Applying $A$ to a vector $u$
- Multiplying that result by $\alpha$
Scalar multiplication of a matrix

- Suppose $A: \mathbb{F}^n \to \mathbb{F}^m$ and $\alpha \in \mathbb{F}$
  - Question: What is $\alpha A$?

- Recall that if $A = (a_1 \cdots a_n)$, then
  $$Au = u_1a_1 + \cdots + u_na_n$$

- Thus, we have that if $A = (a_1 \cdots a_n)$, then
  $$(\alpha A)u = \alpha (Au)$$
  $$= \alpha (u_1a_1 + \cdots + u_na_n)$$
  $$= \alpha (u_1a_1) + \cdots + \alpha (u_na_n)$$
  $$= u_1(\alpha a_1) + \cdots + u_n(\alpha a_n)$$
  $$= (\alpha a_1 \cdots \alpha a_n)u$$

- Thus, $\alpha A$ is every entry of $A$ multiplied by $\alpha$
Example

• For example, $A: \mathbb{R}^6 \rightarrow \mathbb{R}^4$ and $-1.5 \in \mathbb{R}$

\[
A = \begin{pmatrix}
2.3 & -1.2 & -0.5 & 2.5 & 3.1 & -1.3 \\
2.1 & 4.5 & 3.2 & -1.3 & -2.4 & 0.9 \\
4.1 & 1.3 & 1.2 & -1.8 & 1.3 & 0.4 \\
-1.3 & -1.5 & -1.8 & 1.0 & -1.3 & -2.4
\end{pmatrix}
\]

− Question: What is $-1.5 \cdot A$ ?

\[
-1.5 \cdot A = \begin{pmatrix}
-3.45 & 1.8 & 0.75 & -3.75 & -4.65 & 1.3 \\
-3.15 & -6.75 & -4.8 & 1.95 & 3.6 & -1.35 \\
-6.15 & -1.95 & -1.8 & 2.7 & -1.95 & -0.6 \\
1.95 & 2.25 & 2.7 & -1.5 & 1.95 & 3.6
\end{pmatrix}
\]
Adding two linear maps

• Similarly, given linear maps $A, B: \mathcal{U} \rightarrow \mathcal{V}$, we will define the linear map $(A + B): \mathcal{U} \rightarrow \mathcal{V}$ where

$$(A + B)\mathbf{u} = A\mathbf{u} + B\mathbf{u}$$

– That is, the map $A + B$ is that map that has the same effect of
  • Applying both $A$ and $B$ to a vector $\mathbf{u}$, separately
  • Adding the resultant vectors
Addition of two matrices

• Suppose $A, B: \mathbb{F}^n \rightarrow \mathbb{F}^m$
  
  – Question: What is $A + B$ ?

• Recall that if $A = (a_1 \cdots a_n)$ and $B = (b_1 \cdots b_n)$, then
  
  $A\mathbf{u} = u_1a_1 + \cdots + u_na_n$ and $B\mathbf{u} = u_1b_1 + \cdots + u_nb_n$

• Thus, we have that
  
  $(A + B)\mathbf{u} = A\mathbf{u} + B\mathbf{u}$
  
  $= (u_1a_1 + \cdots + u_na_n) + (u_1b_1 + \cdots + u_nb_n)$
  
  $= u_1(a_1 + b_1) + \cdots + u_n(a_n + b_n)$
  
  $= (a_1+b_1 \cdots a_n+b_n)\mathbf{u}$

• Thus, $A + B$ sums the corresponding entries of $A$ and $B$
Example

- For example, $A, B: \mathbb{R}^4 \rightarrow \mathbb{R}^3$

\[
A = \begin{pmatrix}
2.5 & 8.9 & -7.1 & 0.3 \\
6.4 & -5.7 & 9.0 & -1.8 \\
-0.8 & 1.9 & 3.5 & -2.4
\end{pmatrix} \quad B = \begin{pmatrix}
3.9 & 2.5 & 3.8 & -1.6 \\
2.8 & 8.7 & -1.4 & -5.8 \\
-2.0 & -4.6 & 3.7 & -1.9
\end{pmatrix}
\]

- Question: What is $A + B$?

\[
A + B = \begin{pmatrix}
6.4 & 11.4 & -3.3 & -1.3 \\
9.2 & 3.0 & 7.6 & -7.6 \\
-2.8 & -2.7 & 7.2 & -4.3
\end{pmatrix}
\]
Example

• Why did we not just define \( \alpha A \) and \( A + B \) for matrices?
  – First, not all linear maps are matrices
  – Second, we derived the formulas for scalar multiplication and matrix addition based on the general definitions of linear maps
  – Finally, we will later define composition of linear maps, and from the definition, we will deduce matrix-matrix multiplication
    • Hint, we will not be multiplying two \( m \times n \) matrices elementwise
Linear maps are a vector space

- We will show that the collection of all linear maps $A: \mathcal{U} \rightarrow \mathcal{V}$ forms a vector space over $\mathbb{F}$

- For many of the proofs, we will show that one linear comb applied to an arbitrary vector $\mathbf{u}$ equals a different expression for an arbitrary vector $\mathbf{u}$
  - We said two maps $A$ and $B$ are equal if $A\mathbf{u} = B\mathbf{u}$ for all $\mathbf{u}$

- For each step, we will either
  - Apply one of the definitions of a linear map
    $$(\alpha A) \mathbf{u} = \alpha (A\mathbf{u}) \quad \text{or} \quad (A + B) \mathbf{u} = A\mathbf{u} + B\mathbf{u}$$
    or a property of vector spaces
Axiom 1

Theorem

If $A: \mathcal{U} \rightarrow \mathcal{V}$ is a linear map, then $\alpha A$ is a linear map.

Proof:

$$(\alpha A)(\beta \mathbf{u}_1 + \gamma \mathbf{u}_2) = \alpha A(\beta \mathbf{u}_1 + \gamma \mathbf{u}_2)$$

$$= \alpha (\beta A \mathbf{u}_1 + \gamma A \mathbf{u}_2)$$

$$= \alpha (\beta A \mathbf{u}_1) + \alpha (\gamma A \mathbf{u}_2)$$

$$= \beta (\alpha A \mathbf{u}_1) + \gamma (\alpha A \mathbf{u}_2)$$

$$= \beta (\alpha A) \mathbf{u}_1 + \gamma (\alpha A) \mathbf{u}_2$$
Axiom 2

Theorem
If \( A: U \to V \) is a linear map, then \( 1A = A \).

Proof:
\[
(1 \cdot A)u = 1 \cdot Au = Au \]

Theorem

If $A: \mathcal{U} \rightarrow \mathcal{V}$ is a linear map, then $(\alpha \beta) A = \alpha (\beta A)$.

Proof:

$((\alpha \beta) A) \mathbf{u} = (\alpha \beta)(A \mathbf{u})$

$= \alpha (\beta A \mathbf{u})$

$= \alpha ((\beta A) \mathbf{u})$

$= (\alpha (\beta A)) \mathbf{u}$ □
Axiom 4

Theorem

If $A, B: \mathcal{U} \rightarrow \mathcal{V}$ are linear maps, then $A + B$ is a linear map.

Proof:

$$(A + B) (\alpha \mathbf{u}_1 + \beta \mathbf{u}_2) = A(\alpha \mathbf{u}_1 + \beta \mathbf{u}_2) + B(\alpha \mathbf{u}_1 + \beta \mathbf{u}_2)$$

$$= \alpha A \mathbf{u}_1 + \beta A \mathbf{u}_2 + \alpha B \mathbf{u}_1 + \beta B \mathbf{u}_2$$

$$= \alpha (A \mathbf{u}_1 + B \mathbf{u}_1) + \beta (A \mathbf{u}_2 + B \mathbf{u}_2)$$

$$= \alpha (A + B) \mathbf{u}_1 + \beta (A + B) \mathbf{u}_2 \quad \blacksquare$$
Axiom 5

Theorem

If $A, B : \mathcal{U} \rightarrow \mathcal{V}$ are linear maps, then $A + B = B + A$.

Proof:

\[
(A + B) \mathbf{u} = A\mathbf{u} + B\mathbf{u} \\
= B\mathbf{u} + A\mathbf{u} \\
= (B + A) \mathbf{u}
\]
Axiom 6

Theorem

If $A, B, C: \mathcal{U} \to \mathcal{V}$ are linear maps, then $(A + B) + C = A + (B + C)$.

Proof:

$$(A + B) + C = (A + B)u + Cu = (Au + Bu) + Cu = Au + (Bu + Cu) = Au + (B + C)u = (A + (B + C))u$$
Axiom 7

Theorem

If $A, B: \mathcal{U} \rightarrow \mathcal{V}$ are linear maps, $\alpha(A + B) = \alpha A + \alpha B$.

Proof:

$$(\alpha(A + B)) \mathbf{u} = \alpha ((A + B) \mathbf{u})$$

$$= \alpha (A \mathbf{u} + B \mathbf{u})$$

$$= \alpha (A \mathbf{u}) + \alpha (B \mathbf{u})$$

$$= (\alpha A) \mathbf{u} + (\alpha B) \mathbf{u}$$

$$= (\alpha A + \alpha B) \mathbf{u} \blacksquare$$
Axiom 8

Theorem

If $A: \mathcal{U} \rightarrow \mathcal{V}$ is a linear map, $(\alpha + \beta)A = \alpha A + \beta A$.

Proof:

$$((\alpha + \beta)A) \mathbf{u} = (\alpha + \beta)(A \mathbf{u})$$
$$= \alpha(A \mathbf{u}) + \beta(A \mathbf{u})$$
$$= (\alpha A)\mathbf{u} + (\beta A)\mathbf{u}$$
$$= (\alpha A + \beta A)\mathbf{u} \blacksquare$$
Axiom 9

Theorem
The zero map \( O: \mathcal{U} \rightarrow \mathcal{V} \) where \( Ou = 0 \) satisfies \( A + O = A \) for all linear maps \( A: \mathcal{U} \rightarrow \mathcal{V} \).

Proof:

\[
(O + A)u = Ou + Au \\
= 0 + Au \\
= Au
\]
Axiom 10

Theorem

For any $A: \mathcal{U} \to \mathcal{V}$, define $(-A)u = -(Au)$, in which case,

$A + (-A) = O$.

Proof:

$(A + (-A))u = Au + (-A)u$

$= Au + (- (Au))$

$= 0$ $\blacksquare$
Example

- Thus, we showed that the collection of all linear maps $A: \mathcal{U} \to \mathcal{V}$ forms a vector space over $\mathbb{F}$
  - The individual maps are the vectors in this vector space
  - This vector space is written as $\mathcal{L}(\mathcal{U}, \mathcal{V})$
  - For example, $\mathcal{L}(\mathbb{R}^4, \mathbb{R}^3)$ is the vector space of all linear maps that map $\mathbb{R}^4$ to $\mathbb{R}^3$
    - It is the collection of all $3 \times 4$ real-valued matrices
Example

• Recall our theorem: if $A \in \mathcal{L}(\mathbb{R}^4, \mathbb{R}^3)$, then $-A = (-1)A$, so if

$$A = \begin{pmatrix} 2.5 & 8.9 & -7.1 & 0.3 \\ 6.4 & -5.7 & 9.0 & -1.8 \\ -0.8 & 1.9 & 3.5 & -2.4 \end{pmatrix}$$

then automatically,

$$-A = (-1)A = \begin{pmatrix} -2.5 & -8.9 & 7.1 & -0.3 \\ -6.4 & 5.7 & -9.0 & 1.8 \\ 0.8 & -1.9 & -3.5 & 2.4 \end{pmatrix}$$

• Also, the zero element can be found by multiplying any matrix by zero, so

$$0 \cdot A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = O_{3,4}$$
Vector space of linear operators

• If the domain and codomain are equal, the vector space of all linear operators $A: \mathcal{U} \rightarrow \mathcal{U}$ forms a special vector space
  – This vector space is written as $\mathcal{L}(\mathcal{U})$
  – For example, $\mathcal{L}(\mathbb{R}^4)$ is the vector space of all linear maps that map $\mathbb{R}^4$ to itself
    • It is the collection of all $4 \times 4$ real-valued square matrices
Example with derivatives

- The 2\textsuperscript{nd} derivative, the derivative, and the identity operator are linear operators on polynomials, so
  - We may write that \( \frac{d^2}{dt^2}, \frac{d}{dt}, I \in \mathcal{L}(\mathbb{R}[t]) \)
  - You may recall \( \mathbb{R}[t] \) is the vector space of all polynomials with real coefficients of a real variable
- Thus, a linear combination of linear operators is a linear operator:
  \[
  \frac{d^2}{dt^2} + 3 \frac{d}{dt} + 7I
  \]
- If \( p(t) = 5t^2 + 2t - 1 \), then
  \[
  \left( \frac{d^2}{dt^2} + 3 \frac{d}{dt} + 7I \right) p(t) = \frac{d^2}{dt^2} p(t) + 3 \frac{d}{dt} p(t) + 7Ip(t) = 10 + 3(10t + 2) + 7\left(5t^2 + 2t - 1\right) = 35t^2 + 44t + 9
  \]
Norms on matrices

• Can we define a norm on a matrix?
  – How about the square root of the sum of the squares of the entries?
  – Unfortunately, this definition has no practical use, consequently, very few people use it

• How about the following?
  – The 2-norm of a linear map $A$ is the maximum that map stretches the 2-norm of a vector in the domain?

\[
\|\|A\|\|_2 = \max_{\substack{u \in \mathcal{U} \\ u \neq 0}} \left\{ \frac{\|Au\|_2}{\|u\|_2} \right\}
\]

  – Under this definition, $\mathcal{L}(\mathcal{U},\mathcal{V})$ is a normed vector space
Inner product on matrices?

• One may think you could define the inner product on matrices just like we defined the inner product on vectors:
  – It is the sum of the element-wise multiplications of the two matrices...
  – Unfortunately, this inner product has no practical applications, so while one could use this definition, no one does
Summary

• Following this topic, you now
  – Know the collection of all linear maps from one vector space to another forms a vector space
  – Know the definition of scalar multiplication of linear maps, and the addition of linear maps
  – Have reviewed proofs that this is a vector space
    • Recall, the vectors in this vector space are the linear maps!
  – Are aware of the norm of a linear map
8.9 The composition of linear maps

Instructor: Prof. Chrystopher L. Nehaniv, Ph.D.

slides by Douglas Wilhelm Harder, LEL, M.Math.
dwharder@uwaterloo.ca
dwharder@gmail.com
Introduction

• In this topic, we will
  – Define composition of linear maps
  – Examine some of the properties
  – Learn that composition is not commutative
  – Determine how to calculate the composition of matrices
    • That is, matrix-matrix multiplication
  – Learn under which circumstances we can multiply two matrices
Composition of linear maps

- Given one linear map $A: \mathcal{U} \rightarrow \mathcal{V}$ and a second linear map $B: \mathcal{V} \rightarrow \mathcal{W}$, then we can define the composition $BA: \mathcal{U} \rightarrow \mathcal{W}$ where
  \[(BA)u = B(Au)\].

**Theorem**

The composition of linear maps is a linear map.

**Proof:**

\[(BA)(\alpha u_1 + \beta u_2) = B(A(\alpha u_1 + \beta u_2)) = B(\alpha Au_1 + \beta Au_2) = \alpha B(Au_1) + \beta B(Au_2) = \alpha (BA)u_1 + \beta (BA)u_2\]
Composition is associative

Theorem

Given three linear maps \( A: U \rightarrow V, \ B: V \rightarrow W \) and \( C: W \rightarrow X \), then \( (CB)A = C(BA) \).

- That is, composition is associative

Proof:

\[
((CB)A) \ u \ = \ (CB)(Au) \\
= \ C(B(Au)) \\
= \ C((BA)u) \\
= \ (C(BA)) \ u \]

\hfill \square
The composition of linear maps

Composition distributes over map addition

Theorem

Given three linear maps \( A_1, A_2 : U \to V \) and \( B : V \to W \), then \( B(A_1 + A_2) = BA_1 + BA_2 \).

– That is, composition distributes over map addition

Proof:

\[
(B(A_1 + A_2)) \, u = B((A_1 + A_2)u) \\
= B(A_1 \, u + A_2 \, u) \\
= B(A_1 \, u) + B(A_2 \, u) \\
= (BA_1) \, u + (BA_2) \, u \\
= (BA_1 + BA_2) \, u \]

\[\square\]
The composition of linear maps

Composition distributes over map addition

Theorem

Given three linear maps \( A : \mathcal{U} \rightarrow \mathcal{V} \) and \( B_1, B_2 : \mathcal{V} \rightarrow \mathcal{W} \), then \( (B_1 + B_2)A = B_1A + B_2A \).

– That is, composition distributes over map addition

Proof:

\[
((B_1 + B_2)A) \, u = (B_1 + B_2)(Au) \\
= B_1(Au) + B_2(Au) \\
= (B_1A)u + (B_2A)u \\
= (B_1A + B_2A) \, u \]

\[
\square
\]
The composition of linear maps

Composition is bilinear

Theorem

Given linear maps $A: \mathcal{U} \rightarrow \mathcal{V}$ and $B: \mathcal{V} \rightarrow \mathcal{W}$ and a scalar $\alpha$, then $\alpha(BA) = (\alpha B)A = B(\alpha A)$.

That is, composition is bilinear

Proof:

$$(\alpha(BA)) \ u = \alpha (\alpha(BA)u)$$

$$= \alpha B(Au)$$

$$= (\alpha B)(Au)$$

$$= (\alpha B)A)u$$

$$= ((\alpha B)A)u$$

$$= B(\alpha(Au))$$

$$= B((\alpha A)u)$$

$$= B((\alpha A)) \ u \ □$$
Composition is not commutative

- First, if $A: \mathcal{U} \to \mathcal{V}$ and $B: \mathcal{V} \to \mathcal{W}$, then $AB$ is not even defined.
- Second, if $A: \mathcal{U} \to \mathcal{V}$ and $B: \mathcal{V} \to \mathcal{U}$, then $BA: \mathcal{U} \to \mathcal{U}$ while $AB: \mathcal{V} \to \mathcal{V}$.
- Thus, we can only even begin to consider commutativity if $A, B: \mathcal{U} \to \mathcal{U}$.
- To show that composition of linear maps is not commutative, we only must find one pair of linear maps between any two vector spaces we want that do not commute.

\[
\begin{bmatrix}
0 \\
1
\end{bmatrix}
\begin{bmatrix}
1 \\
1
\end{bmatrix}
= \begin{bmatrix}
1
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 \\
1
\end{bmatrix}
\begin{bmatrix}
1 \\
0
\end{bmatrix}
= \begin{bmatrix}
0.5 \\
0.5
\end{bmatrix}
\]
What matrix is the composition of matrices?

• Given a linear map \( A: \mathbb{R}^n \rightarrow \mathbb{R}^m \) and \( B: \mathbb{R}^m \rightarrow \mathbb{R}^\ell \), then we know that \( BA: \mathbb{R}^n \rightarrow \mathbb{R}^\ell \) is a linear map.

• We know that all linear maps between finite-dimensional vector spaces representable by matrices, so
  – If \( A \) is an \( m \times n \) matrix and \( B \) is an \( \ell \times m \) matrix, it follows that and \( BA \) is an \( \ell \times n \) matrix, but what matrix?
What matrix is the composition of matrices?

• Recall that if $A = (a_1 \cdots a_n)$ then $Au = u_1 a_1 + \cdots + u_n a_n$, so
  
  $$(BA)u = B(Au) = B(u_1 a_1 + \cdots + u_n a_n)$$
  
  $$= u_1 Ba_1 + \cdots + u_n Ba_n$$
  
  $$= (Ba_1 \cdots Ba_n)u$$

• Thus, if $A = (a_1 \cdots a_n)$ then $BA = (Ba_1 \cdots Ba_n)$,
  that is, multiply each column vector of $A$ by $B$
What matrix is the composition of matrices?

- For example, consider these two matrices:

\[
A = \begin{pmatrix} 2 & 1 & -3 \\ 0 & 4 & -2 \end{pmatrix} \quad B = \begin{pmatrix} -1 & 3 \\ 2 & 0 \\ 4 & 1 \\ 5 & -3 \end{pmatrix}
\]

- We see \(A: \mathbb{R}^3 \rightarrow \mathbb{R}^2\) and \(B: \mathbb{R}^2 \rightarrow \mathbb{R}^4\), so \(BA: \mathbb{R}^3 \rightarrow \mathbb{R}^4\) and

\[
BA = \begin{pmatrix} -1 & 3 \\ 2 & 0 \\ 4 & 1 \\ 5 & -3 \end{pmatrix} \begin{pmatrix} -1 & 3 \\ 2 & 0 \\ 4 & 1 \\ 5 & -3 \end{pmatrix} = \begin{pmatrix} -2 & 11 & -3 \\ 4 & 2 & -6 \\ 8 & 8 & -14 \\ 10 & -7 & -9 \end{pmatrix}
\]
What matrix is the composition of matrices?

- There is a more algorithmic approach to finding the composition, or matrix-matrix product:

\[
BA = \begin{pmatrix}
-1 & 3 \\
2 & 0 \\
4 & 1 \\
5 & -3 \\
\end{pmatrix}
\begin{pmatrix}
2 & 1 & -3 \\
0 & 4 & -2 \\
\end{pmatrix}
= \begin{pmatrix}
-2 & 11 & -3 \\
4 & 2 & -6 \\
8 & 8 & -14 \\
10 & -7 & -9 \\
\end{pmatrix}
\]
Which matrices can be multiplied?

• Important: If the dimensions of the matrices do not align, it makes no sense to discuss matrix-matrix multiplication
  – You cannot multiply a 5-dimensional vector by a $5 \times 4$ matrix
  – A $5 \times 4$ matrix maps $\mathbb{R}^4$ to $\mathbb{R}^5$
  – Similarly, you cannot multiply a $4 \times 7$ matrix by a $3 \times 4$ matrix, because the one on the right maps $\mathbb{R}^4$ to $\mathbb{R}^3$, but the second maps a 7-dimensional vector to $\mathbb{R}^4$
  – However, you can multiply a $3 \times 4$ matrix by a $4 \times 7$ matrix
Summary

• Following this topic, you now
  – Know the definition of the composition of linear maps
  – Understand it to be associative, mapping over addition of linear maps, and is bilinear
  – Know that it is not commutative
  – Know how to find the composition of matrices:
    • Matrix-matrix multiplication
  – Understand how to determine which matrices can be multiplied
8.10 Linear operators

Instructor: Prof. Chrystopher L. Nehaniv, Ph.D.

slides by Douglas Wilhelm Harder, LEL, M.Math.
dwharder@uwaterloo.ca
dwharder@gmail.com
Introduction

• In this topic, we will
  – Review the properties of linear maps with composition
  – Discuss linear operators with composition
  – Discuss the identity operator
  – Define matrix powering
  – Discuss consequences of noncommutativity
  – Describe matrix polynomials
  – Define inverses
Composition of linear maps

• The collection of all linear operators \( \mathcal{L}(U) \) has scalar multiplication, operator addition, and operator composition, and thus has the following properties:
  
  – The collection forms a vector space under scalar multiplication and operator addition
  
  – The additional axioms associated with operator composition include:

11. Composition is associative: \( C(BA) = (CB)A \)
12. Composition distributes over operator addition from the left
   \( B(A_1 + A_2) = BA_1 + BA_2 \)
13. Composition distributes over operator addition from the right
   \( (B_1 + B_2)A = B_1A + B_2A \)
14. Composition is bilinear: \( \alpha (BA) = (\alpha B)A = B(\alpha A) \)
Composition of linear maps

- Operators, however, have one additional axiom:
  15. There is an identity operator $I$ such that $AI = A = IA$ for all linear operators $A$

Theorem
The linear operator $I$ defined as $Iu = u$ serves as an identity operator.

Proof:

$$
(AI) u = A(Iu) \\
= Au \\
= I(Au) \\
= (IA) u \blacksquare
$$

- Note that this means $I$ commutes with all linear operators $A$
Algebra of linear operators

• A vector space together with a composition that satisfies the five additional axioms we saw define a *unital associative algebra*
  – Thus, the focus of this course will be looking at the
    Unital associative algebra of linear operators mapping a vector space to itself
  – Hence the name of the course: linear algebra

• Generally, however, we will continue to refer to $\mathcal{L}(U)$ as the vector space of linear operators mapping a vector space to itself
  – It is assumed it has a composition of operators and an identity operator
Composition of linear maps

- For $\mathcal{L}(\mathbb{R}^n)$, these operators can be represented as $n \times n$ matrices.
  - The identity operator is the $n \times n$ diagonal matrix:

$$I_n = \begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{pmatrix}$$
Invertibility

• Recall that the multiplicative inverse of 1 is 1
  – Notice that $II = I$?
  – Hence, we have found the compositional inverse of 1 is itself
    • We will simply call this the multiplicative inverse

• For any other $A \in \mathcal{L}(\mathcal{U})$,
  if there exists an $B \in \mathcal{L}(\mathcal{U})$ such that $AB = BA = I$,
  then we will say $A$ is invertible and
denote its multiplicative inverse as $A^{-1}$
Composition of linear maps

- There are two additional field axioms that are not satisfied by linear operators:
  - Operator composition is not communitive
  - Not all linear operators have compositional inverses
Composition of linear maps

• For example, we note

\[
\begin{pmatrix}
0 & -1 \\
2 & 1
\end{pmatrix}
\begin{pmatrix}
1 & 1 \\
0 & 1
\end{pmatrix}
= 
\begin{pmatrix}
0 & -1 \\
2 & 3
\end{pmatrix}
\]

\[
\begin{pmatrix}
1 & 1 \\
0 & 1
\end{pmatrix}
\begin{pmatrix}
0 & -1 \\
2 & 1
\end{pmatrix}
= 
\begin{pmatrix}
2 & 0 \\
2 & 1
\end{pmatrix}
\]

• Also note that

\[
\begin{pmatrix}
0 & 0 \\
0 & 1
\end{pmatrix}
\begin{pmatrix}
a & b \\
c & d
\end{pmatrix}
= 
\begin{pmatrix}
0 & 0 \\
c & d
\end{pmatrix}
\neq 
\begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}
\]

• However, 

\[
\begin{pmatrix}
1 & 2 \\
3 & 4
\end{pmatrix}
\begin{pmatrix}
-2 & 1 \\
1.5 & -0.5
\end{pmatrix}
= 
\begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}
\]

\[
\begin{pmatrix}
-2 & 1 \\
1.5 & -0.5
\end{pmatrix}
\begin{pmatrix}
1 & 2 \\
3 & 4
\end{pmatrix}
= 
\begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}
\]
Powers of operators

- We can also define powers of operators:
  For $A \in \mathcal{L}(\mathcal{U})$ and a non-negative integer $n$, define
  
  $$A^n = \begin{cases} 
  I & n = 0 \\
  AA^{n-1} & n > 0 
  \end{cases}$$

  For an invertible $A \in \mathcal{L}(\mathcal{U})$ and an integer $n$, define
  
  $$A^n = \begin{cases} 
  (A^{-1})^{-n} & n < 0 \\
  I & n = 0 \\
  AA^{n-1} & n > 0 
  \end{cases}$$
Powers of operators

Theorem

If $A \in \mathcal{L}(\mathcal{U})$ and $m$ and $n$ are non-negative integers,
then $A^{m+n} = A^m A^n$

Theorem

If $A \in \mathcal{L}(\mathcal{U})$, $A$ is invertible and $m$ and $n$ are integers,
then $A^{m+n} = A^m A^n$

Theorem

If $A \in \mathcal{L}(\mathcal{U})$ and $m$ and $n$ are non-negative integers,
then $(A^m)^n = A^{mn}$

Theorem

If $A \in \mathcal{L}(\mathcal{U})$, $A$ is invertible and $m$ and $n$ are integers,
then $(A^m)^n = A^{mn}$
Powers of operators

• On the other hand, none of these common simplifications work:
  \[(AB)^2 \neq A^2 B^2\]  \[(A + B)^2 \neq A^2 + 2AB + B^2\]

• However, we still have
  \[(\alpha A)^n = \alpha^n A^n\]

• This is more familiar, but is a consequence of \(AI = A\):
  \[(A + I)^n = A^n + \binom{n}{n-1} A^{n-1} + \binom{n}{n-2} A^{n-2} + \cdots + \binom{n}{2} A^2 + \binom{n}{1} A + I\]
  \[(x + 1)^n = x^n + \binom{n}{n-1} x^{n-1} + \binom{n}{n-2} x^{n-2} + \cdots + \binom{n}{2} x^2 + \binom{n}{1} x + 1\]
Efficient calculations of integer powers

- If $A$ is a matrix, the inefficient way to calculate $A^n$ is to start with $A$ and multiply that by $A$ a total of $n - 1$ times.
- This is much more efficient:

\[
A^n = \begin{cases} 
I & n = 0 \\
A & n = 1 \\
(A^k)^2 & n = 2k \\
A(A^k)^2 & n = 2k + 1
\end{cases}
\]

- If $n$ is even, it requires one matrix-matrix multiplication.
- If $n$ is odd, it requires two.
Efficient calculations of integer powers

• For example, to calculate $A^{49}$, we must:
  – Calculate $A^{24}$, square and multiply by $A$
• To calculate $A^{24}$, we must:
  – Calculate $A^{12}$ and square it
• To calculate $A^{12}$, we must:
  – Calculate $A^6$ and square it
• To calculate $A^6$, we must:
  – Calculate $A^3$ and square it
• To calculate $A^3$, we must:
  – Retrieve $A$, square it and multiply by $A$

• This is a total of 7 matrix-matrix multiplications, and not 48 such multiplications

\[
A^n = \begin{cases} 
  I & n = 0 \\
  A & n = 1 \\
  (A^k)^2 & n = 2k \\
  A(A^k)^2 & n = 2k + 1
\end{cases}
\]
Polynomials on operators

• Now that we have defined powers of linear operators, we can now define polynomials on linear operators.

• Let $A$ be a linear operator, then if we define $p$ to be the polynomial such that

$$p(A) = A^2 + 2A + 6I,$$

we now have a polynomial in $A$.

  – If the linear operators are real, the coefficients must be from $\mathbb{R}$.
  – If the linear operators are complex, the coefficients can come from $\mathbb{C}$. 
Polynomials on operators

• For example, given $p$ defined as $p(A) = A^2 + 2A - 6I$, and $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, we note that

$$A^2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$$

$$2A = 2 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}$$

$$-6I = -6 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -6 & 0 \\ 0 & -6 \end{pmatrix}$$

• Thus, $p(A) = A^2 + 2A - 6I = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} + \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} + \begin{pmatrix} -6 & 0 \\ 0 & -6 \end{pmatrix} = \begin{pmatrix} 3 & 14 \\ 21 & 24 \end{pmatrix}$
Summary

• Following this topic, you now
  – Know about the application of the identity operator
  – Understand about the existence of inverses
  – Know that composition is not commutative and not all operators have compositional inverses
  – Know how to find powers of matrices
  – Understand that you can define polynomials on linear operators

• Important:
  – From here on in, we will refer to the composition of linear operators as multiplication,
    and the compositional inverse as the multiplicative inverse
References

[1] https://en.wikipedia.org/wiki/Linear_map
   #Forming_new_linear_maps_from_given_ones
8.11 Row operations

Instructor: Prof. Chrystopher L. Nehaniv, Ph.D.

slides by Douglas Wilhelm Harder, LEL, M.Math.

dwharder@uwaterloo.ca
dwharder@gmail.com
Introduction

• In this topic, we will
  – Review the row operations on a matrix
  – See that each row operation is a linear operator, and thus can be represented by a square matrix
  – Find these matrices
  – Observe that these matrices are invertible
Row operations

• There are three row operations:
  – Swapping two rows
  – Adding a multiple of one row onto another
  – Multiplying a row by a non-zero scalar

• Previously, we represented these operations by an italicized capital R with a subscript

\[ R_{R_i \leftrightarrow R_j} \quad R_{\alpha R_i \rightarrow R_j} \quad R_{\alpha R_i} \]

  – Recall also that all linear maps are represented by such letters
Swapping two rows

• Suppose that $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, so $A$ is representable by an $m \times n$ matrix
• Recall that the matrix-matrix multiplication $BA$ can be found by multiplying each column vector of $A$ by $B$, so
  \[ BA = (Ba_1 \cdots Ba_n) \]
• Question: How can we swaps two entries of each $a_k$?
• Now, swapping two entries in a vector is a linear mapping, for it doesn’t matter if we swap those entries in $\beta u_1 + \gamma u_2$ or swap the two entries of $u_1$ and $u_2$ and then take the same linear combination of the results
  – The resulting vector will be the same dimension: $m$
  – Thus, the linear operator is an $m \times m$ matrix
• Question: what is the matrix $R_{R_i \leftrightarrow R_j}$?
### Swapping two rows

Given a matrix 

\[ R_{R_i \leftrightarrow R_j} \]

where \( i \) and \( j \) are rows to be swapped. The effect of swapping two rows \( R_i \) and \( R_j \) can be represented as:

\[
\begin{pmatrix}
1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
i & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
j & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
m & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1
\end{pmatrix}
\]
Swapping two rows

- Suppose we want to swap Rows 2 and 5 of this matrix
  - Multiply on the left by $R_{R2\leftrightarrow R5}$

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
3 & 6 & 5 \\
0 & 1 & 7 \\
0 & 4 & 2 \\
0 & 2 & 3 \\
0 & -8 & 2
\end{pmatrix}
= 
\begin{pmatrix}
3 & 6 & 5 \\
0 & -8 & 2 \\
0 & 4 & 2 \\
0 & 2 & 3 \\
0 & 1 & 7
\end{pmatrix}
\]
Swapping two rows

- Now, we can undo the operation of swapping two rows by once again swapping two rows
  - Consequently, this linear operator is its own inverse:

\[
R_{R_i\leftrightarrow R_j}^{-1} = R_{R_i\leftrightarrow R_j}
\]

\[
R_{R_i\leftrightarrow R_j} R_{R_i\leftrightarrow R_j} = I
\]

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 \\
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 \\
\end{pmatrix}
= \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
\end{pmatrix} = I_5
\]
Adding a multiple of one row onto another

- Again, \( A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \) and is represented by an \( m \times n \) matrix
- Again, we have
  \[
  BA = (Ba_1 \cdots Ba_n)
  \]

  - Question: How can we add a multiple of one row onto another?
  - Now, adding a multiple of one entry onto another is linear,
    for we can do this mapping after calculating \( \beta u_1 + \gamma u_2 \)
    or we can do this mapping first on both \( u_1 \) and \( u_2 \) and
    then take the same linear combination of the results
      - The resulting vector will be the same dimension: \( m \)
      - Thus, the linear operator is an \( m \times m \) matrix
  - Question: what is the matrix \( R_{\alpha_{Ri \rightarrow Rj}} \)?
Adding a multiple of one row onto another

$$R_{\alpha R_i \rightarrow R_j} =$$

$$\begin{bmatrix}
1 & \cdots & i & \cdots & j & \cdots & m \\
1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\
\end{bmatrix}$$
Adding a multiple of one row onto another

- Suppose we want to add 0.5 times Rows 2 onto Row 3:
  - Multiply on the left by $R_{0.5R2 \rightarrow R3}$

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0.5 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\begin{pmatrix}
3 & 6 & 5 \\
0 & -8 & 2 \\
0 & 4 & 2 \\
0 & 2 & 3 \\
0 & 1 & 7 \\
\end{pmatrix} =
\begin{pmatrix}
3 & 6 & 5 \\
0 & -8 & 2 \\
0 & 0 & 3 \\
0 & 2 & 3 \\
0 & 1 & 7 \\
\end{pmatrix}
\]
Swapping two rows

• Now, we can undo the operation of adding a multiple of one row onto another?
  – Yes, subtract that same multiple of the one row from the other:

\[
R^{-1}_{\alpha R_i \rightarrow R_j} = R_{-\alpha R_i \rightarrow R_j}
\]

\[
R_{\alpha R_i \rightarrow R_j} R_{-\alpha R_i \rightarrow R_j} = R_{-\alpha R_i \rightarrow R_j} R_{\alpha R_i \rightarrow R_j} = I
\]

\[
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & -0.5 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
= I_5
\]
Multiplying a row by a non-zero scalar

- Again, $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and is represented by an $m \times n$ matrix
- Again, we have

$$BA = (Ba_1 \cdots Ba_n)$$

- Question: How can we multiply a row by a non-zero scalar?
- Now, multiplying an entry by a non-zero scalar is linear, for we can do this mapping after calculating $\beta u_1 + \gamma u_2$ or we can do this mapping first on both $u_1$ and $u_2$ and then take the same linear combination of the results
  - The resulting vector will be the same dimension: $m$
  - Thus, the linear operator is an $m \times m$ matrix
- Question: what is the matrix $R_{\alpha R_i}$?
Multiplying a row by a non-zero scalar

\[
R_{aR_i} =
\begin{pmatrix}
1 & \ldots & i & \ldots & j & \ldots & m \\
1 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \ldots & 1 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\
i & 0 & \ldots & 0 & \alpha & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\
0 & \ldots & 0 & 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \ldots & 0 & 0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 0 \\
j & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\
0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
m & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 1
\end{pmatrix}
\]
Multiplying a row by a non-zero scalar

• Suppose we want to multiply Rows 2 by \(-0.125\):
  – Multiply on the left by \(R_{-0.125R2}\)

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 0.125 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
1 & 2 & 1.666 \\
0 & -8 & 2 \\
0 & 0 & 3 \\
0 & 0 & 3.5 \\
0 & 0 & 7.25
\end{pmatrix}
= \begin{pmatrix}
1 & 2 & 1.666 \\
0 & 1 & -0.25 \\
0 & 0 & 3 \\
0 & 0 & 3.5 \\
0 & 0 & 7.25
\end{pmatrix}
\]
Multiplying a row by a non-zero scalar

• Now, we can undo the operation of multiplying a row by a non-zero scalar?
  – Yes, multiply that row by the multiplicative inverse of that scalar:
    \[ R^{-1}_{\alpha R_i} = R_{\alpha^{-1} R_i} \]

    \[ R_{\alpha^{-1} R_i} R_{\alpha R_i} = R_{\alpha R_i} R_{\alpha^{-1} R_i} = I \]

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & -0.125 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & -8 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
= \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
\end{pmatrix} = I_5
\]
Summary

• Following this topic, you now
  – Know that each row operation can be represented by a matrix
  – Understand that each of these matrices has inverses
References

These slides were prepared using the Cambria typeface. Mathematical equations use Times New Roman, and source code is presented using Consolas. Mathematical equations are prepared in MathType by Design Science, Inc. Examples may be formulated and checked using Maple by Maplesoft, Inc.

The photographs of flowers and a monarch butter appearing on the title slide and accenting the top of each other slide were taken at the Royal Botanical Gardens in October of 2017 by Douglas Wilhelm Harder. Please see https://www.rbg.ca/ for more information.
Disclaimer

These slides are provided for the NE 112 *Linear algebra for nanotechnology engineering* course taught at the University of Waterloo. The material in it reflects the authors’ best judgment in light of the information available to them at the time of preparation. Any reliance on these course slides by any party for any other purpose are the responsibility of such parties. The authors accept no responsibility for damages, if any, suffered by any party as a result of decisions made or actions based on these course slides for any other purpose than that for which it was intended.