

Linear algebra 3

Matrices and Linear Transformations

Martin Russell

School of Computer Science, University of Birmingham

January 23, 2020

Overview

- 1 Matrices
 - Introduction to matrices
 - Matrix multiplication
- 2 Linear transformations
 - Definition
 - Any matrix implements a linear transform
 - A linear transformation is defined by what it does to a basis
 - Every linear transformation is defined by a matrix
- 3 Change of basis
 - Change of basis
 - Orthogonal matrices
- 4 Summary

Matrices

A **matrix** A is just a rectangular grid of numbers:

$$A = \begin{bmatrix} a_{11} & a_{12}, & \cdots & a_{1j}, & \cdots, & a_{1N} \\ a_{21} & a_{22}, & \cdots & a_{2j}, & \cdots, & a_{2N} \\ \vdots & \vdots & \cdots & \vdots, & \vdots, & \vdots \\ a_{i1} & a_{i2}, & \cdots & a_{ij}, & \cdots, & a_{iN} \\ \vdots & \vdots & \cdots & \vdots, & \vdots, & \vdots \\ a_{M1} & a_{M2}, & \cdots & a_{Mj}, & \cdots, & a_{MN} \end{bmatrix} \quad (1)$$

A is an M (number of rows) $\times N$ (number of columns) matrix.
In this course the matrix entries are **real** numbers, $a_{ij} \in \mathbb{R}$

Matrix multiplication

- Given an $M \times N$ matrix A and a $P \times Q$ matrix B , the **matrix product** AB exists if and only if $N = P$. In this case $C = AB$ is the $M \times Q$ matrix whose $(i, j)^{th}$ entry is:

$$c_{ij} = ab_{ij} = \sum_{k=1}^P a_{ik} b_{kj} = a_{i*} \cdot b_{*j} \quad (2)$$

where a_{i*} is the i^{th} row of A and b_{*j} is the j^{th} column of B

$$\begin{bmatrix} a_{11} & \cdots & a_{ik} & \cdots & a_{1P} \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{iP} \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ a_{M1} & \cdots & a_{ik} & \cdots & a_{MP} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \vdots & b_{1N} \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ \vdots & \cdots & b_{kj} & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ b_{P11} & \cdots & b_{Pj} & \vdots & b_{PN} \end{bmatrix} \quad (3)$$

Example

Suppose

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \end{bmatrix}, B = \begin{bmatrix} -7 & 8 \\ 9 & 10 \\ -11 & 12 \end{bmatrix}$$

Then, for example

$$ab_{12} = a_{1*} \cdot b_{*2} = 1 \times 8 + (-2) \times 10 + 3 \times 12 = 24 \quad (4)$$

and

$$AB = \begin{bmatrix} 1 & -2 & 3 \\ 4 & 5 & -6 \end{bmatrix} \begin{bmatrix} -7 & 8 \\ 9 & 10 \\ -11 & 12 \end{bmatrix} = \begin{bmatrix} -58 & 24 \\ 83 & 10 \end{bmatrix} \quad (5)$$

Matrix multiplication

- Matrix multiplication is **not commutative**. In other words, even if AB and BA both exist it is not generally true that $AB = BA$
- The $N \times N$ **identity** matrix is the matrix

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

- If A is $N \times N$ then $AI = IA = A$

Matrices as transformations

If A is an $N \times N$ matrix and $\vec{v} \in \mathbb{R}^N$ is a real N -dimensional vector, then $A\vec{v}$ exists and

$$\vec{w} = A\vec{v} \in \mathbb{R}^N$$

In other words, \vec{w} is an N -dimensional vector and A defines a **transformation** T_A from \mathbb{R}^N to itself:

$$T_A : \mathbb{R}^N \rightarrow \mathbb{R}^N, T_A(\vec{v}) = A\vec{v}$$

Definition

- Let \mathbb{V} and \mathbb{W} be vector spaces
- Let $T : \mathbb{V} \rightarrow \mathbb{W}$ be a transformation.
- So, T is a rule that associates each $\vec{v} \in \mathbb{V}$ with $T(\vec{v}) \in \mathbb{W}$
- T is a *linear* transformation if for $\vec{v}_1, \vec{v}_2 \in \mathbb{V}$ and any scalar λ :

$$T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) \quad (6)$$

$$T(\lambda \vec{v}_1) = \lambda T(\vec{v}_1) \quad (7)$$

- Alternatively

$$T(\lambda \vec{v}_1 + \vec{v}_2) = \lambda T(\vec{v}_1) + T(\vec{v}_2) \quad (8)$$

Example: rotations in \mathbb{R}^2

- In \mathbb{R}^2 , rotation anti-clockwise through an angle θ is realized by multiplication by the rotation matrix

$$R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos\theta \end{bmatrix} \quad (9)$$

- Define a linear transformation $T_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by

$$T_\theta(\vec{v}) = R_\theta \vec{v} \quad (10)$$

Then T_θ is a linear transformation.

- This follows from properties of matrix multiplication, i.e.

$$T_\theta(\lambda \vec{v}_1 + \vec{v}_2) = R_\theta(\lambda \vec{v}_1 + \vec{v}_2) = \lambda R_\theta \vec{v}_1 + R_\theta \vec{v}_2 = \lambda T_\theta(\vec{v}_1) + T_\theta(\vec{v}_2)$$

Any matrix defines a linear transformation

- In fact, *any* $M \times N$ matrix A defines a linear transformation

$$T_A : \mathbb{R}^N \rightarrow \mathbb{R}^M \quad (11)$$

$$\text{by: } T_A(\vec{v}) = A\vec{v}, (\vec{v} \in \mathbb{R}^N) \quad (12)$$

- So, are there linear transforms $T_A : \mathbb{R}^N \rightarrow \mathbb{R}^M$ that are **not** defined by a matrix in this way?
- The answer is “no”. *Every* linear transformation is defined by a matrix. Hence there is a 1-to-1 correspondence between linear transformations $T : \mathbb{R}^N \rightarrow \mathbb{R}^M$ and $M \times N$ matrices

Every linear transformation is defined by a matrix

- Let $T : \mathbb{R}^N \rightarrow \mathbb{R}^M$ be a linear transformation
- Let $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_N\}$ be any basis for \mathbb{R}^N ,
- Let $\vec{v} \in \mathbb{R}^N$. Then $\vec{v} = v_1 \vec{e}_1 + v_2 \vec{e}_2 + \dots + v_N \vec{e}_N$, for some scalars v_1, \dots, v_N
- Since T is linear, $T(\vec{v}) = v_1 T(\vec{e}_1) + v_2 T(\vec{e}_2) + \dots + v_N T(\vec{e}_N)$
- T is *completely* defined by what it does to $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_N\}$
- This holds for *any* basis $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_N\}$

Every linear transformation is defined by a matrix (cont.)

- Let $T : \mathbb{R}^N \rightarrow \mathbb{R}^M$ be a linear transformation
- Now let $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_N\}$ be the standard basis for \mathbb{R}^N , i.e

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \vec{e}_N = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \text{ and let } \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{N-1} \\ v_N \end{bmatrix} \quad (13)$$

- Let E be the $M \times N$ matrix whose n^{th} column is $T(\vec{e}_n)$
- Then,

$$T(\vec{v}) = E\vec{v}, (\vec{v} \in \mathbb{V}) \quad (14)$$

- The linear transformation T is matrix multiplication by E

Proof

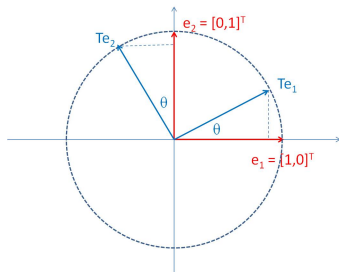
$$E = \begin{bmatrix} T(\vec{e}_1)_1 & T(\vec{e}_2)_1 & \cdots & T(\vec{e}_n)_1 & \cdots & T(\vec{e}_N)_1 \\ T(\vec{e}_1)_2 & T(\vec{e}_2)_2 & \cdots & T(\vec{e}_n)_2 & \cdots & T(\vec{e}_N)_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ T(\vec{e}_1)_n & T(\vec{e}_2)_n & \cdots & T(\vec{e}_n)_n & \cdots & T(\vec{e}_N)_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ T(\vec{e}_1)_M & T(\vec{e}_2)_M & \cdots & T(\vec{e}_n)_M & \cdots & T(\vec{e}_N)_M \end{bmatrix} \quad (15)$$

Proof (continued)

$$\begin{aligned}
 Ev &= \begin{bmatrix} T(\vec{e}_1)_1 & T(\vec{e}_2)_1 & \cdots & T(\vec{e}_n)_1 & \cdots & T(\vec{e}_N)_1 \\ T(\vec{e}_1)_2 & T(\vec{e}_2)_2 & \cdots & T(\vec{e}_n)_2 & \cdots & T(\vec{e}_N)_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ T(\vec{e}_1)_n & T(\vec{e}_2)_n & \cdots & T(\vec{e}_n)_n & \cdots & T(\vec{e}_N)_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ T(\vec{e}_1)_M & T(\vec{e}_2)_M & \cdots & T(\vec{e}_n)_M & \cdots & T(\vec{e}_N)_M \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \\ \vdots \\ v_N \end{bmatrix} \\
 &= \begin{bmatrix} T(\vec{e}_1)_1 v_1 + T(\vec{e}_2)_1 v_2 + \cdots + T(\vec{e}_N)_1 v_N \\ T(\vec{e}_1)_2 v_1 + T(\vec{e}_2)_2 v_2 + \cdots + T(\vec{e}_N)_2 v_N \\ \vdots \\ T(\vec{e}_1)_n v_1 + T(\vec{e}_2)_n v_2 + \cdots + T(\vec{e}_N)_n v_N \\ \vdots \\ T(\vec{e}_1)_M v_1 + T(\vec{e}_2)_M v_2 + \cdots + T(\vec{e}_N)_M v_N \end{bmatrix} = T(\vec{v}) \quad (16)
 \end{aligned}$$

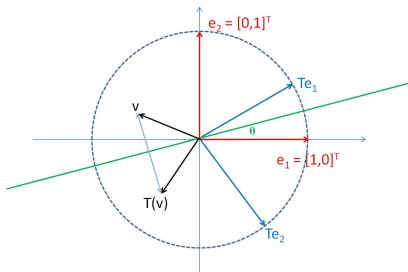
Example 1: Rotations in \mathbb{R}^2

- Suppose that the linear transformation T corresponds to rotation anti-clockwise through an angle θ
- $T(\vec{e}_1) = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$, $T(\vec{e}_2) = \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix}$
- Hence T is implemented by $\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$



Example 2: Reflections in \mathbb{R}^2

- The transformation T corresponds to a reflection through a line through the origin at angle θ to the horizontal axis
- $T(\vec{e}_1) = \begin{bmatrix} \cos(2\theta) \\ \sin(2\theta) \end{bmatrix}$, $T(\vec{e}_2) = \begin{bmatrix} \cos(\frac{\pi}{2} - 2\theta) \\ -\sin(\frac{\pi}{2} - 2\theta) \end{bmatrix} = \begin{bmatrix} \sin(2\theta) \\ -\cos(2\theta) \end{bmatrix}$
- Hence T is implemented by $\begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}$



Summary so far

- There is a 1-to-1 correspondence between:
 - Linear transformations $T : \mathbb{R}^N \rightarrow \mathbb{R}^M$, and
 - $N \times M$ real matrices
- If
 - $T : \mathbb{R}^N \rightarrow \mathbb{R}^M$,
 - $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_N$ is the standard basis for \mathbb{R}^N , and
 - E is the $M \times N$ matrix whose n^{th} column is $T(\vec{e}_n)$, then

$$T(\vec{v}) = E\vec{v}, (\vec{v} \in \mathbb{R}^N) \quad (17)$$

Change of basis

- Let $e = \vec{e}_1, \vec{e}_2, \dots, \vec{e}_N$ be the standard basis for \mathbb{R}^N
- Let $u = \vec{u}_1, \vec{u}_2, \dots, \vec{u}_N$ be another basis for \mathbb{R}^N
- Suppose that we want to transform \mathbb{R}^N so that vectors are represented with respect to u
- The required transformation would satisfy
$$T(\vec{e}_n) = \vec{u}_n, (n = 1, \dots, N)$$
- This transformation is implemented by the $N \times N$ matrix U whose n^{th} column is \vec{u}_n

Orthogonal matrices

- Let U be the $N \times N$ matrix whose n^{th} column is \vec{u}_n , then

$$U^T U = I_N = U U^T. \quad (18)$$

because the i, j^{th} entry of $U^T U$ is $\vec{u}_i \cdot \vec{u}_j$ and $\vec{u}_1, \dots, \vec{u}_N$ is a basis. where I_N is the $N \times N$ identity matrix.

- An $N \times N$ real matrix that satisfies equation (18) is called an **orthogonal** matrix
- An orthogonal matrix U implements a **change of basis** transformation in \mathbb{R}^N :
 - U transforms the standard basis $\vec{e}_1, \dots, \vec{e}_N$ into the new basis $\vec{u}_1, \dots, \vec{u}_N$
 - For any $\vec{v} \in \mathbb{R}^N$, $U\vec{v}$ is the vector of coordinates of \vec{v} with respect to $\vec{u}_1, \dots, \vec{u}_N$

Summary

- Definition of a linear transformation
- A linear transformation is defined by what it does to a basis
- There is a one-to-one correspondence between
 - Linear transformations $T : \mathbb{R}^N \rightarrow \mathbb{R}^M$ and
 - $M \times N$ matrices
- Definition of an orthogonal matrix
- Orthogonal matrices correspond to change-of-basis transformations