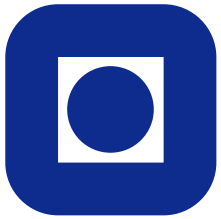

NTNU
Norwegian University of
Science and Technology

TMA4115 - Calculus 3

Lecture 13, Feb 27

Toke Meier Carlsen
Norwegian University of Science and Technology

Spring 2013

Review of last week's lecture

Last week we looked at

- *linear transformations*,
- the *standard matrix* of a linear transformation,
- *onto* linear transformations,
- *one-to-one* linear transformations,
- *applications* of linear models,
- *Maple* and *WolframAlpha*.

NTNU
Norwegian University of
Science and Technology

Today's lecture

We shall

- see how we can *add* and *multiply* matrices,
- define and study *invertible* matrices.

Matrix addition and scalar multiplication

- If A and B are $m \times n$ matrices, then the *sum* $A + B$ is the $m \times n$ matrix whose entries are the sums of the corresponding entries in A and B.
- The sum $A + B$ is defined only when A and B are the same size.
- If r is a scalar and A is an $m \times n$ matrix, then the *scalar multiple* rA is the $m \times n$ matrix whose entries are r times the corresponding entries in A.
- We will use $-A$ to denote the matrix $(-1)A$.

NTNU
Norwegian University of Science and Technology

Examples

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 4 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 2 & -6 \\ 0 & 7 & 1 \end{bmatrix}$, and $C = \begin{bmatrix} 2 & 1 \\ -4 & 3 \end{bmatrix}$.

Then $A + B = \begin{bmatrix} 4 & 4 & -3 \\ -1 & 11 & 1 \end{bmatrix}$ and $-2C = \begin{bmatrix} -4 & -2 \\ 8 & -6 \end{bmatrix}$.

$A + C$ is not defined because A and C have different sizes.

Properties of matrix addition and scalar multiplication

Theorem 1

Let A, B, and C be matrices of the same size, and let r and s be scalars. Then:

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + 0 = A$ where 0 is the matrix of the same size of A whose entries are all zero.
4. $r(A + B) = rA + rB$
5. $(r + s)A = rA + sA$
6. $r(sA) = (rs)A$

NTNU
Norwegian University of Science and Technology

Composition of linear transformations

Let $T : \mathbb{R}^p \to \mathbb{R}^n$ and $S : \mathbb{R}^n \to \mathbb{R}^m$ be linear transformations. Then $S \circ T$ is the transformation from $\mathbb{R}^p \to \mathbb{R}^m$ such that $S \circ T(\mathbf{x}) = S(T(\mathbf{x}))$ for $\mathbf{x} \in \mathbb{R}^p$.
If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^p$ and c is a scalar, then

- $S \circ T(\mathbf{u} + \mathbf{v}) = S(T(\mathbf{u} + \mathbf{v})) = S(T(\mathbf{u}) + T(\mathbf{v})) = S(T(\mathbf{u})) + S(T(\mathbf{v})) = S \circ T(\mathbf{u}) + S \circ T(\mathbf{v})$, and
- $S \circ T(c\mathbf{u}) = S(T(c\mathbf{u})) = S(cT(\mathbf{u})) = cS(T(\mathbf{u})) = cS \circ T(\mathbf{u})$.

So $S \circ T$ is linear. What is the standard matrix of $S \circ T$?

NTNU
Norwegian University of Science and Technology

The standard matrix of $S \circ T$

Let A be the standard matrix of S, let B be the standard matrix of T, and let $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p$ be the columns of B.

Then we have for $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$ in $\mathbb{R}^p$ that

$$\begin{aligned} S \circ T(\mathbf{x}) &= S(T(\mathbf{x})) = S(B\mathbf{x}) \\ &= A(B\mathbf{x}) = A(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_p\mathbf{b}_p) \\ &= x_1 A\mathbf{b}_1 + x_2 A\mathbf{b}_2 + \ldots x_p A\mathbf{b}_p. \end{aligned}$$

It follows that the standard matrix of $S \circ T$ is the $m \times p$ matrix $[A\mathbf{b}_1 \ A\mathbf{b}_2 \ \ldots \ A\mathbf{b}_p]$.

NTNU
Norwegian University of
Science and Technology

Matrix multiplication

Definition

If A is an $m \times n$ matrix, and if B is an $n \times p$ matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_p$, then the product AB is the $m \times p$ matrix whose columns are $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$.
That is,

$$AB = A[\mathbf{b}_1 \ \mathbf{b}_2 \ \ldots \ \mathbf{b}_p] = [A\mathbf{b}_1 \ A\mathbf{b}_2 \ \ldots \ A\mathbf{b}_p].$$

Each column of AB is a linear combination of the columns of A using weights from the corresponding column of B.

NTNU
Norwegian University of
Science and Technology

The standard matrix of the composition of 2 linear transformation

Let $T : \mathbb{R}^p \to \mathbb{R}^n$ and $S : \mathbb{R}^n \to \mathbb{R}^m$ be linear transformations. If A is the standard matrix of S, and B is the standard matrix of T, then AB is the standard matrix of $S \circ T$.

NTNU
Norwegian University of Science and Technology

Example

Let $A = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 & -3 \\ 2 & 4 & 0 \end{bmatrix}$. Let us compute AB.

Solution

$$A\mathbf{b}_1 = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

$$A\mathbf{b}_2 = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 11 \end{bmatrix}$$

$$A\mathbf{b}_3 = \begin{bmatrix} 2 & 1 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}$$

$$\text{So } AB = \begin{bmatrix} 6 & 6 & -6 \\ 4 & 11 & 3 \end{bmatrix}.$$

NTNU
Norwegian University of
Science and Technology

Row—column rule for computing AB

If a product AB is defined, then the entry in row i and column j of AB is the sum of the products of corresponding entries from row i of A and column j of B.
If A is an $m \times n$ matrix, B is an $n \times p$ matrix, a_{ik} denotes the (i,k)-entry in A, b_{kj} denotes the (k,j)-entry in B, and $(AB)_{ij}$ denotes the (i,j)-entry in A, then

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

NTNU
Norwegian University of
Science and Technology

Example

Let $A = \begin{bmatrix} 2 & -2 \\ 1 & -1 \\ 3 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} -2 & 1 \\ 4 & 0 \end{bmatrix}$. Let us compute AB.

Solution

$$\begin{bmatrix} 2 & -2 \\ 1 & -1 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 2\cdot(-2)+(-2)\cdot 4 & 2\cdot 1+(-2)\cdot 0 \\ 1\cdot(-2)+(-1)\cdot 4 & 1\cdot 1+(-1)\cdot 0 \\ 3\cdot(-2)+0\cdot 4 & 3\cdot 1+0\cdot 0 \end{bmatrix}$$

$$= \begin{bmatrix} -12 & 2 \\ -6 & 1 \\ -6 & 3 \end{bmatrix}$$

Properties of matrix multiplication

Theorem 2

Let r be a scalar, let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined. Then:

1. $A(BC) = (AB)C$
2. $A(B + C) = AB + AC$
3. $(B + C)A = BA + CA$
4. $r(AB) = (rA)B = A(rB)$
5. $I_m A = A I_n = A$

NTNU
Norwegian University of
Science and Technology

AB and BA might be different

- BA might not be defined, even if AB is defined (if, for example A is a 2×2 matrix and B is a 2×3 matrix, then AB is defined, but BA is not).
- Even when AB and BA are both defined, then AB and BA might not have the same size (if, for example A is a 3×2 matrix and B is a 2×3 matrix, then AB is a 3×3 matrix and BA is a 2×2 matrix).
- Even when AB and BA are both defined, and AB and BA have the same size, then AB and BA might be different.

NTNU
Norwegian University of
Science and Technology

Warnings

1. In general, $AB \neq BA$ (for instance, $\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} \neq \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$).
2. The cancellation laws do not hold for matrix multiplication. That is, if $AB = AC$, then it is not true in general that $B = C$ (for instance, $\begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -3 & -5 \\ 2 & 1 \end{bmatrix}$).
3. If a product AB is the zero matrix, you cannot conclude in general that either $A = 0$ or $B = 0$ (for instance, $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 2 & -6 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$).

NTNU
Norwegian University of
Science and Technology

Commuting matrices

- If $AB = BA$, then we say that A and B *commute* (*kommutere* in Norwegian).
- A necessary (but not sufficient) condition for A and B to commute is that there are square matrices of the same size.

NTNU
Norwegian University of
Science and Technology

Powers of a matrix

- If A is a square matrix and k is a positive integer, then A^k denotes the product of k copies of A. That is,

$$A^k = \underbrace{A \dots A}_{k \text{ times}}.$$

Example

In a certain region, 5% of a city's population moves to the surrounding suburbs each year, and 3% of the suburban population moves into the city.
We saw last week that if r_i denotes the population in the city in year $2000 + i$, s_i denotes the population in the suburbs in year $2000 + i$ and $\mathbf{x}_i = \begin{bmatrix} r_i \\ s_i \end{bmatrix}$, then $\mathbf{x}_{i+1} = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix} \mathbf{x}_i$.

So $\mathbf{x}_n = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}^n \mathbf{x}_0$.

NTNU
Norwegian University of
Science and Technology

The transpose of a matrix

Definition

Given an $m \times n$ matrix A, the *transpose* of A is the $n \times m$, denoted by A^T, whose columns are formed from the corresponding rows of A.

NTNU
Norwegian University of
Science and Technology

Examples

Let $A = \begin{bmatrix} 2 & 2 \\ -4 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 2 & -7 \\ -1 & -1 & 3 \end{bmatrix}$.

Then $A^T = \begin{bmatrix} 2 & -4 \\ 2 & 5 \end{bmatrix}$ and $B^T = \begin{bmatrix} 3 & -1 \\ 2 & -1 \\ -7 & 3 \end{bmatrix}$.

NTNU
Norwegian University of
Science and Technology

Properties of the transpose

Theorem 3

Let A and B denote matrices whose sizes are appropriate for the following sums and products. Then:

1. $(A^T)^T = A$.
2. $(A + B)^T = A^T + B^T$.
3. $(rA)^T = rA^T$ for any scalar r.
4. $(AB)^T = B^T A^T$.

NTNU
Norwegian University of
Science and Technology

Invertible transformations

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a transformation.

1. T is said to be *invertible* if there exists a transformation $S : \mathbb{R}^m \to \mathbb{R}^n$ such that $S \circ T(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$ and $T \circ S(\mathbf{y}) = \mathbf{y}$ for all $\mathbf{y}$ in $\mathbb{R}^m$.
2. T is invertible if and only if it is onto and one-to-one.
   Proof: Assume that T is invertible and let $S : \mathbb{R}^m \to \mathbb{R}^n$ be a transformation such that $S \circ T(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$ and $T \circ S(\mathbf{y}) = \mathbf{y}$ for all $\mathbf{y}$ in $\mathbb{R}^m$. If $\mathbf{y}$ is a vector in $\mathbb{R}^m$ and we let $\mathbf{x} = S(\mathbf{y})$, then $T(\mathbf{x}) = T(S(\mathbf{y})) = \mathbf{y}$. Thus, T is onto.

NTNU
Norwegian University of
Science and Technology

Invertible transformations (cont.)

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$ and $T(\mathbf{u}) = T(\mathbf{v})$, then $\mathbf{u} = S(T(\mathbf{u})) = S(T(\mathbf{v})) = \mathbf{v}$. Thus, T is one-to-one.
Assume on the other hand that T is onto and one-to-one. Then there is for each $\mathbf{y}$ in $\mathbb{R}^m$ a unique $\mathbf{x}$ in $\mathbb{R}^n$ such that $T(\mathbf{x}) = \mathbf{y}$. Let $S(\mathbf{y}) = \mathbf{x}$. Then S is a transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$ and $T(S(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{y}$ in $\mathbb{R}^m$. Let $\mathbf{x}$ be a vector in $\mathbb{R}^n$. Then $T(S(T(\mathbf{x}))) = T(\mathbf{x})$, and since T is one-to-one, it follows that $S(T(\mathbf{x})) = \mathbf{x}$. Thus, T is invertible.

NTNU
Norwegian University of
Science and Technology

Invertible transformations (cont.)

3. There can at most be one transformation $S : \mathbb{R}^m \to \mathbb{R}^n$ such that $S \circ T(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$ and $T \circ S(\mathbf{y}) = \mathbf{y}$ for all $\mathbf{y}$ in $\mathbb{R}^m$. If such an S exists, then it is called the *inverse* of T and is denoted by T^{-1}.
Proof: Assume that S_1 and S_2 are transformations from $\mathbb{R}^m$ to $\mathbb{R}^n$ such that $S_1 \circ T(\mathbf{x}) = S_2 \circ T(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$ and $T \circ S_1(\mathbf{y}) = T \circ S_2(\mathbf{y}) = \mathbf{y}$ for all $\mathbf{y}$ in $\mathbb{R}^m$. If y is a vector in $\mathbb{R}^m$, then $S_1(\mathbf{y}) = S_1(T(S_2(\mathbf{y}))) = S_2(\mathbf{y})$. This shows that $S_1 = S_2$.

NTNU
Norwegian University of
Science and Technology

Invertible linear transformations

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

1. A necessary (but not sufficient) condition for T to be onto is that $n \geq m$.
   Proof: Let A be the standard matrix of T. Then T is onto if and only if A has a pivot position in every row. If $n < m$, then A has fewer rows than columns, so A cannot have a pivot position in every row, and T is therefore not onto.

Invertible linear transformations (cont.)

2. A necessary (but not sufficient) condition for T to be one-to-one is that $n \leq m$.
   Proof: T is one-to-one if and only if A has a pivot position in every column. If $n > m$, then A has fewer columns than rows, so A cannot have a pivot position in every column, and T is therefore not one-to-one.
3. It follows that a necessary (but not sufficient) condition for T to be invertible is that $n = m$.

NTNU
Norwegian University of Science and Technology

Invertible linear transformations (cont.)

4. If T is invertible, then T^{-1} is a linear transformation.
Proof: Suppose that T is invertible. Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$ and let c be a scalar. Then

$$\begin{aligned} T^{-1}(\mathbf{u}+\mathbf{v}) &= T^{-1}(T(T^{-1}(\mathbf{u}))+T(T^{-1}(\mathbf{v}))) \\ &= T^{-1}(T(T^{-1}(\mathbf{u})+T^{-1}(\mathbf{v}))) \\ &= T^{-1}(\mathbf{u})+T^{-1}(\mathbf{v}) \end{aligned}$$

and

$$\begin{aligned} T^{-1}(c\mathbf{u}) = T^{-1}(cT(T^{-1}(\mathbf{u}))) &= T^{-1}(T(cT^{-1}(\mathbf{u}))) \\ &= cT^{-1}(\mathbf{u}), \end{aligned}$$

NTNU
Norwegian University of Science and Technology

Invertible linear transformations (cont.)

so T^{-1} is linear.

5. If $n = m$ and A is the standard matrix of T, then T is invertible if and only if there exists a matrix C such that $AC = CA = I_n$.
   Proof: Suppose that T is invertible, and let C be the standard matrix of T^{-1}. If $\mathbf{x}$ is a vector in $\mathbb{R}^n$, then
   $$AC\mathbf{x} = T(C\mathbf{x}) = T(T^{-1}\mathbf{x}) = \mathbf{x}$$
   and
   $$CA\mathbf{x} = T^{-1}(A\mathbf{x}) = T^{-1}(T\mathbf{x}) = \mathbf{x}.$$
   It follows that $AC = CA = I_n$.

NTNU
Norwegian University of Science and Technology

Invertible linear transformations (cont.)

Suppose on the other hand that C is a matrix such that $AC = CA = I_n$. Let $S(\mathbf{x}) = C\mathbf{x}$ for each $\mathbf{x}$ in $\mathbb{R}^n$. Then S is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$, and

$$S \circ T(\mathbf{x}) = C(T(\mathbf{x})) = C(A\mathbf{x}) = (CA)\mathbf{x} = I_n\mathbf{x} = \mathbf{x}$$

and

$$T \circ S(\mathbf{x}) = T(C\mathbf{x}) = A(C\mathbf{x}) = (AC)\mathbf{x} = I_n\mathbf{x} = \mathbf{x}.$$

Thus, T is invertible.

NTNU
Norwegian University of
Science and Technology

The inverse of a matrix

- An $n \times n$ matrix A is said to be *invertible* if there exists an $n \times n$ matrix C such that $AC = CA = I_n$.
- There can at most be one $n \times n$ matrix C such that $AC = CA = I_n$. If such a C exists, then it is called the *inverse* of A and is denoted by A^{-1}.
- If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation and A is the standard matrix of T, then T is invertible if and only if A is.
- If T and A are invertible, then A^{-1} is the standard matrix of T^{-1}.
- A matrix that is not invertible is sometimes called a *singular* matrix, and an invertible matrix is called a *nonsingular* matrix.

NTNU
Norwegian University of
Science and Technology

Tomorrow's lecture

Tomorrow we shall

- further study *invertible* matrices,
- look at *the invertible matrix theorem*.

Section 2.3 in "Linear Algebras and Its Applications" (pages 111–116).

NTNU
Norwegian University of
Science and Technology