Slide 1

Linear transformations

- Definition of linear transformations.
- Linear transformations and matrices.
- one-to-one, onto.

Slide 2

Linear transformations

Definition 1 *A function $T : \mathbb{R}^n \to R \subset \mathbb{R}^m$ is called a linear transformation if*

$$\begin{aligned} T(\mathbf{x}_1 + \mathbf{x}_2) &= T(\mathbf{x}_1) + T(x_2), && \forall\, \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n, \\ T(a\mathbf{x}) &= aT(\mathbf{x}), && \forall\, \mathbf{x} \in \mathbb{R}^n,\, \forall a \in \mathbb{R}. \end{aligned}$$

The symbol $\subset$ means "subset of" and $\forall$ means "for all."

- $\mathbb{R}^n$ is called the domain of T.
- $\mathbb{R}^m$ is called the codomain of T.
- $R = T(\mathbb{R}^n) = \{\mathbf{v} \in \mathbb{R}^m : \text{exists } \mathbf{x} \in \mathbb{R}^n, \text{ such that } T(\mathbf{x}) = \mathbf{v}\}$ is called the range of T.

Slide 3

Example: $T : \mathbb{R}^2 \to \mathbb{R}^3$,

$$T\mathbf{x} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_1 + x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} x_2,$$

so $T(\mathbb{R}^2) = \mathbb{R}^2$, that is, the range of T is $\mathbb{R}^2$, a subset of the codomain, R^3.

Theorem 1 *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, with $1 \leq m < \infty$, and $1 \leq n < \infty$. Then, there exists an $m \times n$ matrix A such that*

$$T(\mathbf{x}) = A\mathbf{x}.$$

Slide 4

- Recall: $T(a_1\mathbf{b}_1 + \cdots + a_n\mathbf{b}_n) = a_1 T(\mathbf{b}_1) + \cdots + a_n T(\mathbf{b}_n)$.
- Introduce the n-vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix},$$

then

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n.$$

Slide 5

Proof:

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n), \\ &= x_1 T(\mathbf{e}_1) + \cdots + x_n T(\mathbf{e}_n), \\ &= [T(\mathbf{e}_1), \cdots, T(\mathbf{e}_n)]\,\mathbf{x}, \\ &= A\mathbf{x}, \end{aligned}$$

where

$$A = [\mathbf{a}_1, \cdots, \mathbf{a}_n] = [T(\mathbf{e}_1), \cdots, T(\mathbf{e}_n)]\,,$$

□

Therefore, the column vectors $\mathbf{a}_i$ in A are the images of the $\mathbf{e}_i$ by T, that is, $\mathbf{a}_i = T(\mathbf{e}_i)$, for $i = 1 \cdots n$.

Slide 6

One-to-one, and onto

Definition 2 *$T : \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if the following holds: If $\mathbf{x}_1 \neq \mathbf{x}_2$, then $T(\mathbf{x}_1) \neq T(\mathbf{x}_2)$.*

Definition 3 *$T : \mathbb{R}^n \to \mathbb{R}^m$ is onto if the following holds: For all $\mathbf{b} \in \mathbb{R}^m$ there exists $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{b} = T(\mathbf{x})$.*

Slide 7

Theorem 2 *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.*
T is one-to-one $\Leftrightarrow$ $T(\mathbf{x}) = 0$ has only the trivial solution $\mathbf{x} = \mathbf{0}$.

For the proof, first recall: T linear implies $T(\mathbf{0}) = \mathbf{0}$.

Proof:

($\Rightarrow$) If $\mathbf{x} \neq \mathbf{0}$ and T one-to-one, then $T(\mathbf{x}) \neq T(\mathbf{0}) = \mathbf{0}$, hence $T(\mathbf{x}) \neq \mathbf{0}$ for nonzero $\mathbf{x}$. In other words, $T(\mathbf{x}) = \mathbf{0}$ only has the trivial solution.

($\Leftarrow$) Take $\mathbf{u} - \mathbf{v} \neq \mathbf{0}$. Because $T(\mathbf{x}) = \mathbf{0}$ has only the trivial solution, then $\mathbf{0} \neq T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$. So, we have shown that $\mathbf{u} \neq \mathbf{v}$ implies that $T(\mathbf{u}) \neq T(\mathbf{v})$, and this says that T is one-to-one. $\square$

Slide 8

Theorem 3 *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation with matrix $A = [\mathbf{a}_1, \cdots, \mathbf{a}_n]$. Then the following assertions hold:*
T onto $\Leftrightarrow$ $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ span $\mathbb{R}^m$.
T one-to-one $\Leftrightarrow$ $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ are l.i..

Proof:

T onto means that for all $\mathbf{b} \in \mathbb{R}^m$ there exists $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{b} = T(\mathbf{x}) = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$. This says that $\mathbf{b} \in \text{Span}\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ for all $\mathbf{b} \in \mathbb{R}^m$, so $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ span $\mathbb{R}^m$.

$T(\mathbf{x}) = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$, so $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$ has only the solution $x_1 = \cdots = x_n = 0$. This says that $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ are l.i.. $\square$

Slide 9

Matrix Operations

- Multiplication by a number,
- Addition.
- Matrix multiplication.
  - Properties (Non-commutative).
- Transpose of a matrix.

Slide 10

The origin of Matrix operations

Linear transformations $T : I\!R^n \to I\!R^m$ are functions that happen to be linear.

Because they are functions, the usual operations on functions can be introduced, namely,

- multiplication by a number,
- a pair of appropriate functions can be added,
- a pair of appropriate functions can be *composed*.

The association

$$T, \text{ function } \longleftrightarrow A, \text{ matrix}$$

translates operations on functions into operations on matrices.

Slide 11

Linear combination of matrices

Given the $m \times n$ matrices $A = [\mathbf{a}_1, \cdots, \mathbf{a}_n]$ and $B = [bb_1, \cdots, \mathbf{b}_n]$, and the numbers c, d, the linear combination is defined as

$$cA + dB = [c\mathbf{a}_1 + d\mathbf{b}_1, \cdots, c\mathbf{a}_n + d\mathbf{b}_n].$$

That is, in components,

$$cA + dB = \begin{bmatrix} (c\,a_{11} + d\,b_{11}) & \cdots & (c\,a_{1n} + d\,b_{1n}) \\ \vdots & & \vdots \\ (c\,a_{m1} + d\,b_{m1}) & \cdots & (c\,a_{mn} + d\,b_{mn}) \end{bmatrix}.$$

Slide 12

Multiplication of Matrices

Composition of $S : \mathbb{R}^n \to \mathbb{R}^m$, with $T : \mathbb{R}^m \to \mathbb{R}^\ell$

$$\mathbb{R}^n \xrightarrow{S} \mathbb{R}^m \xrightarrow{T} \mathbb{R}^\ell,$$

$$T \circ S(\mathbf{x}) = T(S(\mathbf{x})),$$

implies the multiplication of the matrix A (associated to T) with the matrix B (associated to S),

$$\mathbf{x} \in \mathbb{R}^n \xrightarrow{B} B\mathbf{x} \in \mathbb{R}^m \xrightarrow{A} A(B\mathbf{x}) \in \mathbb{R}^\ell.$$

$$T(S(\mathbf{x})) \longleftrightarrow A(B\mathbf{x}).$$

Slide 13

Multiplication of Matrices

$$\underset{\ell \times n}{AB} \quad \text{denotes multiplication of} \quad \underset{\ell \times m}{A} \quad \text{with} \quad \underset{m \times n}{B}.$$

And the product is,

$$AB = [A\mathbf{b}_1, \cdots, A\mathbf{b}_n].$$

Slide 14

Matrix operations: Properties

- $A(BC) = (AB)C$,
- $A(B + C) = AB + AC$,
- $(B + C)A = BA + BC$,
- $a(AB) = (aA)B = A(aB)$,
- $I_m A = A = AI_n$.

Notice: For $n \times n$ squared matrices A, B, one has in general that

$$AB \neq BA,$$

that is, the product is not commutative.

Slide 15

Transpose of a matrix

Definition 4 *The transpose of an $m \times n$ matrix A is an $n \times m$ matrix called A^T, whose columns are the rows of A.*

Example:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}, \quad \left(A^T\right)^T = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = A.$$

Slide 16

Properties of the transpose

- $(A^T)^T = A$,
- $(A + B)^T = A^T + B^T$,
- $(aA)^T = aA^T$,
- $(AB)^T = B^T A^T$.

Slide 17

Inverse of a matrix

- Inverse of a matrix.
- Computation of the matrix.

Slide 18

Inverse of a matrix

Definition 5 *An $n \times n$ (squared) matrix A is said to be invertible $\Leftrightarrow$ there exists an $n \times n$ matrix, denoted as A^{-1}, satisfying*

$$(A^{-1})A = I_n, \quad A(A^{-1}) = I_n.$$

Slide 19

Inverse matrix and systems of linear equations.

Theorem 4 *An* $n \times n$ *matrix* A *is invertible* $\Leftrightarrow$ *for all* $\mathbf{b} \in \mathbb{R}^n$ *there exists a unique* $\mathbf{x} \in \mathbb{R}^n$ *solution of* $A\mathbf{x} = \mathbf{b}$.

Proof: ($\Rightarrow$) Define $\mathbf{x} = A^{-1}\mathbf{b}$. Then, $A\mathbf{x} = A(A^{-1})\mathbf{b} = \mathbf{b}$.

If $\mathbf{x}_1$ and $\mathbf{x}_2$ satisfy $A\mathbf{x}_1 = \mathbf{b}$, and $A\mathbf{x}_2 = \mathbf{b}$, then $A(\mathbf{x}_1 - \mathbf{x}_2) = 0$, so $(A^{-1})A(\mathbf{x}_1 - \mathbf{x}_2) = 0$, and then $\mathbf{x}_1 = \mathbf{x}_2$. The solution is unique.

($\Leftarrow$)(sketch)

For all $\mathbf{b} \in \mathbb{R}^n$ exists a unique $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$. This defines a function $S : \mathbb{R}^n \to \mathbb{R}^n$, $S(\mathbf{b}) = \mathbf{x}$.

Assume that this function is linear. (We do not show this here.) Then it has associated a matrix C, such that $C\mathbf{b} = \mathbf{x}$.

Then, $AC\mathbf{b} = A\mathbf{x} = \mathbf{b}$, so $AC = I$.

Finally, $A\mathbf{x} = \mathbf{b}$ implies that $CA\mathbf{x} = C\mathbf{b} = \mathbf{x}$, so $CA = I$. $\square$

Slide 20

Computation of the inverse matrix

- The inverse matrix can be computed with ERO (elementary row operations).
- Start with and augmented matrix $[A|I]$.
- Perform ERO until $[A|I] \to [I|B]$, for some matrix B.
- Then, $B = A^{-1}$. (Otherwise, A is not invertible.)

Slide 21

Claim: Each ERO can be performed by multiplication by an appropriate matrix.

If the ERO given by the matrices $E_1, \cdots, E_k$ transform A into the identity matrix I, then the following equations holds,

$$E_k \cdots E_1 A = I,$$

therefore $E_k \cdots E_1 = A^{-1}$.

The computation of A^{-1} can be done as follows:

$$[A|I] \to [E_1 A|E_1] \to \cdots \to [E_k \cdots E_1 A|E_k \cdots E_1].$$

But the ERO are chosen such that $E_k \cdots E_1 A = I$, so

$$[E_k \cdots E_1 A|E_k \cdots E_1] = [I|E_k \cdots E_1]$$

and the matrix $E_k \cdots E_1$ is precisely equal to A^{-1}.