Linear transformations

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

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

$$
T(x_1 + x_2) = T(x_1) + T(x_2), \quad \forall x_1, x_2 \in \mathbb{R}^n,
$$

$$
T(ax) = aT(x), \quad \forall x \in \mathbb{R}^n, \forall a \in \mathbb{R}.
$$

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

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

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

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

**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 

$$T(x) = Ax.$$ 

- Recall: $T(a_1 b_1 + \cdots + a_n b_n) = a_1 T(b_1) + \cdots + a_n T(b_n)$.

- Introduce the $n$-vectors

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix},$$

then

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 e_1 + \cdots + x_n e_n.$$
Proof:

\[
T(x) = T(x_1e_1 + \cdots + x_ne_n), \\
= x_1T(e_1) + \cdots + x_nT(e_n), \\
= [T(e_1), \ldots, T(e_n)]x, \\
= Ax,
\]

where

\[
A = [a_1, \ldots, a_n] = [T(e_1), \ldots, T(e_n)],
\]

Therefore, the column vectors \(a_i\) in \(A\) are the images of the \(e_i\) by \(T\), that is, \(a_i = T(e_i)\), for \(i = 1 \cdots n\).

---

_Definition 2_ \(T : \mathbb{R}^n \to \mathbb{R}^m\) is one-to-one if the following holds:

If \(x_1 \neq x_2\), then \(T(x_1) \neq T(x_2)\).

_Definition 3_ \(T : \mathbb{R}^n \to \mathbb{R}^m\) is onto if the following holds:

For all \(b \in \mathbb{R}^m\) there exists \(x \in \mathbb{R}^n\) such that \(b = T(x)\).
Theorem 2 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.  
$T$ is one-to-one $\iff T(x) = 0$ has only the trivial solution $x = 0$.

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

Proof:

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

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

---

Theorem 3 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation with matrix $A = [a_1, \ldots, a_n]$. Then the following assertions hold:

$T$ onto $\iff \{a_1, \ldots, a_n\}$ span $\mathbb{R}^m$.

$T$ one-to-one $\iff \{a_1, \ldots, a_n\}$ are l.i.

Proof:

$T$ onto means that for all $b \in \mathbb{R}^m$ there exists $x \in \mathbb{R}^n$ such that $b = T(x) = x_1a_1 + \cdots + x_na_n$. This says that $b \in \text{Span}\{a_1, \ldots, a_n\}$ for all $b \in \mathbb{R}^m$, so $\{a_1, \ldots, a_n\}$ span $\mathbb{R}^m$.

$T(x) = 0$ has only the trivial solution $x = 0$, so $x_1a_1 + \cdots + x_na_n = 0$ has only the solution $x_1 = \cdots = x_n = 0$. This says that $\{a_1, \ldots, a_n\}$ are l.i.
Matrix Operations

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

The origin of Matrix operations

Linear transformations $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ 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

$T$, function $\leftrightarrow A$, matrix

translates operations on functions into operations on matrices.
**Linear combination of matrices**

Given the $m \times n$ matrices $A = [a_1, \cdots, a_n]$ and $B = [b_1, \cdots, b_n]$, and the numbers $c, d$, the linear combination is defined as

$$cA + dB = [ca_1 + db_1, \cdots, ca_n + db_n].$$

That is, in components,

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

---

**Multiplication of Matrices**

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

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

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

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

$$T(S(x)) \longleftrightarrow A(Bx).$$
Multiplication of Matrices

$AB$ denotes multiplication of $A$ with $B$.

And the product is,

$$AB = [Ab_1, \ldots, Ab_n].$$

Matrix operations: Properties

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

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

$$AB \neq BA,$$

that is, the product is not commutative.
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:

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

Properties of the transpose

- \((A^T)^T = A\),
- \((A + B)^T = A^T + B^T\),
- \((aA)^T = aA^T\),
- \((AB)^T = B^T A^T\).
Inverse of a matrix

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

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

$$(A^{-1})A = I_n, \quad A(A^{-1}) = I_n.$$
Inverse matrix and systems of linear equations.

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

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

If \( x_1 \) and \( x_2 \) satisfy \( Ax_1 = b \), and \( Ax_2 = b \), then \( A(x_1 - x_2) = 0 \), so \( (A^{-1})A(x_1 - x_2) = 0 \), and then \( x_1 = x_2 \). The solution is unique.

(\( \Leftarrow \)) (sketch)

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

Assume that this function is linear. (We do not show this here.)

Then it has associated a matrix \( C \), such that \(Cb = x\).

Then, \( ACb = Ax = b \), so \( AC = I \).

Finally, \( Ax = b \) implies that \( CAx = Cb = x \), so \( CA = I \).

---

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

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

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

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 $A^{-1}$. 