Introduction to linear mappings [1.8]

- A transformation or function or mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule which assigns to every $x$ in $\mathbb{R}^n$ a vector $T(x)$ in $\mathbb{R}^m$.

- $\mathbb{R}^n$ is called the domain space of $T$ and $\mathbb{R}^m$ the image space or co-domain of $T$.

- Notation:

$$ T : \mathbb{R}^n \longrightarrow \mathbb{R}^m $$

- $T(x)$ is the image of $x$ under $T$.

Example: Take the mapping from $\mathbb{R}^2$ to $\mathbb{R}^3$:

$$ T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3 $$

$$ x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longrightarrow T(x) = \begin{pmatrix} x_1 + x_2 \\ x_1x_2 \\ x_1^2 + x_2^2 \end{pmatrix} $$

Example: Another mapping from $\mathbb{R}^2$ to $\mathbb{R}^3$:

$$ T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3 $$

$$ x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longrightarrow T(x) = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \\ x_1 + 5x_2 \end{pmatrix} $$

What is the main difference between these 2 examples?

Definition A mapping $T$ is linear if:

(i) $T(u + v) = T(u) + T(v)$ for $u, v$ in the domain of $T$

(ii) $T(\alpha u) = \alpha T(u)$ for all $\alpha \in \mathbb{R}$, all $u$ in the domain of $T$

- The mapping of the second example given above is linear - but not for the first one.

- If a mapping is linear then $T(0) = 0$. (Why?)

Observation: A mapping is linear if and only if

$$ T(\alpha u + \beta v) = \alpha T(u) + \beta T(v) $$

for all scalars $\alpha, \beta$ and all $u, v$ in the domain of $T$.

Prove this.
Given an \( m \times n \) matrix \( A \), consider the special mapping:

\[
T : \mathbb{R}^n \rightarrow \mathbb{R}^m \\
x \mapsto y = Ax
\]

\( \text{Domain} == ??; \text{Image space} == ?? \)

From what we saw earlier ['Properties of the matrix-vector product'] such mappings are linear.

As it turns out:

If \( T \) is linear, there exists a matrix \( A \) such that \( T(x) = Ax \) for all \( x \) in \( \mathbb{R}^n \).

In plain English: ‘A linear mapping can be represented by a matvec’

\( A \) is the representation of \( T \).

Let \( A \) be a square matrix. Is the mapping \( x \rightarrow x + Ax \) linear? If so find the matrix associated with it.

Same questions for the mapping \( x \rightarrow Ax + \alpha x \) - where \( \alpha \) is a scalar.

Express the following mapping from \( \mathbb{R}^3 \) to \( \mathbb{R}^2 \) in matrix/vector form:

\[
\begin{align*}
y_1 &= 2x_1 - x_2 + 1 \\
y_2 &= x_2 - x_3 - 2
\end{align*}
\]

Is this a linear mapping?

Read Section 1.9 and explore the notions of onto mappings ('surjective') and one-to-one mappings ('injective') in the text. You must at least know the definitions.

A mapping is onto if and only if ....

A mapping is one-to-one if and only if ....

\[ \text{Matrix operations} \]

If \( A \) is an \( m \times n \) matrix (\( m \) rows and \( n \) columns) then the scalar entry in the \( i \)th row and \( j \)th column of \( A \) is denoted by \( a_{ij} \) and is called the \((i, j)\)-entry of \( A \).

\[
\begin{bmatrix}
a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\
a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\
a_{m1} & \cdots & a_{mj} & \cdots & a_{mn}
\end{bmatrix}
= A
\]

Column \( j \)

Row \( i \)
The number $a_{ij}$ is the $i$th entry (from the top) of the $j$th column.

Each column of $A$ is a list of $m$ real numbers, which identifies a vector in $\mathbb{R}^m$ called a column vector.

The columns are denoted by $a_1, \ldots, a_n$, and the matrix $A$ is written as $A = [a_1, a_2, \ldots, a_n]$.

The diagonal entries in an $m \times n$ matrix $A$ are $a_{11}, a_{22}, a_{33}, \ldots$, and they form the main diagonal of $A$.

A diagonal matrix is a matrix whose nondiagonal entries are zero.

An important example is the $n \times n$ identity matrix, $I_n$ (each diagonal entry equals one) - Example:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Another important matrix is the zero matrix (all entries are 0). It is denoted by $O$.

Equality of two matrices: Two matrices $A$ and $B$ are equal if they have the same size (they are both $m \times n$) and if their entries are all the same.

$$a_{ij} = b_{ij} \text{ for all } i = 1, \ldots, m, \ j = 1, \ldots, n$$

Sum of two matrices: If $A$ and $B$ are $m \times n$ matrices, then their sum $A + B$ is the $m \times n$ matrix whose entries are the sums of the corresponding entries in $A$ and $B$.

If we call $C$ this sum we can write:

$$c_{ij} = a_{ij} + b_{ij} \text{ for all } i = 1, \ldots, m, \ j = 1, \ldots, n$$

Scalar multiple of a matrix If $r$ is a scalar and $A$ is a matrix, then the scalar multiple $rA$ is the matrix whose entries are $r$ times the corresponding entries in $A$.

$$(\alpha A)_{ij} = \alpha a_{ij} \text{ for all } i = 1, \ldots, m, \ j = 1, \ldots, n$$

Theorem Let $A$, $B$, and $C$ be matrices of the same size, and let $\alpha$ and $\beta$ be scalars. Then:

- $A + B = B + A$
- $(A + B) + C = A + (B + C)$
- $A + 0 = A$
- $\alpha(A + B) = \alpha A + \alpha B$
- $(\alpha + \beta)A = \alpha A + \beta A$
- $\alpha(\beta A) = (\alpha \beta)A$

Prove all of the above equalities
Matrix Multiplication

- When a matrix $B$ multiplies a vector $x$, it transforms $x$ into the vector $Bx$.
- If this vector is then multiplied in turn by a matrix $A$, the resulting vector is $A(Bx)$.
- Thus $A(Bx)$ is produced from $x$ by a composition of mappings—the linear transformations induced by $B$ and $A$.
- Note: $x \rightarrow yA(Bx)$ is a linear mapping (prove this).

Goal: to represent this composite mapping as a multiplication by a single matrix, call it $C$ for now, so that

$$A(Bx) = Cx$$

Assume $A$ is $m \times n$, $B$ is $n \times p$, and $x$ is in $\mathbb{R}^p$.

- Denote the columns of $B$ by $b_1, \cdots, b_p$, and the entries in $x$ by $x_1, \cdots, x_p$. Then:

$$Bx = x_1 b_1 + \cdots + x_p b_p$$

Definition: If $A$ is an $m \times n$ matrix, and if $B$ is an $n \times p$ matrix with columns $b_1, \cdots, b_p$, then the product $AB$ is the matrix whose $p$ columns are $Ab_1, \cdots, Ab_p$. That is:

$$AB = A[b_1, b_2, \cdots, b_p] = [Ab_1, Ab_2, \cdots, Ab_p]$$

Important to remember that:

Multiplication of matrices corresponds to composition of linear transformations.

Operation count: How many operations are required to perform product $AB$?

Denoted by $AB$
Compute $AB$ when 
\[
A = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 2 & -1 \\ 1 & 3 & -2 \end{bmatrix}
\]

Compute $AB$ when 
\[
A = \begin{bmatrix} 2 & -1 & 2 & 0 \\ 1 & -2 & 1 & 0 \\ 3 & -2 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & -1 & -2 \\ 0 & -2 & 2 \\ 2 & 1 & -2 \\ -1 & 3 & 2 \end{bmatrix}
\]

Can you compute $AB$ when 
\[
A = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 2 \\ 1 & 3 \\ -1 & 4 \end{bmatrix}
\]

**Row-wise matrix product**

- Recall what we did with matrix-vector product to compute a single entry of the vector $Ax$.
- Can we do the same thing here? i.e., How can we compute the entry $c_{ij}$ of the product $AB$ without computing entire columns?
- Do this to compute entry $(2, 2)$ in the first example above.
- Operation counts: Is more or less expensive to perform the matrix-vector product row-wise instead of column-wise?

**Properties of matrix multiplication**

**Theorem** Let $A$ be an $m \times n$ matrix, and let $B$ and $C$ have sizes for which the indicated sums and products are defined.

- $A(BC) = (AB)C$ (associative law of multiplication)
- $A(B + C) = AB + AC$ (left distributive law)
- $(B + C)A = BA + CA$ (right distributive law)
- $\alpha(AB) = (\alpha A)B = A(\alpha B)$ for any scalar $\alpha$
- $I_mA = AI_n = A$ (product with identity)

- If $AB = AC$ then $B = C$ (‘simplification’) : True-False?
- If $AB = 0$ then either $A = 0$ or $B = 0$ : True or False?
- $AB = BA$ : True or false??

**Square matrices. Matrix powers**

- Important particular case when $n = m$ - so matrix is $n \times n$
- In this case if $x$ is in $\mathbb{R}^n$ then $y = Ax$ is also in $\mathbb{R}^n$
- $AA$ is also a square $n \times n$ matrix and will be denoted by $A^2$
- More generally, the matrix $A^k$ is the matrix which is the product of $k$ copies of $A$:
\[
A^1 = A; \quad A^2 = AA; \quad \cdots \quad A^k = A \cdots A
\]

- Also true when $k$ or $l$ is zero.

$A_l \times A_k = A_{l+k}$

- For consistency define $A^0$ to be the identity: $A^0 = I_n$.

$A^l \times A^k = A^{l+k}$

- Also true when $k$ or $l$ is zero.
**Transpose of a matrix**

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

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

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- $(\alpha A)^T = \alpha A^T$ for any scalar $\alpha$
- $(AB)^T = B^T A^T$

**Matrix operations: Matrix-vector product (review)**

- $x, y$ are vectors; $y$ is the result of $A \times x$.
- $a_1, a_2, \ldots, a_n$ are the columns of $A$

**Example:**

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \quad x = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix} \quad y =?$$

**Result:**

$$y = -2 \times \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \times \begin{bmatrix} 2 \\ -1 \end{bmatrix} - 3 \times \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -10 \end{bmatrix}$$
Matrix operations: Matrix-Matrix product

When $A$ is $m \times n$, $B$ is $n \times p$, the product $AB$ of the matrices $A$ and $B$ is the $m \times p$ matrix defined as

$$AB = [Ab_1, Ab_2, \ldots, Ab_p]$$

Each $Ab_j$ is a matrix-vector product: the product of $A$ by the $j$-th column of $B$. Matrix $AB$ has dimension $m \times p$.

Can use what we know on matvecs to perform the product.

1. Column form – In words: “The $j$-th column of $AB$ is a linear combination of the columns of $A$, with weights $b_{1j}, b_{2j}, \ldots, b_{nj}$” (entries of $j$-th col. of $B$).

2. If we call $C$ the matrix $C = AB$ what is $c_{ij}$? From above:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ik}b_{kj} + \ldots + a_{in}b_{nj}$$

Fix $j$ and run $i \rightarrow$ column-wise form just seen.

3. Fix $i$ and run $j \rightarrow$ row-wise form.

Example: Get second row of $AB$ in previous example.

$$c_{2j} = a_{21}b_{1j} + a_{22}b_{2j} + a_{23}b_{3j}, \quad j = 1, 2$$

- Can be read as: $c_2 = a_{21}b_1 + a_{22}b_2 + a_{23}b_3$, or in words: row 2 of $C = a_{21}$ (row 1 of $B$) + $a_{22}$ (row 2 of $B$) + $a_{23}$ (row 3 of $B$) = $0$ (row 1 of $B$) + (-1) (row 2 of $B$) + (3) (row 3 of $B$) = $[-10 \quad 8]$