MA110:
Lecture 09

Saurav Bhaumik
Department of Mathematics
IIT Bombay

Spring 2025

Recall: Linear Transformations

Just as we can define a continuous function from a subset of $\mathbb{R}^n$ to $\mathbb{R}^m$, we now define a 'linear' function from a subspace of $\mathbb{R}^{n\times 1}$ to $\mathbb{R}^{m\times 1}$.

Let V be a subspace of $\mathbb{R}^{n\times 1}$, and let W be a subspace of $\mathbb{R}^{m\times 1}$. A **linear transformation** or a **linear map** from V to W is a function $T : V \to W$ which 'preserves' the operations of addition and scalar multiplication, that is, for all $\mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{R}$,

$$T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}) \quad \text{and} \quad T(\alpha\,\mathbf{x}) = \alpha\, T(\mathbf{x}).$$

It follows that if $T : V \to W$ is linear, then $T(\mathbf{0}) = \mathbf{0}$, and T 'preserves' linear combinations of vectors in V, that is,

$$T(\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k) = \alpha_1 T(\mathbf{x}_1) + \cdots + \alpha_k T(\mathbf{x}_k)$$

for all $\mathbf{x}_1, \ldots, \mathbf{x}_k \in V$ and $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$.

Model Example

Let $V := \mathbb{R}^{n\times 1}$, $W := \mathbb{R}^{m\times 1}$ and $\mathbf{A}$ be an $m\times n$ matrix, that is, $\mathbf{A} \in \mathbb{R}^{m\times n}$. Define $T_{\mathbf{A}} : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ by

$$T_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\,\mathbf{x} \quad \text{for } \mathbf{x} \in V.$$

The properties of matrix multiplication show that $T_{\mathbf{A}}$ is linear.

Conversely, suppose $T : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ is linear. We show that $T = T_{\mathbf{A}}$ for some $\mathbf{A} \in \mathbb{R}^{m\times n}$. Let $\mathbf{x} := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{n\times 1}$. Then $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the basic column vectors in $\mathbb{R}^{n\times 1}$. Since T is linear, we obtain

$$T(\mathbf{x}) = x_1 T(\mathbf{e}_1) + \cdots + x_n T(\mathbf{e}_n).$$

Define $\mathbf{c}_k := T(\mathbf{e}_k)$ for $k = 1, \ldots, n$, and $\mathbf{A} := \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{bmatrix}$. Then $T(\mathbf{x}) = x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n = \mathbf{A}\,\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^{n\times 1}$. Thus $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $T = T_{\mathbf{A}}$. (Note: kth column of $\mathbf{A}$ is $T(\mathbf{e}_k)$.)

Thus every linear transformation $T : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ is given by

$$T\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) := \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} \quad \text{for} \quad \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n\times 1},$$

where $a_{11}, \ldots, a_{1n}, \ldots, a_{m1}, \ldots, a_{mn} \in \mathbb{R}$.
Similarly, one can define a linear map $T : \mathbb{R}^{1\times n} \to \mathbb{R}^{1\times m}$, and find that for $\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \in \mathbb{R}^{1\times n}$,

$$T\left(\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}\right) := \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n & \cdots & a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix}.$$

Remark: Let D be an open subset of $\mathbb{R}^{1\times 2}$, $[x_0, y_0] \in D$, and let a function $f : D \to \mathbb{R}$ be differentiable at $[x_0, y_0]$. Then the total derivative of f at $[x_0, y_0]$ is a linear map (which depends on f) given by $T([x, y]) = \alpha x + \beta y$ for $[x, y] \in \mathbb{R}^{1\times 2}$, where $\alpha := f_x(x_0, y_0)$ and $\beta := f_y(x_0, y_0)$.

Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ and $\alpha, \beta \in \mathbb{R}$. Then $\alpha\mathbf{A} + \beta\mathbf{B} \in \mathbb{R}^{m\times n}$ and

$$T_{\alpha\mathbf{A}+\beta\mathbf{B}}(\mathbf{x}) = (\alpha\mathbf{A} + \beta\mathbf{B})\mathbf{x} = \alpha T_{\mathbf{A}}(\mathbf{x}) + \beta T_{\mathbf{B}}(\mathbf{x})$$

for $\mathbf{x} \in \mathbb{R}^{n\times 1}$. We write this as follows:

$$T_{\alpha\mathbf{A}+\beta\mathbf{B}} = \alpha T_{\mathbf{A}} + \beta T_{\mathbf{B}}.$$

Next, let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$. Then $\mathbf{AB} \in \mathbb{R}^{m\times p}$, and

$$T_{\mathbf{AB}}(\mathbf{x}) = (\mathbf{AB})\mathbf{x} = \mathbf{A}(\mathbf{Bx}) = T_{\mathbf{A}}(\mathbf{Bx}) = T_{\mathbf{A}}(T_{\mathbf{B}}(\mathbf{x})) = T_{\mathbf{A}} \circ T_{\mathbf{B}}(\mathbf{x})$$

for $\mathbf{x} \in \mathbb{R}^{p\times 1}$ by the associativity of matrix multiplication. Thus

$$T_{\mathbf{AB}} = T_{\mathbf{A}} \circ T_{\mathbf{B}}.$$

This says that the linear map associated with the product $\mathbf{AB}$ of matrices $\mathbf{A}$ and $\mathbf{B}$ is the composition of the linear maps associated with $\mathbf{A}$ and associated with $\mathbf{B}$ in the same order. This partially justifies the definition of matrix multiplication.

Examples

Let $\mathbf{A} \in \mathbb{R}^{2\times 2}$. Then $T_{\mathbf{A}} : \mathbb{R}^{2\times 1} \to \mathbb{R}^{2\times 1}$.

(i) Let $\mathbf{A} := \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. Then $T_{\mathbf{A}} : \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}} \longmapsto \begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix}^{\mathsf{T}}$.

$T_{\mathbf{A}}$ stretches each vector by a factor of 2.

(ii) Let $\mathbf{A} := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then $T_{\mathbf{A}} : \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}} \longmapsto \begin{bmatrix} x_2 & x_1 \end{bmatrix}^{\mathsf{T}}$.

$T_{\mathbf{A}}$ is the reflection in the line $x_1 = x_2$.

(iii) Let $\mathbf{A} := \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Then $T_{\mathbf{A}} : \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}} \mapsto \begin{bmatrix} -x_1 & -x_2 \end{bmatrix}^{\mathsf{T}}$.

$T_{\mathbf{A}}$ is the reflection in the origin.

(iv) Let $\mathbf{A} := \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, where $\theta \in (-\pi, \pi]$. Then

$T_{\mathbf{A}} : \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}} \longmapsto \begin{bmatrix} x_1\cos\theta - x_2\sin\theta & x_1\sin\theta + x_2\cos\theta \end{bmatrix}^{\mathsf{T}}$.

$T_{\mathbf{A}}$ is the rotation through an angle θ.

These are geometric interpretations of matrices.

While constructing an $m \times n$ matrix which represents a transformation from $\mathbb{R}^{n \times 1}$ to $\mathbb{R}^{m \times 1}$, we made an explicit use of the standard basis of $\mathbb{R}^{n \times 1}$ consisting of the basic column vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ (in this order). Also, we implicitly used the standard basic column vectors $\mathbf{e}_1, \ldots, \mathbf{e}_m$ in $\mathbb{R}^{m \times 1}$ (in this order) when we wrote $\mathbf{A} := \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$.

More generally, let an ordered basis $E := (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ of $\mathbb{R}^{n \times 1}$ and an ordered basis $F := (\mathbf{y}_1, \ldots, \mathbf{y}_m)$ of $\mathbb{R}^{m \times 1}$ be given. Then there are unique $a_{11}, \ldots, a_{1n}, \ldots, a_{m1}, \ldots, a_{mn} \in \mathbb{R}$ such that

$$\begin{array}{rcl} T(\mathbf{x}_1) & = & a_{11}\mathbf{y}_1 + \cdots + a_{j1}\mathbf{y}_j + \cdots + a_{m1}\mathbf{y}_m, \\ \vdots & & \qquad\qquad \vdots \qquad\qquad\quad \vdots \\ T(\mathbf{x}_k) & = & a_{1k}\mathbf{y}_1 + \cdots + a_{jk}\mathbf{y}_j + \cdots + a_{mk}\mathbf{y}_m, \\ \vdots & & \qquad\qquad \vdots \qquad\qquad\quad \vdots \\ T(\mathbf{x}_n) & = & a_{1n}\mathbf{y}_1 + \cdots + a_{jn}\mathbf{y}_j + \cdots + a_{mn}\mathbf{y}_m. \end{array}$$

The $m \times n$ matrix $[a_{jk}]$ is called the **matrix of the linear transformation** $T : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ with respect to the ordered basis $E := (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ of $\mathbb{R}^{n\times 1}$ and the ordered basis $F := (\mathbf{y}_1, \ldots, \mathbf{y}_m)$ of $\mathbb{R}^{m\times 1}$. This matrix is denoted by $\boxed{\mathbf{M}_F^E(T)}$.

Note: The kth column of $\mathbf{M}_F^E(T)$ is $\begin{bmatrix} a_{1k} & \cdots & a_{mk} \end{bmatrix}^\mathsf{T}$, where $T(\mathbf{x}_k) = a_{1k}\mathbf{y}_1 + \cdots + a_{jk}\mathbf{y}_j + \cdots + a_{mk}\mathbf{y}_m$ for $k = 1, \ldots, n$.

The $m \times n$ matrix $\mathbf{M}_F^E(T)$ represents the linear map T in the following sense. For $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$,

$$\begin{aligned} T(\alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n) &= \sum_{k=1}^{n} \alpha_k T(\mathbf{x}_k) = \sum_{k=1}^{n} \alpha_k \left(\sum_{j=1}^{m} a_{jk}\mathbf{y}_j \right) \\ &= \sum_{j=1}^{m} \left(\sum_{k=1}^{n} a_{jk}\alpha_k \right) \mathbf{y}_j = \beta_1\mathbf{y}_1 + \cdots + \beta_m\mathbf{y}_m, \end{aligned}$$

where $\beta_j := \sum_{k=1}^{n} a_{jk}\alpha_k$ for $j = 1, \ldots, m$. Thus

$$T\left(\begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix}\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix}\right)=\begin{bmatrix}\mathbf{y}_1 & \cdots & \mathbf{y}_m\end{bmatrix}\begin{bmatrix}\beta_1\\ \vdots\\ \beta_m\end{bmatrix},$$

$$\text{while}\quad \mathbf{M}_F^E(T)\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix}=\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & \vdots & \vdots\\ a_{m1} & \cdots & a_{mn}\end{bmatrix}\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix}=\begin{bmatrix}\beta_1\\ \vdots\\ \beta_m\end{bmatrix}.$$

Conversely, suppose we are given an $m \times n$ matrix $\mathbf{A}$. Define $T : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ as follows. For $\mathbf{x} := \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n$, let

$$T(\mathbf{x}) := \beta_1\mathbf{y}_1 + \cdots + \beta_m\mathbf{y}_m, \quad \text{where } \begin{bmatrix}\beta_1\\ \vdots\\ \beta_m\end{bmatrix} := \mathbf{A}\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix}.$$

Then T is a linear map, and $\mathbf{M}_F^E(T) = \mathbf{A}$.
In particular, this holds if E and F are the standard ordered bases of $\mathbb{R}^{n\times 1}$ and $\mathbb{R}^{m\times 1}$ respectively.

Examples

(i) Consider the map $T : \mathbb{R}^{2\times 1} \to \mathbb{R}^{3\times 1}$ defined by $T(\mathbf{x}) := \begin{bmatrix} x_1 - x_2 & -x_1 + 2x_2 & x_2 \end{bmatrix}^\mathsf{T}$ for $\mathbf{x} := \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\mathsf{T}$. Then T is a linear map. If $E := (\mathbf{e}_1, \mathbf{e}_2)$ is the standard ordered basis for $\mathbb{R}^{2\times 1}$ and $F := (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ is the standard ordered basis for $\mathbb{R}^{3\times 1}$, then $\mathbf{M}_F^E(T) = \begin{bmatrix} 1 & -1 \\ -1 & 2 \\ 0 & 1 \end{bmatrix}$.

On the other hand, let $E' := (\mathbf{e}_1',\, \mathbf{e}_2')$, where $\mathbf{e}_1' := \mathbf{e}_1$ and $\mathbf{e}_2' := \mathbf{e}_1 + \mathbf{e}_2$, and let $F' := (\mathbf{e}_1',\, \mathbf{e}_2',\, \mathbf{e}_3')$, where $\mathbf{e}_1' := \mathbf{e}_1$, $\mathbf{e}_2' := \mathbf{e}_1 + \mathbf{e}_2$ and $\mathbf{e}_3' := \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$. Then

$$T(\mathbf{e}_1') = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix}^\mathsf{T} = 2\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^\mathsf{T} - \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^\mathsf{T} = 2\mathbf{e}_1' - \mathbf{e}_2',$$

$$T(\mathbf{e}_2') = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^\mathsf{T} = -\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^\mathsf{T} + \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^\mathsf{T} = -\mathbf{e}_1' + \mathbf{e}_3'.$$

Hence $\mathbf{M}_{F'}^{E'}(T) = \begin{bmatrix} 2 & -1 \\ -1 & 0 \\ 0 & 1 \end{bmatrix}$.

(ii) Consider the map $T : \mathbb{R}^{3\times 1} \to \mathbb{R}^{3\times 1}$ defined by $T(\mathbf{x}) :=$ $\begin{bmatrix} 2.9x_1 + 0.6x_2 - 0.1x_3 & 2.9x_1 + 1.6x_2 - 1.1x_3 & 2.5x_1 + x_2 + 1.5x_3 \end{bmatrix}^{\mathsf{T}}$ for $\mathbf{x} := \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^{\mathsf{T}}$. Then T is a linear map.

If $E := (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ is the standard ordered basis for $\mathbb{R}^{3\times 1}$, then

$$\mathbf{M}_E^E(T) = \frac{1}{10}\begin{bmatrix} 29 & 6 & -1 \\ 29 & 16 & -11 \\ 25 & 10 & 15 \end{bmatrix}.$$

On the other hand, let $E' := (\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3')$, where $\mathbf{e}_1' := \begin{bmatrix} -1 & 3 & -1 \end{bmatrix}^{\mathsf{T}}, \mathbf{e}_2' := \begin{bmatrix} 1 & -1 & 3 \end{bmatrix}^{\mathsf{T}}, \mathbf{e}_3' := \begin{bmatrix} 2 & 1 & 4 \end{bmatrix}^{\mathsf{T}}$.
Then it can be checked that

$$T(\mathbf{e}_1') = \mathbf{e}_1', \quad T(\mathbf{e}_2') = 2\,\mathbf{e}_2', \quad T(\mathbf{e}_3') = 3\,\mathbf{e}_3'.$$

Hence $\mathbf{M}_{E'}^{E'}(T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, which is a diagonal matrix!

Remark: We have shown that if $T : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ is linear map, and if E and F are ordered bases of $\mathbb{R}^{n\times 1}$ and $\mathbb{R}^{m\times 1}$ respectively, then T is represented by an $m\times n$ matrix $\mathbf{M}_F^E(T)$ with respect to E and F. Now let V and W be subspaces of dimension n and m of some possibly higher dimensional spaces of vectors, and let T be a linear map from V to W. Even in this case, if E and F are ordered bases of V and W respectively, then the linear map T from V to W is represented, with respect to E and F, by an $m\times n$ matrix. This matrix is denoted by $\boxed{\mathbf{M}_F^E(T)}$.

Thus if $E := (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ and $F := (\mathbf{y}_1, \ldots, \mathbf{y}_m)$, and we let $\mathbf{A} := \mathbf{M}_F^E(T)$, then for $\mathbf{x} = \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n \in V$ and $\mathbf{y} = \beta_1\mathbf{y}_1 + \cdots + \beta_m\mathbf{y}_m \in W$, we see that

$$T(\mathbf{x}) = \mathbf{y} \iff \mathbf{A}\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix} = \begin{bmatrix}\beta_1\\ \vdots\\ \beta_m\end{bmatrix}.$$

Remark: Let V be a subspace of $\mathbb{R}^{n\times 1}$, W be a subspace of $\mathbb{R}^{m\times 1}$, and let $T : V \to W$ be a linear map. Two important subspaces associated with T are as follows.

(i) $\mathcal{N}(T) := \{\mathbf{x} \in V : T(\mathbf{x}) = \mathbf{0}\}$, called the **null space** of T,

(ii) $\mathcal{I}(T) := \{T(\mathbf{x}) : \mathbf{x} \in V\}$, called the **image space** of T.

We note that
a linear map T is one-one $\iff \mathcal{N}(T) = \{\mathbf{0}\}$, and
a linear map T is onto $\iff \mathcal{I}(T) = W$.

Further, if $V := \mathbb{R}^{n\times 1}$, $W := \mathbb{R}^{m\times 1}$, and $\mathbf{A} \in \mathbb{R}^{m\times n}$, then

$$\begin{aligned}\mathcal{N}(T_{\mathbf{A}}) &= \{\mathbf{x} \in \mathbb{R}^{n\times 1} : \mathbf{A}\mathbf{x} = \mathbf{0}\} = \mathcal{N}(\mathbf{A}),\\ \mathcal{I}(T_{\mathbf{A}}) &= \{\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{R}^{n\times 1}\} = \mathcal{C}(\mathbf{A}).\end{aligned}$$

The last equality follows by noting that if $\mathbf{A} = \begin{bmatrix}\mathbf{c}_1 & \cdots & \mathbf{c}_n\end{bmatrix}$, then $\mathbf{A}\mathbf{x} = x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n$ for $\mathbf{x} := \begin{bmatrix}x_1 & \cdots & x_n\end{bmatrix} \in \mathbb{R}^{n\times 1}$.

Example

Let $\mathbf{A} := \begin{bmatrix} 1 & -1 \\ -1 & 2 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 2}$. Then $T_{\mathbf{A}} : \mathbb{R}^{2\times 1} \to \mathbb{R}^{3\times 1}$.

In fact, $T_{\mathbf{A}} : \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}} \longmapsto \begin{bmatrix} x_1 - x_2 & -x_1 + 2x_2 & x_2 \end{bmatrix}^{\mathsf{T}}$ for all $\begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{2\times 1}$. Clearly, $\mathcal{N}(T_{\mathbf{A}}) = \{\mathbf{0}\}$. Also,

$$\mathcal{I}(T_{\mathbf{A}}) = \left\{ \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{3\times 1} : y_1 + y_2 - y_3 = 0 \right\}.$$

To see this, note that $(x_1 - x_2) + (-x_1 + 2x_2) - x_2 = 0$ for all $\begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{2\times 1}$, and if $\begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{3\times 1}$ satisfies $y_1 + y_2 - y_3 = 0$, then we may let $x_1 := y_1 + y_3$, $x_2 := y_3$, so that $x_1 - x_2 = y_1$, $-x_1 + 2x_2 = y_2$ and $x_2 = y_3$, that is, $T_{\mathbf{A}}(\begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathsf{T}}) = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}^{\mathsf{T}}$.

Note: $\mathcal{I}(T_{\mathbf{A}})$ is a plane through the origin $\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^{\mathsf{T}}$ in $\mathbb{R}^{3\times 1}$.

Complex Numbers

In our development of matrix theory, we have so far used real numbers as scalars, and we have considered matrices whose entries are real numbers. Now we introduce an extension of $\mathbb{R}$ which has all the properties that $\mathbb{R}$ has (and one more).

A **complex number** is a 2×2 matrix $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, where $a, b \in \mathbb{R}$. The set of all complex numbers is denoted by $\mathbb{C}$. Addition and multiplication in $\mathbb{C}$ are defined as in $\mathbb{R}^{2\times 2}$. Thus

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} + \begin{bmatrix} c & d \\ -d & c \end{bmatrix} = \begin{bmatrix} a+c & b+d \\ -(b+d) & a+c \end{bmatrix}$$

and

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} c & d \\ -d & c \end{bmatrix} = \begin{bmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{bmatrix}.$$

These algebraic operations possess all the usual properties such as associativity, distributivity and commutativity.

Moreover, the map $a \longmapsto \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$ from $\mathbb{R}$ to $\mathbb{C}$ is one-one.

Hence we identify $a \in \mathbb{R}$ with $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \in \mathbb{C}$. We define

$$i := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \text{so that } i^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

We write $\begin{bmatrix} a & b \\ -b & a \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} a + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} b \in \mathbb{C}$ as $a + ib$.

It follows that $(a + ib) + (c + id) = (a + c) + i(b + d)$ and $(a + ib)(c + id) = (ac - bd) + i(ad + bc)$.

Let $z \in \mathbb{C}$. Then $z = x + iy$ for unique $x, y \in \mathbb{R}$. Then x is called the **real part** of z, and it is denoted by $\mathfrak{R}(z)$, while y is called the **imaginary part** of z, and it is denoted by $\mathfrak{I}(z)$.

The complex number $x - iy$ is called the **conjugate** of $z = x + iy$, and it is denoted by $\overline{z}$.

We shall use complex numbers as scalars and consider matrices whose entries are complex numbers.

An $m \times n$ matrix with complex entries is an element of $\mathbb{C}^{m\times n}$. In particular, a row vector of length n belongs to $\mathbb{C}^{1\times n}$ and a column vector of length m belongs to $\mathbb{C}^{m\times 1}$.

For $\mathbf{A} := [a_{jk}] \in \mathbb{C}^{m\times n}$, define $\mathbf{A}^* := [\overline{a_{kj}}]$. Then $\mathbf{A}^* \in \mathbb{C}^{n\times m}$. It is called the **adjoint** of $\mathbf{A}$. We note that $(\alpha\mathbf{A} + \beta\mathbf{B})^* = \overline{\alpha}\mathbf{A}^* + \overline{\beta}\mathbf{B}^*$ for $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m\times n}$ and $\alpha, \beta \in \mathbb{C}$.

- A square matrix $\mathbf{A} = [a_{jk}]$ is called **self-adjoint** if $\mathbf{A}^* = \mathbf{A}$, that is, if $a_{jk} = \overline{a_{kj}}$ for all j, k.
- A square matrix $\mathbf{A} = [a_{jk}]$ is called **skew self-adjoint** if $a_{jk} = -\overline{a}_{kj}$ for all j, k.

Note: Every diagonal entry of a self-adjoint matrix is real since $a_{jj} = \overline{a_{jj}} \implies a_{jj} \in \mathbb{R}$ for $j = 1, \ldots, n$. On the other hand, the real part of every diagonal entry of a skew self-adjoint matrix is equal to 0 since $a_{jj} = -\overline{a_{jj}} \implies \Re(a_{jj}) = 0$ for $j = 1, \ldots, n$.

Note: If $\mathbf{x} := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^{\mathsf{T}} \in \mathbb{C}^{n\times 1}$ is a column vector, then $\mathbf{x}^* = \begin{bmatrix} \overline{x_1} & \cdots & \overline{x_n} \end{bmatrix} \in \mathbb{C}^{1\times n}$ is a row vector, and $\mathbf{x}^*\mathbf{x} = |x_1|^2 + \cdots + |x_n|^2$. It follows that $\mathbf{x}^*\mathbf{x} = 0 \iff \mathbf{x} = \mathbf{0}$.

A matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ defines a linear transformation from $\mathbb{C}^{n\times 1}$ to $\mathbb{C}^{m\times 1}$, and every linear transformation from $\mathbb{C}^{n\times 1}$ to $\mathbb{C}^{m\times 1}$ can be represented by a matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ (with respect to an ordered basis for $\mathbb{C}^{n\times 1}$ and an ordered basis for $\mathbb{C}^{m\times 1}$).

Similarly, we can consider vector subspaces of $\mathbb{C}^{n\times 1}$, and the concepts of linear dependence of vectors and of the span of a subset carry over to $\mathbb{C}^{n\times 1}$. The Fundamental Theorem for Linear Systems remains valid for matrices with complex entries.

Having thus completed our discussion of solution of a linear system, we shall turn to solution of an 'eigenvalue problem' associated with a matrix. In this development, the role of complex numbers will turn out to be important.

For future use, we define the **absolute value** of a complex number $z = x + iy \in \mathbb{C}$ by

$$|z| := \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}.$$

Then the **triangle inequality** $|z_1 + z_2| \leq |z_1| + |z_2|$ holds for all $z_1, z_2 \in \mathbb{C}$.
(Recall: $\|(x_1, y_1) + (x_2, y_2)\| \leq \|(x_1, y_1)\| + \|(x_2, y_2)\|$ for all $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$.)

Also, note that

$$\max\{|\Re(z)|, |\Im(z)|\} \ \leq \ |z| \ \leq \ |\Re(z)|+|\Im(z)| \quad \text{for all } z \in \mathbb{C}.$$

We shall use complex numbers as scalars and consider matrices whose entries are complex numbers.

The set of all $m \times n$ matrices with entries in $\mathbb{C}$ is denoted by $\mathbb{C}^{m\times n}$. In particular, $\mathbb{C}^{1\times n}$ is the set of all row vectors of length n, while $\mathbb{C}^{m\times 1}$ is the set of all column vectors of length m.

For $\mathbf{A} := [a_{jk}] \in \mathbb{C}^{m\times n}$, define $\mathbf{A}^* := [\overline{a_{kj}}]$. Then $\mathbf{A}^* \in \mathbb{C}^{n\times m}$. It is called the **conjugate transpose** or the **adjoint** of $\mathbf{A}$. Note: $(\alpha\mathbf{A} + \beta\mathbf{B})^* = \overline{\alpha}\mathbf{A}^* + \overline{\beta}\mathbf{B}^*$ for $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m\times n}$ and $\alpha, \beta \in \mathbb{C}$. In case $m = n$, then $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$.

- A square matrix $\mathbf{A} = [a_{jk}]$ is called **Hermitian** or **self-adjoint** if $\mathbf{A}^* = \mathbf{A}$, that is, if $a_{jk} = \overline{a_{kj}}$ for all j, k.
- A square matrix $\mathbf{A} = [a_{jk}]$ is called **skew-Hermitian** or **skew self-adjoint** if $a_{jk} = -\overline{a}_{kj}$ for all j, k.

Note: Every diagonal entry of a self-adjoint matrix is real since $a_{jj} = \overline{a_{jj}} \implies a_{jj} \in \mathbb{R}$ for $j = 1, \ldots, n$. On the other hand, the real part of every diagonal entry of a skew self-adjoint matrix

Note: If $\mathbf{x} := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^\mathsf{T} \in \mathbb{C}^{n\times 1}$ is a column vector, then $\mathbf{x}^* = \begin{bmatrix} \overline{x_1} & \cdots & \overline{x_n} \end{bmatrix} \in \mathbb{C}^{1\times n}$ is a row vector, and $\mathbf{x}^*\mathbf{x} = |x_1|^2 + \cdots + |x_n|^2$. It follows that $\mathbf{x}^*\mathbf{x} = 0 \iff \mathbf{x} = \mathbf{0}$.

A matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ defines a linear transformation from $\mathbb{C}^{n\times 1}$ to $\mathbb{C}^{m\times 1}$, and every linear transformation from $\mathbb{C}^{n\times 1}$ to $\mathbb{C}^{m\times 1}$ can be represented by a matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ (with respect to an ordered basis for $\mathbb{C}^{n\times 1}$ and an ordered basis for $\mathbb{C}^{m\times 1}$).

Similarly, we can consider vector subspaces of $\mathbb{C}^{n\times 1}$, and the concepts of linear dependence of vectors and of the span of a subset carry over to $\mathbb{C}^{n\times 1}$. The Fundamental Theorem for Linear Systems remains valid for matrices with complex entries.

Having thus completed our discussion of solution of a linear system, we shall turn to solution of an 'eigenvalue problem' associated with a matrix. In this development, the role of complex numbers will turn out to be important.