Matrix Transformations
Introduction to Linear Transformations

Applied Linear Algebra
MATH 5112/6012

UNIVERSITY OF Cincinnati

Another look at the matrix product $A\vec{x}$

Let $\vec{a}_j$ be the j^{th} column of some $m \times n$ matrix A, and $\vec{x}$ be in $\mathbb{R}^n$, so

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad \vec{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, \quad \text{and} \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Recall that the product $A\vec{x}$ is defined to be the LC

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n.$$

Since each $\vec{a}_j$ is a vector in $\mathbb{R}^m$, so is $A\vec{x}$. Thus we can define a vector function by the rule $\boxed{\vec{y} = A\vec{x}}$. Here

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \vec{y} = A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n.$$

The matrix transformation $\vec{y} = A\vec{x}$

This defines a *function* from $\mathbb{R}^n$ to $\mathbb{R}^m$; the input variable $\vec{x}$ comes from $\mathbb{R}^n$, it gets multiplied by the matrix A via the formula

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n,$$

and we call the resulting output $\vec{y}$. That is,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \vec{y} = A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n.$$

In linear algebra, functions are usually called *transformations*.

We write $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ to indicate that T is a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, meaning that the input variable $\vec{x}$ comes from $\mathbb{R}^n$ and the resulting output $\vec{y} = T(\vec{x})$ is a vector in $\mathbb{R}^m$.

A transformation (aka function) $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$

Here $\mathbb{R}^n$ is the *domain* of T– where the input variables $\vec{x}$ live, and $\mathbb{R}^m$ is the *codomain* of T– where the resulting output $\vec{y} = T(\vec{x})$ lives.

For each $\vec{x}$ in $\mathbb{R}^n$, $\vec{y} = T(\vec{x})$ is called the *T-image* of $\vec{x}$. If $\mathbb{S}$ is a bunch of vectors in $\mathbb{R}^n$ (i.e., $\mathbb{S} \subset \mathbb{R}^n$), then

$$T(\mathbb{S}) = \left\{\text{all images } T(\vec{x}) \text{ where } \vec{x} \text{ is in } \mathbb{S}\right\}$$

is called the *T-image of* $\mathbb{S}$.

The *range* of T is $\mathcal{R}ng(T) = T(\mathbb{R}^n)$, the set of all images $T(\vec{x})$; $\vec{b}$ is in $\mathcal{R}ng(T)$ if and only if $\vec{b} = T(\vec{x})$ for some $\vec{x}$. Evidently, $\mathcal{R}ng(T)$ is a subset of the codomain of T; $\mathcal{R}ng(T) \subset \mathbb{R}^m$.

An important question is to know which vectors $\vec{b}$ are in the range of T. That is: Given $\vec{b}$, when can we find an $\vec{x}$ in $\mathbb{R}^n$ with $T(\vec{x}) = \vec{b}$?

A matrix transformation $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$

Here we assume that $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ is given by the rule $T(\vec{x}) = A\vec{x}$ for some $m \times n$ matrix A. So, for each $\vec{x}$ in $\mathbb{R}^n$,

$$T(\vec{x}) = A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n$$

where $\vec{a}_j$ is the j^{th} column of A. Thus each image $T(\vec{x})$ is a LC of the columns of A.

Therefore, the *range* of T, which is the set of all images $T(\vec{x})$, is the set of all linear combinations of the columns of A; i.e., the range of T is the *span* of the columns of A.

Look at: Given $\vec{b}$, when can we find an $\vec{x}$ in $\mathbb{R}^n$ with $T(\vec{x}) = \vec{b}$? Here we are just asking whether or not we can solve $A\vec{x} = \vec{b}$.

The *range* of T is exactly all rhs vectors $\vec{b}$ such that $A\vec{x} = \vec{b}$ has a solution.

The range of $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$

Assume $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ is given by $T(\vec{x}) = A\vec{x}$ for some $m \times n$ matrix A.

The *range* of T is $\mathcal{R}ng(T) = T(\mathbb{R}^n)$ (the T-image of $\mathbb{R}^n$).
There are many equivalent ways to view the range of T!

- $\mathcal{R}ng(T)$ is all images $T(\vec{x})$ for $\vec{x}$ in $\mathbb{R}^n$
- $\mathcal{R}ng(T)$ is all $\vec{b}$ in $\mathbb{R}^m$ such that $\vec{b} = T(\vec{x})$ for some $\vec{x}$ in $\mathbb{R}^n$
- $\mathcal{R}ng(T)$ is all $\vec{b}$ in $\mathbb{R}^m$ such that $A\vec{x} = \vec{b}$ has a solution (is consistent)
- $\mathcal{R}ng(T)$ is the *span* of the columns of A