Matrix Transformation

CE40282-1: Linear Algebra
Hamid R. Rabiee and Maryam Ramezani
Sharif University of Technology

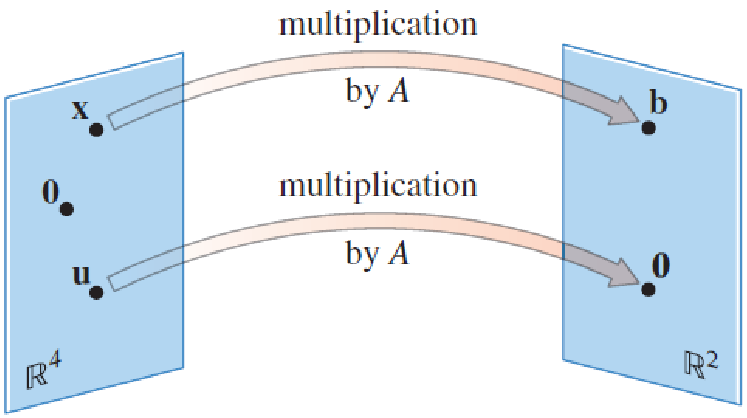
Linear Transformation

- Matrix is a linear transformation: map one vector to another vector

$$A \in \mathbb{R}^{m\times n},\ x \in \mathbb{R}^n,\ y \in \mathbb{R}^m : \qquad y_{m\times 1} = A_{m\times n} x_{n\times 1}$$

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

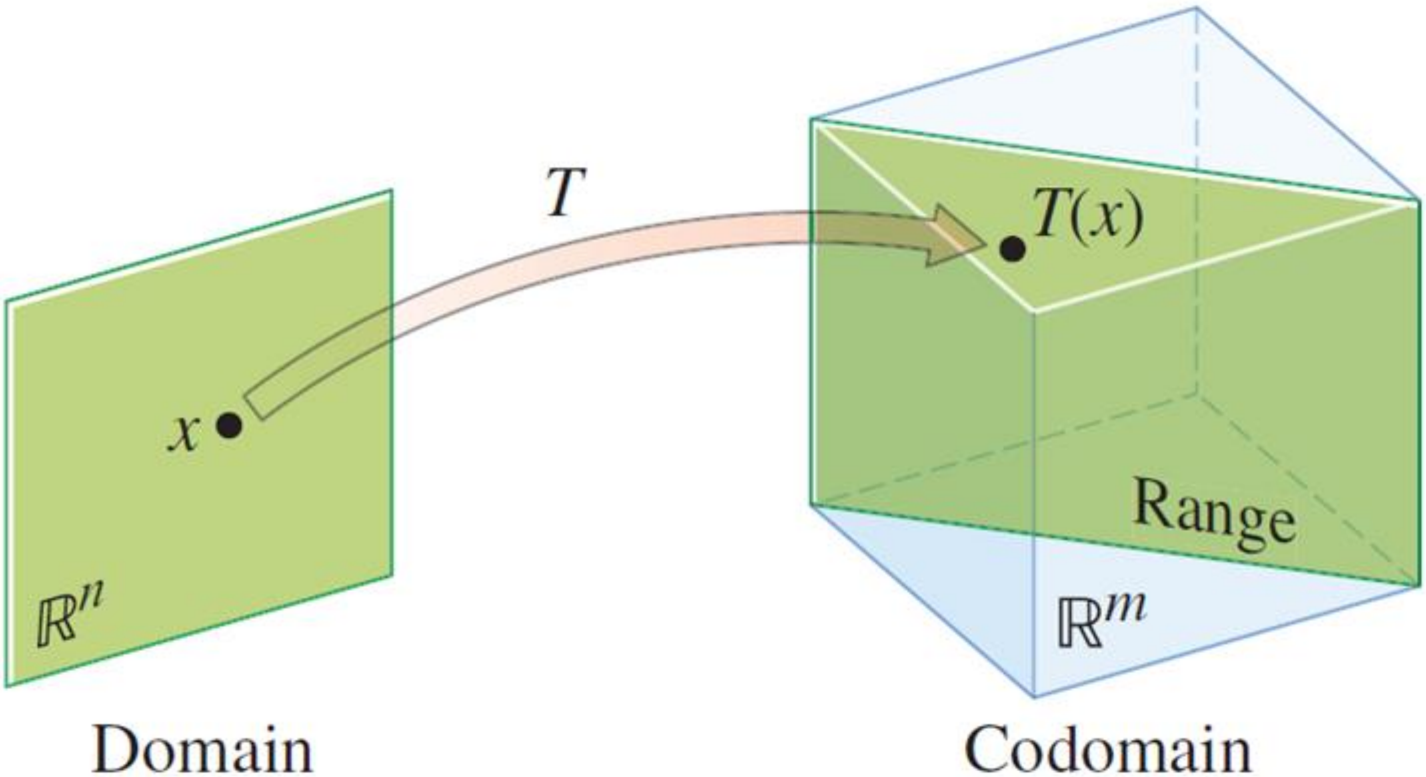
- Input-output



Matrix in applications

- Feature matrix
- Signal matrix
- Correlation matrix

Linear Transformation



Domain, codomain, and range of $T : \mathbb{R}^n \to \mathbb{R}^m$

Linear Transformation

EXAMPLE 1 Let $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$, and define a transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(\mathbf{x}) = A\mathbf{x}$, so that

$$T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$

a. Find $T(\mathbf{u})$, the image of $\mathbf{u}$ under the transformation T.

b. Find an $\mathbf{x}$ in $\mathbb{R}^2$ whose image under T is $\mathbf{b}$.

c. Is there more than one $\mathbf{x}$ whose image under T is $\mathbf{b}$?

d. Determine if $\mathbf{c}$ is in the range of the transformation T.

Linear mapping

- V, W are vector spaces over $\mathbb{F}$.

- A function $T : V \to W$ is called **linear** if

$$
\begin{aligned}
T(u+v) &= T(u) + T(v) && \text{for all } u, v \in V, \\
T(av) &= aT(v) && \text{for all } a \in \mathbb{F} \text{ and } v \in V.
\end{aligned}
$$

Linear mapping

- Example: which are linear mapping?
  - **zero** map $0 : V \to W$
  - **identity** map $I : V \to V$
  - Let $T : \mathcal{P}(\mathbb{F}) \to \mathcal{P}(\mathbb{F})$ be the **differentiation** map defined as $Tp(z) = p'(z)$.
  - $\boxed{\text{Let } T : \mathbb{R}^2 \to \mathbb{R}^2 \text{ be the map given by } T(x, y) = (x - 2y, 3x + y)}$
  - $f(x) = e^x$
  - $f : \mathbb{F} \to \mathbb{F}$ given by $f(x) = x - 1$

Linear mapping

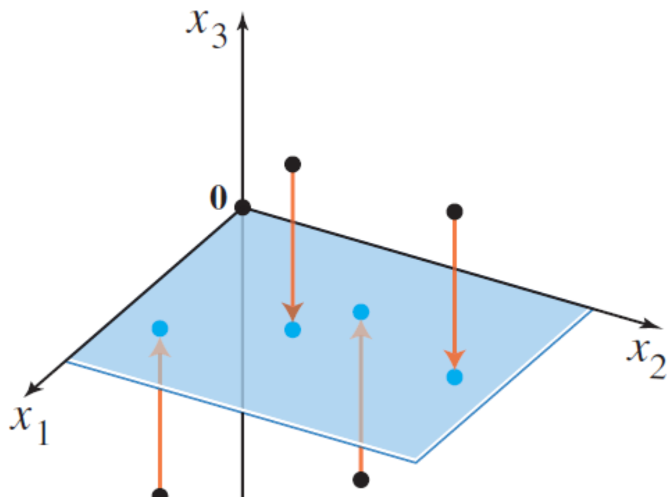
- Theorem

Let $(v_1, \ldots, v_n)$ *be a basis of* V *and* $(w_1, \ldots, w_n)$ *an arbitrary list of vectors in* W. *Then there exists a unique linear map*

$$T : V \to W \quad \text{such that } T(v_i) = w_i.$$

Projection

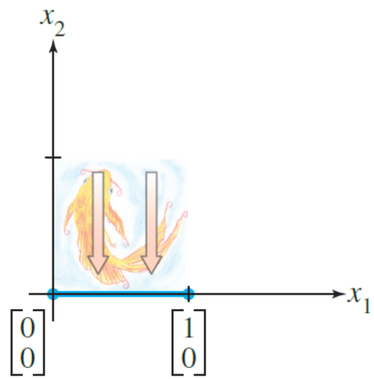
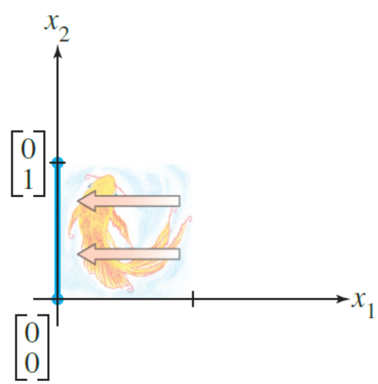
- Example:



If $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ *projects* points in $\mathbb{R}^3$ onto the x_1x_2-plane because

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$

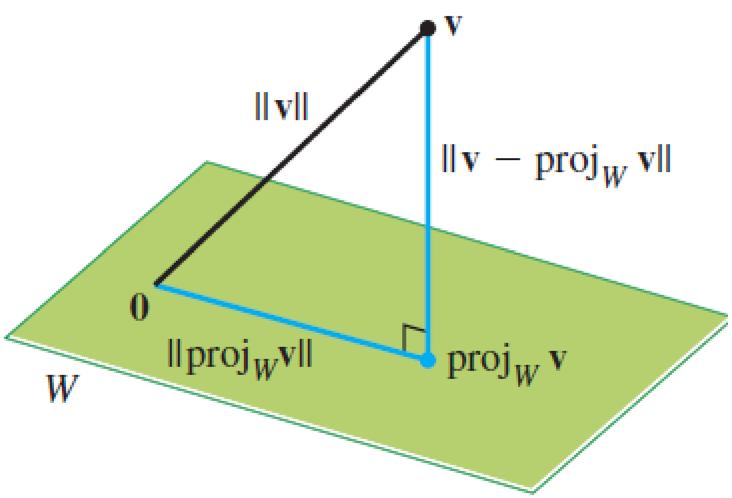
Projection

<table>
<tr><th>Transformation</th><th>Image of the Unit Square</th><th>Standard Matrix</th></tr>
<tr><td>Projection onto the x_1-axis</td><td>
</td><td>$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$</td></tr>
<tr><td>Projection onto the x_2-axis</td><td>
</td><td>$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$</td></tr>
</table>

Projection

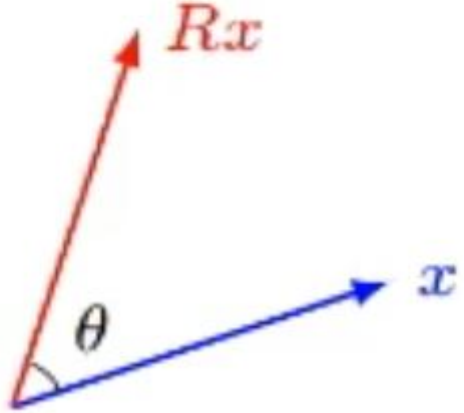
The ***projection*** of a vector $y \in \mathbb{R}^m$ onto the span of $\{x_1, \ldots, x_n\}$ is the vector $v \in \mathrm{span}(\{x_1, \ldots x_n\})$, such that v is as close as possible to y, as measured by the Euclidean norm $\|v - y\|_2$.

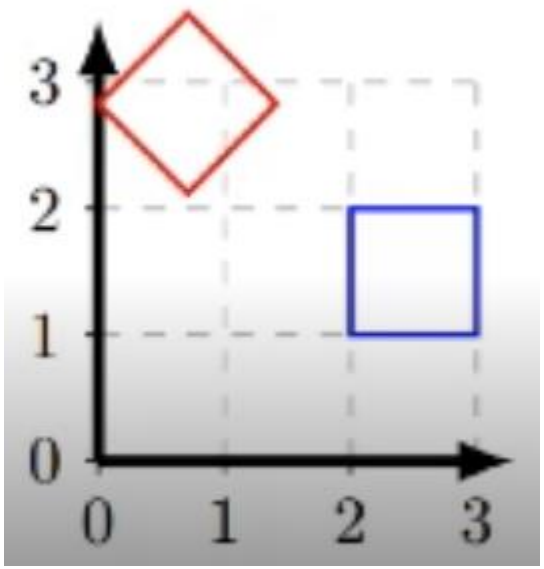
$$\mathrm{Proj}(y; \{x_1, \ldots x_n\}) = \mathrm{argmin}_{v \in \mathrm{span}(\{x_1, \ldots, x_n\})} \|y - v\|_2.$$



Rotation

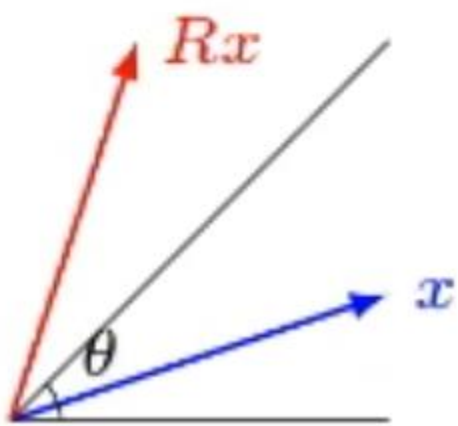
- $$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

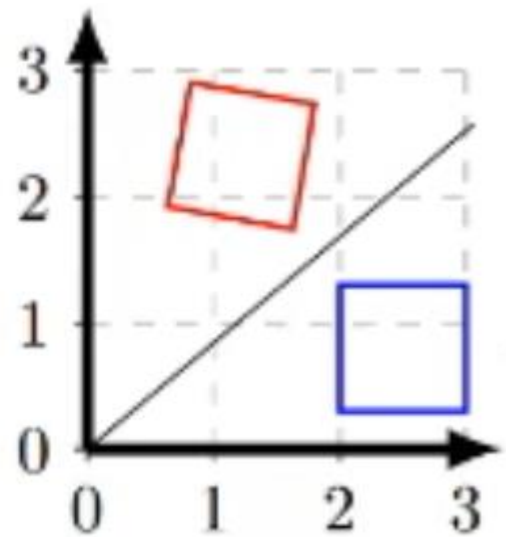




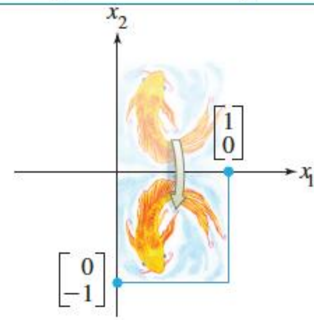
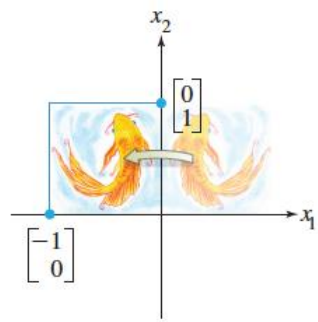
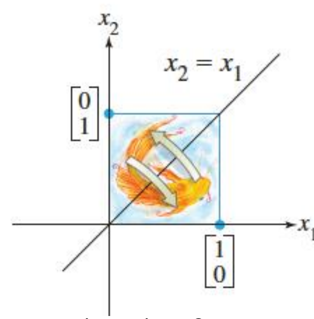
Reflection

- $R = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}$



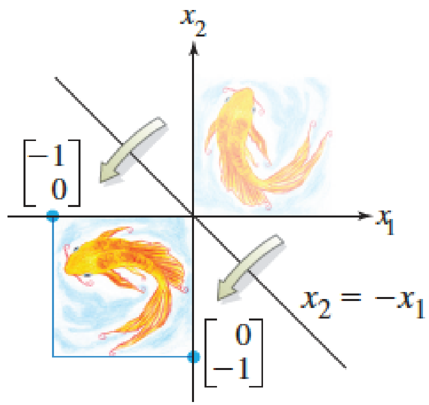


Reflection

Transformation	Image of the Unit Square	Standard Matrix
Reflection through the x_1-axis		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Reflection through the x_2-axis		$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
Reflection through the line $x_2 = x_1$		$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

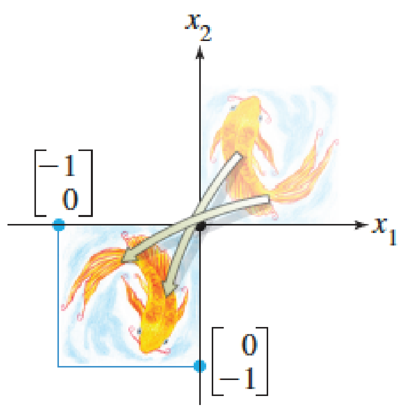
Reflection

Reflection through the line $x_2 = -x_1$



$$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

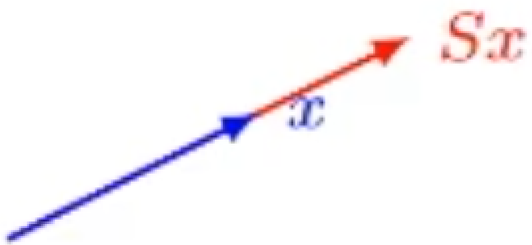
Reflection through the origin

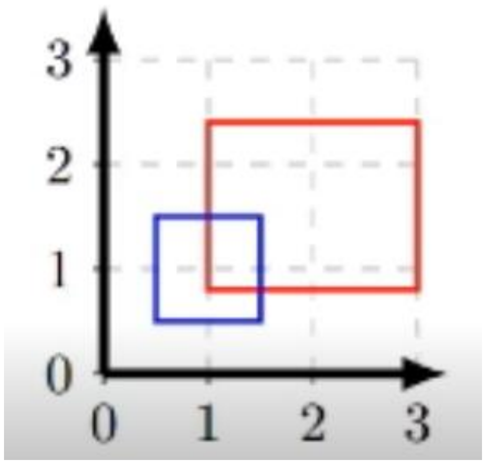


$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Uniform Scaling

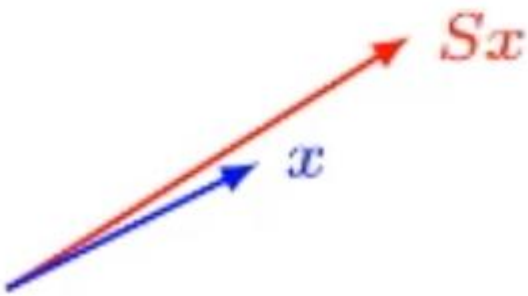
- $$S = sI = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix}$$

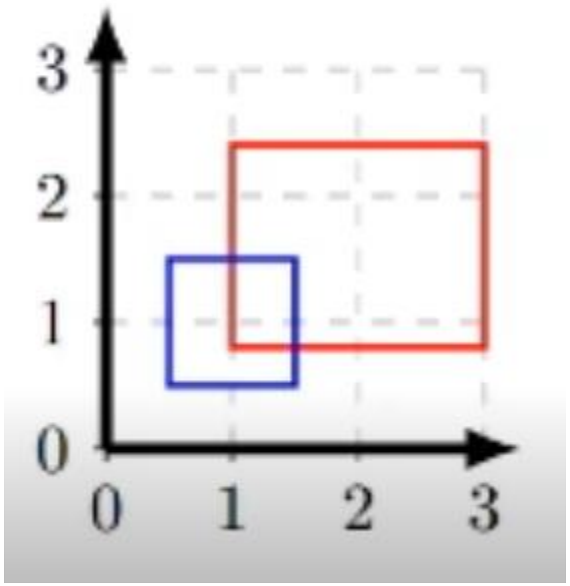




Non-uniform Scaling

- $S = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix}$





Shearing

- Example

Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$. The transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$

sheep

A typical shear matrix is of the form

$$S = \begin{pmatrix} 1 & 0 & 0 & \lambda & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

sheared sheep

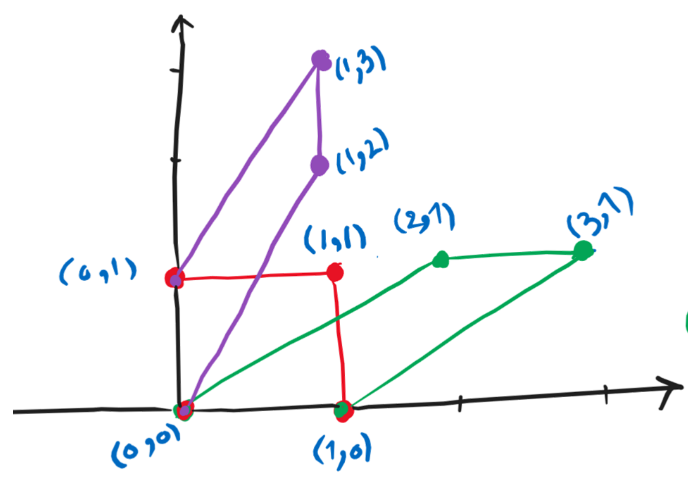
Shearing

A shear parallel to the *x* axis results in $x' = x + \lambda y$ and $y' = y$. In matrix form:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Similarly, a shear parallel to the *y* axis has $x' = x$ and $y' = y + \lambda x$. In matrix form:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$



x_axis $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ y_axis $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$

$(0,1) \to (2,1)$ $(0,1)$
$(0,0) \to (0,0)$ $(0,0)$
$(1,1) \to (3,1)$ $(1,3)$
$(1,0) \to (1,0)$ $(1,2)$

Difference Matrix

- $$D_{(n-1)\times n} = \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & -1 & 1 & 0 \\ 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}$$

$$D : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-1} \quad \Longrightarrow \quad D\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ \vdots \\ x_n - x_{n-1} \end{bmatrix}$$

- Example

$$\begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 3 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} -1-0 \\ 3-(-1) \\ 2-3 \\ 5-2 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ -1 \\ 3 \end{bmatrix}$$

Selectors

- an $m \times n$ *selector matrix*: each row is a unit vector (transposed)

$$A = \begin{bmatrix} e_{k_1}^T \\ \vdots \\ e_{k_m}^T \end{bmatrix}$$

multiplying by A selects entries of x:

$$Ax = (x_{k_1}, x_{k_2}, \ldots, x_{k_m})$$

- $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m \quad \Longrightarrow \quad A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{k_1} \\ x_{k_2} \\ \vdots \\ x_{k_m} \end{bmatrix}$

Selectors

- Example $$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 0 \\ -3 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

  - Selecting first and last elements of vector:
  - Reversing the elements of vector:

Slicing

- Keeping m elements from r to s (m=s-r+1)

$$\begin{bmatrix} 0_{m\times(r-1)} & I_{m\times m} & 0_{m\times(n-s)} \end{bmatrix}$$

- Example: Slicing two first and one last elements:

$$\begin{bmatrix} -1 \\ 2 \\ 0 \\ -3 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$$

Down Sampling

- Down sampling with k: selecting one sample in every k samples

- Example: k=2?

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ \vdots \end{bmatrix} = \begin{bmatrix} x_1 \\ x_3 \\ x_5 \\ \vdots \end{bmatrix}$$

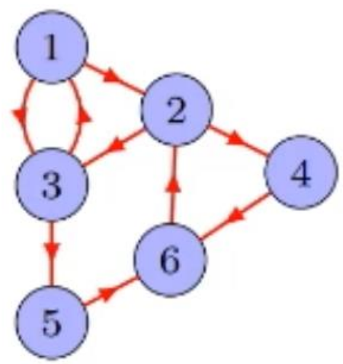
Applications

(i) sin 2A = 2 sin A cos A
(ii) cos 2A = cos² A − sin² A

- Rotation matrix

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \Longrightarrow R^n = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^n = \begin{bmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{bmatrix}$$

- Adjacency matrix



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

$$A^2 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \qquad A^3 = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Multiple Transformation

- $x_{n\times 1} \xrightarrow{A_{m\times n}} y_{m\times 1} \xrightarrow{B_{p\times m}} z_{p\times 1} \Longrightarrow \begin{cases} y = Ax \\ z = By \end{cases} \Longrightarrow z = B(Ax) = BAx$
- Example
  - Difference Matrix

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} \xrightarrow[4\times 5]{D} y = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ x_4 - x_3 \\ x_5 - x_4 \end{bmatrix} \xrightarrow[3\times 4]{D} z = \begin{bmatrix} x_3 - x_2 - (x_2 - x_1) \\ x_4 - x_3 - (x_3 - x_2) \\ x_5 - x_4 - (x_4 - x_3) \end{bmatrix} = \begin{bmatrix} x_3 - 2x_2 + x_1 \\ x_4 - 2x_3 + x_2 \\ x_5 - 2x_4 + x_3 \end{bmatrix}$$

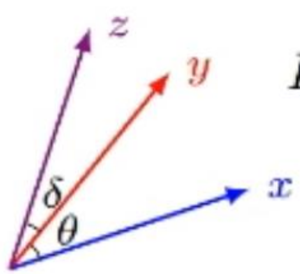
$$x \longrightarrow z \begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{bmatrix}_{3\times 5}$$

$$x \longrightarrow y \longrightarrow z$$

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}_{3\times 4} \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}_{4\times 5} = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{bmatrix}$$

Multiple Transformation

- $x_{n\times 1} \xrightarrow{A_{m\times n}} y_{m\times 1} \xrightarrow{B_{p\times m}} z_{p\times 1} \Longrightarrow \begin{cases} y = Ax \\ z = By \end{cases} \Longrightarrow z = B(Ax) = BAx$
- Example
  - Rotation



$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

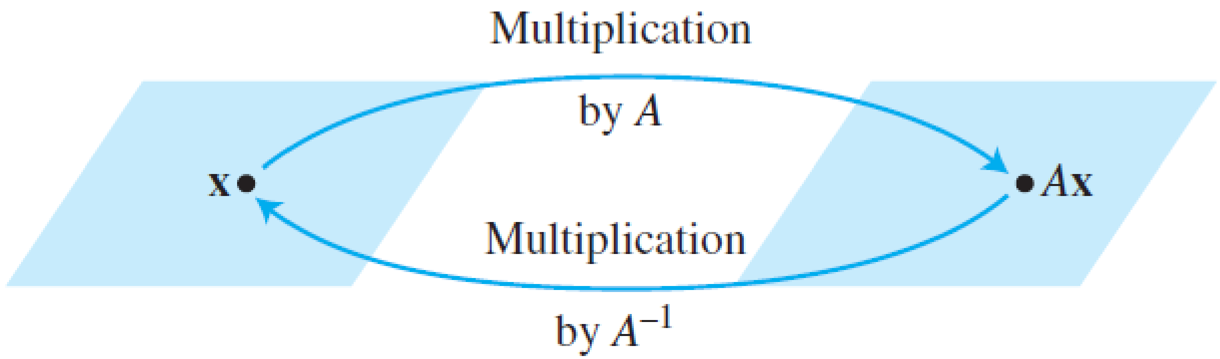
$$x \longrightarrow z \qquad z = R_{\delta+\theta}x \qquad \begin{bmatrix} \cos(\delta+\theta) & -\sin(\delta+\theta) \\ \sin(\delta+\theta) & \cos(\delta+\theta) \end{bmatrix}$$

$$x \longrightarrow y \longrightarrow z \; \begin{cases} y = R_\theta x \\ z = R_\delta y \end{cases} \Longrightarrow z = R_\delta R_\theta x \qquad \begin{bmatrix} \cos\delta & -\sin\delta \\ \sin\delta & \cos\delta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos\delta\cos\theta - \sin\delta\sin\theta & -\cos\delta\sin\theta - \sin\delta\cos\theta \\ \sin\delta\cos\theta + \cos\delta\sin\theta & -\sin\delta\sin\theta + \cos\delta\cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos(\delta+\theta) & -\sin(\delta+\theta) \\ \sin(\delta+\theta) & \cos(\delta+\theta) \end{bmatrix}$$

Invertible Linear Transformations



- Definition:

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be **invertible** if there exists a function $S : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$S(T(\mathbf{x})) = \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n$$
$$T(S(\mathbf{x})) = \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

Invertible Linear Transformations

- Theorem:

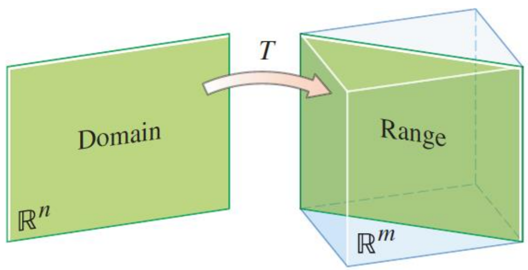
Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let A be the standard matrix for T. Then T is invertible if and only if A is an invertible matrix. In that case, the linear transformation S given by $S(\mathbf{x}) = A^{-1}\mathbf{x}$ is the unique function satisfying

$$S(T(\mathbf{x})) = \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n$$
$$T(S(\mathbf{x})) = \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

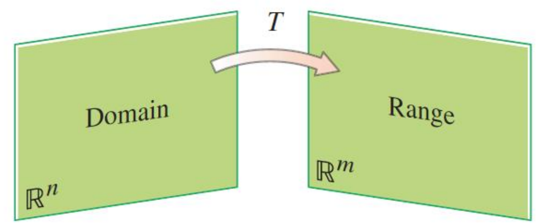
  - Proof in HW3!
- Example:

Mapping

- A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **onto** $\mathbb{R}^m$ if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at least one* $\mathbf{x}$ in $\mathbb{R}^n$.

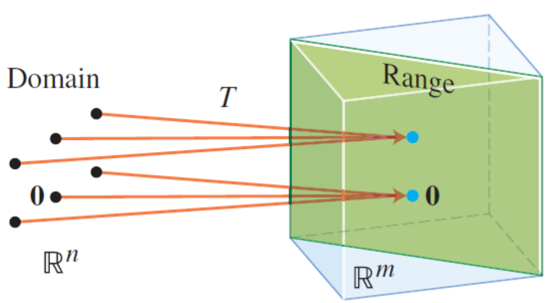


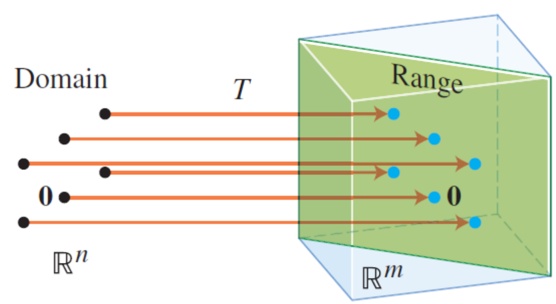
T is *not* onto $\mathbb{R}^m$



T is onto $\mathbb{R}^m$

- A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **one-to-one** if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at most one* $\mathbf{x}$ in $\mathbb{R}^n$.





Let T be the linear transformation whose standard matrix is

$$A = \begin{bmatrix} 1 & -4 & 8 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

Does T map $\mathbb{R}^4$ onto $\mathbb{R}^3$? Is T a one-to-one mapping?

One-to-One Linear Transformation

THEOREM

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

One-to-One Linear Transformation

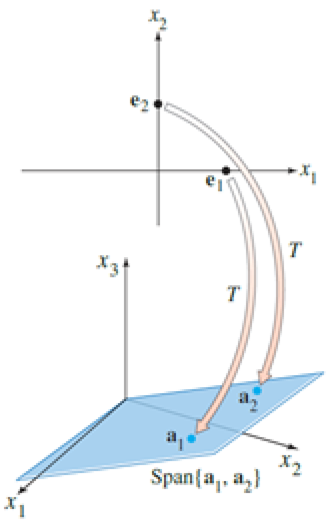
- Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let A be the standard matrix for T. Then:

  a. T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$;

  b. T is one-to-one if and only if the columns of A are linearly independent.

- Example

Let $T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2)$. Show that T is a one-to-one linear transformation. Does T map $\mathbb{R}^2$ onto $\mathbb{R}^3$?



Machine Learning Application

- The central problem in machine learning and deep learning is to meaningfully transform data: in other words, to learn useful representations of the input data at hand — representations that get us closer to the expected output.