

§1.8-1.9: Linear Transformations

Goal: think of the equation $A\mathbf{x} = \mathbf{b}$ in terms of the “multiplication by A ” function: its input is \mathbf{x} and its output is \mathbf{b} .

$$\begin{aligned} 2^2 &= 4 \\ 3^2 &= 9 \end{aligned}$$

Think of this as:

$$\begin{array}{ccc} 2 & \xrightarrow{\text{squaring}} & 4 \\ 3 & \xrightarrow{\text{squaring}} & 9 \end{array}$$

$$\begin{bmatrix} 8 & 3 & -4 \\ 5 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 9 \end{bmatrix}$$

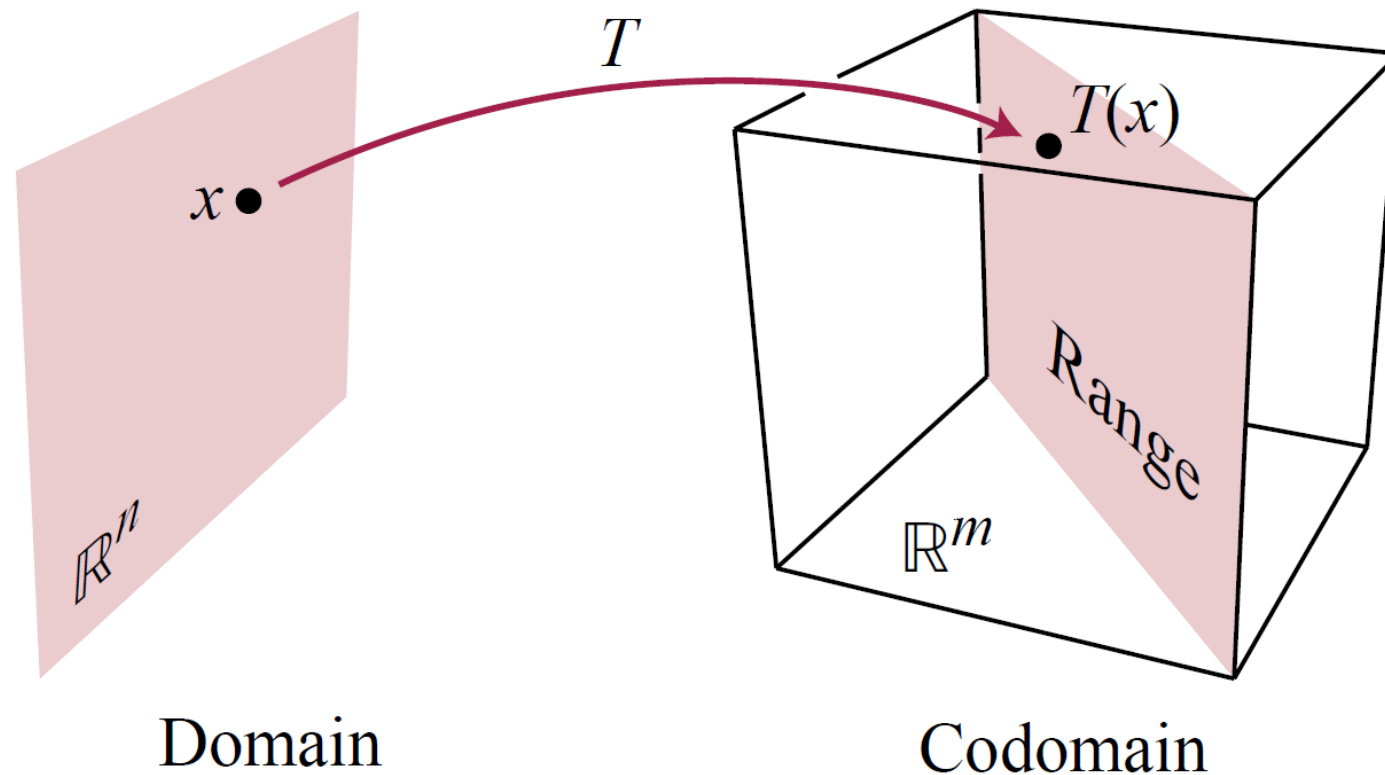
Think of this as:

$$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \xrightarrow{\text{multiply by } A} \begin{bmatrix} 10 \\ 9 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 3 & -4 \\ 5 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \xrightarrow{\text{multiply by } A} \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

Definition: A *function* f from \mathbb{R}^n to \mathbb{R}^m is a rule that assigns to each vector \mathbf{x} in \mathbb{R}^n a vector $f(\mathbf{x})$ in \mathbb{R}^m . We write $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$.



\mathbb{R}^n is the *domain* of f .

\mathbb{R}^m is the *codomain* of f .

$f(x)$ is the *image of x under f* .

The *range* is the set of all images. It is a subset of the codomain.

Example: $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = x^2$.

Its domain = codomain = \mathbb{R} , its range = {zero and positive numbers}.

Examples:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \text{ defined by } f \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} x_1^3 x_2 \\ 2x_2 \\ 0 \end{bmatrix}.$$

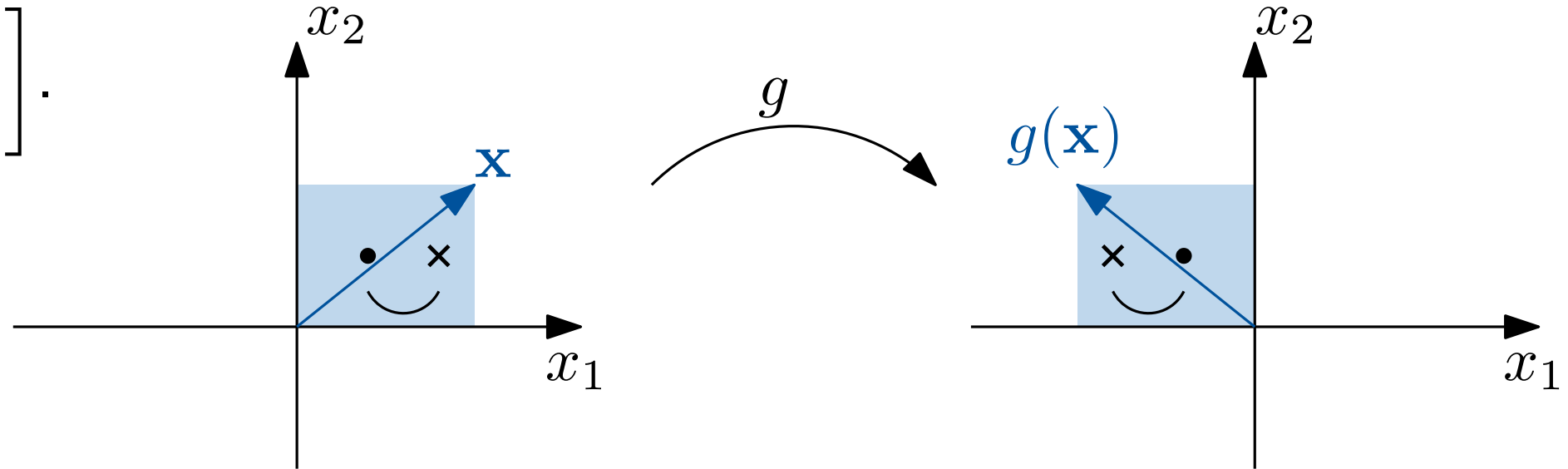
The range of f is the plane $z = 0$.

$$h : \mathbb{R}^3 \rightarrow \mathbb{R}^2, \text{ given by the matrix transformation } h(\mathbf{x}) = \begin{bmatrix} 8 & 3 & -4 \\ 5 & 1 & 2 \end{bmatrix} \mathbf{x}.$$

Examples:

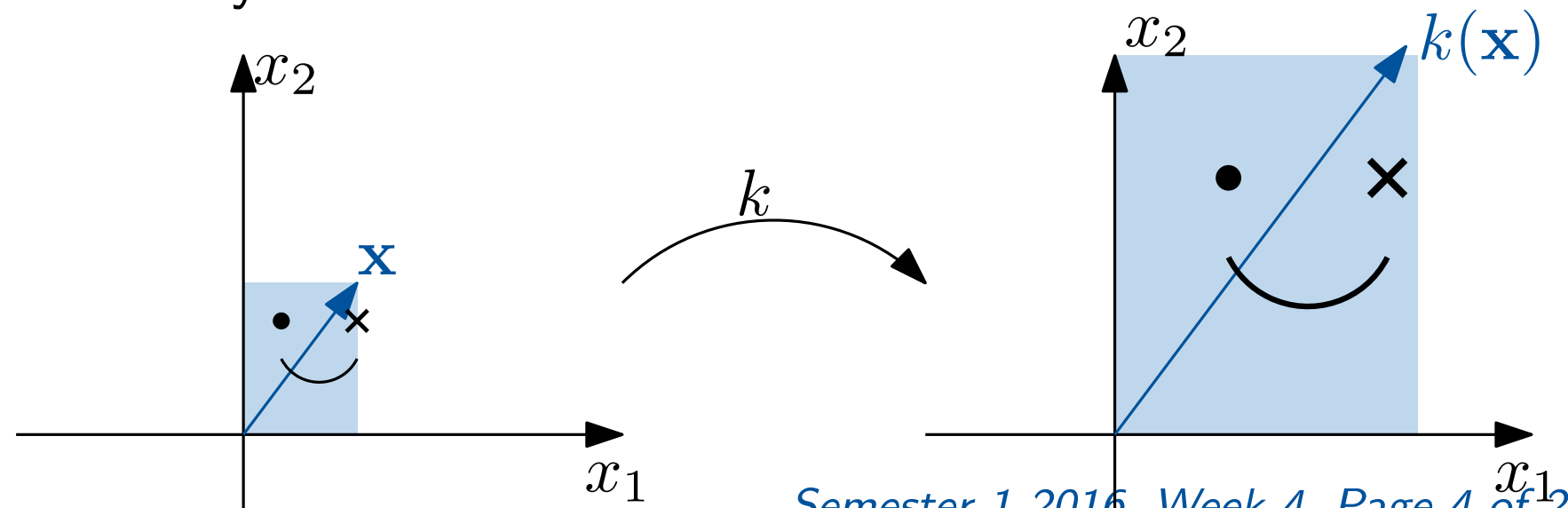
$g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, given by reflection through the x_2 -axis.

$$g \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix}.$$



$k : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, given by **dilation** by a factor of 3.

$$k(\mathbf{x}) = 3\mathbf{x}.$$



In this class, we will concentrate on functions that are **linear**.

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

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all \mathbf{u}, \mathbf{v} in the domain of T ;
2. $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars c and for all \mathbf{u} in the domain of T .

For your intuition: the name “linear” is because these functions preserve lines:
A line through the point \mathbf{p} in the direction \mathbf{v} is the set $\mathbf{p} + s\mathbf{v}$, where s is any number.
If T is linear, then the image of this set is

$$T(\mathbf{p} + s\mathbf{v}) \stackrel{1}{=} T(\mathbf{p}) + T(s\mathbf{v}) \stackrel{2}{=} T(\mathbf{p}) + sT(\mathbf{v}),$$

the line through the point $T(\mathbf{p})$ in the direction $T(\mathbf{v})$.
(If $T(\mathbf{v}) = \mathbf{0}$, then the image is just the point $T(\mathbf{p})$.)

Fact: A linear transformation T must satisfy $T(\mathbf{0}) = \mathbf{0}$.

Proof: Put $c = 0$ in condition 2.

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

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all \mathbf{u}, \mathbf{v} in the domain of T ;
2. $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars c and all \mathbf{u} , in the domain of T .

Example: $f \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} x_1^3 x_2 \\ 2x_2 \\ 0 \end{bmatrix}$ is not linear:

Take $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $c = 2$:

$$f \left(2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = f \left(\begin{bmatrix} 2 \\ 2 \end{bmatrix} \right) = \begin{bmatrix} 16 \\ 4 \\ 0 \end{bmatrix}.$$

$$2f \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = 2 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 16 \\ 4 \\ 0 \end{bmatrix}.$$

So condition 2 is false for f .

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

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all \mathbf{u}, \mathbf{v} in the domain of T ;
2. $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars c and all \mathbf{u} , in the domain of T .

Example: $g \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix}$ (reflection through the x_2 -axis) is linear:

$$1. \quad g \left(\begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix} \right) = \begin{bmatrix} -(u_1 + v_1) \\ u_2 + v_2 \end{bmatrix} = \begin{bmatrix} -u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} -v_1 \\ v_2 \end{bmatrix} = g \left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right) + g \left(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right).$$

$$2. \quad g \left(\begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix} \right) = \begin{bmatrix} -cu_1 \\ cu_2 \end{bmatrix} = c \begin{bmatrix} -u_1 \\ u_2 \end{bmatrix} = cg \left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right).$$

Alternatively, we can combine the two conditions at the same time, and check just one statement: $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$, for all scalars c, d and all vectors \mathbf{u}, \mathbf{v} .

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

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all \mathbf{u}, \mathbf{v} in the domain of T ;
2. $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars c and all \mathbf{u} , in the domain of T .

Alternatively, we can combine the two conditions at the same time, and check just one statement: $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$, for all scalars c, d and all vectors \mathbf{u}, \mathbf{v} .

Example: $k(\mathbf{x}) = 3\mathbf{x}$ (dilation by a factor of 3) is linear:

$$k(c\mathbf{u} + d\mathbf{v}) = 3(c\mathbf{u} + d\mathbf{v}) = 3c\mathbf{u} + 3d\mathbf{v} = k(c\mathbf{u}) + k(d\mathbf{v}).$$

Important Example: All **matrix transformations** $T(\mathbf{x}) = A\mathbf{x}$ are **linear**:

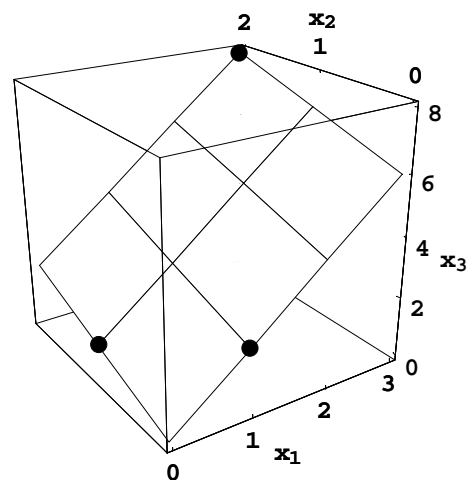
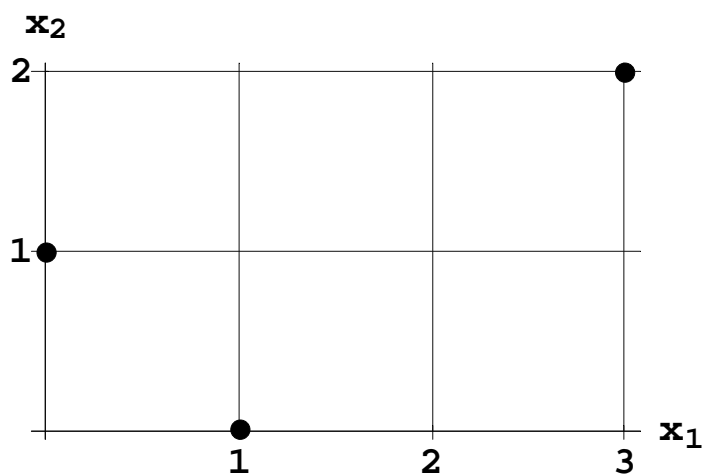
$$T(c\mathbf{u} + d\mathbf{v}) = A(c\mathbf{u} + d\mathbf{v}) = A(c\mathbf{u}) + A(d\mathbf{v}) = cA\mathbf{u} + dA\mathbf{v} = cT(\mathbf{u}) + dT(\mathbf{v}).$$

EXAMPLE: Let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Suppose $T : \mathbf{R}^2 \rightarrow \mathbf{R}^3$ is a linear transformation with

$$T(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Find the image of $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

Solution:



$$T(3\mathbf{e}_1 + 2\mathbf{e}_2) = 3T(\mathbf{e}_1) + 2T(\mathbf{e}_2)$$

In general:

Write \mathbf{e}_i for the vector with 1 in row i and 0 in all other rows.

For example, in \mathbb{R}^3 , we have $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

$\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ span \mathbb{R}^n , and $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n$.

So, if $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear transformation, then

$$T(\mathbf{x}) = T(x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n) = x_1 T(\mathbf{e}_1) + \dots + x_n T(\mathbf{e}_n) = \begin{bmatrix} | & & | \\ T(\mathbf{e}_1) & \dots & T(\mathbf{e}_n) \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Theorem 10: The matrix of a linear transformation: Every linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a matrix transformation: $T(\mathbf{x}) = A\mathbf{x}$ where A is the *standard matrix for T* , the $m \times n$ matrix given by

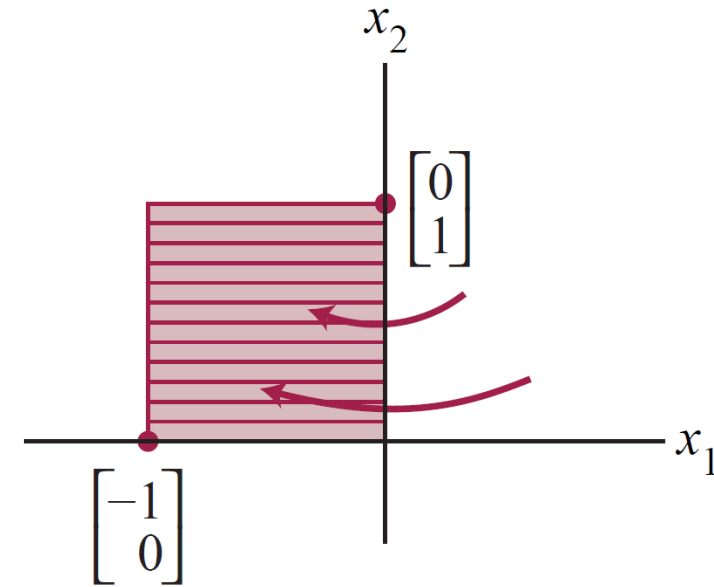
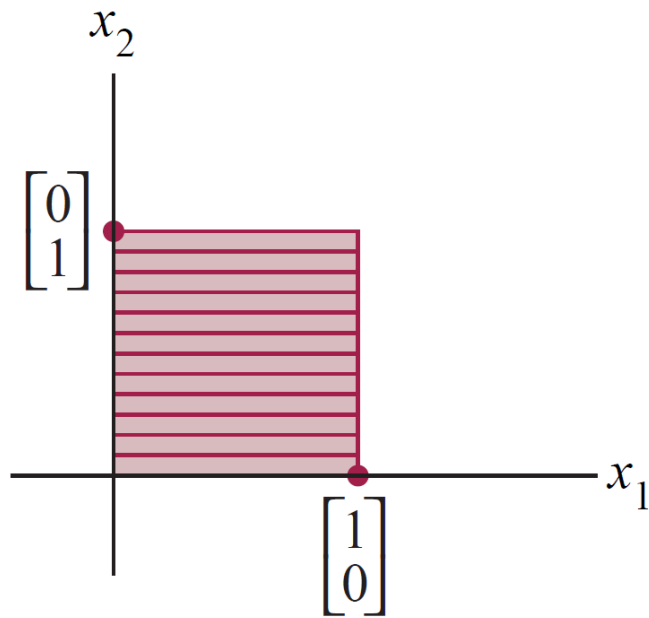
$$A = \begin{bmatrix} | & & | \\ T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \\ | & & | \end{bmatrix}.$$

Example: $k : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, given by *dilation* by a factor of 3, $k(\mathbf{x}) = 3\mathbf{x}$.

$$k(\mathbf{e}_1) = k\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = 3\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \quad k(\mathbf{e}_2) = k\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = 3\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

So the standard matrix of k is $\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$, i.e. $k(\mathbf{x}) = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \mathbf{x}$.

Example: $g \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix}$ (reflection through the x_2 -axis):



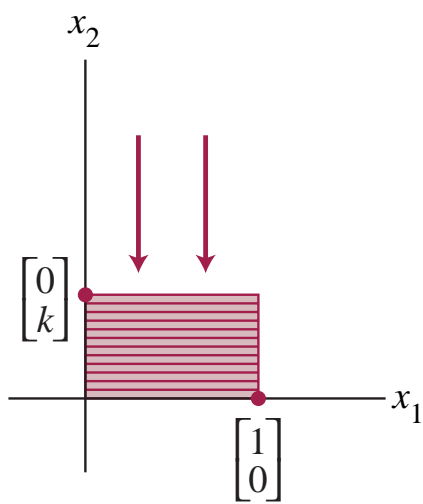
The standard matrix of g is $\begin{bmatrix} | & | \\ g(\mathbf{e}_1) & g(\mathbf{e}_2) \\ | & | \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$.

Indeed, $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix}$.

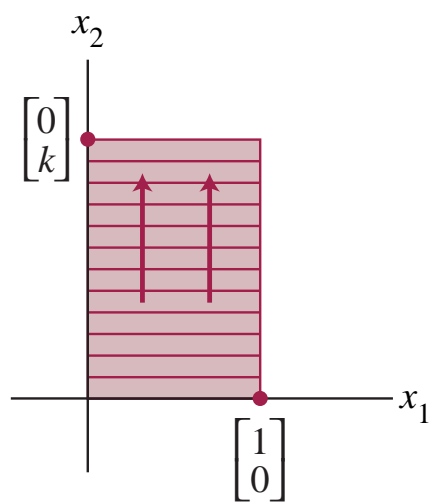
Vertical Contraction and Expansion

Image of the
Unit Square

Standard
Matrix



$$0 < k < 1$$



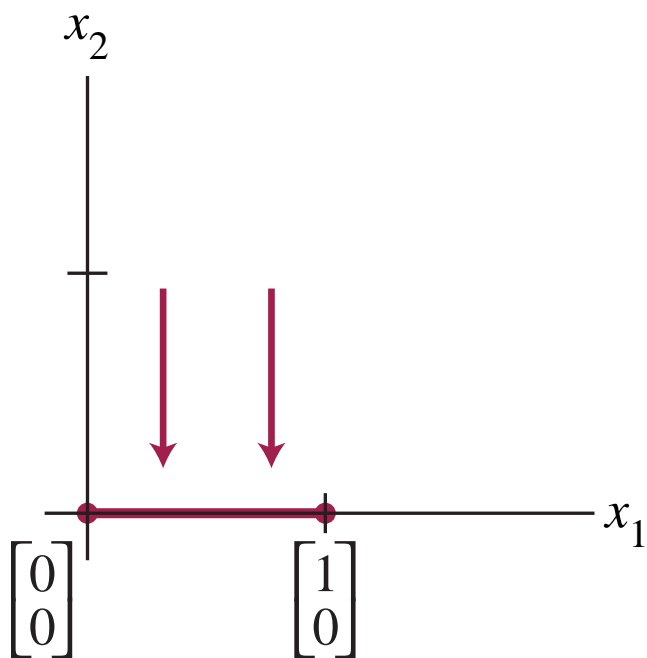
$$k > 1$$



Projection onto the x_1 -axis

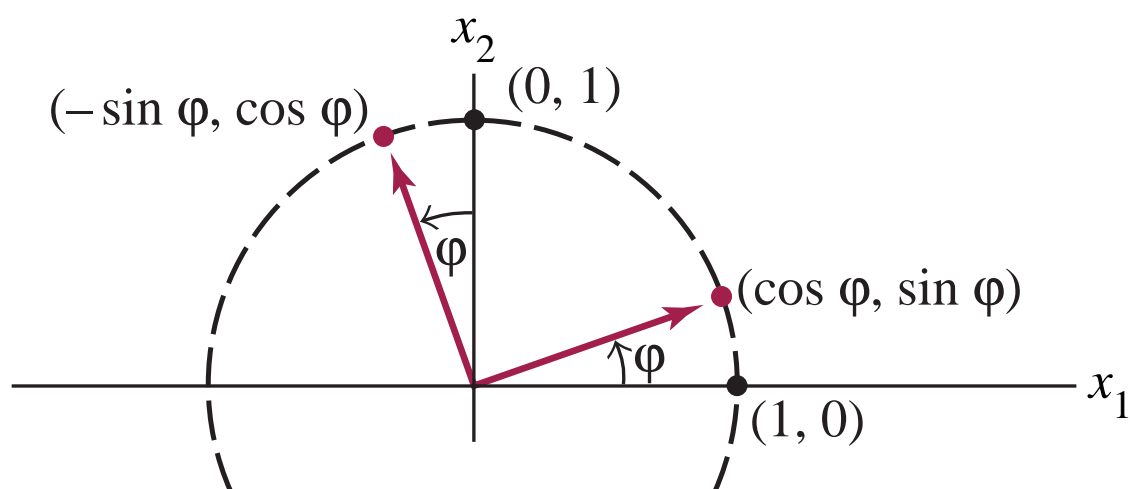
Image of the
Unit Square

Standard
Matrix



$$\begin{bmatrix} \\ \end{bmatrix}$$

EXAMPLE: $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by rotation counterclockwise about the origin through an angle φ :



Definition: A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is *onto* (surjective) if each \mathbf{y} in \mathbb{R}^m is the image of *at least one* \mathbf{x} in \mathbb{R}^n .

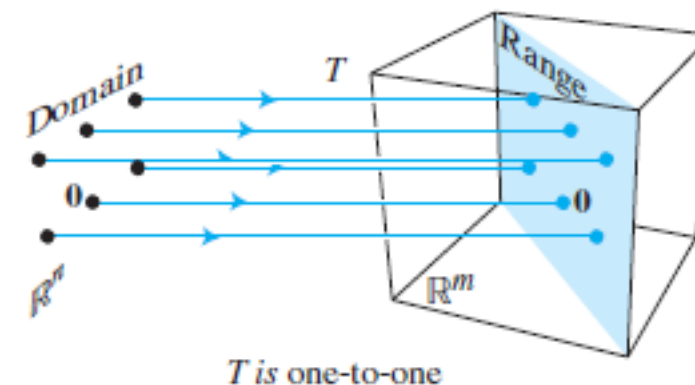
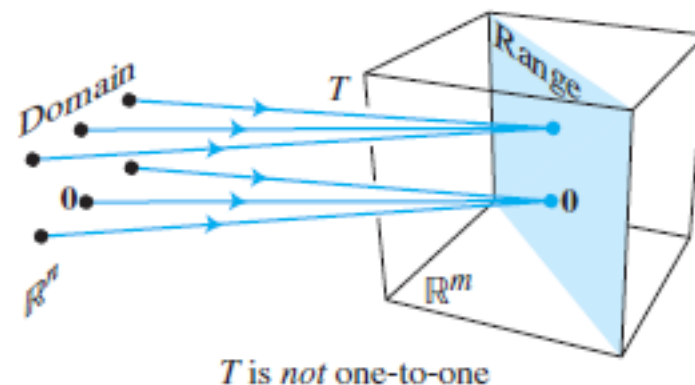
Other ways of saying this:

- The range is all of the codomain \mathbb{R}^m ,
- The equation $f(\mathbf{x}) = \mathbf{y}$ always has a solution.

Definition: A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is *one-to-one* (injective) if each \mathbf{y} in \mathbb{R}^m is the image of *at most one* \mathbf{x} in \mathbb{R}^n .

Other ways of saying this:

- ??? (something that only works for linear transformations),
- The equation $f(\mathbf{x}) = \mathbf{y}$ has no solutions or a unique solution.



Definition: A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is *onto* (surjective) if each \mathbf{y} in \mathbb{R}^m is the image of *at least one* \mathbf{x} in \mathbb{R}^n .

Definition: A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is *one-to-one* (injective) if each \mathbf{y} in \mathbb{R}^m is the image of *at most one* \mathbf{x} in \mathbb{R}^n .

Example: $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, defined by $f \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} x_1^3 x_2 \\ 2x_2 \\ 0 \end{bmatrix}$.

f is not onto, because $f(\mathbf{x}) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ does not have a solution. Indeed, the range of f is the plane $z = 0$.

f is one-to-one: the solution to $f(\mathbf{x}) = \begin{bmatrix} y_1 \\ y_2 \\ 0 \end{bmatrix}$ is $x_2 = \frac{1}{2}y_2$, $x_1 = \sqrt[3]{\frac{2y_1}{y_2}}$.

There is an easier way to check if a linear transformation is one-to-one:

Definition: The *kernel* of a linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the set of solutions to $T(\mathbf{x}) = \mathbf{0}$.

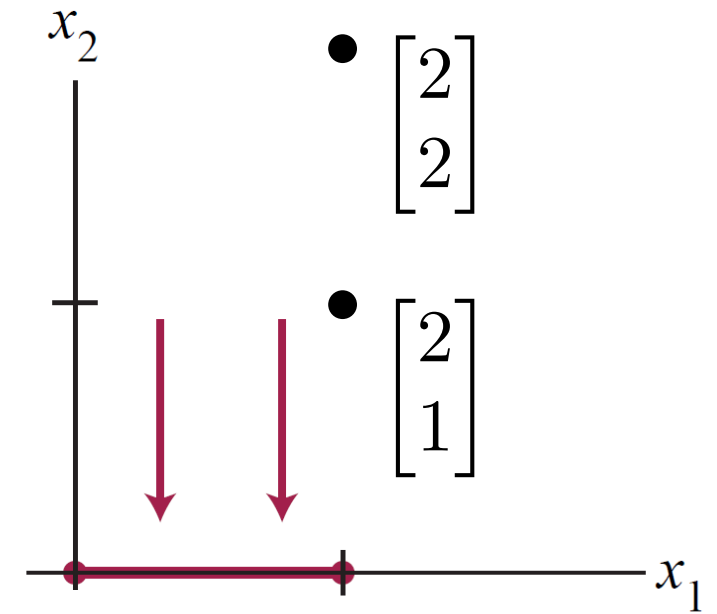
Fact: If $T(\mathbf{v}_1) = T(\mathbf{v}_2)$, then $\mathbf{v}_1 - \mathbf{v}_2$ is in the kernel of T .

Example: Let T be projection onto the x_1 -axis.

The kernel of T is the x_2 -axis.

$$T\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}\right) = T\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

$$\begin{bmatrix} 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \text{ which is in the kernel.}$$



Proof of Fact: If $T(\mathbf{v}_1) = T(\mathbf{v}_2) = \mathbf{y}$, then $T(\mathbf{v}_1 - \mathbf{v}_2) = T(\mathbf{v}_1) - T(\mathbf{v}_2) = \mathbf{y} - \mathbf{y} = \mathbf{0}$.

There is an easier way to check if a linear transformation is one-to-one:

Definition: The *kernel* of a linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the set of solutions to $T(\mathbf{x}) = \mathbf{0}$.

Fact: If $T(\mathbf{v}_1) = T(\mathbf{v}_2)$, then $\mathbf{v}_1 - \mathbf{v}_2$ is in the kernel of T .

Theorem: A linear transformation is *one-to-one* if and only if its *kernel* is $\{\mathbf{0}\}$.

Warning: this only works for linear transformations. For other functions, the solution sets of $f(\mathbf{x}) = \mathbf{y}$ and $f(\mathbf{x}) = \mathbf{0}$ are not related.

Proof:

Suppose T is one-to-one. So $T(\mathbf{x}) = \mathbf{0}$ has at most one solution. Since $\mathbf{0}$ is a solution, it must be the only one. So its kernel is $\{\mathbf{0}\}$.

Suppose the kernel of T is $\{\mathbf{0}\}$. Then, from the Fact, if there are vectors $\mathbf{v}_1, \mathbf{v}_2$ with $T(\mathbf{v}_1) = T(\mathbf{v}_2) = \mathbf{y}$, then $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$, i.e. $\mathbf{v}_1 = \mathbf{v}_2$.

For a linear transformation T whose standard matrix is A

Theorem: Uniqueness of solutions to linear systems: ~~For a matrix A ,~~ the following are equivalent:

- a. $A\mathbf{x} = \mathbf{0}$ has no non-trivial solution.
- b. If $A\mathbf{x} = \mathbf{b}$ is consistent, then it has a unique solution.
- c. The columns of A are linearly independent.
- d. $\text{rref}(A)$ has a pivot in every column (i.e. all variables are basic).
- e. T is a one-to-one function.

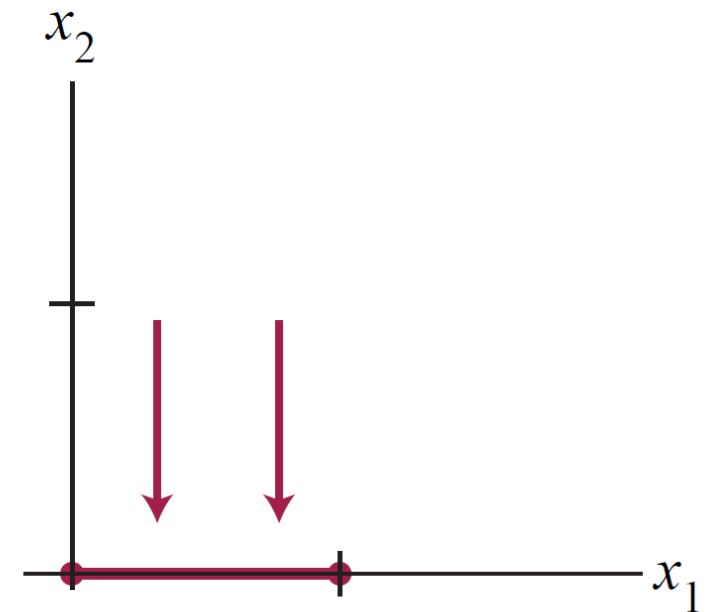
The range of a linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the set of images, i.e. the set of \mathbf{y} in \mathbb{R}^m with $T(\mathbf{x}) = \mathbf{y}$ for some \mathbf{x} .

So, if A is the standard matrix of T , then the range of T is the set of \mathbf{b} for which $A\mathbf{x} = \mathbf{b}$ has a solution.

So the **range** of T is the **span of the columns** of A .

Example: The standard matrix of projection onto the x_1 -axis is $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Its range is the x_1 -axis, which is also $\text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$.



So the **range** of T is the **span** of the columns of A .

For a linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ whose standard matrix is A

Theorem 4: Existence of solutions to linear systems: ~~For an $m \times n$ matrix A ,~~ the following statements are logically equivalent (i.e. for any particular matrix A , they are all true or all false):

- a. For each \mathbf{b} in \mathbb{R}^m , the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
- b. Each \mathbf{b} in \mathbb{R}^m is a linear combination of the columns of A .
- c. The columns of A span \mathbb{R}^m .
- d. $\text{rref}(A)$ has a pivot in every row.
- e. T is onto