§1.8-1.9: Linear Transformations

This week's goal is to 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 $\mathbf{b}$.

Primary One:

$2^2 = 4$
$3^2 = 9$

Primary Four:

Think of this as: $2 \xrightarrow{\text{squaring}} 4$, $3 \xrightarrow{\text{squaring}} 9$

Last week:

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

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

Today:

Think of this as:

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

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

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 $\mathbf{b}$.

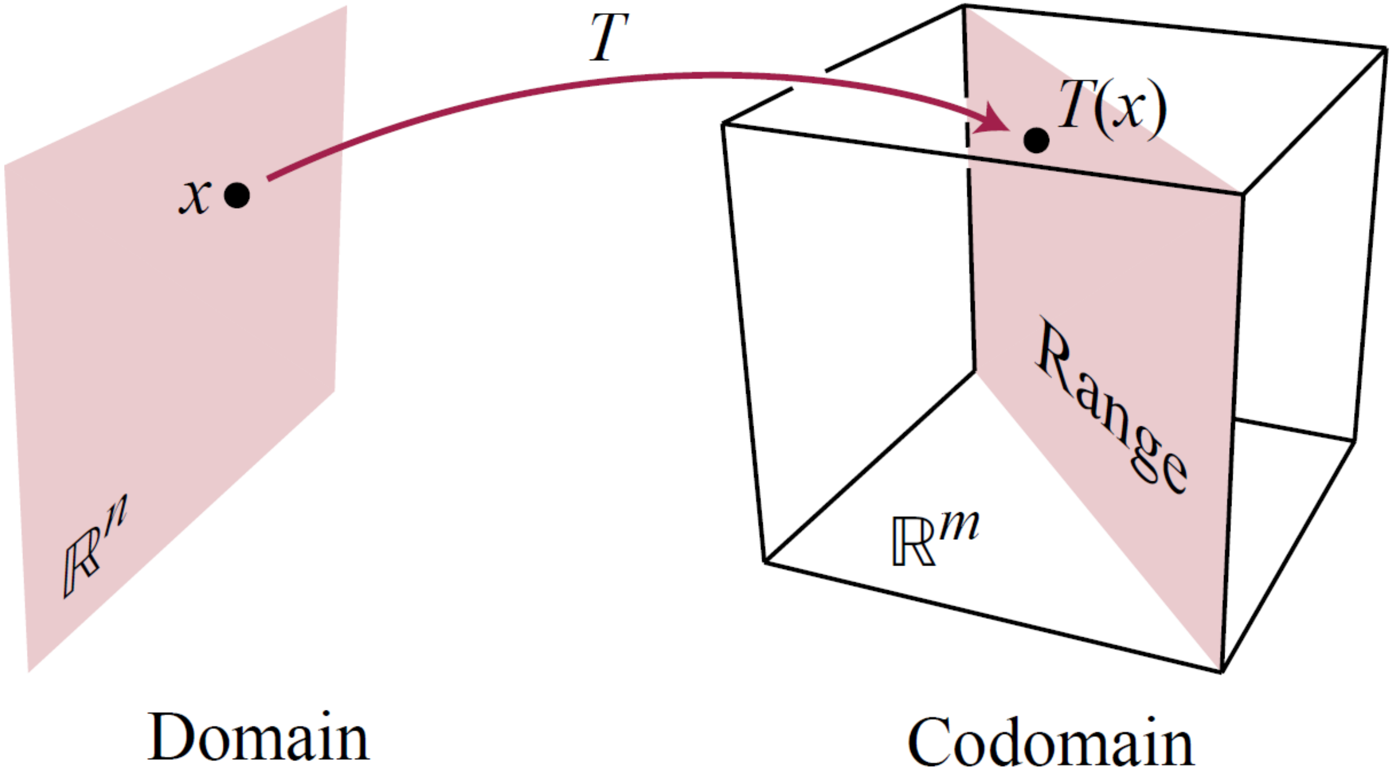
In this class, we are interested in functions that are linear (see p6 for the definition).
Key skills:

i Determine whether a function is linear (p7-9);
(This involves the important mathematical skill of "axiom checking", which also appears in other classes.)

ii Find the standard matrix of a linear function (p12-13);

iii Describe existence and uniqueness of solutions in terms of linear functions (p17-27).

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 $f : \mathbb{R}^n \to \mathbb{R}^m$.



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

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

$f(\mathbf{x})$ is the *image of* $\mathbf{x}$ *under* f.

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

Example: $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$.
Its domain $=$ codomain $= \mathbb{R}$, its range $= \{y \in \mathbb{R} \mid y \geq 0\}$.

Examples:

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

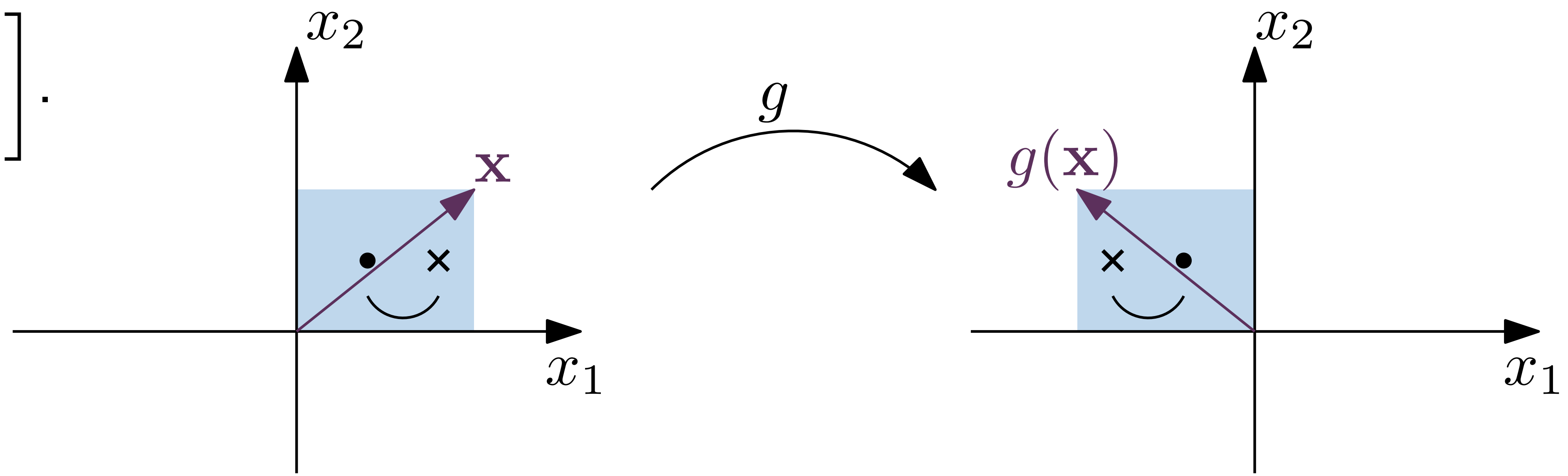
The range of f is the plane $z = 0$ (it is obvious that the range must be a subset of the plane $z = 0$, and with a bit of work (see p19), we can show that all points in $\mathbb{R}^3$ with $z = 0$ is the image of some point in $\mathbb{R}^2$ under f).

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

Geometric Examples:

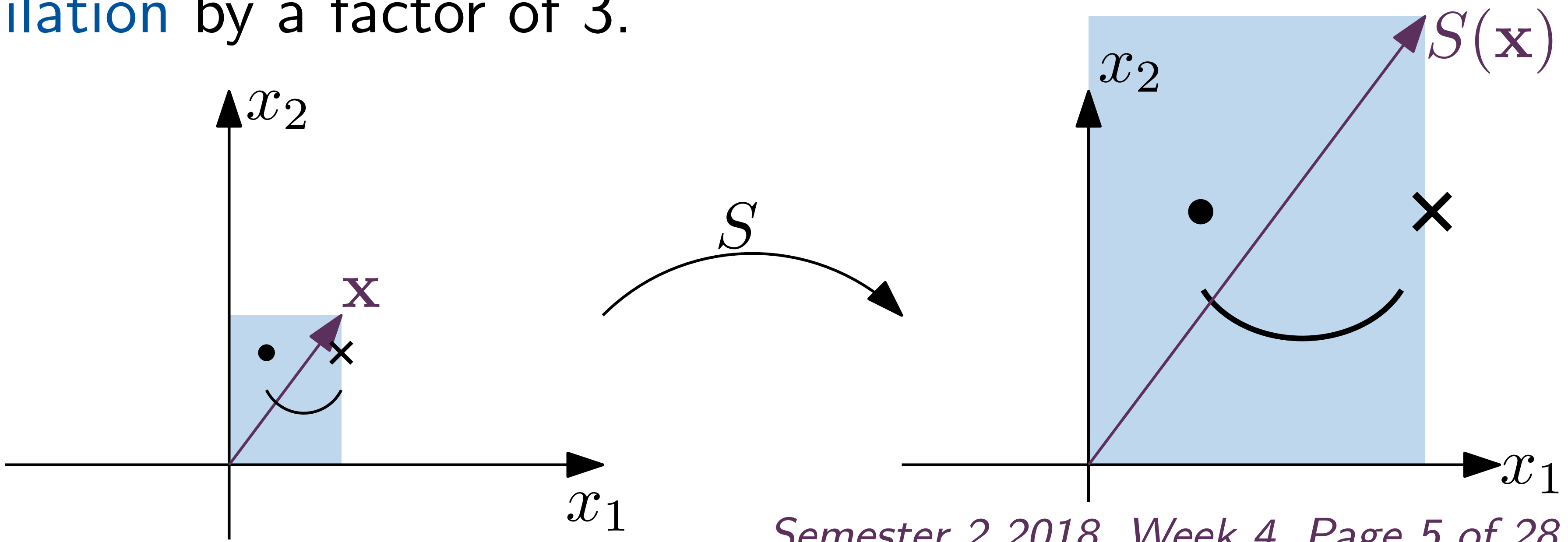
$g : \mathbb{R}^2 \to \mathbb{R}^2$, given by reflection through the x_2-axis.

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



$S : \mathbb{R}^2 \to \mathbb{R}^2$, given by dilation by a factor of 3.

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



In this class, we will concentrate on functions that are linear. (For historical reasons, people like to say "linear transformation" instead of "linear function".)

Definition: A function $T : \mathbb{R}^n \to \mathbb{R}^m$ 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 $T(\mathbf{p})$.)

In this class, we will concentrate on functions that are linear. (For historical reasons, people like to say "linear transformation" instead of "linear function".)

Definition: A function $T : \mathbb{R}^n \to \mathbb{R}^m$ 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

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

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 $T : \mathbb{R}^n \to \mathbb{R}^m$ 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: Is $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_2^3 \\ 2x_1 + x_2 \\ 0 \end{bmatrix}$ linear?

Definition: A function $T : \mathbb{R}^n \to \mathbb{R}^m$ 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: $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_2^3 \\ 2x_1 + x_2 \\ 0 \end{bmatrix}$ is not linear:

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

$$f\left(2\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = f\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 8 \\ 6 \\ 0 \end{bmatrix}.$$

$$2f\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = 2\begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 8 \\ 6 \\ 0 \end{bmatrix}.$$

So condition 2 is false for f.

Exercise: find a $\mathbf{u}$ and a $\mathbf{v}$ to show that condition 1 is also false.

Definition: A function $T : \mathbb{R}^n \to \mathbb{R}^m$ 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. $g\left(\begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}\right) = \begin{bmatrix} -u_1 - v_1 \\ u_2 + v_2 \end{bmatrix} = \begin{bmatrix} -u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} -v_1 \\ v_2 \end{bmatrix} = g\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right) + g\left(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right).$
2. $g\left(\begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix}\right) = \begin{bmatrix} -cu_1 \\ cu_2 \end{bmatrix} = c\begin{bmatrix} -u_1 \\ u_2 \end{bmatrix} = cg\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right).$

Definition: A function $T : \mathbb{R}^n \to \mathbb{R}^m$ 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. $g\left(\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \end{bmatrix}\right) = \begin{bmatrix} -u_1-v_1 \\ u_2+v_2 \end{bmatrix} = \begin{bmatrix} -u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} -v_1 \\ v_2 \end{bmatrix} = g\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right) + g\left(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right).$
2. $g\left(\begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix}\right) = \begin{bmatrix} -cu_1 \\ cu_2 \end{bmatrix} = c\begin{bmatrix} -u_1 \\ u_2 \end{bmatrix} = cg\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right).$

Notice from the previous two examples:
To show that a function is linear, check both conditions for general $\mathbf{u}, \mathbf{v}, c$ (i.e. use variables).
To show that a function is not linear, show that one of the conditions is not satisfied for a particular numerical values of $\mathbf{u}$ and $\mathbf{v}$ (for 1) or of c and $\mathbf{u}$ (for 2).

Definition: A function $T : \mathbb{R}^n \to \mathbb{R}^m$ 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.

For simple functions, 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 $\mathbf{u}, \mathbf{v}$. (Condition 1 is the case $c = d = 1$, condition 2 is the case $d = 0$. Exercise: show that if T satisfies conditions 1 and 2, then T satisfies the combined condition.)

Notice from the previous two examples:
To show that a function is linear, check both conditions for general $\mathbf{u}, \mathbf{v}, c$ (i.e. use variables).
To show that a function is not linear, show that one of the conditions is not satisfied for a particular numerical values of $\mathbf{u}$ and $\mathbf{v}$ (for 1) or of c and $\mathbf{u}$ (for 2).

Definition: A function $T : \mathbb{R}^n \to \mathbb{R}^m$ 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.

For simple functions, 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 $\mathbf{u}, \mathbf{v}$. (Condition 1 is the case $c = d = 1$, condition 2 is the case $d = 0$. Exercise: show that if T satisfies conditions 1 and 2, then T satisfies the combined condition.)

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

$$S(c\mathbf{u} + d\mathbf{v}) = 3(c\mathbf{u} + d\mathbf{v}) = 3c\mathbf{u} + 3d\mathbf{v} = cS(\mathbf{u}) + dS(\mathbf{v}).$$

Definition: A function $T : \mathbb{R}^n \to \mathbb{R}^m$ 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.

For simple functions, 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 $\mathbf{u}, \mathbf{v}$. (Condition 1 is the case $c = d = 1$, condition 2 is the case $d = 0$. Exercise: show that if T satisfies conditions 1 and 2, then T satisfies the combined condition.)

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

$$S(c\mathbf{u} + d\mathbf{v}) = 3(c\mathbf{u} + d\mathbf{v}) = 3c\mathbf{u} + 3d\mathbf{v} = cS(\mathbf{u}) + dS(\mathbf{v}).$$

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

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

In general:

Write $\mathbf{e_i}$ for the vector with 1 in row i and 0 in all other rows.
(So $\mathbf{e}_i$ means a different thing depending on which $\mathbb{R}^n$ we are working in.)

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 $\mathbf{x} = \begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} = x_1\mathbf{e_1} + \cdots + x_n\mathbf{e_n}$.

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

$$T(\mathbf{x}) = T(x_1\mathbf{e_1} + \cdots + x_n\mathbf{e_n}) = x_1T(\mathbf{e_1}) + \cdots + x_nT(\mathbf{e_n}) = \begin{bmatrix} | & & | \\ T(\mathbf{e}_1) & \dots & T(\mathbf{e}_n) \\ | & & | \end{bmatrix} \begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}.$$

Theorem 10: The matrix of a linear transformation: Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ 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

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

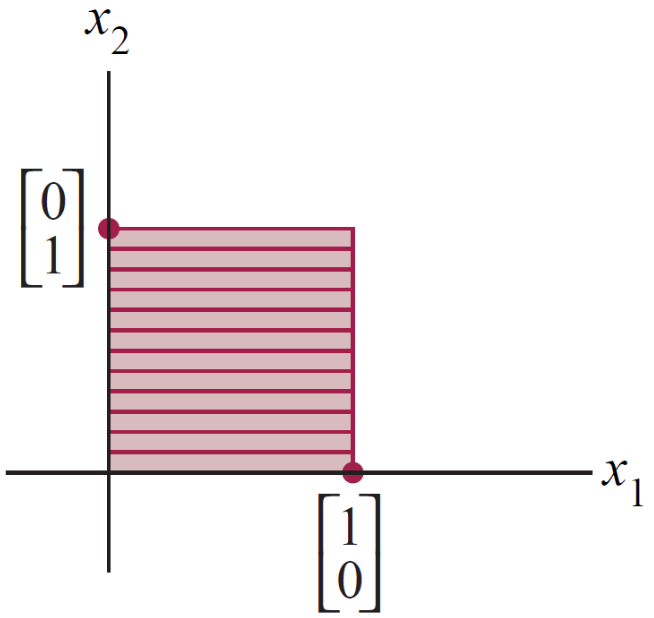
We can think of the standard matrix as a compact way of storing the information about T.

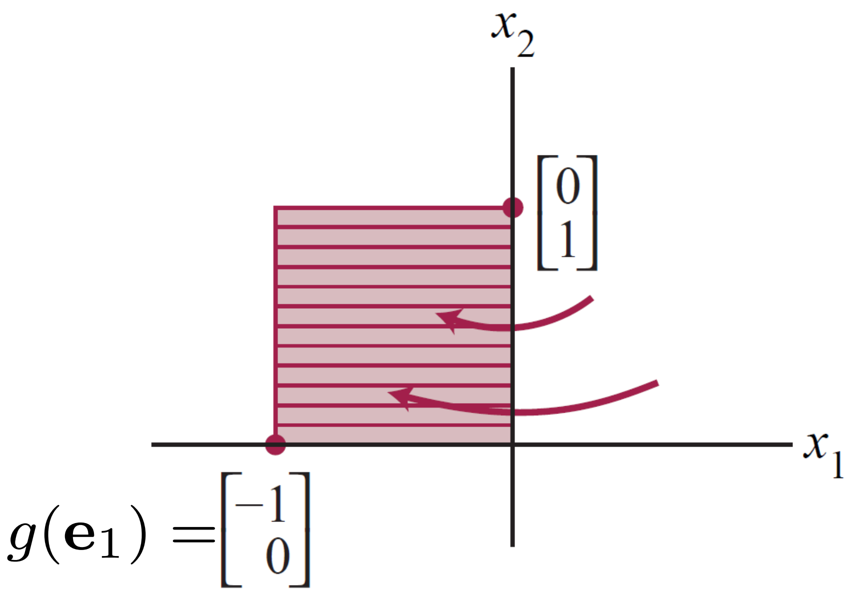
Example: $S : \mathbb{R}^2 \to \mathbb{R}^2$, given by dilation by a factor of 3, $S(\mathbf{x}) = 3\mathbf{x}$.

$$S(\mathbf{e_1}) = S\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = 3\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \quad S(\mathbf{e_2}) = S\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = 3\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

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

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

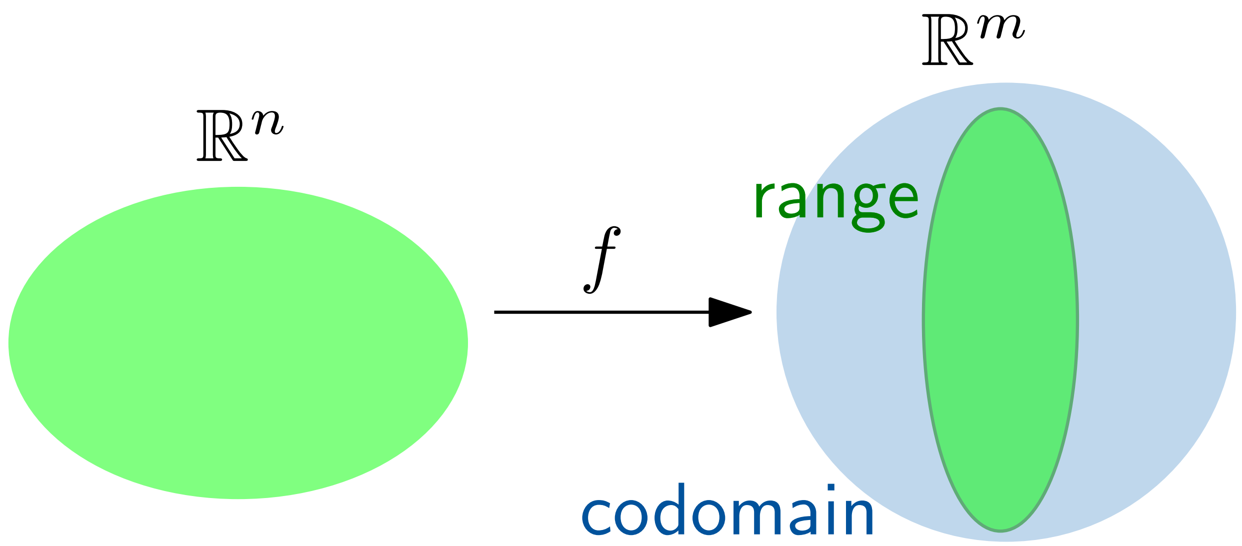
Indeed, $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix}$.

Now we rephrase our existence and uniqueness questions in terms of functions.

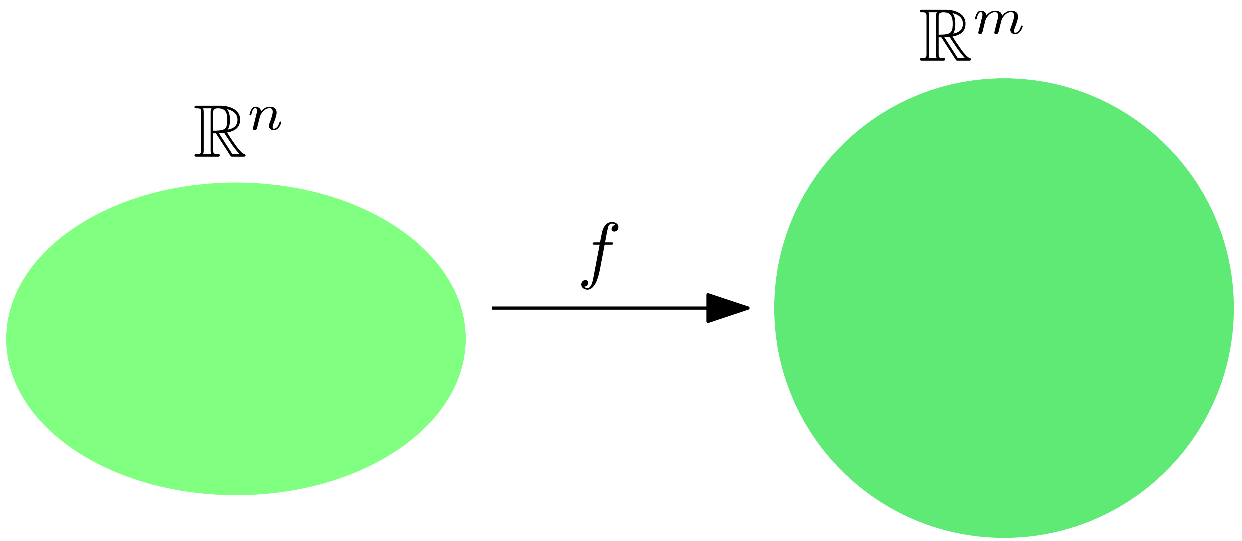
Definition: A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *onto* (surjective) if each $\mathbf{y}$ in $\mathbb{R}^m$ is the image of at least one $\mathbf{x}$ in $\mathbb{R}^n$.

Other ways of saying this:

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



f is not onto



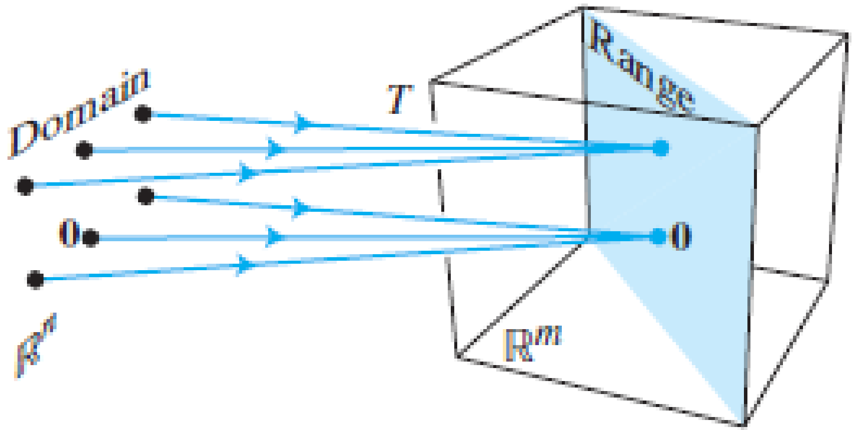
f is onto

Now we rephrase our existence and uniqueness questions in terms of functions.

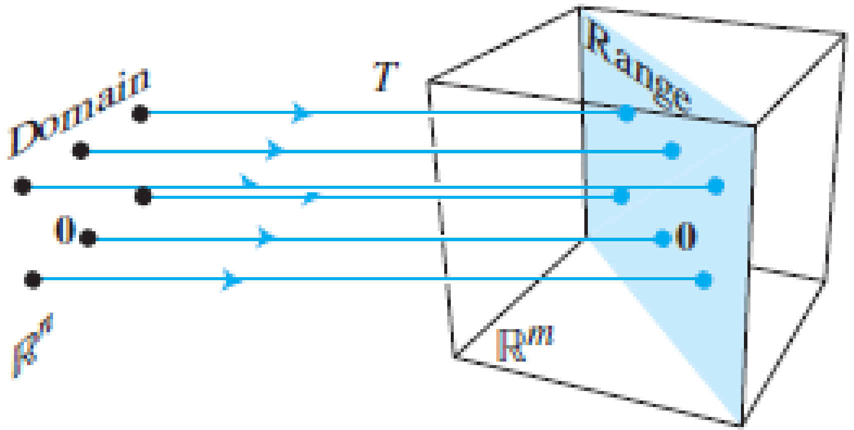
Definition: A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *one-to-one* (injective) if each $\mathbf{y}$ in $\mathbb{R}^m$ is the image of at most one $\mathbf{x}$ in $\mathbb{R}^n$.

Other ways of saying this:

- ??? (something that only works for linear transformations, see p22),
- The equation $f(\mathbf{x}) = \mathbf{y}$ has no solutions or a unique solution.



T is *not* one-to-one



T is one-to-one

Definition: A function $f : \mathbb{R}^n \to \mathbb{R}^m$ 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 $f : \mathbb{R}^n \to \mathbb{R}^m$ 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: $f : \mathbb{R}^2 \to \mathbb{R}^3$, defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_2^3 \\ 2x_1 + x_2 \\ 0 \end{bmatrix}$.

Is f onto? Is f one-to-one?

Definition: A function $f : \mathbb{R}^n \to \mathbb{R}^m$ 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 $f : \mathbb{R}^n \to \mathbb{R}^m$ 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: $f : \mathbb{R}^2 \to \mathbb{R}^3$, defined by $f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_2^3 \\ 2x_1 + x_2 \\ 0 \end{bmatrix}$.

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

f is one-to-one:
if $y_3 \neq 0$, then $f(\mathbf{x}) = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ does not have a solution,

if $y_3 = 0$, then the unique solution to $f(\mathbf{x}) = \begin{bmatrix} y_1 \\ y_2 \\ 0 \end{bmatrix}$ is $x_2 = \sqrt[3]{y_1}$,
$x_1 = \frac{1}{2}(y_2 - x_2) = \frac{1}{2}(y_2 - \sqrt[3]{y_1})$.

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

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

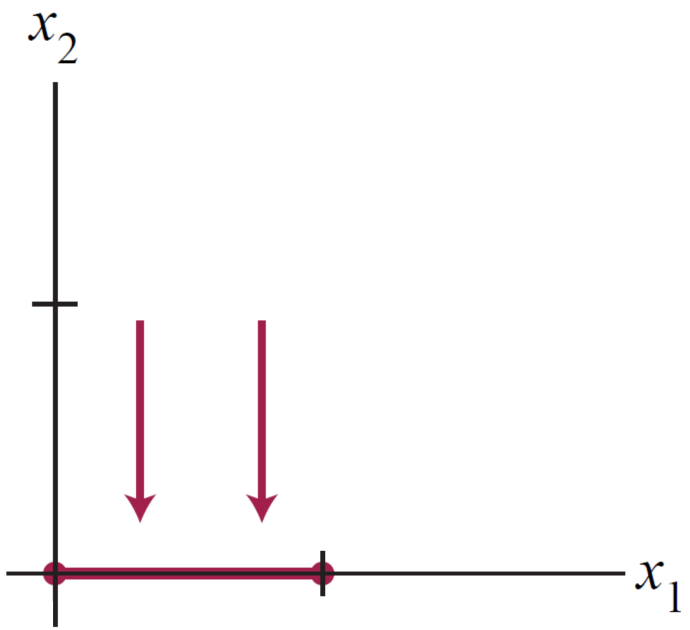
Or, in set notation: $\ker T = \{\mathbf{x} \in \mathbb{R}^n \mid T(\mathbf{x}) = \mathbf{0}\}$.

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

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

Or, in set notation: $\ker T = \{\mathbf{x} \in \mathbb{R}^n \mid T(\mathbf{x}) = \mathbf{0}\}$.

Example: Let T be projection onto the x_1-axis, whose standard matrix is $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ (i.e. $T(\mathbf{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \mathbf{x}$).



The kernel of T is the solution set of $T(\mathbf{x}) = \mathbf{0}$, i.e. the solution set of $\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$. Using the usual algorithm, this solution set is $\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} t \,\middle|\, t \in \mathbb{R} \right\}$, which is the x_2-axis.

It is also clear from the geometric description of projection that the x_2-axis is mapped to the origin.

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

Recall: given $T : \mathbb{R}^n \to \mathbb{R}^m$ a linear transformation, $\ker T = \{\mathbf{x} \in \mathbb{R}^n \mid T(\mathbf{x}) = \mathbf{0}\}$.

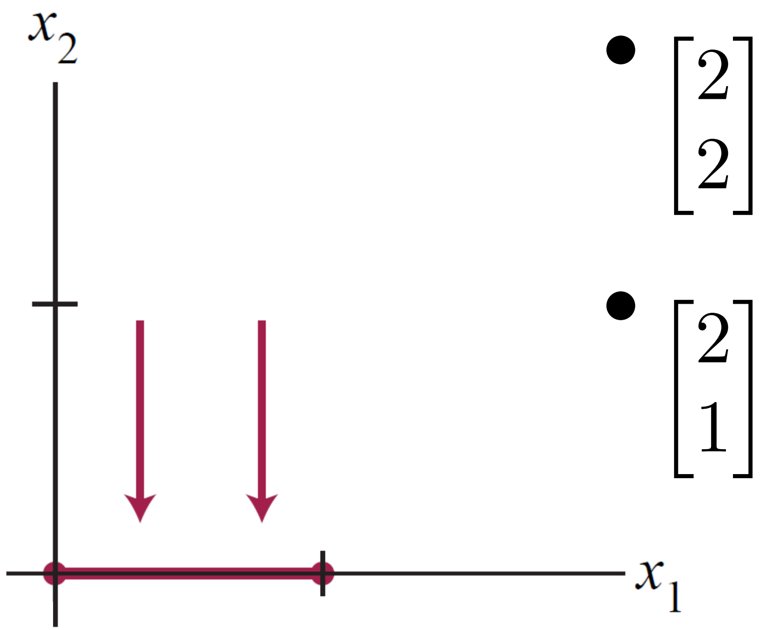
Fact: If $T(\mathbf{v_1}) = T(\mathbf{v_2})$, then $\mathbf{v_1} - \mathbf{v_2}$ is in the kernel of T.

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

The previous page showed that $\ker T$ is the x_2-axis.

Notice that $T\left(\begin{bmatrix}2\\2\end{bmatrix}\right) = T\left(\begin{bmatrix}2\\1\end{bmatrix}\right) = \begin{bmatrix}2\\0\end{bmatrix}$, and

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



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

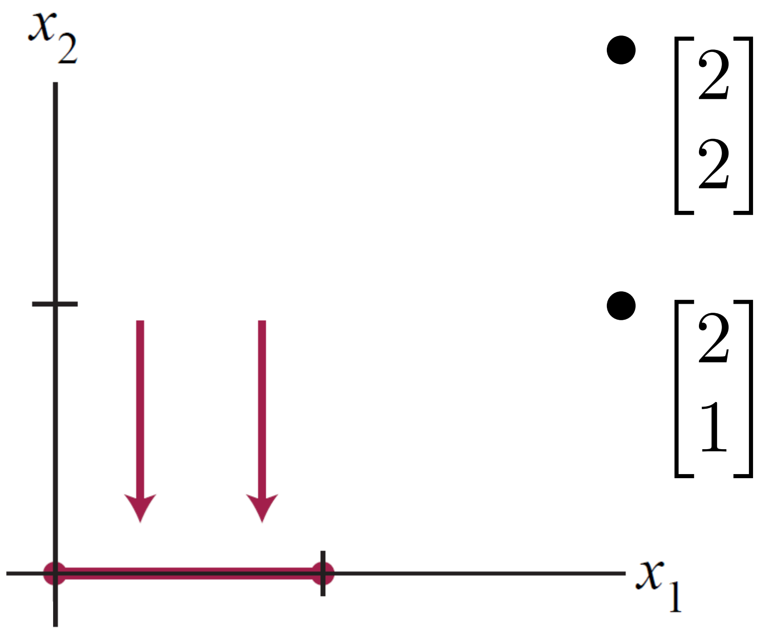
Recall: given $T : \mathbb{R}^n \to \mathbb{R}^m$ a linear transformation, $\ker T = \{\mathbf{x} \in \mathbb{R}^n \mid T(\mathbf{x}) = \mathbf{0}\}$.

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 previous page showed that $\ker T$ is the x_2-axis.

Notice that $T\left(\begin{bmatrix}2\\2\end{bmatrix}\right) = T\left(\begin{bmatrix}2\\1\end{bmatrix}\right) = \begin{bmatrix}2\\0\end{bmatrix}$, and

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



uses linearity of T

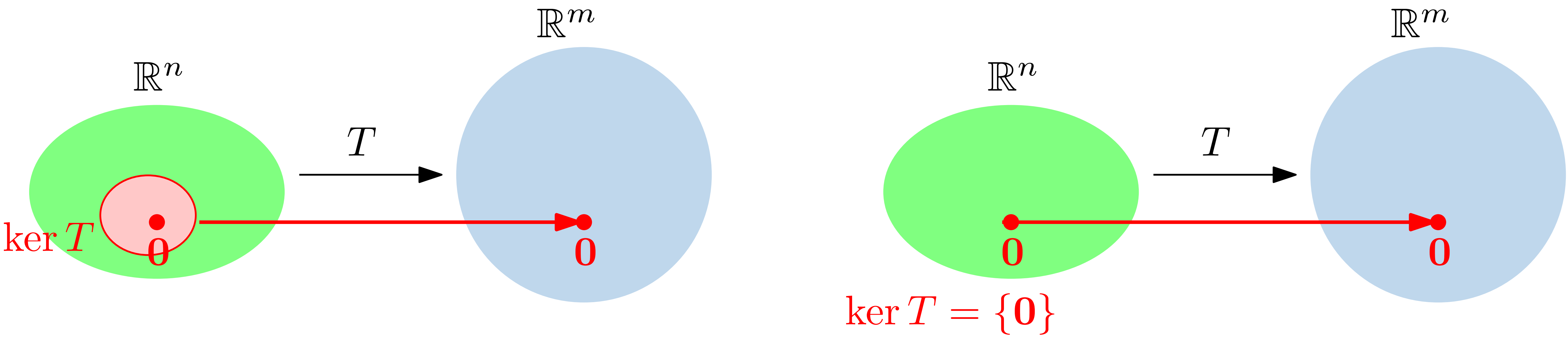
Proof of Fact: If $T(\mathbf{v_1}) = T(\mathbf{v_2}) = \mathbf{y}$, then $T(\mathbf{v_1} - \mathbf{v_2}) = T(\mathbf{v_1}) - T(\mathbf{v_2}) = \mathbf{y} - \mathbf{y} = \mathbf{0}$, so $\mathbf{v_1} - \mathbf{v_2} \in \ker T$.

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

Recall: given $T : \mathbb{R}^n \to \mathbb{R}^m$ a linear transformation, $\ker T = \{\mathbf{x} \in \mathbb{R}^n \mid T(\mathbf{x}) = \mathbf{0}\}$.

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 $\{\mathbf{0}\}$.



T is not one-to-one

T is one-to-one

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

Recall: given $T : \mathbb{R}^n \to \mathbb{R}^m$ a linear transformation, $\ker T = \{\mathbf{x} \in \mathbb{R}^n \mid T(\mathbf{x}) = \mathbf{0}\}$.

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: the theorem is only 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. Taking $\mathbf{y} = \mathbf{0}$ in the definition of one-to-one shows $T(\mathbf{x}) = \mathbf{0}$ has at most one solution. Since $\mathbf{0}$ is a solution (because T is linear), it must be the only one. So its kernel is $\{\mathbf{0}\}$.

Suppose the kernel of T is $\{\mathbf{0}\}$.

T is one-to-one.

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

Recall: given $T : \mathbb{R}^n \to \mathbb{R}^m$ a linear transformation, $\ker T = \{\mathbf{x} \in \mathbb{R}^n \mid T(\mathbf{x}) = \mathbf{0}\}$.

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: the theorem is only 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. Taking $\mathbf{y} = \mathbf{0}$ in the definition of one-to-one shows $T(\mathbf{x}) = \mathbf{0}$ has at most one solution. Since $\mathbf{0}$ is a solution (because T is linear), 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. $\mathbf{v_1} = \mathbf{v_2}$. So if $T(\mathbf{x}) = \mathbf{y}$ has a solution, then the solution is unique, i.e. T is one-to-one.

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

So a matrix transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one if and only if the set of solutions to $A\mathbf{x} = \mathbf{0}$ is $\{\mathbf{0}\}$. This is equivalent to many other things:

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

So a matrix transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one if and only if the set of solutions to $A\mathbf{x} = \mathbf{0}$ is $\{\mathbf{0}\}$. This is equivalent to many other things:

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 (i.e. $\mathbf{x} = \mathbf{0}$ is the only 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. rref(A) has a pivot in every column (i.e. all variables are basic).
e. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
f. The kernel of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is $\{\mathbf{0}\}$.

Notice that e. is in terms of linear transformations, b. is in terms of matrices and linear equations, and they are the same thing.
f. is in terms of linear transformations, a. is in terms of matrices and linear equations, and they are the same thing.

Now let's think about onto and existence of solutions.

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

So, the range of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is the set of $\mathbf{b}$ for which $A\mathbf{x} = \mathbf{b}$ has a solution.

Now let's think about onto and existence of solutions.

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

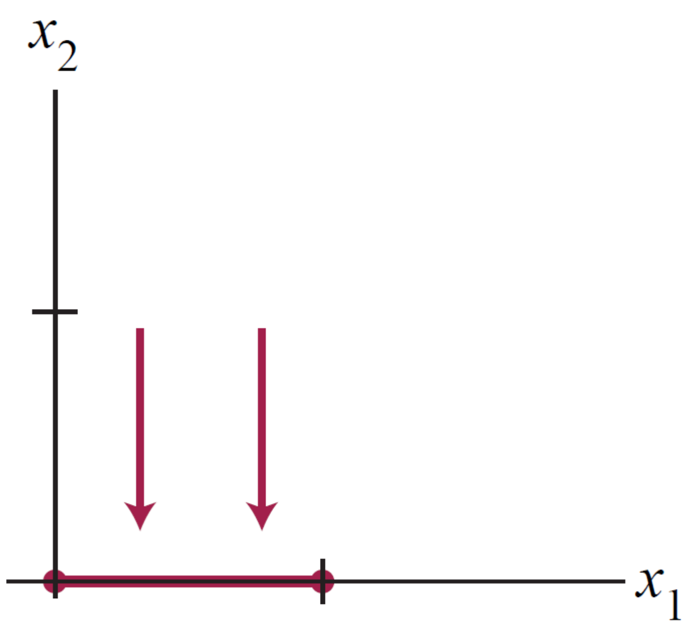
So, the range of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is the set of $\mathbf{b}$ for which $A\mathbf{x} = \mathbf{b}$ has a solution.

So the range of T is the span of the columns of A (see week 2 p17).

Example: Let T be projection onto the x_1-axis, whose standard matrix is $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Its range is the span of the columns of $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, i.e.

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



It is also clear from the geometric description of projection that the set of images is the x_1-axis.

The range of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is the set of $\mathbf{b}$ for which $A\mathbf{x} = \mathbf{b}$ has a solution.
And a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ is onto if and only if its range is all of $\mathbb{R}^m$.
Putting these together: $\mathbf{x} \mapsto A\mathbf{x}$ is onto if and only if $A\mathbf{x} = \mathbf{b}$ is always consistent, and this is equivalent to many things:

The range of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is the set of $\mathbf{b}$ for which $A\mathbf{x} = \mathbf{b}$ has a solution.

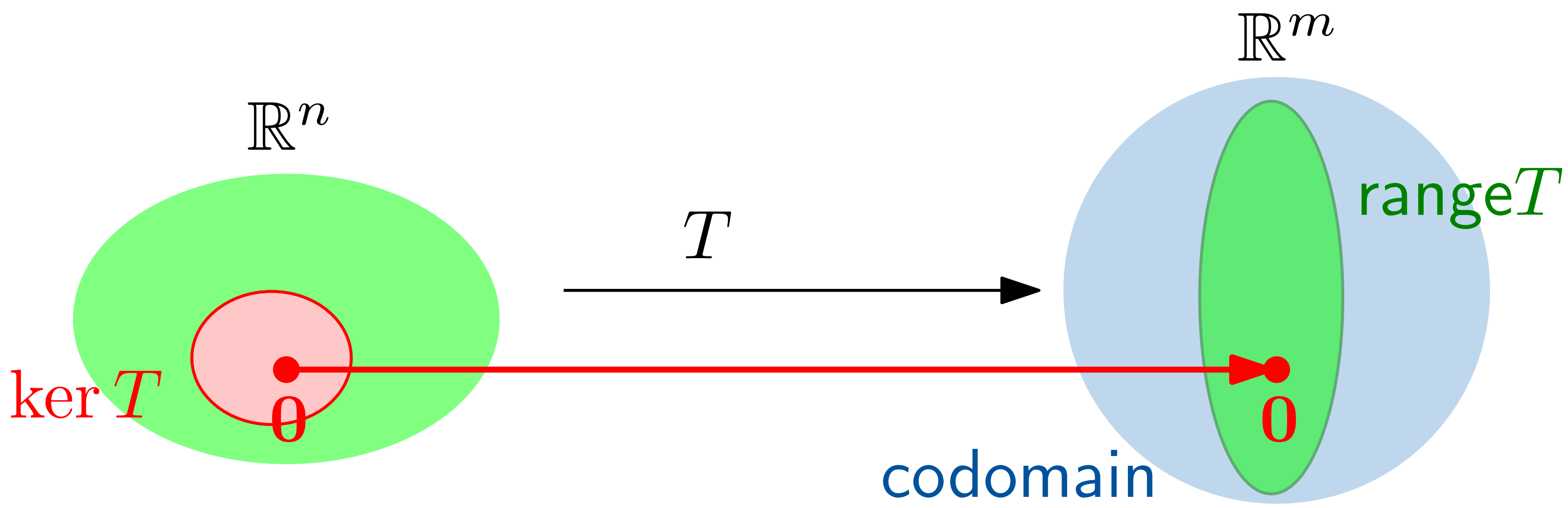
And a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ is onto if and only if its range is all of $\mathbb{R}^m$.

Putting these together: $\mathbf{x} \mapsto A\mathbf{x}$ is onto if and only if $A\mathbf{x} = \mathbf{b}$ is always consistent, and this is equivalent to many things:

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. rref(A) has a pivot in every row.
e. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto.
f. The range of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is $\mathbb{R}^m$.

The range and the kernel on one picture:



Remember from weeks 1-3 that existence and uniqueness are separate, unrelated concepts. Similarly, onto and one-to-one are unrelated:
Exercise 1: think of a linear transformation that is onto but not one-to-one, or both onto and one-to-one, or etc.

Exercise 2: consider the other linear transformations in this week's notes. Are they onto? Are they one-to-one?

Conceptual problems regarding linear independence and linear transformations:

In problems without specific numbers, it's often better not to use row-reduction.
The all-important equation: $T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p)$.

Example: Prove that, if $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent and T is a linear transformation, then $\{T(\mathbf{u}), T(\mathbf{v}), T(\mathbf{w})\}$ is linearly dependent.

Step 1 Rewrite the mathematical terms in the question as formulas.

What we know: there are scalars c_1, c_2, c_3 not all zero with $c_1\mathbf{u} + c_2\mathbf{v} + c_3\mathbf{w} = \mathbf{0}$.
What we want to show: there are scalars d_1, d_2, d_3 not all zero such that $d_1T(\mathbf{u}) + d_2T(\mathbf{v}) + d_3T(\mathbf{w}) = \mathbf{0}$.

Step 2 Fill in the missing steps by rearranging vector equations.

Conceptual problems regarding linear independence and linear transformations:

In problems without specific numbers, it's often better not to use row-reduction.
The all-important equation: $T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p)$.

Example: Prove that, if $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent and T is a linear transformation, then $\{T(\mathbf{u}), T(\mathbf{v}), T(\mathbf{w})\}$ is linearly dependent.

Step 1 Rewrite the mathematical terms in the question as formulas.

What we know: there are scalars c_1, c_2, c_3 not all zero with $c_1\mathbf{u} + c_2\mathbf{v} + c_3\mathbf{w} = \mathbf{0}$.
What we want to show: there are scalars d_1, d_2, d_3 not all zero such that
$d_1T(\mathbf{u}) + d_2T(\mathbf{v}) + d_3T(\mathbf{w}) = \mathbf{0}$.

Step 2 Fill in the missing steps by rearranging vector equations.

Answer: We know $c_1\mathbf{u} + c_2\mathbf{v} + c_3\mathbf{w} = \mathbf{0}$ for some scalars c_1, c_2, c_3 not all zero.
Apply T to both sides: $T(c_1\mathbf{u} + c_2\mathbf{v} + c_3\mathbf{w}) = T(\mathbf{0})$.

Conceptual problems regarding linear independence and linear transformations:

In problems without specific numbers, it's often better not to use row-reduction.
The all-important equation: $T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p)$.

Example: Prove that, if $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent and T is a linear transformation, then $\{T(\mathbf{u}), T(\mathbf{v}), T(\mathbf{w})\}$ is linearly dependent.

Step 1 Rewrite the mathematical terms in the question as formulas.

What we know: there are scalars c_1, c_2, c_3 not all zero with $c_1\mathbf{u} + c_2\mathbf{v} + c_3\mathbf{w} = \mathbf{0}$.
What we want to show: there are scalars d_1, d_2, d_3 not all zero such that $d_1T(\mathbf{u}) + d_2T(\mathbf{v}) + d_3T(\mathbf{w}) = \mathbf{0}$.

Step 2 Fill in the missing steps by rearranging vector equations.

Answer: We know $c_1\mathbf{u} + c_2\mathbf{v} + c_3\mathbf{w} = \mathbf{0}$ for some scalars c_1, c_2, c_3 not all zero.
Apply T to both sides: $T(c_1\mathbf{u} + c_2\mathbf{v} + c_3\mathbf{w}) = T(\mathbf{0})$.
Because T is a linear transformation: $c_1T(\mathbf{u}) + c_2T(\mathbf{v}) + c_3T(\mathbf{w}) = \mathbf{0}$.
Because c_1, c_2, c_3 are not all zero, this is a linear dependence relation among $T(\mathbf{u}), T(\mathbf{v}), T(\mathbf{w})$.