2

Linear Transformations and Matrices

2.1 Opening Remarks

2.1.1 Rotating in 2D

☞ **View at edX**

Let $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be the function that rotates an input vector through an angle θ:
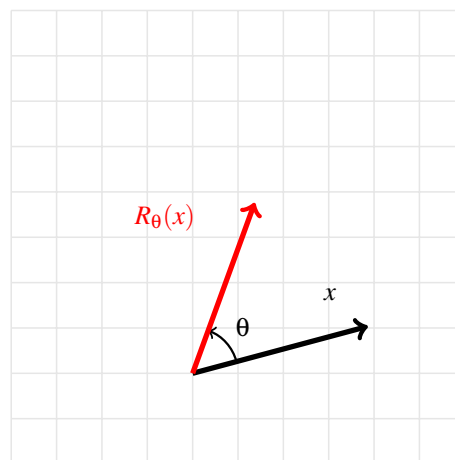

Figure 2.1 illustrates some special properties of the rotation. Functions with these properties are called called linear transformations. Thus, the illustrated rotation in 2D is an example of a linear transformation.
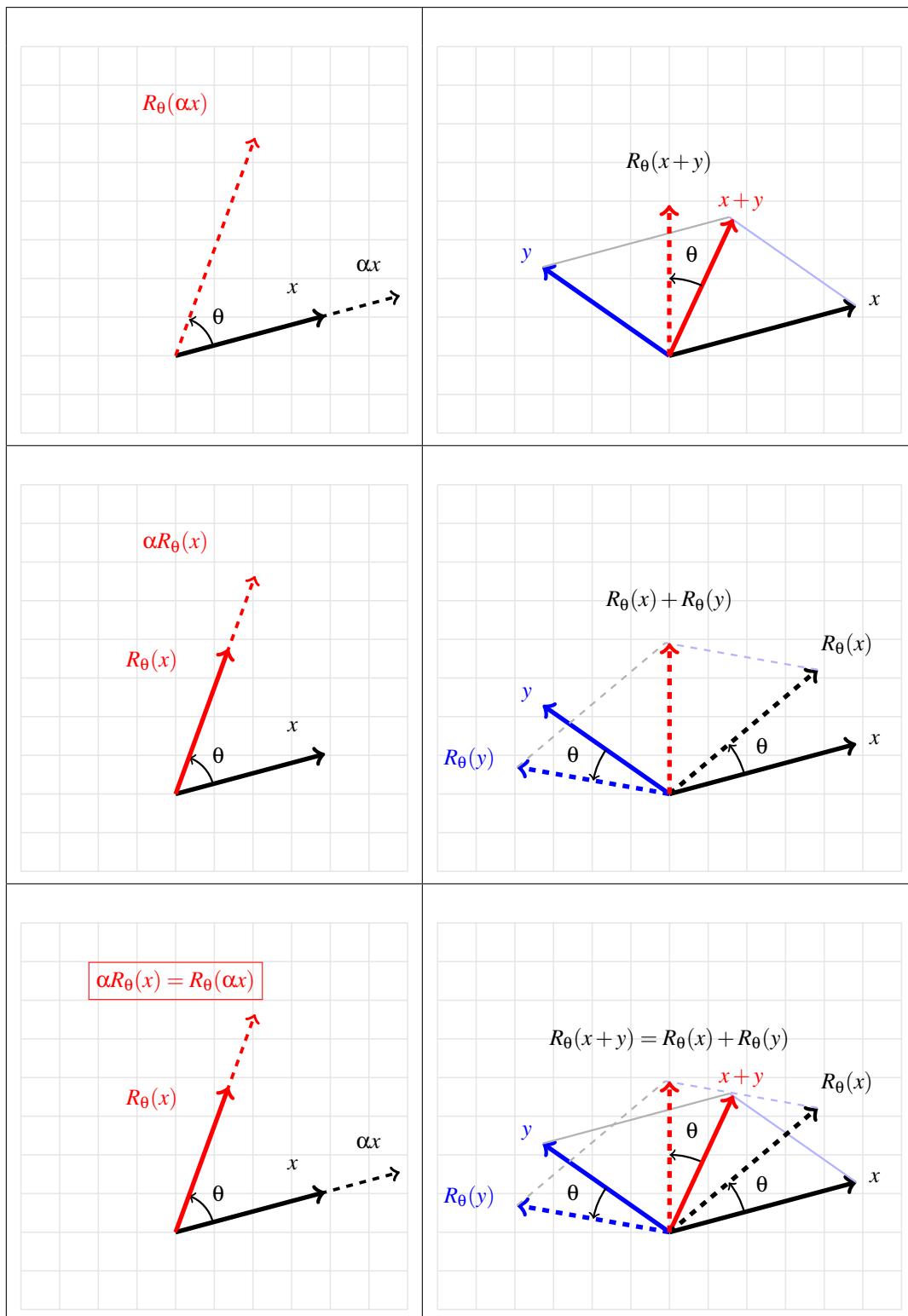
Figure 2.1: The three pictures on the left show that one can scale a vector first and then rotate, or rotate that vector first and then scale and obtain the same result. The three pictures on the right show that one can add two vectors first and then rotate, or rotate the two vectors first and then add and obtain the same result.

Homework 2.1.1.1 A reflection with respect to a 45 degree line is illustrated by
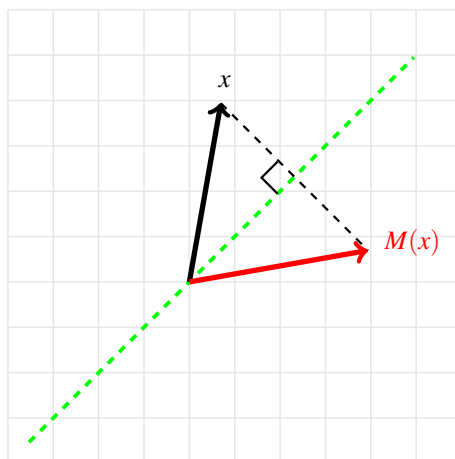
Think of the dashed green line as a mirror and $M : \mathbb{R}^2 \to \mathbb{R}^2$ as the vector function that maps a vector to its mirror image. If $x, y \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$, then $M(\alpha x) = \alpha M(x)$ and $M(x+y) = M(x) + M(y)$ (in other words, M is a linear transformation).

True/False

2.1.2 Outline

2.1.3 What You Will Learn

Upon completion of this unit, you should be able to

- Determine if a given vector function is a linear transformation.

- Identify, visualize, and interpret linear transformations.

- Recognize rotations and reflections in 2D as linear transformations of vectors.

- Relate linear transformations and matrix-vector multiplication.

- Understand and exploit how a linear transformation is completely described by how it transforms the unit basis vectors.

- Find the matrix that represents a linear transformation based on how it transforms unit basis vectors.

- Perform matrix-vector multiplication.

- Reason and develop arguments about properties of linear transformations and matrix vector multiplication.

- Read, appreciate, understand, and develop inductive proofs.
  (Ideally you will fall in love with them! They are beautiful. They don't deceive you. You can count on them. You can build on them. The perfect life companion! But it may not be love at first sight.)

- Make conjectures, understand proofs, and develop arguments about linear transformations.

- Understand the connection between linear transformations and matrix-vector multiplication.

- Solve simple problems related to linear transformations.

2.2 Linear Transformations

2.2.1 What Makes Linear Transformations so Special?

☞ **View at edX**

Many problems in science and engineering involve vector functions such as: $f : \mathbb{R}^n \to \mathbb{R}^m$. Given such a function, one often wishes to do the following:

- Given vector $x \in \mathbb{R}^n$, evaluate $f(x)$; or

- Given vector $y \in \mathbb{R}^m$, find x such that $f(x) = y$; or

- Find scalar λ and vector x such that $f(x) = \lambda x$ (only if $m = n$).

For general vector functions, the last two problems are often especially difficult to solve. As we will see in this course, these problems become a lot easier for a special class of functions called linear transformations.

For those of you who have taken calculus (especially multivariate calculus), you learned that general functions that map vectors to vectors and have special properties can locally be approximated with a linear function. Now, we are not going to discuss what make a function linear, but will just say "it involves linear transformations." (When $m = n = 1$ you have likely seen this when you were taught about "Newton's Method") Thus, even when $f : \mathbb{R}^n \to \mathbb{R}^m$ is not a linear transformation, linear transformations still come into play. This makes understanding linear transformations fundamental to almost all computational problems in science and engineering, just like calculus is.

But calculus is not a prerequisite for this course, so we won't talk about this... :-(

2.2.2 What is a Linear Transformation?

☞ **View at edX**

Definition

Definition 2.1 *A vector function $L : \mathbb{R}^n \to \mathbb{R}^m$ is said to be a linear transformation, if for all $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$*

- *Transforming a scaled vector is the same as scaling the transformed vector:*

$$L(\alpha x) = \alpha L(x)$$

- *Transforming the sum of two vectors is the same as summing the two transformed vectors:*

$$L(x + y) = L(x) + L(y)$$

Examples

Example 2.2 *The transformation $f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) = \begin{pmatrix} \chi_0 + \chi_1 \\ \chi_0 \end{pmatrix}$ is a linear transformation.*

The way we prove this is to pick arbitrary $\alpha \in \mathbb{R}$, $x = \begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}$, and $y = \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix}$ for which we then show that $f(\alpha x) = \alpha f(x)$ and $f(x + y) = f(x) + f(y)$:

- *Show $f(\alpha x) = \alpha f(x)$:*

$$f(\alpha x) = f(\alpha \begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) = f(\begin{pmatrix} \alpha\chi_0 \\ \alpha\chi_1 \end{pmatrix}) = \begin{pmatrix} \alpha\chi_0 + \alpha\chi_1 \\ \alpha\chi_0 \end{pmatrix} = \begin{pmatrix} \alpha(\chi_0 + \chi_1) \\ \alpha\chi_0 \end{pmatrix}$$

and

$$\alpha f(x) = \alpha f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) = \alpha \begin{pmatrix} \chi_0 + \chi_1 \\ \chi_0 \end{pmatrix} = \begin{pmatrix} \alpha(\chi_0 + \chi_1) \\ \alpha\chi_0 \end{pmatrix}$$

Both $f(\alpha x)$ and $\alpha f(x)$ evaluate to the same expression. One can then make this into one continuous sequence of equivalences by rewriting the above as

$$
\begin{aligned}
f(\alpha x) \quad &= \quad f(\alpha \begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) = f(\begin{pmatrix} \alpha\chi_0 \\ \alpha\chi_1 \end{pmatrix}) = \begin{pmatrix} \alpha\chi_0 + \alpha\chi_1 \\ \alpha\chi_0 \end{pmatrix} \\
&= \quad \begin{pmatrix} \alpha(\chi_0 + \chi_1) \\ \alpha\chi_0 \end{pmatrix} = \alpha \begin{pmatrix} \chi_0 + \chi_1 \\ \chi_0 \end{pmatrix} = \alpha f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) = \alpha f(x).
\end{aligned}
$$

- *Show $f(x+y) = f(x) + f(y)$:*

$$f(x+y) = f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix} + \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix}) = f(\begin{pmatrix} \chi_0 + \psi_0 \\ \chi_1 + \psi_1 \end{pmatrix}) = \begin{pmatrix} (\chi_0 + \psi_0) + (\chi_1 + \psi_1) \\ \chi_0 + \psi_0 \end{pmatrix}$$

and

$$
\begin{aligned}
f(x) + f(y). \quad &= \quad f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) + f(\begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix}) = \begin{pmatrix} \chi_0 + \chi_1 \\ \chi_0 \end{pmatrix} + \begin{pmatrix} \psi_0 + \psi_1 \\ \psi_0 \end{pmatrix} \\
&= \quad \begin{pmatrix} (\chi_0 + \chi_1) + (\psi_0 + \psi_1) \\ \chi_0 + \psi_0 \end{pmatrix}.
\end{aligned}
$$

Both $f(x+y)$ and $f(x) + f(y)$ evaluate to the same expression since scalar addition is commutative and associative. The above observations can then be rearranged into the sequence of equivalences

$$
\begin{aligned}
f(x+y) \quad &= \quad f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix} + \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix}) = f(\begin{pmatrix} \chi_0 + \psi_0 \\ \chi_1 + \psi_1 \end{pmatrix}) \\
&= \quad \begin{pmatrix} (\chi_0 + \psi_0) + (\chi_1 + \psi_1) \\ \chi_0 + \psi_0 \end{pmatrix} = \begin{pmatrix} (\chi_0 + \chi_1) + (\psi_0 + \psi_1) \\ \chi_0 + \psi_0 \end{pmatrix} \\
&= \quad \begin{pmatrix} \chi_0 + \chi_1 \\ \chi_0 \end{pmatrix} + \begin{pmatrix} \psi_0 + \psi_1 \\ \psi_0 \end{pmatrix} = f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) + f(\begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix}) = f(x) + f(y).
\end{aligned}
$$

Example 2.3 *The transformation $f(\begin{pmatrix} \chi \\ \psi \end{pmatrix}) = \begin{pmatrix} \chi + \psi \\ \chi + 1 \end{pmatrix}$ is not a linear transformation.*

*We will start by trying a few scalars α and a few vectors x and see whether $f(\alpha x) = \alpha f(x)$. If we find even **one** example such that $f(\alpha x) \neq f(\alpha x)$ then we have proven that f is not a linear transformation. Likewise, if we find even **one** pair of vectors x and y such that $f(x+y) \neq f(x) + f(y)$ then we have done the same.*

A vector function $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if *for all* scalars α and *for all* vectors $x, y \in \mathbb{R}^n$ it is that case that

- $f(\alpha x) = \alpha f(x)$ and

- $f(x+y) = f(x) + f(y)$.

If there is even one scalar α and vector $x \in \mathbb{R}^n$ such that $f(\alpha x) \neq \alpha f(x)$ or if there is even one pair of vectors $x, y \in \mathbb{R}^n$ such that $f(x+y) \neq f(x) + f(y)$, then the vector function f is *not* a linear transformation. Thus, in order to show that a vector function f is *not* a linear transformation, it suffices to find one such counter example.

Now, let us try a few:

- *Let $\alpha = 1$ and* $\begin{pmatrix} \chi \\ \psi \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. *Then*

$$f(\alpha \begin{pmatrix} \chi \\ \psi \end{pmatrix}) = f(1 \times \begin{pmatrix} 1 \\ 1 \end{pmatrix}) = f(\begin{pmatrix} 1 \\ 1 \end{pmatrix}) = \begin{pmatrix} 1+1 \\ 1+1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

and

$$\alpha f(\begin{pmatrix} \chi \\ \psi \end{pmatrix}) = 1 \times f(\begin{pmatrix} 1 \\ 1 \end{pmatrix}) = 1 \times \begin{pmatrix} 1+1 \\ 1+1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

For this example, $f(\alpha x) = \alpha f(x)$, but there may still be an example such that $f(\alpha x) \neq \alpha f(x)$.

- *Let $\alpha = 0$ and* $\begin{pmatrix} \chi \\ \psi \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. *Then*

$$f(\alpha \begin{pmatrix} \chi \\ \psi \end{pmatrix}) = f(0 \times \begin{pmatrix} 1 \\ 1 \end{pmatrix}) = f(\begin{pmatrix} 0 \\ 0 \end{pmatrix}) = \begin{pmatrix} 0+0 \\ 0+1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and

$$\alpha f(\begin{pmatrix} \chi \\ \psi \end{pmatrix}) = 0 \times f(\begin{pmatrix} 1 \\ 1 \end{pmatrix}) = 0 \times \begin{pmatrix} 1+1 \\ 1+1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

For this example, we have found a case where $f(\alpha x) \neq \alpha f(x)$. Hence, the function is not a linear transformation.

Homework 2.2.2.1 The vector function $f(\begin{pmatrix} \chi \\ \psi \end{pmatrix}) = \begin{pmatrix} \chi\psi \\ \chi \end{pmatrix}$ is a linear transformation.

TRUE/FALSE

Homework 2.2.2.2 The vector function $f(\begin{pmatrix} \chi_0 \\ \chi_1 \\ \chi_2 \end{pmatrix}) = \begin{pmatrix} \chi_0 + 1 \\ \chi_1 + 2 \\ \chi_2 + 3 \end{pmatrix}$ is a linear transformation. (This is the same function as in Homework 1.4.6.1.)

TRUE/FALSE

Homework 2.2.2.3 The vector function $f(\begin{pmatrix} \chi_0 \\ \chi_1 \\ \chi_2 \end{pmatrix}) = \begin{pmatrix} \chi_0 \\ \chi_0 + \chi_1 \\ \chi_0 + \chi_1 + \chi_2 \end{pmatrix}$ is a linear transformation. (This is the same function as in Homework 1.4.6.2.)

TRUE/FALSE

Homework 2.2.2.4 If $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then $L(0) = 0$.
(Recall that 0 equals a vector with zero components of appropriate size.)

Always/Sometimes/Never

Homework 2.2.2.5 Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and $f(0) \neq 0$. Then f is not a linear transformation.

True/False

Homework 2.2.2.6 Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and $f(0) = 0$. Then f is a linear transformation.

Always/Sometimes/Never

Homework 2.2.2.7 Find an example of a function f such that $f(\alpha x) = \alpha f(x)$, but for some x, y it is the case that $f(x + y) \neq f(x) + f(y)$. (This is pretty tricky!)

Homework 2.2.2.8 The vector function $f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) = \begin{pmatrix} \chi_1 \\ \chi_0 \end{pmatrix}$ is a linear transformation.

TRUE/FALSE

2.2.3   Of Linear Transformations and Linear Combinations

Now that we know what a linear transformation and a linear combination of vectors are, we are ready to start making the connection between the two with matrix-vector multiplication.

Lemma 2.4 $L : \mathbb{R}^n \to \mathbb{R}^m$ *is a linear transformation if and only if (iff) for all $u, v \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$*

$$L(\alpha u + \beta v) = \alpha L(u) + \beta L(v).$$

Proof:

($\Rightarrow$) Assume that $L : R^n \to R^m$ is a linear transformation and let $u, v \in \mathbb{R}^n$ be *arbitrary* vectors and $\alpha, \beta \in \mathbb{R}$ be *arbitrary* scalars. Then

$\quad\quad L(\alpha u + \beta v)$

$=$ $<$since αu and βv are vectors and L is a linear transformation $>$

$\quad\quad L(\alpha u) + L(\beta v)$

$=$ $<$ since L is a linear transformation $>$

$\quad\quad \alpha L(u) + \beta L(v)$

($\Leftarrow$) Assume that for all $u, v \in \mathbb{R}^n$ and all $\alpha, \beta \in \mathbb{R}$ it is the case that $L(\alpha u + \beta v) = \alpha L(u) + \beta L(v)$. We need to show that

- $L(\alpha u) = \alpha L(u)$.
  This follows immediately by setting $\beta = 0$.

- $L(u + v) = L(u) + L(v)$.
  This follows immediately by setting $\alpha = \beta = 1$.

☞ **View at edX**

Lemma 2.5 *Let $v_0, v_1, \ldots, v_{k-1} \in \mathbb{R}^n$ and let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then*

$$L(v_0 + v_1 + \ldots + v_{k-1}) = L(v_0) + L(v_1) + \ldots + L(v_{k-1}). \tag{2.1}$$

While it is tempting to say that this is simply obvious, we are going to prove this rigorously. When one tries to prove a result for a general k, where k is a natural number, one often uses a "proof by induction". We are going to give the proof first, and then we will explain it.

Proof: Proof by induction on k.

Base case: $k = 1$. For this case, we must show that $L(v_0) = L(v_0)$. This is trivially true.

Inductive step: Inductive Hypothesis (IH): Assume that the result is true for $k = K$ where $K \geq 1$:

$$L(v_0 + v_1 + \ldots + v_{K-1}) = L(v_0) + L(v_1) + \ldots + L(v_{K-1}).$$

We will show that the result is **then** also true for $k = K + 1$. In other words, that

$$L(v_0 + v_1 + \ldots + v_{K-1} + v_K) = L(v_0) + L(v_1) + \ldots + L(v_{K-1}) + L(v_K).$$

$L(v_0 + v_1 + \ldots + v_K)$

$=$ $<$ expose extra term – We know we can do this, since $K \geq 1 >$

$L(v_0 + v_1 + \ldots + v_{K-1} + v_K)$

$=$ $<$ associativity of vector addition $>$

$L((v_0 + v_1 + \ldots + v_{K-1}) + v_K)$

$=$ $< L$ is a linear transformation) $>$

$L(v_0 + v_1 + \ldots + v_{K-1}) + L(v_K)$

$=$ $<$ Inductive Hypothesis $>$

$L(v_0) + L(v_1) + \ldots + L(v_{K-1}) + L(v_K)$

By the Principle of Mathematical Induction the result holds for all k.

The idea is as follows:

- The base case shows that the result is true for $k = 1$: $L(v_0) = L(v_0)$.

- The inductive step shows that if the result is true for $k = 1$, then the result is true for $k = 1 + 1 = 2$ so that $L(v_0 + v_1) = L(v_0) + L(v_1)$.

- Since the result is indeed true for $k = 1$ (as proven by the base case) we now know that the result is also true for $k = 2$.

- The inductive step also implies that if the result is true for $k = 2$, then it is also true for $k = 3$.

- Since we just reasoned that it is true for $k = 2$, we now know it is also true for $k = 3$: $L(v_0 + v_1 + v_2) = L(v_0) + L(v_1) + L(v_2)$.

- And so forth.

2.3 Mathematical Induction

2.3.1 What is the Principle of Mathematical Induction?

☞ **View at edX**

The **Principle of Mathematical Induction** (weak induction) says that *if* one can show that

- (Base case) a property holds for $k = k_b$; *and*

- (Inductive step) if it holds for $k = K$, where $K \geq k_b$, then it is also holds for $k = K + 1$,

then one can conclude that the property holds for all integers $k \geq k_b$. Often $k_b = 0$ or $k_b = 1$.

If mathematical induction intimidates you, have a look at in the enrichment for this week (Section 2.5.2) :Puzzles and Paradoxes in Mathematical Induction", by Adam Bjorndahl.

Here is Maggie's take on Induction, extending it beyond the proofs we do.

If you want to prove something holds for all members of a set that can be defined inductively, then you would use mathematical induction. You may recall a set is a collection and as such the order of its members is not important. However, some sets do have a natural ordering that can be used to describe the membership. This is especially valuable when the set has an infinite number of members, for example, natural numbers. Sets for which the membership can be described by suggesting there is a first element (or small group of firsts) then from this first you can create another (or others) then more and more by applying a rule to get another element in the set are our focus here. If all elements (members) are in the set because they are either the first (basis) or can be constructed by applying "The" rule to the first (basis) a finite number of times, then the set can be inductively defined.

So for us, the set of natural numbers is inductively defined. As a computer scientist you would say 0 is the first and the rule is to add one to get another element. So $0, 1, 2, 3, \ldots$ are members of the natural numbers. In this way, 10 is a member of natural numbers because you can find it by adding 1 to 0 ten times to get it.

So, the Principle of Mathematical induction proves that something is true for all of the members of a set that can be defined inductively. If this set has an infinite number of members, you couldn't show it is true for each of them individually. The idea is if it is true for the first(s) and it is true for any constructed member(s) no matter where you are in the list, it must be true for all. Why? Since we are proving things about natural numbers, the idea is if it is true for 0 and the next constructed, it must be true for 1 but then its true for 2, and then 3 and 4 and 5 ...and 10 and ... and 10000 and 10001 , etc (all natural numbers). This is only because of the special ordering we can put on this set so we can know there is a next one for which it must be true. People often picture this rule by thinking of climbing a ladder or pushing down dominoes. If you know you started and you know where ever you are the next will follow then you must make it through all (even if there are an infinite number).

That is why to prove something using the Principle of Mathematical Induction you must show what you are proving holds at a start and then if it holds (assume it holds up to some point) then it holds for the next constructed element in the set. With these two parts shown, we know it must hold for all members of this inductively defined set.

You can find many examples of how to prove using PMI as well as many examples of when and why this method of proof will fail all over the web. Notice it only works for statements about sets "that can be defined inductively". Also notice subsets of natural numbers can often be defined inductively. For example, if I am a mathematician I may start counting at 1. Or I may decide that the statement holds for natural numbers ≥ 4 so I start my base case at 4.

My last comment in this very long message is that this style of proof extends to other structures that can be defined inductively (such as trees or special graphs in CS).

2.3.2 Examples

☞ **View at edX**
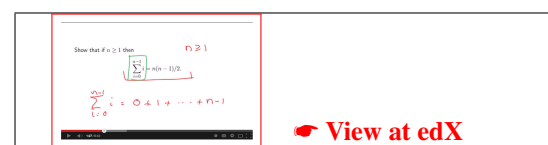
Later in this course, we will look at the cost of various operations that involve matrices and vectors. In the analyses, we will often encounter a cost that involves the expression $\sum_{i=0}^{n-1} i$. We will now show that

$$\sum_{i=0}^{n-1} i = n(n-1)/2.$$

Proof:

Base case: $n = 1$. For this case, we must show that $\sum_{i=0}^{1-1} i = 1(0)/2$.

$\sum_{i=0}^{1-1} i$

= $\qquad\qquad\qquad\qquad\qquad$ < Definition of summation>

0

= $\qquad\qquad\qquad\qquad\qquad$ < arithmetic>

$1(0)/2$

This proves the base case.

Inductive step: Inductive Hypothesis (IH): Assume that the result is true for $n = k$ where $k \geq 1$:

$$\sum_{i=0}^{k-1} i = k(k-1)/2.$$

We will show that the result is then also true for $n = k+1$:

$$\sum_{i=0}^{(k+1)-1} i = (k+1)((k+1)-1)/2.$$

Assume that $k \geq 1$. Then

$\sum_{i=0}^{(k+1)-1} i$

= $\qquad\qquad\qquad\qquad\qquad$ < arithmetic>

$\sum_{i=0}^{k} i$

= $\qquad\qquad\qquad\qquad\qquad$ < split off last term>

$\sum_{i=0}^{k-1} i + k$

= $\qquad\qquad\qquad\qquad\qquad$ < I.H.>

$k(k-1)/2 + k.$

= $\qquad\qquad\qquad\qquad\qquad$ < algebra>

$(k^2 - k)/2 + 2k/2.$

= $\qquad\qquad\qquad\qquad\qquad$ < algebra>

$(k^2 + k)/2.$

= $\qquad\qquad\qquad\qquad\qquad$ < algebra>

$(k+1)k/2.$

= $\qquad\qquad\qquad\qquad\qquad$ < arithmetic>

$(k+1)((k+1)-1)/2.$

This proves the inductive step.

By the Principle of Mathematical Induction the result holds for all n.

As we become more proficient, we will start combining steps. For now, we give lots of detail to make sure everyone stays on board.
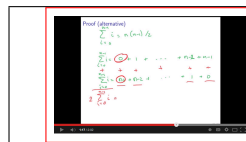
There is an alternative proof for this result which does not involve mathematical induction. We give this proof now because it is a convenient way to rederive the result should you need it in the future.

Proof:(alternative)

$$
\begin{array}{ccccccccccc}
\sum_{i=0}^{n-1} i & = & 0 & + & 1 & + & \cdots & + & (n-2) & + & (n-1) \\
\sum_{i=0}^{n-1} i & = & (n-1) & + & (n-2) & + & \cdots & + & 1 & + & 0 \\
\hline
2\sum_{i=0}^{n-1} i & = & \underbrace{(n-1) & + & (n-1) & + & \cdots & + & (n-1) & + & (n-1)} \\
\end{array}
$$
$$n \text{ times the term } (n-1)$$

so that $2\sum_{i=0}^{n-1} i = n(n-1)$. Hence $\sum_{i=0}^{n-1} i = n(n-1)/2$.

For those who don't like the "$\cdots$" in the above argument, notice that

$$
\begin{array}{rcll}
2\sum_{i=0}^{n-1} i & = & \sum_{i=0}^{n-1} i + \sum_{j=0}^{n-1} j & < \text{algebra} > \\
& = & \sum_{i=0}^{n-1} i + \sum_{j=n-1}^{0} j & < \text{reverse the order of the summation} > \\
& = & \sum_{i=0}^{n-1} i + \sum_{i=0}^{n-1} (n-i-1) & < \text{substituting } j = n-i-1 > \\
& = & \sum_{i=0}^{n-1} (i+n-i-1) & < \text{merge sums} > \\
& = & \sum_{i=0}^{n-1} (n-1) & < \text{algebra} > \\
& = & n(n-1) & < (n-1) \text{ is summed } n \text{ times} >.
\end{array}
$$

Hence $\sum_{i=0}^{n-1} i = n(n-1)/2$.

Homework 2.3.2.1 Let $n \geq 1$. Then $\sum_{i=1}^{n} i = n(n+1)/2$.

Always/Sometimes/Never

Homework 2.3.2.2 Let $n \geq 1$. $\sum_{i=0}^{n-1} 1 = n$.

Always/Sometimes/Never

Homework 2.3.2.3 Let $n \geq 1$ and $x \in \mathbb{R}^m$. Then

$$\sum_{i=0}^{n-1} x = \underbrace{x + x + \cdots + x}_{n \text{ times}} = nx$$

Always/Sometimes/Never

Homework 2.3.2.4 Let $n \geq 1$. $\sum_{i=0}^{n-1} i^2 = (n-1)n(2n-1)/6$.

Always/Sometimes/Never

2.4 Representing Linear Transformations as Matrices

2.4.1 From Linear Transformation to Matrix-Vector Multiplication

☞ View at edX

Theorem 2.6 *Let* $v_o, v_1, \ldots, v_{n-1} \in \mathbb{R}^n$, $\alpha_o, \alpha_1, \ldots, \alpha_{n-1} \in \mathbb{R}$, *and let* $L : \mathbb{R}^n \to \mathbb{R}^m$ *be a linear transformation. Then*

$$L(\alpha_0 v_0 + \alpha_1 v_1 + \cdots + \alpha_{n-1} v_{n-1}) = \alpha_0 L(v_0) + \alpha_1 L(v_1) + \cdots + \alpha_{n-1} L(v_{n-1}). \tag{2.2}$$

Proof:

$L(\alpha_0 v_0 + \alpha_1 v_1 + \cdots + \alpha_{n-1} v_{n-1})$

$=$ $\qquad\qquad\qquad\qquad$ $<$ Lemma 2.5: $L(v_0 + \cdots + v_{n-1}) = L(v_0) + \cdots + L(v_{n-1}) >$

$L(\alpha_0 v_0) + L(\alpha_1 v_1) + \cdots + L(\alpha_{n-1} v_{n-1})$

$=$ $\qquad\qquad\qquad\qquad$ $<$Definition of linear transformation, n times $>$

$\alpha_0 L(v_0) + \alpha_1 L(v_1) + \cdots + \alpha_{k-1} L(v_{k-1}) + \alpha_{n-1} L(v_{n-1}).$

Homework 2.4.1.1 Give an alternative proof for this theorem that mimics the proof by induction for the lemma that states that $L(v_0 + \cdots + v_{n-1}) = L(v_0) + \cdots + L(v_{n-1})$.

Homework 2.4.1.2 Let L be a linear transformation such that

$$L\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \quad \text{and} \quad L\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Then $L\left(\begin{pmatrix} 2 \\ 3 \end{pmatrix}\right) =$

For the next three exercises, let L be a linear transformation such that

$$L\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \quad \text{and} \quad L\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 5 \\ 4 \end{pmatrix}.$$

Homework 2.4.1.3 $L\left(\begin{pmatrix} 3 \\ 3 \end{pmatrix}\right) =$

Homework 2.4.1.4 $L\left(\begin{pmatrix} -1 \\ 0 \end{pmatrix}\right) =$

Homework 2.4.1.5 $L\left(\begin{pmatrix} 2 \\ 3 \end{pmatrix}\right) =$

Homework 2.4.1.6 Let L be a linear transformation such that

$$L(\begin{pmatrix} 1 \\ 1 \end{pmatrix}) = \begin{pmatrix} 5 \\ 4 \end{pmatrix}.$$

Then $L(\begin{pmatrix} 3 \\ 2 \end{pmatrix}) =$

Homework 2.4.1.7 Let L be a linear transformation such that

$$L(\begin{pmatrix} 1 \\ 1 \end{pmatrix}) = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \quad \text{and} \quad L(\begin{pmatrix} 2 \\ 2 \end{pmatrix}) = \begin{pmatrix} 10 \\ 8 \end{pmatrix}.$$

Then $L(\begin{pmatrix} 3 \\ 2 \end{pmatrix}) =$

Now we are ready to link linear transformations to matrices and matrix-vector multiplication.

Recall that any vector $x \in \mathbb{R}^n$ can be written as

$$x = \begin{pmatrix} \chi_0 \\ \chi_1 \\ \vdots \\ \chi_{n-1} \end{pmatrix} = \chi_0 \underbrace{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{e_0} + \chi_1 \underbrace{\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}}_{e_1} + \cdots + \chi_{n-1} \underbrace{\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}}_{e_{n-1}} = \sum_{j=0}^{n-1} \chi_j e_j.$$

Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Given $x \in \mathbb{R}^n$, the result of $y = L(x)$ is a vector in $\mathbb{R}^m$. But then

$$y = L(x) = L\left(\sum_{j=0}^{n-1} \chi_j e_j.\right) = \sum_{j=0}^{n-1} \chi_j L(e_j) = \sum_{j=0}^{n-1} \chi_j a_j,$$

where we let $a_j = L(e_j)$.

The Big Idea. The linear transformation L is completely described by the vectors

$$a_0, a_1, \ldots, a_{n-1}, \quad \text{where } a_j = L(e_j)$$

because for any vector x, $L(x) = \sum_{j=0}^{n-1} \chi_j a_j$.

By arranging these vectors as the columns of a two-dimensional array, which we call the matrix A, we arrive at the observation that the matrix is simply a representation of the corresponding linear transformation L.

Homework 2.4.1.8 Give the matrix that corresponds to the linear transformation $f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) = \begin{pmatrix} 3\chi_0 - \chi_1 \\ \chi_1 \end{pmatrix}.$

Homework 2.4.1.9 Give the matrix that corresponds to the linear transformation $f(\begin{pmatrix} \chi_0 \\ \chi_1 \\ \chi_2 \end{pmatrix}) = \begin{pmatrix} 3\chi_0 - \chi_1 \\ \chi_2 \end{pmatrix}.$

If we let

$$
A = \begin{pmatrix}
\alpha_{0,0} & \alpha_{0,1} & \cdots & \alpha_{0,n-1} \\
\alpha_{1,0} & \alpha_{1,1} & \cdots & \alpha_{1,n-1} \\
\vdots & \vdots & \ddots & \vdots \\
\alpha_{m-1,0} & \alpha_{m-1,1} & \cdots & \alpha_{m-1,n-1}
\end{pmatrix}
$$

$$
\underbrace{}_{a_0} \quad \underbrace{}_{a_1} \quad \underbrace{}_{a_{n-1}}
$$

so that $\alpha_{i,j}$ equals the ith component of vector a_j, then

$$
\begin{aligned}
L(x) &= L(\sum_{j=0}^{n-1}\chi_j e_j) = \sum_{j=0}^{n-1} L(\chi_j e_j) = \sum_{j=0}^{n-1}\chi_j L(e_j) = \sum_{j=0}^{n-1}\chi_j a_j \\
&= \chi_0 a_0 + \chi_1 a_1 + \cdots + \chi_{n-1} a_{n-1} \\
&= \chi_0 \begin{pmatrix}\alpha_{0,0}\\ \alpha_{1,0}\\ \vdots\\ \alpha_{m-1,0}\end{pmatrix} + \chi_1 \begin{pmatrix}\alpha_{0,1}\\ \alpha_{1,1}\\ \vdots\\ \alpha_{m-1,1}\end{pmatrix} + \cdots + \chi_{n-1}\begin{pmatrix}\alpha_{0,n-1}\\ \alpha_{1,n-1}\\ \vdots\\ \alpha_{m-1,n-1}\end{pmatrix} \\
&= \begin{pmatrix}\chi_0\alpha_{0,0}\\ \chi_0\alpha_{1,0}\\ \vdots\\ \chi_0\alpha_{m-1,0}\end{pmatrix} + \begin{pmatrix}\chi_1\alpha_{0,1}\\ \chi_1\alpha_{1,1}\\ \vdots\\ \chi_1\alpha_{m-1,1}\end{pmatrix} + \cdots + \begin{pmatrix}\chi_{n-1}\alpha_{0,n-1}\\ \chi_{n-1}\alpha_{1,n-1}\\ \vdots\\ \chi_{n-1}\alpha_{m-1,n-1}\end{pmatrix} \\
&= \begin{pmatrix}
\chi_0\alpha_{0,0} + & \chi_1\alpha_{0,1} + & \cdots + & \chi_{n-1}\alpha_{0,n-1} \\
\chi_0\alpha_{1,0} + & \chi_1\alpha_{1,1} + & \cdots + & \chi_{n-1}\alpha_{1,n-1} \\
\vdots & \vdots & \vdots & \vdots \\
\chi_0\alpha_{m-1,0} + & \chi_1\alpha_{m-1,1} + & \cdots + & \chi_{n-1}\alpha_{m-1,n-1}
\end{pmatrix} \\
&= \begin{pmatrix}
\alpha_{0,0}\chi_0 + & \alpha_{0,1}\chi_1 + & \cdots + & \alpha_{0,n-1}\chi_{n-1} \\
\alpha_{1,0}\chi_0 + & \alpha_{1,1}\chi_1 + & \cdots + & \alpha_{1,n-1}\chi_{n-1} \\
\vdots & \vdots & \vdots & \vdots \\
\alpha_{m-1,0}\chi_0 + & \alpha_{m-1,1}\chi_1 + & \cdots + & \alpha_{m-1,n-1}\chi_{n-1}
\end{pmatrix} \\
&= \begin{pmatrix}
\alpha_{0,0} & \alpha_{0,1} & \cdots & \alpha_{0,n-1} \\
\alpha_{1,0} & \alpha_{1,1} & \cdots & \alpha_{1,n-1} \\
\vdots & \vdots & \vdots & \vdots \\
\alpha_{m-1,0} & \alpha_{m-1,1} & \cdots & \alpha_{m-1,n-1}
\end{pmatrix}\begin{pmatrix}\chi_0\\ \chi_1\\ \vdots\\ \chi_{n-1}\end{pmatrix} = Ax.
\end{aligned}
$$

Definition 2.7 ($\mathbb{R}^{m\times n}$)
The set of all $m \times n$ real valued matrices is denoted by $\mathbb{R}^{m\times n}$.

Thus, $A \in \mathbb{R}^{m\times n}$ means that A is a real valued matrix of size $m \times n$.

Definition 2.8 (**Matrix-vector multiplication or product**)

Let $A \in \mathbb{R}^{m \times n}$ and $x \in \mathbb{R}^n$ with

$$A = \begin{pmatrix} \alpha_{0,0} & \alpha_{0,0} & \cdots & \alpha_{0,n-1} \\ \alpha_{1,0} & \alpha_{1,0} & \cdots & \alpha_{1,n-1} \\ \vdots & \vdots & & \vdots \\ \alpha_{m-1,0} & \alpha_{m-1,0} & \cdots & \alpha_{m-1,n-1} \end{pmatrix} \quad and \quad x = \begin{pmatrix} \chi_0 \\ \chi_1 \\ \vdots \\ \chi_{n-1} \end{pmatrix}.$$

then

$$\begin{pmatrix} \alpha_{0,0} & \alpha_{0,1} & \cdots & \alpha_{0,n-1} \\ \alpha_{1,0} & \alpha_{1,1} & \cdots & \alpha_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m-1,0} & \alpha_{m-1,1} & \cdots & \alpha_{m-1,n-1} \end{pmatrix} \begin{pmatrix} \chi_0 \\ \chi_1 \\ \vdots \\ \chi_{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_{0,0}\chi_0 + & \alpha_{0,1}\chi_1 + & \cdots + & \alpha_{0,n-1}\chi_{n-1} \\ \alpha_{1,0}\chi_0 + & \alpha_{1,1}\chi_1 + & \cdots + & \alpha_{1,n-1}\chi_{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{m-1,0}\chi_0 + & \alpha_{m-1,1}\chi_1 + & \cdots + & \alpha_{m-1,n-1}\chi_{n-1} \end{pmatrix}. \tag{2.3}$$

2.4.2 Practice with Matrix-Vector Multiplication

Homework 2.4.2.1 Compute Ax when $A = \begin{pmatrix} -1 & 0 & 2 \\ -3 & 1 & -1 \\ -2 & -1 & 2 \end{pmatrix}$ and $x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Homework 2.4.2.2 Compute Ax when $A = \begin{pmatrix} -1 & 0 & 2 \\ -3 & 1 & -1 \\ -2 & -1 & 2 \end{pmatrix}$ and $x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Homework 2.4.2.3 If A is a matrix and e_j is a unit basis vector of appropriate length, then $Ae_j = a_j$, where a_j is the jth column of matrix A.

Always/Sometimes/Never

Homework 2.4.2.4 If x is a vector and e_i is a unit basis vector of appropriate size, then their dot product, $e_i^T x$, equals the ith entry in x, χ_i.

Always/Sometimes/Never

Homework 2.4.2.5 Compute

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}^T \left(\begin{pmatrix} -1 & 0 & 2 \\ -3 & 1 & -1 \\ -2 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = \underline{\quad\quad}$$

Homework 2.4.2.6 Compute

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}^T \left(\begin{pmatrix} -1 & 0 & 2 \\ -3 & 1 & -1 \\ -2 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = \underline{\hspace{1cm}}$$

Homework 2.4.2.7 Let A be a $m \times n$ matrix and $\alpha_{i,j}$ its (i,j) element. Then $\alpha_{i,j} = e_i^T(Ae_j)$.

Always/Sometimes/Never

Homework 2.4.2.8 Compute

- $\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -2 & 3 \end{pmatrix} \left((-2) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) =$

- $(-2) \left(\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) =$

- $\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -2 & 3 \end{pmatrix} \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) =$

- $\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} =$

Homework 2.4.2.9 Let $A \in \mathbb{R}^{m \times n}$; $x, y \in \mathbb{R}^n$; and $\alpha \in \mathbb{R}$. Then
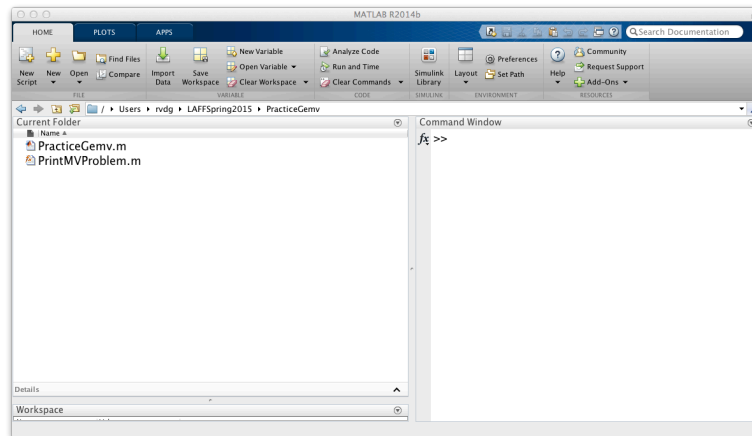
- $A(\alpha x) = \alpha(Ax)$.

- $A(x+y) = Ax + Ay$.

Always/Sometimes/Never

Homework 2.4.2.10 You can practice as little or as much as you want!

Some of the following instructions are for the desktop version of Matlab, but it should be pretty easy to figure out what to do instead with Matlab Online.

Start up Matlab or log on to Matlab Online and change the current directory to Programming/Week02/.

Then type `PracticeGemv` in the command window and you get to practice all the matrix-vector multiplications you want! For example, after a bit of practice my window looks like
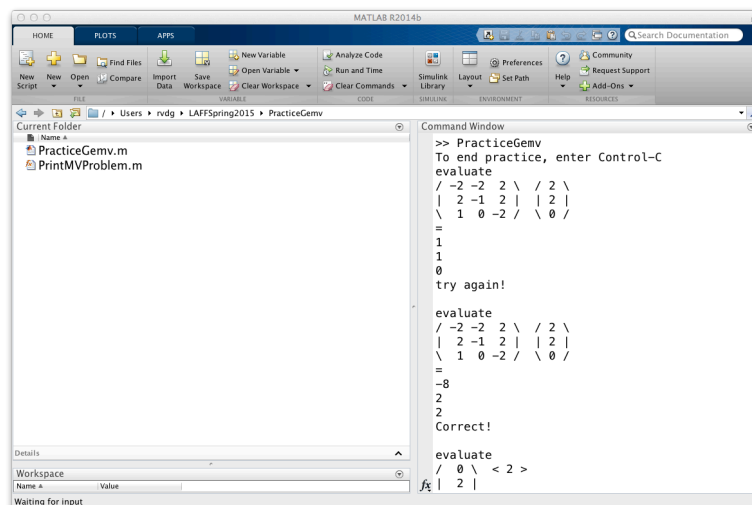
Practice all you want!

2.4.3 It Goes Both Ways

The last exercise proves that the function that computes matrix-vector multiplication is a linear transformation:

Theorem 2.9 *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be defined by $L(x) = Ax$ where $A \in \mathbb{R}^{m \times n}$. Then L is a linear transformation.*

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if it can be written as a matrix-vector multiplication.

Homework 2.4.3.1 Give the linear transformation that corresponds to the matrix

$$\left(\begin{array}{cccc} 2 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{array} \right).$$

Homework 2.4.3.2 Give the linear transformation that corresponds to the matrix

$$\left(\begin{array}{cc} 2 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{array} \right).$$

Example 2.10 We showed that the function $f(\left(\begin{array}{c} \chi_0 \\ \chi_1 \end{array} \right)) = \left(\begin{array}{c} \chi_0 + \chi_1 \\ \chi_0 \end{array} \right)$ is a linear transformation in an earlier example. We will now provide an alternate proof of this fact.

We compute a *possible* matrix, A, that represents this linear transformation. We will then show that $f(x) = Ax$, which then means that f *is* a linear transformation since the above theorem states that matrix-vector multiplications are linear transformations.

To compute a possible matrix that represents f consider:

$$f(\left(\begin{array}{c} 1 \\ 0 \end{array} \right)) = \left(\begin{array}{c} 1+0 \\ 1 \end{array} \right) = \left(\begin{array}{c} 1 \\ 1 \end{array} \right) \quad \text{and} \quad f(\left(\begin{array}{c} 0 \\ 1 \end{array} \right)) = \left(\begin{array}{c} 0+1 \\ 0 \end{array} \right) = \left(\begin{array}{c} 1 \\ 0 \end{array} \right).$$

Thus, *if f is a linear transformation*, then $f(x) = Ax$ where $A = \left(\begin{array}{cc} 1 & 1 \\ 1 & 0 \end{array} \right)$. Now,

$$Ax = \left(\begin{array}{cc} 1 & 1 \\ 1 & 0 \end{array} \right) \left(\begin{array}{c} \chi_0 \\ \chi_1 \end{array} \right) g = \left(\begin{array}{c} \chi_0 + \chi_1 \\ \chi_0 \end{array} \right) = f(\left(\begin{array}{c} \chi_0 \\ \chi_1 \end{array} \right)) = f(x).$$

Hence f is a linear transformation since $f(x) = Ax$.

Example 2.11 In Example 2.3 we showed that the transformation $f(\begin{pmatrix} \chi \\ \psi \end{pmatrix}) = \begin{pmatrix} \chi + \psi \\ \chi + 1 \end{pmatrix}$ is not a linear trans-

formation. We now show this again, by computing a possible matrix that represents it, and then showing that it does *not* represent it.

To compute a possible matrix that represents f consider:

$$f(\begin{pmatrix} 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} 1+0 \\ 1+1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{and} \quad f(\begin{pmatrix} 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} 0+1 \\ 0+1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Thus, *if f is a linear transformation*, then $f(x) = Ax$ where $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$. Now,

$$Ax = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix} = \begin{pmatrix} \chi_0 + \chi_1 \\ 2\chi_0 + \chi_1 \end{pmatrix} \neq \begin{pmatrix} \chi_0 + \chi_1 \\ \chi_0 + 1 \end{pmatrix} = f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) = f(x).$$

Hence f is *not* a linear transformation since $f(x) \neq Ax$.

The above observations give us a straight-forward, fool-proof way of checking whether a function is a linear transformation. You compute a possible matrix and then you check if the matrix-vector multiply always yields the same result as evaluating the function.

Homework 2.4.3.3 Let f be a vector function such that $f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) = \begin{pmatrix} \chi_0^2 \\ \chi_1 \end{pmatrix}$ Then

- (a) f is a linear transformation.

- (b) f is not a linear transformation.

- (c) Not enough information is given to determine whether f is a linear transformation.

How do you know?

Homework 2.4.3.4 For each of the following, determine whether it is a linear transformation or not:

- $f(\begin{pmatrix} \chi_0 \\ \chi_1 \\ \chi_2 \end{pmatrix}) = \begin{pmatrix} \chi_0 \\ 0 \\ \chi_2 \end{pmatrix}.$

- $f(\begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}) = \begin{pmatrix} \chi_0^2 \\ 0 \end{pmatrix}.$

## 2.4.4   Rotations and Reflections, Revisited

☛ View at edX

Recall that in the opener for this week we used a geometric argument to conclude that a rotation $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation. We now show how to compute the matrix, A, that represents this rotation.

Given that the transformation is from $\mathbb{R}^2$ to $\mathbb{R}^2$, we know that the matrix will be a 2×2 matrix. It will take vectors of size two as input and will produce vectors of size two. We have also learned that the first column of the matrix A will equal $R_\theta(e_0)$ and the second column will equal $R_\theta(e_1)$.

We first determine what vector results when $e_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is rotated through an angle θ:
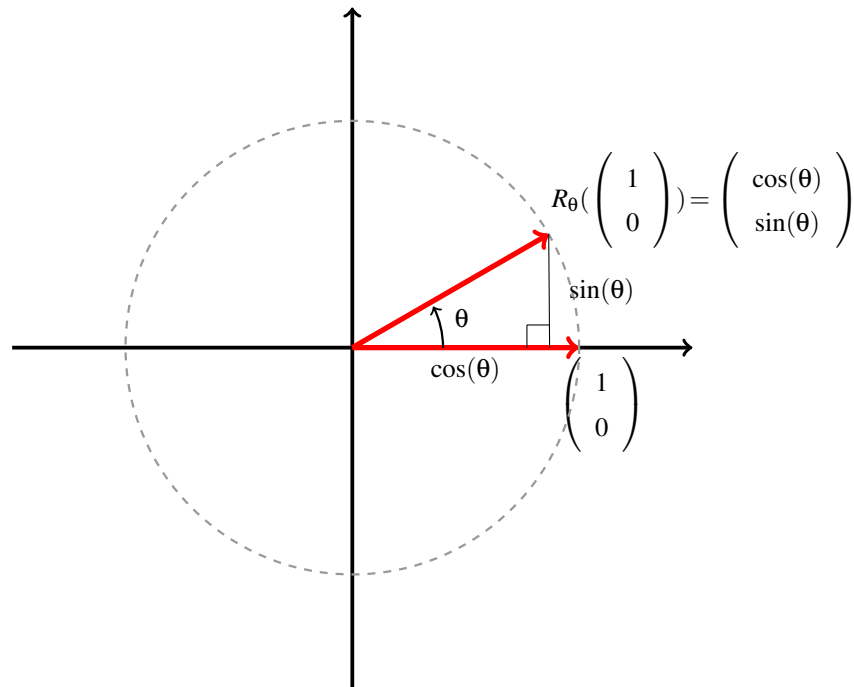
Next, we determine what vector results when $e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is rotated through an angle θ:
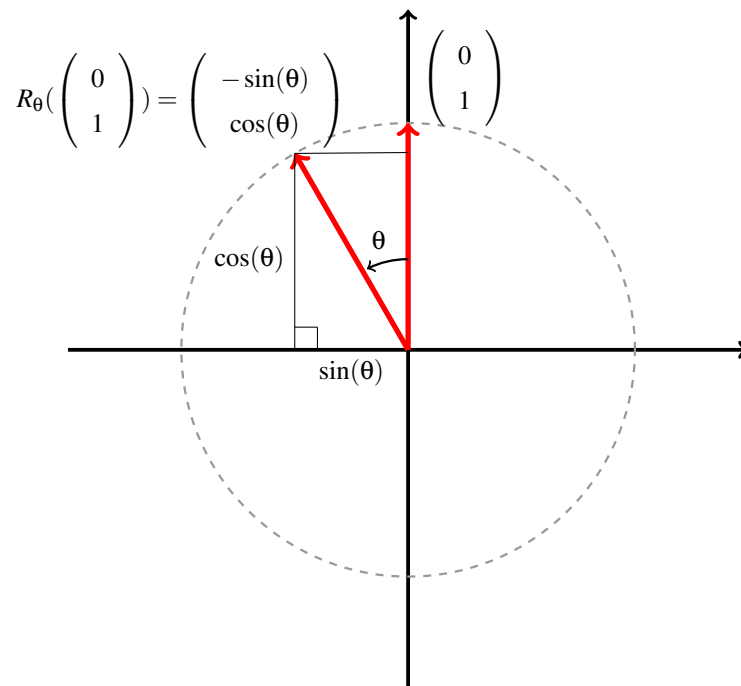
This shows that

$$R_\theta(e_0) = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \quad \text{and} \quad R_\theta(e_1) = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}.$$

We conclude that

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

This means that an arbitrary vector $x = \begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix}$ is transformed into

$$R_\theta(x) = Ax = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \chi_0 \\ \chi_1 \end{pmatrix} = \begin{pmatrix} \cos(\theta)\chi_0 - \sin(\theta)\chi_1 \\ \sin(\theta)\chi_0 + \cos(\theta)\chi_1 \end{pmatrix}.$$

This is a formula very similar to a formula you may have seen in a precalculus or physics course when discussing *change of coordinates*. We will revisit to this later.

Homework 2.4.4.1 A reflection with respect to a 45 degree line is illustrated by
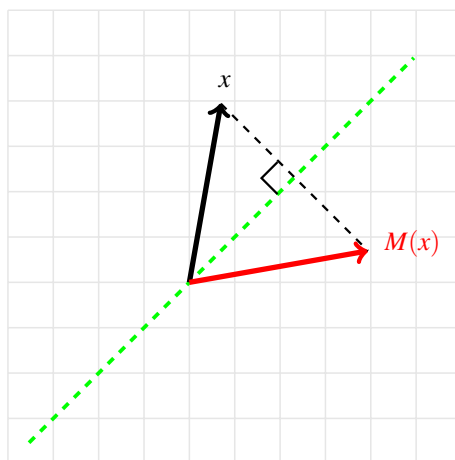
Again, think of the dashed green line as a mirror and let $M : \mathbb{R}^2 \to \mathbb{R}^2$ be the vector function that maps a vector to its mirror image. Evaluate (by examining the picture)

- $M(\begin{pmatrix} 1 \\ 0 \end{pmatrix}) =.$

- $M(\begin{pmatrix} 0 \\ 3 \end{pmatrix}) =.$

- $M(\begin{pmatrix} 1 \\ 2 \end{pmatrix}) =.$
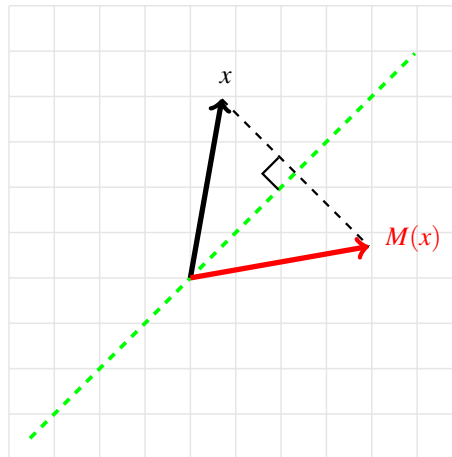

Homework 2.4.4.2 A reflection with respect to a 45 degree line is illustrated by

Again, think of the dashed green line as a mirror and let $M : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the vector function that maps a vector to its mirror image. Compute the matrix that represents M (by examining the picture).

2.5 Enrichment

2.5.1 The Importance of the Principle of Mathematical Induction for Programming

☞ **View at edX**

Read the ACM Turing Lecture 1972 (Turing Award acceptance speech) by Edsger W. Dijkstra:

<p align="center">The Humble Programmer.</p>

Now, to see how the foundations we teach in this class can take you to the frontier of computer science, I encourage you to download (for free)

<p align="center">The Science of Programming Matrix Computations</p>

<p align="center">Skip the first chapter. Go directly to the second chapter. For now, read ONLY that chapter!</p>

Here are the major points as they relate to this class:

- Last week, we introduced you to a notation for expressing algorithms that builds on slicing and dicing vectors.

- This week, we introduced you to the Principle of Mathematical Induction.

- In Chapter 2 of "The Science of Programming Matrix Computations", we

    - Show how Mathematical Induction is related to computations by a loop.
    - How one can use Mathematical Induction to prove the correctness of a loop.
      (No more debugging! You prove it correct like you prove a theorem to be true.)
    - show how one can systematically derive algorithms to be correct. As Dijkstra said:

        Today [back in 1972, but still in 2014] a usual technique is to make a program and then to test it. But: program testing can be a very effective way to show the presence of bugs, but is hopelessly inadequate for showing their absence. The only effective way to raise the confidence level of a program significantly is to give a convincing proof of its correctness. But one should not first make the program and then prove its correctness, because then the requirement of providing the proof would only increase the poor programmer's burden. On the contrary: the programmer should let correctness proof and program grow hand in hand.

To our knowledge, for more complex programs that involve loops, we are unique in having made this comment of Dijkstra's practical. (We have practical libraries with hundreds of thousands of lines of code that have been derived to be correct.)

Teaching you these techniques as part of this course would take the course in a very different direction. So, if this interests you, you should pursue this further on your own.

2.5.2 Puzzles and Paradoxes in Mathematical Induction

Read the article "Puzzles and Paradoxes in Mathematical Induction" by Adam Bjorndahl.

2.6 Wrap Up

2.6.1 Homework

Homework 2.6.1.1 Suppose a professor decides to assign grades based on two exams and a final. Either all three exams (worth 100 points each) are equally weighted or the final is double weighted to replace one of the exams to benefit the student. The records indicate each score on the first exam as χ_0, the score on the second as χ_1, and the score on the final as χ_2. The professor transforms these scores and looks for the maximum entry. The following describes the linear transformation:

$$l\left(\begin{array}{c} \chi_0 \\ \chi_1 \\ \chi_2 \end{array}\right) = \left(\begin{array}{c} \chi_0 + \chi_1 + \chi_2 \\ \chi_0 + 2\chi_2 \\ \chi_1 + 2\chi_2 \end{array}\right)$$

What is the matrix that corresponds to this linear transformation?

If a student's scores are $\left(\begin{array}{c} 68 \\ 80 \\ 95 \end{array}\right)$, what is the transformed score?

2.6.2 Summary

A linear transformation is a vector function that has the following two properties:

- Transforming a scaled vector is the same as scaling the transformed vector:

$$L(\alpha x) = \alpha L(x)$$

- Transforming the sum of two vectors is the same as summing the two transformed vectors:

$$L(x + y) = L(x) + L(y)$$

$L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if (iff) for all $u, v \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$

$$L(\alpha u + \beta v) = \alpha L(u) + \beta L(v).$$

If $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then

$$L(\beta_0 x_0 + \beta_1 x_1 + \cdots + \beta_{k-1} x_{k-1}) = \beta_0 L(x_0) + \beta_1 L(x_1) + \cdots + \beta_{k-1} L(x_{k-1}).$$

A vector function $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if it can be represented by an $m \times n$ **matrix**, which is a very special two dimensional array of numbers (elements).

The **set of all real valued** $m \times n$ **matrices** is denoted by $\mathbb{R}^{m \times n}$.

Let A is the matrix that represents $L : \mathbb{R}^n \to \mathbb{R}^m$, $x \in \mathbb{R}^n$, and let

$$A = \left(\begin{array}{c|c|c|c} a_0 & a_1 & \cdots & a_{n-1} \end{array} \right) \qquad (a_j \text{ equals the } j\text{th column of } A)$$

$$= \begin{pmatrix} \alpha_{0,0} & \alpha_{0,1} & \cdots & \alpha_{0,n-1} \\ \alpha_{1,0} & \alpha_{1,1} & \cdots & \alpha_{1,n-1} \\ \vdots & \vdots & & \vdots \\ \alpha_{m-1,0} & \alpha_{m-1,1} & \cdots & \alpha_{m-1,n-1} \end{pmatrix} \qquad (\alpha_{i,j} \text{ equals the } (i,j) \text{ element of } A).$$

$$x = \begin{pmatrix} \chi_0 \\ \chi_1 \\ \vdots \\ \chi_{n-1} \end{pmatrix}$$

Then

- $A \in \mathbb{R}^{m \times n}$.

- $a_j = L(e_j) = A e_j$ (the jth column of A is the vector that results from transforming the unit basis vector e_j).

- $L(x) = L(\sum_{j=0}^{n-1} \chi_j e_j) = \sum_{j=0}^{n-1} L(\chi_j e_j) = \sum_{j=0}^{n-1} \chi_j L(e_j) = \sum_{j=0}^{n-1} \chi_j a_j$.

-
$$Ax = L(x)$$

$$= \left(\begin{array}{c|c|c|c} a_0 & a_1 & \cdots & a_{n-1} \end{array} \right) \begin{pmatrix} \chi_0 \\ \chi_1 \\ \vdots \\ \chi_{n-1} \end{pmatrix}$$

$$= \chi_0 a_0 + \chi_1 a_1 + \cdots + \chi_{n-1} a_{n-1}$$

$$= \chi_0 \begin{pmatrix} \alpha_{0,0} \\ \alpha_{1,0} \\ \vdots \\ \alpha_{m-1,0} \end{pmatrix} + \chi_1 \begin{pmatrix} \alpha_{0,1} \\ \alpha_{1,1} \\ \vdots \\ \alpha_{m-1,1} \end{pmatrix} + \cdots + \chi_{n-1} \begin{pmatrix} \alpha_{0,n-1} \\ \alpha_{1,n-1} \\ \vdots \\ \alpha_{m-1,n-1} \end{pmatrix}$$

$$= \begin{pmatrix} \chi_0 \alpha_{0,0} + \chi_1 \alpha_{0,1} + \cdots + \chi_{n-1} \alpha_{0,n-1} \\ \chi_0 \alpha_{1,0} + \chi_1 \alpha_{1,1} + \cdots + \chi_{n-1} \alpha_{1,n-1} \\ \vdots \\ \chi_0 \alpha_{m-1,0} + \chi_1 \alpha_{m-1,1} + \cdots + \chi_{n-1} \alpha_{m-1,n-1} \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_{0,0} & \alpha_{0,1} & \cdots & \alpha_{0,n-1} \\ \alpha_{1,0} & \alpha_{1,1} & \cdots & \alpha_{1,n-1} \\ \vdots & \vdots & & \vdots \\ \alpha_{m-1,0} & \alpha_{m-1,1} & \cdots & \alpha_{m-1,n-1} \end{pmatrix} \begin{pmatrix} \chi_0 \\ \chi_1 \\ \vdots \\ \chi_{n-1} \end{pmatrix}.$$

How to check if a vector function is a linear transformation:

- Check if $f(0) = 0$. If it isn't, it is **not** a linear transformation.

- If $f(0) = 0$ then *either*:

    - Prove it is or isn't a linear transformation from the definition:

        * Find an example where $f(\alpha x) \neq \alpha f(x)$ or $f(x+y) \neq f(x) + f(y)$. In this case the function is *not* a linear transformation; or
        * Prove that $f(\alpha x) = \alpha f(x)$ and $f(x+y) = f(x) + f(y)$ for all α, x, y.

    or

    - Compute the *possible* matrix A that represents it and see if $f(x) = Ax$. If it is equal, it is a linear transformation. If it is not, it is not a linear transformation.

Mathematical induction is a powerful proof technique about natural numbers. (There are more general forms of mathematical induction that we will not need in our course.)

The following results about summations will be used in future weeks:

- $\sum_{i=0}^{n-1} i = n(n-1)/2 \approx n^2/2$.

- $\sum_{i=1}^{n} i = n(n+1)/2 \approx n^2/2$.

- $\sum_{i=0}^{n-1} i^2 = (n-1)n(2n-1)/6 \approx \frac{1}{3}n^3$.