

# Math 1600B Lecture 22, Section 2, 3 Mar 2014

## Announcements:

**Read** Sections 4.0 and 4.1 for next class. Work through recommended [homework questions](#).

**Office hour:** today, 1:30-2:30 and Wednesday, 10:30-11:15, MC103B.

**Tutorials:** Quiz this week covers Sections 3.5 and 3.6.

**Help Centers:** Monday-Friday 2:30-6:30 in MC 106.

## Review of last lecture: Section 3.6: Linear Transformations

Given an  $m \times n$  matrix  $A$ , we can use  $A$  to transform a column vector in  $\mathbb{R}^n$  into a column vector in  $\mathbb{R}^m$ . We write:

$$T_A(\vec{x}) = A\vec{x} \quad \text{for } \vec{x} \text{ in } \mathbb{R}^n$$

Any rule  $T$  that assigns to each  $\vec{x}$  in  $\mathbb{R}^n$  a unique vector  $T(\vec{x})$  in  $\mathbb{R}^m$  is called a **transformation** from  $\mathbb{R}^n$  to  $\mathbb{R}^m$  and is written  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

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

1.  $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$  for all  $\vec{u}$  and  $\vec{v}$  in  $\mathbb{R}^n$ , and
2.  $T(c\vec{u}) = cT(\vec{u})$  for all  $\vec{u}$  in  $\mathbb{R}^n$  and all scalars  $c$ .

**Theorem 3.30:** Let  $A$  be an  $m \times n$  matrix. Then  $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is a linear transformation.

**Theorem 3.31:** Let  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  be a linear transformation. Then  $T = T_A$ , where

$$A = [T(\vec{e}_1) \mid T(\vec{e}_2) \mid \cdots \mid T(\vec{e}_n)]$$

The matrix  $A$  is called the **standard matrix** of  $T$  and is written  $[T]$ .

**Example 3.58:** Let  $R_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be rotation by an angle  $\theta$  counterclockwise about the origin. Show that  $R_\theta$  is linear and find its standard matrix.

**Solution:** A geometric argument shows that  $R_\theta$  is linear.

To find the standard matrix, we note that

$$R_\theta \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \quad \text{and} \quad R_\theta \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}$$

Therefore, the standard matrix of  $R_\theta$  is  $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ .

## New linear transformations from old

If  $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$  and  $S : \mathbb{R}^n \rightarrow \mathbb{R}^p$ , then  $S(T(\vec{x}))$  makes sense for  $\vec{x}$  in  $\mathbb{R}^m$ . The **composition** of  $S$  and  $T$  is the transformation  $S \circ T : \mathbb{R}^m \rightarrow \mathbb{R}^p$  defined by

$$(S \circ T)(\vec{x}) = S(T(\vec{x})).$$

If  $S$  and  $T$  are linear, it is easy to check that this new transformation  $S \circ T$  is automatically linear.

**Theorem 3.32:**  $[S \circ T] = [S][T]$ .

## New material

**Applet:** This [java applet](#) illustrates  $T_A$  for any  $2 \times 2$   $A$  you provide.

Try it with:

$$[R_{90^\circ}] = \begin{bmatrix} \cos(90) & -\sin(90) \\ \sin(90) & \cos(90) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

$$[F] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$[R_{60^\circ}] = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 0.5 & -0.866 \\ 0.866 & 0.5 \end{bmatrix}$$

**Example 3.61:** Find the standard matrix of the transformation that rotates  $90^\circ$  counterclockwise and then reflects in the  $x$ -axis. How do  $F \circ R_{90}$  and  $R_{90} \circ F$  compare? On board.

**Example:** It is geometrically clear that  $R_\theta \circ R_\phi = R_{\theta+\phi}$ . This tells us that

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

This implies some trigonometric identities. For example, looking at the top-left entry, we find that

$$\cos(\theta + \phi) = \cos(\theta) \cos(\phi) - \sin(\theta) \sin(\phi)$$

Other trig identities also follow.

Note that  $R_0$  is rotation by zero degrees, so  $R_0(\vec{x}) = \vec{x}$ . We say that  $R_0$  is

the **identity transformation**, which we write  $I : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ . Similarly,  $R_{360} = I$ .

Since  $R_{120} \circ R_{120} \circ R_{120} = R_{360} = I$ , we must have  $[R_{120}]^3 = [I] = I$ .

This is how I came up with the answer  $[R_{120}] = \begin{bmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix}$  to the challenge problem in Lecture 15.

Our new point of view about matrix multiplication gives us a **geometrical** way to understand it!

## Inverses of Linear Transformations

Since  $R_{60} \circ R_{-60} = R_0 = I$ , it follows that  $[R_{60}][R_{-60}] = I$ . So  $[R_{-60}] = [R_{60}]^{-1}$ . See Example 3.62 for details.

**Definition:** Let  $S$  and  $T$  be linear transformations from  $\mathbb{R}^n$  to  $\mathbb{R}^n$ . Then  $S$  and  $T$  are **inverse transformations** if  $S \circ T = I$  and  $T \circ S = I$ . When this is the case, we say that  $S$  and  $T$  are **invertible** and are **inverses**.

The same argument as for matrices shows that an inverse is unique when it exists, so we write  $S = T^{-1}$  and  $T = S^{-1}$ .

**Theorem 3.33:** Let  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be a linear transformation. Then  $T$  is invertible if and only if  $[T]$  is an invertible matrix. In this case,  $[T^{-1}] = [T]^{-1}$ .

The argument is easy and is essentially what we did for  $R_{60}$ .

**Question:** Is projection onto the  $x$ -axis invertible?

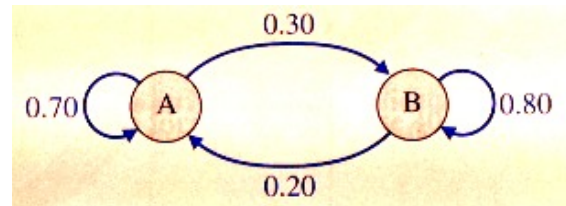
**Question:** Is reflection in the  $x$ -axis invertible?

**Question:** Is translation a linear transformation?

---

## Section 3.7: Markov Chains

**Example 3.64:** 200 people are testing two brands of toothpaste, Brand A and Brand B. Each month they are allowed to switch brands. The research firm observes the following:



- Of those using Brand A in a given month, 70% continue in the following month and 30% switch to B.
- Of those using Brand B in a given month, 80% continue in the following month and 20% switch to A.

This is called a **Markov chain**. There are definite states, and from each state there is a **transition probability** for moving to another state at each time step. These probabilities are constant and depend only on the current state.

Suppose at the start that 120 people use Brand A and 80 people use Brand B. Then, in the next month,

$$0.70(120) + 0.20(80) = 100 \quad \text{will use Brand A}$$

and

$$0.30(120) + 0.80(80) = 100 \quad \text{will use Brand B}$$

This is a matrix equation:

$$\begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \begin{bmatrix} 120 \\ 80 \end{bmatrix} = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$$

Write  $P$  for the **transition matrix** and  $\vec{x}_k$  for the **state vector** after  $k$  months have gone by. Then  $\vec{x}_{k+1} = P\vec{x}_k$ . So

$$\vec{x}_2 = P\vec{x}_1 = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \begin{bmatrix} 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 90 \\ 110 \end{bmatrix}$$

and we see that there are 90 people using Brand A and 110 using Brand B after 2 months.

We can also work with the percentage of people using each brand. Then

$$\vec{x}_0 = \begin{bmatrix} 120/200 \\ 80/200 \end{bmatrix} = \begin{bmatrix} 0.60 \\ 0.40 \end{bmatrix} \text{ and } P\vec{x}_0 = \begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix}. \text{ Vectors with}$$

non-negative components that sum to 1 are called **probability vectors**

Note that  $P$  is a **stochastic matrix**: this means that it is square and that each column is a probability vector.

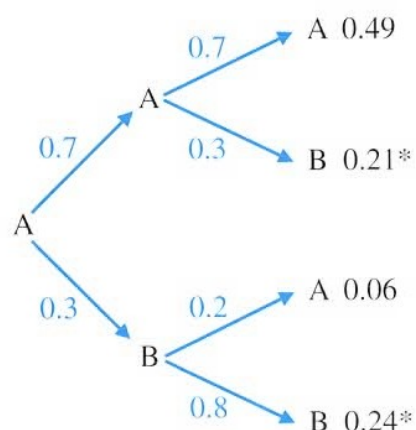
The column indices of  $P$  correspond to the current state and the row indices correspond to the next state. The entry  $P_{ij}$  is the probability that you transition from state  $j$  to state  $i$  in one time step, where we now label the states with numbers.

**Multiple steps:** Can we compute the probability that we go from state  $j$  to state  $i$  in **two** steps?

Well,  $x_{k+2} = Px_{k+1} = P^2x_k$ , so the matrix  $P^2$  computes this transition:

$$P^2 = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.55 & 0.30 \\ 0.45 & 0.70 \end{bmatrix}$$

So the probability of going from Brand A to Brand B after two steps is  $(P^2)_{21} = 0.45 = 0.21 + 0.24$ .



More generally,  $(P^k)_{ij}$  is the probability of going from state  $j$  to state  $i$  in  $k$  steps.

**Long-term behaviour:** By multiplying by  $P$ , you can show that the state evolves as follows:

$$\begin{bmatrix} 0.60 \\ 0.40 \end{bmatrix}, \begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix}, \begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix}, \begin{bmatrix} 0.425 \\ 0.575 \end{bmatrix}, \begin{bmatrix} 0.412 \\ 0.588 \end{bmatrix}, \begin{bmatrix} 0.406 \\ 0.594 \end{bmatrix}, \\ \begin{bmatrix} 0.403 \\ 0.597 \end{bmatrix}, \begin{bmatrix} 0.402 \\ 0.598 \end{bmatrix}, \begin{bmatrix} 0.401 \\ 0.599 \end{bmatrix}, \begin{bmatrix} 0.400 \\ 0.600 \end{bmatrix}, \begin{bmatrix} 0.400 \\ 0.600 \end{bmatrix}, \dots$$

with 40% of the people using Brand A in the long run. Since

$$\begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$$

once we reach this state, we don't leave. A state  $\vec{x}$  with  $P\vec{x} = \vec{x}$  is called a **steady state vector**. We'll prove below that every Markov chain has a steady state vector!

Here's how to find it. We want to find  $\vec{x}$  such that  $(I - P)\vec{x} = \vec{0}$ . The augmented system is

$$[I - P \mid \vec{0}] = \left[ \begin{array}{cc|c} 0.30 & -0.20 & 0 \\ -0.30 & 0.20 & 0 \end{array} \right]$$

which reduces to

$$\left[ \begin{array}{cc|c} 1 & -2/3 & 0 \\ 0 & 0 & 0 \end{array} \right]$$

The solution is

$$x_1 = \frac{2}{3}t, \quad x_2 = t$$

We'd like a probability vector, so  $\frac{2}{3}t + t = 1$  which means that  $t = 3/5$ .

This gives  $\vec{x} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$  as we found above.