Princeton University MAT 202 Spring 2008

Willie W. Wong

Feb 11 - 15

Feb 11, 2008

- Remind students that Adrian Banner's review sessions starts this week: Tuesdays 7:30pm - 9:30pm in Fine 314
- Motivation: recall Linear Combination. First chapter: given a matrix A and a vector $\boldsymbol{b}$, want to know the vector $\boldsymbol{x}$ such that $A\boldsymbol{x} = \boldsymbol{b}$... or, whether $\boldsymbol{b}$ is a linear combination of the column vectors of A. Now we want to think about the case where we are given some pairs $(\boldsymbol{x}, \boldsymbol{b})$ and see how we can find an A that transforms $\boldsymbol{x}$ into $\boldsymbol{b}$.
- The topic is Linear Transformations
  - A *transformation* T from a set X to a set Y is a rule that assigns to every $x \in X$ an element $y = T(x) \in Y$. We call the set X the *domain* or the *source* of the transformation, and the set Y the *co-domain* or the *target* of the transformation.
  - To express that T is a transformation from X to Y, we write
  $$T : X \to Y$$
  - To expression that T takes an element x to an element y, we write
  $$T(x) = y$$
  or, sometimes we drop the parenthesis and write
  $$Tx = y$$
  We call y the image of x under T.
  - (Draw picture.)
  - If the set X is $\mathbb{R}^M$ for some whole number M, and the set Y is $\mathbb{R}^N$ for some N, then in both of the spaces we have the concept of lines and the concept of the origin (the point $(0, 0, \ldots, 0)$). If a transformation
  $$T : \mathbb{R}^M \to \mathbb{R}^N$$

is such that the image of any line is still a line, and the image of the origin is still the origin, then we say that the transformation T is *linear*.

- Above is a geometric definition. It says that essentially, a linear transformation is one that preserves the concept of lines and the concept of the origin. The geometric concept of lines in $\mathbb{R}^M$ is the same as the algebraic concept of addition of vectors. The geometric concept of an origin is the same as the algebraic concept of multiplication by a scalar number. So we can equivalently say that a linear transformation is one that preserves the concept of vector-addition and the concept of multiplication of a vector by a scalar number.
- So the equivalent definition for a linear transformation $T : \mathbb{R}^M \to \mathbb{R}^N$ is that given x_1, x_2 two vectors in $\mathbb{R}^M$, and k a scalar, that T satisfies

$$T(x_1 + kx_2) = T(x_1) + kT(x_2)$$

  And this is the definition that we will use.
- Examples:
  * The transformation $T : \mathbb{R} \to \mathbb{R}$ given by $T(x) = x^2$ is NOT linear. Since

$$4 = T(2) = T(1+1) \neq T(1) + T(1) = 1 + 1 = 2$$

    Notice that for a transformation, not every element in the co-domain can be written as the image of an element in the domain: the negative numbers in $\mathbb{R}$ cannot be the image of this T. The subset of the elements in the co-domain which are images of elements in the domain is called the *range* of the transformation.
  * The transformation $T : \mathbb{R}^2 \to \mathbb{R}$ given by

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = 3x + 4y + 1$$

    is not linear. We can check this in several ways. The geometric definition requires that the image of the origin be the origin. The origin in $\mathbb{R}^2$ is $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and the origin in $\mathbb{R}$ is just 0. But

$$T(\begin{bmatrix} 0 \\ 0 \end{bmatrix}) = 3 \cdot 0 + 4 \cdot 0 + 1 = 1 \neq 0$$

    so T is not linear. You can also check that this is not linear by explicitly showing that

$$T(\begin{bmatrix} 1 \\ 1 \end{bmatrix}) \neq T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) + T(\begin{bmatrix} 0 \\ 1 \end{bmatrix})$$

  * Let $T : \mathbb{R} \to \mathbb{R}^3$ be

$$T(x) = \begin{bmatrix} 3x \\ -5x \\ x \end{bmatrix}$$

then T is linear. As we can check that

$$T(x+ky) = \begin{bmatrix} 3x+3ky \\ -5x-5ky \\ x+ky \end{bmatrix} = \begin{bmatrix} 3x \\ -5x \\ x \end{bmatrix} + k \begin{bmatrix} 3y \\ -5y \\ y \end{bmatrix} = T(x) + kT(y)$$

- Remember we talked about a linear combination of vectors last week. So now let's do something that looks a little bit stupid: we write

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

- In general, we can write

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} = x_1 \boldsymbol{e}_1 + x_2 \boldsymbol{e}_2 + \cdots + x_M \boldsymbol{e}_M$$

  where

$$\boldsymbol{e}_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

  with the "1" in the ith row. We call the $\boldsymbol{e}_i$'s the *standard vectors* in $\mathbb{R}^M$.

- What's the point of splitting up a perfectly good vector into bits, you ask. The answer is so we can better understand the linear transformation of a vector. Suppose we are given a vector $\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_M\boldsymbol{e}_M$. We are interested in knowing how it transforms under T. Using the definition of linearity, we know that

$$T(\boldsymbol{x}) = T(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_M\boldsymbol{e}_M) = x_1T(\boldsymbol{e}_1) + x_2T(\boldsymbol{e}_2) + \cdots + x_MT(\boldsymbol{e}_M)$$

  Does this not look awfully familiar?

- Yes, the expression looks just like the first method of multiplying a vector to a matrix. Recall that we can write a $N \times M$ matrix as M separate column vectors, each of size N, stacked next to each other, so that the matrix $A = \begin{bmatrix} \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_M \end{bmatrix}$. Then we we multiply A by the vector $\boldsymbol{x}$, we end up with

$$A\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_M\boldsymbol{v}_M$$

- So the punch-line, every linear transformation

$$T : \mathbb{R}^M \to \mathbb{R}^N$$

can be written as an $N \times M$ matrix. In particular, given the standard vectors $\boldsymbol{e}_i$ of $\mathbb{R}^M$, we can write the matrix A corresponding to T as

$$T(\boldsymbol{x}) = A \cdot \boldsymbol{x} = \begin{bmatrix} T(\boldsymbol{e}_1) & \ldots & T(\boldsymbol{e}_M) \end{bmatrix}$$

i.e., the matrix A corresponding to a linear transformation T is formed by stacked the size-N column vectors corresponding to the images of the standard vectors next to each other.

- Now, remember that multiplication of a matrix by a vector is also linear in the sense that

$$A \cdot (\boldsymbol{x} + k\boldsymbol{y}) = A\boldsymbol{x} + kA\boldsymbol{y}$$

we see that the transformation from $\mathbb{R}^M$ to $\mathbb{R}^N$ given by the multiplication of an $N \times M$ matrix is naturally a linear transformation.

- So.... we have another equivalent definition of a linear transformation: T is a linear transformation from $\mathbb{R}^M$ to $\mathbb{R}^N$ if and only if the transformation can be written as a multiplication by an $N \times M$ matrix. We will use this definition and the linearity definition above interchangeably.

- Examples

  - Let a, b, c, d be the points in $\mathbb{R}^2$ given by $(1, 1), (1, -1), (-1, -1), (-1, 1)$. Is the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ which satisfies

  $$Ta = b, Tc = a, Tb = d, Td = c$$

  a linear map? The answer is no. Notice that $a = -c = (-1)c$. But

  $$Ta = b \neq -a (= c =)(-1)Tc$$

  So T is not linear.

  - Given a vector $\boldsymbol{w}$ in $\mathbb{R}^M$. The transformation from $\mathbb{R}^M \to \mathbb{R}$ given by

  $$T(\boldsymbol{x}) = \boldsymbol{w} \cdot \boldsymbol{x}$$

  where the two multiplies by the dot product is a linear transformation. That is because of our second method of multiplying matrix against vectors: $\boldsymbol{w} \cdot \boldsymbol{x}$ can be written as multiplying the vector $\boldsymbol{x}$ to a $1 \times M$ matrix whose ith column is the ith entry of $\boldsymbol{w}$.

– The projection

$$T(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} x \\ y \end{bmatrix}$$

from $\mathbb{R}^3 \to \mathbb{R}^2$ is linear, and its corresponding 2×3 matrix is

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

– Another linear map is

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} x - y \\ 2y \\ 4x + 6y \end{bmatrix}$$

Its matrix is

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 4 & 6 \end{bmatrix}$$

– The map $T : \mathbb{R}^M \to \mathbb{R}^M$ give by

$$T(\boldsymbol{x}) = \boldsymbol{x}$$

is called the *identity transformation*. Its corresponding matrix is

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

with 1s on the diagonal and 0s everywhere else. This matrix comes up often enough that we'll give it a name: it is called the *identity matrix* on $\mathbb{R}^M$ and is written I_M.

FEB 13, 2008

- Lots of definitions this week.
- The length of a vector $\boldsymbol{v}$ is given by

$$|\boldsymbol{v}| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$$

the square root of its dot product against itself. An *unit vector* is a vector with length 1. For an arbitrary non-zero vector $\boldsymbol{v}$, its corresponding unit vector is given by $\frac{1}{|\boldsymbol{v}|}\boldsymbol{v}$.

- The dot product between two non-zero vectors measures how "parallel" they are. In particular, if the dot product between $\boldsymbol{u}$ and $\boldsymbol{v}$ is 0, then they are perpendicular.

- Remember that a linear transformation corresponds, geometrically, to transformations that sends lines to lines and the origin to the origin. If we look at the linear transformation from the plane $\mathbb{R}^2$ to itself, there are essentially five types, as illustrated below.[1]

- A scaling transformation is given by

$$T(\boldsymbol{x}) = k\boldsymbol{x}$$

  for some positive number k. Its transformation matrix is given by

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

  In the illustration, $k = 2$.

- Projection: given a line L (through the origin) in the plane, we ask "how much of a given vector $\boldsymbol{v}$ is parallel to the line L?" Projection unto a line L is to "squash away" all components that are perpendicular to L, and keeping the component that is parallel to L. Let $\boldsymbol{u}$ be an unit vector of L, the formula for the projection is

$$T(\boldsymbol{x}) = (\boldsymbol{x} \cdot \boldsymbol{u})\boldsymbol{u}$$

  If $\boldsymbol{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ (since $|\boldsymbol{u}| = 1$, $a^2 + b^2 = 1$), the corresponding matrix is

$$A = \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}$$

- (Draw pictures of projections of figures)

- Using the projection, given a unit vector $\boldsymbol{u}$, we can write any vector $\boldsymbol{v}$ as the sum of two parts: the part that is parallel to $\boldsymbol{u}$ and the part that is perpendicular. The parallel part is from the projection: $(\boldsymbol{v} \cdot \boldsymbol{u})\boldsymbol{u}$. The perpendicular part we'll just write as $\boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{u})\boldsymbol{u}$. You can easily check that

$$[\boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{u})\boldsymbol{u}] \cdot \boldsymbol{u} = 0$$

  and

$$[\boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{u})\boldsymbol{u}] + (\boldsymbol{v} \cdot \boldsymbol{u})\boldsymbol{u} = \boldsymbol{v}$$

[1] I lied a little bit here. Most of the linear transformations of the plane to itself can be represented as successive compositions of the five operations listed below. There's one that cannot: stretching a given direction by a factor k. The corresponding matrix has the numbers $(1, 1, \ldots, k, 1, \ldots, 1)$ on the diagonal and 0 everywhere else. Pedagogically this one is not included here because, as we shall see later on, this transformation changes the eigenvalues of the matrix, and is, in some sense, not allowed. Since this transformation corresponds to the Gauss-Jordan elimination operator of "multiplying a row by a non-zero constant factor", we see that Gauss-Jordan elimination cannot be completely encoded in the following five transformations. The operation of swapping rows is represented by the rotation or the reflection, and the operation of adding one row to another is the shear transformation.

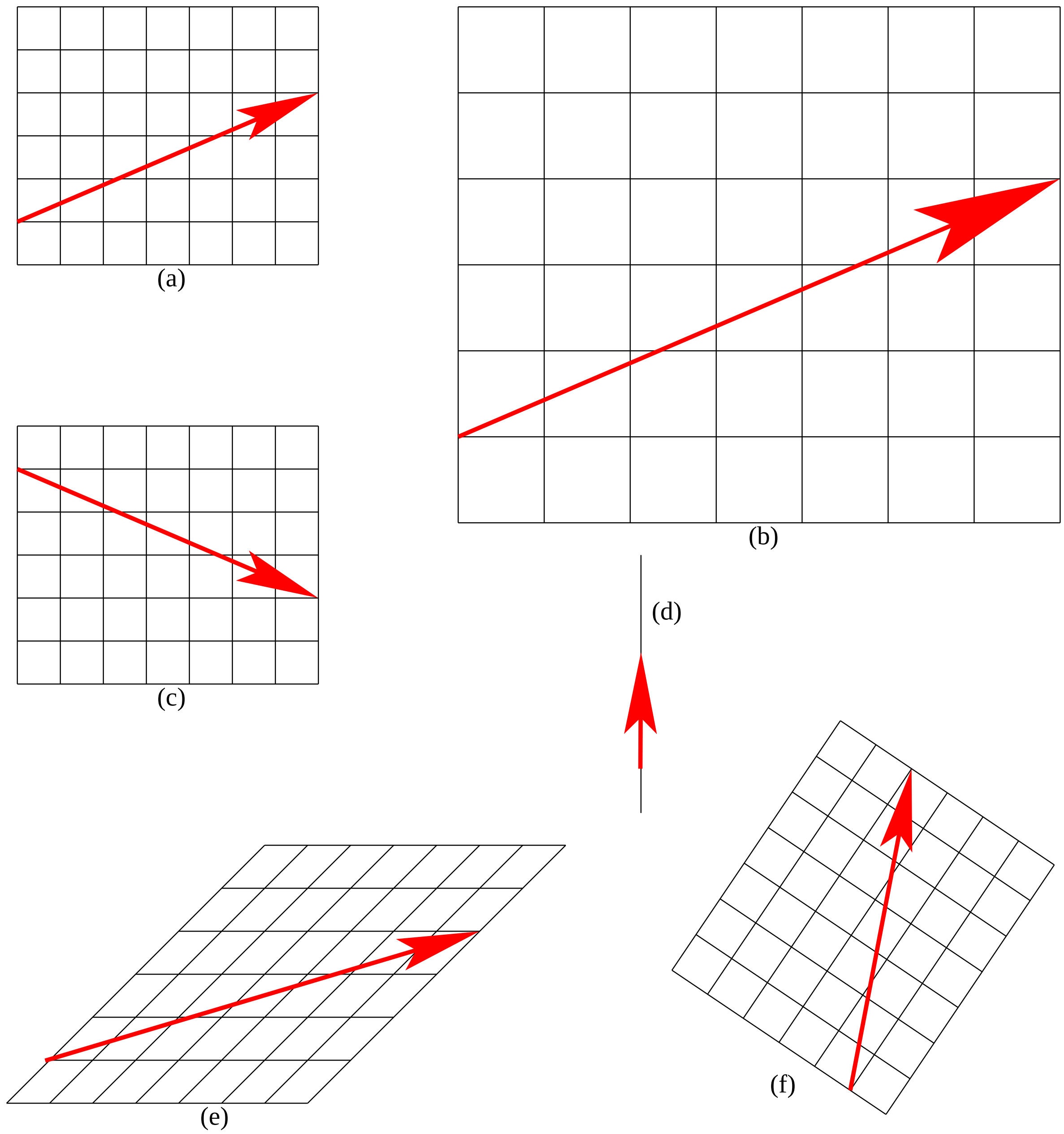

Figure 1: Examples of planar linear transformations: (a) is the original image; (b) is a scaling by a factor of 2; (c) is a reflection about a horizontal line; (d) is the projection unto a vertical axis; (e) is a shear of the original, where the point (x, y) is send to $(x + y, y)$; and (f) is a counter-clockwise rotation of the original by $56°$.

- Reflection: suppose we have a line L in the plane, and we want to reflect our figure over that line L. How do we do it?
  - First let's look at the case when L is the horizontal axis. To reflect over the horizontal, or x-, axis, we send the y coordinate of a figure to $-y$. So the transformation looks like

    $$T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} x \\ -y \end{bmatrix}$$

    With the associated matrix looking like

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

  - For reflection over the line L (which has its unit vector $\boldsymbol{u}$), we first decompose a vector $\boldsymbol{v}$ into the parts parallel to and perpendicular to $\boldsymbol{u}$ like above. To reflect over L, we keep the part that is parallel the same, and negate the part that is perpendicular to $\boldsymbol{u}$. So we have the transformation is given by

    $$T(\boldsymbol{x}) = (\boldsymbol{x} \cdot \boldsymbol{u})\boldsymbol{u} + (-1)[\boldsymbol{x} - (\boldsymbol{x} \cdot \boldsymbol{u})\boldsymbol{u}] = 2(\boldsymbol{x} \cdot \boldsymbol{u})\boldsymbol{u} - \boldsymbol{x}$$

  - Again, supposing the unit vector $\boldsymbol{u} = \begin{bmatrix} a \\ b \end{bmatrix}$, the associated matrix looks like

    $$A = \begin{bmatrix} 2a^2 - 1 & 2ab \\ 2ab & 2b^2 - 1 \end{bmatrix}$$

- Shear: a shear is achieved by slicing up the plane into parallel lines, and then rearranging the lines by shifting each successive line a little bit more.
  - A vertical shift is given by the matrix

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

  - A horizontal shift is given by

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

  - A commonly seen example is the *Galilean transform*. A point in space-time is transformed by

    $$T(\begin{bmatrix} x \\ t \end{bmatrix}) = \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} = \begin{bmatrix} x + vt \\ t \end{bmatrix}$$

    This transformation relates vectors as seen in by two different observers, one moving at a velocity v relative to the other.

- In general, one recognizes a shear transformation when the transformation fixes some direction ($T(\boldsymbol{w}) = \boldsymbol{w}$), and that an arbitrary vector is shifted by a multiple of that direction ($T(\boldsymbol{x}) - \boldsymbol{x}$ is in the direction of $\boldsymbol{w}$).
- (This bullet will not be on the test. It is not to be taught in class. Included for the reader's benefit.) Given unit vectors $\boldsymbol{u}, \boldsymbol{w}$ such that they are perpendicular $\boldsymbol{u} \cdot \boldsymbol{w} = 0$. And given a non-zero constant k, the shear transform that fixes the $\boldsymbol{w}$ direction is given by
$$T(\boldsymbol{x}) = \boldsymbol{x} + k(\boldsymbol{x} \cdot \boldsymbol{u})\boldsymbol{w}$$
If $\boldsymbol{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} b \\ -a \end{bmatrix}$ (up to multiplying $\boldsymbol{w}$ by (-1), this is the canonical form of two unit vectors in a plane that are perpendicular.), then the associated matrix is
$$A = \begin{bmatrix} 1 + kab & kb^2 \\ -ka^2 & 1 - kab \end{bmatrix}$$

- Rotation: the matrix for rotation of an angle θ in the counterclockwise direction is
$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$
Notice that A has the form
$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$
with $a^2 + b^2 = 1$. Any matrix that has this form is a matrix of a rotation.

- By applying various of the above transformations successively, we can get more complicated linear transformations. Notice: the order matters! For example, consider a counterclockwise rotation by $90°$ and a vertical shear. Starting with the unit vector in the horizontal direction, if we first rotate then shear, the result is the unit vector in the vertical direction (since a vertical shear does not change a vector in the vertical direction). But if we first shear then rotate, we'll get a vector sitting in the second quadrant.

- Generalizations to higher dimensions:
  - The scaling transformation is just a constant times the identity transformation, both of which is available in higher dimensions, so extending it to higher dimensions is easy.
  - Similarly, the projection expression
$$T(\boldsymbol{x}) = (\boldsymbol{x} \cdot \boldsymbol{u})\boldsymbol{u}$$
for some unit vector $\boldsymbol{u}$ is also available in higher dimensions, but the corresponding matrix becomes more complicated.

– The reflection transformation is slightly tricky. The idea is that once you are given a vector and a line to reflect the vector over, there are two cases: the vector is parallel to the line or not. In the first case, since the vector is parallel to the line, its reflection is itself, so nothing it to be done. If the vector is not on the line, then the vector and the line together defines at least three non-colinear points, which then define a plane. The reflection of the vector is defined as the reflection *in* that plane over the line. While the geometrical interpretation is slightly more complicated, the algebraic expression is still the same:

$$T(\boldsymbol{x}) = 2(\boldsymbol{x} \cdot \boldsymbol{u})\boldsymbol{u} - \boldsymbol{x}$$

and, as for the projection, the associated matrix is complicated.

– The shear and rotation transformations are much more complicated. The general algebraic expression given for the shear (which is not covered in class) above in terms of two perpendicular unit vectors can be used to generalize the shear transformation to higher dimensions. The rotation is more complicated because one also need to specify which axis to rotate around. (But illustrate with rotation about a fixed coordinate axis, which is simple.)

FEB 15, 2008

i. Some stuff I forgot to mention last time

Two notations for linear transformations of the plane: we write

$$proj_L \boldsymbol{x}$$

for projection onto the line L. If $\boldsymbol{u}$ is a unit vector of L, then

$$proj_L \boldsymbol{x} = (\boldsymbol{x} \cdot \boldsymbol{u})\boldsymbol{u}$$

We write

$$ref_L \boldsymbol{x}$$

for reflection about the line L. So again, using $\boldsymbol{u}$ to denote a unit vector of L, we have

$$ref_L \boldsymbol{x} = 2(\boldsymbol{x} \cdot \boldsymbol{u})\boldsymbol{u} - \boldsymbol{x}$$

ii. Inverse of a linear transformation

- First, recall the inverse of a function/transformation/map
  – (Draw standard picture with sets and maps.)
  – Remember that a transformation (or a map) is an association between two sets, X, the domain, and Y, the codomain, with every elements in X associated to an image.

- A transformation is invertible if the association can go the other way, i.e., every element in Y is the image of only one element in x.
- There are two ways for a transformation to *fail* to be invertible: 1) not every element in the codomain is the image of an element in the domain (give example) or 2) some element in the codomain is the image of more than one element in the domain (give example).
- If the transformation $T : X \to Y$ is invertible, we write T^{-1} to denote the inverse transformation defined from $Y \to X$. For every $y = T(x)$, we have that $T^{-1}(y) = x$. In other words, $T^{-1}(y)$ is the *one and only one* $x \in X$ *satisfying* $T(x) = y$.
- We have
$$T(T^{-1}(y)) = y \ , \quad T^{-1}(T(x)) = x$$
- If T is invertible, T^{-1} is also invertible, with
$$(T^{-1})^{-1} = T$$
- Example: $T : \mathbb{R} \to \mathbb{R}$,
$$T(x) = x^3$$
is invertible. Neither of the possible points of failure are problems.
- Example:
$$T(x) = x^3 - x^2$$
is not invertible. It fails the second possible point of failure: the points $0, 1$ both satisfy $T(0) = 0$ and $T(1) = 0$. So $0 \in Y$ is an example of a point that does not have a unique pre-image in X.
- Definition: given a transformation $T : X \to Y$, the pre-image of a point $y \in Y$ is the set of all points $X \in X$ with $T(x) = y$. So the criteria for T to be invertible is that every element in Y has one and only one pre-image.
- Example: $T : \mathbb{R} \to \mathbb{R}$
$$T(x) = x^2$$
fails both criteria: $T(x) = T(-x)$, so every positive number has two pre-images. On the other hand, every negative image has no pre-image.
- Fact: if the transformation only fails the first criteria, i.e., if there exists y such that the pre-image of y is empty, but for every element y with a pre-image, the pre-image is unique, then we can make the transformation invertible if we shrink the codomain to remove all the extra points.
- For example, if we consider $T(x) = x^2$ as a map from the positive real numbers to itself, then T is invertible.

• Now, let's restrict our attention to linear transformations. Remember that the linear transformation $T : \mathbb{R}^M \to \mathbb{R}^N$ can be associated to an $N \times M$ matrix A. We consider whether T is invertible by the dimension of A.

- $N > M$: if $N > M$, the rref of the coefficient matrix A must have rows in the bottom that consist of all zeros. (Remember, the number of pivots/dependent variables of A is at most M. Since $N > M$, there must be rows without pivots. And so they have to be all zero.) So, if we choose $\boldsymbol{b}$ carefully in the beginning, the rref of the augmented matrix can be made (for good choices of $\boldsymbol{b}$) to have a row that looks like

$$\begin{bmatrix} 0 & 0 & \dots & 0 & | & 1 \end{bmatrix}$$

  and therefore is inconsistent. This means that we can find $\boldsymbol{b} \in \mathbb{R}^N$ such that there is no solution to

$$A\boldsymbol{x} = \boldsymbol{b}$$

  which means that the linear transformation T fails the first criterion for invertible.

- $N < M$: if $N < M$, remember that we showed last Friday that

$$A\boldsymbol{x} = \boldsymbol{b}$$

  has either zero or infinite solutions for every single $\boldsymbol{b}$ (since there must be free parameters). So the linear transform cannot be invertible (it must fail at least one of the criteria).

- $N = M$: if $N = M$, remember that the system

$$A\boldsymbol{x} = \boldsymbol{b}$$

  has a unique solution if and only if the rank of A is equal to M. This in particular means that *the linear transformation T is invertible if the rref of its associated matrix A is the identify matrix I_M.*

- Question: of the five transformations on Wednesday (scaling, projection, reflection, rotation, shear), which are invertible?

- Now, if T is a linear transformation, and T is invertible, then T^{-1} is also a linear transformation. (Since T takes lines to lines and the origin to the origin, and does so uniquely, the pre-image of a line is still a line, and the pre-image of the origin is still the origin...)

- The criteria for a linear transformation $T : \mathbb{R}^M \to \mathbb{R}^M$ to be invertible can be simplified. Notice that for any linear transformation

$$A\mathbf{0} = \mathbf{0}$$

  using linearity, we only need to check to see whether the 0 vector is the unique vector with $A\boldsymbol{x} = 0$, i.e., for the linear transformation given by a square matrix, the matrix is *not invertible* if and only if there are infinitely many solutions to $A\boldsymbol{x} = 0$.

- Since T^{-1} is a linear transformation, it has an associated matrix. We write it as A^{-1}, and call it the *inverse matrix of A*.

- A square matrix A is said to be invertible if its associated linear transformation T is invertible. And its inverse matrix is A^{-1}. (We restricted to square matrices since we have already seen that non-square matrices cannot have linear transformations that are invertible.)

- What does it mean for T to be invertible? Consider the system of equations

$$A\boldsymbol{x} = \boldsymbol{y}$$

  if A is invertible, we can write

$$\boldsymbol{x} = A^{-1}\boldsymbol{y}$$

  i.e., we can solve for $\boldsymbol{x}$ uniquely in terms of $\boldsymbol{y}$.

- How to find the inverse? Example: is the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{bmatrix}$$

  invertible, and what is its inverse? (Show the diagonalization process.)

- This process can be succintly described as doing Gauss-Jordan elimination on the $M \times 2M$ matrix

$$M = \begin{bmatrix} A & | & I_M \end{bmatrix}$$

  A is invertible when

$$rref(M) = \begin{bmatrix} I_M & | & B \end{bmatrix}$$

  for some B, and in this case $A^{-1} = B$. If, on ther other hand, the left half of $rref(M)$ fails to be I_M, then A is not invertible. (Notice that the left half of rref(M) is rref(A).)

- Example:

$$\begin{bmatrix} 1 & 2 & 2 & | & 1 & 0 & 0 \\ 1 & 3 & 1 & | & 0 & 1 & 0 \\ 1 & 1 & 3 & | & 0 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 2 & | & 1 & 0 & 0 \\ 0 & 1 & -1 & | & -1 & 1 & 0 \\ 0 & -1 & 1 & | & -1 & 0 & 1 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} 1 & 0 & 4 & | & 3 & -2 & 0 \\ 0 & 1 & -1 & | & -1 & 1 & 0 \\ 0 & 0 & 0 & | & -2 & 1 & 1 \end{bmatrix}$$

  So the matrix is not invertible

- Example:

$$\begin{bmatrix} 1 & 1 & 2 & 3 & | & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & | & 0 & 1 & 0 & 0 \\ 2 & 2 & 5 & 4 & | & 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 & | & 0 & 0 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 2 & 3 & | & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & | & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & | & -2 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 & | & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 2 & 3 & | & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & | & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -2 & | & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & | & 0 & 3 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & | & 5 & -20 & -2 & -7 \\ 0 & 1 & 0 & 0 & | & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & | & -2 & 6 & 1 & 2 \\ 0 & 0 & 0 & 1 & | & 0 & 3 & 0 & 1 \end{bmatrix}$$

and the matrix is invertible with the inverse given.

- Example: for which k is the following invertible?

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & k & 1 \\ 0 & 1 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & k-4 & 1 \\ 0 & 1 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & k-4 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 2 \\ 0 & 0 & 9-2k \end{bmatrix}$$

We don't need to compute the full $M \times 2M$ matrix since we are only asking about invertibility, and not about finding the actual inverse. So all we need to look for is whether the rref is equal to the identity matrix. In this case, as long as $9 - 2k \neq 0$, we can divide the last row by $9 - 2k$ and finish the Gauss-Jordan elimination to get the identity matrix. So if $9 - 2k$ is not 0, the matrix is invertible. On the other hand, if $9 - 2k$ is 0, the above matrix is in rref, and so the matrix is not invertible.

- Formula for 2×2 matrices. Start with $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

  – If a, c are both 0, then the matrix is not invertible: since G-J can never give you a leading 1 in the first column. So we can assume that at least one of a and c is non-zero. Since G-J allows swapping of rows, we can assuming that a is non-zero.

  – Then divide the first row by a. And continue with G-J

$$\begin{bmatrix} 1 & \frac{b}{a} & \vdots & \frac{1}{a} & 0 \\ c & d & \vdots & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & \frac{b}{a} & \vdots & \frac{1}{a} & 0 \\ 0 & d - \frac{bc}{a} & \vdots & -\frac{c}{a} & 1 \end{bmatrix}$$

  So as long as the quantity $d - \frac{bc}{a} \neq 0$, we can continue the G-J elimination. Since $a \neq 0$, we multiply the relation and write it as $ad - bc \neq 0$. Notice that in the case where $a = c = 0$, this quantity $ad - bc$ is also 0. This means that as long as $ad - bc \neq 0$, we can continue all the way past this step.

  – Then we finish the G-J

$$\Rightarrow \begin{bmatrix} 1 & \frac{b}{a} & \vdots & \frac{1}{a} & 0 \\ 0 & 1 & \vdots & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & \vdots & \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ 0 & 1 & \vdots & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}$$

- The quantity $ad - bc$ is called the ***determinant*** of the matrix. Suppose the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we write $\det(A) = ad - bc$.
- Conclusion: given a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.
    * A is invertible if and only if its determinant $\det(A) = ad - bc \neq 0$.
    * and when A is invertible its inverse is

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Homework for this week

2.1: 4, 6, 10, 19, 20, 35

2.2: 4, 10, 12, 36

2.3: 10, 20, 26, 30, 35