2/17/16

HW3 due now.
Exam 1 Fri Mar 4 in class.

Today: HW3 Discussion.

Problem 1': Why do I say that a hyperplane is "flat"?

Let $\vec{a}$ be a vector in n-dimensional space and let b be a constant. Then the equation

$$\boxed{\vec{a} \cdot \vec{x} = b}$$

defines the <u>hyperplane</u> perpendicular to $\vec{a}$ that has minimum distance $b / \|\vec{a}\|$ from the origin.

Suppose that $\vec{x}_1$ & $\vec{x}_2$ are two points on this hyperplane. That is, suppose that the equations

$$\vec{a} \cdot \vec{x}_1 = b \quad \& \quad \vec{a} \cdot \vec{x}_2 = b$$

are both true.

Then I claim that the midpoint

$$\frac{1}{2}(\vec{x}_1 + \vec{x}_2)$$

is also on the hyperplane.

Proof: We have

$$\begin{aligned} \vec{a} \cdot \left(\frac{1}{2}(\vec{x}_1 + \vec{x}_2)\right) &= \frac{1}{2}\, \vec{a} \cdot (\vec{x}_1 + \vec{x}_2) \\ &= \frac{1}{2}\left(\vec{a} \cdot \vec{x}_1 + \vec{a} \cdot \vec{x}_2\right) \\ &= \frac{1}{2}(b + b) \\ &= \frac{1}{2}(2b) = b \quad \checkmark \end{aligned}$$

///

Note that this same fact is <u>not</u> true for curvy shapes like a paraboloid or the surface of a sphere. [The midpoint of two points on the surface of a sphere will be <u>inside</u> the sphere, <u>not</u> on the surface.]

More generally, the set of points

$$s\vec{x}_1 + t\vec{x}_2 \quad \text{with} \quad s + t = 1$$

is the unique <u>line</u> in n-dimensional space containing the two points $\vec{x}_1$ & $\vec{x}_2$.

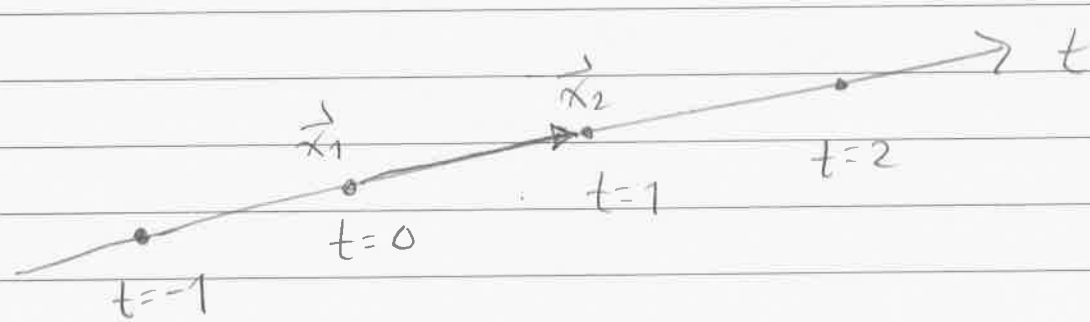
Q: If it's a line, why does it have two free parameters?

A: It doesn't! The equation $s+t=1$ means that $s = 1-t$, so we can express the line as

$$s\vec{x}_1 + t\vec{x}_2 = (1-t)\vec{x}_1 + t\vec{x}_2$$

$$= \vec{x}_1 + t(\vec{x}_2 - \vec{x}_1).$$

This is the line containing <u>the point</u> $\vec{x}_1$ and parallel <u>to the vector</u> $\vec{x}_2 - \vec{x}_1$.



Now, if $\vec{x}_1$ & $\vec{x}_2$ are two points on the hyperplane $\vec{a} \cdot \vec{x} = b$, I claim that the whole line $\vec{x}_1 + t(\vec{x}_2 - \vec{x}_1)$ lives in the hyperplane.

Proof: Assume that $\vec{a} \cdot \vec{x}_1 = b$ & $\vec{a} \cdot \vec{x}_2 = b$.
Then for all values of t we have

$$\vec{a} \cdot \left(\vec{x}_1 + t(\vec{x}_2 - \vec{x}_1) \right)$$

$$= \vec{a} \cdot \vec{x}_1 + t\, \vec{a} \cdot (\vec{x}_2 - \vec{x}_1)$$

$$= \vec{a} \cdot \vec{x}_1 + t\left(\vec{a} \cdot \vec{x}_1 - \vec{a} \cdot \vec{x}_2 \right)$$

$$= b + \cancel{t(b-b)} = b. \checkmark$$

This is really what I mean when I say that a hyperplane is "flat".

But even more is true. Suppose that we have a system of hyperplanes

$$\vec{a}_1 \cdot \vec{x} = b_1, \ \vec{a}_2 \cdot \vec{x} = b_2, \ \ldots, \ \vec{a}_m \cdot \vec{x} = b_m.$$

If the two points $\vec{x}_1$ & $\vec{x}_2$ lie in the intersection of the hyperplanes, then they lie in each individual hyperplane, so the line $\vec{x}_1 + t(\vec{x}_2 - \vec{x}_1)$ lies in each individual hyperplane, so the line lies in the intersection of the hyperplanes.

We conclude that any intersection of hyperplanes is also "flat".

In particular [see HW 1(b)], if 25 hyperplanes in 12-dimensional space meet at two given points $\vec{x}_1$ & $\vec{x}_2$ then they also meet at

every point of the line $\vec{x}_1 + t(\vec{x}_2 - \vec{x}_1)$!

[including, for example, the midpoint of $\vec{x}_1$ & $\vec{x}_2$ (when $t = 1/2$)]

Starting on Friday we will begin developing a languge that makes it much easier to say these things: the languge of "matrix algebra".

Problem 3'. Consider the system

$$\begin{cases} x + y + z = 2 & ① \\ x + 2y + z = 3 & ② \\ 2x + 3y + 2z = c & ③. \end{cases}$$

The planes ① & ② meet in a line L, which we can find by reducing the subsystem of equations ① & ②:

$$\begin{cases} x + y + z = 2 & ① \to ① \\ 0 + y + 0 = 1 & ② \to ② - ① \end{cases}$$

$$\begin{cases} x + 0 + z = 1 & ① \to ① - ② \\ 0 + y + 0 = 1 & ② \to ② \end{cases}$$

Let $t := z$. Then the line L is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1-t \\ 1 \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

The question is: how does the plane ③ meet the line L?

Well, there are three possible cases:

i) (3) contains the line L,

ii) (3) intersects L at a point,

iii) (3) never meets L (i.e. the plane (3) is parallel to the line L).

On HW3 you found that

i) happens when $c = 5$,

ii) never happens,

iii) happens when $c \neq 5$.

See the HW3 solutions for a beautiful picture of the situations i) & iii).

2/19/16

HW 4 : TBA.

Exam 1 Fri Mar 4 in class.

Our discussion of Gaussian elimination is done and now we will move on to a new topic. Actually, it's the same topic but written is a new language:

the language of "matrix algebra".

Recall that the central problem of linear algebra is to solve a system of m linear equations in n unknowns, which we can write as follows:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \qquad\qquad\qquad\qquad \vdots \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

However, this notation is quite cumbersome. So far we have seen two ways to simplify it.

1. Row Picture.

When thinking of the system as an intersection of m hyperplanes in n-dimensional space, we can rewrite the ith equation as

$$\boxed{\vec{a}_{i*} \cdot \vec{x} = b_i}$$

where $\vec{a}_{i*}$ is the ith row vector of the system and $\vec{x}$ is the vector of variables:

$$\vec{a}_{i*} = \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix} \quad \& \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

We have seen that this is the unique hyperplane in n-dimensional space that

- is perpendicular to $\vec{a}_{i*}$
- has distance $b_i / \|\vec{a}_{i*}\|$ from $\vec{0}$.

Then we can express the system of hyperplanes as

$$\begin{cases} \vec{a}_{1*} \cdot \vec{x} = b_1 \\ \vec{a}_{2*} \cdot \vec{x} = b_2 \\ \quad \vdots \\ \vec{a}_{m*} \cdot \vec{x} = b_m \end{cases}$$

///

2. Column Picture.

When thinking of the system as a linear combination of n vectors in m-dimensional space, we can rewrite it as a single vector equation

$$\boxed{x_1 \vec{a}_{*1} + x_2 \vec{a}_{*2} + \cdots + x_n \vec{a}_{*n} = \vec{b}}$$

where $\vec{a}_{*j}$ is the j^{th} column vector of the system and $\vec{b}$ is the vector of constants:

$$\vec{a}_{*j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad \& \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \\ b_m \end{pmatrix}.$$

///

But both of these notations still use a lot of symbols. Wouldn't it be nice if we could

- simplify the notation even further,
- express the row & column pictures simultaneously?

This is exactly what the language of matrix algebra does for us. In this language we will express the system simply as

$$\boxed{A\vec{x} = \vec{b}}$$

Where A is a rectangular array called the matrix of coefficients:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

To write it out fully, we have

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_m \end{pmatrix}$$

To emphasize that m & n may be different, here is a concrete example from HW3 Problem 4. We can rewrite the system of 3 equations in 6 unknowns

$$\begin{cases} 0 + x_2 + 0 + x_4 - x_5 - 4x_6 = -1 \\ x_1 + 2x_2 - x_3 + 4x_4 - x_5 - 4x_6 = 3 \\ x_1 + 2x_2 - x_3 + 4x_4 + 0 - x_6 = 5 \end{cases}$$

as a single "matrix equation":

$$\begin{pmatrix} 0 & 1 & 0 & 4 & -1 & -4 \\ 1 & 2 & -1 & 4 & -1 & -4 \\ 1 & 2 & -1 & 4 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix}.$$

So what, you may ask. So we can express a linear system with a small number of symbols:

$$A\vec{x} = \vec{b}.$$

But does this actually help us to <u>solve</u> linear systems?

Well, yes it does. Like any good notation, this one gives us a new point of view and suggests new questions we can ask. For example:

The expression "$A\vec{x}$" on the left looks kind of like "multiplation". This suggests that maybe we could also "divide" to get

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

$$\Longrightarrow \quad \vec{x} = \text{"}\frac{1}{A}\vec{b}\text{"}$$

and that would be pretty cool....

In fact we *will* learn how to do something like this but it will take us several weeks to make sense of it.

But don't feel bad. It took the human race thousands of years to take this step, so several weeks is actually pretty fast.

The whole key to the language of matrix algebra is the concept of

"matrix multiplication".

Stay tuned.

2/22/16

HW 4 due next Mon Feb 29
Exam 1 next Fri Mar 4.

Last time I introduced the "matrix" notation, in which we rewrite a system of linear equations

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \qquad\qquad \vdots \qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

as a single "matrix equation"

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

and then we replace each matrix and vector by a single symbol to write

$$\boxed{A\vec{x} = \vec{b}}$$

///

☆ Definition: A matrix is just a rectangular (or square) array of numbers. If the matrix has m rows and n columns we say it has shape $m \times n$.

We can also think of a vector with n components as a matrix of shape $n \times 1$. (By convention we always think of vectors as column matrices.)

[Remark: The word "matrix" (Latin for "womb") was first applied to a rectangle of numbers by James Joseph Sylvester in 1850. Pretty recently!]

///

Thus the matrix notation is really just a rule for "multiplying" an $m \times n$ matrix A by an $n \times 1$ matrix/vector $\vec{x}$ to obtain an $m \times 1$ matrix/vector that we call "$A\vec{x}$".

Example: Let $A = \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 2 & 3 & 4 \end{pmatrix}}_{2 \times 3}$ & $\vec{x} = \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{3 \times 1}$.

Then by definition we have

$$\begin{pmatrix} 1 & 0 & -1 \\ 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 0 - y \\ 2x + 3y + 4z \end{pmatrix}.$$

$$2 \times 3 \quad \underset{\text{match}}{} \quad 3 \times 1 \qquad\qquad 2 \times 1$$

As a special case we can multiply a $1 \times n$ matrix/row by an $n \times 1$ matrix/column to obtain a 1×1 matrix/number.

Example:

$$\begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_1 x_1 + a_2 x_2 + a_3 x_3 \end{pmatrix}.$$

Yes, you are right. This is just a fancy way to write the dot product of vectors.

To be explicit, we will define the concept of the "transpose" of a matrix.

☆ Definition: Let A be an $m \times n$ matrix with number a_{ij} in the i^{th} row & j^{th} column (we call this the (i,j) entry of the matrix):

$$A = \begin{array}{cc} & \begin{array}{ccccc} \phantom{a_{11}} & & j & & \phantom{a_{1n}} \end{array} \\ i & \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}}_{n \text{ columns}} \end{array} \left.\vphantom{\begin{pmatrix} a \\ a \\ a \\ a \\ a \end{pmatrix}}\right\} m \text{ rows}$$

Then we define the <u>transpose</u> matrix A^T as the matrix of shape $n \times m$ with entry a_{ij} in the j^{th} <u>row</u> & i^{th} <u>column</u>:

$$A^T = \begin{array}{cc} & \begin{array}{ccccc} \phantom{a_{11}} & & i & & \phantom{a_{m1}} \end{array} \\ j & \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{i1} & \cdots & a_{m1} \\ \vdots & & \vdots & & \vdots \\ a_{1j} & \cdots & a_{ij} & \cdots & a_{mj} \\ \vdots & & \vdots & & \vdots \\ a_{1n} & \cdots & a_{in} & \cdots & a_{mn} \end{pmatrix}}_{m \text{ columns}} \end{array} \left.\vphantom{\begin{pmatrix} a \\ a \\ a \\ a \\ a \end{pmatrix}}\right\} n \text{ rows}$$

Pretty basic idea, but it's useful because it allows us to turn the dot product of vectors into a product of matrices.

Consider two vectors (i.e. column matrices)

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad \& \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Then if we think of a 1×1 matrix as just a number, we have

$$\vec{a} \cdot \vec{x} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

$$= \left(a_1 x_1 + \cdots + a_n x_n \right)$$

$$= \left(a_1 \cdots a_n \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \vec{a}^T \vec{x} .$$

In summary,

$$\boxed{\vec{a} \bullet \vec{x} = \vec{a}^T \vec{x}}$$

dot product matrix product.

Now we have an effective way to describe the two pictures of matrix notation.

Let A be an $m \times n$ and let $\vec{x}$ be an $n \times 1$ matrix/column/vector.

1. Column Picture.

Let $\vec{a_{*j}}$ be the j^{th} column vector of A (which has shape $m \times 1$). Then we have

$$A\vec{x} = \underbrace{x_1 \vec{a_{*1}} + x_2 \vec{a_{*2}} + \cdots + x_n \vec{a_{*n}}}_{\text{a "linear combination" of the columns of } A.}$$

2. Row Picture.

Let $\vec{a}_{i*}$ be the i^{th} row vector (which we think of as an $n\times 1$ column, because we think of <u>every</u> vector as a column). Then

$$A\vec{x} = \begin{pmatrix} \vec{a}_{1*}^{\,T}\,\vec{x} \\ \vec{a}_{2*}^{\,T}\,\vec{x} \\ \vdots \\ \vec{a}_{m*}^{\,T}\,\vec{x} \end{pmatrix}.$$

///

It's very important that you understand these two different ways to compute $A\vec{x}$.

Example: Let $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & -2 & 1 \\ 3 & 1 & 0 & -2 \end{pmatrix}$ and

$$\vec{x} = (1,2,3,4) = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}^T.$$

Compute the product $A\vec{x}$ in two ways.

1. Column Picture.

$$A\vec{x} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & -2 & 1 \\ 3 & 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

$$= 1\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} + 3\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} + 4\begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} + \begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ -6 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 4 \\ -8 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ -3 \end{pmatrix}.$$

2. Row Picture.

$$A\vec{x} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & -2 & 1 \\ 3 & 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

$$= \begin{pmatrix} (1\ 0\ 1\ 0)\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \\ (0\ 2\ -2\ 1)\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \\ (3\ 1\ 0\ -2)\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1+0+3+0 \\ 0+4-6+4 \\ 3+2+0-8 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ -3 \end{pmatrix}$$

same answer ✓

So far we have only defined the product

$$A\vec{x}$$

when A is an $m \times n$ matrix and $\vec{x}$ is an $n \times 1$ column. Next time we will think about how to define the product

$$\text{"}AB\text{"}$$

when B is a more general kind of matrix. Buckle your seatbelts; it will involve a radically new point of view.

2/24/16

HW 4 due Mon
Exam 1 Fri Mar 4 in class.
[I'll distribute practice exams
this Friday.]

We have decided two write a system of
m linear equations in n unknowns as
a single matrix equation

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

where

- A is an $m \times n$ matrix of coefficients
- $\vec{x}$ is an $n \times 1$ matrix of variables
- $\vec{b}$ is an $m \times 1$ matrix of constants

Accordingly, we have two different ways
to view the matrix product "$A\vec{x}$",
coming from the row & column pictures
of the linear system. To be specific, let

$\vec{a}_{i*}$ = ith row vector of A,
$\vec{a}_{*j}$ = jth column vector of A.

1. Row Picture.

The ith entry of the vector $A\vec{x}$ is

$$\vec{a_{i*}} \cdot \vec{x}.$$

2. Column Picture

The vector $A\vec{x}$ is a linear combination of the column vectors of A,

$$A\vec{x} = x_1 \vec{a_{*1}} + x_2 \vec{a_{*2}} + \cdots + x_n \vec{a_{*n}}.$$

////

Today we will consider the following question:

☆ Is it possible to define ~~a~~ the "product" of two matrices

"AB"

when B is not just a single column?

A quick glance at Wikipedia (or a textbook) shows that the answer is <u>yes</u>.

You will also see that the product of matrices looks like a mess of symbols. So here's a better question:

☆ Why would we *want* to multiply matrices, and how *should* we think about the definition of matrix multiplication?

The reason we *want* to multiply matrices is to give us new tools for *solving* systems of equations:

$$\text{“} A\vec{x} \text{''} = \vec{b} \implies \vec{x} = \text{“} \frac{1}{A}\vec{b} \text{''} \;??$$

But to understant what matrix multiplication *should* be requires a radically new and modern point of view (first stated by Arthur Cayley in 1858).

[Remark: Is 1858 modern? Yes, by mathematical standards. The Calculus was invented in the 1660s.]

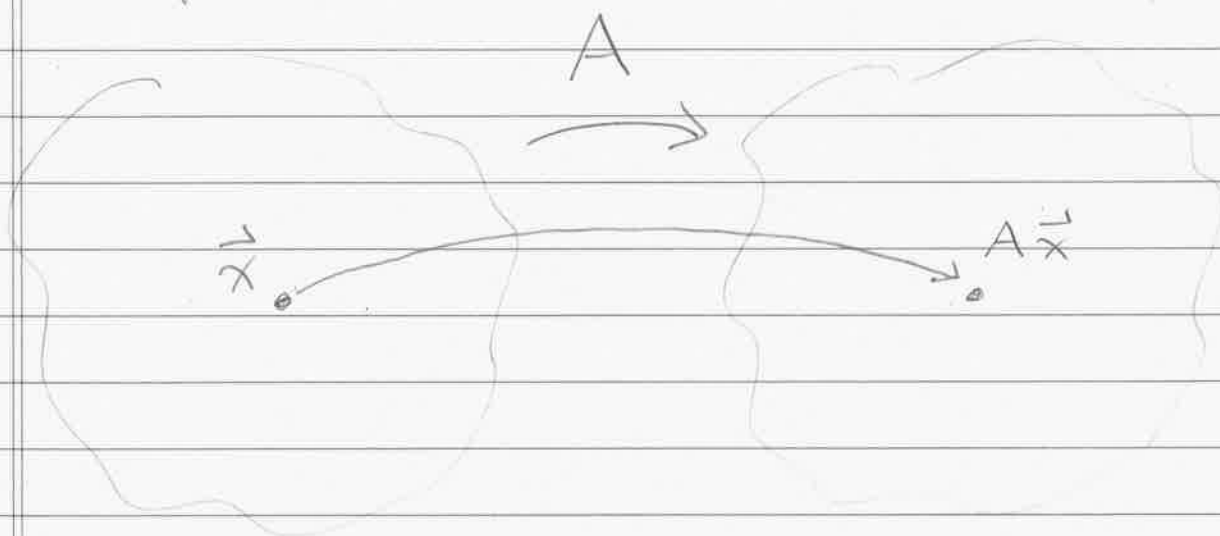
☆ Modern Point of View:

We will think of an $m \times n$ matrix A as a <u>function</u> that accepts a vector $\vec{x}$ with n coordinates and spits out a vector $A\vec{x}$ with m coordinates.

Picture:

n-dimensional space

m-dimensional space.



Thus if A is <u>square</u> $(m=n)$, we can think of A as a function sending vectors to vectors in the <u>same space</u>

Let's try some examples.

Example: The 2×2 matrix

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

sends points in the plane to points in the plane. What does it <u>do</u> to the points?

Given a point $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, the function I sends it to the point

$$I\vec{x} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$= x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Thus I sends every point to itself! We call this the "identity function" (or the "do-nothing function" on the Cartesian plane.

Example: How about the function

$$F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ?$$

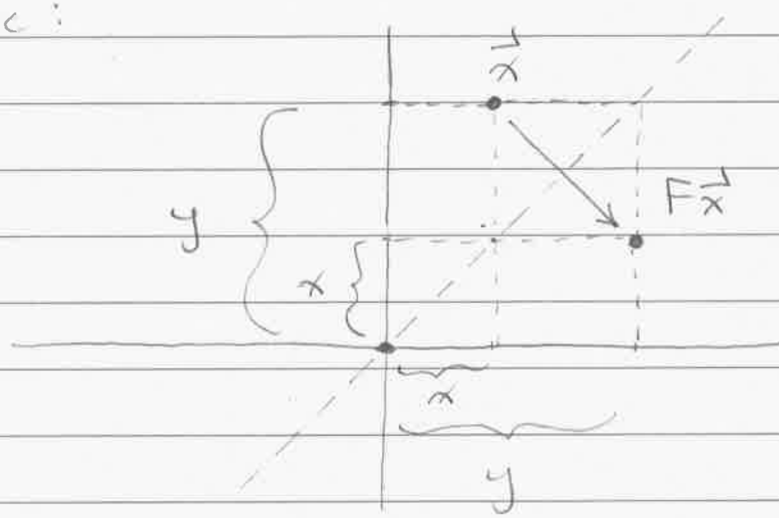
The Function F sends the point $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ to

$$F\vec{x} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= x\begin{pmatrix} 0 \\ 1 \end{pmatrix} + y\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ x \end{pmatrix} + \begin{pmatrix} y \\ 0 \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}.$$

It switches the two coordinates.
Geometrically, we can think of this as a <u>reflection</u> across the line of slope 1.

Picture:



[Remark: F is for "Flip" or "reFlection".]

And what happens if we do F twice in succession?

Let's check:

$$F(F\vec{x}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right]$$

$$= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y \\ x \end{pmatrix}$$

$$= y \begin{pmatrix} 0 \\ 1 \end{pmatrix} + x \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ y \end{pmatrix} + \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} = \vec{x}.$$

Does that surprise you? NO. If we reflect across the line and <u>then</u> reflect <u>again</u> we get back where we started. ///

To summarize these examples: For any point $\vec{x}$ in the plane we have

- $I\vec{x} = \vec{x}$
- $F(F\vec{x}) = \vec{x}$,

and hence

$$F(F\vec{x}) = I\vec{x}.$$

Now I'm really tempted to rearrange the parentheses and write

(*) $$(FF)\vec{x} = I\vec{x},$$

but is that allowed? Well, no because the expression "FF" is not defined.

OK, no problem. Let's just <u>define</u> $FF := I$,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

And now the equation (*) is perfectly true.

Congratulations: We just "multiplied" two 2×2 matrices to obtain another 2×2 matrix. <u>And</u> we understand what it <u>means</u>.

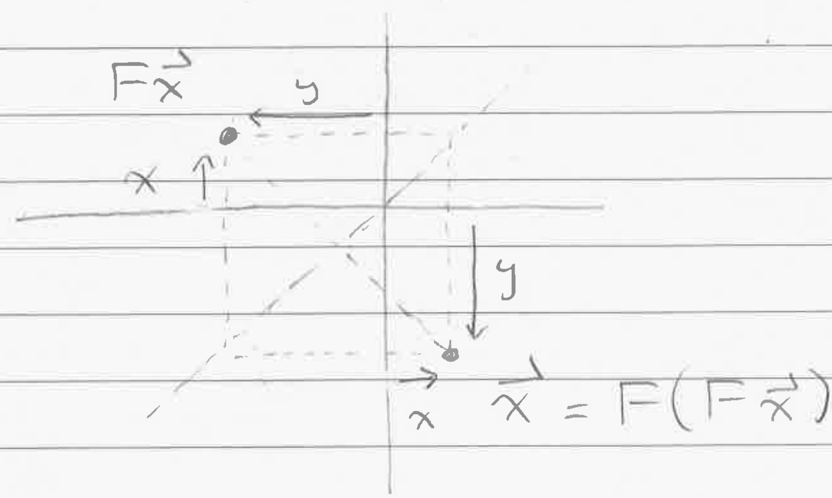
"Reflecting across the same line twice is the same as doing nothing once."

2/24/16

HW 4 due Mon
Review session Wed
Exam 1 Fri.

Last time we saw our first example of a matrix product "AB" when B is not a single column.

Recall: The matrix $F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ sends the point $\vec{x} = (x, y)$ to the point $F\vec{x} = (y, x)$, which geometrically is a reflection across the line of slope 1:



If we perform the reflection again then we arrive back where we started:

$$F(F\vec{x}) = \vec{x}.$$

Thus "doing F twice" is the same as "doing nothing once" and we know that the "doing nothing" function is represented by the identity matrix

$$I\vec{x} = x$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

So we conclude that for all points $\vec{x}$ in the plane we have

$$F(F\vec{x}) = I\vec{x}.$$

and this makes it perfectly clear how we should define the matrix $F^2 = FF$: we just rearrange the parentheses,

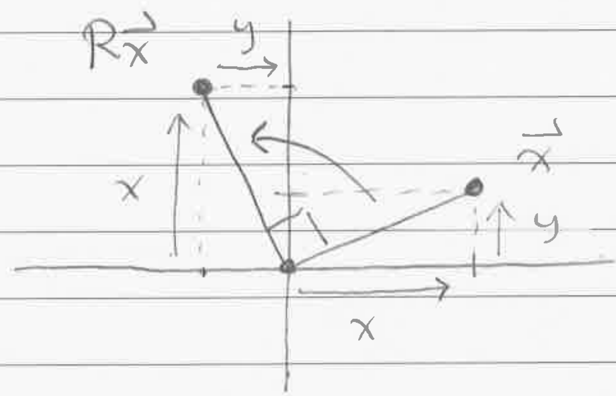
$$(FF)\vec{x} = I\vec{x}$$

and declare that $FF = I$, i.e.,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Another Example: Let R be the function that rotates each vector in the plane counterclockwise by $90°$. Can we represent R as a matrix?

Note that for all points $\vec{x} = (x, y)$ we must have $R\vec{x} = (-y, x)$ because of the following picture:



That is, for all $\vec{x} = (x, y)$ we must have

$$R\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} = \begin{pmatrix} 0x - 1y \\ 1x + 0y \end{pmatrix}$$

and there is a unique matrix that does this:

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Can we compute the matrix $R^2 = RR$?

Sure: For all $\vec{x} = (x, y)$ we have

$$R\left(R\begin{pmatrix} x \\ y \end{pmatrix}\right) = R\begin{pmatrix} -y \\ x \end{pmatrix} = \begin{pmatrix} -(x) \\ (-y) \end{pmatrix}.$$

So we must have

$$(RR)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix} = \begin{pmatrix} -1x + 0y \\ 0x - 1y \end{pmatrix}$$

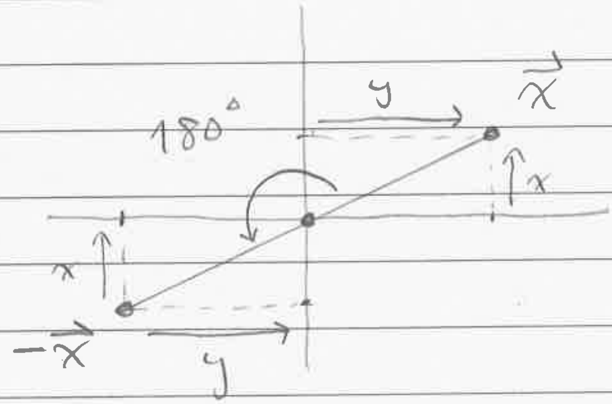
The unique solution is

$$R^2 = RR = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = -I.$$

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

Does this make sense? We know that rotating c.c.w. <u>twice by 90°</u> is the same as rotating c.c.w. <u>once by 180°</u>.

And <u>does</u> the matrix $-I$ rotate the plane c.c.w. by 180° ?



$-I\vec{x} = -\vec{x}$

Yes it does!

[Remark: We are starting to see some kind of "matrix algebra" emerging. Can you predict what the matrix R^3 is without doing any more computations?]

Ok, now let's try to compute the product of two <u>general</u> 2x2 matrices:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \& \quad B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}.$$

For all points $\vec{x} = (x, y)$ we have

$$A(B\vec{x}) = A \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$= A \begin{pmatrix} a'x + b'y \\ c'x + d'y \end{pmatrix}$$

$$= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a'x + b'y \\ c'x + d'y \end{pmatrix}$$

$$= \begin{pmatrix} a(a'x + b'y) + b(c'x + d'y) \\ c(a'x + b'y) + d(c'x + d'y) \end{pmatrix}$$

$$= \begin{pmatrix} (aa' + bc')x + (ab' + bd')y \\ (ca' + dc')x + (cb' + dd')y \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}}_{\text{call this } C} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$= C\vec{x}.$$

[No thinking was involved here; all of this was forced on me by the original definition of matrix $\times$ vector.]

We conclude that for all $\vec{x}$ we have

$$A(B\vec{x}) = C\vec{x},$$

thus we should define the matrix AB so that the equation

$$(AB)\vec{x} = C\vec{x}$$

is true for all $\vec{x}$. In other words, we <u>should</u> define

$$AB := C.$$

$$\bigstar \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} := \begin{pmatrix} aa' + bd' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$$

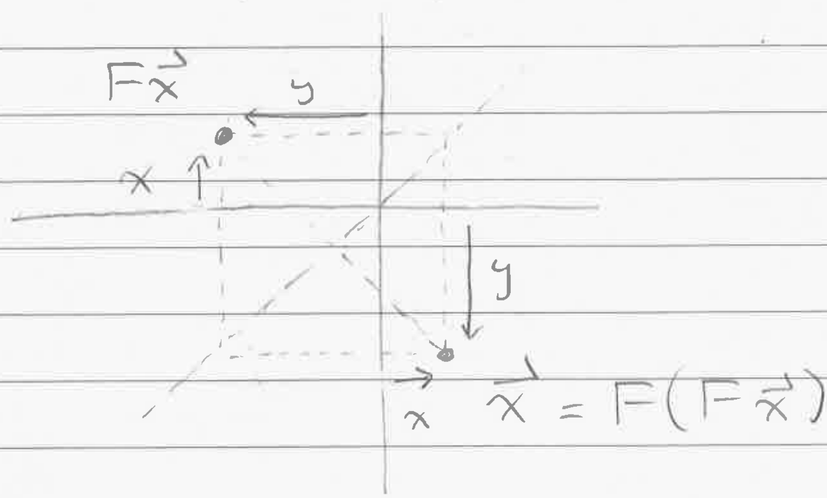
and this is how we <u>will</u> define it.

2/24/16

HW 4 due Mon
Review session Wed
Exam 1 Fri.

Last time we saw our first example of a matrix product "AB" when B is not a single column.

Recall: The matrix $F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ sends the point $\vec{x} = (x, y)$ to the point $F\vec{x} = (y, x)$, which geometrically is a reflection across the line of slope 1:



If we perform the reflection again then we arrive back where we started:

$$F(F\vec{x}) = \vec{x}.$$

Thus "doing F twice" is the same as "doing nothing once" and we know that the "doing nothing" function is represented by the identity matrix

$$I\vec{x} = x$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

So we conclude that for all points $\vec{x}$ in the plane we have

$$F(F\vec{x}) = I\vec{x}.$$

and this makes it perfectly clear how we should <u>define</u> the matrix $F^2 = FF$: we just rearrange the parentheses,

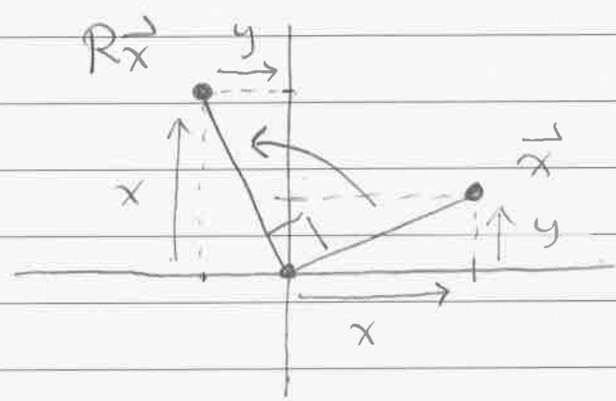
$$(FF)\vec{x} = I\vec{x}$$

and declare that $FF = I$, i.e.,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Another Example: Let R be the function that <u>rotates</u> each vector in the plane <u>counterclockwise</u> by $90°$. Can we represent R as a matrix?

Note that for all points $\vec{x} = (x, y)$ we must have $R\vec{x} = (-y, x)$ because of the following picture:



That is, for all $\vec{x} = (x, y)$ we must have

$$R\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} = \begin{pmatrix} 0x - 1y \\ 1x + 0y \end{pmatrix}$$

and there is a unique matrix that does this:

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Can we compute the matrix $R^2 = RR$?

Sure: For all $\vec{x} = (x, y)$ we have

$$R\left(R\begin{pmatrix} x \\ y \end{pmatrix}\right) = R\begin{pmatrix} -y \\ x \end{pmatrix} = \begin{pmatrix} -(x) \\ (-y) \end{pmatrix}.$$

So we must have

$$(RR)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix} = \begin{pmatrix} -1x + 0y \\ 0x - 1y \end{pmatrix}$$

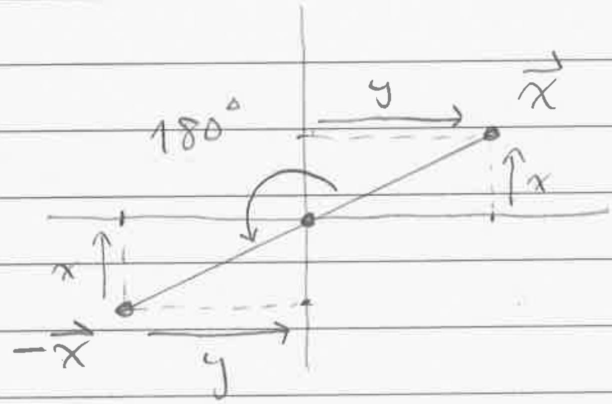
The unique solution is

$$R^2 = RR = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = -I.$$

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

Does this make sense? We know that rotating c.c.w. <u>twice by $90°$</u> is the same as rotating c.c.w. <u>once by $180°$</u>.

And <u>does</u> the matrix $-I$ rotate the plane c.c.w. by $180°$?



$-I\vec{x} = -\vec{x}$

Yes it does!

[Remark: We are starting to see some kind of "matrix algebra" emerging. Can you predict what the matrix R^3 is without doing any more computations?]

Ok, now let's try to compute the product of two general 2x2 matrices:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \& \quad B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}.$$

For all points $\vec{x} = (x, y)$ we have

$$A(B\vec{x}) = A \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$= A \begin{pmatrix} a'x + b'y \\ c'x + d'y \end{pmatrix}$$

$$= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a'x + b'y \\ c'x + d'y \end{pmatrix}$$

$$= \begin{pmatrix} a(a'x + b'y) + b(c'x + d'y) \\ c(a'x + b'y) + d(c'x + d'y) \end{pmatrix}$$

$$= \begin{pmatrix} (aa' + bc')x + (ab' + bd')y \\ (ca' + dc')x + (cb' + dd')y \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}}_{\text{call this } C} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$= C\vec{x}.$$

[No thinking was involved here; all of this was forced on me by the original definition of matrix × vector.]

We conclude that for all $\vec{x}$ we have

$$A(B\vec{x}) = C\vec{x},$$

thus we should define the matrix AB so that the equation

$$(AB)\vec{x} = C\vec{x}$$

is true for all $\vec{x}$. In other words, we <u>should</u> define

$$AB := C.$$

$\star$ $$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} := \begin{pmatrix} aa' + bd' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$$

and this is how we <u>will</u> define it.

2/29/16

HW4 due now.

Review session on Wed

[I'll also distribute solutions for the 2010 practice exam on Wed.]

Exam 1 in class on Fri

- closed book
- cheaters will receive 0 pts.

Today: HW4 Discussion.

Let A be an $m \times n$ matrix and let B be an $n \times p$ matrix. Then for all $p \times 1$ vectors $\vec{x}$ we can define an $m \times 1$ vector as follows:

$$\underset{n \times p}{B} \; \underset{p \times 1}{\vec{x}} = n \times 1 \text{ vector.}$$

$$\underset{m \times n}{A} \, (\underset{n \times 1}{B\vec{x}}) = m \times 1 \text{ vector.}$$

Jargon: Let $\mathbb{R}$ denote the set of "real numbers", i.e., numbers that have decimal expansion. Then we will write

$$\mathbb{R}^n := \text{ordered } n\text{-tuples of real numbers}$$

and we will think of this as

"n-dimensional Cartesian space". //

Now we can draw the following schematic diagram:

$$\begin{array}{ccccc} \mathbb{R}^p & \xrightarrow{B} & \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^m \\ \vec{x} & \rightsquigarrow & B\vec{x} & \rightsquigarrow & A(B\vec{x}). \end{array}$$

But maybe there is one single matrix C that could perform the same composition of functions in one step:

$$\begin{array}{ccc} \mathbb{R}^p & \xrightarrow{C} & \mathbb{R}^m \\ \vec{x} & \rightsquigarrow & C\vec{x}. \end{array}$$

Well, we know that there certainly is a function that does this:

$$\begin{tikzcd} \mathbb{R}^p \arrow[r, "B"] \arrow[rr, bend right, "A \circ B"'] & \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m \end{tikzcd}$$

Recall that $A \circ B$ (say "A follows B") is the function obtained by first doing B then doing A. The only question is whether the function $A \circ B$ from $\mathbb{R}^p$ to $\mathbb{R}^n$ can be represented by a matrix (necessarily a matrix of shape $m \times p$).

And we know by now that it can. So we make the following definition.

☆ Definition: Given an $m \times n$ matrix A and an $n \times p$ matrix B we define

"AB"

to be the $m \times p$ matrix the represents the composite function $A \circ B$.

In short, we define the matrix "AB" that for all $\vec{x}$ in $\mathbb{R}^p$ the following equation is true:

$$(AB)\vec{x} := A(B\vec{x}).$$

So that's it. Now you've seen the definition of matrix multiplication.

Example (HW 4 Problem 3):

Compute the matrix product AB when

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & -1 & 0 \end{pmatrix} \ \& \ B = \begin{pmatrix} 3 & 1 \\ 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

For all $\vec{x} = (x, y)$ we must have

$$(AB)\begin{pmatrix} x \\ y \end{pmatrix} = A\left(B\begin{pmatrix} x \\ y \end{pmatrix}\right)$$

$$= \begin{pmatrix} 1 & 0 & 2 \\ 1 & -1 & 0 \end{pmatrix}\left[x\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right]$$

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

$$= (3x+y)\begin{pmatrix} 1 \\ 1 \end{pmatrix} + x\begin{pmatrix} 0 \\ -1 \end{pmatrix} + (-y)\begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 3x + y - 2y \\ 3x + y - x \end{pmatrix} = \begin{pmatrix} 3x - 1y \\ 2x + 1y \end{pmatrix}$$

$$= \begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

so we conclude that

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

///

OK, but that's pretty tedious. Isn't there some kind of shortcut or trick for multiplying matrices?

Sure there is. Last time we used the definition to compute the product of two general 2×2 matrices:

$$(*) \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}.$$

If you can memorize this formula then you can get to the answer much more quickly. The only problem is: how can we memorize the formula?

Here's a helpful computational trick.

If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ & $B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$

note from the formula $(*)$ that

$$AB = \begin{pmatrix} (a \ b)\begin{pmatrix} a' \\ c' \end{pmatrix} & (a \ b)\begin{pmatrix} b' \\ d' \end{pmatrix} \\ (c \ d)\begin{pmatrix} a' \\ c' \end{pmatrix} & (c \ d)\begin{pmatrix} b' \\ d' \end{pmatrix} \end{pmatrix}.$$

$$= \begin{pmatrix} (\text{1st row } A) \cdot (\text{1st col } B) & (\text{1st row } A) \cdot (\text{2nd col } B) \\ (\text{2nd row } A) \cdot (\text{1st col } B) & (\text{2nd row } A) \cdot (\text{2nd col } B) \end{pmatrix}$$

The nice trick is that it works for matrices arbitrary shape.

☆ TRICK for computing matrix products.

Let A & B be matrices. If the product is defined then we have

$$\boxed{(i,j) \text{ entry of } AB = (i\text{th row of } A) \circ (j\text{th col of } B)}$$

Let's try the trick on the example from HW4 Problem 3. We have

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

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

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

$$= \begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix} \checkmark \quad \text{It works!}$$

While were at it, let's compute BA.

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

$$= \begin{pmatrix} 3+1 & 0-1 & 6+0 \\ 1+0 & 0+0 & 2+0 \\ 0-1 & 0+1 & 0+0 \end{pmatrix}$$

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

You can take my word for it that this is the correct answer;

I'm not going to compute it the long way using the definition

$$B(A\vec{x}) = (BA)\vec{x}.$$

In fact, we'll *never* compute it the long way again! [Unless I specifically ask you to do so on an exam problem :)]

///

Remark: Note that the matrices AB & BA above are both defined and they are both square, but of different *sizes*.

[see HW2 Problem 4(b).]

We saw on HW2 Problem 5 that even if AB & BA are both defined and *have the same size*, they are still not necessarily equal.