

Section 5.3 Applications of the Determinant

The determinant of a matrix can tell us a lot about a linear transformation, as we have seen, when non-zero we get everything in the unifying theorem we have been building up over the semester. However we can also use it to solve problems directly, understand error propagation, and changes in geometry after a linear transformation.

Cramer's rule

Let A be an invertible $n \times n$ matrix. Then the components of the unique solution to $A\vec{x} = \vec{b}$ are given by:

$$x_i = \frac{\det(A_i)}{\det(A)} \text{ for } i = 1, \dots, n$$

Where A_i has the i th column of A replaced with the vector $\vec{b}$

Example: Solve the system: $\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$$x_1 = \frac{\begin{vmatrix} 2 & 2 \\ 1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix}} = \frac{0}{(1)(1) - (-1)(2)} = 0$$

$$x_2 = \frac{\begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix}} = \frac{2}{2} = 1$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Section 5.3 Applications of the Determinant

This is easy for a 2x2 maybe even a 3x3, but determinants get harder to calculate as we go.

```

octave:18> A=randi([-5 5], 4,4)
A =
    2    4   -3    3
    1   -4    2   -4
   -4   -5    3    3
   -3   -3    1   -3

octave:22> b=randi([-4 6], 4,1)
b =
    2
    6
    0
    4

octave:23> A1=[b,A(:,2:4)]
A1 =
    2    4   -3    3
    6   -4    2   -4
    0   -5    3    3
    4   -3    1   -3

octave:26> A2=[A(:,1),b,A(:,3:4)]
A2 =
    2    2   -3    3
    1    6    2   -4
   -4    0    3    3
   -3    4    1   -3

octave:27> A3=[A(:,1:2),b,A(:,4)]
A3 =
    2    4    2    3
    1   -4    6   -4
   -4   -5    0    3
   -3   -3    4   -3

octave:28> A4=[A(:,1:3),b]
A4 =
    2    4   -3    2
    1   -4    2    6
   -4   -5    3    0
   -3   -3    1    4

octave:29> x1=det(A1)/det(A)
x1 = 64/75
octave:30> x2=det(A2)/det(A)
x2 = -296/75
octave:31> x3=det(A3)/det(A)
x3 = -27/5
octave:32> x4=det(A4)/det(A)
x4 = -1/25

octave:33> x=[x1;x2;x3;x4]
x =
    64/75
   -296/75
   -27/5
   -1/25

octave:34> A*x
ans =
    2
    6
    0
    4

octave:35> inv(A)*b
ans =
    64/75
   -296/75
   -27/5
   -1/25
  
```

Why does Cramer's Rule work?

Consider replacing the i^{th} column of the identity matrix with our solution

$$I_i = \begin{bmatrix} 1 & 0 & \dots & x_1 & \dots & 0 \\ 0 & 1 & 0 & x_2 & \dots & \vdots \\ \vdots & 0 & \ddots & \vdots & \dots & 0 \\ 0 & \dots & 0 & x_i & \dots & 0 \\ 0 & \dots & 0 & \vdots & \dots & 0 \\ 0 & 0 & 0 & x_n & \dots & 1 \end{bmatrix} \quad \text{Then } AI_i = A \begin{bmatrix} e_1 & e_2 & \dots & \vec{x} & \dots & e_n \end{bmatrix} = \begin{bmatrix} Ae_1 & Ae_2 & \dots & A\vec{x} & \dots & Ae_n \end{bmatrix} = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{b} & \dots & \vec{a}_n \end{bmatrix} = A_i$$

Then $\det(A_i) = \det(AI_i) = \det(A) \det(I_i)$ or just that $\det(I_i) = \frac{\det(A_i)}{\det(A)}$ So, what is $\det(I_i)$?

Just expand across the i^{th} row!

$$\begin{bmatrix} 1 & 0 & \dots & x_1 & \dots & 0 \\ 0 & 1 & 0 & x_2 & \dots & \vdots \\ \vdots & 0 & \ddots & \vdots & \dots & 0 \\ 0 & \dots & 0 & x_i & \dots & 0 \\ 0 & \dots & 0 & \vdots & \dots & 0 \\ 0 & 0 & 0 & x_n & \dots & 1 \end{bmatrix} \rightarrow 0 + \dots + x_i(-1)^{i+i} \det(I) + 0 + \dots + 0 = x_i(-1)^{2i}(1) = x_i = \det(I_i) = \frac{\det(A_i)}{\det(A)}$$

Section 5.3 Applications of the Determinant

If A is an invertible matrix, then the inverse of A may be calculated as,

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$$

Where $\text{adj}(A) = [C_{ji}]$, the matrix made of the cofactors of A with i and j switched.

Proof:

$$\begin{aligned} A(\text{adj}(A)) &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}C_{11} + \cdots + a_{1n}C_{1n} & a_{11}C_{21} + \cdots + a_{1n}C_{2n} & \cdots & a_{11}C_{n1} + \cdots + a_{1n}C_{nn} \\ a_{21}C_{11} + \cdots + a_{2n}C_{1n} & a_{21}C_{21} + \cdots + a_{2n}C_{2n} & \cdots & a_{21}C_{n1} + \cdots + a_{2n}C_{nn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}C_{11} + \cdots + a_{nn}C_{1n} & a_{n1}C_{21} + \cdots + a_{nn}C_{2n} & \cdots & a_{n1}C_{n1} + \cdots + a_{nn}C_{nn} \end{bmatrix} \end{aligned}$$

Recall that for the expansion

$$a_{i1}C_{ji} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn}$$

If $i \neq j$ $a_{i1}C_{ji} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn} = 0$ and if $i = j$ $a_{i1}C_{ji} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn} = \det(A)$

So our multiplicative result is just

$$\begin{bmatrix} \det(A) & 0 & \cdots & 0 \\ 0 & \det(A) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \det(A) \end{bmatrix} = \det(A) \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \det(A) I$$

So $A(\text{adj}(A)) = \det(A) I$

Thus

$$A \left(\frac{1}{\det(A)} \text{adj}(A) \right) = \frac{1}{\det(A)} A(\text{adj}(A)) = \frac{1}{\det(A)} \det(A) I = I$$

A^{-1}

Section 5.3 Applications of the Determinant

Because of this factor of $\frac{1}{\det(A)}$ on our inverse, we can use the determinant as a quick gauge of how much we should be worried about measurement error in an applied problem.

Suppose you have a physical system modeled by $A\vec{x} = \vec{b}$ where $\vec{b}$ is a measurable quantity and A relates the physical quantities we want, stored in $\vec{x}$, to the measurable quantity $\vec{b}$. Further, consider the fact that no measurement is perfect and suppose the error in a given measurement is given by $\vec{e}$.

That is, we want to solve the system $A\vec{x} = \vec{b}$ but we only have the system $A\vec{x} = \vec{b} + \vec{e}$ and we may not know $\vec{e}$!

Let's consider 2 cases one where $|\det(A)| \ll 1$ (very small) and one where $|\det(A)| \gg 1$ (very big)

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$$

$$A\vec{x} = \vec{b} + \vec{e} \rightarrow \vec{x} = A^{-1}(\vec{b} + \vec{e}) = \underbrace{A^{-1}\vec{b}}_{\vec{x}_T} + \underbrace{A^{-1}\vec{e}}$$

$$= \vec{x}_T + \frac{1}{\det(A)} \text{adj}(A) \vec{e} = \vec{x}_T + \text{adj}(A) \left(\frac{1}{\det(A)} \vec{e} \right)$$

Big

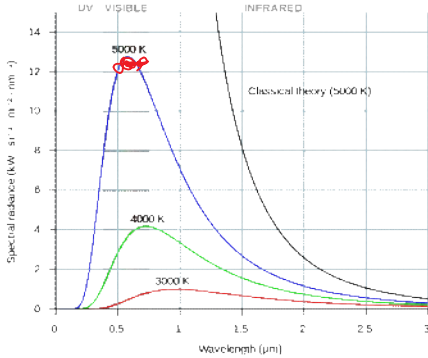
Section 5.3 Applications of the Determinant

In class FUN! Grab your laptops and head to <https://octave-online.net/bucket~RueujNQheWZQYFXvVo3kDL>
The link is also in your Heelmail sent out earlier.

Sea Ice Concentration Retrieval from Passive Microwave Brightness Temperature (TB)

Everything with a nonzero temperature emits light!

Planks Law



Planks Law

$$B(\nu, T) = \frac{2h\nu^3}{c^2} \left(e^{\frac{h\nu}{kT}} - 1 \right)^{-1}$$

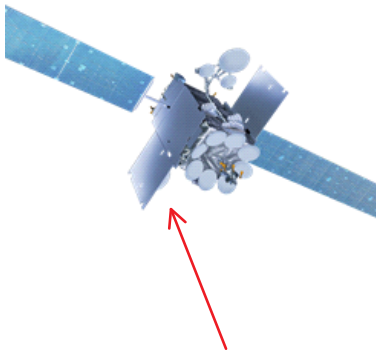
For microwaves: $\frac{h\nu}{kT} \ll 1$

$$B(\nu, T) = \frac{2h\nu^3}{c^2} \left(1 + \frac{h\nu}{kT} - 1 \right)^{-1}$$

Rayleigh-Jeans Approximation

$$\tilde{B}(\nu, T) = \frac{2kT\nu^2}{c^2}$$

Satellite
Radiance intensities
at different
frequencies
measured



Spectral Emissivity

$$\epsilon(\nu, T) = \frac{I(\nu)}{\tilde{B}(\nu, T)}$$

Brightness Temperature

$$TB = I(\nu) \left(\frac{2K\nu^2}{c^2} \right)$$

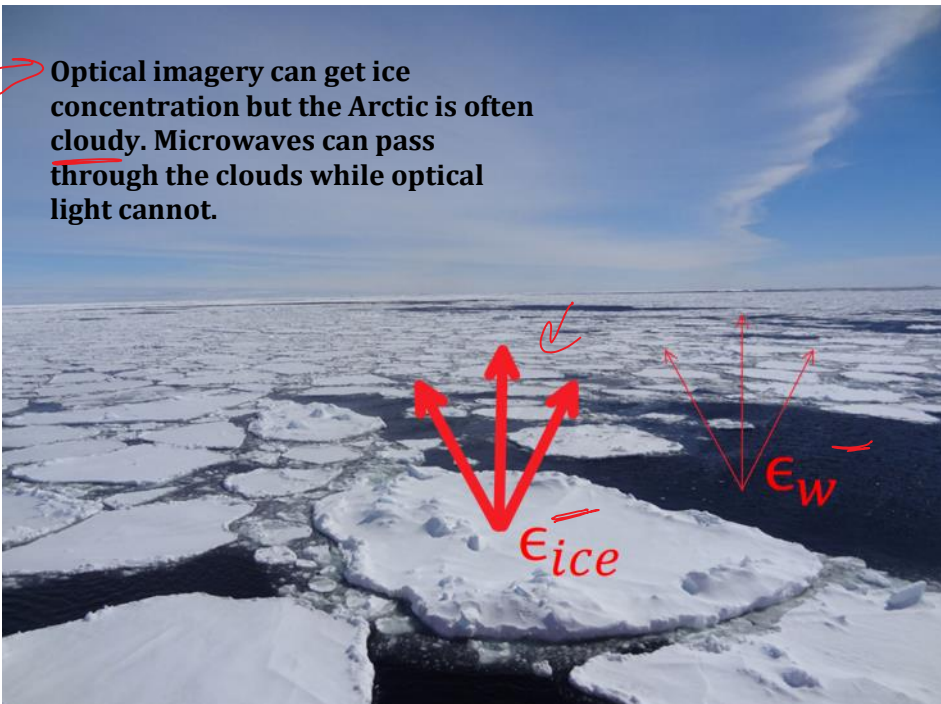
Measured by satellite...

Note that also.... $\epsilon(\nu, T) = \frac{TB}{T}$

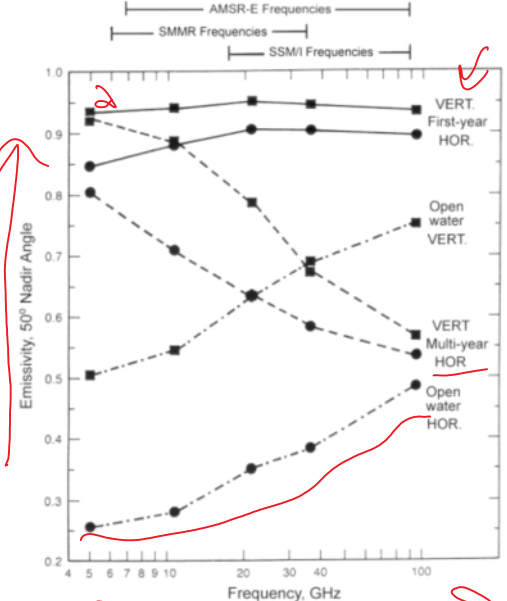
Here ν is frequency
and T temperature

$$TB = \epsilon(\nu, T)T$$

Optical imagery can get ice concentration but the Arctic is often cloudy. Microwaves can pass through the clouds while optical light cannot.



Sea ice is significantly more emissive than open water at microwave frequencies.



Massom et al

Section 5.3 Applications of the Determinant

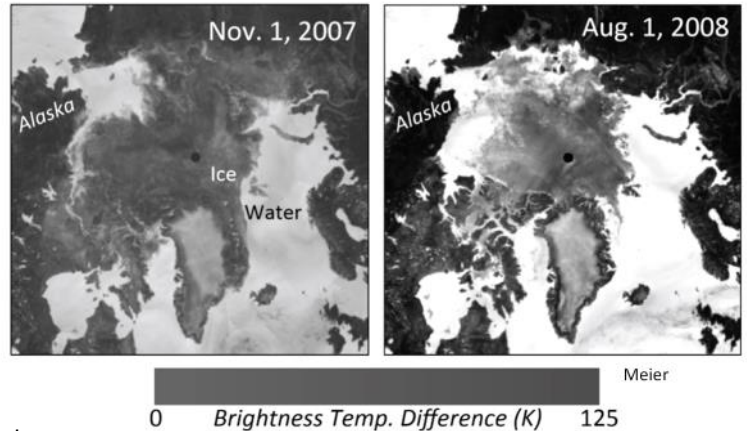
Suppose we have 40% sea ice concentration and that we have data from a passive microwave satellite which measures intensities at the 10 and 20 horizontal GHz range.

At the 10 GHz frequency $\epsilon_{ice} = 0.95$, $\epsilon_w = 0.3$

At the 20 GHz frequency $\epsilon_{ice} = 0.95$, $\epsilon_w = 0.35$

$$I(\nu) = \frac{2KT\nu^2}{c^2} \epsilon(\nu, T)$$

Intensity from emissivity



$C_{ice}I_{ice}(\nu) + C_wI_w(\nu) = I_s(\nu)$ Linear combination of ice/water intensities should add up to measured intensity

$$C_{ice} \frac{2KT\nu^2}{c^2} \epsilon_{ice}(\nu, T_1) + C_w \frac{2KT\nu^2}{c^2} \epsilon_w(\nu, T_2) = I_s(\nu) \text{ Assume their temperatures are very close. } T_1 = T_2$$

$$C_{ice}\epsilon_{ice}(\nu) + C_w\epsilon_w(\nu) = \frac{c^2}{2KT\nu^2} I_s(\nu) = \hat{I}_s(\nu) \text{ Rewrite our equation}$$

Using the values from the charts we could then say

$$0.94C_{ice} + 0.3C_w = \hat{I}_{s10} \text{ at 10 GHz channel}$$

$$0.95C_{ice} + 0.35C_w = \hat{I}_{s20} \text{ at 20 GHz channel}$$

$$\begin{bmatrix} 0.94 & 0.3 \\ 0.95 & 0.35 \end{bmatrix} \begin{bmatrix} C_{ice} \\ C_w \end{bmatrix} = \begin{bmatrix} \hat{I}_{s10} \\ \hat{I}_{s20} \end{bmatrix}$$

$$\begin{bmatrix} C_{ice} \\ C_w \end{bmatrix} = \left(\begin{bmatrix} 0.94 & 0.3 \\ 0.95 & 0.35 \end{bmatrix} \right)^{-1} \begin{bmatrix} \hat{I}_{s10} \\ \hat{I}_{s20} \end{bmatrix}$$

So we have a linear system we could solve! This is not at all how it is done in practice for reasons that will become apparent.

Hopefully by now you have gotten access to the octave bucket. In there we have some data....

Section 5.3 Applications of the Determinant

Something interesting about all of this...

```
octave:16> A=[0.94999999 0.3; 0.95 0.3]
A =
    0.95000    0.30000
    0.95000    0.30000
```

```
octave:17> det(A)
ans =   -3.0000e-09
```

```
octave:18> Ctrue=[0.8; 0.2]
Ctrue =
    0.80000
    0.20000
```

```
octave:19> Itrue=A*Ctrue
Itrue =
    0.82000
    0.82000
```

Let's Add a 10% error or so...

```
octave:20> Ie10=Itrue+[.08 ; .08]
Ie10 =
    0.90000
    0.90000
```

Invert for the concentrations.....

```
octave:21> inv(A)*Ie10
ans =
    0.80000
    0.46667
```

Not Bad!!! Did we just get lucky?
Probably not, let's try another set of errors.

Let's just add a tiny bit of error, less error than before, but differently.

```
octave:22> Ie10=Itrue+[.0001 ; -.0001]
Ie10 =
    0.82010
    0.81990
```

```
octave:23> inv(A)*Ie10
ans =
```

```
-1.9999e+04
 6.3334e+04
```

Why such a difference!?!?!?

Section 5.3 Applications of the Determinant

Well, it had to do specifically with the type of error we added and our specific matrix.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \left(\begin{bmatrix} C_{ice} \\ C_w \end{bmatrix} + \begin{bmatrix} C_{ice}^a \\ C_w^a \end{bmatrix} \right) = \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$\begin{bmatrix} C_{ice}^a \\ C_w^a \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$\begin{bmatrix} C_{ice}^a \\ C_w^a \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} de_1 - be_2 \\ -ce_1 + ae_2 \end{bmatrix}$$

```
octave:16> A=[0.94999999 0.3; 0.95 0.3]
```

```
A =
```

```
0.95000    0.30000
0.95000    0.30000
```

what a, b, c, d were...

```
octave:17> det(A)
```

```
ans =   -3.0000e-09
```

$$\begin{bmatrix} C_{ice}^a \\ C_w^a \end{bmatrix} = \frac{-1}{3 \times 10^{-9}} \begin{bmatrix} 0.3e_1 - 0.3e_2 \\ -0.95e_1 + 0.949e_2 \end{bmatrix}$$

What happens when $e_1 = e_2$ here?

The upshot is, when the error was the same the fact that our columns were almost linearly dependent meant that the resulting error in concentration was very small, since that sum in the vector would be very close to zero when $e_1 = e_2$. Note in fact, it was for the ice concentration! However when we added very small error, but different numbers, the sum in the vector was not all that close to zero and one over that tiny determinant blew up the error in the retrieved concentration. It should be noted, we could have an $n \times n$ system for which only two of the columns are very similar, in that case this effect would be mitigated by the other columns. All of this points to a need to understand how the matrix itself will affect error in these kinds of calculations, we will expand on this when we talk about the singular value decomposition.

Section 5.3 Applications of the Determinant

Area and Volume Scaling Under Linear Transformations

Let S be the unit square in the first quadrant of $\mathbb{R}^2$, and let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ with $T(\vec{x}) = A\vec{x}$. If $S' = T(S)$ is the image of S under T , then

$$\text{area}(S') = |\det(A)|$$

Let's look at simple shear example:

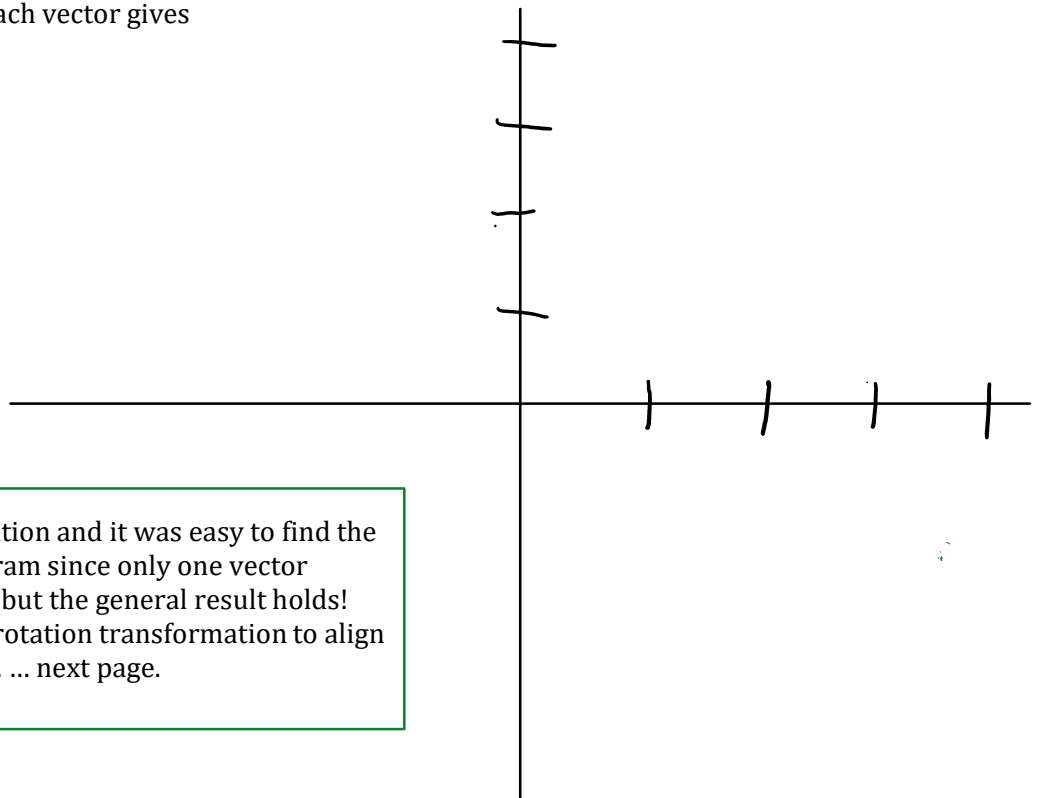
$$A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$$

The vectors that define the unit square are $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Applying the transformation to each vector gives

$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$$



In this example there was no rotation and it was easy to find the height of the resulting parallelogram since only one vector changed its angle with the x-axis, but the general result holds! The trick to see this is to apply a rotation transformation to align one of the vectors with the x-axis. ... next page.

To make use of the rotation first we need the fact that

$$\det \left(\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \right) = \cos(\theta) \cos(\theta) - \sin(\theta) (-\sin(\theta)) = \cos^2(\theta) + \sin^2(\theta) = 1$$

Section 5.3 Applications of the Determinant

What do I mean? Well first apply T....

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$$

Then Rotate by angle theta to align one of the vectors with the x-axis

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

A new transformation combining the two would, by design, have the form:

$$\begin{bmatrix} a' & b' \\ 0 & d' \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a' \\ 0 \end{bmatrix} \quad \begin{bmatrix} a' & b' \\ 0 & d' \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b' \\ d' \end{bmatrix}$$

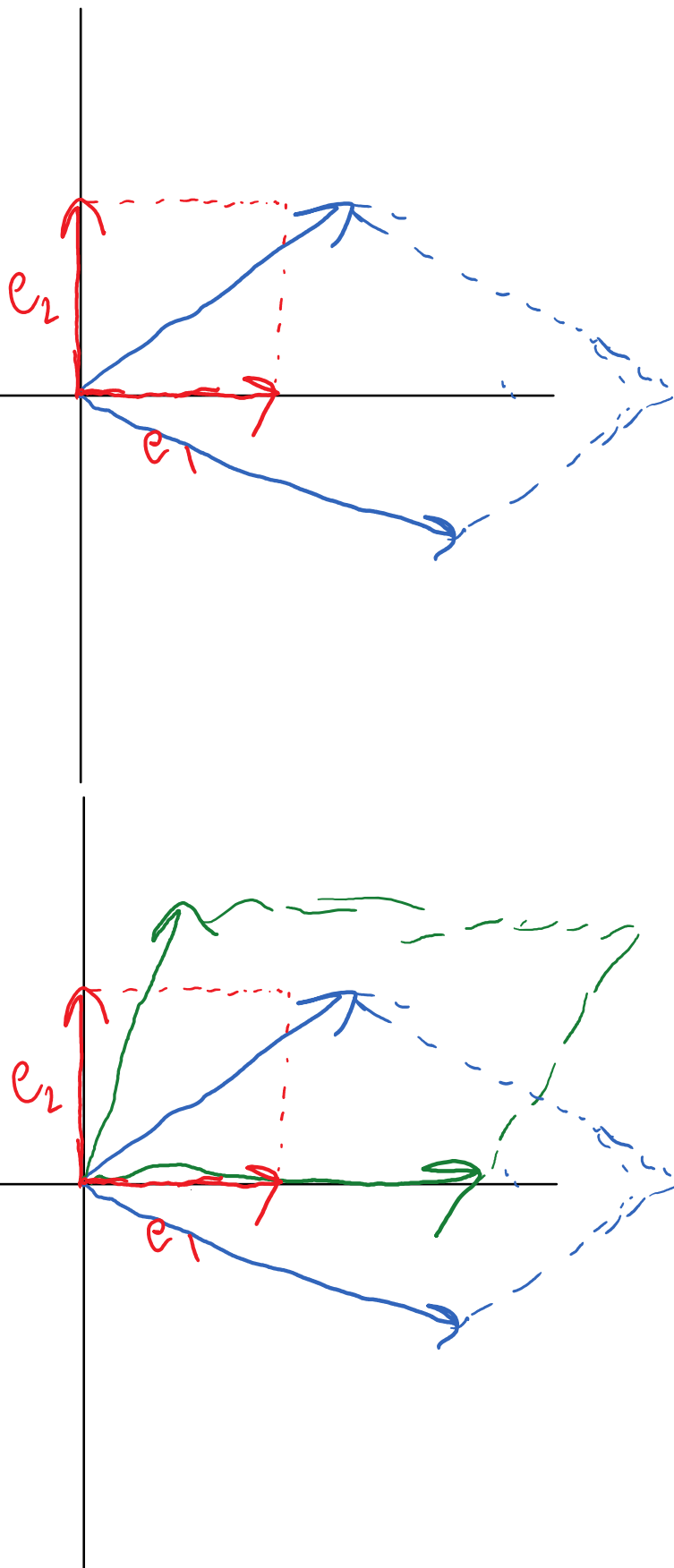
For this case the area of the green parallelogram, which is the same as the blue one would just be:

$$\det \left(\begin{bmatrix} a' & b' \\ 0 & d' \end{bmatrix} \right) = a'd'$$

Just as we had before

But.....

$$\det \left(\begin{bmatrix} a' & b' \\ 0 & d' \end{bmatrix} \right) = \det \left(\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = \det \left(\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \right) \det \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = (1) \det \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = \det(A)$$



Section 5.3 Applications of the Determinant

Let D be a region of finite Area in $\mathbb{R}^2$, and suppose that $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ with $T(\vec{x}) = A\vec{x}$. If $T(D)$ is the image of D under T , then

$$\text{area}(T(D)) = |\det(A)| \cdot \text{area}(D)$$

The argument here is that we can decompose our region into a bunch of squares of small sizes and take a limit. It should be clear that for a non unit square S' , $\text{area}(T(S')) = |\det(A)|\text{area}(S')$ since we would shrink the unit square by a given scalar r on for each vector.

Suppose we just scale the unit square by r

$$T(e_1) = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad T(e_2) = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Then by what we proved before,

$$\text{area}(T(S)) = |\det(A)| = r^2 \text{ which is the area of a square with sidelengths } r$$

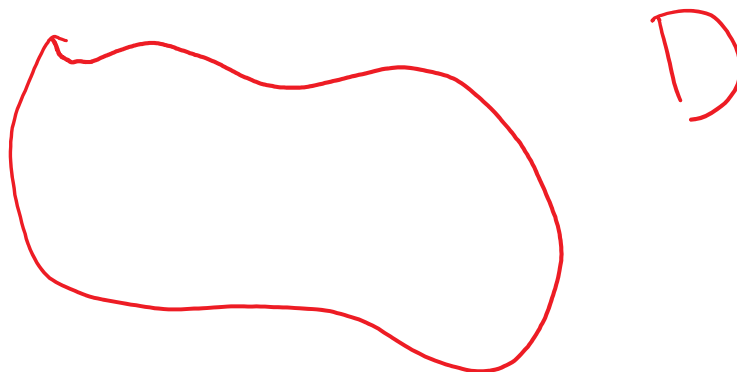
So for another transformation T on a square of side length r , S' , we could think of an equivalent transformation on the unit square where we shrink it first.

$$\hat{T} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$$

So for whatever this T looks like we do know that:

$$\text{area}(T(S')) = \text{area}(\hat{T}(S)) = \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\right) = \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \det\left(\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\right) = r^2 \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)$$

Armed with this we can use squares to approximate the area of a region D and take a limit.



Section 5.3 Applications of the Determinant

Let D be a region of finite volume in R^n , and suppose that $T: R^n \rightarrow R^n$ with $T(\vec{x}) = A(\vec{x})$. If $T(D)$ denotes the image of D under T , then

$$\text{volume}(T(D)) = |\det(A)| \cdot \text{volume}(D)$$

The arguments to prove this are similar to what we did in R^2 we can start with the unit square in R^n and then argue about different sizes of cubes and finally play a limit game again.

However, there is some visual intuition to this that is hard to draw or explain, so if there is time in lecture I would like us to watch the following video together, if not I very much suggest you watch it at home.

https://www.youtube.com/watch?v=Ip3X9LOh2dk&list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab&index=6

I would also recommend that you watch the whole linear algebra series associated with it!