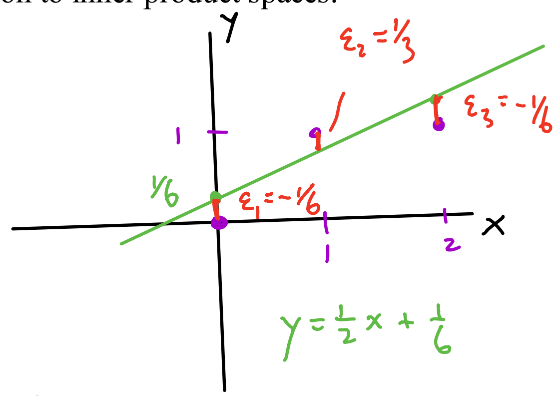
Math 2270-002 Week 14 notes
We will not necessarily finish the material from a given day's notes on that day. We may also add or subtract some material as the week progresses, but these notes represent an in-depth outline of what we plan to cover. These notes cover material in 6.6-6.8, 7.1-7.2

Mon Nov 26
- 6.6-6.7 Power laws and least squares for log-log data; introduction to inner product spaces.



Announcements: today
- power law example for HW
- inner product spaces

Warm-up Exercise: Find the linear regression line for the three points $\{(0,0), (1,1), (2,1)\}$ i.e. the least squares solution to

$$m\begin{bmatrix}0\\1\\2\end{bmatrix} + b\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\1\\1\end{bmatrix}$$

want (m, b)

$$m\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix} + b\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = \begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}$$

Sketch.

$$\begin{bmatrix}m\\b\end{bmatrix} = \begin{bmatrix}1/2\\1/6\end{bmatrix}$$

$$\begin{bmatrix}0 & 1\\1 & 1\\2 & 1\end{bmatrix}\begin{bmatrix}m\\b\end{bmatrix} = \begin{bmatrix}0\\1\\1\end{bmatrix}$$

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

$A^TA\vec{x} = A^T\vec{b}$

least sqrs "sol"

$$\begin{bmatrix}0 & 1 & 2\\1 & 1 & 1\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 1\\2 & 1\end{bmatrix}\begin{bmatrix}m\\b\end{bmatrix} = \begin{bmatrix}0 & 1 & 2\\1 & 1 & 1\end{bmatrix}\begin{bmatrix}0\\1\\1\end{bmatrix}$$

$$\begin{bmatrix}5 & 3\\3 & 3\end{bmatrix}\begin{bmatrix}m\\b\end{bmatrix} = \begin{bmatrix}3\\2\end{bmatrix}$$

$m = 1/2$, $b = 1/6$

$(A^TA)^T = A^T(A^T)^T$
$= A^TA$

since $(AB)^T = B^TA^T$

Review and preview:

On Wednesday before Thanksgiving we were discussing Section 6.6, *Applications to linear models,* which discusses how *least squares solutions to linear systems*, Section 6.5, are applied to linear models in experimental sciences and statistics. Today (on Monday) we'll apply these ideas to power laws, by looking at an example related to Kepler's Laws. The Matlab scripts should be helpful for your homework problem this week about a height-weight power law for humans.

Then in the rest of Monday's class and thru Wednesday we'll discuss Section 6.7-6.8, *Inner product spaces and applications*. Inner product spaces are vector spaces which additionally pocess the algebraic equivalent of the $\mathbb{R}^n$ dot product. As a result, inner product spaces have analgous geometry related to orthogonality, projection, angles. These ideas have amazing applications, for example to Fourier series (important in applied mathematics and physics) and as a fun and recent special case, image and audio compression algorithms.

On Friday we'll begin Chapter 7, which is about "The Spectral Theorem" and its applications. The spectral theorem is the fact that symmetric matrices A (ones for which $A^T = A$) are always diagonalizable, and with real eigenvalues, and furthermore so that the $\mathbb{R}^n$ eigenbasis can always be chosen to be ortho-normal. There are many interesting and important applications of the Spectral Theorem and we should have time to discuss some of them in the remaing days of our course.

Example like HW

Astronomical example As you may know, Isaac Newton was motivated by Kepler's (observed) Laws of planetary motion to discover the notions of velocity and acceleration, i.e. differential calculus and then integral calculus, along with the inverse square law of planetary acceleration around the sun.....from which he deduced the concepts of mass and force, and that the universal inverse square law for gravitatonal attraction was the ONLY force law depending only on distance between objects, which was consistent with Kepler's observations! Kepler's three observations were that

(1) Planets orbit the sun in ellipses, with the sun at one of the ellipse foci.
(2) A planet sweeps out equal areas from the sun, in equal time intervals, independently of where it is in its orbit.
(3) The square of the period of a planetary orbit is directly proportional to the cube of the orbit's semi-major axis.

So, for *roughly circular (elliptical) orbits*, Keplers third law translates to the statement that the period P is related to the radius r, by the equation $P = b\, r^{1.5}$, for some proportionality constant b. (And $b = 1$ in earth-centric units below.) Let's see if that's consistent with the following data:

Planet	mean distance r from sun (in astronomical units where 1=dist to earth)	Orbital period t (in earth years)
Mercury	0.387	0.241
Earth	1.	1.
Jupiter	5.20	11.86
Uranus	19.18	84.0
Pluto	39.53	248.5

The Matlab script works with this data and first finds the least squares fit to the log-log data. Then it creates two figures, the first of which is the linear regression line plotted along with a scatter plot of the log-log data. The second plot is the graph of the power law for the period as a function of radius, $P = C\, r^m$ along with a scatter plot of the original radius vs. period data. want to find power m const C.

Step 1: Matlab is finding the least squares solution to the ln-ln data, $Y = m\, X + B$ (see next page):

$$\begin{array}{l} \log(.387) \rightarrow \\ \log(1) \rightarrow \\ \log(5.20) \rightarrow \\ \vdots \\ \\ \end{array} \begin{bmatrix} -.9493 & 1 \\ 0 & 1 \\ 1.6487 & 1 \\ 2.9539 & 1 \\ 3.6771 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} -1.4230 \\ 0 \\ 2.4732 \\ 4.4308 \\ 5.5154 \end{bmatrix} . \begin{array}{l} \leftarrow \log(.241) \\ \leftarrow \log 1 \\ \vdots \\ \\ \\ \end{array}$$

$$A\, \hat{\underline{x}} = \hat{\underline{b}}$$

$$A^T A\, \underline{x} = A^T\, \underline{b}$$

$$\hat{\underline{x}} = \left(A^T A\right)^{-1} A^T\, \underline{b} \approx \begin{bmatrix} 1.5 \\ 0 \end{bmatrix}.$$

$P = C r^m$

$\log P = \log C + \log r^m$

$(\log = \ln$ for us & Matlab$)$

$\log P = \log C + m \log r$

$\text{"}Y\text{"} = m X + B$

Step 1:

```
%power laws example, with Kepler's laws:
radii=[.387; 1; 5.2; 19.18; 39.53]  %units=multiples of earth radius
periods=[.241; 1; 11.86; 84; 248.5] %units=multiples of earth period
% vectors containing the (natural) "log" of the origional ones:
log_radii=log(radii)
log_periods=log(periods)
% first column of A is log of radii, second column is "ones"
A=[log_radii,ones(5,1)]
%Matlab automatically computes least squares solution, which will be [m,b]'
ans1=linsolve(A,log_periods)
power=ans1(1) %extract slope "m", which will be power law power
intercept=ans1(2) %extract intercept "b"; Power law is P=exp(b)*r^m.
```

A

5x2 double

	1	2
1	-0.9493	1
2	0	1
3	1.6487	1
4	2.9539	1
5	3.6771	1
6		

$$A \begin{bmatrix} m \\ b \end{bmatrix} = \text{log\_periods}$$

log_periods

5x1 double

	1
1	-1.4230
2	0
3	2.4732
4	4.4308
5	5.5154
6	

ans1

2x1 double

	1
1	1.4998
2	4.8689e-04
3	
4	

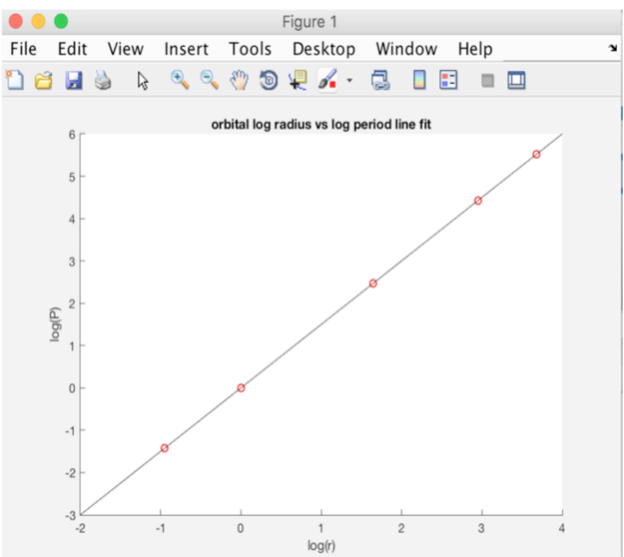
steps 2, 3: figure with linear regression line and log-log data; and figure with original data and power law graph.

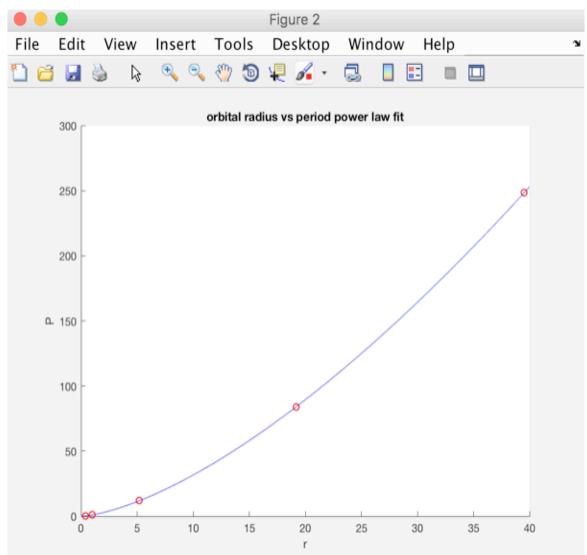
last portion of Matlab script:

```
figure(1)
log_r=linspace(-2,4,100)
scatter(log_radii,log_periods,'red')
xlabel('log(r)')
ylabel('log(P)')
title('orbital log radius vs log period line fit')
hold on
plot(log_r,ans1(2)+ans1(1)*log_r,'black')
hold off

figure(2)
r=linspace(0,40,100)
scatter(radii,periods,'red')
xlabel('r')
ylabel('P')
title('orbital radius vs period power law fit')
hold on
% the "r." applies the power to each entry of the vectors of r's.
plot(r,exp(intercept)*r.^power,'blue')
hold off
```

outputs:





Flow chart of dot product development in $\mathbb{R}^n$

$$\vec{x}\cdot\vec{y} := \sum_{i=1}^{n} x_i y_i$$

algebra

a) $\vec{x}\cdot\vec{y} = \vec{y}\cdot\vec{x}$ (symmetry)

b) $\vec{x}\cdot(\vec{y}+\vec{z}) = \vec{x}\cdot\vec{y} + \vec{x}\cdot\vec{z}$

$\vec{x}\cdot(k\vec{y}) = k\,\vec{x}\cdot\vec{y}$ (linear in each factor)

c) $\vec{x}\cdot\vec{x} \geq 0$; $\vec{x}\cdot\vec{x} = 0$ iff $\vec{x}=\vec{0}$ (positive)

An <u>inner product space</u> is a (real scalar) vector space V together with an inner product $\langle \cdot,\cdot \rangle$ which gives a real number for each pair of vectors, s.t. the following axioms hold, $\forall f,g,h \in V$, $k \in \mathbb{R}$:

a) $\langle f,g\rangle = \langle g,f\rangle$

b) $\langle f, g+h\rangle = \langle f,g\rangle + \langle f,h\rangle$

$\langle f, kg\rangle = k\langle f,g\rangle$

c) $\langle f,f\rangle \geq 0$. $\langle f,f\rangle = 0$ iff $f=0$

From these algebra axioms the entire concept chart on the left also holds, for finite dimensional subspaces W.

from algebra...

$\|\vec{x}\|^2 = \vec{x}\cdot\vec{x}$

magnitude (norm)

$\vec{x}\cdot\vec{z} = 0$

orthogonal

$\text{proj}_L \vec{x} = (\vec{x}\cdot\vec{u})\vec{u} = \dfrac{\vec{x}\cdot\vec{v}}{\|\vec{v}\|^2}\vec{v}$

$L = \text{span}\{\vec{v}\}$

$\vec{x}\cdot\vec{y} = \|\vec{x}\|\|\vec{y}\|\cos\theta$ (angles)

distance from $\vec{x}$ to $\vec{y}$ $= \|\vec{y}-\vec{x}\|$

$\|\vec{x}+\vec{z}\|^2 = \|\vec{x}\|^2 + \|\vec{z}\|^2$

Pythagorean Theorem

orthonormal basis for $W \subset \mathbb{R}^n$, coords for such basis

Gram-Schmidt algorithm

$|\vec{x}\cdot\vec{y}| \leq \|\vec{x}\|\|\vec{y}\|$ Cauchy Schwarz inequality

$\|\vec{x}+\vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$

triangle inequality (for estimates)

(we didn't really use this in this class)

projection onto W

$\mathbb{R}^n = W \oplus W^\perp$

i.e. $\vec{x} = \underset{\in W}{\text{proj}_W \vec{x}} + \underset{\in W^\perp}{\vec{z}}$, uniquely

nearest point to $\vec{b} \in \mathbb{R}^n$ in W is $\text{proj}_W \vec{b}$

Least squares solutions to $A\vec{x} = \vec{b}$

$A\hat{x} = \text{proj}_W \vec{b}$

data fitting

linear regression

polynomial fits

Examples of function space inner products:

$$V = \{ f : [a, b] \rightarrow \mathbb{R} \; s.t. \; f \text{ is continuous} \} := C([a, b]).$$

$$\langle f, g \rangle := \int_a^b f(t)\, g(t)\, dt \qquad \text{(or some fixed positive multiple of this integral).}$$

Exercise 1) Check the algebra requirements a), b), c) for an inner product.

a) $\langle f, g \rangle \overset{?}{=} \langle g, f \rangle$ $\qquad \int_a^b f(t) g(t)\, dt = \int_a^b g(t) f(t)\, dt$

b) $\langle f, g+h \rangle \overset{?\checkmark}{=} \langle f, g \rangle + \langle f, h \rangle$ $\qquad \int_a^b f(t)\left(g(t)+h(t)\right) dt \overset{\checkmark}{=} \int_a^b f(t) g(t)\, dt + \int_a^b f(t) h(t)$

$\langle f, cg \rangle \overset{\checkmark}{=} c \langle f, g \rangle$ $\quad$ c const. $\qquad$: study $\langle f, cg \rangle = \int_a^b f(t)\, c\, g(t)\, dt$

c) $\langle f, f \rangle \geq 0$ $\qquad \langle f, f \rangle = \int_a^b f(t)^2\, dt \geq 0$

$= 0$ only if $f \equiv 0$. $\qquad = 0$ iff $f \equiv 0$ (f is cont).

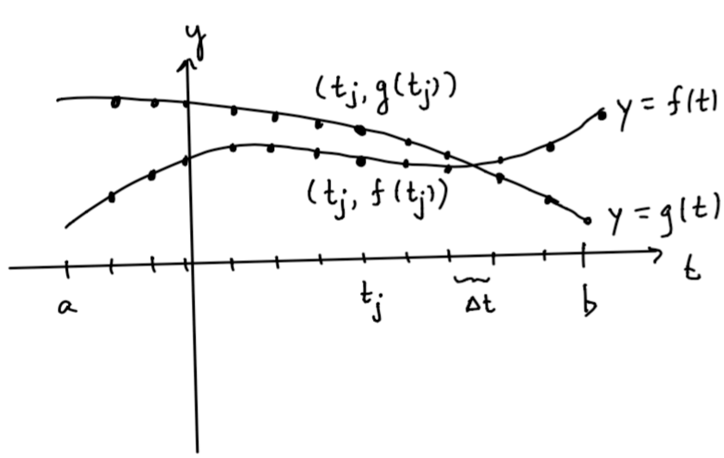
This inner product $\langle f, g \rangle$ is not so different from the $\mathbb{R}^n$ dot product if you think of Riemann sums: Let

$$\Delta t = \frac{b-a}{n}; \qquad t_j = a + j\Delta t, \, j = 1, 2, .. n.$$

Then

$$\langle f, g \rangle = \int_a^b f(t) g(t)\, dt = \lim_{n \to \infty} \sum_{j=1}^{n} f(t_j)\, g(t_j) \Delta t$$

$$= \lim_{n \to \infty} \left(\begin{bmatrix} f(t_1) \\ f(t_2) \\ : \\ f(t_n) \end{bmatrix} \cdot \begin{bmatrix} g(t_1) \\ g(t_2) \\ : \\ g(t_n) \end{bmatrix} \Delta t \right).$$



Example For the inner product on $C[-1, 1]$ given by

$$\langle f, g\rangle := \int_{-1}^{1} f(t)g(t)\, dt$$

If one applies Gram-Schmidt to the the set $\{1, t, t^2, t^3, \ldots.\}$ one creates the (normalized) *Legendre polynomials* which have an interesting entry at Wikipedia. Projecting a continuous function f onto

$$W_n = span\{1, t, t^2, \ldots t^n\}$$

will create polynomi~~c~~al approximations, that improve in the sense that

$$\lim_{n\to\infty} \left\| f - proj_{W_n} f \right\|^2 = 0.$$

Exercise 2 Find the first three Legendre polynomials by using Gram-Schmidt on the functions

$$\{f_1(t) = 1, f_2(t) = t, f_3(t) = t^2\}.$$

(In your homework you will carry this one step further.)

$$\|f_1\|^2 = \langle f_1, f_1\rangle = \int_{-1}^{1} 1\cdot 1\, dt = 2 \qquad \|f_1\| = \sqrt{2}.$$

$$u_1 = \frac{f_1}{\|f_1\|} = \frac{1}{\sqrt{2}}$$

$$u_1(t) = \frac{1}{\sqrt{2}}.$$

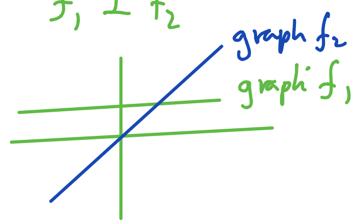
$$\vec{z}_2 = \vec{v}_2 - (\vec{v}_2 \cdot \vec{u}_1)\vec{u}_1$$

$$= f_2 - \frac{\langle f_2, f_1\rangle}{\|f_1\|^2} f_1$$

$$\langle f_2, f_1\rangle = \int_{-1}^{1} 1\cdot t\, dt = 0$$

$$f_1 \perp f_2$$

$$z_2 = f_2$$



$$\{\vec{u}_1, \vec{u}_2\} = \left\{\frac{1}{\sqrt{2}}, \frac{f_2}{\|f_2\|}\right\}$$

to be continued.