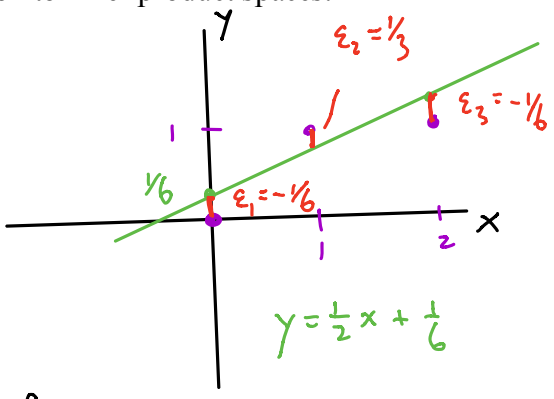
Math 2270-002 Week 14 notes

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"

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

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

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

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

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):

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

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

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

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

$P = C r^m$

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

(log = ln for us & Matlab)

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

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

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	

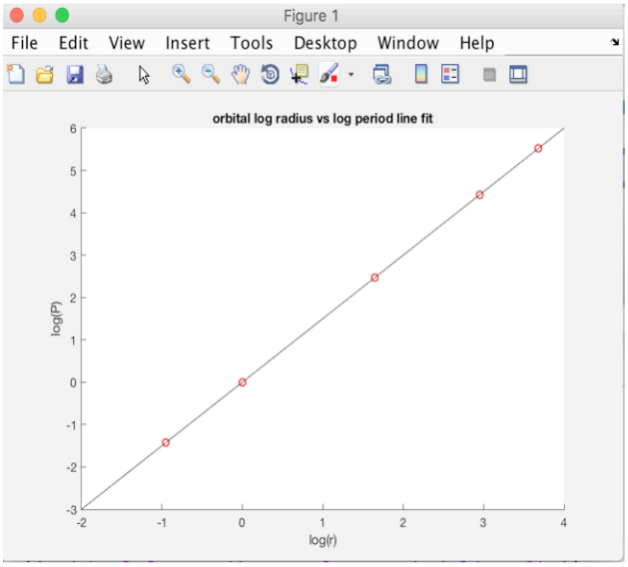
steps 2, 3: figure with linear regression line and log-log data; and figure with original data and power law graph.

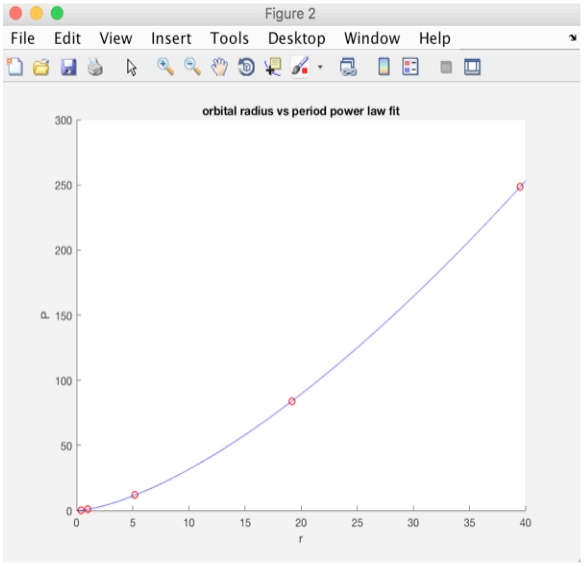
last portion of Matlab script:

```
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 inner product space 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, $\forall f, g, h \in V$, $k \in \mathbb{R}$:

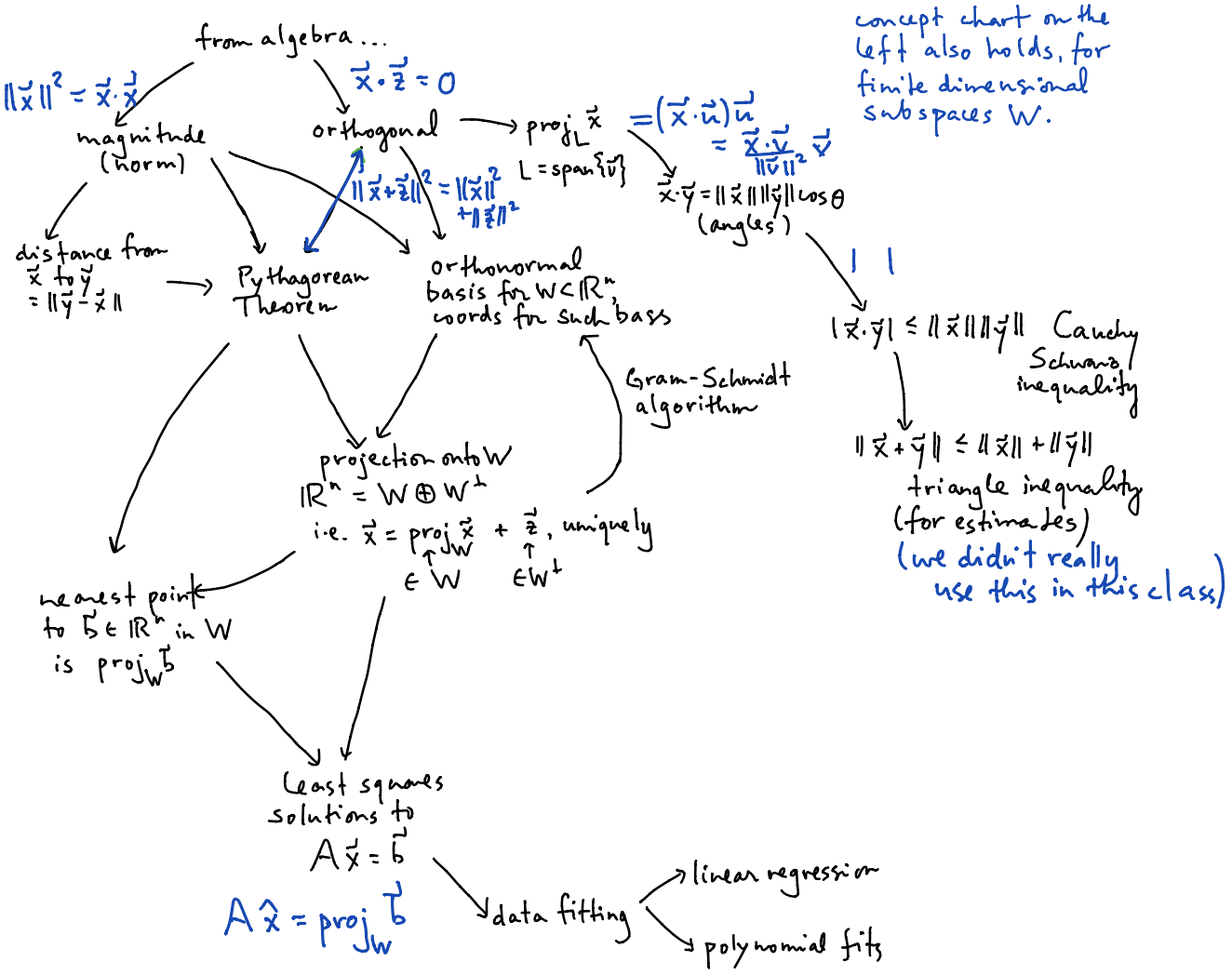
a) $\langle f, g \rangle = \langle g, f \rangle$

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.



Examples of function space inner products:

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

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

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

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

b) $\langle f, g+h \rangle \stackrel{?}{=} \langle f, g \rangle + \langle f, h \rangle$ ✓ $\quad \int_a^b f(t)\,(g(t)+h(t))\,dt \stackrel{✓}{=} \int_a^b f(t)g(t)\,dt + \int_a^b f(t)h(t)$

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

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

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

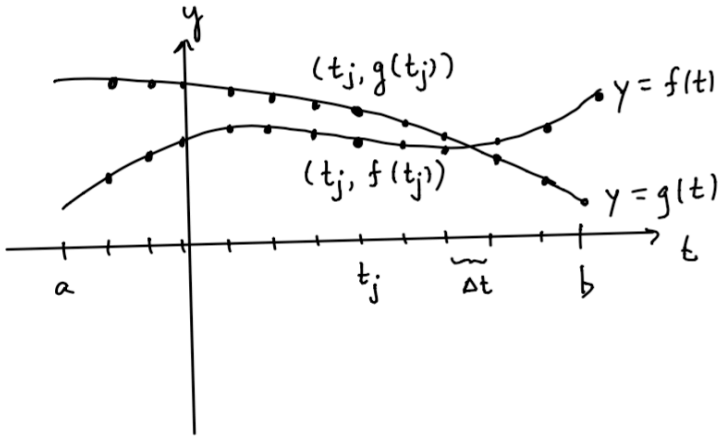
This inner product $\langle f, g \rangle$ is not so different from the $\mathbb{R}^n$ dot product if you think of Riemann sums: Let

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

Then

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

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



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 polynomial 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.)

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

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

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

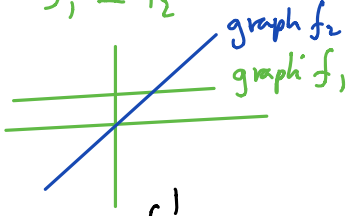
$\vec{z}_2 = \vec{v}_2 - (\vec{v}_2\cdot\vec{u}_1)\vec{u}_1$

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

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

$z_2 = f_2$

$f_1 \perp f_2$



$\{f_1, f_2\}$ orthogonal & $\{\vec{u}_1, \vec{u}_2\} = \left\{\frac{1}{\sqrt{2}}, \frac{f_2}{\|f_2\|}\right\}$ orthonormal

to be continued.

$z_3 = f_3 - \frac{\langle f_3, f_1\rangle}{\|f_1\|^2} f_1 - \frac{\langle f_3, f_2\rangle}{\|f_2\|^2} f_2$

$\|f_2\|^2 = \langle f_2, f_2\rangle = \int_{-1}^{1} t^2\,dt = \frac{t^3}{3}\Big]_{-1}^{1} = \frac{2}{3}$

$\langle f_3, f_1\rangle = \int_{-1}^{1} t^2\cdot 1\,dt = 2/3$

$\|f_2\| = \sqrt{\frac{2}{3}}$

$\langle f_3, f_2\rangle = \int_{-1}^{1} \underbrace{t^2\cdot t}_{t^3}\,dt = 0$ because t^3 is odd.

$\{u_1, u_2\} = \left\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}\,t\right\}$.

$z_3 = f_3 - \frac{2}{3}\frac{1}{2} f_1 - 0$

$z_3(t) = t^2 - \frac{1}{3}\cdot 1$

Orthogonal set $\left\{1, t, t^2 - \frac{1}{3}\right\}$

Tues Nov 27

- 6.7-6.8 Inner product spaces and introduction to Fourier series.

Announcements: I'll make a CANVAS "assignment" where you can upload the matlab script for this HW

- finish Monday example
- start Fourier series

Warm-up Exercise: Show (a) $\frac{1}{\pi}\int_{-\pi}^{\pi}(\sin 2x)(\sin 5x)\,dx = 0$ $\quad \langle \sin 2x, \sin 5x\rangle = 0$

(b) $\frac{1}{\pi}\int_{-\pi}^{\pi}\sin^2 2x\,dx = 1$ $\quad \|\sin 2x\|^2 = 1$

We'll use these facts today!

Hint $(\sin a)(\sin b) = \frac{1}{2}[\cos(a-b) - \cos(a+b)]$ $\quad a \neq b$

$\sin^2 a = \frac{1}{2}(1-\cos 2a)$

about "Fourier series inner product" is for 2π-periodic functions (or, functions defined on $[-\pi,\pi]$)

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

$\cos a\cos(-b) - (\sin a)\sin(-b)$

$-(\cos a\cos b - \sin a\sin b)$

$\cos(-b) = \cos b$

$\sin(-b) = -\sin b$

$= 2\sin a\sin b$ ✓

$\int_{-\pi}^{\pi}(\sin 2x)(\sin 5x)\,dx$

$= \int_{-\pi}^{\pi}\frac{1}{2}[\cos(-3x) - \cos 7x]\,dx$

$= \frac{1}{2}\left[\frac{1}{3}\sin 3x - \frac{1}{7}\sin 7x\right]_{-\pi}^{\pi} = 0 - 0$

$\sin n\pi = 0,\ n \in \mathbb{Z}.$

(b) $\frac{1}{\pi}\int_{-\pi}^{\pi}\sin^2 2x\,dx = \frac{1}{\pi}\cdot 2\pi \cdot$ average val. of $\sin^2 2x$ $= 1.$ ($\sin^2 2x$ has average value $1/2$)

or. $\frac{1}{\pi}\int_{-\pi}^{\pi}\frac{1}{2}(1-\cos 4x)\,dx$

$= \frac{1}{\pi 2}\left[x - \frac{\sin 4x}{4}\right]_{-\pi}^{\pi} = \frac{1}{\pi 2}[2\pi] = 1.$

Fourier series

or piecewise continuous, or equivalently 2π-periodic piecewise continuous fcns on $\mathbb{R}$

Define the Fourier series inner product on $C[-\pi, \pi]$ by

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

The infinite set of functions

$$\left\{ \frac{1}{\sqrt{2}}, \cos(t), \sin(t), \cos(2t), \sin(2t), \ldots, \sin(nt), \cos(nt), \ldots \right\}$$

is already orthonormal! Thus begins the subject of *Fourier Series*. (See Wikipedia.)

To show the ortho-normality properties one applies the following trig identities, which follow from the addition angle formulas

$$\cos(mt)\cos(nt) = \frac{1}{2}[\cos((m+n)t) + \cos((m-n)t)]$$

$$\cos^2(nt) = \frac{1}{2}[\cos(2nt) + 1]$$

$$\sin(mt)\sin(nt) = \frac{1}{2}[-\cos((m+n)t) + \cos((m-n)t)]$$

$$\sin^2(nt) = \frac{1}{2}[-\cos(2nt) + 1]$$

$$\cos(mt)\sin(nt) = \frac{1}{2}[\sin((m+n)t) + \sin((-m+n)t)]$$

Exercise 1 verify how ortho-normality follows from these identities, in at least a few cases.

Let's check that $f_1(t) = \frac{1}{\sqrt{2}}$ satisfies $\|f_1\| = 1$.

$\|f_1\|^2 = \langle f_1, f_1 \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}}\, dt = \frac{1}{\pi} \frac{1}{2} [2\pi] = 1$ ✓.

the rest of the orthogonality and unit length checking is just like warm-up exercise

Let $V_n := span\left\{\frac{1}{\sqrt{2}}, \cos(t), \sin(t), \cos(2t), \sin(2t), \ldots \cos(nt), \sin(nt)\right\}$ be the $2n+1$ dimensional subspace spanned by the first $2n+1$ of these functions. A deep theorem says that if $f \in C[-\pi, \pi]$ (actually, f only needs to be piecewise continous), then

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

Because we have an orthonormal basis for V_n the projection formula is easy to write down:

$$proj_{V_n} f = \left\langle f, \frac{1}{\sqrt{2}} \right\rangle \frac{1}{\sqrt{2}} + \langle f, \cos(t)\rangle \cos(t) + \langle f, \sin(t)\rangle \sin(t) + \ldots + \langle f, \cos(nt)\rangle \cos(nt) + \langle f, \sin(nt)\rangle \sin(nt)$$

$$= \frac{1}{2} \langle f, 1 \rangle 1 + \langle f, \cos(t)\rangle \cos(t) + \langle f, \sin(t)\rangle \sin(t) + \ldots + \langle f, \cos(nt)\rangle \cos(nt) + \langle f, \sin(nt)\rangle \sin(nt)$$

It is customary to write

$$a_0 = \langle f, 1\rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)\, dt$$

$$a_k = \langle f, \cos(kt)\rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(kt)\, dt$$

$$b_k = \langle f, \sin(kt)\rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(kt)\, dt.$$

Then

$$proj_{V_n} f = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kt) + \sum_{k=1}^{n} b_k \sin(kt).$$

There is a theorem that infinite series converges to $f(t)$ pointwise at values of t where f is differentiable, and to the average of right and left hand limits at jump discontinuities, so we also often consider the infinite Fourier series

$$f(t) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kt) + \sum_{k=1}^{\infty} b_k \sin(kt).$$

$$f(t) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kt) + \sum_{k=1}^{\infty} b_k \sin(kt).$$

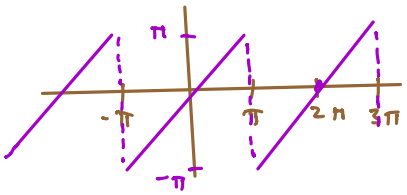
$$a_0 = \langle f, 1\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,dt \qquad a_k = \langle f, \cos(kt)\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(kt)\,dt$$

$$b_k = \langle f, \sin(kt)\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(kt)\,dt.$$

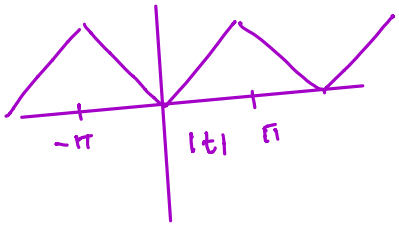
Exercise 2: Define $f(t) = t$, on the interval $-\pi < t < \pi$. Show

$$t \sim 2\sum_{k=1}^{\infty} \frac{(-1)^{n+1}}{n}\sin(nt)$$

sawtooth



in final HW assignment you will study F.S. for "tent fcn"



$f(t) = t \quad -\pi < t < \pi$

$a_0 = \langle f, 1\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} t\cdot 1\,dt = 0$

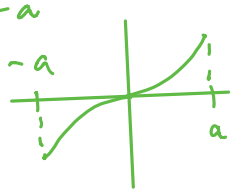
$a_n = \langle f, \cos nt\rangle$

$= \frac{1}{\pi}\int_{-\pi}^{\pi} t\cos nt\,dt = 0$

if $g(t)$ is odd $(g(-t) = -g(t))$

e.g. polys with only odd powers

then $\int_{-a}^{a} g(t)\,dt = 0$



$g(t) = t\cos nt$

$g(-t) = (-t)\underbrace{\cos(-nt)}_{\cos nt}$

$= -g(t)$

g is odd.

$b_n = \langle f, \sin nt\rangle$

$= \frac{1}{\pi}\int_{-\pi}^{\pi} \underbrace{t}_{\text{odd}}\underbrace{\sin nt}_{\text{odd}}\,dt$

$= \frac{1}{\pi}\,2\int_{0}^{\pi} t\sin nt$

even

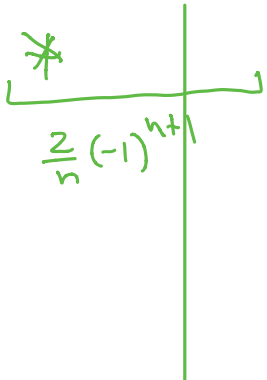
$g(t)$ is even, $h(t)$ is odd

then $g(t)h(t)$ is odd.

$g(-t)h(-t) = g(t)[-h(t)]$

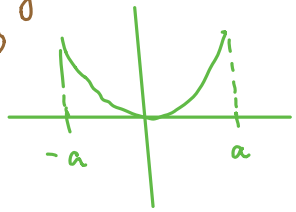
$= -g(t)h(t)$

odd · odd = even



if $g(t)$ is even

$\int_{-a}^{a} g(t)\,dt = 2\int_{0}^{a} g(t)$



continue on Wed.

$$b_n = \frac{2}{\pi}\int_0^{\pi} \underbrace{t}_{u} \underbrace{\sin nt \, dt}_{dv}$$

$$\int u\,dv = uv - \int v\,du$$

$du = dt \qquad v = -\frac{1}{n}\cos nt$

$$= \frac{2}{\pi}\left[t\left(-\frac{1}{n}\cos nt\right)\Big]_0^{\pi} + \int_0^{\pi} +\frac{1}{n}\cos nt \, dt \right]$$

$$= \frac{2}{\cancel{\pi}}\left[\cancel{\pi}\left(-\frac{1}{n}\underbrace{\cos n\pi}\right) - 0 + \cancel{\frac{1}{n^2}\sin nt\Big]_0^{\pi}} \right]$$

$\to 0$

$\sin n\pi = 0$

$\sin 0 = 0$

$$= \begin{cases} 1 & n \text{ even} \\ -1 & n \text{ odd} \end{cases}$$

$$= (-1)^n$$

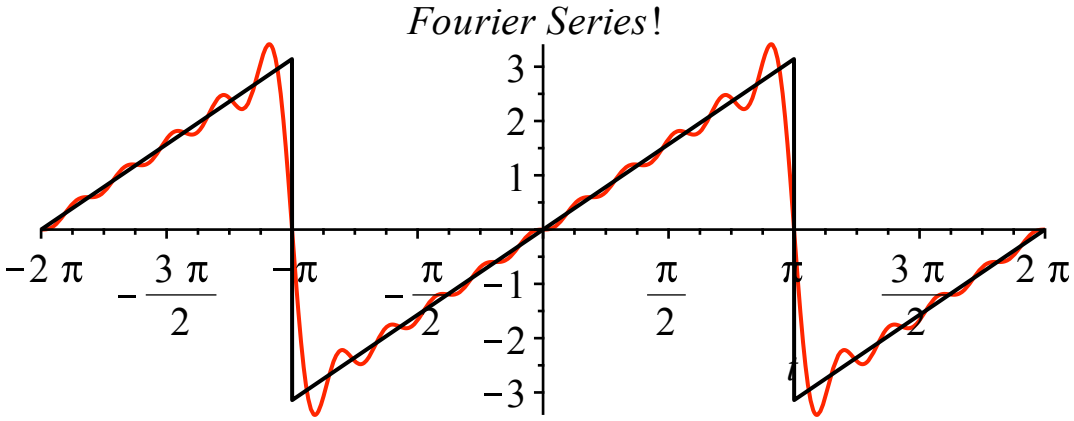
$$= \underbrace{\frac{2}{n}(-1)^{n+1}}_{b_n}$$

$$t \sim \sum_{n=1}^{\infty} \frac{2}{n}(-1)^{n+1}\sin nt$$

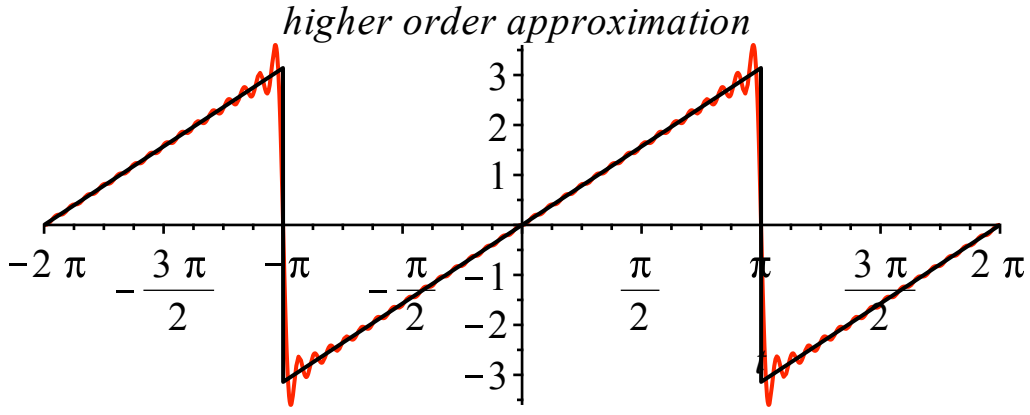
$$= 2\left(\sin t - \frac{1}{2}\sin 2t + \frac{1}{3}\sin 3t - \dots\right)$$

$$proj_{V_{10}} f(t):$$

```
> with(plots):
  plot1 := plot(t + 2·π − 2·π·Heaviside(t + π) − 2·π·Heaviside(t − π), t = −2·π..2·π, color
       = black):
  plot2 := plot(2·Σ_{n=1}^{10} (−1)^{n+1}·sin(n·t)/n, t = −2·π..2·π, color = red):
  display({plot1, plot2}, title = 'Fourier Series!');
```




```
> plot3 := plot(2·Σ_{n=1}^{30} (−1)^{n+1}·sin(n·t)/n, t = −2·π..2·π, color = red):
  display({plot1, plot3}, title = 'higher order approximation');
```



Wed Nov 28

- 6.8 Fourier series and applications

Announcements:

- finish F.S. example
- show & tell about .jpg image compression
- quiz : find least squares line fit

Warm-up Exercise: Suppose $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is an orthonormal set in an inner product space. Find a formula for

$$\| a_1\vec{u}_1 + a_2\vec{u}_2 + a_3\vec{u}_3 \|^2 = a_1^2 + a_2^2 + a_3^2$$

in terms of the weights a_1, a_2, a_3.

(Recall, $\|\vec{x}\|^2 = \langle \vec{x}, \vec{x} \rangle$)

$$\| a_1\vec{u}_1 + a_2\vec{u}_2 + a_3\vec{u}_3 \|^2 = \langle a_1\vec{u}_1 + a_2\vec{u}_2 + a_3\vec{u}_3,\ a_1\vec{u}_1 + a_2\vec{u}_2 + a_3\vec{u}_3 \rangle$$

(linearity)

$$= \langle a_1\vec{u}_1 + a_2\vec{u}_2 + a_3\vec{u}_3,\ a_1\vec{u}_1 \rangle + \langle a_1\vec{u}_1 + a_2\vec{u}_2 + a_3\vec{u}_3,\ a_2\vec{u}_2 \rangle + \langle \cdots, a_3\vec{u}_3 \rangle$$

$$= a_1^2 \langle \vec{u}_1, \vec{u}_1 \rangle + \langle a_2\vec{u}_2, a_1\vec{u}_1 \rangle + \langle a_3\vec{u}_3, a_1\vec{u}_1 \rangle + a_2^2 + a_3^2$$

($\langle \vec{u}_1, \vec{u}_1 \rangle = 1$; $\langle \vec{u}_2, \vec{u}_1 \rangle = 0$; $\langle \vec{u}_3, \vec{u}_1 \rangle = 0$)

$$= a_1^2 + a_2^2 + a_3^2$$

As part of the deep theorem about Fourier series, as long as f is piecewise continuous,

$$\left\| proj_{V_N} f - f \right\| \to 0 \quad \text{and} \left\| proj_{V_N} f \right\| \to \| f \| .$$

Recall, the norm that we get from the Fourier series inner product is

$$\| g \|^2 = \frac{1}{\pi} \int_{-\pi}^{\pi} g(t)^2 \, dt.$$

Now,

$$proj_{V_n} f = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(k t) + \sum_{k=1}^{n} b_k \sin(k t)$$

So

$$\left\| proj_{V_n} f \right\|^2 = \left\| \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(k t) + \sum_{k=1}^{n} b_k \sin(k t) \right\|^2$$

$$= \left\langle \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(k t) + \sum_{k=1}^{n} b_k \sin(k t), \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(k t) + \sum_{k=1}^{n} b_k \sin(k t) \right\rangle$$

$$= \left\langle \frac{a_0}{2}, \frac{a_0}{2} \right\rangle + \sum_{k=1}^{n} a_k^2 + \sum_{k=1}^{n} b_k^2 = \frac{a_0^2}{2} + \sum_{k=1}^{n} a_k^2 + \sum_{k=1}^{n} b_k^2,$$

because the cross terms in the expanded inner product cancel out - since the basis vectors we've chosen for V_n are orthonormal:

$$V_n := span\left\{ \frac{1}{\sqrt{2}}, \cos(t), \sin(t), \cos(2 t), \sin(2 t), ... \cos(n t), \sin(n t) \right\}$$

As an application, for our function $f(t) = t$,

$$\| f \|^2 = \frac{1}{\pi}\int_{-\pi}^{\pi} t^2 \, dt = \frac{2}{\pi}\int_{0}^{\pi} t^2 \, dt = \frac{2}{3}\pi^2.$$

Since

$$t \sim t \sim 2\sum_{k=1}^{\infty} \frac{(-1)^{n+1}}{n}\sin(n\,t)$$

It must be that

$$\frac{2}{3}\pi^2 = 4\sum_{k=1}^{\infty} \frac{1}{n^2}$$

$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{25} + \cdots$ $\qquad \sum_{k=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$

This magic formula is true (which is sort of amazing), although you may not have seen it before:

```
> sum(1/k^2, k = 1 .. infinity);
```

$$\frac{1}{6}\pi^2 \qquad \textbf{(1)}$$

Image compression show and tell... I'll demo the matlab code

Math 3150 2D Fourier Series Field Trip Project

Due date: Monday, Nov. 28 after Thanksgiving holiday.

Two-dimensional Fourier series can be used to perform image processing and data compression, and is a salient example of how an set of orthogonal basis functions can be used to approximate functions. Suppose $f(x, y)$ is defined on the region $(x, y) \in [0, L] \times [0, H]$ and represents a grey scale image. For each point (x, y) the greyscale value ranges from zero (black) to unity (white) $f \in [0, 1]$. The orthogonal basis set we use is a 2D sine series:

$$\phi_{n,m}(x, y) = \sin\left(\frac{n\pi}{L}x\right) \sin\left(\frac{m\pi}{H}y\right),$$

.jpg uses cosine series instead.

where n and m both range from $1, 2, 3, \ldots$. The values $\frac{n\pi}{L}$ and $\frac{m\pi}{H}$ represent the horizontal and vertical spatial frequencies. The approximate image is the double sum orthogonal projection:

$$\hat{f}_{N,M}(x, y) = \sum_{n=1}^{N} \sum_{m=1}^{M} B_{n,m}\phi_{n,m}(x, y),$$

where N and M represent the sum truncation and the values $\frac{N\pi}{L}$ and $\frac{M\pi}{H}$ represent the maximum horizontal and vertical spatial frequencies available to represent the image—any image feature that has higher spatial frequency, such as sharp areas of contrast, fine texture, etc, cannot be represented. The Fourier coefficients are

$$B_{n,m} = \frac{\langle f, \phi_{n,m}\rangle}{\langle \phi_{n,m}, \phi_{n,m}\rangle}, \quad \langle g, h\rangle = \int_0^L \int_0^H g(x, y)h(x, y)dydx.$$

The goal of the project is to assess the qualitative nature of Fourier image processing in two experiments.

Experiment 1:

The first experiment you will show a subject (a friend that has not seen the full image) a Fourier-decomposed image of a car/truck that you take with your cell phone camera. I advise to use a "square" Instagram-ready image, and compress it to 256X256 pixels, which is most easily accomplished by re-sending the picture to yourself by email and compressing it to the "small" size for sending. The automobile image should be centered in the fame, and fill approximately 1/2-3/4 the width of the frame. It should be a random car parked on a street, or something, that's from a common brand and model that's recognizable to most people, or at least your friend. Your friend should not know anything about the picture at all and don't tell them anything. Show your friend successively higher truncated Fourier compressed images of the car until your friend correctly guesses (1) that its a automobile of some type, and then (2) guesses the make and/or model type. Start by showing your friend the Fourier compressed image at $N = M = 10$, then ascend $N = 20, 30, 40, 50, \ldots$ until he/she gets both (1) and then (2). At each stage, record you friend's response and report your results, including the images and the compressed images.

Experiment 2:

Take two pictures, both 256X256 as described above. One of the images should be a "natural scene," which should be interpreted broadly as naturescapes, or varied urban cityscapes—the point is that it should contain a mix lots of things in the image, both foreground and background, objects with lots of different sizes in the frame—and the other should be a somewhat boring picture of a single human-made material—e.g., a wall of bricks, patterned fabric, things of a regular or repeated nature to it; be sure that you fit several repetitions of the pattern in your picture.. Get inventive with what you choose. We will compare the two pictures' Fourier coefficients $B_{n,m}$. Natural images have been commonly reported to have squared Fourier coefficients that decay with a power law:

$$B_{n,m}^2 \sim \frac{1}{n^\gamma} \quad \text{or} \quad \sim \frac{1}{m^\gamma},$$

where γ is typically in the range between 1.7 and 2.3. That is, γ is usually around 2. Why do we examine squared Fourier coefficients? Its because squared values are associated with the energy in the image through Parseval's identity:

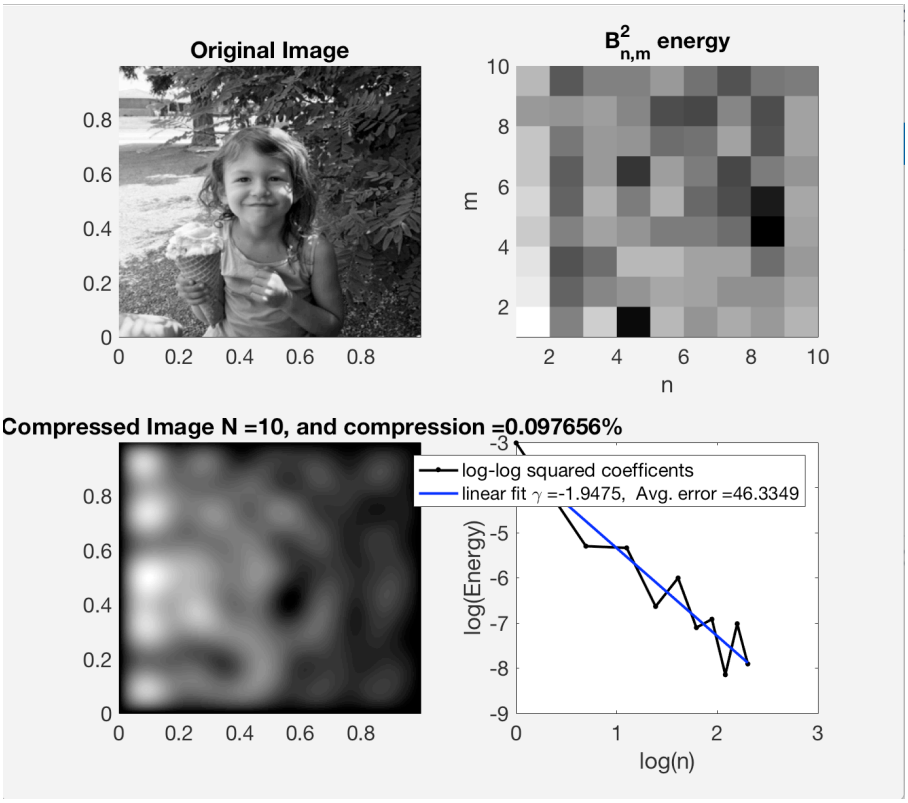
$$Energy = \int_0^L \int_0^H |f(x,y)|^2 dydx = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} B_{n,m}^2.$$

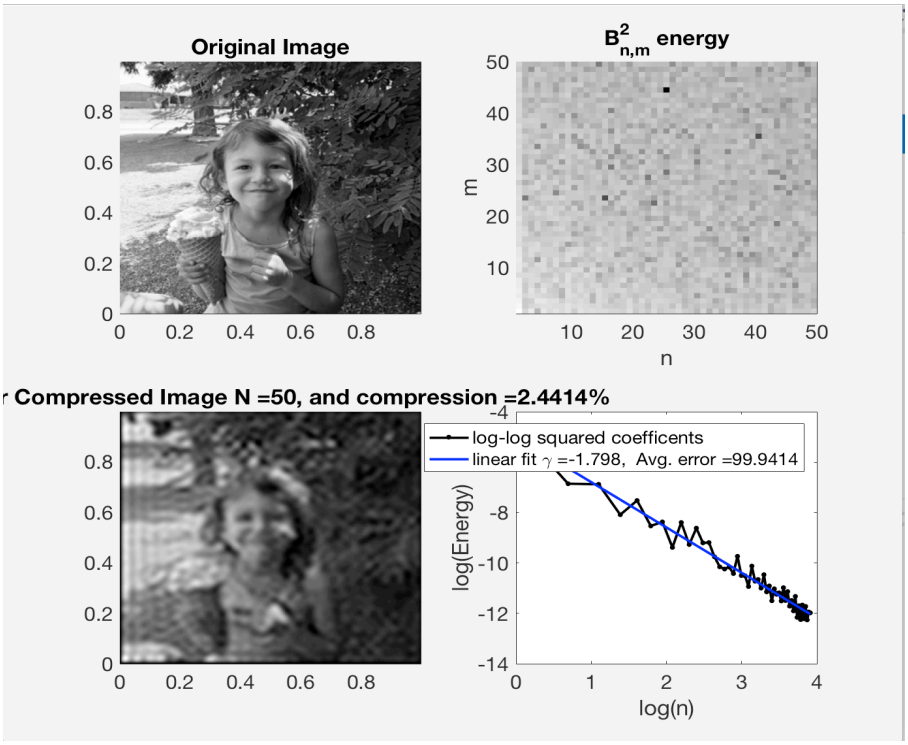
In the accompanying code, the coefficients $B_{n,m}$ are computed and represented as a matrix, and rendered of the squared values of the $B_{n,m}$. The energy spectrum is a log-log plot of the average of vertical and horizontal average energies: $\frac{1}{2} avg_m(B_{n,m}^2) + \frac{1}{2} avg_m(B_{m,n}^2) = b_n^2$—this gives an estimate of the spectral energy at each spatial frequency $n\pi$ per unit image length L. If $b_n^2 \sim \frac{1}{n^\gamma}$, then taking the log of both sides we get:

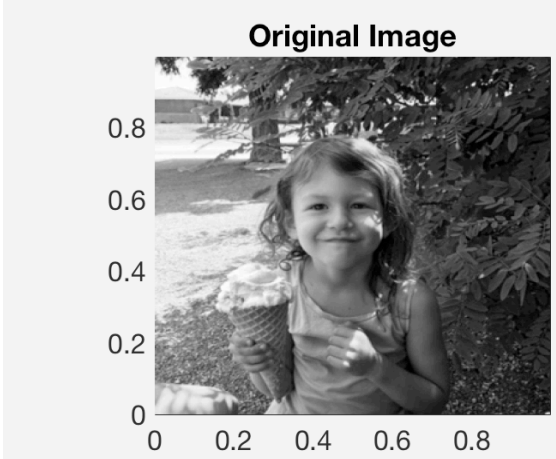
$$\ln(b_n^2) \sim \ln\left(\frac{1}{n^\gamma}\right) = -\gamma \ln(n).$$

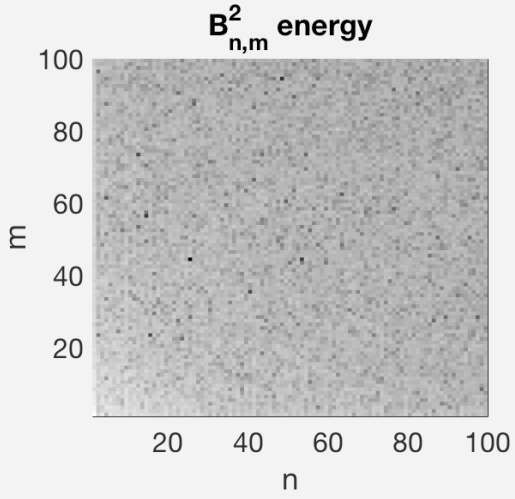
That is, the log squared coefficient averages will be linearly related to the log of n with slope $-\gamma$. The code performs a linear curve fitting on the log-log data and finds the best-fit γ-value as our estimate. For the two images you choose, record the γ-estimates and report them in your results along with your images. Use a large truncation $N = 100$ value—it may take a while. Report the gamma-values you find, and the standard deviations of the linear fit.

How to use the code: Put the images you want to analyze in a file folder with the .m code given with this experiment. Type in the code the file name of the image you want to examine and edit the code to set an N value for your truncation. There is a variable called "experiment", which you set to 1 or 2, respectively. The code will output figures. Figure 1 will give you the Nth Fourier truncated image in greyscale. Figure 2 will output the results for experiment 2.

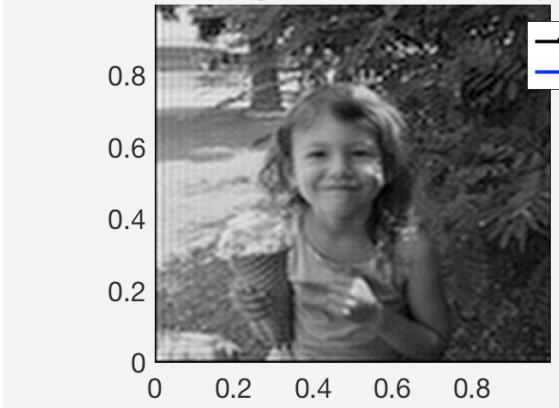

Original Image
$B^2_{n,m}$ energy
m
n
Compressed Image N =10, and compression =0.097656%
log-log squared coefficents
linear fit γ =-1.9475, Avg. error =46.3349
log(Energy)
log(n)

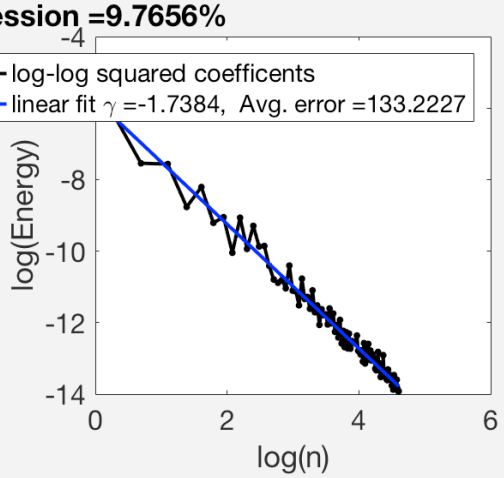

Original Image
$B^2_{n,m}$ energy
m
n
Compressed Image N =50, and compression =2.4414%
log-log squared coefficents
linear fit γ =-1.798, Avg. error =99.9414
log(Energy)
log(n)





Compressed Image N =100, and compression =9.7656%





Fri Nov 30

- Symmetric matrices and the spectral theorem, 7.1-7.2

Announcements:

- PCA ← ap. (principal component analysis)
  Prof Tom Alberts guest lecture on Tues., will (at least by end of lecture) indicate the ideas behind genetic geneology)

Warm-up Exercise: Find eigenvalues and eigenvectors for the matrix

$$A = \begin{bmatrix} 2 & 5/2 \\ 5/2 & 2 \end{bmatrix}$$

Note that the eigenspaces are orthogonal (we need this eigendata in today's class)

$$|A-\lambda I| = \begin{vmatrix} 2-\lambda & 5/2 \\ 5/2 & 2-\lambda \end{vmatrix} = (\lambda-2)^2 - \left(\tfrac{5}{2}\right)^2 = \left(\lambda-2-\tfrac{5}{2}\right)\left(\lambda-2+\tfrac{5}{2}\right)$$
$$= \left(\lambda-\tfrac{9}{2}\right)\left(\lambda+\tfrac{1}{2}\right).$$
$$= 0 \text{ for } \lambda = \tfrac{9}{2}, -\tfrac{1}{2}.$$

$$E_{\lambda=\frac{9}{2}} = \text{span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$$

$$\begin{bmatrix} 2 & 5/2 \\ 5/2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 9/2 \\ 9/2 \end{bmatrix} = \tfrac{9}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \checkmark$$

$$E_{\lambda=-\frac{1}{2}} = \text{span}\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$$

Recall that the transpose operation swaps rows with columns, and vise verse. These properties arose from the actual definition for A^T, which was

$$entry_{i\,j}\, A^T = entry_{j\,i}\, A.$$

The $i\,j$ and $j\,i$ locations on a matrix are reflections across the diagonal of each other. (This is the matrix version of the $\mathbb{R}^2$ reflection across the line $x_2 = x_1$ that we've encountered several times in this course.) See how this plays out for the matrix A below, by finding the transpose three ways: Turning rows into columns; turning columns into rows; reflecting across the diagonal.

$$A = \begin{bmatrix} 1 & 2 & 7 \\ -1 & 3 & 2 \\ 9 & 4 & 2 \end{bmatrix}$$

$$A^T = \begin{bmatrix} 1 & -1 & 9 \\ 2 & 3 & 4 \\ 7 & 2 & 2 \end{bmatrix}$$

$$A^T = \begin{bmatrix} 1 & -1 & 9 \\ 2 & 3 & 4 \\ 7 & 2 & 2 \end{bmatrix}$$

reflect across diag

$$A^T = \begin{bmatrix} 1 & -1 & 9 \\ 2 & 3 & 4 \\ 7 & 2 & 2 \end{bmatrix}$$

Def A square matrix is *symmetric* if and only if $A^T = A$.

Exercise 1 Which of the following matrices is symmetric, and which is not?

1a)

$$B := \begin{bmatrix} 4 & 2 & 1 \\ 2 & 0 & -2 \\ 1 & -2 & 7 \end{bmatrix}$$

yes

1b)

$$C := \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & -2 \\ 2 & -2 & 3 \end{bmatrix}$$

no

The *Spectral Theorem* asserts that all $n \times n$ symmetric matrices A (with real number entries) are diagonalizable, with n linearly independent real eigenvectors and associated eigenvalues. Furthermore, eigenvectors with different eigenvalues are *automatically* orthogonal. (For eigenspaces with dimension greater than one, one can use Gram Schmidt to create orthonormal bases). Thus, the eigenvector basis of $\mathbb{R}^n$ can be chosen to be orthonormal. In otherwords, we may express

$$A\,P = P\,\mathrm{D}$$

as

$A = P^{-1} D P = P^T D P$! $D = P^{-1}AP = P^TAP$

where P is an orthogonal matrix which can also be interpreted as a change of basis matrix. Let's see how this plays out in an example. This will forshadow sections 7.1-7.2. You'll notice that we're using major concepts and ideas from throughout the course, which is not a bad way to be reviewing course material at this point of the semester.

Example

1 Consider the curve in $\mathbb{R}^2$ defined implicitly as the solution set to the equation

$$2\,x_1^{\,2} + 2\,x_2^2 + 5\,x_1\,x_2 = 1.$$

Can you identify the curve as a conic section? Can you graph it? Note the $x_1\,x_2$ cross-term!

2 Does the function $f(x_1, x_2) = 2\,x_1^2 + 2\,x_2^2 + 5\,x_1\,x_2$ have a local maximum or local minimum at $(x_1, x_2) = (0, 0)$? Note, the gradient

$$\nabla f = \left[f_{x_1}, f_{x_2}\right] = \left[4\,x_1 + 5\,x_2,\, 4\,x_2 + 5\,x_1\right] = [0, 0] \text{ at the point } (0, 0),$$

so the origin is at least a candidate for a local max or min.

Exercise 1a. Check that we can rewrite the quadratic expression as

$$2\,x_1^2 + 2\,x_2^2 + 5\,x_1\,x_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2x_1 + \frac{5}{2}x_2 \\ \frac{5}{2}x_1 + 2x_2 \end{bmatrix}$$

$$x_1\left(2x_1 + \tfrac{5}{2}x_2\right) + x_2\left(\tfrac{5}{2}x_1 + 2x_2\right)$$

$$= 2x_1^2 + 5x_1x_2 + 2x_2^2 \quad \checkmark$$

Note, in general, if $\underline{\boldsymbol{v}}, \underline{\boldsymbol{w}} \in \mathbb{R}^n$ and if A is an $n \times n$ matrix then

$$\underline{\boldsymbol{v}}^T_{1 \times n} A_{n \times n} \underline{\boldsymbol{w}}_{n \times 1}$$

is a 1×1 matrix, i.e. a scalar. And its value is

$$\underline{\boldsymbol{v}}^T A \underline{\boldsymbol{w}} = \sum_{i=1}^{n} v_i \left(entry_i \left(A \underline{\boldsymbol{w}} \right) \right) = \sum_{i=1}^{n} v_i \left(\sum_{j=1}^{n} a_{ij} w_j \right) = \sum_{i,j=1}^{n} a_{ij} v_i w_j \, .$$

So given a quadratic expression ("quadratic form") in any number of variables $\left(x_1, x_2, \ x_n\right)$ one can rewrite the quadratic form as

$$\underline{\boldsymbol{x}}^T A \underline{\boldsymbol{x}} = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

and one can choose to make A a symmetric matrix, as we did in our specific example. by splitting cross terms symmetrically in the off-diagonal entries of A.

e.g. $5x_1^2 - 2x_2^2 + x_3^2 + 4x_1x_2 - x_1x_3 + 7x_2x_3$ "quadratic form" in 3 variables

$$= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 5 & 2 & -\frac{1}{2} \\ 2 & -2 & \frac{7}{2} \\ -\frac{1}{2} & \frac{7}{2} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Exercise 1b Find the eigenvalues and eigenvectors for the matrix we're using to express our quadratic expression.

$$2x_1^2 + 2x_2^2 + 5x_1x_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

warm-up!

Solution: $E_{\lambda=-\frac{1}{2}} = span\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$ ✓

$$E_{\lambda=\frac{9}{2}} = span\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}.$$

$\begin{bmatrix} 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0$

$E_{\lambda=-1/2} \perp E_{\lambda=9/2}$

Was it an accident that the two eigenvectors were orthogonal? No. Here's why that will always be true as long as the eigenvalues are different, for any symmetric matrix of arbitrary size:

Theorem Let A be symmetric, and let

$$A\,\underline{v} = \lambda_1 \underline{v} \qquad A\,\underline{w} = \lambda_2\,\underline{w}$$

with $\lambda_1 \neq \lambda_2$. then $\underline{v} \perp \underline{w}$.

Proof: Consider the scalar

$$\underline{v}^T A\,\underline{w}.$$

On one hand

$$\underline{v}^T (A\,\underline{w}) = \underline{v}^T \left(\lambda_2\,\underline{w}\right) = \lambda_2\,\underline{v}\cdot\underline{w}\,.$$

On the other hand

$$\left(\underline{v}^T A\right)\underline{w} = \left(\underline{v}^T A^T\right)\underline{w} = (A\,\underline{v})^T\,\underline{w} = \lambda_1\,\underline{v}^T\,\underline{w} = \lambda_1\,\underline{v}\cdot\underline{w}\,.$$

So

$$\lambda_2\,\underline{v}\cdot\underline{w} = \lambda_1\,\underline{v}\cdot\underline{w}.$$

So as long as $\lambda_1 \neq \lambda_2$ we deduce that the eigenvectors $\underline{v}, \underline{w}$ satsify $\underline{v}\cdot\underline{w} = 0$!

Was it an accident that the eigenvalues were real #'s?

NO: recall if A has real entries, if

$$A(\vec{u}+i\vec{v}) = (a+bi)(\vec{u}+i\vec{v})$$

we checked $\quad A(\vec{u}-i\vec{v}) = (a-bi)(\vec{u}-i\vec{v})$

$\vec{u}+i\vec{v}$ complex eigenvector $a+bi$ " eigenval.

Consider $\quad (\vec{u}-i\vec{v})^T A(\vec{u}+i\vec{v})$

$$(a+bi)(\vec{u}+i\vec{v})$$

$$(a+bi)\left[(\vec{u}-i\vec{v})\cdot(\vec{u}+i\vec{v})\right]$$

$$(a+bi)\left(\|\vec{u}\|^2+\|\vec{v}\|^2\right)$$

$$(\vec{u}-i\vec{v})^T A^T(\vec{u}+i\vec{v})$$

$$\left[A(\vec{u}-i\vec{v})\right]^T$$

$$(a-bi)(\vec{u}-i\vec{v})^T(\vec{u}+i\vec{v})$$

$$(a-bi)\left(\|\vec{u}\|^2+\|\vec{v}\|^2\right)$$

so $b=0$!

Continuing example ...

$$2x_1^2 + 2x_2^2 + 5x_1x_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and for

$$A = \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix}; \quad E_{\lambda=-\frac{1}{2}} = span\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}, \; E_{\lambda=\frac{9}{2}} = span\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}.$$

This suggests creating an orthonormal eigenbasis! And we'll order the eigenvectors so that the corresponding orthogonal matrix is a rotation and not a reflection (by making the determinant of the matrix $+1$ instead of -1).

$$B = \left\{\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}\right\} \qquad P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$AP = PD$$

$A = PDP^{-1}$

$D = P^{-1}AP = P^TAP$

Note that P is the change of coordinates matrix,

$$P = \; _{E \leftarrow B}P$$

where as always,

$$E = \left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}.$$

For $\underline{v} \in \mathbb{R}^2$ write $\underline{v} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ in standard coordinates and $[\underline{v}]_B = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. So the two coordinate systems are related by

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

$$[\underline{v}]_E = P\,[\underline{v}]_B.$$

Do algebra! For A symmetric and $P^T A P = \mathrm{D}$, the *quadratic form*

$$\underline{\boldsymbol{x}}^T A \, \underline{\boldsymbol{x}}$$

can be *diagonalized* with the change of variables $\underline{\boldsymbol{x}} = P \, \underline{\boldsymbol{y}}$:

$$\underline{\boldsymbol{x}}^T A \, \underline{\boldsymbol{x}} = \underline{\boldsymbol{y}}^T P^T A \, P \, \underline{\boldsymbol{y}} = \underline{\boldsymbol{y}}^T \mathrm{D} \, \underline{\boldsymbol{y}} = \sum_{i=1}^{n} \lambda_i \, y_i^2 .$$

In our example,

$$2\,x_1^2 + 2\,x_2^2 - 5\,x_1 x_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} \frac{9}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad (\text{because } P^T A P = \mathrm{D})$$

$$= \frac{9}{2} y_1^{\,2} - \frac{1}{2} y_2^{\,2} .$$

So the original curve with equation

$$2x_1^2 + 2x_2^2 + 5x_1x_2 = 1$$

in the standard coordinate system has equation

$$\frac{9}{2}{y_1}^2 - \frac{1}{2}{y_2}^2 = 1$$

with respect to the rotated coordinate system!

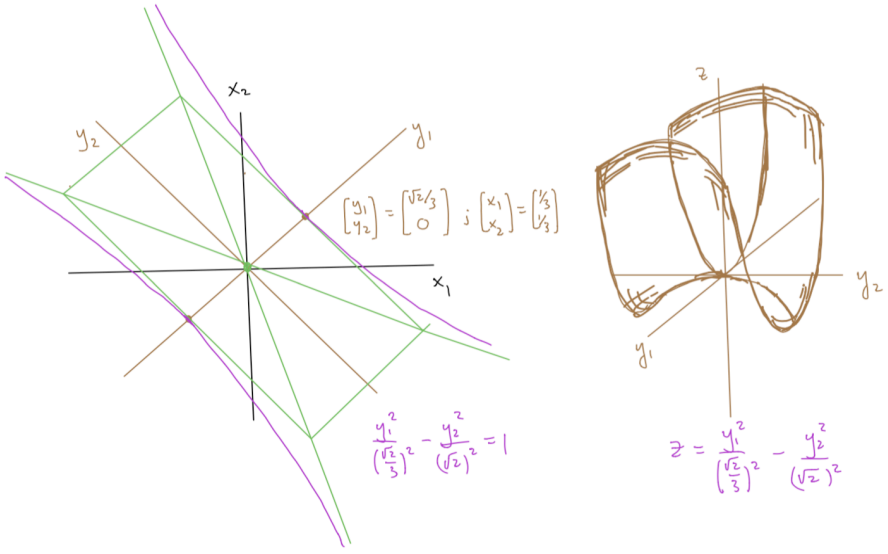
Answer to 1a) This curve is a hyperbola! In the rotated coordinate system its equation is

$$\frac{{y_1}^2}{\left(\frac{\sqrt{2}}{3}\right)^2} - \frac{{y_2}^2}{(\sqrt{2})^2} = 1.$$

Answer to 1b) No! $f(x_1, x_2) = 2x_1^2 + 2x_2^2 + 5x_1x_2$ does not have a local min or max at $(0, 0)$. The origin is a saddle point, because in the rotated coordinate system

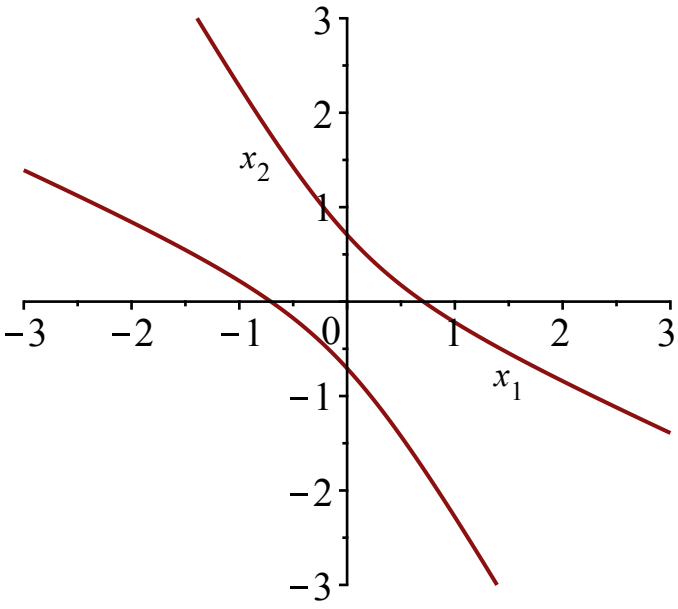
$$f(y_1, y_2) = \frac{9}{2}{y_1}^2 - \frac{1}{2}{y_2}^2.$$

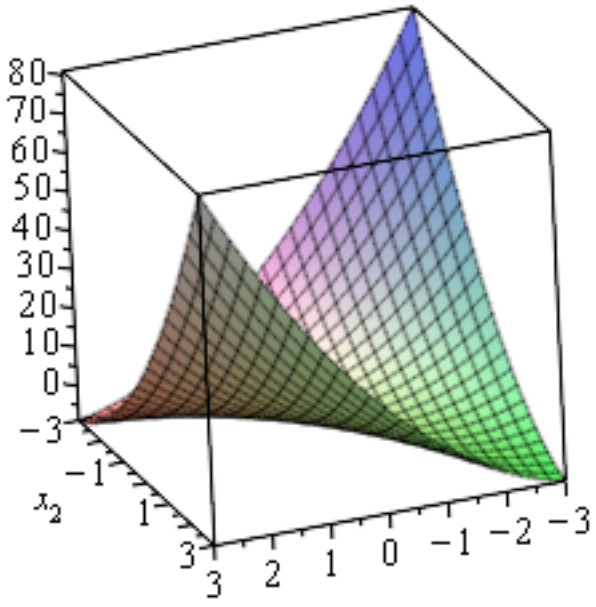
Hand sketches:



Maple verification: To be continued ....

```
> with(plots):
  implicitplot(2·x1^2 + 2·x2^2 + 5·x1·x2 = 1, x1 = -3..3, x2 = -3..3, grid = [200, 200]);
```




```
> plot3d(2·x1^2 + 2·x2^2 + 5·x1·x2, x1 = -3..3, x2 = -3..3);
```