Math 2270-004 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 6.8, 7.1-7.2, with some supplementary material. The Friday notes are not yet included.

Mon Apr 16
- 6.8 Truncated Fourier series as projection of functions via an orthonormal basis of sinusoidal functions; Fourier series in two variables and the idea behind jpg image compression, show and tell.

Announcements:
- does Wolfram alpha work with large matrices? anything else online?
- careful Fourier series
- jpeg compression.

Warm-up Exercise: well, you could review the dot product, inner product flow chart...

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) $\vec{x}\cdot(\vec{y}+\vec{z}) = \vec{x}\cdot\vec{y} + \vec{x}\cdot\vec{z}$ (linear in each factor)

$\vec{x}\cdot(k\vec{y}) = k\,\vec{x}\cdot\vec{y}$

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

from algebra...

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

magnitude (norm)

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

orthogonal

$\text{proj}_L \vec{x}$

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

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

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

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)

projection onto W

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

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

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

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

data fitting

linear regression

polynomial fits

power laws.

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.

Example for the inner product on $C[-\pi, \pi]$ given by

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

$$\left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}\,dt = \frac{1}{\pi}\frac{1}{2}\int_{-\pi}^{\pi} 1\,dt = \frac{2\pi}{2\pi} = 1$$

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

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

$$\cos^2(n\,t) = \frac{1}{2}[\cos(2\,n\,t) + 1]$$

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

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

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

Exercise verify how ortho-normality follows from these identities.

$\langle \cos mt, \cos nt\rangle = 0 \qquad m \neq n.$

$$\frac{1}{\pi}\int_{-\pi}^{\pi}\cos(mt)\cos(nt)\,dt = \frac{1}{\pi}\int_{-\pi}^{\pi}\frac{1}{2}\left(\cos(m+n)t + \cos(m-n)t\right)dt$$

$$= \frac{1}{\pi}\left[\frac{1}{2}\left(\frac{\sin(m+n)t}{m+n} + \frac{\sin(m-n)t}{m-n}\right)\right]_{-\pi}^{\pi}$$

$$= \frac{1}{\pi}(0-0) = 0$$

$$\|\cos nt\|^2 = \langle \cos nt, \cos nt\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi}\underbrace{\frac{1}{2}[\cos 2nt + 1]}_{\cos^2 nt}\,dt = \frac{1}{\pi}\frac{1}{2}\left(\frac{\sin 2nt}{2n} + t\right)\Bigg]_{-\pi}^{\pi}$$

$$= \frac{1}{\pi}\frac{1}{2}\left(0 + \pi - (0 - \pi)\right)$$

$$= \frac{2\pi}{2\pi} = 1$$

Let $V_n := span\left\{\frac{1}{\sqrt{2}}, \cos(t), \sin(t), \cos(2\,t), \sin(2\,t), ... \cos(n\,t), \sin(n\,t)\right\}$ be the $2\,n+1$ dimensional subspace spanned by the first $2\,n+1$ 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:

$proj_W \vec{x} = (\vec{x}\cdot\vec{u}_1)\vec{u}_1 + (\vec{x}\cdot\vec{u}_2)\vec{u}_2 + \dots + (\vec{x}\cdot u_N)\vec{u}_N$

$$proj_{V_n} f = \left\langle f, \frac{1}{\sqrt{2}}\right\rangle \frac{1}{\sqrt{2}} + \underbrace{\langle f, \cos(t)\rangle}_{a_1} \cos(t) + \underbrace{\langle f, \sin(t)\rangle}_{b_1} \sin(t) + ... + \langle f, \cos(n\,t)\rangle \cos(n\,t) + \langle f, \sin(n\,t)\rangle \sin(n\,t)$$

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

We write

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

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

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

Then

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

The infinite series converges to $f(t)$ pointwise at places 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

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

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

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

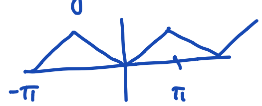
$$b_k = \langle f, \sin(k\,t)\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\sin(k\,t)\,\mathrm{d}t.$$

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

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

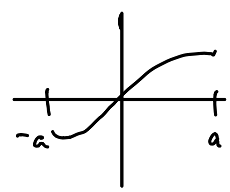
n

On HW $g(t) = |t| \quad -\pi < t < \pi$



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

$$\boxed{\begin{aligned} &g(t) = t \text{ is } \underline{\text{odd}} \\ &\quad g(-t) = -g(t) \end{aligned}}$$



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

finish on Monday!

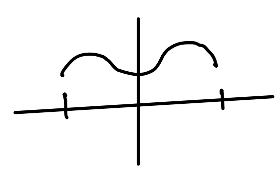
$$a_n = \langle f, \cos nt\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt$$

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

integral of odd fcn over interval $[-a, a]$ is zero.

odd fcn   even fcn

$h(t)$ even means $h(-t) = h(t)$



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

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

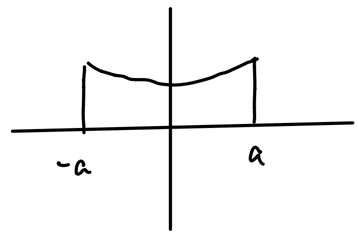
$$g(-t)h(-t) = -g(t)h(t)$$

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

if $g(t), h(t)$ are odd
then $g(t)\,h(t)$ is even

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

if $f(t)$ even $\quad \displaystyle\int_{-a}^{a} f(t)\,dt = 2\int_0^a f(t)\,dt$



$$b_n = \frac{1}{\pi} 2 \int_0^{\pi} \underbrace{t}_{u}\ \underbrace{\sin nt\, dt}_{dv}$$

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

$$= \frac{2}{\pi}\left[uv - \int v\,du \right]$$

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

$$\downarrow \quad \cancel{-\frac{\sin nt}{n^2}\Big]_0^{\pi}} \to 0$$

$$= \frac{2}{\cancel{\pi}}\left[\cancel{\pi}\left(-\frac{\cos n\pi}{n}\right) - 0\right]$$

$$= \frac{2}{n}\Big(-\underbrace{\cos n\pi}_{(-1)^n}\Big)$$

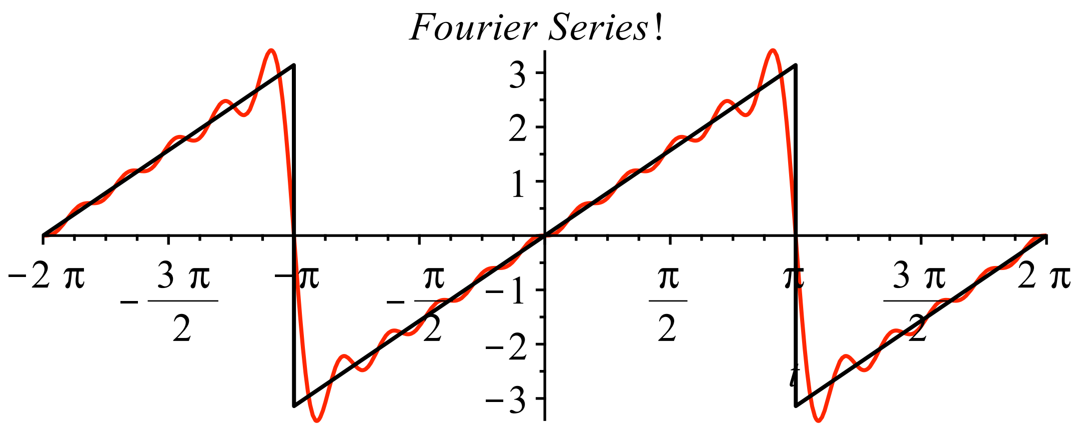
$$\boxed{b_n = \frac{2}{n}(-1)^{n+1}}$$

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

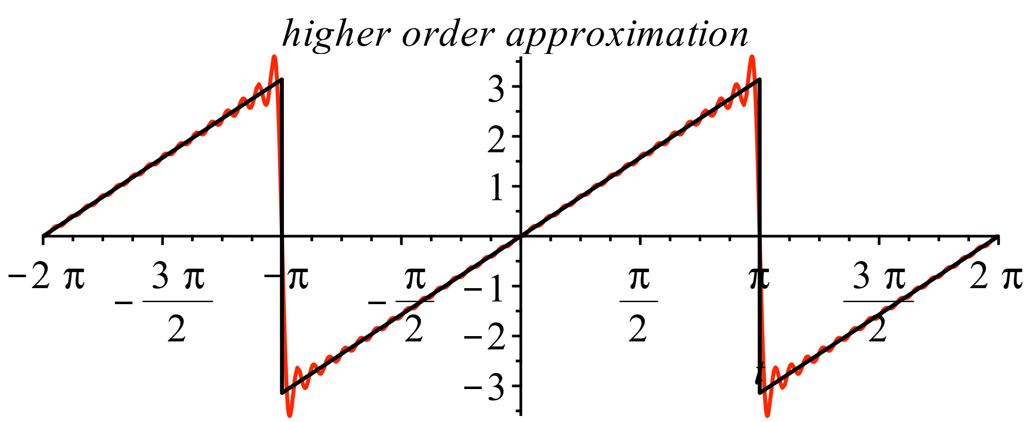
$$\| t - \mathrm{proj}_{J_n} t \| \to 0$$

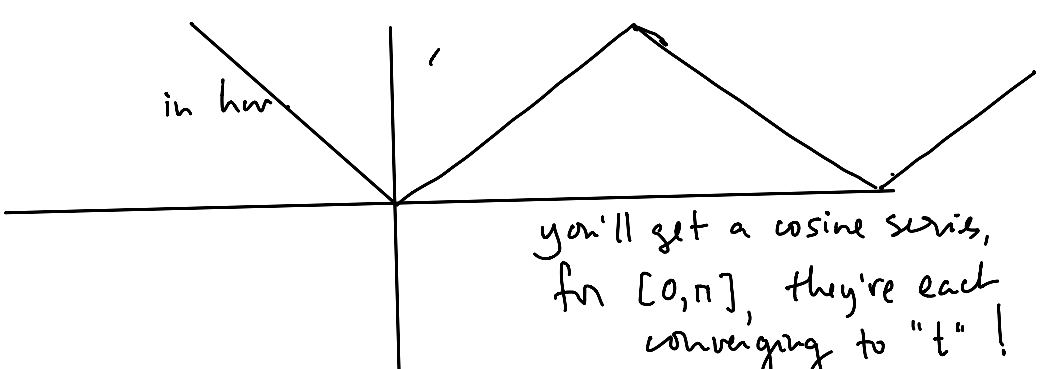
$$proj_{V_{10}}\, f(t):$$

> $with(plots):$
$plot1 := plot(t + 2\cdot\pi - 2\cdot\pi\cdot \mathrm{Heaviside}(t+\pi) - 2\cdot\pi\cdot\mathrm{Heaviside}(t-\pi), t = -2\cdot\pi..2\cdot\pi, color = black):$
$$plot2 := plot\left(2\cdot\sum_{n=1}^{10}(-1)^{n+1}\cdot\frac{\sin(n\cdot t)}{n}, t=-2\cdot\pi..2\cdot\pi, color = red\right):$$
$display(\{plot1, plot2\}, title = \text{'Fourier Series!'});$



$$> plot3 := plot\left(2\cdot\sum_{n=1}^{30}(-1)^{n+1}\cdot\frac{\sin(n\cdot t)}{n}, t=-2\cdot\pi..2\cdot\pi, color = red\right):$$
$display(\{plot1, plot3\}, title = \text{'higher order approximation'});$





you'll get a cosine series,
for [0, π], they're each
converging to "t"!

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

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

$\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ orthonormal

$\to (c_1\vec{u}_1 + c_2\vec{u}_2 + c_3\vec{u}_3) \cdot (c_1\vec{u}_1 + c_2\vec{u}_2 + c_3\vec{u}_3)$

$= c_1^2 + c_2^2 + c_3^2$

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

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

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

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

$$\int t^2 = \frac{t^3}{3}$$

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

$$\| f \|^2 = \frac{1}{\pi}\int_{-\pi}^{\pi} t^2 \, \mathrm{d}t = \frac{2}{\pi}\int_{0}^{\pi} t^2 \, \mathrm{d}t = \frac{2}{3}\pi^2. \qquad \frac{2}{\pi}\,\frac{t^3}{3}\Big]_0^{\pi} = \frac{2}{3}\pi^2$$

Since

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

$$b_n = \frac{2(-1)^{n+1}}{n} \qquad b_n^2 = \frac{4}{n^2}$$

It must be that

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

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots = \sum_{k=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

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

> $\sum_{k=1}^{\infty} \frac{1}{k^2}$;

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

·jpg image compression show & tell. More info at Wikipedia

Prof. Nesse

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 actually uses cosine series

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.

Tues Apr 17

- Symmetric matrices and the spectral theorem, 7.1-7.2

Chapter 7: spectral theorem & applications

Announcements:

Warm-up Exercise: Find eigenvalues & eigenspace bases for

$$A = \begin{bmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 0 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 0 \end{bmatrix} = A$$

$$|A - \lambda I| = (\lambda-2)(\lambda+1).$$

$$A - \lambda I = \begin{bmatrix} 1-\lambda & \sqrt{2} \\ \sqrt{2} & -\lambda \end{bmatrix}$$

$$|A-\lambda I| = \lambda^2 - \lambda - 2 = (\lambda-2)(\lambda+1)$$

$$E_{\lambda=2}: \quad \left[\begin{array}{cc|c} -1 & \sqrt{2} & 0 \\ \sqrt{2} & -2 & 0 \end{array}\right]$$

$$Row_2 = -\sqrt{2}\, Row_1$$

$$\left[\begin{array}{cc|c} -1 & \sqrt{2} & 0 \\ 0 & 0 & 0 \end{array}\right]$$

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

$$E_{\lambda=-1} \quad \left[\begin{array}{cc|c} 2 & \sqrt{2} & 0 \\ \sqrt{2} & 1 & 0 \end{array}\right]$$

$$Row_1 = \sqrt{2}\, Row_2$$

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

$$\begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -\sqrt{2} \end{bmatrix} = 0 \,!$$

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.

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

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

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

③ reflect across diag.

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

<u>Def</u> A square matrix is *symmetric* if and only if $A^T = A$.

<u>Exercise 1</u> Which of the following matrices is symmetric, and which is not?

1a)

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

B is symmetric

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

$$AP = PD$$

$$\cancel{A = P^{-1} D P = P^T D P}$$

$$A = PDP^{-1} = PDP^T$$

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

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 all of 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

$$2x^2 + 2y^2 + 5xy = 1.$$

Can you identify the curve as a conic section? Can you graph it? Note the xy term!

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

$$\nabla f = [f_x, f_y] = [4x + 5y, 4y + 5x] = [0, 0] \text{ at the point } (0, 0),$$

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

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

$$2x^2 + \boxed{2y^2} + 5xy = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & \boxed{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

$$= [x, y] \begin{bmatrix} 2x + \frac{5}{2}y \\ \frac{5}{2}x + \boxed{2}y \end{bmatrix}$$

$$= x\left(2x + \frac{5}{2}y\right) + y\left(\frac{5}{2}x + 2y\right)$$

$$2x^2 + \frac{5}{2}xy + \frac{5}{2}yx + 2y^2$$

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

$$\underline{\boldsymbol{v}}^T{}_{1 \times \cancel{n}} A_{\cancel{n} \times \cancel{n}} \underline{\boldsymbol{w}}_{\cancel{n} \times 1} = [\vec{v}^T A \vec{w}]_{1\times 1} \quad \text{scalar}$$

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 (x_1, x_2, x_n) one can rewrite the quadratic form as

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

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

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

a_{11} for x_1^2

a_{22} for x_2^2

a_{33} for x_3^2

e.g. $x_1^2 + 4x_2^2 + 2x_3^2 + 2x_1x_2 - 1x_1x_3 + 5x_2x_3$

"quadratic form"

$$= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1/2 \\ 1 & 4 & 5/2 \\ -1/2 & 5/2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$1\, x_1x_2 + 1\, x_2x_1$

$-\frac{1}{2} x_1x_3 - \frac{1}{2} x_3x_1$

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

$$2x^2+2y^2+5xy=\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

Check eigendata

$$\begin{bmatrix} 2 & 5/2 \\ 5/2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}=\begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix}=-\frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \checkmark \qquad A\vec{v}=\lambda\vec{v}$$

$$\begin{bmatrix} 2 & 5/2 \\ 5/2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}=\begin{bmatrix} 9/2 \\ 9/2 \end{bmatrix}=\frac{9}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \checkmark \qquad \begin{bmatrix} 1 \\ -1 \end{bmatrix}\cdot\begin{bmatrix} 1 \\ 1 \end{bmatrix}=0$$

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

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

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: Let A be symmetric, and let

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

with $\lambda_1 \neq \lambda_2$. Because $A^T = A$, we claim that

$$\underline{w} \cdot A\,\underline{v} = A\,\underline{w} \cdot \underline{v}.$$

(in general $\vec{w} \cdot A\vec{v} = A^T\vec{w} \cdot \vec{v}$)

One way to see this is by noting

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

• $(AB)^T = B^T A^T$

Since the result of this operation is a scalar, it equals its transpose:

$$\boxed{\underbrace{\left(\underline{w}^T A\,\underline{v}\right)}_{[\ ]_{1\times 1}}} = \left(\underline{w}^T A\,\underline{v}\right)^T = \underline{v}^T A^T \underline{w} = \underline{v}^T A\,\underline{w} = \boxed{\underline{v} \cdot A\,\underline{w}\,.}$$

$A^T = A$

But

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

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

$\lambda_1\, \vec{v} \cdot \vec{w} = \lambda_2\, \vec{v} \cdot \vec{w}$

$\lambda_1 \neq \lambda_2 \Rightarrow \vec{v} \cdot \vec{w} = \vec{0}.$

So, since $\lambda_1 \neq \lambda_2$ is must be that $\underline{v} \cdot \underline{w} = 0$!

* And a special fact for 2 × 2 symmetric matrices and eigenvectors in $\mathbb{R}^2$: If $A\,\underline{v} = \lambda\,\underline{v}$ for $\underline{v} \neq 0$ let $\underline{w} \perp \underline{v}$. Then $\underline{w}$ is automatically an eigenvector:

$$\underline{w} \cdot A\,\underline{v} = \underline{w} \cdot (\lambda\,\underline{v}) = \lambda\,\underline{w} \cdot \underline{v} = 0.$$

true but don't need it,

So

$$0 = \underline{w} \cdot A\,\underline{v} = \underline{v} \cdot A\,\underline{w} \Rightarrow \underline{v} \perp A\,\underline{w} \Rightarrow A\,\underline{w} \in span\{\underline{w}\}$$

because we're in $\mathbb{R}^2$. So $\underline{w}$ is also an eigenvector, automatically.

Theorem: Spectral theorem for 2×2 symmetric matrices

Let $= A = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$. $A - \lambda I = \begin{bmatrix} a-\lambda & c \\ c & b-\lambda \end{bmatrix}$

$$|A - \lambda I| = (\lambda - a)(\lambda - b) - c^2$$
$$(a-\lambda)(b-\lambda)$$
$$= \lambda^2 - (a+b)\lambda + ab - c^2$$
$$\lambda = \frac{(a+b) \pm \sqrt{(a+b)^2 - 4(ab - c^2)}}{2}$$
$$= \frac{(a+b) \pm \sqrt{a^2 + 2ab + b^2 - 4ab + 4c^2}}{2}$$

Continuing ...

Else $(a-b)^2+4c^2=0$

$\Rightarrow a=b,\ c=0$

$A=\begin{bmatrix} a & 0 \\ 0 & a\end{bmatrix}$ & $\{\vec{e}_1,\vec{e}_2\}$ o.n. eigenbasis

$$\lambda = \frac{a+b \pm \sqrt{(a-b)^2+4c^2}}{2}$$

real if $(a-b)^2+4c^2>0$ then $\lambda_1 \neq \lambda_2$, $\vec{v}_1 \perp \vec{v}_2$

So $\left\{ \frac{\vec{v}_1}{\|\vec{v}_1\|}, \frac{\vec{v}_2}{\|\vec{v}_2\|} \right\}$ o.n. basis for $\mathbb{R}^2$.

$$2x^2+2y^2+5xy = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

and for

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

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).

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

$\vec{x}^T A \vec{x} = 1$

$$\cancel{A = P^{-1} D P = P^T D P}$$

$AP = DP$

$\boxed{P^{-1}AP = D}$

$\boxed{P^TAP = D} = \begin{bmatrix} 9/2 & 0 \\ 0 & -1/2 \end{bmatrix}$

Note

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

where as always,

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

For $\underline{\boldsymbol{v}} \in \mathbb{R}^2$ write $\underline{\boldsymbol{v}} = \begin{bmatrix} x \\ y \end{bmatrix}$ in standard coordinates and $[\underline{\boldsymbol{v}}]_B = \begin{bmatrix} x' \\ y' \end{bmatrix}$. (The text uses $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ respectively.) So the two coordinate systems are related by

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

$\vec{x} = P\vec{y}$

Do algebra!

$$2x^2 + 2y^2 - 5xy = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\vec{x}^T A \vec{x}$$

$$\vec{x} = P\vec{y}$$

$$= \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & +\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

$$\vec{y}^T \underbrace{P^T A P}_{\substack{= \\ D}} \vec{y}$$

$$= \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} \frac{9}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \quad \text{(because } P^T AP = \text{D)}$$

$$= \frac{9}{2}(x')^2 - \frac{1}{2}(y')^2.$$

So the original curve with equation

$$2x^2 + 2y^2 + 5xy = 1 \qquad \text{standard coords}$$

in the standard coordinate system has equation

$$\frac{9}{2}(x')^2 - \frac{1}{2}(y')^2 = 1 \qquad \mathcal{B}\text{-coords}$$

with respect to the rotated coordinate system!

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

$$\frac{(x')^2}{\left(\sqrt{\frac{2}{3}}\right)^2} - \frac{(y')^2}{(\sqrt{2})^2} = 1.$$

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

$$f(x', y') = \frac{9}{2}(x')^2 - \frac{1}{2}(y')^2.$$

Old pictures from when I could still sketch well:

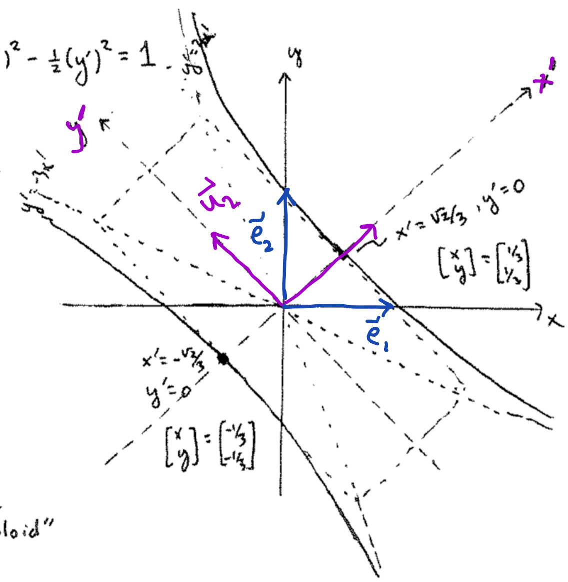
$$\{\vec{u}_1, \vec{u}_2\} = \left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \right\}$$

$$= [x' \;\; y'] \begin{bmatrix} 9/2 & 0 \\ 0 & -1/2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{9}{2}(x')^2 - \frac{1}{2}(y')^2 = 1$$

ans to 1

Curve is a hyperbola! $\dfrac{(x')^2}{(\sqrt{2}/3)^2} - \dfrac{(y')^2}{(\sqrt{2})^2} = 1$

$$\frac{x'^2}{a^2} - \frac{y'^2}{b^2}$$

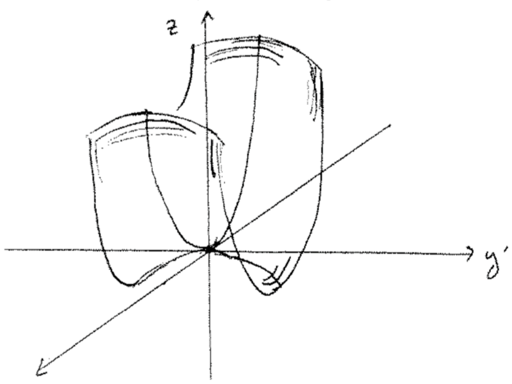


NO! not a local min!

ans to 2 $\quad z = 2x^2 + 2y^2 + 5xy$

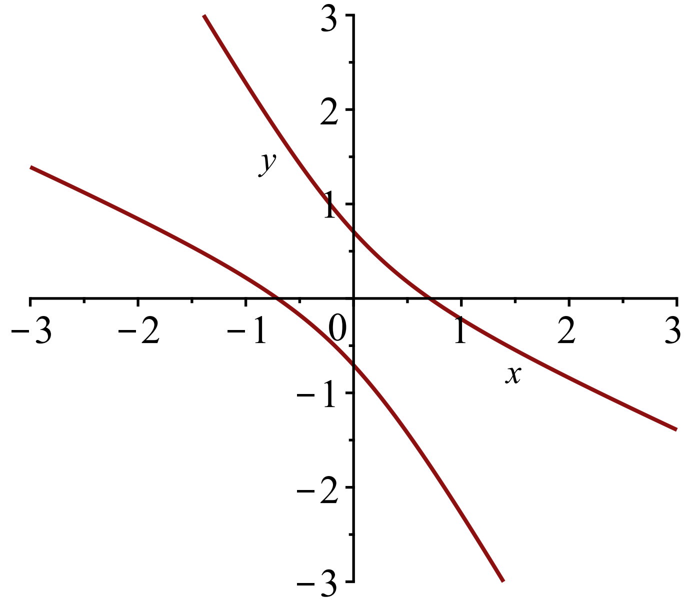
$z = \frac{9}{2}(x')^2 - \frac{1}{2}(y')^2$ Saddle surface!

"parabolic hyperboloid"

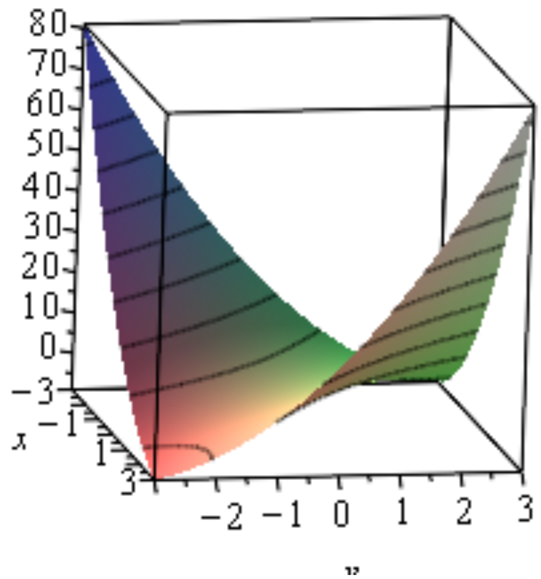


Maple verification: To be continued ....

> $with(plots):$
$implicitplot(2\cdot x^2 + 2\cdot y^2 + 5\cdot x\cdot y = 1, x = -3..3, y = -3..3, grid = [200, 200]);$



> $plot3d(2\cdot x^2 + 2\cdot y^2 + 5\cdot x\cdot y, x = -3..3, y = -3..3);$



Wed Apr 18

- 7.1-7.2 Diagonalizing quadratic forms and surfaces and curves defined implicitly with quadratic equations, via the spectral theorem continued; with proof of spectral theorem appended.

Announcements:
- Guest lecture on principal component analysis (§7.5) Friday (Prof. Tom Alberts)
- prep notes

Warm-up Exercise: Compute

$$\underset{A}{\begin{bmatrix} 1 & -2 \\ 2 & 0 \end{bmatrix}} \underset{B}{\begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & 1 \end{bmatrix}} = \begin{bmatrix} 1 & -4 & 1 \\ 2 & 0 & 6 \end{bmatrix}$$

"inner product"

then compute $\sum_{j=1}^{2} [col_j(A)][row_j(B)]$

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

$$= \begin{bmatrix} 1 & 0 & 3 \\ 2 & 0 & 6 \end{bmatrix} + \begin{bmatrix} 0 & -4 & -2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -4 & 1 \\ 2 & 0 & 6 \end{bmatrix}$$

"outer product"

Spectral Theorem Let A be an $n \times n$ symmetric matrix. Then all of the eigenvalues of A are real, and there exists an orthonormal eigenbasis $B = \{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \dots \underline{\boldsymbol{u}}_n\}$ consisting of eigenvectors for A. Eigenspaces with different eigenvalues are automatically orthogonal to each other. If any eigenspace has dimension greater than 1, its orthonormal basis may be constructed via *Gram − Schmidt*.   (Proof of spectral theorem at end of today's notes.)

Diagonalization of quadratic forms: Let

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

for a symmetric matrix A, with real entries. A symmetric $\Rightarrow$ by the spectral theorem there exists an orthonormal eigenbasis $B = \{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \dots \underline{\boldsymbol{u}}_n\}$.

For the corresponding orthogonal matrix

$$P = [\underline{\boldsymbol{u}}_1 | \underline{\boldsymbol{u}}_2 | \cdots | \underline{\boldsymbol{u}}_n]$$

$$D = P^T A P$$

$AP = PD$

$P^{-1} A P = D$

$P^{-1} = P^T$

where D is the diagonal matrix of eigenvalues corresponding to the eigenvectors in P. And we have

$$\underline{\boldsymbol{x}} = P \underline{\boldsymbol{y}}$$

where $\underline{\boldsymbol{y}} = [\underline{\boldsymbol{x}}]_B$ and $P = P_{E \leftarrow B}$. Thus

$\vec{x}^T = \vec{y}^T P^T$

$$Q(\underline{\boldsymbol{x}}) = \underline{\boldsymbol{x}}^T A \underline{\boldsymbol{x}}$$

$$= \underline{\boldsymbol{y}}^T P^T A P \underline{\boldsymbol{y}} = \underline{\boldsymbol{y}}^T D \underline{\boldsymbol{y}}$$

$$= \sum_{i=1}^{n} \lambda_i y_i^2 .$$

So by the orthogonal change of variables all cross terms have been removed. Applications include conic curves, quartic surfaces, multivariable second derivative test, singular value decomposition theorem, and more.

Material we need for Prof. Alberts' guest lecture Friday on Principal Component Analysis. (The text discusses most of this background material in 7.1, 7.2)

Definition: The quadratic form $Q(\underline{\boldsymbol{x}}) = \sum_{i,j=1}^{n} a_{ij} x_i x_j = \underline{\boldsymbol{x}}^T A \underline{\boldsymbol{x}}$ (for A a symmetric matrix) is called *positive definite* if

$$Q(\underline{\boldsymbol{x}}) > 0 \quad \text{for all } \underline{\boldsymbol{x}} \neq \underline{\boldsymbol{0}}.$$

From the previous page, we see that this is the same as saying that all of the eigenvalues of A are positive.

Theorem: The "outer product" way of computing the matrix product $A\,B$. (Section 2.4 topic on partitioned matrices that we skipped....our usual way is with dot product or rows of A with columns of B, aka an "inner product").

(1) first, notice that the product of an $m \times 1$ column vector with a $1 \times n$ row vector is an $m \times n$ matrix:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \end{bmatrix} = \begin{bmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \\ a_3 b_1 & a_3 b_2 \end{bmatrix}.$$

(1) Let $A_{m \times p}$ and $B_{p \times n}$. Express A in terms of its columns, and B in terms of its rows:

$$A = \begin{bmatrix} | & | & & | \\ \underline{\boldsymbol{a}}_1 & \underline{\boldsymbol{a}}_2 & \ldots & \underline{\boldsymbol{a}}_p \\ | & | & & | \\ | & | & & | \end{bmatrix} \qquad B = \begin{bmatrix} ---\underline{\boldsymbol{b}}_1--- \\ ---\underline{\boldsymbol{b}}_2--- \\ \vdots \\ ---\underline{\boldsymbol{b}}_p--- \end{bmatrix}.$$

Then

$$A\,B = \sum_{j=1}^{p} \underline{\boldsymbol{a}}_j \underline{\boldsymbol{b}}_j.$$

We can illustrate the general proof by considering the example in which A and B are each 3 × 3: Look column by column in the output of each expression to verify the identity, using the the linear combination form of matrix times vector, for $A\,B$:

$$A\,\mathrm{col}_1(\vec{B}) = b_{11}\,\mathrm{col}_1(A) + b_{21}\,\mathrm{col}_2(A) + b_{31}\,\mathrm{col}_3(A)$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} & b_{13} \end{bmatrix} + \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix}\begin{bmatrix} b_{21} & b_{22} & b_{23} \end{bmatrix} + \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix}\begin{bmatrix} b_{31} & b_{32} & b_{33} \end{bmatrix}.$$

actual proof with formula

via dot product: $\mathrm{entry}_{k\ell}\, AB = \mathrm{row}_k(A)\cdot \mathrm{col}_\ell(B) = \sum_{j=1}^{p} a_{kj}\, b_{j\ell}$

Same!

via outer product: $\mathrm{entry}_{k\ell} \begin{bmatrix} \vec{a}_j \end{bmatrix}\begin{bmatrix} - \vec{b}_j - \end{bmatrix} = a_{kj}\, b_{j\ell}$

$$\sum_{j=1}^{p} \mathrm{entry}_{k\ell} \begin{bmatrix} \vec{a}_j \end{bmatrix}\begin{bmatrix} - \vec{b}_j - \end{bmatrix} = \sum_{j=1}^{p} a_{kj}\, b_{j\ell}$$

$$\parallel$$

$$\mathrm{entry}_{k\ell} \sum_{j=1}^{p} \begin{bmatrix} \vec{a}_j \end{bmatrix}\begin{bmatrix} - \vec{b}_j - \end{bmatrix}$$

Spectral decomposition for symmetric matrices. Let $A_{n \times n}$ be symmetric (and positive definite, for the applications Prof. Alberts will talk about on Friday). Order the eigenvalues as

$$\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n > 0$$

and let

$$\{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \cdots, \underline{\boldsymbol{u}}_n\}$$

be a corresponding orthonormal eigenbasis of $\mathbb{R}^n$. Let P be the orthogonal matrix

$$P = [\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \ldots \underline{\boldsymbol{u}}_n]$$

with

$$A\,P = P\,D$$

where D is the diagonal matrix with diagonal entries $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n > 0$.

$$D = P^T A P$$

$$A P = P D$$

$$A P P^T = P D P^T$$

$$A = P D P^T$$

$$(P^{-1} = P^T)$$

Then

$$A = P\,D\,P^T$$

$$= \underbrace{\begin{bmatrix} | & | & & | \\ \underline{\boldsymbol{u}}_1 & \underline{\boldsymbol{u}}_2 & \ldots & \underline{\boldsymbol{u}}_n \\ | & | & & | \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \ldots & \ldots & 0 \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}} \begin{bmatrix} ---\underline{\boldsymbol{u}}_1^T--- \\ ---\underline{\boldsymbol{u}}_2^T--- \\ \vdots \\ ---\underline{\boldsymbol{u}}_n^T--- \end{bmatrix}$$

$$= \begin{bmatrix} | & | & & | \\ \lambda_1 \underline{\boldsymbol{u}}_1 & \lambda_2 \underline{\boldsymbol{u}}_2 & \ldots & \lambda_n \underline{\boldsymbol{u}}_n \\ | & | & & | \\ | & | & & | \end{bmatrix} \begin{bmatrix} ---\underline{\boldsymbol{u}}_1^T--- \\ ---\underline{\boldsymbol{u}}_2^T--- \\ \vdots \\ ---\underline{\boldsymbol{u}}_n^T--- \end{bmatrix}$$

$$A = \lambda_1 \underline{\boldsymbol{u}}_1 \underline{\boldsymbol{u}}_1^T + \lambda_2 \underline{\boldsymbol{u}}_2 \underline{\boldsymbol{u}}_2^T + \ldots + \lambda_n \underline{\boldsymbol{u}}_n \underline{\boldsymbol{u}}_n^T$$

"Principal component analysis" makes use of the fact that if only a few of the eigenvalues of A are large and the rest are near zero, then the corresponding leading terms in the expression above are a good approximation for the matrix A.

Remark: There's slick way to see this spectral decomposition matrix identity that doesn't use the outer product but uses our work on projection instead:

$$\underline{\boldsymbol{x}} = (\underline{\boldsymbol{x}} \cdot \underline{\boldsymbol{u}}_1)\underline{\boldsymbol{u}}_1 + (\underline{\boldsymbol{x}} \cdot \underline{\boldsymbol{u}}_2)\underline{\boldsymbol{u}}_2 + \ \cdots \ (\underline{\boldsymbol{x}} \cdot \underline{\boldsymbol{u}}_n)\underline{\boldsymbol{u}}_n$$

$$\Rightarrow A\,\underline{\boldsymbol{x}} = (\underline{\boldsymbol{x}} \cdot \underline{\boldsymbol{u}}_1)\lambda_1\underline{\boldsymbol{u}}_1 + (\underline{\boldsymbol{x}} \cdot \underline{\boldsymbol{u}}_2)\lambda_2\underline{\boldsymbol{u}}_2 + \quad (\underline{\boldsymbol{x}} \cdot \underline{\boldsymbol{u}}_n)\lambda_n\underline{\boldsymbol{u}}_n$$

$$= \lambda_1\, \underline{\boldsymbol{u}}_1\left(\underline{\boldsymbol{u}}_1^T\, \underline{\boldsymbol{x}}\right) + \lambda_2\underline{\boldsymbol{u}}_2\left(\underline{\boldsymbol{u}}_2^T\underline{\boldsymbol{x}}\right) + \quad + \lambda_n\underline{\boldsymbol{u}}_n\left(\underline{\boldsymbol{u}}_n^T\underline{\boldsymbol{x}}\right)$$

$$A\,\underline{\boldsymbol{x}} = \left[\ \lambda_1\, \underline{\boldsymbol{u}}_1\underline{\boldsymbol{u}}_1^T + \lambda_2\, \underline{\boldsymbol{u}}_2\, \underline{\boldsymbol{u}}_2^T + \ldots + \lambda_n\underline{\boldsymbol{u}}_n\, \underline{\boldsymbol{u}}_n^T\right]\underline{\boldsymbol{x}}\,.$$

Since this is true for all $\underline{\boldsymbol{x}}$, (in particular for the standard basis vectors, which lets us recover the columns of A) we deduce

$$A = \lambda_1\, \underline{\boldsymbol{u}}_1\underline{\boldsymbol{u}}_1^T + \lambda_2\, \underline{\boldsymbol{u}}_2\, \underline{\boldsymbol{u}}_2^T + \ldots + \lambda_n\underline{\boldsymbol{u}}_n\, \underline{\boldsymbol{u}}_n^T\,.$$

Testing spectral decomposition in a small example:

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

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

$$\underline{\boldsymbol{u}}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \underline{\boldsymbol{u}}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\lambda_1 \underline{\boldsymbol{u}}_1 \underline{\boldsymbol{u}}_1^T + \lambda_2 \underline{\boldsymbol{u}}_2 \underline{\boldsymbol{u}}_2^T = \frac{9}{2} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix} - \frac{1}{2} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \end{bmatrix}$$

$$= \frac{9}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \frac{1}{4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix} !!!!$$

We didn't finish this page last week

Definition A square $n \times n$ matrix Q is called *orthogonal* if its columns are ortho-normal. (You can read more about orthogonal matrices at e.g. Wikipedia.)

Theorem. Let Q be an orthogonal matrix. Then

a) $Q^{-1} = Q^T$.

$$\text{entry}_{ij}\, Q^T Q = [\text{Row}_i\, Q^T]\begin{bmatrix}\text{Col}_j Q\end{bmatrix} = \text{Col}_i Q \cdot \text{col}_j Q = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

$$\Rightarrow Q^T Q = I$$

$$\Rightarrow Q Q^T = I \text{ also (prev. Chapter).}$$

i.e. $Q^{-1} = Q^T$

b) The rows of Q are also ortho-normal.

$$Q Q^T = I$$

$$ij \text{ entry: } \text{row}_i Q \cdot \underbrace{\text{col}_j Q^T}_{\text{row}_j(Q)} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

c) the transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$T(\underline{\boldsymbol{x}}) = Q\, \underline{\boldsymbol{x}}$$

preserves dot products and magnitudes, (so also volumes, since cubes generated by perpendicular vectors will be transformed into equal-volume cubes). In other words, for all $\underline{\boldsymbol{x}}, \underline{\boldsymbol{y}} \in \mathbb{R}^n$,

$$T(\underline{\boldsymbol{x}}) \cdot T(\underline{\boldsymbol{y}}) = \underline{\boldsymbol{x}} \cdot \underline{\boldsymbol{y}}$$

$$\|T(\underline{\boldsymbol{x}})\| = \|\underline{\boldsymbol{x}}\|.$$

d) The only matrix transformations $T : \mathbb{R}^n \to \mathbb{R}^n$ that preserve dot products are orthogonal transformations. (These transformations are often referred to as *isometries*.)

Example Identify and sketch the surface defined implicitly by

$$x_1^2 + x_2^2 - 2x_3^2 - 2x_1x_2 - 4x_1x_3 - 4x_2x_3 = 8.$$

Exercise 1) Find the symmetric matrix so that

$$x_1^2 + x_2^2 - 2x_3^2 - 2x_1x_2 - 4x_1x_3 - 4x_2x_3 = \underline{\mathbf{x}}^T A\,\mathbf{x}.$$

Recall that

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

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & -1 & -2 \\ -1 & 1 & -2 \\ -2 & -2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

If we found the matrix correctly technology tells us that

$$E_{\lambda=-2} = span\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\},\ E_{\lambda=2} = span\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right\},\ E_{\lambda=4} = span\left\{\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}\right\}.$$

(positively oriented in this order)

$$x_1^2 + x_2^2 - 2x_3^2 - 2x_1x_2 - 4x_1x_3 - 4x_2x_3 = 8$$

$$\underline{\boldsymbol{x}}^T A\,\underline{\boldsymbol{x}} = 8$$

For

$$P = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix}$$

$$P^T A P = D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$\underline{\boldsymbol{x}}^T A\,\underline{\boldsymbol{x}} = 8$$

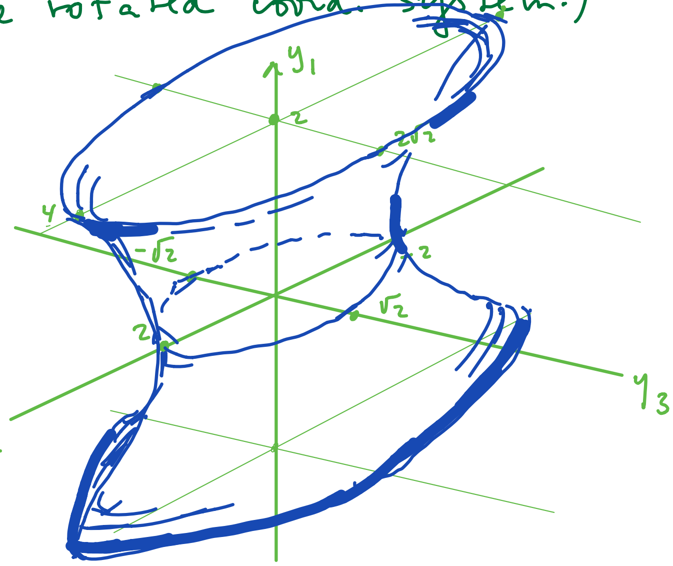
$$\underline{\boldsymbol{y}}^T P^T A\,P\,\underline{\boldsymbol{y}} = 8$$

$$\underline{\boldsymbol{y}}^T \mathrm{D}\,\underline{\boldsymbol{y}} = 8$$

$$-2y_1^2 + 2y_2^2 + 4y_3^2 = 8. \qquad 2y_2^2 + 4y_3^2 = 8 + 2y_1^2$$

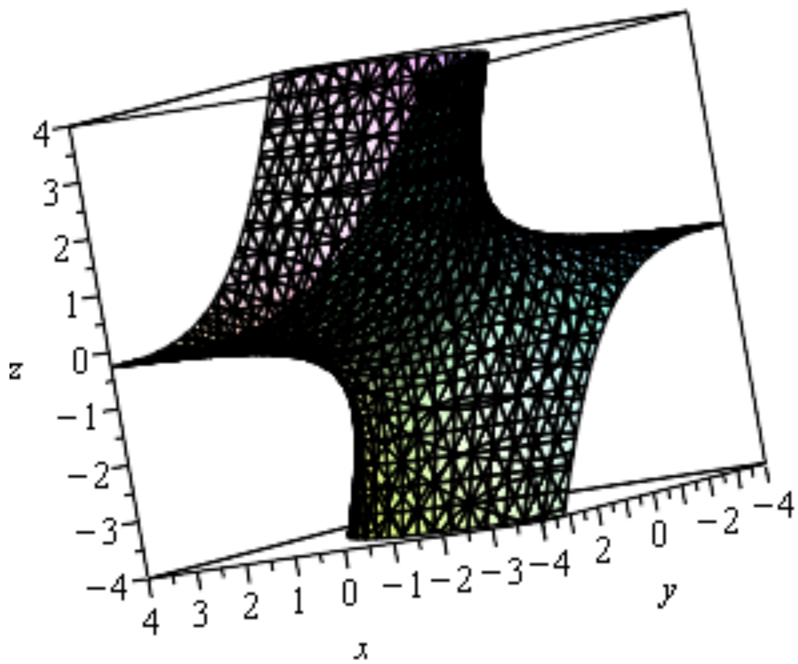
We can try to sketch this in the rotated coordinate system. It's an elliptic hyperboloid of one sheet. :-)

(Notice that when all is said & done, we only needed the eigenvalues of A to sketch the quadric surface with respect to the rotated coord. system.)



> $with(plots):$

$implicitplot3d(x^2 + y^2 - 2\,z^2 - 2\,x \cdot y - 4 \cdot x \cdot z - 4 \cdot y \cdot z = 8, x = -4..4, y = -4..4, z = -4..4, grid = [20, 20, 20]);$



eigenvalues{{1,-1,-2},{-1,1,-2},{-2,-2,2}}

Browse Examples Surprise Me

Input:

eigenvalues $\begin{pmatrix} 1 & -1 & -2 \\ -1 & 1 & -2 \\ -2 & -2 & 2 \end{pmatrix}$

Open code

Results: Step-by-step solution

$\lambda_1 = 4$

$\lambda_2 = -2$

$\lambda_3 = 2$

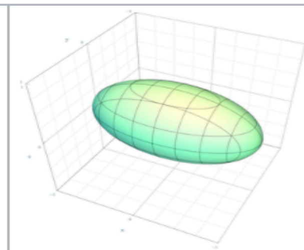
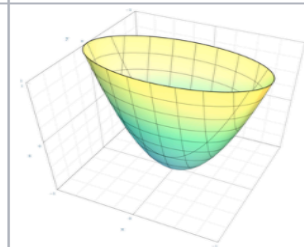
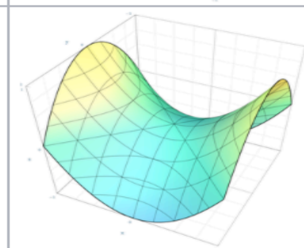
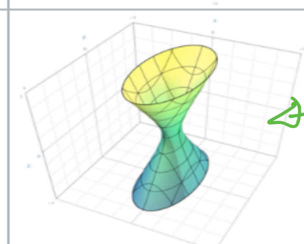
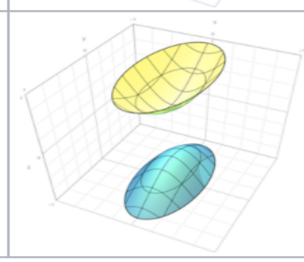
Corresponding eigenvectors: Step-by-step solution

$v_1 = (-1, -1, 2)$

$v_2 = (1, 1, 1)$

$v_3 = (-1, 1, 0)$

from Wikipedia, "quadric surfaces". There is also a Wikipedia page on conic sections.

Non-degenerate real quadric surfaces		
Ellipsoid	$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$	
Elliptic paraboloid	$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-z=0$$	
Hyperbolic paraboloid	$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-z=0$$	
Elliptic hyperboloid of one sheet	$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$	
Elliptic hyperboloid of two sheets	$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1$$	

Spectral Theorem Let $A_{n\times n}$ be a real, symmetric matrix.
Then $\exists$ an orthonormal $\mathbb{R}^n$ basis made of eigenvectors of A, $\mathcal{B} = \{\vec u_1, \vec u_2, \ldots \vec u_n\}$ $\quad A\vec u_j = \lambda_j \vec u_j$.
Thus for $S = [\vec u_1 | \vec u_2 | \cdots \vec u_n]$,

$$S^T A S = \begin{bmatrix} \lambda_1 & & O \\ & \lambda_2 \ddots & \\ O & & \lambda_n \end{bmatrix} \text{ is diagonal.}$$

proof

1. On Monday we showed that if $\lambda_1 \neq \lambda_2$ are real eigenvalues of A, with eigenvectors $\vec v_1, \vec v_2 \neq \vec 0$

$$A\vec v_1 = \lambda_1 \vec v_1 \qquad\qquad \text{Then } \vec v_1 \perp \vec v_2$$
$$A\vec v_2 = \lambda_2 \vec v_2$$

proof $v_2^T A \vec v_1 = v_2^T(\lambda_1 \vec v_1) = \lambda_1 \vec v_2 \cdot \vec v_1$

$$\begin{aligned} &\| \\ (v_2^T A^T)\vec v_1 &= (A\vec v_2)^T \vec v_1 \\ &= \lambda_2 \vec v_2^T \vec v_1 \\ &= \lambda_2 \vec v_2 \cdot \vec v_1 \end{aligned}$$

So $\vec v_2 \cdot \vec v_1 = 0$ ∎

We also showed that for $A_{2\times 2}$ symmetric, either A is already diagonal (a multiple of I, in fact), or A has 2 distinct real eigenvalues $\Rightarrow A$ diagonalizable. By 1 the eigenvectors are $\perp$, so normalize to get orthonormal eigenbasis.

2. All eigenvalues of A are real: Let $\lambda = a + bi$ be any root of $\overset{p(\lambda)}{\cancel{f_A(\lambda)}}$, and let $\vec u + i\vec v$ be a corresponding non-zero eigenvector.

$$A(\vec u + i\vec v) = (a+ib)(\vec u + i\vec v)$$

take conjugates: $A(\vec u - i\vec v) = (a-ib)(\vec u - i\vec v)$.

Now consider

$$\underbrace{(\vec u - i\vec v)^T A}(\vec u + i\vec v) = (\vec u - i\vec v)^T(a+ib)(\vec u + i\vec v) = (a+ib)\left[(\vec u - i\vec v)^T(\vec u + i\vec v)\right]$$
$$= (a+ib)\left[(\vec u - i\vec v)\cdot(\vec u + i\vec v)\right]$$
$$= (a+ib)\left(\|\vec u\|^2 + \|\vec v\|^2\right)$$

$$\|$$
$$\left((\vec u - i\vec v)^T A^T\right)(\vec u + i\vec v)$$
$$\|$$
$$[A(\vec u - i\vec v)]^T(\vec u + i\vec v) = (a-ib)(\vec u - i\vec v)^T(\vec u + i\vec v)$$
$$= (a-ib)\left(\|\vec u\|^2 + \|\vec v\|^2\right)$$

$\Rightarrow b = 0!$

∎

thus $\overset{p(\lambda)}{\cancel{f_A(\lambda)}}$ factors completely over $\mathbb{R}$.

- If it has n distinct roots, then alg mult = geom mult = 1, all evects w different evals are $\perp$ by 1, and normalize to ~~get~~ orthonormal eigenbasis
- Otherwise it's a little harder: (In practice, if λ_i has alg & geom mult $k_i > 1$ just Gram-~~Sm~~ Schmidt its eigenbasis)

3. General proof, by induction:
Spectral Theorem true for $n=1$ (1×1 matrices are diagonal)
$n=2$ (we checked yesterday).

Inductive step:

- Assume all $(n-1)\times(n-1)$ symmetric matrices are diagonalizable with an orthogonal matrix (with eigenbasis columns).

- Now let $A_{n\times n}$ symmetric

Let λ_1 be any root of $\cancel{f_A(\lambda)}$ $p(\lambda)$. λ_1 is real by 2.

Let $\vec{u}_1$ be a unit eigenvector

$$A\vec{u}_1 = \lambda_1\vec{u}_1, \quad \|\vec{u}_1\| = 1.$$

Complete to an orthonormal basis

$$\mathcal{B}_0 = \{\vec{u}_1, \underbrace{\vec{v}_2, \vec{v}_3, \ldots \vec{v}_n}_{\text{probably not eigenvectors}}\} \text{ of } \mathbb{R}^n$$

$$S_0 = \left[\vec{u}_1 \middle| \vec{v}_2 \middle| \ldots \vec{v}_n\right] \qquad S_0^T S_0 = I$$

$S_0^T A S_0$ is symmetric (take its transpose!)

1st column is $\begin{bmatrix}\lambda_1\\0\\\vdots\\0\end{bmatrix}$

so $$S_0^T A S_0 = \left[\begin{array}{c|c}\lambda_1 & \quad 0 \quad \\ \hline 0 & \\ 0 & B \\ \vdots & \\ 0 & \end{array}\right]$$

$B_{(n-1)\times(n-1)}$ is symmetric, so by induction hypothesis $\exists$ S_1 orthog, with $S_1^T B S_1 = \begin{bmatrix}\lambda_2 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix}$ (λ_i's need not be distinct!)

Thus
$$\left[\begin{array}{c|c}1 & 0\\ \hline 0 & S_1^T\end{array}\right]\underbrace{\left[\begin{array}{c|c}\lambda_1 & 0\\ \hline 0 & B\end{array}\right]}_{S_0^T A S_0}\left[\begin{array}{c|c}1 & 0\\ \hline 0 & S_1\end{array}\right] = \begin{bmatrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n\end{bmatrix}$$

so $$S^T A S = D$$

$$S = S_0\left[\begin{array}{c|c}1 & 0\\ \hline 0 & S_1\end{array}\right] \quad \text{orthog (product of orthog)}$$

∎