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...
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: $\forall \mathbf{f}, \mathbf{g}, \mathbf{h} \in V, k \in \mathbb{R}$:

1. $\langle \mathbf{f}, \mathbf{g} \rangle = \langle \mathbf{g}, \mathbf{f} \rangle$ (symmetry)
2. $\langle \mathbf{f}, \mathbf{g} + \mathbf{h} \rangle = \langle \mathbf{f}, \mathbf{g} \rangle + \langle \mathbf{f}, \mathbf{h} \rangle$ (linearity in each factor)
3. $\langle k \mathbf{f}, \mathbf{g} \rangle = k \langle \mathbf{f}, \mathbf{g} \rangle$
4. $\langle \mathbf{f}, \mathbf{f} \rangle \geq 0, \langle \mathbf{f}, \mathbf{f} \rangle = 0 \text{ iff } \mathbf{f} = \mathbf{0}$ (positive)

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

$\| \mathbf{x} \| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$ (norm)
Example for the inner product on $C[-\pi, \pi]$ given 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:

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

Exercise verify how ortho-normality follows from these identities.

$$\langle \cos(m \, t), \cos(n \, t) \rangle = 0 \quad m \neq n.$$
Let \( V_n := \text{span} \left\{ \frac{1}{\sqrt{2}}, \cos(t), \sin(t), \cos(2t), \sin(2t), \ldots, \cos(nt), \sin(nt) \right\} \) 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 continuous), then

\[
\lim_{n \to \infty} \| f - \text{proj}_{V_n} f \| = 0.
\]

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

\[
\text{proj}_{V_n} f = \begin{pmatrix} f, \\ \frac{1}{\sqrt{2}} \end{pmatrix} \frac{1}{\sqrt{2}} + \langle f, \cos(t) \rangle \cos(t) + \langle f, \sin(t) \rangle \sin(t) + \ldots + \langle f, \cos(nt) \rangle \cos(nt) + \langle f, \sin(nt) \rangle \sin(nt).
\]

We 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

\[
\text{proj}_{V_n} f = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kt) + \sum_{k=1}^{n} b_k \sin(kt).
\]

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

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

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

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

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

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

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

\[
g(t) = t \text{ is odd} \quad g(-t) = -g(t)
\]

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

Finish on Monday!

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

\[
= \frac{1}{\pi} \int_{-\pi}^{\pi} t \cos nt \, dt
\]

\[
\text{odd function} \quad \text{even function}
\]

\[
h(t) \text{ even means } h(-t) = h(t)
\]

\[
g(t) \text{ odd}
\]

\[
g(t)h(t) \text{ is odd}
\]

\[
g(-t)h(-t) = -g(t)h(t)
\]
\[ b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} t \sin(nt) \, dt \]

- if \( g(t), h(t) \) are odd
  - \( g(t)h(t) \) is even
    - \( g(-t)h(-t) = (-g(t))(-h(t)) \)

- if \( f(t) \) even
  - \( \int_{-a}^{a} f(t) \, dt = 2 \int_{0}^{a} f(t) \, dt \)

\[ b_n = \frac{1}{\pi} \int_{0}^{\pi} t \sin nt \, dt \]

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

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

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

\[ = \frac{2}{n} \left( -\cos nt \right) \]

\[ b_n = \frac{2}{n} (-1)^{n+1} \]

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

\[ \| t - \text{proj}_n t \| \to 0 \]
proj$_v f(t)$:

```maple
with(plots):
plot1 := plot(t + 2 - 2, Heaviside(t + 1) - 2, t = -2..2, color = black):
plot2 := plot(2 * sum((-1)$n+1 * sin(n*t), t = -2..2, color = red)):
display({plot1, plot2}, title = 'Fourier Series!');
```

**Fourier Series!**

```maple
plot3 := plot(2 * sum((-1)$n+1 * sin(n*t), t = -2..2, color = red)):
display({plot1, plot3}, title = 'higher order approximation');
```

**higher order approximation**

In this way, you'll get a cosine series, for [0, n], they're each converging to "t"!
As part of the deep theorem about Fourier series, as long as \( f \) is piecewise continuous,

\[
\left\| \text{proj}_N f \right\| \to 0 \quad \text{and} \quad \left\| \text{proj}_N f \right\| \to \left\| f \right\|. 
\]

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,

\[
\text{proj}_N f = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kt) + \sum_{k=1}^{n} b_k \sin(kt)
\]

So

\[
\left\| \text{proj}_N f \right\|^2 = \left\| \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kt) + \sum_{k=1}^{n} b_k \sin(kt) \right\|^2
\]

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

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

\[
V_n := \text{span} \left\{ \frac{1}{\sqrt{2}}, \cos(t), \sin(t), \cos(2t), \sin(2t), \ldots, \cos(nt), \sin(nt) \right\}
\]
As an application, for our function \( f(t) = t \),

\[
\int t^2 \, dt = \left. \frac{t^3}{3} \right|_0^\pi = \frac{\pi^3}{3}.
\]

Since

\[
\|f\|^2 = \frac{1}{\pi} \int_{-\pi}^{\pi} t^2 \, dt = \frac{2}{\pi} \int_0^\pi t^2 \, dt = \frac{2}{3} \pi^2.
\]

It must be that

\[
\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.
\]

This magic formula is true (which is sort of amazing), although you may not have seen it before:
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:
\[
\phi_{n,m}(x, y) = \sin \left( \frac{n \pi}{L} x \right) \sin \left( \frac{m \pi}{H} y \right),
\]
where \( n \) and \( m \) both range from 1, 2, 3, …. 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
\[
B_{n,m} = \frac{\langle f, \phi_{n,m} \rangle}{\langle \phi_{n,m}, \phi_{n,m} \rangle} = \langle g, h \rangle = \int_{0}^{L} \int_{0}^{H} g(x, y) h(x, y) dy dx.
\]

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

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

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: 

$$\frac{1}{2} \text{avg}_m(B_{n,m}^2) + \frac{1}{2} \text{avg}_m(B_{m,n}^2) = b_n^2$$

—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 $\gamma$-value as our estimate. For the two images you choose, record the $\gamma$-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 $N$th Fourier truncated image in greyscale. Figure 2 will output the results for experiment 2.
Warm-up Exercise:

Find eigenvalues and eigenspace bases for

\[ A = \begin{bmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 0 \end{bmatrix} \quad \text{and} \quad A^T = \begin{bmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 0 \end{bmatrix} = A \]

\[ (A-\lambda I) = (\lambda-2)(\lambda+1). \]

\[ E_{\lambda=2} = \begin{bmatrix} -1 & \sqrt{2} \\ \sqrt{2} & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \]

\[ \text{Row}_1 = -\sqrt{2} \text{ Row}_1 \]

\[ \text{Row}_1 = -1 \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \]

\[ E_{\lambda=2} = \text{span} \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix} \]

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

\[ (A-\lambda I) = (\lambda^2 - \lambda + 2) \]

\[ E_{\lambda=-1} = \begin{bmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \]

\[ \text{Row}_1 = \sqrt{2} \text{ Row}_2 \]

\[ \text{Row}_1 = \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \]

\[ E_{\lambda=-1} = \text{span} \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix} \]

\[ E_{\lambda=-1} = \text{span} \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix} \]

\[ \begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -\sqrt{2} \\ 1 \end{bmatrix} = 0 ! \]
Recall that the transpose operation swaps rows with columns, and vice versa. 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.

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

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

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

**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) 

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

$B$ is symmetric

1b) 

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

$\text{No}$
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

\[
A P = P D
\]

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

\[
\nabla f = [f_x, f_y] = [4x + 5y, 4y + 5x] = [0, 0]
\]

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

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

\[
= \begin{bmatrix} x, y \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}
\]

\[
= x(2x + \frac{5}{2}y) + y(\frac{5}{2}x + 2y) + 2x^2 + \frac{5}{2}xy + \frac{5}{2}yx + 2y^2
\]
Note, in general, if \( \mathbf{v}, \mathbf{w} \in \mathbb{R}^n \) and if \( A \) is an \( n \times n \) matrix then
\[
\begin{align*}
\mathbf{v}^T A \mathbf{w} & = \sum_{i=1}^{n} v_i \left( \text{entry}_i(A \mathbf{w}) \right) \\
& = \sum_{i=1}^{n} v_i \left( \sum_{j=1}^{n} a_{ij} w_j \right) \\
& = \sum_{i,j=1}^{n} a_{ij} v_i w_j.
\end{align*}
\]
is a \( 1 \times 1 \) matrix, i.e. a scalar. And its value is
\[
\mathbf{v}^T A \mathbf{w} = \sum_{i=1}^{n} v_i \left( \text{entry}_i(A \mathbf{w}) \right) = \sum_{i,j=1}^{n} a_{ij} v_i w_j.
\]

So given a quadratic expression ("quadratic form") in any number of variables \( (x_1, x_2, \ldots, x_n) \) one can rewrite the quadratic form as
\[
x^T A x
\]
and one can choose to make \( A \) a symmetric matrix, as we did in our specific example, by splitting cross terms symmetrically.

So given a quadratic expression ("quadratic form") in any number of variables \( (x_1, x_2, \ldots, x_n) \) one can rewrite the quadratic form as
\[
x^T A x
\]
and one can choose to make \( A \) a symmetric matrix, as we did in our specific example, by splitting cross terms symmetrically.
Exercise 1a  Find the eigenvalues and eigenvectors for the matrix we're using to express our quadratic expression.

\[
2x^2 + 2y^2 + 5xy = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 5/2 \\ 5/2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.
\]

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

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

Check:

\[
\begin{bmatrix} 2 & 5/2 \\ 5/2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad A\vec{v} = \lambda\vec{v}
\]

\[
\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}, \quad \begin{bmatrix} 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0.
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

$$A \mathbf{v} = \lambda_1 \mathbf{v} \quad A \mathbf{w} = \lambda_2 \mathbf{w}$$

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

$$\mathbf{w} \cdot A \mathbf{v} = A \mathbf{w} \cdot \mathbf{v}.$$  

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

One way to see this is by noting

$$\begin{pmatrix} \mathbf{w} \cdot A \mathbf{v} \\ \mathbf{A}^T \mathbf{v} \end{pmatrix} = \begin{pmatrix} \mathbf{w}^T A \mathbf{v} \\ \mathbf{v}^T A \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{v}^T A \mathbf{w} \\ \mathbf{v}^T A \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{v} \cdot A \mathbf{w} \\ \mathbf{A}^T \mathbf{v} \end{pmatrix}.$$  

But

$$\begin{pmatrix} \mathbf{w} \cdot A \mathbf{v} \\ \mathbf{A}^T \mathbf{v} \end{pmatrix} = \begin{pmatrix} \mathbf{w} \cdot \lambda_1 \mathbf{v} \\ \lambda_1 \mathbf{v} \cdot \mathbf{w} \end{pmatrix} = \begin{pmatrix} \lambda_1 \mathbf{w} \cdot \mathbf{v} \\ \lambda_2 \mathbf{w} \cdot \mathbf{v} \end{pmatrix}.$$  

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

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

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

So

$$0 = \mathbf{w} \cdot A \mathbf{v} = \mathbf{v} \cdot A \mathbf{w} \Rightarrow \mathbf{v} \perp A \mathbf{w} \Rightarrow A \mathbf{w} \in \text{span}\{\mathbf{w}\}$$

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

Theorem: Spectral theorem for $2 \times 2$ symmetric matrices

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} \Rightarrow A - \lambda I = \begin{pmatrix} a-\lambda & c \\ c & b-\lambda \end{pmatrix}$$

$$|A - \lambda I| = (\lambda - a)(\lambda - b) - c^2$$

$$= \lambda^2 - (a+b)\lambda + ab - c^2$$

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

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

\[ \begin{aligned}
(\mathbf{a} - \mathbf{b})^2 + 4 \mathbf{c}^2 &= 0 \\
\Rightarrow \quad \mathbf{a} &= \mathbf{b}, \quad \mathbf{c} = 0
\end{aligned} \]

\[ \mathbf{A} = \begin{bmatrix} 2 & 5 \\ 0 & 2 \end{bmatrix} \quad \text{and for} \quad \mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

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).

\[ \mathbf{A} = \mathbf{P} \mathbf{D} \mathbf{P}^{-1} \]

Note

\[ \mathbf{P} = \mathbf{P}_E \leftarrow \mathbf{B} \]

where as always,

\[ \mathbf{E} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

For \( \mathbf{y} \in \mathbb{R}^2 \) write \( \mathbf{y} = \begin{bmatrix} x \\ y \end{bmatrix} \) in standard coordinates and \( [\mathbf{y}]_B = \begin{bmatrix} x' \\ y' \end{bmatrix} \). (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

\[ \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ \sqrt{2} & \sqrt{2} \\ 1 & -1 \\ \sqrt{2} & \sqrt{2} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \]
Do algebra!

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

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

\[
= \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} \frac{9}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \quad \text{(because } P^TAP = D)\]

\[
= \frac{9}{2} (x')^2 - \frac{1}{2} (y')^2.
\]
So the original curve with equation
\[ 2x^2 + 2y^2 + 5xy = 1 \]
in the standard coordinate system has equation
\[ \frac{9}{2} (x')^2 - \frac{1}{2} (y')^2 = 1 \]
with respect to the rotated coordinate system!

Answer to 1a) This curve is a hyperbola! In the rotated coordinate system its equation is
\[ \frac{(x')^2}{(\sqrt{2})^2} - \frac{(y')^2}{(\sqrt{2})^2} = 1. \]

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:
Maple verification: To be continued ....

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

![2D plot](image1)

```maple
> plot3d(2*x^2 + 2*y^2 + 5*x*y, x = -3..3, y = -3..3);
```

![3D plot](image2)
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 by Prof. Tom Alberts (Fri)
- Prep notes

**Warm-up Exercise:**

Compute
\[
\begin{bmatrix}
1 & -2 \\
2 & 0 \\
0 & 2 \\
2 & 1 \\
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 3 \\
0 & 2 & 1 \\
\end{bmatrix}
= \begin{bmatrix}
1 & -4 & 1 \\
2 & 0 & 6 \\
\end{bmatrix}
\]

Then compute
\[
\sum_{j=1}^{2} [\text{col}_j(A)][\text{row}_j(B)]
\]

\[
= \begin{bmatrix} 1 \\ 2 \end{bmatrix}[1, 0, 3] + \begin{bmatrix} -2 \\ 0 \end{bmatrix}[0, 2, 1]
= \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}
\]

"inner 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 $B = \{ \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \}$ 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 

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

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

For the corresponding orthogonal matrix

$$P = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix}$$

$$D = P^T A P$$

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

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

where $\mathbf{y} = [\mathbf{x}]_B$ and $P = P \leftrightarrow B$. Thus

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} = \mathbf{y}^T P^T A P \mathbf{y} = \mathbf{y}^T D \mathbf{y} = \sum_{i=1}^{n} \lambda_i y_i^2.$$

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 \( Q(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j = x^T A x \) (for \( A \) a symmetric matrix) is called **positive definite** if

\[
Q(x) > 0 \quad \text{for all } x \neq 0.
\]

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:

\[
A = \begin{bmatrix}
a_1 & a_2 & \ldots & a_p \\
\vdots & & & \vdots \\
| & | & \ldots & | \\
| & | & \ldots & | \\
| & | & \ldots & | \\
\end{bmatrix}, \\
B = \begin{bmatrix}
\begin{array}{c}
\vdots \\
- & - & b_1 & - & - \\
- & - & b_2 & - & - \\
\vdots \\
- & - & b_p & - & - \\
\end{array}
\end{bmatrix}.
\]

Then

\[
A B = \sum_{j=1}^{p} a_j b_j.
\]
We can illustrate the general proof by considering the example in which $A$ and $B$ are each $3 \times 3$: Look column by column in the output of each expression to verify the identity, using the linear combination form of matrix times vector, for $A B$:

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

Actual proof with formula:

- **Dot Product**:
  
  $\text{entry}_{jk \in \mathbb{R}} \begin{bmatrix} a_{j1} & \cdots & a_{jn} \end{bmatrix} \cdot \begin{bmatrix} b_{1j} \\ \cdots \\ b_{nj} \end{bmatrix} = \sum_{j=1}^{n} a_{nj} b_{kj}$

- **Outer Product**:
  
  $\sum_{j=1}^{n} \begin{bmatrix} \tilde{a}_{j1} & \cdots & \tilde{a}_{jn} \end{bmatrix} \cdot \begin{bmatrix} -\tilde{b}_{1j} \\ \cdots \\ -\tilde{b}_{nj} \end{bmatrix} = \sum_{j=1}^{n} a_{nj} b_{kj}$

Same!
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

\[
\{ \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \}
\]

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

\[
P = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]
\]

with

\[
A P = P D
\]

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

Then

\[
A = P D P^T
\]

\[
= \begin{bmatrix}
\mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_n \\
\mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_n \\
\mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_n \\
\end{bmatrix}
\begin{bmatrix}
\lambda_1 & 0 & \ldots & 0 \\
0 & \lambda_2 & \ldots & 0 \\
0 & 0 & \ldots & \lambda_n \\
\end{bmatrix}
\begin{bmatrix}
\mathbf{u}_1^T \\
\mathbf{u}_2^T \\
\vdots \\
\mathbf{u}_n^T \\
\end{bmatrix}
\]

"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:

\[ x = (x \cdot u_1)u_1 + (x \cdot u_2)u_2 + \cdots (x \cdot u_n)u_n \]

\[ \Rightarrow Ax = (x \cdot u_1)\lambda_1 u_1 + (x \cdot u_2)\lambda_2 u_2 + (x \cdot u_n)\lambda_n u_n \]

\[ = \lambda_1 u_1^T x + \lambda_2 u_2^T x + \cdots + \lambda_n u_n^T x \]

\[ A x = [\lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \cdots + \lambda_n u_n u_n^T] x. \]

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

\[ A = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \cdots + \lambda_n u_n u_n^T. \]
Testing spectral decomposition in a small example:

\[
A = \begin{bmatrix}
2 & 5 \\
5 & 2 \\
\end{bmatrix}
\]

\[
E = \text{span} \begin{bmatrix}
1 \\
1 \\
\end{bmatrix} \quad \lambda = \frac{9}{2} \\
E = \text{span} \begin{bmatrix}
1 \\
-1 \\
\end{bmatrix} \quad \lambda = -\frac{1}{2}
\]

\[
u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \nu_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}
\]

\[
\lambda_1 \nu_1 \nu_1^T + \lambda_2 \nu_2 \nu_2^T = \frac{9}{2} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \frac{1}{2} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}
\]

\[
= \frac{9}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \frac{1}{4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix}
\]
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) \( Q^{-1} = Q^T \).

\[
\begin{align*}
\text{entry } &_{ij} Q^T Q = [\text{Row } i \cdot Q^T] [\text{Col } j Q] \\
\implies & Q^T Q = I \\
\implies & QQ^T = I \quad \text{(prev. chap.)}
\end{align*}
\]

i.e. \( Q^{-1} = Q^T \)

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

\[
Q Q^T = I
\]

\[
\begin{align*}
\text{i}^\prime \text{j} \text{ entry: } & \frac{\text{row } i Q \cdot \text{col } j Q^T}{\text{row } i Q^T} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}
\end{align*}
\]

g) the transformation \( T : \mathbb{R}^n \to \mathbb{R}^n \) given by

\[
T(x) = Q x
\]

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 \( \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \),

\[
T(\mathbf{x}) \cdot T(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}
\]

\[
||T(\mathbf{x})|| = ||\mathbf{x}||.
\]

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

\[ x_1^2 + x_2^2 - 2x_3^2 - 2x_1x_2 - 4x_1x_3 - 4x_2x_3 = x^T A x. \]

Recall that

\[ x^T A x = \sum_{i,j=1}^{n} a_{ij} x_i x_j. \]

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

If we found the matrix correctly technology tells us that

\[ E_{\lambda=2} = \text{span} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad E_{\lambda=-2} = \text{span} \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}, \quad E_{\lambda=4} = \text{span} \begin{bmatrix} -1 \\ -1 \end{bmatrix}. \]

(positively oriented in this order)
\[ x_1^2 + x_2^2 - 2x_3^2 - 2x_1x_2 - 4x_1x_3 - 4x_2x_3 = 8 \]

\[ \mathbf{x}^T A \mathbf{x} = 8 \]

For

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

\[ P^T A P = D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 4 \end{bmatrix} \]

\[
\mathbf{x}^T A \mathbf{x} = 8 \\
\mathbf{y}^T P^T A P \mathbf{y} = 8 \\
\mathbf{y}^T D \mathbf{y} = 8
\]

\[-2y_1^2 + 2y_2^2 + 4y_3^2 = 8. \quad 2y_2^2 + y_3^2 = 8 + 2y_1^2\]

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):

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

eigenvalues([1, 1, -2], [1, 1, -2], [2, 2, 2])

$\lambda_1 = 4$

$\lambda_2 = -2$

$\lambda_3 = 2$

Corresponding eigenvectors:

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

$v_2 = (1, 1, 1)$

$v_3 = (-1, 1, 0)$
from Wikipedia, "quadric surfaces". There is also a Wikipedia page on conic sections.

<table>
<thead>
<tr>
<th>Non-degenerate real quadric surfaces</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ellipsoid</strong></td>
</tr>
<tr>
<td><strong>Elliptic paraboloid</strong></td>
</tr>
<tr>
<td><strong>Hyperbolic paraboloid</strong></td>
</tr>
<tr>
<td><strong>Elliptic hyperboloid of one sheet</strong></td>
</tr>
<tr>
<td><strong>Elliptic hyperboloid of two sheets</strong></td>
</tr>
</tbody>
</table>
Spectral Theorem

Let \( A \in \mathbb{R}^{n \times n} \) be a real, symmetric matrix.

Then \( \exists \) an orthonormal \( \mathbb{R}^n \) basis made of
eigenvectors of \( A \), \( \mathcal{B} = \{ \vec{v}_1, \vec{v}_2, \ldots \vec{v}_n \} \)

Thus for \( S = [\vec{v}_1 | \vec{v}_2 | \ldots | \vec{v}_n] \),

\[
S^T A S = \begin{bmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n
\end{bmatrix}
\]

is diagonal.

Proof

1. On Monday we showed that if \( \lambda_1 \neq \lambda_2 \) are real eigenvalues of \( A \),
   with eigenvectors \( \vec{v}_1, \vec{v}_2 \), then

\[
A \vec{v}_1 = \lambda_1 \vec{v}_1,
A \vec{v}_2 = \lambda_2 \vec{v}_2
\]

and

\[
\text{prob. } v_1^T A v_1 = v_1^T (A v_1) = \lambda_1 \vec{v}_1^T \vec{v}_1
\]

Thus, \( \vec{v}_1, \vec{v}_2 \) are orthonormal.

We also showed that
for \( A \in \mathbb{R}^{2 \times 2} \) 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 \( \lambda = \alpha + \beta i \) be any root of \( p(\lambda) \), and let \( \vec{u} + i \vec{v} \) be a corresponding
   non-zero eigenvector.)

\[
A (\vec{u} + i \vec{v}) = (\alpha + \beta i)(\vec{u} + i \vec{v})
\]

Take conjugates:

\[
A (\vec{u} - i \vec{v}) = (\alpha - \beta i)(\vec{u} - i \vec{v}).
\]

Now consider

\[
(\vec{u} - i \vec{v})^T A (\vec{u} + i \vec{v}) = (\vec{u} - i \vec{v})^T (\alpha + \beta i)(\vec{u} + i \vec{v}) = (\alpha + \beta i)[(\vec{u} - i \vec{v})^T (\vec{u} + i \vec{v})]
\]

\[
= (\alpha + \beta i)[(\vec{u} - i \vec{v}) \cdot (\vec{u} + i \vec{v})]
\]

\[
= (\alpha + \beta i)(\vec{u}^T \vec{u} + \vec{v}^T \vec{v})
\]

\[
= (\alpha + \beta i)(1 \| \vec{u} \|^2 + 1 \| \vec{v} \|^2)
\]

Thus \( p(\lambda) \) factors completely over \( \mathbb{R} \).

- If it has \( n \) distinct roots, then alg mult = geom mult = 1, all
  eigenvectors are \perp by 1, and normalizing to get orthonormal
  eigenbasis

- Otherwise it's a little harder: (in practice, if \( \lambda_i \) has alg mult \( \geq 1 \)
  just Gram-Schmidt its eigenbasis.)
General proof, by induction:
Spectral Theorem true for \( n=1 \) (1x1 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 eigenvectors columns).

Now let \( A_{nn} \) symmetric.

Let \( \lambda_1 \) be any root of \( \det(A) \). \( \lambda_1 \) is real by (2).

Let \( \vec{u}_1 \) be a unit eigenvector
\[ A\vec{u}_1 = \lambda_1 \vec{u}_1, \quad \|\vec{u}_1\|_2 = 1. \]

Orthonormal basis:
Complete to an orthonormal basis
\[ \begin{align*}
B_0 &= \{ \vec{u}_1, \vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n \} \quad \text{of } \mathbb{R}^n \\
S_0 &= \begin{bmatrix} \vec{u}_1 & \vec{v}_2 & \ldots & \vec{v}_n \end{bmatrix} \\
S_0^T S_0 &= I
\end{align*} \]

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

1st column is \[ \begin{bmatrix} \lambda_1 \\ 0 \end{bmatrix} \]

so
\[ S_0^T A S_0 = \begin{bmatrix} \lambda_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} B \end{bmatrix} \]

Thus
\[ S_0^T A S_0 = \begin{bmatrix} \lambda_1 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ B \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ S_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} S_1 \end{bmatrix} \]

so
\[ S_0^T A S_0 = D \]

\[ S = S_0 \begin{bmatrix} \lambda_1 \\ \vdots \\ 0 \end{bmatrix} \text{ orthog (product of orthog)} \]