1. Let $A = \begin{bmatrix} 7.5 & -7 & 0 \\ 3.5 & -3 & 0 \\ 7 & -7 & 0.5 \end{bmatrix}$.
   (a) Find all eigenvalues of A and their multiplicity.
   (b) Find a basis for each eigenspace of A.
   (c) Is A diagonalizable? If so, diagonalize A. If not, explain why not.
   (d) Compute $\lim_{n\to\infty} A^n$ or explain why it doesn't exist.
   (e) X is a matrix with eigenvalues $2, 3, 3$. Does $\lim_{n\to\infty} X^n$ exist? Does it depend on whether X is diagonalizable?

2. (a) Prove that for a matrix B if char(B) has distinct roots (that is, each root has multiplicity 1), then B is diagonlaizable.
   (b) Prove that if C is an invertible matrix, then $(C^T)^{-1} = (C^{-1})^T$.
   (c) Prove that if B is diagonalizable, then $\det(B) = \det(B^T)$. This is true for all square matrices, but you only need to prove it for diagonalizable ones.

3. McDonalds recently negotiated a large purchasing deal for fish, chicken, and beef. They have agreed to purchase 40 million tons of fish, 40 million tons of chicken, and 100 million tons of beef. As such, they want to create an advertising campaign to ensure that consumers eat the correct portion of each meat product.

   After paying to have a commercial produced, McDonalds collects the following data: After watching the commercial once, a person who initially wanted fish now has a 10% chance of buying a fish product, a 60% chance of buying a beef product, and a 30% chance of buying a chicken product; after watching, a person who initially wanted to buy a beef product has a 20% chance of buying a fish product, a 60% chance of buying a beef product, and a 20% chance of buying a chicken product; after watching, a person who initially wanted to by a chicken product has a 40% chance of buying a fish product, a 50% chance of buying a beef product, and a 10% chance of buying a chicken product.

   (a) If the vector $\vec{e}_1$ represents a person who wants to buy a fish product, $\vec{e}_2$ represents a person who wants to buy a beef product, and $\vec{e}_3$ represents a person who wants to buy a chicken product, find a matrix M such that $M\vec{e}_i$ gives the probability of buying fish, beef, or chicken after watching the commercial once.
   (b) Compute the eigenvalues and eigenvectors of M.
   (c) Assume that a fish product takes 50 grams of fish, a beef product takes 50 grams of beef, and a chicken product takes 50 grams of chicken. Further, assume that each time a person watches the commercial it has the same impact (i.e., watching the commercial twice means the likelihood of buying a particular product is given by M^2). If McDonalds ensures that the average customer sees the commercial 3000 times, what are the relative proportions of fish, beef, and chicken McDonalds expects to sell?
   (d) Should McDonalds run the ad? Does the initial population's preferences for fish, beef, or chicken matter? *Explain your reasoning.*

4. We are often interested in bounding error. As we are all familiar with, if we make a measurement and find it to be 7 ± 0.1, if we amplify our result by a factor of two, the error will double and we should report 14 ± 0.2,

   This one-dimensional example is straightforward. However, things are more complicated when we deal with multiple dimensions. In higher dimensions, your error takes the form of an error vector, and your transformation may be a matrix.

   Let $\vec{m}_a$ be the actual value of a quantity. We say that $\vec{m}$ is a measurement of $\vec{m}_a$ within tolerance ϵ if $\|\vec{m}_a - \vec{m}\| \leq \epsilon$. Another way of saying that is

$$\vec{m}_a = \vec{m} + \vec{v} \qquad \text{where } \|\vec{v}\| < \epsilon.$$

Here, $\vec{v}$ is our *error vector*. If we apply linear transformation T, we have that $T\vec{m}_a = T\vec{m} + T\vec{v}$ and so our error goes from $\vec{v}$ to $T\vec{v}$. The question becomes, can we find a number k so that if $\|\vec{v}\| < \epsilon$, then $\|T\vec{v}\| < k\epsilon$. The smallest such number is called the *norm* of the matrix T, and we write $\|T\| = k$.

(a) Explain why $\|T\| = \max\{\|T\vec{u}\| : \vec{u} \text{ is a unit vector}\}$.

(b) Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$. Use Matlab/octave to create a 2×360 matrix C whose column vectors are uniformly distributed points on the unit circle (Hint: a command like `[cos(1:10)]` will produce the vector $\begin{bmatrix}\cos 1 & \cos 2 & \cdots & \cos 10\end{bmatrix}$).

Graph C and AC on the same plot. Notice that C should be a circle (you may need to rescale your axes with `xlim([-4 4]); ylim([-4 4])` to get C to look like a circle) and AC should be an ellipse. Explain how the major and minor axes of the ellipse AC relate to $\|A\|$.

(c) Numerically estimate $\|A\|$. If $\vec{v}$ is an error vector with $\|\vec{v}\| = 0.01$, give an upper bound on $\|A^{30}\vec{v}\|$.

(d) Suppose B is a diagonalizable 2×2 matrix with eigenvalues λ_1 and λ_2 satisfying $|\lambda_1|, |\lambda_2| \leq 1$. Come up with a conjecture for what an upper bound for $\|B\|$ might be. Then, numerically experiment. Try to explain your findings. (Hint: if you want a "random" matrix with particular eigenvalues, you might consider something like `r=rand(2); B=r*D*r^(-1)` for a well-chosen `D`). Make sure to be good scientists and seek for evidence to *disprove* your hypothesis.

5. Let's blow the lid off. It's time to tiptoe into the realm of infinite dimensions. Using the idea of *sampling*, we can easily turn a function into a finite dimensional vector. Let $f : [0,1] \to \mathbb{R}$ be a continuous function, and let $s_n : \{\text{continuous functions}\} \to \mathbb{R}^n$ be the function that creates a vector out of a function by sampling that function at n equally spaced points. That is,

$$s_n(f) = \begin{bmatrix} f(0) \\ f(\frac{1}{n}) \\ f(\frac{2}{n}) \\ \vdots \\ f(\frac{n-1}{n}) \end{bmatrix}.$$

(a) Using s_n, we can now carry over many of our familiar operations from $\mathbb{R}^n$ to the space of functions. Let $\mathcal{P}_n$ be the set of polynomials of degree at most n. We can attempt to induce an *inner product* on $\mathcal{P}_n$ as follows,

$$\langle \cdot, \cdot \rangle_n : \{\text{continuous functions}\} \to \mathbb{R} \qquad \text{defined by} \qquad \langle p, q \rangle_n = s_n(p) \cdot s_n(q).$$

Show that $\langle \cdot, \cdot \rangle_n$ induces an inner product on $\mathcal{P}_{n-1}$ but not on $\mathcal{P}_n$ (Hint: a degree n polynomial is uniquely determined by $n+1$ points).

(b) Recall that with an inner product $\langle \cdot, \cdot \rangle_n$, we define an induced norm via $\|f\|_{\langle \cdot, \cdot \rangle_n} = \sqrt{\langle f, f \rangle_n}$. Our goal is to find an induced norm that works for all polynomials simultaneously. One idea might be to try to take the limit of $\langle \cdot, \cdot \rangle_n$, however, this runs into problems, because $\lim_{n\to\infty} \langle 1, 1 \rangle_n = \infty$. Let's fix this by *normalizing*.

Define

$$\langle a, b \rangle_{\lim} = \lim_{n\to\infty} \frac{1}{n} \langle a, b \rangle_n$$

and let $\| \cdot \|_{\lim}$ be the induced norm. Compute $\|f\|_{\lim}$ for $f(x) = 1$, $f(x) = x$, and $f(x) = x^2$. (Hint: you already know how to do this! Think about Riemann sums!)

(c) It turns out $\langle \cdot, \cdot \rangle_{\text{lim}}$ is an inner product on all continuous functions on $[0, 1]$. Thus, we can use it to do projections and find angles just like we would in $\mathbb{R}^n$.

Let $\mathcal{B} = \{1, x, x^2, x^3\}$ be a basis for $\mathcal{P}_3$. Apply the Gram-Schmidt process to $\mathcal{B}$ (read your textbook or Wikipedia if we haven't done it in class yet) to find an orthonormal basis for $\mathcal{P}_3$.

(d) Use your basis from the previous part to find $\text{proj}_{\mathcal{P}_3} \frac{1}{1+10x^2}$. Plot $\frac{1}{1+10x^2}$ and your projection on the same axes. Is it a good approximation? How does it compare with the degree-3 Taylor approximation? (You may use a computer to estimate the coefficients).

(e) THE FOURIER BASIS. The Fourier basis is one of the most important bases for the space of functions. Let $\mathcal{B}_n = \{1, \sin 2\pi t, \cos 2\pi t, \sin 4\pi t, \cos 4\pi t, \sin 6\pi t, \cos 6\pi t, \ldots, \sin 2n\pi t, \cos 2n\pi t\}$. Apply the Gram-Schmidt process to $\mathcal{B}_n$ to obtain the orthonormal Fourier basis $\mathcal{F}_n$.

(f) Let D be differentiation. Is the Fourier basis an eigenbasis for D? How about for the second derivative D^2? Write out $[D]_{\mathcal{F}_7}$ and $[D^2]_{\mathcal{F}_7}$.

(g) Using complex numbers, the Fourier basis can be written in terms of exponentials as $\mathcal{F}_n = \{e^{2k\pi it} : 0 \le k \le n\}$. Show that each element in the complex Fourier basis is an eigenvector of differentiation and give its corresponding eigenvalue. Then marvel at how projecting a function onto the complex Fourier basis looks a lot like taking the Laplace transform (it isn't the same, it's of course called the *Fourier* transform, but it's darn close).

6. AN ULTRA-VIOLET CATASTROPHE. We can use the Fourier basis to quantify the Heisenberg uncertainty principle! And, because I'm sure you're feeling computational, we'll do so numerically with Matlab/octave.

(a) Create a 129×129 matrix F whose columns are the basis elements in $\mathcal{F}_{64}$ evaluated at the points $\mathbf{x} = (0/129, 1/129, 2/129, \ldots, 128/129)$. If you plot the first three columns of F with the command `plot(F(:,1:3))` you should see the graphs of $y = 1$, $y = \sqrt{2}\sin 2\pi t$ and $y = \sqrt{2}\cos 2\pi t$. If you'd like to see F in all its glory, you can use the command `imshow(F)`.

Compute $F^T F$ and explain your result.

(b) Since F changes basis from the Fourier basis to the standard basis, F^{-1} is almost F^T and goes from the standard basis to the Fourier basis. (In your computations, use `F'` instead of `F^(-1)` for numerical stability reasons.) Create four column vectors

$$\vec{v}_0 = 3\sin(4\pi\mathbf{x}) \qquad \vec{v}_{0.2} = 3\sin(4\pi(\mathbf{x} - 0.2))$$

$$\vec{v}_{0.5} = 3\sin(4\pi(\mathbf{x} - 0.5)) \qquad \vec{v}_{0.7} = 3\sin(4\pi(\mathbf{x} - 0.7))$$

where $\mathbf{x}$ ranges over the values specified in part 1. Use F^{-1} to write these vectors in the Fourier basis. What frequencies in the Fourier basis have non-zero coefficients? Does this make sense?

(c) Recall that $a \sin t + b \cos t$ is a shifted sine wave with amplitude $\sqrt{a^2 + b^2}$. Using just the coefficients from the Fourier basis found in part 2, verify that $\vec{v}_i$ is a shifted sine wave with amplitude 3 for each $i \in \{0, .2, .5, .7\}$.

(d) According to Einstein, the energy contained in a wave of frequency k is proportional to ka where a is the amplitude of the wave. Since we're doing math, we will assume the energy of a wave of frequency k and amplitude a is *equal* to ka. Suppose a particle is described in the Fourier basis by

$$[\vec{p}]_{\mathcal{F}_{64}} = \begin{bmatrix} 1 & 1 & 1.5 & 2 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Compute the energy of $\vec{p}$.

(e) We are now ready to see the Heisenberg uncertainty principle in action. Using the command `s=4; xs=(0:128)/129; s*normpdf(s*(xs-0.5))` we can create a bell curve centered at 0.5 with standard deviation $1/s$. As `s` increases, the bell curve will become more and more concentrated. Using your knowledge of the Fourier basis, compute the wave-energy contained in bell curves with standard deviations $1/4, 1/40, 1/400, 1/4000$. If a particle has a finite amount of energy, is there a limit to how concentrated in one spot it can be?

(f) The command `a=.3; b=.4; xs=(0:128)/129; 1/(b-a)*(xs > a).*(xs < b)` will create the distribution of a particle that is contained in the range $[a, b]$ with 100% certainty. Compute the energy of such a particle. Does changing the width of the interval affect the amount of energy in the particle? How does the energy contained in this particle compare to that in a bell-curve-distributed particle?

(g) Let $\vec{p}$ be a particle strictly contained in $[a, b]$. We can attempt to see what it would look like if this particle existed in nature by forcing its energy to be low. That is, if we compute $F^{-1}\vec{p}$, we can then use the use a truncation function to set the coefficients of all high-frequency sine and cosine waves to zero. Let $t : \{\text{ vectors in the basis } \mathcal{F}_{64}\} \to \{\text{ vectors in the basis } \mathcal{F}_{64}\}$ defined so that $t(\vec{v})$ sets the coefficients of every sine and cosine wave with frequency higher than 25 to zero and leaves the rest of the coefficients alone.

Now, we might say that $F(t(F^{-1}\vec{p}))$ is what $\vec{p}$ would "look like" if we forced it to have finite energy. For $[a, b] = [.3, .4]$ and $[a, b] = [.4, .6]$ graph a picture of what $\vec{p}$ would "look like."