Topic of this homework: Pythagorean triples, Pell’s equation, Fibonacci sequence

Deliverable: Answers to problems

1 Pythagorean triples

Euclid’s formula for the Pythagorean triples \( a, b, c \) is:

\[
a = p^2 - q^2, \quad b = 2pq, \quad c = p^2 + q^2.
\]

1. What condition(s) must hold for \( p \) and \( q \) such that \( a, b, \) and \( c \) are always positive and nonzero?

2. Solve for \( p \) and \( q \) in terms of \( a, b \) and \( c \).

3. The ancient Babylonians (c2000BEC) cryptically recorded \((a,c)\) pairs of numbers on a clay tablet, archeologically denoted Plimpton-322.

To Do: Find \( p \) and \( q \) for the first five pairs of \( a \) and \( c \) from the tablet entries:

<table>
<thead>
<tr>
<th>a</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>119</td>
<td>169</td>
</tr>
<tr>
<td>3367</td>
<td>4825</td>
</tr>
<tr>
<td>4601</td>
<td>6649</td>
</tr>
<tr>
<td>12709</td>
<td>18541</td>
</tr>
<tr>
<td>65</td>
<td>97</td>
</tr>
</tbody>
</table>

4. Based on Euclid’s formula, show that \( c > (a,b) \).

5. What happens when \( c = a \)?

6. Is \( b + c \) a perfect square? Discuss.

2 Pell’s equation

Pell’s equation is one of the most historic (i.e., important) equations of Greek number theory, because it was used to show that \( \sqrt{2} \in \mathbb{I} \). We seek integer solutions

\[
x^2 - Ny^2 = 1.
\]

As shown in Lec 8 of the lecture notes, the solutions \( x_n, y_n \) for the case of \( N = 2 \) are given by the 2x2 matrix recursion

\[
\begin{bmatrix}
x_{n+1} \\
y_{n+1}
\end{bmatrix} = 1J \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix}
\]

with \([x_0, y_0]^T = [1, 0]^T\) and \( 1J = \sqrt{-1} = e^{j\pi/2} \). It follows that the general solution to Pell’s equation for \( N = 2 \) is

\[
\begin{bmatrix} x_n \\ y_n \end{bmatrix} = (e^{j\pi/2})^n \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}
\]
To calculate solutions to Pell’s equation using the matrix equation above, we must calculate

\[ A^n = e^{\pi n/2} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} = e^{\pi n/2} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \]

which becomes tedious for \( n > 2 \), since it requires \( n \times 2 \times 2 \) matrix multiplications.

**Diagonalization of a matrix ("eigenvalue/eigenvector decomposition"):** As derived in Appendix C of the lecture notes, the most efficient way to compute \( A^n \) is to diagonalize the matrix \( A \), by finding its eigenvalues and eigenvectors.

The eigenvalues \( \lambda_k \) and eigenvectors \( \vec{e}_k \) of a square matrix \( A \) are related by

\[ A\vec{e}_k = \lambda_k \vec{e}_k, \quad (1) \]

such that multiplying an eigenvector \( \vec{e}_k \) of \( A \) by the matrix \( A \) is the same as multiplying by a scalar, \( \lambda_k \in \mathbb{C} \) (the corresponding eigenvalue). The complete eigenvalue problem may be written as

\[ AE = \Lambda. \]

If \( A \) is a \( 2 \times 2 \) matrix,\(^1\) the matrices \( E \) and \( \Lambda \) (of eigenvectors and eigenvalues, respectively) are

\[ E = \begin{bmatrix} \vec{e}_1 & \vec{e}_2 \end{bmatrix} \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \]

Thus, the matrix equation \( AE = \begin{bmatrix} A\vec{e}_1 & A\vec{e}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \vec{e}_1 & \lambda_2 \vec{e}_2 \end{bmatrix} \equiv \Lambda \) contains Eq. 1 for each eigenvalue-eigenvector pair.

The diagonalization of the matrix \( A \) refers to the fact that the matrix of eigenvalues, \( \Lambda \), has non-zero elements only on the diagonal. The key result is found by post-multiplication of the eigenvalue matrix by \( E^{-1} \), giving

\[ AEE^{-1} = A = \Lambda E^{-1}. \quad (2) \]

If we now take powers of \( A \), the \( n^{th} \) power of \( A \) is

\[ A^n = (\Lambda E^{-1})^n = \Lambda E^{-1} \Lambda E^{-1} \cdots \Lambda E^{-1} = \Lambda E^{-1}. \quad (3) \]

This is a very powerful result, because the \( n^{th} \) power of a diagonal matrix is extremely easy to calculate:

\[ \Lambda^n = \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} \]

Thus, from Eq. 3 we can calculate \( A^n \) using only two matrix multiplications

\[ A^n = \Lambda^n E^{-1}. \]

\(^1\)These concepts may be easily extended to higher dimensions.
Finding the eigenvalues: The eigenvalues $\lambda_k$ are determined by Eq. 1, by factoring out $\vec{e}_k$

$$A\vec{e}_k = \lambda_k \vec{e}_k$$

$$(A - \lambda_k I)\vec{e}_k = 0.$$ 

Matrix $I = [1, 0; 0, 1]^T$ is the identity matrix, having the dimensions of $A$, with elements $\delta_{ij}$ (i.e., diagonal elements $\delta_{11,22} = 1$ and off-diagonal elements $\delta_{12,21} = 0$).

The vector $\vec{e}_k$ is not zero, yet when operated on by $A - \lambda_k I$, the result must be zero. The only way this can happen is if the operator is degenerate (has no solution), that is

$$\det(A - \lambda I) = \det \begin{bmatrix} (a_{11} - \lambda) & a_{12} \\ a_{21} & (a_{22} - \lambda) \end{bmatrix} = 0. \tag{4}$$

This means that the two equations have the same slope (the equation is degenerate).

This determinant equation results in a second degree polynomial in $\lambda$

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0,$$

the roots of which are the eigenvalues of the matrix $A$.

Finding the eigenvectors: An eigenvector $\vec{e}_k$ can be found for each eigenvalue $\lambda_k$ from Eq. 1,

$$(A - \lambda_k I)\vec{e}_k = 0.$$ 

The left side of the above equation becomes a column vector, where each element is an equation in the elements of $\vec{e}_k$, set equal to 0 on the right side. These equations are always degenerate, since the determinant is zero. Thus the two equations have the same slope.

Solving for the eigenvectors is often confusing, because they have arbitrary magnitudes, $||\vec{e}_k|| = \sqrt{\vec{e}_k \cdot \vec{e}_k} = \sqrt{e_{k,1}^2 + e_{k,2}^2} = d$. From Eq. 1, you can only determine the relative magnitudes and signs of the elements of $\vec{e}_k$, so you will have to choose a magnitude $d$. It is common practice to normalize each eigenvector to have unit magnitude ($d = 1$).

To do:

**Hint:** Use Matlab’s function $[E, \text{Lambda}] = \text{eig}(A)$ to check your results!

1. Solutions to Pell’s equation were used by the Pythagoreans to explore the value of $\sqrt{2}$. Explain why Pell’s equation is relevant to $\sqrt{2}$.
2. Find the first 3 values of $[x_n, y_n]^T$ by hand and show that they satisfy Pell’s equation for $N=2$.
3. By hand, find the eigenvalues $\lambda_\pm$ of the $2 \times 2$ Pell’s equation matrix

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$$

4. By hand, show that the matrix of eigenvectors, $E$, is

$$E = \begin{bmatrix} \vec{e}_+ & \vec{e}_- \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} -\sqrt{2} & \sqrt{2} \\ 1 & 1 \end{bmatrix}$$

5. Using the eigenvalues and eigenvectors you found for $A$, verify that

$$E^{-1}AE = \Lambda = \begin{bmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{bmatrix}$$

6. Now that you have diagonalized $A$ (Equation 3), use your results for $E$ and $\Lambda$ to solve for the $n = 10$ solution $[x_{10}, y_{10}]^T$ to Pell’s equation with $N = 2$. 


3 The Fibonacci sequence

The Fibonacci sequence is famous in mathematics, and has been observed to play a role in the mathematics of genetics. Let $x_n$ represent the Fibonacci sequence,

$$x_n = x_{n-1} + x_{n-2},$$  \hspace{1cm} (5)

where the current output sample, $x_n$, is equal to the sum of the previous two inputs. This is a ‘discrete time’ recurrence relation. To solve for $x_n$, we require some initial conditions. In this exercise, let us define $x_0 = 1$ and $x_{n<0} = 0$. This leads to the Fibonacci sequence \{1, 1, 2, 3, 5, 8, 13, \ldots\} for $n = 0, 1, 2, 3, \ldots$

Here we seek the general formula for $x_n$. Like the Pell’s equation, Eq. 5 has a recursive, eigen decomposition solution. To find it we must recast $x_n$ as a 2x2 matrix relation, and then proceed as we did for the Pell case.

1. Show that Eq. 5 is equivalent to the 2 $\times$ 2 matrix equation

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = A \begin{bmatrix} x_{n-1} \\ y_{n-1} \end{bmatrix}. \hspace{1cm} A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$  \hspace{1cm} (6)

and that the Fibonacci sequence $x_n$ as described above may be generated by

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = A^n \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \hspace{1cm} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

What is the relationship between $y_n$ and $x_n$?

2. Write a Matlab program to compute $x_n$ using the matrix equation above (you don’t need to turn in your code). Test your code using the first few values of the sequence. Using your program, what is $x_{40}$?

Note: to make your program run faster, consider using the eigen decomposition of $A$, described by Eq. 3 from the Pell’s equation problem.

3. Using the eigen decomposition of the matrix $A$ (and a lot of algebra), it is possible to obtain the general formula for the Fibonacci sequence,

$$x_n = \frac{1}{\sqrt{5}} \left[ \left( \frac{1 + \sqrt{5}}{2} \right)^{n+1} - \left( \frac{1 - \sqrt{5}}{2} \right)^{n+1} \right].$$  \hspace{1cm} (7)

What are the eigenvalues $\lambda_{\pm}$ of the matrix $A$? How is the formula for $x_n$ related to these eigenvalues?

4. Consider Eq. 7 in the limit as $n \to \infty$...

(a) What happens to each of the two terms $[(1 \pm \sqrt{5})/2]^{n+1}$?

(b) What happens to the ratio $x_{n+1}/x_n$?

5. Prove that$^2$

$$\sum_{1}^{N} f_n^2 = f_N f_{N+1}.$$

---

$^2$I found this problem on a worksheet for Math 213 midterm (213practice.pdf).
6. Replace the Fibonacci sequence with

\[ x_n = \frac{x_{n-1} + x_{n-2}}{2}, \]

such that the value \( x_n \) is the average of the previous two values in the sequence.

(a) What matrix \( A \) is used to calculate this sequence?
(b) Modify your computer program to calculate the new sequence \( x_n \). What happens as \( n \to \infty \)?
(c) What are the eigenvalues of your new \( A \)? How do they relate to the behavior of \( x_n \) as \( n \to \infty \)?

\textit{Hint: you can expect the closed-form expression for \( x_n \) to be similar to Eq. 7.}

7. Now consider

\[ x_n = \frac{x_{n-1} + 1.01x_{n-2}}{2}. \]

(a) What matrix \( A \) is used to calculate this sequence?
(b) Modify your computer program to calculate the new sequence \( x_n \). What happens as \( n \to \infty \)?
(c) What are the eigenvalues of your new \( A \)? How do they relate to the behavior of \( x_n \) as \( n \to \infty \)?

\textit{Hint: you can expect the closed-form expression for \( x_n \) to be similar to Eq. 7.}