EECS 16A Spring 2016

Designing Information Devices and Systems I

Elad Alon, Babak Ayazifar

Homework 13

This homework is due May 3, 2016, at Noon.

1. **Homework process and study group**

   Who else did you work with on this homework? List names and student ID's. (In case of hw party, you can also just describe the group.) How did you work on this homework?
   Working in groups of 3-5 will earn credit for your participation grade.

   Solution: I worked on this homework with...

   I first worked by myself for 2 hours, but got stuck on Problem 5 so I went to office hours on...

   Then I went to homework party for a few hours, where I finished the homework.

2. **Mechanical Problems**

   (a) Compute the determinant of $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

   Draw what this matrix does to the unit square in 2D. Compute the area of the resulting shape using standard geometric arguments.

   Solution:

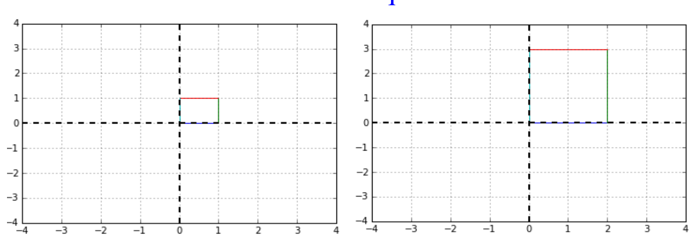
   We can use the form of a 2x2 determinant from lecture:

   $$det(\begin{bmatrix} a & b \\ c & d \end{bmatrix}) = ad - bc$$

   So:

   $$det(\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}) = 2 \cdot 3 - 0 = 6$$

   The matrix transforms the unit square as

   By inspection, we see that we have a rectangle, with area $2 \cdot 3 = 6$.

   For our discussion of determinant, this second approach — the direct computation of the area — is more fundamental. This is what we wanted the determinant to mean.

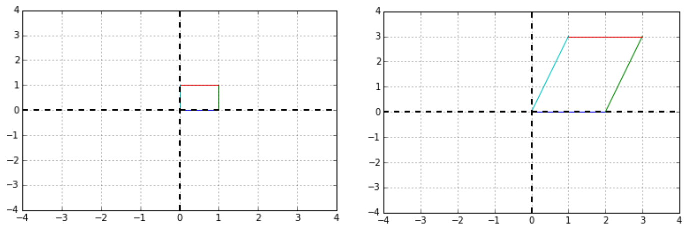
   (b) Compute the determinant of $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$

   Draw what this matrix does to the unit square in 2D. Compute the area of the resulting shape using standard geometric arguments.

Solution:

$$det(\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}) = 2 \cdot 3 - 1 \cdot 0 = 6$$

The matrix transforms the unit square as:



The unit square transforms into the above parallelogram. The area of this parallelogram is its base (2) times its height (3), so the area is 6, matching the determinant. (The area could also be computed by any other geometric argument, for example by breaking the parallelogram into two triangles and a rectangle).

(c) Compute the determinant of $\begin{bmatrix} -4 & 0 & 0 & 0 \\ 0 & 17 & 0 & 0 \\ 0 & 0 & -31 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

Solution:

This is a diagonal matrix. So what it does is to scale each axis by the appropriate term on the diagonal. Consequently, since for us the determinant is defined to be the oriented hypervolume of the unit hypercube as transformed by the matrix, we immediately know that the determinant is 4216 since that is the product of the diagonal entries. The negative signs here represent flips — reflections about that particular axis. These should flip the sign of the oriented hypervolume (this is what makes it oriented). However, the two flips cancel.

Alternatively, we can use Gaussian elimination (which if you remember from lecture, was justified by the more fundamental volume-based definition of the determinant). We use Gaussian elimination on the matrix to reduce it to the identity, keeping track of all changes made:

i. Multiply R1 by $-\frac{1}{4}$. Denote $c_1 = -\frac{1}{4}$ for bookkeeping.

$$A = \begin{bmatrix} -4 & 0 & 0 & 0 \\ 0 & 17 & 0 & 0 \\ 0 & 0 & -31 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 17 & 0 & 0 \\ 0 & 0 & -31 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

ii. Multiply R2 by $\frac{1}{17}$. Denote $c_2 = \frac{1}{17}$ for bookkeeping.

$$A_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 17 & 0 & 0 \\ 0 & 0 & -31 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -31 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

iii. Multiply R3 by $-\frac{1}{31}$. Denote $c_3 = -\frac{1}{31}$ for bookkeeping.

$$A_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -31 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

iv. Multiply R4 by $\frac{1}{2}$. Denote $c_4 = \frac{1}{2}$ for bookkeeping.

$$A_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

v. The determinant of this last matrix is known to be 1. Now we trace our steps, knowing that scaling a row by a constant c_i scales the determinant by c_i. Suppose the initial matrix is A. Then, the following must be true:

$det(I) = c_4 \cdot det(A_4) = c_4 \cdot c_3 \cdot det(A_3) = c_4 \cdot c_3 \cdot c_2 \cdot det(A_2) = c_4 \cdot c_3 \cdot c_2 \cdot c_1 \cdot det(A)$

We know $det(I) = 1$, so we substitute and solve for $det(A)$:

$det(I) = c_4 \cdot c_3 \cdot c_2 \cdot c_1 \cdot det(A) = \frac{1}{2} \cdot \frac{-1}{31} \cdot \frac{1}{17} \cdot \frac{-1}{4} \cdot det(A) = \frac{1}{4216} det(A)$

$det(A) = 4216 \cdot det(I) = 4216 \cdot 1 = 4216$

so the final answer is 4216. Notice that this is exactly the product of the diagonal entries, which is what we saw from the definition must be true for matrices that have zeros everywhere except on the main diagonal.

(d) Find the eigenvalues and the eigenspaces of the following matrix A:

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$$

What is the inner product between the two eigenvectors?

Solution: To find the eigenvalues, we need to solve $det(A - \lambda I) = 0$

$$det(\begin{bmatrix} 2-\lambda & 3 \\ 3 & 4-\lambda \end{bmatrix}) = (2-\lambda)\cdot(4-\lambda) - 3\cdot 3 = 8 - 6\lambda + \lambda^2 - 9 = 0$$

$$\lambda^2 - 6\lambda - 1 = 0 \implies \lambda = \frac{6 \pm \sqrt{6^2 - 4\cdot 1\cdot(-1)}}{2} = \frac{6 \pm \sqrt{40}}{2}$$

Which gives the following two eigenvalues:

$\lambda_1 = 3 + \sqrt{10}$

$\lambda_2 = 3 - \sqrt{10}$

Eigenspace for λ_1: We need to compute the nullspace of $A - \lambda_1 I$. The matrix $A - \lambda_1 I$ is:

$$\begin{bmatrix} 2-(3+\sqrt{10}) & 3 \\ 3 & 4-(3+\sqrt{10}) \end{bmatrix} = \begin{bmatrix} -1-\sqrt{10} & 3 \\ 3 & 1-\sqrt{10} \end{bmatrix}$$

Performing Gaussian elimination consists of multiplying the first row by $\frac{3}{-1-\sqrt{10}}$ and then setting $R2 = R2 - R1$

$$\begin{bmatrix} -1-\sqrt{10} & 3 \\ 3 & 1-\sqrt{10} \end{bmatrix} \sim \begin{bmatrix} 3 & \frac{9}{-1-\sqrt{10}} \\ 3 & 1-\sqrt{10} \end{bmatrix} \sim \begin{bmatrix} 3 & \frac{9}{-1-\sqrt{10}} \\ 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{-3}{1+\sqrt{10}} \\ 0 & 0 \end{bmatrix}$$

The null space of this matrix $span(\begin{bmatrix} \frac{3}{1+\sqrt{10}} \\ 1 \end{bmatrix})$, so this is exactly the λ_1 eigenspace.

Eigenspace for λ_2:

The matrix $A - \lambda_2 I$ is:

$$\begin{bmatrix} 2-(3-\sqrt{10}) & 3 \\ 3 & 4-(3-\sqrt{10}) \end{bmatrix} = \begin{bmatrix} -1+\sqrt{10} & 3 \\ 3 & 1+\sqrt{10} \end{bmatrix}$$

Performing Gaussian elimination consists of multiplying the first row by $\frac{3}{-1+\sqrt{10}}$ and then setting

$R2 = R2 - R1$

$$\begin{bmatrix} -1+\sqrt{10} & 3 \\ 3 & 1+\sqrt{10} \end{bmatrix} \sim \begin{bmatrix} 3 & \frac{9}{\sqrt{10}-1} \\ 3 & 1+\sqrt{10} \end{bmatrix} \sim \begin{bmatrix} 3 & \frac{9}{\sqrt{10}-1} \\ 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{3}{\sqrt{10}-1} \\ 0 & 0 \end{bmatrix}$$

The null space of this matrix is $span(\begin{bmatrix} \frac{-3}{-1+\sqrt{10}} \\ 1 \end{bmatrix})$, so this is the λ_2 eigenspace.

Now, notice that the eigenvectors $\vec{v}_1 = \begin{bmatrix} \frac{3}{1+\sqrt{10}} \\ 1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} \frac{-3}{-1+\sqrt{10}} \\ 1 \end{bmatrix}$, corresponding to (different) eigenvalues λ_1 and λ_2 are orthogonal: $\langle v_1, v_2 \rangle = 0$. In fact, the inner-product of any vector $\vec{x}$ in the λ_1-eigenspace and any vector $\vec{y}$ in the λ_2 eigenspace will be zero (since all $\vec{x}$ can be written $\vec{x} = a\vec{v}_1$, and similarly $\vec{y} = b\vec{v}_2$ for some $a, b \in \mathbb{R}$).

It is always true that eigenvectors for distinct eigenvalues are linearly independent, but in the case of real *symmetric* matrices, it is also the case that eigenvectors for distinct eigenvalues are *orthogonal*. Try it for other symmetric matrices![1]

Orthogonality will play a bigger role in the second half of this course.

3. Mechanical Change of Basis

All calculations in this problem are intended to be done by hand, but you can use a computer to check your work.

(a) Consider two bases for $\mathbb{R}^2$.

$$A = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \qquad B = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\} \tag{1}$$

Suppose $\vec{x}_A$ represents the coordinates of a vector $\vec{x}$ in basis A and $\vec{x}_B$ represents the coordinates of the same vector in basis B. Write the coordinate transformation that converts $\vec{x}_A$ to $\vec{x}_B$. That is find the matrix T such that

$$\vec{x}_B = T\vec{x}_A \tag{2}$$

Solution:

We will abuse notation to use A and B to represent matrices with columns from the bases A and B respectively. That is

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \tag{3}$$

For a given vector $\vec{x}$, we have that

$$\vec{x} = A\vec{x}_A = B\vec{x}_B \tag{4}$$

[1] The proof of this fact is actually elementary but perhaps is not intuitively satisfying. Consider two different eigenspaces corresponding to $\lambda_1 \neq \lambda_2$. Choose eigenvectors $\vec{v}_1, \vec{v}_2$ from the corresponding eigenspaces so that $S\vec{v}_1 = \lambda_1\vec{v}_1$ and $S\vec{v}_2 = \lambda_2\vec{v}_2$.
Now consider the inner product $\vec{v}_1^T S\vec{v}_2 = (S^T\vec{v}_1)^T\vec{v}_2 = (S\vec{v}_1)^T\vec{v}_2$.
But $\vec{v}_1^T S\vec{v}_2 = \vec{v}_1^T(\lambda_2\vec{v}_2) = \lambda_2\vec{v}_1^T\vec{v}_2$. Similarly, $(S\vec{v}_1)^T, \vec{v}_2 = \lambda_1\vec{v}_1^T\vec{v}_2$. So we know that $(\lambda_1 - \lambda_2)\vec{v}_1^T\vec{v}_2 = 0$. Since $\lambda_1 \neq \lambda_2$, we are done since that means that $\vec{v}_1^T\vec{v}_2 = 0$.

Thus the matrix T is given by

$$T = B^{-1}A = \frac{1}{2-1}\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \tag{5}$$

(b) Consider two bases for $\mathbb{R}^2$.

$$A = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \qquad B = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\} \tag{6}$$

Suppose $\vec{x}_A$ represents the coordinates of a vector $\vec{x}$ in basis A and $\vec{x}_B$ represents the coordinates of the same vector in basis B. Write the coordinate transformation that converts $\vec{x}_A$ to $\vec{x}_B$. That is find the matrix T such that

$$\vec{x}_B = T\vec{x}_A \tag{7}$$

Solution:

Again let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \tag{8}$$

For a given vector $\vec{x}$, we have that

$$\vec{x} = A\vec{x}_A = B\vec{x}_B \tag{9}$$

Thus the matrix T is given by

$$T = B^{-1}A = \frac{1}{2-1}\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix} \tag{10}$$

(c) Consider two bases for $\mathbb{R}^3$.

$$A = \left\{ \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \right\} \qquad B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \right\} \tag{11}$$

Suppose $\vec{x}_A$ represents the coordinates of a vector $\vec{x}$ in basis A and $\vec{x}_B$ represents the coordinates of the same vector in basis B. Write the coordinate transformation that converts $\vec{x}_A$ to $\vec{x}_B$. That is find the matrix T such that

$$\vec{x}_B = T\vec{x}_A \tag{12}$$

Hint: What do you notice about A and B that will simplify this calculation?

Solution:

Again, we will abuse notation to use A and B to represent matrices with columns from the bases A and B respectively. That is

$$A = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & -1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \tag{13}$$

For a given vector $\vec{x}$, we have that

$$\vec{x} = A\vec{x}_A = B\vec{x}_B \tag{14}$$

Thus the matrix T is given by

$$T = B^{-1}A \tag{15}$$

Since the columns of B are orthonormal, $B^{-1} = B^T$. (This is also true for A.) As a result we can compute

$$T = B^{-1}A = B^TA \tag{16}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & -1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/2 & 1/\sqrt{2} & 1/2 \\ 1/2 & -1/\sqrt{2} & 1/2 \end{bmatrix} \tag{17}$$

(Actually in this case, B also happens to be symmetric so $B = B^T = B^{-1}$, but in general $B \neq B^{-1}$ for orthonormal matrices.) Since A and B are both orthonormal, T will also be orthonormal as well. (The product of orthonormal matrices is orthonormal.) T is an example of a coordinate transformation between orthonormal coordinate frames which can always be represented as an orthonormal matrix. (Check this!)

4. Mechanical Diagonalization

All calculations in this problem are intended to be done by hand, but you can use a computer to check your work.

(a) Diagonalize the matrices A and B, i.e. compute P_A, P_A^{-1}, D_A, P_B, P_B^{-1}, and D_B such that $A = P_A D_A P_A^{-1}$ and $B = P_B D_B P_B^{-1}$ and the D matrices are diagonal with the eigenvalues along the diagonal.

$$A = \begin{bmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix} \tag{18}$$

given that A has eigenvalues $\{1,\ 2\}$ and B has eigenvalues $\{1,\ -1\}$

Solution: We want to write A as $A = P_A D_A P_A^{-1}$ where D_A is diagonal. We first compute the eigenvectors of A. For eigenvalues 1 we get

$$\text{Nullspace of } \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{bmatrix} \right) = \text{Nullspace of } \left(\begin{bmatrix} -1/2 & -1/2 \\ -1/2 & -1/2 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{19}$$

and for eigenvalue 2 we get

$$\text{Nullspace of } \left(\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{bmatrix} \right) = \text{Nullspace of } \left(\begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{20}$$

P_A is a matrix whose columns are the eigenvectors of A. (Note: the norm of the columns of P_A doesn't matter. Why not?) Thus we can write P_A and P_A^{-1} as

$$P_A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \qquad P_A^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \tag{21}$$

Thus we can write A in diagonal form and check it

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{bmatrix} \tag{22}$$

For B we can compute the eigenvectors

$$\text{Nullspace of } \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix} \right) = \text{Nullspace of } \left(\begin{bmatrix} 0 & 0 \\ 2 & 2 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{23}$$

$$\text{Nullspace of } \left(\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix} \right) = \text{Nullspace of } \left(\begin{bmatrix} -2 & 0 \\ 2 & 0 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{24}$$

Again, P_B is a matrix whose columns are the eigenvectors of A. Thus we can write P_B and P_B^{-1} as

$$P_B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \qquad P_B^{-1} = \frac{1}{1} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{25}$$

Thus we can write B in diagonal form and check it

$$B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix} \tag{26}$$

(b) Diagonalize the matrix

$$A = \begin{bmatrix} 1/2 & 1/2 & -1/2 \\ -1/2 & 3/2 & 1/2 \\ -1 & 1 & 1 \end{bmatrix} \tag{27}$$

given that A has eigenvalues 1, 2, and 0.

Solution:

First we compute the eigenvectors of A given the eigenvalues.

For the eigenvalue of 2 the eigenvector spans

$$\text{Nullspace of } \left(\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} - \begin{bmatrix} 1/2 & 1/2 & -1/2 \\ -1/2 & 3/2 & 1/2 \\ -1 & 1 & 1 \end{bmatrix} \right) = \text{Nullspace of } \left(\begin{bmatrix} 3/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \\ 1 & -1 & 1 \end{bmatrix} \right) \tag{28}$$

By inspection we get the eigenvector is $[\ 0,\ 1,\ 1\ ]^T$. (If we didn't notice that the summing the last two columns gives 0, we could have solved using Gaussian Elimination.)

For the eigenvalue of 1 the eigenvector spans

$$\text{Nullspace of } \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1/2 & 1/2 & -1/2 \\ -1/2 & 3/2 & 1/2 \\ -1 & 1 & 1 \end{bmatrix} \right) = \text{Nullspace of } \left(\begin{bmatrix} 1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \\ 1 & -1 & 0 \end{bmatrix} \right) \tag{29}$$

By inspection, again we get that the eigenvector is $[\ 1,\ 1,\ 0\ ]^T$.

For the eigenvalue of 0, the eigenvector simply spans the nullspace of A.

$$\text{Nullspace of}\left(\begin{bmatrix}1/2 & 1/2 & -1/2\\ -1/2 & 3/2 & 1/2\\ -1 & 1 & 1\end{bmatrix}\right) \tag{30}$$

Again by inspection, we get the eigenvector is $[\ 1,\ 0,\ 1\ ]^T$.
We now write a matrix P whose columns are the eigenvectors of A.

$$P = \begin{bmatrix}0 & 1 & 1\\ 1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix} \tag{31}$$

We compute P^{-1} using Gaussian Elimination

$$\left[\begin{array}{ccc|ccc}1 & 0 & 0 & 0 & 1 & 1\\ 0 & 1 & 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 1 & 0 & 1\end{array}\right] \underbrace{\Rightarrow}_{\substack{\text{Switching}\\ \text{row order}}} \left[\begin{array}{ccc|ccc}0 & 0 & 1 & 1 & 0 & 1\\ 0 & 1 & 0 & 1 & 1 & 0\\ 1 & 0 & 0 & 0 & 1 & 1\end{array}\right] \underbrace{\Rightarrow}_{R2\leftarrow R2-R1} \left[\begin{array}{ccc|ccc}0 & 0 & 1 & 1 & 0 & 1\\ 0 & 1 & -1 & 0 & 1 & -1\\ 1 & 0 & 0 & 0 & 1 & 1\end{array}\right] \tag{32}$$

$$\underbrace{\Rightarrow}_{\substack{R3\leftarrow R3-R2\\ R3\leftarrow \frac{1}{2}R3}} \left[\begin{array}{ccc|ccc}0 & 0 & 1 & 1 & 0 & 1\\ 0 & 1 & -1 & 0 & 1 & -1\\ 1/2 & -1/2 & 1/2 & 0 & 0 & 1\end{array}\right] \underbrace{\Rightarrow}_{\substack{R1\leftarrow R1-R3\\ R2\leftarrow R2+R3}} \left[\begin{array}{ccc|ccc}-1/2 & 1/2 & 1/2 & 1 & 0 & 0\\ 1/2 & 1/2 & -1/2 & 0 & 1 & 0\\ 1/2 & -1/2 & 1/2 & 0 & 0 & 1\end{array}\right] \tag{33}$$

Thus

$$P^{-1} = \begin{bmatrix}-1/2 & 1/2 & 1/2\\ 1/2 & 1/2 & -1/2\\ 1/2 & -1/2 & 1/2\end{bmatrix} \tag{34}$$

(It's a good idea to check that $PP^{-1} = I$ which they do.)
We can now write A in it's diagonal form as

$$A = PDP^{-1} = \begin{bmatrix}0 & 1 & 1\\ 1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}\begin{bmatrix}2 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}-1/2 & 1/2 & 1/2\\ 1/2 & 1/2 & -1/2\\ 1/2 & -1/2 & 1/2\end{bmatrix} \tag{35}$$

(It's a good idea to multiply it out to check that you get A.)

5. Row Operations and Determinants

In this question we explore the effect of row operations on the determinant of a matrix. Recall from lecture that scaling a row by a will increase the determinant by a, and adding a multiple of one row to another will not change the determinant. The determinant of an identity matrix is 1 (corresponding to the volume of a unit hypercube).

(a) An upper triangular matrix is a matrix with zero below its diagonal. For example a 3×3 upper triangular matrix is :

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ 0 & b_2 & b_3 \\ 0 & 0 & c_3 \end{bmatrix}$$

By considering row-operations and what they do to a determinant, argue that the determinant of a general $n \times n$ upper-triangular matrix is the product of its diagonal entries, if they are non-zero. For example, the determinant of the 3×3 matrix above is $a_1 \times b_2 \times c_3$ if $a_1, b_2, c_3 \neq 0$.

Solution: A $n \times n$ upper-triangular matrix looks like:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,n} \end{bmatrix}$$

For every row i, divide it by $a_{i,i}$. Then we get 1s on the diagonal.

$$A' = \begin{bmatrix} 1 & \frac{a_{1,2}}{a_{1,1}} & \cdots & \frac{a_{1,n}}{a_{1,1}} \\ 0 & 1 & \cdots & \frac{a_{2,n}}{a_{2,2}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

The determinant of this new matrix is reduced by $1/(a_{1,1} \times \cdots \times a_{n,n})$ times:

$$\det A' = \frac{\det A}{a_{1,1} \times \cdots \times a_{n,n}}$$

Finally, starting from the last row, subtract multiples of the row from the ones above it, so that we get the $n \times n$ identity matrix I_n. This does not change the determinant since we are subtracting rows from each other. Thus:

$$1 = \det I_n = \det A' = \frac{\det A}{a_{1,1} \times \cdots \times a_{n,n}}$$

$$\det A = a_{1,1} \times \cdots \times a_{n,n}$$

(b) If the diagonal of an upper-triangular matrix has a zero entry, argue that its determinant is still the product of its diagonal entries.

Solution: If an upper-triangular matrix has a zero in its diagonal, it cannot be row-reduced into an identity matrix, this means that that its rows are linearly dependent. Therefore its determinant is zero, which is the product of all diagonal entries (since one of them is zero).

(c) Find a formula for the determinant of a general 3×3 matrix using Gaussian elimination, by keeping track of what each row operation does to the determinant of the matrix while reducing it to the identity matrix.

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

You may assume that the matrix is structured so that no division by zero occurs in your calculations (this simplifies the proof at the expense of getting a less general result). After simplification, you should get a summation with 6 terms.

Solution: First we row-reduce the matrix into an upper-triangular matrix:

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \xrightarrow[R_3 \leftarrow R_3 - R_1 c_1/a_1]{R_2 \leftarrow R_2 - R_1 b_1/a_1} \begin{bmatrix} a_1 & a_2 & a_3 \\ 0 & b_2 - b_1 a_2/a_1 & b_3 - b_1 a_3/a_1 \\ 0 & c_2 - c_1 a_2/a_1 & c_3 - c_1 a_3/a_1 \end{bmatrix}$$

$$\xrightarrow{R_3 \leftarrow R_2 \frac{c_2 - c_1 a_2/a_1}{b_2 - b_1 a_2/a_1}} \begin{bmatrix} a_1 & a_2 & a_3 \\ 0 & b_2 - b_1 a_2/a_1 & b_3 - b_1 a_3/a_1 \\ 0 & 0 & c_3 - c_1 a_3/a_1 - (b_3 - b_1 a_3/a_1)\frac{c_2 - c_1 a_2/a_1}{b_2 - b_1 a_2/a_1} \end{bmatrix}$$

We did not change the determinant of A in the process of row-reduction since we only subtracted rows from each other. From the first part, we know that the determinant of an upper-triangular matrix is the product of its diagonals, thus:

$$\begin{aligned}
\det A &= a_1 \times \left(b_2 - \frac{b_1 a_2}{a_1}\right) \times \left(c_3 - \frac{c_1 a_3}{a_1} - \left(b_3 - \frac{b_1 a_3}{a_1}\right) \frac{c_2 - c_1 a_2/a_1}{b_2 - b_1 a_2/a_1}\right) \\
&= a_1 \left(b_2 - \frac{b_1 a_2}{a_1}\right)\left(c_3 - \frac{c_1 a_3}{a_1}\right) - a_1 \left(b_3 - \frac{b_1 a_3}{a_1}\right)\left(c_2 - \frac{c_1 a_2}{a_1}\right) \\
&= \frac{1}{a_1}\left((a_1 b_2 - a_2 b_1)(a_1 c_3 - a_3 c_1) - (a_1 b_3 - a_3 b_1)(a_1 c_2 - c_1 a_2)\right) \\
&= \frac{1}{a_1}\left(a_1^2 b_2 c_3 - a_1 a_3 b_2 c_1 - a_1 a_2 b_1 c_3 + a_2 a_3 b_1 c_1 - a_1^2 b_3 c_2 + a_1 a_2 b_3 c_1 + a_1 a_3 b_1 c_2 - a_2 a_3 b_1 c_1\right) \\
&= a_1 b_2 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_1 b_3 c_2 - a_2 b_1 c_3 - a_3 b_2 c_1
\end{aligned}$$

(d) Do you see a pattern in the formula of the 3×3 determinant you derived in the previous part? How many times do the row indices a, b, c appear in each term of the determinant? How many times do the column indices 1, 2, 3 appear in each term of the determinant?

Solution: There is one of each a, b, c in each term, and one of each 1, 2, 3 in each term.

Extra Note: This is a consequence of another definition of the determinant. Each term in the sum is a product of three entries in the original matrix. These three entries are never in the same row or column. For example, if we consider $a_2 b_3 c_1$, their positions in A can be represented by the following matrix:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

This is exactly a permutation matrix, representing one permutation of 3 items (the one that brings $2 \to 1$, $3 \to 2$, $1 \to 3$). In fact, the determinant is the sum of the terms in A corresponding to all permutations of 3 items, of which there are 6 of them. You might also wonder about the signs present in the sum. How are they related to a permutation? It turns out that the sign is related to the number of swaps needed to get to the permutation. If an odd number of swaps is needed, the sign is negative, and if an even number of swaps is needed, the sign is positive. For example, to get the permutation corresponding to $a_2 b_3 c_1$, we need two swaps. To get to the permutation corresponding to $a_1 b_3 c_2$ ($1 \to 1$, $3 \to 2$, $2 \to 3$), we need a single swap.

6. Spectral Mapping and the Fibonacci Sequence

One of the most useful things about diagonalization is it allows us to easily compute polynomial functions of matrices. This in turn lets us do far more, including solving many linear recurrence relations. This problem shows you how this can be done for the Fibonacci numbers, but you should notice that the same exact technique can apply far more generally.

Suppose we have a matrix A that can be diagonalized as

$$A = PDP^{-1} = \begin{bmatrix} | & & | \\ \vec{p}_1 & \cdots & \vec{p}_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} | & & | \\ \vec{p}_1 & \cdots & \vec{p}_n \\ | & & | \end{bmatrix}^{-1} \tag{36}$$

where D is a diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_n$ on the diagonal and P is a matrix whose columns $\vec{p}_1, \ldots, \vec{p}_n$ are the eigenvectors of A.

(a) **Write out A^N in terms of P,P^{-1}, and D and simplify it as much as you can.** You should be able to show that you can write A^N as

$$A^N = PD^NP^{-1} = \begin{bmatrix} | & & | \\ \vec{p}_1 & \cdots & \vec{p}_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1^N & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \lambda_n^N \end{bmatrix} \begin{bmatrix} | & & | \\ \vec{p}_1 & \cdots & \vec{p}_n \\ | & & | \end{bmatrix}^{-1} \tag{37}$$

What does this say about any polynomial function of A?

Solution:

$$A^N = \underbrace{A \times A \times \cdots \times A}_{\times N} \tag{38}$$

$$= PD\underbrace{P^{-1} \times P}_{I}D\underbrace{P^{-1} \times P}_{I}DP^{-1} \times \cdots \times PD\underbrace{P^{-1} \times P}_{I}DP^{-1} \tag{39}$$

$$= PD^NP^{-1} \tag{40}$$

Since D is diagonal with the eigenvalues of $\lambda_1, \ldots, \lambda_n$ on the diagonal, we can easily compute

$$D^N = \begin{bmatrix} \lambda_1^N & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \lambda_n^N \end{bmatrix} \tag{41}$$

We note that this implies that we can evaluate any polynomial function of A by simply applying that polynomial to the eigenvalues of A. For example, can write

$$3A^{10} - 2A^9 = P\left(3D^{10}\right)P^{-1} + P\left(-2D^9\right)P^{-1} \tag{42}$$

$$= P\left(3D^{10} - 2D^9\right)P^{-1} \tag{43}$$

$$= P\begin{bmatrix} 3\lambda_1^{10} - 2\lambda_1^9 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 3\lambda_n^{10} - 2\lambda_n^9 \end{bmatrix}P^{-1} \tag{44}$$

and similarly for any polynomial.

(b) This idea that for diagonalizable matrices you can raise a matrix to any power by simply raising it's eigenvalues to that power is part of the **spectral mapping theorem**. We will now illustrate the power of this theorem to compute analytical expressions for numbers in the famous Fibonacci sequence.

Take a look at the Wikipedia article and find a cool fact about Fibonacci numbers to report!

Solution:

Give yourself credit for any cool fact. One of our favorites is "This means that every positive integer can be written as a sum of Fibonacci numbers, where any one number is used once at most."

(c) The Fibonacci sequence can be constructed according to the following relation. The Nth number in the Fibonacci sequence, F_N is computed by adding the previous two numbers in the sequence together

$$F_N = F_{N-1} + F_{N-2} \tag{45}$$

We select the first two numbers in the sequence to be $F_1 = 0$ and $F_2 = 1$ and then we can compute the following numbers as

$$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ldots \tag{46}$$

Notice that we can write the operation of computing the next Fibonacci numbers from the previous two using matrix multiplication

$$\begin{bmatrix} F_N \\ F_{N-1} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}}_{A} \begin{bmatrix} F_{N-1} \\ F_{N-2} \end{bmatrix} \tag{47}$$

Do you see why? Notice also that we could use powers of A to compute Fibonacci numbers starting from the original two, 0 and 1.

$$\begin{bmatrix} F_N \\ F_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{N-2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{48}$$

Diagonalize A and use Equation (37) to show that

$$F_N = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{N-1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{N-1} \tag{49}$$

is an analytical expression for the Nth Fibonacci number.

Note that A has eigenvalues and eigenvectors

$$\left\{\lambda_1 = \frac{1+\sqrt{5}}{2},\ \lambda_2 = \frac{1-\sqrt{5}}{2}\right\} \qquad \left\{\vec{p}_1 = \begin{bmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{bmatrix},\ \vec{p}_2 = \begin{bmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{bmatrix}\right\} \tag{50}$$

Feel free to use the 2×2 inverse formula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{51}$$

Solution: First, we diagonalize A

$$P = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix} \qquad P^{-1} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ -1 & \frac{1+\sqrt{5}}{2} \end{bmatrix} \tag{52}$$

$$A = PDP^{-1} \tag{53}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \left(\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ -1 & \frac{1+\sqrt{5}}{2} \end{bmatrix} \right) \tag{54}$$

Then, we have that F_N is equal to the first element of $A^{N-2} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

$$F_N = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{N-2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{55}$$

$$= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}^{N-2} \left(\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ -1 & \frac{1+\sqrt{5}}{2} \end{bmatrix} \right) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{56}$$

$$= \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{N-2} & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^{N-2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{bmatrix} \tag{57}$$

$$= \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^{N-1} - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2} \right)^{N-1} \tag{58}$$

(d) **(Bonus In-Scope)** Generalize what you found to a procedure that will give you, in principle, expressions for many linear recurrence relations that are recursively defined as $S_{n+k} = \sum_{i=0}^{k-1} \alpha_i S_{n+i}$ for some coefficients $\vec{\alpha}$ and initial conditions $[S_{k-1}, S_{k-2}, \ldots, S_0]^T = \vec{s}_0$.

Do this by setting up the appropriate matrix A and then invoking a computation of its eigenvalues and eigenvectors. And then using the results. (Feel free to assume diagonalizability of the resulting matrix, although there are some important cases when that does not hold.)

Solution:

Similar to Equation (47), we can write this recursive relationship as

$$\begin{bmatrix} S_{n+k} \\ S_{n+k-1} \\ S_{n+k-2} \\ \vdots \\ S_{n+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha_{k-1} & \alpha_{k-2} & \cdots & \alpha_1 & \alpha_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}}_{A} \begin{bmatrix} S_{n+k-1} \\ S_{n+k-2} \\ \vdots \\ S_{n+1} \\ S_n \end{bmatrix} \tag{59}$$

We can then write S_N (for $N > k-1$) as

$$S_N = \vec{e}_1^T A^{N-(k-1)} \vec{s}_0 \tag{60}$$

where $\vec{e}_1^T = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$ We can write $A = PDP^{-1}$ where D is a diagonal matrix with the eigenvalues of A on the diagonal $(\lambda_1, \ldots, \lambda_n)$ and the columns of P are the corresponding eigenvectors of A, $(\vec{p}_1, \ldots, \vec{p}_n)$. The above expression then becomes

$$S_N = \vec{e}_1^T P D^{N-(k-1)} P^{-1} \vec{s}_0 \tag{61}$$

If we define

$$P = \begin{bmatrix} | & & | \\ \vec{p}_1 & \cdots & \vec{p}_n \\ | & & | \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} - & \vec{q}_1^T & - \\ & \vdots & \\ - & \vec{q}_n^T & - \end{bmatrix} \tag{62}$$

we can write out S_N explicitly as

$$S_N = \vec{e}_1^T \vec{p}_1 \lambda_1^{N-(k-1)} \vec{q}_1^T \vec{s}_0 + \cdots + \vec{e}_1^T \vec{p}_n \lambda_n^{N-(k-1)} \vec{q}_n^T \vec{s}_0 \tag{63}$$

$$= \sum_{i=1}^{n} \vec{e}_1^T \vec{p}_i \lambda_i^{N-(k-1)} \vec{q}_i^T \vec{s}_0 \tag{64}$$

(e) **(Bonus Out-Of-Scope)** Take a closer look at the matrix A you constructed. Notice that it has a very special structure. Can you express $\det(A-\lambda I)$ as an explicit polynomial in λ that depends only on the α_i above?

Induction is helpful here, as is noticing the following facts about determinants that come from their fundamental nature as oriented volumes. (You should be able to derive these facts or at least see why these are true.)

- $\det([A,\vec{q}])$ is a linear function of $\vec{q}$, the last column of the matrix whose determinant is being taken. (Hint: think about volume. It must be zero if that last column is in the subspace spanned by A. And the volume is just a constant times the length of the component of $\vec{q}$ orthogonal to the subspace spanned by A whenever that subspace is $n-1$ dimensional. Because it is an oriented volume, the determinant only depends linearly on an A-dependent one-dimensional projection of $\vec{q}$.)
- $\det(\begin{bmatrix} A & \vec{0} \\ \vec{b}^T & 1 \end{bmatrix}) = \det(A)$. (i.e. The volume of a shape with a thickness of 1 is just the area in the lower-dimensiononal projection perpendicular to the direction of thickness. Similar intuition to the item above.)
- There is an oddity when the dimension n of the matrix is even. In those cases the sign flips when we try to do the same thing for the first row. $\det(\begin{bmatrix} \vec{b}^T & 1 \\ A & \vec{0} \end{bmatrix}) = (-1)^{n-1}\det(A)$. (i.e. This can be thought of as an aspect that comes from the "oriented" nature of the volume being computed. Cyclically shifting the columns (vectors) by 1 causes the sign of the oriented volume to flip when there are an even number of vectors. This can be viewed as shift by 1 of the natural $\det(\begin{bmatrix} 1 & \vec{b}^T \\ \vec{0} & A \end{bmatrix}) = \det(A)$.)
- The determinant of an upper-triangular matrix is just the product of the diagonal entries.

Solution:

We want to calculate

$$\det(A-\lambda I) = \det\left(\begin{bmatrix} \alpha_{k-1}-\lambda & \alpha_{k-2} & \cdots & \alpha_1 & \alpha_0 \\ 1 & -\lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -\lambda \end{bmatrix}\right) \tag{65}$$

Using the first hint, we label the right most column of the matrix $\vec{q}$ and notice that

$$\vec{q} = \begin{bmatrix} \alpha_0 \\ 0 \\ \vdots \\ 0 \\ -\lambda \end{bmatrix} = \alpha_0 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} + -\lambda \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{66}$$

and thus using the fact that $\det(\cdot)$ is a linear function of $\vec{q}$ we get that

$$\det(A-\lambda I) = \alpha_0 \det\left(\begin{bmatrix} \alpha_{k-1}-\lambda & \alpha_{k-2} & \cdots & \alpha_1 & 1 \\ 1 & -\lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}\right) + (-\lambda)\det\left(\begin{bmatrix} \alpha_{k-1}-\lambda & \alpha_{k-2} & \cdots & \alpha_1 & 0 \\ 1 & -\lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 1 \end{bmatrix}\right) \tag{67}$$

Applying the second and third hints, we get that

$$\det(A-\lambda I) = \alpha_0(-1)^{k-1} \det\left(\begin{bmatrix} 1 & -\lambda & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}\right) + (-\lambda)\det\left(\begin{bmatrix} \alpha_{k-1}-\lambda & \alpha_{k-2} & \cdots & \alpha_2 & \alpha_1 \\ 1 & -\lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -\lambda \end{bmatrix}\right) \tag{68}$$

By the fourth hint, the determinant of the first matrix is simply 1. Notice also, that the second matrix has the same form as the initial matrix we started with except it has dimension $k-1 \times k-1$ (instead of $k \times k$). So, informally, if we think of the original matrix as A_k and the smaller submatrix as A_{k-1}, we have the recursion:

$$\det(A_k - \lambda I) = \alpha_0(-1)^{k-1} - \lambda \det(A_{k-1} - \lambda I) \tag{69}$$

$$= \alpha_0(-1)^{k-1} + \lambda(-\det(A_{k-1} - \lambda I)) \tag{70}$$

At this point, formally, we could proceed by induction. We will do that at the end, just to show you how that is done. Alternatively, we can just repeatedly apply the same set of arguments to get that:

$$\det(A_k - \lambda I) = \alpha_0(-1)^{k-1} + \lambda(-\det(A_{k-1} - \lambda I)) \tag{71}$$

$$= \alpha_0(-1)^{k-1} + \lambda(-(\alpha_1(-1)^{k-2} + \lambda(-\det(A_{k-2} - \lambda I)))) \tag{72}$$

$$= \alpha_0(-1)^{k-1} + \lambda(\alpha_1(-1)^{k-1} + \lambda \det(A_{k-2} - \lambda I)) \tag{73}$$

Notice that after expanding twice, we get something with the same parity. If k is even, $k-2$ is even. Similarly, if k is odd, $k-2$ is odd. So this expansion will just keep happening this way forever until it ends with a final term that is either a 1x1 or a 2x2 matrix. If it ends with a 2x2 matrix, then we know that k was even to start with and all the $(-1)^{k-1}$ terms are just -1. In that case, notice that the 2x2 case has negative signs too and has determinant $-\lambda\alpha_{k-1} - \alpha_{k-2} + \lambda^2$. If it ends with a 1x1 matrix, then we know that k was odd to start with and all the $(-1)^{k-1}$ terms are just $+1$. In that case, notice that the 1x1 case has positive signs too and has determinant $\alpha_{k-1} - \lambda$. In both cases, the sign of the highest degree λ is flipped from all the others. Writing this out:

$$\det(A - \lambda I) = \alpha_0(-1)^{k-1} + \lambda\left[\alpha_1(-1)^{k-1} + \lambda\left[\alpha_2(-1)^{k-1} + \lambda\left[\alpha_3(-1)^{k-1} + \dots\right.\right.\right. \tag{74}$$

$$\left.\left.\left. + \lambda\left[\alpha_{k-2}(-1)^{k-1} + \lambda\left[\alpha_{k-1}(-1)^{k-1} - \lambda(-1)^{k-1}\right]\right]\cdots\right]\right]\right] \tag{75}$$

Thus we have that

$$\det(A - \lambda I) = (-1)^{k-1}\left[\alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \cdots + \alpha_{k-1}\lambda^{k-1} - \lambda^k\right] \tag{76}$$

The above represents how we can figure out the answer without knowing what it is. But once we get it, we can just quickly prove it is correct by induction. It is clearly the right expression for the base case $k = 1$. Then by induction, we see that:

$$\begin{aligned}
\det(A_k - \lambda I) &= \alpha_0(-1)^{k-1} + \lambda(-\det(A_{k-1} - \lambda I)) && (77)\\
&= \alpha_0(-1)^{k-1} + \lambda(-1(-1)^{k-2}\left[\alpha_1 + \alpha_2\lambda + \alpha_3\lambda^2 + \cdots + \alpha_{k-1}\lambda^{k-2} - \lambda^{k-1}\right]) && (78)\\
&= (-1)^{k-1}(\alpha_0 + \lambda\left[\alpha_1 + \alpha_2\lambda + \alpha_3\lambda^2 + \cdots + \alpha_{k-1}\lambda^{k-2} - \lambda^{k-1}\right]) && (79)\\
&= (-1)^{k-1}(\alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \alpha_3\lambda^3 + \cdots + \alpha_{k-1}\lambda^{k-1} - \lambda^k) && (80)
\end{aligned}$$

which is what we wanted to show. So by induction, we are done.

7. Your Own Problem Write your own problem related to this week's material and solve it. You may still work in groups to brainstorm problems, but each student should submit a unique problem. What is the problem? How to formulate it? How to solve it? What is the solution?