Revisiting Coordinates, Eigenvectors and Diagonalization

- Think about what you did in your group project on Friday. You found a polynomial with the property that when you differentiate it and multiply by $2t - 1$ you get a number times the polynomial you started with. You did that by translating the whole question into the language of matrices. This is one of the useful things about coordinates. It puts the elements of any vector space on equal footing with the vectors in $\mathbb{R}^n$ which we can answer questions about using row reduction and determinants!

- But, that is not the only use of coordinates. Sometimes, it is useful to choose better coordinates even for $\mathbb{R}^n$. In particular, if you are going to be working a lot with a particular matrix $A$, then it would be wise to use a basis made up entirely of eigenvectors of $A$. I would like to explain why. Along the way, this will give us a chance to review and dig deeper into many of the things we have already seen. (All of this material is covered in the book in sections we have already seen. It is not “new”.)

- Matrices Can Mix Up the Coordinates: Let’s consider a $2 \times 2$ matrix and a vector in $\mathbb{R}^2$:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad x = \begin{bmatrix} x \\ y \end{bmatrix}.$$  

We can think of $x$ as being a point in the plane, with coordinates $(x, y)$. Notice that if we “transform” it by multiplying by the matrix $A$, the new $x$-coordinate (which is $ax + by$) depends on $x$ and $y$, as does the new $y$ coordinate ($cx + dy$). This can make it a little hard to think about the geometry of the transformation.

- Diagonal Matrices are the Ones that Don’t: But, if $b = c = 0$, then $A$ is a diagonal matrix and $Ax$ is just $(ax, dy)$. This makes it easy. The $x$-coordinate is scaled by $a$ and the $y$-coordinate is scaled by $d$ and that’s it.

- Coordinates are a Choice: The association of coordinates to a vector (or a point in the plane) are something we have control over. We can decide to assign the coordinates differently. As we’ve seen, coordinates for a vector space are determined by a choice of basis. Let’s see how that works. Think of the vector $v = [2, 3]^T$. The reason we associate those numbers to it is because we are thinking in terms of the basis $S = \{e_1, e_2\}$ and $v = 2e_1 + 3e_2$. More generally, we can picture a “grid” with vertices at every point $ae_1 + be_2$ where $a$ and $b$ are integers. This makes it easy for us to look at a vector and determine what its coordinates are relative to the basis $S$.

- Picturing Other Coordinates: But, we can do the same with another choice of basis as well. Suppose, for instance, we choose

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$
If we draw a grid with vertices at each point that is a linear combination of these two with integer coefficients, we get a new coordinate grid. Just tilt your head a bit to the right and picture the line with slope 1 to be the new "y-axis" and the line perpendicular to it to be the new "x-axis." The vector in blue in the figure below is the same vector \( v \) from above. But, now we can see that relative to the basis \( B \) it has coordinates \(-1/2\) and \( 5/2 \) because it is \(-1/2\) times the first basis element plus \( 5/2 \) the second. (Again, tilt your head to see how this makes sense.)

![Figure](image.png)

- **Another Example:** The previous example is a good first one because the basis vectors in \( B \) are just the ones from \( S \) rotated and scaled by the same amount, leaving the "grid" square. But, there is no reason this must be the case. In the figure below, we can see that the same vector \( v \) which had coordinates 2 and 3 relative to \( S \) has coordinates \( 1/3 \) and \( 4/3 \) relative to

\[
C = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\} = \{c_1, c_2\}.
\]

![Figure](image.png)

- **A Basis of Eigenvectors Diagonalizes:** The matrix \( A = \begin{bmatrix} 0 & 2 \\ -2 & 5 \end{bmatrix} \) has both vectors in \( C \) as eigenvectors. (The first basis vector \( c_1 \) has eigenvalue \( \lambda = 1 \) and \( c_2 \) has eigenvalue \( \lambda = 4 \).) Think about what this means about the action of \( A \) on the \( C \)-coordinates of a vector. If \( w = a_1 c_1 + a_2 c_2 \) then \( A w = a_1 c_1 + 4a_2 c_2 \). In other words, it leaves the first coordinate the same and multiplies the second by 4. Note that it does not mix them up. This is reflected in the fact that the \( C \)-matrix representing the linear transformation \( w \mapsto A w \) is diagonal. (The first column would be \( \begin{bmatrix} 1 & 0 \end{bmatrix}^T \) because the image of \( c_1 \) is \( c_1 \) and the second
column would be $[0 \ 4]^{T}$ because the image of $c_2$ is $4c_2$. It also is reflected in the fact that $A = PDP^{-1}$ where $P$ is the matrix whose columns are the basis $C$ and $D$ is the diagonal matrix we constructed above. This is why a matrix is diagonalizable if and only if the vector space has a basis made up entirely of its eigenvectors!

- I hope that helps you to understand why we were talking about matrix representations, eigenvectors and diagonalizability last week. If there is more time, I will say a few words about the application of these ideas to dynamical systems and the geometric interpretation of complex eigenvalues. However, I will not assign any homework from the two sections on that material. So, the homework below is just the homework from last Friday's handout repeated.

**Homework**

This is the same homework assignment that was listed in last Friday's handout. No new problems are being assigned today:

**Section 5.4:** Read the section and then do these problems: 1, 2, 5, 6*, 7, 9–13, 14*, 15 16, 23, 25

---

**Material from Sections 5.5 and 5.6 (not to be tested)**

- When a mathematician calls something a dynamical system, they mean that it is something that changes in time according to mathematical rules. There are many different kinds of dynamical systems, but we will especially be interested in the ones that take place in a vector space where a vector at one moment in time is turned into its next “state” by being multiplied by a matrix. That is, we consider a vector $v_0$ which is the initial state of the system. One unit of time later it is $v_1 = Av_0$, then $b_2 = Ab_1$ and in general at time $t = k$ it is $v_k = A^k v_{k-1}$.

- **An Ecological Example:** Let $R_k$ be the number of rats (in thousands) and $O_k$ be the number of owls in a certain large forest at time $t = k$ months. From these two numbers we can figure out how many rats and owls there will be the next month:

  $$O_{k+1} = .5O_k + .4R_k \quad R_{k+1} = -.104O_k + 1.1R_k.$$  

  (We’re ignoring the question of whether there is enough food and space for the rats, just to keep things simple.)

- (How to read these formulas and make sense of them: If there were no rats, half of the owls would die each month until none were left. Similarly, if there are no owls then the rat population will grow by 10% a month. On the other hand, one owl eats 104 rats a month and if there are enough of those then the owl population will grow too.)

- This is a dynamical system that can be written in the form specified in the first paragraph because

  $$\begin{bmatrix} O_{k+1} \\ R_{k+1} \end{bmatrix} = \begin{bmatrix} .5 & .4 \\ -.104 & 1.1 \end{bmatrix} \begin{bmatrix} O_k \\ R_k \end{bmatrix} = A^{k+1} \begin{bmatrix} O_0 \\ R_0 \end{bmatrix}.$$  

  **Question 1:** So if you have 25 owls and 12 thousand rats one month, how many of each will be there the following month?
• Even though each decreases at first, if we look over a longer time period, the situation turns around and the populations start increasing! Moreover, the trajectory of the sequence of vectors starts looking very much like a straight line. What’s going on?

• Since the answer has to do with eigenvalues and eigenvectors, let’s find the eigenvalues and eigenvectors for this matrix $A$. Moreover, even though I’m usually a “stickler” for being mathematical precise, I will allow for rounding off here, since in the mathematical model there is certainly some approximation going on already. (E.g. there cannot be 10.5 owls for real!)

   - The characteristic polynomial of $A$ is $p(\lambda) = 0.5916 - 1.6\lambda + \lambda^2 = (\lambda - 1.02)(\lambda - 0.58)$.
   - $v_1 = \begin{bmatrix} 10 \\ 13 \end{bmatrix}$ is an eigenvector for $\lambda = 1.02$ and $v_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ is an eigenvector for $0.58$.

• What does this mean for the dynamical systems? It means that the ratio of 10 owls to 13 thousand rats is one that will persist, as is the ratio of 5 owls to 1000 rats. In each case, the ratio will stay the same from month to month. But in one case the populations will both increase and in the other case they will both decrease.

• More interestingly (and more importantly) is the question of what happens to a vector that is not an eigenvector. Since these eigenvectors form a basis for $\mathbb{R}^2$, any vector is a linear combination of them. So, let’s consider as a pretty reasonable example the vector $v_3 = 0.5v_1 + 2v_2$.

   - Note that $v_3$ is not an eigenvector because $Av_3 = (1.02)(0.5)v_1 + (0.58)(2)v_2$. More generally, $A^k v_3 = (1.02)^k(0.5)v_1 + (0.58)^k(2)v_2$.

   - More importantly, we can see that the $v_2$ term is going to disappear as $k$ gets large because $(0.58)^k$ gets tiny while the $(1.02)^k$ gets bigger and bigger.

---

**Key Fact:** The leading eigenvector (the one with the largest absolute value) is the most important one. In particular, in the long run, a linear combination of eigenvectors always approaches an eigenvector for the eigenvalue with the largest absolute value.

• This is what happened in our opening example with the rats and owls. Even though there was a decrease initially, in the long run the ratio of 10 to 13 wins out. (Unless the initial population ratio is exactly 5 to 1.)

• **Complex Eigenvalues:** The matrix $A$ in the example above had real eigenvalues, so we can imagine what that would mean for the population. But what could it possibly mean when the eigenvalues are complex?!? As we will see, even complex eigenvalues tell you a lot about dynamics.

   - (Quick review of complex numbers: $i^2 = -1$ is the only thing we need to add to the reals, $z = a + bi$ is a complex number, $\overline{z} = a - bi$ is its complex conjugate. We also use $\text{Re} z = a$ and $\text{Im} z = b$. Now if $p(x)$ is a polynomial of degree $n$ then $p(x) = (x - z_1)(x - z_2) \cdots (x - z_n)$ for some roots $z_j$.)

**Question 2:** Find the eigenvalues and eigenvectors of

$\begin{bmatrix} 0.5 & 0.6 \\ 0.75 & 1.1 \end{bmatrix}$. 

(Answer: $\lambda = .8 \pm .6i$. Now subtract this from the diagonal of $A$ to get a matrix which must have rank one since there are nonzero elements in its null space. This means we can basically consider either row in trying to solve and ignore the other since we know they are linearly dependent! Thus $\bar{x} = [-.4 \pm .8i]$ are the corresponding eigenvectors!)

- **Theorem:** If $A$ is a real matrix then if $\lambda$ is an eigenvalue with eigenvector $x$ then $\bar{\lambda}$ is an eigenvalue with eigenvector $\bar{x}$!

  - The trajectories of any initial vector under the dynamical system with this matrix $A$ are elliptical. How could we tell this from looking at the eigenvalues? What other information do they provide?

- $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ is a matrix which has $\lambda = a \pm bi$ as eigenvalues and acts by rotating by the angle $(a, b)$ makes with the axis and dilating by $r = \sqrt{a^2 + b^2}$ (so spirals out if $r > 1$, in if $r < 1$ and stays on the circle if $r = 1$).

- More generally, if $A$ is any $2 \times 2$ matrix with eigenvalue $\lambda = a + bi$ then $A = PCP^{-1}$ where $P = [\text{Re } v \ \text{Im } v]$ and $A$ is the matrix from the previous bullet. The only effect of the $P$ is to turn circles into ellipses, but we can still tell how it is going to rotate and whether it will spiral out or in from $a$ and $b$. Thus, the eigenvalues still give us dynamical information even if they are complex.