Lab 10: Eigenvalues and Eigenvectors, Convexity Solutions

EECS 245, Spring 2026 at the University of Michigan

due for completion at 11:59PM Ann Arbor Time on Monday, June 15th, 2026

Name: ______________________________

uniqname: ______________________________

Each lab worksheet will contain several activities, some of which will involve writing code and others that will involve writing math on paper. To receive credit for a lab, you must complete as many of the activities as you can in 2 hours and submit a PDF of your work to Gradescope. We will provide specific instructions on how to submit programming activities (e.g. submitting the notebook or including a screenshot of some output).

Feel free to work with others in the course, but you must submit individually.

Recap: Eigenvalues and Eigenvectors

Let $A = \begin{bmatrix} 6 & 3 \\ 3 & -2 \end{bmatrix}$.

- An **eigenvector** of A is a non-zero vector $\vec{v}$ such that $A\vec{v} = \lambda\vec{v}$ for some scalar λ. The scalar λ is called the **eigenvalue** corresponding to $\vec{v}$. For A's eigenvectors, multiplying by A is equivalent to multiplying by a scalar.
- The **characteristic polynomial** of A is given by $p(\lambda) = \det(A - \lambda I)$.

  $$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 6-\lambda & 3 \\ 3 & -2-\lambda \end{vmatrix} = (6-\lambda)(-2-\lambda) - 3 \cdot 3 = \lambda^2 - 4\lambda - 21 = (\lambda+3)(\lambda-7)$$

- The eigenvalues of A are the roots of the characteristic polynomial, so $\lambda_1 = -3$ and $\lambda_2 = 7$.
  - The eigenvector $\vec{v}_1$ satisfies $A\vec{v}_1 = -3\vec{v}_1$.

    $$\begin{bmatrix} 6 & 3 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = -3 \begin{bmatrix} a \\ b \end{bmatrix} \implies b = -3a$$

    So any vector of the form $\begin{bmatrix} a \\ -3a \end{bmatrix}$ ($a \neq 0$) is an eigenvector of A corresponding to the eigenvalue -3. We could pick $\boxed{\vec{v}_1 = \begin{bmatrix} 2 \\ -6 \end{bmatrix}}$.
  - The eigenvector $\vec{v}_2$ satisfies $A\vec{v}_2 = 7\vec{v}_2$. Another way to find it is to solve for the null space of $A - 7I = \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}$. One vector in nullsp$(A - 7I)$ is $\boxed{\vec{v}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}}$.

Activity 1: Introduction

For each 2×2 matrix A below:

(i) Find the characteristic polynomial of A, and use it to find the eigenvalues of A.

(ii) Find one eigenvector for each eigenvalue of A. Verify that each eigenvector is indeed an eigenvector of A by multiplying it by A.

(iii) **By hand (not using Python or Desmos)**, draw a picture (like the one in Chapter 9.1 titled **Visualizing the eigenvectors of A**) with vectors $\vec{v}_1, A\vec{v}_1, \vec{v}_2, A\vec{v}_2$ as arrows (where $\vec{v}_1$ and $\vec{v}_2$ are the eigenvectors you found above).

a) $A = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$

Solution:

(i) The characteristic polynomial of A is

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 3-\lambda & 0 \\ 0 & 4-\lambda \end{vmatrix} \\ &= (3-\lambda)(4-\lambda) \end{aligned}$$

So, the eigenvalues of A are $\lambda_1 = 3$ and $\lambda_2 = 4$. A is a diagonal matrix, which means its eigenvalues are its diagonal entries.

(ii) Matching examples we've seen in class (in particular, **this example** from Chapter 9.2), an eigenvector for $\lambda_1 = 3$ is $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and an eigenvector for $\lambda_2 = 4$ is $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$; indeed, $A\vec{v}_1 = 3\vec{v}_1$ and $A\vec{v}_2 = 4\vec{v}_2$.

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

Solution:

(i) The characteristic polynomial of A is

$$\begin{aligned}\det(A-\lambda I) &= \begin{vmatrix}3-\lambda & 4\\ 4 & 3-\lambda\end{vmatrix}\\ &= (3-\lambda)^2-16\\ &= \lambda^2-6\lambda+9-16\\ &= \lambda^2-6\lambda-7\\ &= (\lambda+1)(\lambda-7)\end{aligned}$$

So, the eigenvalues of A are $\lambda_1=-1$ and $\lambda_2=7$.

(ii)
- The eigenvector $\vec{v}_1$ corresponding to $\lambda_1=-1$ satisfies $A\vec{v}_1=-1\vec{v}_1$.

$$\begin{bmatrix}3 & 4\\ 4 & 3\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}=-1\begin{bmatrix}a\\ b\end{bmatrix}$$

The first component implies $3a+4b=-a \implies 4a+4b=0 \implies a=-b$. So, an eigenvector $\vec{v}_1$ is $\begin{bmatrix}1\\ -1\end{bmatrix}$ (though there are infinitely many other choices; any scalar multiple of $\begin{bmatrix}1\\ -1\end{bmatrix}$ is also an eigenvector for eigenvalue -1). Note that using the second component would give us the same relationship. To verify that we correctly found the eigenvector-eigenvalue pair, multiplying A by $\begin{bmatrix}1\\ -1\end{bmatrix}$ gives

$$A\begin{bmatrix}1\\ -1\end{bmatrix}=\begin{bmatrix}3-4\\ 4-3\end{bmatrix}=\begin{bmatrix}-1\\ 1\end{bmatrix}=-1\begin{bmatrix}1\\ -1\end{bmatrix}$$

(We will omit this step moving forward.)

- The eigenvector $\vec{v}_2$ corresponding to $\lambda_2=7$ satisfies $A\vec{v}_2=7\vec{v}_2$.

$$\begin{bmatrix}3 & 4\\ 4 & 3\end{bmatrix}\begin{bmatrix}c\\ d\end{bmatrix}=7\begin{bmatrix}c\\ d\end{bmatrix}$$

The first component implies $3c+4d=7c \implies 4c-4d=0 \implies c=d$. So, an eigenvector $\vec{v}_2$ is $\begin{bmatrix}1\\ 1\end{bmatrix}$.

To conclude,

$$\boxed{\lambda_1=-1,\vec{v}_1=\begin{bmatrix}1\\ -1\end{bmatrix}, \qquad \lambda_2=7,\vec{v}_2=\begin{bmatrix}1\\ 1\end{bmatrix}}$$

Activity 2: Rapid Fire

The goal of this activity is to practice spotting eigenvalues and characteristic polynomials quickly. Two quick facts:

- The **sum** of the eigenvalues of a matrix is equal to the **trace** of the matrix (which is the sum of the diagonal entries).
- The **product** of the eigenvalues of a matrix is equal to the **determinant** of the matrix.

a) A 2×2 matrix A has $\text{trace}(A) = 5$ and $\det(A) = 6$. What are the eigenvalues of A?

> **Solution:** $\lambda_1 = 2, \lambda_2 = 3$.
> The eigenvalues of A must sum to 5 and multiply to 6. The only possible solution is $\lambda_1 = 2$ and $\lambda_2 = 3$, which indeed satisfy both conditions. There are infinitely many **matrices** that satisfy these conditions, but the eigenvalues in all of them are 2 and 3.

b) A non-invertible 2×2 matrix has an eigenvalue of 5. What is its characteristic polynomial?

> **Solution:** $p(\lambda) = \lambda^2 - 5\lambda$.
>
> Let A be the matrix in question. Since A is not invertible, 0 is one of its eigenvalues. Since it's a 2×2 matrix, it can only have two eigenvalues, so its two eigenvalues are 0 and 5. So, its characteristic polynomial is a polynomial with roots 0 and 5, i.e.
>
> $$p(\lambda) = (\lambda - 0)(\lambda - 5) = \lambda^2 - 5\lambda$$

c) A 3×3 matrix A has $\det(A) = 20$ and two unique **positive integer** eigenvalues, one of which is repeated twice. In other words, $p(\lambda)$ has the form

$$p(\lambda) = (\lambda - \lambda_1)^2(\lambda - \lambda_2)$$

(λ_1 has an **algebraic multiplicity** of 2. This is a term we'll see more in tomorrow's lecture and Chapter 9.4.)
What are all possible values of λ_1 and λ_2?

> **Solution:** $\lambda_1 = 2, \lambda_2 = 5$ or $\lambda_1 = 1, \lambda_2 = 20$.
>
> The product of A's eigenvalues — including arithmetic multiplicities — is equal to the determinant of A. Since λ_1 has $\text{AM}(\lambda_1) = 2$, it must appear twice in the product of the eigenvalues. So,
>
> $$\lambda_1^2 \lambda_2 = 20$$
>
> We're told that λ_1 and λ_2 are integers. Note that $20 = 2 \cdot 10 = 2 \cdot 2 \cdot 5$. The only possible solutions are $\lambda_1 = 2$ and $\lambda_2 = 5$ (which means $\lambda_1^2\lambda_2 = (2)^2 \cdot 5 = 20$) or $\lambda_1 = 1$ and $\lambda_2 = 20$ (which means $\lambda_1^2\lambda_2 = (1)^2 \cdot 20 = 20$).

Activity 3: Quadratic Forms Return

Open Desmos in 3D mode at **desmos.com/3d** and write $z = x^2 + 2bxy + 16y^2$. This should show you a 3D surface along with a slider for b. Drag the slider to see how the shape of the surface changes for different b's. You should notice that depending on the value of b, the surface may or may not have a global minimum. Let's explore!

a) z is a quadratic form, $f(\vec{x}) = \vec{x}^T A\vec{x}$, where $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and A is a symmetric matrix. Find A.

Solution: $A = \begin{bmatrix} 1 & b \\ b & 16 \end{bmatrix}$.

To see where this comes from, let's expand $f(\vec{x}) = \vec{x}^T A\vec{x}$:

$$\begin{aligned} f(\vec{x}) &= \vec{x}^T A\vec{x} \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & b \\ b & 16 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} x + by \\ bx + 16y \end{bmatrix} \\ &= x(x + by) + y(bx + 16y) \\ &= x^2 + bxy + bxy + 16y^2 \\ &= x^2 + 2bxy + 16y^2 \end{aligned}$$

So, $A = \begin{bmatrix} 1 & b \\ b & 16 \end{bmatrix}$.

b) For a vector-to-scalar function $f : \mathbb{R}^n \to \mathbb{R}$, the **Hessian** of f, denoted $\nabla^2 f$, is the $n \times n$ matrix of second partial derivatives of f. Find $\nabla^2 f$ for $f(\vec{x}) = \vec{x}^T A\vec{x}$.

Solution: $\nabla^2 f = \begin{bmatrix} 2b & 2b \\ 2b & 32 \end{bmatrix}$.

To find the second partial derivatives of f, we first find the first partial derivatives:

$$f(\vec{x}) = \vec{x}^T A\vec{x} = x^2 + 2bxy + 16y^2$$
$$\frac{\partial f}{\partial x} = 2x + 2by$$
$$\frac{\partial f}{\partial y} = 2bx + 32y$$

Since there are two first partial derivatives, there are $2 \times 2 = 4$ second partial derivatives. We can find them all by taking the partial derivatives of the first partial derivatives:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}(2x + 2by) = 2$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y}(2x + 2by) = 2b$$
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial x}(2bx + 32y) = 2b$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}(2bx + 32y) = 32$$

Note that $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$, which is a general property of second partial derivatives — it doesn't matter which order we take the derivatives in. So, the Hessian is

$$\nabla^2 f = \begin{bmatrix} 2 & 2b \\ 2b & 32 \end{bmatrix}$$

Notice that this is just $2A$!

c) A symmetric matrix A is **positive semidefinite** (PSD) if $\vec{v}^T A\vec{v} \geq 0$ for all $\vec{v} \in \mathbb{R}^n$. In English, this says that A is positive semidefinite if the quadratic form $f(\vec{v}) = \vec{v}^T A\vec{v}$ is always non-negative for all $\vec{v} \in \mathbb{R}^n$. Two relevant facts:

- A differentiable vector-to-scalar function f is **convex** if its Hessian is PSD.
- A symmetric matrix A is PSD if and only if all of its eigenvalues are non-negative.

Using the facts above, find the range of values b for which f is convex, and verify your answer by dragging the slider on Desmos.

Solution: As long as $-4 \leq b \leq 4$, the Hessian is PSD, and f is convex.

Recall from the previous part that the Hessian is

$$\nabla^2 f = \begin{bmatrix} 2 & 2b \\ 2b & 32 \end{bmatrix} = 2A$$

f is convex if and only if $2A$ is PSD, which means that both of $2A$'s eigenvalues are non-negative. Let λ_1 and λ_2 be $2A$'s eigenvalues. Then, we need $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$. How do we ensure this?

Let's recall the facts about the trace and determinant introduced at the start of Activity 2:

- The sum of $2A$'s eigenvalues, $\lambda_1 + \lambda_2$, is equal to the trace of $2A$, trace$(2A)$, which is the sum of the diagonal entries of $2A$. Here, this is $2 + 32 = 34$.
- The product of $2A$'s eigenvalues, $\lambda_1\lambda_2$, is equal to the determinant of $2A$, det$(2A)$, which is $2 \cdot 32 - (2b)^2 = 64 - 4b^2$.

The fact that $\lambda_1 + \lambda_2 = 34$ means that no matter what b is, at least one of the eigenvalues is non-negative, since you can't add two negative numbers and get a positive sum.

So, let's focus our attention on the product of the eigenvalues, $\lambda_1\lambda_2 = 64 - 4b^2$. We need this to be ≥ 0, because if it were negative, then one of the eigenvalues would be negative (since multiplying a positive number by a negative number gives a negative number).

Using this reasoning, we need $64 - 4b^2 \geq 0$, which simplifies to $4b^2 \leq 64$, which simplifies to $b^2 \leq 16$, which simplifies to $-4 \leq b \leq 4$. If $|b| > 4$, then one of the eigenvalues would be negative, $2A$ would not be PSD, and f would not be convex.

Notice that in this case, the "test" for convexity ended up simplifying to checking whether the determinant of A is non-negative. That is true, in general, for quadratic forms $f(\vec{x}) = \vec{x}^T A\vec{x}$ where A is a 2×2symmetric matrix. But if A is 3×3 or larger, the determinant test alone isn't sufficient; it's possible to have a positive determinant and trace but a negative eigenvalue in a 3×3 matrix. What **is** true in general is that if A is a symmetric $n \times n$ matrix, then $f(\vec{x}) = \vec{x}^T A\vec{x}$ is convex if and only if all of A's eigenvalues are non-negative.

Activity 4: Understanding Complex Proofs

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. It turns out that the function $g(\vec{x})$, defined by

$$g(\vec{x}) = f(A\vec{x} + \vec{b})$$

for some $n \times n$ matrix A and vector $\vec{b} \in \mathbb{R}^n$, is also convex, no matter what A and $\vec{b}$ are. We're not going to ask you to prove this on your own: instead, we'll give you a proof and ask you questions to ensure you understand it.

Our **goal** is to show that $g((1-t)\vec{x} + t\vec{y}) \leq (1-t)g(\vec{x}) + tg(\vec{y})$, for all $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $t \in [0, 1]$. We'll start with the "left-hand side" of the definition, and try and leverage f's convexity.

$$\begin{aligned}
g((1-t)\vec{x} + t\vec{y}) &= f\left(A\left((1-t)\vec{x} + t\vec{y}\right) + \vec{b}\right) && (1)\\
&= f\left((1-t)A\vec{x} + tA\vec{y} + \vec{b}\right) && (2)\\
&= f\left((1-t)(A\vec{x} + \vec{b}) + t(A\vec{y} + \vec{b})\right) && (3)\\
&\leq (1-t)f(A\vec{x} + \vec{b}) + tf(A\vec{y} + \vec{b}) && (4)\\
&= \boxed{(1-t)g(\vec{x}) + tg(\vec{y})} && (5)
\end{aligned}$$

---

a) In which line did we use the fact that f is convex?

> **Solution:** Line 4. We simplified the original expression to the form of the formal definition's left side in line 3, and line 4 is where we connect it back to the right side.

b) How did we move from line (1) to line (2), i.e. $f\left(A\left((1-t)\vec{x} + t\vec{y}\right) + \vec{b}\right) = f\left((1-t)A\vec{x} + tA\vec{y} + \vec{b}\right)$?

> **Solution:** Distributing A by left multiplying it to the terms in the parentheses.

c) How did we move from line (2) to line (3), i.e. $f\left((1-t)A\vec{x} + tA\vec{y} + \vec{b}\right) = f\left((1-t)(A\vec{x} + \vec{b}) + t(A\vec{y} + \vec{b})\right)$?

> **Solution:** Add $t\vec{b} - t\vec{b}$ to the input expression, that way we can increase the number of terms without changing the value of the expression.

Recall, $g(\vec{x}) = f(A\vec{x} + \vec{b})$, where A is an $n \times n$ matrix and $\vec{x}, \vec{b} \in \mathbb{R}^n$. On the last page, we showed that if f is convex, then g is convex.

Now, let's explore what happens if f is **strictly** convex. Recall, this means that for all (non-equal) $\vec{x}$ and $\vec{y}$ in its domain, and for any $t \in (0, 1)$,

$$f((1-t)\vec{x} + t\vec{y}) < (1-t)f(\vec{x}) + tf(\vec{y})$$

d) Suppose $\text{rank}(A) = n$. Explain why it's impossible for $A\vec{x}+\vec{b} = A\vec{y}+\vec{b}$ for two different vectors $\vec{x}$ and $\vec{y}$.

> **Solution:** If $\text{rank}(A) = n$, then the columns of A are linearly independent, so $A\vec{x}$ and $A\vec{y}$ must be different for any $\vec{x} \neq \vec{y}$.

e) Suppose $\text{rank}(A) < n$. Explain why it's possible for $g(\vec{x}) = g(\vec{y})$ for two different vectors $\vec{x}$ and $\vec{y}$. *Hint: Think about nullsp*(A).

> **Solution:** If $\text{rank}(A) < n$, then A's columns are linearly dependent, so $A\vec{x}$ and $A\vec{y}$ can be the same vector. In that case, $f(A\vec{x} + \vec{b}) = f(A\vec{y} + \vec{b}) \to g(\vec{x}) = g(\vec{y})$.

f) Using the above reasoning, explain why if f is strictly convex, then g is strictly convex if $\text{rank}(A) = n$, and is (not strictly) convex if $\text{rank}(A) < n$.

> **Solution:** We can show this with a proof by cases. In both cases, we'll start from line 4 of the proof on the previous page, but with f being strictly convex.
> Case 1: $\text{rank}(A) = n$
>
> $$\begin{aligned} g((1-t)\vec{x} + t\vec{y}) &< (1-t)f(A\vec{x} + \vec{b}) + tf(A\vec{y} + \vec{b}) \\ &< (1-t)g(\vec{x}) + tg(\vec{y}) \end{aligned}$$
>
> Case 2: $\text{rank}(A) < n$
> We know from part **e)** that it's possible for $g(\vec{x}) = g(\vec{y})$. Using proof by contradiction, assume that g is strictly convex.
>
> $$\begin{aligned} g((1-t)\vec{x} + t\vec{y}) &< (1-t)f(A\vec{x} + \vec{b}) + tf(A\vec{y} + \vec{b}) \\ &< (1-t)g(\vec{x}) + tg(\vec{y}) \\ &< (1-t)g(\vec{x}) + tg(\vec{x}) \\ &< g(\vec{x}) \end{aligned}$$
>
> This is a contradiction, because if $t = 0$, then the left side of the inequality is $g(\vec{x})$, leaving us with $g(\vec{x}) < g(\vec{x})$.

g) What were your thoughts on this type of activity, where we give you a proof and ask you questions about it?

○ Hated it ○ Didn't like it ○ Neutral ○ Liked it ○ Loved it