Topic(s): spectral theorem, quadratic forms, matrix powers

Theorem 24.1.1. Let A be an $n \times n$ matrix and $\mathbf{v}_1, \ldots, \mathbf{v}_r$ a collection of eigenvectors for A with respective eigenvalues $\lambda_1, \ldots, \lambda_r$ (i.e., $A\mathbf{v}_j = \lambda_j \mathbf{v}_j$ for all j).

1. If the r eigenvalues are pairwise different, then the collection of r vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ in $\mathbb{R}^n$ is linearly independent, so $r \leq n$. Thus, an $n \times n$ matrix cannot have more than n different eigenvalues.

2. If A is symmetric and $\lambda_i \neq \lambda_j$, then $\mathbf{v}_i \cdot \mathbf{v}_j = 0$. That is, for a *symmetric* $n \times n$ matrix, eigenvectors for different eigenvalues are orthogonal to each other.

Theorem 24.1.4 (Spectral Theorem). Let A be a *symmetric* $n \times n$ matrix. There is an orthogonal basis $\mathbf{w}_1, \ldots, \mathbf{w}_n$ of $\mathbb{R}^n$ consisting of eigenvectors for A. The corresponding eigenvalues are all of the eigenvalues for A (i.e. if $\mathbf{w}_j$ has eigenvalue λ_j, then any eigenvalue of A equals some λ_j).

The collection of eigenvalues of a square matrix is called its **spectrum**. The spectrum plays key roles in physics and data science!

Even if there is a basis of eigenvectors in $\mathbb{R}^n$ with distinct eigenvalues, beyond the symmetric case, these eigenvectors are *not* all pairwise orthogonal. The Spectral Theorem assures us that such unfortunate situations never occur for symmetric matrices.

Given an orthogonal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ of eigenvectors of a symmetric $n \times n$ matrix A, the Fourier formula is

$$\mathbf{v} = \sum_{i=1}^{n} \left(\frac{\mathbf{v} \cdot \mathbf{w}_i}{\mathbf{w}_i \cdot \mathbf{w}_i} \right) \mathbf{w}_i.$$

In other words, we can express any n-vector as a linear combination of orthogonal eigenvectors! This is really useful. Today, we will use this to solve problems such as determining when $q_A(\mathbf{v}) = \mathbf{v}^\intercal A\mathbf{v}$ is positive or negative (this is very important because this tells us how functions behave near a critical point) and describing the behavior of $A^m\mathbf{v}$ for large values of m.

Example 1. Consider $A = \begin{bmatrix} 13 & -3 \\ -3 & 5 \end{bmatrix}$. Find the eigenvectors and associated eigenvalues of A, and compute $q_A(x, y)$.

We compute $P_A(\lambda) = \lambda^2 - 18\lambda + 56$, and so, $\lambda = 14, 4$.

For $\lambda = 14$, $A - 14I_2 = \begin{bmatrix} -1 & -3 \\ -3 & -9 \end{bmatrix}$, and so, $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ is a 14-eigenvector.

For $\lambda = 4$, $A - 4I_2 = \begin{bmatrix} 9 & -3 \\ -3 & 1 \end{bmatrix}$, and so, $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is a 4-eigenvector.

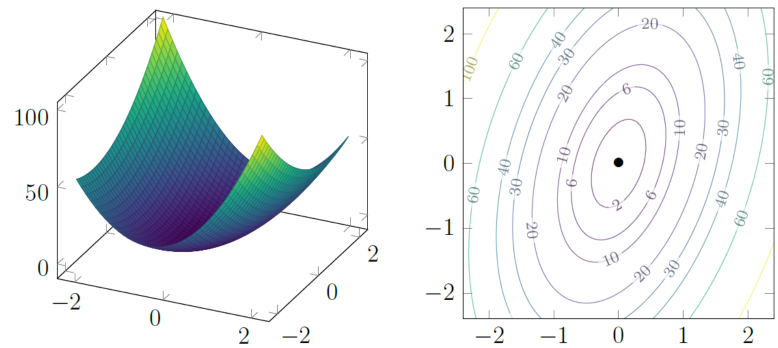
Note that the two eigenvectors are orthogonal!

Also, $q_A(x, y) = 13x^2 - 6xy + 5y^2$.

If A is symmetric, we have seen that the eigenvectors are orthogonal. Hence, if we take the associated *unit* eigenvectors, $\mathbf{w}_1' = \mathbf{w}_1/\|\mathbf{w}_1\|$ and $\mathbf{w}_2' = \mathbf{w}_2/\|\mathbf{w}_2\|$, then $\{\mathbf{w}_1', \mathbf{w}_2'\}$ is an *orthonormal basis*. If we can write $\mathbf{v} = x\mathbf{e}_1 + y\mathbf{e}_2 \in \mathbb{R}^2$ as $\mathbf{v} = t_1'\mathbf{w}_1' + t_2'\mathbf{w}_2'$, then

$$q(x, y) = q_A(\mathbf{v}) = q_A(t_1'\mathbf{w}_1' + t_2'\mathbf{w}_2') = \lambda_1 {t'}_1^2 + \lambda_2 {t'}_2^2.$$

Explicitly, $t_j' = \mathbf{v} \cdot \mathbf{w}_j'$. The amazing thing here is that when we express $\mathbf{v}$ in terms of an orthogonal basis of eigenvectors for the symmetric matrix A, the value $q(\mathbf{v}) = q_A(\mathbf{v})$ at $\mathbf{v}$ contains no cross-terms. In other words, if we rotate our orthonormal reference frame from the standard basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ to the basis $\{\mathbf{w}_1', \mathbf{w}_2'\}$, then the expression for q becomes "$ax_1^2 + bx_2^2$."



For an $n \times n$ symmetric matrix A and the associated quadratic form $q_A(\mathbf{x}) = \mathbf{x}^\intercal A\mathbf{x}$, we say A and q_A are

- **positive-definite** if $q_A(\mathbf{v}) > 0$ for all $\mathbf{v} \neq \mathbf{0}$, and **positive-semidefinite** if $q_A(\mathbf{v}) \geq 0$ for all $\mathbf{v} \neq \mathbf{0}$;
- **negative-definite** if $q_A(\mathbf{v}) < 0$ for all $\mathbf{v} \neq \mathbf{0}$, and **negative-semidefinite** if $q_A(\mathbf{v}) \leq 0$ for all $\mathbf{v} \neq \mathbf{0}$;
- **indefinite** if $q_A(\mathbf{v})$ takes both positive and negative values as $\mathbf{v}$ varies.

Example 2. Determine whether $A = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} -3 & 0 \\ 0 & -4 \end{bmatrix}$ are positive-definite, negative-definite, or indefinite.

$q_A(x, y) = 2x^2 + 5y^2$ is positive-definite.

$q_B(x, y) = -3x^2 - 4y^2$ is negative-definite.

Example 3. Determine whether $A = \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix}$ is positive-definite or positive-semidefinite, or neither.

$q_A(x, y) = x^2 - 6xy + 9y^2 = (x - 3y)^2$ is positive-semidefinite, but not positive-definite.

If a symmetric $n \times n$ matrix A is positive-definite or negative-definite, then A is *invertible*. We can test the definiteness of the quadratic form by noting that if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues associated with the orthogonal eigenvectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$, then for $\mathbf{v} = \sum_{i=1}^n t_i \mathbf{w}_i$,

$$q_A(\mathbf{v}) = \mathbf{v} \cdot (A\mathbf{v}) = \mathbf{v} \cdot \left(\sum_{i=1}^n t_i A \mathbf{w}_i \right) = \mathbf{v} \cdot \left(\sum_{i=1}^n t_i \lambda_i \mathbf{w}_i \right) = \sum_{i=1}^n \lambda_i t_i (\mathbf{v} \cdot \mathbf{w}_i) = \sum_{i=1}^n \lambda_i t_i (t_i \mathbf{w}_i \cdot \mathbf{w}_i).$$

This gives rise to the **diagonalization formula**:

$$q_A(\mathbf{v}) = \sum_{i=1}^n \lambda_i (\mathbf{w}_i \cdot \mathbf{w}_i) t_i^2.$$

Note that all the cross-terms disappear when we write everything in terms of the orthogonal basis $\mathbf{w}_1, \ldots, \mathbf{w}_n$ of eigenvectors of A.

Example 4. The matrix $A = \begin{bmatrix} 16 & -2 & -6 \\ -2 & 19 & -3 \\ -6 & -3 & 27 \end{bmatrix}$ has eigenvectors $\mathbf{w}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{w}_2 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$, and $\mathbf{w}_3 = \begin{bmatrix} -2 \\ -1 \\ 5 \end{bmatrix}$ with eigenvalues 12, 20, and 30, respectively. Determine the nature of q_A.

$q_A(x, y, z) = 16x^2 + 19y^2 + 27z^2 - 4xy - 12xz - 6yz$ is... ???

If $\vec{v} = t_1\vec{w}_1 + t_2\vec{w}_2 + t_3\vec{w}_3$, then

$$q_A(\vec{v}) = \lambda_1(\vec{w}_1 \cdot \vec{w}_1) t_1^2 + \lambda_2(\vec{w}_2 \cdot \vec{w}_2) t_2^2 + \lambda_3(\vec{w}_3 \cdot \vec{w}_3) t_3^2$$
$$= 72t_1^2 + 100t_2^2 + 900t_3^2.$$

by the diagonalization formula. Hence, q_A is positive-definite.

Example 5. Consider $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$ and compute $q_A(x, y)$. Determine the nature of q_A.

Note that $q_A(x, y) = x^2 + 4xy - y^2$ and $P_A(\lambda) = \lambda^2 + \lambda - 6$. Hence, $\lambda = -3, 2$.

- $A + 3I_2 = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \Rightarrow \vec{w}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$
- $A - 2I_2 = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} \Rightarrow \vec{w}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

If $\vec{v} = t_1\vec{w}_1 + t_2\vec{w}_2$, then $q_A(\vec{v}) = -15t_1^2 + 10t_2^2$. Thus, q_A is indefinite.

Proposition 24.2.10. A symmetric $n \times n$ matrix A is

(a) positive-definite when its eigenvalues are all positive.

(b) negative-definite when its eigenvalues are all negative.

(c) indefinite when some eigenvalues are positive and some are negative.

(d) positive-semidefinite but not positive-definite when all eigenvalues are greater than or equal to 0 with some equal to 0.

(e) negative-semidefinite but not negative-definite when all eigenvalues are less than or equal to 0 with some equal to 0.

Example 6. Suppose A is a matrix and $A\mathbf{v} = 51\mathbf{v}$ for some vector $\mathbf{v}$. What is $A^2\mathbf{v}$? $A^3\mathbf{v}$? $A^{51}\mathbf{v}$?

We see that

$$A^2\vec{v} = A(A\vec{v}) = A(51\vec{v}) = 51(A\vec{v}) = (51)^2\vec{v}$$
$$A^3\vec{v} = A(A^2\vec{v}) = (51)^2 A\vec{v} = (51)^3\vec{v}$$
$$A^{51}\vec{v} = (51)^{51}\vec{v}.$$

In general, if $\mathbf{v}$ is an eigenvector of an $n \times n$ matrix A with eigenvalue λ, then

$\lambda \neq 0$ b/c A inv. $$A^r\mathbf{v} = \lambda^r\mathbf{v}.$$

Note that if A is invertible with $A\mathbf{v} = \lambda\mathbf{v}$, then $\mathbf{v}$ is also an eigenvector for A^{-1} with eigenvalue λ^{-1}.

$$\vec{v} = A^{-1}A\vec{v} = A^{-1}(\lambda\vec{v}) = \lambda(A^{-1}\vec{v}) \Rightarrow A^{-1}\vec{v} = \lambda^{-1}\vec{v}.$$

Example 7. Recall, from Example 5, that $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$ has eigenvectors $\mathbf{w}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{w}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ with eigenvalues 2 and -3, respectively. Compute $A^4 \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

Note that $\begin{bmatrix} 3 \\ 4 \end{bmatrix} = 2\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Hence,

$$A^4\begin{bmatrix} 3 \\ 4 \end{bmatrix} = 2A^4\begin{bmatrix} 2 \\ 1 \end{bmatrix} + A^4\begin{bmatrix} -1 \\ 2 \end{bmatrix} = 2 \cdot 2^4\begin{bmatrix} 2 \\ 1 \end{bmatrix} + (-3)^4\begin{bmatrix} -1 \\ 2 \end{bmatrix}$$
$$= \begin{bmatrix} 64 \\ 32 \end{bmatrix} + \begin{bmatrix} -81 \\ 162 \end{bmatrix} = \boxed{\begin{bmatrix} -17 \\ 194 \end{bmatrix}}.$$

Solve the equation $A\mathbf{x} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

Similarly as above,

$$\vec{x} = A^{-1}\begin{bmatrix} 3 \\ 4 \end{bmatrix} = 2A^{-1}\begin{bmatrix} 2 \\ 1 \end{bmatrix} + A^{-1}\begin{bmatrix} -1 \\ 2 \end{bmatrix} = 2 \cdot \frac{1}{2}\begin{bmatrix} 2 \\ 1 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} -1 \\ 2 \end{bmatrix}$$
$$= \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix} = \boxed{\begin{bmatrix} \frac{7}{3} \\ \frac{1}{3} \end{bmatrix}}.$$

Note that we did not need A^{-1} to solve the system!

Recall again that $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$ has eigenvectors $\mathbf{w}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{w}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ with respective eigenvalues 2 and -3.
It turns out that there is an exact formula for A^m:

$$A^m = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}^m = \frac{1}{5} \begin{bmatrix} 4 \cdot 2^m + (-3)^m & 2 \cdot (2^m - (-3)^m) \\ 2 \cdot (2^m - (-3)^m) & 2^m + 4(-3)^m \end{bmatrix}.$$

Note that the bases of the powers, 2 and -3, are the eigenvalues of A. The importance of this is that *the eigenvalue with the largest magnitude controls the behavior for large m*. If we put out a factor of $(-3)^m$, we get

$$A^m = \frac{(-3)^m}{5} \begin{bmatrix} 4 \cdot (2/(-3))^m + 1 & 2((2/(-3))^m - 1) \\ 2 \cdot ((2/(-3))^m - 1) & (2/(-3))^m + 4 \end{bmatrix}.$$

The highlighted quantities go to 0 as $m \to \infty$.

Hence, we can approximate this by

$$A^m \approx (-3)^m \begin{bmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{bmatrix}$$

The matrix on the right is equal to $\dfrac{1}{\mathbf{w} \cdot \mathbf{w}} \mathbf{w}\mathbf{w}^\intercal$, where $\mathbf{w} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ is an eigenvector for the eigenvalue -3.

Theorem 24.4.1 (Interpretation of Spectral Theorem via a matrix decomposition). Let A be a symmetric $n \times n$ matrix with orthogonal eigenvectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$, having corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. Let W be the $n \times n$ matrix whose columns are the respective unit eigenvectors

$$\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \ldots, \frac{\mathbf{w}_n}{\|\mathbf{w}_n\|}.$$

Then, $W^\intercal = W^{-1}$ (i.e., W is an *orthogonal* matrix), and

$$A = WDW^\intercal = WDW^{-1},$$

where $D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$ with the diagonal entries being the corresponding eigenvalues.

Thus,

$$\begin{aligned} A^m = (WDW^\intercal)^m &= (WDW^{-1})^m \\ &= (WD\cancel{W^{-1}})(\cancel{W}D\cancel{W^{-1}})(\cancel{W}D\cancel{W^{-1}}) \cdots (\cancel{W}DW^{-1}) \\ &= WD^mW^{-1} = WD^mW^\intercal. \end{aligned}$$

Proposition 24.4.2. For a symmetric $n \times n$ matrix A, if there is an eigenvalue λ whose absolute value exceeds that of all other eigenvalues (called a **dominant eigenvalue**) and if the solutions to $A\mathbf{x} = \lambda\mathbf{x}$ constitute a line (as happens whenever there are n different eigenvalues), then for large m, we have

$$A^m \approx \left(\frac{\lambda^m}{\mathbf{w} \cdot \mathbf{w}} \right) \mathbf{w}\mathbf{w}^\intercal,$$

for any eigenvector $\mathbf{w}$ of A with eigenvalue λ.

Example 8. In this exercise, we use eigenvalues to understand the behavior of the Fibonacci sequence $\{f_1, f_2, f_3, \ldots\}$ that is defined by $f_1 = 1, f_2 = 1$, and $f_n = f_{n-1} + f_{n-2}$ for $n > 2$ (so the sequence begins $1, 1, 2, 3, 5, \ldots$. Consider the matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, which has the property that $A \begin{bmatrix} f_{n-2} \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} f_{n-1} \\ f_n \end{bmatrix}$ for $n \geq 3$.

(a) Check that $\lambda_1 = \dfrac{1+\sqrt{5}}{2}$ and $\lambda_2 = \dfrac{1-\sqrt{5}}{2}$ are the two eigenvalues of A. (λ_1 is often denoted ϕ and called the **golden ratio.**)

The characteristic polynomial of A is $P_A(\lambda) = \lambda^2 - \lambda - 1$. Hence, the eigenvalues of A are

$$\lambda = \frac{1 \pm \sqrt{1+4}}{2} = \frac{1 \pm \sqrt{5}}{2}.$$

(b) Explain why $A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix}$ for $n \geq 1$.

We see that $A^1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$ and $A^2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} f_2 \\ f_1 + f_2 \end{bmatrix} = \begin{bmatrix} f_2 \\ f_3 \end{bmatrix}$.

Thus, $A \begin{bmatrix} f_{n-1} \\ f_n \end{bmatrix} = \begin{bmatrix} f_n \\ f_{n-1} + f_n \end{bmatrix} = \begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix}$, and so, $A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} = A^{n-1} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \cdots = \begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix}$.

(c) Find an eigenvector for each of the two eigenvalues λ_1 and λ_2, and write $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ as a linear combination of the two eigenvectors.

$$A - \lambda_1 I_2 = \begin{bmatrix} \frac{-1-\sqrt{5}}{2} & 1 \\ 1 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \Rightarrow \vec{w}_1 = \begin{bmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ \varphi \end{bmatrix}$$

$$A - \lambda_2 I_2 = \begin{bmatrix} \frac{-1+\sqrt{5}}{2} & 1 \\ 1 & \frac{1+\sqrt{5}}{2} \end{bmatrix} \Rightarrow \vec{w}_2 = \begin{bmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ 1-\varphi \end{bmatrix}$$

(d) Using parts (b) and (c), give a closed form expression for f_n involving powers of the eigenvalues.

Since $\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \vec{w}_1 - \frac{1}{\sqrt{5}} \vec{w}_2$, $\begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix} = A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^n \vec{w}_1 - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2} \right)^n \vec{w}_2$. So,

$$\boxed{f_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right]}.$$

(e) Using the fact that $|\lambda_2| < 1$, find an approximation for f_n in terms of $\phi = \lambda_1$ when n is large.

By Prop. 24.4.2,

$$A^n \approx \left(\frac{\lambda^n}{\vec{w}_1 \cdot \vec{w}_1} \right) \vec{w}_1 \vec{w}_1^T = \frac{\varphi^n}{1+\varphi^2} \begin{bmatrix} 1 \\ \varphi \end{bmatrix} \begin{bmatrix} 1 & \varphi \end{bmatrix} = \frac{\varphi^n}{1+\varphi^2} \begin{bmatrix} 1 & \varphi \\ \varphi & \varphi^2 \end{bmatrix}.$$

Thus, $A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} \approx \frac{\varphi^n}{1+\varphi^2} \begin{bmatrix} \varphi \\ \varphi^2 \end{bmatrix}$, and so, $\boxed{f_n \approx \frac{\varphi^{n+1}}{1+\varphi^2}}$. This is actually a pretty good estimate for $n \geq 8$!