02-680 Module 10

Essentials of Mathematics and Statistics

Aidan Jan

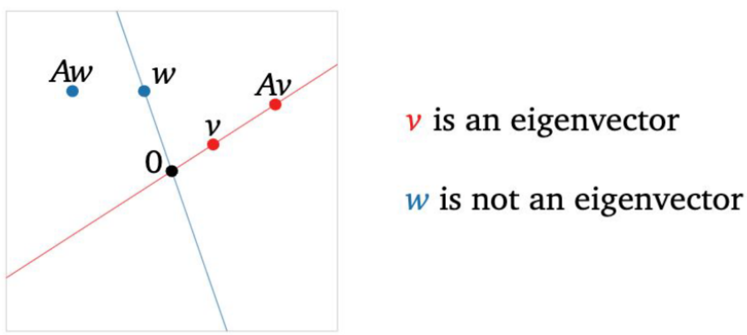
November 11, 2025

Eigenvalues and Eigenvectors

For a matrix $A \in \mathbb{R}^{n \times n}$, we define

- an **eigenvalue** $\lambda \in \mathbb{R}$, and
- an **eigenvector** $x \in \mathbb{R}^n \setminus 0$ such that
- $Ax = \lambda x$

To say that $Av = \lambda v$ means that Av and λv are **collinear** with the origin. So, an eigenvector of A is a nonzero vector v such that Av and v lie on the same line through the origin. In this case, Av is a scalar multiple of v; the eigenvalue is the scaling factor.



For example, let

$$A = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \lambda = 5$$

Then, we can calculate both Ax and λx.

$$Ax = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \end{bmatrix}$$

$$\lambda x = 5 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \end{bmatrix}$$

We can verify eigenvectors by doing this multiplication and seeing if Aw is a scalar multiple of w.

Finding Eigenelements

Suppose we have a matrix we want to find the eigenelements for. Consider $A = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}$ as an example. The eigenvectors must fit the form: $Av = \lambda v$, so we can write the equation:

$$\begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Now we can simplify:

$$= \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$= \left(\begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This will only have **nontrivial** solutions when this left matrix is **singular**. In other words, the determinant must be equal to zero.

The general process to solve for eigenvectors and eigenvalues is as follows:

1. Subtract lambda from the main diagonal of matrix A.
2. Find the determinant of the resulting matrix, and solve for all the possible values of λ.
   - The polynomial we get in terms of λ while solving for them is called the **characteristic equation** (or characteristic polynomial) of A.
3. Plug in the solutions of λ and solve for the vectors.

For example, let $A = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}$. Then, we first find the determinant with the λ's.

$$\det\left(\begin{bmatrix} 3-\lambda & 4 \\ 2 & 1-\lambda \end{bmatrix} \right)$$

When we simplify, we get the characteristic equation $(3-\lambda)(1-\lambda)-(4)(2) = 0$, which simplifies to $\lambda^2 - 4\lambda - 5$.

Solving, we get $\lambda = -1, 5$. Now we can plug the values in:

$$\begin{bmatrix} 3-(-1) & 4 \\ 2 & 1-(-1) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 4 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 3-(5) & 4 \\ 2 & 1-(5) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} -2 & 4 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Solving for v_1 and v_2 for both cases gives the eigenvectors $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Linear Independence

The eigenvectors $x_1, x_2, \ldots, x_n$ of a matrix $A \in \mathbb{R}^{n \times n}$ with n **distinct** eigenvalues $\lambda_1 \neq \lambda_2 \neq \ldots \neq \lambda_n$ are linearly independent.

- the eigenvectors form a **basis** in $\mathbb{R}^n$
- By the spectral theorem, **if** $A \in \mathbb{R}^{n \times n}$ is **symmetric**, there exists an **orthonormal basis** of the corresponding vector space V consisting of eigenvectors of A, and each eigenvector is real.
  - $A = V \wedge V^T$ where $V \in \mathbb{R}^{n \times n}$ is an orthonormal matrix of eigenvectors of A, and $\wedge \in \mathbb{R}^{n \times n}$ is a real diagonal matrix of eigenvalues.
- If a matrix is symmetric (e.g., $A = A^T$), then all eigenvalues are real and all eigenvectors are orthonormal.
- For any matrix $B \in \mathbb{R}^{m \times n}$, $B^T B$ is a square, symmetric, **positive semidefinite**.
- A matrix M is "positive semi-definite" if and only if $x^T M x \geq 0 \quad \forall x \in \mathbb{R}^n$.

Characteristic Polynomial

Let A be an $n \times n$ matrix. The characteristic polynomial of A is the function $f(\lambda)$ given by

$$f(\lambda) = \det(A - \lambda I_n)$$

Finding the characteristic polynomial means computing the determinant of the matrix $A - \lambda I_n$, whose entries contain the unknown λ.

- Basically, subtract λ from each element of the main diagonal, and find the determinant.

The characteristic polynomial can be used to compute **eigenvalues**. Eigenvalues are roots of the characteristic polynomial.

- Setting the characteristic polynomial to zero (e.g., $\det(A - \lambda I_n) = 0$) gives the **characteristic equation**. This is the equation that you solve to find the eigenvalues of matrix A.

Note that

- the **determinant** is the product of all the eigenvalues.
- the **rank** of A is equal to the number of non-zero eigenvalues.
- If A is **nonsingular**, then $1/\lambda i$ is an eigenvalue of A^{-1} with the same associated original eigenvector.

Triangular Matrices

Triangular matrices (those where all of the entries above/below the diagonal ard 0) make computing both a determinant and eigenelements **easier**.

- The determinant becomes the product of the entries on the diagonal.
- The characteristic equation of the matrix is then the product of the diagonals in $A - \lambda I_n$, and thus the eigenvalues are the original values in the diagonal of A.

A matrix $T \in \mathbb{R}^{n \times n}$ is **triangular** if all values on one side of the diagonal are 0. There are two types: upper-triangular and lower-triangular.

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

upper-triangular

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

lower-triangular

Additional Terminology

- Two matrices A and $A' \in \mathbb{R}^{n \times n}$ are **equivalent** if there exists nonsingular matrices

$$S \in \mathbb{R}^{n \times n}, T \in \mathbb{R}^{m \times m} \text{ such that } A' = T^{-1}AS$$

- Two matrices B and $B' \in \mathbb{R}^{n \times n}$ are **similar** if there exists a nonsingular matrix $P \in \mathbb{R}^{n \times n}$ such that $B' = P^{-1}BP$.
- Note that all similar matrix pairs are equivalent, but not all equivalent matrix pairs are similar. Or, "similar" $\Rightarrow$ "equivalent".

Eigendecomposition and Diagonalization

Diagonalization

A diagonal matrix $D \in \mathbb{R}^{n \times n}$ has the value 0 in all off-diagonal locations:

$$\begin{bmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_n \end{bmatrix}$$

A matrix $A \in \mathbb{R}^{n \times n}$ is **diagonalizable** if it is similar to a diagonal matrix $D \in \mathbb{R}^{n \times n}$.

- If there exists a nonsingular $P \in \mathbb{R}^{n \times n}$ such that $D = P^{-1}AP$.

Let $A \in \mathbb{R}^{n \times n}$, let $\lambda_1, \ldots, \lambda_n$ be scalars, and let $p_1, \ldots, p_n$ be vectors in $\mathbb{R}^n$. Define

- $P = [p_1, \ldots, p_n]$ (here the p_i's are the columns of a matrix) and
- $D \in \mathbb{R}^{n \times n}$ be a diagonal matrix with diagonal values $\lambda_1, \ldots, \lambda_n$.

Then, $AP = PD$ if and only if

- $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A and
- $p_1, \ldots, p_n$ are the corresponding eigenvectors.

If we create a new matrix $X \in \mathbb{R}^{n \times n}$ where each column is one of the eigenvectors of A, then create $\wedge \in \mathbb{R}$

- $n \times n$ which contains the eigenvalues on the diagonal (i.e., $\wedge = \langle \lambda_1, \lambda_2, \ldots, \lambda_n \rangle I_n$)
- it turns out that $AX = X\wedge$ because it satisfies all of the eigenelement sets simultaneously.

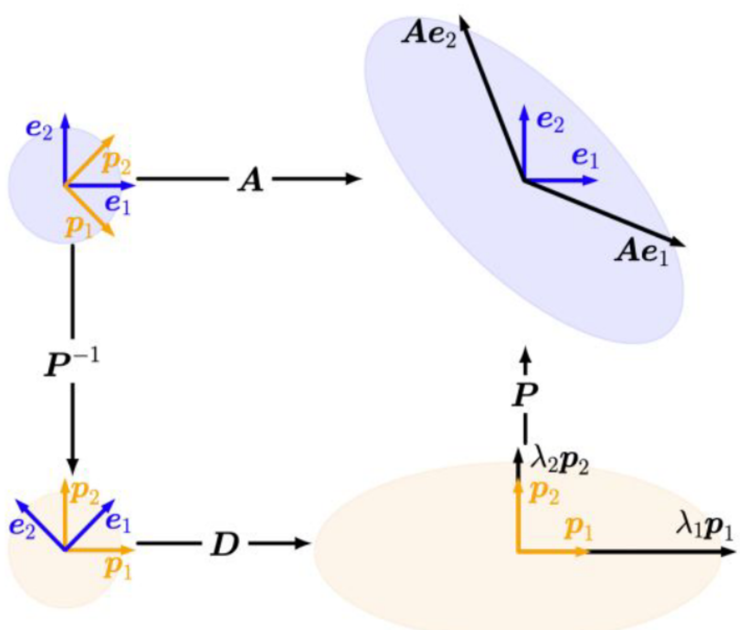
If the eigenvectors are linearly independent then X is invertible. If X is invertible,

$$A = X \wedge X^{-1}$$

We call A diagonalizable if it can be rewritten this way. This is sometimes also called an eigendecomposition.

Eigendecompositions

A square matrix $A \in \mathbb{R}^{n \times n}$ can be factored into $A = PDP^{-1}$ where $P \in \mathbb{R}^{n \times n}$ and D is a diagonal matrix whose entries are the eigenvalues of A, if and only if the eigenvalues form a basis of $\mathbb{R}^n$.



Singular Value Decomposition (SVD)

The form $A = X \wedge X^{-1}$ is similar to what is known as the Singular Value Decomposition (SVD) of a **rectangular** matrix $R \in \mathbb{R}^{n \times n}$ as

$$R = USV$$

(sometimes Σ is used in place of S but this is a reserved character in this class) where $U \in \mathbb{R}^{n \times n}$, $V \in \mathbb{R}^{m \times m}$ and

$$S \in \left\{ \widehat{\operatorname{diag}}_{n \times m}(x) | x \in \mathbb{R}^{\min(n,m)} \right\} \subset \mathbb{R}^{n \times m}$$

Here, the

$$\widehat{\operatorname{diag}}_{n \times m}(\langle x_1, x_2, \ldots, x_m \rangle) = \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_m \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \text{ or } \begin{bmatrix} x_1 & 0 & \cdots & 0 & 0 & \cdots \\ 0 & x_2 & \cdots & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & x_m & 0 & \cdots \end{bmatrix}$$

It produces the left matrix when $n > m$ and the right when $m > n$. That is, it creates a diagonal matrix padded with 0 rows or columns to make it the correct size.

Meaning of SVD

Let's assume you are in the situation where you have n users who reviewed m movies and those ratings are in matrix $E \in \mathbb{R}^{n \times m}$. If you can find the SVD of the matrix such that $E = FGH$ such that

- $F \in \mathbb{R}^{n \times m}$, $H \in \gtrdot \times \gtrdot$ and G is a diagonal m-dimension square matrix each of these represents something about your set.
- F and H show commonalities between users or movies respectively, and G is a connection matrix.

Another example would be gene correlations, assume you have multiple assays of a gene's expressions in various conditions. If you can decompose that matrix using the theory above you'd have a correlation matrix of genes, correlation of assays, and a relationship matrix. The details about finding this are beyond the scope of this course.

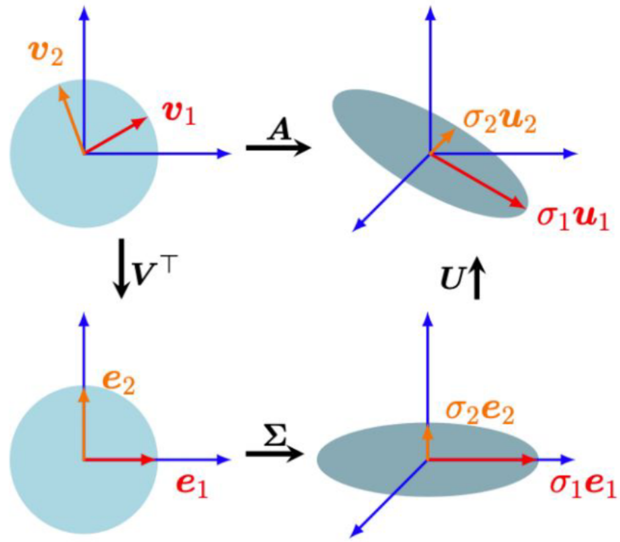
SVD Theorem

Let $A \in \mathbb{R}^{m \times n}$ be a rectangular matrix of rank $r \in [0, \min(m, n)]$. The SVD of A is a decomposition of the form, where

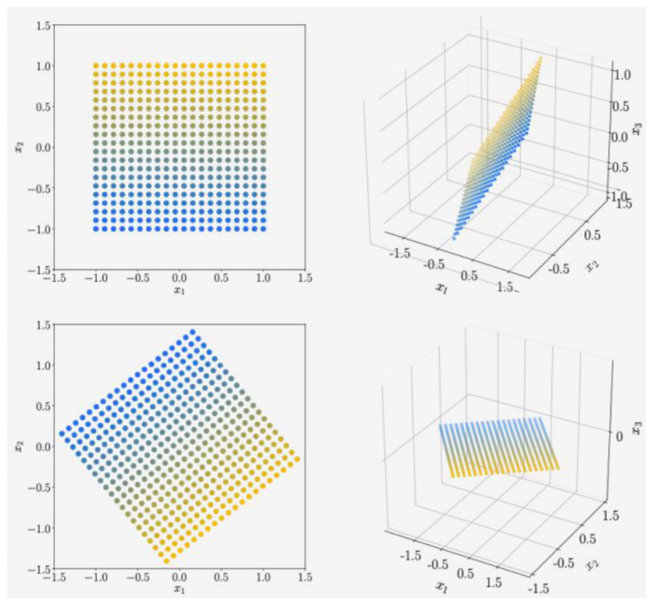
- $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are orthogonal matrices and
- $\Sigma \in \mathbb{R}^{m \times n}$ is a matrix with $\Sigma_{ii} = \sigma_i$ and all other values equal 0.

It is worth noting that the columns of U $(u_1, u_2, \ldots, u_m)$ and V $(v_1, v_2, \ldots, v_n)$ are called the left- and right-singular vectors, respectively. Additionally, the σ_i's are called **singular values** and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$

Intuition



$$\boldsymbol{A} = \begin{bmatrix} 1 & -0.8 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \boldsymbol{U\Sigma V}^\top$$
$$= \begin{bmatrix} -0.79 & 0 & -0.62 \\ 0.38 & -0.78 & -0.49 \\ -0.48 & -0.62 & 0.62 \end{bmatrix} \begin{bmatrix} 1.62 & 0 \\ 0 & 1.0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -0.78 & 0.62 \\ -0.62 & -0.78 \end{bmatrix}$$



Constructing SVD

Recall that in Eigendecomposition for symmetric positive semidefinite matrices, we had $S = S^T = PDP^T$. And for SVD we have $S = U\Sigma V^T$.

- if $V = U = P$ and $\Sigma = D$, then they are the same.

Recall that we can make a square positive semidefinite from any matrix multiplied by its transpose: $A^T A$. Now, we need to find U to be the orthonormal matrix paired with V^T such that:

$$u_i = \frac{Av_i}{\|Av_i\|} = \frac{1}{\sqrt{\lambda_i}} Av_i = \frac{1}{\sigma_i} Av_i \Longleftrightarrow Av_i = \sigma_i u_i \quad i = 1, \dots, \min(m, n)$$

Note that

$$(Av_i)^T (Av_i) = v_i^T (A^T A) v_i = v_i^T (\lambda_i v_i) = \lambda_i (v_i^T v_i) = \lambda_i$$