# MFML SEC9 : Lecture 4

This guide delves into the heart of matrices, exploring how they transform space and how we can deconstruct them into their fundamental components. We begin with determinants—a key to understanding a matrix's impact on volume and solvability—before unlocking the secrets of eigenvalues and eigenvectors, the intrinsic directions and scaling factors of linear transformations. These concepts are pivotal for powerful techniques in data analysis, physics, engineering, and machine learning.

# Contents

# 1 Determinants: The Measure of a Transformation's Power

Before we explore the special directions (eigenvectors) of a matrix, we need a way to quantify a fundamental aspect of the transformation it represents: how much does it scale space? And does it collapse space into a lower dimension? This is the role of the **determinant**.

**Intuition: What Does a Determinant Tell Us?**

Imagine a matrix transforming a simple unit square (in 2D) or a unit cube (in 3D).

- The **absolute value of the determinant** tells you the factor by which the area (in 2D) or volume (in 3D) of that initial shape changes after the transformation. A determinant of 2 means the area/volume doubles. A determinant of 0.5 means it halves.
- A **determinant of zero** is a critical case! It means the transformation squashes the space into a lower dimension (e.g., a 2D square is flattened into a line or a point; a 3D cube into a plane, line, or point). When this happens, the matrix is **singular** (not invertible) – you can't undo the transformation because information is lost.
- The **sign of the determinant** (in 2D and 3D) indicates whether the transformation preserves or reverses the orientation of space. A positive determinant means the orientation is preserved (e.g., a counter-clockwise arrangement of vectors remains counter-clockwise). A negative determinant means the orientation is flipped (like looking at a mirror image).

The determinant encapsulates the "oomph" of a matrix in a single number.

## 1.1 Defining Determinants: Minors and Cofactors

While the geometric intuition is powerful, we need a formal way to calculate determinants. For a general $n \times n$ matrix $A$, the determinant, denoted $\det(A)$ or $|A|$, can be defined recursively using minors and cofactors.

**Definition 1. Minor and Cofactor**

Let $A$ be an $n \times n$ matrix.

- The **minor** $M_{ij}$ associated with the element $a_{ij}$ (the element in the $i$-th row and $j$-th column) is the determinant of the $(n-1) \times (n-1)$ submatrix formed by deleting the $i$-th row and $j$-th column of $A$.

- The **cofactor** $C_{ij}$ associated with $a_{ij}$ is given by $C_{ij} = (-1)^{i+j}M_{ij}$. The $(-1)^{i+j}$ term creates a "checkerboard" pattern of signs.

The determinant is then calculated by the **cofactor expansion** (or Laplace expansion) along any row or any column.

**Definition 2. Determinant by Cofactor Expansion**

For an $n \times n$ matrix $A$:

- Expansion along row $i$: $\det(A) = \sum_{j=1}^{n} a_{ij}C_{ij} = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$
- Expansion along column $j$: $\det(A) = \sum_{i=1}^{n} a_{ij}C_{ij} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$

The result is the same regardless of which row or column you choose.

**Base Cases for the Recursion:**

- For a $1 \times 1$ matrix $A = [a]$, $\det(A) = a$.
- For a $2 \times 2$ matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the determinant is $\det(A) = ad - bc$. (You can verify this using cofactor expansion: $\det(A) = a \cdot C_{11} + b \cdot C_{12} = a \cdot (-1)^{1+1}\det([d]) + b \cdot (-1)^{1+2}\det([c]) = ad - bc$.)

**Example 1. Calculating a $3 \times 3$ Determinant**

Let $A = \begin{pmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{pmatrix}$. Let's calculate $\det(A)$ by expanding along the first row.

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

The elements are $a_{11} = 3, a_{12} = 1, a_{13} = 0$. Now, the cofactors:

- $C_{11} = (-1)^{1+1}M_{11} = 1 \cdot \det \begin{pmatrix} -4 & 3 \\ 4 & -2 \end{pmatrix} = ((-4)(-2) - (3)(4)) = 8 - 12 = -4$.
- $C_{12} = (-1)^{1+2}M_{12} = -1 \cdot \det \begin{pmatrix} -2 & 3 \\ 5 & -2 \end{pmatrix} = -1 \cdot ((-2)(-2) - (3)(5)) = -1 \cdot (4 - 15) = -1(-11) = 11$.
- $C_{13} = (-1)^{1+3}M_{13} = 1 \cdot \det \begin{pmatrix} -2 & -4 \\ 5 & 4 \end{pmatrix} = ((-2)(4) - (-4)(5)) = -8 - (-20) = -8 + 20 = 12$.

So, $\det(A) = 3(-4) + 1(11) + 0(12) = -12 + 11 + 0 = -1$.

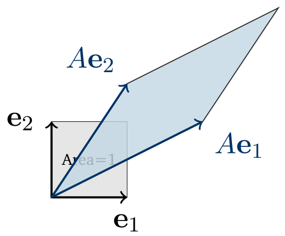
**Choosing the Best Row/Column for Expansion**

To simplify calculations, always choose a row or column with the most zeros. Each zero term in the expansion saves you from calculating a cofactor!

## 1.2 Geometric Interpretation of Determinants

**Determinants as Area Scaling Factors**

**2D: Area of Parallelogram**




*The absolute value of the determinant,* $|\det(A)|$*, measures how much a transformation* $A$ *scales area (as shown in the 2D example) or, more generally,* $n$*-dimensional volume. For the 2D example with* $A = \begin{pmatrix} 2 & 1 \\ 1 & 1.5 \end{pmatrix}$ *(where* $A\mathbf{e}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ *and* $A\mathbf{e}_2 = \begin{pmatrix} 1 \\ 1.5 \end{pmatrix}$*),* $\det(A) = 2(1.5) - 1(1) = 2$*, so the area is doubled. A zero determinant means the transformation collapses space into a lower dimension.*

**Key Geometric Insights:**

- **Columns as Transformed Basis Vectors:** The columns of an $n \times n$ matrix $A$ are the images of the standard basis vectors ($\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$) under the transformation $A$. That is, the $j$-th column of $A$ is $A\mathbf{e}_j$.
- **Determinant and Volume:** The absolute value of $\det(A)$ is the volume of the $n$-dimensional parallelepiped whose edges are the column vectors of $A$ (or, equivalently, the images of the standard basis vectors).
- **Linear Dependence and Zero Determinant:** If the column vectors of $A$ are linearly dependent, they cannot span an $n$-dimensional space. The parallelepiped they form will be "flat" or "squashed" into a lower dimension, having zero $n$-dimensional volume. This is precisely when $\det(A) = 0$. This also implies that the rows are linearly dependent.
- **Invertibility:** A matrix $A$ is invertible if and only if $\det(A) \neq 0$. This makes perfect sense: if $\det(A) = 0$, the transformation collapses space, and there's no way to uniquely reverse it. If $\det(A) \neq 0$, the transformation doesn't lose dimensionality,

so it can be undone.

- **Properties:**
  - $\det(AB) = \det(A)\det(B)$ (The scaling factor of a composite transformation is the product of individual scaling factors).
  - $\det(A^T) = \det(A)$ (Transposing doesn't change the volume scaling).
  - $\det(A^{-1}) = 1/\det(A)$ (The inverse transformation scales volume by the reciprocal factor).
  - If $A$ is triangular, $\det(A)$ is the product of its diagonal entries. (This is a huge computational shortcut!)

This understanding of determinants as a measure of how a transformation scales space and an indicator of invertibility is crucial for grasping the concept of eigenvalues.

# 2 The Essence of a Matrix: Eigenvalues and Eigenvectors

Every square matrix $A$ describes a linear transformation. It takes a vector $\mathbf{x}$ and maps it to a new vector $A\mathbf{x}$. This transformation can stretch, shrink, rotate, shear, or reflect space. But within this complex dance, are there any vectors that, after the transformation, simply get scaled without changing their original direction (span)?

Yes. These are the **eigenvectors**, and they reveal the true "axes of transformation" or the "characteristic directions" of the matrix. The amount by which they are scaled is given by their corresponding **eigenvalues**.

**Intuition: Finding the "Grain" or "Axes" of the Transformation**

Imagine a transformation as a dynamic process acting on space:

- **Spinning Globe:** If you spin a globe, vectors along the axis of rotation don't change their direction; they are eigenvectors. If the globe isn't also expanding or contracting, their eigenvalue is 1. Other vectors (e.g., pointing from the center to a city not on the poles) will change direction.
- **Stretching Dough:** If you stretch a piece of dough uniformly in one direction and compress it in another, vectors aligned with the stretch direction are eigenvectors (eigenvalue $> 1$), and vectors aligned with the compression direction are also eigenvectors (eigenvalue between 0 and 1). Vectors in other directions will have their direction changed, pointing more towards the stretch direction.
- **Shearing a Deck of Cards:** When you push the top of a deck of cards, the horizontal vectors within the cards that are parallel to the push remain horizontal (eigenvectors, eigenvalue 1). Vertical vectors get tilted (not eigenvectors).

An **eigenvector** $\mathbf{x}$ is a special non-zero vector whose direction is invariant under the transformation $A$. It only gets scaled by a factor $\lambda$, its **eigenvalue**. A positive eigenvalue means scaling in the same direction. A negative eigenvalue means scaling in the opposite direction (a flip plus scaling).

This core idea is captured in the defining equation:

**Definition 3. Eigenvector and Eigenvalue**

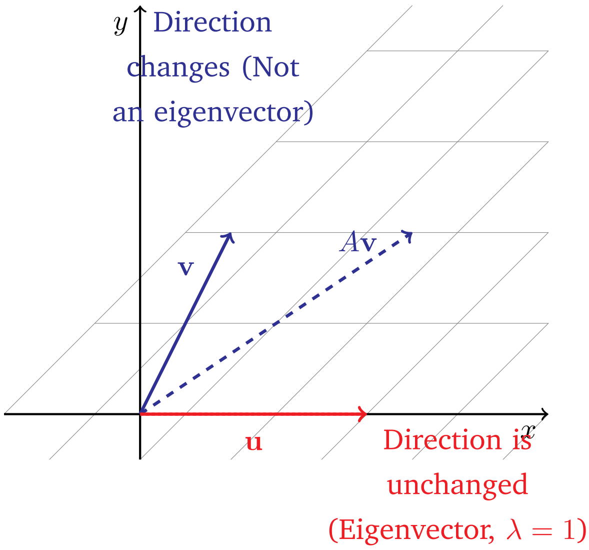
An **eigenvector** of an $n \times n$ matrix $A$ is a **nonzero** vector $\mathbf{x}$ such that for some scalar $\lambda$:

$$A\mathbf{x} = \lambda\mathbf{x}$$

The scalar $\lambda$ is called the **eigenvalue** corresponding to the eigenvector $\mathbf{x}$.

*Note: The zero vector $\mathbf{0}$ is trivially mapped to $\mathbf{0}$ by any matrix ($A\mathbf{0} = \mathbf{0}$), so it satisfies $A\mathbf{x} = \lambda\mathbf{x}$ for any $\lambda$. However, it doesn't define a direction and is explicitly excluded from being an eigenvector.*

**A Visual Interpretation of Shear**




*A shear transformation $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. The red vector $\mathbf{u}$ (along the x-axis) is an eigenvector because its direction is unchanged ($A\mathbf{u} = \mathbf{u}$, so $\lambda = 1$). The blue vector $\mathbf{v}$ is not an eigenvector; its direction is altered by the shear.*

**Example 2. Verifying Eigenvectors**

Let $A = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}$, $\mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. Are $\mathbf{u}$ and $\mathbf{v}$ eigenvectors of $A$?

**1. Test vector $\mathbf{u}$:** We multiply $A$ by $\mathbf{u}$:

$$A\mathbf{u} = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix} \begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} 1(6) + 6(-5) \\ 5(6) + 2(-5) \end{bmatrix} = \begin{bmatrix} 6 - 30 \\ 30 - 10 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix}$$

Is the result $A\mathbf{u}$ a scalar multiple of $\mathbf{u}$? We can see that $\begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4 \begin{bmatrix} 6 \\ -5 \end{bmatrix}$. Since $A\mathbf{u} = -4\mathbf{u}$, the equation $A\mathbf{u} = \lambda\mathbf{u}$ holds with $\lambda = -4$. So, $\mathbf{u}$ **is** an eigenvector of $A$ with eigenvalue $\lambda = -4$.

**2. Test vector $\mathbf{v}$:** We multiply $A$ by $\mathbf{v}$:

$$A\mathbf{v} = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1(3) + 6(-2) \\ 5(3) + 2(-2) \end{bmatrix} = \begin{bmatrix} 3 - 12 \\ 15 - 4 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix}$$

Is $\begin{bmatrix} -9 \\ 11 \end{bmatrix}$ a scalar multiple of $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$? If it were, there would be a $\lambda$ such that $-9 = \lambda(3)$ and $11 = \lambda(-2)$. The first equation implies $\lambda = -3$. The second implies $\lambda = -11/2$. Since these $\lambda$ values are different, there is no single scalar $\lambda$ that satisfies the condition. Therefore, $\mathbf{v}$ is **not** an eigenvector of $A$.

# 3 How to Find Eigenvalues and Eigenvectors

We need a systematic way to find these characteristic directions and scaling factors. The method cleverly transforms the problem $A\mathbf{x} = \lambda\mathbf{x}$ into finding the roots of a polynomial.

## 3.1 The Characteristic Equation: The Key to the Kingdom

We start with the defining equation: $A\mathbf{x} = \lambda\mathbf{x}$. Our goal is to find $\lambda$ and the corresponding non-zero $\mathbf{x}$. Let's rearrange the equation:

$$\begin{aligned} A\mathbf{x} - \lambda\mathbf{x} &= \mathbf{0} \\ A\mathbf{x} - \lambda I\mathbf{x} &= \mathbf{0} \quad \text{(where } I \text{ is the identity matrix, same size as } A\text{)} \\ (A - \lambda I)\mathbf{x} &= \mathbf{0} \end{aligned}$$

This equation, $(A - \lambda I)\mathbf{x} = \mathbf{0}$, is a homogeneous system of linear equations. It always has the trivial solution $\mathbf{x} = \mathbf{0}$. But by definition, eigenvectors must be **non-zero**. So, we are looking for non-trivial (non-zero) solutions for $\mathbf{x}$.

A non-trivial solution $\mathbf{x}$ exists if and only if the matrix $(A - \lambda I)$ is singular (i.e., not invertible). And a matrix is singular if and only if its determinant is zero. This is the crucial insight that links eigenvalues to determinants!

**Why the Determinant Must Be Zero for Eigenvectors**

The equation $M\mathbf{x} = \mathbf{0}$ (where $M = A - \lambda I$) asks: "Is there a non-zero vector $\mathbf{x}$ that the matrix $M$ maps to the zero vector?"

- If $M$ is invertible (i.e., $\det(M) \neq 0$), it means $M$ doesn't collapse space. The only vector it can map to $\mathbf{0}$ is $\mathbf{0}$ itself ($\mathbf{x} = M^{-1}\mathbf{0} = \mathbf{0}$). This is the trivial solution, which we don't want for eigenvectors.
- If $M$ is **not** invertible (i.e., $\det(M) = 0$), it means the transformation $M$ "collapses" space into a lower dimension. For example, a 2D matrix $M$ with $\det(M) = 0$ might map the entire plane onto a single line (or even a point). This "collapse" means that an entire line or plane of vectors (called the null space or kernel of $M$) gets mapped to the zero vector.

For a non-zero eigenvector $\mathbf{x}$ to exist, the matrix $(A - \lambda I)$ must perform such a collapse. Thus, we require its determinant to be zero: $\det(A - \lambda I) = 0$. This equation allows us to find the special values of $\lambda$ (the eigenvalues) for which such a collapse occurs.

This gives us our master method for finding eigenvalues.

**Definition 4. The Characteristic Equation**

The eigenvalues of an $n \times n$ matrix $A$ are the solutions $\lambda$ to the **characteristic equation**:

$$\det(A - \lambda I) = 0$$

When we compute $\det(A - \lambda I)$, we get a polynomial in $\lambda$ of degree $n$. This is called the **characteristic polynomial** of $A$. Its roots are the eigenvalues.

Once we find an eigenvalue $\lambda$, we find the corresponding eigenvectors by solving the system $(A - \lambda I)\mathbf{x} = \mathbf{0}$. The set of all solutions (including the zero vector) forms a subspace called the **eigenspace** for that $\lambda$. Any non-zero vector in this eigenspace is an eigenvector.

**Example 3. Finding Eigenvalues and Eigenspaces**

Let's find the eigenvalues and corresponding eigenspaces of $A = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}$.

**Step 1: Form the characteristic equation** $\det(A - \lambda I) = 0$. First, construct $A - \lambda I$:

$$A - \lambda I = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-\lambda & 6 \\ 5 & 2-\lambda \end{pmatrix}$$

Now, compute its determinant:

$$\begin{aligned} \det(A - \lambda I) &= (1-\lambda)(2-\lambda) - (6)(5) \\ &= 2 - \lambda - 2\lambda + \lambda^2 - 30 \\ &= \lambda^2 - 3\lambda - 28 \end{aligned}$$

The characteristic equation is $\lambda^2 - 3\lambda - 28 = 0$.

**Step 2: Solve the characteristic equation to find eigenvalues** $\lambda$. This is a quadratic equation. We can factor it:

$$(\lambda - 7)(\lambda + 4) = 0$$

The solutions are $\lambda_1 = 7$ and $\lambda_2 = -4$. These are the eigenvalues of $A$.

**Step 3: For each eigenvalue, find the corresponding eigenspace by solving** $(A - \lambda I)\mathbf{x} = \mathbf{0}$.

For $\lambda_1 = 7$: We need to solve $(A - 7I)\mathbf{x} = \mathbf{0}$.

$$A - 7I = \begin{pmatrix} 1-7 & 6 \\ 5 & 2-7 \end{pmatrix} = \begin{pmatrix} -6 & 6 \\ 5 & -5 \end{pmatrix}$$

The system is $\begin{pmatrix} -6 & 6 \\ 5 & -5 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. This gives the equations:

- $-6x_1 + 6x_2 = 0 \implies x_1 = x_2$
- $5x_1 - 5x_2 = 0 \implies x_1 = x_2$

Both equations tell us $x_1 = x_2$. Let $x_2 = t$ (a free parameter). Then $x_1 = t$. The eigenvectors are of the form $\mathbf{x} = \begin{bmatrix} t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ for $t \neq 0$. The eigenspace for $\lambda_1 = 7$ is the set of all scalar multiples of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, which is *span* $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$. A basis for this eigenspace is $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$.

For $\lambda_2 = -4$: We need to solve $(A - (-4)I)\mathbf{x} = (A + 4I)\mathbf{x} = \mathbf{0}$.

$$A + 4I = \begin{pmatrix} 1+4 & 6 \\ 5 & 2+4 \end{pmatrix} = \begin{pmatrix} 5 & 6 \\ 5 & 6 \end{pmatrix}$$

The system is $\begin{pmatrix} 5 & 6 \\ 5 & 6 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. This gives the equation $5x_1 + 6x_2 = 0$. So, $x_1 = -\frac{6}{5}x_2$. Let $x_2 = s$ (a free parameter). Then $x_1 = -\frac{6}{5}s$. The eigenvectors are of the form $\mathbf{x} = \begin{bmatrix} -6/5s \\ s \end{bmatrix} = s \begin{bmatrix} -6/5 \\ 1 \end{bmatrix}$ for $s \neq 0$. To get an eigenvector with integer components, we can choose $s = 5$ (or $s = -5$ as in the original example). If $s = 5$, $\mathbf{x} = \begin{bmatrix} -6 \\ 5 \end{bmatrix}$. If $s = -5$, $\mathbf{x} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$ (this matches $\mathbf{u}$ from Example **??**). The eigenspace for $\lambda_2 = -4$ is *span* $\left\{ \begin{bmatrix} 6 \\ -5 \end{bmatrix} \right\}$. A basis for this eigenspace is $\left\{ \begin{bmatrix} 6 \\ -5 \end{bmatrix} \right\}$.

## 3.2 A Shortcut for Triangular Matrices

For triangular matrices (either upper or lower), finding eigenvalues is delightfully simple!

**Theorem 1. Eigenvalues of a Triangular Matrix**

The eigenvalues of a triangular matrix (upper or lower) are the entries on its main diagonal.

**Why is this true? A Moment of Clarity!**

Let $A$ be an $n \times n$ upper triangular matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

Then $A - \lambda I$ is also an upper triangular matrix:

$$A - \lambda I = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ 0 & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} - \lambda \end{pmatrix}$$

Recall from our discussion on determinants that the determinant of a triangular matrix is simply the product of its diagonal entries. So, the characteristic equation $\det(A - \lambda I) = 0$ becomes:

$$(a_{11} - \lambda)(a_{22} - \lambda)\dots(a_{nn} - \lambda) = 0$$

The solutions to this equation (the eigenvalues) are transparently $\lambda_1 = a_{11}, \lambda_2 = a_{22}, \dots, \lambda_n = a_{nn}$. The diagonal entries themselves are the eigenvalues! The same logic applies to lower triangular matrices. This is a beautiful consequence of determinant properties.

# 4 The Spectral Theorem: The Crown Jewel for Symmetric Matrices

We now arrive at one of the most profound and practically useful results in linear algebra, especially impactful in fields like data science, quantum mechanics, and engineering. This theorem applies to a special, yet very common, class of matrices: **symmetric matrices**.

**Definition 5. Symmetric Matrix**

A square matrix $A$ is **symmetric** if it is equal to its transpose, i.e., $A = A^T$. This means that $a_{ij} = a_{ji}$ for all $i$ and $j$. Visually, the matrix is symmetric across its main diagonal. Example: $S = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 5 & 3 \\ 0 & 3 & -4 \end{pmatrix}$ is symmetric.

Symmetric matrices represent transformations that do not have any "hidden" rotational component beyond simple reflections (which can be part of scaling by a negative eigenvalue). They embody pure stretches or compressions along some set of perpendicular axes. This inherent geometric "honesty" is reflected in their remarkable properties.

**Theorem 2. The Spectral Theorem for Real Symmetric Matrices**

If an $n \times n$ matrix $A$ is symmetric (and has real entries), then:

1. All of its eigenvalues $\lambda$ are **real** numbers (no complex eigenvalues). This is a huge simplification!
2. Eigenvectors corresponding to **distinct** eigenvalues are **orthogonal**.
3. There exists an **orthonormal basis** for $\mathbb{R}^n$ consisting of $n$ eigenvectors of $A$. This means $A$ is always **diagonalizable**, and spectacularly so, using an orthogonal matrix.

**The "Spectrum" in Spectral Theorem**

The term "spectrum" of a matrix refers to its set of eigenvalues. The Spectral Theorem is so named because it describes the nature of these eigenvalues (real) and the structure of the eigenvectors (orthonormal basis) for symmetric matrices. It's like finding the pure frequencies (spectrum) that compose a complex sound wave.

This theorem is a cornerstone. It guarantees that for any symmetric transformation, we can find a perfect, perpendicular set of axes (an orthonormal basis for $\mathbb{R}^n$) along which the transformation acts simply as scaling. These axes are precisely the eigenvectors of the matrix.

## 4.1 Eigendecomposition: Deconstructing Symmetric Transformations

The most powerful consequence of the Spectral Theorem is that any real symmetric matrix $A$ can be factored (or "decomposed") into:

$$A = Q\Lambda Q^T$$

Where:

- $Q$ is an **orthogonal matrix** whose columns are the $n$ orthonormal eigenvectors of $A$. Recall that for an orthogonal matrix, $Q^{-1} = Q^T$. This means its columns are unit vectors and mutually orthogonal.
- $\Lambda$ (capital Lambda) is a **diagonal matrix** whose diagonal entries are the eigenvalues of $A$, in the same order as their corresponding eigenvectors in $Q$.

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

This is called the **eigendecomposition** or **spectral decomposition** of $A$.

**Example 4. Eigendecomposition of a Symmetric Matrix**

Let's find the eigendecomposition $A = Q\Lambda Q^T$ for the symmetric matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

**1. Find Eigenvalues:** $A - \lambda I = \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix}$. $\det(A - \lambda I) = (2-\lambda)^2 - 1^2 = (2-\lambda-1)(2-\lambda+1) = (1-\lambda)(3-\lambda) = 0$. Eigenvalues are $\lambda_1 = 3, \lambda_2 = 1$. (Both real, as expected).

**2. Find Eigenvectors:** <u>For $\lambda_1 = 3$:</u> $(A - 3I)\mathbf{x} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies -x_1 + x_2 = 0 \implies x_1 = x_2$. An eigenvector is $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Normalizing it:

$\mathbf{q}_1 = \frac{1}{\sqrt{1^2+1^2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

For $\lambda_2 = 1$: $(A - 1I)\mathbf{x} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies x_1 + x_2 = 0 \implies x_1 = -x_2$.

An eigenvector is $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ (or $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$). Normalizing it: $\mathbf{q}_2 = \frac{1}{\sqrt{(-1)^2+1^2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

Note: Check orthogonality: $\mathbf{q}_1 \cdot \mathbf{q}_2 = (1/\sqrt{2})(-1/\sqrt{2}) + (1/\sqrt{2})(1/\sqrt{2}) = -1/2 + 1/2 = 0$. They are indeed orthogonal!

**3. Construct $Q$ and $\Lambda$:** $Q$ has $\mathbf{q}_1, \mathbf{q}_2$ as columns: $Q = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$. $\Lambda$ has $\lambda_1, \lambda_2$ on the diagonal: $\Lambda = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$.

**4. Verify $A = Q\Lambda Q^T$:** $Q^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$.

$$
\begin{aligned}
Q\Lambda Q^T &= \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \\
&= \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 3/\sqrt{2} & 3/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \\
&= \begin{pmatrix} (1/\sqrt{2})(3/\sqrt{2}) + (-1/\sqrt{2})(-1/\sqrt{2}) & (1/\sqrt{2})(3/\sqrt{2}) + (-1/\sqrt{2})(1/\sqrt{2}) \\ (1/\sqrt{2})(3/\sqrt{2}) + (1/\sqrt{2})(-1/\sqrt{2}) & (1/\sqrt{2})(3/\sqrt{2}) + (1/\sqrt{2})(1/\sqrt{2}) \end{pmatrix} \\
&= \begin{pmatrix} 3/2 + 1/2 & 3/2 - 1/2 \\ 3/2 - 1/2 & 3/2 + 1/2 \end{pmatrix} = \begin{pmatrix} 4/2 & 2/2 \\ 2/2 & 4/2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = A \quad \checkmark
\end{aligned}
$$

The decomposition holds!

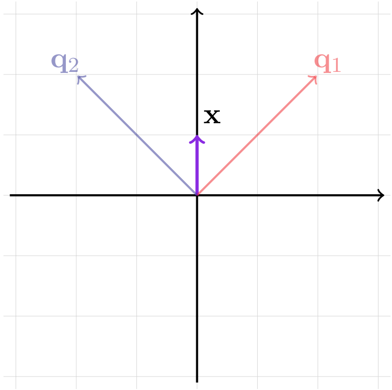
**Geometric Interpretation of Eigendecomposition: $A\mathbf{x}$**

The transformation $A\mathbf{x}$ by a symmetric matrix $A$ can be understood as a sequence of three simpler geometric operations. We use $A = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)$ from Example **??**. Eigenvalues $\lambda_1 = 3, \lambda_2 = 1$. Eigenvectors $\mathbf{q}_1 = (1/\sqrt{2}, 1/\sqrt{2})$ (Red), $\mathbf{q}_2 = (-1/\sqrt{2}, 1/\sqrt{2})$ (Blue). Let's transform the vector $\mathbf{x} = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ (Purple).

The matrix $Q = \left(\begin{smallmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{smallmatrix}\right)$ represents a counter-clockwise rotation by $45°$. Thus, $Q^T = \left(\begin{smallmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{smallmatrix}\right)$ represents a clockwise rotation by $45°$.

**1. Original Space**

Vector $\mathbf{x}$ in std. basis

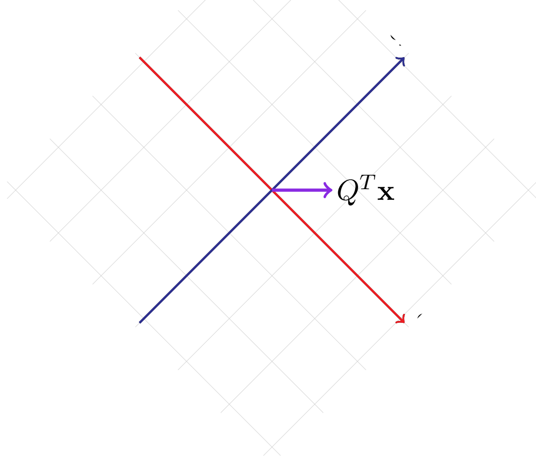



$\xrightarrow{Q^T}$

To Eigenbasis

**2. In Eigen-Basis**

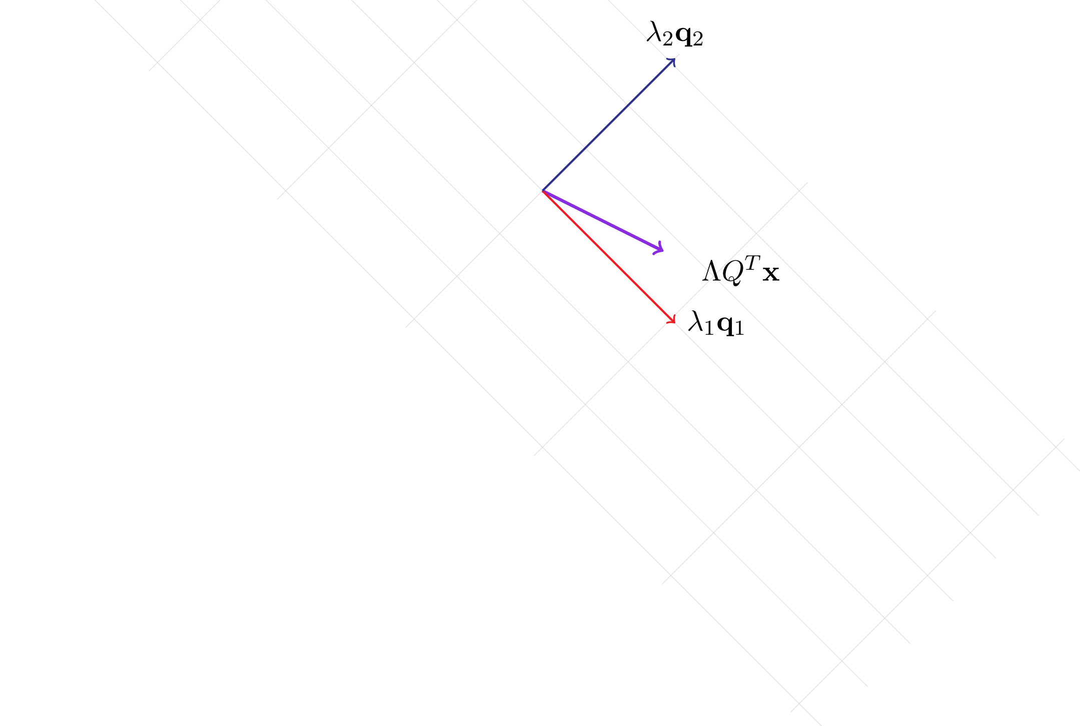
$(Q^T\mathbf{x})$




$\xrightarrow{\Lambda}$

Scale

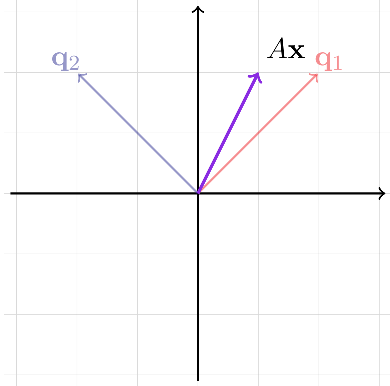
**3. Scaled in Eigen-Basis**
$(\Lambda Q^T\mathbf{x})$




**4. Final Transformation**
$(A\mathbf{x} = Q\Lambda Q^T\mathbf{x})$

$\xrightarrow{Q}$

To Standard Basis




*The three fundamental steps: 1. ($Q^T\mathbf{x}$): The vector $\mathbf{x}$ is expressed in terms of the eigenbasis $\{\mathbf{q}_1, \mathbf{q}_2\}$. This is like rotating our perspective so the new axes are the eigenvectors. 2. ($\Lambda(Q^T\mathbf{x})$): In this eigenbasis, the transformation is a simple scaling along these new axes by the eigenvalues. The space itself is stretched/shrunk along*

*$\mathbf{q}_1$ and $\mathbf{q}_2$. 3. ($Q(\Lambda Q^T\mathbf{x})$): The scaled vector is transformed back to the original standard basis. This yields the final result $A\mathbf{x}$. This decomposition is incredibly powerful as it breaks down a complex transformation (like $A$) into a rotation, a simple axis-aligned scaling, and a rotation back.*

## 4.2 The Big Picture: Why Eigendecomposition Matters

Eigenvalues, eigenvectors, and especially the eigendecomposition of symmetric matrices are not just mathematical curiosities. They are fundamental building blocks for many advanced techniques and provide deep insights into data and systems.

- **Principal Component Analysis (PCA):** This is a cornerstone of dimensionality reduction in machine learning and data analysis.
  - We often construct a symmetric **covariance matrix** (or correlation matrix) from our high-dimensional data. This matrix describes how different features in the data vary together.
  - The eigenvectors of this covariance matrix are called the **principal components**. They point in the directions of maximum variance in the data. The first principal component is the direction capturing the most variance, the second (orthogonal to the first) captures the next most, and so on.
  - The eigenvalues tell us *how much* variance exists along each principal component.
  - By keeping only the eigenvectors (principal components) corresponding to the largest eigenvalues, we can project the data onto a lower-dimensional subspace while retaining the most significant information (variance). This simplifies models, speeds up computation, and helps in visualization. The Spectral Theorem guarantees these principal components are orthogonal and the eigenvalues are real, making the interpretation clean.
- **Understanding Quadratic Forms and Ellipsoids:** A quadratic form like $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ (where $A$ is symmetric) can be simplified by changing to the eigenbasis. In this basis, $f(\mathbf{x}') = \sum \lambda_i (x_i')^2$. If $A$ is positive definite (all $\lambda_i > 0$), the equation $\mathbf{x}^T A\mathbf{x} = c$ defines an ellipsoid whose axes are aligned with the eigenvectors of $A$, and the lengths of the semi-axes are inversely related to the square roots of the eigenvalues. This is crucial in optimization, physics (e.g., inertia tensor), and statistics (e.g., confidence ellipses).

- **Solving Systems of Differential Equations:** Eigenvalues and eigenvectors are key to solving systems of linear ordinary differential equations of the form $\frac{d\mathbf{y}}{dt} = A\mathbf{y}$. The solutions often involve terms like $e^{\lambda t}\mathbf{v}$, where $\lambda$ is an eigenvalue and $\mathbf{v}$ is an eigenvector. The stability of such systems is determined by the signs of the real parts of the eigenvalues.
- **The Gram Matrix ($M^T M$ or $MM^T$):** For *any* real matrix $M$ (not necessarily square or symmetric), the matrices $M^T M$ and $MM^T$ are always symmetric and positive semidefinite (meaning their eigenvalues are $\geq 0$).
  - The eigenvalues of $M^T M$ are the squares of the singular values of $M$. The eigenvectors of $M^T M$ are the right singular vectors of $M$.
  - The eigenvalues of $MM^T$ are also the squares of the singular values of $M$. The eigenvectors of $MM^T$ are the left singular vectors of $M$.
  - This connection is fundamental to Singular Value Decomposition (SVD), another powerful matrix factorization technique that generalizes eigendecomposition to non-square matrices.

  The Spectral Theorem ensures that these important Gram matrices can be analyzed through their real, non-negative eigenvalues and orthogonal eigenvectors. This is vital in linear regression (normal equations involve $X^T X$), manifold learning, and many optimization problems.
- **Quantum Mechanics:** Observables (like energy, momentum, spin) are represented by Hermitian operators (the complex analogue of symmetric matrices). The eigenvalues of these operators are the possible measured values of the observable, and the eigenvectors are the states corresponding to those values. The Spectral Theorem's generalization to Hermitian matrices is foundational here.

The ability to decompose a symmetric matrix into $Q\Lambda Q^T$ means we can understand its action as a coordinate change to a "natural" basis (the eigenbasis), a simple scaling along these natural axes, and then a change back to the original coordinates. This simplifies complex interactions into their most fundamental components.

# 5 A Look Beyond: Further Properties and Considerations

## 5.1 Diagonalizability

The decomposition $A = PDP^{-1}$ (where $D$ is diagonal) is called **diagonalization**. A matrix $A$ is **diagonalizable** if such a $P$ (whose columns are eigenvectors) and $D$ (whose diagonal entries are eigenvalues) exist. This requires $A$ to have $n$ linearly independent eigenvectors to form the columns of $P$.

- The Spectral Theorem guarantees that all **real symmetric matrices are diagonalizable** using an orthogonal matrix $Q$ (so $P = Q$, and $P^{-1} = Q^T$). This is a very strong and convenient form of diagonalizability.
- What about non-symmetric matrices?
  - An $n \times n$ matrix is guaranteed to be diagonalizable if it has $n$ **distinct** eigenvalues. This is because eigenvectors corresponding to distinct eigenvalues are always linearly independent.
  - If a non-symmetric matrix has **repeated eigenvalues** (i.e., an eigenvalue with algebraic multiplicity greater than 1), it *might not* be diagonalizable. It is diagonalizable if and only if the geometric multiplicity (dimension of the eigenspace) for every eigenvalue equals its algebraic multiplicity (how many times it's a root of the characteristic polynomial).
  - The shear matrix $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ is a classic example of a non-diagonalizable matrix. Its only eigenvalue is $\lambda = 1$ (algebraic multiplicity 2). The eigenspace for $\lambda = 1$ is $span\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$, which has dimension 1 (geometric multiplicity 1). Since $1 \neq 2$, it doesn't have enough linearly independent eigenvectors to form a basis for $\mathbb{R}^2$.

## 5.2 Complex Eigenvalues from Real Matrices

While symmetric matrices always have real eigenvalues, non-symmetric real matrices can have complex eigenvalues.

- If a real matrix $A$ has a complex eigenvalue $\lambda = a + bi$ (where $b \neq 0$) with corresponding eigenvector $\mathbf{v}$, then its complex conjugate $\bar{\lambda} = a - bi$ is also an eigenvalue with corresponding eigenvector $\bar{\mathbf{v}}$. So, complex eigenvalues for real matrices always appear in **conjugate pairs**.

- Geometrically, a complex eigenvalue $\lambda = a \pm bi$ for a real $2 \times 2$ matrix often corresponds to a **rotation combined with scaling**.
  - Consider a pure rotation matrix $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Its eigenvalues are $\cos\theta \pm i\sin\theta = e^{\pm i\theta}$. Unless $\theta$ is $0$ or $\pi$ (resulting in $I$ or $-I$), these eigenvalues are complex. A rotation doesn't have any real vector that stays in the same direction (unless it's a trivial rotation).
  - If a matrix has eigenvalues $re^{\pm i\theta}$, the $r$ factor implies scaling, and $e^{\pm i\theta}$ implies rotation.
- The eigenvectors corresponding to these complex eigenvalues will also have complex entries. They define invariant subspaces (often planes in $\mathbb{R}^3$) where the transformation acts as a rotation-scaling.

**The Power of Abstraction**

Eigenvalues and eigenvectors are concepts that generalize far beyond simple matrices of real numbers. They appear in abstract linear algebra (linear operators on vector spaces), functional analysis (eigenfunctions of differential operators), and many areas of physics and engineering. The core idea remains the same: finding special "directions" or "states" that are preserved (up to scaling) by a transformation or operator.

## Concluding Thoughts: The Elegance of Unveiling Matrices

This journey through determinants, eigenvalues, and eigenvectors has hopefully illuminated the profound structures hidden within matrices. These are not merely arrays of numbers; they are operators that describe fundamental geometric and algebraic transformations. Understanding how to deconstruct them, particularly through eigendecomposition and the insights offered by the Spectral Theorem, empowers us to simplify complex problems, extract meaningful information from data, and appreciate the underlying elegance of linear algebra in describing the world around us. The ability to find a matrix's "true axes" and understand its action along them is a cornerstone of countless applications across science and engineering.