Introduction to Machine Learning

Lecture 5: Linear Algebra (Matrices)

Prof. Subrahmanya Swamy Peruru
Department of Electrical Engineering
IIT Kanpur

Recap

Vectors

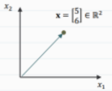
Example

We work mostly with $V = \mathbb{R}^d$.

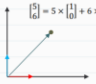
A vector:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} \in \mathbb{R}^d.$$

$\mathbf{x} = \begin{bmatrix} 5 \\ 6 \end{bmatrix} \in \mathbb{R}^2$



$\begin{bmatrix} 5 \\ 6 \end{bmatrix} = 5 \times \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 6 \times \begin{bmatrix} 0 \\ 1 \end{bmatrix}$



$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Linear Combination

Example

Given the vectors $\mathbf{x}_1, \dots, \mathbf{x}_k$, any vector

$$\mathbf{x} = \sum_{i=1}^{k} a_i \mathbf{x}_i, \quad a_i \in \mathbb{R}$$

is their linear combination.

$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$

$$\mathbf{x} = 2 \times \mathbf{x}_1 + 3 \times \mathbf{x}_2$$
$$= 2 \times \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3 \times \begin{bmatrix} 0 \\ 5 \end{bmatrix}$$
$$= \begin{bmatrix} 2 \\ 19 \end{bmatrix}$$

(Informally) A vector space is a set of vectors where linear combinations of vectors stay within the set.

Basis and Dimension

A set $\{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ is a basis of a vector space U if:

1. $U = \text{Span}\,(\{\mathbf{b}_1, \dots, \mathbf{b}_m\})$
2. it is linearly independent

Then every $\mathbf{x} \in U$ has a *unique* representation

$$\mathbf{x} = \sum_{i=1}^{m} c_i \, \mathbf{b}_i.$$

The vector $[\mathbf{x}]_{\mathcal{B}} = (c_1, \dots, c_m)^\top$ are the coordinates in basis $\mathcal{B}$.

Related facts:

Dimension of a vector space, $\dim(U)$ = number of vectors in any basis

Let U be any finite-dimensional vector space with $\dim(U) = m$, Then

- Any linearly independent set with exactly m vectors is a basis.
- Any set with fewer than m vectors is not a basis (it does not span U).
- No linearly independent set can have more than m vectors.

Coordinates of a vector will depend on Basis

What will be the coordinates of the vector $\mathbf{x} = (2,2)^\top$ using the standard basis $\mathcal{E} = \{(1,0)^\top, (0,1)^\top\}$?

$$(2,2)^T = 2 \times (1{,}0)^T + 2 \times (0{,}1)^T$$

So $[\mathbf{x}]_{\mathcal{E}} = (2,2)^\top$.

Changing the Basis:

Let $\mathbf{b}_1 = (1{,}1)^\top$ and $\mathbf{b}_2 = (1,-1)^\top$ ($\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ is another basis of $\mathbb{R}^2$)

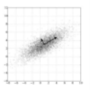
Represent $\mathbf{x} = (2,2)^\top$ as $\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2$:

$$c_1(1{,}1) + c_2(1,-1) = (c_1 + c_2,\ c_1 - c_2) = (2,2).$$

Solve:

$$c_1 + c_2 = 2, \quad c_1 - c_2 = 2 \Rightarrow \quad c_1 = 2,\ c_2 = 0.$$

So $[\mathbf{x}]_{\mathcal{B}} = (2,0)^\top$.



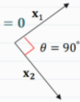
PCA

Remark:

- The coordinates of the vector will depend on the choice of basis
- Careful choice of basis leads to compact representations

Orthogonal Vectors

$$< \mathbf{x}_1, \mathbf{x}_2 > = \mathbf{x}_1^\top \mathbf{x}_2 = 0$$



A set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \dots \mathbf{x}_n\}$ are orthogonal if their pair-wise inner products is zero:

$$< \mathbf{x}_i, \mathbf{x}_j > = \mathbf{x}_i^\top \mathbf{x}_j = 0, \qquad \text{for } i \neq j$$

$$\cos\theta = \frac{\mathbf{x}_1^\top \mathbf{x}_2}{\|\mathbf{x}_1\|\|\mathbf{x}_2\|} = 0$$

$$\Rightarrow \theta = 90^\circ$$

Related Facts:

A zero vector is orthogonal to any vector since $\mathbf{x}^\top \mathbf{0} = \sum_i 0 \times x_i = 0$

Is there any relation between Orthogonality and Linear Independence?

For any set of non-zero vectors,

Orthogonal vectors $\Rightarrow$ **Linearly Independent**
Linearly Independent $\not\Rightarrow$ Orthogonal vectors

Prove this!

Norms: Definition and Examples

A **norm** $\|\cdot\|$ satisfies:

- Non-negative: $\|\mathbf{x}\| \geq 0,$
- Zero length only for the zero vector: $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0},$
- Scaling property: $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|,$
- Triangle inequality $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|.$

Examples:

$$\|\mathbf{x}\|_2 = \sqrt{\sum_i x_i^2}, \quad \|\mathbf{x}\|_1 = \sum_i |x_i|, \quad \|\mathbf{x}\|_\infty = \max_i |x_i|.$$

Let $\mathbf{x} = (3,-4,12)^\top$.

$$\|\mathbf{x}\|_2 = \sqrt{3^2 + (-4)^2 + 12^2} = 13.$$
$$\|\mathbf{x}\|_1 = |3| + |-4| + |12| = 19,$$
$$\|\mathbf{x}\|_\infty = \max\{3,4,12\} = 12.$$

Intuition (Helpful to prevent Over fitting):

L2: encourages small energy

L1: encourages sparsity;

L∞: controls worst-case coordinate.

Recap

Matrix Multiplication: Column View

Matrix Multiplication can be viewed as a linear combination of the column vectors

$$\boxed{A\mathbf{x} = \begin{bmatrix} 2 & 4 \\ 2 & 3 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 4x_2 \\ 2x_1 + 3x_2 \\ 3x_1 + 7x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2}$$

where $\mathbf{v_1} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{v_2} = \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix}$ are the column vectors.

$A\mathbf{x}$ is a linear combination of columns

This is the a useful viewpoint in ML.

Lecture Content

1 Linear Transformation

- Matrix multiplication as Linear map
- Column / Null space
- Rank-Nullity

2 Eigen Decomposition

- Orthogonal matrix
- Eigen values and vectors
- Diagonalization using EVD

What is a Linear Transformation?

In this course, we deal with mappings of the form

$$T : \mathbb{R}^m \to \mathbb{R}^n,$$

which take an input vector with m real numbers and produce an output vector with n real numbers.

Definition. The mapping T is called a ***linear transformation*** if for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ and any scalars $\alpha, \beta \in \mathbb{R}$,

Scaling property:

$$T(\alpha\mathbf{x}) = \alpha\, T(\mathbf{x})$$

Additivity property:

$$T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$$

Key takeaway.

Taking a linear combination before or after transformation gives the same result

$$T(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\, T(\mathbf{x}) + \beta\, T(\mathbf{y})$$

Exercise: Show that $T(\mathbf{0}) = \mathbf{0}$ for any linear transformation

Linear Transformations Are Easy to Analyze

Big idea.

- You do *not* need to know $T(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^m$.
- It is sufficient to know $T(\mathbf{x})$ for basis vectors.

Every vector $\mathbf{x} = [\mathrm{x}_1, \mathrm{x}_2, \dots x_m]^T \in \mathbb{R}^m$ can be written using the standard basis:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_m\mathbf{e}_m,$$

where $\mathbf{e}_i$ is the vector with a 1 in the i^{th} position and 0 elsewhere.

What linearity gives us.

$$T(\mathbf{x}) = x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) + \cdots + x_m T(\mathbf{e}_m).$$

Interpretation.

Any input vector is built from basis vectors

Linearity allows us to transform each piece independently

Knowing $\mathrm{T}(\mathbf{e}_1), \dots, \mathrm{T}(\mathbf{e}_m)$ is sufficient to know $\mathrm{T}(\mathrm{x})$ for all x

This observation allows us to study linear transformation using matrices.

Matrix Representation of $T(\mathbf{x})$

Store the transformed basis vectors as columns of a matrix:

$$A = \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_m) \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{n \times m}.$$

Then, any $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_m\mathbf{e}_m$

$$T(\mathbf{x}) = x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) + \cdots + x_m T(\mathbf{e}_m)$$

$$= \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_m) \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

$$\boxed{T(\mathbf{x}) = A\mathbf{x}}$$

Important remark: The matrix representation changes with the choice of basis

Example

Let $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ be defined by $T(\mathbf{x}) = T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ 2x_1 \\ 3x_2 \end{bmatrix}$. Compute its matrix representation.

Solution:
Verify that $T(\mathbf{x})$ is a indeed a linear transformation (check scaling and additivity property).

Find $T(\mathbf{x})$ for basis vectors

$$T(\mathbf{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 - 0 \\ 2 \times 1 \\ 3 \times 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \qquad T(\mathbf{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}.$$

Then store them as columns in a matrix.

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 0 & 3 \end{bmatrix}, \qquad T(\mathbf{x}) = A\mathbf{x}.$$

Exercise:
Find the matrix representation if we take $\{[1,1]^T, [1,-1]^T\}$ as the basis of $\mathbb{R}^2$
instead of the standard basis $\{[1,0]^T, [0,1]^T\}$.

Column Space and Rank of a Matrix

Column space: The span of the columns (i.e., all linear combinations of the columns).

$$A = \begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \boldsymbol{a}_m \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{n \times m}$$

$$\mathrm{col}(A) = \{A\mathbf{x}\colon \mathbf{x} \in \mathbb{R}^m\} = \mathrm{span}\{\mathbf{a}_1, \dots, \mathbf{a}_m\} \subseteq \mathbb{R}^n$$

Rank: Dimension of the column space (i.e., the number of linearly independent columns)

Example: Find column space and rank of the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \end{bmatrix}.$$

Columns: $\mathbf{a}_1 = (1,2,0)^\top$,$\mathbf{a}_2 = (2,4,0)^\top$.
Thus

$$\mathrm{col}(A) = \mathrm{span}\{(1,2,0)^\top, (2,4,0)^\top\}.$$

Since $(\mathbf{2}, \mathbf{4}, \mathbf{0})^{\boldsymbol{T}} = \mathbf{2} \times (\mathbf{1}, \mathbf{2}, \mathbf{0})^\top$, they are linearly dependent.
Hence, there is only **one linearly independent column**

$$\Rightarrow \mathrm{rank}(A) = 1.$$

Null Space

The null space of a matrix is the set of all **input vectors that get mapped to the zero** vector

$$\mathrm{null}(A) = \{\mathbf{x} \in \mathbb{R}^m : A\mathbf{x} = \mathbf{0}\} \subseteq \mathbb{R}^m.$$

Example: Find null space of the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \end{bmatrix}$.

Solution: solve $A\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

$$x_1 + 2x_2 = 0,\ 2x_1 + 4x_2 = 0 \Rightarrow x_1 = -2x_2.$$

So any vector of the form $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ belongs to the null space.

$$\mathrm{null}(A) = \mathrm{span}\left\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\}, \quad \dim(\mathrm{null}(A)) = 1.$$

Rank-Nullity Theorem

The number of columns of a matrix **equals the sum** of its
rank *(dimension of the column space)*
and nullity *(dimension of the null space).*

For $A \in \mathbb{R}^{n \times m}$:

$$\mathbf{rank}(\boldsymbol{A}) + \mathbf{dim}\big(\mathbf{null}(\boldsymbol{A})\big) = \mathbf{number\ of\ columns.}$$

Key observation:

$$\mathrm{rank(A)} \ < \ \text{number of columns} \Rightarrow \ \mathrm{dim(null(A))} \ > \ 0 \ \Rightarrow \text{null space contains non} - \text{zero vectors}$$

There exist **non-zero vectors** $\tilde{\mathbf{x}} \in \mathcal{N}(A)$, i.e.,

$$A(\mathbf{x} + \tilde{\mathbf{x}}) = A\mathbf{x}.$$

Thus, adding a null-space vector does not change the output.

Null Space Example: Different Inputs, Same Output

Consider $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \end{bmatrix}$, $\quad \mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

$$\operatorname{rank}(A) = 1,$$
$$\dim(\operatorname{null}(A)) = 1,$$
$$\text{and number of columns} = 2.$$

Null-space direction. $null(A) = span\left\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\}$. Let $\tilde{\mathbf{x}} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

Same output for $\mathbf{x}$ and $(\mathbf{x} + \tilde{\mathbf{x}})$.

$$A\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \qquad A(\mathbf{x} + \tilde{\mathbf{x}}) = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 + 2 \\ -2 + 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

$$\boxed{A(\mathbf{x} + \tilde{\mathbf{x}}) = A\mathbf{x}} \quad \text{since} \quad A\tilde{\mathbf{x}} = \mathbf{0}.$$

Orthogonal Matrices

$Q \in \mathbb{R}^{n\times n}$ is orthogonal if:

$$\boxed{Q^\top Q = I}$$

Equivalently, the **columns of Q form an orthonormal basis** for $\mathbb{R}^n$.

Key properties:

- Inverse equals transpose: $Q^{-1} = Q^\top$
- Preserves length: For all $\mathbf{x} \in \mathbb{R}^n$, $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$

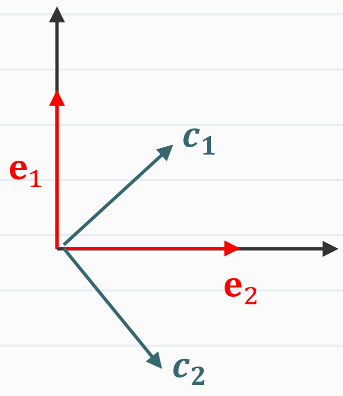
Geometric meaning: rotation/reflection.

Example: $Q = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

Verify $Q^\top Q = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = I$.

Orthogonal Matrices as Coordinate Transforms

Orthogonal matrices are used for changing coordinates between orthonormal bases



Standard basis in $\mathbb{R}^2$:

$$\mathbf{e}_1 = \begin{bmatrix}1\\0\end{bmatrix}, \qquad \mathbf{e}_2 = \begin{bmatrix}0\\1\end{bmatrix}.$$

New orthonormal basis:

$$\mathbf{c}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \qquad \mathbf{c}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}.$$

Form the orthogonal matrix with the *new basis vectors as columns*:

$$Q = [\mathbf{c}_1 \quad \mathbf{c}_2] = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}, \qquad Q^\top Q = I.$$

Coordinate conversion rules:

$$\boxed{[\mathbf{x}]_{\text{std}} = Q\,[\mathbf{x}]_{\text{new}}} \quad \text{and} \quad \boxed{[\mathbf{x}]_{\text{new}} = Q^\top[\mathbf{x}]_{\text{std}}}$$

Example

Example vector (geometric vector): $\mathbf{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

Coordinates in the standard basis:

$$\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \mathbf{2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \mathbf{2} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \Rightarrow \qquad [\mathbf{x}]_{\text{std}} = \begin{bmatrix} \mathbf{2} \\ \mathbf{2} \end{bmatrix}.$$

Coordinates in the new basis:

$$\begin{bmatrix} 2 \\ 2 \end{bmatrix} = (\mathbf{2\sqrt{2}}) \times \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) + \mathbf{0} \times \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right) \qquad \Rightarrow \qquad [\mathbf{x}]_{\text{new}} = \begin{bmatrix} \mathbf{2\sqrt{2}} \\ \mathbf{0} \end{bmatrix}.$$

Key idea: The vector $\mathbf{x}$ is the same, only the coordinate description changes.

Check on the example.

New → Standard:

$$Q[\mathbf{x}]_{\text{new}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 2\sqrt{2} \\ 2\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = [\mathbf{x}]_{\text{std}}.$$

Standard → New:

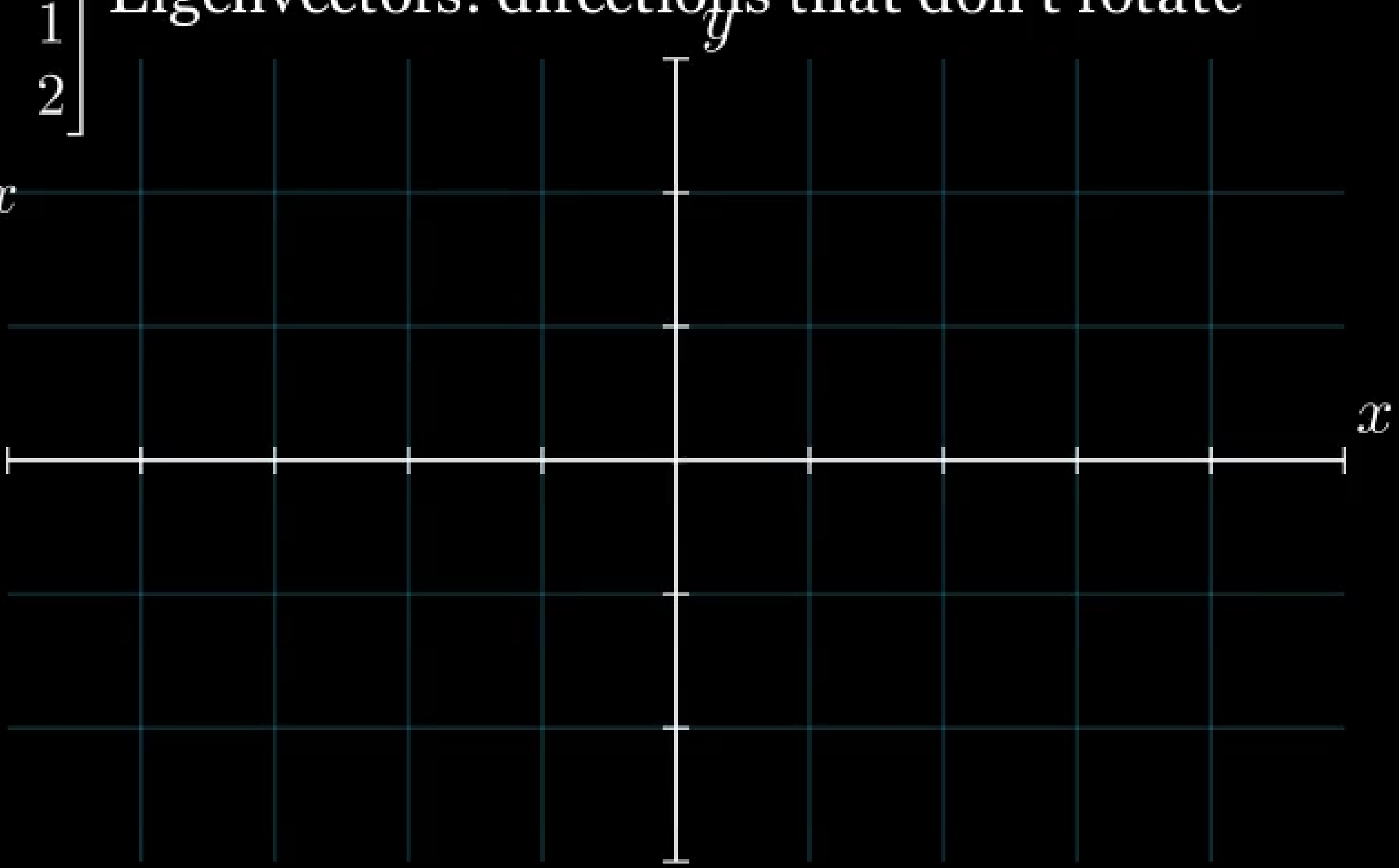
$$Q^{\top}[\mathbf{x}]_{\text{std}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix} = [\mathbf{x}]_{\text{new}}$$

Eigen Vectors

Eigenvectors: directions that don't rotate

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

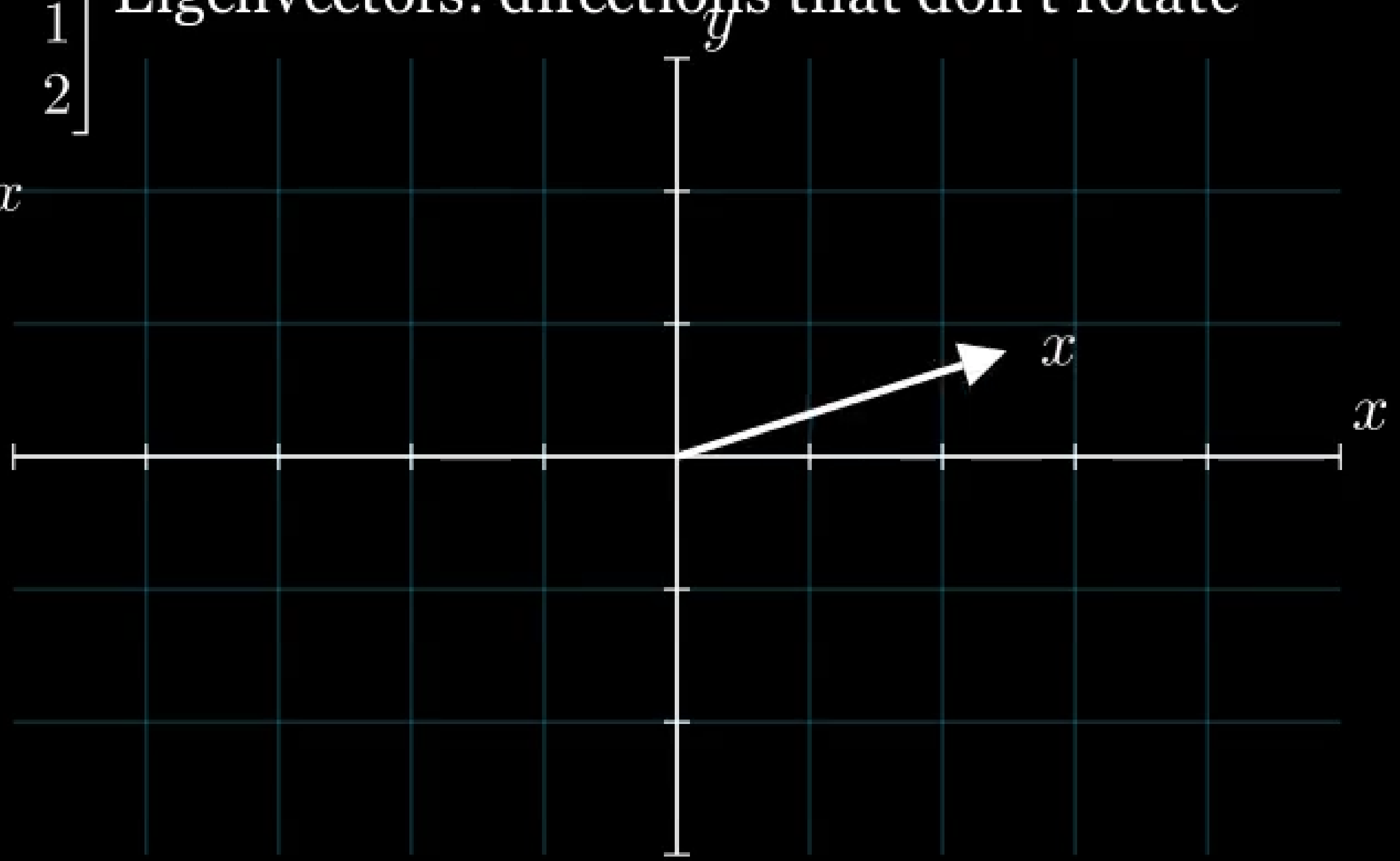
$$Ax = \lambda x$$



Eigenvectors: directions that don't rotate

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

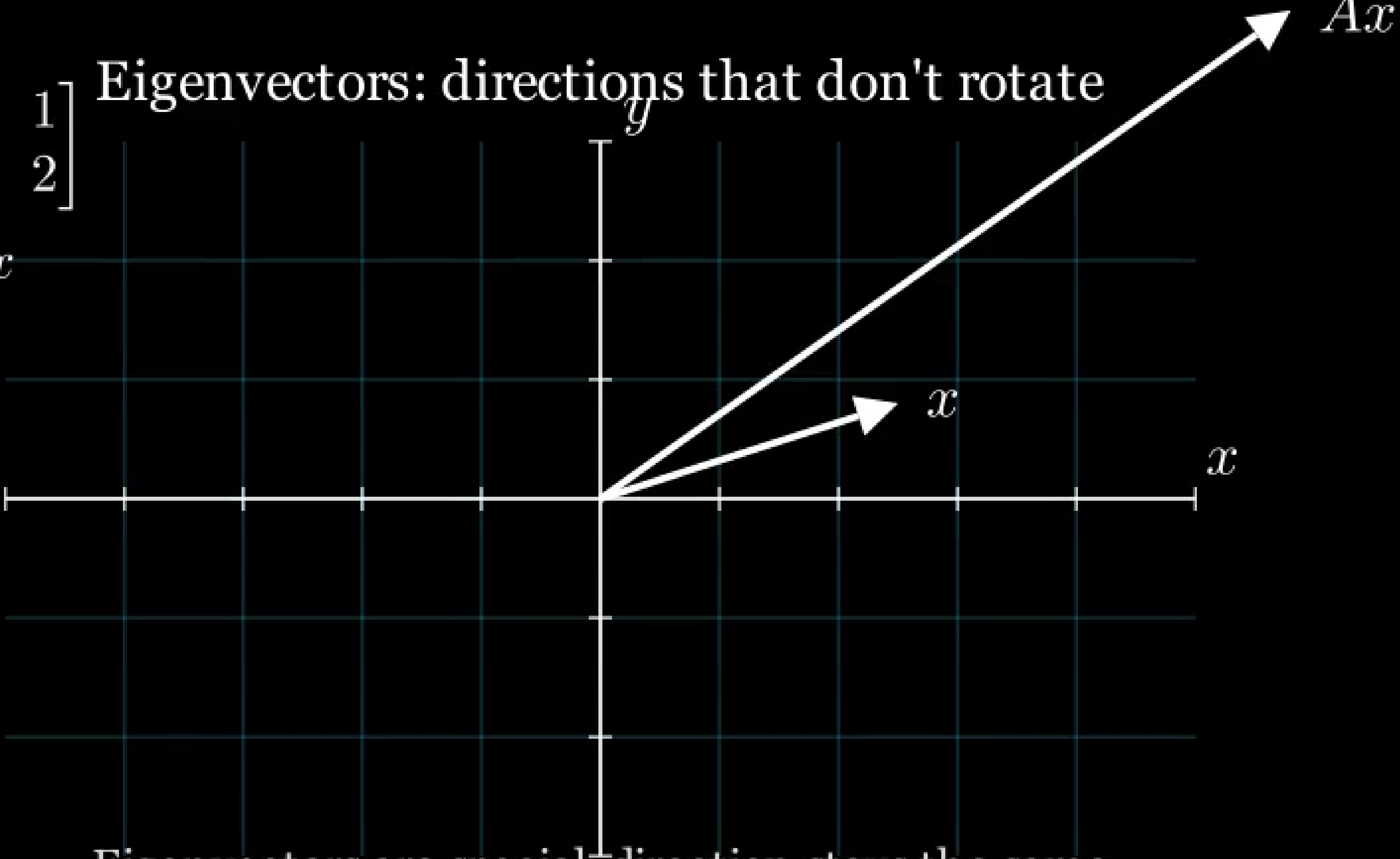
$$Ax = \lambda x$$



Eigenvectors: directions that don't rotate

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$Ax = \lambda x$$

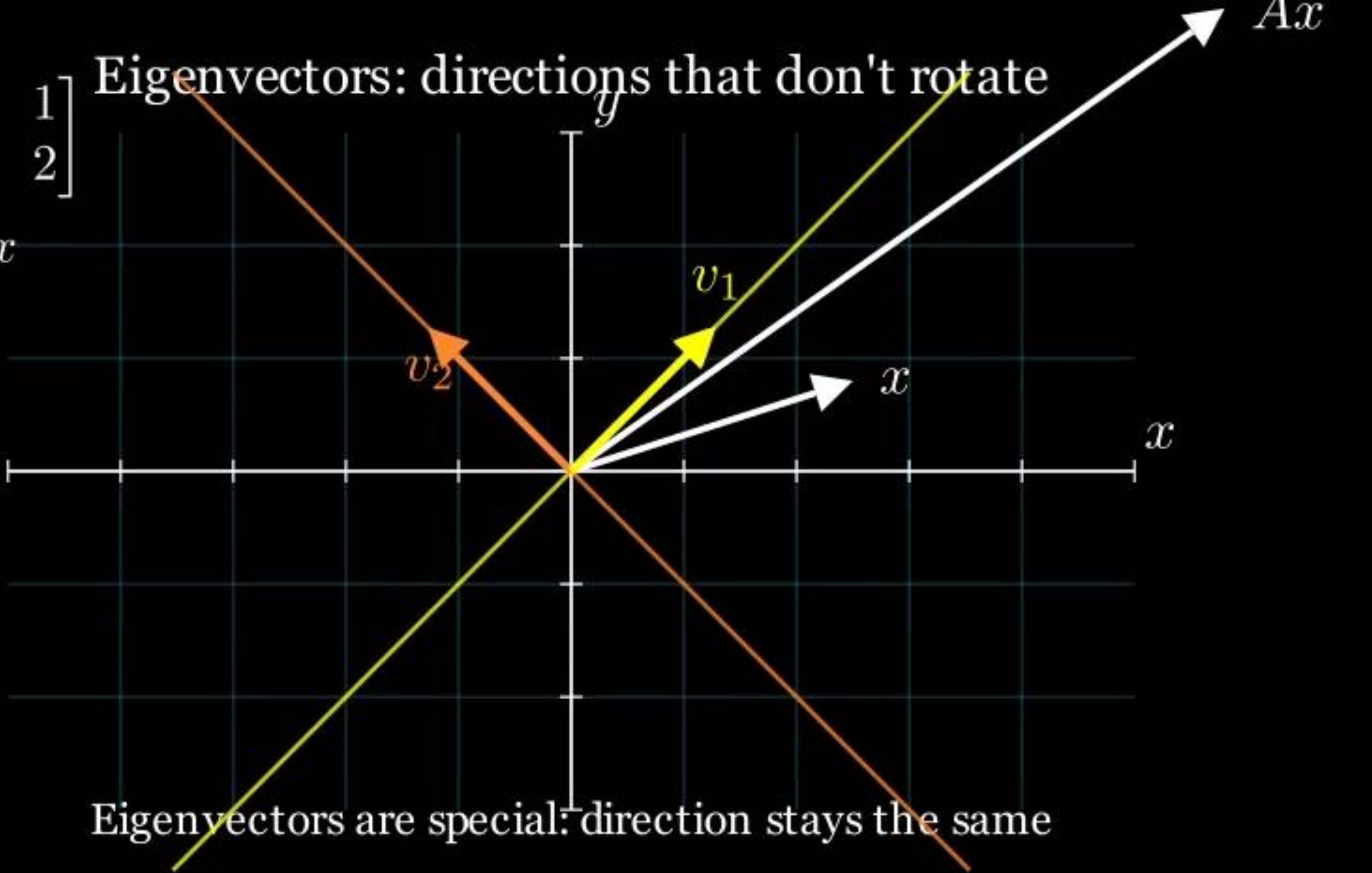


Eigenvectors are special: direction stays the same

Eigenvectors: directions that don't rotate

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$Ax = \lambda x$$



Eigenvectors are special: direction stays the same

Eigenvalues and Eigenvectors

$\mathbf{x} \neq 0$ is an eigenvector if

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

Eigenvalues are the roots of the characteristic equation:

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

Eigenvectors represent special directions that are preserved by a matrix, while eigenvalues indicate the corresponding scaling.

If $\mathbf{x}$ is an eigenvector, then any nonzero scalar multiple of $\mathbf{x}$ is also an eigenvector.

Worked Example: Eigenvalues/Eigenvectors

Find eigen vectors and eigen values of $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

Solution:

$$\det(A - \lambda I) = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3.$$

Solving for Eigen values: $\det(A - \lambda I) = \lambda^2 - 4\lambda + 3 = 0$ gives the eigen values $\lambda = 3,\ 1$.

Eigenvectors:
(i) For $\lambda = 3$:

$$(A - 3I)\mathbf{x} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 + x_2 \\ x_1 - x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Hence

$$-x_1 + x_2 = 0 \ \Rightarrow\ x_1 = x_2 \ \Rightarrow\ \mathbf{x} = x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

(ii) For $\lambda = 1$: (similarly)

$$(A - I)\mathbf{x} = \mathbf{0} \ \Rightarrow\ x_1 = -x_2 \ \Rightarrow\ \mathbf{x} = x_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

So, any vector along the direction $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a eigen vector with eigen value 3.

Similarly, any vector along the direction $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigen vector with eigen value 1.

Diagonal Matrices

Definition:

A matrix is called *diagonal* if all its off-diagonal entries are zero.

$$D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}.$$

Observations:

1. Each coordinate of the vector is scaled independently when multiplied by a diagonal matrix.

$$D\mathbf{x} = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} d_1 x_1 \\ d_2 x_2 \end{bmatrix}.$$

2. The diagonal entries are its eigen values, i.e., $\lambda_1 = \mathrm{d}_1,\ \lambda_2 = \mathrm{d}_2$

3. The standard basis vectors, $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, are its eigenvectors.

Exercise: Verify the above claims on eigen values and vectors.

Thank You