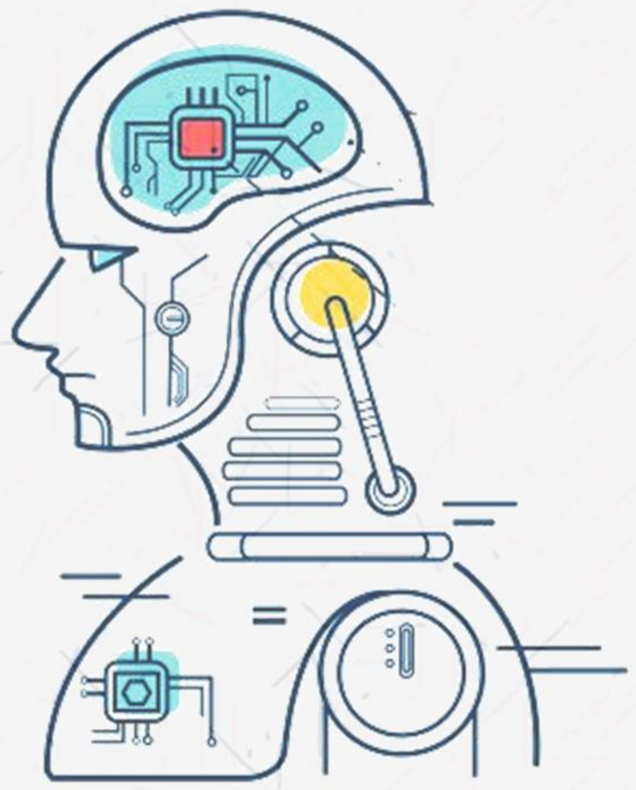
وزارة الإتصالات
وتكنولوجيا المعلومات



Linear Algebra for Data Science

ITI AI TALENTS HUB

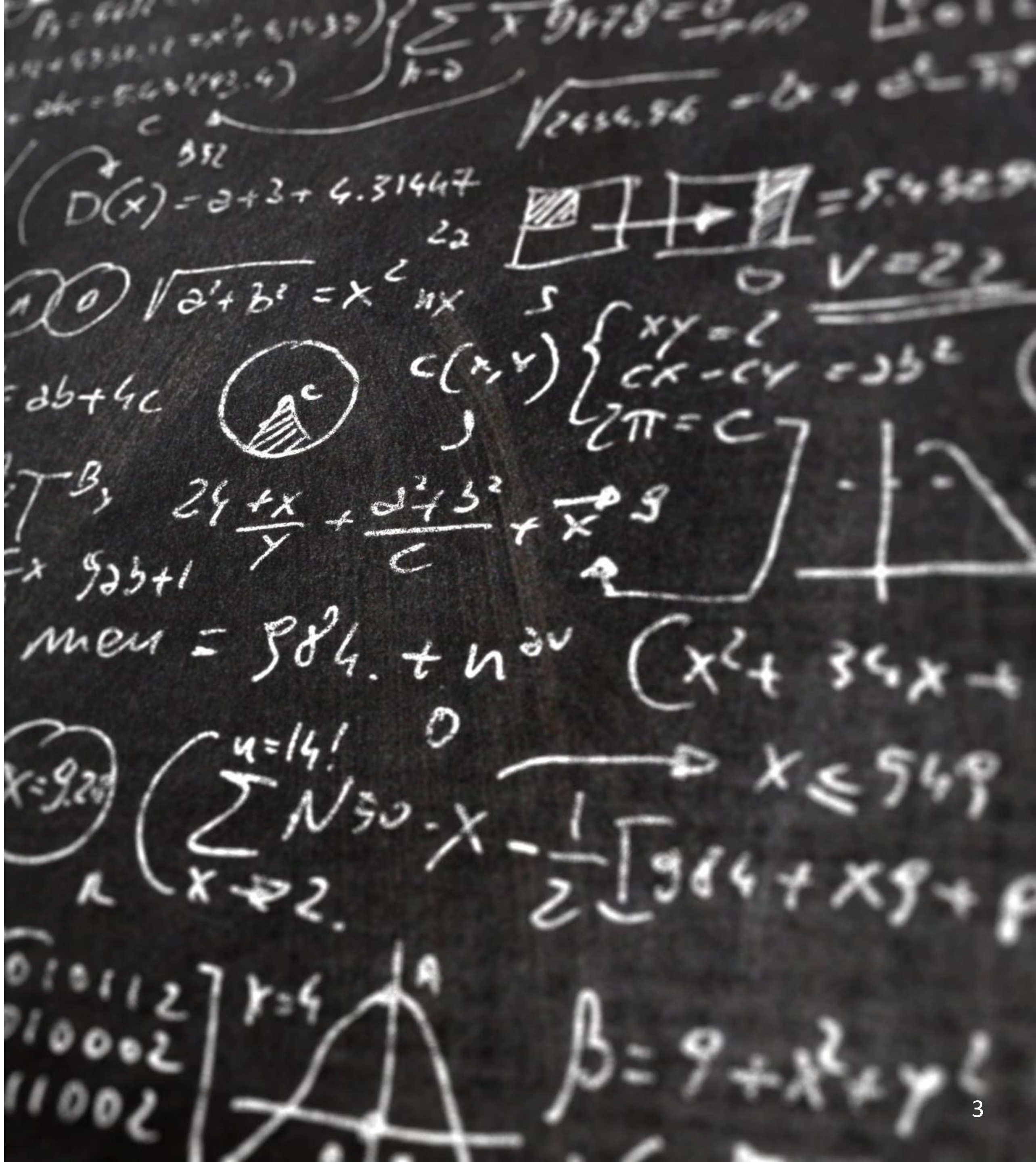
EPITA
ÉCOLE D'INGÉNIEURS EN INFORMATIQUE

Linear Algebra for data science

Summary of Session 4 – Nov. 2025

Session 4 contents

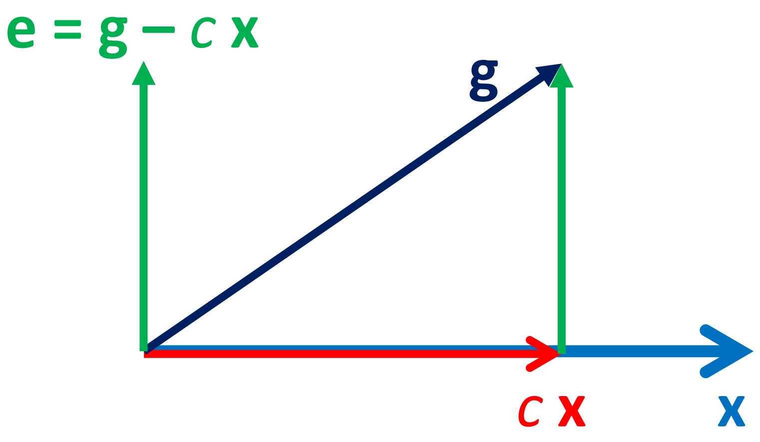
- Gram–Schmidt process.
- Eigenvectors and Eigenvalues
- Changing to Eigen Basis (diagonalization).



Introduction to orthogonalization

Introduction, Vector Projection, Vector component along another vector

Projection of a Vector into another Vector



$$proj_x(\boldsymbol{g}) = c\mathbf{x} = \frac{\mathbf{g}\cdot\mathbf{x}}{\mathbf{x}\cdot\mathbf{x}}\mathbf{x}$$

- The vector *c***x** is the projection of the vector **g** into the vector **x**

$$c\boldsymbol{x} = proj_x(\boldsymbol{g})$$

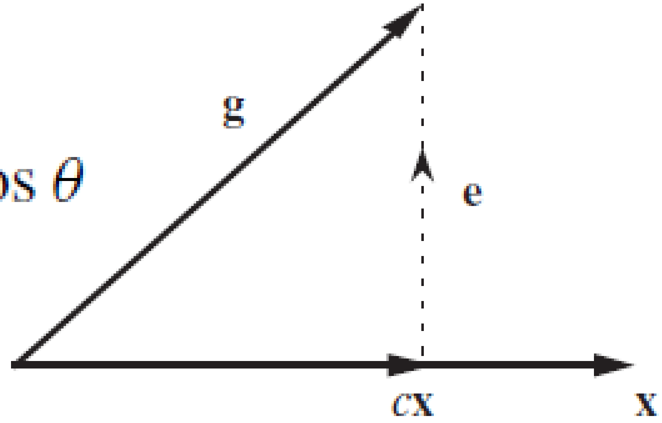
- *c* is the component of the vector **g** along the vector **x**

$$c = \frac{\boldsymbol{g}\cdot\boldsymbol{x}}{\boldsymbol{x}\cdot\boldsymbol{x}}$$

Component of a vector along another vector and cosine similarity

- Remember ‘Cosine Similarity’?
- Two vectors **g** and **x** are <u>similar</u> if **g** has a large component along **x**.

$$c||\mathbf{x}|| = ||\mathbf{g}||\cos\theta$$



$$c\mathbf{x} = \frac{\mathbf{g}\cdot\mathbf{x}}{\mathbf{x}\cdot\mathbf{x}}\mathbf{x}$$

$$\boxed{\begin{gathered} c||\mathbf{x}||^2 = ||\mathbf{g}||\,||\mathbf{x}||\cos\theta = <\mathbf{g}, \mathbf{x}> \\[2em] c = \frac{<\mathbf{g}, \mathbf{x}>}{<\mathbf{x}, \mathbf{x}>} = \frac{1}{||\mathbf{x}||^2} <\mathbf{g}, \mathbf{x}> \end{gathered}}$$

Ref: (The Oxford Series in Electrical and Computer Engineering) Ding, Zhi, Lathi, Bhagwandas Pannalal - Modern digital and analog communication systems-Oxford University Press, USA (2018_2019)

Gram-Schmidt Orthonormalization

Orthonormalization ➔ Orthogonalization + Normalization

Gram-Schmidt Orthonormalization of a Vector Set

- Remember: the dimensionality of a vector space is th… independent vectors in the space.
- In an N-dimensional space $\mathbb{R}^n$, there can be no more … ›rs.
- It is always possible to find a set of *n* vectors that are i…
- Starting from any *n* independent vectors we c… vectors that is orthogonal.

$$\mathbf{y}_1 = \mathbf{x}_1$$

$$\mathbf{y}_2 = \mathbf{x}_2 - \frac{<\mathbf{x}_1, \mathbf{x}_2>}{\|\mathbf{x}_1\|^2}\mathbf{x}_1$$

$$= \mathbf{x}_2 - \frac{<\mathbf{y}_1, \mathbf{x}_2>}{\|\mathbf{y}_1\|^2}\mathbf{y}_1$$

$$\mathbf{y}_3 = \mathbf{x}_3 - \text{sum of projections of } \mathbf{x}_3 \text{ on } \mathbf{y}_1 \text{ and } \mathbf{y}_2$$

$$= \mathbf{x}_3 - \frac{<\mathbf{y}_1, \mathbf{x}_3>}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \frac{<\mathbf{y}_2, \mathbf{x}_3>}{\|\mathbf{y}_2\|^2}\mathbf{y}_2$$

$$\boxed{\mathbf{y}_j = \mathbf{x}_j - \sum_{k=1}^{j-1} \frac{<\mathbf{y}_k, \mathbf{x}_j>}{\|\mathbf{y}_k\|^2}\mathbf{y}_k \qquad j = 2, 3, \ldots, N}$$

Gram-Schmidt Orthogonalization of a Vector Set

- Starting with three independent vectors **x1**, **x2**, **x3**, *we* intend to construct three orthogonal vectors **y1**, **y2**, **y3**.
- Start by choosing **y1** = **x1**.
- The next vector **y2** must be perpendicular to **y1**. Take **x2** *and subtract its projection along* **y1**. This leaves the perpendicular part.
- Repeat for **x3.**
- These results can be extended to an *n*-dimensional space.

$$\mathbf{y}_1 = \mathbf{x}_1$$

$$\mathbf{y}_2 = \mathbf{x}_2 - \frac{<\mathbf{x}_1, \mathbf{x}_2>}{\|\mathbf{x}_1\|^2}\mathbf{x}_1 = \mathbf{x}_2 - \frac{<\mathbf{y}_1, \mathbf{x}_2>}{\|\mathbf{y}_1\|^2}\mathbf{y}_1$$

$$\mathbf{y}_3 = \mathbf{x}_3 - \text{sum of projections of } \mathbf{x}_3 \text{ on } \mathbf{y}_1 \text{ and } \mathbf{y}_2 = \mathbf{x}_3 - \frac{<\mathbf{y}_1, \mathbf{x}_3>}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \frac{<\mathbf{y}_2, \mathbf{x}_3>}{\|\mathbf{y}_2\|^2}\mathbf{y}_2$$

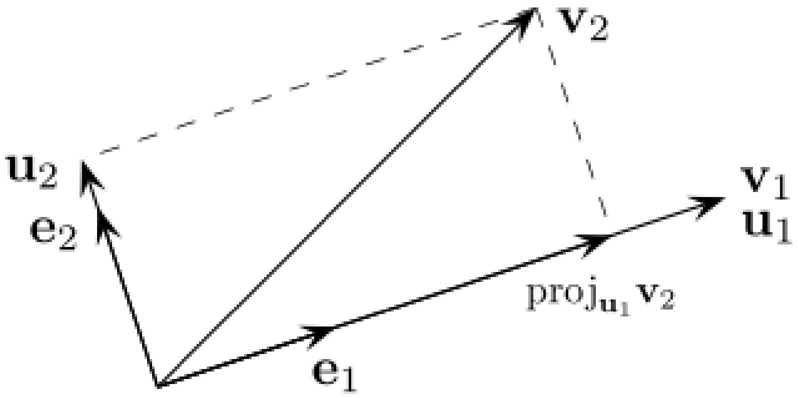
$$\boxed{\mathbf{y}_j = \mathbf{x}_j - \sum_{k=1}^{j-1} \frac{<\mathbf{y}_k, \mathbf{x}_j>}{\|\mathbf{y}_k\|^2}\mathbf{y}_k \qquad j = 2, 3, \ldots, N}$$

Normalization

- The resulting vector set **y1**, **y2**, **y3**, is not yet an orthonormal set.
- The orthonormal set can be obtained by normalizing the lengths of the respective vectors.

$$\hat{\mathbf{y}}_k = \frac{\mathbf{y}_k}{||\mathbf{y}_k||}$$

Example, Euclidean Space



Consider the following set of vectors in $\mathbf{R}^2$ (with the conventional inner product)

$$S = \left\{ \mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\}.$$

Now, perform Gram–Schmidt, to obtain an orthogonal set of vectors:

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_2) = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - \mathrm{proj}_{\left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - \frac{8}{10} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 6/5 \end{bmatrix}.$$

We check that the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are indeed orthogonal:

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \left\langle \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -2/5 \\ 6/5 \end{bmatrix} \right\rangle = -\frac{6}{5} + \frac{6}{5} = 0,$$

noting that if the dot product of two vectors is 0 then they are orthogonal.

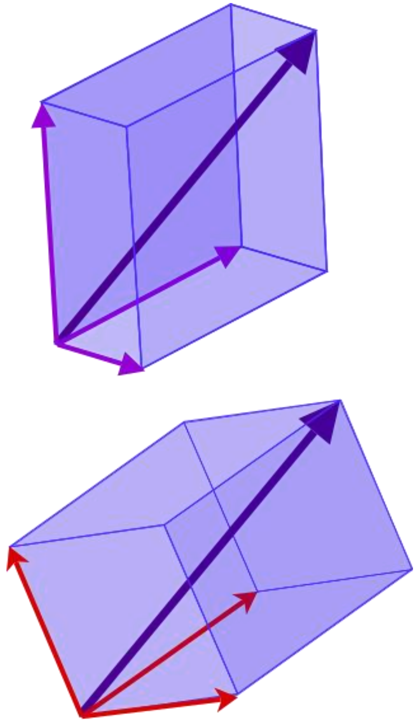
For non-zero vectors, we can then normalize the vectors by dividing out their sizes as shown above:

$$\mathbf{e}_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mathbf{e}_2 = \frac{1}{\sqrt{\frac{40}{25}}} \begin{bmatrix} -2/5 \\ 6/5 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

Change of Basis

Change of Basis Matrix



Any set of two linearly independent vectors $\{\hat{e}_1, \hat{e}_2\}$ can serve as a basis for $\mathbb{R}^2$. We can write any vector $\vec{v} \in \mathbb{R}^2$ as a linear combination of these basis vectors $\vec{v} = v_1\hat{e}_1 + v_2\hat{e}_2$.

Note the *same* vector $\vec{v}$ corresponds to different coordinate pairs depending on the basis used: $\vec{v} = (v_x, v_y)$ in the standard basis $B_s \equiv \{\hat{\imath}, \hat{\jmath}\}$, and $\vec{v} = (v_1, v_2)$ in the basis $B_e \equiv \{\hat{e}_1, \hat{e}_2\}$. Therefore, it is important to keep in mind the basis with respect to which the coefficients are taken, and if necessary specify the basis as a subscript, e.g., $(v_x, v_y)_{B_s}$ or $(v_1, v_2)_{B_e}$.

Converting a coordinate vector from the basis B_e to the basis B_s is performed as a multiplication by a *change of basis* matrix:

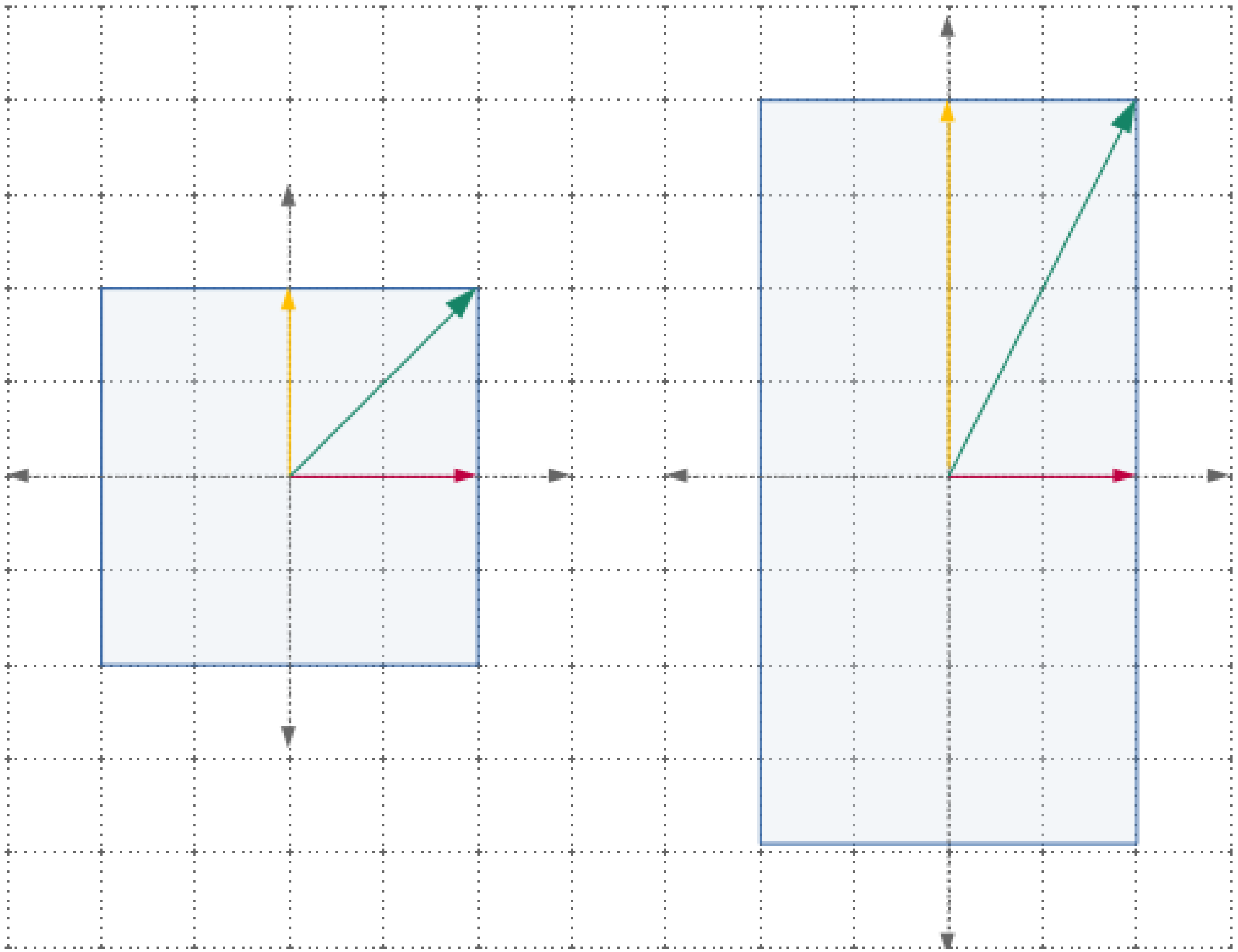
$$\begin{bmatrix}\vec{v}\end{bmatrix}_{B_s} = {}_{B_s}\begin{bmatrix} & \mathbb{1} & \end{bmatrix}_{B_e} \begin{bmatrix}\vec{v}\end{bmatrix}_{B_e} \quad \Leftrightarrow \quad \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \hat{\imath}\cdot\hat{e}_1 & \hat{\imath}\cdot\hat{e}_2 \\ \hat{\jmath}\cdot\hat{e}_1 & \hat{\jmath}\cdot\hat{e}_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

https://minireference.com/static/tutorials/linear_algebra_in_4_pages.pdf

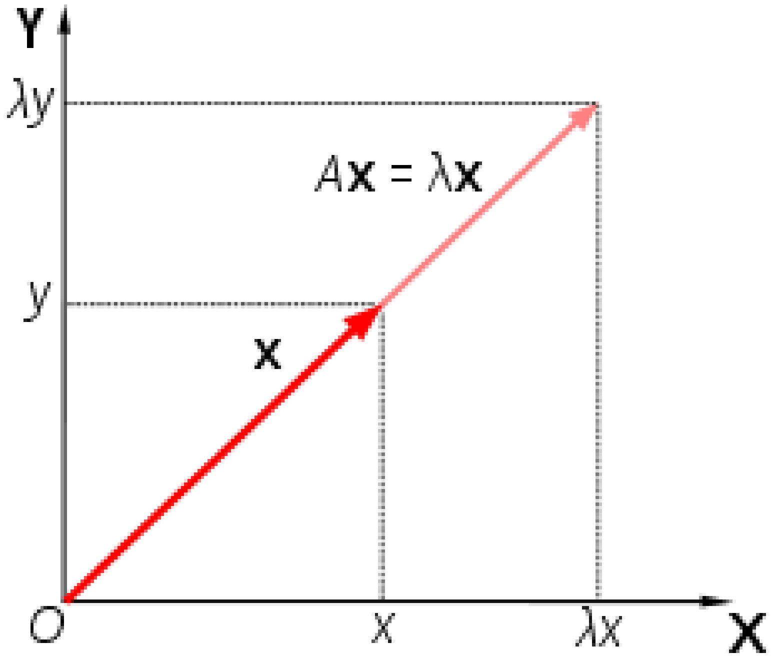
Eigenvectors and Eigenvalues

Eigenvalues and Eigen Vectors of a Matrix

- Loosely speaking, **eigenvectors** are vectors that do not change their span after transformation
- An eigenvector is considered the characteristic of transformation.
- **Eigenvalues** are the values associated with eigenvectors after transformation.



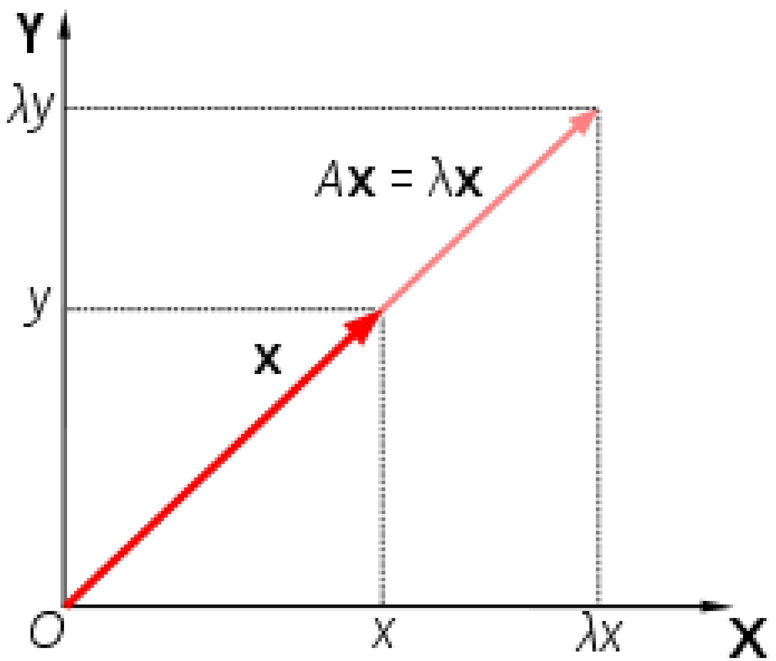
Eigenvalues and Eigen Vectors of a Matrix



- Formally, In linear algebra, an **eigenvector** or **characteristic vector** of a linear transformation is a nonzero vector that changes by a scalar factor when that linear transformation is applied to it. The corresponding **eigenvalue**, often denoted by λ is the factor by which the eigenvector is scaled.

- Geometrically, an **eigenvector**, corresponding to a real nonzero eigenvalue, points in a direction in which it is stretched by the transformation and the **eigenvalue** is the factor by which it is stretched. If the eigenvalue is negative, the direction is reversed. Loosely speaking, in a multidimensional vector space, the eigenvector is not rotated.

Finding the Eigenvalues and Eigen Vectors of a Matrix



- Suppose we have a transformation matrix A and we apply this transformation to vector x. This will be equivalent to stretching (or diminishing) the vector x by a scalar factor λ (λ can be positive or negative).

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

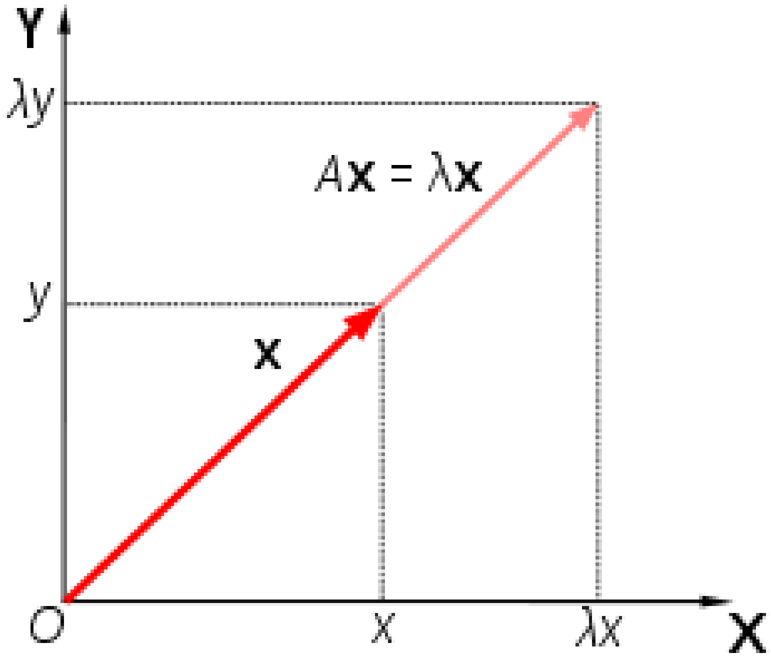
$$\mathbf{(A-\lambda I)x=0}$$

- A is n × n matrix and x is n dimensional vector

$$\mathbf{|A-\lambda I|=0}$$

- The above equation has a non-zero solution iff the determinant of the matrix (A − λI) is zero.

Finding the Eigenvalues and Eigen Vectors of a Matrix



- Evaluating this determinant gives the characteristic polynomial i.e.,

|A−λI|=0

$$\text{if } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\left| \begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0$$

$$\lambda^2 - (a+d)\lambda + ad - bc = 0$$

- The solution of this equation gives the eigenvalues. Put these eigenvalues in the original expression to get their corresponding eigenvectors.

Eigenvalues and eigenvectors example

$$A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \quad \det \begin{bmatrix} .8-\lambda & .3 \\ .2 & .7-\lambda \end{bmatrix} = \lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} = (\lambda - 1)\left(\lambda - \frac{1}{2}\right).$$

I factored the quadratic into $\lambda - 1$ times $\lambda - \frac{1}{2}$, to see the two eigenvalues $\boldsymbol{\lambda} = \mathbf{1}$ and $\boldsymbol{\lambda} = \frac{\mathbf{1}}{\mathbf{2}}$. For those numbers, the matrix $A - \lambda I$ becomes *singular* (zero determinant). The eigenvectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are in the nullspaces of $A - I$ and $A - \frac{1}{2}I$.

$(A - I)\boldsymbol{x}_1 = 0$ is $A\boldsymbol{x}_1 = \boldsymbol{x}_1$ and the first eigenvector is $(\mathbf{.6},\ \mathbf{.4})$.

$(A - \frac{1}{2}I)\boldsymbol{x}_2 = 0$ is $A\boldsymbol{x}_2 = \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{x}_2$ and the second eigenvector is $(\mathbf{1},\ \mathbf{-1})$:

$$\boldsymbol{x}_1 = \begin{bmatrix} .6 \\ .4 \end{bmatrix} \quad \text{and} \quad A\boldsymbol{x}_1 = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}\begin{bmatrix} .6 \\ .4 \end{bmatrix} = \boldsymbol{x}_1 \quad (A\boldsymbol{x} = \boldsymbol{x} \text{ means that } \lambda_1 = 1)$$

$$\boldsymbol{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad A\boldsymbol{x}_2 = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} .5 \\ -.5 \end{bmatrix} \quad (\text{this is } \tfrac{1}{2}\boldsymbol{x}_2 \text{ so } \lambda_2 = \tfrac{1}{2}).$$

If $\boldsymbol{x}_1$ is multiplied again by A, we still get $\boldsymbol{x}_1$. Every power of A will give $A^n\boldsymbol{x}_1 = \boldsymbol{x}_1$. Multiplying $\boldsymbol{x}_2$ by A gave $\frac{1}{2}\boldsymbol{x}_2$, and if we multiply again we get $(\frac{1}{2})^2$ times $\boldsymbol{x}_2$.

Eigenvalues relationship to the trace and the determinant

$$\det(A) = \prod_{i=1}^{n} \lambda_i$$

$$\mathrm{Trace}(A) = \sum_{i=1}^{n} \lambda_i$$

- Example:

$$A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$$

$$\lambda_1 = 1, \lambda_2 = 0.5$$

$$\mathrm{Tr}(A) = ?$$

$$\det(A) = ?$$

Practice

- Find the eigenvalues and eigen vectors for the transformation matrices discussed in the following slides.

Rotation

Example: We want to create a matrix $\mathbf{A}_\alpha$ that performs a rotation by an angle α.

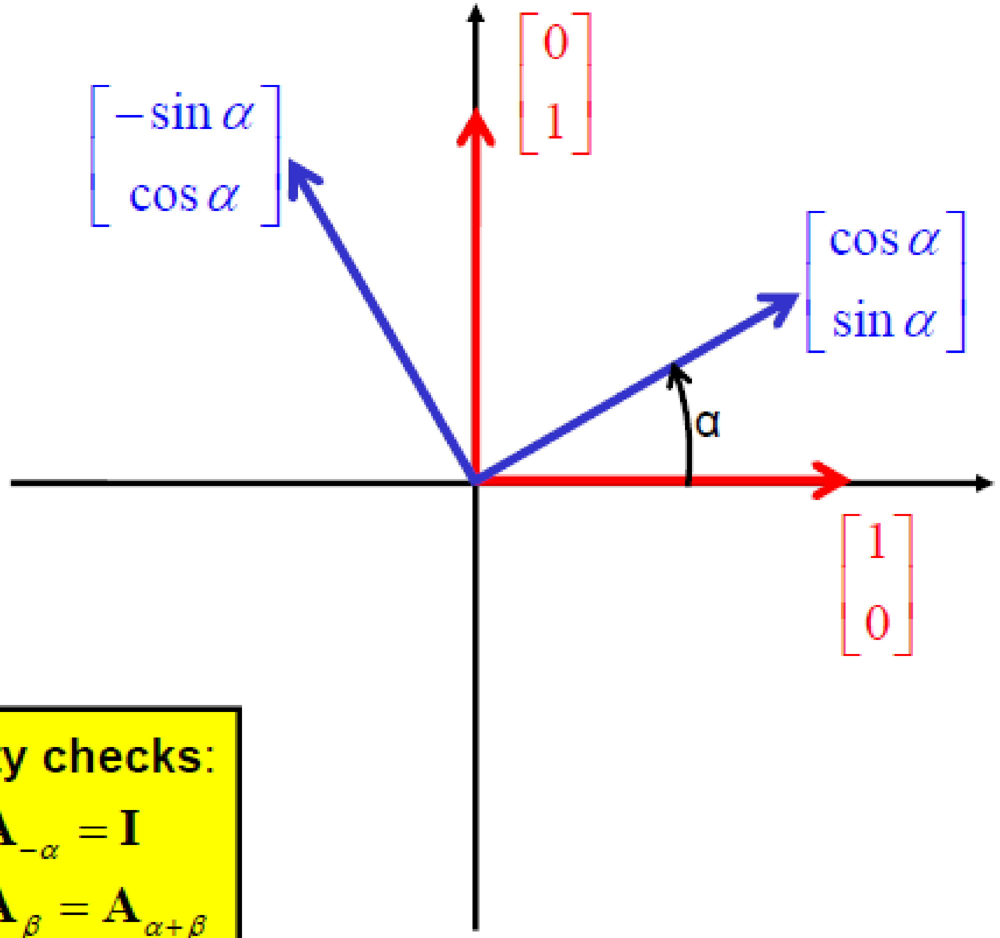
First determine what hapens to the basis vectors.

$$\mathbf{A}_\alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1,1} \\ a_{2,1} \end{bmatrix} = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$$

$$\mathbf{A}_\alpha \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{1,2} \\ a_{2,2} \end{bmatrix} = \begin{bmatrix} -\sin\alpha \\ \cos\alpha \end{bmatrix}$$

This gives us the two columns of $\mathbf{A}_\alpha$ directly:

$$\mathbf{A}_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$



Sanity checks:

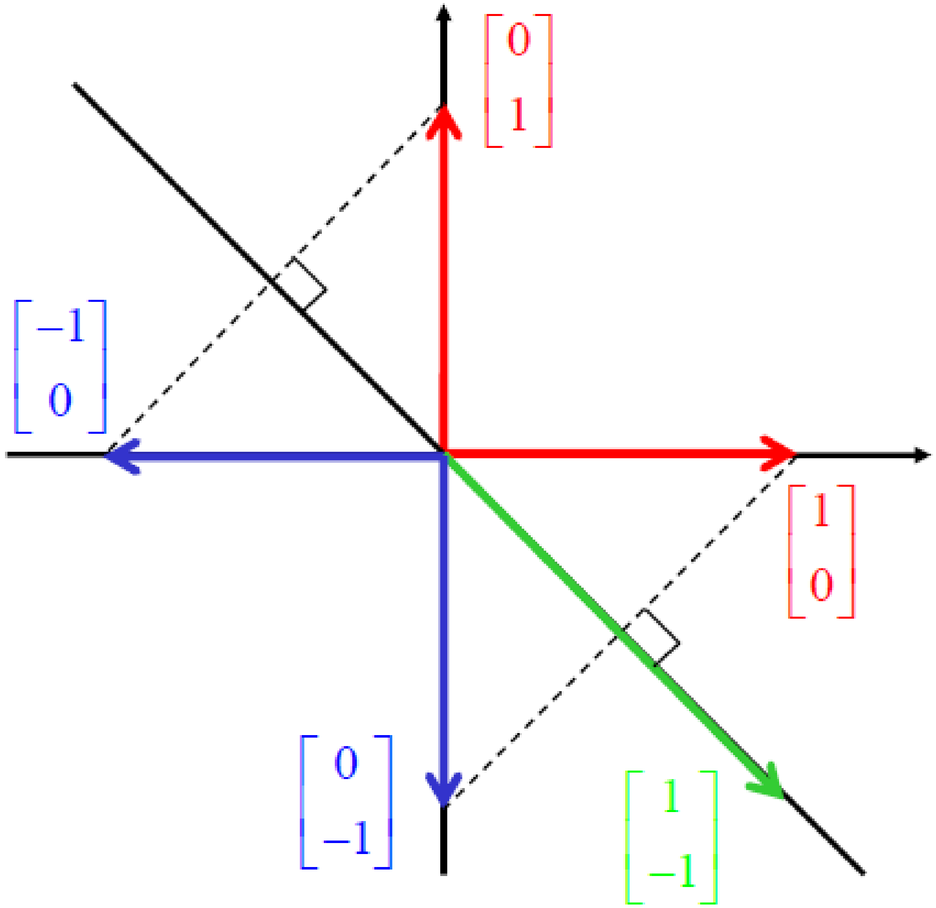
$$\mathbf{A}_\alpha \mathbf{A}_{-\alpha} = \mathbf{I}$$

$$\mathbf{A}_\alpha \mathbf{A}_\beta = \mathbf{A}_{\alpha+\beta}$$

Reflection

Example: We want to create a matrix **A** that reflects vectors onto in the space (line) spanned by the vector [1 -1]^T.

First determine what hapens to the basis vectors.



This gives us the two columns of **A** directly:

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

Sanity check:

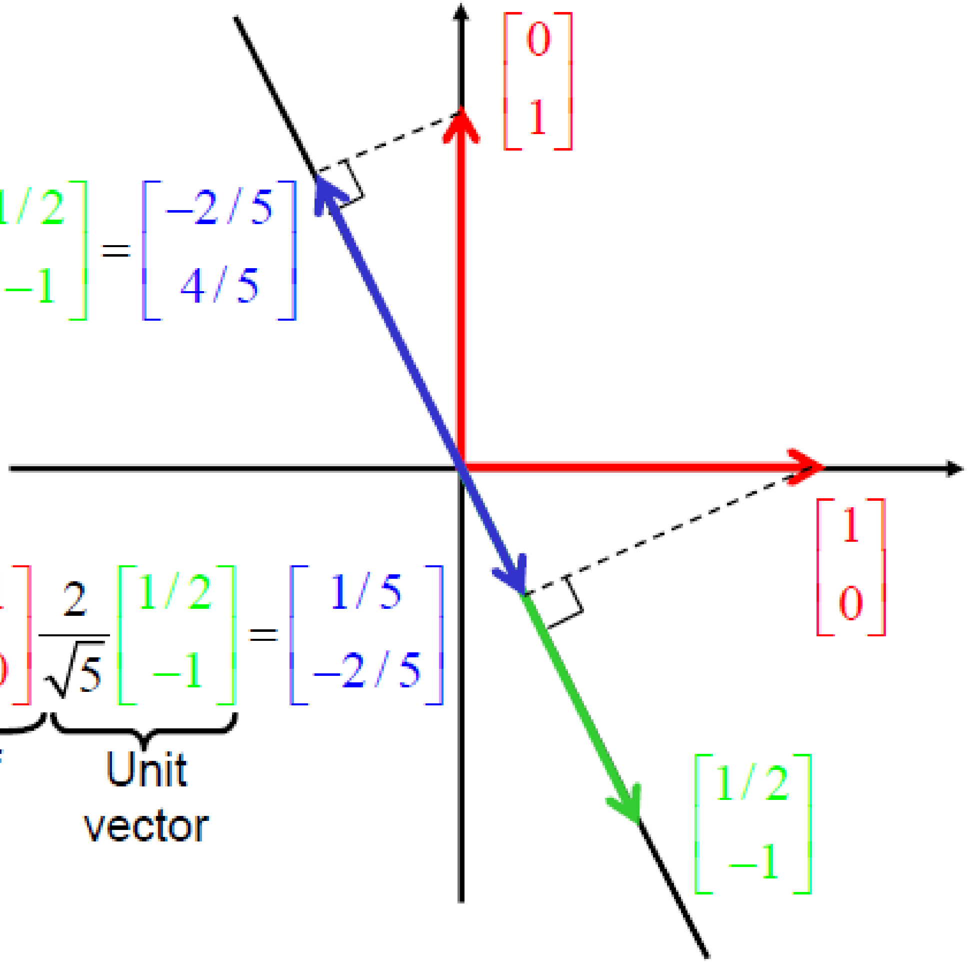
$$\mathbf{A}^2 = \mathbf{I}$$

Projection

Example: We want to create a matrix **A** that projects vectors onto the space (line) spanned by the vector [1/2 -1]^T.

First determine what hapens to the basis vectors.

$$\frac{2}{\sqrt{5}}\begin{bmatrix}1/2\\-1\end{bmatrix}^T\begin{bmatrix}0\\1\end{bmatrix}\frac{2}{\sqrt{5}}\begin{bmatrix}1/2\\-1\end{bmatrix}=\begin{bmatrix}-2/5\\4/5\end{bmatrix}$$



This gives us the two columns of **A** directly:

$$\mathbf{A}=\begin{bmatrix}1/5 & -2/5\\-2/5 & 4/5\end{bmatrix}$$

$$\underbrace{\frac{2}{\sqrt{5}}\begin{bmatrix}1/2\\-1\end{bmatrix}^T\begin{bmatrix}1\\0\end{bmatrix}}_{\text{"Length" of projection}}\underbrace{\frac{2}{\sqrt{5}}\begin{bmatrix}1/2\\-1\end{bmatrix}}_{\text{Unit vector}}=\begin{bmatrix}1/5\\-2/5\end{bmatrix}$$

Sanity check:

$$\mathbf{A}^k=\mathbf{A}$$

Diagonalization

Eigendecomposition, eigenvalues and eigenvectors.

Diagonalization

$$\mathbf{A} = \begin{bmatrix} | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_N \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} - & \mathbf{u}_1^H & - \\ & \vdots & \\ - & \mathbf{u}_N^H & - \end{bmatrix}$$

Diagonalization

An n-square matrix A is similar to a diagonal matrix D if and only if A has n linearly independent eigenvectors. In this case, the diagonal elements of D are the corresponding eigenvalues and $D = P^{-1}AP$, where P is the matrix whose columns are the eigenvectors.

Diagonalization Suppose the n by n matrix A has n linearly independent eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. Put them into the columns of an ***eigenvector matrix*** $\boldsymbol{X}$. Then $X^{-1}AX$ is the ***eigenvalue matrix*** Λ:

Eigenvector matrix X
Eigenvalue matrix Λ

$$X^{-1}AX = \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}. \qquad (1)$$

Diagonalizing a Matrix

1 The columns of $AX = X\Lambda$ are $A\boldsymbol{x}_k = \lambda_k \boldsymbol{x}_k$. The eigenvalue matrix Λ is diagonal.

2 $\boldsymbol{n}$ independent eigenvectors in X diagonalize A $\boxed{\boldsymbol{A = X\Lambda X^{-1}} \text{ and } \boldsymbol{\Lambda = X^{-1}AX}}$

3 The eigenvector matrix X also diagonalizes all powers A^k: $\boxed{\boldsymbol{A^k = X\Lambda^k X^{-1}}}$

4 Solve $\boldsymbol{u}_{k+1} = A\boldsymbol{u}_k$ by $\boldsymbol{u}_k = A^k\boldsymbol{u}_0 = X\Lambda^k X^{-1}\boldsymbol{u}_0 =$ $\boxed{\boldsymbol{c_1(\lambda_1)^k x_1} + \cdots + \boldsymbol{c_n(\lambda_n)^k x_n}}$

5 No equal eigenvalues $\Rightarrow$ X is invertible and A can be diagonalized.

Equal eigenvalues $\Rightarrow$ A *might* have too few independent eigenvectors. Then X^{-1} fails.

6 Every matrix $C = B^{-1}AB$ has the **same eigenvalues** as A. These C's are "**similar**" to A.

Diagonalizing a Matrix

Not all *NxN* matrices are diagonalizable, since not all posesses *N* linearly independent eigenvectors.

Example: The matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Triangular with zeros on the diagonal:
$\lambda_1 = \lambda_2 = 0$
(algebraic multiplicity 2)

is "defective" in the sense that it is of size 2x2 but only has "one" eigenvector.

Do not match!

The non-zero solutions to

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

are all on the form

$$\mathbf{x} = \begin{bmatrix} c \\ 0 \end{bmatrix}$$

All eigenvectors are multiples of $[\ 1\ 0\ ]^T$.
(geometric multiplicity 1)

Applications

- Powers of a Matrix
- Principal component analysis
- Markov Chains
- Power Control
- Min-Max algorithms
- Signal subspace in Random signal processing
- Whitening (decoupling) of correlated signals/noise

Example

Example 1 This A is triangular so its eigenvalues are on the diagonal: $\lambda = 1$ and $\lambda = 6$.

Eigenvectors go into X $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{5} \\ 0 & \mathbf{6} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{6} \end{bmatrix}$$

$$\boldsymbol{X^{-1} \quad A \quad X \quad = \quad \Lambda}$$

In other words $A = X\Lambda X^{-1}$. Then watch $A^2 = X\Lambda X^{-1} X\Lambda X^{-1}$. So $\boldsymbol{A^2}$ **is** $\boldsymbol{X\Lambda^2 X^{-1}}$.

A^2 has the same eigenvectors in X and squared eigenvalues in Λ^2.

Questions?