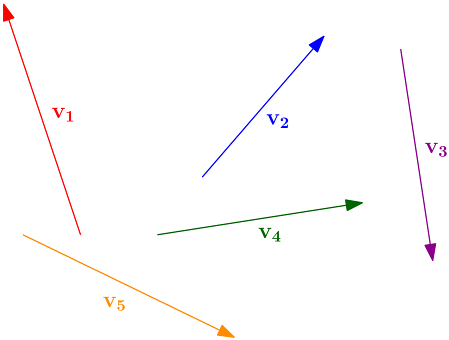
Linear Algebra Review

- Vector
- Vector Operations
- Linear Combination
- Span, Bases, Linear Independence
- Length of Vectors
- Dot Product
- Angle between Vectors
- Projection
- Linear Functions
- Linear Transformation
- Scaling, Mirror, Shear, Rotation, Projection
- Composition of Linear Transformations
- Determinent and Inverse
- Change of Bases
- Transformation in Different Bases
- Eigenvalue and Eigenvectors
- Eigenbases and Diagnoalization
- Power of Matrices
- Random Walk and Markov Chain

Imdad ullah Khan

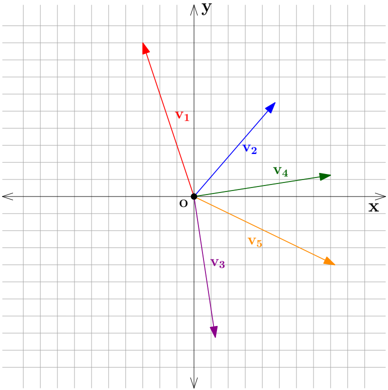
Vector

- Arrows in n-dim space $\mathbb{R}^n$
- Objects with length and directions
- Technically, they are called *free vectors*



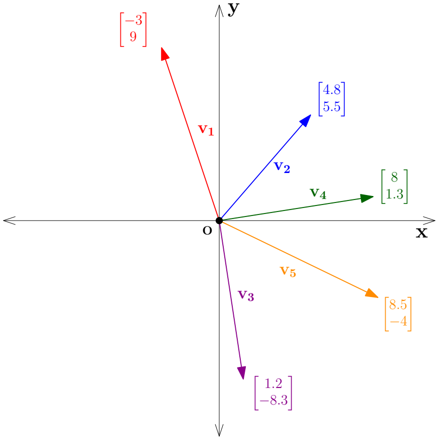
Vector

- A coordinate system ▷ Origin and unit length defined
- Look at vectors with tails fixed at the origin
- Displacement in coordinates from the origin



Vector

- A sequence of n numbers, array (ordered list) of numbers
- Bijection: n-length real sequences $\leftrightarrow$ fixed-tail arrows



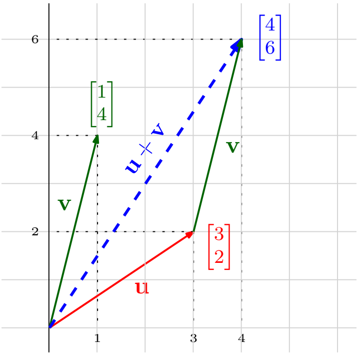
Vector

- *n*-dimensional objects
- Pairwise addition, well defined
- Scalar multiplication (multiplication with real number)

$$\underset{\mathbf{A}}{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}} + \underset{\mathbf{B}}{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}} = \underset{\mathbf{A}+\mathbf{B}}{\begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix}} \qquad x \underset{\mathbf{A}}{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}} = \underset{x\mathbf{A}}{\begin{bmatrix} xa_1 \\ xa_2 \\ \vdots \\ xa_n \end{bmatrix}}$$

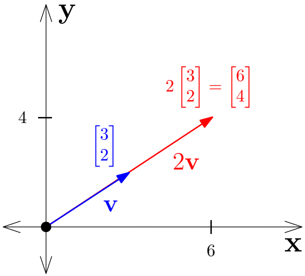
Vector Operations: Addition

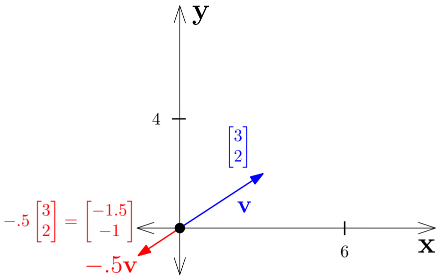
- Vectors addition defined numerically $\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1+v_1 \\ \vdots \\ u_n+v_n \end{bmatrix}$
- Geometrically it is the cumulative displacement from origin in each dimension by following the vectors with tip-to-tail joining



Vector Operations: Scaling

- Vector scaling defined numerically $x \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} xu_1 \\ \vdots \\ xu_n \end{bmatrix}$
- Geometrically it is the arrow scaled by a factor of x

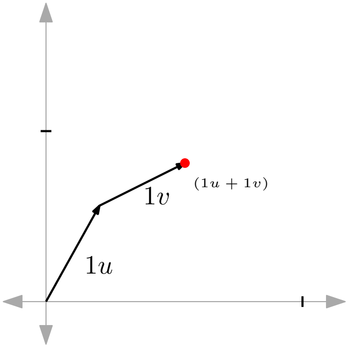




- Vectors subtraction is just combining scaling and addition

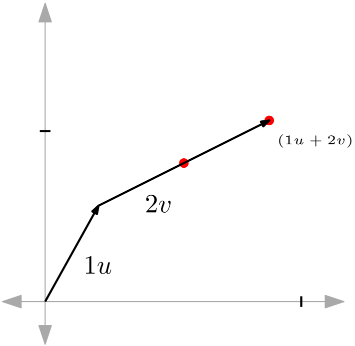
Vector Operations: Linear Combination

- Algebraically and geometrically a combination of scaling & addition
- $x \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + y \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} xu_1+yv_1 \\ \vdots \\ xu_n+yv_n \end{bmatrix}$
- linear combination $\because$ for fixed x and changing y, $x\mathbf{u} + y\mathbf{v}$ gives a line



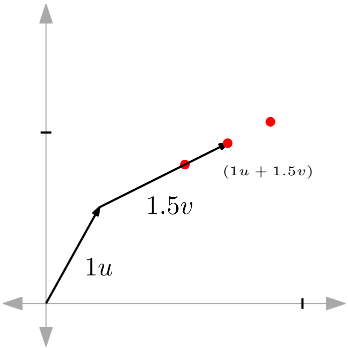
Vector Operations: Linear Combination

- Algebraically and geometrically a combination of scaling & addition
- $x \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + y \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} xu_1+yv_1 \\ \vdots \\ xu_n+yv_n \end{bmatrix}$
- linear combination $\because$ for fixed x and changing y, $x\mathbf{u} + y\mathbf{v}$ gives a line



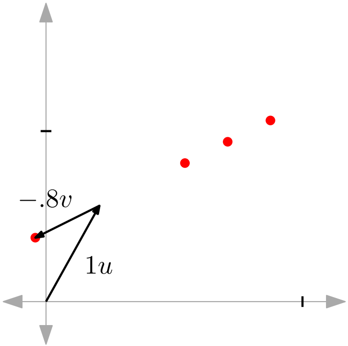
Vector Operations: Linear Combination

- Algebraically and geometrically a combination of scaling & addition
- $x \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + y \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} xu_1 + yv_1 \\ \vdots \\ xu_n + yv_n \end{bmatrix}$
- linear combination $\because$ for fixed x and changing y, $x\mathbf{u} + y\mathbf{v}$ gives a line



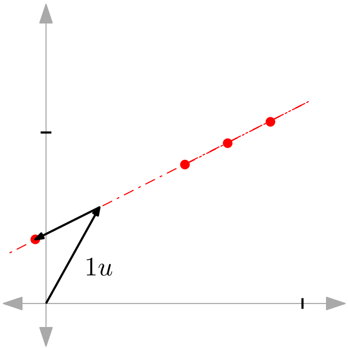
Vector Operations: Linear Combination

- Algebraically and geometrically a combination of scaling & addition
- $x \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + y \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} xu_1+yv_1 \\ \vdots \\ xu_n+yv_n \end{bmatrix}$
- linear combination $\because$ for fixed x and changing y, $x\mathbf{u} + y\mathbf{v}$ gives a line



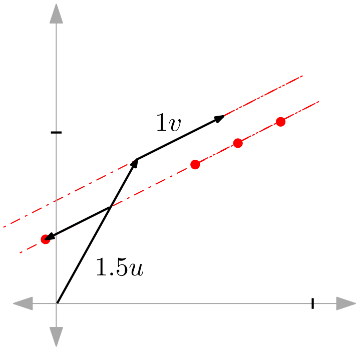
Vector Operations: Linear Combination

- Algebraically and geometrically a combination of scaling & addition
- $x \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + y \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} xu_1+yv_1 \\ \vdots \\ xu_n+yv_n \end{bmatrix}$
- linear combination $\because$ for fixed x and changing y, $x\mathbf{u} + y\mathbf{v}$ gives a line



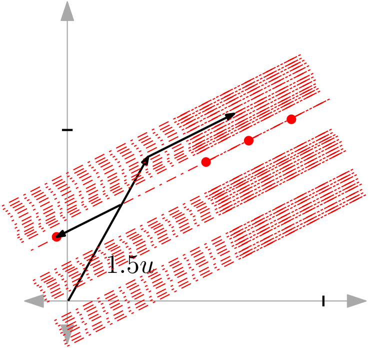
Vector Operations: Linear Combination

- Algebraically and geometrically a combination of scaling & addition
- $x \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + y \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} xu_1 + yv_1 \\ \vdots \\ xu_n + yv_n \end{bmatrix}$
- linear combination $\because$ for fixed x and changing y, $x\mathbf{u} + y\mathbf{v}$ gives a line



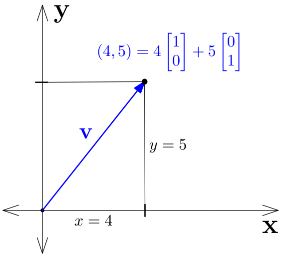
Vector Operations: Linear Combination

- Algebraically and geometrically a combination of scaling & addition
- $x \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + y \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} xu_1+yv_1 \\ \vdots \\ xu_n+yv_n \end{bmatrix}$
- linear combination $\because$ for fixed x and changing y, $x\mathbf{u} + y\mathbf{v}$ gives a line



Vector: Standard Bases

- $\mathbf{e_1} = \hat{\mathbf{i}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e_2} = \hat{\mathbf{j}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are standard basis vectors in $\mathbb{R}^2$
- A vector $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ is two scalars expressing how much this vector scales the standard basis vectors $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$
- Each vector in $\mathbb{R}^2$ is a linear combination of $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$

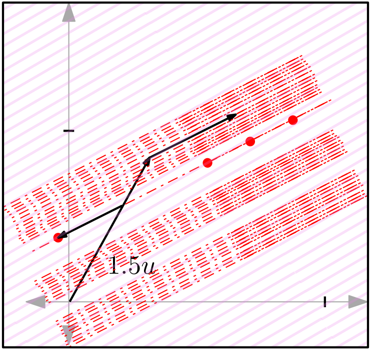


Vector: Standard Bases

- In $\mathbb{R}^n$, $\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\mathbf{e_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, $\ldots, \mathbf{e_n} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$
- The standard bases are unit vectors along the axes
- A vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$ is $\mathbf{v} = v_1\mathbf{e_1} + v_2\mathbf{e_2} + \ldots + v_n\mathbf{e_n}$

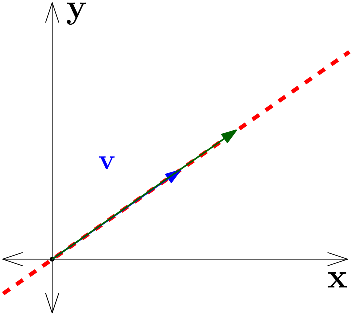
Vector: Different Bases

- Take two different vectors **u** and **v** ($\neq \hat{\mathbf{i}}, \hat{\mathbf{j}}$)
- Consider all linear combinations of **u** and **v**
- Try all combinations of scalars x and y, and check $x\mathbf{u} + y\mathbf{v}$
- Which vectors can you get? In most cases, you get all vectors in $\mathbb{R}^2$



Vector: Span, Bases and linear independence

- Take two different vectors **u** and **v** ($\neq \hat{\mathbf{i}}, \hat{\mathbf{j}}$)
- Span: space of vectors we get as linear combination of **u** and **v**
- Generally it is $\mathbb{R}^2$, or **u** and **v** line up $\implies$ it is a 1-dim subspace of $\mathbb{R}^2$
- **u** and **v** are linearly dependent, otherwise linearly independent
- Or when $\mathbf{u} = \mathbf{v} = \mathbf{0}$, then we get 0-dim subspace

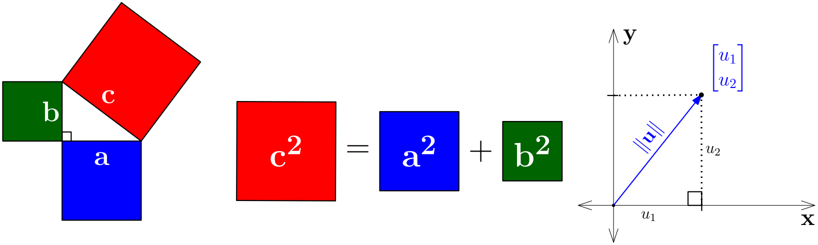


Vector: Span, Bases and linear independence

- Span of a vector $\mathbf{v} \in \mathbb{R}^2$ (actually any space) is a line (unless $\mathbf{v} = \mathbf{0}$)
- Span of 2 vectors in $\mathbb{R}^3$ is a plane (unless they line up)
- Span of 3 vectors in $\mathbb{R}^3$ is the whole $\mathbb{R}^3$ (unless one vector is in the plane spanned by the other two)
- Technically given k vectors if a vector can be removed without reducing the span, then they are linearly dependent
- That is if one vector can be expressed as linear combination of the others, then they are linearly dependent
- Otherwise, they are linearly independent, *every vector really add another dimension*
- Basis of a vector space (or a space) is a set of linearly independent vectors that spans the whole space

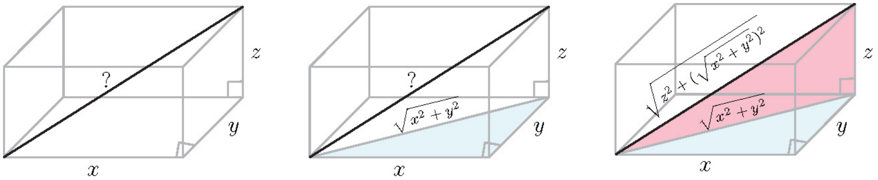
Vector: Length of vectors

- Length of **u**, denoted by $\|\mathbf{u}\|$, comes from the Pythagoras theorem



- $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, then $\|\mathbf{u}\|^2 = u_1^2 + u_2^2$
- $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2}$
- For $\mathbf{u} \in \mathbb{R}^n$, $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2} = \sqrt{\sum_{i=1}^{n} u_i^2}$
- By inductively applying the Pythagoras theorem $n - 1$ times

Vector: Length of vectors



- $\mathbf{u} = \begin{bmatrix} x \ y \ z \end{bmatrix}^T$ is diagonal of the cube
- $\mathbf{u}' = \begin{bmatrix} x \ y \ 0 \end{bmatrix}^T$ is a vector in the $x - y$ plane
- length of base and perpendicular is u_1 and u_2, so $\|\mathbf{u}\| = \sqrt{x^2 + y^2}$
- $\mathbf{u}$ makes a right triangle $\mathbf{u}'$ (base) and $\begin{bmatrix} 0 \ 0 \ z \end{bmatrix}$ (perpendicular)
- So $\|\mathbf{u}\| = \sqrt{\|\mathbf{u}'\|^2 + z^2} = \sqrt{x^2 + y^2 + y^2}$
- by a second application of the Pythagoras theorem

Vector: Unit Vector

- A vector $\mathbf{u}$ is called a unit vector, if $\|\mathbf{u}\| = 1$
- For any vector $\mathbf{u}$ we can get the unit vector in the direction of $\mathbf{u}$ by scaling it to have length 1

$$\hat{\mathbf{u}} = \frac{\mathbf{u}}{\|\mathbf{u}\|}$$

- Verify that $\hat{\mathbf{u}}$ has length 1

Vector: Dot Product

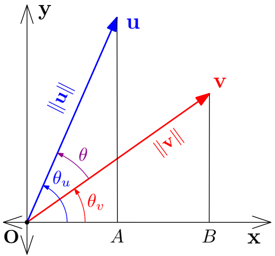
$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^t \mathbf{v} = \begin{bmatrix} u_1 & \dots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + \ldots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

- It takes two vectors and returns a scalar (function)
- Also called inner product, scalar product, projection product
- Many names because it is a really fundamental operation
- Many concepts can be expressed in terms of dot-product
- Note that $\langle \mathbf{u}, \mathbf{u} \rangle = \mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2$ (length of vectors from dot-product)

Angle between vectors

- Angle θ between vectors $\mathbf{u}$ and $\mathbf{v}$ is related to their dot-product
- Let $\mathbf{u}$ and $\mathbf{v}$ make angles θ_u and θ_v resp. with $\mathbf{e_1}$ or x-axis

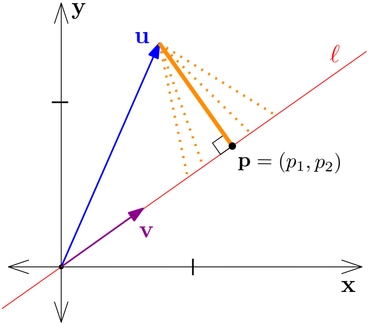
- From $\triangle OAU$
  - $\sin\theta_u = \frac{u_2}{\|\mathbf{u}\|}$ and $\cos\theta_u = \frac{u_1}{\|\mathbf{u}\|}$
- From $\triangle OBV$
  - $\sin\theta_v = \frac{v_2}{\|\mathbf{v}\|}$ and $\cos\theta_v = \frac{v_1}{\|\mathbf{v}\|}$



- $\cos\theta \; = \; \cos(\theta_u - \theta_v) \; = \; \cos\theta_v \cos\theta_u + \sin\theta_v \sin\theta_u$
- $\cos\theta \; = \; \frac{u_1}{\|\mathbf{u}\|}\frac{v_1}{\|\mathbf{v}\|} + \frac{u_2}{\|\mathbf{u}\|}\frac{v_2}{\|\mathbf{v}\|} \; = \; \frac{u_1 v_1 + u_2 v_2}{\|\mathbf{u}\|\|\mathbf{v}\|} \; = \; \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}$
- What happens if we (negatively) scale one or both vectors

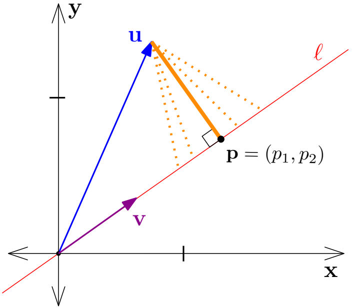
Projection

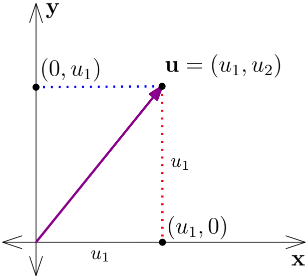
- Let **v** be a unit vector, let ℓ be a line in the direction of **v**
- Find the point **p** on ℓ that is closest to a vector **u**
- The line connecting **u** to **p** is perpendicular to **v**
- Otherwise **p** will not be the closest point (Pythagoras theorem)
- The point (vector) **p** is called the the projection of **u** on **v**



Projection

- Let **v** be a unit vector, let ℓ be a line in the direction of **v**
- Find the point **p** on ℓ that is closest to a vector **u**
- The point (vector) **p** is called the the projection of **u** on **v**
- The line connecting **u** to **p** is perpendicular to **v**
- Finding projection of **v** on the standard basis vectors is easy

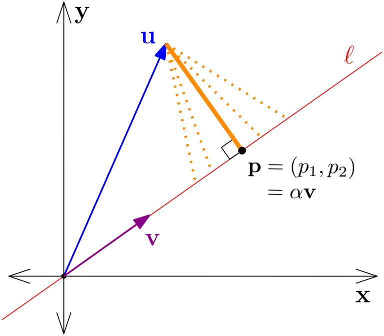




Dot product and Projection

- Find the projection **p** of **u** on **v**
- For general vectors we derive it from dot product
- **p** is just scaled vector **v**, $p = a\mathbf{v}$, find that scalar a

- $\mathbf{u} - p = \mathbf{u} - a\mathbf{v}$ is perpendicular on **v**
  - $\mathbf{v} \cdot \mathbf{v} - a\mathbf{v} = 0$
- Hence $\mathbf{v} \cdot \mathbf{u} - \mathbf{v} \cdot a\mathbf{v} = \mathbf{v} \cdot \mathbf{u} - a\mathbf{v} \cdot \mathbf{v} = 0$
- Which means $a\mathbf{v} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- $a = \dfrac{\mathbf{v} \cdot \mathbf{u}}{\mathbf{v} \cdot \mathbf{v}} = \dfrac{\mathbf{v} \cdot \mathbf{u}}{\|\mathbf{v}\|}$



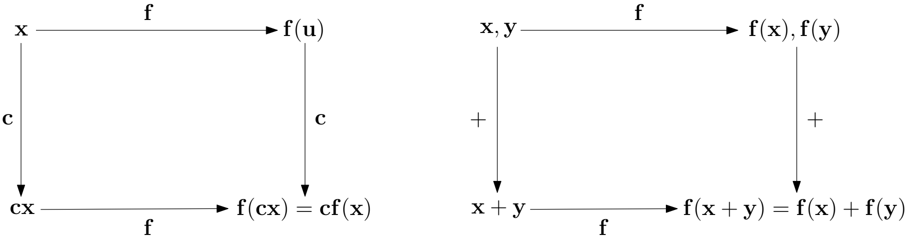
Orthogonal Vectors, Orthonormal Basis

- $\mathbf{u}$ and $\mathbf{v}$ are called **orthogonal**, if $\mathbf{u} \cdot \mathbf{v} = 0$
- They are perpendicular to each other, angle θ between them is 90°
- $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta = \|\mathbf{u}\|\|\mathbf{v}\| \cos 90^\circ = 0$
- If $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, then they are linearly independent
- If $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ are pairwise orthogonal, they are all linearly independent
- If bases of a space are all pairwise orthogonal, then they are called orthogonal bases
- If they are unit vectors, they are called **orthonormal basis**
- Verify that the standard bases make orthonormal bases of $\mathbb{R}^n$

Linear Functions

A function $f : \mathbb{R} \mapsto \mathbb{R}$ is called linear if

1. $f(cx) \; = \; c\,f(x)$
2. $f(x+y) \; = \; f(x) + f(y)$



Linear Functions

A function $f : \mathbb{R} \mapsto \mathbb{R}$ is called linear if

1. $f(cx) = cf(x)$
2. $f(x + y) = f(x) + f(y)$

A shorter version:

A function $f : \mathbb{R} \mapsto \mathbb{R}$ is linear if

1. $f(ax + by) = af(x) + bf(x)$

- These imply that $f(0) = 0$
- Generally, functions of the form $g(x) = ax + b$ are called linear, which doesn't necessarily imply $g(0) = 0$
- Functions like $g(\cdot)$ are technically and correctly called affine functions, which are linear functions followed by a translation

Dot Product as Linear Functions

For a fixed vector $\mathbf{a} \in \mathbb{R}^n$, define $f_{\mathbf{a}} : \mathbb{R}^n \mapsto \mathbb{R}^1$ as follows

$$f_{\mathbf{a}}(\mathbf{x}) \;:=\; \langle \mathbf{a}, \mathbf{x} \rangle \;=\; \mathbf{a} \cdot \mathbf{x}$$

$f_{\mathbf{a}}$ is a linear function from $\mathbb{R}^n$ to $\mathbb{R}^1$

In fact, it can be shown that these are the only functions that are linear

$$\mathbf{a} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$f_{\mathbf{a}}(4\mathbf{x} + 5\mathbf{y}) \;=\; \begin{bmatrix} 3 \\ 4 \end{bmatrix} \cdot \left(\begin{bmatrix} 4 * 2 \\ 4 * 3 \end{bmatrix} + \begin{bmatrix} 5 * 1 \\ 5 * 2 \end{bmatrix} \right) = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 13 \\ 22 \end{bmatrix} = 39 + 88 = 127$$

$$4f_{\mathbf{a}}(\mathbf{x}) + 5f_{\mathbf{a}}(\mathbf{y}) \;=\; 4 * \underbrace{\begin{bmatrix} 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix}}_{18} + 5 * \underbrace{\begin{bmatrix} 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{11} = 4 * 18 + 5 * 11 = 127$$

Linear Functions on Euclidean Space

- For linear functions of the form $\mathbb{R}^n \mapsto \mathbb{R}^m$, for $m > 1$

  ▷ vector functions - functions that output vectors in $\mathbb{R}^m$

- Extend the notion of dot product as linear function as follows:

For m fixed vectors $\mathbf{a_1}, \mathbf{a_2}, \ldots \mathbf{a_m} \in \mathbb{R}^n$, define $f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}} : \mathbb{R}^n \mapsto \mathbb{R}^m$ as:

$$f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}}(\mathbf{x}) \; := \; \begin{bmatrix}\mathbf{a_1}\cdot\mathbf{x} & \mathbf{a_2}\cdot\mathbf{x} & \ldots & \mathbf{a_m}\cdot\mathbf{x}\end{bmatrix}^T$$

$f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}}$ is a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$

- Again it can be shown that these are the only functions that are linear

- $f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}}$ is represented by $m \times n$ matrix $T_f = \begin{bmatrix} \text{---} & \mathbf{a_1} & \text{---} \\ \text{---} & \mathbf{a_2} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a_m} & \text{---} \end{bmatrix}$

- Evaluated by matrix-vector product $f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}}(\mathbf{x}) = T_f\,\mathbf{x}$

  ▷ $T_f\,\mathbf{x}$ is $n \times 1$ vector

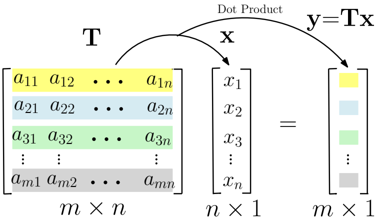
Linear Functions on Euclidean Spaces

For m fixed vectors $\mathbf{a_1}, \mathbf{a_2}, \ldots \mathbf{a_m} \in \mathbb{R}^n$, define $f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}} : \mathbb{R}^n \mapsto \mathbb{R}^m$ as:

$$f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}}(\mathbf{x}) := \begin{bmatrix} \mathbf{a_1} \cdot \mathbf{x} & \mathbf{a_2} \cdot \mathbf{x} & \ldots & \mathbf{a_m} \cdot \mathbf{x} \end{bmatrix}^T$$

$f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}}$ is a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$

- $f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}}$ is represented by $m \times n$ matrix $T_f = \begin{bmatrix} — & \mathbf{a_1} & — \\ — & \mathbf{a_2} & — \\ & \vdots & \\ — & \mathbf{a_m} & — \end{bmatrix}$
- Evaluated by matrix-vector multiplication $f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}}(\mathbf{x}) = T_f\,\mathbf{x}$



Linear Functions on Euclidean Spaces

For m fixed vectors $\mathbf{a_1}, \mathbf{a_2}, \ldots \mathbf{a_m} \in \mathbb{R}^n$, define $f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}} : \mathbb{R}^n \mapsto \mathbb{R}^m$ as:

$$f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}}(\mathbf{x}) := \begin{bmatrix} \mathbf{a_1} \cdot \mathbf{x} & \mathbf{a_2} \cdot \mathbf{x} & \ldots & \mathbf{a_m} \cdot \mathbf{x} \end{bmatrix}^T$$

- $T_f = \begin{bmatrix} \text{---} & \mathbf{a_1} & \text{---} \\ \text{---} & \mathbf{a_2} & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{a_m} & \text{---} \end{bmatrix}$ $\qquad f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}}(\mathbf{x}) = T_f\,\mathbf{x}$

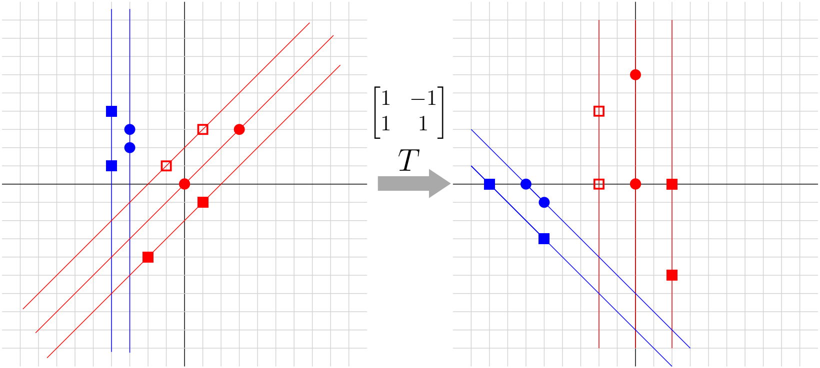
$$\mathbf{a_1} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \ \mathbf{a_2} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad T = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}$$

$$T(4\mathbf{x} + 5\mathbf{y}) = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \left(\begin{bmatrix} 8 \\ 12 \end{bmatrix} + \begin{bmatrix} 5 \\ 10 \end{bmatrix} \right) = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 13 \\ 22 \end{bmatrix} = \begin{bmatrix} 127 \\ 48 \end{bmatrix}$$

$$4T\mathbf{x} + 5T\mathbf{y} = 4 \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} + 5 \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 72 \\ 28 \end{bmatrix} + \begin{bmatrix} 55 \\ 20 \end{bmatrix} = \begin{bmatrix} 127 \\ 48 \end{bmatrix}$$

Linear Functions on Euclidean Spaces

- Geometrically, linear functions (matrix-vector multiplications)
- maps the $\mathbf{0}$ vector (origin) to $\mathbf{0}$
- maps any straight line to a straight lines
- maps any set of parallel lines to a set of parallel lines



Matrices as Linear Transform

- Linear functions, multiplications of $m \times n$ matrices with $n \times 1$ vectors output $m \times 1$ vectors
- For any $m \times n$ matrix T, $\mathbf{y} = T\mathbf{x}$ is a linear function $\mathbb{R}^n \mapsto \mathbb{R}^m$
- Generally called **linear transformation**, because we are interested in **how it transforms the whole space ($\mathbb{R}^n$)**
  - and not in evaluating output on specific inputs
  - or its properties as a function (injective, surjective, bijective etc.)
- Just a few quick terminology (while we still call it functions)
- Linear functions on Euclidean space are also called linear maps
- When $m = n$ (same $\mathbb{R}^n \mapsto \mathbb{R}^n$), they are called linear operators
- When the function is bijective (the corresponding matrix is invertible), they are called linear isomorphisms

Matrices as Linear Transform

Meaning of rows of a matrix A as a linear transform

Recall standard bases of $\mathbb{R}^n$ (unit vectors along the axes)

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e_n} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

- They help write awkward and wordy things concisely and precisely

Matrices as Linear Transform: Rows

Meaning of rows of a matrix A as a linear transform

- They help write awkward and wordy things concisely and precisely
- $\mathbf{e_i}^T A$ is the i^{th} row of A $$\begin{bmatrix}0 & 1 & 0\end{bmatrix}\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix} = \begin{bmatrix}4\\5\\6\end{bmatrix}$$
- $\mathbf{e_i}^T A$ is $\mathbf{a_i}$ in the definition of the function $f_{\mathbf{a_1},\mathbf{a_2},\ldots,\mathbf{a_m}}$ corresponding to A
- $\mathbf{e_i}^T A$ describes how to compute the i^{th} coordinate of result, $\mathbf{y} = A\mathbf{x}$

  $\triangleright$ $\mathbf{y}(i) = \mathbf{e_i}^T A \cdot \mathbf{x}$

Matrices as Linear Transform

Meaning of columns of a matrix A as a linear transform

- $A\mathbf{e_i}$ is the i^{th} column of A $\quad \begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix}\begin{bmatrix}0\\1\\0\end{bmatrix} = \begin{bmatrix}2\\5\\8\end{bmatrix}$

- $A\mathbf{e_i}$ is the vector in R^n where $\mathbf{e_i}$ maps to

- So the columns of A are the locations in the range space ($\mathbb{R}^m$), where the standard bases map to by the transform A

- This is the most important concept to understand

Matrices as Linear Transform

Meaning of columns of a matrix A as a linear transform

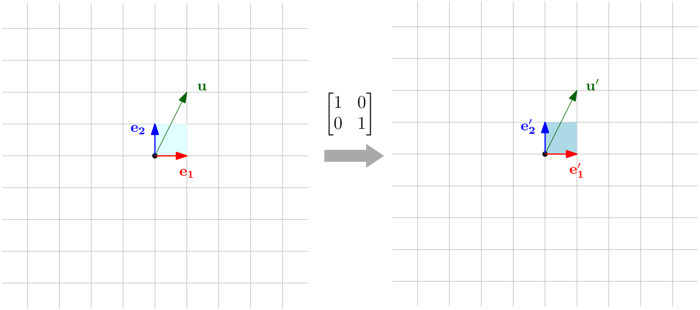
- The columns of A are the locations in the range space ($\mathbb{R}^m$), where the standard bases map to by the transform A
- A linear transform is completely described by knowing where it maps the basis vectors
- Follows from linearity, as $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is actually $\mathbf{u} = u_1\mathbf{e_1} + u_2\mathbf{e_2}$
- $A\mathbf{u} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} au_1+bu_2 \\ cu_1+du_2 \end{bmatrix}$, By linearity
- $A\mathbf{u} = A(u_1\mathbf{e_1} + u_2\mathbf{e_2}) = u_1A\mathbf{e_1} + u_2A\mathbf{e_2} = u_1 \begin{bmatrix} a \\ c \end{bmatrix} + u_2 \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} au_1+bu_2 \\ cu_1+du_2 \end{bmatrix}$
- Under A, the image of $\mathbf{u} = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}^T$ is a linear combination of images of basis vectors $(A\mathbf{e_1}, \ldots, A\mathbf{e_n})$ with coefficients $u_1, \ldots, u_n$

Common Linear Transformation

- We discuss some common transformation to master the concepts
- They are fundamental to computer graphics, image processing, computer vision and other CS disciplines
- In these fields, they mostly need affine transformation, which, as mentioned earlier, is linear transformation followed by translation
- We mainly focus on linear operators ($\mathbb{R}^n \mapsto \mathbb{R}^n$) with $n = 2$, but will mention some others to highlight certain concepts
- We discussed that a linear transformation (matrix) is completely described by its columns - images of standard bases vectors
- We will mainly just show the transformed bases vectors and the image of the 1×1 square in the first quadrant

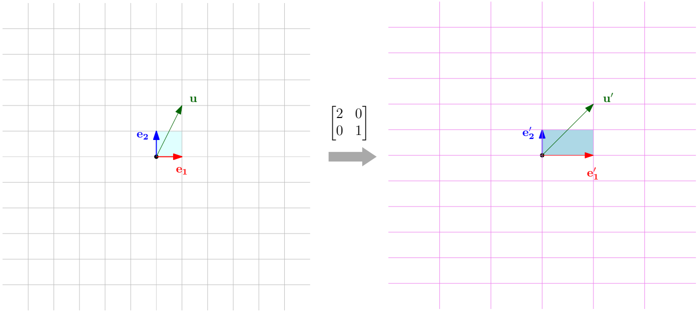
Linear Transformation: Identity

- $A = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ does not change any vectors
- $\mathbf{e}_1' = A\mathbf{e}_1 = \mathbf{e}_1$ and $\mathbf{e}_2' = A\mathbf{e}_2 = \mathbf{e}_2$
- For $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix} = x\mathbf{e}_1 + y\mathbf{e}_2, \quad A\mathbf{u} = x\mathbf{e}_1' + y\mathbf{e}_2' = \mathbf{u}$
- The space does not change, the unit square remains the same



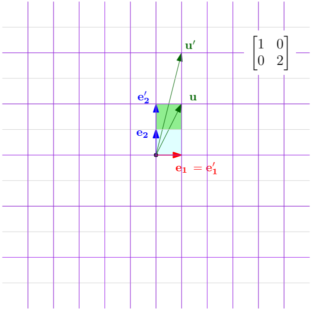
Linear Transformation: Horizontal Scaling

- $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ stretches each vector by a factor of 2 horizontally
- $\mathbf{e_1'} = A\mathbf{e_1} = 2\mathbf{e_1}$ and $\mathbf{e_2'} = A\mathbf{e_2} = \mathbf{e_2}$
- For $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix} = x\mathbf{e_1} + y\mathbf{e_2}$, $A\mathbf{u} = x\mathbf{e_1'} + y\mathbf{e_2'} = \begin{bmatrix} 2x \\ y \end{bmatrix}$
- grid changes, unit square becomes 2×1 rectangle



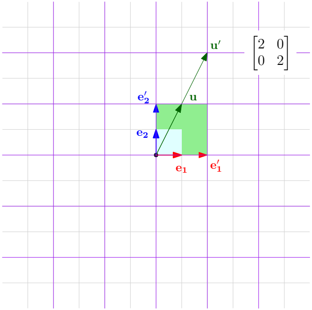
Linear Transformation: Vertical Scaling

- $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ stretches each vector by a factor of 2 vertically
- $\mathbf{e_1'} = A\mathbf{e_1} = \mathbf{e_1}$ and $\mathbf{e_2'} = A\mathbf{e_2} = 2\mathbf{e_2}$
- For $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix} = x\mathbf{e_1} + y\mathbf{e_2}$, $A\mathbf{u} = x\mathbf{e_1'} + y\mathbf{e_2'} = \begin{bmatrix} x \\ 2y \end{bmatrix}$
- grid changes, unit square becomes 1×2 rectangle



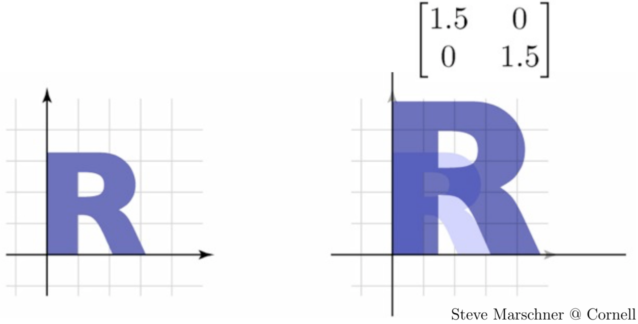
Linear Transformation: Uniform Scaling

- $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ stretches each vector by a factor of 2 in both directions
- $\mathbf{e}_1' = A\mathbf{e}_1 = 2\mathbf{e}_1$ and $\mathbf{e}_2' = A\mathbf{e}_2 = 2\mathbf{e}_2$
- For $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix} = x\mathbf{e}_1 + y\mathbf{e}_2$, $\quad A\mathbf{u} = x\mathbf{e}_1' + y\mathbf{e}_2' = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$
- grid changes, unit square is uniformly stretched by a factor of 2



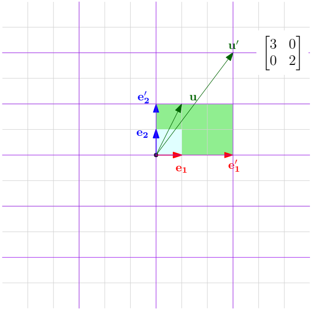
Linear Transformation: Uniform Scaling

Uniform Scaling Application



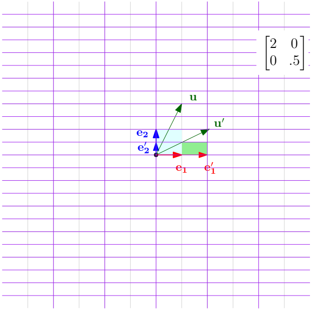
Linear Transformation: Non-Uniform Scaling

- $A = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$ stretches vectors by factors 3 and 2
- $\mathbf{e_1'} = A\mathbf{e_1} = 3\mathbf{e_1}$ and $\mathbf{e_2'} = A\mathbf{e_2} = 2\mathbf{e_2}$
- For $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix} = x\mathbf{e_1} + y\mathbf{e_2}$, $A\mathbf{u} = x\mathbf{e_1'} + y\mathbf{e_2'} = \begin{bmatrix} 3x \\ 2y \end{bmatrix}$
- grid changes, unit square becomes a 3×2 rectangle



Linear Transformation: Non-Uniform Scaling

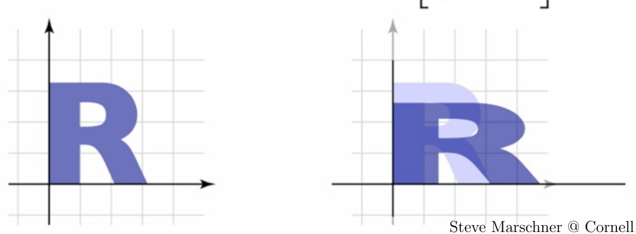
- $A = \begin{bmatrix} 2 & 0 \\ 0 & .5 \end{bmatrix}$ stretches vectors by factor of 3 and $1/2$
- $\mathbf{e_1'} = A\mathbf{e_1} = 2\mathbf{e_1}$ and $\mathbf{e_2'} = A\mathbf{e_2} = 1/2\mathbf{e_2}$
- For $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix} = x\mathbf{e_1} + y\mathbf{e_2}$, $A\mathbf{u} = x\mathbf{e_1'} + y\mathbf{e_2'} = \begin{bmatrix} 2x \\ y/2 \end{bmatrix}$
- grid changes, unit square becomes a $2 \times 1/2$ rectangle



Linear Transformation: Non-Uniform Scaling

Non-Uniform Scaling Application

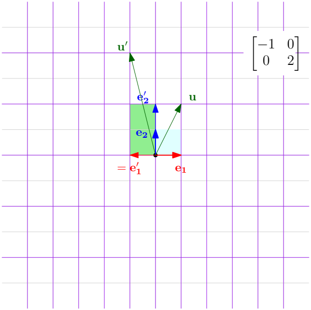
$$\begin{bmatrix} 1.5 & 0 \\ 0 & 0.8 \end{bmatrix}$$



Steve Marschner @ Cornell

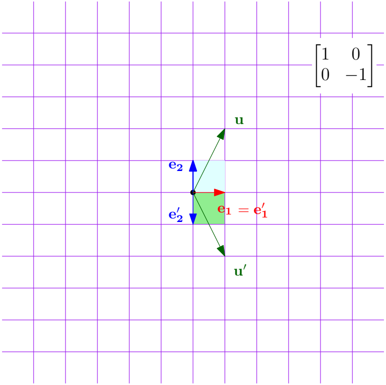
Linear Transformation: Negative Scaling

- $A = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$ stretches each vector by a factor of -1 horizontally and by a factor of 2 vertically
- $\mathbf{e_1'} = A\mathbf{e_1} = -1\mathbf{e_1}$ and $\mathbf{e_2'} = A\mathbf{e_2} = 2\mathbf{e_2}$
- For $\mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix} = x\mathbf{e_1} + y\mathbf{e_2}$, $A\mathbf{u} = x\mathbf{e_1'} + y\mathbf{e_2'} = \begin{bmatrix} -x \\ 2y \end{bmatrix}$
- grid changes, unit square becomes a 1×2 rectangle but flipped across



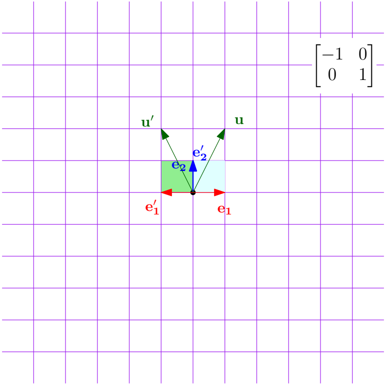
Linear Transformation: Horizontal Mirror

- $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ reflects each vector across vertical axis
- grid stays the same with different orientation, unit square is mirrored through horizontal axis



Linear Transformation: Vertical Mirror

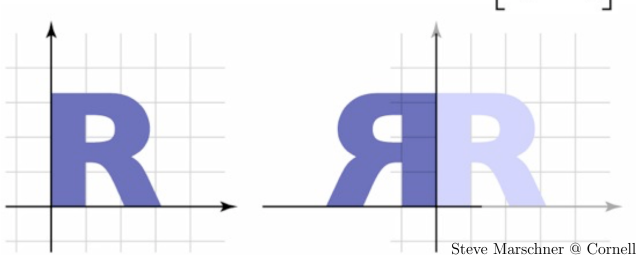
- $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ reflects each vector across vertical axis
- grid stays the same with different orientation, unit square is mirrored through horizontal axis



Linear Transformation: Vertical Mirror

Reflection/Mirror Application

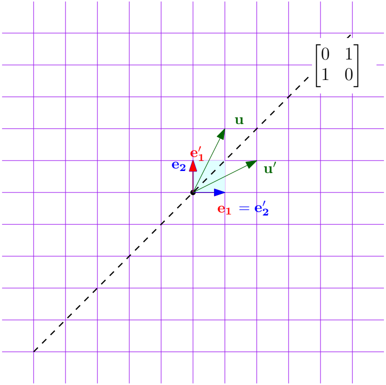
$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$



Steve Marschner @ Cornell

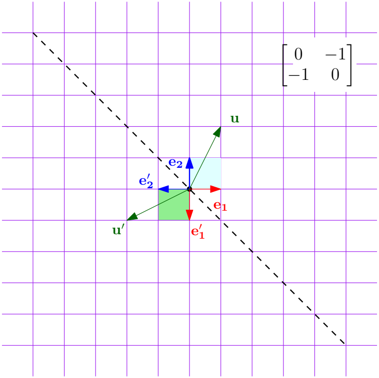
Linear Transformation: Diagonal Mirror

- $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ reflects each vector across 45° mirror
- grid stays the same with different orientation, unit square is mirrored through 45° mirror



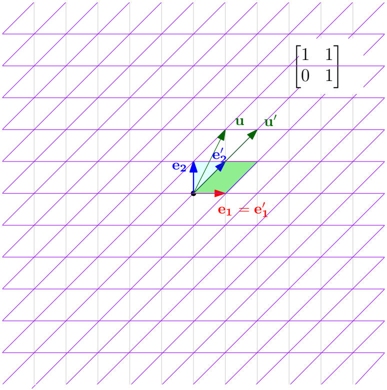
Linear Transformation: Other Diagonal Mirror

- $A = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ reflects each vector across 45˚ mirror
- grid changes, unit square is mirrored through the other diagonal mirror



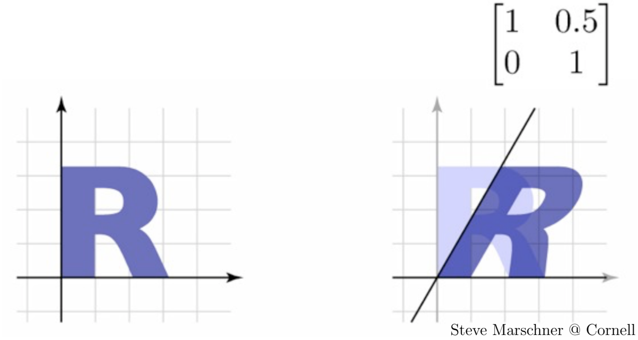
Linear Transformation: Horizontal Shear

- $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ leaves horizontal dimension intact and skew each vector in vertical dimension (horizontal shear)
- unit square becomes a parallelogram



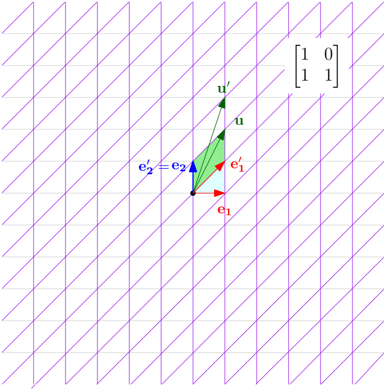
Linear Transformation: Horizontal Shear

Horizontal Shear Application



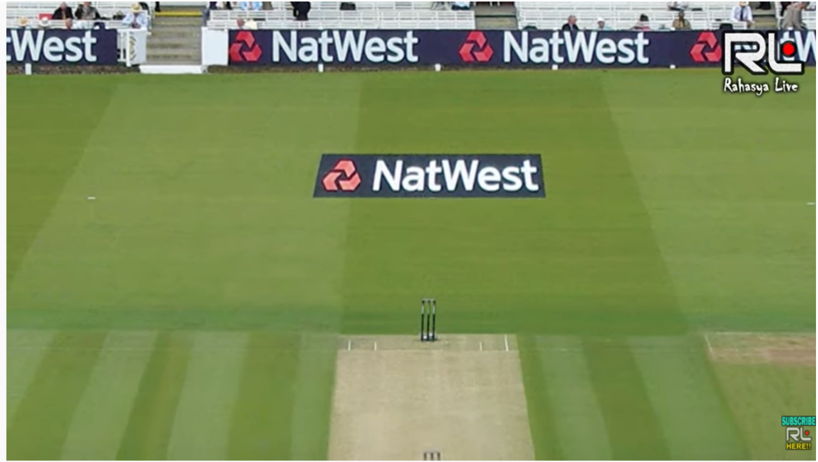
Linear Transformation: Vertical Shear

- $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ leaves vertical dimension intact and skew each vector in horizontal dimension (horizontal shear)
- unit square becomes a parallelogram



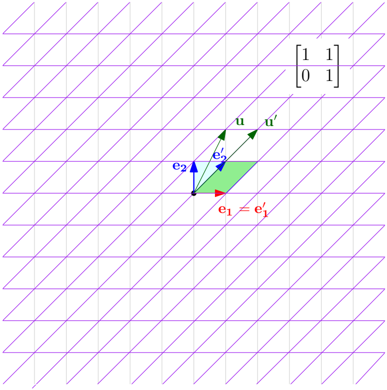
Linear Transformation: Vertical Shear

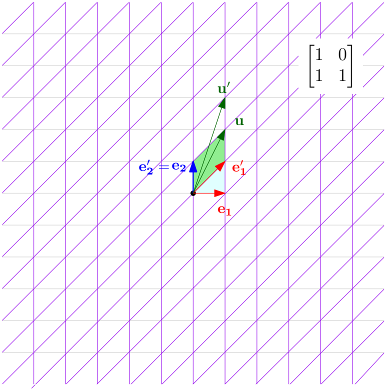
Vertical Shear Application



Linear Transformation: Shear

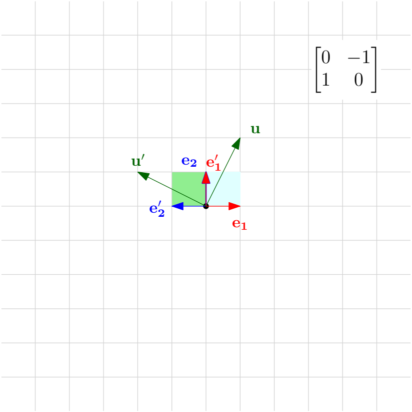
- $A = \begin{bmatrix} 1 & 1 \\ 0 & s \end{bmatrix}$ vertical shear and $A = \begin{bmatrix} s & 0 \\ 1 & 1 \end{bmatrix}$ horizontal shear
- unit square becomes a parallelogram





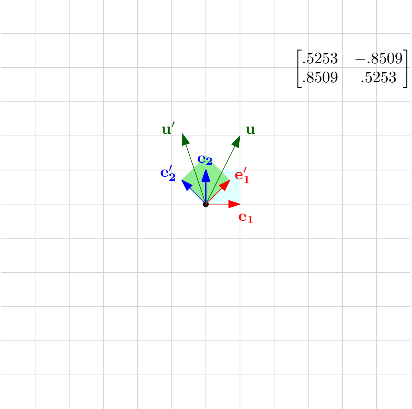
Linear Transformation: Rotation

- $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ rotates every vector by 90˚ clockwise
- unit square rotates to the adjacent unit square



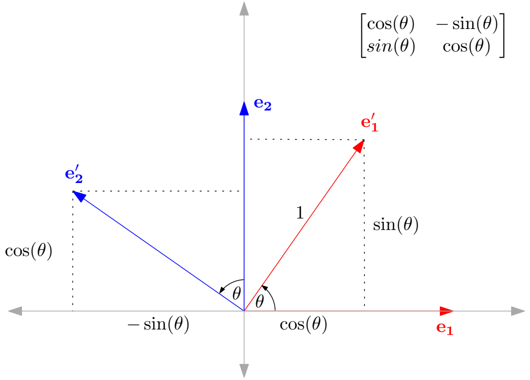
Linear Transformation: Rotation

- $A = \begin{bmatrix} .5253 & -.8509 \\ .8509 & .5253 \end{bmatrix}$ rotates every vector by 45˚ clockwise
- unit square rotates by 45˚



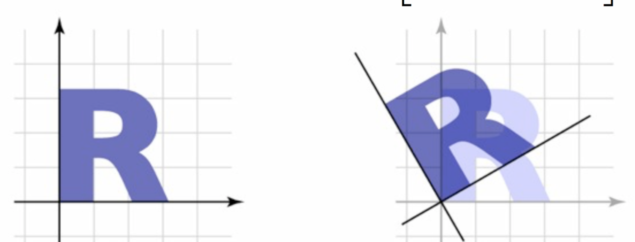
Linear Transformation: Rotation

- $A = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ sin(\theta) & \cos(\theta) \end{bmatrix}$ rotates every vector by θ clockwise
- unit square rotates by θ clockwise



Rotation Applications

$$\begin{bmatrix} 0.866 & -0.5 \\ 0.5 & 0.866 \end{bmatrix}$$

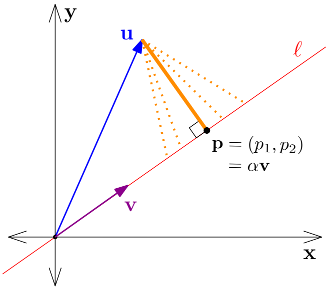


Steve Marschner @ Cornell

Linear Transformation: Projection

- Let $\mathbf{v}$ be a vector, let ℓ be a line in the direction of $\mathbf{v}$
- Projection of $\mathbf{u}$ on ℓ (or on $\mathbf{v}$) is the point $\mathbf{p}$ on ℓ that is closest to $\mathbf{u}$

- $\mathbf{p}$ is scaled vector $hat(v)$ $\quad p = a\hat{\mathbf{v}}$
  - ▷ a : scalar projection or projection length
- $\mathbf{u} - p = \mathbf{u} - a\hat{\mathbf{v}}$ is perpendicular on $\hat{\mathbf{v}}$
  - $\mathbf{v} \cdot \mathbf{v} - a\mathbf{v} = 0$
- Hence $\mathbf{v} \cdot \mathbf{u} - \mathbf{v} \cdot a\mathbf{v} = \mathbf{v} \cdot \mathbf{u} - a\mathbf{v} \cdot \mathbf{v} = 0$
- Which means $a\mathbf{v} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- $a = \dfrac{\mathbf{v} \cdot \mathbf{u}}{\mathbf{v} \cdot \mathbf{v}} = \dfrac{\mathbf{v} \cdot \mathbf{u}}{\|\mathbf{v}\|}$



The vector projection, $\mathbf{p}$ is given by $\mathbf{p} = \dfrac{\mathbf{v} \cdot \mathbf{u}}{\|\mathbf{v}\|}\hat{\mathbf{v}} = \dfrac{\mathbf{v} \cdot \mathbf{u}}{\|\mathbf{v}\|}\dfrac{\mathbf{v}}{\|\mathbf{v}\|} = \dfrac{\mathbf{v} \cdot \mathbf{u}}{\|\mathbf{v}\|^2}\mathbf{v}$

Linear Transformation: Projection

The vector projection, $\mathbf{p}$ is given by $\mathbf{p} = \frac{\mathbf{v}\cdot\mathbf{u}}{\|\mathbf{v}\|}\hat{\mathbf{v}} = \frac{\mathbf{v}\cdot\mathbf{u}}{\|\mathbf{v}\|}\frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{v}\cdot\mathbf{u}}{\|\mathbf{v}\|^2}\mathbf{v}$

- For unit vector $\hat{\mathbf{v}}$, the vector projection, $\mathbf{p}$ of $\mathbf{u}$ on $\hat{\mathbf{v}}$ is $\mathbf{p} = (\mathbf{u}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}}$

$$\mathbf{p} = (\mathbf{u}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}} = \left(\begin{bmatrix} x \\ y \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix}\right)\begin{bmatrix} a \\ b \end{bmatrix} = (xa + yb)\begin{bmatrix} a \\ b \end{bmatrix}$$

$$= \begin{bmatrix} xa^2 + yab \\ xab + yb^2 \end{bmatrix} = \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

- $A = \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}$ projects every vector onto the unit vector $\begin{bmatrix} a \\ b \end{bmatrix}$

Composition of Linear Transformation

Any image processing operation (linear) can be described as combination of the above elementary transformation

Composing transformations

- Want to transform an object, then transform it some more

$$\mathbf{u} \mapsto g(\mathbf{u}) \mapsto f(g(\mathbf{u})) \quad := \quad (f \circ g)(\mathbf{u})$$

- Represent $(f \circ g)(\cdot)$ using same representation as for f and g (matrix) ▷ ("f compose g")
- Let S and T be the corresponding matrices for f and g, resp.
- $f(\mathbf{u}) = S\mathbf{u}$ and $g(\mathbf{u}) = T\mathbf{u}$
- $f \circ g(\mathbf{u}) = ST\mathbf{u}$

Composition of Linear Transformation

90° rotation followed by horizontal shear

$$\underbrace{S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{shear} \qquad \underbrace{T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{rotation}$$

$$\mathbf{e_1'} = T\mathbf{e_1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{e_2'} = T\mathbf{e_2} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

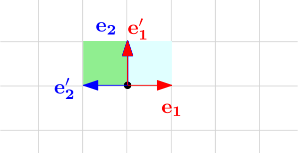
$$S\mathbf{e_1'} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$S\mathbf{e_2'} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

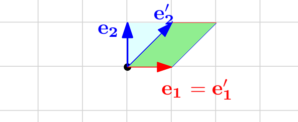
$$ST = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$$

Composition of Linear Transformation

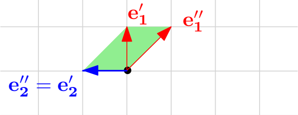
$$S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$



$$T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

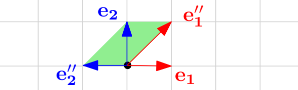


$$T\mathbf{e'_1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$



$$T\mathbf{e'_2} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

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



Composition of Linear Transformation

- Transforming first by T then by S is the same as transforming by ST
- In general, composition is not commutative
- Generally, $ST \neq TS$
- Note that $S \circ T$, is applying T first and S second
- We can compose many transformation $S \circ T \circ R$

Simultaneous Equations: Solving $A\mathbf{x} = \mathbf{b}$

Consider the following scenario

- ISB metro has 3 bridges, 4 stations, 20km length and cost is 20b
- Lahore metro has 2 bridges, 6 stations, 27km length and cost is 27b
- Multan metro has 3 bridges, 5 stations, 22km length and cost is 24b
- You want another metor with 4 bridges, 5 stations and 25km length, what will be the cost?
- If we have cost per bridge, per station, per km then we can solve it

$$\begin{aligned} 3b + 4s + 20\ell &= 20 \\ 2b + 6s + 27\ell &= 27 \\ 3b + 5s + 22\ell &= 24 \end{aligned} \implies \begin{bmatrix} 3 & 4 & 20 \\ 2 & 6 & 27 \\ 3 & 5 & 22 \end{bmatrix} \begin{bmatrix} b \\ s \\ \ell \end{bmatrix} = \begin{bmatrix} 20 \\ 27 \\ 24 \end{bmatrix} := A\mathbf{x} = \mathbf{b}$$

Which vector $\mathbf{x}$ the transformation A maps to $\mathbf{b}$? (the reverse question)

Simultaneous Equations: Solving $A\mathbf{x} = \mathbf{b}$

Solving $A\mathbf{x} = \mathbf{b}$

For a matrix A, let A^{-1} be a matrix such that

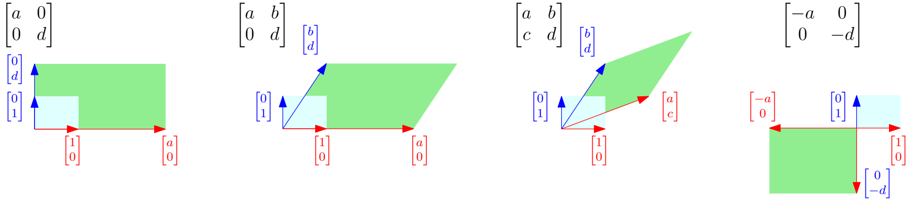
$$A^{-1}A = \mathbb{I}$$

Composing A^{-1} with A gives solution to $A\mathbf{x} = \mathbf{b}$

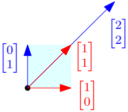
$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \implies I\mathbf{x} = A^{-1}\mathbf{b}$$

A^{-1} is called the inverse of A, if we can find it then we can solve $A\mathbf{x} = \mathbf{b}$

Linear Transformation: Determinant and Inverse



The area of this new parallelogram (the transformed unit square) *ad − bc* in 2*d* is called the determinant of the matrix *A*, *det*(*A*)



$$det\left(\begin{bmatrix}1 & 2\\1 & 2\end{bmatrix}\right) = 0$$

- Columns of *A* are linearly dependent $\implies$ determinant is 0
- This matrix is not invertible

Change of Bases

If $B = \{\mathbf{b_1}, \mathbf{b_2}, \ldots, \mathbf{b_n}\}$ is a **basis** for $\mathbb{R}^n$, then any vector $\mathbf{x} \in \mathbb{R}^n$

- can be expressed uniquely as $\mathbf{x} = \beta_1\mathbf{b_1} + \beta_2\mathbf{b_2} + \ldots + \beta_n\mathbf{b_n}$
- the scalars $\beta_1, \beta_2, \ldots, \beta_n$ are the coordinates of $\mathbf{x}$ w.r.t the basis B
- $\mathbf{x}$ is denoted by $\mathbf{x}_B = \left[\beta_1, \beta_2, \ldots, \beta_n\right]_B^T$

Let A be the standard basis, $A = \{\mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_n}\}$

Let $\mathbf{x}_A := \begin{bmatrix}\alpha_1 & \alpha_2 & \ldots & \alpha_n\end{bmatrix}_A^T$

To find coordinates of $\mathbf{x}$ w.r.t B, i.e. $\mathbf{x}_B = \begin{bmatrix}\beta_1 & \beta_2 & \ldots & \beta_n\end{bmatrix}_B^T$

Solve the linear system of equations $\mathbf{x} = \beta_1\mathbf{b_1} + \beta_2\mathbf{b_2} + \ldots + \beta_n\mathbf{b_n}$

Change of Bases

Let $\mathbf{x}_A := \begin{bmatrix}\alpha_1 & \alpha_2 & \dots & \alpha_n\end{bmatrix}_A^T$

To find coordinates of $\mathbf{x}$ w.r.t B, i.e. $\mathbf{x}_B = \begin{bmatrix}\beta_1 & \beta_2 & \dots & \beta_n\end{bmatrix}_B^T$

Solve the linear system of equations $\mathbf{x} = \beta_1\mathbf{b_1} + \beta_2\mathbf{b_2} + \ldots + \beta_n\mathbf{b_n}$

B: the matrix with basis vectors as columns, $\implies$ B is invertible

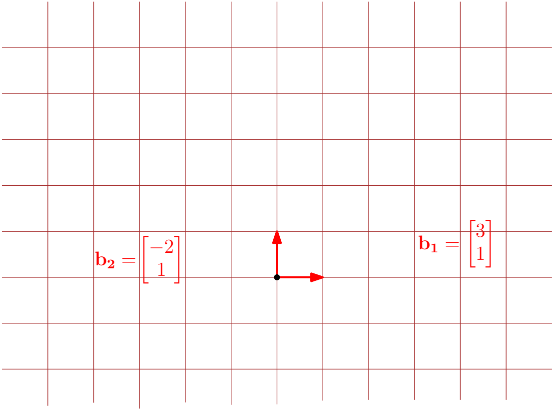
$$\begin{bmatrix} | & | & & | \\ \mathbf{b_1} & \mathbf{b_2} & \dots & \mathbf{b_n} \\ | & | & & | \end{bmatrix} \begin{bmatrix}\beta_1 \\ \vdots \\ \beta_n\end{bmatrix}_B = \begin{bmatrix}\alpha_1 \\ \vdots \\ \alpha_n\end{bmatrix}_A$$

$\begin{bmatrix}2 & 3\end{bmatrix}_B$ means go 2 and 3 steps in directions $\mathbf{b_1}$ and $\mathbf{b_2}$. We need to know $\mathbf{b_1}$ and $\mathbf{b_2}$ in coordinate system of A. Because in B's coordinates they are $\begin{bmatrix}1 & 0\end{bmatrix}_B^T$ and $\begin{bmatrix}0 & 1\end{bmatrix}_B^T$

$$\begin{bmatrix} | & | & & | \\ \mathbf{b_1} & \mathbf{b_2} & \dots & \mathbf{b_n} \\ | & | & & | \end{bmatrix}^{-1} \begin{bmatrix}\alpha_1 \\ \vdots \\ \alpha_n\end{bmatrix}_A = \begin{bmatrix}\beta_1 \\ \vdots \\ \beta_n\end{bmatrix}_B$$

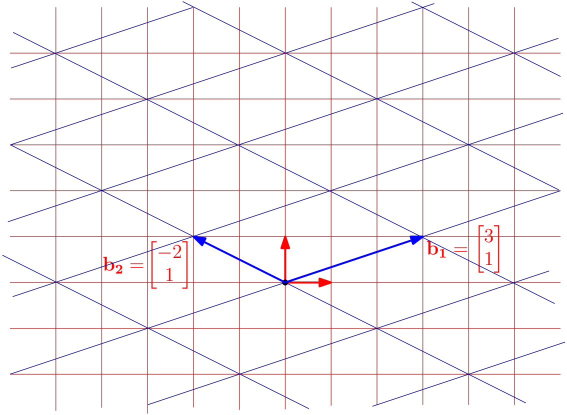
Change of Bases

$$\mathbf{b}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \mathbf{b_2} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$



Change of Bases

$$\mathbf{b}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \mathbf{b_2} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$



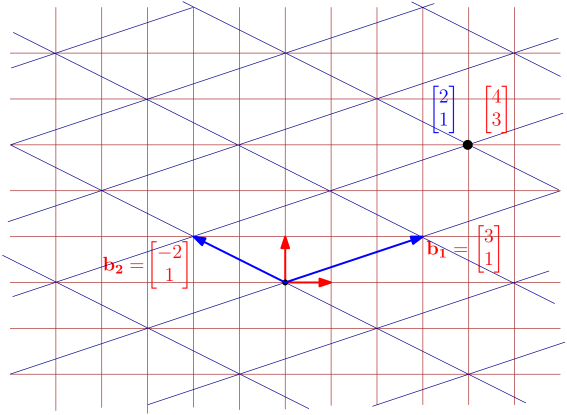
Change of Bases

$$\mathbf{b_1} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \mathbf{b_2} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 3 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 3 & -2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} .2 & .4 \\ -.2 & .6 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$



Transformation in different Bases

- Apply transformation T to vector $\mathbf{x}_B$
- T is given in coordinate system of A, $\underline{\text{we cannot do } T\mathbf{x}_B}$
- Previously we translated vector from one coordinates system to other
- Now we need to do it for transformation

$$\underbrace{B^{-1}\,\underbrace{T\,\underbrace{B\,\mathbf{x}_B}_{\mathbf{x}_A}}_{\mathbf{x}'_A}}_{\mathbf{x}'_B}$$

- Let T_B be the transformation in B coordinate system then

$$T_B = B^{-1}TB$$

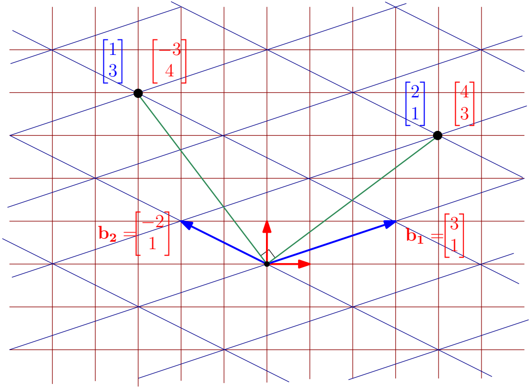
- By the same reasoning

$$T = BT_BB^{-1}$$

Transformation in different Bases

$$\underbrace{B^{-1}\, \underbrace{T\, \underbrace{B\, \underbrace{\mathbf{x}_B}_{}}_{\mathbf{x}_A}}_{\mathbf{x}'_A}}_{\mathbf{x}'_B}$$

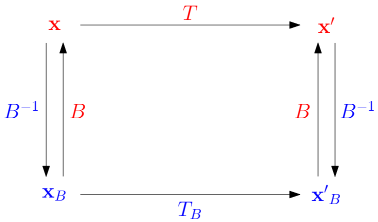
$$\begin{bmatrix} .2 & .4 \\ -.2 & .6 \end{bmatrix} \underbrace{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{90^\circ \text{ rotation}} \begin{bmatrix} 3 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$



Transformation in different Bases

- Translation of vectors and linear transformation between standard bases and another basis B
- Vectors in B are bases vectors (linearly independent) i.e. B is invertible

$$B = \begin{bmatrix} | & | & & | \\ \mathbf{b_1} & \mathbf{b_2} & \dots & \mathbf{b_n} \\ | & | & & | \end{bmatrix}$$

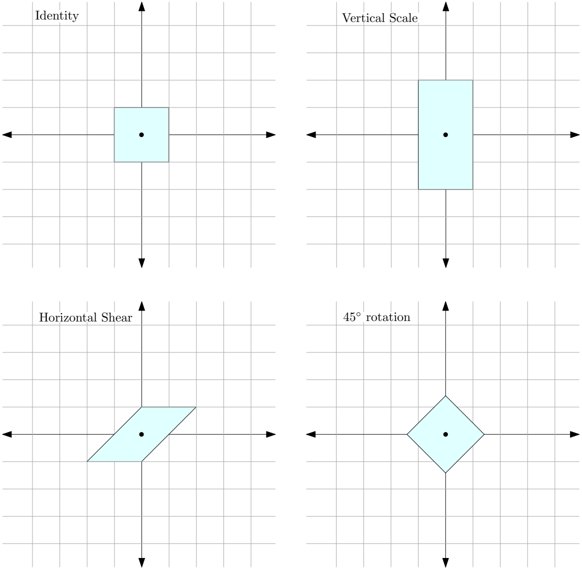


$$T_B = B^{-1} T B \qquad T = B T_B B^{-1}$$

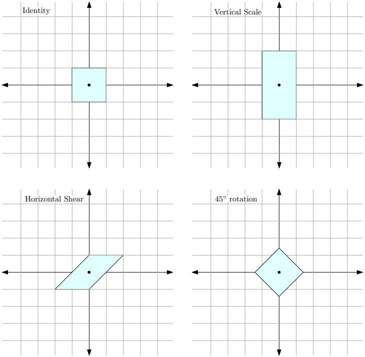
Eigenvalue and Eigenvectors

- Eigenvalue/eigenvectors are extremely important concepts related to linear transformation
- Has fundamental applications in large graph analysis
  - Google's pagerank algorithm and Ask's HITS algorithm
  - Spectral clustering
  - Matrix decomposition
  - Recommender systems
  - Diffusion Processes and Immunization
  - Dynamic systems and many more

Eigenvalue and Eigenvectors: Definition

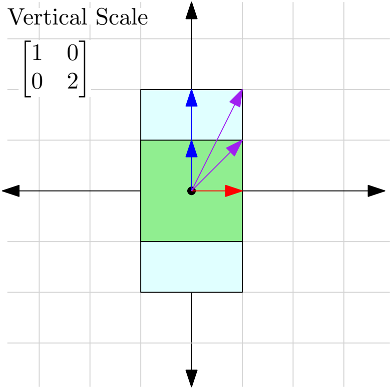


Eigenvalue and Eigenvectors: Definition



- Recall matrices as linear transformation and our view of how the whole space is transformed
- We visualize transformation of the space by observing transformation of the "unit square" (2×2 square centered at the origin)
- Notice some vectors do not change their directions with transformation

Eigenvalue and Eigenvectors: Definition



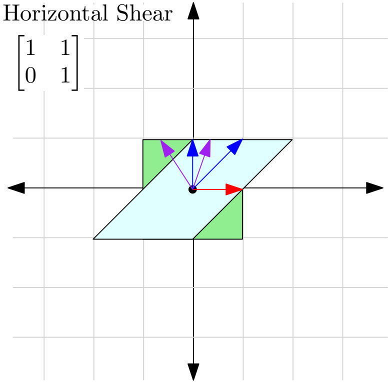
- $\begin{bmatrix}1\\0\end{bmatrix}$ does not change direction or size
- $\begin{bmatrix}0\\1\end{bmatrix}$ does not change direction, size is doubled
- $\begin{bmatrix}1\\1\end{bmatrix}$ changes direction and size

The horizontal and vertical vectors are special, they are called eigenvectors

Horizontal vector size does not change so the corresponding eigenvalue is 1

Vertical vector's size is doubled so the corresponding eigenvalue is 2

Eigenvalue and Eigenvectors: Definition

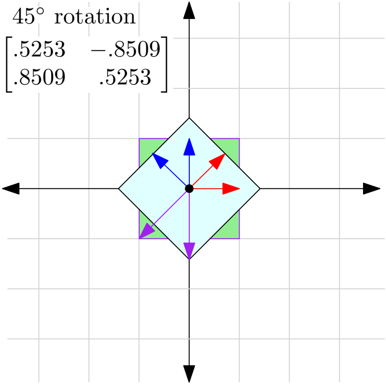


- $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ does not change direction or size
- $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ changes direction and size
- $\begin{bmatrix} -.6 \\ 1 \end{bmatrix}$ changes direction and size

The horizontal vector is special called eigenvector

Horizontal vector size does not change so the corresponding eigenvalue is 1

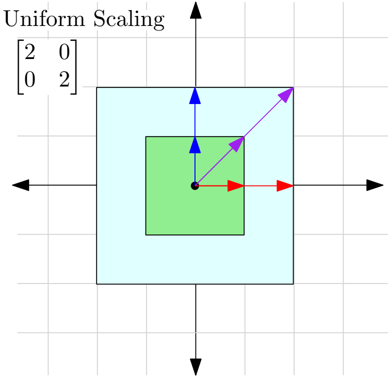
Eigenvalue and Eigenvectors: Definition



- $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ rotates by 45°
- $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ rotates by 45°
- $\begin{bmatrix} -.6 \\ 1 \end{bmatrix}$ rotates by 45°

All vectors change their span

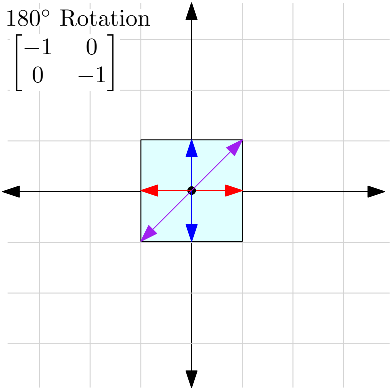
Eigenvalue and Eigenvectors: Definition



- $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ does not change span and size is doubled
- $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ does not change span and size is doubled
- $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ does not change span and size is doubled

All vectors stay on their spans and sizes are doubled

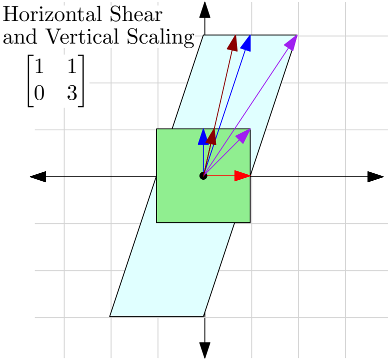
Eigenvalue and Eigenvectors: Definition



- $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ does not change span and size is scaled by -1
- $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ does not change span and size is scaled by -1
- $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ does not change span and size is scaled by -1

All vectors stay on their spans and sizes are doubled

Eigenvalue and Eigenvectors: Definition



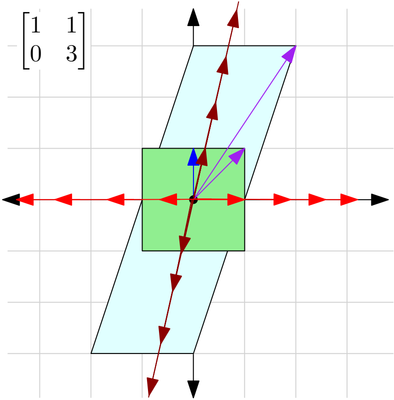
All other vectors change their span

- $\begin{bmatrix}1\\0\end{bmatrix}$ does not change span and size
- $\begin{bmatrix}0\\1\end{bmatrix}$ changes its span and size
- $\begin{bmatrix}.4472\\.8944\end{bmatrix}$ does not change span and size is increased

Eigenvalue and Eigenvectors: Computation

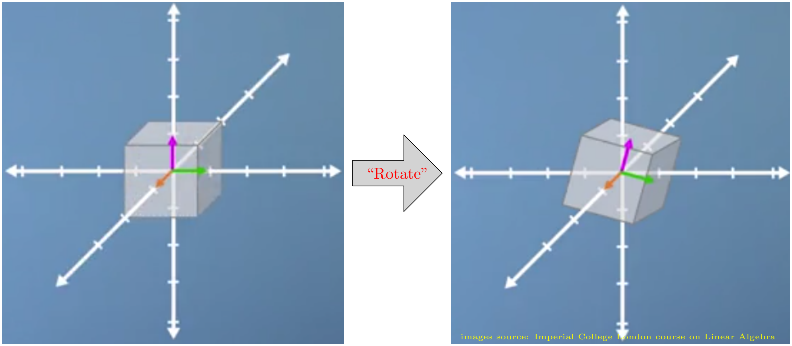
- eigen (German) means "self"' ' or "characteristic"
- eigenvectors := "self vectors" or "characteristic vectors"
- Transform the space
- Find vectors that remain on the same span (these are eigenvectors)
- Measure how their lengths have changed (corresponding eigenvalues)
- Clearly, cannot do it geometrically, think of higher dimensions
- For a square matrix A, solve $A\mathbf{x} = \lambda\mathbf{x}$ for $\mathbf{x}$
  - $\mathbf{x}$ is a vector that stays on its span, just scales by a factor of λ
  - There is no change of direction (span) of $\mathbf{x}$
  - Solutions $\mathbf{x}$'s are called eigenvectors of A
  - λ is called the eigenvalue corresponding to $\mathbf{x}$

Eigenvalue and Eigenvectors: Definition



- By linearity, vectors on a line map to a line, all vectors on the span of an eigenvectors are also eigenvectors

Eigenvalue and Eigenvectors: Definition



- In 2d rotation all vectors change their spans (except 180˚ rotation)
- In 3d x-axis and y-axis change their spans but z-axis does not
- These are eigenvectors of this rotation
- Physically, this is the axis of rotation

Eigenvalue and Eigenvectors: Computation

$$[\mathbf{x}, \lambda] \text{ is an eigen pair} \quad \Leftrightarrow \quad A\mathbf{x} = \lambda\mathbf{x}$$

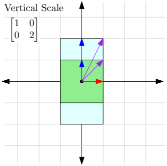
- LHS is matrix-vector product, RHS is scalar-vector product
- Convert RHS to $\lambda\mathbb{I}\mathbf{x}$ ($\lambda\mathbb{I}$ is the uniform scaling matrix)
- This makes the math work but does not change the meaning

$$[\mathbf{x}, \lambda] \text{ is an eigen pair} \quad \Leftrightarrow \quad A\mathbf{x} - \lambda\mathbb{I}\mathbf{x} = \mathbf{0} \quad \Leftrightarrow \quad (A - \lambda\mathbb{I})\mathbf{x} = \mathbf{0}$$

- $\mathbf{x} = \mathbf{0}$ is a trivial solution (no length or direction)
- We want $\mathbf{x}$ that is mapped to $\mathbf{0}$ by the linear transform $(A - \lambda\mathbb{I})$
- A transformation maps a non-zero vector to $\mathbf{0}$ only if it's determinant is 0
- $\therefore$ we find λ such that $det(A - \lambda\mathbb{I}) = 0$
- Once we get the transformation, solve the system of linear equation to $(A - \lambda\mathbb{I})\mathbf{x} = \mathbf{0}$ to find $\mathbf{x}$

Eigenvalue and Eigenvectors: Computation

- $det\left(\begin{bmatrix}1-\lambda & 0\\ 0 & 2-\lambda\end{bmatrix}\right) = (1-\lambda)(2-\lambda)$
- $(1-\lambda)(2-\lambda) = 0 \implies \lambda = 1$ or $\lambda = 2$



$$\underbrace{\begin{bmatrix}1-1 & 0\\ 0 & 2-1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}}_{@\lambda=1:}$$

$$\implies \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\implies \begin{bmatrix}0\\ x_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix} \implies \mathbf{x} = \begin{bmatrix}t\\ 0\end{bmatrix}$$

$[1, \begin{bmatrix}t\\ 0\end{bmatrix}]$ is an eigenpair

$$\underbrace{\begin{bmatrix}1-2 & 0\\ 0 & 2-2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}}_{@\lambda=2:}$$

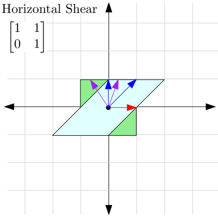
$$\implies \begin{bmatrix}-1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\implies \begin{bmatrix}-x_1\\ 0\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix} \implies \mathbf{x} = \begin{bmatrix}0\\ t\end{bmatrix}$$

$[2, \begin{bmatrix}0\\ t\end{bmatrix}]$ is an eigenpair

Eigenvalue and Eigenvectors: Computation

- $det\left(\begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix}\right) = (1-\lambda)^2$
- $(1-\lambda)^2 = 0 \implies \lambda = 1$



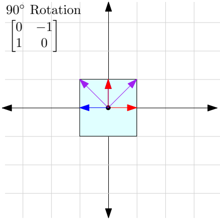
$$\underbrace{\begin{bmatrix} 1-1 & 1 \\ 0 & 1-1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{@\lambda=1:}$$

$$\implies \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \begin{bmatrix} x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \mathbf{x} = \begin{bmatrix} t \\ 0 \end{bmatrix}$$

$[1, \begin{bmatrix} t \\ 0 \end{bmatrix}]$ is an eigenpair

Eigenvalue and Eigenvectors: Computation

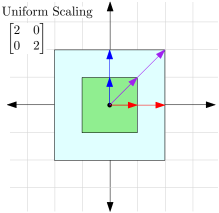
- $det\left(\begin{bmatrix} 0-\lambda & -1 \\ 1 & 0-\lambda \end{bmatrix}\right) = (0-\lambda)^2$ -(1)(-1)
- $(-\lambda)^2 + 1 = 0 \implies \lambda^2 = -1$
- No real λ as solution



Hence no real eigenvectors

Eigenvalue and Eigenvectors: Computation

- $det\left(\begin{bmatrix} 2-\lambda & 0 \\ 0 & 2-\lambda \end{bmatrix}\right) = (2-\lambda)^2$
- $(2-\lambda)^2 = 0 \implies \lambda = 2$



$$\underbrace{\begin{bmatrix} 2-2 & 0 \\ 0 & 2-2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{@\lambda=2:}$$

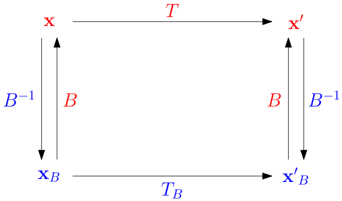
$$\implies \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \mathbf{x} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

All vectors are eigenvectors with eigenvalue 2

Transformation in different Bases

- Translation of vectors and linear transformation between standard bases and another bases B
- Vectors in B are basis vectors (linearly independent) B is invertible

$$B = \begin{bmatrix} | & | & & | \\ \mathbf{b_1} & \mathbf{b_2} & \dots & \mathbf{b_n} \\ | & | & & | \end{bmatrix}$$

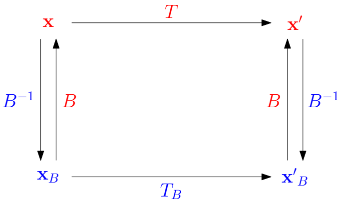


$$T_B = B^{-1} T B \qquad T = B T_B B^{-1}$$

Eigenbases: Diagonalization

- Let T be a $n \times n$ linear transformation
- Let $B = \{\mathbf{b_1}, \ldots, \mathbf{b_n}\}$ be bases - vectors in B are eigenvectors of T
- For $1 \leq i \leq n, \quad T\mathbf{b_i} = \lambda_i \mathbf{b_i}$
- Note there must be n vectors in B

$$B = \begin{bmatrix} | & | & & | \\ \mathbf{b_1} & \mathbf{b_2} & \ldots & \mathbf{b_n} \\ | & | & & | \end{bmatrix}$$



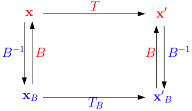
$$T_B = B^{-1} T B \qquad T = B T_B B^{-1}$$

- How does $T\mathbf{x}$ looks like in eigenbasis?

Eigenbases: Diagonalization

- Let T be a $n \times n$ linear transformation
- Let $B = \{\mathbf{b_1}, \ldots, \mathbf{b_n}\}$ be bases - vectors in b are eigenvectors of T
- For $1 \leq i \leq n$, $\quad T\mathbf{b_i} = \lambda_i \mathbf{b_i}$

$T_B = B^{-1}TB$



$T = BT_BB^{-1}$

- How does $T\mathbf{x}$ looks like in eigenbasis?

$$T\mathbf{x} = T(\alpha_1\mathbf{e_1} + \ldots + \alpha_n\mathbf{e_n}) \quad = \quad T(\beta_1\mathbf{b_1} + \ldots + \beta_n\mathbf{b_n})$$

$$= \beta_1 T\mathbf{b_1} + \ldots + \beta_n T\mathbf{b_n} \quad = \quad \beta_1\lambda_1\mathbf{b_1} + \ldots + \beta_n\lambda_n\mathbf{b_n}$$

$$= \begin{bmatrix} | & | & & | \\ \mathbf{b_1} & \mathbf{b_2} & \ldots & \mathbf{b_n} \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} \quad = \quad BD\mathbf{x_B} \quad = \quad BDB^{-1}\mathbf{x}$$

Eigenbases: Diagonalization

- Let T be a $n \times n$ linear transformation
- Let $B = \{\mathbf{b_1}, \ldots, \mathbf{b_n}\}$ be bases - vectors in b are eigenvectors of T
- For $1 \leq i \leq n, \quad T\mathbf{b_i} = \lambda_i \mathbf{b_i}$

$$B = \begin{bmatrix} | & | & & | \\ \mathbf{b_1} & \mathbf{b_2} & \ldots & \mathbf{b_n} \\ | & | & & | \end{bmatrix} \qquad D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$T\mathbf{x} = BDB^{-1}\mathbf{x}$$

- Very easy to take T to a higher power (compose it many times)
- $T = BDB^{-1}$
- $T^2 = BDB^{-1}BDB^{-1} = BDDB^{-1} = BD^2B^{-1}$
- $T^3 = BD^2B^{-1}BDB^{-1} = BD^2DB^{-1} = BD^3B^{-1}$
- $T^4 = BD^3B^{-1}BDB^{-1} = BD^3DB^{-1} = BD^4B^{-1}$
- $T^k = \ldots = BD^kB^{-1}$

Eigenbases: Diagonalization

- Let T be a $n \times n$ linear transformation
- Let $B = \{\mathbf{b_1}, \ldots, \mathbf{b_n}\}$ be bases - vectors in b are eigenvectors of T
- For $1 \leq i \leq n, \quad T\mathbf{b_i} = \lambda_i \mathbf{b_i}$

$$B = \begin{bmatrix} | & | & & | \\ \mathbf{b_1} & \mathbf{b_2} & \ldots & \mathbf{b_n} \\ | & | & & | \end{bmatrix} \qquad D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$T\mathbf{x} \ = \ BDB^{-1}\mathbf{x}$$

- $T^k \ = \ BD^kB^{-1}$

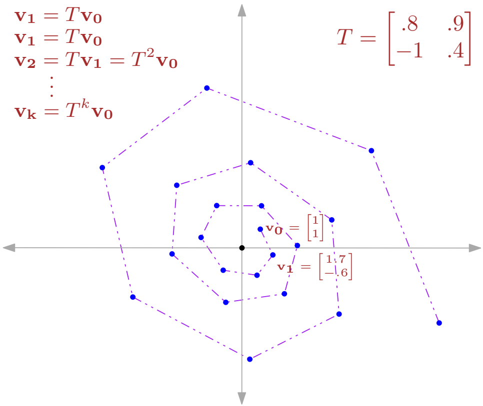
$$D^k = \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix}$$

Powers of Matrices:

Suppose T represents the change in location of a particle per second

$$\mathbf{v_1} = T\mathbf{v_0}$$
$$\mathbf{v_1} = T\mathbf{v_0}$$
$$\mathbf{v_2} = T\mathbf{v_1} = T^2\mathbf{v_0}$$
$$\vdots$$
$$\mathbf{v_k} = T^k\mathbf{v_0}$$

$$T = \begin{bmatrix} .8 & .9 \\ -1 & .4 \end{bmatrix}$$



Find location of the particle after two weeks

Powers of Matrices:

Fibonacci numbers F_n, $0, 1, 1, 2, 3, 5, 8, 13, 21, \ldots$

$$F_n = \begin{cases} 0 & \text{if } n = 0 \\ 1 & \text{if } n = 1 \\ F_{n-2} + F_{n-1} & \text{if } n \geq 2 \end{cases}$$

Let $T = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$

$\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$ $\quad$ $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 3 \\ 2 \end{smallmatrix}\right]$ $\quad$ $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} 3 \\ 2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 5 \\ 3 \end{smallmatrix}\right]$ $\quad$ $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} 5 \\ 3 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 8 \\ 5 \end{smallmatrix}\right]$ $\quad$ $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} 8 \\ 5 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 13 \\ 8 \end{smallmatrix}\right]$

$$\begin{bmatrix} F_{k+2} \\ F_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} \qquad \begin{bmatrix} F_{k+2} \\ F_{k+1} \end{bmatrix} = T^k \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$F_k = \frac{\lambda_1^k - \lambda_2^k}{\lambda_1 - \lambda_2} = \frac{(1+\sqrt{5})^k - (1-\sqrt{5})^k}{2^k\sqrt{5}}$$

Powers of Matrices:

First order linear recurrence relation

$$x_{t+1} = ax_t$$
$$x_0 = 3$$

Coupled system of recurrence relations

$$x_{t+1} = 3x_t + 5y_t$$
$$y_{t+1} = 4x_t - 2y_t$$
$$x_0 = 2, \ y_0 = 3$$

Model many practical scenarios in population dynamics, economics, epidemiology, computing, signal processing

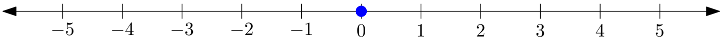
Let $\mathbf{u_t} = \begin{bmatrix} x_t \\ y_t \end{bmatrix}$

$\mathbf{u_0} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

$T = \begin{bmatrix} 3 & 5 \\ 4 & -2 \end{bmatrix}$

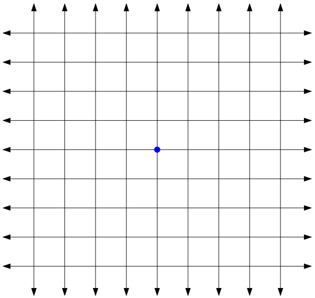
- $\mathbf{u_1} = T\mathbf{u_0}$
- $\mathbf{u_2} = T\mathbf{u_1} = TT\mathbf{u_0} = T^2\mathbf{u_0}$
- $\mathbf{u_3} = T\mathbf{u_2} = TT^2\mathbf{v_0} = T^3\mathbf{u_0}$
- $\vdots$
- $\mathbf{u_k} = T^k\mathbf{u_0}$

Random Walk

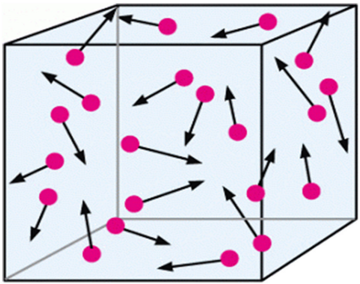


- Suppose the blue dot starts at 0
- At every step if it is at number i, then with probability $1/2$ it goes $i+1$ and and with probability $1/2$ it goes to $i-1$
- How many steps would it take to reach 6 or -8?
- What is root mean squared distance the • covers in n steps?
- Many possible extensions
- Lazy walks: with prob. $1/2$ stay at i, move to $i \pm 1$ each prob $1/4$
- Biased walks: with prob. $3/4$ move to $i+1$ and $1/4$ move to $i-1$
- Biased walks: with prob. $1/2$ move to $i+b$ and $1/2$ move to $i-1$
- Models many things: stock prices fluctuations, gambling outcomes, team results in a game's season, molecules movements

Random Walk Generalizations



At every step • goes $\{Up, Down, Left, Right\}$ with probability $1/4$



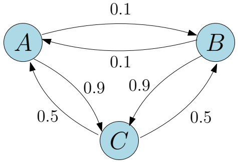
- Random walk on grid
- Random walk in space, often called Brownian motion
  - Model movements of particles in liquid or gas. The particle undertake random walk caused by momentum imparted to it by molecules in random directions

Random Walk on Graphs

- Let $G = (V, E)$ be a graph or digraph
- Let $d(u)$ be the degree of $u \in v$
- A random walker starts at some vertex $v_0 \in V$
- At every step if the walker is at vertex u, it picks randomly moves to a random (out) neighbor of u
- The probability that current vertex is u and next vertex is $v \in N(u)$ is $1/d(u)$ or $1/d^+(u)$ (for digraphs)

Markov Chain

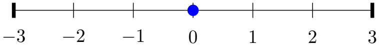
- A Markov chain is a stochastic process defined on finite number of states
- The changes of state of system are called transition
- Transitions probabilities b/w states are given in transition matrix T
- Let X_n be the state of the system at time n
- $T(i,j) := Pr[X_{n+1} = i | X_n = j]$: prob. that system goes from state j to i
- $0 \leq T(i,j) \leq 1$ and columns sum to 1 ▷ column-stochastic
- Memoryless process: $T(i,j)$ does not depend on the history of transitions ▷ Markovian property
- Given present state, the past and future states are independent



$$\begin{bmatrix} 0.0 & 0.1 & 0.5 \\ 0.1 & 0.0 & 0.5 \\ 0.9 & 0.9 & 0.0 \end{bmatrix}$$

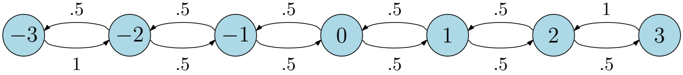
Markov Chain

- Bounded Random Walk on integers $\{-3, \ldots, 3\}$



- The $\color{blue}\bullet$ begins at 0
- If $\color{blue}\bullet$ is at ± 3, then with prob. 1 it goes to ± 2
- If $\color{blue}\bullet$ is at $i \neq \pm 3$, then with prob. .5 it goes to $i \pm 1$

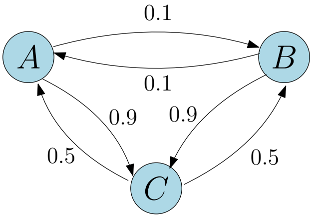
	−3	−2	−1	0	1	2	3
−3	0	.5	0	0	0	0	0
−2	1	0	.5	0	0	0	0
−1	0	.5	0	.5	0	0	0
0	0	0	.5	0	.5	0	0
1	0	0	0	.5	0	.5	0
2	0	0	0	0	.5	0	1
3	0	0	0	0	0	.5	0



Language Recognition System

- Smartphones next words suggestions use language generation
- The first i words are typed, what will be the $(i+1)$st word?
- Model language generation as a Markov chain ▷ not realistic
- States correspond to last used words (say vocabulary has 1000 words)
- Transition probabilities $p_{w_i w_j} := Pr[w_j|w_i] := \dfrac{freq(w_i w_j)}{freq(w_i)}$
- Estimate the 1000×1000 probabilities from a large text corpus
- Probability of generating a text $w_1 w_2 w_3 w_4 w_5$ is $p_{w_1} p_{w_1 w_2} p_{w_2 w_3} p_{w_3 w_4} p_{w_4 w_5}$
- p_{w_i} is (empirical prob) frequency of w_i as first word in the corpus
- Can extend it by estimating $p_{w_i w_j w_k} := Pr[w_k|w_i w_j]$

Markov Chain

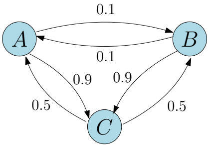


$$\begin{bmatrix} 0.0 & 0.1 & 0.5 \\ 0.1 & 0.0 & 0.5 \\ 0.9 & 0.9 & 0.0 \end{bmatrix}$$

- Instead of thinking that the system is in a given state at time t, consider
- a vector $\mathbf{x}$ specifying probability distribution of system being in all states
- $\mathbf{x}^{(\mathbf{t})}$ is probability distribution at time t, ${\mathbf{x}^{\mathbf{t}}}_i \geq 0$, $\sum_i {\mathbf{x}^{\mathbf{t}}}_i = 1$
- $\mathbf{x}^{(\mathbf{t+1})} = T\mathbf{x}^{(\mathbf{t})}$
- By Markovian property, probability of going from j to i in two steps is $\sum_k T(k,j)T(i,k) = T^2(i,j)$
- probability of going from j to i in s steps is $T^s(i,j)$

Markov Chain

- $\mathbf{x^{(t)}}$: prob. distribution at time t
- $\mathbf{x^{(t+1)}} = T\mathbf{x^{(t)}}$



$$\begin{bmatrix} 0.0 & 0.1 & 0.5 \\ 0.1 & 0.0 & 0.5 \\ 0.9 & 0.9 & 0.0 \end{bmatrix}$$

A distribution π is a stationary distribution for Markov chain T, if

$$T\pi = \pi$$ ▷ eigenvector of T with eigenvalue 1

- The largest eigenvalue of a column stochastic real matrix is real ($\lambda_1 = 1$)

A markov chain is *ergodic* if there is a unique stationary distribution π and for any initial distribution $\mathbf{x}$ we have

$$\lim_{t\to\infty} M^t\mathbf{x} = \pi$$ ▷ always converges to π