Lecture 11: Projections

Tuesday, September 29, 2015 9:30 AM

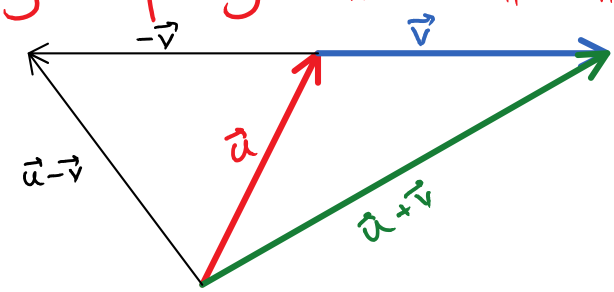
Admin: Projections Ch. 5.13

Recall:

Inner product: $\vec{u}\cdot\vec{v} = \sum_i u_i^* v_i$ — complex conjugate

Length: $\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}}$

Triangle inequality: $\|\vec{u}+\vec{v}\| \le \|\vec{u}\| + \|\vec{v}\|$

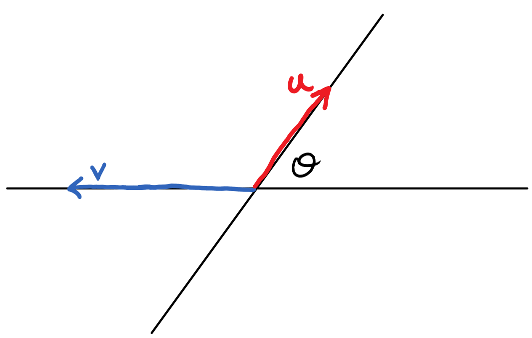


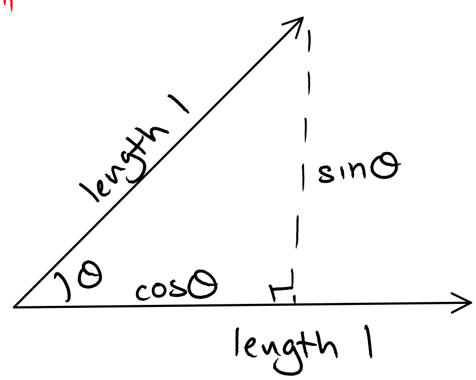
Corollary: $\|u\| \le \|v\| + \|u-v\|$

$\Rightarrow \|u-v\| \ge \|u\| - \|v\|$

Angles: angle between lines Span(u) and Span(v)

$$\cos\theta = \frac{|\vec{u}\cdot\vec{v}|}{\|u\|\cdot\|v\|}$$





Cauchy-Schwarz inequality: $|\vec{u}\cdot\vec{v}| \le \|\vec{u}\|\cdot\|\vec{v}\|$

(with equality iff collinear)

Example: $\sum_{i=1}^{n} u_i = (1,1,\dots,1)\cdot\vec{u}$

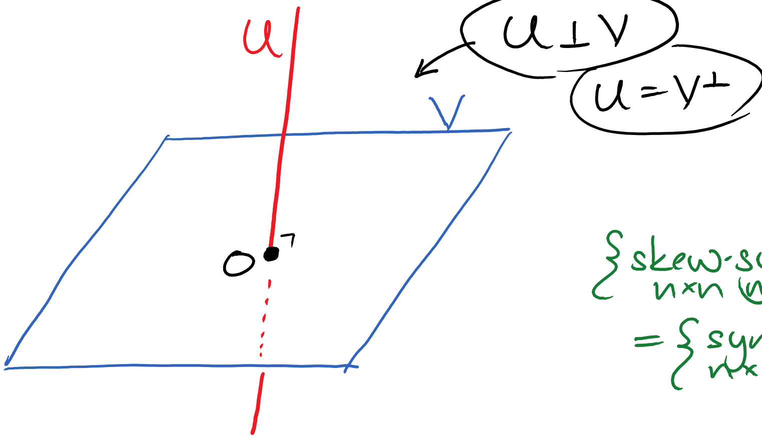
$\le \|(1,\dots,1)\|\cdot\|\vec{u}\|$

$= \sqrt{n}\cdot\|\vec{u}\|$

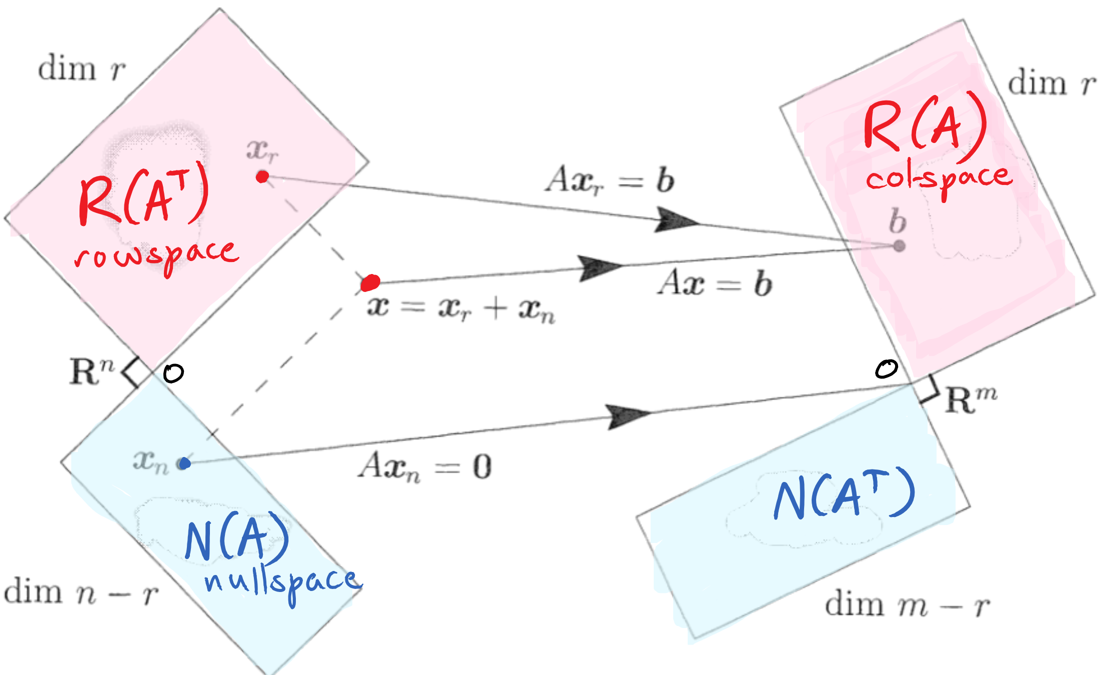
Orthogonal vectors: $\vec{u}\perp\vec{v} \iff \vec{u}\cdot\vec{v} = 0$

Orthogonal subspaces & orthogonal complements:

Orthogonal subspaces & orthogonal complements:



$\left\{ \begin{matrix} \text{skew-symmetric} \\ n\times n \text{ matrices} \end{matrix} \right\} = \left\{ \begin{matrix} \text{symmetric} \\ n\times n \text{ matrices} \end{matrix} \right\}^{\perp}$



$N(A) = R(A^T)^{\perp}$ $\qquad$ $N(A^T) = R(A)^{\perp}$

Consequences:

* $\{x \mid Ax = b\}$ is an $(n-r)$-dim affine space.
* Any subspace $\mathcal{U}$ can be described either
  - by a basis for $\mathcal{U}$ (dim $\mathcal{U}$ vectors), or
  - by a basis for $\mathcal{U}^{\perp}$ (n-dim $\mathcal{U}$ vectors)

each one a constraint on $\mathcal{U}$

This is very important for

Error-correcting codes: Vector spaces over finite fields

Let $V = \{0,1\}^n$ over $\mathbb{F}_2$ (addition/multiplication mod 2)

e.g., $V = \{(0,0),(0,1),(1,0),(1,1)\}$ for $n=2$

- Lengths don't mean much
  $\|(1,1)\| \overset{?}{=} \sqrt{1^2+1^2} = \sqrt{2}$
  but $\frac{1}{\sqrt{2}}(1,1) \notin V$; you can't renormalize
- Angles don't mean anything
  there's no Euclidean geometry here
- But orthogonality still makes sense!
  $\vec{x} \perp \vec{y}$ if $\sum_{i=1}^{n} x_i y_i = 0 \mod 2$
  Note: $(1,1) \perp (1,1)$!!
  any even-weight vector is orthogonal to itself

$\Rightarrow$ The Rank-Nullity Theorem still holds!

Definition: A binary linear error-correcting code is a subspace of $\{0,1\}^n$.

Example: 3-bit repetition code

$$0 \mapsto 000$$
$$1 \mapsto 111$$

code $= \{(0,0,0),(1,1,1)\}$ 1-dimensional subspace

$= R(A^T)$ for $A = (1 \;\; 1 \;\; 1)$

generator matrix

$$N(A) = R(A^T)^\perp$$
$$= R\left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}\right) = \left\{ \begin{matrix} (0,0,0), (0,1,1), \\ (1,1,0), (1,0,1) \end{matrix} \right\}$$

parity check matrix

2 dimensional

$\leadsto$ a linear code can be specified by either the generating matrix or the parity-check matrix

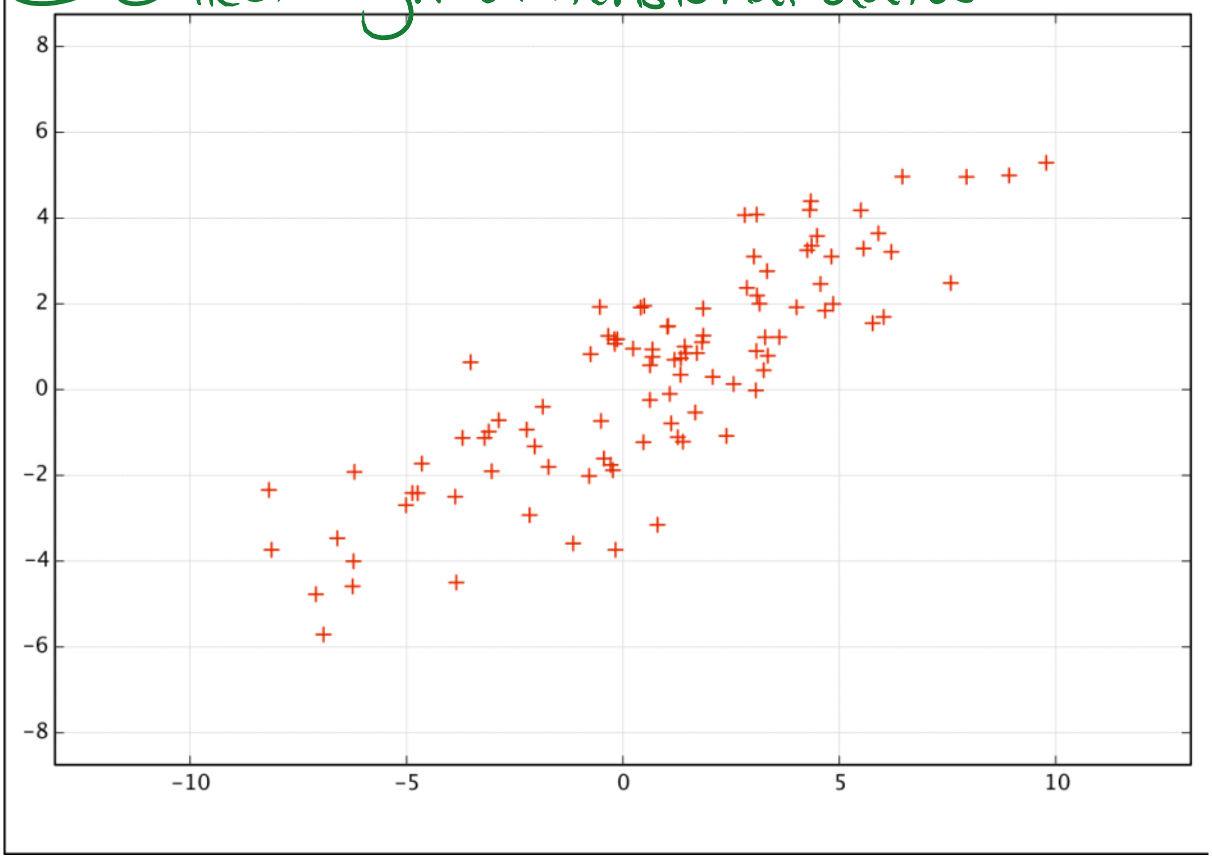
Today: Projections and orthogonal bases

PROJECTIONS

① Motivation: Principal component analysis (PCA)

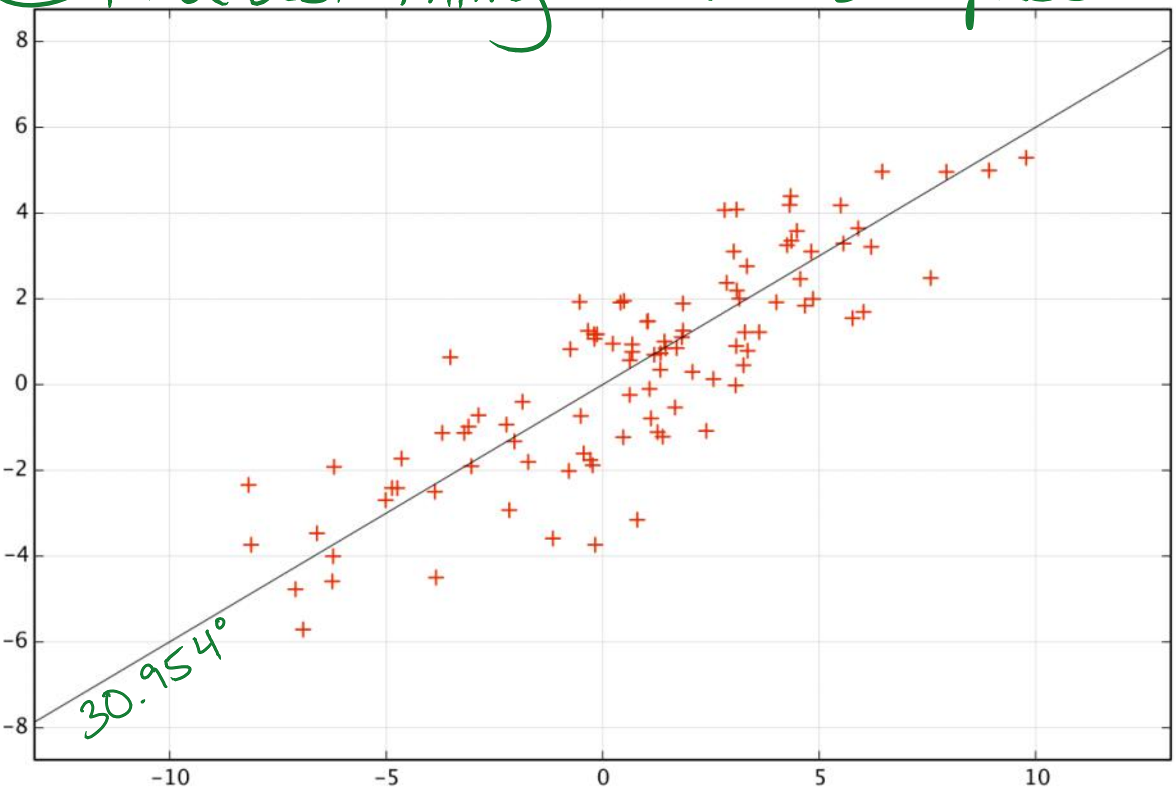
Example:

① Collect high-dimensional data



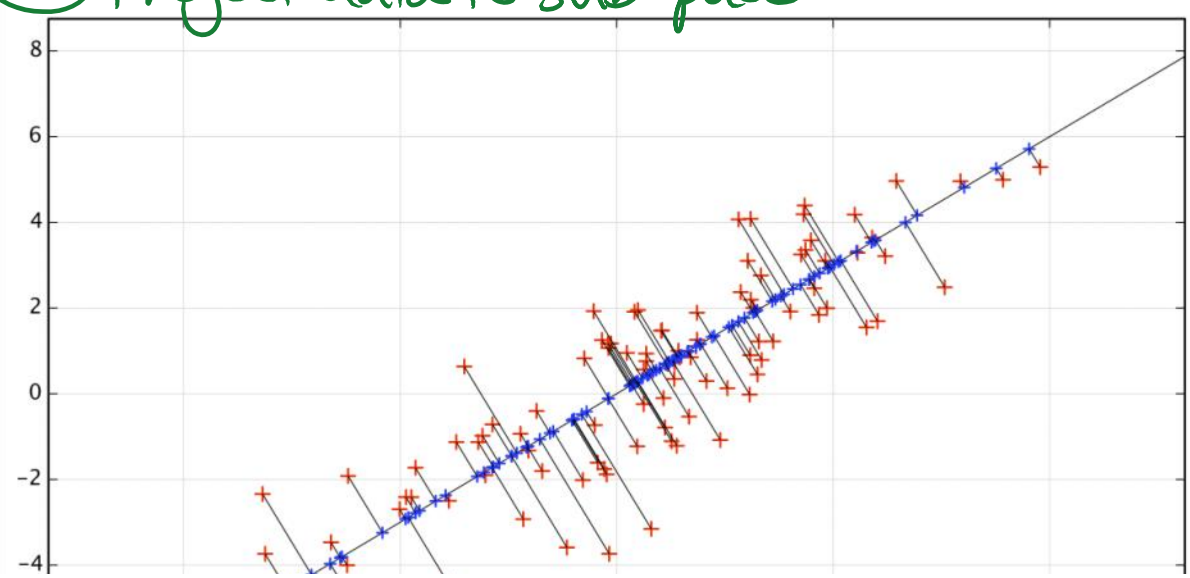
```
n = 100; largerstddev = 5; rotationangle = 30;
c = cosd(rotationangle); s = sind(rotationangle);
data = [c,-s; s,c] * [largerstddev,0; 0,1] * randn(2, n);
```

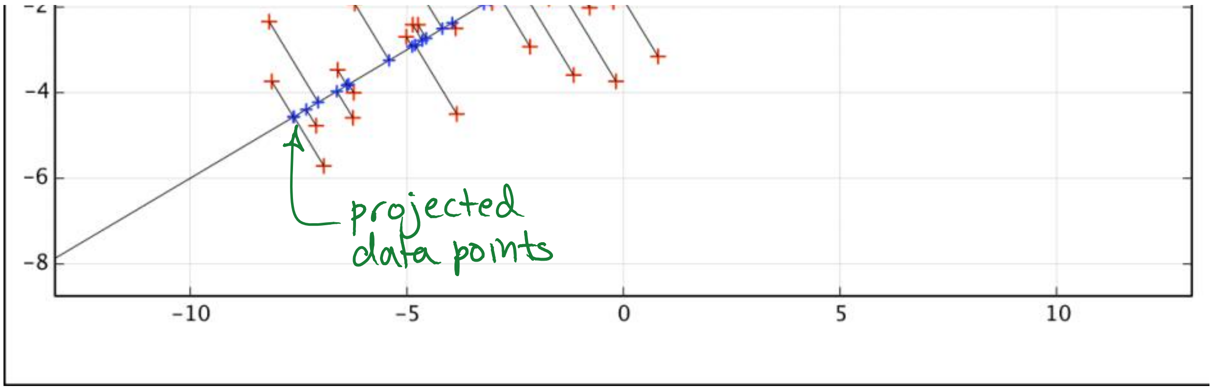
② Find best-fitting low-dim'l subspace




```
# run PCA to extract best-fitting line
[coeff,score] = princomp(data');
slope = coeff(2,1) / coeff(1,1);
fitangle = atand(slope)
```

③ Project data to subspace






```
projecteddata = coeff(:,1) * (coeff(:,1)' * data);
```

Applications: Machine learning, statistics, data analysis, compression, ...

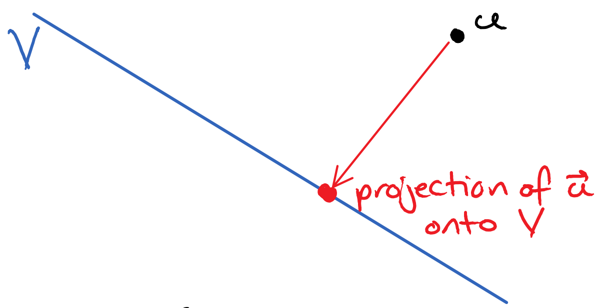
We'll cover PCA when we get to the singular-value decomp.

Problem: How can we make all these PROJECTIONS?

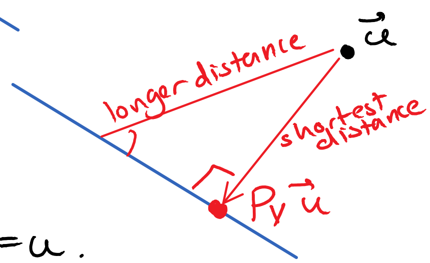
Definition: Let V be a subspace of $\mathbb{R}^n$ (or $\mathbb{C}^n$).

The orthogonal projection onto V maps

any point $u \in \mathbb{R}^n$ $\longmapsto$ the closest point in V



Observe: 1. $(P_V\vec{u} - \vec{u}) \perp V$.



2. If $u \in V$, $P_V u = u$.
3. $P_V^2 = P_V$

3. $P_V^2 = P_V$

$\left(\text{since } P_V \vec{u} \in V, \quad P_V(P_V \vec{u}) = P_V \vec{u} \right)$

4. Projections are <u>linear transformations</u>

$\left(P_V(\alpha \vec{u}) = \alpha P_V \vec{u}, \; P_V(\vec{u}+\vec{w}) = P_V \vec{u} + P_V \vec{w} \right)$

5. $I - P_V = P_{V^\perp}$ projection onto $V^\perp$

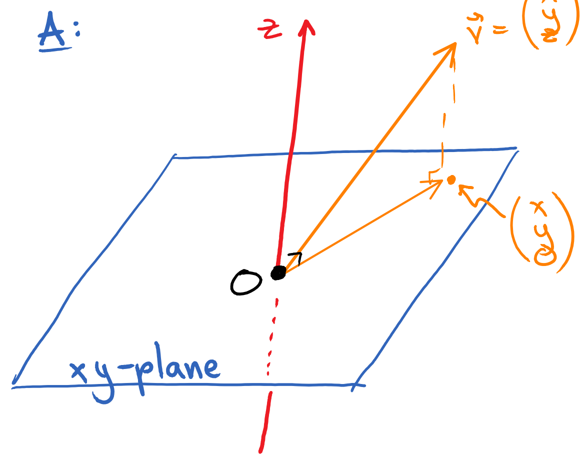
where $V^\perp = \{\text{all vectors } \perp \text{ to } V\}$, the "orthogonal complement" of V. Recall $\dim(V^\perp) = n - \dim(V)$.

We'll prove these properties later!

<u>Intuition</u>: Projection to a coordinate subspace

Q: What is the projection from $\mathbb{R}^3$ to the xy-plane?

A:



The matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

discards the z-coordinate of any point.

<u>More generally</u>

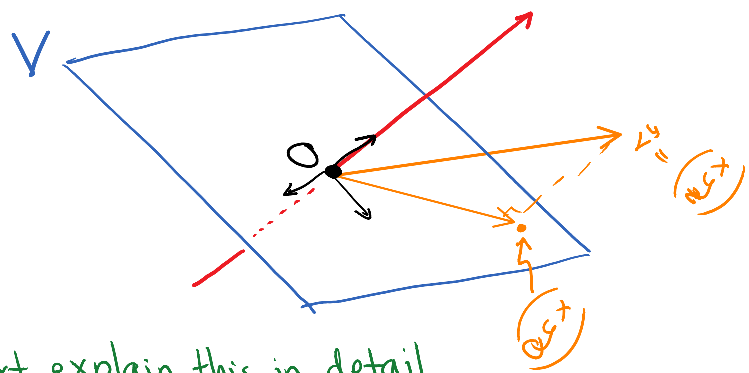
$$\begin{array}{l} k \\ n-k \end{array} \left(\begin{array}{c|c} \begin{matrix} 1 & & \\ & 1 & \\ & & \ddots \; 1 \end{matrix} & 0 \\ \hline 0 & \begin{matrix} 0 & & \\ & 0 & \\ & & \ddots \; 0 \end{matrix} \end{array} \right)$$

projects onto the subspace $\text{Span}\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k\}$ in $\mathbb{R}^n$.

<u>More generally</u>:

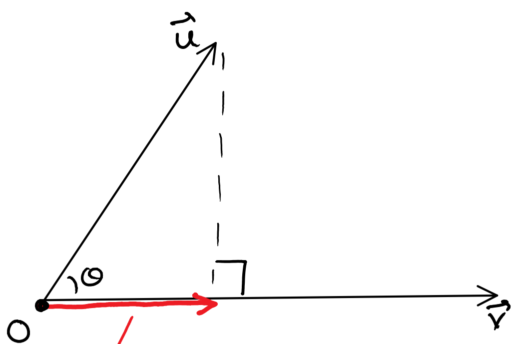
Projecting onto an arbitrary subspace V $\longleftrightarrow$

- Change basis to orthogonal coordinate system starting with basis for V
- Discard the coordinates $\perp$ to V



We'll next explain this in detail...

Example: Projection onto a line



projection of $\vec{u}$ onto the line Span$(\vec{v})$ $= (\vec{v} \cdot \vec{u})\,\vec{v}$ if $\|\vec{v}\| = 1$

Lemma: If $\|\vec{v}\| = 1$, then the projection onto the line Span$(\vec{v})$ is a linear trans. given by $\boxed{\vec{v}\,\vec{v}^T}$

an $n \times n$ matrix if $\vec{v}$ is $n \times 1$

– since $(\vec{v}\,\vec{v}^T)\vec{u} = \vec{v}(\vec{v}^T\vec{u}) = \vec{v}(\underbrace{\vec{v} \cdot \vec{u}}_{\|\vec{u}\|\cos\theta})$ ✓

Example: $\mathbb{R}^3$:

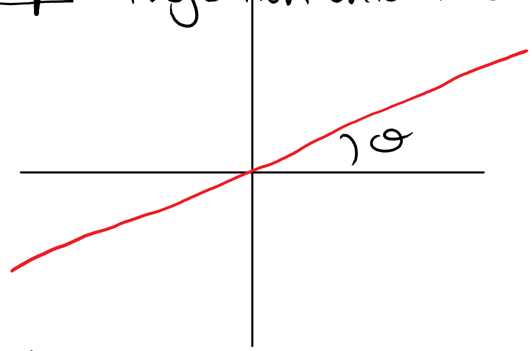
$$\vec{e}_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\vec{e}_y\vec{e}_y^T = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\vec{e}_y\vec{e}_y^T \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ \beta \\ 0 \end{pmatrix}$$

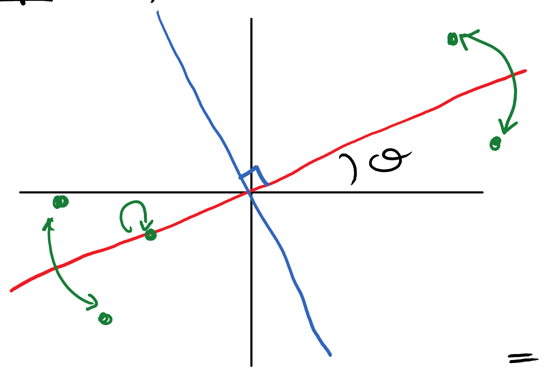
Example: Projection onto the line at angle θ

Example: Projection onto the line at angle θ



$$\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}(\cos\theta, \sin\theta) = \underbrace{\begin{pmatrix}\cos^2\theta & \cos\theta\sin\theta\\ \sin\theta\cos\theta & \sin^2\theta\end{pmatrix}}_{P_\theta}$$

Example: Reflection about the line at angle θ



$$P_\theta - (I - P_\theta)$$

P_θ: leaves vector in red space unchanged

$(I-P_\theta)$: puts a minus sign on vectors in the orthogonal, blue line

$$= 2P_\theta - I = \begin{pmatrix}2c^2-1 & 2cs\\ 2cs & 2s^2-1\end{pmatrix}$$

$$= \begin{pmatrix}\cos 2\theta & \sin 2\theta\\ \sin 2\theta & -\cos 2\theta\end{pmatrix}$$

Example: Projection onto $(1,1,\dots,1)$

$$P = \frac{1}{\sqrt{n}}\begin{pmatrix}1\\ 1\\ \vdots\\ 1\end{pmatrix}\cdot\frac{1}{\sqrt{n}}(1 \ 1 \ \cdots \ 1) = \frac{1}{n}\begin{pmatrix}1 & 1 & \cdots & 1\\ 1 & 1 & \cdots & 1\\ \vdots & \vdots & \ddots & \vdots\\ 1 & 1 & \cdots & 1\end{pmatrix}$$

Observe: $Pv = (\text{avg. of } v\text{'s coords})\cdot\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$.

Example: Projection to $(1,1,\dots,1)^\perp = R\left(\begin{pmatrix}-2 & 1 & & \\ 1 & -2 & 1 & \\ & 1 & -2 & 1\\ & & 1 & \ddots\end{pmatrix}\right)$.

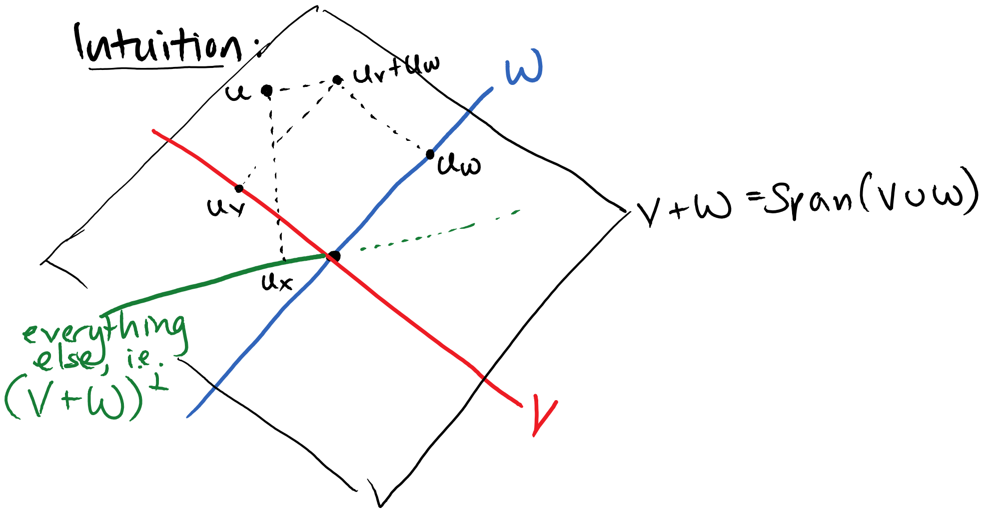
$= I - P$

Problem: How can we construct the projector onto higher-dim. subspaces?

$$\boxed{\text{Key property: If } V \perp W,\quad P_V + P_W = P_{V+W}.}$$

Intuition:

Intuition:



any vector $\vec{u}$ can be expanded as

$$\vec{u} = \vec{u}_V + \vec{u}_W + \vec{u}_X$$

where $\vec{u}_V \in V$, $\vec{u}_W \in W$, $\vec{u}_X \in X$

$$\overset{\shortparallel}{P_V \vec{u}} \qquad \overset{\shortparallel}{P_W \vec{u}} \qquad \overset{\shortparallel}{P_X \vec{u}}$$

$$P_{V+W}\vec{u} = \vec{u}_V + \vec{u}_W$$

Examples & counterexamples:

$$P_{e_1} = e_1 e_1^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$P_{e_2} = e_2 e_2^T = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$P_{e_1} + P_{e_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = P_{x\text{-}y \text{ plane}}$$

This does not work if V is not $\perp$ to W:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix} = ? \text{ not a projection } (?^2 \neq ?)$$

It only works because $e_1 \perp e_2$!

$(c, s) = (\cos\theta, \sin\theta)$

$$\begin{pmatrix} c \\ s \end{pmatrix}\begin{pmatrix} c & s \end{pmatrix} + \begin{pmatrix} s \\ -c \end{pmatrix}\begin{pmatrix} s & -c \end{pmatrix}$$

$$= \begin{pmatrix} c^2 & cs \\ cs & s^2 \end{pmatrix} + \begin{pmatrix} s^2 & -cs \\ -cs & c^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \checkmark$$

$\Rightarrow$ To project onto V, it is enough to have a basis of pairwise orthogonal vectors for V...

Pairwise Orthogonal Sets of Vectors

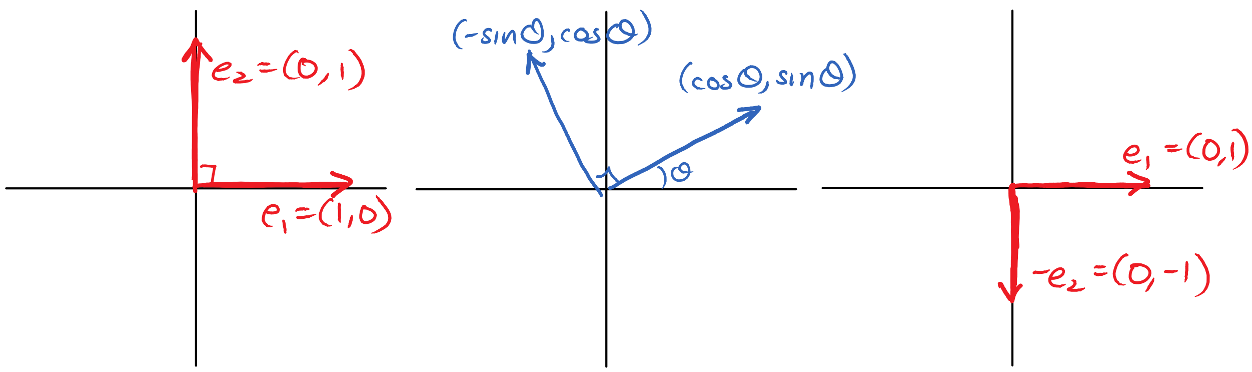
Examples:

① Standard basis in $\mathbb{R}^n$

$$\vec{e}_1 = (1, 0, \dots, 0), \ \dots, \ \vec{e}_n = (0, \dots, 0, 1)$$

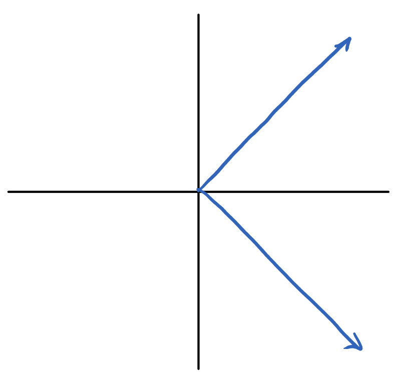
$$\vec{e}_i \cdot \vec{e}_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

② Rotations/reflections of the standard basis



③ Hadamard basis

$$\tfrac{1}{\sqrt{2}}(1,1) \quad , \quad \tfrac{1}{\sqrt{2}}(1,-1)$$



= Fourier basis mod 2

(foundation for Fourier analysis of boolean functions)

change-of-basis matrix $H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

Hadamard basis in dimension 4:

$$H_4 = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

* all entries either $\pm\frac{1}{\sqrt{n}}$
* columns are orthogonal pairwise

$$H_4 = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

In dimensions $n = 2^k$:

$$H_{2^k} = \frac{1}{\sqrt{2}}\begin{pmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{pmatrix}$$

Applications: Coding, control theory, statistics, quantum computing, Fourier analysis of boolean functions

Conjecture (1936): If n is a multiple of 4, then there is a Hadamard basis for $\mathbb{R}^n$.

Open: Find a Hadamard basis for $n = 668$.

④ Haar wavelet basis

$(1, 1, 1, 1)$
$(1, 1, -1, -1)$
$(1, 1, 0, 0)$
$(0, 0, 1, -1)$

for $\mathbb{R}^4$

$(1, 1, 1, 1, 1, 1, 1, 1)$
$(1, 1, 1, 1, -1, -1, -1, -1)$
$(1, 1, -1, -1, 0, 0, 0, 0)$
$(1, -1, 0, 0, 0, 0, 0, 0)$
$(0, 0, 1, -1, 0, 0, 0, 0)$
$(0, 0, 0, 0, 1, 1, -1, -1)$
$(0, 0, 0, 0, 1, -1, 0, 0)$
$(0, 0, 0, 0, 0, 0, 1, -1)$

for $\mathbb{R}^8$

- used in image compression

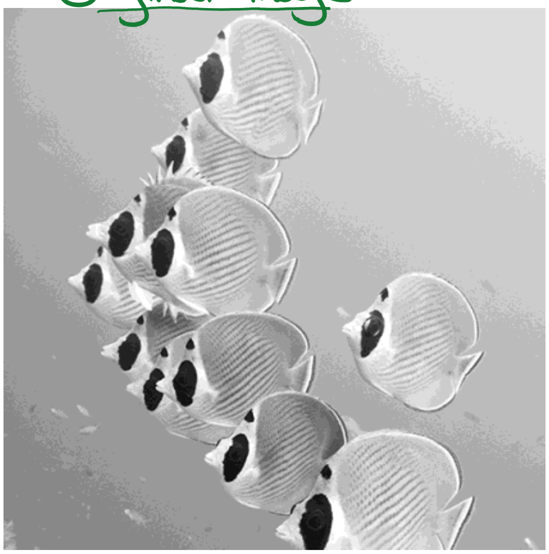
a vector in $\mathbb{R}^8$ with few jumps will have a **sparse** representation in this basis, e.g.,

$$\begin{aligned}(1, 1, 1, 0, -2, -2, -2, -2) &= -\tfrac{5}{8}(1, 1, 1, 1, 1, 1, 1, 1) \\ &+ \tfrac{11}{8}(1, 1, 1, 1, -1, -1, -1, -1) \\ &+ \tfrac{1}{4}(1, 1, -1, -1, 0, 0, 0, 0) \\ &+ \tfrac{1}{2}(0, 0, 1, -1, 0, 0, 0, 0)\end{aligned}$$

after discarding 0.9n

$+\frac{1}{2}(0,0,1,-1,0,0,0,0)$

Original image



... after discarding 0.9n smallest-magnitude Haar basis coordinates

(5) Fourier basis for $\mathbb{C}^n$

-used everywhere!

$(\vec{v}_0, \vec{v}_1, \ldots, \vec{v}_{n-1})$

where $\vec{v}_j = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \exp\left(\frac{2\pi i}{n} \cdot jk\right) \vec{e}_k$

$$[I]_{\substack{\text{Fourier} \\ \to \text{std basis}}} = \frac{1}{\sqrt{n}} \begin{array}{c} \\ \vec{e}_0 \\ \vec{e}_1 \\ \\ \\ \\ \vec{e}_{n-1} \end{array} \begin{array}{c} \begin{array}{ccccc} \vec{v}_0 & \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots \\ 1 & \omega & \omega^2 & \omega^3 & \cdots \\ 1 & \omega^2 & \omega^4 & \omega^6 & \cdots \\ 1 & \omega^3 & \omega^6 & \omega^9 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \omega^{3(n-1)} & \cdots \end{pmatrix} \end{array}$$

where $\omega = \exp\left(\frac{2\pi i}{n}\right)$

$= \cos\frac{2\pi}{n} + i \sin\frac{2\pi}{n}$

$$\Rightarrow \vec{v}_j \cdot \vec{v}_{j'} = \frac{1}{n} \sum_{k,k'=0}^{n-1} \exp\left(\frac{2\pi i}{n}(-jk + j'k')\right) \underbrace{(\vec{e}_k \cdot \vec{e}_{k'})}_{\delta_{kk'}}$$

$$= \frac{1}{n} \sum_{k=0}^{n-1} \exp\left(\frac{2\pi i}{n} k(j'-j)\right)$$

case $j=j'$: $\vec{v}_j \cdot \vec{v}_j = \frac{1}{n}(1+1+\cdots+1) = 1$

case $j \neq j'$: (geometric series ...

case $j \neq j'$:

$$\vec{v}_j \cdot \vec{x}_{j'} = \frac{1}{n} \cdot \left(\begin{array}{l} \text{geometric series} \\ 1 + \omega^{j'-j} + \omega^{2(j'-j)} + \cdots + \omega^{(n-1)(j'-j)} \end{array} \right)$$

where $\omega = e^{\frac{2\pi i}{n}} = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$

$$1 + r + r^2 + \cdots + r^{n-1} = \frac{1-r^n}{1-r}$$

$= 0$ since $\omega^n = 1$

⑥ Various bases for functions

e.g., Hermite, Laguerre, Chebyshev

sines and cosines, ...

There are lots of named examples because

Sets of pairwise orthogonal vectors are nice!

$\{\vec{v}_1, \ldots, \vec{v}_n\}$ with $\vec{v}_i \cdot \vec{v}_j = 0$ for $i \neq j$