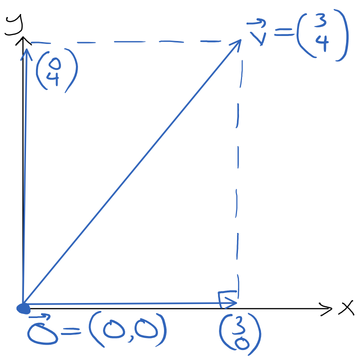
Lecture 1: Geometry in R^n (class)

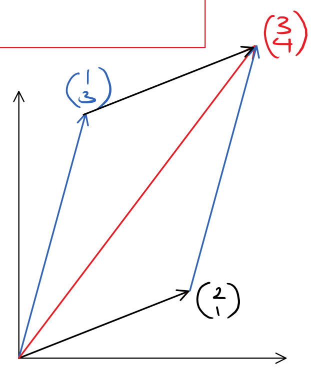
Reading:  2-3  5

1. Vectors lengths & angles, projections
2. Lines, planes, hyperplanes
   dimension, codimension
   angles & orthogonaity
3. High-dimensional geometry

Randomized projections, Johnson-Lindenstrauss
Finite fields

VECTORS





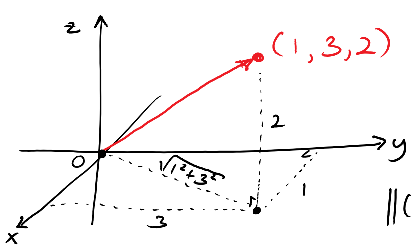
LENGTHS

- The length of a vector $\vec{v} \in \mathbb{R}^n$ (or $\mathbb{C}^n$) is

$$\|v\| = \sqrt{|v_1|^2 + \cdots + |v_n|^2}.$$

Example:

(*Note: We will define other lengths later.)



$$\|(1,3,2)\| = \sqrt{1^2+3^2+2^2} = \sqrt{14}$$

Observe: • For $v \in \mathbb{R}^n$, $\|v\|^2 = v^T v$

$= (\text{———} v^T \text{———}) \begin{pmatrix} | \\ v \\ | \end{pmatrix}$

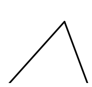
• Scaling: $\|\alpha \cdot \vec{v}\| = |\alpha| \cdot \|\vec{v}\|$

$\Rightarrow \dfrac{v}{\|v\|}$ is a unit vector

• For $v \in \mathbb{C}^n$, $\|v\|^2 = \bar{v}^T v$,

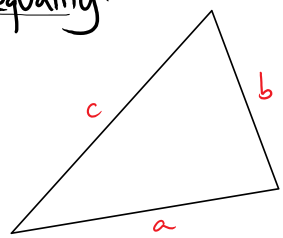
e.g. $\|(1,i)\|^2 = 1^2 + |i|^2 = 2$

Triangle inequality:

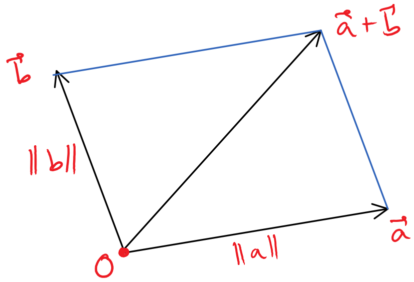


$a+b \geq c$
$a+c \geq b$

Triangle inequality:

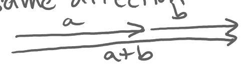


$a+b \geq c$
$a+c \geq b$
$b+c \geq a$



$$\boxed{\|\vec{a}+\vec{b}\| \leq \|\vec{a}\| + \|\vec{b}\|}$$

with equality if and only if a and b point in the same direction



Inner Products & Angles

- The **inner product** of two vectors $\vec{u},\vec{v} \in \mathbb{R}^n$ is

$$\boxed{\vec{u}\cdot\vec{v} = u^T v = \sum_{i=1}^{n} u_i v_i} \in \mathbb{R}$$

For two complex vectors $\vec{u},\vec{v} \in \mathbb{C}^n$,
$$\vec{u}\cdot\vec{v} = \sum_i \bar{u}_i v_i \in \mathbb{C}.$$
(Other notation: dot product, scalar product, $\langle u | v \rangle$, (u,v).)

Observe:
- $\vec{u}\cdot\vec{v} = (\vec{v}\cdot\vec{u})^*$ ← complex conjugate
- $\vec{u}\cdot\vec{u} = \|\vec{u}\|^2$
- It is "bilinear":

$\vec{u}\cdot(\vec{v}+\vec{w}) = \vec{u}\cdot\vec{v} + \vec{u}\cdot\vec{w}$ $\qquad$ $(\vec{u}+\vec{v})\cdot\vec{w} = \vec{u}\cdot\vec{w} + \vec{v}\cdot\vec{w}$

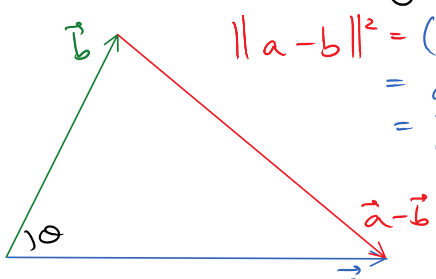
$\vec{u}\cdot(\alpha\vec{v}) = \alpha(\vec{u}\cdot\vec{v})$ $\qquad$ $(\alpha\vec{u})\cdot\vec{v} = \alpha^*(\vec{u}\cdot\vec{v})$

- The **angle** between $\vec{u},\vec{v} \neq \vec{0}$ is

$$\boxed{\cos^{-1}\left(\frac{\vec{u}}{\|u\|}\cdot\frac{\vec{v}}{\|v\|}\right)} \in [0,\pi]$$

They are **orthogonal** if $\vec{u}\cdot\vec{v} = 0$ (angle $= \pi/2$)
AKA perpendicular, $\vec{u} \perp \vec{v}$

Observe: This definition agrees with what we know from geometry:



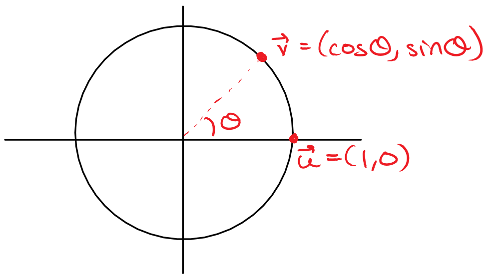
$$\|a-b\|^2 = (a-b)\cdot(a-b)$$
$$= a\cdot a + b\cdot b - 2a\cdot b$$
$$= \|a\|^2 + \|b\|^2 - 2\|a\|\,\|b\|\cos\theta$$

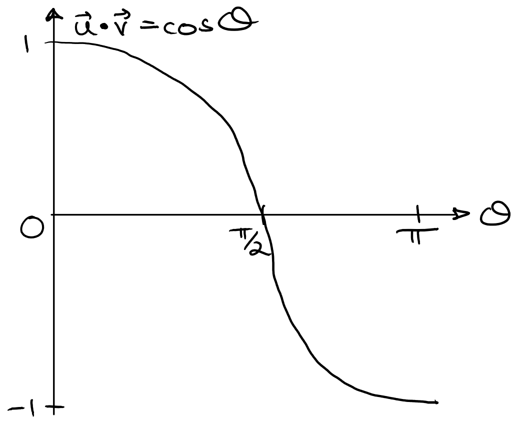
$\vec{a}$

- **Cauchy-Shwarz inequality:**

$$|\vec{u}\cdot\vec{v}| \leq \|\vec{u}\|\cdot\|\vec{v}\|$$

Example:



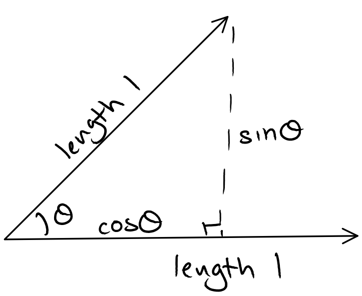


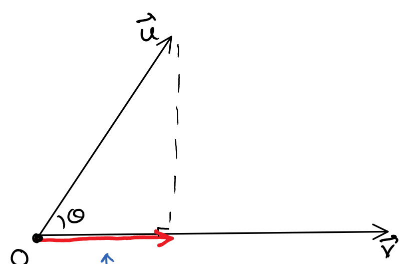
$\vec{u}\cdot\vec{v} = \|\vec{u}\|\cdot\|\vec{v}\|$ if they point in the same direction

$\vec{u}\cdot\vec{v} = -\|\vec{u}\|\cdot\|\vec{v}\|$ if they point in opposite directions

otherwise, $|\vec{u}\cdot\vec{v}| < \|\vec{u}\|\cdot\|\vec{v}\|$ strictly

Observe: The magnitude $|\vec{u}\cdot\vec{v}|$ measures the "overlap" between $\vec{u}$ and $\vec{v}$:





projection of $\vec{u}$ in the direction of $\vec{v}$

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

$$= \boxed{\frac{(\vec{u}\cdot\vec{v})}{\|\vec{v}\|^2}\,\vec{v}}$$

– the projection is 0 if $\vec{u}\cdot\vec{v} = 0$, or $\theta = \frac{\pi}{2}$

$\Rightarrow$ If $\|\vec{v}\| = 1$,

Projection of $\vec{u}$ in the direction of $\vec{v}$

$$= (\vec{u}\cdot\vec{v})\,\vec{v}$$

- Matlab commands:

```
>> v = randn(5, 1)

v =

   -0.5416
   -0.4010
    1.2395
   -0.9094
    0.3267
```

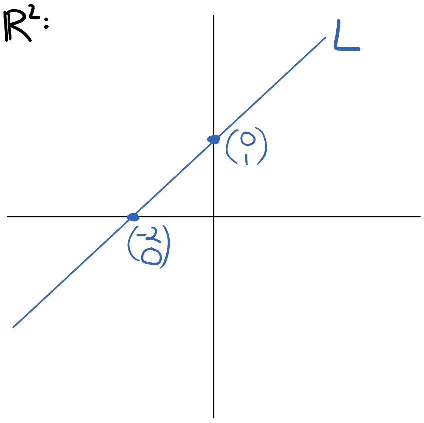
```
>> sqrt(v' * v)

ans =

    1.7100
```

```
>> sqrt(dot(v, v))

ans =

    1.7100
```

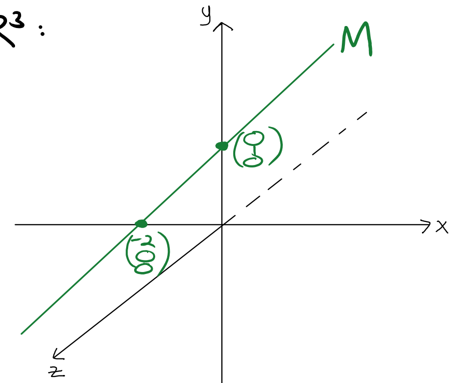
```
>> norm(v)

ans =

    1.7100
```

Lines, Planes, Hyperplanes

$\mathbb{R}^2$:



$\mathbb{R}^3$:



Two ways of specifying a line:

Constructive/parametric

2 points determine a line

$$L = \left\{ \begin{pmatrix} -2 \\ 0 \end{pmatrix} + t \begin{pmatrix} 2 \\ 1 \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

e.g., $t = 0$ gives $\begin{pmatrix} -2 \\ 0 \end{pmatrix}$

$t = 1$ gives $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$$= \left\{ \begin{pmatrix} -1 \\ 1/2 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\}$$

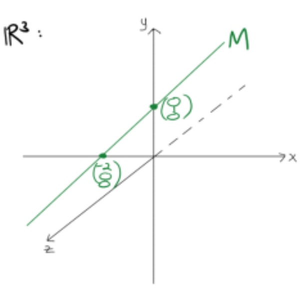
Via constraints

$$L = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,\middle|\, y = \underbrace{\tfrac{1}{2}}_{\text{slope}} x + \underbrace{1}_{\text{y-axis intercept}} \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,\middle|\, x - 2y = -2 \right\}$$

$$= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,\middle|\, 10x - 20y = -20 \right\}$$

$\mathbb{R}^3$:

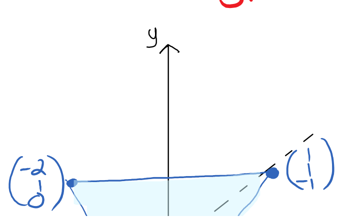


$$M = \left\{ \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix} + \underbrace{t}_{\text{one degree of freedom}} \left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix} \right] \,\middle|\, t \in \mathbb{R} \right\}$$

$$M = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,\middle|\, \begin{array}{c} x - 2y = -2 \\ z = 0 \end{array} \right\}$$

2 <u>independent</u> equations determine a line in 3D

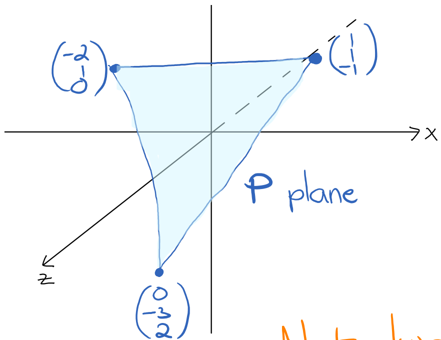
Planes and hyperplanes



Constructive:

3 points determine a 2D plane

($d+1$ points determine a d-dimensional plane)



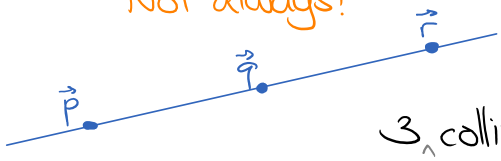
(d+1 points determine a d-dimensional plane)

$$P = \left\{ \vec{p} + s(\vec{q}-\vec{p}) + t(\vec{r}-\vec{p}) \;\middle|\; s,t \in \mathbb{R} \right\}$$

$s=0, t=0 \to \vec{p}$
$s=1, t=0 \to \vec{q}$
$s=0, t=1 \to \vec{r}$

2 degrees of freedom
$\Rightarrow$ 2D plane

Not always!



3 (or more) collinear points $\Rightarrow$ line

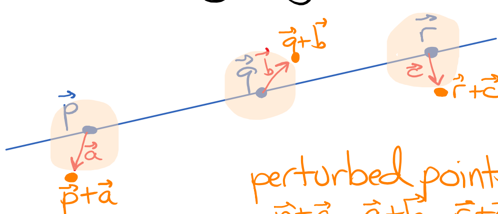
(this is called a degeneracy)

d+1 nondegenerate points $\Rightarrow$ d-dimensional plane

Observe: ① In $\mathbb{R}^n$, a set of up to n+1 "random" points will almost certainly be nondegenerate.

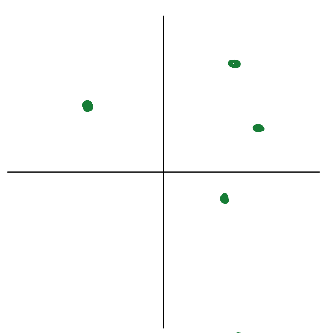
e.g., in $\mathbb{R}^2$, $\mathbb{P}[\text{3 random pts are collinear}] = 0$.

② Randomly perturbing points will usually break any degeneracy (if the dim. is enough)



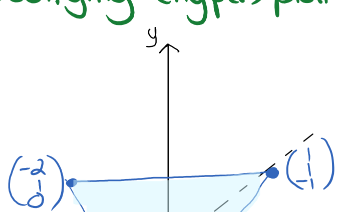
perturbed points
$\vec{p}+\vec{a}, \vec{q}+\vec{b}, \vec{r}+\vec{c}$
are no longer collinear
$\Rightarrow$ determine 2D plane

③ Too many points will always be degenerate.



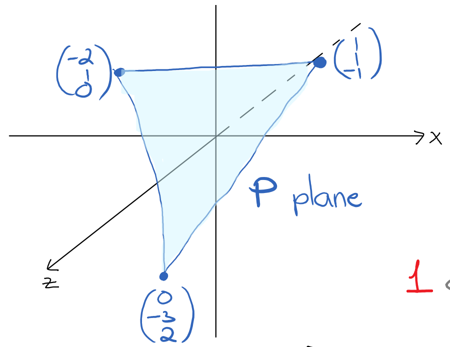
e.g., 4 points in $\mathbb{R}^2$
do not give a 3D plane;
there isn't space!
(they must be coplanar)

Specifying (hyper)planes via constraints:



$$P = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x+2y+3z=0 \right\}$$

Observe:



Observe:

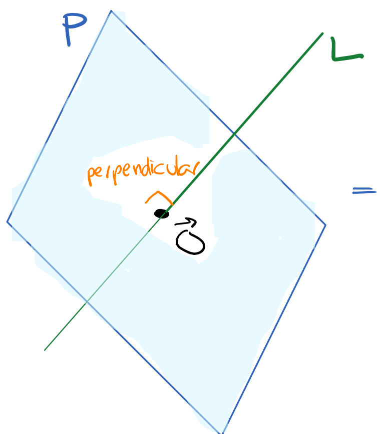
① 1 equation in $\mathbb{R}^3$
$\Rightarrow$ P has dimension $3-1=2$
("P has codimension 1")

1 eqn. in $\mathbb{R}^n$ $\Rightarrow$ $n-1$ dimensional plane

② $x+2y+3z = \begin{pmatrix}x\\y\\z\end{pmatrix}\cdot\begin{pmatrix}1\\2\\3\end{pmatrix}$

$$\therefore\ P=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{R}^3 \,\middle|\, x+2y+3z=0\right\} = \left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{R}^3 \,\middle|\, \begin{pmatrix}x\\y\\z\end{pmatrix}\cdot\begin{pmatrix}1\\2\\3\end{pmatrix}=0\right\}$$

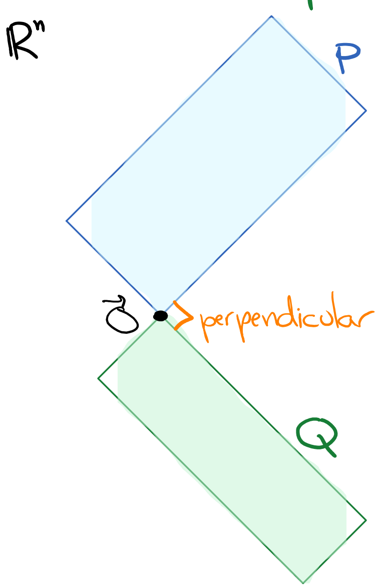
$$=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix} \,\middle|\, \begin{pmatrix}x\\y\\z\end{pmatrix}\perp\begin{pmatrix}1\\2\\3\end{pmatrix}=0\right\}$$



$=\{$set of vectors perpendicular to $\begin{pmatrix}1\\2\\3\end{pmatrix}\}$
or $\perp$ to $\begin{pmatrix}2\\4\\6\end{pmatrix}$, or $\begin{pmatrix}-1\\-2\\-3\end{pmatrix}$, or...

= set of vectors perpendicular to the line through $\vec{0}$ and $\begin{pmatrix}1\\2\\3\end{pmatrix}$

Planes and complements, in general



In $\mathbb{R}^n$, a d-dimensional hyperplane P through $\vec{0}$ can be specified either

- constructively: by giving d nondegen. points in P
- via constraints: by specifying Q with $n-d$ points, and saying that P is everything $\perp$ to Q

Example: codimension 2

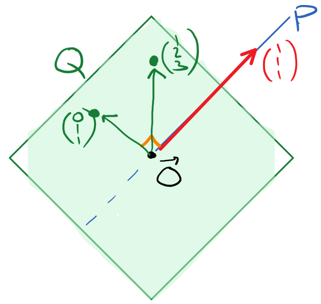
Let P be the set of solutions to

$$\begin{aligned} x+2y+3z&=0 \\ y+z&=0 \end{aligned} \quad \updownarrow_{-3} \longleftrightarrow \quad \begin{aligned} x-y&=0 \\ y+z&=0 \end{aligned}$$

$$P=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{R}^3 \,\middle|\, \begin{pmatrix}1&2&3\\0&1&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\right\}$$

$$=\left\{t\begin{pmatrix}1\\1\\-1\end{pmatrix} \,\middle|\, t\in\mathbb{R}\right\}$$ = set of points perpendicular to Q,

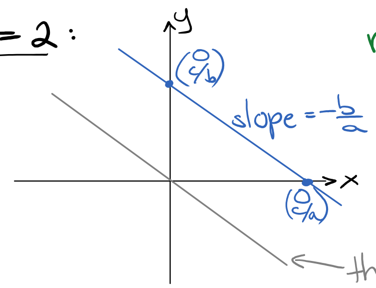
$$= \left\{ t\begin{pmatrix}1\\1\\1\end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\}$$ = set of points perpendicular to Q, the plane through $\vec{0}$, $\begin{pmatrix}1\\2\\3\end{pmatrix}$, $\begin{pmatrix}0\\1\\1\end{pmatrix}$



Geometry of m equations on n variables

n=1: one variable (trivial)

n=2:

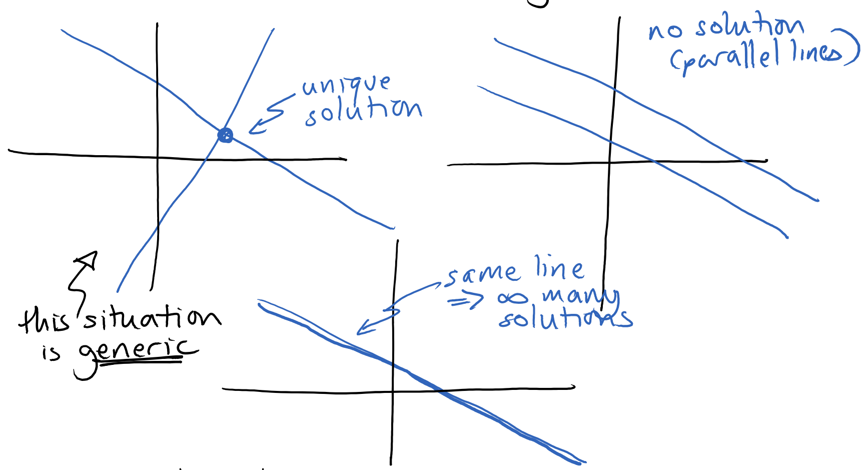


m = 1

$$ax + by = c$$

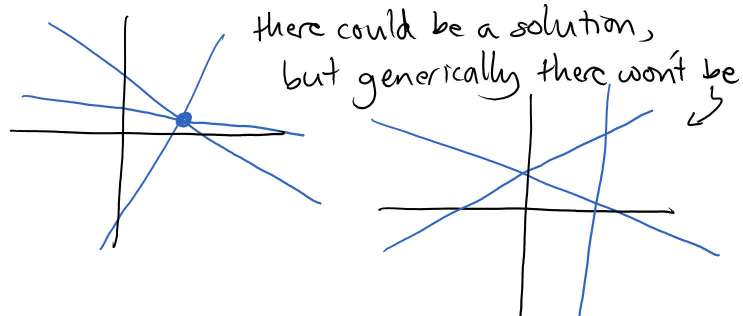
a line

m = 2 equations: $a_{11}x + a_{12}y = c_1$, $a_{21}x + a_{22}y = c_2$ } two lines

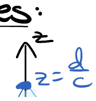


m = 3: three lines in $\mathbb{R}^2$

there could be a solution, but generically there won't be



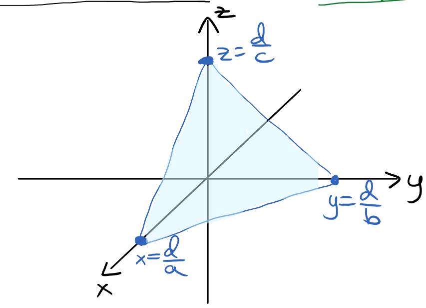
n = 3 variables:



m = 1 equation:

$$ax + by + cz = d$$

n=3 variables:



m=1 equation:

$$ax+by+cz=d$$

determines a 2D plane!

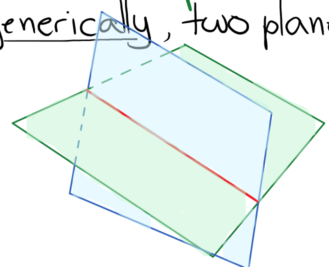
Special case: $a_{11}=0$

$\Rightarrow x_1$ doesn't matter

$\Rightarrow$ plane is parallel to x_1 axis

m=2 equations:

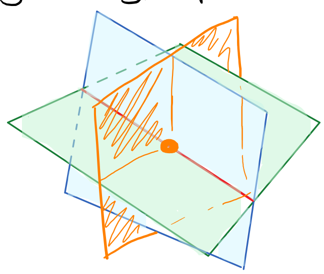
generically, two planes intersect in a line



but parallel planes intersect nowhere,
or the planes could be the same

m=3 equations:

generically, a plane & a line intersect in a point



n variables:

equation $\longleftrightarrow$ (n−1)-dimensional hyperplane
(codimension 1)

generically, the intersection of

m equations $\longleftrightarrow$ (n−m)-dim. plane
(codimension m)

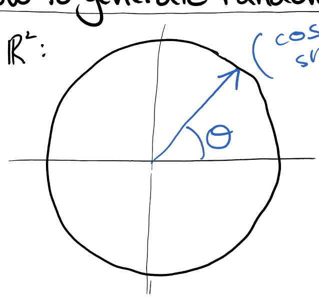
n equations $\longleftrightarrow$ a point (0 dimensional)

Geometrically, what does Gaussian elimination correspond to?

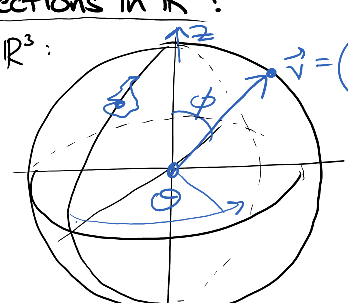
HIGH-DIMENSIONAL VECTORS

How to generate random directions in $\mathbb{R}^n$?

$\mathbb{R}^2$: $\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}$



$\mathbb{R}^3$: $\vec{v} = (\sin\phi\cos\theta,\ \sin\phi\sin\theta,\ \cos\phi)$



"direction" = unit vector $\|\vec{v}\| = 1$

= point on the unit sphere

$\mathbb{R}^4$: $(\sin\phi \sin\theta \sin\psi,\ \sin\phi \sin\theta \cos\psi,\ \sin\phi \cos\theta,\ \cos\phi)$

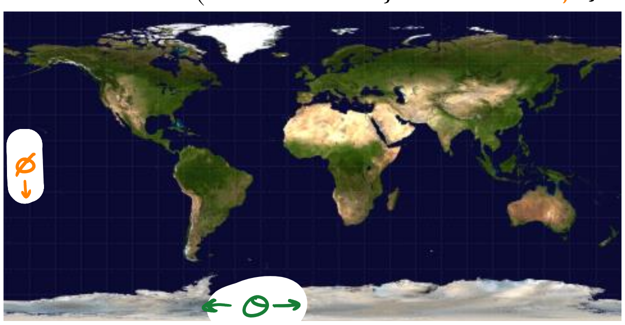
```
>> help rand
 rand Uniformly distributed pseudorandom numbers.
    R = rand(N) returns an N-by-N matrix containing pseudorandom values drawn
    from the standard uniform distribution on the open interval(0,1).  rand(M,N)
    or rand([M,N]) returns an M-by-N matrix.  rand(M,N,P,...) or
    rand([M,N,P,...]) returns an M-by-N-by-P-by-... array.  rand returns a
    scalar.  rand(SIZE(A)) returns an array the same size as A.
```

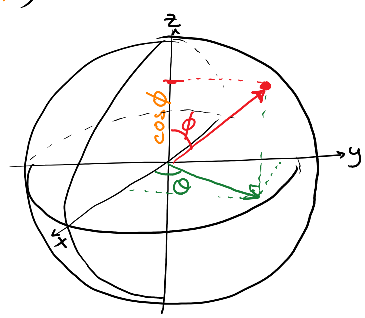
$\mathbb{R}^2$: $\theta = 2\pi \cdot \text{rand}(1)$

$\vec{v} = (\cos\theta, \sin\theta)$ ✓

$\mathbb{R}^3$: ① Use longitude θ and latitude ϕ?

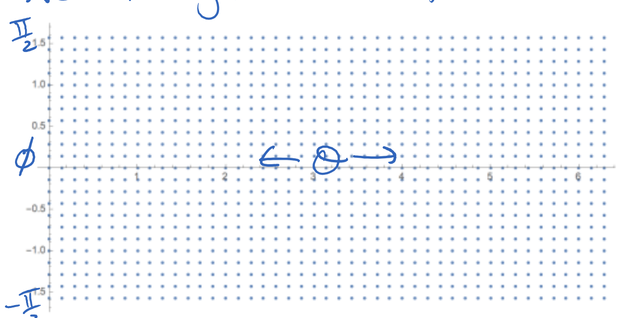
$\vec{v} = (\cos\theta \cdot \sin\phi,\ \sin\theta \cdot \sin\phi,\ \cos\phi)$

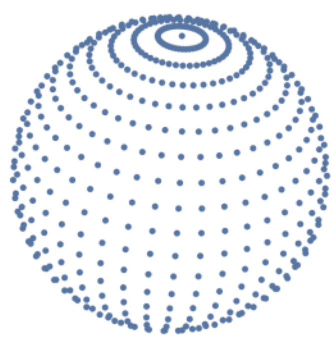




$\mathbb{R}^4$: $(\cos\psi \cos\theta \sin\phi,\ \sin\psi \cos\theta \sin\phi,\ \sin\theta \sin\phi,\ \cos\phi)$, etc.

No! This gives a non-uniform distribution



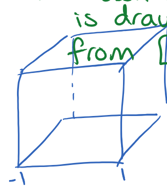


(points near the poles occur too often)

② Pick 3 coordinates at random, and renormalize?

```
>> v = 2 * rand(3, 1) - 1

v =

    0.8268
    0.2647
   -0.8049

>> v = v / norm(v)

v =

    0.6984
    0.2236
   -0.6799
```

← each coordinate is drawn uniformly from $[-1, 1]$.

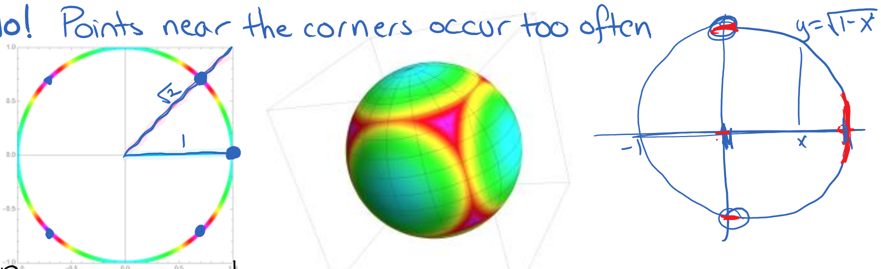


← renormalize so $\|\vec{v}\| = 1$

No! Points near the corners occur too often

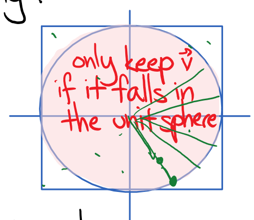
```
0.2236
-0.6799
```

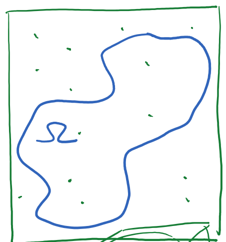
No! Points near the corners occur too often



③ Rejection sampling? https://en.wikipedia.org/wiki/Rejection_sampling

```
n = 18;
length = Inf;
while length > 1
  v = 2 * rand(n,1) - 1;
  length = norm(v);
end
v = v / length;
```



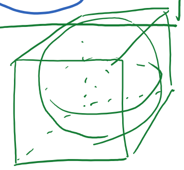


This works, but is FAR too slow.

$$\mathbb{P}_{\vec{v}\in_R[-1,1]^n}\left[\|\vec{v}\|\le 1\right] = \frac{\text{volume(unit sphere)}}{\text{volume}([-1,1]^n)}$$

$$= \frac{\frac{\pi^{n/2}}{(n/2)!}}{2^n} \quad \text{if } n \text{ is even}$$

(https://en.wikipedia.org/wiki/N-sphere#Volume_and_surface_area)



$$\sim \frac{\left(\frac{\pi}{\frac{n}{2e}}\right)^{n/2}}{2^n} \quad \text{(Stirling's approximation)}$$

https://en.wikipedia.org/wiki/Stirling%27s_approximation

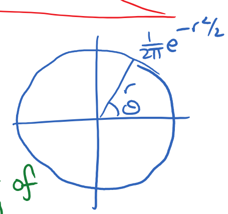
$$= \left(\frac{\pi e}{2n}\right)^{n/2} \longrightarrow 0 \quad \text{exponentially fast in } n.$$

Better: Use the normal distribution

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$\Rightarrow \underbrace{p(x)\,p(y)}_{p(z)} = \frac{1}{(2\pi)^{2/2}} e^{-(x^2+y^2)/2}$$

$x^2+y^2 = \|(x,y)\|^2$ independent of the angle $\tan^{-1}\frac{y}{x}$

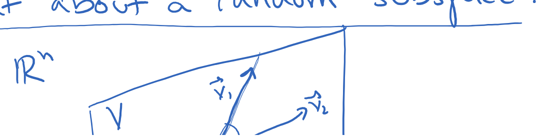


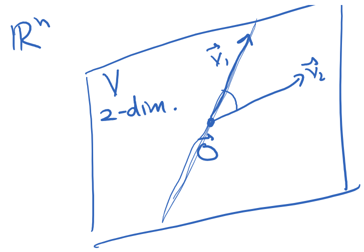

```
>> help randn
 randn Normally distributed pseudorandom numbers.
    R = randn(N) returns an N-by-N matrix containing pseudorandom values drawn
    from the standard normal distribution.  randn(M,N) or randn([M,N]) returns
    an M-by-N matrix. randn(M,N,P,...) or randn([M,N,P,...]) returns an
    M-by-N-by-P-by-... array. randn returns a scalar.  randn(SIZE(A)) returns
    an array the same size as A.
```

```
>> n = 5; v = randn(n, 1); v = v / norm(v)

v =

   -0.0912
    0.0721
    0.7527
    0.5825
   -0.2839
```

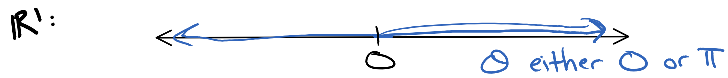
✓

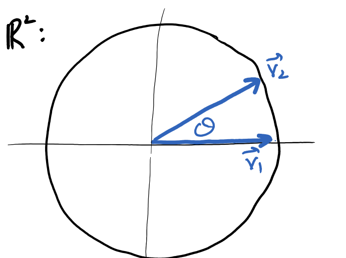
What about a random subspace?



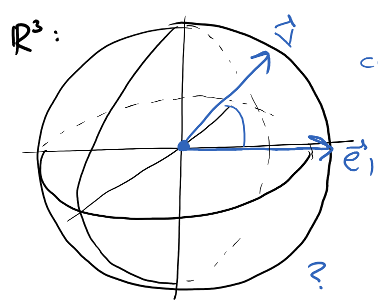


High-dimensional intuition: What is the angle between two random unit vectors in $\mathbb{R}^n$?





all angles $\theta \in [0, \pi]$ are equally likely



$\cos^{-1}(\vec{v} \cdot \vec{e}_1)$
$= \cos^{-1}(v_1)$

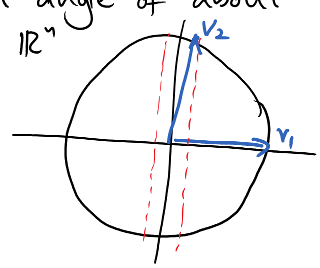
in $\mathbb{R}^n$?

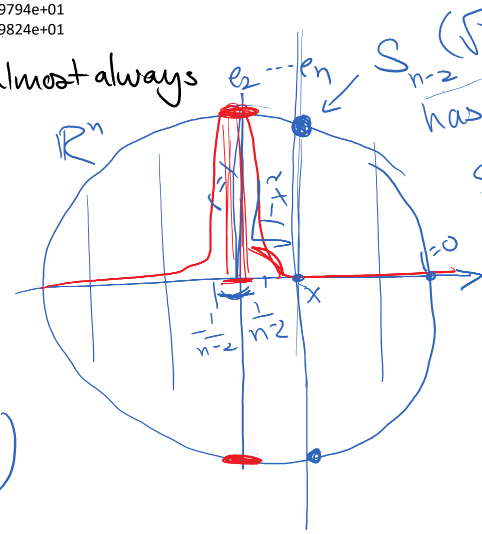
```
trials = 100;
for pow = 1:15
 n = 2^pow;
 minangle = 180;
 for i = 1:trials
  x = randn(n,1);
  y = randn(n,1);
  angle = 180/pi * acos(x'*y / sqrt((x'*x) * (y'*y)));
  minangle = min(angle, minangle);
 endfor
 disp([n, minangle]);
endfor
```

} two random n-dim vectors

n	minangle
2.00000	0.87482
4.0000	16.8946
8.0000	36.7159
16.000	54.869
32.000	57.955
64.000	75.011
128.000	78.256
256.000	79.392
512.000	83.195
1024.000	84.732
2048.000	85.614
4096.000	87.519
8192.000	87.944
1.6384e+04	8.8889e+01
3.2768e+04	8.9228e+01
6.5536e+04	8.9405e+01
1.3107e+05	8.9468e+01
2.6214e+05	8.9661e+01
5.2429e+05	8.9794e+01
1.0486e+06	8.9824e+01

$\Rightarrow$ In high dimensions, random vectors are almost always at an angle of about 90 degrees.

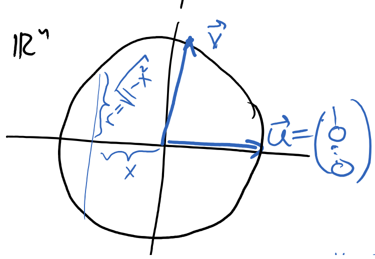




$S_{n-2}\left(\sqrt{1-x^2}\right)$ has surface area $c \cdot r^{n-2}$

$\left(1 - \frac{1}{n}\right)^n \approx \frac{1}{e}$

Why?

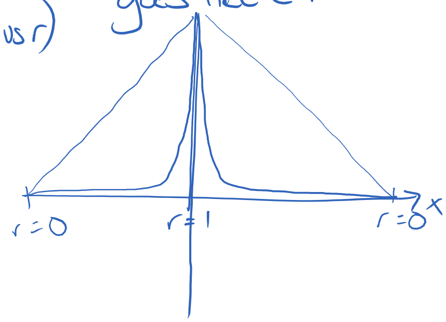


$\theta = \cos^{-1}(v_1)$

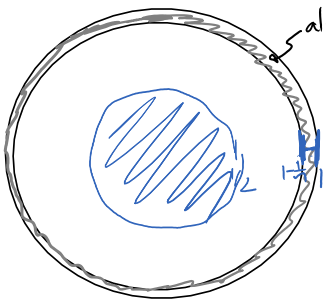
surface area goes like $c \cdot r^{n-2}$

n	$\text{Vol(ball of radius } r)$
1	$2r$
2	πr^2
3	$\frac{4}{3}\pi r^3$
$\vdots$	
n	$c \cdot r^n$

surface area goes like $c \cdot r^{n-2}$



① Almost all the volume of the unit ball is near the surface



almost everything!

$V(\frac{1}{2}, n) = \frac{1}{2^n} V(1,n)$

$\frac{V(1-\frac{1}{n})}{V(1,n)} = (1-\frac{1}{n})^n \approx \frac{1}{e} \approx 30\%$

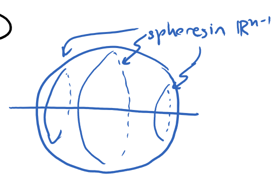
Why? $V(r,n) = v_n \cdot r^n$

$\Rightarrow$ fraction of unit ball with length $\leq r$

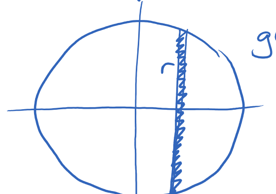
$$= \frac{V(r,n)}{V(1,n)} = \frac{v_n \cdot r^n}{v_n} = r^n$$

In a high-dimensional game of darts, it is really hard to get a bullseye!

②



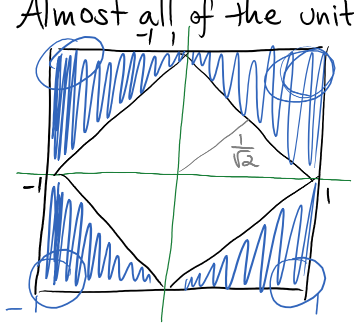
the total surface area here goes like r^{n-2}



$\Rightarrow$ since r^{n-2} drops so quickly as r drops from 1 to 0, almost all the surface area is concentrated in the middle (along the "equator")

Similarly:

Almost all of the unit hypercube is near the corners.



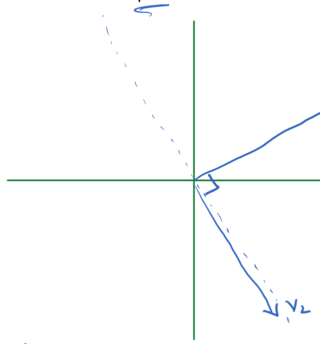
volume(outer cube) $= 2^n$

volume (inscribed cube) $= \left(\frac{2}{\sqrt{n}}\right)^n$

$\Rightarrow$ ratio of volumes $= \frac{1}{n^{n/2}} \xrightarrow[n\to\infty]{} 0$

Question: How many pairwise orthogonal vectors can we fit in

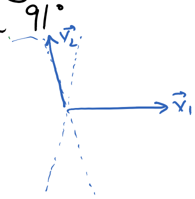
		Answer:
$\mathbb{R}^2$	?	2
$\mathbb{R}^3$	?	3
$\mathbb{R}^n$	?	n

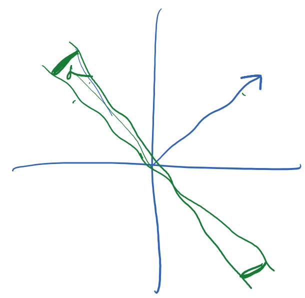


no room for a third vector orthogonal to $\vec{v}_1$ and $\vec{v}_2$

(we'll prove this later)

Trivia: How many pairwise almost-orthogonal vectors can we fit in $\mathbb{R}^n$?
(say, the angle between any pair of vectors is between 89° and 91°)





Answer:

$\mathbb{R}^2$: 2 vectors

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

⋮

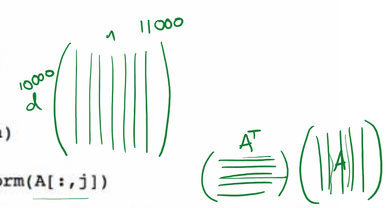
$\mathbb{R}^n$ (for n large): exp(n) vectors !!!

Totally counterintuitive!

Again because for n large, the angle between two randomly chosen directions is close to 90° (with high probability).

Example:

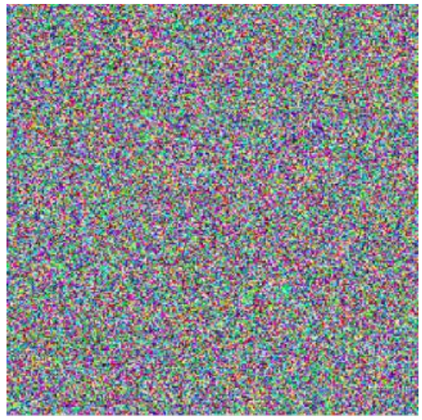
```
import numpy as np

d = 10000   # dimensions
n = 11000   # vectors

A = np.random.randn(d, n)
for j in range(n):
  A[:,j] /= np.linalg.norm(A[:,j])

dots = A.T.dot(A)  # pairwise dot products v_i . v_j
for j in range(n):
  dots[j,j] = 0    # we don't care about v_i . v_i = 1

# least and largest angles between *any* pairs v_i, v_j
180/np.pi * np.arccos((np.max(dots), np.min(dots)))

array([86.87586219, 93.10179547])
```



$O(n^2 d)$ time

Careful! Vectors from real data often do not "look like" random vectors

Example:



```
>> n = 256; image = rand(n, n, 3);
>> imshow(image)
```

RGB components

What's the difference?

one big difference: look from different directions....