Lecture 7

Monday, September 18, 2023 10:02 AM

Normal Distribution (Gaussian)

$$f(\vec{x} \mid \vec{\mu}, \Sigma) = \frac{1}{(\sqrt{2\pi})^d} \frac{1}{\sqrt{|\Sigma|}} e^{-\frac{(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})}{2}}$$

$\vec{\mu}$: mean, $\mathbb{R}^d$

Σ: cov, $d \times d$

$$|\Sigma| \equiv \det(\Sigma) = \prod_{i=1}^{d} \lambda_i \quad \text{product of all eigenvalues}$$

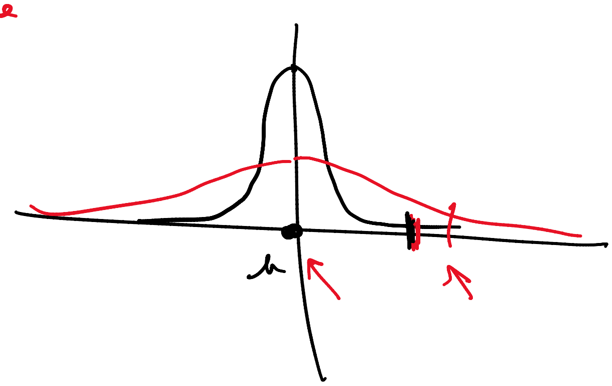
$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$$

$$\det\left(\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}\right) = a \cdot \det\begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \det\begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \det\begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

1d

$$P(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

distance



$$f(\mu) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2}}$$

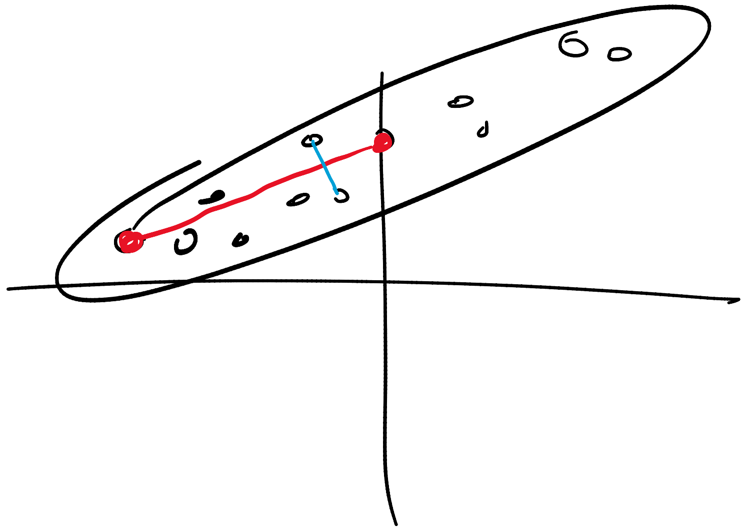
Mahalanobis Distance: $(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})$

Mahalanobis Distance : $(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})$

Euclidean Distance $\quad (\vec{x}-\vec{\mu})^T \, \underline{I} \, (\vec{x}-\vec{\mu})$

$$= \|\vec{x}-\vec{\mu}\|^2$$

generalization of squared distance



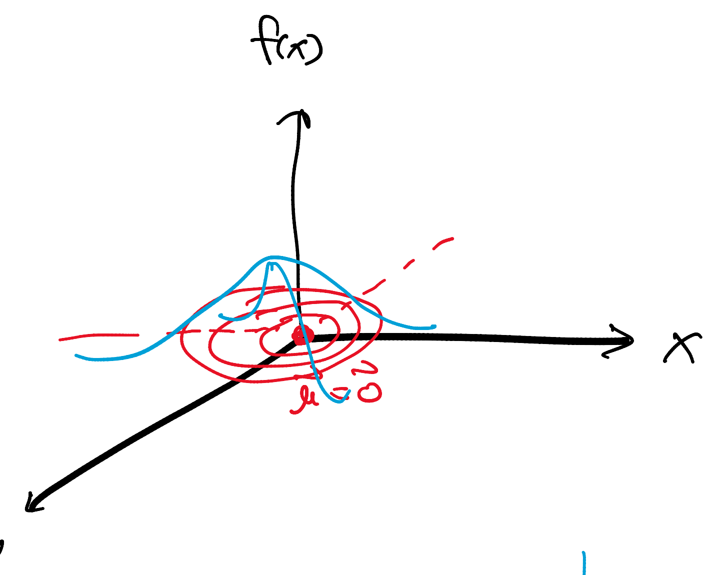
$$Z = \frac{x-\mu}{\sigma}$$

Z-score

how many standard deviations apart

Geometry of Normal

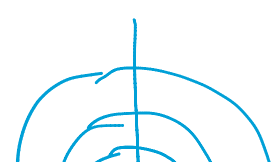
$\vec{x} \in \mathbb{R}^2$



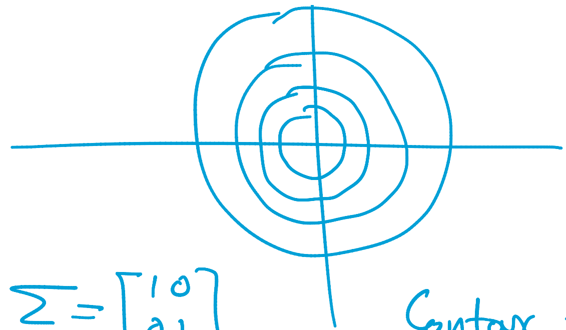
Standard Normal

$\vec{\mu} = 0$

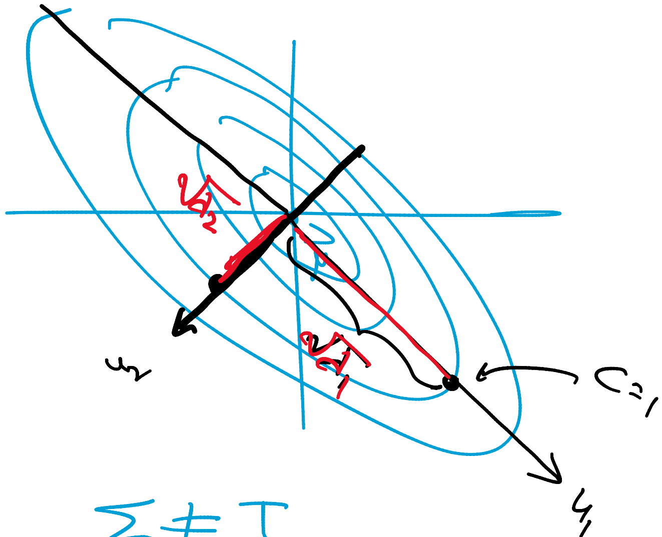
$$\Sigma = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



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



$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ Contour plot

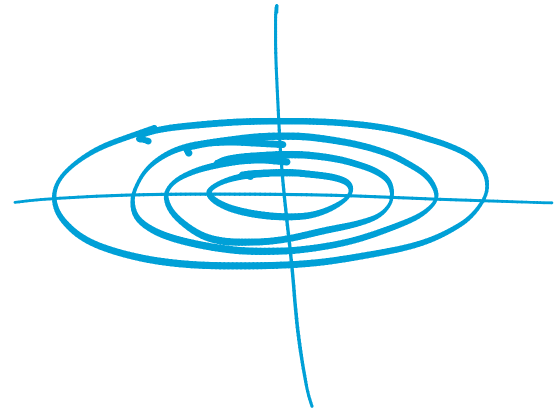


$\Sigma \neq I$

non-diagonal

full cov

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$



$\Sigma \neq I$

but diagonal matrix

$\sigma_{ij} = 0$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

$$\boxed{(\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu}) = 1}$$

equation for an ellipsoid in d- dimensions

eigenvectors give the primary axes of the ellipsoid

and the eigenvalue is related to the length of the axis

$\sqrt{\lambda_i}$

$$\Sigma = U \Lambda U^T = \underbrace{\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_d \\ | & | & & | \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_d \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_d^T \end{bmatrix}}_{U^T}$$

$$\Sigma^{-1} = (U \Lambda U^T)^{-1} = (U^T)^{-1} \Lambda^{-1} U^{-1} = (U^{-1})^{-1} \Lambda^{-1} U^T$$
$$= U \Lambda^{-1} U^T$$

U is orthogonal $\equiv U^{-1} = U^T$

$$\Sigma = U \Lambda U^T = U \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & \ddots & \lambda_d \end{bmatrix} U^T$$

$$\Sigma^{-1} = U \Lambda^{-1} U^T = U \begin{bmatrix} 1/\lambda_1 & & 0 \\ & 1/\lambda_2 \ddots & \\ 0 & & 1/\lambda_d \end{bmatrix} U^T$$

$$(\vec{x} - \vec{\mu}) \Sigma^{-1} (\vec{x} - \vec{\mu}) = 1$$

$\Rightarrow$ quadratic equation in $\vec{x}$ $\equiv$ equation of an ellipsoid

$$\Sigma \begin{array}{l} \nearrow U \\ \searrow \Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 \ddots & \\ 0 & & \lambda_d \end{pmatrix} \end{array}$$

Cov

diagonal

$$\Sigma = U \Lambda U^T$$

$$\Sigma = U \Lambda U^T$$

If one of the eigenvalues is 0

$\lambda_k = 0$

1) for all $a > k$ $\lambda_a = 0$

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots \geq \lambda_k \geq \lambda_{k+1} \cdots \geq \lambda_d \geq 0$$

$$0 \longrightarrow 0 \longrightarrow 0$$

2) $\det(\Sigma) = \prod_{i=1}^{d} \lambda_i = 0$

3) Σ^{-1} does not exist

since $\Sigma^{-1} = U \Lambda^{-1} U^T$

$$= U \begin{bmatrix} 1/\lambda_1 & & & & \\ & 1/\lambda_2 & & & \\ & & \ddots & & \\ & & & 1/0 & \\ & & & & 1/0 \ddots \end{bmatrix} U^T$$

λ_k

undefined!

4) we can still compute a pseudo inverse (when $\lambda_k = 0$)

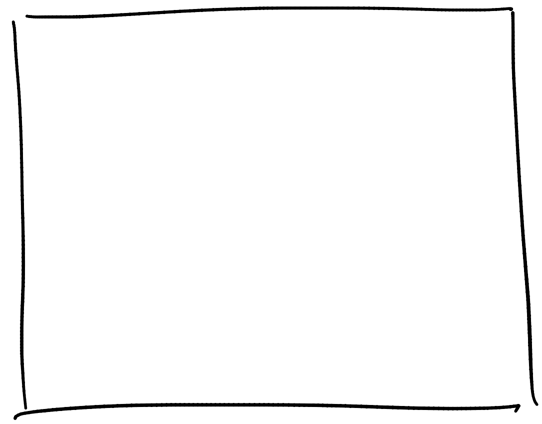
Use the $\underbrace{\text{first}\ k-1}_{\text{largest}}$ eigen vectors & value to compute the inverse

High Dimensional Data space / objects

← d →

items

Users



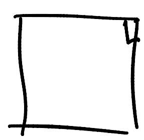
D-dim geometric objects

1) Hypercube 0d

1d

l

2d



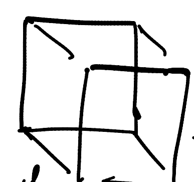
l

l

Square

x-y

3d



l

l

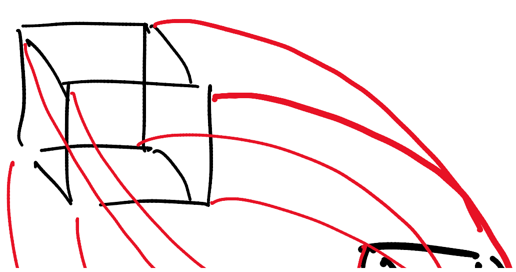
l

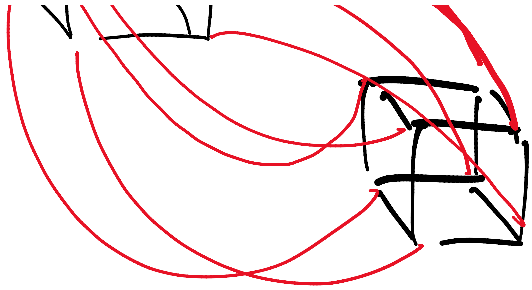
Cube

x-y-z

hypercube 4d

x-y-z-t





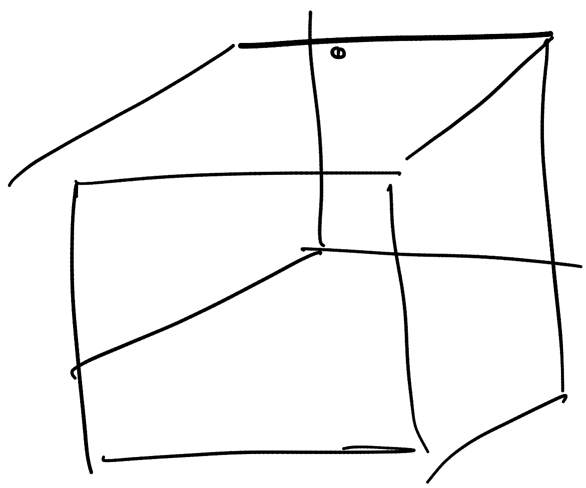
all segments $= \ell$

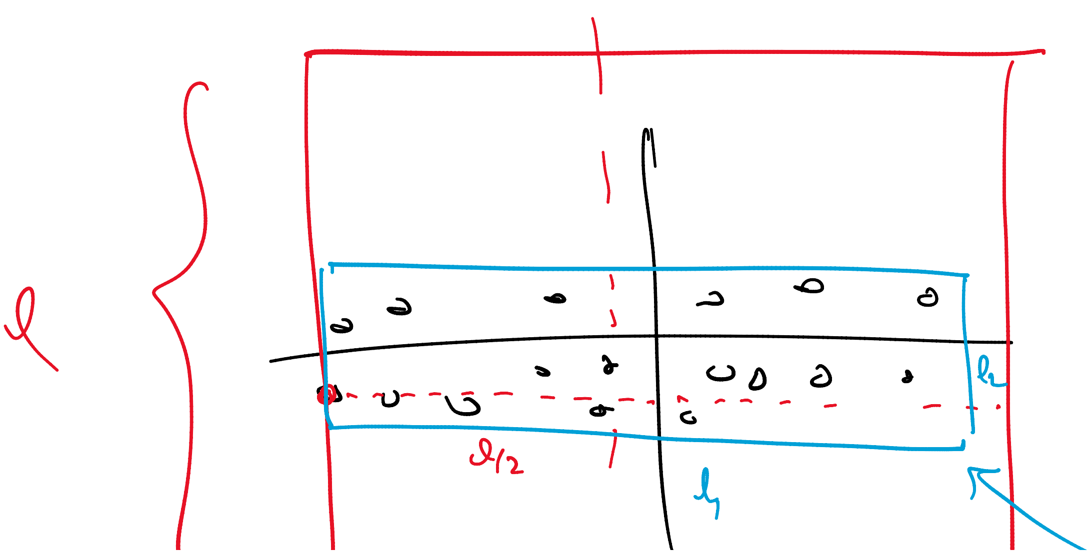
$H_d(\ell) \equiv$ hypercube in d-dim with length ℓ

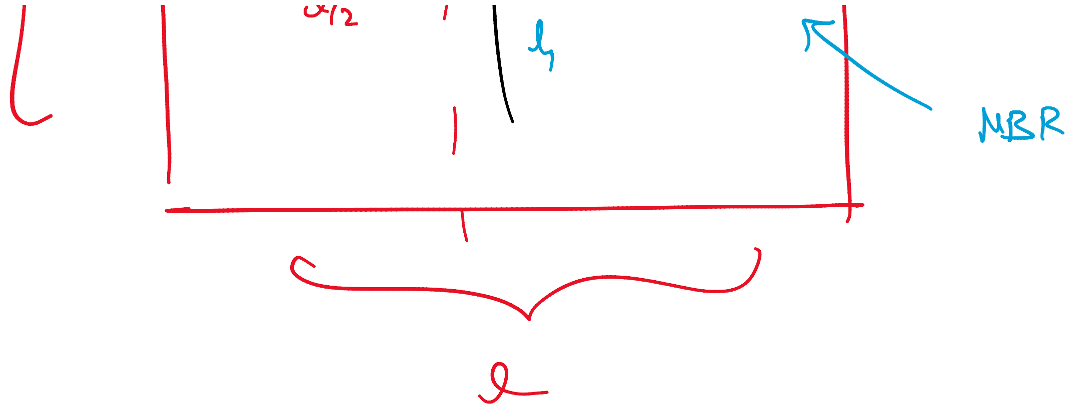
$$\mathrm{Vol}(H_d(\ell)) = \ell^d$$

MBH

Minimum bounding hypercube

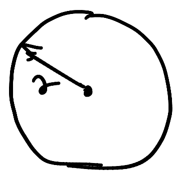






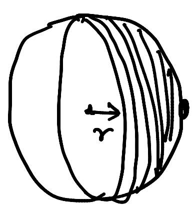
Hypersphere $\longrightarrow$ hyperellipsoid

Circle 2d:



$x-y$

Sphere 3d:



$x-y-z$

$\longrightarrow$ hyperball

$$B_d(r) = \{ \vec{x} \mid \|\vec{x}\|_2 \leq r \}$$

all points within distance r

hyper sphere 4d!

hypersphere

$$S_d(r) = \{ \vec{x} \mid \|\vec{x}\| = r \}$$

exactly at distance r

surface of the hyperball

$$\mathrm{Vol}(S_d(r)) = \mathrm{Vol}(B_d(r))$$

2d Circle $= \pi r^2$

Sphere (3d) $= \left(\frac{4}{3}\pi\right) r^3$

Sphere (3d) $= \left(\frac{1}{3}\pi\right) r$

d-dim

$$\mathrm{vol}\left(S_d(r)\right) = \underbrace{\left(\frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}\right)}_{} r^d$$

$$= k_d \, r^d$$

$k_2 = \pi$

$k_3 = \frac{4}{3}\pi$

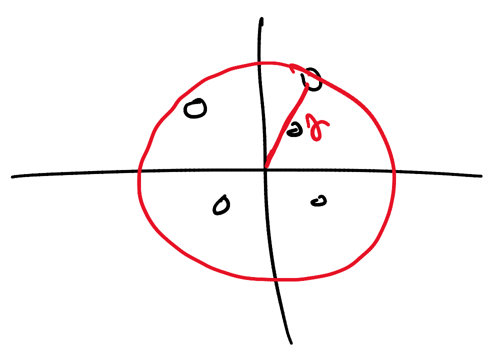
$\vdots$

$$\Gamma\left(\frac{d}{2}+1\right) = \begin{cases} \left(\frac{d}{2}\right)! & \text{if } d \text{ is even} \\ \sqrt{\pi}\left(\frac{d!!}{2^{(d+1)/2}}\right) & \text{if } d \text{ is odd} \end{cases}$$

for integer arguments

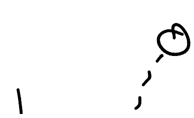
$$n! = n(n-1)(n-2)(n-3)\cdots 3 \cdot 2 \cdot 1$$

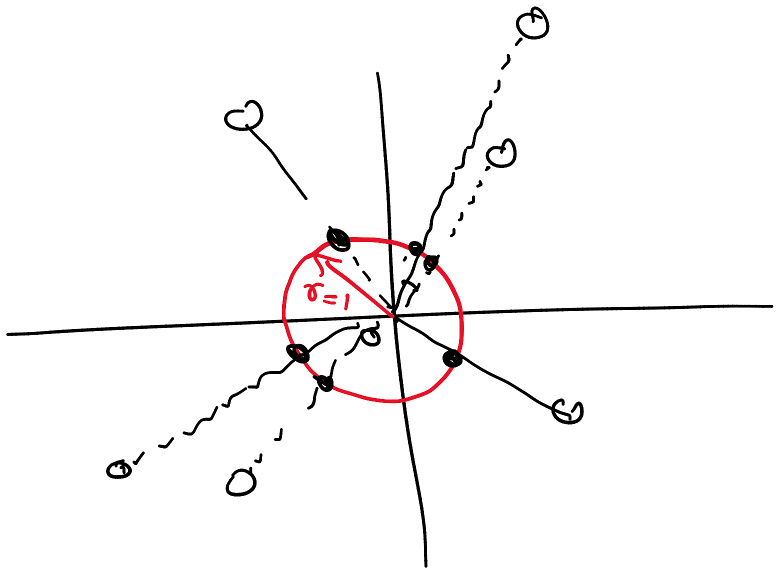
$$n!! = n(n-2)(n-4)\cdots \quad 1$$



$r = \max\{\|\vec{x}\|\}$

Project onto the unit hypersphere!

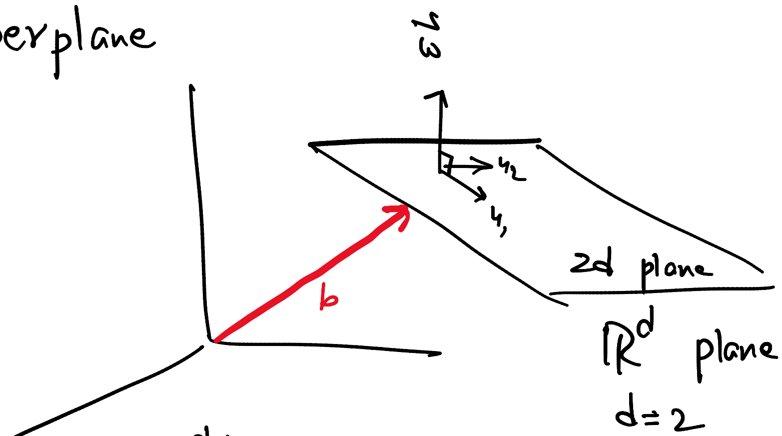




$$\underbrace{\text{dot product}}_{\text{Similarity}}$$

---

hyperplane



$\mathbb{R}^d$ plane

$d=2$

$\vec{w}$ normal vector

i.e. $\perp$ to all vectors on the plane

$\mathbb{R}^{d+1} = \mathbb{R}^3$

$$\boxed{h(\vec{x}) = \vec{w}^T \vec{x} + b = 0}$$

$\mathbb{R}^d$

$\vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{pmatrix}$ $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}$ $b =$ scalar

hyperplane: set of all points $\vec{x}$ that satisfy

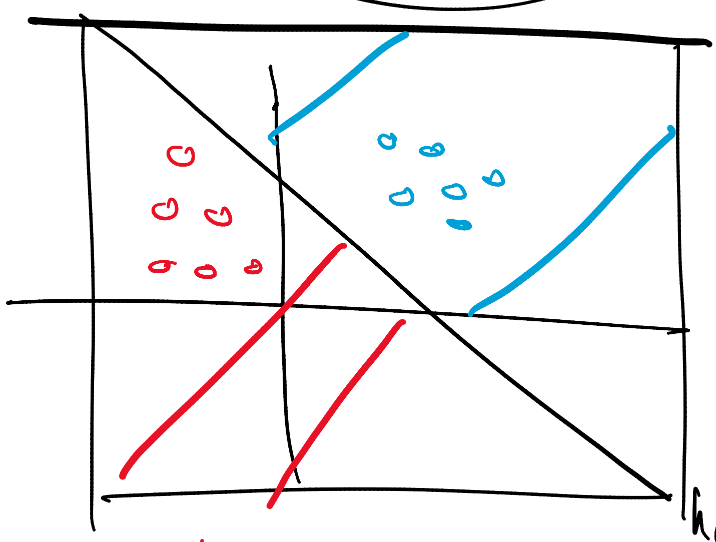
$$h(\vec{x}) = 0$$

$$\Rightarrow \vec{w}^T\vec{x} + b = 0$$

$$= w_1x_1 + w_2x_2 + \cdots + w_dx_d + b = 0$$

$$\Rightarrow \boxed{w_1x_1 + w_2x_2 + \cdots + w_dx_d = -b}$$

$$\vec{w}^T\vec{x} = -b$$



$h(\vec{x})$ = line = 1d hyperplane

$h(\vec{x}) = 0$

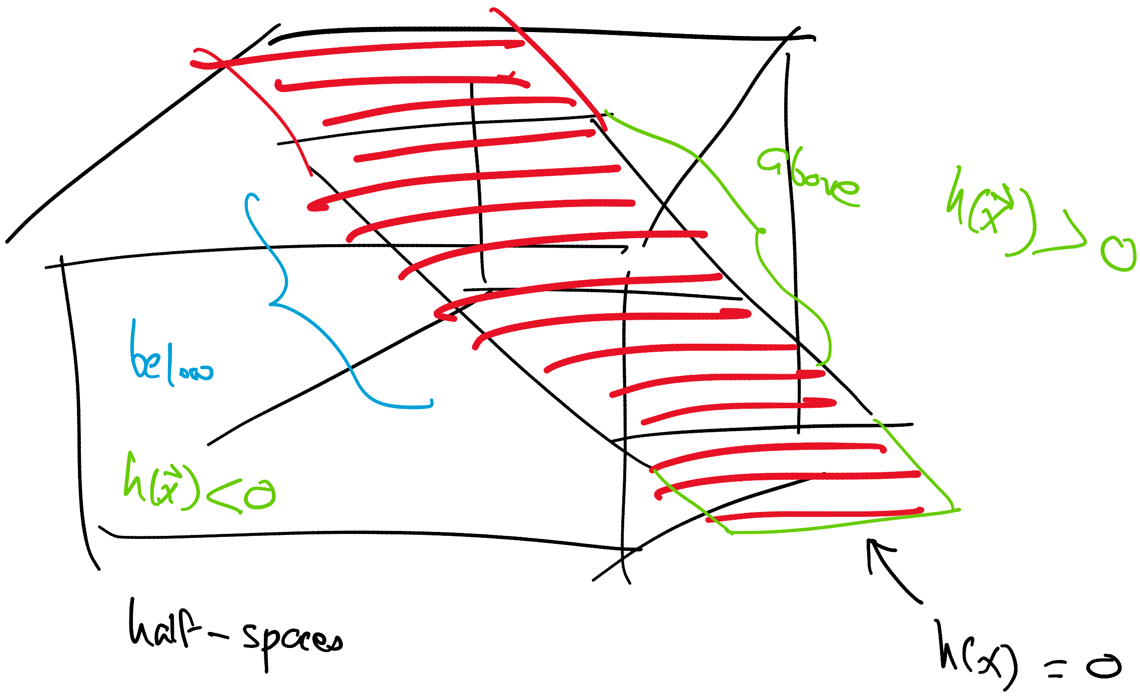
$\underbrace{\text{binary Classifier}}_{\text{Model}} \Rightarrow$

$$\underbrace{h(\vec{z})}_{\vec{w}^T\vec{z}+b} = \begin{cases} 1 & \text{if } h(x) > 0 \\ -1 & \text{if } h(x) < 0 \\ 0 & \text{if } h(x) = 0 \end{cases}$$

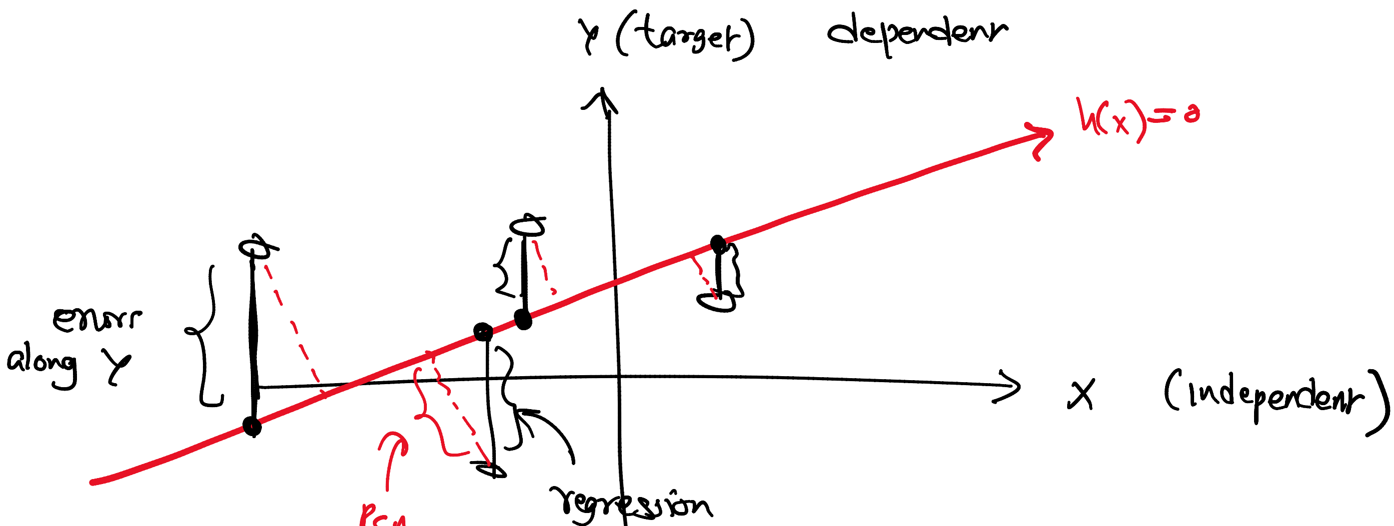
No class or tie-breaking

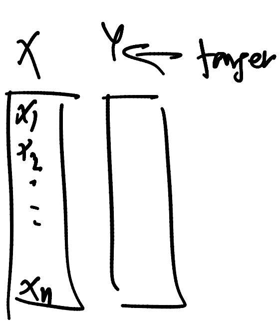
$C_1 = +1$ (red)

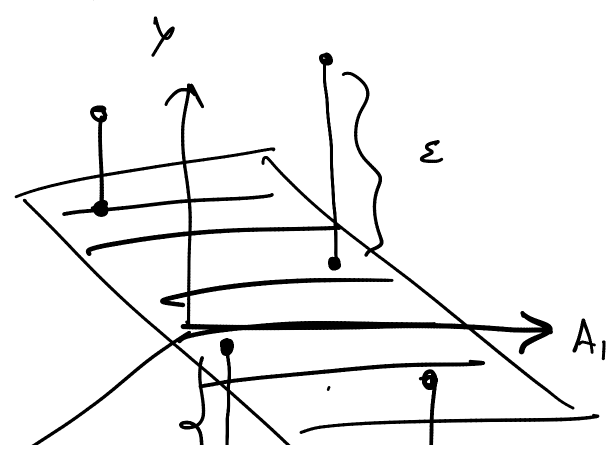
$C_2 = -1$ (blue)

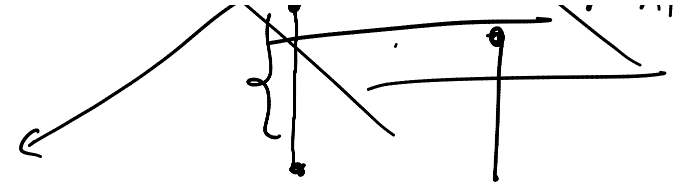


hyperplanes ≡ classifiers ≡ discriminator









d_2

Wierdness of high dimensions

Unit hypersphere $\quad S_d(1)$
$r = 1$

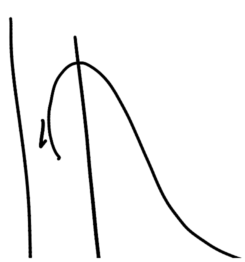
$$\mathrm{Vol}(S_d(1)) = \left(\frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}\right) \overset{1}{\cancel{r}}{}^{d}$$

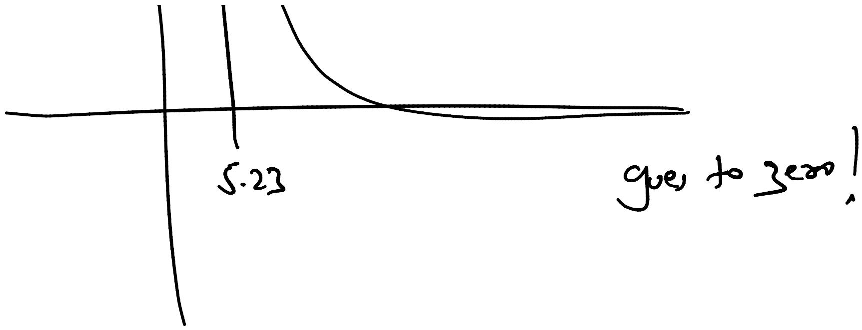
if $\quad d$ is even

$$\mathrm{vol}(S_d(1)) = \frac{\pi^{d/2}}{\left(d/2\right)!} \approx \frac{(\pi)^{d/2}}{(d/2)^{d/2}}$$

$$\lim_{d \to \infty} \left(\frac{\pi}{d/2}\right)^{d/2} = \underbrace{\left(\frac{2\pi}{d}\right)}_{}{}^{d/2}$$

$$(<1)^{d/2} \longrightarrow 0$$

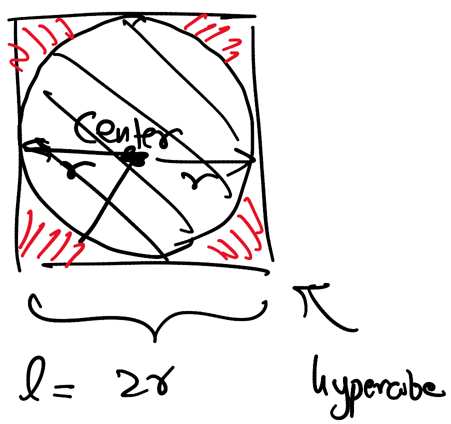




Unit hypersphere decay to zero!

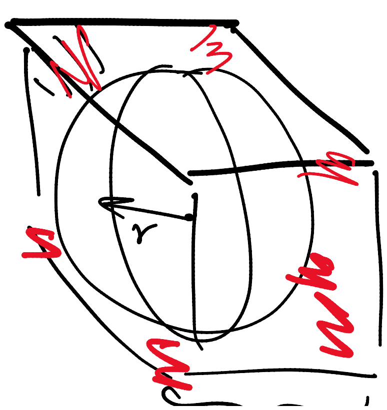
↳ you cannot sample a point within the hypersphere!

'Corners' ← all volume is in the corners



inscribed hypersphere within the hypercube

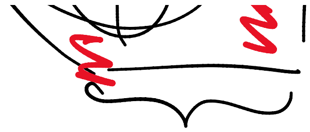
2D Sphere 78.5% Corner = 21.5%



3D Sphere 52.4% Corner = 47.6%

↓

$d \to \infty$



$l = 2r$

$d \to \infty$

sphere 0%. Corner = 100%.

$$\lim_{d\to\infty} \frac{Vol(S_d(r))}{Vol(H_d(2r))} = \frac{\pi^{d/2}/\Gamma(\frac{d}{2}+1)\ \cancel{r^d}}{2^d\cancel{r^d}} \longleftarrow (2r)^d$$

$$\lim_{d\to\infty} = \frac{\pi^{d/2}}{2^d \cdot \Gamma(\frac{d}{2}+1)} \longrightarrow 0$$