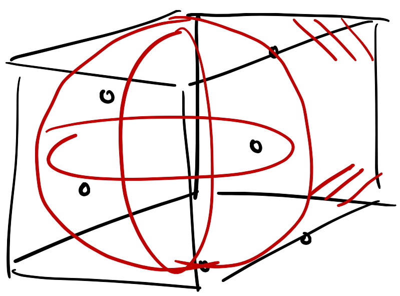


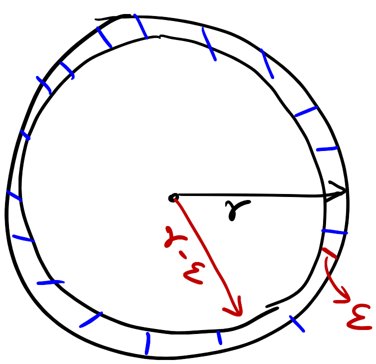
$\mathbb{R}^d$

there will be no points in the largest inscribed hypersphere

all points are in 2^d <u>corners</u>

---

all points are on the <u>surface</u>



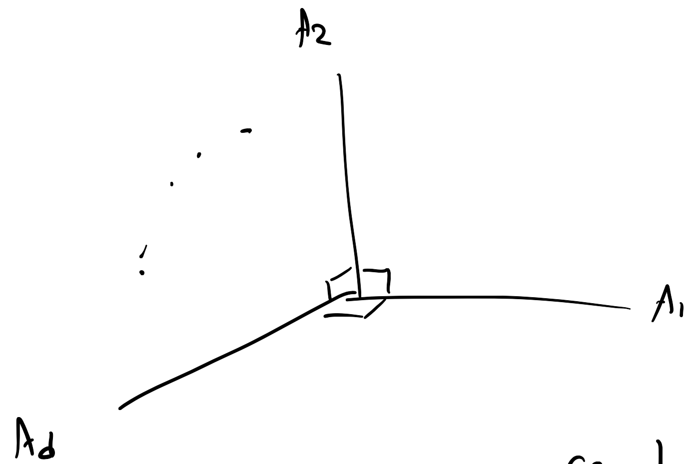
$d \to \infty$

$$\frac{\text{vol in the shell}}{\text{vol of outer hypersphere}} = 1$$

$\varepsilon > 0$

$$\frac{\text{vol}\left(S_d(r)\right) - \text{vol}\left(S_d(r-\varepsilon)\right)}{\text{vol}\left(S_d(r)\right)}$$

$$\lim_{d\to\infty} \frac{\cancel{K_d}\, r^d - \cancel{K_d}\,(r-\varepsilon)^d}{\cancel{K_d}\, r^d} = 1 - \left(\frac{r-\varepsilon}{r}\right)^d = 1 - \left(\overset{0}{\cancel{1-\frac{\varepsilon}{r}}}\right)^d$$

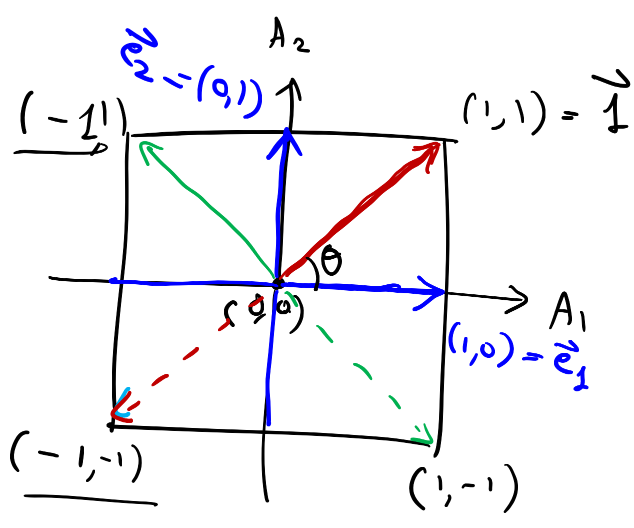


$d \leftarrow$ original orthogonal axes

as $d \to \infty$

there are additional 2^{d-1} "orthogonal" axes!

2d:



Standard basis $\{\vec{e}_1, \vec{e}_2\}$ for $\mathbb{R}^2$

2^d corners

$\frac{2^d}{2}$ axe $= 2^{d-1}$

$\cos\theta$ between $\vec{1}, \vec{e}_1 = 0$

$\Rightarrow \theta = 90^\circ$ as $\underline{\underline{d \to \infty}}$

$\mathbb{R}^d$: $\vec{1} = (\underbrace{1 \;\; 1 \;\; 1 \;\; \cdots \;\; 1}_{d})$ $\quad \vec{e}_1 = (\underbrace{1, 0, 0, \ldots, 0}_{d})$

$$\|\vec{1}\|^2 = d \qquad \|\vec{e}_1\|^2 = 1$$

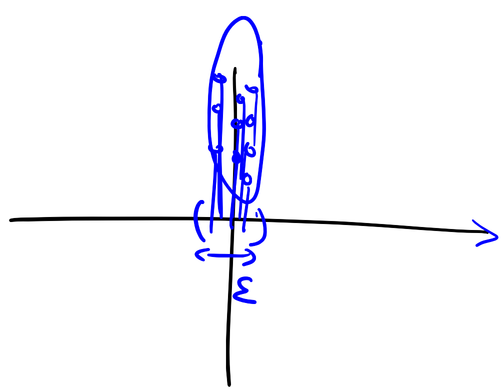
$$\|\vec{1}\| = \sqrt{d}$$

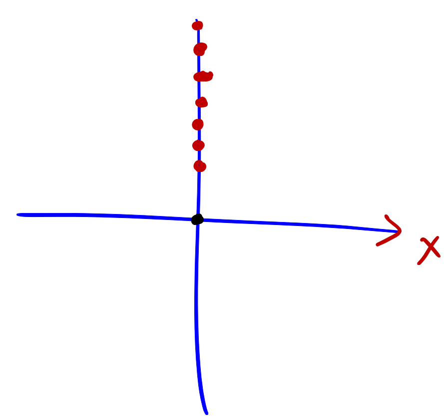
$$\lim_{d\to\infty} \boxed{\cos\theta = \left(\frac{\vec{1}}{\|\vec{1}\|}\right)^T \left(\frac{\vec{e}_1}{\|\vec{e}_1\|}\right) = \frac{\vec{1}^T \vec{e}_1}{\sqrt{d}\cdot 1} = \frac{1}{\sqrt{d}}} = 0$$

$$d + \underbrace{2^{d-1}}_{\substack{\text{w.r.t} \\ \text{original} \\ d \text{ axes}}} + \underbrace{\cdots}_{} +$$

higher order

i.e.





Normal Distribution

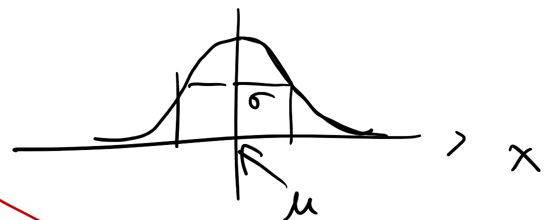
↳ parametric density function

$\vec{\mu}$ — sample mean

Σ — sample cov matrix

1d:

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



Multivariate : $\mathbb{R}^d$

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

$$|\Sigma| = \det(\Sigma) = \underbrace{\prod_{i=1}^{d} \lambda_i}_{\text{generalized variance}} \quad (\text{product of eigenvalues})$$

vs

$$\sum_{i=1}^{d} \lambda_i = \text{totvar}(D) = \sum_{i=1}^{d} \sigma_i^2$$

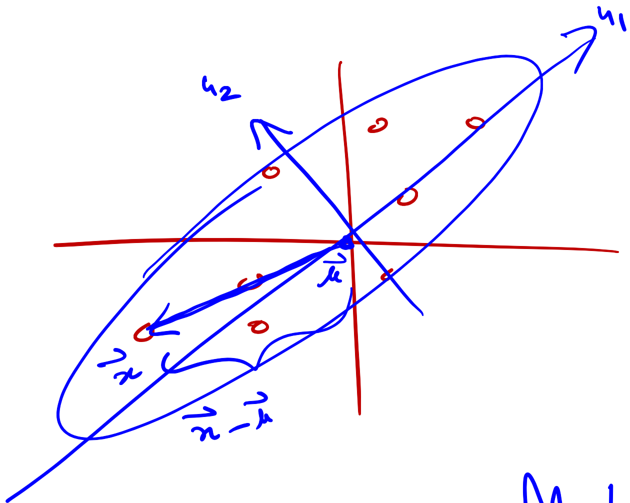
$\Sigma^{-1} \equiv$ precision matrix

Inverse of cov Σ

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

$$\forall k \quad \Sigma^k = U \Lambda^k U^T$$

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



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

Euclidean Distance

Mahalanobis Distance

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

takes variance & covariance into account

$$= \vec{\bar{x}}^T \Sigma^{-1} \vec{\bar{x}}$$ centered data point

$$= \vec{z}^T \Sigma^{-1} \vec{z} \qquad \vec{z} = \vec{x} - \vec{\mu}$$

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

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

$$U = \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_d \\ | & | & & | \end{bmatrix}$$

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

$$= \vec{a}^T \Lambda^{-1} \vec{a}$$

$$U^T \vec{z} = \left(\vec{u}_1^T z \;\; \vec{u}_2^T z \;\; \ldots \;\; \vec{u}_d^T z\right)^T$$

$$= \left(a_1 \;\; a_2 \;\; \ldots \;\; a_d\right)^T$$

$$= \vec{a}$$

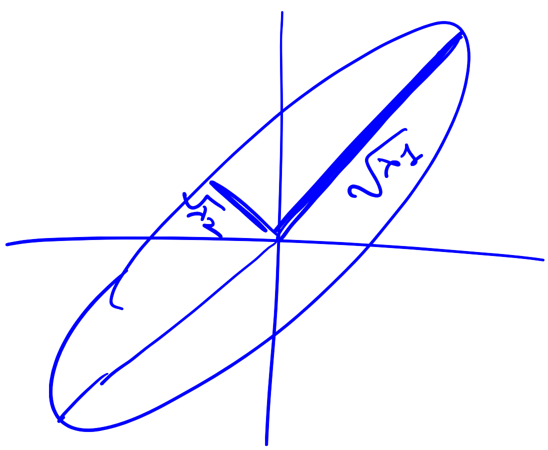
projected point in new basis

$$= (a_1, a_2 \ldots a_d) \begin{bmatrix} 1/\lambda_1 & & & \\ & 1/\lambda_2 & & \\ & & \ddots & \\ & & & 1/\lambda_d \end{bmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{pmatrix}$$

$$= \sum_{i=1}^{d} \frac{a_i^2}{\lambda_i} = 1$$

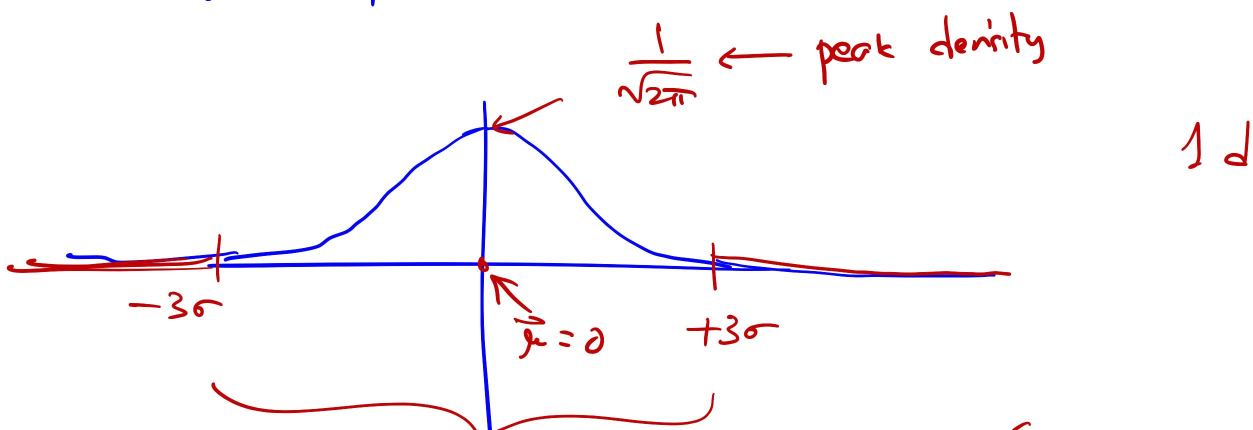
scalar

equation of hyperellipsoid centred at $\vec{0}$ with $\sqrt{\lambda_i}$ as the length of the axes



$d \to \infty$, what happens for Multivariate Normal

all of the probability mass migrate to the tails!



99.9% of the points lie in $(\mu - 3\sigma, \mu + 3\sigma)$

$\vec{\mu} = 0 \qquad \Sigma = I$

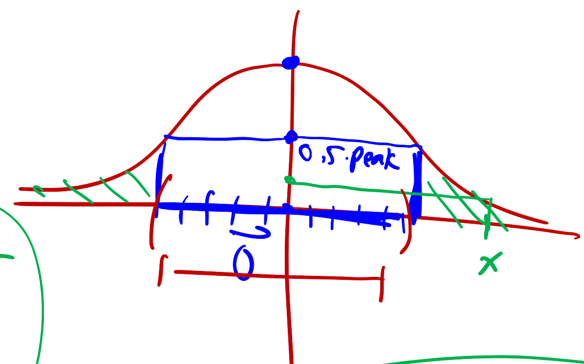
$\underbrace{\vec{\mu} = 0 \qquad \Sigma = I}_{\text{Standard Multivariate Normal}}$

$$\text{peak density} = \underset{\substack{\uparrow \\ \vec{\mu}}}{f(\vec{0})} = \frac{1}{(\sqrt{2\pi})^d} \cdot \frac{1}{\sqrt{|\Sigma|}}$$

$$= \frac{1}{(\sqrt{2\pi})^d}$$

Q: what is the probability of finding a point $\vec{x}$ within α fraction of the peak density

$$\frac{f(\vec{x})}{f(\vec{0})} \geq \alpha = 0.5$$



$\Sigma = I, \quad \vec{\mu} = 0 \quad \Sigma^{-1} = I$

$$\frac{f(\vec{x})}{f(\vec{0})} = \frac{\cancel{\frac{1}{(\sqrt{2\pi})^d}} \frac{1}{\cancel{\sqrt{|\Sigma|}}^{\,1}} \, e^{-\frac{\|\vec{x}\|^2}{2}}}{\cancel{\frac{1}{(\sqrt{2\pi})^d}}}$$

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

$$(\vec{x} - \vec{0})^T I (\vec{x} - \vec{0})$$

$$\vec{x}^T \vec{x} = \|\vec{x}\|^2$$

$$\frac{f(\vec{x})}{f(\vec{0})} = \boxed{e^{-\frac{\|\vec{x}\|^2}{2}} \geq \alpha}$$

$$\|\vec{x}\|^2 \leq -2\ln\alpha$$

$$p \equiv P\left(\boxed{\|\vec{x}\|^2 \leq -2\ln\alpha}\right) \qquad 0 < \alpha < 1$$

to find this : Monte carlo simulations

$$\|\vec{x}\|^2 = \underbrace{\sum_{i=1}^{d} x_i^2}$$

has a χ^2 distribution with d dof (degrees of freedom)

chi-squared

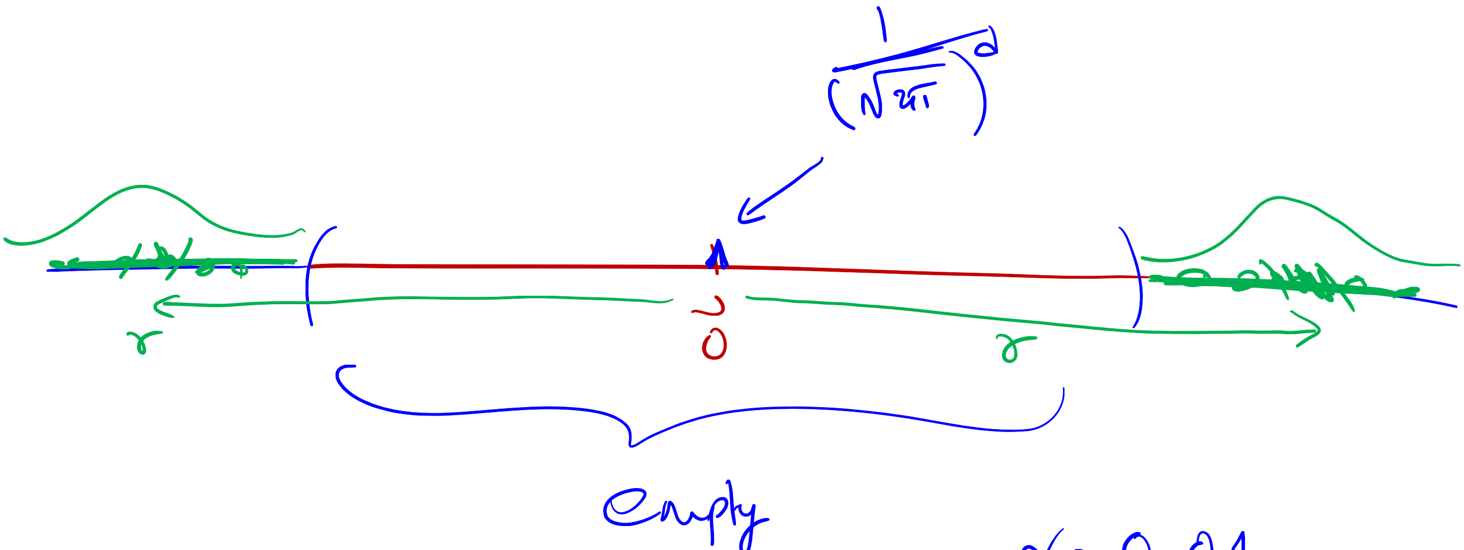
$d = 1$, $P = 76.1\%$ $\alpha = 0.5$

$d = 2$ $= 50\%$

$d = 3$ $= 29\%$

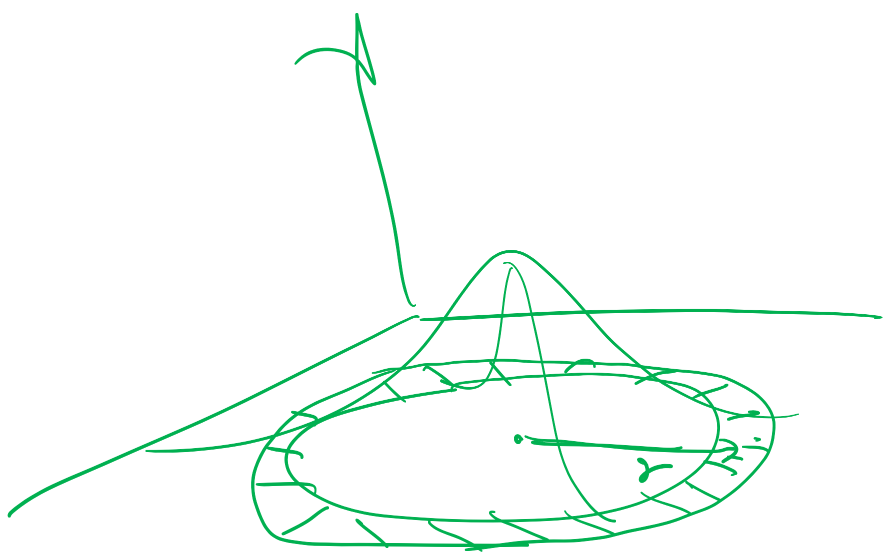
$= \vdots$

$$\lim_{d \to \infty} P = 0$$



$\alpha = 0.01$

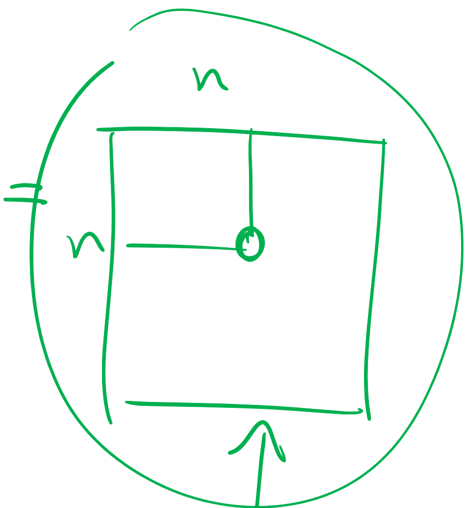
$P = 0$



D ($n \times d$) — data matrix

$\Sigma = d \times d$ — Pairwise dims

$K =$

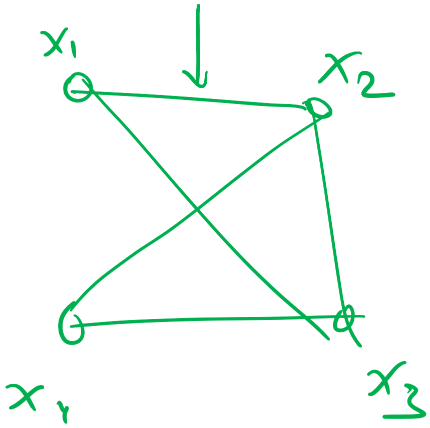


Complete graph of pairwise similarities

$$K = \{ k(\vec{x}_i, \vec{x}_j) \}$$

K: pairwise similarity matrix

k is some kernel function



K to be PSD