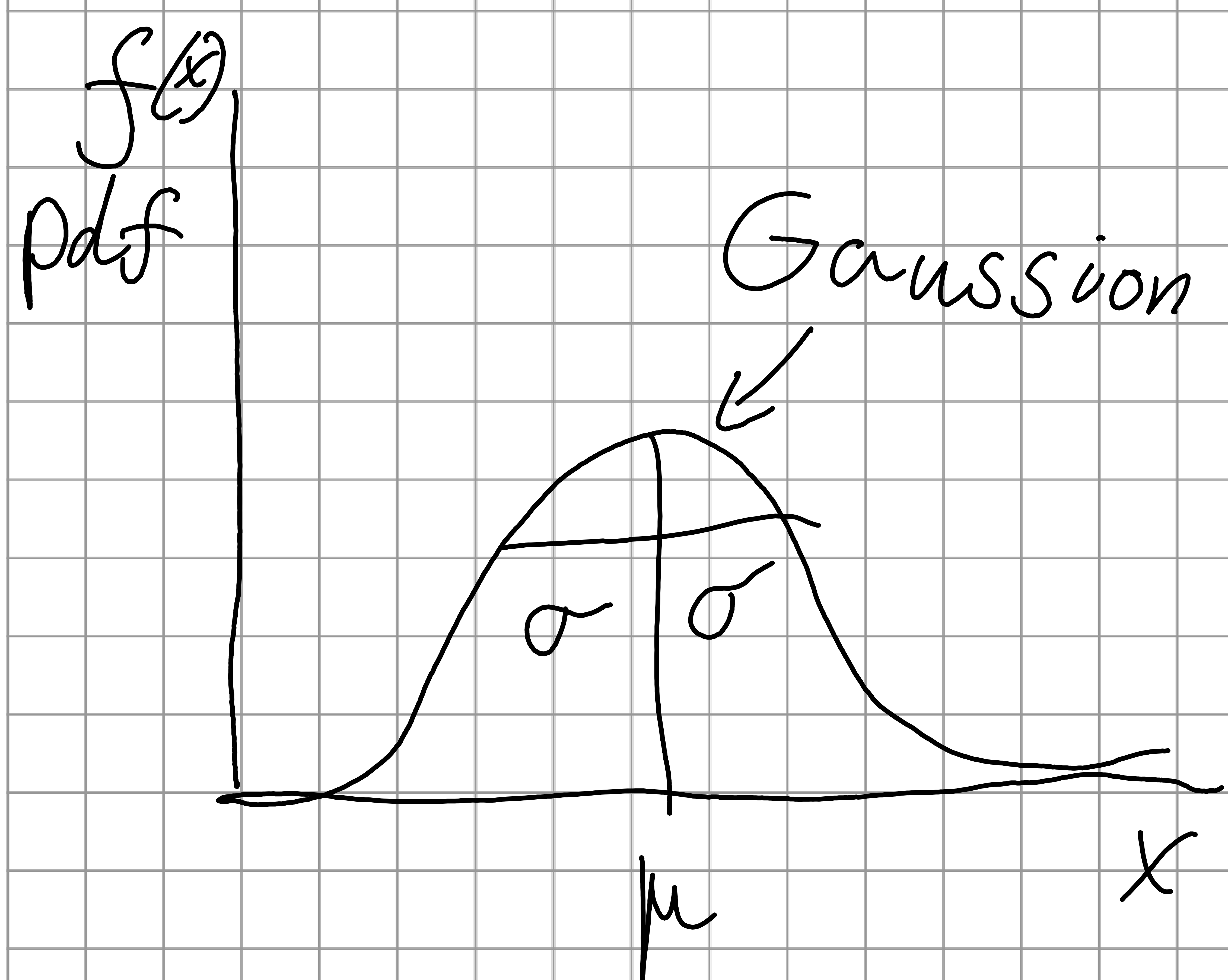
MVN

Multivariate Normal Distribution

1D $X \sim N(\mu, \sigma^2)$



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

$E[X] = \mu$

$V[X]\,\sigma^2$

Standard normal

$Z \sim N(0,1)$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad z \in \mathbb{R}$$

normal pdf

pD

$$\underset{(p\times 1)}{X} = [X_1\ X_2\ \ldots\ X_p]^T$$

1D exponent $-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 =$

$= -\frac{1}{2}(x-\mu)^T \frac{1}{\sigma}(x-\mu)$

pD —"— $-\frac{1}{2}\underbrace{(x-\mu)^T}_{(1\times p)} \underset{(p\times p)}{\Sigma^{-1}} \underset{(p\times 1)}{(x-\mu)}$

pdf: $f(\underset{(p\times 1)}{x}) =$

$= 2\pi^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}, x \in \mathbb{R}^p$

Notation $\underset{(p\times 1)}{X} \sim N_p(\underset{(p\times 1)}{\mu}\ \underset{(p\times p)}{\Sigma})$

$E[X] = \mu$

$\mathrm{cov}[X] = \Sigma$

Constant Density Contours

$$(x-\mu)^T \Sigma^{-1} (x-\mu) = \text{constant} = C$$

$x^T A x$:

- parabolas
- hyperbolas — dependent on A
- ellipses

Σ positive semidefinite

$\Updownarrow$

Symmetric

$\Updownarrow$

$\forall \lambda_i \geq 0$

all

(Hyper) ellipsoids

$p=2 \rightarrow$ ellipses
$p=3 \rightarrow$ ellipsoid
$p>3$ (hyper) ellipsoid

Centered at $x=\mu$
p dimensions

axis directions - determined by eigenvectors of Σ, e_i, Σ
$i=1,\dots,p$

half axis length -"- eigenvalues of Σ $\sqrt{c}\sqrt{\lambda_i}$, $i=1\dots T$

$$EVD = \Sigma = P \cdot \Lambda \cdot P^T$$

$$\begin{bmatrix} \uparrow & \uparrow \\ e_i & e_p \\ \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_p \end{bmatrix}$$

$$\underset{p \times p}{\Sigma} = \sum_{i=1}^{p} \lambda_i \cdot e_i \cdot e_i^T$$

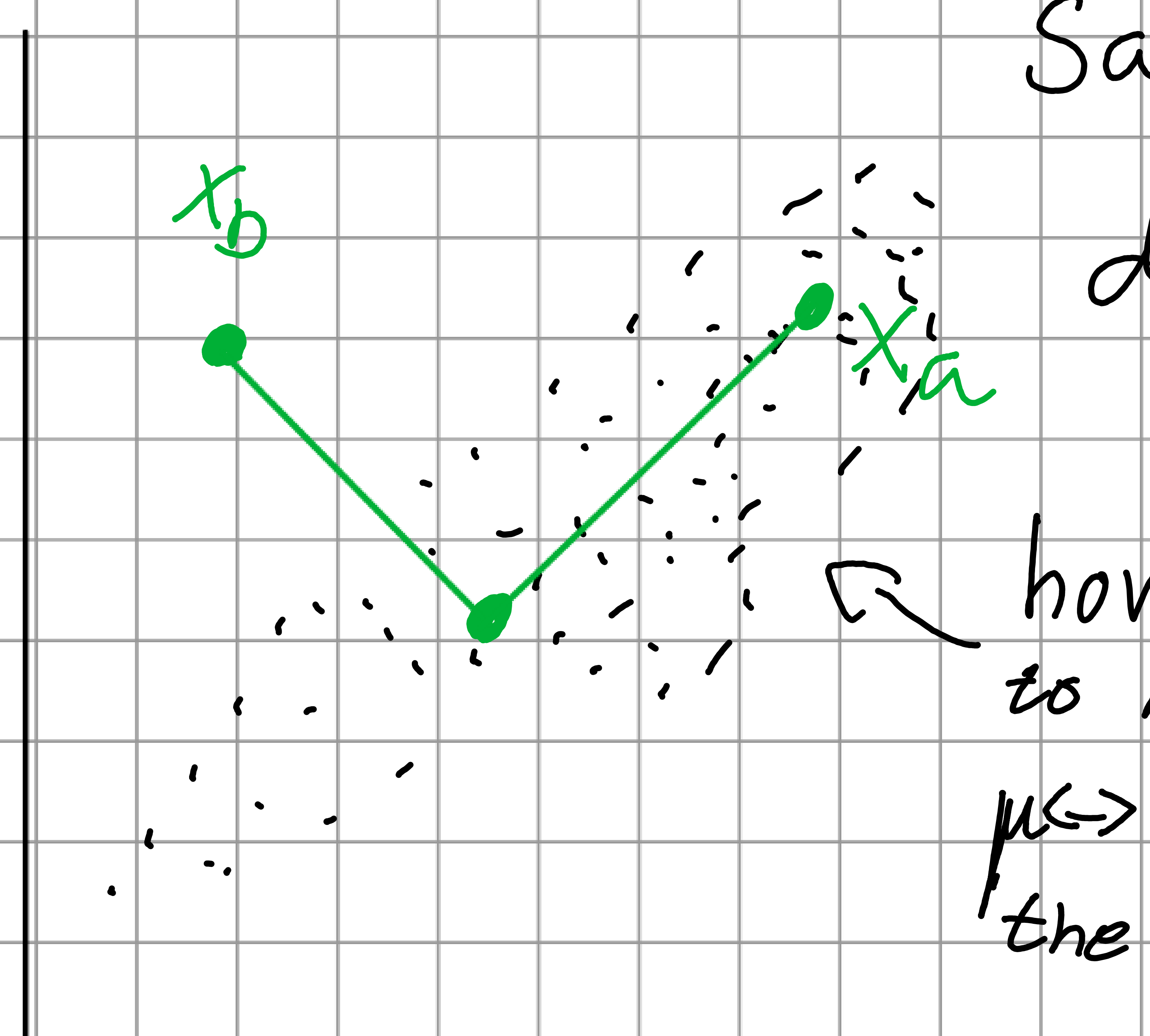
Statistical (Mahalonobis) Distance

$$d_m^2(x, \mu) = (x-\mu)^T \Sigma^{-1} (x-\mu)$$

squared distance between x, μ

$$|x_a - \mu| = |x_b - \mu|$$

Same euclidian distance



how likely are we to make an observation. $\mu \leftrightarrow x_a$ is more likely, as the variance is greater.

$$d_m^2(x_b - \mu) > d^2(x_a - \mu)$$

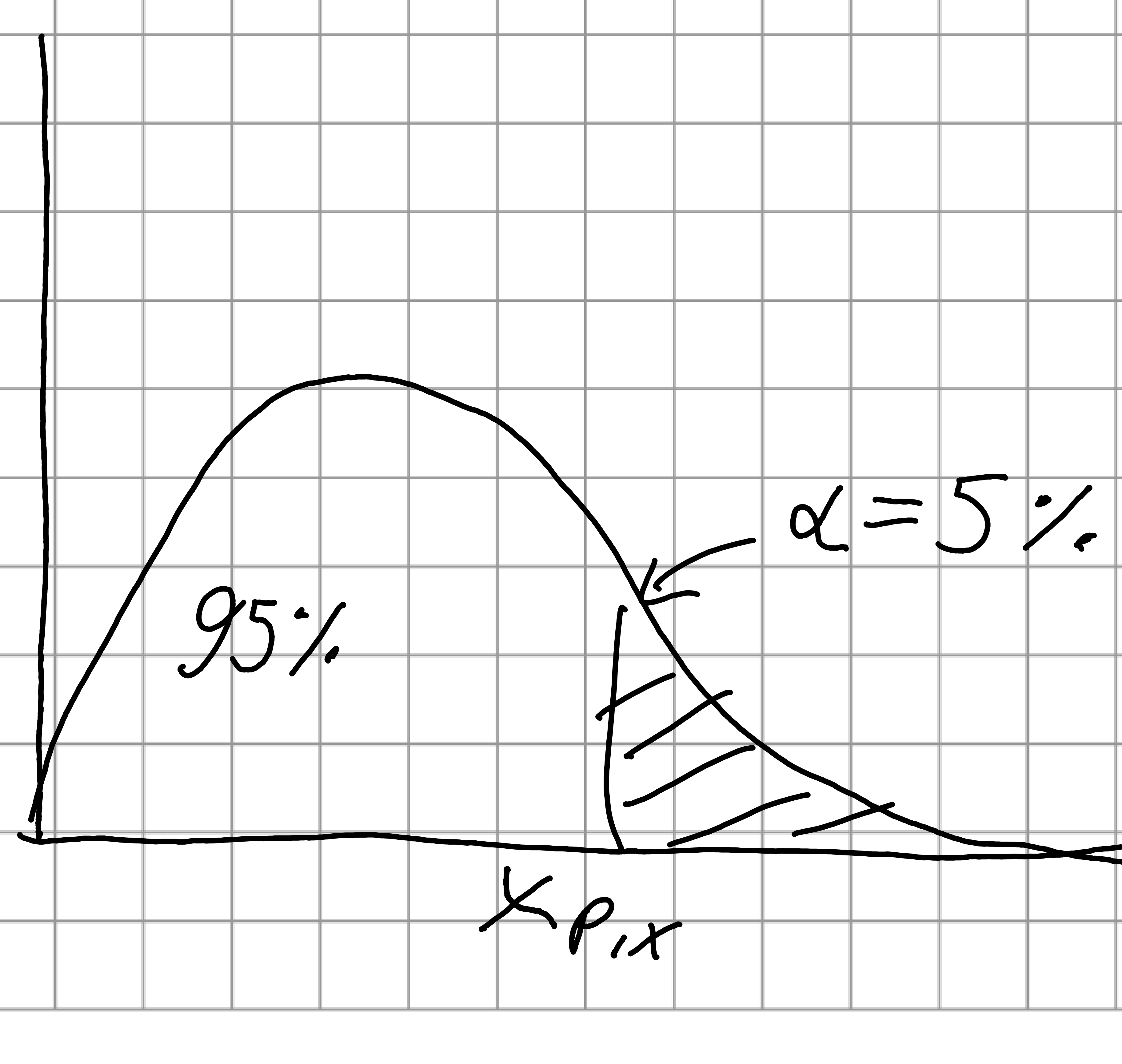
$$\boxed{\begin{aligned} &X \sim N_p(\mu, \Sigma) \\ &\downarrow \\ &d_m^2(x,\mu) \sim \chi_p^2 \end{aligned}}$$

100(1-α)% Confidence Region of observation.

$$d_m^2(x,\mu) = (x-\mu)^T \Sigma^{-1} (x-\mu) \le \chi^2_{p,\alpha}$$

Matlab: chi2inv

χ_p^2



Standard MVN

$$X \sim N_p(\mu, \Sigma)$$

matlab "eye"

$$\underset{p\times 1}{Z} \overset{\text{def}}{=} \Sigma^{-\frac{1}{2}}(X-\mu) \sim N_p(0, I)$$

$$Z \sim N_p(0, I)$$

$$X = \Sigma^{\frac{1}{2}} \cdot Z + \mu \sim N_p(\mu, \Sigma)$$

= distributed as

Central Limit Theorem

$X_1, X_2 \ldots X_n$ iid

independent

identically

distributed

$X_i \in \mathbb{R}$

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_i \overset{\text{approx}}{\sim} N_p\left(\mu, \frac{\Sigma}{n}\right)$$

with: $E[X] = \mu$

$\mathrm{cov}[X] = \Sigma$

Estimation of MVN parameters

μ, Σ

obs $X_1, \dots, X_n$ iid $X_j \sim N_p(\mu, \Sigma)$

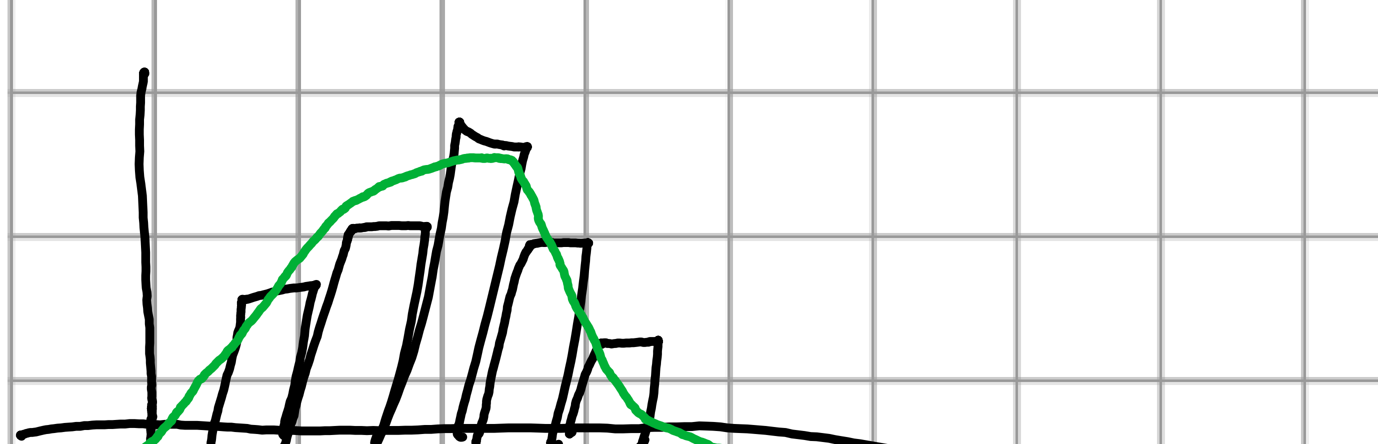
Maximum likelyhood estimate

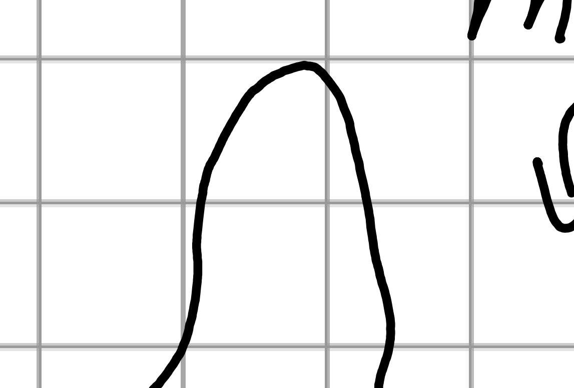
Likelihood function (pdf value)

we try to fit a pdf to our dataset. We try alot to see which one fits best.

pdf

$$L(\mu, \Sigma) \overset{def}{=} \text{"} p(\text{given obs} \mid \mu, \Sigma) \text{"}$$





maximize $L(\mu, \Sigma) \rightarrow (\hat{\mu}_{ML}, \hat{\Sigma}_{ML})$

$$\hat{\mu}_{ML} = \bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j \sim N_p\left(\mu, \frac{\Sigma}{n}\right)$$

$$E\left[\hat{\mu}_{ML}\right] = \mu \quad \text{unbiased}$$

expected value

if we do it infinitly many times

$$Cov(\hat{\mu}_{ML}) = \frac{\Sigma}{n} \rightarrow \text{consistent}$$

more samples means more consistent estimator $\rightarrow$ less uncertainty

$$\hat{\Sigma}_{ML} = \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X})(X_j - \bar{X})^T$$

$$E\left[\hat{\Sigma}_{ML}\right] = \frac{n-1}{n}\Sigma$$

$$S = \hat{\Sigma}$$