

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{\int P(D|\theta)P(\theta)d\theta}$$


is  $\hat{\theta}_{MLE} : \text{POINT}$

is  $\hat{\theta}_{MAP} : \text{POINT}$

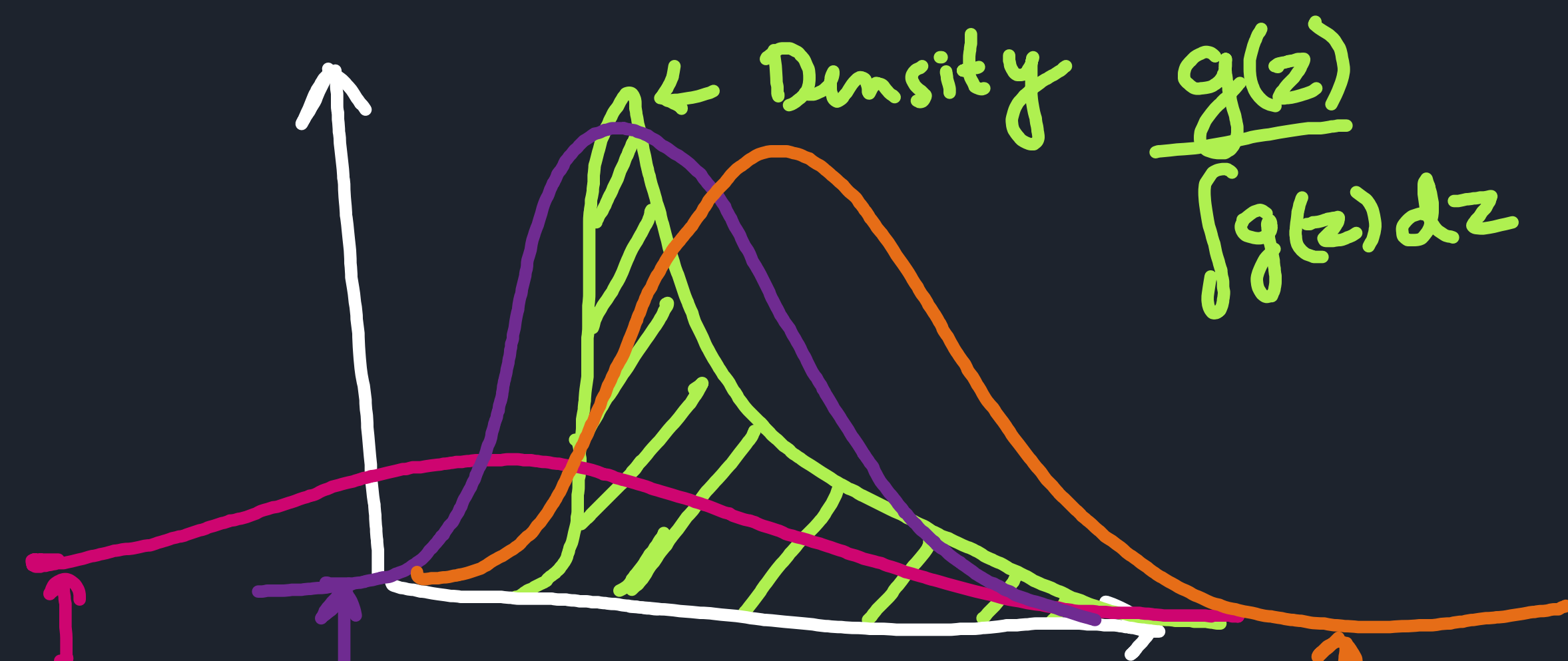
(ii)  $P(\theta|D) : D_1$  DISTRIBUTION:

DENOMINATOR  
FOR LIN REG.

$$\int P(D|\theta)P(\theta)d\theta$$



# Laplace Appx.



Density  $\frac{g(z)}{\int g(z) dz}$

Approximate  
density  
w/  
NORMAL

$N_1$

$N_2$   
(MODE  
MATCHES)

$N_3$   
• 'LOWER' TO TARGET  
VARIATIONAL  
APPX.



$$N\left(\begin{bmatrix} \mu_0 \\ \mu_1 \end{bmatrix}, \Sigma\right)$$

$$; \begin{bmatrix} \mu_0 \\ \mu_1 \end{bmatrix} = \text{MODE}(\text{TARGET})$$

# Taylor Series

Approx  $f(x)$ , e.g.  $\cos(x)$  around  $x=0$

$$g(x) = c_0 + c_1 x + c_2 x^2 + \dots$$

$$g(x) = f(0) + \frac{f'(x)|_{x=0}}{1!} x + \frac{f''(x)|_{x=0}}{2!} x^2 + \dots$$

Appx.  $f(x)$  around  $x=x_0$

$$g(x) = \frac{f(x_0) + f'(x)|_{x=x_0} \cdot (x-x_0)}{1!} + \frac{f''(x)|_{x=x_0} (x-x_0)^2}{2!} + \dots$$

(Q)  $f(x) = \cos(x)$  around  $x=\pi$

$$= \frac{\cos(\pi)}{1!} (x-\pi) - \frac{\sin(\pi)}{2!} (x-\pi)^2 + \dots$$

$$= -1 + \frac{1}{2} (x-\pi)^2 + \dots$$

Taylor Series for M.V. Input at  $x=x_0$

$$g(x) = f(x_0) + \frac{1}{1!} (x-x_0)^T \nabla f(x_0)$$

$$+ \frac{1}{2!} (x-x_0)^T \nabla^2 f(x_0) (x-x_0) + \dots$$

Q)  $f(x) = \cos x_1 + \cos x_2$

around  $x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\cos 0 + \cos 0 + \frac{1}{1!} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \frac{1}{2!} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \nabla^2 f \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\vec{\nabla} f \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial}{\partial x_1} (-\sin x_1) & \frac{\partial}{\partial x_1} (-\sin x_2) \\ \frac{\partial}{\partial x_2} (-\sin x_1) & \frac{\partial}{\partial x_2} (-\sin x_2) \end{bmatrix}$$

$$= \begin{bmatrix} -\cos x_1 & 0 \\ 0 & -\cos x_2 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 \\ -\cos x_1 & -\cos x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\therefore g(x) = 1 + 1 - \frac{1}{2} [x_1 \ x_2] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 2 - \frac{1}{2} [x_1 \ x_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 2 - \frac{1}{2} (x_1^2 + x_2^2)$$



$$\hat{\theta}_{MAP} = \underset{\theta}{\operatorname{argmax}} P(\theta|D)$$

$$P(\theta|D) = ?$$

$$= \frac{P(D|\theta)P(\theta)}{P(D)} = \frac{P(D,\theta)}{\int P(D,\theta)d\theta}$$

$$= \frac{e^{\log P(D,\theta)}}{\int e^{\log P(D,\theta)} d\theta}$$

$$\rightarrow f(\theta) = \log P(D,\theta)$$

$$\Theta_{MAP} = \underset{\theta}{\operatorname{argmax}} P(\theta|D) = \underset{\theta}{\operatorname{argmax}} P(D|\theta)P(\theta)$$

$\searrow$   
 $P(D|\theta)$

$$f(\theta) = \log P(D, \theta)$$

Approx. w/ polynomial around  $\theta = \theta_0$

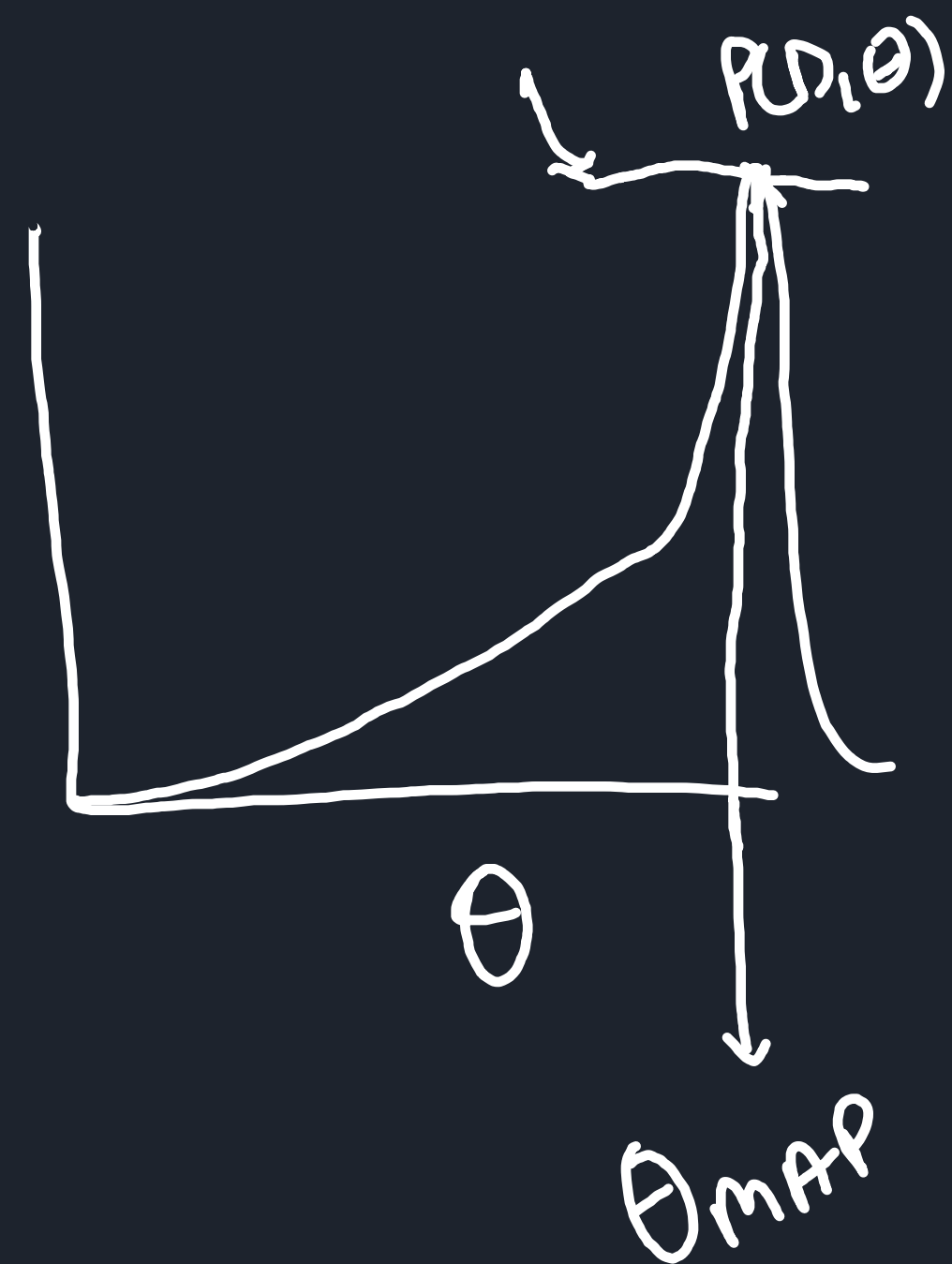
$$f(\theta) = \log(P, \theta_0) + 1(\theta - \theta_0)^T \nabla f|_{\theta_0} + \frac{1}{2}(\theta - \theta_0)^T \nabla^2 f|_{\theta_0} (\theta - \theta_0) + \dots$$

At  $\theta = \theta_{MAP}$

$$\nabla f|_{\theta_{MAP}} = 0$$

$\therefore f(\theta)$  around  $\theta = \theta_{MAP}$

$$= \log(P, \theta_0) - \frac{1}{2}(\theta - \theta_0)^T (-\nabla^2 f|_{\theta_{MAP}}) (\theta - \theta_0)$$



$$\nabla f|_{\theta = \theta_{MAP}} = 0$$

$$P(\theta|D) = \cancel{e^{\log(D, \theta)}} \cdot e^{\frac{1}{2}(\theta - \theta_{\text{MAP}})^T (-\nabla^2 \log(D, \theta)) (\theta - \theta_{\text{MAP}})}$$

$$\cancel{\int e^{\log(D, \theta)} \cdot e^{\frac{1}{2}(\theta - \theta_{\text{MAP}})^T (\dots \dots)} d\theta}$$

$$= N(\theta | \theta_{\text{MAP}}, \Sigma)$$

$\mu, \Sigma$

$$\Sigma^{-1} = -\nabla^2 \log(D, \theta) \Big|_{\theta_{\text{MAP}}}$$

Multi Variate Normal

$$e^{\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)}$$

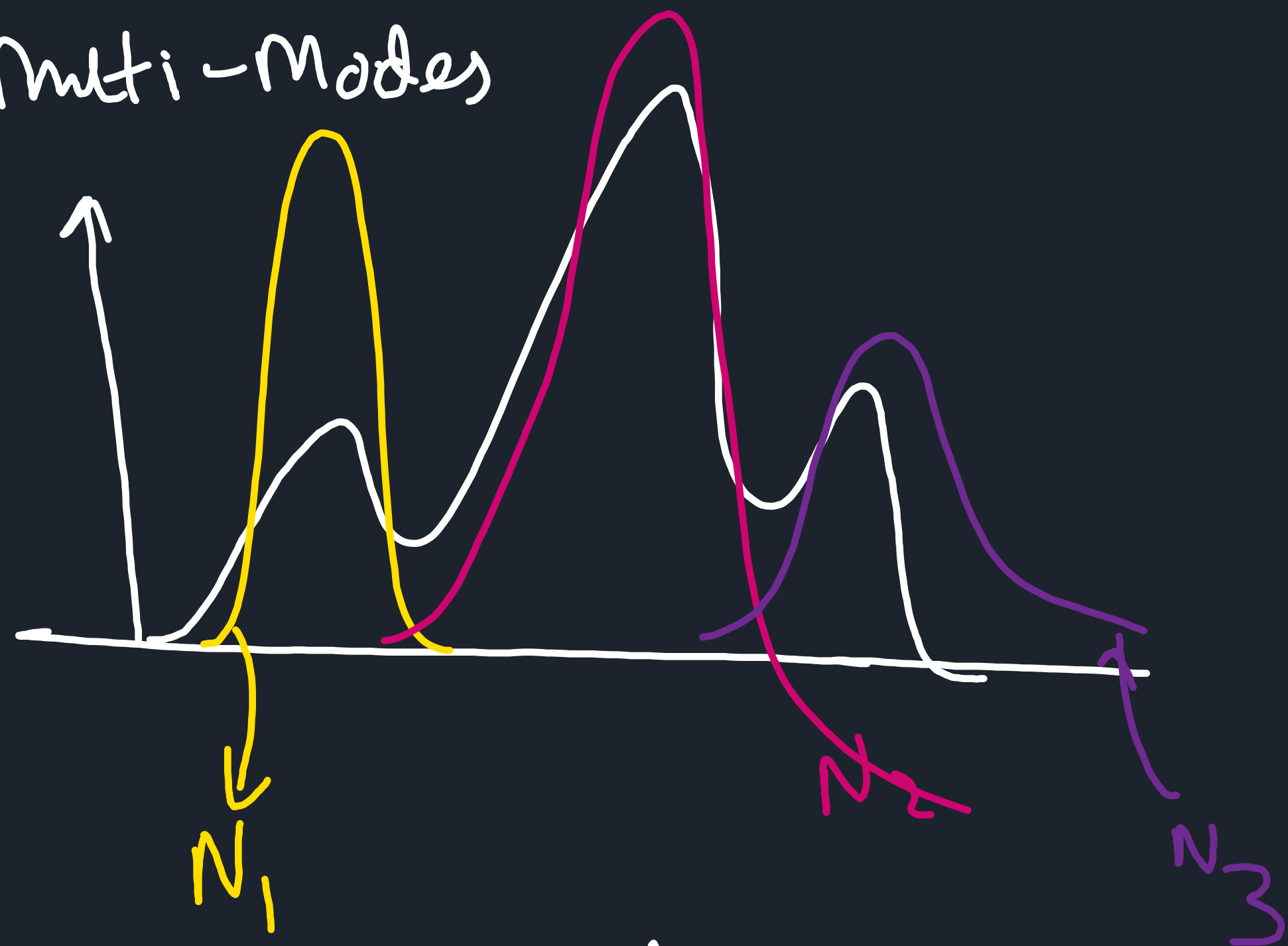
$$\int e^{\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)} d\mathbf{x}$$



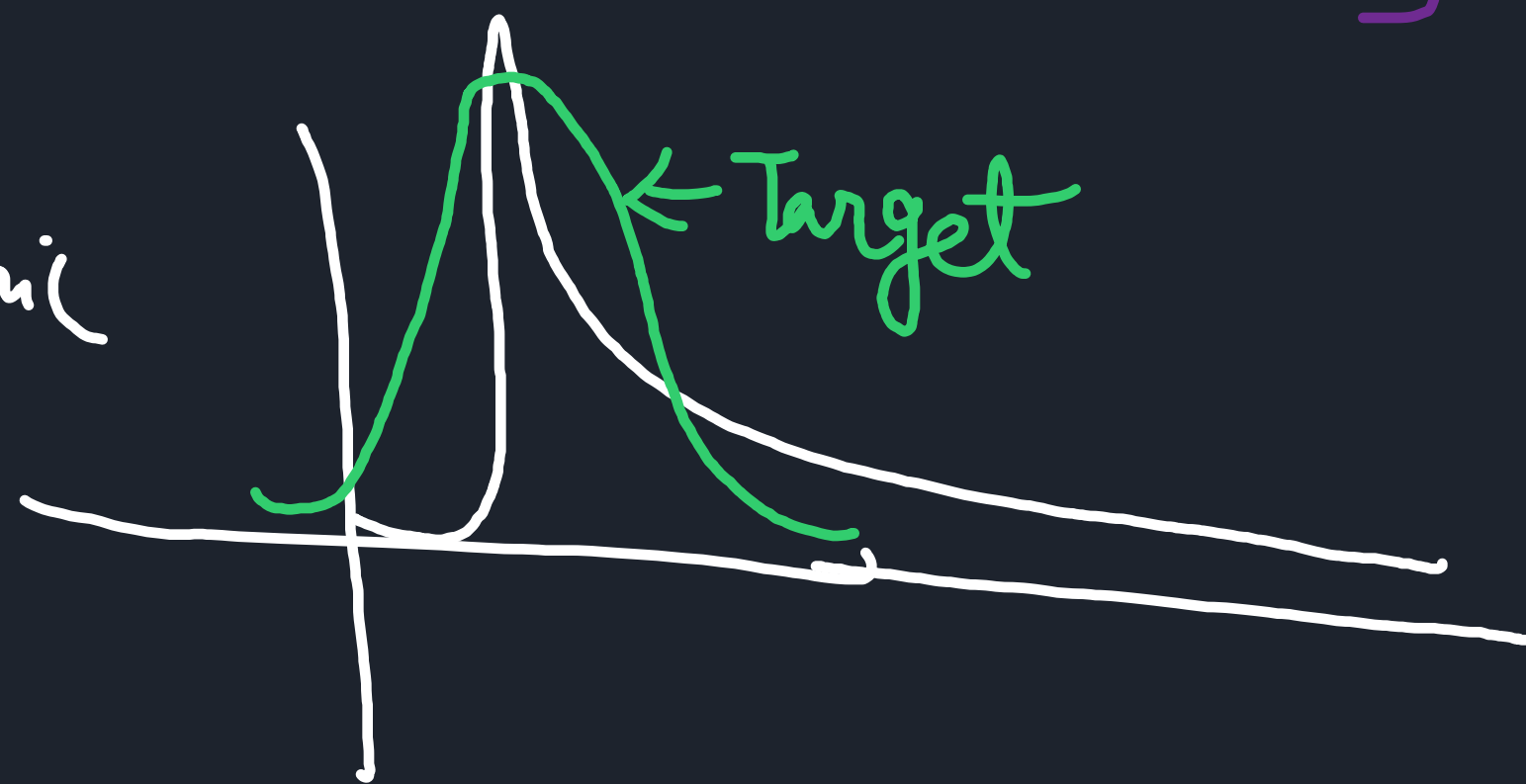
(i) Find  $\theta_{MAP}$

(ii) Find  $\Sigma$  of Appx. NORMAL

(i) Multi-Modes



(ii) Non-symmetric



(iii) Support Mismatch

