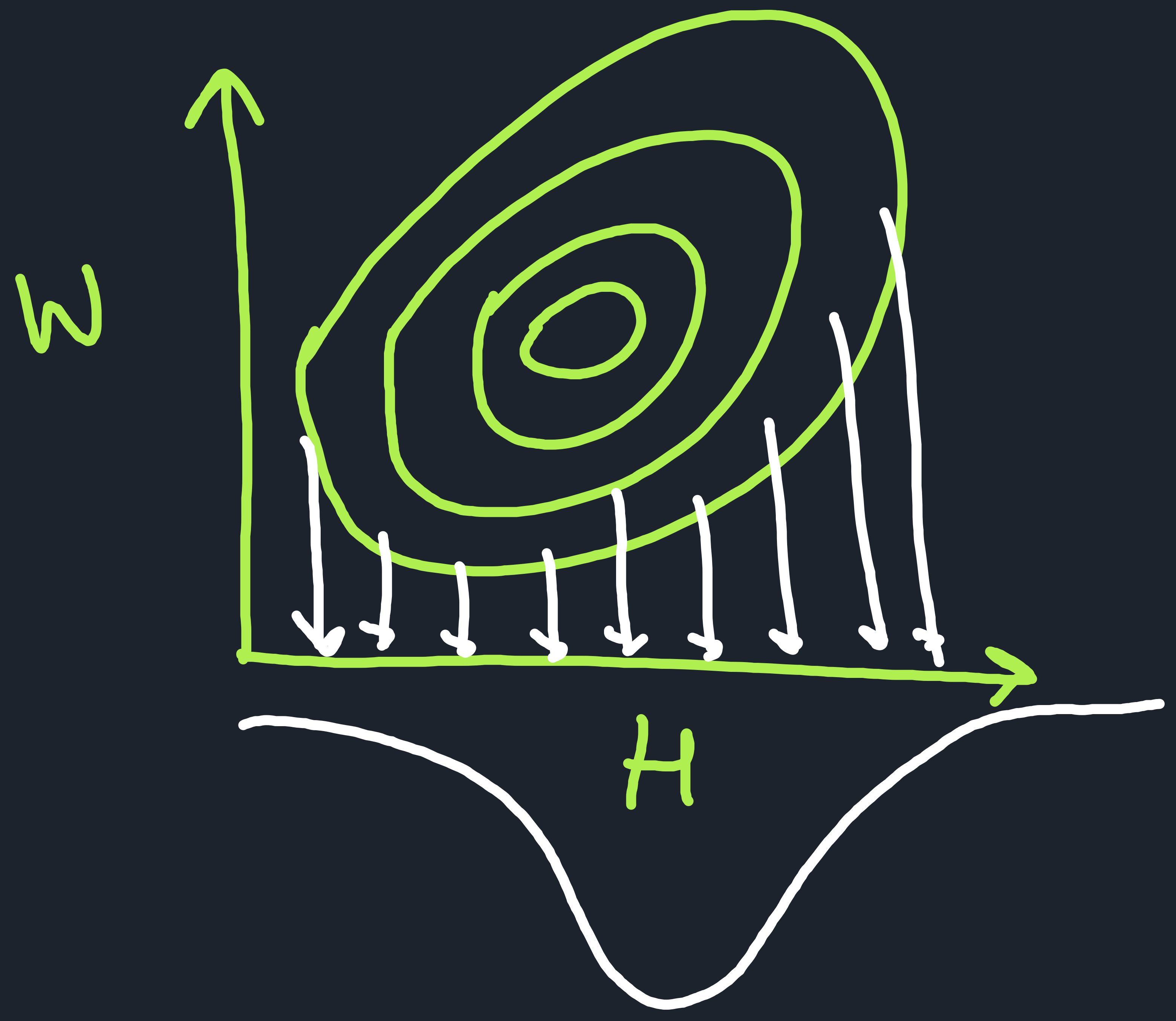
Marginalization



$$P(H,W) = N_2(\mu, \Sigma)$$

$$P(H) = ?$$

$$= \int_{W_L}^{W_H} P(H,W)\,dW = \int_{-\infty}^{\infty} P(H,W)\,dW$$

Prior Predictive Distribution

$P(\theta)$: Prior ✓

D : Dataset

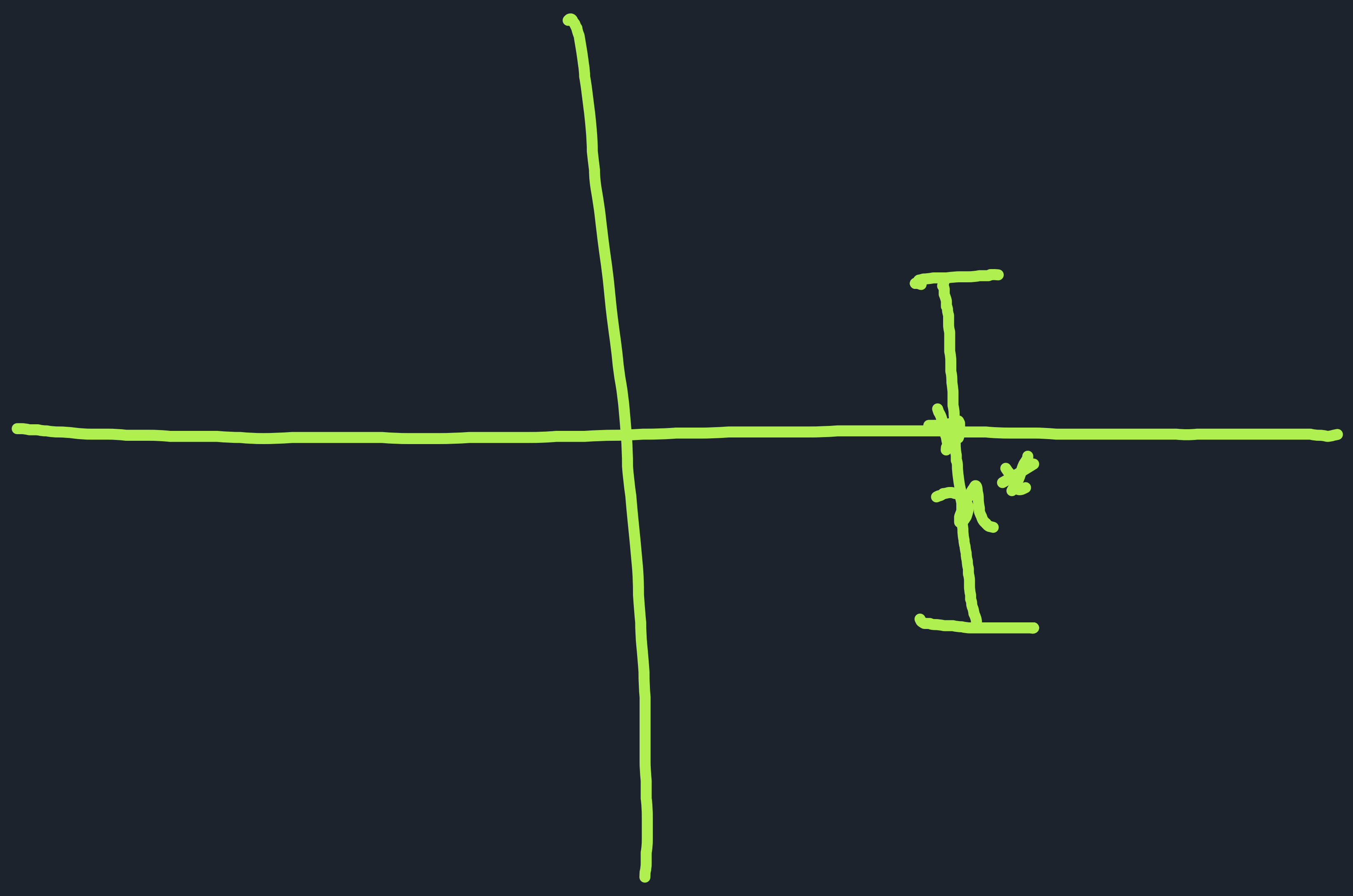
$P(\theta|D)$: Posterior

$P(D|\theta)$: Likelihood

$$\boxed{P(y^* \mid x^*)}$$

Test i/p

o/p



Prior Predictive

$P(y^*|x^*)$

Posterior Predictive

$P(y^*|x^*, D)$

$$P(y^*|x^*) = \int \underbrace{P(y^*, \theta | x^*)}_{\text{Apply Bayes here}} d\theta$$

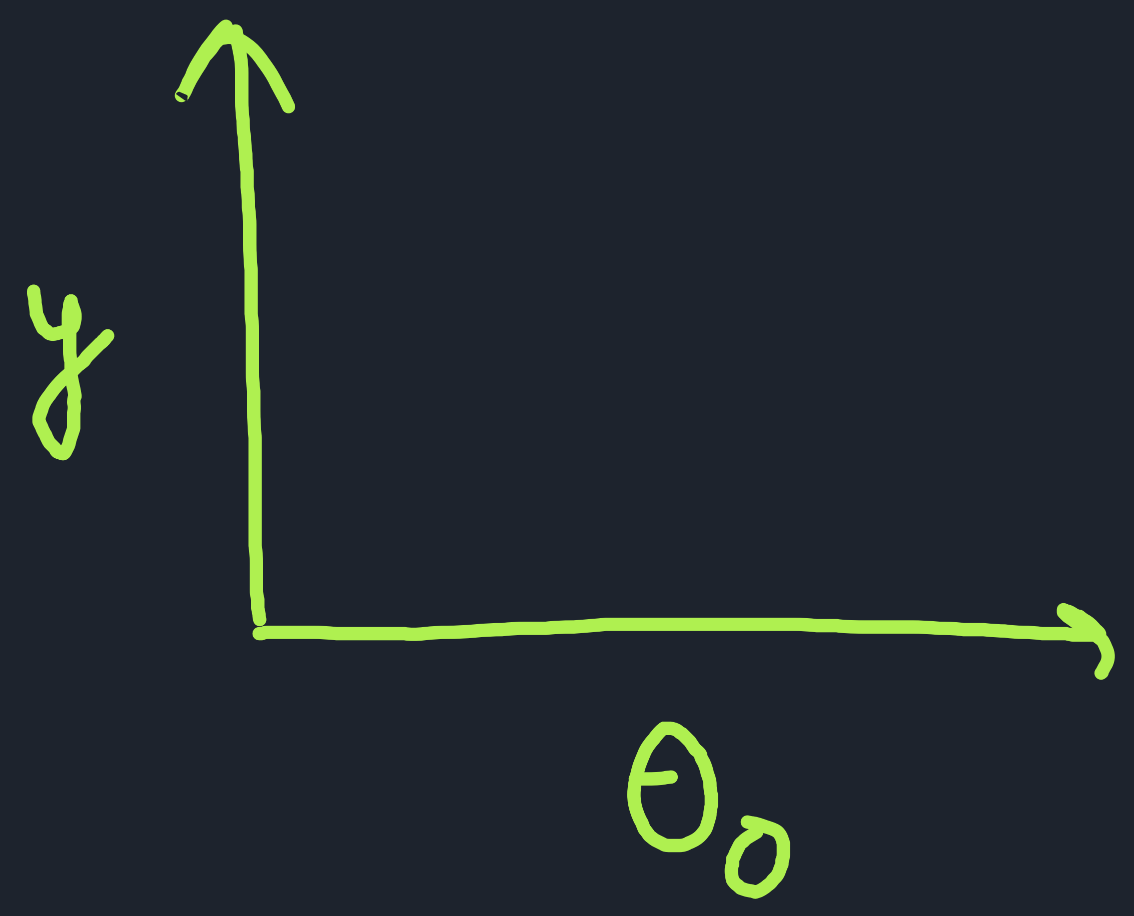
$P(A,B) = P(A|B)P(B)$

$P(A,B|C) = P(A|B,C).P(B|C)$

$$= \int P(y^*|\theta, x^*).P(\theta|x^*)d\theta$$

$$\boxed{P(y^*|x^*) = \int P(y^*|\theta, x^*).P(\theta)d\theta}$$ (as θ does not depend on x^*)

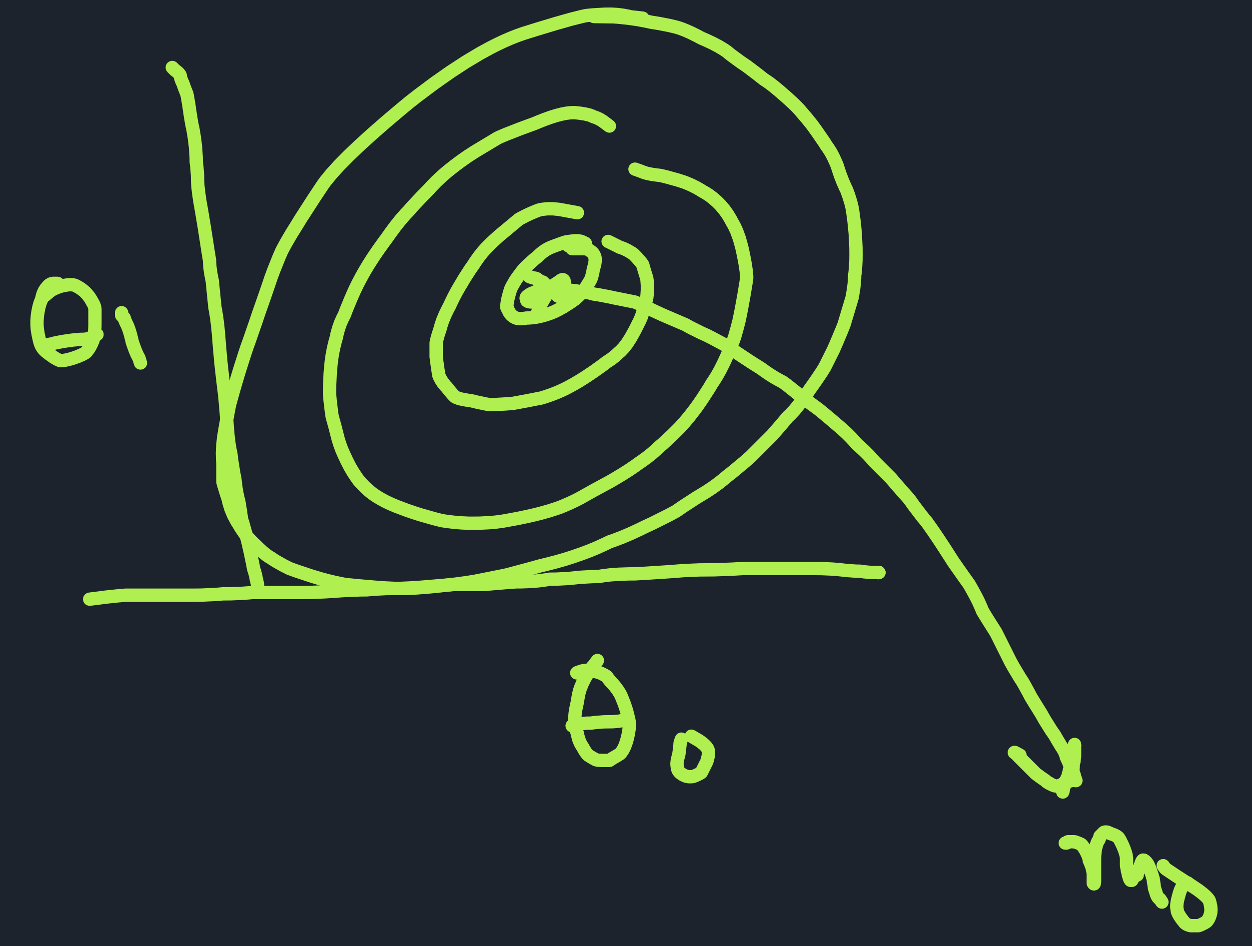
$$= \int f(\theta)P(\theta)d\theta$$



$$y \sim N(x^T\theta, \sigma^2)$$

Prior predictive for linear regression

$$p(\theta) = N(\overset{\mathbb{R}^D}{m_0}, \overset{\mathbb{R}^{D\times D}}{S_0})$$



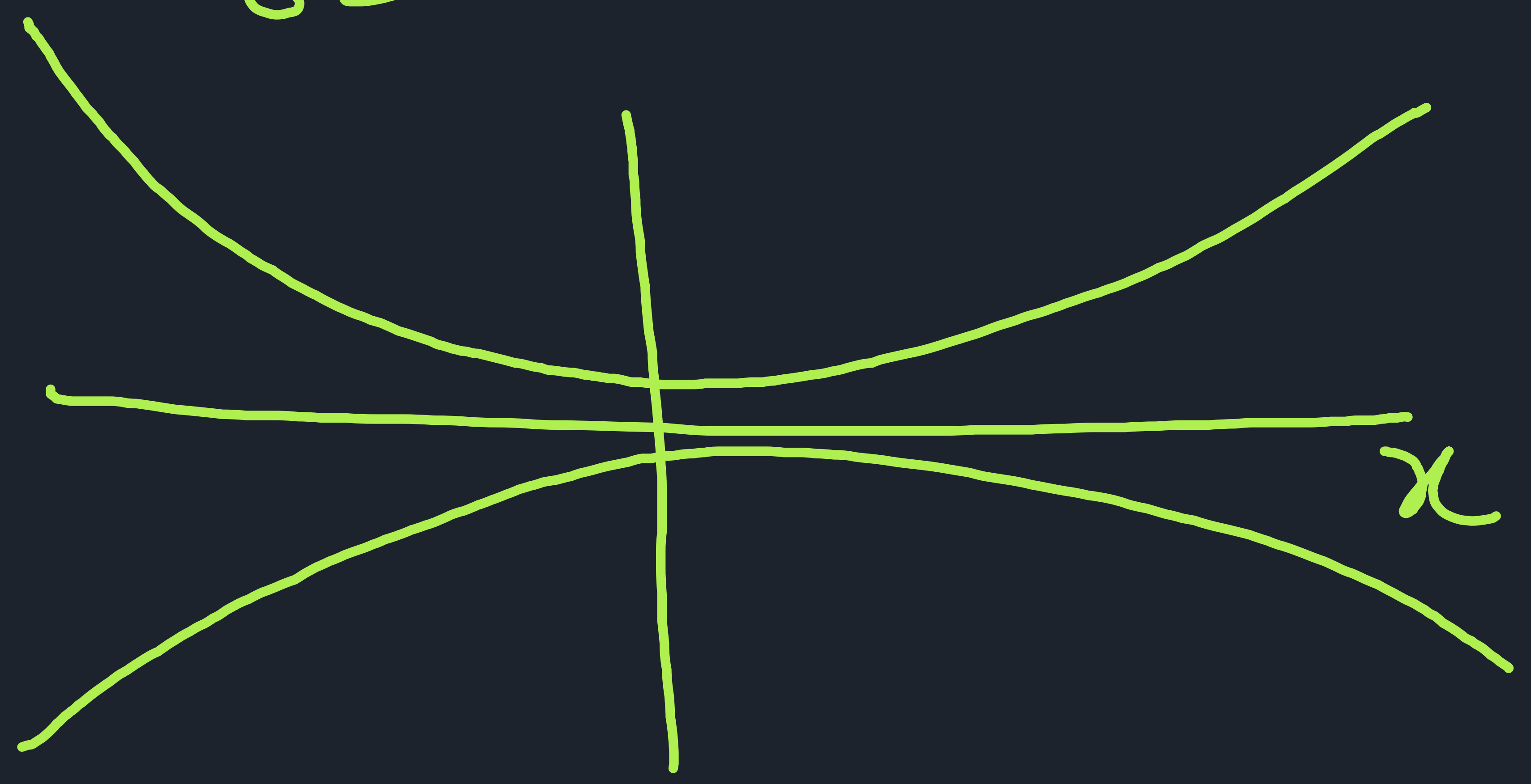
$$p(\overset{\mathbb{R}}{y^*} \mid \overset{\mathbb{R}^D}{x^*}) = ?$$

$$= \int \underbrace{p(y^* \mid x^*, \theta)}_{N(x^{*T}\theta,\, \sigma^2)} \underbrace{p(\theta)}_{N(m_0,\, \underbrace{S_0}_{D\times D})} d\theta \quad = N\left(x^{*T} m_0,\ \sigma^2 + x^{*T} S_0 x^*\right)$$

Case I

$$p(\theta) = N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\right)$$

$\sigma = 1$




```
σ = 1
m0 = jnp.array([0,0])
S0 = jnp.array([[1,0],[0,1]])
prior = tfd.MultivariateNormalFullCovariance(m0, S0)

x* = jnp.linspace(-2,2,100)
ppd = tfd.Normal(x*^T m0, σ^2 + x*^T S0 x*)
ppd_μ = ppd.mean
σ = ppd.scale
plt.fill_between(x*, ppd_μ-2σ, ppd_μ+2σ)
```

Posterior Predictive Distribution

$$p(y^*|x^*,D) = ?$$

$$= \int p(y^*,\theta|x^*,D)\,d\theta = \int p(y^*|x^*,D,\theta)\cdot p(\theta|x^*,D)\,d\theta$$

$$= \int \underbrace{p(y^*|x^*,\theta)}_{\text{Likelihood}}\cdot \underbrace{p(\theta|D)}_{\text{Posterior}}\,d\theta$$

Post. Pred. for Lin Reg.

$$P(\theta|D) = ? = N(m_N, S_N)$$

$$P(y^*|x^*, D) = N\left(x^{*T} m_N, \sigma^2 + x^{*T} S_N x^*\right)$$

Posterior Predictive for Logistic Regression

$$\int f(\theta)\,.\,p(\theta)\,d\theta$$

$$P(y^*|x^*, D) = \int P(y^*|x^*, \theta)\,.\,P(\theta|D)\,d\theta$$

Bernoulli $\left(P = \sigma(x^{*T}\theta)\right)$

Normal under Laplace Approx

Intractable

Monte-Carlo Sampling

$$I = \int f(x)\, p(x)\, dx$$

$$= \mathop{E}_{p}[f(x)]$$

$$= \frac{\sum_{i=1}^{S} f(x_i)}{S} \quad \text{where } x_i \sim p \text{ i.i.d.}$$

$X \sim N(0,1)$

$E[X] = ?$

$$= \frac{\sum_{i=1}^{S} x_i}{S} \quad ; x_i \sim N(0,1) \text{ i.i.d.}$$

seed=1

S=10

```
x = tfd.Normal(0,1)

x_samples = x.sample(seed, S)

...  μ̂ = x_samples.mean()
```