Today

1) Maximum Likelihood vs MAP Estimation (ML vs MAP)

(Maxm-a-Posteriori) ← MAP

2) Introduction to Random Processes

Prior vs Posterior vs Likelihood

(Unknown) X - - - - correlated with - - - - (Observed) Y - - - → Goal: Estimate X from Y.

↑ data that we collect/observe

Prior → PDF/PMF of unknown X → $f_X(x)$.

Posterior distribution → Conditional PDF of X given that we observe $Y = y$. → $f_{X|Y}(x|y)$

Likelihood (distrib.) → Conditional PDF of Y if the unknown data was $X = x$ → $f_{Y|X}(y|x)$

From last lecture: estimate θ using $(x_1, \dots x_n)$ → Y. observations.

ML estimation $\hat{\theta}_{ML} = \arg\max_{\theta} f_{X_1,\dots X_n}(x_1,\dots x_n|\theta)$

ML estimate of X from Y?

$$\hat{x}_{ML} = \arg\max_{x} f_{Y|X}(y|x).$$

MAP estimate of X from Y?

$$\hat{x}_{MAP} = \arg\max_{x} f_{X|Y}(x|y)$$

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$$f_{X|Y}(x|y) = \frac{f_X(x) \times f_{Y|X}(y|x)}{f_Y(y)}$$

observe y → maximize

constant.

$y = 3 \Rightarrow$

$Y = 3$

Prior. Likelihood.

$$\Rightarrow \hat{x}_{MAP} = \arg\max_{x} f_X(x) \times f_{Y|X}(y|x)$$

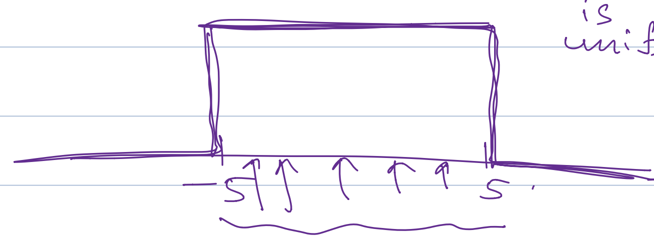
VS

$$\hat{x}_{ML} = \arg\max_{x} f_{Y|X}(y|x)$$

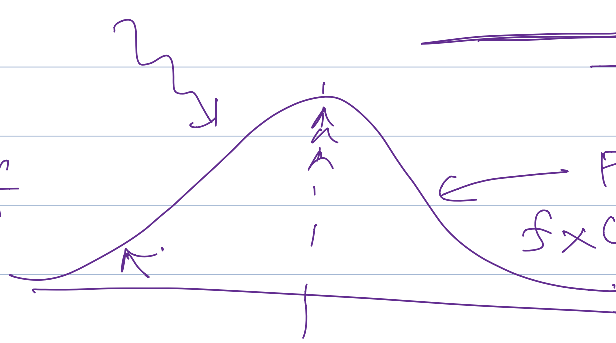
When are MAP & ML estimation equivalent??

if the prior $f_X(x) = \text{const} \;\; \forall x$.

→ Prior of X is uniform



if



Example: $\underbrace{X}_{\text{unknown}} \sim \mathcal{N}(0, \sigma_X^2)$ ((W, X) are independent)

$\longrightarrow Y = X + \underbrace{W}_{\text{noise}} \sim \mathcal{N}(0, \sigma_W^2)$

(1) Find the ML estimate of X given $Y = y$ is observed.

(2) " " MAP " " " " " .

$\underbrace{f_X(x)}_{\text{Prior}}$, $\underbrace{f_{X|Y}(x|y)}_{\text{Posterior}}$, $\underbrace{f_{Y|X}(y|x)}_{\text{likelihood}}$.

OK ??

$$\longrightarrow \underbrace{f_X(x)}_{\text{Prior}} = \frac{1}{\sqrt{2\pi}\sigma_X} \cdot e^{-\frac{x^2}{2\sigma_X^2}}$$

$$\longrightarrow \underbrace{f_{Y|X}(y|x)}_{\text{(Likelihood)}} = \frac{1}{\sqrt{2\pi}\,\sigma_W} e^{-\frac{(y-x)^2}{2\sigma_W^2}}$$

$$(Y \mid X = x) = \underbrace{x + W}\sim \mathcal{N}(x, \sigma_W^2)$$

$$E[x+W] = x + \underbrace{E[W]}_{=0} = x.$$

$$\hat{x}_{ML} \overset{?}{=} \arg\max_{x} \frac{1}{\sqrt{2\pi}\sigma_W} e^{-\frac{(y-x)^2}{2\sigma_W^2}}$$

$$\boxed{\hat{x}_{ML} \overset{?}{=} y.}$$

$$\max_{\cdot} f_{X|Y}(x|y) \iff \max_{\cdot} \underbrace{f_X(x)} \times f_{Y|X}(y|x)$$

$$\hat{x}_{MAP} = \arg\max_{x} \frac{1}{(\sqrt{2\pi})^2 \sigma_X \sigma_W} e^{-\frac{x^2}{2\sigma_X^2}} \cdot e^{-\frac{(y-x)^2}{2\sigma_W^2}}$$

$$= \arg\max_{x} e^{-\frac{1}{2}\underbrace{\left[\frac{x^2}{\sigma_X^2} + \frac{(y-x)^2}{\sigma_W^2}\right]}}$$

$$= \arg\min_{x} \underbrace{\left[\frac{x^2}{\sigma_X^2} + \frac{(y-x)^2}{\sigma_W^2}\right]}.$$

$$\frac{d(\ldots)}{dx} = \frac{2x}{\sigma_X^2} - \frac{2(y-x)}{\sigma_W^2} = 0.$$

$$\frac{x}{\sigma_X^2} = \frac{y-x}{\sigma_W^2}.$$

$$x\left(\frac{1}{\sigma_X^2} + \frac{1}{\sigma_W^2}\right) = \frac{y}{\sigma_W^2}.$$

$$\boxed{\hat{x}_{MAP} = \left(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_W^2}\right) y.}$$