Analytical inference

Posterior distribution

Posterior distribution

Likelihood

Prior

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$$

Evidence

Posterior distribution

$$P(\theta|X) = \frac{\overbrace{P(X|\theta)}^{\text{Likelihood}}\overbrace{P(\theta)}^{\text{Prior}}}{\underbrace{P(X)}_{\text{Evidence}}}$$

What is $P(X)$?

Posterior distribution

Likelihood

Prior

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$$

Evidence

What is $P(X)$?

Van Gogh Starry night

Posterior distribution

Likelihood

Prior

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$$

Evidence

What is $P(X)$?

Van Gogh Starry night

Van Gogh, Starry night over the Rhone

Maximum a posteriori

Maximum a posteriori

Maximum a posteriori

$$\theta_{\mathrm{MP}} = arg\max_{\theta} P(\theta|X)$$

Maximum a posteriori

$$\theta_{\mathrm{MP}} = arg\max_{\theta} P(\theta|X)$$

$$\theta_{\mathrm{MP}} = arg\max_{\theta} \frac{P(X|\theta)P(\theta)}{P(X)}$$

Maximum a posteriori

$$\theta_{\mathrm{MP}} = arg\max_{\theta} P(\theta|X)$$

$$\theta_{\mathrm{MP}} = arg\max_{\theta} \frac{P(X|\theta)P(\theta)}{P(X)}$$

$$\theta_{\mathrm{MP}} = arg\max_{\theta} P(X|\theta)P(\theta)$$

Maximum a posteriori

$$\theta_{\mathrm{MP}} = arg \max_{\theta} P(\theta|X)$$

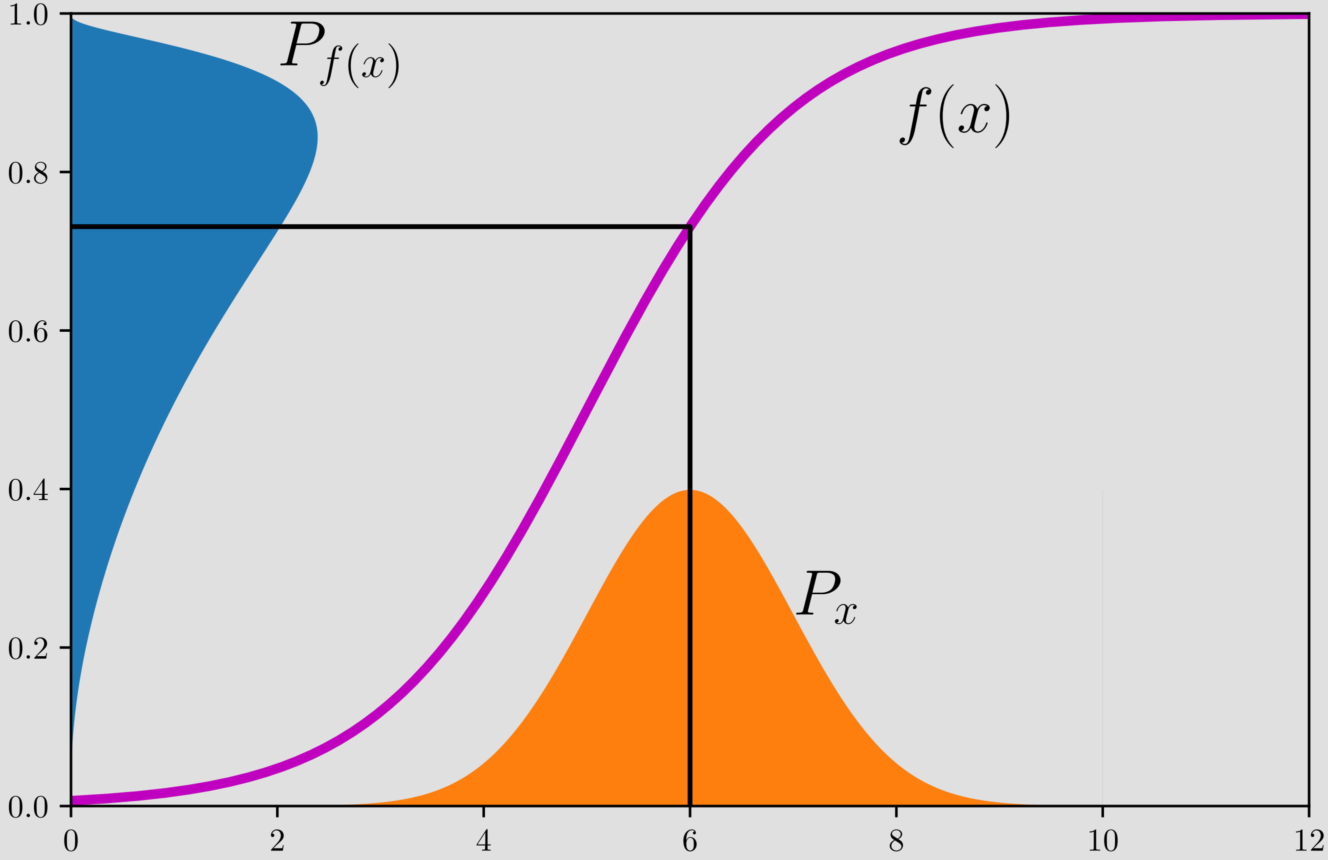
$$\theta_{\mathrm{MP}} = arg \max_{\theta} \frac{P(X|\theta)P(\theta)}{P(X)}$$

$$\theta_{\mathrm{MP}} = arg \max_{\theta} P(X|\theta)P(\theta)$$

Optimization problem!

MAP: problems

Not invariant to reparametrization



MAP: problems

Can't use as prior

$$P_k(\theta) = \frac{P(x_k|\theta)P_{k-1}(\theta)}{P(x_k)}$$

MAP: problems

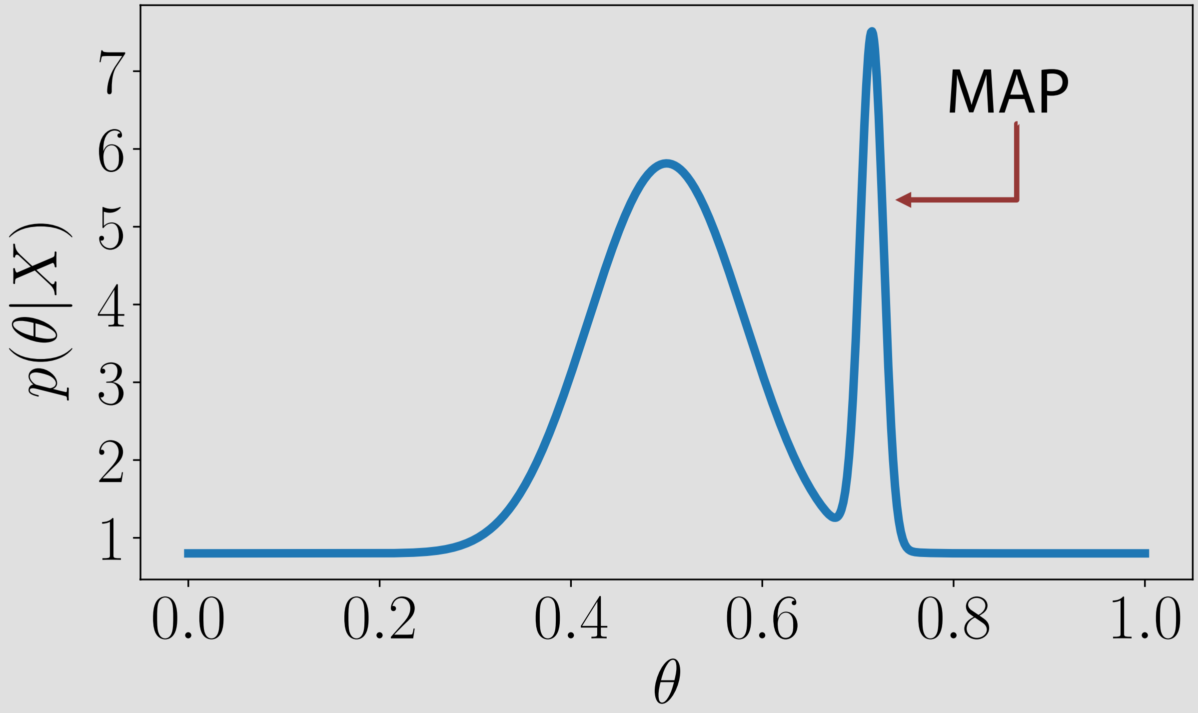
Can’t use as prior

$$P_k(\theta) = \frac{P(x_k|\theta)P_{k-1}(\theta)}{P(x_k)}$$

$$P_k(\theta) = \frac{P(x_k|\theta)\delta(\theta - \theta_{\mathrm{MP}})}{P(x_k)} = \delta(\theta - \theta_{\mathrm{MP}})$$

MAP: problems

MAP is a solution to $L(\theta) = \mathbb{I}[\theta \neq \theta^*] \to \min_\theta$



MAP: problems

Objectives	Solution
$L(\theta) = \mathbb{I}[\theta \neq \theta^*] \to \min\limits_{\theta}$	Mode
$L(\theta) = \mathbb{E}(\theta - \theta^*)^2 \to \min\limits_{\theta}$	Mean
$L(\theta) = \mathbb{E}\lvert\theta - \theta^*\rvert \to \min\limits_{\theta}$	Median

MAP: problems

Can't compute credible regions

$$\theta_{\mathrm{MP}} = 12.53$$

MAP: problems

Can't compute credible regions

$$\theta_{\mathrm{MP}} = 12.53$$

$$\theta_{\mathrm{MP}} = 12.53 \pm 1000$$

$$\theta_{\mathrm{MP}} = 12.53 \pm 0.001$$

Summary

Pros:

- Easy to compute

Cons:

- Not invariant to reparametrization
- Can’t use as a prior
- Finds untypical point
- Can’t compute credible intervals