



Computer Science Year 2

Algorithms & Data

Estimation, Regression, Classification

Prof Alin Achim



🔥 Example 2

- Assume that the likelihood function is exponential:

$$p(x|\theta) = \begin{cases} \theta \exp(-\theta x), & \text{for } x > 0 \\ 0, & \text{for } x < 0 \end{cases}$$

and that the prior pdf is

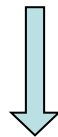
$$p(\theta) = \begin{cases} \lambda \exp(-\lambda\theta), & \text{for } \theta > 0 \\ 0, & \text{for } \theta < 0 \end{cases}$$

$$\hat{\theta}_{MAP} = ?$$

🔥 Example 2 - Exponential prior

- The MAP estimator can be found as follows:-

$$\hat{\theta}(x) = \operatorname{argmax} [\ln p(x|\theta) + \ln p(\theta)]$$



$$\begin{aligned}\hat{\theta}(x) &= \operatorname{argmax} [\ln[\theta \exp(-\theta x)] + \ln[\lambda \exp(-\lambda \theta)]] = \\ &= \operatorname{argmax} [\ln \theta - \theta x + \ln \lambda - \lambda \theta]\end{aligned}$$

- Differentiating with respect to θ

$$\frac{d}{d\theta} [\ln \theta - \theta x + \ln \lambda - \lambda \theta] = \frac{1}{\theta} - x - \lambda$$

- Setting equal to 0 yields

$$\hat{\theta} = \frac{1}{x + \lambda}$$

🔥 Example 3 - Gaussian prior

- Assume that the likelihood function is

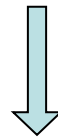
$$p(x|\theta) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp - \frac{(x - \theta)^2}{2\sigma_n^2}$$

and that the prior pdf is

$$p(\theta) = \frac{1}{\sigma \sqrt{2\pi}} \exp - \frac{\theta^2}{2\sigma^2}$$

- The MAP estimator can be found as follows:-

$$\hat{\theta}(x) = \operatorname{argmax} [\ln p(x|\theta) + \ln p(\theta)]$$



$$\hat{\theta}(x) = \operatorname{argmax} \left[-\frac{(x - \theta)^2}{2\sigma_n^2} - \frac{\theta^2}{2\sigma^2} \right]$$

🔥 Example 3 (continued)

$$\hat{\theta}(x) = \operatorname{argmax} \left[-\frac{(x - \theta)^2}{2\sigma_n^2} - \frac{\theta^2}{2\sigma^2} \right]$$

- Differentiating with respect to θ

$$\frac{d}{d\theta} \left[-\frac{(x - \theta)^2}{2\sigma_n^2} - \frac{\theta^2}{2\sigma^2} \right] = \frac{x - \theta}{\sigma_n^2} - \frac{\theta}{\sigma^2}$$

- Setting equal to 0 yields

$$\frac{x - \theta}{\sigma_n^2} = \frac{\theta}{\sigma^2}$$

- Finally:-

$$\hat{\theta} = \frac{\sigma^2}{\sigma^2 + \sigma_n^2} x$$

🔥 Example 4: Laplace Prior - soft thresholding

- Assume that the likelihood function is

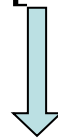
$$p(x|\theta) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp - \frac{(x - \theta)^2}{2\sigma_n^2}$$

and that the prior pdf is

$$p(\theta) = \frac{1}{\sigma \sqrt{2}} \exp - \frac{\sqrt{2}|\theta|}{\sigma}$$

- The MAP estimator can be found as follows:-

$$\hat{\theta}(x) = \operatorname{argmax} [\ln p(x|\theta) + \ln p(\theta)]$$



$$\hat{\theta}(x) = \operatorname{argmax} \left[-\frac{(x - \theta)^2}{2\sigma_n^2} - \frac{\sqrt{2}|\theta|}{\sigma} \right]$$

🔥 Example 4: Laplace Prior - soft thresholding

$$\hat{\theta}(x) = \operatorname{argmax} \left[-\frac{(x - \theta)^2}{2\sigma_n^2} - \frac{\sqrt{2}|\theta|}{\sigma} \right]$$

- Differentiating with respect to θ

$$\frac{d}{d\theta} \left[-\frac{(x - \theta)^2}{2\sigma_n^2} - \frac{\sqrt{2}|\theta|}{\sigma} \right] = \frac{x - \theta}{\sigma_n^2} - \frac{\sqrt{2}}{\sigma} \operatorname{sign}(\theta)$$

- Setting equal to 0 yields

$$\hat{\theta} = x - \sqrt{2} \frac{\sigma_n^2}{\sigma} \operatorname{sign}(\theta)$$

- Finally:-

$$\hat{\theta} = \operatorname{sign}(x) \left(|x| - \sqrt{2} \frac{\sigma_n^2}{\sigma} \right)_+, \text{ where } (g)_+ = \begin{cases} 0, & g < 0 \\ g, & \text{otherwise} \end{cases}$$

🔥 Soft thresholding and other Bayesian Estimators

- Behaviour comparison for four different estimators
 - Cauchy thresholding (CT)
 - Soft thresholding (ST)
 - Amplitude scale invariant Bayes estimator (ABE)

