MLE vs. MAP Estimator with Gaussian Prior

- consider the single measurement
- $z = x + w$ of the unknown parameter x in the presence of the additive measurement noise w,
- assumed to be a normally (Gaussian) distributed random variable with mean zero and variance σ^2 $w \sim \mathcal{N}(0, \sigma^2)$
- first assume x is an unknown constant (no prior information about it is available)
- the likelihood function of x

$$\Lambda(x) = p(z|x) = \mathcal{N}(z; x, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-x)^2}{2\sigma^2}}$$

$$\hat{x}^{\mathrm{ML}} = \arg\max_x \Lambda(x) = z$$

- the peak or mode of the above equation occurs at $x = z$

MLE vs. MAP Estimator with Gaussian Prior

- Next assume that the prior information about the parameter is that x is Gaussian with mean $\bar{x}$ and variance σ_0^2

$$p(x) = \mathcal{N}(x; \bar{x}, \sigma_0^2)$$

  assume x is independent of w

- the posterior pdf of x conditioned on the observation z is

$$p(z|x) = \frac{p(z|x)p(x)}{p(z)} = \frac{1}{c} e^{-\frac{(z-x)^2}{2\sigma^2} - \frac{(x-\bar{x})^2}{2\sigma_0^2}}$$

  where $c = 2\pi\sigma\sigma_0 p(z)$ is the normatlization constant indepedent of x

- normalization constant which guarantees that the pdf integrates to unity

MLE vs. MAP Estimator with Gaussian Prior

- it can be shown that the posterior pdf of x is (i.e., Gaussian)

$$p(x|z) = \mathcal{N}[x; \xi(z), \sigma_1^2] = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\xi(z))^2}{2\sigma_1^2}}$$

$$\xi(z) \triangleq \frac{\sigma^2}{\sigma_0^2 + \sigma^2}\bar{x} + \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2} z = \bar{x} + \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2}(z - \bar{x})$$

$$\sigma_1^2 \triangleq \frac{\sigma_0^2 \sigma^2}{\sigma_0^2 + \sigma^2}$$

- maximization of $p(x|z)$ with respect to x yields

$$\hat{x}^{\text{MAP}} = \xi(z)$$

- $\xi(z)$ is the MAP estimator for the random parameter x