1. Start by deriving the MAP estimate of the Normal distribution parameters μ and σ^2

Given data

$$D = (x_1, x_2, x_3, \dots x_n), \; x_i \in R^n$$

Assume a joint distribution $P(D, \theta)$; where θ is random variable

$\hat{\theta}_{MAP} = \arg_\theta \max P(\theta/D)$; where $P(\theta/D)$ is the posterior distribution on θ given the data D.

By bayes rule.

$$\hat{\theta}_{MAP} = \arg_\theta \max \frac{P(D|\theta) . P(\theta)}{P(D)}$$

where $P(\theta)$ = Prior knowledge of θ, $P(D|\theta)$ = Likelihood

$P(D)$ = does not depend on θ

$$\hat{\theta}_{MAP} = \arg_\theta \max P(D|\theta) . P(\theta)$$

Taking log

$$\hat{\theta}_{MAP} = \arg_\theta \max \log [P(D|\theta) + \log P(\theta)]$$

$$\hat{\theta}_{MAP} = \arg_\theta \max [\Sigma \log P(x_i|\theta) + \log P(\theta)]$$

where $\log P(\theta)$ is the Regularisation term.

$X_1, X_2, X_3 \dots X_n \sim N(\mu, \sigma^2)$ is i.id given μ;

$\mu \sim N(V, \beta^2)$. Assuming σ^2, β^2, V are known... and $(\theta = \mu)$

$$\hat{\mu}_{MAP} = \arg_\mu \max \sum_{i=1}^{n} \log P(x_i|\mu, \sigma^2) + \log P(\mu)$$

$$\hat{\mu}_{MAP} = \arg_\mu \max \sum_{i=1}^{n} \log \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}} + \log \frac{1}{\sqrt{2\pi\beta^2}} e^{\frac{-(V-\mu)^2}{2\beta^2}}$$

Solving for μ

$$\hat{\mu}_{MAP} = \frac{\sigma^2 V + \frac{n}{n}\beta^2 \sum_{i=1}^{n} x_i}{\sigma^2 + n\beta^2} = \frac{\sigma^2 V + n\beta^2 \sum_{i=1}^{n} \frac{x_i}{n}}{\sigma^2 + n\beta^2} = \frac{\sigma^2 V + n\beta^2 \bar{x}}{\sigma^2 + n\beta^2}$$

where $\bar{x}$ is the sample Mean

So therefore as we derived earlier $\bar{\mu} = \bar{\mu}_{MLE}$

Above equation can also be written as,

$$= \frac{\sigma^2}{\sigma^2 + n\beta^2} V + \frac{n\beta^2}{\sigma^2 + n\beta^2} \bar{\mu} \longrightarrow (1)$$

where

$$C_1 = \frac{\sigma^2}{\sigma^2 + n\beta^2} \text{ and } C_2 = \frac{n\beta^2}{\sigma^2 + n\beta^2} \text{ are constant}$$

From equation(1) is convex combination of the prior mean V and sample mean $\bar{\mu}$.

Also $C_1 + C_2 = 1$

Also from the above equation we can prove that

if $n \rightarrow \infty$;

$$\hat{\mu}_{MAP} = \hat{\mu}_{MLE} = \bar{\mu}$$

Scanned with CamScanner