Maximum Likelihood Estimation (MLE)

$\mathcal{X}$: training data set

θ : parameters

$$\theta^*_{MLE} = \arg\max_{\theta} p(\mathcal{X} \mid \theta)$$

$$p(\mathcal{X} \mid \theta) = \prod_{i=1}^{N} p(x_i \mid \theta)$$

Maximum a Posteriori Estimation (MAP)

$$\theta^*_{MAP} = \arg\max_{\theta} p(\theta \mid \mathcal{X})$$

$$= \arg\max_{\theta} \frac{p(\mathcal{X} \mid \theta)\, p(\theta)}{p(\mathcal{X})}$$

Parametric Regression :

$$y = \underbrace{f(x)}_{\text{underlying process}} + \underbrace{\epsilon}_{\text{noise}}$$

$$x_{N+1} \rightarrow \hat{y}_{N+1} = ?$$

$$f(x_{N+1}) \rightarrow \hat{y}_{N+1} \leftarrow g(x_{N+1}|\theta)$$

$\Rightarrow$ approximate $f(x)$ with $g(x|\theta)$

$E[x] = \gamma$

$E[x+c] = \gamma + c$

$VAR[x] = k^2$

$VAR[x+c] = k^2$

Assumptions:

I. $p(\epsilon) \sim N(\epsilon; 0, \sigma^2)$

II. $p(y|x) \sim N(y; g(x|\theta), \sigma^2)$

$y|x \Rightarrow f(x) + \epsilon$

$$y|x \Rightarrow \underbrace{g(x|\theta)}_{\text{constant}} + \underbrace{\epsilon}_{\text{R.V.}}$$

$$E[y|x] = E[\underbrace{g(x|\theta)}+\epsilon]$$

$$= E[g(x|\theta)] + E[\epsilon]$$

$$= g(x|\theta) + 0$$

$$VAR[y|x] = VAR[g(x|\theta)+\epsilon]$$

$$= VAR[\epsilon] = \sigma^2$$

$x \sim N(x; 0, 9)$

$x+5 \sim N(x; 5, 9)$