Question : 3

(A.) The likelihood function for a heteroscedastic setting for a single data point with input x_n and output t_n.

$$L(w \mid x_n, t_n) = p(t_n \mid x_n, w)$$

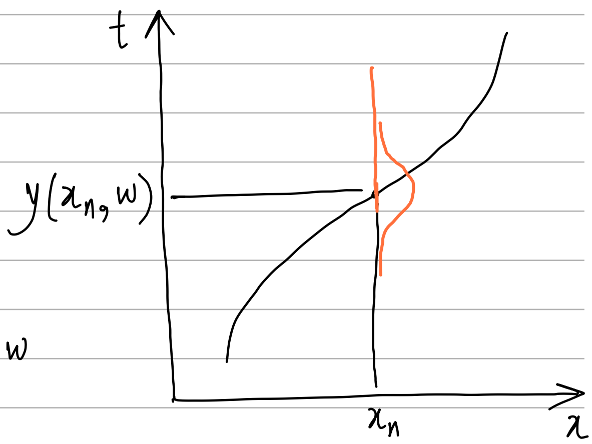
w = some parameter of given model.

→ likelihood function = probability of observed output t_n given the input x_n and parameter of model.

→ Taking Curve fitting problem into consideration. Given the value of x_n, the corresponding value of t_n has a Gaussian distribution.
mean = $y(x, w)$

$$\therefore \; p(t_n \mid x_n, w, \sigma^2) = \mathcal{N}(t_n \mid y(x_n, w), \beta^{-1})$$

precision parameter (β) = inverse of variance. (σ^2)



∴ Expression of likelihood function

$$L(w, \sigma^2 \mid x_n, t_n) = p(t_n \mid x_n, w, \sigma^2)$$

→ Now introducing the prior to the distribution over the polynomial coefficients w

$$p(w)$$

(B.) objective function considered for ML estimation

$$L(w, \sigma^2 \mid x_n, t_n) = \prod_{n=1}^{N} \mathcal{N}(t_n \mid y(x_n, w), \beta^{-1})$$

$$\mathcal{N}(t_n \mid y(x_n, w), \beta^{-1}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(t_n - w^T\phi(x_n))^2\right)$$

objective function considered for MAP estimation

Posterior = likelihood * Prior

$$L(w, \sigma^2 \mid x_n, t_n, p(w)) = \prod_{n=1}^{N} \mathcal{N}(t_n \mid y(x_n, w), \beta^{-1}) * p(w)$$

(C.) Each data point t_n is associated with a weighted factor $r_n > 0$.

→ Maximizing the Logarithm of likelihood function (obj. function for ML)

$$\ln L(w, \sigma^2 / x_n, t_n) = -\frac{1}{2} \sum_{n=1}^{N} \left\{ \frac{t_n - w^T\phi(x_n)}{\sigma^2} \right\}^2 + \frac{N}{2} \ln\beta - \frac{N}{2} \ln(2\pi)$$

Neglecting the constant terms and converting the $\frac{1}{\sigma^2}$ as r_n.

∴ Sum of squares error function becomes.

$$E(w) = \frac{1}{2} \sum_{n=1}^{N} r_n \left\{ t_n - w^T\phi(x_n) \right\}^2$$

→ Now $\frac{dE(w)}{dw} = 0$ & finding the solution for w.

$$0 = \frac{1}{2} \sum_{n=1}^{N} r_n \left(2\left(t_n - w^T\phi(x_n)\right) - w^T \right)$$

Expression of w that minimize the error function.

$$w = (\lambda I + \phi\phi^T)^{-1} \phi t$$