1 Find w

Observations: $x = (x_1, x_2, ..., x_n)^T$
Target: $t = (t_1, t_2, ..., t_n)^T = $ y(x, W) $+ \epsilon$
Model: $y(x, w) = w_0 + x_1 w_1 + x_2 w_2 + ... + x_n w_n$

$$t = y(x, w) + N(0, \beta^{-1}) t = N(y(x, w), \beta^{-1}) \ with \ \beta = \sigma^{-1}$$
$$p(t) = N(t|y(x, w), \beta^{-1})$$
$$\rightarrow p(t_n) = N(t_n|y(x_n, w), \beta^{-1})$$
$$\rightarrow p(t|x, w, \beta) = \prod_{n=1}^{N} N(t_n|y(x_n, w), \beta^{-1})$$
$$N(t_n|y(x, w), \beta^{-1}) = \frac{1}{\sqrt{2\pi\beta^{-1}}} e^{-(t_n - y(x,w))^2 \beta/2}$$

Using Maximum Likelihood, we have:

$$p(t|x, w, \beta) \rightarrow max$$
$$log(\prod_{n=1}^{N} e^{-(t_n - y(x,w))^2 \beta/2}) \rightarrow max$$
$$\sum_{n=1}^{N}(t_n - y(x, w))^2 \rightarrow min$$

We have: x $= \begin{bmatrix} 1 & x_1 \\ ... & \\ 1 & x_n \end{bmatrix}$, t $= \begin{bmatrix} t_1 \\ ... \\ t_n \end{bmatrix}$, w $= \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} \rightarrow y = \begin{bmatrix} y_1 \\ ... \\ y_n \end{bmatrix} = \begin{bmatrix} w_1 x_1 \\ ... \\ w_n x_n \end{bmatrix} = xw$

$\rightarrow t - y = \begin{bmatrix} t_1 - y_1 \\ ... \\ t_n - y_n \end{bmatrix} \rightarrow ||t - y||_2^2 = \sum_{n=1}^{N}(t_n - y_n)^2$

$$L = ||t - xw||_2^2$$
$$\frac{\sigma L}{\sigma w} = -2x^T(t - xw) = 0$$
$$x^T t = x^T xw$$
$$w = (x^T x)^{-1} x^T t$$

Chi Nguyen **Homework 3**

2 $\mathbf{X}^T X$ *inverse when* X *full rank*

When X full rank, $Xa = 0$ has the trivial solution a $= 0$
$\rightarrow$ The reduced row echelon form of X is I
$\rightarrow \prod_{i=1}^{N} E_i X = I$ with E_i is the matrix that X reduced row
$\rightarrow$X $= \prod_{N}^{i=1} E_i^{-1} I \rightarrow X$ *inverse* $\rightarrow$X$^T X$ inverse