Linear Regression - ~~EM~~ MLE

①

5/8/26

Can do EM for
- Latent weight
- Multiple lines (line membership)
- Missing data

$$y = w^T x + \varepsilon$$

$$\varepsilon \sim \mathcal{N}(0, \sigma^2)$$

~~$y = m \times 1$~~ ~~$x = m \times n$~~ ~~$w = n \times m$~~

$y = n \times 1$

$X = n \times m$

$w = 1 \times m$

$$p(y \mid x, w, \sigma^2) = \mathcal{N}(y \mid w^T x, \sigma^2)$$

$$\mathcal{N}(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]$$

$$\ln \mathcal{N}(\mu, \sigma^2) = -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{(x-\mu)^2}{2\sigma^2}$$

$$= -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma^2 - \frac{(x-\mu)^2}{2\sigma^2}$$

$$= -\frac{1}{2}\ln 2\pi - \ln\sigma - \frac{(x-\mu)^2}{2\sigma^2}$$

$$\ln p(y \mid x, w, \sigma^2) = -\frac{1}{2}\ln 2\pi - \ln\sigma - \frac{(y - w^T x)^2}{2\sigma^2}$$

$$\ln \mathcal{L}(w, \sigma^2) = \ln \prod^n p(y \mid x, w, \sigma^2) = \sum^n \ln p(y \mid x, w, \sigma^2)$$

$$\Rightarrow \sum^n -\frac{1}{2}\ln 2\pi - \ln\sigma - \frac{(y - w^T x)^2}{2\sigma^2}$$

$$\frac{\partial \mathcal{L}}{\partial w} = \frac{\partial}{\partial w}\sum^n -\frac{(y - w^T x)^2}{2\sigma^2} = \frac{-1}{2\sigma^2}\sum^n \frac{\partial}{\partial w}(y - w^T x)^2$$

$$\frac{\partial}{\partial w}(y - w^T x)^2 = 2(y - w^T x)(-x)$$

$$\frac{\partial \mathcal{L}}{\partial w} = \frac{-1}{2\sigma^2}\sum^n -2(y - w^T x)(x) = \frac{1}{\sigma^2}\sum^n (y - w^T x)(x)$$

$$\frac{\partial \mathcal{L}}{\partial w} = 0 \Rightarrow \frac{1}{\sigma^2}\sum^n (y - w^T x)(x) = 0 \Rightarrow \sum^n (y - w^T x)(x) = 0$$

$$\Rightarrow \sum^n yx - \sum^n (w^T x)x = 0 \Rightarrow \sum^n yx = \sum^n (w^T x)(x)$$

$$\frac{\partial \mathcal{L}}{\partial w} = 0 \Rightarrow \sum^{n} y x = \sum^{n} (w^T x)(x) \Rightarrow X^T \vec{y} = X^T X w \Rightarrow \text{Normal equations.}$$

$$\boxed{w = (X^T X)^{-1} X^T \vec{y}}$$

$$\frac{\partial \mathcal{L}}{\partial \sigma} = \frac{\partial}{\partial \sigma} \sum \underbrace{-\frac{1}{2}\ln 2\pi}_{\frac{\partial}{\partial \sigma} = 0} - \ln \sigma - \frac{(y - w^T x)^2}{2\sigma^2}$$

$$\Rightarrow \sum^{n} \frac{\partial}{\partial \sigma} - \ln \sigma - (y - w^T x)^2 \frac{\partial}{\partial \sigma} \frac{\cancel{2}\sigma^{-2}}{2}$$

$$= \sum^{n} \frac{-1}{\sigma} - (y - w^T x)^2 \left(-\frac{\cancel{4}\sigma^{-3}}{2} \right) = 0$$

$$\Rightarrow \sum^{n} -\frac{1}{\sigma} = \sum^{n} \frac{(y - w^T x)^2}{-\cancel{4}\sigma^3} \Rightarrow \frac{N}{\sigma} = \frac{1}{\cancel{4}\sigma^3} \sum^{n} (y - w^T x)^2$$

$$\Rightarrow \cancel{4} N \sigma^2 = \sum^{n} (y - w^T x)^2$$

$$\sigma = \frac{\sqrt{\sum^{n} (y - w^T x)^2}}{\sqrt{\cancel{4} N}} = \frac{1}{\cancel{2}\sqrt{N}} \sqrt{\sum^{n} (y - w^T x)^2}$$

$$\boxed{\sigma^2 = \frac{1}{N} \sum^{n} (y - w^T x)^2}$$