Gaussian Basics

$$p(x_i | \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i - \mu)^2}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x_i-\mu)^2}{E[(x-\mu)^2]}}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{\Delta x_i^2}{E[\Delta x^2]}}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x_i - E[x])^2}{E[(x-E[x])^2]}}$$

$$\sigma^2 = E[(x-\mu)^2] = \int (x-\mu)^2 p(x)\, dx$$

$$\mu = E[x]$$

Ratio of one squared deviation to mean squared deviation

→ Relative squared deviation from the mean squared deviation aka. variance

Likelihood of the Linear Model

$$p(\underline{y} \mid \underline{X}, \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_i - x_i\mu)^2}$$

Independence assumption

$$= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma_2}\sum_{i=1}^{n}(y_i - x_i\mu)^2}$$

$$p(\underline{y} \mid \underline{X}, \underline{\mu}, \sigma^2) = \left(\frac{1}{\sqrt{2\pi(\sigma I)^2}}\right)^n e^{-\frac{1}{2}(\underline{y} - \underline{X}\underline{\mu})^T(\sigma^2 I)^{-1}(\underline{y} - \underline{X}\mu)}$$

Prior

$$p(\underline{\mu}) = (2\pi\underline{\sigma_0}^2)^{-\frac{1}{2}} e^{-\frac{1}{2}(\mu - \underline{\mu_0})^T(\underline{\sigma_0^2 I})^{-1}(\underline{\mu} - \underline{\mu_0})}$$

$$\boxed{\mu = \theta}$$

$$p(\underline{\theta}) = (2\pi \underline{V_0}^2)^{-\frac{1}{2}} e^{-\frac{1}{2}(\underline{\theta} - \underline{\theta_0})^T \underline{V_0}^{-1}(\underline{\theta} - \underline{\theta_0})}$$

Posterior of θ

$$p(\underline{\theta} \mid \underline{y}, \underline{X}, \underline{\sigma}) \propto p(\underline{y} \mid \underline{X}, \underline{\theta}, \underline{\sigma}) \cdot p(\underline{\theta})$$

$$\propto e^{-\frac{1}{2}(\underline{y} - \underline{X}\underline{\theta})^T(\underline{\sigma^2 I})^{-1}(y - \underline{X}\underline{\theta})} \cdot e^{-\frac{1}{2}(\theta - \theta_0)^T \underline{V}^{-1}(\theta - \theta_0)}$$

$$p(\underline{\theta}\mid \underline{y},\underline{X},\underline{\sigma}) \propto p(\underline{y}\mid \underline{X},\underline{\theta},\underline{\sigma})\cdot p(\underline{\theta})$$

$$\propto e^{-\frac{1}{2}(\underline{y}-\underline{X}\underline{\theta})^T(\sigma^2\underline{I})^{-1}(\underline{y}-\underline{X}\underline{\theta})} \cdot e^{-\frac{1}{2}(\theta-\theta_0)^T\underline{V}^{-1}(\theta-\theta_0)}$$

$$\propto e^{-\frac{1}{2}\left[(y-X\theta)^T(\sigma^2 I)^{-1}(y-X\theta) + (\theta-\theta_0)^T V_0^{-1}(\theta-\theta_0)\right]}$$

$$\propto e^{-\frac{1}{2}\left[\underbrace{\boxed{y^T(\sigma^2 I)^{-1}y}}_{\text{const}}\underbrace{\boxed{-2y^T(\sigma^2 I)^{-1}X\theta}}_{\theta}\underbrace{\boxed{+\theta^T X^T(\sigma^2 I)^{-1}X\theta + \theta^T V_0^{-1}\theta}}_{\theta\text{-squared}}\underbrace{\boxed{-2\theta_0^T V_0^{-1}\theta}}_{\theta}\underbrace{\boxed{+\theta_0^T V_0^{-1}\theta_0}}_{\text{const}}\right]}$$

$$\propto e^{-\frac{1}{2}\left[\text{const} + \theta^T\underbrace{\left(X^T(\sigma^2 I)^{-1}X + V_0^{-1}\right)}_{V_n^{-1}}\theta - 2\left(y^T(\sigma^2 I)^{-1}X + \theta_0 V_0^{-1}\right)\theta\right]}$$

$$\propto e^{-\frac{1}{2}\left(\text{const} + \theta^T V_n^{-1}\theta - 2\left(\frac{y^T X}{\sigma^2} + \theta_0 V_0^{-1}\right)\theta\right)}$$

$$\propto e^{-\frac{1}{2}\left(\text{const}_2 + \underbrace{\theta^T V_n^{-1}\theta}\ \boxed{\underbrace{-2\theta_n^T V_n^{-1}\theta} + 2\theta_n^T V_n^{-1}\theta + \underbrace{\theta_n^T V^{-1}\theta_n}} - 2\left(\frac{y^T X}{\sigma^2} + \theta_0 V_0\right)\theta\right]}$$

$$\propto e^{-\frac{1}{2}\left[\text{const}_2 + (\theta-\theta_n)^T V_n^{-1}(\theta-\theta_n)\right.} \quad \underbrace{+2\left(\theta_n^T V_n^{-1} - \frac{y^T X}{\sigma^2} - \theta_0 V_0\right)\theta}_{=0 \text{ when}}\ ]$$

$$\theta_n = V_n\left[V_0^{-1}\theta_0 + X^T y/\sigma^2\right]$$

$$\propto e^{-\frac{1}{2}\left[(\theta-\theta_n)V_n^{-1}(\theta-\theta_n)\right]}$$