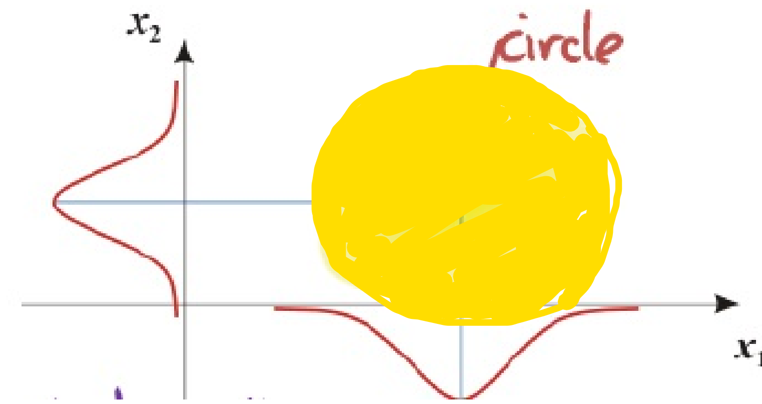
Bivariate Gaussian distribution example

Assume we have two ***independent*** *univariate Gaussian variables*

$x_1 = \mathcal{N}(\mu_1, \sigma^2)$ *and* $x_2 = \mathcal{N}(\mu_2, \sigma^2)$

Their joint distribution $p(x_1, x_2)$ *is:* $\Sigma = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$



$$p(x_1, x_2) = p(x_1) \cdot p(x_2) \quad | \text{ if independend}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_1-\mu)^2} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_2-\mu)^2}$$

$$= \text{const} \cdot e^{-\frac{1}{2}\left[(x_1-\mu)(\sigma^2)^{-1}(x_1-\mu) + (x_2-\mu)(\sigma^2)^{-1}(x_2-\mu)\right]}$$

$$\begin{bmatrix}(x_1-\mu) & (x_2-\mu)\end{bmatrix} \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}^{-1} \begin{bmatrix}(x_1-\mu) \\ (x_2-\mu)\end{bmatrix}$$

$$\begin{bmatrix}(x_1-\mu) & (x_2-\mu)\end{bmatrix} \begin{bmatrix}(\sigma^2)^{-1}(x_1-\mu) \\ (\sigma^2)^{-1}(x_2-\mu)\end{bmatrix}$$

$$P(y \mid \hat{y}, \sigma) = \prod \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) e^{-\frac{1}{2\sigma^2}(y_i - \hat{y}_i)^2}$$

product to sum

$$P(y \mid \hat{y}, \sigma) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{\sum -\frac{1}{2\sigma^2}\left[(y_i - \hat{y}_i)^2\right]}$$

sum to linear algebra

$$P(y \mid \hat{y}, \sigma) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\left[(\underline{y} - \underline{\hat{y}})^T(\underline{y} - \underline{\hat{y}})\right]}$$

$$\boxed{\underline{\hat{y}} = \underline{X\theta}}$$

$$\left[P(y \mid X, \theta, \sigma)\right] = \left[(2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}(\underline{y} - \underline{X\theta})^T(\underline{y} - \underline{X\theta})}\right]$$

$\log(\ldots)$

$$\ell(y \mid X, \theta, \sigma) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}(\underline{y} - \underline{X\theta})^T(\underline{y} - \underline{X\theta})$$

$$= -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\left(\underline{y}^T\underline{y} - 2y^TX\theta + \theta^TX^TX\theta\right)$$

$$\frac{d\ell(y \mid X, \theta, \sigma)}{d\theta} = 0 - \frac{1}{2\sigma}\left[0 + 2y^TX - 2X^TX\theta\right] \overset{!}{=} 0$$

$$\cancel{2}X^TX\theta = \cancel{2}y^TX$$

! MLE

$$\boxed{\theta_{ML} = (X^TX)^{-1}y^TX}$$

MLE for σ

$$\boxed{\frac{1}{\sigma^2} I = \Lambda}$$

S_μ : Scatter Matrix

$$\ell(y|X,\theta,\sigma) = \frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\overbrace{(y-X\theta)^T(y-X\theta)}^{S_\mu}$$

$$= \frac{n}{2}\log(\Lambda) - \frac{1}{2}\, tr(S_\mu \Lambda)$$

$$\boxed{\frac{d\, tr(AX)}{dX} = A^T}$$

Λ : precision Matrix

$$\frac{d\ell(y|X,\theta,\sigma)}{d\Lambda} = \frac{n}{2}\Lambda^{-T} - \frac{1}{2} S_\mu^T$$

$$0 \overset{!}{=} \frac{n}{2}\Lambda^{-T} - \frac{1}{2} S_\mu^T$$

$$\frac{n}{\cancel{2}}\Lambda^{-T} = \frac{1}{\cancel{2}} S_\mu^T$$

$$\boxed{\Lambda^{-T} = \frac{S_\mu'}{n} = \frac{(y-X\theta)^T(y-X\theta)}{n} = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i\theta)^2}$$