Covariance,Covariance matrix,Gaussian distribution

Covariance(공분산)

2개의 확률변수의 상관정도를 나타내는 값이다.

K variables are observed together

- k variables are observed together

$$\left(x_1^{(1)}, x_2^{(1)}, \cdots, x_k^{(1)}\right), \left(x_1^{(2)}, x_2^{(2)}, \cdots, x_k^{(2)}\right), \cdots, \left(x_1^{(n)}, x_2^{(n)}, \cdots, x_k^{(n)}\right)$$

$$X_i = \left\{x_i^{(1)}, x_i^{(2)}, \cdots, x_i^{(n)}\right\}$$

- Expectation and variance

$$E(X_i) = \mu_{X_i} = \frac{1}{n}\sum_{j=1}^{n} x_i^{(j)}$$

$$V(X_i) = \sigma_{X_i}^2 = E(X_i^2) - \mu_{X_i}^2 = E\left((X_i - \mu_{X_i})^2\right) = \frac{1}{n}\sum_{j=1}^{n}\left(x_i^{(j)} - \mu_{X_i}\right)^2$$

여기서 K는 K-dimensional을 말한다.(k차원=축)

- Covariance

$$\mathrm{cov}(X_i, X_j) = \sigma^2_{X_i,X_j} = \sigma^2_{i,j} = E(X_i X_j) - \mu_{X_i}\mu_{X_j}$$
$$= E\left((X_i - \mu_{X_i})(X_j - \mu_{X_j})\right) = \frac{1}{n}\sum_{m=1}^{n}\left(x_i^{(m)} - \mu_{X_i}\right)\left(x_j^{(m)} - \mu_{X_j}\right)$$

- Compare with variance

$$V(X_i) = \sigma^2_{X_i} = E(X_i^2) - \mu^2_{X_i} = E\left((X_i - \mu_{X_i})^2\right) = \frac{1}{n}\sum_{j=1}^{n}\left(x_i^{(j)} - \mu_{X_i}\right)^2$$

축과(dimension) 축 사이의 covariance

위 사진에서 i번째 축과 j번째 축 사이의 covariance

Meaning of Covariance

- How much two random variables(두차원,축) change together.
  - When x increases, y also increases and when x decreases, y also decreases—> covariance is plus
  - When x increases, y decreases and when x decreases, y increases—> covaraiance is minus

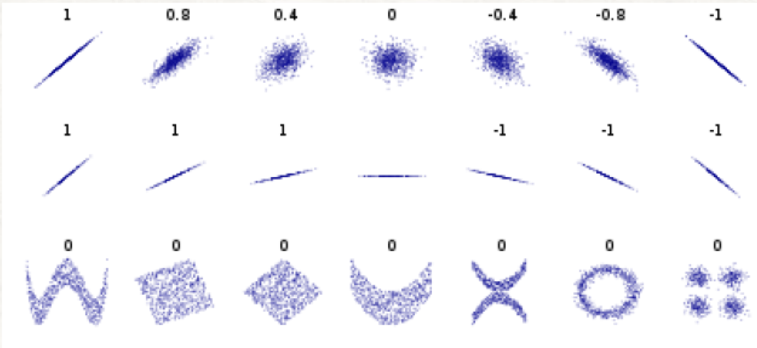
- The sign of the covariance shows the tendency in the linear relationship between the variables.
- The magnitude of the covariance(공분산크기) is not that easy to interpret.(공분산의 크기를 따지기 어렵다)
  - if zero, uncorrelated.

Pearson correlation coefficient(피어슨 상관계수)

변수 X 와 Y 간의 선형 상관 관계를 계량화한 수치다 .

피어슨 상관 계수는 코시-슈바르츠 부등식에 의해 +1과 -1 사이의 값을 가지며, +1은 완벽한 양의 선형 상관 관계, 0은 선형 상관 관계 없음, -1은 완벽한 음의 선형 상관 관계를 의미한다. 일반적으로 상관관계는 피어슨 상관관계를 의미한다.==normalized covariance

$$\rho_{X,Y} = corr(X,Y) = \frac{\sigma^2_{X,Y}}{\sigma_X \sigma_Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$



#코시 슈바르츠 부등식

코시 슈바르츠 부등식

$$(a^2 + b^2)(x^2 + y^2) \geq (ax + by)^2$$

(ay = bx일 때 등호 성립)

Covariance,Covariance Matrix

- k variables are observed together

$$\mathbf{x}_1 = \left(x_1^{(1)}, x_2^{(1)}, \cdots, x_k^{(1)}\right), \mathbf{x}_2 = \left(x_1^{(2)}, x_2^{(2)}, \cdots, x_k^{(2)}\right), \cdots, \mathbf{x}_n = \left(x_1^{(n)}, x_2^{(n)}, \cdots, x_k^{(n)}\right)$$

- Covariance Matrix

$$\Sigma = \begin{pmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,k} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \sigma_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k,1} & \cdots & \cdots & \sigma_{k,k} \end{pmatrix}$$

$$\mathbf{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})^T \bullet (\mathbf{x}_i - \boldsymbol{\mu})$$

$$\boldsymbol{\mu} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i$$

  - Symmetric because $\sigma_{i,j} = \sigma_{j,i}$

covariance Matrix

- Covariance Matrix

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma^2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \sigma^2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \sigma^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma^2 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \sigma^2 \end{pmatrix}$$

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma^2_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \sigma^2_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \sigma^2_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma^2_{m-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & \sigma^2_m \end{pmatrix}$$

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma^2_1 & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{12} & \sigma^2_2 & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1m} & \sigma_{2m} & \cdots & \sigma^2_m \end{pmatrix}$$

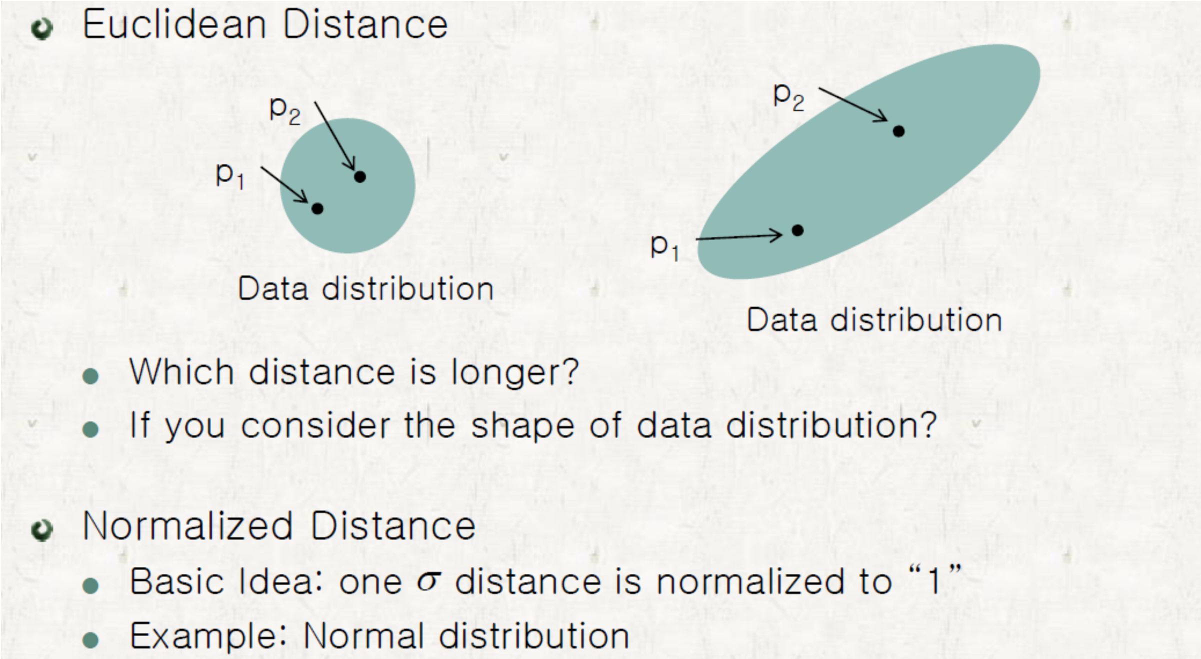
두번째 사진은 데이터의 축과 평행

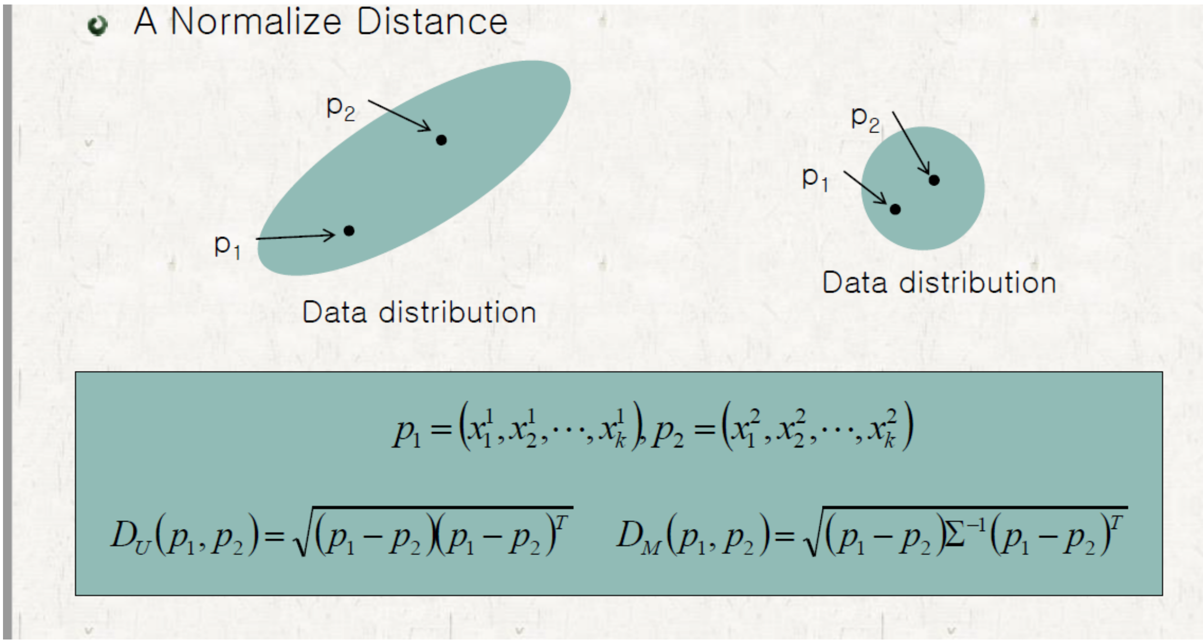
세번째 사진은 데이터의 축과 평행X

분포를 표준화 하기 위해 covariance 사용!!

Mahalanobis Distance

Normailze Distance





행벡터:1xm

열벡터:mx1

example

- Example: Variables are not correlated

$$p_1 = \left(x_1^{(1)}, x_2^{(1)}\right)$$
$$p_2 = \left(x_1^{(2)}, x_2^{(2)}\right)$$

Data distribution

Data distribution

$$\Sigma = \begin{pmatrix} \sigma_{1,1}^2 & 0 \\ 0 & \sigma_{2,2}^2 \end{pmatrix}$$

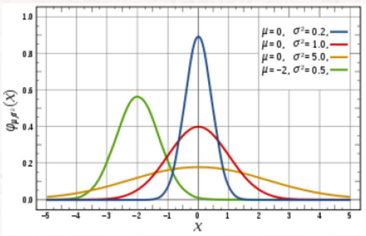
$$\Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_{1,1}^2} & 0 \\ 0 & \frac{1}{\sigma_{2,2}^2} \end{pmatrix}$$

$$D_M(p_1, p_2) = \sqrt{(p_1 - p_2)\Sigma^{-1}(p_1 - p_2)^T}$$
$$= \sqrt{\left(x_1^{(1)} - x_1^{(2)}, x_2^{(1)} - x_2^{(2)}\right) \begin{pmatrix} \frac{1}{\sigma_{1,1}^2} & 0 \\ 0 & \frac{1}{\sigma_{2,2}^2} \end{pmatrix} \begin{pmatrix} x_1^{(1)} - x_1^{(2)} \\ x_2^{(1)} - x_2^{(2)} \end{pmatrix}}$$
$$= \sqrt{\left(\frac{x_1^{(1)} - x_1^{(2)}}{\sigma_{1,1}}\right)^2 + \left(\frac{x_2^{(1)} - x_2^{(2)}}{\sigma_{2,2}}\right)^2}$$

Gaussian Distribution

- One dimensional Gaussian Distribution

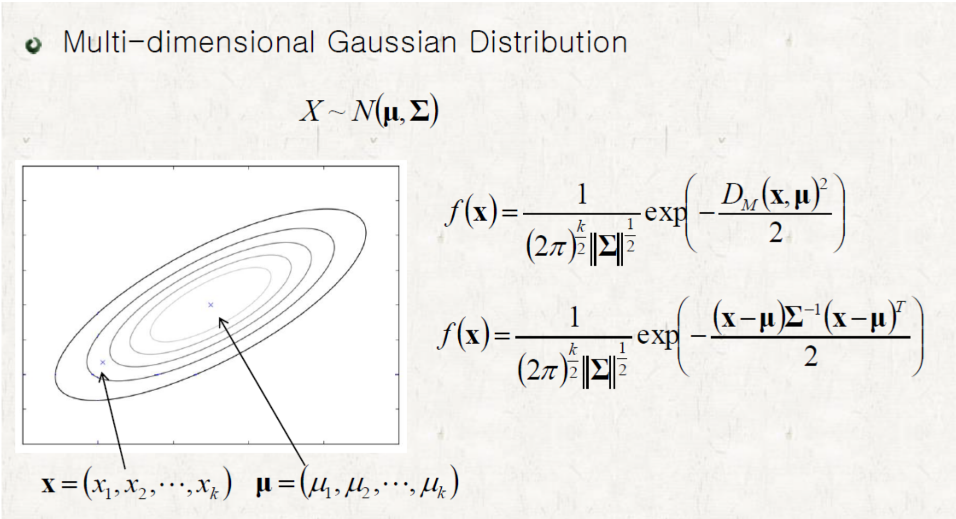
$$X \sim N(\mu, \sigma)$$



Mahalanobis distance between x and μ

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}\right)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{D_M(x, \mu)^2}{2}\right)$$



f(x): x(데이터)가 멀티 가우시안 분포에서 생성될 확률