$$N(\mu,\Sigma) = \frac{1}{(2\pi)^{d/2}\,|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$

실제 학습분포를 사용하게면 굉장히 어렵기 때문에 정규분포로 가정. → 실제학습 분포와의 차이로 인한 오류 발생

Class w_i 가 $N(\mu_i, \Sigma_i)$ 를 따른다고 가정하면,

w_i의 우도, $P(x|w_i) = N(\mu_i, \Sigma_i)$ 이고,

$P(x|w_i)P(w_i)$에 ln을 취한것을 분별함수 $g_i(x)$라 하면,

$$g_i(x) = \ln N(\mu_i, \Sigma_i) + \ln P(w_i)$$

$$= -\frac{d}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}(x-\mu_i)^T\Sigma_i^{-1}(x-\mu_i) + \ln P(w_i)$$

$$= -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}\left(x^T\Sigma_i^{-1}x - x^T\Sigma^{-1}\mu_i - \mu_i^T\Sigma^{-1}x + \mu_i^T\Sigma^{-1}\mu_i\right) + \ln P(w_i)$$

$$= -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}(x^T\Sigma^{-1}x - 2\mu_i^T\Sigma^{-1}x + \mu_i^T\Sigma^{-1}\mu_i) + \ln P(w_i)$$ → 변수 x에 대한 2차식.

두 class w_i와 w_j의 decision boundary는

$$g_i(x) = g_j(x)$$

이때, $g_{ij}(x) = g_i(x) - g_j(x)$ 를 결정경계함수라고 정의.

선형분별 분석 (Linear Discriminant Analysis, LDA)

- 모든 부류의 공분산 행렬이 같을 때. ($\Sigma_i \to \Sigma$)

$$g_i(x) = -\frac{1}{2}(2\mu_i^T\Sigma^{-1}x - \mu_i^T\Sigma^{-1}\mu_i + 2\ln P(w_i)) - \underbrace{\frac{1}{2}(x^T\Sigma^{-1}x + d\ln 2\pi + \ln|\Sigma|)}_{i\text{에 무관항}}$$

$$= (\Sigma^{-1}\mu_i)^T x + \left(\ln P(w_i) - \frac{1}{2}\mu_i^T\Sigma^{-1}\mu_i\right)$$

$$= w_i^T x + b_i$$ → x에 대한 1차식

결정 경계 함수 $g_{ij}(x)$에 대입하면,

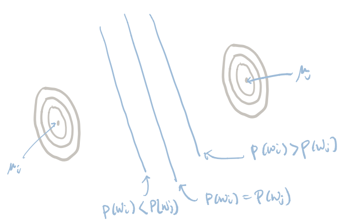
$$g_{ij}(x) = g_i(x) - g_j(x)$$

$$= (\Sigma^{-1}(\mu_i - \mu_j))^T x + \left(\ln P(w_i) - \ln P(w_j) - \frac{1}{2}\mu_i^T\Sigma^{-1}\mu_i + \frac{1}{2}\mu_j^T\Sigma^{-1}\mu_j\right)$$

$$= (\Sigma^{-1}(\mu_i - \mu_j))^T \left(x - \left(\frac{1}{2}(\mu_i + \mu_j) - \frac{\mu_i - \mu_j}{(\mu_i - \mu_j)^T\Sigma^{-1}(\mu_i - \mu_j)}\ln\frac{P(w_i)}{P(w_j)}\right)\right)$$

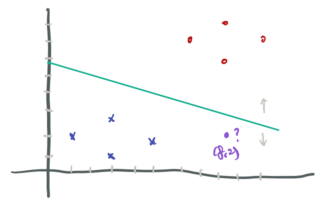
$$= w^T(x - x_0)$$

→ x_0를 지나고 w^T에 수직인 초평면.



Ex. $\omega_1 : (1,2)^T, (3,1)^T, (5,2)^T, (3,3)^T$

$\omega_2 : (6,6)^T, (8,5)^T, (10,6)^T, (8,7)^T$



$\rightarrow$

$$\mu_1 = \frac{1}{4}\left(\begin{pmatrix}1\\2\end{pmatrix}+\begin{pmatrix}3\\1\end{pmatrix}+\begin{pmatrix}5\\2\end{pmatrix}+\begin{pmatrix}3\\3\end{pmatrix}\right) = \begin{pmatrix}3\\2\end{pmatrix}$$

$$\mu_2 = \frac{1}{4}\left(\begin{pmatrix}6\\6\end{pmatrix}+\begin{pmatrix}8\\5\end{pmatrix}+\begin{pmatrix}10\\6\end{pmatrix}+\begin{pmatrix}8\\7\end{pmatrix}\right) = \begin{pmatrix}8\\6\end{pmatrix}$$

$$\Sigma_1 = \begin{pmatrix}8/3 & 0\\ 0 & 2/3\end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix}8/3 & 0\\ 0 & 2/3\end{pmatrix}$$

$$g_{12}(x) = \left(\Sigma^{-1}(\mu_1-\mu_2)\right)^T x + \left(\ln P(\omega_1) - \ln P(\omega_2) - \frac{1}{2}\mu_i^T\Sigma^{-1}\mu_i + \frac{1}{2}\mu_j^T\Sigma^{-1}\mu_j\right)$$

$$= \left(\begin{pmatrix}3/8 & 0\\ 0 & 3/2\end{pmatrix}\begin{pmatrix}3-8\\2-6\end{pmatrix}\right)^T x + \left(\ln\frac{P(\omega_1)}{P(\omega_2)} - \frac{1}{2}(3\ 2)\begin{pmatrix}3/8 & 0\\ 0 & 3/2\end{pmatrix}\begin{pmatrix}3\\2\end{pmatrix} + \frac{1}{2}(8\ 6)\begin{pmatrix}3/8 & 0\\ 0 & 3/2\end{pmatrix}\begin{pmatrix}8\\6\end{pmatrix}\right)$$

$$= (-15/8\ \ -6)\begin{pmatrix}x_1\\x_2\end{pmatrix} + \left(\ln\frac{P(\omega_1)}{P(\omega_2)} + 34.3125\right)$$

$$= -15/8\, x_1 - 6x_2 + \ln\frac{P(\omega_1)}{P(\omega_2)} + 34.3125$$

$P(\omega_1)=P(\omega_2)=0.5$ 인경우,

decision boundary $\rightarrow$ $5x_1 + 16x_2 - 91.5 = 0$. $\rightarrow$ $x_2 = -\frac{5}{16}x_1 + 5.72$

벡터 $\Sigma^{-1}(\mu_1-\mu_2)$에 수직

$\rightarrow (-15/8\ \ -6)$ $\rightarrow$ $(\mu_1-\mu_2)$를 Σ^{-1}로 선형변환.

Ex 2. $\omega_1 : (1,2)^T, (3,1)^T, (5,2)^T, (3,3)^T$

$\omega_2 : (7,6)^T, (8,4)^T, (9,6)^T, (8,8)^T$

최소 거리 분류기

Mahalanobis distance : $\left((x-\mu_i)^T\Sigma^{-1}(x-\mu_i)\right)^{1/2}$

Euclidean distance : $\left((x-\mu_i)^T(x-\mu_i)\right)^{1/2} = \|x-\mu_i\|$

Naïve Bayesian classifier

모든 Feature들이 독립이라는 가정.

$\rightarrow$ $P(x|\omega_i) = \prod_{j=1}^{d} P(x_j|\omega_i)$