LDA는 정규분포를 가정하며, 각 클래스는 평균이 다르더라고 분산은 동일함.

→ 각 클래스의 PDF는 아래와 같음.

$f_k(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu_k)^2\right)$

최종적인 LDA의 모델은 아래와 같음.

$\rightarrow p_k(x) = \frac{\pi_k \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu_k)^2\right)}{\sum_{l=1}^{K} \pi_l \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu_l)^2\right)}$

위나, 모든 클래스에 대해서 분산과 상수도 동일하기 때문에 클래스 k에 따라 변화하는 식만 남게 되면, 아래와 같다.

$\rightarrow \pi_k \exp\left(-\frac{1}{2\sigma^2}(x-\mu_k)^2\right)$

이 값을 로그변환 하면 아래와 같다.

$\rightarrow \log \pi_k - \frac{x^2}{2\sigma^2} + \frac{\mu_k x}{\sigma^2} - \frac{\mu_k^2}{2\sigma^2}$

이 중 $\frac{x^2}{2\sigma^2}$는 모든 클래스에 대해 동일하므로 최종적으로는 아래만 비교하여 가장 큰 값을 갖는 클래스로 x를 분류하게 된다.

$\rightarrow \log \pi_k + \frac{\mu_k x}{\sigma^2} - \frac{\mu_k^2}{2\sigma^2}$

최종 판별함수는 아래와 같다.

$\rightarrow \delta_k(x) = x\frac{\mu_k}{\sigma^2} - \frac{\mu_k^2}{2\sigma^2} + \log(\pi_k)$

이 형태가 선형 함수이기 때문에 선형 판별이라 부르는 것임.

LDA의 결정경계

K가 2인 경우 두확률이 동일하게 50%인 지점이 있고 이를 표현하면 아래와 같다.

$\rightarrow \pi_1 = \pi_2 = 0.5$.

$\rightarrow \delta_1(x) = \delta_2(x)$가 되며, 이를 계산하면 $x = (\mu_1 + \mu_2)/2$가 된다.

즉, 결정경계값은 $(\mu_1+\mu_2)/2$임.

LDA의 파라미터(π) 추정

최대우도 측정법보다 간단하게 접근이 가능함.

$\rightarrow \hat{\mu}_k = \frac{1}{n_k} \sum_{i:y_i=k} x_i$

$\hat{\sigma}^2 = \frac{1}{n-K} \sum_{k=1}^{K} \sum_{i:y_i=k} (x_i - \hat{\mu}_k)^2$

$\hat{\pi}_k = \frac{n_k}{n}$

다변수 LDA.

$X = (X_1, \cdots X_p)$이고, 조건을 가지지만 분포함수로 생성되었다고 가정함 가지는 분포는 아래와 같다.

$X \sim N(\mu, \Sigma)$, 평균은 μ, 분산은 Σ임.

$\rightarrow \mu = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_p] \end{bmatrix}$

$\Sigma = E[XX^T] = \begin{bmatrix} E[X_1^2] & \cdots & E[X_1X_p] \\ \vdots & \ddots & \vdots \\ E[X_pX_1] & \cdots & E[X_p^2] \end{bmatrix}$

다변량 가지지만 분포의 PDF는 아래와 같다.

$\rightarrow f(x) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$

다변수 LDA의 판별함수

$\delta_k(x) = x^T\Sigma^{-1}\mu_k - \frac{1}{2}\mu_k^T\Sigma^{-1}\mu_k + \log \pi_k$

마찬가지로 X에 대한 선형함수의 형태임.

다변수 LDA의 결정경계

$\delta_k(x) = \delta_\ell(x)$를 통해 x값을 찾게 되면 상수가 아닌 선형함수의 형태로 드러남.

→ 

→ 데이터 포인트는 조건을 가지지만 분포함수로 같기 때문에 공분산은 다르고 공분산이 동일함.

→ 결정경계는 선의 형태로 드러남.