Topics covered so far:

(1) Bayes decision theory
(2) Minimizing error probability
(3) Minimizing average risk
(4) Discriminant functions & decision surfaces
(5) Bayesian Classification for normal distribution
- 1D Gaussian pdf
- 2D Gaussian pdf (Shape of isocurves etc.)
- ND Gaussian pdf
- Bayesian classification for 2D normal pdf
  - Decision Hyperplanes (Diagonal, Nondiagonal Σ)
  - Minimum distance classifiers
    - Euclidean distance-based
    - Mahalanobis distance-based.

(6) Estimation of Unknown pdfs
- Maximum Likelihood (ML) parameter Estimation
- Maximum a Posteriori Probability (MAP) parameter Estimation

ML parameter Estimation:

$$P(\omega_i|\vec{x}) = \frac{P(\vec{x}|\omega_i)\,P(\omega_i)}{\cancel{P(\vec{x})}}$$

If $p(\omega_i)$ are equal
$\Rightarrow$ Maximum likelihood problem

For different $p(\omega_i)$
$\Rightarrow$ MAP problem

$p(\vec{x}|\omega_i)$

Assume: likelihood pdf is normally distributed.

$\vec{\mu}, \Sigma$

$\{x_1, x_2, \ldots x_N\}$

$$\vec{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2$$

(A1) $p(\vec{x} \mid \omega_i ; \theta_i) \Rightarrow p(\vec{x} ; \theta_i)$

(A2) let $\vec{x}_1, \vec{x}_2, \vec{x}_3, \dots, \vec{x}_N$ be the random samples drawn from $p(\vec{x};\theta)$

let $X = \left[\vec{x}_1, \vec{x}_2, \dots \vec{x}_N \right] \Rightarrow \underline{p(X;\theta)}$

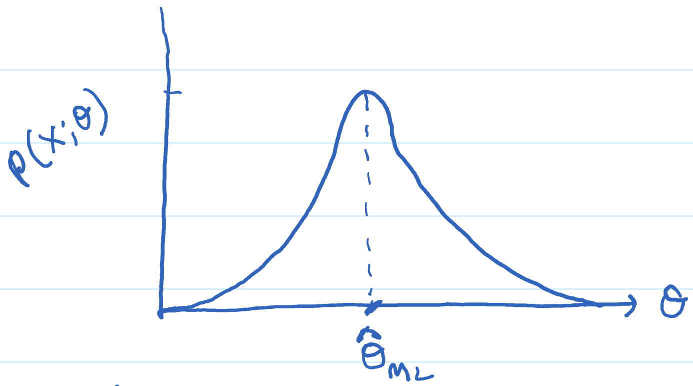
$$p(X;\theta) \triangleq p(\vec{x}_1, \vec{x}_2, \dots \vec{x}_N ; \theta)$$

Assume that all samples are statistically independent

$$\Rightarrow p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N ; \theta) = \prod_{k=1}^{N} p(\vec{x}_k ; \theta)$$

$$\boxed{\hat{\theta}_{ML} = \arg\max_{\theta} \left(\prod_{k=1}^{N} p(\vec{x}_k ; \theta) \right)}$$

let θ: single unknown parameter.



$$\Rightarrow \frac{\partial}{\partial \theta} \left(\prod_{k=1}^{N} p(\vec{x}_k ; \theta) \right) = 0$$

$$\text{let } L(\vec{\theta}) = \ln \left(\prod_{k=1}^{N} p(\vec{x}_k ; \theta) \right)$$

monotonically increasing function.

$$\Rightarrow \frac{dL(\vec{\theta})}{d\vec{\theta}} = \sum_{k=1}^{N} \frac{\partial}{\partial \theta} \left(\ln (p(\vec{x}_k ; \theta)) \right) = 0$$

$$\Rightarrow \boxed{\frac{\partial L(\theta)}{\partial \theta} = \sum_{n=1}^{N} \frac{1}{p(\vec{x}_n;\theta)} \frac{\partial}{\partial \theta}\left(p(\vec{x}_n;\theta)\right) = 0}$$

P1 Let $p(x_n, \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x_n-\mu)^2}{2\sigma^2}\right)$

Compute σ^2_{ML}

P2 Let $p(x_n, \mu) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x_n-\mu)^2}{2\sigma^2}\right)$

Compute μ_{ML}