CSC4008 Homework 2&3

3. Math Proof

(1) $$D^2(X_0|C) = \sum_{X_i \in C} \|X_0 - X_i\|_2^2 = \sum_{X_i \in C} \|X_0 - X_c + X_c - X_i\|_2^2$$

$$= \sum_{X_i \in C} \|X_0 - X_c\|_2^2 + 2\|X_0 - X_c\|_2 \|X_c - X_i\|_2 + \|X_c - X_i\|_2^2$$

$$= \sum_{X_i \in C} \|X_0 - X_c\|_2^2 + \|X_c - X_i\|_2^2$$

$$= nd(X|C) + Var(x)$$

$\therefore$ Q.E.D

(2) the two classifier are equivalent if X is one-dimensional or the covariance matrix of x is $\sigma^2 I$.

1. if x is one-dimensional

Then for each data point: $X \sim N_c(\mu_c, \sigma^2)$

here σ is the pooled variance

$$\therefore f(X = x_0, C) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_0 - x_c)^2\right)$$

$$\mu_c = \frac{1}{N_c} \sum_{X_i \in C} X_i$$

we want the largest $f(x, c)$, which is equivalent to the smallest $x_0 - \mu_c$

For centroid classifier: $d(x_0, c) = \|X_0 - \mu_c\|$

we still want the small $x_0 - \mu_c$

$\therefore$ the two classifiers are equivalent

2. if the covariance matrix of x is $\sigma^2 I$

Then for each data point: $X \sim N_c(\mu_c, \sigma^2 I)$

$$f(X = x_0, C) = \frac{1}{(2\pi)^{\frac{D}{2}} \sigma I} \exp\left(-\frac{1}{2}(x_0 - \mu_c)^T \sigma^2 I (x_0 - \mu_c)\right)$$

$$\mu_c = \frac{1}{N_c} \sum_{X_i \in C} X_i$$

we want the largest $f(x, c)$, which is equivalent to the smallest $x_0 - \mu_c$

For centroid classifier: $d(x_0, c) = \|x_0 - \mu_c\|$

we still want the small $x_0 - \mu_c$

$\therefore$ the two classifiers are equivalent

$\therefore$ Q.E.D.