Comp 254, Image Processing and Analysis
Spring 2002, Prof. Guido Gerig

Solutions to Assignment 1: Statistical Pattern Recognition

Required Readings: Duda and Hart, Chapter 2

I. Theoretical Problems

Problem 1: Normal Density Functions

$$
\begin{aligned}
p(x|\omega_i) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu_i}{\sigma})^2} \\
P(\omega_1) &= P(\omega_2) = \frac{1}{2}
\end{aligned}
$$

The probability of error is the probability that there exists another category than the one that is decided for:

$$
P(\text{error}|x) = \left\{ \begin{array}{l} P(\omega_1|x) \\ P(\omega_2|x) \end{array} \right. \text{ and decision for } \begin{array}{l} \omega_2 \\ \omega_1 \end{array}
$$

where $P(\omega_i|x)$ is obtained using Bayes' rule. A decision with minimum loss is obtained using the symmetric 0-1 loss function, i.e. a decision for class ω_i if $P(\omega_i|x) \geq P(\omega_j|x)$.
Let's look at the case $\mu_1 < \mu_2$. This means with given $p(x|\omega_i)$ and $P(\omega_i)$ that we decide for ω_1 if $x \leq a' = \frac{(\mu_1+\mu_2)}{2}$ and for ω_2 otherwise. The probability of error thus becomes:

$$
\begin{aligned}
P_e &= \int_{-\infty}^{\infty} p(\text{error}, x)dx = \int_{-\infty}^{\infty} P(\text{error}|x)p(x)dx \\
&= \int_{-\infty}^{a'} P(\omega_2|x)p(x)dx + \int_{a'}^{\infty} P(\omega_1|x)p(x)dx \\
&= 2\int_{a'}^{\infty} P(\omega_1|x)p(x)dx \qquad \text{(symmetry)} \\
&= 2\int_{a'}^{\infty} p(x|\omega_1)P(\omega_1)dx \qquad \text{(Bayes)} \\
&= \int_{a'}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu_1}{\sigma})^2} dx
\end{aligned}
$$

The integral can't be solved using elementary calculation. The sought limites, however, can be estimated. We substitute $t = \frac{x-\mu_1}{\sigma}, \quad dx = \sigma dt$ and obtain:

$$
\begin{aligned}
a &= \frac{\mu_2 - \mu_1}{2\sigma} \qquad \text{(new limit)} \\
P_e &= \int_{a}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \leq \frac{1}{\sqrt{2\pi}a} e^{-\frac{a^2}{2}} \to 0, \text{ if } a \to \infty,
\end{aligned}
$$

and thus $\lim_{a\to\infty} P_e = 0$.

Problem 2: Decision with minimum risk

Assumptions:

- c categories $\Omega = \{\omega_1, \ldots \omega_c\}$
- $c+1$ decisions $\mathcal{A} = \{\alpha_i, \ldots \alpha_{c+1}\}$. Decision α_i $(i \leq c)$ means it is assigned to the category ω_i; decision α_{c+1} means rejection.
- Cost matrix:

$$\Lambda = \begin{array}{c|cccc} & \omega_1 & \omega_2 & \ldots & \omega_c \\ \hline \alpha_1 & 0 & \lambda_s & \ldots & \lambda_s \\ \alpha_2 & \lambda_s & 0 & \ldots & \lambda_s \\ \vdots & \vdots & & \ddots & \\ \alpha_c & \lambda_s & \lambda_s & \ldots & 0 \\ \alpha_{c+1} & \lambda_r & \lambda_r & \ldots & \lambda_r \end{array}$$

(a) conditional risk for decisions α_i:

$$\begin{aligned} R(\alpha_i|x) &= \sum_{j=1}^{c} \lambda(\alpha_i|\omega_j) P(\omega_j|x), \quad (\forall i \leq c) \\ &= \sum_{j=1, j\neq i}^{c} \lambda_s P(\omega_j|x) = \lambda_s(1 - P(\omega_i|x)) \\ R(\alpha_{c+1}|x) &= \lambda_r \end{aligned}$$

With the assumption

$$P(\omega_i|x) \geq P(\omega_j|x), \quad (\forall j \leq c)$$

it follows that

$$\lambda_s(1 - P(\omega_i|x)) \leq \lambda_s(1 - P(\omega_j|x))$$

and thus

$$R(\alpha_i|x) \leq R(\alpha_j|x) \quad (\forall j \leq c).$$

It also follows from the minimum risk decision $Min[R(\alpha_i|x), \lambda_r]$ that $P(\omega_i|x) \geq 1 - \frac{\lambda_r}{\lambda_s}$:

$$\begin{aligned} R(\alpha_i|x) &= \lambda_s(1 - P(\omega_i|x)) \\ &\leq \lambda_s(1 - (1 - \frac{\lambda_r}{\lambda_s})) \\ &\leq \lambda_r = R(\alpha_{c+1}|x) \end{aligned}$$

q.e.d.

(b) If $\lambda_r = 0$, we get

$$R(\alpha_i|x) = \lambda_s(1 - P(\omega_i|x)) \geq 0$$
$$Min[R(\alpha_i|x), \lambda_r] = \lambda_r = 0$$

Therefore, we always reject (rejection is cheaper) except in the case where $P(\alpha_i|x) = 1$.

(c) With $\lambda_r > \lambda_s$ we get:

$$R(\alpha_i|x) = \lambda_s(1 - P(\omega_i|x)) \leq \lambda_s < \lambda_r.$$

In this case no rejection occurs. We always decide for one of the categories.

d) Show that the following discriminant functions are optimal

$$g_i(x) = \begin{cases} p(x|w_i)P(w_i) & i = 1, \cdots, c \\ \frac{\lambda_s - \lambda_r}{\lambda_s} \sum_{j=1}^{c} p(x|w_j)P(wj) & i = c+1 \end{cases}$$

see solution 2a:

$$\begin{aligned} R(\alpha_i|x) &= \lambda_s(1 - P(\omega_i|x)) \\ R(\alpha_{c+1}|x) &= \lambda_r \end{aligned}$$

then, replace $P(\omega_i|x)$ by $p(x|w_i)P(w_i)/p(x)$ using the Bayes rule. Multiplying both equations by $p(x)$ you get (replacing risk R by monotonic functions $G = f(R)$):

$$\begin{aligned} G^1(\alpha_i|x) &= \lambda_s p(x) - \lambda_s p(x|w_i)P(w_i) \\ G^1(\alpha_{c+1}|x) &= \lambda_r p(x) \end{aligned}$$

subtracting $\lambda_s p(x)$ from both equations and dividing by λ_s yields:

$$\begin{aligned} G^2(\alpha_i|x) &= -p(x|w_i)P(w_i) \\ G^2(\alpha_{c+1}|x) &= -(\frac{\lambda_s - \lambda_r}{\lambda_s})p(x) \end{aligned}$$

Switching the sign, these functions can be used as discriminant functions for a minimum risk decision using the MAX decision rule.

e) Sketch discriminant functions and the decision regions

Use the results obtained in a) to d) to sketch the functions.