Chapter 2 problem 4

Consider the minimax criterion for a two-category classification problem.

(a) Fill in the steps of the derivation of Eq. 23.

(f) Assume $p(x|\omega_1) \sim N(5, 1)$ and $p(x|\omega_2) \sim N(6, 1)$. Without performing any explicit calculations, determine $x^*$ for the minimax criterion. Explain your reasoning.

Problem 4 refers to the following section from section 2.3.1 of the text, dealing with a two-class classifier:

...we let $R_1$ denote that (as yet unknown) region in feature space where the classifier decides $\omega_1$ and likewise for $R_2$ and $\omega_2$, and then we write our overall risk Eq. 12 in terms of conditional risks:

$$ R = \int_{R_1} [\lambda_{11} p(\omega_1)p(x|\omega_1) + \lambda_{12} p(\omega_2)p(x|\omega_2)] \, dx $$

$$ + \int_{R_2} [\lambda_{21} p(\omega_1)p(x|\omega_1) + \lambda_{22} p(\omega_2)p(x|\omega_2)] \, dx. \quad (22) $$

We use the fact that $P(\omega_2) = 1 - P(\omega_1)$ and that $\int_{R_1} p(x|\omega_1) \, dx = 1 - \int_{R_2} p(x|\omega_1) \, dx$ to rewrite the risk as:

$$ R(P(\omega_1)) = \lambda_{22} + (\lambda_{12} - \lambda_{22}) \int_{R_1} p(x|\omega_1) \, dx $$

$$ + P(\omega_1) \left[ (\lambda_{11} - \lambda_{22}) + (\lambda_{21} - \lambda_{11}) \int_{R_2} p(x|\omega_1) \, dx - (\lambda_{12} - \lambda_{22}) \int_{R_1} p(x|\omega_2) \, dx \right]. \quad (23) $$

Chapter 2 problem 11

Suppose that we replace the deterministic decision function $\alpha(x)$ with a randomized rule, namely, one giving the probability $P(\alpha_i|x)$ of taking action $\alpha_i$ upon observing $x$.

(a) Show that the resulting risk is given by

$$ R = \int \left[ \sum_{i=1}^{n} R(\alpha_i|x) P(\alpha_i|x) \right] p(x) \, dx. $$

(b) In addition, show that $R$ is minimized by choosing $P(\alpha_i|x) = 1$ for the action $\alpha_i$ associated with the minimum conditional risk $R(\alpha_i|x)$, thereby showing that no benefit can be gained from randomizing the best decision rule.

(c) Can we benefit from randomizing a suboptimal rule? Explain.

Chapter 2 problem 28

Two random variables $x$ and $y$ are called statistically independent if $p(x,y|\omega) = p(x|\omega)p(y|\omega)$.

(a) Prove that if $x_i - \mu_i$ and $x_j - \mu_j$ are statistically independent (for $i \neq j$), the $\sigma_{ij}$ as defined in Eq. 43 is 0.

(b) Prove that the converse is true for the Gaussian case.

Equation 43 is $\sigma_{ij} = \mathcal{E}[(x_i - \mu_i)(x_j - \mu_j)].$