Classification based on discriminant functions.

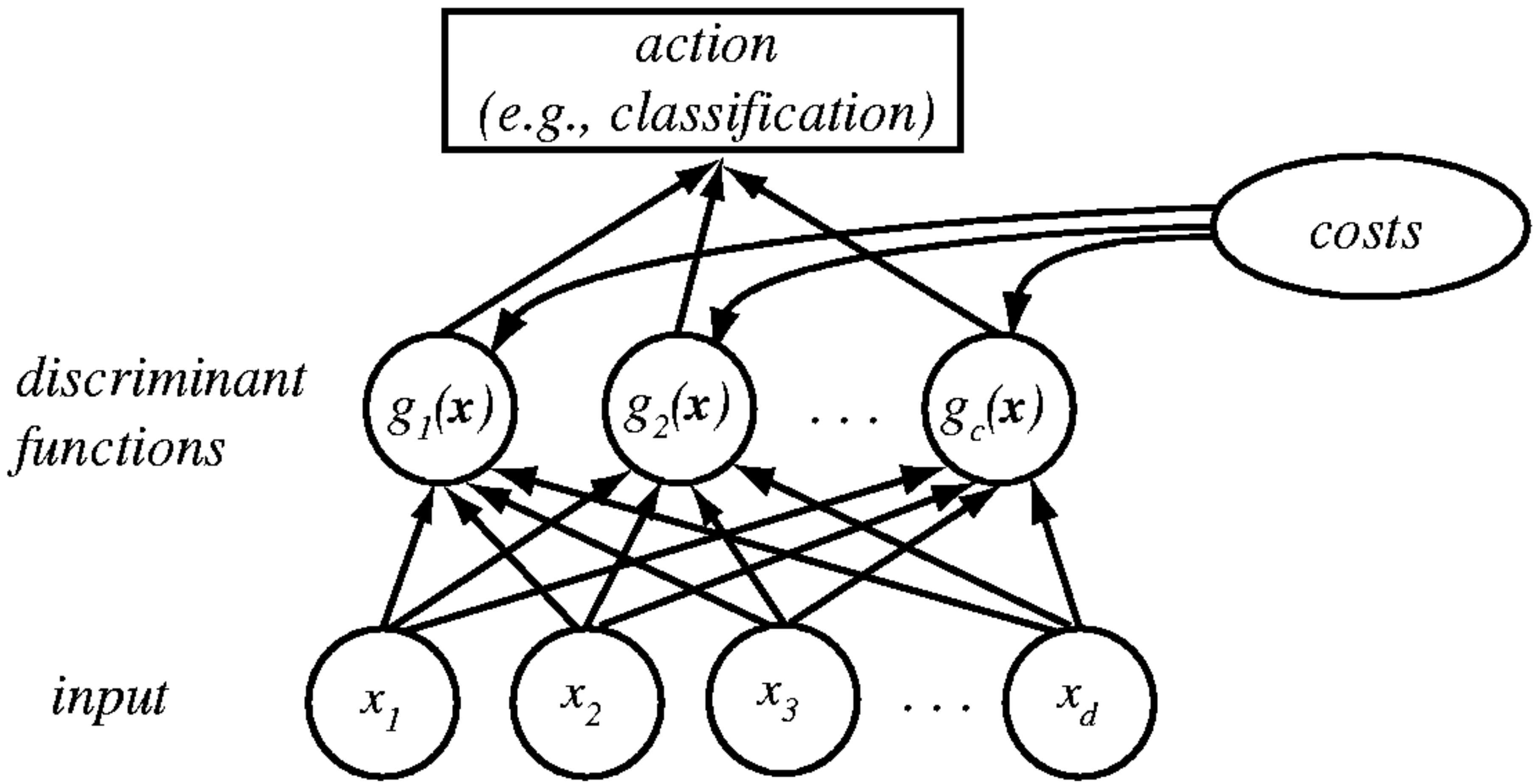
Compute values of discriminant functions $g_i(x)$ and assign to x the class corresponding to the discriminant function with the largest value.

The classifier assigns a feature vector x to class ω_i if

$$g_i(\mathrm{x}) > g_j(\mathrm{x}), \quad \text{for all } j \neq i$$

where $g_i(\mathrm{x})$ is the set of discriminant functions.

General structure of a classifier



from Duda, Hart, Stork (2001) Pattern classification

Examples of discriminant functions

Risks:

$$g_i(\mathrm{x}) = -R(\alpha_i \mid \mathrm{x})$$

Posterior probabilities:

$$g_i(\mathrm{x}) = P(\omega_i \mid \mathrm{x})$$

Freedom to change discriminant functions

The choice of discriminant functions is not unique.

In general: If every discriminant function $g_i(\mathrm{x})$ is replaced by a monotonically increasing function $f(g_i(\mathrm{x}))$, the classification result does not change.

Examples of changing of discriminant functions

Motivation: Other quantities, simpler to understand or to compute, lead to identical classification results.

$$g_i(\mathrm{x}) = P(\omega_\mathrm{i} \mid \mathrm{x}) = \frac{p(\mathrm{x} \mid \omega_i) P(\omega_i)}{\sum_{j=1}^{c} p(\mathrm{x} \mid \omega_i) P(\omega_i)}$$

$$g'_i(\mathrm{x}) = p(\mathrm{x} \mid \omega_i) P(\omega_i)$$

$$f(g'_i(\mathrm{x})) = \ln p(\mathrm{x} \mid \omega_i) + \ln P(\omega_i)$$

(ln is useful in the important case of normal distribution)

Dichotomizer

In a two-class problem, a ‘dichotomizer’ is defined

$$g(\mathrm{x}) = g_1(\mathrm{x}) - g_2(\mathrm{x})$$

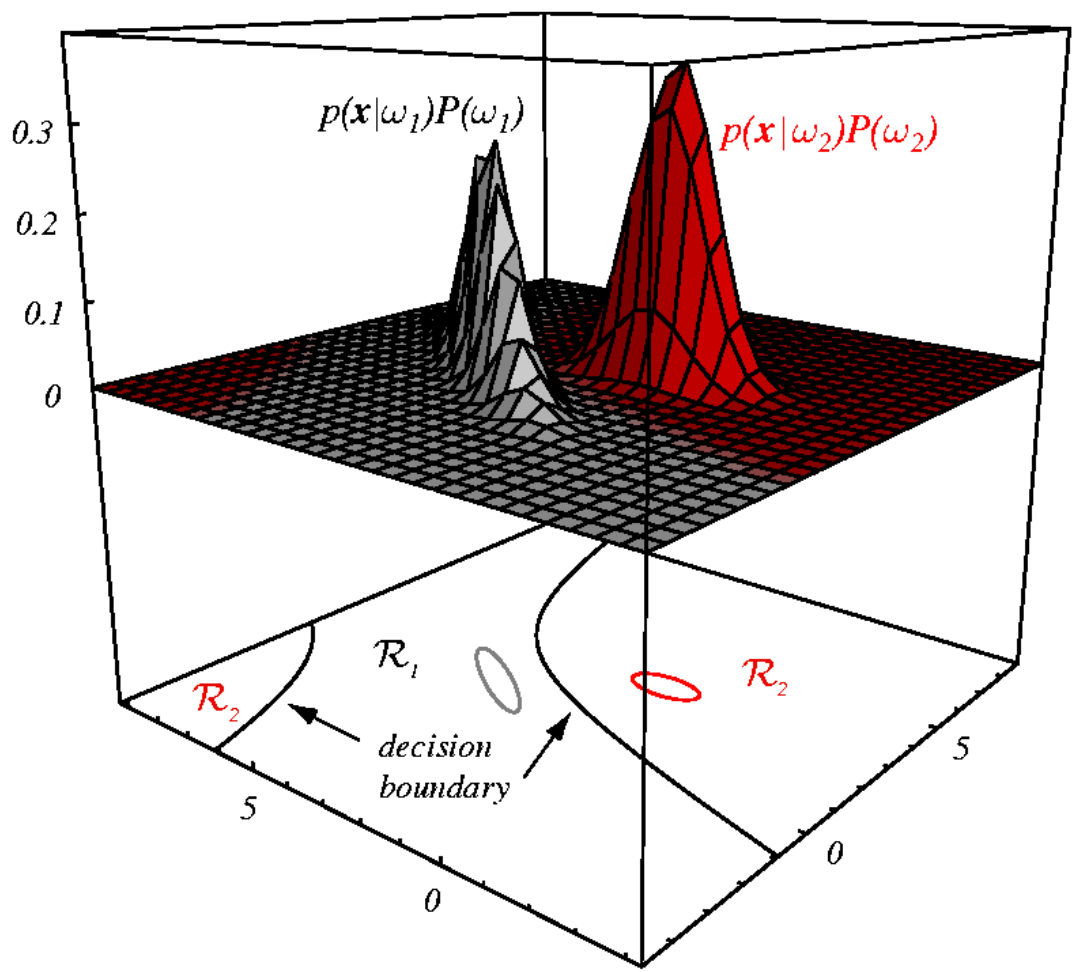
$$Decide\ \omega_1\ \ if\ \ g(\mathrm{x}), > 0\ and\ \ \omega_2\ otherwise.$$

Alternative (sometimes more convenient) forms:

$$g(\mathrm{x}) = P(\omega_1 \mid \mathrm{x}) - P(\omega_2 \mid \mathrm{x})$$

$$g(\mathrm{x}) = \ln \frac{p(\mathrm{x} \mid \omega_1)}{p(\mathrm{x} \mid \omega_2)} + \ln \frac{P(\omega_1)}{P(\omega_2)}$$

Decision boundary for a dichtomizer



from Duda, Hart, Stork (2001) Pattern classification