Computing decision boundaries

Wagner L. Truppel

January 2001

Abstract

We derive the functional form of the decision boundary that separates two classes.

For simplicity, we'll abuse of the mathematical notation in two ways: vector symbols are dropped, and no distinction is made between column vectors and row vectors.

Let $x = (x_1, x_2, \ldots, x_d)$ be a feature vector and c_i denote a class $(1 \leq i \leq m)$. The *decision boundary* between classes c_i and c_j (with $i \neq j$) is the locus of vectors having equal probabilities of being assigned to either of those classes, that is, the locus of vectors x_{ij} such that $p(c_i|x_{ij}, M) = p(c_j|x_{ij}, M)$. Using Bayes' rule, this implies the equality

$$p(x_{ij}|c_i, M)\, p(c_i|M) = p(x_{ij}|c_j, M)\, p(c_j|M)\,,$$

where all probabilities are conditioned on the model M we're using. In order to simplify the notation, we'll drop the ij subscript, since all the quantities to follow are dependent on the particular choice of classes. Since the model assumes multivariate normal distributions for each class,

$$\begin{aligned} p(x|c_i, M) &= \left[\frac{|\Sigma_i^{-1}|}{(2\pi)^d} \right]^{1/2} \exp\left[-\frac{1}{2}\, r_i^2(x) \right], \quad \text{with} \\ r_i^2(x) &= (x - \mu_i) \cdot \Sigma_i^{-1} \cdot (x - \mu_i)\,, \end{aligned}$$

the equation defining the decision boundary can be written as

$$r_i^2(x) - r_j^2(x) = 2\,\ln \left[\frac{p(c_i|M)\, |\Sigma_i^{-1}|^{1/2}}{p(c_j|M)\, |\Sigma_j^{-1}|^{1/2}} \right] \equiv \xi\,. \tag{1}$$

Clearly, ξ is just a number that can be computed from the class data. Allowing for the symmetric nature of the covariance matrix Σ, the left–hand–side of this equation can be rewritten as

$$\begin{aligned} r_i^2(x) - r_j^2(x) &\;=\; x \cdot \mathsf{M} \cdot x - 2\,v \cdot x + \delta\,, \quad \text{with} \\ \mathsf{M} &\;\equiv\; \Sigma_i^{-1} - \Sigma_j^{-1}\,, \\ v &\;\equiv\; \left(\mu_i \cdot \Sigma_i^{-1}\right) - \left(\mu_j \cdot \Sigma_j^{-1}\right), \quad \text{and} \\ \delta &\;\equiv\; \left(\mu_i \cdot \Sigma_i^{-1} \cdot \mu_i\right) - \left(\mu_j \cdot \Sigma_j^{-1} \cdot \mu_j\right). \end{aligned} \tag{2}$$

Note that M is a symmetric matrix, v a vector, and δ a number, all computable from the class data. The right–hand–side is a quadratic form on x with a symmetric real matrix and, thus, can be diagonalized through an orthogonal transformation, *even if* M *is singular*. However, when M is singular, the path to the solution is a little more complicated, so for now we'll assume that M is invertible. Defining the vectors $u = \mathsf{M}^{-1} \cdot v$ and $y = x - u$, we find

$$r_i^2(x) - r_j^2(x) = \left(y \cdot \mathsf{M} \cdot y\right) - \left(v \cdot \mathsf{M}^{-1} \cdot v\right) + \delta\,.$$

Combining this result with that of Eq. (**??**), we get $\left(y \cdot \mathsf{M} \cdot y\right) = s$, where

$$s \equiv \left(v \cdot \mathsf{M}^{-1} \cdot v\right) + \left(\xi - \delta\right).$$

The final step in the process of diagonalizing the quadratic form in question is to diagonalize M itself. Transforming y into z through $y = \mathsf{R} \cdot z$, where R is the orthogonal matrix diagonalizing M (that is, $\mathsf{D} \equiv \tilde{\mathsf{R}} \cdot \mathsf{M} \cdot \mathsf{R}$ is diagonal, and also $\tilde{\mathsf{R}} = \mathsf{R}^{-1}$), we obtain

$$\sum_{i=1}^{d} \mathsf{D}_{ii}\, z_i^2 = s\,, \tag{3}$$

which describes a multi–dimensional conic section in z-space. In terms of the original x-space, the points on this hypersurface will have coordinates satisfying $x = \mathsf{R} \cdot z + \mathsf{M}^{-1} \cdot v$.

The actual hyperconic described by Eq. (**??**) depends on the signs of s and of M's eigenvalues. In the specific case of two dimensions, the possible solutions are spelled out in the table below. Sign conditions not listed in the table correspond to situations where no solutions exist, or to situations where M is a singular matrix.

sign of s	sign of D_{11}	sign of D_{22}	result in (z_1, z_2) space
neg	neg	neg	ellipse:
pos	pos	pos	$\lvert\mathsf{D}_{11}\rvert\, z_1^2 + \lvert\mathsf{D}_{22}\rvert\, z_2^2 = \lvert s\rvert$
neg	neg	pos	hyperbola:
pos	pos	neg	$\lvert\mathsf{D}_{11}\rvert\, z_1^2 - \lvert\mathsf{D}_{22}\rvert\, z_2^2 = \lvert s\rvert$
neg	pos	neg	hyperbola:
pos	neg	pos	$\lvert\mathsf{D}_{22}\rvert\, z_2^2 - \lvert\mathsf{D}_{11}\rvert\, z_1^2 = \lvert s\rvert$
zero	neg	pos	slanted lines:
zero	pos	neg	$z_2 = \pm\,\lvert\mathsf{D}_{11}/\mathsf{D}_{22}\rvert^{1/2}\, z_1$
zero	neg	neg	point (origin):
zero	pos	pos	$z_1 = z_2 = 0$

Let's now look at the cases when M is a singular matrix. Again, let R be the orthogonal matrix diagonalizing M. Then, $\mathsf{M} = \mathsf{R}\cdot\mathsf{D}\cdot\tilde{\mathsf{R}}$, and the first line of Eq. (**??**) now reads:

$$r_i^2(x) - r_j^2(x) = x\cdot\mathsf{R}\cdot\mathsf{D}\cdot\tilde{\mathsf{R}}\cdot x - 2\,v\cdot x + \delta = \xi\,,$$

which can be rewritten as

$$\begin{aligned} s &= z\cdot\mathsf{D}\cdot z - 2\,u\cdot z\,, \quad \text{with} \\ z &\equiv \tilde{\mathsf{R}}\cdot x\,, \quad u \equiv v\cdot\mathsf{R}\,, \quad \text{and} \quad s \equiv (\xi - \delta)\,. \end{aligned}$$

In the particular case of two dimensions, this expands into

$$\mathsf{D}_{11}\, z_1^2 + \mathsf{D}_{22}\, z_2^2 - 2\,u_1\, z_1 - 2\,u_2\, z_2 = s\,. \tag{4}$$

There are two cases to consider now: only one of the eigenvalues of M is zero, or both are zero. In either case, we also have to take into account whether or not u_1 and u_2 vanish, as well as the signs of the various quantities. These possibilities produce the table below:

D_{11}	D_{22}	u_1	u_2	s	result in (z_1, z_2) space
zero	zero	zero	non-zero	any	horizontal line: $z_2 = -s/(2\,u_2)$
zero	zero	non-zero	zero	any	vertical line: $z_1 = -s/(2\,u_1)$
zero	zero	non-zero	non-zero	any	slanted line: $z_2 = -(s + 2\,u_1 z_1)/(2\,u_2)$
zero	non-zero	zero	zero	zero	line (z_1 axis): $z_2 = 0$
non-zero	zero	zero	zero	zero	line (z_2 axis): $z_1 = 0$
zero	pos	zero	zero	pos	horizontal lines: $z_2 = \pm\,\lvert s/\mathsf{D}_{22}\rvert^{1/2}$
zero	neg	zero	zero	neg	
pos	zero	zero	zero	pos	vertical lines: $z_1 = \pm\,\lvert s/\mathsf{D}_{11}\rvert^{1/2}$
neg	zero	zero	zero	neg	
zero	non-zero	zero	non-zero	$\geq -\dfrac{u_2^2}{\mathsf{D}_{22}}$	horizontal lines: $z_2 = (u_2 \pm \sqrt{u_2^2 + s\,\mathsf{D}_{22}})/\mathsf{D}_{22}$
non-zero	zero	non-zero	zero	$\geq -\dfrac{u_1^2}{\mathsf{D}_{11}}$	vertical lines: $z_1 = (u_1 \pm \sqrt{u_1^2 + s\,\mathsf{D}_{11}})/\mathsf{D}_{11}$
non-zero	zero	non-zero	non-zero	any	parabola: $z_2 = (\mathsf{D}_{11}\,z_1^2 - 2\,u_1\,z_1 - s)/(2\,u_2)$
zero	non-zero	non-zero	non-zero	any	parabola: $z_1 = (\mathsf{D}_{22}\,z_2^2 - 2\,u_2\,z_2 - s)/(2\,u_1)$

■