

Statistical Methods of Machine Learning

Assignment 2

Martin Grunbaum
Martin Jorgensen
Thomas Barnholdt

February 28, 2014

II.1

II.1.1

Linear Discriminant Analysis with $m > 2$ classes is done by creating a discriminant function for each class and running each function on all points we wish to classify. The discrimination functions calculates the posterior propability that a point belongs to its class, and so the one with the highest propability is chosen. The discrimination function for a class k looks like this:

$$\delta_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \ln \Pr(Y = C_k)$$

The prior distribution is calculated like so

$$\Pr(Y = C_k) = \ell_k / \ell$$

Where ℓ is the number of elements in the training data, and ℓ_k is the number of elements of class k in the training data. μ_k is the mean of a given class, while Σ as the covariance matrix for each class added together and normalized. They're calculated like so:

$$\mu_k = \frac{1}{\ell_k} \sum_{(x,y) \in S_k} x$$
$$\Sigma = \frac{1}{\ell - m} \sum_{k=1}^m \sum_{(x,y) \in S_k} (x - \mu_k)(x - \mu_k)^T$$

Where S_k is the points in the training set that corresponds to the class k and m is the number of classes.

Using this method we managed to get an XXX% accuracy on the training data and XXX% for the test data.

References