CS771: Machine learning: tools, techniques, applications
Assignment #3: Parametric methods and discriminant functions

Due by: 27-2-2015, 23.00 8-2-2015
MM: 205

1. Suppose the $d-$sized vector $\mathbf{x}$ in a population with two classes is normally distributed as: $\mathbf{x} \sim \mathcal{N}(\mu, \Sigma_j),\ j = 1, 2$ where $\Sigma_j = \sigma_j^2[(1-\rho_j)I + \rho_j \mathbf{1}\mathbf{1}^{\mathbf{T}}]$, where $\mathbf{1}$ is a vector of all 1s and I is the identity matrix. Show that the Bayes discriminant function is given by upto within a constant by:
$-\frac{1}{2}(c_{11} - c_{12})b_1 + \frac{1}{2}(c_{21} - c_{22})b_2$, where
$b_1 = (\mathbf{x} - \mu)(\mathbf{x} - \mu)^T$
$b_2 = (\mathbf{1}^T(\mathbf{x} - \mu))^2$
$c_{1j} = [\sigma_j^2(1-\rho_j)]^{-1}$
$c_{2j} = \rho_j[\sigma_j^2(1-\rho_j)(1+(d-1)\rho_j)]^{-1}$

[25]

2. Let $p(x)$ be modelled as mixture $p(x) = \sum_{j=1}^{g} \pi_j p(x|\mu_j, m_j)$, where $p(x|\mu, m)$ is given by a gamma distribution with mean μ and order parameter m as follows:

$$p(x|\mu, m) = \frac{m}{(m-1)!\mu}\Big(\frac{mx}{\mu}\Big)^{m-1} exp\big(-\frac{mx}{\mu}\big)$$

Derive the EM parameter update equations for π_j, μ_j and m_j.
(Note: Many parametrizations of the gamma distribution are available. This one is used in radar based ship detection.)

[30]

3. Generate three datasets (train, validation and test sets) for the three-class, 21 variable, waveform data (Breiman et al. Classification and Regression Trees, 1984).

$$x_i = uh_1(i) + (1-u)h_2(i) + \epsilon_i \quad \text{(class } \omega_1)$$
$$x_i = uh_1(i) + (1-u)h_3(i) + \epsilon_i \quad \text{(class } \omega_2)$$
$$x_i = uh_2(i) + (1-u)h_3(i) + \epsilon_i \quad \text{(class } \omega_3)$$

where $i = 1 \ldots 21$; u is uniformly distributed on $[0, 1]$; ϵ_i are normally distributed with zero mean and unit variance; and the h_i are shifted triangular waveforms:

$$h_1(i) = max(6 - |i - 11|, 0)$$
$$h_2(i) = h_1(i-4)$$
$$h_3(i) = h_1(i+4)$$

Assume equal class priors.

(a) Construct a three-component mixture model for each class using a common covariance matrix across components and classes. Investigate different values (systematically) for the means and covariance matrix and choose a model based on the validation set error rate.

For this model, evaluate the classification error on the test set.

(b) Compare the results in 3a) with a linear discriminant classifier and a quadratic discriminant classifier constructed using the training set and evaluated on the test set.

[30 (data set gen.)+60+(30+30)=150]