ECE 09.455 / 09.555
Theory & Applications / Advanced Topics in Pattern Recognition

Homework 2

Due: Sep 25, 2009

Most extremely very ultra mega useful hint: Work as groups (of no more then 2), and start early. Typical completion time for this homework is about a week. Considering that you have other classes, you are given two weeks to complete. Have each member work on two or so questions, and then combine your results. Make sure that every member of the group understands the solution of every question, since these questions may and will appear in the quizzes and exams, and I may randomly ask anyone to solve a typical question on the board.

You will NOT get it done, if you start late!

1. The following will be useful throughout the semester
   a. Write a general function (you may use Matlab's build in commands) to generate random samples from N(**μ,Σ**) in d-dimensions.
   b. Write a procedure of the discriminant of the following form
   $$g_i(\mathbf{x}) = -\frac{1}{2}\left[(\mathbf{x}-\mu)^T \cdot \Sigma_i^{-1} \cdot (\mathbf{x}-\mu)\right] - \frac{d}{2}\ln 2\pi - \frac{1}{2}\ln\left|\Sigma_i\right| + \ln P(\omega_i)$$
   c. Write a procedure for computing the Euclidean distance between two vectors of d-dimensions
   d. Write a procedure for computing the Mahalanobis distance between a point **x** and some mean vector **μ**, given a covariance matrix **Σ.**
2. Given the normally distributed data in the table below,
   a. Design a classifier using only the first feature x_1, if the prior probabilities are $P(\omega_1)= P(\omega_2)=1/2$, $P(\omega_3)=0$.
   b. What is the empirical error?
   c. Repeat the above process for two features, x_1 and x_2.
   d. Repeat for all three features.
   e. Repeat for all classes (with equal priors) using all features.
3. Again, for the same data, and equal priors,
   a. compute the Mahalanobis distances between the following points and all category means: $(1,2,1)^T$, $(5,3,2)^T$, $(0, 0, 0)^T$, $(1,0,0)^T$,
   b. Classify these points
   c. What happens if the priors are now $P(\omega_2)= P(\omega_3)=0.1$, $P(\omega_1)=0.8$

4. Generate a multi-class (three or four class) artificial dataset with various Gaussian distributions (of each of the three types) and demonstrate the Bayesian classifier. Calculate its theoretical and empirical error. Choose examples that demonstrate your understanding. Show decision boundaries by plotting the classifier decisions on a grid that encompasses the entire feature space.

5. Implement the naïve Bayes classifier from scratch and then compare your results to that of Matlab's built-in implementation. Generate graphs similar to those shown in the lecture slides. Use different means, covariance matrices, prior probabilities (indicated by relative data size for each class) to demonstrate that your implementations are correct. Specifically, use different distributions to generate the data and then use the native distribution, as well as the Gaussian to implement the naïve Bayes. How tolerant is Gaussian based naïve Bayes to incorrect selection of distribution? How about the violations of the independence?

6. For grad students only: Read / learn and implement Bayesian Networks. **(floating deadline – mid October)**

Duda, Hart & Stork, Pattern Classification, 2/e Wiley, 2000

	ω_1			ω_2			ω_3		
sample	x_1	x_2	x_3	x_1	x_2	x_3	x_1	x_2	x_3
1	−5.01	−8.12	−3.68	−0.91	−0.18	−0.05	5.35	2.26	8.13
2	−5.43	−3.48	−3.54	1.30	−2.06	−3.53	5.12	3.22	−2.66
3	1.08	−5.52	1.66	−7.75	−4.54	−0.95	−1.34	−5.31	−9.87
4	0.86	−3.78	−4.11	−5.47	0.50	3.92	4.48	3.42	5.19
5	−2.67	0.63	7.39	6.14	5.72	−4.85	7.11	2.39	9.21
6	4.94	3.29	2.08	3.60	1.26	4.36	7.17	4.33	−0.98
7	−2.51	2.09	−2.59	5.37	−4.63	−3.65	5.75	3.97	6.65
8	−2.25	−2.13	−6.94	7.18	1.46	−6.66	0.77	0.27	2.41
9	5.56	2.86	−2.26	−7.39	1.17	6.30	0.90	−0.43	−8.71
10	1.03	−3.33	4.33	−7.50	−6.32	−0.31	3.52	−0.36	6.43