**CS 520 — Artificial Intelligence**
Written Exercises #1
**Out: Feb 12, 2001**
Due: Feb 21, 2001

Problem 0. I continue to encourage you to implement the algorithms for OPTIMUM and EXECUTE from the decision models notes, in the programming language of your choice. Meanwhile, here are some quick exercises to help think about the material of the last couple lectures. All questions can be answered with a sketch and/or at most two sentences of English explanation.

Problem 1. In class, and on page 25 of Bishop, a diagram of the following sort is used to illustrate the probability of error for Bayesian classifier deciding between $C1$ and $C2$ on the basis of a noisy measurement x.
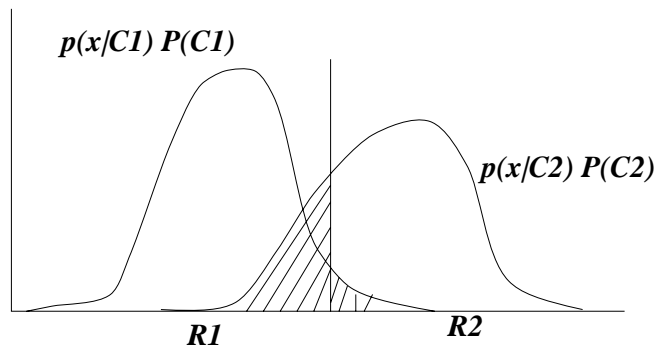

By contrast, show a corresponding overlaid graph of $P(C1|x)$ and $P(C2|x)$. What property of a graph of $P(C1|x)$ and $P(C2|x)$ makes it uninformative of the probability of error of the classifier?

Problem 2. We've seen that a Naive Bayes model assigns log probability of class membership proportional to the distance above a plane corresponding to the class; the normal to the plane is a weight vector, as in the diagram below (where the box represents a feature space, weight vector $w1$ corresponds to class $C1$ and weight vector $w2$ corresponds to class $C2$).
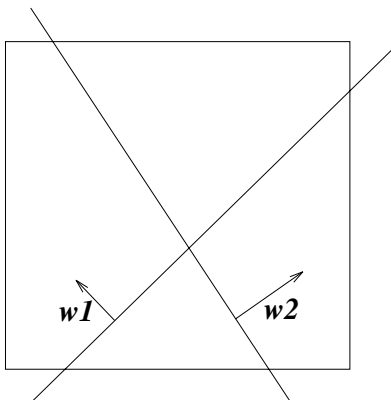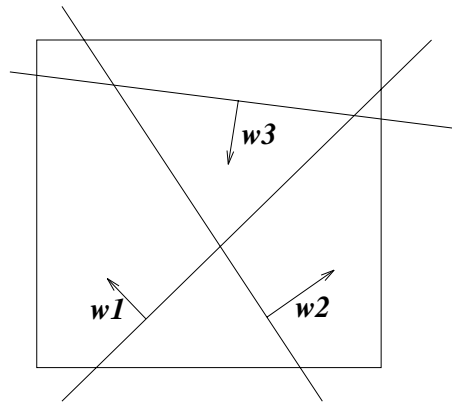

Use a diagram (and an explanatory sentence) to indicate geometrically how these planes determine a plane decision boundary for a Naive Bayes classifier (with decision regions on either side).

Problem 3. Now do the same for this case.
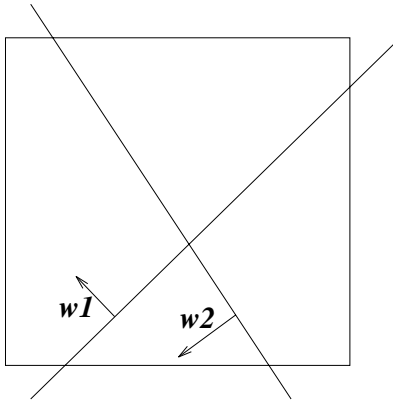
That is, use a diagram to indicate geometrically how these planes determine a plane decision boundary for a Naive Bayes classifier (with decision regions on either side). What is different?

Problem 4. Now consider a three-way classification problem, as schematized by this feature-space diagram:

As in problem 2, each PAIR of classes determines a plane in this feature space. What do the regions on either side of these planes correspond to?

Use a diagram (and an explanatory sentence) to indicate geometrically how these planes determine the decision boundaries and decision regions for a Naive Bayes classifier.

Problem 5. It turns out that—ignoring the labels for the features and the labels of the classes—there are only six qualitatively different rules possible for splitting the eight possible observations of three binary features into two categories each covering four members. The possibilities are shown below.
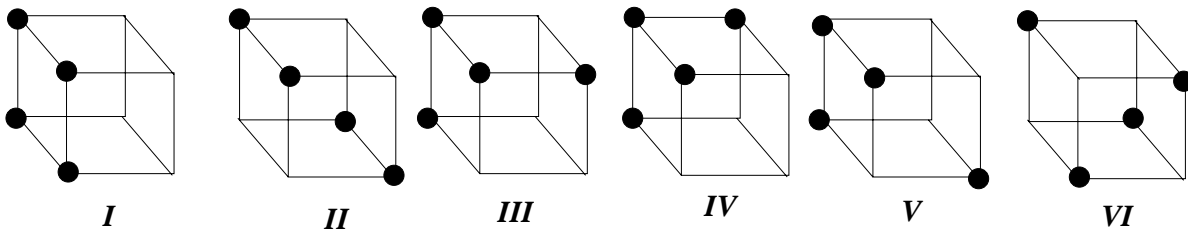
In which cases might the feature values be conditionally independent, given the class?