Pattern Recognition with Applications to Biomedical Images
Independent Study in Mathematics – CSUN Spring 2006
Module 4 – The 'brute force' approach

P. Perona & K. F. Stevenson

January 22, 2007

Consider the linear classifiers you used in the previous weeks. As you may remember, you trained by hand. Could you have automated the process of selecting the parameters $\mathbf{w}$?

The 'brute force' approach is simple:

1. Sample systematically a large number of meaningful decision boundaries parametrized by $\mathbf{w}$
2. For each, measure the classification errors on the training set, e.g. $\epsilon(\mathbf{w}) = \sum_{k,i} \delta(y(x_i; \mathbf{w}), y_i)$
3. Pick the best classifier: $\mathbf{w}^* = \arg\min_{\mathbf{w}} \epsilon(\mathbf{w})$

In order to sample the space of all linear boundaries in 2D you may wish to place the origin in the center of the data, scale the axes so that the data are contained in the $[-11]^2$ square, and pick 16 regularly spaced orientations in $[02\pi]$, and for each orientation pick 16 regularly spaced distances from the origin in the $[-11]$ interval. Notice that this parametrization of lines in the plane is different from the one you used last week, so you will have to convert the parameters between one and the other to re-use your old code (how come you had 3 parameters and now you have 2? Comment on this).

Implement such a brute force approach and compare the results with what you obtained by selecting classifiers by hand. Did you do better than the machine?

Does the 'brute force' approach work in general? How many samples would you have needed if your feature space was 16-dimensional and you had wished to explore 10 samples per parameter? What if the feature space had been 100-dimensional?