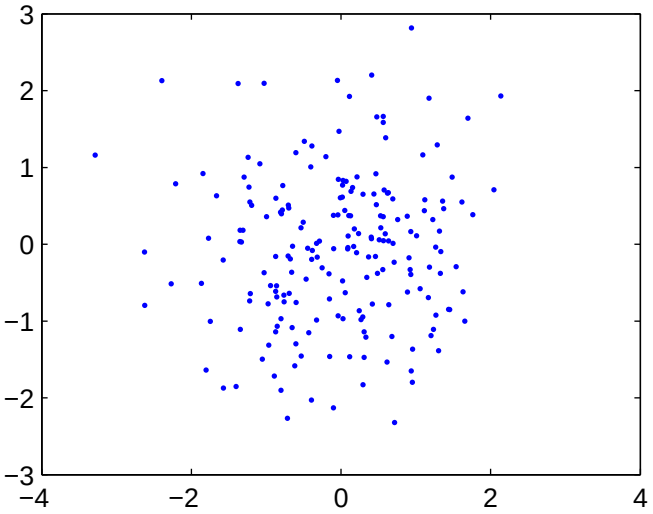
T-61.5010 Information visualization

Problem set 5. Tue, 24th Feb., 2009, 10-12 T2

This problem sheet has two (2) pages.

1. Consider the 2 dimensional data presented in the figure below (uncorrelated gaussian noise with $\mu = 0$ and $\sigma = 1$).



   How would you project this cloud of points to one dimension (onto a line, that is) such that

   (a) precision

   (b) recall

   is preserved? Is it possible to create a projection where both measures are good?

2. Explain the technique of *Manifold embedding* in terms understandable to your grandmother.

3. Take the Iris data set of the previous exercises and study its 2-dimensional PCA projection. Compare this to the projection you obtain by using some other method of your choice, e.g., Sammon's Mapping, Laplacian Eigenmap or other.

4. Select one of the previous lectures and come up with *three* suitable examination questions related to its contents. The first question should be about

briefly explaining 4-6 new concepts appearing in the lecture. The second question should be of the form "Write a 1-2 page essay about XYZ.", where XYZ is some of the broader topics the lecture is about. The final question should be similar to the ones discussed in the weekly problem sessions. Some of your questions might appear in the examination! To get the *bonus point* for this problem, you need to *hand in your answer*! (i.e. the sample questions).