T-61.5010 Information visualization

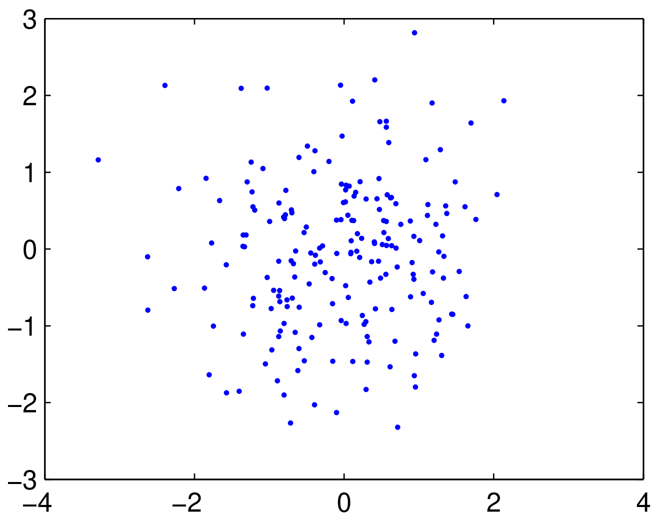
Exercise 5. Mon Feb. 27, 2006, 10-12 T2

- Dimensionality reduction methods

1. Pick one of the nontrivial MDS models (see slide 30 of the lecture on Feb. 21.) and write out the complete stress function to optimize. Explain how the model you chose affects the final visualization. What situations would benefit from the chosen model?

2. Download the iris dataset from the course website (`http://www.cis.hut.fi/Opinnot/T-61.5010/2006/problems.shtml`) and study its 2-dimensional PCA projection with and without removing the mean from the data when calculating the covariance matrix. Note that Matlabs `cov` function subtracts the mean automatically!

3. 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 such that

   a) continuity

   b) trustworthiness

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

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 will appear in the examination! You can get *bonus points* for this problem *only by handing in your answer*!