T-61.3050 PROBLEMS 6/2007

In T1 on 19 October 2007 at 10 o'clock.

You should solve the problems before the problem session and give the solved problems to the assistant. Please write clearly and leave a wide (left or right) margin. The solutions should be stapled together **with a cover sheet** containing your name, student number and the numbers of problems you have solved.

For the problems where a "correct" solution exists (math and algorithm questions) the assistant will present one possible solution during the session. In some cases the questions do not have a single correct answer, but the idea is that you think about the problem and are prepared to discuss it with the assistant and other students during the session.

See `http://www.cis.hut.fi/Opinnot/T-61.3050/2007/problems` for up-to-date information of the problem sessions.

This problem sheet has two pages.

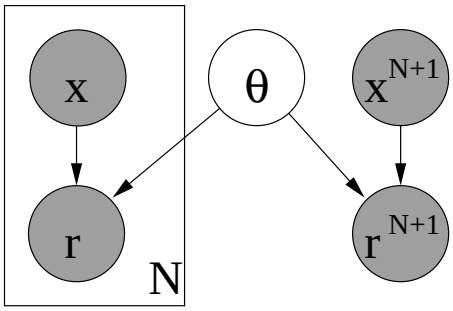

Figure 1: See exercise 1.

1. (a) Write down the joint distribution corresponding to the model in Figure 1.

   (b) Using the sum and product rules, derive an expression for $p(r^{N+1} \mid x^{N+1}, \mathcal{X})$ from the joint distribution, where the training data is given by $\mathcal{X} = \{(r^t, x^t)\}_{t=1}^N$.

   (c) Write down the log-posterior (used in the MAP estimate) using Gaussian likelihood and prior, $p(r \mid x, \theta) = N(g(x \mid \overline{w}), \sigma^2)$, $p(\theta) = \prod_{i=0}^k p(w_i)$ and $p(w_i) = N(0, 1/\lambda)$, where $g(x \mid \overline{w})$ is a polynomial regressor, $g(x \mid \overline{w}) = \sum_{i=0}^k w_i x^i$, and the parameters θ are given by $\theta = (w_0, \ldots, w_k, \sigma^2)$.

2. Use the data set for Problem 3 of the 2nd problem session to compare the AIC and BIC model selection criteria for polynomials of different degree.

3. Download the modified Iris-data set from the course webpage (`http://www.cis.hut.fi/Opinnot/T-61.3050/2007/problems#6`) and divide it to a training- and test-set so that you have 25 examples of both classes both for training and testing. Assume the class-conditional densities are gaussian and use the multivariate classification approach to construct classifiers for this data. Try the *quadratic discriminant*, *linear discriminant*, *naive Bayes* and *nearest mean* classifiers.