Datasta Tietoon, Autumn 2010

EXAM 12. 1. 2011 (note: problems in Finnish on the reverse side)

1.
A convolution filter is given using the formula

$$g_k = \sum_{m=-\infty}^{\infty} f_m h_{k-m},$$

where f_m is the (discrete) input signal, h_n is the filter sequence, and g_k is the output signal. Derive the output signal g_k when

$$f_0 = f_1 = 1, \quad f_m = 0 \text{ otherwise;} \tag{1}$$
$$h_n = a^n, n \geq 0, \quad h_n = 0 \text{ otherwise} \tag{2}$$

where a is a constant, $0 < a < 1$. Plot the output signal when $a = 0.5$.

2.
Assume two classes for a scalar variable x. The class densities $p(x|\omega_1), p(x|\omega_2)$ are gaussian such that both have mean value 0 but different standard deviations σ_1, σ_2. The prior probabilities are $P(\omega_1), P(\omega_2)$. Plot the densities. Derive the decision boundaries of the Bayes classifier.

3.
Assume 5 vectors $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_5$, whose mutual distances $d(\mathbf{x}_i, \mathbf{x}_j)$ are given as the following matrix

$$D = \begin{bmatrix} 0 & 4 & 9 & 6 & 5 \\ 4 & 0 & 1 & 8 & 7 \\ 9 & 1 & 0 & 2 & 3 \\ 6 & 8 & 2 & 0 & 1 \\ 5 & 7 & 3 & 1 & 0 \end{bmatrix}$$

Perform hierarchical clustering for the vectors and plot the clustering tree, assuming that the distance between two clusters is equal to the distance between the two closest vectors in the two clusters. What is the best clustering into 3 clusters?

4.
When searching for relevant pages from the Web, a method based on hubs and authorities is sometimes used. Describe the basic principle of this method and give the formulas for computing the weights.

5.
Answer one of the following essay questions that are associated with the Matlab exercise:
A) "Eigenfaces" and the use of eigenvalues for clustering face images
B) k-nearest neighbor classifier.