$SST = \frac{1}{N}\sum (t_n - \hat{t})^2$

$SSE = \frac{1}{N}\sum (t_n - y_n)^2$

Exam Pattern Recognition
Date: Tuesday, March 15, 2011
Time: 14.00-17.00

General Remarks

1. Hand in your answers to part A (statistical pattern recognition) and part B (geometrical pattern recognition) on separate sheets of paper.
2. Put your name and student number on every sheet.
3. It is not allowed to consult books, notes, telephone, etc., or someone else's answers. You are allowed to use a calculator.
4. Always show how you arrived at the result of your calculations. Always explain your answer, used symbols, etc.; be precise.
5. Answers may be given in Dutch or in English.

Part A: Statistical Pattern Recognition

Question 1 Short Questions (8 points)

(a) In neural networks, what is an activation function?

(b) In linear regression, what is R^2?

(c) Give two parameters used in training support vector machines that influence the complexity of the model. For each parameter, explain how it influences model complexity.

$R^2 = \frac{SSE}{SST}$

(d) Give the steps of the backpropagation algorithm used in training neural networks.

Question 2 Logistic Regression (12 points)

We analyse data on 200 men taken from the Los Angeles Heart Study conducted under supervision of John M. Chapman, UCLA. The data consist of four variables:

Abbreviation	Variable	Units
Ag	Age	in years
Ch	Cholesterol	milligrams per DL
W	Weight	pounds
Cnt	Coronary incident	1 if an incident had occurred in the previous ten years; 0 otherwise

We estimate the model

$$p(\mathrm{Cnt}=1) = \frac{\exp(w_0 + w_1\mathrm{Ag} + w_2\mathrm{Ch} + w_3\mathrm{W})}{1+\exp(w_0 + w_1\mathrm{Ag} + w_2\mathrm{Ch} + w_3\mathrm{W})}$$

with maximum likelihood. This yields the following results

$$w_0 = -9.25 \qquad w_1 = 0.053 \qquad w_2 = 0.007 \qquad w_3 = 0.018$$

(a) We note that w_1 has a positive sign. Is this surprising? Explain.

(b) According to this model, what is the probability that someone had a coronary incident in the past ten years, given that he or she is 50 years old, has a cholesterol level of 350 and weighs 200 pounds.

(c) Use the fitted model to give a simple classification rule.

$w_0 + w_1 Ag + w_2 Ch + w_3 W > 0$

w_{kj} j_k

4 + 1 + 3   (2-4)

-2

4

(2-4)(2-5) + (4-4) + 2 · 2

-2 · -2 + ... + 2 · 3

-2 · -3

6   + 4

Question 3 Discriminant Analysis (15 Points)

We are given the following training data:

x_1	x_2	t
2	2	1
4	6	1
6	7	1
8	5	2
10	6	2
12	10	2

Here t is the output variable (the class or group label) and x_1 and x_2 are the input variables. Assume that $\mathbf{x} = [x_1 \;\; x_2]^{\mathrm{T}}$ follows a bivariate normal distribution in each group, and that the two groups have the same covariance matrix.

(a) Estimate the group mean vectors $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, and the group prior probabilities $p(t = 1)$ and $p(t = 2)$.

(b) Estimate the common group covariance matrix $\boldsymbol{\Sigma}$.

(c) Compute the linear discriminant functions $a_1(\mathbf{x})$ and $a_2(\mathbf{x})$ based on the estimates obtained under (a) and (b).

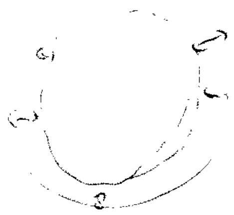
(d) On the basis of the observed data, is the use of *linear* (as opposed to *quadratic*) discriminant analysis justified?

Question 4 Estimation (15 points)

A student has decided that the least squares estimator is too much trouble to compute. He recalls from his geometry class that two points determine a line, so he chooses two points from a sample of size N and draws a line through them. He proposes to use the slope of this line as an estimator of w_1 in the simple linear regression model. Algebraically, if the two points are (t_1, x_1) and (t_2, x_2), then the estimator is

$$w_1^* = \frac{t_2 - t_1}{x_2 - x_1}$$

Assume that all the usual assumptions of the linear regression model apply, that is



1. $t_i = w_0 + w_1 x_i + \varepsilon_i$, with $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$, $i = 1, \ldots, N$.
2. t_i, t_j independent for $i \neq j$.
3. The x_i are not random variables, but fixed constants.
   In particular, x_1 and x_2 are *fixed* in repeated samples.

(a) Show that w_1^* is an *unbiased* estimator of w_1, that is, $\mathbb{E}[w_1^*] = w_1$.

(b) Derive the variance of w_1^*.

It is given that $\hat{w}_1$, the least squares estimator of w_1, is unbiased. Its variance is given by

$$\mathrm{var}(\hat{w}_1) = \frac{\sigma^2}{\sum (x_i - \bar{x})^2}.$$

(c) Convince the student that his estimator is not as good as the least squares estimator.

Part B: Geometrical Pattern Recognition

1. (Annulus) 10 points

   What is the smallest width annulus, and how can it be used to recognize circles in a finite point set in the plane?

2. (Exact point sets) 10 points

   Give an algorithm for 1-1, approximate finite point set matching with a noise region of radius at most ϵ, under rigid motion in the plane. What is the time complexity?

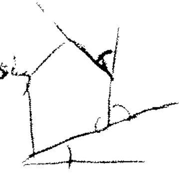
3. (Voting algorithm) 15 points

   Explain how "Pose clustering" (also called "Generalized Hough transform") works. What is the time complexity? Is it for exact or approximate, partial or complete, and 1-1 or n-m matching?

4. (Polygon matching) 15 points

   (a) Explain what the cumulative turning angle function is.

   (b) Give an algorithm to compute the dissimilarity between two planar polygons on the basis of the cumulative turning angle function.



$$\frac{w_0 + w_1 x_2 - w_0 - w_1 x_1}{x_2 - x_1} = \frac{w x_2 - w x_1}{x_2 - x_1} = w$$

Formulas

Estimated linear discriminant function:

$$a_k(\mathbf{x}) = \bar{\mathbf{x}}_k^{\mathrm{T}} \hat{\boldsymbol{\Sigma}}_{\text{pooled}}^{-1} \mathbf{x} - \tfrac{1}{2} \bar{\mathbf{x}}_k^{\mathrm{T}} \hat{\boldsymbol{\Sigma}}_{\text{pooled}}^{-1} \bar{\mathbf{x}}_k + \ln \frac{N_k}{N},$$

with

$$\hat{\boldsymbol{\Sigma}}_{\text{pooled}} = \sum_{k=1}^{K} (N_k/N) \hat{\boldsymbol{\Sigma}}_k$$

$$E\left(\frac{t_2 - t_1}{x_2 - x_1}\right) = w_1$$

10

7

$$\begin{bmatrix} 1\tfrac{1}{6} & -5/6 \\ -5/6 & 2/3 \end{bmatrix}$$

$5\frac{5}{6}$ $-3\frac{2}{3}$

$[4 \quad 5]$ $1/2$ 0

$[10 \quad 7]$ $5\frac{5}{6}$ $-3\frac{2}{3}$

$1/2$

x_1

0

x_2

4

$1/2 \quad 0$ $1/2\, x_1$

5

$1/2 \quad 0$ 2

1 voor alle puntparen in S_1 : l_1 en l_2

2. voor alle paren van punten in S_2 : z_1 en z_2

3 Bereken een transformatie die $l_1 = t(z_1)$ en $l_2 = t(z_2)$ geeft

4 Voer de transformatie uit op S_2 : $T = t(S_2)$

5 Tel het aantal punten in T dat binnen een afstand ε van een punt in S_1 ligt.

6 Voor 1-1 matching is extra controle nodig