AARHUS UNIVERSITET

Computer Engineering

Term:	**Ordinary exam Fall 2019**
Examination:	**Optimization and Data Analytics**
Date:	**16. December 2019**
Duration:	**3 hours, 9-12**

1 cover plus paper for draft and fair copy will be handed out.

Digital Exam The exam questions will be available in “Digital Exam”, and your exam answers must be handed in via “Digital Exam”. Handwritten parts of the exam answers must be digitized and attached to your exam paper. In “Digital Exam”, the exam answers must be uploaded in PDF format. You will get a receipt, immediately after you have uploaded correctly. Remember to upload within the time limit, if you exceed the time limit, you must send in an application for an exemption. It is your own responsibility to have knowledge of the rules for electronic hand in and if necessary, to be able to download to your own USB memory key in the unlikely event that Blackboard is down.

REMEMBER name and study number on all pages and in the file name when you upload (pdf).

Aids All aids, like Computer, mathematical software, books, notes, pen and paper and internet is allowed. No form of communication or file sharing is allowed during the exam.

Exercise 1

Consider the following optimization problem:

Maximize: $f(x_1, x_2) = 4x_1 + 2x_2 + x_3$

Subject to: $x_1 \leq 1;$ (i)

$4x_1 + 1x_2 \leq 4;$ (ii)

$8x_1 + +4x_2 + 1x_3 \leq 16;$ (iii)

and $x_1 \geq 0; x_2 \geq 0;$

a) Write out the simplex tableaux for the problem and show the first step needed to bring in a new variable into the solution (e.g. argue what column and row to choose, and what elementary operations are needed for the first reductions). Find the maximum for f on the feasible set (you may use MatLab).

b) Write out the dual problem for the optimization problem above and find the solution of this. Explain how to find the solution to the dual problem, and explain how to interpret the solution.

Exercise 2

Consider the function defined $g(x_1, x_2) = x_1^3 - 3x_1 - x_2^2$

a) Find the critical points for the function g.

b) Classify each critical point as a local maximum, local minimum, or saddle.

Now let $f(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 1)^2$ and $h(x_1, x_2) = (x_1 - 1)^2 + (x_2 + 2)^2 - 1$.

c) Find the minimum and maximum of f subject to the constraint $h(x_1, x_2) = 0$ (argue for your calculations).

Exercise 3

Answer the following with **ONE sentence** per question.

a) Are particle swarm optimization methods guaranteed to find the global optimum if the search is run for enough iterations (Yes/No)?

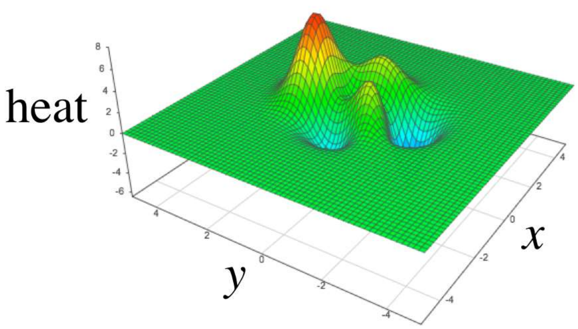
b) In the context of optimization, is the global optimum necessarily unique (Yes/No)?

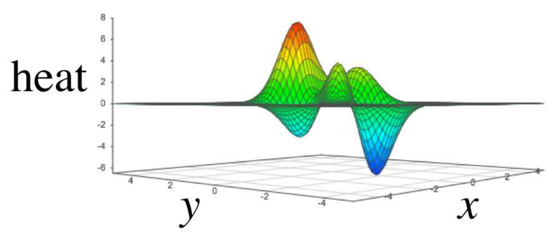
c) In simulated annealing, what are the two conditions that make the search algorithm more likely to accept a **worse** candidate solution?

d) What is discrete optimization? Name one example of a well-known discrete optimization problem.

e) What is the main difference between Newton and Quasi-Newton methods?

f) We are helping a cat find the warmest location in the room to have a sleep. Each location in the room is represented by two integer coordinates (*x*,*y*) such that *x* and *y* are each in the interval [-7, 7]. Every location gives a real-value heat score in the range [-8, 8] described by a function *heat*. The distribution of heat across the room according to function *heat* is partially illustrated in the graphs below, e.g. *heat*(2,2) = 8.





We want to use a **genetic algorithm** to find the warmest location in the room, i.e. the (x,y) coordinates that have the highest "heat" score.

i) Develop a representation of a candidate solution location as a chromosome. Explain your representation, and give one example of a chromosome with the corresponding (x,y) coordinate.

ii) Define your **objective** function *f* for population size *N* so that you can evaluate each candidate solution. Give an example of *f* using a population of two candidate solutions (x1,y1), (x2,y2) where:
heat(x1,y1) = 4
heat(x2,y2) = -4

iii) Define your **fitness** function *F* for population size *N* so that you can evaluate each candidate solution. Give an example of *F* using a population of two candidate solutions (x1,y1), (x2,y2) where:
heat(x1,y1) = 4,
heat(x2,y2) = -4

(exercise 3 continues on the next page)

Exercise 3 (cont.)

iv) Give an example of single-point cross-over using your chromosomes (choose an example that clearly demonstrates this).

v) Given an initial population P of size N, explain the THREE main steps of your genetic algorithm search for ONE generation (about one sentence per algorithm step).

Exercise 4

A classification problem is formed by two classes. We are given a set of 2-dimensional data **X** = [$\mathbf{x}_1$ $\mathbf{x}_2$ … $\mathbf{x}_5$]:

$$\boldsymbol{X} = \begin{bmatrix} 0 & 1 & 3 & 3 & 4 \\ 1 & 1 & 1 & 2 & 2 \end{bmatrix},$$

each belonging to one of the two classes, as indicated in the class label vector **l** = [l_1 l_2 … l_5]:

$$l = [1 \quad 1 \quad 2 \quad 2 \quad 2].$$

Using the above (training) vectors and the corresponding class labels, classify the following vectors:

$$\mathbf{x}_6 = [1 \quad 2]^T, \quad \mathbf{x}_7 = [3 \quad 0]^T, \quad \mathbf{x}_8 = [2 \quad 1]^T$$

using:

a) The Nearest Class Centroid (NCC) classifier
b) The Nearest Neighbor Classifier (using only one neighbor)
c) The Bayes-based classification scheme, where:

$$p(\boldsymbol{x}_i|c_k) = \frac{\|\boldsymbol{x}_i - \boldsymbol{m}_k\|_2^{-2}}{\sum_{m=1}^{K}\|\boldsymbol{x}_i - \boldsymbol{m}_l\|_2^{-2}},$$

where $\boldsymbol{m}_k$ is the class mean vector of class c_k, k=1,2 and $\|v\|_2^{-2} = 1 / (v^T v)$.

d) Compare (qualitatively) the decision functions obtained by using the NCC classifier and the above Bayes-based classification scheme.

Exercise 5

The two classes of a binary classification problem are formed by the blue and red samples plotted in Figure 3.

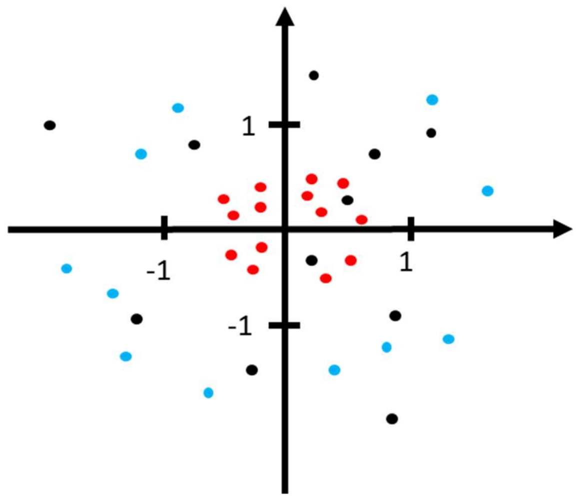


Fig. 3: Training samples of a two-class classification problem

a) Would you use Linear Discriminant Analysis in order to discriminate the two classes in a one-dimensional space? If yes, why. If no, why not?

b) Describe how we can use a Generalized Linear Discriminant Function in order to classify the test samples (plotted as black dots) using a linear classifier. Draw the linear classification function in the multi-dimensional feature space where linear classification is performed. Draw the corresponding classification function in the original feature space (the one shown in Figure 3).

c) Draw a neural network with one hidden layer that can be used for the above classification problem. In your drawing include all the parameters and activation functions.

d) Explain (with equations) why we cannot use linear activation functions for the neurons of the neural network you drew in question (c) for solving the problem in Figure 3.

Exercise 6

In a two-class classification problem, the distribution of the class-conditional probabilities $p(x|c_k)$, k=1,2 is given in Figure 1.

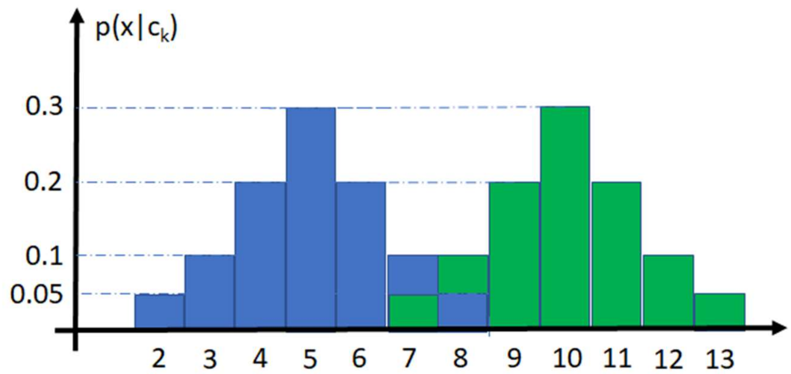


Fig. 1: Class-conditional probabilities of class 1 (blue) and class 2 (green)

The number of samples in class 1 and class 2 is equal to 100 and 200, respectively.

a) Classify (using the trained classifier) the following vectors (test) samples:

$$x_1 = 3, \quad x_2 = 7, \quad x_3 = 8 \text{ and } x_4 = 9$$

b) Consider that the classification risk for the two classes is given by the matrix:

$$\Lambda = \begin{bmatrix} 0 & 0.3 \\ 0.2 & 0 \end{bmatrix}$$

where $\Lambda_{ij} = \lambda((\alpha_i|c_k)$ is the risk of taking action α_i while the correct class is c_k. What is the classification result for the (test) samples x_i, i=1,…,4?