Computational Statistics
Exam, Fall 2023

Department of Computer and Information Science (IDA), Linköping University

March 19, 2024, 8:00-13:00

Course:	732A90 Computational Statistics
Teachers:	Krzysztof Bartoszek, Frank Miller (examiner)
Allowed aids:	Printed course books, 100 page computer document
Provided aids:	material in zip file exam_material_732A90.zip
Grades:	A = [18, 20] points, B = [16, 18) points, C = [14, 16) points, D = [12, 14) points, E = [10, 12) points, F = [0, 10) points.
Instructions:	Provide a detailed report that includes plots, conclusions and interpretations. If you are unable to include a plot in your solution file clearly indicate the section of R code that generates it. Give motivated answers to the questions. If an answer is not motivated, the points are reduced. Provide all necessary codes in an appendix. In a number of questions you are asked to do plots. Make sure that they are informative, have correctly labelled axes, informative axes limits and are correctly described. Points may be deducted for poorly done graphs. If you have problems with creating a pdf you may submit your solutions in text files with unambiguous references to graphics and code that are saved in separate files. There are TWO assignments (with sub–questions) to solve. Provide a separate solution file for each assignment. Include all R code that was used to obtain your answers in your solution files. Make sure it is clear which code section corresponds to which question. If you also need to provide some hand–written derivations please number each page according to the pattern: Question number . page in question number, i.e. Q1.1, Q1.2, . . ., Q2.1, Q2.2, ... Name your solution files as: [your exam account id] [own file description].[format]

Note: If you are not able to solve a part of a question, you can anyway try to show how you would solve the subsequent parts with explaining and providing code examples and you might receive partial points.

Assignment 1 (10p)

The task is to find a **global maximum** of the function

$$f(x) = \exp(-x^2/100) \cdot (2 - 0.1 \cdot \cos(3\pi(x + 0.1))).$$

Question 1.1 (2p)

Plot the function f.

Compare between the Newton-algorithm and the simulated annealing algorithm when this function should be maximized:

- What is the main problem here for the Newton algorithm and how can you deal with it?
- Discuss one advantage and one disadvantage of similating annealing versus Newton when applied for this function.

Question 1.2 (4p)

Write your own simulated annealing algorithm. The starting x-value and starting temperature should be parameters of your function. You can use a fixed iteration number, e.g. 10000, and you should decrease the temperature during the iterations. You need to decide how the x-value for the candidate in the next iteration is chosen based on the current iteration; describe how you did this and why.

Remember that the simulating annealing algorithm is often described for a minimization problem, while you have a maximization problem here.

Question 1.3 (4p)

Compare the performance of the simulated annealing algorithm for two different starting temperatures when using the starting x-value $x = 5$. Run the algorithm several times for each starting temperature for this comparison and check how many of the runs successfully found the global maximum.

If you failed to implement your own simulated annealing algorithm in the previous Part 1.2, you can do this Part 1.3 by using the SANN method of the `optim`*-function. The starting temperature can be changed via the argument* `control`*, see the helppage for* `optim`*. Recall that the default of* `optim` *is minimization and that you have here a maximization problem.*

Assignment 2 (10p)

Let $\mathbf{X} = (X_1, X_2)$ be a random variable (vector) with bivariate normal distribution $N(\mathbf{0}, \Sigma)$ which has density

$$f(\mathbf{x}) = \frac{1}{2\pi \det(\Sigma)^{1/2}} \exp(-\mathbf{x}^T \Sigma^{-1} \mathbf{x}/2)$$

for vector $\mathbf{x} = \binom{x_1}{x_2}$. We consider the covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

i.e., the correlation of X_1 and X_2 is $\rho \in [-1, 1]$.

The conditional distribution of X_1 given $X_2 = x_2$ is $N(\rho x_2, 1 - \rho^2)$. The conditional distribution of X_2 given $X_1 = x_1$ is $N(\rho x_1, 1 - \rho^2)$.

The task is here to generate a random sample following this bivariate normal distribution.

Question 2.1 (4p)

Implement a Gibbs sampler to generate a random sample of size n for X with this bivariate normal distribution.

If you fail to implement the Gibbs sampler, you can use another method of your choice to generate a bivariate normal distribution, and you might obtain up to 2 points for this part. Further, you can then use this other method to generate the samples in 2.2.

Question 2.2 (3p)

Use first the identity matrix as covariance matrix, $\Sigma = I$, i.e., $\rho = 0$. Generate $n = 1000$ samples $\binom{X_1}{X_2}$ with the Gibbs sampler and plot them (figure with x_1 on the horizontal and x_2 on the vertical axis).

Use now the following covariance matrix with correlation $\rho = 0.998$:

$$\Sigma = \begin{pmatrix} 1 & 0.998 \\ 0.998 & 1 \end{pmatrix}.$$

Generate 1000 samples $\binom{X_1}{X_2}$ with the Gibbs sampler and plot them.

Discuss the results: How good represents the sample the distribution in both cases, and why or why not are there differences between the two cases with regard to the quality of the results.

Question 2.3 (1p)

Based on your generated samples, provide an approximation for $P(X_1 + X_2 > 0)$ in both cases.

Question 2.4 (2p)

Instead of using Gibbs sampling, suggest an alternative algorithm which can be used to generate such a bivariate normal distribution and describe it (you need not to implement the alternative here in 2.4). Discuss advantages or disadvantages of the alternative in the two cases for Σ of Part 2.2.