Stockholms universitet

Computational Statistics
Advanced Course, 7.5 HEC, Spring 2022
Home Assignment 1 until January 27

Mahmood Ul Hassan and Frank Miller, Department of Statistics

January 20, 2022

You are supposed to solve these problems individually without any cooperation. Record your answers clearly. If you cite material from other sources, or use intellectual ideas or code from others, point out this clearly with stating the source. Please include your `R` computer code and relevant output in your solution. Send an email to `mahmood.ul-hassan@stat.su.se` with your solutions until the above deadline with a file name like `CompStat2022HA1_[Your name].pdf` and extra file(s) with similar name with your computer code.

1 Solving linear equations and matrix calculation

An estate agency is selling houses in a town. A dataset with the selling price data (y, in 1000 SEK) from 29 sold houses and with the following potential influence factors is in Athena (`VillorSW.csv` in the folder Assignments):

- x1 = size of living space (square meters, m^2),
- x2 = lot size (square meters, m^2),
- x3 = whether the house is in a specific area which is more attractive (x3=1) or not (x3=0),
- x4 = age of the house (years).

Ignore the additional variable z in the dataset. The selling price should be used as dependent variable and $x_1, \ldots, x_4$ as independent variables; additionally an intercept should be in the model. We want therefore to fit the regression model

$$Y = X\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{1}$$

where Y is the vector of the dependent variable, X the design matrix (with a first column consisting of 1's for the intercept), and ϵ the vector of errors. Fitting model (1) with the least square (LS) method is equivalent to solving

$$(X^\top X)\hat{\boldsymbol{\beta}} = X^\top Y. \tag{2}$$

a. Compute the necessary matrices for the LS system (2). Report the matrices and check whether the matrix $(X^\top X)$ is non-singular.

b. Find the solution for $\hat{\boldsymbol{\beta}}$ in (2).

c. Apply a standard `R`-function (for example `lm`) to fit the regression model and report the parameter estimates.

d. Compare your results in b. and c. and compare the computing times for both methods. Since the computation is quick in both cases, write a loop around each computation to repeat the same calculation e.g. 10000 times and note the time used for it.

2 Optimization of a general function of one variable

Consider the function

$$g(x) = \frac{\log(x+1)}{x^{3/2}+1}.$$

a. Plot the function $g(x)$ in the interval $[0, 4]$. What's your guess for the maximum point?

b. Compute the derivative $g'(x)$ of $g(x)$; recall the quotient rule that for $g(x) = \frac{u(x)}{v(x)}$, the derivative is $g'(x) = \frac{u'(x)v(x)-u(x)v'(x)}{(v(x))^2}$. Plot g' in $[0, 4]$, and add a horizontal reference line at 0 to the plot.

c. Apply the bisection method to g' in order to find a local maximum of g in $[0, 4]$. Use the absolute convergence criterion to assess the algorithm's convergence.

d. Apply the secant method to g' in order to find a local maximum of g in $[0, 4]$. Use the absolute convergence criterion to assess the algorithm's convergence. Compare with the result in c.

3 Optimization in experimental design

Given a cubic regression model $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \epsilon_i$ for four observations $x_1 = -1, x_2 = -a, x_3 = a, x_4 = 1$, the model can be written in matrix notation as:

$$Y = X\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

with

$$X = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & -a & a^2 & -a^3 \\ 1 & a & a^2 & a^3 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}.$$

The matrix $X^\top X$ is called information matrix for the parameter-vector β. In experimental design, a design is called D-optimal if it maximizes the determinant of the information matrix.

a. Compute the function $f(a) = \det(X^\top X)$ manually (report some steps of your derivation) such that you obtain a polynomial with variable a.

b. Program the function $f(a)$ in `R` both using the `det`-function, as well as using your result from a. avoiding the `det`-function.

c. Plot f for $a \in [0, 1]$ with both functions you have written in b. and check if they coincide by plotting their difference. Interpret the cases $a = 0$ and $a = 1$.

d. In `R`, one can use the function `optimize` to find a local optimum of a function within a given interval. Use this function to compute the local maximum of $f(a)$ on the interval $[0, 1]$.

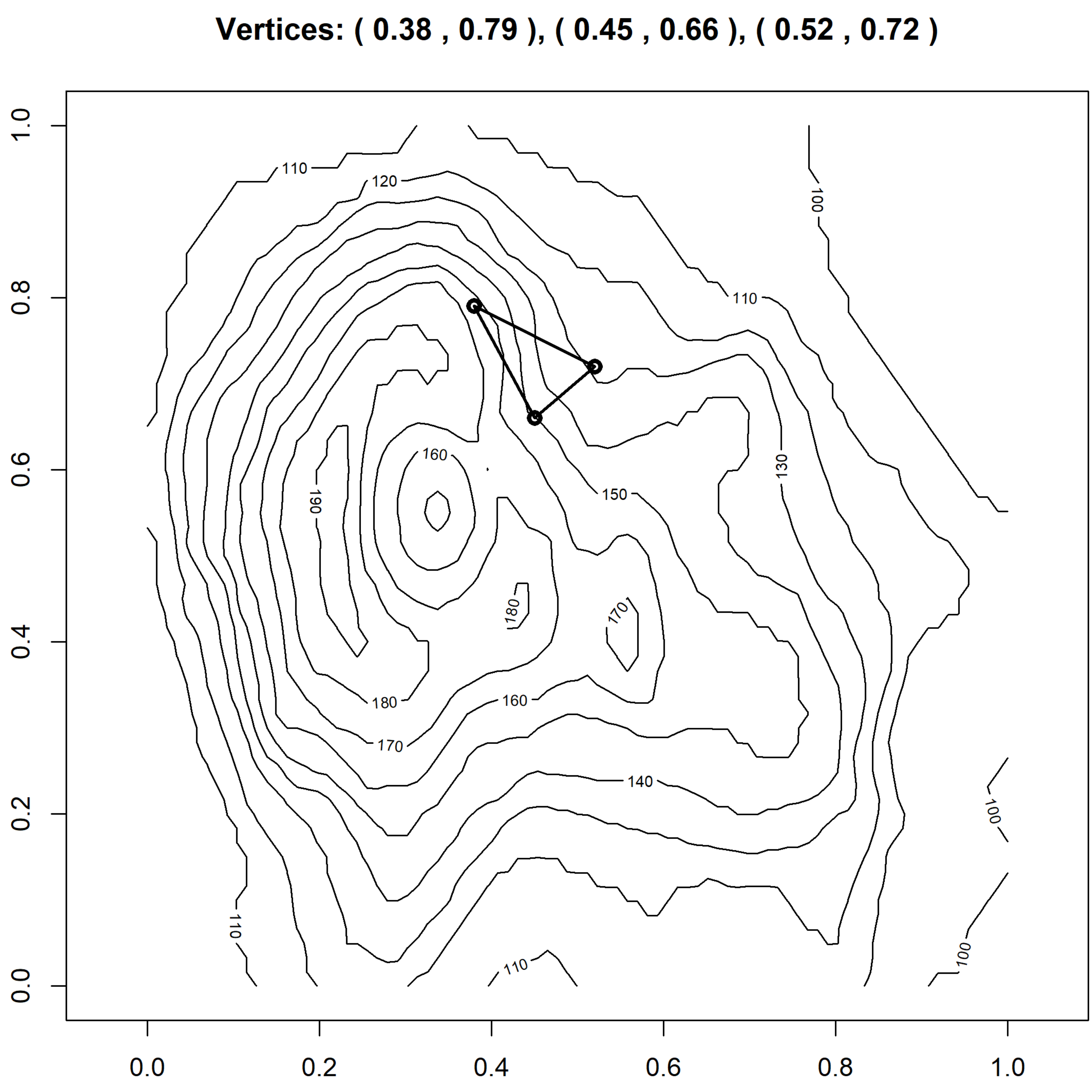
(Remark: One can calculate the maximum exactly based on the calculation in a. which requires a little more work, but you have not to do it here.)

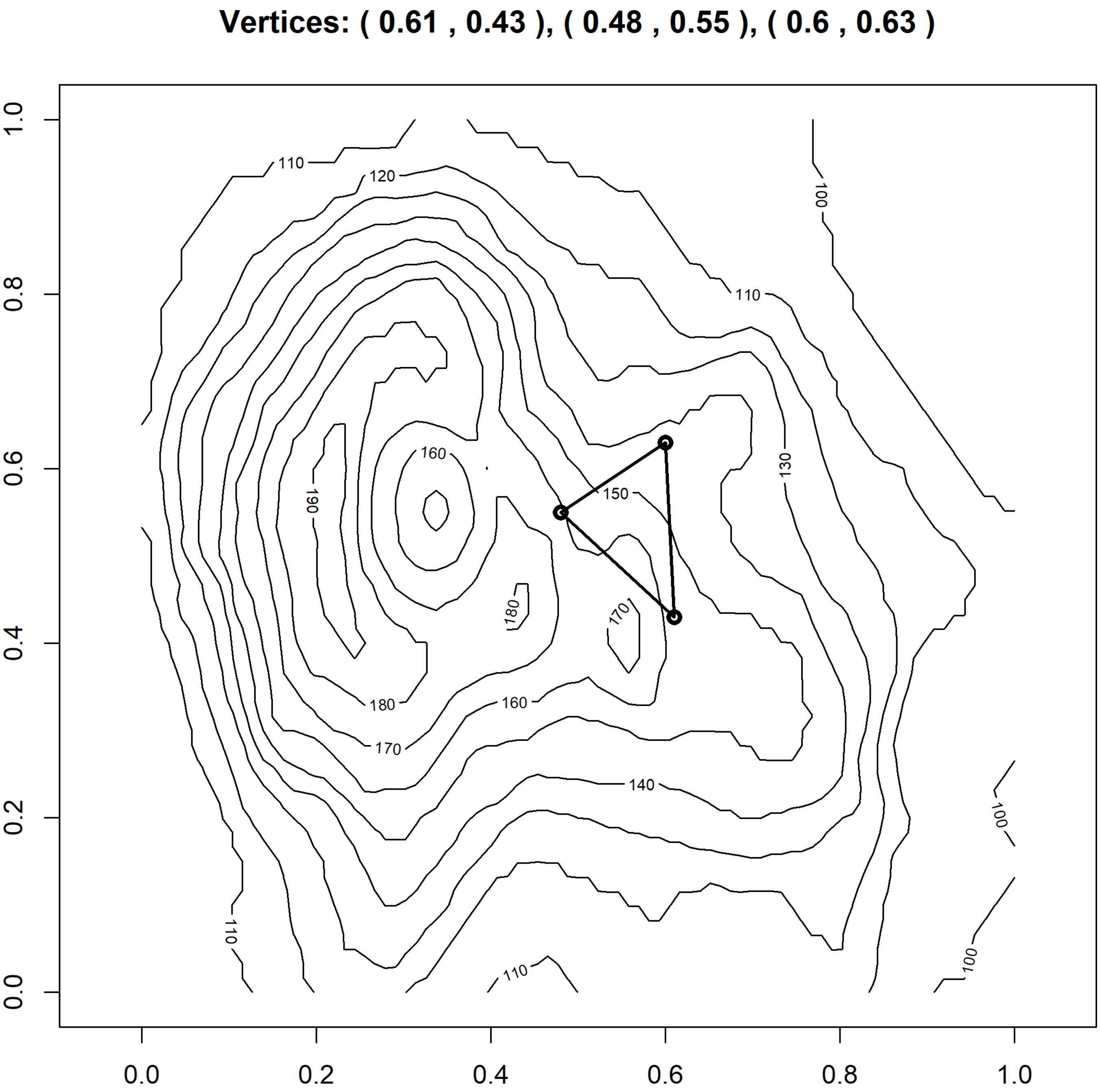
4 Maximization with Nelder-Mead algorithm

On the next pages, you see two contour plots of the same two-dimensional function where it is desired to find the maximum. In each plot, a triangle is drawn which represents a Nelder-Mead starting triangle.

Use the description of the algorithm in the course book Givens and Hoeting (2013), page 47-48, to construct **the next two** iterations (next two triangles) in the two pictures. Use the following standard Nelder-Mead values: $(\alpha_R, \alpha_E, \alpha_C, \alpha_S) = (1, 2, 0.5, 0.5)$.

Explain in detail the steps used. Draw the triangles in the plot (you can add it either electronically in the pdf or in the png-version placed in Athena; or you can print it out, draw with pencil and scan it; if no of the possibilities before works for you, read and report the coordinates of the vertices in your solution).


Vertices: (0.38 , 0.79), (0.45 , 0.66), (0.52 , 0.72)
110
120
100
110
130
150
160
190
180
170
180
160
170
140
100
110
110
100
0.0
0.2
0.4
0.6
0.8
1.0


Vertices: (0.61 , 0.43), (0.48 , 0.55), (0.6 , 0.63)
0.0
0.2
0.4
0.6
0.8
1.0
100
110
120
130
140
150
160
170
180
190