LUISS Guido Carli

Bachelor's Degree in Management and Computer Science
Quantitative Models for Data Science – A.Y. 2022/2023

Problem sheet n.5 – *Critical points, least square problems.*

Exercise 1. Discuss whether the function $f(x,y) = x^2 + 3y^2 - 2xy + 12x - 7y + 2$ admits a global minimum, and find it in that case. What about a global maximum?

Exercise 2. For each of the following functions

(a) $f(x,y) = x^3 + y^3 - 6xy$,
(b) $f(x,y) = \log(y^2 - x) + 2x$,
(c) $f(x,y) = \dfrac{x-y}{x^2+1}$,
(d) $f(x,y) = x(x+y)\mathrm{e}^{x-y}$,
(e) $f(x,y) = (2x + y - 1)^3$,
(f) $f(x,y) = e^x - x\log(y)$.
(g) $f(x,y) = 4x^2y + x - 4y$
(h) $f(x,y) = \sqrt{x^2 + y^2 + 1}$
(i) $f(x,y) = \log(x - y + 1) + x^2 + 2y$
(j) $f(x,y) = ye^{x^2+y}$
(k) $f(x,y) = \dfrac{xy - x}{x + 2}$
(l) $f(x,y) = \dfrac{x}{1 + x^2 + y^2}$.

1. Find all critical points.
2. Determine the nature of the critical points (i.e. decode whether the critical points are local maxima, local minima or saddle points).
3. Determine whether local minima and maxima are also global.

Exercise 3. Let $\mathbf{x} = \begin{bmatrix} 2 \\ -3 \\ 3 \\ 1 \\ -4 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 4 \\ 1 \\ 2 \\ 2 \\ -1 \end{bmatrix}$ be two vectors of data that you believe to be linearly correlated.

1. Find the design matrix X, then verify that the square matrix X^TX is invertible, and compute $(X^TX)^{-1}$.
2. Find the vector of the coefficients $\widehat{\beta} = (X^TX)^{-1}X^T\mathbf{y}$ and the fitted values $\widehat{\mathbf{y}} = X\widehat{\beta}$.
3. Draw points representing the data and the regression line.
4. Compute the residuals $\widehat{\varepsilon} = \mathbf{y} - \widehat{\mathbf{y}}$, then compute the residual sum of squares $\|\widehat{\varepsilon}\|^2$.
5. Predict the value of y for $x_1 = -1$ and $x_2 = 7$, which prediction should you trust the most?

Exercise 4. GoodLife is a recently established insurance company. As of today, it has 6 customers, whose personal data have been meticolously collected. Among data, we can find:

	Cust. 1	Cust. 2	Cust. 3	Cust. 4	Cust. 5	Cust. 6
Hours of physical exercise per day	1.5	2.5	1	3	0.5	4
N° of doctor appointments per year	6	3	4	1	5	1

In order to price its products, the company needs to predict the number of doctor appointments per year, and seeks to find a simple linear model that uses just the hours of physical exercise per day as predictor.

1. Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points above, with x representing the daily hours of physical exercise and y representing the the number of doctor appointments per year.
2. Compute the residual associated to each data point.
3. Draw the data points and the least-squares line.
4. A potential customer is interviewed and says that they exercise 3.5 hours a day. Can you predict their number of of doctor appointments per year?