LINKÖPING UNIVERSITY

Advanced Computational Statistics – Spring 2025
Assignment for Lecture 2

Sebastian Mair, sebastian.mair@liu.se,
Department of Computer and Information Science, Linköpings University

March 12, 2025

Perform the solutions individually and send your report **until March 31** by email to me. Try to keep this deadline. You will then be assigned to perform a peer-review which you have to hand in until April 14.

If you have problems with it, there will be a final deadline on September 30 for all assignments.

Problem 2.1

Consider a linear regression model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ for a design matrix $\mathbf{X} \in \mathbb{R}^{n \times p}$ of full rank p with $n > p$. The interest is in the least squares solution, i.e., the function

$$g(\mathbf{b}) = \frac{1}{n}\|\mathbf{X}\mathbf{b} - \mathbf{y}\|_2^2$$

should be minimised. Show that g is L-smooth and m-strongly convex. Present expressions for L and m.

Problem 2.2

We consider an ML estimation for simple logistic regression

$$p(x) = P(Y = 1|x) = \frac{1}{1 + \exp(-\beta_0 - \beta_1 x)}.$$

a. Program a stochastic gradient descent algorithm with a fixed step size α and a predefined total number of iterations T for simple logistic regression. Your program should also plot the computed $(\beta_0^{(t)}, \beta_1^{(t)})$ in each iteration t ($t = 0, 1, \ldots, T$) such that you can monitor the search path.

b. Analyze the dataset `logist.txt` (homepage; first column is x, second column y) with your algorithm using the starting value $(\beta_0^{(0)}, \beta_1^{(0)}) = (0.2, 0.5)$. Choose the total number of iterations T and the step size α. You might need to test different options first to come to a good choice. Explain why you have chosen these values T and α.

Problem 2.3

Verify the formula

$$f(x^k) = \frac{1}{2}\mathbb{E}_{\omega_1,\ldots,\omega_k,\omega}\left[\left(\frac{1}{k}\sum_{j=1}^{k}\omega_j - \omega\right)^2\right] = \frac{1}{2k}\sigma^2 + \frac{1}{2}\sigma^2, \tag{5.14}$$

given that the mean of the random variable ω is μ and its variance is σ^2. Note that the random variables $\omega_1, \ldots, \omega_k, \omega$ all follow the same distribution, and all random variables in this expression are independent.

References

Wright, S. J. and Recht, B. (2022). *Optimization for data analysis.* Cambridge University Press.