Exercise and Homework Round 5

These exercises (except for the last) will be gone through on **Friday, October 14, 12:15–14:00** in the exercise session. The last exercise is a homework which you should return via mycourses by **Friday, November 4 at 12:00**.

Exercise 1. (Gradient descent with line search)

It is also possible to use line search with gradient descent algorithm. Write down the pseudo-code for this algorithm when an exact line search (grid search) is used.

Exercise 2. (Gauss–Newton with linear-search for scalar models)

Let us consider a scalar nonlinear model

$$y = g(x) + r \tag{1}$$

and the corresponding least squares cost function

$$J(x) = (y - g(x))^2. \tag{2}$$

Further assume that $J(x)$ has a unique global minimum x^*.

(a) Write down the equation for scaled Gauss–Newton iteration (with scaling parameter γ) for this model.

(b) Assume that if we start the iterations from point $\hat{x}^{(0)}$ such that the derivative at that point $g'(\hat{x}^{(0)}) \neq 0$. Show that there exists a scaling parameter γ such that the algorithm reaches x^* on a single step.

(c) What does this imply on the (exact) line-search Gauss–Newton algorithm convergence for this model? Does this conclusion extend to multidimensional models?

Exercise 3. (Gauss–Newton with line search and Levenberg–Marquardt

Consider the model

$$y_n = g(\mathbf{x}) + r_n, \tag{3}$$

where $n = 1, \ldots, N$, $r_n \sim \mathcal{N}(0, R)$, and

$$g(\mathbf{x}) = \begin{bmatrix} \alpha \sqrt{x_1} \\ \beta \sqrt{x_2} \end{bmatrix}. \tag{4}$$

(a) Simulate data from this model with suitable parameters and implement Gauss–Newton algorithm with exact line search for minimizing the cost function.

(b) Implement a Levenberg–Marquardt algorithm for this model.

(c) Discuss the relative convergence rate of the methods.

Homework 5 (DL Friday, November 4 at 12:00)

Implement Gauss–Newton with line search for minimizing the cost function $J(x) = (1.1 - \sin(x))^2$. Use grid search with grid $\gamma \in [0, 1/10, 2/10, \ldots, 1]$. *Hint: Beware of the singularity of the derivative at the minimum.*