TTK4135 Optimization and Control
Spring 2019

Norwegian University of Science and Technology
Department of Engineering Cybernetics

Exercise 9
Hints

Problem 1 (50 %) Unconstrained Optimization

It can be fairly difficult and time consuming to write a working implementation of all three methods. If you are stuck, you can look at the file steepest_descent_suggestion.m for a rough idea of how an implementation *might* look. You can use this file as a framework for implementing the steepest-descent method, but you might prefer to write your own. Once you think you have a working implementation, you can use the file plot_iter_rosenbrock.m to check that the variables move towards the minimizer $x^* = [1,\ 1]^\top$.

When choosing the initial step length α_0 for line search at each iteration k, use the formula on page 59 for the steepest-descent direction, and $\alpha_0 = 1$ for the Newton and Quasi-Newton directions (as described in Section 2.2 in the textbook for the Newton method).

For the BFGS method, the initial guess of the inverse Hessian, H_0, can be crucial. A finite difference approximation based on the gradients is probably a good approach.

Problem 2 (5 %) Cholesky Factorization

Problem 3 (10 %) Gradient Calculation

a The expression you derive for the approximated gradient depends on the perturbation ϵ. (The two elements of the approximated gradient will not have the same ϵ dependence.)

b When you calculate the analytical expression for the gradient, you will see that each element differs from the approximation by a function of ϵ.

c How does the approximation improve numerically as ϵ goes to zero? Are there relevant things to say about the sensitivity to ϵ?

Problem 4 (30 %) The Nelder-Mead Method

e A formula for the average function value can be found in Equation (9.29) in the textbook. This must be used with the now familiar definition of convexity (Equation (1.4) in the textbook). Convexity can tell us something about what happens to the function value when shrinkage is performed (shrinkage is described in Algorithm 9.5). This can again be used to derive an inequality that describes the function values at the points of the simplex after shrinkage. This inequality can then be used to argue whether or not the average function value has decreased after shrinkage.