Division of Computing Science
Department of Information Technology
Uppsala University

Written examination in 1TD184 Optimization

- Date: 2023-08-25, 08.00-13.00
- Allowed tools: Pocket calculator, one A4 paper with notes (computer typed, font size minimum 10 pt).
- Maximum number of points: 36 (18 to pass).
- All assumptions and answers *must* be motivated for full points.

(1) Explain the following concepts: 1) Compass search, 2) Gauss-Newton method. (4p)

(2) Consider the linear least-squares data fitting problem (LLSQ). For a linear function $b = x_1 t + x_2$, x_1 and x_2 are our optimization variables defining the function. There are m data points $(t_1, b_1), (t_2, b_2), \ldots, (t_m, b_m)$. If there is a perfect fit, then there exist x_1 and x_2 such that $b_i = x_1 t_i + x_2, i = 1, \ldots, m$. This is however very unlikely, hence there will be some errors. Given x_1 and x_2, the error for data point i is the absolute difference between $x_1 t_i + x_2$ and b_i. The objective function is to minimize the sum of the squares of errors.

- Formulate LLSQ mathematically, and apply the first-order optimality condition to show that solving LLSQ amounts to solving a linear equation system. (2p)
- Show that LLSQ is a convex optimization problem. (2p)

(3) True or false (answers *must* be motivated):

a) For function $f(x)$, a direction p may be a descent direction if $p^T \nabla f(x) = 0$. (1p)

b) For constrained optimization, a local optimum is always a stationary point of the Lagrangian function. (1p)

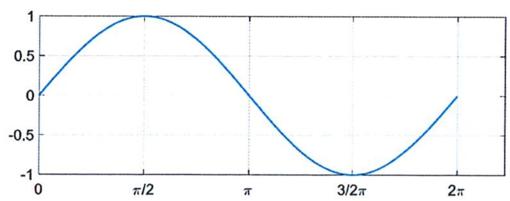
c) An optimization method designed for equality cannot be used for inequality constraints. (1p)

d) Sequential quadratic programming (SQP) does not guarantee improvement in each iteration. (1p)

(4) Consider unconstrained optimization.

- Quasi-Newton methods strive for a balance between computational complexity and convergence speed. Explain the mechanisms used to achieve this balance. (2p)
- Suppose we perform the following: In each iteration, we run steepest descent, Newton, and Quasi-Newton (with line search for each of them), in parallel on multiple CPU cores. We then select the best solution from the three, and repeat this process. What could be the advantage(s) and potential issue(s) of this approach? (2p)

(5) Consider the sine function $\sin(x)$ in interval $[0, 2\pi]$, illustrated below. (We use radians here for x, instead of degree.)



- For the interval $[0, 2\pi]$, show that $\frac{\pi}{2}$ and $\frac{3}{2}\pi$ are stationary points, and these two are local maximum and minimum, respectively[1]. (2p)
- Consider $x = \pi$. Compute the (steepest) descent direction and the Newton direction. If any of them does not "exist", could you provide an intuition of why? (2p)
- Explain the basic idea of Armijo line search with backtracking. Denote by μ ($0 < \mu < 1$) the Armijo parameter tuning the line function. Consider $x = \pi$ and search direction $p = 1.0$. Give one value of μ such that line search has no chance of finding the minimum. Next, for the μ value you have, give one value of the initial step size, such that at least one backtracking will take place. (3p)

(6) Consider the following problem, where Q is a 2×2 matrix $\begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$, coefficient vector $c = (c_1, c_2)^{\mathrm{T}}$, and variable vector $x = (x_1, x_2)^{\mathrm{T}}$,

$$\begin{aligned} \min \quad & \frac{1}{2} x^{\mathrm{T}} Q x + c^{\mathrm{T}} x \\ & x_1 + x_2 = 1 \end{aligned}$$

- Build the Lagrangian function to express the first-order necessary optimality condition, and show that the condition leads to a linear equation system. Is the number of unknowns equal to the number of equations? (3p)
- Suppose now Q is the identity matrix. Are you able to obtain a solution from the linear equation system? If yes, is this a minimum? If no, what is the underlying reason? (2p)
- Suppose all elements of Q are zeros, then the objective function becomes $\min c^{\mathrm{T}} x$, and the problem is a linear program (LP). Does optimal solution exist to this LP? Does your answer hold also for the general LP $\{\min c^{\mathrm{T}} x, Ax = b\}$ where A is of dimension $m \times n$ (with $m < n$)? (2p)

(7) We have the following LP and its dual, where coefficient $b > 0$.

(Primal) LP

$$\begin{aligned} \min \quad & x_1 + 2x_2 + \cdots + nx_n \\ & 2x_1 + 3x_2 + \cdots + (n+1)x_n \geq b \\ & x_1, x_2, \ldots, x_n \geq 0 \end{aligned}$$

Dual

$$\begin{aligned} \max \quad & by \\ & 2y \leq 1 \\ & 3y \leq 2 \\ & \ldots \\ & (n+1)y \leq n \\ & y \geq 0 \end{aligned}$$

- Utilize the structure of the dual LP to solve it, and argue why the solution you obtain is indeed the optimal one to the dual. (2p)
- Use the dual optimum to solve the primal LP. Which variable is the basic variable in the primal optimum? Does this basic feasible solution satisfy the reduced cost criterion of the simplex method? (2p)
- Consider the two-dimensional case with $n = 2$ for the primal LP. Illustrate the solution space for $b = 6$. Specify all extreme points, and verify if the extreme point with the minimum value corresponds to what you have obtained above. (2p)

Good Luck!

[1] The derivative of $\sin(x)$ is $\cos(x)$, and the derivative of $\cos(x)$ is $-\sin(x)$.