Exam Continuous Optimization

22 January 2024, 14.00–17.00

This closed-book exam consists of 5 questions. Please start each question on a new page, write legibly, and hand in your work with the solutions in the correct order. In total, you can obtain 90 points. The final grade is $1 + \#\text{points}/10$ rounded to the nearest integer. Good luck!

1. Let $\phi(y) = -\sum_{i=1}^{m} \log(y_i)$ and consider the optimization problem

$$\begin{array}{ll} \text{minimize} & \phi(y) \\ \text{subject to} & y = b - Ax. \end{array}$$

   in the variables $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$.

   (a) (5 points) Show this problem is convex.

   (b) (15 points) Derive and simplify the Lagrange dual problem.

2. Consider the optimization problem

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{i=1}^{n} u_i(x_i) \\ \text{subject to} & \displaystyle\sum_{i=1}^{n} x_i \leq B \\ & x_i \geq 0,\ i = 1, \ldots, n, \end{array}$$

   where B is a positive scalar and the functions $u_i \colon \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, n$, are convex and twice continuously differentiable.

   (a) (15 points) Suppose x^* is an optimal solution. Use the KKT conditions to show that the value of $u_i'(x_i^*)$ is the same for each i with $x_i^* \neq 0$. Here u_i' is the derivative of u_i.

   (b) (5 points) Suppose we apply the barrier method to solve this optimization problem. Why do we not need to use the Phase I problem here?

3. (15 points) Consider the function

$$F(x_1, x_2) = x_1^3 e^{x_1^2 + x_2}$$

   where we view addition, multiplication, the exponential function, and taking the second or third power as elementary functions. Show how $\nabla F(2, -4)$ is computed using reverse-mode automatic differentiation by drawing the appropriate diagrams.

4. (a) (10 points) Apply Fibonacci search to the function $f(x) = |x - e|$, where $e = 2.7182\ldots$, with initial bracket $[0, 4]$ and $k = 5$ evaluations. At which points is f evaluated and in what order? What is the final bracket you obtain?

(b) (5 points) Suppose we start golden section search with a bracket of size 1. How many function evaluations do you need (roughly) to reach machine precision; that is, after how many function evaluations is the bracket size approximately 10^{-16}?

5. Let f be a twice continuously differentiable function. Fix $\alpha \in (0, 1/2)$ and $\beta \in (0, 1)$. Let x be the current point and Δx a descent direction at x. Recall that with backtracking line search we find the step size by first trying $t = \beta$, then $t = \beta^2$, then $t = \beta^3$, etc., until the following inequality holds: z

$$f(x + t\Delta x) \leq f(x) + \alpha \nabla f(x)^{\mathsf{T}}(t\Delta x).$$

(a) (10 points) Use Taylor's theorem to show that backtracking line search terminates after finitely many iterations, and show that the objective value never increases when we use a descent method with backtracking line search to minimize f.

(b) (10 points) We now assume that there are strictly positive scalars m and M with

$$mI \preceq \nabla^2 f(x) \preceq MI$$

for all x. It can be shown that

$$f(y) \geq f(x) - \frac{1}{2m}\|\nabla f(x)\|_2^2$$

for all x and y and that at each point backtracking line search will terminate with either $t = 1$ or $t \geq \beta/M$. Use this to show that gradient descent with backtracking line search converges at least linearly with rate

$$1 - \min\{2\alpha m, 2\alpha\beta m/M\}.$$

End of test