The Chinese University of Hong Kong, Shenzhen
SDS · School of Data Science

Andre Milzarek · Summer Semester 2020

MAT 3007 – Optimization

Exercise Sheet 6

Exercise E6.1 (Multiple Choice – Descent Directions):
Answer the following multiple choice questions and decide whether the statements are true or false. Give short explanations of your answer.

a) Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable, $x^k \in \mathbb{R}^n$, and let d^k be a descent direction of f at x^k. We consider the function $\phi(\alpha) := f(x^k + \alpha d^k) - f(x^k)$, $\alpha \geq 0$.

Which of the following statements is true?

☐ 1. It holds that $\phi(0) < 0$.

☐ 2. It holds that $\phi'(0) = 0$.

☐ 3. We have $\phi(0) = 0$.

☐ 4. It holds that $\phi'(0) = \nabla f(x^k + \alpha d^k)^\top d^k$.

☐ 5. ϕ is a function from $\mathbb{R}^n$ to $\mathbb{R}$.

☐ 6. ϕ is a function from $\mathbb{R}_+$ to $\mathbb{R}$.

☐ 7. We have $\phi'(0) < 0$.

☐ 8. It holds that $\phi'(0) = \nabla f(x^k)^\top d^k$.

b) The derivative f' of the function $f(x) = \sin(x)$ is Lipschitz continuous.

☐ True. ☐ False.

c) Let $x^0 \in \mathbb{R}^n$ be an initial point with $\nabla f(x^0) \neq 0$ and consider the point $x^1 = x^0 - \nabla f(x^0)$. Then, it holds that $f(x^1) < f(x^0)$.

☐ True. ☐ False.

d) We consider the function

$$f(x) = -x_1 + 4x_2 + 10x_1^2 - 9x_1x_2 - 5x_2^2.$$

Let us set $x = (1, 1)^\top$. Which of the following statements is true?

☐ 1. $d = (0, 0)^\top$ is a descent direction of f at x.

☐ 2. $d = (0, 0)^\top$ is not a descent direction of f at x.

☐ 3. $d = (1, 1)^\top$ is a descent direction of f at x.

☐ 4. $d = (1, 1)^\top$ is not a descent direction of f at x.

☐ 5. $d = (1.5, 1)^\top$ is a descent direction of f at x.

☐ 6. $d = (1.5, 1)^\top$ is not a descent direction of f at x.

☐ 7. $d = (0, 1)^\top$ is a descent direction of f at x.

☐ 8. $d = (0,1)^\top$ is not a descent direction of f at x.

Exercise E6.2 (Multiple Choice – Step Sizes):
Answer the following multiple choice questions and decide whether the statements are true or false. Try to give short explanations of your answer.

a) We use a general descent method (Lecture 16, slide 21) to minimize a continuously differentiable function f. We consider the k-th iterate x^k and suppose that $\nabla f(x^k) \neq 0$. We choose the direction $d^k = -3\nabla f(x^k)$. Which of the following statements is true?

   ☐ 1. d^k is a descent direction.

   ☐ 2. For the step size $\alpha = 1$, we have $f(x^k + \alpha d^k) < f(x^k)$.

   ☐ 3. For the step size $\alpha = \frac{1}{14}$, it holds that $f(x^k + \alpha d^k) < f(x^k)$.

   ☐ 4. For the step size $\alpha = \frac{1}{7}$, it holds that $f(x^k + \alpha d^k) < f(x^k)$.

   ☐ 5. For the step size $\alpha = \frac{1}{3}$, we have $f(x^k + \alpha d^k) < f(x^k)$.

   ☐ 6. We can not ensure the condition $f(x^k + \alpha d^k) < f(x^k)$ for the different step sizes $\alpha \in \{1, \frac{1}{3}, \frac{1}{7}, \frac{1}{14}\}$. This depends on f.

   ☐ 7. There is $\varepsilon > 0$ such that $f(x^k + \alpha d^k) < f(x^k)$ for all $\alpha \in (0, \varepsilon]$.

b) Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable, $x^k \in \mathbb{R}^n$, and let d^k be a descent direction of f at x^k. We again consider the function $\phi(\alpha) := f(x^k + \alpha d^k) - f(x^k)$, $\alpha \geq 0$.

   For a given parameter $\gamma \in (0,1)$, a step size $\alpha \geq 0$ is said to satisfy the Armijo condition if $\phi(\alpha) \leq \gamma\alpha \cdot \phi'(0)$. Backtracking or the Armijo line search technique tries to find the largest step size $\alpha \in \{1, \sigma, \sigma^2, \sigma^3, ...\}$, $\sigma \in (0,1)$, that satisfies the Armijo condition. Exact line search determines the step size $\bar{\alpha}$ as global minimizer of ϕ: $\bar{\alpha} = \arg\min_{\alpha \geq 0} \phi(\alpha)$.

   Consider Figure 1(a). Which of the following statements is true?

   ☐ 1. Backtracking returns the step size $\alpha = 1$.

   ☐ 2. Suppose that $\sigma = \frac{1}{2}$. Then, backtracking returns the step size $\alpha = 1$.

   ☐ 3. We can not guarantee that Armijo line search returns the step size $\alpha = 1$, this depends on the choice of σ.

   ☐ 4. The Armijo condition is satisfied for all $\alpha \in [0, 1.25]$.

   ☐ 5. The exact step size $\bar{\alpha}$ lies in the interval $[0.5, 1]$.

   ☐ 6. We can not ensure that the exact step size $\bar{\alpha}$ lies in the interval $[0.5, 1]$.

   ☐ 7. The exact step size satisfies $\bar{\alpha} \geq 3$.

   Consider Figure 1(b). Which of the following statements is true?

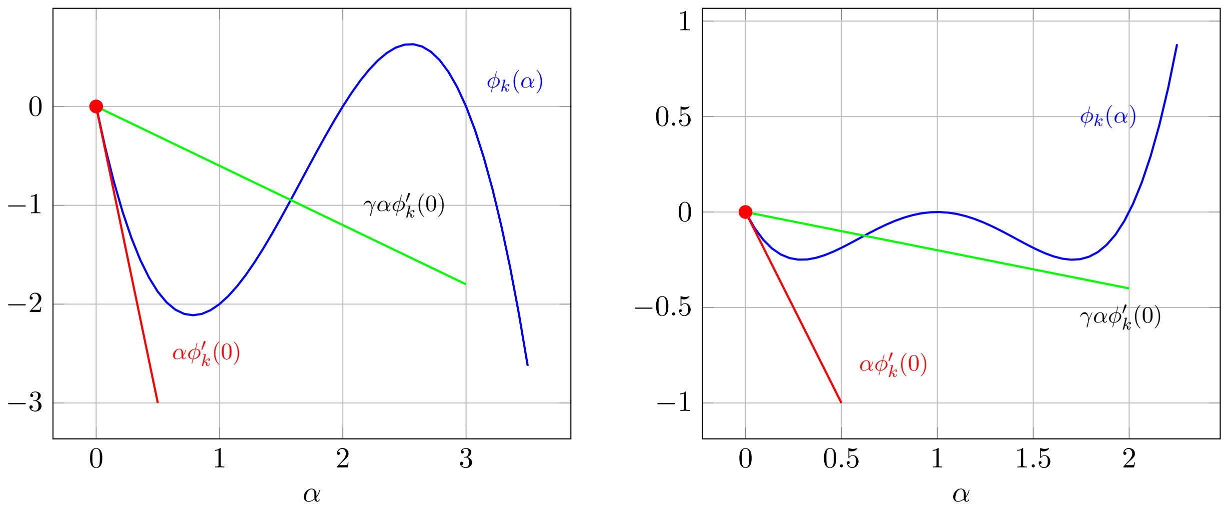
   ☐ 1. Backtracking returns the step size $\alpha = 1$.

   ☐ 2. Suppose that $\sigma = \frac{1}{2}$. Then, backtracking returns the step size $\alpha = 1$.

   ☐ 3. Suppose that $\sigma = \frac{1}{2}$. Then, backtracking returns the step size $\alpha = \frac{1}{2}$.

   ☐ 4. The Armijo condition is satisfied for all $\alpha \in [0, 0.75]$.

   ☐ 5. If $\phi(\alpha) \to \infty$ for $\alpha \to \infty$, then the exact step size $\bar{\alpha}$ lies in the interval $[0, 0.5]$.



(a) Example 1 (b) Example 2

Figure 1: Multiple choice & step sizes

☐ 6. If $\sigma \le \frac{1}{2}$, then Armijo line search returns the step size $\alpha = \sigma$.

c) If the gradient descent method with Armijo line search terminates after k steps with $\nabla f(x^k) = 0$, then x^k is a local minimum of f.

☐ True. ☐ False.

Exercise E6.3 (Convex Sets and Convex Functions):
In this exercise, we investigate convexity of sets and functions.

a) Let $A \in \mathbb{R}^{n\times n}$ be a given symmetric and positive semidefinite matrix and consider the set $\Omega := \{x \in \mathbb{R}^n : x^\top Ax \le 0\}$. Show that the set Ω is convex.

b) Show that the function $f : \mathbb{R}^2_{++} \to \mathbb{R}$, $f(x) = x_1^2 - 2x_1x_2 + x_2^2 - \ln(x_1x_2)$ is convex on $\mathbb{R}^2_{++} := \{x \in \mathbb{R}^2 : x > 0\}$.

c) Determine whether the function $f(x) = -x_1^2 - x_2^2 - 2x_3^2 + x_1x_2$ is convex or concave.

d) Let $r : \mathbb{R}^n \to \mathbb{R}$ be a norm on $\mathbb{R}^n$ and consider the mapping $f(x) := (r(x))^2$. Show that f is a convex function.

Assignment A6.1 (Convex Compositions): (approx. *20* points)
Either prove or find a counterexample for each of the following statements (you can assume that all functions are twice continuously differentiable if needed):

a) If $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are convex, then the composition $f \circ g : \mathbb{R}^n \to \mathbb{R}$, $(f \circ g)(x) = f(g(x))$ is convex.

b) Let $\Omega \subset \mathbb{R}^n$ be a convex set and suppose that $g : \Omega \to \mathbb{R}$ is convex and $f : I \to \mathbb{R}$ is convex and nondecreasing where $I \supseteq g(\Omega)$ is an interval containing $g(\Omega)$. Then, $f \circ g$ is convex.

c) In part b) assume that f is instead concave and nonincreasing. Then $f \circ g$ is convex.

d) If $f : \mathbb{R} \to \mathbb{R}$ is increasing and non-negative, then $x \mapsto xf(x)$ is convex on $\mathbb{R}_+$.

Assignment A6.2 (Convex Sets, Functions, and Problems): (approx. *30* points)
In this exercise, we study convexity of various sets and functions.

a) Verify whether the following sets are convex or not and explain your answer!

$$\Omega_1 = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : x^\top x \le t^2\},$$
$$\Omega_2 = \{x \in \mathbb{R}^n : \|x-a\|_2 \le \|x-b\|_2\}, \quad a, b \in \mathbb{R}^n,\ a \ne b.$$

b) Show that the hyperbolic set $\{x \in \mathbb{R}^2_+ : x_1x_2 \ge 1\}$ is convex, where $\mathbb{R}^2_+ := \{x \in \mathbb{R}^2 : x \ge 0\}$.

Hint: Rewrite the condition "$x_1x_2 \ge 1$" in a suitable way.

c) Let $A \in \mathbb{R}^{n\times n}$ be a positive semidefinite matrix. Show that the mapping $f(x) = \sqrt{x^\top Ax + 1}$ is convex.

d) Show that the optimization problem

$$\begin{array}{ll} \text{minimize} & -x_1 - x_2 + \max\{x_3, x_4\} \\ \text{s.t.} & (x_1 - x_2)^2 + (x_3 + 2x_4)^4 \le 5 \\ & x_1 + 2x_2 + x_3 + 2x_4 \le 6 \\ & x_1, x_2, x_3, x_4 \ge 0 \end{array}$$

is convex and write a CVX code that solves it. Use CVX to compute the optimal solution.

Assignment A6.3 (Bisection and Golden Section Method): (approx. *25* points)
We consider the optimization problem

$$\min_x\ f(x) := -\frac{1}{(x-1)^2}\left[\ln(x) - \frac{2(x-1)}{x+1}\right] \qquad \text{s.t.} \qquad x \in [1.5, 4.5].$$

a) Implement the golden section method to solve this problem and output a solution with accuracy at least 10^{-5}.

b) Consider the minimization problem

$$\min_{x\in\mathbb{R}}\ g(x) \qquad \text{s.t.} \qquad x \in [0,1],$$

where g is given by $g(x) := e^{-x} - \cos(x)$. Solve this problem using the bisection and the golden section method. Compare the number of iterations required to recover a solution in $[0,1]$ with accuracy less or equal than 10^{-5}.

Assignment A6.4 (Implementing the Gradient Descent Method): (approx. *25* points)
Implement the gradient descent method that was presented in the lecture in `MATLAB` or `Python` to solve the optimization problem:

$$\min_{x\in\mathbb{R}^2}\ f(x) = e^{1-x_1-x_2} + e^{x_1+x_2-1} + x_1^2 + x_1x_2 + x_2^2 + 2x_1 - 3x_2.$$

The following input parameters should be considered:

- $x^0 = (0,0)^\top$ – the initial point.
- `tol` $= 10^{-5}$ – the tolerance parameter. The method should stop whenever the current iterate x^k satisfies the criterion $\|\nabla f(x^k)\| \le$ `tol`.

- $\sigma, \gamma = \frac{1}{2}$ – parameters for backtracking and the Armijo condition.

The method should return the final iterate x^k that satisfies the stopping criterion.

a) Implement the gradient descent method using both exact line search and the backtracking (Armijo) line search method. Test your implementation using the given parameter choices and report the number of iterations, the final objective function value, and the point to which the methods converged.

b) Adjust your code such that the iterates x^k and gradient values $\|\nabla f(x^k)\|$ are saved and returned. Suppose that x^* is the limit of the sequence (x^k) and plot the sequences $(\|x^k - x^*\|)_k$ and $(\|\nabla f(x^k)\|)_k$ (with the iteration number k as x-axis) using a logarithmic scale. Which type of convergence can be observed?

c) Plot a figure of the solution path for each of the two line search techniques (similar to the one in the lecture slides).