Math 164 - Assignment #3, Fall 2024

- **Due Date:** 11:59 pm on October 25th, 2024

---

1. Perform two iterations leading to the minimization of

$$f(\boldsymbol{x}) = x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_1^2 + x_2^2 + 3 \tag{1}$$

using the steepest descent method with the starting point $\boldsymbol{x}^{(0)} = \mathbf{0}$. Also determine an optimal solution analytically.

2. Consider the two sequences $\{x^{(k)}\}$ and $\{y^{(k)}\}$ defined iteratively as follows:

$$x^{(k+1)} = ax^{(k)} \tag{2}$$

$$y^{(k+1)} = \left(y^{(k)}\right)^b \tag{3}$$

where $a \in \mathbb{R}, b \in \mathbb{R}, 0 < a < 1, b > 1, x^{(0)} \neq 0, y^{(0)} \neq 0$, and $\left|y^{(0)}\right| < 1$.

(a) Derive a formula for $x^{(k)}$ in terms of $x^{(0)}$ and a. Use this to deduce that $x^{(k)} \to 0$.

(b) Derive a formula for $y^{(k)}$ in terms of $y^{(0)}$ and b. Use this to deduce that $y^{(k)} \to 0$.

(c) Find the order of convergence of $\{x^{(k)}\}$ and the order of convergence of $\{y^{(k)}\}$.

(d) Calculate the smallest number of iterations k such that $\left|x^{(k)}\right| \leq c\left|x^{(0)}\right|$, where $0 < c < 1$. *Hint*: The answer is in terms of a and c. You may use the notation $\lceil z \rceil$ to represent the smallest integer not smaller than z.

(e) Calculate the smallest number of iterations k such that $\left|y^{(k)}\right| \leq c\left|y^{(0)}\right|$, where $0 < c < 1$.

(f) Compare the answer of part e with that of part d, focusing on the case where c is very small.

3. This exercise explores a zero-finding algorithm. Suppose that we wish to solve the equation $\boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}$, where

$$\boldsymbol{h}(\boldsymbol{x}) = \begin{bmatrix} 4 + 3x_1 + 2x_2 \\ 1 + 2x_1 + 3x_2 \end{bmatrix} \tag{4}$$

Consider using an algorithm of the form $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - a\boldsymbol{h}\left(\boldsymbol{x}^{(k)}\right)$, where α is scalar constant that does not depend on k.

(a) Find the solution of $\boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}$.

(b) Find the largest range of values of α such that the algorithm is globally convergent to the solution of $\boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}$.

(c) Assuming that α is outside the range of values in part b, give an example of an initial condition $\boldsymbol{x}^{(0)}$ of the form $[x_1, 0]^\top$ such that the algorithm is guaranteed not to satisfy the descent property.

4. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(\boldsymbol{x}) = \frac{3}{2}\left(x_1^2 + x_2^2\right) + (1 + a)x_1x_2 - (x_1 + x_2) + b \tag{5}$$

where a and b are some unknown real-valued parameters,

(a) Write the function f in the usual multivariable quadratic form.

(b) Find the largest set of values of a and b such that the unique global minimizer of f exists, and write down the minimizer (in terms of the parameters a and b).

(c) Consider the following algorithm:

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}-\frac{2}{5} \nabla f\left(\boldsymbol{x}^{(k)}\right)$$

Find the largest set of values of a and b for which this aigorithm converges to the global minimizer of f for any initial point $\boldsymbol{x}^{(0)}$.

5. Let $f: \mathbb{R} \rightarrow \mathbb{R}, f \in C^{3}$, with first derivative $f^{\prime}$, second derivative $f^{\prime \prime}$, and unique minimizer x^{*}. Consider a fixed-step-size gradient algorithm

$$x^{(k+1)}=x^{(k)}-\alpha f^{\prime}\left(x^{(k)}\right) .$$

Suppose that $f^{\prime \prime}\left(x^{*}\right) \neq 0$ and $\alpha=1 / f^{\prime \prime}\left(x^{*}\right)$. Assuming that the algorithm converges to x^{*}. show that the order of convergence is at least 2 .

6. Let $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ be given by $f(\boldsymbol{x})=\frac{1}{2} \boldsymbol{x}^{\top} \boldsymbol{Q} \boldsymbol{x}-\boldsymbol{x}^{\top} \boldsymbol{b}$, where $\boldsymbol{b} \in \mathbb{R}^{n}$ and $\boldsymbol{Q}$ is a real symmetric positive definite $n \times n$ matrix. Suppose that we apply the steepest descent method to this function, with $\boldsymbol{x}^{(0)} \neq \boldsymbol{Q}^{-1} \boldsymbol{b}$. Show that the method converges in one step, that is, $\boldsymbol{x}^{(1)}=\boldsymbol{Q}^{-1} \boldsymbol{b}$, if and only if $\boldsymbol{x}^{(0)}$ is chosen such that $\boldsymbol{g}^{(0)}:=\boldsymbol{Q} \boldsymbol{x}^{(0)}-b$ is an eigenvector of $\boldsymbol{Q}$.