ESE 415: Optimization (FL2024)
https://wustl.instructure.com/courses/132084

Instructor: Shen Zeng
September 30, 2024

Midterm Exam

Name: Solutions

Student ID:

There are **4 problems** that add up to 110 points.
The total for this Midterm Exam is set to 100 points,
i.e., there are 10 implicit bonus points.

Exam period: **11:30am to 12:50pm**.

(Leave blank! To be filled in by grader.)

Problem	Points	Grader
1	/ 30	
2	/ 20	
3	/ 20	
4	/ 40	
Total:		

1. (30 pts) Consider the optimization problem

$$\begin{array}{ll} \text{minimize} & f(x_1, x_2) = (x_2 - x_1^2)^2 - x_1^2 \\ \text{subject to} & |x_1| \le 10, \quad x_2 \in \mathbb{R}. \end{array}$$

(a) (10 pts) Compute both the gradient and the Hessian matrix of f.

(b) (10 pts) Determine the stationary points of f and check whether the stationary points could be (local) minimizers.

(c) (10 pts) Observing the special structure of the cost functional f, it is possible to locate the global minimum of f for the given constrained optimization problem. Find $x^\star$ and also provide the associated optimal cost $f(x^\star)$.

(a) $f(x_1, x_2) = x_2^2 - 2x_1^2 x_2 + x_1^4 - x_1^2$

$\rightarrow (\nabla f)(x) = \begin{pmatrix} -4x_1x_2 + 4x_1^3 - 2x_1 \\ 2x_2 - 2x_1^2 \end{pmatrix}$

$\rightarrow (Hf)(x) = \begin{pmatrix} -4x_2 + 12x_1^2 - 2, & -4x_1 \\ -4x_1, & 2 \end{pmatrix}$

(b) $(\nabla f)(x) \overset{!}{=} 0 \rightarrow$ solving for x gives stationary pt

from 2nd eq'n: $2x_2 - 2x_1^2 = 0 \Rightarrow x_2 = x_1^2$

plugging $x_2 = x_1^2$ into 1st eq'n yields

$\cancel{-4x_1^3} + \cancel{4x_1^3} - 2x_1 = 0 \Rightarrow \underline{x_1 = 0} \Rightarrow \underline{x_2 = x_1^2 =}$

$\rightarrow (0,0)$ only stationary point

$(Hf)(0,0) = \begin{pmatrix} -2, & 0 \\ 0, & 2 \end{pmatrix} \Rightarrow$ by 2nd order necessary condition, $x^* = (0,0)$ NOT a minimizer!

(c) Goal: keep $(x_2 - x_1^2)^2$ as zero while making $-x_1^2$ as negative as possible.

2. (20 pts) Consider the function

$$f(x_1, x_2) = \frac{1}{2}(x_1 + x_2^2)^2 .$$

Let $x^0 = (0,1)^\top$ and $v = (1,-1)^\top$ and consider the iteration $x^1 = x^0 + \gamma v$. This problem concerns an exact line search method for f in the given setting.

(a) (10 pts) Compute $(\nabla f)(x^0)$ and verify that v is a *descent direction* at x^0.

(b) (10 pts) Examine $f(x^0 + \gamma v)$ with x^0 and v as specified above and show that

$$f(x^0 + \gamma v) = \frac{1}{2}\Big(\big(\gamma - \frac{1}{2}\big)^2 + \frac{3}{4}\Big)^2.$$

Is there an optimal γ for the exact line search? Provide its value if it does exist.

(a) $f(x_1, x_2) = \frac{1}{2}(x_1 + x_2^2)^2 = \frac{1}{2}(x_1^2 + 2x_1x_2^2 + x_2^4)$

$(\nabla f)(x) = \begin{pmatrix} x_1 + x_2^2 \\ 2x_1x_2 + 4x_2^3 \end{pmatrix}$ evaluated at $\binom{0}{1}$: $\binom{1}{4}$

$(\nabla f)(x^0) \cdot v = \binom{1}{4} \cdot \binom{1}{-1} = 1 - 4 = -3 < 0$

$\Rightarrow$ v is a descent direction ✓

(b) $f(x^0 + \gamma v) = f\left(\binom{0}{1} + \binom{\gamma}{-\gamma}\right) = f\binom{\gamma}{1-\gamma} = \frac{1}{2}(\gamma + (1-\gamma)^2)^2$

$= \frac{1}{2}(\gamma + 1 - 2\gamma + \gamma^2)^2 = \frac{1}{2}(\underline{\gamma^2 - \gamma + 1})^2$ ✓

Observe: $(\gamma - \frac{1}{2})^2 + \frac{3}{4} = \gamma^2 - \gamma + \frac{1}{4} + \frac{3}{4} = \underline{\gamma^2 - \gamma + 1}$

$\Rightarrow f(x^0 + \gamma v) = \frac{1}{2}\Big(\underbrace{(\gamma - \frac{1}{2})^2}_{\geq 0} + \frac{3}{4}\Big)^2$

For $f(x^0 + \gamma v)$ to be as small as possible, we need $(\gamma - \frac{1}{2})^2 = 0$, i.e., $\underline{\gamma = \frac{1}{2}}$.

3. (20 pts) Consider the convergent sequence given by

$$z_{k+1} = \frac{z_k^2 + 1}{2z_k}.$$

Initializing with $z_0 > 0$, results in convergence of the sequence to $z^\star = 1$. This fact can be assumed as given and does not have to be verified.

Show that the convergence of (z_k) to $z^\star = 1$ is at a *quadratic* rate.

Consider: $\dfrac{|z_{k+1} - 1|}{(z_k - 1)^2} = \dfrac{\left|\frac{z_k^2+1}{2z_k} - 1\right|}{(z_k-1)^2}$

recognize as $(z_k-1)^2$

$$= \frac{\left|\frac{z_k^2+1-2z_k}{2z_k}\right|}{(z_k-1)^2} = \frac{\frac{\cancel{(z_k-1)^2}}{2z_k}}{\cancel{(z_k-1)^2}}$$

$$= \frac{1}{2z_k} \xrightarrow{k\to\infty} \frac{1}{2}$$

In Summary: $\limsup_{k\to\infty} \dfrac{|z_{k+1} - z^*|}{(z_k - z^*)^2} = \dfrac{1}{2}$

i.e. $p = 2$ & $\beta = \frac{1}{2} < \infty$

→ Quadratic convergence

4. (40 pts: 5 pts each) Determine whether the following statements are TRUE or FALSE. Explanations are not strictly required but could lead to partial credit in cases where the answer is incorrect but the explanation includes relevant and correct lines of thought.

(a) A local optimizer of a function $f : \mathbb{R} \to \mathbb{R}$ is always a global minimizer of an appropriate restriction of f to a smaller domain $A \subset R$, i.e., $g : A \to \mathbb{R}$ defined by $g(x) = f(x)$ for all $x \in A \subset R$, with restricted domain A suitably chosen.

☒ True ☐ False

(b) If $x^\star \in \mathbb{R}^n$ is a critical point of a function $f \in C^2(\mathbb{R}^n, \mathbb{R})$ and $(Hf)(x^\star) \succcurlyeq 0$, then $x^\star$ is a minimizer of f.

☐ True ☒ False

(c) If $x^\star \in \mathbb{R}^n$ is a local minimizer of $f \in C^2(\mathbb{R}^n, \mathbb{R})$, then $(\nabla f)(x^\star) = 0$ and $(Hf)(x^\star) \succcurlyeq 0$.

☒ True ☐ False

(d) If for all $v \in \mathbb{R}^n$ it holds that $(D_v f)(x^\star) = 0$, then $x^\star$ is a critical point.

☒ True ☐ False

(e) If for all $v \in \mathbb{R}^n$, the function $h_v(t) := f(x^\star + tv)$ has a global minimum at $t = 0$, then $x^\star$ is a global minimizer of f.

☒ True ☐ False

(f) For $f : \mathbb{R}^2 \to \mathbb{R}$, if $\bar{x} = 0$ is a strict global minimizer of $x \mapsto f(x, 0)$ and $\bar{y} = 0$ is a strict global minimizer of $y \mapsto f(0, y)$, then $(\bar{x}, \bar{y}) = (0, 0)$ is a minimizer of f.

☐ True ☒ False

(g) For a quadratic cost $f(x) = x^\top Q x$ with $Q \succcurlyeq 0$, the origin is a global minimum.

☒ True ☐ False

(h) For a quadratic cost functional $f(x) = x^\top Q x$ with $Q \succcurlyeq 0$, the fixed-step gradient method with the step-size chosen as in the theorem in the lecture will result in the algorithm to produce a sequence of iterates that converge to a local minimum.

☐ True ☒ False