Fall 2023 Qualifying Exam
OPTIMIZATION

September 18, 2023

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book your code number. *Do not write your name on any answer book.* On *one* of your books list the numbers of *all* the questions answered.
3. Return all answer books in the folder provided. Additional answer books are available if needed.

SPECIFIC INSTRUCTIONS:

Answer all 4 questions.

POLICY ON MISPRINTS AND AMBIGUITIES:

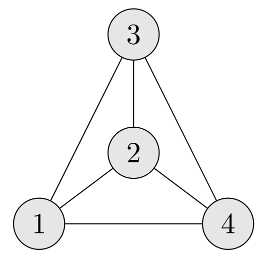
The Exam Committee tries to proofread the exam as carefully as possible. Nevertheless, the exam sometimes contains misprints and ambiguities. If you are convinced a problem has been stated incorrectly, mention this to the proctor. If necessary, the proctor can contact a representative of the area to resolve problems. In any case, you should indicate your interpretation of the problem in your written answer. Your interpretation should be such that the problem is nontrivial.

1. **Optimization Modeling**

Two-Out-Of-Three in Triangles. This sounds like a made up problem, but it arose as part of the question-writer's research. We are given an undirected graph $G = (V, E)$. The graph contains a collection of triangles $\mathcal{T} = \{T_1, T_2, \dots T_q\}$ (cliques of size 3). We would like to select a *maximum* subset of the graph's triangles so that the nodes can be partitioned into two sets $(S, V \setminus S)$ with the property that each selected triangle has *exactly* two nodes in S.

For example, consider the clique of size 4. It has four triangles:

$$\mathcal{T} = \{\{1,2,3\}, \{1,2,4\}, \{1,3,4\}, \{2,3,4\}\}.$$



If we select triangles $\{1,2,3\}$ and $\{1,3,4\}$, then by setting $S = \{1,3\}$, each of the two selected triangles has exactly two nodes in S. However, it is impossible to select three of these triangles and partition the nodes to satisfy the required property. Thus, the optimal solution to this instance has cardinality 2.

(a) Write an integer programming formulation for the general version of this problem. Be sure to clearly state the decision variables of the model and give their definition.

2. **Linear Optimization**

Let $P \subseteq \mathbb{R}^n$ be a polyhedron in standard form. Suppose that from any vertex of P we are only allowed to move to an adjacent vertex. Define the distance $d(x, y)$ between two vertices x and y of P as the *minimum* number of such moves required to reach y starting from x. The diameter $\delta(P)$ of P is then defined as the *maximum* of $d(x, y)$ over all pairs (x, y) of vertices of P.

(i) Explain why $\delta(P)$ is a *lower* bound on the *maximum* number of pivots the simplex algorithm might perform for minimizing a linear function over P starting from a given vertex, where the maximum is taken over all possible linear functions and starting vertices.

(ii) Let $c \in \mathbb{R}^n$ and let $\bar{x}$ be a vertex of P that is not optimal to the problem

$$\begin{aligned} &\min\ c^T x \\ &\text{s.t. } x \in P. \end{aligned} \tag{1}$$

Prove that $\bar{x}$ has an adjacent vertex $\tilde{x}$ with $c^T \tilde{x} < c^T \bar{x}$.

Hint: Apply the simplex method for solving (1).

(iii) Suppose P is a $(0, 1)$-polytope, i.e., a bounded polyhedron whose vertices have entries equal either to 0 or to 1. Let $\bar{x}$ be a vertex of P. Consider the problem

$$\begin{aligned} &\min\ x_n \\ &\text{s.t. } x \in P. \end{aligned} \tag{2}$$

Use problem (2) and the result in part (ii) to prove that, if $\bar{x}_n = 1$, then either $P \subseteq \{x \in \mathbb{R}^n : x_n = 1\}$ or there exists a vertex $\tilde{x}$ adjacent to $\bar{x}$ with $\tilde{x}_n = 0$.

(iv) Suppose P is a $(0, 1)$-polytope. For $i \in \{0, 1\}$ we define the polytopes in $\mathbb{R}^{n-1}$

$$P^i = \{(x_1, \ldots, x_{n-1}) : (x_1, \ldots, x_{n-1}, i) \in P\}.$$

Let x and y be two vertices of P with $x_n = y_n = i$, $i \in \{0, 1\}$. It is known that x and y are adjacent in $P \cap \{x \in \mathbb{R}^n : x_n = i\}$ —thus also in P— if and only if $(x_1, \ldots, x_{n-1})$ and $(y_1, \ldots, y_{n-1})$ are adjacent in P^i.

Prove that $\delta(P) \leq n$.

Hint: Proceed by induction on n. For the base case, let $n = 1$. Then either $P = \{0\}$ or $P = \{1\}$. For the inductive argument, consider two vertices whose distance is $\delta(P)$ and use the result shown in (iii).

3. **Integer Optimization**

For $x \in \{0,1\}^n$, we denote by $N(x)$ the set of the n vectors in $\{0,1\}^n$ that have exactly one component different from the corresponding component of x. For example, if x is the zero vector, then $N(x)$ is the set of vectors $e_1, e_2, \ldots, e_n$ of the standard basis of $\mathbb{R}^n$.

(a) Prove that, if a linear inequality $ax \geq \beta$ is not valid for the zero vector, but is valid for $e_1, e_2, \ldots, e_n$, then it is valid for every other vector in $\{0,1\}^n$.

Hint: First show that in this case we have $a_i \geq \beta > 0$ *for every* $i = 1, \ldots, n$.

(b) Let $\bar{x} \in \{0,1\}^n$. Prove that, if a linear inequality $ax \geq \beta$ is not valid for $\bar{x}$, but is valid for $N(\bar{x})$, then it is valid for every other vector in $\{0,1\}^n$.

Hint: It can follow from (a) using a change of variables, or you can prove it directly.

(c) Let O be the set of vectors in $\{0,1\}^n$ that have an odd number of ones. Let $Ax \geq b$ so that $O = \{x \in \mathbb{R}^n : Ax \geq b\} \cap \{0,1\}^n$. Prove that the number of inequalities in $Ax \geq b$ is at least 2^{n-1}.

Hint: Apply (b) to each vector $\bar{x} \in \{0,1\}^n \setminus O$.

4. **Nonlinear Optimization**

Let $f : \mathbb{R}^n \to \mathbb{R}$ be an L-smooth (possibly nonconvex) function, where $L \in (0, \infty)$. Suppose that f attains its minimum on $\mathbb{R}^n$ and let $x_* \in \arg\min_{x \in \mathbb{R}^n} f(x)$ be its arbitrary minimizer.

Consider standard gradient descent algorithm with step size $1/L$, that is: the algorithm starts with some $x_0 \in \mathbb{R}^n$, and for each $k \geq 0$ it updates its iterate as $x_{k+1} = x_k - \frac{1}{L}\nabla f(x_k)$.

(a) Prove that for each $k \geq 0$, we have

$$\min_{0 \leq i \leq k} \|\nabla f(x_i)\|_2^2 \leq \frac{2L(f(x_0) - f(x_*))}{k+1}.$$

(b) Now suppose that f satisfies the following property: for any $x \in \mathbb{R}^n$, it holds that

$$f(x_*) \geq f(x) + \nabla f(x)^\top (x_* - x). \tag{3}$$

Prove that under the property from Eq. (3), we have: for all $k \geq 1$,

$$f(x_{k+1}) - f(x_*) \leq \frac{L\|x_0 - x_*\|_2^2}{2k}.$$

(c) Alternative to Eq. (3), now suppose that f satisfies the Polyak-Łojasiewicz inequality, meaning that there exists a positive constant μ such that for all $x \in \mathbb{R}^n$,

$$\frac{1}{2\mu}\|\nabla f(x)\|_2^2 \geq f(x) - f(x_*). \tag{4}$$

Prove that in this case we have: for all $k \geq 1$,

$$f(x_k) - f(x_*) \leq \left(1 - \frac{\mu}{L}\right)^k (f(x_0) - f(x_*)).$$

(d) Given $\epsilon > 0$, bound the number of iterations that the gradient descent algorithm takes to find a point x_i such that $\|\nabla f(x_i)\|_2 \leq \epsilon$, under settings and convergence bounds from (a)–(c).