Fall 2024 Qualifying Exam
OPTIMIZATION

September 16, 2024

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. Global climate change has already had observable effects on the environment. Offshore wind power is the use of wind farms constructed in bodies of water, usually in the ocean, to harvest wind energy to generate electricity. An offshore wind farm is a collection of wind turbines placed at sea to take advantage of the strong offshore winds. These strong winds produce more electricity, but offshore wind farms are more expensive to install and operate than those on land.

   In our example, a wind farm of four turbines is being built off the east coast of Milwaukee. There is a power station on the coast where all the electricity must be transferred to be distributed to the electric grid. There is the possibility of building one or two transfer stations in the wind farm where the power from several turbines can be collected and transferred along a single cable to the shore. Each turbine can be connected to at most one transfer station. The locations of the turbines, and the possible locations and cost to build of the transfer stations are known. The maximum amount of energy that can be generated from each turbine is also known.

   There are three factors we must consider when installing the cables. First, there is a fixed cost to lay a cable on the lake floor. This cost is proportional to the distance between the two stations the cable connects. Second, we must consider how much current will flow through the cables. There are several choices of cable capacity that can be used. Connections that carry large currents need thick cables. Thick cables are more expensive than thin cables.

   Construct an optimization problem to decide which transfer stations should be built and which cables should be laid to connect the wind farm power network at a minimum cost while ensuring that the maximum amount of energy generated by the turbines can be delivered to the power station.

   (a) Describe the data (sets, parameters, etc) that you need to construct your model.

   (b) Write down all the variables that you need to determine the configuration of the power delivery network.

   (c) Write constraints to ensure that a cable can only be laid to/from a transfer station if that transfer station is built.

   (d) Write down constraints to implement the fixed cost of laying a cable.

   (e) Write down constraints to determine and limit flow from a turbine to one transfer station.

   (f) Ensure your model only chooses one type of cable for each connection.

   (g) Note one extension you might consider to improve this model.

2. **Linear Optimization**

Consider the following primal/dual pair of linear programming problems, where (P) denotes the primal and (D) denotes the dual:

$$\begin{array}{llll} \min & c'x & & \\ \text{s.t.} & Ax = b & & (P) \\ & x \geq 0 & & \end{array} \qquad \begin{array}{llll} \max & p'b & & \\ \text{s.t.} & p'A \leq c'. & & (D) \end{array}$$

We have $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and A is a $m \times n$ real matrix with linearly independent rows. We denote by A_i the i-th column of A, for $i \in \{1, \dots, n\}$. Let $\{B(1), \dots, B(m)\}$ be the basic indices and denote by B the submatrix of A consisting of the basic columns $A_{B(1)}, \dots, A_{B(m)}$. The reduced costs are $\bar{c}_i = c_i - c_B' B^{-1} A_i$ for $i \in \{1, \dots, n\}$. We assume that B yields a dual basic solution p that is dual feasible.

(i) Identify $m - n$ primal variables that are zero valued in the primal basic solution associated to B and write the primal basic solution associated to B.

(ii) Identify m dual constraints that are active in the dual basic solution associated to B and write the formula for the dual basic solution p associated to B.

(iii) What is the sign of the reduced costs? Justify your answer.

(iv) Let $d \in \mathbb{R}^m$. We want to compute the maximum stepsize θ that we can take along d so that $p + \theta d$ remains dual feasible. Prove that this stepsize is ∞ if $d' A_i \leq 0$ for all $i \in \{1, \dots, n\}$ and that otherwise it is equal to $\theta^* = \min_{\{i : d' A_i > 0\}} \frac{\bar{c}_i}{d' A_i}$.

(v) Let j be an index in $\{i : d' A_i > 0\}$ such that $\theta^* = \frac{\bar{c}_j}{d' A_j}$. Prove that $(p + \theta^* d)' A_j = c_j$.

(vi) Let $h \in \{1, \dots, m\}$. We want $B(h)$ to leave the basis. Prove that the only direction $d \in \mathbb{R}^m$ such that $(p + d)' A_{B(i)} = c_{B(i)}$ for all $i \in \{1, \dots, m\} \setminus \{h\}$ and $(p + d)' A_{B(h)} + 1 = c_{B(h)}$ is $d' = -e_h' B^{-1}$. (Note: e_h denotes the unit vector with all entries being zero, except the entry in position h, which is one).

(vii) Consider the direction $d' = -e_h' B^{-1}$. What is the change in the dual objective if we move from p to $p + d$?

3. **Integer Optimization**

Let $N = \{1, \ldots, n\}$ represent a set of locations. We consider formulations of the *hub location and assignment* problem in which a subset of size p of these nodes are to be chosen as hubs, and all remaining nodes must be assigned to a hub. Define $c_{ijkm} \geq 0$ as the cost to send material from node $i \in N$ to node $j \in N$ if node $i \in N$ is assigned to hub $k \in N$ and node $j \in M$ is assigned to hub $m \in M$. This problem can be formulated as the following quadratic integer program.

$$\begin{aligned}
\min \quad & \sum_{i\in N}\sum_{j\in N:j>i}\sum_{k\in N}\sum_{m\in N} c_{ijkm}x_{ik}x_{jm} & & \\
\text{s.t.} \quad & \sum_{k\in N} x_{kk} = p & & (1)\\
& x_{ik} \leq x_{kk}, & \forall i, k \in N (i \neq k) \quad & (2)\\
& \sum_{k\in N} x_{ik} = 1, & \forall i \in N \quad & (3)\\
& x_{ik} \in \{0, 1\}, & \forall i, k \in N \quad & (4)
\end{aligned}$$

In this formulation, the variables x_{kk} indicate whether ($x_{kk} = 1$) or not location $k \in N$ is selected as a hub, and variables x_{ik}, for $i \neq k$ indicate whether ($x_{ik} = 1$) or not location $i \in N$ is assigned to hub $k \in N$.

(a) Write an integer *linear* programming re-formulation of this problem that uses decision variables y_{ijkm} to represent the product terms $x_{ik}x_{jm}$ for $i, k, j, m \in N (j > i)$. (If you re-use any of the variables and constraints from the original formulation, you do not have to re-write them – just refer to them via their equation numbers above.)

(b) Show that the following equations are valid for your formulation:

$$\sum_{k\in N} y_{ijkm} = x_{jm}, \quad \forall i \in N, j \in N(j > i), m \in N$$

(c) Consider now an alternative mixed-integer linear formulation in which variables s_i for $i \in N$ are introduced, the objective is changed to:

$$\min \quad \sum_{i\in N} s_i$$

and the variables x_{ik} and constraints (1)–(4) are included as in part (a), and the following constraints are added:

$$\begin{aligned}
& s_i \geq \sum_{j\in N:j>i}\sum_{m\in N} c_{ijkm}(x_{ik} + x_{jm} - 1), \quad \forall i, k \in N & (5)\\
& s_i \geq 0, \quad \forall i \in N. & (6)
\end{aligned}$$

Show that this is a correct mixed-integer formulation for the problem. (Hint: Use constraints (3) and for each $i \in N$ consider separately the cases $x_{ik} = 1$ and $x_{ik} = 0$.)

(d) Let z_a^{LP} be the optimal value of LP relaxation of the formulation from part (a), and z_c^{LP} be the optimal value of the LP relaxation of this formulation. Show that $z_a^{LP} \geq z_c^{LP}$.

4. **Nonlinear Optimization**

Throughout this question, $f : \mathbb{R}^d \to \mathbb{R}$ is a convex and continuously differentiable function with a unique minimizer x^*. Given $r \geq 1$ and $\mu > 0$, we say that a function f (r, μ)-sharp if

$$f(x) - f(x^*) \geq \frac{\mu}{r} \|x - x^*\|^r, \qquad \forall x \in \mathbb{R}^d,$$

where $\|\cdot\|$ is the ℓ_2 norm.

(a) Suppose that f is $(2, \mu)$-sharp and L-smooth (i.e., ∇f is L-Lipschitz). Assume that $\frac{L}{\mu}$ is a positive integer. Consider gradient descent with stepsize $\frac{1}{L}$:

$$x_{k+1} = x_k - \frac{1}{L}\nabla f(x_k). \tag{7}$$

We proved in class that the iterates satisfy

$$f(x_k) - f(x^*) \leq \frac{L\|x_0 - x^*\|^2}{2k}, \qquad \forall k \geq 1. \tag{8}$$

Use (8) to prove the following linear convergence result:

$$f(x_{2n\frac{L}{\mu}}) - f(x^*) \leq \frac{f(x_0) - f(x^*)}{2^n}, \qquad \forall n \geq 1.$$

(b) Suppose instead that f is $(1, \mu)$-sharp and M-Lipschitz for some $M > \mu$. Assuming the optimal value $f(x^*)$ is known, we consider Polyak gradient method:

$$x_{k+1} = x_k - \eta_k \nabla f(x_k) \quad \text{where } \eta_k = \frac{f(x_k) - f(x^*)}{\|\nabla f(x_k)\|^2}. \tag{9}$$

Prove the following linear convergence result:

$$\|x_{k+1} - x^*\|^2 \leq \left(1 - \frac{\mu^2}{M^2}\right) \|x_k - x^*\|^2, \qquad \forall k \geq 0.$$

(c) Still assume that f is $(1, \mu)$-sharp and M-Lipschitz, and in addition that $\frac{\mu}{M} \leq \sqrt{\frac{1}{2}}$. Without knowing $f(x^*)$, we consider gradient descent with geometrically decaying stepsizes: starting from some x_0 with $\|x_0 - x^*\| \leq 1$,

$$x_{k+1} = x_k - \alpha_k \frac{\nabla f(x_k)}{\|\nabla f(x_k)\|} \quad \text{where } \alpha_k = \lambda q^k \tag{10}$$

with $\lambda = \frac{\mu^2}{4M^2}$ and $q = \sqrt{1 - \left(\frac{\mu}{M}\right)^2}$. Prove the following linear convergence result:

$$\|x_k - x^*\|^2 \leq \frac{\mu^2}{16M^2} q^{2k}, \qquad \forall k \geq 1.$$

(Hint: first prove $\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - 2\lambda q^k \frac{\mu}{M} \|x_k - x^*\| + \lambda^2 q^{2k}, \forall k \geq 0.$)