Spring 2020 Qualifying Exam
OPTIMIZATION

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**

The National Agricultural Bank (NAB) decides to establish an early retirement scheme for fifteen employees who are taking early retirement. These employees will retire over a period of seven years starting from the next year. To finance this early retirement scheme, the bank decides to invest in bonds for this seven-year period. The necessary amounts to cover the pre-retirement leavers are given below:

Year	1	2	3	4	5	6	7
Amount (1000 $)	1000	600	640	480	760	1020	950

These amounts have to be paid at the beginning of every year, and year 1 starts right now, so we have an immediate obligation of $1,000,000 to cover the employees taking early retirement at the start of the planning horizon.

For the investments, the bank decides to take three different types of bonds: SNCF bonds, Fujitsu bonds, and Treasury bonds. The money that is not invested in these bonds is placed as savings with a guaranteed yield of 3.2%. Information concerning the (annual) yield, durations and value of the bonds is given below:

Bond	Value ($1000)	Interest Rate	Duration
SNCF	1.0	7.0 %	5 years
Fujitsu	0.8	7.0 %	4 years
Treasury	0.5	6.5 %	6 years

It is only possible to buy integer number of bonds and the invested capital remains locked in for the total duration of the bond. The capital invested in bonds is returned at the end of the bond's duration. Each year, the interest on the savings and the interest from the purchased bonds is returned. For example, suppose that you bought 10 Fujitsu bonds. This would cost you $10 \times \$800 = \$8000$ right now, and at the beginning of years 2, 3, 4, and 5, you would receive $7\% \times \$8000 = \$560$ in interest. Also in the beginning of year 5 (end of year 4), you would receive back the initial $8000 that could be used to meet retirement obligations.

The person in charge of the retirement plan decides to buy bonds at the beginning of the first year, but not in the remaining years. How should she organize the investments in order to spend the least amount of money (right now) to cover the projected retirement plan?

(a) Write an optimization model that determines the optimal number of each bond type to buy and the amount to invest in the savings account in order to minimize the current investment and meet obligations. If you wish to write a general model, you may use the following notation:

- Time periods (years) $T = \{1, 2, \ldots, 7\}$
- Set of Bonds $B = \{$SNCF, Fujitsu, Treasury$\}$
- Obligation Amounts (\$): $h_t, \forall t \in T$
- Bond Prices/Value (\$): $p_b, \forall b \in B$
- Annual rate of return of bonds (%): $r_b, \forall b \in B$
- Duration of each bond (years): $d_b, \forall b \in B$
- Risk free rate of return (%): $\kappa = 3.2\%$.

(b) Suppose the bank could pay a bonus of 1% of the value additionally to retirees if they delay retirement by one year. For example, the bank could decide not to meet its current \$1,000,000 obligation to its retirees and instead pay these retirees $(1.01) \times \$1,000,000 = \$1,010,000$ at the start of year two. Retirees can only be deferred for 1 year. For instance, you cannot pay the people who want to retire now $(1.01) \times (1.01) \times \$1,000,000$ in two years. The people retiring at the beginning of year 7 cannot be deferred. Extend your model to determine if this saves money or not, and in which years the incentive should be offered? You may do this instance in the general model or modeling language. But be sure to clearly define any new decision variables and equations.

2. **Linear Optimization**

Consider the linear programming problem (P) defined as

$$\min_{x,M} \quad M \tag{1}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} x_j = 1 \tag{2}$$

$$\sum_{j=1}^{n} a_{ij} x_j \leq M \qquad \text{for } i = 1, \ldots, m \tag{3}$$

$$x_j \geq 0 \qquad \text{for } j = 1, \ldots, n, \tag{4}$$

where a_{ij} is the ij-th entry of a given matrix $A \in \mathbb{R}^{m \times n}$.

(a) Denote by N and y_i, $i = 1, \ldots, m$, the dual variables associated to the primal constraints (2) and (3), respectively. Write the dual (D) of (P) and bring it into a form where the dual vector $y \in \mathbb{R}^m$ is constrained to be nonnegative.

(b) Show that both (P) and (D) have a finite optimum that is attained at a vertex of their feasible sets.

(c) Prove that for each feasible solution (x, M) to (P) and for each feasible solution (y, N) to (D) it holds

$$N \leq \sum_{i=1}^{m} \sum_{j=1}^{n} y_i a_{ij} x_j \leq M.$$

(d) Let $\bar{x}$ and $\bar{y}$ be feasible solutions to (P) and to (D), respectively. Derive closed form solutions to the problems

$$\begin{aligned} \min_{x} \quad & \sum_{i=1}^{m} \sum_{j=1}^{n} \bar{y}_i a_{ij} x_j \\ \text{s.t.} \quad & \sum_{j=1}^{n} x_j = 1 \qquad \text{(P')} \\ & x \geq 0, \end{aligned} \qquad \begin{aligned} \max_{y} \quad & \sum_{i=1}^{m} \sum_{j=1}^{n} y_i a_{ij} \bar{x}_j \\ \text{s.t.} \quad & \sum_{i=1}^{m} y_i = 1 \qquad \text{(D')} \\ & y \geq 0. \end{aligned}$$

(e) Prove Von Neumann's Minimax Theorem for a two-person zero-sum game:

$$\max_{y \in Y} \min_{x \in X} \sum_{i=1}^{m} \sum_{j=1}^{n} x_j a_{ij} y_i = \min_{x \in X} \max_{y \in Y} \sum_{i=1}^{m} \sum_{j=1}^{n} x_j a_{ij} y_i,$$

where $X = \{x \in \mathbb{R}^n | x \geq 0, \sum_{j=1}^{n} x_j = 1\}$ and $Y = \{y \in \mathbb{R}^m | y \geq 0, \sum_{i=1}^{m} y_i = 1\}$.

3. **Integer Optimization**

Let $D = (V, A)$ be a complete digraph with n nodes, and denote by $\mathcal{C}$ the set of directed cycles in D with k nodes, for $k \in \{2, 3, \dots, n-2\}$. Let P be defined as the set of vectors $x \in \mathbb{R}^A$ satisfying the following constraints:

$$\sum_{ij \in A : j \neq i} x_{ij} = 1 \qquad \forall i \in V \qquad \text{(outdegree equations)}, \tag{5}$$

$$\sum_{ji \in A : j \neq i} x_{ji} = 1 \qquad \forall i \in V \qquad \text{(indegree equations)}, \tag{6}$$

$$\sum_{a \in C} x_a \leq |C| - 1 \qquad \forall C \in \mathcal{C} \qquad \text{(cycle inequalities)}, \tag{7}$$

$$0 \leq x_a \leq 1 \qquad \forall a \in A. \tag{8}$$

Denote by P_I the convex hull of the integral vectors in P.

(a) Show that that P_I is the *asymmetric traveling salesman polytope,* i.e., the convex hull of the Hamiltonian cycles in D.

(b) Derive the lifting coefficient for variable x_{13} in the inequality $x_{12} + x_{23} + x_{31} \leq 2$. I.e., find the largest coefficient α_{13} such that the inequality $x_{12} + x_{23} + x_{31} + \alpha_{13} x_{13} \leq 2$ is valid for P_I.

(c) Find all the sequential liftings of the 3-cycle inequality $x_{12} + x_{23} + x_{31} \leq 2$.

(d) Give a valid inequality for P_I that is a nonsequential lifting of $x_{12} + x_{23} + x_{31} \leq 2$. Hint: Recall the following inequalities, that are valid for P_I:

$$\sum_{ij \in A,\ i,j \in S} x_{ij} \leq |S| - 1 \qquad \forall \emptyset \subset S \subset V \qquad \text{(subtour elimination inequalities).}$$

(e) Give an upper bound on the Chvátal rank of the nonsequential lifting obtained in (c).

4. **Nonlinear Optimization**

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function with Lipschitz continuous gradient, with Lipschitz constant L and a minimizer x^*. We know that for any $x, z \in \mathbb{R}^n$ that

$$f(z) \le f(x) + \nabla f(x)^T (z - x) + \frac{L}{2}\|z - x\|^2. \tag{9}$$

(a) Show (by minimizing both sides of (9) with respect to z) that for any $x \in R^n$ we have

$$f(x) - f(x^*) \ge \frac{1}{2L}\|\nabla f(x)\|^2.$$

(b) Prove the following *co-coercivity* property: For any $x, y \in \mathbb{R}^n$, we have

$$[\nabla f(x) - \nabla f(y)]^T (x - y) \ge \frac{1}{L}\|\nabla f(x) - \nabla f(y)\|^2.$$

Hint: Apply part (a) to the following two functions:

$$h_x(z) := f(z) - \nabla f(x)^T z, \quad h_y(z) := f(z) - \nabla f(y)^T z.$$

(c) Suppose in addition to the properties of f given above, it is also *strongly* convex, with modulus of convexity $m > 0$. That is, for all $x, z \in \mathbb{R}^n$, we have

$$f(z) \ge f(x) + \nabla f(x)^T (z - x) + \frac{m}{2}\|z - x\|^2. \tag{10}$$

It is known that the function $q(x) := f(x) - \frac{m}{2}\|x\|^2$ is convex with $L - m$-Lipschitz continuous gradients. By applying the co-coercivity property of part (b) to this function q, show that the following property holds:

$$[\nabla f(x) - \nabla f(y)]^T (x - y) \ge \frac{mL}{m + L}\|x - y\|^2 + \frac{1}{m + L}\|\nabla f(x) - \nabla f(y)\|^2.$$

Fall 2019 Qualifying Exam
OPTIMIZATION

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam, your code number, and the question answered in that book. On *one* of your books list the numbers of *all* the questions answered. *Do not write your name on any answer book.*
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 during the *first hour* of the exam. 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. **Linear Programming**

Recall the definitions of 1-norm and ∞-norm of a vector $x \in \mathbb{R}^n$:

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|, \qquad \|x\|_\infty := \max\{|x_i| : i = 1, \dots, n\}.$$

Consider the following optimization problem:

$$\begin{aligned} &\min\ c^\top x \\ &\text{s.t.}\ \ \|Ax + b\|_1 \le 1. \end{aligned} \tag{1}$$

In this formulation, the decision variables are $x \in \mathbb{R}^n$, and the given data consists of $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

(a) Formulate this problem as a linear program (LP) in inequality form and prove that your LP formulation is equivalent to problem (1).
Hint 1: You may use additional variables.
Hint 2: Recall that you can show that two maximization problems (A) and (B) are equivalent by showing: (i) For any feasible solution of (A) there is a feasible solution of (B) with objective value not lower, (ii) For any feasible solution of (B) there is a feasible solution of (A) with objective value not lower.

(b) Derive the dual LP, and show that it is equivalent to the problem

$$\begin{aligned} &\max\ b^\top z - \|z\|_\infty \\ &\text{s.t.}\ \ A^\top z + c = 0. \end{aligned} \tag{2}$$

What is the relation between the optimal value of z of problem (2) and the optimal value of the variables in the dual LP derived?

(c) Let x be feasible for (1) (i.e., $\|Ax + b\|_1 \le 1$) and let z be feasible for (2) (i.e., $A^\top z + c = 0$). Using only the weak duality theorem, what can you argue about the relation between $c^\top x$ and $b^\top z - \|z\|_\infty$?

2. **Optimization Modeling**

An *economy* consists of *sectors*. You can think of a sector as a process that consumes resources at the start of the year and produces other resources at the end of the year. We can also choose an *activity level* for each sector, which determines how much consumption and production happens in each sector.

Example. Here is an example with two sectors and three resources:

- Sectors: $\{\text{house-building}, \text{road-building}\}$
- Resources: $\{\text{wheat}, \text{brick}, \text{ore}\}$

At an activity level of 1, suppose we have the following:

- House-building consumes (1 brick, 1 ore) and produces (2 wheat, 2 brick, 2 ore).
- Road-building consumes (1 wheat, 1 brick) and produces (1 ore, 2 brick).

We can think of the consumption and production levels above as *rates*. To find the actual consumption and production, we multiply by the activity level. For example, if we choose an activity level of 100 for house-building and 50 for road-building in Year 1, then our economy behaves as follows:

- Total resources consumed by all sectors: (50 wheat, 150 brick, 100 ore)
- Total resources produced by all sectors: (200 wheat, 300 brick, 250 ore)

Every year, we must choose activity levels for the sectors, with the goal of making every sector *grow*. That is, we want the activity level of each sector to increase compared to the previous year. However, we cannot grow too fast: the consumption of a given year cannot exceed the production in the previous year. For example, if we decided to triple our activity levels for Year 2, this would require consuming (150 wheat, 450 brick, 300 ore). This is not possible because we produced an insufficient amount of brick and ore (300 and 250 respectively) in the previous year.

For this problem, we will look at optimizing growth over a two-year planning period. Specifically, we assume:

- There are m resources $i = 1, \dots, m$ and n sectors $j = 1, \dots, n$.
- For sector j, resource i is consumed at a rate b_{ij} and produced at a rate a_{ij}. These are fixed quantities known ahead of time.
- Sector j has activity level $x_j^{(1)}$ in Year 1 and $x_j^{(2)}$ in Year 2. These activity levels are things we must decide on.

- Year 1 activity levels are strictly positive and Year 2 activity levels are nonnegative.
- The consumption in Year 2 must not exceed the production in Year 1.

Define the *growth rate* of sector j as $x_j^{(2)}/x_j^{(1)}$. Our objective will be to maximize the *minimum growth rate*, which is the growth rate of the sector with the smallest growth rate. This ensures that every sector is growing. Finally, here are the problems:

(a) Formulate the above as an optimization problem. That is, specify the parameters, decision variables, constraints, and objective function.

(b) The formulation from Part (a) includes the strict inequalities $x_j^{(1)} > 0$ for $j = 1, \ldots, n$. Explain why strict inequalities are generally undesirable in an optimization model. Also explain how and why the strict inequalities can be replaced by $x_j^{(1)} \geq 1$ without any loss of generality.

(c) Suppose we want to know whether it's possible to achieve a minimum growth rate of r. Explain how the problem of Part (a) can be reformulated as a linear program where r appears as a parameter.

(d) Consider maximizing the *total annual growth* $\left(\sum_{j=1}^n x_j^{(2)}\right)/\left(\sum_{j=1}^n x_j^{(1)}\right)$ instead. In this scenario, we allow activity levels in Year 1 to be zero, so long as the total activity in Year 1 is strictly positive. Formulate this problem as a linear program.

3. **Integer Optimization**

Let $S = \{v^1, \dots, v^T\} \subseteq \mathbb{R}^n$ be a finite (possibly huge) set of points and consider the optimization problem:

$$z^* = \min\ c^\top x \tag{3}$$
$$\text{s.t. } Ax \geq b \tag{4}$$
$$x \in S \tag{5}$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$. Assume that this optimization problem has a feasible solution, so that z^* is finite. For $\lambda \in \mathbb{R}^m_+$, define:

$$z(\lambda) = \lambda^\top b + \min\ (c^\top - \lambda^\top A)x$$
$$\text{s.t. } x \in S$$

Finally, define:

$$z^{LD} = \max\{z(\lambda) : \lambda \in \mathbb{R}^m_+\}$$

The number of points that will be allocated to each part when grading are given in brackets at the beginning of each part.

(a) [2 pts] Show that $z(\lambda) \leq z^*$ for any $\lambda \in \mathbb{R}^m_+$.

(b) [4 pts] Recall that conv(S) is notation for the convex hull of S. Show that

$$z^{LD} = \min\ c^\top x$$
$$\text{s.t. } Ax \geq b$$
$$x \in \text{conv}(S).$$

[Hint: Start by formulating the problem $\max\{z(\lambda) : \lambda \in \mathbb{R}^m_+\}$ as a linear program, possibly by adding additional decision variable(s).]

(c) [2 pts] Provide an example where $z^{LD} < z^*$. [Hint: this can be done with $n = 1$ and a set S containing two points.]

(d) [1 pt] Suppose $S = \{x \in \mathbb{Z}^n_+ : Dx \geq d\}$. Use the result from part (b) to show that the $z^{LD} \geq z^{LP}$, where z^{LP} is the optimal value of the linear programming relaxation of (3)-(5):

$$z^{LP} = \min\ c^\top x$$
$$\text{s.t. } Ax \geq b$$
$$Dx \geq d$$
$$x \in \mathbb{R}^n_+$$

(e) [1 pt] Using the definition of S from part (d), suppose that the matrix D is totally unimodular and d is an integer vector. Again using results from previous parts, argue that in this case $z^{LD} = z^{LP}$.

4. **Nonlinear Optimization**

Consider the inexact Newton method applied to the problem of minimizing a twice Lipschitz continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$. The search direction p_k at iteration $k = 0, 1, 2, \ldots$ satisfies the formula

$$H_k p_k + g_k = r_k,$$

where $H_k = \nabla^2 f(x_k)$, $g_k = \nabla f(x_k)$, and the residual vector r_k satisfies

$$\|r_k\| \le \eta_k \|\nabla f(x_k)\|,$$

for some $\eta_k \in [0, \bar{\eta}]$, where $\bar{\eta} \in [0, 1)$.

Assume that there is a local minimizer x^* at which second-order sufficient conditions are satisfied, that is, $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite. Assume that the smallest eigenvalue of $\nabla^2 f(x^*)$ is μ and the largest is L, with $0 < \mu < L$.

Consider the full-step version of the method, in which we define $x_{k+1} = x_k + p_k$ (so that $g_{k+1} = \nabla f(x_{k+1})$).

(a) Show that there are positive quantities $\rho > 0$ and $M_1 > 0$ and $M_2 > 0$ such that if $\|x_k - x^*\| \le \rho$, we have that

$$\|p_k\| \le M_1 \|g_k\|, \quad \frac{\|g_{k+1}\|}{\|g_k\|} \le \eta_k + M_2 \|g_k\|.$$

(You do not need to find specific values for ρ, M_1, and M_2, just explain why such quantities must exist.)

(b) Consider the Taylor series expansion

$$f(x_{k+1}) = f(x_k) + g_k^T p_k + \frac{1}{2} p_k^T H_k p_k + O(\|p_k\|^3).$$

Express the sum of the first- and second-order terms $\left(g_k^T p_k + \frac{1}{2} p_k^T H_k p_k\right)$ in terms of H_k^{-1}, g_k, and r_k.

(c) Using the result of (b), find a condition on r_k of the form $\|r_k\| \le T \|g_k\|$ (for some $T > 0$) that guarantees $g_k^T p_k + \frac{1}{2} p_k^T H_k p_k \le 0$ when x_k is sufficiently close to x^*. Express T in terms of L, μ.

(d) Similarly to part (c), find a value $\bar{T}$ such that when $\|r_k\| \le \bar{T} \|g_k\|$ and x_k is sufficiently close to x^*, p_k is a descent direction for f, that is, $g_k^T p_k < 0$.

Fall 2018 Qualifying Exam
OPTIMIZATION

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam, your code number, and the question answered in that book. On *one* of your books list the numbers of *all* the questions answered. *Do not write your name on any answer book.*
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 during the *first hour* of the exam. 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. **Linear Programming**

Recall the definitions of 1-norm and ∞-norm of a vector $x \in \mathbb{R}^n$:

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|, \qquad \|x\|_\infty := \max\{|x_i| : i = 1, \ldots, n\}.$$

Consider the following optimization problem:

$$\begin{aligned} &\min\ c^\top x \\ &\text{s.t.}\ \ \|Ax + b\|_1 \leq 1. \end{aligned} \tag{1}$$

In this formulation, the decision variables are $x \in \mathbb{R}^n$, and the given data consists of $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

(a) Formulate this problem as a linear program (LP) in inequality form and prove that your LP formulation is equivalent to problem (1).
Hint 1: You may use additional variables.
Hint 2: Recall that you can show that two maximization problems (A) and (B) are equivalent by showing: (i) For any feasible solution of (A) there is a feasible solution of (B) with objective value not lower, (ii) For any feasible solution of (B) there is a feasible solution of (A) with objective value not lower.

(b) Derive the dual LP, and show that it is equivalent to the problem

$$\begin{aligned} &\max\ b^\top z - \|z\|_\infty \\ &\text{s.t.}\ \ A^\top z + c = 0. \end{aligned} \tag{2}$$

What is the relation between the optimal value of z of problem (2) and the optimal value of the variables in the dual LP derived?

(c) Let x be feasible for (1) (i.e., $\|Ax + b\|_1 \leq 1$) and let z be feasible for (2) (i.e., $A^\top z + c = 0$). Using only the weak duality theorem, what can you argue about the relation between $c^\top x$ and $b^\top z - \|z\|_\infty$?

2. **Optimization Modeling**

An *economy* consists of *sectors*. You can think of a sector as a process that consumes resources at the start of the year and produces other resources at the end of the year. We can also choose an *activity level* for each sector, which determines how much consumption and production happens in each sector.

Example. Here is an example with two sectors and three resources:

- Sectors: $\{\text{house-building}, \text{road-building}\}$
- Resources: $\{\text{wheat}, \text{brick}, \text{ore}\}$

At an activity level of 1, suppose we have the following:

- House-building consumes (1 brick, 1 ore) and produces (2 wheat, 2 brick, 2 ore).
- Road-building consumes (1 wheat, 1 brick) and produces (1 ore, 2 brick).

We can think of the consumption and production levels above as *rates*. To find the actual consumption and production, we multiply by the activity level. For example, if we choose an activity level of 100 for house-building and 50 for road-building in Year 1, then our economy behaves as follows:

- Total resources consumed by all sectors: (50 wheat, 150 brick, 100 ore)
- Total resources produced by all sectors: (200 wheat, 300 brick, 250 ore)

Every year, we must choose activity levels for the sectors, with the goal of making every sector *grow*. That is, we want the activity level of each sector to increase compared to the previous year. However, we cannot grow too fast: the consumption of a given year cannot exceed the production in the previous year. For example, if we decided to triple our activity levels for Year 2, this would require consuming (150 wheat, 450 brick, 300 ore). This is not possible because we produced an insufficient amount of brick and ore (300 and 250 respectively) in the previous year.

For this problem, we will look at optimizing growth over a two-year planning period. Specifically, we assume:

- There are m resources $i = 1, \ldots, m$ and n sectors $j = 1, \ldots, n$.
- For sector j, resource i is consumed at a rate b_{ij} and produced at a rate a_{ij}. These are fixed quantities known ahead of time.
- Sector j has activity level $x_j^{(1)}$ in Year 1 and $x_j^{(2)}$ in Year 2. These activity levels are things we must decide on.

- Year 1 activity levels are strictly positive and Year 2 activity levels are nonnegative.
- The consumption in Year 2 must not exceed the production in Year 1.

Define the *growth rate* of sector j as $x_j^{(2)}/x_j^{(1)}$. Our objective will be to maximize the *minimum growth rate*, which is the growth rate of the sector with the smallest growth rate. This ensures that every sector is growing. Finally, here are the problems:

(a) Formulate the above as an optimization problem. That is, specify the parameters, decision variables, constraints, and objective function.

(b) The formulation from Part (a) includes the strict inequalities $x_j^{(1)} > 0$ for $j = 1, \ldots, n$. Explain why strict inequalities are generally undesirable in an optimization model. Also explain how and why the strict inequalities can be replaced by $x_j^{(1)} \geq 1$ without any loss of generality.

(c) Suppose we want to know whether it's possible to achieve a minimum growth rate of r. Explain how the problem of Part (a) can be reformulated as a linear program where r appears as a parameter.

(d) Consider maximizing the *total annual growth* $\left(\sum_{j=1}^n x_j^{(2)}\right)/\left(\sum_{j=1}^n x_j^{(1)}\right)$ instead. In this scenario, we allow activity levels in Year 1 to be zero, so long as the total activity in Year 1 is strictly positive. Formulate this problem as a linear program.

3. **Integer Optimization**

Let $S = \{v^1, \ldots, v^T\} \subseteq \mathbb{R}^n$ be a finite (possibly huge) set of points and consider the optimization problem:

$$z^* = \min\ c^\top x \tag{3}$$
$$\text{s.t. } Ax \geq b \tag{4}$$
$$x \in S \tag{5}$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$. Assume that this optimization problem has a feasible solution, so that z^* is finite. For $\lambda \in \mathbb{R}^m_+$, define:

$$z(\lambda) = \lambda^\top b + \min\ (c^\top - \lambda^\top A)x$$
$$\text{s.t. } x \in S$$

Finally, define:

$$z^{LD} = \max\{z(\lambda) : \lambda \in \mathbb{R}^m_+\}$$

The number of points that will be allocated to each part when grading are given in brackets at the beginning of each part.

(a) [2 pts] Show that $z(\lambda) \leq z^*$ for any $\lambda \in \mathbb{R}^m_+$.

(b) [4 pts] Recall that conv(S) is notation for the convex hull of S. Show that

$$z^{LD} = \min\ c^\top x$$
$$\text{s.t. } Ax \geq b$$
$$x \in \text{conv}(S).$$

[Hint: Start by formulating the problem $\max\{z(\lambda) : \lambda \in \mathbb{R}^m_+\}$ as a linear program, possibly by adding additional decision variable(s).]

(c) [2 pts] Provide an example where $z^{LD} < z^*$. [Hint: this can be done with $n = 1$ and a set S containing two points.]

(d) [1 pt] Suppose $S = \{x \in \mathbb{Z}^n_+ : Dx \geq d\}$. Use the result from part (b) to show that the $z^{LD} \geq z^{LP}$, where z^{LP} is the optimal value of the linear programming relaxation of (3)-(5):

$$z^{LP} = \min\ c^\top x$$
$$\text{s.t. } Ax \geq b$$
$$Dx \geq d$$
$$x \in \mathbb{R}^n_+$$

(e) [1 pt] Using the definition of S from part (d), suppose that the matrix D is totally unimodular and d is an integer vector. Again using results from previous parts, argue that in this case $z^{LD} = z^{LP}$.

4. **Nonlinear Optimization**

Consider the inexact Newton method applied to the problem of minimizing a twice Lipschitz continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$. The search direction p_k at iteration $k = 0, 1, 2, \ldots$ satisfies the formula

$$H_k p_k + g_k = r_k,$$

where $H_k = \nabla^2 f(x_k)$, $g_k = \nabla f(x_k)$, and the residual vector r_k satisfies

$$\|r_k\| \le \eta_k \|\nabla f(x_k)\|,$$

for some $\eta_k \in [0, \bar{\eta}]$, where $\bar{\eta} \in [0, 1)$.

Assume that there is a local minimizer x^* at which second-order sufficient conditions are satisfied, that is, $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite. Assume that the smallest eigenvalue of $\nabla^2 f(x^*)$ is μ and the largest is L, with $0 < \mu < L$.

Consider the full-step version of the method, in which we define $x_{k+1} = x_k + p_k$ (so that $g_{k+1} = \nabla f(x_{k+1})$).

(a) Show that there are positive quantities $\rho > 0$ and $M_1 > 0$ and $M_2 > 0$ such that if $\|x_k - x^*\| \le \rho$, we have that

$$\|p_k\| \le M_1 \|g_k\|, \quad \frac{\|g_{k+1}\|}{\|g_k\|} \le \eta_k + M_2 \|g_k\|.$$

(You do not need to find specific values for ρ, M_1, and M_2, just explain why such quantities must exist.)

(b) Consider the Taylor series expansion

$$f(x_{k+1}) = f(x_k) + g_k^T p_k + \frac{1}{2} p_k^T H_k p_k + O(\|p_k\|^3).$$

Express the sum of the first- and second-order terms $\left(g_k^T p_k + \frac{1}{2} p_k^T H_k p_k\right)$ in terms of H_k^{-1}, g_k, and r_k.

(c) Using the result of (b), find a condition on r_k of the form $\|r_k\| \le T \|g_k\|$ (for some $T > 0$) that guarantees $g_k^T p_k + \frac{1}{2} p_k^T H_k p_k \le 0$ when x_k is sufficiently close to x^*. Express T in terms of L, μ.

(d) Similarly to part (c), find a value $\bar{T}$ such that when $\|r_k\| \le \bar{T} \|g_k\|$ and x_k is sufficiently close to x^*, p_k is a descent direction for f, that is, $g_k^T p_k < 0$.

Fall 2017 Qualifier Exam:
OPTIMIZATION

September 18, 2017

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam, your code number, and the question answered in that book. On *one* of your books list the numbers of *all* the questions answered. *Do not write your name on any answer book.*
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 during the *first hour* of the exam. 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. **Linear Programming**

   Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ and $Q = \{x \in \mathbb{R}^n : Cx \leq d\}$ be two non-empty polyhedra.

   (a) Write a linear programming formulation that solves the problem:

   $$\min\{\|x - y\|_1 : x \in P, y \in Q\}$$

   where $\|z\|_1 = \sum_{i=1}^{n} |z_i|$ is the 1-norm.

   (b) Write the dual of the formulation you wrote in part (a).

   (c) Justify that both the primal and dual problems have an optimal solution (you may use the strong duality theorem).

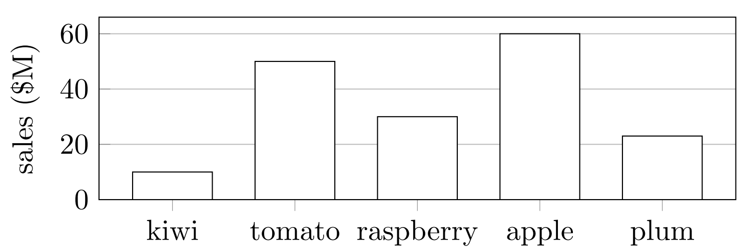
   (d) Using the above primal/dual pair of linear programs, show that if $P \cap Q = \emptyset$, then there exists a vector $p \in \mathbb{R}^n$ such that $p^\top x < p^\top y$ for all $x \in P$ and $y \in Q$. [Hint: the vector p can be defined using an optimal dual solution.]

---

2. **Modeling**

   It is common knowledge that words/objects/entities have color associations. For example, *anger* is often associated with the color red. These associations are not one-to-one mappings, e.g. *strawberry* is also associated with the color red. The associations are not unique either; *apple* can be associated with red or green, and if we're talking about the company Apple Inc., the associations will be different still!

   You are given a bar graph where each bar represents a different entity and your task is to choose colors to use for each of the bars. For example, the graph might look like the one below:



   Your task is to choose colors for the bars in the graph so that each chosen color has a strong association with the category it represents. Suppose the labels for the bars in the graph are $\{b_1, \ldots, b_m\}$ and the colors at your disposal are $\{c_1, \ldots, c_n\}$. You have access to a dataset of color-category association strengths. The data is in the form of a table:

Category \ Color	c_1	c_2	$\cdots$	c_n
b_1	a_{11}	a_{12}	$\ldots$	a_{1n}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
b_m	a_{m1}	a_{m2}	$\ldots$	a_{mn}

So a_{ij} is association strength between category b_i and color c_j. We'll assume all the data are normalized so that $0 \le a_{ij} \le 1$ and $\sum_j a_{ij} = 1$. In other words, you can think of the i^{th} row of the table as a distribution over colors for the category b_i. We'll assume $n \ge m$, so there are more colors than categories.

(a) Suppose we would like to assign the colors to the categories in a way that maximizes the total association strength of all pairs. For example, if we associate b_1 with c_6 and b_2 with c_1, then the total association strength is $a_{16} + a_{21}$. Note: you cannot assign the same color to two different categories. Formulate this optimization problem as a linear program that depends on the data $\{a_{ij}\}$. Be sure to explain why your model is correct and describe the variables, constraints, and objective function.

(b) The approach of minimizing total association strength doesn't work as well when several categories all have similar color association profiles. Instead, we'll look for a way to assign colors to categories such that the chosen pair has a high association strength and the non-chosen pairs have a low association strength. To this effect, we modify the objective such that, if color c_j is associated with category b_i, this contributes h_{ij} to the objective, where

$$h_{ij} = a_{ij} - \tau \max_{k \ne i} a_{kj}$$

Here, $\tau \ge 0$ is a parameter and the max is taken over all $k \in \{1, \dots, i-1, i+1, \dots, m\}$. The net effect of using such an objective is that b_i should be strongly associated with c_j and at the same time, c_j should not be strongly associated with any of the other $b_k's$ for $k \ne i$. How should you modify your linear program to account for this new objective?

(c) Picking a different τ in the formula for h_{ij} generally leads to a different optimal assignment of colors to categories. Prove that when $m = 2$ and $n = 2$ (two categories and two colors), all values of $\tau \ge 0$ lead to the same solution.

---

3. **Integer Optimization**

Given an undirected graph $G = (V, E)$ and a weight function $w : E \to \mathbb{R}$ on the edges, a *matching* $M \subseteq E$ is a subset of pairwise disjoint edges of G (i.e., every node of G is contained in at most one edge of M). The weight $w(M)$ of a matching M is defined as the sum of the weights of the edges in M, namely

$$w(M) = \sum_{e \in M} w_e.$$

In this setting the *maximum weight matching problem* asks to find a matching in the graph with maximum weight.

(a) Explain how the maximum weight matching problem can be solved in polynomial time if G is bipartite.

The *greedy algorithm* for the maximum weight matching problem proceeds as follows:

- Set $M := \emptyset$.
- Set $A := E$.
- While $A \neq \emptyset$, do:
  - Let e be an edge in A with highest weight.
  - Add e to M.
  - Remove from A all edges adjacent to e.
- Return M.

(b) Show an example in which the greedy algorithm does not find a matching with maximum weight.

(c) We now consider a restricted version of the maximum weight matching problem in which the weights of all edges are 1, hence a maximum matching M is simply a matching with maximum cardinality $|M|$. Notice that the greedy algorithm in this case chooses an arbitrary edge from A in every iteration.

Let OPT be the cardinality of the optimal solution and let M_g be the output of the greedy algorithm. Show that $\frac{|M_g|}{OPT} \geq 0.5$ (In other words, the matching that the greedy algorithm finds is at least half the size of an optimum one).

(Hint: Consider the relationship between the edges in a greedy matching and those in an optimal matching.)

(d) For every $n \in \mathbb{Z}_+$ give a graph with at least n vertices for which the greedy algorithm could possibly yield a matching with $\frac{|M_g|}{OPT} = 0.5$.

---

4. **Nonlinear Optimization**

(a) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice continuously differentiable function such that $f(x) \geq 0$ for all $x \in \mathbb{R}^n$. Define the function $g(x) = \frac{1}{2}f(x)^2$, and consider the following two problems:

$$\min_x f(x), \tag{F}$$

$$\min_x g(x). \tag{G}$$

Verify that the first order necessary conditions for these two problems are equivalent, that is, x^* satisfies first-order necessary conditions for (F) if and only if x^* satisfies first-order necessary conditions for (G). (Hint: Consider the case of $f(x^*) = 0$ carefully.)

(b) Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ has Lipschitz continuous gradient, that is, there is $L > 0$ such that

$$\|\nabla f(y) - \nabla f(z)\|_2 \le L\|y - z\|_2 \quad \text{for all } y, z \in \mathbb{R}^n.$$

Suppose in addition that $f(x) \ge \bar{f}$ for all x, and for some $\bar{f} > -\infty$. Consider the following short-step steepest descent method:

$$x^{k+1} = x^k - \frac{1}{L}\nabla f(x^k), \quad k = 0, 1, 2, \dots .$$

Show that the following three inequalities hold:

$$f(x^{k+1}) \le f(x^k) - \frac{1}{2L}\|\nabla f(x^k)\|_2^2, \quad k = 0, 1, 2, \dots,$$

$$\sum_{k=0}^{T-1} \|\nabla f(x^k)\|_2^2 \le 2L\left[f(x^0) - \bar{f}\right], \quad \text{for all } T \ge 1,$$

$$\min_{k=0,1,\dots,T-1} \|\nabla f(x^k)\|_2 \le \sqrt{\frac{2L[f(x^0) - \bar{f}]}{T}}, \quad \text{for all } T \ge 1.$$

Cite explicitly any theorems you use in proving these results. (Hint: Prove these three inequalities in sequence, using each one to prove the next in the sequence.)

Fall 2016 Qualifier Exam:
OPTIMIZATION

September 19, 2016

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam, your code number, and the question answered in that book. On *one* of your books list the numbers of *all* the questions answered. *Do not write your name on any answer book.*
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 during the *first hour* of the exam. 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. Suppose that

$$g(x) := \min_y c^T y \text{ subject to } Ay = b + Dx,\ y \geq 0$$

where $x \in \mathbf{R}^n$, $y \in \mathbf{R}^p$, c is a given p vector, A a given $m \times p$ matrix, b is a given m vector and D a given $m \times n$ matrix. Assume that

$$0 = \min_y c^T y \text{ subject to } Ay = 0,\ y \geq 0$$

and

$$\{z : z = Ay,\ y \geq 0\} = \mathbf{R}^m.$$

(a) Give an example of a matrix A and a vector c that satisfy the assumptions.

(b) For a given x, write down the dual of the problem defining $g(x)$.

(c) Under the assumptions, show that $g(x)$ is finite for all $x \in \mathbf{R}^n$.

(d) Under the assumptions, show that g is convex on $\mathbf{R}^n$.

(e) What is another property that g has?

---

2. Let $D = (V, A)$ be a directed graph, and let $w : A \to \mathbb{R}$ be a weight vector. The weight of a subset B of A is defined as $w(B) := \sum_{a \in B} w_a$. Consider the problem of finding a maximum-weight subset $B \subseteq A$ such that no node of V is at the same time the head of an arc in B and the tail of another arc in B.

(a) Formulate this problem as a $0, 1$ linear program.

(b) Is the polyhedron defined by the natural linear programming relaxation of your $0, 1$ linear set integral? Provide a justification of your answer either way.

(c) If the answer to part (b) is "no", give an additional class of inequalities, which is not implied by the inequalities of your natural linear programming relaxation, but which is valid for all its $0, 1$ solutions. (Hint: Think about certain directed cycles in D.)

3. The city of Sol, wanting to stay true to its name, is assessing the feasibility of using solar panels to provide power for its inhabitants during the warm summer months, when air conditioning costs are the highest. A multi-year survey of the inhabitants' energy habits as well as the hourly solar availability revealed the following aggregate data representing a typical summer day:

Hour of day:	1	2	3	$\cdots$	24
Hourly demand (MWh):	d_1	d_2	d_3	$\cdots$	d_{24}
Solar intensity:	s_1	s_2	s_3	$\cdots$	s_{24}

Unfortunately, demand doesn't always align itself with the solar cycle. For example, there is less sun in the evening, when the demand for energy is highest. The strategy is as follows:

- Purchase some number (N_p) of solar panels. Each solar panel costs C_p dollars and provides a maximum amount of MWh equal to the solar intensity. For example, 150 solar panels can provide up to $150s_1$ MWh during the first hour. Note that this is the *maximum* yield of the panels. We can always have the panels produce less if need be.
- Purchase some number (N_b) of batteries. Each battery costs C_b dollars and can be used to store up to 1 MWh of power, which may be used at a later time. Each battery has an hourly efficiency of 98%, so each hour the batteries lose 2% of their total stored energy. It is required that the batteries end the day with the same level of charge that they had when the day started.

Now, the problems:

(a) Write an optimization model (define and explain the variables, constraints, and objective) that would solve the problem of figuring out how many solar panels N_p and batteries N_b should be purchased so that total cost is minimized and the city will generate enough power to meet the hourly demands with a 5% buffer. In other words, we would like to be able to provide up to $1.05d_t$ of power at time t. What sort of optimization problem is this?

(b) Sol followed your recommendation and purchased N_b batteries and N_p solar panels. To reduce wear and tear on the batteries, Sol would like to operate their system in a way that minimizes the number of times the batteries are charged or discharged. This time, we will run our system with no buffer (we must exactly meet the hourly energy demands). How can this be accomplished? What sort of optimization problem is this?

4. Consider solving the problem $\min_{x\in\mathbb{R}^n} f(x)$, where f is Lipschitz continuously differentiable, with $\|\nabla f(x) - \nabla f(y)\|_2 \le L\|x-y\|_2$ for all x, y. As a consequence of Taylor's theorem, we have for all $x, d \in \mathbb{R}^n$ and all $\alpha > 0$ that

$$f(x+\alpha d) \le f(x) + \alpha \nabla f(x)^T d + \frac{1}{2} L \alpha^2 \|d\|^2. \quad \textbf{(Qbound)}$$

Suppose we use a line-search method to minimize f, with steps of the form $x_{k+1} = x_k + \alpha_k d_k$.

(a) Consider the Gauss-Southwell choice for search direction $d_k = -[\nabla f(x_k)]_{i_k} e_{i_k}$, where e_{i_k} is the unit vector $(0,\dots,0,1,0,\dots,0)^T$ with the 1 in position i_k, where

$$i_k := \arg\max_{i=1,2,\dots,n} |[\nabla f(x_k)]_i|.$$

Find positive values of $\bar{\epsilon}$, γ_1, and γ_2 such that this d_k satisfies conditions

$$-d_k^T \nabla f(x_k) \ge \bar{\epsilon}\|\nabla f(x_k)\|_2 \|d_k\|_2, \quad \gamma_1 \|\nabla f(x_k)\|_2 \le \|d_k\|_2 \le \gamma_2 \|\nabla f(x_k)\|_2.$$

(b) Suppose the search direction d_k is a descent direction, satisfying the conditions

$$-d_k^T \nabla f(x_k) \ge \bar{\epsilon}\|\nabla f(x_k)\|_2 \|d_k\|_2, \quad \|d_k\|_2 \ge \gamma_1 \|\nabla f(x_k)\|_2.$$

for positive $\bar{\epsilon}$ and γ_1. Suppose that we use a backtracking procedure to select α_k, where we try in turn $\alpha_k = \bar{\alpha}, \bar{\alpha}/2, \bar{\alpha}/4, \dots,$ for some $\bar{\alpha} > 0$, stopping when the following sufficient decrease condition is satisfied:

$$f(x_k + \alpha_k d_k) \le f(x_k) + c_1 \alpha_k d_k^T \nabla f(x_k), \quad \textbf{(SuffDecr)}$$

for some constant $c_1 \in (0,1)$. Find a value $\Delta_1 > 0$ such that when no backtracking is required (that is, the value $\alpha_k = \bar{\alpha}$ satisfies condition **(SuffDecr)**), we have

$$f(x_{k+1}) \le f(x_k) - \Delta_1 \|\nabla f(x_k)\|_2^2,$$

(c) Consider the same setup as in (b), but suppose now that bactracking is required, so that $\alpha_k < \bar{\alpha}$. That is, condition **(SuffDecr)** is satisfied by α_k but is violated when α_k is replaced by $2\alpha_k$, which was the previous value tried in the backtracking process. Find a positive lower bound on α_k.

(d) Considering again the same setup as in (b) and (c), and using the lower bound derived in (c), find a value $\Delta_2 > 0$ such that when backtracking is required, we have

$$f(x_{k+1}) \le f(x_k) - \Delta_2 \|\nabla f(x_k)\|_2^2,$$

Thus, find a value $\Delta > 0$ such that regardless of whether backtracking is needed or not, we have

$$f(x_{k+1}) \le f(x_k) - \Delta \|\nabla f(x_k)\|_2^2.$$

Spring 2016 Qualifier Exam:
OPTIMIZATION

February 1, 2016

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam, your code number, and the question answered in that book. *Do not write your name on any answer book.*
3. Return all answer books in the folder provided. Additional answer books are available if needed.

SPECIFIC INSTRUCTIONS:

Answer all 4 questions to the best of your abilities.

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 during the *first hour* of the exam. 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. Let k be an integer with $1 \leq k < n$. Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be defined by

$$f(x) = \sum_{i=1}^{k} x_{[i]},$$

where $x_{[i]}$ is the ith largest value in the vector x.

   (a) Show that f is convex.

   (b) Show that for any $x \in \mathbb{R}^n$, we have

$$f(x) = \max_{y} x^T y \text{ s.t.} \sum_{j=1}^{n} y_j = k, \ 0 \leq y_i \leq 1, i = 1, 2, \ldots, n.$$

   (c) For any $\alpha \in \mathbb{R}$ show that $f(x) \leq \alpha$ if and only if there exist $\lambda \in \mathbb{R}^n_+$ and $u \in \mathbb{R}$ such that

$$ku + \sum_{j=1}^{n} \lambda_j \leq \alpha, \ u + \lambda_i \geq x_i, i = 1, 2, \ldots, n.$$

   (d) Prove that a solution to

$$\max_{x} \sum_{i=1}^{n} x_i \text{ s.t. } f(x) \leq \alpha$$

   is $x_i = \frac{\alpha}{k}, i = 1, 2, \ldots, n$.

2. Given an undirected, connected, not necessarily complete graph $G = (V, E)$, with $|V| = n$, $|E| = m$, and a weight vector $w \in \mathbb{R}^m$, we are interested in formulating optimization models for finding *spanning trees* of G with specific properties. Answer the following questions:

   (a) How many edges are in each spanning tree of G?

   (b) Using (only) the decision variables $x_e \in \mathbb{Z} \ \forall e \in E$, write an integer programming model that finds a *minimum weight* spanning tree of G.

   (c) Write an integer programming formulation that maximizes the number of nodes in the spanning tree that have degree exactly 2. You may use additional decision variables. And the following jibberish may come in handy for you:

   - $\delta = 1 \Rightarrow \sum_{j \in N} a_j x_j \leq b \Leftrightarrow \sum_{j \in N} a_j x_j + M\delta \leq M + b$
   - $\sum_{j \in N} a_j x_j \leq b \Rightarrow \delta = 1 \Leftrightarrow \sum_{j \in N} a_j x_j - (m - \epsilon)\delta \geq b + \epsilon$
   - $\delta = 1 \Rightarrow \sum_{j \in N} a_j x_j \geq b \Leftrightarrow \sum_{j \in N} a_j x_j + m\delta \geq m + b$
   - $\sum_{j \in N} a_j x_j \geq b \Rightarrow \delta = 1 \Leftrightarrow \sum_{j \in N} a_j x_j - (M + \epsilon)\delta \leq b - \epsilon$

   (d) Of course, there may be *many* spanning trees that have the maximum number of degree 2 nodes. Discuss how you might find, out of all of these, the tree with the maximum number of degree 2 nodes that has minimum weight with respect to the weight vector w.

   (e) A graph is connected if and only if there is a flow of value one from each node $v \in V \setminus \{r\}$ to an arbitrary *root* node $r \in V$. Use this fact to write another formulation for the minimum-weight spanning tree problem. You may use decision variables besides x_e.

3. Let

$$S := \{x \in \{0,1\}^4 : 90x_1 + 35x_2 + 26x_3 + 25x_4 \leq 138\}.$$

(a) Show that

$$S = \{x \in \{0,1\}^4 : 2x_1 + x_2 + x_3 + x_4 \leq 3\},$$

and

$$\begin{aligned} S = \{x \in \{0,1\}^4 : 2x_1 + x_2 + x_3 + x_4 &\leq 3 \\ x_1 + x_2 + x_3 &\leq 2 \\ x_1 + x_2 + x_4 &\leq 2 \\ x_1 + x_3 + x_4 &\leq 2\}. \end{aligned}$$

(b) Can you rank these three formulations in terms of the tightness of their linear relaxations, when $x \in \{0,1\}^4$ is replaced by $x \in [0,1]^4$? Show any strict inclusion.

4. Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is a strongly convex function with uniformly Lipschitz continuous gradient. That is, there exist constants γ and L with $0 < \gamma < L$ such that

$$f(y) \leq f(x) + \nabla f(x)^T (y - x) + \frac{L}{2}\|y - x\|_2^2, \tag{1a}$$

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{\gamma}{2}\|y - x\|_2^2, \tag{1b}$$

for all $x, y \in \mathbb{R}^n$.

(a) Prove that f attains its minimizer at a unique point x^*, with $\nabla f(x^*) = 0$. (You may use the fact that if f is bounded below over a compact set C, it attains its minimum on C, that is, there is a point $x^* \in C$ such that $f(x^*) = \inf_{y \in C} f(y)$.)

(b) Consider the short-step steepest descent procedure for minimizing f:

$$x^{k+1} = x^k - \alpha \nabla f(x^k), \quad \text{where } \alpha \equiv 1/L.$$

Prove that the sequence $\{f(x^k)\}_{k=0,1,\ldots}$ converges to $f(x^*)$ at a linear rate, in particular,

$$[f(x^{k+1}) - f(x^*)] \leq \left(1 - \frac{\gamma}{L}\right)[f(x^k) - f(x^*)], \quad k = 0, 1, 2, \ldots. \tag{2}$$

Hint: Use (1a) to prove that

$$f(x^{k+1}) \leq f(x^k) - \frac{1}{2L}\|\nabla f(x^k)\|^2,$$

and use (1b) with $x = x^k$ to show that

$$f(x^*) \geq f(x^k) - \frac{1}{2\gamma}\|\nabla f(x^k)\|^2.$$

(c) Prove that the same convergence rate holds if we obtain x^{k+1} by an exact line search along $-\nabla f(x^k)$, that is,

$$x^{k+1} = x^k - \alpha_k \nabla f(x^k), \quad \text{where } \alpha_k = \arg\min_{\alpha} f(x^k - \alpha \nabla f(x^k)).$$

Spring 2015 Qualifier Exam:
OPTIMIZATION

February 2, 2015

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam, your code number, and the question answered in that book. On *one* of your books list the numbers of *all* the questions answered. *Do not write your name on any answer book.*
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 during the *first hour* of the exam. 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. (a) Consider the linear programming problem of minimizing $c^T x$ over a bounded polyhedron $P \subseteq \mathbf{R}^n$, subject to additional equality constraints $a_i^T x = b_i$, $i = 1, 2, \ldots, L$. Assume that the feasible set for this problem is nonempty. Show that there exists an optimal solution that is a convex combination of at most $L + 1$ extreme points of P.

   (b) Consider a modified version of the problem in (a) which is the same except for the addition of M further *inequality* constraints $c_j^T x \le d_j$, $j = 1, 2, \ldots, M$. Show that the solution is again expressible as a convex combination of the extreme points of P, involving at most $\min(L + M + 1, K)$ extreme points, where K is the number of extreme points in P.

2. A farmer went to market and purchased a fox, a goose, and a bag of beans. On his way home, the farmer came to the right bank of a river and rented a boat. But in crossing the river by boat, the farmer could carry only himself and a single one of his purchases - the fox, the goose, or the bag of the beans. If left together, the fox would eat the goose, or the goose would eat the beans. The farmer's challenge was to carry himself and his purchases to the left bank of the river, leaving each purchase intact.

   We formulate this as a binary programming problem the objective of which is to minimize the time required to complete the crossing. Here are the set and variable declarations for the model:

```
set     i       Items                    /fox, goose, beans/,
        t       Crossings                /t0*t5/,
        a(i,i)  Items to be kept apart   /fox.goose, goose.beans/;

alias (i,j);

variable        TIME    Objective function (crossings to completion);

binary
variables       L(i,t)  Item i is on the left bank at the start of round t,
                R(i,t)  Item i is on the right bank at the start of round t,
                X(i,t)  Item i is in boat from right to left bank during round t,
                Y(i,t)  Item i is in boat from left to right bank during round t,
                B(i,t)  Item i is left on right bank during round t,
                Z(t)    Transportation is still in progress during round t;
```

   The GAMS code produces a solution report as follows:

```
----     87 PARAMETER report  Summary report

                      R           L           B           X           Y
```

t0.fox	1		1		
t0.goose	1			1	
t0.beans	1		1		
t1.fox	1			1	
t1.goose		1			1
t1.beans	1		1		
t2.fox		1			
t2.goose	1		1		
t2.beans	1			1	
t3.fox		1			
t3.goose	1			1	
t3.beans		1			
t4.fox		1			
t4.goose		1			
t4.beans		1			
t5.fox		1			
t5.goose		1			
t5.beans		1			

Write down the formulation for this problem in terms of the data objects and variables above. You can express the objective and constraints in algebraic form, or in `GAMS` code, whichever you prefer. Your model should include constraints that determine the following conditions:

(a) Which items are on the left bank at the start of cycle $t + 1$?

(b) Which items are on the right bank at the start of cycle $t + 1$?

(c) Which items are left behind on the right bank during cycle t?

(d) Which items cannot be left on the right bank during cycle t?

(e) Until the farmer is finished, selected items cannot be left together on the left bank.

3. Let n be a positive integer, and let b be a rational number with $0 < b < n$. Consider the mixed-integer set:

$$X = \Big\{x \in \mathbb{R}^n_+, y \in \{0,1\}^n : \sum_{j=1}^{n} x_j \le b,\ x_j \le y_j,\ j = 1, \ldots, n\Big\}.$$

Let $f = b - \lfloor b \rfloor$. Define the inequalities:

$$\sum_{j \in S} x_j + f \sum_{j \in S} (1 - y_j) \le b, \quad \forall S \subseteq \{1, \ldots, n\} \text{ with } |S| = \lceil b \rceil. \tag{1}$$

(a) Show that the inequalities (1) are valid for X. (Hint: consider two cases, either $y_j = 1$ for all $j \in S$, or $y_k = 0$ for some $k \in S$.)

(b) Consider the special case with $n = 2$ and $b = 0.5$. Write down one of the inequalities (1) for this case, and show that it is facet-defining.

(c) Considering again the case of general n, given a vector $(\bar{x}, \bar{y}) \in \mathbb{R}^{2n}_{+}$, formulate a binary integer program that can be used to find a violated inequality of the form (1) if one exists, or else shows no violated inequality exists.

(d) Explain why the binary integer program in part (c) can be solved as a linear program.

(e) Use the results of parts (c) and (d) to derive a compact linear programming formulation, in an extended variable space, which models the set of $(x, y) \in \mathbb{R}^{2n}_{+}$ that satisfy the inequalities (1). The number of constraints and variables in this formulation should be linear in n. (Specifically, if $P = \{(x, y) \in \mathbb{R}^{2n}_{+} : (1)\}$, you should define a polyhedron Q in a higher variable space, having $O(n)$ variables and inequalities, such that $\text{proj}_{(x,y)}(Q) = P$.)

4. Consider the continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, whose gradient has Lipschitz constant L (that is, $\|\nabla f(y) - \nabla f(z)\| \leq L\|y - z\|$ for all $y, z \in \mathbb{R}^n$). Suppose that f is bounded below on $\mathbb{R}^n$ and in fact that f has the coercive property that $f(x) \to \infty$ as $\|x\| \to \infty$. Consider an algorithm for minimizing f that generates an iteration sequence $\{x^k\}_{k=0,1,2,\ldots}$ according to the following formula:

$$x^{k+1} = x^k + \alpha_k p^k, \quad k = 0, 1, 2, \ldots,$$

where $\alpha_k \geq 0$ is selected to be a global minimizer of f along the direction p^k from x^k. The direction p^k has the form

$$p^k = - \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{\partial f(x^k)}{\partial x_{i_k}} \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

where the nonzero element occurs in the i_k position, and the index i_k is chosen in such a way that

$$\left| \frac{\partial f(x^k)}{\partial x_{i_k}} \right| \geq \frac{1}{10} \|\nabla f(x^k)\|_\infty = \frac{1}{10} \max_{j=1,2,\ldots,n} \left| \frac{\partial f(x^k)}{\partial x_j} \right|.$$

(a) Show that for all $\alpha > 0$, we have

$$f(x^k + \alpha p^k) \leq f(x^k) + \alpha (p^k)^T \nabla f(x^k) + L\alpha^2 \|p^k\|_2^2.$$

(b) Show that $\lim_{k\to\infty} \nabla f(x^k) = 0$.

(c) Show that all accumulation points of this algorithm are stationary points of f.

Spring 2014 Qualifier Exam:
OPTIMIZATION

February 3, 2014

GENERAL INSTRUCTIONS:

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

SPECIFIC INSTRUCTIONS:

Answer 4 out of 5 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 during the *first hour* of the exam. 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. You have a set J of jobs that must be scheduled within a set of time periods $\mathcal{T} = \{1, \ldots, T\}$. Each job has an integer processing time $p_j > 0$. When being processed, jobs use capacity from a set I of machines. In particular, when job $j \in J$ is being processed it requires $a_{ij} > 0$ units of the capacity of machine $i \in I$. At any point in time, the total available capacity of machine $i \in I$ is denoted by $b_i > 0$. Jobs can be processed simultaneously, provided the machine capacity constraints are not exceeded, but jobs cannot be interrupted (i.e., once a job starts it will be in process for p_j consecutive time periods). In the following questions, you should use (at least) the following "time-indexed" decision variables to determine the start-time of the jobs.

   - $x_{jt} = 1$ if job $j \in J$ starts at time $t \in \{1, 2, \ldots, T - p_j\}$, $x_{jt} = 0$ otherwise.

   (a) Write a linear integer programming formulation to minimize the sum of start times of the jobs.

   (b) Suppose now each job $j \in J$ has an earliest start-time $r_j \geq 1$ and a latest completion time $D_j \leq T$. One possible way to model these restrictions is with the constraints:

   $$r_j \leq \sum_{t \in \mathcal{T}} t x_{jt} \leq D_j - p_j, \quad j \in J. \tag{1}$$

   However, these constraints would lead to a weak linear programming relaxation. Provide a *different* way to model the start-time and completion time restrictions, and argue why your model is preferred.

   (c) Now suppose that you find it is not feasible to schedule all the jobs so that they are completed by their latest completion time D_j. Thus, you wish to relax this constraint, and instead penalize lateness in the objective. Specifically, if job $j \in J$ completes at time $t > D_j$, then this will be penalized by $(D_j - t)^2$. Modify your model to replace the objective with the objective of minimizing the sum of penalties. (Your model must remain an integer *linear* program.)

   (d) Now consider a pair of jobs j and k that have a precedence relationship: job k cannot begin processing until job j completes its processing. This can be modeld with the constraint

   $$\sum_{t \in \mathcal{T}} t x_{jt} + p_j \leq \sum_{t \in \mathcal{T}} t x_{kt}. \tag{2}$$

   However, this constraint again leads to a weak linear programming relaxation. Provide an alternative model for this restriction that would lead to a better relaxation. (Hint: the formulation should involve inequalities that only have coefficients 0 or 1 on the decision variables). You *do not* have to provide a proof that the LP relaxation of your formulation is better than (2).

2. Consider the parametric linear program PLP(θ):

$$\begin{array}{rllllllll} z(\theta) = \min & 4x_1 & +2x_2 & & +x_4 & & & \\ \text{s.t.} & x_1 & & -x_3 & +x_4 & & & =\theta \\ & x_1 & +x_2 & & & & & =\theta \\ & x_1 & & -x_3 & & -x_5 & & =1 \\ & & x_2 & & & & +x_6 & =1 \\ & x_1, & x_2, & x_3, & x_4, & x_5, & x_6 & \geq 0 \end{array}$$

(a) Write the dual linear program of PLP(θ).

(b) Determine $z(\theta)$ $\forall \theta \in \mathbb{R}$.

(c) Is $z(\theta)$ a convex function, a concave function, or neither?

(d) Is your answer from part (c) generally true? That is, for the general parametric linear program, where the right-hand side vector b is a treated as a parameter, is the value function $z(b)$,

$$z(b) \stackrel{\text{def}}{=} \min_{x \geq 0} \{c^\top x \mid Ax = b\},$$

a convex function of b, a concave function of b, or neither? Provide a proof of your claim.

(e) Let $x_4^*(\theta)$ be an optimal value for the decision variable x_4^* as a function of the parameter θ in PLP(θ). (For PLP(θ), $x_4^*(\theta)$ is a single-valued mapping (a function)). Determine $x_4^*(\theta)$ $\forall \theta \in \mathbb{R}$.

(f) Is $x_4^*(\theta)$ a convex function, a concave function, or neither?

---

3. Let $S = \{x^1, x^2, \ldots, x^t\}$ be a finite set of points in $\mathbf{R}^n$, and let $P = \{x \in \mathbf{R}^n : Ax \leq b\}$ be a non-empty polytope where A is an $m \times n$ matrix and $b \in \mathbf{R}^n$. Let $\hat{x} \in \mathbf{R}^n$ be given. Formulate a linear program that determines whether or not $\hat{x} \in \text{conv}(S \cup P)$, and if not identifies a valid inequality for $\text{conv}(S \cup P)$ that is violated by $\hat{x}$. State how the linear program answers the question and prove that it provides a correct answer (this requires proving something in its two possible statements: $\hat{x} \in \text{conv}(S \cup P)$ and $\hat{x} \notin \text{conv}(S \cup P)$).

---

4. Assume that X is a bounded polyhedron. Let $x^0 \in X$ be given and for $k \geq 0$ let $\bar{x}^k$ be defined as an extreme point of X satisfying

$$\bar{x}^k \in \arg\min_{x \in X} \nabla f(x^k)^T (x - x^k),$$

and suppose that x^{k+1} is a stationary point of the optimization problem

$$\min_{x \in X^k} f(x),$$

where X^k is the convex hull of x^0 and the extreme points $\bar{x}^0, \ldots, \bar{x}^k$.

(a) Write down the definition of X^k explicitly.

(b) Define the notion of a stationary point for $\min_{x \in X} f(x)$.

(c) Under what conditions is $\bar{x}^k$ well defined. Identify a (known) algorithm that can determine $\bar{x}^k$.

(d) Show that there exists a finite integer k such that the above method finds a stationary point of f over X.

---

5. (a) Consider the following semidefinite program, in standard form:

$$\min_{X \in S\mathbb{R}^{2\times 2}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \bullet X \text{ subject to } \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \bullet X = 1, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \bullet X = 0, \quad X \succeq 0.$$

Write down the dual of this problem, and find the complete primal and dual solution sets, together with the optimal objective value for both problems.

(b) Consider the following semidefinite program, in standard form:

$$\min_{X \in S\mathbb{R}^{2\times 2}} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \bullet X \text{ subject to } \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \bullet X = 0, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \bullet X = 2, \quad X \succeq 0.$$

Write down the dual of this problem, and find the complete primal and dual solution sets, together with the optimal objective value for both problems.

(c) Give sufficient conditions on a primal-dual pair of semidefinite programs that guarantees that the solutions sets of both problems are *nonempty* and *bounded* and have equal objective value. Are these conditions satisfied by the problems in parts (a) and (b)?