Fall 2013 Qualifier Exam: OPTIMIZATION

September 16, 2013

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. In this problem, we will attempt to solve the quadratic program:

$$\min_{x \geq 0} f(x) \stackrel{\text{def}}{=} \frac{1}{2} x^T M x + (r - ae)^T x \tag{P}$$

where we are given parameters $a \in \mathbb{R}_+, r \in \mathbb{R}^n$, $e \in \mathbb{R}^n$ is a vector of all ones, and the matrix M has the value 2 for its diagonal elements, and 1 everywhere else. Specifically, $M = J + I$, where $J = ee^T \in \mathbb{R}^{n \times n}$ is the matrix of all ones, and $I \in \mathbb{R}^{n \times n}$ is the identity matrix.

(a) Show that $M = J + I$ is positive-definite.

(b) Write the KKT conditions for (P).

(c) Assume that $r_1 \leq r_2 \leq \ldots \leq r_n$. Establish the following monotonicity property for an optimal solution x^* to (P):

$$x_i^* = 0 \Rightarrow x_j^* = 0 \; \forall j > i.$$

(d) Show that

$$M^{-1} = \frac{1}{n+1} \begin{bmatrix} n & -1 & \cdots & -1 \\ -1 & n & \ldots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n \end{bmatrix}$$

(e) Let $\mathcal{R}$(P) be the problem (P) without the non-negativity constraints $x \geq 0$. Write a closed-form solution (formula) for an optimal solution to $\mathcal{R}$(P).

(f) Using all these results, determine a *closed-form* solution for x^*, an optimal solution to (P).

Answer:

(a) $M = J + I$ is strictly positive definite because the quadratic form

$$x^T (J + I) x = x^T J x + x^T x$$

is positive for all $x \neq 0$.

(b) The KKT system for (P) is

$$\begin{aligned} x &\geq 0 \\ Mx + r - ae &= \mu \\ \mu &\geq 0 \\ \mu^T x &= 0 \end{aligned} \tag{1}$$

(c) Proof is by contradiction. Suppose $x_i = 0$ and $x_j > 0$ for some $j > i$ and some optimal solution x. By complementary slackness, $\mu_j = 0$, and the ith and jth rows of the equations (1) are

$$\left(\sum_{k=1,k\neq i}^{n} x_k\right) + r_i - a = \mu_i$$
$$\left(\sum_{k=1,k\neq i}^{n} x_k\right) + x_j + r_i - a = 0$$

Substituting $\mu_i - r_i + a = \sum_{k=1,k\neq i}^{n} x_k$ into the jth row gives

$$\mu_i - r_i + a + x_j + r_j - a = 0, \text{ or}$$
$$u_i + x_i = r_i - r_j \tag{2}$$

The left hand side of (2) is > 0, and the right hand side is ≤ 0, a contradiction.

(d) It is easy to show that $MM^{-1} = I$

(e) If the problem is unconstrained, then the optimal solution occurs at the point x where $\nabla f(x) = 0$. This implies that

$$Mx + r - ae = 0, \text{ or } x = M^{-1}(ae - r)$$

If you do the multiplication using the closed form for M^{-1}, you get that

$$x_i^* = \frac{1}{n+1}\left(a - (n+1)r_i + \sum_{j=1}^{n} r_j\right).$$

(f) Using the monotonicty property proved in [c], we characterize the solution to the KKT system of (P) in terms of p, the number of components of x_i^* that are positive. Writing (1) gives

$$\left[\sum_{j=1}^{p} x_j^* + x_i^*\right] + r_i - a = 0 \quad \forall i \leq p \tag{3}$$
$$\left[\sum_{j=1}^{p} x_j^*\right] + r_i - a = \mu_i^* \quad \forall i > p \tag{4}$$

The equation (3) are p equations in p unknowns, and they have the closed form solution

$$x_i^* = \frac{1}{(p+1)}\left[a - (p+1)r_i + \sum_{j=1}^{p} r_j\right] \quad \forall i \leq p \tag{5}$$

One can also show then by substituting (5) into (4) that

$$\mu_i^* = \frac{-1}{p+1}\left[a - (p+1)r_i + \sum_{j=1}^{p} r_j\right] \quad \forall i > p \tag{6}$$

Define

$$S_j = \sum_{i=i}^{j} r_i$$
$$\kappa_i(p) = a - (p+1)r_i + S_p$$

Thus what we have shown is that for x_i^*, μ_i^* to define an optimal solution, we need to choose the number of positive components p^* of x to be such that

$$\kappa_i(p^*) \geq 0 \quad \forall i \leq p^*$$
$$\kappa_i(p^*) \leq 0 \quad \forall i > p^*.$$

$\diamond$

---

2. Consider the multi-item lot-sizing problem formulated below:

$$\min_{x,y,s} \sum_{i=1}^{m}\sum_{t=1}^{n}(p_{it}x_{it} + h_{it}s_{it} + f_{it}y_{it}) \tag{7}$$
$$\text{s.t.} \sum_{i=1}^{m} C_{it}y_{it} \leq B_t, \quad t = 1,\ldots,n \tag{8}$$
$$s_{i,t-1} + x_{it} - s_{it} = d_{it}, \quad t = 1,\ldots,n,\ i = 1,\ldots,m \tag{9}$$
$$x_{it} \leq D_{itn}y_{it}, \qquad t = 1,\ldots,n,\ i = 1,\ldots,m \tag{10}$$
$$s_{it} \geq 0, x_{it} \geq 0, y_{it} \in \{0,1\},\ t = 1,\ldots,n,\ i = 1,\ldots,m. \tag{11}$$

where $s_{i0} = 0$ is a constant, all data (B_i, C_{it}, d_{it}, p_{it}, h_{it}, f_{it}) is positive and $D_{it\ell} = \sum_{j=t}^{\ell} d_{ij}$ for any $1 \leq t \leq \ell$. For each item i, define

$$X_i = \{x_i \in \mathbf{R}_+^n, y_i \in \{0,1\}^n, s_i \in \mathbf{R}_+^n : (9) - (10).\}$$

For each i, the ℓ-S inequalities

$$\sum_{j\in S} x_{ij} \leq \sum_{j \in S} D_{ij\ell}y_{ij} + s_{i\ell}, \quad \forall S \subseteq \{1,\ldots,\ell\}, \ell \in \{1,\ldots,n\} \tag{12}$$

are valid for $\text{conv}(X_i)$. Furthermore, it is known that the convex hull of X_i is given by:

$$\text{conv}(X_i) = \big\{x_i \in \mathbf{R}_+^n, y_i \in [0,1]^n, s_i \in \mathbf{R}_+^n : (9) - (10), (12)\big\}.$$

Let z_1^{LP} be the value of the LP relaxation of (7) – (11) (i.e., where $y_{it} \in \{0,1\}$ is replaced by $y_{it} \in [0,1]$) and let z_2^{LP} be the value of the LP relaxation of (7) – (11) augmented with the additional inequalities (12).

(a) Write down the Lagrangian relaxation problem obtained by relaxing constraints (8) and the associated Lagrangian dual problem.

(b) Can the Lagrangian relaxation problem from part (a) be solved by solving a set of smaller subproblems? If so, describe in words what these subproblems represent. If not, explain why not.

(c) Let w_1^{LD} be the optimal value of the Lagrangian dual associated with the relaxation you wrote in part (a). How does w_1^{LD} compare to z_1^{LP}? Choose one answer and explain. Note: here and in the following sub-questions, you may apply (without proof) known Lagrangian duality theory in your explanation.

i. $w_1^{LD} = z_1^{LP}$.
ii. $w_1^{LD} \geq z_1^{LP}$ and inequality might be strict.
iii. $w_1^{LD} \leq z_1^{LP}$ and inequality might be strict.
iv. Based on the given information either $w_1^{LD} > z_1^{LP}$ or $w_1^{LD} < z_1^{LP}$ is possible.

(d) How does w_1^{LD} compare to z_2^{LP}? Choose one answer and explain:

i. $w_1^{LD} = z_2^{LP}$.
ii. $w_1^{LD} \geq z_2^{LP}$ and inequality might be strict.
iii. $w_1^{LD} \leq z_2^{LP}$ and inequality might be strict.
iv. Based on the given information either $w_1^{LD} > z_2^{LP}$ or $w_1^{LD} < z_2^{LP}$ is possible.

(e) Now consider an alternative Lagrangian relaxation problem in which constraints (9) are relaxed (and (8) are not relaxed). Let w_2^{LD} be the optimal value of the associated Lagrangian dual problem. How does w_2^{LD} compare to z_2^{LP}? Choose one answer and explain:

i. $w_2^{LD} = z_2^{LP}$.
ii. $w_2^{LD} \geq z_2^{LP}$ and inequality might be strict.
iii. $w_2^{LD} \leq z_2^{LP}$ and inequality might be strict.
iv. Based on the given information either $w_2^{LD} > z_2^{LP}$ or $w_2^{LD} < z_2^{LP}$ is possible.

Answer:

(a) Let $\pi \in \mathbf{R}_+^n$ be the multipliers for relaxing (8). We obtain the Lagrangian relaxation problem:

$$\begin{aligned} L_1(\pi) = \min \ & \sum_{i=1}^{m}\sum_{t=1}^{n}(p_{it}x_{it} + h_{it}s_{it} + f_{it}y_{it}) + \sum_{t=1}^{n}\pi_t\Big(\sum_{i=1}^{m} C_{it}y_{it} - B_t\Big) \\ \text{s.t. } & (9) - (11). \end{aligned}$$

The associated Lagrangian dual problem is:

$$\max\{L_1(\pi) : \pi \in \mathbf{R}_+^n\}$$

(b) Yes, the problem decomposes into a separate single-item lot sizing problems for each item.

(c) $w_1^{LD} \geq z_1^{LP}$ and inequality might be strict. Explanation: This inequality holds for any Lagrangian dual problem. By Lagrangian duality theory, we know that:

$$w_1^{LD} = \min\{\sum_{i=1}^{m}\sum_{t=1}^{n}(p_{it}x_{it} + h_{it}s_{it} + f_{it}y_{it}) : (8), (x_i, s_i, y_i) \in \text{conv}(X_i), i = 1, \ldots, m\}.$$

Therefore, inequality might be strict because the basic LP relaxation of the single-item sets X_i does not define the convex hull conv(X_i)..

(d) $w_1^{LD} = z_2^{LP}$. The LP relaxation with the addition of the ℓ-S inequalities *does* define the convex hull of the single-item lot sizing sets, and so the bounds are identical.

(e) Lagrangian duality theory shows that:

$$w_2^{LD} = \min\{\sum_{i=1}^{m}\sum_{t=1}^{n}(p_{it}x_{it} + h_{it}s_{it} + f_{it}y_{it} : y_{\cdot,t} \in \text{conv}(Y_t), (9) - (10), x \geq 0, s \geq 0\}$$

where for each t, $Y_t = \{y_{\cdot,t} \in \{0,1\}^m : \sum_{i=1}^m C_{it}y_{it} \leq B_t\}$ is the single period knapsack set. The constraints $y_{\cdot,t} \in \text{conv}(Y_t)$ are potentially stronger than (8) plus $y \in [0,1]^m$, and so it's possible that $w_2^{LD} > w_1^{LD}$. On the other hand, the constraints $(x_i, s_i, y_i) \in \text{conv}(X_i)$ are potentially stronger than (9) - (10) and $x_i \geq 0$, $s_i \geq 0$, $y_i \in [0,1]^m$ and so the reverse inequality is also possible. (Not asked for, but the obvious formulation better than both of them is to also include the ℓ-S inequalities in the second Lagrangian dual.)

◇

---

3. Let f be twice continuously differentiable. Suppose that x^* is a local minimum such that for all x in an open sphere S centered at x^*, we have, for some $m > 0$,

$$m\,\|d\|^2 \leq d^T\nabla^2 f(x)d, \quad \forall d \in \mathbb{R}^n.$$

Show that for every $x \in S$, we have

$$\|x - x^*\| \leq \frac{\|\nabla f(x)\|}{m}, \quad f(x) - f(x^*) \leq \frac{\|\nabla f(x)\|^2}{m}.$$

Hint: use the relation

$$\nabla f(y) = \nabla f(x) + \int_0^1 \nabla^2 f(x + t(y - x))(y - x)dt.$$

Answer: We have

$$\nabla f(x) - \nabla f(x^*) = \int_0^1 \nabla^2 f(x^* + t(x - x^*))(x - x^*)dt$$

and since $\nabla f(x^*) = 0$, we obtain

$$(x - x^*)^T \nabla f(x) = \int_0^1 (x - x^*)^T \nabla^2 f(x^* + t(x - x^*))(x - x^*) dt \geq m \int_0^1 \|x - x^*\|^2 \, dt.$$

Using the Cauchy-Schwartz inequality $(x - x^*)^T \nabla f(x) \leq \|x - x^*\| \, \|\nabla f(x)\|$, we have

$$m \int_0^1 \|x - x^*\|^2 \, dt \leq \|x - x^*\| \, \|\nabla f(x)\| \,,$$

and

$$\|x - x^*\| \leq \frac{\|\nabla f(x)\|}{m}.$$

Now define for all scalars t,

$$F(t) = f(x^* + t(x - x^*)).$$

We have

$$F'(t) = (x - x^*)^T \nabla f(x^* + t(x - x^*))$$

and

$$F''(t) = (x - x^*)^T \nabla^2 f(x^* + t(x - x^*))(x - x^*) \geq m \|x - x^*\|^2 \geq 0.$$

Thus F' is an increasing function, and $F'(1) \geq F'(t)$ for all $t \in [0, 1]$. Hence

$$\begin{aligned} f(x) - f(x^*) &= F(1) - F(0) = \int_0^1 F'(t) dt \\ &\leq F'(1) = (x - x^*)^T \nabla f(x) \\ &\leq \|x - x^*\| \, \|\nabla f(x)\| \leq \frac{\|\nabla f(x)\|^2}{m}, \end{aligned}$$

where in the last step we used the result shown earlier. $\diamond$

---

4. Consider the following constrained optimization problem

$$\textbf{(A)} \quad \min_x f(x) \;\text{ s.t. }\; c_i(x) \geq 0, \;\; i = 1, 2, \ldots, m,$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $c_i : \mathbb{R}^n \to \mathbb{R}$ are smooth functions. Consider the following reformulation of (A) that makes use of "squared slack variables:"

$$\textbf{(B)} \quad \min_{x,s} f(x) \;\text{ s.t. }\; c_i(x) - s_i^2 = 0, \;\; i = 1, 2, \ldots, m.$$

Evidently any solution of

(a) Write down the Karush-Kuhn-Tucker (KKT) conditions for both (A) and (B).

(b) If x^* is a KKT point for (A), verify that we can obtain a KKT point for (B) by setting x^* to the same value and defining $s_i^* = \sqrt{c_i(x^*)}$.

(c) If (x^*, s^*) is a KKT point for (B), is it true that x^* must be a KKT point for (A)? Explain.

(d) Write down the linear independence constraint qualification (LICQ) conditions for both (A) and (B). (Use $\mathcal{A}$ to denote the set of active indices in (A), that is, $\mathcal{A} := \{i = 1, 2, \ldots, m \,:\, c_i(x^*) = 0\}$.)

(e) Given a KKT point x^* for (A) at which LICQ holds, is it true that the LICQ conditions for (B) are satisfied at the KKT point for (B) constucted as in part (b)?

(f) Write down the Mangasarian-Fromovitz constraint qualifications (MFCQ) for both (A) and (B).

(g) Given a KKT point x^* for (A) at which MFCQ holds, is it true that the MFCQ conditions for (B) are satisfied at the KKT point for (B) constucted as in part (b)?

Answer:

(a) For (A): There exists $\lambda^* \in \mathbb{R}^m$ such that

$$\textbf{(A)} \quad \nabla f(x^*) - \sum_{i=1}^{m} \lambda_i^* \nabla c_i(x^*) = 0, \quad 0 \le \lambda_i^* \perp c_i(x^*) \ge 0, \; i = 1, 2, \ldots, m.$$

For (B): There exists $\lambda^* \in \mathbb{R}^m$ such that

$$\textbf{(B)} \quad \nabla f(x^*) - \sum_{i=1}^{m} \lambda_i^* \nabla c_i(x^*) = 0, \quad 2 s_i^* \lambda_i^* = 0, \quad c_i(x^*) - (s_i^*)^2 = 0, \;\; i = 1, 2, \ldots, m.$$

(b) If x^* is a KKT point for (A), we have $\lambda^* \in \mathbb{R}^m$ that satisfies the conditions above for (A). Defining s^* as in the question, it is obvious that (x^*, s^*, λ^*) satisfy the conditions for (B) above.

(c) No. If (x^*, s^*, λ^*) satisfies the conditions for (B), it is clear that x^* will be feasible for (A) and that the required complementarity conditions involving λ_i^* and $c_i(x^*)$ are satisfied. Moreover, we will have $c_i(x^*) \ge 0$. However, there is no guarantee that the λ_i^* are nonnegative, as required by the KKT conditions for (A).

(d) For (A), we need that $\{\nabla c_i(x^*) \,:\, i \in \mathcal{A}\}$ is a linearly independent set. For (B), since all constraints are equalities (and hence all are active), we require the following set of vectors in $\mathbb{R}^{m+n}$ to be linearly independent:

$$\left\{ \begin{bmatrix} \nabla f_i(x^*) \\ -2 s_i^* e_i \end{bmatrix} \,:\, i = 1, 2, \ldots, m \right\},$$

where e_i is the ith unit vector in $\mathbb{R}^m$.

(e) Yes. Suppose we have w_i, $i = 1, 2, \ldots, m$ such that

$$\sum_{i=1}^{m} \begin{bmatrix} \nabla f_i(x^*) \\ -2 s_i^* e_i \end{bmatrix} w_i = 0.$$

For $i \notin \mathcal{A}$, we have $s_i^* > 0$, so by examining the $n+i$ component of this summation, we conclude that $w_i = 0$. The first n components of the summation therefore yield

$$\sum_{i \in \mathcal{A}} \nabla c_i(x^*) w_i = 0,$$

which implies that $w_i = 0$ for all $i \in \mathcal{A}$, by the LICQ conditions for (A).

(e) Using the active set notation $\mathcal{A}$, the MFCQ conditions for (A) are that there is a vector $v \in \mathbb{R}^n$ such that $\nabla c_i(x^*)^T v > 0$ for all $i \in \mathcal{A}$. Formulation (B) has only equality constraints, so the MFCQ condition here is simply that the Jacobian is linearly independent — the same condition as LICQ desribed in the answer to part (d).

(f) No. In order for MFCQ for (B) to hold, we need the set of active constraint gradients $\{\nabla c_i(x^*) : i \in \mathcal{A}\}$ to be linearly independent. This is not guaranteed by MFCQ for (A).

$\diamond$