Fall 2012 Qualifier Exam:
OPTIMIZATION

September 24, 2012

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 6 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. Consider the following convex quadratic program with a single equality constraint, nonnegativity constraints, and a diagonal Hessian:

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} x^T Q x + c^T x \text{ subject to } a^T x = 1, \;\; x \geq 0, \tag{1}$$

where $a \in \mathbb{R}^n$ is a vector with all positive entries, and Q is a diagonal matrix with all positive diagonal entries.

(a) Suppose we drop the bounds $x \geq 0$ from the formulation (1). Write down the KKT conditions for the resulting simplified problem, and use them to deduce the solution x in closed form.

(b) Returning to the full problem (1), write down the KKT conditions, denoting the Lagrange multiplier for the constraint $a^T x = 1$ by λ.

(c) Fixing the value of λ in these KKT conditions, find the value of x_i, $i = 1, 2, \ldots, n$ that satisfies these conditions as an explicit function of λ. (Use the notation $x_i(\lambda)$, $i = 1, 2, \ldots, n$ to denote these values.)

(d) Show that the function $t : \mathbb{R} \to \mathbb{R}$ defined by

$$t(\lambda) = a^T x(\lambda) - 1 = \sum_{i=1}^{n} a_i x_i(\lambda) - 1$$

is a monotonic piecewise linear function of λ, and identify the breakpoints of this function (the points where the slope changes discontinuously).

Answer:

(a) The KKT conditions for the simplified problem are

$$Qx + c - \lambda a = 0, \;\; a^T x = 1.$$

Since A is positve definite diagonal, we can use the first condition to express x explicitly in terms of λ: $x = Q^{-1}(\lambda a - c)$. Substituting into the constraint, we have $a^T Q^{-1}(\lambda a - c) = 1$, from which we deduce that $\lambda = (1 + a^T Q^{-1} c)/(a^T Q^{-1} a)$. We thus obtain

$$x = Q^{-1} \left[\frac{1 + a^T Q^{-1} c}{a^T Q^{-1} a} a - c \right].$$

(b)

$$0 \leq Qx + c - \lambda q \perp x \geq 0, \quad a^T x = 1.$$

(c)

$$x_i(\lambda) = \max\left(0, (a_i \lambda - c_i)/Q_{ii}, 0\right), \;\; i = 1, 2, \ldots, n.$$

(d) Note that each $x_i(\lambda)$ in (c) is a piecewise linear function of λ with two pieces separated by a breakpoint located at c_i/a_i. We have

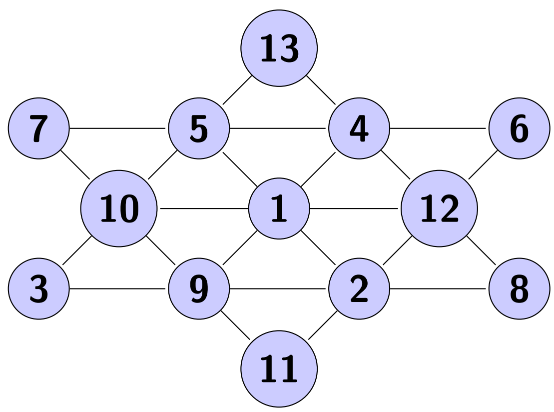
$$x_i(\lambda) = \begin{cases} 0 & \text{for } \lambda \le c_i/a_i \\ (a_i\lambda - c_i)/Q_{ii} & \text{for } \lambda \ge c_i/a_i. \end{cases}$$

Since $a_i > 0$ for all i, the function $t(\lambda)$ inherits the properties of piecewise linearity and monotonicity, and also inherits the breakpoints c_i/a_i, $i, 1, 2, \ldots, n$.

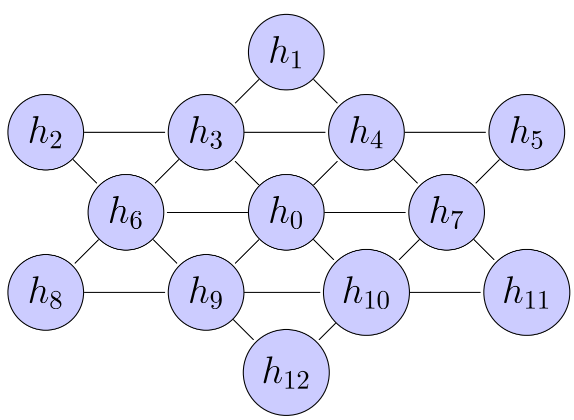
◇

---

2. The Christmas board game "22" involves a board with 13 holes and 13 pegs which fit in the holes. The pegs are numbered from 1 to 13. Holes are situated at the 12 intersection points on a six-pointed star and in the center of the star. To play the game, a peg is inserted in each hole. A winning configuration is one in which the sum of values for each of the six outer triangles sums to 22. Here, for example, is a winning assignment:



In your solution, use the following indexing scheme to reference the game board holes:



(a) Determine which variables are needed to provide a solution to the game?

(b) Define a mapping $H(t)$ that provides the subset of "holes" used in triangle t and use this to write down the "22" constraint. Note that t will range from 1 to 6, indicating each of the "outer" triangles.

(c) Write the full mathematical (or GAMS) model which finds a solution to the game.

(d) Suppose this model is solved for one solution. Determine an additional constraint that would eliminate just this solution, and enable the model to be rerun to find another solution.

(e) What techniques could you use to remove "equivalent solutions" from within your model search? Provide two constraints that remove such "symmetries" from your search.

(f) Write pseudo-code (GAMS or similar for example) that shows the sequence of model solves that will find all solutions to the game.

Answer: $title Christmas board game

option limrow=0, limcol=0, solprint=off; $offlisting

set h /h0*h12/; set tri /t1*t6/; set p /p1*p13/; set map(tri,h) / t1.(h1,h3,h4), t2.(h2,h3,h6), t3.(h4,h5,h7), t4.(h6,h8,h9), t5.(h7,h10,h11), t6.(h9,h10,h12) /; set inner(h) /h3,h4,h6,h7,h9,h10/; set cuts /c1*c1000/; set cut(cuts); set soln(cuts,h,p);

equations sumup(tri), onepeg(p), oneper(h), dummy, sym1(h), sym2, remove(cuts); variables obj; binary variables x(h,p);

sumup(tri).. sum((h,p)$map(tri,h), p.ord*x(h,p)) =e= 22;

onepeg(p).. sum(h, x(h,p)) =e= 1;

oneper(h).. sum(p, x(h,p)) =e= 1;

* fix inner permutation so h3 has minimum value sym1(h)$(inner(h) and not sameas(h,'h3')).. sum(p, p.ord*x('h3',p)) + 1 =l= sum(p, p.ord*x(h,p));

* inner permutation can go both ways: disallow sym2.. sum(p, p.ord*x('h4',p)) + 1 =l= sum(p, p.ord*x('h6',p));

* exclude solutions via cut to generate all solutions remove(cut).. sum((h,p)$(inner(h) and soln(cut,h,p)), x(h,p)) =l= 5;

dummy.. obj =e= 0;

model xmas22 /all/;

alias(c,cuts); scalar done /0/; cut(cuts) = no; soln(cuts,h,p) = no;

* fix to improve computing time marginally * x.fx('h3',p)$(p.ord gt 6) = 0;

loop(c$(not done), solve xmas22 using mip min obj; if (xmas22.modelstat eq 1, soln(c,h,p)$(x.l(h,p) gt 0.9) = 1; cut(c) = yes; else done = 1;););

* run using command options: ps=9999999 pw=240 parameter allsols(cuts,h); allsols(cut,h) = sum(p$soln(cut,h,p), p.ord); option allsols:0:1:1; display allsols;

* Total solutions is then dimension of cut * 12 * for each solution, move h3 to another inner hole * for each solution, flip the ordering of the permutation in inner holes ◇

---

3. In this problem, we will consider the feasible region of a *chance-constrained* problem:

$$X = \{x \in \mathbb{R}^n \mid \mathbb{P}[g_i(x, \xi) \geq 0 \ \forall i = 1, \ldots, m] \geq 1 - \epsilon\},$$

with each constraint function

$$g_i : \mathbb{R}^n \times \mathbb{R}^d \to \bar{\mathbb{R}},$$

and ξ being a random vector on a probability space $(\Omega, \Sigma, \mathbb{P})$.

(a) The set X is in general not convex. Give a simple example of constraints $g_i(x, \xi)$ and probability space $(\Omega, \Sigma, \mathbb{P})$, where X is *not* a convex set. Prove that your example set X is not convex.

(b) Now suppose the feasible region takes the form

$$X = \{x \in \mathbb{R}^n \mid \mathbb{P}(\xi^T x \geq b) \geq 1 - \epsilon\},$$

where $\xi \in \mathbb{R}^n$ is a **normally distributed** random vector with mean μ and covariance matrix Q. Show that X is convex if $\epsilon < 0.5$.

(c) In this concrete example, we will consider a production/distribution problem with a set J of customers whose (random) demand $d_j(\xi)$ must be met from a set of facilities I. Let x_{ij} be the amount of product shipped from $i \in I$ to $j \in J$. Suppose that the random demand for customer j comes from a discrete distribution; namely, that the demand of customer j in scenario $s \in S$ is d_{js} with probability p_s, for a finite set of scenarios S. We must choose the distribution amounts x_{ij} before the demands d_{js} are known. We would like to impose the constraint that the probability that *all* customers get their demand met is at least $1 - \epsilon$. Demonstrate how to model this using binary variables.

Answer:

(a) Here's one simple example. Let $\xi = (\xi_1, \xi_2)^T \in \mathbb{R}^2$ have two outcomes:

$$\Omega = \left\{ \begin{pmatrix} 0.25 \\ 0.75 \end{pmatrix}, \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix} \right\},$$

with each outcome having probability 0.5 of occuring. Consider the set

$$X = \{x \in \mathbb{R}^2 \mid x_1 \geq \xi_1, x_2 \geq \xi_2\}.$$

The points $y = (0.25, 0.75)^T$ and $z = (0.75, 0.25)$ are both in X. However $0.5y+0.5z = (0.5, 0.5)^T \notin X$, so X is not convex.

(b) We argue convexity by first doing a standard transformation to the standard normal:

$$\begin{aligned}\mathbb{P}(\xi^T x \geq b) &= \mathbb{P}\left(\frac{\xi^T x - \mu^T x}{\sqrt{x^T Q x}} \geq \frac{b - \mu^T x}{\sqrt{x^T Q x}}\right) \\ &= 1 - \Phi\left(\frac{b - \mu^T x}{\sqrt{x^T Q x}}\right).\end{aligned}$$

So

$$\mathbb{P}(\xi^T x \geq b) \geq 1 - \epsilon$$

when

$$1 - \Phi\left(\frac{b - \mu^T x}{\sqrt{x^T Q x}}\right) \geq 1 - \epsilon,$$

or when

$$\Phi\left(\frac{b - \mu^T x}{\sqrt{x^T Q x}}\right) \leq \epsilon.$$

The standard normal cdf is invertible, so we can say that

$$X = \{x \in \mathbb{R}^n \mid b - \mu^T x - \Phi^{-1}(\epsilon)\sqrt{x^T Q x} \leq 0\}.$$

It remains to argue that X is convex. Note that $\Phi^{-1}(\epsilon) < 0$ for $\epsilon < 0.5$, and since $Q \succeq 0$, the function $\sqrt{x^T Q x} = \sqrt{x^T L L^T x} = \sqrt{u^T u} = \|u\|_2$ for some matrix L. The $\|\cdot\|_2$ function is convex, so $b - \mu^T x - \Phi^{-1}(\epsilon)\sqrt{x^T Q x}$ is convex, and thus so is X.

(c) For each scenario $s \in S$, we need to introduce binary indicator variables z_s that will be 1 if

$$\sum_{i \in I} x_{ij} < d_{js} \text{ for some } j \in J.$$

We add the following two classes of constraints:

$$\begin{aligned}\sum_{i \in I} x_{ij} - d_{js} z_s \geq d_{js} \quad & \forall j \in J, \forall s \in S \\ \sum_{s \in S} p_s z_s \leq \epsilon.\end{aligned}$$

◇

---

4. Let X be the set of $(x, y) \in \{0,1\}^n \times \{0,1\}^{n(n-1)/2}$ that satisfy

$$y_{ij} \leq x_i, \quad \forall 1 \leq i < j \leq n \tag{2}$$

$$y_{ij} \leq x_j, \quad \forall 1 \leq i < j \leq n \tag{3}$$

$$x_i + x_j - y_{ij} \leq 1, \quad \forall 1 \leq i < j \leq n. \tag{4}$$

(a) Use Gomory-Chvátal rounding to show that the inequality

$$x_i + x_j + x_k \leq y_{ij} + y_{jk} + y_{ik} + 1 \tag{5}$$

is valid for conv(X) for any $1 \leq i < j < k \leq n$.

(b) Show that the inequality

$$y_{ij} + y_{ik} \leq x_i + y_{jk} \tag{6}$$

is valid for conv(X) for any $1 \leq i < j < k \leq n$. (You do not have to use Gomory-Chvátal rounding for this question, but you may if you wish.)

(c) Now, suppose that $n = 3$. Prove the specific case of inequality (6),

$$y_{12} + y_{13} \leq x_1 + y_{23}, \tag{7}$$

is facet-defining for conv(X). You may take as given the fact that $\dim(\text{conv}(X)) = 6$, i.e., conv(X) is full-dimensional.

Answer:

(a) Add inequality (4) for $\{i, j\}$, $\{i, k\}$ and $\{j, k\}$ yields

$$2x_i + 2x_j + 2x_k - y_{ij} - y_{jk} - y_{ik} \leq 3.$$

Divide by two yields,

$$x_i + x_j + x_k - (1/2)y_{ij} - (1/2)y_{jk} - (1/2)y_{ik} \leq 3/2.$$

Round down the coefficients on the left-hand side yields,

$$x_i + x_j + x_k - y_{ij} - y_{jk} - y_{ik} \leq 3/2.$$

Now, round down the right-hand side yields,

$$x_i + x_j + x_k - y_{ij} - y_{jk} - y_{ik} \leq 1.$$

Rearranging yields (5).

(b) I give two possible correct answers. First, is a direct case-by-case argument. Suppose $x_i = 0$. Then, by (2), it holds that $y_{ij} = 0$ and $y_{ik} = 0$, and hence

$$y_{ij} + y_{ik} = 0 \leq x_i + y_{jk}$$

and so (6) holds. Now suppose $x_i = 1$. If $x_j = 0$, then $y_{ij} = 0$ by (3), and hence

$$y_{ij} + y_{ik} = y_{ik} \leq 1 = x_i \leq x_i + y_{jk}$$

and so (6) holds. A symmetric argument works if $x_k = 0$. If $x_j = 1$ and $x_k = 1$, then by (4), it holds that $y_{jk} = 1$, and hence

$$y_{ij} + y_{ik} \leq 2 = x_i + y_{jk}$$

and so (6) holds.

The second argument is based on Gomory-Chvátal rounding. First, add inequalities (2) and (3) for (i, j) and for (i, k), which yields

$$2y_{ij} + 2y_{ik} - 2x_i - x_j - x_k \leq 0.$$

Now add inequality (4) corresponding to (j, k) to obtain

$$2y_{ij} + 2y_{ik} - 2x_i - y_{jk} \leq 1.$$

Now divide by two:

$$y_{ij} + y_{ik} - x_i - (1/2)y_{jk} \leq 1/2.$$

Round down the coefficients:

$$y_{ij} + y_{ik} - x_i - y_{jk} \leq 1/2.$$

Finally, round down the right-hand side:

$$y_{ij} + y_{ik} - x_i - y_{jk} \leq 0.$$

Rearranging yields the inequality (6).

(c) We must provide six affinely independnet points in conv(X) that satisfy the inequality as an equation. Consider the six points below:

x_1	x_2	x_3	y_{12}	y_{13}	y_{23}	$y_{12} + y_{13}$	$x_1 + y_{23}$
0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	1	0	0	0	0	0
1	1	0	1	0	0	1	1
1	0	1	0	1	0	1	1
1	1	1	1	1	1	2	2

As verified in this table, each of these points satisfies (7) as an equation. In addition, noting that (2) – (4) model the conditions that $y_{ij} = x_i x_j$, it is easy to check that these points are in X and hence are in conv(X). Finally, subtracting the first point from the others doesn't change them, and hence it is sufficient to show the final 5 points are linearly independent. Note from their arrangement in the table that these points form a lower-triangular matrix with ones on the diagonal, and hence these points have full row-rank, and thus are linearly independent. (Depending on the order the points are given in, row and column swaps may be necessary to make it obvious the points are linearly independent.)

◇

5. (a) Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is convex and concave. Show that f must be an affine function.
   (b) Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is convex and bounded above. Show that f must be a constant function.
   (c) Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is strongly convex and Lipschitz, meaning that there is a constant L such that $|f(x) - f(y)| \leq L||x - y||$ for all x and y. Show no such f exists.

   Answer:

   (a) If f is both convex and concave, $f(x) = f(0) + v^T x$ for all x and some $v \in \partial f(0)$.
   (b) If f is not constant, then there is a line that goes off to infinity that lower bounds the function contradicting boundedness.
   (c) Assuming Lipschitz gives: $|f(x) - f(y)| < L||x - y||$ for some constant L. Assuming strongly convex and Lipschitz gradients gives:

   $$|f(x) - f(y)| > -|\nabla f(x)^T (x - y)| + \ell/2||x - y||^2 > (\ell/2||x - y|| - \|\nabla f(x)\|)||x - y|| \,.$$

   Putting these together gives

   $$\ell/2||x - y|| < \|\nabla f(x)\| + L$$

   for all x and y in $\mathbb{R}^d$. Plugging in $x = 0$ and y sufficiently large gives a contradiction.

◇

---

6. Consider the following optimization problem, which is parametrized by the scalar α:

$$P(\alpha) : \qquad \min_{x \in \mathbb{R}^n} f(x) \text{ subject to } p^T x \leq \alpha, \tag{8}$$

where $f : \mathbb{R}^n \to R$ is a smooth, strongly convex function and p is a nonzero vector in $\mathbb{R}^n$. We denote the optimal objective value for this problem by $\phi(\alpha)$, and note that the problem (8) has a unique minimizer $x(\alpha)$ for each $\alpha \in \mathbb{R}$.

   (a) Show that ϕ is a convex, decreasing function of α.
   (b) Show that ϕ is a continuous function of α.
   (c) Show that there is a threshold value $\bar{\alpha}$ such that $\phi(\alpha) = \phi(\bar{\alpha})$ for all $\alpha \geq \bar{\alpha}$ while $\phi(\alpha) > \phi(\bar{\alpha})$ for all $\alpha < \bar{\alpha}$. (Hint: Consider the unconstrained global minimizer x^* of $f(x)$.)
   (d) Consider the following related problem, in which $\lambda \geq 0$ is a parameter:

   $$\min_{z \in \mathbb{R}^n} f(z) + \lambda p^T z, \tag{9}$$

   where f and p are the same as in (8). Show that the point $z(\lambda)$ that solves (9) is identical to the solution $x(\alpha)$ of (8) if we set $\alpha = p^T z(\lambda)$.

Answer:

(a) Consider two values $\alpha_1 < \alpha_2$ and set $\alpha := \rho\alpha_1 + (1-\rho)\alpha_2$, for some $\rho \in [0,1]$. The point $x_\rho := \rho x(\alpha_1) + (1-\rho)x(\alpha_2)$ is feasible for $P(\alpha)$, and we have

$$\phi(\alpha) \leq f(x_\rho) \leq \rho f(x(\alpha_1)) + (1-\rho)f(x(\alpha_2)) = \rho\phi(\alpha_1) + (1-\rho)\phi(\alpha_2),$$

proving convexity. The fact that ϕ is a decreasing function of α follows immediately from the fact that the feasible set for $P(\alpha)$ increases as α increases.

(b) Continuity follows from monotonicity, convexity, and the fact that ϕ has domain $\mathbb{R}$. We can prove from first principles as follows. Given any α and small $\epsilon > 0$ note that $x(\alpha) + \epsilon p/(p^T p)$ is feasible for $P(\alpha + \epsilon)$, so

$$\phi(\alpha + \epsilon) \leq f(x(\alpha) + \epsilon p/(p^T p)) = f(x(\alpha)) + O(\epsilon) = \phi(\alpha) + O(\epsilon).$$

Similarly, $x(\alpha + \epsilon) - \epsilon p/(p^T p)$ is feasible for $P(\alpha)$, so

$$\phi(\alpha) \leq f(x(\alpha + \epsilon) - \epsilon p/(p^T p)) = \phi(\alpha + \epsilon) + O(\epsilon).$$

By combining these bounds, we have $|\phi(\alpha) - \phi(\alpha + \epsilon)| = O(\epsilon)$ for all $\epsilon \geq 0$, implying continuity of ϕ at α.

(c) Set $\bar{\alpha} = f(x^*)$, where x^* is the unconstrained minimized of f. Note that x^* is feasible for $P(\alpha)$ for all $\alpha \geq \bar{\alpha}$, so $\phi(\alpha) = \phi(\bar{\alpha})$. For $\alpha < \bar{\alpha}$ we must have $\phi(\alpha) > \phi(\bar{\alpha})$, since otherwise we would have $x(\alpha) \neq x^*$ with $f(x(\alpha)) \leq f(x^*)$, which is not possible since x^* is the global minimizer of f, unique because f is convex.

(d) The unique solution $z(\lambda)$ of (9) satisfies $\nabla f(z(\lambda)) + \lambda p = 0$. Comparing with the KKT conditions for (8), namely,

$$\nabla f(x) - \gamma p = 0, \;\; 0 \leq \gamma \perp \alpha - p^T x \geq 0,$$

we see that $x = z(\lambda)$ and $\gamma = \lambda$ satisfies these conditions, when α is defined as in the question. Thus $z(\lambda)$ is the unique solution of $P(\alpha)$ when $\alpha = p^T z(\lambda)$.

◇