Prof. P.A. Parrilo Fall 2008

Optimization Methods
6.255/15.093
Final exam

Date Given: December 17th, 2008
Date Due: You have **three hours** (academic time, 170 minutes) to complete the exam.

P1. [***30 pts***] Classify the following statements as true or false. All answers must be well-justified, either through a short explanation, or a counterexample. Unless stated otherwise, all LP problems are in standard form.

(a) Assume Q is positive definite. Then, if two vectors are Q-conjugate then they are orthogonal.

(b) For a uncapacitated min-cost flow problem specified by integer data, there always exist integer optimal flows.

(c) For a quadratic function, the steepest descent method (with exact line search) converges quadratically.

(d) For the problem

$$\text{minimize } x \qquad \text{subject to } y \leq x^3, \quad y \geq 0,$$

the gradients of the constraints satisfy the linear independence constraint qualification (LICQ).

(e) For an integer program, the value obtained by Lagrangean relaxation is never worse than that obtained by the LP relaxation.

(f) If $f(x)$ and $g(x)$ are convex univariate functions, then so is $2f(x) - 3g(x)$.

(g) For unconstrained minimization of a differentiable function, the condition $\nabla f(\mathbf{x}) = 0$ is necessary for global optimality.

(h) For a constrained optimization problem, if a point $\bar{\mathbf{x}}$ is feasible and satisfies the KKT conditions, then it is a local minimum.

(i) If all the reduced costs are nonnegative, then the current basis is dual feasible.

(j) If the set $\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq 0\}$ is convex, then the function $f(\mathbf{x})$ is convex.

Solution: **(a)** FALSE. A simple counterexample is

$$Q = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

The matrix Q is positive definite, we have $v_1' Q v_2 = 0$ (they are Q-conjugate), but v_1 and v_2 are not orthogonal.

(b) TRUE. If the problem is bounded below, this is a consequence of the fact that BFSs correspond to tree solutions.

(c) FALSE. The steepest descent method has *linear* convergence, not quadratic.

(d) FALSE. The only local minimum is $(0, 0)$. At this point, the gradient of the constraints are linearly dependent.

(e) TRUE. This is essentially Theorem 11.4 in the book, and was covered in Lecture 14. It follows directly from the fact that the feasible set of the Lagrangean relaxation is a subset of that of the LP relaxation.

(f) FALSE. Just take for instance $f(x) = 0$, $g(x) = x^2$. Then $f(x) - g(x) = -3x^2$, which is not convex.

(g) TRUE. The vanishing gradient condition ensures that small perturbations cannot decrease the optimal value (up to linear terms).

(h) FALSE. Since the KKT are essentially first-order conditions, they can only "see" the linear part of the objective and constraints, and are thus unable to ensure even local optimality (unless further conditions, such as convexity, are imposed). As a simple example, consider the objective function $f(x) := -x^2$, and the feasible set defined by $g(x) := x \leq 0$. Clearly, the origin $x = 0$ is a KKT point (since $\nabla f(x) + u_1 \cdot \nabla g(x) = 0 + 0 \cdot 1 = 0$), but it is a local maximum (not a minimum).

(i) TRUE. Dual feasibility is equivalent to the nonnegativity of the reduced costs.

(j) FALSE. Consider for instance the univariate function $f(x) = x^4 - x^2 - \epsilon$, for some small $\epsilon > 0$ (plot it!). The sublevel set $\{x \in \mathbb{R} : f(x) \leq 0\}$ is a closed interval (and thus convex) but the function is not convex (the Hessian at the origin is negative definite).

P2. [***25 pts***] This problem is about finding the largest Euclidean ball contained inside a given polyhedron. Consider a (nonempty) polyhedron P in $\mathbb{R}^n$ described by m linear inequalities, i.e., of the form

$$P = \{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{a}_i'\mathbf{x} \leq b_i, \quad i = 1, \ldots, m\},$$

where $\mathbf{a}_i \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$. For simplicity, assume the inequalities have been scaled so that $||\mathbf{a}_i|| = 1$. We define the Euclidean ball $B(\mathbf{x}_0, R)$ with center at $\mathbf{x}_0$ and radius R as the set $B(\mathbf{x}_0, R) := \{\mathbf{x} \in \mathbb{R}^n \,:\, ||\mathbf{x} - \mathbf{x}_0|| \leq R\}$.

(a) Prove that the ball $B(\mathbf{x}_0, R)$ is contained in the half-space defined by $\mathbf{a}_i'\mathbf{x} \leq b_i$ if and only if $\mathbf{a}_i'\mathbf{x}_0 + R \leq b_i$. (Hint: Draw a picture. What is the normal vector of the corresponding hyperplane?)

(b) Write an LP formulation for finding the largest ball contained in the polyhedron P.

(c) Write the dual of your LP formulation.

(d) Write the complementarity slackness conditions for this LP problem.

(e) Show that the optimality conditions are equivalent to the following geometric interpretation:

> A ball $B(\mathbf{x}_0, R)$ is optimal if and only if the center $\mathbf{x}_0$ lies inside the convex hull of those contact points where the ball touches the hyperplanes $\mathbf{a}_i'\mathbf{x} = b_i$.

Here you can assume that the problem is nontrivial (i.e., the radius of the largest ball is finite and nonzero).

(f) Assume now that the $\mathbf{a}_i$ are nonnegative vectors. Show using the dual LP that in this case, the polyhedron P must contain balls of arbitrarily large radius. Is this true for all unbounded polyhedra (regardless of the assumption on the $\mathbf{a}_i$)?

Solution: **(a)** The ball $B(\mathbf{x}_0, R)$ in contained inside the half-space $\mathbf{a}_i'\mathbf{x} \leq b_i$ if and only if the point $\mathbf{x}_0 + R\mathbf{a}_i$ is in the half-space. Using the normalization condition, this yields:

$$\mathbf{a}_i'(\mathbf{x}_0 + R\mathbf{a}_i) \leq b_i \qquad \Leftrightarrow \qquad \mathbf{a}_i'\mathbf{x}_0 + R \leq b_i.$$

(b) The ball is contained inside the polyhedron if and only if it is contained in all the halfspaces. From this, we can write the primal LP formulation:

$$\text{maximize } R \qquad \text{subject to} \qquad \begin{cases} \mathbf{a}_i'\mathbf{x}_0 + R & \leq & b_i, \quad i = 1, \ldots, m \\ R & \geq & 0 \end{cases}$$

where the decision variables are $\mathbf{x}_0$ and R. This problem is always feasible, since P is nonempty.

(c) The dual LP problem is

$$\text{minimize } \sum_{i=1}^{m} \lambda_i b_i \qquad \text{subject to} \qquad \begin{cases} \sum_{i=1}^m \lambda_i \mathbf{a}_i & = & \mathbf{0} \\ \sum_{i=1}^m \lambda_i & \geq & 1 \\ \lambda_i & \geq & 0 \end{cases}.$$

(d) The optimality conditions for a linear programming problem are primal feasibility, dual feasibility, and complementary slackness (or equivalently, zero duality gap). The complementary slackness conditions take the form:

$$R \cdot \left(\textstyle\sum_{i=1}^m \lambda_i - 1\right) = 0, \qquad \lambda_i \cdot (\mathbf{a}_i'\mathbf{x}_0 + R - b_i) = 0.$$

(the additional condition $(\mathbf{x}_0)_j \cdot (\sum_{i=1}^m \lambda_i \mathbf{a}_i)_j = 0$ is automatically satisfied for all dual feasible λ).

(e) If the optimal R is nonzero, then we must have $\sum_{i=1}^m \lambda_i = 1$. The ith primal constraint is active if the ball $B(\mathbf{x}_0, R)$ touches the hyperplane $\mathbf{a}_i'\mathbf{x} = b_i$ (if the ith primal constraint is inactive, the corresponding λ_i must be equal to zero). Thus, we have:

$$\sum_{i \in \mathcal{A}} \lambda_i \mathbf{a}_i = \mathbf{0}, \qquad \sum_{i \in \mathcal{A}} \lambda_i = 1, \qquad \lambda_i \geq 0, \quad i \in \mathcal{A},$$

where $\mathcal{A}$ is the set of active primal constraints. These can be equivalently rewritten as:

$$\sum_{i \in \mathcal{A}} \lambda_i (\mathbf{x}_0 + R\mathbf{a}_i) = \mathbf{x}_0, \qquad \sum_{i \in \mathcal{A}} \lambda_i = 1, \qquad \lambda_i \geq 0, \quad i \in \mathcal{A}.$$

The geometric interpretation of this condition is clear: the center $\mathbf{x}_0$ of the ball must be inside the convex hull of the points $\mathbf{x}_0 + R\mathbf{a}_i$, which are exactly those where the ball $B(\mathbf{x}_0, R)$ touches the hyperplanes.

(f) If all the $\mathbf{a}_i$ are nonnegative, then the dual LP is obviously infeasible. Since the primal LP is always feasible, then it must be unbounded (and thus, there are balls of arbitrarily large radius R inside the polyhedron P).

The conclusion is not true for all unbounded polyhedra. For instance, the polyhedron $\{(x, y) \in \mathbb{R}^2 : -1 \leq x \leq 1\}$ is unbounded, but the largest ball inside it has radius 2.

P3. [***20 pts***] Consider a discrete-time linear dynamical system of the form

$$x_{k+1} = Ax_k + bu_k, \qquad k = 0, \dots, N-1,$$

where $x_k \in \mathbb{R}^n$, the control input $u_k \in \mathbb{R}$, and $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^{n \times 1}$ are given.

The *total fuel consumption* F is given by the expression

$$F = \sum_{k=0}^{N-1} f(u_k),$$

where $f : \mathbb{R} \to \mathbb{R}$ describes the fuel consumption at stage k as a function of the control input u_k. For this problem, we assume that $f(u)$ is a piecewise linear function, of the form:

$$f(u) := \begin{cases} 2u & u \geq 0 \\ -u & u \leq 0. \end{cases}$$

Our goal in this problem is to choose values for the control inputs $u_0, \dots, u_{N-1}$, in such a way that the state x_N at time N is close to a given desired final state x_F, and the total fuel consumption is small. More concretely, we want to find control inputs that minimize $F + \alpha \cdot ||x_N - x_F||_\infty$, for some given fixed parameter α, where $\alpha \geq 0$ (this parameter gives a tradeoff between total fuel consumption and accuracy of the final state). Assume that the initial state is equal to x_0, and let $J(x_0)$ denote the optimal cost with this initial state.

(a) Formulate the minimum fuel optimal control problem as a linear program.

(b) Propose a dynamic programming formulation for this problem. Describe clearly all the components of the solution, and how to solve it.

(c) Prove, using either formulation, that the optimal cost $J(x_0)$ is a convex function of the initial state x_0.

(d) Fix the initial state x_0. What can you say about the optimal cost $J(x_0)$, as a function of the parameter α?

Solution: **(a)** The problem admits the following LP formulation:

$$\min_{x_k,u_k,\gamma_k,\epsilon} \alpha\cdot\epsilon + \sum_{k=0}^{N-1}\gamma_k \qquad \text{s.t.} \qquad \begin{cases} Ax_k + bu_k = x_{k+1} & k = 0,\dots,N-1 \\ 2u_k \le \gamma_k, \quad -u_k \le \gamma_k & k = 0,\dots,N-1 \\ -\epsilon \le (x_N)_i - (x_F)_i \le \epsilon & i = 1,\dots,n \end{cases}$$

where the decision variables are $\{x_1,\dots,x_N,u_0,\dots,u_{N-1},\gamma_0,\dots,\gamma_{N-1},\epsilon\}$.

(b) A DP formulation can be easily obtained, since the problem is directly in a form suitable for a Bellman-type iteration. Define the value function $J_k(x_k)$ to be the optimal cost starting at stage k from the state x_k. Then, we have the recursion and final condition:

$$J_k(x_k) = \min_{u_k}[f(u_k) + J_{k+1}(Ax_k + bu_k)], \qquad J_N(x_N) = \alpha||x_N - x_F||_\infty.$$

We can solve this recursion backwards in time, thus obtaining $J(x_0)$ as an explicit function of x_0.

(c) We will prove this using either formulation:

1. For the LP formulation, notice that x_0 appears in the right-hand side of the constraints, with the problem being a minimization. Thus, by dualizing, it follows that the optimal solution is a convex function of x_0 (since it is the maximum of a finite set of affine functions of x_0).
2. From the DP formulation, notice that $J_N(x_N)$ is convex (since it is a norm) and piecewise linear. Furthermore, if J_{k+1} is convex then so is J_k, since it is a partial minimization of a convex function over a convex set. Thus, $J(x_0)$ is a convex function of x_0.

(d) By a similar argument as the one used above (for the LP case), the cost $J(x_0)$ is a piecewise-linear concave function of α. It can also be seen that it is an increasing function of α, since the objective function is an increasing function of α over the feasible set (notice that $\epsilon \ge 0$).

P4. [***25 pts***] This problem discusses the *0-1 multiple knapsack* (MKP) problem. Consider a set of n items, and m knapsacks (with $m \leq n$). Let p_j, w_j denote the profit and weight of item j, and c_i be the capacity of the ith knapsack. Given the n items, the objective is to find m *disjoint* subsets, so that the total profit of all the selected items is as large as possible, and with each subset being assigned to a knapsack subject to the respective weight constraint. (In terms of the formulation described in class, think of a band of m thieves that try to maximize their collective profits). For simplicity, assume all p_j, w_j, c_i are positive integers.

(a) Give an integer programming formulation of the multiple knapsack problem. For this, define binary variables x_{ij}, such that x_{ij} is equal to one if item j is assigned to knapsack i, or zero otherwise. Clearly express the objective function, and the contraints.

(b) Using your formulation, show that an upper bound on the optimal cost of the MKP problem can be obtained by considering a single 0-1 knapsack problem, with a knapsack of capacity equal to $\sum_{i=1}^{m} c_i$. Explain in detail the relationship between the two formulations, and why the inequality holds.

(c) Propose a Lagrangean relaxation for the IP formulation in item (a). Show that for a suitable choice of the dualized constraints, the Lagrangean dual can be reduced to m independent 0-1 single knapsack problems. What common feature do these problems have?

(d) Describe very clearly (but at a high level) how would you use all the information above to fully solve a multiple knapsack problem (in particular: how would you solve the subproblems? how would you find the optimal Lagrange multipliers?).

Solution: **(a)** The IP formulation of MKP is:

$$\text{maximize} \sum_{i=1}^{m}\sum_{j=1}^{n} p_j x_{ij} \qquad \text{subject to} \qquad \begin{cases} \sum_{j=1}^{n} w_j x_{ij} & \leq & c_i & i = 1,\dots,m \\ \sum_{i=1}^{m} x_{ij} & \leq & 1 & j = 1,\dots,n \\ x_{ij} & \in & \{0,1\} \end{cases}$$

(b) Clearly, any feasible solution of the MKP problem yields a feasible solution for the single knapsack problem of capacity $\sum_{i=1}^{m} c_i$ (just put all the items together in a big knapsack!). Indeed, the latter corresponds to the formulation:

$$\text{maximize} \sum_{j=1}^{n} p_j z_j \qquad \text{subject to} \qquad \begin{cases} \sum_{j=1}^{n} w_j z_j & \leq & \sum_{i=1}^{m} c_i \\ z_j & \in & \{0,1\} \end{cases}$$

Any feasible solution of the first problem yields a feasible solution of the second (with the same cost), via $z_j := \sum_{i=1}^{m} x_{ij}$. Thus, the value of the MKP problem is always less than or equal to than the value of the single-knapsack problem.

(c) We propose a Lagrangean formulation by dualizing the n constraints $\sum_{i=1}^{m} x_{ij} \leq 1$, with some fixed nonnegative multipliers λ_i. We have then (notice that we are maximizing):

$$\max \sum_{i=1}^{m}\sum_{j=1}^{n} p_j x_{ij} - \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{m} x_{ij} - 1 \right) \qquad \text{s.t.} \qquad \begin{cases} \sum_{j=1}^{n} w_j x_{ij} & \leq & c_i & i = 1,\dots,m \\ x_{ij} & \in & \{0,1\} \end{cases}$$

The objective can be rewritten as:

$$\sum_{i=1}^{m}\sum_{j=1}^{n} (p_j - \lambda_j) x_{ij} + \sum_{j=1}^{n} \lambda_j,$$

and provides an upper bound on the value of the MKP. Its value can be obtained by solving m independent single 0-1 knapsack problems with profits, weights, and capacity given by $(\mathbf{p} - \lambda, \mathbf{w}, c_i)$, i.e. :

$$\max \sum_{j=1}^{n} (p_j - \lambda_j) x_{ij} \qquad \text{s.t.} \qquad \begin{cases} \sum_{j=1}^{n} w_j x_{ij} & \leq & c_i \\ x_{ij} & \in & \{0,1\} \end{cases}$$

Interestingly, these m problems all have the same weights and profits, but differ in the capacity c_i.

(d) We can fully solve a multiple knapsack problem using a *branch and bound* scheme (for instance, branching on whether a particular item is on a knapsack or not), where we use the Lagrangian relaxation approach discussed above to compute an upper bound for every subproblem. The single 0-1 knapsacks could be solved by dynamic programming (as explained during the lecture), which would enable solving all of them "simultaneously," since the DP formulation obtains solutions for all values of the right-hand side. For the minimization over λ_i (i.e., the computation of the optimal dual multipliers), we could use a subgradient method as discussed in class. The use of branch and bound allows us to close any possible duality gap, in case the Lagrangian relaxation is not exact.