EE127 - EE227AT 5/14/2015
L. El Ghaoui

Final: Solutions

1. (*10 points*) A project consisting of n different tasks can be represented as a directed graph with n arcs and m nodes. The arcs represent the tasks. The nodes represent precedence relations: If arc k starts at node i and arc j ends at node i, then task k cannot start before task j is completed. Node 1 only has outgoing arcs. These arcs represent tasks that can start immediately and in parallel. Node m only has incoming arcs. When the tasks represented by these arcs are completed, the entire project is completed.

   We can fully describe the network with the so-called arc-node incidence matrix, which is the $m \times n$ matrix defined as

$$A_{ij} = \begin{cases} 1 & \text{if arc } j \text{ starts at node } i, \\ -1 & \text{if arc } j \text{ ends at node } i, \\ 0 & \text{otherwise.} \end{cases}, \quad 1 \le i \le m, \;\; 1 \le j \le n.$$

   We are interested in computing an optimal schedule, that is, in assigning an optimal start time and a duration to each task. The variables in the problem are are $v \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, which are defined as follows.

   - y_k is the duration of task k, for $k = 1, \ldots, n$. The variables y_k must satisfy the constraints $\alpha_k \le y_k \le \beta_k$. We also assume that the cost of completing task k in time y_k is given by $c_k(\beta_k - y_k)$. This means there is no cost if we we use the maximum allowable time β_k to complete the task, but we have to pay if we want the task finished more quickly.
   - v_j is an upper bound on the completion times of all tasks associated with arcs that end at node j. Thus, these variables must satisfy the relations

$$v_j \ge v_i + y_k \text{ if arc } k \text{ starts at node } i \text{ and ends at node } j.$$

   Our goal is to minimize the sum of the completion times of the entire project, plus the total cost. Formulate the problem as an LP.

   Solution: The total completion time is $v_m - v_1 = -e^T v$, where $e = (1, 0, \ldots, 0, -1)$. The inequalities

$$v_j \ge v_i + y_k \text{ if arc } k \text{ starts at node } i \text{ and ends at node } j.$$

   can be written as $A^T v + y \le 0$.

   The problem then writes

$$\min_{y,v} \; e^T v + c^T(\beta - y) \;:\; A^T v + y \le 0, \;\; \alpha \le y \le \beta.$$

2. (*10 points*) A retailer wishes to optimize the prices of its products based on estimated demand (estimated amount of sales). The demand D_i for product $i \in \{1, \ldots, n\}$ is modeled as

$$D_i(p_i) = b_i - g_i(p_i - p_i^r)$$

where p_i is the price of the product, p_i^r is a reference price (say, the manufacturer's suggested price), b_i is the corresponding demand, and $g_i > 0$ is a "price sensitivity". (The model assumes that the demand decreases as price increases, which is usually the case.) For a vector of prices $p \in \mathbb{R}^n$, the revenue is given by $R(p) := p^T D(p)$, and the profit if $P(p) := (p - p^0)^T D(p)$, with p^0 the vector of purchase prices. The pricing problem is to maximize revenue, subject to non-negativity of the price vector; a lower bound P_{low} on the profit; and inventory constraints, which translate as upper and lower bounds D_{up}, D_{low} on the demand.

(a) Show how to formulate the problem as an optimization problem. Make sure to define precisely the constraints, the variables, and the objective function.

(b) Is the problem you have obtained convex? Discuss.

Solution:

(a) The problem writes

$$\max_p \sum_{i=1}^n p_i(b_i - g_i(p_i - p_i^r)) \quad : \quad \begin{array}{l} \sum_{i=1}^n (p_i - p_i^0)(b_i - g_i(p_i - p_i^r)) \geq P_{\text{low}}, \\ D_{\text{low}} \leq b_i - g_i(p_i - p_i^r) \leq D^{\text{up}}, \quad p \geq 0. \end{array}$$

(b) The problem is convex, in fact a QCQP, since $g > 0$.

3. (*10 points*) We consider a portfolio optimization problem, of the form

$$p^* = \max_{w \in \mathcal{W}} \hat{r}^T w - \frac{1}{2} w^T D w,$$

where $\hat{r} \in \mathbb{R}^n$ is the vector of expected returns of n different assets (*e.g.*, stocks), and $D = \mathbf{diag}(\sigma_1^2, \ldots, \sigma_n^2)$ the (diagonal) covariance matrix, with $\sigma_i > 0$ the corresponding standard deviation of asset i. Here, $w \in \mathbb{R}^n$ is a vector that contains the proportions of a given budget to be allocated to each asset, and $\mathcal{W} = \{w \geq 0 \, : \, w^T \mathbf{1} = 1\}$, with $\mathbf{1}$ the vector of ones.

(a) Show that, for any scalars $\rho \in \mathbb{R}$ and $\sigma > 0$, we have

$$\psi := \max_{\omega \geq 0} \rho\omega - \frac{\sigma^2}{2}\omega^2 = \frac{1}{2\sigma^2}\rho_+^2,$$

where $\rho_+ = \max(0, \rho)$, and with *unique* optimal point $\omega^* = \rho_+/\sigma^2$. Carefully argue your proof. *Hint:* distinguish the case $\rho \leq 0$ from $\rho > 0$, and for each case, show that the RHS is an upper bound, and that it is attained.

(b) Using duality, with the Lagrangian

$$\mathcal{L}(w, \nu) = \hat{r}^T w - \frac{1}{2} w^T D w + \nu \left(1 - w^T \mathbf{1}\right)$$

show that the optimal value p^* can be expressed as the optimal value of a one-dimensional problem:

$$p^* = \min_{\nu} \nu + \frac{1}{2} \sum_{i=1}^n \frac{(r_i - \nu)_+^2}{\sigma_i^2}.$$

Make sure to justify any use of strong duality. *Hint:* use part 3a.

(c) Explain how to recover a primal optimal point w^* based on a dual optimal point $\nu*$.

(d) *This is a bonus question, worth an extra* 5 *points.* Assume that the covariance matrix is not diagonal anymore, but of the form $C = D + ff^T$, with $f \in \mathbb{R}^n$. Show that the problem can be reduced to a two-dimensional problem, which you will detail.

Solution:

(a) If $\rho \leq 0$, then $\psi \leq 0$. The zero upper bound is attained with the unique point $\omega = 0$. Hence, $\psi = 0$ in that case. If $\rho > 0$, then since

$$\psi \leq \max_{\omega} \rho\omega - \frac{\sigma^2}{2}\omega^2 = \frac{1}{2\sigma^2}\rho^2,$$

the upper bound is attained with the feasible (unique) point $\omega^* = \rho/\sigma^2 (\geq 0)$. Hence, $\psi = \rho^2/(2\sigma^2)$ in that case. This proves the result.

(b) We have

$$p^* = \max_{w \geq 0} \min_{\nu} \hat{r}^T w - \frac{1}{2} w^T D w + \nu \left(1 - w^T \mathbf{1}\right)$$

Strong duality holds, since the original problem is convex and strictly feasible. We obtain $p^* = d^*$, with

$$\begin{aligned} d^* &= \min_{\nu} \max_{w \geq 0} \hat{r}^T w - \frac{1}{2} w^T D w + \nu \left(1 - w^T \mathbf{1}\right) \\ &= \min_{\nu} \nu + \max_{w \geq 0} (\hat{r} - \nu \mathbf{1})^T w - \frac{1}{2} w^T D w \\ &= \min_{\nu} \nu + \frac{1}{2} \sum_{i=1}^{n} \frac{\max(0, \hat{r}_i - \nu)^2}{\sigma_i^2}, \end{aligned}$$

where we have used part 3a.

(c) Since for each ν, the solution to

$$\max_{w \geq 0} (\hat{r} - \nu \mathbf{1})^T w - \frac{1}{2} w^T D w$$

is unique, and given by

$$w^*(\nu) = (\hat{r}_i - \nu)_+ / \sigma_i^2, \quad i = 1, \ldots, n,$$

we conclude that, if ν^* is optimal for the dual problem, then $w^*(\nu^*)$ is optimal for the primal problem.

(d) We start with

$$p^* = \max_{w \geq 0} \hat{r}^T w - \frac{1}{2}(w^T D w + z^2) \; : \; z = f^T w.$$

Again this problem is convex and strictly feasible, therefore strong duality holds. Using the Lagrangian

$$\mathcal{L}(w, \nu, \mu) = \hat{r}^T w - \frac{1}{2}(w^T D w + z^2) + \nu \left(1 - w^T \mathbf{1}\right) + \mu(z - f^T w)$$

easily leads to the dual formulation

$$\min_{\nu} \nu + \frac{1}{2}\mu^2 + \frac{1}{2} \sum_{i=1}^{n} \frac{\max(0, \hat{r}_i - \nu - \mu f_i)^2}{\sigma_i^2}.$$

We recover the optimal w, z as above, from a unicity argument.

4. (*10 points*) Let $A \in \mathbb{R}^{m \times n}$, $y \in \mathbb{R}^m$ and $\mu > 0$. Consider the problem

$$\min_x \|Ax - y\|_1 + \mu \|x\|_2.$$

(a) Express the problem in standard SOCP format.

(b) Find a dual to the problem. *Hint:* use the fact that, for any vector z:

$$\max_{u \,:\, \|u\|_2 \le 1} u^T z = \|z\|_2, \quad \max_{u \,:\, \|u\|_\infty \le 1} u^T z = \|z\|_1.$$

(c) Does strong duality hold? *Hint:* apply Sion's theorem.

(d) Assume A is 100×10^6. Which problem would you solve, the primal or the dual? Justify your answer carefully.

Solution:

(a) The problem writes

$$\min_{x,z,t} z^T \mathbf{1} + \mu t \;:\; t \ge \|x\|_2, \;\; z_i \ge |(Ax - y)_i|, \;\; i = 1, \ldots, m. \tag{1}$$

(b) Based on the hint, we use the Lagrangian

$$\mathcal{L}(x, u, v) = u^T (Ax - y) + v^T x,$$

which is such that

$$p^* = \min_x \max_{u,v} \left\{ \mathcal{L}(x, u, v) \;:\; \|u\|_\infty \le 1, \;\; \|v\|_2 \le \mu \right\}. \tag{2}$$

Exchanging min and max leads to the dual:

$$p^* \ge d^* = \max_{u,v} g(u, v),$$

with g the dual function

$$g(u, v) = \min_x \mathcal{L}(x, u, v) = \begin{cases} -u^T y & \text{if } A^T u + v = 0, \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem writes

$$d^* = \max_u -u^T y \;:\; A^T u + v = 0, \;\; \|u\|_\infty \le 1, \;\; \|v\|_2 \le \mu.$$

We can eliminate v:

$$d^* = \max_u -u^T y \;:\; \|u\|_\infty \le 1, \;\; \|A^T u\|_2 \le \mu.$$

(c) Strong duality holds, due to the application of Sion's theorem to the expression (2).

(d) The dual problem writes

$$d^* = \max_u \; -u^T y \;:\; \|u\|_\infty \leq 1, \;\; u^T K u \leq \mu,$$

with $K = AA^T$ a 100×100 matrix. In this form, the dual problem is an SOCP with 100 variables and 101 constraints. In contrast, the primal problem in SOCP format (1) has $\sim 10^6$ variables and 101 constraints. Therefore, the dual form is much better.