EE 127 / EE 227AT 10/25/18
L. El Ghaoui

HW #4 Solutions

Exercise 1 (Quadratic inequalities) Consider the set defined by the following inequalities

$$(x_1 \geq x_2 - 1 \text{ and } x_2 \geq 0) \text{ or } (x_1 \leq x_2 - 1 \text{ and } x_2 \leq 0).$$

1. Draw the set. Is it convex?
2. Show that it can be described as a single quadratic inequality of the form $q(x) = x^\top Ax + 2b^\top x + c \leq 0$, for matrix $A = A^\top \in \mathbb{R}^{2,2}$, $b \in \mathbb{R}^2$ and $c \in \mathbb{R}$ which you will determine.
3. What is the convex hull of this set?

Solution 1

1. The set is not convex, as shown by counterexample. $(0,1)$ and $(-2,0)$ both belong to the set, but the midpoint $(-1,1/2)$ does not.
2. Within this set, $x_1 - x_2 + 1 \geq 0$ when $x_2 \geq 0$ and $x_1 - x_2 + 1 \leq 0$ when $x_2 \leq 0$. It follows that $q(x) = x_2(x_2 - x_1 - 1) \leq 0$ if and only if it is in the set. Expressing $q(x)$ in the desired form:

$$\begin{aligned} q(x) &= x_2^2 - x_1x_2 - x_2 \\ \implies q(x) &= x^\top Ax + 2b^\top x + c \geq 0, \end{aligned}$$

   where
$$A = \begin{bmatrix} 0 & -1/2 \\ -1/2 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ -1/2 \end{bmatrix}, \quad c = 0.$$

3. The convex hull of the set is the whole space, $\mathbb{R}^2$.

Exercise 2 (Formulating problems as LPs or QPs) Formulate the problem

$$p_j^* \doteq \min_x \ f_j(x),$$

for different functions f_j, $j = 1, \ldots, 4$, as QPs or LPs, or, if you cannot, explain why. In our formulations, we always use $x \in \mathbb{R}^n$ as the variable, and assume that $A \in \mathbb{R}^{m,n}$, $y \in \mathbb{R}^m$. If you obtain an LP or QP formulation, make sure to put the problem in standard form, stating precisely what the variables, objective, and constraints are.

$$\begin{array}{rcl} f_1(x) & = & \|Ax - y\|_\infty + \|x\|_1 \\ f_2(x) & = & \|Ax - y\|_2^2 + \|x\|_1 \\ f_3(x) & = & \|Ax - y\|_2^2 - \|x\|_1 \\ f_4(x) & = & \|Ax - y\|_2^2 + \|x\|_1^2 \end{array}$$

Solution 2

1. For p_1^*, we replace the ℓ_∞ norm with constraints on the maximum absolute value of each element of $Ax - y$, and rewrite absolute values as linear constraints. This gives us the LP formulation:

$$p_1^* = \min_{x,t,z} \ t + \mathbb{1}^T z \ : \quad \begin{array}{l} z_i \geq x_i \geq -z_i, \ \ i = 1, \ldots, n \\ t \geq (Ax - y)_i \geq -t, \ \ i = 1, \ldots, m. \end{array}$$

2. For p_2^*, we obtain the convex QP

$$p_2^* = \min_{x,z} \ x^\top (A^\top A) x - 2y^\top Ax + y^\top y + \mathbb{1}^T z \ : \quad z_i \geq x_i \geq -z_i, \ \ i = 1, \ldots, n.$$

3. For p_3^*, the problem is not convex. Consider for instance the special case with $n = 1$, $A = 1$, $y = 0$: plot $f_3(x) = x^2 - |x|$ to verify it is not convex. In general, the objective of the *difference* of two convex functions is not necessarily convex.

4. For p_4^*, we have the convex QP

$$p_4^* = \min_{x,z} \ x^\top (A^\top A) x - 2y^\top Ax + y^\top y + \left(\sum_{i=1}^n z_i \right)^2 \ : \quad z_i \geq x_i \geq -z_i, \ \ i = 1, \ldots, n.$$

   Notice that $\left(\sum_{i=1}^n z_i\right)^2 = z^T Q z$, where we define Q as an $n \times n$ matrix of all ones. Thus, our problem in standard form is:

$$p_4^* = \min_{x,z} \ x^\top (A^\top A) x - 2y^\top Ax + y^\top y + z^T Q z \ : \quad z_i \geq x_i \geq -z_i, \ \ i = 1, \ldots, n.$$

Exercise 3 (Regularized least-squares problem with low-rank data) Consider the problem

$$p^* \doteq \min_x \ \|Ax - y\|_2^2 + \lambda \|x\|_2^2$$

where $A \in \mathbb{R}^{m,n}$, $y \in \mathbb{R}^m$ and the penalty parameter $\lambda \geq 0$ are given. In this exercise, we show that if A is low-rank, and a low-rank decomposition is known, then we can exploit that structure to greatly reduce the computational burden involved.

1. Assume that A has rank $r << \min(n, m)$, and that a low-rank decomposition of A is of the form $A = LR^T$ is known, with $L \in \mathbb{R}^{m,r}$, $R \in \mathbb{R}^{n,r}$ both full rank. Explain how to reduce the problem to one with r variables only.

2. Discuss the complexity of the method and compare it to that of a "direct" approach where the low-rank structure is ignored.

3. In the notebook `lpqp_low_rank_ls.ipynb`, we have generated some random data for $r = 10, m = 1000, n = 1000$. Implement the direct approach and compare it with the low-rank approach.

4. We investigate how this approach can be used even when the data is full-rank. In particular, we generated a full-rank data matrix A for $m = 300, n = 300$ in the second part of the notebook `lpqp_low_rank_ls.ipynb` but for which most of the variance is explained in the first $\hat{r}$ eigenvalues. In that particular case, we provide you with code that approximates your matrix A with $\hat{A}$ of rank r for all r, solves the low-rank problem based on your previous code, and plots the error p^* as a function of r. Use this graph to make a guess about $\hat{r}$ and interpret.

Solution 3

1. With $A = LR^T$, we decompose the variable x as

   $$x = Ru + v, \quad R^T v = 0,$$

   with $u \in \mathbb{R}^r$.

   The problem becomes

   $$\min_{u,v} \ \|L(R^T R)u - y\|_2^2 + \lambda(\|Ru\|_2^2 + \|v\|_2^2).$$

   The optimal v is zero, and the problem is reduced to one with r variables only.

   $$\min_u \ \|L(R^T R)u - y\|_2^2 + \lambda \|Ru\|_2^2. \tag{1}$$

   The solution expresses as

   $$u^* = (\lambda R^T R + (R^T R)L^T L(R^T R))^{-1}(R^T R)L^T y,$$

   and the optimal variable x is

   $$x^* = Ru^* = R(\lambda R^T R + (R^T R)L^T L(R^T R))^{-1}(R^T R)L^T y. \tag{2}$$

2. The solution to the original problem writes

$$x^* = (\lambda I + A^T A)^{-1} A^T y.$$

The cost of computing the above expression is as follows:

- Forming the $n \times n$ matrix $A^T A$, with $A \in \mathbb{R}^{m,n}$: $O(mn^2)$.
- Forming the vector $A^T y$: $O(nm)$.
- Solving the $n \times n$ system $(\lambda I + A^T A)x = A^T y$: $O(n^3)$.

The complexity of solving the original regularized least-squares problem is thus $O(mn^2 + n^3)$.

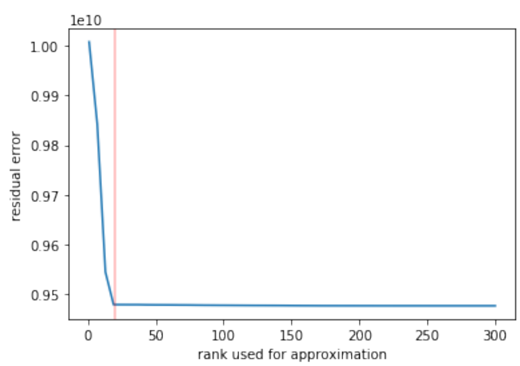
For the reduced problem, the cost is as follows.

- Forming the $r \times r$ matrix $R^T R$, with $R \in \mathbb{R}^{n,r}$: nr^2.
- Forming $L(R^T R)$, with $L \in \mathbb{R}^{m,r}$: $O(mr^2)$.
- Solving the problem via the formula (2): $O(r^3 + r(m+n))$.

The total complexity goes as $O((m+n)r^2)$ (assuming $r << \min(m,n)$). This is a dramatic reduction: the problem's complexity is now linear in m, n (for fixed r).

3. See the notebook `lpqp_low_rank_ls_sols` for full detail. Essentially, the low-rank solver is at least 50x faster than the full-rank method.

4. We observe that after 20 dimensions, the error does not improve much. This suggests that $\hat{r} = 20$.



Exercise 4 (A Boolean problem of maximal margin) We have to decide which items to sell among of a collection of n items, with given selling prices $s_i > 0$, $i = 1, \ldots, n$. Selling implies a transaction cost $c_i > 0$ for each item i; the total transaction cost is the sum of these costs, plus a fixed amount $d > 0$ (say, $d = 1$). We are interested in the following decision problem: which items need to be sold, in order to maximize the margin (revenue-to-cost ratio). *Assume a transaction will occur for all items we choose to sell.*

1. Show that the problem can be formalized as

$$p^* \doteq \max_{x \in \{0,1\}^n} f(x), \quad f(x) \doteq \frac{s^\top x}{1 + c^\top x}.$$

2. Show that the problem admits a one-dimensional formulation:

$$p^* = \min_{t \geq 0} t \ : \ t \geq \mathbf{1}^\top (s - tc)_+.$$

   Here z_+, for a vector $z \in \mathbb{R}^n$, denotes the vector with components $\max(0, z_i)$, $i = 1, \ldots, n$. *Hint:* for given $t \geq 0$, express the condition $f(x) \leq t$ for every $x \in \{0,1\}^n$ in simple terms.

3. Let us assume the inequality constraint $t \geq \mathbf{1}^\top (s - ct)_+$ is active at optimum. How can you recover an optimal solution x^* from an optimal value t^* for the above problem?

Solution 4

1. Let us parametrize our decision with a Boolean vector $x \in \{0,1\}^n$, with $x_i = 1$ if item i is sold, $x_i = 0$ if not. The total revenue is then $s^\top x$, while the total transaction cost is $1 + c^\top x$. This proves the result.

2. For given $t \geq 0$, the condition $f(x) \leq t$ for every $x \in \{0,1\}^n$ is equivalent to

$$\forall\, x \in \{0,1\}^n \ : \ s^\top x \leq t(1 + c^\top x),$$

   or:

$$t \geq \max_{x \in \{0,1\}^n} (s - ct)^\top x = \mathbf{1}^\top (s - ct)_+.$$

   The desired formulation follows.

3. At optimum, the inequality constraint $t \geq \mathbf{1}^\top (s - ct)_+$ is active, hence

$$t^* = \max_{x \in \{0,1\}^n} (s - ct^*)^\top x.$$

   Therefore, there exists a vector $x \in \{0,1\}^n$ such that

$$\max_{x \in \{0,1\}^n} (s - ct^*)^\top x = (s - ct^*)^\top x^* = t^*.$$

By construction, x^* achieves the optimal value:

$$t^* = \frac{s^\top x^*}{1 + c^\top x^*}.$$

Since x^* is feasible, it is optimal.

We can identify x^* more precisely, setting

$$\forall i = 1, \ldots, m \;:\; x_i^* = \begin{cases} 1 & \text{if } s_i > t^* c_i, \\ 0 & \text{otherwise.} \end{cases}$$

Exercise 5 (Robust production plans) Recall the drug production problem discussed in lectures 10 (slide 22) and 13. A company produces two kinds of drugs, DrugI and DrugII, containing a specific active agent A, which is extracted from raw materials purchased on the market. There are two kinds of raw materials, RawI and RawII, which can be used as sources of the active agent. The related production, cost and resource data are given next. The goal is to find the production plan which maximizes the profit of the company.

Let $x_{\rm rmDrugI}$, $x_{\rm rmDrugII}$ denote the amounts (in 1000 of packs) of Drug I and II produced, while $x_{\rm rmRawI}$, $x_{\rm rmRawII}$ denote the amounts (in kg) of raw materials to be purchased. We seek to find the production and raw material amounts that maximize profit (revenue minus cost). Furthermore, we are subject to constraints on the balance of the active agent, storage, manpower, equipment, and budget, and constrain all variables to be non-negative.

Putting this together, we get the LP:

$$\min_x c^T x \quad : Ax \leq b,\ x \geq 0.$$

where $x = (x_{\rm RawI}, x_{\rm RawII}, x_{\rm DrugI}, x_{\rm DrugII})$. For details on how this optimization problem is formulated, please see slide 22 in lecture 10.

We now assume that the vector c, which represents cost minus revenue, is subject to scenario uncertainty. Precisely, c is only known to belong to a given set of the form

$$\mathcal{C} = \left\{ c^{(k)} = \hat{c} + \rho\delta c^{(k)},\ \ k = 1, \ldots, K \right\},$$

where $\hat{c} \in \mathbb{R}^4$ is the nominal value, $\delta c^{(k)}$ is the k-th scenario and $\rho \geq 0$ is a measure of uncertainty. To any given candidate solution $x \in \mathbb{R}^4$, we define the worst-case cost as

$$p^{\rm wc}(x) := \max_{c \in \mathcal{C}} c^T x.$$

We have loaded the data for this problem in the notebook `drug_prod.ipynb`.

1. Solve for the nominal problem's solution $x^{\rm nom}$, in which $c = \hat{c}$, and compute the profit. Compute the worst-case profit of this plan.

2. Derive the robust counterpart of the problem, based on optimizing the worst-case cost. Compare the robust solution to the nominal one, in terms of nominal cost and in terms of worst-case cost, and comment.

3. Plot the elements of the robust solution as the uncertainty level ρ grows. What is the smallest value of ρ for which the robust solution is zero?

Solution 5

1. We solve the LP

$$\min_x c^T x \text{ s.t. } Ax \le b, x \ge 0$$

to get $x^{\text{nom}} = [9.9034e^{-6}, 438.79, 17.552, 2.617e^{-10}]$. In this case, the profit would be \$8819.7. The worst-case profit of the plan is \$1515.5

2. We formulate the robust problem as

$$\begin{aligned} \min_x \quad & \max_{c \in C} c^T x \\ \text{s.t.} \quad & Ax \le b \\ & x \ge 0 \end{aligned}$$

Suppose we hold x constant. Using the epigraph trick, we can rewrite $\max_{c \in C} c^T x$ as

$$\min_t t \text{ s.t. } c^T x \le t \ \ \forall c \in C$$

by introducing a slack variable t. Substituting, our robust problem is now

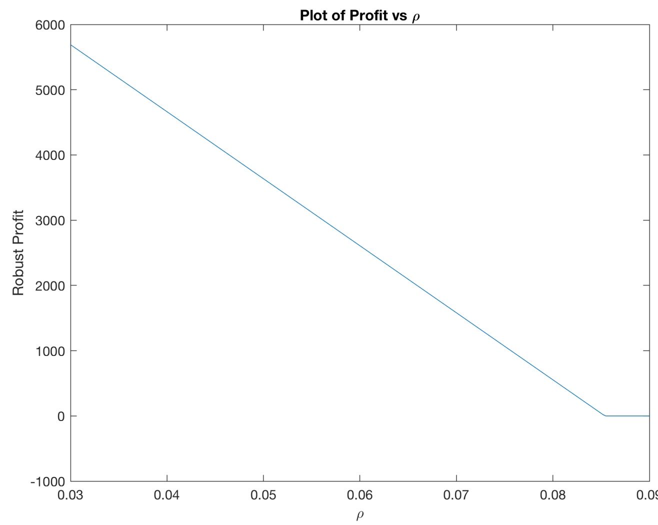
$$\begin{aligned} \min_{x,t} \quad & t \\ \text{s.t.} \quad & c^T x \le t, \quad c \in C \\ & Ax \le b \\ & x \ge 0. \end{aligned}$$

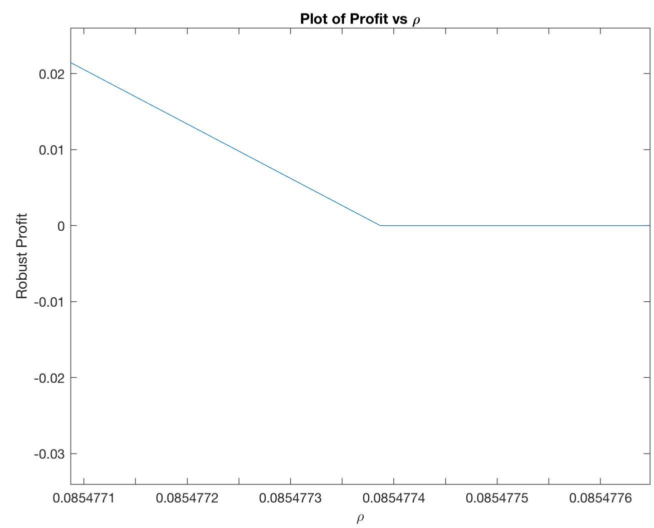
This is a LP with solution is $x^{\text{rob}} = [877.19, 1.862e^{-6}, 17.54, 7.909e^{-9}]$. The profit made in this case is \$5689.9. We see that due to the uncertainty of the cost vector c, the profit is lower in the robust case than in the nominal case. Below is a table summarizing the generated results:

nom	rob
9.9034e-06	877.19
438.79	1.8629e-06
17.552	17.544
2.6174e-10	7.9091e-09

nom	rob	wc
8819.7	5689.9	1515.5

3. Below, the robust profit is plotted as ρ grows. We see that the smallest value of ρ where the profit is 0 is approximately 0.0854.


Plot of Profit vs ρ
Robust Profit
ρ
6000
5000
4000
3000
2000
1000
0
-1000
0.03
0.04
0.05
0.06
0.07
0.08
0.09


Plot of Profit vs ρ
Robust Profit
ρ
0.02
0.01
0
-0.01
-0.02
-0.03
0.0854771
0.0854772
0.0854773
0.0854774
0.0854775
0.0854776

Exercise 6 (Robust linear programming) In this problem we will consider a version of linear programming under uncertainty.

1. Let $x \in \mathbb{R}^n$ be a given vector. Prove that $x^T y \leq \|x\|_1$ for all y such that $||y||_\infty \leq 1$. Is this inequality tight? (Is there always a y such that the equality holds ?)

Let us focus now on a LP in standard form:

$$\begin{aligned} &\min_x c^T x \\ \text{s.t. } & a_i^T x \leq b_i, \quad i = 1, ..., m \end{aligned} \tag{3}$$

Consider the set of linear inequalities in (3). Suppose you don't know the coefficients a_i exactly. Instead you are given nominal values $\overline{a}_i$, and you know that the actual coefficient vectors satisfy $\|a_i - \overline{a}_i\|_\infty \leq \rho$ for a given $\rho > 0$. In other words, the actual coefficients a_{ij} can be anywhere in the intervals $[\overline{a}_{ij} - \rho, \overline{a}_{ij} + \rho]$. or equivalently, each vector a_i can lie anywhere in a rectangle with corners $\overline{a}_i + v$ where $v \in \{-\rho, \rho\}^n$. The set of inequalities that constrain problem 3 must be satisfied for all possible values of a_i; i.e., we replace these with the constraints

$$a_i^T x \leq b_i \;\; \forall a_i \in \{\overline{a}_i + v \;\; | \; \|v\|_\infty \leq \rho\} \;\; i = 1, ..., m. \tag{4}$$

A straightforward but very inefficient way to express this constraint is to enumerate the 2^n corners of the rectangle of possible values a_i and to require that

$$\overline{a}_i^T x + v^T x \leq b_i \;\; \forall v \in \{-\rho, \rho\}^n \;\; i = 1, ..., m.$$

2. Use the previous result to show that (4) is in fact equivalent to the much more compact set of nonlinear inequalities

$$\overline{a}_i^T x + \rho \|x\|_1 \leq b_i, \quad i = 1, ..., m.$$

We now would like to formulate the uncertainty in the LP we introduced. We are therefore interested in situations where the coefficient vectors a_i are uncertain, but satisfy bounds $\|a_i - \overline{a}_i\|_\infty \leq \rho$ for given $\overline{a}_i$ and ρ. We want to minimize $c^T x$ subject to the constraint that the inequalities $a_i^T x \leq b_i$ are satisfied for *all* possible values of a_i.

We call this a *robust LP* :

$$\begin{aligned} &\min_x c^T x \\ \text{s.t. } & a_i^T x \leq b_i, \quad \forall a_i \in \{\overline{a}_i + v \mid \|v\|_\infty \leq \rho\} \quad i = 1, ..., m. \end{aligned} \tag{5}$$

3. Using the result from part 2, express the above optimization problem as an LP.

Solution 6

1. Our optimization problem is:

$$\max_y x^T y : \|y\|_\infty \leq 1$$

   The optimal value is bounded by $\|x\|_1$. We have that $\max_{y_i} x_i y_i = |x_i|$, where $y_i^* = \text{sign}(x_i)$. This implies that $y^* = sign(x)$. This is a tight inequality, as $x^T sign(y) = \|x\|_1$.

2. Reformulating the inequality, we have:

$$\overline{a}_i^T x + \max_v v^T x \leq b_i \;\; \forall v \in [-\rho, \rho]^n \;\; i = 1, ..., m.$$

   This inequality can be written as:

$$\overline{a}_i^T x + \max_v \rho v^T x \leq b_i \;\; \forall v \in [-1, 1]^n \;\; i = 1, ..., m.$$

   or equivalently as

$$\overline{a}_i^T x + \rho \max_{\|v\|_\infty = 1} v^T x \leq b_i \;\; \forall i = 1, ..., m.$$

   From part 1, we have $\max_{\|y\|_\infty \leq 1} x^T y = \|x\|_1$, which implies:

$$\overline{a}_i^T x + \rho \|x\|_1 \leq b_i, \quad i = 1, ..., m.$$

3. We can rewrite the constraint as:

$$\begin{aligned} &\min_x c^T x \\ \text{s.t.} \;\; &(\overline{a}_i + v)^T x \leq b_i, \quad i = 1, ..., m. \end{aligned}$$

   Which is equivalent to:

$$\begin{aligned} &\min_x c^T x \\ \text{s.t.} \;\; &\overline{a}_i^T x + \max_v \rho v^T x \leq b_i, \;\; \forall v \in \{-1, 1\}^n, \quad i = 1, ..., m. \end{aligned}$$

   Using our constraint from part 2, we can express the nonlinear optimization as:

$$\begin{aligned} &\min_x c^T x \\ \text{s.t.} \;\; &\overline{a}_i^T x + \rho \|x\|_1 \leq b_i, \;\; i = 1, ..., m. \end{aligned}$$

We can express this optimization problem as an LP by introducing variables t_i:

$$\begin{aligned} &\min_{x,t} c^T x \\ \text{s.t. } &\overline{a}_i^T x + \rho \sum_i t_i \leq b_i, \quad i = 1, ..., m. \\ &x_i \leq t_i \;\; i = 1, ..., m. \\ &- x_i \leq t_i \;\; i = 1, ..., m. \end{aligned} \tag{6}$$