EE5121: Convex Optimization

CVX Assignment

Instructions:

- **Deadline:** Solutions need to be submitted via Moodle on or before 9th May, 2019.
- **Submission Format:** Submit a single .zip file named <rollno>.zip with the following contents:
  1. One .m file for each question (with comments explaining the code) named <rollno>_<questionnumber>.m.
  2. A report (named <rollno>.pdf). In the report, please include a screenshot of the MATLAB command window for each question, clearly showing the optimal value and other details output by CVX. For each question, the final optimization formulation should be provided in the report.

Questions

1. **(Recovering a piecewise constant signal from its noisy measurement)** Consider a source which generates piecewise constant signals. Suppose you receive a noisy measurement of a signal generated by this source (given in the file "piecewise_constant_data.mat"). Your goal is to recover the original piecewise constant signal as accurately as possible. Formally, let x be the piecewise constant signal generated by the source, y be the data received by you, $\hat{x}$ be your estimate of x, and $e := y - \hat{x}$ be the error between the received data and your estimate of the generated signal. Your goal is to minimize the squared error, i.e, the 2-norm of e and to have as few 'jumps' in the estimated signal $\hat{x}$ as possible (since the generated signal x is piecewise constant). Formulate this as a convex optimzation problem as follows.

   Construct an appropriate objective function which minimizes the weighted sum of (a) the 2-norm of e, (b) the number of jumps in $\hat{x}$. Formulate this as an SOCP (second-order conic program). Solve the problem using CVX. Adjust the weights in the objective function until you get a signal $\hat{x}$ which has no more than (approximately) 20 jumps. (Note: This is not a constraint but simply a criterion to know when to stop adjusting the weights.) Give the optimal value of the 2-norm of e and plot a graph of the optimal $\hat{x}$ obtained.

   Hints: (i) To minimize the number of jumps in $\hat{x}$, try to make use of a matrix of the form

$$A = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix}.$$

   (ii) Suppose $\text{card}(x)$ denotes the number of nonzero entries in a vector x. It can be approximated by $||x||_1$ to bring the problem in convex form.

2. We consider the selection of n non-negative activity levels, denoted by $x_1, \dots, x_n$. These activities consume m resources, which are limited. Activity j consumes A_{ij} of resource i, where A_{ij} are given. The total resource consumption is additive, so the total of resource i consumed is $c_i = \sum_{j=1}^{n} A_{ij} x_j$ (Ordinarily we have $A_{ij} \geq 0$, i.e., activity j consumes resource i. But we allow the possibility that $A_{ij} \leq 0$, which means that activity j actually generates resource i as a by-product.) Each resource consumption is limited: we must have $c_i \leq c_i^{max}$, where c_i^{max} are given. Each activity generates revenue, which is a piecewise-linear concave function of the activity level:

$$r_j = \begin{cases} p_j x_j, & 0 \leq x_j \leq q_j \\ p_j q_j + p_j^{disc}(x_j - q_j), & x_j \geq q_j \end{cases}$$

   Here $p_j \geq 0$ is the basic price, $q_j \geq 0$ is the quantity discount level, and p_j^{disc} is the quantity discount price, for (the product of) activity j. (We have $0 \leq p_j^{disc} \leq p_j$) The total revenue is the sum of the revenues associated with each activity, i.e., $\sum_{j=1}^{n} r_j(x_j)$ The goal is to choose activity levels that maximize the total revenue

while respecting the resource limits. Show how to formulate this problem as an LP.

Problem data

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 3 & 1 \\ 0 & 3 & 1 & 1 \\ 2 & 1 & 2 & 5 \\ 1 & 0 & 3 & 2 \end{bmatrix}, \quad c^{max} = \begin{bmatrix} 100 \\ 100 \\ 100 \\ 100 \\ 100 \end{bmatrix}, \quad p = \begin{bmatrix} 3 \\ 2 \\ 7 \\ 6 \end{bmatrix}, \quad p^{disc} = \begin{bmatrix} 2 \\ 1 \\ 4 \\ 2 \end{bmatrix}, \quad q = \begin{bmatrix} 4 \\ 10 \\ 5 \\ 10 \end{bmatrix}$$

Use CVX to determine the optimal activity levels, the revenue generated by each one, and the total revenue generated by the optimal solution. Also, give the average price per unit for each activity level, i.e., the ratio of the revenue associated with an activity, to the activity level. (These numbers should be between the basic and discounted prices for each activity.) Comment on the solution you find.

3. **[Matrix Completion Problem: Netflix Data]** Matrix completion problems refer to a class of problems in which one aims at recovering a matrix X from incomplete information about its entries with some apriori knowledge and assumptions.

   A famous example in this class is the Netflix Problem. This is a problem arising in a recommender system, where rows in matrix X represent users and columns represent movies. Users are given the opportunity of giving a rating for the movies. However, each user rates only a few movies and hence only a few entries of the X matrix are known. Yet, it would be interesting to complete this matrix by "guessing" the missing entries, so that Netflix might recommend movies that a user may like. Here, one may observe that users' preferences are a function of a few factors (such as genre, protagonist, director, etc.), hence rank of matrix X is not more than the independent factors that were used in the rating. This suggests the idea that X can be completed by finding the matrix of minimum rank which is compatible with the available entries. The optimization problem can be written as:

$$\begin{aligned} &\text{Minimize } \operatorname{rank}(\hat{X}) \\ &\text{Subject to } \hat{X}_{ij} = X_{ij} \; \forall (i,j) \in J \end{aligned} \tag{1}$$

   where J is the set of indices with known elements of X with $|J| <$ No of users $\times$ No of movies. Now this is not a convex problem as rank is not a convex function. However, we can relax this problem when $X \succeq 0$ into minimizing the trace of X which is a convex function. We can use the following result to apply this relaxation for any matrix X.

   "Given $X \in \mathbb{R}^{m \times n}$, rank $X \leq r$ if and only if $\exists\, Y \in \mathbf{S}^m$ and $Z \in \mathbf{S}^n$ such that rank$Y$ + rank$Z \leq 2r$ and $\left[\begin{smallmatrix} Y & X \\ X^T & Z \end{smallmatrix}\right] \succeq 0$, where Y and Z are symmetric matrices."

   Using the relaxation and this result, re-formulate the optimization problem in Eq. (1) as an SDP and write a CVX code for it. An incomplete matrix X is provided in the file "Ratings.mat" which is sampled from data released by Netflix. It consists of ratings between 1-5 given by 50 users for 38 movies. A zero element in that matrix indicates no rating given. Complete this matrix X by solving the SDP and mention the rank of the completed matrix along with the SDP formulation, in the report. The completed matrix may contain decimal values as well and no integer constraints are required.

   Note: Refer to the CVX documentation for coding an SDP problem in CVX.