EE5121: Convex Optimization

Programming Assignment

Instructions:

- **Deadline:** Solutions need to be submitted via Moodle before May 15th, 12pm
- **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.

Practice questions:

[Reformulating constraints in CVX] The following question is for you to get acquainted with CVX, and *need not be submitted*. Each of the following CVX code fragments describes a convex constraint on the scalar variables x, y, and z, but violates the CVX rule set, and so is invalid. Hence, re-write each one in an equivalent form that conforms to the CVX rule set. In your reformulations, you can use linear equality and inequality constraints, and inequalities constructed using CVX functions. You can also introduce additional variables, or use LMIs. You can check your reformulations by creating a small problem that includes these constraints, and solving it using CVX. Your test problem does not have to be feasible; it is enough to verify that CVX processes your constraints without error.

1. `norm([x + 2*y, x - y]) == 0`
2. `square(square(x + y)) <= x - y`
3. `1/x + 1/y <= 1; x >= 0; y >= 0`
4. `norm([max(x,1), max(y,2)]) <= 3*x + y`
5. `x*y >= 1; x >= 0; y >= 0`
6. `(x + y)^2/sqrt(y) <= x - y + 5`
7. `x^3 + y^3 <= 1; x >= 0; y >= 0`
8. `x + z <= 1 + sqrt(x*y - z^2); x >= 0; y >= 0`

Questions:

1. You are given a set of $m = 50$ points $(x_i, y_i), i = 1, \ldots, m$ in $\mathbb{R}^2$ in the file `circle_fit.m`. Use `CVX` to find a circle

$$(x_c - x)^2 + (y_c - y)^2 = r^2 \tag{1}$$

that best fits the given set of points in the least-squares sense. This can be done by finding x_c, y_c, r which solve the problem

$$\underset{x_c, y_c, r}{\text{minimize}} \sum_{i=1}^{m} \left((x_c - x_i)^2 + (y_c - y_i)^2 - r^2 \right)^2. \tag{2}$$

(a) Show that the above problem is equivalent to

$$\underset{z\in\mathbb{R}^3}{\text{minimize}} \ ||Az-b||^2 \tag{3}$$

by appropriately defining $z \in \mathbb{R}^3$ in terms of x_c, y_c, r and $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$ in terms of the data points $(x_i, y_i), i = 1, \ldots, m$.

(b) Solve (3) using `CVX`. Then, plot the data points and the circle you found using `CVX` in the same figure. You may use the following code.

```
t = linspace(0, 2*pi, 1000);
plot(x, y, 'o', r*cos(t) + x_c, r*sin(t) + y_c, '-');
```

Include in your report the above plot and the optimal values of x_c, y_c, r.

2. Consider the problem

$$\underset{x_1,x_2}{\text{minimize}} \ f(x) = e^{x_1+3x_2-0.1} + e^{x_1-3x_2-0.1} + e^{-x_1-0.1}. \tag{4}$$

Use $x^0 = (-1, 0.7)$ as the starting point and $||\nabla f(x)|| \leq 0.01$ as the stopping criterion and find the optimal value of the problem using the following methods.

(a) Gradient descent with fixed step-size of 0.1

(b) Gradient descent with backtracking line-search with parameters $\alpha = 0.1, \beta = 0.5$ (Algorithm 9.2 in Convex Opt. by Boyd)

(c) Newton's method with backtracking line-search with same parameters as above

For each method, include in your report the optimal value x^* and a plot of the trajectory $x^k \in \mathbb{R}^2, k = 0, 1, \ldots$. You may use the following code to plot the trajectory after downloading the file `draw_line.m`.

```
plot(x(1,1),x(2,1),'rx')
hold on
for i = 2:end
plot(x(1,i),x(2,i),'rx')
draw_line(x(:,i-1),x(:,i))
hold on
end
```

3. Consider the *boolean* optimization problem

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & c^T x \\ \text{subject to} \quad & Ax \leq b, \\ & x_i \in \{0,1\}, i = 1, \ldots, n, \end{aligned} \tag{5}$$

where $A \in \mathbb{R}^{m\times n}$.

(a) Show that this problem is equivalent to the problem

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & c^T x \\ \text{subject to} \quad & Ax \leq b, \\ & x_i(1-x_i) = 0, i = 1, \ldots, n. \end{aligned} \tag{6}$$

(b) Is problem (6) convex? Provide justification.

(c) Find the dual of problem (6).

Hint: Let $\lambda \in \mathbb{R}^m$ be the dual variable for the constraint $Ax \leq b$ and $\mu \in \mathbb{R}^n$ be the dual variable for the constraint $x_i(1-x_i) = 0, i = 1, \ldots, n$. First, show that the dual problem is given by

$$\begin{aligned} \underset{\lambda\in\mathbb{R}^m,\mu\in\mathbb{R}^n}{\text{maximize}} \quad & -b^T\lambda - \frac{1}{4}\sum_{i=1}^{n} \frac{(c_i + a_i^T\lambda - \mu_i)^2}{\mu_i} \\ \text{subject to} \quad & \lambda \geq 0, \mu \geq 0, \end{aligned} \tag{7}$$

where $a_i \in \mathbb{R}^m$ is the ith column of A. Then, for each i in the summation above, consider two cases—$c_i + a_i^T \lambda \geq 0$ and $c_i + a_i^T \lambda < 0$. Then, show that the dual problem can be simplified to

$$\begin{array}{ll} \underset{\lambda \in \mathbb{R}^m}{\text{maximize}} & -b^T \lambda + \sum_{i=1}^{n} \min\{0, c_i + a_i^T \lambda\} \\ \text{subject to} & \lambda \geq 0. \end{array} \tag{8}$$

(d) Argue that the dual problem (8) is convex.

(e) You are given a data for the above problem in the file `boolean.m`. Solve the dual problem with this data using `CVX`.

(f) What can you say about the optimal value d^* of the dual compared to that of the original problem p^*? (No explanation required, just write the relation.)

4. A linear dynamical system has the form

$$x(t+1) = Ax(t) + Bu(t) + w(t), \;\; t = 1, ..., T-1,$$

where $x(t) \in \mathbb{R}^n$, is the state, $u(t) \in \mathbb{R}^m$, is the input signal, and $w(t) \in \mathbb{R}^n$ is the process noise, at time t. We assume the process noises are *IID* $N(0, W)$, where $W > 0$ is the covariance matrix. The matrix $A \in \mathbb{R}^{n \times n}$ is called the dynamics matrix or the state transition matrix, and the matrix $B \in \mathbb{R}^{n \times m}$ is called the input matrix.

You are given accurate measurements of the state and input signal, *i.e.*, $x(1), ..., x(T), \;\; u(1), ..., u(T-1)$, and W is known.

Consider the error in estimation of $\hat{A}$ and $\hat{B}$ is given by the term

$$\sum_{t=1}^{T-1} ||(x(t+1) - \hat{A}x(t) - \hat{B}u(t)||_2^2$$

(a) Minimize the error in estimation.

(b) Now consider the trade-off between the estimation error and the sparsity of $\hat{A}$ and $\hat{B}$. Carry out your method on the data found in *Q4_data.mat*. Vary the trade-off parameter (say λ from 10 to 70 with interval of 2) associated with the sparsity of matrices. Plot the trade-off curve (pareto frontier). Plot the curve for the two objectives that you have in the objective function (sum of number of non zero elements in $\hat{A}$ and $\hat{B}$ vs estimation error), If you are to pick a point on the optimal-trade off curve which would you choose? Provide the justification accordingly.

To judge whether an entry of $\hat{A}$ (or $\hat{B}$) is nonzero, you can use the test $|\hat{A_{ij}}| \geq 0.01$ (or$|\hat{B_{ij}}| \geq 0.01$).

(Hints: The objective function be written in terms of cardinality of matrix *e.g,* card(X) is the cardinality (the number of nonzero entries) of matrix X. (Note: This is not convex problem))

5. *Uniform approximation:* Consider a function $f(x) = \sin(x)$ on the interval $I = [-\pi \quad \pi]$. we wish to approximate it uniformly by a polynomial of order k on N equally spaced grid points. Code the corresponding LP and record the following observations. plot the trajectory of the sine function and obtained polynomial function.

Finally, when you get to the problem. Use the data given in *Q5_data*.

(Hint: reduce the objective function to 1-norm minimization and in turn convert to LP)

6. Consider a correlation matrix C which has been manipulated by adding some random noise, resulting in a matrix $\hat{C}$. We remark that to be a valid correlation matrix, a matrix must be symmetric, positive semidefinite, and have all diagonal entries equal to one. The resulting matrix $\hat{C}$ is symmetric and has ones on the diagonal, but $\hat{C}$ is not positive semidefinite.

Suppose we have

$$\hat{C} = \begin{bmatrix} 1 & -0.76 & 0.07 & -0.96 \\ -0.76 & 1 & 0.18 & 0.07 \\ 0.07 & 0.18 & 1 & 0.41 \\ -0.96 & 0.07 & 0.41 & 1 \end{bmatrix}$$

Recover the original matrix by finding the nearest correlation matrix to $\hat{C}$ in Frobenius norm (i.e., the correlation matrix C that minimizes $||C - \hat{C}||_F$). Give your optimal solution.