Optimization Models: Homework #3

Due on October 10, 2025 at 11:59pm

Zachary Brandt
zbrandt@berkeley.edu

Problem 1

Consider an aerial system that moves in $\mathbb{R}^3$ according to the dynamics

$$x(k+1) = Ax(k) + Bu(k), \quad k = 0, 1, 2, 3 \tag{1}$$

where $x(k) \in \mathbb{R}^3$ is the position of the system at time $k \in \{0, 1, 2, 3, 4\}$ and $u(k) \in \mathbb{R}$ is the scalar input applied to the system at time k. Assume that the initial position $x(0)$ is equal to $[0\ \ 0\ \ 0]^T$. Given a target position $x_d \in \mathbb{R}^3$, the goal is to design the input sequence $u(0), u(1), u(2), u(3)$ to take the system to the target position x_d at time $k = 4$, i.e., $x(4) = x_d$.

i) Find a matrix $H \in \mathbb{R}^{3\times 4}$ in terms of A and B with the property that

$$x(4) = H \begin{bmatrix} u(0) \\ u(1) \\ u(2) \\ u(3) \end{bmatrix} \tag{2}$$

ii) Assume that

$$A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \tag{3}$$

Show that the vector $[1\ \ 1\ \ 0]^T$ belongs to $\mathcal{N}(H^T)$ (note: you are allowed to use a calculator to compute H, but you cannot use a calculator or a computer code to study the null space of H^T and the analysis should be done by hand).

iii) Again, consider the system parameters given in (3). By studying the relationship between $\mathcal{N}(H^T)$ and $\mathcal{R}(H)$, prove that there is no sequence of inputs that can take the system to the position $x_d = [1\ \ 1\ \ 0]^T$ at time 4.

iv) Again, consider the system parameters given in (3). By finding $\mathcal{N}(H^T)$ and using the relations $\mathcal{N}(H^T) \perp \mathcal{R}(H)$ and $\mathcal{N}(H^T) \oplus \mathcal{R}(H) = \mathbb{R}^3$, show that there exists a sequence of inputs to take the system to the position x_d at time 4 if and only if x_d belongs to the set

$$\{x \in \mathbb{R}^3 \mid x_1 + x_2 = 0\} \tag{4}$$

v) **(Coding)** Now, assume that

$$A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \tag{5}$$

The goal is to find a sequence of inputs such that the total energy $u(0)^2 + u(1)^2 + u(2)^2 + u(3)^2$ is minimized and yet the system arrives at the target position $x_d = [3\ \ 2\ \ 2]^T$ at time 4. Formulate this as an optimization problem and write a code in CVX to solve the problem numerically. Plot the optimal trajectory (i.e., plot the optimal values of the points $x(0), ..., x(4)$ in $\mathbb{R}^3$ and then connect each point to the next one (such as $x(1)$ to $x(2)$)).

vi) **(Coding)** Consider the safety set

$$\mathcal{S} = \{x \in \mathbb{R}^3 \mid -3.3 \le x_i \le 3.2, \quad i = 1, 2, 3\} \tag{6}$$

Assume that the state $x(k)$ must always stay in the safety set $\mathcal{S}$ for $k = 0, 1, ..., 4$. Redo Part (v) under this additional constraint and find the optimal input sequence. Compares the optimal trajectories and optimal energies (objective values) obtained in Parts (v) and (vi).

Problem 1 continued on next page. . .

Solution

i) To find a matrix H in terms of A and B, I'll solve for $x(k)$ at times 1, 2, 3, and 4

$$\begin{aligned}
x(1) &= Ax(0) + Bu(0) = Bu(0) \\
x(2) &= Ax(1) + Bu(1) = ABu(0) + Bu(1) \\
x(3) &= Ax(2) + Bu(2) = A^2Bu(0) + ABu(1) + Bu(2) \\
x(4) &= Ax(3) + Bu(3) = A^3Bu(0) + A^2Bu(1) + ABu(2) + Bu(3)
\end{aligned}$$

Therefore, the matrix H is

$$H = \begin{bmatrix} A^3B & A^2B & AB & B \end{bmatrix}$$

ii) To show that the vector $[1 \ \ 1 \ \ 0]^\top$ belongs to $\mathcal{N}(H^\top)$, I'll first compute H using the given A and B matrices

$$\begin{aligned}
AB &= \begin{bmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \\
A^2B &= \begin{bmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 1 \end{bmatrix} \\
A^3B &= \begin{bmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ -4 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ -8 \\ 1 \end{bmatrix}
\end{aligned}$$

Therefore, the matrix H is

$$H = \begin{bmatrix} A^3B & A^2B & AB & B \end{bmatrix} = \begin{bmatrix} 8 & 4 & 2 & 1 \\ -8 & -4 & -2 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Now, I'll compute $H^\top$ and multiply it by the vector $[1 \ \ 1 \ \ 0]^\top$

$$H^\top = \begin{bmatrix} 8 & -8 & 1 \\ 4 & -4 & 1 \\ 2 & -2 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

$$H^\top \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 8(1) + -8(1) + 1(0) \\ 4(1) + -4(1) + 1(0) \\ 2(1) + -2(1) + 1(0) \\ 1(1) + -1(1) + 1(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Since the result is the zero vector, the vector $[1 \ \ 1 \ \ 0]^\top$ belongs to $\mathcal{N}(H^\top)$.

iii) There is no sequence of inputs that can take the system to the position $x_d = [1 \ \ 1 \ \ 0]^\top$ at time 4 because x_d is not in the range of H. Since $[1 \ \ 1 \ \ 0]^\top$ is in the null space of $H^\top$, it is orthogonal to every vector in the range of H ($\mathcal{R}(H) \perp \mathcal{N}(H^\top)$). If there were a sequence of inputs that could take the system to the position x_d at time 4, then x_d would be in the range of H and would have to be orthogonal to itself, which is a contradiction since x_d is not the zero vector.

iv) To show that there exists a sequence of inputs to take the system to the position x_d at time 4 if and only if x_d belongs to the set $\{x \in \mathbb{R}^3 \mid x_1 + x_2 = 0\}$, I'll first find the null space of $H^\top$. From part (ii), we know $[1, 1, 0]^\top \in \mathcal{N}(H^\top)$. Since H is 3×4 and has rank 2 (the first two rows are linearly independent), we have $\dim(\mathcal{N}(H^\top)) = 3 - 2 = 1$. Therefore, $\mathcal{N}(H^\top) = \text{span}\{[1, 1, 0]^\top\}$. Using $\mathcal{N}(H^\top) \perp \mathcal{R}(H)$, a vector $x_d \in \mathcal{R}(H)$ is reachable if and only if $x_d \perp v$ for all $v \in \mathcal{N}(H^\top)$. Since $\mathcal{N}(H^\top) = \text{span}\{[1, 1, 0]^\top\}$, we need:

$$x_d^\top \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 0$$

which simplifies to

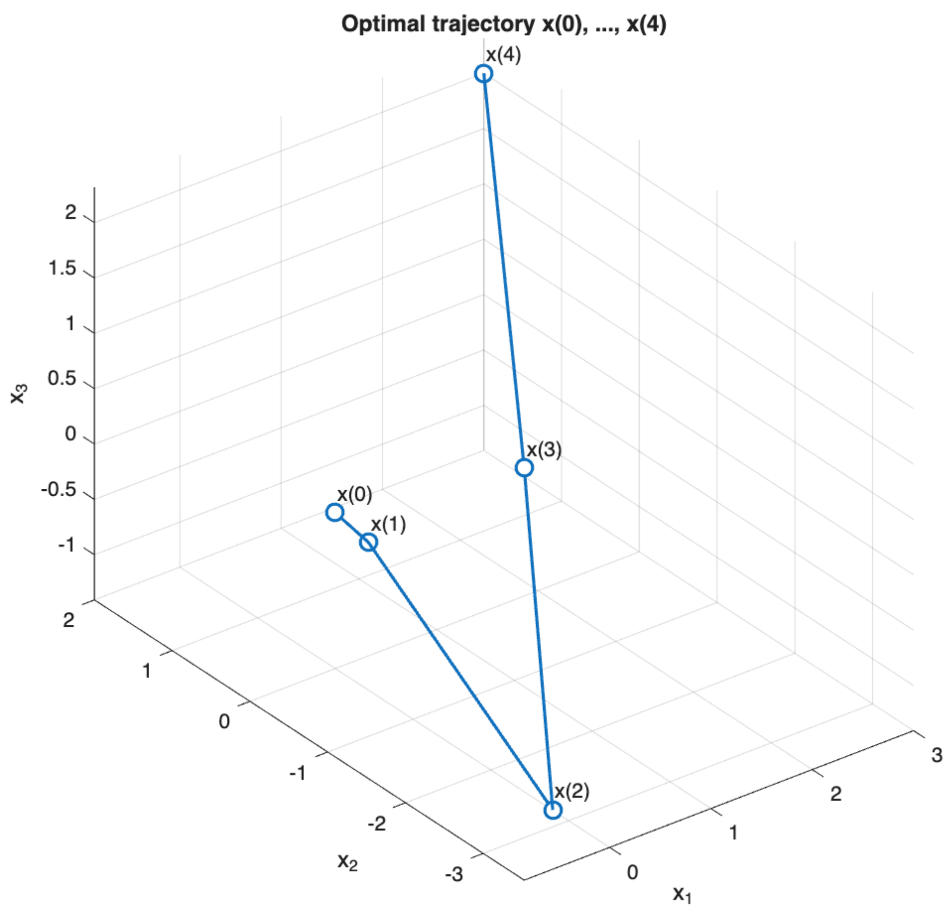
$$x_1 + x_2 = 0$$

where x_1, x_2, x_3 are the components of x_d. Therefore, there exists a sequence of inputs to reach x_d at time 4 if and only if $x_d \in \{x \in \mathbb{R}^3 \mid x_1 + x_2 = 0\}$.

v) Here is the optimization problem for minimizing total energy with $u(0), ..., u(3) \in \mathbb{R}^3$ as arguments

$$\begin{aligned} \min_{\boldsymbol{u}} \quad & u(0)^2 + u(1)^2 + u(2)^2 + u(3)^2, \\ \text{s.t.} \quad & x_d - H\boldsymbol{u} = 0 \end{aligned}$$

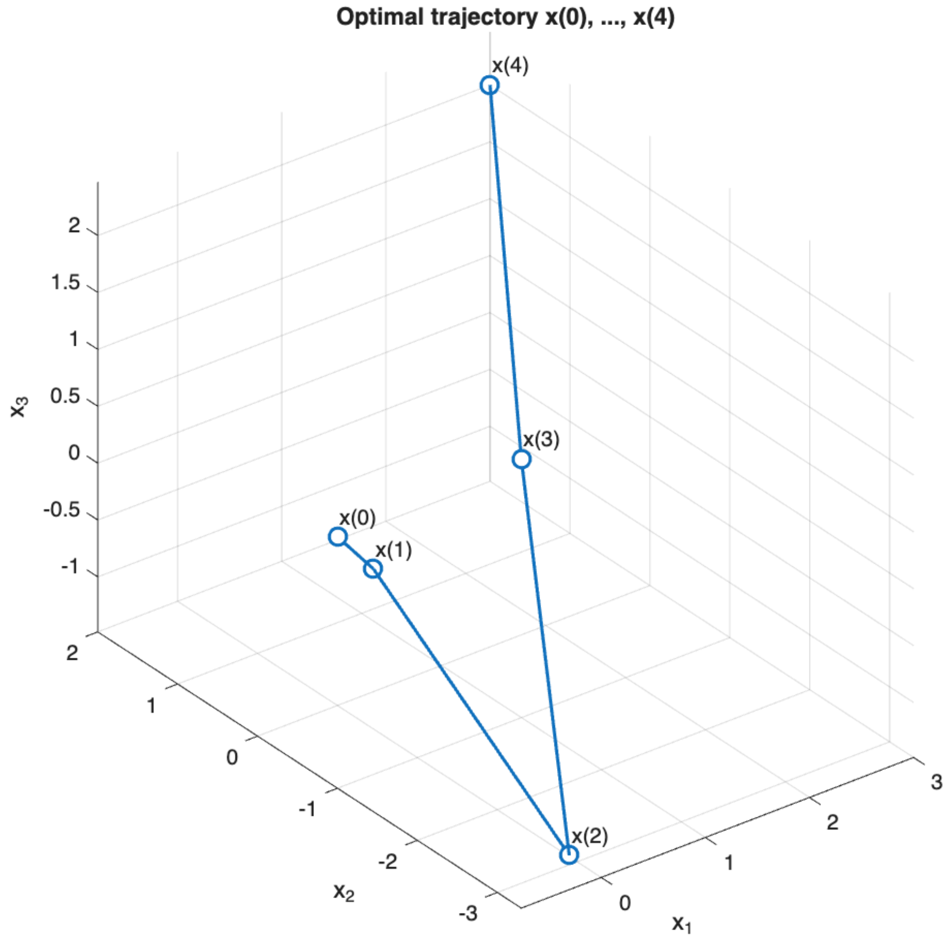
where $\boldsymbol{u} = [u(0)\ u(1)\ u(2)\ u(3)]^\top$ and $x_d = [3\ 2\ 2]^\top$. I wrote a script in `problem1v.m` to solve the problem numerically in CVX. Below is a plot of the optimal trajectory.



vi) Here is the optimization problem for minimizing total energy with $u(0), ..., u(3) \in \mathbb{R}^3$ as arguments

$$\begin{aligned} \min_{\boldsymbol{u}} \quad & u(0)^2 + u(1)^2 + u(2)^2 + u(3)^2, \\ \text{s.t.} \quad & x_d - H\boldsymbol{u} = 0, \\ & -3.3 \leq x_i(k) \leq 3.2, \quad i = 1,2,3, \quad k = 0,1,2,3,4 \end{aligned}$$

where $\boldsymbol{u} = [u(0)\ u(1)\ u(2)\ u(3)]^\top$ and $x_d = [3\ 2\ 2]^\top$. I wrote a script in `problem1vi.m` to solve the problem numerically in CVX. Below is a plot of the optimal trajectory.



Problem 2

(Coding) Consider the matrix A and vector x^* defined as

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}, \qquad x^* = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{7}$$

Define $b = Ax^* + v$ where $v \in \mathbb{R}^6$ is some measurement noise. Assume that the user has no access to x^* and aims to learn x^* from the measurement vector b. We consider two different estimators to learn x^*:

$$l_1 \text{ estimator:} \qquad \min_x \|Ax - b\|_1, \tag{8a}$$

$$l_2 \text{ estimator:} \qquad \min_x \|Ax - b\|_2 \tag{8b}$$

Given a solution $\hat{x}$ obtained from any of the above estimators, we define the estimation error $e = \|\hat{x} - x^*\|_2$ (note that the error is always computed with respect to the l_2-norm no matter which estimator is used for obtaining $\hat{x}^*$). Assume that the noise v is in the form

$$v = \begin{bmatrix} t_1 \\ 0 \\ 0 \\ 0 \\ t_2 \\ 0 \end{bmatrix} \tag{9}$$

where t_1 and t_2 are constants that belong to the discrete set $\{-2, -1.9, -1.8, ..., -0.1, 0, 0.1, ..., 1.8, 1.9, 2\}$ (the increment is 0.1).

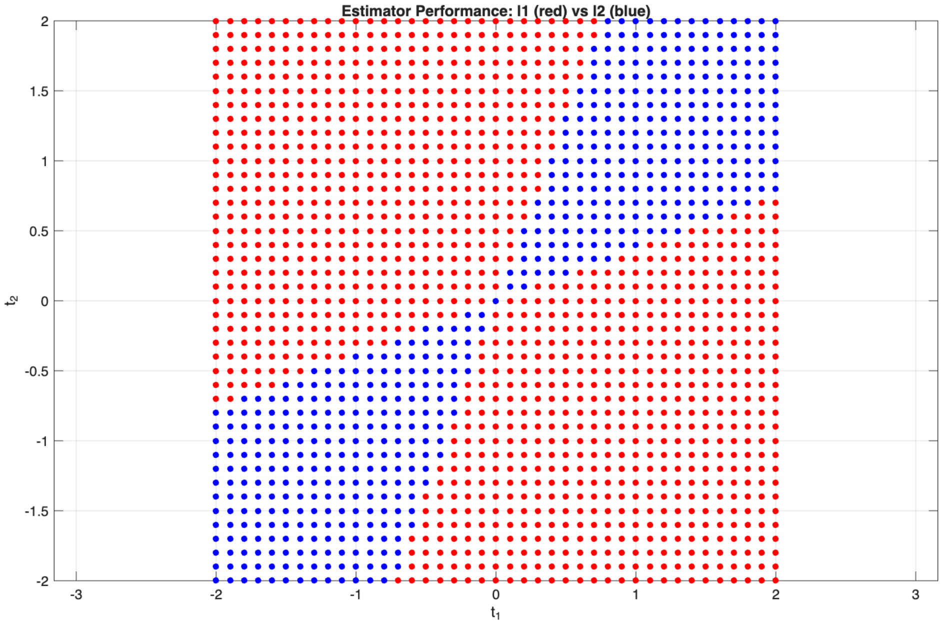
i) For each possible value of the pair (t_1, t_2), solve the l_1 and l_2 estimators in CVX and record the corresponding estimation errors (note: there are 41×41 possibilities for (t_1, t_2)).

ii) Draw a grid in $\mathbb{R}^2$ obtained as follows: For each possible value of the pair (t_1, t_2), we put a symbol in the location (t_1, t_2) in $\mathbb{R}^2$, where the symbol is a small red circle if the l_1 estimator gives the lowest estimation error and is a small blue circle if the l_2 estimator gives the lowest estimation error (note: if the estimation errors for both estimators are the same, use the blue circle). Analyze the plot and report your observations.

Solution

I wrote a script in `problem2.m` to solve the problem numerically in CVX. The script iterates through each possible pair (t_1, t_2), solves both the l_1 and l_2 estimators, and records the estimation errors. For each pair (t_1, t_2), the script then compares the estimation errors and records which estimator gives the lower error. Finally, the script generates a plot in $\mathbb{R}^2$ indicating which estimator performs better for each possible pair (t_1, t_2), where a small red circle indicates that the l_1 estimator gives the lowest estimation error and a small blue circle indicates that the l_2 estimator gives the lowest or the same estimation error as that of the l_1 estimator.

Problem 2 continued on next page. . .



Above is the plot generated by the script. There is a clear pattern in the performance of the two estimators. The l_1 estimator (red circles) performs better in the upper-left quadrant where t_1 is negative and t_2 is positive, and in the lower-right quadrant where t_1 is positive and t_2 is negative. The l_2 estimator (blue circles) dominates in the remaining regions, particularly in the lower-left and upper-right quadrants where t_1 and t_2 have the same sign, as well as in a thin band in the center, around the origin, connecting the two larger regions.

Problem 3

Consider the optimization problem

$$\min_{x \in \mathbb{R}^2} \quad e^{x_1+x_2} + 2x_1^2 + 2x_2^2 - x_1 x_2 - \sin(x_1 + x_2) \tag{10}$$

By analyzing the gradient and Hessian of the objective function, prove that $x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a global minimum of the optimization problem (Hint: A matrix $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is positive definite if and only if $a > 0$ and $ac - b^2 > 0$).

Proof

To prove that x^* is a global minimum, the following first-order and second-order conditions must be true:

$$\nabla f(x^*) = 0 \quad \text{and} \quad H_f(x^*) \succ 0$$

For the first condition, the gradient of the objective function is:

$$\nabla f(x) = \begin{bmatrix} e^{x_1+x_2} + 4x_1 - x_2 - \cos(x_1 + x_2) \\ e^{x_1+x_2} + 4x_2 - x_1 - \cos(x_1 + x_2) \end{bmatrix}$$

Evaluating at $x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$:

$$\nabla f(x^*) = \begin{bmatrix} e^0 + 0 - 0 - \cos(0) \\ e^0 + 0 - 0 - \cos(0) \end{bmatrix} = \begin{bmatrix} 1 - 1 \\ 1 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Therefore, x^* could be a local minimum. For the second-order condition, the Hessian is:

$$H_f(x) = \begin{bmatrix} e^{x_1+x_2} + 4 + \sin(x_1 + x_2) & e^{x_1+x_2} - 1 + \sin(x_1 + x_2) \\ e^{x_1+x_2} - 1 + \sin(x_1 + x_2) & e^{x_1+x_2} + 4 + \sin(x_1 + x_2) \end{bmatrix}$$

Evaluating at $x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$:

$$H_f(x^*) = \begin{bmatrix} e^0 + 4 + \sin(0) & e^0 - 1 + \sin(0) \\ e^0 - 1 + \sin(0) & e^0 + 4 + \sin(0) \end{bmatrix} = \begin{bmatrix} 1 + 4 + 0 & 1 - 1 + 0 \\ 1 - 1 + 0 & 1 + 4 + 0 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$$

$H_f(x^*)$ is positive definite if and only if $a > 0$ and $ac - b^2 > 0$. Here, $a = 5$, $b = 0$, and $c = 5$:

$$a = 5 > 0$$

$$ac - b^2 = 5 \cdot 5 - 0^2 = 25 > 0$$

Therefore, $H_f(x^*)$ is positive definite and the second-order condition is true as well. So $x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a global minimum.