CONTROL & ESTIMATION THEORY

SHIHAOCHEN

1. Problem 1

Consider the discrete-time linear dynamical system $x_{t+1} = \gamma A x_t + B u_t$, where $\gamma \neq 0$. Suppose that the pair (A, B) is controllable.Show that (γA, B) is controllable.Be careful: this exercise says "controllable," not "reachable." Do not use the controllability matrix.

Solotion:

The controllability matrix is $C_n = [A^{n-1}B \;\; A^{n-2}B \;\; \cdots \;\; B]$.

For the pair (A,B) is controllable we can steer its state to the origin from any initial state x in n steps: $\phi(n; x; u_n) = 0$,must be solvable for u_n for any $x \in \mathbb{R}^n$.

Which means

$$A^n x + C_n u_n = 0 \Leftrightarrow C_n u_n = -A^n x \tag{1}$$

This is true iff $rangeC_n = rangeA^n$.Thus $rangeC_n = rangeA^n = range\gamma A^n$

So for the pair (γA, B) is controllable due to

$$C_n u_n = -\gamma A^n x \Leftrightarrow \gamma A^n x + C_n u_n = 0 \tag{2}$$

2. Problem 2

2.1. In Handout 1, Section 1.1.6, we stated that "Lyapunov stability does not imply attractivity" and that "Attractivity does not imply Lyapunov stability." Demonstrate this with a couple of examples.

Solotion:

▷ First example for *Lyapunov stability* $\not\Rightarrow$ *Attractivity*:

$$x_{t+1} = x_t \tag{3}$$

This system is Lyapunov stable since Equation(3) over a positive invariant set X is a constant function for every $\epsilon > 0$ there is a $\delta > 0$.Such that

$$x_0 \in X \cap \mathcal{B}_\delta \Longrightarrow \phi(t; x_0) \in \mathcal{B}_\epsilon, \forall t \in \mathbb{N}. \tag{4}$$

But it is not globally attractive since as $t \to \infty, \phi(t; x_0) \nrightarrow 0$.

▷ Second example for *Attractivity* $\not\Rightarrow$ *Lyapunov stability*:

$$x_{t+1} = \begin{cases} -2x_t, & if\ |x_t| < 1 \\ 0, & otherwise \end{cases} \tag{5}$$

(Shihao Chen) Email addresses: schen45@qub.ac.uk

Thanks Dr. Pantelis Sopasakis. Version 0.0.1. Last updated: August 4, 2022.

For all $x_0 \in$ (-1,1) with $x_0 \neq 0$,the trajectory x_t will exist (-1,1) in finite time and $x_t \to 0$ as $t \to \infty$.Thus this system (5) is globally attractive.

But the origin is not stable since the trajectories will not stay close to zero.

2.2. Consider a dynamical system $x_{t+1} = f(x_t)$ with $f(0) = 0$ and a positive invariant set X $\subseteq \mathbb{R}^n$. Is it possible that the origin is GAS over X, but not LAS? Justify your answer.

Solotion:

For the system (6) over a positive invariant set $X = [0, \infty]$,the origin is GAS over X since it is both Lyapunov stable,globally attractive and x_t approaches 0 much closely and converges wherever it starts from on the right of set X,for the left of set X it does not converge.

$$x_{t+1} = \begin{cases} 0.5x_t, & if\ x_t \geq 0 \\ 2x_t, & if\ x_t < 0 \end{cases} \tag{6}$$

But it is not locally sttractive over X since it is not for all $x_0 \in X \cap \mathcal{B}_\eta$ there is $\eta > 0$ such that $\phi(t; x_0) \to 0$ as $t \to \infty$.

Thus it is not LAS.

3. Problem 3

(This is from Handout 1) Find a GES linear system, $x_{t+1} = Ax_t$,with $x_t \in \mathbb{R}^2$ such that the Euclidean unit ball,$\mathcal{B} = \{x \in \mathbb{R}^2 : \|x\| \leq 1\}$, is not an invariant set.

Solotion: The Python code: Please Click Here.

The unit ball is not invariant for the system (7) starts from $x_0 \in \mathcal{B}$ since there is a point (x_2) in the unit ball such that $\begin{bmatrix} 0.2 & -0.5 \\ 0.4 & 1.5 \end{bmatrix} x_2$ (x_3) is not in the unit ball as shown in Fig. 1.

$$x_{t+1} = \begin{bmatrix} 0.2 & -0.5 \\ 0.4 & 1.5 \end{bmatrix} x_t \tag{7}$$

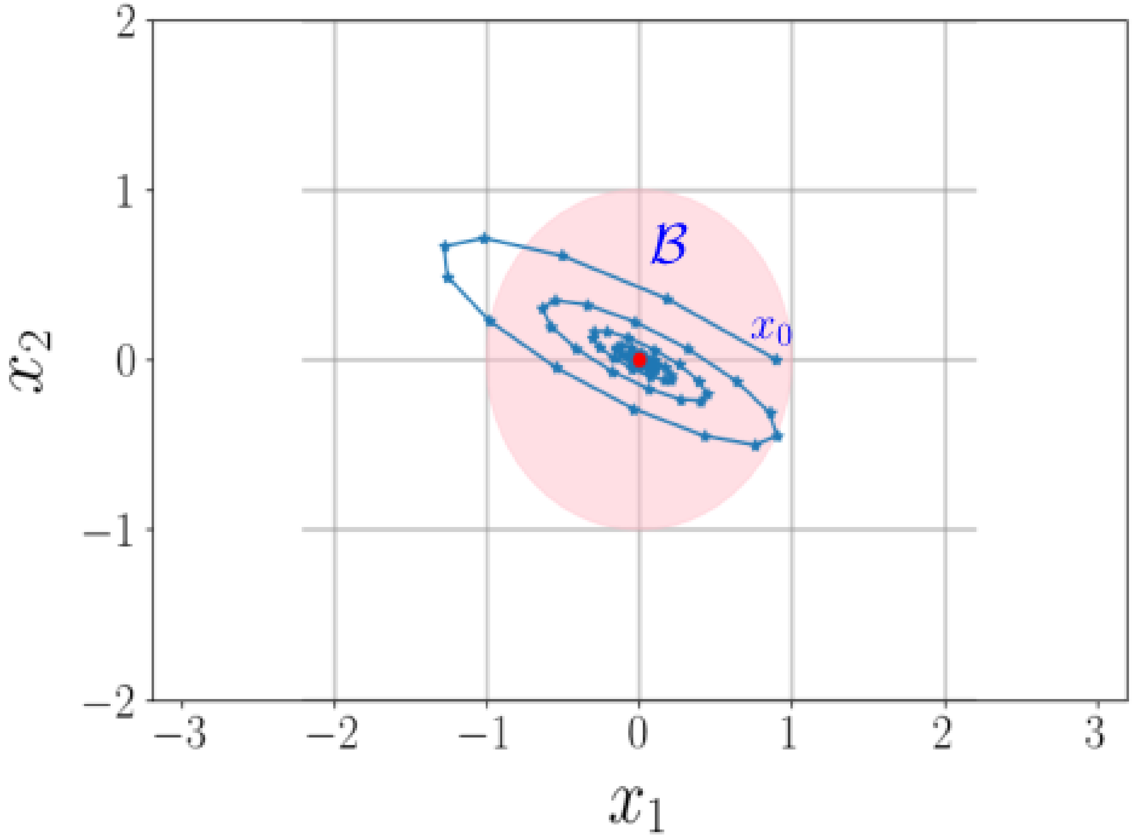


FIGURE 1. Not an invariant set

4. PROBLEM 4

4.1. Check whether the origin is globally exponentially stable (GES) for the discrete-time linear dynamical system

$$x_{t+1} = \begin{bmatrix} 0.9 & 0.1 & 0 \\ -0.1 & 0.9 & 0 \\ 0 & 0 & -0.3 \end{bmatrix} x_t \tag{8}$$

using **Theorem 1.1:(Stability of Linear Systems)**:the origin is GES for the linear system $x_{t+1} = Ax_t$ over ($\mathbb{R}^n$) iff all the eigenvalues of A are within the unit circle (otherwise it is not attractive).

Solotion:

$A = \begin{bmatrix} 0.9 & 0.1 & 0 \\ -0.1 & 0.9 & 0 \\ 0 & 0 & -0.3 \end{bmatrix}$

Thus $\lambda_1 = 0.9 + 0.1i, \lambda_2 = 0.9 - 0.1i, \lambda_3 = -0.3$.

Their modulus for $\lambda_1, \lambda_2, \lambda_3$ are $0.9055, 0.9055, 0.3$ respectively and they are all less than 1 so that this system 8 is GES.

4.2. If it is GES, determine a Lyapunov function.

Solotion:

The linear system $x_{t+1} = Ax_t$ is GES iff given $Q \in \mathbb{S}^n_{++}$ there exists a $P \in \mathbb{S}^n_{++}$ that satisfies the Lyapunov equation

$$A^\intercal PA - P = -Q \tag{9}$$

The P that solves this equation is unique in $\mathbb{S}^n_{++}$ and $V(x) = x^\intercal Px$ satisfies the conditions of Lyapunov's theorem.

For this linear system 8.

Using $Q = I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ that the following matrix is the unique solution of the Lyapunov equation: $P = \begin{bmatrix} 5.55556 & 0 & 0 \\ 0 & 5.55556 & 0 \\ 0 & 0 & 1.0989 \end{bmatrix}$

So $V(x) = x^\intercal Px = x^\intercal \begin{bmatrix} 5.55556 & 0 & 0 \\ 0 & 5.55556 & 0 \\ 0 & 0 & 1.0989 \end{bmatrix} x.$

4.3. Determine an invariant set for this system that is contained in the Euclidean unit ball, $\mathcal{B} = \{x \in \mathbb{R}^3 : \|x\| \leq 1\}$.

Solotion:

For origin of the linear system 8 is GES,the Lyapunov function 9 satisfies the following bounds:

$$\lambda_{min}(P)\|x\|^2 \leq V(x) \leq \lambda_{max}(P)\|x\|^2 \tag{10}$$

which is

$$0.3\|x\|^2 \leq V(x) \leq 0.9055\|x\|^2 \tag{11}$$

Since $V(x)$ satisfies LDC.For $V : X \to [0, \infty]$.Then $Y = x \in X : V(x) \leq 0.9055$ is a invariant for this system 8.

5. Problem 5

Consider the following discrete-time linear dynamical system

$$\begin{bmatrix} a_{t+1} \\ b_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{1}{5}a_t + cos(b_t) - 1 \\ \frac{1}{2}e^{a_t}b_t \end{bmatrix} \tag{12}$$

The state of the system is the vector $x_t = [a_t\ b_t]^\intercal \in \mathbb{R}^2$.

5.1. Show that the origin is an equilibrium point of this dynamical system.

Solotion:

With $x_0 = [0\ 0]^\intercal$ that $\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = x_1 = \begin{bmatrix} \frac{1}{5}0 + cos(0) - 1 \\ \frac{1}{2}e^0 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Thus the origin is an equilibrium point of this dynamical system since $[0\ 0]^\intercal = f([0\ 0]^\intercal)$.

5.2. Use Theorem 2.6 (in Handout 2) to tell whether the system is locally exponentially stable.

Solotion:

The system state is $x_t = (a_t, b_t)$.The Jacobian is

$$Jf(x) = \begin{bmatrix} \frac{1}{5} & -\sin(b) \\ \frac{1}{2}e^a b & \frac{1}{2}e^a \end{bmatrix} \tag{13}$$

So

$$A = Jf(0) = \begin{bmatrix} \frac{1}{5} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \tag{14}$$

The eigenvalues of A are $\frac{1}{5}$ and $\frac{1}{2}$,so $\rho(A) = \frac{1}{2} < 1$, therefore, the origin is LES for the system 12.

5.3. Use Python to simulate this system starting from various initial states close to the origin. Include a plot of the system trajectories in your report.

Solotion: The Python code: Please Click Here.

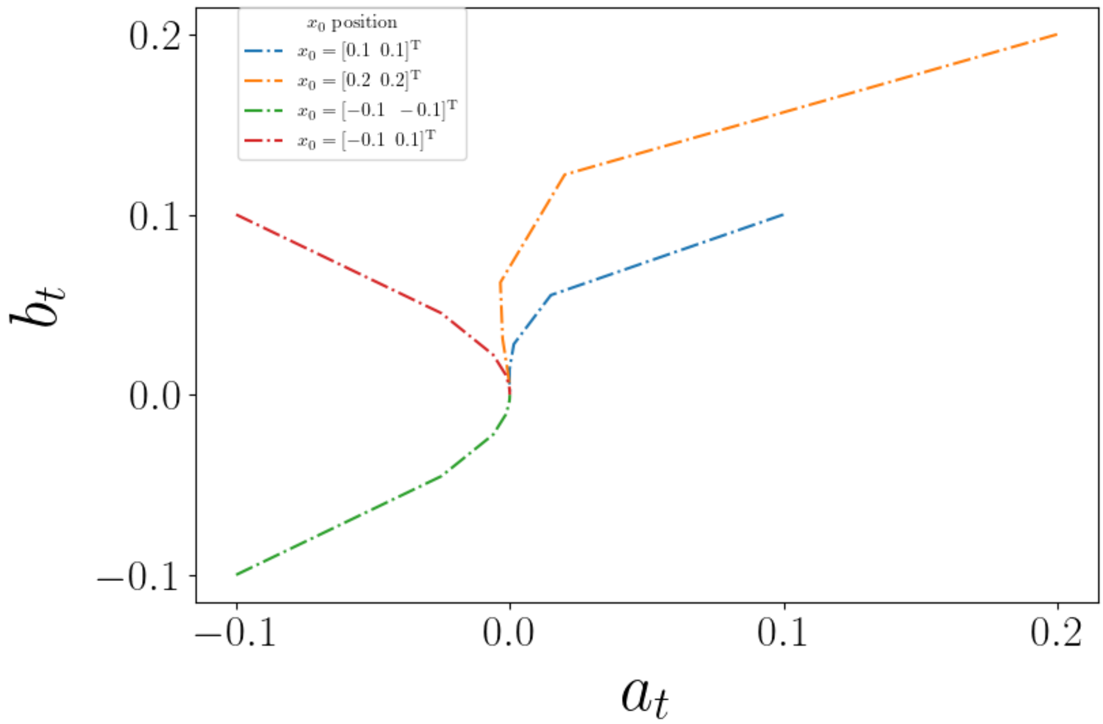


FIGURE 2. Simulation from various initial states

6. PROBLEM 6

6.1. Prove that the function $f : (0, \infty) \to \mathbb{R}$ with $f(x) = x^2 + x - 2022 \log x$ is convex.

Solotion:

The gradient for this function is $\nabla f(x) = 2x + 1 - 2022\frac{1}{x \ln 10}$ and its Hessian is $\nabla^2 f(x) = 2 + 2022\frac{\ln 10 + \frac{x}{10}}{x^2 \ln^2 10}$.

Since $x \in (0, \infty)$ that $\ln 10 + \frac{x}{10}$ and $x^2 \ln^2 10$ are both greater than 0,thus $\nabla^2 f(x) > 0$,so this function is convex.

6.2. Use Python to solve the minimisation problem

$$\underset{1 \leq x \leq 100}{\text{minimise}}\, x^2 + x - 2022 \log x, \tag{15}$$

that is, determine a minimiser and the optimal value of this problem. Include a link to your code and report your results in your report.

Solotion: The Python code: Please Click Here.

The results are shown below as Fig.3.

```
C:\Users\ChenShihao\AppData\Local\Programs\Python\Python39\python.exe C:
Optimal value: -2211.702100393303
Minimiser: [20.70562308]
Whether the iteration terminates successfully: True
Reason for iteration termination: Optimization terminated successfully

Process finished with exit code 0
```

FIGURE 3. Simulation results

7. PROBLEM 7

Give examples of

7.1. A convex function $f : \mathbb{R} \to [0, \infty)$ which has a finite infimum but no minimisers.

Solotion: $f(x) = \exp(x)$ whose infimum is 0, but has no minimisers.

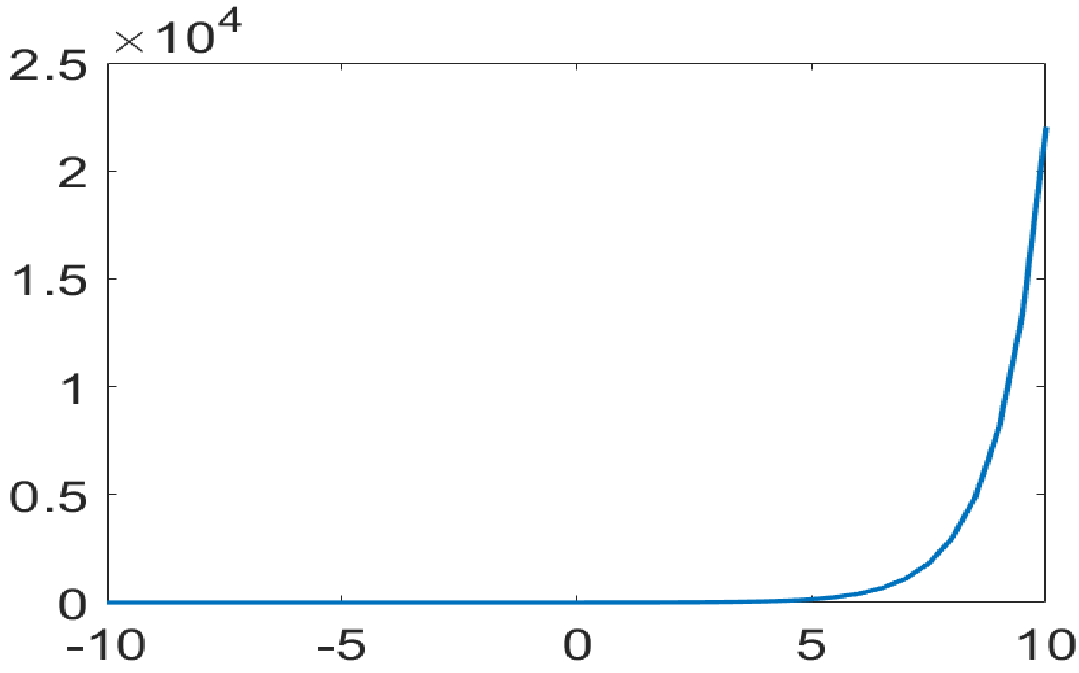


FIGURE 4. $f(x) = \exp(x)$ Simulation

7.2. A convex function $f : \mathbb{R} \to \mathbb{R}$ that has infinitely many minimisers.
Solotion:

$$f(x) = \begin{cases} -x & x < -2 \\ 0 & -2 \leq x \leq 2 \\ x & x > 2 \end{cases} \tag{16}$$

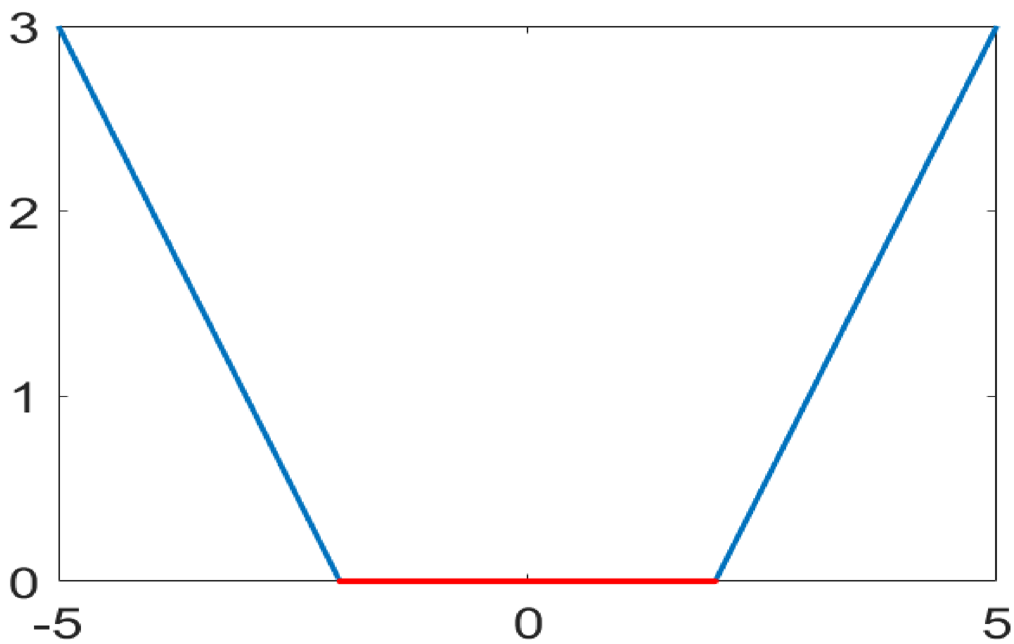


FIGURE 5. $f(x)$(16) Simulation

7.3. A differentiable function $f : \mathbb{R} \to \mathbb{R}$ that has a single stationary point that is not a minimiser.

Solotion: $f(x) = x^3$ whose single stationary point is (0,0).

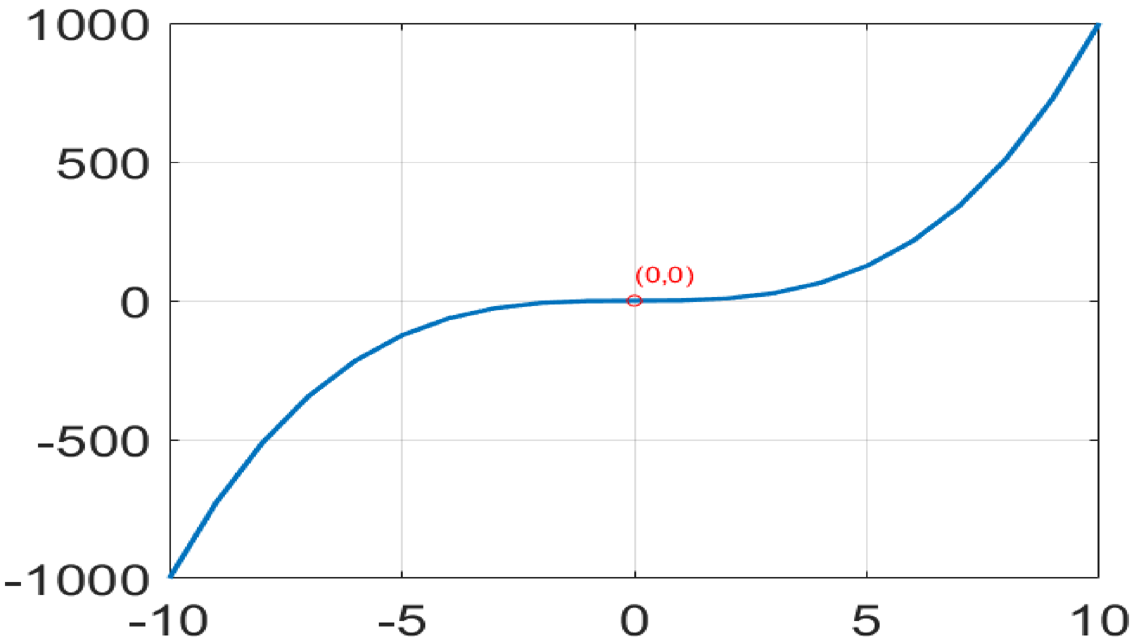


FIGURE 6. $f(x) = x^3$ Simulation

7.4. A differentiable function $f : \mathbb{R} \to \mathbb{R}$ that has infinitely many stationary points none of which is a minimiser.

Solotion: $f(x) = x - \sin x$ whose infinitely many stationary points are in set $\{(x, y)|x = y = 2k\pi, k \in \mathbb{Z}\}$.

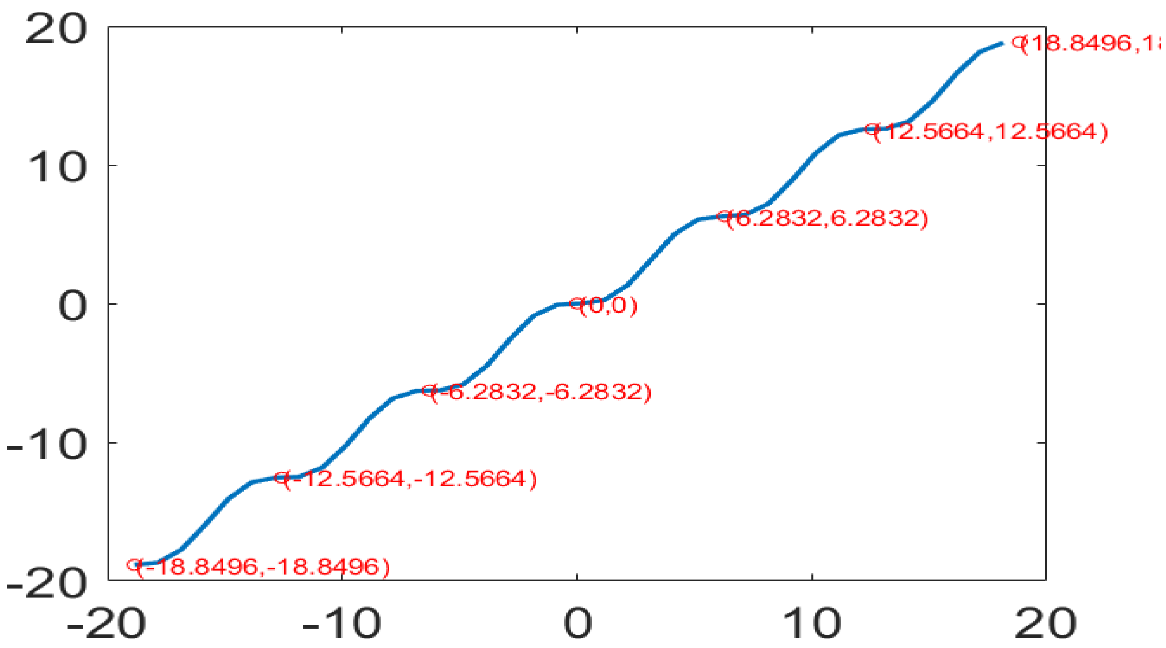


FIGURE 7. $f(x) = x - \sin x$ Simulation

7.5. A twice differentiable function that is not constant and has infinitely many minimisers.

Solotion: $f(x) = \sin(x)$ whose minimisers are in set $\{x|x = -\frac{\pi}{2}+2k\pi, k \in \mathbb{Z}\}$.

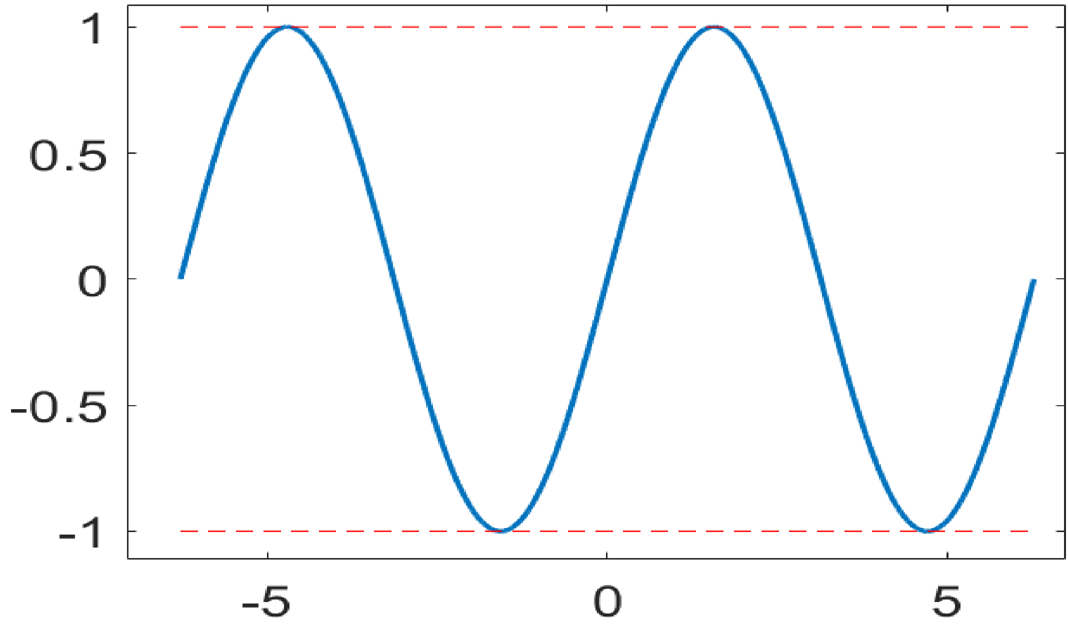


FIGURE 8. $f(x) = \sin(x)$ Simulation

8. PROBLEM 8

8.1. Use the KKT theorem to solve the following problem

$$\operatorname*{minimise}_{x\in\mathbb{R}^{100}} 2\|x\|^2 + 3\sum_{i=1}^{50} x_i + 2022, \tag{17a}$$

$$\text{subject to:} \sum_{i=1}^{100}(-1)^i x_i = 100. \tag{17b}$$

Provide a detailed derivation in your report.

Solotion:

The cost function is

$$f(x) = 2\|x\|^2 + 3\sum_{i=1}^{50} x_i + 2022$$

and the constraints are

$$g(x) = \sum_{i=1}^{100}(-1)^i x_i - 100$$

The Lagrangian $\mathscr{L}$ is

$$\begin{aligned}\mathscr{L}(x,\lambda) &= f(x)+\lambda g(x)\\ &= 2\|x\|^2+3\sum_{i=1}^{50}x_i+2022+\lambda(\sum_{i=1}^{100}(-1)^i x_i-100)\end{aligned}$$

the gradient of $\mathscr{L}$ given by

$$\begin{aligned}\nabla\mathscr{L}(x,\lambda) &= \begin{bmatrix}\nabla_x\mathscr{L}(x,\lambda)\\ \nabla_\lambda\mathscr{L}(x,\lambda)\end{bmatrix}\\ &= \begin{bmatrix}4x+b-c\lambda\\ c^\intercal x-100\end{bmatrix}\end{aligned}$$

where

$$b=[\underbrace{3\;3\;3\cdots3}_{50}\overbrace{0\cdots0\;0\;0}^{50}]^\intercal$$

$$c=[\underbrace{-1\;1\;-1\cdots1\;-1\;1}_{100}]^\intercal$$

So that

$$x^\star=\frac{c\lambda-b}{4} \tag{18}$$

and

$$c^\intercal x^\star-100=0$$

From 18

$$\frac{c^\intercal(c\lambda-b)}{4}-100=0$$

Thus

$$\lambda=\frac{400+c^\intercal b}{\|c\|^2}$$

Back to 18

$$\begin{aligned}x^\star &= \frac{c\frac{400+c^\intercal b}{\|c\|^2}-b}{4}\\ &= \frac{c(400+c^\intercal b)}{4\|c\|^2}-\frac{b}{4}\end{aligned}$$

8.2. Solve the same problem using Python (e.g., CVXPy) and check whether the solution obtained using Python is the same as the solution you obtained in Question 8.1. In your report you should provide a link to your code and report your results.

Solotion: The Python code: Please Click Here.

```
C:\Users\ChenShihao\AppData\Local\Programs\Python\Python39\python.exe C:/
x_star_cp:
[-1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25
 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25
 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25
 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25
 -1.75  0.25 -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.
 -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.
 -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.
 -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.
 -1.    1.   -1.    1.  ]
f_star_cp:
2165.7500000000005
x_star:
[-1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25
 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25
 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25
 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25 -1.75  0.25
 -1.75  0.25 -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.
 -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.
 -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.
 -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.   -1.    1.
 -1.    1.   -1.    1.  ]
f_star:
2165.75
```

FIGURE 9. Simulation results

9. PROBLEM 10

9.1. Give examples of discrete-time linear dynamical systems, $x_{t+1} = Ax_t + Bu_t$, with $x_t \in \mathbb{R}^2$ and $u_t \in \mathbb{R}$, which are

1.Controllable, but not reachable

Solotion:

$x_{t+1} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_t$

2.Stabilisable, but not controllable

Solotion:

$$x_{t+1} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} x_t + 0 \cdot u_t$$

9.2. Show that the following system is reachable

$$x_{t+1} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix} x_t + \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 3 \end{bmatrix} u_t, \tag{19}$$

and determine a (finite) sequence of control actions that drives the state from the initial state x = (0, 1, -2) to the target state x' = (2, 1, 3).

Solotion:

The controllability matrix,$C_3 = [A^2B \;\; AB \;\; B]$,of this system 19 is

$$C_3 = \begin{bmatrix} 6 & 10 & 5 & 10 & 1 & 1 \\ 4 & -14 & 2 & -6 & 2 & 0 \\ 2 & 20 & -1 & 4 & 0 & 3 \end{bmatrix}$$

and the rank of C_3 is 3.So this system 19 is reachable.

We need to solve the equation $C_3 \cdot u = x' - A^3 x$,which is,

$$\begin{bmatrix} 6 & 10 & 5 & 10 & 1 & 1 \\ 4 & -14 & 2 & -6 & 2 & 0 \\ 2 & 20 & -1 & 4 & 0 & 3 \end{bmatrix} \cdot u = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 6 & 7 & 12 \\ -6 & 3 & -16 \\ 8 & -2 & 12 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}$$

$$\begin{bmatrix} 6 & 10 & 5 & 10 & 1 & 1 \\ 4 & -14 & 2 & -6 & 2 & 0 \\ 2 & 20 & -1 & 4 & 0 & 3 \end{bmatrix} \cdot u = \begin{bmatrix} 19 \\ -34 \\ 29 \end{bmatrix}$$

So,$u = \begin{bmatrix} -1.22185226 \\ 1.32254203 \\ -0.21643201 \\ 1.51523816 \\ -0.53635475 \\ -0.42817361 \end{bmatrix}$, thus,we need to apply the control actions $u_0 = \begin{bmatrix} -1.22185226 \\ 1.32254203 \end{bmatrix}$, $u_1 = \begin{bmatrix} -0.21643201 \\ 1.51523816 \end{bmatrix}$ and $u_2 = \begin{bmatrix} -0.53635475 \\ -0.42817361 \end{bmatrix}$

The Python code: Please Click Here.

10. Problem 11

Consider the Riccati recursion

$$P_{t+1} = Q + A^\intercal P_t A - A^\intercal P_t B(R + B^\intercal P_t B)^{-1} B^\intercal P_t A,$$

where

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 2 \\ -1 \end{bmatrix},$$

$$Q = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$$R = 2,$$

$$P_0 = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix},$$

10.1. Is it true that this Riccati recursion converges? Justify your answer and check all conditions carefully. As a hint, note that Q can be written as $Q = qq^\intercal$ with $q = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Solotion:

The controllability matrix,C_2 of (A, B) is,

$$C_2 = [AB \;\; B] = \begin{bmatrix} 3 & 2 \\ 2 & -1 \end{bmatrix}$$

The rank of C_2 is 2,so (A, B) is controllable.

Since $Q = q^\intercal q$ with $q = \begin{bmatrix} 1 & 1 \end{bmatrix}$.

$$\begin{aligned} \mathcal{O} &= \begin{bmatrix} q \\ qA \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix} \end{aligned}$$

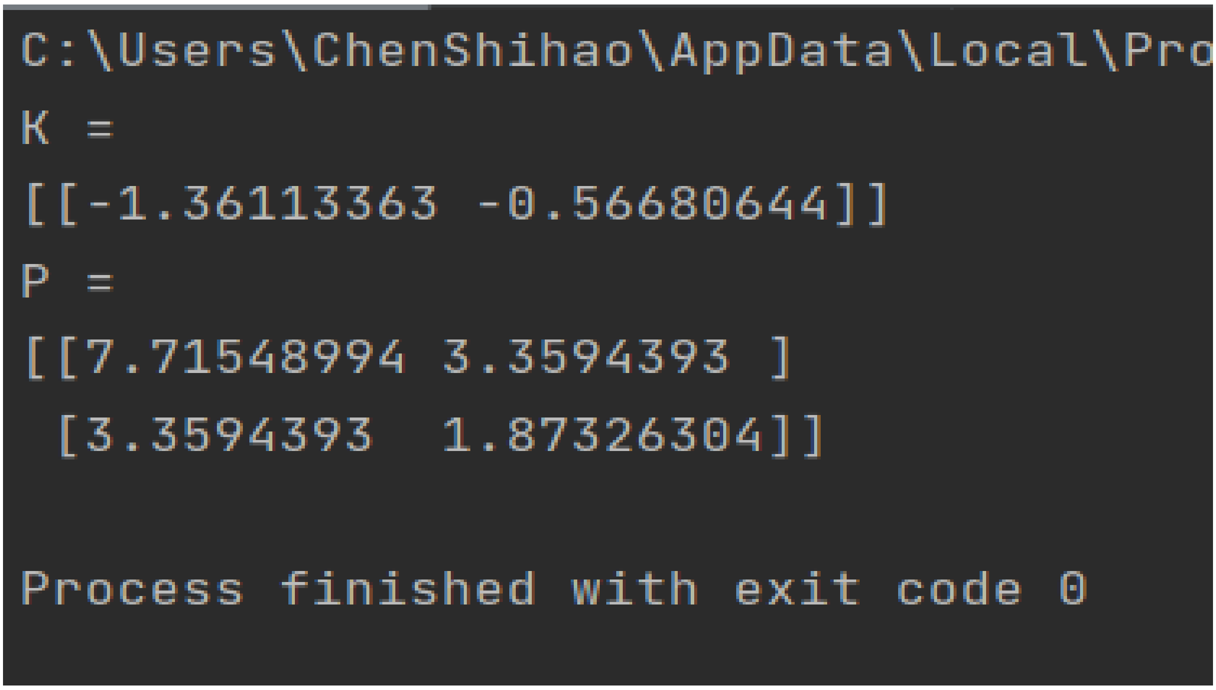
The rank of $\mathcal{O}$ is 2,so (A, q) is observable.

And $Q \in \mathbb{S}_+, R \in \mathbb{S}_{++}, P_0 \in \mathbb{S}_+$

So $P = \lim_{t\to\infty} P_t$ exists which means this Riccati recursion converges.

10.2. Use control.dare in Python to determine the limit $\lim_{t\to\infty} P_t$. Include this limit in your report and provide a link to your code.

Solotion: The Python code: Please Click Here.



```
C:\Users\ChenShihao\AppData\Local\Pro
K =
[[-1.36113363 -0.56680644]]
P =
[[7.71548994 3.3594393 ]
 [3.3594393  1.87326304]]

Process finished with exit code 0
```

FIGURE 10. Simulation results

11. PROBLEM 14

Let X be a continuous random variable with pdf

$$p_X(x) = \begin{cases} \frac{2}{x^2}, \text{for } x \in [1,2), \\ 0, \text{otherwise} \end{cases}$$

11.1. Determine P$[X \geq 1.5]$

Solotion:

$$\begin{aligned} P[X \geq 1.5] &= \int_{1.5}^{\infty} p_X(x)dx \\ &= \int_{1.5}^{2} \frac{2}{x^2}dx + \int_{2}^{\infty} 0dx \\ &= \int_{1.5}^{2} \frac{2}{x^2}dx \\ &= -\frac{2}{x}\Big|_{1.5}^{2} \\ &= \frac{1}{3} \end{aligned}$$

11.2. Determine $\mathbb{E}[X]$ and Var$[X]$

Solotion:

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} x p_X(x) dx \\
&= \int_1^2 \frac{2}{x^2} dx \\
&= -\frac{2}{x}\bigg|_1^2 \\
&= 1
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}[(X - \mathbb{E}[X^2])^2] \\
&= \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 p_X(x) dx \\
&= \int_1^2 (x-1)^2 \frac{2}{x^2} dx \\
&= \int_1^2 \frac{2x^2 - 4x + 2}{x^2} dx \\
&= \int_1^2 2 - \frac{4}{x} + \frac{2}{x^2} dx \\
&= -4 \ln x\bigg|_1^2 + \frac{-2}{x}\bigg|_1^2 \\
&= -4 \ln 2 + 1
\end{aligned}$$

11.3. Determine $\mathbb{E}[X^3]$ using the law of the unconscious statistician
Solotion:

$$\begin{aligned}
\mathbb{E}[X^3] &= \int_{-\infty}^{\infty} x^3 p_X(x) dx \\
&= \int_1^2 x^3 \frac{2}{x^2} dx \\
&= x^2\bigg|_1^2 \\
&= 3
\end{aligned}$$

12. Problem 16

Alice has two coins in her pocket, a fair coin (heads on one side and tails on the other side) and a two-headed coin (heads on both sides). She picks one at random from her pocket, tosses it and obtains heads. What is the probability that she flipped the fair coin?

Solotion:

P[fair coin]=P[two-headed coin]=50%,,P[heads|fair coin]=50%,P[heads|two-headed coin]=1,such that

$$\begin{aligned}\mathrm{P[heads]} &= \mathrm{P[heads|fair\ coin]} \cdot \mathrm{P[fair\ coin]} + \\ &\quad \mathrm{P[heads|two\text{-}headed\ coin]} \cdot \mathrm{P[two\text{-}headed\ coin]} \\ &= 50\% \cdot 50\% + 1 \cdot 50\% \\ &= 75\%\end{aligned}$$

$$\begin{aligned}\mathrm{P[fair\ coin|heads]} &= \frac{\mathrm{P[heads|fair\ coin]} \cdot \mathrm{P[fair\ coin]}}{\mathrm{P[heads]}} \\ &= \frac{50\% \cdot 50\%}{75\%} \\ &= \frac{1}{3}\end{aligned}$$

13. Problem 17

You have a coin that has probability p of landing heads and probability 1 - p of landing tails. You believe that the coin is fair. To that end, you treat p as a random variable with expectation $\mathbb{E}[p] = 0.5$ and variance $\mathrm{Var}[p] = 0.0119$; you may choose any appropriate prior distribution for p that has such an expectation and variance.

You flip the coin 25 times and you obtain the following observations:

Heads, Heads, Heads, Tails, Heads, Heads, Tails, Heads, Heads, Heads, Heads, Tails, Heads, Heads, Tails, Tails, Heads, Heads, Tails, Tails, Heads, Heads, Heads, Heads, Heads.(18H,7T)

13.1. What is the posterior distribution of p given these measurements? Include a plot in your report.

Solotion:

We have,

$$\mathbb{E}[p] = \frac{\alpha}{\alpha + \beta} = 0.5$$

$$\mathrm{Var}[p] = \frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2} = 0.0119$$

Thus,

$$\alpha = \beta = 10$$

The posterior distribution is,

$$
\begin{aligned}
p|x_1, x_2, x_3, \ldots, x_{25} &\sim \text{Beta}(\sum_{i=1}^{10} x_i + \alpha, 25 - \sum_{i=1}^{10} x_i + \beta) \\
&= \text{Beta}(18 + 10, 25 - 18 + 10) \\
&= \text{Beta}(28, 17)
\end{aligned}
$$

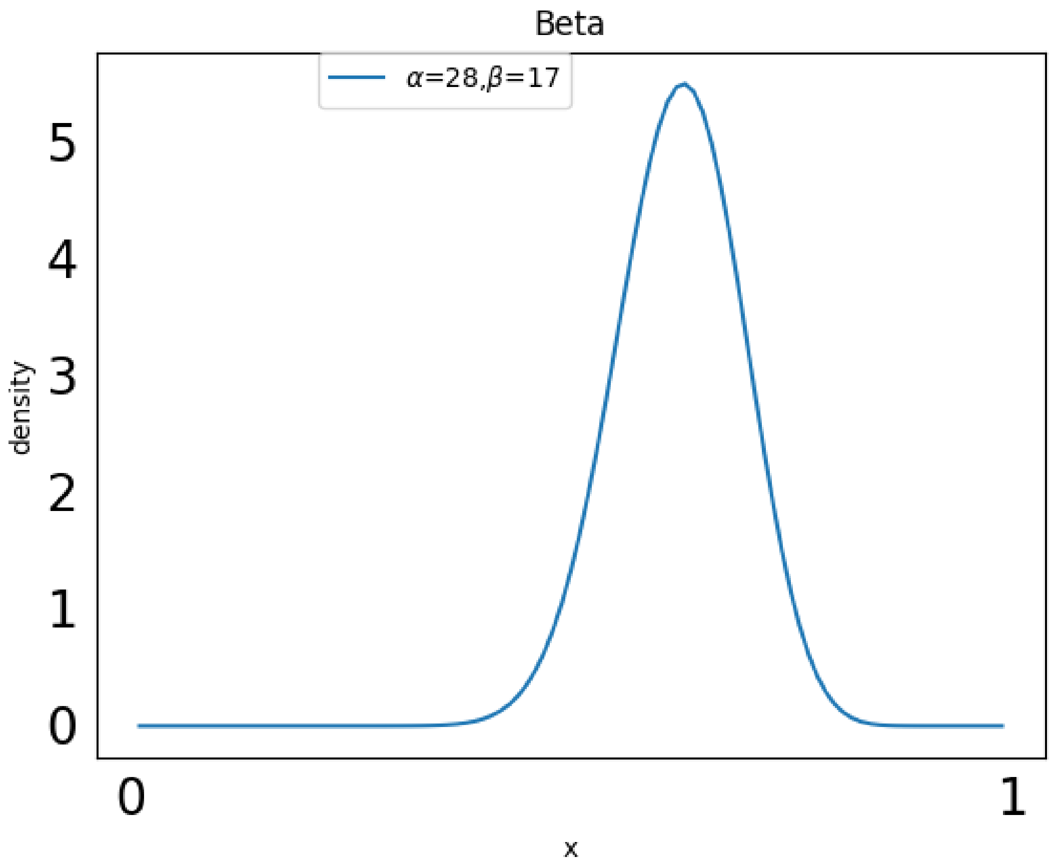


FIGURE 11. Beta Distribution

13.2. What is the maximum a posteriori (MAP) estimate of p given these measurements?

Solotion:

$$\hat{\theta}_{MAP} = \frac{28 - 1}{28 + 17 - 2} = \frac{27}{43}$$

13.3. What is the posterior mean given these measurements?

Solotion:

14. Problem 18

Let N be the number of customers that enter a store in a given day and we know that $\mathbb{E}[N] = 20$ and $\text{Var}[N] = 25$. Let X_i be the amount of money that the i-th customer spends and assume that the X_i are independent random variables and independent of N. Assume that $\mathbb{E}[X_i] = £20$ and $\text{Var}[X_i] = £100$ for all $i = 1,..., N$. The Y be the store's total sales,

$$Y = \sum_{i=1}^{N} X_i.$$

14.1. Determine $\mathbb{E}[Y]$ using the law of total expectation
Solotion:

$$\begin{aligned}
\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|N]] &= \mathbb{E}[\mathbb{E}[\sum_{i=1}^{N} X_i|N]] \\
&= \mathbb{E}[\sum_{i=1}^{N} \mathbb{E}[X_i|N]] \\
&= \mathbb{E}[20N] = 20\mathbb{E}[N] \\
&= 400
\end{aligned}$$

14.2. Determine $\text{Var}[Y]$ using the law of total variance
Solotion:

$$\text{Var}[\mathbb{E}[Y|N]] = \text{Var}[20N] = 400 \cdot 25 = 10000$$

$$\begin{aligned}
\text{Var}[Y|N] &= \text{Var}[[\sum_{i=1}^{N} X_i|N]] \\
&= \sum_{i=1}^{N} \text{Var}[X_i|N] \\
&= 100N
\end{aligned}$$

$$\begin{aligned}
\text{Var}[Y] &= \mathbb{E}[\text{Var}[Y|N]] + \text{Var}[\mathbb{E}[Y|N]] \\
&= \mathbb{E}[100N] + 10000 \\
&= 100 \cdot 20 + 10000 \\
&= 12000
\end{aligned}$$

15. Problem 25

Describe the extended Kalman filter and the moving horizon estimation methodologies (you are ex- pected to provide technical details). What are the merits and limitations of each approach?
Solotion:

*Extended Kalman Filter:

Extended Kalman filtering can be used to solve problems in nonlinear cases by Taylor expansion. Obviously, the tangent at the point with the largest possible value in the original distribution should be used, which is the tangent at the mean, as shown in 12. The result (dashed line) is closer to the mean variance of the actual result.

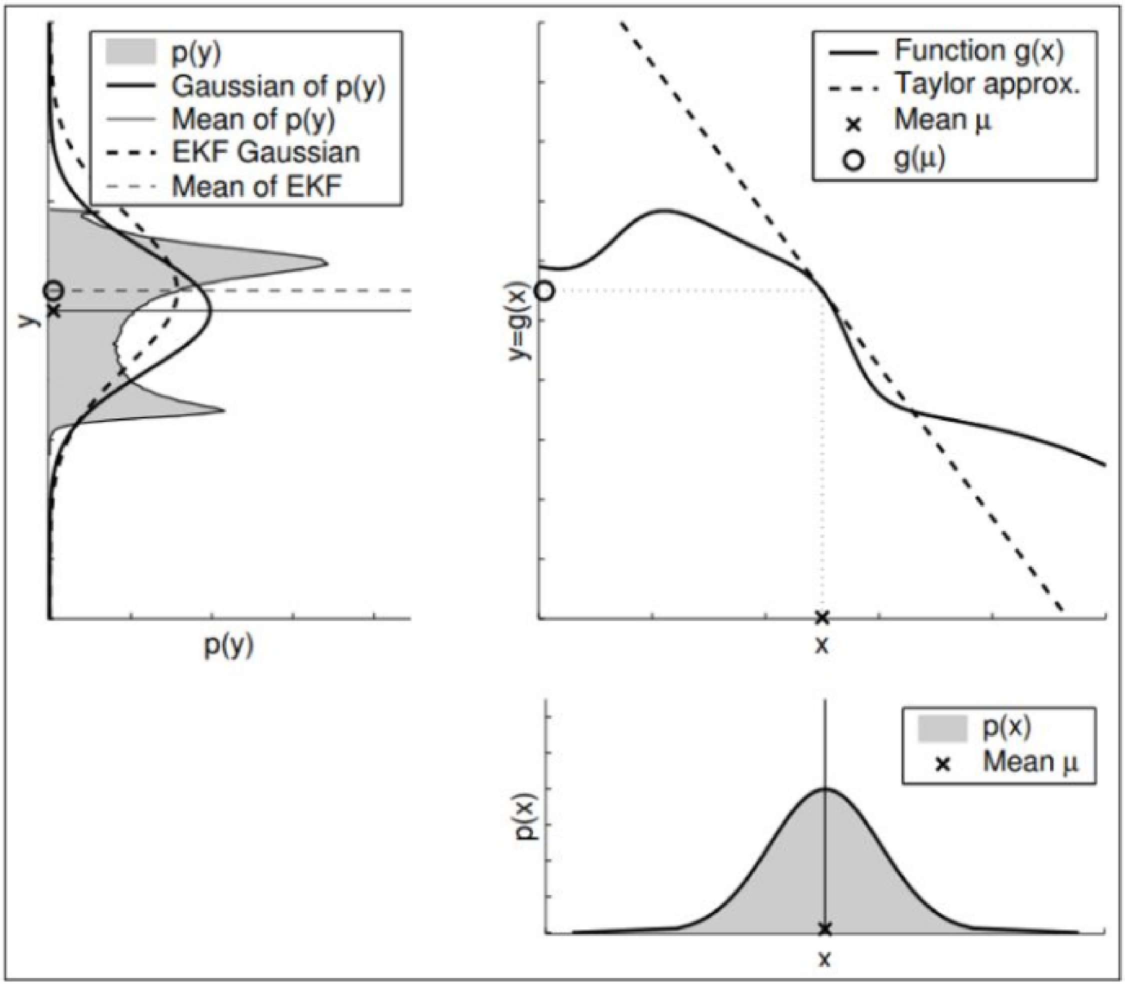


FIGURE 12. Extended Kalman Filter

For EKF,linearize at a time when the nearest state x_{k-1} , the nearest input u_{k-1} , and 0 noise are known,

Equation of motion,

$$x_k = f_{k-1}(x_{k-1}, u_{k-1}, w_{k-1}) \approx f_{k-1}(\hat{x}_{k-1}, u_{k-1}, 0)$$
$$+\frac{\partial f_{k-1}}{\partial x_{k-1}}|_{\hat{x}_{k-1},u_{k-1},0}(x_{k-1} - \hat{x}_{k-1}) + \frac{\partial f_{k-1}}{\partial w_{k-1}}|_{\hat{x}_{k-1},u_{k-1},0} w_{k-1}$$

Observation equation,

$$y_k = h_k(x_k, v_k) \approx h_k(\hat{x}_k, 0) + \frac{\partial h_k}{\partial x_k}|_{\hat{x}_k,0}(x_k - \hat{x}_k) + \frac{\partial h_k}{\partial v_k}|_{\hat{x}_k,0} v_k$$

The four Jocabian matrices are as follows,

$$F_{k-1} = \frac{\partial f_{k-1}}{\partial x_{k-1}}|_{\hat{x}_{k-1},u_{k-1},0}$$
$$L_{k-1} = \frac{\partial f_{k-1}}{\partial w_{k-1}}|_{\hat{x}_{k-1},u_{k-1},0}$$
$$H_k = \frac{\partial h_k}{\partial x_k}|_{\hat{x}_k,0}$$
$$M_k = \frac{\partial h_k}{\partial v_k}|_{\hat{x}_k,0}$$

Linearized motion model,

$$x_k = f_{k-1}(\hat{x}_{k-1}, u_{k-1}, 0) + F_{k-1}(x_{k-1} - \hat{x}_{k-1}) + L_{k-1}w_{k-1}$$

Linearized measurement model,

$$y_k = h_k(\hat{x}_k, 0) + H_k(x_k - \hat{x}_k) + M_k v_k$$

Prediction,

$$\hat{x}_k = f_{k-1}(\hat{x}_{k-1}, u_{k-1}, 0)$$
$$\hat{P}_k = F_{k-1}\hat{P}_{k-1}F_{k-1}^\intercal + L_{k-1}Q_{k-1}L_{k-1}^\intercal$$

Optimal gain,

$$K_k = \hat{P}_k H_k^\intercal (H_k \hat{P}_k H_k^\intercal + M_k R_k M_k^\intercal)^{-1}$$

Correction,

$$\hat{x}_k = \hat{x}_k + K_k(y_k - h_k(\hat{x}_k, 0))$$
$$\hat{P}_k = (1 - K_k F_k)\hat{P}_k$$

Linearization is performed by computing a local linear approximation of the nonlinear function with respect to the selected operating point. Linearization relies on computing the Jacobian matrix (consisting of the first partial derivatives of the function)

From the main idea of EKF, it is not difficult to see that since only the first-order approximation is used, in subsequent iterations, the error may become larger and larger. This is also its disadvantage.

*Moving Horizon Estimation:

Moving horizon estimation is an optimization-based state-estimation technique where the current state of the system is inferred based on a finite sequence of past measurements.

In comparison to more traditional state-estimation methods, e.g. the extended Kalman filter (EKF), MHE will often outperform the former in terms of estimation accuracy. This is especially true for non-linear dynamical systems, which are treated rigorously in MHE and where the EKF is known to work reliably only if the system is almost linear during updates.

Another advantage of MHE is the possible incorporation of further constraints on estimated variables. These can be used to enforce physical bounds, e.g. fractions between 0 and 1.

All of this comes at the cost of additional computational complexity.

The basic idea of MHE is to estimate the state trajectory using only the past N measurements, where N is referred to the length of horizon window. An nonlinear programming (NLP) problem is formulated at each sample time where a batch of past measurements is introduced and the states are estimated simultaneously. The horizon window allows incorporation of both frequent and infrequent measurements where the slow ones can be introduced as they are available.

The MHE formulation is shown as,

$$\min_{z_l - N} n\|z_l - N - \overline{z}_l - N\|^2 W_z + \sum_{l=0}^{N} \|y_l - \overline{y}_{l+k-N}\|^2 W_y \tag{20a}$$

$$\text{s.t.} z_{l+1} = f(z_l, u_l) + w_l \tag{20b}$$

$$y_l = h(z_l) + v_l \tag{20c}$$

where the process is assumed to be modelled by Eq. 20b 20c,z_l is the vector of states, u_l is the vector of inputs,y_l is the vector of measurements,$w_l$is the vector of unknown disturbances, and v_l is the vector of measurement noise. The noise variables are assumed to be independent of states and $w_l \sim N(0, Q_l)$ and $v_l \sim N(0, R_l).W_z, W_y$ represent the inverse of the covariance matrices of the prior state estimates $\overline{z}_{l-N}$ and measurement $\overline{y}_{l+k-N}$, respectively.The state vector is denoted as $z_l = [x_l, p_l]$,wherex_l are the reactant concentrations, and p_l are kinetic parameters and $r_a(j)$.It is also assumed the bounds and constraints of states can be further added in the NLP formulation.

16. Problem 22

Consider the system

$$x_{t+1} = Ax_t + Bu_t + Gw_t,$$
$$y_t = Cx_t + v_t,$$

where (i) x_0 is a random variable and (iv) $x_0 \sim \mathcal{N}(\widetilde{x}_0, P_0)$, (ii) $\mathbb{E}[w_t] = 0$ and $\mathbb{E}[v_t] = 0$ for all $t \in \mathbb{N}$,(iii) $x_0, (w_t)_t$ and $(v_t)_t$ are mutually independent random variables, (iii) w_t and v_t are normallydistributed and $\mathbb{E}[w_t w_t^\intercal] = Q, \mathbb{E}[v_t v_t^\intercal] = R$.Determine the covariance matrix $P_{t,l}$ =Cov(x_t, x_l) for $t, l \in \mathbb{N}$,as a function of P_0, A, G, Q, t and l.

Solotion:

Assume $l \geq t$.If $l = t$,

$$P_{t,t} = \mathrm{Cov}(x_t, x_t) = \mathrm{Var}(x_t)$$

We have,

$$P_{0,0} = P_0 \tag{21a}$$
$$P_{t+1} = AP_tA^\intercal + Q \tag{21b}$$

If $l = t + 1$,

$$\begin{aligned}
\mathrm{Cov}(x_t, x_{t+1}) &= \mathrm{Cov}(x_t, Ax_t + Bu_t + Gw_t) \\
&= \mathrm{Cov}(x_t, Ax_t) + \mathrm{Cov}(x_t, Bu_t) + \mathrm{Cov}(x_t, Gw_t) \\
&= \mathrm{Cov}(x_t, x_t) \cdot A^\intercal + \mathrm{Cov}(x_t, x_t) \cdot B^\intercal + \mathrm{Cov}(x_t, x_t) \cdot G^\intercal \\
&= P_t \cdot A^\intercal + P_t \cdot B^\intercal + P_t \cdot G^\intercal
\end{aligned}$$

And,

$$\mathrm{Cov}(x_t, x_{t+2}) = P_t \cdot (A^\intercal)^2 + P_t \cdot (B^\intercal)^2 + P_t \cdot (G^\intercal)^2 \tag{22a}$$
$$\mathrm{Cov}(x_t, x_{t+3}) = P_t \cdot (A^\intercal)^3 + P_t \cdot (B^\intercal)^3 + P_t \cdot (G^\intercal)^3 \tag{22b}$$
$$\vdots$$
$$\mathrm{Cov}(x_t, x_{t+n}) = P_t \cdot (A^\intercal)^n + P_t \cdot (B^\intercal)^n + P_t \cdot (G^\intercal)^n \tag{22c}$$

Where,

$$P_1 = AP_0A^\intercal + Q \tag{23a}$$
$$P_2 = AP_1A^\intercal + Q = A(AP_0A^\intercal + Q)A^\intercal + Q = A^2P_0(A^\intercal)^2 + AQA^\intercal + Q \tag{23b}$$
$$P_3 = A^3P_0(A^\intercal)^3 + A^2Q(A^\intercal)^2 + AQA^\intercal + Q \tag{23c}$$
$$\vdots$$
$$P_t = A^tP_0(A^\intercal)^t + A^{t-1}Q(A^\intercal)^{t-1} + A^{t-2}Q(A^\intercal)^{t-2} + \cdots + A^{t-t}Q(A^\intercal)^{t-t} + Q \tag{23d}$$

Combining with 22c and 23d,

$$\mathrm{Cov}(x_t, x_l) =$$