Reinforcement Learning and Optimal Control
Project 1
Yash Patel(yp2285)

Part 1 - Setting up

1. Discretize the system dynamics using the method seen in class - write the time discretization step as Δt (use symbols not numbers for the mass, etc).

$$Z_{n+1} = Z_n + \Delta t(dy/dt) \text{ ---------------------------- (A)}$$

Based on the above equation (A) -

$$x_{n+1} = x_n + \Delta t V_{xn} \text{ ---------------------------- (B)}$$

$$y_{n+1} = y_n + \Delta t V_{yn} \text{ ---------------------------- (C)}$$

$$V_{xn+1} = V_x + \Delta t(-(u1 + u2)sin(\theta))/m \text{ ---------------------------- (D)}$$

$$V_{yn+1} = V_y + \Delta t((u1 + u2)cos(\theta) - mg)/m \text{ ---------------------------- (E)}$$

$$\theta_{n+1} = \theta_n + \Delta t\omega_n \text{ ---------------------------- (F)}$$

$$\omega_{n+1} = \omega_n + \Delta t * r * (u1 - u2)/I \text{ ---------------------------- (G)}$$

2. Assume that the robot starts at an arbitrary position $x(0) = x$, $y(0) = y$ and $\theta(0) = 0$ with 0 velocities. Compute $u1^*$ and $u2^*$ such that the robot stays at this position forever after (you may test your answer using the simulation below).

From the above equations B,C,D,E,F,G, substituting V_x V_y $\omega_n = 0$

$$0 = 0 + \Delta t(-(u1 + u2)sin(\theta))/m \text{ --------------------(H)}$$

$$0 = 0 + \Delta t((u1 + u2)cos(\theta) - mg)/m \text{ ---------------(I)}$$

$$\theta_{n+1} = 0 \text{ ----------------------------(J)}$$

Solving the equation H,I,J -

$$u1^* = u2^* = m * g / 2$$

3. Analyzing the system dynamics, is it possible to move in the x direction while keeping θ=0. Explain why?

   From equation D when substituting $\theta = 0$: the sine part will give the value 0. So it's not possible.

4. Analyzing the system dynamics, is it possible to have the system at rest with θ=π/2 (i.e. have the quadrotor in a vertical position)? Explain why.

   If theta = pi/2, the control u1 and u2 will be zero as seen in equation E. Due to this a force of -mg will act downwards and the quadrotor cannot be controlled.

PART 2 - LQR to stay in PLACE

1. Linearize the dynamics

   Using taylor series we have,

$$A = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right]_{\mathbf{x}^*, u^*} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_1^*, \dots, x_n^*, u^*) & \cdots & \frac{\partial f_1}{\partial x_n}(x_1^*, \dots, x_n^*, u^*) \\ \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial x_1}(x_1^*, \dots, x_n^*, u^*) & \cdots & \frac{\partial f_n}{\partial x_n}(x_1^*, \dots, x_n^*, u^*) \end{bmatrix},$$

$$B = \left[\frac{\partial \mathbf{f}}{\partial u}\right]_{\mathbf{x}^*, u^*} = \begin{bmatrix} \frac{\partial f_1}{\partial u}(x_1^*, \dots, x_n^*, u^*) \\ \vdots \\ \frac{\partial f_n}{\partial u}(x_1^*, \dots, x_n^*, u^*) \end{bmatrix},$$

Based on the above equation we partial differentiate and compute A and B

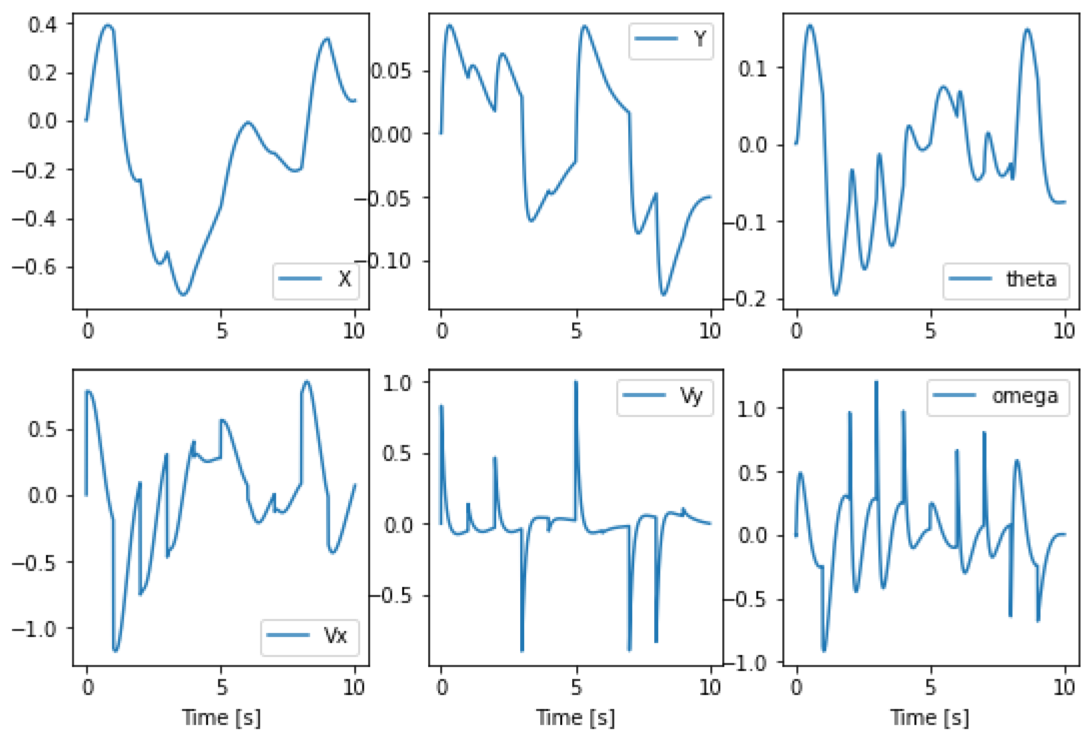
$$A = \begin{bmatrix} 1 & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t(\frac{-(u_1+u_2)cos\theta}{m}) & 0 \\ 0 & 0 & 1 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & 1 & \Delta t(\frac{-(u_1+u_2)sin\theta}{m}) & 0 \\ 0 & 0 & 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ \Delta t(\frac{-sin\theta}{m}) & \Delta t(\frac{-sin\theta}{m}) \\ 0 & 0 \\ \Delta t(\frac{cos\theta}{m}) & \Delta t(\frac{cos\theta}{m}) \\ 0 & 0 \\ \Delta t(\frac{r}{I}) & \Delta t(\frac{-r}{I}) \end{bmatrix}$$

The system of equations can be represented as -

$$z_{n+1} = \begin{bmatrix} x_{n+1} \\ v_{x_{n+1}} \\ y_{n+1} \\ v_{y_{n+1}} \\ \theta_{n+1} \\ \omega_{n+1} \end{bmatrix} = \begin{bmatrix} x_n + \Delta t V_{xn} \\ v_x + \Delta t \dfrac{-(u_1 + u_2)}{m} sin\theta \\ y_n + \Delta t V_{yn} \\ v_y + \Delta t \dfrac{(u_1 + u_2)cos\theta - mg}{m} \\ \theta_n + \Delta t \omega_n \\ \omega_n + \Delta t \dfrac{r(u1 - u2)}{I} \end{bmatrix}$$

$$x_{n+1} = A(z - z^*) + B(u - u^*)$$

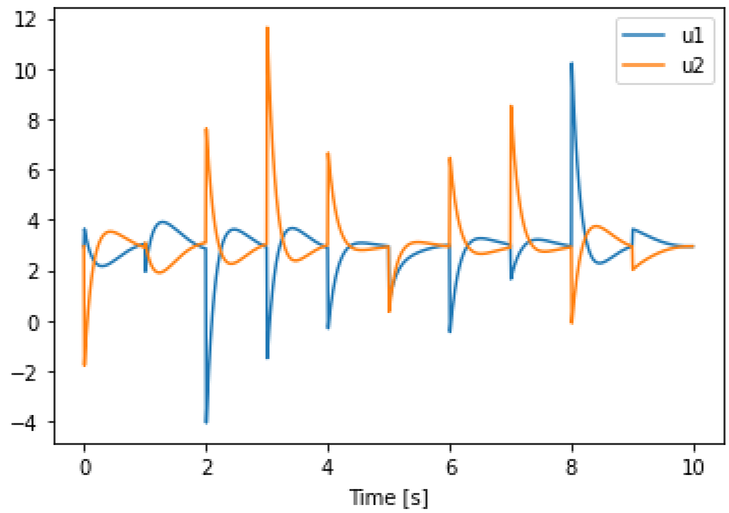
5. Plots (with disturbance)



The controller is able to counter the disturbance. THe controller design is based on the infinite horizon LQR, where the optimal policy is solved using the contol law -

$$u = K * z + u^*$$

Where, K = gain, z= state, u^* = control value to keep robot at rest, u = optimal control value



PART 3 - following Trajectory using Linearized dynamics

1. The quadrotor dynamics is designed about the point z^* and u^* due it tries to remain close to origin.
2. The desired trajectory to be followed by the quadrotor is circle i.e. according to the circle equation-

$$x^2 + y^2 = 1$$

In terms of sin and cosine -

$$x = sin(t)$$

$$y = cos(t)$$

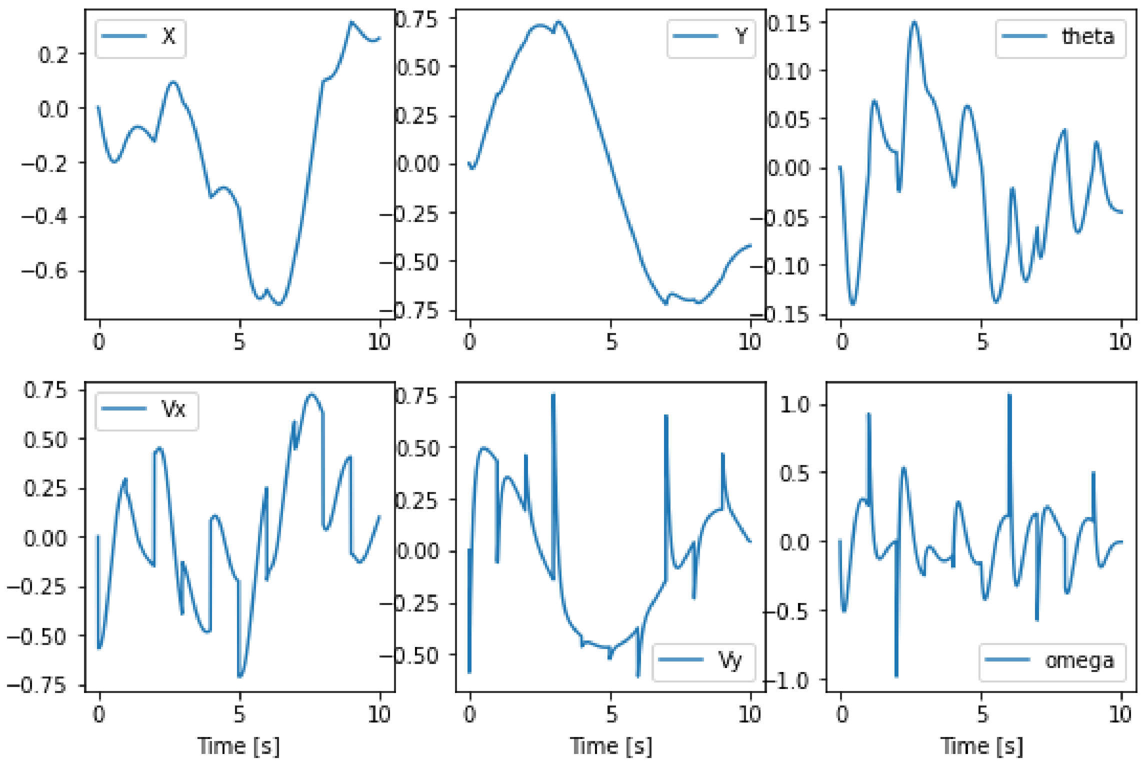
$$sin^2(t) + cos^2(t) = 1$$

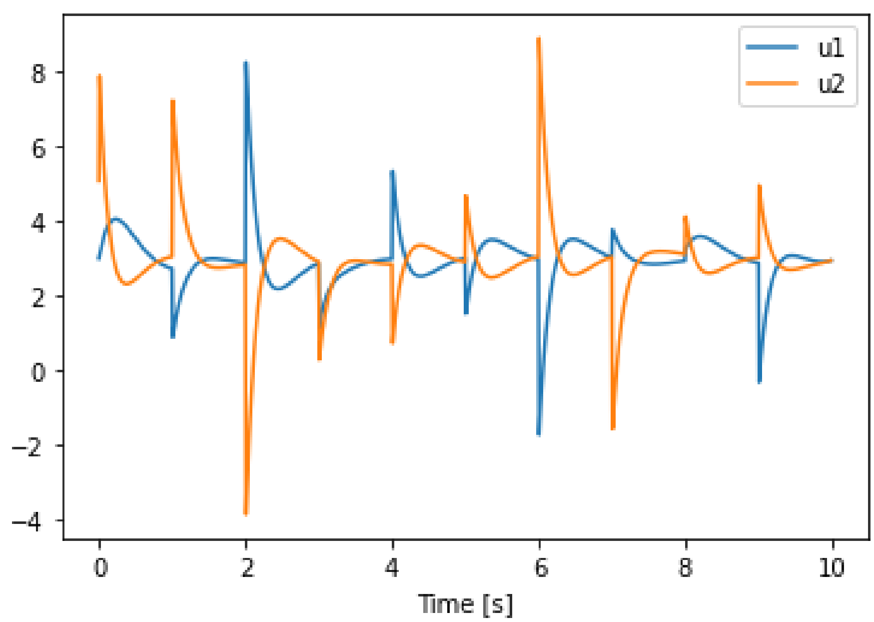
Riccati equation is used to find the optimal policy given by -

$$\bar{u}_t = K_t \bar{x}_t + k_t$$

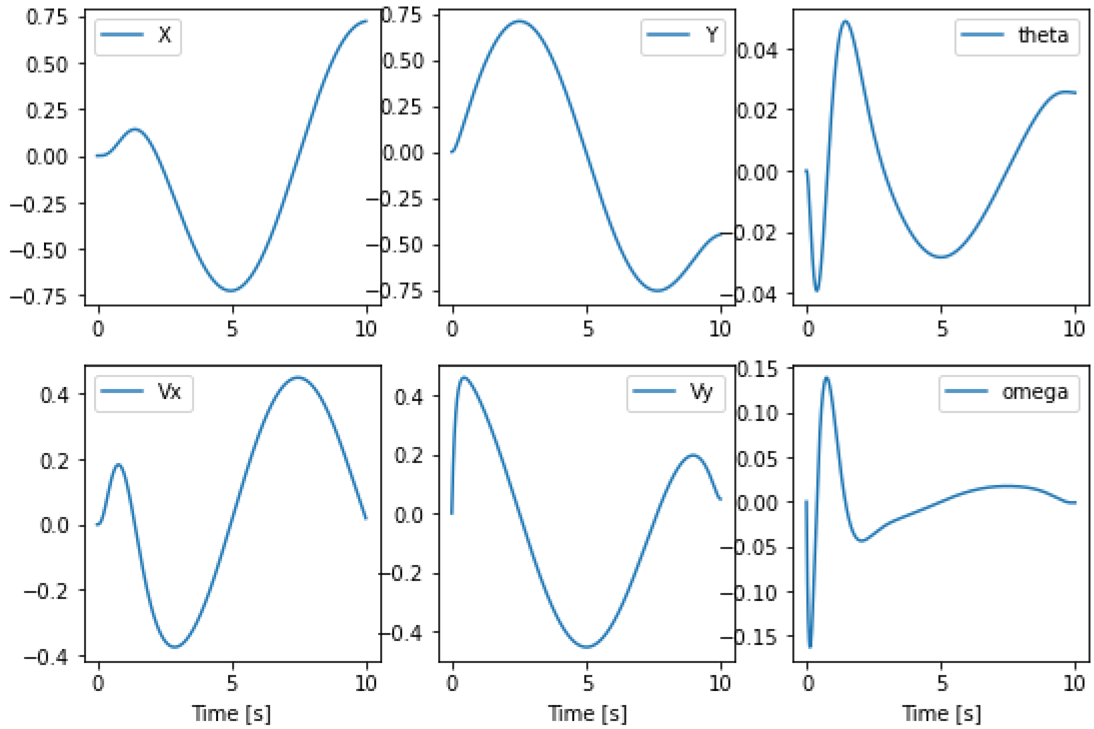
3. The tracking of theta is not exact, since the motion of the quadrotor is against gravity.

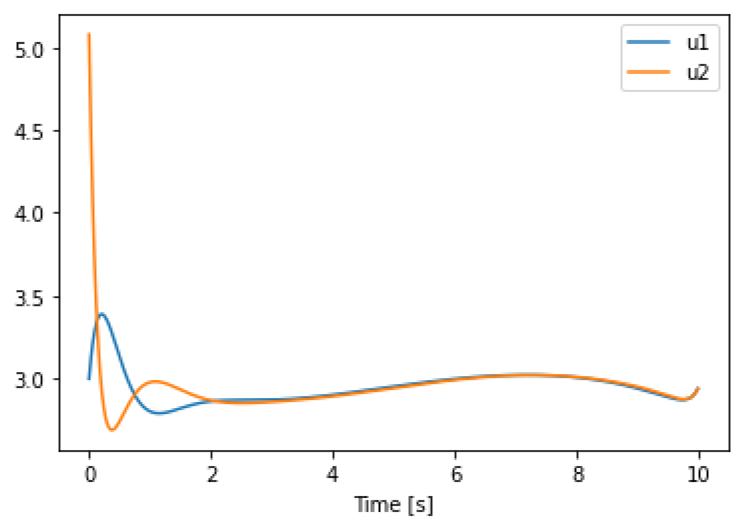
4. Plots (with disturbance)





Plots (without disturbance)





Moving in a planar motion (x-y) has dependency on theta value as well. Therefore it is not entirely possible to keep the value of theta = 0 . The controller makes the quadrotor follow the trajectory even when disturbance is present.