List of Contents (sections and sub-section with page numbers)

1. **Introduction**
   1. Motivation: Linear Quadratic regulator(LQR) and difficulty is choosing Q and R matrices (design matrices).
   2. Application: Drones! Benefits of faster response.
2. **Literature survey**: *Analyse advantages and disadvantages of literature in each section below.*
   1. LQR: its application to various systems (including drones).
      1. Survey of methods for choosing design matrices in LQR.
   * Example: Bryson’s rule, reinforcement learning (RL).
   * Note: Mention that using LQR with RL still guarantees optimality whereas using standalone RL agent to produce control inputs provides no such guarantee
     1. Methods for combining LQR with Proportional Integral Derivative (PID) control.
   * This should include Integral Square Error (ISE) optimization
   1. Reinforcement Learning for continuous action space
   * Include RL algorithm used in code: <https://arxiv.org/abs/1509.02971>
3. **Theory**
   1. Quadrotor dynamics, state space equation
   2. LQR
   3. PID control
      1. Deriving PID gains using ISE optimization
   4. Reinforcement Learning background
4. **Methodology (can make a separate section for benchmarks)**
   1. Control method for drone: LQR/ LQR+ PD control
   2. Bryson’s rule for selecting design matrices
      1. Changing all diagonal elements
      2. Changing only 4 diagonal elements.
   3. Reinforcement learning for selecting design matrices
      1. Task formulation: state, transition, action and reward in our context. (mention we are tuning only 4 diagonal elements)
      2. Actor, Critic neural network architecture. More details about the implementation like batch normalization.
5. **Numerical Experiments**

Few words about Numerical Experiments setup: Training and test conditions and how set point or trajectory is generated.

* 1. Test 1: Reaching a set point.

Results should include:

1. Training curve for RL agent: mean reward, reward per episode, number of failures, average length of episode.
2. Number of failures while training with RL agent. (the drone should be able to reach the set point more consistently as the training episodes increase)
3. ISE of all 4 methods over 100 random tests. Report number of failures when testing with RL agent.
4. Sample plot with all 4 methods.
   1. Test 2: Following a trajectory
5. Training curve for RL agent.
6. Number of failures while training with RL agent. (the drone should be able to reach the desired goal more consistently as the training episodes increase)
7. ISE of all 4 methods over 100 random tests. Report number of failures when testing with RL agent.
8. Sample plot with all 4 methods.
   1. Test 3: TODO: any additional tests.

This could include:

1. giving extra features like next two waypoints to the RL agent.
2. Different reward functions:
   1. Penalty due to limits on maximum input values (due to limit on maximum torque generated by motors).
   2. Penalty due to singularity of Q matrix: The Q matrix in LQR should well behaved.
3. Robustness of RL agent to noise in sensor measurements or model.
4. Robustness and response of RL agent to sudden change in trajectory due to obstacle.
5. **Conclusion**

Scale input

16×64

Fully connected Layer

FC 1

Input

1×12

(State input)

1×12

Scale the input vector

Concatenation

Layer

1×1

Range: [-1000,1000]

1×16

64×64

Fully connected Layer

FC 2

64×64

Fully connected Layer

FC 3

1×1

Range: [0,1]

64×1

Fully connected Layer

Scale output

Tanh Layer

FC 4

Scale Action

Action

1×4

Scale the action vector

1×4

(Action input)

64×64

Fully connected Layer

1×24

Scale the input vector

4×4

Range: [0.01,100]

4×4

Range: [0,1]

4×64

Fully connected Layer

FC 3

Scale output

Sigmoid Layer

FC 4

1×24

(State input + goal input)

24×64

Fully connected Layer

FC 2

Input

Scale input

FC 1

64×64

Fully connected Layer

y

u

-

x

K

Figure 1.

|  |  |
| --- | --- |
|  | Action-value function in Reinforcement Learning |
|  | State of control system |
|  | Action given by RL agent |
|  | Probability function |
|  | Parameters of policy function in RL |
|  | Policy function in RL |
|  | Cosine rotation matrix |
|  | Transfer matrix |
|  | Mass of quadrotor |
|  | Identity matrix |
|  | Linear velocity of quadrotor |
|  | Rotational velocity of quadrotor wrt body frame |
|  | External force acting on quadrotor |
|  | External torque acting on quadrotor |
|  | Rotation speed of rotors |
|  | Thrust force |
|  | roll torque |
|  | pitch torque |
|  | yaw torque |
|  |  |
|  |  |
|  |  |
|  | Gyroscopic torque |
|  | Moment of inertia of propellor |
|  |  |
|  | Angular rates wrt body frame |
| and | Principal moments of inertia |
|  |  |
|  | Proportional gain value in PID controller |
|  |  |
|  |  |
|  | Gradient of vector |
|  | Transition function in RL |
|  | Discount factor in RL |
|  | Reward function in RL |
|  | Expectation |