Model Free Reinforcement Learning

Application to leader follower

Amr Elhussein
Advisor: Dr. Suruz Miah

Department of Electrical and Computer Engineering
Bradley University
1501 W. Bradley Avenue
Peoria, IL, 61625, USA

Friday, October 11, 2019

BRADLEY University

Outline

- Reminders
- Cost Function
- Model Free RL Algorithm
- How to calculate P
- Results

Reminders

Objective

Our Goal is to find optimum policy that eleminates the tracking error

$$u = [v, \gamma] \tag{1}$$

$$\mathbf{e}_k = \begin{bmatrix} x_k^{[\ell]} - x_k - d\cos\theta_k' \\ y_k^{[\ell]} - y_k - d\sin\theta_k' \\ \theta_k' - \theta_k \end{bmatrix} \tag{2}$$

Cost Function

Minimize

$$J(\mathbf{u}) = \sum_{k=0}^{\infty} \frac{1}{2} \left[\mathbf{e}_k^T \mathbf{Q} \mathbf{e}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_\mathbf{k} \right] \tag{3}$$

Algorithm

Input: Sampling time T, $\mathbf{e}_0$, $\mathbf{Q}$, $\mathbf{R}$, and threshold ε
Output: Optimal error tracjectory $\mathbf{e}_k$, for $k = 0, 1, \ldots$
begin
  $k = 0, r = 0$ [h]Discrete time and policy indices;
  $\eta = (n + m)(n + m + 1)/2$
  Initialize $\mathbf{P}^{[0]}$ [h]RH and positive definite **repeat**[h]Main timing loop
    Find $\mathbf{e}_{k+1}$;
    Compute policy $\mathbf{u}_{k+1}^{[r]}$;
    if $[(k + 1) \text{ modulo } \eta] == 0$ **then**
      $r \leftarrow r + 1$[h]Update policy
      Solve for critic weights $\mathbf{w}$; Construct $\mathbf{P}^{[r]}$ from critic weight vector $\mathbf{w}$
      if $\|\mathbf{P}^{[r]} - \mathbf{P}^{[r+1]}\| < \varepsilon$ **then**
        Set $\mathbf{u}_{k+1}^{*} \leftarrow \mathbf{u}_{k+1}^{[r]}$
      else
        $k \leftarrow k + 1$
      end
    else
      $k \leftarrow k + 1$
    end
  until *Forever or tracking error is zero*
end

How to calculate P matrix

Least Square

$$w = (\mathbf{\Lambda^T \Lambda})^{-1} \mathbf{\Lambda^T v} \tag{4}$$

gradient descent

$$w^{r+1} = w^r - \mathbf{l_c \Lambda^T (\Lambda w^r - v)} \tag{5}$$

Results

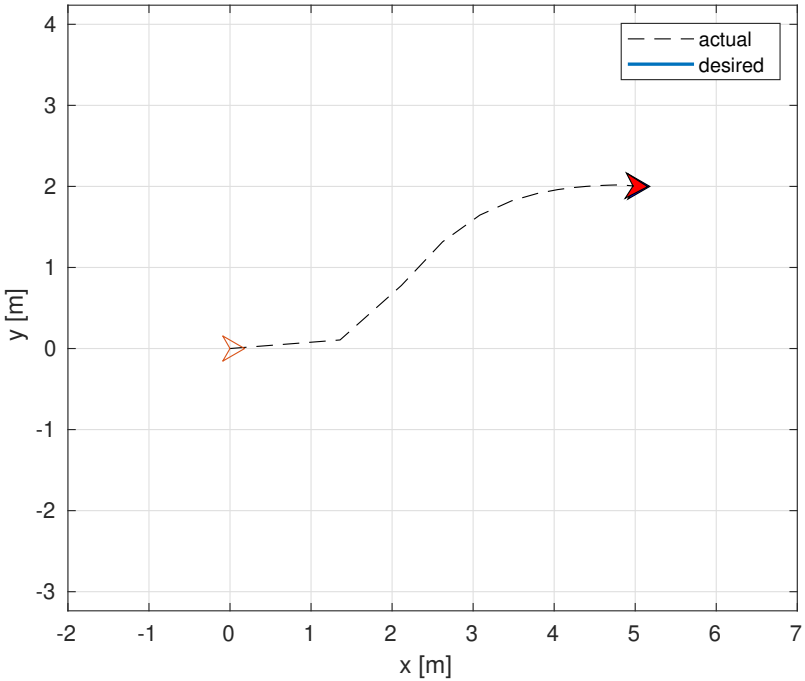

Figure: Trajectory

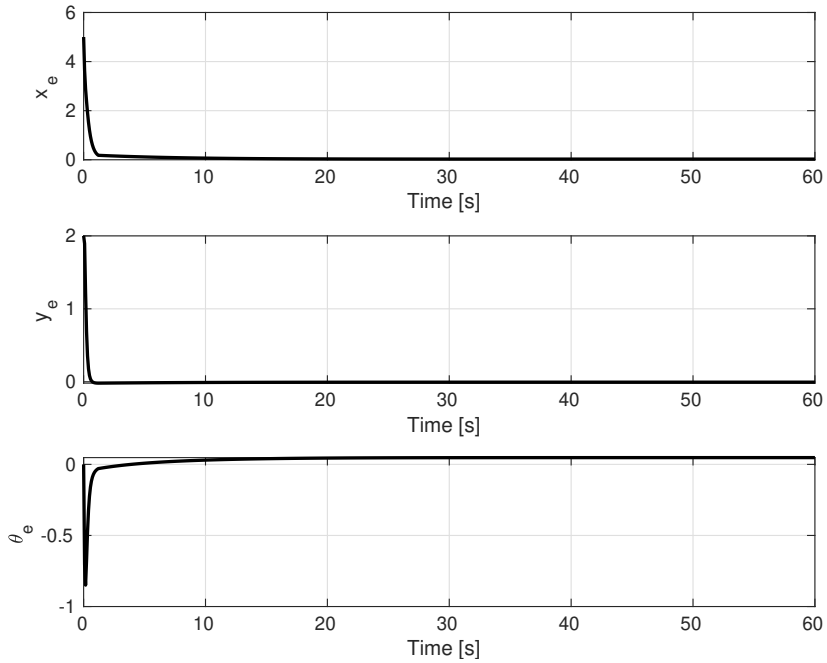

Figure: follower position error

Results

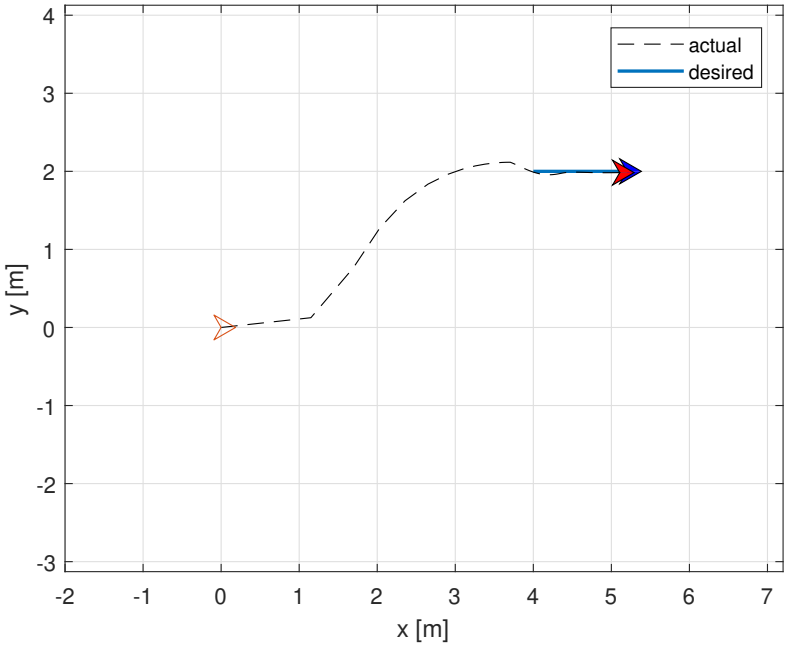

Figure: Trajectory

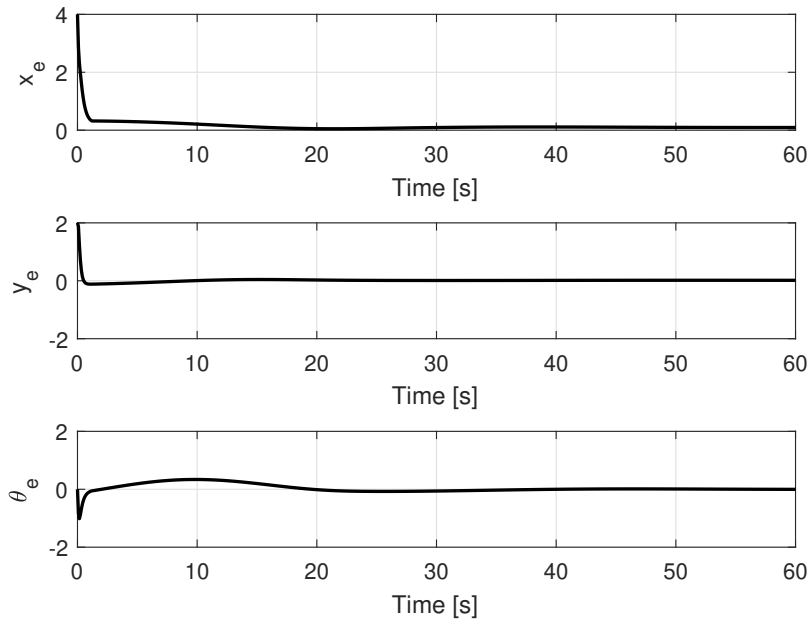

Figure: follower position error

Results

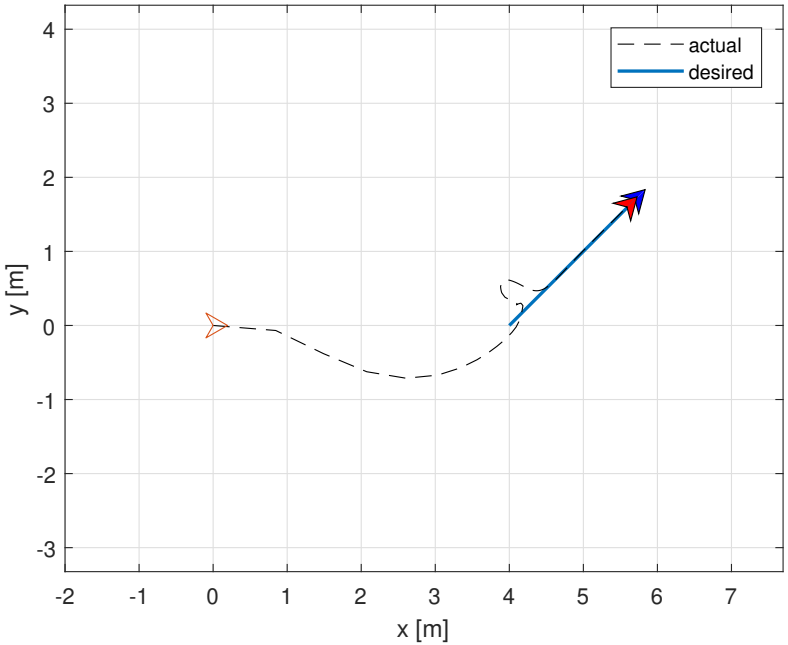

Figure: Trajectory

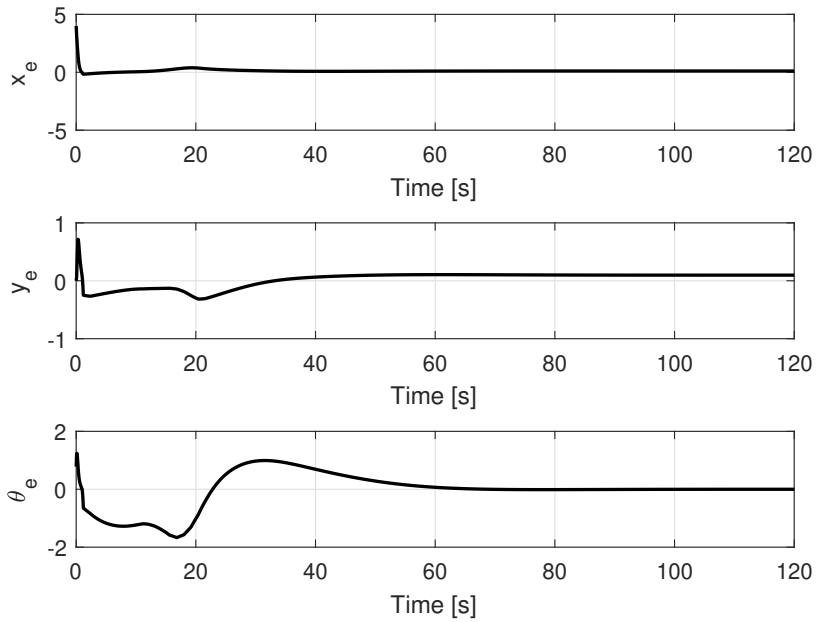

Figure: follower position error

Questions?