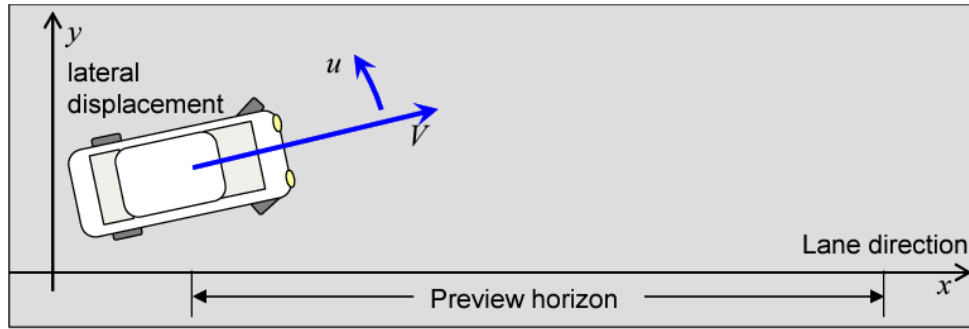


Άσκηση 5: Νευροασαφής Έλεγχος και Εφαρμογές 2021-2022

Μέρος A (Model Predictive Control)

Στην άσκηση αυτή θα σχεδιάσουμε έναν προβλεπτικό ελεγκτή για τη θέση ενός αυτοκινήτου μέσα στη λωρίδα του. Έστω a η γωνία μεταξύ της ταχύτητας του οχήματος και της κατεύθυνσης x της λωρίδας κυκλοφορίας.



Η δυναμική του οχήματος δίνεται από τις σχέσεις

$$\dot{x} = V \cos a$$

$$\dot{y} = V \sin a$$

$$\dot{a} = u$$

όπου u είναι ανάλογο με τη γωνία στροφής του τιμονιού.

Θέλοντας να επικεντρώσουμε στη θέση του αυτοκινήτου μέσα στη λωρίδα, θα εξετάσουμε μόνο τη δεύτερη και τρίτη εξίσωση. Γραμμικοποιώντας και διακριτοποιώντας το χρόνο, έχουμε τη δυναμική

$$x_{k+1} = Ax_k + Bu_k,$$

$$A = \begin{bmatrix} 1 & 0 \\ VT_s & 1 \end{bmatrix}, \quad B = \begin{bmatrix} T_s \\ 0.5VT_s^2 \end{bmatrix},$$

$$\text{όπου } x_k = \begin{bmatrix} a_k & y_k \end{bmatrix}^T.$$

Κατασκευάστε έναν ελεγκτή MPC που σε κάθε βήμα να λύνει το

$$\text{Minimize } J = \lambda_1 |y_{t+3}| + \lambda_2 |a_{t+3}| + \sum_{k=0}^T (\lambda_1 |y_{t+k}| + \lambda_2 |a_{t+k}| + |u_{t+k}|)$$

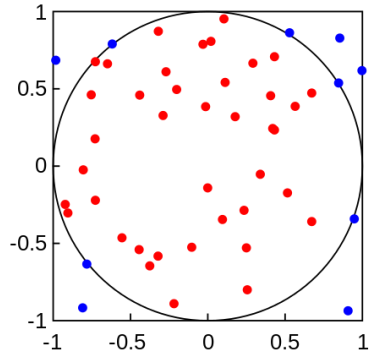
$$\text{subj. to } |u_{t+k}| \leq U$$

$$\begin{bmatrix} a_{t+k+1} \\ y_{t+k+1} \end{bmatrix} = A \begin{bmatrix} a_{t+k} \\ y_{t+k} \end{bmatrix} + Bu_k$$

Μπορείτε να πειραματιστείτε με διάφορες τιμές για το T , λ_1, λ_2 . Θεωρήστε ότι $U=1$, $T_s=0.1$ και $V=1$.

Μέρος Β (Υπολογισμός του π με Monte Carlo Προσομοίωση).

Στο θέμα αυτό θα εξετάσουμε μια Monte Carlo μέθοδο για τον υπολογισμό του π , παίρνοντας τυχαίες μεταβλητές, ομοιόμορφα κατανομημένες στο $[0,1]^2$ και εξετάζοντας το πόσες από αυτές βρίσκονται μέσα στο μοναδιαίο κύκλο.



Μέρος Γ (Τυχαίος Περίπατος με έλεγχο: Άγνωστο Σύστημα).

Εδώ θα επαναλάβουμε το Θέμα 3 από την Άσκηση 4, υποθέτοντας ότι η δυναμική του συστήματος και το κόστος δεν είναι γνωστά.

Θεωρήστε ένα στοχαστικό σύστημα (controlled Markov chain) του οποίου η κατάσταση μπορεί να πάρει τις τιμές $1, \dots, 10$. Ο έλεγχος παίρνει δύο τιμές $+1$ και -1 . Για $u=+1$, η κατάσταση μεταβαίνει μια θέση δεξιότερα με πιθανότητα 50% και παραμένει η ίδια με πιθανότητα 50%. Αντίστοιχα για $u=-1$ η κατάσταση μεταβαίνει μια θέση αριστερότερα με πιθανότητα 50% και παραμένει η ίδια με πιθανότητα 50%. Έστω το κριτήριο κόστους

$$J = \sum_{k=0}^{\infty} \alpha^k g(x_k),$$

Το g περιγράφεται στον παρακάτω πίνακα:

x	1	2	3	4	5	6	7	8	9	10
$g(x)$	1	2	3	4	5	4	2	0	1	2

1. Υποθέστε ότι ο ελεγκτής είναι δοσμένος. Γράψτε ένα πρόγραμμα που να βρίσκει τροχιές του συστήματος.
2. Υλοποιήστε το Q-learning για αυτό το σύστημα. Εξετάστε το την επίπτωση των διαφόρων παραμέτρων στην ταχύτητα σύγκλισης.