

### L3. Proiectarea reguletoarelor optimale după stare cu ajutorul algoritmilor *Policy Iteration (PoIt)* și *Value Iteration (VI)*

Lucrarea ilustrează calculul unui regulator optimal după stare folosind algoritmi Policy Iteration (PoIt) și Value Iteration (VI) reprezentativi pentru tehnicile de învățare prin recompensă.

Fie procesul condus dat prin funcția de transfer  $H(s) = 1/[(1 + 0.5s)(1 + 0.2s)]$  care conduce la un model discretizat de ordinul doi pe stare, prin discretizare cu perioada de eșantionare de  $T_s = 0.1\text{sec}$  folosind metoda realizării invariante la semnal treaptă (r.i.s.t.). Se cere proiectarea unui regulator optimal liniar după stare care să asigure minimizarea funcției de cost (FC) de forma

$$J(\mathbf{x}_0) = \sum_{i=0}^{\infty} \{ \mathbf{x}_i^T \mathbf{Q} \mathbf{x}_i + u_i^T \mathbf{R} u_i \}, \mathbf{Q} = \mathbf{I}_2 = \text{diag}(1,1), \mathbf{R} = 0.1.$$

Un model transformat de tip intrare-stare-ieșire pentru procesul de mai sus, de tip formă canonică controlabilă<sup>1</sup>, obținut cu ajutorul funcției *tf2ss* din Matlab, este de forma

$$\begin{cases} \begin{pmatrix} x_{1,k+1} \\ x_{2,k+1} \end{pmatrix} = \begin{pmatrix} 1.4253 & -0.4966 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{1,k} \\ x_{2,k} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u_k \\ y_k = \begin{pmatrix} 0.0398 & 0.0315 \end{pmatrix} \begin{pmatrix} x_{1,k} \\ x_{2,k} \end{pmatrix} \end{cases} \Leftrightarrow \begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_d \mathbf{x}_k + \mathbf{B}_d u_k, \\ y_k = \mathbf{C} \mathbf{x}_k \end{cases} \quad (1)$$

Pentru a garanta calculul unui regulator liniar după stare care să minimizeze FC  $J$ , în primul pas se verifică faptul că perechea  $(\mathbf{A}_d, \mathbf{B}_d)$  este stabilizabilă<sup>2</sup>. Întrucât rangul matricei de controlabilitate de dimensiune 2x2 aferente perechii  $(\mathbf{A}_d, \mathbf{B}_d)$  este doi, controlabilitatea perechii implică stabilizabilitatea. Deci există cel puțin un regulator liniar după stare de forma  $u_k = \mathbf{K}^T \mathbf{x}_k$  care să asigure un sistem de reglare automată (SRA) în buclă închisă asimptotic stabil, care rezultă într-o valoare finită a FC  $J$ .

#### I) Calculul regulatorului liniar optimal după stare folosind ecuații algebrice matriceale Riccati.

Este recunoscută în problema de minimizare a FC  $J$  problema regulatorului liniar pătratic (LQR) cu orizont de timp infinit în cazul omogen staționar de timp discret. Problema de optimizare (PO) poate fi reformulată

<sup>1</sup> Pentru o f.d.t. monovariabilă de forma  $H(s) = \frac{b_n s^n + \dots + b_1 s + b_0}{a_n s^n + \dots + a_1 s + a_0}$ , o formă canonică controlabilă (Preitl et al., 2009) conduce la

$$\text{matricele pe stare} \quad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ -\frac{a_0}{a_n} & -\frac{a_1}{a_n} & \dots & \dots & -\frac{a_{n-1}}{a_n} \end{bmatrix} \in R^{n \times n}, \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \\ \frac{1}{a_n} \end{bmatrix} \in R^{n \times 1}, \mathbf{C} = [b_0 - a_0 \frac{b_n}{a_n} \dots b_{n-1} - a_{n-1} \frac{b_n}{a_n}], \mathbf{D} = b_n / a_n.$$

Conform documentației Matlab, nu există consens în literatură pe denumirea formelor canonice pe stare și că transformarea canonică controlabilă ce rezultă în urma apelării funcției *tf2ss* nu produce matricele pe stare de mai sus.

<sup>2</sup> Perechea  $(\mathbf{A}, \mathbf{B})$  caracteristică sistemului liniar invariant în timp (1) este stabilizabilă dacă există un regulator liniar după stare  $u_k = \mathbf{K}^T \mathbf{x}_k$  caracterizat de matricea  $\mathbf{K}$  care, pentru sistemul de reglare în buclă închisă  $\mathbf{A} + \mathbf{B}\mathbf{K}^T$  face ca  $\lim_{k \rightarrow \infty} \|\mathbf{x}_k\| = 0$ . Echivalent, există matricea  $\mathbf{K}$  care face ca valorile proprii ale matricei sistemului închis să fie în semiplanul complex stâng:  $\sigma(\mathbf{A} + \mathbf{B}\mathbf{K}^T) \subset C_{(-\infty, 0)}$ . Controlabilitatea perechii  $(\mathbf{A}, \mathbf{B})$  implică stabilizabilitatea sa și poate fi verificată direct în Matlab folosind funcția *ctrb*.

$$\mathbf{K}^* = \arg \min_{\mathbf{K}} J(\mathbf{x}_0, \mathbf{K}), \quad (2)$$

$$s.l. \quad \mathbf{x}_{k+1} = \mathbf{A}_d \mathbf{x}_k + \mathbf{B}_d u_k.$$

Cu ajutorul funcției Matlab *dlqr*, se calculează soluția  $\mathbf{S} = [-2.2078 \ -0.0663; -0.0663 \ 1.0236]$  aferentă ecuației algebrice matriceale Riccati de forma  $\mathbf{A}_d^T \mathbf{S} \mathbf{A}_d - \mathbf{S} - \mathbf{A}_d^T \mathbf{S} \mathbf{B}_d (\mathbf{R} + \mathbf{B}_d^T \mathbf{S} \mathbf{B}_d)^{-1} \mathbf{B}_d^T \mathbf{S} \mathbf{A}_d + \mathbf{Q} = \mathbf{0}$ . Este cunoscut faptul că FC optimală este pătratică  $J^*(\mathbf{x}_k) = \mathbf{x}_k^T \mathbf{S} \mathbf{x}_k$ . Regulatorul liniar pătratic optimal este calculat ca  $\mathbf{K}^* = -(\mathbf{R} + \mathbf{B}_d^T \mathbf{S} \mathbf{B}_d)^{-1} \mathbf{B}_d^T \mathbf{S} \mathbf{A}_d = [-1.3348 \ 0.4751]^T$ .

## II) Calculul regulatorului liniar optimal după stare folosind algoritmul PoIt în varianta bazată pe model.

În baza exemplului anterior de calcul al regulatorului liniar pătratic optimal, va fi exemplificată procedura de calcul a aceluiași regulator, utilizând algoritmul PoIt în varianta bazată pe model cunoscut al procesului (Radac, 2019).

Pentru amorsarea algoritmului PoIt, este necesar un regulator inițial liniar după stare admisibil<sup>3</sup>. Un astfel de regulator inițial poate fi obținut, de exemplu, prin alocarea polilor SRA (a spectrului matricei  $\mathbf{A}_d + \mathbf{B}_d \mathbf{K}^T$  a sistemului închis). Fie polii alocați reali în timp continuu  $s_1 = -10, s_2 = -9$ , transformați în timp discret cu formula  $e^{sT_s}$  la valoarea  $P = [z_1 = 0.3679, z_2 = 0.4066]^T$ . Apoi, funcția *place*( $\mathbf{A}_d, \mathbf{B}_d, \mathbf{P}$ ) din Matlab conduce la regulatorul liniar după stare inițial  $\mathbf{K}_0 = [0.6508 \ -0.3470]^T$ .

În continuare, va fi evaluată funcția de cost aferentă unui regulator oarecare  $\mathbf{u}_k = \mathbf{K}^T \mathbf{x}_k$ . Ecuația Bellman afirmă că pentru orice regulator admisibil, este valabil că

$$J^{\mathbf{K}}(\mathbf{x}_k) = \mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k + J^{\mathbf{K}}(\mathbf{x}_{k+1}). \quad (3)$$

Se poate arăta că, pentru un regulator liniar după stare  $\mathbf{u}_k = \mathbf{K}^T \mathbf{x}_k$ , pentru  $J^{\mathbf{K}}$  pe un orizont de timp infinit,  $J^{\mathbf{K}}(\mathbf{x}_k)$  este de forma  $J^{\mathbf{K}}(\mathbf{x}_k) = \mathbf{x}_k^T \mathbf{M} \mathbf{x}_k$ . Folosind ecuația de dinamică  $\mathbf{x}_{k+1} = \mathbf{A}_d \mathbf{x}_k + \mathbf{B}_d \mathbf{u}_k$  în (3) precum și expresia legii de reglare  $\mathbf{u}_k = \mathbf{K}^T \mathbf{x}_k$ , se deduce că

$$\mathbf{x}_k^T \mathbf{M} \mathbf{x}_k = \mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k + \mathbf{x}_k^T (\mathbf{A}_d + \mathbf{B}_d \mathbf{K}^T)^T \mathbf{M} (\mathbf{A}_d + \mathbf{B}_d \mathbf{K}^T) \mathbf{x}_k, \quad (4)$$

de unde rezultă ecuația matriceală algebrică în necunoscuta  $\mathbf{M}$ :

$$\mathbf{M} - \mathbf{Q} - \mathbf{K} \mathbf{R} \mathbf{K}^T - (\mathbf{A}_d + \mathbf{B}_d \mathbf{K}^T)^T \mathbf{M} (\mathbf{A}_d + \mathbf{B}_d \mathbf{K}^T) = \mathbf{0}. \quad (5)$$

Spre exemplu, pentru regulatorul inițial admisibil  $\mathbf{K}_0 = [0.6508 \ -0.3470]^T$ , rezolvând ecuația (5), FC aferentă utilizării acestui regulator pe un orizont de timp infinit va avea o formă pătratică caracterizată de matricea simetrică  $\mathbf{M} = [3.7907 \ -0.4016; -0.4016 \ 1.0968]$ . Ecuația (5) nu depinde de starea  $\mathbf{x}_k$ .

Pașului de evaluare al FC îi urmează, în cadrul algoritmului PoIt, cel de îmbunătățire al regulatorului. Condiția de optimalitate care duce la valoarea minimă a FC, scrisă pe baza ecuației Bellman este

$$J^*(\mathbf{x}_k) = \min_{\mathbf{u}_k} \{ \mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k + J(\mathbf{x}_{k+1}) \}, \quad (6)$$

și indică operația care trebuie efectuată pentru a obține regulatorul optimal. Pentru a calcula regulatorul îmbunătățit, se calculează comanda care minimizează membrul drept al ecuației (6) ca soluția  $\bar{\mathbf{u}}_k$  care anulează gradientul FC în raport cu comanda

<sup>3</sup> Care face ca SRA să fie asimptotic stabil.

$$\frac{\partial J(\mathbf{x}_k)}{\partial \mathbf{u}_k} = \frac{\partial \{\mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k + \mathbf{x}_{k+1}^T \mathbf{M} \mathbf{x}_{k+1}\}}{\partial \mathbf{u}_k} = 0 \quad \Leftrightarrow \quad \frac{\partial \{\mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k + \mathbf{x}_k^T \mathbf{A}_d^T \mathbf{M} \mathbf{A}_d \mathbf{x}_k + 2\mathbf{x}_k^T \mathbf{A}_d^T \mathbf{M} \mathbf{B}_d \mathbf{u}_k + \mathbf{u}_k^T \mathbf{B}_d^T \mathbf{M} \mathbf{B}_d \mathbf{u}_k\}}{\partial \mathbf{u}_k} = 0 \quad \Leftrightarrow, \quad (7)$$

$$2\mathbf{R} \mathbf{u}_k + 2\mathbf{B}_d^T \mathbf{M}^T \mathbf{A}_d \mathbf{x}_k + 2\mathbf{B}_d^T \mathbf{M} \mathbf{B}_d \mathbf{u}_k = 0 \quad \Leftrightarrow$$

$$\bar{\mathbf{u}}_k = -(\mathbf{R} + \mathbf{B}_d^T \mathbf{M} \mathbf{B}_d)^{-1} \mathbf{B}_d^T \mathbf{M}^T \mathbf{A}_d \mathbf{x}_k = \bar{\mathbf{K}} \mathbf{x}_k,$$

din ultima linie a relației putând fi dedusă expresia matricei  $\bar{\mathbf{K}} = -(\mathbf{R} + \mathbf{B}_d^T \mathbf{M} \mathbf{B}_d)^{-1} \mathbf{B}_d^T \mathbf{M}^T \mathbf{A}_d$  a regulatorului îmbunătățit.

Prin alternarea relațiilor (5) și (7), utilizând  $\bar{\mathbf{K}}$  din (5) în (7) pentru aflarea matricei  $\mathbf{M}$  care caracterizează FC aferentă regulatorului  $\bar{\mathbf{u}}_k = \bar{\mathbf{K}}^T \mathbf{x}_k$ , algoritmul Polt converge în doar câteva iterații la regulatorul optimal, pentru exemplul procesului de ordin doi la valoarea  $\mathbf{K}^* = [-1.3348 \quad 0.4751]^T$ . Merită remarcată similitudinea expresiei  $\bar{\mathbf{K}} = -(\mathbf{R} + \mathbf{B}_d^T \mathbf{M} \mathbf{B}_d)^{-1} \mathbf{B}_d^T \mathbf{M}^T \mathbf{A}_d$  din pasul de îmbunătățire a regulatorului, cu expresia de calcul a regulatorului optimal  $\mathbf{K}^* = -(\mathbf{R} + \mathbf{B}_d^T \mathbf{S} \mathbf{B}_d)^{-1} \mathbf{B}_d^T \mathbf{S} \mathbf{A}_d$ . Practic locul matricei simetrice  $\mathbf{S}$  este luat de matricea simetrică  $\mathbf{M}$  care se actualizează iterativ.

### III) Calculul regulatorului liniar optimal după stare folosind algoritmul VI în varianta fără model (model-free).

Pentru proces considerat anterior, va fi exemplificată procedura de calcul a aceluiași regulator optimal liniar după stare, utilizând algoritmul VI în varianta fără model cunoscut al procesului (Radac, 2019).

Este evidentă, atât în cadrul soluției regulatorului LQR bazată pe rezolvarea ecuațiilor algebrice matriceale Riccati cât și în cadrul soluției bazată pe algoritmul Polt, necesitatea cunoașterii matricelor  $\mathbf{A}_d, \mathbf{B}_d$  aferente reprezentării pe stare a procesului condus. Nu doar în ecuația algebrică matriceală (5), dar și în pasul de îmbunătățire al regulatorului când trebuie calculată derivata membrului drept al ecuației Bellman în raport cu comanda. Există însă posibilitatea de a ajunge la regulatorul optimal fără a cunoaște matricele modelului pe stare al procesului.

În acest sens, se introduce FC extinsă notată  $\Theta^K(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k + J^K(\mathbf{x}_{k+1})$  care reprezintă costul aferent utilizării comenzii  $\mathbf{u}_k$  la pasul curent  $k$ , după care întreg orizontul de timp rămas utilizează legea de reglare  $\mathbf{u}_k = \mathbf{K}^T \mathbf{x}_k$ . Practic,  $\Theta^K(\mathbf{x}_k, \mathbf{u}_k = \mathbf{K}^T \mathbf{x}_k) = J^K(\mathbf{x}_k)$  iar FC extinsă respectă de asemenea ecuația Bellman aferentă:

$$\Theta^K(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k + \Theta^K(\mathbf{x}_{k+1}, \mathbf{u}_{k+1} = \mathbf{K}^T \mathbf{x}_{k+1}). \quad (8)$$

Avantajul substanțial al FC extinse este acela că permite calculul direct al unui regulator îmbunătățit prin minimizarea directă a FC extinse:  $\bar{\mathbf{u}}_k = \arg \min_{\mathbf{u}_k} \Theta^K(\mathbf{x}_k, \mathbf{u}_k)$ . FC extinsă optimală este cea care respectă ecuația de optimalitate Bellman

$$\Theta^*(\mathbf{x}_k, \mathbf{u}_k) = \min_{\mathbf{K}} \{\mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k + \Theta^K(\mathbf{x}_{k+1}, \mathbf{u}_{k+1} = \mathbf{K}^T \mathbf{x}_{k+1})\}, \quad (9)$$

din care poate fi dedusă expresia regulatorului optimal  $\mathbf{K}^* = \arg \min_{\mathbf{K}} \Theta^*(\mathbf{x}_k, \mathbf{u}_k)$  precum și valoarea FC optimale originale  $J^*(\mathbf{x}_k) = \Theta^*(\mathbf{x}_k, \mathbf{u}_k = (\mathbf{K}^*)^T \mathbf{x}_k)$ .

Presupunând că pentru un regulator liniar după stare oarecare de forma  $\mathbf{u}_k = \mathbf{K}^T \mathbf{x}_k$ , FC originală este o funcție pătratică de stare de forma  $J^K(\mathbf{x}_k) = \mathbf{x}_k^T \mathbf{M} \mathbf{x}_k$ , se poate arăta, folosind ecuația modelului pe stare al procesului  $\mathbf{x}_{k+1} = \mathbf{A}_d \mathbf{x}_k + \mathbf{B}_d \mathbf{u}_k$  că FC extinsă este de asemenea pătratică, de forma

$$\Theta^K(\mathbf{x}_k, \mathbf{u}_k) = \underbrace{\begin{pmatrix} \mathbf{x}_k^T & \mathbf{u}_k^T \end{pmatrix}}_{\mathbf{z}_k^T} \begin{pmatrix} \mathbf{Q} + \mathbf{A}_d^T \mathbf{M} \mathbf{A}_d & \mathbf{B}_d^T \mathbf{M} \mathbf{A}_d \\ \mathbf{A}_d^T \mathbf{M} \mathbf{B}_d & \mathbf{R} + \mathbf{B}_d^T \mathbf{M} \mathbf{B}_d \end{pmatrix} \begin{pmatrix} \mathbf{x}_k \\ \mathbf{u}_k \end{pmatrix}. \quad (10)$$

Pentru cazul procesului de ordin doi prezentat anterior, FC extinsă poate fi explicitată la forma

$$\Theta^K(\mathbf{x}_k, \mathbf{u}_k) = (x_{1,k}^2 \ x_{2,k}^2 \ u_k^2 \ x_{1,k}x_{2,k} \ x_{1,k}u_k \ x_{2,k}u_k) \cdot (\theta_1 \dots \theta_6)^T = \Phi^T(\mathbf{x}_k, \mathbf{u}_k)\boldsymbol{\theta}, \quad (11)$$

unde este evidentă parametrizarea liniară a FC extinse pătratică în raport cu argumentele, matricea  $\Phi^T(\mathbf{x}_k, \mathbf{u}_k)$  fiind matricea funcțiilor de bază de aproximare. Această parametrizare conduce, în cazul general, la rescrierea ecuației Bellman pentru FC extinsă în forma

$$\Phi^T(\mathbf{x}_k, \mathbf{u}_k)\boldsymbol{\theta}^K = \mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k + \Phi^T(\mathbf{x}_{k+1}, \mathbf{u}_{k+1}) = \mathbf{K}^T \mathbf{x}_{k+1} \boldsymbol{\theta}^K, \quad (12)$$

unde indicele superior  $\mathbf{K}$  din parametrizarea  $\boldsymbol{\theta}^K$  ilustrează faptul că FC extinsă liniar parametrizată este aferentă regulatorului  $\mathbf{K}$ .

Având la dispoziție multiple **tranziții de stare** de forma  $(\mathbf{x}_k, \mathbf{u}_k, \mathbf{x}_{k+1})$  culese de la proces prin interacțiune cu acesta pe baza unor comenzi  $\mathbf{u}_k$  generate de **orice formulă** (inclusiv aleator) și având la dispoziție expresia unui regulator liniar după stare oarecare  $\mathbf{K}$ , se poate recurge la:

- evaluarea FC extinse aferente regulatorului  $\mathbf{K}$  rezolvând în raport cu  $\boldsymbol{\theta}^K$  în sensul celor mai mici pătrate sistemul supradeterminat de ecuații

$$\begin{pmatrix} \Phi^T(\mathbf{x}_1, \mathbf{u}_1) - \Phi^T(\mathbf{x}_2, \mathbf{u}_2 = \mathbf{K}^T \mathbf{x}_2) \\ \dots \\ \Phi^T(\mathbf{x}_N, \mathbf{u}_N) - \Phi^T(\mathbf{x}_{N+1}, \mathbf{u}_{N+1} = \mathbf{K}^T \mathbf{x}_{N+1}) \end{pmatrix} \boldsymbol{\theta}^K = \begin{pmatrix} \mathbf{x}_1^T \mathbf{Q} \mathbf{x}_1 + \mathbf{u}_1^T \mathbf{R} \mathbf{u}_1 \\ \dots \\ \mathbf{x}_N^T \mathbf{Q} \mathbf{x}_N + \mathbf{u}_N^T \mathbf{R} \mathbf{u}_N \end{pmatrix}, \quad (13)$$

care este o abordare specifică algoritmului PoIt, precum și la

- îmbunătățirea estimatei FC extinse aferente regulatorului  $\mathbf{K}_{j-1}$  de la iterația anterioară rezolvând în raport cu  $\boldsymbol{\theta}^{K_j}$  în sensul celor mai mici pătrate sistemul supradeterminat de ecuații

$$\begin{pmatrix} \Phi^T(\mathbf{x}_1, \mathbf{u}_1) \\ \dots \\ \Phi^T(\mathbf{x}_N, \mathbf{u}_N) \end{pmatrix} \boldsymbol{\theta}^{K_j} = \begin{pmatrix} \mathbf{x}_1^T \mathbf{Q} \mathbf{x}_1 + \mathbf{u}_1^T \mathbf{R} \mathbf{u}_1 + \Phi^T(\mathbf{x}_2, \mathbf{u}_2 = \mathbf{K}_{j-1}^T \mathbf{x}_2) \boldsymbol{\theta}^{K_{j-1}} \\ \dots \\ \mathbf{x}_N^T \mathbf{Q} \mathbf{x}_N + \mathbf{u}_N^T \mathbf{R} \mathbf{u}_N + \Phi^T(\mathbf{x}_{N+1}, \mathbf{u}_{N+1} = \mathbf{K}_{j-1}^T \mathbf{x}_{N+1}) \boldsymbol{\theta}^{K_{j-1}} \end{pmatrix}, \quad (14)$$

care este o abordare specifică algoritmului VI.

De remarcat că ambele relații (13), (14) conduc la evaluări ale FC extinse aferente unui regulator, fără a necesita cunoașterea modelului pe stare al procesului. De asemenea, numărul  $N$  al tranzițiilor de stare care construiesc sistemele de ecuații (13), (14) trebuie să fie cel puțin egal cu numărul parametrilor din vectorul  $\boldsymbol{\theta}^K$ , pentru a avea o soluție unică. De asemenea, tranzițiile de stare trebuie să difere pe cât posibil, cerință care implică o explorare bună a spațiului stării și al comenzii.

În continuare, va fi discutat algoritmul VI fără model cunoscut al procesului, care nu necesită un regulator inițial admisibil pentru amorsare, ci doar existența unui set de date care conține tranzițiile de stare culese de la proces. Pasul de îmbunătățire al regulatorului, pas bazat pe o parametrizare  $\boldsymbol{\theta}^{K_j}$  curentă a FC extinse, este ușor ilustrabil pentru FC extinse pătratice, cum sunt cele de forma (11) aferente proceselor liniare cu regulator liniar după stare. Astfel, în cazul procesului particular studiat pentru care este valabilă expresia FC extinse din (11), pasul de îmbunătățire al regulatorului rezultat ca soluție a ecuației  $\partial \Theta^K(\mathbf{x}_k, \mathbf{u}_k) / \partial \mathbf{u}_k = 0$  devine

$$\frac{\partial [(x_{1,k}^2 \ x_{2,k}^2 \ u_k^2 \ x_{1,k}x_{2,k} \ x_{1,k}u_k \ x_{2,k}u_k) \cdot (\theta_1^{K_j} \dots \theta_6^{K_j})^T]}{\partial \mathbf{u}_k} = 0 \Leftrightarrow \bar{\mathbf{u}}_k = - \left( \frac{\theta_5^{K_j}}{2\theta_3^{K_j}} \ \frac{\theta_6^{K_j}}{2\theta_3^{K_j}} \right) \begin{pmatrix} x_{1,k} \\ x_{2,k} \end{pmatrix} = \mathbf{K}_j^T \mathbf{x}_k, \quad (15)$$

de unde rezultă și matricea regulatorului liniar după stare îmbunătățit

$$\mathbf{K}_j = - \left( \frac{\theta_5^{K_j}}{2\theta_3^{K_j}} \ \frac{\theta_6^{K_j}}{2\theta_3^{K_j}} \right)^T. \quad (16)$$

Practic, alternând relațiile (14) și (16), pornind de la un regulator inițial  $\mathbf{K}_0$  oarecare (nu neapărat admisibil) și de la o parametrizare inițială  $\boldsymbol{\theta}^{K_0}$  oarecare a FC extinse, algoritmul VI converge la regulatorul optimal  $\mathbf{K}^*$ .

Pentru exemplul procesului de ordin doi studiat, fie regulatorul inițial  $\mathbf{K}_0 = [1 \ 1]^T$  precum și parametrizarea inițială a FC extinse  $\boldsymbol{\theta}^{\mathbf{K}_0} = [1 \ 1 \ 1 \ 1 \ 1]^T$ . Fie setul de tranziții de stare cules prin interacțiune cu procesul de forma

$$\begin{aligned} \mathbf{x}_1 &= [0.5377 \ 1.8339]^T, u_1 = 0.7254, \mathbf{x}_2 = [0.5810 \ 0.5377]^T, \\ \mathbf{x}_2 &= [-2.2588 \ 0.8622]^T, u_2 = -0.0631, \mathbf{x}_3 = [-3.7106 \ -2.2588]^T, \\ \mathbf{x}_3 &= [0.3188 \ -1.3077]^T, u_3 = 0.7147, \mathbf{x}_4 = [1.8184 \ 0.3188]^T, \\ \mathbf{x}_4 &= [-0.4336 \ 0.3426]^T, u_4 = -0.2050, \mathbf{x}_5 = [-0.9931 \ -0.4336]^T, \\ \mathbf{x}_5 &= [3.5784 \ 2.7694]^T, u_5 = -0.1241, \mathbf{x}_6 = [3.6007 \ 3.5784]^T, \\ \mathbf{x}_6 &= [-1.3499 \ 3.0349]^T, u_6 = 1.4897, \mathbf{x}_7 = [-1.9413 \ -1.3499]^T, \end{aligned} \quad (17)$$

unde tranzițiile de stare nu sunt succesive în timp (deși acest lucru este permis). Prima iterație VI conduce, prin rezolvarea (14) la  $\boldsymbol{\theta}^{\mathbf{K}_1} = [17.2204 \ 1.7398 \ 3.1000 \ -6.7295 \ 13.5516 \ -2.9795]^T$ , iar apelarea (16) conduce la  $\mathbf{K}_1 = [-2.1857 \ 0.4806]^T$ . Apoi, câteva iterații VI conduc asimptotic către regulatorul liniar după stare optimal  $\mathbf{K}^* = [-1.3348 \ 0.4751]^T$ . Se constată convergența practică a algoritmului VI.

Conceptele prezentate mai sus care stau la baza algoritmilor PoIt și VI pot fi extinse relativ ușor și la clase de procese neliniare. În cazul abordărilor fără model cunoscut al procesului condus, problema explorării eficiente (acoperire bună cu un număr cât mai mic de tranziții de stare) a spațiului stării și al comenzii devine critică.

#### Tematica lucrării:

1) Fie procesul în timp continuu redat de ecuația pe stare  $\dot{x}(t) = -2x(t) + 2u(t)$  pentru o perioadă de eșantionare de 0.2 sec, folosind metoda r.i.s.t. Sistemul în timp discret rezultat este  $x_{k+1} = A_d x_k + B_d u_k = 0.6703 x_k + 0.3297 u_k$ .

a) Să se calculeze regulatorul liniar după stare optimal (precum și valoarea FC aferente acestuia) care din orice stare inițială minimizează FC de orizont infinit de timp

$J(x_k) = \sum_{i=k}^{\infty} x_i^T Q x_i + u_i^T R u_i, Q = 1, R = 0.1$ , folosind algoritmul PoIt în varianta bazată pe model. Să se simuleze răspunsul SRA cu regulatorul optimal calculat, pornind de la condiția inițială  $x_0 = 1$ .

b) Folosind modelul pe stare al procesului, generați un număr suficient de tranziții de stare care să permită calculul regulatorului optimal care minimizează aceeași FC, folosind algoritmul VI fără model, pornind de la valori inițiale arbitrare ale regulatorului inițial  $\mathbf{K}_0$  și ale parametrizării  $\boldsymbol{\theta}^{\mathbf{K}_0}$  aferente FC extinse.

2) Fie procesul condus în timp continuu redat prin schema bloc informațională de mai jos, cu valorile parametrilor  $R_a = 4.2 \ \Omega$ ,  $L_a = 14.28 \text{ mH}$ ,  $k_e = k_m = 0.33$ ,  $J = 0.71 \times 10^{-3} \text{ kg} \cdot \text{m}^2$ ,  $k_p = 35 \text{ V/V}$ .

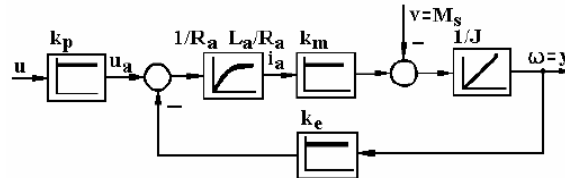


Fig. 1. Procesul condus. (Preitl et al., 2009)

a) Pentru un model discretizat pe stare la o perioadă de eșantionare de 0.01 sec, în condiții de perturbație nulă, să se calculeze regulatorul liniar după stare optimal (precum și valoarea FC aferente acestuia) care din orice stare inițială minimizează FC de orizont infinit de timp

$J(\mathbf{x}_k) = \sum_{i=k}^{\infty} \mathbf{x}_i^T \mathbf{Q} \mathbf{x}_i + u_i^T \mathbf{R} u_i, \mathbf{Q} = \mathbf{I}_2, \mathbf{R} = 0.1$ , folosind algoritmul PoIt în varianta bazată pe model.

b) Folosind modelul pe stare al procesului, generați un număr suficient de tranziții de stare care să permită calculul regulatorului optimal care minimizează aceeași FC, folosind algoritmul VI fără model cunoscut al procesului, pornind de la valori inițiale arbitrare ale regulatorului inițial  $\mathbf{K}_0$  și ale parametrizării  $\boldsymbol{\theta}^{K_0}$  aferente FC extinse.

*Observație:* Reglatoarele după stare rezultate vor fi testate prin simulare. Sunt permise două variante de testare în buclă închisă: atât reglare la zero, pornind din condiții inițiale pe stare nenule la alegere, cât și prin introducerea unei mărimi de referință. În cel de-al doilea caz, regulatorul după stare va fi testat în condițiile unei intrări de referință de tip treaptă care comută din 0 în 1 la momentul de tip 0 sec, și a unei intrări de perturbație de tip treaptă de amplitudine 0.2 care acționează la 0.6 sec. Vor fi notate observațiile privind comportarea sistemului de reglare.

### **Bibliografie**

Radac, M.-B., Învățare prin recompensă. Note de curs, Universitatea Politehnica Timișoara, 2019.

Preitl, St., Precup, R.-E., Preitl, Zs.: Structuri și algoritmi pentru conducerea automată a proceselor, vol. 2, Editura Orizonturi Universitare, Timișoara, 2009.

Soluție. Pentru cerința 1, predictorii vor fi calculați în cadrul unui script Matlab care în prealabil discretizează funcția de transfer a procesului de la intrare la ieșire pentru o perioadă de eșantionare de 0.01 sec și o aduce la forma ARX cunoscută:  $A(q^{-1})y(k) = B(q^{-1})u(k-1)$ . Calculul fiecărui predictor se face rezolvând ecuațiile diofantice aferente. Fie acest script redat mai jos:

```

1 COD 1
2 % -----date_init.m----- %
3 kp=35;Ra=4.2;La=14.28e-3;ke=0.33;km=0.33;J=0.71e-3;
4 KP=kp/ke;radacini=roots([J*La/(km*ke) J*Ra/(km*ke) 1]);
5 T1=-1/radacini(2);T2=-1/radacini(1);
6
7 %-----%
```

### Definiții stabilizabilitate

1. Theorem The real matrix  $A$  is a stability matrix if and only if for any given real symmetric positive definite (r.s.p.d.) matrix  $Q$  the solution  $P$  of the continuous Lyapunov matrix equation (18) is also positive definite.
2. The linear time invariant control system (19) is stabilizable if there exists an  $m \times n$  matrix  $K$  such that  $A+BK$  is a stability matrix.

Perechea  $(A,B)$  caracteristică sistemului liniar invariant în timp este stabilizabilă dacă există un regulator liniar după stare  $u_k = Kx_k$  care, pentru sistemul de reglare în buclă închisă  $A + BK$  face ca  $\lim_{k \rightarrow \infty} \|x_k\| = 0$ .

Controlabilitate perechii  $(A,B)$  implică stabilizabilitatea și poate fi verificată direct în Matlab folosind funcția *ctrb*.

### **Bibliografie.**

- [1] L. Wang, *Model Predictive Control System Design and Implementation Using MATLAB®*, Springer, 2009.
- [2] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press , 2004.
- [3] C. L. Byrne, *A First Course in Optimization*, University of Massachusetts Lowell.
- [4] D. G. Luenberger, *Optimization By Vector Space Methods*, John Wiley & Sons, Inc, 1969.
- [5] G. A. Terejanu, *Extended Kalman Filter Tutorial*, University at Buffalo, Buffalo, NY 14260.
- [6] G. Welch, G. Bishop, *An Introduction to the Kalman Filter*, University of North Carolina at Chapel Hill, 2001.
- [7] D.-A. Duțescu, *Controlul poziției pentru un sistem aerodinamic neliniar cu două rotoare folosind reglare predictivă bazată pe model*, Lucrare de licență, Universitatea Politehnica Timișoara, 2016.