

STRATEGII DE ÎNVĂȚARE MULTIMODEL PENTRU CONDUCEREA PROCESELOR NELINIARE

Conf. dr. ing. Nicolae Constantin

Universitatea Politehnica București
email: nicu @ lnx.cib.pub.ro.

Rezumat: Această lucrare prezintă metodologii locale pentru modelarea și conducerea sistemelor dinamice, în condițiile în care se dispune de o cantitate redusă de date intrare-ieșire din proces. Pentru selectarea celei mai bune configurații a modelului, se utilizează o tehnică bazată pe memorare locală. Identificarea modelului și validarea acestuia se face printr-o tehnică recursivă, împreună cu o metodă statistică, în vederea selectării celei mai bune variante. Conducerea în zonele de operare, aflate departe de punctele de echilibru, este realizată prin utilizarea combinată a tehnicilor de liniarizare locală și a teoriei conducerii optime. Studiul de caz pentru aprecierea performanțelor metodei propuse se referă la aspecte de conducere, în cadrul unui bioreactor.

Cuvinte cheie: algoritmi de învățare, modelare locală, cross-validare, conducere optimală.

1. Introducere

Problema modelării proceselor, utilizând o cantitate limitată de date, obținute din observații directe asupra proceselor, constituie obiectul unor domenii de cercetare, pornind de la regresie neliniară și ajungând până la identificarea sistemelor și „machine learning”.

În literatura dedicată acestui subiect, se disting două direcții principale, care utilizează metode bazate pe memorare locală, respectiv globală. Modelarea globală construiește, din setul de date disponibile din proces, un singur model funcțional. Această categorie de metode a fost utilizată în modelarea pe bază de rețele neurale și alte tehnici de regresie statistică neliniară.

Setul de date disponibil, corespunzând, în general, măsurătorilor asupra evoluției procesului, este utilizat în cadrul unui algoritm de învățare pentru a determina un model care să reflecte fidel corespondența intrare-ieșire. Apoi, se renunță la setul de date și este păstrat doar modelul obținut.

Algoritmii pe bază de memorare locală procesează setul de date, până la obținerea informației dorite pentru predicție sau modelare locală. Abordarea clasică a modelării locale este reprezentată de metoda celui mai apropiat vecin – „nearest neighbor method” [2]. În cadrul acestei abordări, se construiește o bază de date, formată din măsurătorile asupra intrărilor și ieșirilor din proces, și se încearcă, prin interpolare, obținerea unei estimări a unui nou punct de funcționare, utilizând o vecinătate a punctului curent de operare.

O nouă caracteristică a abordării locale o reprezintă adoptarea unor proceduri statistice pentru identificarea aproximării locale. Un exemplu în acest sens îl constituie pachetul statistic PRESS, care permite implementarea eficientă a metodelor de validare de tip „leave-one-out” - LOO [7] și care permite evaluarea performanțelor de generalizare a modelelor liniare locale.

Paradigma „multimodel” sau modele multiple descrie o arhitectură constituită din câteva submodele împreună cu un mecanism care să combine ieșirile acestor submodele. Să notăm cu y_i ieșirea corespunzătoare modelului i . Valabilitatea relativă a submodelului i se poate exprima sub forma unei funcții de partiționare sau de interpolare g_i , unde $g_i \in [0, 1]$.

Dacă submodelul i este foarte precis pentru o anumită zonă din domeniul de operare, atunci valoarea funcției g_i va fi apropiată de unitate [8]. Dacă se consideră că modelul este compus din N submodele, ieșirea modelului poate fi scrisă sub forma:

$$Y = F[(y_1, g_1), (y_2, g_2), \dots, (y_N, g_N)]. \quad (1)$$

Deși există instrumente pentru aproximarea unei funcții neliniare cu o suficient de bună precizie, ținând seama de existența unor regiuni diferite de operare, abordarea multimodel se poate dovedi benefică în următoarele condiții [8]:

- din datele intrare-ieșire din proces, este dificilă construirea precisă a unui model global;
- în cadrul modelării, pot fi combinate diferite tehnici de modelare;
- în cadrul modelării, pot fi utilizate simultan modele liniare sau neliniare;
- cunoștințele apriori pot exista doar pentru anumite condiții de operare, situație în care vor putea fi utilizate pentru descrierea comportării sistemului în regiunile respective cu ajutorul unor modele mecaniciste.

Procedeul de combinare a unor modele diferite, având ca obiectiv descrierea unui singur proces, poate fi denumit modelare hibridă – „hybrid modelling”. Structura modelului poate fi interpretată atât în mod calitativ în termenii regiunilor de operare, cât și cantitativ, în termenii modelelor individuale. În cadrul unor regiuni diferite de operare, mărimile de intrare, care sunt relevante pentru identificarea procesului, pot fi și ele diferite. Abordarea multimodel face posibilă utilizarea unor intrări diferite pentru fiecare domeniu de operare. Această proprietate se dovedește cu adevărat utilă în cadrul proceselor multivariabile, unde importanța variabilelor pentru identificare se modifică de la o regiune de operare la alta.

Implementarea strategiilor de conducere multimodel, în cadrul unor structuri predictive de conducere, este recomandată datorită numărului redus de parametri ai modelului. Adaptarea on-line a submodelelor nu necesită o condiție globală de excitație persistentă.

Compromisul „bias-variance” reprezintă un aspect important pentru strategia de tip «divide and conquer». În general, împărțirea setului de date în subseturi pentru identificarea procesului conduce la creșterea varianței și descreșterea polarizării. Există câteva modalități de a controla aceasta strategie. În primul rând, diferitele modele pot fi construite astfel încât precizia de aproximare și complexitatea modelului să fie diferite. În al doilea rând, utilizarea împărțirilor ‘soft’ a datelor ușurează dilema polarizare - varianță. Spre deosebire de partițiile ‘hard’ cele ‘soft’ sau ‘fuzzy’ permit modelului dintr-o regiune să fie influențat de datele din regiunile vecine.

Complexitatea modelului utilizat într-o regiune poate crește o dată cu creșterea numărului de date din aceea regiune. Acest fapt reduce polarizarea și, în același timp, varianța deoarece fiecare model este construit potrivit datelor din regiunea de operare corespunzătoare.

În al treilea rând, pentru aproximarea submulțimilor, pot fi utilizate funcții constante sau liniare, pe porțiuni. Aceste funcții permit o varianță mai redusă în detrimentul unei polarizări ridicate.

În cadrul acestei lucrări, este propusă o metodologie de identificare, bazată pe o procedură de optimizare, în vederea selectării celui mai performant model local, dintr-o mulțime de variante posibile. În procedura de modelare locală convențională, trebuiau făcute o serie de opțiuni de către analistul de date, conform unor criterii euristice și a unor presupuneri apriori.

În lucrare, este propusă o procedură de selectare automată, care să caute configurația optimă a modelului, pe baza analizei proprietăților statistice ale fiecărui model. Se propune un nou algoritm pentru estimarea, în mod recursiv, a performanțelor modelului, printr-o metodă de „cross-validare”. De asemenea, se prezintă și o tehnică incrementală de calcul statistic, pe baza metodei celor mai mici pătrate, în variantă recursivă. În plus, pentru compararea performanțelor a două modele candidate, se utilizează un test statistic performant, pe baza erorilor de „cross-validare”.

În lucrare, se utilizează tehnici de conducere optimală, parametrizate cu valorile returnate de estimatori locali, de tip liniar. Se propune o structură de conducere, cu lege de reglare de tip predictiv, care întoarce la fiecare moment de eșantionare prima componentă a unei secvențe optimale, calculată printr-o procedură de optimizare, pe baza de gradient. Problema de optimizare este formulată ca o minimizare a unei funcții de cost returnată în urma unei simulări înainte a modelului identificat, iar gradientul este calculat pe baza ecuației Riccati discrete, pentru sisteme liniare variante în timp.

Lucrarea are următoarea organizare. În secțiunea a 2-a, se prezintă tehnici de modelare pe baza unor proceduri iterative de selecție. Problematika privind tehnicile de conducere predictivă este prezentată în secțiunea a 3-a. În secțiunea a 4-a, se prezintă rezultatele simulării unui proces de conducere.

2. Identificarea procesului utilizând modele locale

Identificarea modelului unui proces, pe baza determinărilor experimentale, solicită o bogată experiență în utilizarea tehnicilor de învățare din partea analistului. În numeroase situații, setul de date disponibile din proces este limitat și, în același timp, se solicită o predicție precisă a evoluțiilor viitoare. Adesea, informația referitoare la ordinul modelului, la structura acestuia, la tipul variabilelor relevante, nu este disponibilă. Procesul de învățare constă dintr-o procedură „trial- and-error”, pe durata căreia modelul este ajustat la setul de date disponibil. Estimarea valorii necunoscute a funcției este efectuată prin focalizarea atenției asupra vecinătății punctului în care este solicitată estimarea.

Se consideră o mappare necunoscută $f: \mathcal{R}^m \rightarrow \mathcal{R}$ ca un set de N date $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$. Aceste date pot fi structurate într-o matrice X de dimensiune $[N \times m]$, și într-un vector y de dimensiune $[N \times 1]$. Fiind dat un punct curent x_q , în continuare se prezintă modul de obținere a predicției valorii $y_q = f(x_q)$. În primul rând, pentru fiecare cuplu (x_i, y_i) , se calculează o pondere w_i în funcție de distanța $d(x_i, x_q)$ de la punctul curent

x_q la punctul x_i . Apoi fiecare linie din X și respectiv y este înmulțită cu ponderea corespunzătoare rezultând variabilele $Z = WX$ și $v = Wy$, cu W matrice diagonală, având elementele de pe diagonală egale cu $W_{ii} = w_i$. În final se calculează un model de regresie ponderat prin rezolvarea ecuației $(Z^T Z) u = Z^T v$.

Predicția valorii $f(x_q)$ se obține prin evaluarea modelului în punctul curent

$$\hat{y} = x_q^T (Z^T Z)^{-1} Z^T v \quad (2)$$

În mod uzual, în cazul în care se adoptă maniera de modelare locală, trebuie luate o serie de decizii privind modelul (numărul de vecini, funcția de ponderare, familia parametrică, criteriul pentru estimarea parametrilor).

Abordarea clasică este extinsă cu o metodă din domeniul analizei statistice liniare, ce selectează automat configurația potrivită. Pe baza performanțelor fiecărui model, acestea pot fi comparate în vederea alegerii celei mai bune variante. Etapele generale ale acestei abordări pot fi rezumate astfel:

1. sarcina învățării unei corespondențe intrare-ieșire este descompusă într-o serie de probleme de estimare liniară;
2. fiecare problemă de estimare este tratată ca o problemă de optimizare, în spațiul modelelor alternative;
3. capacitatea de estimare a fiecărui model candidat este evaluată cu ajutorul performanțelor de „cross-validare”, calculate pe baza metodelor statistice.

Pentru a îmbunătăți eficiența acestor etape, se propun în cadrul metodei de învățare doi algoritmi:

- algoritm recursiv, pentru estimarea parametrică și „cross-validarea” fiecărui model local; această metodă evită reluarea evaluării fiecărui model de la zero și reduce semnificativ efortul de calcul;
- test statistic pentru a compara performanțele a două modele alternative; testul statistic nu consideră doar valorile medii ale erorilor de „cross - validare”, ci și distribuțiile acestora.

Prezentarea va fi orientată către procedura de „cross - validare” de tip „leave-one-out” - LOO [7] - care calculează o estimare a erorii de predicție în x_q .

Se definește eroarea de „cross-validare” de tip „leave-one-out”, corespunzătoare modelului i , $e^{cv}(i)$ ca diferență între y_i și predicția dată de modelul local, centrat în x_q și utilizând toate datele disponibile cu excepția (x_i, y_i) .

O estimare a erorii de predicție în x_q se poate obține prin medierea erorilor $e^{cv}(i)$, fiecare fiind ponderată corespunzător distanței $d(x_i, x_q)$.

În cazul considerării modelului local liniar, „cross - validarea” de tip LOO - „mean square error cross-validation” - poate fi efectuată fără a recalcula parametrii de regresie pentru fiecare exemplu exclus, utilizând versiunea locală a statisticii PRESS [1]:

$$MSE^{cv}(x_q) = \frac{1}{\sum_i w_i^2} \sum_i \left(\frac{v_i - z_i^T (Z^T Z)^{-1} Z^T v}{1 - z_i^T (Z^T Z)^{-1} z_i} \right)^2 = \frac{1}{\sum_i w_i^2} \sum_i [w_i e^{cv}(i)]^2 \quad (3)$$

Unul din cei mai importanți parametri ce trebuie ajustați într-o configurație locală este dimensiunea din vecinătatea x_q , în care funcția $f(\cdot)$ poate fi aproximată printr-un model local liniar.

Un astfel de parametru poate fi pus în legătură cu numărul de exemple de antrenare, care sunt din regiunea de antrenare. Alegerea numărului de vecini n din jurul punctului curent x_q pentru a fi utilizați în regresia locală este funcție de compromisul polarizare /varianță.

Adăugarea unor date suplimentare conduce la reducerea varianței și la creșterea polarizării. Pe de altă parte, reducerea numărului de puncte considerate în cadrul vecinătății punctului curent poate reduce polarizarea în detrimentul unei varianțe crescute a estimatorului. Astfel, pot fi considerate modele de complexitate diferită, fiecare utilizând un număr diferit de măsurători, iar procedura de „cross-validare leave-one-out” poate fi utilizată pentru comparare și selectarea celor pentru care eroarea de predicție este mai redusă. Este adoptată o metodă incrementală, bazată pe tehnici liniare recursive.

Această metodă este utilizată pentru obținerea parametrilor modelului construit pentru $n1$ dintre cei mai apropiați vecini, prin actualizarea parametrilor modelului bazat pe n vecini. Algoritmul recursiv are următoarea formă:

$$P_{n+1} = P_n - \frac{P_n x_{n+1} x_{n+1}^T P_n}{1 + x_{n+1}^T P_n x_{n+1}} \quad (4)$$

$$e_{n+1} = y_{n+1} - x_{n+1}^T \theta_n$$

$$\theta_{n+1} = \theta_n + P_{n+1} x_{n+1} e_{n+1}$$

Astfel, se obține valoarea erorii LOO sub forma:

$$e_{n+1}^{cv}(i) = \frac{y_i - x_i^T \theta_{n+1}}{1 - x_i^T P_{n+1} x_i} \quad (5)$$

unde (x_{n+1}, y_{n+1}) este al $(n+1)$ -lea vecin cel mai apropiat de punctul curent, P_n reprezintă aproximarea recursivă a matricii $(Z^T Z)^{-1}$, θ_n reprezintă parametrii modelului considerând n cei mai apropiați vecini, iar valorile $e_n^{cv}(i), 1 \leq i \leq n$ compun vectorul En al erorilor „leave-one-out”.

3. Proiectarea legii de comandă optimală

Pentru problema de conducere, se consideră reprezentarea generală în spațiul stărilor a unui sistem discret neliniar:

$$x(k+1) = g[x(k), u(k)] + v(k) \quad (6)$$

$$y(k) = h[x(k)] = w(k)$$

unde x reprezintă vectorul de stare, v, w secvențele de perturbație respectiv zgomot, g, f - funcții neliniare. În ipoteza îndeplinirii condițiilor de controlabilitate și observabilitate, reprezentarea pe stare a sistemului în buclă poate fi scrisă sub forma:

$$x(k+1) = F[\varphi(k)]x(k) + v(k) \quad (7)$$

Considerăm problema de control optimal pentru sistemul discret neliniar (6) peste un orizont finit de timp. Funcția de cost de tip pătratic se consideră scrisă sub forma:

$$J = \frac{1}{2} x(t_f)^T P_f x(t_f) + \frac{1}{2} \sum_{k=0}^{t_f} [x(k)^T u(k)^T \begin{bmatrix} Q_k & M_k \\ M_k^T & R_k \end{bmatrix} \begin{bmatrix} x(k) \\ u(k) \end{bmatrix}] \quad (8)$$

în care Q_k, M_k, R_k, P_f reprezintă termeni de ponderare definiți apriori.

În continuare, se consideră traiectoria stărilor unui sistem dinamic la intrarea căruia se aplică secvența de comandă $U = [u(0), u(1), \dots, u(t_f-1)]$. Se presupune că sistemul poate fi liniarizat în jurul fiecărei stări. Dacă sunt neglijate erorile reziduale, datorate aproximării în serie Taylor de ordinul întâi, comportarea sistemului liniar, de-a lungul unui traiectorii generice este similară cu cea a unui sistem liniar variant în timp:

$$\begin{aligned} x(k+1) &= F[\varphi(k)]x(k) + G[\varphi(k)]u(k) + K[\varphi(k)] = \\ &= F_k x(k) + G_k u(k) + K_k \end{aligned} \quad (9)$$

unde F_k, G_k, K_k reprezintă parametrii sistemului liniarizat în jurul punctului curent.

Soluția pentru sistemul liniar variant în timp se obține sub forma secvențelor de comandă optimală, la fiecare pas de eșantionare:

$$u(k) = -(R_k + G_k^T P_{k+1} G_k)^{-1} (M_k^T + G_k^T P_{k+1} F_k) x(k) \quad (10)$$

unde P_k reprezintă soluția ecuației „backward” Riccati:

$$\begin{aligned} P_k &= Q_k + F_k^T P_{k+1} F_k - (M_k + F_k^T P_{k+1} G_k) \\ &\quad \times (R_k + G_k^T P_{k+1} G_k)^{-1} (M_k^T + G_k^T P_{k+1} F_k) \end{aligned} \quad (11)$$

având condiția finală $P(t_f) = P_f$.

Soluția optimală se obține prin rezolvarea ecuațiilor Euler-Lagrange, condiții necesare și suficiente pentru optimalitate considerând fixată valoarea finală a timpului.

$$0 = \frac{\partial H_k}{\partial u_k} = x_k^T M_k + u_k^T R_k + \lambda_{k+1}^T G_k \quad (12)$$

$$\lambda_k^T = \frac{\partial H_k}{\partial x_k} = x_k^T Q_k + u_k^T M_k^T + \lambda_{k+1}^T F_k \quad (13)$$

$$\lambda_f^T = x_f^T P_f \quad (14)$$

cu funcția Hamiltonian de forma:

$$H_k = J + \lambda_{k+1}^T [F_k x(k) + G_k u(k)] \quad (15)$$

Matricile F_k și G_k pot fi obținute, la fiecare pas, prin liniarizarea dinamicii sistemului de-a lungul traiectoriei, considerând la intrare secvența de comandă. Algoritmul pentru conducerea optimală neliniară este formulat ca o problemă de optimizare pe bază de gradient și se bazează pe liniarizarea locală a sistemului.

Algoritmul caută secvența de intrare:

$$U^{opt} = \arg \min_{U^i} J(U^i) \quad (16)$$

care minimizează funcția de cost, definită pe un orizont finit de-a lungul a t_f pași.

Funcția de cost $J(U^i)$ pentru secvența generică U^i este calculată prin simularea, pentru t_f pași, a modelului obținut prin metoda de identificare locală. Gradientul funcției de cost, în raport cu secvența U^i , se obține din (10).

Etaplele acestei proceduri de optimizare sunt următoarele:

- simularea înainte a modelului, având aplicată la intrare secvența de comandă U^i ;
- liniarizarea sistemului simulat de-a lungul traiectoriei rezultate;
- evaluarea funcției de cost rezultate $J(U^i)$;
- calculul gradientului funcției de cost, în raport cu secvența simulată;
- actualizarea secvenței de comandă, pe baza algoritmului de gradient.

După aplicarea algoritmului de căutare și obținerea secvenței optime U^{opt} , numai prima componentă a acestei secvențe este aplicată către proces (strategia de conducere pe orizont îndepărtat – „receding horizon control strategy” [5]. Modelul obținut reprezintă un estimator care predictează comportarea sistemului supus acțiunii secvenței de intrare generice, și întoarce o aproximare liniară a dinamicii sistemului.

Prin procedeul de învățare, este posibilă liniarizarea sistemului în puncte aflate la distanță mare de echilibru. Abordarea sistemului variant face posibilă utilizarea unei strategii de conducere liniară chiar dacă sistemul operează în cadrul unor regimuri liniare diferite.

Pentru sistemele liniare variabile în timp, încă nu sunt demonstrate proprietățile de stabilitate pentru secvența de conducere optimală. Acest formalism are avantajul că poate fi extins cu ușurință la sisteme multivariabile. În același timp, permite evaluarea secvenței de comandă, considerând un orizont extins. Acest formalism poate ține seama și de incertitudinile care afectează modelul. În cadrul proiectării legii de comandă, se presupune că parametrii obținuți în cadrul modelelor locale reprezintă descrierea reală a comportării locale (principiul echivalenței sigure – „certainty equivalence principle”).

Totuși, aceasta reprezintă o presupunere restrictivă, ce necesită un grad de precizie suficient de mare al aproximării. Teoria conducerii optimale poate furniza o soluție pentru această limitare. De fapt, teoria conducerii optimale stocastice oferă o soluție formală problemei, în cazul considerării incertitudinilor parametrice. În plus, în cadrul procedurilor de modelare pot fi calculate descrieri statistice ale parametrilor estimați.

4. Rezultate simulări. Studiu de caz

În cadrul studiului de caz, se consideră un bioreactor, și anume, o problemă test, întâlnită adesea în conducerea neliniară a proceselor [3,9]. Bioreactorul reprezintă un rezervor ce conține apă, nutrienți și celule

biologice. Nutrienții și celulele sunt introduse în rezervor și supuse unui proces de amestecare uniformă. Starea sistemului este caracterizată de numărul de celule x_1 și cantitatea de nutrienți x_2 . Ecuatiile ce guvernează desfășurarea procesului în cadrul bioreactorului pot fi scrise sub forma:

$$\begin{aligned}\dot{x}_1 &= -x_1 u + x_1(1-x_2)e^{x_2/\gamma} \\ \dot{x}_2 &= -x_2 u + x_1(1-x_2)e^{x_2/\gamma} \left(\frac{1+\beta}{1+\beta-x_2} \right)\end{aligned}\quad (17)$$

unde valorile parametrilor sunt $\beta = 0.02$ și $\gamma = 0.48$. Obiectivul este de a stabili sistemul în jurul stării instabile $(x_1, x_2) = (0.2107, 0.726)$ prin aplicarea unui semnal de comandă la fiecare 0.5 secunde.

Orizontul algoritmului descris în secțiunea anterioară este ales $t_f = 5$. Condițiile inițiale de operare sunt stabilite printr-o procedură de inițializare cu valori aleatoare. Considerând aplicarea unei secvențe de intrare, aleasă în mod aleator, se construiește o bază de date, formată din 1000 de determinări. Actualizarea acestei baze de date se face on-line, pe măsură ce o nouă pereche intrare-ieșire este calculată cu ajutorul sistemului simulat. În figura 1a, se prezintă ieșirea celor două variabile de stare controlate, iar în figura 1b se arată evoluția semnalului de comandă.

Dificultatea problemelor de conducere a bioreactorului se datorează neliniarităților prezente, precum și faptului că variații reduse ale parametrilor pot conduce la o evoluție instabilă. Rezultatele simulărilor arată că este posibilă conducerea unor sisteme complexe pe un domeniu extins de operare, utilizând o cantitate redusă de date și cunoștințe reduse asupra dinamicii procesului.

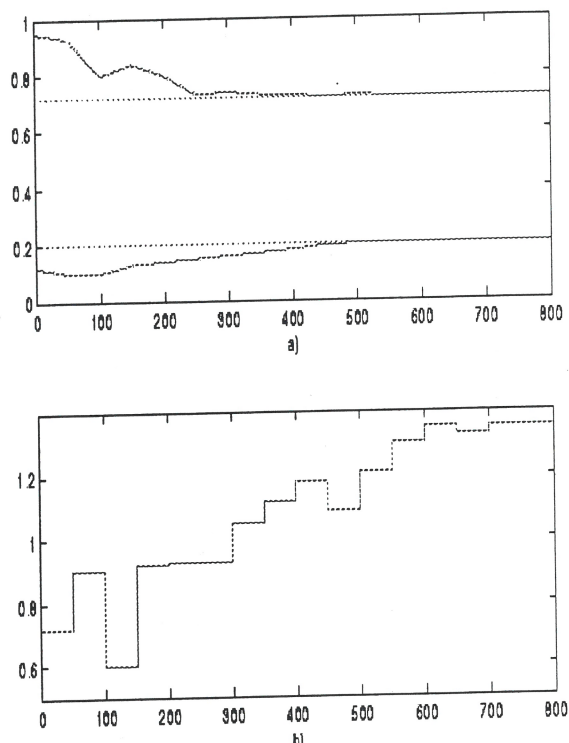


Figura 1. Evoluția stărilor (a) și evoluția comenzii (b).

5. Concluzii

Tehnicile bazate pe memorare locală reprezintă instrumente eficiente, utilizate în metodele de învățare cu ajutorul unei cantități reduse de date și observarea comportării locale a sistemelor neliniare.

În această lucrare, se propune un nou algoritm pentru îmbunătățirea performanțelor acestor tehnici bazate pe memorare locală. Acest algoritm se bazează pe o versiune recursivă a procedurii de „cross-validare” și pe selecția modelului pe baza unor metode statistice.

În cadrul proiectării regulatorului, tehnicile de comandă optimală au fost parametrizate cu valorile calculate de estimatorul liniar. Metoda propusă poate fi aplicată fără modificări majore și pentru cazul sistemelor multivariabile.

Bibliografie

1. **ATKESON C.G., A.W. MOORE, S. SCHAAL:** Locally Weighted Learning. În: Artificial Intelligence Review, 11(1-5), 1997, pp. 11-73.
2. **ATKESON C.G., MOORE A.W. S. SCHAAL:** Locally Weighted Learning for Control. În: Artificial Intelligence Review, 11(1-5), 1997, pp. 75-113.
3. **BERSINI H., V. GORRINI:** A Simplification of the Back-Propagation-Through-Time Algorithm for Optimal Neurocontrol. În: IEEE Trans. on Neural Networks, 8(3), 1996, pp. 437-441.
4. **BOTTOU L., V.N. VAPNIK:** Local Learning Algorithms. În: Neural Computation, 4(6), 1992, pp. 888-900.
5. **CLARKE D.W.:** Advances in Model-Based Predictive Control. Oxford University Press., 1994.
6. **CONSTANTIN, N., I. DUMITRACHE:** Model Predictive Control Based on Multimodel Estimation. În: Proc. of Int. Conf. on Control Systems and Computer Science, 2003, pp. 354-359.
7. **EFRON B., R.J. TIBSHIRANI:** An Introduction to the Bootstrap. New York: Chapman and Hall., 1993.
8. **JOHANSEN T.A. AND FOSS B.A.** Constructing NARMAX Models Using ARMAX Models. În: International Journal of Control, 58, 1993, pp. 1125-1153.
9. **MILLER W.T., R.S. SUTTON, P.J. WERBOS (eds):** Neural Networks for Control, The MIT Press., 1990.
10. **MOORE A.W., J. SCHNEIDER, K. DENG:** Efficient Locally Weighted Polynomial Regression Predictions. În: Proc. of the 1997 International Machine Learning Conference, Morgan Kaufmann Publishers, 1997.
11. **MURRAY-SMITH R., K. HUNT:** Local Model Architectures for Nonlinear Modelling and Control. Pages 61-82 of: Hunt K.J., Irwin G.R. & Warwick K. (eds), Neural Network Engineering in Dynamic Control Systems, Springer Verlag, 1995.
12. **MURRAY-SMITH, R., T. A. JOHANSEN:** Multiple Model Approaches to Modelling and Control, Taylor and Francis, 1997.
13. **NARENDRA K.S., J. BALAKRISHNAN:** Adaptive Control Using Multiple Models. În: IEEE Trans. on Automatic Control, 42(2), 1997, pp. 171-187.
14. **SCHAAL S., C. G. ATKESON:** Robot Juggling: Implementation of Memory-Based Learning. În: IEEE Control Systems, February, 1994, pp. 57-71.
15. **SCHOTT K.D., B. W. BEQUETTE:** Multiple Model Adaptive Control, Murray-Smith R. & Johansen T.A. (eds), Taylor and Francis, pp. 269-291.
16. **TANAKA K., M. SUGENO:** Stability Analysis and Design of Fuzzy Control Systems. În: Fuzzy Sets and Systems, 45, 1992, pp. 135-156.