Nové metody a postupy v oblasti přístrojové techniky, automatického řízení a informatiky
2010
Ústav přístrojové a řídicí techniky ČVUT v Praze, Odbor přesné mechaniky a optiky
Technická 4, 166 07 Praha 6

TREND KVADRATICKÉHO ADAPTIVNÍHO MODELU PRO AUTOMATICKÉ ŘÍZENÍ

Trend of Quadratic Adaptive Model for Automatic Control

Ivo Bukovský

Abstrakt: Článek vyzdvihuje zajímavé a přitom známé konvexní vlastnosti kvadratického polynomu, jako důležitého mezičlánku mezi lineárními modely a silně nelineárními modely jako jsou neuronové sítě. Charakter jediného řešení a přitom silné nelineární aproximační schopnosti jsou vhodné pro aplikace v automatickém řízení.

Key words:Kvadratický polynom, kvadratické neuronová jednotka, konvexní chybové kriterium, globální řešení, adaptivní modelování, adaptivní řízení.

1. Úvod

V průmyslových realizacích je kladen důraz na spolehlivost navržených modelovacích a řídících algoritmů a průhlednost (řešitelnost) těchto algoritmů je stále rozhodujícím faktorem při volbě řešení a realizace. Lineární modely a regulátory (analyticky odvozované či numericky optimalizované) jsou proto stále běžnou formou řešení a implementace. Existence jediného řešení je obrovskou výhodou lineárních systémů a umožňuje dobře analyticky řešit úlohy či dobře optimalizovat lineární modely, které potom mohou jednoznačně konvergovat ke globálnímu minimu.

Nelineární modely a algoritmy nelineárních regulátorů jsou dnes díky počítačové podpoře středem zájmu odborných publikací. I když jsou nelineární metody modelování a řízení v odborné literatuře velmi četné, nelze si nepovšimnout často spíše teoretického charakteru řady takovýchto prací (řízená soustava je model, výsledky matematického odvození jsou ověřeny jen simulačně, a podobně). Nevyzpytatelnost silně nelineárních systémů (existence lokálních minim (řešení), citlivost na správnost odhadu nelineárního modelu, nastavení parametrů, a pod.) a relativně velká matematická náročnost odborných článků v kvalitních publikacích dnes zabraňují inženýrům běžně implementovat nejnovější postupy.

Nelineární algoritmy s dnes již neodmyslitelnou počítačovou podporou patří do oblasti, která se dnes často označuje termínem „computational intelligence“(CI). Do této oblasti spadají také umělé neuronové sítě, fuzzy logika, genetické algoritmy a řada dalších (kvadratické programování, swarm inteligence, imune-based systems...). Nové metody automatického řízení by měly být významným ale i rozumným aplikačním vyustěním nově vyvíjených algoritmů CI.

Kvadratický polynom je dávno znám například z Taylorova rozvoje druhého řádu v aproximacích nelineárních systémů. V tomto článku je ukázáno, že kvadratický polynom má konvexní vlastnosti kombinující výhody jediného řešení lineárních systémů a zdatnost v aproximovaci nelineárních systémů. Článek upozorňuje na skutečnost, že kvadratický polynom zasluhuje pozornost jako důležitý mezičlánek mezi omezenými lineárními systémy a prakticky nepříliš vhodnými nelineárními systémy a při tom jeho struktura i implementační náročnost je nižší v porovnání s např. neuronovými sítěmi.

2. Kvadratický polynom v oboru umělých neuronových sítí

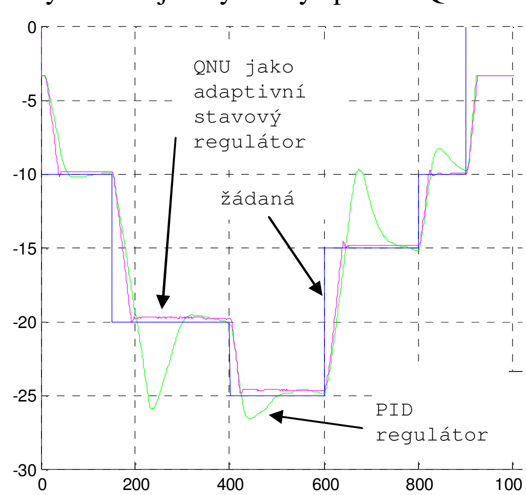
Z hlediska odborné literatury zapadá kvadratický polynom v oboru neuronových sítí do dvou podobných směrů, a to jak do polynomiálních neuronových sítí (PNN-Polynomial Neural Networks) tak i do neuronových sítí s nelinearitami vyšších řádů (HONN – Higher Order Nonlinear Neural Networks) např. [1]–[11].

V roce 2003 se autor článku připojil ke skupině (Gupta et al.) zabývající se implementací kvadratického a kubického polynomu jako samostatných neuronových jednotek (neuronů) [12]–[14]. Následný vývoj, terminologie a nová klasifikace nekonvenčních neuronových jednotek včetně kvadratické neuronové jednotky (QNU-Quadratic Neural Unit) jsou zaznamenány v publikacích [15]–[22].

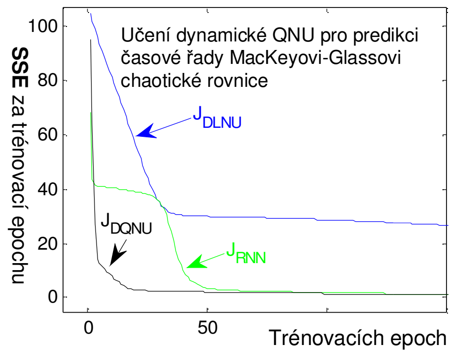
Z pohledu PNN lze QNU chápat jako speciální případ polynomiální neuronové sítě, zatímco z pohledu HONN lze na QNU nahlížet jako na samostatný stavební prvek sítě sestávající se z dílčích QNU jejichž synaptické vazby (agregace vstupů a zpětných vazeb) jsou kvadratické polynomy. Tab.1 v příloze shrnuje základní formy QNU a základní algoritmy učení. Shrnutí našich nejdůležitějších výsledků je publikováno v pracích [23]–[27];

3. Konvexní charakter QNU

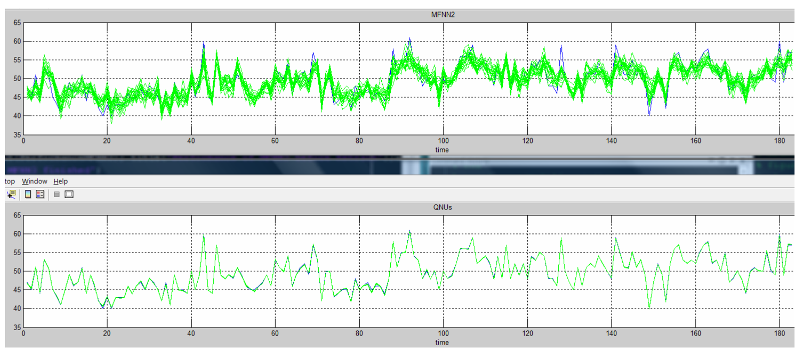
Pro ilustraci jsou uvedeny následující výsledky aplikací QNU.



Obr. 2: QNU jako adaptivní stavový regulátor dokáže výborně řídit systém batyskaf v celém rozsahu žádané veličiny i při relativně velké nepřesnosti modelu soustavy, kterým je opět QNU (dipl. práce L. Smetana [24]).

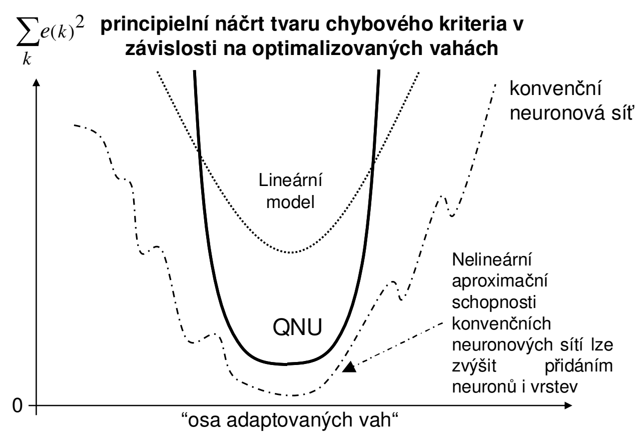


Obr. 3: Dynamická QNU při učení konverguje do blízkosti svého globálního minima a přesněji než lineární adaptivní prediktor DLNU.



Obr. 3: Trénování predikce doby výdechů pacienta: Sítě MLP (nahoře) se z různých počátečních podmínek natrénují nejednotně, zatímco statická QNU (dole) se natrénovala z různých počátečních vždy se stejným průběhem výstupů v čase a během mnohem menšího počtu epoch a velmi přesně.

Výsledky našich aplikací a experimentálních analýz statických a dynamických QNU [23]–[27] potvrzují následující zřejmý charakter chybového kriteria při optimalizaci kvadratického neuronu (statického i dynamického).



Obr. 4: Chybové kriterium pro optimalizaci kvadratického modelu má konvexní charakter (bez většího počtu lokálních minim) s plochým dnem, tj. QNU se rychle optimalizuje do blízkosti glob. minima.

4. Shrnutí

Článek úvodem prezentuje zařazení kvadratické neuronové jednotky v dnešním oboru umělých neuronových sítí a v odborné literatuře. Dále článek stručně demonstruje vhodné aplikační vlastnosti QNU související s konvexním charakterem kvadratického polynomu, tj. dobré nelineární aproximační vlastnosti a při tom spolehlivou konvergenci do blízkosti globálního minima a to při použití dnes již obyčejných optimalizačních metod (Tab. 1), které dokážou zvládnout i studenti třetího ročníku bakalářského studia.

Poděkování

Tento článek byl podpořen grantem SGS ČVUT č. SGS10/252/OHK2/3T/12.

Odkazy

[1] A. G. Ivakhnenko, “Polynomial Theory of Complex Systems,” *IEEE Tran. on Systems. Man, and Cybernetics*. Vol. SMC-1 (4). pp. 364-378, 1971.

[2] Y. Shin and J. Ghosh, “The Pi-sigma Network: An Efficient Higher-order Neural Network for Pattern Classification and Function Approximation,” *Proc. Int. Joint Conf. on Neural Networks* (pp. 13-18) , 1991.

[3] R. W. Softky and D. M. Kammen, “Correlations in high dimensional or asymmetrical data sets: Hebbian neuronal processing,” *Neural Networks*, 4, pp. 337-347, 1991.

[4] J. G. Taylor and S. Commbes, “Learning higher order correlations,” *Neural Networks*, 6, pp. 423-428, 1993.

[5] M. S. Chen and M. T. Manry, “Conventional modeling of the multilayer perceptron using polynomial basis functions,” *IEEE Trans. Neural Networks*, vol. 4, pp. 164–166, Jan. 1993.

[6] W. Schmidt, and J. Davis, “Pattern recognition properties of various feature spaces for higher order neural networks,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15, pp. 795-801, 1993.

[7] E. Kosmatopoulos, M. Polycarpou, M. Christodoulou, and P. Ioannou, “High-Order Neural Network Structures for Identification of Dynamical Systems,” *IEEE Trans. on Neural Networks*, vol. 6, no. 2, pp. 422-431, March 1995.

[8] N. Y. Nikolaev and H. Iba, “Learning Polynomial Feedforward Neural Network by Genetic Programming and Backpropagation”: *IEEE Trans. on Neural Networks*, vol. 14, no. 2, pp.337-350, March, 2003.

[9] N. Y. Nikolaev and H. Iba, *Adaptive Learning of Polynomial Networks: Genetic Programming, Backpropagation and Bayesian Methods,* Series: Genetic and Evolutionary Computation Springer, New York, XIV, 316 p. , ISBN: 0-387-31239-0, 2006.

[10] *Artificial Higher Order Neural Networks for Economics and Business*, ed. by Ming Zhang, Christopher Newport University, USA, IGI-Global, ISBN13: 9781615207114 EISBN13: 9781615207121, 2010

[11] *Artificial Higher Order Neural Networks for Computer Science and Engineering: Trends for Emerging Applications*, ed. by Ming Zhang, Christopher Newport University, USA, Information Science Reference, ISBN: 978-1-59904-897-0, 2010.

[12] Redlapalli, S., Song. K. –Y., and Gupta, M., M.: “Development of Quadratic Neural Unit With Applications to Pattern Classification”, *The Fourth International Symposium on Uncertainty Modeling and Analysis ISUMA 2003*, IEEE Computer Society, Maryland USA, ISBN 0-7695-1997-0, 2003, pp.141-146.

[13] Song, K. -Y., Redlapalli, S. and Gupta, M., M.: "Cubic Neural Units for Control Applications", *The Fourth International Symposium on Uncertainty Modeling and Analysis ISUMA 2003*, IEEE Computer Society, Maryland USA, ISBN 0-7695-1997-0, 2003, pp.324-329.

[14] Bukovsky I., S. Redlapalli and M. M. Gupta : Quadratic and Cubic Neural Units for Identification and Fast State Feedback Control of Unknown Non-Linear Dynamic Systems, *Fourth International Symposium on Uncertainty Modeling and Analysis ISUMA 2003*, IEEE Computer Society, 2003, Maryland USA, ISBN 0-7695-1997-0, p.p.330-334

[15] Bukovsky I., Bila J. : Development of Higher Order Nonlinear Neural Units for Evaluation of Complex Static and Dynamic Systems, *Proceedings of Workshop 2004, Part A*, March 2004, Vol.8, Special Issue, Czech Technical University, Czech Republic, Prague, pp. 372-373, ISBN 80-01-02945-X

[16] Bukovský, I. – Bíla, J.: Nelineární dynamické neuronové jednotky pro paralelní manipulátor TRIPOD. In: *Sborník ze semináře VZ MSM 212200008* [CD-ROM]. Praha: ČVUT, Fakulta strojní, 2004, díl 1, s. 66–68. ISBN 80-01-03105-5.

[17] Bukovsky, I.: Extended Dynamic Neural Architectures HONNU with Minimum Number of Neural Parameters for Evaluation of Nonlinear Dynamic Systems (in Czech), In: *New Methods and Approaches in the Fields of Control Technology, Automatic Control, and Informatics*, Czech Technical University, Prague, 2005, s. 93-97. ISBN 80-01-03240-X.

[18] Bukovsky, I., Bila, J.: „Basic Classification of Nonconventional Artificial Neural Units"(In Czech), Proceedings of Seminar Nove Hrady, Czech Technical University in Prague, FME, ISBN: 978-80-01-03747-8, Czech Republic, 2007, pp. 76-80.

[19] Bukovsky, I. : *Modeling of Complex Dynamic Systems by Nonconventional Artificial Neural Architectures and Adaptive Approach to Evaluation of Chaotic Time Series*, Ph.D. THESIS, Faculty of Mechanical Engineering, Czech Technical University in Prague (defended September 7, 2007).

[20] Bukovsky, I., Hou, Z-G., Gupta, M., M., Bila, J.: "Foundation of Notation and Classification of Nonconventional Static and Dynamic Neural Units", accepted paper for special section on neural networks for *ICCI 2007, The 6th IEEE International Conference on Cognitive Informatics*, California , USA, 2007, ISBN: 978-1-4244-1328-7.

[21] Bukovsky, I., Hou, Z-G., Bila, J., Gupta, M., M.: „Foundation of Nonconventional Neural Units and their Classification", *International Journal of Cognitive Informatics and Natural Intelligence (IJCiNi)*, 2(4), October-December 2008, IGI Publishing, Hershey PA, USA, pp.29-43, ISSN 1557-3958.

[22] Bukovsky, I., Bila. J,.: "Foundation and Classification of Nonconventional Neural Units and Paradigm of Nonsynaptic Neural Interaction" in *Discoveries and Breakthroughs in Cognitive Informatics and Natural Intelligence* within the series of the Advances in Cognitive Informatics and Natural Intelligence (ACINI), ed. Y. Wang, Information Science Reference, Hershey PA, USA, 2010. ISBN: 978-1-60566-902-1, p.508-523

[23] Bila, J., Bukovsky, I., Jura. J.: "Review of Development of Nonconventional Neural Architectures at the Czech Technical University in Prague", *10th WSEAS Int. Conf. on Neural Networks (NN'09)*, pp.138-146, Prague, Czech Republic, 2009. ISSN: 1790-5109, ISBN: 978-960-474-065-9 (best conference paper).

[24] Ladislav Smetana: *Nelineární neuro-regulátor pro úlohy automatického řízení*, Diplomová práce, ČVUT FS, 2008.

[25] Ivo Bukovsky, Martin Lepold and Jiri Bila: „Quadratic Neural Unit and its Network in Validation of Process Data of Steam Turbine Loop and Energetic Boiler", *2010 IEEE WCCI International Joint Conference on Neural Networks*, Barcelona, (accepted paper).

[26] Ivo Bukovsky, Kei Ichiji, Noriyasu Homma, Makoto Yoshizawa and Rodriguez Ricardo:"Testing Potentials of Dynamic Quadratic Neural Unit for Prediction of Lung Motion during Respiration for Tracking Radiation Therapy" ", *2010 IEEE WCCI International Joint Conference on Neural Networks*, Barcelona, (accepted paper).

[27] Ivo Bukovsky, Noriyasu Homma, Ladislav Smetana, Ricardo Rodriguez, Martina Mironovova, Stanislav Vrana: „Quadratic Neural Unit is a Good Compromise between Linear Models and Neural Networks for Industrial Applications", *ICCI 2010 The 9th IEEE International Conference on Cognitive Informatics*, July 7-9, 2010, Beijing, China, (accepted paper).

Příloha

Tab. 1: Základní formy QNU a algoritmy učení.

QNU	Matematická struktura	Algoritmus učení
Statická QNU	y_n... výstup neuronu $x_1, x_2, \ldots, x_n$ ... vnější vstupy W... horní trojúhelníková váhová matice $\mathbf{x}=\begin{bmatrix}1\\x_1\\\vdots\\x_n\end{bmatrix},\ \mathbf{W}=\begin{bmatrix}w_{00} & w_{01} & \cdots & w_{0n}\\0 & w_{11} & \cdots & w_{1n}\\\vdots & \ddots & \ddots & \vdots\\0 & \cdots & 0 & w_{nn}\end{bmatrix}$ $y_n=\sum_{i=0}^{n}\sum_{j=i}^{n}x_i x_j w_{ij}=\mathbf{x}^{\mathbf{T}}\mathbf{W}\,\mathbf{x}$	<u>Levenberg-Marquardt (globální optimalizace)</u> N...počet vzorků (délka dat) $\Delta w_{ij}=-(\mathbf{j}_{ij}{}^T\cdot\mathbf{j}_{ij}+\frac{1}{\mu})^{-1}\cdot\mathbf{j}_{ij}{}^T\cdot\mathbf{e}$ $\quad\mathbf{j_{ij}}=\frac{\partial\mathbf{y_n}}{\partial w_{ij}}=\begin{bmatrix}x_i(1)\,x_j(1)\\x_i(2)\,x_j(2)\\\vdots\\x_i(N)\,x_j(N)\end{bmatrix}$ $\mathbf{e}=\left[e(1)\quad e(2)\quad\cdots\quad e(N)\right]^T$ <u>Gradient Descent (lokální optimalizace)</u> k ...číslo vzorku $w_{ij}(k+1)=w_{ij}(k)+\Delta w_{ij}(k)$ $\Delta w_{ij}(k)=-\frac{1}{2}\mu\,e(k)^2=\mu\,e(k)\,x_i(k)\,x_j(k)$
Diskrétní dynamická QNU	$\mathbf{x}=\begin{bmatrix}1\\y_n(k+n_s-1)\\y_n(k+n_s-2)\\\vdots\\y_n(k+1)\\x_1(k)\\\vdots\\x_m(k)\end{bmatrix}$, $y_n(k+n_s)=\mathbf{x}^{\mathbf{T}}\mathbf{W}\,\mathbf{x}$ y_r... reálná hodnota typicky pro predikci: $x_1(k)=y_r(k)$ $x_m(k)=$ $=y_r(k-m+1)$	<u>RTRL (lokální optimalizace)</u> $\Delta w_{ij}(k)=\mu\,e(k)\,\frac{\partial y_n(k+n_s)}{\partial w_{ij}}$ $\frac{\partial y_n(k+n_s)}{\partial w_{ij}}=\frac{\partial(\mathbf{x}^T\mathbf{W}\mathbf{x})}{\partial w_{ij}}=\mathbf{j_{ij}}^T\,\mathbf{W}\,\mathbf{x}+x_i x_j+\mathbf{x}^T\mathbf{W}\,\mathbf{j_{ij}}$ $\mathbf{j_{ij}}=\frac{\partial\mathbf{x}}{\partial w_{ij}}=\left[0\ \frac{\partial y_n(k+n_s-1)}{\partial w_{ij}}\ \frac{\partial y_n(k+n_s-2)}{\partial w_{ij}}\cdots\frac{\partial y_n(k+1)}{\partial w_{ij}}\ 0\cdots0\right]^T$
Spojitá dynamická QNU (příklad dynamiky 2-ho řádu)	$\mathbf{x}=\begin{bmatrix}1\\y_n(t)\\y_n'(t)\\x_1(t)\\\vdots\\x_m(t)\end{bmatrix},\quad y_n''(t)=\mathbf{x}^{\mathbf{T}}\mathbf{W}\,\mathbf{x}$	<u>RTRL (lokální optimalizace)</u> $\Delta w_{ij}(k)=\mu\,e(t)\,\frac{\partial y_n(t)}{\partial w_{ij}}$ $\frac{\partial y_n(t)}{\partial w_{ij}}=\frac{\partial}{\partial w_{ij}}\iint(\mathbf{x}^T\mathbf{W}\mathbf{x})dt^2=\iint\left(\mathbf{j_{ij}}^T\,\mathbf{W}\,\mathbf{x}+x_i x_j+\mathbf{x}^T\mathbf{W}\,\mathbf{j_{ij}}\right)dt^2$ $\mathbf{j_{ij}}=\left[0\ \frac{\partial y_n(t)}{\partial w_{ij}}\ \frac{\partial y_n'(t)}{\partial w_{ij}}\ 0\cdots0\right]^T,\quad\frac{\partial y_n''(t)}{\partial w_{ij}}=\mathbf{j_{ij}}^T\,\mathbf{W}\,\mathbf{x}+x_i x_j+\mathbf{x}^T\mathbf{W}\,\mathbf{j_{ij}}$