ADAPTIVE NONLINEAR MODEL PREDICTIVE CONTROL BASED ON WIENER MODEL

DOKOUPIL, J. & PIVONKA, P.

Abstract: *This article deals with a nonlinear model predictive control design (NMPC) with closed loop identification which applies numerical optimization using the Levenberg-Marquardt method in iterative batch mode adaptation. The proposed approach enables asymptotic tracking of a reference trajectory by a prediction of a general nonlinear model. Investigation of the properties of the adaptive NMPC is performed using the Wiener nonlinear model which is considered to be suitable for representing a wide range of nonlinear process behavior. Although it requires little more effort in development than a standard pseudolinear model from the output error class, it offers better approximation of systems with highly nonlinear gains. The work therefore also seeks to formulate the optimal prediction of Wiener model output in both state space and input-output representation.*

Key words: *nonlinear model predictive control, time-varying systems, Wiener model, Levenberg-Marquardt method*

Authors´ data: Dipl.-Ing. **Dokoupil**, J[akub]; Univ.Prof. Dipl.-Ing. **Pivonka**, P[etr], Brno University of Technology, Kolejni 2906/4, 612 00, Brno, Czech Republic, xdokou05@stud.feec.vutbr.cz, pivonka@feec.vutbr.cz

This Publication has to be referred as: Dokoupil, J[akub] & Pivonka, P[etr] (2011). Adaptive Nonlinear Model Perdictive Control Based on Wiener Model, Chapter 33 in DAAAM International Scientific Book 2011, pp. 417-424, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-901509-84-1, ISSN 1726-9687, Vienna, Austria
DOI: 10.2507/daaam.scibook.2011.33

1. Introduction

NMPC falls within a class of controllers optimizing future plant evolution through the use of an explicit mathematical process description. The most widely used approach is linear model predictive control implementation (MPC) (Maeder et al., 2002) applied to a linear or Wiener model quadratic problem (Enso & Kaddour, 2002). Employing MPC together with Wiener model, which is able to better effect a general nonlinear process, requires a calculation of inversion of static nonlinearity in the feedback loop. Furthermore, given its linear and nonlinear parts, parameterization of the Wiener model might not be unique, which might lead to an incorrect synthesis of MPC.

The aim of this work is to generate the design of a predictive controller directly optimizing the Wiener model prediction and also to ensure tracking capability of time-varying parameters. Desired properties of adaptivity are achieved by suppressing an influence of old data to new parameter estimates by incorporating of finite data horizon technique. Optimal sequence of action values is newly found based on a LM search direction (Verheagen & Verdult, 2007). Some works, such as (Ekman, 2008), deal with the design of MPC for bilinear model.

The next area of research will be focused on the design of a nonlinear predictive controller and nonlinear process identification with a higher order of linear approximation and verification of adaptive NMPC using the Hammerstein-Wiener model and its variations.

2. Wiener Model

One of suitable candidates to describe a black-box process behavior is the Wiener model. It consists of a linear dynamic followed in series by a static nonlinearity. The Wiener's model structure directly describing input-output relation can be mathematically expressed as follows

$$z_k = \varphi_{k,\theta}\theta + v_k \tag{1}$$

$$y_k = g(z_k, \eta) \tag{2}$$

where y_k is a measured output, the vectors $\theta \in \mathbb{R}^{n_\theta}$ and $\eta \in \mathbb{R}^{n_\eta}$ represent a vectors of possibly time-varying parameters determining a linear and a nonlinear part of the model. The unmeasurable internal variable z_k is given by a linear combination of measured and estimated signals contained in the regressor vector $\varphi_{k,\theta} \in \mathbb{R}^{n_\theta}$ and by the stochastic disturbance v_k. The variable v_k represents disturbing noises and omnipresent modeling imperfections. v_k is ideally having the properties of zero mean white noise with known covariance matrix R_k, ie

$$v_k \sim (0, R_k) \tag{3}$$

$$E\{v_k v_j^T\} = R_k \delta_{k-j} \tag{4}$$

where δ_{k-j} is the Kronecker delta function given by

$$\delta_{k-j} = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases} \tag{5}$$

The static nonlinearity $g(\cdot, \eta)$ considered here is in a form of polynom of order n_η

$$g(z_k, \eta_k) = \sum_{i=1}^{n_\eta} [\eta]_i z_k^i \tag{6}$$

Equivalently, the Wiener model can be expressed in a state space representation as follows

$$x_{k+1} = Ax_k + Bu_k + Gv_k \tag{7}$$

$$z_k = Cx_k + v_k \tag{8}$$

$$y_k = g(z_k, \eta) \tag{9}$$

where u_k is the input variable, $x_k \in \mathbb{R}^{n_x}$ is the state of the linear dynamics represented by an innovative state space model. We assume throughout that matrices (A, B, G, C) ensure the reachability and the observability of linear dynamics.

3. Nonlinear Model Predictive Control

The role of predictive control can be interpreted as a requirement for tracking the reference trajectory by a predicted model output. In order to find the best model prediction which would use quadratic cost function based on a known sequence of input excitation $\{u_k\}$ and the knowledge of the initial state vector x_{k_0}, the optimal state predictor minimizing the cost function (10) is built first

$$\hat{x}_{k+i|k} = \arg\min_{\hat{x}} E\{[x_{k+i} - \hat{x}]^T [x_{k+1} - \hat{x}]\} \tag{10}$$

which results in

$$\hat{x}_{k+i|k} = A^{i-1}(A - GC)x_k + A^{i-1}Gy_k + \sum_{j=1}^{i} A^{i-j}Bu_{k+j-1} \tag{11}$$

Optimal i-step ahead output predictor shall be probably found by solving the following problem

$$\hat{y}_{k+i|k} = \arg\min_{\hat{y}} E\{[y_{k+i} - \hat{y}]^T [y_{k+i} - \hat{y}]\} \tag{12}$$

by substituting for y_{k+i} from the equation (9) and by replacing the vectors x_{k+i} and η by theirs best possible estimates $\hat{x}_{k+i|k}$ and $\hat{\eta}$, we obtain

$$\hat{y}_{k+i|k} = g\left(C\hat{x}_{k+i|k}, \hat{\eta}\right) \tag{13}$$

rewritten (13) into a compact matrix form results in

$$\hat{Z}_{k+N,N} = \mathcal{O}_N x_k + \mathcal{K}_N y_k + \mathcal{T}_N U_{k+N-1,N} \tag{14}$$

$$\hat{Y}_{k+N,N} = g\left(\hat{Z}_{k+N,N}, \hat{\eta}\right) \tag{15}$$

Vectors $\hat{Y}_{k+N,N}$ and $U_{k+N-1,N}$ will henceforth be referred to as vectors of predicted outputs and vectors of inputs. Vector indices will be written following these rules: the first entry of subscript refers to the time index of the vector's top entry and the second entry of subscript refers to the number of predicted values. The matrix $\mathcal{O}_N$ will be henceforth referred to as the extended observability matrix, where the subscript corresponds (as in the case of matrices $\mathcal{K}_N$ and $\mathcal{T}_N$) to the number of predicted values. Individual matrices take a form

$$\hat{Y}_{k+N,N} = [\hat{y}_{k+N|k} \quad \cdots \quad \hat{y}_{k+2|k} \quad \hat{y}_{k+1|k}]^T \tag{16}$$

$$\hat{Z}_{k+N,N} = [\hat{z}_{k+N|k} \quad \cdots \quad \hat{z}_{k+2|k} \quad \hat{z}_{k+1|k}]^T \tag{17}$$

$$U_{k+N-1,N} = [u_{k+N-1} \quad \cdots \quad u_{k+1} \quad u_k]^T \tag{18}$$

$$\mathcal{O}_N = \begin{bmatrix} CA^{N-1}(A-GC) \\ \vdots \\ CA(A-GC) \\ C(A-GC) \end{bmatrix} \quad \mathcal{K}_N = \begin{bmatrix} CA^{N-1}G \\ \vdots \\ CAG \\ CG \end{bmatrix} \tag{19}$$

$$\mathcal{T}_N = \begin{bmatrix} CB & \cdots & CA^{N-2}B & CA^{N-1}B \\ & \ddots & \vdots & \vdots \\ & & CB & CAB \\ & & & CB \end{bmatrix} \tag{20}$$

The goal is to minimize the square vector of errors $E_{r,k+N,N,U}$ in prediction horizon $\hat{Y}_{k+N,N}$ of the reference trajectory $W_{k+N,N}$ (Muske & Badgwell, 2002) which consists of sequences of the desired values

$$E_{r,k+N,N,U} = W_{k+N,N} - \hat{Y}_{k+N,N} \tag{21}$$

$$W_{k+N,N} = [w_{k+N} \quad \cdots \quad w_{k+2} \quad w_{k+1}]^T \tag{22}$$

and, at the same time, to penalize the invested control energy. As the problem considered here is nonlinear, given the properties of the model, its direct analytical solution is not tractable. Numerical optimization is in this case achieved using the technique of linearization in the neighborhood of $U^{i-1}_{k+N-1,N}$. For the sake of clarity we will employ the following notation $U^{i-1} \leftarrow U^{i-1}_{k+N-1,N}$ and $E_{r,U^i} \leftarrow E_{r,k+N,N,U^i}$.

For the cost function and its approximation by a first-order Taylor series expressed in a matrix form, we write

$$V_N\left(u_{k+N-1}^{i-1} + \delta u_{k+N-1}^{i}, \cdots, u_k^{i-1} + \delta u_k^{i}\right) = \left\|E_{r,U^i}\right\|_{Q_{e,N}}^2 + \left\|U^i\right\|_{Q_{u,N}}^2 \approx \left\|E_{r,U^{i-1}} + \left(\frac{\partial E_{r,U^{i-1}}^T}{\partial U^{i-1}}\right)^T \delta U^i\right\|_{Q_{e,N}}^2 + \left\|U^{i-1} + \delta U^i\right\|_{Q_{u,N}}^2 \quad (23)$$

in which $Q_{e,N}$ and $Q_{u,N}$ are positive definite weighting matrices and with the requirement of distance reduction $\delta U^i = U^i - U^{i-1}$. A numerically more attractive approach, which uses the Levenberg-Marquardt techniques, will be achieved by expansion of the cost function (23) by adding a restrictive condition which will require minimal variation in δU^i. By formulating a multicriterial problem and applying the Lagrange multiplier λ which works with a set of positive real numbers, we obtain

$$V_{N,\lambda}\left(u_{k+N-1}^{i-1} + \delta u_{k+N-1}^{i}, \cdots, u_k^{i-1} + \delta u_k^{i}\right) = V_N\left(u_{k+N-1}^{i-1} + \delta u_{k+N-1}^{i}, \cdots, u_k^{i-1} + \delta u_k^{i}\right) + \lambda\left\|\delta U^i\right\|_2^2 \quad (24)$$

The linear problem of least-squares is solved in relation to the operation of δU^i in the linear approximation of the cost function $V_{N,\lambda}$

$$\left\{\delta u_{k+N-1}^{*i}, \cdots, \delta u_k^{*i}\right\} = \arg\min_{\delta u_{k+N-1}^i, \cdots, \delta u_k^i} V_{N,\lambda} \quad (25)$$

and as a result we obtain a relation for iterative actualization of a sequence of control actions for $i = 1,2,\cdots,i_{max}$ (where i_{max} is the desired number of update iterations)

$$U^i = U^{i-1} + \left[\Psi_{r,,U^{i-1}}^T Q_{e,N} \Psi_{r,,U^{i-1}} + Q_{u,N} + \lambda I\right]^{-1} \left[\Psi_{r,,U^{i-1}}^T Q_{e,N} E_{r,U^{i-1}} - Q_{u,N} U^{i-1}\right] \quad (26)$$

where $\Psi_{r,U^{i-1}} = \left[\frac{\partial \hat{Y}_{k+N,N}^T}{\partial U^{i-1}}\right]^T$

4. Wiener Model Identification

Numerical optimization based on gradient search direction (Verheagen & Verdult, 2007), which is our case, utilizes local linearization of an investigated problem. In order to build an optimal predictor of the linearized model we perform a Taylor series expansion of the equation (2) around $\theta = \bar{\theta}$, $\eta = \bar{\eta}$ and $v_k = 0$ to obtain the following

$$y_k \approx g\left(\varphi_{k,\bar{\theta}}^T \bar{\theta} + v_k, \bar{\eta}\right) + \begin{pmatrix} \left.\frac{\partial g\left(\varphi_{k,\theta}^T \theta + v_k, \eta\right)}{\partial \theta}\right|_{\theta=\bar{\theta}, \eta=\bar{\eta}, v_k=0} \\ \left.\frac{\partial g\left(\varphi_{k,\theta}^T \theta + v_k, \eta\right)}{\partial \eta}\right|_{\theta=\bar{\theta}, \eta=\bar{\eta}, v_k=0} \end{pmatrix}^T \begin{pmatrix} \theta - \bar{\theta} \\ \eta - \bar{\eta} \end{pmatrix}$$

$$+ \left(\left.\frac{\partial g\left(\varphi_{k,\theta}^T \theta + v_k, \eta\right)}{\partial v_k}\right|_{\theta=\bar{\theta}, \eta=\bar{\eta}, v_k=0} \right)^T v_k \tag{27}$$

and hence assuming $\theta = \bar{\theta}$ and $\eta = \bar{\eta}$ for the optimal predictor satisfying

$$\hat{y}_{k|k-1} = \arg\min_{\hat{y}} E\{[y_k - \hat{y}]^T [y_k - \hat{y}]\} \tag{28}$$

the following notation will be valid

$$\hat{y}_{k|k-1} = g\left(\varphi_{k,\theta}^T \theta, \eta\right) \tag{29}$$

To evaluate the quality of a model, we compare the set of predicted outputs $\hat{Y}_{k,N,\theta_k,\eta_k}$ with the corresponding measured outputs $Y_{k,N}$ in the following prediction error cost function

$$V_N(\theta_k) = \arg\min_{\theta_k} \left\| Y_{k,N} - \hat{Y}_{k,N,\theta_k,\eta_k} \right\|_{Q_N}^2 \tag{30}$$

in which the matrix Q_N ensures exponentialy forgetting of information in a data window

$$Q_N = \begin{bmatrix} \lambda_e^0 & & \\ & \ddots & \\ & & \lambda_e^{N-1} \end{bmatrix} \tag{31}$$

The optimal values of $\hat{\theta}_k^i$ and $\hat{\eta}_k^i$ will be identified solving the problem given by linearization of the constrained cost function (30) in accordance with the Levenberg-Marquardt search direction

$$\delta \begin{pmatrix} \hat{\theta}_k^i \\ \hat{\eta}_k^i \end{pmatrix} = \arg \min_{\delta \begin{pmatrix} \theta_k^i \\ \eta_k^i \end{pmatrix}} \left(V_N \left(\begin{pmatrix} \theta_k^{i-1} \\ \eta_k^{i-1} \end{pmatrix} + \delta \begin{pmatrix} \theta_k^i \\ \eta_k^i \end{pmatrix} \right) + \lambda V_\delta \left(\delta \begin{pmatrix} \theta_k^i \\ \eta_k^i \end{pmatrix} \right) \right)$$

$$\approx \arg\min_{\delta \begin{pmatrix} \theta_k^i \\ \eta_k^i \end{pmatrix}} \left\| \begin{bmatrix} E_{k,N,\theta_k^{i-1},\eta_k^{i-1}} \\ 0 \end{bmatrix} + \begin{bmatrix} \left(\frac{\partial E_{k,N,\theta_k^{i-1},\eta_k^{i-1}}^T}{\partial \theta_k^{i-1}} \right)^T \\ \sqrt{\lambda} I \end{bmatrix} \delta \begin{pmatrix} \theta_k^i \\ \eta_k^i \end{pmatrix} \right\|_{\begin{bmatrix} Q_N & 0 \\ 0 & I \end{bmatrix}}^2 \tag{32}$$

where $\hat{E}_{k,N,\theta_k^{i-1},\eta_k^{i-1}}$ is the vector of a prediction errors

$$\hat{E}_{k,N,\hat{\theta}_k^{i-1},\hat{\eta}_k^{i-1}} = \begin{bmatrix} y_k \\ \vdots \\ y_{k-N+2} \\ y_{k-N+1} \end{bmatrix} - \begin{bmatrix} g\left(\varphi^T_{k,\hat{\theta}_k^{i-1}}\hat{\theta}_k^{i-1}, \hat{\eta}_k^{i-1}\right) \\ \vdots \\ g\left(\varphi^T_{k-N+2,\hat{\theta}_k^{i-1}}\hat{\theta}_k^{i-1}, \hat{\eta}_k^{i-1}\right) \\ g\left(\varphi^T_{k-N+1,\hat{\theta}_k^{i-1}}\hat{\theta}_k^{i-1}, \hat{\eta}_k^{i-1}\right) \end{bmatrix} = Y_{k,N} - \hat{Y}_{k,N,\hat{\theta}_k^{i-1},\hat{\eta}_k^{i-1}} \quad (33)$$

Solving the problem (32), following iterative scheme for $i = 1,2,\cdots,i_{max}$ shall be found

$$\begin{pmatrix} \hat{\theta}_k^i \\ \hat{\eta}_k^i \end{pmatrix} = \begin{pmatrix} \theta_k^{i-1} \\ \eta_k^{i-1} \end{pmatrix} + \left(\Psi^T_{k,N,\hat{\theta}_k^{i-1},\eta_k^{i-1}} Q_N \Psi_{k,N,\hat{\theta}_k^{i-1},\eta_k^{i-1}} + \lambda I\right)^{-1} \Psi^T_{k,N,\hat{\theta}_k^{i-1},\eta_k^{i-1}} Q_N \hat{E}_{k,N,\theta_k^{i-1},\eta_k^{i-1}} \quad (34)$$

where $\Psi_{k,N,\hat{\theta}_k^{i-1},\eta_k^{i-1}} = \left[\frac{\partial \hat{Y}^T_{k,N,\theta_k^{i-1},\eta_k^{i-1}}}{\partial \begin{pmatrix} \hat{\theta}_k^{i-1} \\ \eta_k^{i-1} \end{pmatrix}}\right]^T$ denotes the gradient of $\hat{Y}^T_{k,N,\theta_k^{i-1},\eta_k^{i-1}}$ with respect to $\hat{\theta}_k^{i-1}$ and η_k^{i-1}. The initialization and reinicialization of the iterative procedure (34) is given by

$$\hat{\theta}_{k_0}^0 = E\{\theta_{k_0}\} \quad (35)$$

$$\hat{\theta}_k^0 = \hat{\theta}_{k-1}^{i_{max}} \quad (36)$$

5. Simulation Example

In this section, adaptive NMPC practical properties are examined via simulation example. The Wiener system to be controlled has static nonlinearity given by $g(z_k,\eta) = z_k - 0.5z_k^2 + 0.1z_k^3$ and its corresponding linear part by $\frac{y(s)}{u(s)} = \frac{2}{(10s+1)(s+1)^2}$, which is sampled with period $T_s = 0.2s$. The result of the Wiener process control by suggested adaptive NMPC is shown in Fig. 1., where for both explicit parts $\lambda = 0.01$ was chosen, length of a model prediction $N = 100$, number of a measurements $N = 300$, $Q_{e,N} = I$, $Q_{u,N} = 0.015I$.

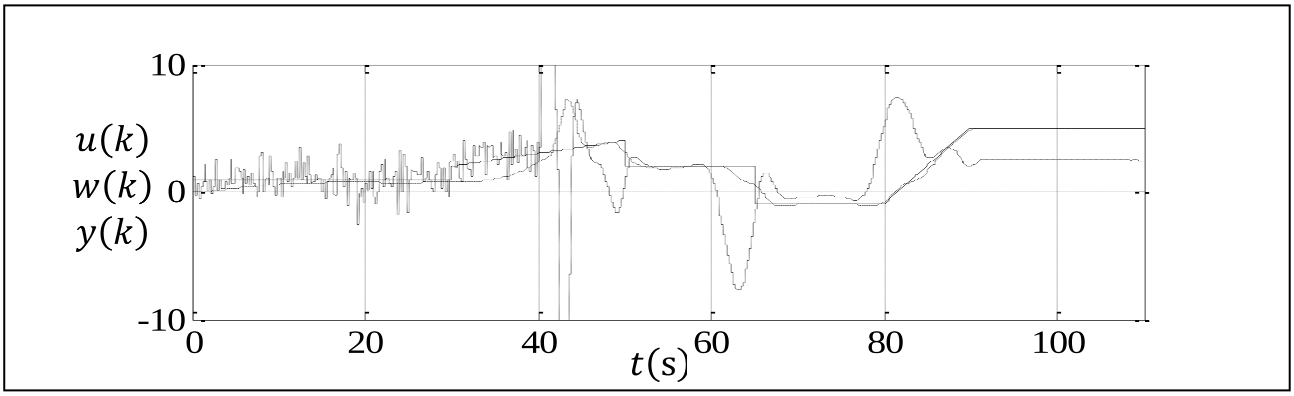

Fig. 1. Direct application of adaptive NMPC to the Wiener process control. Dashed line $u(k)$, solid line $w(k)$, dot-dashed line $y(k)$.

6. Conclusion

This paper presents and develops the strategy of adaptive NMPC design which is based on LM search direction. The simulations result shows that the proposed approach is effective in the tracking capability of referencing trajectory of the Wiener model without having to calculate the inversion of static nonlinearity. It should be mentioned, that the theoretical aspects of algorithm provides its implementation in the cases, when other types of nonlinear model structures (as Hammerstein-Wiener etc.) or just saturation (caused by analog converters) are introduced.

7. Acknowledgements

This work was supported as a part of the solution of Czech Science Foundation GAČR project No. 102/09/1680 Control Algorithm Design by Means of Evolutionary Approach and by the Czech Ministry of Education in the frame of MSM 0021630529 Intelligent Systems in Automation.

8. References

Ekman, M. (2008). Bilinear Black-Box Identification and MPC of the Activated Sludge Process. *Journal of Process Control*, Vol. 18, No. 7, 643-653, ISSN 0959-1524

Enso, I.; Kaddour, N. (2002). *Advanced Process Identification*, Marcel Dekker, Inc., ISBN 0-8247-0648-X, New York

Maeder, U.; Borrelli, F. & Morari, M. (2002). Linear Offset-Free Model Predictive Control. *Automatica*, Vol. 42, No. 10, 2214-2222, ISSN 0005-1098

Muske, K.; Badgwell, T. (2002). Disturbance Modeling for Offset-Free Linear Model Predictive Control. *Journal of Process Control*, Vol. 12, No. 5, 617-632, ISSN 0959-1524

Verheagen, M.; Verdult, V. (2007). *Filtering and System Identification: a least squares approach*, Cambridge University Press, ISBN 0-521-87512-9, Cambridge