Stable neural network based model predictive control

Krzysztof Patan and Józef Korbicz

Institute of Control and Computation Engineering
University of Zielona Góra

Introduction

- Model Predictive Control (MPC) – modern control strategy
- MPC derives a control signal by solving at each sampling time finite horizon open-loop optimal control problem
- Predictive control algorithms are able to consider constraints imposed on both controls and process outputs (states)
- Stability problems

unconsidered nonlinearities, e.g. inequalities imposed on process variables, may result in degraded performance of the closed-loop control and may lead to stability problems

Stability of MPC strategies

- Rich literature about linear/nonlinear MPC represented in the state-space (survey paper: Mayne et al. 2000)
  1. Cost in the form of a Lyapunov candidate function
     - terminal constraint (Keerthi and Gilbert, 1988)
     - infinite output prediction horizon (Keerthi and Gilbert 1988)
     - terminal cost function (Rawlings and Muske, 1993)
     - terminal costraint set methods (Scokaert, Mayne and Rawlings, 1999)
  2. Requirement that the state is decreasing in some norm (Bemporad, 1998)
- High level of maturity

- Stability of MPC using GPC concept for linear systems
  - infinite horizon GPC (GPC^{∞}) (Scokaert and Clarke, 1994)
  - Constrained Receding-Horizon Predictive Control (CRHPC) (Clarke and Scatollini, 1991)
  - Stable Generalized Predictive Control (SGPC) (Gossner, Kouvaritakis and Rossiter, 1997)
  - min-max GPC (Kim, Kwon and Lee, 1998)
- This paper proposes nonlinear predictive control using a dynamic neural network
- The stability is investigated checking the monotonicity of the cost – extension of the approach proposed by Scokaert and Clarke, (1994)
  - predictor is nonlinear
  - prediction horizon is finite
  - control horizon is not greater than the prediction horizon

k.patan@issi.uz.zgora.pl

Nonlinear MPC

- Cost based on the GPC criterion

$$J = \sum_{i=N_1}^{N_2} e^2(k+i) + \rho \sum_{i=1}^{N_u} \Delta u^2(k+i-1)$$

where $e(k+1) = r(k+i) - \hat{y}(k+i)$

$r(k+i)$ – the future reference signal

$\hat{y}(k+i)$ – the prediction of future outputs

$\Delta u(k+i-1) = u(k+i-1) - u(k+i-2)$

$\Delta u(k+i-1)$ – control change

ρ – the factor penalizing changes in the control signal

- Constraints on control moves

$$\Delta u(k+i) = 0, \quad N_u \leqslant i \leqslant N_2 - 1$$

- Constraints on process variable v

$$\underline{v} \leqslant v(k+j) \leqslant \overline{v}, \quad \forall j \in [0, N_v]$$

where N_v – constraint horizon
$\underline{v}$ – lower limits
$\overline{v}$ – upper limits

- Terminal constraints, e.g.

$$e(k+N_p+j) = 0, \quad \forall j \in [1, N_c],$$

where N_c – terminal constraint horizon

Neural predictor

- Prediction can be done by successive recursion of a one-step ahead nonlinear model
- One-step ahead prediction

$$\hat{y}(k+1) = f(y(k), \ldots, y(k-n_a+1), u(k), \ldots, u(k-n_b+1)) \qquad (1)$$

  where n_a and n_b represent number of past outputs and inputs, respectively
- Function f can be realized using dynamic neural network
- i-step ahead prediction

$$\hat{y}(k+i) = f(y(k+i-1), \ldots, y(k+i-n_a), u(k+i-1), \ldots u(k+i-n_b)) \quad (2)$$

- Measurements of the output are available up to time k – one should substitute predictions for actual measurements since these do not exist

$$y(k+i) = \hat{y}(k+i), \quad \forall i > 1$$

k.patan@issi.uz.zgora.pl

Problem definition

Let us redefine the nonlinear model predictive control based on the following open-loop optimization problem

$$\boldsymbol{u}(k) \triangleq \min_{\boldsymbol{u}} J \tag{3a}$$

$$\text{s.t.} \quad e(k+N_2+j) = 0, \quad \forall j \in [1, N_c], \tag{3b}$$

$$\Delta u(k+N_u+j) = 0, \quad \forall j \geqslant 0, \tag{3c}$$

$$\underline{u} \leqslant u(k+j) \leqslant \overline{u}, \quad \forall j \in [0, N_u-1], \tag{3d}$$

where N_c – the terminal constraints horizon
$\underline{u}$ – lower control bound
$\overline{u}$ – upper control bound

 k.patan@issi.uz.zgora.pl

Stability conditions

Proposition

The nonlinear model predictive control system (3) using the predictor (2) is asymptotically stable if the following conditions are satisfied:

i) $\rho \neq 0$,

ii) $N_c = \max\left[n_a + 1, \max\left[0, n_b + N_u - N_2\right]\right]$,

regardless the choice of N_1, N_2, and N_u.

Sketch of the proof

The cost function at time k has the form:

$$J(k) = \sum_{i=N_1}^{N_2} e^2(k+i) + \rho \sum_{i=1}^{N_u} \Delta u^2(k+i-1)$$

$\boldsymbol{u}(k)$ – is the optimal control at time k
$\boldsymbol{u}^*(k+1)$ – the suboptimal control postulated at time $k+1$

if $\boldsymbol{u}(k) = [u(k), u(k+1), \dots u(k+N_u-1)]^T$ then
$\boldsymbol{u}^*(k+1) = [u(k+1), \dots, u(k+N_u-1), u(k+N_u-1)]^T$

$$J^*(k+1) = \sum_{i=N_1+1}^{N_2+1} e^2(k+i) + \rho \sum_{i=2}^{N_u} \Delta u^2(k+i-1)$$

Sketch of the proof

The cost function at time k has the form:

$$J(k) = \sum_{i=N_1}^{N_2} e^2(k+i) + \rho \sum_{i=1}^{N_u} \Delta u^2(k+i-1)$$

$\boldsymbol{u}(k)$ – is the optimal control at time k
$\boldsymbol{u}^*(k+1)$ – the suboptimal control postulated at time $k+1$

if $\boldsymbol{u}(k) = [u(k), u(k+1), \dots u(k+N_u-1)]^T$ then
$\boldsymbol{u}^*(k+1) = [u(k+1), \dots, u(k+N_u-1), u(k+N_u-1)]^T$

$$J^*(k+1) = \sum_{i=N_1+1}^{N_2+1} e^2(k+i) + \rho \sum_{i=2}^{N_u} \Delta u^2(k+i-1)$$

 k.patan@issi.uz.zgora.pl

Sketch of the proof

The cost function at time k has the form:

$$J(k) = \sum_{i=N_1}^{N_2} e^2(k+i) + \rho \sum_{i=1}^{N_u} \Delta u^2(k+i-1)$$

$\boldsymbol{u}(k)$ – is the optimal control at time k
$\boldsymbol{u}^*(k+1)$ – the suboptimal control postulated at time $k+1$

if $\boldsymbol{u}(k) = [u(k), u(k+1), \ldots u(k+N_u-1)]^T$ then
$\boldsymbol{u}^*(k+1) = [u(k+1), \ldots, u(k+N_u-1), u(k+N_u-1)]^T$

$$J^*(k+1) = \sum_{i=N_1+1}^{N_2+1} e^2(k+i) + \rho \sum_{i=2}^{N_u} \Delta u^2(k+i-1)$$

 k.patan@issi.uz.zgora.pl

Sketch of the proof

Difference of cost $J(k)$ and $J^*(k+1)$

$$J^*(k+1) - J(k) = e^2(k + N_2 + 1) - e^2(k + N_1) - \rho\Delta u^2(k)$$

Sketch of the proof

Difference of cost $J(k)$ and $J^*(k+1)$

$$J^*(k+1) - J(k) = \underbrace{e^2(k+N_2+1)}_{=0} - e^2(k+N_1) - \rho\Delta u^2(k)$$

Sketch of the proof

Difference of cost $J(k)$ and $J^*(k+1)$

$$J^*(k+1) - J(k) = -e^2(k+N_1) - \rho\Delta u^2(k)$$

Sketch of the proof

Difference of cost $J(k)$ and $J^*(k+1)$

$$J^*(k+1) - J(k) = -e^2(k+N_1) - \rho\Delta u^2(k) \leqslant 0$$

 k.patan@issi.uz.zgora.pl

Sketch of the proof

Difference of cost $J(k)$ and $J^*(k+1)$

$$J^*(k+1) - J(k) = -e^2(k+N_1) - \rho\Delta u^2(k) \leqslant 0$$

Tracking error equality constraints hold for all $j \geqslant 1$ if:

i) $N_c = n_a + 1$, assuming that $n_a \geqslant n_b + N_u - N_2$,

ii) $N_c = n_b + N_u - N_2, \quad N_c > 0$, assuming that $n_a < n_b + N_u - N_2$.

k.patan@issi.uz.zgora.pl

Sketch of the proof

Difference of cost $J(k)$ and $J^*(k+1)$

$$J^*(k+1) - J(k) = -e^2(k+N_1) - \rho\Delta u^2(k) \leqslant 0$$

Tracking error equality constraints hold for all $j \geqslant 1$ if:

i) $N_c = n_a + 1$, assuming that $n_a \geqslant n_b + N_u - N_2$,

ii) $N_c = n_b + N_u - N_2, \quad N_c > 0$, assuming that $n_a < n_b + N_u - N_2$.

Setting the constraint horizon on the value:

$$N_c = \max\left[n_a + 1, \max\left[0, n_b + N_u - N_2\right]\right]$$

quarantees that tracking error equality constraints hold not only for $j \in [1, N_c]$ but for all $j \geqslant 1$

k.patan@issi.uz.zgora.pl

Sketch of the proof

Moreover $\boldsymbol{u}^*(k+1)$ satisfies all constraints at time $k+1$, and subsequently the vector $\Delta\boldsymbol{u}^*(k+1)$ also satisfies constraints

Assuming $\boldsymbol{u}(k+1)$ as the optimal solution of the optimization problem time $k+1$ then

$$J(k+1) \leqslant J^*(k+1)$$

and

$$\Delta J(k+1) = J(k+1) - J(k) \leqslant -e^2(k+N_1) - \rho\Delta u^2(k)$$

Finally, for $\rho \neq 0$ the cost is monotonically decreased with respect to time and the control system is stable ■

 k.patan@issi.uz.zgora.pl

Sketch of the proof

Moreover $\boldsymbol{u}^*(k+1)$ satisfies all constraints at time $k+1$, and subsequently the vector $\Delta\boldsymbol{u}^*(k+1)$ also satisfies constraints

Assuming $\boldsymbol{u}(k+1)$ as the optimal solution of the optimization problem time $k+1$ then

$$J(k+1) \leqslant J^*(k+1)$$

and

$$\Delta J(k+1) = J(k+1) - J(k) \leqslant -e^2(k+N_1) - \rho\Delta u^2(k)$$

Finally, for $\rho \neq 0$ the cost is monotonically decreased with respect to time and the control system is stable ■

k.patan@issi.uz.zgora.pl

Sketch of the proof

Moreover $\boldsymbol{u}^*(k+1)$ satisfies all constraints at time $k+1$, and subsequently the vector $\Delta\boldsymbol{u}^*(k+1)$ also satisfies constraints

Assuming $\boldsymbol{u}(k+1)$ as the optimal solution of the optimization problem time $k+1$ then

$$J(k+1) \leqslant J^*(k+1)$$

and

$$\Delta J(k+1) = J(k+1) - J(k) \leqslant -e^2(k+N_1) - \rho\Delta u^2(k)$$

Finally, for $\rho \neq 0$ the cost is monotonically decreased with respect to time and the control system is stable ■

Constrained optimization

Let us recall the optimization problem as follows:

$$\boldsymbol{u}(k) \triangleq \min_{\boldsymbol{u}} J(\boldsymbol{u}) \tag{4a}$$

$$\text{s.t.} \quad h_{1_i}(\boldsymbol{u}) = 0, \quad \forall i \in [1, N_c], \tag{4b}$$

$$h_{2_i}(\boldsymbol{u}) = 0, \quad \forall i \geqslant 0 \tag{4c}$$

$$g_{1_i}(\boldsymbol{u}) \leqslant 0, \quad \forall i \in [0, N_u - 1], \tag{4d}$$

$$g_{2_i}(\boldsymbol{u}) \leqslant 0, \quad \forall j \in [0, N_u - 1], \tag{4e}$$

where $h_{1_i}(\boldsymbol{u}) = e(k + N_2 + i)$
$h_{2_i}(\boldsymbol{u}) = \Delta u(k + N_u + i)$
$g_{1_i}(\boldsymbol{u}) = u(k + i) - \overline{u}$
$g_{2_i}(\boldsymbol{u}) = \underline{u} - u(k + i)$

 k.patan@issi.uz.zgora.pl

Let us define Lagrangian

$$L(\boldsymbol{u}, \boldsymbol{\mu_1}, \boldsymbol{\mu_2}, \boldsymbol{\mu_3}) = J(\boldsymbol{u}) + \sum_{i=1}^{N_c} \mu_{1_i} h_{1_i}(\boldsymbol{u}) + \sum_{i=1}^{N_u-1} \mu_{2_i} g_{1_i}(\boldsymbol{u}) + \sum_{i=1}^{N_u-1} \mu_{3_i} g_{2_i}(\boldsymbol{u}) \quad (5)$$

with Lagrange multiplier vectors $\boldsymbol{\mu_1}$, $\boldsymbol{\mu_2}$, $\boldsymbol{\mu_3}$

Transformation of the original problem to its unconstrained form

$$\bar{J}(\boldsymbol{u}) = J(\boldsymbol{u}) + \mu \sum_{i=1}^{N_c} h_{1_i}^2(\boldsymbol{u}), \quad (6)$$

Objective – to solve an unconstrained problem:

$$\boldsymbol{u}(k) \triangleq \min_{\boldsymbol{u}} \quad \bar{J}(\boldsymbol{u}), \quad (7)$$

where μ is suitably large constant

If $\boldsymbol{u}_\mu$ is a solution of the problem (5), and excluding inequality constraints (4d) and (4e) it can be shown that as $\mu \to \infty$ there obtains $\boldsymbol{u}_\mu \to \boldsymbol{u}^*$, where $\boldsymbol{u}^*$ is solution of (4)

Let us define Lagrangian

$$L(\boldsymbol{u}, \boldsymbol{\mu_1}, \boldsymbol{\mu_2}, \boldsymbol{\mu_3}) = J(\boldsymbol{u}) + \sum_{i=1}^{N_c} \mu_{1_i} h_{1_i}(\boldsymbol{u}) + \sum_{i=1}^{N_u-1} \mu_{2_i} g_{1_i}(\boldsymbol{u}) + \sum_{i=1}^{N_u-1} \mu_{3_i} g_{2_i}(\boldsymbol{u}) \quad (5)$$

with Lagrange multiplier vectors $\boldsymbol{\mu_1}$, $\boldsymbol{\mu_2}$, $\boldsymbol{\mu_3}$

Transformation of the original problem to its unconstrained form

$$\bar{J}(\boldsymbol{u}) = J(\boldsymbol{u}) + \mu \sum_{i=1}^{N_c} h_{1_i}^2(\boldsymbol{u}), \quad (6)$$

Objective – to solve an unconstrained problem:

$$\boldsymbol{u}(k) \triangleq \min_{\boldsymbol{u}} \quad \bar{J}(\boldsymbol{u}), \quad (7)$$

where μ is suitably large constant

If $\boldsymbol{u}_\mu$ is a solution of the problem (5), and excluding inequality constraints (4d) and (4e) it can be shown that as $\mu \to \infty$ there obtains $\boldsymbol{u}_\mu \to \boldsymbol{u}^*$, where $\boldsymbol{u}^*$ is solution of (4)

Let us define Lagrangian

$$L(\boldsymbol{u},\boldsymbol{\mu_1},\boldsymbol{\mu_2},\boldsymbol{\mu_3}) = J(\boldsymbol{u}) + \sum_{i=1}^{N_c} \mu_{1_i} h_{1_i}(\boldsymbol{u}) + \sum_{i=1}^{N_u-1} \mu_{2_i} g_{1_i}(\boldsymbol{u}) + \sum_{i=1}^{N_u-1} \mu_{3_i} g_{2_i}(\boldsymbol{u}) \quad (5)$$

with Lagrange multiplier vectors $\boldsymbol{\mu_1}$, $\boldsymbol{\mu_2}$, $\boldsymbol{\mu_3}$

Transformation of the original problem to its unconstrained form

$$\bar{J}(\boldsymbol{u}) = J(\boldsymbol{u}) + \mu \sum_{i=1}^{N_c} h_{1_i}^2(\boldsymbol{u}), \quad (6)$$

Objective – to solve an unconstrained problem:

$$\boldsymbol{u}(k) \triangleq \min_{\boldsymbol{u}} \quad \bar{J}(\boldsymbol{u}), \quad (7)$$

where μ is suitably large constant

If $\boldsymbol{u}_\mu$ is a solution of the problem (5), and excluding inequality constraints (4d) and (4e) it can be shown that as $\mu \to \infty$ there obtains $\boldsymbol{u}_\mu \to \boldsymbol{u}^*$, where $\boldsymbol{u}^*$ is solution of (4)

 k.patan@issi.uz.zgora.pl

Let us define Lagrangian

$$L(\boldsymbol{u}, \boldsymbol{\mu_1}, \boldsymbol{\mu_2}, \boldsymbol{\mu_3}) = J(\boldsymbol{u}) + \sum_{i=1}^{N_c} \mu_{1_i} h_{1_i}(\boldsymbol{u}) + \sum_{i=1}^{N_u-1} \mu_{2_i} g_{1_i}(\boldsymbol{u}) + \sum_{i=1}^{N_u-1} \mu_{3_i} g_{2_i}(\boldsymbol{u}) \quad (5)$$

with Lagrange multiplier vectors $\boldsymbol{\mu_1}$, $\boldsymbol{\mu_2}$, $\boldsymbol{\mu_3}$

Transformation of the original problem to its unconstrained form

$$\bar{J}(\boldsymbol{u}) = J(\boldsymbol{u}) + \mu \sum_{i=1}^{N_c} h_{1_i}^2(\boldsymbol{u}), \quad (6)$$

Objective – to solve an unconstrained problem:

$$\boldsymbol{u}(k) \triangleq \min_{\boldsymbol{u}} \quad \bar{J}(\boldsymbol{u}), \quad (7)$$

where μ is suitably large constant

If $\boldsymbol{u}_\mu$ is a solution of the problem (5), and excluding inequality constraints (4d) and (4e) it can be shown that as $\mu \to \infty$ there obtains $\boldsymbol{u}_\mu \to \boldsymbol{u}^*$, where $\boldsymbol{u}^*$ is solution of (4)

Inequality constraints handling

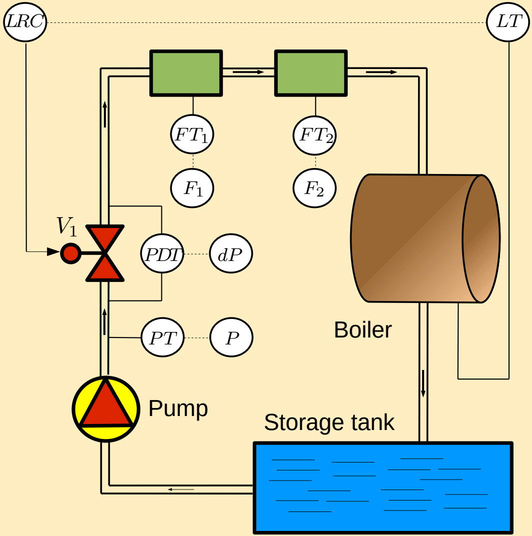
Solution projection

for $i := 0$ **to** $N_u - 1$ **do**
    if $u(k+i) > \overline{u}$ **then**
        $u(k+i) := \overline{u}$
    else if $u(k+i) < \underline{u}$ **then**
        $u(k+i) := \underline{u}$
    end if
end for

- Projection of the solution onto a feasible region
- This simple solution can deteriorate the optimal solution but quarantees that inqequality constraints stay satisfied

 k.patan@issi.uz.zgora.pl

Tank unit

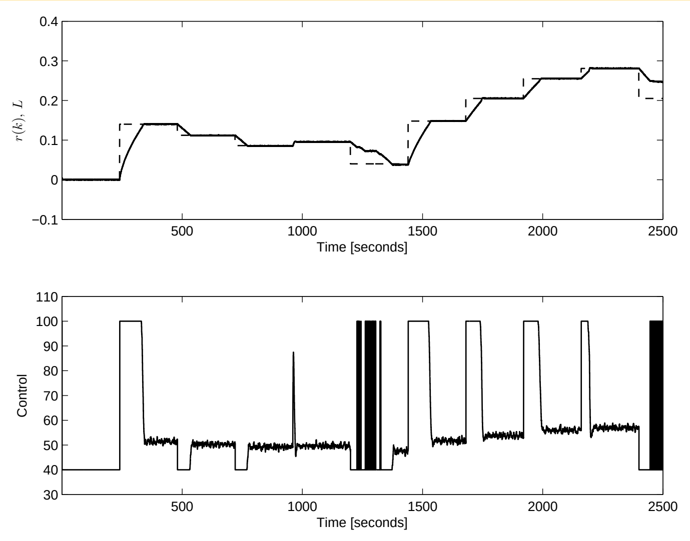


CV – control value
dP – pressure difference on the valve V_1
P – pressure before the valve V_1
$F1$ – flow (electromagnetic flowmeter)
$F2$ – flow (Vortex flowmeter)
L – water level in the boiler

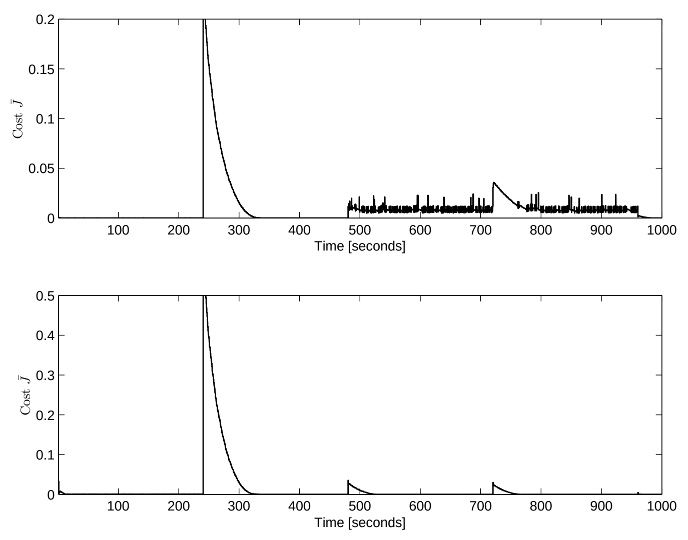
Experiments

- best performing model (NNOE): one input (CV), one output (L), 7 tangensoidal neurons in the hidden layer, one linear output neuron, the number of input delays $n_b = 2$ and the number of delayed outputs $n_a = 2$
- prediction horizon $N_p = 15$
- control horizon $N_u = 2$
- constraint horizon $N_c = 3$
- penalty factor $\rho = 10^{-6}$
- upper control bound $\overline{u} = 100$
- lower control bound $\underline{u} = 40$

Process output (solid) and reference signal (dashed) (upper graph), the control signal (lower graph)



Evolution of the cost function without terminal constraint (upper graph), with terminal constraints (lower graph).



Concluding remarks

- The presented neural network based MPC quarantees the stable work of the control system
- The proposed numerical solution is very simple to implement and no time consuming
- Unfortunately the presented solution can cause a ringing effect in the control directly caused by control projection onto the feasible region
- This effect can be eliminated using a more robust constrained optimization procedure