Functional Adaptive Controller for MIMO Systems with Dynamic Structure of Neural Network

Ladislav KRÁL, Ivo PUNČOCHÁŘ, Jindřich DUNÍK

Department of Cybernetics
University of West Bohemia
Pilsen, Czech Republic

10^{th} International PhD Workshop on Systems and Control
Young Generation Viewpoint
Hluboká nad Vltavou, September 22-26, 2009

Contents

Introduction

Overview

- Adaptive control of nonlinear stochastic systems
- Modeling of nonlinear systems using neural networks (e.g. radial basis function, multilayer perceptron)
- Functional adaptive control – nonlinear functions and parameters of the system are unknown
- Basic approaches to adaptive control
  1. certainty equivalence control
  2. cautious control
  3. DUAL CONTROL
     - ➠ estimation of the neural network parameters
     - ➠ structure optimization of the neural network
     - ➠ dual control design

Introduction – approaches, motivation and goal

Dual control design

- Several different dual control methods: Inovation Dual Control (IDC), Bicriterial Dual Control (BDC), Wide-sense dual control, ...
- Linear systems with unknown parameters are mostly considered
- Only IDC (*Fabri and Kadirkamanathan '01*) and BDC (*Šimandl '05*) were used for nonlinear systems with unknown functions where BDC achieves better results
- Both these works on the functional adaptive control are limited to single-input single-output (SISO) systems and functional adaptive control for multivariable stochastic systems has not been studied yet
  ➠ motivation

Goal

To design a functional adaptive controller for a nonlinear stochastic discrete-time MIMO system where a neural network with dynamically optimized structure serves as a model of a system

Problem statement

Nonlinear stochastic discrete-time system

$$\boldsymbol{y}_k = \boldsymbol{f}(\boldsymbol{x}_{k-1}) + \boldsymbol{G}(\boldsymbol{x}_{k-1})\boldsymbol{u}_{k-1} + \boldsymbol{e}_k,$$

vector $\boldsymbol{f}(\boldsymbol{x}_{k-1})$ and matrix $\boldsymbol{G}(\boldsymbol{x}_{k-1})$ contain unknown nonlinear functions

$\boldsymbol{x}_{k-1} \triangleq [\boldsymbol{y}_{k-p}^T, \ldots, \boldsymbol{y}_{k-1}^T, \boldsymbol{u}_{k-1-s}^T, \ldots, \boldsymbol{u}_{k-2}^T]^T$ is known measurable state

$\boldsymbol{y}_k = [y_k^{(1)}, \ldots, y_k^{(n)}]^T$ is output

$\boldsymbol{u}_k = [u_k^{(1)}, \ldots, u_k^{(m)}]^T$ is input

$\boldsymbol{e}_k = [e_k^{(1)}, \ldots, e_k^{(n)}]^T$ is additive white noise, pdf $\mathcal{N}\{\boldsymbol{0}, \boldsymbol{\Xi}\}$

Bicriterial dual controller

$$\boldsymbol{u}_k = \boldsymbol{h}_k(\boldsymbol{r}_{k+1}, \boldsymbol{I}_k)$$

output $\boldsymbol{y}_k$ should follow reference signal $\boldsymbol{r}_k = [r_k^{(1)}, \ldots, r_k^{(n)}]^T$

$\boldsymbol{I}_k$ contains information received up to time k

Bicriterial dual controller – basic idea

The bicriterial dual controller design is based on two separate criteria. Each of those criteria introduces one of opposing aspects between estimation and control: **caution** and **probing**.

The caution control component

$$J_k^c = E\Big\{(\boldsymbol{y}_{k+1} - \boldsymbol{r}_{k+1})^T \boldsymbol{Q}_{k+1} (\boldsymbol{y}_{k+1} - \boldsymbol{r}_{k+1}) + \boldsymbol{u}_k^T \boldsymbol{S}_{k+1} \boldsymbol{u}_k | \boldsymbol{I}_k\Big\},$$

$$\boldsymbol{u}_k^c = \underset{\boldsymbol{u}_k}{\operatorname{argmin}}\ J_k^c$$

The probing control component

$$J_k^a = -E\Big\{(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1})^T \boldsymbol{W}_{k+1} (\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1}) | \boldsymbol{I}_k\Big\}$$

$$\Omega_k = [\boldsymbol{u}_k^c - \boldsymbol{\delta}_k, \boldsymbol{u}_k^c + \boldsymbol{\delta}_k]$$

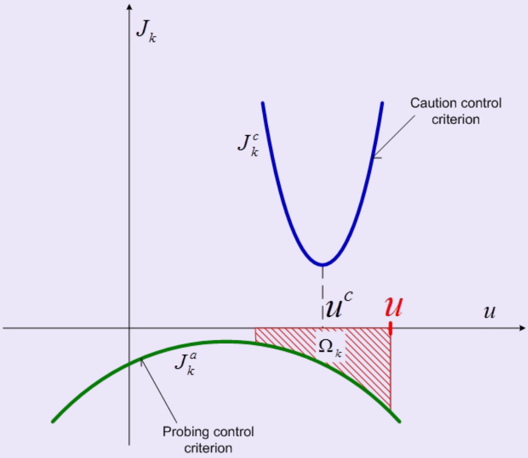
$$\boldsymbol{\delta}_k = \boldsymbol{\eta}\ \mathrm{Tr}(\boldsymbol{P}_{k+1|k})$$

The final control

$$\boldsymbol{u}_k = \underset{\boldsymbol{u}_k \in \Omega_k}{\operatorname{argmin}}\ J_k^a.$$

Bicriterial dual controller – graphical interpretation

Graphical interpretation for single input systems



Bicriterial dual controller – cont'd

Bicriterial dual controller

- Computational demands
  - Caution component – unconstrained minimization of convex function (analytical computation)
  - Probing component – constrained minimization of concave function (vertex enumeration)
- $\boldsymbol{u}_k = \boldsymbol{h}_k(\boldsymbol{\eta}, \boldsymbol{r}_{k+1}, \hat{\boldsymbol{\theta}}_{k+1|k}, \boldsymbol{P}_{k+1|k}) \Rightarrow \boldsymbol{\eta}$ - *designer parameter*

  $\Rightarrow \boldsymbol{r}_{k+1}$ - *known variables*

  $\Rightarrow \hat{\boldsymbol{\theta}}_{k+1|k}, \boldsymbol{P}_{k+1|k}$ - *estimation*

Neural network – model choice

Model of the system

- The unknown nonlinear functions $\boldsymbol{f}(\boldsymbol{x}_{k-1})$ and $\boldsymbol{G}(\boldsymbol{x}_{k-1})$ are approximated by Multi-Layer Perceptron (MPL) networks ➠ **model**
- There are various structures of neural network for MIMO systems
- Recommendation ➠ two neural networks $\hat{f}^{(i)}$, $\hat{g}^{(i\cdot)}$ for each of n outputs $y_k^{(i)}$ of the system

$$\begin{aligned}
\hat{\boldsymbol{y}}_k &= \hat{\boldsymbol{f}}(\boldsymbol{x}_{k-1}, \boldsymbol{w}_k^f, \boldsymbol{c}_k^f) + \hat{\boldsymbol{G}}(\boldsymbol{x}_{k-1}, \boldsymbol{w}_k^g, \boldsymbol{c}_k^g)\boldsymbol{u}_{k-1} \\
\hat{y}_k^{(i)} &= \hat{f}^{(i)} + \textstyle\sum_{j=1}^{m} \hat{g}^{(ij)} u_{k-1}^{(j)}, \quad \text{for } i = 1, \ldots, n \\
\hat{f}^{(i)} &= (\boldsymbol{c}_k^{f_i})^T \phi^{f_i}(\boldsymbol{x}_{k-1}^a, \boldsymbol{w}_k^{f_i}) \\
\hat{g}^{(ij)} &= (\boldsymbol{c}_k^{g_{ij}})^T \phi^{g_i}(\boldsymbol{x}_{k-1}^a, \boldsymbol{w}_k^{g_i})
\end{aligned}$$

$$\boldsymbol{\Theta}_k = \left[(\boldsymbol{c}_k^f)^T, (\boldsymbol{w}_k^f)^T, (\boldsymbol{c}_k^g)^T, (\boldsymbol{w}_k^g)^T\right]^T \Rightarrow \hat{\boldsymbol{\Theta}}_{k+1|k},\ \boldsymbol{P}_{k+1|k} = ?$$

Neural network – parameter estimation

Estimation model

- Neural network can be rewritten into state space estimation model

$$\boldsymbol{\Theta}_{k+1} = \boldsymbol{\Theta}_k$$
$$\boldsymbol{y}_k = \hat{\boldsymbol{f}}(\boldsymbol{x}_{k-1}, \boldsymbol{w}_k^f, \boldsymbol{c}_k^f) + \hat{\boldsymbol{G}}(\boldsymbol{x}_{k-1}, \boldsymbol{w}_k^g, \boldsymbol{c}_k^g)\boldsymbol{u}_{k-1} + \boldsymbol{e}_k$$

- The measurement equation is nonlinear
- It is possible to use non-linear estimation methods - Extended Kalman Filter (EKF)
- Prior information about parameters given by pdf $\mathcal{N}\{\hat{\boldsymbol{\Theta}}_{0|-1}, \boldsymbol{P}_{0|-1}\}$

Neural network – dynamic structure optimization

Optimization of the neural network structure is performed on-line by pruning insignificant connections from the neural network

Three basic steps of the optimization algorithm

- Check whether the neural network is already trained using prediction error ε_k

$$\Delta_k = \left| \frac{1}{k+1} \Sigma_{t=0}^{k} \varepsilon_t^2 - \frac{1}{k} \Sigma_{t=0}^{k-1} \varepsilon_t^2 \right|$$

- If the prediction error is steady sort the parameters of the neural network according their "significancy" E_i

$$E_i = \frac{\hat{\theta}_i^2}{P_i}$$

- Try to set to zero (i.e. leave out) as many insignificant parameters as possible

$$T = \frac{1}{k+1} (\hat{\boldsymbol{\Theta}}_{[1,N]} - \hat{\boldsymbol{\Theta}})^T \mathbf{P}^{-1} (\hat{\boldsymbol{\Theta}}_{[1,N]} - \hat{\boldsymbol{\Theta}})$$

Bicriterial dual control algorithm

Algorithm

At the beginning

- initialization

At each time instant k

- **step 1:** measurement of the output $\boldsymbol{y}_k$ of the system
- **step 2:** estimation of neural network parameters by EKF
- **step 2:** dynamic optimization of neural network structure
- **step 3:** generation of input $\boldsymbol{u}_k$ using bicriterial dual approach

$k \rightarrow k + 1$

Numerical example

Benchmark system with two inputs and two outputs

$$y_k^{(1)} = \frac{0.7 y_{k-1}^{(1)} y_{k-2}^{(1)}}{1 + (y_{k-1}^{(1)})^2 + (y_{k-2}^{(2)})^2} + \frac{0.1 u_{k-1}^{(2)}}{1 + 3(y_{k-2}^{(1)})^2 + (y_{k-1}^{(2)})^2} + u_{k-1}^{(1)} + 0.25 u_{k-2}^{(1)} + 0.5 u_{k-2}^{(2)} + e_k^{(1)},$$

$$y_k^{(2)} = \frac{0.5 y_{k-1}^{(2)} \sin y_{k-2}^{(2)}}{1 + (y_{k-1}^{(2)})^2 + (y_{k-2}^{(1)})^2} + 0.5 u_{k-2}^{(2)} + 0.3 u_{k-2}^{(1)} + u_{k-1}^{(2)} \left(0.1 u_{k-2}^{(2)} - 1.5 \right) + e_k^{(2)},$$

Two controllers were compared

- Bicriterial dual controller with static structure (BDC stat)
- Bicriterial dual controller with dynamic structure (BDC dynam)

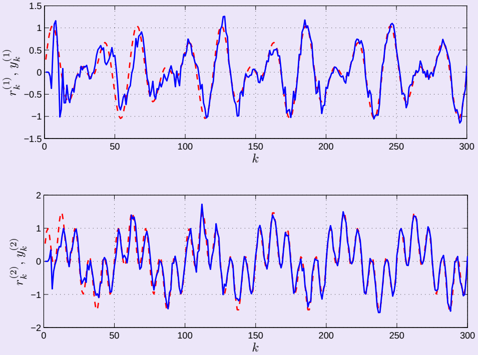
The results - numerical interpretation

The quality of control is measured by the mean of sums of square errors between reference value $r_{kj}^{(i)}$ and system output $y_{kj}^{(i)}$ over 100 trials: $\hat{V} = \frac{1}{100}\sum_{i=1}^{2}\sum_{j=1}^{100}\sum_{k=1}^{200}(y_{kj}^{(i)} - r_{kj}^{(i)})^2$

	$\hat{V}$	$\mathrm{cov}(\hat{V})$	$n\theta$	time [s]
BDC stat	27.8	15.8	590	57.2
BDC dynam	26.5	18.2	112	45.5

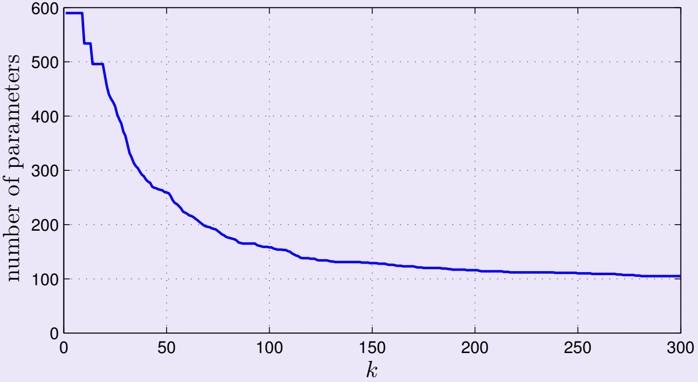
The results – graphical interpretation

Typical output of the system (output - blue and reference - red)



The results - graphical interpretation (cont'd)

Number of the neural network parameters



Conclusion

☞

- ✯ The bicriterial dual controller for non-linear stochastic MIMO systems was designed.
- ✯ The model of the system is given by the multilayer perceptron network.
- ✯ The extended Kalman filter was applied for the on-line parameter estimation of the derived estimation model.
- ✯ In order to avoid the problem with choice of the neural network structure, an on-line dynamic structure optimization algorithm of the network was utilized.
- ✯ The proposed dual adaptive controller with dynamic structure has lower computational demands and comparable control quality in comparison with controller that utilizes static structure of the neural network.