Available online at www.sciencedirect.com
ScienceDirect
Engineering Applications of Artificial Intelligence 21 (2008) 1035–1055
Engineering Applications of
ARTIFICIAL INTELLIGENCE
www.elsevier.com/locate/engappai

Nonlinear system identification: From multiple-model networks to Gaussian processes

Gregor Gregorčič[a,*], Gordon Lightbody[b]

[a]AVL List GMBH, Hans-List-Platz 1, A-8020 Graz, Austria
[b]Department of Electrical Engineering, University College Cork, Ireland

Received 25 March 2005; received in revised form 24 October 2007; accepted 8 November 2007
Available online 28 January 2008

Abstract

Neural networks have been widely used to model nonlinear systems for control. The curse of dimensionality and lack of transparency of such neural network models has forced a shift towards local model networks and recently towards the nonparametric Gaussian processes approach. Assuming common validity functions, all of these models have a similar structure. This paper examines the evolution from the radial basis function network to the local model network and finally to the Gaussian process model. A simulated example is used to explain the advantages and disadvantages of each structure.

Keywords: Nonlinear system identification; Radial basis function network; Local model network; Network structure; Gaussian processes

1. Introduction

On a global industrial scale it is necessary to increase efficiency and thereby to reduce operating costs. It is also important to minimise environmental and health impacts, through new industrial approaches, as well as an enhancement of resources. To achieve this goal, it is essential to upgrade existing approaches and introduce new, breakthrough technologies, based on high-tech artificial intelligence solutions. Increased automation of industry will improve consistency and quality, which will consequently reduce the error margins in production.

This will improve efficiency, cut production costs and reduce pollution. The main contributing factor towards the improved automation of industry is the development of new advanced control strategies.

Many of the processes in the chemical, petrochemical, food processing, pharmaceutical and biochemical industry are nonlinear in nature. The nonlinearities arise from the dynamics in chemical reactions, thermodynamic relationships, etc. This effectively means that such processes are relatively complex and difficult to control. Control approaches in which conventional techniques, designed for linear systems are applied, can result in a high degree of uncertainty being incorporated within the design procedure. This leads to robustness problems, in which the controlled system can only operate satisfactory within a restricted region. Conventional PID controllers must be conservatively tuned in order to ensure closed-loop stability over the full range of operating conditions (Seborg, 1994a, b). In turn this may mean that the plant will not operate at high efficiency.

To improve efficiency and flexibility more advanced techniques, such as model-based control, have been proposed. If the nonlinear system can be accurately modelled (Billings, 1980), then model-based techniques such as internal model control (Economou et al., 1986; Hunt and Sbarbaro, 1991; Nahas et al., 1992; Brown et al., 1997) and model predictive control (Richalet, 1993) can be applied to improve the closed-loop performance. In these strategies, the model of the plant is included directly within the control structure. In these schemes the efficiency and quality of the control is strongly related to the accuracy of the model which represents the plant.

*Corresponding author.
E-mail address: gregor.gregorcic@avl.com (G. Gregorčič).

0952-1976/$ - see front matter
doi:10.1016/j.engappai.2007.11.004

Many modelling techniques which can be applied in model-based control strategies have been proposed over the past years. Due to the complexity of industrial processes first-principle models and their analytical approximations cannot usually be efficiently applied and hence a more flexible type of model needs to be utilised.

Neural networks (Narendra and Parthasarathy, 1990; Pollard et al., 1992; DeVeaux et al., 1993) and fuzzy logic-based models (Takagi and Sugeno, 1985; Hunt et al., 1996a; Babuška and Verbruggen, 2003) are widely used in industrial applications (Morris et al., 1994; Aoyama et al., 1995; Lightbody and Irwin, 1997). Their characteristics of learning and nonlinear representation have been employed in many control applications (Hunt and Sbarbaro, 1991; Jang and Sun, 1995; Abonyi, 2003). These types of models are known as black-box models which are mainly identified using input/output data. The main difficulty related to such black-box models is the lack of transparency, which means that the resulting model does not provide any physical knowledge about the underlying system. Consequently, this lack of transparency hinders the incorporation of prior engineering knowledge about the system into the model. As well as a lack of transparency, both neural network and fuzzy models have proven to be less useful with large data sets and the multi-dimensional operating space, which is usually the case in industry. This is known as the curse of dimensionality.

Multiple-model approaches for the modelling and control of nonlinear systems were proposed in the last decade (Murray-Smith and Johansen, 1997) and have significantly improved the transparency of the models and reduced the curse of dimensionality. Here a global process model is formed by the combination or blending of a number of local models, which have been identified over the operating range of the process (Johansen and Foss, 1993, 1995, 1997). It is obvious that this type of network will be able to accurately model the system only close to the points, where the local models have been identified. It is to be hoped that this type of network can model accurately the system close to the points where the local models were identified. However, in most cases, for practical reasons, local models used in control system applications are identified only around equilibrium points and away from those points the accuracy of the network can decrease rapidly due to unmodelled off-equilibrium dynamics. The off-equilibrium dynamics problem has shifted research towards investigation of probabilistic Bayesian-based modelling methods for use in control applications (Murray-Smith et al., 1999).

The Gaussian process (GP) prior as a Bayesian approach for the modelling of nonlinear systems has been in use in the statistics community for many years (von Mises, 1964). Until recently this technique was ignored by most of the engineering community. This GP approach offers the ability to model the system better in off-equilibrium regions, where a limited amount of data is available. The idea is to place a prior on the space of functions, instead of tuning a large number of network parameters. The number of tuneable parameters is dramatically reduced compared to the number of parameters of a conventional neural network. The strongest advantage of the GP approach in comparison with traditional models is that the GP model provides an analytic expression of the model uncertainty. This makes the GP model very attractive for model-based control strategies.

This paper explains the similarities and differences between radial basis function network (RBFN), local model network (LMN), fuzzi and GP prior models. Assuming Gaussian radial validity functions, it is shown that the general structures of these types of models are closely related. The discussion focuses on various types of structures, training algorithms, number of adjustable parameters, transparency, the importance of the selection of the scheduling vector and the problem of off-equilibrium dynamics. An examination of the interpolation between the models in the LMN is given along with a discussion on the choice of the actual realisation of the network. The advantages and drawbacks of each type of model are highlighted using a simulated example.

The work presented here also aims to create a link between the probabilistic Bayesian-based modelling methods and well-established parametrical modelling techniques. An overview of the GP prior as a Bayesian-based modelling alternative is given. A relationship between the RBFN, multiple models and GPs is explained in the light of control engineering. At this point it is important to emphasise that the presented modelling techniques are not limited for use only in control systems but can also be successfully applied in a much broader range of engineering as well as nonengineering applications.

2. Multiple-model networks

Research in the area of multiple models has rapidly expanded over the past years, with numerous paradigms developed, training algorithms proposed and applications demonstrated. An extensive overview of existing multiple-model approaches to modelling and control is given in Murray-Smith and Johansen (1997).

In general, a nonlinear dynamic system can be represented as a multi-dimensional surface above the operating space $\Re^{\mathrm{d}}$. In multiple-modelling approaches, it is assumed that the surface can be modelled using a combination of weighted models, combined to form a global model (Murray-Smith and Johansen, 1997). This modelling structure could be represented by

$$\hat{y}(k+1) = \sum_{i=1}^{M} \varphi_i(k)\mathscr{M}_i(k), \tag{1}$$

where $\hat{y}(k+1)$ is the one-step ahead predicted output, M is the number of local models, $\varphi_i(k)$ is the ith validity function and $\mathscr{M}_i(k)$ is the ith local model output. For visualisation convenience, it is assumed in this paper that the nonlinear