Computers and Chemical Engineering 108 (2018) 31–46

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Enhancement of modifier adaptation scheme via feedforward decision maker using historical disturbance data and deep machine learning

Dong Hwi Jeong, Jong Min Lee*

School of Chemical and Biological Engineering, Institute of Chemical Processes, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea

ARTICLE INFO

Article history:
Received 16 February 2017
Received in revised form 18 August 2017
Accepted 18 August 2017
Available online 24 August 2017

Keywords:
Modifier adaptation
Model-plant mismatch
Deep neural network
KKT conditions

ABSTRACT

Most advanced processes struggle to reduce the production cost under constraints. For this, an iterative optimization method called modifier adaptation has been utilized due to its ability to ensure the necessary conditions of optimality even under model-plant mismatch. However, the optimization performance may be degraded by the disturbance which may significantly change the true optimum. In this study, a feedforward decision maker is designed to deal with disturbances in advance and compensate the limitation of feedback scheme of the conventional modifier adaptation. It is constructed by historical data and deep machine learning, and combined with the modifier adaptation. When disturbances occur, the decision maker provides an initial point close to the true optimum by exploiting the historical data. As the information is accumulated, a better initial point for modifier adaptation is obtained. Constrained optimization of numerical example and run-to-run bioprocess are illustrated to validate the utility of the proposed method.

1. Introduction

Chemical and biological processes are characterized by complex dynamics with constraints and many possible operating modes to handle (Marchetti et al., 2009). In operating these processes, enhancing productivity while satisfying the constraints is one of the key issues in global competitions (François and Bonvin, 2013a,b; Bunin et al., 2011). Thus various process optimization techniques have been introduced and applied for this purpose. Typical process optimization uses mathematical models and numerical procedures to compute an optimal solution (Chachuat et al., 2009). However, the time and resources required to develop an accurate process model incurs a significant cost in general (Jia et al., 2015). Moreover, uncertainties like exogeneous disturbances and their unknown effects on the process make modelling more challenging and often lead to model-plant mismatch (Camacho et al., 2007). If this mismatch occurs, open-loop optimization cannot guarantee optimal performance and the constraints must be revised by additional control layer in most cases (Camacho et al., 2015; Gao and Engell, 2005). Thus, the techniques that use measurements to offset the uncertainty and lead the process to real plant's optimum, such as real time optimization (RTO), have recently received much attention (François and Bonvin, 2013a,b).

A classical two-step approach updates model parameters and performs optimization (François and Bonvin, 2013a,b). The idea is to repeatedly estimate the model parameters and use the updated model to compute a new operating point via optimization. In this method, the updated model is expected to yield a better description of the real plant near the current operating mode (Marchetti et al., 2009). However, the two-step approach works well only if (1) there exists little structural model-plant mismatch and (2) the changes in operating conditions provide sufficient excitations for parameter estimation. However, these are difficult to satisfy in practice (Marchetti et al., 2009).

Instead of updating model parameters, discrepancy between model and plant can be directly reflected in the model. A method called integrated system optimization and parameter estimation (ISOPE) incorporates a gradient modification term into the objective function to achieve convergence to a Karush–Kuhn–Tucker or KKT point that satisfies necessary condition of plant's optimality (Gao and Engell, 2005). Chachuat et al. (2009) recently introduced a different way to compensate model-plant mismatch by adding zero and first-order modification terms to constraints as well as to objective. This class of approaches are called modifier adaptation, where the necessary

* Corresponding author.
E-mail addresses: cacaca26@snu.ac.kr (D.H. Jeong), jongmin@snu.ac.kr (J.M. Lee).

http://dx.doi.org/10.1016/j.compchemeng.2017.08.005

conditions of optimality of the plant and model can match (Jia et al., 2015). The main advantage of modifier adaptation lies in its proven ability to converge to KKT point of a plant under model-plant mismatch arising from structural mismatch as well as parametric uncertainty (François and Bonvin, 2013a,b). There have been successful applications of modifier adaptation schemes to several case studies involving chemical and biological processes under uncertainty, where true plant optima could be achieved against model-plant mismatches (Jia et al., 2015, 2016).

In real processes, however, disturbances may occur frequently and their magnitudes can be large enough to shift the process far from an optimal point. Although the conventional modifier adaptation can handle disturbances to a certain extent, it may not adequately deal with the large changes of optimal point due to the disturbance because it adjusts the model-based optimization using the plant output measurements in a post-processing fashion like feedback control. Therefore, such disturbances need to be handled proactively with a feedforward structure to provide optimal operating points in a timely fashion.

With recent advances in sensor technology, a large amount and different kinds of data are available and bring tremendous opportunities for various applications. Moreover, it is also shifting paradigms in many areas towards data-driven discovery (Chen and Lin, 2014). As the amount of available data increases, machine learning that utilizes data and provides insightful analysis has become a key technique (Chen and Lin, 2014). It has already been utilized in various fields such as content filtering on social networks and recommendation system on e-commerce websites (LeCun et al., 2015). It can also be used for identifying images, transcribing sound data, matching news with users' interests (Michalski et al., 2013).

Various types of machine learning techniques have been introduced and improved for these purposes. The first one is support vector machine or SVM. It has advantages of low computational cost and accessible optima due to convex quadratic optimization. It is proven that SVM can effectively tackle problems with small samples, and has been utilized in many fields such as artificial intelligence, pattern recognition and machine learning (Shang et al., 2014). However, there exists a remaining problem that the computational complexity grows exponentially with the number of training samples (Shang et al., 2014). Artificial neural networks or ANN is an overlapped structure of neural units which mimics the way of a brain to solve problems with large clusters of neurons connected by axons. The most common types of ANN are multi-layer perceptron and radial basis function networks (Hagan et al., 1996). It has been popularly used for function approximation and pattern recognition (Hagan et al., 1996). Despite the successful applications of SVM and ANN, however, they have a disadvantage that the interpretation capability still remains a major problem for complex systems (Bengio et al., 2007; LeCun et al., 2015; Shang et al., 2014). This limitation is caused by no allowance for latent variable subspace (Shang et al., 2014). Especially, ANN suffers from an uncontrolled convergence speed and local optima. Moreover, parameters of ANN with more than two layers are difficult to optimize using traditional gradient descent schemes (Bengio et al., 2007). In order to overcome these limitations, it was suggested to pre-process data by unsupervised learning that eliminates the noise and possibly reduces its searching dimensionality (Hinton et al., 2006; Bengio et al., 2007). This has opened up a new area of machine learning, called deep learning.

In contrast to the conventional machine learning techniques having shallow architectures, the deep learning refers to a class of machine learning that uses supervised and/or unsupervised strategies to automatically learn hierarchical representations in deep architectures (Chen and Lin, 2014). One of the earliest deep learning schemes, the deep belief network by Hinton, can be summarized as follows: first, pre-train one layer at a time by unsupervised data and restricted Bolzman machine to preserve information from the input; then do fine-tuning the whole network with respect to the criterion of interest (Hinton et al., 2006). This approach was shown to perform better than those trained exclusively with back-propagation (Bengio et al., 2007). In this way, machine learning can contain a deep structure of more than two hidden layers, which allows for expressing complex relationships (Srivastava et al., 2014). Since the deep belief network was introduced, deep learning technology has evolved to an economically viable level in many aspects and is being utilized in many industrial applications including MNIST handwriting challenge (Ciregan et al., 2012), face detection (Sun et al., 2014), speech recognition (Hinton et al., 2012), natural language processing (Collobert and Weston, 2008) and soft sensor (Shang et al., 2014). Global companies have heavily invested in research related to deep learning; Apple's Siri, Google's translator, street view and image search engine, Android's voice recognition, Facebook's recommended friends and news feed system, etc. (Chen and Lin, 2014).

Thanks to such advances in machine learning technology, we propose a modifier adaptation scheme combined with feedforward decision maker to handle the issue of disturbances. This utilizes historical data of process and deep learning techniques. In this decision maker, the disturbance information in the past serves as an input and the suboptimal or optimal strategy against the disturbance serves as an output. A starting point for modifier adaptation, expected to be near the plant optimal point, is determined in a feedforward manner for each of disturbances. By starting the update near the plant optimal point with the feedforward decision maker, optimization performance is expected to improve compared with conventional modifier adaptation which handles disturbances only in a post-processing sense. In addition, continuous updates of learning model allows for prompt and optimal handling of the same disturbance experienced before.

The rest of the paper is organized as follows: Section 2 introduces a problem of model-plant mismatch, KKT conditions and the modifier adaptation. Section 3 deals with a starting point issue of the iterative optimization, deep learning techniques and proposes a new modifier adaptation combined with the feedforward decision maker. Section 4 presents illustrative examples of numerical and bio-process optimization problems and shows the validity of the proposed approach. Finally, Section 5 provides concluding remarks.

2. Modifier adaptation

2.1. Model-plant mismatch and suboptimality

Consider the plant equation based optimization:

$$\begin{aligned} u_p^* := \ & \operatorname{argmin}_u \ \Phi_p(u) \\ & \text{s.t. } G_p(u) \leq 0 \end{aligned} \tag{1}$$