Journal of Process Control 30 (2015) 69–79

Contents lists available at ScienceDirect

Journal of Process Control

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

Using horizon estimation and nonlinear optimization for grey-box identification☆

Alf J. Isaksson[a,b,*], Johan Sjöberg[b,c], David Törnqvist[d], Lennart Ljung[b], Manon Kok[b]

[a] ABB AB, Corporate Research, SE-721 78 Västerås, Sweden
[b] Dept. Electrical Eng., Linköping University, SE-581 83 Linköping, Sweden
[c] Volvo Construction Equipment, SE-631 85 Eskilstuna, Sweden
[d] SenionLab AB, Teknikringen 7, 583 30 Linköping, Sweden

ARTICLE INFO

Article history:
Received 19 May 2014
Received in revised form
23 December 2014
Accepted 24 December 2014
Available online 20 February 2015

Keywords:
System identification
State estimation
Parameter estimation
Optimization
Nonlinear systems
Kalman filtering
Moving horizon estimation
Model predictive control

ABSTRACT

An established method for grey-box identification is to use maximum-likelihood estimation for the nonlinear case implemented via extended Kalman filtering. In applications of (nonlinear) model predictive control a more and more common approach for the state estimation is to use moving horizon estimation, which employs (nonlinear) optimization directly on a model for a whole batch of data. This paper shows that, in the linear case, horizon estimation may also be used for joint parameter estimation and state estimation, as long as a bias correction based on the Kalman filter is included. For the nonlinear case two special cases are presented where the bias correction can be determined without approximation. A procedure how to approximate the bias correction for general nonlinear systems is also outlined.

1. Introduction

This paper ultimately deals with parameter estimation in nonlinear models. What triggers our interest is nonlinear model predictive control [9]. While linear model predictive control has long been an established industrial area, nonlinear MPC has only found industrial applications more recently, see for example [8,22,23]. Partly, the limited applicability is related to modelling and state estimation challenges. Since in MPC an optimization framework is already in place, it is natural to deploy this also for the state estimation. This results in so-called moving horizon estimation (MHE).

This paper will focus on some aspects that are important when MHE of states is combined with parameter estimation of model parameters. These parameters could be part of linear black box model parameterizations, but could also be selected unknown parameters of physical significance in a linear or nonlinear model. Estimation of such unknown parameters is known as *Grey-box-identification*, see e.g. [3] or [18]. Nonlinear Grey-box identification is typically approached in an off-line setting by maximum likelihood techniques. If process noise is present in the model, this requires the difficult nonlinear prediction problem to be handled. Typical approaches for this are approximate solutions with extended Kalman filter, as detailed in [3] or with particle filtering, e.g. [28].

We study here a different approach, namely to extend the common moving horizon state estimation solution [1,6,7,14,19,20,24,25] with parameter estimation. Such an approach has also been discussed in, for instance, [17,30]. This tempting approach has some fallacies, however, since biased parameter estimates may result if proper attention is not paid to the criterion formulation. We explain the root of this effect and show how it can be handled in the linear model case. That also suggests a way to treat the problem in the general nonlinear case. We illustrate the approach with an application to a nonlinear drumboiler model.

2. Model predictive control

As the name model predictive control indicates a crucial element of an MPC application is the model on which the control is based.

☆ This paper is an extension of a keynote paper, [13], presented at DYCOPS 2013 by the first author, which in turn refers to a conference paper, [12], presented at SSS'09 by the first four authors.

* Corresponding author at: ABB AB, Corporate Research, SE-721 78 Västerås, Sweden. Tel.: +46 705323236.
E-mail address: alf.isaksson@se.abb.com (A.J. Isaksson).

http://dx.doi.org/10.1016/j.jprocont.2014.12.008
0959-1524/

Therefore, before a controller can be implemented a model has to be established. There are two main alternatives available for obtaining the model.

- Deriving a model from first principles using laws of physics, chemistry, etc.; so-called white-box modelling
- Estimating an empirical model from experimental data; black-box modelling

In general, a white-box model becomes a DAE

$$0 = f(\dot{x}(t), x(t), u(t))$$
$$y(t) = h(x(t), u(t))$$

where $y(t)$ denotes the measured process variables, and $u(t)$ the manipulated variable, i.e. the output of the MPC. Finally the internal variable $x(t)$ is what is usually referred to as the state of the system.

A black-box model on the other hand, is typically linear, but most often also discrete in time

$$x_{k+1} = Ax_k + Bu_k$$
$$y_k = Cx_k$$

Here the integer k denotes the k:th time index for which the signal value is available, i.e. at time kT_s, where T_s is the sampling interval. Hence, we have for example $x_k = x(kT_s)$.

The core of MPC is optimization. In each iteration of the control, i.e. any time a new measurement is collected, two optimization problems have to be solved (both using the model as an equality constraint); one using past data to estimate the current state vector $x(t)$ and one to optimize the future control variables. When solving the forward optimization problem a number of future values of the manipulated variables are calculated. However, only the values at the first time instant are transmitted to the underlying process. At the next time instant the optimizations are repeated, with the optimization windows shifted one time step. This is known as receding horizon control, and is in fact what makes this a feedback control method. Performing optimization just once would correspond to open-loop control. The emphasis of this paper is on the second step – the state estimation – which will be presented in some more detail in the next subsection.

2.1. State estimation

For the state estimation the optimization target is to obtain the best estimate of the internal variable x using knowledge of y and u, to be used as starting point for the forward optimization. This can be done using a Kalman filter (for an old classic see [2]) – or if the model is nonlinear an extended Kalman filter see [15] – where stochastic modelling of the process and measurement noises is applied. A Kalman filter is a recursive method, meaning that it takes only the most recent values of y_k and u_k to update the previous estimate $\hat{x}_{k-1}$ to obtain the new $\hat{x}_k$. Hence, it does not actually solve an optimization problem on-line. Kalman filtering is done in a statistical framework by adding process noise and measurement noise to the discrete-time state space system given in the previous section.

$$x_{k+1} = Ax_k + Bu_k + w_k \tag{1a}$$
$$y_k = Cx_k + v_k \tag{1b}$$

where w_k and v_k are white Gaussian noises with covariance matrices Q and R respectively. Since we want to use certain quantities in the calculation later, the complete set of Kalman filter equations is given below:

$$S_k = CP_{k|k-1}C^T + R \tag{2a}$$
$$K_k = P_{k|k-1}C^T S_k^{-1} \tag{2b}$$

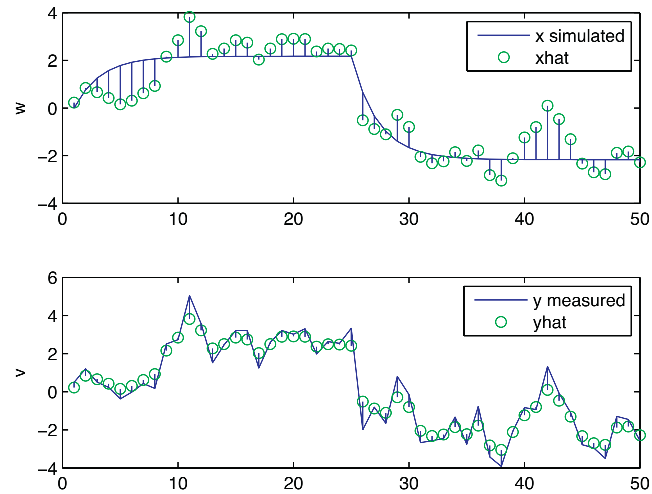

Fig. 1. Illustration of moving horizon estimation.

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - C\hat{x}_{k|k-1}) \tag{2c}$$
$$P_{k|k} = (I - K_k C)P_{k|k-1} \tag{2d}$$
$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k \tag{2e}$$
$$P_{k+1|k} = AP_{k|k}A^T + Q \tag{2f}$$

With access to more computational power, a much newer and increasingly popular approach is to use so-called moving horizon estimation (MHE). In MHE the introduced process and measurement noises are used as slack variables in an optimization formulation. If the model is nonlinear, these slack variables are usually introduced in a discretized version of the model.

$$x_{k+1} = g(x_k, u_k) + w_k$$
$$y_k = h(x_k, u_k) + v_k$$

Moving horizon estimation then corresponds to minimizing

$$V = (x_{k-M+1} - \hat{x}_{k-M+1})^T P^{-1}(x_{k-M+1} - \hat{x}_{k-M+1}) + \sum_{n=k-M+1}^{k} w_n^T Q^{-1} w_n + v_n^T R^{-1} v_n \tag{3}$$

with respect to all states x_n within the horizon and possibly subject to constraints as, for example,

$$x_{\min} \le x_n \le x_{\max}$$

Here P, Q and R are weight matrices used for tuning of the estimator, which have a similar interpretation and importance as the estimate and noise covariance matrices in Kalman filtering. In the minimization w_n and v_n are replaced by expression with x_n, x_{n+1}, y_n and u_n according to (3).

As indicated in its name, the optimization for moving horizon estimation is typically done over a horizon of data $[t-(M-1)T_s, t]$, where t is the current measurement time. Since this time interval is in the past, we assume access to historic values of the applied manipulated variables u_k and the measured process variables y_k. The first penalty term in the criterion, P^{-1}, is called the *arrival cost*. It is to create a link from one optimization window to the next, where $\hat{x}_{k-M+1}$ denotes the estimate for this particular time instant from the optimization run at the previous cycle. Fig. 1 tries to illustrate the MHE optimization which is a weighted sum of the vertical bars in the upper and lower plots.

MHE for estimating the states is an extensively studied field, see [26] for a thorough overview. Much of the literature is related