OPTIMAL PARAMETRIC SENSITIVITY CONTROL FOR A FED BATCH REACTOR.

J.D. Stigter*, **K.J. Keesman***

Systems and Control Group
Wageningen University and Research Centre
Bomenweg 4
6703 HD Wageningen
THE NETHERLANDS
Fax: +31 317 484819
e-mail: hans.stigter@user.aenf.wau.nl

Keywords: Identification of Non-Linear Systems, Identification Methods, Control and Optimization, Bio-technical Processes

Abstract

The paper presents a method to derive an optimal parametric sensitivity controller for optimal estimation of a set of parameters in an experiment. The method is demonstrated for a fed batch bio-reactor case study for optimal estimation of the saturation constant K_S and, albeit intuitively, the parameter combination $\frac{\mu_{max}X}{Y}$ where μ_{max} is the maximum growth rate [g/min], Y is the yield coefficient [g/g], and X is the (constant) biomass [g].

1 Introduction

A model structure f may essentially be viewed as a mapping of the input signals (u), via the states (x) and parameters (θ), to the system outputs (y). In parametric models the output sensitivity with respect to a (set of) parameter(s) θ determines whether a parameter can be estimated from the input/output data. If the sensitivity of y with respect to θ is small or even zero, then the instruments may not be well chosen or the input sequence $u(t)$ may not excite the parametric sensitivities sufficiently. Since our aim is to reconstruct the true parameter values from input/output observations it is intuitively appealing to focus on what information of the parameter vector θ is observable in the output signal through 'off-line' or 'on-line' calculation of the sensitivities $\frac{dx}{d\theta}$. Dötsch and van den Hof (1996) have shown that, for a linear system model, identifiability can be interpreted in terms of controllability and observability of the linear system model Σ augmented with the system of associated sensitivities driven by the state $x(t)$ and (possible) input $u(t)$.

It is also known that the gradient $\frac{dy}{d\theta}$ provides access to a measure for the information content of the specific dataset. More specifically, the gradient appears in the well known Fisher information matrix (FIM)

$$\Lambda(t_f) = \int_0^{t_f} \frac{\partial y(\tau)}{\partial \theta}^T Q^{-1} \frac{\partial y(\tau)}{\partial \theta} d\tau \tag{1}$$

where Q is a weighting matrix, usually in diagonal form with each diagonal element inversely proportional to the variance of the associated measurement noise of sensor i ($i = 1, \ldots, m$). The FIM provides a measure (after some criterion has been chosen 'a priori') of the information content of the data for the specific experiment conducted. A natural question is then, of course, how the input sequence $u(t)$ can be chosen in such a way that the parameters can be optimally estimated. This is the well known problem of experimental design which is a classical problem in the identification literature, (Norton, 1986; Ljung, 1987).

1.1 Definitions

Let the model structure be defined by the dynamical equation

$$\dot{x}(t) = f(x(t), u(t), \theta) \tag{2}$$

where $x(t)$ is the n-dimensional state vector, $u(t)$ the r-dimensional input vector, θ the q-dimensional vector of model parameters, and f the, possibly non-linear, model structure that maps the inputs, via the states and model parameters, to the outputs of the system model. The question of identifiability of the parameter set θ can be stated more explicitly by the question 'Given a model structure f: Is it possible to distinguish *uniquely* the values of the parameter vector θ on the basis of a carefully designed experiment? While this question may be answered by existing methods, e.g. the Taylor series expansion, (Pohjanpalo, 1978), our concern here is to find a feedback law that optimally controls the associated system of parametric sensitivities which can be derived straightforwardly from the model (2), leading to

$$\dot{x}_\theta(t) \approx \frac{\partial f}{\partial x} x_\theta(t) + \frac{\partial f}{\partial \theta} \tag{3}$$

Surprisingly enough the idea of finding such a feedback law on the basis of the model equation (2) augmented with the sensitivity functions (3) has hardly been pursued in the identification literature.

2 Optimal Sensitivity Control

2.1 A Fed-Batch Experiment

In the following we will derive a singular controller that maximally excites the sensitivity of the state with respect to the parameter θ (assuming a directly observable state x). In other words, we will maximally excite the sensitivity $\frac{dy}{d\theta}$ to a specific parameter in the model structure f so that the gradient $\frac{dy}{d\theta}$ is optimally present in the data set generated by the designed experimental setup.

The thought experiment taken here as an example is a fed batch bio-reactor experiment which can be used, for example, to determine the respiration rate of a population of bacteria feeding on a supplied substrate, (Vanrolleghem et al., 1995; Dochain et al., 1995). Observation of the substrate concentration can then be used to determine certain combinations of characteristic bio-kinetic parameters in terms of process yield Y in grams of biomass per gram of substrate, maximum specific growth rate μ_{max} in grams of biomass per minute, and the saturation constant K_S in grams per litre. Optimal input profiles (in terms of a Fisher design criterium) for the bio-kinetic parameters have been obtained for this setup in an 'ad hoc' manner Versyck (2000). To avoid a measurement problem of substrate (encountered in wastewater treatment) the study focuses on directly observable substrates such as sugar and ammonium. While observation of other substrates is difficult to achieve in practice this does not form a serious constraint for the method proposed here since the method can be extended to include, for example, oxygen and/or oxygen uptake rate (OUR) observations instead of substrate.

Assume a bio-reactor with biomass (X) that grows on the substrate, continuously fed into the reactor *dynamically* with a dynamical feed rate $u(t)$ in grams per litre per minute. Further assume that the growth process does not contribute substantially to the biomass X over the time-span considered (which is in the order of magnitude of several minutes) so that the biomass X may be assumed constant over the interval $[t_0, t_f]$ where t_f marks the end of the experiment. The growth dynamics include Monod kinetics so that the dynamical model reads

$$\dot{x}_1(t) = -\frac{\mu_{max}X}{Y}\frac{x_1(t)}{K_s + x_1(t)} + u(t) \tag{4}$$

where $x_1(t)$ is the substrate concentration in the bio-reactor. The reactor is assumed completely mixed and, as said, the substrate $x_1(t)$ is assumed to be directly observable. The parameters considered in the following are the combination $\frac{\mu_{max}X}{Y}$ and the parameter K_S which is sufficient.

It is known that the parameter K_S is most difficult to estimate from a set of observations (Holmberg and Ranta, 1982; Vanrolleghem et al., 1995) since this parameter is strongly correlated with the parameter μ_{max}. Define $x_2(t) \triangleq \frac{dx_1(t)}{dK_S}$ and $x_3(t) \triangleq \frac{dx_1(t)}{d(\mu_{max}X/Y)}$. The sensitivities of the parameter K_S and the parameter combination $\frac{\mu_{max}X}{Y}$ are derived as

$$\dot{x}_2(t) = \frac{\mu_{max}X}{Y}\frac{(x_1(t) - K_S x_2(t))}{(K_S + x_1(t))^2} \tag{5}$$

$$\dot{x}_3(t) = \frac{\mu_{max}X}{Y}\frac{K_S}{(K_S + x_1(t))^2}x_3(t) + \frac{x_1(t)}{K_S + x_1(t)} \tag{6}$$

In the sequel we will focus on finding a singular controller for optimal excitation of the sensitivity $\frac{dx_1}{dK_S}$ and we will assume that the combination $\frac{\mu_{max}X}{Y}$ is known for reasons of simplicity. It could be added that this assumption is in tune with the practical conditions since an obvious strategy to estimate the combination $\frac{\mu_{max}X}{Y}$ is to saturate the reactor completely with substrate so that the growth model is in the linear regime.

2.2 A Non-Linear Singular Control Problem

In order to optimally excite the sensitivity $x_2(t) = \frac{dx_1}{dK_S}$ we maximize $x_2^2(t)$ and define the following Hamiltonian

$$\mathcal{H}(x(t), \lambda(t)) \triangleq -x_2^2(t) + \begin{pmatrix} \lambda_1(t) & \lambda_2(t) \end{pmatrix} \cdot \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} \tag{7}$$

where $\lambda_1(t)$ and $\lambda_2(t)$ are the co-states defined as

$$\dot{\lambda}_1(t) = -\frac{\partial\mathcal{H}}{\partial x_1} = \frac{\mu_{max}X}{Y}\left\{\frac{K_S\lambda_1(t)}{(K_S + x_1(t))^2} + \frac{[x_1(t) - K_S(1 + 2x_2(t))]\lambda_2(t)}{(K_S + x_1(t))^3}\right\} \tag{8}$$

$$\dot{\lambda}_2(t) = -\frac{\partial\mathcal{H}}{\partial x_2} = \frac{\mu_{max}X}{Y}\frac{K_S\lambda_2(t)}{(K_S + x_1(t))^2} + 2x_2(t) \tag{9}$$

Since $\mathcal{H}$ does not explicitly depend on time a first integral of the problem gives $\mathcal{H} = constant$. Also, since the final time t_f is assumed unknown and no terminal conditions are specified (determining the value of the co-states at t_f) this constant can be assumed equal to zero. Since the problem is linear in the control variable $u(t)$ a singular control law that minimizes the Hamiltonian $\mathcal{H}$ over all possible input sequences $u(t)$ can be derived by setting (A. E. Bryson (Jr.), 1999)

$$\forall i \in \mathbb{N}: \quad \frac{d^i}{dt^i}\frac{d\mathcal{H}}{du} = 0 \tag{10}$$

In order to determine $u(t)$ explicitly only two differentiations are needed. For $i = 0$ we get $\lambda_1(t) = 0$ so that $\lambda_1(t)$ is the switching function for this problem. Consistency between the condition $\mathcal{H} = 0$ and $\frac{d}{dt}\frac{d\mathcal{H}}{du} = 0$ also eliminates $\lambda_2(t)$, so that from the case $i = 1$, the singular arc condition (or interior boundary condition) can be derived as

$$x_1(t) = K_S(1 + 2x_2(t)) \tag{11}$$

Finally, the case $i = 2$ determines the optimal input $u^\star(t)$ as

$$u^\star(t) = \frac{\mu_{max}X}{Y} \tag{12}$$

This surprisingly simple 'feedback law', together with the singular condition (11), determines a trajectory in state space on which K_S can be optimally identified. The condition (11) determines when to switch from a 'bang' input to the singular control law (12).

The optimal control law $u^\star(t)$ allows a reduction of the original set of equations (*on a singular arc*) giving

$$\dot{x}_1^\star(t) = \frac{\mu_{max}X}{Y}\frac{K_S}{K_S + x_1^\star(t)} \tag{13}$$

$$\dot{x}_2^\star(t) = \frac{\mu_{max}X}{Y}\frac{(x_1^\star(t) - K_S x_2^\star(t))}{(K_S + x_1^\star(t))^2} \tag{14}$$

where $x_i^\star$ denotes the optimal trajectory in state space on a singular arc, i.e. given $\frac{d\mathcal{H}}{du} = 0$. It is immediately clear that the optimal controlled system does not have a point of equilibrium and, therefore, is never in steady state. From the initial condition $x_2(0) \triangleq 0$ it is easily deduced that $x_1(0) = K_S$ immediately puts the system in singular mode so that the optimal control law $u^\star$ can be applied at the very beginning of the experiment. One can solve equation (13), given the initial condition $x_1^\star(0) = K_S$, analytically giving

$$x_1^\star(t) = -K_S + \sqrt{2K_S\frac{\mu_{max}X}{Y}t + 4K_S^2} \tag{15}$$

Since the solution $\begin{pmatrix} x_1^\star(t) \\ x_2^\star(t) \end{pmatrix}$ 'lives' on a singular arc for which condition (11) holds an analytical solution for $x_2^\star(t)$ follows immediately as

$$x_2^\star(t) = -1 + \sqrt{\frac{\mu_{max}X}{2K_S Y}t + 1} \tag{16}$$

Of course, the singular arc condition (11) depends on K_S meaning that in order to determine the locus of the singular arc satisfying this condition, knowledge of the parameter K_S is needed. Violation of constraint (11) will therefore be investigated in more detail for sensitivity to errors in an estimate of K_S at time t_0. This, together with some other simulation results, will be presented in the next section.

3 Results and Discussion

Since the parameter K_S is not 'a priori' known it is important to investigate the sensitivity of (13)–(14) to the initial condition $x^\star(0) = \begin{pmatrix} K_S \\ 0 \end{pmatrix}$. It was found through numerical simulation that an increase or decrease of $x_1(0)$ does not substantially change the trajectory of the sensitivity $x_2(t)$, (see figure 1). The figure demonstrates that, indeed, the initial condition $x_1(0) = K_S$ does not introduce a major difference for the evolution of the sensitivity $\frac{dx_1}{dK_S}$ for initial conditions $x_1(0) = 2K_S$ and $x_1(0) = \frac{1}{2}K_S$. In a second numerical simulation exercise the system (4)–(5) was simulated for three constant input values, namely $u(t) = \frac{1}{2}u^\star$, $u(t) = u^\star$, and $u(t) = \frac{3}{2}u^\star$. The resulting sensitivities are plotted in figure 2 from which it can clearly be observed that, indeed, the optimal input $u^\star = \frac{\mu_{max}X}{Y}$ excites the information content for estimation of the parameter K_S best.

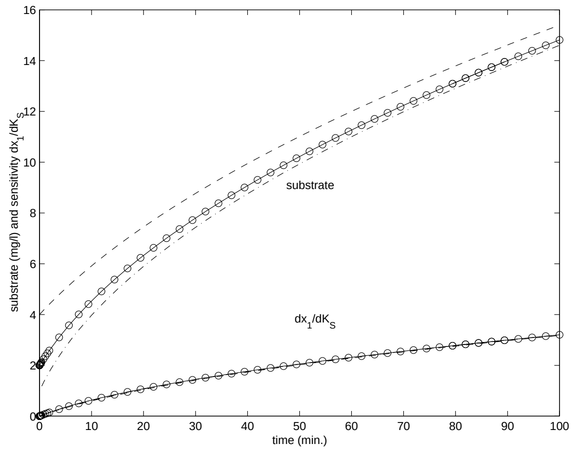

Figure 1: Substrate and sensitivity $\frac{dx_1(t)}{dK_S}$ for $\mu_{max} = 0.5$, $X = 1.0$, $K_S = 2.0$, and $Y = 0.75$.

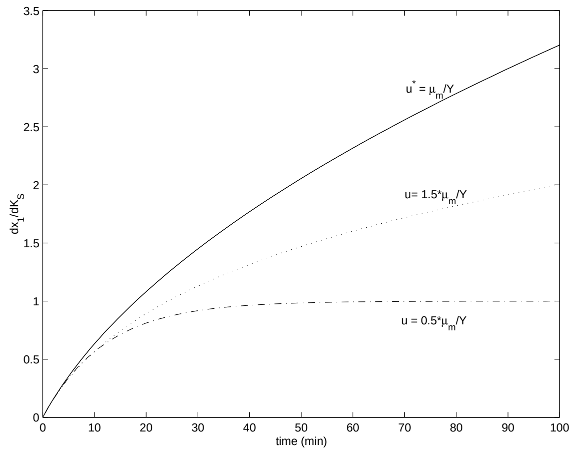

Figure 2: Sensitivity $x_2(t)$ for three different *constant* inputs, namely $u(t) = \frac{1}{2}\frac{\mu_{max}X}{Y}$, $u(t) = \frac{3}{2}\frac{\mu_{max}X}{Y}$, and $u^\star(t) = \frac{\mu_{max}X}{Y}$.

It could finally be noted that the optimal solution for optimal identification of K_S may well be used in an identification experiment where *both* $\frac{\mu_{max}X}{Y}$ and K_S are to be estimated. Intuitively (but also through simulation) one can show that an initial injection of substrate (first 'bang') will excite the sensitivity $x_3(t)$ substantially. The experiment could therefore be organized as follows:

(i) Apply a 'bang', i.e. $u(t) = u_{max}$, at the beginning of the experiment in order to identify the parameter combination $\frac{\mu_{max}X}{Y}$.

(iii) Apply a second 'bang', i.e. $u(t) = 0$, after a short period of time t_s and observe whether the singularity condition (11) holds.

(ii) Once the singularity condition (11) is satisfied, switch to the control $u^\star = \frac{\mu_{max}X}{Y}$ and estimate K_S.

The above algorithm was simulated for a switching time t_s

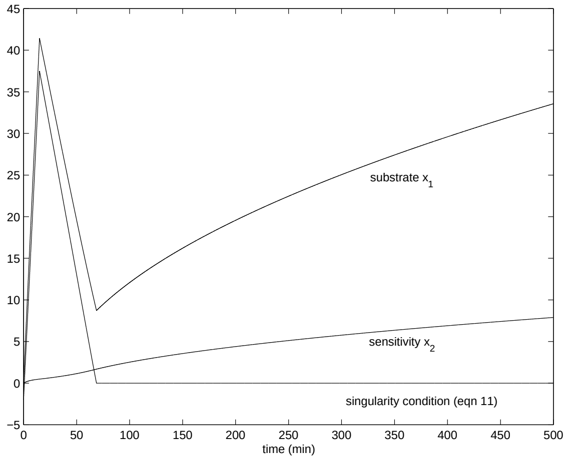

Figure 3: A simulated identification experiment.

(from the first 'bang' to the second 'bang') of 15 minutes for which the results are presented in figure 3. First, a bang-input $u_{max} = 5\frac{\mu_{max}X}{Y}$ was applied and after $t_s = 15$ min this controller was shut off. The resulting trajectory forms an excellent starting point for an optimal experiment in which the parameter combination $\frac{\mu_{max}X}{Y}$ and K_S are to be estimated.

4 Conclusions

A first attempt to include parametric sensitivities in the control loop has been made for a fed-batch reactor. The study has lead to a satisfactory controller for estimation of K_S, assuming Monod kinetics for the substrate consumption. The simple control law $u^\star(t) = \frac{\mu_{max}X}{Y}$ gives satisfactory results in the sense that it optimally excites the sensitivity of substrate with respect to the parameter K_S. The simplicity of the controller allows simple implementation in an 'on-line' estimation scheme in which the estimator controls its own input in order to find the parameter values with maximal sensitivity or minimum uncertainty. For the case of a more complicated sensitivity controller for which the analytical solution is not easy to derive one should consider a numerical scheme that optimizes the sensitivities on basis of a gradient $\frac{d\mathcal{H}}{du}$ (A. E. Bryson (Jr.), 1999).

References

A. E. Bryson (Jr.) (1999). *Dynamic Optimization*. Addison Wesley.

Dochain, D., Vanrolleghem, P. A., and Daele, M. V. (1995). Structural Identifiability of Biokinetic Models of Activated Sludge Respiration. *Water Research*, 29(11):2571–2578.

Dötsch, H. G. M. and van den Hof, P. M. J. (1996). Test for Local Structural Identifiability of High-order Non-linearly Parametrized State Space Models. *Automatica*, 32(6):875–883.

Frieden, B. R. (1998). *Physics from Fisher Information: A Unification*. Cambridge University Press.

Holmberg, A. and Ranta, J. (1982). Procedures for Parameter and State Estimation of Microbial Growth Process Models. *Automatica*, 18(2):181–193.

Ljung, L. (1987). *System Identification – Theory for the User*. Prentice Hall.

Norton, J. (1986). *An Introduction to Identification*. Academic Press.

Pohjanpalo, H. (1978). System Identifiability Based on the Power Series Expansion of the Solution. *Mathematical Biosciences*, 41:21–33.

Vanrolleghem, P. A., Daele, M. V., and Dochain, D. (1995). Practical identifiability of a biokinetic model of activated sludge respiration. *Water Research*, 29(11):2561–2570.

Versyck, K. J. (2000). *Dynamic Input Design for Optimal Estimation of Kinetic Parameters in Bioprocess Models*. PhD thesis, Katholieke Universiteit Leuven.