Available online at www.sciencedirect.com
ScienceDirect
automatica
Automatica 43 (2007) 178–183
www.elsevier.com/locate/automatica

Technical communique

Integrating virtual reference feedback tuning into a unified closed-loop identification framework☆

Antonio Sala*

Department of Systems Engineering and Control, Universidad Politécnica de Valencia, Camino de Vera, 14, E-46022 Valencia, Spain

Received 20 June 2005; received in revised form 3 July 2006; accepted 11 August 2006
Available online 2 October 2006

Abstract

This paper compares particular instances of control-relevant identification, closed-loop identification, and controller identification (either in an actual loop or in a virtual reference feedback tuning approach, VRFT) problems. Significant similarities appear between them which allow, in particular, alternative formulations of the VRFT methodology in a differentiable setting.

Keywords: System identification; Controller tuning; Data-based controller design; VRFT

1. Introduction

This paper discusses computation of the gradient of an identification (ID) cost index in four closed-loop scenarios. On one hand, closed-loop process ID (Gilson & Van den Hof, 2005; Ljung, 1999) and a control-relevant ID setting are considered (Hjalmarsson, 2005; Van den Hof, 1998). On the other hand, data-based controller adjustment procedures are discussed: closed-loop controller ID (Landau et al., 2001), virtual reference feedback tuning (VRFT) (Campi et al., 2002; Campi & Savaresi, 2006; Sala & Esparza, 2005), and iterative feedback tuning (IFT) (Hjalmarsson et al., 1998). The four scenarios offer different perspectives about the same problem and cross-fertilisation is possible. In particular, the VRFT approach and that in Landau et al. (2001) may be considered the same. One of the objectives of the paper is extending Landau's and VRFT techniques: the unified approach includes the procedures in Campi and Savaresi (2006) and outlines other possible alternatives.

The structure of the paper is as follows: next section discusses the notation, identification objectives and calculation of the derivatives of a closed-loop operator. Section 3 discusses four ID scenarios pinpointing the common aspects of all of them, followed by a conclusion section.

2. Preliminaries and notation

Let us consider a nonlinear mapping $N : \mathscr{S}_u \to \mathscr{S}_y$, i.e., $y = Nu$, where signals y, u (known by assumption) belong to Hilbert spaces $\mathscr{S}_u$, $\mathscr{S}_y$ with scalar product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. By far, the most common signal spaces used in practical system ID (Ljung, 1999) are a finite set of real-valued input and output samples obtained at regular intervals, $u = \{u_0, u_1, \dots, u_N\}$, $y = \{y_0, y_1, \dots, y_N\}$, being the norm $\|y\| = \sqrt{\sum_{i=0}^N y_i^2}$.

Consider now a parameterised nonlinear model $y_m = T(\psi, u)$, where ψ is a set of adjustable parameters. A cost index, J, will be defined as the squared norm of the so-called filtered *prediction error* $\varepsilon = F(y_m - y)$ (least-squares criteria), where F is a suitable linear filter:[1]

$$J(\psi) = \tfrac{1}{2}\|\varepsilon\|^2 = \tfrac{1}{2}\|F(y_m - y)\|^2. \tag{1}$$

Nonlinear system ID will be understood as minimising $J(\psi)$. Gradient computation is at the root of many optimisation techniques (Luenberger, 2003), which may be used in order to find a (maybe local) minimiser ψ^{opt}. In order to compute the

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Jay H. Lee under the direction of Editor André Tits.

* Tel.: +34 96 387 9770; fax: +34 96 387 9579.
E-mail address: asala@isa.upv.es.

0005-1098/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2006.08.005

[1] Unless otherwise stated, $F = I$ (identity) will be assumed in the sequel.

gradient, the derivative of J with respect to each adjustable parameter ψ_i is given by[2]

$$\frac{\mathrm{d}J}{\mathrm{d}\psi_i} = \left\langle F(y_m - y), F\frac{\mathrm{d}y_m}{\mathrm{d}\psi_i}\right\rangle. \tag{2}$$

The term $\mathrm{d}y_m/\mathrm{d}\psi_i = \mathrm{d}T(\psi, u)/\mathrm{d}\psi_i$ will be denoted as *model derivative*. Computation of $\mathrm{d}y_m/\mathrm{d}\psi_i$ may require gradient propagation through time (Werbos, 1990), if model T is a recurrent one.[3] When fitting N data samples, F in (2) is a $N \times N$ Toeplitz convolution matrix formed with impulse response coefficients of F (Burrus, 1972), and $\mathrm{d}y_m/\mathrm{d}\psi_i$ is an $N \times 1$ column vector.

If there exists an ideal ψ^* such that $y = T(\psi^*, u)$ and ψ is close enough to ψ^* ($y_m - y$ is small), then

$$y_m - y = T(\psi, u) - T(\psi^*, u) \approx \left.\frac{\mathrm{d}T}{\mathrm{d}\psi_i}\right|_{\psi}(\psi_i - \psi_i^*), \tag{4}$$

where only increments of parameter ψ_i have been considered. On the sequel, the subscript i will be removed and the derivatives with respect to ψ must be understood to be with respect to any arbitrarily chosen element of ψ.

Open-loop ID: In a control context, a basic scenario is the so-called prediction-oriented ID (Ljung, 1999) of a plant $y = Pu$: a data set u, y is available and $y_m = P(\theta, u)$ ("output error" model structure, OE). As u does not depend on θ, the gradient of J, (2), in this case is

$$\frac{\mathrm{d}J}{\mathrm{d}\theta} = \left\langle F(P(\theta, u) - y), F\frac{\partial P}{\partial \theta}(\theta, u)\right\rangle. \tag{5}$$

Closed-loop ID: The set of scenarios considered in next section involve a closed-loop with a plant P and a controller C, given by widely used basic equations: $e = r - y$, $y = Pu$, $u = Ce$, being r a reference signal. Under mild assumptions, $y = PC(r - y)$ implicitly defines a closed-loop map $y = Mr$, and $e = Sr = (I - M)r$. Also, $e = r - y = r - PCe$, i.e., $(I + PC)e = r$ hence

$$S = (I + PC)^{-1}, \quad M = PC(I + PC)^{-1}. \tag{6}$$

In general, $y \neq SPCr$; basically, only in the linear PC case $M = (I + PC)^{-1}PC$ is also true, as (10) later shows.

[2] Derivative notation $\mathrm{d}/\mathrm{d}\psi$ will denote the total derivative taking into account *implicit* dependences on ψ of all arguments of a function (applying chain rule if needed). Partial derivatives $\partial/\partial\psi$ will account only for *explicit* dependences on argument ψ (e.g., in (1), dependence on ψ of J is implicit).

[3] For instance, consider a nonlinear discrete-time model

$$y_m(t + 1) = 0.5y_m(t) + \psi y_m(t)^2 u(t). \tag{3}$$

Then, $\mathrm{d}y_m(t+1)/\mathrm{d}\psi = \partial y_m(t+1)/\partial\psi + (\partial y_m(t+1)/\partial y_m(t))(\mathrm{d}y_m(t)/\mathrm{d}\psi) = y_m(t)^2u(t) + (0.5 + 2\psi y_m(t)u(t))\,\mathrm{d}y_m(t)/\mathrm{d}\psi$ is a recurrent equation for computing the required model derivatives. In closed-loop, as $u(t)$ itself depends on past values of $y_m(t)$ and hence, indirectly on ψ, the expression would need the addition of $\psi y_m(t)^2\mathrm{d}u(t)/\mathrm{d}\psi$, where $\mathrm{d}u(t)/\mathrm{d}\psi$ would come from the loop equations, see (8)–(10).

Considering a model with parameterisations $P(\theta, u)$ and $C(\mu, e)$, its output can be expressed as

$$y_m = P(\theta, C(\mu, r - y_m)) = P_\theta C_\mu(r - y_m), \tag{7}$$

where shorthand $P_\theta u = P(\theta, u)$, $C_\mu e = C(\mu, e)$ have been introduced. Differentiability of plants and controllers with respect to all of its arguments will be assumed. Linearity of plants and controllers, when applicable, will refer to the second argument (input signals) and not to the parameterisations. Some linearisations will be denoted by $\bar{M} = \partial M/\partial r$, $\bar{S} = I - \bar{M}$, $\bar{P} = \partial P/\partial u$, $\bar{C} = \partial C/\partial e$. For instance, when trying to fit N input–output data samples, $\bar{P}$ is an $N \times N$ matrix with elements $\bar{P}_{ij} = \mathrm{d}y_m(i)/\mathrm{d}u(j)$ which may be considered the convolution coefficients of a linear time-varying system.[4]

On the following, the notation M_θ, M_μ will refer to the achieved closed-loop map when a plant P_θ is operating with a fixed controller or when a controller C_μ is operating with a fixed plant (the "real" one), respectively, whereas M will denote a target closed-loop map for control design. Notation S_θ, $\bar{S}_\theta$, $\bar{M}_{\theta,\mu}$, etc. (with obvious meaning) will also be used.

The following results, obtained from (7) via chain rule and implicit function theorem, will be later used for gradient calculations (dependence on parameters not subject to derivative calculations will be omitted):

$$\frac{\mathrm{d}y_m}{\mathrm{d}\theta} = \left(I + \frac{\partial P_\theta}{\partial u}\frac{\partial C}{\partial e}\right)^{-1}\frac{\partial P_\theta u}{\partial \theta} = \bar{S}_\theta\frac{\partial P_\theta u}{\partial \theta}, \tag{8}$$

$$\frac{\mathrm{d}y_m}{\mathrm{d}\mu} = \left(I + \frac{\partial P}{\partial u}\frac{\partial C_\mu}{\partial e}\right)^{-1}\frac{\partial P}{\partial u}\frac{\partial C_\mu e}{\partial \mu} = \bar{S}_\mu\bar{P}\frac{\partial C_\mu e}{\partial \mu}, \tag{9}$$

$$\frac{\mathrm{d}y_m}{\mathrm{d}r} = \bar{M} = (I + \bar{P}\bar{C})^{-1}\frac{\partial P}{\partial u}\frac{\partial C}{\partial e} = \bar{S}\bar{P}\bar{C}. \tag{10}$$

For instance, (9) is obtained from derivation of $y = PC_\mu e$ as follows: $\mathrm{d}y_m/\mathrm{d}\mu = (\partial P/\partial u)(\mathrm{d}u/\mathrm{d}\mu) = \bar{P}(\partial C_\mu e/\partial\mu + \bar{C}_\mu(\mathrm{d}e/\mathrm{d}\mu)) = \bar{P}(\partial C_\mu e/\partial\mu + \bar{C}_\mu(-\mathrm{d}y_m/\mathrm{d}\mu))$; (10) from $\mathrm{d}y_m/\mathrm{d}r = \bar{P}\bar{C}(I - \mathrm{d}y_m/\mathrm{d}r)$.

3. Closed-loop identification scenarios

Scenario 1: *closed-loop plant ID*: Consider data r, u, y, obtained in closed-loop with a known controller C in operation, in order to adjust some parameters of a plant model P_θ. In that case, $y_m = M_\theta r$ is readily obtained by simulating the closed-loop $y_m = P_\theta(C(r - y_m))$, and model derivatives in (2) are given by (8)

$$\frac{\mathrm{d}J}{\mathrm{d}\theta} = \left\langle M_\theta r - y, \bar{S}_\theta\frac{\partial P_\theta u}{\partial \theta}\right\rangle. \tag{11}$$

[4] As an example, for system (3),

$$\frac{\mathrm{d}y_m(i)}{\mathrm{d}u(j)} = \begin{cases} \psi y_m(j)^2, & i = j + 1, \\ (0.5 + 2\psi y_m(i-1)u(i-1))\frac{\mathrm{d}y_m(i-1)}{\mathrm{d}u(j)}, & i > j + 1, \\ 0, & i < j + 1. \end{cases}$$