Miroslav Kárný; Ivan Nagy; Josef Böhm; Alena Halousková

Design of spline-based self-tuners

Kybernetika, Vol. 26 (1990), No. 1, 17--30

Persistent URL: http://dml.cz/dmlcz/124221

KYBERNETIKA — VOLUME 26 (1990), NUMBER 1

DESIGN OF SPLINE-BASED SELF-TUNERS

MIROSLAV KÁRNÝ, IVAN NAGY, JOSEF BÖHM, ALENA HALOUSKOVÁ

The paper describes a way of constructing digital adaptive controllers for continuous-time linear systems. The key step of the design is the simultaneous spline approximation of both signals and operator kernels in the integral (convolutional) description of the controlled system. Having the freedom to choose nonequidistant spline nodes, a possibility of better modelling — under model-complexity restrictions and for short sampling periods — is gained.

With a suitable spline basis chosen, recursive identification and minimization of integral quadratic loss can be solved by reliable algorithms developed for regression model and discrete time quadratic loss. The conceptual feasibility of the approach is demonstrated (even for potentially infinite control horizon) to yield implementable adaptive controllers.

1. INTRODUCTION

A significant part of contemporary research into digital self-tuning controllers is oriented towards continuous-time modelling of the system under control [7], [6], [16]. The primary aim is to improve the quality of control by taking into account the continuous nature of controlled systems and applying feedback and/or feedforward actions with the highest possible sampling rate.

Implementation of adequate solutions to these and associated problems depends heavily on appropriate connections between the continuous-time reality and the discrete-time nature of digital data handling. The cited papers are examples of the progress made in this respect. They share, however, a common drawback: they rely on ARMAX-type models and must deal with inherent difficulties related to the MA-part of this model (no finite-dimensional statistics for its recursive estimation, no systematic method for fixing it beforehand).

This paper can be taken as an alternative, hopefully more flexible, way of addressing digital self-tuning control of continuous-time systems. The approach has been developed in connection with a particular application case [11] in paper industry. The key problems of modelling and approximation are solved as follows:

- The convolution-type non-parametric linear model of a stochastic controlled process is used as underlying process description. Among equivalent input-output descriptions, the version is chosen the noise of which becomes practically white after sampling.
- A finite parametrization needed for practical use is achieved by
  - approximating both signals and operator kernels by splines with a finite support;
  - assuming that the process history has a practical impact on a limited time span only.

In this way, the sampled system model converts to regression-type models with regressors formed by *filtered* input-output samples. Consequently, the fact that digital LQG self-tuners based on regression models have achieved the stage of practical applicability can be exploited:

- The gained model can be identified by recursive least squares.
- The *integral* quadratic performance index converts (under the adopted assumptions) in a discrete quadratic loss with a non-diagonal weighting. For it, suitable optimization techniques are available [10], even for constrained inputs [2, 3].

The ARMAX-based solutions referred above can be interpreted and treated similarly. Our design is, however, felt to be conceptually straightforward and applicable to non-standard cases like systems with distributed parameters.

A common difficult point of the cited and proposed solutions is the choice of data filters. In both cases, they can be chosen a priori by using structure determination theory [8]. Properties of the spline bases, which determine the filters in our case, are, however, more tightly connected to those of observable signals in the time domain. It leads us to the conjecture that the choice of the proper spline base will be simpler than that of the noise covariance in the ARMAX model.

Approximating the involved signals by splines, we have joined the stream of attempts which exploit various, typically orthogonal, types of expansions for control and/or identification purposes (see e.g. [4], [5]). However, according to our best knowledge:

- Nobody has had the audacity to *approximate both signals and operators* at the same time. The joint approximation adopted here seems to be more universal and internally consistent than the combination of splines with a state space model [4].
- No expansion-based version of self-tuners (in the sense of the cited papers) has been worked out. Their on-line nature has to be especially dealt with: a finite number of parameters can only be estimated using recursively computable finite-dimensional statistics. The regression model we have arrived at is the most important case with this property.

The key steps of modelling and approximation are explained in detail. It should enable the user to tailor the proposed controller to various versions of LQG selftuner design. The optimization part of the controller is touched only because the major changes (comparing to the usual LQG design) concern identification.

For simplicity, the discussion is restricted to SISO systems. The basic notions are illustrated using first degree splines which are both simple and of direct practical use.

2. PROBLEM, KEY ASSUMPTIONS AND DESIGN RESTRICTIONS

The *constructive* problem is addressed in the paper:

Design a prototype of the adaptive digital controller of a continuous-time linear time-invariant system optimizing the integral-quadratic performance index under the following key assumptions:

- the relation of the system output to the system input and process noise is affine describable in terms of convolution operators;
- the controlled process history has a practical impact on a limited future only;
- all functions involved in system model can be approximated by properly chosen splines in the sense that the sampled process noise corrupted by approximation error can be modelled by stationary gaussian process.

The design is performed under commonly acceptable restrictions:

- the separation of identification and control design is enforced;
- the control design is performed for a potentially infinite horizon.

3. PRELIMINARIES

Essential concepts and facts are summarized in this section. The notation is introduced at the same time.

Common formal agreements

Simultaneous use of operators, continuous-time dependent vectors, functions, discrete-time dependent samples etc. makes the notation somewhat complicated. Noting the following rules should assist the reader:

$Y(\cdot)$, $y(t)$, $u(t)$, $f_i(t)$, $a(t - \tau)$ – functions of continuous time t or τ are identified by $(\cdot)$, (t), $(t - \tau)$ type arguments;

$a(k)$, $F(k)$ – samples related to the discrete time step k (sampling time t_k) will be denoted by a (k) type argument;

$\mathbf{A}$, $\bar{\mathbf{B}}$, $\mathbf{b}$ – bold Latin Roman symbols denote operators;

$\boldsymbol{w}$, $\boldsymbol{f}(t)$, $\boldsymbol{F}$ – lower-case (upper-case) of bold Italics symbols denote column vectors (matrices) respectively;

w_i, $f_i(t)$, F_{ij} – subscripts indicate a particular entry in an array;

$\mathscr{H}^U$, w^y, n^y – superscripts refer to the functions to which the indexed symbols are related (the transposition is denoted by $'$).

The relation of the chosen system model to the general description of affine systems

The general description of an affine time-invariant continuous-time stochastic controlled system has the form

$$\bar{\mathbf{A}}Y(\cdot) + \bar{\mathbf{B}}U(\cdot) + \bar{\mathbf{C}}E(\cdot) = 0 \tag{1}$$

where $\bar{\mathbf{A}}$, $\bar{\mathbf{B}}$, $\bar{\mathbf{C}}$ are *general* affine, time-invariant, causal operators acting on function spaces $\mathscr{H}^Y$, $\mathscr{H}^U$, $\mathscr{H}^E$ of the system output $Y(.)$, input $U(\cdot)$ and white noise $E(\cdot)$, respectively.

The description (1) is not unique – the equation can be multiplied by an operator. Thus, a theoretically equivalent model

$$\mathbf{A}Y(\cdot) + \mathbf{B}U(\cdot) + \bar{E}(\cdot) = 0 \quad \text{with} \quad \mathbf{A} = \mathbf{C}^*\bar{\mathbf{A}}\,, \quad \mathbf{B} = \mathbf{C}^*\bar{\mathbf{B}} \tag{2}$$

can be considered. We suppose that there is such an operator $\mathbf{C}^*$ for which the transformed noise signal

$$\bar{E}(\cdot) = \mathbf{C}^*\bar{\mathbf{C}}E(\cdot) \tag{3}$$

becomes white discrete process when sampled with the shortest technically feasible period. Under this assumption, the latter form of the model will be used as the basic system model from here onwards.

Convolution form of causal operators

The operator $\mathbf{A}$ acting on the signal $Y(\cdot)$ at the time $t \in [0, T]$ will be assumed in the convolution form

$$\mathbf{A}Y(t) = O^A + \int_0^t A(\tau)\, Y(t - \tau)\, \mathrm{d}\tau \tag{4}$$

where $T < \infty$ is the horizon of interest, the offset O^A reflects the nondecreasing influence of the initial conditions and $A(\tau) \in \mathscr{H}^A$ is a causal ($A(\tau) = 0$ for $\tau < 0$) smooth kernel.

The operator $\mathbf{B}$ is described in the same manner with the characteristics O^B and $B(\cdot)$.

Remark. A more careful modelling of the influence of system initial conditions is possible, by assuming time-dependent O^A in the form

$$\text{const}_1^A + \text{const}_2^A \int_0^t A^O(\tau)\, \mathrm{d}\tau$$

with a kernel $A^O(\cdot)$ and similarly for O^B. For simplicity, this straightforward extension will not be treated here.

Splines

We shall approximate functions defined on time interval $[0, T]$ from various function spaces $\mathscr{H}^X$ (distinguished by the superscript $X \in \{Y, U, A, B\}$). For each of them, approximating functions will be chosen from a properly selected set of splines. Where helpful, symbols related to approximating functions adjoint to respective spaces will be distinguished by the superscript $x \in \{y, u, a, b\}$.

For defining splines of the degree m we shall divide the interval $[0, T]$ by the nodes $\{t_i\}_{i=0}^{N}$

$$0 = t_0 < t_1 < \ldots < t_N = T \tag{5}$$

A function $x(t)$ is called spline of the degree m and of the defect $d \in \{1, 2, \ldots, m\}$ [13] iff

- $x(t)$ is a polynomial in t of the degree at most m on every open subinterval (t_{i-1}, t_i);
- the derivatives of $x(t)$ on the entire interval $[0, T]$ are continuous as long as their orders are at most $m - d$.

The linear space spanned over splines of the degree m defined by the above grid has the dimension $dim = (N + 1)\, d + m + 1$ [13]. Thus, there is a basis having dim linearly independent members, say $f_i(t)$, such that any spline $x(t)$ from the discussed span can be expressed as

$$x(t) = \sum_{i=0}^{dim-1} w_i f_i(t) = w' f(t) \tag{6}$$

where the weights w_i are uniquely determined. The base functions as well as the weights are ordered into the dim-vectors

$$f(t) = [f_0(t), \ldots, f_{dim-1}(t)]' , \quad w = [w_0, \ldots, w_{dim-1}]' . \tag{7}$$

In order to stress that the control horizon is unbounded, the vectors (matrices), the dimensions of which grow linearly with T, will be called *potentially* (semi)infinite, and when written entry-wise the following notation will be used

$$x' = [x_0, \ldots, x_i, \ldots] . \tag{8}$$

4. APPROXIMATION OF SYSTEM MODEL

Motivation

For self-tuners, the model (1) as well as its formal equivalent (2) are supposed to be unknown and identified when the control takes place. At least for identification, a finite parametrization is needed. It means a class of finitely parametrized models has to be chosen within which there is "sufficiently" good approximant of the real infinite-dimensional systems. Obviously, the good approximants of the model (1) need not be satisfactory for the model (2) and vice versa. Thus, the discussed system descriptions might be quite non-equivalent when performing the inductive step from reality to mathematics.

The spline approximation of the version (2) has been chosen as:

- the need for approximative recursive identification (connected with MA-part of ARMAX models) is avoided;

- the number of the estimated parameters is expected to decrease when approximating the "ratios" **A**, **B** by splines with nonequidistant nodes (comparing to the traditional sum of pre-specified exponentials);
- smoothness degree of the signal part taken as useful (worth to control) is explicitly in the designer's hands (a time domain counterpart of cross-over frequency?!).

Outline of the used approximation

The input-output signals $Y(t)$, $U(t)$ will be interpolated by splines $y(t)$, $u(t)$. The approximation of the linear operators **A**, **B** in the system model (2) will rely on the assumption that they are convolutions specified by the kernels $A(\cdot)$, $B(\cdot)$. The approximating operators **a**, **b** are assumed to be of the same convolution form as **A**, **B**. Their kernels $a(t)$, $b(t)$ will be constructed as the spline approximations of the kernels $A(t)$, $B(t)$.

The spline bases $\{f_i^x\}$ for $x \in \{y, u, a, b\}$ approximating the functions from the signal spaces $\mathscr{H}^Y$, $\mathscr{H}^U$ and from the kernel spaces $\mathscr{H}^A$, $\mathscr{H}^B$ can differ in the grids dividing the interval $[0, T]$ as well as in the degrees and in the defects employed.

Thus, all functions are approximated by appropriate splines

$$y(t) = w^{y'} f^y(t)\,, \quad u(t) = w^{u'} f^u(t) \tag{9}$$

$$a(t) = w^{a'} f^a(t)\,, \quad b(t) = w^{b'} f^b(t)\,. \tag{10}$$

Remark. Note that the approximation of operators used relies on the assumption that *the approximated kernels are smooth*, i.e. above all without Dirac δ-functions. If these are present it is sufficient to treat them as additional terms in operators expression. The extension to this case is straightforward and will not be treated here.

Choice of base functions

The choice of the basis is determined by the requirements on the approximation to be performed. We are searching for the basis with members having finite (possibly shortest) intervals (supports) for which they are nonzero, for which the weights w_i in (6) are easy to determine and which have possibly a small defect. The ordering of the requirements expresses our current preferences over them.

The finite length of the supports (required even for $T \to \infty$) is of vital importance: we shall rely on the obvious consequence that under this condition only a finite bounded number of supports of different base functions have a nonempty intersection with a given one. Similarly, the determination of the weights must require only a finite number of neighboring data samples.

The smoothness of the approximants, i.e. the relevant defect, should correspond to the smoothness of the approximated functions. At present, we have not found sufficient reasons to insist strictly on the assumption that the spaces $\mathscr{H}^Y$, $\mathscr{H}^U$, $\mathscr{H}^A$, $\mathscr{H}^B$ contain only the functions with continuous higher-order derivatives.

In view of the listed preferences, we have chosen *interpolating splines* with *finite support*, fulfilling the condition

$$f_i(t_j) = \delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{otherwise}. \end{cases} \tag{11}$$

This requirement implies that the weights in (6) become values of the function $x(t)$

$$w_i = x(t_i) \tag{12}$$

It is a simple algebraic exercise to show that splines of the recommended odd degree [13] $m = 2r - 1$ have the necessary defect equal to r and the supports are pairs of neighboring subintervals specified by nodes t_i. For a defect greater than 1, the selected splines do not form a whole basis of the span.

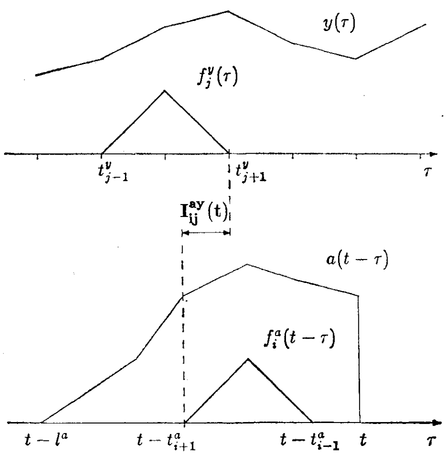

Fig. 1. Intersection of supports of base functions $f_j^y(.), f_i^a(.)$ used for approximating the output $y(\tau)$ and the kernel $a(t - \tau)$ of the convolution operator $\mathbf{a}(.) = \int_0^t a(t - \tau)\,(.)\,d\tau$.

In the first-degree case, the basis is unique and consists of "hat"-like functions $f_i(t)$ on (t_{i-1}, t_i) for $i = 1, \ldots, N - 1$ (see Fig. 1). The functions f_0, f_N are "half-hats" restricted to the intervals $[0, t_1]$ and $[t_{N-1}, T]$.

Remark. The inclusion of the approximating functions from outside the subspan cannot improve the quality of interpolation. A loss in smoothness of approximating functions is currently not considered to be important, but this attitude may be changed after further research.

Signal approximation

Choosing (as usual) the same sampling rate for all signals with a fixed sampling period t_s, it is natural to take sampling points as grid points of the corresponding splines, i.e.

$$t_i^u = it_s\,, \; t_i^y = it_s + \text{shift of input-output sampling} \tag{13}$$

The weights of the signal approximants $\{w_i^y, w_i^u\}$ are then directly measurable and coincide with the signals $Y(t)$, $U(t)$ at sampling points

$$w^y = [w_0^y, \ldots, w_i^y, \ldots]' = [y(t_0^y), \ldots, y(t_i^y), \ldots]' = [Y(t_0^y), \ldots, Y(t_i^y), \ldots]' \quad (14)$$

$$w^u = [w_0^u, \ldots, w_i^u, \ldots]' = [u(t_0^u), \ldots, u(t_i^u), \ldots]' = [U(t_0^u), \ldots, U(t_i^u), \ldots]'. \quad (15)$$

Because of the fixed finite distance between respective sampling nodes, the number of grid points N^y, N^u is proportional to the assumed horizon. As the case $T \to \infty$ is assumed, the numbers of grid points are potentially infinite, too. From here onwards we shall stress this fact by formally setting $N^y = N^u = \infty$.

Operator approximation

The approximation of the kernels is based on the assumption, that approximants are interrelated by the same equation as the original one, i.e.

$$\mathbf{a}y(\cdot) + \mathbf{b}u(\cdot) + e(\cdot) = 0 \quad (16)$$

where $e(\cdot)$ represents the stochastic term $\bar{E}(\cdot)$ modified by approximation error, which has to be made virtually negligible.

As the number of the input-output samples is potentially infinite the *finite-memory assumption* is crucial for restricting the necessary operations to a fixed finite amount even for a potentially infinite control horizon. We rely on the assumption that the kernels have (*practically*) finite bounded support so that they can be approximated by the splines with a fixed finite number of nodes ($n^a + 1$ resp. $n^b + 1$)

$$a(t) = \sum_{i=0}^{n^a} w_i^a f_i^a(t), \quad b(t) = \sum_{i=0}^{n^b} w_i^b f_i^b(t) \quad (17)$$

with constant weights

$$w^a = [w_0^a, w_1^a, \ldots, w_{r^a}^a]', \quad w^b = [w_0^b, w_1^b, \ldots, w_{n^b}^b]'. \quad (18)$$

The necessary quality of approximation of both kernels can be achieved by choosing appropriate spline supports and (non-equidistant) grids. Having on mind the aim of the paper, instead of analyzing the precision of the approximation (closely related to the so-called L-splines) a ready algorithmic tool for support and grid choice is referred in the next section.

Design model

By combining the above steps, the model (16) converts into the following (linear-in-parameters w^a, w^b) model

$$\sum_{i=0}^{n^a} \sum_{j=0}^{\infty} w_i^a\, y(t_j^y) \int_0^t f_i^a(t-\tau) f_j^y(\tau)\, \mathrm{d}\tau +$$

$$+ \sum_{i=0}^{n^b} \sum_{j=0}^{\infty} w_i^b\, u(t_j^u) \int_0^t f_i^b(t-\tau) f_j^u(\tau)\, \mathrm{d}\tau + O^A + O^B + e(t) = 0 \quad (19)$$

which will be used as *design* model (its parameters are estimated and control synthesis performed for it with parameters replaced by their estimates).

We put (19) into matrix form

$$w^{a'} F^{ay}(t)\, w^y + w^{b'} F^{bu}(t)\, w^u + o + e(t) = 0 \tag{20}$$

where we have introduced the "joint" offset

$$o = O^A + O^B\,. \tag{21}$$

The semi-infinite matrices F^{ay}, F^{bu} are of the type $(n^a + 1, \infty)$, $(n^b + 1, \infty)$, respectively, with generic (i, j)th entries

$$F_{ij}^{ay}(t) = \int_{I_{ij}{}^{ay}(t)} f_i^a(t - \tau)\, f_j^y(\tau)\, d\tau \tag{22}$$

$$F_{ij}^{bu}(t) = \int_{I_{ij}{}^{bu}(t)} f_i^b(t - \tau)\, f_j^u(\tau)\, d\tau \tag{23}$$

where $I_{ij}^{**}(t)$ is the actual intersection of the integrand supports.

Due to the finite length of the spline supports, only a finite bounded number of them have a nonempty intersection. This implies that each matrix $F(t)$ contains a finite bounded number of nonzero entries (the position of their bounded cluster shifts with time, cf. also Fig. 2).

Consequently, for the selected base functions, the signals $y(t)$, $u(t)$ at time t can be shown to be linear combinations of a *fixed finite number* of the measured *past* samples and at most of a *single future* sample. It can be seen from the explicit expression for the generic interval $I_{ij}^{ay}(t)$

$$I_{ij}^{ay}(t) = \left(\max\left(t_{j-1}^y, t - t_{i+1}^a\right), \min\left(t_{j+1}^y, t - t_{i-1}^a\right)\right) \tag{24}$$

which implies

$$I_{ij}^{ay}(t) = \emptyset \quad \text{for} \quad t_{j+1}^y < t - t_{i+1}^a \quad \text{or} \quad t_{j-1}^y > t - t_{i-1}^a\,. \tag{25}$$

The formulae for I_{ij}^{bu} are analogous.

In the above expressions, the relative positions of indices i, j and of the time t correspond to schematic Figure 1 which illustrates the above statements. There, the symbol l^a denotes the length of the kernel support in terms of signal sampling units.

5. IDENTIFICATION

Due to the continuous-time nature of the model, identification can be performed for any sequence of identification moments, say $\{t_k^I\}_{k=0}^{\infty}$, with increments guaranteeing the whiteness of the sampled noise

$$\{\bar{e}(k)\}_{k=0}^{\infty} = \{e(t_k^I)\}_{k=0}^{\infty}\,. \tag{26}$$

If the relative positions of the identification-moments and interpolating nodes $\{t_k^y, t_k^u\}_{k=0}^{\infty}$ are invariant then the *nonzero entries* of the matrices $F^{ay}(k) = F^{ay}(t_k^I)$, $F^{bu}(k) = F^{bu}(t_k^I)$ *do not depend on* k and can be precomputed. It can be verified by inspecting their definitions (22), (23).

Introducing the $(n^a + 1)$ and $(n^b + 1)$ *vectors* of the filtered outputs and inputs

$$\bar{\boldsymbol{y}}(k) = \boldsymbol{F}^{ay}(k)\, \boldsymbol{w}^y \quad \bar{\boldsymbol{u}}(k) = \boldsymbol{F}^{bu}(k)\, \boldsymbol{w}^u \quad k = 0, 1, 2 \ldots \tag{27}$$

and the $(n^a + n^b + 3)$ vector of unknown coefficients

$$\boldsymbol{\theta} = [\boldsymbol{w}^{a\prime}, \boldsymbol{w}^{b\prime}, o]' , \tag{28}$$

we arrive at the standard regression model

$$\boldsymbol{\theta}'\, \boldsymbol{d}(k) + \bar{e}(k) = 0\,, \quad k = 0, 1, 2, \ldots \tag{29}$$

with the $(n^a + n^b + 3)$ data vector

$$\boldsymbol{d}(k) = (\bar{\boldsymbol{y}}(k)', \bar{\boldsymbol{u}}(k)', 1)' \tag{30}$$

Figure 2 illustrates the relation between the measured data samples and the filtered

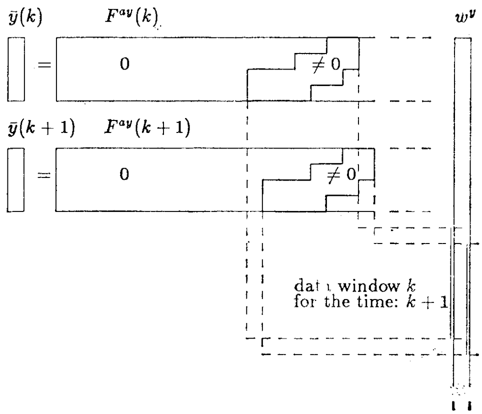

Fig. 2. Time evolution of the filtered data $\bar{y}(k)$.

signals. A small section of the semiinfinite data vector is needed only. The framed nonzero part of the $\boldsymbol{F}(\cdot)$ matrix remains the same, its position changes as it is shown schematically.

The model (29) is not unique – any of the parameters or the noise dispersion can be normalized to unity. The normalized parameter $\tilde{\boldsymbol{\theta}}$ is introduced (with $(n^a + n^b + 2)$ entries) by setting, as usual,

$$w_0^a = -1\,. \tag{31}$$

In this way, the regressand $y_0(k)$ and the regressor $z(k)$ are made explicit in the data vector $\boldsymbol{d}(k)$:

$$\bar{y}_0(k) = [[\bar{y}_1(k), \ldots, \bar{y}_{n^a}(k)]\,, \ \bar{\boldsymbol{u}}(k)', 1]\, \tilde{\boldsymbol{\theta}} + \bar{e}(k) = \tilde{\boldsymbol{\theta}}'\, z(k) + \bar{e}(k)\,. \tag{32}$$

Assuming a normal distribution for the sampled noise, the full strength of the Bayesian approach is available for the identification of the model. This means:

– an algorithm formally equivalent to the recursive least squares is optimal for gaining available information about the system [15];

- the efficient technique of restricted forgetting [14] for tracking slowly varying weights is immediately applicable;
- prior information about unknown parameters can be built in the estimation using the available theory [15], [9];
- the theory and algorithms are prepared allowing a prior data-based choice of the optimal structure of the data vector [8], [12].

Remarks.

1. The last item is of vital importance as we can rely on efficient tools for determining both the "degrees" n^a, n^b and the position of the grid points for the kernel approximation.

2. From the identification point of view, the availability of the vector of the filtered output $\bar{y}(k)$ defines the boundary between past and future. The time-delay in identification introduced due to the spline non-causality is at most a single sampling period.

6. CONTROL SYNTHESIS

Using the continuous model, we shall optimize the expected value $\mathsf{E}[\cdot]$ of the *integral* quadratic performance index J

$$J = \frac{1}{T}\int_0^T [Y^2(t) + q\, U^2(t)]\, \mathrm{d}t \tag{33}$$

specified by the potentially infinite horizon T and by the input penalty $q > 0$.

The selected regulation problem serves as a prototype for optimization: features such as a varying set point [10] (also modelled by splines), discounting etc. as well as extensions (hard bounds on the inputs) [2], [3] of the LQ design are directly applicable.

The control synthesis is performed assuming

- practically negligible approximation errors for all functions involved (relying heavily on the structure determination);
- complete knowledge of the parameter $\theta = [w^{a\prime}, w^{b\prime}, o]'$ (adopting the certainty-equivalence suboptimal control strategy usual for self-tuning controllers).

Using the first assumption, the loss function (33) can be written in terms of the approximative functions $y(t)$, $u(t)$

$$J = \frac{1}{T}\int_0^T [y^2(t) + q\, u^2(t)]\, \mathrm{d}t\,. \tag{34}$$

The signals involved are of the form (9), so we find that

$$J = w^{y\prime} Q^y w^y + w^{u\prime} Q^u w^u \tag{35}$$

where the penalty matrices Q^y, Q^u are evaluated in the following obvious way

$$Q^y = \frac{1}{T}\int_0^T f^y(t) f^{y\prime}(t) \mathrm{d}t \qquad Q^u = \frac{q}{T}\int_0^T f^u(t) f^{u\prime}(t)\, \mathrm{d}t. \tag{36}$$

These (potentially) infinite-dimensional penalty matrices are clearly positive semi-definite, symmetric *tridiagonal* (just two base functions have nonempty intersection of supports)

$$Q^x_{ij} = \begin{cases} \int_{t^x_{i-1}}^{t^x_{i+1}} (f^x_i(t))^2 \, \mathrm{d}t & \text{for} \quad i = j \\ \int_{t_i}^{t_{i+1}} f^x_i(t) f^x_{i+1}(t) \, \mathrm{d}t & \text{for} \quad |j - i| = 1 \\ 0 & \text{otherwise.} \end{cases} \quad \text{for} \quad x = y, u \tag{37}$$

Thus, the continuous criterion (33) converts into the discrete-time one (35) which attaches weights to the input-output samples w^u, w^y (cf. (12)). These samples are, according to the second assumption, related through the linear equation (32) with known coefficients. Introducing the "filtered" parameters

$$\bar{a}'(k) = w^{a'} F^{ay}(k), \quad \bar{b}'(k) = w^{b'} F^{bu}(k), \tag{38}$$

we can write (33) explicitly in terms of the sampled data. The filtered parameters are (potentially) infinite-dimensional vectors which depend on sampling times (cf. Fig. 3).

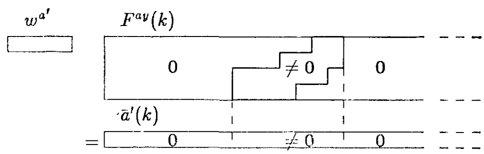

Fig. 3. The extension of the estimated weight $\bar{w}^{a'}$ to filtered parameters $\bar{a}'(k)$ used in the optimization.

They have, however, a finite number of *time-invariant nonzero entries* which shift with time. This is implied by the fact that the matrices $F^{ay}(k), F^{by}(k)$ have at most $n^y + 1 = IFIX(l^a/t_s + 3), n^u + 1 = IFIX(l^b/t_s + 3)$ nonzero, time-invariant columns respectively ($IFIX(\cdot)$ denotes the integer part of the argument; l_a, l_b are lengths of the kernel supports). Thus, the model for discrete samples relates $y(t^I_k)$, $y(t^I_{k-1})$, … …, $y(t^I_{k-n^y})$ and $u(t^I_k)$, $u(t^I_{k-1})$, …, $u(t^I_{k-n^u})$.

The evaluation of factors for tridiagonal positive definite matrices can be performed recursively with negligible computational demand per row. Thus, we can assume that factorized versions [1] of penalty matrices are available. Consequently, the factorized, discrete-time linear quadratic optimization for regression [10] generates the optimal discrete input samples. Using the relation (9) the continuous input is designed.

Remarks.

1. The use of nondiagonal penalty matrices in SISO system is somewhat unusal but fully justified in our case. Strong connections to data prefiltering discussed in [16] can be detected.

2. The definition (38) can be viewed as a special way of generating an overparametrized regression model (see Fig. 3). This observation could be used for inspecting relations between the described controller and the class labelled as predictive control [17].

3. A decrease of computational demands could be achieved by penalizing filtered signals (27) instead of the measured data. This possibility will be tested in the future.

7. CONCLUSIONS

The paper describes a way of constructing self-tuning controllers of a continuous-time-modelled system. It can be taken as an alternative to former attempts of designing digital self-tuners which respect continuous nature of the controlled process.

The key step of our design is the simultaneous approximation of signals and of the operator kernels by properly chosen splines with a finite support. In the paper, the conceptual feasibility of the approach is demonstrated yielding easily implementable adaptive controllers which promise

- an increase of the control quality due to improved modelling and a possibility of achieving shorter sampling periods than is usual for discrete-time controllers of a limited complexity;
- a computerized prior choice of data prefilters (left up to now to user's responsibility in other solutions).

Moreover, the used convolutional model provides a very general description of time-invariant linear system. The adopted way of its approximation is felt to be conceptually straightforward and can be applied directly to other cases, e.g. to 2D systems [11]. At the same time, smoothness of that signal part which is taken as worth to control comes explicitly to the designer's hands (a time domain couterpart of cross-over frequency?).

The paper reports a relatively early stage of the research which has produced applicable controllers but which is by no means complete. It will be necessary

- to perform extensive simulation and pilot-plant tests to gain experience of the sensitivity of the controllers to mismodelling (the simulation experience is encouraging);
- to assess the possibility of using smoothing splines instead of interpolating ones (the methodology developed in [18] seems to be applicable);
- to make the modelling of noise term more rigorous (an imbedding of noise space to a span over polynomials of unbounded order is studied to this purpose);

– to decrease the computational demands of the synthesis part (a direct spline approximation of performance index might be appropriate for this).

(Received April 21, 1989.)

REFERENCES

[1] G. J. Bierman: Factorization Methods for Discrete Sequential Estimation. Academic Press, New York 1977.

[2] J. Böhm and M. Kárný: Self-tuning regulators with restricted inputs. Kybernetika *18* (1982), 6, 529–544.

[3] J. Böhm: LQ self-tuners with signal level constraints. 7th Symp. Identification and Parameter Estimation, York, UK, preprints *1* (1985), 131–137.

[4] Ching-Tien and C. Yi-Shyong: Piecewise linear polynomial functions and applications to an analysis and parameter identification. Internat. J. Systems Sci. *18* (1987), 10, 1919–1929.

[5] L. Ching-Tien and C. Yi-Shyong: Operational matrices of piecewise linear polynomial functions with application to linear time-varying systems. Internat. J. Systems Sci. *18* (1987), 10, 1931–1942.

[6] P. J. Gawthrop: Continuous-time self-tuning control. A unified approach. In: Proc. of 2nd IFAC Workshop on Adaptive Systems in Control and Signal Processing, Lund, Sweden (1986), 19–24.

[7] G. C. Goodwin: Some observations on robust estimation and control. 7th Symp. Identification and Parameter Estimation, York, UK, preprints *1* (1985), 851–859.

[8] M. Kárný: Algorithms for determining the model structure of a controlled system. Kybernetika *19* (1973), 2, 164–178.

[9] M. Kárný: Quantification of prior knowledge about global characteristics of linear normal model. Kybernetika *20* (1984), 5, 376–385.

[10] M. Kárný, A. Halousková, J. Böhm, R. Kulhavý and P. Nedoma: Design of linear quadratic adaptive control: Theory and algorithms for practice. Supplement to Kybernetika *21* (1985), No. 3–6.

[11] M. Kárný, A. Halousková and I. Nagy: Modelling, identification and adaptive control of cross-direction basis weight of paper sheets. Internat. Conf. CONTROL 88, Oxford 1988, 159–164.

[12] M. Kárný and R. Kulhavý: Structure determination of regression-type models for adaptive prediction and control. In: Bayesian Analysis of Time Series and Dynamic Models (J. C. Spall, ed.), Marcel Dekker, New York 1988.

[13] N. P. Korneichuk: Splines in the Approximation Theory (in Russian). Nauka, Moscow 1984.

[14] R. Kulhavý: Restricted exponential forgetting in real-time identification. Automatica *23* (1987), 5, 598–600.

[15] V. Peterka: Bayesian system identification. In: Trends and Progress in System Identification (P. Eykhoff, ed.), Pergamon Press, Oxford 1981, 239–304.

[16] V. Peterka: Algorithms for LQG self-tuning control. In: Proc. of 2nd IFAC Workshop on Adaptive Systems in Control and Signal Processing Lund, 1986, Sweden, 13–18.

[17] J. Richalet: Why predictive control? 10th IFAC World Congress on Automatic Control 1987, preprint vol. 11. 7.

[18] R. Kohn and C. F. Ansley: Equivalence between Bayesian smoothness priors and optimal smoothing for function estimation. In: Bayesian Analysis of Time Series and Dynamic Models (J. C. Spall, ed.), Marcel Dekker, New York 1988.

Ing. Miroslav Kárný, CSc., Ing. Ivan Nagy, CSc., Ing. Josef Böhm, CSc., Ing. Alena Halousková, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.