Data Interpolation, Causality Structure, and System Identification[1]

WILLIAM A. PORTER[2]

Department of Electrical and Computer Engineering,
University of Michigan, Ann Arbor, Michigan 48104

Given a finite set $\{(x_i, y_i)\}$ of ordered pairs from $X \times Y$ where X, Y are Hilbert spaces over the same field, there are numerous techniques for constructing a function, f, on X to Y such that $f(x_i) = y_i$. However, when X, Y have a causality structure and f must be causal then the data interpolation problem is much more complicated. In this paper two interpolation methods, namely linear interpolation and interpolation via generalized Lagrange polynomials are considered. It is shown that these techniques can be modified to accomodate the causality constraint. The development is indicative of the modifications that must be made in any existing data interpolation algorithm if causal interpolation is required.

1. INTRODUCTION

In this paper the term system is synonomous with function. The range and domain of our function lying in a Hilbert resolution space $\{H, P^t\}$. It is assumed that our system is known only by experimental measurements, that is, a set of input–output pairs $E = \{(x_i, y_i), i = 1,..., n\}$ exist from experimentation. The primary problem is the following.

PRIMAL PROBLEM. Given the set E, construct a function f which (a) is causal and (b) $y_i = f(x_i)$, $i = 1, 2,..., n$.

Any solution to this problem is called an *identification* of the system. Of course the system *per se* is *not* identified, we have merely constructed a mathematical model which is consistent with existing data.

217

Without the causality constraint the primal problem has both linear and nonlinear solutions in abundance. Two such solutions will be summarized here.

Let H denote our Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and suppose that the input set $\{x_i\}$ is linearly independent. Then there exists a dual set $\{x_i^+\}$ with the properties (see [Porter, 1966] Section 3.2)

  (i) $\langle x_i^+, x_j \rangle = \delta_{ij}$
  (ii)   span $(\{x_i^+\})$ = span($\{x_i\}$).

The dual set satisfying both (i) and (ii) is in fact unique. We construct a linear map T by the formula

$$Tu = \sum_1^n y_i \langle x_i^+, u \rangle. \tag{1}$$

It is readily verified that T is a linear solution to the primal problem.

A second seolution is supplied by Prenter [1971] who generalized the Lagrange and Hermite interpolating polynomials to Banach spaces. In our setting it suffices to form the functionals

$$\psi_i(u) = \frac{\prod\limits_{k \neq i}^n \langle u - x_k, x_i - x_k \rangle}{\prod\limits_{k \neq i}^n \| x_i - x_k \|^2} \qquad i = 1,...,n.$$

It is easily verified that

$$\psi_i(x_j) = \delta_{ij}.$$

An immediate consequence is that the map

$$f(u) = \sum_1^n y_i \psi_i(u) \tag{2}$$

solves the primal problem.

It is apparent that our two solutions are not unique. Indeed if g is any function on H whose null space includes the set $\{x_i\}$ then g may be added to f, or if g is linear added to T, thereby constructing other solutions, respectively, linear solutions, to the primal problem. On the other hand no system can actually be identified without testing every input in its domain and hence the abundance of solutions to the primal problem should not disturb us.

The following sections of this paper analyze the constructions of Eqs. (1) and (2) when causality (definition to follow) is a constraint.

2. MATHEMATICAL PRELIMINARIES

The causality structure of interest in this paper is interwoven with the concept of a Hilbert resolution space. This has been thoroughly dealt with in [Porter and Zahm, 1969; DeSantis and Porter, 1973; Gohberg and Krein, 1970; Saeks, 1973 and Porter, 1973] and our review here will touch on only the features necessary to proceed with the analysis.

Let H denote a Hilbert space and ν a linearly ordered set. Without loss of generality we assume that ν has a minimal element, t_0, and a maximal element, t_∞, respectively. A family $\mathbb{R} = \{P^t \colon t \in \nu\}$ of orthoprojectors on H is a resolution of the identity if

   (i) $P^{t_0}(H) = 0$, $P^{t_\infty}(H) = H$ and $P^k(H) \supseteq P^l(H)$ whenever $k \geqslant l$

   (ii) $\mathbb{R}$ is strongly closed.

In (ii) we mean that if $\{P^j\}$ is a sequence of $\mathbb{R}$ such that $P^j x \to Px$, where P is an orthoprojector then $P \in \mathbb{R}$. Our results have an easy interpretation if ν is discrete. We focus attention on the more difficult continuous case. We say that $\mathbb{R}$ is *complete* if for every $k < l$, $k, l \in \nu$ there exists $m \in \nu$ such that $k < m < l$ and $P^l \subset P^m \subset P^l$. This means that for $t' \in \nu$

$$\lim_{t \to t'} \|(P^t - P^{t'})x\| = 0 \qquad \text{all} \quad t' \in \nu, \quad x \in H.$$

An integral type notation will also prove useful. For $y \in H$ arbitrary and m a scalar valued function on ν the integral I,

$$I = \int m(s)\, dP(s)\, y,$$

is interpreted in the following way (see [Porter, 1969], [DeSantis and Porter, 1973], and [Gohberg and Krein, 1970]). If ν is complete then with Ω a finite set $\{t_j \colon t_{j-1} < t_j, \; t_0 = t_0, \; t_N = t_\infty\}$

$$I = \lim_{\Omega \text{ refines}} \sum m(t_k')(P^{t_k} - P^{t_{k-1}})y,$$

where the limit is taken as Ω is refined. This limit exists and is unique whenever m is sectionally continuous. Indeed if the structure of ν is upgraded to that of a measure space the integral I can be justified for any square integrable m over this measure space (see [Masani, 1968], Section 5).

It is convenient at this point to note that the causality constraint requires certain consistency conditions on the experimental data. For any causal function f (see [Porter, 1969], [DeSantis and Porter, 1973; Gohberg and Krein, 1970])

$$P^t u = P^t v \Rightarrow P^t f(u) = P^t f(v), \qquad t \in \nu.$$

Thus an obvious condition on the experimental data is that

$$P^t x_i = P^t x_j \Rightarrow P^t y_i = P^t y_j, \qquad t \in \nu. \tag{3}$$

Moreover in the linear realization since $0 = T(0)$ the stronger condition

$$P^t x_i = 0 \Rightarrow P^t y_i = 0 \qquad t \in \nu \tag{4}$$

must hold.

Other consistency conditions also appear natural. In the linear model if the input set is linearly dependent then the output set must submit to the same zero linear combination. In the polynomic case the inputs must be distinct. These observations lead to the following definition.

DEFINITION 1.   The primal problem is said to be *well posed* if

$$\{P^t x_i \colon i = 1,..., n\}$$

is linearly independent for all $t \neq t_0$.

The order relation on the projection family assures that the rank of the set $\{P^t x_i\}$ is nondecreasing with increasing t. The well-posed condition then focuses on the immediate neighborhood of t_0. The apparant severity of this assumption is removed in Section 6.

3. THE LINEAR MODEL

As a prelude to the causal linear identification let us consider a map $\eta \colon \nu \to H$. If ν is complete then η may be continuous (norm) that is

$$\lim_{t \to a} \| \eta(t) - \eta(a) \| = 0, \qquad \alpha \in \nu.$$

We may then think of η as a path in H.

Suppose now that $N = \{\eta_i \colon i = 1,..., n\}$ is a family of paths in H. N is said to be *nondegenerate* if the set $\{\eta_i(t) \colon i = i,..., n\}$ is linearly independent for

every $t \in \nu$. If ν has the structure of a measure space then the nondegenerate condition may be modifiable a.e.

For fixed $t \in \nu$ let $\{\eta_i(t)\}$ be linearly independent. Then there exists a dual set $\{\eta_i^+(t)\}$ having the properties

$$\langle \eta_i^+(t), \eta_j(t) \rangle = \delta_{ij} \tag{5}$$
$$\mathrm{span}\{\eta_i^+(t)\} = \mathrm{span}\{\eta_i(t)\}.$$

Indeed assuming coefficients $\{\alpha_{ij}\}$ such that

$$\eta_i^+ = \sum_1^n \alpha_{ij} \eta_j$$

and employing Eq. (5) we arrive at the matrix identity

$$[\alpha_{ij}] = [\langle \eta_j, \eta_i \rangle]^{-1},$$

the invertibility of $[\langle \eta_j, \eta_i \rangle]$ being a well-known consequence of the linear independence of $\{\eta_i\}$. For instance with $n = 2$

$$\det[\langle \eta_j, \eta_i \rangle] = \| \eta_1 \|^2 \| \eta_2 \|^2 - |\langle \eta_1, \eta_2 \rangle|^2$$

which, using the Cauchy inequality, shows that linear independence is necessary and sufficient.

Suppose N is a nondegenerate set of continuous paths. Then

$$\lim_{\epsilon \to 0} \{ \langle \eta_i(t + \epsilon), \eta_j(t + \epsilon) \rangle - \langle \eta_i(t), \eta_j(t) \rangle \}$$
$$= \lim_{\epsilon \to 0} \{ \langle \eta_i(t + \epsilon) - \eta_i(t), \eta_i(t + \epsilon) - \eta_i(t) \rangle$$
$$+ \langle \eta_i(t + \epsilon) - \eta_i(t), \eta_j(t) \rangle - \langle \eta_i(t), \eta_j(t + \epsilon) - \eta_j(t) \rangle \}$$
$$\leqslant \lim_{\epsilon \to 0} \{ \| \eta_i(t + \epsilon) - \eta_i(t) \|^2 + \| \eta_i(t + \epsilon) - \eta_i(t) \| \cdot \| \eta_j(t) \|$$
$$+ \| \eta_i(t) \| \cdot \| \eta_j(t + \epsilon) - \eta_j(t) \| \}$$

which shows that $t \to \langle \eta_i(t), \eta_j(t) \rangle$ is a continuous function. Similarly the determinant, as a sum of products of continuous functions is continuous as is every cofactor of $[\langle \eta_j, \eta_i \rangle]$. It follows then that the functions $N^+ = \{\eta_i^+(t)\}$ inherit continuity from $N = \{\eta_i(t)\}$. (For convenience we assume det bounded away from zero.) Finally if $h \in H$ is arbitrary then $t \to \langle \eta_j^+(t), h \rangle$ is a continuous scalar valued map.

Return ·now to the resolution family $\{P^t : t \in \nu\}$. For arbitrary $x \in H$ the map $t \to P^t x$ is a continuous path with $0 \leqslant \| P^t x \| \leqslant \| x \|$, and $\| P^t x \|$ is

monotone nondecreasing. For a linearly independent set $\{x_i\}$ of H we define the $t \to H$ functions

$$\hat{\eta}_i(t) = P^t x_i \qquad i = 1, 2, ..., n.$$

PROPOSITION (a). *If for $t' \in \nu$ $\{\hat{\eta}_i(t')\}$ is linearly independent then $\{\hat{\eta}_i(t)\}$ is linearly independent for all $t' \leqslant t \in \nu$.*

Proof. $\sum \alpha_i P^t x_i = 0 \Rightarrow \sum \alpha_i P^{t'} x_i + \sum \alpha_i (P^t - P^{t'}) x_i = 0$ however since Range $(P^{t'}) \perp$ Range $(P^t - P^{t'})$ we have $\sum \alpha_i P^{t'} x_i = 0$, $\Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$

PROPOSITION (b). *If for $t \in \nu$ $\{\hat{\eta}_i(t)\}$ is linearly dependent then $\{\hat{\eta}_i(t')\}$ is linearly dependent for all $t' \leqslant t$.*

Proof [same style as Proposition (a)]. Suppose now that the primal problem is well posed. The set $\{\eta_j\}$ is constructed according to the formula

$$\eta_j(t) = \frac{P^t x_j}{\| P^t x_j \|} \qquad t \in \nu, \quad j = 1, ..., n \tag{6}$$

and consists of a family of continuous paths on the unit ball with linearly independent values at all times.

The dual set $\{\eta_j^+\}$ consists of continuous paths with linearly independent values. We shall make use, however, of the set $\{Q_j\}$ computed by

$$Q_j(t) = \| P^t x_j \|^{-1} \eta_j^+(t) \qquad t \in \nu. \tag{7}$$

PROPOSITION (c). *The set $\{Q_j\}$ consists of continuous paths with the property*

    (i) $\langle Q_j(t), x_i \rangle = \delta_{ij}$

    (ii) $P^\beta u = 0 \Rightarrow \langle Q_j(t), u \rangle = 0 \quad t \leqslant \beta.$

Both of the stated properties follow by inspection.

Using the continuity of Q_j we see that the transformation

$$T_j u = \int dP(s) \, y_j \langle Q_j(s), u \rangle \qquad u, y_j \in H \tag{8}$$

is well defined (see Section 2) linear and bounded.

THEOREM 1. *The well-posed problem has the causal linear solution*

$$T = \sum_1^n T_j$$

where T_j, Q_j are given by Eqs. (8) and (7), respectively.

Proof. The linearity of T is evident. Property (i) of Proposition (c) guarantees that $y_j = Tx_j$, $j = 1,..., n$. To see that T is causal notice that

$$P^\beta T_j u = P^\beta \int dP(s) \, y_j \langle Q_j(s), u \rangle$$

$$= \int_{s \leqslant \beta} dP(s) \, y_j \langle Q_j(s), u \rangle$$

$$= \int_{s \leqslant \beta} dP(s) \, y_j \langle Q_j(s), P^\beta u \rangle$$

$$= P^\beta T_j P^\beta u,$$

where property (ii) was used.

4. THE POLYNOMIC SOLUTION

The polynomic solution to the well-posed problem is akin in spirit to the linear solution. With $\mathscr{F}$ the scalar field of H let K denote the set of all functions $f : \nu \times H \to \mathscr{F}$. We are interested in the n^2 elements $m_{ij} \in K$, $i, j = 1,..., n$ defined by

$$m_{ij}(u, t) = \begin{cases} \dfrac{\langle P^t(u - x_j), (x_i - x_j) \rangle}{\| P^t(x_i - x_j) \|^2} & u \in H, \quad t_0 \neq t \in \nu \\[4mm] 0 & t = t_0 . \end{cases} \tag{9}$$

We note that if $\{x_j\}$ are from a well-posed problem the m_{ij} are well defined. The relevant properties of m_{ij} are the following:

(i) $P^t u = P^t v \Rightarrow m_{ij}(u, \beta) = m_{ij}(v, \beta)$ $\beta \leqslant t$

(ii) $m_{ij}(0, t) - m_{ji}(0, t) = 1$ $t_0 \neq t \in \nu$

(iii) $m_{ij}(x_i, t) = 1$ $t_0 \neq t \in \nu$

(iv) $m_{ij}(x_j, t) = 0$ $t \in \nu$

(v) $m_{ij}(u, t)$ is continuous except at t_0 .

The first four properties follow by inspection. Property (v) assumes that ν is complete and follows from the identity

$$m_{ij}(u, t'') - m_{ij}(u, t') = \frac{\langle \triangle(u - x_j), (x_i - x_j) \rangle}{\| \triangle(x_i - x_j) \|^2 + \| P^{t'}(x_i - x_j) \|^2}$$
$$- \frac{\| \triangle(x_i - x_j) \|^2 \, \langle P^{t'}(u - x_j), (x_i - x_j) \rangle}{\| P^{t'}(x_i - x_j) \|^2 \cdot \| P^{t'}(x_i - x_j) \|^2} ,$$

where $\triangle = P^{t''} - P^{t'}$.

The m_{ij} functions are intermediary to the n functions M_i which are defined by

$$M_i(u, t) = \prod_{\substack{j \neq i \\ j=1}}^{n} m_{ij}(u, t) \qquad i = 1,\dots, n. \tag{10}$$

The relevant properties of M_i are inherited from the m_{ij} and include

(i') $P^t u = P^t v \Rightarrow M_i(u, \beta) = M_i(v, \beta) \quad \beta \leqslant t$

(ii') $M_i(x_j, t) = \delta_{ij}$

(iii') $M_i(u, t)$ is continuous except at t_0.

Once again these properties presume that the $\{x_j\}$ come from a well-posed problem and that v is complete.

THEOREM 2. *The causal Lagrange interpolation of the well-posed problem is given by*

$$f(u) = \sum_{i=1}^{n} \int dP(s)\, y_i M_i(u, s), \tag{11}$$

where M_i is defined in Eq. (10).

Proof. Property (i') provides the causality of f. Property (ii') assures $y_i = f(u_i)$ and Property (iii') gurantees the existence of the integral of Eq. (11).

5. COMPARATIVE EXAMPLE

It is constructive to compare the two data interpolation techniques by means of a concrete example. For this let $v = (0, 1]$ and $H = L_2(0, 1)$ equipped with the (truncation) projections

$$(P^t x)(s) = \begin{cases} x(s) & s \leqslant t \\ 0 & s > t \end{cases} \qquad t, s \in (0, 1]. \tag{12}$$

It is easily verified that $\{H, P^t\}$ is a Hilbert resolution space.

Suppose now that three input–output pairs have been observed. To be explicit let

$$\begin{aligned} x_1(t) &= 1 & y_1(t) &= \sin t \\ x_2(t) &= t & y_2(t) &= \cos t \\ x_3(t) &= t^2 & y_3(t) &= e^t, \quad t \in v. \end{aligned}$$

Using Eq. (12) and mundane computations it follows that

$$\| P^t x_1 \|^2 = \int_0^t 1^2 \, d\beta = t \qquad t \in \nu$$

$$\| P^t x_2 \|^2 = \int_0^t \beta^2 \, d\beta = (1/3) \, t^3 \qquad t \in \nu$$

$$\| P^t x_3 \|^2 = \int_0^t \beta^4 \, d\beta = (1/5) \, t^5 \qquad t \in \nu$$

$$\| P^t (x_1 - x_2) \|^2 = \int_0^t (\beta - 1)^2 \, d\beta = (1/3) \, t^3 - t^2 + t \qquad t \in \nu \tag{13}$$

$$\| P^t (x_1 - x_3) \|^2 = \int_0^t (\beta^2 - 1)^2 \, d\beta = (1/5) \, t^5 - (2/3) \, t^3 + t \qquad t \in \nu$$

$$\| P^t (x_2 - x_3) \|^2 = \int_0^t (\beta^2 - \beta)^2 \, d\beta = (1/5) \, t^5 - (1/2) \, t^4 + (1/3) \, t^3 \qquad t \in \nu$$

which verifies that the problem is well posed.

For convenience we treat the somewhat shorter polynomial realization first. The computation of Eq. (9) is very explicit, for example,

$$m_{12}(u, t) = \| P^t (x_1 - x_2) \|^{-2} \int_0^t [u(\beta) - \beta][1 - \beta] \, d\beta, \qquad t \in \nu.$$

Using these formulas and Eq. (13), the functions $M_i(u, t)$ of Eq. (10) are explicitly given by

$$M_1(u, t) = \| P^t (x_1 - x_2) \|^{-2} \cdot \| P^t (x_1 - x_3) \|^{-2}$$
$$\cdot \left\{ \int_0^t [u(\beta) - \beta][1 - \beta] \, d\beta \right\} \cdot \left\{ \int_0^t [u(\beta) - \beta^2][1 - \beta^2] \, d\beta \right\}$$

$$M_2(u, t) = \| P^t (x_1 - x_2) \|^{-2} \cdot \| P^t (x_2 - x_3) \|^{-2}$$
$$\cdot \left\{ \int_0^t [u(\beta) - \beta][1 - \beta] \, d\beta \right\} \left\{ \int_0^t [u(\beta) - \beta^2][1 - \beta^2] \, d\beta \right\} \tag{14}$$

$$M_3(u, t) = \| P^t (x_1 - x_3) \|^{-2} \cdot \| P^t (x_2 - x_3) \|^{-2}$$
$$\cdot \left\{ \int_0^t [u(\beta) - 1][\beta^2 - 1] \, d\beta \right\} \left\{ \int_0^t [u(\beta) - \beta][\beta^2 - \beta] \, d\beta \right\}.$$

It remains only to evaluate Eq. (11). For every $u \in H$ the functions $M_i(u, t)$ are continuous in t, thus the integral of Eq. (11) exists and is well defined as

we have noted earlier. In the present example each y_i is itself a continuous function. Using the above resolution space structure it follows easily that

$$dP(s)\, y_i = \lim_{\epsilon \to 0}(P^{s+\epsilon} - P^{s-\epsilon})\, y_i = \begin{cases} 0 & t \neq s \\ y_i(t) & t = s \end{cases}$$

hence

$$dP(s)\, y_i M_i(u, s) = y_i(s)\, M_i(u, s), \qquad s \in \nu.$$

It follows also that $f(u) \in H$ is the continuous function whose value at $t \in \nu$ is given by

$$[f(u)](t) = \sin(t)\, M_1(u, t) + \cos(t)\, M_2(u, t) + e^t M_3(u, t). \qquad (15)$$

By obvious manipulations on Eq. (14) Eq. (15) can be reorganized into a (nonlinear) integral kernal form.

The computations in the linear interpolator are just as straightforward although somewhat more arduous. Using Eq. (13) we first identify $\{\eta_i(t)\}$. Recalling that $\eta_i(t)$ for fixed $t \in \nu$ is an element of $L_2(0, 1)$ we have

$$\begin{aligned}
[\eta_1(t)](\beta) &= \begin{cases} t^{-1/2} & \beta \leqslant t \\ 0 & \beta > t \end{cases} \\[2mm]
[\eta_2(t)](\beta) &= \begin{cases} (3)^{\frac{1}{2}}\, t^{-3/2}\beta & \beta \leqslant t \\ 0 & \beta > t \end{cases} \qquad (16) \\[2mm]
[\eta_3(t)](\beta) &= \begin{cases} (5)^{\frac{1}{2}}\, t^{-5/2}\beta^2 & \beta \leqslant t \\ 0 & \beta > t. \end{cases}
\end{aligned}$$

The second step is to compute $N = [\langle \eta_i(t), \eta_j(t) \rangle]$ which in general will depend on t (but not β). In the present case N is constant and given by

$$N = \begin{bmatrix} 1 & (3)^{\frac{1}{2}}/2 & (5)^{\frac{1}{2}}/3 \\ (3)^{\frac{1}{2}}/2 & 1 & (15)^{\frac{1}{2}}/4 \\ (5)^{\frac{1}{2}}/3 & (15)^{\frac{1}{2}}/4 & 1 \end{bmatrix}.$$

The inverse of this matrix is given by (we spare the reader some details)

$$N^{-1} = \begin{bmatrix} 9 & -12(3)^{\frac{1}{2}} & 6(5)^{\frac{1}{2}} \\ -12(3)^{\frac{1}{2}} & 64 & -12(15)^{\frac{1}{2}} \\ 6(5)^{\frac{1}{2}} & -12(15)^{\frac{1}{2}} & 36 \end{bmatrix}.$$

The rows of N^{-1} contain the coefficients α_{ij} needed in the formula

$$\eta_i^+(t) = \sum_{j=1}^{3} \alpha_{ij} \eta_j(t) \qquad i = 1, 2, 3 \qquad t \in \nu.$$

For instance

$$\eta_1^+(t) = 9\eta_1(t) - 12(3)^{\frac{1}{2}}\eta_2(t) + 6(5)^{\frac{1}{2}}\eta_3(t) \qquad t \in v.$$

Recognizing that $\eta_i(t)$ for fixed $t \in v$ still depends on another variable and noting that $[Q_i(t)](\beta) = \| P^t x_i \|^{-1}[\eta_i^+(t)](\beta)$ it follows directly that (here $t \neq 0$)

$$[Q_1(t)](\beta) = \begin{cases} 3t^{-1}[3 - 12\beta t^{-1} + 10\beta^2 t^{-2}] & \beta \leqslant t \\ 0 & \beta > t \end{cases}$$

$$[Q_2(t)](\beta) = \begin{cases} 12t^{-2}[-3 + 16\beta t^{-1} - 15\beta^2 t^{-2}] & \beta \leqslant t \\ 0 & \beta > t \end{cases} \qquad (17)$$

$$[Q_3(t)](\beta) = \begin{cases} 30t^{-3}[1 - 6\beta t^{-1} + 6\beta^2 t^{-2}] & \beta \leqslant t \\ 0 & \beta > t. \end{cases}$$

Turning now to Eq. (8) we have, for instance, that

$$\langle Q_1(t), u \rangle = \int_0^t 3t^{-1}[3 - 12\beta t^{-1} + 10\beta^2 t^{-2}] u(\beta) \, d\beta$$

and from our above observations,

$$\int dP(t) \, y_1 \langle Q_1(t), u \rangle = \int_0^t \frac{3y_1(t)}{t} [3 - 12\beta t^{-1} + 10\beta^2 t^{-2}] u(\beta) \, d\beta.$$

To draw our results together let $w(t, \beta)$ be defined on $(0, 1]^2$ by

$$w(t, \beta) = \frac{3 \sin t}{t} \left[3 - 12 \left(\frac{\beta}{t} \right) + 10 \left(\frac{\beta}{t} \right)^2 \right]$$

$$+ \frac{12 \cos t}{t^2} \left[-3 + 16 \left(\frac{\beta}{t} \right) - 15 \left(\frac{\beta}{t} \right)^2 \right]$$

$$+ \frac{30e^t}{t^3} \left[1 - 6 \left(\frac{\beta}{t} \right) + 6 \left(\frac{\beta}{t} \right)^2 \right].$$

Then the linear interpolation T, is given by

$$(Tu)(t) = \int_0^t w(t, \beta) \, u(\beta) \, d\beta. \qquad (18)$$

The reader may wish it verify directly that $y_i = Tx_i\, i = 1, 2, 3$. The linearity and causality of T follows from Eq. (17).

6. Extensions, Remarks, and Conclusion

The example of Section 5 provides the basis for a comparison of the two causal data interpolation techniques. First the Lagrange interpolator of Eq. (15) is arrived at through computations which are apparently easier in that linearity requires the inversion of the matrix N. Secondly, the Lagrange interpolator adjusts more readily to additional input–output pairs. Suppose that Eqs. (15) and (18) are in hand and an additional input–output pair (x^*, y^*) then become available. The existing functions $\{m_{ij}\}$ are undisturbed while some additional ones are computed and the functions $\{M_i\}$ are adjusted easily by including the requisite additional factor. In the linear interpolator N changes and N^{-1} must be recomputed.

Suppose now that $x^* = \sum \alpha_i x_i$ and that linearity is not violated, that is $y^* = \sum \alpha_i y_i$. The linear interpolator requires no adjustment, the polynomic interpolator does. Of course if linearity is violated, linear interpolation must be adjusted to a "best fit" context.

Our attention now turns to ill-posed problems. Since the data can fail to be well posed in a variety of ways we shall be content to touch on two of the adjustments. Suppose that the input pair (x_1, y_1) violates the well-posed condition, in particular suppose that for

$$\gamma = \inf\{t: P^t x_1 \neq 0\}$$

we have $t_0 < \gamma$. Both interpolation methods can proceed with the size of the data set changing at $t = \gamma$. The result is a possible time discontinuity in the functions $\{Q_j\}$ and $\{M_j\}$. This complicates but does not destroy either procedure.

In the well-posed condition of Definition 1, an apparently severe constraint is imposed, namely that $\{P^t x_i, i = 1,..., n\}$ is assumed linearly independent for all $t \neq t_0$. This assumption can be alleviated without difficulty. To see how this can be achieved suppose that $t_1^* = \inf\{t \in \nu: \text{rank}(\text{span}\{P^t x_i\}) = n\}$, and let $n_1 = \text{rank}(\text{span}\{P^{t_1^*} x_i\})$. Now assuming $t_1^* \neq t_0$ and $n_1 \neq 0$ we continue $t_2^* = \inf\{t \in \nu: \text{rank}(\text{span}\{P^t x_i\}) \leqslant n_1\}$ while $n_2 = \text{rank}(\text{span}\{P^{t_2^*} x_2\})$, until the interval ν is divided into at most n parts with the rank constant on the parts into at most n parts with the rank constant on the parts and monotone increasing. The linear solution to the primal problem proceeds as before except the number of pairs (x_i, y_i) considered is cut down to the rank number of the interval in question (always choosing a basis). Since linearity is assumed *not* to be violated, the excluded data will still be reproduced by the linear solution.

Finally, the example of Section 5 benefited greatly from the occurrence of a constant matrix N. The conditions under which this occurs are rather complex and will be dealt with in the appendix.

<h2>APPENDIX</h2>

In the comparative example of Section 5 the matrix $N = [\langle \eta_i(t), \eta_j(t) \rangle]$ turned out to be independent of t. As this result is unexpected, we explore it fully in this appendix.

Consider the power functions $x_j(t) = t^j$, $t \in [0, T]$, $j = 0, 1,..., n - 1$. By direct inspection we see that

$$\| P^t x_j \|^2 = \int_0^t \beta^{2j} \, d\beta = (2j + 1)^{-1} t^{2j+1}$$

and hence the η_j functions, in this case, take the form

$$\eta_j(\beta, t) = \begin{cases} (2j + 1)^{\frac{1}{2}} t^{\frac{1}{2}} \left(\dfrac{\beta}{t} \right)^j & 0 \leqslant \beta \leqslant t \\ 0 & t < \beta \leqslant T. \end{cases} \quad t \in (0, T] \qquad (19)$$

A consequence of this is that $N_{ij}(t)$ is independent of t. In fact

$$N_{ij} = \langle \eta_i(\cdot, t), \eta_j(\cdot, t) \rangle = [(2j + 1)(2i + 1)]^{\frac{1}{2}} t^{-1} \int_0^t (\beta/t)^{i+j} \, d\beta$$

$$= [(2j + 1)(2i + 1)]^{\frac{1}{2}} \int_0^1 \alpha^{i+j} \, d\alpha$$

$$= \left[\frac{(2j + 1)(2i + 1)}{(i + j + 1)^2} \right]^{\frac{1}{2}} \quad i, j = i,..., n.$$

Thus any collection of power functions has the interval invariance property. In retrospect, Eq. (19) shows that the power functions have the property

$$x(\beta) = \lambda_x t^{-1/2} x(\beta/t) \| P^t x \|, \qquad \beta \leqslant t \in (0, T]. \qquad (20)$$

If x, y satisfy Eq. (20) then

$$\int_0^t x(\beta) y(\beta) \, d\beta = \lambda_x \lambda_y \| P^t x \| \cdot \| P^t y \| \int_0^1 x(\alpha) y(\alpha) \, d\alpha \qquad (21)$$

holds and the interval invariance property follows. This discovery, however, is of no assistance in moving beyond the power functions as we see in the lemma.

LEMMA 1.   *x satisfies Eq. (20) if and only if x is a power function.*

Proof.   It is convenient to use the integral form of Eq. (20)

$$x(\beta) = \lambda \left[\int_0^t x^2(\alpha) \, d\alpha \right]^{\frac{1}{2}} t^{-\frac{1}{2}} x \left(\frac{\beta}{t} \right) \qquad \beta \leqslant t \in (0, T].$$

The continuity (and differentiability) of the multiplier of $x(\beta/t)$ for $t > 0$ can be seen to imply (fix β) the differentiability of x. Differentiating with respect to β results in

$$\dot{x}(\beta) = \lambda \left[\int_0^t x^2(\alpha) \, d\alpha \right]^{\frac{1}{2}} t^{-\frac{3}{2}} \dot{x} \left(\frac{\beta}{t} \right), \qquad \beta \leqslant t.$$

Assuming $x(\beta) \neq 0$ and dividing we have

$$\frac{\dot{x}(\beta)}{x(\beta)} = \left(\frac{1}{t} \right) \frac{\dot{x}(\beta/t)}{x(\beta/t)} \qquad \beta \leqslant t \in (0, T]$$

or for $\beta \neq 0$

$$\beta \frac{\dot{x}(\beta)}{x(\beta)} = \left(\frac{\beta}{t} \right) \frac{\dot{x}(\beta/t)}{x(\beta/t)} \qquad \beta \leqslant t \in (0, T].$$

The independence of variables β, t implies that both sides of this last equality are constant, hence

$$\dot{x}(\beta) = \frac{k}{\beta} x(\beta)$$

for some k, the solution to which is $x(\beta) = c\beta^k$.

Interval Invariance Criteria

In view of Lemma 1 it remains only to examine Eq. (21). Our next lemma provides a necessary and sufficient condition.

LEMMA 2.   *In $L_2(0, T)$, $\langle P^t x, P^t y \rangle = \lambda \| P^t x \| \cdot \| P^t y \|$, $t \in [0, T]$ if and only if $x(t) \| P^t y \| = \gamma y(t) \| P^t x \|$, $t \in [0, T]$.*

Proof. It suffices to consider $\lambda > 0$ in which case $\langle P^t x, P^t y \rangle$ is non-negative and we may square both sides of the hypothesis

$$\left[\int_0^t x(s)\, y(s)\, ds \right]^2 = \lambda^2 \int_0^t x^2(s)\, ds \cdot \int_0^t y^2(s)\, ds, \qquad t \in (0, T]. \quad (22)$$

Differentiating both sides we have

$$2x(t)\, y(t) \int_0^t x(s)\, y(s)\, ds = \lambda^2 \left\{ x^2(t) \int_0^t y^2(s)\, ds + y^2(t) \int_0^t x^2(s)\, ds \right\}.$$

Using the hypothesis of the lemma

$$2\lambda x(t)\, y(t)\, \| P^t x \| \cdot \| P^t y \| = \lambda^2 \{ \| P^t y \|^2 x^2(t) + \| P^t x \|^2 y^2(t) \} \qquad t \in (0, T]$$

rearranging terms we have on $(0, T]$

$$[\lambda \| P^t y \| x(t)]^2 - [2 \| P^t x \| y(t)][\lambda \| P^t y \| x(t)] + [\lambda \| P^t x \| y(t)]^2 = 0.$$

Using the quadratic formula we have

$$\lambda \| P^t y \| x(t) = (1 \pm (1 - \lambda^2)^{\frac{1}{2}}) \| P^t x \| y(t), \qquad t \in (0, T].$$

This establishes Lemma 2 and using the symmetry of Eq. (22) in x, y we have moreover that

$$\gamma = (\lambda^{-1} \pm (\lambda^{-1} - 1)^{\frac{1}{2}}) \pm 1.$$

The case $\lambda = 0$ results easily in $x(t)\, y(t) = 0$, $t \in [0, T]$ and for $\lambda < 0$ we replace x by $-x$.

To illustrate the use of Lemma 2 let us prove that

LEMMA 3. *If $y(t) = t^j$, $t \in [0, T]$ for arbitrary j then $x \in L_2\,(0, T)$ satisfies the hypothesis of Lemma 2 if and only if $x(t) = \alpha t^m$ for some α, m.*

Proof. For $y(t) = t^j$, $t \in [0, T]$ we have from Eq. (19) that

$$\frac{y(t)}{\| P^t y \|} = (2j + 1)^{\frac{1}{2}}\, t^{-\frac{1}{2}} \qquad t \in (0, T].$$

In view of Lemma 2 it suffices to determine all $x \in L_2(0, T)$ satisfying

$$x(t) = \gamma \left[(2j + 1)\, t^{-1} \int_0^t x^2(s)\, ds \right]^{\frac{1}{2}}.$$

Since the right-hand side is positive we can square both sides without loss of generality and setting $z(t) = x^2(t)$, $t \in (0, T]$ we have

$$tz(t) = \gamma^2(2j + 1) \int_0^t z(s)\, ds, \qquad t \in (0, T].$$

Differentiation yields

$$z(t) + t\dot{z}(t) = \gamma^2(2j + 1)\, z(t)$$

or rearranging

$$\dot{z}(t) = \mu t^{-1} z(t), \qquad t \in (0, T]$$

where $\mu = (2j + 1)\,\gamma^2 - 1$. The unique solution to this last equation is

$$z(t) = \lambda t^\mu = x^2(t), \qquad t \in (0, T].$$

Since $x(t) \geqslant 0$ the lemma follows.

In the applications that give rise to this Appendix the condition $\| P^t x \| \neq 0$, $t > 0$ arises naturally. We shall adopt this condition in the following Corollary to Lemma 2 which is a constructive technique for generating interval invariant sets.

COROLLARY 1.   *In $L_2(0, T)$ the relations $\langle P^t x, P^t y \rangle = \lambda \| P^t x \| \cdot \| P^t y \|$, $t \in [0, T]$; $\| P^t x \| \neq 0$, $\| P^t y \| \neq 0$, $t > 0$ hold if and only if*

$$x(t) = \beta y(t) \left[\int_0^t y^2(s)\, ds \right]^{(\lambda-1)/2} \qquad t \in (0, T] \tag{23}$$

holds for some scalar β.

Proof.   Noting that $x^2(t)[\int_0^t x^2(s)\, ds] = d \ln [\int_0^t x^2(s)\, ds$ we have by squaring the necessary and sufficient equation of Lemma 2

$$\frac{x^2(t)}{\| P^t x \|^2} = \lambda \frac{y^2(t)}{\| P^t y \|^2} \Leftrightarrow d \ln \left[\int_0^t x^2(s)\, ds \right] = \lambda d \ln \left[\int_0^t y^2(s)\, ds \right]$$

$$\Leftrightarrow d \ln \left[\int_0^t x^2(s)\, ds \Big/ \left(\int_0^t y^2(s)\, ds \right)^\lambda \right] = 0$$

$$\Leftrightarrow \int_0^t x^2(s)\, ds = e^c \left[\int_0^t y^2(s)\, ds \right]^\lambda. \tag{24}$$

Equation (23) follows by taking derivatives, then square roots while noting that $x(t)$ and $y(t)$ have the same sign.

Equation (24) reveals how close the power functions and the interval invariance property are tied together. By inspection we see that the following corollary holds.

COROLLARY 2.   *The set $\{x_i\} \subset L_2(0, T)$ is interval invariant if and only if the set $\{\hat{x}_i\}$, where $\hat{x}_i(t) = \| P^t x_i \|^2$ are powers of $\hat{x}_1$.*

EXAMPLE.   Suppose that we wish to construct an interval invariant set containing $y(t) = e^t$, $t \in [0, T]$. We have immediately

$$\int_0^t y^2(s) \, ds = \tfrac{1}{2}[e^{2t} - 1] \qquad t \in [0, T],$$

and any x satisfying

$$x(t) = \beta e^t [e^{2t} - 1]^{(\lambda-1)/2} \qquad t \in (0, T]$$

for $|\lambda| < 1$ is a suitable addition to the set.

REFERENCES

PORTER, W. A. (1966), "Modern Foundations of Systems Engineering," The Macmillan Co., Inc., New York.

PRENTER, P. M. (1971), Lagrange and Hermite interpolation in Banach space, *J. Approx. Theory* **4**, 419–432.

PORTER, W. A. AND ZAHM, C. (1969), "Basic Concepts in System Theory," *Systems Engineering Lab. Tech. Rep. No.* **44**, The University of Michigan.

DESANTIS, R. M. AND PORTER, W. A. (1973), On time related properties of nonlinear systems, *SIAM J. Appl. Math.* **24**, 188–206.

GOHBERG, I. Z. AND KREIN, M. G. (1970), Theory of Volterra operators in Hilbert space and applications, *Amer. Math. Soc.* **24**.

SAEKS, R. (1973), "Resolution Space Operators and Systems," Lecture Notes, Springer-Verlag, New York.

PORTER, W. A. (1971), Nonlinear systems in Hilbert resolution spaces, *Int. J. Contr.* **13**, 593–602.

MASANI, P. (1968), Orthogonally scattered measures, *Advan. Math.* **2**, 61–117 (Amer. Math. Soc.).