ARKIV FÖR MATEMATIK Band 6 nr 12

Read 28 October 1964

A fix-point theorem with econometric background

Part I. The theorem

By Herman O. A. Wold

The theorem provides a new approach to a problem that has originated in econometrics, namely the statistical estimation of parameters of so-called interdependent systems. We formulate the problem generally as one of orthogonal projection in Hilbert vector space, an interpretation based on the definition of interdependent systems in terms of conditional expectations; [5–7]. The subspace on which the vectors are projected is spanned by vectors some of which are given, whereas others—and this is a nonlinear feature of the problem—are unknown and to be determined by the projection. The problem is solved by an iterative least squares procedure that rests on a new application of the principle of contraction mapping.

Part I presents the fix-point theorem and the ensuing iterative procedure. Part II gives illustrations and comments, taking up three specifications of the components of the given vectors y, z: (i) As vectors in Euclidean space R_h; (ii) as random variates, making y, z a multivariate random distribution; (iii) as statistical data, interpreted as a sample from the distribution under (ii). Extensions in various directions are briefly outlined. It is emphasized that the assumptions of the approach broaden the scope of interdependent systems.

For the mathematical groundwork utilized in Part I, see [2 and 4], and, especially as regards contraction mapping, [3].

*

From October 1963 onwards some twenty numerical experiments with the iterative procedure have been carried out at the Computing Centre of the Royal Polytechnic Institute of Norway, Trondheim. I wish to express my deep gratitude to Mr. H. Michalsen, Programming Consultant, and Mr. N. Sanders, Director of the Computing Centre, for their valuable help in the numerical work. My theoretical work on the present paper has to a large extent been carried through at the Univ. of Wisconsin when visiting for six months 1964 at the Mathematics Research Center and the Social Systems Research Institute. I feel greatly indebted to numerous friends and colleagues for valuable discussions and comments. For specific amendments of the first draft of Part I, I am indebted to Professor T. H. Hildebrandt, Univ. of Michigan; Professor J. Karamata, Univ. of Geneva and Univ. of Wisconsin; and Prof. L. Rall, Univ. of Wisconsin.

1. We consider a linear vector space H^n with elements

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \{x_1, \ldots, x_n\}; \quad x \in H,$$

where the components x_i belong to a Hilbert space H with real-valued inner products (x_i, x_k). We define spaces H and H^n as metric; thus

$$d(x_i, x_k) = \| x_i - x_k \| = (x_i - x_k, x_i - x_k)^{\frac{1}{2}} \tag{1}$$

is the distance between two elements x_i, x_k in H, and we take

$$D(x, y) = \max (\| x_1 - y_1 \|, \ldots, \| x_n - y_n \|) \tag{2}$$

to be the distance between two vectors $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ in H^n. As is well known, H and H^n are complete in the metrics (1) and (2); [3, 4].

2. Let

$$y = \{y_1, \ldots, y_n\}, \quad n \geqslant 2 \tag{3}$$

be a fixed vector in H^n; let

$$z = \{z_1, \ldots, z_m\}, \quad m \geqslant 1 \tag{4}$$

be a fixed vector in H^m; let $H(z)$ be the subspace of H spanned by the components $z_1, \ldots, z_m$; and let $H^n(z)$ be the subspace of H^n spanned by vectors u the components of which belong to $H(z)$,

$$u = \{u_1, \ldots, u_n\}; \quad u_i \in H(z), \quad u \in H^n(z).$$

Given a set of entries i, k, to be called *positions*, say

$$\left.\begin{array}{lll} i=1; & k=p_1(1), & p_1(2), \ldots, p_1(k_1) \\ i=2; & k=p_2(1), & p_2(2), \ldots, p_2(k_2) \\ \cdots & \cdots & \cdots \\ i=a; & k=p_a(1), & p_a(2), \ldots, p_a(k_a) \end{array}\right\} \tag{5}$$

the matrix

$$\Pi = [\pi_{ik}], \quad i = 1, \ldots, a; \quad k = 1, \ldots, b$$

is called a *position matrix* if

$$\pi_{ik} = 0 \quad \text{for} \quad k \neq p_i(1), \ldots, p_i(k_i); \quad i = 1, \ldots, a$$

whereas π_{ik} may take nonzero values in the positions (5).

We shall represent y in the form

$$y = \mathrm{B}u + \Gamma z + \varepsilon, \quad u \in H^n(z) \tag{6}$$

subject to the following assumptions **2 a–e**. We see that B and Γ are position matrices, with positions specified by (8 b–c) and (9 b), whereas (10)–(11) specifies the numerical values of the elements β_{ip}, λ_{iq} in these positions.

2 a. The vector u is to be specified later, and z is the fixed vector (4).

2 b. The representation of each component y_i involves a prescribed selection of components u_p and z_q, as follows,

$$y_i = \sum_{j=1}^{n} \beta_{ij} u_j + \sum_{k=1}^{m} \gamma_{ik} z_k = \varepsilon_i; \quad i = 1, \ldots, n, \tag{7}$$

where
$$\left.\begin{array}{l} \beta_{ij} = 0 \quad \text{for} \quad j \neq p, \\ p = p_i(1), p_i(2), \ldots, p_i(j_i); \quad j_i \leqslant n-1, \\ p \neq i; \quad \text{hence} \quad \beta_{ii} = 0; \end{array}\right\} \tag{8 a–c}$$

and
$$\left.\begin{array}{l} \gamma_{ik} = 0 \quad \text{for} \quad k \neq q, \\ q = q_i(1), q_i(2), \ldots, q_i(k_i); \quad k_i \leqslant m. \end{array}\right\} \tag{9 a–b}$$

2 c. Let $H_i = H_i(u_p, z_q)$ be defined as the subspace of H spanned by the elements u_p, z_q specified in (8 b) and (9 b). Since $u \in H^n(z)$,

$$H_i \subset H(z) \quad i = 1, \ldots, n.$$

Writing
$$y^* = \{y_1^*, \ldots, y_n^*\} = \mathrm{B}u + \Gamma z \tag{10}$$

we define each component y_i^* as the projection of y_i on H_i. Hence

$$y^* \in H^n(z)$$

with
$$y_i = y_i^* + \varepsilon_i; \quad y_i^* \in H_i, \quad \varepsilon_i \perp H_i \quad i = 1, \ldots, n. \tag{11}$$

2 d. The matrix
$$I - \mathrm{B}$$

is assumed to be nonsingular.

2 e. Finally, we assume

$$k_i \geqslant 1 \quad \text{for at least one } i \quad (i = 1, \ldots, n). \tag{12}$$

3. Let y and z remain fixed. Taking

$$u \in H^n(z)$$

to be an arbitrary vector in $H^n(z)$, the projection (10) defines a mapping of $H^n(z)$ into itself, say Au, where A is the mapping operator defined by

$$Au = \mathrm{B}u + \Gamma z. \tag{13}$$

Lemma. *For the transformation (13) to be a contraction mapping,*

$$D(Au, Av) \leqslant \alpha D(u, v), \quad \alpha < 1 \tag{14}$$

it is sufficient that no component y_i belongs to H_i.

Proof. We consider first the special case when representation (6) for each i involves just one element u_p and one element z_q. The contraction principle (14) then is equivalent to the inequality

$$\max_{i=1}^{n} \| \beta_{i1}^{(u)} u_p + \gamma_{i1}^{(u)} z_q - \beta_{i1}^{(v)} v_p - \gamma_{i1}^{(v)} z_q \|^2 \leqslant \alpha^2 \max_{i=1}^{n} \| u_i - v_i \|^2; \quad \alpha^2 < 1, \tag{15}$$

where the superscripts indicate that the projection (10) gives coefficients that depend upon u_p and v_p, respectively. Accordingly, we shall show

$$Q^2 = N/D \leqslant \alpha^2, \quad \alpha^2 < 1, \tag{16}$$

where, with some change of notation for easy writing, the numerator is given by

$$N = \max_{i=1}^{n} \| \beta_{iu} u_p + \gamma_{iu} z_q - \beta_{iv} v_p - \gamma_{iv} z_q \|^2 \tag{17}$$

and the denominator
$$D = \max_{i=1}^{n} \| u_i - v_i \|^2. \tag{18}$$

We adopt two simplifying devices which involve no loss of generality:

(i) We take u_p and v_p to belong to the orthogonal complement of z_q in $H(z)$,

$$u_p, v_p \subset H(z) \ominus z_q; \quad i = 1, \ldots, n. \tag{19}$$

To see that this involves no restriction, let u_p' and v_p' be arbitrary in $H^n(z)$. Then two elements u_p, v_p that satisfy (19) are uniquely determined by the decompositions

$$u_p' = u_p + \lambda_1 z_q, \quad v_p' = v_p + \lambda_2 z_q,$$

where
$$u_p, v_p \perp z_q.$$

Hence $Au_p = Au_p'$ and $Av_p = Av_p'$. This gives, in obvious notation,

$$Q' = \frac{D(Au_p', Av_p')}{D(u_p', v_p')} = \frac{D(Au_p, Av_p)}{\| u_p - v_p + (\lambda_1 - \lambda_2) z_q \|} = \frac{D(Au_p, Av_p)}{D(u_p, v_p) + \| (\lambda_1 - \lambda_2) z_q \|}.$$

Writing $Q'(0)$ for the value taken by Q' when $\lambda_1 = \lambda_2 = 0$, we infer

$$Q' \leqslant Q'(0) = Q \tag{20}$$

showing that it will suffice to consider elements u_p, v_p that satisfy (19).

(ii) We normalize the components y_i so as to have unit norms,

$$\| y_i \| = 1; \quad i = 1, \ldots, n \tag{21}$$

and assume

$$\| u_i \| = \| v_i \| = \| z_k \| = 1; \quad \begin{cases} i = 1, \ldots, n \\ k = 1, \ldots, m. \end{cases} \tag{22}$$

Thus prepared, let for z given by (4) and for arbitrary $u, v \in H^n(z)$

$$y_i = \beta_u u_p + \beta_v v_p + \gamma z_p + \delta_i; \quad \delta_i \perp u_p, v_p, z_q \tag{23}$$

be the projection of y_i on the subspace spanned by u_p, v_p, z_q. The general theory of orthogonal projection gives

$$\begin{vmatrix} 1 & (y_i, u_p) & (y_i, v_p) & (y_i, z_q) \\ (y_i, u_p) & 1 & (u_p, v_p) & 0 \\ (y_i, v_p) & (u_p, v_p) & 1 & 0 \\ (y_i, z_q) & 0 & 0 & 1 \end{vmatrix} = \| \delta_i \|^2 \cdot \begin{vmatrix} 1 & (u_p, v_p) & 0 \\ (u_p, v_p) & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \geqslant 0. \tag{24}$$

Hence the following relation, which is pivotal in the proof,

$$\frac{(y_i, u_p)^2 + (y_i, v_p)^2 - 2 (y_i, u_p)(y_i, v_p)(u_p, v_p)}{1 - (u_p, v_p)^2} = 1 - (y_i, z_q)^2 - \| \delta_i \|^2 \geqslant 0. \tag{25}$$

Our interest will focus on nearly coincident vectors u, v; that is, vectors with $(u_i, v_i) \cong 1$. We write for brevity

$$\Delta_i = 1 - (u_i, v_i), \quad i = 1, \ldots, n,$$

Hence

$$\Delta_i \geqslant 0, \quad i = 1, \ldots, n.$$

For the denominator in (25) we obtain

$$1 - (u_p, v_p)^2 = 2 \Delta_p \left[1 - \frac{\Delta_p}{2} \right] \leqslant 2 \Delta_p.$$

The squared norm that appears in the denominator (18) is

$$\| u_i - v_i \|^2 = (u_i - v_i, u_i - v_i) = 2 [1 - (u_i, v_i)] = 2 \Delta_i. \tag{26}$$

Next some preparatory reduction of the numerator (17). The general formulas for orthogonal projection in Hilbert space give the coefficient vector

$$\{\beta_{iu}, \gamma_{iu}\} = [\{u_p, z_q\} [u_p, z_q]]^{-1} \{(y_i, u_p), (y_i, z_q)\} \tag{27}$$

and the projection component

$$(Au)_i = \beta_{iu} u_p + \gamma_{iu} z_q = [u_p, z_q] \{\beta_{iu}, \gamma_{iu}\}. \tag{28}$$

As an example of the reductions to follow we note

$$((Au)_i, (Av)_i) = [(y_i, u_p), (y_i, z_q)] [\{u_p, z_q\} [u_p, z_q]]^{-1}$$
$$\{u_p, z_q\} [v_p, z_q] [\{v_p, z_q\} [v_p, z_q]]^{-1} \{(y_i, v_p), (y_i, z_q)\}$$
$$= [(y_i, u_p), (y_i z_q)] \begin{bmatrix} (u_p, v_p) & 0 \\ 0 & 1 \end{bmatrix} \{(y_i, v_p), (y_i, z_q)\}$$
$$= (y_i, u_p) (y_i, v_p) (u_p, v_p) + (y_i, z_q)^2. \tag{29}$$

Forming the squared norm that appears in the numerator (17), we obtain

$$\| (Au - Av)_i \|^2 = ((Au)_i, (Au)_i) - 2 ((Au)_i, (Av)_i) + ((Au)_i, (Av)_i)$$
$$= (y_i, u_p)^2 + (y_i, v_p)^2 - 2 (y_i, u_p) (y_i, v_p) (u_p, v_p). \tag{30}$$

Making use of (25),

$$\| (Au - Av)_i \|^2 = (1 - (u_p, v_p)^2) (1 - y_i, z_q)^2 - \| \delta_i \|^2)$$
$$= 2 \Delta_p \left(1 - \frac{\Delta_p}{2}\right) (1 - (y_i, z_q)^2 - \| \delta_i \|^2). \tag{31}$$

Turning now to Q^2 as defined by (16)–(18), we obtain by (26) and (31)

$$Q^2 = \frac{\max_{i=1}^n \Delta_p (1 - \Delta_p/2) (1 - (y_i, z_q)^2 - \| \delta_i \|^2)}{\max (\Delta_1, \ldots, \Delta_n)}$$
$$\leqslant \frac{(\max_{i=1}^n \Delta_p) \max_{i=1}^n (1 - (y_i, z_q)^2 - \| \delta_i \|^2)}{\max (\Delta_1, \ldots, \Delta_n)}.$$

Hence, since all Δ_p are contained among $\Delta_1, \ldots, \Delta_n$,

$$Q^2 \leqslant \max_{i=1}^n (1 - (y_i, z_q)^2 - \| \delta_i \|^2). \tag{32}$$

On the assumptions of the lemma y_i does not belong to H_i. Hence

$$\| \delta_i \|^2 > 0, \quad \text{say} \quad \| \delta_i \|^2 \geqslant c > 0,$$

which establishes (14) with

$$\alpha = \sqrt{1 - c} < 1. \tag{33}$$

The lemma is proved in the special case of one u_p and one z_q.

Proceeding to the general case, the lemma will be established if we can again show (15), this time interpreting the terms as vector products, as follows,

$$\left.\begin{aligned} \beta_{i1}^{(u)} u_p &= \textstyle\sum_{j=1}^{j_i} \beta_{ij}^{(u)} u_p; \quad & \beta_{i1}^{(v)} v_p &= \textstyle\sum_{j=1}^{j_i} \beta_{ij}^{(v)} v_p; \\ \gamma_{i1}^{(u)} z_q &= \textstyle\sum_{k=1}^{k_i} \gamma_{ik}^{(u)} z_q; \quad & \gamma_{i1}^{(v)} z_q &= \textstyle\sum_{k=1}^{k_i} \gamma_{ik}^{(v)} z_q. \end{aligned}\right\} \tag{34}$$

Without impairing the generality, we adopt the following normalizations and assumptions: (i) The two vectors u_p, v_p are of dimensionality j_i. (ii) We impose (19), (21), (22 a) and (22 c), all of which allow straightforward extension to the general case (34). (iii) Making use of a standard device in transformation theory, [1], we assume in generalization of (22) that u_p and v_p have been subjected to two linear transformations that make

$$\|u_{p(a)}\| = 1, \ \max \|v_{p(a)}\| \leqslant 1; \quad a = 1, \ldots, j_i \tag{35 a–b}$$

$$(u_{p(a)}, u_{p(b)}) = (v_{p(a)}, v_{p(b)}) = 0; \quad a \neq b; \quad a, b = 1, \ldots, j_i \tag{36 a–b}$$

$$(u_{p(a)}, v_{p(b)}) = f_a \cdot \delta_{ab} \qquad a, b = 1, \ldots, j_i \tag{37}$$

where δ is Kronecker's delta, and $f_1, \ldots, f_{j_i}$ are scalar factors.

On the appropriate vector interpretation, (23)–(25) extend to the general case. In (24) the determinants are to be interpreted as partitioned. The inner products either become vectors, for example

$$\{(y_i, u_p)\}, \quad p = p_i(1), \ldots, p_i(j_i)$$

instead of (y_i, u_p); or matrices, for example the $j_i \times j_i$ diagonal matrix

$$[(u_{p(a)}, v_{p(b)})], \quad p = p_i(1), \ldots, p_i(j_i)$$

instead of (u_p, v_p). The unit in the middle of the denominator matrix becomes the diagonal matrix $[(u_{p(a)}, u_{p(b)})]$, and similarly in the numerator matrix. On this interpretation, relation (24) extends to the general case, and making use of (37) we obtain the following generalization of (25),

$$\sum_p \frac{(y_i, u_p)^2 + (y_i, v_p)^2 - 2(y_i, u_p)(y_i, v_p)(u_p, v_p)}{(v_p, v_p) - (u_p, v_p)^2} = \sum_p \frac{N_p}{D_p} = 1 - \sum_q (y_i, z_q)^2 - \|\delta_i\|^2. \tag{38}$$

This relation gives, making use of (35 b) in the third inference,

$$\begin{aligned} \sum_p ((y_i, u_p)^2 + (y_i, v_p)^2 - 2(y_i, u_p)(y_i, v_p)(u_p, v_p)) &= \sum_p \frac{N_p}{D_p} D_p \\ \leqslant (\max D_p) \sum_p \frac{N_p}{D_p} &\leqslant (\max[1 - (u_p, v_p)^2]) \sum_p \frac{N_p}{D_p} \\ = 2\left(\max_p \Delta_p \left(1 - \frac{\Delta_p}{2}\right)\right) &(1 - \sum_q (y_i, z_q)^2 - \|\delta_i\|^2) \\ \leqslant 2(1 - \sum_q (y_i, z_q)^2 - \|\delta_i\|^2) &\max_p \Delta_p. \end{aligned} \tag{39}$$

Formula (26) remains the same. In (27)–(29) the vector interpretation calls for no formal change. Instead of (29) we obtain in the general case

$$((Au)_i,\,(Av)_i) = \sum_{a=1}^{k_i}(y_i,\,z_{q(a)})^2 + \sum_{a=1}^{j_i}\sum_{b=1}^{j_i}(y_i,\,u_{p(a)})\,(y_i,\,v_{p(b)})\,(u_{p(a)},\,v_{p(b)})$$

and (30) generalizes as follows,

$$\|(Au - Av)_i\|^2 = \sum_a (y_i,\,u_{p(a)})^2 + \sum_a (y_i,\,v_{p(a)})^2 - 2\sum_a\sum_b (y_i,\,u_{p(a)})\,(y_i,\,v_{p(b)})\,(u_{p(a)},\,v_{p(b)}).$$

In virtue of (37) this gives

$$\|(Au - Av)_i\|^2 = \sum_p [(y_i,\,u_p)^2 + (y_i,\,v_p)^2 - 2\,(y_i,\,u_p)\,(y_i,\,v_p)\,(u_p,\,v_p)]. \tag{40}$$

Forming now Q^2, making use of (26) and (40), we obtain

$$Q^2 = \frac{\max_{i=1}^n \sum_p [(y_i,u_p)^2 + (y_i,v_p)^2 - 2\,(y_i,u_p)\,(y_i,v_p)\,(u_p,v_p)]}{2\max(\Delta_1,\,\ldots,\,\Delta_n)}.$$

By (39), the numerator is bounded upwards by the maximum of $2\,\Delta_p$ multiplied by the maximum of the last member in (38). This conclusion links up with the argument in (32), and so the rest of the proof is a straightforward extension of (33). We obtain relation (14), specified as in (33) with $0 < c \leqslant \|\delta_i\|^2$. The lemma is proved.

4. Theorem. *Let y and z be given vectors (3)–(4), and let the relation (6) be subject to the specifications 2 a–d, where B, Γ are position matrices with unknown positional coefficients β_{ip}, γ_{iq}. Then there exists one and only one representation*

$$y = \mathrm{B}\,y^* + \Gamma\,z + \varepsilon; \quad y^* \subset H^n(z), \quad \varepsilon_i \perp y_p^*, z_q \tag{41}$$

with

$$y^* = \mathrm{B}\,y^* + \Gamma\,z. \tag{42}$$

The vector y^* is uniquely determined by the following iterative procedure, where the initial vector $y^{(0)}$ can be chosen arbitrarily in $H^n(z)$,

$$\begin{cases} y = \mathrm{B}^{(1)}y^{(0)} + \Gamma^{(1)}z + \varepsilon^{(1)}, & \varepsilon_i^{(1)} \perp y_p^{(0)}, z_q, \\ y^{(1)} = \mathrm{B}^{(1)}y^{(0)} + \Gamma^{(1)}z, & i = 1, \ldots, n; \end{cases}$$

$$\begin{cases} y = \mathrm{B}^{(2)}y^{(1)} + \Gamma^{(2)}z + \varepsilon^{(2)}, & \varepsilon_i^{(2)} \perp y_p^{(1)}, z_q, \\ y^{(2)} = \mathrm{B}^{(2)}y^{(1)} + \Gamma^{(2)}z, & i = 1, \ldots, n; \end{cases}$$

$$\cdots\cdots\cdots$$

$$\begin{cases} y = \mathrm{B}^{(s)}y^{(s-1)} + \Gamma^{(s)}z + \varepsilon^{(s)}, & \varepsilon_i^{(s)} \perp y_p^{(s-1)}, z_q, \\ y^{(s)} = \mathrm{B}^{(s)}y^{(s-1)} + \Gamma^{(s)}z, & i, \ldots, n; \end{cases} \tag{43}$$

$$\cdots\cdots\cdots$$

$$y^* = \lim_{s\to\infty} y^{(s)}. \tag{44}$$

If we disregard the exceptional case when the elements

$$y_p^*, z_q \text{ with } p = p_i(1), \ldots, p_i(j_i); \quad q = q_i(1), \ldots, q_i(k_i) \tag{45}$$

are linearly interrelated for some fixed $i\,(i = 1, \ldots, n)$, *the matrices are* B, Γ *are uniquely determined by*

$$\mathrm{B} = \lim_{s\to\infty} \mathrm{B}^{(s)}; \quad \Gamma = \lim_{s\to\infty} \Gamma^{(s)}. \tag{46 a–b}$$

Proof. Up to (45) the theorem follows from our lemma as an immediate corollary of the general fix-point theorem of contraction mapping: The equation

$$y^* = Ay^* \tag{47}$$

has one and only one solution in $H^n(z)$, and the solution is given by the iterative procedure (43)–(44). For the sake of completeness we recapitulate the proof, following [3].

Let $y^{(0)}$ be an arbitrary vector in $H^n(z)$. Set $y^{(1)} = Ay^{(0)}$; $y^{(2)} = Ay^{(1)} = A^2 y^{(0)}$, and in general let $y^{(s)} = Ay^{(s-1)} = A^s y^{(0)}$. We shall show that the sequence $y^{(s)}$ satisfies the Cauchy criterion. In fact, for any $t > s$

$$\begin{aligned} D(y^{(s)}, y^{(t)}) &= D(A^s y^{(0)}, A^t y^{(0)}) \leqslant \alpha^s D(y^{(0)}, y^{(t-s)}) \\ &\leqslant \alpha^s [D(y^{(0)}, y^{(1)}) + D(y^{(1)}, y^{(2)}) + \ldots + D(y^{(t-s-1)}, y^{(t-s)})] \\ &\leqslant \alpha^s D(y^{(0)}, y^{(1)})(1 + \alpha + \alpha^2 + \ldots + \alpha^{t-s-1}) \leqslant \alpha^s D(y^{(0)}, y^{(1)})/(1-\alpha). \end{aligned} \tag{48}$$

Since $\alpha < 1$ this quantity is arbitrarily small for sufficiently large s. Hence the sequence $y^{(s)}$ is fundamental, and since $H^n(z)$ is complete it follows that $\lim_{s\to\infty} y^{(s)}$ exists, We set $y^* = \lim_{s\to\infty} y^{(s)}$. Then by virtue of the continuity of the mapping A we infer $Ay^* = A \lim_{s\to\infty} y^{(s)} = \lim_{s\to\infty} Ay^{(s)} = \lim_{s\to\infty} y^{(s+1)} = y^*$.

Thus, the iteration (43) converges, and thereby the existence of a fix-point y^* is proved. We shall now prove its uniqueness. If $Ay^* = y^*$, $Ay^{**} = y^{**}$, then $D(y^*, y^{**}) \leqslant \alpha D(y^*, y^{**})$, where $\alpha < 1$; this implies $D(y^*, y^{**}) = 0$; that is, $y^* = y^{**}$.

The general fix-point theorem thus ensures the limiting relation (44). It remains to prove (46 a–b). These relations readily follow if we disregard the case when the variables (45) are interrelated. First, the inner products

$$(y^*, z), (y^*, y^*), (y^*, y) \tag{49}$$

will be uniquely determined by the limiting element y^*. Second, writing

$$y = \mathrm{B}^* y^* + \Gamma^* z + \varepsilon^*, \quad \varepsilon^* \perp y^*, z$$

for the orthogonal projection of y on the linear manifold spanned by y^* and z, the ensuing normal equations will give us B^* and Γ^* in terms of the inner products (y_i, z_k), (z_j, z_k) and (49), and in the case under consideration the coefficients $\beta_{ip}^*, \gamma_{iq}^*$ will for every i be uniquely determined. Finally, the resulting B^* and Γ^* will be the limits of $\mathrm{B}^{(s)}$ and $\Gamma^{(s)}$, giving

$$\beta = \beta^*; \quad \Gamma = \Gamma^*$$

since the elements of matrices β and Γ are continuous functions of y^*. The theorem is proved.

Remark. Once the limiting vector y^* has been obtained by (44), Gram's criterion can be used to find out for any fixed i whether or not the elements (45) are linearly interrelated. Forming the $(j_i + k_i) \times (j_i + k_i)$ matrix or inner products of the elements (45), the criterion for linear independence is that the matrix should be of rank $j_i + k_i$.

As to criteria that can be applied prior to the iterative procedure, we note that the following obvious and very simple conditions are necessary for linear independence.

$$d_i = k_i; \quad d \geqslant j_i + k_i \quad i = 1, \ldots, n, \tag{50}$$

where d_i is the dimensionality of vector $z_q\,[q = q_i(1), \ldots, q_i(k_i)]$, and d the dimensionality of vector z.

5. The iterative procedure (43) can be exploited for further information on the representation (41)–(42). We note:

(i) According to assumption **2d** the matrix $(I - \beta)$ is nonsingular. Hence our theorem gives as an immediate corollary

$$y^* = \Omega z \ \text{ with } \ \Omega = \{\omega_{ik}\} = (I - \beta)^{-1}\,\Gamma. \tag{51}$$

Further we note the following formula, which follows from (42) by iterated substitutions of the right-hand member into itself,

$$y^* = \Gamma z + \beta\,\Gamma z + \ldots + \beta^{s-1}\,\Gamma z + \beta^s y^*. \tag{52}$$

It is instructive to compare with the following relation, which follows by substitution from the iterative formula (43),

$$y^{(s)} = \Gamma^{(s)} z + \beta^{(s)}\,\Gamma^{(s-1)} z + \ldots + \beta^{(s)}\,\beta^{(s-1)} \ldots \beta^{(2)}\,\Gamma^{(1)} z + \beta^{(s)}\,\beta^{(s-1)} \ldots \beta^{(1)}\,y^{(0)}. \tag{53}$$

In case all eigenvalues of $(I - \beta)$ lie inside the periphery of the unit circle, (51) gives

$$y^* = \lim_{s\to\infty} (\Gamma z + \beta\,\Gamma z + \ldots + \beta^s\,\Gamma z) = \Gamma z + \beta\,\Gamma z + \ldots + \beta^s\,\Gamma z + \ldots. \tag{54}$$

The limiting relation $y^{(s)} \to y^*$ is then a matter of term by term convergence in (53) and (54).

The following inner products involving y_i^* are readily obtained from (51),

$$\left.\begin{aligned} (y_i^*, z_k) &= \textstyle\sum_{a=1}^{m} \quad \omega_{ia}\,(z_a, z_k), \\ (y_i^*, y_k) &= \textstyle\sum_{a=1}^{m} \quad \omega_{ia}\,(z_a, y_k), \\ (y_i^*, y_k^*) &= \textstyle\sum\sum_{a,b=1}^{m} \quad \omega_{ia}\,\omega_{ab}\,(z_b, z_k), \end{aligned}\right\} \tag{55 a–c}$$

where $i, \ k = 1, \ldots, n$.

As to the exceptional case when assumption **2 d** is not fulfilled, we note that if $(I - \mathrm{B})$ is singular, the representation (41)–(42) will not be unique. In fact, let in this case λ be an eigenvalue of $(I - \mathrm{B})$ and $y^{(\lambda)}$ the corresponding eigenvector, giving

$$(I - \mathrm{B})\, y^{(\lambda)} = \lambda\, y^{(\lambda)} = \| y^{(\lambda)} \| \, \Gamma z,$$

where $\| y^{(\lambda)} \|$ is an arbitrary positive number. Hence y allows a plurality of representations of type

$$y = \Omega^{(\lambda)} z \quad \text{with} \quad \Omega^{(\lambda)} = \lambda^{-1} \| y^{(\lambda)} \| \, \Gamma z. \tag{56}$$

Further we note that (56) may be interpreted as a plurality of representations of type (41), say

$$y = \mathrm{B}^* y^* + \Gamma^* z + \varepsilon$$

with

$$\mathrm{B}^* = 0; \quad \Gamma^* = \lambda^{-1} \| y^{(\lambda)} \| \, \Gamma z.$$

(ii) Having obtained the limiting element y^*, the orthogonal complement ε will be given by

$$\varepsilon = y - y^*. \tag{57}$$

As to inner products that involve ε_i, we know from (11)

$$\left.\begin{aligned} (\varepsilon_i, z_q) &= 0; \quad q = q_i(1), \ldots, q_i(k_i), \\ (\varepsilon_i, y_p^*) &= 0; \quad p = p_i(1), \ldots, p_i(j_i), \end{aligned}\right\} \tag{58 a–b}$$

for $i = 1, \ldots, n$. Hence

$$(\varepsilon_i, y_i^*) = 0; \quad i = 1, \ldots, n \tag{59}$$

since we know from (42) that y_i^* is linear in the elements y_p^*, z_q that appear in (58). According to (51), the inner products (59) are the same as

$$(\varepsilon_i, \textstyle\sum_{i=1}^{m} \omega_{ia} z_a) = 0; \quad i = 1, \ldots, n.$$

By (57),

$$\left.\begin{aligned} (\varepsilon_i, z_k) &= (y_i, z_k) - (y_i^*, z_k), \\ (\varepsilon_i, y_k) &= (y_i, y_k) - (y_i^*, y_k), \\ (\varepsilon_i, y_k^*) &= (y_i, y_k^*) - (y_i^*, y_k^*), \\ (\varepsilon_i, \varepsilon_k) &= (y_i, y_k) - (y_i, y_k^*), - (y_i^*, y_k) + (y_i^*, y_k^*), \end{aligned}\right\} \tag{60 a–d}$$

where $i, k = 1, \ldots, n$. According to (59), formula (60 c) simplifies for $i = k$. Formula (60 d), too, becomes simpler for $i = k$:

$$\| \varepsilon_i \|^2 = \| y_i \|^2 - \| y_i^* \|^2; \quad i = 1, \ldots, n. \tag{61}$$

REFERENCES

1. ANDERSON, T. W., *An introduction to multivariate statistical analysis.* New York, Wiley and Sons, 1958.
2. FADDE'EVA, V. N., *Computational methods of linear algebra.* New York, Dover Publ., 1959.
3. KOLMOGOROV, A. N., and FOMIN, S. V., *Elements of the theory of functions and functional analysis*, 1–2. Rochester, N. Y., Graylock, 1957, 1958.
4. THRALL, R. M., and TORNHEIM, L., *Vector spaces and matrices*, New York, Wiley and Sons, 1957.
5. WOLD, H., "*Ends and means in econometric model building. Basic considerations reviewed.*" In *Probability and statistics. The Harald Cramér volume*, ed. U. Grenander, pp. 355–434. Stockholm, Almqvist and Wiksell, 1959; New York, Wiley and Sons, 1960.
6. ——, "Unbiased predictors." *Proceedings of the Fourth Berkeley Symposium of Mathematical Statistics and Probability.* Vol. 1, pp. 719–761. Berkeley, Univ. of California Press, 1961.
7. ——, "Forecasting by the chain principle." In *Econometric model building. Essays on the causal chain approach*, ed. H. Wold, pp. 1–34. Amsterdam, North-Holland Publication Company, 1964.

Tryckt den 27 juli 1965

Uppsala 1965. Almqvist & Wiksells Boktryckeri AB