A fix-point theorem with econometric background

Part II. Illustrations. Further developments

By HERMAN O. A. WOLD

6. An illustration. We take y and z to be vectors with $n = 2$ and $m = 4$ components, and specify the representation (41)–(42) as follows, giving it in the developed form (7),

$$\left. \begin{aligned} y_1 &= \beta_{12} y_2^* + \gamma_{11} z_1 + \gamma_{12} z_2 + \varepsilon_1 \\ y_2 &= \beta_{21} y_1^* + \gamma_{23} z_3 + \gamma_{24} z_4 + \varepsilon_2 \end{aligned} \right\} \tag{62 a–b}$$

with

$$\varepsilon_1 \perp y_2^*, z_1, z_2; \quad \varepsilon_2 \perp y_1^*, z_3, z_4$$

and

$$\left. \begin{aligned} y_1^* &= \beta_{12} y_2^* + \gamma_{11} z_1 + \gamma_{12} z_2 \\ y_2^* &= \beta_{21} y_1^* + \gamma_{23} z_3 + \gamma_{24} z_4 \end{aligned} \right\} \tag{63 a–b}$$

We shall consider four interpretations of our fix-point theorem as applied to this system of relations. The four versions of the vectors y, z are:

A. The basic specification as vectors with components in Hilbert space H;

B. Vectors in Euclidean space R_h, with inner products interpreted as angular cosines;

C. Vectors of random variates; that is, multivariate probability distributions, with inner products interpreted as theoretical covariances;

D. Multivariate observations, in the form of time series or cross section data, with inner products interpreted as observed covariances.

7. Version A. Given the vectors

$$y = \{y_1, y_2\}, \quad z = \{z_1, z_2, z_3, z_4\} \tag{64 a–b}$$

the problem is to determine the vectors

$$y^* = \{y_1^*, y_2^*\} \tag{65}$$

and the coefficients

$$\beta_{12}, \gamma_{11}, \gamma_{12}; \quad \beta_{21}, \gamma_{23}, \gamma_{24} \tag{66}$$

so as to satisfy the following relationships: First, writing

$$Ay_1 = \beta_{12} y_2^* + \gamma_{11} z_1 + \gamma_{12} z_2 \left.\right\}$$
$$Ay_2 = \beta_{21} y_1^* + \gamma_{23} z_3 + \gamma_{24} z_4 \left.\right\}$$
(67 a–b)

Ay_1 is the projection of y_1 on the space H_1 spanned by y_2^*, z_1, z_2, and Ay_2 is the projection of y_2 on the space H_2 spanned by y_1^*, z_3, z_4; second, Ay_1 and Ay_2 coincide with y_1^* and y_2^*, respectively, giving

$$Ay_1 = y_1^* = \beta_{12} y_2^* + \gamma_{11} z_1 + \gamma_{12} z_2 \left.\right\}$$
$$Ay_2 = y_2^* = \beta_{21} y_1^* + \gamma_{23} z_3 + \gamma_{24} z_4 \left.\right\}$$
(68 a–b)

We specify those inner products of the components y_i, z_k that are required for solving the problem by the iterative procedure (43), namely the two squared norms

$$(y_1, y_1) = 111.75; \quad (y_2, y_2) = 101 \tag{69}$$

and in Table 1 the eight mixed products (y_i, z_k) and the ten different products (z_i, z_k).

<div align="center">*Table 1*</div>

	$k = 1$	$k = 2$	$k = 3$	$k = 4$
(y_1, z_k)	35	17.5	32	13
(y_2, z_k)	19	22	38	12
(z_1, z_k)	22	2	5	3
(z_2, z_k)	2	6	11	1
(z_3, z_k)	5	11	22	2
(z_4, z_k)	3	1	2	6

We note that the inner product (y_1, y_2) is redundant for the procedure.

As is readily verified, the selective indices (8 b)–(9 b) and the vectors y, z thus specified satisfy the conditions (50) of our procedure. In the present case,

$$j_1 = j_2 = 1, \quad k_1 = k_2 = 2, \quad d = 4 \tag{70}$$

The iterative procedure converges, and we quote the following results:

(a) The coefficients β, γ as given in the limit (46) by the iterative procedure,

$$\beta_{12} = 0.5; \quad \gamma_{11} = \gamma_{12} = 1; \quad \beta_{21} = 0.4; \quad \gamma_{23} = \gamma_{24} = 1 \tag{71}$$

(b) We see from (71) that the matrix $(I - \beta)$ is nonsingular. Hence (51) gives the vector y^* in the following form, with coefficients obtained in terms of the limiting values (71),

$$y_1^* = 1.25 (z_1 + z_2) + 0.625 (z_3 + z_4) \left.\right\}$$
$$y_2^* = 0.5 (z_1 + z_2) + 1.25 (z_3 + z_4) \left.\right\}$$
(72 a–b)

We note that the eigenvalues of $(I - \beta)$ lie inside the unit circle; they are

$\pm (0.2)^{\frac{1}{2}}$. Hence the expansion (52) is valid, and gives the same coefficients ω_{ik} as (72 a–b).

(c) Finally, the inner products that involve y_i^* or ε_i. Calculated in accordance with formulas (55)–(61), the ensuing products are presented in

Table 2

	$k=1$	$k=2$	$k=3$	$k=4$
(y_1^*, z_k)	35	17.5	35	10
(y_2^*, z_k)	22	19	38	12
(y_1^*, y_k)	93.75	82.5	—	—
(y_2^*, y_k)	82.5	83	—	—
(ε_1, z_k)	0	0	-3	3
(ε_2, z_k)	-3	3	0	0
(ε_1, y_k^*)	0	0	(y_1^*, y_1^*)	(y_2^*, y_2^*)	(y_1^*, y_2^*)
(ε_2, y_k^*)	0	0	93.75	83	82.5
(ε_1, y_k)	18	*	$(\varepsilon_1, \varepsilon_1)$	$(\varepsilon_2, \varepsilon_2)$	$(\varepsilon_1, \varepsilon_2)$
(ε_2, y_k)	*	18	18	18	*

Nonexisting entries are barred in Table 2. Big zeros refer to inner products that by (58)–(59) are necessarily equal to zero; small zeros are products that happen to be zero in the present special case. The three stars (*) mark inner products that cannot be evaluated since (y_1, y_2) has not been specified.

8. Version B. In (62)–(63) each component $y_i = y_{it}$ $(i = 1, 2)$ and $z_k = z_{kt}$ $(k = 1, 2, 3, 4)$ is now a vector in Euclidean space R_h $(t = 1, ..., h)$. We choose $h = 6$, and specify the vectors as shown in Table 3 a.

Table 3 a. Version B: Given vectors. *Table 3 b.* Deduced vectors.

t	y_{1t}	y_{2t}	z_{1t}	z_{2t}	z_{3t}	z_{4t}	y_{1t}^*	y_{2t}^*	ε_{1t}	ε_{2t}
1	-5	-2	-4	0	0	0	-5	-2	0	0
2	6.75	6.5	1	1	1	1	3.75	3.5	3	3
3	-1.25	-2.5	0	0	0	-2	-1.25	-2.5	0	0
4	0.75	3.5	1	1	2	0	3.75	3.5	-3	0
5	3.75	0.5	2	0	1	1	3.75	3.5	0	-3
6	-5	-6	0	-2	-4	0	-5	-6	0	0

The problem is to construct two vectors y_1^*, y_2^* in accordance with (62)–(63). To paraphrase, we seek y_1^*, y_2^* such that y_1^* is the orthogonal projection of y_1 on the space spanned by y_2^*, z_1, z_2 and y_2^* is the projection of y_2 on the space spanned by y_1^*, z_3, z_4.

Version B coincides with A, inasmuch as the inner products in (69) and Table 1 are nothing else than the corresponding products of vectors in Table 3 a. Hence the iterative procedure (43)–(44) gives precisely the same results in version B as have been presented for version A; see (71)–(72) and Table 2. In addition, we obtain from (63) and (51) the vectors y_1^*, y_2^*, and from (62) the vectors $\varepsilon_1, \varepsilon_2$; these results of the iterative procedure are shown in Table 3 b.

For version B we have used very simple vectors z_{kt}, ε_{it}. Thanks to the simple specification of the vectors z_{kt}, ε_{it} it can be readily checked that the representation of vectors y_{1t}, y_{2t} in the form (62)–(63) is unique, in accordance with our theorem.

Vectors y_{1t}, y_{2t} give $\qquad (y_1, y_2) = 91.5 \qquad\qquad (73)$

By (60), this gives $\qquad (\varepsilon_1, y_2) = (\varepsilon_2, y_1) = (\varepsilon_1, \varepsilon_2) = 9 \qquad (74)$

In version A the inner product (y_1, y_2) has not yet been specified. Now if we supplement version A by assuming (73), this will again give us (74). The three missing inner products in Table 2 will thereby be specified.

9. Version C. We rewrite (62)–(63) with different notation for the vectors y, y^*, z:

$$\left. \begin{aligned} \eta_1 &= \beta_{12}\eta_2^* + \gamma_{11}\zeta_1 + \gamma_{12}\zeta_2 + \varepsilon_1 \\ \eta_2 &= \beta_{21}\eta_1^* + \gamma_{23}\zeta_3 + \gamma_{24}\zeta_4 + \varepsilon_3 \end{aligned} \right\} \qquad (75)$$

with

$$\left. \begin{aligned} \eta_1^* &= \beta_{12}\eta_2^* + \gamma_{11}\zeta_1 + \gamma_{12}\zeta_2 \\ \eta_2^* &= \beta_{21}\eta_1^* + \gamma_{23}\zeta_3 + \gamma_{24}\zeta_4 \end{aligned} \right\} \qquad (76)$$

We shall interpret (75)–(76) as relations between random variates $\eta_1, \eta_2, \zeta_1, \zeta_2, \zeta_3, \zeta_4$ that are subject to a joint probability distribution such that the second order moments are given by the corresponding inner products; in symbols,

$$(y_i, y_i) = E(\eta_i^2); \quad (y_i, y_k) = E(\eta_i \eta_k); \quad (y_i, z_k) = E(\eta_i \zeta_k); \quad \text{etc.} \qquad (77)$$

We specify the second order moments so as to give the same inner products as in Illustration A, that is, by (69) and Table 1. The problem now is to determine the random variates η_1^*, η_2^* and the parameters (66) so that (75 a) will be the theoretical regression of η_1 upon $\eta_2^*, \zeta_1, \zeta_2$ and (75 b) the theoretical regression of η_2 upon $\eta_1^*, \zeta_3, \zeta_4$ and at the same time, in accordance with (76), the residual-free part of the regression (75 a) will coincide with η_1^* and the residual-free part of the regression (75 b) will coincide with η_2^*.

The iterative procedure (43)–(44) applies, and on the appropriate interpretation in terms of random variates the results are again given by (71)–(72) and Table 2.

In version C, thus far, we have not specified the mixed second order moment $E(\eta_1 \eta_2) = (y_1, y_2)$. If $E(\eta_1 \eta_2)$ is specified by (73) we obtain (74) and thereby the missing items in Table 2.

The theoretical regression (75 a–b) being based on the joint probability distribution of the variates η_i, ζ_k we note that the iterative procedure covers the general case when the regression is defined as linear moment regression; see Ref. 9. Less general definitions will be covered as special cases. With reference to the general representation (41)–(42) dealt with by our theorem, we quote

two definitions of linear regression (i)–(ii) that are special cases of linear moment regression (iii).

(i) *Normal distribution.* The joint probability distribution of the six variates η_i, ζ_k is a multivariate normal distribution. Hence the regressions (75) are strictly linear, and

$$\eta_i^* = E(\eta_i \mid \eta_p^*, \zeta_q) = \sum_{j=1}^{n} \beta_{ij} \eta_p^* + \sum_{k=1}^{m} \gamma_{ik} \zeta_q \tag{78}$$

The distribution of the resulting variates η_i^*, ε_k will likewise be jointly normal.

(ii) *Linear predictors.* More generally than (i), it is assumed that the theoretical regressions (41) are strictly linear in the sense of (78), but the distribution of η_i, ζ_k need not be normal, nor will it follow that the distribution of the resulting η_i^*, ε_k is normal. The residual-free part of the regression (41) is for each i said to be a (linear) *eo ipso predictor* of η_i, or briefly a *predictor* of η_i. For a review of regression analysis from the point of view of predictors, with special regard to multirelation models, see Refs. 5–7.

(iii) *Linear moment regression.* Given the product moments of $\eta_i, \eta_j^*, \zeta_k$ the coefficients β, γ of the moment regression (41) make the variance of ε_i the smallest possible for each i. The regression need not be strictly linear in the sense of (78), nor need the joint distribution of η_i, ζ_k be normal.

10. Version D. In (62)–(63) we now interpret each component y_i, z_k as a time series, say

$$y_i = y_{i1}, \ldots, y_{iT}; \quad z_k = z_{k1}, \ldots, z_{kT} \tag{79}$$

where $i = 1, 2$ and $k = 1, 2, 3, 4$. The problem is to determine the coefficients β, γ and two time series

$$y_1^* = y_{11}^*, \ldots, y_{1T}^*; \quad y_2^* = y_{21}^*, \ldots, y_{2T}^* \tag{80}$$

so that (62 a) gives the least squares regression of y_1 on y_2^*, z_1, z_2 and (62 b) the least squares regression of y_2 on y_1^*, z_3, z_4, at the same time as the residual-free part of regression (62 a) coincides with y_1^*, and the residual-free part of regression (62 b) with y_2^*.

Thus specified, version D is altogether equivalent to version B, taking $h = T$ to be the dimension of the Euclidean space R_h. New aspects of the problem emerge if we adopt the specification of version C, let (62)–(63) be a multirelation stochastic model (75)–(76), and interpret the observed time-series (79) as generated in accordance with the theoretical model (75)–(76). To follow up this interpretation we rewrite (62) with different notation for the coefficients β, γ and the residual ε,

$$\left. \begin{array}{l} y_{1t} = b_{12} y_{2t}^* + g_{11} z_{1t} + g_{12} z_{2t} + e_{1t} \\ y_{2t} = b_{21} y_{1t}^* + g_{23} z_{3t} + g_{24} z_{4t} + e_{2t} \end{array} \right\} \tag{81}$$

and similarly for (63),

Table 4a. Version *D*: Given time-series, obtained by artificial sampling.

t	y_{1t}	y_{2t}	z_{1t}	z_{2t}	z_{3t}	z_{4t}
1	− 2.42	.82	− 2.07	1.01	2.63	− .07
2	− .03	.43	.47	− .33	− 2.14	− 1.11
3	7.56	8.85	− .09	2.05	3.64	1.53
4	2.01	4.46	.06	1.25	2.06	− .06
5	− 4.88	− 4.01	.82	− .61	− .81	− 1.48
6	7.56	6.46	.82	.97	1.53	1.92
7	− 1.93	− 4.69	2.21	− 1.31	− 2.06	− .73
8	− 1.15	− .34	.05	.15	.33	− .59
9	9.01	6.12	4.47	.96	1.76	.55
10	2.26	2.63	.18	.47	.65	.48
11	− 3.04	− 3.55	− 1.06	− .76	− .90	.27
12	− 7.02	− 5.99	− 2.29	− 1.27	− 2.89	− 1.33
13	− 6.80	− 6.70	− 3.86	− 1.87	− 3.81	.59
14	− 3.75	− 1.75	− 3.75	.45	1.02	− .83
15	1.66	1.95	− .08	.70	1.70	.36
16	− 3.36	− 2.28	− 1.34	− .41	− .85	− .36
17	1.71	3.74	.28	.68	1.63	1.65
18	5.16	3.36	2.17	.57	1.37	.76
19	− 4.53	− 1.72	− 4.15	− .25	− 1.46	− .48
20	3.80	− 1.97	3.05	− 1.49	− 2.11	2.00
21	− 1.54	− 3.54	.91	− 1.30	− 1.95	.82
22	8.23	9.55	1.52	2.36	4.14	.58
23	.09	− 5.11	.78	− 1.30	− 2.68	− 1.41
24	1.23	4.16	− .32	1.21	2.13	.14
25	.07	− .93	.26	− .19	.03	.18
26	− 4.35	− 4.26	− 1.56	− .97	− 1.80	− .29
27	− 2.29	− 1.95	.44	− .23	− .55	− 1.34
28	6.73	5.46	3.00	1.60	3.75	.10
29	7.28	7.03	4.96	1.16	2.57	1.22
30	− 4.93	− 2.73	− 2.69	− .50	− 1.07	.01
31	1.06	2.36	.69	.84	1.51	− .65
32	− 8.19	− 10.35	− 2.07	− 2.03	− 3.49	− 1.78
33	− .94	− 2.29	− .74	− .98	− 1.93	.88
34	− 6.27	− 4.91	− 3.67	− .57	− 1.93	− 2.17
35	− .61	− 4.79	1.08	− .71	− .48	− 1.24
36	2.33	− .88	− .27	− .39	− .41	.83
37	− 1.01	3.00	− 2.51	.69	.32	.25
38	7.74	8.25	3.38	1.57	2.54	.64
39	− 1.86	− 3.65	− 1.77	− 1.13	− 1.87	1.05
40	2.34	− .28	.99	.31	1.24	− .51

Table 4b. Expected endogenous variables, and corresponding residuals: Theoretical vs. estimated values.

t	$\overset{*}{\eta}_{1t}$	$\overset{*}{\eta}_{2t}$	$\overset{*}{y}_{1t}$	$\overset{*}{y}_{2t}$	ε_{1t}	ε_{2t}	e_{1t}	e_{2t}
1	.27	2.67	.81	3.60	− 2.69	− 1.86	− 3.23	− 2.78
2	− 1.86	− 3.99	− 2.06	− 4.73	1.83	4.42	2.03	5.16
3	5.69	7.45	5.28	7.63	1.88	1.40	2.28	1.22
4	2.88	3.15	2.69	3.32	− .87	1.30	− 0.68	1.14
5	− 1.16	− 2.74	− 0.79	− 2.87	− 3.72	− 1.26	− 4.09	− 1.14
6	4.40	5.21	4.07	4.72	3.16	1.25	3.49	1.74
7	− .61	− 3.03	− 0.31	− 3.82	− 1.32	− 1.65	− 1.62	− 0.87
8	.08	− .22	.29	− .20	− 1.23	− .12	− 1.44	− 0.14
9	8.23	5.61	7.41	4.72	.78	.51	1.60	1.40
10	1.52	1.74	1.50	1.52	.74	.88	0.76	1.11
11	− 2.67	− 1.70	− 1.94	− 1.75	− .37	− 1.85	− 1.10	− 1.80
12	− 7.09	− 7.06	− 6.27	− 7.11	.07	1.06	− 0.75	1.12
13	− 9.19	− 6.90	− 7.95	− 7.13	2.39	.20	1.15	0.43
14	− 4.00	− 1.40	− 3.19	− .39	.25	− .34	− 0.56	− 1.36
15	2.06	2.89	2.22	3.07	− .41	− .94	− 0.56	− 1.12
16	− 2.94	− 2.38	− 2.41	− 2.37	− .42	.10	− 0.95	0.09
17	3.25	4.58	3.30	4.41	− 1.53	− .84	− 1.59	− 0.67
18	4.75	4.03	4.52	3.58	.40	− .67	0.64	− 0.22
19	− 6.71	− 4.62	− 6.01	− 4.29	2.18	2.90	1.48	2.57
20	1.88	.65	2.09	− .81	1.92	− 2.62	1.71	−1.16
21	− 1.19	− 1.60	− .69	− 2.38	− .35	− 1.94	− 0.85	− 1.16
22	7.79	7.84	7.06	7.96	.44	1.71	1.17	1.59
23	− 3.21	− 5.37	− 2.82	− 5.96	3.30	.27	2.91	0.85
24	2.53	3.28	2.45	3.53	− 1.30	.88	− 1.22	0.63
25	.22	.30	.54	.13	− .14	− 1.23	− 0.47	− 1.06
26	− 4.46	− 3.87	− 3.73	− 3.98	.11	− .39	− 0.62	− 0.28
27	− .91	− 2.25	− .71	− 2.37	− 1.37	.31	− 1.58	0.42
28	8.16	7.12	7.68	7.15	− 1.42	− 1.66	− 0.95	− 1.69
29	10.01	7.79	9.18	6.91	− 2.73	− .76	− 1.90	0.12
30	− 4.66	− 2.92	− 3.89	− 2.74	− .28	.19	− 1.04	0.01
31	2.45	1.85	2.34	1.93	− 1.39	.51	− 1.28	0.43
32	− 8.42	− 8.64	− 7.19	− 8.59	.24	− 1.71	− 1.00	− 1.76
33	− 2.81	− 2.17	− 2.28	− 2.73	1.87	− .11	1.34	0.44
34	− 7.86	− 7.25	− 7.07	− 6.73	1.59	2.34	0.80	1.82
35	− .61	− 1.96	− .11	− 2.02	− .01	− 2.83	− 0.50	− 2.77
36	− .56	.19	− .14	− .08	2.88	− 1.07	2.47	− 0.80
37	− 1.91	− .19	− 1.75	− .02	.90	3.19	0.74	3.02
38	8.17	6.44	7.28	5.85	− .43	1.81	0.46	2.40
39	− 4.14	− 2.47	− 3.30	− 2.77	2.28	− 1.18	1.44	− 0.88
40	2.08	1.56	2.26	1.66	.26	− 1.84	0.08	− 1.94

$$y_{1t}^* = b_{12} y_{2t}^* + g_{11} z_{1t} + g_{12} z_{2t} \left.\right\}$$
$$y_{2t}^* = b_{21} y_{2t}^* + g_{23} z_{3t} + g_{24} z_{4t} \left.\right\}$$

(82)

In this notation, corresponding theoretical and empirical concepts will be represented by Greek and Latin letters respectively.

The time-series (79) used in version D are shown in Table 4 a.[1] The series have been constructed so as to form an artificial sample, with $T = 40$, of a joint probability distribution for the variables $\eta_1, \eta_2, \zeta_1, \zeta_2, \zeta_3, \zeta_4$. The distribution is multivariate normal with zero means,

$$E(\eta_1) = E(\eta_2) = E(\zeta_1) = E(\zeta_2) = E(\zeta_3) = E(\zeta_4) = 0$$

and with variances and covariances specified in accordance with (69), (73) and Table 1.

Applying the iterative estimation procedure (43)–(44) to the model (75) and the data in Table 4 a, we obtain the results reported in Tables 4 b and 5. In Table 5 we have included the theoretical standard errors of the parameter estimates as calculated on the basis of the subsequent large-sample formula (102). The largest absolute deviation between theoretical and estimated parameters occurs for γ_{12}, but we see that the deviation is less than twice the standard error. The largest relative deviation occurs for γ_{23}, where the deviation is 3.3 times the standard error.

Table 5. The fix-point estimation (43)–(46) as applied to model (75) and the data in Table 4 b.

Relation (75 a)			Relation (75 b)		
Parameter		Standard error	Parameter		Standard error
Theoretical	Estimated		Theoretical	Estimated	
$\beta_{12} = .5$	.563	.11	$\beta_{21} = .4$	.269	.09
$\gamma_{11} = 1$	.917	.13	$\gamma_{23} = 1$	1.391	.12
$\gamma_{12} = 1$	.374	.36	$\gamma_{24} = 1$	.891	.18

As to the convergence of the iteration procedure (43), the calculations reported in Table 5 were carried on until all parameters were stable with accuracy .00001. This happened in the ninth iteration, showing that the speed of convergence was fairly rapid.

The fix-point approach opens up a host of problems to explore about its sampling properties. For the moment we leave it at this, the limited purpose of illustration (62) and the four versions A–D being to provide a numerical illustration of the iterative procedure (43), and to check that it works in accordance with the theory in Part I. We proceed to some comments on the fix-point approach, with special regard to its scope and to a comparison with other aspects of interdependent systems.

[1] The computational work in generating the artificial sample was performed by Mr. K. Revankar, and in applying the fix-point estimation by Mr. H. Michalsen.

11. Comments

In the literature on interdependent systems the structural form of the system is usually defined in terms of current endogenous variables η and predetermined variables ζ, say

$$\eta = \beta\eta + \Gamma\zeta + \varepsilon^* \tag{83}$$

where the notation is chosen so as to emphasize the similarity and difference relative to interdependent systems (41) as dealt with in the present paper.[2] The system

$$\left. \begin{array}{l} \eta_1 = \beta_{12}\eta_2 + \gamma_{11}\zeta_1 + \gamma_{12}\zeta_2 + \varepsilon_1^* \\ \eta_2 = \beta_{21}\eta_1 + \gamma_{23}\zeta_3 + \gamma_{24}\zeta_4 + \varepsilon_2^* \end{array} \right\} \tag{84}$$

is the special case of (83) that corresponds to the system (62) that serves as illustration in sections 7 to 10.

We repeat that the fix-point procedure (43)–(46) for the estimation of inter-dependent system is not based on (83) but on the respecified version (41) of the structural form. The versions (41) and (83) of the structural form have numerically the same parameters; hence the estimation problem is precisely the same for the two versions. Thus when the parameters β, γ have been estimated by the fix-point procedure (43)–(46), the resulting parameters may also serve as estimates for the version (83). It will be noted that the residuals ε^* in (83) are linear expressions of the residuals ε in (41),

$$\varepsilon^* = (I - \beta)\varepsilon \tag{85}$$

We see from (51 a) that the reduced form of the system is the same for both of the versions (41) and (83) of the structural form, namely,

$$\eta = \Omega\zeta + \varepsilon = (I - \beta)^{-1} \Gamma\zeta + \varepsilon \tag{86}$$

To put it otherwise, the respecified version (41) of the structural form has the same residual elements $\varepsilon_1, \ldots, \varepsilon_n$ as the reduced form (86), but this is in general not so for the customary version (83) of the structural form.

In the following brief comments on the fix-point approach, section 11.1 fo-cusses upon the position of the approach relative to other estimation techniques, whereas section 11.2 deals with the rationale of version (41) relative to the customary version (83) of the structural form.

11.1 We shall bring into relief two fundamental differences between the fix-point iteration (43) relative to other techniques for the estimation of interdependent systems: (i) The fix-point method stays with the structural form (41), making no use of the reduced form (86); (ii) The fix-point method broadens the scope

---

[2] The reader is assumed to be oriented in the literature on interdependent systems. Initiated by T. Haavelmo in 1943, Ref. 12, the approach rapidly came to the fore. Refs. 13, 18 are fundamental works on the theoretical side, Refs. 15, 17 on the applied side. Refs. 14, 21 are recent textbook treatments. For the comparison of interdependent systems vs. causal chains, and for further references, see Refs. 6, 7.

of interdependent systems, inasmuch as the properties (58 a–b) of the residuals are more general than the assumptions imposed on the residuals in the current theory of interdependent systems.

11.1.1. A key feature of the theory of interdependent systems, emphasized ever since its beginnings, is that the parameter estimation will in general be inconsistent if performed by ordinary least squares regression (OLS; Refs. 12, 18). Much of the theory has been concerned with the development of consistent estimation techniques, among those the methods of limited information maximum likelihood (LIML; Refs. 8, 18), full information maximum likelihood (FIML; Refs. 11, 18), and two stage least squares regression (TSLS; Ref. 26). The LIML, FIML, TSLS and related estimation methods are known to work when applied to small size systems, and as is equally well known they meet with severe difficulties when it comes to medium size and large size systems. As emphasized by L. Klein, Ref. 16, the common source of these difficulties is that all of the methods at issue involve an auxiliary, prerequisite estimation of the reduced form of the system. In the reduced form each relation will in general involve all the predetermined variables of the entire model; hence for medium size and large systems the relations of the reduced form will make veritable "mammoth regressions", to quote L. Klein's apt description (Ref. 16; the oral discussion). The statistical treatment of these mammoth regressions and the reverse transformation back to the structural form will be marred by a complex of interwoven difficulties where collinearities, autocorrelations and shortage of degrees of freedom are main ingredients.

The fix-point method (43)–(46) is not touched by these difficulties since it works all the time on the structural form. As a matter of fact, my principal incentive for a fresh approach to the estimation problem of interdependent system was to avoid the passage to the reduced form. What makes this possible is that the respecification (41) of the customary structural form (83) gives the estimation problem a new twist, inasmuch as it takes the form of the fix-point problem (42). The lemma in section 3 shows that the transformation Au that defines the fix-point is a contraction mapping, so as a corollary to the general fix-point theorem of contraction mapping, Ref. 3, the fix-point (42) and thereby the solution to the estimation problem is provided by the iteration (43)–(44).

11.1.2. According to the current theory of interdependent systems, in the rest of this paper to be referred to as the classic theory, each residual $\varepsilon_1^*, ..., \varepsilon_n^*$ is assumed to be uncorrelated with all predetermined variables $\zeta_1, ..., \zeta_m$ of the entire system. We see that relations (85) imply that on the classic assumptions not only each $\varepsilon_i^*, ..., \varepsilon_n^*$ but also each residual $\varepsilon_i, ..., \varepsilon_n$ of the respecified structural form (41) will be uncorrelated with all of the predetermined variables. Now according to (58 a–b) the fix-point method (43)–(46) will give residuals $\varepsilon_i, ..., \varepsilon_n$ each of which is uncorrelated with some but in general not all of the predetermined variables. For illustration, reference is made to model (75) and Table 2, where ε_1 is correlated with ζ_3 and ζ_4, and ε_2 correlated with ζ_1 and ζ_2.

Thus with regard to the properties of the residuals, the fix-point approach constitutes a generalization of the current theory of interdependent systems, an extension that will be referred to as generalized interdependent (GEID) systems.

The theoretical foundations of the fix-point method (43)–(44) may be summarized as follows.

a. We seek a representation

$$\eta = \eta^* + \varepsilon \tag{87}$$

such that η^* is linear in the predetermined variables,

$$\eta^* = \sum_{k=l}^{m} \omega_k \zeta_k \tag{88}$$

and such that the structural form (41) is exactly satisfied by η^*, giving

$$\eta^* = \beta\eta^* + \Gamma\zeta \tag{89}$$

b. Relation (89) constitutes the fix-point formulation of the estimation problem. We note that no residuals ε_i or ε_i^* enter in this formulation.

c. According to the lemma in section 3 and the ensuing theorem in section 4, the fix-point η^* specified by (89) will under general conditions exist, be unique, and be delivered by the iteration (43)–(44).

The resulting residuals ε_i will for each i, according to (58), be uncorrelated with those variables η_p^* and ζ_q that occur in the ith relation of the structural form (41). We see that this makes as many noncorrelation properties as there are coefficients in the right-hand members of the structural form. Otherwise expressed, the noncorrelation properties are the same in number as the product moments of type

$$(\eta_i, \eta_p^*), (\eta_i, \zeta_q) \qquad i = 1, \ldots, n \tag{90}$$

where as before η_p^*, ζ_q are the variables that occur in the right-hand member of the ith relation of the structural form.

d. Owing to the uniqueness of the parameters β, γ delivered by the fix-point method (43)–(46) it will in general not be possible to impose more noncorrelation properties upon the residuals than those given by (58). Thus if the reduced form (86) involves more coefficients than the structural form (41), and this is so if the system is overidentified (see Ref. 14), one or more of the residuals ε_i will in general be correlated with one or more of the predetermined variables ζ_k. Again the situation is illustrated by system (75), which is overidentified, and where ε_1 is correlated with ζ_3 and ζ_4.

e. An equivalent restatement of the situation in overidentified systems is as follows: There are nm product moments (η_i, ζ_k), whereas the moments (90) are fewer in number. The moments (90) will determine the parameters of the model, and thereby all of the product moments (η_i, ζ_k). Hence there are one or more moments (η_i, ζ_k) that are uniquely determined by the moments (90). Thus if we interpret the product moments (η_i, ζ_k) as points in a Euclidean space R_{nm}, the moments (90) will form a space, say R^*, which has lower dimensionality than R_{nm}. Each point in R^* will provide complete specification of all parameters β, γ of the system.

f. The argument a–e may be developed as follows.

Definition. If the moments (90) are less than nm in number, we distinguish between classic (CLID) and generalized (GEID) interdependent systems (41) according as each residual ε_i is assumed to be uncorrelated with (a) all of the predetermined variables ζ_k, or (b) only those variables η_p^*, ζ_q that occur in the ith relation of the system.

The extension from CLID to GEID systems has deepgoing implications for the theory of interdependent systems. We do not propose to take up here a general discussion of the rationale of this extension. The following comments will be limited to more or less immediate conclusions regarding similarities and differences between the two types of system.

g. The fix-point method (43)–(46) applies to CLID systems, since these constitute a special case of GEID systems.

Any *just identified* interdependent system (see Ref. 14) is a CLID system.

h. The FIML, LIML, TSLS and other methods based on the classic assumptions are consistent when applied to CLID systems.

When applied to the population (probability distribution) defined by a CLID system, the TSLS method will coincide with the fix-point method (43) if we take $\eta^{(0)}$ to be the first stage of the TSLS procedure. The second stage of the TSLS method will then coincide with $\eta^{(1)}$, and the iteration (43) will stop in the second step, giving $\eta^{(1)} = \eta^*$.

If applied to the sample, not to the population, the fix-point estimate (44) will in general not be the same as the corresponding TSLS estimate.

j. The FIML, LIML, TSLS and other estimates based on the reduced form and the classic assumptions will in general be inconsistent when applied to GEID systems.

Illustration. We return to model (75) and the corresponding population as defined by the product moments (η_i, η_k), (η_i, ζ_k), (ζ_i, ζ_k) specified in (69) and Table 1. Applying the TSLS method to this population, we obtain

$$\left.\begin{aligned}
\eta_1 &= .816\ \eta_2^* + .921\ \zeta_1 - .382\ \zeta_2 + u_1 \\
\eta_2 &= .587\ \eta_1^* + .833\ \zeta_3 + .451\ \zeta_4 + u_2
\end{aligned}\right\} \tag{91}$$

Comparing with the true values (71) we see that the inconsistency of the TSLS estimates is quite considerable. The worst case is the coefficient of ζ_2, where the TSLS estimate deviates by -138%.

k. We shall say that an interdependent system (41) is *invariant to inversion* if the system allows interchange of the current endogenous variables from the right to the left member, subject to the corresponding algebraic transformation of parameters and residuals. In the current theory of interdependent systems, the structural form is often treated as invariant to inversion; whether and why this is permissible is not quite clear from the literature. We note that GEID systems in general are not invariant to inversion, neither w.r.t. the parameters, nor w.r.t. the residuals. CLID systems are invariant to inversion w.r.t. the parameters, but in general not w.r.t. the residuals.

Again returning to the system (62), the following two illustrations refer to the model (75) as specified in the population, not the estimate (81) as obtained from the sample.

Illustration 1. Let us interchange η_1 and η_2 in system (75), without change of the moments in (69) and Table 1. Applying the fix-point method (43) after the inversion we obtain the following estimates,

$$
\left.
\begin{aligned}
\eta_1 &= .859 \; \eta_2^* + .074 \; \zeta_3 + .946 \; \zeta_4 + \vartheta_1 \\
\eta_2 &= .986 \; \eta_1^* - .282 \; \zeta_1 + .233 \; \zeta_2 + \vartheta_2
\end{aligned}
\right\}
\tag{92}
$$

We see that the model is far from invariant to inversion. In case of invariance the coefficient of, say, η_2^* should be $(.4)^{-1} = 2.5$, and the coefficient of ζ_1 should be -2.0. The coefficients of $\zeta_2, \zeta_3, \zeta_4$ should not as here be positive.

Illustration 2. If (75) were a CLID system, that is, if both ε_1 and ε_2 were uncorrelated with all of $\zeta_1, \zeta_2, \zeta_3, \zeta_4$, would the coefficients of the system (92) be obtained from (75) by formal inversion? In CLID systems (75) the coefficients are invariant to inversion in this sense (so that the coefficient of, say, η_2^* in (92) would be 2.5). Well to note, we must not expect that the residuals participate in the inversion; if they did, we should have

$$
\vartheta_1 = -2.5 \, \varepsilon_2; \quad \vartheta_2 = -2 \, \varepsilon_1
\tag{93}
$$

whereas CLID systems (75) in actual fact give

$$
\vartheta_1 = \varepsilon_1; \quad \vartheta_2 = \varepsilon_2
\tag{94}
$$

We see that the situation in this respect is the same as when a GEID system is transformed from the structural form (41) to the reduced form (86), inasmuch as the two forms have the same residuals ε.

11.2. The notions

    (i) *realizable* models, and
    (ii) *minimum-delay* models

make categories of fundamental relevance for the rationale of forecasting models (Refs. 23, 24, 32). We shall now in all brevity comment upon interdependent systems from the point of view of these notions.

11.2.1. *Classic (83) vs. respecified (41) interdependent systems*

When written in the customary version (83), an interdependent system in general is *not realizable*. A case in point is system (84) if $\beta_{12} \neq 0$ and $\beta_{21} \neq 0$. In fact, if we interpret the model as describing a filter with input and output variables, we see from the right-hand members that $\eta_1, \eta_2, \zeta_1, \zeta_2, \zeta_3, \zeta_4$ are input variables, and from the left-hand members that η_1, η_2 are output variables. Hence

η_1 and η_2 represent both input and output, a situation that is not technically realizable.

When written in the respecified form (41) an interdependent system is realizable. In system (75), for example, all input variables $\eta_1^*, \eta_2^*, \zeta_1, \zeta_2, \zeta_3, \zeta_4$ are predetermined variables or — cf. (88)—linear combinations of the predetermined variables, whereas η_1, η_2 are output variables. This is a perfectly realizable situation.

It is instructive to note that the residuals ε^* of the classic version (83) may be smaller than in the respecified version (41). In model (75), for example, we infer from Table 3 b

$$\sigma^2(\varepsilon_1) = \sigma^2(\varepsilon_2) = 3 \tag{95}$$

whereas (84) gives, as is readily verified,

$$\sigma^2(\varepsilon_1^*) = 2.25; \quad \sigma^2(\varepsilon_2^*) = 2.28 \tag{96}$$

To paraphrase, if we could use the classic version (84 a) to forecast η_1 for given η_2, and at the same time use (84 b) to forecast η_2 for given η_1 (a situation that clearly is not realizable) it would be possible to obtain smaller forecasting error variances than if we use (75 a) to forecast η_1 for given η_2^* and use (75 b) to forecast η_2 for given η_1^*.

11.2.2. *Interdependent systems vs. causal chains.*

A multirelation system (83) is called a *causal chain*, or a *recursive system* if the position matrix β is subdiagonal, that is, if

$$\beta_{ik} = 0; \quad k \geqslant i, \quad i = 1, \dots, n \tag{97}$$

The notion of *minimum delay* replaces and extends the engineering notion of minimum-phase, Ref. 23; whereas the minimum-phase concept applies only to unirelation models, the minimum-delay concept applies both to unirelation and multirelation systems. The notion of minimum delay will here be applied in the following situation:

Suppose we know the past values of the endogenous variables, say

$$\eta_{is} \quad \text{with} \quad s = t-1, t-2, \dots \tag{98}$$

and the past and present values of the predetermined variables $\zeta_1, \dots, \zeta_m$. On the basis of (98) and the predetermined variables we form a sequence of conditional forecasts of the current endogenous variables $\eta_{1t}, \dots, \eta_{nt}$, in symbols,

$$\text{Pred } \eta_{it} = L(\text{Pred } \eta_{i-1,t}, \dots, \text{Pred } \eta_{1t}; \zeta_{1s}, \dots, \zeta_{ms}) \tag{99}$$

where $i = 1, \dots, n; s \leqslant t$, and L is a linear expression formed in accordance with the given model. The model is called minimum delay if the error of the forecast (99) has the smallest possible variance for $i = 1, \dots, n$.

We note: (i) For causal chain systems the parameters can be specified so that the model is minimum delay. (ii) An interdependent system (41) in general is not minimum delay.

As to (i), reference is made to the recent work of E. A. Robinson, Refs. 23, 24. As to (ii), we see that the statement is true since the forecast (99) does not exploit the information that may lie in the current residuals $\varepsilon_{1t}, \varepsilon_{2t}, ..., \varepsilon_{i-1,t}$.

12. Further developments

The results in sections 3–5 can be extended in several directions. The subsequent statements do not aim at the greatest possible generality, and the proofs will only be briefly indicated.

12.1. *Sampling properties of the fix-point estimates*

The fix-point method (43)–(46) provides simultaneous estimation of the parameters β, γ and the expectations η^*. The sampling properties of the ensuing estimates involves several groups of problems; hereunder:

(i)   Consistency of the parameter estimates b, g.

(ii) Large-sample standard errors and sampling covariances of the parameter estimates b, g.

(iii) Large-sample standard deviations and product moments of the estimated expectations y^*.

(iv) Consistency, large-sample standard errors and sampling covariances of the estimated expectations y^*.

In dealing with the problems under (i)–(iv) it makes for simplicity that the fix-point estimates (44) may in the limit be regarded as least squares regressions,

$$y_i = \sum_p b_{ip} y_p^* + \sum_q g_{iq} z_q + e_i; \quad i = 1, ..., n \qquad (100)$$

with the limiting expectations y^* as ordinary explanatory variables.

As to (i) we see that the consistency of ordinary regression estimates, Ref. 28, readily extends to the fix-point estimates. The fix-point estimates b, g are determined by the sampling product moments $(y_i, y_k), (y_i, z_k), (z_i, z_k)$. Thus if the sample is ergodic in the sense that each of these product moments tends to the corresponding product moment in the population, it follows that the estimates b, g tend to the parameters β, γ that define the population.

As to (ii) we note that the estimates y_p^* involve sampling errors of the same order or magnitude as those in b_{ip} and g_{iq}. Hence in (100) the terms $b_{ip} y_p^*$ will have to be treated as twin terms

$$b_{ip} \eta_p^* + \beta_{ip} y_p^* \qquad (101)$$

On the assumption that all residuals $\varepsilon_{is}, \varepsilon_{kt}$ are uncorrelated for $s \neq t$, the resulting matrix of large sample variances and covariances of b, g is

$$E[T(a - \alpha)(a - \alpha)'] \sim s^2(e_i^*) \left(\frac{1}{T} XX'\right)^{-1} \qquad (102)$$

where we have used the abridged notation

$$a = \{b_{ip}, g_{iq}\}; \quad x = \{y_{ip}^*, z_{iq}\}; \quad x_i = [x_{i1}, \ldots, x_{iT}]; \quad X = [x_i] \tag{103}$$

and T is the number of observations in the sample.

Formula (102) allows extension to the case of autocorrelated residuals ε_i. The ensuing deductions are of the type given for ordinary least squares regression by H. Wold, Refs. 27, 31, for the case of exogenous explanatory variables, and by E. Lyttkens, Refs. 19, 20, for the case of exogenous and/or endogenous variables.

It will be noted that (102) coincides with H. Theil's large-sample formula for the standard errors of TSLS estimates for the parameters of CLID systems, Ref. 26. The following formula for the large-sample standard errors of TSLS estimates of CLID parameters is asymptotically equivalent to the formula given by H. Theil, Ref. 26, and has been obtained by U. Norlén by a different line of argument,

$$E[T(a - \alpha)(a - \alpha)'] \sim s^2(e_{i(2)}^*) \left(\frac{1}{T} X_{(1)} X_{(1)}'\right)^{-1} \tag{104}$$

where the matrix $X_{(1)}$ involves the first stage estimates of η^*, and $e_{i(2)}^*$ are the second stage residuals as transformed in accordance with (85).

The problems under (iii)–(iv) are related to those under (i)–(ii), lead to similar deductions, and are on the whole simpler. The ensuing formulas will be reported elsewhere, in connection with Monte Carlo studies of the sampling properties of the fix-point method (43)–(46). An interesting aspect is the distinction between two situations, namely (a) when the observations $z_{1t}, \ldots, z_{nt}$ of the predetermined variables are the same in the various replications of the sample, and (b) when the observations of the predetermined variables are generated anew for each replication.

12.2. *Interdependent systems that involve identities*

By definition, a relation between random variables is called an *identity* if it is deterministic (residual-free), and the coefficients are known *a priori*. Any identity that involves one or more current endogenous variables can be used to eliminate one of these variables from a model which includes the identity.

The lemma and theorem of Part I have been given for interdependent systems in the standard form (41), which does not involve identities. The lemma and theorem allow straightforward extension to systems that involve one or more identities. We shall here consider a special case that illustrates the following features:

(i) For GEID systems, but not for CLID systems, the numerical specification will in general depend upon which endogenous variable we choose to eliminate by the use of an identity. (ii) The choice specifies the variable in the identity that is assumed to be determined by the other variables, and this causal specification is part of the subject matter specification of the entire model. (iii) The multiplicity of different versions of the model resulting from different choices of variables to be eliminated increases quite rapidly with the number of identities.

Illustration. Systems of the following type are standard material in the literature on interdependent systems,

$$
\left.\begin{aligned}
\eta_c &= \beta_c \, \eta + \gamma_{c1} \, \zeta_1 + \gamma_{c2} \, \zeta_2 + \varepsilon_c^* \\
\eta_i &= \beta_i \, \eta + \gamma_{i3} \, \zeta_3 + \gamma_{i4} \, \zeta_4 + \varepsilon_i^* \\
\eta &= \eta_c + \eta_i
\end{aligned}\right\}
\tag{105}
$$

where $\eta_c =$ consumption, $\eta_i =$ investment, $\eta =$ gross national product.

We respecify the behavioural relations (105 a–b) in accordance with (41),

$$
\left.\begin{aligned}
\eta_c &= \beta_c \, \eta^* + \gamma_{c1} \, \zeta_1 + \gamma_{c2} \, \zeta_2 + \varepsilon_c \\
\eta_i &= \beta_i \, \eta^* + \gamma_{i3} \, \zeta_3 + \gamma_{i4} \, \zeta_4 + \varepsilon_i
\end{aligned}\right\}
\tag{106 a–b}
$$

and correspondingly for the identity (105 c),

$$
\eta^* = \eta_c^* + \eta_i^*, \quad \text{giving} \quad \varepsilon = \eta - \eta^* = \varepsilon_c + \varepsilon_i
\tag{106 c}
$$

The identity (105 c) with (106 c) allows us to eliminate any of the three endogenous variables η, η_c, η_i. Eliminating η, we obtain

$$
\left.\begin{aligned}
\eta_c &= \beta_c(\eta_c^* + \eta_i^*) + \gamma_{c1} \, \zeta_1 + \gamma_{c2} \, \zeta_2 + \varepsilon_c \\
\eta_i &= \beta_i(\eta_c^* + \eta_i^*) + \gamma_{i3} \, \zeta_3 + \gamma_{i4} \, \zeta_4 + \varepsilon_i
\end{aligned}\right\}
\tag{107}
$$

or alternatively,

$$
\left.\begin{aligned}
\eta_c &= \beta_c' \, \eta_i^* + \gamma_{c1}' \, \zeta_1 + \gamma_{c2}' \, \zeta_2 + \varepsilon_c' \\
\eta_i &= \beta_i' \, \eta_c^* + \gamma_{i3}' \, \zeta_3 + \gamma_{i4}' \, \zeta_4 + \varepsilon_i'
\end{aligned}\right\}
\tag{108}
$$

If instead we eliminate η_c we obtain

$$
\left.\begin{aligned}
\eta - \eta_i^* &= \beta_c'' \, \eta^* + \gamma_{c1}'' \zeta_1 + \gamma_{c2}'' \zeta_2 + \varepsilon'' \\
\eta_i &= \beta_i'' \, \eta^* + \gamma_{i3}'' \zeta_3 + \gamma_{i3}'' \zeta_4 + \varepsilon_i''
\end{aligned}\right\}
\tag{109}
$$

or alternatively,

$$
\left.\begin{aligned}
\eta &= \beta''' \, \eta_i^* + \gamma_1''' \zeta_1 + \gamma_2''' \zeta_2 + \varepsilon''' \\
\eta_i &= \beta_i''' \, \eta^* + \gamma_{i3}''' \zeta_3 + \gamma_{i4}''' \zeta_4 + \varepsilon_i'''
\end{aligned}\right\}
\tag{110}
$$

Thirdly, elimination of η_i leads to systems analogous to (109)–(110).

The fix-point method applies to (107)–(110) as well as the third pair of systems. For version (107) the mapping (13) reads

$$
\left.\begin{aligned}
(Au)_c &= \beta_c(u_c + u_i) + \gamma_{c1} \, \zeta_1 + \gamma_{c2} \, \zeta_2 \\
(Au)_i &= \beta_i(u_c + u_i) + \gamma_{i3} \, \zeta_3 + \gamma_{i4} \, \zeta_4
\end{aligned}\right\}
\tag{111}
$$

and for (109),

$$
\left.\begin{aligned}
(Au)_y - u_i &= \beta_c'' u_y + \gamma_{c1}'' \zeta_1 + \gamma_{c2}'' \zeta_2 \\
(Au)_i &= \beta_i'' u_y + \gamma_{i3}'' \zeta_3 + \gamma_{i4}'' \zeta_4
\end{aligned}\right\}
\tag{112}
$$

When eliminating a variable from a system by the use of an identity, the ensuing versions of the system may or may not be numerically equivalent. In

the present case, for example, systems (107) and (108) are equivalent, whereas (108) and (110) are not. As to this last point, we note regarding (108) that ζ_1 and ζ_2 in general will be uncorrelated with ε_1 and correlated with ε_2, which implies that $\varepsilon = \varepsilon_c + \varepsilon_i$ in general will be correlated with ζ_1 and ζ_2. In (110), on the other hand, ε'' will be uncorrelated with ζ_1 and ζ_2, in contrast to our conclusion regarding the residual in (108).

12.3. *Various extensions*

12.3.1. *Generalization of the lemma*

In section 3 the lemma has been proved on the assumption that each relation of the system involves a nonvanishing residual ε. The lemma extends to certain systems where one or more ε_i vanish, a simple case being the following.

Given a system (41) where the position matrix β picks out just one variable y_p^* for each relation, we shall say that the variables $y_1 \ldots y_n$ form a *loop of interdependence* if the relation for y_1 involves y_n^* and the relation for y_i involves $y_{i-1}^* (i = 2, \ldots, n)$; more generally, the variables y_i will be said to form a loop of interdependence if the above condition is fulfilled after a suitable renumbering. Then for the lemma to be valid it is sufficient that the variables $y_1, \ldots, y_n$ form a loop of interdependence, and as least one of the residuals ε_i has positive norm, $\| \varepsilon_i \| \geqslant c > 0$.

The above extension of the lemma requires only a slight change in the argument (30)–(33) if we adopt the following vector distance instead of (2),

$$D(x, y) = \left\{ \sum_{i=1}^{n} d^2(x_i, y_i) \right\}^{1/2} \tag{113}$$

To repeat from section 1, the Hilbert space H under consideration is assumed to have real-valued inner products. Making use of standard arguments, the lemma and the theorem extend to the case of complex-valued inner products.

12.3.2. *Systems with nonlinear features*

The fix-point method is distribution-free in the same sense as ordinary least squares regression, since each step in the iteration (43) is distribution-free. Hence in a linear system (43) the fix-point method allows us to perform nonlinear transformations of the variables, introducing for example y_i^2 or $\log y_i$ instead of y_i. In this respect the fix-point method is similar to the TSLS method, and different from the FIML method.

As to nonlinearities in the coefficients, a situation like

$$y_1 = \beta y_2^* + \beta^2 z_1 + \varepsilon$$

can be handled by a Lagrange multiplier λ,

$$y_1 = \beta_1 y_2^* + \beta_2 z_1 + \lambda(\beta_2 - \beta_1^2) + \varepsilon$$
$$\beta_2 = \beta_1^2$$

The fix-point y^* will belong to the linear space (88). To assess the requisite coefficients numerically, we must solve a polynomial equation system in several variables.

Nonlinearities in the form of products, quotients etc. of variables involve more serious difficulties, for example

$$y_1 = \beta y_2^* z_1 + \varepsilon$$

We see that the fix-point formulation (42) here involves a transformation (13) that is nonlinear and accordingly does not transform the linear space (88) into itself. Under favourable circumstances we may seek for approximate solutions by suitable linearizations. For example, writing

$$z_1 = m\left(1 + \varphi \cdot \frac{\sigma}{m}\right) \tag{114}$$

and similarly for the other variables, we see that if the distribution of each variable is closely concentrated around its mean, we can linearize the product in (114) by omitting the term that involves the product of two factors φ.

REFERENCES

8. ANDERSON, T. W., and RUBIN, H., "Estimation of the parameters of a single equation in a complete system of stochastic equations." *Annals of Mathematical Statistics*, 20, 46–63 (1949).
9. CRAMÉR, H., *Mathematical methods of statistics*. Uppsala, Almqvist och Wiksells, 1945. New York, Wiley and Sons, 1946.
10. DAVIES, J., NORLÉN, U., STEWARD, D., and WOLD, H., "Generalized interdependent systems: Parameter estimation by iterative least squares." *Systems Formulation and Methodology Workshop Paper* 6413, Social Systems Research Institute, Univ. of Wisconsin, 1964.
11. EISENPRESS, H., *Forecasting by Econometric Systems*. IBM Data Systems Division, Library Services Dept., New York, 1962.
12. HAAVELMO, T., "The statistical implications of a system of simultaneous equations." *Econometrica*, 11, 1–12 (1963).
13. HOOD, W. C., and KOOPMANS, T. C., eds., *Studies in econometric method*. New York, Wiley and Sons, 1953.
14. JOHNSTON, J., *Econometric methods*. New York, McGraw-Hill, 1963.
15. KLEIN, L. R., *Economic fluctuations in the United States 1921–1941*. New York, Wiley and Sons, 1950.
16. ——, "Problems in the estimation of interdependent systems." Forthcoming in the *Transactions des Entretiens de Monaco 1964*. Centre International d'Étude des Problèmes Humains, Monaco. (In press.)

17. KLEIN, L. R., and GOLDBERGER, A., *An econometric model of the United States 1929–1952.* Amsterdam, North-Holland Publ. Co, 1955.
18. KOOPMANS, T. C., ed., *Statistical inference in dynamic economic models.* New York, Wiley and Sons, 1950.
19. LYTTKENS, E., "Standard errors of regression coefficients in the case of autocorrelated residuals." Paper in *Ref. 25,* pp. 38–60 (1963).
20. ——, "Standard errors of regression coefficients by autocorrelated residuals." Paper in *Ref. 30,* pp. 169–228 (1963).
21. MALINVAUD, E., *Méthodes statistiques de l'économétrie.* Paris: Dunod, 1964.
22. MICHALSEN, H., *BESILS-metoden, Estimering av parametre for interdependente systemer.* SINTEF, Trondheim, Norway, 1964.
23. ROBINSON, E. A., *Random vawelets and cybernetic systems.* London, Griffin, 1962.
24. ——, "Structural properties of stationary stochastic processes with applications." Paper in *Ref. 25,* pp. 170–196 (1963).
25. ROSENBLATT, M., ed., *Proceedings of the Symposium on Time Series Analysis, Brown Univ. June 1962.* New York, Wiley and Sons, 1963.
26. THEIL, H., *Economic forecasts and policy.* Amsterdam, North-Holland Publ. Co., 1958.
27. WOLD, H., "On least square regression with autocorrelated variables and residuals." *Bulletin de l'Institut International de Statistique, 32,* no. 2, 277–289 (1950).
28. ——, "On the consistency of least squares regression." *Sankhyā,* Ser. A, *25,* 211–215 (1963).
29. ——, "A fix-point theorem with econometric background. I." *Arkiv för Matematik, 6,* pp. 209–220 (1965).
30. ——, ed., *Econometric model building. Essays on the causal chain approach.* Amsterdam, North-Holland Publ. Co., 1964.
31. ——, in association with L. Juréen, *Demand analysis. A study in econometrics.* Stockholm, Gebers, 1952. New York, Wiley and Sons, 1953.
32. WOLD, H., and ROBINSON, E. A., "Minimum-delay structure of least-squares and eo-ipso predicting systems for stationary stochastic processes". Appendix in *Ref. 24,* pp. 192–196, 1963.