COWLES COMMISSION DISCUSSION PAPER: STATISTICS: NO. 328

Computational Methods Used in Limited Information Treatment of a Set of Linear Stochastic Difference Equations

by Jean Bronfenbrenner
Herman Chernoff

February 22, 1949

A. Notation

We consider a system of equations of the form

$$B y' + \Gamma z' + \mathcal{E}' = u',$$

and we write a single equation of the system in the form

$$\beta y^{*\prime} + \gamma z^{*\prime} + \mathcal{E} = u.$$

The following symbols will be used in describing the computational procedure for treating this system of equations:

y_t = vector of endogenous variables (at time t)

z_t = vector of predetermined variables (at time t)

u_t = vector of stochastic disturbances (at time t)

y_t^* = vector of endogenous variables appearing in a single equation (at time t)

z_t^* = vector of predetermined variables appearing in a single equation (at time t)

$x_t = (y_t \quad z_t)$

$M_{yy} = T \sum_t y_t' y_t - \sum_t y_t' \sum_t y_t$ = moment matrix of the endogenous variables. (M_{zy}, M_{zz}, M_{z*y*}, M_{z*z*} analogously defined.)

m = vector of sample means of variables appearing in a single equation

A. Notation (continued)

a = vector of adjustment factors

T = size of sample

H = number of elements in y^*

K^* = number of elements in z^*

$F = H + K^*$

K = number of elements in z

$K^{**} = K - K^*$

Vectors normally are considered to be row vectors; and a column vector is written as the transpose of the corresponding row vector (e.g., β' is a column vector).

The subscript "i" attached to a vector denotes the i^{th} element of the vector. The subscript "ij" attached to a matrix denotes the ij^{th} element of the matrix.

The left subscript "1" attached to a vector indicates the omission of the first element of the vector, i.e., ${}_1\beta = (\beta_2, \beta_3 \cdots, \beta_k)$. The left subscripts "1 0," "0 1," and "1 1" attached to a matrix indicate respectively the omission of the first row, the omission of the first column, and the omission of the first row and column.

If u is a variable vector, the symbol $\sigma[u]$ denotes the covariance matrix of the elements of u.

B. Mathematical outline

1. Estimation of the parameters of a single equation

a) $W_{y^*y^*} = M_{y^*y^*} - M_{y^*z}M_{zz}^{-1}M_{zy^*}$ 1/

1/ This matrix may be interpreted as that part of the variance and covariance of the elements of y^* which is not explained by z.

B. Mathematical outline (continued)

b) $P^{*\prime} = M^{-1}_{z^*z^*}M_{z^*y^*}$

c) $B_{y^*y^*} = M_{y^*z}M^{-1}_{zz}M_{zy^*} - M_{y^*z^*}P^{*\prime}$ 1/

d) $A = B^{-1}_{y^*y^*}W_{y^*y^*}$

e) $\hat{\beta}' =$ solution of

$(A - \lambda I)\beta' = 0$

where λ is largest characteristic value of A.

f) $\hat{\gamma}' = - P^{*\prime}\,\hat{\beta}'$

g) $\hat{\varepsilon} = - \frac{1}{T}(\hat{\beta}\sum_t y^{*\prime}_t + \hat{\gamma}\sum_t z^{*\prime}_t) = -(\hat{\beta} \quad \hat{\gamma})m'$

h) $(b \quad c \quad e) = \frac{1}{\hat{\beta}_1}(\hat{\beta} \quad \hat{\gamma} \quad \hat{\varepsilon})$ 2/

i) $s^2 = \frac{\sum_t u_t^2}{T - F}, \qquad \sum_t u_t^2 = (1 + \frac{1}{\lambda})\hat{\beta}W_{y^*y^*}\hat{\beta}'/T(\hat{\beta}_1)^2$

j) $\frac{\delta^2}{s^2} = \frac{\sum_t (u_t - u_{t-1})^2}{\sum_t u_t^2}\left(1 + \frac{1}{T - F - 1}\right)$ 3/

1/ This matrix may be interpreted as that part of the variance and co-variance of the elements of y^* which is **explained by z but not by z*.**

2/ This step gives the normalization $b_1 = 1$.

3/ $\left|\frac{\delta^2}{s^2} - 2\right|$ measures the serial correlation of the calculated residual.

See Von Neumann, J. and B. I. Hart, "Tabulation of the Probabilities for the Ratio of the Mean Square Successive Difference to the Variance," Annals of Mathematical Statistics, 1942, p. 207.

B. Mathematical outline (continued)

2. Covariances of the estimates of the parameters of a single equation

a) $C = (1+\nu)\beta W_{y^*y^*}\beta'$ [1/] where $\nu = \frac{1}{\lambda}$

b) $F_{\beta\beta} = \left[{}_{11}\left(B_{y^*y^*} - \frac{\nu}{\beta W_{y^*y^*}\beta'} W_{y^*y^*}\beta'\beta W_{y^*y^*}\right)\right]^{-1}$

c) $F'_{\beta\gamma} = {}_{01}\left[P^{\dagger\prime}\right] F_{\beta\beta}$

d) $F_{\gamma\gamma} = {}_{01}\left[P^{\dagger\prime}\right] F_{\beta\gamma} + M^{-1}_{z^*z^*}$

e) $\hat{\sigma}[(\beta \;\; \gamma)] = \frac{C}{T-F} \begin{pmatrix} F_{\beta\beta} & -F_{\beta\gamma} \\ & F_{\gamma\gamma} \end{pmatrix}$ [1/]

f) $$\hat{\sigma}[(b \;\; c \;\; e)] = \begin{pmatrix} \frac{1}{\beta_1^2}\sigma[(\beta \;\; \gamma)] & -\frac{1}{\beta_1^2}\sigma[(\beta \;\; \gamma)]\,({}_1m)' \\ & ({}_1m)\frac{1}{\beta_1^2}\sigma[(\beta \;\; \gamma)]\,({}_1m)' + \frac{\sum_t u_t^2}{T(T-F)} \end{pmatrix}$$

C. Computational outline

A schematic representation of the computation procedures follows in Tables I, II, and III.

For purposes of actual computation it is usually convenient to

---

1/ The carets in $\hat{\beta}$ and $\hat{\gamma}$ will henceforth be omitted. It will be clear from the context whether parameters or estimates are referred to. Similarly for $\hat{\sigma}$.

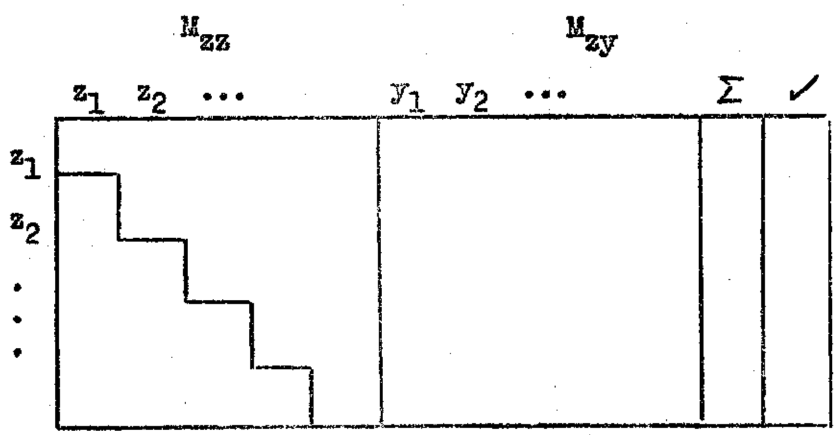
C. <u>Computational outline</u> (continued)

adjust each time series by multiplying it by an appropriate power of ten. Unless otherwise indicated (by the presence of a superscript "n"), the matrices and vectors in the following outline refer to the adjusted time series $x_{ti} = a_i x_{ti}^{(n)}$.

The columns headed "Σ" and "✓" are summation and check columns. Symmetric matrices with elements omitted below the main diagonal are represented by a zigzag diagonal line as in M_{zz} below.

Table I

Computations Used for all Equations of the System



Doolittle Forward Solution

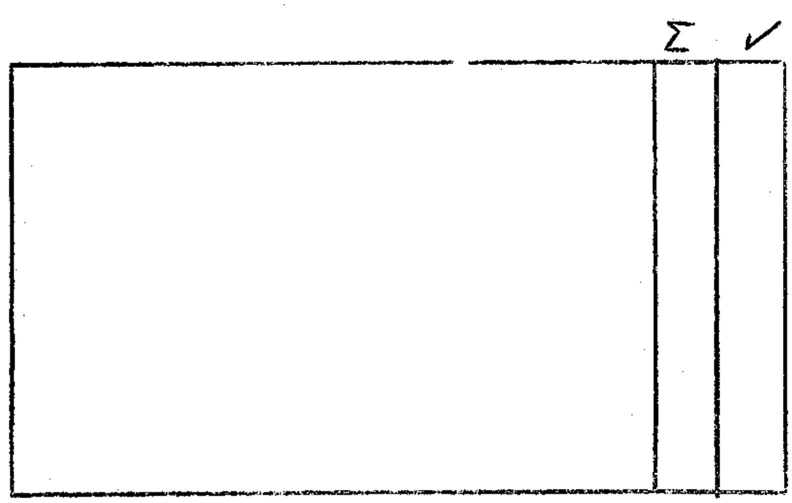
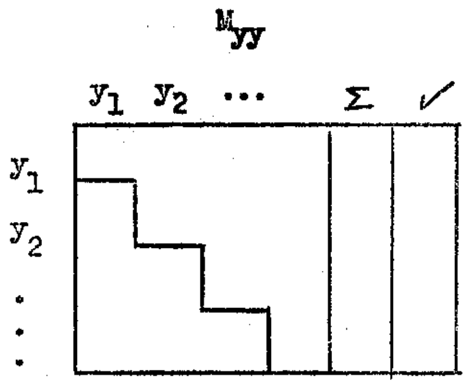
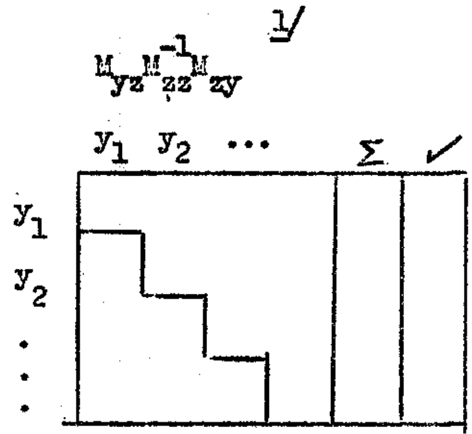


Table I (continued)





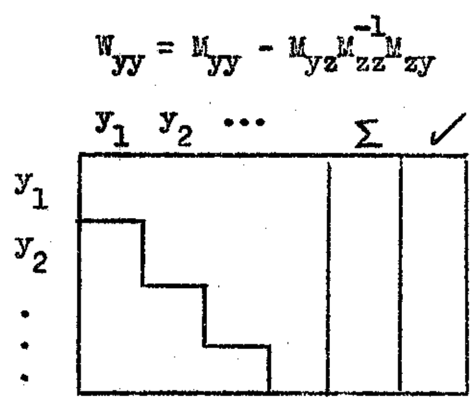
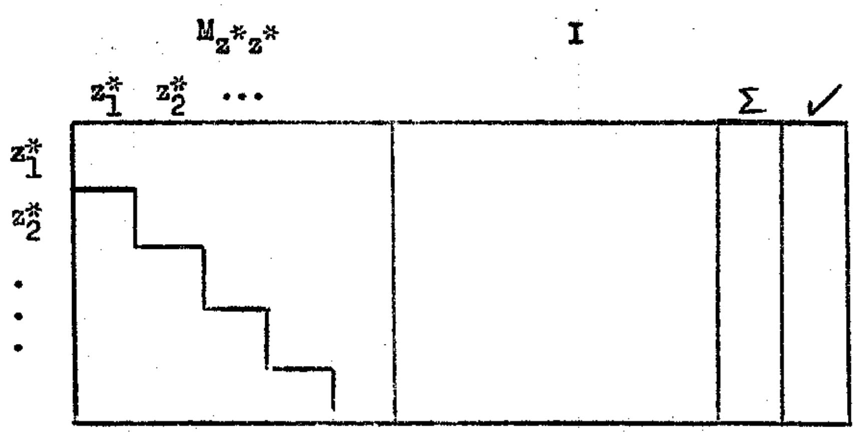


Table II

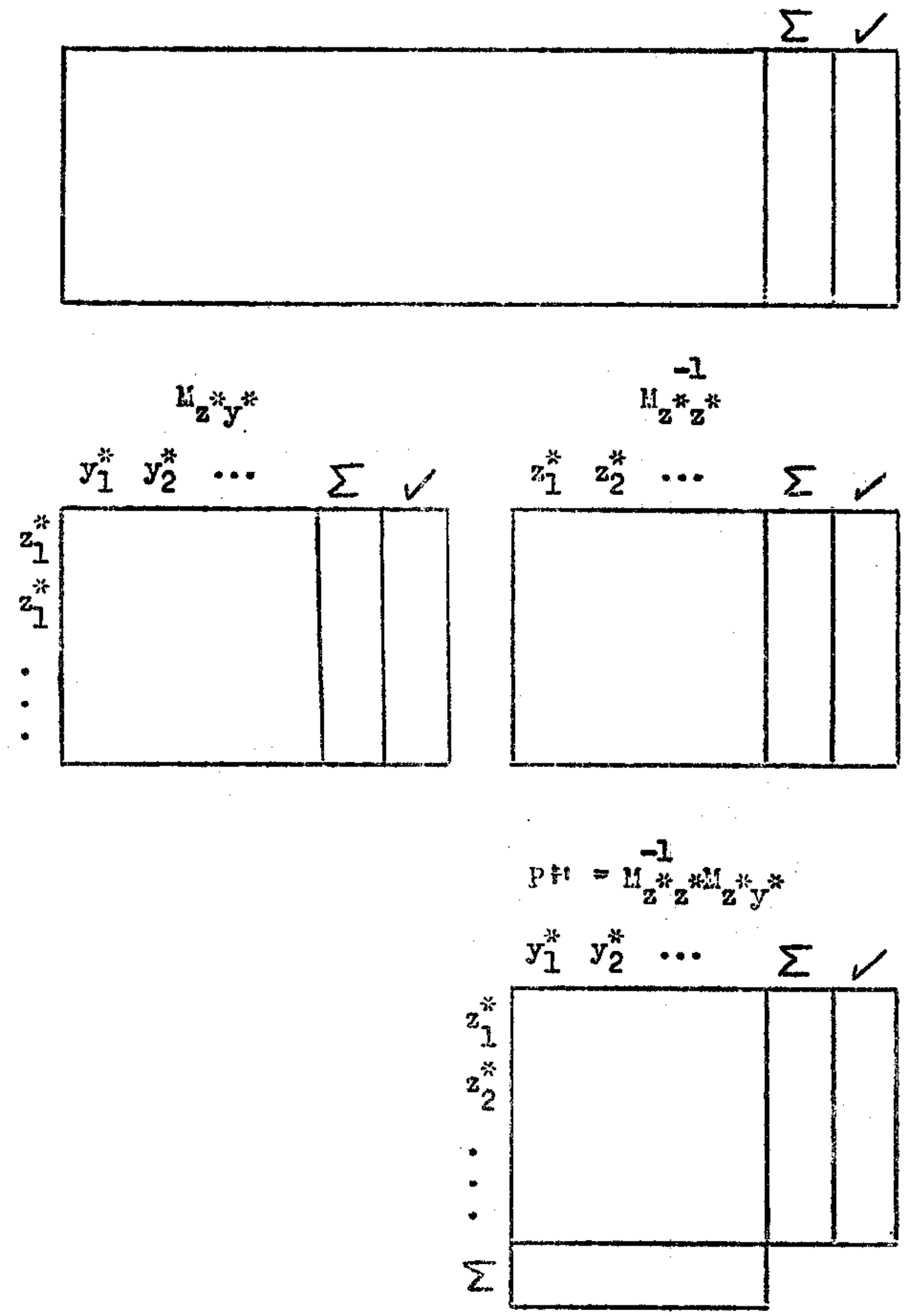
Computations for a Single Equation



1/ This matrix may be obtained directly from the forward solution in the following way. In the forward solution let a_{ij} and b_{ij} be the $(K + j)$th elements of the i-th odd row and the i-th even row, respectively. (Here

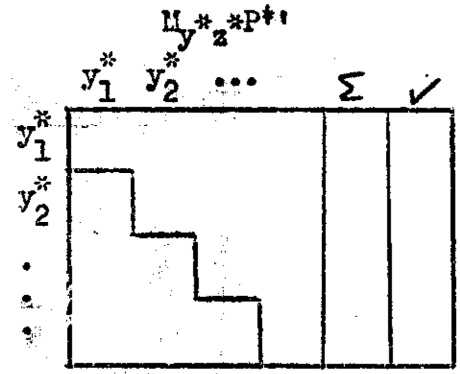
Table II (continued)

Doolittle Forward Solution



K is the order of M_{zz}.) Then the mn-th element of $M_{yz}M_{zz}^{-1}M_{zy}$ is given by $\sum_{t} a_{im}b_{in}$.

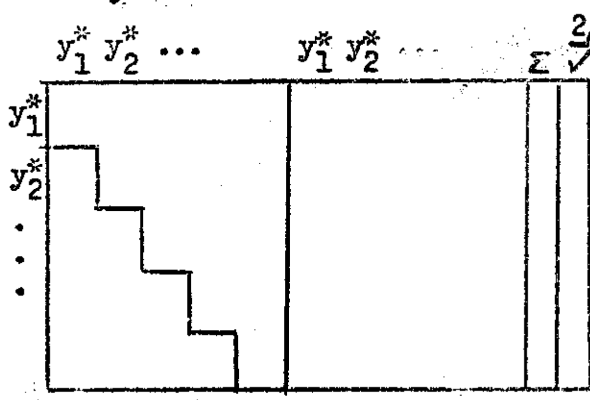
Table II (continued)



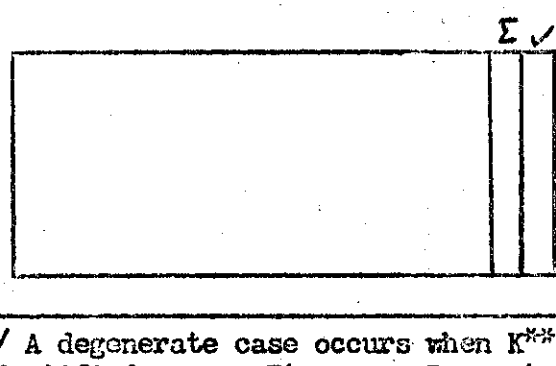
<u>Case I</u>: H > 2, $B_{y^*y^*}$ nonsingular 1/

$B_{y^*y^*} =$
$M_{y^*z}M_{zz}^{-1}M_{zy^*}$
$- M_{y^*z^*}P^{*\prime}$

$W_{y^*y^*}$



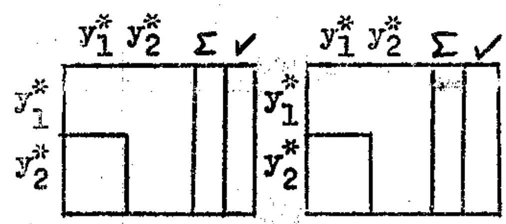
Doolittle Forward Solution



<u>Case II</u>: H = 2, $B_{y^*y^*}$ nonsingular 1/

$B_{y^*y^*} =$
$M_{y^*z}M_{zz}^{-1}M_{zy^*}$
$- M_{y^*z^*}P^{*\prime}$

$W_{y^*y^*}$



$$p_1 = |W| = W_{11}W_{22} - W_{12}^2$$

$$p_2 = W_{11}B_{22} + W_{22}B_{11} - 2W_{12}B_{12}$$

$$p_3 = |B| = B_{11}B_{22} - B_{12}^2$$

$$\lambda = \frac{1}{2p_3}\left(p_2 + \sqrt{p_2^2 - 4p_1p_3}\right)$$

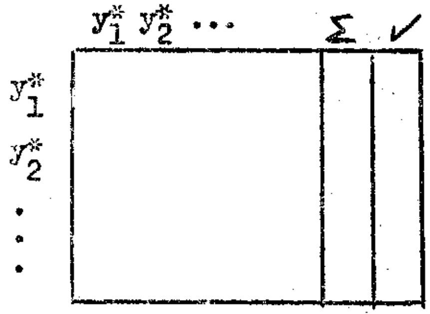
1/ A degenerate case occurs when $K^{**} = H - 1$. This is known as the just-identified case. It causes $B_{y^*y^*}$ to become singular, and requires a modification in procedure which is discussed in the Appendix, Note 1. Also if H = 1 the limited information method is equivalent to that of least squares.

2/ Check is Σ of $M_{y^*y^*}$ − Σ of $M_{y^*z^*}P^{*\prime}$.

Table II (continued)

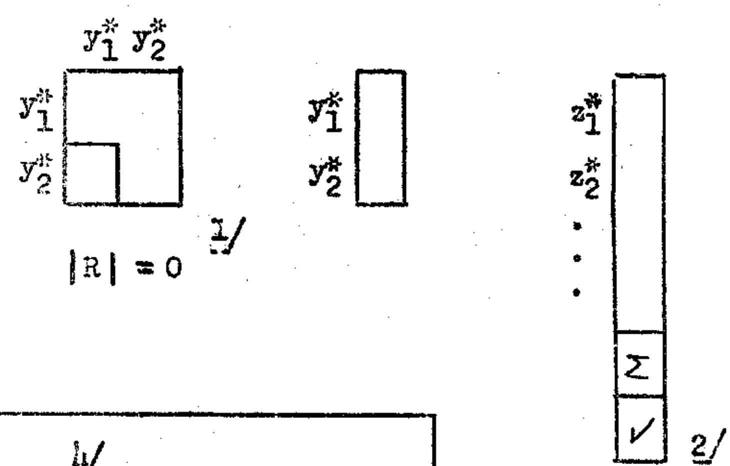
Case I (continued)

$A = B^{-1}_{y^*y^*}W_{y^*y^*}$

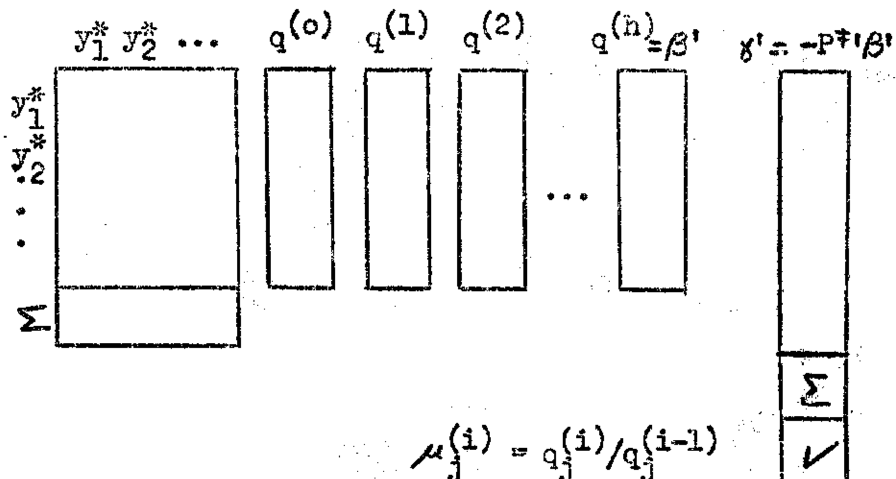


Case II (continued)

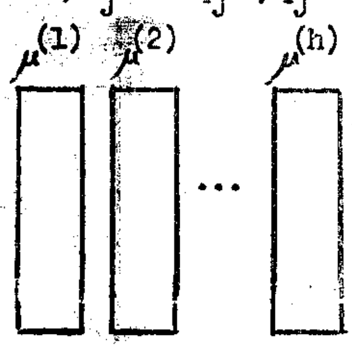
$R_{y^*y^*} = W_{y^*y^*} - \lambda B_{y^*y^*}$ $\qquad \beta' = \begin{pmatrix} R_{12} \\ -R_{11} \end{pmatrix}$ $\qquad \gamma' = -P^{\dagger\prime}\beta'$



A' 3/ $\qquad q^{(i)} = Aq^{(i-1)}, \quad i \neq 0$ 4/



$\mu_j^{(i)} = q_j^{(i)}/q_j^{(i-1)}$



In case convergence is slow, see the Appendix, Note 2.

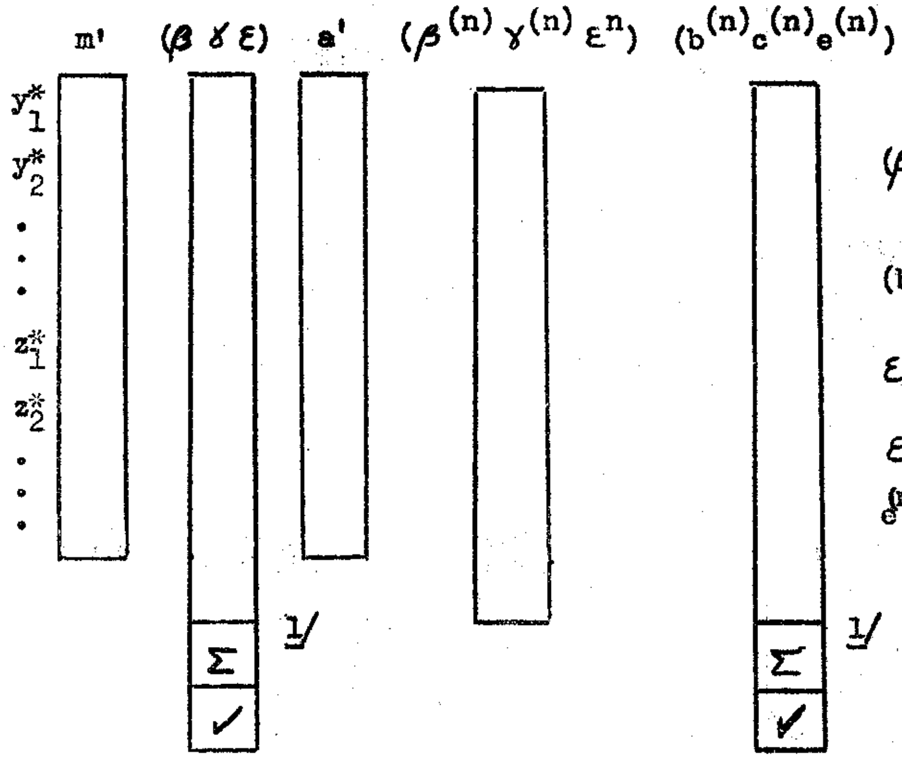
1/ This checks computation of λ as well as of R.

2/ The check here is obtained by using the column sums of $P^{\dagger\prime}$.

3/ A is rewritten here in transposed form to permit column by column multiplication in computing the $q^{(i)}$.

4/ $q^{(o)}$ = initial approximation.

Table II (continued)



$$(\beta^{(n)}\gamma^{(n)})_i = a_i(\beta\,\gamma)_i$$

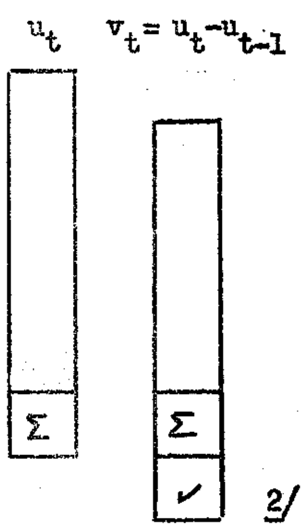
$$(b^{(n)}c^{(n)}) = \frac{1}{\beta_1^{(n)}}\,(\beta^{(n)}\gamma^{(n)})$$

$$\varepsilon = -(\beta\;\gamma)m'$$

$$\varepsilon^{(n)} = \varepsilon$$

$$e^{(n)} = \varepsilon/\beta_1^{(n)}$$

$$u_t = (b^{(n)}c^{(n)}e^{(n)})(y^{*(n)}z^{*(n)}1)' = \left[(\beta\;\gamma\;\varepsilon)(y^*z^*1)'\right]/\beta_1^{(n)}$$



$\sum_t u_t^2 = \square$ 3/

$\sum_t u_t u_{t-1} = \square$

$\sum_t v_t^2 = \square\;\checkmark$ 4/

$$s^2 = \frac{\sum_t v_t^2}{T - F} = \square$$

$s = \square$

$$\frac{\delta^2}{s^2} = \frac{\sum_t v_t^2}{\sum_t u_t^2}\left(\frac{T - F}{T - F - 1}\right) = \square$$

1/ Summation excluding e.

2/ $\sum_t v_t = u_T - u_1$

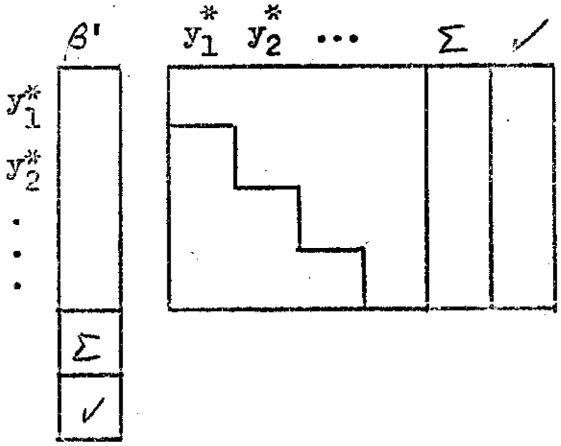
3/ It is sometimes convenient to compute this from the formula $\sum_t u_t^2 = (1 + 1/\lambda)(\beta W \beta')/T(\beta_1^{(n)})^2$. (See C^* on following page.)

4/ $\sum_t v_t^2 = 2\sum_t u_t^2 - 2\sum_t u_t u_{t-1} - u_1^2 - u_T^2$

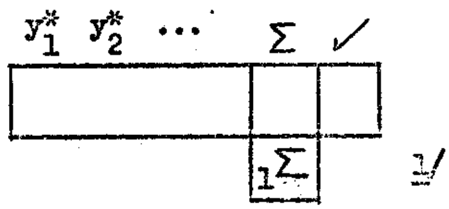
Table III

Computations for Covariances of Estimates for a Single Equation

$W_{y^*y^*}$

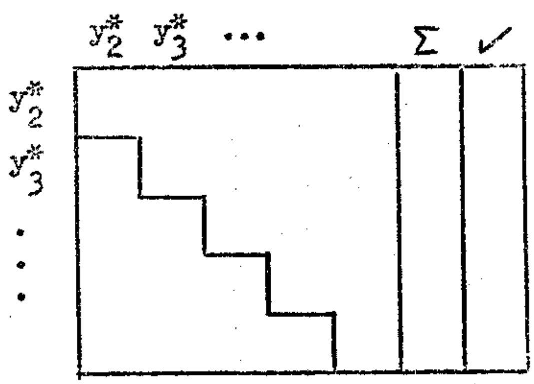


$\beta W_{y^*y^*}$



$\beta W_{y^*y^*} \beta' = \square$

$_{11}[(\beta W_{y^*y^*})'(\beta W_{y^*y^*})]$



1/ This is the summation omitting the first element. It is used in checking $_{11}[(\beta W_{y^*y^*})'(\beta W_{y^*y^*})]$.

Table III (continued)

λ	=	
$\nu = 1/\lambda$	=	
$1 + \nu$	=	
$\dfrac{\nu}{\beta W_{y^*y^*}\beta'}$	=	
$C = (1+\nu)\beta W_{y^*y^*}\beta'$	=	
a_1	=	
$C^* = \dfrac{C}{(T - F)(\beta_1 a_1)^2}$	=	

$_{11}B_{y^*y^*}$

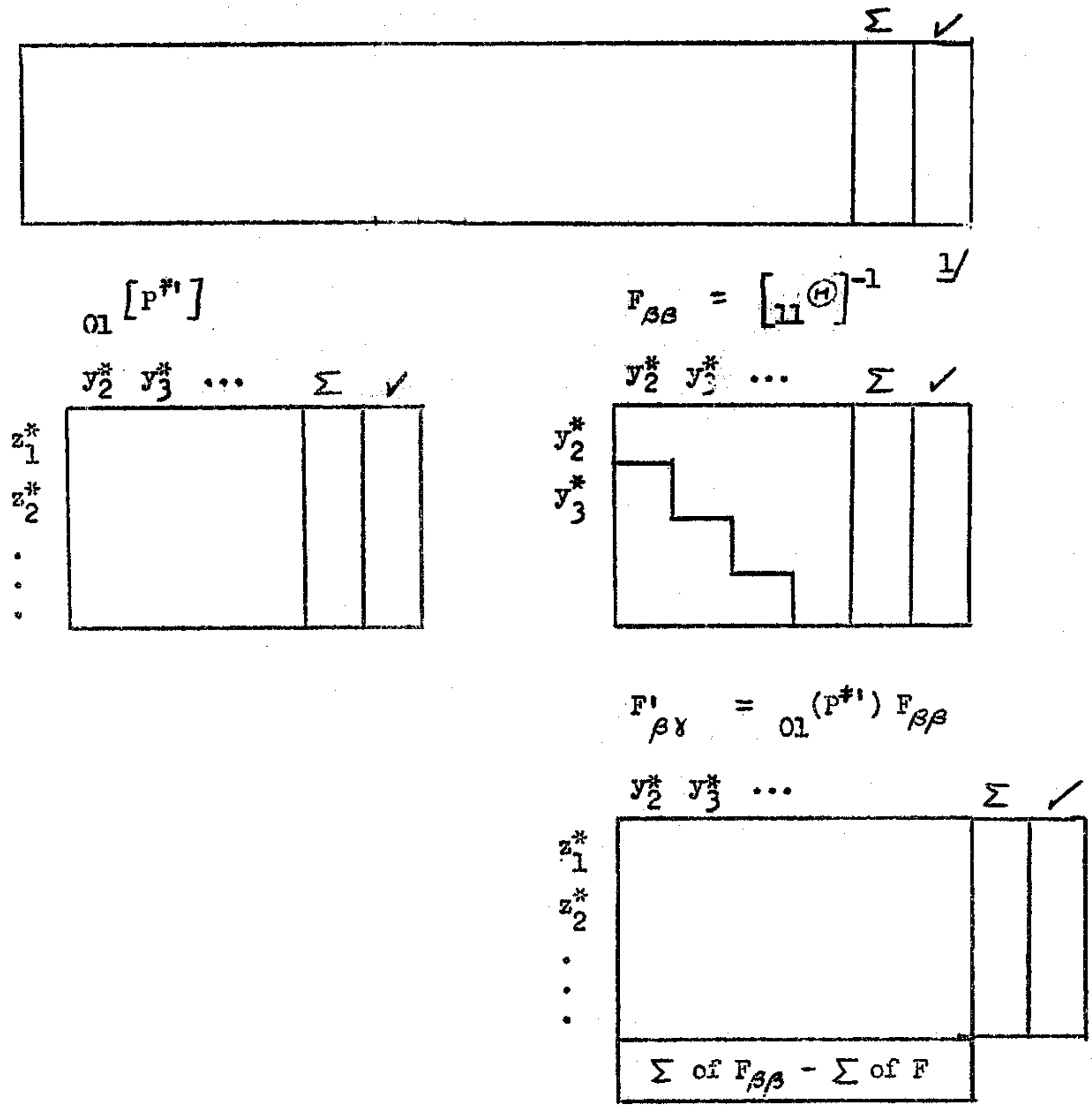
	y_2^*	y_3^*	...	Σ	✓
y_2^*					
y_3^*					
.					
.					

$$_{11}\Theta = {}_{11}\left[B_{y^*y^*} - \frac{\nu}{\beta W_{y^*y^*}\beta'}\,(W_{y^*y^*}\beta'\beta W_{y^*y^*})\right]$$

	$_{11}\Theta$: y_2^*	y_3^*	...	Σ	✓	I	Σ
y_2^*							
y_3^*							
.							
.							
.							

Table III (continued)

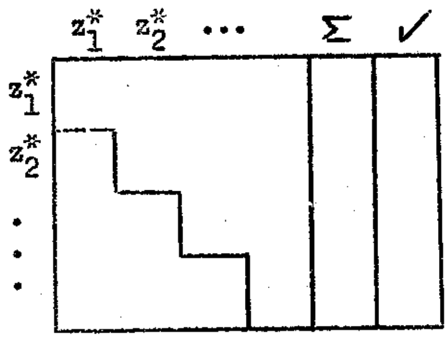
Doolittle Forward Solution



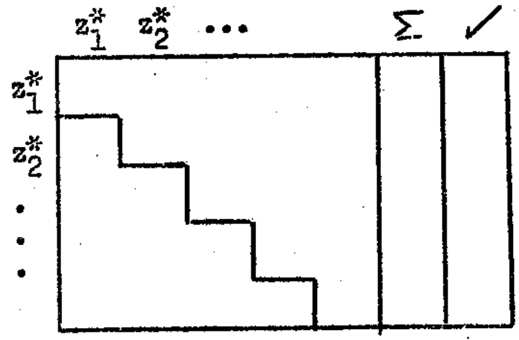
1/ The subscript β will, contrary to our usual notation, frequently represent a row corresponding to $(y_2^* \; y_3^* \ldots)$.

Table III (continued)

$M^{-1}_{z^*z^*}$



$U = {}_{01}(P^{*\prime})\, F_{\beta\gamma}$



$F_{\gamma\gamma} = U + M^{-1}_{z^*z^*}$

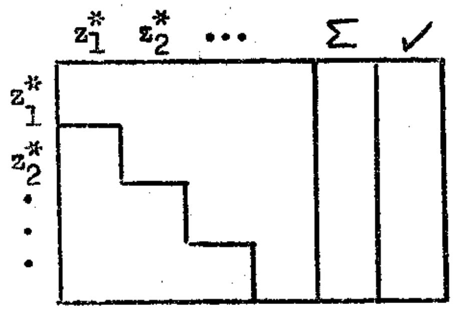
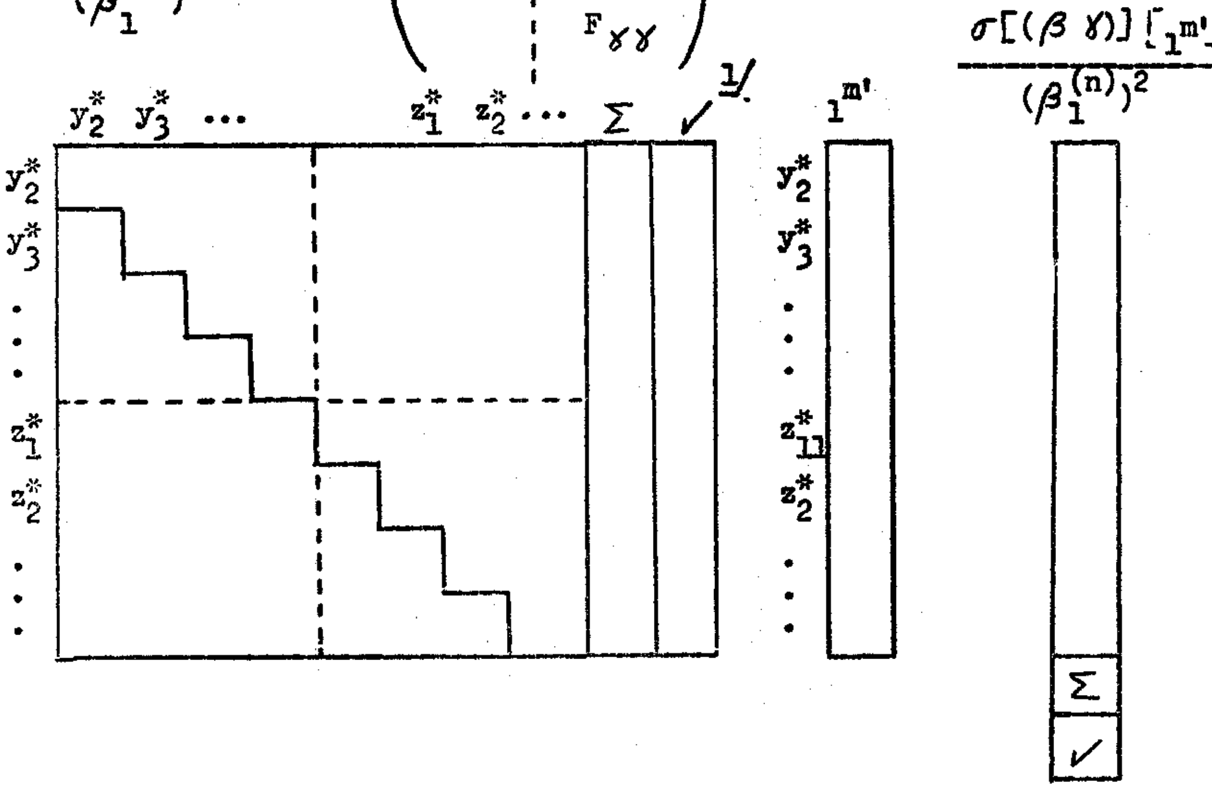


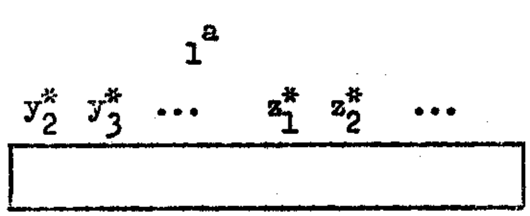
Table III (continued)

$$\frac{\sigma[(\beta\gamma)]}{(\beta_1^{(n)})^2} = C^* \left(\begin{array}{c|c} F_{\beta\beta} & -F_{\beta\gamma} \\ \hline & F_{\gamma\gamma} \end{array}\right)$$



$$[{}_1m] \; \frac{1}{(\beta_1^{(n)})^2} \; \sigma[(\beta\gamma)] \; [{}_1m'] = \square$$

$$\frac{\sum_t u_t^2}{T(T-F)} = \frac{C^*}{T^2} = \square$$

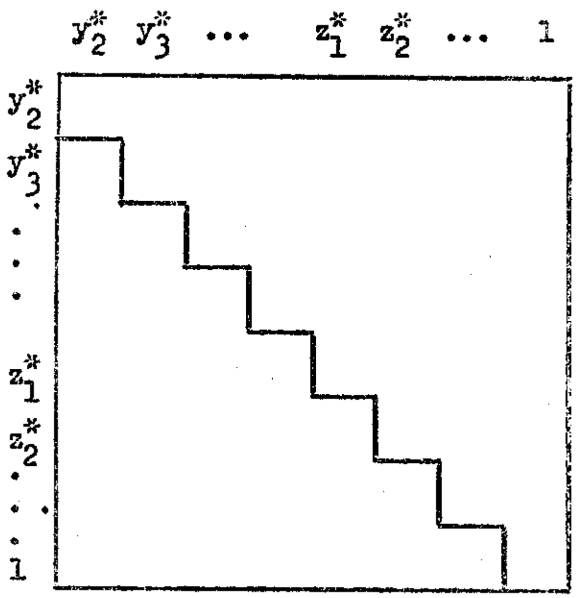


1/ Check for upper block is $C^*(\Sigma$ of $F_{\beta\beta} - \Sigma$ of $F_{\beta\gamma})$. Check for lower block is $C^*(\Sigma$ of $F_{\gamma\gamma} - \Sigma$ of $F'_{\beta\gamma})$.

Table III (continued)

$\sigma\left[(b\ c\ e)^{(n)}\right] =$

$$\begin{pmatrix} a_2 & 0 & \cdots & 0 \\ 0 & a_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & 1 \end{pmatrix} \left(\begin{array}{c|c} \dfrac{\sigma[(\beta\ \gamma)]}{(\beta_1^{(n)})^2} & -\dfrac{\sigma[(\beta\ \gamma)]\ [1^{m'}]}{(\beta_1^{(n)})^2} \\ \hline & \dfrac{[1^{m}]\ \sigma[(\beta\ \gamma)]\ [1^{m'}]}{(\beta_1^{(n)})^2} + \dfrac{C^*}{T^2}\ \text{1/} \end{array}\right) \begin{pmatrix} a_2 & 0 & \cdots & 0 \\ 0 & a_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & 1 \end{pmatrix} \text{ 2/}$$



1/ Our moments were defined by $M_{xx} = T\sum_t x_t'x_t - (\sum_t x_t)'(\sum_t x_t)$. If more conventional moments $\tilde{M}_{xx} = (1/T^2)M_{xx}$ were used, the only change in the computational formulae would be to replace C^*/T^2 by C^*. If $\tilde{M}_{xx} = (1/T)M_{xx}$ were used, C^*/T^2 would be replaced by C^*/T.

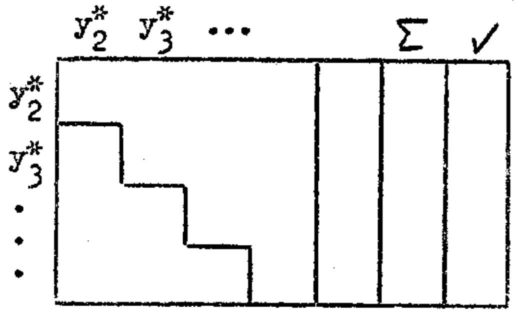
2/ The matrix composed of the adjustment factors is used to readjust a matrix. The last element is 1 because that corresponds to the constant term which has been tacked on.

Appendix

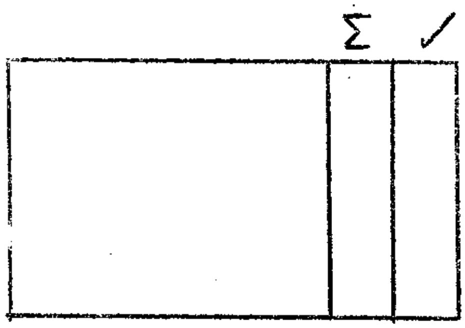
Note I: Procedure in Just-Identified Case ($K^{**} = H - 1$).

Let $B = M_{y^*z}M_{zz}^{-1}M_{zy^*} - M_{y^*z^*}P^{\ddagger\prime}$ as usual and let B_1 stand for the first column of B minus its last (H^{th}) element.

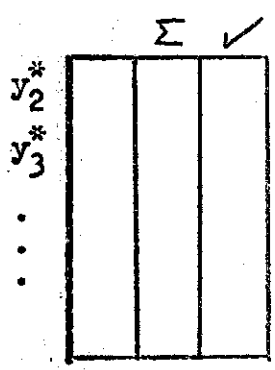
$${}_{11}(B_{y^*y^*}) \qquad - B_1$$

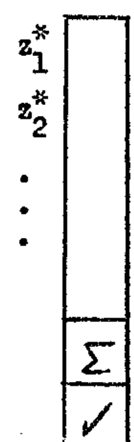


Doolittle Forward Solution



$${}_1\beta' = -\left[{}_{11}(B_{y^*y^*})\right]^{-1} B_1 \qquad\qquad \gamma' = -P^{\ddagger\prime}\begin{pmatrix}1\\ {}_1\beta'\end{pmatrix}$$





[Here $\beta = (1 \quad {}_1\beta)$ is subject to the normalization $\beta_1 = 1$ but refers to adjusted rather than unadjusted time series. It is therefore necessary to normalize again after obtaining $(\beta^{(n)} \gamma^{(n)})$ in order to find $(b^{(n)} c^{(n)})$].

It is also possible to obtain β directly from $M_{zz}^{-1} M_{zy} = P$. Let $P^{**'}$ be the H x K submatrix in the upper right-hand corner of P! Then if $\beta = (1 \quad {}_1\beta)$, we may obtain ${}_1\beta$ from the equation

$${}_1\beta \; {}_{10}(P^{**'}) = -(P_{11}^{**'} P_{12}^{**'} \ldots P_{1K^{**}}^{**'}).$$

Since $H - 1 = K^{**}$, the matrix ${}_{10}(P^{**'})$ is square (and so has rank $H - 1$ with probability 1). Thus there is a unique solution for ${}_1\beta$.

This procedure is not actually used because the matrix $P = M_{zz}^{-1} M_{zy}$ is not available without additional computing cost. While $M_{yz} M_{zz}^{-1} M_{zy}$ must always be computed, it is more convenient to do this directly without finding $M_{zz}^{-1} M_{zy}$ as an intermediate step.

Note II: Procedure in Case of Slow Convergence.

Two procedures may be applied to speed up convergence: (1) matrix squaring and (2) the Aitken acceleration technique.

(1) If we replace the matrix A by the matrix A^2 in obtaining the $q^{(i)}$, then each iteration will be as effective as two iterations using A. Whether or not it will be profitable to square A depends on the cost of squaring, which varies with the order H of A, and on the number of iterations which would otherwise be expected to be necessary. In general the squaring of A is advised whenever the number of additional iterations is expected to exceed $2(H + 1)$.

The number of additional iterations R may be estimated in the following way. Let i be the number of iterations already performed, k the number of decimal places of accuracy desired, and n the number of decimal places to which the elements of $\mu^{(i)}$ agree. Let λ_1 and λ_2 be the largest and second largest characteristic roots, respectively. Then

$$R \approx \frac{n - k}{\log_{10} \frac{\lambda_2}{\lambda_1}} \approx \frac{n - k}{\log_{10} \left(\frac{\mu^{i} - \mu^{i-1}}{\mu^{i-1} - \mu^{i-2}} \right)}$$

When matrix squaring is used, the final iteration should be made with A rather than A^2 as a check against possible mistakes in the squaring of A.

In cases of extremely slow convergence A^4 or higher powers of A might be used.

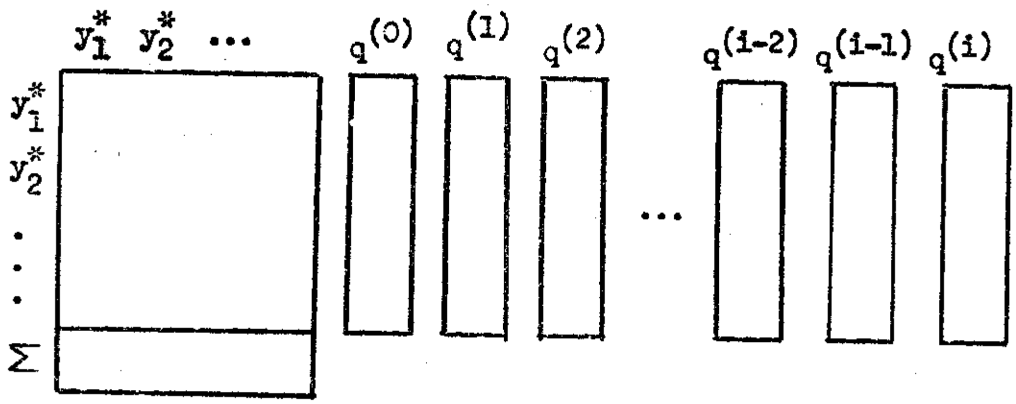
(2) In cases where the rate of convergence is too rapid to make matrix squaring profitable but is still slower than desired, the Aitken acceleration technique is sometimes useful.[1/] The procedure in this case is shown below.

1/ A. C. Aitken, "Studies in Practical Mathematics," Proc. Roy. Soc. Edinburgh, Vol. 57, p. 269.

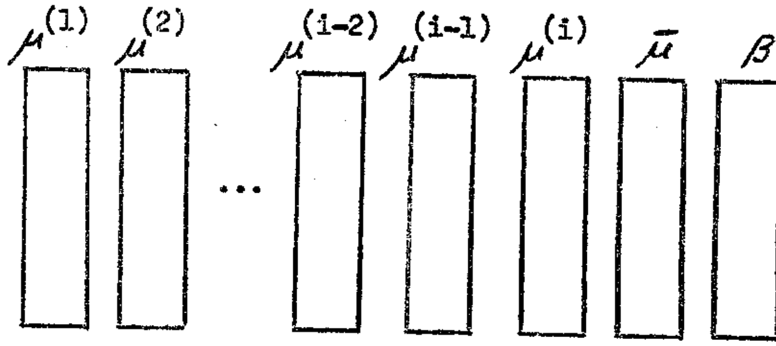
Note II (continued)

Appendix

A' $q^{(i)} = Aq^{(i-1)}, \quad i \neq 0$



$$\mu_j^{(i)} = q_j^{(i)}/q_j^{(i-1)}$$



$$\bar{\mu}_j = \frac{\mu_j^{(i)}\mu_j^{(i-2)} - (\mu_j^{(i-1)})^2}{\mu_j^{(i)} + \mu_j^{(i-2)} - 2\mu_j^{(i-1)}} \approx \lambda$$

$$\beta_j' = \frac{q_j^{(i)}q_j^{(i-2)} - (q^{(i-1)})^2}{\frac{1}{\lambda}q_j^{(i)} + \lambda q_j^{(i-2)} - 2q_j^{(i-1)}}$$

This technique may be applied when $\dfrac{\mu^{(i)} - \mu^{(i-1)}}{\mu^{(i-1)} - \mu^{(i-2)}}$ seems to be converging and when the elements of $\mu^{(i)}$ already agree to four or five significant figures. The effect is then to reduce the influence of the second highest characteristic value.

The procedure is very sensitive to rounding errors, and as many significant figures as possible must be retained.