

Hans-Jürg Büttler

1. Einleitung

Jeder empirisch arbeitende Oekonom, der von Hand, mittels bestehendem oder selbstgeschriebenem Computer-Programm eine Regressionsrechnung, im folgenden als Ausgleichsrechnung bezeichnet, durchführt, sollte sich Rechenschaft über das gewählte Rechenverfahren geben in bezug auf die Genauigkeit oder den Stellenverlust, auf die Rechenzeit und schliesslich für den Computer auf die Speicherplatzbedürfnisse. In diesem Aufsatz möchte ich zwei Rechenverfahren, nämlich dasjenige von CHOLESKY und der SCHMIDT'schen Orthogonalisierung vorstellen, die sich besonders für die einstufige (ordinary least squares), zweistufige (two-stage least squares) und dreistufige (three-stage least squares) Ausgleichsrechnung eignen, da sie sowohl dem Problem der Minimierung der Residuenquadrate angepasst sind als auch eine hohe Genauigkeit aufweisen, geringe Rechenzeit und wenig Speicherplatz benötigen.

Die ~~drei folgenden~~ Beispiele sollen die Eigenschaften der beiden Verfahren verdeutlichen. Anschliessend werde ich die beiden für die einstufige Ausgleichsrechnung entwickelten Verfahren vorstellen und ihre Erweiterung für die zwei- und dreistufige Ausgleichung zeigen. Die Schlussfolgerungen fassen alle Vorteile gegenüber den üblichen Inversionsverfahren zusammen.

2. Drei Beispiele

Erstes Beispiel:

Als erstes Beispiel dient das von KLEIN [1] aufgestellte Modell I für die Jahre 1921 bis 1941, für welches ich mit dem Verfahren von CHOLESKY, im folgenden VC abgekürzt, und der SCHMIDT'schen Orthogonalisierung, abgekürzt VSO, die dreistufige Ausgleichung (three-stage least squares estimation) durchgeführt habe. Zum Vergleich gebe ich die Ergebnisse von ROTHENBERG und LEENDERS [2]. Diejenigen von ZELLNER und THEIL [3] können leider nicht gebraucht werden, da, wie schon ROTHENBERG und LEENDERS [2] nachgewiesen haben, die Kovarianzmatrix der Residuen falsch ist. Die Bezeichnungen der Variablen sind diejenigen von GOLDBERGER [4].

Tafel I Dreistufige Ausgleichung für KLEINS Modell I

Gleichung	Variablen	Rothenberg und Leenders		Verf. von Cholesky*		Verf. der Orthogonalisierung*	
		Koeffizienten	Standardabweichungen	Koeff.	Stand.abw.	Koeff.	Stand.abw.
C	1	16.6913	1.3137	16.4408	1.3046	16.4408	1.3046
	P	0.1232	0.1082	0.1249	0.1081	0.1249	0.1081
	P ⁻¹	0.1695	0.1005	0.1631	0.1004	0.1631	0.1004
	W	0.7819	0.0387	0.7901	0.0379	0.7901	0.0379
I	1	34.0981	7.5257	28.1779	6.7938	28.1779	6.7938
	P	-0.1073	0.1700	-0.0131	0.1619	-0.0131	0.1619
	P ⁻¹	0.8343	0.1587	0.7557	0.1529	0.7557	0.1529
	K ₋₁	-0.2227	0.0361	-0.1949	0.0325	-0.1949	0.0325
W*	1	1.8149	1.1214	1.7972	1.1159	1.7972	1.1159
	E	0.4052	0.0316	0.4005	0.0318	0.4005	0.0318
	E ⁻¹	0.1757	0.0346	0.1813	0.0342	0.1813	0.0342
	A	0.1454	0.0283	0.1497	0.0279	0.1497	0.0279

* Die Ergebnisse des VC und VSO stimmen mit jenen überein, die von THEIL [13] auf vier Stellen veröffentlicht wurden.

ROTHENBERG und LEENDERS [2] haben die Rechnung mit ... Stellen von Hand nach dem Verfahren von durchgeführt. Der Stellenverlust für die Investitionsfunktion I betrug demzufolge mindestens ... Stellen, so dass kein Koeffizient numerisch gesichert ist. Das VC und VSO habe ich in PASCAL [5] programmiert und auf der CDC 6400 der ETH-Zürich, die über eine 14-stellige Mantisse verfügt, rechnen lassen. Der Stellenverlust für das VC betrug nur 3 Stellen, so dass die Koeffizienten und Standardabweichungen nach beiden Verfahren für 11 wesentliche Stellen übereinstimmen.

Zweites Beispiel:

Um die Kosten oder indirekt die Rechenzeit, die Genauigkeit und den Speicherbedarf der beiden Verfahren für ein grösseres Modell zu ermitteln, habe ich ein Vierteljahresmodell* mit 19 Gleichungen, davon 15 Verhaltensgleichungen, 25 vorbestimmten Variablen, einer Stichprobe von 36 Werten und 53 zu schätzenden Parametern mit der dreistufigen Ausgleichung gerechnet. Die Ergebnisse sind in Tafel 2 ersichtlich:

Tafel 2 Dreistufige Ausgleichung für ein Modell mit 19 Gleichungen und 53 zu schätzenden Parametern auf der CDC 6400

Merkmale	Cholesky	Orthogonalisierung
Rechenkosten	29.69 Fr.	91.29 Fr.
Rechenzeit einschliesslich Kompilationszeit des Pascal-Compilers	33.89 Sek.	73.86 Sek.
benötigte Feldlänge (oktal bzw. in K)	52!000=21 K	125!000≈43 K
grösster Stellenverlust	5	≤5
grösste Ordnung einer Matrix	53 x 53	375 x 53

Bei einer Rechnung auf einem IBM-Computer mit einfacher Genauigkeit, d. h. mit einer 7-stelligen Mantisse, wären die Ergebnisse nach dem VC nur für die ersten 2 Stellen numerisch gesichert.

Drittes Beispiel:

Um das Verhalten dieser beiden Verfahren gegenüber Multikollinearität zu prüfen, habe ich folgende Gleichung geschätzt:

$$Q = a_0 + a_1 \cdot Y + a_2 \cdot P$$

Dabei könnte Q für die Menge, P für den Preis und Y für das Einkommen stehen. Für die Variablen Q, Y und P habe ich eine Stichprobe von 15 fiktiven

* Es handelt sich um ein nominelles Vierteljahresmodell der Schweiz, das Studenten im Rahmen eines Seminars bei Frau Prof. Schelbert-Syfrig an der Universität Zürich erarbeitet haben.

Werten angenommen. Anfänglich wurden 9 Werte der Variablen Y und P genau linear abhängig gewählt, der Quotient betrug 1.25, für die restlichen 6 Werte wurde der Quotient leicht geändert. In den nächsten Schritten wurde die lineare Abhängigkeit zwischen Y und P ständig erhöht, bis schliesslich ein Wert den Quotienten 1.2499 und die restlichen Werte den Quotienten 1.25 hatten. Zusätzlich zum VC und VSO wurde die Rechnung noch mit der CRAMER-Regel durchgeführt. Tafel 3 zeigt das unterschiedliche Verhalten für die drei Verfahren:

Tafel 3 Rechnung mit CRAMER, VC und VSO auf der CDC 6400

Verfahren	A = Anzahl genau linear abhängiger Werte der Variablen Y und P S = grösster Stellenverlust					
	A	S	A	S	A	S
CRAMER	9	7	12	9	14	13
VC	9	5	12	7	14	10
VSO	9	<5	12	<5	14	<5

Die Ergebnisse bestätigen die theoretischen Ueberlegungen (SCHWARZ [6]), dass das VSO noch brauchbare Ergebnisse liefern kann, wenn alle anderen Verfahren, einschliesslich die üblichen Inversionsmethoden, versagen.

3. Die einstufige Ausgleichung

Gegeben sei die Regressionsgleichung

$$(1) \quad X \cdot b + y = r \quad , wo$$

X : Datenmatrix $[1..T, 1..K]$ der K Regressoren mit T Beobachtungen einschliesslich konstantes Glied. (Im Sinne der Programmiersprachen PASCAL wird die Ordnung der Matrix nicht $(T \times K)$ geschrieben, weil dadurch bei Teilmatrizen auch der Ort festgelegt werden kann.)

b : Vektor $[1..K]$ der K zu schätzenden Parameter.

y : Datenvektor $[1..T]$ des Regressanden

r : Residuenvektor $[1..T]$

Das GAUSSsche Prinzip verlangt die Minimierung des Quadrates der Länge des Residuenvektors r . Also:

$$(2) \quad * \quad r^T \cdot r \equiv (r, r) = (X \cdot b + y, X \cdot b + y) \longrightarrow \text{minimieren}$$

Nach Satz 2.1 von SCHWARZ [6] ist dies gleichbedeutend mit der Auflösung des symmetrisch-definiten Gleichungssystems:

* Ein hochgestelltes T bedeutet "transformiert".

$$(3) \quad (X^T \cdot X)b + X^T y = 0, \text{ oder abgekürzt}$$

$$(3a) \quad A \cdot b + c = 0.$$

3.1 Das Verfahren von CHOLESKY

Die üblichen Inversionsverfahren wie das Austauschschrittverfahren (STIEFEL [7]) oder jene nach GAUSS-JORDAN-RUTISHAUSER [8], nach GAUSS-SEIDEL und GAUSS-SOUTHWELL (ZURMÜHL [9]) sind nicht auf die speziellen Eigenschaften von **A** zugeschnitten. Diesen ist in bezug auf Genauigkeit, Rechenzeit und Speicherplatz das VC mindestens ebenbürtig. Das VC, für das man den Algorithmus bei ZURMÜHL [9] oder SCHWARZ [6] nachlesen kann, geht im Prinzip folgendermassen vor sich:

1. Die positiv-definite und symmetrische Matrix **A** wird durch einen expliziten Rechenvorgang in eine Rechtsdreiecksmatrix **R** zerlegt, so dass $A = R^T \cdot R$. Gl. (3a) wird dann zu

$$(4) \quad R^T(R \cdot b) + c = 0$$

Führt man den Hilfsvektor $h = R \cdot b$ ein, so kann

2. das Gleichungssystem (4) nacheinander aufgelöst werden. Zuerst

$$(5) \quad R^T \cdot h + c = 0$$

durch Vorwärtseinsetzen, und gewinnt **h** bei gegebenem **c**, sodann

$$(6) \quad R \cdot b - h = 0$$

durch Rückwärtseinsetzen, und gewinnt **b** bei vorher bestimmtem **h**.

Das VC hat auch den Vorteil, dass man die Varianzen der geschätzten Parameter direkt berechnen kann, ohne dass die ganze Kehrmatrix bekannt sein muss. Bekanntlich ist die Kovarianzmatrix der Parameter

$$(7) \quad S = s^2 \cdot (X^T X)^{-1} = s^2 \cdot A^{-1}, \text{ wo } s^2 \text{ die Varianz der Residuen.}$$

Die gesuchten Standardfehler sind die Wurzeln aus den Diagonalelementen. Diese können direkt berechnet werden. Es gilt:

$$(8) \quad A^{-1} = R^{-1}(R^{-1})^T$$

R^{-1} erhält man durch einen expliziten Prozess.

Mit $B \equiv A^{-1}$ und $W \equiv R^{-1}$ sind die Diagonalelemente von A^{-1} :

$$(9) \quad b_{kk} = \sum_{j=k}^K w_{kj}^2 \quad (k := 1..K)$$

Oft ist es angezeigt, auch bei Computer-Programmen eine Skalierung der beobachteten Werte vorzunehmen, um die Konditionszahl von **A** klein zu halten, d. h. den Stellenverlust zu verringern (SCHWARZ [6]).

3.2 Die SCHMIDT'sche Orthogonalisierung

Die Orthogonalisierung umgeht die GAUSS'schen Normalgleichungen (3a), zeichnet sich aber trotz mathematischer Äquivalenz zum VC durch eine grössere numerische Stabilität aus (SCHWARZ [6]). Das VSO, für welches man den Algorithmus bei ZURMÜHL [9] oder SCHWARZ [6] nachlesen kann, geht im Prinzip folgendermassen vor sich:

1. Die Datenmatrix X wird mittels des SCHMIDT'schen Orthogonalisierungsverfahrens zerlegt in

$$(10) \quad X = C \cdot R \quad , \text{ wo } C [1..T, 1..K] \text{ aus } K \text{ orthonormierten Vektoren } c_i \text{ besteht, so dass}$$

$$c_i^T \cdot c_j = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases} \quad (i, j := 1..K)$$

R ist eine Rechtsdreiecksmatrix. Wie SCHWARZ [6] zeigt, ist diese identisch mit der Rechtsdreiecksmatrix, die man mit dem VC aus $X^T \cdot X$ gewinnt.

2. Dann wird ein Vektor f berechnet, der sich bei der Orthogonalisierung des Regressanden bezüglich C ergibt:

$$(11) \quad f = C^T \cdot y \quad , \quad f: [1..K]$$

3. Schliesslich folgt die Lösung durch Rückwärtseinsetzen in

$$(12) \quad R \cdot b + f = 0$$

Die Residuen erhält man ohne zusätzlichen Aufwand aus

$$(13) \quad r = y - C \cdot f$$

4. Die zweistufige Ausgleichung

Die i -te Strukturgleichung ($i := 1..M$) eines Gleichungssystems sei

$$(14) \quad y_i = Y_i \gamma_i + X_i \beta_i + r_i \quad , \text{ wo}$$

y_i : Vektor $[1..T]$ des i -ten Regressanden.

Y_i : Datenmatrix $[1..T, 1..M_i^*]$ der endogenen Variablen, die rechts vom Gleichheitszeichen auftreten, mit $M_i^* < M$.

γ_i : Parametervektor $[1..M_i^*]$ der endogenen Variablen

X_i : Datenmatrix $[1..T, 1..K_i^*]$ der vorbestimmten Variablen, die rechts vom Gleichheitszeichen auftreten mit $K_i^* < K$.

β_i : Parametervektor $[1..K_i^*]$ der vorbestimmten Variablen

r_i : Residuenvektor $[1..T]$ der i -ten Strukturgleichung.

Zudem sei die notwendige Bedingung der Identifikation $K_i^{**} > M_i^*$ erfüllt, wo K_i^{**} die Anzahl der in der i -ten Strukturgleichung ausgelassenen vorbestimmten Variablen bedeutet. Das ganze Modell hat M Gleichungen mit $M+K$ Variablen, und es gilt gemäss Definition: $K = K_i^* + K_i^{**}$ und $M = M_i^* + M_i^{**} + 1$ für die i -te Gleichung.

Bekanntlich kann die zweistufige Ausgleichung auf drei Arten ermittelt werden:

- (a) Durch einfache Ausgleichung werden die endogenen Variablen auf die vorbestimmten Variablen regressiert (reduzierte Form). Mit der so geschätzten Datenmatrix \hat{Y}_i werden wiederum mittels einstufiger Ausgleichung die Parameter aus Gl. (14) geschätzt.
- (b) Die in (a) gewonnene Schätzung \hat{Y}_i wird zusammen mit den vorbestimmten Variablen X_i als Instrumentalvariablen für Gleichung (14) verwendet.
- (c) Gl. (14) wird mit X_i^T multipliziert. Auf diese neue Gleichung wird die verallgemeinerte einstufige Ausgleichung nach AITKEN angewandt. Dieses Verfahren leitet zur dreistufigen Ausgleichung über.

Mit $Z_i = [Y_i | X_i]$ erhält man die zu Gl. (3) bzw. (3a) analogen GAUSSschen Normalgleichungen (GOLDBERGER [4]):

$$(15) \quad [Z_i^T X (X^T X)^{-1} X^T Z_i] \cdot b_i - Z_i^T X (X^T X)^{-1} X^T y_i = 0 \quad \text{oder abgekürzt}$$

$$(15a) \quad A_i \cdot b_i - c_i = 0.$$

4.1 Verfahren von CHOLESKY

Aus Gleichung (15) ist ersichtlich, dass A_i in (15a) positiv definit* und symmetrisch ist, d. h. das VC ist anwendbar. Die zwei Stufen sind:

1. Stufe: Zerlegung $X^T X = R^T \cdot R$ und Bildung der reduzierten Form nach dem VC. Dies braucht für alle Strukturgleichungen nur ein Mal gemacht zu werden.
 $(X^T X)^{-1} X^T Y = R^{-1} (R^T)^{-1} X^T Y$
2. Stufe: Berechnung der Ausgangsform (15a) und Lösung mit dem VC.

Interessiert man sich für die Kovarianzmatrix der Koeffizienten, die man für die Kovarianzmatrix der Koeffizienten der abgeleiteten reduzierten Form (derived reduced form nach GOLDBERGER [4] oder solved reduced form nach CHRIST [10]) benötigt, so hat das VC den grossen Vorteil, dass die Lösungen der Systeme (15a) jeweils in der oberen Hälfte der Matrix A : array [I..AP, I..AP] of real (AP = Anzahl der zu schätzenden Koeffizienten) durchgeführt, während die Kehrmatrizen A_i^{-1} laufend in der unteren Hälfte von A gespeichert werden können. Für die Berechnung der Kovarianzmatrix der Koeffizienten für die ein- und zweistufige Ausgleichung siehe THEIL [11] und für jene der abgeleiteten reduzierten Form siehe GOLDBERGER/NAGAR/ODEH [12].

* Dies ist eine notwendige Bedingung für AITKENS Ausgleichung.

4.2 Orthogonalisierung

Mit der Zerlegung nach dem VSO kann geschrieben werden $\mathbf{X} = \mathbf{C} \cdot \mathbf{R}$ und damit $(\mathbf{X}^T \cdot \mathbf{X})^{-1} = \mathbf{R}^{-1} \cdot (\mathbf{C}^T)^{-1}$. Eingesetzt in Gl. (15) folgt:

$$(16) \quad (\mathbf{Z}_i^T \cdot \mathbf{X} \cdot \mathbf{R}^{-1}) ((\mathbf{R}^{-1})^T \cdot \mathbf{X}^T \cdot \mathbf{Z}_i) \cdot b_i - (\mathbf{Z}_i^T \cdot \mathbf{X} \cdot \mathbf{R}^{-1}) ((\mathbf{R}^{-1})^T \cdot \mathbf{X}^T \cdot \mathbf{y}_i) = 0$$

Die mit runden Klammern verdeutlichte Aufteilung zeigt nun, dass Gl. (16) als Normalform der Gleichung

$$(17) \quad [(\mathbf{R}^{-1})^T \cdot \mathbf{X}^T \cdot \mathbf{Z}_i] \cdot b_i - (\mathbf{R}^{-1})^T \cdot \mathbf{X}^T \cdot \mathbf{y}_i = (\mathbf{R}^{-1})^T \cdot \mathbf{X}^T \cdot \mathbf{r}_i \quad \text{oder abgekürzt}$$

$$(17a) \quad \hat{\mathbf{Z}}_i \cdot b_i - \hat{\mathbf{y}}_i = \hat{\mathbf{r}}_i \quad [1..K, 1..K_i^* + M_i^*]$$

aufgefasst werden kann, die als Ausgangsgleichung für die Orthogonalisierung dient. Die Orthogonalisierung geht auch in zwei Stufen vor sich:

1. Stufe: Orthonormierung der Datenmatrix \mathbf{X} [1..T, 1..K] aller vorbestimmten Variablen. Dies braucht für alle Strukturgleichungen nur ein Mal gemacht zu werden. (Analog zur reduzierten Form)
2. Stufe: Berechnung der Ausgangsform (17) für die i-te Strukturgleichung und Lösung mit dem VSO.

4.3 Eigenschaften der zweistufigen Ausgleichung

Die Praktiker haben festgestellt, dass sich die zweistufige Schätzung kaum merklich von der einstufigen unterscheidet, wenn die Anzahl der vorbestimmten Variablen gross ist. Dies hat meines Erachtens zwei Gründe:

1. Durch die kleinen Stichproben wird der Freiheitsgrad für die reduzierte Form sehr klein, so dass $\hat{\mathbf{Y}} \rightarrow \mathbf{Y}$. Exakte Konvergenz tritt dann ein, wenn die Anzahl der vorbestimmten Variablen gleich dem Stichprobenumfang ist; d. h. $K = T$.
2. Wie ein Test gezeigt hat, gibt die Regression beliebiger Zufallsfolgen erstaunlich hohe Korrelationskoeffizienten bei "hohen" Freiheitsgraden, deshalb ist für die reduzierte Form mit kleinen Abweichungen von $\hat{\mathbf{Y}}$ zu \mathbf{Y} zu rechnen, wo vielfach durch gelagte endogene Grössen starke Multikollinearität besteht. Das Ergebnis des Zufallszahlentests ist in Abbildung 1 dargestellt:

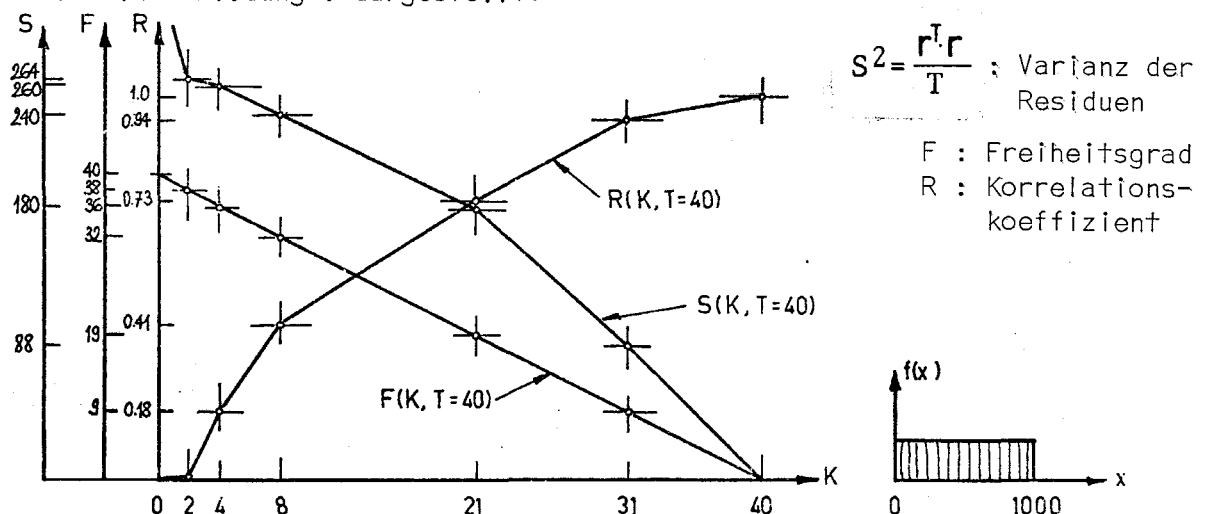


Abb. 1: Zufallszahlen mit gleichmässiger Dichtefunktion $f(x)$. K = Anzahl Regressoren, Stichprobenumfang $T = 40$.

5. Die dreistufige Ausgleichung

Bekanntlich wird in der dreistufigen Ausgleichung das Modell gleichzeitig geschätzt. Die zu (14) analoge Form ist:

$$(18) \quad y = Z \cdot b + r \quad , \text{ mit}$$

$$y^T = [y_1^T, \dots, y_i^T, \dots, y_M^T] \quad : \text{Datenmatrix } [1..M \cdot T] \text{ der } M \text{ Regressanden.}$$

$$Z = \begin{bmatrix} Z_1 & & 0 \\ & \ddots & \\ 0 & & Z_M \end{bmatrix} [1..M \cdot T, 1.. \sum_{i=1}^M (M_i^* + K_i^*)] \quad \text{um die Diagonale gegliederte Datenmatrizen } Z_i \text{ der } M \text{ Strukturgleichungen.}$$

$$b^T = [b_1^T, \dots, b_i^T, \dots, b_M^T] \quad : \text{Parametervektor } [1.. \sum_{i=1}^M (M_i^* + K_i^*)]$$

$$r^T = [r_1^T, \dots, r_i^T, \dots, r_M^T] \quad : \text{Residuenvektor } [1..M \cdot T]$$

Die zu Gl. (3) bzw. (15) analogen GAUSSschen Normalgleichungen sind (GOLDBERGER [4]):

$$(19) \quad (Z^{*T} [\hat{\Sigma}^{-1} \otimes (X^T X)^{-1}] Z^*) \cdot b - Z^{*T} [\hat{\Sigma}^{-1} \otimes (X^T X)^{-1}] y^* = 0 \quad , \text{ wo}$$

$$\Xi = \begin{bmatrix} X & & 0 \\ & \ddots & \\ 0 & & X \end{bmatrix} \quad , \quad Z^* = \Xi^T Z \quad , \quad y^* = \Xi^T y \quad , \quad r^* = \Xi^T r \quad \text{und}$$

$\hat{\Sigma}$ die Kovarianzmatrix der Residuen aus der zweistufigen Ausgleichung bedeuten.

In abgekürzter Schreibweise:

$$(19a) \quad A \cdot b - c = 0$$

Gl. (19) wurde mittels ALTKENS verallgemeinerter Ausgleichung gewonnen, d. h. notwendigerweise muss $\Omega^{-1} \equiv \hat{\Sigma}^{-1} \otimes (X^T X)^{-1}$ positiv definit sein. Diese Eigenschaft von Ω^{-1} bzw. Ω kann mit folgender Zerlegung gezeigt werden:

$$(20) \quad \Omega = D^T D = (\Xi^T \tilde{E}^T) (\tilde{E} \Xi) = \Xi^T (\hat{\Sigma} \otimes I) \Xi = \hat{\Sigma} \otimes (X^T X)$$

Es bedeuten:

$$\Xi : [1..M \cdot T, 1..M \cdot K] \quad \text{s. Gl. (19)}$$

$$\tilde{E} : [1..T \cdot T, 1..M \cdot T]$$

$$I : [1..T, 1..T] \quad \text{Einheitsmatrix}$$

$$D : [1..T \cdot T, 1..M \cdot K]$$

und

Gleichung (22) kann analog zu Gl. (16) als Normalform der Gleichung

$$(23) \quad [(R^{-1})^T \Xi^T Z] \cdot b - (R^{-1})^T \Xi^T y = (R^{-1})^T \Xi^T r \quad \text{oder abgekürzt}$$

$$(23a) \quad \tilde{Z} \cdot b - \tilde{y} = \tilde{r} \quad , \quad \tilde{Z}: [1..M \cdot K, 1.. \sum_{i=1}^M (M_i^* + K_i^*)]$$

aufgefasst werden, die als Ausgangsgleichung für die Orthogonalisierung dient.

D in (20) erfordert bei grösseren Modellen einen zu hohen Speicherplatzbedarf, der für das zweite Beispiel durch die CDC 6400 nicht mehr gedeckt werden konnte. Wenn man die spezielle Form von **A** in (19a) und von \tilde{Z} in (23a) ausnützt, kann mit der Ausgangsgleichung

Ausgangsgleichung

$$(23b) \quad H \cdot b - d = \bar{r} \quad , \quad H: [1..M \cdot K, 1.. \sum_{i=1}^M (M_i^* + K_i^*)]$$

der Speicherplatzbedarf erheblich reduziert werden.

H hat die Form:

$$H = \begin{array}{|c|c|c|c|} \hline \hat{r}_{11} \hat{z}_1 & 0 & \dots & 0 \\ \hline \hat{r}_{21} \hat{z}_1 & \hat{r}_{22} \hat{z}_2 & \dots & 0 \\ \hline \vdots & \vdots & \dots & \vdots \\ \hline \hat{r}_{M1} \hat{z}_1 & \hat{r}_{M2} \hat{z}_2 & \dots & \hat{r}_{Mk} \hat{z}_M \\ \hline \end{array} \begin{array}{l} K \\ 2 \cdot K \\ (M-1) \cdot K \\ M \cdot K \end{array}$$

\hat{z}_i : siehe Gl. (17a) ($i:=1..M$)

\hat{r}_{ik} sind die Elemente aus der Rechtsdreiecksmatrix $(\hat{R}^{-1})^T$, die bestimmt wird aus:

$$\hat{\Sigma} = \hat{E}^T \cdot \hat{E} \quad , \quad \text{mit} \quad \hat{E} = \frac{1}{\sqrt{r}} E$$

Orthogonalisierung von $\hat{\mathbf{E}} [1..T, 1..M]$ gibt:

$$\hat{\mathbf{E}} = \mathbf{C} \cdot \hat{\mathbf{R}} \quad \text{und daraus} \quad \hat{\Sigma}^{-1} = \hat{\mathbf{R}}^{-1} \cdot (\hat{\mathbf{R}}^{-1})^T$$

\mathbf{d} hat die Form:

$$\mathbf{d} = \begin{array}{|l} \hat{r}_{11} \cdot \hat{y}_1 \\ \hline \hat{r}_{21} \cdot \hat{y}_1 + \hat{r}_{22} \cdot \hat{y}_2 \\ \hline \vdots \\ \hline \hat{r}_{M1} \cdot \hat{y}_1 + \hat{r}_{M2} \cdot \hat{y}_2 + \dots + \hat{r}_{MM} \cdot \hat{y}_M \end{array} \begin{array}{l} k \\ 2 \cdot k \\ \\ (M-1) \cdot k \\ M \cdot k \end{array}$$

\hat{y}_i : siehe Gl. (17a) ($i:=1..M$)

Einfache Rechnung zeigt, dass tatsächlich $\mathbf{A} = \mathbf{H}^T \cdot \mathbf{H}$ und $\mathbf{c} = \mathbf{H}^T \cdot \mathbf{d}$. Der Residuenvektor $\bar{\mathbf{r}}$ in (23b) ist gleich wie \mathbf{d} aufgebaut, nur dass anstelle der \hat{y}_i die \hat{r}_i stehen. Die drei Stufen sind:

1. und 2. Stufe: wie unter 4.2 beschrieben
3. Stufe: Orthogonalisierung von $\hat{\mathbf{E}}$ und Berechnung von $(\hat{\mathbf{R}}^{-1})^T$. Bestimmung der Ausgangsform (23b) und Lösung mit dem VSO.

6. Schlussfolgerungen

6.1 Genauigkeit

Das VC und VSO sind den üblichen Lösungsverfahren in bezug auf die Genauigkeit mindestens ebenbürtig. Der Stellenverlust dürfte auch bei Computern eine Rolle spielen, wo zum Beispiel nur mit einer 7-stelligen Mantisse gerechnet wird. Die Ergebnisse des zweiten Beispiels wären dann für das VC nur noch für die ersten **zwei** Stellen numerisch gesichert, nach einem üblichen Verfahren möglicherweise für weniger als **zwei** Stellen. Das VSO dürfte dann noch brauchbare Ergebnisse liefern, wenn die anderen Verfahren versagen, insbesondere bei Multikollinearität.

6.2 Rechenzeit

In bezug auf die Schnelligkeit ist das VC den anderen Verfahren überlegen, weil es die Symmetrie ausnützt.

6.3 Speicherplatz

Der Bedarf an Speicherplatz ist beim VC geringer als beim VS0. Interessiert man sich für die Kovarianzmatrix der Koeffizienten der abgeleiteten reduzierten Form, so bietet das VC eine zusätzliche Ersparnis an Speicherplatz.

Literaturnachweis

- [1] L. R. Klein:
"Economic Fluctuations in the United States, 1921 - 1941", Cowles Commission Monograph II, John Wiley & Sons Inc., New York 1950.
- [2] T. J. Rothenberg, C. T. Leenders:
"Efficient Estimation of Simultaneous Equation Systems",
Econometrica Vol. 32, January-April 1964, S. 57 - 76.
- [3] A. Zellner, H. Theil:
"Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations",
Econometrica Vol. 30, January 1962, S. 54 - 78.
- [4] A. S. Goldberger:
"Econometric Theory", John Wiley & Sons Inc., New York 1964.
- [5] N. Wirth:
"The Programming Language Pascal", Acta informatica I, pp. 35 - 63,
Springer 1971 oder in Berichte der Fachgruppe Computer-Wissenschaften,
ETH, November 1970, Heft 1.
- [6] H. R. Schwarz, H. Rutishauser, E. Stiefel:
"Numerik symmetrischer Matrizen", B. G. Teubner, Stuttgart 1968.
- [7] E. Stiefel:
"Einführung in die numerische Mathematik", B. G. Teubner, Stuttgart
1965.
- [8] "Programm GJR", Programmsammlung des ETH-Rechenzentrums.
- [9] Rudolf Zurmühl:
"Matrizen", Springer, Berlin 1961.
- [10] Carl F. Christ:
"Econometric Models and Methods", John Wiley & Sons Inc., New York 1966.
- [11] H. Theil:
"Economic Forecasts and Policy", North-Holland, Amsterdam 1959.
- [12] A. S. Goldberger, A. L. Nagar, H. S. Odeh:
"The Covariance Matrices of Reduced-form Coefficients and of Forecasts
for a Structural Econometric Model", Econometrica, Vol. 29, S. 556 -
573, October 1961.
- [13] H. Theil:
"Principles of Econometrics", North-Holland Publishing Company, Amsterdam
1971.