

Das Programmsystem

U N I R U N

Eine Darstellung seiner Methodik

von Edwin DEUTSCH

Forschungsbericht Nr. 68

Juni 1972

INHALT

1. Einleitung
2. Der inhaltliche Rahmen
3. Die Programmstruktur
4. Die Programme und ihre methodischen Grundlagen
 - 4.1. UNIRUN
 - 4.2. UNICOM
 - 4.3. UNIMAX
 - 4.4. UNIOLS
 - 4.5. UNIORF
 - 4.6. UNIGEN
 - 4.7. Die Lagprogramme
 - 4.7.1. UNILAG
 - 4.7.2. UNIFAG
 - 4.7.3. Die Interpretation von Lagstrukturen
 - 4.8. UNISIM
 - 4.8.1. Einleitung und Methode
 - 4.8.2. Aussagemöglichkeiten von UNISIM
5. Weitere Entwicklungen

1. Einleitung

Das Programmsystem UNIRUN stellt in seiner Gesamtheit einen Modellfall dar, methodische Überlegungen mit den Anforderungen der Datenverarbeitung in Einklang zu bringen. UNIRUN verarbeitet ökonomische Methoden, und hat zum Ziel, einzelne Regressionsgleichungen nach verschiedenen Kriterien zu testen. Für den Ökonomen können diese Kriterien nicht rein statistischer oder formaler Natur sein; er muss entscheiden, welche Inhalte er der Struktur der Gleichungen zugrundelegt. Ebenso wichtig ist jedoch auch die Beurteilung nach statistischen Kriterien.

Das Programmsystem UNIRUN ist nun so gestaltet, daß der Benutzer in einem Arbeitsgang seine Hypothesen über die Struktur von ökonomischen oder anderen Zusammenhängen am statistischen Material prüfen und sich über die Qualität der Ergebnisse informieren kann. Die Möglichkeiten welche das System UNIRUN in dieser Hinsicht bietet, werden im 2. Abschnitt dargelegt und im 4. Abschnitt im einzelnen behandelt. Die Darstellung der dabei verwendeten Methoden bildet die Hauptaufgabe der vorliegenden Arbeit. Deshalb werden auch Ergebnisse diskutiert und - vom Standpunkt des Verfassers - sinnvolle Interpretationen erörtert.

UNIRUN soll auch den Anforderungen der Datenverarbeitung gerecht werden. Es war ein Ziel, die Programme möglichst übersichtlich und flexibel zu gestalten. Darum zeichnet sich das System durch einen hierarchischen Aufbau der Unterprogramme aus, die sich wie Bausteine zu ganzen Programmen zusammenfügen. Entsprechende Haupt- und Unterprogramme besitzen die gleiche Struktur. Die Eingabe der Befehle und Daten ist genormt. Die Grundzüge der Struktur werden im Abschnitt 3 erläutert.

Die Programme des Systems sind vorerst in Fortran II geschrieben, weil an der bestehenden Anlage des IHS keine andere Sprache verfügbar ist. Die Anwendbarkeit wird dadurch jedoch nicht eingeschränkt, weil die Programme leicht umzuprogrammieren bzw. zu ersetzen sind. Außerdem wird das System in naher Zukunft in Fortran V umgeschrieben und auf der neuen Großrechenanlage des Wirtschafts- und Sozialwissenschaftlichen Rechenzentrums verwendbar sein. Damit werden auch einige Beschränkungen aufgehoben, denen das System derzeit noch

unterliegt, weiters sollen neue Entwicklungen berücksichtigt werden. Aus begreiflichen Gründen wurde deshalb im Rahmen dieser Arbeit auf eine Dokumentation verzichtet und nur die Methode des Systems beschrieben. Nur zur Orientierung sind die Namen der Programme angegeben, welche die Programmbeschreibungen enthalten.

Ich möchte an dieser Stelle einigen Mitgliedern des Instituts für Höhere Studien, im besonderen Herrn Erich Wang (dst. Institut für Wirtschaftsforschung), von dem einige Subroutinenstammen, für ihre Diskussion und Vorarbeit danken.

2. Der inhaltliche Rahmen

Wir betrachten vorerst eine Menge M von Zeitreihen $\{x_{it}\}$,
 $i = 1, 2, \dots, m$ und $t = 1, 2, \dots, T$.

Wir wollen nun eine Zeitreihe y_t mit den vorhandenen Zeitreihen x_{it}
 erklären und suchen nach einer Gleichung

$$y_t = \sum_{i=1}^n \beta_i x_{n(i)t} + \varepsilon_t \quad (2.1.)$$

Dabei stellen die Indizes $n(i)$ eine Auswahl der Indizes $i = 1, 2, \dots, m$
 dar. Die Schätzwerte für β bezeichnen wir mit $\hat{\beta}_i$, die errechneten
 Residuen mit e_t .

Der Ausdruck $n(i)$ darf natürlich nicht mit der Anzahl n von erklärenden
 Variablen verwechselt werden.

Wir fordern nun für die Regressionsgleichung (1) "optimale" Eigen-
 schaften, d. h. sie soll

- a) ökonomisch ausreichend interpretierbar sein,
- b) statistisch ausreichend gesichert sein.

Die Programme von UNIRUN sollen den Benutzer in die Lage ver-
 setzen, auf folgende Fragen eine zufriedenstellende Antwort zu erhalten:

- a) Welche Teilmenge ECM gibt die richtige Auswahl an
 erklärenden Variablen wieder, d. h. stellt die Gleichung

$$y_t = \sum \beta_i x_{n(i)t} + \varepsilon_t$$

eine richtige Struktur dar, mit $E = \{x_{n(i)t}\}$.

Die kombinatorischen Programme UNICOM und UNIMAX
 liefern eine erste Auswahl vertretbarer Strukturen.

- b) Die Berechnung einer einzelnen Gleichung erfolgt mit UNIOLS,
 wenn man mit Daten in Abweichungsform rechnen will.
- c) Das Programm UNIORF ermittelt die Regressionsergebnisse
 mit den Daten in Originalform, es ist auch eine Unter-
 drückung der Konstante möglich.
- d) Bei a-priori Annahmen über das Zufallglied verwenden wir die
 verallgemeinerte Regressionsmethode in UNIGEN.

- e) Die Untersuchung einer Regressionsgleichung nach vorhandenen Lags erfolgt mit UNILAG und UNIPAG.
- f) Den Effekt einer Störung der Daten auf die Koeffizienten untersuchen wir mit UNISIM.

Die vorhandenen Programme dienen also im speziellen dazu, um ein genaues Bild vom Verhalten der Koeffizienten einer Regressionsgleichung zu erhalten. Denn die Konfidenzintervalle und Schwankungsbreiten der Koeffizienten hängen ab

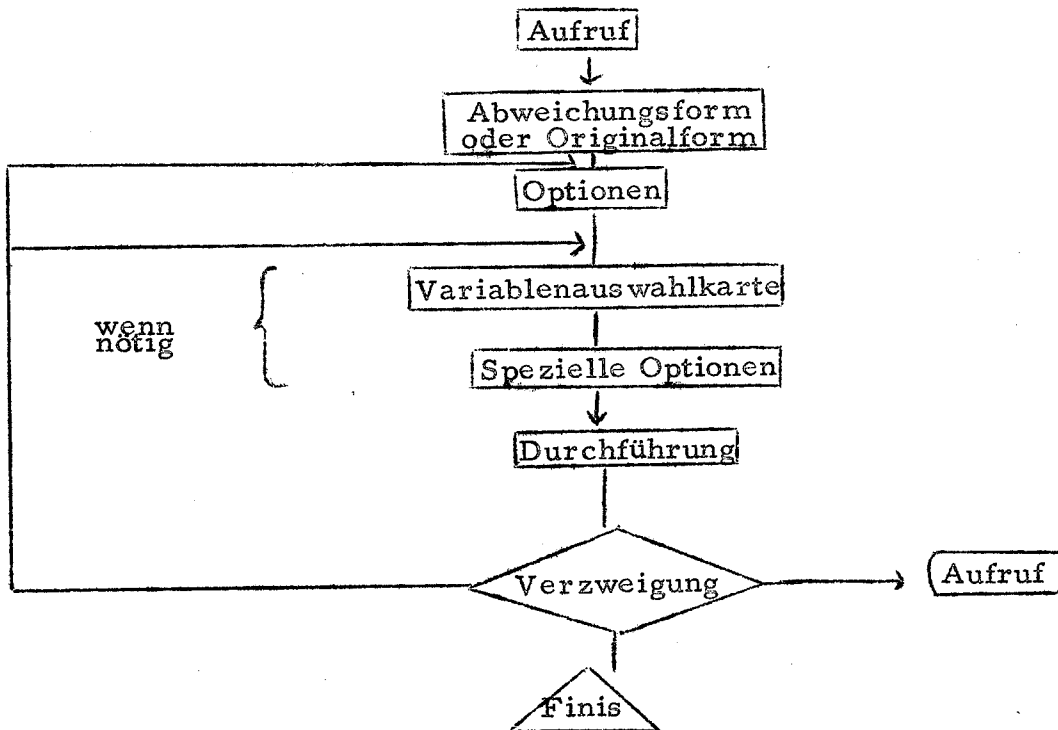
- von der Verteilung des Zufallsgliedes
- von vorhandenen Lags
- von der Güte des Datenmaterials.

Im allgemeinen sind die Schätzungen nicht einmal erwartungstreu. Im Abschnitt 4.8. werden die Bedingungen untersucht, unter denen bei Einzelgleichungen die Koeffizienten nicht mehr erwartungstreue Schätzfunktionen sind.

Auch können solche Schätzverzerrungen nicht beseitigt werden, deren Lösung der Theorie der simultanen Schätzverfahren vorbehalten bleibt. Das Ziel der vorliegenden Arbeit ist jedoch die optimale Schätzung von Einzelgleichungen. Eine Kenntnis der Struktur von Einzelgleichungen ist auf jeden Fall nützlich. Auf diese Weise mag das System UNIRUN die Ergebnisse verbessern helfen, trotz aller Beschränkungen, denen die Theorie der Ökonometrie unterliegt.

3. Die Programmstruktur

Der Titel des Systems ist zugleich der Name des Programmes, welches allen anderen vorangeht: UNIRUN übernimmt die Daten vom Datenträger (Disk oder Karten) und transformiert sie auf gewünschte Weise. Die anderen Programme übernehmen diese Daten. Sie haben alle dieselbe Struktur:



Abweichungsform oder Originalform:

Je nach Programm wird mit Daten in Abweichungsform

$$x'_{it} = x_{it} - \frac{1}{T} \sum_{T=1}^T x_{it} = x_{it} - \bar{x}_i$$

oder mit Daten in Originalform

$$x_{it} \text{ bzw. } x_{it} = x'_{it} + \bar{x}_i$$

gerechnet.

Optionen

Jedem Programm wird eine Liste von Optionen zugeteilt, die in Unterprogrammen mit den Namen COMOPT, LAGOPT, MAXOPT, etc. aufgerufen werden. Eine Option besteht aus einem Kennwort und nachfolgenden Worten bzw. Parametern. Das Ende der Optionen bildet immer das Kennwort ENDOPT.

Variablenauswahlkarte

Diese Karte trifft eine Auswahl aus der vorliegenden Grundmenge M von Variablen:

$$M \longrightarrow E, \text{ bzw. } \{x_{it}\} \longrightarrow \{x_{n(j)}, t\}$$

mit $i = 1, 2, \dots, m$ $j = 1, 2, \dots, n.$

Die Nummern i werden den Variablen in der Reihenfolge des Einlesens in UNIRUN zugeteilt. Die Variablenauswahlkarte enthält eine Eins-Lochung in jeder Spalte, deren Nummer der Nummer der auszuwählenden Variablen entspricht.

Spezielle Optionen

Manche Programme lassen für jede Teilmenge E noch eine Folge von Optionen zu. In diesen Fällen ist die Beschreibung im entsprechenden Unterprogramm genau zu beachten. Der Name des Unterprogrammes steht im Kommentar zum Hauptprogramm verzeichnet.

Verzweigung

Das Unterprogramm SWITCH übernimmt eine Karte, welche den weiteren Ablauf des Programmes bestimmt. Diese Karte ist wahlweise

- NEUOPT führt zu neuer Option
Variablenauswahlkarte
- UNICOM, UNIGEN, etc... führt zu neuen Programmen
- FINIS beendet den gesamten Programm-
ablauf.

Das Programmsystem UNIRUN unterliegt leider derzeit noch einigen Beschränkungen, die in der Neufassung in FORTRAN V weitgehend gelockert werden.

Einschließlich der endogenen Variablen sind überall höchstens 30 (in UNIORF höchstens 29) Variable zugelassen.

Die Länge der transformierten Zeitreihe, die Residuenlänge, darf höchstens 150 betragen. Das Produkt (Anzahl der Zeitreihen) x (Residuenlänge) darf höchstens 1000 betragen.

In den Teilmengen der erklärenden Variablen dürfen höchstens 9
(in UNIGEN und in UNIORF außer bei Konstantenunterdrückung je 8)
Variable aufscheinen.

Die Parameter und Felder behalten überall gleiche Namen. Die
Namen der Subroutinen stehen zum Teil in Beziehung zum überge-
ordneten Programmnamen.

Die Programme sind einheitlich dokumentiert.

4. Die einzelnen Programme und ihre methodischen Grundlagen

4.1. UNIRUN

Dieses Programm besorgt den Aufruf des Systems, die Eingabe der Daten und deren Transformation. Die Daten werden zuerst über einen bestimmten Zeitbereich eingelesen, und dann je nach Wunsch einer bestimmten Transformation unterworfen:

a) Logarithmieren:

$$x_{it} := \log_e x_{it} \quad t = 1, 2, \dots, T \quad (4.1.1.)$$

b) Keine Transformation gemäß c, d, oder e.

c) Absolute Differenzen k-ter Ordnung:

$$x_{it} := x_{it} - x_{it-k} \quad t = \Theta, \Theta+1, \dots, T \quad (4.1.2.)$$

und $\Theta = k+1$

d) Relative Differenzen k-ter Ordnung:

$$x_{it} := \frac{x_{it} - x_{it-k}}{x_{it-k}} \quad t = \Theta, \Theta+1, \dots, T \quad (4.1.3.)$$

und $\Theta = k+1$

e) Akzelerator von b oder c:

$$x_{it} := x_{it} - x_{it-k} \quad t = \Theta, \Theta+1, \dots, T \quad (4.1.4.)$$

$\Theta = 2k+1$

Mit $\Theta = 1$ bzw. Θ aus c, d, e erfolgt dann:

f) Lagbildung k-ter Ordnung: (4.1.5.)

$$x_{it} := x_{it-k} \quad t = \Theta+k, \Theta+k+1, \dots, T$$

Wir setzen nun $\Theta := \Theta+k$ (wenn kein Lag gebildet wird, ist dabei $k = \Theta$).

Die transformierten Zeitreihen werden nun zu einem Zeitpunkt t_r , der vom Benutzer autonom bestimmt und angegeben wird, hinten abgeschnitten und damit auf gleiche Residuenlänge l_zr gebracht. Der Index t hat dann den Bereich $t = 1, 2, \dots, l_zr$ (statt $t = t_r, t_r+1, \dots, T$). Wenn $t_r \leq \Theta$, dann erfolgt eine Fehlermeldung.

In folgenden wird statt l_zr wieder das Symbol T verwendet:

$$t = 1, 2, \dots, T.$$

Zuletzt ist noch ein Normieren möglich:

g) Normieren:

$$x_{it} := \frac{x_{it} - \bar{x}_i}{\sqrt{\frac{1}{T-1} \sum_{\tau=1}^T (x_{i\tau} - \bar{x}_i)^2}} \quad (4.1.6.)$$

$$\text{mit } \bar{x}_i = \frac{1}{T} \sum_{\tau=1}^T x_{i\tau}$$

Folgende Daten werden dann auf die anderen Programme übertragen:

NAMEVK Feld der Namen der Zeitreihen
 INDCTR Feld der Variablencodes (Zeitreihen-
 indikatoren)
 NEND Anzahl der Variablen
 LZR Residuenlänge
 X Feld der transformierten Zeit-
 reihen
 JAFORM=0 logische Größe, die hier anzeigt, daß
 die Daten nicht in Abweichungsform
 stehen.
 XSUM=0 Feld der Mittelwerte der Zeitreihen
 in Abweichungsform (wird hier Null
 gesetzt).

Programmbeschreibung siehe UNIRUN

4.2. UNICOM

Berechnet alle Kombinationen von erklärenden Variablen mit einer (zuletzt eingegebenen) endogenen Variablen mit OLS (siehe unter 4.4.)

Die Kombinationen erfolgen mit einem Modulo - 2 Zähler. Ihre Anzahl ist demnach $2^{\text{nend} - 1}$. Diese bei hohem nend exorbitante Zahl kann durch die Einschränkung auf Regressionen gewisser Ordnung stark herabgedrückt werden. Ebenso ist es möglich, durch Angabe geeigneter Schranken für R^2 , proz und $\hat{\sigma}$ (siehe 4.4.) nur eine gewisse Auswahl von Regressionen ausdrucken zu lassen. Die gleichen Auswahlmöglichkeiten bestehen im übrigen auch für UNIMAX, UNILAG und UNIPAG.

Programmbeschreibung siehe UNICOM und COMOPT.

4.3. UNIMAX

Berechnung aller Kombinationen von erklärenden Variablen mit einer zuletzt angegebenen endogenen Variablen nach einem Schema:

FIX	VAR	VAR	VAR	VAR	VAR	
						1
						2
						3
						4
						5
						6

Die Kombinationen erfolgen nun so:

Es werden nicht die Variablen, sondern die Kästchen kombiniert, dafür aber alle (höchstens 5) Variablen in jedem Kästchen aufgerufen.

Es werden alle (vorhandenen, d.h. angegebenen) Spalten kombiniert, und die Kästchen einer Spalte substituiert. Die erste Spalte taucht, wenn angegeben, in jeder Kombination auf.

Die erste Spalte heißt FIX (fixe Gruppe).

Die folgenden Spalten heißen VAR (variable Gruppe).

Die Kästchen einer Spalte heißen Untergruppen, die variablen Gruppen und Untergruppen werden numeriert.

Nennen wir die Anzahl der angegebenen variablen Gruppen mit $nvar$, die Anzahl der Kästchen der variablen Gruppe i mit $n(i)$ dann lautet die Anzahl k aller Kombinationen des Schemas

a) wenn keine fixe Gruppe vorhanden:

$$K = \prod_{i=1}^{nvar} (n(i) + 1) \quad (4.3.1.)$$

b) wenn eine fixe Gruppe mit m Kästchen (Untergruppen) vorhanden ist:

$$K = m \cdot \prod_{i=1}^{nvar} (n(i) + 1) \quad (4.3.2.)$$

Die Angabe der Variablen einer Untergruppe erfolgt mit der Variablenauswahlkarte Das Schema kann beliebig angefüllt werden. Unter den Optionen ist die Angabe der Konstantenunterdrückung möglich.

Programmbeschreibung siehe UNIMAX, MAXOPT, und DEFKAT.

4.4. UNIOLS

UNIOLS berechnet folgende Resultate der gewöhnlichen Kleinstquadratmethode und druckt sie aus:

geg. seien $\{x_{n(i)t}\}$, die erklärenden Variablen, und y_t , die endogene Variable, und der Modellansatz

$$y_t = \alpha + \sum_{i=1}^n \beta_i x_{n(i),t} + \epsilon_t \quad (4.4.1.)$$

mit $i = 1, 2, \dots, n$

$t = 1, 2, \dots, T$

$$x = \begin{bmatrix} x_{n(1),1} & \dots & x_{n(n),1} \\ \dots & \dots & \dots \\ x_{n(1),T} & \dots & x_{n(n),T} \end{bmatrix}$$

$n(n)$ bedeutet die Nummer der n -ten Zeitreihe aus der Teilmenge E .

Dann wird berechnet (mit den Daten in Abweichungsform)

$$\hat{\beta} = (X'X)^{-1} X'Y \quad (4.4.2.)$$

$$\hat{\alpha} = \bar{Y} - \bar{X}\hat{\beta} \quad (4.4.3.)$$

$$R^2 = \frac{X'Y\hat{\beta}}{Y'Y} \quad (4.4.4.)$$

$$e = Y - X\hat{\beta} \quad (4.4.5.)$$

Die Prozentwerte der Standardabweichung der Konstante sollen $proza$ heißen, die Prozentwerte der Standardabweichung des Koeffizienten $\hat{\beta}_i$ heißen $proz_i$.

$$\text{proza} = 100 \cdot \frac{\hat{s}}{\sqrt{2}} \cdot \sqrt{\left(\frac{1}{T} + \bar{X}'(X'X)^{-1}\bar{X}\right)} \quad (4.4.6.)$$

$$\text{proz}_i = 100 \cdot \frac{\hat{s}}{|\hat{\beta}_i|} \cdot \sqrt{(X'X)^{-1}_{ii}} \quad (4.4.7.)$$

$$\text{Mit } \hat{s}^2 = \frac{e'e}{T - n - 1} \quad (4.4.8.)$$

$$dw = \frac{\sum (e_t - e_{t-1})^2}{\sum e_t^2} \quad (4.4.9.)$$

Wie bei den Programmen unter 4.5., 4.6., und 4.7., ist es möglich die Residuen auszudrucken.

Programmbeschreibung siehe UNIOLS und OLSOPT.

4.5. UNIORF

UNIORF arbeitet mit den Daten in Originalform und berechnet die Regressionsergebnisse.

a) Konstantenunterdrückung:

UNIORF rechnet mit folgenden Formeln:

(4.4.2.)

Für R^2 erfolgt aus technischen Gründen folgende Umwandlung:

$$\begin{aligned} R^2 &= \frac{[\sum Y_t \hat{Y}_t - T\bar{Y}\hat{\bar{Y}}]^2}{[\sum Y_t^2 - T\bar{Y}^2][\sum \hat{Y}_t^2 - T\hat{\bar{Y}}^2]} = & (4.5.1.) \\ &= \frac{[X'Y\hat{\beta} - T\bar{Y}^2 + \bar{Y}(\sum e_t)]^2}{[\sum Y_t^2 - T\bar{Y}^2][\sum Y_t^2 - T\bar{Y}^2 + \sum e_t^2 - \frac{1}{T}(\sum e_t)^2 + 2(\bar{Y}\sum e_t - \sum Y_t e_t)]} \end{aligned}$$

weitere

(4.4.5.), (4.4.7.) (4.4.8.), (4.4.9.).

b) gewöhnliche Kleinstquadratmethode ohne Bedingungen:

An die Zeitreihen wird ein Eins-vektor angefügt und die Anzahl der erklärenden Variablen n automatisch um Eins erhöht:

$$X = \begin{bmatrix} X_{n(1), 1} & \dots & X_{n(n-1), 1} & X_{n, 1} & =1 \\ \vdots & & \vdots & \vdots & \\ X_{n(1), T} & \dots & X_{n(n-1), T} & X_{n, T} & =1 \end{bmatrix}$$

Dann rechnet UNIORF Mit den Formeln:

(4.4.2.)

$$\hat{\beta} = \hat{\beta}_n \tag{4.5.2.}$$

$$R^2 = \frac{X' Y \hat{\beta} - T \bar{Y}^2}{\sum Y_t^2 - T \bar{Y}^2} \tag{4.5.3.}$$

(4.4.5.)

$$\text{proza} = 100 \cdot \frac{\hat{S}}{|Z|} \cdot \sqrt{(X'X)^{-1}_{n,n}} \tag{4.5.4.}$$

(4.4.7.)

$$\hat{S}^2 = \frac{e'e}{T - n} \tag{4.5.5.}$$

(4.4.9.)

Durch die Option "Konstante" wird a. gerechnet, sonst erfolgt automatisch b.

Programmbeschreibung siehe UNIORF, ORFOPT.

4.6. UNIGEN

UNIGEN behandelt die verallgemeinerte Kleinstquadratmethode, welche die Kenntnis der Varianz- Kovarianzmatrix Ω des Fehlergliedes voraussetzt.

Folgende Formeln werden berechnet und deren Resultate ausgedruckt:

$$\hat{\beta} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} Y \tag{4.6.1.}$$

wobei die Zeitreihen y_t und $\{x_{n(i)t}\}$ in Originalform vorliegen.

$$i = 1, 2, \dots, n$$

$$t = 1, 2, \dots, T$$

$x_{n(n), t}$ ist ein Einsvektor.

$$\hat{\alpha} = \hat{\beta}_n \tag{4.6.2.}$$

$$e = Y - X\hat{\beta} \tag{4.6.3.}$$

$$R^2 = 1 - \frac{e' \Omega^{-1} e}{Y' \Omega^{-1} Y} \tag{4.6.4.}$$

Dieses R^2 ist mit dem R^2 aus der gewöhnlichen Kleinstquadratmethode nicht unmittelbar vergleichbar. Es gilt nämlich (mit y in Abweichungsform)

$$R_{OLS}^2 = 1 - \frac{e' e}{y' y} \tag{4.6.5.}$$

und dabei $\sum e_t = 0$.

Für die verallgemeinerte Kleinstquadratmethode gilt die letzte Beziehung nicht, außerdem werden für y Originalwerte und nicht Werte in Abweichungsform verwendet.

Damit ist R_{GEN}^2 auch höher als R_{OLS}^2 .

Um trotzdem einen Vergleich der Korrelationskoeffizienten zu schaffen, wird auch ein verzerrtes R^2 berechnet:

$$VR^2 = 1 - \frac{e' e}{y' y} \tag{4.6.6.}$$

mit Werten von y in Abweichungsform. Natürlich bringt VR^2 verzerrte Resultate, wenn sich Ω stark von I unterscheidet.

$$\hat{s}^2 = \frac{e' \Omega^{-1} e}{T - n} \tag{4.6.7.}$$

$$dw = \frac{\sum (e_t - e_{t-1})^2}{\sum e_t^2} \tag{4.6.8.}$$

$$\text{proz}_i = 100 \cdot \frac{\hat{s}}{|\beta_i|} \sqrt{(X' \Omega^{-1} X)_{ii}^{-1}} \tag{4.6.9.}$$

$$\text{proza} = \text{proz}_n \tag{4.6.10.}$$

Folgende Fälle der Matrix Ω sind möglich:

- a) Die Matrix Ω wird eingelezen
- b) Autoregressiver Prozess erster Ordnung der Fehlerglieder:

$$\begin{aligned} \epsilon_t &= \rho \epsilon_{t-1} + u_t, & E u_t &= 0 \\ & & E u_t u_t' &= \sigma^2 I_T. \end{aligned}$$

damit erhalten wir

(4.6.11.)

$$\Omega^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho & 0 & \dots & 0 \\ -\rho & 1+\rho^2 & -\rho & & \\ 0 & & -\rho & & \\ & & & 1+\rho^2 & -\rho \\ 0 & \dots & 0 & -\rho & 1 \end{bmatrix}$$

ρ wird entweder eingegeben oder aus berechneten Residuen geschätzt:

$$\hat{\rho} = \frac{T \cdot \sum (e_t - \bar{e})(e_{t-1} - \bar{e})}{(T - n) \cdot \sum (e_t - \bar{e})^2} \quad (4.6.12.)$$

$$\bar{e} = \frac{1}{T} \sum e_t$$

- c) Bildung absoluter Differenzen K - ter Ordnung:

$$\sigma^2 \Omega = \sigma^2$$

$$\begin{bmatrix} 2 & 0 & \dots & 0 & -1 & & 0 & \dots & 0 \\ 0 & 2 & & & & & & & \\ \vdots & & \ddots & & & & & & \\ -1 & & & & & & & & \\ 0 & & & & & & & & \\ \vdots & & & & & & & & \\ 0 & \dots & 0 & -1 & & & 0 & \dots & 0 \end{bmatrix}$$

K-1

(4.6.13.)

(4.6.11.) folgt aus der Entwicklung von

$$E \left\{ \begin{bmatrix} u_{k+1} - u_1 \\ u_{k+2} - u_2 \\ \vdots \\ u_T - u_{T-k} \end{bmatrix} (u_{k+1} - u_1, u_{k+2} - u_2, \dots, u_T - u_{T-k}) \right\}$$

unter der Annahme von $E u_t = 0$
 $E u_t u_t = \sigma^2 I_T$

Programmbeschreibung siehe UNIGEN, GENOPT.

4.7. Die Lagprogramme

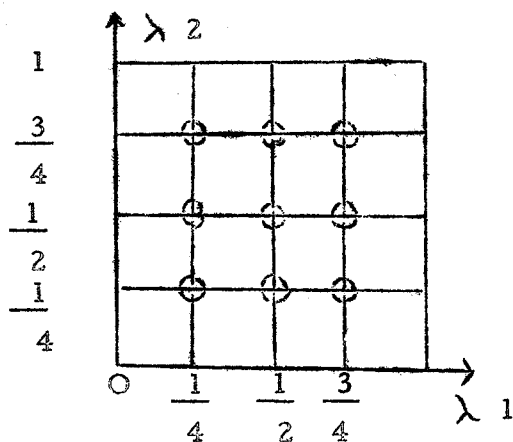
Die Lagprogramme dienen zur Bestimmung von optimalen Lags einer oder mehrerer Variablen, wobei die Lags mit linearer Interpolation (siehe 4.7.1. UNILAG) und mit Splineinterpolation (siehe 4.7.2. UNIPAG) gebildet werden können. Vorerst wird die Methode beschrieben, auf welche Weise die Lagkombinationen hergestellt werden.

Gegeben sei eine Menge $\{x_{n(i)t}\}$ von erklärenden Variablen.

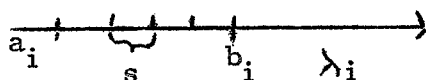
Der Einfachheit halber bezeichnen wir die Variablen aus dieser Menge, die nicht gelagt werden, mit $\{z_{j,t}\}$. Wir unterscheiden dann Variable mit fixem Lag μ_i : $\{x_{i,t-\mu_i}\}$ und Variable mit variablen Lag λ_i : $\{x_{i,t-\lambda_i}\}$; sind 1 Variable mit variablen Lag vorhanden, dann suchen die Lagprogramme auf einem vorgeschriebenen Gitter der λ_i alle Lagkombinationen, mit denen die OLS-Gleichungen gebildet werden:

$$y_t = \sum_{i=1}^l \beta_i x_{i,t-\lambda_i} + \sum_{i=l+1}^{l+m} \beta_i x_{i,t-\mu_i} + \sum_{i=l+m+1}^n \beta_i z_{it} + \varepsilon_t \quad (4.7.1.)$$

Die Lagkombinationen erfolgen auf einem 1 - dimensionalen Gitter, für $l = 2$ also etwa



Das Laggitter kann auf zwei Arten hergestellt werden. Entweder durch die Angabe der Optionen, Semester, Quartale, oder Monate: Dann werden die Lags im Intervall $[0, 1]$ mit entsprechenden Schritten gesucht. Diese Kombinationen erfolgen für alle ausgewählten Variablen mit variablen Lag. Die zweite Möglichkeit besteht darin, für alle Variablen mit variablen Lag die Anzahl der Schritte und die Intervalle anzugeben:



$$s = \frac{b_i - a_i}{n}, \text{ mit } \lambda_i \in [a_i, b_i], \text{ und} \quad (4.7.2.)$$

der Schrittzahl n .

Welche Lagkombination dann ausgewählt wird, unterliegt dem Ermessen des Rechnenden. Eine Diskussion über dieses Thema erfolgt in 4.7.3. Lagstrukturen.

Beschreibungen dieser Programme:

UNILAG, LAGOPT, LAGSPK bzw.
UNIPAG, LAGOPT, LAGSPK.

4.7.1. UNILAG

Die Lagbildung im Programm UNILAG erfolgt mit linearer Interpolation.

Gegeben sei eine Variable $x_{n(i),t}$, und ein Lag λ_i .

Dann lautet die gelagte Variable x_{it}^L :

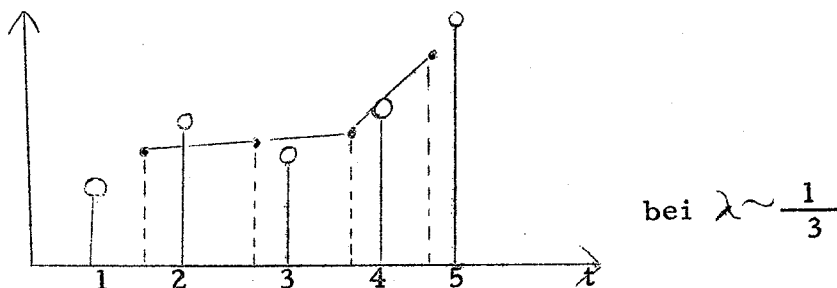
$$x_{i,t}^L := x_{n(i),t+1} - \lambda_i := (1 - \lambda_i) x_{n(i),t+1} + \lambda_i x_{n(i),t}$$

$$t = 1, 2, \dots, T - 1 \quad (4.7.3.)$$

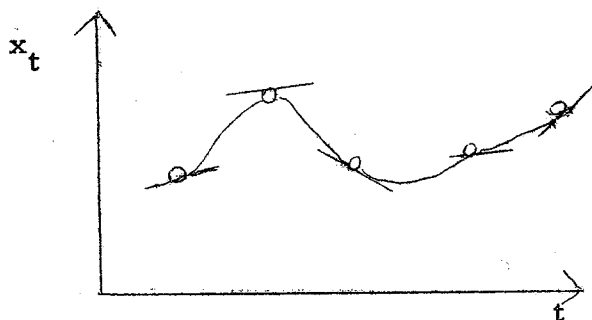
Dann wird die Formel (4.7.1.) mit der gewöhnlichen Kleinstquadratmethode ausgewertet. Die Residuenlänge wird automatisch um Eins verkürzt.

4.7.2. UNIPAG

Der Nachteil der linearen Interpolation bei UNILAG liegt am Verlust an Information, der durch die Glättung der Zeitreihen entsteht.



Der Gedanke von UNIPAG ist es, den Verlauf der Zeitreihe durch eine möglichst glatte Kurve, d. h. ein Polynom zu approximieren. Dafür eignen sich besonders die kubischen Splines. Diese Kurven sind Polynome 3. Grades, welche zwischen je 2 Interpolationspunkten verlaufen. Jedes Polynom hat mit dem benachbarten Polynom den Berührungspunkt und die Tangente gemeinsam.



Wir führen nun die "Zéitreehenfunktion" $s(t)$ mit stetigem Argument ein. Sie soll den kubischen Spline darstellen, und den folgenden Bedingungen genügen:

$$s(k) = x_k \text{ für } k = 1, 2, \dots, T$$

Auf eine Auswahl von Zeitreehen bezogen würde er lauten:

$$s_i(k) = x_{n(i),k} \text{ für } k = 1, 2, \dots, T \text{ und } i = 1, 2, \dots, n.$$

Ableitung des kubischen Spline $s(t)$.

Wir betrachten sofort als Spezialfall eine äquidistante Teilung ω im Intervall $[1, T]$:

$$t_k := k \text{ für } k = 1, 2, \dots, T$$

P_3 sei die Menge der Polynome vom Grad 3. $C_2[1, T]$ sei der Raum der 2 mal stetig differenzierbaren Funktionen auf dem Intervall $[1, T]$.

Darum soll hier als Definition für den Spline s gelten:

$$s \in C_2[1, T] \text{ mit } s \in P_3 \text{ für } t \in [t_{k-1}, t_k] \text{ oder } t \in [k-1, k].$$

Der Spline lautet damit für $t \in [k-1, k]$:

$$\begin{aligned} s(t) = & -s_k'' \frac{(t-k)^3}{6} + s_k'' \frac{(t-k+1)^3}{6} + \\ & + \left(x_k - \frac{s_k''}{6} \right) (t-k+1) - \left(x_{k-1} - \frac{s_{k-1}''}{6} \right) (t-k). \end{aligned} \quad (4.7.4.)$$

Wir setzen nun statt t den Wert $k-\lambda$ ein:

$$s(k-\lambda) = \frac{s_{k-1}''}{6} \lambda^3 + \frac{s_k''}{6} (1-\lambda)^3 + \\ + \left(x_k - \frac{s_k''}{6}\right) (1-\lambda) + \left(x_{k-1} - \frac{s_{k-1}''}{6}\right) \lambda$$

mit $k = 2, 3, \dots, T$

Mit sp_k statt $\frac{s_k''}{6}$ lautet dieser Ausdruck endgültig

$$s(k-\lambda) = sp_{k-1} \lambda^3 + sp_k (1-\lambda)^3 + \\ + (x_k - sp_k) (1-\lambda) + (x_{k-1} - sp_{k-1}) \lambda. \quad (4.7.5.)$$

Unter der Voraussetzung, daß wir $s'(1)$ und $s'(T)$ bereits kennen, erhalten wir T Gleichungen für die T Unbekannten

$$sp_1, sp_2, \dots, sp_T \quad ; \quad \text{mit } x_1' = s'(1) \text{ und } x_T' = s'(T)$$

$$\begin{aligned} 2sp_1 + sp_2 &= x_2' - x_1' - x_1' \\ sp_1 + 4sp_2 + sp_3 &= x_3' - 2x_2' + x_1' \end{aligned} \quad (4.7.6.)$$

$$\begin{aligned} sp_T - 2 + 4sp_{T-1} + sp_T &= x_T' - 2x_{T-1}' + x_{T-2}' \\ sp_{T-1} + 2sp_T &= (x_T' - x_{T-1}') + x_T' \end{aligned}$$

Die rechte Seite des tridiagonalen symmetrischen Gleichungssystems bezeichnen wir mit e . Die Werte für x_1' und x_T' definieren wir durch den Anstieg des Lagrangepolynoms 2. Grades durch die ersten bzw. letzten drei Interpolationspunkte.

a) linke Seite:

$$\psi(t) = x_1 \frac{(t-2)(t-3)}{(1-2)(1-3)} + x_2 \frac{(t-1)(t-3)}{(2-1)(2-3)} + x_3 \frac{(t-1)(t-2)}{(3-1)(3-2)}$$

$$\varphi(t) = \frac{x_1}{2}(t^2 - 5t + 6) - x_2(t^2 - 4t + 3) + \frac{x_3}{2}(t^2 - 3t + 2)$$

$$\varphi'(t) = t(x_1 - 2x_2 + x_3) + \left(-\frac{5}{2}x_1 + 4x_2 - \frac{3}{2}x_3\right)$$

$$\varphi'(1) = -\frac{3}{2}x_1 + 2x_2 - \frac{1}{2}x_3$$

$$e_1 = x_2 - x_1 - x_1' = x_2 - x_1 + \frac{3}{2}x_1 - 2x_2 + \frac{1}{2}x_3$$

$$e_1 = \frac{x_1 - 2x_2 + x_3}{2} = \frac{e_2}{2} \quad (4.7.7.)$$

b) rechte Seite

$$\begin{aligned} \varphi(t) &= x_{T-2} \frac{(t-T+1)(t-T)}{(-1)(-2)} + x_{T-1} \frac{(t-T+2)(t-T)}{(1)(-1)} + \\ &+ x_T \frac{(t-T+2)(t-T+1)}{(2)(1)} \end{aligned}$$

$$\varphi'(t) = \frac{x_{T-2}}{2}(2t - 2T + 1) - x_{T-1}(2t - 2T + 2) + \frac{x_T}{2}(2t - 2T + 3)$$

$$\varphi'(T) = \frac{x_{T-2}}{2} - 2x_{T-1} + \frac{3}{2}x_T$$

$$e_T = -x_T + x_{T-1} + x_T' = -x_T + \frac{3}{2}x_T + x_{T-1} - 2x_{T-1} + \frac{1}{2}x_{T-2}$$

$$e_T = \frac{x_T - 2x_{T-1} + x_{T-2}}{2} = \frac{e_{T-1}}{2} \quad (4.7.8.)$$

Die Lösung des Gleichungssystems erfolgt mit einem Einzelschrittverfahren. Der Algorithmus wird in der Dokumentation, Subroutine TRISYM, genau beschrieben.

Mit den erhaltenen Werten der sp_k kann jeder beliebige Wert der Zeitreihenfunktion im Intervall $[1, T]$ ermittelt werden. Die Auswertung der Regressionen erfolgt dann wie am Anfang von 4.7. beschrieben.

Die angenehme Eigenschaft der kubischen Splines sind ihre Minimal-eigenschaften bezüglich der zweiten Ableitungen.

Sei Φ folgende Menge von Funktionen:

$$\Phi := \left\{ \varphi \in C_2 [1, T] \mid \bigwedge_{1 \leq k \leq T} \varphi(k) = x_k \right\}$$

s sei der oben erhaltene Spline mit

$$s(k) = x_k. \text{ Dann gilt für alle } \varphi$$

$$\sqrt{\int_1^T s''(t)^2 dt} \leq \sqrt{\int_1^T \varphi''(t)^2 dt}. \quad (4.7.9.)$$

Damit werden unnötige und fehlerhafte Oszillationen der Zeitreihenfunktion ausgeschaltet.

Es ist jedoch notwendig noch zu prüfen, ob die Zeitreihenfunktion einen sinnvollen Kurvenverlauf darstellt. Die kubische Interpolation kann nützlich sein, etwa bei konjunkturellen Verlauf der Zeitreihe, oder sinnlos sein, etwa bei starken zufälligen Schwankungen der Zeitreihe. Deshalb werden die Daten der interpolierten Zeitreihe jeweils ausgedruckt.

4.7.3. Die Interpretation von Lagstrukturen

Es liegt nahe, unter allen Gleichungen auf einem gegebenen Laggitter jene auszuwählen, welche die besten Resultate in bezug auf die statistischen Kriterien R^2 und p liefert (p bedeutet Prozentwerte der Standardabweichung). Wir suchen also nach Extremwerten für R^2 und p . Dabei müssen wir beachten, daß diese Schätzfunktionen Zufallsvariable sind, daher gilt dasselbe auch für

$$\frac{dR^2}{d\lambda} \quad \text{und} \quad \frac{dp}{d\lambda},$$

wenn wir die Maxima in bezug auf den Lag suchen. Aus diesem Grund scheint es angebracht, die Ergebnisse von UNILAG und UNIPAG einer besonders sorgfältigen Untersuchung zu unterziehen.

Empirische Ergebnisse mit linearer Interpolation haben vermuten lassen daß zwischen den Größen, R^2 , p und $\hat{\beta}$ ein ursächlicher Zusammenhang besteht: Der Extremwert einer dieser Größen bedingt (mit gewisser Wahrscheinlichkeit) einen Extremwert der anderen. Die folgende Untersuchung soll dieses Resultat theoretisch festigen, und wir betrachten vorerst nur den Fall der linearen Interpolation, und beschränken uns auf den einfachen Modellansatz

$$\tilde{y}_t = \alpha + \beta \tilde{x}_{t-\lambda} + \varepsilon_t, \quad t = 1, 2, \dots, T \quad (4.7.10.)$$

\sim bedeutet Variable in Originalform.

Der Wert \tilde{x}_0 soll vorhanden sein, Wir notieren ab jetzt

$$y_t = \tilde{y}_t - \frac{1}{T} \sum_{t=1}^T \tilde{y}_t \quad (4.7.11.)$$

$$\begin{aligned} x_t^\lambda &= \tilde{x}_{t-\lambda} - \frac{1}{T} \sum_{t=1}^T \tilde{x}_{t-\lambda} = \\ &= (1-\lambda) \tilde{x}_t + \lambda \tilde{x}_t - \bar{x}(\lambda). \end{aligned} \quad (4.7.12.)$$

Wir verwenden noch die Schreibweise

$$\begin{aligned} y' y &= \sum_{t=1}^T y_t^2 & t = 1, 2, \dots, T. \\ x' x &= \sum_{t=1}^T x_t^2 \\ x' y &= \sum_{t=1}^T x_t y_t \end{aligned}$$

Die Schätzfunktionen sind rationale Funktionen in λ :

$$\hat{\beta} = \frac{x^{iL} y}{x^{iL} x^L} = \hat{\beta}(\lambda) \quad (4.7.13.)$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}(\lambda) \bar{x}(\lambda) = \hat{\alpha}(\lambda) \quad (4.7.14.)$$

$$R^2 = \frac{(x^{iL} y)^2}{x^{iL} x^L y y} = R^2(\lambda) \quad (4.7.15.)$$

$$p = \text{loo} \cdot \frac{s}{|\hat{\beta}|} \cdot \frac{1}{\sqrt{x^{iL} x^L}} = p(\lambda). \quad (4.7.16.)$$

Es gilt nun der Satz 4.1.

Wenn R^2 für das Modell (4.7.10.) einen Extremwert besitzt, dann gilt das auch für p und umgekehrt.

$$\frac{dR^2}{d\lambda} = 0 \leftrightarrow \frac{dp}{d\lambda} = 0$$

Beweis:

$$R^2 = \frac{(x^{iL} y)^2}{y'y \cdot x^{iL} x^L}$$

$$\begin{aligned} \frac{dR^2}{d\lambda} &= \frac{1}{y'y} \left[\frac{2x^{iL} y}{x^{iL} x^L} \frac{d}{d\lambda} x^{iL} y - \frac{(x^{iL} y)^2}{(x^{iL} x^L)^2} \cdot \frac{d}{d\lambda} x^{iL} x^L \right] = \\ &= \frac{1}{y'y} \left[2\hat{\beta} \frac{d}{d\lambda} x^{iL} y - \hat{\beta} \frac{d}{d\lambda} x^{iL} x^L \right] = 0 \end{aligned} \quad (4.7.17.)$$

laut Annahme.

o.B.d.A. soll nun $\hat{\beta} > 0$ gelten.

p lässt sich umformen:

$$p = \frac{\text{loo}}{\sqrt{T-2}} \cdot \frac{\sqrt{y'y} \sqrt{x^{iL} x^L} \cdot \sqrt{1-R^2}}{x^{iL} y}$$

$$\begin{aligned} \frac{dp}{d\lambda} &= \frac{\text{loo}}{\sqrt{T-2}} \frac{\sqrt{y'y} \sqrt{x^{iL} x^L}}{x^{iL} y} \cdot \frac{1}{2\sqrt{1-R^2}} \left(-\frac{dR^2}{d\lambda} \right) - \\ &- \frac{\text{loo}}{\sqrt{T-2}} \frac{\sqrt{y'y} \sqrt{x^{iL} x^L} \cdot \sqrt{1-R^2}}{(x^{iL} y)^2} \cdot \frac{d}{d\lambda} x^{iL} y + \\ &+ \frac{\text{loo}}{\sqrt{T-2}} \frac{\sqrt{y'y} \sqrt{1-R^2}}{x^{iL} y} \cdot \frac{1}{2\sqrt{x^{iL} x^L}} \cdot \frac{d}{d\lambda} x^{iL} x^L \end{aligned}$$

Laut Annahme ist der erste Ausdruck Null. Substitution der folgenden Ausdrücke ergibt wieder

$$\begin{aligned}
 \frac{dp}{d\lambda} &= \frac{100s}{\hat{\beta}} \left[-\frac{\frac{d}{d\lambda} x^{IL} y}{x^{IL} y \sqrt{x^{IL} x^L}} + \frac{\frac{d}{d\lambda} x^{IL} x^L}{2x^{IL} x^L \sqrt{x^{IL} x^L}} \right] = \\
 &= p \left[-\frac{1}{\hat{\beta}} \frac{\frac{d}{d\lambda} x^{IL} y}{x^{IL} x^L} + \frac{1}{2} \frac{\frac{d}{d\lambda} x^{IL} x^L}{x^{IL} x^L} \right] = \\
 &= -\frac{p}{2\hat{\beta} x^{IL} x^L} \left[2\hat{\beta} \frac{d}{d\lambda} x^{IL} y - \hat{\beta}^2 \frac{d}{d\lambda} x^{IL} x^L \right]
 \end{aligned}
 \tag{4.7.18.}$$

Ein Vergleich mit (4.7.17) ergibt die Identität der Klammerausdrücke. Wenn man beachtet, daß die Momente von x^L und y nicht identisch verschwinden, läßt sich aus

$$\frac{dp}{d\lambda} = 0 \quad \text{wieder} \quad \frac{dR^2}{d\lambda} = 0 \quad \text{schließen.}$$

Damit ist Satz 4.1. gezeigt.

Wir wenden uns nun der asymptotischen Betrachtungsweise zu, und formulieren folgenden bekannten Satz, dessen Annahmen in Hinblick auf das vorgegebene Modell (4.7.10.) etwas modifiziert wurden:

Satz 4.2.

Folgende Annahmen sollen gelten:

1. $E(\varepsilon_t) = 0$
 $E(\varepsilon_t \varepsilon_t') = \sigma^2 I_T$
2. \tilde{x}^L ist nichtstochastisch, und nicht identisch Null.

3. Der Grenzwert der Matrix

$$\begin{bmatrix} 1 & \bar{x}(\lambda) \\ \bar{x}(\lambda) & \frac{x'x}{T} \end{bmatrix}^{-1}, \quad \text{soll existieren.}$$

(Daraus folgt $\lim_{T \rightarrow \infty} \left(\frac{1}{T} x'x \right)^{-1} = c < \infty$) (4.7.19.)

Dann ist $\hat{\beta}$ von (4.7.13.) eine konsistente Schätzfunktion für $\beta(\lambda)$.

Natürlich ist das wahre β eine Funktion von λ . Wir benützen in folgenden die Konsistenz von $\hat{\beta}(\lambda)$ und damit die Beschränktheit. Ebenso ist auch $R^2(\lambda)$ beschränkt.

Satz 4.3.

Annahme: $\lim_{T \rightarrow \infty} \frac{\frac{d}{d\lambda} \frac{x'x}{T}}{\frac{x'x}{T}} = 0.$ (4.7.20.)

Unter dieser Bedingung gilt: $\frac{dR^2}{d\lambda} = 0 \rightarrow \text{plim} \frac{d\hat{\beta}}{d\lambda} = 0$
 $\frac{d\hat{\beta}}{d\lambda} = 0 \rightarrow \text{plim} \frac{dR^2}{d\lambda} = 0$

Beweis:

$$R^2 = \hat{\beta}' \frac{x'x}{y'y}$$

$$\frac{dR^2}{d\lambda} = 2 \hat{\beta}' \frac{d\hat{\beta}}{d\lambda} \frac{x'x}{y'y} + \hat{\beta}' \cdot \frac{1}{y'y} \cdot \frac{d}{d\lambda} x'x$$

$$\frac{dR^2}{d\lambda} = 0 \rightarrow \frac{d\hat{\beta}}{d\lambda} = -\frac{1}{2} \hat{\beta}' \frac{\frac{d}{d\lambda} x'x}{x'x}$$

$$\begin{aligned} \frac{d\hat{\beta}}{d\lambda} = 0 \rightarrow \frac{dR^2}{d\lambda} &= \hat{\beta}' \frac{1}{y'y} \cdot \frac{d}{d\lambda} x'x = \\ &= R^2 \cdot \frac{\frac{d}{d\lambda} x'x}{x'x} \end{aligned}$$

Mit der obigen Annahme und dem Satz von Slutsky gilt

$$\text{plim} \frac{d\hat{\beta}}{d\lambda} = -\frac{1}{2} \beta(\lambda) \lim \frac{\frac{d}{d\lambda} x'x}{x'x} = 0$$

$$\left| \text{plim} \frac{dR^2}{d\lambda} \right| \leq \left| \text{plim} \frac{\frac{d}{d\lambda} x'' x^L}{x'' x^L} \right| = \left| \frac{\lim \frac{d}{d\lambda} x'' x^L}{x'' x^L} \right| = 0$$

Damit ist der Satz bewiesen.

Die Annahme von Satz 4.3. erfordert noch eine kurze Diskussion.

Der Ausdruck (4.7.20.) fordert, daß bei hinreichend hoher Zahl von

Beobachtungen der Einfluß einer Verzögerung nicht mehr spürbar wird:

Der relative Zuwachs der Quadratsumme bei einer Verzögerung um $\Delta\lambda$

soll verschwinden. Diese Bedingung ist nicht immer erfüllt, etwa

wenn die Beobachtungen \tilde{x}_t^L regelmäßig um ihren Mittelwert $\bar{x}(\lambda)$

schwanken. Bei annähernd gleichmäßiger Verteilung ist die Bedingung

jedoch leicht erfüllt. Wir betrachten den Sonderfall

$$x_t^L := \frac{t}{T}, \quad t = 1, 2, \dots, T$$

$$\begin{aligned} \frac{d}{d\lambda} x'' x^L &= \frac{d}{d\lambda} \sum_{t=1}^T ((1-\lambda)\tilde{x}_t + \lambda\tilde{x}_{t-1} - \bar{x}(\lambda))^2 = \\ &= - \sum_{t=1}^T x_t^L \cdot (\tilde{x}_t - \tilde{x}_{t-1} + \frac{d}{d\lambda} \bar{x}(\lambda)) \quad (4.7.21.) \end{aligned}$$

$$\tilde{x}_t - \tilde{x}_{t-1} = \frac{1}{T} (t - t + 1) = \frac{1}{T}$$

$$\frac{d}{d\lambda} \bar{x}(\lambda) = \text{konst}(T) \quad \text{siehe (4.7.22.)}$$

$$\frac{d}{d\lambda} x'' x^L = - \left(-\frac{1}{T} - \text{konst}(T) \right) \sum_{t=1}^T x_t^L \equiv 0$$

weil die x_t^L in Abweichungsform stehen. Umso stärker wegen (4.7.19)

gilt daher die Annahme (4.7.20.). Auch wenn die x_t^L unregelmäßig

verteilt sind, wird die Annahme noch sicher zutreffen.

Satz 4.4.

Falls $\frac{d\hat{\beta}}{d\lambda} = 0$, gilt $\text{plim} \frac{d\hat{\alpha}}{d\lambda} = 0$,

und umgekehrt: $\frac{d\hat{\alpha}}{d\lambda} = 0 \rightarrow \text{plim} \frac{d\hat{\beta}}{d\lambda} = 0$.

Beweis:

$$\hat{\mathcal{L}} = \bar{y} - \hat{\beta}(\lambda) \cdot \bar{x}(\lambda)$$

$$\frac{d\hat{\mathcal{L}}}{d\lambda} = \frac{d\hat{\beta}(\lambda)}{d\lambda} \cdot \bar{x}(\lambda) - \hat{\beta}(\lambda) \frac{d\bar{x}(\lambda)}{d\lambda}$$

$\bar{x}(\lambda)$ ist endlich.

$$\frac{d}{d\lambda} \bar{x}(\lambda) = \frac{d}{d\lambda} \frac{1}{T} \sum_{t=1}^T \tilde{x}_t = \frac{d}{d\lambda} \cdot \frac{1}{T} \sum_{t=1}^T ((1-\lambda)\tilde{x}_t + \lambda\tilde{x}_{t-1}) =$$

$$= \frac{-1}{T} (x_T - x_1). \quad (4.7.22.)$$

$$\lim_{T \rightarrow \infty} \frac{d}{d\lambda} \bar{x}(\lambda) = - (x_T - x_1) \lim_{T \rightarrow \infty} \frac{1}{T} = 0. \quad (4.7.23.)$$

Damit ist der Satz bewiesen.

Bisher haben wir nur den Fall der linearen Interpolation behandelt.

Für die Spline-Interpolation gilt im Gegensatz zu (4.7.12.), wenn man die Glieder in (4.7.5.) zusammenfaßt, mit einer Funktion $g_t(\lambda)$:

$$x_t^L = s(t-\lambda) - \bar{x}(\lambda) = \tilde{x}_t(1-\lambda) + \tilde{x}_{t-1}\lambda + g_t(\lambda) - \bar{x}(\lambda)$$

$$(4.7.24.)$$

Die Werte sp_t sind Linearkombinationen der \tilde{x}_t und daher beschränkt. Wegen (4.7.5.) sind daher auch die $\frac{d}{d\lambda} g_t(\lambda)$ beschränkt. Wenn wir weiterhin regelmäßige Oszillation der \tilde{x}_t ausschließen, gilt die Annahme (4.7.20.).

Weiters folgt aus (4.7.6.), daß

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum sp_t = 0 \text{ und damit}$$

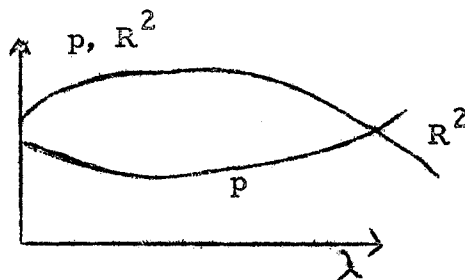
$$\lim_{T \rightarrow \infty} \frac{1}{T} \frac{d}{d\lambda} \sum_t g_t(\lambda) = 0 \quad (4.7.25.)$$

Damit gilt der

Satz 4.5.

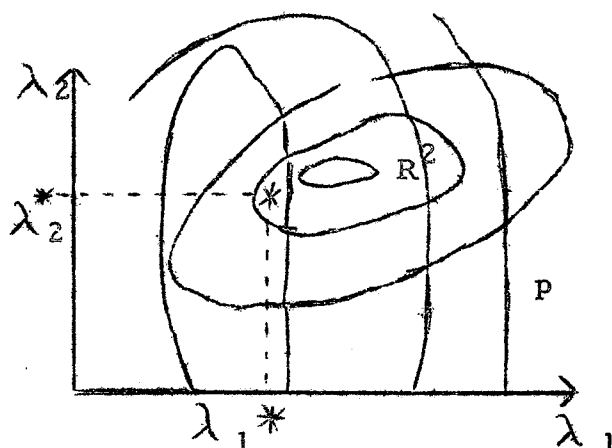
Die Sätze 4.1., 4.3. und 4.4. gelten auch für die Spline- Interpolation. Für die praktische Beurteilung von Lagstrukturen ergeben sich nun folgende Konsequenzen:

1. Die Ergebnisse der Sätze 4.1., 4.3., 4.4., und 4.5. gelten abgeschwächt auch für die multiple Regression.
2. Wenn eine optimale Gleichung für einen Lag $\lambda_i > 0$ existiert, dann läßt sich das Vorhandensein eines solchens Lags bereits durch eine Einzelregression mit der entsprechenden Variablen nachweisen:
Die Kurven der $R^2(\lambda)$ und $p(\lambda)$ sehen ungefähr so aus:



3. Es ist unbedingt empfehlenswert, die gesamte Lagstruktur ins Auge zu fassen, und die Veränderung der statistischen Größen auf dem gesamten Gitter zu berücksichtigen. Nur dann ist es möglich, einen "zufälligen" Extremwert von R^2 oder p von einem tatsächlichen zu unterscheiden.

Ein Schaubild der Niveaulinien läßt die Lagstruktur dabei deutlich hervortreten:



λ_1^* und λ_2^* wären die optimalen Lags. Auch die Schwankungsbreite von $\hat{\beta}(\lambda)$ ist sehr aufschlußreich.

4. UNILAG und UNIPAG suchen eine optimale 2 - Punkt - Lag-
verteilung in einem vorgegebenen Intervall. Es ist daher keinesfalls
ratsam, mit diesen Programmen zu extrapolieren, d.h. einen Lag
zu wählen, der weit über das $[0, \bar{1}]$ - Intervall hinausgeht
(z. B. einen Lag $\lambda = 1.5$, wenn $\lambda = 0$ und $\lambda = 1$ die Stützstellen
der Interpolation darstellen).

5. Die Trennschärfe zur Wahl des optimalen λ darf nicht über-
schätzt werden. Der optimale λ läßt sich im allgemeinen nur in einem
Intervall von 2-4 Monaten lokalisieren.

6. Wenn eine nichtsignifikante Konstante vorliegt, dann verschwindet
sie meistens in der Nähe des optimalen Punktes von selbst.

4.8. UNISIM

4.8.1. Einleitung und Methode

Bei der Analyse eines ökonometrischen Modells begegnen wir stets zwei grundsätzlichen Schwierigkeiten: Erstens können wir (im Sinn der mathematischen Statistik) niemals eine Struktur mit absoluter Sicherheit verifizieren, und zweitens wissen wir nur in wenigen Fällen genau über die Art der auftretenden Fehler Bescheid. Es ist evident, dass der Modellansatz

$$y = X\beta + \varepsilon \quad (4.8.1.)$$

die Gefahr der Einseitigkeit in sich birgt. Sind die Fehler ε_t tatsächlich nur die summierten und vernachlässigbaren Elemente im Modellansatz? Oder stellen die ε Fehler in den Messwerten der endogenen Variablen dar? Aber dann ist nicht einzusehen, warum die Beobachtungsmatrix X nicht stochastisch, das heisst in diesem Fall messfehlerfrei sein sollte. Tatsächlich lässt sich zeigen, dass Fehler in den erklärenden Variablen eine systematische Schätzverzerrung hervorrufen. Im allgemeinen treten Schätzverzerrungen immer dann auf, wenn im Ansatz (4.8.1.) die Matrix $E(X'\varepsilon)$ nicht gleich der Nullmatrix ist. Es ist natürlich auch in diesem Fall möglich, die Kleinstquadratmethode anzuwenden; es bleibt nur die Frage, ob etwa ein β^* , welches den Ausdruck

$$S = E_{y,x} \left\{ (y - X\beta^*)' (y - X\beta^*) \right\} \quad (4.8.2.)$$

minimiert, mit dem β des vorweggenommenen Modellansatzes übereinstimmt. Für genaue Details sei im weiteren etwa auf Schönfeld II, § 8 verwiesen. Wir erwähnen hier nur kurz zwei wesentliche Sätze, welche das Problem der Schätzverzerrung betreffen.

Satz 4.6.

Es gelte der Ansatz (4.8.1.)

Die Matrix $E \left\{ (X, \varepsilon)' (X, \varepsilon) \right\}$ soll existieren, $E(X'X)$ sei nicht singulär.

$\beta = \beta^*$ dann und nur dann, wenn $E(X'\varepsilon) = 0$.

Statt (X, ε) schreiben wir nun $\{x(t); \varepsilon_t\}$, wobei dieser Klammerausdruck einen stationären stochastischen Prozess darstellt, der Momente bis zur 4. Ordnung besitzt und gewisse Ergodizitätsbedingungen erfüllt. Die Matrix M_{xx} der zweiten Momente sei nicht singulär:

$$M_{xx} = E \{x(t) x'(t)\}$$

Weiter sei

$$M_{xx} = \frac{1}{T} \sum x(t) x'(t) \quad (4.8.3.)$$

$$m_{xy} = \frac{1}{T} \sum x(t) y'(t)$$

Dann gilt der folgende

Satz 4.7.

Für jede Lösung $\hat{\beta}_T$ der Normalgleichungen

$$M_{xx} \hat{\beta}_T = m_{xy} \quad (4.8.4.)$$

gilt dann $\text{plim}_{T \rightarrow \infty} \hat{\beta}_T = \beta^*$, mit β^* von (4.8.2.).

Damit konvergiert bei wachsenden Stichprobenumfang der Schätzwert $\hat{\beta}$ gegen ein β^* , das laut Satz 4.6. nicht unbedingt mit dem angenommenen β übereinstimmen muss.

Um Schätzverzerrungen zu vermeiden, verwendet man sogenannte "Instrumente": Man führt neue Variable ein, um methodische Schätzverzerrungen zu vermeiden. Oft wird jedoch dadurch der Modellansatz nur noch komplizierter. Zur Lösung des Fehler-in-den-Variablen-Modells bietet sich ein Verfahren mit orthonormaler Regression und Eigenwertbestimmung an. Die Bestimmung des kleinsten Eigenwertes ist jedoch mit grossen relativen Fehlern verbunden.

In dieser Situation wurde vom Verfasser ein einfaches Simulationsverfahren entworfen, welches dem Ökonometriker einige Informationen über das Verfahren der Schätzwerte $\hat{\beta}$ liefern soll. Im Anschluss an die Darstellung der Methode folgt noch eine Diskussion der Aussage-

möglichkeiten und der Beschränkungen.

Gegeben sei die Struktur

$$y_t = \sum \beta_i x_{n(i),t} + \varepsilon_t \quad (4.8.6.)$$

Von ihr wollen wir annehmen, dass sie die richtige ist. Die Matrix $E(X'X)$ ist jedoch im allgemeinen unbekannt.

Wir bilden nun aus den vorhandenen Beobachtungen neue, gestörte Variable:

$$\psi_t = y_t + |y_t| \cdot \xi_t \quad (4.8.7.)$$

$$\varphi_{it} = x_{n(i),t} + |x_{n(i),t}| \cdot \eta_{it} \quad (4.8.8.)$$

Die ξ_t und η_{it} sind normalverteilte Zufallsvariable:

$$\xi_t \text{ ist } N(0, \xi^2) \quad (4.8.9.)$$

$$\eta_{it} \text{ ist } N(\mu_i, \rho_i^2) \quad (4.8.10.)$$

Wir definieren noch zur Einfachheit

$$v_t = |y_t| \xi_t \quad (4.8.11.)$$

$$u_{it} = |x_{n(i),t}| \cdot \eta_{it} \quad (4.8.12.)$$

Dabei sind die v_t und u_{it} Zufallsvariable mit variablen Mittelwerten und Streuungen. Die Absolutbeträge in (4.8.7.) und (4.8.8.) dienen dazu, um einen angegebenen positiven oder negativen Mittelwert v bzw. μ_i in den gestörten Variablen fortwirken zu lassen, gleich ob die Beobachtungen y_t bzw. $x_{n(i),t}$ positiv oder negativ sind. Die Aufgabe der Verteilungsparameter v und μ_i erfolgt in positiven oder negativen Promillewerten, die Angabe von ξ und ρ_i in positiven Promillewerten.

Die Anzahl von Simulationen ist beliebig wählbar. Ebenso kann man nach beliebigen npart Schritten die Zwischenergebnisse ausdrucken lassen.

UNISIM liefert folgende Resultate:

Wir bezeichnen die Matrix der gestörten Beobachtungen mit (ψ, ϕ) . Der Vektor $\hat{\beta}_s$ wird mit Hilfe der gewöhnlichen Kleinstquadratmethode berechnet:

$$\hat{\beta}_s = (\phi' \phi)^{-1} \phi' \psi \quad (4.8.13.)$$

Der Vektor der Mittelwerte über n totale Simulationen lautet:

$$\hat{\bar{\beta}}_s = \frac{1}{n_{total}} \sum_{\text{Simul.}} \hat{\beta}_s \quad (4.8.14.)$$

Die Streuung

$$s_{\beta} = \sqrt{\frac{1}{n_{total}} \sum_{\text{Simul.}} (\hat{\beta}_s - \hat{\bar{\beta}})^2} \quad (4.8.15.)$$

Der Mittelwert der Determinante $(\phi' \phi)^{-1}$:

$$\Delta = \frac{1}{n_{total}} \sum_{\text{Simul.}} \text{Det}(\phi' \phi)^{-1} \quad (4.8.16.)$$

Der Ausdruck von (4.8.15.) erfolgt in Promillewerten, sonst in Absolutbeträgen.

Nach Wunsch kann man eine Tabelle der empirischen Häufigkeitsverteilung der $\hat{\beta}_s$ in einem vorgegebenen Intervall erhalten.

Programmbeschreibung siehe UNISIM, SIMOPT und PERTUR.

4.8.2. Aussagemöglichkeiten von UNISIM

Vorerst sollen die Beschränkungen erwähnt werden, denen die Simulationsmethode unterliegt.

1. Es ist keinesfalls möglich, die Schätzverzerrung $\beta - \beta^*$ im Sinne von Satz 4.6 festzustellen.
2. Die normalverteilten Zufallsvariablen v_t und u_{it} bedingen in gewisser Weise die numerischen Resultate.

Folgende Aussagen scheinen nach Meinung des Verfassers sinnvoll:

3. Die Grösse Δ (4.8.16) ist besonders in jenen Fällen interessant, wo $(X'X)$ bereits eine fast singuläre Matrix darstellt. Mit UNISIM kann daher eine vermutete Multikollinerität leicht aufgedeckt werden.
4. Es ist möglich, den Einfluss von Fehlern in den Daten auf die Koeffizienten $\hat{\beta}$ festzustellen. Auf diese Weise können leicht stabile und instabile Strukturen unterschieden werden.
5. Systematische Fehler in einer Zeitreihe, wie etwa eine stete Unterschätzung der wahren Messwerte, beeinflussen auf verwickelte Weise das Verhältnis der Komponenten $\beta_1 : \beta_2 : \dots : \beta_n$. Mit UNISIM können solche Relationen wie $\frac{\hat{\beta}_i}{\hat{\beta}_j}$ untersucht werden. Bei Datenfehlern ohne systematische Abweichung muss man allerdings voraussetzen, dass v bzw. μ_i Null ist.
6. Wenn $\hat{\beta}_s$ als Zufallsvariable in den Simulationen aufgefasst wird, kann die empirische Verteilung einer Komponente $\hat{\beta}_{s,i}$ betrachtet werden. Die Häufigkeitsverteilung der $\hat{\beta}_{s,i}$ ähnelt sicher der Gestalt der Dichtefunktion einer normalverteilten Zufallsvariablen. In diesem Zusammenhang interessiert jedoch nur die Tatsache, ob die Glockenkurve breit oder schmal ist.

5. Weiter Entwicklungen

Das System stellt, wie eingangs erwähnt, einen Modellfall dar. In seinem bisherigen Umfang bildet es ein Fundament für weitere vielfältige Entwicklungen. Einige davon sollen an dieser Stelle erwähnt werden:

Einbau verschiedener Tests, im besonderen einseitige und zweiseitige t- Tests.

Heteroskedastie

Modifikationen der Schätzfunktionen R^2 und dw .

Autoregressive Prozesse

Lagverteilungen.

Nichtlineare Regressionen

(Bestehende Programme sollen eingebaut werden [3]).

Quadratische Programmierung.

Auswahl spezieller Gleichungen und Bildung der reduzierten Form.

An eine Verwirklichung dieser Vorhaben ist in naher Zukunft gedacht.

LITERATUR

- [1] JOHNSTON J., econometric methods, international student edition, Mc Graw-Hill Book Company, Inc., 1963
- [2] SCHÖNFELD P., Methoden der Ökonometrie, Band I und II, Verlag Franz Vahlen GmbH, 1969 bzw. 1971
- [3] FLEISSNER P., und HIETLER H., Nichtlineare Ökonometrie, Programme und Anwendungen, Forschungsbericht Nr. 69/1972, Institut für Höhere Studien, Wien.