SESO
STUDIECENTRUM VOOR ECONOMISCH EN SOCIAAL ONDERZOEK

PROGRAMMABESCHRIJVINGEN

D E E L I

E. Borghers

Werknota 7321

September 1973

Universitaire Faculteiten St.-Ignatius
Prinsstraat 13 - 2000 Antwerpen
D/1973/1169/8

De bedoeling van deze werknota is het samenbrengen van een reeks programmabeschrijvingen.

De hier beschreven programma's werden oorspronkelijk opgesteld voor specifieke probleemstellingen in het kader van SESO-werkzaamheden. Omwille van hun meestal ruimere intrinsieke mogelijkheden werden deze programma's echter zô gewijzigd dat een veralgemeend gebruik mogelijk werd gemaakt.

Vermeld dient nog te worden dat in deze reeks vier SUBROUTINES van W. NONNEMAN werden opgenomen, nl. nr. 10 t/m 13.

E. Borghers
14 september 1973.

INHOUD

SUBROUTINES

PROGRAMMA'S

SUBRØUTINE EXVAR (B, SX, NX, E)

1. Omschrijving en methode

De uiteindelijke bedoeling van het algemeen lineair regressiemodel

$$y_t = \beta_o + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_k x_{kt} + \varepsilon_t \qquad (t=1,2,\dots,T)$$

is, b.m.v. het minimeren van

$$S = \sum_{t=1}^{T} \varepsilon_t^2$$

een zo groot mogelijk gedeelte van de variantie van y te verklaren. Het aandeel van elk der verklarende variabelen (excl. de constante) in deze totale verklarende variantie wordt niet enkel bepaald door de geschatte coëfficiënten maar tevens door de variantie van de betreffende verklarende variabele.

Het relatieve aandeel van elk der onafhankelijke variabelen in de totale verklaarde variantie kan dan ook worden weergegeven door de volgende index

$$e_j = \frac{|\hat{\beta}_j| \; \sigma_{x_j}}{\sum_{j=1}^{k} |\hat{\beta}_j| \; \sigma_{x_j}} \qquad (j=1,2,\dots,k)$$

waarin

e_j = het relatieve aandeel van de $j^{\underline{de}}$ verklarende variabele voorstelt

σ_{x_j} = de standaarddeviatie is van de $j^{\underline{de}}$ onafhankelijke variabele.

2. Beschrijving van de parameters

INPUT | B = Real array
(B(I) moet de geschatte coëfficiënt bevatten voor de $I^{\underline{de}}$ verklarende variabele)
DIMENSIØN B(NX)

SX = Real array
(SX(I) moet de standaarddeviatie bevatten van de I$^{\underline{de}}$ verklarende variabele)
DIMENSIØN SX(NX)

NX = Integer constant/variable
Aantal verklarende variabelen exclusief de constante
Restrictie: NX $\geq$ 2

ØUTPUT
E = Real array
(E(I) bevat het relatieve aandeel in de totale verklaarde variantie van de I$^{\underline{de}}$ onafhankelijke variabele.)
DIMENSIØN E(NX)

3. Opmerkingen

Indien ook een constante term werd geschat moet deze op de laatste plaats staan of m.a.w. als laatste element van de real array B. In dit geval geldt echter DIMENSIØN B(NX+1).

Benodigde subroutines: GEEN

SUBRØUTINE GMNRM (V, VN, K)

1. Omschrijving en methode

In deze subroutine wordt een variantie-covariantie-matrix V genormaliseerd tot een matrix V^n en dit door elke covariantie (de niet diagonaalelementen) uit te drukken in functie van de bijbehorende varianties (de diagonaalelementen), nl.

$$v_{ij}^n = \frac{v_{ij}}{\sqrt{v_{ii}}\ \sqrt{v_{jj}}} \qquad (i,j = 1,2, \ldots,k)$$

De resulterende correlatiematrix heeft het grote voordeel dat bij de interpretatie geen rekening meer dient gehouden te worden met de verscheidenheid in gekozen waarnemingseenheden en dimensies.

2. Beschrijving van de parameters

INPUT

- V = Real array

  De variantie-covariantiematrix. Deze matrix staat in vectornotatie met "0" als "storage mode" d.w.z. "general".

  DIMENSIØN V(K*K)

- K = Integer constant/variable

  Aantal rijen en aantal kolommen van de matrices V en VN

ØUTPUT

- VN = Real array

  De genormaliseerde variantie-covariantiematrix. Deze matrix staat in vector notatie met "0" als "storage mode" d.w.z. "general".

  DIMENSIØN VN(K*K).

3. Opmerkingen

Benodigde subroutines: GEEN.

SUBRØUTINE GRAF1(X,N,UNIT)

1. Omschrijving en methode

Tekenen van een grafiek van een set gegeven functie-waarden met zelf te kiezen eenheid op de verticale as.

2. Beschrijving van de parameters

X - Real array

(X(I) moet de functie-waarde voor het I-de punt bevatten)

N - Integer constant/variable

Aantal functie-waarden. Restrictie: $N \leq 50$

UNIT - Real constant/variable

Eenheid op de verticale as. Restrictie: $UNIT \geq 0.0001$

3. Output

Elk punt van de grafiek wordt voorgesteld door "*".
De functie-waarden op de verticale as worden afgedrukt in FØRMAT (E15.6)
Zorg zelf voor nieuwe bladzijde of blanco regels.

4. Opmerkingen

Het aanroepende programma moet voorzien worden van een CØMMØN-statement, nl.

CØMMØN MX, MY

Waarin

MX - Integer constant

"Logical unit number" voor de output

MY - Integer constant

"Logical unit number" voor de input.

SUBRØUTINE GRAF2(X,N,M)

1. Omschrijving en methode

Tekenen van een grafiek van een set gegeven functie-waarden op een zelf te kiezen aantal regels. De eenheid op de verticale as wordt bekomen na schaling.

2. Beschrijving van de parameters

X - Real array
(X(I) moet de functie-waarde voor het I-de punt bevatten)

N - Integer constant/variable
Aantal functie-waarden. Restrictie: $N \leqslant 50$

M - Integer constant/variable
Aantal regels output

3. Output

Elk punt van de grafiek wordt voorgesteld door "*"
De functie-waarden op de verticale as worden afgedrukt in FØRMAT (E15.6). Het effectief aantal regels output bedraagt M+3 regels. Zorg zelf voor nieuwe bladzijde of blanco regels.

4. Opmerkingen

Het aanroepende programma moet voorzien worden van een CØMMØN-statement, nl.

CØMMØN MX,MY

waarin

MX - Integer constant
"Logical unit number" voor de output.

MY - Integer constant
"Logical unit number" voor de input.

SUBRØUTINE GRAF3(Y,X,N,UNIT)

1. Omschrijving en methode

Tekenen van een grafiek van twee sets van functie-waarden met zelf te kiezen eenheid op de verticale as.

2. Beschrijving van de parameters

Y - Real array
    (Y(I) moet de functie-waarde voor het I-de punt van de eerste functie bevatten)

X - Real array
    (X(I) moet de functie-waarde voor het I-de punt van de tweede functie bevatten)

N - Integer constant/variable
    Aantal functie-waarden. Restrictie: N≤25

UNIT - Real constant/variable
    Eenheid op verticale as. Restrictie: UNIT≥0.0001

3. Output

Elk punt van de grafiek m.b.t. de eerste functie wordt voorgesteld door "*" terwijl elk punt van de grafiek m.b.t. de tweede functie wordt voorgesteld door "X".

Bij samenvallende functie-waarden prevaleert de Y-waarde.

De functie-waarden op de verticale as worden afgedrukt in FØRMAT (E15.6)

Zorg zelf voor nieuwe bladzijde of blanco regels.

4. Opmerkingen

Het aanroepende programma moet voorzien worden van een CØMMØN-statement, nl.

    CØMMØN MX,MY
    waarin
    MX - Integer constant
        "Logical unit number" voor de output
    MY - Integer constant
        "Logical unit number" voor de input.

SUBRØUTINE GRAF4(Y,X,N,M)

1. Omschrijving en methode

Tekenen van een grafiek van twee sets van functie-waarden op een zelf te kiezen aantal regels.
De eenheid op de verticale as wordt bekomen na schaling.

2. Beschrijving van de parameters

Y - Real array
(Y(I) moet de functie-waarde voor het I-de punt van de eerste functie bevatten)
X - Real array
(X(I) moet de functie-waarde voor het I-de punt van de tweede functie bevatten)
N - Integer constant/variable
Aantal functie-waarden. Restrictie: $N \leq 25$
M - Integer constant/variable
Aantal regels output.

3. Output

Elk punt van de grafiek m.b.t. de eerste functie owordt voorgesteld door "±" terwijl elk punt van de grafiek m.b.t. de tweede functie wordt voorgesteld door "X".
Bij samenvallende functie-waarden prevaleert de Y-waarde.
De functie-waarden op de verticale as worden afgedrukt in FØRMAT (E15.6).
Het effectief aantal regels output bedraagt M+3 regels.
Zorg zelf voor nieuwe bladzijde of blanco regels.

4. Opmerkingen

Het aanroepende programma moet voorzien worden van een CØMMØN-statement, nl.

```
CØMMØN MX,MY
```

waarin
MX - Integer constant
"Logical unit number" voor de output
MY - Integer constant
"Logical unit number" voor de input.

SUBRØUTINE GRAF5(NØ,A,N,M,NL,NS)

1. Omschrijving en methode

Deze subroutine is een gewijzigde versie van de IBM-subroutine PLØT(*). De wijzigingen hebben betrekking op:

input: de digits (blanco, 1,2,...,9), nodig voor het afdrukken, hoeven niet ingelezen te worden.

output:
- de hoofding (chart number) wordt niet afgedrukt.
- de elf waarden op de horizontale as worden eveneens niet afgedrukt.
- de waarden op de verticale as worden afgedrukt in FØRMAT (F4.0)!
- de grafische voorstelling wordt slechts afgedrukt over 50 i.p.v. 100 kolommen.

2. Beschrijving van de parameters

NØ - Volgnummer van de grafische voorstelling. Dit "chart number" wordt niet afgedrukt.

A - Matrix bestaande uit de grafisch voor te stellen variabelen. Een variabele per kolom. De eerste kolom bevat de basisvariabele. De matrix staat in vectornotatie.

N - Aantal rijen van matrix A (d.w.z. het aantal waarnemingen voor elk der variabelen).

M - Aantal kolommen van matrix A (d.w.z. het aantal variabelen inclusief de basisvariabele). Restrictie: $M \leq 10$.

NL - Aantal regels waarover de grafische voorstelling verdeeld wordt.

NS - Code voor het sorteren van de basisvariabele in stijgende orde van grootte.

NS=o Het sorteren is niet vereist. De basisvariabele is reeds geordend.

NS=1 Het sorteren is wel vereist.

(*) IBM Application Program, 1130 Scientific Subroutine Package, Programmers Manual, Fifth Edition, June 1970, p.156.

3. Output

Elk punt van de K-de niet basisvariabele, d.i. de variabele uit de (K+1)-de kolom van matrix A, wordt afgebeeld door het cijfer K. De punten van de laatste niet basisvariabel, d.i. de variabele uit de N-de kolom van matrix A, worden echter afgebeeld door "Y". De waarden op de verticale as worden afgedrukt in FØRMAT(F4.0) Zorg zelf voor nieuwe bladzijde of blanco regels.

4. Opmerkingen

Het aanroepende programma dient voorzien te worden van een CØMMØN-statement, nl.

```
CØMMØN MX,MY
waarin
MX - Integer constant
     "Logical unit number" voor de output
MY - Integer constant
     "Logical unit numer" voor de input
```

Wanneer men als basisvariabele de gehele getallen neemt van 1 t/m N en wanneer verder geldt dat N=NL en NS=o dan is deze subroutine zeer geschikt voor het tekenen van een grafiek van maximaal negen sets van functiewaarden (vb. tijdreeksen).

SUBRØUTINE GRAF6 (NØ, A, N, M, NL, NS)

1. Omschrijving en methode

Deze subroutine is een gewijzigde versie van de IBM-subroutine PLØT (*). De wijzigingen hebben betrekking op:

input: de digits (blanco, 1, 2, ..., 9), nodig voor het afdrukken, hoeven niet ingelezen te worden daar de grafische voorstelling gebeurt b.m.v. speciale tekens i.p.v. met cijfers (zie verder).

output:
- de hoofding (chart number) wordt niet afgedrukt
- de elf waarden op de horizontale as worden eveneens, niet afgedrukt
- de waarden op de verticale as worden afgedrukt in FØRMAT (F4.0)!
- de grafische voorstelling wordt afgedrukt over 50 of 100 kolommen (zie verder)

2. Beschrijving van de parameters

NØ: Volgnummer van de grafische voorstelling. Dit "chart number" wordt niet afgedrukt

A - Matrix bestaande uit de grafisch voor te stellen variabelen. Een variabele per kolom. De eerste kolom bevat de basisvariabele. De matrix staat in vectornotatie.

N - Aantal rijën van matrix A (d.w.z. het aantal waarnemingen voor elk der variabelen).

M - Aantal kolommen van matrix A (d.w.z. het aantal variabelen inclusief de basisvariabele). Restrictie: M≤5

NL - Aantal regels waarover de grafische voorstelling verdeeld wordt.

NS - Code voor het sorteren van de basisvariabele in stijgende orde van grootte.

NS=0 Het sorteren is niet vereist. De basisvariabele is reeds geordend.

NS=1 Het sorteren is wel vereist.

(*) IBM Application Program, 1130 Scientific Subroutine Package, Programmers Manual, Fifth Edition, June 1970, p. 156.

3. Output

Met deze subroutine kunnen dus maximaal vier variabelen grafisch voorgesteld worden. De grafische voorstelling gebeurt b.m.v. de speciale tekens: "+", "x", ".", "±". De punten van de laatste variabele, d.i. de variabele uit de N-de kolom van matrix A, worden evenwel steeds afgebeeld b.m.v. "±".
De waarden op de verticale as worden afgedrukt in FØRMAT (F4.0).
De grafische voorstelling wordt afgedrukt over

50 kolommen indien NL $\leq$ 100

100 kolommen indien NL > 100

Zorg zelf voor nieuwe bladzijde of blanco regels.

4. Opmerkingen

Het aanroepende programma dient voorzien te worden van een CØMMØN-statement, nl. CØMMØN MX, MY
waarin

MX - Integer constant
"Logical unit number" voor de output

MY - Integer constant
"Logical unit number" voor de input.

Wanneer men als basisvariabele de gehele getallen neemt van 1 t/m N en wanneer verder geldt dat N=NL en NS=0 dan is deze subroutine zeer geschikt voor het tekenen van een grafiek van maximaal vier sets van functiewaarden (vb. tijdreeksen).

SUBRØUTINE GRAFH(Z,N,IN)

1. Omschrijving en methode

Van een set getallen wordt, bij een gegeven aantal intervallen, de frequentieverdeling bepaald en onder de vorm van een histogram grafisch weergegeven.

De intervallen zijn beneden gesloten en boven open. Het kleinste getal is de ondergrens van het eerste interval en het grootste getal is de bovengrens van het laatste interval.

2. Beschrijving van de parameters

Z - Real array
(Z(I) moet het I-de getal bevatten)

N - Integer constant/variable
Aantal elementen van Z-array, d.w.z. het totaal aantal getallen

IN - Integer constant/ variable
Aantal intervallen. Restrictie: IN$\leq$20

3. Output

Voor de grafische voorstelling werd gebruik gemaakt van de IBM-subroutine HIST (*). De afgedrukte output is dan ook identiek aan de output van deze subroutine. Enkel de hoofding werd weggelaten. Zorg zelf voor nieuwe bladzijde of blanco zegels.

4. Opmerkingen

Het aanroepende programma moet voorzien worden van een CØMMØN-statement, nl.

CØMMØN MX, MY

waarin

MX - Integer constant
"Logical unit number" voor de output

MY - Integer constant
"Logical unit number" voor de input.

(*) IBM Application Program, 1130 Scientific Subroutine Package Programmers Manual, Fifth Edition, June 1970, p.145.

SUBRØUTINE TRCØN (X,N,C,S)

1. Omschrijving en methode

Deze subroutine berekent de gesommeerde gekwadrateerde afwijkingen van een reeks van N waarnemingen t.o.v. het gemiddelde van deze reeks, nl.

$$S = \sum_{t=1}^{N} (X_t - C)^2 \qquad (t=1,N)$$

waarin

$$C = \frac{1}{N} \sum_{t=1}^{N} X_t$$

Het gemiddelde C kan beschouwd worden als een constante trend.

2. Beschrijving van de parameters

INPUT
- X - Real array
  X(I) moet de Ide waarneming van de array bevatten.
  DIMENSIØN: X(N)
- N - Integer constant/variable
  Aantal elementen van de array X

ØUTPUT
- C - Gemiddelde van de array X
- S - Gesommeerde gekwadrateerde afwijkingen t.o.v. het gemiddelde C

3. Opmerkingen

Geen.

SUBRØUTINE TREXP (X, N, A, B, S)

1. Omschrijving en methode

Deze subroutine berekent b.m.v. de methode der kleinste kwadraten de parameters α en β van de exponentiële trendvergelijking

$$X_t = \beta e^{\alpha t} \qquad (t=1,N)$$

alsook de gesommeerde gekwadrateerde afwijkingen

$$S = \sum_{t=1}^{N} (X_t - \hat{X}_t)^2$$

2. Beschrijving van de parameters

INPUT
- X - Real array
  X(I) moet de Ide waarneming van de array bevatten
  DIMENSIØN : X(N)
- N - Integer constant/variable
  Aantal elementen van de array X

ØUTPUT
- A - Kleinste kwadraten schatter voor α
- B - Kleinste kwadraten schatter voor β
- S - Gesommeerde gekwadrateerde afwijkingen t.o.v. de berekende trendwaarde.

3. Opmerkingen

Alle waarnemingen van de array X moeten positief zijn (om reden van de logarithmering).

SUBRØUTINE TRLIN (X, N,A, B,S)

1. Omschrijving en methode

Deze subroutine berekent b.m.v. de methode der kleinste kwadraten de parameters α en β van de lineaire trendvergelijking

$$X_t = \alpha t + \beta \qquad\qquad (t=1,N)$$

Alscok de gesommeerde gekwadrateerde afwijkingen

$$S = \sum_{t=1}^{N} (X_t - \hat{X}_t)^2$$

2. Beschrijving van de parameters

INPUT
- X - Real array
  - X(I) moet de Ide waarneming van de array bevatten
  - DIMENSIØN : X(N)
- N - Integer constant/variable
  - Aantal elementen van de array X.

ØUTPUT
- A - Kleinste kwadraten schatter voor α.
- B - Kleinste kwadraten schatter voor β.
- S - Gesommeerde gekwadrateerde afwijkingen t.o.v. de berekende trendwaarde.

3. Opmerkingen

Geen.

SUBRØUTINE TRPAR (X, N, A, B, G, S)

1. Omschrijving en methode

Deze subroutine berekent b.m.v. de methode der kleinste kwadraten de parameters α, β en γ van de parabolische trendvergelijking

$$X_t = \alpha t^2 + \beta t + \gamma \qquad (t=1,N)$$

alsook de gesommeerde gekwadrateerde afwijkingen

$$S = \sum_{t=1}^{N} (X_t - \hat{X}_t)^2$$

2. Beschrijving van de parameters

INPUT
- X - Real array
  X(I) moet de Ide waarneming van de array bevatten
  DIMENSIØN : X(N)
- N - Integer constant/variable
  Aantal elementen van de array X

ØUTPUT
- A - Kleinste kwadraten schatter van α
- B - Kleinste kwadraten schatter van β
- G - Kleinste kwadraten schatter van γ
- S - Gesommeerde gekwadrateerde afwijkingen t.o.v. de trendwaarde.

3. Opmerkingen

Geen

SUBRØUTINE TRPRT (PR, T, K, N)

1. Omschrijving en methode

In deze subroutine wordt voor elk van de partiële correlatiecoëfficiënten uit matrix PR een grootheid berekend die, onder zekere voorwaarde, t-verdeeld is. De voorwaarde waaraan voldaan zou moeten worden heeft betrekking op de matrix met de variabelen die geleid heeft tot de partiële correlatiecoëfficiënten. Deze matrix, bestaande uit de N waarnemingen voor elk van de K variabelen, moet namelijk multinormaal verdeeld zijn. De gebruikte formule luidt

$$t_{ij} = \frac{r_{ij,\ldots,k} \sqrt{N-k}}{\sqrt{1-r^2_{ij,\ldots,k}}} \qquad (i,j = 1,2, \ldots,k)$$

waarin

t_{ij} = t-verdeeld is met (N-k) vrijheidsgraden

$r_{ij,\ldots,k}$ = de partiële correlatiecoëfficiënt voorstelt tussen variabele i en j

N = het aantal waarnemingen is voor elk van de k-variabelen

k = het aantal variabelen is.

2. Beschrijving van de parameters

INPUT
- PR = Real array
  De matrix met partiële correlatiecoëfficiënten. Deze matrix staat in vector notatie met "0" als "storage mode" d.w.z. "general".
  DIMENSIØN: PR(K*K)
- K = Integer constant/variable
  Aantal rijen en aantal kolommen van de matrices PR en T
- N = Integer constant/variable
  Aantal waarnemingen op elk van de K variabelen

ØUTPUT
- T = Real array
  De matrix met de t-waarden. Deze matrix staat in vector-notatie met "0" als "storage mode" d.w.z. "general".
  DIMENSIØN: T(K*K).

3. Opmerkingen

Benodigde subroutines: LØC

SUBRØUTINE TRRPR (R, PR, K)

1. Omschrijving en methode

In deze subroutine worden partiële correlatiecoëfficiënten berekend b.m.v. de elementen van de inverse correlatiematrix. De formule luidt

$$r_{ij,\ldots,k} = \frac{-\gamma_{ij}}{\sqrt{\gamma_{ii}}\,\sqrt{\gamma_{jj}}} \qquad (i,j = 1,2, \ldots, k)$$

waarin

$r_{ij,\ldots,k}$ = de partiële correlatiecoëfficiënt voorstelt, tussen variabele i en j

γ_{ij} = het $(i,j)^{\underline{de}}$ element is van de inverse correlatiematrix

2. Beschrijving van de parameters

INPUT
- R = Real array
  De inverse van de correlatiematrix. Deze matrix staat in vector notatie met "0" als "storage mode" d.w.z. "general".
  DIMENSIØN: R(K*K)
- K = Integer constant/variable
  Aantal rijen en aantal kolommen van de matrices R en PR.

OUTPUT
- PR = Real array
  De matrix met partiële correlatie-coëfficiënten. Deze matrix staat in vector notatie met "0" als "storage mode" d.w.z. "general".
  DIMENSIØN: PR(K*K)

3. Opmerkingen

Benodigde subroutines: LØC

SUBRØUTINE TSTAT (Y,YC,N,K,NC,E,U,S2,ARØ)

1. Omschrijving en methode

De bedoeling van deze subroutine is additionele informatie te verschaffen m.b.t. de regressievergelijking

$$y = X\beta + \varepsilon$$

en meer bepaald in die situatie waarin we enkel beschikken over de waargenome en berekende waarden van de te verklaren variabele.

2. Beschrijving van de parameters

INPUT

Y = Real array
De te verklaren variabele.
(Y(I) moet de Ide waarneming bevatten van de variabele Y)
DIMENSIØN Y(N)

YC = Real array
De berekende waarden voor de te verklaren variabele.
(YC(I) moet de Ide waarneming bevatten van de berekende Y-variabele)
DIMENSIØN YC(N)

N = Integer constant/variable
Aantal waarnemingen voor de te verklaren variabele Y

K = Integer constant/variable
Aantal verklarende variabelen, exclusief de constante, welke geleid hebben tot de berekende YC-waarden van de te verklaren variabele Y.

NC = 0 De regressievergelijking bevatte geen constante.
= 1 De regressievergelijking bevatte wel een constante.

ØUTPUT

E = Real array
De berekende residuën
(E(I)=YC(I)-Y(I))
DIMENSIØN E(N)

U = Real array
De berekende gestandaardizeerde residuën
DIMENSIØN U(N)

S2 = Real variable
De berekende restvariantie

ARØ = Real array
Deze array bevat acht schattingen voor de auto-correlatiecoëfficiënt (Markov-schema van de eerste orde) nl.: RØG, CRØG, RØL, CRØL, RØM, CRØM, RØS, CRØS. (In deze volgorde !!!)
Wat de juiste definitie van deze schatters betreft kan verwezen worden naar de outputbeschrijving.
DIMENSIØN ARØ(8)

3. Output

De output van deze subroutine is zowel qua inhoud als qua lay-out identiek aan deze welke beschreven staat in

BORGHERS E., "A Summary Statement of the General Linear Model" (With a Fortran Program), SESO-werknota, 7106, 58pp.

en wel meer bepaald op bladzijde 45 (Page 6) en 46 (Page 7).

4. Opmerkingen

Benodigde subroutines:

- GAMMA (*)
- WSNS (**)

Het aanroepende programma moet voorzien worden van een CØMMØN-statement, nl. CØMMØN MX, MY
waarin

MX - Integer constant
"Logical Unit number" voor de output.

MY - Integer constant
"Logical Unit number" voor de input.

De benodigde geheugenlengte voor deze subroutine is vrij groot, nl.

CØMMØN : 2
VARIABLES: 122
PRØGRAM : 1296

(*) IBM Application Program, 1130 Scientific Subroutine Package, Programmes Manual, Fifth Edition, June 1970, p.

(**) Rekencentrum - subroutine.

SUBRØUTINE VTRNS (NC, X, Y, CØNST, N, NER)

1. Omschrijving en methode

Deze subroutine maakt het mogelijk een gegeven real array (variabele) te transformeren op een van de volgende 24 mogelijkheden:

Code	Transformatie	
NC=1	X(I)=1./X(I)	
=2	X(I)=X(I)**2	
=3	X(I)=X(I)**3	
=4	X(I)=SQRT(X(I9))	
=5	X(I)=ALØG(X(I))	($\log_e X_i$)
=6	X(I)=ALØG(X(I))*0.4343	($\log_{10} X_i$)
=7	X(I)=CØS(X(I))	
=8	X(I)=SIN(X(I))	
=9	X(I)=ABS(X(I))	
=10	X(I)=2.718282**X(I)	(e^{X_i})
=11	X(I)=X(I)+CØNST	
=12	X(I)=X(I)-CØNST	
=13	X(I)=X(I)*CØNST	
=14	X(I)=X(I)/CØNST	
=15	X(I)=X(I)**CØNST	
=16	X(I)=X(I)+Y(I)	
=17	X(I)=X(I)-Y(I)	
=18	X(I)=X(I)*Y(I)	
=19	X(I)=X(I)/Y(I)	
=20	X(I)=X(I)**Y(I)	
=21	X(I)=X(I)-X(I-1)=ΔX(I)	
=22	X(I)=(X(I)-X(I-1))-(X(I-1)-X(I-2))=Δ^2X(I)	
=23	X(I)=((X(I)-X(I-1))/X(I-1))*100.=$\mathring{X}$(I)	
=24	X(I)=$\mathring{X}$(I)-$\mathring{X}$(I-1)	

2. Beschrijving der parameters

INPUT

NC = Integer constant/variable

Codenummer voor de gekozen transformatie

X = Real array

De te transformeren variabele (X(I) moet de Ide waarneming bevatten van variabele X)

DIMENSIØN: X(N)

Y = Real array

Variabele noodzakelijk bij de transformatie van X volgens de transformatiecodes 16 t/m 20.

(Y(I) moet de Ide waarneming bevatten van variabele Y)

DIMENSIØN: Y(N)

Indien deze hulpvariabele niet gebruikt wordt moet wel deze parameter ingevuld worden maar dient de Y-array echter niet gedimmensioneerd te worden.

CØNST = Real constant

Constante noodzakelijk bij de transformatie van X volgens de transformatiecodes 11 t/m 15.

Indien deze constante niet gebruikt wordt dan moet deze nochtans toch als een dummy parameter ingevuld worden.

N = Integer constant/variable

Aantal waarnemingen van zowel de X- als de Y-variabele.

ØUTPUT

X = Real array

De getransformeerde variabele (X(I) bevat de Ide waarneming van de getransformeerde X-variabele)

DIMENSIØN: X(N)

NER = Integer constant = $I_4 I_3 I_2 I_1$

"Error Code"

$I_4 I_3$ geeft het nummer aan van het element van de X-array waarvoor de transformatie $I_2 I_1$ niet kon uitgevoerd worden.

3. Bemerkingen

Bij het verlaten van de subroutine staat het eerste nieuw gedefiniëerde element van de X-array op de tweede plaats voor wat de transformaties 21 en 23 betreft en op de derde plaats voor wat de transformaties 22 en 24 betreft. De getransformeerde array heeft dus zijn oorspronkelijke lengte N behouden. Zorg dus zelf, indien nodig, voor het weglaten van het eerste, respectievelijk tweede element van de output array X.

De benodigde geheugenlengte voor deze subroutine is vrij groot,

nl. CØMMØN : 0
VARIABLES : 10
PRØGRAM : 1120

PRØGRAMMA DS2ØD

1. Omschrijving en methode

Dit programma geeft de algemene oplossing voor het homogene gedeelte van

(1) $$y_t + \beta_1 y_{t-1} + \beta_2 y_{t-2} = \delta$$

d.i. een differentievergelijking van de tweede orde met vaste coëfficiënten.

De karakteristieke vergelijking van het homogene gedeelte van deze differentievergelijking

(2) $$y_t + \beta_1 y_{t-1} + \beta_2 y_{t-2} = 0$$

kan geschreven worden als

(3) $$Z^2 + \beta_1 Z + \beta_2 = 0$$

De algemene oplossing van de differentievergelijking kan voorgesteld worden door

(4) $$y_t = y_t^h + y_t^p$$

waarin de eerste term van het rechter lid het homogene gedeelte van de oplossing is en de tweede term de particuliere oplossing voorstelt.

De vorm van het homogene gedeelte y_t^h van de algemene oplossing is echter afhankelijk van het al dan niet gelijk zijn van de wortels uit de karakteristieke vergelijking (3). De twee gevallen met hun respectievelijke oplossing zijn

- twee gelijke wortels ($m_1 = m_2 = m$)

(5) $$y_t^h = (C_1 + C_2 t)m^t$$

- twee ongelijke wortels ($m_1 \neq m_2$)

(6) $$y_t^h = C_1 m_1^t + C_2 m_2^t$$

waarin C_1 en C_2 constanten voorstellen die afhankelijk zijn van de aanvangstoestand.

Indien echter de discriminant uit de karakteristieke vergelijking negatief is ($\beta_1^2 < 4\beta_2$) dan zijn de wortels niet enkel ongelijk maar bovendien ook nog toegevoegd complex nl.

(7) $$m_1 = r(\cos\phi + i \sin\phi) = a + ib$$
$$m_2 = r(\cos\phi - i \sin\phi) = a - ib$$

Ook de constanten C_1 en C_2 kunnen dan als toegevoegd complex geschreven worden als

(8) $$C_1 = K(\cos\alpha + i \sin\alpha)$$
$$C_2 = K(\cos\alpha - i \sin\alpha)$$

Gebruik makend van het theorema van DE MOIVRE en de definitie voor het vermenigvuldigen van complexe getallen kan (6) herschreven worden als

(9) $$y_t^h = 2Kr^t \cos(\phi t + \alpha)$$
of
(10) $$y_t^h = r^t(C_1' \cos\phi t + C_2' \sin\phi t)$$

waarin r de modulus voorstelt en C_1', C_2' en K constanten zijn, afhankelijk van de begintoestand.

2. Input

De input bestaat uit een parameterkaart en een datakaart.

* Parameterkaart (FØRMAT(2I5))

Kol.

1 - 5 NØ = Integer constant

NØ>0 Nummer van het probleem

NØ=0 CALL EXIT

6 -10 NGRAF = Integer constant

NGRAF = 1 Er worden 50 functiewaarden berekend en op grafiek uitgetekend.

NGRAF = 0 Geen berekende functiewaarden en geen grafische voorstelling.

* Datakaart (FØRMAT(4F10.0))

Kol.

1 -10 BETA1 = Real constant

De coëfficiënt β_1 van y_{t-1} (zie (1))

Indien de term y_{t-1} ontbreekt dan is BETA1=C

11-20 BETA2 = Real constant

De coëfficiënt β_2 van y_{t-2} (zie (1))

21-30 Y0 = Real constant

De waarde van y_t voor t=o (Aanvangstoestand voor het tweede orde auto-regressieve schema)

31-40 Y1 = Real constant

De waarde van y_t voor t=1.

3. Output

Bij de beschrijving van de output dient een onderscheid gemaakt te worden voor het geval waarin de discriminant positief is ($\beta_1^2>4\beta_2$) en het geval waarin de discriminant negatief is ($\beta_1^2<4\beta_2$).

3.1. Discriminant is positief

- Hoofding met nummer van het probleem (NØ)
- De coëfficiënt van y_{t-1} nl. β_1 (zie (1))
- De coëfficiënt van y_{t-2} nl. β_2 (zie (1))
- De waarde van de discriminant
- De eerste wortel m_1 (zie (6))
- De tweede wortel m_2 (zie (6))

- De constante C_1 (zie (6))
- De constante C_2 (zie (6))

Indien NGRAF=1 dan worden op een tweede bladzijde de 50 functiewaarden, berekend volgens de formule

$$(6) \qquad y_t^h = C_1 m_1^t + C_2 m_2^t \qquad (t = 0, 1, 2, \ldots, 49)$$

afgedrukt onder de hoofding "FUNCTIØN VALUES" en op een derde bladzijde uitgetekend onder de hoofding "FUNCTIØN GRAPH".

3.2. Discriminant is negatief

- Hoofding met nummer van het probleem (NØ)
- De coëfficiënt van y_{t-1} nl. β_1 (zie (1))
- De coëfficiënt van y_{t-2} nl. β_2 (zie (1))
- De waarde van de discriminant
- Het reëele gedeelte van de wortels n.l. a (zie (7))
- Het immaginaire gedeelte van de wortels n.l. b (zie (7))
- De modulus r (zie (7), (9) en (10))
- De waarde van $\cos\phi$ (zie (7), (9) en (10))
- De waarde van $\sin\phi$ (zie (7), (9) en (10))
- De hoekfrequentie $\phi=2\Pi/P$ (zie (7), (9) en (10))
- De frequentie $\lambda=1/P$
- De periode $P=2\Pi/\phi$
- De constante C_1' (zie (10))
- De constante C_2' (zie (10))
- De constante K (zie (9))
- Beginamplitudo y_0^h (zie (9) en (10))
- De phase α (zie (9)).

Indien NGRAF=1 dan worden op een tweede bladzijde de 50 functiewaarden, berekend volgens de formule

$$(9) \qquad y_t^h = 2Kr^t\cos(\phi t+\alpha) \qquad (t = 0, 1, 2, \ldots, 49)$$

of

$$(10) \qquad y_t^h = r^t(C_1'\cos\phi t + C_2'\sin\phi t) \qquad (t = 0, 1, 2, \ldots, 49)$$

afgedrukt onder de hoofding "FUNCTIØN VALUES" en op een derde bladzijde uitgetekend onder de hoofding "FUNCTIØN GRAPH".

4. Opmerkingen

Het programma "DS2ØD" is een programma dat b.m.v. "CALL LINK" instructies ingebouw werd in een sequentie van programma's. Het kan nochtans, op de hier beschreven wijze, toch AFZONDERLIJK gebruikt worden op voorwaarde evenwel dat "SWITCH 1" in "ØN" stand geplaatst wordt alvorens het programma te starten.

Benodigde subroutines: GETA, GRAF2 (*)

(*) SESO-subroutines

PRØGRAMMA RCØAN

1. Omschrijving en methode

Een vaak voorkomend probleem is het toetsen of het lineair verband tussen een variabele y en een set van K X-variabelen significant verschillend is in M onafhankelijke steekproeven.

Als uitgangspunt voor dit probleem geldt volgend algemeen model

$$(1) \qquad y_{ij} = \beta_{iK} + \sum_{\ell=1}^{K-1} \beta_{i\ell}\, x_{ij\ell} + \varepsilon_{ij} \qquad \begin{cases} i = 1, 2, \ldots, M \\ j = 1, 2, \ldots, N_i \\ \ell = 1, 2, \ldots, K-1 \end{cases}$$

De meest gebruikelijke benadering bij dit probleem gebeurt in drie stadia: In een eerste fase worden de residuële varianties vergeleken. Blijkt uit deze toetsing dat de homogeniteit van de restvarianties niet kan verworpen worden dan, en slechts dan, heeft het zin de regressiecoëfficiënten zelf te toetsen. De tweede fase bestaat in het toetsen van de "slopes" of m.a.w. de regressiecoëfficiënten maar dan exclusief de constanten. In een derde en laatste stadium wordt dan getoetst of de constanten (de "intercepts") significant verschillend zijn.

1.1. Nadere toelichting bij de eerste fase

Voor twee steekproeven (M=2) gebeurt het toetsen van de residuële varianties b.m.v. de volgende F-toets.

$$(2) \qquad F_{d_2}^{d_1} = \frac{MS_1}{MS_2}$$

Voor meer dan twee steekproeven (M>2) wordt de volgende χ^2-toets gebruikt.

$$(3) \qquad \chi^2_{(M-1)} = \frac{T}{C}$$

waarin $T = 2.3026\{(\Sigma d_i)\log(\frac{\Sigma SS_i}{\Sigma d_i}) - \Sigma d_i \log MS_i\}$

$$C = 1 + \frac{1}{3(M-1)}\ \{\Sigma\frac{1}{d_i} - \frac{1}{\Sigma d_i}\}$$

en waarin alle sommaties over i lopen van 1 t/m M.
Verder geldt voor (2) en (3)

$2.3026 = \log_e 10$

N_i = aantal waarnemingen op al de variabelen van de i^{de} vergelijking

K = aantal verklarende variabelen inclusief de constante term

M = aantal vergelijkingen of m.a.w. het aantal onafhankelijke steekproeven

$d_i = N_i - K$

SS_i = gesommeerde gekwadrateerde residuën (Sum of Squares) van de afzonderlijk geschatte i^{de} vergelijking

$MS_i = SS_i/d_i$ (Mean Squares)

1.2. Nadere toelichting bij de tweede fase

Deze tweede fase bestaat in het toetsen van de volgende hypothesen:

1. De M vergelijkingen verschillen van elkaar zowel in de constante term als in de overige (K-1) coëfficiënten.
2. De M vergelijkingen verschillen van elkaar enkel in de constante term terwijl de overige (K-1) coëfficiënten als gemeenschappelijk kunnen beschouwd worden.

De eerste hypothese kan, voor de M vergelijkingen samen, symbolisch voorgesteld worden als

(4) $\quad y = Z\gamma + U$

waarin y = $(Nx1)$ vector bestaande uit de M subvectoren van elk $(N_i x1)$

Z = $(Nx(MxK))$ blokdiagonaal matrix met op de hoofddiagonaal de M X_i matrices ($i = 1, 2, \ldots, M$)

= $((MxK)x1)$ vector bestaande uit de M subvectoren van elk $(N_i x1)$

U = $(Nx1)$ vector bestaande uit de M subvectoren van elk $(N_i x1)$.

of in zijn geschatte vorm als:

(5) $\quad y = Z\hat{\gamma} + r$

De tweede hypothese uit deze tweede fase kan symbolisch voorgesteld worden als

(6) $\quad y = D\delta + X\alpha + U$

waarin y = $(Nx1)$ vector bestaande uit de M subvectoren van elk $(N_i x1)$

D = $(Nx(M-1))$ matrix met dummy's $D' = (D_1 D_2 \ldots D_M)'$
De $(i-1)$de kolom van de $(N_i x(M-1))$ submatrix D_i is een eenheidsvector terwijl de overige kolommen nulvectoren zijn.

δ = $(M-1)$ vector
$\delta = (\delta_2 \delta_3 \ldots \delta_M)$

X = (NxK) matrix bestaande uit de M submatrices van elk $(N_i xK)$

α = $(Kx1)$ vector

U = $(Nx1)$ vector bestaande uit de M subvectoren van elk $(N_i x1)$.

of in zijn geschatte vorm als:

(7) $\quad y = D\hat{\delta} + X\hat{\alpha} + e$

Gebruik makend van deze notatie kan voor deze tweede fase de volgende covariantie-analyse tabel opgesteld worden:

	"Sum of Squares"	d.f.	"Mean Squares"
Hypothese 1	$SS_1 = r'r$	$d_1 = N - (MxK)$	$MS_1 = SS_1/d_1$
Verschil	$SS_2 = e'e - r'r$	$d_2 = (M-1)(K-1)$	$MS_2 = SS_2/d_2$
Hypothese 2	$SS_3 = e'e$	$d_3 = N-M-(K-1)$	$MS_3 = SS_3/d_3$

De vooropgestelde hypothese, d.w.z. zelfde "slopes" maar verschillende "intercepts", kan dan getoetst worden b.m.v. de volgende F-toets:

$$(8) \qquad F^{d_2}_{d_1} = \frac{SS_2/d_2}{SS_1/d_1}$$

1.3. Nadere toelichting bij de derde fase

Deze derde fase bestaat in het toetsen van volgende hypothesen:

1. De M vergelijkingen verschillen van elkaar enkel in de constante term terwijl de overige (K-1) coëfficiënten als gemeenschappelijk kunnen beschouwd worden.
2. Voor elk van de M vergelijkingen zijn zowel de "intercepts" als de "slopes" niet significant van elkaar verschillend.

De eerste van deze twee hypothesen werd reeds symbolisch weergegeven bij de toelichting van de tweede fase (zie (6) en (7)).

De tweede hypothese kan symbolisch voorgesteld worden als

$$(9) \qquad y = X\beta + U$$

waarin y = (Nx1) vector bestaande uit de M subvectoren van elk $(N_i x1)$

X = (NxK) matrix bestaande uit de M submatrices van elk $(N_i xK)$

β = (Kx1) vector

U = (Nx1) vector bestaande uit de M subvectoren van elk $(N_i x1)$

of in zijn geschatte vorm als

$$(10) \qquad y = X\hat{\beta} + e$$

Gebruik makend van deze notatie kan voor deze derde fase de volgende covariantie-analyse tabel opgesteld worden:

	"Sum of Squares"	d.f.	"Mean Squares"
Hypothese 2	$SS_3 = e'e$	$d_3 = N-M-(K-1)$	$MS_3 = SS_3/d_3$
Verschil	$SS_4 = s's-e'e$	$d_4 = M-1$	$MS_4 = SS_4/d_4$
Hypothese 3	$SS_5 = s's$	$d_5 = N-K$	$MS_5 = SS_5/d_5$

De vooropgestelde hypothese, d.w.z. verschillende "intercepts" maar gelijke "slopes", wordt dan getoetst b.m.v. de volgende F-toets:

(11) $$F_{d_3}^{d_4} = \frac{SS_4/d_4}{SS_3/d_3}$$

Voor verdere referentie m.b.t. de methode zie:

SNEDECOR G. & COCHRAN W., "Statistical Methods", The Iowa State University Press, Ames, Iowa, 1967, pp. 593, Chapter XIV.

JOHNSTON J., "Econometric Methods", McGraw-Hill, New-York, 1972, 2nd. Edition, pp. 192-207.

Voor een toepassing zie:

PHLIPS L., "Effects of Industrial Concentration", Contributions to Economic Analysis 74, North-Holland Publ. Company, Amsterdam, 1971, pp. 18-25.

2. Input

De input van data geschiedt niet via kaarten maar wel via het schijvengeheugen. Hieruit volgt dat de voor dit programma benodigde data allereerst op de daartoe gereserveerde schijf geplaatst moeten worden b.m.v. het programma "ØNDSK" (1). Wat het programma "RCØAN" betreft is de capaciteit van de data-bank "DBANK", die voorzien is op deze gereserveerde schijf, volledig beschikbaar.

De input van de parameters bestaat uit een algemene parameterkaart gevolgd door zoveel parameterkaarten als er vergelijkingen zijn.

* Algemene parameterkaart (FØRMAT (16I5))

Kol.		
1 - 5	NO	= Nummer van het probleem
6 - 10	K	= Aantal verklarende variabelen, inclusief de constante Restrictie: K $\leq$ 10
11 - 15	N	= Aantal vergelijkingen Restrictie: M $\leq$ 10

* Parameterkaart voor elk van de M vergelijkingen (FØRMAT (16I5))

Kol.		
1 - 5	N(I)	= Aantal waarnemingen voor elk van de variabelen van vergelijking I. Restrictie: N(I) $\leq$ 50
6 - 10	KEYS(1)	= Sectornummer van de eerste verklarende variabele van vergelijking I
⋮	⋮	
6+5*(K-1) t/m 10+5*(K-1)	KEYS(K)	= Sectornummer van de laatste verklarende variabele van vergelijking I of m.a.w. de sectornummer van de constante van de Ide vergelijking
6+5*K t/m 10+5*K	KEYS(K+1)	= Sectornummer van de te verklaren variabele van vergelijking I

---

(1) BORGHERS E., "A Summary Statement of the General Linear Model" (With a Fortran Program), SESO-werknota, 7106, p. 56.

Het totaal aantal parameterkaarten bedraagt dus (M+1). Deze input wordt zo dikwijls herhaald als er te verwerken problemen zijn.
Sluitkaart: blanco.

3. Output

* Bladzijde 1

- Hoofding - Nummer van het probleem (NØ)
  - Aantal verklarende variabelen, inclusief de constante (NX)
  - Aantal vergelijkingen (NE)
- Voor elk van de M vergelijkingen
  - Volgnummer (EQ)
  - Aantal waarnemingen (NT)
  - Sectornummer van de te verklaren variabele (Y)
  - Sectornummers van de K verklarende variabelen (X(K)).
    De constante staat op de laatste plaats.

* Bladzijde 2

- Hoofding: idem als bladzijde 1
- Voor elk van de M vergelijkingen (zie (4) en (5))
  - Volgnummer
  - De schatting van de K coëfficiënten
    De constante staat op de laatste plaats
- Op de (M+1)de regel
  - "CØMB"
  - De K geschatte coëfficiënten van de gecombineerde regressievergelijking (zie (9) en (10))
    De constante staat op de laatste plaats.
- Voor elk van de M vergelijkingen (zie (6) en (7))
  - Volgnummer
  - De schatting van de K coëfficiënten
    De wijzigende constante staat op de laatste plaats.

± Bladzijde 3 (Eerste fase) (zie paragraaf 1.1)

- Hoofding: "Testing the homogeneity of the residual variances"
- Voor elk van de M vergelijkingen
  - Volgnummer
  - Vrijheidsgraden ($DF_i = N_i - K$)
  - "Sum of Squares" (SS_i)
  - "Mean Squares" ($MS_i = SS_i/DF_i$)
- Op de $(M+1)^{de}$ regel
  - "TØT"
  - Vrijheidsgraden (ΣDF_i)
  - "Sum of Squares" (ΣSS_i)
  - "Mean Squares" ($\Sigma SS_i/\Sigma DF_i$)
- Toets op de homogeniteit van de residuele varianties m.a.w.
  - χ^2-toets met bijbehorende vrijheidsgraad voor de situatie $M > 2$ (zie (3))
  - F-toets met bijbehorende vrijheidsgraden voor de situatie $M = 2$ (zie (2))

± Bladzijde 4 (Tweede fase) (zie paragraaf 1.2)

- Hoofding: "Comparison of Slopes"
- Covariantie-analyse tabel bestaande uit vrijheidsgraden (DF), "Sum of Squares" (SS) en "Mean Squares" (MS) en dit voor
  - Hypothese 1 (zie (4) en (5))
    ("Total Separate")
  - Verschil tussen hypothese 1 en 2
    ("Diff. betw. Slopes")
  - Hypothese 2(zie (6) en (7))
    ("Common Slopes")
- F-toets met bijbehorende vrijheidsgraden (zie (8))

* Bladzijde 5 (derde fase) (zie paragraaf 1.3)

- Hoofding: "Comparison of Intercepts"
- Covariantie-analyse tabel bestaande uit vrijheidsgraden (DF), "Sum of Squares" (SS) en "Mean Squares" (MS) en dit voor
  - Hypothese 2 (zie (6) en (7)) ("Common Slopes")
  - Verschil tussen hypothese 2 en 3 ("Diff. betw. Intercepts")
  - Hypothese 3 (zie (9) en (10)) ("Combined")
- F-toets met bijbehorende vrijheidsgraden (zie (11))

4. Opmerkingen

Voor dit programma gelden de volgende restricties:

$N_i \leqslant 50 \quad (i = 1, 2, \ldots, M)$

$K \leqslant 10$

$M \leqslant 10$

$KEYS(I) \leqslant 100 \quad (I = 1, 2, \ldots, K)$

$(M*K+1) \leqslant 100$ voor $N_i \leqslant 25$

$(M*K+1) \leqslant 50$ voor $N_i \leqslant 50$

Deze laatste twee restricties zijn zô gedefinieerd dat voor elk van de M vergelijkingen het fictieve adres KEYS(K), d.w.z. het sectornummer waarop de constante werd geplaatst, hetzelfde is. Hieruit volgt dat de lengte van de array van deze constante gelijk moet zijn aan het maximum aantal waarnemingen dat in deze analyse voorkomt (max N_i voor $i = 1, 2, \ldots, M$)

Benodigde subroutines :

-MINV, MATA, GMADD, GTPRD, GMPRD, MSTR, SCLA, LØC (*)

-PUTA, GETA (**)

(*) IBM Application Program, 1130 Scientific Subroutine Package, Programmers Manual, Fifth Edition, June 1970.

(**) SESO-subroutines.

PRØGRAMMA RØØT1

1. Omschrijving en methode

Berekenen van de eigenwaarden van een reële, niet noodzakelijk symmetrische doch vierkante matrix A.
De matrix A wordt getransformeerd tot een matrix van HESSENBERG waaruit de karakteristieke polynoom wordt afgeleid. De wortels van deze polynoom, die ook complex kunnen zijn, geven dan de eigenwaarden van de oorspronkelijke matrix A.
Voor het berekenen van de reëele en complexe wortels van de reëele polynoom werd gebruik gemaakt van de IBM-subroutine PØLRT (*).
Voor methode en verdere referentie, zie:

WILKINSON J.H., "The Algebraic Eigenvalue Problem", Clarendon Press, Oxford, 1965, pp.662. Chapter VI and VII.

FLEISSNER P. & HIETLER K., "Stabilität linearer ökonometrischer Modelle (2. Teil)", Computing, Vol.10, Fasc.1-2, 1972, pp.33-62.

Het programma bestaat uit twee afzonderlijke deelprogramma's die aan elkaar geschakeld werden met CALL-LINK-instructies.

2. Input

De input bestaat uit een parameterkaart en een set datakaarten.

- Parameterkaart FØRMAT (2I5)

Kol.	Var.	
1-5	NA	Orde van de vierkante matrix A. Restrictie: NA<25
6-10	NE	Aantal elementen van matrix A, verschillend van nul. Restrictie: NE<625.

---

(1) IBM Application Program, 1130 Scientific Subroutine Package, Programmers Manual, Fifth Edition, June 1970, p. 120.

- Datakaart FØRMAT (2I5, F13.0)

Kol.	Var.	
1-5	NR	Rijnummer
6-10	NK	Kolomnummer
11-13	E	Matrixelement

Voor elk element uit matrix A verschillend van nul een nieuwe datakaart.
Het totaal aantal inputkaarten bedraagt dus (NE+1)

3. Output

- Matrix 1: Inputmatrix
- Matrix 2: Matrix van Hessenberg
- Char. Polynom: Vector met de (NA+1) coëfficiënten van de karakteristieke polynoom, geordend van de kleinste tot de grootste macht.
- Het nummer, het reëele en imaginaire gedeelte van de wortels van de karakteristieke polynoom.

  Tussen de hoofding en eigenlijke output wordt de "error code" van de subroutine PØLRT afgedrukt, nl.

IER=o	geen fout
IER=1	NA<1
IER=2	NA>36
IER=3	Na 500 iteraties voor ieder van de 5 startwaarden werd geen wortel gevonden.
IER=5	De coëfficiënten van de hoogste graadsterm is nul.

4. Opmerkingen

4.1. Opmerkingen m.b.t. de input

De reden waarom de NE van nul verschillende elementen van matrix A element per element worden ingelezen is tweërlei. Het berekenen van de eigenwaarden van een reëele, niet symmetrische doch vierkante matrix A komt voornamelijk voor bij het onderzoek naar de stabiliteit van de "FINAL-FORM" als uiteindelijke oplossing van econometrische modellen (*). Bij een dergelijke toepassing bestaat de matrix A voornamelijk uit nullen. Een tweede reden heeft betrekking op de handigheid bij het onderzoek naar de sensiviteit van enkele gewijzigde matrixelementen.

4.2. Opmerkingen m.b.t. de resultaten

- De aandacht dient gevestigd te worden op de problemen die zich kunnen voordoen bij eigenwaarden met een multipliciteit die groter is dan één. De subroutine PØLRT blijkt namelijk niet steeds geschikt te zijn om op een behoorlijk nauwkeurige wijze de wortels van de karakteristieke polynoom te bepalen voor het geval deze wortels meerdere malen voorkomen. Een reden voor deze onnauwkeurigheid is niet enkel de gebruikte lengte van de mantisse ("ØNE WØRD INTEGERS") doch vooral de aard van het algorithme zelf.
- Deze opmerking kan het best geïllustreerd worden door volgende voorbeelden:

Voorbeeld I

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

met als reële eigenwaarden:

theoretisch	berekend
λ_1 = 1.0	λ_1 = 0.100021 E 01
λ_2 = 1.0	λ_2 = 0.999784 E 00
λ_3 = 0.0	λ_3 = 0.000000 E 00

(*) THEIL H. & BOOT J.: "The Final Form of Econometric Equation Systems", Review of the International Stаtistical Institute, 30, 1962, PP.136-152. FLEISSNER P. & HIETLER K., "Stabilität linearer ökonometrischer Modelle (1.Teil)", Computing, Vol.9, Fasc.4, 1972, pp.293-315. FLEISSNER P. & HIETLER K., "Stabilität linearer ökonometrischer Modelle (2.Teil)", Computing, Vol.10, Fasc.1-2, 1972, pp.33-62.

Voorbeeld II

$$A = \begin{bmatrix} 7.0 & 5.0 & -2.0 \\ 1.0 & 11.0 & -2.0 \\ -0.5 & 0.0 & 6.0 \end{bmatrix}$$

met als reële eigenwaarden:

theoretisch	berekend
$\lambda_1 = 12$	$\lambda_1 = 0.120000$ E 02
$\lambda_2 = 6$	$\lambda_2 = 0.600245$ E 01
$\lambda_3 = 6$	$\lambda_3 = 0.599754$ E 01

Voorbeeld III

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

met als reële eigenwaarden:

theoretisch	berekend
$\lambda_1 = 1$	$\lambda_1 = 0.110657$ E 01
$\lambda_2 = 1$	$\lambda_2 = 0.110134$ E 01
$\lambda_3 = 1$	$\lambda_3 = 0.108463$ E 01
$\lambda_4 = 1$	$\lambda_4 = 0.109473$ E 01
$\lambda_5 = 1$	$\lambda_5 = 0.109946$ E 01
$\lambda_6 = 1$	$\lambda_6 = 0.111303$ E 01

- Merk op dat voorbeeld III, in tegenstelling tot de voorbeelden I en II, de berekende λ's zelfs niet voldoen aan $\Sigma\lambda_i = tr(A)$.

- Zoals blijkt uit volgende kleine testvoorbeelden voldoet de geprogrammeerde methode blijkbaar wel wanneer geen eigenwaarden voorkomen met een multipliciteit groter dan één.

Voorbeeld IV

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$$

met als eigenwaarden

theoretisch	berekend Re	berekend Im
$\lambda_1 = \frac{1}{2} + \frac{1}{2}\sqrt{3}\ i$	0.500000 E 00	0.866025 E 00
$\lambda_2 = \frac{1}{2} - \frac{1}{2}\sqrt{3}\ i$	0.500000 E 00	-0.866025 E 00

<u>Voorbeeld V</u>

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$$

met als reële eigenwaarden

theoretisch	berekend
$\lambda_1 = 5$	λ_1 = 0.500000 E 01
$\lambda_2 = 0$	λ_2 = 0.000000 E 00

<u>Voorbeeld VI</u>

$$A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

met als eigenwaarden

theoretisch	berekend Re	berekend Im
$\lambda_1 = i$	0.375166 E-11	0.100000 E 01
$\lambda_2 = -i$	0.375166 E-11	-0.100000 E 01
$\lambda_3 = -1$	-0.100000 E 01	0.000000 E 00

PRØGRAMMA TLQØM

1. Omschrijving en methode

De voornaamste bedoeling van dit programma is het verstrekken van informatie m.b.t. het al dan niet aanwezig zijn van multicollineariteit in het lineaire regressieprobleem

$$y_j = \beta_o + \sum_{i=1}^{K} \beta_i \, x_{ij} + \varepsilon_j \qquad \begin{cases} j=1,2,\ldots,N \\ i=1,2,\ldots,K \end{cases}$$

Voor een beschrijving van de gebruikte methode en formules kan verwezen worden naar

BORGHERS E., "A Summary Statement of the General Linear Model" (With a Fortran Program), SESO-werknota, 7106, pp. 8-L6.

2. Input

De input van data geschiedt niet via kaarten maar wel via het schijvengeheugen. Hieruit volgt dat de voor dit programma benodigde data allereerst op de daartoe gereserveerde schijf geplaatst moet worden b.m.v. het programma "ØNDSK" (*). Wat het programma "TLQØM" betreft is de capaciteit van de data-bank "DBANK", die voorzien is op deze gereserveerde schijf, volledig beschikbaar (d.w.z. 2.500 reals).

De input van de parameters bestaat uit 2 parameterkaarten.

* Parameterkaart 1 (FØRMAT (16I5))

Kol.		
1 - 5	NO =	Nummer van het probleem
6 -10	K =	Aantal verklarende variabelen, exclusief de constante Restrictie: $2 \leq K \leq 10$
11-15	N =	Aantal waarnemingen voor elk van de variabelen Restrictie: $N \leq 200$

(*) BORGHERS E., op.cit., p. 56.

* Parameterkaart 2 (FØRMAT (16I5))

Kol.	
1 - 5	KEYS(1) = Sectornummer van de eerste verklarende variabele
⋮	⋮
1+5*(K-1) t/m 5+5*(K-1)	KEYS(K) = Sectornummer van de K^{de}, d.w.z. de laatste, verklarende variabele
1+5*K t/m 5+5*K	KEYS(K+1) = Sectornummer van de te verklaren variabele

Deze input wordt zo dikwijls herhaald als er te verwerken problemen zijn.

Sluitkaart : blanco

3. Output

- Hoofding: "Multicoll. information"
  - Nummer van het probleem (NO)
  - Aantal verklarende variabelen exclusief de constante (NX)
  - Aantal waarnemingen op al de variabelen (NT)
  - Sectornummers voor elk van de K verklarende variaben len (X(I)=KEYS(I)) en voor de te verklaren variebele (Y=KEYS(K+1))

- Matrix 90: Gemiddelden en standaardafwijkingen
  - Kolom 1: Op de I^{de} rij staat het gemiddelde van de I^{de} verklarende variabele. Het gemiddelde van de te verklaren variabele staat op de $(K+1)^{de}$ rij.
  - Kolom 2: Op de I^{de} rij staat de standaardafwijking van de I^{de} verklarende variabele. De standaard-afwijking van de te verklaren variabele staat op de $(K+1)^{de}$ rij.

- Matrix 91: Correlatie coëfficiënten tussen te verklaren en verklarende variabelen. Op de I^{de} rij staat de correlatiecoëfficiënt tussen de I^{de} verklarende variabele en de te verklaren variabele.

- Matrix 92: Correlatiematrix van de verklarende variabelen. Het $(I,J)^{de}$ element is de correlatiecoëfficiënt tussen de I^{de} en de J^{de} verklarende variabele.

- Matrix 93: De eerste K elementen geven, in dalende orde van grootte, de eigenwaarde weer van matrix 92.

  Het $(K+1)^{de}$ element, d.i. het laatste element, is de vierkantswortel van de gesommeerde gekwadrateerde eigenwaarden van de correlatiematrix van de verklarende variabelen.

- Matrix 94: Het eerste element geeft de waarde aan van de determinant van de correlatiematrix van de verklarende variabelen.
  Het tweede element geeft de waarde aan van de χ^2-toets op deze determinantwaarde.
  Het aantal vrijheidsgraden voor deze toets staat vermeldt als derde element.

- Matrix 95: Partiële correlatiecoëfficiënten van de verklarende variabelen.
  Het $(I,J)^{de}$ element is de partiële correlatiecoëfficiënt tussen de I^{de} en de J^{de} verklarende variabele.

- Het aantal vrijheidsgraden voor de t-waarden van matrix 96 (DF).

- Matrix 96: Het $(I,J)^{de}$ element van deze matrix geeft de t-waarde aan waarmee het $(I,J)^{de}$ element van matrix 95, d.i. de partiële correlatiecoëfficiënt tussen de I^{de} en de J^{de} verklarende variabele, getoetst kan worden op zijn significantie.

- Het aantal vrijheidsgraden voor de F-waarden van de tweede kolom van matrix 97.

- Matrix 97: - Kolom 1: Op de I^{de} rij staat de multiple correlatiecoëfficiënt tussen de I^{de} X-variabele als te verklaren variabele en de overige X-variabelen als verklarende variabelen.
  - Kolom 2: Op de I^{de} rij staat de F-toets waarmee de multiple correlatiecoëfficiënt van de I^{de} rij van de eerste kolom op zijn significantie getoetst kan worden.

4. Opmerkingen

Het programma "TLQØM" is een programma dat b.m.v. "CALL LINK" instructies ingebouwd werd in een sequentie van programma's. Het kan nochtans, op de hier beschreven wijze, toch AFZONDERLIJK gebruikt worden op voorwaarde evenwel dat "SWITCH 1" in "ØN" stand geplaatst wordt alvorens het programma te starten.

Benodigde subroutines:

- MINV,EIGEN,DCPY,XCPY,MSTR (*)
- WSNS (**)
- GETA,XØUT,TRPRT,TRRPR (***)

(*) IBM Application Program, 1130 Scientific Subroutine Package, Programmers Manual, Fifth Edition, June 1970.

(**) Rekencentrum - subroutines.

(***) SESO - subroutines.