

EEN ANALYSE VAN VERKIEZINGSUITSLAGEN MET BEHULP VAN CORRESPONDENTIEANALYSE

Peter van der Heijden *)

Samenvatting: Met behulp van correspondentieanalyse wordt een 3-weg matrix geanalyseerd. Het betreft hier verkiezingsuitslagen: een matrix van jaren (2) bij stemdistricten (58) bij partijen (10 en 20). Hierbij worden verschillende methoden gebruikt om de 3-weg matrix als 2-weg matrix te analyseren; d.m.v. het bekijken van elk jaar afzonderlijk, d.m.v. concateneren, en d.m.v. het analyseren van een zgn. 'bimarginale' matrix. De resultaten van de analyses worden gepresenteerd.

Trefwoorden: Correspondentieanalyse, categorische data, longitudinaal onderzoek, verkiezingen.

1. Inleiding en probleemstelling

Een verkiezingsuitslag kan weergegeven worden in een frequentietabel, waarbij de kolommen staan voor partijen, en de rijen voor bijvoorbeeld provincies, steden, districten, wijken.

De kolomtotalen leveren dan het aantal stemmen per partij op, de rijtotalen het aantal stemmen dat per district, stad etcetera is uitgebracht.

Indien men een bepaalde rij wil vergelijken met andere rijen, wordt dit gemakkelijker als de frequentietabel is omgezet in een tabel met percentages, waarbij de percentages van elke rij optellen tot 100. Hetzelfde kan natuurlijk, vice versa, gedaan worden voor de kolommen. Men vergelijkt dan partijen met elkaar.

*) Peter van der Heijden is als student-assistent werkzaam bij de vakgroep Datatheorie FSW/RUL, Hooigracht 15, Leiden. Hij wil hierbij de andere leden van de vakgroep danken voor het waardevolle commentaar op eerdere versies.

Indien men uitslagen van verschillende verkiezingen heeft, kan men op soortgelijke wijze veranderingen in kolommen en rijen over die verkiezingen bekijken. Hoe groter de afmetingen van de frequentietabel echter, en hoe meer verkiezingen men tegelijk wenst te bekijken, des te moeilijker is het een globaal overzicht te krijgen van de verkiezingsuitslagen.

In dit artikel willen we laten zien hoe correspondentieanalyse, uitgevoerd middels het computerprogramma ANACOR dat ontwikkeld is door de vakgroep Data-theorie (FSW) te Leiden, een oplossing biedt voor dit probleem.

Beschreven wordt één voorbeeld: de uitslagen van de Tweede Kamerverkiezingen van 1981 en de Provinciale Statenverkiezingen van 1982 in de stembedistricten van de gemeente Leiden. De nadruk komt hierbij te liggen op de gehanteerde methode. Allereerst wordt beschreven wat correspondentieanalyse inhoudt.

2. Correspondentieanalyse

Correspondentieanalyse is de Nederlandse naam voor wat in het Frans wordt aangeduid met Analyse des Correspondances. Het betreft een techniek die de relatie tussen rijen en kolommen van een matrix zichtbaar maakt. De rijen en kolommen worden afgebeeld als punten in een p-dimensionele ruimte met behulp van zgn. χ^2 -afstanden. Dit gebeurt als volgt:

als f_{ij} de frequentie is van de cel behorende bij rij i en kolom j , en de rijmarginalen worden weergegeven door $f_{i.}$, de kolommarginalen door $f_{.j}$, en $n=f_{..}$ is de som van de frequenties van de matrix, dan is de χ^2 voor een matrix

$$\chi^2 = \sum_i f_{i.} \sum_j \frac{(f_{ij}/f_{i.} - f_{.j}/n)^2}{f_{.j}/n}$$

Deze χ^2 is te splitsen voor de rijen: $\chi^2 = \sum_i f_{i.} x_i^2$. x_i^2 is de χ^2 -afstand tussen rij i en de kolommarginalen. Dit brengt ons tot de χ^2 -afstand δ_{ik} tussen elk paar van de rijen i en k in de matrix:

$$\delta_{ik}^2 = \sum_j \frac{(f_{ij}/f_{i.} - f_{kj}/f_{k.})^2}{f_{.j}/n}$$

Hierbij geldt dat $\delta_{ik} \geq 0$, $\delta_{ii} = 0$ en $\delta_{ik} + \delta_{kl} \geq \delta_{il}$.

Als rijen i en k proportioneel zijn, dan is $\delta_{ik} = 0$.

De coördinaten van de rijpunten krijgen we door de singuliere waarden decompositie van de matrix van δ_{ik} 's te nemen. Vervolgens plaatsen we de kolompunten in het gewogen gemiddelde van de rijpunten. Dit principe staat

bekend onder de naam 'reciprocal averaging' (Richardson (1933), genoemd in Horst (1935)) en 'le principe baricentrique' (Benzécri (1973)), en wordt ook gebruikt bij principale componenten analyse.

In het hierbovenstaande mogen rijen en kolommen verwisseld worden. Dat de rijpunten dan in het gewogen gemiddelde komen te liggen van de kolommen, is slechts een kwestie van normalisatie. In de analyses is gekozen voor een zgn. symmetrische aanpak: op elke dimensie liggen de rijen (kolommen) in het gewogen gemiddelde van de kolommen (rijen) gedeeld door de bijbehorende singuliere waarden. De symmetrie van de aanpak blijkt hieruit, dat, indien de matrix met frequenties gekanteld wordt, dezelfde oplossing wordt gevonden.

Het bovenstaande leidt tot het volgende: een rijpunt wordt dichter bij een kolompunt gelegd naarmate de frequentie in cel f_{ij} groter is dan op grond van de marginalen van die cel verwachte waarde $e_{ij} = f_{i.} \cdot f_{.j} / n$. Is f_{ij} kleiner dan e_{ij} , dan worden de punten van rij i en kolom j uiteen gelegd. Lijken patronen van twee rijen veel op elkaar, dan zullen deze punten in het plaatje dicht bij elkaar worden geplaatst. Hetzelfde geldt voor de kolommen. Is er sprake van een unieke relatie tussen een rij en een kolom, dan worden deze niet alleen dicht bij elkaar geplaatst, maar ook meer naar de periferie.

Ligt een rijpunt in het centrum, dan is f_{ij} voor de gehele rij ongeveer gelijk aan e_{ij} . Hetzelfde geldt uiteraard voor een kolompunt.

Een standaardwerk op het gebied van correspondentieanalyse is geschreven door Benzécri e.a. (1973). Door Hill (1974) is dit samengevat. Een recentere samenvatting van correspondentieanalyse en gerelateerde technieken is te vinden in Gifi (1981). Eveneens van de hand van Gifi (1982) is een User's Guide bij het programma ANACOR, waarin de methode en enkele voorbeelden worden beschreven.

In KM werd correspondentieanalyse kort beschreven door A.Z. Israëls e.a. (1981).

3. Leidse stemdistricten 1981 (TK) en 1982 (PS)

Er zijn twee frequentietabellen, weergegeven in tabel 1 (uitslag 1981) en tabel 2 (uitslag 1982). Beide zijn verkrijgbaar bij het Bureau Statistiek van het stadhuis van de gemeente Leiden.

In de tabellen is het aantal kolommen verschillend: in 1981 (TK) deden achttien politieke partijen mee, in 1982 slechts acht partijen. In beide

Tabel 1: Tweede Kamer 1981. Verkiezingsuitslagen per stemdistrict.
Aantal op elke partij uitgebrachte stemmen.

STEM-DISTRIKT	PVDA	CDA	VVD	D-66	SGP	PPR	CPN	GPV	PSP	RVP	DS	RK PN	LN	SP	RPF	IKB	EVP	SOS	ON-	AFWE- GELDIG	ZIG- TAAL
1.	295	179	544	181	4	71	29	2	104	2	10	1	1	6	5	15	10	1	8	145	167
2.	282	193	320	149	7	66	44	5	89	1	1	3	3	5	8	7	1	1	8	157	133
3.	419	154	399	123	8	4	28	4	121	1	1	3	1	5	7	2	3	5	19	152	146
4.	280	167	362	156	4	73	41	5	129	1	4	1	1	13	7	2	1	1	18	152	130
5.	383	147	253	164	2	8	4	4	121	1	4	3	3	13	7	2	1	1	13	152	130
6.	380	169	245	172	2	86	48	1	114	1	1	3	3	13	7	2	1	1	18	155	155
7.	297	235	386	154	2	6	37	1	118	1	1	3	1	2	5	4	4	1	11	191	155
8.	395	222	395	158	13	27	37	1	153	1	1	3	1	2	5	4	4	1	11	270	153
9.	368	222	395	156	13	27	37	1	153	1	1	3	1	2	5	4	4	1	11	270	153
10.	338	268	338	137	12	27	37	1	153	1	1	3	1	2	5	4	4	1	11	270	153
11.	488	268	488	192	9	42	44	2	66	2	2	1	1	17	7	15	6	4	19	221	143
12.	493	265	488	133	15	23	38	1	66	2	2	1	1	17	7	15	6	4	19	221	143
13.	474	255	474	183	12	23	38	1	66	2	2	1	1	17	7	15	6	4	19	221	143
14.	241	172	355	161	1	50	16	1	77	1	1	1	1	10	16	16	1	1	10	270	153
15.	297	200	395	171	1	50	16	1	77	1	1	1	1	10	16	16	1	1	10	270	153
16.	332	141	332	141	1	50	16	1	77	1	1	1	1	10	16	16	1	1	10	270	153
17.	597	157	597	182	4	17	54	1	55	5	5	1	2	13	6	1	15	17	17	410	137
18.	615	125	615	94	1	16	80	2	55	5	5	1	2	13	6	1	15	17	17	410	137
19.	613	125	613	94	1	16	80	2	55	5	5	1	2	13	6	1	15	17	17	410	137
20.	602	147	602	83	4	15	60	1	55	5	5	1	2	13	6	1	15	17	17	410	137
21.	413	185	413	73	3	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
22.	693	189	693	49	3	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
23.	578	154	578	115	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
24.	415	154	415	115	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
25.	594	200	594	185	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
26.	594	200	594	185	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
27.	268	291	268	155	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
28.	349	273	349	166	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
29.	288	313	288	181	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
30.	288	313	288	181	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
31.	288	313	288	181	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
32.	288	313	288	181	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
33.	288	313	288	181	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
34.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
35.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
36.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
37.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
38.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
39.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
40.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
41.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
42.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
43.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
44.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
45.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
46.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
47.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
48.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
49.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
50.	382	265	382	151	1	20	61	4	16	2	2	1	1	10	16	12	2	1	10	236	164
TOTAAL	23436	12816	12242	8485	446	2114	1966	263	2574	80	427	100	53	481	463	25	317	16	499	10840	7643

Tabel 2: Provinciale Staten 1982. Aantal op elke partij uitgebrachte stemmen

STEN- DISTR	PVDA	CDA	VVD	D'66	SGP	PPR/CPN	SP	RPV/ GPV	ON- GELD	AFVE- ZIG	TO- TAAL
1.	193	119	512	143	7	203	4	16	6	560	1763
2.	151	151	288	137	10	167	12	11	8	371	1337
3.	295	131	185	73	10	77	16	1	24	583	1583
4.	204	123	321	135	12	156	17	17	13	580	1452
5.	248	122	230	186	4	234	8	12	11	585	1552
6.	175	160	350	180	7	218	17	11	11	516	1560
7.	211	155	265	199	9	199	10	18	10	649	1629
8.	254	153	184	94	13	144	4	14	6	512	1388
9.	215	201	236	81	11	177	8	17	2	396	1244
10.	287	218	118	73	13	109	17	13	13	589	1436
11.	338	187	187	120	9	134	14	16	11	585	1479
12.	316	200	243	198	20	172	16	21	12	580	1493
13.	269	155	357	117	14	172	5	7	6	310	1278
14.	173	226	358	115	14	188	16	16	6	314	1349
15.	197	164	73	54	11	117	20	7	5	792	1373
16.	185	116	59	48	6	132	15	14	7	509	1091
17.	255	130	86	49	4	118	19	17	14	558	1240
18.	384	118	43	36	7	179	15	8	15	525	1230
19.	356	115	77	26	7	182	13	5	8	591	1322
20.	335	115	73	56	9	171	24	2	2	576	1322
21.	228	187	46	51	5	89	9	17	7	376	1037
22.	262	132	114	55	5	107	26	14	11	554	1324
23.	212	133	114	58	11	176	6	17	17	485	1180
24.	344	175	119	61	17	132	24	15	14	635	1538
25.	293	328	202	95	13	85	16	11	11	519	1573
26.	192	250	388	107	12	163	16	8	10	481	1340
27.	249	230	218	97	15	106	16	18	10	481	1340
28.	202	237	350	81	15	177	17	20	8	323	1185
29.	171	291	253	154	10	181	3	18	14	330	1440
30.	320	301	215	100	29	189	3	14	12	348	1391
31.	239	222	137	56	9	62	17	8	44	489	1209
32.	283	193	97	65	27	96	16	20	6	389	1192
33.	322	138	102	79	6	80	9	8	9	363	1106
34.	198	139	266	121	10	103	7	9	15	325	1243
35.	250	112	116	58	15	157	13	5	9	625	1885
36.	34	166	83	107	10	177	18	12	10	326	1337
37.	224	194	231	181	22	174	10	13	15	650	1447
38.	216	200	144	92	15	174	10	24	14	445	1372
39.	175	277	292	108	10	69	12	14	14	382	1334
40.	297	214	108	68	4	60	4	15	15	481	1266
41.	317	212	131	97	5	92	10	13	14	566	1457
42.	253	200	255	102	3	93	4	29	29	586	1542
43.	322	237	141	114	14	82	7	13	13	521	1455
44.	216	232	377	114	18	166	7	12	12	315	1288
45.	207	162	278	126	5	160	8	12	5	303	1315
46.	220	154	284	143	3	99	6	26	5	236	1189
47.	173	189	171	144	9	97	3	18	3	299	1559
48.	248	133	186	183	2	116	13	10	9	772	1759
49.	220	137	470	198	6	83	4	6	7	372	1554
50.	226	187	385	175	3	174	2	9	9	283	1198
51.	155	155	172	319	4	95	12	11	11	393	1415
52.	215	172	319	180	5	93	4	17	13	276	982
53.	225	156	158	56	4	73	4	17	13	276	982
TOTAAL	14204	10394	12026	5413	516	6555	583	729	541	27164	78125

Verklaring van de voor de politieke partijen gebruikte afkortingen:

1. PVDA : PARTIJ VAN DE ARBEID
2. CDA : CHRISTEN DEMOCRATISCH APPEL
3. VVD : VOLKSPARTIJ VOOR VRIJHEID EN DEMOCRATIE
4. D'66 : DEMOCRATEN 1966
5. SGP : STAATKUNDIG GEREFORMEERDE PARTIJ
6. PPR : POLITIEKE PARTIJ RADIKALEN
7. CPN : COMMUNISTISCHE PARTIJ VAN NEDERLAND
8. GPV : GEREFORMEERD POLITIEK VERBOND
9. PSP : PACIFISTISCH SOCIALISTISCHE PARTIJ
10. RVP : RECHTSE VOLKSPARTIJ
11. DS'70 : DEMOCRATISCH SOCIALISTEN 1970
12. RKPN : ROOMS KATHOLIEKE PARTIJ NEDERLAND
13. LN : LEEFBAAR NEDERLAND
14. SP : SOCIALISTISCHE PARTIJ
15. RPF : REFORMATISCHE POLITIEKE FEDERATIE
16. IKG : INTERNATIONALE KOMMUNISTENBOND
17. EVP : EVANGELISCHE VOLKSPARTIJ
18. SOS : SAMEN OVERLEVEN OF OTERVEN
19. KL.LINKS: PPR, PSP EN CPN

tabellen is er sprake van een kolom 'ongeldig', voor ongeldig gemaakte stembiljetten, en een kolom 'afwezig', voor de mensen die niet van hun stemrecht gebruik hebben gemaakt.

De verklaring van de voor de politieke partijen gebruikte afkortingen staat naast tabel 2.

Beide tabellen hebben achteenvijftig stemdistricten. Geografisch gezien horen de volgende stemdistricten bij elkaar: 1 t/m 9; 10 t/m 14; 15 en 16; 17 t/m 27; 28 t/m 38; 39 t/m 50; 51 t/m 58. Deze clusters van stemdistricten zullen in de resultaten van de analyses in grote lijnen zijn terug te vinden.

Achtereenvolgens wordt behandeld de analyse van 1981 en 1982 apart, de analyse van 1981 en 1982 tesamen, en de berekening van betrouwbaarheidsintervallen voor de kolompunten (partijen).

3.1 De analyse van de verkiezingen 1981 en 1982 afzonderlijk

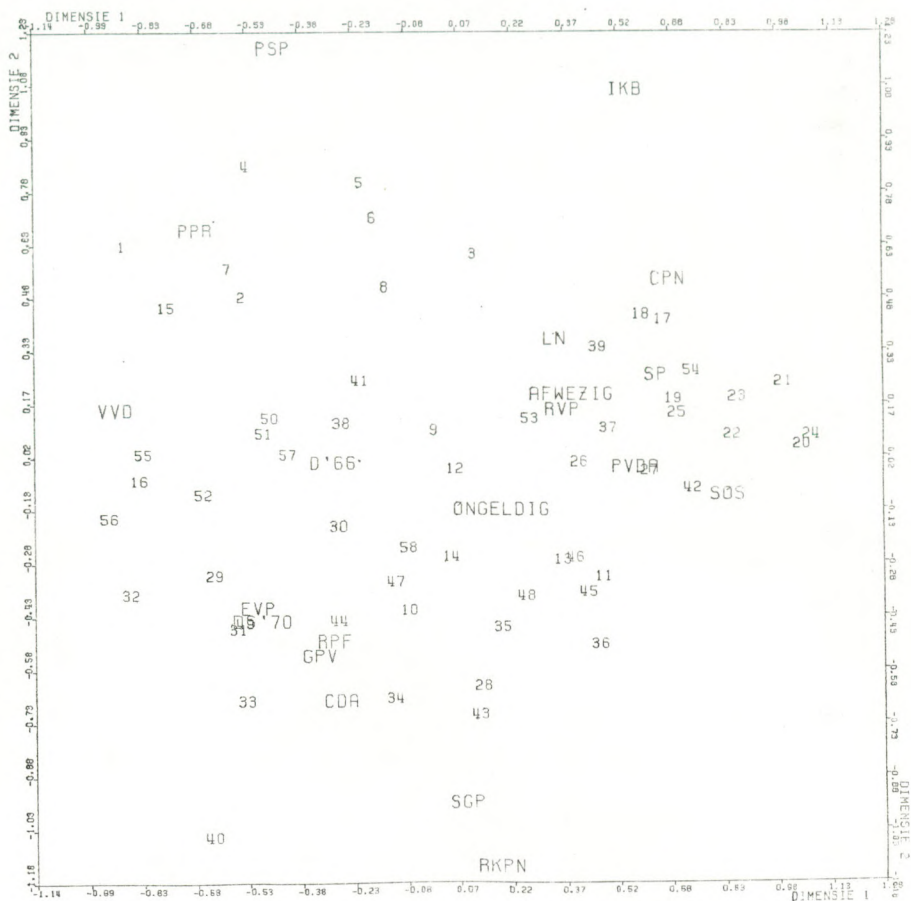
Indien we met het programma ANACOR correspondentieanalyse doen op tabel 1 (1981) krijgen wij het volgende plaatje (zie figuur 1). De stemdistricten hebben hier dezelfde nummers als in de tabel. Hetzelfde geldt voor de labels van de partijen.

De stemdistricten 1 t/m 9, liggend in de binnenstad, zijn dicht te vinden bij de PSP en de PPR. Dit ligt waarschijnlijk voor een belangrijk deel aan de hier wonende studenten. Districten 15 en 16, horend bij een ietwat goede buurt, liggen dicht bij de VVD. Districten 17 t/m 27, horend bij een verkrotte buurt, liggen voornamelijk rond de PVDA en de CPN. Districten 28 t/m 38, die in een oudere goeide buurt liggen, is te vinden bij de christelijke partijen. Stemdistricten 51 t/m 58, horend bij een nieuwbouwwijk met voornamelijk duurdere woningen, liggen voor het grootste deel in de buurt van D'66 en de VVD. Districten 39 t/m 50, ook liggend in een nieuwbouwwijk, zijn zo'n beetje bij alle grote partijen in de buurt te vinden. De niet-stemmers wonen vooral in districten waar veel op de PVDA wordt gestemd, en wonen relatief minder in wijken waar veel CDA en VVD wordt gestemd.

We kunnen ook individuele districten bekijken. Zo ligt district 1 in de buurt van de VVD en PPR, PSP. Als we naar tabel 1 kijken, zien we dat in dit district op deze partijen hoger dan het stedelijk gemiddelde is gescoord.

De verkiezingen van 1982 laten wat de plaats van de stemdistricten onder-

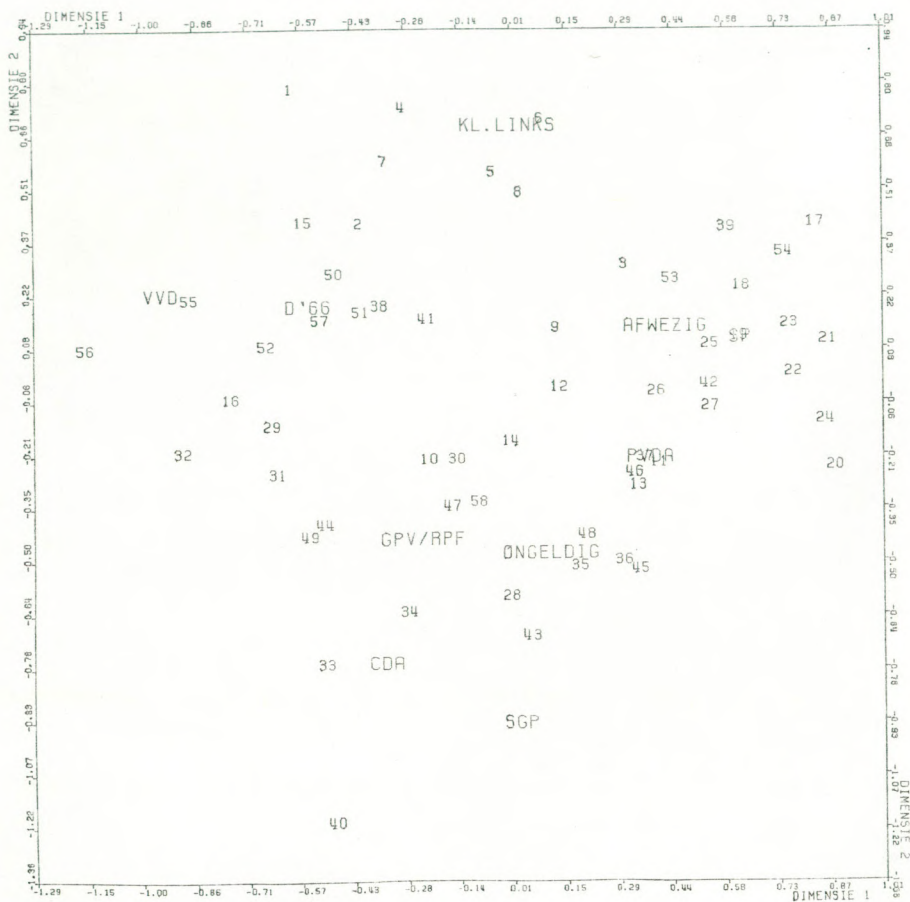
Figuur 1: Verkiezingen Tweede Kamer 1981



STEMDISTRIKTEN (GENUMMERD) EN PARTIJEN (AFGEKORT)

ling betreft een vrijwel identiek beeld zien (figuur 2). Dit geldt niet voor de partijen: het punt afwezig ligt dicht bij de stemdistricten uit de verkrotte buurt, de PVDA ligt iets verder van dezelfde districten af. VVD en D'66 liggen dicht bij elkaar. D'66 lijkt verloren te hebben in de verkrotte buurt, of gewonnen in de VVD-districten (is namelijk vanaf de verkrotte buurt naar de 'VVD-achtige' districten bewogen). Tot zover de analyses van de verkiezingen 1981 en 1982 afzonderlijk.

Figuur 2: Verkiezingen Provinciale Staten 1982



STEMDISTRIKTEN (GENUMMERD) EN PARTIJEN (AFGEKORT)

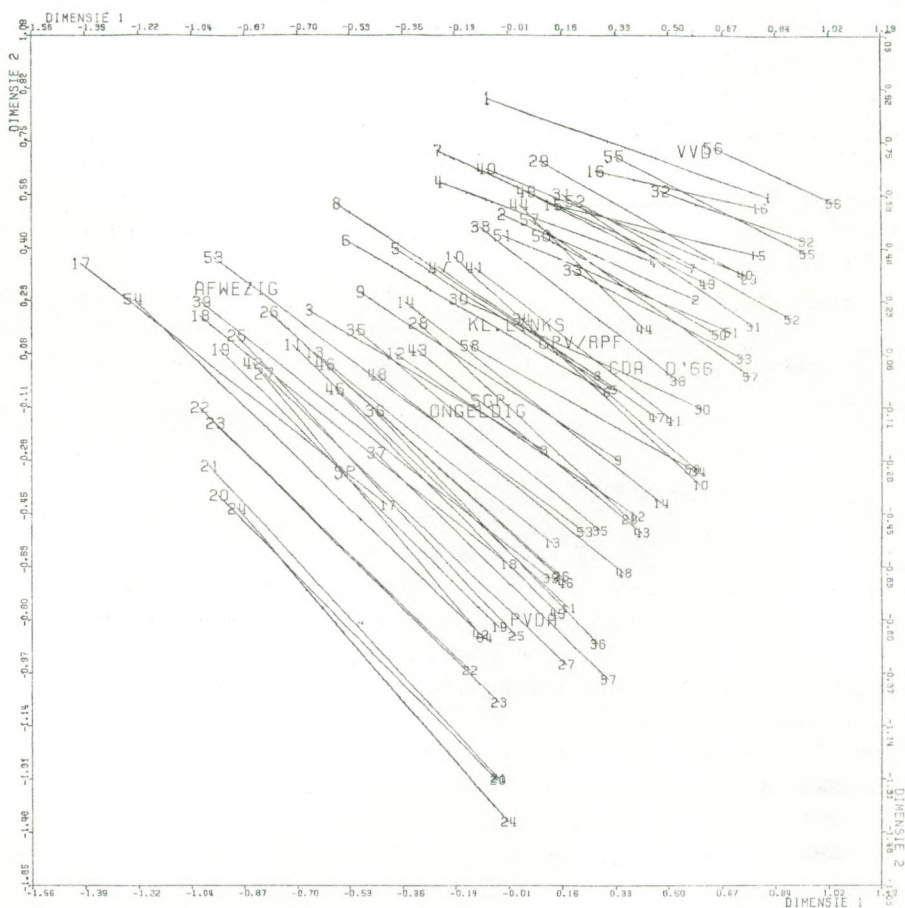
3.2 De analyse van de verkiezingen 1981 en 1982 tesamen

Hoe kunnen we bekijken of er verschuivingen zijn opgetreden in de uitslagen van 1982 ten opzichte van die van 1981? Hier hebben we drie manieren voor. Twee behelzen het aan elkaar plakken van de tabellen. De derde manier is de bimarginale tabel te analyseren. Deze drie manieren zullen we hieronder achtereenvolgens beschrijven.

3.2.1 Tabellen onder elkaar

Het is mogelijk de tabel van 1982 onder die van 1981 te zetten. Vóór we dit doen, moeten we zorgen dat deze tabellen een gelijk aantal kolommen, met identieke betekenis per kolom, hebben. Hiertoe nemen we voor 1981 de partijen PSP, CPN en PPR tesamen, evenals RPF en de GPV. De kolommen RVP, DS'70, RKPN, LN, IKB, EVP en SOS laten we weg. Dit is bij elkaar 2,2% van de stemmen van 1981.

Figuur 3: Verkiezingen Tweede Kamer 1981 en Provinciale Staten 1982



STEMDISTRIKTEN (GENUMMERD; KLEIN 1981, GROOT 1982)
EN PARTIJEN (AFGEKORT)

De nieuwe tabel bestaat nu uit 116 districten (ieder district twee maal) en tien partijen. We kunnen nu bekijken hoe voor elk district het punt van 1981 zich verhoudt tot dat van 1982. Dit is te zien in figuur 3. De districten zijn met lijnen verbonden.

Alle districten maken een sprong in dezelfde richting, namelijk vanaf de PVDA naar AFWEZIG. Dit komt doordat in alle districten de PVDA in 1981 veel minder stemmen heeft gekregen dan in 1982, en AFWEZIG veel meer stemmen. Indien er een lijn wordt getrokken van het punt PVDA naar het punt afwezig, zien we dat de lijnen die de bij elkaar horende stemdistricten verbinden, hieraan parallel lopen. De oorzaak hiervan is dat wijken zowel van de PVDA af moeten bewegen als naar AFWEZIG toe. Dit kan alleen door de lijn tussen 1981 en 1982 van een stemdistrict parallel aan de lijn van de PVDA naar AFWEZIG te laten lopen.

Indien alle partijen op deze lijn geprojecteerd worden, zien we de partij- en ongeveer in de volgorde winst verlies hierop liggen:

AFWEZIG - (VVD-KLEIN LINKS) - (SGP-RPF/GPV-ONGELDIG) - CDA - D'66 - PVDA.

Bovendien zien we de PVDA, AFWEZIG en de VVD extreem liggen: zij halen hun stemmen vooral uit een specifieke stemdistrictengroep. Dit in tegenstelling tot het CDA en KLEIN LINKS.

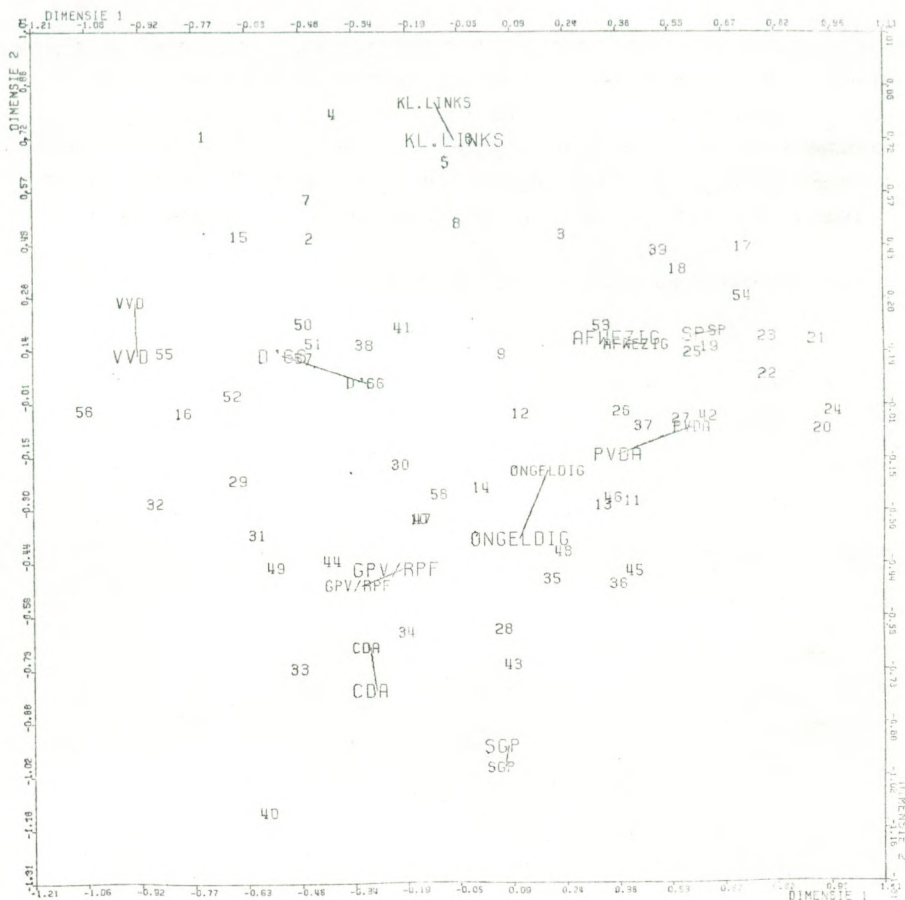
Vergelijken we figuur 3 met figuur 1 en 2, dan blijkt door de lengte van de lijnen dat de verandering in de stemdistricten waar veel PVDA, en veel niet gestemd wordt, veel groter is dan in de districten waar veel VVD wordt gestemd.

3.2.2 Tabellen naast elkaar

Plaatsen we de tabel van 1981 met 10 partijen naast die van 1982, en analyseren we deze tabel, dan krijgen we een plaatje waarin we partijpunten met elkaar kunnen verbinden (figuur 4). We zien dat de PVDA haar stemmen in 1982 minder specifiek uit de verkrotte buurt heeft gekregen, en omdat we weten dat de PVDA minder stemmen heeft gekregen, duidt dit op verlies van de PVDA in deze buurt.

Klein links is meer naar het centrum van het plaatje verschoven, en D'66 meer naar de VVD: daar D'66 heeft verloren duidt dit erop dat zij afhankelijker is geworden van de stemmen uit 'VVD-achtige' districten. Het punt AFWEZIG is vrijwel blijven liggen: dit betekent dat het frequentiepatroon van de thuisblijvers per district voor de twee verkiezingen redelijke overeenkomst vertoont, en dus dat in de meeste districten een rond de twee en een half maal zo groot aantal thuisblijvers woont in 1982 als

Figuur 4: Verkiezingen Tweede Kamer 1981 en Provinciale Staten 1982



STEMDISTRIKTEN (GENUMMERD)
EN PARTIJEN (AFGEKORT; KLEIN 1981, GROOT 1982)

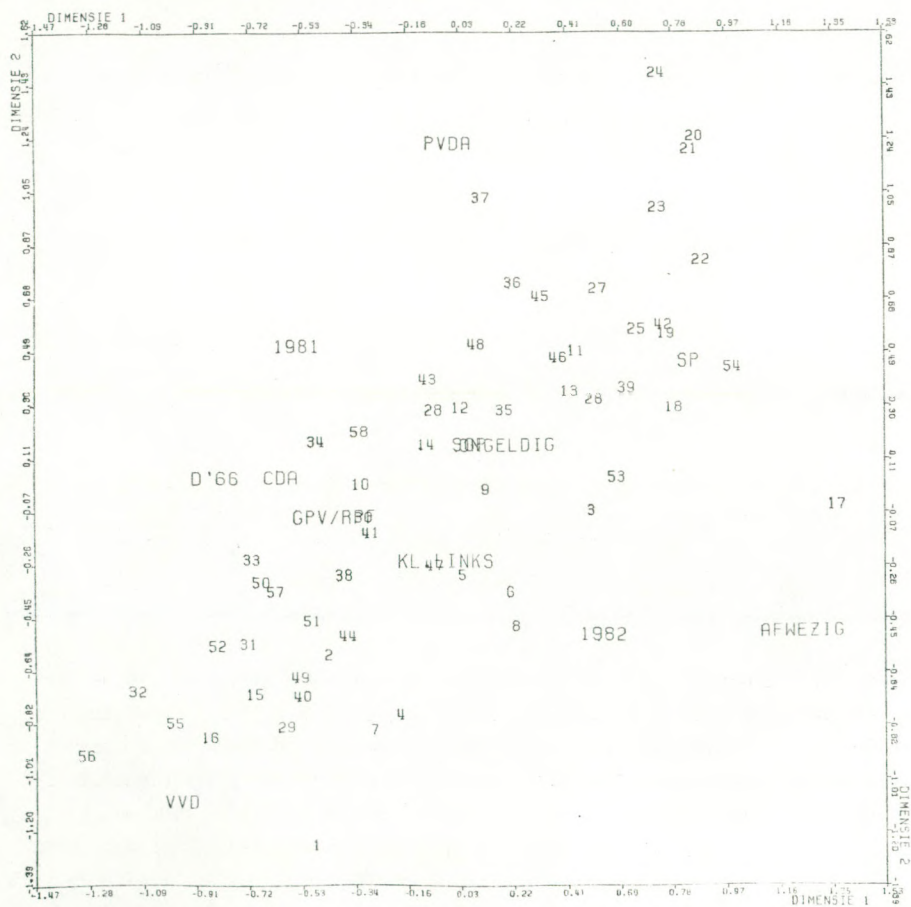
in 1981 (zie tabellen 1 en 2). Daar thuisblijven vooral gebeurt in de verkrotte buurt (daar ligt het punt immers bij), en de PVDA daar hoog scoorde, verklaart dit gedeeltelijk het verlies van de PVDA. Het CDA is wat haar stemmen betreft meer afhankelijk geworden van haar specifieke stemdistricten, aangezien het CDA-punt van 1982 meer naar de periferie ligt dan het CDA-punt van 1981. ONGELDIG heeft de grootste verschuiving ondergaan. Deze verschuiving zegt echter niet zoveel, daar dit slechts op een klein aantal stemmers betrekking heeft.

3.2.3 Bimarginale matrix

Een zogenaamde bimarginale matrix construeert men door als kolommen en rijen identieke categoriën te nemen, en wel de jaren (2), de partijen (10) en de stembedistricten (58).

De bimarginale matrix heeft dus 70 rijen en 70 kolommen, namelijk zowel voor de rijen als de kolommen de categoriën 1981, 1982, PVDA t/m AFWEZIG, district 1 t/m 58. De oorspronkelijke 3-weg matrix kun je op drie manieren optellen

Figuur 5: Verkiezingen 1981 en 1982: de bimarginale matrix



STEMDISTRIKTEN (GENUMMERD), PARTIJEN (AFGEKORT) EN JAREN (JAARTAL)

tot een 2-weg matrix: over de jaren, over de districten en over de partijen. Alle drie op deze wijze verkregen 2-weg matrices zijn terug te vinden in de bimarginale matrix. Er zitten ook lege cellen in de matrix: bijvoorbeeld de cel behorend bij rij 1981 en kolom 1982: een stem kan niet tegelijk zowel in 1981 als in 1982 uitgebracht zijn.

Het resultaat van de analyse van deze 70 bij 70 matrix is te zien in figuur 5 (NB. SGP en ONGELDIG overlappen elkaar).

Bij 1982 ligt het punt afwezig en KL. LINKS, deze hebben in 1982 meer stemmen gekregen dan in 1981. Het omgekeerde geldt voor de PVDA, D'66 en het CDA: deze hebben in 1982 minder stemmen gekregen. De PVDA en het punt AFWEZIG liggen verder uiteen dan op andere plaatjes.

In figuur 5 kunnen we zien dat in bijvoorbeeld district 37 de PVDA relatief minder heeft verloren dan in andere districten die ook in de verkrotte buurt liggen, en méér mensen naar de stembus zijn gegaan. In districten 17, 18, 53 en 54 is relatief méér niet gestemd, en tegelijkertijd is het aantal PVDA-stemmen relatief meer gedaald dan in andere wijken.

3.3 Betrouwbaarheidsintervallen

Indien we aannemen dat de data zijn op te vatten als een willekeurige steekproef uit een onbekende populatie, dan volgen de celfrequenties een multinomiale verdeling. Onder aanname van dit model is het mogelijk de momenten (gemiddelde, variantie) van de parameters te berekenen.

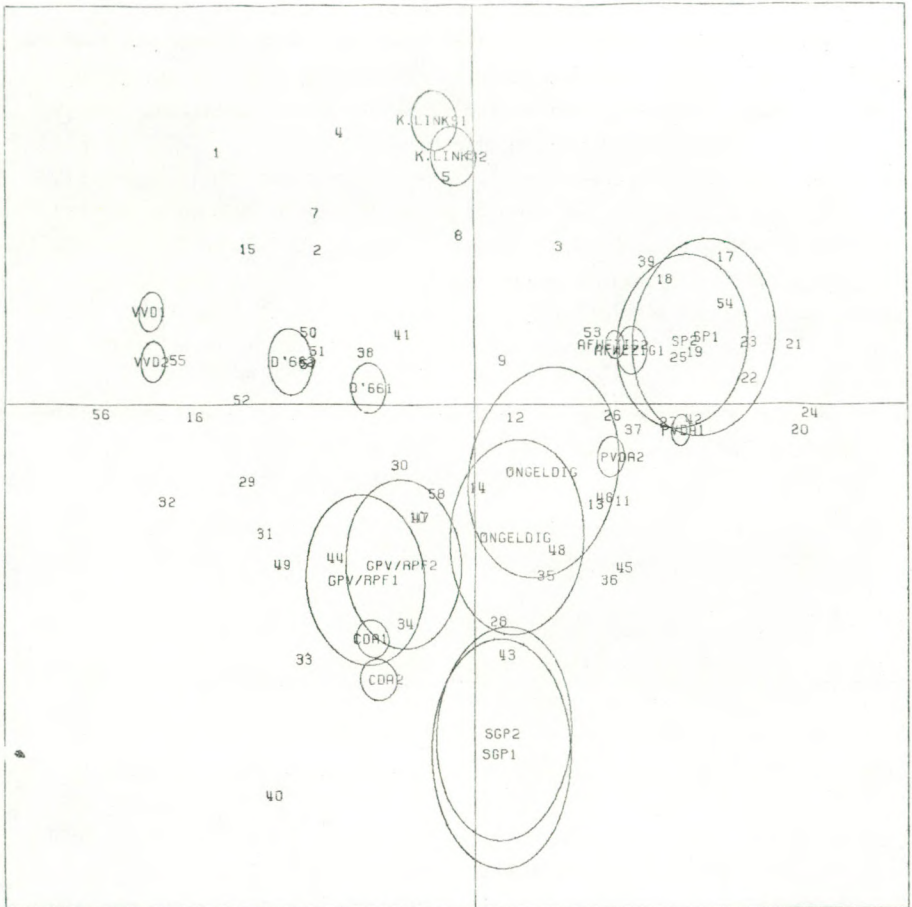
Het computerprogramma ANACOR kan de varianties van de singuliere waarden en de rij- en kolomscores berekenen, en de correlaties tussen scores van dezelfde parameter in verschillende dimensies. Zo bestaat de mogelijkheid de plaatjes van rij- en kolompunten van betrouwbaarheidsintervallen te voorzien: er kunnen ellipsen getekend worden die het gebied aangeven waarbinnen elk populatiepunt met 95% zekerheid ligt. (Het programma levert slechts de scores met behulp waarvan de ellipsen zijn te construeren, niet de ellipsen zelf).

Wat betreft de aanname van een willekeurige steekproef uit een onbekende populatie, moeten we ons iets voorstellen als dat de districten staan voor een grotere groep vergelijkbare districten uit bijvoorbeeld Nederland.

We hebben de betrouwbaarheidsgebieden van de partijpunten van figuur 4 als voorbeeld genomen (zie figuur 6).

Wat men kan zien is dat grotere partijen kleinere ellipsen hebben dan de kleinere partijen, en dus stabielere op hun plaats liggen. Ook van de ver-

Figuur 6: Figuur 4 opnieuw, ellipsen rondom de partijen



schuivingen van deze partijen kunnen we dus zekerder zijn: de verschuivingen van de PVDA, het CDA, de VVD en D'66 zijn alle significant: de betrokken gebieden overlappen niet.

3.4 Naar welk plaatje te kijken?

Figuur 1 t/m 5 zijn te beschouwen als vijf verschillende weergaven van de 3-weg matrix. Alle vijf leggen ze echter andere accenten. Deze zullen we kort aanstippen.

Figuur 1 en 2 geven de verkiezingsuitslag van 1981 en 1982 afzonderlijk weer. Indien men slechts in deze afzonderlijke uitslagen geïnteresseerd is, zijn deze figuren voldoende.

Figuren 3, 4 en 5 geven verschuivingen weer.

Indien men interactie veronderstelt tussen wijken en jaren, kijkt men naar figuur 3. Men kan bijvoorbeeld zien dat sommige wijken meer verschoven zijn dan andere, of in een andere richting, en daar conclusies uit trekken.

Hetzelfde kan gezegd worden over de interactie tussen partijen en jaren.

Indien men deze veronderstelt, kijkt men naar figuur 4.

In figuur 5 is de afbeelding van de variabelen te vinden zonder dat a priori interactieeffecten zijn verondersteld. Indien men niet zozeer geïnteresseerd is in verschuivingen van wijken en partijen, kan men deze figuur bekijken.

Correspondentieanalyse is op te vatten als een exploratieve dataanalysetechniek. Het is aan te bevelen terug te vallen op de oorspronkelijke frequentietabellen in geval van twijfel bij de interpretatie van de plaatjes.

4. Verder onderzoek

De hier gehanteerde methode kan op veel terreinen voor de analyse van 2-weg of 3-weg matrixen worden toegepast. Zo is het bij verkiezingsonderzoek interessant om meer dan twee jaren in te voeren. Men kan ook iets anders als rijen nemen, bijvoorbeeld plaatsen, provincies. Ook zou men slechts één of twee partijen kunnen nemen, en vervolgens verschuivingen van wijken per verkiezing kunnen bestuderen.

Een theoretischer tak van verder onderzoek is de relatie tussen loglineaire analyse en correspondentieanalyse: bij loglineaire analyse veronderstelt men interactie, en vervolgens worden de verwachte frequenties vergeleken met de waargenomen frequenties, en gekeken hoe goed de fit van het model is; bij correspondentieanalyse veronderstelt men óók interactie (bij een matrix van meer dan twee dimensies) en onderzoekt men vervolgens de (on)afhankelijkheid in een geconstrueerde 2-weg matrix.

De exacte relatie tussen loglineaire analyse en correspondentieanalyse is nog niet duidelijk. Aan het uitzoeken hiervan wordt momenteel gewerkt.

Literatuur

Benzécri, J.P. & Collaborateurs,

1973, L'analyse des Données, II L'analyse des Correspondances. Dunod
Paris-Bruxelles-Montréal

Gifi, A.,

1981, Non-linear multivariate analysis, Department of Datatheory,
University of Leyden

Gifi, A.,

1982, ANACOR User's guide, Department of Datatheory, University of Leyden

Hill, M.O.

1974, Correspondence analysis: a neglected method. Applied Statistics 23,
pp. 340-354

Horst, P

1935, Measuring complex attitudes. J. Soc. Psychol. 6, pp. 369-374

Israëls, A.Z.; Bethelam, J.P.; van Driel, J.; Janssen, M.E.; Pannekoek, J.;
de Ree, S.J.M.; Sikkels, D.,

1981, Multivariate analysemethoden voor discrete variabelen.
Kwantitatieve methoden 2, pp. 128-133