P. Doreian, N.P. Hummon, <u>Modelling Social Processes</u>, Elsevier, New York, 1976, 172 pagina's, prijs ca. *f* 31,--.

Doreian en Hummon hebben evenals Oud een systeem-methodologie voor ogen voor Sociaal-Wetenschappelijk Onderzoek. Zij gaan er echter van uit dat modellen moeten aansluiten bij (te formaliseren) problemen en stellen zich niet op Oud's standpunt. Oud's standpunt is het, om te laten zien dat de (Sociaal-Wetenschappelijke) LISREL-benadering voor modelbouw ook te formuleren is als een (klassiek) systeemtheoretisch model. Doreian an Hummon (hierna aangehaald door middel van 'D en H') storen zich veel minder aan bestaande formuleringen, zij creëren (waar nodig) nieuwe modellen. Een voorbeeld van hun werkwijze volgt. Om een verbale formulering van de situatie met betrekking tot overwerk in een fabriek te formaliseren, gebruiken D en H observaties, die onder andere luiden: 'de situatie was zō, dat als er minder overgewerkt werd, de mannen ontslag namen, omdat zij te weinig zouden verdienen. Maar als er meer overgewerkt moest worden, dan was er onder de mannen meer absenteïsme'. Zorgvuldige bewerking door D en H van dergelijke formuleringen leidt tot een model $x' = Ax$, waarin de vector $x = x(t)$ als eerste component x_1 het percentage voorstelt van het aantal werkuren dat afvloeit op dag t, terwijl de tweede component x_2 de hoeveelheid overuren voorstelt die gemaakt wordt op dag t. Fitten van het model op data betekent dus schatten van de 2x2- matrix A. Er zijn enige moeilijkheden verbonden aan een rechtstreekse schatting van A, daarom ontwikkelen D en H een aparte schattingsmethode van A door deze eerst op handige wijze te parametrizeren. Deze parametrizering wordt klungelig uitgevoerd, dat doet echter geenszins afbreuk aan de greep die D en H op dit model hebben. Ik zal de parametrizering hier iets beter uitleggen. De heuristiek voor de parametrizering is de observatie dat in het tweedimensionale vlak (x_1,x_2) - de betekenis van x_1, respectievelijk x_2 is hiervoor uitgelegd - de data vrijwel op een ellips liggen.
D en H berekenen de parameters van de ellips en vervolgens daarmee de parameters in de matrix A van hun oorspronkelijke model. Na zes pagina's rekenwerk komen zij tot:

$$A = \begin{pmatrix} 1.078 & -0.179 \\ 12.61 & -1.078 \end{pmatrix}$$

Het procédé dat D en H volgen kan veel inzichtelijker en eenvoudiger, mits men genegen is enige elementaire lineaire algebra te hanteren. Als voorbeeld:

De matrix A heeft de karakteristieke vergelijking in λ

$$\det(A - \lambda I) = |A - \lambda I| = \lambda^2 - b\lambda + c = 0$$
met $b = a_{11} + a_{22}$ en $c = a_{11}a_{22} - a_{12}a_{21} = \det(A)$

De karakteristieke waarden, ook wel eigenwaarden genaamd van deze vergelijking zijn:

$$\lambda_{1,2} = (b \pm \sqrt{b^2 - 4c})/2$$

We weten echter op grond van een plot van de data in het fasevlak dat deze een ellips vormen en daarmee een vrijwel cyclisch (dus stabiel) proces voorstellen, zodat onmiddellijk volgt dat de eerste-orde coëfficiënt $b = 0$ en dat de eigenwaarden puur imaginair zijn,

$$\lambda_{1,2} = \pm\, i\sqrt{\det(A)}, \text{ met } \det(A) > 0$$

terwijl bijbehorende eigenvectoren zijn:

$$x^1 = (a_{12}, \lambda_1 - a_{11}) \text{ en } x^2 = (a_{12}, \lambda_2 - a_{11})$$

Transformeren we daarmee de matrix A, dan komt er een diagonaalmatrix met op de hoofddiagonaal de complexe eigenwaarden λ_1 en λ_2. De twee oplossingen zijn:

$$x_1(t) = D_1 \cos(t\sqrt{c} - D_2)$$

$$x_2(t) = -D_1 \sin(t\sqrt{c} - D_2)$$

De periodiciteit T van de data blijkt uit figuur 2.2 ongeveer 6 jaar te zijn. Omdat zowel x_1 als x_2 het 'golfgetal' $\sqrt{c} = \sqrt{\det(A)}$ hebben, vinden we $\det(A) = 2\pi/T \sim 1.10$, terwijl D en H vinden $\det(A) = 1.095$.

Met behulp van Residuentheorie [1], toegepast op het onderhavige lineaire stelsel zijn de algemene oplossingen voor de variabelen $x_1(t)$ en $x_2(t)$ gemakkelijk te vinden.

$$x_1(t) = (2\,\lambda_1)^{-1}\,(C_1(\lambda_1 - a_{22}) + C_2 a_{12})\,e^{t\lambda_1} + {}$$
$$+ (2\,\lambda_2)^{-1}\,(C_1(\lambda_2 - a_{22}) + C_2 a_{12})\,e^{t\lambda_2},$$

$$x_2(t) = (2\lambda_1)^{-1}\,(C_2(\lambda_1 - a_{11}) + C_1 a_{21})\,e^{t\lambda_1} + {}$$
$$+ (2\lambda_2)^{-1}\,(C_2(\lambda_2 - a_{11}) + C_1 a_{21})\,e^{t\lambda_2}.$$

Vergelijking van coëfficiënten in de diverse oplossingen geeft

$$D_1^2 = -a_{12}(C_1^2 a_{21} + 2C_1C_2a_{22} - C_2^2 a_{12})/c$$
$$D_2 = \arctan\,((C_2a_{12} - C_1a_{22})/C_1\sqrt{c})$$

Met behulp van de dataset (niet gepubliceerd in D en H's boek) van een zestal punten zijn de x_1 en x_2 nu te fitten en vervolgens daaruit de a_{ij} te vinden.

Bij dit voorbeeld wil ik het laten, hopend de indruk achter gelaten te hebben dat D en H een zeer goede greep op mathematische modelbouw hebben en vervolgens een minder goede greep op de daarvoor benodigde wiskunde.

Literatuur: [1] Smirnov, A course of higher mathematics, III/2, Pergamon Press, 1964.

H. Koppelaar.