CWI Syllabi

Managing Editors

J.W. de Bakker (CWI, Amsterdam)
M. Hazewinkel (CWI, Amsterdam)
J.K. Lenstra (CWI, Amsterdam)

Editorial Board

W. Albers (Maastricht)
P.C. Baayen (Amsterdam)
R.J. Boute (Nijmegen)
E.M. de Jager (Amsterdam)
M.A. Kaashoek (Amsterdam)
M.S. Keane (Delft)
J.P.C. Kleijnen (Tilburg)
H. Kwakernaak (Enschede)
J. van Leeuwen (Utrecht)
P.W.H. Lemmens (Utrecht)
M. van der Put (Groningen)
M. Rem (Eindhoven)
A.H.G. Rinnooy Kan (Rotterdam)
M.N. Spijker (Leiden)

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

The CWI is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

CWI Syllabus 10

Vacantiecursus 1986
Matrices



Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

ISBN 90 6196 304 4

Copyright © 1986, Mathematisch Centrum, Amsterdam
Printed in the Netherlands

INHOUD

Inleiding
F. van Schagen
Matrices, wat doe je er alzo mee?
F. van der Blij 1
Matrices en meetkunde
J. van de Craats 25
Matrices en grafen
W.H. Haemers 47
Lineaire modellen voor dynamische verschijnselen
J.M. Schumacher 61
Matrices en Markovprocessen
O.J. Boxma 81
Matrices en statistiek
A.J. Bosch 105
Lineair programmeren, in het bijzonder transportproblemen
M. Kindt 117

INLEIDING

Matrices zijn het centrale thema van deze syllabus. In de verschillende bijdragen wordt steeds voor een ander deel van de wiskunde een voorbeeld van het gebruik van matrices gegeven. Daarbij worden de benodigde eigenschappen van matrices niet a priori bekend verondersteld.

In de eerste bijdrage zal blijken wat matrices eigenlijk zijn. Men kan zeggen dat matrices rechthoekige schemas zijn van objecten die men kan optellen en vermenigvuldigen. Die objecten kunnen dus (reële, complexe, positieve) getallen zijn maar ook andere "dingen". Een essentiele eigenschap van matrices is dat men ze kan optellen en vermenigvuldigen mits de afmetingen "passen". In de eerste bijdrage zal verder ook iets van het gebruik van matrices in de algebra worden getoond.

De tweede bijdrage handelt over toepassing van matrices in de meetkunde. Al spoedig worden daarbij de platgetreden paden van de klassieke lineaire algebra verlaten en belanden we in de meetkunde over eindige lichamen. Matrices waarbij de elementen uit een eindig lichaam komen, spelen hier een rol.

In de derde bijdrage staat de grafentheorie centraal. De matrices die dan optreden zijn bijvoorbeeld verbindingmatrices van grafen. Maar ook wordt hier gebruik gemaakt van zogenaamde blokmatrices. Dat laatste is niets anders dan een matrix waarvan de elementen op zich weer matrices zijn. Dit sluit goed aan bij het idee dat de eis die we aan de elementen van een matrix opleggen is, dat die elementen vermenigvuldigd en opgeteld kunnen worden, wat met matrices, mits van geschikte afmetingen, inderdaad het geval is.

De vierde bijdrage behandelt ondermeer een toepassing van matrices in de differentiaalvergelijkingen en de differentievergelijkingen. Opnieuw treedt een ander soort matrices op. Ditmaal zijn de objecten veeltermen. Dat leidt tot weer andere vragen en inzichten over matrices. Maar ook hier weer zijn de matrices niet het doel maar slechts het middel.

In de vijfde bijdrage staan de Markovprocessen en in het bijzonder Markovketens centraal. De matrices die daar bij optreden hebben de bijzondere eigenschap dat de elementen reële getallen zijn tussen 0 en 1 en dat de som van de elementen uit één rij gelijk is aan 1. Die matrices hebben eigenschappen waaruit voor Markovketens belangrijke conclusies getrokken kunnen worden.

In de zesde bijdrage wordt via de matrices een verband gelegd tussen de statistiek enerzijds en de lineaire algebra en meetkunde anderzijds. Er worden conclusies getrokken over verbanden tussen stochastische grootheden op basis overwegingen uit de lineaire algebra, meetkunde en matrixtheorie. Opnieuw spelen blokmatrices een rol.

In de laatste bijdrage wordt een toepassing van de matrixrekening in een klasse van lineaire programmeringsproblemen gegeven. Er wordt door middel van

matrixalgorithmen een bewerkelijk probleem gereduceerd tot iets wat met een aanvaardbare inspanning kan worden opgelost.

De zeven bijdragen betreffen ieder een ander deel van de wiskunde. Toch zal verschillende malen blijken dat er, misschien onverwachte, verbanden zijn. Helaas blijven ook belangrijke onderwerpen buiten de beschouwing. Men denke bijvoorbeeld aan de numerieke wiskunde en de vele matrixalgorithmen die daar optreden.

F. van Schagen

MATRICES, WAT DOE JE ER ALZO MEE?

F. van der Blij
Rijksuniversiteit Utrecht

1. Rijtjes 3

2. Matrices 4

3. Een voorbeeld 5

4. Matrices en vectoren 7

5. Lineaire afbeeldingen 10

6. Lineaire afbeeldingen van een ruimte in zichzelf 14

7. Groepen 18

MATRICES, WAT DOE JE ER ALZO MEE?

§1. Rijtjes

Volgens sommigen begint het allemaal met verzamelingen.
En dan zijn genummerde verzamelingen net weer iets anders. We beperken ons nu even tot eindige verzamelingen. We defineren een *rijtje* als een afbeelding van de verzameling van de eerste n natuurlijke getallen $\{1, 2, 3, \ldots, n\}$ in een of andere verzameling V. Als a de afbeelding is spreken we van het rijtje met n elementen $k \to a_k$; $a_k \in V$, $k \in \{1, 2, \ldots, n\}$.

Meestal spreken we gewoon nederlands en noemen $(a_1, a_2, a_3, \ldots, a_n)$ een rijtje met n elementen. Wat kunnen we met zo'n rijtje doen? Bijvoorbeeld de nummering veranderen, het rijtje omschikken. We bestuderen dan permutaties. We willen echter meer met rijtjes doen.

Laat in V een optelling gedefiniëerd zijn, met de gebruikelijke rekenregels, een nulelement en bij ieder element een tegengestelde.
Dan definiëren we voor twee rijtjes $(a_1, a_2, a_3, \ldots a_n)$ en $(b_1, b_2, b_3, \ldots b_n)$ met evenveel elementen de som a + b door
$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$.

Met deze definitie gelden de gewone rekenregels, het rijtje $(0, 0, 0, \ldots, 0)$ is het nul element en de tegengestelde is het rijtje van alle tegengestelden.

Het ligt voor de hand het produkt van een natuurlijk getal en een rijtje te definiëren door

$$\lambda \cdot (a_1, a_2, a_3, \ldots, a_n) = (\lambda a_1, \lambda a_2, \lambda a_3, \ldots, \lambda a_n).$$

Wanneer in V een (scalaire) vermenigvuldiging met rationale, reële of complexe getallen gedefiniëerd is, definiëren we vermenigvuldigingen van rijtjes met zulke getallen door bovenstaande formule.

Wanneer in V een vermenigvuldiging met de gewone rekenregels, eenheidselement en inverse gedefiniëerd is, is het niet algemeen mogelijk deze structuur naar de rijtjes over te brengen.
De voor de hand liggende poging

$(a_1, a_2, \ldots, a_n) \cdot (b_1, b_2, \ldots, b_n) = (a_1 b_1, a_2 b_2, \ldots, a_n b_n)$

vereist dat (1, 1, ... 1) het eenheidselement is. Maar dan hebben vele elementen ongelijk 0, zoals bijvoorbeeld (1, 0, ..., 1) geen inverse.
Daarom voeren we nu nog geen produkt van twee rijtjes in.

§ 2. Matrices

We gaan nu funkties van twee variabele bezien. De eerste variabele heeft als domein de deelverzameling {1, 2, 3, ... m} van de natuurlijke getallen, de tweede variabele de deelverzameling {1, 2, 3, ..., n}. Het bereik is een deelverzameling van een verzameling V. Het beeld van het paar (i, j) onder de afbeelding a geven we aan met a_{ij}.

We schrijven alle beelden in een tabel

$$(a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Zo'n tabel noemen we een *matrix* A, soms geven we met ${}^{m}A^{n}$ aan dat de matrix m rijen en n kolommen heeft.

Matrices met m = 1 noemen we *rijvectoren*, we schrijven deze vaak als $(u_1, u_2, \ldots, u_n)$. Matrices met n = 1 noemen we *kolomvectoren*), deze schrijven we vaak als

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \cdot \\ \cdot \\ v_m \end{pmatrix}$$

De som van twee matrices ${}^{m}A^{n}$ en ${}^{m}B^{n}$ met gelijke aantallen rijen en ge-

lijke aantallen kolommen definiëren we, in het geval dat in de beeldverzameling V een optelling gedefiniëerd is door

$$(a + b)_{ij} = a_{ij} + b_{ij}.$$

En skalaire vermenigvuldiging wordt op voor de hand liggende manier gedefiniëerd.

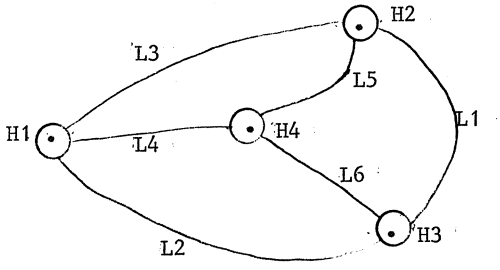
Net als bij rijtjes ligt de vraag naar de vermenigvuldiging van twee matrices voor de hand. De definitie

$$(a \cdot b)_{ij} = a_{ij} \cdot b_{ij}$$

heeft niet zoveel toepassingen. In § 4 zullen we een geheel andere en erg nuttige vermenigvuldiging voor matrices invoeren.

§ 3. Een voorbeeld

We bestuderen een stelsel tramlijnen met haltes. We beginnen met vier haltes H1, H2, H3 en H4 en met zes tramlijnen L1, L2, L3, L4, L5, L6. We geven een geschematiseerde plattegrond.



We geven een schema in matrixvorm.

	L1	L2	L3	L4	L5	L6
H1	•	*	*	*	•	•
H2	*	•	*	•	*	•
H3	*	*	•	•	•	*
H4	•	•	•	*	*	*

Een sterretje op de plaats (i, j) geeft aan dat de halte Hi ligt aan de tramlijn Lj. Waar geen sterretje staat zetten we een punt. Het schema is een matrix ${}^{4}T^{6}$ met elementen in V = {•, *}.

Met het zelfde gemak hadden we de situatie kunnen beschrijven met een andere matrix.

	H1	H2	H3	H4
L1	•	*	*	•
L2	*	•	*	•
L3	*	*	•	•
L4	*	•	•	*
L5	•	*	•	*
L6	•	•	*	*

We noemen deze matrix ${}^{6}S^{4}$ de *gespiegelde* of *getransponeerde* van ${}^{4}T^{6}$. We merken op dat de getransponeerde van de getransponeerde weer de oorspronkelijke matrix is. Verder is de getransponeerde van ${}^{m}A^{n}$ een matrix ${}^{n}B^{m}$ met $a_{ij} = b_{ji}$, (i = 1, 2, ... m), (j = 1, 2, ... n).

We geven de getransponeerde van een matrix C aan met C^{t}. We merken nog op dat de getransponeerde van een rijvector een kolomvector is en omgekeerd.

We keren terug naar ons voorbeeld van de tramlijnen. We willen het stelsel nu zo uitbreiden dat iedere twee haltes door precies één tramlijn verbonden worden en iedere twee tramlijnen precies één halte gemeen hebben.

Nu heeft L1 met L4 geen halte gemeen, we verlengen de lijnen en voegen een halte H5 toe. Analoog gaan we te werk om L2 en L5 een gemeenschappelijke halte H6 te geven en om L3 en L6 een halte H7 te geven. Nu moet nog L7 aangelegd worden, die H5, H6 en H7 verbindt.
De aangevulde matrix T wordt nu:

	L1	L2	L3	L4	L5	L6	L7
H1	•	*	*	*	•	•	•
H2	*	•	*	•	*	•	•
H3	*	*	•	•	•	*	•
H4	•	•	•	*	*	*	•
H5	*	•	•	*	•	•	*
H6	•	*	•	•	*	•	*
H7	•	•	*	•	•	*	*

Optellen van zulke matrices heeft natuurlijk geen zin, wat zou * + * moeten betekenen? We kunnen natuurlijk wel de nummerking van de haltes en/of van de tramlijnen veranderen. Dan worden de rijen en/of kolommen onderling verwisseld.

Een aardige vraag is of zo'n nummering mogelijk is dat de matrix *symmetrisch* wordt, dat wil zeggen gelijk aan zijn getransponeerde. In het algemeen heet een vierkante matrix A symmetrisch als voor alle i en j geldt $a_{ij} = a_{ji}$.

Opgave voor kenners. Vervangen we • door 0 en * door 1 dan krijgen we een matrix met gehele elementen. Wanneer we vooruit grijpen naar de klassieke matrix vermenigvuldiging uit de lineaire algebra (zie § 4) kunnen we het kwadraat van deze matrix berekenen. Heeft de uitkomst een betekenis voor ons tramnet?

§ 4. Matrices en vectoren

We verzinnen een verhaaltje over een diëtiste. Een maaltijd wordt samengesteld uit een voorgerecht, een hoofdgerecht en een dessert. In ieder komen voedingsstoffen E, K, M en V voor. Hoewel didaktisch onverantwoord, verzinnen we even de hoeveelheden voedingsstoffen per gram voor ieder gerecht. (U bent immers toch wiskundigen!) Het resultaat geven we in een tabel.

	Voorgerecht	Hoofdgerecht	Dessert
E	3	5	1
K	4	2	6
M	2	1	0
V	3	7	5

Wanneer de maaltijd bestaat uit a gram voorgerecht, b gram hoofdgerecht en c gram dessert, vinden we voor de totale hoeveelheid voedingsstoffen.

E: 3a + 5b + c
K: 4a + 2b + 6c
M: 2a + b
V: 3a + 7b + 5c

We introduceren nu de matrix $H = {}^{4}H^{3}$ door

$$H = \begin{pmatrix} 3 & 5 & 1 \\ 4 & 2 & 6 \\ 2 & 1 & 0 \\ 3 & 7 & 5 \end{pmatrix},$$

en de kolomvector m door

$$m = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

De hoeveelheid voedingsstoffen is ook een kolomvector

$$v = \begin{pmatrix} 3a + 5b + c \\ 4a + 2b + 6c \\ 2a + b \\ 3a + 7b + 5c \end{pmatrix}$$

We kunnen de berekening van v bij gegeven H en m door een zogenaamde *matrix-vermenigvuldiging* definiëren

$$\begin{pmatrix} 3 & 5 & 1 \\ 4 & 2 & 6 \\ 2 & 1 & 0 \\ 3 & 7 & 5 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3a + 5b + c \\ 4a + 2b + 6c \\ 2a + b \\ 3a + 7b + 5c \end{pmatrix},$$

kortweg

$$H \cdot m = v$$

Het algoritme is bekend en duidelijk.

We geven nu meteen de algemene definitie van het *matrix produkt* van twee matrices ${}^{m}A^{n}$ en ${}^{p}B^{q}$. We definiëren dit alleen als n = p. Het produkt is een matrix ${}^{m}C^{q}$ gedefiniëerd door

$$C_{\alpha\beta} = \sum_{\gamma=1}^{n} a_{\alpha\gamma}\, b_{\gamma\beta}$$

Dit produkt is niet commutatief. Met enige moeite is wel de associativiteit te bewijzen, in de volgende paragraaf ontdekken we een interpretatie van de matrixvermenigvuldiging waardoor de associativiteit triviaal wordt.

Met I geven we een symmetrische, vierkante matrix aan waarvoor $i_{\alpha\alpha} = 1$ en $i_{\alpha\beta} = 0$ voor $\alpha \neq \beta$. We merken op dat

$${}^{m}A^{n} \cdot {}^{n}I^{n} = {}^{m}A^{n} \text{ en } {}^{m}I^{m} \cdot {}^{m}A^{n} = {}^{m}A^{n}.$$

Met enig rekenen is te bewijzen dat

$$(A \cdot B)^{t} = B^{t} \cdot A^{t}.$$

In ons voorbeeld is duidelijk dat

$$v^{t} = m^{t} \cdot H^{t},$$

met rijvectoren v^{t} en m^{t}.

Als de voedingsstoffen per gewichtseenheid resp. e, k, m en v joules (energie) bevatten, zien we dat het voorgerecht 3e + 4k + 2m + 3v joules bevat, het hoofdgerecht 5e + 2k + m + 7v joules en het dessert e + 6k + 5v joules. We kunnen een vector

$$g = \begin{pmatrix} e \\ k \\ m \\ v \end{pmatrix}$$ invoeren en een vector $$z = \begin{pmatrix} 3e + 4k + 2m + 3v \\ 5e + 2k + m + 7v \\ e + 6k + 5v \end{pmatrix}$$

dan geldt

$$g^t \cdot H = z^t \text{ of } H^t \cdot g = z$$

Het produkt

$$g^t \cdot H \cdot m = z^t \cdot m = g^t \cdot v = m^t \cdot H^t \cdot g$$

heeft ook een betekenis. Welke?

§ 5. Lineaire afbeeldingen

In het vervolg kiezen we voor de verzameling V een lichaam K, meestal dat van de reële of complexe getallen.

Een verzameling L heet een *Lineaire ruimte* over K, als voor ieder tweetal elementen a en b uit L een som $a + b$ en voor ieder element $a \in L$ en iedere $\lambda \in K$ een element $\lambda \cdot a \in L$ gedefiniëerd is, waarbij de volgende regels gelden.

Laat $a, b, c \in L$ en $\lambda, \mu \in K$

$(a + b) + c = a + (b + c)$

$a + b = b + a.$

Er is een $0 \in L$ met $0 + a = a + 0 = a$.

Bij iedere $a \in L$ is een $(-a)$ met $a + (-a) = (-a) + a = 0$.

$\lambda \cdot (a + b) = \lambda \cdot a + \lambda \cdot b$

$(\lambda + \mu) \cdot a = \lambda \cdot a + \mu \cdot a$

$\lambda \cdot (\mu \cdot a) = (\lambda \mu) \cdot a$

$1 \cdot a = a$

We beperken ons tot eindig dimensionale lineaire ruimten, dat wil zeggen lineaire ruimten waarvoor een eindig aantal elementen $e_1, e_2, \ldots, e_n$ bestaan, zodat ieder element éénduidig als lineaire combinatie van deze elementen te schrijven is, dus $a = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$. Zo'n stelsel $e_1, e_2, \ldots, e_n$ noemen we een *basis* van de lineaire ruimten, het getal

n heet de *dimensie* van de ruimte.

We beschouwen twee lineaire ruimten L en M met dimensie n en m resp. Een afbeelding $A : L \to M$ heet lineair als

$$A(u + v) = A(u) + A(v), \qquad u, v \in L$$
$$A(\lambda u) = \lambda A(u). \qquad u \in L, \lambda \in K$$

Laat $e_1, e_2, \ldots, e_n$ een basis van L zijn en $f_1, f_2, \ldots, f_m$ een basis van m zijn. Stel $a \in L$ dan geldt $a = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$ zodat a eenduidig, bij deze basiskeuze, te beschrijven is met een rijvector $(x_1, x_2, \ldots, x_n)$.

De beelden van de basisvectoren zijn te schrijven als lineaire combinaties van de basis vectoren van M.

$$A(e_j) = \sum_{i=1}^{m} a_{ij} f_i \quad (j = 1, 2, \ldots, n)$$

Bij vast gekozen basis van L en M is de lineaire afbeelding éénduidig vastgelegd door de matrix

$$(a_{ij}) = {}_mA^n = A.$$

We schrijven soms, wanneer de basiskeuze vast ligt zowel voor de afbeelding als voor de matrix A.

Kiezen we in L een andere basis $g_1, g_2, \ldots, g_n$ met

$$g_\alpha = \sum_{j=1}^{n} u_{j\alpha} e_j,$$

dan geldt

$$\begin{aligned} A(g_\alpha) &= \sum_{j=1}^{n} u_{j\alpha} A(e_j) \\ &= \sum_{j=1}^{n} u_{j\alpha} \sum_{i=1}^{m} a_{ij} f_i \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} u_{j\alpha} \right) f_i \\ &= \sum_{i=1}^{m} b_{i\alpha} f_i \end{aligned}$$

Bij deze nieuwe basiskeuze in L hoort bij de lineaire afbeelding A een nieuwe matrix

$$(b_{i\alpha}) = {}_mB_n = B$$

We merken op dat $b_{i\alpha} = \sum_{j=1}^{n} a_{ij}\, u_{j\alpha}$, dus de matrix B is juist het matrixprodukt (zoals in de vorige § ingevoerd) van de matrices A en U:

$${}_mB_n = {}_mA_n \cdot {}_nU_n \cdot$$

We merken op dat de vierkante matrix U een bijzondere eigenschap heeft. Omdat zowel $(e_1, e_2, \ldots, e_n)$ als $(g_1, g_2, \ldots, g_n)$ een basis voor L is bestaan er getallen $z_{\alpha\beta}$ zodat

$$e_\beta = \sum_{j=1}^{n} z_{\alpha\beta}\, g_\alpha. \qquad (*)$$

En dus

$$e_\beta = \sum_{\alpha=1}^{n} z_{\alpha\beta} \sum_{j=1}^{n} u_{j\alpha}\, e_j$$

$$e_\beta = \sum_{j=1}^{n} \sum_{\alpha=1}^{n} u_{j\alpha}\, z_{\alpha\beta}\, e_j$$

Omdat de voorstelling van elementen met de basis éénduidig is geldt

$$\sum_{\alpha=1}^{n} u_{j\alpha}\, z_{\alpha\beta} = \begin{cases} 1 \text{ als } j = \beta \\ 0 \text{ als } j \neq \beta \end{cases}$$

Anders gezegd $U \cdot Z = I$

De matrix U heeft dus een inverse. Men bewijst eenvoudig dat ook

$$Z \cdot U = I$$

Een vierkante matrix, die een inverse heeft noemen we *regulier*
Kiezen we in M een andere basis $h_1, h_2, \ldots, h_m$ met

$$f_i = \sum_{\beta=1}^{m} w_{\beta i}\, h_\beta \qquad \text{(Let op het verschil met *)}$$

dan geldt

$$A(e_j) = \sum_{i=1}^{m} a_{ij} \sum_{\beta=1}^{m} W_{\beta i} h_\beta$$
$$= \sum_{\beta=1}^{m} \left(\sum_{i=1}^{m} W_{\beta i} a_{ij} \right) h_\beta$$
$$= \sum_{\beta=1}^{m} C_{\beta j} h_\beta$$

Zodat nu de matrix C bij de lineaire afbeelding hoort. We merken op dat C het matrixprodukt van de matrices W en A is, dus

$$^{m}C^{n} = {}^{m}W^{m} \cdot {}^{m}A^{n}$$

De matrix W heeft weer een inverse, is dus ook regulier.

Combinatie van de twee basistransformaties geeft voor de matrix van de lineaire afbeelding van L in M de vorm

$$D = W \cdot A \cdot U$$

Zonder bewijs vermelden we de eenvoudig te bewijzen

Stelling

In L en M zijn bases te kiezen, zodat bij de lineaire afbeelding een matrix $D = (d_{ij})$ hoort, die de gedaante heeft

$$d_{ij} = 1 \text{ voor } i = 1, 2, \ldots, k$$
$$d_{ij} = 0 \text{ voor alle andere paren } (i, j).$$

Het getal k heet de *rang* van de lineaire afbeelding. Als A een matrix is, die hoort bij een lineaire afbeelding met rang k, zeggen we ook dat de matrix A de rang k heeft. We merken op dat k gelijk is aan het maximale aantal lineair onafhankelijke kolommen en evenzo aan het maximale aantal lineair onafhankelijke rijen van de matrix A.

We formuleren de stelling nog iets anders. Bij een matrix $A = {}^{m}A^{n}$ van rang k zijn twee reguliere matrices $U = {}^{n}U^{n}$ en $W = {}^{m}W^{m}$ te vinden zodat

$$W \cdot A \cdot U = \begin{pmatrix} I & N_1 \\ N_2 & N_3 \end{pmatrix},$$

waarin $I = {}^{k}I^{k}$ en de alle elementen van de matrices N gelijk aan 0 zijn.

Bij een lineaire afbeelding van L in M is het *beeld* in M een lineaire deelruimte van M. De verzameling originelen van het nul element van M vormt een lineaire deelruimte in L, die we de *kern* van de afbeelding noemen.
Als afbeelding van een n-dimensionale ruimte in een n-dimensionale ruimte de rang k heeft geldt de volgende stelling:
De dimensie van de kern is n - k.
De dimensie van het beeld is k .

Laten L, M en N drie lineaire ruimten zijn met dimensies n, m en r resp. Laat A een lineaire afbeelding: L → M en B een lineaire afbeelding M → N zijn:
Dan is de samenstelling B • A een lineaire afbeelding L → N.

Laten in L, M en N bases gekozen zijn, zodat bij A en B de matrices ${}^{m}A^{n}$ en ${}^{r}B^{m}$ horen. Laat bij B • A de matrix ${}^{r}C^{n}$ horen. Dan geldt, zoals eenvoudig is na te rekenen, ${}^{r}C^{n} = {}^{r}B^{m} \cdot {}^{m}A^{n}$. Bij de samengestelde afbeelding B • A hoort het matrixprodukt B • A. Hieruit volgt nu direct de associativiteit van de matrix vermenigvuldiging.

We vermelden nog dat een lineaire afbeelding van L in M dan en slechts dan een bijectie is als n = m = k. In dit geval is er een lineaire inverse afbeelding. Hoort bij de gegeven afbeelding bij baseskeuze de matrix, dan hoort bij de inverse afbeelding de matrix A^{-1} met $A^{-1} \cdot A = A \cdot A^{-1} = I$.

§ 6. Lineaire afbeeldingen van een ruimte in zichzelf

Laat L een lineaire ruimte met dimensie n zijn. Laat A een lineaire afbeelding L → L zijn. Bij gegeven basis keuze in L hoort bij A een vierkante matrix.

Deze matrix heeft rang n dan en slechts dan als A een bijectie is. In dit geval heeft A een inverse. We noemden de matrix A dan regulier. Een vierkante matrix die niet regulier is heet *singulier*.

Bij een lineaire afbeelding van L in L geldt dat de som van de dimensies van kern en beeld gelijk is aan n. Bij een reguliere afbeelding heeft

de kern dimensie 0 en het beeld dimensie n.

In een lineaire ruimte L met dimensie n is aan ieder n-tal element een *volume* toe te kennen. Dit volume is een element van K, wanneer de n elementen lineair onafhankelijk zijn is dit volume ongelijk 0 en omgekeerd.

Bij een lineaire afbeelding A: L $\to$ L worden alle volumina met eenzelfde factor uit K vermenigvuldigd. Deze factor heet de determinant van de afbeelding. De determinant van een vierkante matrix is de determinant van de bijbehorende afbeelding. De determinant van een matrix is dan en slechts dan ongelijk 0 als de matrix regulier is.

Er is een bekend, maar nogal gecompliceerd, algorithme om van een gegeven vierkante matrix de determinant uit te rekenen.

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \ ,$$

$$\det \begin{pmatrix} a & b & c \\ p & q & r \\ x & y & z \end{pmatrix} = aqz - ary - bpz + brx + cpy - cqx \ .$$

Bij de berekening van de determinant komen alleen optellingen, aftrekkingen en vermenigvuldigingen voor, geen delingen. Hierdoor is de determinant ook te definiëren als de matrix elementen uit een willekeurige commutatieve ring gekozen zijn. We vermelden in dit verband nog dat als A een vierkante matrix is met elementen uit een ring R, er een vierkante matrix $\tilde{A}$ met elementen uit R bestaat zodat

$$\tilde{A} \cdot A = A \cdot \tilde{A} = \det (A) \cdot I \ .$$

Als $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ dan $\tilde{A} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$

Algemeen geldt $\det (A \cdot B) = \det (A) \cdot \det (B)$ en $\det (A^t) = \det (A)$. Verder $\det (I) = 1$. Dus

$$\det (\tilde{A}) = [\det (A)]^{n-1}$$

Bestaat er ook voor lineaire afbeeldingen van L in zichzelf een eenvoudige matrix vorm, te verkrijgen door geschikte basiskeuze. Omdat we nu

slechts één basis te kiezen hebben gaat het anders als in de vorige §.

Wanneer we de procedure voor basistransformatie uit de vorige § toepassen op afbeeldingen $L \to L$ zien we dat in dit geval $W = U^{-1}$. Hoort bij een gegeven lineaire afbeelding een matrix A dan hoort er na basistransformatie de matrix

$$U^{-1} \cdot A \cdot U$$

bij. (Merk op dat $\det (U^{-1} A U) = \det (A)$.)

De normaalvorm, na geschikte basiskeuze, voor de matrix is nu gecompliceerder. We noemen een vierkante matrix 'speciaal-boven-driehoekig' als hij de gedaante

$$\begin{pmatrix} \alpha & 1 & 0 & 0 & \cdot & \cdot & 0 & 0 \\ 0 & \alpha & 1 & 0 & \cdot & \cdot & 0 & 0 \\ 0 & 0 & \alpha & 1 & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & \alpha & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \alpha & 1 \\ 0 & 0 & 0 & 0 & \cdot & \cdot & 0 & \alpha \end{pmatrix}$$

heeft.

Een nogal moeilijk te bewijzen stelling is

Stelling: (Jordan)

Bij een gegeven lineaire afbeelding $A: L \to L$ is steeds een basis te kiezen zodat de matrix van A bij deze basis keuze bestaat uit een aantal langs de diagonaal gelegen speciale-boven-driehoeks matrices. De diagonaal elementen liggen in de algebraïsche afsluiting van K.

Een voorbeeld:

$$\begin{pmatrix} \alpha & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \gamma & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta \end{pmatrix}.$$

Samengevat is dus bij iedere vierkante matrix A een reguliere matrix U te vinden zodat U^{-1} A U een combinatie van speciale bovendriehoeksmatrices is (normaalvorm van Jordan).

We kunnen bovenstaande stelling in verband brengen met de vraag of bij een gegeven lineaire afbeelding A: L → L er een element $x \in L$ en een $\lambda \in K$ bestaat met

$$A x = \lambda x , \quad x \neq 0.$$

Zo'n element x heet een *eigenvector* van A bij *eigenwaarde* λ. Bij veel toepassingen van lineaire algebra blijken de eigenvectoren een belangrijke rol te spelen.
We schrijven nu

$$(A - \lambda I)\, x = 0.$$

De eigenvector ligt dus in de kern van $(A - \lambda I)$. Omdat de eigenvector ongelijk het nulelement moet zijn, moet dus $A - \lambda I$ singlier zijn, dus $\det (A - \lambda I) = 0$. Dit geeft een n-de graads vergelijking voor λ. De wortels van deze vergelijking zijn de eigenwaarden. Bij iedere eigenwaarde λ vormen de elementen x, waarvoor $Ax = \lambda x$ een lineaire deelruimte van L, bij de λ behorende *eigen ruimte*. Heeft een afbeelding precies n *verschillende* eigenwaarden, dan is de eigen ruimte bij iedere eigenwaarde één dimensionaal.

We merken op dat A en U^{-1} A U dezelfde eigenwaarden hebben, immers

$$\det (U^{-1} A U - \lambda I) = \det (U^{-1} (A - \lambda I) U) = \det (A - \lambda I).$$

In de Jordan normaalvorm zijn de eigenwaarden direct af te lezen. Uit de algorithme voor de determinant volgt dat de eigenwaarden van een matrix in de Jordan normaalvorm juist de diagonaal elementen van de matrix zijn.

Van een algemene 2 x 2 matrix zijn de eigenwaarden de wortels van een vierkantsvergelijking.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}$$

$$\det (A - \lambda I) = \lambda^2 - (a+d)\,\lambda + ad - bc.$$

In het algemene geval is het produkt van de eigenwaarden gelijk aan de determinant en de som van de eigenwaarden gelijk aan de som van de diagonaal elementen van de matrix.
Dit volgt uit eigenschappen van wortels van algebraïsche vergelijkingen en het determinant algorithme.
De som van de eigenwaarde heet het *spoor* van de matrix.
De eigenvectoren bij een gegeven eigenwaarde kan men vinden door lineaire vergelijkingen op te lossen; zij vormen de kern van $A - \lambda I$.

§ 7. Groepen

We beschouwen vierkante matrices met de vermenigvuldiging als compositie. Wanneer we ons beperken tot matrices met determinant ongelijk 0, hebben deze een inverse. Alle matrices met n rijen en kolommen met determinant ongelijk 0 vormen een groep, de algemene lineaire groep. De matrices met determinant 1 vormen een ondergroep, de speciale lineaire groep. De diagonaal matrices, met determinant ongelijk 0 vormen een andere ondergroep.

De orthogonale matrices (met reële elementen) en de unitaire matrices (met complexe elementen) vormen eveneens ondergroepen.
Met matrices zijn dus velerlei groepen te bouwen. Is iedere groep ook met een verzameling matrices te beschrijven? Zogenaamde matrix representaties vormen een belangrijk hulpmiddel in de groepen theorie.

We beginnen met een voorbeeld. We beschouwen de groep van alle permutaties van drie elementen A, B en C. Deze heeft 3! = 6 elementen. Is er een groep van 6 matrices die isomorf is met deze groep?

We proberen een meetkundige aanpak en beschouwen een gelijkzijdige driehoek A, B, C in het platte vlak met A = (1, 0), B = $(-\frac{1}{2}, \frac{1}{2}\sqrt{3})$,

$C = (-\frac{1}{2}, -\frac{1}{2}\sqrt{3})$.
Een permutatie van A, B en C komt neer op een lineaire transformatie in het vlak, die te beschrijven is met een 2 x 2 matrix.
De permutatie (A B C) → (A C B) hoort bij de spiegeling.

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

De permutatie (A B C) → (B C A) hoort bij een rotatie over 120°, dus bij

$$\begin{pmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix}$$

Zo vinden we een groep van zes matrices die isomorf is met de symmetrische groep met 6 elementen.

We kunnen de driehoek echter ook in $\mathbb{R}_3$ leggen, via A = (1, 0, 0), B = (0, 1, 0), C = (0, 0, 1).
De permutaties kunnen dan voorgesteld worden door 3 x 3 matrices, behorende bij lineaire transformaties in de $\mathbb{R}_3$.

Zo wordt (A B C) → (A C B) afgebeeld op

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

en (A B C) → (B C A) op

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

enzovoorts.

Duidelijk is dat iedere symmetrische groep op deze manier gerepresenteerd kan worden door matrices in een ruimte met even veel dimensies als het aantal te permuteren objecten. Kan het echter ook door een groep matrices met kleinere afmeting?

We merken op dat als een groep G isomorf is met een groep matrices

$\{I, A_1, A_2, \ldots, A_k\}$, de groep ook isomorf is met $\{I, U^{-1} A_1 U, U^{-1} A_2 U, \ldots, U^{-1} A_k U\}$ met een willekeurige inverteerbare matrix U. Geeft dit mogelijkheden om alle representerende matrices op een eenvoudige gedaante te brengen? Het komt natuurlijk neer op geschikt basiskeuze in de representerende ruimte.

Bij de inbedding van A B C als driehoek in de $\mathbb{R}_3$ is duidelijk dat het vlak $x + y + z = 1$ bij alle transformaties invariant blijft. Listig zal het dus zijn één coorditaal as loodrecht op dit vlak te kiezen. Nemen we hiervoor de eerste en kiezen we de tweede en derde as in het vlak dan krijgen alle representerende matrices de vorm

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix}$$

En de rechter onderhoek vormt een representatie door 2 x 2 matrices.

We kunnen ook het geval bestuderen dat de matrices een groep vormen die slechts homomorf met G is, de afbeelding van G naar de groep matrices is dan geen bijectie.
De symmetrische groep kunnen we homomorf afbeelden op de groep van twee 1 x 1 matrices, namelijk (1) en (-1), door aan de even permutaties 1 en aan de oneven permutaties -1 toe te voegen.

We bespreken als voorbeeld nog de symmetrische groep met 4! = 24 elementen. We noemen de elementen A, B, C, D.
Door A = (1, 0, 0, 0), B = (0, 1, 0, 0), C = (0, 0, 1, 0), D = (0, 0, 0, 1) vinden we direct een representatie door 24 matrices met 4 rijen en 4 kolommen. De ruimte $x_1 + x_2 + x_3 + x_4 = 1$ is invariant, enz. We kunnen ook een regelmatig viervlak in $\mathbb{R}_3$ leggen en de transformaties die dit viervlak in zichzelf vervoeren. Zo krijgen we een representatie door 3 x 3 matrices.

We kunnen het viervlak ook op een vlak V plaatsen en daarna in de richting van één van de ribben die niet in V ligt op V projecteren. De permutaties van de hoekpunten A, B, C en D (waarbij D niet in V ligt) worden nu afgebeeld op permutaties vanA, B, C, die we boven al door 2 x 2 matrices representeerden. Maar als de projectie evenwijdig met A D is zal de permutatie (A B C D) → (D B C A) op de identieke matrix worden afgebeeld. We hebben nu dus geen isomorfe representatie meer, maar wel een homomorfe. Iedere groep is als een groep van matrices te representeren.

Laten $a_1, a_2, \ldots, a_n$ de elementen van de groep zijn. De afbeelding $P_x : a_k \rightarrow x\, a_k$ is een permutatie van de elementen van de groep. Deze permutaties van N elementen vormen een ondergroep van alle permutaties van N elementen. De laatste was te representeren met N x N matrices. Dus is de eerste ook met N x N matrices te representeren. De opdracht is nu natuurlijk te proberen de representatie met behulp van invariante deelruimten op te splitsen in representaties in ruimten met lagere dimensie.

MATRICES EN MEETKUNDE

J. van de Craats
Koninklijke Militaire Academie, Breda

Affiene ruimte en vectorruimte 26

Affiene bases 28

Coördinatentransformaties 28

Affiene afbeeldingen 29

Affiene ruimten in het algemeen 30

Opnieuw vrije vectoren 32

Affiene meetkunde 34

Affien regelmatige veelhoeken 35

Regelmatige veelhoeken in een willekeurig affien vlak 36

Een paar eenvoudige gevallen 37

Het spoor 38

Het middelpunt van een regelmatige veelhoek 39

De gevallen $K=\mathbb{R}$ en $K=\mathbb{C}$ 39

Eindige affiene vlakken 40

Literatuur 42

Opgaven 42

MATRICES EN MEETKUNDE

Denk je aan matrices in de meetkunde, dan denk je in de eerste plaats aan transformaties die met matrices kunnen worden beschreven. Dus aan lineaire transformaties van lineaire ruimten, of, zoals ze ook wel worden genoemd, vectorruimten. Vectoren, uitgedrukt in coördinaten ten opzichte van een zekere basis, worden met behulp van een matrix in andere vectoren getransformeerd. Dus moet er allereerst een vectorruimte zijn. Hoe komen we daar aan in de meetkunde?

Laten we als uitgangspunt gewoon de ruimte om ons heen nemen. Hoe maak je daar een vectorruimte van? We gaan bij de fysica te rade. Hoe komen fysici aan vectoren? Grootheden als krachten, snelheden, versnellingen, momenten stellen zij vaak voor door pijlen, waarmee ze de grootte en de richting ervan aangeven. Vaak hebben gelijk gerichte, even lange pijlen dezelfde fysische betekenis; fysici spreken dan van "vrije" vectoren. Een vrije vector is dus eigenlijk een equivalentieklasse van pijlen.

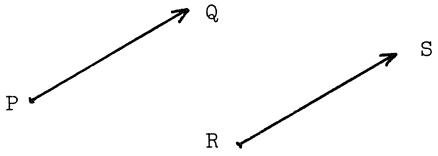
Elke pijl ligt natuurlijk volledig vast door z'n beginpunt P en z'n eindpunt Q. We nemen daarom als notatie voor zo'n pijl het geordende paar (P,Q). De vector (het woordje "vrij" laten we verder weg) die bij (P,Q) hoort, dus de equivalentieklasse van alle pijlen die dezelfde lengte en richting hebben als (P,Q), noteren we als $\overrightarrow{PQ}$. Dus $(P,Q) \backsim (R,S) \iff \overrightarrow{PQ} = \overrightarrow{RS}$.

We hebben nu vectoren. Maar om een echte vectorruimte te krijgen, moet er nog een structuur gedefinieerd worden: een vectoroptelling en een scalaire vermenigvuldiging. En daarvoor is essentiëel dat de zojuist gedefinieerde vectoren een paar meetkundige grondeigenschappen bezitten. De eerste gaat over de

"pijlenvoorraad" binnen een equivalentieklasse:

Bij elk drietal punten P, Q, R is er precies één punt S zo, dat $(P,Q) \backsim (R,S)$, m.a.w., zo, dat $\overrightarrow{PQ} = \overrightarrow{RS}$.

Dit is de precieze uitdrukking van het begrip "vrije vector". Slordig gezegd: elke vector kun je in elk willekeurig punt op precies één manier laten aangrijpen.



Dan de vectoroptelling en de scalaire vermenigvuldiging. U weet allemaal hoe dat met pijlen gaat. Maar vectoren zijn <u>equivalentieklassen</u> van pijlen. Je mag dus vectoroptelling wel met pijlen definiëren, bijvoorbeeld via de kop-aan-staart-constructie, maar dan moet je ook aantonen dat het resultaat niet van de representantenkeuze afhangt. En voor de scalaire vermenigvuldiging geldt iets dergelijks. Op welke meetkundige premissen is dit eigenlijk gebaseerd? Heeft U daar wel eens bij stilgestaan?

We zullen die kwestie hier laten rusten. Maar we merken wel op dat dit voorbeeld van vrije vectoren in de fysica model heeft gestaan voor het abstracte begrip vectorruimte. De axioma's voor een vectorruimte die U in alle boeken over Lineaire Algebra in de eerste paragraaf aantreft, en die ook in de openingsvoordracht van deze cursus vermeld staan, zijn niets anders dan een opsomming van de basiseigenschappen van dit prototype van een vectorruimte.

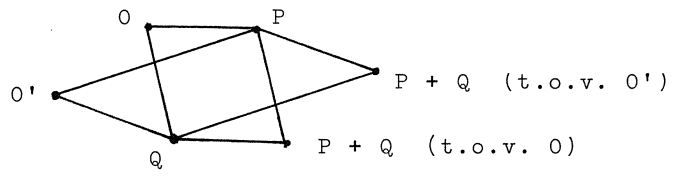
<u>Affiene ruimte en vectorruimte</u>

Heeft U overigens gemerkt dat we nog helemaal geen oorsprong gekozen hebben? De ruimte is nog steeds volkomen homogeen: alle punten zijn volstrekt gelijkwaardig. En zo hoort het natuurlijk ook in de meetkunde. Waar zitten de vectoren dan? Ik zeg het nogmaals: vectoren zijn equivalentieklassen van puntenparen. De nulvector, bijvoorbeeld, is de klasse die uit

alle paren (P,P) van twee gelijke punten bestaat. Wat we eigenlijk hier beschrijven, is een zg. affiene ruimte. Dat is een puntverzameling met daaraan vastgekoppeld een equivalentie-relatie op de puntenparen, en een vectorruimtestructuur op de daardoor vastgelegde verzameling van equivalentieklassen. Dat klinkt, zo geformuleerd, nogal abstract en geleerd, maar met de toelichting van hierboven begrijpt U misschien toch wat we bedoelen. De vectorruimte bevindt zich als het ware nog een beetje op de achtergrond. Vectoren zijn equivalentieklassen van geordende puntenparen. Vooral die "equivalentieklassen" maken de zaak weinig tastbaar.

Maar we kunnen het best wat concreter maken. Dat doen we ook in de praktijk bijna altijd als we een meetkundig probleem onder handen hebben waaraan gerekend moet worden. Dan kiezen we een geschikt punt O als oorsprong. En daarna nemen we uit elke equivalentieklasse precies dié pijl die in O begint. In iets vagere bewoordingen: we beroven de "vrije" vectoren van hun vrijheid doordat we ze allemaal in O laten aangrijpen. Dan ligt elke vector $\overrightarrow{OP}$ volledig vast door z'n eindpunt P, en via de identificatie van $\overrightarrow{OP}$ met P worden de punten in de ruimte dan zèlf vectoren die opgeteld kunnen worden en met scalairen kunnen worden vermenigvuldigd. De affiene ruimte is nu echt een vectorruimte geworden.

Maar opgelet, de ruimte als puntverzameling is z'n homogeniteit kwijtgeraakt. Er is één bijzonder punt geschapen, namelijk de oorsprong O. Die correspondeert met de nulvector, en onderscheidt zich structureel van alle andere punten. In dit verband moet U zich ook goed realiseren dat de vectorruimte-structuur die er nu op de ruimte als puntverzameling is gelegd, afhankelijk is van de keuze van O. Kijkt U maar:



Affiene bases

Meestal zijn we met deze verstoring van de homogeniteit nog niet klaar. We kiezen nòg drie bijzondere punten E_1, E_2, E_3, die niet samen met O in een vlak mogen liggen. Zo ontstaat een basis van de vectorruimte. Elk punt P kan dan van coördinaten ten opzichte van die basis worden voorzien, en daarmee is onze ruimte een $\mathbb{R}^3$ geworden: elk punt correspondeert met een geordend rijtje van drie reële getallen, en de vectoroptelling en de scalaire vermenigvuldiging corresponderen met de coördinaatsgewijze optelling en scalaire vermenigvuldiging in $\mathbb{R}^3$. Nu kan er in $\mathbb{R}^3$ worden gerekend. De affiene ruimte is dus helemaal niet a priori een $\mathbb{R}^3$ - ik hoor sommigen wel eens "de ruimte"en "$\mathbb{R}^3$" als synoniemen gebruiken - het wordt pas een $\mathbb{R}^3$ na de keuze van een zg. affiene basis, d.w.z. een viertal punten (O, E_1, E_2, E_3). Die punten mogen niet in één vlak liggen, maar verder kunnen ze geheel vrij gekozen worden.

Coördinatentransformaties

De vrijheid die we hebben om O, en vervolgens E_1, E_2 en E_3 te kiezen, moeten we zorgvuldig benutten, want iedereen weet dat rekenen in een onhandig gekozen coördinatenstelsel de eenvoudigste sommetjes in een hopeloze janboel van onoverzichtelijke rekenpartijen kan veranderen. Maar omdat die vrijheid van keuze van zo'n affiene basis er is, moeten we natuurlijk ook weten wat er gebeurt als we op een andere affiene basis (O', E_1', E_2', E_3') overstappen. En daar komt dan voor het eerst een matrix om de hoek kijken.

Als (x_1,x_2,x_3) de coördinaten zijn van een punt X met betrekking tot de basis (O, E_1, E_2, E_3), en als (x_1',x_2',x_3') de coördinaten zijn van datzelfde punt X ten opzichte van een andere affiene basis (O', E_1', E_2', E_3'), dan is er zoals U weet een betrekking van de vorm

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + b_1 \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + b_2 \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + b_3 \end{aligned}$$

waarbij de getallen a_{ij} en b_k alleen van de basisovergang afhangen. In matrixnotatie met vectoren in kolomvorm:

$$\underline{x}' = A\,\underline{x} + \underline{b} \ .$$

De matrix A is daarbij inverteerbaar, m.a.w. de determinant det(A) is niet nul.

Hoe vind je de matrix A en de vector $\underline{b}$? In het algemeen zullen de coördinatenvectoren $\underline{o}'$, $\underline{e}_1'$, $\underline{e}_2'$, $\underline{e}_3'$ van de nieuwe basispunten O', E_1', E_2', E_3' ten opzichte van de oude basis bekend zijn. En de coördinatenvectoren van deze punten t.o.v. de nieuwe basis zijn respectievelijk $(0,0,0)^t$, $(1,0,0)^t$, $(0,1,0)^t$, $(0,0,1)^t$. Bijgevolg geldt

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = A\,\underline{o}' + \underline{b}\ , \quad \begin{pmatrix}1\\0\\0\end{pmatrix} = A\,\underline{e}_1' + \underline{b}$$

$$\begin{pmatrix}0\\1\\0\end{pmatrix} = A\,\underline{e}_2' + \underline{b}\ , \quad \begin{pmatrix}0\\0\\1\end{pmatrix} = A\,\underline{e}_3' + \underline{b}$$

dus

$$A^{-1}\begin{pmatrix}1\\0\\0\end{pmatrix} = \underline{e}_1' - \underline{o}'\ , \quad A^{-1}\begin{pmatrix}0\\1\\0\end{pmatrix} = \underline{e}_2' - \underline{o}'\ , \quad A^{-1}\begin{pmatrix}0\\0\\1\end{pmatrix} = \underline{e}_3' - \underline{o}'.$$

De matrix A^{-1} bevat blijkbaar als kolommen de verschilvectoren $\underline{e}_i' - \underline{o}'$. Een <u>matrixinversie</u> geeft dus de gezochte matrix A. De resterende onbekende, de vector $\underline{b}$, is dan ook onmiddellijk gevonden: $\underline{b} = -A\,\underline{o}'$.

<u>Affiene afbeeldingen</u>

We hebben een matrix gebruikt als hulpmiddel om een coördinatentransformatie tot stand te brengen. YAGLOM spreekt van de "alias-functie" van de vergelijking $\underline{x}' = A\,\underline{x} + \underline{b}$: de punten X blijven onveranderd, maar ze krijgen een andere naam, d.w.z. andere coördinaten. We kunnen die vergelijking echter ook anders interpreteren, namelijk als een <u>afbeelding</u> van $\mathbb{R}^3$ op zichzelf, die daarmee ook een afbeelding α van de affiene ruimte op zichzelf met zich meebrengt. Dit is de zg. "alibi-interpretatie": we blijven met dezelfde affiene basis (O, E_1, E_2, E_3) werken, maar voegen aan elk punt X met coördinaten $\underline{x}$ nu een

punt $\alpha(X) = X'$ met coördinaten $\underline{x}' = A\,\underline{x} + \underline{b}$ toe. Dit is een affiene afbeelding, dat wil zeggen dat de affiene structuur behouden blijft: equivalente puntenparen gaan over in equivalente puntenparen, en op de onderliggende vectorruimte van equivalentieklassen werkt de afbeelding als een lineaire afbeelding. Gaat U maar na.

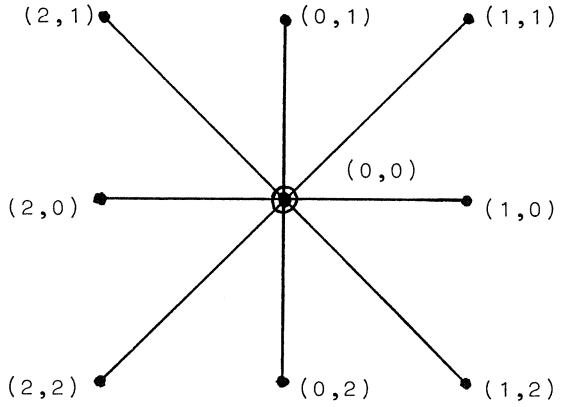
Er zijn natuurlijk allerlei bijzondere affiene afbeeldingen in de ruimte. We hebben het nog maar even gehad over lengte, en helemaal nog niet over hoeken. In affiene termen of in vectortermen moet je daarvoor nog wat extra structuur introduceren, bijvoorbeeld een inwendig product. Als je dat hebt, kun je ook orthonormale bases definiëren, en isometrieën en gelijkvormigheidstransformaties onderzoeken. Dat leidt dan al snel tot orthogonale en symmetrische matrices, waarmee we weer terug zijn bij het hoofdonderwerp van deze cursus. Maar ik wil vandaag deze min of meer platgetreden paden vermijden, en met U terugkeren naar het brede terrein van de affiene meetkunde pur sang. Wèl gaan we ons losmaken van dat ene concrete voorbeeld van de driedimensionale "fysische" reële affiene ruimte.

Affiene ruimten in het algemeen

Bij vectorruimten bent U al aan een abstracte definitie gewend. Elke niet-lege puntverzameling V in de wiskunde waarop een optelling en een scalaire vermenigvuldiging gedefinieerd zijn die voldoen aan de axioma's, heet een vectorruimte. De scalairen kunnen bovendien nog genomen worden uit een willekeurig lichaam K. We spreken dan van een vectorruimte over K. De wiskunde zit barstens vol met allerlei soorten vectorruimtem. Meestal over $\mathbb{R}$ of over $\mathbb{C}$, maar tegenwoordig zijn in de Discrete Wiskunde ook vectorruimten over eindige lichamen belangrijk geworden, vooral natuurlijk de vectorruimten over het lichaam $\mathbb{F}_2$ dat slechts uit de twee elementen 0 en 1 bestaat.

Ook affiene ruimten over een willekeurig lichaam K kunnen we axiomatisch definiëren. Het pad daarvoor is eigenlijk hierboven al geëffend. Omdat U vertrouwd bent met vectorruimten, is het 't gemakkelijkst om van een abstracte vectorruimte V over een lichaam K uit te gaan. Werkt U liever wat concreter, dan kunt

weer een voorbeeld in gedachten nemen. Maar nu eens een ander dan hierboven. Bijvoorbeeld als lichaam K het lichaam $\mathbb{F}_3$ bestaande uit de drie elementen 0, 1 en 2 met optelling en vermenigvuldiging modulo 3. Voor de vectorruimte V zouden we dan $\mathbb{F}_3^2$ kunnen nemen, dat wil zeggen de verzameling van alle geordende paren (x,y) met x en y uit $\mathbb{F}_3$. Hier is een schematisch plaatje van V:

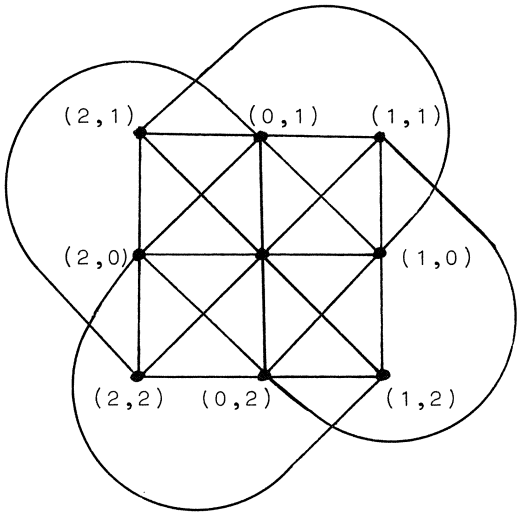


De lineaire deelruimten zijn ook aangegeven. Het zijn de nulvector $\{(0,0)\}$ op zichzelf, vier ééndimensionale deelruimten:
$\{(0,0), (1,0), (2,0)\}$, $\{(0,0),(0,1), (0,2)\}$,
$\{(0,0), (1,1), (2,2)\}$, $\{(0,0),(1,2), (2,1)\}$,
en natuurlijk de gehele ruimte V. Let wel, de getekende lijnen zijn slechts bedoeld om aan te geven welke drietallen punten samen een ééndimensionale deelruimte vormen.

Van meetkundig standpunt uit bezien is dit echter nogal onbevredigend. De nulvector neemt als punt duidelijk een zeer bijzondere positie in. We zouden liever een homogene puntverzameling hebben. En als "lijnen" zouden we er ook graag drietallen bijnemen als $\{(2,1), (0,1) (1,1)\}$ of $\{(1,1), (1,0), (1,2)\}$, kortom, alle deelverzamelingen van de vorm

$$\{(x,y) \in \mathbb{F}_3^2 \mid ax + by = c\}$$

voor zekere a, b, c $\in \mathbb{F}_3$ met a en b niet beide nul, en c ongelijk aan nul. Er zijn nog precies 8 van die "ééndimensionale lineaire



variëteiten". Hierboven hebben we ze erbij getekend. Nu ziet het er wèl homogeen uit: vier lijnen door elk punt, drie punten op elke lijn, en elk punt en elke lijn past precies op dezelfde manier in het geheel. Maar nog steeds is de vectorruimtestructuur aan één bepaald punt opgehangen. Wat is daar aan te doen?

Opnieuw vrije vectoren

De truc is om weer met pijlen te gaan werken, dat wil zeggen met geordende puntenparen. Of liever gezegd, met equivalentieklassen van pijlen. Ziet U al pijlen die onderling equivalent genoemd zouden mogen worden? Wat dacht U bijvoorbeeld van de pijlen ((0,1), (1,1)) en ((2,2), (0,2)), of van ((2,0), (0,1)) en ((0,2), (1,0)) ? Het gaat gewoon om verschilvectoren: we definiëren $(P,Q) \sim (R,S)$ als $Q - P = S - R$. Die verschilvectoren zijn inderdaad welbepaald, want V is immers een vectorruimte (bedenk ook dat we in dit voorbeeld modulo 3 moeten rekenen!). Er ontstaat zo een nette equivalentierelatie, en de equivalentieklassen kunnen we weer "vrije vectoren" noemen. Hoe tel je ze op? Natuurlijk als volgt:

$$\overrightarrow{((0,1),(1,1))} + \overrightarrow{((2,0),(0,1))} = \overrightarrow{((0,1),(1,1))} + \overrightarrow{((1,1),(2,2))} = \overrightarrow{((0,1),(2,2))}.$$

Dit is de gewone kop-aan-staart-constructie. U verifieert onmiddellijk dat deze definitie onafhankelijk is van de keuze van de representanten. De scalaire vermenigvuldiging is ook simpel. Bijvoorbeeld:

$2 \cdot \overrightarrow{((0,1),(1,1))} = \overrightarrow{((0,2),(2,2))}$.

De algemene definities liggen direct voor de hand. Op deze manier wordt de vectorruimtestructuur van V overgeplant naar de verzameling van vrije vectoren. Dat wordt de "onderliggende" vectorruimte, een vectorruimte op de achtergrond die isomorf is met V.

Laten we voordat we verder gaan met affiene bases en coördinaten, nog even recapituleren hoe je in het algemeen van een willekeurige vectorruimte V over een lichaam K een affiene ruimte maakt: als punten neem je gewoon de elementen van V, maar de vectorruimtestructuur plant je over naar de verzameling van vrije vectoren, dat wil zeggen de verzameling van alle equivalentieklassen van geordende puntenparen. De equivalentierelatie wordt daarbij als volgt gedefinieerd:

$(P,Q) \backsim (R,S)$ als $Q - P = S - R$.

Affiene deelruimten zijn natuurlijk de deelverzamelingen die op zichzelf beschouwd al een affiene ruimte zijn. Het is niet moeilijk om na te gaan dat dit precies de "lineaire variëteiten" in V zijn, de al dan niet "verschoven" lineaire deelruimten.

Een (k+1)-tal punten $(0, E_1, \ldots, E_k)$ noemen we affien onafhankelijk als de equivalentieklassen van de pijlen $(0,E_i)$ lineair onafhankelijk zijn in de onderliggende vectorruimte. We spreken van een affiene basis als ze een (lineaire) basis van die vectorruimte vormen. Onder de affiene coördinaten van een punt X t.o.v. zo'n basis verstaan we de coördinaten van de vector $\overrightarrow{(0,X)}$ met betrekking tot de basis $\overrightarrow{(0,E_1)}, \ldots, \overrightarrow{(0,E_k)}$ in de onderliggende vectorruimte. Die coördinaten vormen een rijtje $(x_1, \ldots, x_k)$ met $x_i \in K$. Door de keuze van een affiene basis wordt de affiene ruimte dus getransformeerd in de vectorruimte

$$K^k = \{ (x_1, \ldots, x_k) \mid x_i \in K, i = 1, \ldots, k \}$$

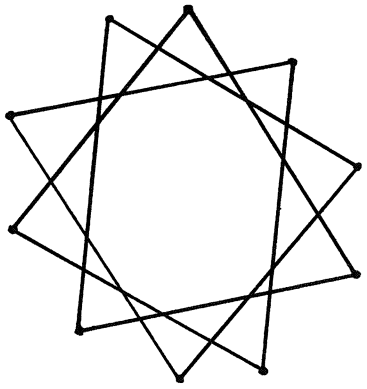
en daarin kan gerekend worden. Dat rekenen proberen we natuurlijk

zo simpel mogelijk te houden door de basis handig te kiezen.

Affiene meetkunde

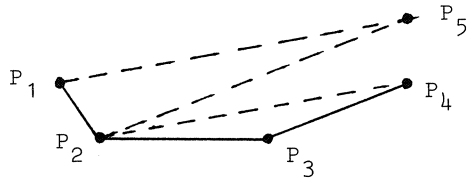
Het bovenstaande overziende concluderen we dat, los van enige meetkundige achtergrond, uit elke axiomatisch gedefinieerde abstracte vectorruimte over een willekeurig lichaam een affiene ruimte kan worden gemaakt. Bijvoorbeeld uit de vectorruimte $\mathbb{F}_3^2$, zoals we hebben gezien. Maar eveneens uit de vectorruimte $\mathbb{R}^3$, die ook helemaal axiomatisch gedefinieerd wordt als de verzameling van geordende drietallen reële getallen. En reële getallen kunnen geheel los van elke meetkundige vooronderstelling worden geconstrueerd. De reële affiene driedimensionale ruimte kan dus geheel zonder meetkundige axioma's worden gedefinieerd. Het is logisch gezien een wiskundig model dat goed bruikbaar blijkt te zijn bij het beschrijven van de fysische ruimte om ons heen.

Affiene meetkunde is de term die we gebruiken om de studie te beschrijven van de structuur van affiene ruimten in het algemeen. Ons woordgebruik daarbij is sterk beïnvloed door dat ene prototype waarmee alles begonnen is, de reële "fysische" affiene ruimte, met daarin punten, lijnen, vlakken, veelhoeken, veelvlakken, etc. Bij die studie speelt het rekenen met coördinaten natuurlijk een belangrijke rol. En vanzelf komen dan ook matrices (met getallen uit het lichaam K) tevoorschijn. We zullen dat illustreren aan de hand van een "case study": een onderwerp getiteld affien regelmatige veelhoeken.



Affien regelmatige veelhoeken

We keren weer even terug naar de ruimte om ons heen, en de bekende euclidische meetkunde er in. Als je een regelmatige n-hoek onderwerpt aan een scheve parallelprojectie, gaat er veel van de regelmaat verloren. De hoeken en de zijden zijn niet meer onderling gelijk. Maar er blijft ook iets behouden. Evenwijdige lijnenparen gaan bij zo'n projectie altijd weer in evenwijdige lijnenparen over. Teken je in een regelmatige n-hoek alle diagonalen, dan zie je nogal wat evenwijdige paren lijnstukken. Die moeten dus allemaal in de projectie ook evenwijdig blijven. Deze evenwijdigheidseigenschappen zijn zelfs tot op zekere hoogte karakteristiek voor parallelprojecties van regelmatige n-hoeken. Denkt U zich maar eens in dat we van zo'n projectie van (bijvoorbeeld) een regelmatige tienhoek vier opvolgende hoekpunten P_1, P_2, P_3 en P_4 op het bord hebben staan, en dat de rest per ongeluk is uitgeveegd. Dan kunnen we P_5 terugvinden door een lijn te tekenen door P_2 evenwijdig aan P_3P_4, en die te snijden met de lijn door P_1 evenwijdig aan P_2P_4. Het snijpunt moet dan P_5 zijn.



En hebben we P_5, dan is op dezelfde manier uit het viertal P_2, P_3, P_4, P_5 het zesde punt P_6 te vinden, enzovoorts. Vier opvolgende hoekpunten en een eenvoudig constructievoorschrift leggen de geprojecteerde regelmatige tienhoek dus volledig vast. En dat geldt natuurlijk niet alleen voor tienhoeken, maar in het algemeen voor alle n-hoeken.

Vier beginpunten. Kunnen die zo maar willekeurig gekozen worden? Natuurlijk niet. Zeker zal ook moeten gelden dat P_1P_4 en P_2P_3 evenwijdig zijn. Dus $P_1P_2P_3P_4$ moet een trapezium zijn. En verder? Je kunt natuurlijk met een willekeurig trapezium beginnen, en kijken wat er gebeurt. Misschien kom je na een aantal stappen weer op het uitgangspunt terug. Maar misschien ook niet. Het wordt tijd om de zaak eens analytisch aan te pakken, anders komen we niet verder.

Regelmatige veelhoeken in een willekeurig affien vlak

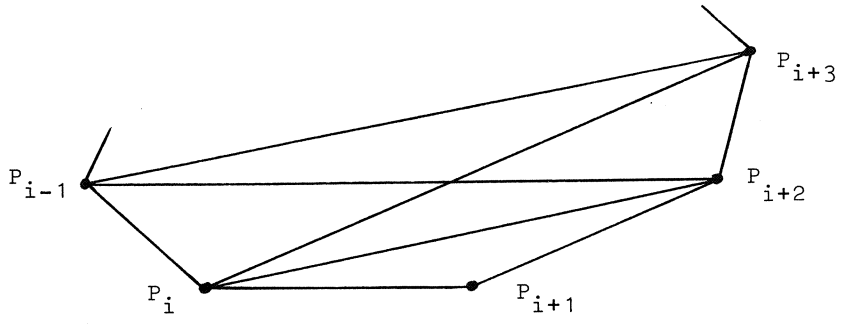
Laat V een affien vlak zijn over een lichaam K. Een n-hoek is een cyclisch geordend rijtje van n punten P_1, P_2, ..., P_n. We zullen steeds veronderstellen dat n minstens 3 is, en dat niet alle punten op één lijn liggen. Verder laten we alle euclidische achtergronden over regelmatige veelhoeken rusten. We definiëren affien regelmatige veelhoeken op de volgende manier:

DEFINITIE: de n-hoek P_1, ..., P_n heet (affien) regelmatig als

1) $P_iP_{i+1} \parallel P_{i-1}P_{i+2}$ voor alle i, en

2) $P_iP_{i+2} \parallel P_{i-1}P_{i+3}$ voor alle i

(we gebruiken de cyclische notatie: $P_{i+n} = P_i$ voor alle i. Dus $P_o = P_n$, $P_{-1} = P_{n-1}$, etc.)



Natuurlijk is het helemaal niet zeker dat deze definitie iets oplevert dat een beetje aansluit bij ons intuïtieve idee van wat een "affien regelmatige veelhoek" zou moeten zijn. Maar het bovenstaande geeft misschien wel enige hoop dat we - althans in

het reële affiene vlak - krijgen wat we hebben willen.

Een paar eenvoudige gevallen

Eerst maar eens een paar simpele consequenties van de definitie. Hoe zit het als bijvoorbeeld n = 3 ? Dan staat er voor i = 1:

1) $P_1P_2 \parallel P_3P_3$ en

2) $P_1P_3 \parallel P_3P_1$.

Voor i = 2 en i = 3 krijgen we iets dergelijks. Maar dat zijn lege eisen: er is altijd aan voldaan, als we tenminste onder P_3P_3 gewoon het punt P_3 verstaan, en afspreken dat elk punt evenwijdig is met elke lijn. Inderdaad is het niet zo gek om elke driehoek affien regelmatig te noemen. In het reële geval kun je elke driehoek krijgen als projectie van een regelmatige driehoek. In het algemeen is elke driehoek ook een affiene basis van het affiene vlak. Elke driehoek is dus ook door een affiene transformatie in elke andere driehoek over te voeren.

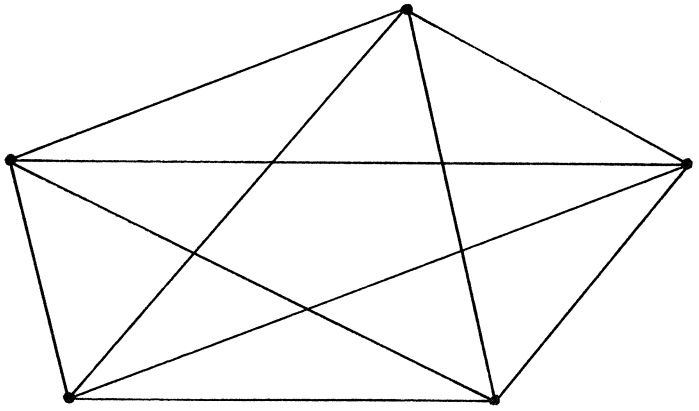
Wat zijn de eisen bij vierhoeken? Neem eerst maar weer i=1.

1) $P_1P_2 \parallel P_4P_3$

2) $P_1P_3 \parallel P_4P_4$.

Voor de andere drie waarden van i krijg je net zo iets. Nu is telkens de tweede eis leeg. En de eerste eis komt er op neer dat de vierhoek een parallellogram moet zijn. Parallellogrammen zijn blijkbaar de affien regelmatige vierhoeken.

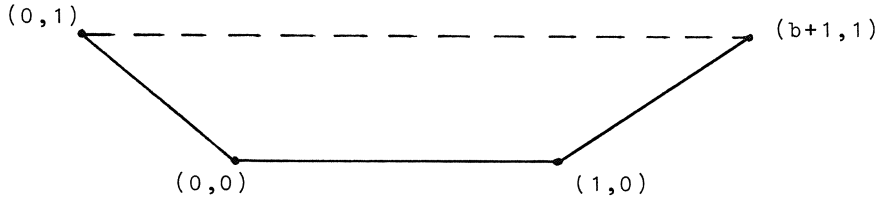
Voor vijfhoeken kunnen de voorwaarden vertaald worden in de eis dat elk van de vijf diagonalen evenwijdig moet zijn aan de tegenover liggende zijde.



Het spoor

Nu de analytische aanpak. Laat $(P_1, \ldots, P_n)$ een affien regelmatige n-hoek zijn. Nog steeds veronderstellen we dat $n \geq 3$ is, dat niet alle punten op één lijn liggen, en dat alle punten verschillend zijn. We gebruiken weer de cyclische notatie.

Stel dat P_{-1}, P_0, P_1 niet op één lijn liggen. Kies deze drie punten dan als affiene basis, en neem de coördinaten zo, dat $P_{-1} = (0,1)$, $P_o = (0,0)$ en $P_1 = (1,0)$. Dan moet P_2 ergens op de lijn $x_2 = 1$ liggen, dus $P_2 = (b+1,1)$ voor zeker getal $b \in K$. (Waarom we b+1 schrijven en niet b, zal zodadelijk blijken.)



We zien dat de gevallen b = -1 en b = 0 corresponderen met n = 3 (de driehoek) en n = 4 (de regelmatige vierhoek, oftewel het parallellogram). Voordat we verder gaan merken we op dat er natuurlijk een affiene afbeelding α is die het drietal (P_{-1},P_o,P_1) overvoert in het drietal (P_o,P_1,P_2), namelijk (in het gekozen coördinatenstelsel):

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} b & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Deze afbeelding α is een speciale affiene afbeelding in de letterlijke zin van het woord: de determinant van het lineaire deel (gegeven door de 2x2 matrix) is gelijk aan 1. U ziet ook dat α inderdaad een lineaire afbeelding induceert op de vectorruimte van de verschilvectoren.

De affiene afbeelding α wordt slechts door één parameter bepaald: het getal b. Dat getal is ook het spoor van de matrix A die bij α hoort. Het legt bovendien de plaats van P_2 vast op de lijn door P_{-1} evenwijdig aan P_oP_1.

Wat zou de afbeelding α trouwens doen met het punt P_2 ? We weten dat P_{-1} overgaat in P_o, P_o in P_1, P_1 in P_2. Het is niet moeilijk te bewijzen wat U natuurlijk al vermoedt: $\alpha(P_2) = P_3$ en zo verder: α verwisselt alle hoekpunten cyclisch! Dat betekent ook dat α^n alle punten weer op hun plaats brengt, en omdat $P_{-1}P_oP_1$ een affiene basis is, moet dus gelden dat α^n de identieke afbeelding is. De <u>orde</u> van α in de groep van alle affiene transformaties is dus precies n.

<u>Het middelpunt van een regelmatige veelhoek</u>

Laten we eens kijken of α invariante punten heeft. Als (x_1,x_2) de coördinaten van zo'n punt zijn, moet dus gelden

$$x_1 = b\,x_1 - x_2 + 1$$
$$x_2 = x_1$$

dus

$$x_1 = x_2 = (2 - b)^{-1}.$$

Natuurlijk moeten we $b \neq 2$ veronderstellen. Het geval $b = 2$ zullen we apart bekijken.

Als $b \neq 2$ dan is $M = ((2-b)^{-1},(2-b)^{-1})$ het <u>enige</u> invariante punt van α. We noemen M dan het <u>middelpunt</u> van de regelmatige veelhoek. Het ligt voor de hand om dan op nieuwe affiene coördinaten over te stappen door M als oorsprong te kiezen. Nemen we verder $P_1 = (1,0)$ en $P_2 = (0,1)$, dan wordt α:

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

(bedenk dat de determinant en het spoor niet veranderen door coördinatentransformaties!).

<u>De gevallen $K = \mathbb{R}$ en $K = \mathbb{C}$</u>

Stel dat we te maken hebben met een reëel of een complex affien vlak. Omdat $\alpha^n = id$, moeten de twee (complexe) eigenwaarden voldoen aan $\lambda^n = 1$, en omdat hun product gelijk is aan de determinant, dus gelijk is aan 1, zijn het complex geconjugeerde primitieve n-de machts eenheidswortels, d.w.z., ze zijn van de

vorm $e^{\pm i\theta}$ met $\theta = \frac{2k\pi}{n}$ en g.g.d.(k,n) = 1.

En het spoor? Dat is de som van de eigenwaarden, dus $b = e^{i\theta} + e^{-i\theta} = 2\cos\theta$. Hiermee zien we welke waarden van b in aanmerking komen voor regelmatige n-hoeken in het reële of complexe affiene vlak. En kiezen we dan wederom nieuwe coördinaten, nu met M = (0,0), P_1 = (1,0), P_2 = ($\cos\theta$, $\sin\theta$), dan moet α blijkbaar van de vorm zijn:

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

(wéér geldt dat de determinant 1 moet zijn, en het spoor is $b = 2\cos\theta$).

Nu herkennen we α : het is (t.o.v. het standaardinproduct in dit coördinatenstelsel) de gewone rotatie om M over een hoek θ, en de affien regelmatige n-hoek is de gewone euclidische regelmatige n-hoek (eventueel een ster-n-hoek als k groter dan 1 is).

Het geval b = 2 zouden we nog apart bekijken. In de oorspronkelijke coördinaten is α dan:

$$\begin{aligned} x_1' &= 2x_1 - x_2 + 1 \\ x_2' &= x_1 \end{aligned}$$

Hieruit volgt $x_1' - x_2' = x_1 - x_2 + 1$. Omdat P_o = (0,0) betekent dit dat P_i moet liggen op de lijn $x_1 - x_2 = i$. Voor variërende i staat hier een serie evenwijdige lijnen. Maar als dit allemaal zo is, dan kan $P_n = P_o$ niet waar zijn, althans niet als K = $\mathbb{R}$ of K = $\mathbb{C}$.

Eindige affiene vlakken

Het kan natuurlijk wèl als K een eindig lichaam is, en dat is eigenlijk veel interessanter. Want wat is er dan aan de hand? Stel K = $\mathbb{F}_q$, een lichaam met q elementen. Uit de algebra weten we dat q dan de macht van een priemgetal moet zijn, dus $q = p^r$ met p een priemgetal. De gehele opzet van het bovenstaande blijft geldig. In het bijzonder de uitgangssituatie met de

coördinatenkeuze $P_{-1} = (0,1)$, $P_o = (0,0)$, $P_1 = (1,0)$ en $P_2 = (b+1,1)$. Er zijn nu precies q mogelijke posities voor het punt P_2 op de lijn $x_2 = 1$, corresponderende met de q verschillende waarden van het spoor b. Elke b geeft automatisch een zekere regelmatige n-hoek, want bij affiene vlakken over eindige lichamen heeft elke affiene afbeelding een eindige orde. We kunnen het bijvoorbeeld eens nagaan voor het geval dat $K = \mathbb{F}_3$. We hebben al gezien dat $b = -1$ (=2) een driehoek, en $b = 0$ een vierhoek (parallellogram) geeft. Blijft over $b = 1$. Dan is $P_{-1} = (0,1)$, $P_o = (0,0)$, $P_1 = (1,0)$, $P_2 = (2,1)$ en (steeds α toepassen) $P_3 = (2,2)$, $P_4 = (1,2)$ en $P_5 = (0,1) = P_{-1}$. Het is blijkbaar een regelmatige zeshoek. In het affiene vlak over $\mathbb{F}_3$ bestaan blijkbaar alleen maar regelmatige driehoeken, vierhoeken en zeshoeken.

Op net zo'n manier kunnen we het geval $K = \mathbb{F}_5$ behandelen. Als $b = -1$ (= 4) dan is $n = 3$; als $b = 0$ dan is $n = 4$, en op grond van het slot van de vorige paragraaf moet $b = 2$ wel $n = 5$ geven (want op elke lijn $x_1 - x_2 = i$ ligt een hoekpunt van de n-hoek). Net als zopas geeft $b = 1$ de regelmatige zeshoek ($n = 6$), en ons rest dus nog slechts het geval $b = 3$. De berekeningen kunnen dan aanzienlijk versneld worden als we over gaan op nieuwe coördinaten met het middelpunt als de oorsprong, en $P_1 = (1,0)$, $P_2 = (0,1)$. Hierboven hebben we gezien dat α dan de volgende gedaante krijgt ($b = 3$):

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Hiermee berekenen we $P_3 = (4,3)$, $P_4 = (2,3)$, $P_5 = (2,1)$, $P_6 = (4,0)$, $P_7 = (0,4)$, $P_8 = (1,2)$, $P_9 = (3,2)$, $P_{10} = (3,4)$ en $P_{11} = (1,0) = P_1$. Dit is dus een regelmatige tienhoek.

Voor $K = \mathbb{F}_7$ vindt men op dezelfde manier dat met $b = 0, 1, \ldots, 6$ corresponderen resp. $n = 4, 6, 7, 8, 8, 14, 3$. Iets lastiger worden de berekeningen als q een echte macht van een priemgetal is. Voor $q = 2^3 = 8$ zijn de mogelijke waarden voor n: 4, 3, 7 en 9. De zevenhoeken en negenhoeken komen dan ook nog elk in drie verschillende soorten voor.

Al deze, min of meer experimenteel verkregen gegevens vragen natuurlijk om een algemeen resultaat. Hier is het:

STELLING (J.van de CRAATS & J.SIMONIS, 1986): In het affiene vlak over $\mathbb{F}_q$ ($q = p^r$, p priem) bestaan regelmatige n-hoeken dan en slechts dan als n ($\geq$ 3) voldoet aan een van de volgende voorwaarden:

- n = p,
- n = 2p,
- n is een deler van q-1,
- n is een deler van q+1.

Bij het bewijs spelen de eigenwaarden van het lineaire deel van α (in $\mathbb{F}_q$ of een uitbreidingslichaam van $\mathbb{F}_q$) de hoofdrol.

Gevraagd naar z'n mening over de rol van matrices in de meetkunde schijnt een beroemd wiskundige eens te hebben opgemerkt dat de wiskunde er altijd beter van wordt als je matrices zoveel mogelijk tracht te vermijden. Met deze enigszins provocerende uitspraak bedoelde hij natuurlijk dat je rekenpartijen zoveel mogelijk uit de weg moet gaan, en moet proberen coördinaatvrije formuleringen en resultaten te verkrijgen. Toch zijn matrices, mits met beleid gehanteerd, een fraai hulpmiddel om meetkundige problemen op te lossen. Ik hoop daarvan een boeiend voorbeeld te hebben gegeven.

Literatuur

H.S.M.COXETER, Introduction to Geometry, John Wiley, New York, 2nd ed., 1969

J.van de CRAATS & J.SIMONIS, Affinely regular polygons, verschijnt binnenkort

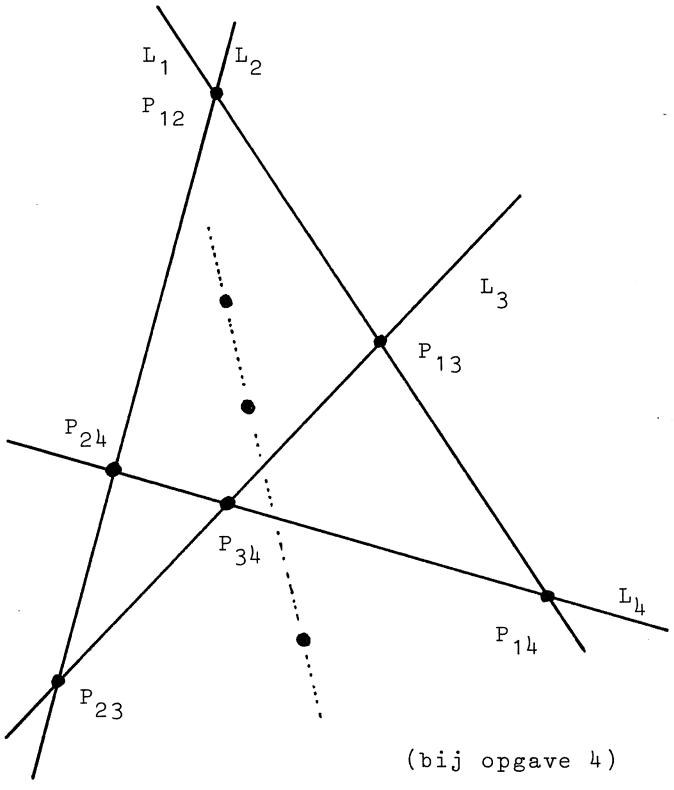
Opgaven:
Als toegift nog een paar oefenvraagstukken voor de geïnteresseerde toehoorder.

1. Bewijs dat in een affien regelmatige zeshoek de drie diagonalen die overstaande hoekpunten verbinden, door één punt gaan.
2. Bewijs dat de hoekpunten van de regelmatige tienhoek in het vlak over $\mathbb{F}_5$ afwisselend op twee parallelle lijnen liggen. Probeer dit resultaat te generaliseren.

3. Bewijs: zijn in een vijfhoek vier diagonalen evenwijdig met vier zijden, dan is de vijfde diagonaal evenwijdig met de vijfde zijde.
4. Gegeven zijn vier lijnen L_1, L_2, L_3, L_4 die elkaar in zes verschillende punten $P_{ij} = L_i \cap L_j$ $(i \neq j)$ snijden. Bewijs dat de middens van de drie lijnstukken $P_{12}P_{34}$, $P_{13}P_{24}$ en $P_{14}P_{23}$ op één lijn liggen.
5. Gegeven zijn vijf verschillende punten P_1, P_2, P_3, P_4, P_5 in de ruimte, en stel dat t een wortel is van de vergelijking $t^2 = t + 1$. Bewijs dat de vijf punten

$$Q_i = \frac{1}{2+t}\,(P_{i-1} + t\,P_i + P_{i+1})$$

de hoekpunten vormen van een vlakke affien regelmatige vijfhoek.



(bij opgave 4)

MATRICES EN GRAFEN

W.H. Haemers
Katholieke Hogeschool Tilburg

1. De ongelijkheid van Fisher 47

2. Grafen en eigenwaarden 50

3. Moore grafen 54

Referenties 57

MATRICES EN GRAFEN

In dit hoofdstuk zullen we met behulp van matrices enige resultaten afleiden voor grafen en andere combinatorische structuren. Matrixtheorie is een zeer belangrijk hulpmiddel in de combinatoriek. Er zijn belangrijke resultaten die puur combinatorisch van aard zijn, maar die zonder gebruik van matrixtheorie onoplosbaar lijken. In veel gevallen blijkt dan het bewijs met behulp van matrices verrassend elementair te zijn. In dit hoofdstuk worden enige voorbeelden van dit verschijnsel behandeld.

We gebruiken de volgende notatie. Een matrix waarvan elk element 1 is noteren we als J, en O staat voor de matrix waarvan elk element 0 is. Een vector waarvan elke coördinaat 1 is (een kolom van J dus) noteren we als j.

§ 1 DE ONGELIJKHEID VAN FISHER

Beschouw het voorbeeld van de tramlijnen en haltes uit Hoofdstuk 1. Dit tramnet voldoet aan de eigenschap dat elke twee haltes op precies een tramlijn liggen (tramreizigers hoeven dus niet over te stappen of te twijfelen over de te nemen lijn). Bovendien heeft elke lijn evenveel haltes en ligt elke halte op precies evenveel lijnen. Men kan zich afvragen onder welke voorwaarden zo'n tramnet te realiseren is en wat voor eigenschappen er kunnen worden afgeleid. Een belangrijke eigenschap is de ongelijkheid van Fisher die zegt dat als niet alle haltes op een tramlijn liggen, er tenminste zoveel tramlijnen als haltes zijn. Het bewijs van deze ongelijkheid is een mooie illustratie van het gebruik van matrices. We zullen in deze paragraaf dit bewijs dan ook geven.

Behalve Fishers ongelijkheid zijn er nog andere voorwaarden bekend die nodig zijn voor het bestaan van zo'n tramlijnen stelsel. Ook zijn er verschillende constructiemethoden bekend. Maar lang niet alles is hiermee gedekt. Een van de grote open problemen in dit vakgebied is de vraag of er zo'n tramlijnennet bestaat met 100 haltes en 10 haltes per lijn (het zogeheten affiene vlak van de orde 10).

Het tramnet is een speciaal geval van een blok-design. Een blok-design met parameters v,b,k,r en λ ($1 \leqslant k \leqslant v-1$) is een verzameling V met v elementen samen met een collectie deelverzamelingen van V, blokken geheten (b is het

aantal blokken), zodanig dat aan de volgende eisen is voldaan.

i. Elk blok heeft k elementen,

ii. elk element van V zit in precies r blokken,

iii. elk paar elementen van V komt voor in precies λ blokken.

Het stelsel tramlijnen uit Hoofdstuk 1 is een voorbeeld van een blok-design met parameters $v = b = 7$, $k = r = 3$ en $\lambda = 1$. Overigens is de bovenstaande definitie overbepaald. Met behulp van teltechnieken kan eis (ii) uit (i) en (iii) worden afgeleid. Blok-designs worden ondermeer gebruikt bij statistische proeven. De consumentenbond wil bijvoorbeeld de smaak testen van 7 merken cola. Daartoe krijgt een aantal proefpersonen ieder 3 merken cola om op smaak te vergelijken. Om de test zo eerlijk mogelijk te houden is het van belang dat elk merk even vaak wordt geproefd en dat elk tweetal merken door evenveel proefpersonen wordt vergeleken. Het tramlijnennet uit Hoofdstuk 1 geeft een mogelijk schema voor deze cola test.

Bij een blok-design hoort een incidentiematrix N. De rijen van N zijn de elementen van V en de kolommen van N zijn de blokken. Verder is N als volgt gedefinieerd.

$$(N)_{ij} = \begin{cases} 1 & \text{als element } i \text{ in blok } j \text{ zit,} \\ 0 & \text{anders.} \end{cases}$$

We zullen eis (i) t/m (iii) vertalen in termen van matrix N. Eis (i) betekent dat elke kolomsom van N gelijk is aan k, dus

i'. $N^t j = k j.$

Evenzo is eis (ii) gelijkwaardig met

ii'. $N j = r j.$

Eis (iii) betekent dat $(NN^t)_{ij} = \lambda$ voor $i \neq j$, immers $(NN^t)_{ij}$ is het inwendig product van rij i met rij j en dus gelijk aan het aantal malen dat er op overeenkomstige posities een 1 staat. Hieruit volgt

iii'. $N N^t = \lambda J + (r - \lambda) I.$

De stelling van Fisher is een direct gevolg van (iii').

STELLING 1. Voor een blok-design met parameters v,b,k,r en λ geldt

$$b \geqslant v.$$

BEWIJS. We zullen laten zien dat rang N = v. Hieruit volgt meteen de ongelijkheid want de rang van een matrix is niet groter dan het aantal kolommen. We definieren

$$N' = \frac{1}{r-\lambda} N^t - \frac{\lambda}{r(r-\lambda)} J.$$

Het is niet moeilijk in te zien dat $r-\lambda \neq 0$ als $0 < k < v$, zoals vereist in de definitie. Gebruik makend van (ii') en (iii') vinden we

$$NN' = \frac{1}{r-\lambda} NN^t - \frac{r\lambda}{r(r-\lambda)} J = \frac{\lambda}{r-\lambda} J + I - \frac{\lambda}{r-\lambda} J = I.$$

Een fundamentele eigenschap van de rang zegt dat de rang van het product van matrices niet groter is dan de rang van elke factor. Daarom geldt

$$\text{rang } N \geqslant \text{rang } NN'.$$

Met rang NN' = rang I = v en rang $N \leqslant v$ volgt hieruit dat rang N = v. Hiermee is de ongelijkheid van Fisher bewezen.

Bovenstaande afleiding wijkt iets af van de gebruikelijke bewijsmethode. Meestal wordt bewezen dat $NN^t = \lambda J + (r-\lambda)I$ niet singulier is, waaruit de ongelijkheid volgt op dezelfde manier. Het voordeel van onze aanpak is dat er meteen conclusies getrokken kunnen worden voor het geval v = b. Het blokdesign wordt dan symmetrisch genoemd. Inderdaad, als v = b dan is N' de inverse matrix van N. Er geldt dus

$$NN' = N'N = I.$$

Met (i') volgt

$$N'N = \frac{1}{r-\lambda} N^t N - \frac{k\lambda}{r(r-\lambda)} J.$$

Er geldt dat $k = r$ als $b = v$, want het totaal aantal enen in N is zowel gelijk aan bk als aan vr. Dus volgt nu

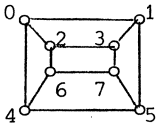
$$N^t N = J + (r-\lambda)\, I.$$

Dit betekent dat N^t ook de incidentiematrix is van een blok-design. Anders gezegd, de doorsnede van elk tweetal blokken bevat precies λ elementen.

§ 2 GRAFEN EN EIGENWAARDEN

Beschouw de volgende graaf

G_1 :



Bij deze graaf maken we de zogeheten verbindingsmatrix A als volgt

A_1 :

	0	1	2	3	4	5	6	7
0	0	1	1	0	1	0	0	0
1	1	0	0	1	0	1	0	0
2	1	0	0	1	0	0	1	0
3	0	1	1	0	0	0	0	1
4	1	0	0	0	0	1	1	0
5	0	1	0	0	1	0	0	1
6	0	0	1	0	1	0	0	1
7	0	0	0	1	0	1	1	0

Dus $(A)_{ij} = 1$ als punt i en j verbonden zijn en $(A)_{ij} = 0$ als $i = j$ of als i en j niet verbonden zijn. Als we de punten van G_1 in een andere volgorde zetten krijgen we een andere verbindingsmatrix. Bijvoorbeeld

A_1' :

	0	3	5	6	7	4	2	1
0	0	0	0	0	0	1	1	1
1	0	0	0	0	1	0	1	1
2	0	0	0	0	1	1	0	1
3	0	0	0	0	1	1	1	0
4	0	1	1	1	0	0	0	0
5	1	0	1	1	0	0	0	0
6	1	1	0	1	0	0	0	0
7	1	1	1	0	0	0	0	0

Het is duidelijk dat A_1 en A_1' dezelfde eigenwaarden hebben, immers $\det(A_1 - x\,I) = \det(A_1' - x\,I)$. In het algemeen geldt dat de eigenwaarden van de verbindingsmatrix van een graaf niet afhangen van de volgorde van de pun-

ten. We spreken daarom vaak kortweg over de eigenwaarden van een graaf. De liefhebber kan verifiëren dat de eigenwaarden van G_1 zijn 3,1,-1 en -3 met multipliciteiten 1,2,2 en 1 respectievelijk. In dit hoofdstuk beschouwen we alleen eindige grafen, ongericht en zonder lussen of meervoudige verbindingen. Dit houdt in dat de verbindingsmatrix symmetrisch is met 0 op de diagonaal en 0 of 1 elders. De eigenwaarden zijn dan reëel en we noteren ze als $\lambda_1 \geqslant \ldots \geqslant \lambda_v$, waarin v het aantal punten van de graaf is. We zullen nu een aantal begrippen introduceren.

Een graaf heet k-*regulier* als elk punt met precies k punten verbonden is. Bijvoorbeeld G_1 is 3-regulier.

Een (niet lege) graaf heet *bipartiet* als de punten in twee deelverzamelingen verdeeld kunnen worden, zodanig dat er binnen de beide deelverzamelingen geen verbindingen bestaan. Graaf G_1 bijvoorbeeld is bipartiet, hetgeen duidelijk wordt geïllustreerd door matrix A_1' .

Een graaf heet *samenhangend* als vanuit elk punt naar elk ander punt een wandeling is via de verbindingen. Graaf G_1 is bijvoorbeeld samenhangend.

Veronderstel nu we hebben een k-reguliere graaf G met verbindingsmatrix A. Dan is elke rijsom (en kolomsom) van A gelijk aan k. Hieruit volgt

$$A\,j = k\,j.$$

Dus k is een eigenwaarde van G met eigenvector j. De volgende stelling laat zien dat k een bijzondere eigenwaarde is.

STELLING 2. Voor een k-reguliere graaf G geldt

i. $\lambda_1 = k$ met eigenvector j,
ii. als G samenhangend is dan heeft k multipliciteit 1, dus $\lambda_2 < k$,
iii. $-\lambda_v \leqslant k$,
iv. als G bipartiet is, dan is $-\lambda_v = k$.

BEWIJS. (1): Laat $x = (x_1,\ldots,x_v)^t$ een eigenvector zijn voor de eigenwaarde λ. Laat x_p de coördinaat van x zijn die in absolute waarde het grootst is. Vervang x door -x indien $x_p < 0$. Dan geldt het volgende

$$\lambda x_p = (\lambda x)_p = (Ax)_p = a_{p1}x_1 + \ldots + a_{pv}x_v \leqslant$$
$$\leqslant a_{p1}x_p + \ldots + a_{pv}x_p = kx_p .$$

Dus $\lambda \leqslant k$ voor elke eigenwaarde λ.

(ii): Stel $\lambda = k$. In bovenstaande afleiding geldt dan gelijkheid, dus moet $x_i = x_p$ als $a_{pi} \neq 0$, ofwel als i en p verbonden zijn. Evenzo moet $x_j = x_i = x_p$ als j en i en i en p verbonden zijn, enzovoort. Dus $x_i = x_p$ voor alle i als G samenhangend is. Maar dan geldt $x = x_p j$. Dus is de multipliciteit van k gelijk aan 1.

(iii): Gebruik makend van $|a_{p1}x_1 + \ldots + a_{pv}x_v| \leqslant a_{p1}|x_1| + \ldots + a_{pv}|x_v|$, maar voor de rest analoog aan (i) vinden we

$$|\lambda| x_p \leqslant k x_p .$$

(iv): Als G bipartiet is kunnen we A als volgt nemen

$$A = \begin{pmatrix} O & A_{12} \\ A_{21} & O \end{pmatrix} .$$

Uit $Aj = kj$ volgt $A_{12}j = kj$ en $A_{21}j = kj$. Hiermee vinden we

$$\begin{pmatrix} O & A_{12} \\ A_{21} & O \end{pmatrix} \begin{pmatrix} j \\ -j \end{pmatrix} = -k \begin{pmatrix} j \\ -j \end{pmatrix} ,$$

dus -k is eigenwaarde van A.

Een belangrijke stelling over matrices met niet-negatieve elementen is die van Perron-Frobenius. Bovenstaande stelling is een direct gevolg hiervan. Er volgt echter meer. Zo gelden (ii) en (iv) ook omgekeerd: Als $\lambda_2 < k$ dan is G samenhangend, en als $-\lambda_v = k$ dan is G bipartiet.

Met behulp van eigenwaarden kunnen vaak ongelijkheden worden afgeleid met betrekking tot deelstrukturen van een graaf. Een voorbeeld van zo'n deelstructuur is een cokliek. Een *cokliek* (of onafhankelijke verzameling) is een deelverzameling van de punten van een graaf waarbinnen geen punten verbonden zijn. Bijvoorbeeld de punten 0,3,5 en 6 vormen een cokliek in G_1.

STELLING 3. Het aantal punten van een cokliek van een k-reguliere graaf met v punten is ten hoogste gelijk aan

$$v \frac{k - \lambda_v}{-\lambda_v} .$$

BEWIJS. Laat c het aantal punten van de cokliek zijn. We kunnen de verbindingsmatrix A als volgt nemen

$$A = \begin{pmatrix} O & A_{12} \\ A_{21} & A_{22} \end{pmatrix} ,$$

waarin O c rijen en c kolommen heeft. We maken nu de 2 bij 2 matrix B zo dat b_{ij} de gemiddelde rijsom is van A_{ij} en $b_{11} = 0$. De rijsom van A_{12} is k. Het aantal enen in A_{12} is gelijk aan kc. Dus de rijsom van $A_{21} = A_{12}^t$ is gemiddeld $\frac{kc}{v-c}$. De gemiddelde rijsom van A_{21} en A_{22} samen is k, dus

$$B = \begin{pmatrix} 0 & k \\ \frac{kc}{v-c} & k - \frac{kc}{v-c} \end{pmatrix} .$$

Nu is er een stelling die zegt dat elke eigenwaarde μ van B voldoet aan

$$\lambda_1 \geqslant \mu \geqslant \lambda_v .$$

De eigenwaarden μ_1 en μ_2 van B zijn k en $-\frac{kc}{v-c}$, immers Bj = kj en spoor B = $= \mu_1 + \mu_2 = k - \frac{kc}{v-c}$. De ongelijkheid $\lambda_1 \geqslant k \geqslant \lambda_v$ levert niets op, maar $-\frac{kc}{v-c} \geqslant \geqslant \lambda_v$ leidt tot de gewenste ongelijkheid.

We zullen het nut van dit type ongelijkheden illustreren met een voorbeeld uit de coderingstheorie.

Een ruimteschip zendt op haar reis door het zonnestelsel berichten naar de aarde. Elk bericht is opgebouwd uit rijtjes bits ter lengte n. Door ruis bestaat er een kans dat een bericht fout op aarde aankomt. Daarom gebruikt men voor de berichten niet alle mogelijke 2^n rijtjes, maar slechts een slim gekozen deelverzameling hieruit. Om dit duidelijk te maken definieren we de *afstand* tussen twee rijtjes als het aantal coördinaten waarop ze verschillen. Men kiest nu de zinvolle rijtjes (codewoorden genoemd) zo dat de onderlinge

afstand tussen elk tweetal codewoorden tenminste een vast getal d is. Als d = 3 bijvoorbeeld betekent dit dat als er een codewoord door ruis op een positie is verminkt, het nog steeds duidelijk is welk codewoord het geweest is. Men zegt dan dat de code één fout kan verbeteren. We maken nu een graaf G. De punten van G zijn alle 2^n rijtjes. Twee punten zijn verbonden als hun afstand kleiner is dan d. De codewoorden vormen dan een cokliek in G. Bijvoorbeeld graaf G_1 is zo'n graaf voor n = 3 en d = 2. De binaire representaties van de labels geven de bijbehorende rijtjes. Als codewoorden kunnen we dan de punten 0,3,5 en 6 nemen. We stellen ons nu de vraag hoeveel codewoorden er maximaal genomen kunnen worden als d = 2 voor willekeurige n. Alle rijtjes met een even aantal enen voldoen. Dit levert 2^{n-1} codewoorden. We zullen met behulp van Stelling 3 laten zien dat het niet beter kan. Behalve alle rijtjes met een even aantal enen zijn ook de rijtjes met een oneven aantal enen onderling niet verbonden. G is dus bipartiet. G is ook k-regulier (met $k = n$). Dus $\lambda_1 = k$ en $\lambda_v = -k$ volgens Stelling 2. Stelling 3 geeft nu dat het aantal codewoorden ten hoogste gelijk is aan

$$v \frac{k - (-k)}{-(-k)} = \tfrac{1}{2}v = 2^{n-1} .$$

Ook voor d > 2 zijn met eigenwaardentechnieken op vergelijkbare manier grenzen af te leiden voor het maximaal aantal codewoorden.

§ 3 MOORE GRAFEN

Een Moore graaf is een k-reguliere graaf die voldoet aan de volgende twee eigenschappen.

i. Voor elk paar verbonden punten bestaat er geen punt dat met beide punten verbonden is.

ii. Voor elk paar niet verbonden punten is er precies één punt met beide punten verbonden.

Deze eigenschappen houden in dat er in maximaal twee stappen van elk punt naar elk ander punt gegaan kan worden, terwijl het totaal aantal verbindingen zo klein mogelijk is. Dergelijke grafen zijn van nut voor telecommunicatienetten waarbij aan de ene kant zuinig moet worden omgesprongen met transmissiemiddelen, terwijl aan de andere kant het aantal tussenstations klein moet zijn. Voorbeelden van Moore grafen zijn:



De tweede graaf staat bekend als de Petersen graaf.

Het is niet moeilijk in te zien dat het aantal punten van een Moore graaf gelijk is aan $k^2 + 1$. We laten dit als oefening over aan de lezer. De stelling waaraan deze paragraaf is gewijd zegt dat er maar weinig Moore grafen zijn.

STELLING 4. Voor een Moore graaf geldt $k = 1, 2, 3, 7$ of 57.

BEWIJS. Laat A de verbindingsmatrix zijn van een Moore graaf G. Dan geldt

(1) $\qquad Aj = kj.$

Verder is $(A^2)_{ij} = (AA^t)_{ij}$ het inwendig product van rij i en rij j van A en dus gelijk aan het aantal punten dat met i èn j verbonden is. Hieruit volgt

$$A^2 = J - A - I + kI = J - A + (k-1)I \ .$$

Dit is te herschrijven als

(2) $\qquad (A - s_1 I)(A - s_2 I) = J \ , \text{ met } s_{1,2} = -\tfrac{1}{2} \pm \tfrac{1}{2}\sqrt{4k-3} \ .$

Uit (1) en (2) volgt

(3) $\qquad (A - s_1 I)(A - s_2 I)(A - kI) = 0 \ .$

Dit betekent dat s_1, s_2 en k de enige mogelijke eigenwaarden van A zijn. Inderdaad, vermenigvuldig (3) rechts met een eigenvector x behorend bij een eigenwaarde λ, dan volgt

$$(\lambda - s_1)(\lambda - s_2)(\lambda - k) = 0 \ .$$

Dus $\lambda = s_1$, $\lambda = s_2$, of $\lambda = k$. De multipliciteit van k is gelijk aan 1, zoals bijvoorbeeld volgt uit Stelling 2. Laat m_1 en m_2 de multipliciteiten zijn van s_1 en s_2 respectievelijk. De som der multipliciteiten is het aantal punten van G. Dus

(4) $$m_1 + m_2 + 1 = k^2 + 1 .$$

De som der eigenwaarden is gelijk aan spoor A = 0, dus

(5) $$m_1 s_1 + m_2 s_2 + k = 0 .$$

Uit (4) en (5) volgt

$$m_2(s_2 - s_1) = -k - s_1 k^2 .$$

Substitutie van s_1 en s_2 en vermenigvuldigen met $\sqrt{4k-3}$ leidt tot

(6) $$2m_2(4k-3) = k(\sqrt{4k-3}(k-2) + k(4k-3)).$$

Daar $2m_2(4k-3)$ een geheel getal is, geldt k = 2 of 4k-3 is een kwadraat. Noem $4k-3 = w^2$, dan is $k = \frac{1}{4}(w^2+3)$. Uit formule (6) volgt dan

$$32m_2 w = (w^2+3)(w^2-5+4kw) = w^4+4kw^3-2w^2+12kw-15 .$$

Dus w is een deler van 15. Daarom is w = 1, 3, 5 of 15 en dus k = 1, 3, 7, of 57. Samen met de eerder gevonden mogelijkheid k = 2 levert dit de bewering van de stelling.

Voor k = 1, 2, 3 en 7 is bekend dat een Moore graaf bestaat. Voor k = 1, 2 en 3 zijn dit: twee verbonden punten, de vijfhoek en de Petersen graaf, respectievelijk. Ook is bekend dat er voor k = 1, 2, 3 en 7 maar precies een Moore graaf bestaat. Het bestaan van een 57-reguliere Moore graaf echter is nog een onopgelost probleem.

Tot slot merken we op dat grafen met veel structuur zoals Moore grafen erg belangrijk zijn voor de bestudering van eindige groepen. Stelling 4 heeft dan ook gevolgen in de groepentheorie.

REFERENTIES

Voor verdere bestudering van dit onderwerp wordt verwezen naar de volgende boeken.

1. N.L. Biggs, *Algebraic Graph Theory*, Cambridge Univ. Press, Cambridge 1974.

2. P.J. Cameron & J.H. van Lint, *Graphs, Codes and Designs*, London Math. Soc. Lecture Notes 43, Cambridge Univ. Press, Cambridge 1980.

3. D.M. Cvetcović, M. Doob & H. Sachs, *Spectra of Graphs*, Academic Press, New York 1980.

4. W.H. Haemers, *Eigenvalue Techniques in Design and Graph Theory*, Math. Centre Tract 121, Mathematisch Centrum, Amsterdam 1980.

LINEAIRE MODELLEN VOOR DYNAMISCHE VERSCHIJNSELEN

J.M. Schumacher
Centrum voor Wiskunde en Informatica, Amsterdam

1. Dynamische verschijnselen 61

2. Lineaire modellen 62

3. De matrix van een stelsel differentiaalvergelijkingen 64

4. Rang: (on)afhankelijkheid van vergelijkingen 66

5. De determinant: het aantal oplossingen van een vierkant stelsel 67

6. Inverteerbaarheid: transformaties van stelsels 68

7. Het vegen van een polynoommatrix 69

8. Open systemen: stelsels met meer onbekenden dan vergelijkingen 73

9. Literatuur 76

Lineaire modellen voor dynamische verschijnselen

1. Dynamische verschijnselen

Een verschijnsel waarvan we de ontwikkeling in de tijd willen beschrijven heet wel een 'dynamisch verschijnsel'; het adjektief 'statisch' wordt dan gereserveerd voor zaken waarbij de tijdsparameter niet in de beschouwingen wordt betrokken. De trilling van een brugdek die wordt opgewekt door een peloton marcherende soldaten is een dynamisch verschijnsel; het berekenen van de vervorming van hetzelfde brugdek onder het gewicht van een stilstaande vrachtwagen is een probleem in de statica. De invloed van de prijs van een vat olie op de rentestand kan in statische termen worden beschreven, als men het voor de analyse voldoende vindt om aan te nemen dat er een vast verband tussen deze twee gegevens bestaat; men kan ook geïnteresseerd zijn in beschouwingen over het vliegwieleffekt dat gepaard gaat met een plotselinge prijsverhoging of prijsverlaging van de olie, en dan is een dynamische beschrijving vereist.

In deze bijdrage zal aandacht worden besteed aan mathematische modellen voor dynamische verschijnselen. We gaan er van uit dat de te beschrijven grootheden kunnen worden weergegeven door reële getallen. Speciale aandacht gaat uit naar situaties waarbij in één probleem meerdere veranderlijken een rol spelen; in een macro-economische context zouden dat bijvoorbeeld kunnen zijn de de inflatie, de rentestand, de werkloosheid, en het bruto nationaal produkt. Voor de beschrijving van een trillende snaar zou men de verplaatsing en de snelheid van alle punten op de snaar als variabelen kunnen nemen, en dit zou dan leiden tot een beschrijving in termen van een oneindig aantal variabelen; hieronder zullen we ons echter beperken tot situaties waarbij het aantal veranderlijken eindig is.

Het mathematische hulpmiddel bij uitstek om processen van veranderingen te beschrijven is de theorie van de *differentiaalvergelijkingen*. De beweging van

een massa die aan een veer hangt wordt bijvoorbeeld beschreven door de vergelijking

$$mx''(t) = -kx(t)$$

waarbij m en k positieve konstanten zijn, en waarbij $x(t)$ de verplaatsing van de massa uit de evenwichtsstand op tijdstip t aangeeft. Een oplossing van deze vergelijking wordt gegeven door

$$x(t) = \sin(k/m)^{\frac{1}{2}}t\,;$$

inderdaad vertoont deze funktie het slingerende gedrag dat van een massa aan een veer verwacht kan worden. Een minder realistisch aspekt van de oplossing is dat de beweging altijd onverminderd doorgaat; in werkelijkheid zal de amplitude afnemen, en na verloop van zekere tijd zal de massa tot stilstand komen. De bovenstaande differentiaalvergelijking geeft alleen een *model*; het beste dat we hiervan kunnen hopen is dat het de te beschrijven verschijnselen weergeeft met een zekere graad van nauwkeurigheid en gedurende een bepaalde periode. Men kan een hogere graad van nauwkeurigheid en/of een langere geldigheidsduur bereiken door het model ingewikkelder te maken. In de praktijk is de keuze van een model vaak een compromis tussen enerzijds de gewenste nauwkeurigheid en geldigheidsduur, en anderzijds de mate van hanteerbaarheid.

2. Lineaire modellen

Een goed hanteerbare en daarom veel gebruikte modelklasse is die van de *lineaire* modellen, die worden gespecificeerd door lineaire relaties op te geven tussen de variabelen en hun afgeleiden. Het massa-veer-model van hierboven is een voorbeeld van een lineair model. Een voorbeeld van een model dat *niet* lineair is wordt gegeven door de vergelijking van de mathematische slinger:

$$Ml\alpha''(t) = \sin(\alpha(t))$$

waarin M het gewicht is dat aan de slinger hangt, l de lengte van de slinger voorstelt, en α de uitwijkingshoek is. De niet-lineariteit wordt hier veroorzaakt door het verschijnen van de sinusfunktie die werkt op de variabele $\alpha(t)$. Men kan natuurlijk opmerken dat voor kleine waarden van deze variabele de uitdrukking $\sin(\alpha(t))$ met goede benadering vervangen kan worden door $\alpha(t)$; de differentiaalvergelijking hierboven gaat dan over in

$$Ml\alpha''(t) = \alpha(t),$$

hetgeen in essentie hetzelfde is als de eerder gegeven massa-veer-vergelijking. Inderdaad kan men aantonen dat de oplossingen van de *gelineariseerde* vergelijking een benadering vormen voor de oplossingen van de oorspronkelijke vergelijking; de aanname van 'kleine uitwijkingen' is echter wel essentieel. In het algemeen kan men zeggen dat lineaire modellen goed voldoen als het gaat om beschrijving van situaties in de buurt van een evenwichtspunt; wat 'in de buurt van' precies betekent hangt sterk van de aard van de toepassing af. Niet-lineaire modellen zijn van belang voor het beschrijven van gedrag tussen

evenwichtspunten in (bijvoorbeeld het overwippen van stand '0' naar stand '1' van een transistor die fungeert als logisch element in een schakeling op een chip) of van periodiek gedrag (zoals het slaan van uw hart). Met een overmaat aan generalisatie zou men kunnen zeggen dat niet-lineaire modellen vooral populair zijn bij fysici en biologen, terwijl lineaire modellen veelvuldig in de handen verkeren van ingenieurs en econometristen. Aan niet-lineaire dynamische systemen is in het kader van de vakantiecursussen in eerdere jaren onder meer aandacht besteed door J. Grasman.[1] Hieronder zullen we proberen aan te tonen dat er ook (zelfs) met stelsels lineaire differentiaalvergelijkingen wiskundig heel wat aardigs valt te beleven.

In het bovenstaande hebben we het gehad over differentiaalvergelijkingen; daarmee wordt verondersteld dat de waarden van de variabelen beschikbaar zijn op een volledig tijdsinterval. Het komt nogal eens voor dat dit echter niet het geval is, bijvoorbeeld omdat er maar eens in de zoveel tijd gegevens beschikbaar zijn (bv. de werkloosheid in Nederland, in maandcijfers), of omdat het een verschijnsel betreft dat slechts met tussenpozen optreedt (het aantal broedende vinkenparen in Nederland). We zeggen dan dat er sprake is van een *diskrete tijdsparameter* (in tegenstelling tot de situatie met dóórlopende gegevens, die wordt gekarakteriseerd door een *continue tijdsparameter*). Processen 'in diskrete tijd' worden beschreven met *differentievergelijkingen*. Stel dat u in het bezit bent van een aantal konijnen met de volgende voortplantingseigenschappen: de nuljarigen planten zich niet voort, de éénjarigen produceren per paar twee nakomelingen, de tweejarigen doen hetzelfde, en de ouderen werpen niets. Dan kunt u het aantal nuljarigen in het jaar j, $x_0(j)$, berekenen met behulp van de volgende differentievergelijking:

$$x_0(j) = x_0(j-1) + x_0(j-2).$$

Zo kunt u dus de reeks van Fibonacci verkrijgen (als u er tenminste van af ziet om helemaal vooraan te beginnen; daar is namelijk niet aan de modelveronderstellingen voldaan). Zulke differentievergelijkingen kunnen voor een groot deel op dezelfde voet worden behandeld als differentiaalvergelijkingen, door de bewerking 'differentiatie' te vervangen door de bewerking 'verschuiving (met één tijdsstap)'. In het bijzonder gaat alles wat hieronder gezegd zal worden over stelsels differentiaalvergelijkingen mutatis mutantis ook op voor stelsels van differentievergelijkingen.

Een laatste inleidende opmerking dient gemaakt te worden over de mogelijkheid van *variabele parameters*. Als in het hierboven genoemde slingervoorbeeld het gewicht aan de slinger wordt gevormd door een zak zand, en in de zak bevindt zich een gaatje waardoor het zand langzaam wegloopt, dan is één van de parameters die in de vergelijking optreedt veranderlijk in de tijd. Dit is een situatie die we hieronder niet zullen beschouwen: we zullen kijken naar *stelsels van gewone lineaire differentiaalvergelijkingen met konstante coëfficienten*. We

1. J. Grasman, "Matrices en de theorie van dynamische systemen", in: *Hewet-plus wiskunde*, CWI Syllabus 1 (Vakantiecursus 1984), pp. 63-74.

bestuderen dus verschijnselen die aan verandering onderhevig zijn, maar we veronderstellen dat de wetten die de veranderlijkheid beschrijven zelf niet veranderen.

3. De matrix van een stelsel differentiaalvergelijkingen

We beginnen met te kijken naar één differentiaalvergelijking in één onbekende. Met 'differentiaalvergelijking' bedoelen we nu — en deze vernauwing van terminologie zal hieronder ook steeds worden toegepast — een homogene lineaire gewone differentiaalvergelijking met konstante reële coëfficienten, dus iets van de vorm

$$a_k x^{(k)}(t) + a_{k-1} x^{(k-1)}(t) + \cdots a_1 x'(t) + a_0 x(t) = 0$$

waarbij $x(t)$ de onbekende is, en de getallen $a_0,\ldots,a_k$ gegeven konstanten zijn. Een voorbeeld is de massa-veer-vergelijking van boven, die na een kleine herschrijving luidt:

$$mx''(t) + kx(t) = 0.$$

Vergelijkingen van dit type kunnen worden beschreven door middel van een bijbehorend *polynoom* in een (vooralsnog formele) variabele s. Het polynoom dat hoort bij de massa-veer-vergelijking is dan

$$p_{m\text{-}v}(s) = ms^2 + k,$$

en bij de hierboven gegeven algemene vergelijking hoort het algemene polynoom

$$p(s) = a_k s^k + a_{k-1} s^{k-1} + \cdots + a_1 s + a_0.$$

De schrijfwijze

$$p(s)x = 0$$

vatten we dan op als verkorte notatie voor die differentiaalvergelijking.

Het op deze manier associeren van een polynoom aan een differentiaalvergelijking kan men opvatten als een puur boekhoudkundige handeling — het op een andere manier beschrijven van hetzelfde —, maar interessant is natuurlijk met name dat als we de variabele s opvatten als een complexe variabele en het geassocieerde polynoom dus zien als een 'echt' polynoom, de eigenschappen hiervan belangrijke informatie geven over de oplossingen van de differentiaalvergelijking. Zoals: als λ een reële wortel is van het geassocieerde polynoom, dan is $e^{\lambda t}$ een oplossing van de differentiaalvergelijking. Als $\lambda \pm \mu i$ een stel complexe wortels van het polynoom is, dan zijn de funkties $e^{\lambda t}\sin \mu t$ en $e^{\lambda t}\cos \mu t$ beide oplossingen van de differentiaalvergelijking. En sterker nog, alle oplossingen van de differentiaalvergelijking kunnen worden geschreven als konstante lineaire kombinaties van de oplossingen die op deze manier worden gevonden. Het aantal lineair onafhankelijke oplossingen is precies gelijk aan het aantal wortels van het polynoom, en dat getal is weer gelijk aan de *graad* van het polynoom, dus de hoogste macht van s die voorkomt (met een coëfficient ongelijk 0).

Het polynoom van de massa-veer-vergelijking, $ms^2 + k$, is een tweedegraads polynoom met als wortels het complex toegevoegde paar $\pm(k / m)^{1/2} i$. Particuliere oplossingen van de vergelijking zijn dus $\sin(k / m)^{1/2} t$ en $\cos(k / m)^{1/2} t$, en de algemene oplossing van de vergelijking kan worden geschreven als

$$a \sin(k / m)^{1/2} t + b \cos(k / m)^{1/2} t$$

waarbij a en b willekeurige reële getallen zijn.

Het succes van deze benadering voor *één* vergelijking in *één* variabele leidt als vanzelf tot de gedachte dat een soortgelijke aanpak ook van nut zou moeten zijn voor *stelsels* van vergelijkingen in *meerdere* variabelen. Laten we, om een voorbeeld te hebben, aan onze massa-met-veer een tweede massa-met-veer vastmaken. Om het schrijfwerk te beperken nemen we de beide massa's en ook de beide veerkonstanten aan elkaar gelijk. De vergelijkingen worden dan

$$mx_1''(t) = -2kx_1(t) + kx_2(t)$$
$$mx_2''(t) = -kx_2(t) + kx_1(t).$$

Omdat we nu meerdere vergelijkingen in meerdere onbekenden hebben, moeten we hiermee niet een enkel polynoom associeren maar een *matrix* waarvan de elementen polynomen zijn:

$$\begin{bmatrix} ms^2 + 2k & -k \\ -k & ms^2 + k \end{bmatrix}.$$

Als we deze matrix, volgens de gewone regels, toepassen op de staande vector met componenten $x_1(t)$ en $x_2(t)$, en we vatten s weer op als 'differentiatie', dan krijgen we inderdaad het bovenstaande stelsel terug. Dus: *met een stelsel differentiaalvergelijkingen associeren we een matrix waarvan de elementen polynomen zijn.* Een dergelijke matrix noemt men een *polynoommatrix*; deze terminologie heeft het voordeel dat ze voor de hand liggend is. De schrijfwijze

$$P(s)x = 0$$

vatten we op als verkorte notatie voor een stelsel van differentiaalvergelijkingen in meerdere onbekenden; hier stelt x dus een *vector* van tijdsfunkties voor.

De vraag doet zich meteen voor: bestaat er weer zo'n nauw verband tussen het stelsel differentiaalvergelijkingen en de bijbehorende polynoommatrix? We zullen zien dat dat verband er inderdaad is, maar dat er rekening moet worden gehouden met allerlei verschijnselen. Ten eerste weten we uit de gewone matrixtheorie al dat werken met matrices soms lijkt op werken met getallen (optellen, vermenigvuldigen), maar dat er toch ook grote verschillen zijn: er moet op de afmetingen worden gelet, en bijvoorbeeld de kwestie van inverteerbaarheid is voor matrices heel wat moeilijker dan voor getallen. Bovendien hebben we te maken met matrices waarvan de elementen geen *getallen* zijn, maar *polynomen*. De wereld van de polynoommatrices kent een gemengde cultuur: er spelen begrippen een rol die afkomstig zijn uit de 'gewone' matrixtheorie (rang, determinant, inverteerbaarheid), maar ook zijn ideeën van belang die typisch te maken hebben met polynomen (nulpunten,

grootste gemene delers, rationale funkties).[1] Hieronder zullen we enige voortbrengselen van deze gemengde cultuur wat nader bekijken.

4. Rang: (on)afhankelijkheid van vergelijkingen

De (rij)rang van een matrix over de reële getallen wordt gedefinieerd als het maximale aantal lineair onafhankelijke rijen. Dat wil zeggen, de rang van een $p \times m$-matrix is k als onder ieder $(k+1)$-tal rijen dat uit de matrix wordt gekozen er minstens één is die geschreven kan worden als lineaire kombinatie van de andere, en als dat niet zo is wanneer er slechts k rijen mogen worden gekozen. Dit begrip is van belang in verband met onafhankelijkheid van vergelijkingen. De drie vergelijkingen

$$\begin{aligned} 3x + 4y + z &= 0 \\ 2x + y + 3z &= 0 \\ 4x + 7y - z &= 0 \end{aligned}$$

zijn bijvoorbeeld niet onafhankelijk: de derde vergelijking kan geschreven worden als tweemaal de eerste minus de tweede. Daaruit volgt dat ieder drietal getallen x, y, z dat aan de eerste twee vergelijkingen voldoet, automatisch ook aan de derde voldoet. De drie vergelijkingen kunnen dus tot twee worden gereduceerd (maar niet tot minder dan twee). Dit komt tot uitdrukking in het feit dat de matrix van het stel vergelijkingen, die te voorschijn komt als we de vergelijkingen herschrijven in de matrixvorm

$$\begin{bmatrix} 3 & 4 & 1 \\ 2 & 1 & 3 \\ 4 & 7 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

de rang 2 heeft.

De rijrang van een polynoommatrix definieren we in principe op dezelfde manier, maar we zijn iets voorzichtiger in de formulering. We zullen zeggen dat een stel rijen *afhankelijk* is als er een lineaire kombinatie bestaat van de rijen, met polynomiale coëfficienten (niet alle gelijk aan 0), die de nulrij oplevert. De *rijrang* van een polynoommatrix is dan het maximale aantal onafhankelijke rijen in de matrix.

Rijen in een polynoommatrix kunnen afhankelijk zijn in de juist gedefinieerde zin zonder dat één van de rijen geschreven kan worden als lineaire kombinatie (met polynomiale coëfficienten) van de andere. Bekijk bijvoorbeeld de volgende polynoommatrix:

$$\begin{bmatrix} s+1 & 2s+2 \\ s & 2s \end{bmatrix}.$$

1. Uit deze beschrijving moet niet de konklusie worden getrokken dat het gaat om twee culturen die vroeger niet gemengd waren. Integendeel, al vanaf het begin van de matrixtheorie (in de eerste helft van de vorige eeuw) is er volop aandacht geweest voor matrices waarvan de elementen uit een *ring* komen (bv. gehele getallen zijn, of polynomen). Waarschijnlijk pas na de tweede wereldoorlog

De rijen van deze matrix zijn afhankelijk, immers

$$s[s+1 \quad 2s+2] - (s+1)[s \quad 2s] = [0 \quad 0].$$

Toch is het niet zo dat één van de rijen geschreven kan worden als veelvoud van de andere. Volgens onze definitie is de rijrang van bovenstaande matrix gelijk aan 1. Het is niet moeilijk om na te gaan dat dit overeenstemt met het aantal 'onafhankelijke vergelijkingen' in het met de matrix geassocieerde stelsel. De vergelijkingen zijn:

$$x_1' + x_1 + 2x_2' + 2x_2 = 0$$
$$x_1' + 2x_2' = 0.$$

Door de tweede vergelijking af te trekken van de eerste kan dit stelsel worden herschreven tot

$$x_1 + 2x_2 = 0$$
$$x_1' + 2x_2' = 0.$$

Maar nu is het duidelijk dat alle oplossingen die aan de eerste vergelijking voldoen ook aan de tweede vergelijking zullen voldoen, zodat de tweede vergelijking dus overbodig is. Ook in het algemeen kan men bewijzen dat het aantal onafhankelijke vergelijkingen in een stelsel differentiaalvergelijkingen precies wordt gegeven door de rijrang van de bijbehorende matrix. En net zoals dat bij konstante matrices het geval is, is de rijrang van een polynoommatrix altijd gelijk aan de (overeenkomstig gedefinieerde) kolomrang.

5. De determinant: het aantal oplossingen van een vierkant stelsel

Hierboven is gesteld dat het aantal onafhankelijke oplossingen van één differentiaalvergelijking in één onbekende gelijk is aan de graad van het bijbehorende polynoom. Hoe zit dat voor stelsels? Als voorbeeld bekijken we het volgende stelsel:

$$x_1''(t) + x_1(t) + x_2'(t) = 0$$
$$x_1'(t) + x_2(t) = 0.$$

Hoeveel oplossingen heeft dit, denkt u? Wel, we kunnen de tweede vergelijking een keer differentieren en hem dan van de eerste aftrekken; dit leidt tot de konklusie dat $x_1(t) = 0$. Met dit gegeven blijkt dan uit de tweede vergelijking dat ook $x_2(t) = 0$. Met andere woorden, het stelsel heeft uitsluitend de triviale oplossing; het aantal onafhankelijke oplossingen is nul.

Men kan aantonen dat het aantal onafhankelijke oplossingen van een stelsel van n onafhankelijke differentiaalvergelijkingen in n onbekenden wordt bepaald door de *determinant* van de bijbehorende matrix. Zoals bekend wordt de determinant van een matrix over de reële getallen berekend op een systematische, zij het gecompliceerde wijze; bij de bewerkingen hoeft evenwel alleen opgeteld, afgetrokken en vermenigvuldigd te worden, en niet gedeeld (zie ook de bijdrage van Van der Blij in deze bundel). Dezelfde bewerkingen

werd de belangstelling voor matrices over de reële getallen zo sterk overheersend.

toepassen op een polynoommatrix levert daarom een polynoom als resultaat op. Bijvoorbeeld, de determinant van een twee-bij-twee-matrix wordt gegeven door

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

Dus de determinant van een twee-bij-twee polynoommatrix is

$$\det \begin{bmatrix} a(s) & b(s) \\ c(s) & d(s) \end{bmatrix} = a(s)d(s) - b(s)c(s).$$

Passen we deze regel toe op de matrix van het als voorbeeld gebruikte stelsel, dan krijgen we

$$\det \begin{bmatrix} s^2+1 & s \\ s & 1 \end{bmatrix} = s^2 + 1 - s^2 = 1.$$

We zien: *het aantal onafhankelijke oplossingen is gelijk aan de graad van de determinant.* Dit is in het algemeen waar; een bewijs zal hieronder worden gegeven (in §7).

6. Inverteerbaarheid: transformaties van stelsels

Een vierkante polynoommatrix A heet *inverteerbaar* als er een polynoommatrix B bestaat die zodanig is dat $AB = BA = I$, de eenheidsmatrix. Net zoals dat bij matrices over de reële getallen het geval is is een polynoommatrix inverteerbaar dan en slechts dan als zijn determinant inverteerbaar is — alleen is deze voorwaarde voor polynomen heel wat beperkender dan voor reële getallen. Een getal heeft een inverse dan en slechts dan als het niet gelijk is aan 0; voor een polynoom $p(s)$ geldt echter dat er een polynoom $q(s)$ bestaat zó dat $p(s)q(s) = 1$ dan en slechts dan als $p(s) = p_0$, een konstante ongelijk nul. Een inverteerbare polynoommatrix wordt ook wel eens een *unimodulaire* matrix genoemd.[1] Zo'n matrix hoeft niet konstant te zijn, zoals blijkt uit het in de vorige paragraaf gebruikte voorbeeld.

Als de matrix van een stelsel differentiaalvergelijkingen unimodulair is, dan heeft het stelsel alleen de nuloplossing. Dit is niet zo'n interessante situatie. Unimodulaire matrices ontlenen hun belang meer aan de volgende stelling: *als de matrix van een stelsel differentiaalvergelijkingen van links wordt vermenigvuldigd met een unimodulaire matrix, dan blijft de verzameling van oplossingen onveranderd.* Het is niet moeilijk om in te zien waarom dit zo is. Als $P(s)$ een $p \times m$-polynoommatrix is, en $U(s)$ is een $p \times p$ unimodulaire matrix, dan volgt uit $P(s)x = 0$ zeker dat ook $U(s)P(s)x = 0$. Dus iedere oplossing van het stelsel gegeven door $P(s)$ is ook een oplossing van het stelsel gegeven door

1. Dit is naar analogie met de *getallen van Gauss*, d.w.z. complexe getallen van de vorm $a + bi$ waarbij a en b gehele getallen zijn. De inverse van een dergelijk getal is weer van dezelfde vorm dan en slechts dan als zijn modulus gelijk is aan 1.

$U(s)P(s)$. Maar het omgekeerde is ook waar, want als we de inverse van $U(s)$ aangeven met $V(s)$, dan volgt uit $U(s)P(s)x = 0$ dat ook $V(s)U(s)P(s)x = 0$, dus $P(s)x = 0$.

Dit betekent dat we linksvermenigvuldiging met unimodulaire matrices kunnen gaan toepassen als middel om stelsels differentiaalvergelijkingen te vereenvoudigen, volgens (in principe) hetzelfde procédé als in gebruik is voor stelsels lineaire algebraïsche vergelijkingen. Rechtsvermenigvuldiging met unimodulaire matrices kan ook worden gebruikt; dit korrespondeert met transformaties op de variabelen. De vraag doet zich dan voor tot welke standaardvorm een polynoommatrix kan worden herleid door links- en rechtsvermenigvuldiging met unimodulaire matrices. Dit werd reeds in 1861 beantwoord door H. J. S. Smith,[1] die bewees dat een standaardvorm kan worden verkregen waarin alleen de elementen op de hoofddiagonaal ongelijk aan nul zijn, en waarbij bovendien elk element op de hoofddiagonaal een deler is van het volgende diagonaalelement. Deze zogenaamde 'Smith-vorm' is een zeer belangrijk theoretisch hulpmiddel in de studie van polynoommatrices.

7. Het vegen van een polynoommatrix

Een goede methode om een stelsel van lineaire algebraïsche vergelijkingen op te lossen is gebaseerd op het 'vegen' van de bijbehorende matrix (ook wel *Gauss-eliminatie* genoemd). Dit levert meestal veel minder werk op dan, bijvoorbeeld, het bepalen van een inverse met behulp van de regel van Cramer. Operaties die in het veegproces worden toegelaten zijn elementaire rijtransformaties (d. w. z. linksvermenigvuldiging met inverteerbare matrices) en verwisseling van kolommen; het laatste komt eenvoudig neer op het hernummeren van de variabelen. Het proces verloopt ongeveer[2] als volgt. Aannemende dat de matrix waarmee we te maken hebben niet geheel uit nullen bestaat (anders wordt de zaak wel heel eenvoudig), kunnen we door verwisseling van rijen en kolommen een element ongelijk nul in de linkerbovenhoek brengen. Door dan de eerste rij telkens met een geschikte konstante te vermenigvuldigen en van de volgende rijen af te trekken kunnen we ervoor zorgen dat er in de eerste kolom onder het element linksboven alleen maar nullen komen te staan. De eerste rij en de eerste kolom zijn nu klaar, en we richten onze aandacht vervolgens op de (2,2)-positie in de matrix. Opnieuw kunnen we, door (zo nodig) rijen en kolommen te verwisselen, ervoor zorgen dat hier een element ongelijk nul komt te staan. (Als tenminste niet inmiddels alle elementen buiten de eerste rij nul zijn; als dat wel zo is, dan stopt het proces.) Daarmee kunnen we dan weer alle elementen op de derde en hogere posities in

1. H. J. S. Smith, "On systems of linear indeterminate equations and congruences", Philos. Trans. Roy. Soc. London, vol. 151 (1861-1862), pp. 293-326 = Coll. Math. Papers, pp. 367-406. Overigens had Smith het hierin eigenlijk over matrices met gehele getallen als elementen; deze situatie is echter volledig vergelijkbaar met die waarin polynomen als elementen optreden. Smith' belangstelling voor het onderwerp kwam voort uit de studie van kwadratische vormen met gehele coëfficienten.
2. Er kunnen nog variaties en verfijningen worden toegevoegd, bijvoorbeeld om het effekt van afrondfouten zo klein mogelijk te houden.

de tweede kolom nul gaan maken. Dan gaan we naar de (3,3)-positie toe, enzovoorts. Het eindresultaat is een matrix in 'bovendriehoeksvorm', en het bijbehorende stelsel kan nu gemakkelijk 'van achter naar voor' worden opgelost.

Precies dezelfde methode kan worden toegepast voor stelsels differentiaalvergelijkingen, mits we antwoord kunnen geven op de volgende vraag: gegeven een kolomvector waarvan de elementen polynomen zijn (niet alle gelijk aan nul), hoe kunnen we een unimodulaire matrix vinden die deze vector omzet in een vector waarvan nog maar één element (het bovenste) ongelijk nul is? Omdat we de elementen 'om beurten' kunnen behandelen is het voldoende als we deze vraag kunnen beantwoorden voor een vector ter lengte 2. Dus: gegeven twee polynomen $p(s)$ en $q(s)$, vind een 2×2 unimodulaire matrix zó dat

$$\begin{bmatrix} a(s) & b(s) \\ c(s) & d(s) \end{bmatrix} \begin{bmatrix} p(s) \\ q(s) \end{bmatrix} = \begin{bmatrix} g(s) \\ 0 \end{bmatrix}.$$

Het wordt als opgave aan de lezer gelaten om te bewijzen dat uit bovenstaande relatie volgt dat $g(s)$ een *grootste gemene deler* van $p(s)$ en $q(s)$ moet zijn. Voor onze doeleinden mogen we aannemen dat $p(s)$ en $q(s)$ geen (niet-triviale) gemeenschappelijke faktoren hebben; waren die er wel, dan zouden we die immers kunnen uitdelen en in het rechterlid zetten. We kunnen dan gebruik maken van de volgende belangrijke eigenschap van polynomen: *als $p(s)$ en $q(s)$ twee polynomen zonder gemeenschappelijke nulpunten zijn, dan bestaan er polynomen $a(s)$ en $b(s)$ zó dat $a(s)p(s) + b(s)q(s) = 1$.* Met deze polynomen is ons probleem meteen opgelost: neem maar $a(s)$ en $b(s)$ zoals aangegeven, en neem $c(s) = -q(s)$ en $d(s) = p(s)$. Dan geldt

$$\begin{bmatrix} a(s) & b(s) \\ c(s) & d(s) \end{bmatrix} \begin{bmatrix} p(s) \\ q(s) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

en de zo gedefinieerde matrix is unimodulair want

$$\det \begin{bmatrix} a(s) & b(s) \\ c(s) & d(s) \end{bmatrix} = a(s)p(s) + b(s)q(s) = 1.$$

We zullen de bovenstaande eigenschap nu bewijzen volgens de strenge eisen van het scheepsrecht.

Bewijs 1. Herinnert u zich de *breuksplitsing* nog? Als u een rationale funktie opgegeven krijgt, dus een breuk waarvan de teller en noemer polynomen zijn, dan kunt u de breuk 'splitsen'. Dat wil zeggen: u schrijft hem als een som van rationale funkties waarbij de noemers van de vorm $(s-\lambda)^k$ of $(s^2+\lambda s+\mu)^k$ zijn, korresponderend met de ontbinding in faktoren van de noemer van de gegeven funktie. Dus, bijvoorbeeld,

$$\frac{1}{s^2-1} = \frac{0.5}{s-1} - \frac{0.5}{s+1}.$$

We voeren nu zo'n breuksplitsing uit op de funktie $1/p(s)q(s)$. In het resultaat krijgen we twee groepen termen: een groep die samenhangt met faktoren uit de ontbinding van $p(s)q(s)$ die afkomstig zijn van $p(s)$, en een groep die op dezelfde manier afstamt van $q(s)$. (Omdat, volgens de aanname, $p(s)$ en $q(s)$ geen gemeenschappelijke faktoren hebben, kunnen we dit onderscheid inderdaad zo maken.) We brengen de eerste groep onder één noemer: die noemer wordt dan precies weer $p(s)$, en de teller noemen we $b(s)$. Op dezelfde manier schrijven we de tweede groep als $a(s)/q(s)$. We hebben dan

$$\frac{1}{p(s)q(s)} = \frac{b(s)}{p(s)} + \frac{a(s)}{q(s)}.$$

Vermenigvuldig nu met $p(s)q(s)$ — en daar is het gewenste resultaat.

Bewijs 2. Herinnert u zich de *polynoominterpolatie* nog? Er bestaan formules die een polynoom opleveren dat op gegeven punten in het complexe vlak gegeven waarden aanneemt (eventueel mogen op deze punten zelfs ook nog de waarden van de eerste tot en met k-de afgeleide worden voorgeschreven). We willen nu polynomen $a(s)$ en $b(s)$ zoeken die zodanig zijn dat $a(s)p(s) + b(s)q(s) = 1$, anders gezegd: we zoeken een polynoom $a(s)$ zó dat $1 - a(s)p(s)$ een veelvoud is van $q(s)$. Dat wil zeggen dat $1 - a(s)p(s)$ nulpunten moet hebben overal waar $q(s)$ ze heeft, en met tenminste dezelfde multipliciteit. Dus als λ een nulpunt is van $q(s)$ met multipliciteit k, dan moet gelden

$$1 - a(\lambda)p(\lambda) = 0,$$

en als $k \geqslant 2$ ook

$$a'(\lambda)p(\lambda) + a(\lambda)p'(\lambda) = 0,$$

etcetera. De aanname was dat $p(s)$ en $q(s)$ geen nulpunten gemeenschappelijk hebben, dus dat betekent in het bijzonder dat $p(\lambda)$ niet nul is. Dus de eis voor $a(s)$ wordt: $a(\lambda) = 1/p(\lambda)$, plus eisen voor de hogere afgeleiden van $a(s)$ in λ, die op soortgelijke wijze kunnen worden afgeleid. Alle nulpunten van $q(s)$ langslopend komen we zo aan een serie interpolatie-eisen voor $a(s)$. We weten dat hieraan kan worden voldaan; dus daarmee is het probleem ook opgelost.

Bewijs 3. Herinnert u zich de stelling uit de lineaire algebra nog, die zegt: *de dimensie van het beeld van een lineaire afbeelding is gelijk aan de dimensie van de ruimte vanwaar hij vertrekt, minus de dimensie van zijn kern*? Laten we zeggen dat $p(s)$ de graad n heeft, en $q(s)$ de graad m. De verzameling van alle polynomen met graad ten hoogste gelijk aan k is een lineaire ruimte met dimensie $k+1$. We kijken nu naar de verzameling van paren $(a(s), b(s))$ waarbij $a(s)$ een polynoom is van graad ten hoogste m, en $b(s)$ een polynoom van graad ten hoogste n. Die verzameling van paren is dus een lineaire ruimte van dimensie $n+m+2$. De afbeelding die aan het paar $(a(s), b(s))$ het polynoom $a(s)p(s) + b(s)q(s)$ toevoegt is een lineaire afbeelding, die gaat van een ruimte van dimensie $n+m+2$ naar een ruimte van dimensie $n+m+1$ (de verzameling van polynomen van graad ten hoogste $n+m$). Wat is de kern van

deze afbeelding, d.w.z. welke paren worden op 0 afgebeeld? Uit $a(s)p(s) + b(s)q(s) = 0$ volgt

$$\frac{a(s)}{b(s)} = -\frac{q(s)}{p(s)}.$$

We hadden aangenomen dat $p(s)$ en $q(s)$ geen faktoren gemeenschappelijk hebben, zodat de breuk $q(s)/p(s)$ niet vereenvoudigd kan worden; bovendien is de graad van $a(s)$ niet hoger dan de graad van $q(s)$ (nl. m). Dus uit de bovenstaande formule volgt dat moet gelden $a(s) = -\lambda q(s)$ en $b(s) = \lambda p(s)$ voor een zeker getal λ, oftewel $(a(s), b(s)) = \lambda(-q(s), p(s))$. Met andere woorden: de dimensie van de kern van onze afbeelding is 1. De bovengenoemde stelling uit de lineaire algebra laat dan zien dat deze afbeelding *surjectief* is; in het bijzonder wordt ook het polynoom 1 als waarde aangenomen. Er bestaat dus een paar $(a(s), b(s))$ zó dat $a(s)p(s) + b(s)q(s) = 1$, waarmee onze bewering ten derden male is bewezen.[1]

Het 'vegen' van een polynoommatrix leidt tot een stelsel waarvan de matrix in 'bovendriehoeksvorm' staat; dat wil zeggen, beneden de hoofddiagonaal staan alleen maar nullen. Als het aantal kolommen (n) groter is dan het aantal rijen (k), dan zijn voor de laatste $n-k$ variabelen willekeurige funkties in te vullen. Verder kan dan de variabele $x_k(t)$ worden gevonden als oplossing van een gewone differentiaalvergelijking met gegeven rechterlid; deze oplossing zal in het algemeen niet uniek zijn — er worden willekeurig te kiezen konstantes geïntroduceerd in een aantal dat gelijk is aan de graad van het polynoom op de (k,k)-de plaats in de matrix. Vervolgens wordt de variabele $x_{k-1}(t)$ gevonden, opnieuw als oplossing van een differentiaalvergelijking met gegeven rechterlid, met introduktie van een aantal willekeurige konstantes dat gelijk is aan de graad van het polynoom op de $(k-1, k-1)$-de plaats; etcetera. Uiteindelijk zal de gehele verzameling van oplossingen geschreven kunnen worden als een uitdrukking waarin in het algemeen zowel een aantal willekeurige funkties als een aantal willekeurige konstanten voorkomen.

In het bijzonder kunnen we de situatie bestuderen waarin we n onafhankelijke vergelijkingen in n onbekenden hebben. Er zijn dan geen willekeurige funkties te kiezen in de oplossing, alleen willekeurige konstanten. Als we de geveegde matrix aangeven met $Q(s)$, dan is volgens de boven gemaakte opmerkingen het aantal geïntroduceerde konstantes (dus het aantal onafhankelijke oplossingen) gelijk aan de som van de graden van de diagonaalelementen van $Q(s)$. Omdat $Q(s)$ een bovendriehoeksmatrix is is zijn determinant gelijk aan het produkt van de diagonaalelementen; dus de graad van $\det Q(s)$ is gelijk aan de som van de graden van de diagonaalelementen.

1. Om de polynomen $a(s)$ en $b(s)$ in deze opzet expliciet uit te rekenen, ligt het het meest voor hand *bases* te kiezen in de relevante lineaire ruimtes, en het geheel als een matrixprobleem te schrijven. Zo zouden we komen op het gebruik van *konstante* matrices om beweringen over *polynomen* te doen; maar dat is weer een ander chapiter.

Het aantal onafhankelijke oplossingen is derhalve gelijk aan de graad van $\det Q(s)$. Nu is de matrix $Q(s)$ ontstaan uit de matrix van het oorspronkelijke stelsel, die we $P(s)$ zullen noemen, door het toepassen van veegoperaties; dus er is een unimodulaire matrix $U(s)$ zó dat $Q(s) = U(s)P(s)$. Daaruit volgt dat

$$\det Q(s) = \det U(s) \cdot \det P(s).$$

We weten echter dat de determinant van een unimodulaire matrix een konstante ongelijk nul is; hieruit blijkt dat de graad van de determinant van $P(s)$ gelijk is aan de graad van de determinant van $Q(s)$, en dus ook gelijk is aan het aantal onafhankelijke oplossingen. Daarmee is de aan het eind van §5 gedane bewering bewezen.

8. Open systemen: stelsels met meer onbekenden dan vergelijkingen

Een stelsel waarvan het aantal onafhankelijke vergelijkingen gelijk is aan het aantal onbekenden valt op te vatten als een 'gesloten' systeem. De algemene oplossing kan worden geschreven als een lineaire kombinatie van een eindig aantal funkties, en er kan een unieke oplossing worden verkregen door een geschikte beginvoorwaarde op te leggen. Eenmaal in gang gezet beweegt het systeem zich autonoom voort, zonder onderhevig te zijn aan invloeden van buiten.

Anders is het, wanneer het aantal onbekenden groter is dan het aantal onafhankelijke vergelijkingen. De betekenis van een dergelijk stelsel is eigenlijk dat het een (dynamische) *relatie* legt tussen diverse grootheden. Als eenvoudig voorbeeld kan de wet van Newton genoemd worden: $F(t) = mx''(t)$. Dit is één vergelijking in twee onbekenden; hij beschrijft de relatie tussen de kracht die op een deeltje werkt en de positie van het deeltje, beide als funktie van de tijd. Een ander voorbeeld wordt geleverd door de ARMA-modellen (in diskrete tijd), die in de econometrie populair zijn:

$$y(t) + a_1 y(t-1) + \cdots + a_k y(t-k) =$$
$$= b_0 e(t) + b_1 e(t-1) + \cdots + b_l e(t-l).$$

Aan de variabelen $e(t)$ wordt vaak opgelegd dat het onderling onafhankelijke identiek verdeelde stochastische grootheden zijn;[1] het model dient dan als verklaring van het gedrag van de variabele $y(t)$. Een derde voorbeeld komt van het begrip '*n-port*' uit de theorie van de elektrische netwerken. Een n-port is een elektrisch netwerk waarbij tussen n stellen aansluitpunten spanningen en stromen kunnen worden voorgeschreven of gemeten. Ook hier zijn de vergelijkingen kleiner in aantal dan de onbekenden; ze dienen dan ook niet om een of andere unieke oplossing te bepalen, maar om de relaties tussen de stromen en spanningen aan de 'ports' te specificeren.

Bij het beschrijven van dynamische relaties tussen variabelen doen zich vele interessante diskussiepunten voor. Eén daarvan betreft het begrip *causaliteit*.

1. Zodat we dus eigenlijk $\underline{e}(t)$ zouden moeten schrijven, en ook $\underline{y}(t)$.

Neem bijvoorbeeld de eenvoudigste situatie waarin er sprake is van één vergelijking in twee onbekenden. Welke van de twee variabelen zullen we aanduiden als 'oorzaak', en welke als 'gevolg'? Als we in diskrete tijd werken is het betrekkelijk gemakkelijk om een kriterium aan te geven. In de vergelijking $x(t+1) = y(t)$ moet y als oorzaak en x als gevolg worden gezien, als we tenminste willen vasthouden aan de gedachte dat de oorzaak aan het gevolg voorafgaat. Meer in het algemeen zullen we in de vergelijking

$$a_k x(t+k) + a_{k-1} x(t+k-1) + \cdots + a_0 x(t) =$$
$$= b_l y(t+l) + b_{l-1} y(t+l-1) + \cdots + b_0 y(t)$$

de variabele x als oorzaak willen zien als k kleiner is dan l, terwijl y als oorzaak zal worden gezien als l kleiner is dan k. Is l gelijk aan k, dan lijkt er geen reden te zijn om de richting van de causaliteit vast te leggen. (Merk overigens op dat in het hierboven genoemde ARMA-model de keuze van parameters zodanig is, dat de positie van e als 'oorzaak' en y als 'gevolg' automatisch is verzekerd.)

Voor systemen in continue tijd is misschien meer diskussie mogelijk, maar algemeen wordt toch de opvatting aangehangen dat de richting van causaliteit dient samen te vallen met die van integratie. Dus in de vergelijking $x'(t) = y(t)$ vatten we x op als gevolg, en y als oorzaak, of — als we het gebruik van zwaarbeladen termen willen vermijden — we noemen y de *ingang* en x de *uitgang*. In het algemeen, als we een differentiaalvergelijking in twee onbekenden hebben die in polynoomnotatie kan worden geschreven als $p(s)x - q(s)y = 0$, dan noemen we x de uitgang als de graad van $p(s)$ hoger is dan die van $q(s)$, terwijl we y de uitgang noemen als het omgekeerde het geval is; zijn de twee graden aan elkaar gelijk, dan laten we beide mogelijkheden open. Het ligt voor de hand om met de vergelijking $p(s)x - q(s)y = 0$ de rationale funktie $p(s)/q(s)$ te associeren, die de variabele y 'uitdrukt' in termen van de variabele x. [1] Deze funktie wordt de *overdrachtsfunktie* genoemd; de overdrachtsfunktie heet *causaal* als de graad van de teller ten hoogste gelijk is aan de graad van de noemer, dus als hij de overgang uitdrukt *van* de ingang *naar* de uitgang. Een rationale funktie waarvan de graad van de teller kleiner is dan of gelijk is aan de graad van de noemer wordt ook wel een *eigenlijk rationale* funktie genoemd.

Dan nu het algemene geval, met meerdere vergelijkingen en meerdere onbekenden. Laten we kijken naar een stelsel van k onafhankelijke vergelijkingen in m onbekenden, waarbij m groter is dan k:

$$P(s)w = 0.$$

Het volgende kan nu worden bewezen. De m onbekenden die zijn samengevat

1. Onder zekere (beperkende) voorwaarden kan hieraan een konkrete betekenis worden gegeven met behulp van de *Laplace-transformatie*. Voor onze doeleinden is het voldoende de breuk $p(s)/q(s)$ te beschouwen als een formele uitdrukking waarin dus niet bijvoorbeeld onder en boven de deelstreep faktoren tegen elkaar mogen worden weggestreept.

in de vector w kunnen worden verdeeld in twee groepen: een groep van k variabelen die we opnemen in een vector y, en een groep van $m-k$ variabelen die we schrijven als een vector u. Deze indeling kan op zo'n manier gebeuren dat de vergelijking $P(s)w = 0$ kan worden herschreven tot

$$Q(s)y + R(s)u = 0$$

waarbij de determinant van de $k \times k$-matrix $Q(s)$ niet nul is zodat deze matrix een *rationale* (niet noodzakelijk polynomiale!) inverse $Q^{-1}(s)$ heeft, met bovendien de eigenschap dat de elementen van de *overdrachtsmatrix* $Q^{-1}(s)P(s)$ alle eigenlijk rationale funkties zijn. Met andere woorden: ook in het geval dat de relaties tussen een aantal variabelen worden gespecificeerd door een *stelsel* differentiaalvergelijkingen kunnen we de variabelen onderscheiden in twee groepen op zo'n manier dat de overdrachtsmatrix van de ene groep naar de andere een matrix is van eigenlijk rationale funkties, zodat we met recht de ene groep kunnen beschrijven als 'ingangen' en de andere als 'uitgangen'. Nog korter gezegd: in ieder lineair stelsel van differentiaalvergelijkingen valt een *causaliteitsstruktuur* aan te geven. Overigens hoeft die struktuur niet uniek bepaald te zijn, zoals we al zagen in de situatie met één vergelijking in twee onbekenden.

Een andere belangrijke beschrijvingsvorm voor stelsels met meer variabelen dan vergelijkingen kan worden gegeven door 'interne variabelen' te introduceren: die zijn op te vatten als nieuwe variabelen die worden gebruikt om de vergelijkingen op een andere manier te kunnen schrijven. In het bijzonder kan worden bewezen dat er voor stelsels van de vorm $Q(s)y + R(s)u = 0$, waarbij $\det Q(s)$ ongelijk nul is en $Q^{-1}(s)R(s)$ een matrix van eigenlijk rationale funkties is, er altijd een variabele $x(t)$ kan worden geïntroduceerd op zo'n manier dat het stelsel kan worden herschreven tot de vorm

$$x'(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t).$$

waarbij A, B, C en D *konstante* matrices zijn.[1] Deze representatie heet de *toestandsrepresentatie*; de variabele $x(t)$ heet de *toestand* van het systeem. De toestand representeert het 'geheugen' van het systeem: het gedrag van de uitgang vanaf een zeker tijdstip t_0 wordt volledig vastgelegd door de waarde van de toestand op het tijdstip t_0 plus het gedrag van de ingang vanaf t_0.

In een stochastische context kan het begrip 'toestand' nog op een andere manier worden geïnterpreteerd. We roepen nog even het ARMA-model uit de econometrie in herinnering:

$$\underline{y}(t) + a_1\underline{y}(t-1) + \cdots + a_k\underline{y}(t-k) =$$

1. 'Herschreven' wordt hier bedoeld in de volgende zin: voor elk paar van funkties $(u(t), y(t))$ dat voldoet aan het stelsel $Q(s)y + P(s)u = 0$ bestaat er een funktie $x(t)$ zodanig dat het drietal $(u(t), x(t), y(t))$ voldoet aan het stel vergelijkingen boven, en omgekeerd, als een drietal $(u(t), x(t), y(t))$ voldoet aan deze vergelijkingen, dan voldoet het paar $(u(t), y(t))$ aan het stelsel waar we mee begonnen.

$$= b_0\underline{e}(t) + b_1\underline{e}(t-1) + \cdots + b_l\underline{e}(t-l),$$

waarbij de variabelen $\underline{e}(t)$ voor verschillende tijdstippen identiek verdeelde en onafhankelijke stochastische grootheden zijn. Het is al eerder opgemerkt dat in dit model de variabele $\underline{e}(t)$ altijd als de ingang kan worden gezien, en we kunnen dus een toestandsrepresentatie opschrijven in de vorm

$$\underline{x}(t+1) = A\underline{x}(t) + B\underline{e}(t)$$
$$\underline{y}(t) = C\underline{x}(t) + D\underline{e}(t).$$

Uit de eerste vergelijking valt af te leiden dat het proces $\underline{x}(t)$ een *Markovproces*[1] is. Het proces $\underline{y}(t)$ dat door het ARMA-model wordt beschreven heeft in het algemeen niet de Markoveigenschap. We kunnen de toestandrepresentatie dus zien als het opdelven van een 'onderliggend Markovproces'.

Over de betekenis van interne variabelen kan men verschillend denken. Bij de beschrijving van een n-port komen ze op natuurlijke wijze naar voren, als men de vergelijkingen opstelt voor de componenten van het netwerk; er komt een niet onaanzienlijke inspanning aan te pas om de vergelijkingen te herschrijven op zo'n manier dat ze alleen nog maar in termen van de portvariabelen zijn gesteld. Aan de andere kant: een econometrist die werkt met een ARMA-model waarvan de parameters door schatting uit waargenomen tijdreeksen zijn bepaald zal een toestandsvariabele opvatten als een gekonstrueerd objekt dat niet in een direkte relatie tot de werkelijkheid staat. Deze variatie in gezichtspunten heeft ertoe geleid dat er in de literatuur verschillende kriteria zijn geformuleerd om te beoordelen wanneer een stelsel mag worden opgevat als een herschrijving van een ander stelsel.[2] De wiskundige wrijft zich hierbij natuurlijk in de handen: aan hem (of haar) de schone taak om precies duidelijk te maken wat de relaties zijn tussen deze kriteria, en hoe, met gebruikmaking van welk stel herschrijfregels men ook verkiest, een gegeven stelsel kan worden getransformeerd naar geschikte standaardvormen.

9. Literatuur

Besprekingen van stelsels gewone lineaire differentiaalvergelijkingen met konstante coëfficienten vindt men vooral in oudere tekstboeken, zoals

E. L. Ince, *Ordinary Differential Equations*, Dover, New York, 1956 (herdruk van het origineel uit 1926)

R. A. Frazer, W. J. Duncan, A. R. Collar, *Elementary Matrices*, Cambridge University Press, 1938.

Het laatstgenoemde boek gaat uitvoerig in op polynoommatrices, evenals

1. Zie de bijdrage van Boxma in deze bundel. Merk evenwel op, dat de toestandsruimte (de ruimte waarin de vector $\underline{x}(t)$ zich bevindt) hier niet eindig is.
2. Het hier weergegeven standpunt is ontleend aan: J. C. Willems, "System theoretic models for the analysis of physical systems", *Ricerche di Automatica* 10 (1979), pp. 71-106.

P. Lancaster, *Lambda-matrices and Vibrating Systems*, Pergamon Press, Oxford, 1966.

Hoofdstuk VI in

F. R. Gantmacher, *Matrix Theory*, Chelsea, New York, 1959 (vertaling van het Russische origineel uit 1954)

is eveneens gewijd aan polynoommatrices. Meer in het algemeen kan men kijken naar matrices waarvan de elementen uit een ring komen. Boeken op dit gebied zijn:

C. C. MacDuffee, *The Theory of Matrices*, Chelsea, New York, 1956 (herdruk van het origineel uit 1933)

M. Newman, *Integral Matrices*, Academic Press, New York, 1972.

Een polynoommatrix kan ook worden geschreven als een matrixpolynoom, d. w. z. als een polynoom waarvan de coëfficienten (konstante) matrices zijn. Dit gezichtspunt wordt uitvoerig belicht in

I. Gohberg, P. Lancaster, L. Rodman, *Matrix Polynomials*, Academic Press, New York, 1982.

Het onderdeel van de toegepaste wiskunde dat zich in het bijzonder bezighoudt met stelsels waarin meer variabelen dan vergelijkingen voorkomen wordt *systeemtheorie* genoemd. Boeken op dit terrein waarin intensief wordt gewerkt met polynoommatrices zijn onder meer:

H. H. Rosenbrock, *State Space and Multivariable Theory*, Wiley, New York, 1970

W. A. Wolovich, *Linear Multivariable Systems*, Springer, New York, 1974

V. Kučera, *Discrete Linear Control: The Polynomial Equation Approach*, Wiley, Londen, 1979

T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, N. J., 1980

F. M. Callier, C. A. Desoer, *Multivariable Feedback Systems*, Springer, New York, 1982

H. Blomberg, R. Ylinen, *Algebraic Theory for Multivariable Linear Systems*, Academic Press, Londen, 1983.

MATRICES EN MARKOVPROCESSEN

O.J. Boxma
Centrum voor Wiskunde en Informatica, Amsterdam

1. Inleiding 81

2. Markovketens 83

3. Matrices van overgangskansen in Markovketens 87

4. Matrices en toestandskansen in Markovketens 90

5. Aanvullingen en verwijzingen 101

Literatuur 102

Matrices en Markovprocessen

1. Inleiding

Een wezenlijk kenmerk van veel verschijnselen in het dagelijks leven is *onzekerheid.* In de kansrekening zijn verscheidene modellen ontwikkeld om zulke, deels door het toeval bepaalde, verschijnselen te analyseren. Het model van A.A. Markov neemt hierbij een centrale plaats in. Waar zijn voorgangers onzekerheid als een rij *onafhankelijke* "experimenten" hadden gemodelleerd, introduceerde Markov een eenvoudige vorm van *afhankelijkheid:* elk volgend experiment is (uitsluitend) afhankelijk van de uitkomst van het zojuist voltooide experiment. Stochastische processen met een dergelijke afhankelijkheidsstructuur worden *Markovprocessen* genoemd. Markovprocessen geven enerzijds de essentie van veel verschijnselen uit het dagelijks leven redelijk natuurgetrouw weer, en hebben anderzijds een zo eenvoudige structuur dat een uitvoerige wiskundige analyse mogelijk is.

We geven eerst een voorbeeld van een Markovproces, alvorens tot een meer formele behandeling over te gaan.

Voorbeeld 1

Stel dat voor een wiskundevak twee goede leerboeken beschikbaar zijn, A en B. Een bepaalde leraar kiest elk jaar één van deze twee boeken. Hij kiest boek A met kans 0.8 als hij vorig jaar ook al boek A had gebruikt, en hij kiest boek A met kans 0.3 als hij vorig jaar boek B had gebruikt. We kunnen dit procédé als een Markovproces met twee toestanden en discrete tijdsparameter modelleren:

Beschouw de jaren 0,1,2,... als discrete tijdstippen (in jaar 0 wordt het vak voor het eerst gegeven), en onderscheid voor elk jaar twee toestanden A en B, corresponderend met gebruik van boek A respectievelijk B in dat jaar. In de

hierboven gebruikte terminologie is de keuze van een boek een experiment. De uitkomst van een experiment hangt uitsluitend af van de uitkomst van het vorige experiment, dus van de keuze van het boek in het voorgaande jaar.

Stel dat in jaar 0 boek A is gekozen. Dan wordt volgens het gegeven in jaar 1 met kans 0.8 boek A gekozen. Wat is de kans dat boek A in jaar 2 wordt gekozen? En in jaar 10? En in een oneindig ver in de toekomst gelegen jaar? Maakt het voor deze laatste vraag verschil als in jaar 0 met boek B was begonnen? Als men mag veronderstellen dat veel wiskundeleraren elk jaar voor deze zelfde keuze staan, en hun keuze volgens hetzelfde procédé bepalen, dan zijn bovenstaande vragen interessant voor een uitgever.

Iets formeler zijn we geinteresseerd in de volgende vragen. Stel dat in jaar 0 met kans p_0 boek A wordt gekozen, en dat in jaar $m+1$ met kans α respectievelijk β boek A wordt gekozen als in jaar m boek A respectievelijk B is gebruikt. Voor welke waarden van de kansen p_0, α en β bestaat een limietkansverdeling voor de kans op keuze van boek A in jaar m, $m\to\infty$? En zo een limietverdeling bestaat, wat is deze limietverdeling, en hoe snel wordt hij benaderd?

De theorie van de Markovprocessen leert, hoe de vragen uit voorbeeld 1 op snelle en elegante wijze kunnen worden beantwoord. Het maakt daarbij in principe niet uit of het aantal toestanden 2 is, zoals in het voorbeeld, of een willekeurig geheel getal $k\geqslant 2$.

De theorie van de Markovprocessen maakt sterk gebruik van matrixrekening. De matrix van *overgangskansen* $P = (p_{ij})$, met p_{ij} de overgangskans van toestand i in jaar m naar toestand j in jaar $m+1$, speelt een centrale rol. In het voorbeeld is (identificeer toestanden A en 1, en toestanden B en 2):

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}.$$

We zullen tonen hoe, zowel in het geval van voorbeeld 1 als bij ingewikkelder problemen, vragen over de kansverdeling na m stappen, over de limietkansverdeling voor $m\to\infty$, en over de snelheid van convergentie kunnen worden beantwoord door analyse van de overgangsmatrix P, en in het bijzonder van de eigenwaarden en eigenvectoren van P.

De opzet van dit verhaal is als volgt. Paragraaf 2 geeft een inleiding in de theorie der Markovprocessen met discrete tijdsparameter, de zg. *Markovketens.* Paragraaf 3 is gewijd aan overgangskansen in Markovketens, terwijl paragraaf 4 toestandskansen (kansverdelingen waarin de begintoestand geen invloed heeft) beschouwt; in beide paragrafen speelt de overgangsmatrix P een grote rol. In paragraaf 5 worden enige aspecten van Markovprocessen genoemd die tot dan niet aan de orde zijn gekomen; deze paragraaf bevat ook verscheidene literatuurverwijzingen voor verdere studie. De behandelde theorie wordt toegelicht aan de hand van eenvoudige voorbeelden, waarbij het voorbeeld van de twee wiskundeboeken als een rode draad door het verhaal loopt.

2. MARKOVKETENS

We veronderstellen dat de lezer bekend is met het begrip *stochastische variabele [3].* In dit verhaal zullen stochastische variabelen worden onderstreept.

Zij t een parameter die waarden aanneemt in een verzameling T (bv. $\{0,1,2,...\}$ of $[0,\infty)$); t zal steeds een tijdsparameter zijn. Zij $\underline{x}_t$ een stochastische variabele voor elke $t \in T$. We verkrijgen zodoende een familie $\{\underline{x}_t,\ t \in T\}$ van stochastische variabelen. Zo'n familie heet een *stochastisch proces.* De elementen van T heten tijdspunten, en T is een lineaire verzameling, aftelbaar of niet-aftelbaar. Als $T = \{t: t \in [0,\infty)\}$, dan heet het stochastische proces $\{\underline{x}_t,\ t \in T\}$ een proces met *continue tijdsparameter;* als $T = \{t: t = t_n,\ n = 0,1,2,...\}$ met $t_0 \leqslant t_1 \leqslant t_2 \leqslant ...$, dan is sprake van een proces met *discrete tijdsparameter.*

Men kan een stochastisch proces beschouwen als de probabilistische tegenhanger van een deterministisch proces, en net als bij deterministische processen is men geinteresseerd in haar tijdsafhankelijke structuur en haar gedrag voor grote waarden van de tijdsparameter. Vanzelfsprekend zijn stochastische processen veel gecompliceerder dan deterministische processen, en zal men dan ook in het algemeen minder gedetailleerde uitspraken kunnen doen.

Men kan een deterministisch proces ook beschouwen als een gedegenereerd stochastisch proces, waarin het "random" karakter geheel ontbreekt. Anderzijds is een stochastisch proces met een volledig "random" karakter, een proces $\{\underline{x}_t,\ t \in [0,\infty)\}$ waarbij $\underline{x}_{t_1}, \cdots, \underline{x}_{t_n}$ onafhankelijke stochastische variabelen zijn voor alle n en elke deelverzameling $\{t_1, \ldots, t_n\}$ van T, nauwelijks interessant (tenzij T een discrete verzameling is); zo'n proces is te onregelmatig en bevat geen enkel verband tussen verleden, heden en toekomst. Enig verband tussen deze grootheden in een stochastisch proces kan het proces veel bruikbaarder maken als model voor toepassingen.

Markovketens zijn stochastische processen waarbij de toekomst, *gegeven het heden,* niet afhangt van het verleden. De Russische wiskundige A.A. Markov (1856-1922) voerde de naar hem genoemde Markovketens in bij zijn onderzoek naar de afwisseling van klinkers en medeklinkers in de Russische literatuur, i.h.b. in Pushkin's "Eugene Onegin" (Markov's artikel [6] is opgenomen in een appendix van Howard [4]). Markovketens vormen reeds geruime tijd een belangrijk hoofdstuk in de stochastiek, niet in het minst doordat zeer veel toepassingen worden gevonden in zulke uiteenlopende gebieden als natuurkunde, sociologie, biologie en economie.

We zullen nu een formele definitie van Markovketens geven. Zij in het vervolg $T = \{0,1,2,...\}$. Zij S de toestandsruimte van het stochastische proces $\{\underline{x}_n,\ n = 0,1,...\}$, dwz. de ruimte van mogelijke realisaties van de stochastische variabelen $\underline{x}_n$, $n = 0,1,...$; S zij aftelbaar. S kan bijvoorbeeld de verzameling positieve gehele getallen zijn, of de verzameling $\{A,B\}$ van twee boeken. In het vervolg geeft Pr een kans (Eng.: Probability) aan.

DEFINITIE 1
Het stochastische proces $\{\underline{x}_n,\ n=0,1,...\}$ is een Markovketen als voor elk geheel getal $n \geqslant 0$ en elke $i_k \in S,\ k=0,1,...,n+1$,

$$Pr\{\underline{x}_{n+1}=i_{n+1}|\underline{x}_n=i_n,\underline{x}_{n-1}=i_{n-1},\ldots,\underline{x}_0=i_0\} = \tag{2.1}$$
$$Pr\{\underline{x}_{n+1}=i_{n+1}|\underline{x}_n=i_n\}.$$

(Uiteraard moet $Pr\{\underline{x}_n=i_n,\underline{x}_{n-1}=i_{n-1},\ldots,\ \underline{x}_0=i_0\} > 0$ zijn, opdat het linkerlid gedefinieerd is; en dan is ook het rechterlid gedefinieerd.)

In woorden: als in de rij $\{\underline{x}_n,\ n=0,1,...\}$ de realisatie van $\underline{x}_n$ op tijdstip n bekend is, dan hangt de conditionele kans van de toekomstige realisatie van $\underline{x}_{n+1}$ slechts af van het heden $\underline{x}_n$, en niet van het verleden $\underline{x}_{n-1},\ldots,\underline{x}_0$ (verder gaat het geheugen van de keten niet).

VOORBEELD 1 (VERVOLG)
In het wiskundeboekenvoorbeeld uit paragraaf 1 geldt: $\underline{x}_n$ is het boek dat in jaar n in gebruik is; de toestandsruimte $S = \{A,B\}$, terwijl $T = \{0,1,...,n,...\}$. Er geldt voor elk geheel getal $n \geqslant 0$ en alle $i_k \in S,\ k=0,1,...,n-1$:

$$Pr\{\underline{x}_{n+1}=A|\underline{x}_n=A,\underline{x}_{n-1}=i_{n-1},\ldots,\underline{x}_0=i_0\} =$$
$$Pr\{\underline{x}_{n+1}=A|\underline{x}_n=A\} = 0.8,$$

en analoog

$$Pr\{\underline{x}_{n+1}=A|\underline{x}_n=B,\underline{x}_{n-1}=i_{n-1},\ldots,\underline{x}_0=i_0\} = 0.3,$$
$$Pr\{\underline{x}_{n+1}=B|\underline{x}_n=A,\underline{x}_{n-1}=i_{n-1},\ldots,\underline{x}_0=i_0\} = 0.2,$$
$$Pr\{\underline{x}_{n+1}=B|\underline{x}_n=B,\underline{x}_{n-1}=i_{n-1},\ldots,\underline{x}_0=i_0\} = 0.7.$$

Enige eigenschappen van Markovketens (zie Cohen [2], hoofdstuk I.2)
Definitie 1 impliceert (gebruik elementaire eigenschappen van conditionele kansen):

$$Pr\{\underline{x}_{n+2}=i_{n+2},\ \underline{x}_{n+1}=i_{n+1}|\underline{x}_k=i_k,\ k=0,1,...,n\}Pr\{\underline{x}_k=i_k,\ k=0,1,...,n\} \tag{2.2}$$
$$Pr\{\underline{x}_k=i_k,\ k=0,1,...,n+2\} =$$
$$Pr\{\underline{x}_{n+2}=i_{n+2}|\underline{x}_k=i_k,\ k=0,1,...,n+1\}Pr\{\underline{x}_k=i_k,\ k=0,1,...,n+1\} =$$
$$Pr\{\underline{x}_{n+2}=i_{n+2}|\underline{x}_{n+1}=i_{n+1}\}\ Pr\{\underline{x}_{n+1}=i_{n+1}|\underline{x}_n=i_n\}Pr\{\underline{x}_k=i_k,\ k=0,1,...,n\}.$$

Derhalve geldt (als de betreffende conditionele kansen bestaan, wat in het vervolg zal worden aangenomen):

$$Pr\{\underline{x}_{n+2}=i_{n+2},\ \underline{x}_{n+1}=i_{n+1}|\underline{x}_k=i_k,\ k=0,1,...,n\} = \tag{2.3}$$
$$Pr\{\underline{x}_{n+2}=i_{n+2}|\underline{x}_{n+1}=i_{n+1}\}\ Pr\{\underline{x}_{n+1}=i_{n+1}|\underline{x}_n=i_n\}.$$

Het rechterlid van (2.3) hangt niet af van $\underline{x}_k = i_k$, $k = 0,1,...,n-1$, zodat

$$Pr\{\underline{x}_{n+2} = i_{n+2}, \underline{x}_{n+1} = i_{n+1} | \underline{x}_k = i_k, k = 0,1,...,n\} = \tag{2.4}$$
$$Pr\{\underline{x}_{n+2} = i_{n+2}, \underline{x}_{n+1} = i_{n+1} | \underline{x}_n = i_n\}.$$

Formules (2.3) en (2.4) laten zich eenvoudig generaliseren tot:

$$Pr\{\underline{x}_h = i_h, h = n+1,...,n+m | \underline{x}_k = i_k, k = 0,1,...,n\} = \tag{2.5}$$
$$Pr\{\underline{x}_h = i_h, h = n+1,...,n+m | \underline{x}_n = i_n\} =$$
$$\prod_{k=1}^{m} Pr\{\underline{x}_{n+k} = i_{n+k} | \underline{x}_{n+k-1} = i_{n+k-1}\},$$

voor alle gehele getallen $n \geqslant 0$, $m \geqslant 1$, en voor alle $i_k \in S$, $k = 0,1,...,n+m$.
Het tweede gelijkteken impliceert (neem $n = 0$ in (2.5)):

$$Pr\{\underline{x}_h = i_h, h = 0,1,...,m\} = \tag{2.6}$$
$$Pr\{\underline{x}_0 = i_0\} Pr\{\underline{x}_h = i_h, h = 1,2,...,m | \underline{x}_0 = i_0\} =$$
$$Pr\{\underline{x}_0 = i_0\} \prod_{k=1}^{m} Pr\{\underline{x}_k = i_k | \underline{x}_{k-1} = i_{k-1}\}.$$

Dit leidt tot een belangrijke conclusie:
Als in een Markovketen $\{\underline{x}_n, n = 0,1,...\}$ de *beginverdeling* (de kansverdeling van $\underline{x}_0$) en alle *éénstaps overgangskansen*

$$Pr\{\underline{x}_{n+1} = i_{n+1} | \underline{x}_n = i_n\}$$

bekend zijn, dan kan elke eindig-dimensionale gezamenlijke kansverdeling van de Markovketen eenduidig worden bepaald (en i.h.b. de kansverdeling van elk der stochastische variabelen $\underline{x}_n$, $n = 1,2,...$).
Met iets andere woorden: voor de beschrijving van een Markovketen is het voldoende de beginverdeling en alle éénstaps overgangskansen te geven.

VRAAG 1
Mag in (2.5) een deel van het verleden $(\underline{x}_{n-1}, \ldots, \underline{x}_0)$ worden weggelaten?

ANTWOORD:
Ja, dit volgt onmiddellijk uit het feit dat het tweede en derde lid van (2.5) geen termen $\underline{x}_{n-1}, \ldots, \underline{x}_0$ bevatten.

In de eerste identiteit van (2.5) mag ook een deel van de toekomst worden weggelaten. Er geldt in het bijzonder voor alle gehele getallen $n \geqslant 0$, $m \geqslant 1$,

$$Pr\{\underline{x}_{n+m} = i_{n+m} | \underline{x}_k = i_k, k = 0,1,...,n\} = Pr\{\underline{x}_{n+m} = i_{n+m} | \underline{x}_n = i_n\}. \tag{2.7}$$

We bewijzen (2.7) voor $m = 2$; generalisaties kunnen volledig analoog worden bewezen. Omdat

$$\bigcup_{i \in S} \{\underline{x}_{n+1} = i\} = \Omega,$$

waarbij Ω de zekere gebeurtenis is, kunnen we schrijven:

$$Pr\{\underline{x}_{n+2}=i_{n+2}|\underline{x}_k=i_k,\ k=0,1,...,n\} = \tag{2.8}$$

$$Pr\{\underline{x}_{n+2}=i_{n+2},\ \bigcup_{i\in S}\underline{x}_{n+1}=i|\underline{x}_k=i_k,\ k=0,1,...,n\} =$$

$$\sum_{i\in S} Pr\{\underline{x}_{n+2}=i_{n+2},\ \underline{x}_{n+1}=i|\underline{x}_k=i_k,\ k=0,1,...,n\} =$$

$$\sum_{i\in S} Pr\{\underline{x}_{n+2}=i_{n+2},\ \underline{x}_{n+1}=i|\underline{x}_n=i_n\} =$$

$$Pr\{x_{n+2}=i_{n+2},\ \bigcup_{i\in S}\underline{x}_{n+1}=i|\underline{x}_n=i_n\} = Pr\{\underline{x}_{n+2}=i_{n+2}|\underline{x}_n=i_n\}.$$

De conditionele kans

$$Pr\{\underline{x}_{n+2}=i_{n+2}|\underline{x}_n=i_n\},$$

heet de *tweestaps overgangskans* van de Markovketen op tijdstip n; algemeen heet de conditionele kans

$$Pr\{\underline{x}_{n+m}=i_{n+m}|\underline{x}_n=i_n\},$$

de *m-staps overgangskans* van de Markovketen op tijdstip n.

We hebben in deze paragraaf het begrip Markovketen gedefinieerd, en we hebben getoond hoe deze definitie leidt tot algemene formules als (2.5), welke ons in staat stellen conditionele kansen uit te drukken in éénstaps overgangskansen van de Markovketen. Tot besluit van deze paragraaf, en als voorbereiding op paragraaf 3, laten we zien hoe (h+m)-staps overgangskansen kunnen worden uitgedrukt in h-staps en m-staps overgangskansen (voor willekeurige gehele getallen $h\geqslant 1$, $m\geqslant 1$). Een elementaire eigenschap van conditionele kansen impliceert dat

$$Pr\{\underline{x}_{n+h+m}=i_{n+h+m},\ \underline{x}_{n+m}=i_{n+m}|\underline{x}_n=i_n\} =$$

$$Pr\{\underline{x}_{n+h+m}=i_{n+h+m}|\underline{x}_{n+m}=i_{n+m},\ \underline{x}_n=i_n\}Pr\{\underline{x}_{n+m}=i_{n+m}|\underline{x}_n=i_n\}.$$

Omdat $\{\underline{x}_n,\ n=0,1,...\}$ een Markovketen is, mogen we schrijven (vgl. vraag 1):

$$Pr\{\underline{x}_{n+h+m}=i_{n+h+m}|\underline{x}_{n+m}=i_{n+m},\ \underline{x}_n=i_n\} =$$

$$Pr\{\underline{x}_{n+h+m}=i_{n+h+m}|\underline{x}_{n+m}=i_{n+m}\}.$$

Uit bovenstaande twee relaties volgt:

$$Pr\{\underline{x}_{n+h+m}=i_{n+h+m},\ \underline{x}_{n+m}=i_{n+m}|\underline{x}_n=i_n\} = \tag{2.9}$$

$$Pr\{\underline{x}_{n+h+m}=i_{n+h+m}|\underline{x}_{n+m}=i_{n+m}\}\ Pr\{\underline{x}_{n+m}=i_{n+m}|\underline{x}_n=i_n\}.$$

Sommatie van (2.9) over alle $i_{n+m}\in S$ geeft:

$$Pr\{\underline{x}_{n+h+m}=i_{n+h+m}|\underline{x}_n=i_n\} = \tag{2.10}$$

$$\sum_{i_{n+m}\in S} Pr\{\underline{x}_{n+h+m}=i_{n+h+m}|\underline{x}_{n+m}=i_{n+m}\}Pr\{\underline{x}_{n+m}=i_{n+m}|\underline{x}_n=i_n\},$$

voor alle gehele getallen $n \geqslant 0$, $h \geqslant 1$, $m \geqslant 1$, en alle i_n, i_{n+m}, $i_{n+h+m} \in S$. Formule (2.10) geeft aan hoe meerstaps overgangskansen rechtstreeks kunnen worden berekend uit éénstaps overgangskansen. Deze formule heet de vergelijking van Chapman-Kolmogorov.

VOORBEELD 1 (VERVOLG)
Uit (2.10) volgt:

$$\begin{aligned} Pr\{\underline{x}_2 = A | \underline{x}_0 = A\} &= Pr\{\underline{x}_2 = A | \underline{x}_1 = A\} Pr\{\underline{x}_1 = A | \underline{x}_0 = A\} + \\ &\quad Pr\{\underline{x}_2 = A | \underline{x}_1 = B\} Pr\{\underline{x}_1 = B | \underline{x}_0 = A\} = \\ &\quad 0.8 \times 0.8 + 0.3 \times 0.2 = 0.7. \end{aligned}$$

Een soortgelijke berekening, of het feit dat

$$Pr\{\underline{x}_2 = A | \underline{x}_0 = A\} + Pr\{\underline{x}_2 = B | \underline{x}_0 = A\} = 1,$$

leert dat

$$Pr\{\underline{x}_2 = B | \underline{x}_0 = A\} = 0.3.$$

Evenzo

$$Pr\{\underline{x}_2 = A | \underline{x}_0 = B\} = 0.45, \quad Pr\{\underline{x}_2 = B | \underline{x}_0 = B\} = 0.55.$$

In de volgende paragraaf zullen we tonen hoe de matrixrekening, uitgaande van (2.10), een snelle en elegante bepaling van meerstaps overgangskansen mogelijk maakt.

3. MATRICES VAN OVERGANGSKANSEN IN MARKOVKETENS

De éénstaps overgangskansen $Pr\{\underline{x}_{n+1} = j | \underline{x}_n = i\}$ van een Markovketen $\{\underline{x}_n, n = 0, 1, ...\}$ kunnen in principe afhangen van n. Een dergelijke "tijdsafhankelijkheid" geeft een Markovketen meestal een zo gecompliceerde structuur, dat een exacte mathematische analyse niet meer mogelijk is. We beperken ons in dit verhaal tot Markovketens waarvan de overgangskansen niet van de tijd afhangen. Dit zijn zowel vanuit praktisch als theoretisch oogpunt de belangrijkste Markovketens.

DEFINITIE 2
De Markovketen $\{\underline{x}_n, n = 0, 1, ...\}$ heeft *stationaire* overgangskansen als voor alle $i, j \in S$ en alle gehele getallen $m \geqslant 1, n \geqslant 0$,

$$Pr\{\underline{x}_{n+m} = j | \underline{x}_n = i\}$$

niet afhangt van n.

Een nodige en voldoende voorwaarde opdat de Markovketen $\{\underline{x}_n,\ n=0,1,...\}$ stationaire overgangskansen heeft, is dat alle eenstaps overgangskansen $Pr\{\underline{x}_{n+1}=j|\underline{x}_n=i\}$ niet afhangen van n (nodig: triviaal; voldoende: uit (2.10), met volledige inductie naar h en m, startend met $h=1$, $m=1$).

Bij Markovketens met stationaire overgangskansen kunnen we de m-staps overgangskansen als volgt noteren: voor alle $i,j \in S$, en alle niet-negatieve gehele getallen n en m, is

$$p_{ij}^{(m)} := Pr\{\underline{x}_{n+m}=j|\underline{x}_n=i\}; \tag{3.1}$$

hierbij is

$$p_{ij}^{(0)}=1 \quad als \quad i=j,$$
$$p_{ij}^{(0)}=0 \quad als \quad i\neq j.$$

We voeren de *matrix* van m-staps overgangskansen in:

$$P^{(m)} := (p_{ij}^{(m)}), \quad m=0,1,...; \tag{3.2}$$

ten slotte noteren we

$$p_{ij} := p_{ij}^{(1)}, \quad P := P^{(1)}. \tag{3.3}$$

De kansrekening leert dat de matrices $P^{(m)}$, $m=0,1,...$, de volgende eigenschappen hebben:
(i) alle elementen zijn niet-negatief;
(ii) de rijsommen zijn gelijk aan een.

Een dergelijke vierkante matrix heet een *stochastische matrix.* De matrix van overgangskansen van een Markovketen is derhalve een stochastische matrix.

We keren terug tot formule (2.10), die (h+m)-staps overgangskansen uitdrukt in h-staps en m-staps overgangskansen. Een zeer belangrijke constatering is, dat (2.10) precies de berekening van een element van een matrix bij vermenigvuldiging van twee matrices weergeeft; in matrixvorm zegt (2.10):

$$P^{(h+m)} = P^{(h)}\, P^{(m)}, \quad h,m=0,1,... . \tag{3.4}$$

In het bijzonder merken we op:

$$P^{(2)} = P^{(1)}\, P^{(1)} = P\, P = P^2;$$

en algemener

$$P^{(m)} = P^m. \tag{3.5}$$

We hebben hiermee een fundamentele eigenschap van Markovketens met stationaire overgangskansen afgeleid: *de matrix van m-staps overgangskansen is de m-de macht* (t.a.v. de matrixvermenigvuldiging) *van de matrix van eenstaps overgangskansen.*

VOORBEELD 1 (VERVOLG)
In het wiskundeboekenvoorbeeld is

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}.$$

Herhaalde matrixvermenigvuldiging van P met zichzelf leert:

$$P^{(2)} = P^2 = \begin{bmatrix} 0.7 & 0.3 \\ 0.45 & 0.55 \end{bmatrix},$$

$$P^{(3)} = P^3 = \begin{bmatrix} 0.65 & 0.35 \\ 0.525 & 0.475 \end{bmatrix},$$

$$P^{(4)} = P^4 = \begin{bmatrix} 0.625 & 0.375 \\ 0.5625 & 0.4375 \end{bmatrix},$$

$$P^{(5)} = P^5 = \begin{bmatrix} 0.6125 & 0.3875 \\ 0.58125 & 0.41875 \end{bmatrix},$$

enz.
Merk op dat elke term (kans) in deze matrices een interpretatie heeft als som van kansen, gesommeerd over alle mogelijkheden die leiden tot het gewenste resultaat. Zo komt $Pr\{\underline{x}_4 = A | \underline{x}_0 = A\} = 0.625$ tot stand door sommatie over de conditionele kansen op de 8 mogelijkheden AAAA, AABA, ..., BBBA voor $\underline{x}_1, \ldots, \underline{x}_4$, gegeven $\underline{x}_0 = A$.

Limietgedrag
In bovenstaand voorbeeld gaan de twee rijen van $P^{(m)}$ steeds meer op elkaar lijken naarmate m groter wordt. Met andere woorden, het verschil tussen $Pr\{\underline{x}_m = A | \underline{x}_0 = A\}$ en $Pr\{\underline{x}_m = A | \underline{x}_0 = B\}$ wordt steeds kleiner, evenals het verschil tussen $Pr\{\underline{x}_m = B | \underline{x}_0 = A\}$ en $Pr\{\underline{x}_m = B | \underline{x}_0 = B\}$. De interpretatie is: naarmate de jaren verstrijken wordt de invloed van de keuze uit jaar 0 steeds zwakker. Dit is in het voorbeeld niet verwonderlijk. Maar hebben alle Markovketens met stationaire overgangskansen deze eigenschap? De "meeste" wel, maar we zullen in paragraaf 4, waar we nader op dit verschijnsel ingaan, ook Markovketens tegenkomen waarbij de invloed van de begintoestand nooit verdwijnt.

VOORBEELD 1 (VERVOLG)
Eenvoudig kan worden bewezen, dat in ons voorbeeld het absolute verschil tussen termen in verschillende rijen maar in dezelfde kolom telkens halveert; er geldt:

$$\lim_{m \to \infty} P^{(m)} = \begin{bmatrix} 0.6 & 0.4 \\ 0.6 & 0.4 \end{bmatrix}.$$

Introduceer

$$P^{(\infty)} := \lim_{m\to\infty} P^{(m)}.$$

De rijen van $P^{(\infty)}$ geven de kansverdeling van de boekenkeus na oneindig veel jaren, ongeacht de beginkeus.

4. Matrices en toestandskansen in Markovketens

Tot dusverre hebben we slechts gesproken over overgangskansen, dus over conditionele kansen. Dikwijls is men echter vooral geinteresseerd in de kans dat een Markovketen zich in een bepaalde toestand bevindt op een tijdstip, *zonder* dat de begintoestand gegeven is. We spreken in dit geval over een *toestandskans;* de toestandskans dat de Markovketen zich op tijdstip m in toestand j bevindt wordt aangegeven met $\pi_j(m)$ en is gedefinieerd als

$$\pi_j(m) := Pr\{\underline{x}_m = j\}, \quad j \in S,\ m = 0,1,\ldots . \tag{4.1}$$

Er bestaat het volgende verband met overgangskansen:

$$\pi_j(m) = \sum_{i\in S} Pr\{\underline{x}_0 = i\} Pr\{\underline{x}_m = j | \underline{x}_0 = i\} = \sum_{i\in S} \pi_i(0) P_{ij}^{(m)}, \tag{4.2}$$

of in matrixnotatie, met de rijvector $\pi(m) := (\pi_1(m), \pi_2(m), \ldots)$:

$$\pi(m) = \pi(0) P^m. \tag{4.3}$$

Deze paragraaf is gewijd aan toestandskansen in Markovketens. Eerst moeten we enige nieuwe begrippen over toestanden introduceren.

Definitie 3

Een Markovketen heet *niet-reduceerbaar* als elke toestand bereikbaar is vanuit elke andere toestand, dwz. als voor elk paar toestanden (i,j) er een geheel getal m_0 (eventueel afhankelijk van i en j) bestaat, zodat

$$p_{ij}^{(m_0)} > 0. \tag{4.4}$$

Anders is sprake van een *reduceerbare Markovketen.*

Definitie 4

Een deelverzameling S_1 van de toestandsruimte S van een Markovketen heet *gesloten* als alle éénstaps overgangskansen van toestanden uit S_1 naar toestanden uit S_1^c, het complement van S_1 ten opzichte van S, nul zijn.
Als S_1 uit één toestand, zeg i, bestaat dan heet die toestand i *absorberend.*

Uit de voorgaande definities volgt:

(i) toestand i heet absorberend als en alleen als $p_{ii} = 1$.
(ii) Als S gesloten is en geen echte deelverzameling bevat die gesloten is, dan is de Markovketen niet-reduceerbaar.

We hebben gezien dat, in een Markovketen met stationaire overgangskansen, $p_{jj}^{(m)}$ de kans aangeeft dat een toestand j na m stappen weer optreedt. Het is dikwijls ook belangrijk, de kans te bepalen dat een toestand *voor het eerst* weer optreedt na m stappen. Definieer voor $j \in S$, $m = 1,2,...,$

$$f_{jj}^{(m)} := Pr\{\textit{eerste terugkeer naar toestand } j \textit{ treedt op na } m \textit{ stappen}\}. \tag{4.5}$$

Vanzelfsprekend is

$$f_{jj} := \sum_{m=1}^{\infty} f_{jj}^{(m)} = Pr\{\textit{ooit vindt een terugkeer naar toestand } j \textit{ plaats}\}. \tag{4.6}$$

We beperken ons in het vervolg tot niet-reduceerbare Markovketens. Het blijkt zinvol te zijn een klassificatie te geven van de toestanden van een (niet-reduceerbare) Markovketen. De terugkeerkansen f_{jj} spelen hierbij een grote rol.

DEFINITIE 5
Toestand j van een Markovketen heet *recurrent* als $f_{jj} = 1$;
toestand j van een Markovketen heet *transiënt* als $f_{jj} < 1$.
Als het enig mogelijke aantal stappen waarna de toestand j weer kan optreden gelijk is aan γ, 2γ, 3γ,... (met $\gamma > 1$ en zo groot mogelijk), dan heet toestand j *periodiek* met periode γ; als $\gamma = 1$ dan heet toestand j *aperiodiek.*

Als een toestand recurrent is, kunnen we spreken over de *gemiddelde terugkeertijd* m_j van toestand j:

$$m_j := \sum_{m=1}^{\infty} m f_{jj}^{(m)}. \tag{4.7}$$

We kunnen nu de toestandsklassificatie nog verder doorvoeren:

DEFINITIE 6
Toestand j van een Markovketen heet *positief-recurrent* als $f_{jj} = 1$ en $m_j < \infty$;
toestand j van een Markovketen heet *nul-recurrent* als $f_{jj} = 1$ en $m_j = \infty$.

Zonder bewijs vermelden we twee centrale stellingen uit de theorie der Markovketens (zie [3]). Stelling 1 geeft een verband tussen de klassificaties van toestanden van een niet-reduceerbare Markovketen (alle toestanden van zo'n Markovketen blijken van hetzelfde type te zijn; dit verklaart ook het nut van deze klassificatie). in stelling 2 keren we terug naar het begrip toestandskans.

STELLING 1
De toestanden van een niet-reduceerbare Markovketen zijn of alle transiënt, of alle nul-recurrent, of alle positief-recurrent.
Als een toestand periodiek is met periode γ, dan zijn alle toestanden periodiek met periode γ.
Als de toestandsruimte S van de niet-reduceerbare Markovketen een eindig

aantal toestanden heeft, dan zijn alle toestanden van S positief-recurrent.

DEFINITIE 7
Een Markovketen waarvan alle toestanden positief-recurrent zijn heet een positief-recurrente Markovketen.
Een Markovketen waarvan alle toestanden aperiodiek zijn heet een aperiodieke Markovketen.
Analoog voor nul-recurrente, transiënte en periodieke Markovketens.

STELLING 2
Beschouw een niet-reduceerbare, positief-recurrente, aperiodieke Markovketen (met stationaire overgangskansen). Er geldt:

$$\lim_{m\to\infty} p_{ij}^{(m)} = \pi_j, \quad i,j \in S; \tag{4.8}$$

$\{\pi_j, j \in S\}$ is een kansverdeling, die eenduidig bepaald is als de op één genormeerde niet-negatieve oplossing van het stelsel vergelijkingen

$$z_j = \sum_{i\in S} z_i p_{ij}, \quad j \in S. \tag{4.9}$$

Bovendien geldt:

$$\pi_j = \lim_{m\to\infty} \pi_j(m), \quad j \in S, \tag{4.10}$$

en

$$\pi_j = 1/m_j, \quad j \in S. \tag{4.11}$$

In paragraaf 3 merkten we op dat in de "meeste" Markovketens alle rijen van $P^{(m)}$ gelijk worden voor $m\to\infty$ - met als interpretatie dat de invloed van de begintoestand in m-staps overgangskansen voor grote waarden van m verdwijnt. Uit (4.8) blijkt dat dit inderdaad het geval is in de klasse van niet-reduceerbare, positief-recurrente, aperiodieke Markovketens; de rijen van $P^{(\infty)}$ blijken dan alle bovendien precies gelijk te zijn aan de rijvector, π, die de limietverdeling van toestandskansen weergeeft: $\pi_j = \lim_{m\to\infty} \pi_j(m)$.
Stelling 2 toont dat het voor de bepaling van π_j niet nodig is P tot steeds hogere machten te verheffen; een veel eenvoudiger methode is oplossing van het stelsel vergelijkingen (4.9). Het feit dat π_j, $j \in S$ aan (4.9) voldoet kan snel als volgt worden ingezien. (4.3) impliceert (in vectornotatie):

$$\pi(m+1) = \pi(0)P^{m+1} = \pi(0)P^m P = \pi(m)P, \quad m = 0,1,..., \tag{4.12}$$

zodat als $\pi_j = \lim_{m\to\infty} \pi_j(m)$ bestaat, π_j voldoet aan (4.9).

VOORBEELD 1 (VERVOLG)
In het wiskundeboekenvoorbeeld is de Markovketen niet-reduceerbaar, terwijl de toestandsruimte eindig is. Stelling 1 leert nu dat de Markovketen positief-recurrent is. Vanzelfsprekend is de Markovketen ook aperiodiek; stelling 2 is derhalve toepasbaar. Volgens deze stelling kan de limietverdeling van de toestandskansen, aan te duiden met de rijvector $\pi = (\pi_A, \pi_B)$, worden gevonden door oplossing van het stelsel vergelijkingen

$$\pi_A = 0.8\pi_A + 0.3\pi_B,$$
$$\pi_B = 0.2\pi_A + 0.7\pi_B,$$
$$\pi_A + \pi_B = 1.$$

Merk op dat de eerste twee vergelijkingen afhankelijk zijn; optelling geeft een identiteit (zoals algemeen het geval is bij (4.9)). De eerste (of tweede) vergelijking geeft

$$0.2\pi_A = 0.3\pi_B,$$

zodat uit de derde vergelijking volgt:

$$\pi_A = 0.6, \quad \pi_B = 0.4,$$

(in overeenstemming met het resultaat aan het eind van paragraaf 3).

OPMERKING 1
Een niet-reduceerbare, positief-recurrente, aperiodieke Markovketen heet "ergodisch". Men drukt het verdwijnen van de invloed van de begintoestand in ergodische Markovketens wel uit door te zeggen: de Markovketen nadert, als $m \to \infty$, tot een toestand van statistisch evenwicht. In dit verband is het goed te wijzen op een fysische interpretatie van de limiettoestandskansen. Het is duidelijk dat in een toestand van statistisch evenwicht geldt: de kans dat het proces zich *voor* de volgende stap in toestand j bevindt is even groot als de kans dat het proces zich *na* de volgende stap in toestand j bevindt. De eerste kans is gelijk aan π_j, en de tweede kans is gelijk aan de som, over alle mogelijke toestanden $i \in S$, van de kansen π_i maal de overgangskansen p_{ij} van toestand i naar toestand j:

$$\pi_j = \sum_{i \in S} \pi_i p_{ij}.$$

Vergelijk nu met (4.9). Dit "evenwichtsargument" is niet slechts toepasbaar op een bepaalde toestand, maar ook op de overgang tussen twee toestanden. In voorbeeld 1 moet er evenwicht bestaan tussen de "kansstroom" van A naar B en die van B naar A:

$$\pi_A p_{AB} = \pi_B p_{BA},$$

oftewel

$$0.2\pi_A = 0.3\pi_B,$$

wat tezamen met de normeringsvoorwaarde $\pi_A + \pi_B = 1$ leidt tot de juiste waarden van π_A en π_B.

VOORBEELD 2 (ZIE HOWARD [4])
Een taxichauffeur heeft zijn werkterrein in drie aangrenzende dorpen, die we voor het gemak met 1, 2 en 3 zullen aangeven. De ervaring leert dat de afhankelijkheid van opeenvolgende bestemmingen door een Markovketen kan worden beschreven, met onderstaande matrix van overgangskansen:

$$P = \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.1 & 0.8 & 0.1 \\ 0.4 & 0.4 & 0.2 \end{bmatrix}.$$

Met andere woorden, als de vorige rit in dorp 1 was geeindigd, dan blijft de volgende rit met kans 0.3 in dorp 1 en voert met kans 0.2 resp. 0.5 naar dorp 2 resp. dorp 3. Omdat alle overgangskansen positief zijn is de Markovketen niet-reduceerbaar. Volgens stelling 1 zijn alle toestanden positief-recurrent. Eenvoudig is in te zien dat alle toestanden aperiodiek zijn. Stelling 2 is nu toepasbaar; deze stelling leert dat de limietverdeling van de toestandskansen bestaat en eenduidig is vastgelegd als de oplossing van het stelsel vergelijkingen

$$\begin{aligned} \pi_1 &= 0.3\pi_1 + 0.1\pi_2 + 0.4\pi_3, \\ \pi_2 &= 0.2\pi_1 + 0.8\pi_2 + 0.4\pi_3, \\ \pi_3 &= 0.5\pi_1 + 0.1\pi_2 + 0.2\pi_3, \end{aligned}$$

en

$$\pi_1 + \pi_2 + \pi_3 = 1.$$

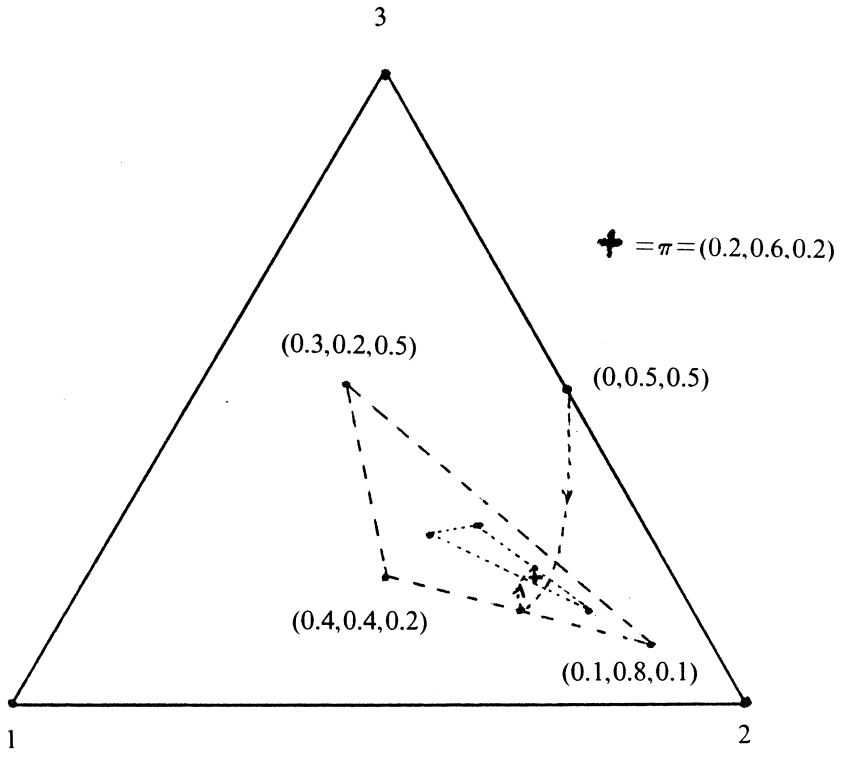
De oplossing luidt: $\pi_1 = 0.2$, $\pi_2 = 0.6$, $\pi_3 = 0.2$.
(Probeer zelf de evenwichtsargumenten uit opmerking 1 eens op dit voorbeeld toe te passen.) Volgens stelling 2 duurt het gemiddeld $m_1 = 1/\pi_1 = 5$ ritten, voor de taxi in dorp 1 terugkeert.
We geven de matrices $P^{(0)}$, $P^{(1)}$, $P^{(2)}$, $P^{(3)}$:

$$P^{(0)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$P^{(1)} = \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.1 & 0.8 & 0.1 \\ 0.4 & 0.4 & 0.2 \end{bmatrix},$$

$$P^{(2)} = \begin{bmatrix} 0.31 & 0.42 & 0.27 \\ 0.15 & 0.70 & 0.15 \\ 0.24 & 0.48 & 0.28 \end{bmatrix},$$

$$P^{(3)} = \begin{bmatrix} 0.243 & 0.506 & 0.251 \\ 0.175 & 0.650 & 0.175 \\ 0.232 & 0.544 & 0.224 \end{bmatrix}.$$



Figuur 1

Merk op dat de chauffeur die in dorp 1 startte, na één rit met kans 0.2 in dorp 2 zit, na 2 ritten met kans 0.42, na drie ritten met kans 0.506, en na oneindig veel ritten met kans 0.6. Voor dit laatste resultaat hoeven we de matrix P niet tot steeds hogere machten te verheffen; de voorgaande theorie leert dat $P^{(\infty)}$ wordt gegeven door

$$P^{(\infty)} = \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.2 & 0.6 & 0.2 \\ 0.2 & 0.6 & 0.2 \end{bmatrix}.$$

Inkrimping

Het telkens terugkerende verschijnsel van convergentie van overgangskansen en toestandskansen, als de tijd cq. het aantal stappen voortschrijdt, kan op grafische wijze worden gerepresenteerd. Beschouw hiertoe een gelijkzijdige driehoek met hoogte 1 (zie figuur 1). Een eigenschap van de gelijkzijdige driehoek is, dat de som van de lengten van de loodlijnen vanuit een inwendig punt op de drie zijden gelijk is aan de hoogte van de driehoek (dus hier gelijk aan 1). Derhalve kan elk inwendig punt van de driehoek een toestandsvector representeren (met som van de kansen gelijk aan 1) voor de Markovketen uit het taxivoorbeeld. Dat voorbeeld aanhoudend nummeren we de hoekpunten van de driehoek 1, 2 en 3. De drie hoekpunten representeren de vectoren (1,0,0), (0,1,0) en (0,0,1). Als de beginvector $\pi(0)=(1,0,0)$, dan is de toestandsvector na 1 stap (0.3,0.2,0.5); bij (0,1,0) hoort (0.1,0.8,0.1) en bij (0,0,1) hoort (0.4,0.4,0.2). De drie corresponderende punten markeren de gestreepte driehoek in figuur 1. Het is snel in te zien dat elk punt binnen de gelijkzijdige driehoek met hoogte 1 (dus elke begintoestandsvector) na één stap wordt afgebeeld op een punt binnen de kleinere, gestreepte, driehoek. Deze driehoek wordt op haar beurt bij de tweede stap van de Markovketen (de tweede rit van de taxi) op een nog kleinere driehoek afgebeeld (gestippeld in figuur 1). De hoekpunten van deze derde driehoek zijn de rijvectoren van $P^{(2)}$. De inkrimping gaat bij elke stap verder, met als convergentiepunt het punt behorend bij de limiettoestandsvector (0.2,0.6,0.2)!

Merk op dat de oriëntatie van de driehoeken telkens omklapt. Dit is duidelijk te zien als we als beginvector (0,0.5,0.5) nemen, halverwege de hoekpunten 2 en 3. Dit punt blijft telkens het middelpunt van een zijde, maar verschijnt tevens telkens aan de andere kant van de loodlijn uit hoekpunt 2. Wat is de achtergrond van die oriëntatiewisseling? We zullen deze vraag spoedig beantwoorden.

VOORBEELD 3

Stel dat bij het taxivoorbeeld de volgende matrix van overgangskansen optreedt:

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

De corresponderende Markovketen is wederom niet-reduceerbaar en positief-recurrent, maar elke toestand is nu periodiek met periode 3 (ga na). Het is duidelijk dat in dit geval de invloed van de begintoestand nooit verdwijnt; als de eerste rit in dorp 1 start, dan is de chauffeur $3m$ ritten later weer in dorp 1, $m = 1,2,\ldots$.

VOORBEELD 4
Stel dat bij het taxivoorbeeld de volgende matrix van overgangskansen optreedt:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.3 & 0.2 & 0.5 \end{bmatrix}.$$

De bijbehorende Markovketen is *reduceerbaar.* Toestanden 1 en 2 zijn absorberend (zie definitie 4), en toestand 3 is transiënt. Een eenvoudige berekening leert dat

$$P^{(\infty)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.6 & 0.4 & 0 \end{bmatrix}.$$

Als u , analoog aan figuur 1, de inkrimpende driehoeken tekent zult u merken dat in dit geval convergentie naar een lijn (i.p.v. naar een punt) optreedt: de lijn tussen de hoekpunten 1 en 2.

De rijen van $P^{(\infty)}$ zijn niet identiek: de limietverdeling van de toestandskansen hangt af van de begintoestand. Als de taxi in dorp 1 of 2 is gestart, dan blijft hij daar altijd. Als de taxi in dorp 3 is gestart, dan eindigt hij in dorp 1 resp. 2 met kansen 0.6 resp. 0.4.

Probeer zelf voor het wiskundeboekenvoorbeeld de convergentie grafisch voor te stellen.

Eigenwaarden en eigenvectoren van de matrix van overgangskansen
Uit (4.9) blijkt dat de limiettoestandsvector π (en alle rijen van $P^{(\infty)}$) van een niet-reduceerbare, positief-recurrente, aperiodieke Markovketen gelijk is aan de linkereigenvector van de matrix P, behorende bij de eigenwaarde 1. Merk op dat de bijbehorende rechtereigenvector (een veelvoud van) de eenheidsvector $1 := (1,\ldots,1)^t$ is; alle rijsommen van P zijn immers gelijk aan één (opm.: de index t geeft de getransponeerde van een vector aan; de rechtereigenvectoren zijn kolomvectoren). We zullen tonen dat bij de bepaling van een expliciete uitdrukking voor $p_{ij}^{(m)}$ ook de andere eigenwaarden en eigenvectoren van P een essentiële rol spelen. Een interessante bonus van de bestudering van eigenwaarden en eigenvectoren van P is een analytische verklaring van het meetkundige inkrimpingsverschijnsel uit figuur 1. Voor dit alles gebruiken we de stelling van Perron-Frobenius, de basisstelling voor *niet-negatieve* matrices, dwz. matrices met uitsluitend niet-negatieve elementen. Alvorens een vorm

van deze stelling te kunnen formuleren, hebben we nog één begrip nodig. De analogie tussen niet-negatieve matrices en Markovketens benadrukkend noemen we een vierkante, niet-negatieve matrix $T = (t_{ij})$ *niet-reduceerbaar* als voor elk indicespaar (i,j) er een positief geheel getal m_0 (eventueel afhankelijk van i en j) bestaat, zodat

$$t_{ij}^{(m_0)} > 0,$$

(vgl. definitie 3).

STELLING 3
(stelling van Perron-Frobenius; zie Seneta [7])
Zij $T = (t_{ij})$ een niet-negatieve, niet-reduceerbare $k \times k$ matrix. Dan bestaat er een eigenwaarde r van T zodat

(i) r is reëel en positief;
(ii) met r kunnen een linker- en rechtereigenvector met uitsluitend positieve elementen worden geassocieerd;
(iii) $r \geqslant |\lambda|$ voor elke eigenwaarde λ;
(iv) $\min \sum_{j=1}^{k} t_{ij} \leqslant r \leqslant \max \sum_{j=1}^{k} t_{ij}$.

Omdat de rijsommen van een stochastische matrix gelijk zijn aan één, volgt uit stelling 3:

GEVOLG 1
Als T een stochastische matrix is, dan is $r = 1$ een eigenwaarde; alle overige eigenwaarden zijn in absolute waarde hoogstens gelijk aan één.

We beperken ons in het vervolg weer tot niet-reduceerbare Markovketens, met $k < \infty$ toestanden; volgens stelling 1 is zo'n Markovketen positief-recurrent. De stelling van Perron-Frobenius impliceert dat, voor een niet-reduceerbare, positief-recurrente Markovketen, alle eigenwaarden in absolute waarde hoogstens gelijk zijn aan één. Slechts als de keten periodiek is, zijn er meerdere eigenwaarden in absolute waarde gelijk aan één (bereken de eigenwaarden bij de matrix in voorbeeld 3). Zij P de overgangskansmatrix van de Markovketen, en zij $\lambda_1, \ldots, \lambda_k$ de eigenwaarden van P; $\lambda_1 = 1 \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_k|$. Voor het gemak nemen we aan dat alle eigenwaarden multipliciteit één hebben; de theorie behoeft enige aanpassing in het algemene geval (zie Feller [3]). Zij x_j en y_j de linker- en rechtereigenvector van P behorend bij eigenwaarde λ_j, dwz. er geldt:

$$x_j P = \lambda_j x_j, \qquad P y_j = \lambda_j y_j. \tag{4.13}$$

Zij X de matrix met als rijen $x_1, \ldots, x_k$, en zij Y de matrix met als kolommen $y_1, \ldots, y_k$. De lineaire algebra leert: als alle eigenwaarden verschillend zijn, vormen de linkereigenvectoren een lineair onafhankelijk stelsel zodat X niet-singulier is en een inverse heeft; analoog voor Y. We kunnen de eigenvectoren zodanig kiezen dat $Y = X^{-1}$. Merk hiertoe op dat voor $i \neq j$ (dus $\lambda_i \neq \lambda_j$),

$$x_i y_j = \frac{1}{\lambda_i} x_i P y_j = \frac{\lambda_j}{\lambda_i} x_i y_j,$$

zodat

$$x_i y_j = 0, \quad i \neq j.$$

De eigenvectoren van P zijn op een multiplicatieve factor na bepaald; we kunnen en zullen de eigenvectoren zodanig kiezen dat

$$x_j y_j = 1, \quad j = 1,...,k.$$

Het bovenstaande impliceert dat

$$Y = X^{-1}. \tag{4.14}$$

We introduceren nog enige matrices: de diagonaalmatrix

$$\Lambda := \begin{bmatrix} \lambda_1 & 0 & . & . & 0 \\ 0 & \lambda_2 & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & \lambda_k \end{bmatrix}, \tag{4.15}$$

en de k matrices

$$B_j := y_j x_j, \quad j = 1,...,k. \tag{4.16}$$

Nu kan een zeer nuttige *spectraalrepresentatie* van de matrices $P^{(m)}$ worden gegeven:

STELLING 4

$$P^{(m)} = P^m = Y \Lambda^m X = \sum_{j=1}^{k} \lambda_j^m B_j, \quad m = 0,1,... . \tag{4.17}$$

BEWIJS

De eerste gelijkheid volgt uit (3.5). De tweede gelijkheid volgt uit (4.14) en de identiteit

$$PY = Y\Lambda,$$

welke op haar beurt volgt uit (4.13). De derde gelijkheid volgt door uitschrijven van $Y\Lambda^m X$ (zie (4.16)).

OPMERKING 2
Stelling 4 impliceert in het geval dat $|\lambda_2|<1$: er bestaat een positieve constante C zodat

$$|p_{ij}^{(m)} - p_{ij}^{(\infty)}| = |p_{ij}^{(m)} - \pi_j| \leqslant C|\lambda_2|^m, \quad i,j=1,...,k, \quad m=1,2,... ; \quad (4.18)$$

$P^{(m)}$ convergeert meetkundig naar een matrix $P^{(\infty)}$ met als rijen de limiettoestandsvectoren π, en de convergentiesnelheid wordt bepaald door de één-na-grootste eigenwaarde van P.

VOORBEELD 1 (VERVOLG)
De eigenwaarden van de matrix

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix},$$

zijn: $\lambda_1=1$ en $\lambda_2=0.5$. De bijbehorende eigenvectoren, genormeerd zodat $x_j y_j=1$, zijn:
$x_1 = (0.6,0.4), y_1 = (1,1)^t, x_2 = (1,-1), y_2 = (0.4,-0.6)^t$.
B_1 is (vanzelfsprekend, gezien de betekenis van x_1 en de vorm van y_1) gelijk aan de al aan het eind van paragraaf 3 verkregen matrix $P^{(\infty)}$; B_2 wordt gegeven door

$$B_2 = \begin{bmatrix} 0.4 & -0.4 \\ -0.6 & 0.6 \end{bmatrix}.$$

Stelling 4 levert nu de volgende spectraalrepresentatie van de m-staps overgangsmatrix $P^{(m)}$ op: voor $m=0,1,...,$

$$P^{(m)} = \begin{bmatrix} 0.6 & 0.4 \\ 0.6 & 0.4 \end{bmatrix} + (0.5)^m \begin{bmatrix} 0.4 & -0.4 \\ -0.6 & 0.6 \end{bmatrix}. \quad (4.19)$$

Formule (4.19) bevestigt precies de resultaten die we in de vorige paragrafen hadden verkregen cq. vermoed: $P^{(m)}$ convergeert naar $P^{(\infty)}$, een matrix met identieke rijen (0.6 0.4), en het verschil tussen de elementen van $P^{(m)}$ en $P^{(\infty)}$ wordt elke volgende stap gehalveerd. Merk op dat B_2 onmiddellijk volgt uit de relatie (zie (4.17) voor $m=0$) $B_1+B_2=I$.

VOORBEELD 2 (VERVOLG)
We keren terug naar het taxivoorbeeld. De eigenwaarden van de matrix

$$P = \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.1 & 0.8 & 0.1 \\ 0.4 & 0.4 & 0.2 \end{bmatrix}$$

zijn: $\lambda_1=1$, $\lambda_2=0.5$ en $\lambda_3=-0.2$. De bijbehorende eigenvectoren, genormeerd zodat $x_j y_j=1$, zijn:

$$x_1 = (0.2,0.6,0.2), y_1 = (1,1,1)^t, x_2 = (-\frac{1}{35},\frac{2}{35},-\frac{1}{35}), y_2 = (-13,7,-8)^t,$$

$$x_3 = (\frac{3}{7},\frac{1}{7},-\frac{4}{7}), y_3 = (1,0,-1)^t.$$

Na enig rekenwerk volgt voor $m = 0, 1, \ldots$:

$$P^{(m)} = \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.2 & 0.6 & 0.2 \\ 0.2 & 0.6 & 0.2 \end{bmatrix} + (0.5)^m \begin{bmatrix} \frac{13}{35} & -\frac{26}{35} & \frac{13}{35} \\ -\frac{1}{5} & \frac{2}{5} & -\frac{1}{5} \\ \frac{8}{35} & -\frac{16}{35} & \frac{8}{35} \end{bmatrix} + \tag{4.20}$$

$$+ (-0.2)^m \begin{bmatrix} \frac{3}{7} & \frac{1}{7} & -\frac{4}{7} \\ 0 & 0 & 0 \\ -\frac{3}{7} & -\frac{1}{7} & \frac{4}{7} \end{bmatrix}.$$

Ditmaal zijn er twee "inkrimpings"bijdragen. Men kan bewijzen dat de oppervlakten van de driehoeken telkens inkrimpen met een factor die gelijk is aan $|det(P)|$. Volgens een stelling uit de matrixrekening is (zie ook paragraaf 6 van de bijdrage van Van der Blij in deze bundel):

$$|det(P)| = |det(Y\Lambda Y^{-1})| = |det(\Lambda)| = |\prod \lambda_i|.$$

De omkering in oriëntatie van de inkrimpende driehoeken, die werd geconstateerd in Figuur 1, wordt veroorzaakt door het feit dat $det(P)$ negatief is.

5. Aanvullingen en verwijzingen

De theorie der Markovprocessen is een zeer rijk en uitgebreid onderdeel van de theorie der stochastische processen. We hebben in dit verhaal slechts het topje van de ijsberg getoond: enkele basisbegrippen van Markovprocessen met discrete tijdsparameter, de zg. Markovketens, met sterke nadruk op de relatie met de matrixrekening.

We hebben ons grotendeels beperkt tot aperiodieke ketens, maar dat is geen al te wezenlijke beperking. Het onderscheid tussen recurrente en transiënte Markovketens is veel essentiëler. Voor een bespreking van transiënte ketens verwijzen we naar Chung [1], een standaardwerk over Markovprocessen.

De theorie der Markovprocessen met *continue* tijdsparameter vertoont een sterke analogie met die der Markovketens. Markovprocessen $\{x_t, t \in T\}$ met continue tijdsparameter zijn stochastische processen met de eigenschap dat, gegeven de huidige toestand x_{t_0}, geen aanvullende informatie betreffende het proces vóór t_0 de conditionele kans op het optreden van een willekeurige toestand x in de toekomst kan beinvloeden. Zie Chung [1] of Cohen [2] voor een formele definitie, een inleiding in de theorie en verdere verwijzingen. [2] is verder hoofdzakelijk gewijd aan een belangrijk toepassingsgebied van Markovprocessen, de wachtrijtheorie.

Een interessante uitbreiding van Markovprocessen wordt gevormd door de

Markovbeslissingsprocessen. Dat zijn Markovprocessen waarbij de overgangskansen kunnen afhangen van eerder genomen beslissingen; zie Howard [5]. Howard's boeken hebben een zelfde moeilijkheidsgraad als dit artikel; [1-3] zijn geavanceerder. De behandeling van Wessels en Van Nunen [8] is elementair en bevat veel aardige voorbeelden.

In dit verhaal hebben we enige malen gelijkenissen en verbanden aangegeven tussen de theorie van Markovketens en die van vierkante, niet-negatieve matrices. Voor een zeer gedetailleerde bespreking van zulke matrices, en de toepassing ervan op Markovketens, zij verwezen naar Seneta [7].

LITERATUUR

1. K.L. CHUNG (1967). *Markov Chains with Stationary Transition Probabilities* (Springer, Berlin).
2. J.W. COHEN (1982). *The Single Server Queue* (North-Holland Publ. Cy., Amsterdam).
3. W. FELLER (1957). *An Introduction to Probability Theory and its Applications, Vol. I* (Wiley, New York).
4. R.A. HOWARD (1971). *Dynamic Probabilistic Systems; Volume I: Markov Models* (Wiley, New York).
5. R.A. HOWARD (1971). *Dynamic Probabilistic Systems, Vol. II: Semi-Markov and Decision Processes* (Wiley, New York).
6. A.A. MARKOV (1907). *Extension of the limit theorems of probability theory to a sum of variables connected in a chain,* The Notes of the Imperial Academy of Sciences of St. Petersburg VIII Series, Physio-Mathematical College, Vol. XXII, No. 9.
7. E. SENETA (1981). *Non-negative Matrices and Markov Chains* (Springer, Berlin).
8. J. WESSELS EN J. VAN NUNEN (1983). *Processen uit het dagelijks leven,* Euclides 58, 202-218.

MATRICES EN STATISTIEK

A.J. Bosch
Technische Hogeschool Eindhoven

1. Enkele punten uit de lineaire algebra 105

2. Enkele punten uit de mathematische statistiek 106

3. Lineaire regressie-analyse 107

4. Correlaties 110

Literatuur 114

MATRICES EN STATISTIEK

1. Enkele punten uit de lineaire algebra

In deze paragraaf worden enkele definities en stellingen uit de lineaire algebra in herinnering gebracht.

Algemeen: beschouwd worden alleen matrices met reële getallen als elementen.
Matrices geven we aan met hoofdletters A, B, H, X e.d..
I is de eenheidsmatrix;
A^t is de getransponeerde van A, d.w.z. $(A^t)_{ij} = a_{ji}$.
Vektoren worden met kleine letters aangegeven.
a^t is een rij-, a een kolomvektor. Vektoren onderstrepen we dus niet.

1.1. Er geldt $(AB)^t = B^tA^t$; $(A^{-1})^t = (A^t)^{-1} =: A^{-t}$.
$\text{rang}(A^tA) = \text{rang}\, A$.

1.2. $\langle A \rangle$ is de kolommenruimte van A, dat is de lineaire ruimte opgespannen door de kolommen van A; $Ax \in \langle A \rangle$.

1.3. H is een orthogonale matrix als $H^tH = HH^t = I$.

1.4. Als $Ax = \lambda x$ met $x \neq 0$, dan heet x eigenvektor van A met eigenwaarde λ.

1.5. Zij D een lineaire deelruimte $\subset \mathbb{R}^n$. Het orthogonale complement van D is $D^\perp := \{x \in \mathbb{R}^n \mid x \perp D\}$.

1.6. Het inwendig produkt van 2 vektoren in $\mathbb{R}^n$ is $(a,b) = a^tb$.
Voor de lengte geldt $\|a\|^2 = a^ta$.
De hoek tussen a,b wordt gedefinieerd via $\cos\varphi = \dfrac{(a,b)}{\|a\|\,\|b\|}$.

Symmetrische matrices: een matrix heet symmetrisch als $A^t = A$.

1.7. Is A symmetrisch, dan is er een orthogonale matrix H (bestaande uit orthonormale eigenvektoren) en een reële diagonaalmatrix Λ (bestaande uit de bijbehorende eigenwaarden) zodat $A = H\Lambda H^t$.

1.8. Een symmetrische matrix A heet positief definiet ($A > 0$) als voor $x \neq 0$, $x^tAx > 0$.
Een symmetrische matrix A heet semi positief definiet ($A \geqslant 0$) als $x^tAx \geqslant 0$.
$A > B$ wil zeggen $A,B \geqslant 0$ en $A - B > 0$.

1.9. Is B van volle kolomrang dan geldt: $A > 0 \Rightarrow B^tAB > 0$.
Bewijs: $x^tB^tABx = y^tAy > 0$ waarbij $y = Bx \neq 0$ als $x \neq 0$.

1.10. Is $A > 0$ dan zijn alle eigenwaarden positief.
Bewijs: $A = H\Lambda H^t$ dus met 1.9 is $\Lambda > 0$ d.w.z. alle $\lambda > 0$.

1.11. Is $A > 0$, dan ook $A^{-1} > 0$.
Bewijs: $A^{-1} = H\Lambda^{-1}H^t$; $\Lambda > 0$ dus ook Λ^{-1} en weer met 1.9 dus ook $A^{-1} > 0$.

1.12. $A^tA \geqslant 0$.
Bewijs: $x^tA^tAx = (Ax)^tAx = \|Ax\|^2 \geqslant 0$.

1.13. Is $A > 0$, dan is er een $B > 0$ zodat $A = B^2$.
Bewijs: $A = H\Lambda H^t = H\Lambda^{\frac{1}{2}}H^tH\Lambda^{\frac{1}{2}}H^t = B^2$ met $B = H\Lambda^{\frac{1}{2}}H^t > 0$.

1.14. Een matrix P is een <u>orthogonale projector</u> als $P^tP = P$ (dus P is symmetrisch en idempotent).

1.15. Er geldt $P \geqslant 0$, d.w.z. P is semi positief definiet. Dit volgt direct uit 1.12.

1.16. $P_X = X(X^tX)^{-1}X^t$ is een orthogonale projector op $\langle X \rangle$.
Bewijs: P_X is symmetrisch, idempotent; $P_XX = X$ en $\operatorname{rang} P_X = \operatorname{rang} X$.

1.17. $I - P_X$ is een orthogonale projector op $\langle X \rangle^{\perp}$.

2. Enkele punten uit de mathematische statistiek

Bij de behandeling van de lineaire regressie-analyse zullen we enkele basisbegrippen uit de mathematische statistiek nodig hebben.

2.1. Stochasten geven we weer met onderstreepte, latijnse letters. Parameters met griekse letters.
We weten: het gemiddelde $\mu = \mathcal{E}\underline{x}$
de covariantie $\operatorname{cov}(\underline{x},\underline{y}) = \mathcal{E}\{(\underline{x} - \mu_x)(\underline{y} - \mu_y)\}$
de variantie van $\underline{x}$, $\operatorname{var}\underline{x} = \sigma^2 = \mathcal{E}(\underline{x} - \mu)^2$.
Als $\underline{x}$ en $\underline{y}$ onderling onafhankelijk zijn is $\operatorname{cov}(\underline{x},\underline{y}) = 0$. (Niet omgekeerd!)
De <u>correlatiecoëfficiënt</u> tussen $\underline{x}$ en $\underline{y}$ is $\rho(\underline{x},\underline{y}) = \dfrac{\operatorname{cov}(\underline{x},\underline{y})}{\sigma_x\sigma_y}$.
Er geldt $|\rho| \leqslant 1$; En $|\rho| = 1$ als $\underline{y} = \alpha\underline{x} + \beta$.
ρ is een maat voor <u>lineaire</u> afhankelijkheid.

2.2. De verwachting van een stochastische matrix (of vektor) is de matrix

(vektor) van zijn verwachtingen: $\mathcal{E}\underline{x} = (\mathcal{E}\underline{x}_{ij})$.
Stel $\underline{x} = (\underline{x}_1,\ldots,\underline{x}_p)^t$ dan is $\mu = \mathcal{E}\underline{x} = (\mathcal{E}\underline{x}_1,\ldots,\mathcal{E}\underline{x}_p)^t = (\mu_1,\ldots,\mu_p)^t$.

$\operatorname{var}\underline{x} = \Sigma^2 = \mathcal{E}(\underline{x}-\mu)(\underline{x}-\mu)^t =$

$$= \begin{pmatrix} \operatorname{var}\underline{x}_1 & \operatorname{cov}(\underline{x}_1,\underline{x}_2) & \cdots\cdots & \operatorname{cov}(\underline{x}_1,\underline{x}_p) \\ \operatorname{cov}(\underline{x}_2,\underline{x}_1) & \ddots & & \vdots \\ \vdots & & \ddots & \\ \operatorname{cov}(\underline{x}_p,\underline{x}_1) & \cdots\cdots\cdots\cdots & & \operatorname{var}\underline{x}_p \end{pmatrix}$$

m.a.w. $(\Sigma^2)_{ij} = \operatorname{cov}(\underline{x}_i,\underline{x}_j)$.
Σ^2 wordt de variantie-covariantie matrix genoemd van de vektor $\underline{x}$.

2.3. $\operatorname{cov}(\underline{x},\underline{y}) = \mathcal{E}(\underline{x}-\mu_x)(\underline{y}-\mu_y)^t$.

$\operatorname{cov}(A\underline{x},B\underline{y}) = \mathcal{E}(A\underline{x}-A\mu_x)(B\underline{y}-B\mu_y)^t = A\operatorname{cov}(\underline{x},\underline{y})B^t$.
In het bijzonder: $\operatorname{var}A\underline{x} = A\operatorname{var}\underline{x}\,A^t = A\Sigma^2A^t$.
$\operatorname{cov}(\underline{x}+B\underline{y},\underline{z}) = \operatorname{cov}(\underline{x},\underline{z}) + B\operatorname{cov}(\underline{y},\underline{z})$.

2.4. $\Sigma^2 \geqslant 0$ nl. voor $\alpha \neq 0$ is $\operatorname{var}(\alpha^t\underline{x}) = \alpha^t\Sigma^2\alpha \geqslant 0$.
$\Sigma^2 > 0$ als de $\underline{x}_1,\ldots,\underline{x}_p$ lineair onafhankelijk zijn.

2.5. De <u>correlatiematrix</u> van $\underline{x} = (\underline{x}_1,\ldots,\underline{x}_p)^t$ is gedefinieerd als

$$R := S^{-1}\Sigma^2S^{-1} \quad \text{waarbij } (S)_{ii} = \sqrt{\operatorname{var}\underline{x}_i} \text{ en } (S)_{ij} = 0 \text{ voor } i \neq j\,.$$

2.6. Een <u>schatter</u> voor een parameter θ is een funktie van de waarnemingen.
Een schatter $\underline{t}$ voor θ heet <u>zuiver</u> als $\mathcal{E}\underline{t} = \theta$.

2.7. Beschouw de vektorruimte van ééndimensionale stochasten ($\underline{x}$, $\underline{y}$ e.d.).
Onderstel alle hebben gemiddelde 0 dus $\mu_x = \mu_y = 0$ etc..
Dan voeren we als volgt een inprodukt in: $(\underline{x},\underline{y}) := \mathcal{E}(\underline{x}\underline{y})$.
Dit geeft $\|\underline{x}\|^2 = \mathcal{E}\underline{x}^2 = \operatorname{var}\underline{x}$.

En $\cos(\underline{x},\underline{y}) = \rho(\underline{x},\underline{y})$. Dus $\underline{x} \perp \underline{y}$ betekent $\rho(\underline{x},\underline{y}) = 0$.

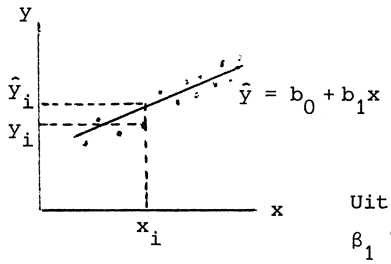
3. Lineaire regressie-analyse

3.1. Enkelvoudige lineaire regressie

In deze voordracht ligt het accent op de matrixrekening. Er zal dus niet te diep op de statistische aspecten worden ingegaan.

Stel we hebben twee variabelen, η en x, in de relatie $\eta = \beta_0 + \beta_1 x$. De parameters β_0 en β_1 zijn echter onbekend. Uit paren metingen aan (x,η) willen we β_0 en β_1 schatten.

We nemen aan dat x, de onafhankelijke variabele, zonder fout gemeten kan worden. De afhankelijke variabele η kan slechts met een fout gemeten worden. Zo ontstaat het model:



$$\underline{y}_i = \beta_0 + \beta_1 x_i + \underline{e}_i \ ,$$

$$i = 1,\ldots,n \ ; \ \mathcal{E}\underline{e}_i = 0 \ .$$

Uit deze n paren waarnemingen willen we β_0 en β_1 "zo goed mogelijk" schatten. Hiertoe gebruiken we de methode der kleinste kwadraten. Daarbij wordt $f(\beta_0,\beta_1) := \sum_i (y_i - \beta_0 - \beta_1 x_i)^2$ geminimaliseerd naar β_0 en β_1.

Differentiatie naar β_0 en β_1 geeft: $\frac{\partial f}{\partial \beta_0} = -2 \Sigma (y_i - \beta_0 - \beta_1 x_i) = 0$

$$\frac{\partial f}{\partial \beta_1} = -2 \Sigma (y_i - \beta_0 - \beta_1 x_i) x_i = 0 \ .$$

Hieruit kunnen we b_0,b_1 (schattingen voor β_0,β_1) berekenen.
De regressievergelijking wordt dan: $\hat{y} = b_0 + b_1 x$. Deze lijn past "het beste" bij de puntenwolk, in die zin dat elke andere rechte lijn een grotere Σr_i^2 (restkwadratensom) geeft. Hierin is $r_i := y_i - \hat{y}_i$ het i^e-residu.
Enkelvoudig betekent dat y slechts van één variabele afhangt.
Lineair betekent, lineair in de parameters.
Hangt y van meerdere variabelen af, dan is het veel duidelijker de regressie-analyse m.b.v. matrixrekening te behandelen.

3.2. Meervoudige lineaire regressie-analyse

De variabele η hangt nu van meerdere variabelen $x_1,\ldots,x_p$ af:

$$\eta = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p \ .$$

Ondersteld is weer dat de x_i zonder fout gemeten kunnen worden. Uit waarnemingen willen we de onbekende parameters β_i schatten. Het model is nu

$$\underline{y}_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + \underline{e}_i \ , \qquad i = 1,\ldots,n \ .$$

In matrixnotatie luidt dit:

$$\underline{y} = X\beta + \underline{e}$$

waarin

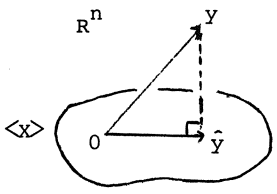
$\underline{y} = (\underline{y}_1,\ldots,\underline{y}_n)^t$ de responsevektor

$\beta = (\beta_0,\ldots,\beta_p)^t$ de parametervektor (die geschat moet worden)

$$X_{n\times p} = \begin{pmatrix} 1\, x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ 1\, x_{n1} & \cdots & x_{np} \end{pmatrix}$$ de designmatrix. Ondersteld is rang $X = p+1$.

$\underline{e} = (\underline{e}_1,\ldots,\underline{e}_n)^t$ de storingsvektor; $\mathcal{E}\underline{e} = 0$ en $\operatorname{var}\underline{e} = \Sigma^2$.

De regressievergelijking is nu $\hat{y} = Xb$. Hoe bepalen we b en $\hat{y}$?



We zien dat $\hat{y} \in \langle X\rangle$. Bovendien is $\|y-\hat{y}\|^2$ minimaal. Dit betekent dat $y - \hat{y} \perp \langle X\rangle$ oftewel $\hat{y}$ is de loodrechte projectie van y op $\langle X\rangle$.
Dus $\hat{y} = P_X y = X(X^tX)^{-1}X^ty$. Dit kunnen we ook direkt afleiden: $y - Xb \perp \langle X\rangle$ d.w.z. $X^t(Y - Xb) = 0$, de normaalvergelijkingen. Hieruit volgt: $X^ty = X^tXb$ en $b = (X^tX)^{-1}X^ty$, $\hat{y} = Xb = X(X^tX)^{-1}X^ty$.

Opmerking: $\sum_i r_i = \sum_i (y_i - \hat{y}_i) = 0$ daar $y - \hat{y} \perp \langle X\rangle$, dus ook $\perp(1,\ldots,1)$.

$\underline{b}$ heet de kleinste kwadratenschatter voor β; deze is zuiver nl.
$\mathcal{E}\underline{b} = \mathcal{E}(X^tX)^{-1}X^t\underline{y} = (X^tX)^{-1}X^tX\beta = \beta$.

Een interessante vraag dringt zich op: is er een "betere" (lineaire, zuivere) schatter voor β? Hiertoe definiëren we: $\tilde{\underline{b}}$ is een betere schatter dan $\underline{b}$ als $\operatorname{var}\underline{b} \geqslant \operatorname{var}\tilde{\underline{b}}$.
Beschouw eens $\tilde{\underline{b}} := (X^t\Sigma^{-2}X)^{-1}X^t\Sigma^{-2}\underline{y}$. Deze is ook lineair en zuiver, nl.

$$\mathcal{E}\tilde{\underline{b}} = (X^t\Sigma^{-2}X)^{-1}X^t\Sigma^{-2}X\beta = \beta\ .$$

Deze blijkt beter te zijn dan $\underline{b}$, nl.

$$\operatorname{var}\underline{b} = (X^tX)^{-1}X^t\Sigma^2X(X^tX)^{-1}$$

en

$$\operatorname{var}\tilde{\underline{b}} = (X^t\Sigma^{-2}X)^{-1}X^t\Sigma^{-2}\Sigma^2\Sigma^{-2}X(X^t\Sigma^{-2}X)^{-1} = (X^t\Sigma^{-2}X)^{-1}$$

$$\operatorname{var}\underline{b} - \operatorname{var}\tilde{\underline{b}} = (X^tX)^{-1}X^t\Sigma^2X(X^tX)^{-1} - (X^t\Sigma^{-2}X)^{-1} =$$

$$= (X^tX)^{-1}X^t\Sigma\,[I - \Sigma^{-1}X(X^t\Sigma^{-2}X)^{-1}X^t\Sigma^{-1}]\,\Sigma X(X^tX)^{-1} =$$

$$=: B[I - A]B^t \text{ waarin } B := (X^tX)^{-1}X^t\Sigma \text{ van volle rijrang}$$

en

$$A := P_{\Sigma^{-1}X} \geqslant 0\,, \text{ en dus ook } I - A \geqslant 0\,.$$

Met (1.9) volgt dus $\operatorname{var}\underline{b} \geqslant \operatorname{var}\tilde{\underline{b}}$.
$\tilde{\underline{b}}$ heet de <u>gewogen</u> kleinste kwadratenschatter voor β. Deze blijkt de "beste lineaire zuivere schatter" voor β te zijn.
Eenvoudig is na te gaan dat ook $\operatorname{var}\hat{\underline{y}} \leqslant \operatorname{var}\underline{y}$.

4. Correlaties

In de meervoudige regressie-analyse bestudeerden we één variabele, afhankelijk van veel andere (verklarende) variabelen.
Nu beschouwen we twee deelverzamelingen van variabelen:

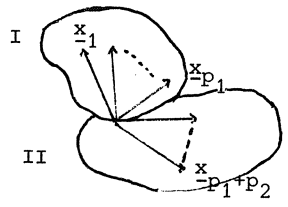
I de te verklaren variabelen $\underline{x}_1,\ldots,\underline{x}_{p_1}$. In één vektor: $\underline{x}^{(1)} = (\underline{x}_1,\ldots,\underline{x}_{p_1})^t$.

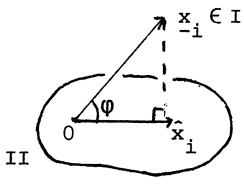
II de verklarende variabelen $\underline{x}_{p_1+1},\ldots,\underline{x}_{p_1+p_2}$. In één vektor:

$\underline{x}^{(2)} = (\underline{x}_{p_1+1},\ldots,\underline{x}_{p_1+p_2})^t$.

Tevens noteren we $\operatorname{var}\underline{x}^{(1)} = \Sigma_{11}$; $\operatorname{var}\underline{x}^{(2)} = \Sigma_{22}$; $\operatorname{cov}(\underline{x}^{(1)},\underline{x}^{(2)}) = \Sigma_{12}$.

Samengevat $\underline{x} = (\underline{x}^{(1)t},\underline{x}^{(2)t})^t$ met $\operatorname{var}\underline{x} = \Sigma^2 = \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}$.





We zoeken in II een lineaire combinatie die maximaal gecorreleerd is met $\underline{x}_i \in I$.
Deze noteren we als $\hat{\underline{x}}_i$.

Algemeen is $\hat{\underline{x}}^{(1)} = B\underline{x}^{(2)}$. We zien dat $\underline{x}_i - \hat{\underline{x}}_i \perp \underline{x}^{(2)}$. Dus

$$\operatorname{cov}(\underline{x}^{(1)} - B\underline{x}^{(2)}, \underline{x}^{(2)}) = 0 = \Sigma_{12} - B\Sigma_{22} \quad \text{oftewel} \quad B = \Sigma_{12}\Sigma_{22}^{-1}.$$

$$\operatorname{var}\hat{\underline{x}}^{(1)} = \operatorname{var} B\underline{x}^{(2)} = B\Sigma_{22}B^t = \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} =: \Sigma_{1\cdot 2}.$$

$$\operatorname{var}(\underline{x}^{(1)} - \hat{\underline{x}}^{(1)}) = \operatorname{var}\underline{x}^{(1)} + \operatorname{var}\hat{\underline{x}}^{(1)} - 2\operatorname{cov}(\underline{x}^{(1)}, \hat{\underline{x}}^{(1)}) =$$

$$= \Sigma_{11} + \Sigma_{1\cdot 2} - 2\Sigma_{12}B^t = \Sigma_{11} - \Sigma_{1\cdot 2} =: \Sigma_{11\cdot 2}.$$

Er geldt dus:

$$\operatorname{var}\underline{x}^{(1)} = \operatorname{var}\hat{\underline{x}}^{(1)} + \operatorname{var}(\underline{x}^{(1)} - \hat{\underline{x}}^{(1)}) \quad \text{oftewel} \quad \Sigma_{11} = \Sigma_{1\cdot 2} + \Sigma_{11\cdot 2}.$$

We definiëren de volgende correlatiecoëfficiënten:

$$\rho_{ij} := \frac{(\Sigma_{11})_{ij}}{\sqrt{(\Sigma_{11})_{ii}(\Sigma_{11})_{jj}}} = \rho(\underline{x}_i, \underline{x}_j), \text{ de } \underline{\text{enkelvoudige}} \text{ correlatiecoëfficiënt tussen } \underline{x}_i \text{ en } \underline{x}_j.$$

$$\rho_{ij\cdot 2} := \frac{(\Sigma_{11\cdot 2})_{ij}}{\sqrt{(\Sigma_{11\cdot 2})_{ii}(\Sigma_{11\cdot 2})_{jj}}} = \rho(\underline{x}_i - \hat{\underline{x}}_i, \underline{x}_j - \hat{\underline{x}}_j), \text{ de } \underline{\text{partiële}} \text{ correlatiecoëfficiënt tussen } \underline{x}_i \text{ en } \underline{x}_j \in I, \text{ bij gegeven } x^{(2)}.$$

$$\rho_{i\cdot 2} := \frac{(\Sigma_{1\cdot 2})_{ii}}{\sqrt{(\Sigma_{11})_{ii}(\Sigma_{1\cdot 2})_{ii}}} = \rho(\underline{x}_i, \hat{\underline{x}}_i), \text{ de } \underline{\text{multiple}} \text{ correlatiecoëfficiënt tussen } \underline{x}_i \in I \text{ en } \underline{x}^{(2)}.$$

Laten we deze eens nader bekijken:

1) Stel $\underline{x}_i$ en $\underline{x}_j$ zijn ongecorreleerd met $\underline{x}^{(2)}$, dan is $\rho_{ij\cdot 2} = \rho_{ij}$ nl. dan zijn de i^e en j^e rij van Σ_{12} geheel nul, dus $(\Sigma_{1\cdot 2})_{ii} = (\Sigma_{1\cdot 2})_{jj} = (\Sigma_{1\cdot 2})_{ij} = 0$.

2) Evident is $\rho_{i\cdot 2} \geqslant \rho_{ik}$ voor $\underline{x}_k \in \underline{x}^{(2)}$.

3) Is $\underline{x}_i$ ongecorreleerd met $\underline{x}^{(2)}$, dan is $\rho_{i\cdot 2} = 0$.

4) $$\frac{\operatorname{var}(\underline{x}_i | x^{(2)})}{\operatorname{var}\underline{x}_i} = \frac{\|\underline{x}_i - \hat{\underline{x}}_i\|^2}{\|x_i\|^2} = \sin^2\varphi = 1 - \cos^2\varphi = 1 - \rho_{i\cdot 2}^2$$

oftewel $\operatorname{var}(\underline{x}_i | x^{(2)}) = (1 - \rho_{i\cdot 2}^2)\operatorname{var}\underline{x}_i \leqslant \operatorname{var}\underline{x}_i$.

Algemeen geldt: $\Sigma_{11\cdot 2} \leqslant \Sigma_{11}$.

We werken het een en ander eens uit voor $p_1 + p_2 = 3$. We noteren $\text{var}\, \underline{x}_i = \sigma_{ii}$; $\text{cov}(\underline{x}_i, \underline{x}_j) = \sigma_{ij}$. Stel $\underline{x}^{(1)} = (\underline{x}_1, \underline{x}_2)^t$ en $\underline{x}^{(2)} = \underline{x}_3$, dus $\underline{x} = (\underline{x}_1, \underline{x}_2, \underline{x}_3)^t$

$$\Sigma_{11} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} ; \quad \Sigma_{12} = \begin{pmatrix} \sigma_{13} \\ \sigma_{23} \end{pmatrix} ; \quad \Sigma_{22} = \sigma_{33} ; \quad \Sigma^2 = \left(\begin{array}{cc|c} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \hline \sigma_{31} & \sigma_{32} & \sigma_{33} \end{array} \right) ;$$

Dan is $\Sigma_{1\cdot 2} = \dfrac{1}{\sigma_{33}} \begin{pmatrix} \sigma_{13}^2 & \sigma_{13}\sigma_{23} \\ \sigma_{13}\sigma_{23} & \sigma_{23}^2 \end{pmatrix}$ en

$$\Sigma_{11\cdot 2} = \begin{pmatrix} \sigma_{11} - \sigma_{13}^2/\sigma_{33} & \sigma_{12} - \sigma_{13}\sigma_{23}/\sigma_{33} \\ \sigma_{21} - \sigma_{13}\sigma_{23}/\sigma_{33} & \sigma_{22} - \sigma_{23}^2/\sigma_{33} \end{pmatrix} .$$

Hieruit volgt: $\rho_{12\cdot 3} = \dfrac{\rho_{12} - \rho_{13}\rho_{23}}{\sqrt{(1-\rho_{13}^2)(1-\rho_{23}^2)}}$;

Analoog kunnen we afleiden: $\rho_{2\cdot 13}^2 = \dfrac{\rho_{12}^2 + \rho_{23}^2 - 2\rho_{13}\rho_{23}\rho_{12}}{1-\rho^2 13}$.

Beide formules zijn eenvoudig meetkundig af te leiden uit Fig. 1. Probeert U het eens! Er geldt nl.:

$$\rho_{12\cdot 3} = \cos\delta = \frac{\cos\gamma - \cos\alpha\cos\beta}{\sqrt{(1-\cos^2\alpha)(1-\cos^2\beta)}} \quad \text{en}$$

$$\rho_{2\ 13}^2 = \cos^2\varphi = \frac{\cos^2\gamma + \cos^2\beta - 2\cos\alpha\cos\beta\cos\gamma}{\sqrt{1-\cos^2\alpha}} .$$

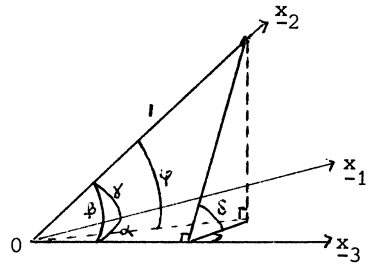


Fig. 1.

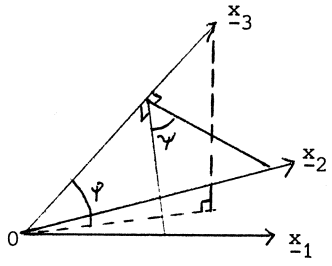


Fig. 2.

Roken en longkanker

Stel dat uit waarnemingen is gebleken dat het aantal sigaretten $\underset{-}{x}_1$ dat men per dag rookt hoog gecorreleerd is met de kans op longkanker ($\underset{-}{x}_2$). Is het sigaretten roken dan wel de boosdoener? M.a.w. als men minder gaat roken, neemt de kans op longkanker dan ook af of anders gezegd, is er een oorzakelijk verband?

Gemeten is dat ρ_{12} hoog is, dus de hoek tussen $\underset{-}{x}_1$ en $\underset{-}{x}_2$ is vrij klein. Onderstel nu dat er een andere stochast $\underset{-}{x}_3$ is (ons nog onbekend) die met beide hoog gecorreleerd is. Dat betekent dat $\rho_{3\cdot12}$ groot is oftewel dat $\underset{-}{x}_3$ dicht bij het vlak ligt opgespannen door $\underset{-}{x}_1$ en $\underset{-}{x}_2$ (nl. $\cos\varphi = \rho_{3\cdot12}$, zie Fig. 2).

Bekijken we nu eens $\rho_{12\cdot3}$ (= $\cos\psi$). Dan kan het zo zijn dat $\cos\psi = 0$ (of zelfs negatief) m.a.w. dat $\underset{-}{x}_1$ en $\underset{-}{x}_2$ ongecorreleerd zijn ná eliminatie van $\underset{-}{x}_3$. In dat geval heeft een wijziging in de rookgewoonten (mits men $\underset{-}{x}_3$ dus constant houdt) géén invloed op de kans op longkanker.

Dit voorbeeld toont dus aan dat, wil men een causaal verband tussen twee stochasten aantonen, de onbekende stochasten, waarvan $\underset{-}{x}_1$ en $\underset{-}{x}_2$ mogelijk afhankelijk zijn, gedurende de waarnemingen constant moeten zijn.

Is de proefopzet dus onnauwkeurig gebeurd, dan kan men gemakkelijk tot foute conclusies komen.

Litteratuur

1. Anderson,T.W., An introduction to multivariate statistical analysis,Wiley (1958).
2. Press, S.James, Applied Multivariate Analysis, Holt,Rinehart and Winston (1972).
3. Zurmühl R., Matrizen und ihre anwendungen, teil 1, Springer Verlag, Berlin (1984).

LINEAIR PROGRAMMEREN, IN HET BIJZONDER TRANSPORTPROBLEMEN

M. Kindt
Vakgroep onderzoek wiskundeonderwijs
& onderwijs computercentrum
Rijksuniversiteit Utrecht

1. Een transportvraagstuk 117

2. Het transportprobleem in algemene vorm 124

3. De hoekpuntenstelling 129

4. Hoekpunt-algoritmen 135

5. Een test voor het optimale hoekpunt 141

6. Naar een optimaal hoekpunt 145

7. Toepassingen 151

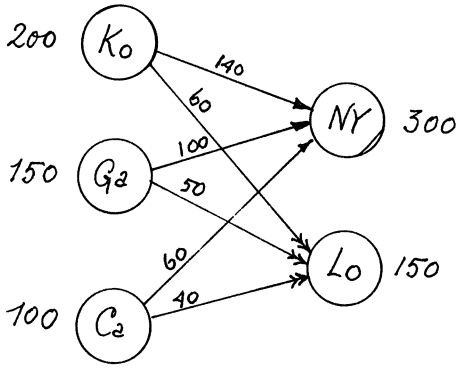
Literatuur 162

LINEAIR PROGRAMMEREN, IN HET BIJZONDER TRANSPORTPROBLEMEN

§ 1. Een transport vraagstuk

'Een oliemaatschappij heeft een voorraad van 200.000 barrels in Koeweit, 150.000 in Galveston en 100.000 in Caracas.
Een klant in New York heeft 300.000 barrels besteld. Een tweede klant in Londen wil de overige 150.000 barrels afnemen.'

Aldus begint een eindexamenopgave v.w.o.wiskunde A in 1983.
Een *transportprobleem*, waarbij het er om gaat een verstandig schema op te stellen voor het vervoer van barrels naar klanten.
Zo'n schema is bijvoorbeeld:



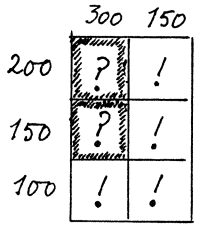
figuur 1

Of in matrixvorm:

	NY	Lo	
Ko	140	60	200
G2	100	50	150
C2	60	40	100
	300	150	

Er zijn zeer veel verschillende transportverdelingen mogelijk bij bovenstaande gegevens; zelfs als men ervan uitgaat dat de barrels per gehele

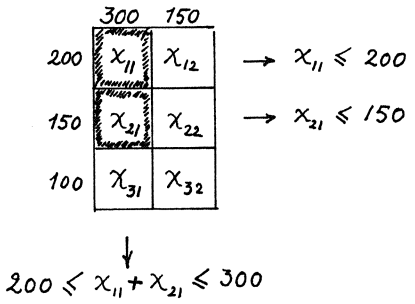
duizendtallen worden verscheept, zijn er om en nabij 10000 toelaatbare schema's!
Bij vulling van twee cellen in twee verschillende rijen van de matrix, ligt het schema onwrikbaar vast vanwege de knellende voorwaarde omtrent de totalen in rijen en kolommen.



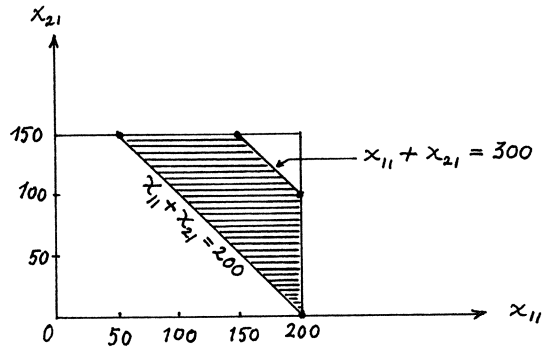
Die getallen, zeg x_{11} en x_{21}, mogen niet al te groot zijn.
Drie beperkingen die direct in het oog springen, zijn :

$$x_{11} \leq 200,\ x_{21} \leq 150 \text{ en } x_{11} + x_{21} \leq 300.$$

Een vierde beperking, $x_{11} + x_{21} \geq 200$ volgt uit het feit dat x_{31} ten hoogste 100 kan zijn

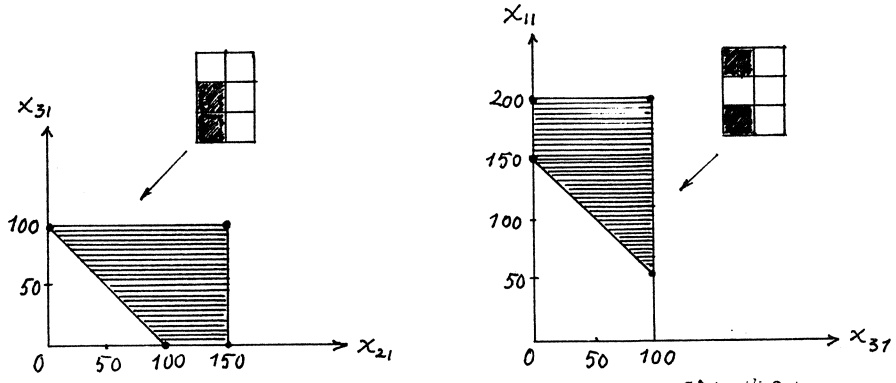


Het feit dat er precies twee (betrekkelijk) vrije keuzen zijn bij het opstellen van een transportschema, maakt nu dat elk schema zich kan laten vertegenwoordigen door een punt in het coördinatenvlak.
De verzameling toelaatbare schema's kan dan als volgt worden weergegeven:



figuur 2

Bij een andere keuze van de twee bepalende cellen in de transportmatrix, kan een ander plaatje ontstaan.
Voor de leraar die het werk moet corrigeren is het aardig om te weten dat er twaalf mogelijkheden zijn om de twee *beslissingsvariabelen* te kiezen. Het aantal verschillende plaatjes valt mee; op spiegeling of rotatie na zijn er drie gedaanten van het *toelaatbare gebied*.
Behalve de hiervoor getekende nog:



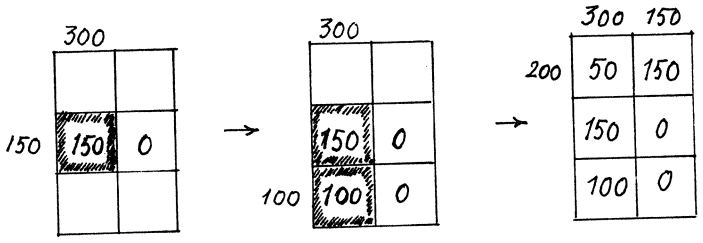
figuur 3

Merk op dat de drie gebieden dezelfde oppervlakte bezitten; de oppervlakte is namelijk een maat van het aantal toelaatbare schema's.

Het woord beslissingsvariabele is al gevallen, maar er valt nog weinig te beslissen over het te kiezen transportschema als de vervoersprijzen niet bekend zijn.
Op het betreffende wiskunde A examen was nog de volgende kostenmatrix (in dollarcenten per barrel) gegeven:

	NY	Lo
Ko	38	35
Gr	10	22
Ca	18	25

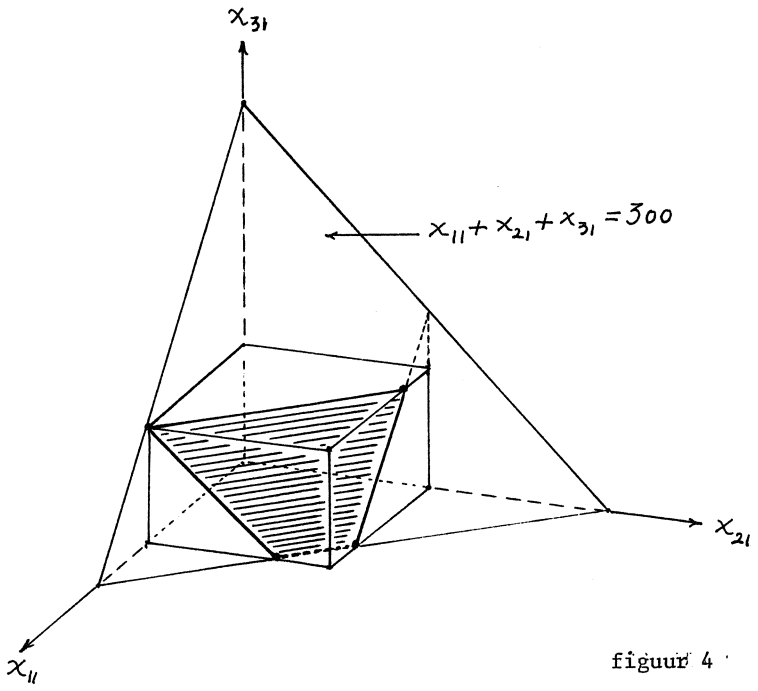
Natuurlijk met de vraag naar dàt transportschema, waarbij de totale kosten minimaal zijn.
Als je zo'n vraagstuk met gezond verstand wilt aanpakken, lijkt het voor de hand te liggen om zoveel mogelijk barrels tegen een zo laag mogelijke prijs te vervoeren.
Dit levert dan op:



Inderdaad blijkt dit het goedkoopste schema te zijn, maar zoals later zal blijken, is dit een kwestie van geluk.
Een waterdichte methode levert een plaatje van het toelaatbare gebied (figuur 2). Het totale kostenbedrag is, aangenomen dat de vervoerskosten evenredig oplopen met het aantal barrels, een eerste graadsfunctie van x_{11} en x_{21}. Die functie, waarvan het domein een vierhoek is, neemt zijn extreme waarden aan in een hoekpunt. Deze zogenaamde *hoekpuntenstelling* is meetkundig in te zien door aan een kaartje van iso-kostenlijnen te denken of aan een grafiek in de ruimte, boven het domein.
Berekening van de kosten in elk van de vier hoekpunten of een schets van de iso-kostenlijnen, leert dat (50, 150) het optimale punt is, inderdaad de representant van het voorbedachte schema.

Deze aanpak van het probleem is niet alleenzaligmakend. Men zou ook met drie variabelen, zeg x_{11}, x_{21} en x_{31}, kunnen werken met als beperkingen: $0 \leq x_{11} \leq 200$, $0 \leq x_{21} \leq 150$, $0 \leq x_{31} \leq 100$ en $x_{11} + x_{21} + x_{31} = 300$.
Zo wordt een toelaatbaar gebied bepaald in $\mathbb{R}^3$ dat meetkundig kan worden

opgevat als de doorsnede van een blok en een vlak (figuur 4).
Algebraïsch eleganter, maar meetkundig wat lastiger



figuur 4

Ook nu gaat het om het minimaliseren van eerste graadsfunctie (van x_{11}, x_{21} en x_{31}) op een vierhoekig domein en biedt de hoekpuntenstelling uitkomst.

Van lieverlee zijn we in hogere sferen terechtgekomen en waarom dan niet meteen doorgestoten naar de 6e dimensie?
Het transportschema kent immers zes open plaatsen, en elke schema is dus op te vatten als een punt in $\mathbb{R}^6$.
De beperkende voorwaarden hebben, afgezien van het niet-negatief zijn van de variabelen x_{ij} ($i = 1, 2, 3$; $j = 1, 2$) de gedaante van een stelsel *vergelijkingen*:

$$\begin{cases} x_{11} + x_{12} = 200 \\ x_{21} + x_{22} = 150 \\ x_{31} + x_{32} = 100 \\ x_{11} + x_{21} + x_{31} = 300 \\ x_{12} + x_{22} + x_{32} = 150 \end{cases}$$

en dat is meteen het voordeel van deze aanpak.
In matrixvorm ziet het stelsel er zó uit:

$$\left[\begin{array}{cccccc|c} 1 & 1 & 0 & 0 & 0 & 0 & 200 \\ 0 & 0 & 1 & 1 & 0 & 0 & 150 \\ 0 & 0 & 0 & 0 & 1 & 1 & 100 \\ 1 & 0 & 1 & 0 & 1 & 0 & 300 \\ 0 & 1 & 0 & 1 & 0 & 1 & 150 \end{array}\right]$$

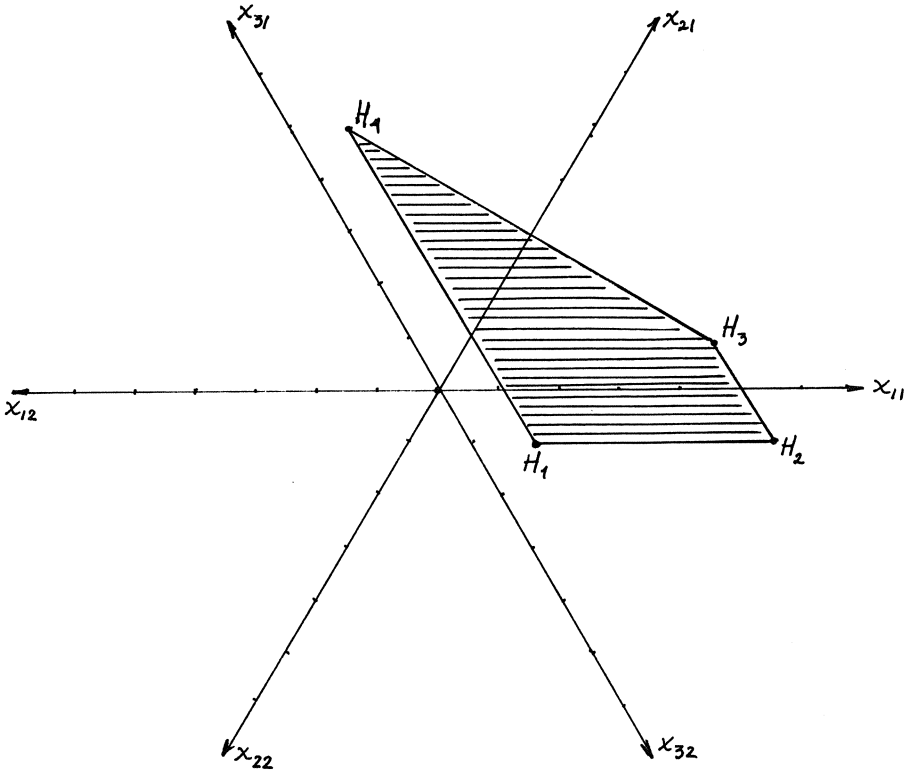
Gemakkelijk is na te gaan dat de oplossingsruimte van het stelsel 2-dimensionaal is, een vlak in $\mathbb{R}^6$.
Een parametervoorstelling van dit vlak is:

$$\begin{pmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{pmatrix} = \begin{pmatrix} 0 \\ 200 \\ 0 \\ 150 \\ 300 \\ -200 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

Dit vlak in $\mathbb{R}^6$ moet nu worden gesneden met het eerste '64-tant' van het 6-dimensionaal assenstelsel; en zo ontstaat een veelhoek als domein.
Elk hoekpunt van die veelhoek bevat tenminste twee nullen in zijn coördinatenrijtje.
Kies twee van de zes x_{ij}'s nul en de overige vier zijn bepaald! Omdat noch x_{11} en x_{12}, noch x_{21} en x_{22}, noch x_{31} en x_{32} tegelijk nul kunnen zijn, is de keuze niet geheel vrij. Er blijven $\binom{6}{3} - 3 = 12$ mogelijkheden over, die echter niet allemaal hoekpunten opleveren, vanwege het negatief zijn van een der niet-nullen.
Uiteindelijk blijken er vierhoekpunten te zijn (niet zo erg verrassend), die als matrix genoteerd er zó uit ziet:

$$H_1 = \begin{array}{|c|c|} \hline 200 & 0 \\ \hline 0 & 150 \\ \hline 100 & 0 \\ \hline \end{array} \qquad H_2 = \begin{array}{|c|c|} \hline 200 & 0 \\ \hline 100 & 50 \\ \hline 0 & 100 \\ \hline \end{array} \qquad H_3 = \begin{array}{|c|c|} \hline 150 & 50 \\ \hline 150 & 0 \\ \hline 0 & 100 \\ \hline \end{array} \qquad H_4 = \begin{array}{|c|c|} \hline 50 & 150 \\ \hline 150 & 0 \\ \hline 100 & 0 \\ \hline \end{array}$$

Het 6-dimensionale plaatje in projectie:



figuur 5

Ook nu blijkt een van de hoekpunten (H4) het goedkoopste transportschema op te leveren.

Het graantransportprobleem is aldus van diverse kanten belicht en leidt, meetkundig gezien, steeds tot het bepalen van de hoekpunten van een vlakke vierhoek (trapezium).
De examenopgave van destijds besluit met een vraag over acht depots en zes afnemers, gegeven het feit dat de totale vraag precies is afgestemd op de totale voorraad.
De oplossingsmethode met vergelijkingen voert dan meteen tot een deelverzameling van $\mathbb{R}^{48}$: een 35-dimensionaal *polytoop*. Dit domein heeft weliswaar eindig veel hoekpunten, namelijk punten met 35 (eventueel meer) nullen in een coördinatenmatrix van 48, maar dat zijn er duidelijk veel te veel om

op te noemen.
De klassieke 'simplexmethode' (zie bijvoorbeeld [1]) werkt hier met een tableau van 48 bij 13; en hoewel er slechts een fractie van het aantal hoekpunten hoeft te worden aangedaan, kan de methode met deze aantallen al aardig tijdrovend zijn.
Het specifieke karakter van de beperkende voorwaarden bij het transportprobleem biedt echter de mogelijkheid tot een sneller algoritme.

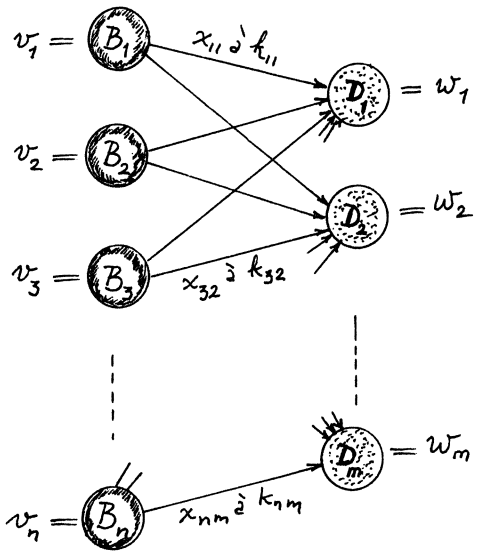
Opgave 1: Ga na dat de oppervlakte van de vierhoek $H_1H_2H_3H_4$ in $\mathbb{R}^6$ gelijk is aan $2000\sqrt{3}$, dus $2\sqrt{3}$ keer zo groot is als de oppervlakte van het domein in figuur 2. Hoe rijmt dat met het idee dat de oppervlakte van het domein maatgevend is van het aantal toelaatbare oplossingen?

§ 2. Het transportprobleem in algemene vorm

Het in § 1 besproken examenvraagstuk is een prototype van het transportprobleem.
Bij het transportprobleem is er sprake van
- een aantal *bronnen*, zeg $B_1, B_2, \ldots, B_n$
- een aantal *bestemmingen* of *doelen*: $D_1, D_2, \ldots, D_m$
- de *voorraad* V_i van een zeker artikel in B_i
- de gewenste hoeveelheid (*vraag*) W_j van dat artikel in D_j
- de *transportkosten* k_{ij}, nodig om een zekere hoeveelheid van dat artikel van B_i naar D_j te brengen.

De omvang van het transport van B_i naar D_j wordt voorgesteld door het niet-negatieve getal x_{ij}.

Schematisch:



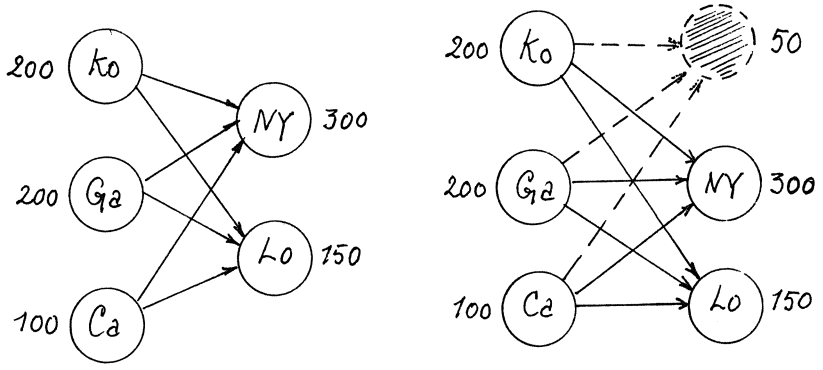
Voor i = 1, .., n; j = 1, .., m:

$x_{ij} \geq 0$
$k_{ij} > 0$
$V_i > 0$
$W_j > 0$

figuur 6

De aanname hierbij is dat de transportkosten van B_i naar D_j evenredig zijn met de omvang van het transport, zodat de totale transportprijs gelijk is aan $\sum_{i,j} k_{ij} x_{ij}$
Het doel van dit alles is om de x_{ij} zo te kiezen dat $\sum_{i,j} k_{ij} x_{ij}$ minimaal is.

Een eerste zorg is of voorraad en vraag precies op elkaar zijn afgestemd, met andere woorden of $\sum_i V_i$ en $\sum_j W_j$ aan elkaar gelijk zijn.
Is dat niet het geval, dan kan met een kuntsgreep die afstemming worden geforceerd.
Ter illustratie het voorbeeld van § 1, waarbij Galveston over 200 duizend barrels in plaats van 150 duizend beschikt, terwijl alle overige gegevens ongewijzigd zijn. Door aan de beide bestemmingen N en L een fictieve bestemming toe te voegen waarnaar het totale overschot (= 50 duizend barrels) in gedachten, en dus gratis, wordt vervoerd, wordt bereikt dat $\sum_i V_i = \sum_j W_j$.



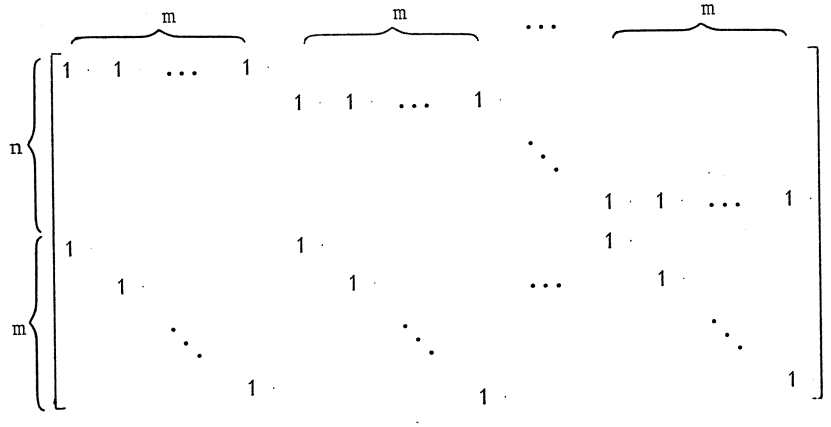
figuur 7

Overtreft de totale vraag de totale voorraad, dan kan een 'tekort-bron' als 'dummy' worden toegevoegd aan de verzameling bronnen. De kostprijs van het fictieve transport naar de beide bestemmingen wordt uiteraard als nihil beschouwd.

Aanname voor het vervolg: $\sum_i V_i = \sum_j W_j$

De n.m variabelen x_{ij} moeten voldoen aan de n + m vergelijkingen $\sum_j x_{ij} = V_i$ en $\sum_i x_{ij} = W_j$

De coefficientenmatrix van dit stelsel bevat louter enen en nullen en heeft de vorm



met nullen op de lege plekken.

De som van de eerste n rijen = de som van de laatste m rijen, dus de rang van de matrix is kleiner dan n + m.

Anderzijds ontstaat na weglating van bijvoorbeeld de eerste rij een onafhankelijk stelsel rijvectoren.
Conclusie: de rang van de matrix is n + m -1.
De voorwaarde $\sum_i V_i = \sum_j W_j$ garandeert dat de oplossingsruimte van het stelsel niet leeg is.
In verband met het vervolg is het handig om de punten van de nm-dimensionale reële ruimte waarin zich alles afspeelt, voor te stellen door n x m-matrices.
De ruimte van die matrices wordt genoteerd als $\mathbb{R}^{n,m}$.
Stel nu $X = (x_{ij})$ en $K = (k_{ij})$.
De totale kostprijs van het transportschema X wordt gevonden via een soort 'inwendig produkt' van K en X:

$$(K \cdot X) = \sum_{i,j} k_{ij}\, x_{ij} \quad *)$$

Het transportprobleem, wordt nu bondig geformuleerd als:

> Minimaliseer de lineaire funktie $X \to (K \cdot X)$ onder de voorwaarden: $\sum_j x_{ij} = V_i$, $\sum_i x_{ij} = W_j$ en $x_{ij} \geq 0$

De voorwaarden met betrekking tot de sommen van rijen en kolommen van X bepalen een lineaire variëteit L in $\mathbb{R}^{n,m}$. Uit wat hiervoor over de rang van de coëfficiëntenmatrix van het stelsel voorwaarden is gezegd, volgt dat de dimensie van L gelijk is aan d = nm - n - m + 1.
L is voor te stellen door een parametervoorstelling:

$$X = S + \lambda_1 R_1 + \lambda_2 R_2 + \ldots + \lambda_d R_d$$

($S, R_1, R_2, \ldots, R_d \in \mathbb{R}^{n,m}$; $R_1, R_2, \ldots, R_d$ lineair onafhankelijk)

De doorsnede T van L met $\{X \in \mathbb{R}^{n,m} | x_{ij} \geq 0\}$ is een polytoop begrensd door zijruimten van dimensie d-1. Een hoekpunt van het toelaatbare gebied T wordt gerepresenteerd door een matrix met tenminste d nullen, of, wat op hetzelfde neerkomt met ten hoogste n + m - 1 positieve cellen.

*) Deze vorm kan ook met het 'gewone' matrix produkt worden beschreven:

$$(K \cdot X) = \text{Spoor } (K^t X)$$

Uit de onafhankelijkheid van $R_1, R_2, \ldots, R_d$ volgt dat er ten hoogste één punt van het toelaatbare gebied is, met nullen in d uitgekozen cellen.

Bekijk als voorbeeld opgave 2 examen wiskunde A 1986, waarin sprake is van vier filialen van een warenhuis die videorecorders van twee depots betrekken.
De numerieke gegevens zijn:

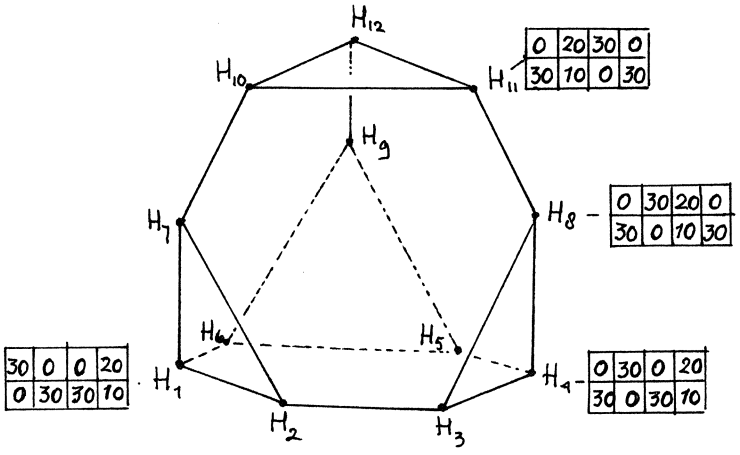
$K =$

12	15	15	9
15	17	12	10

$X =$

30	30	30	30	
?	?	?	?	50
?	?	?	?	70

Uit n = 2, m = 4 volgt: dimensie T = 8 − 2 − 4 + 1 = 3. Elk hoekpunt van T bevat dus ten minste drie nullen. Een klein beetje combinatoriek leert dat er $\binom{4}{2} \cdot \binom{2}{1} = 12$ hoekpunten zijn. Een projectieplaatje van T:



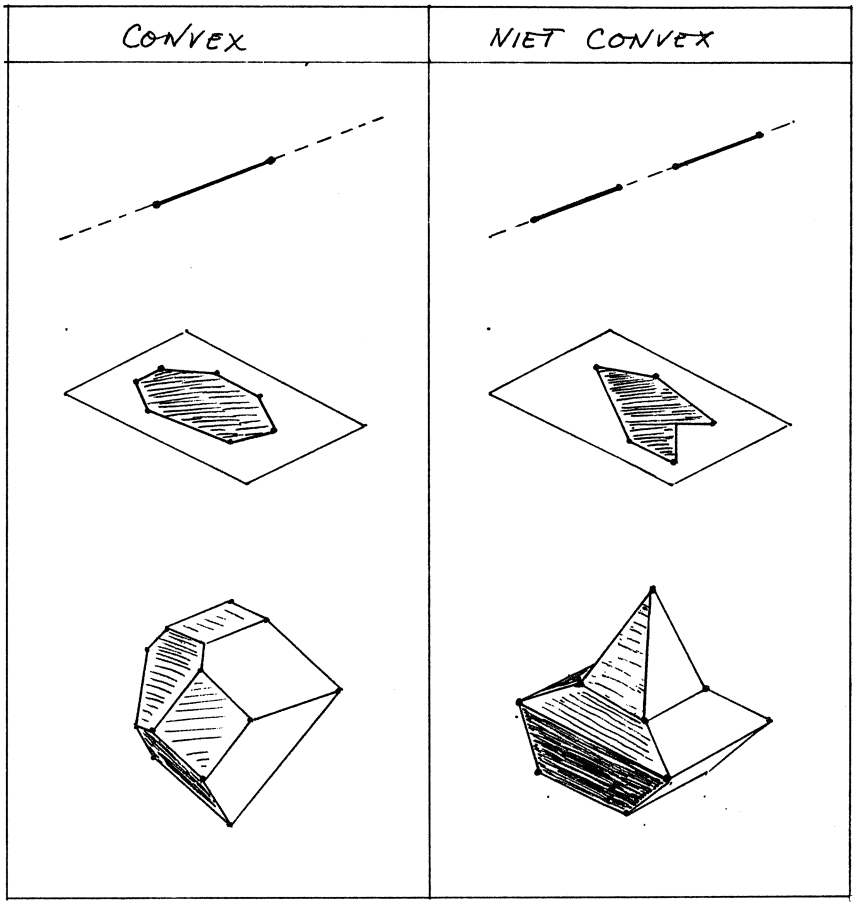
figuur 8

Opgave 2: Ga na dat verwisseling van twee (ongelijke) kolommen in de matrix van een hoekpunt van figuur 8 leidt naar een van de drie buurpunten van dat hoekpunt.

Opgave 3: De dimensie nm − n − n + 1 van het toelaatbare gebied T bij een transportprobleem met n bronnen en m bestemmingen laat zich rechtstreeks afleiden uit de graaf (figuur 6). Hoe?

§ 3. De hoekpuntenstelling

Een belangrijk gegeven bij lineaire programmeringsproblemen in het algemeen, en transportproblemen in het bijzonder, is het feit dat het toelaatbare gebied T *convex* is.

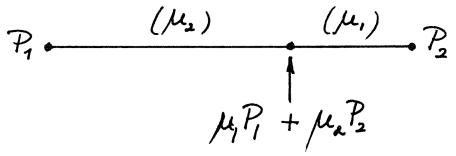


Definitie:

Een deelverzameling C van de n-dimensionale ruimte $\mathbb{R}^n$ *is convex, als met elk tweetal punten* P_1, P_2 *van C ook alle punten* $\mu_1 P_1 + \mu_2 P_2$ *met* $\mu_1 \geq 0$ *en* $\mu_2 \geq 0$ *en* $\mu_1 + \mu_2 = 1$ *tot C behoren.*

Meetkundig gezegd: met P_1 en P_2 behoort het lijnsegment P_1P_2 tot C.
De lineaire combinatie $\mu_1 P_1 + \mu_2 P_2$ ($\mu_1 \geqq 0$, $\mu_2 \geqq 0$, $\mu_1 + \mu_2 = 1$) wordt

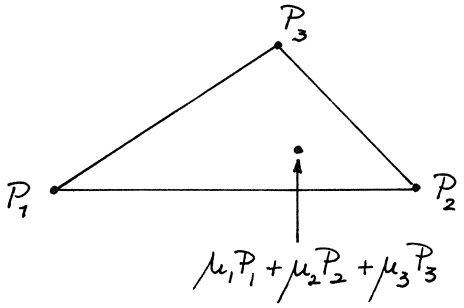
een *gewogen gemiddelde* van P_1 en P_2 genoemd of een *lineair convexe combinatie* (l.c.c) van P_1 en P_2



figuur 10

Als in P_1 en P_2 massa's μ_1 en μ_2 worden geplaatst, is $\mu_1 \cdot P_1 + \mu_2 \cdot P_2$ het zwaartepunt van het stelsel van 2 massa's.
Analoog kan men spreken van een gewogen gemiddelde van k punten ($k > 2$) in C; als C convex is behoort dat punt ook tot C.

Opgave 4: De 'gewichten' μ_1 en μ_2 in de l.c.c. $\mu_1 \cdot P_1 + \mu_2 \cdot P_2$ hebben een meetkundige betekenis (zie figuur 10). Wat is de meetkundige betekenis van de gewichten μ_1, μ_2, μ_3 in de l.c.c. $\mu_1 \cdot P_1 + \mu_2 \cdot P_2 + \mu_3 \cdot P_3$?

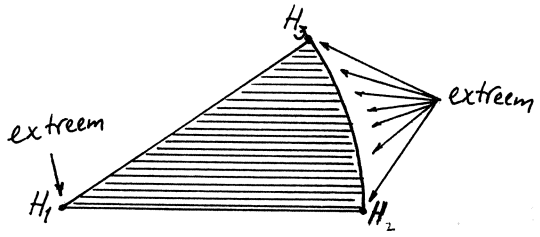


figuur 11

Opgave 5: Bewijs dat voor $k \geq 3$ geldt:
Als $P_1, P_2, \ldots, P_k$ punten zijn van een convexgebied C, dan is elke l.c.c. van $P_1, P_2, \ldots, P_k$ een punt van C.

In de voorgaande paragraaf is wat lichtzinnig over de hoekpunten van het toelaatbare gebied T gesproken, maar nu is de tijd gekomen om wat zwaarder op de hand te zijn.
Een kenmerk van een hoekpunt van een convex polytoop is dat het een *extreem* punt van dat polytoop is, dat wil zeggen dat het niet tussen twee

andere punten van het gebied ligt. Ook bij niet-polytopische convexe gebieden kan men van extreme punten spreken (zie figuur 12)



figuur 12

Een paar andere manieren om de extremiteit van een punt H in een convex gebied C te definiëren:
- er zijn geen punten P_1, P_2 in C, verschillend van H, zodat H het midden is van P_1 P_2;
- elk lijnsegment in C dat H bevat, heeft H als randpunt;
- $C \setminus \{H\}$ is convex.

Merk op dat de definitie van extreme punten beperkt is tot convexe gebieden. Bij een niet-convex polytoop kan men dus wel van hoekpunt, maar niet van extreem punt spreken. Overigens kan een hoekpunt van zo'n polytoop wel of niet de extremiteits-eigenschap hebben.

Bij het transportprobleem is het toelaatbare gebied T een *convex* polytoop in $\mathbb{R}^{n,m}$. T is bovendien *begrensd* en *gesloten*.
De convexiteit van T volgt uit:
- een lineaire variëteit is convex;
- het gebied $\{X \in \mathbb{R}^{n,m} | x_{ij} \geq 0\}$ is convex;
- de doorsnede van twee convexe gebieden is convex.

De begrensdheid en de geslotenheid volgt eveneens direct uit de toelaatbaarheidsvoorwaarden.

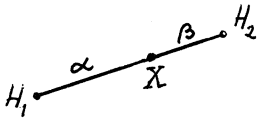
Let nu op de punten van T met tenminste d (= nm - n - m + 1) nullen, dus ten hoogste n + m - 1 positieve coördinaten.
Dat dit extreme punten zijn van T volgt gemakkelijk.
Immers: als H zo'n punt is en $H = \frac{1}{2}P + \frac{1}{2}Q$ (met P, Q in T), dan betekent dit dat P en Q elk zeker ook nullen hebben in dezelfde cellen als H, want negatieve coördinaten zijn verboden in T. Al eerder is vastgesteld dat er niet meer dan één punt in T is met nullen op d voorgeschreven plaatsen, dus

P = Q = H.
De extreme punten (hoekpunten) van T worden in het vervolg H_1, H_2, ..., H_k genoemd.
Er geldt nu:

T is het convexe omhulsel van H_1, H_2, ..., H_k

dat wil zeggen
(1) Elke 1.c.c. van H_1, H_2, ..., H_k behoort tot T
(2) Elk punt van T is 1.c.c. van H_1, H_2, ..., H_k

Bewijs:
(1) volgt onmiddellijk uit de convexiteit van T en uit opgave 5.
(2) wordt aangetoond met volledige inductie naar de dimensie d van T:
d = 1



figuur 13

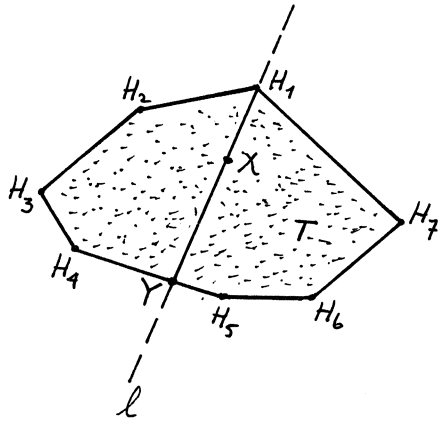
Een begrensde, gesloten en convexe 1-dimensionale verzameling kan niet anders dan een lijnsegment zijn.
Is X een punt van het segment H_1 H_2 met afstanden α en β tot H_1 en H_2, dan geldt:

$$X = \frac{\beta}{\alpha + \beta} H_1 + \frac{\alpha}{\alpha + \beta} H_2$$

en (2) is bewezen.

Inductieveronderstelling: (2) geldt als dim T = d - 1
Stel nu dim T = d en X in T.
Als $X = H_1$, dan $X = 1 \cdot H_1 + 0 \cdot H_2 + \ldots + 0 \cdot H_k$
Als $X \neq H_1$, dan bepalen X en H_1 een rechte lijn l.
$l \cap T$ in convex, gesloten en begrensd, kortom een lijnsegment met randpunten op de rand van T.
Het ene randpunt van $l \cap T$ is H_1, immers H_1 is een extreem punt. Het andere randpunt Y ligt in een d - 1-dimensionaal randpolytoop van T en is krachtens de inductieveronderstelling een 1.c.c. van de hoekpunten van dat rand-

polytoop, ofwel van een deelcollectie van $\{H_1, H_2, \ldots, H_k\}$.



figuur 14

Dus: $Y = \alpha_2 \cdot H_2 + \ldots + \alpha_k \cdot H_k$ $(\alpha_2 \geq 0, \ldots, \alpha_k \geq 0, \alpha_2 + \ldots + \alpha_k = 1)$
Omdat X tussen H_1 en Y ligt volgt:

$$X = \lambda \cdot H_1 + \mu \cdot Y \quad (\lambda \geq 0, \mu \geq 0, \lambda + \mu = 1)$$
$$= \lambda \cdot H_1 + \mu\, \alpha_2 \cdot H_2 + \ldots + \mu\, \alpha_k \cdot H_k$$

Dat dit een l.c.c. is volgt uit:

$$\lambda \geq 0,\ \mu\, \alpha_2 \geq 0, \ldots, \mu\, \alpha_k \geq 0 \text{ en}$$
$$\lambda + \mu\, \alpha_2 + \ldots + \mu\, \alpha_k = \lambda + \mu\, (\alpha_2 + \ldots \alpha_k) = \lambda + \mu \cdot 1 = 1$$

Opmerking:
De hier bewezen stelling is een bijzonder geval van de stelling van Krein-Milman, die zegt dat het convex omhulsel van de verzameling extreme punten van een compact convex gebied C samenvalt met C.
Uir deze stelling volgt direct de belangrijke hoekpuntenstelling voor transportproblemen:

Een lineaire functie f van het toelaatbare gebied T naar $\mathbb{R}$ heeft als bereik een gesloten en begrensd interval in $\mathbb{R}$,
Minimum en maximum van f worden aangenomen in een hoekpunt van T.

Bewijs:

Stel $m = \text{Min}\{f(H_1), \ldots, f(H_k)\}$

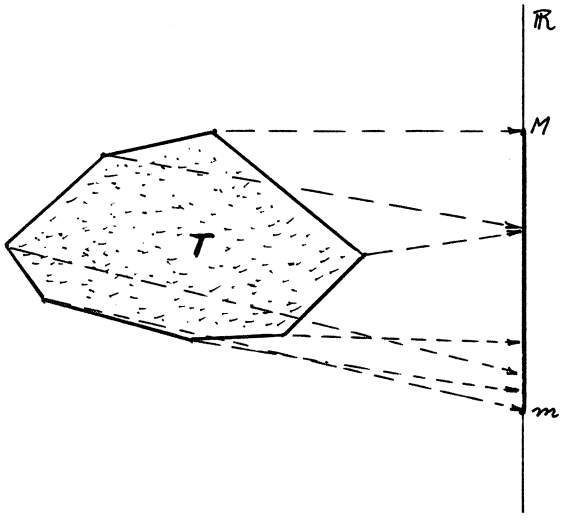
$M = \text{Max}\{f(H_1), \ldots, f(H_k)\}$

Voor elke X in T geldt:

$X = \mu_1 \cdot H_1 + \ldots + \mu_k \cdot H_k$ $(\mu_1 \geq 0, \ldots, \mu_k \geq 0, \mu_1 + \ldots + \mu_k = 1)$ en

$f(X) = \mu_1 \cdot f(H_1) + \ldots + \mu_k \cdot f(H_k)$.

Het bereik van f is dus de verzameling gewogen gemiddelden van de waarden $f(H_1), \ldots, f(H_k)$ en dat is precies het inteval [m, M].



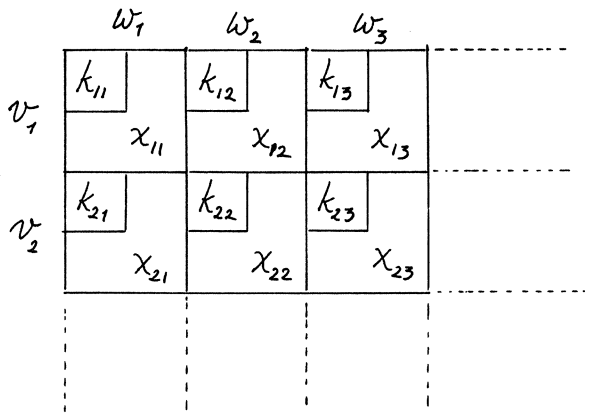
Opgave 6: Het convex omhulsel van 3 punten in $\mathbb{R}^2$ die niet collineair zijn (4 punten in $\mathbb{R}^3$ die niet copleniair zijn) is een driehoek (viervlak), het inwendige meegerekend.

Het convex omhulsel van n + 1 punten in $\mathbb{R}^n$, zeg $P_1, \ldots, P_{n+1}$, die niet in één hypervlak liggen, is een gebied dat wordt begrensd door een *simplex*.

Bewijs dat elk punt van een simplex-gebied met hoekpunten $P_1, \ldots, P_{n+1}$ een *uniek bepaalde* l.c.c. van $P_1, \ldots, P_{n+1}$ is.

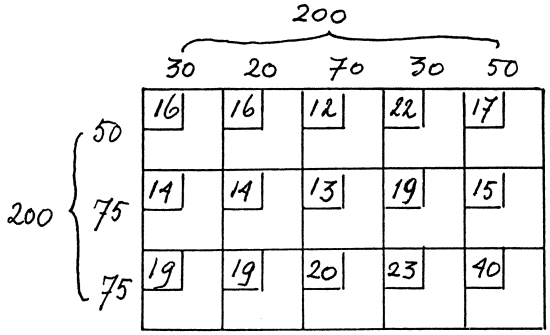
§ 4. Hoekpunt-algoritmen

Bij het zoeken naar een optimale oplossing van het transportprobleem komen slechts de hoekpunten van T in aanmerking.
Er bestaan bestaan diverse snelle algoritmen om een hoekpunt van T te vinden. Vervolgens kan met een tamelijk eenvoudige procedure worden nagegaan of dit hoekpunt optimaal is. Als dat niet het geval is, en de kans daarop is zeer groot, kan er vanuit het al gevonden hoekpunt een nieuw hoekpunt worden gevonden met lagere kosten.
Daarvan wordt dan weer nagegaan of het optimaal is. Zo niet, dan weer een volgend hoekpunt bepaald.
Kortom een iteratief proces, dat in een eindig aantal stappen naar het gezochte optimale punt leidt. Bij dit proces wordt vaak gewerkt met een *dubbel-matrix*, waarvan de cellen zijn gevuld of dienen te worden gevuld met de kostprijzen k_{ij} en de transporthoeveelheden x_{ij}. Aan de buitenkant worden de getallen V_i en W_j vermeld



De eerste stap van het optimaliseringsproces is het vinden van een eerste hoekpunt, een startpunt.
De eenvoudigste methode om zo'n startpunt te vinden is de zogenaamde *Noord-West-hoekregel*

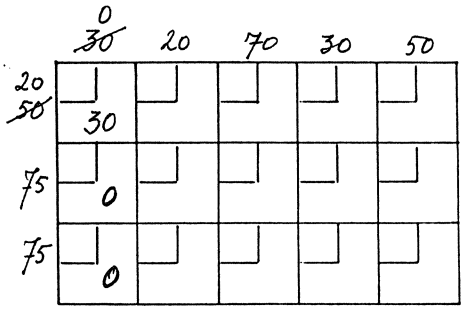
Bekijk onderstaand transportprobleem met n = 3, m = 5:



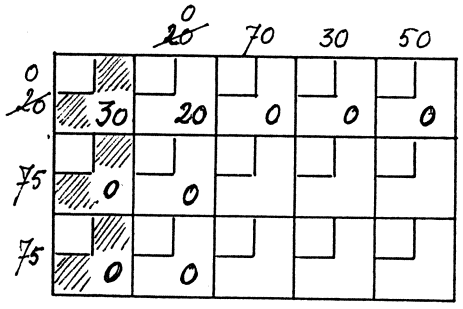
Het idee van de NW-hoekregel is als volgt:

- vul de cel in de NW-hoek (dus cel 1,1) met het minimum van de getallen V_1 en W_1; in het voorbeeld: 30;
- de overige celinhouden in de 1e rij of 1e kolom (of eventueel beiden) liggen dan vast; in het voorbeeld moet de eerste *kolom* vanwege het kolomtotaal 30 met nullen worden gevuld;
- er blijft nu een kleinere matrix over (in het voorbeeld een 4 x 3-matrix) waarvan het randgetal bij 1e rij of 1e kolom wordt verminderd met de vastgelegde transporthoeveelheid in de NW-cel;
- op de kleinere matrix wordt hetzelfde principe toegepast: in de NW-cel het grootst mogelijk transport en de rest van rij of kolom vullen met nullen.

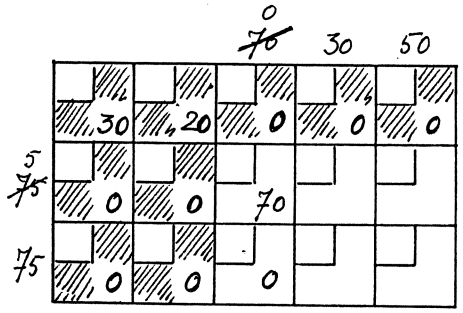
Na een eindig aantal stappen levert dit een element van T op met tenminste 8 nullen, kortom een hoekpunt.



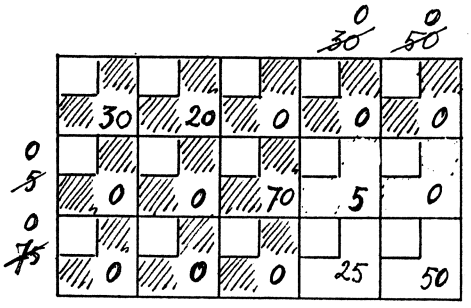
Stap 1
1e kolom gevuld



Stap 2
1e rij èn 2e kolom gevuld



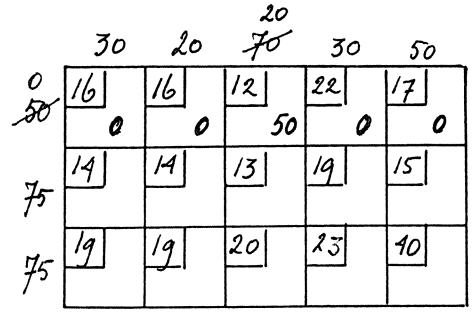
Stap 3
3e kolom gevuld



Stap 4
2e rij gevuld en de laatste rij ligt vast

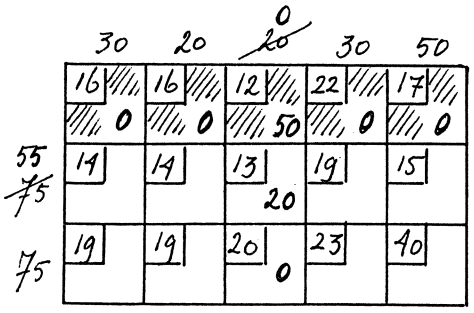
Het gevonden hoekpunt (met één extra nul!) levert een kostprijs op van:
30•16 + 20•16 + 70•13 + 10•5 + 25•23 + 50•40 = 4380
Bij de NW-hoekmethode is onderweg niet op de prijzen gelet.
Een meer prijsbewuste strategie is de volgende:

- zoek de laagste prijs in het tableau en maak de cel met de laagste prijs zo vol als mogelijk.
- na completering van rij of kolom (al naar gelang welk randgetal het kleinst is) blijft er een deelmatrix over, waarin opnieuw de laagste prijs wordt gezocht, enz.

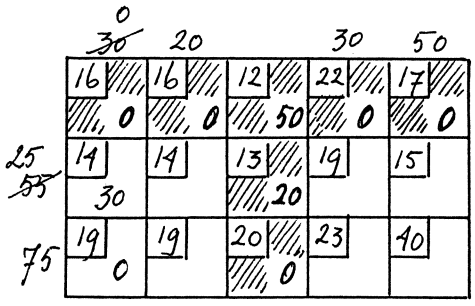


Stap 1

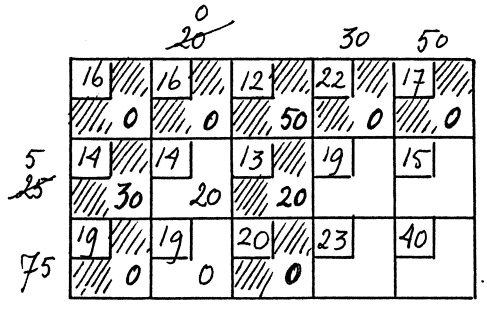
(12 is de laagste prijs)



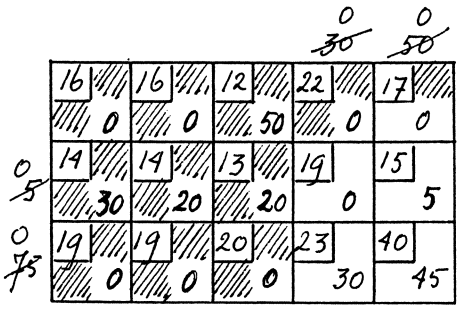
Stap 2



Stap 3



Stap 4

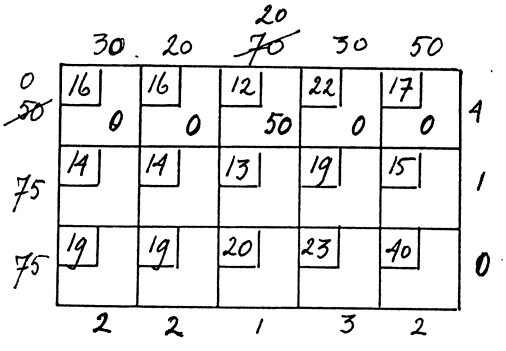


Stap 5

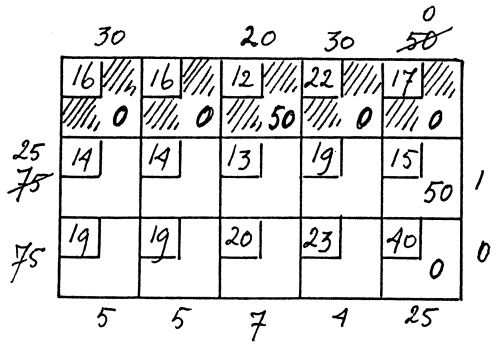
Inderdaad levert dit een goedkoper hoekpunt op met als totale prijs:
50•12 + 30•14 + 20•14 + 20•13 + 5•15 + 30•23 + 45•40 = 4125

Deze, toch wat naieve poging om het minimum te vinden, is niet de beste startmethode. Wat geraffineerder is de zogenaamde approximatie-methode van Vogel.
Het idee hierbij is:

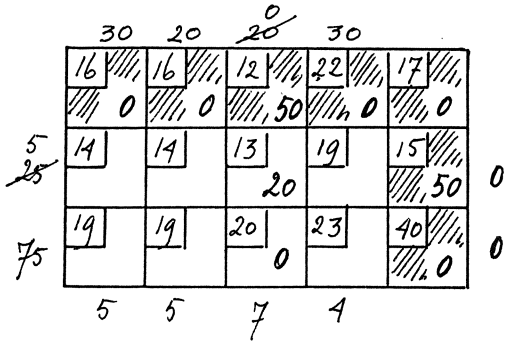
- bereken van elke rij èn kolom het *verschil* van de laagste twee voorkomende prijzen;
- zoek die rij of kolom waarvan dit verschil het grootst is; vul daarna de cel met de laagste prijs in die rij of kolom (met andere woorden: zoek de rij of kolom waarin de keuze van de laagste prijs het grootste voordeel oplevert);
- herhaal dit procede voor de resterende matrix met aangepaste randgetallen en aangepaste verschillen.



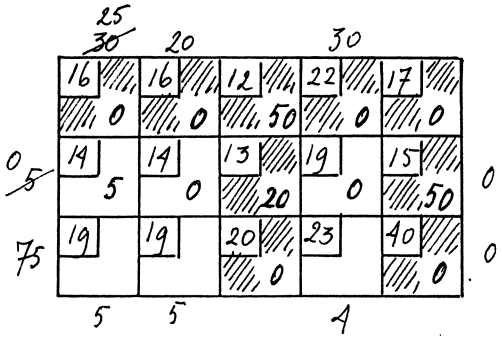
Stap 1



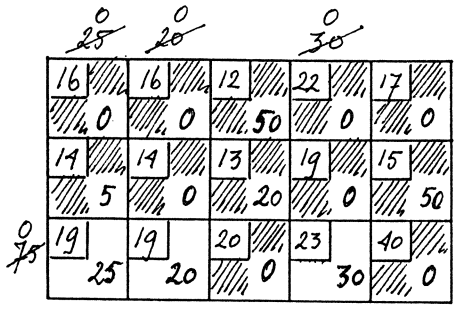
Stap 2



Stap 3



Stap 4



Stap 5

De totale prijs is nu:
$50 \cdot 12 + 5 \cdot 14 + 20 \cdot 13 + 50 \cdot 15 + 25 \cdot 19 + 20 \cdot 19 + 30 \cdot 23 = 3225$
Een aanzienlijke verbetering ten opzichte van de beide vorige situaties.

Opgave 7: Pas de drie hoekpuntalgoritmen toe op het examenvraagstuk van 1986. Ga na of een (of meer) van die methoden de minimale kostprijs oplevert.

§ 5. Een test voor het optimale hoekpunt

De approxmatie-methode van Vogel levert in het voorbeeld van § 4 een relatief lage kostprijs op.
De vraag is nu of die prijs het absolute minimum is.
In deze § wordt een zogenaamde *minimum-test* behandeld. Daarbij wordt uitgegaan van het volgende eenvoudige principe:

> *Is de transportprijs alleen afhankelijk van de bron, dus onafhankelijk van de bestemming, dan is de totale kostprijs constant.*
> *Hetzelfde is het geval als de transportprijs onafhankelijk is van de bron.*

In matrixtaal:
Als de rijen (of de kolommen) van een kostenmatrix constant zijn, is de totale kostprijs constant.
Stel nu:

$$A = \begin{bmatrix} a_1 & a_1 & \cdots & a_1 \\ a_2 & a_2 & \cdots & a_2 \\ \vdots & & & \\ a_n & a_n & \cdots & a_n \end{bmatrix} \quad , \quad B = \begin{bmatrix} b_1 & b_2 & \cdots & b_m \\ b_1 & b_2 & \cdots & b_m \\ \vdots & & & \\ b_1 & b_2 & \cdots & b_m \end{bmatrix}$$

beide in $\mathbb{R}^{n,m}$
De sommatrix $N = A + B$, opgevat als kostenmatrix, levert dan een constante totale kostprijs op:
$(N \cdot X) = (A \cdot X) + (B \cdot X) = \sum_i a_i V_i + \sum_j b_j W_j$

Anders gezeg:
bij een transportprobleem, waarbij $k_{ij} = a_i + b_j$ (i = 1, .., n; j = 1, .., m) is elk transportschema optimaal.
De kans dat een transportprobleem een dergelijke kostenmatrix heeft, is weliswaar niet bijster groot, maar toch blijkt bovengenoemde bewering van groot belang te zijn.
We zullen een matrix N van het type $(a_i + b_j)$ namelijk als een soort kortingsmatrix gebruiken met betrekking tot een gegeven kostenmatrix K.
Uit het voorgaande volgt dat de kostenmatrices K en K - N dezelfde optimale punten van T opleveren, immers (K•X) en ((K - N)•X) verschillen een constante.

In het uitgewerkte voorbeeld van § 4 werd met de Vogel-methode het hoekpunt H gevonden:

$H =$

0	0	50	0	0
5	0	20	0	50
25	20	0	30	0

bij een kostenmatrix

$K =$

16	16	12	22	17
14	14	13	19	15
19	19	20	23	40

De opzet is nu om de matrix N zo te bepalen dat in de cellen van K - N nullen komen op dié plaatsen die overeenkomen met cellen in H waar juist geen nullen staan:

$K-N=$

?	?	0	?	?
0	?	0	?	0
0	0	?	0	?

hetgeen betekent: ((K - N)•H) = 0

Als nu de andere cellen van K - N *positief of nul zijn geldt voor elke* $X \in T$: $((K - N) \cdot X) \geq 0$ *en is* H *een hoekpunt met minimale kostprijs.*
De minimale kostprijs is dan: $(K \cdot H) = (N \cdot H)$

Bevat K - N daarentegen negatieve cellen dan is het de moeite waard te zoeken naar een hoekpunt met lagere kostprijs.

Hoe nu N te vinden bij gegeven H?
In bovenstaande situatie geldt blijkbaar:

$N =$

?	?	12	?	?
14	?	13	?	15
19	19	?	23	?

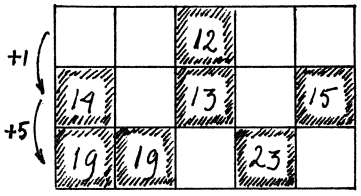
De getallen in de cellen zijn sommen $a_i + b_j$ ($i = 1, 2, 3$; $j = 1, 2, .., 5$)
De getallen a_i zijn op een constante c na bepaald.
Als a_i en b_j voldoen, voldoen $a_i - c$ en $b_j + c$ ook!
Dat betekent dat een der a_i (of een der b_j) willekeurig kan worden gekozen, waaruit de andere a_i en b_j dan volgen.
Kies bijvoorbeeld $a_1 = 0$. Er komt dan:

	+ 13	13	12	17	19
$a_1 = 0$	?	?	12	?	?
1	14	?	13	?	15
6	19	19	?	23	?

$\longrightarrow N =$

13	13	12	17	19
14	14	13	18	15
19	19	18	23	20

Merk op dat elke rij (kolom) van N uit elke andere rij(kolom) verkregen kan worden door optelling met een constante. Dat maakt dat de matrix N bepaald kan worden zonder gebruik te maken van de randgetallen a_i en b_j.



$\longrightarrow N =$

13	13	12	17	14
14	14	13	18	15
19	19	18	23	20

De 'gekorte-kostenmatrix' wordt:

$K - N =$

3	3	0	5	3
0	0	**0**	1	0
0	0	2	0	20

Omdat K - N geen negatieve getallen bevat is H een optimaal hoekpunt; de minimale kostprijs is: (K•H) = (N•H) = 3225.

Er doet zich een complicatie voor als het gevonden hoekpunt meer dan d nullen, dus minder dan n + m - 1 positieve cellen heeft. De matrix N is namelijk bepaald door precies n + m - 1 cellen, waarbij nog gegeven is dat elke rij en elke kolom tenminste één positief getal bevat (hetgeen uiteraard bij elementen van T het geval is).

In het geval van te veel nullen in de matrix H moet onderscheid worden gemaakt tussen de 'extra-nullen' en de nullen die het gevolg zijn van het gehanteerde systeem om tot een hoekpunt te komen.

Dit kan het eenvoudigst worden duidelijk gemaakt aan de hand van een voorbeeld.

Met de NW-hoekregel (zie § 4) werd gevonden

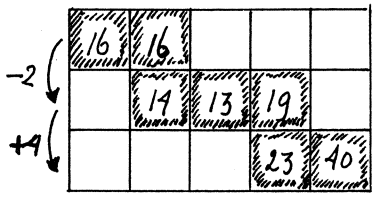
$H =$

30	20	0	0	0
0	0*	70	5	0
0	0	**0**	25	50

Met het vullen van de cel 2,1 raakten zowel de eerste rij als de tweede kolom vol. Een van de twee 'buurnullen' van 20 kan als extra nul worden beschouwd; hier is de nul in 2,2 als extra gemerkt.

Bij het opstellen van de matrix N wordt de gemerkte nul als positief meegeteld.

Dus:



$\rightarrow N =$

16	16	15	21	38
14	14	13	19	36
18	18	17	23	40

Gevolg:

$K - N =$

0	0	-3	1	-21
0	0	0	0	-21
1	1	3	0	0

Omdat H geen optimaal hoekpunt is bevat K - N negatieve cellen.

Opgave 8: Bepaal bij elk van de hoekpunten van figuur 5 de bijbehorende K - N en ga nadat H_4 inderdaad het optimale hoekpunt is.

§ 6. Naar een optimaal hoekpunt

Als bij een zeker hoekpunt van T de bijbehorende matrix K - N negatieve cellen heeft, leeft de verwachting dat het transportschema dat met dat hoekpunt correspondeert, verbeterd kan worden.
De vraag is natuurlijk: hoe?
Als voorbeeld kiezen we het hoekpunt bij het voorbeeld van § 4 dat is gevonden via het principe: vul steeds de cel met de laagste kosten.
Het resultaat was:

$H_1 =$

0	0	50	0	0
30	20	20	0	5
0	0	0	30	45

De bijbehorende 'gekorte-kostenmatrix' is $K - N_1$.

$K - N_1 =$

3	3	0	25	3
0	0	0	21	0
-20	-20	-18	0	0

De vorm $((K - N_1) \cdot X)$ bevat drie negatieve coëfficiënten;

$((K - N_1) \cdot X) = 3x_{11} + 3x_{12} + 25x_{14} + 3x_{15} + 21x_{25} - 20x_{31} - 20x_{32} - 18x_{33}$

Het doel is nu om een nieuw hoekpunt H_2 te vinden zodat $((K - N_1) \cdot H_2)$ negatief is en dus $(K \cdot H_2) < (K \cdot H_1)$.
Dit kan worden bereikt door in één van de drie cellen van H_1, die correspondeert met een negatief element van $K - N_1$, de nul te wijzigen in een positief getal. In dit geval kiezen we daartoe cel 3,1. Verhoging van die cel moet echter worden gecompenseerd door verlaging van een *positieve cel* in de eerste kolom; hier dus cel 2,1.

0	0	50	0	0
30$^-$	20	20	0	5
0$^+$	0	0	30	45

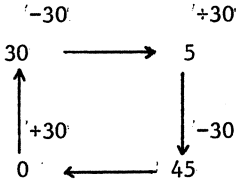
De verlaging van cel 2,1 moet worden gecompenseerd in de tweede rij, waarbij echter de cel met 0 buiten schot moeten blijven!
De verhoging in een van de cellen van rij 2 moet weer worden opgevangen met een verlaging in een passende kolom.
Zo ontstaat een *pad* van afwisselend '+' en '-' in de matrix, dat uiteindelijk na een even aantal stappen weer terugkeert bij de cel van waaruit is gestart.
In dit geval kan worden volstaan met een 4-stappen-pad
cel 3,1 → cel 2,1 → cel 2,5 → cel 3,5 → cel 3,1:

0	0	50	0	0
30$^-$	20	20	0	5$^+$
0$^+$	0	0	30	45$^-$

De waarde waarmee de cellen beurtelings worden verhoogd en verlaagd ligt vast door twee eisen:

- er mogen geen negatieve cellen in H ontstaan;
- de nul in startcel wordt maximaal opgehoogd om de kostprijs zoveel mogelijk te drukken.

Hier:



Zo wordt het nieuwe hoekpunt H_2 gevonden.

$H_2 =$

0	0	50	0	0
0	20	20	0	35
30	0	0	30	15

H_2 stemt in 7 nullen overeen met H_1 en is een buurpunt van H_1.
Bij H_2 worden N_2 en $K - N_2$ bepaald:

$N_2 =$

-7	13	12	-3	19	+1
-6	19	13	-2	15	
19	39	38	23	40	+25

$K - N_2 =$

23	3	0	25	3
20	0	0	21	0
0	-20	-18	0	0

Er is vooruitgang geboekt. Niet zo zeer omdat er nog maar twee negatieve cellen zijn, maar de totale kostprijs is gedaald met 30•20 = ƒ600,-

Merk op:

$(K \cdot H_1) - ((K - N_1) \cdot H_2) = (N_1 \cdot H_1) + (K \cdot H_2) - (N_1 \cdot H_2) = (K \cdot H_2) = -600$

Omdat de nieuwe kostenmatrix $K - N_2$ nog negatieve cellen heeft, wordt het proces herhaald.

$H_2 =$

0	0	5	0	0
0	20^-	20	0	35^+
30	0^+	0	30	15^-

$H_3 =$

0	0	5	0	0
0	5	20	0	50
30	15	0	30	0

H_3 is, zoals in § 5 al is aangetoond, een optimaal hoekpunt ($K - N_3$ bevat geen negatieve cellen). Bij de stap $H_2 \rightarrow H_3$ is de kostprijs gedaald met $15 \cdot 20 = 300$, zodat de minimale kostprijs $3525 - 300 = 3225$ bedraagt.

Nogmaals het proces, maar nu uitgaande van het minder gunstige hoekpunt:

$H_1 =$

30	20	0	0	0
0	0^*	70	5	0
0	0	0	25	50

Eerder werd gevonden:

$K - N_1 =$

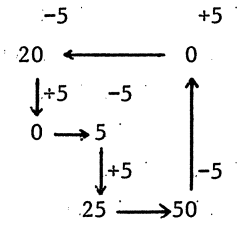
0	0	-3	1	-21
0	0	0	0	-21
1	1	3	0	0

Het pad van plussen en minnen (met start in cel 1,5) bestaat uit 6 stappen:

$H_1 =$

30	20$^-$	0	0	0$^+$
0	0^{+*}	70	5$^-$	0
0	0	0	25$^+$	50$^-$

De waarde waarmee de nul in cel 1,5 wordt opgehoogd is gelijk aan de kleinste cel die met een minteken is voorzien!
Zodat:



en

$H_2 =$

30	15	0	0	5
0	5	70	0	0
0	0	0	30	45

Merk op dat H_2 één nul minder heeft dan H_1, maar H_1 kon, met één nul extra een stootje velen! H_2 is met zijn acht nullen zeker een hoekpunt van T.
De verwerking van dit algorithme:
$H_1 \rightarrow K - N_1 \rightarrow H_2 \rightarrow K - N_2 \rightarrow \ldots$
kan uitstekend geschieden met behulp van een microcomputer. Hieronder het resultaat van de verwerking op een BBC.

```
TRANSPORTMATRIX (STAP 1 )          KOSTENMATRIX
 30   20    0    0    0  | 50      0    0   -3    1  -21
  0    *   70    5    0  | 75      0    0    0    0  -21
  0    0    0   25   50  | 75      1    1    3    0    0
-------------------------
 30   20   70   30   50

KOSTEN:  4380

TRANSPORTMATRIX (STAP 2 )          KOSTENMATRIX
 30   15    0    0    5  | 50      0    0   -3   22    0
  0    5   70    0    0  | 75      0    0    0   21    0
  0    0    0   30   45  | 75    -20  -20  -18    0    0
-------------------------
 30   20   70   30   50

KOSTEN:  4275

TRANSPORTMATRIX (STAP 3 )          KOSTENMATRIX
  0   15    0    0   35  | 50     20    0   -3   22    0
  0    5   70    0    0  | 75     20    0    0   21    0
 30    0    0   30   15  | 75      0  -20  -18    0    0
-------------------------
 30   20   70   30   50

KOSTEN:  3675

TRANSPORTMATRIX (STAP 4 )          KOSTENMATRIX
  0    *    0    0   50  | 50      0    0   -3    2    0
  0    5   70    0    0  | 75      0    0    0    1    0
 30   15    0   30    0  | 75      0    0    2    0   20
-------------------------
 30   20   70   30   50

KOSTEN:  3375

TRANSPORTMATRIX (STAP 5 )          KOSTENMATRIX
  0    0    *    0   50  | 50      3    3    0    5    0
  0    5   70    0    0  | 75      0    0    0    1   -3
 30   15    0   30    0  | 75      0    0    2    0   17
-------------------------
 30   20   70   30   50

KOSTEN:  3375

TRANSPORTMATRIX (STAP 6 )          KOSTENMATRIX
  0    0   50    0    0  | 50      3    3    0    5    3
  0    5   20    0   50  | 75      0    0    0    1    0
 30   15    0   30    0  | 75      0    0    2    0   20
-------------------------
 30   20   70   30   50

KOSTEN:  3225

OPTIMALE OPLOSSING GEVONDEN
```

Opgave 9: Wat opvalt is dat de overgang van H_4 naar H_5 geen verbetering oplevert.
Hoe is dat te verklaren?

Samengevat ziet het algoritme voor het vinden van een optimaal transportschema er zó uit:

I Zoek een eerste hoekpunt H_1 van het toelaatbare gebied.
(NW-hoekregel, minimumregel, Vogel, ...)
II Bepaal de bijbehorende N_1 en $K - N_1$. Ga na of $K - N_1$ negatieve cellen bevat. Is dat niet het geval, dan is H_1 een optimaal punt en $(K \cdot H_1) = (N_1 \cdot H_1)$ de minimale prijs.
III Bevat $K - N_1$ negatieve cellen, maak dan een pad van plussen en minnen in de matrix van H_1 en construeer daarmee een nieuw hoekpunt H_2.
IV Ga met $K - N_2$ na of H_2 optimaal is.
Zo niet, herhaal III met vervanging van H_1 door H_2.

Na een eindig aantal stappen wordt het gestelde doel bereikt. De existentie van een gesloten pad van plussen en minnen in de matrix, waarin buiten het startpunt geen 'echte' nullen zijn opgenomen, kan worden bewezen met wat grafentheorie (Zie [4]).
Bij ontaarde gevallen, dat wil zeggen, gevallen waarbij de extreme waarden in meer dan één hoekpunt wordt bereikt (en dan ook in randpunten die geen hoekpunt zijn!) kan het voorkomen dat $K - N$ negatieve cellen heeft, terwijl toch het minimum al is bereikt.

§ 7. Toepassingen

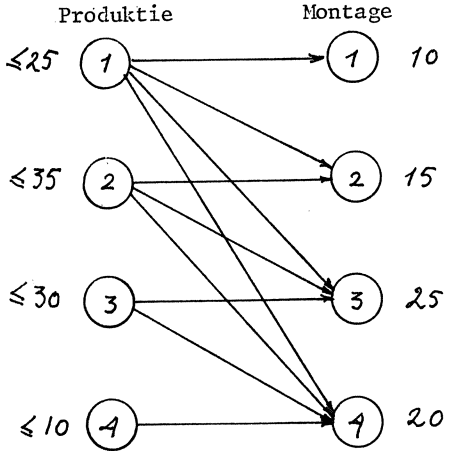
a. *Een produktieschema in de vliegtuigindustrie*
Het bedrijf 'Breeder' bouwt passagiersvliegtuigen voor diverse luchtvaartmaatschappijen, o.a. vliegtuigen van het type Relationship (R 25).
De laatste fase in het produktieproces van de R 25 is de fabricage en de montage van straalmotoren.
Breeder heeft contracten afgesloten, waarin zij zich verplicht een aantal vliegtuigen in de nabije toekomst af te leveren. In verband hiermee wordt de fabricage van staalmotoren voor de komende vier maanden nauwkeurig gepland.

Om de contractuele afspraken te kunnen nakomen moeten er in de komende maanden achtereenvolgens 10, 15, 25 en 20 motoren worden afgeleverd ter montage. De produktiemogelijkheden ten aanzien van de motoren zijn afhankelijk van andere activiteiten in de fabriek, zoals onderhoud en reparatie werkzaamheden. De maximale mogelijke produktie voor de komende vier maanden is resp. 25, 35, 30 en 10 motoren. De produktiekosten variëren per maand; in miljoenen guldens resp. 1,08; 1,11; 1,10; 1,13. Vanwege de variatie in produktiekosten kan het de moeite lonen motoren een of meer maanden eerder te produceren dan het schema aangeeft. Het nadeel hiervan is, dat motoren die te vroeg worden afgelevert, moeten worden opgeslagen tegen f15000,- per maand (inbegrepen het renteverlies op uitgegeven kapitaal).
De produktiemanager wil een schema opstellen voor de komende vier maanden, waarbij het totaal van de produktiekosten en voorraadkosten minimaal is.

Oplossing
De essentiele gegevens zijn:

Maand	Montage	Max.prod.	Prod.kosten	Voorraadkosten
1	10	25	1,08	-
2	15	35	1,11	0,015
3	25	30	1,10	0,015
4	20	10	1,13	0,015

Verrassenderwijs kan het probleem worden vertaald in een transportprobleem met, in eerste instantie, 4 bronnen (= 'produktiemaanden') en 4 bestemmingen (= 'montage maanden')



Omdat het totale produktiepotentieel (= 100 motoren) groter is dan het aantal benodigde motoren voor de montage, wordt een vijfde *dummy bestemming* D ingevoerd.
Er ontstaat zo een transportprobleem met n = 4, m = 5, waarbij x_{ij} het aantal afgeleverde straalmotoren is in maand i ter montage in maand j.
Omdat produktie achteraf niet mogelijk is geldt: $x_{ij} = 0$ als $i > j$ (in de graaf zijn de hiermee corresponderende puntenparen i, j niet verbonden).
De kosten k_{ij} zijn in de gevallen $i \leq j$ direct af te leiden uit de gegevens omtrent produktie en voorraadkosten.

$$K = \begin{bmatrix} 1 & 2 & 3 & 4 & D & \\ 1{,}080 & 1{,}095 & 1{,}110 & 1{,}125 & - & 1 \\ - & 1{,}110 & 1{,}125 & 1{,}140 & - & 2 \\ - & - & 1{,}100 & 1{,}115 & - & 3 \\ - & - & - & 1{,}130 & - & 4 \end{bmatrix}$$

In verband met het te volgen oplossingsprocedé moeten de lege plekken in K worden gevuld.
De 5e kolom (kosten voor een fictieve bestemming) moet uiteraard louter uit nullen bestaan.
Voor de driehoek onder de lijn i = j staat vast dat de corresponderende cellen in elk toelaatbaar schema X nullen zullen zijn. Daarom nemen we voor k_{21}, k_{31}, k_{32}, k_{41}, k_{42}, k_{43} een zeer groot willekeurig getal, genaamd M, zeer veel groter dan alle andere elementen van k. Daarmee wordt een oplossing geforceerd met nullen op de plaatsen die corresponderen met de M-cellen.
Deze truc wordt wel 'the big M'-methode genoemd.
Zo wordt de kostenmatrix:

$$K = \begin{bmatrix} 1{,}080 & 1{,}095 & 1{,}110 & 1{,}125 & 0 \\ M & 1{,}110 & 1{,}125 & 1{,}140 & 0 \\ M & M & 1{,}100 & 1{,}115 & 0 \\ M & M & M & 1{,}130 & 0 \end{bmatrix}$$

Startend met een hoekpunt met nullen op de onderdriehoek levert de microcomputer na 4 stappen het minimum 77,30 en een bijbehorend optimale schema

TRANSPORTMATRIX (STAP 1)

10	15	0	0	0	:	25
0	*	25	10	0	:	35
0	0	0	10	20	:	30
0	0	0	0	10	:	10
10	15	25	20	30		

KOSTEN: 77.9

TRANSPORTMATRIX (STAP 2)

10	15	0	0	0	:	25
0	*	25	0	10	:	35
0	0	0	20	10	:	30
0	0	0	0	10	:	10
10	15	25	20	30		

KOSTEN: 77.65

TRANSPORTMATRIX (STAP 3)

10	15	0	0	0	:	25
0	*	15	0	20	:	35
0	0	10	20	0	:	30
0	0	0	0	10	:	10
10	15	25	20	30		

KOSTEN: 77.4

TRANSPORTMATRIX (STAP 4)

10	15	0	0	0	:	25
0	*	5	0	30	:	35
0	0	20	10	0	:	30
0	0	0	10	0	:	10
10	15	25	20	30		

KOSTEN: 77.3

OPTIMALE OPLOSSING GEVONDEN

KOSTENMATRIX

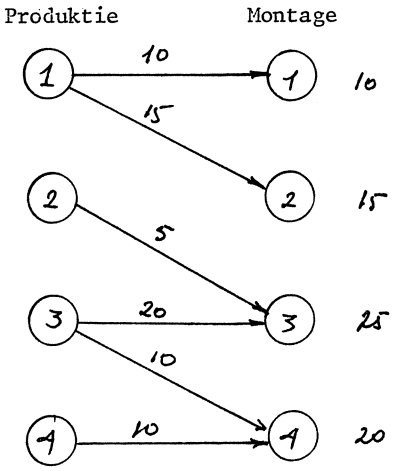
.000	.000	.000	.000	-.010
M	.000	.000	.000	-.025
M	M	.000	.000	.000
M	M	M	.015	.000

.000	.000	.000	.025	.015
M	.000	.000	.025	.000
M	M	-.025	.000	.000
M	M	M	.015	.000

.000	.000	.000	.000	.015
M	.000	.000	.000	.000
M	M	.000	.000	.025
M	M	M	-.010	.000

.000	.000	.000	.000	.015
M	.000	.000	.000	.000
M	M	.000	.000	.025
M	M	M	.000	.010

Vertaald in een graaf:



Opgave 10: Ga na dat dit transportprobleem verschillende optimale schema's heeft.

b. *Een toewijzingprobleem*

Er wordt een sportvijfkamp georganiseerd, waaraan een aantal dorpen deelnemen. Elk dorp vaardigt 5 atleten af, die elk zullen uitkomen op één van de onderdelen zwemmen, hardlopen, gewichtheffen, kanovaren en wielrennen. Een van de deelnemende dorpen heeft zich serieus op het treffen voorbereid. De vijf in aanmerking komende atleten hebben gedurende een trainingsperiode van een paar maanden regelmatig proeven van bekwaamheid in elk van de onderdelen moeten afleggen. Op basis van een waarderingssysteem per sport, is aan elk van de deelnemers een cijfer voor ieder onderdeel toegekend. Het resultaat is af te lezen in een tabel:

		zwemmen	hardlopen	gewichth.	kanovaren	wielrennen
deel-nemers	A	8	5	7	4	5
	B	9	7	8	3	6
	C	7	3	2	1	5
	D	1	6	9	7	7
	E	4	3	6	7	9
	F	3	4	7	6	3

De ploeg wordt zo sterk mogelijk geacht als de som van vijf waarderingscijfers op vijf verschillende onderdelen maximaal is.
Het probleem wordt als volgt vertaald in een transportprobleem: nummer de deelnemers en de sportonderdelen en stel:

x_{ij} = 1 als deelnemer i uitkomt op onderdeel j

x_{ij} = 0 als deelnemer i niet op onderdeel j uitkomt

Elke toelaatbare matrix $X = (x_{ij})$ is een matrix met rij- en kolomtotalen gelijk aan 1.
Is W de waarderingsmatrix, dan is het zaak om die X te vinden, waarvan (W•X) maximaal is. Het transportalgoritme, met K = - W, levert de oplossing:

TRANSPORTMATRIX (STAP 1)

*	1	*	Ø	*	¦	1
1	Ø	Ø	Ø	Ø	¦	1
Ø	*	Ø	1	Ø	¦	1
Ø	Ø	1	Ø	Ø	¦	1
Ø	Ø	Ø	Ø	1	¦	1
1	1	1	1	1		

KOSTENMATRIX

Ø	Ø	Ø	-1	Ø
Ø	-1	Ø	1	Ø
-1	Ø	3	Ø	-2
9	1	Ø	-2	Ø
8	6	5	Ø	Ø

KOSTEN: -33

TRANSPORTMATRIX (STAP 2)

*	1	*	Ø	*	¦	1
1	Ø	Ø	Ø	Ø	¦	1
Ø	Ø	Ø	1	*	¦	1
Ø	Ø	1	Ø	Ø	¦	1
Ø	Ø	Ø	Ø	1	¦	1
1	1	1	1	1		

KOSTENMATRIX

Ø	Ø	Ø	-3	Ø
Ø	-1	Ø	-1	Ø
1	2	5	Ø	Ø
9	1	Ø	-4	Ø
8	6	5	-2	Ø

KOSTEN: -33

TRANSPORTMATRIX (STAP 3)

*	1	*	Ø	Ø	¦	1
1	Ø	Ø	Ø	Ø	¦	1
Ø	Ø	Ø	1	*	¦	1
Ø	Ø	1	*	Ø	¦	1
Ø	Ø	Ø	Ø	1	¦	1
1	1	1	1	1		

KOSTENMATRIX

Ø	Ø	Ø	1	4
Ø	-1	Ø	3	4
-3	-2	1	Ø	Ø
9	1	Ø	Ø	4
4	2	1	-2	Ø

KOSTEN: -33

TRANSPORTMATRIX (STAP 4)

Ø	1	*	Ø	Ø	¦	1
1	Ø	Ø	Ø	Ø	¦	1
*	Ø	Ø	1	*	¦	1
Ø	Ø	1	*	Ø	¦	1
Ø	Ø	Ø	Ø	1	¦	1
1	1	1	1	1		

KOSTENMATRIX

3	Ø	Ø	1	4
Ø	-4	-3	Ø	1
Ø	-2	1	Ø	Ø
12	1	Ø	Ø	4
7	2	1	-2	Ø

KOSTEN: -33

TRANSPORTMATRIX (STAP 5)

Ø	*	1	Ø	Ø	¦	1
Ø	1	Ø	Ø	Ø	¦	1
1	Ø	Ø	*	*	¦	1
Ø	Ø	*	1	Ø	¦	1
Ø	Ø	Ø	Ø	1	¦	1
1	1	1	1	1		

KOSTENMATRIX

3	Ø	Ø	1	4
4	Ø	1	4	5
Ø	-2	1	Ø	Ø
12	1	Ø	Ø	4
7	2	1	-2	Ø

KOSTEN: -37

TRANSPORTMATRIX (STAP 6)

Ø	*	1	Ø	Ø	¦	1
Ø	1	Ø	Ø	Ø	¦	1
1	*	Ø	Ø	*	¦	1
Ø	Ø	*	1	Ø	¦	1
Ø	Ø	Ø	Ø	1	¦	1
1	1	1	1	1		

KOSTENMATRIX

1	Ø	Ø	1	2
2	Ø	1	4	3
Ø	Ø	3	2	Ø
10	1	Ø	Ø	2
7	4	3	Ø	Ø

KOSTEN: -37

OPTIMALE OPLOSSING GEVONDEN

Opgave 11: Er is nu sprake van zes kandidaten voor de vijfkamp met van elke kandidaat een lijstje van vijf waarderingscijfers. De opgave is opnieuw om een zo sterk mogelijke ploeg samen te stellen, waarbij dus een van de kandidaten moet afvallen.
Hoe kan dit probleem als transportprobleem worden gemodelleerd?

Opgave 12: Moeilijker ligt het als bij vijf deelnemers en vijf sportonderdelen wordt bepaald dat elke deelnemer op precies twee verschillende onderdelen moet uitkomen, en dat per onderdeel twee atleten per dorp

meedoen.
Waarom kan deze situatie niet eenvoudig met het transportalgoritme worden aangepakt?

Opmerking:
Een methode die hier met vrucht zou kunnen worden toegepast is de zogenaamde 'branch-and-bound' techniek, speciaal ontworpen voor zogenaamde geheeltallije programeringsproblemen. (zie [2])

c. *Waterdistributie*

De Metro Water Distribution (MWD) is een agentschap dat zorg draagt voor de watervoorziening van een groot district.
Omdat het district nogal droog is moet de MWD water invoeren vanuit een ander district. De bronnen van dit geïmporteerde water zijn de Colombo River, de Sacron River en de White River. De MWD koopt en verkoopt het water van deze bronnen aan gebruikers in het district met als voornaamste klanten de vier gemeenten Bardoo, Los Diablos, San Pedro en Turtletown . Het is mogelijk elk van de plaatsen te voorzien van water, afkomstig van elk van de drie rivieren, met dien verstande dat er geen faciliteiten zijn om Turtletown te voorzien van water uit de White River.
De geografische ligging van de vier plaatsen maakt dat de transportkosten per plaats/rivier nogal variëren.

Kosten ($ per 1000 m^3)

	Bardoo	Los Diablos	San Pedro	Turtletown	Voorraad
C.River	16	13	22	17	50
S.River	14	13	19	15	60
W.River	19	20	23	-	50
Minimale behoefte	30	70	0	10	↑ in 10^9 m^3
Vraag	50	70	30	∞	←

In de tabel zijn ook af te lezen de watervoorraden (per rivier) en de minimale behoefte alsmede de vraag (per plaats).
Zo wenst Bardoo $50 \cdot 10^9$ m^3 water, terwijl die plaats zou kunnen volstaan met $30 \cdot 10^9$ m^3 om in de eerste behoefte te voorzien. San Pedro beschikt zelf over een waterbron (getuige de nul in de kolom van SP) en Turtletown wil

zoveel water afnemen als mogelijk is (∞).
De prijs per 1000 m³ die de MWD berekent aan de vier gemeenten is voor elke gemeente hetzelfde.
De directie van de MWD heeft zich tot doel gesteld de complete beschikbare hoeveelheid water van de drie rivieren te betrekken en die zo te verdelen dat de kosten van de MWD zo laag mogelijk zijn. De MWD is bovendien verplicht tenminste aan de eerste behoefte van elke gemeente te voldoen.

Oplossing:
Dit is een transportprobleen met de drie rivieren als bronnen en de gemeenten als bestemmingen. Voordat het transportalgoritme kan worden toegepast, moet er eerst wat geschaafd worden aan de gegevens.
Allereerst de vraag (= ∞) van TT. Als aan de eerste behoeften van B, LD en SP wordt voldaan, blijft er voor TT ten hoogste $60 \cdot 10^9$ m³ water beschikbaar. De vraag ∞ wordt daarom vervangen door 60.
Let nu op de totale vraag (= 210) en het totale aanbod (= 160). Omdat de vraag het aanbod overtreft wordt een dummy-bron (D) toegevoegd.
Tenslotte het streepje in de gegeven tabel (geen water van WR naar TT); hier biedt een 'big M' uitkomt.
Het transportschema lijkt zo te worden:

	B	LD	SG	TT	
CR	16	13	22	17	50
SR	14	13	19	15	60
WR	19	20	23	M	50
D	0	0	0	0	50
	50	70	30	60	

Als op dit schema klakkeloos het transportalgoritme zou worden toegepast is er alle kans dat er een oplossing uitrolt waarbij niet aan de eerste behoefte van B of LD is voldaan.
Voor SP speelt dit niet (minimale behoefte = 0). Voor TT geldt dat de vraag (= 60) de dummy-voorraad (= 50) met 10 overtreft, zodat elke toelaatbare oplossing gegarandeerd aan de eerste behoefte (= 10) voldoet.
LD's eerste behoefte is weliswaar gelijk aan de vraag (= 70), maar het zou kunnen dat een oplossingschema water van de dummy-bron naar LD zou doen toevloeien. Dit kan worden voorkomen met een big M.
Nu nog B. Hier geldt dat van de vraag (= 50) niet meer dan 20 afkomstig mag zijn van de dummy-bron. Door B te splitsen in twee bestemmingen: B min en B extra kan een derde big M garant staan dat B tenminste $30 \cdot 10^9$ m^3 water krijgt toegewezen.
Kortom:

	B_{min}	B_{extra}	LD	SG	TT	
CR	16	16	13	22	17	50
SR	14	14	13	19	15	60
WR	19	19	20	23	M	50
D	M	0	M	0	0	50
	30	20	70	30	60	

De microcomputer doet de rest:

```
TRANSPORTMATRIX (STAP 1 )          KOSTENMATRIX
 30    0   20    0    0  ! 50      0    4    0   10    9
  0    0    0    *   60  ! 60     -9   -5   -7    0    0
  0    *   50    0    0  ! 50     -4    0    0    4    M
  0   20    0   30    0  ! 50      M    0    M    0    4
-------------------------
 30   20   70   30   60

KOSTEN: 2640

TRANSPORTMATRIX (STAP 2 )          KOSTENMATRIX
 30    0   20    0    0  ! 50      0    4    0   10    0
  *    0    0    0   60  ! 60      0    4    2    9    0
  0    *   50    0    0  ! 50     -4    0    0    4    M
  0   20    0   30    0  ! 50      M    0    M    0   -5
-------------------------
 30   20   70   30   60

KOSTEN: 2640

TRANSPORTMATRIX (STAP 3 )          KOSTENMATRIX
 10    0   40    0    0  ! 50      0    4    0    5    0
 20    0    0    0   40  ! 60      0    4    2    4    0
  0   20   30    0    0  ! 50     -4    0    0   -1    M
  0    0    0   30   20  ! 50      M    5    M    0    0
-------------------------
 30   20   70   30   60

KOSTEN: 2540

TRANSPORTMATRIX (STAP 4 )          KOSTENMATRIX
  0    0   50    0    0  ! 50      4    4    0    9    4
 20    0    0    0   40  ! 60      0    0   -2    4    0
 10   20   20    0    0  ! 50      0    0    0    3    M
  0    0    0   30   20  ! 50      M    1    M    0    0
-------------------------
 30   20   70   30   60

KOSTEN: 2500

TRANSPORTMATRIX (STAP 5 )          KOSTENMATRIX
  0    0   50    0    0  ! 50      4    4    0    7    2
  0    0   20    0   40  ! 60      2    2    0    4    0
 30   20    *    0    0  ! 50      0    0    0    1    M
  0    0    0   30   20  ! 50      M    3    M    0    0
-------------------------
 30   20   70   30   60

KOSTEN: 2460

OPTIMALE OPLOSSING GEVONDEN
```

Het starthoekpunt is hier betrekkelijk willekeurig gekozen met nullen op de M-plaatsen. De Vogel approximatie zou hier meteen het optimale punt hebben opgeleverd!

Literatuur

[1] Wagner, H.M., *Principles of Operations Research*, Londen 1972.
Een standaardwerk op dit gebied.

[2] Hillier, F.S. en G.J. Lieberman, *Introduction to Operations Research*, San Francisco 1974.
Een aantrekkelijk boek met vele uitgewerkte voorbeelden en aandacht voor modelvorming.

[3] Dantzig, G.B., *Linear Programming and Extensions*, Princeton 1963.
Het klassieke werk over lineair programmeren.

[4] Henderiks, R.E.D., *Lineaire Programmering*, Utrecht 1972.
Beknopte, sterk wiskundig gerichte behandeling van de simplexmethode.

[5] Lange, J.de en M.Kindt, *Lineair programmeren*, Culemborg 1986.
Leerlingtekst voor wiskunde A.

[6] Het gebruikte computerprogramma voor het transportalgoritme is ontworpen door Heleen Verhage, Utrecht 1986.

MC SYLLABI

1.1 F. Göbel, J. van de Lune. *Leergang besliskunde, deel 1: wiskundige basiskennis.* 1965.

1.2 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 2: kansberekening.* 1965.

1.3 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 3: statistiek.* 1966.

1.4 G. de Leve, W. Molenaar. *Leergang besliskunde, deel 4: Markovketens en wachttijden.* 1966.

1.5 J. Kriens, G. de Leve. *Leergang besliskunde, deel 5: inleiding tot de mathematische besliskunde.* 1966.

1.6a B. Dorhout, J. Kriens. *Leergang besliskunde, deel 6a: wiskundige programmering 1.* 1968.

1.6b B. Dorhout, J. Kriens, J.Th. van Lieshout. *Leergang besliskunde, deel 6b: wiskundige programmering 2.* 1977.

1.7a G. de Leve. *Leergang besliskunde, deel 7a: dynamische programmering 1.* 1968.

1.7b G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7b: dynamische programmering 2.* 1970.

1.7c G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7c: dynamische programmering 3.* 1971.

1.8 J. Kriens, F. Göbel, W. Molenaar. *Leergang besliskunde, deel 8: minimaxmethode, netwerkplanning, simulatie.* 1968.

2.1 G.J.R. Förch, P.J. van der Houwen, R.P. van de Riet. *Colloquium stabiliteit van differentieschema's, deel 1.* 1967.

2.2 L. Dekker, T.J. Dekker, P.J. van der Houwen, M.N. Spijker. *Colloquium stabiliteit van differentieschema's, deel 2.* 1968.

3.1 H.A. Lauwerier. *Randwaardeproblemen, deel 1.* 1967.

3.2 H.A. Lauwerier. *Randwaardeproblemen, deel 2.* 1968.

3.3 H.A. Lauwerier. *Randwaardeproblemen, deel 3.* 1968.

4 H.A. Lauwerier. *Representaties van groepen.* 1968.

5 J.H. van Lint, J.J. Seidel, P.C. Baayen. *Colloquium discrete wiskunde.* 1968.

6 K.K. Koksma. *Cursus ALGOL 60.* 1969.

7.1 *Colloquium moderne rekenmachines, deel 1.* 1969.

7.2 *Colloquium moderne rekenmachines, deel 2.* 1969.

8 H. Bavinck, J. Grasman. *Relaxatietrillingen.* 1969.

9.1 T.M.T. Coolen, G.J.R. Förch, E.M. de Jager, H.G.J. Pijls. *Colloquium elliptische differentiaalvergelijkingen, deel 1.* 1970.

9.2 W.P. van den Brink, T.M.T. Coolen, B. Dijkhuis, P.P.N. de Groen, P.J. van der Houwen, E.M. de Jager, N.M. Temme, R.J. de Vogelaere. *Colloquium elliptische differentiaalvergelijkingen, deel 2.* 1970.

10 J. Fabius, W.R. van Zwet. *Grondbegrippen van de waarschijnlijkheidsrekening.* 1970.

11 H. Bart, M.A. Kaashoek. H.G.J. Pijls, W.J. de Schipper, J. de Vries. *Colloquium halfalgebra's en positieve operatoren.* 1971.

12 T.J. Dekker. *Numerieke algebra.* 1971.

13 F.E.J. Kruseman Aretz. *Programmeren voor rekenautomaten; de MC ALGOL 60 vertaler voor de EL X8.* 1971.

14 H. Bavinck, W. Gautschi, G.M. Willems. *Colloquium approximatietheorie.* 1971.

15.1 T.J. Dekker, P.W. Hemker, P.J. van der Houwen. *Colloquium stijve differentiaalvergelijkingen, deel 1.* 1972.

15.2 P.A. Beentjes, K. Dekker, H.C. Hemker, S.P.N. van Kampen, G.M. Willems. *Colloquium stijve differentiaalvergelijkingen, deel 2.* 1973.

15.3 P.A. Beentjes, K. Dekker, P.W. Hemker, M. van Veldhuizen. *Colloquium stijve differentiaalvergelijkingen, deel 3.* 1975.

16.1 L. Geurts. *Cursus programmeren, deel 1: de elementen van het programmeren.* 1973.

16.2 L. Geurts. *Cursus programmeren, deel 2: de programmeertaal ALGOL 60.* 1973.

17.1 P.S. Stobbe. *Lineaire algebra, deel 1.* 1973.

17.2 P.S. Stobbe. *Lineaire algebra, deel 2.* 1973.

17.3 N.M. Temme. *Lineaire algebra, deel 3.* 1976.

18 F. van der Blij, H. Freudenthal, J.J. de Iongh, J.J. Seidel, A. van Wijngaarden. *Een kwart eeuw wiskunde 1946-1971, syllabus van de vakantiecursus 1971.* 1973.

19 A. Hordijk, R. Potharst, J.Th. Runnenburg. *Optimaal stoppen van Markovketens.* 1973.

20 T.M.T. Coolen, P.W. Hemker, P.J. van der Houwen, E. Slagt. *ALGOL 60 procedures voor begin- en randwaardeproblemen.* 1976.

21 J.W. de Bakker (red.). *Colloquium programmacorrectheid.* 1975.

22 R. Helmers, J. Oosterhoff, F.H. Ruymgaart, M.C.A. van Zuylen. *Asymptotische methoden in de toetsingstheorie; toepassingen van naburigheid.* 1976.

23.1 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 1.* 1976.

23.2 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 2.* 1977.

24.1 P.J. van der Houwen. *Numerieke integratie van differentiaalvergelijkingen, deel 1: eenstapsmethoden.* 1974.

25 *Colloquium structuur van programmeertalen.* 1976.

26.1 N.M. Temme (ed.). *Nonlinear analysis, volume 1.* 1976.

26.2 N.M. Temme (ed.). *Nonlinear analysis, volume 2.* 1976.

27 M. Bakker, P.W. Hemker, P.J. van der Houwen, S.J. Polak, M. van Veldhuizen. *Colloquium discretiseringsmethoden.* 1976.

28 O. Diekmann, N.M. Temme (eds.). *Nonlinear diffusion problems.* 1976.

29.1 J.C.P. Bus (red.). *Colloquium numerieke programmatuur, deel 1A, deel 1B.* 1976.

29.2 H.J.J. te Riele (red.). *Colloquium numerieke programmatuur, deel 2.* 1977.

30 J. Heering, P. Klint (red.). *Colloquium programmeeromgevingen.* 1983.

31 J.H. van Lint (red.). *Inleiding in de coderingstheorie.* 1976.

32 L. Geurts (red.). *Colloquium bedrijfsystemen.* 1976.

33 P.J. van der Houwen. *Berekening van waterstanden in zeeën en rivieren.* 1977.

34 J. Hemelrijk. *Oriënterende cursus mathematische statistiek.* 1977.

35 P.J.W. ten Hagen (red.). *Colloquium computer graphics.* 1978.

36 J.M. Aarts, J. de Vries. *Colloquium topologische dynamische systemen.* 1977.

37 J.C. van Vliet (red.). *Colloquium capita datastructuren.* 1978.

38.1 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part I.* 1979.

38.2 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part II.* 1979.

39 O.J. Vrieze, G.L. Wanrooy. *Colloquium stochastische spelen.* 1978.

40 J. van Tiel. *Convexe analyse.* 1979.

41 H.J.J. te Riele (ed.). *Colloquium numerical treatment of integral equations.* 1979.

42 J.C. van Vliet (red.). *Colloquium capita implementatie van programmeertalen.* 1980.

43 A.M. Cohen, H.A. Wilbrink. *Eindige groepen (een inleidende cursus).* 1980.

44 J.G. Verwer (ed.). *Colloquium numerical solution of partial differential equations.* 1980.

45 P. Klint (red.). *Colloquium hogere programmeertalen en computerarchitectuur.* 1980.

46.1 P.M.G. Apers (red.). *Colloquium databankorganisatie, deel 1.* 1981.

46.2 P.G.M. Apers (red.). *Colloquium databankorganisatie, deel 2.* 1981.

47.1 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60: general information and indices.* 1981.

47.2 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 1: elementary procedures; vol. 2: algebraic evaluations.* 1981.

47.3 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3A: linear algebra, part I.* 1981.

47.4 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3B: linear algebra, part II.* 1981.

47.5 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 4: analytical evaluations; vol. 5A: analytical problems, part I.* 1981.

47.6 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 5B: analytical problems, part II.* 1981.

47.7 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 6: special functions and constants; vol. 7: interpolation and approximation.* 1981.

48.1 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 1.* 1982.

48.2 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 2.* 1982.

49 T.H. Koornwinder (ed.). *The structure of real semisimple Lie groups.* 1982.

50 H. Nijmeijer. *Inleiding systeemtheorie.* 1982.

51 P.J. Hoogendoorn (red.). *Cursus cryptografie.* 1983.

CWI SYLLABI

1 Vacantiecursus 1984: *Hewet - plus wiskunde.* 1984.

2 E.M. de Jager, H.G.J. Pijls (eds.). *Proceedings Seminar 1981-1982. Mathematical structures in field theories.* 1984.

3 W.C.M. Kallenberg, et al. *Testing statistical hypotheses: worked solutions.* 1984.

4 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 1.* 1984.

5 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 2.* 1984.

6 P.J.M. Bongaarts, J.N. Buur, E.A. de Kerf, R. Martini, H.G.J. Pijls, J.W. de Roever. *Proceedings Seminar 1982-1983. Mathematical structures in field theories.* 1985.

7 Vacantiecurus 1985: *Variatierekening.* 1985.

8 G.M. Tuynman. *Proceedings Seminar 1983-1985. Mathematical structures in field theories, Vol.1 Geometric quantization.* 1985.

9 J. van Leeuwen, J.K. Lenstra (eds.). *Parallel computers and computations.* 1985.

10 Vacantiecursus 1986: *Matrices.* 1986.