TUUT VOOR CULTUURTECHNIEK EN WATERHUISHOUDING
NOTA No.147 d.d.16 oktober 1962

Een meetkundige toelichting op het oplossen van normaalvergelijkingen[1)]

Het schema van CHOLESKI en een toepassing bij de niet-lineaire vereffening

Ph.Th.Stol

[1)] Uitwerking van een inleiding over het oplossen van normaalvergelijkingen met behulp van computers gehouden voor het Colloquim Electronisch Rekenen te Wageningen, d.d.2 juli 1962.

165/1062/20

1. INLEIDING

De vergelijkingen waaruit in een regressieprobleem de regressie-coëfficiënten berekend worden heten de normaalvergelijkingen. Deze vergelijkingen hebben de eigenschap dat het schema ervan een symmetrische matrix vormt. Oplossingsmethoden voor deze vergelijkingen die gebruikmaken van de symmetrie leiden veelal tot overzichtelijke rekenschema's. In Nota 134 (Lineaire regressie) werd door KAMIL het rekenschema van CHOLESKI algebraïsch behandeld.

In het volgende zal het oplossen van normaalvergelijkingen meetkundig toegelicht worden waarbij in de eerste plaats het schema van CHOLESKI ter sprake komt. Tevens zal een eigenschap van dit schema behandeld worden die toepassing kan vinden bij het probleem van de niet-lineaire vereffening.

Uitgegaan zal worden van een systeem met 3 onafhankelijke variabelen. In de figuren zal met een dimensie minder volstaan worden. Toepassing op meer dimensies kan plaatsvinden door naar analogie, de verkregen resultaten uit te breiden.

2. HET OPSTELLEN VAN DE NORMAALVERGELIJKINGEN

Stel dat gevraagd wordt de constanten te bepalen in

$$a_1x + a_2y + a_3z = q \qquad (2.1)$$

waarin x, y, z en q vectoren (kolommen) van bijvoorbeeld waarnemingsuitkomsten voorstellen[1]. De vector q zal in het algemeen niet uit een lineaire combinatie volgens (2.1) verkregen kunnen worden tengevolge van toevallige afwijkingen. Wel zal gelden dat een vector q^* een lineaire combinatie is van de basisvectoren x, y en z zodat

$$q^* = a_1x + a_2y + a_3z \qquad (2.2)$$

en de vraag is (zie fig. 1) de vector $(q-q^*)$ zo klein mogelijk te maken. Aan deze vraag wordt voldaan indien simultaan geldt

$$(q-q^*) \perp x,\ y,\ z \qquad (2.3)$$

of in woorden: wanneer de verschilvector loodrecht op de basisvectoren staat. Wordt (2.2) gesubstitueerd in (2.3) dan vormen de uitgeschreven voorwaarden de normaalvergelijkingen (Nota 134 pag.5) waarvan het schema luidt

A				c
a_1	a_2	a_3	=	1
xx	xy	xz		xq
yx	yy	yz		yq
zx	zy	zz		zq

(2.4)

De matrix van het stelsel wordt A genoemd, de bekende termen worden door de vector c voorgesteld, het stelsel vergelijkingen moet naar a_1, a_2 en a_3 opgelost worden

---

[1] Kleine letters stellen in het volgende vectoren voor, hoofdletters stellen matrices voor. Kleine letters met een index hebben de betekenis van kentallen van een vector. De schaalfactoren h_i zijn scalaire grootheden.

De elementen van de matrix zijn geschreven als het inprodukt van de vectoren x, y, z en q, waarvoor xy = yx enz.

Wordt nog gedefinieerd de matrix X zodat

$$X = (x \quad y \quad z)$$

en

$$b = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \qquad \text{terwijl} \qquad c = \begin{pmatrix} xq \\ yq \\ zq \end{pmatrix}$$

dan kan (2.2) geschreven worden

$$q^* = Xb \tag{2.2'}$$

en dus

$$q - Xb \perp X \tag{2.3'}$$

zodat (2.4) de gedaante krijgt

$$^tXXb = {}^tXq \tag{2.4'}$$

of kortheidshalve

$$Ab = c \tag{2.5}$$

waarin A de matrix van het stelselvergelijkingen voorstelt waarvan de oplossing is

$$b = A^{-1}c \tag{2.6}$$

Om tot deze oplossing te komen moet dus de matrix A geïnverteerd worden.

Het stelselvergelijkingen (2.4) kan eenvoudiger opgelost worden indien de basis (x, y, z) vervangen wordt door een orthogonale basis (x, y_0, z_0) met de eigenschappen

$$\begin{aligned} y_0 &\perp x \\ z_0 &\perp y_0, x \end{aligned} \tag{2.7}$$

Dit houdt dus in dat $(q-q^*)$ nu minimaal wordt voor

$$(q-q^*) \perp x, y_0, z_0 \tag{2.8}$$

waaruit voor de normaalvergelijkingen volgt:

$$\begin{array}{cccccc} a_1 & a_2 & a_3 & = & & 1 \\ \hline xx & xy & xz & & & xq \\ 0 & y_0y & y_0z & & & y_0q \\ 0 & 0 & z_0z & & & z_0q \end{array} \tag{2.9}$$

De "back-solution" verloopt nu als volgt:

$$\begin{aligned} a_3 &= z_0q/z_0z \\ a_2 &= \left[\, y_0q - (y_0z)a_3 \,\right]/y_0y \\ a_1 &= \left[\, xq - (xz)a_3 - (xy)a_2 \,\right]/xx \end{aligned} \tag{2.10}$$

Onderwerp van de volgende paragrafen is nu de elementen van de matrix (2.9) uit te drukken in die van de matrix (2.4).

3. HET ORTHOGONALISEREN VAN DE BASISVECTOREN

Voor het orthogonaliseren van de basis (x, y, z) wordt uitgegaan van de vector x. Deze blijft onveranderd. Vervolgens wordt in het door x en y opgespannen vlak een lineaire combinatie gevormd zodanig dat (zie fig. 2)

$$y_o = (y + \alpha x) \perp x \qquad (3.1)$$

en dus

$$xy + \alpha xx = 0$$

waaruit volgt

$$\alpha = -\frac{xy}{xx}$$

De georthogonaliseerde vector y is nu uitgedrukt in de bekende vectoren x en y namelijk

$$\boxed{y_o = y - \frac{xy}{xx}\, x} \qquad (3.2)$$

Vervolgens wordt in de ruimte opgespannen door (x, y_o, z) de vector z georthogonaliseerd zodanig dat

$$z_o = z + \alpha x + \beta y_o \perp x,\ y_o$$

zodat, gebruikmakend van (9)

$$\begin{aligned} zx + \alpha xx &= 0 \\ zy_o + \beta y_o y_o &= 0 \end{aligned}$$

waaruit volgt

$$\alpha = -\frac{xz}{xx}$$

$$\beta = -\frac{y_o z}{y_o y_o}$$

en dus

$$\boxed{z_o = z - \frac{xz}{xx}x - \frac{y_o z}{y_o y_o}\, y_o} \qquad (3.3)$$

waarmee z_o in reeds bekende vectoren is uitgedrukt

De verschilvector $(q-q^*)$ volgt op overeenkomstige wijze. Deze vector staat namelijk volgens (2.8) loodrecht op de orthogonale basis en, het proces voortzettend wordt dus (zie fig. 1)

$$(q-q^*) = \boxed{q_o = q - \frac{xq}{xx}\, x - \frac{y_o q}{y_o y_o}\, y_o - \frac{z_o q}{z_o z_o}\, z_o} \qquad (3.4)$$

Het kwadraat van de lengte van deze vector is de som van kwadraten van de afwijkingen voorgesteld door S^2.

4. DE HERLEIDING VAN DE DIAGONAALELEMENTEN

De diagonaalelementen van (2.9) vertonen nog de eigenschap dat daarin de "oude" vectoren y en z voorkomen. Deze gemengde inprodukten kunnen als volgt herleid worden door de definitie van het inprodukt toe te passen.

$$y_o y = |y_o|\,|y|\cos\varphi = |y_o|\,|y_o| = y_o y_o \tag{4.1}$$

De juistheid hiervan volgt ook uit figuur 2. De noemers van (2.10) kunnen nu dus eveneens in deze vorm gebracht worden.

5. DE BETREKKING TUSSEN DE INPRODUKTEN

De inprodukten uit (2.9) kunnen als volgt in reeds bekende vectoren uitgedrukt worden. Vermenigvuldig beide leden van (3.2) met y, dan komt er

$$y_o y = yy - \frac{xy}{xx}\,xy$$

Met (4.1) volgt dan voor $y_o y_o$ de betrekking

$$y_o y_o = y_o y = yy - \frac{xy}{xx}\,xy \tag{5.1}$$

Vervolgens

$$y_o z = yz - \frac{xy}{xx}\,xz \tag{5.2}$$

$$y_o q = yq - \frac{xy}{xx}\,xq \tag{5.3}$$

Het nieuwe inprodukt ontstaat dus uit het oude door een aftrekterm in rekening te brengen. Uit (3.3) volgt nu op overeenkomstige wijze

$$z_o z_o = z_o z = zz - \frac{xz}{xx}\,xz - \frac{y_o z}{y_o y_o}\,y_o z \tag{5.4}$$

enz. met tenslotte tevens per definitie

$$q_o q_o = S^2 \tag{5.5}$$

Omgekeerd kunnen de inprodukten van de matrix (2.4) weer in die van de nieuwe matrix (2.9) uitgedrukt worden. Hiertoe worden de uitdrukkingen als volgt cyclisch gemaakt door eerst uit te schrijven

$$xx = \frac{xx}{\sqrt{xx}}\,\frac{xx}{\sqrt{xx}} \tag{5.6}$$

$$xy = \frac{xx}{\sqrt{xx}}\,\frac{xy}{\sqrt{xx}} \tag{5.7}$$

enz.

Uit (5.1) volgt dan

$$yy = \frac{xy}{\sqrt{xx}}\,\frac{xy}{\sqrt{xx}} + \frac{y_o y_o}{\sqrt{y_o y_o}}\,\frac{y_o y_o}{\sqrt{y_o y_o}} \tag{5.8}$$

en uit het stelsel (5.2) tot en met (5.4) op dezelfde wijze

$$yz = \frac{xy}{\sqrt{xx}}\,\frac{xz}{\sqrt{xx}} + \frac{y_o y_o}{\sqrt{y_o y_o}}\,\frac{y_o z}{\sqrt{y_o y_o}} \tag{5.9}$$

enz.

Het blijkt dat in deze vorm algemeen gesteld kan worden (zie hiervoor (5.6) t/m (5.9))

$$\begin{pmatrix} \frac{xx}{\sqrt{xx}} & 0 & 0 \\ \frac{xy}{\sqrt{xx}} & \frac{y_oy_o}{\sqrt{y_oy_o}} & 0 \\ \frac{xz}{\sqrt{xx}} & \frac{y_oz}{\sqrt{y_oy_o}} & \frac{z_oz_o}{\sqrt{z_oz_o}} \end{pmatrix} \begin{pmatrix} \frac{xx}{\sqrt{xx}} & \frac{xy}{\sqrt{xx}} & \frac{xz}{\sqrt{xx}} \\ 0 & \frac{y_oy_o}{\sqrt{y_oy_o}} & \frac{y_oz}{\sqrt{y_oy_o}} \\ 0 & 0 & \frac{z_oz_o}{\sqrt{z_oz_o}} \end{pmatrix} = \begin{pmatrix} xx & xy & xz \\ xy & yy & yz \\ xz & yz & zz \end{pmatrix}$$

Inderdaad is dus hierdoor de uitgangsmatrix (A) gesplitst in twee gelijke matrices (T) zodat

$$^tTT = A \tag{5.10}$$

6. HET REKENSCHEMA VAN CHOLESKI

De in Nota 134 besproken algebraïsche eigenschappen van de gevonden matrix T houden in dat men T verkregen kan denken door de matrix A, zie vergelijking (2.4), voor te vermenigvuldigen met een nog onbekende matrix D.

Schrijft men de kolom van bekende termen overeenkomstig (2.5) als c en de eenheidsmatrix als I en vormt men hieruit een matrix (A, c, I) dan wordt hieruit door voorvermenigvuldiging met D verkregen

$$(DA,\ Dc,\ DI) = (T,\ c^*,\ D)$$

Herleid men dus A zodanig dat T ontstaat, dan gaat de matrix I onder dezelfde herleiding over in D.

Verdere eigenschappen waren

$$^tDD = A^{-1} \tag{6.1}$$

en

$$^tDc^* = {}^tD(Dc) = {}^tDDc = A^{-1}c = b \tag{6.2}$$

waarin b de kolom met gevraagde coëfficiënt voorstelt en dus de oplossing van het stelselvergelijkingen (2.4) representeert. Zie (2.6).

Als uitgangsvorm dient dus de matrix (A, c, I) en wel in de volgende gedaante:

A			c	I		
a_1	a_2	a_3	=1	e_1	e_2	e_3
xx	xy	xz	xq	1		
	yy	yz	yq	0	1	
		zz	zq	0	0	1
			qq			

form (6.3)

Deze matrix moet nu gebracht worden in de volgende vorm

T = DA			$c^* = Dc$	D		
a_1	a_2	a_3	= 1	d_1	d_2	d_3
$\frac{xx}{\sqrt{xx}}$	$\frac{xy}{\sqrt{xx}}$	$\frac{xz}{\sqrt{xx}}$	$\frac{xq}{\sqrt{xx}}$	d_{11}		
	$\frac{y_o y_o}{\sqrt{y_o y_o}}$	$\frac{y_o z}{\sqrt{y_o y_o}}$	$\frac{y_o q}{\sqrt{y_o y_o}}$	d_{21}	d_{22}	
		$\frac{z_o z_o}{\sqrt{z_o z_o}}$	$\frac{z_o q}{\sqrt{z_o z_o}}$	d_{31}	d_{32}	d_{33}

form (6.4)

$$\boxed{\frac{q_o q_o}{\sqrt{q_o q_o}}} = S$$

Deling van elke rij (vergelijking) door de respectieve wortelvormen verandert de uitkomst voor a_1, a_2, a_3 niet. In schema wordt nu het rekenvoorschrift aangegeven in figuur 3.

Steeds wordt een oude term, aangegeven met een cirkel, verminderd met een som van produkten uit de kolommen welke zijn aangegeven met pijlen. Vergelijk dit met de formules (5.1) tot en met (5.4)
Tenslotte wordt door de wortel uit dediagonaalterm gedeeld.

Dit proces wordt ook op de matrix I toegepast. Er ontstaat dan achtereenvolgens

$$d_{11} = \frac{1}{\sqrt{xx}}$$

$$d_{21} = \left[0 - \frac{xy}{\sqrt{xx}}\,\frac{1}{\sqrt{xx}}\right] / \sqrt{y_o y_o}$$

$$d_{22} = \frac{1}{\sqrt{y_o y_o}}$$

enz.

De overige elementen van de matrix D zijn gelijk aan 0.

De oplossing b wordt nu uit (6.4) verkregen door de kolommen van D met de kolomvector c^* te vermenigvuldigen zoals in (6.2) werd aangegeven.

7. TOEPASSING BIJ DE NIET-LINEAIRE VEREFFENING

In de Nota's 138 en 139 werd een nadere beschouwing gegeven over de niet-lineaire vereffening. Hierbij werd als werkmethode onder meer de mogelijkheid geopperd een aantal parameters tijdens de bewerking als constant te beschouwen en op de overige parameters te vereffenen. Bij een volgende bewerkingsslag kunnen deze twee groepen parameters van rol verwisselen.

Stel dat bijvoorbeeld de paramters a, b en c gevraagd worden zodat bijvoorbeeld

$$y = f(x;\ a,\ b,\ c)$$

De matrix waaruit de normaalvergelijkingen worden samengesteld heeft nu de volgende vectoren tot basis:

$$\left(\frac{\partial y}{\partial a} \; \frac{\partial y}{\partial b} \; \frac{\partial y}{\partial c} \right)$$

Te berekenen zijn nu de correcties Δa, Δb, Δc.

Stel dat de parameter c constant gehouden wordt. De matrix waaruit de normaalvergelijkingen worden samengesteld heeft voor dit geval tot basis de vectoren,

$$\left(\frac{\partial y}{\partial a} \; \frac{\partial y}{\partial b} \right)$$

en berekend moeten worden Δa en Δb.

In de terminologie van het reeds behandelde betekent dit dat alleen de oplossing van a_1 en a_2 gevraagd wordt uit het schema

T		c^*	D	
a_1	a_2	$= 1$	d_1	d_2
$\frac{xx}{\sqrt{xx}}$	$\frac{xy}{\sqrt{xx}}$	$\frac{xq}{\sqrt{xx}}$	d_{11}	
	$\frac{y_oy_o}{\sqrt{y_oy_o}}$	$\frac{y_oq}{\sqrt{y_oy_o}}$	d_{21}	d_{22}

form (7.1)

De eerste rij bevat elementen die ook voorkomen in (6.4) terwijl de 2e rij op identieke wijze als aangegeven in figuur 3 uit de eerste rij herleid wordt. Dit houdt in dat ook de d_{21} en d_{22} in (7.1) gelijk zullen zijn aan die in (6.4). Is dus het gehele schema (6.4) doorgerekend dan volgt de oplossing van (7.1) uit het gedeelte van (6.4) dat luidt:

$$\begin{pmatrix} d_{11} & d_{21} \\ 0 & d_{22} \end{pmatrix} \begin{pmatrix} \frac{xq}{\sqrt{xx}} \\ \frac{y_oq}{\sqrt{y_oy_o}} \end{pmatrix}$$

Er kan nu een algemene regel worden opgesteld waaruit achtereenvolgens de oplossing wordt verkregen van

de 1e	vergelijking met	1	onbekende		
de 1e en de 2e	"	"	2	"	n
de 1e t/m de 3e	"	"	3	"	n

Deze regel luidt dat de oplossing wordt verkregen door van de matrixvermenigvuldiging (6.2)

$$^tDc^* = b$$

steeds de deelresultaten te noteren en wel

$$\begin{pmatrix} d_{11} & d_{21} & d_{31} \\ 0 & d_{22} & d_{32} \\ 0 & 0 & d_{33} \end{pmatrix} \begin{pmatrix} c_1^* \\ c_2^* \\ c_3^* \end{pmatrix} =$$

$$\begin{pmatrix} d_{11}c_1^* & d_{11}c_1^* + d_{21}c_2^* & d_{11}c_1^* + d_{21}c_2^* + d_{31}c_3^* \\ 0 & d_{22}c_2^* & d_{22}c_2^* + d_{32}c_3^* \\ 0 & 0 & d_{33}c_3^* \end{pmatrix} \to \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

De eerste kolom geeft de oplossing voor de 1e vergelijking met 1 onbekende, de tweede kolom de oplossing voor de eerste twee vergelijkingen met twee onbekenden enz.

Wordt nu bijvoorbeeld gevraagd naar de oplossing van de volgende gevallen

a_2 alleen

a_2, a_3 gezamenlijk

a_1, a_2, a_3 totale oplossing

dan dienen de <u>kolommen</u> en <u>rijen</u> in (2.4) zodanig gerangschikt te worden dat de volgorde wordt

$$(a_2 \quad a_3 \quad a_1 = 1 \quad e_1 \quad e_2 \quad e_3)$$

Dit houdt in dat de <u>rijen</u> in de eindoplossing achtereenvolgens worden

$$\begin{pmatrix} a_2 & a_2 & a_2 \\ 0 & a_3 & a_3 \\ 0 & 0 & a_1 \end{pmatrix}$$

waarmee alle gevraagde oplossingen uit één rekenschema zijn verkregen.

8 HERLEIDING VAN DE UITGANGSMATRIX A (4a)

De matrix A kan herleid worden tot een matrix waarvan de diagonaalelementen 1 zijn door alle inprodukten door de lengten van de samenstellende vectoren te delen. De elementen stellen nu de cosinussen van de hoeken tussen de basisvectoren x, y en z voor. Voor een (2 x 2) matrix wordt dit

$$\begin{pmatrix} \dfrac{xx}{\sqrt{xx}\sqrt{xx}} & \dfrac{xy}{\sqrt{xx}\sqrt{yy}} \\ \dfrac{xy}{\sqrt{xx}\sqrt{yy}} & \dfrac{yy}{\sqrt{yy}\sqrt{yy}} \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} 1 & m_{xy} \\ m_{xy} & 1 \end{pmatrix} = M \qquad (8.1)$$

Worden de lengten van de vectoren x en y aangeduid met de schaalfactoren h_1 en h_2 dan wordt gedefinieerd

$$\begin{pmatrix} \sqrt{xx} & 0 \\ 0 & \sqrt{yy} \end{pmatrix} = \begin{pmatrix} h_1 & 0 \\ 0 & h_2 \end{pmatrix} = H \quad \text{en } H^{-1} = \begin{pmatrix} \frac{1}{h_1} & 0 \\ 0 & \frac{1}{h_2} \end{pmatrix}$$

Nu ontstaat M uit A (zie (2.4) door respectievelijk elke rij en elke kolom van A door de lengte van de vector die op het snijpunt van die rij en kolom voorkomt te delen.

In matrix notatie is het vermenigvuldigen van rijen met een constante een vóórvermenigvuldiging, het vermenigvuldigen van kolommen met een constante een návermenigvuldiging en dus

$$M = H^{-1} A H^{-1} \tag{8.2}$$

Het effect van deze vermenigvuldigingen is verschillend, namelijk vergelijking (2.5)

$$Ab = c$$

kan voorvermenigvuldigd worden met H^{-1} zodat er komt

$$H^{-1}Ab = H^{-1}c$$

Dit betekent dat elke vergelijking met een bepaald bedrag vermenigvuldigd wordt, waardoor de oplossing niet verandert. De kolom van bekenden is nu getransformeerd in een nieuwe kolom.

Het navermenigvuldigen van A met H^{-1} doet echter de schaal van de gevraagde constanten veranderen namelijk

$$H^{-1}AH^{-1}\ (Hb) = H^{-1}c \tag{8.3}$$

Dit houdt in dat elke kolom met een bepaald bedrag vermenigvuldigd wordt, zodat nu opgelost worden de onbekenden

$$Hb = \begin{pmatrix} \sqrt{xx} & 0 \\ 0 & \sqrt{yy} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} a_1\ \sqrt{xx} \\ a_2\ \sqrt{yy} \end{pmatrix} \underset{=}{\overset{def}{=}} b' \tag{8.4}$$

Hierop moet na het bereiken van de eindoplossing dus nog herleid worden. Uit (8.3) volgt namelijk met gebruik van (8.2) dat

$$Mb' = H^{-1}c \ \overset{def}{=} c' \tag{8.5}$$

en dus

$$b' = M^{-1}c' \tag{8.6}$$

In het schema van CHOLESKI wordt deze uitkomst dus verkregen met

$$D(M,\ c',\ I) = (T,\ (c')^*,\ D)$$

met als oplossing volgens (6.2)

$$b' = {}^tD(c')^* = M^{-1}c'$$

Dit wordt met (8.4) tot b herleid door voorvermenigvuldigen van het resultaat met H^{-1} dus

$$b = H^{-1}b' = H^{-1}(M^{-1}c')$$

De vorm tussen haken is de "inverse maal de (getransformeerde) onbekende".

Wordt de kolom met bekende termen c eveneens nog gedeeld door de lengte van de vector q voorgesteld door $\sqrt{qq} = h_q$, dan gaat c over in een kolom van cosinussen van hoeken tussen respectievelijk x, y en q zodat

$$\begin{pmatrix} c_1 \\ \\ c_2 \end{pmatrix} \rightarrow \begin{pmatrix} \dfrac{xq}{\sqrt{xx}\sqrt{qq}} \\ \\ \dfrac{yq}{\sqrt{yy}\sqrt{qq}} \end{pmatrix}$$

Ook deze bewerking houdt een herleiding van de gevraagde constanten in namelijk uitgaande van (8.3)

$$H^{-1}AH^{-1}\ (Hb) = h_q\ (\frac{1}{h_q}\ H^{-1}c) \stackrel{def}{=\!=} h_q c'' \qquad (8.7)$$

De oplossing wordt

$$\frac{1}{h_q}\ Hb \stackrel{def}{=\!=} b'' = (M^{-1}c'') \qquad (8.8)$$

waarin de vorm tussen haken weer de "inverse maal de (getransformeerde) bekende" voorstelt. Tenslotte is dus(Nota 134 pag.4: de betrekking tussen α_j en β_j)

$$b = h_q H^{-1}\ b''$$

Opgemerkt wordt dat nog geldt volgens (8.2)

$$M^{-1} = (H^{-1}A\ H^{-1})^{-1} = H\ A^{-1}\ H$$

zodat

$$A^{-1} = H^{-1}\ M^{-1}H^{-1}$$

waarmee de inverse van A uit M^{-1} berekend kan worden.

9. DE CORRELATIEMATRIX

Men kan (2.4) ontstaan denken uit de algemene vergelijking

$$u = a_1v_1 + a_2v_2 + a_3v_3 + a_4 \qquad (9.1)$$

door een verschuiving naar het zwaartepunt toe te passen en te schrijven met $\bar{v}_i$ als gemiddelde waarden van v_i enz.

$$(u-\bar{u}) = a_1(v_1-\bar{v}_1) + a_2(v_2-\bar{v}_2) + a_3(v_3-\bar{v}_3) \qquad (9.2)$$

waaruit dus volgt

$$a_4 = \bar{u} - a_1\bar{v}_1 - a_2\bar{v}_2 - a_3\bar{v}_3 \qquad (9.3)$$

waarmee a_4 uit de coëfficiënten a_i en de gemiddelde waarden van u en v_i berekend kan worden. De vergelijking (9.2) gaat nu over in een van het type (2.4) door te stellen

$$\begin{aligned} u-\bar{u} &= q \\ v_1-\bar{v}_1 &= x \end{aligned} \qquad (9.4)$$

enz.

Wordt nu het schema A van de normaalvergelijkingen volgens (2.4) gebracht in de gedaante van (8.1) dan gaat M over in de correlatiematrix met correlatiecoëfficiënten

$$r_{v_1v_2} = \frac{(v_1-\bar{v}_1)(v_2-\bar{v}_2)}{\sqrt{(v_1-\bar{v}_2)^2}\sqrt{(v_2-\bar{v}_2)^2}} = \frac{xy}{|x|\,|y|} = r_{xy}$$

Alle betrekkingen met M uit paragraaf 8 gelden nu op overeenkomstige wijze met betrekking tot de correlatiematrix.

De multipele correlatiecoëfficiënt die als maat gebruikt kan worden voor de aanpassing tussen u en $u(v_i)$ volgens (9.1) is de gewone correlatiecoëfficiënt tussen u en u^* en dus tussen q en q^* de toegepaste herleiding (9.4) FISHER, 1958, pag. 258; KENNEY and KEEPING, 1956, pag.308)

Dit houdt in dat volgens de definitie van het inpnodukt

$$R = \frac{qq^*}{|q||q^*|} = \frac{|q||q^*|\cos\varphi}{|q||q^*|} = \frac{|q^*|}{|q|}$$

wat meetkundig uit figuur 1 volgt daar $R = \cos\varphi$

Er geldt dus

$$\boxed{R^2 = \frac{q^*q^*}{qq} = \frac{q^*q^*}{h_q^2}} \tag{9.5}$$

De multipele correlatiecoëfficiënt kan als volgt uit de verkregen uitkomsten met het CHOLESKI schema verkregen worden. Uit (2.5) volgt, geschreven als matrixprodukt

$$^t bc = {}^t bAb$$

en met (2.4')

$$= {}^t b\, {}^t XXb$$

$$= {}^t(Xb)(Xb)$$

en dus met (2.2') voor de teller van (9.5)

$$\boxed{{}^t bc = q^*q^*}$$

Met de herleiding volgens (8.2) wordt deze uitkomst met (8.4) en (8.5)

$${}^t b'c' = {}^t(Hb)(H^{-1}c)$$

$$= {}^t bH\, H^{-1}c$$

en dus voor de teller van (9.5)

$$\boxed{{}^t b'c' = q^*q^*}$$

Tenslotte kan R^2 rechtstreeks verkregen worden uit het schema waarin q als eenheidsvector voorkomt en wel met (8.7) en (8.8)

$${}^t b''c'' = {}^t\left(\frac{1}{h_q}Hb\right)\left(\frac{1}{h_q}H^{-1}c\right)$$

$$= \frac{1}{h_q^2}\, {}^t b\, {}^t H\, H^{-1}c$$

waarin volgens paragraaf 8 geldt dat ${}^t H = H$ en dus

$$\boxed{{}^t b''c'' = \frac{1}{h_q^2}\, {}^t bc = \frac{q^*q^*}{h_q^2} = R^2}$$

10. DE DETERMINANT VAN A

Volgens een eigenschap van het vermenigvuldigen van determinanten volgt uit (5.10)

$$|A| = |{}^{t}T||T|$$

Aangezien T slechts een driehoeksmatrix is volgt dat de determinant van T de waarde heeft die gelijk is aan het produkt van de diagonaalelementen en dus

$$|A| = (xx)(y_o y_o)(z_o z_o)$$

Indien twee richtingen x en y samenvallen, of bijna samenvallen, zal (zie fig. 2) voor de lengte van y_o gelden

$$|y_o| \approx 0$$

waaruit dan tevens volgt

$$|A| \approx 0$$

zodat er geen; of een niet nauwkeurig te bepalen, inverse bestaat.

De mogelijkheid kan zich voordoen dat $|A| < 0$. In dat geval bevat T imaginaire elementen op de hoofddiagonaal. Een en ander houdt in dat de oriëntatie van het assenstelsel gespiegeld is (BIJL en SALET pag.51 en 70).

11. DE MEETKUNDIGE VOORSTELLING VAN HET INVERTEREN

In figuur 4 en 5 wordt een meetkundige voorstelling van de matrixinversie gegeven.
Uitgegaan is van het hypothetische geval dat opgelost moet worden

$$\begin{aligned} a_1 + 3a_2 &= 4 \\ 3a_1 + 2a_2 &= 5 \end{aligned} \qquad (11.1)$$

of in schema

$$\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$$

$$A^{-1}.\, b = c$$

Dit kan nu zo opgevat worden dat de vector b door transformatie overgaat ($\rightarrow$) in de vector c volgens

$$b \rightarrow c = Ab$$

De oplossing van (11.) is

$$b = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

zodat gesteld kan worden dat de vector b in de parameterruimte (fig. 4.a) afgebeeld wordt op c in de beeldruimte (fig. 4.b) door dezelfde lineaire combinatie (1x en 1x) van de basisvectoren, de kolommen van A, te nemen.

De basis van de beeldruimte kan nu herleid worden op eenheidsvectoren door het toepassen van lineaire combinaties op de kolomvectoren van A. Worden op de eenheidsvectoren van de parameterruimte dezelfde lineaire combinaties toegepast dan ontstaat daaruit het inverse stelsel (fig. 5)

Achtereenvolgens wordt verkregen:

	beeldruimte	parameterruimte	bewerking
figuur 4b	$A = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$,	$I = \frac{1}{7}\begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix}$,	uitgangsvorm
			2e kolom maal 1/2
figuur 4c	$\begin{pmatrix} 1 & 1\frac{1}{2} \\ 3 & 1 \end{pmatrix}$,	$\frac{1}{7}\begin{pmatrix} 7 & 0 \\ 0 & 3\frac{1}{2} \end{pmatrix}$	
			2e kolom maal 3
figuur 4d	$\begin{pmatrix} 1 & 4\frac{1}{2} \\ 3 & 3 \end{pmatrix}$,	$\frac{1}{7}\begin{pmatrix} 7 & 0 \\ 0 & 10\frac{1}{2} \end{pmatrix}$	
			2e kolom min 1e
	$\begin{pmatrix} 1 & 3\frac{1}{2} \\ 3 & 0 \end{pmatrix}$,	$\frac{1}{7}\begin{pmatrix} 7 & -7 \\ 0 & 10\frac{1}{2} \end{pmatrix}$	
			2e kolom maal $\frac{2}{7}$
figuur 4e	$\begin{pmatrix} 1 & 1 \\ 3 & 0 \end{pmatrix}$,	$\frac{1}{7}\begin{pmatrix} 7 & -2 \\ 0 & 3 \end{pmatrix}$	
			1e kolom min 2e
figuur 4f	$\begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix}$,	$\frac{1}{7}\begin{pmatrix} 9 & -2 \\ -3 & 3 \end{pmatrix}$	
			1e kolom maal 1/3, kolommen verwisselen
figuur 4g	$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$,	$A^{-1} = \frac{1}{7}\begin{pmatrix} -2 & 3 \\ 3 & -1 \end{pmatrix}$	

Nu kan men zich de rol van de ruimten verwisseld denken. De vector c wordt nu afgebeeld op b in de parameterruimte (fig. 5). Deze transformatie kan berekend worden aangezien

$$c \rightarrow b = A^{-1}c$$

De kolomvectoren van A^{-1} staan loodrecht op die van A, het inprodukt van de oorspronkelijke met de nieuwe vectoren is 1 zodat inderdaad van een invers stelsel gesproken kan worden ($AA^{-1} = I$).

Uit deze beschouwingswijze volgt dat de oplossing algemeen is, dat wil zeggen dat voor elke kolom van bekenden c het stelsel (11.2) rechtstreeks opgelost kan worden door de inverse transformatie toe te passen.

Opgemerkt wordt nog dat de geldigheid van het gegeven beeld alleen voor symmetrische matrices volledig opgaat.

12. DE MEETKUNDIGE VOORSTELLING VAN HET ELIMINEREN

Een andere mogelijkheid tot het oplossen van (11.2) is de onbekenden a_1 en a_2 te elimineren. Nu wordt er met de rijen (vergelijkingen) gewerkt. De methode berust op het optellen en aftrekken van onderling dezelfde lineaire combinaties (in dit geval weer (1x en 1x) waardoor deze gehandhaafd blijven. De matrix waarop de bewerking wordt uitgevoerd wordt nu aangevuld met de vector c zodat (zie fig. 6a)

$$(A, c) = \begin{pmatrix} 1 & 3 & 4 \\ 3 & 2 & 5 \end{pmatrix}$$

Achtereenvolgens ontstaan nu

figuur 6b $\begin{pmatrix} 1 & 3 & 4 \\ 1 & 2/3 & 5/3 \end{pmatrix}$, 2e kentallen maal 1/3.

figuur 6c $\begin{pmatrix} 1 & 3 & 4 \\ 1 & 2/3 & 5/3 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 1 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 4 \\ 0 & -7/3 & -7/3 \end{pmatrix}$

figuur 6d $\begin{pmatrix} 1 & 3 & 4 \\ 0 & 1 & 1 \end{pmatrix}$, 2e kentallen maal $(-\frac{3}{7})$

figuur 6e $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$, 1e kentallen min 3 maal 2e

waaruit het stelsel van tweemaal één vergelijking met één onbekende volgt

$$a_1 \quad = 1$$
$$a_2 = 1$$

De oplossing is alleen bruikbaar voor de gegeven kolom met bekenden c. Voor een nieuwe kolom met bekenden moet de bewerking herhaald worden.

13. DE MEETKUNDIGE VOORSTELLING VAN HET OPLOSSEN

Volledigheidshalve wordt nog de oplossing van het stelselvergelijkingen (11.1) in Cartesiaanse voorstelling gegeven.

Wordt namelijk (11.1) als een tweetal rechte lijnen voorgesteld dan zal (SCHREK, pag. 41)

$$(a_1 + 3a_2 - 4) + \lambda(3a_1 + 2a_2 - 5) = 0$$

een lijnenbundel voorstellen waarvan elk exemplaar door het snijpunt van beide gegeven lijnen (11.1) zal gaan. Een exemplaar van deze bundel is ook die lijn waarvoor $\lambda = \frac{1}{3}$. Deze lijn looptevenwijdig aan de a_1-as en er resteert een vergelijking in a_2. Op deze wijze kunnen ook voor meer vergelijkingen met meer onbekenden alle assen (variabelen) achtereenvolgens geëlimineerd worden.

LITERATUUR

BIJL J. en H.J.H.SALET, (1958) - Analytische meetkunde II. Delft (I.C.W. 11/99)

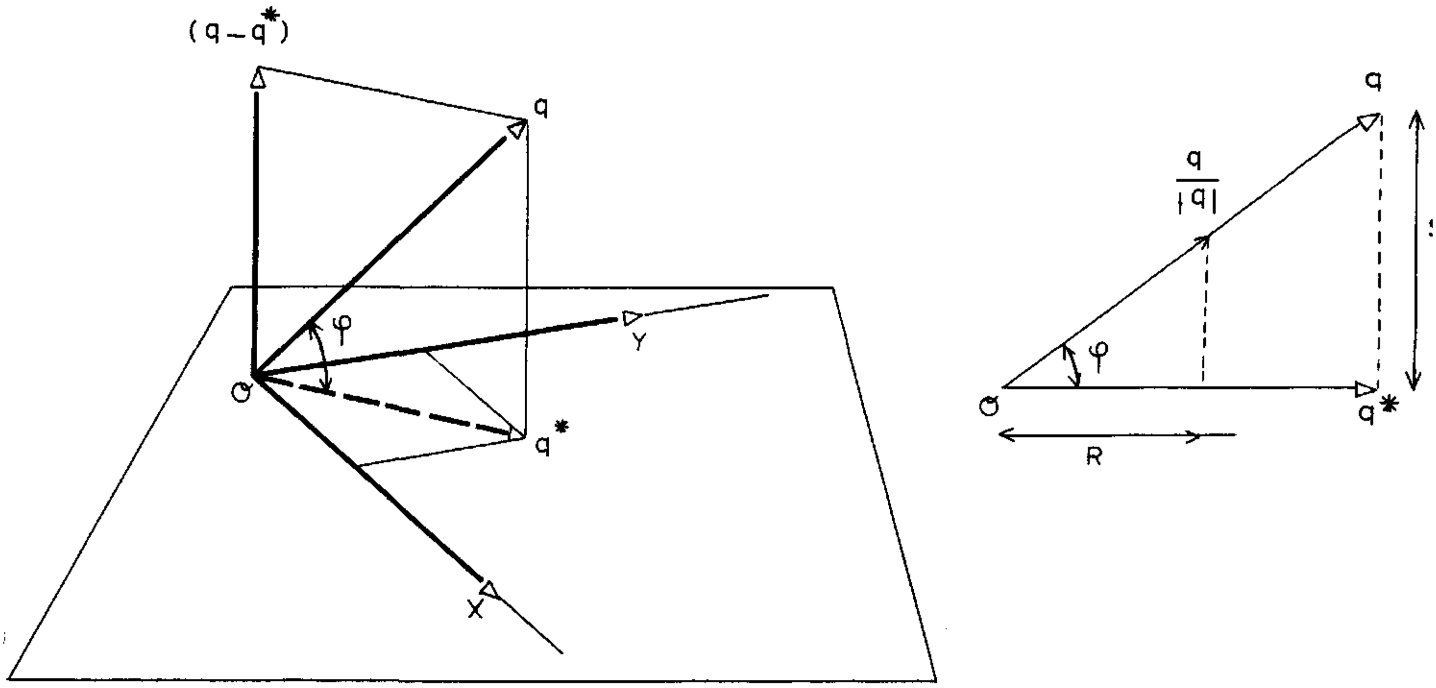
FISHER, R.A.(1958) - Statistical Methods for Research Workers, London (I.C.W. 11/103)

KAMIL,L.P., (1962) - Lineaire regressie, I.C.W. Nota 134

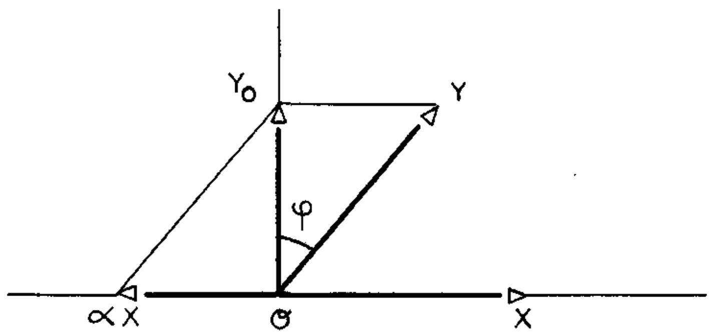
KENNEY,J.F. and E.S.KEEPING,(1956)- Mathematics of Statistics dl.I. N.Y. (I.C.W. 11/35)

SCHREK,D.J.E., Beginselen der analytische meetkunde. Groningen. (I.C.W. 11/66)

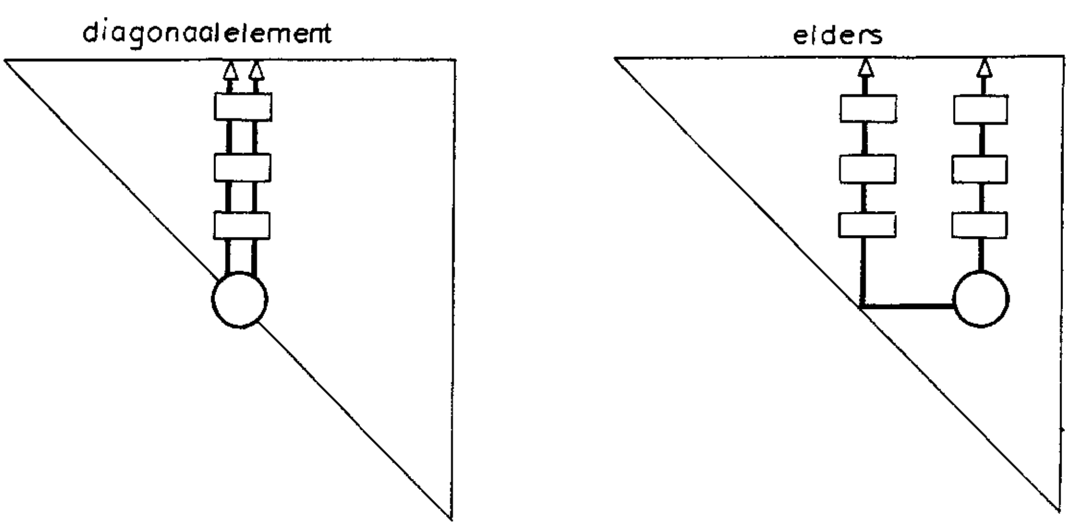
STOL,Ph.Th. (1962) - Een meetkundige beschouwing over de niet-lineaire vereffening. I.C.W. Nota 138



figuur 1

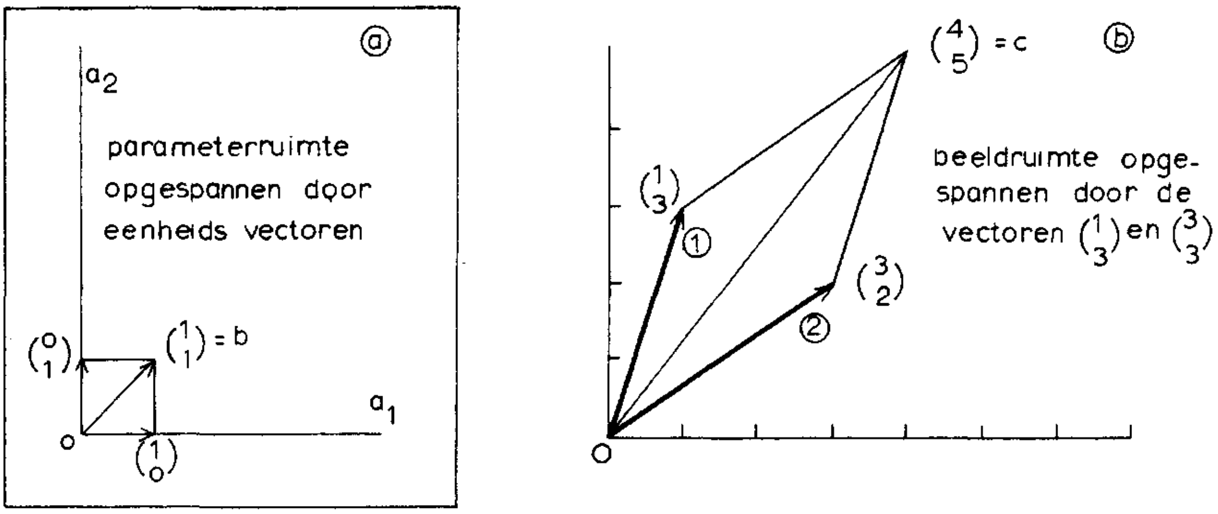


figuur 2

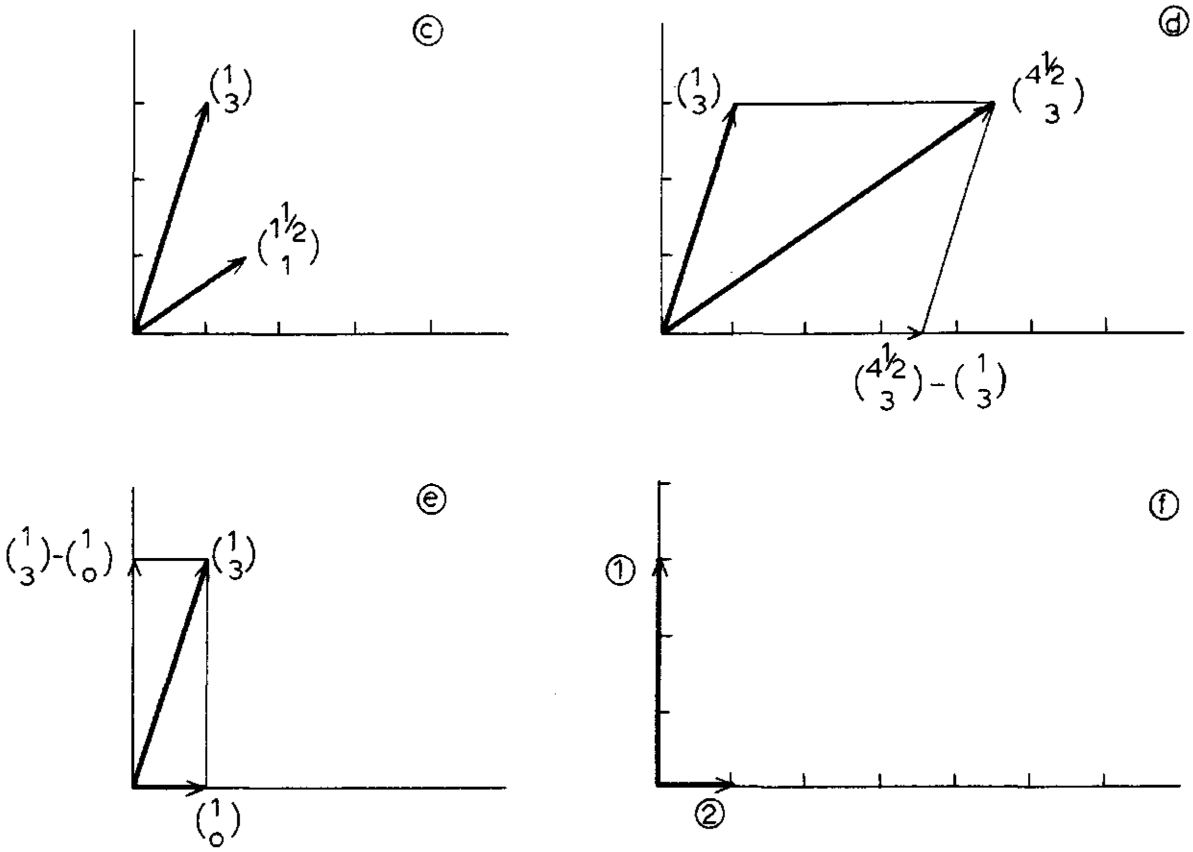


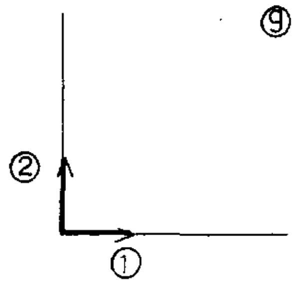
figuur 3

figuur 4



eenheidsvectoren $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ gaan over in $\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$

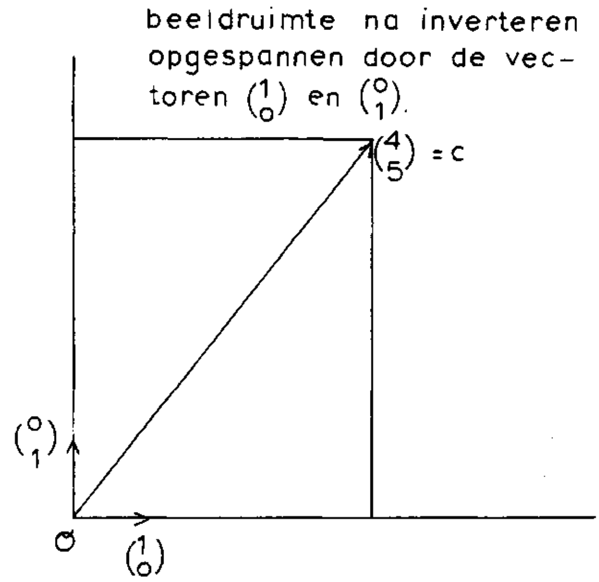




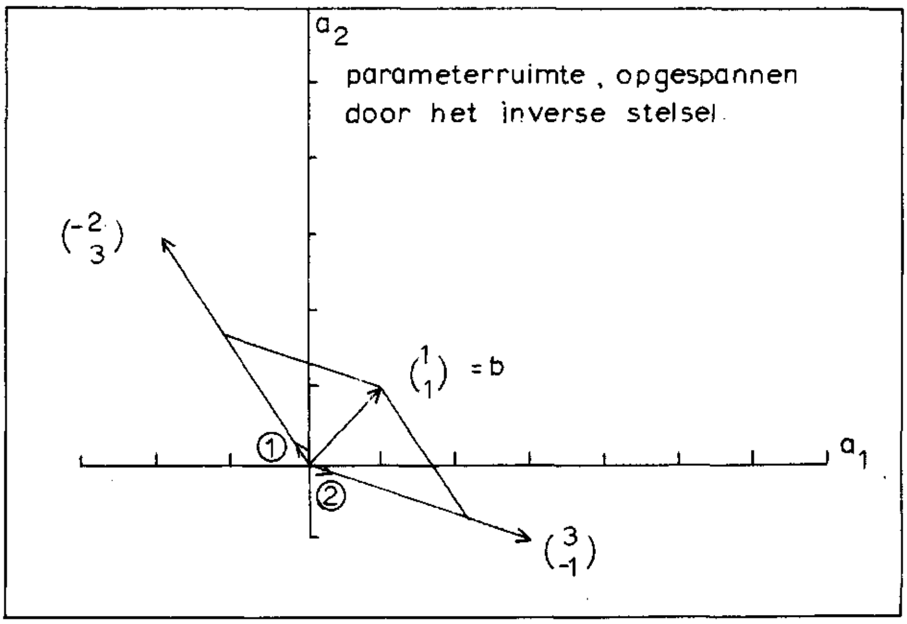
herleiding van

$\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$ in $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

figuur 5



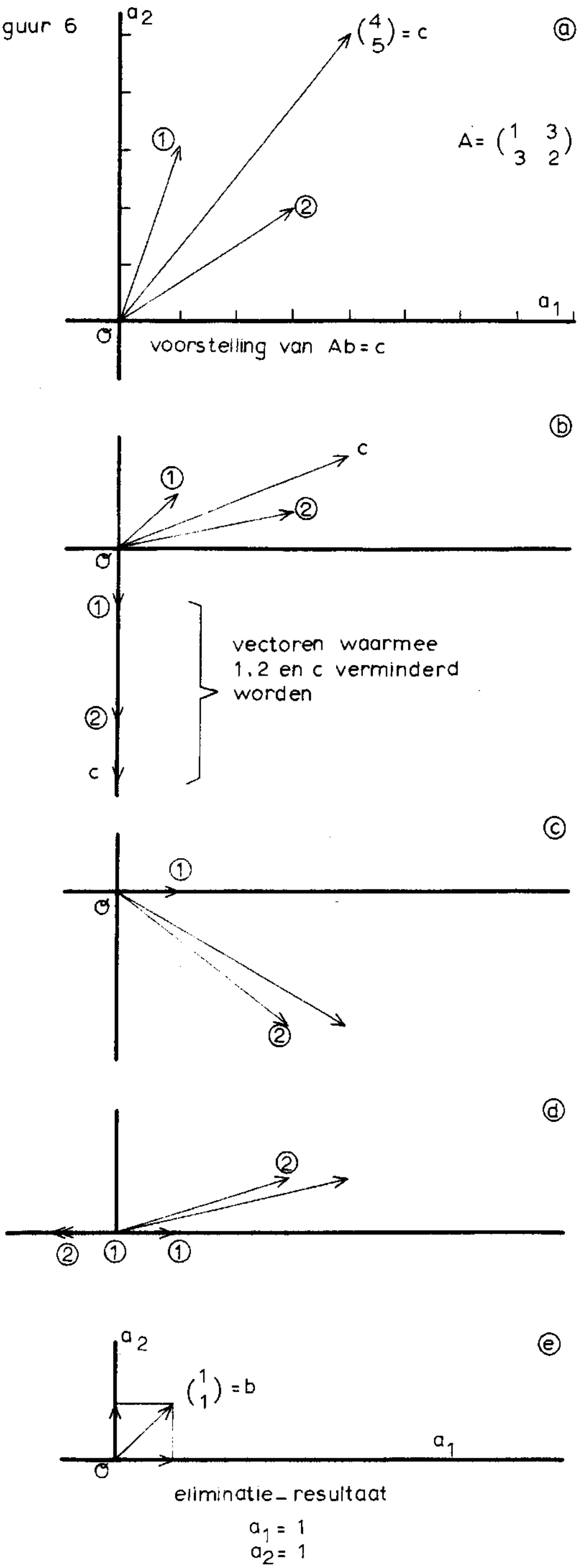
inverse transformatie van $\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$



eenheidsvectoren $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ gaan over in $\frac{1}{7}\begin{pmatrix} -2 & 3 \\ 3 & -1 \end{pmatrix}$.

de vector $\binom{4}{5}$ gaat over in $\binom{1}{1}$.

figuur 6



a
a_2
$\binom{4}{5} = c$
$A = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$
1
2
a_1
0
voorstelling van Ab = c
b
c
1
2
0
1
2
c
vectoren waarmee 1,2 en c verminderd worden
c
1
0
2
d
2
2
1
1
e
a_2
$\binom{1}{1} = b$
a_1
0
eliminatie-resultaat
$a_1 = 1$
$a_2 = 1$