Stichting
MATHEMATISCH CENTRUM
2e Boerhaavestraat 49
Amsterdam

R 179

STROKING

M.L. Potters

1952

STROKING

1. In verband met een opdracht werden twee grootheden tegelijk aan een topstroking onderworpen. Beide waren gegeven als functie van A_t en H_t, op gelijke intervallen, als volgt:

H_t : 5(5) 25 $\qquad H_t \leqslant A_t$

A_t : 5(1) 25

Een subtabellatie moest verkregen worden met A_t en z als parameters, waarbij $H_t = A_t \sin 2.5\, z$. Daartoe werd getracht de functies voor te stellen door vierde-graadspolynomen:

$$(1) \quad u(A_t,z) = P_{44}(A_t,z) = \sum_{p,q} \beta_{pq} \; A_t^p \, z^q \qquad (0 \leqslant p+q \leqslant 4),$$

De 15 coefficienten β_{pq} werden bepaald met behulp van de methode der kleinste quadraten, hetgeen leidde tot 15 lineaire vergelijkingen in β_{pq}:

$$(2) \quad \sum_n \sum_{p,q} \beta_{pq} \; A_t^{p+i} \, z^{q+k} = \sum_n \; u(A_t,z) \; A_t^i \, z^k \qquad (0 \leqslant i+k \leqslant 4).$$

Het aantal gegeven functiewaarden n = 55.

In matrix-notatie luiden de vergelijkingen (2)

$$(3) \qquad \mathcal{A} \cdot \mathcal{X} = \mathcal{C}$$

De matrix $\mathcal{A}$ met elementen van de vorm $A_t^{p+i} \, z^{q+k}$ is quadratisch (15 x 15) en symmetrisch.

$\mathcal{X}$ is de kolom der te bepalen coefficienten β_{pq}.

$\mathcal{C}$ is de kolom van de rechterleden der vergelijkingen.

De oplossing van (3) kan nu geschreven worden als

$$(4) \qquad \mathcal{X} = \mathcal{A}^{-1} \cdot \mathcal{C}$$

De matrix $\mathcal{A}^{-1}$ is voor beide functies dezelfde; als $\mathcal{A}$ dus eenmaal geïnverteerd zou zijn, kon elke op deze wijze gegeven functie gemakkelijk gestrookt worden.

2. De inversie van $\mathcal{A}$ werd volgens een der gebruikelijke methoden uitgevoerd (Fox, L., 1950, J.R.Statist.Soc., Ser.B, 12, 120).
De elementen van $\mathcal{A}$ liepen zeer uiteen in grootte en namen van links boven naar rechts onder een aantal machten van 10 toe:

$$a_{11} = 5.5 \cdot 10^1$$

$$a_{15,15} = 1.6 \cdot 10^{13} .$$

Ze werden geschreven als $a_{mn} = \mathcal{E}.10^s$ met $1 \leqslant \mathcal{E} < 10$ en $\mathcal{E}$ in 7 decimalen. Voor $\mathcal{A}$ was dus $1 \leqslant S \leqslant 13$. Voor de inverse matrix bleek $-10 \leqslant S \leqslant 3$, terwijl 4 tot 6 decimalen werden verkregen.

De bekende termen werden eveneens in deze vorm geschreven , waarna volgens (4) door scalaire vermenigvuldiging van $\mathcal{C}$ met de kolommen van $\mathcal{A}^{-1}$ de coefficienten β_{pq} verkregen konden worden, bijv.

$$\beta_{00} = \sum_{r=1}^{15} (a^{-1})_{r,1} c_r = \sum_{r=1}^{15} \mathcal{E}_r.10^{s_r} \cdot \eta_r.10^{t_r} =$$

$$= \sum_{r=1}^{15} \mathcal{E}_r \cdot \eta_r.10^{s_r+t_r} = \sum_{r=1}^{15} \mathcal{E}'_r \cdot \eta'_r \cdot 10^5$$

Het bleek, dat $s_r + t_r = 4$ à 5; de factor 10^5 werd daarom voor het somteken geplaatst en een eventuele factor 10 in $\mathcal{E}'_r\, \eta'_r$ ondergebracht. Nu was echter $\sum_{r=1}^{15} \mathcal{E}'_r \eta'_r$ tot in 5 decimalen nul.

Eenzelfde verschijnsel trad voor de andere coefficienten op. Gezien de vrij grote precisie, waarmee het proces was uitgevoerd, was dit een vreemd resultaat. Het berekenen van een zesde decimaal zou geen zin gehad hebben, omdat het dan verkregen resultaat van de grootte-orde der afrondingsfouten zou zijn.

3. Om een inzicht in dit verschijnsel te krijgen, werd eerst een soortgelijk probleem behandeld in één dimensie. Een der grootheden was n.l. nagenoeg onafhankelijk van H_t en dus van z. Getracht werd nu een polynoom

(5) $$P_4(A_t) = b_0 + b_1 A_t + b_2 A_t^2 + b_3 A_t^3 + b_4 A_t^4$$

te leggen door de functiewaarden bij $H_t = 5$ (21 stuks). Voor deze coefficienten zal ongeveer gelden $b_k = \beta_{k,0}$ $(k = 0,1,\ldots,4)$. Dit leidde tot de vergelijkingen

(6) $$\mathcal{B} \cdot y = d$$

waarin $\mathcal{B}$ een matrix van (5 x 5) is, analoog aan $\mathcal{A}$
y de kolom der coefficienten b_k $(k = 0,\ldots,4)$
d de kolom der bekende termen, analoog aan $\mathcal{C}$

De oplossing van (6) is

(7) $$y = \mathcal{B}^{-1} \cdot d$$

De berekening werd tweemaal uitgevoerd:

a) met de matrix B, waarvan het kleinste element $b_{11} = 21$
" grootste " $b_{55} = 5.10^{11}$

b) met de matrix B^*, uit B verkregen door zijn rijen en kolommen met zódanige machten van 10 te vermenigvuldigen, dat de elementen van ongeveer gelijke grootte-orde werden en wel was $1 \leqslant b_{mn} \leqslant 22$.

Deze behandeling komt er op neer, dat i.p.v. (7) de oplossing geschreven wordt

$$y = M \cdot B^{*-1}(Md) \tag{8}$$

waarin M een diagonaalmatrix is, waarvan de elementen machten van 10 zijn.

Beide berekeningen voerden tot een resultaat, dat minder overeenstemt, naarmate de coefficienten "hoger" zijn. Er werden 5 tot 7 cijfers meegenomen. Op grond hiervan kan men een relatieve fout van 10% voor de lagere tot 100% voor de hogere coefficienten verwachten.

			Berekening met matrix	
			B (oorspronkelijke)	B^* (getransformeerde)
coefficienten	b_0		0.196	0.156
	b_1		$0.757.10^{-1}$	$0.963.10^{-1}$
	b_2		$0.499.10^{-2}$	$0.195.10^{-2}$
	b_3		$-\ 0.150.10^{-3}$	$0.379.10^{-4}$
	b_4		$0.135.10^{-5}$	$-\ 0.176.10^{-5}$
aansluiting aan gegeven waarden			$P_4(A_t)$	
	$A_t = 5$	$u(A_t)=0.7$	0.68	0.69
	$A_t = 25$	$u(A_t)=3.7$	3.39	3.68

De tweede methode geeft dus beter aan de gegevens aansluitende waarden. De controle-checks werden bij beide berekeningen uitgevoerd; bij de getransformeerde matrix was de afwijking $\pm$ 0.25, bij de eerste werd de check zinloos, doordat de significante cijfers wegvielen. De tweede methode lijkt het dus in nauwkeurigheid te winnen van de eerste.

Berekening met een derde-graadsvorm bracht niets nieuws aan het licht.

4. Gezien dit voorlopige resultaat, leek het de moeite waard, de matrix A opnieuw te inverteren, nu echter, na hem de in 3. genoemde behandeling met machten van 10 te hebben later ondergaan. Hierdoor

werd A getransformeerd tot A^*

	A^*	A^{*-1}
het kleinste element	0.42	49
" grootste "	40.8	$1.9 \cdot 10^6$

De elementen werden nu steeds met een constant aantal decimalen geschreven i.p.v. met een constant aantal cijfers, zoals eerst. En wel in A^* met 6, in A^{*-1} met 4 decimalen.

Hoewel ook nu bij de berekening der coefficienten β_{pq} een aantal significante cijfers tegen elkaar wegvielen, bleef er voldoende nauwkeurigheid over om van een zinvol resultaat te kunnen spreken.

De β_{pq} werden berekend in 3 à 4 decimalen en geven een bevredigende aansluiting aan de gegeven waarden, zoals uit onderstaande voorbeelden blijkt. (De vierde decimaal is dus niet verantwoord geweest).

A_t	25	25	25	25	25	5
z	4.6	9.4	14.7	21.3	36.0	36.0
$P_{44}(A_t,z)$	3.696	3.691	3.686	3.687	3.702	0.699
$u(A_t,z)$	3.7	3.7	3.7	3.7	3.7	0.7

Het is nu wel duidelijk, waarom in 2. geen resultaat verkregen kon worden: de coefficienten β_{pq} hebben pas significante cijfers na de 5 decimalen, die vroeger berekend waren.

Zo is bijv. $\beta_{00} = -0.0586 = -0.000000586 . 10^5$

terwijl gevonden werd $\quad 0.00000 \quad .10^5$.

Waren indertijd drie decimalen meer meegenomen, dus was ε in 10 in plaats van in 7 decimalen geschreven, dan zou er inderdaad een uitkomst verkregen zijn. Er is dus wel zeer veel van de precisie in dat proces verloren gegaan: om dezelfde nauwkeurigheid te bereiken als bij de herhaalde berekening, zouden 5 cijfers meer nodig geweest zijn!

5. Uit het bovenstaande blijkt wel, dat de nauwkeurigheid, waarmee matrix-processen verricht kunnen worden, zeer sterk van de bouw van de matrix afhangt. Om dit quantitatief te kunnen uitdrukken, is het begrip "conditie van een matrix" ingevoerd (Turing, A.M., 1948, Quart.J.Mech.Applied Math., 1, 287, § 7).

Voor de conditie kan men als "maat" definieren:

a) het conditiegetal N

$$N = \frac{1}{n} N(A)\, N(A^{-1})$$

waarin n = orde van de matrix

$$N(A) = \text{norm van } A = \Big(\sum_{ij} a_{ij}^2 \Big)^{\frac{1}{2}} .$$

Heeft men op te lossen de vergelijkingen

$$A x_0 = b$$

maar gebruikt men i.p.v. de exacte matrix A de benadering $S + A$, waarbij dus S de matrix der afrondingsfouten is, dan geldt voor de oplossing x van

$$(S+A)x = b$$

dat, in eerste benadering

$$x = x_0 - A^{-1} S x_0$$

$A^{-1} S x_0$ is dus de kolom der fouten in de oplossing.

Neemt men het gemiddelde hiervan over de elementen van matrix en oplossing en over alle mogelijke matrices S, dan vindt men:

(9) $$\frac{\text{gem.fout in coeff.van oplossing}}{\text{gem.coeff.van opl.}} = N \, \frac{\text{gem.fout in element van } A}{\text{gem. element in } A}$$

waarbij met "gemiddelde" bedoeld wordt de wortel uit het gemiddelde kwadraat.

Met andere woorden: de relatieve statistische fout wordt N keer zo groot.

b) het conditiegetal M

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

waarin $M(A) = \max |a_{i,j}|$.

Dit conditiegetal is voor grote matrices vlugger te berekenen dan N en loopt enigszins parallel aan N. Er is echter niet zo gemakkelijk een statistische betekenis aan toe te kennen.

Ter illustratie volgen hier enige conditiegetallen N

orthogonale matrix	$N = 1$	(de best mogelijke)
"willekeurige" matrix	$N = \sqrt{n}$	
singuliere matrix	$N = \infty$	(de slechtst mogelijke)

enkele matrices van orde 2:

$$A_1 = \begin{Vmatrix} 1 & 1-\delta \\ 1-\delta & 1 \end{Vmatrix} \qquad N = \frac{2-2\delta+\delta^2}{\delta(2-\delta)} \qquad \begin{aligned} &= 1 \text{ voor } \delta = 1 \\ &= \infty \text{ voor } \delta = 0 \end{aligned}$$

$$A_2 = \begin{Vmatrix} a & b \\ b & c \end{Vmatrix} \qquad N = \frac{a^2+2b^2+c^2}{2|b^2-ac|}$$

Als $ac-b^2 = \text{Det}\,(A_2) = 0$ is A_2 singulier en inderdaad $N = \infty$

$$A_3 = \left\| \begin{matrix} a & bn \\ bn & cn^2 \end{matrix} \right\| \qquad N = \frac{a^2n^{-2}+2b^2+c^2n^2}{2|b^2-ac|}$$

Voor a, b, c $\cong 1$

en $n \gg 1$ is A_3 van het in 2. beschreven type. Bij benadering is dan $N \cong n^2$ zodat A_3 slechter van conditie is dan A_2.
Anderzijds kan men berekenen, welke n de conditie zo goed mogelijk maakt. Dit blijkt het geval te zijn voor $n = \sqrt{\left|\frac{a}{c}\right|}$, waarbij

$$N = \frac{|ac|+|b^2|}{|ac-b^2|} .$$

Voor de in 2, 3 en 4 beschreven matrices gelden de volgende conditiegetallen:

	N	M
B de (5 x 5) matrix voor transformatie	$0.806 \cdot 10^{12}$	$1.893 \cdot 10^{13}$
B^* " " " na "	$2.078 \cdot 10^{5}$	$3.045 \cdot 10^{6}$
A " (15 x 15) " voor "	$2 \cdot 10^{15}$	$4 \cdot 10^{17}$
A^* " " " na "	$2.175 \cdot 10^{7}$	$1.176 \cdot 10^{9}$

In beide gevallen blijkt dus door de transformatie van rijen en kolommen met machten van 10 de conditie veel beter te zijn geworden.

Berekent men de statistische fout in de uitkomst, dan vindt men volgens de formule (9):

relatieve fout in β_{pq} berekend met oorspronkelijke matrix $2 \cdot 10^7$
" " getransformeerde " 1 .

Dat de werkelijke nauwkeurigheid beter is dan deze waarschijnlijke waarden, blijkt wel uit de tabel in 4, doch het verschil in nauwkeurigheid van beide rekenprocessen, blijkt hieruit wel duidelijk.

6. Conclusie.

Wanneer men bij het oplossen van

$$A \cdot X = C$$

een matrix A heeft, waarvan de elementen van zeer uiteenlopende grootte-orden zijn, verdient het aanbeveling, in plaats hiervan op te lossen de vergelijking

$$A^* \cdot X^* = C^* \quad \text{waarin} \quad A^* = M \cdot A \cdot M$$
$$X^* = M^{-1} X$$
$$C^* = M \cdot C$$

De oplossing is dan $X = M \cdot A^{*-1} \cdot C^*$

M wordt zo gekozen, dat de elementen van A^* van dezelfde orde van grootte zijn. De voordelen hiervan zijn:

a) men kan met een constant aantal decimalen werken en behoeft geen machten van 10 meer neer te schrijven.
b) de checks kunnen nauwkeuriger uitgevoerd worden
c) daar de conditie van A^* beter is dan die van A wordt de precisie van het rekenproces groter en kan men, van minder decimalen uitgaande, toch een nauwkeurige einduitkomst krijgen

7. Opmerkingen.

Voor de topstroking van de twee functies $u(A_t,z)$ werd uitgegaan van een algemeen vierdegraadspolynoom $P_{44}(A_t,z)$ met 15 coefficienten. Tegenover het voordeel van eenzelfde A voor al dergelijke functies staat echter, dat bij gebruikmaking van de eigenschappen van deze gegeven functies, misschien met minder coefficienten volstaan had kunnen worden.

Een der grootheden f is gelijk aan nul voor $z = 36$, terwijl voor $A_t \to 0$, $H_t \to 0$ de functie blijkbaar ook $\to 0$ is. Bovendien schijnt

$$\lim_{H_t \to 0} \left(\frac{\partial f}{\partial H_t}\right)_{A_t} = 0$$

zodat de volgende vorm geprobeerd zou kunnen worden, met 6 coefficienten:

$$P_{44}(A_t,z) = A_t(z-36)\left\{(z+36)(a_0 + a_1 A_t + a_2 A_t^2) + \right.$$
$$\left. + x^2 (b_{00} + b_{10} A_t + b_{01} z)\right\}.$$

8. Op de volgende pagina's treft men aan de matrices B, B^{-1}, d met het eerste stel oplossingen y ; B^*, B^{*-1}, d^* en het hiermee verkregen stel coefficienten y ; A, A^{-1} en C , waarbij geen oplossing verkregen werd; A^*, A^{*-1}, C^* met de kolom X van de hiermee verkregen coefficienten β_{pq}.

B

21	315	5495	10552.10	21533.10^{2}
315	5495	10552.10	21533.10^{2}	45734.10^{3}
5495	10552.10	21533.10^{2}	45734.10^{3}	99888.10^{4}
10552.10	21533.10^{2}	45734.10^{3}	99888.10^{4}	22268.10^{6}
21533.10^{2}	45734.10^{3}	99888.10^{4}	22268.10^{6}	50421.10^{7}

B^{-1}

7.50710	-1.86322	$1.49142.10^{-1}$	$-4.5999.10^{-3}$	$0.446302.10^{-4}$
-1.86322	$5.72202.10^{-1}$	$-5.90700.10^{-2}$	$5.0621.10^{-4}$	$-3.76062.10^{-5}$
$1.49142.10^{-1}$	$-5.90700.10^{-2}$	$7.6151.10^{-3}$	$-38.8742.10^{-5}$	$6.8033.10^{-6}$
$-0.45999.10^{-2}$	$2.50621.10^{-3}$	$-38.8742.10^{-5}$	$22.2802.10^{-6}$	$-4.21536.10^{-7}$
$0.446302.10^{-4}$	$-3.76062.10^{-5}$	$68.0331.10^{-7}$	$-42.1536.10^{-8}$	$8.3613.10^{-9}$

d

44.5000	$7.846.10^{2}$	$15.1818.10^{3}$	$3.11430.10^{5}$	$6.63935.10^{6}$

y

0.19621	$0.75730.10^{-1}$	$0.49910.10^{-2}$	$-0.14991.10^{-3}$	$0.13518.10^{-5}$

B^*

2.1000	3.1500	5.4950	1.0552	2.1533
3.1500	5.4950	10.5525	2.1533	4.5734
5.4950	10.5525	21.5329	4.5734	9.9888
1.0552	2.1533	4.5734	0.9989	2.2268
2.1533	4.5734	9.9888	2.2268	5.0421

B^{*-1}

136.3990	- 374.5096	342.8447	-1277.2048	166.3100
- 374.5096	1126.5573	- 1134.1242	4677.9644	- 681.0923
342.8447	-1134.1242	1260.9176	-5727.6937	913.8963
-1277.2048	4677.9644	- 5727.6937	28286.0852	-4842.9840
166.3100	- 681.0923	913.8963	-4842.9840	875.3195

d^*

4.4500	7.8460	15.1818	3.1143	6.6393

y

0.1560	0.09625	0.001945	$0.3786 \cdot 10^{-4}$	$-0.1762 \cdot 10^{-5}$

A

$5,5.10^{1}$	$9,75\ .10^{2}$	$9,0947134.10^{2}$	$1,8925\ .10^{4}$	$1,5549715.10^{4}$	1,9
	$1,8925.10^{4}$	$1,5549715.10^{4}$	$3,88875\ .10^{5}$	$2,9666928.10^{5}$	3,2
		$1,9464302.10^{4}$	$2,9666928.10^{5}$	$3,2443908.10^{5}$	4,9
			$8,299165\ .10^{6}$	$6,0445013.10^{6}$	6,1
				$6,1319162.10^{6}$	8,0
					1,4

A^{-1}

$1,645829.10^{3}$				
$-3,680318.10^{2}$	$8,981053.10^{1}$			
$-4,144521.10^{1}$	$3,913439$	$5,1031154$		
$2,841637.10^{1}$	$-7,214540$	$-1,193176\ .10^{-1}$	$5,926308\ .10^{-1}$	
$1,0769268.10^{1}$	$-2,314237$	$-3,341254\ .10^{-1}$	$1,715324\ .10^{-1}$	7,743
$-2,138951$	$8,493362.10^{-1}$	$-2,375879\ .10^{-1}$	$-7,639927\ .10^{-2}$	-1,204
$-9,577838.10^{-1}$	$2,489671.10^{-1}$	$4,454455\ .10^{-4}$	$-2,078998\ .10^{-2}$	-5,559
$-5,263253.10^{-1}$	$1,254314.10^{-1}$	$7,374579\ .10^{-3}$	$-9,774381\ .10^{-3}$	-3,629
$-6,915704.10^{-2}$	$5,404628.10^{-3}$	$9,219076\ .10^{-3}$	$-1,9690852.10^{-5}$	-6,6054
$1,043362.10^{-1}$	$-3,273100.10^{-2}$	$4,971001\ .10^{-3}$	$2,767966\ .10^{-3}$	6,8076
$1,205038.10^{-2}$	$-3,18257\ .10^{-3}$	$2,05291\ .10^{-5}$	$2,69276\ .10^{-4}$	6,7476
$7,75369\ .10^{-3}$	$-1,93181\ .10^{-3}$	$-5,1000\ .10^{-5}$	$1,54005\ .10^{-4}$	5,1883
$2,70442\ .10^{-3}$	$-5,67936\ .10^{-4}$	$-9,2054\ .10^{-5}$	$4,0704\ .10^{-5}$	2,0961
$-5,9555\ .10^{-4}$	$2,6649\ .10^{-4}$	$-8,6849\ .10^{-5}$	$-2,4991\ .10^{-5}$	-2,3336
$-1,1964\ .10^{-3}$	$3,5104\ .10^{-4}$	$-3,9120\ .10^{-5}$	$-2,9005\ .10^{-5}$	-8,1756

C

$1,395.10^{2}$	$2,7274000.10^{3}$	$2,2221231.10^{3}$	$5,6326400.10^{4}$	$4,2740528.10^{4}$	4,

A^*

5,500000	9,750000	9,094713	1,892500	1,554972	1,946430	3,888750
	18,925000	15,549715	3,888750	2,966693	3,244391	8,299165
		19,464302	2,966693	3,244391	4,959831	6,044501
			0,829916	0,604450	0,613192	1,819129
				0,613192	0,809990	1,284463
					1,416308	1,246038
						4,067135

A^{*-1}

49,0324					
6702,3956	-24037,7469				
- 4828,4324	4577,2090	5957,5732			
-78801,5759	224260,4322	-14179,0269	-1921369,429		
10759,6995	27108,2576	-38907,3030	- 441374,7142	198896,1262	
32571,5419	-49607,6904	-27920,9010	331793,5482	211520,6983	118
32536,7413	-84652,9151	751,5449	688759,2502	205484,2417	-107
5873,0718	-28042,7816	8678,6922	296271,1044	- 13263,8844	- 64
- 9378,6935	9937,3329	10683,8560	- 34737,9911	- 82310,9137	- 43
- 8103,0780	14101,3109	5916,0146	- 106268,8277	- 45471,1135	- 25
- 4641,5402	11521,9064	227,7282	- 90363,2888	- 31436,6445	13
- 1647,8569	5213,6125	- 614,1399	- 49299,3062	- 4367,4854	7
439,1179	388,0687	- 1073,8921	- 9880,2324	7421,1223	4
1198,0873	- 1816,6213	- 1003,1204	11606,8684	8605,0761	3
7096,3273	-13220,8638	- 4742,2074	105521,5630	35037,1306	22

C^*

13,950000	27,274000	22,221231	5,632640	4,274053	4,635541	12,064054