

LINEAIRE ALGEBRA 1

Arno van den Essen
met aanvullingen door
Bernd Souvignier

voor het 1e jaar wiskunde en natuurkunde

Instituut WiNSt
Radboud Universiteit Nijmegen

Inhoudsopgave

1	Stelsels lineaire vergelijkingen	5
1.1	Inleiding	5
1.2	Gauss-eliminatie	6
1.3	Toepassingen	9
2	Matrices	19
2.1	Inleiding	19
2.2	Matrix rekening	19
2.3	Toepassingen	25
2.4	Inverteerbare matrices	30
3	Determinanten	41
3.1	Inleiding	41
3.2	Definitie en eigenschappen van de determinant	44
3.3	Toepassingen	51
4	Eigenwaarden en eigenvectoren	57
4.1	Inleiding	57
4.2	Definitie van eigenwaarden en eigenvectoren	57
4.3	Toepassingen	60

Hoofdstuk 1

Stelsels lineaire vergelijkingen

1.1 Inleiding

Op het eind van de 18e eeuw had men het vermoeden dat er tussen Mars en Jupiter een tot dan toe onbekende planeet moest zijn. Op 1 januari 1801 vond de Siciliaanse astronoom Giuseppe Piazzi (1746-1826) deze “planeet” die hij *Ceres* noemde ter ere van de beschermheilige van Sicilië. Het grote publiek was erg geïnteresseerd in deze ontdekking. In die tijd was het aantal planeten in ons zonnestelsel nog steeds onderwerp van debatten van filosofen en kerkleiders (de planeten Neptunus en Pluto werden pas in 1846 resp. 1930 ontdekt). Zeer recentelijk (in de zomer van 2006) haalde de discussie over het aantal planeten ook weer een groot publiek, toen tijdens een conferentie van astronomen Pluto zijn status als planeet kwijt raakte en vanaf nu officieel een *dwerfplaneet* is. Aanleiding hiervoor was de vondst van het hemelsobject *Eris* (voormalig 2003 *UB₃₁₃* en inofficieel *Xena*) in het jaar 2005. Eris beweegt op een baan buiten de baan van Pluto rond de zon en heeft zelfs een grotere diameter dan Pluto. Men kwam uiteindelijk overeen, objecten zo als Pluto en Eris nu dwergplaneten te noemen, en ook Ceres die tot nog toe een astroïde of planetoïde heette, valt met zijn diameter van 950 km in deze categorie.

Piazzi kon de baan van Ceres 40 dagen volgen, maar toen raakte men Ceres kwijt omdat Piazzi ziek werd. Gauss¹, die toen 24 was, kon echter op grond van een paar waarnemingen de baan van Ceres berekenen. Zijn berekeningen waren zó precies dat de Duitse astronoom Wilhelm Olbers (1758-1840) de astroïde op 31 december 1801 terug vond!

Om de baan van Ceres te berekenen gebruikte Gauss de eerste wet van Kepler² die zegt dat iedere planeet zich in een ellipsvormige baan rond de zon beweegt met de zon in één der brandpunten. De algemene vergelijking van een ellips wordt gegeven door de formule

$$ax^2 + bxy + cy^2 + dx + ey + f = 0,$$

d.w.z. de punten (x, y) die aan deze vergelijking voldoen, liggen op de ellips en omgekeerd.

Door het invullen van 6 waarnemingen d.w.z. 6 punten (x_i, y_i) vind je dan 6 (lineaire) vergelijkingen in de onbekenden a, b, c, d, e en f . Om dit stelsel vergelijkingen op te lossen ontwikkelde Gauss een systematische methode die nu Gauss-eliminatie genoemd wordt. Deze is het onderwerp van de volgende paragraaf (zie ook opgave 1.2.7, waar zal blijken dat zelfs 5 waarnemingen voldoende zijn!).

¹Carl Friedrich Gauss, 1777-1855, Duitse wiskundige

²Johannes Kepler, 1571-1636, Duitse wiskundige en astronoom

1.2 Gauss-eliminatie

Een rechte lijn in het xy -vlak kan worden voorgesteld door een vergelijking van de vorm $a_1x + a_2y = b$, waarbij a_1, a_2 en b reële getallen zijn en a_1 en a_2 niet beide nul. Een vergelijking van deze vorm heet een *lineaire vergelijking* in de variabelen x en y . De term *lineair* geeft hierbij aan dat in de vergelijking geen producten van de variabelen zo als x^2 , y^3 of xy voorkomen. De ellips vergelijking uit de inleiding is dus niet lineair in x en y , maar wel in de onbekenden a, b, c, d, e, f .

Definitie 1.2.1

- i) Een *lineaire vergelijking* in n variabelen $x_1, \dots, x_n$ is een vergelijking van de vorm

$$a_1x_1 + \dots + a_nx_n = b$$

waarbij $a_1, \dots, a_n, b$ reële getallen zijn. De variabelen $x_1, \dots, x_n$ noemen we ook wel *onbekenden*.

- ii) Een eindig aantal lineaire vergelijkingen in de variabelen $x_1, \dots, x_n$ heet een *stelsel* (of systeem) *lineaire vergelijkingen* in de variabelen $x_1, \dots, x_n$.
 iii) Een n -tal getallen $(s_1, \dots, s_n)$ heet een *oplossing* van het stelsel als $x_1 = s_1, \dots, x_n = s_n$ een oplossing is van iedere vergelijking van dit stelsel.

Voorbeeld 1.2.2

- i) Het stelsel lineaire vergelijkingen

$$4x_1 - x_2 + 3x_3 = -1$$

$$3x_1 + x_2 + 9x_3 = -4$$

heeft $(1, 2, -1)$ als oplossing omdat deze waarden aan beide vergelijkingen voldoet. Daarentegen is $(1, 8, 1)$ géén oplossing omdat deze waarde alleen aan de eerste vergelijking voldoet.

- ii) De vergelijking $x_1^2 + x_2^2 - 1 = 0$ (die een cirkel van straal 1 beschrijft) is geen lineaire vergelijking.

Ons doel in deze paragraaf is een methode te beschrijven (de reeds eerder genoemde Gauss-eliminatie) om bij een willekeurig stelsel van $m \geq 1$ lineaire vergelijkingen in $n \geq 1$ variabelen alle oplossingen te vinden. M.a.w. we zullen de volgende stelsels bekijken

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

$$\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m$$

waarbij de a_{ij} en de b_j reële getallen zijn.

Vraag: Hoe vind je een systematische methode om zulke stelsels lineaire vergelijkingen op te lossen?

Antwoord: Bestudeer eerst eenvoudige getallen en leer ervan!

Het eenvoudigste geval is natuurlijk $n = 1$ en $m = 1$ m.a.w. de vergelijking $ax = b$. Hieraan zien we al iets: als $a \neq 0$ heeft de vergelijking precies één oplossing, namelijk $x = b/a$.

Echter als $a = 0$ zijn er 2 gevallen, namelijk als $b = 0$ in welk geval iedere $x \in \mathbb{R}$ een oplossing is en het geval $b \neq 0$ in welk geval er geen enkele oplossing $x \in \mathbb{R}$ bestaat!

Hoe dan ook, het geval $n = 1$, $m = 1$ begrijpen we volkomen.

Een volgend geval is $m = 1$ en $n = 2$: $a_1x_1 + a_2x_2 = b$, waarbij we $a_1 \neq 0$ aannemen (wat gebeurt er als $a_1 = 0$?). Kies dan x_2 willekeurig in $\mathbb{R}$, zeg $x_2 = t$. Dan volgt $a_1x_1 = b - a_2t$ en dus $x_1 = \frac{b-a_2t}{a_1}$. Dus is de algemene oplossing (s_1, s_2) van $a_1x_1 + a_2x_2 = b$ van de vorm

$$(s_1, s_2) = \left(\frac{b - a_2t}{a_1}, t \right) = \left(\frac{b}{a_1}, 0 \right) + t \left(\frac{-a_2}{a_1}, 1 \right)$$

met $t \in \mathbb{R}$ willekeurig. De vrije variabele t heet een *parameter*.

Opgave 1.2.3 Zij $n \geq 2$, $a_1, \dots, a_n, b \in \mathbb{R}$ en $a_1 \neq 0$. Vind alle oplossingen van

$$a_1x_1 + \dots + a_nx_n = b.$$

Het volgende geval wat we willen onderzoeken is $n = 2$ en $m = 2$. Bijvoorbeeld

$$\begin{aligned} x + y &= 4 \\ 2x + 5y &= 11 \end{aligned}$$

Het *cruciale* idee is nu de $2x$ term uit de tweede vergelijking te elimineren door -2 keer de eerste vergelijking bij de tweede vergelijking op te tellen! Het resultaat is dat we het volgende nieuwe stelsel vinden

$$\begin{aligned} x + y &= 4 \\ 3y &= 3 \end{aligned}$$

Zoals men eenvoudig nagaat is iedere oplossing van het eerste stelsel een oplossing van het nieuwe stelsel en omgekeerd. We kunnen dus net zo goed het nieuwe stelsel proberen op te lossen. Echter dit stelsel is *eenvoudiger*! Immers de laatste vergelijking geeft meteen $y = 1$ en dit invullend in de eerste vergelijking geeft $x = 3$.

Dit voorbeeld laat zien dat we het geval $n = 2$, $m = 2$ aankunnen! Maar ook het geval $m = 2$, $n \geq 2$ d.w.z. 2 vergelijkingen met 2 of meer variabelen is nu te doen, zoals we aan de hand van het volgende voorbeeld laten zien.

Voorbeeld 1.2.4 Zoek alle oplossingen van het stelsel

$$\begin{aligned} x + y + 2z &= 9 \\ 3x + 6y - 5z &= 0 \end{aligned}$$

Oplossing. Om de $3x$ -term in de tweede vergelijking kwijt te raken tellen we -3 keer de eerste vergelijking bij de tweede op. Dit levert het volgende stelsel vergelijkingen op

$$\begin{aligned} x + y + 2z &= 9 \\ 3y - 11z &= -27 \end{aligned}$$

Kies nu in de tweede vergelijking voor z een willekeurige waarde, zeg $z = t \in \mathbb{R}$. Dan volgt $3y = -27 + 11t$ m.a.w. $y = -9 + \frac{11}{3}t$. Deze waarden van y en z invullend in de eerste vergelijking geeft

$$x + (-9 + \frac{11}{3}t) + 2t = 9 \text{ m.a.w. } x = 18 - \frac{17}{3}t.$$

Dus de algemene oplossing is: $(s_1, s_2, s_3) = (18 - \frac{17}{3}t, -9 + \frac{11}{3}t, t)$ met $t \in \mathbb{R}$ willekeurig. Door voor $t = 3$ te kiezen vinden we de “mooie” oplossing $(1, 2, 3)$.

Het zal de lezer nu misschien zijn duidelijk geworden hoe het verder gaat. Voor alle duidelijkheid nog een voorbeeld met $n = 3$ en $m = 3$.

Voorbeeld 1.2.5 Zoek alle oplossingen van het stelsel

$$\begin{aligned} x + y + 2z &= 9 \\ 2x + 4y - 3z &= 1 \\ 3x + y + z &= 8 \end{aligned}$$

Oplossing. Om de $2x$ -term uit de tweede vergelijking weg te werken, tellen we -2 keer de eerste vergelijking bij de tweede op. Dit levert

$$\begin{aligned} x + y + 2z &= 9 \\ 2y - 7z &= -17 \\ 3x + y + z &= 8 \end{aligned}$$

Om de $3x$ -term uit de derde vergelijking weg te werken tellen we -3 keer de eerste vergelijking bij de derde op en vinden

$$\begin{aligned} x + y + 2z &= 9 \\ 2y - 7z &= -17 \\ -2y - 5z &= -19 \end{aligned}$$

De laatste twee vergelijkingen zijn nu 2 vergelijkingen in 2 onbekenden. Die kunnen we dus oplossen! Daarvoor tellen we de tweede vergelijking op bij de derde en vinden

$$\begin{aligned} x + y + 2z &= 9 \\ 2y - 7z &= -17 \\ -12z &= -36 \end{aligned}$$

Nu kunnen we weer van onder naar boven oplossen en vinden $z = 3$, $y = 2$ (uit $2y - 21 = -17$) en tenslotte $x = 1$ (uit $x + 2 + 6 = 9$).

Opmerking 1.2.6 Soms is het handig vergelijkingen te verwisselen of een vergelijking met een constante $\neq 0$ te vermenigvuldigen. Bekijk bijvoorbeeld het volgende stelsel

$$\begin{aligned} 2y - 8z &= 8 \\ x - 2y + z &= 0 \\ -4x + 5y + 9z &= 0 \end{aligned}$$

Verwissel eerst de eerste vergelijking met de tweede met als doel de x -term in de eerste vergelijking te krijgen waarmee je dan vervolgens de $-4x$ term uit de derde vergelijking kunt wegwerken. Daarna kun je de tweede vergelijking met $\frac{1}{2}$ vermenigvuldigen zodat deze $y - 4z = 4$ wordt. Met deze vergelijking kun je dan de $5y$ -term uit de laatste vergelijking wegwerken. Ga dit zelf na en los het stelsel op!

Samengevat: In bovenstaande Gauss-eliminatie methode hebben we 3 typen bewerkingen gebruikt, nl.

- 1) Tel bij een vergelijking een veelvoud van een andere vergelijking op.
- 2) Vermenigvuldig een vergelijking met een constante $\neq 0$.
- 3) Verwissel twee vergelijkingen.

Opgave 1.2.7 Bepaal de ellips die gaat door de volgende 5 punten: $(2, -1)$, $(2, -5)$, $(5, -3)$, $(-1, -3)$ en $(1, \frac{4}{3}\sqrt{2} - 3)$.

Opgave 1.2.8 Vertaal het volgende puzzeltje in een stelsel lineaire vergelijkingen en vind de oplossing.

- Vandaag is Annie twee jaar jonger dan Ben en Cees samen.
- Over vijf jaar is Annie twee keer zo oud als Ben.
- Twee jaar geleden was Ben half zo oud als Cees.

Hoe oud zijn Annie, Ben en Cees?

1.3 Toepassingen

Heel veel problemen in allerlei takken van wetenschap zijn te herleiden tot het oplossen van stelsels lineaire vergelijkingen. We zullen in deze paragraaf een aantal voorbeelden bekijken.

a) Grafieken door voorgeschreven punten

Stel je hebt twee verschillende reële getallen c_1 en c_2 en twee reële getallen d_1 en d_2 . Dan is het bekend dat er precies één rechte van de vorm $y = a_0 + a_1x$ door de punten (c_1, d_1) en (c_2, d_2) gaat. Iets minder bekend is het feit dat er door drie punten (c_1, d_1) , (c_2, d_2) , (c_3, d_3) met c_1, c_2, c_3 allen verschillend precies één parabool van de vorm $y = a_0 + a_1x + a_2x^2$ gaat. Deze resultaten vormen slechts een speciaal geval van de navolgende stelling die we verderop zullen bewijzen.

Stelling 1.3.1 Voor n verschillende reële getallen $c_1, \dots, c_n$ en n willekeurige reële getallen $d_1, \dots, d_n$ is er precies een veelterm $a_0 + a_1x + \dots + a_{n-1}x^{n-1}$ van graad $\leq n-1$ met $a_1, \dots, a_{n-1}$ in $\mathbb{R}$ zodat de grafiek $y = a_0 + a_1x + \dots + a_{n-1}x^{n-1}$ door de punten $(c_1, d_1), \dots, (c_n, d_n)$ gaat.

Het vinden van deze kromme kunnen we m.b.v. Gauss-eliminatie doen. We laten dit aan de hand van een voorbeeld zien.

Voorbeeld 1.3.2 Bepaal een veelterm $a_0 + a_1x + a_2x^2$ van graad 2 zodanig dat de grafiek $y = a_0 + a_1x + a_2x^2$ door de punten $(1, 6)$, $(2, 3)$ en $(3, 2)$ gaat.

Oplossing. Omdat $x = 1$, $y = 6$ op de kromme ligt vinden we dat $a_0 + a_1 + a_2 = 6$. Net zo vinden we door substitutie van $x = 2$ en $y = 3$ dat $a_0 + 2a_1 + 4a_2 = 3$ en tenslotte

$a_0 + 3a_1 + 9a_2 = 2$ omdat $(3, 2)$ op de kromme ligt. We moeten dus a_0, a_1, a_2 oplossen uit het stelsel

$$\begin{aligned} a_0 + a_1 + a_2 &= 6 \\ a_0 + 2a_1 + 4a_2 &= 3 \\ a_0 + 3a_1 + 9a_2 &= 2 \end{aligned}$$

Door de eerste van de tweede en de eerste van de derde vergelijking af te trekken vinden we

$$\begin{array}{rcl} a_0 + a_1 + a_2 &= & 6 \\ a_1 + 3a_2 &= & -3 \\ 2a_1 + 8a_2 &= & -4 \end{array} \quad \text{ofwel} \quad \begin{array}{rcl} a_0 + a_1 + a_2 &= & 6 \\ a_1 + 3a_2 &= & -3 \\ a_1 + 4a_2 &= & -2 \end{array}$$

Door nu nog de tweede van de derde vergelijking af te trekken vinden we

$$\begin{aligned} a_0 + a_1 + a_2 &= 6 \\ a_1 + 3a_2 &= -3 \\ a_2 &= 1 \end{aligned}$$

Dus $a_2 = 1$, $a_1 = -3 - 3 = -6$ en $a_0 = 6 - 1 - (-6) = 11$. De gezochte kromme is dus $y = 11 - 6x + x^2$.

Opmerking 1.3.3 (De Lagrange³ interpolatie polynomen.)

Het unieke polynoom uit stelling 1.3.1 kan ook als volgt gevonden worden: bekijk $p_i(x) = (x - c_1) \dots \widehat{(x - c_i)} \dots (x - c_n)$, $1 \leq i \leq n$, waarbij $\widehat{}$ betekent dat de factor $x - c_i$ moet worden weggelaten. Dan geldt dat $p_i(c_j) = 0$ als $j \neq i$ en $p_i(c_i) = (c_i - c_1) \dots \widehat{(c_i - c_i)} \dots (c_i - c_n)$ welke $\neq 0$ is omdat $c_i \neq c_j$ als $j \neq i$. Voor iedere $1 \leq i \leq n$ definiëren we dat

$$L_i(x) := \frac{1}{p_i(c_i)} p_i(x).$$

Dan voldoet het polynoom

$$P_{c,d}(x) := d_1 L_1(x) + \dots + d_n L_n(x)$$

aan $P_{c,d}(c_i) = d_i$ voor iedere i (ga na!).

Het polynoom $P_{c,d}(x)$ heet een *Lagrange interpolatie polynoom*.

Tot slot van deze subparagraaf nog een toepassing die valt in de categorie “recreational mathematics”.

Gegeven een rijtje zoals $1, 2, 3, 4, 5, \dots$

Vraag: Maak het rijtje af en nog beter, geef een formule $f(n)$ die de n -de term van het rijtje beschrijft.

Bovenstaand rijtje is natuurlijk “duidelijk”, de volgende term is 6 en de formule is $f(n) = n$. Maar is dat ook zo? Antwoord: natuurlijk niet! Immers, volgens Opmerking 1.3.3 is het makkelijk een polynoom P te maken met graad ≤ 5 zodanig dat $P(1) = 1$, $P(2) = 2$, $P(3) = 3$, $P(4) = 4$, $P(5) = 5$ maar $P(6) = 321!$ namelijk neem in Opmerking 1.3.3

³Joseph Louis Lagrange, 1736-1813, Franse wiskundige

$(c_1, d_1) = (1, 1), (c_2, d_2) = (2, 2), \dots, (c_5, d_5) = (5, 5)$ en $(c_6, d_6) = (6, 321)$. Dan is $P := P_{c,d}$ zo'n polynoom!

Nu zul je misschien denken, ja maar bij een beetje “natuurlijk” probleem gebeurt zoiets niet. M.a.w. als je de regelmaat al na een paar termen ziet dan zal het wel zo doorgaan. Als je dat denkt bekijk dan de volgende vraag.

Vraag 1.3.4 Stel je hebt $n \geq 2$ verschillende punten op een cirkel. Als je alle n punten met elkaar verbindt via rechte lijnen (de zgn. koorden), in hoeveel stukken wordt de binnenkant van de cirkel dan maximaal verdeeld? Noem dit aantal $a(n)$.

Het is eenvoudig in te zien dat $a(2) = 2$, $a(3) = 4$ en $a(4) = 8$. Het lijkt dus duidelijk wat de regelmaat is, $a(5)$ zal 16 zijn. En inderdaad dat klopt! Nu zijn we helemaal zeker, een bewijs is “eigenlijk al niet meer nodig!” Maar dan $\dots a(6) = 31!!$ De verwachte formule $a(n) = 2^{n-1}$ klopt dus niet. Maar welke dan wel?

Het verrassende resultaat is dat $a(n)$ gegeven wordt door het unieke polynoom met graad ≤ 4 dat door de 5 punten $(2, 2), (3, 4), (4, 8), (5, 16)$ en $(6, 31)$ gaat.

Uitrekenen (bijvoorbeeld met Gauss-eliminatie of Opmerking 1.3.3) levert dat $a(n) = 1 + \frac{1}{2}n(n-1) + \frac{1}{4!}n(n-1)(n-2)(n-3)$ ($4! = 1.2.3.4$).

(Voor meer informatie zie: J. Glen, The Quest for the lost Region, Mathematics Teaching 43, pp. 23-25, 1968.)

b) Magische vierkanten (driehoeken, ...)

Bekijk het getallen vierkant

$$\begin{pmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{pmatrix}$$

Dit is een voorbeeld van een zgn. *magisch vierkant*. Het heet zo omdat het heel bijzondere eigenschappen heeft nl. de som van alle getallen in iedere rij = de som van alle getallen in iedere kolom = de som van alle getallen op ieder van de hoofddiagonalen = 15. Bovendien zijn de getallen in dit vierkant de opeenvolgende natuurlijke getallen beginnende met 1 tot het vierkant vol is.

Magische vierkanten werden voor het eerst ontdekt in 2200 voor Christus in China. Bovenstaand vierkant kreeg de naam Lo-Shu. Het eerste 4 bij 4 magische vierkant verscheen in 1100 in India:

$$\begin{pmatrix} 7 & 12 & 1 & 14 \\ 2 & 13 & 8 & 11 \\ 16 & 3 & 10 & 5 \\ 9 & 6 & 15 & 4 \end{pmatrix}$$

Dit vierkant is nog magischer omdat ook de “gebroken diagonalen” dezelfde som hebben zoals bijvoorbeeld $2 + 3 + 15 + 14 = 16 + 6 + 1 + 11 = 2 + 12 + 15 + 5$ enz.

Cornelius Agrippa (1486-1535) maakte magische vierkanten van orde 3, 4, 5, 6, 7, 8 en 9. Als astroloog verbond hij ze met de (toen) bekende 7 planeten!

In de 17e eeuw begonnen verschillende Franse wiskundigen methodes te ontwikkelen om magische vierkanten te maken. Rond 1838 waren er al zo'n 880 verschillende 4 bij 4 magische

vierkanten bekend (spiegelingen en draaiingen niet meegerekend). Inmiddels is er een gigantische literatuur over magische vierkanten en zijn er clubs die zich met magische vierkanten bezighouden; allen beweren 's wereld's meest magische magisch vierkant gemaakt te hebben (surf maar eens op het Web!).

I.p.v. te eisen dat de getallen in zo'n vierkant alleen de opvolgende natuurlijke getallen beginnende bij 1 zijn, zullen we *van nu af onderstellen* dat de getallen in het vierkant *willekeurige reële getallen zijn*.

De vraag die we zullen bekijken is: hoe vinden we al zulke 3 bij 3 magische vierkanten?

Voordat we ons met deze vraag gaan bezighouden bekijken we eerst een soortgelijk maar eenvoudiger probleem.

Magische driehoeken

Zij D de verzameling der driehoeken
$$\begin{array}{ccccc} & & a & & \\ & b & & f & \\ c & & d & & e \end{array}$$
 waarbij a, b, c, d, e en f reële getallen

zijn. Zo'n driehoek heet magisch als de som van alle getallen op iedere zijde hetzelfde is.

Bijvoorbeeld de driehoek

$$\begin{array}{ccccc} & & 1 & & \\ & 2 & & 6 & \\ 3 & & 4 & & -1 \end{array}$$

is magisch want de som der getallen op iedere zijde is 6.

Vraag. Hoe zien alle magische driehoeken eruit?

Antwoord. De driehoek
$$\begin{array}{ccccc} & & a & & \\ & b & & f & \\ c & & d & & e \end{array}$$
 is magisch d.e.s.d.a. de variabelen a, b, c, d, e, f aan het volgende stelsel lineaire vergelijkingen voldoen

$$\begin{aligned} a + b + c &= c + d + e \\ a + b + c &= a + f + e \end{aligned}$$

m.a.w. d.e.s.d.a. $a + b = d + e$ en $b + c = f + e$ waaruit volgt dat $b = f + e - c$ en $a = d + e - (f + e - c) = d + c - f$.

M.a.w. alle magische driehoeken zijn van de vorm

$$\begin{array}{ccccc} & & d + c - f & & \\ & f + e - c & & f & \\ c & & d & & e \end{array}$$

De magische driehoek
$$\begin{array}{ccccc} & & 1 & & \\ & 2 & & 6 & \\ 3 & & 4 & & -1 \end{array}$$
 vinden we terug door $c = 3$, $d = 4$, $e = -1$ en $f = 6$ te nemen.

Magische vierkanten

Laten we nu terug keren naar de 3 bij 3 magische vierkanten. Laten $x_1, x_2, \dots, x_9$ reële getallen zijn. Dan is een vierkant

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & x_8 & x_9 \end{pmatrix}$$

magisch d.e.s.d.a.

$$\begin{array}{llll} x_1 + x_2 + x_3 = x_4 + x_5 + x_6 & \text{én} & x_4 + x_5 + x_6 = x_7 + x_8 + x_9 & \text{én} \\ x_7 + x_8 + x_9 = x_1 + x_4 + x_7 & \text{én} & x_1 + x_4 + x_7 = x_2 + x_5 + x_8 & \text{én} \\ x_2 + x_5 + x_8 = x_3 + x_6 + x_9 & \text{én} & x_3 + x_6 + x_9 = x_1 + x_5 + x_9 & \text{én} \\ x_1 + x_5 + x_9 = x_3 + x_5 + x_7. & & & \end{array}$$

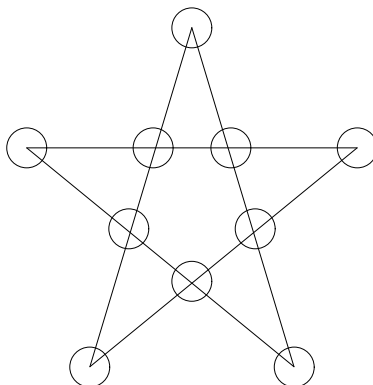
We krijgen dus een stelsel van 7 lineaire vergelijkingen in 9 onbekenden. M.b.v. Gauss-eliminatie is het dan mogelijk alle oplossingen te vinden. Na wat meer rekenwerk als in voorgaand voorbeeld blijkt dan dat alle 3 bij 3 magische vierkanten van de volgende vorm zijn

$$\begin{pmatrix} a - c & a + b + c & a - b \\ a - b + c & a & a + b - c \\ a + b & a - b - c & a + c \end{pmatrix}$$

Het beroemde magische vierkant Lo-Shu vinden we terug door $a = 5$, $b = 3$ en $c = 1$ te kiezen.

Natuurlijk is het met bovenstaande methode mogelijk ook alle 4 bij 4, 5 bij 5, ... magische vierkanten te vinden, alhoewel het rekenwerk wel zal toenemen. Ook kun je magische kubussen maken (de definitie ligt voor de hand). Tot slot:

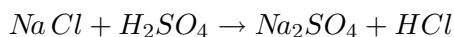
Opgave 1.3.5 Bepaal alle magische pentagrammen (waarbij natuurlijk de som van de vier getallen op ieder van de rechte lijnen hetzelfde moet zijn):



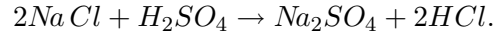
Figuur 1.1: Schema voor een magisch pentagram

c) Het in evenwicht brengen van chemische vergelijkingen

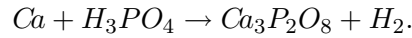
In de meeste voorbeelden van scheikunde op school is het in evenwicht brengen van chemische vergelijkingen een eenvoudige zaak. Zoals bij de reactie



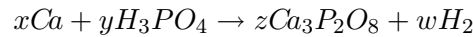
is het duidelijk dat de juiste reactie is



Laten we daarom een iets minder eenvoudig voorbeeld nemen:



We zoeken dus naar coëfficiënten, x, y, z, w natuurlijke getallen zodat de vergelijking



klopt. Door resp. de Ca , H , P en O atomen te vergelijken vinden we

$$x = 3z, 3y = 2w, y = 2z \text{ en } 4y = 8z.$$

Deze oplossen levert $x = 3$, $y = 2$, $z = 1$ en $w = 3$ (dit is de kleinste oplossing in de natuurlijke getallen).

Opgave 1.3.6 Breng de volgende vergelijkingen in evenwicht.

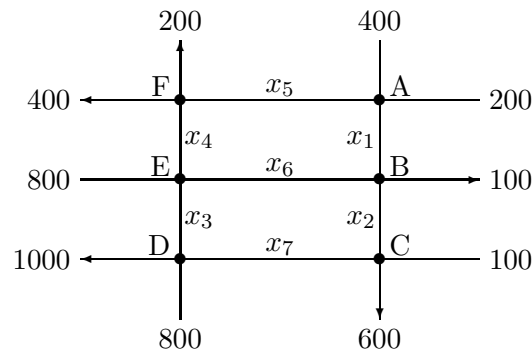
- i) $KMnO_4 + H_2SO_4 + KBr \rightarrow K_2SO_4 + Br_2 + MnSO_4 + H_2O$.
- ii) $As_2S_3 + H_2O + HNO_3 \rightarrow NO + H_3AsO_4 + H_2SO_4$.

d) Netwerk analyse

Onderstaand schema geeft aan de verkeersstroom in het centrum van Jacksonville in Florida. De getallen 400 enz. geven aan het aantal voertuigen per uur dat passeert in de aangegeven richting (alle straten zijn éénrichtings verkeer) gedurende de spitsuren. De aantallen voertuigen die passeren in de overige straten zijn aangegeven met $x_1, x_2, \dots, x_7$.

Als regel veronderstellen we:

- (*) al het verkeer dat een kruispunt nadert, verlaat het ook weer.



Figuur 1.2: Verkeersstromen in Jacksonville

Stel nu dat we op het traject tussen C en D werkzaamheden aan de weg moeten verrichten en we willen dat de verkeersstroom op dit traject zo minimaal mogelijk is. Wat is de kleinste mogelijke verkeersstroom op het traject CD zodat er geen verkeersopstoppen ontstaan en hoe kunnen we die gewenste toestand bereiken?

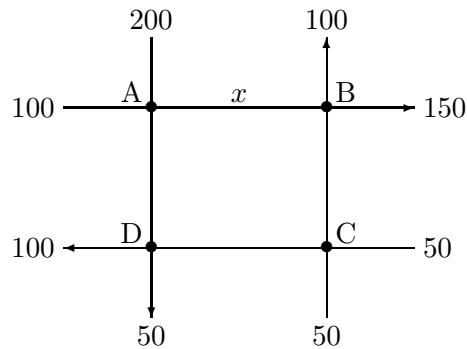
Oplossing. Bekijk wat uit (*) volgt op het kruispunt A : binnenkomend verkeer $400 + 200$, weggaand verkeer $x_1 + x_5$ m.a.w. we krijgen de vergelijking $x_1 + x_5 = 600$. Bij B vinden we: binnenkomend $x_1 + x_6$, uitgaand $x_2 + 100$ m.a.w. we krijgen de vergelijking $x_1 + x_6 = x_2 + 100$. Zo doorgaand vinden we het volgende stelsel vergelijkingen:

$$\begin{aligned} x_1 + x_5 &= 600, & x_1 - x_2 + x_6 &= 100, & x_2 - x_7 &= 500, \\ -x_3 + x_7 &= 200, & -x_3 + x_4 + x_6 &= 800, & x_4 + x_5 &= 600. \end{aligned}$$

M.b.v. Gauss eliminatie kunnen we dit stelsel oplossen en vinden dan:

$x_1 = -x_6 + x_7 + 600$, $x_2 = x_7 + 500$, $x_3 = x_7 - 200$, $x_4 = -x_6 + x_7 + 600$ en $x_5 = x_6 - x_7$. Omdat verkeersstromen ≥ 0 moeten zijn, zien we uit de derde vergelijking dat $x_7 \geq 200$ en dus is de stroom tussen C en D minimaal als $x_7 = 200$. In dat geval is $x_3 = 0$. Omgekeerd, als $x_3 = 0$ dan is $x_7 = 200$. M.a.w. als we het traject tussen E en D afsluiten dan is $x_3 = 0$ en dus $x_7 = 200$, de minimale verkeersstroom op het traject tussen C en D . De andere stromen zijn dan: $x_1 = -x_6 + 800$, $x_2 = 700$, $x_4 = -x_6 + 800$ en $x_5 = x_6 - 200$, waarbij $200 \leq x_6 \leq 800$.

Opgave 1.3.7 Alle straten in onderstaand schema zijn éénrichtings verkeer straten. Wat is de kleinste mogelijke verkeersstroom x op het traject tussen A en B ?



Figuur 1.3: Verkeersstromen

Historische opmerkingen

1. Traditioneel was algebra de kunst van het oplossen van vergelijkingen of stelsels vergelijkingen. Het woord algebra komt van het Arabische woord al-jabr wat reparatie (van gebroken delen) betekent! De term werd voor het eerst in een wiskundige betekenis gebruikt door Mohammed al-khowarizmi (780-850) die in het “Huis van Wijsheid” in Bagdad werkte (en van wiens naam het woord *algoritme* is afgeleid).

Lineaire algebra is de kunst van het oplossen van lineaire vergelijkingen of stelsels lineaire vergelijkingen.

2. Carl Friedrich Gauss (1777-1855) werd geboren in een arme arbeiders klasse familie in Brunswick (Duitsland) en stierf in Göttingen als de beroemdste wiskundige van de wereld!

Als kind was hij al uitzonderlijk: op een dag, hij was toen nog geen drie, was zijn vader bezig de lonen voor de arbeiders die onder zijn leiding werkten klaar te maken zat Gauss in een hoekje van de kamer. Na een lange berekening zei Gauss tegen z'n vader dat

er een fout in de berekening zat en gaf de juiste uitkomst, die hij uit z'n hoofd had uitgerekend. Tot grote verbazing van z'n ouders bleek zijn uitkomst juist te zijn!

Gauss leverde belangrijke bijdragen in o.a. getaltheorie, wiskundige astronomie, wiskundige geografie, statistiek, differentiaal meetkunde en magnetisme. In zijn dagboeken zijn nog heel veel ongepubliceerde resultaten te vinden. Hij wordt beschouwd als (een van) de grootste wiskundigen uit de moderne tijd.

3. Gauss-eliminatie. Deze methode voor het oplossen van stelsels lineaire vergelijkingen werd ruim twee duizend jaar geleden door de Chinezen gebruikt. In het achtste hoofdstuk van het Chinese werk “Boeken der negen hoofdstukken over de wiskunst” (Chin-chang Suan-Shu) rond 200 voor Christus kwam bijvoorbeeld het volgende probleem voor dat met “Gauss-eliminatie” werd opgelost:

De opbrengst van een schoof minderwaardig graan, twee schoven gemiddeld graan en drie schoven superieur graan is 39 tou (een bronzen schaal in de Chou dynastie, 900-225 voor Christus). De opbrengst van een schoof minderwaardig graan, drie schoven gemiddeld graan en twee schoven superieur graan is 34 tou. De opbrengst van drie schoven minderwaardig graan, twee schoven gemiddeld graan en een schoof superieur graan is 26 tou. Wat is de opbrengst van een schoof minderwaardig, gemiddeld en superieur graan?

4. Joseph Louis Lagrange (1736-1813) was een Frans-Italiaans wiskundige en astronoom geboren te Turijn. Hij bracht het grootste deel van zijn carrière door in Berlijn en Parijs. Lagrange werd aangetrokken door wiskunde en astronomie na het lezen van een werk van de astronoom Halley (1656-1742). Deze berekende in 1682 de baan van de na hem genoemde komeet en voorspelde dat hij in 1758 weer zou langskomen. Toen Lagrange 16 was, begon hij zichzelf wiskunde te leren en toen hij 19 was, werd hij Professor aan de Koninklijke artillerieschool te Turijn. Het volgende jaar loste hij een aantal beroemde problemen op met nieuwe methoden die uiteindelijk leidden tot een nieuwe tak van wiskunde: de variatie rekening.

Deze methode, en Lagrange's toepassing op de hemel mechanica waren zo indrukwekkend dat hij op 25-jarige leeftijd beschouwd werd als de grootste op dat moment levende wiskundige.

Een van zijn belangrijkste werken is “Mécanique Analytique” waarin hij de gehele mechanica terugbracht tot een paar formules waaruit alles kon worden afgeleid.

Napoleon (1769-1821) was een groot bewonderaar van Lagrange en overlaadde hem met onderscheidingen. Lagrange was een bescheiden man. Hij is met eer begraven in het beroemde Panthéon in Parijs.

Enkele vooruitblikken

1. In dit hoofdstuk hebben we gekeken naar stelsels lineaire vergelijkingen. Meer algemeen kan men kijken naar stelsels van de vorm

$$\begin{aligned} f_1(x_1, \dots, x_n) &= 0 \\ &\vdots \\ f_m(x_1, \dots, x_n) &= 0 \end{aligned}$$

waarbij iedere $f_i(x_1, \dots, x_n)$ een veelterm in $x_1, \dots, x_n$ met coëfficiënten in $\mathbb{R}$ (of $\mathbb{C}$) is d.w.z. een eindige som van termen van de vorm $c_{i_1, \dots, i_n} x_1^{i_1} \dots x_n^{i_n}$ waarbij $c_{i_1, \dots, i_n}$ een reëel (of complex) getal is en iedere $i_j \geq 0$. Dus bijvoorbeeld $7x_1^2x_2^2 + 3x_2^4 + x_1$ is een voorbeeld van een veelterm (niet lineair) in x_1 en x_2 .

De studie van zulke stelsels vergelijkingen gebeurt in de *algebraïsche meetkunde*. Recentelijk (1964) is er door de Oostenrijkse wiskundige Bruno Buchberger (1942-) een uitbreiding van de Gauss-eliminatie methode gevonden naar stelsels veelterm vergelijkingen; de bijbehorende theorie heet *Gröbner basis theorie*. Het idee is dat je door middel van een soort veepproces (*S*-polynomen) het stelsel vergelijkingen vervangt door een “mooier” stelsel veelterm vergelijkingen, waaraan je bijvoorbeeld de oplossingen eenvoudig kunt aflezen. Meer hierover kun je te weten komen in colleges over *Computer algebra*.

2. De Lagrange interpolatie polynomen kunnen als volgt gegeneraliseerd worden naar meerdere variabelen. Laat $c_1, \dots, c_n$ verschillende vectoren in $\mathbb{R}^m$ zijn en $d_1, \dots, d_n$ gegeven reële getallen. Dan kun je bewijzen dat er een polynoom $P_{c,d}(x_1, \dots, x_m)$ met coëfficiënten in $\mathbb{R}$ bestaat zodanig dat $P_{c,d}(c_i) = d_i$ voor iedere i . Dit resultaat is een speciaal geval van de zgn. Chinese reststelling, welke in het college *Ring en Lichamen* behandeld wordt.

Hoofdstuk 2

Matrices

2.1 Inleiding

In het vorige hoofdstuk behandelden we de Gauss-eliminatie methode waarmee we stelsels lineaire vergelijkingen leerden oplossen. We telden vergelijkingen bij anderen op enz.

Het is, zeker bij grote stelsels, erg onhandig om steeds de variabelen $x_1, \dots, x_n$ te moeten opschrijven. Immers, het enige wat er bij die bewerkingen verandert zijn de coëfficiënten die voor de $x_1, \dots, x_n$ staan en de b_i 's. Het is daarom efficiënter alleen de coëfficiënten en de kolom der b_i 's in een getallen rechthoek te zetten (matrix geheten) en met de rijen van die matrix te werken.

Echter matrices op zich blijken ook heel nuttige objecten te zetten. In dit hoofdstuk zullen we verschillende bewerkingen op matrices definiëren, zoals een optelling en een vermenigvuldiging. Het zal duidelijk worden dat ze algemeen toepasbaar zijn voor het bestuderen van problemen van heel uiteenlopende aard (archeologie, biologie, economie, enz.).

2.2 Matrix rekening

Definitie 2.2.1 Een $m \times n$ matrix A is een getallenrechthoek bestaande uit m rijen en n kolommen, die tussen twee haakjes ingeklemd is.

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

In de matrix staat het element a_{ij} in de i -de rij en j -de kolom m.a.w. de eerste index geeft aan in welke rij het element staat en de tweede index in welke kolom. Bovenstaande matrix noteren we ook als (a_{ij}) .

Voorbeeld 2.2.2 Zij A de 2×3 matrix $\begin{pmatrix} 5 & 0 & 1 \\ 8 & 7 & 6 \end{pmatrix}$. Dan $a_{21} = 8$, $a_{12} = 0$ en $a_{23} = 6$.

Zij $m, n \geq 1$. De verzameling der $m \times n$ matrices noteren we als $M_{m,n}(\mathbb{R})$ of $\text{Mat}_{m,n}(\mathbb{R})$ of $M_{mn}(\mathbb{R})$. Als $m = n$ schrijven we $\text{Mat}_n(\mathbb{R})$ of $M_n(\mathbb{R})$. I.p.v. $M_{n,1}(\mathbb{R})$ schrijven we $\mathbb{R}^n$. Dus $\mathbb{R}^n$ is de verzameling van alle kolommen van n -tallen reële getallen.

Twee matrices uit $M_{m,n}(\mathbb{R})$ kunnen we op de voor de hand liggende manier optellen, namelijk zij $A = (a_{ij})$ en $B = (b_{ij})$ in $M_{m,n}(\mathbb{R})$. We definiëren $A + B$ als de $m \times n$ matrix met op de plaats (i, j) het element $a_{ij} + b_{ij}$. Ook definiëren we voor $c \in \mathbb{R}$ de matrix cA als de $m \times n$ matrix met op de plaats (i, j) het element $c \cdot a_{ij}$. I.h.b. kunnen we dus de elementen uit $\mathbb{R}^n$ optellen en vermenigvuldigen met elementen uit $\mathbb{R}$.

Opgave 2.2.3 Zij $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$, allen in $\mathbb{R}^n$. Laat zien dat iedere

$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ te schrijven is als

$$x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n.$$

Matrices kunnen worden gebruikt om stelsels lineaire vergelijkingen korter te schrijven. Daartoe bekijk het stelsel

$$(2.1) \quad \begin{array}{rcl} a_{11}x_1 + \dots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + \dots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n & = & b_m \end{array}$$

Zij $A := (a_{ij})$ de matrix der coëfficiënten in (2.1), $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ en $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$. We definiëren nu het product Ax van de $m \times n$ matrix A en de $n \times 1$ matrix x door

$$(2.2) \quad Ax = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}$$

We kunnen dan (2.1) herschrijven als $Ax = b$. Uit (2.2) volgt i.h.b.

$$(2.3) \quad Ae_j = \text{de } j\text{-de kolom van } A.$$

Formule (2.2) is een voorbeeld van het vermenigvuldigen van matrices, immers $A \in M_{m,n}(\mathbb{R})$ en $x \in M_{n,1}(\mathbb{R})$.

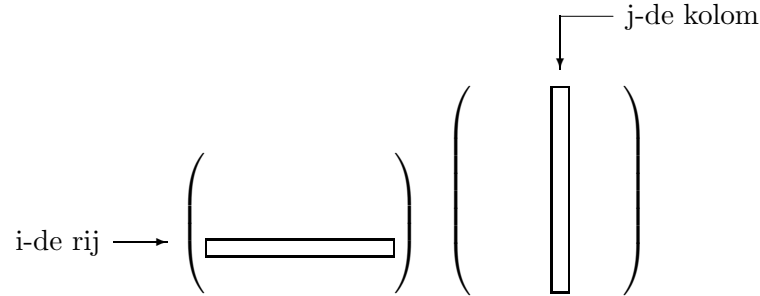
We zullen nu meer algemeen AB definiëren waarbij $A \in M_{m,n}(\mathbb{R})$ en $B \in M_{n,p}(\mathbb{R})$. Zij daartoe B_j de j -de kolom van B . Dan is volgens (2.2) AB_j een welgedefinieerde kolom uit $\mathbb{R}^m$.

Definitie 2.2.4 Zij $A \in M_{m,n}(\mathbb{R})$ en $B \in M_{n,p}(\mathbb{R})$. Dan is AB de $m \times p$ matrix met als j -de kolom AB_j m.a.w. $AB = (AB_1 \ AB_2 \ \dots \ AB_p)$.

Gevolg 2.2.5 Zij $A \in M_{m,n}(\mathbb{R})$ en $B \in M_{n,p}(\mathbb{R})$. Dan is AB de $m \times p$ matrix met op de plaats (i, j) het element $a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}$.

Bewijs. $(AB)_{ij}$ staat in de j -de kolom van AB m.a.w. in AB_j en wel in de i -de rij ervan. Omdat $B_j = \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix}$ is $(AB)_{ij}$ gelijk aan het i -de element uit de kolom $A \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix}$. Pas dan (2.2) toe met $x_1 = b_{1j}, \dots, x_n = b_{nj}$. $\square$

M.a.w. het element $(AB)_{ij}$ is te verkrijgen door de i -de rij van A te “vermenigvuldigen” met de j -de kolom van B . Hiervoor lopen we gelijktijdig over de i -de rij van A en de j -de kolom van B . De elementen van A en B waar we tegelijkertijd op staan worden vermenigvuldigd en deze producten worden bij het doorlopen opgeteld. Schematisch is dit in Figuur 2.1 weergegeven.



Figuur 2.1: Schema voor het verkrijgen van $(AB)_{ij}$

Voorbeeld 2.2.6 Zij $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ en $B = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 3 & -2 \end{pmatrix}$. Bepaal AB .

Oplossing.

$$\begin{aligned} (AB)_{11} &= (1 \ 2 \ 3) \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = 1 \cdot 2 + 2 \cdot (-1) + 3 \cdot 3 = 9 \\ (AB)_{21} &= (4 \ 5 \ 6) \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = 4 \cdot 2 + 5 \cdot (-1) + 6 \cdot 3 = 21 \end{aligned}$$

Net zo vinden we $(AB)_{12} = 1 \cdot 0 + 2 \cdot 1 + 3 \cdot (-2) = -4$ en $(AB)_{22} = 4 \cdot 0 + 5 \cdot 1 + 6 \cdot (-2) = -7$. Dus

$$AB = \begin{pmatrix} 9 & -4 \\ 21 & -7 \end{pmatrix}$$

Opmerking 2.2.7 We hebben dus alleen een matrixvermenigvuldiging gedefinieerd voor twee matrices met dezelfde “binnendimensies” d.w.z. $m \times n$, $n \times p$. Het resultaat is een $m \times p$ matrix en dat zijn precies de “buitendimensies”.

Opgave 2.2.8 Laat zien dat de enige keuze voor de definitie van de matrix AB is die zoals gegeven is in 2.2.4 als je formule (2.2) aanneemt en eist dat $(AB)e_j = A(Be_j)$ voor alle $1 \leq j \leq p$.

Rekenregels voor matrices

Matrices gedragen zich in veel opzichten zoals reële getallen. Stelling 2.2.13 hieronder geeft een overzicht van de belangrijkste eigenschappen. Echter voor we deze “bekende” eigenschappen formuleren, geven we eerst twee belangrijke verschillen.

Voor ieder tweetal reële getallen a en b hebben we dat $ab = ba$; we zeggen dat de vermenigvuldiging *commutatief* is. Voor matrices is de vermenigvuldiging echter niet commutatief.

Voorbeeld 2.2.9 $AB \neq BA$.

Zij $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ en $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Dan $AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ en $BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Opmerking 2.2.10 Het feit dat matrix vermenigvuldiging niet commutatief is, ligt ten grondslag aan één van de meest fundamentele ontdekkingen uit de moderne natuurkunde (quantum mechanica, 1925) namelijk Heisenberg’s *onzekerheids principe* dat zegt dat het onmogelijk is van een deeltje zowel snelheid als plaats met willekeurige precisie te bepalen!

Definitie 2.2.11 De $m \times n$ matrix met $a_{ij} = 0$ voor alle i en j noemen we de *nulmatrix* en noteren deze met $0_{m,n}$ of gewoon 0 als er geen misverstand over de afmetingen kan zijn. Omdat voor ieder $A \in M_{m,n}(\mathbb{R})$ geldt dat $A + 0_{m,n} = 0_{m,n} + A = A$ (ga dit na), speelt de nulmatrix dezelfde rol bij het optellen als de gewone nul in de reële getallen.

Voorbeeld 2.2.12 De berekeningen in Voorbeeld 2.2.9 laten nu een verder afwijkend gedrag van matrices zien: bij reële getallen volgt uit de relatie $ab = 0$ dat a of b nul zijn. Echter hebben we net gezien dat $BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ wel de nulmatrix is, terwijl geen van de matrices A of B de nulmatrix is.

Zoals aangekondigd gaan er ook een groot aantal “gewone” rekenregels door voor matrices. In de navolgende stelling duiden de letters A, B, C steeds matrices aan. We nemen ook steeds aan dat de afmetingen van de matrices zó zijn dat de aangegeven bewerkingen (optelling en vermenigvuldiging) welgedefinieerd zijn.

Stelling 2.2.13 De volgende rekenregels gelden voor matrices.

- i) $A + B = B + A$ (commutativiteit van de optelling)
- ii) $A + (B + C) = (A + B) + C$ (associativiteit van de optelling)
- iii) $A(BC) = (AB)C$ (associativiteit van de vermenigvuldiging)
- iv) $A(B + C) = AB + AC$ (links distributiviteit)
- v) $(B + C)A = BA + CA$ (rechts distributiviteit)

Opgave 2.2.14 Bewijs de regels i), ii), iv) en v).

Het bewijs van iii) is wat meer rekenwerk. Een overzichtelijk bewijs (zonder veel rekenwerk) zal verderop gegeven worden.

De gespiegelde van een matrix

Definitie 2.2.15 Zij $A \in M_{m,n}(\mathbb{R})$. De *gespiegelde* matrix van A , of ook wel de *getransponeerde* matrix van A , genoteerd A^t of A^{tr} (of soms A') is de $n \times m$ matrix gedefinieerd door $(A^t)_{ij} = A_{ji}$. M.a.w. de kolommen van A^t zijn de overeenkomstige rijen van A .

Voorbeeld 2.2.16 Zij $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. Dan is $A^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

De elementen a_{ii} van een matrix heten de *diagonaal elementen* van A . We zien dus dat A^t verkregen is uit A door spiegeling in de *hoofddiagonaal* van A d.w.z. de verzameling der diagonaal elementen.

Als $A^t = A$ dan i.h.b. $n = m$. We zeggen dan dat A *symmetrisch* is.

Stelling 2.2.17 Zij A en B $m \times n$ matrices en C een $n \times p$ matrix. Dan gelden

- i) $(A + B)^t = A^t + B^t$ en $(cA)^t = cA^t$ voor alle $c \in \mathbb{R}$.
- ii) $(AC)^t = C^t A^t$ (let op de volgorde!).
- iii) $(A^t)^t = A$.

Bewijs. We bewijzen ii): $((AC)^t)_{ij} = (AC)_{ji} = \sum_{k=1}^n a_{jk} c_{ki}$. (*)

Verder

$$\begin{aligned} (C^t A^t)_{ij} &= (i\text{-de rij } C^t) \cdot (j\text{-de kolom } A^t) = \underset{\text{als rij}}{(i\text{-de kolom } C)} \cdot \underset{\text{als kolom}}{(j\text{-de rij } A)} \\ &= (c_{1i}, \dots, c_{ni}) \begin{pmatrix} a_{j1} \\ \vdots \\ a_{jn} \end{pmatrix} = \sum_{k=1}^n c_{ki} a_{jk} = \sum_{k=1}^n a_{jk} c_{ki} \end{aligned} \quad (**)$$

Uit (*) en (**) volgt dat $((AC)^t)_{ij} = (C^t A^t)_{ij}$ voor alle i, j . Dus $(AC)^t = C^t A^t$.

De rest van stelling 2.2.17 volgt m.b.v. Opgave 2.2.18. $\square$

Opgave 2.2.18 Bewijs dat i) en iii) uit stelling 2.2.17 gelden.

Diagonaal matrices

Definitie 2.2.19 Zij $A \in M_n(\mathbb{R})$. Dan heet A een *diagonaalmatrix* als $a_{ij} = 0$ voor alle (i, j) met $i \neq j$.

Een handige notatie voor diagonaalmatrices is

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}).$$

Speciale diagonaalmatrices zijn

- i) de *eenheidsmatrix* $I_n = \text{diag}(1, 1, \dots, 1)$ die op de diagonaal alleen maar enen heeft;
- ii) de *scalair matrices* $c \cdot I_n = \text{diag}(c, c, \dots, c)$. I.h.b. is $0 \cdot I_n$ de nulmatrix in $M_n(\mathbb{R})$.

Voorbeeld 2.2.20 $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \pi \cdot I_2 = \begin{pmatrix} \pi & 0 \\ 0 & \pi \end{pmatrix},$

Opgave 2.2.21 Zij $A \in M_{m,n}(\mathbb{R})$. Bewijs dat $AI_n = A$ en $I_m A = A$.

Intermezzo: Volledige inductie

Bij een aantal bewijzen zullen we gebruik maken van een belangrijke techniek, de *volledige inductie* (of kort *inductie*).

Stel we willen bewijzen dat een uitspraak $U(n)$ die van een natuurlijk getal $n \in \mathbb{N}$ afhangt voor alle $n \in \mathbb{N}$ waar is. Dan zegt het principe van de volledige inductie dat het voldoende is om de volgende twee dingen aan te tonen:

1. De uitspraak $U(1)$ is waar, d.w.z. de uitspraak is waar voor $n = 1$. Dit noemen we soms ook de *inductiebasis*.
2. Als de uitspraak $U(n)$ voor een zekere n waar is (*inductieaanname*), dan is ook de uitspraak $U(n+1)$ voor het volgende natuurlijk getal waar. Dit noemen we de *inductiestap*.

We kunnen de volledige inductie als een soort wedstrijd aanzien tussen degene (**B**) die beweert dat de uitspraak waar is en een scepticus (**S**) die hieraan twijfelt.

B: De uitspraak $U(n)$ is waar voor alle $n \in \mathbb{N}$.

S: Daar geloof ik niks van, laat me een bewijs zien.

B: Geef me dan een $m \in \mathbb{N}$ waarvoor je denkt dat hij niet waar is (of ik hem niet kan bewijzen).

S: Laat het me dan voor $m = 4711$ zien.

B: Volgens de inductiebasis is $U(1)$ waar, met de inductiestap volgt dan dat ook $U(2)$ waar is, dan weer met de inductiestap dat $U(3)$ waar is enz. Uiteindelijk volgt met de inductiestap dat $U(4710)$ waar is en dan doe ik een laatste inductiestap en heb bewezen dat ook $U(4711)$ waar is.

Het zal duidelijk zijn dat deze dialoog niet van de keuze $m = 4711$ afhangt. Voor een $m \in \mathbb{N}$ wordt de uitspraak bewezen door met de inductiebasis aan te tonen dat $U(1)$ waar is en vervolgens door herhaaldelijke toepassing van de inductiestap dat $U(2), U(3), \dots, U(m-1), U(m)$ waar zijn.

Opmerking 2.2.22

- i) Soms moet een uitspraak niet voor alle $n \in \mathbb{N}$, maar bijvoorbeeld voor alle $n \geq 2$ of voor alle $n \geq -5$ bewezen worden. Het enige verschil is dat in zo'n geval de inductiebasis niet voor $n = 1$ maar voor $n = 2$ of $n = -5$ gelegd wordt. De rest van het inductie principe gaat gewoon door.
- ii) In sommige gevallen wordt voor de inductiestap niet alleen maar benodigd dat $U(n)$ waar is om $U(n+1)$ te bewijzen, maar ook dat $U(n-1), U(n-2), \dots, U(2), U(1)$ waar zijn, d.w.z. dat de uitspraak waar is voor *alle waarden die kleiner zijn dan $n+1$* .

Maar ook dit is in het inductie principe al bevat, want op het tijdstip dat we $U(n+1)$ willen bewijzen, weten we al dat $U(1), U(2), \dots, U(n)$ waar zijn.

Voorbeeld 2.2.23 Zij x een reëel getal met $x + \frac{1}{x} \in \mathbb{Z}$. Toon aan dat $x^n + \frac{1}{x^n} \in \mathbb{Z}$ voor alle $n \in \mathbb{N}$.

Oplossing. i) Volgens de aanname is de uitspraak waar voor $n = 1$ en ze is ook waar voor $n = 0$, want $x^0 + \frac{1}{x^0} = 2$.

ii) Stel nu dat $x^n + \frac{1}{x^n} \in \mathbb{Z}$ (en dat de uitspraak ook waar is voor alle $m \leq n$). Er geldt

$$x^{n+1} + \frac{1}{x^{n+1}} = \left(x + \frac{1}{x}\right)\left(x^n + \frac{1}{x^n}\right) - \left(x^{n-1} + \frac{1}{x^{n-1}}\right)$$

en omdat volgens de inductieaanname $x + \frac{1}{x} \in \mathbb{Z}$, $x^n + \frac{1}{x^n} \in \mathbb{Z}$ en $x^{n-1} + \frac{1}{x^{n-1}} \in \mathbb{Z}$ is ook de linkerkant een geheel getal.

Opgave 2.2.24 De *Fibonacci rij* $(F_n)_{n \in \mathbb{N}}$ is gedefinieerd door: $F_0 := 0$, $F_1 := 1$, $F_{n+1} := F_{n-1} + F_n$ voor $n \geq 1$. Toon aan dat

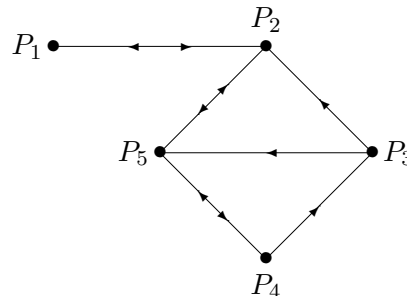
$$\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n.$$

2.3 Toepassingen

a) Grafentheorie en een voorbeeld uit de Sociologie

Stel je hebt een communicatie netwerk tussen 5 stations $P_1, \dots, P_5$. Sommige stations zijn verbonden door een 2-richtings communicatie en sommige door één richting. In Figuur 2.2 zie je een voorbeeld van zo'n netwerk.

De lijnen geven directe verbindingen aan en de pijltjes erin de richting(en) van de verbindingen. Zo is er tussen stations P_1 en P_2 een 2-richtings communicatie.



Figuur 2.2: Communicatie netwerk

Station P_4 kan een boodschap sturen naar P_1 via de stations P_3 en P_2 of via P_5 en P_2 . Dit communicatie netwerk is een voorbeeld van een *gerichte graaf*.

Definitie 2.3.1

- i) Een *gerichte graaf* G is een eindige verzameling van punten $P_1, \dots, P_n$ tesamen met gerichte lijnstukken die zekere paren van punten verbinden.
- ii) Een *pad* tussen twee punten is een serie van gerichte lijnstukken waarlangs je (in de richting van het lijnstuk) van het ene punt naar het andere kunt lopen.
- iii) De *lengte* van een pad is het aantal lijnstukken in zo'n pad. Een pad van lengte n heet een *n-pad*.

In bovenstaand voorbeeld kunnen we op verschillende manieren een boodschap van P_3 naar P_1 sturen. Bijvoorbeeld via het 2-pad $P_3 \rightarrow P_2 \rightarrow P_1$, maar ook via het 3-pad $P_3 \rightarrow P_5 \rightarrow P_2 \rightarrow P_1$. In dit voorbeeld is het eenvoudig om aan de tekening af te lezen wat de kortste verbinding tussen twee punten is. Echter als je met grote netwerken te maken hebt is dat niet meer mogelijk.

We zullen nu laten zien hoe matrixtheorie ons in staat stelt de lengte van het kortste pad te vinden. Daartoe maken we bij een gerichte graaf zijn zgn. adjacency matrix (welke de graaf volledig bepaalt).

Definitie 2.3.2 Beschouw een gerichte graaf G met (hoek)punten $P_1, \dots, P_n$. De *adjacency matrix* $A(G)$ van deze graaf is zó dat

$$a_{ij} = \begin{cases} 1 & \text{als er een directe verbinding van } P_i \text{ naar } P_j \text{ is;} \\ 0 & \text{anders.} \end{cases}$$

De adjacency matrix A van Figuur 2.2 is

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

Bijvoorbeeld $a_{12} = 1$ omdat er een directe verbinding van P_1 naar P_2 is; $a_{13} = 0$ omdat er geen directe verbinding van P_1 naar P_3 is.

De adjacency matrix beschrijft volkomen het netwerk, m.a.w. uit de matrix kunnen we Figuur 2.2 (of een equivalente schets van het netwerk) weer terug vinden. Het voordeel is dat we een matrix aan een computer kunnen geven en die dan berekeningen kunnen laten maken die nodig zijn om de lengte van het kortste pad tussen twee punten te vinden. Hoe kunnen we echter m.b.v. A de lengte van het kortste pad vinden? Het antwoord op deze vraag wordt gegeven door de volgende stelling

Stelling 2.3.3 Zij G een gerichte graaf met n punten $P_1, \dots, P_n$. Zij A de adjacency matrix van G . Dan is het aantal m -paden van P_i naar P_j precies gelijk aan $(A^m)_{ij}$.

Bewijs. We bewijzen dit met volledige inductie.

- i) Volgens de definitie van de adjacency matrix geldt de uitspraak voor $m = 1$.
- ii) Voor de duidelijkheid laten we eerst de inductiestap van $m = 1$ naar $m = 2$ zien en vervolgens het algemene geval van m naar $m + 1$:
 a_{i1} is het aantal directe verbindingen van P_i naar P_1 en a_{1j} is het aantal verbindingen van P_1 naar P_j . Dan is $a_{i1}a_{1j}$ het aantal 2-paden van P_i naar P_j via P_1 . Net zo is $a_{i2}a_{2j}$ het aantal 2-paden van P_i naar P_j via P_2 . Totaal is het aantal 2-paden van P_i naar P_j dus gelijk aan $a_{i1}a_{1j} + a_{i2}a_{2j} + \dots + a_{in}a_{nj}$ en dat is precies $(A^2)_{ij}$!
 ii') Laten we nu $(m+1)$ -paden bekijken van P_i naar P_j . Vat een $(m+1)$ -pad op als een m -pad gevolgd door een 1-pad. Het aantal m -paden van P_i naar P_1 is volgens de inductieaanname gelijk aan $(A^m)_{i1}$. Het aantal 1-paden van P_1 naar P_j is a_{1j} . Dus het aantal $(m+1)$ -paden van P_i naar P_j via P_1 is gelijk aan $(A^m)_{i1}a_{1j}$. Net zo vinden we dat het aantal $(m+1)$ -paden van P_i naar P_j via P_2 gelijk is aan $(A^m)_{i2}a_{2j}$. Totaal is dan het aantal $(m+1)$ -paden

van P_i naar P_j gelijk aan $(A^m)_{i1}a_{1j} + (A^m)_{i2}a_{2j} + \dots + (A^m)_{in}a_{nj}$ wat precies gelijk is aan $(A^m \cdot A)_{ij} = (A^{m+1})_{ij}$.

Volgens het inductie principe geldt de stelling dus voor iedere m . $\square$

We laten nu zien hoe deze stelling gebruikt kan worden om de lengte van het kortste pad van P_1 naar P_3 te vinden.

Het aantal 1-paden van P_1 naar P_3 wordt gegeven door a_{13} . Er geldt (zie de matrix onder definitie 2.3.2) $a_{13} = 0$. Het aantal 2-paden van P_1 naar P_3 wordt gegeven door $(A^2)_{13}$. Uitrekenen geeft

$$A^2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 2 \end{pmatrix}$$

Dus $(A^2)_{13} = 0$ m.a.w. er zijn geen 2-paden van P_1 naar P_3 . Uitrekenen van A^3 levert dat ook $(A^3)_{13} = 0$, dus ook geen 3-paden van P_1 naar P_3 . Tenslotte vinden we $(A^4)_{13} = 1$. Dus is er precies één 4-pad van P_1 naar P_3 . Dus blijktbaar is de lengte van het kortste pad van P_1 naar P_3 gelijk aan 4.

Natuurlijk konden we in ons eenvoudige voorbeeld dat ook al meteen aan Figuur 2.2 zien: de snelste manier om een boodschap van P_1 naar P_3 te sturen is via het 4-pad $P_1 \rightarrow P_2 \rightarrow P_5 \rightarrow P_4 \rightarrow P_3$.

Opmerking 2.3.4

- i) In het volgende hoofdstuk zullen we een methode geven om snel de lengte van het kortste pad van een punt P_i naar een punt P_j te berekenen zonder dat je $A, A^2, \dots$ moet uitrekenen.
- ii) Er is echter geen methode bekend om uit A ook een pad te vinden van P_i naar P_j met de kleinste afstand (behalve proberen, maar dat zal bij grote aantallen punten te lang duren)!

Opgave 2.3.5 Zij A de adjacency matrix van een gerichte graaf. Laat zien dat het aantal paden van lengte *hoogstens* m van P_i naar P_j gelijk is aan het element c_{ij} in de matrix $C = A + A^2 + \dots + A^m$.

Een andere nuttige matrix die we bij een gerichte graaf kunnen maken is de *afstandsmatrix* $D = (d_{ij})$. Deze is als volgt gedefinieerd

$$d_{ij} = \begin{cases} \text{de lengte van het kortste pad van } P_i \text{ naar } P_j; \\ 0 \text{ als } i = j; \\ x \text{ als er geen pad van } P_i \text{ naar } P_j \text{ is} \end{cases}$$

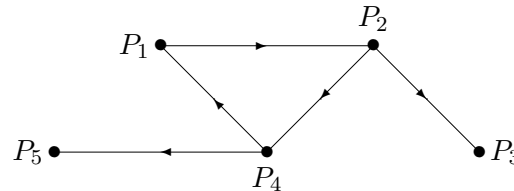
d.w.z. x symboliseert een oneindig grote afstand.

In bovenstaand voorbeeld hebben we dus berekend dat $d_{13} = 4$. Met nog wat meer werk vinden we

$$D = \begin{pmatrix} 0 & 1 & 4 & 3 & 2 \\ 1 & 0 & 3 & 2 & 1 \\ 2 & 1 & 0 & 2 & 1 \\ 3 & 2 & 1 & 0 & 1 \\ 2 & 1 & 2 & 1 & 0 \end{pmatrix}$$

Een voorbeeld uit de Sociologie

Beschouw een groep van 5 mensen, zeg $P_1, \dots, P_5$. Een socioloog wil weten welke van de 5 personen het meeste invloed heeft op de groep. Daartoe geeft hij ieder lid van de groep een formulier. Op dit formulier moet je je naam invullen en de naam van die persoon aan wiens mening je de meeste waarde hecht. Stel dat de uitkomsten de volgende zijn: P_1 zegt P_4 , P_2 zegt P_1 , P_3 zegt P_2 , P_4 zegt P_2 en P_5 zegt P_4 . De invloed gaat dus van rechts naar links (P_4 beïnvloed P_1 , P_1 beïnvloed P_2 enz). Dit geven we weer in de “dominantie graaf”: de groepsleden zijn de hoekpunten en de directe invloed wordt door een gericht lijnstuk weergegeven.



Figuur 2.3: Dominantie graaf

Vorm de afstandsmatrix D en tel de elementen in iedere rij op

$$D = \begin{pmatrix} 0 & 1 & 2 & 2 & 3 \\ 2 & 0 & 1 & 1 & 2 \\ x & x & 0 & x & x \\ 1 & 2 & 3 & 0 & 1 \\ x & x & x & x & 0 \end{pmatrix} \begin{matrix} 8 \\ 6 \\ 4x \\ 7 \\ 4x \end{matrix}$$

In dit voorbeeld zijn 1-paden directe invloed; 2-paden, 3-paden enz. corresponderen met indirecte invloed. We mogen dus redelijkerwijs aannemen dat geldt:

- Des te kleiner de afstand tussen P_i en P_j , des te groter de invloed die P_i op P_j heeft.

De som van de elementen in de i -de rij geeft de totale afstand van P_i tot de andere hoekpunten aan, dus de invloed op de hele groep. Dit leidt dan tot de volgende interpretatie van de rijssommen:

- Des te kleiner de som in de rij i is, des te groter is de invloed van P_i op de groep.

We zien dus dat P_2 de meeste invloed op de groep heeft! Deze informatie kan natuurlijk gebruikt worden als je als buitenstaander invloed op de groep wilt uitoefenen: je beïnvloed P_2 !

b) Een toepassing uit de archeologie

Een van de problemen waar een archeoloog mee te maken krijgt is het volgende: hij onderzoekt een groot aantal graven met daarin allerlei soorten voorwerpen. Hoe kun je die verschillende graven ordenen in chronologische volgorde? Een van de eersten die dit probleem onderzocht was Sir Flinders Petrie (1853-1942). Op het eind van de 19-de eeuw gebruikt hij de gegevens van ongeveer 900 graven om ze in chronologische volgorde te plaatsen. De methode die hij daarbij gebruikte zullen we nu behandelen. Hij ging uit van de volgende veronderstelling:

- Twee graven die veel overeenkomstige voorwerpen bevatten liggen hoogst waarschijnlijk dicht bij elkaar in tijd dan twee graven die weinig overeenkomstige voorwerpen bevatten.

Aan de hand hiervan maken we het volgende wiskundige model: nummer de graven $1, 2, 3, \dots$ en de verschillende soorten aardewerk met $1, 2, 3, \dots$. Zij nu A de matrix gedefinieerd door

$$a_{ij} = \begin{cases} 1 & \text{als graf } i \text{ aardewerk van type } j \text{ bevat;} \\ 0 & \text{anders.} \end{cases}$$

De volgende stelling zegt dan hoe we informatie uit deze matrix kunnen halen.

Stelling 2.3.6 Het element g_{ij} van de matrix $G = AA^t$ is gelijk aan het aantal type aardewerk dat de graven i en j gemeen hebben.

Bewijs.

$$\begin{aligned} g_{ij} &= \text{het element in de } i\text{-de rij en } j\text{-de kolom van } G \\ &= (i\text{-de rij van } A) \cdot (j\text{-de kolom van } A^t) \\ &= (i\text{-de rij van } A) \cdot (j\text{-de rij van } A) \\ &= a_{i1}a_{j1} + a_{i2}a_{j2} + \dots + a_{in}a_{jn}. \end{aligned}$$

Iedere term is 0 of 1. Bijvoorbeeld de term $a_{i2}a_{j2} = 1$ d.e.s.d.a. beide $a_{i2} = 1$ en $a_{j2} = 1$ m.a.w. als i en j beide aardewerk van type 2 bevatten. Dus het aantal enen in bovenstaande uitdrukking voor g_{ij} is precies het aantal type aardewerk dat de graven i en j gemeen hebben. $\square$

Dus: hoe groter g_{ij} hoe dichter de graven i en j bij elkaar liggen. Door dan de elementen g_{ij} te bekijken kan de archeoloog de chronologische volgorde van de graven vaststellen. Laten we deze methode aan de hand van een voorbeeld illustreren.

Stel dat de volgende matrix de 3 soorten aardewerk die in 4 graven voorkomt beschrijft

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Dus bijvoorbeeld $a_{13} = 1$ betekent dat graf 1 aardewerk van type 3 bevat en $a_{23} = 0$ betekent dat graf 2 geen aardewerk van type 3 bevat. We vinden dan

$$G = AA^t = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Omdat de matrix G symmetrisch is ($g_{ij} = \text{aantal type aardewerk dat } i \text{ en } j \text{ gemeen hebben} = \text{aantal type aardewerk dat } j \text{ en } i \text{ gemeen hebben} = g_{ji}$) bekijken we alleen de elementen boven de diagonaal (de elementen op de diagonaal geven de aantallen type aardewerk in de verschillende graven aan). Dus kijken we naar

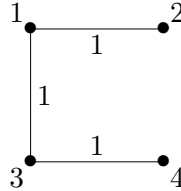
$$g_{12} = 1, g_{13} = 1, g_{14} = 0, g_{23} = 0, g_{24} = 0, g_{34} = 1.$$

Omdat $g_{12} = 1$ hebben graven 1 en 2 één type aardewerk gemeen. We geven dit aan met $1 - 2$. Bekijk nu graf 3. Omdat $g_{13} = 1$ en $g_{23} = 0$ ligt 3 dichtbij 1 en 3 niet dichtbij 2. Dus krijgen we $3 - 1 - 2$. Bekijk nu graf 4. Opmerken dat $g_{24} = 0$ en $g_{34} = 1$ vinden we $4 - 3 - 1 - 2$. De wiskunde zegt ons niets over de tijdsvolgorde. D.w.z. er zijn twee

mogelijkheden $4 \rightarrow 3 \rightarrow 1 \rightarrow 2$ of $4 \leftarrow 3 \leftarrow 1 \leftarrow 2$. Vaak weet de archeoloog van de uiteinden van zo'n diagram wel de chronologische volgorde en dan kan hij daaruit die van alle graven bepalen.

In de praktijk is de matrix G groot en kan de volgorde niet altijd zo eenvoudig (en ondubbelzinnig) verkregen worden als in ons voorbeeld. In Petries geval, die zoals gezegd 900 graven onderzocht, zou het een 900 bij 900 matrix zijn.

Vaak is het handig de matrix G in een *overeenstemmings graaf* te vertalen, waarbij de punten met de graven corresponderen en de punten i en j door een verbinding met label g_{ij} verbonden zijn als $g_{ij} > 0$. Voor het voorbeeld laat Figuur 2.4 deze graaf zien.



Figuur 2.4: Overeenstemmings graaf

Om de volgorde te bepalen, moet een pad door de overeenstemmings graaf gevonden worden die alle punten precies één keer bevat en zo dat de som der labels in de verbindingen in dit pad zo groot mogelijk is.

Er zijn meer matrix- en grafentechnieken ontwikkeld om zulke matrices te onderzoeken (voor de geïnteresseerde lezer: zie “Some Problems and Methods in Statistical Archeology” door David G. Kendall in World Archeology, 1, 1969, pp. 61-76).

2.4 Inverteerbare matrices

Een speciale klasse van matrices zijn de vierkante matrices. Een belangrijke subklasse daarvan vormen de zgn. inverteerbare matrices.

Definitie 2.4.1 Een $n \times n$ matrix A heet *inverteerbaar* als er een $n \times n$ matrix B bestaat zodanig dat $AB = I_n = BA$. De matrix B heet een *inverse* van A . Als zo'n matrix B niet bestaat heet A *singulier* of gewoon *niet inverteerbaar*.

Voorbeeld 2.4.2 i) De matrix $A = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$ is inverteerbaar, immers men rekent eenvoud-

dig na dat de matrix $B = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$ een inverse van A is.

ii) De matrix $A := \begin{pmatrix} 2 & -6 \\ -1 & 3 \end{pmatrix}$ is singulier, immers als er een B bestaat met $BA = I_2$ dan i.h.b. $BA \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Echter $A \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ en dus $BA \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, een tegenspraak. Dus zo'n B bestaat niet m.a.w. A is singulier.

Men kan zich afvragen of, indien een matrix A een inverse matrix B heeft, deze uniek is (net zoals er bij ieder reëel getal $a \neq 0$ maar precies een reëel getal b bestaat met $ba = 1$, namelijk $b = \frac{1}{a}$). Het antwoord is ja en wordt gegeven door de volgende stelling.

Stelling 2.4.3 (Uniciteit inverse.) Als B en C beiden inverse van A zijn, dan $B = C$.

Bewijs. Omdat B een inverse van A is geldt $BA = I$. Vermenigvuldig deze vergelijking van rechts met C . Dit geeft $(BA)C = I.C = C$. Anderzijds geldt $(BA)C = B(AC) = B.I = B$ en dus $B = C$. $\square$

We kunnen daarom nu spreken over *de* inverse van een inverteerbare matrix. Als A inverteerbaar is, duiden we zijn inverse aan met A^{-1} . We hebben dus

$$AA^{-1} = I = A^{-1}A.$$

Gevolg 2.4.4 Zij A een inverteerbare $n \times n$ matrix. Dan is er voor iedere $b \in \mathbb{R}^n$ precies één oplossing x uit $\mathbb{R}^n$ die voldoet aan $Ax = b$, namelijk $x = A^{-1}b$.

Bewijs. i) Zij $b \in \mathbb{R}^n$. Als $x \in \mathbb{R}^n$ voldoet aan $Ax = b$, dan vind je door van links met A^{-1} te vermenigvuldigen dat $A^{-1}(Ax) = A^{-1}b$. Echter $A^{-1}(Ax) = (A^{-1}A)x = I_n x = x$, dus $x = A^{-1}b$.

ii) Omgekeerd $A(A^{-1}b) = (AA^{-1})b = I_n b$, dus $x = A^{-1}b$ voldoet aan $Ax = b$. $\square$

Gevolg 2.4.5 Als A een inverteerbare $n \times n$ matrix is dan geldt: als $Ax = 0$, dan $x = 0$.

We zullen in LA2 bewijzen dat ook de volgende omkering van 2.4.5 geldt:

Stelling 2.4.6 Zij A een $n \times n$ matrix. Als uit $Ax = 0$ volgt dat $x = 0$, dan is A inverteerbaar.

Elementaire matrices

In sectie 1.2 gebruikten we drie bewerkingen om stelsels lineaire vergelijkingen te vereenvoudigen. De overeenkomstige bewerkingen kunnen we op de rijen van een matrix A toepassen: ze heten de *elementaire rij operaties*. Dit zijn de volgende bewerkingen.

- 1) Zij $a \in \mathbb{R}$ en $i \neq j$. Tel a keer de j -de rij van A op bij de i -de rij van A .
- 2) Zij $a \in \mathbb{R} \setminus \{0\}$ en i een rijnummer. Vermenigvuldig de i -de rij van A met a .
- 3) Zij $i \neq j$. Verwissel de i -de en de j -de rij van A .

Het toepassen van deze rij operaties op een matrix heet *vegen* van de matrix (omdat je vaak tot nullen “veegt”).

Definitie 2.4.7 Een *elementaire matrix* is een matrix die uit I_n verkregen kan worden d.m.v. één elementaire rij operatie. De matrices die uit I_n d.m.v. de operaties 1) resp. 2) resp. 3) worden verkregen, noteren we als $E_{ij}(a)$ resp. $D_i(a)$ resp. P_{ij} .

Voorbeeld 2.4.8 In $M_3(\mathbb{R})$:

$$E_{12}(7) = \begin{pmatrix} 1 & 7 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_{32}(-5) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 1 \end{pmatrix}, \quad D_2(3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Iedere elementaire matrix is inverteerbaar en zijn inverse is weer elementair. Dit volgt expliciet m.b.v. de volgende opgave.

Opgave 2.4.9 Laat zien dat voor de elementaire matrices de volgende relaties gelden.

$$E_{ij}(-a)E_{ij}(a) = I_n, \quad D_i\left(\frac{1}{a}\right)D_i(a) = I_n \quad (a \neq 0), \quad P_{ij}P_{ij} = I_n.$$

M.b.v. het volgende lemma kunnen we heel veel inverteerbare matrices aanmaken.

Lemma 2.4.10 Als A en B inverteerbaar zijn is AB het ook en er geldt $(AB)^{-1} = B^{-1}A^{-1}$.

Bewijs. Zij $C := B^{-1}A^{-1}$. Dan $C(AB) = B^{-1}A^{-1}(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$. Net zo geldt $(AB)C = I$, dus AB is inverteerbaar met inverse $C = B^{-1}A^{-1}$. $\square$

Uit lemma 2.4.10 en opgave 2.4.9 volgt dan eenvoudig (met inductie over de lengte van het product) dat ieder eindig product van elementaire matrices inverteerbaar is. Echter de omkering geldt ook m.a.w.:

Stelling 2.4.11 Iedere inverteerbare matrix is een eindig product van elementaire matrices.

Het bewijs van deze stelling maakt gebruik van het volgende lemma.

Lemma 2.4.12 Zij A een $n \times m$ matrix. Dan geldt

- i) $E_{ij}(a)A$ is de matrix verkregen uit A door a keer de j -de rij van A bij de i -de rij van A op te tellen.
- ii) $D_i(a)A$ is de matrix verkregen uit A door de i -de rij van A met a te vermenigvuldigen.
- iii) $P_{ij}A$ is de matrix verkregen uit A door de i -de rij en de j -de rij van A te verwisselen.

Opgave 2.4.13 Zij $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$. Pas op A de elementaire rij operatie toe die 7 keer de eerste rij van A bij de tweede rij van A optelt. Het resultaat is dan

$$\begin{pmatrix} a_1 & a_2 \\ a_3 + 7a_1 & a_4 + 7a_2 \end{pmatrix}.$$

Ga na dat je dit resultaat ook krijgt als je $E_{21}(7)A$ uitrekent.

Bewijs van stelling 2.4.11. We zullen het volgende bewijzen:

- (*) Iedere inverteerbare $n \times n$ matrix is d.m.v. een eindig aantal elementaire rij operaties op de vorm I_n te brengen.

Als we (*) bewezen hebben dan volgt de stelling: immers dan bestaan er volgens 2.4.12 elementaire matrices $E_1, \dots, E_s$ zodanig dat $E_s \dots E_1 A = I_n$. Door nu van links met $E_1^{-1} \dots E_s^{-1}$ te vermenigvuldigen vinden we dat $A = E_1^{-1} \dots E_s^{-1}$. Omdat volgens 2.4.9 iedere E_i^{-1} ook weer een elementaire matrix is, volgt de stelling.

We geven nu een (constructief) bewijs van (*) met inductie naar n . Het geval $n = 1$ is duidelijk, laat nu $n \geq 2$ zijn. Omdat A inverteerbaar is, kan de eerste kolom niet alleen maar nullen bevatten, want anders zou $Ae_1 = 0$ zijn, en dan is $A^{-1}Ae_1 = 0$, in tegenspraak tot $A^{-1}Ae_1 = I_n e_1 = e_1$. Stel dus dat $a_{i1} \neq 0$, dan verruilen we de eerste en de i -de rij en er geldt:

$$P_{i1}A = \begin{pmatrix} a_{i1} \neq 0 & \dots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

(voor het gemak zullen we de elementen in de getransformeerde matrix steeds weer met a_{ij} noteren). Nu delen we de eerste rij door a_{i1} , dit geeft

$$D_1\left(\frac{1}{a_{i1}}\right)P_{i1}A = \begin{pmatrix} 1 & \dots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

We kunnen de eerste kolom nu tot nullen vegen door voor $i = 2, 3, \dots, n$ het a_{i1} -voud van de eerste rij van de i -de rij af te trekken:

$$E_{21}(-a_{21})D_1\left(\frac{1}{a_{11}}\right)P_{i1}A = \begin{pmatrix} 1 & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

enzovoorts tot

$$E_{n1}(-a_{n1}) \dots E_{21}(-a_{21})D_1\left(\frac{1}{a_{11}}\right)P_{i1}A = \begin{pmatrix} 1 & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ 0 & a_{32} & \dots & a_{3n} \\ \vdots & & & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{pmatrix} = B$$

Door met elementaire matrices te vermenigvuldigen hebben we A dus getransformeerd tot de matrix

$$B = \left(\begin{array}{c|ccc} 1 & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{array} \right) = \left(\begin{array}{c|ccc} 1 & a_{12} & \dots & a_{1n} \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \begin{array}{c} \\ B_{n-1} \\ \\ \end{array} \right)$$

Omdat A inverteerbaar is en de elementaire matrices inverteerbaar zijn, is ook B inverteerbaar, dus is er een matrix

$$B^{-1} = \left(\begin{array}{c|ccc} * & * & \dots & * \\ * & & & \\ \vdots & & & \\ * & & & \end{array} \begin{array}{c} \\ C_{n-1} \\ \\ \end{array} \right)$$

met $BB^{-1} = I_n$, d.w.z.

$$\left(\begin{array}{c|ccc} 1 & a_{12} & \dots & a_{1n} \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \begin{array}{c} \\ B_{n-1} \\ \\ \end{array} \right) \left(\begin{array}{c|ccc} * & * & \dots & * \\ * & & & \\ \vdots & & & \\ * & & & \end{array} \begin{array}{c} \\ C_{n-1} \\ \\ \end{array} \right) = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \begin{array}{c} \\ I_{n-1} \\ \\ \end{array} \right)$$

Door naar het blok rechts onder te kijken, zien we dat $B_{n-1}C_{n-1} = I_{n-1}$, want de eerste rij in B^{-1} wordt met de nullen in de eerste kolom van B vermenigvuldigd en levert dus geen bijdrage. De $(n-1) \times (n-1)$ matrix B_{n-1} is dus inverteerbaar.

Volgens inductie nemen we nu aan dat we voor $(n-1) \times (n-1)$ matrices al hebben aangetoond dat ze producten van elementaire matrices zijn, en de corresponderende operaties om B_{n-1} tot I_{n-1} te transformeren kunnen we nu op B toepassen, dit geeft:

$$E_s \dots E_1 B = \left(\begin{array}{c|ccc} 1 & a_{12} & \dots & a_{1n} \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \begin{array}{c} \\ I_{n-1} \\ \\ \end{array} \right) = \left(\begin{array}{c|ccc} 1 & a_{12} & \dots & a_{1n} \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{array} \right) = C$$

Ten slotte kunnen we de matrix C door voor $i = 2, 3, \dots, n$ het a_{1i} -voud van de i -de rij van de eerste af te trekken tot I_n transformeren, er geldt:

$$E_{12}(-a_{12})C = \left(\begin{array}{c|cccc} 1 & 0 & a_{13} & \dots & a_{1n} \\ \hline 0 & 1 & & & \\ \vdots & & \ddots & & \\ 0 & & & & 1 \end{array} \right)$$

enzovoorts tot

$$E_{1n}(-a_{1n}) \dots E_{12}(-a_{12})C = \left(\begin{array}{c|cccc} 1 & 0 & \dots & 0 \\ \hline 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{array} \right)$$

□

Opgave 2.4.14 Zij A een inverteerbare matrix. Bewijs dat A^t een inverteerbare matrix is met $(A^t)^{-1} = (A^{-1})^t$.

Voorbeeld 2.4.15 Zij $A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$. Schrijf A als product van elementaire matrices.

Oplossing: Tel -1 keer de 1-ste rij op bij de 2-de. Dit geeft $A' := \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$. Dus met 2.4.12 zien we: $E_{21}(-1)A = A'$. Vermenigvuldig de tweede rij van A' met -1 . Dit geeft $A'' = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ en dus met 2.4.12 $D_2(-1)A' = A''$ en dus $D_2(-1)E_{21}(-1)A = A''$. Tel tenslotte -2 keer de tweede rij van A'' op bij de eerste rij van A'' . Dit geeft $A''' := I_3$ en dus met 2.4.12 $E_{12}(-2)A'' = I_3$ en dus

$$E_{12}(-2)D_2(-1)E_{21}(-1)A = I_3.$$

Zoals opgemerkt in ii) van het bewijs van 2.4.11 volgt hieruit dat

$$A = E_{21}(1)D_2(-1)E_{12}(2).$$

Een algoritme om A^{-1} te berekenen

Zij A een inverteerbare $n \times n$ matrix. In het bewijs van stelling 2.4.11 hebben we laten zien dat we A door een eindig aantal elementaire rij operaties op de vorm I_n kunnen brengen. Voor de bijbehorende elementaire matrices $E_1, \dots, E_s$ geldt dan

$$(**) \quad E_s \dots E_1 A = I_n \text{ m.a.w. } A^{-1} = E_s \dots E_1.$$

Bekijk nu de $n \times 2n$ matrix (A, I_n) waarvan de eerste n kolommen die van A zijn en de laatste n die van I_n . Pas nu op de matrix (A, I_n) dié elementaire rij operaties toe die A to I_n herleiden m.a.w. zodat $E_s \dots E_1 A = I_n$. Dan

$$E_s \dots E_1 (A, I_n) = (E_s \dots E_1 A, E_s \dots E_1 I_n) = (I_n, A^{-1}) \quad (\text{volgens } (**)).$$

M.a.w. als we de matrix (A, I_n) zó vegen dat de linker $n \times n$ matrix de identiteit wordt, dan vind je dat de rechter matrix vanzelf A^{-1} wordt!

Voorbeeld 2.4.16 Zij $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$. Bereken A^{-1} .

Oplossing. Vorm $(A, I_3) = \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$.

Trek de eerste rij van de tweede en ook van de derde af. Dit geeft

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

Vermenigvuldig de tweede rij met -1 en tel hem dan 2 keer op bij de derde rij. Dit geeft

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

Vermenigvuldig de laatste rij met $\frac{1}{2}$ en tel hem dan -2 keer op bij de tweede en -3 keer bij de eerste. Dit geeft

$$\begin{pmatrix} 1 & 2 & 0 & -\frac{1}{2} & 3 & -\frac{3}{2} \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & \frac{1}{2} & -1 & \frac{1}{2} \end{pmatrix}.$$

Tel tenslotte de tweede rij -2 keer op bij de eerste rij. Dit geeft

$$\begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & \frac{1}{2} & -1 & \frac{1}{2} \end{pmatrix}.$$

Dus

$$A^{-1} = \begin{pmatrix} -\frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 1 & -1 \\ \frac{1}{2} & -1 & \frac{1}{2} \end{pmatrix}.$$

De inverse matrix in het geval $n = 2$

Voor 2×2 -matrices kunnen we zelfs in het algemeen bepalen, of een matrix inverteerbaar is en de inverse dan ook aangeven.

Stelling 2.4.17 De 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is inverteerbaar d.e.s.d.a. $ad - bc \neq 0$. In dit geval geldt $A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

Bewijs. i) We laten eerst zien dat A niet inverteerbaar is als $ad - bc = 0$. Er geldt

$$A \begin{pmatrix} d \\ -c \end{pmatrix} = \begin{pmatrix} ad - bc \\ 0 \end{pmatrix} \quad \text{en} \quad A \begin{pmatrix} -b \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ -bc + ad \end{pmatrix}$$

De nulmatrix is zeker niet inverteerbaar en als A niet de nulmatrix is, is ten minste een van $\begin{pmatrix} d \\ -c \end{pmatrix}$ en $\begin{pmatrix} -b \\ a \end{pmatrix}$ niet nul, maar geeft als resultaat bij vermenigvuldiging met A wel nul als $ad - bc = 0$. Volgens 2.4.5 is A dus niet inverteerbaar.

ii) Stel nu dat $ad - bc \neq 0$. We passen het net beschreven algoritme toe op de matrix $(A, I_2) = \begin{pmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{pmatrix}$.

iii) We nemen eerst aan dat $a \neq 0$. Deel de eerste rij door a en tel hem dan $-c$ keer bij de tweede op. Dit geeft

$$\begin{pmatrix} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & d - \frac{bc}{a} & -\frac{c}{a} & 1 \end{pmatrix}$$

Wegens $ad - bc \neq 0$ kunnen we de tweede rij met $\frac{a}{ad-bc}$ vermenigvuldigen, dit geeft

$$\begin{pmatrix} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}$$

Tenslotte tellen we $-\frac{b}{a}$ keer de tweede rij bij de eerste op, dit geeft

$$\begin{pmatrix} 1 & 0 & \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}$$

In het geval $a \neq 0$ en $ad - bc \neq 0$ is A dus inverteerbaar met inverse $A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

iv) Voor het geval $a = 0$ moet $b \neq 0$ en $c \neq 0$ gelden, want anders zou $ad - bc = 0$ zijn. We verruilen de eerste en de tweede rij en vermenigvuldigen de nieuwe eerste rij met $\frac{1}{c}$. Dit geeft

$$\begin{pmatrix} 1 & \frac{d}{c} & 0 & \frac{1}{c} \\ 0 & b & 1 & 0 \end{pmatrix}$$

Nu vermenigvuldigen we de tweede rij met $\frac{1}{b}$ en tellen vervolgens $-\frac{d}{c}$ keer de tweede rij bij de eerste op. Dit geeft

$$\begin{pmatrix} 1 & 0 & \frac{-d}{bc} & \frac{1}{c} \\ 0 & 1 & \frac{1}{b} & 0 \end{pmatrix}$$

Ook in dit geval blijkt A dus inverteerbaar te zijn en de inverse komt overeen met de boven gevonden inverse met $a = 0$ ingevuld. $\square$

Een toepassing in de cryptografie

Cryptografie is het proces van coderen en decoderen van boodschappen. Het woord komt van het Griekse woord “kryptos” wat “verborgen” betekent. Reeds de oude Grieken verstuurden geheime boodschappen. Tegenwoordig gebruiken regeringen allerlei hoog ontwikkelde technieken voor het coderen en decoderen van boodschappen. Een type code die moeilijk te breken is maakt gebruik van een grote *codings matrix*. De ontvanger decodeert de boodschap m.b.v. de inverse van de matrix, welke de *decodings matrix* heet. We laten aan de hand van een voorbeeld zien hoe deze methode werkt.

Bekijk de boodschap

PREPARE TO ATTACK

Gebruik de volgende coderings matrix

$$\begin{pmatrix} -3 & -3 & -4 \\ 0 & 1 & 1 \\ 4 & 3 & 4 \end{pmatrix}$$

Iedere letter van het alfabet geven we een getal. Voor het gemak geven we A een 1, B een 2, C een 3. enz. en een “spatie” geven we een 27. Onze boodschap wordt dan

$$\begin{array}{cccccccccccccccccccc} P & R & E & P & A & R & E & * & T & O & * & A & T & T & A & C & K \\ 16 & 18 & 5 & 16 & 1 & 18 & 5 & 27 & 20 & 15 & 27 & 1 & 20 & 20 & 1 & 3 & 11 \end{array}$$

Omdat we bovenstaande 3×3 matrix gaan gebruiken om de boodschap te coderen, splitsen we de getallen boodschap op in kolommen van lengte 3 d.w.z.

$$\begin{pmatrix} 16 \\ 18 \\ 5 \end{pmatrix} \begin{pmatrix} 16 \\ 1 \\ 18 \end{pmatrix} \begin{pmatrix} 5 \\ 27 \\ 20 \end{pmatrix} \begin{pmatrix} 15 \\ 27 \\ 1 \end{pmatrix} \begin{pmatrix} 20 \\ 20 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 11 \\ 27 \end{pmatrix}.$$

We vermenigvuldigen nu ieder van deze kolommen met de coderings matrix. We krijgen

$$\begin{pmatrix} -3 & -3 & -4 \\ 0 & 1 & 1 \\ 4 & 3 & 4 \end{pmatrix} \begin{pmatrix} 16 & 16 & 5 & 15 & 20 & 3 \\ 18 & 1 & 27 & 27 & 20 & 11 \\ 5 & 18 & 20 & 1 & 1 & 27 \end{pmatrix} = \begin{pmatrix} -122 & -123 & -176 & -130 & -124 & -150 \\ 23 & 19 & 47 & 28 & 21 & 38 \\ 138 & 139 & 181 & 145 & 144 & 153 \end{pmatrix} = \underline{B}.$$

De kolommen van de matrix $\underline{B}$ geven de gecodeerde boodschap. Deze boodschap wordt als volgt verstuurd

$$-122, 23, 138, -123, 19, 139, -176, 47, 181, -130, 28, 145, -124, 21, 144, -150, 38, 153.$$

Om deze boodschap te decoderen, schrijft de ontvanger deze reeks getallen als een serie van 3×1 kolommen en herhaalt voorgaande constructie met de *inverse* van de coderings matrix, de decoderings matrix. Deze is

$$\begin{pmatrix} 1 & 0 & 1 \\ 4 & 4 & 3 \\ -4 & -3 & -3 \end{pmatrix}.$$

Om de boodschap te decoderen vermenigvuldigen we

$$\begin{pmatrix} 1 & 0 & 1 \\ 4 & 4 & 3 \\ -4 & -3 & -3 \end{pmatrix} \underline{B} = \begin{pmatrix} 16 & 16 & 5 & 15 & 20 & 3 \\ 18 & 1 & 27 & 27 & 20 & 11 \\ 5 & 18 & 20 & 1 & 1 & 27 \end{pmatrix}.$$

De kolommen van de laatste matrix, achter elkaar geschreven, geven de oorspronkelijke boodschap

$$\begin{array}{cccccccccccccccccccc} 16 & 18 & 5 & 16 & 1 & 18 & 5 & 27 & 20 & 15 & 27 & 1 & 20 & 20 & 1 & 3 & 11 & 27 \\ P & R & E & P & A & R & E & * & T & O & * & A & T & T & A & C & K & * \end{array}$$

Historische opmerkingen

1. De term *matrix* is voor het eerst in de wiskundige literatuur genoemd in een artikel van Joseph Sylvester (1814-1897) in 1850. Voor Sylvester was een matrix een getallenrecht-hoek waaruit je kleine vierkante deelmatrices kon nemen om daarvan de determinant te nemen (zie volgend hoofdstuk).

Sylvester was geboren in Londen en werd een van de grootste algebraïci van de 19-de eeuw. Toen hij veertien was studeerde hij aan de Universiteit van Londen onder leiding van de beroemde Engelse wiskundige Augustus DeMorgan (1806-1871). Hij kreeg zijn graad aan de Universiteit van Cambridge in 1837. In 1841 werd hij Professor aan de Universiteit van Virginia, maar door zijn afschuw van de slavernij verbleef hij daar maar kort en verliet Amerika. In 1871, nadat Abraham Lincoln (1809-1865) op 1 januari 1863 de vrijheidsverklaring der slaven had afgekondigd, keerde hij naar de Verenigde Staten terug als hoogleraar aan de Johns Hopkins universiteit. In de tussenliggende periode werkte hij tien jaar als advocaat, gedurende welke tijd hij Arthur Cayley ontmoette. Ook hij was vijftien jaar wiskundedocent aan de Royal Military Academy te Woolwich. Sylvester was ook een enthousiast dichter en veel van zijn wiskundig werk wordt voorafgegaan door eigen gedichten.

2. *Matrixvermenigvuldiging* komt oorspronkelijk voort uit het samenstellen van lineaire substituties in het werk *Disquisitiones Arithmeticae* van Gauss in 1801, in samenhang met de studie van kwadratische vormen, d.w.z. uitdrukkingen van de vorm $ax^2 + bxy + cy^2$. Gauss vermeldde niet expliciet de matrixvermenigvuldiging; dat werd gedaan door zijn student Ferdinand Gotthold Eisenstein (1823-1852), die de notatie $A \times B$ voor matrixvermenigvuldiging invoerde.
3. Het begrip *inverse van een matrix* verscheen voor het eerst in 1855 in een artikel van Arthur Cayley (1821-1895). Hij maakte het meer expliciet in 1859 in een artikel getiteld "A memoir on the theory of matrices". Hij beschreef daarin de basis eigenschappen van matrices, door op te merken dat de meeste van die eigenschappen voortkomen uit het bestuderen van stelsels lineaire vergelijkingen. I.h.b. komt de inverse voort uit het idee x, y, z uit het stelsel

$$\begin{aligned} X &= ax + by + cz \\ Y &= a'x + b'y + c'z \\ Z &= a''x + b''y + c''z \end{aligned}$$

op te lossen in termen van X , Y en Z . Cayley geeft een expliciete constructie.

Arthur Cayley studeerde in 1842 af aan het Trinity College te Cambridge, maar kon geen geschikte onderwijsbaan als wiskundige vinden. Daarom studeerde hij net als Sylvester rechten en werd in 1849 advocaat. Gedurende zijn veertien-jarige advocaatschap schreef hij ongeveer 300 wiskundige artikelen. Tenslotte werd hij in 1863 hoogleraar te Cambridge waar hij bleef tot zijn dood. Hij was in Cambridge de man die het bestuur ertoe overhaalde ook vrouwen als student toe te laten. Gedurende zijn rechtenstudie ontmoette hij Sylvester: hun gesprekken tijdens de volgende veertig jaar waren enorm vruchtbaar voor de vooruitgang van de algebra. Tijdens zijn leven schreef Cayley ongeveer duizend artikelen in wiskunde, theoretische dynamica en mathematische astronomie.

Enkele vooruitblikken

In plaats van matrices met coëfficiënten in $\mathbb{R}$ of $\mathbb{C}$ kan men ook kijken naar matrices waarvan de matrix elementen gehele getallen zijn of veeltermen in meerdere variabelen. Meer algemeen kan men kijken naar matrices waarvan de matrix elementen komen uit een zgn. *commutatieve ring* R , d.w.z. een verzameling waarin je kunt optellen, aftrekken en vermenigvuldigen (zie het college *Ring en Lichamen*). Een voorbeeld is de ring der gehele getallen of een *veelterm ring* over $\mathbb{R}$. Ook kan men dan spreken over elementaire matrices waarbij we dan bij de matrices van de vorm $D_i(a)$ moeten eisen dat a een eenheid in R is d.w.z. er bestaat een b in R met $ab = 1$.

Een verrassing is dat stelling 2.4.11 niet meer doorgaat voor een willekeurige commutatieve ring R . Als bijvoorbeeld R de veeltermring $\mathbb{R}[x, y]$ der veeltermen in twee variabelen x en y met coëfficiënten in $\mathbb{R}$ is, dan is de matrix

$$\begin{pmatrix} 1 - xy & -y^2 \\ x^2 & 1 + xy \end{pmatrix}$$

inverteerbaar over $\mathbb{R}[x, y]$ (wat is zijn inverse?). Maar in 1966 bewees Cohn dat deze matrix niet te schrijven is als een eindig product van elementaire matrices. Daarentegen bewees Suslin in 1977 dat iedere inverteerbare $n \times n$ matrix met veeltermcoëfficiënten wel een product van elementaire matrices is als $n \geq 3$. Deze stellingen behoren tot een tak van de wiskunde die *K*-theorie heet.

Hoofdstuk 3

Determinanten

3.1 Inleiding

In deze paragraaf bekijken we het stelsel $Ax = b$ waarbij A een $n \times n$ matrix is. We zagen in voorgaande dat, als A inverteerbaar is er precies één oplossing van dit stelsel bestaat. Men heeft eeuwen geprobeerd oplossingsformules voor zulke stelsels te vinden. Deze werden uiteindelijk gevonden in de 18-de eeuw en zijn beschreven door de regel van Cramer¹. Het blijkt dat iedere x_i een breuk is met een gemeenschappelijke noemer die een ingewikkelde uitdrukking is in de matrix elementen van A . Deze gemeenschappelijke noemer kreeg de naam “determinant van de matrix A ”. De determinant is een van de meest wonderlijke objecten, of zoals Jean Dieudonné (1906-1992) het noemt de “Deus ex machina”, van de lineaire algebra. Hij duikt overal op de meest onverwachte plaatsen op en is een onmisbaar gereedschap geworden in de meeste wetenschappen.

Om een idee te krijgen hoe de oplossingsformules van de x_i 's eruit zien gaan we eerst wat eenvoudige gevallen bekijken:

$n = 1$:

$$ax = b$$

Dan is $x = \frac{b}{a}$ de oplossingsformule, als $a \neq 0$.

$n = 2$:

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2\end{aligned}$$

We maken een herleiding tot het geval “ $n = 1$ ” door x_2 te elimineren: vermenigvuldig de eerste vergelijking met a_{22} en de tweede met a_{12} en trek het resultaat van de eerste af. Je vindt dan

$$(a_{22}a_{11} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2.$$

Dus $x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{22}a_{11} - a_{12}a_{21}}$, als de noemer niet nul is.

Net zo volgt $x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{22}a_{11} - a_{12}a_{21}}$, als de noemer niet nul is.

¹Gabriel Cramer, 1704-1752, Zwitserse wiskundige

Voor $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ definiëren we dan: de *determinant* van A als

$$\det A := a_{22}a_{11} - a_{12}a_{21}$$

m.a.w. $\det A$ is de gemeenschappelijke noemer van de breuken in de oplossingsformules voor x_1 en x_2 .

Kijkend naar de oplossingsformule voor x_1 zien we dat ook de teller van deze formule een determinant is nl.

$$\det \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix},$$

d.w.z. de determinant van de matrix ontstaan uit A door de eerste kolom van A te vervangen door de rechterzijde van het stelsel vergelijkingen, dus door de kolom $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$.

Net zo zien we dat de teller van de oplossingsformule voor x_2 precies gelijk is aan

$$\det \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix},$$

m.a.w. de determinant van de matrix ontstaan uit A door de tweede kolom van A te vervangen door $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$.

Keer nu even terug naar het geval $n = 1$. Dan zie je dat als je $\det(a) = a$ definieert, dan is de oplossingsformule voor x :

$$x = \frac{\det(b)}{\det(a)}$$

m.a.w. ook hier is de teller ontstaan door de eerste (en tevens enige) kolom van de matrix (a) te vervangen door de kolom (b) en dan de determinant “te nemen”. Samengevat: we vinden dezelfde structuur als bij $n = 2$!

De vraag is nu: zou deze structuur ook voor $n = 3$ (en hoger) gelden?

We kijken naar het geval $n = 3$, m.a.w. we bekijken het stelsel

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

We willen proberen een herleiding tot het geval $n = 2$ te maken, m.a.w. we gaan x_3 elimineren.

Vermenigvuldig daartoe de eerste vergelijking met a_{23} en de tweede met a_{13} en trek dan de tweede van de eerste verkregen vergelijking af. Dit geeft

$$(a_{23}a_{11} - a_{13}a_{21})x_1 + (a_{23}a_{12} - a_{13}a_{22})x_2 = a_{23}b_1 - a_{13}b_2.$$

Noem de nieuwe coëfficiënten van x_1 en x_2 resp. a'_{11} en a'_{12} en het nieuwe rechterlid b'_1 , d.w.z.

$$a'_{11} = a_{23}a_{11} - a_{13}a_{21}, \quad a'_{12} = a_{23}a_{12} - a_{13}a_{22}, \quad b'_1 = a_{23}b_1 - a_{13}b_2.$$

Net zo, vermenigvuldig de eerste vergelijking met a_{33} , de derde met a_{13} en trek dan de derde van de eerste af. Dit geeft

$$(a_{33}a_{11} - a_{13}a_{31})x_1 + (a_{33}a_{12} - a_{13}a_{32})x_2 = a_{33}b_1 - a_{13}b_3.$$

Noem de nieuwe coëfficiënten van x_1 en x_2 resp. a'_{21} en a'_{22} en het nieuwe rechterlid b'_2 , d.w.z.

$$a'_{21} = a_{33}a_{11} - a_{13}a_{31}, \quad a'_{22} = a_{33}a_{12} - a_{13}a_{32}, \quad b'_2 = a_{33}b_1 - a_{13}b_3.$$

We vinden dan het volgende stelsel vergelijkingen in twee onbekenden:

$$\begin{aligned} a'_{11}x_1 + a'_{12}x_2 &= b'_1 \\ a'_{21}x_1 + a'_{22}x_2 &= b'_2 \end{aligned}$$

Aannemende dat $a'_{11}a'_{22} - a'_{12}a'_{21} \neq 0$ vinden we dan volgens de oplossingsformules voor het geval $n = 2$:

$$x_1 = \frac{b'_1a'_{22} - b'_2a'_{12}}{a'_{11}a'_{22} - a'_{12}a'_{21}} \quad \text{en} \quad x_2 = \frac{b'_2a'_{11} - b'_1a'_{21}}{a'_{11}a'_{22} - a'_{12}a'_{21}}.$$

Uitschrijven van de teller en noemer van x_1 geeft:

$$(3.1) \quad \begin{aligned} b'_1a'_{22} - b'_2a'_{12} &= a_{13}(b_1a_{22}a_{23} + b_3a_{23}a_{12} + b_2a_{13}b_{32} - b_1a_{23}a_{32} - b_2a_{12}a_{33} - b_3a_{22}a_{13}) \\ a'_{11}a'_{22} - a'_{12}a'_{21} &= a_{13}(a_{11}a_{22}a_{33} + a_{31}a_{23}a_{12} + a_{21}a_{13}a_{32} - a_{11}a_{23}a_{32} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13}) \end{aligned}$$

In de formule van x_1 vallen de beide factoren a_{13} tegen elkaar weg. De noemer van de overblijvende formule voor x_1 noemen we dan, na analogie met de gevallen $n = 1$ en $n = 2$, de determinant van de coëfficiënten matrix A m.a.w.

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} := a_{11}a_{22}a_{33} + a_{31}a_{13}a_{12} + a_{21}a_{13}a_{32} - a_{11}a_{23}a_{32} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13}.$$

De vraag is nu: geldt dezelfde soort oplossingsformule als bij $n = 2$ (en $n = 1$) m.a.w. is de teller van de formule $x_1 = \frac{T}{\det A}$ te schrijven als de determinant van de matrix ontstaan

uit A door de eerste kolom van A te vervangen door de rechterzijde $\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ van het stelsel

vergelijkingen. Kijkend naar de formules 3.1 zien we dat de teller T inderdaad uit $\det A$ is ontstaan door a_{11} te vervangen door b_1 , a_{21} door b_2 en a_{31} door b_3 , m.a.w. het antwoord op onze vraag is Ja!

Eenzelfde rekenpartij levert inderdaad dat

$$x_i = \frac{\det A_i(b)}{\det A}$$

waarbij $A_i(b)$ de matrix is ontstaan uit A door de i -de kolom van A te vervangen door de

rechterzijde $\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, voor $i = 1, 2, 3$.

Voor het geval $n = 3$ hebben we dus ook oplossingsformules!

Hoe gaan we nu verder? Bekijk daartoe opnieuw de teller T van de formule voor x_1 , d.w.z.

$$\begin{aligned} T = \det \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix} &= b_1a_{22}a_{33} + b_3a_{23}a_{12} + b_2a_{13}a_{32} - b_1a_{23}a_{32} - b_2a_{12}a_{33} - b_3a_{22}a_{13} \\ &= b_1(a_{22}a_{33} - a_{23}a_{32}) + b_2(a_{13}a_{32} - a_{12}a_{33}) + b_3(a_{23}a_{12} - a_{22}a_{13}). \end{aligned}$$

Bekijk de uitdrukkingen die achter b_1 resp. b_2 resp. b_3 staan:

$$\begin{aligned} b_1 : \quad a_{22}a_{33} - a_{23}a_{32} &= \text{de determinant van de } 2 \times 2 \text{ matrix ontstaan uit } A \\ &\quad \text{de eerste rij en eerste kolom van } A \text{ te schrappen.} \\ b_2 : \quad a_{13}a_{32} - a_{12}a_{33} &= - \text{de determinant van de } 2 \times 2 \text{ matrix ontstaan uit } A \\ &\quad \text{door de tweede rij en eerste kolom van } A \text{ te schrappen.} \\ b_3 : \quad a_{23}a_{12} - a_{22}a_{13} &= \text{de determinant van de } 2 \times 2 \text{ matrix ontstaan uit } A \\ &\quad \text{door de derde rij en de eerste kolom van } A \text{ te schrappen.} \end{aligned}$$

Samengevat zien we

$$\det \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix} = \sum_{i=1}^3 (-1)^{i+1} b_i \det A_{(i1)}$$

waarbij $A_{(i1)}$ de 2×2 matrix is ontstaan uit A door de i -de rij en eerste kolom van A te schrappen. Als we nu voor b_i invullen a_{i1} m.a.w. vervang b_1 door a_{11} , b_2 door a_{21} en b_3 door a_{31} , dan zien we

$$(3.2) \quad \det A = \sum_{i=1}^3 (-1)^{i+1} a_{i1} \det A_{(i1)}.$$

We zien dus dat de determinant van de 3×3 matrix A hebben uitgedrukt m.b.v. de elementen uit de eerste kolom van A en de determinanten van zekere 2×2 matrices. Dit geeft ons een manier om de determinant van een $n \times n$ matrix inductief te definiëren.

3.2 Definitie en eigenschappen van de determinant

In deze paragraaf zullen we voor iedere $n \times n$ matrix de determinant ervan definiëren en verschillende eigenschappen ervan geven. Wiskundigen hebben zo'n twee eeuwen met determinanten geworsteld alvorens de theorie die nu behandeld gaat worden helemaal compleet was. Het zal de lezer dan ook niet verbazen dat sommige bewijzen nogal technisch zijn. Om de lijn van het verhaal niet uit het oog te verliezen zou de lezer het bewijs van stelling 3.2.5 even over kunnen slaan (maar bij de tweede doorgang natuurlijk wel lezen).

Definitie 3.2.1 We definiëren de determinant van een $n \times n$ matrix met inductie naar n :

- i) Als $n = 1$ en $A := (a)$ definiëren we: $\det A := a$.
- ii) Zij nu $n \geq 2$ en neem aan dat we voor $(n-1) \times (n-1)$ matrices de determinant ervan al gedefinieerd hebben.

Voor $1 \leq i, j \leq n$ zij $A_{(ij)}$ de matrix ontstaan uit A door de i -de rij en j -de kolom van A te schrappen. Dus $A_{(ij)}$ is een $(n-1) \times (n-1)$ matrix en dus is vanwege de inductieaanname $\det A_{(ij)}$ een welgedefinieerd reëel getal welke de *minor* van A op de plaats (i, j) heet.

- iii) Definieer nu

$$\tilde{a}_{ij} := (-1)^{i+j} \det A_{(ij)}.$$

Dit element uit $\mathbb{R}$ heet de *cofactor* van A op de plaats (i, j) .

- iv) Naar aanleiding van formule 3.2 uit sectie 3.1 definiëren we nu de *determinant* $\det A$ van een $n \times n$ matrix A door

$$\det A = \sum_{i=1}^n a_{i1} \tilde{a}_{i1} = \sum_{i=1}^n (-1)^{i+1} a_{i1} A_{(i1)}$$

waarbij $\tilde{a}_{i1}$ de cofactor van A op de plaats $(i, 1)$ is.

Opgave 3.2.2 Laat zien dat uit definitie 3.2.1 volgt dat $\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = a_1 b_2 - a_2 b_1$.

Voorbeeld 3.2.3 Zij $A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$. Dan geldt:

$$A_{(11)} = \begin{pmatrix} b_2 & c_2 \\ b_3 & c_3 \end{pmatrix}, \text{ dus } \tilde{a}_{11} = (-1)^{1+1} \det \begin{pmatrix} b_2 & c_2 \\ b_3 & c_3 \end{pmatrix} = b_2 c_3 - b_3 c_2;$$

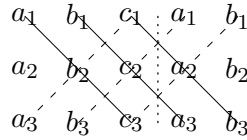
$$A_{(21)} = \begin{pmatrix} b_1 & c_1 \\ b_3 & c_3 \end{pmatrix}, \text{ dus } \tilde{a}_{21} = (-1)^{2+1} \det \begin{pmatrix} b_1 & c_1 \\ b_3 & c_3 \end{pmatrix} = b_1 c_3 - b_3 c_1;$$

$$A_{(31)} = \begin{pmatrix} b_1 & c_1 \\ b_2 & c_2 \end{pmatrix}, \text{ dus } \tilde{a}_{31} = (-1)^{3+1} \det \begin{pmatrix} b_1 & c_1 \\ b_2 & c_2 \end{pmatrix} = b_1 c_2 - b_2 c_1.$$

Hieruit volgt

$$\det A = a_1(b_2 c_3 - b_3 c_2) - a_2(b_1 c_3 - b_3 c_1) + a_3(b_1 c_2 - b_2 c_1).$$

Een mogelijke manier om het berekenen van de determinant van een 3×3 matrix makkelijk te onthouden is het volgende plaatje:



Schrijf de eerste twee kolommen nog eens rechts naast de matrix en teken dan de drie diagonalen (doorgetrokken lijnen) en de drie dwarsdiagonalen (stippellijnen). De determinant is dan de som van de producten op de drie diagonalen min de producten op de drie dwarsdiagonalen.

Waarschuwing: Dit schema voor 3×3 matrices geldt niet voor grotere matrices.

Lemma 3.2.4 Zij I_n de $n \times n$ eenheidsmatrix. Dan geldt $\det I_n = 1$.

Bewijs. Dit zien we rechtstreeks met volledige inductie in. Voor $n = 1$ is dit juist de definitie van de determinant. Zij nu $n \geq 2$ en neem aan dat al bewezen is dat $\det I_{n-1} = 1$. Er geldt $\det I_n = \sum_{i=1}^n a_{i1} \tilde{a}_{i1} = a_{11} \tilde{a}_{11} = 1 \cdot (-1)^2 \det I_{n-1} = 1$. $\square$

We leiden nu een aantal eigenschappen af waaraan determinanten voldoen. Om deze eigenschappen te kunnen beschrijven voeren we eerst wat handige notaties in.

Notatie: De rijen van een $n \times n$ matrix A noteren we als $a_1, \dots, a_n$ en i.p.v. $\det A$ schrijven we soms $d(A)$ of $d(a_1, \dots, a_n)$. Verder als $1 \leq i \leq n$ en b is een willekeurige rij van lengte n , dan noteren we de matrix die ontstaat uit A door de i -de rij van A te vervangen door b als $A_i[b]$.

Stelling 3.2.5 Zij A een $n \times n$ matrix, laten b en c rijen van lengte n zijn en noteer met $D_i(a)$, P_{ij} en $E_{ij}(c)$ de elementaire matrices zo als in sectie 2.4 gedefinieerd. Dan geldt:

- D1. $\det(A_i[b + c]) = \det(A_i[b]) + \det(A_i[c])$.
- D2. $\det D_i(a)A = a \det A$ voor iedere a in $\mathbb{R}$, m.a.w. de determinant van een matrix wordt met a vermenigvuldigd, als een rij met a vermenigvuldigd wordt.
- D3. $\det A = 0$ als A twee gelijke rijen bevat.
- D4. $\det P_{ij}A = -\det A$, voor alle $i \neq j$, m.a.w. de determinant van een matrix verandert van teken als je twee rijen ervan verwisselt.
- D5. $\det E_{ij}(c)A = \det A$, voor alle $i \neq j$ en alle c in $\mathbb{R}$, m.a.w. de determinant van een matrix verandert niet als je bij de i -de rij c keer de j -de rij optelt.

Bewijs. We bewijzen alle eigenschappen met inductie naar n . Als $n = 1$ zijn de beweringen D1 en D2 duidelijk en ook voor $n = 2$ gaat men (door expliciet uitschrijven van de determinant) eenvoudig na dat D1 t/m D5 waar zijn. Zij nu $n > 2$ en neem aan dat we de beweringen D1 t/m D5 voor $n - 1$ bewezen hebben.

- D1. Schrijf $B := A_i[b]$, $C := A_i[c]$ en $D := A_i[b + c]$. We moeten dan bewijzen dat $\det D = \det B + \det C$. Bekijk

$$(3.3) \quad \det D = d_{i1}\tilde{d}_{i1} + \sum_{h=1, h \neq i}^n d_{h1}\tilde{d}_{h1}.$$

Merk nu op dat als $h \neq i$ de matrix $D_{(h1)}$ aan de inductieaanname voldoet (ga na!). Hieruit volgt dat $\det D_{(h1)} = \det B_{(h1)} + \det C_{(h1)}$. Door met $(-1)^{h+1}$ te vermenigvuldigen vinden we

$$(3.4) \quad \tilde{d}_{h1} = \tilde{b}_{h1} + \tilde{c}_{h1} \text{ als } h \neq i.$$

Verder zien we dat als $h \neq i$ dan

$$(3.5) \quad d_{h1} = b_{h1} = c_{h1} (= a_{h1})$$

want A , B , C en D verschillen alleen maar in de i -de rij. Uit 3.4 en 3.5 volgt dus

$$(3.6) \quad \sum_{h=1, h \neq i}^n d_{h1}\tilde{d}_{h1} = \sum_{h=1, h \neq i}^n d_{h1}(\tilde{b}_{h1} + \tilde{c}_{h1}) = \sum_{h=1, h \neq i}^n b_{h1}\tilde{b}_{h1} + \sum_{h=1, h \neq i}^n c_{h1}\tilde{c}_{h1}.$$

Aan de andere kant geldt voor $h = i$ dat $d_{i1} = b_{i1} + c_{i1}$ en $\tilde{d}_{i1} = \tilde{b}_{i1} = \tilde{c}_{i1} (= \tilde{a}_{i1})$ en dus

$$(3.7) \quad d_{i1}\tilde{d}_{i1} = (b_{i1} + c_{i1})\tilde{d}_{i1} = b_{i1}\tilde{b}_{i1} + c_{i1}\tilde{c}_{i1}.$$

Uit 3.3, 3.6 en 3.7 volgt D1.

- D2. Deze eigenschap wordt als oefening aan de lezer overgelaten.
- D3. Laat $1 \leq i < j \leq n$ en neem aan dat $a_j = a_i$. Zij $h \neq i$ en $h \neq j$. Volgens de inductieaanname is $\det A_{(h1)} = 0$, want $A_{(h1)}$ bevat twee gelijke rijen. Dus $\tilde{a}_{(h1)} = 0$ en dus $\det A = a_{i1}\tilde{a}_{i1} + a_{j1}\tilde{a}_{j1} = a_{i1}(\tilde{a}_{i1} + \tilde{a}_{j1})$ want $a_{i1} = a_{j1}$. We moeten dus bewijzen dat geldt

$$(3.8) \quad \tilde{a}_{i1} + \tilde{a}_{j1} = 0.$$

Voor $1 \leq p \leq n$ definiëren we nu a'_p als de rij van lengte $n - 1$ ontstaan uit a_p door het eerste element weg te laten. Dan geldt $\tilde{a}_{i1} + \tilde{a}_{j1} = (-1)^{i+1}d_1 + (-1)^{j+1}d_2$ waarbij

$$\begin{aligned} d_1 &:= d(a'_1, \dots, a'_{i-1}, a'_{i+1}, \dots, a'_{j-1}, a'_i, a'_{j+1}, \dots, a'_n) \quad \text{en} \\ d_2 &:= d(a'_1, \dots, a'_{i-1}, a'_i, a'_{i+1}, \dots, a'_{j-1}, a'_{j+1}, \dots, a'_n). \end{aligned}$$

Merk op dat de matrix met de rijen als in d_2 ontstaan is uit de matrix met rijen als in d_1 door $(j - 1) - i$ verwisselingen uit te voeren, namelijk a'_i gaat van de $(j - 1)$ -ste plaats naar de i -de plaats en iedere verwisseling geeft volgens de inductieaanname voor D4 een minteken. Dus $d_2 = (-1)^{j-1-i}d_1$ en dus $\tilde{a}_{i1} + \tilde{a}_{j1} = (-1)^{i+1}d_1 + (-1)^{j+1}(-1)^{j-1-i}d_1 = 0$, waarmee 3.8 bewezen is en dus D3.

D4. Eerst een notatie: als b en c willekeurige rijen zijn van lengte n en $i \neq j$ dan schrijven we $d[b, c]$ i.p.v. $d(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_{j-1}, c, a_{j+1}, \dots, a_n)$. Volgens D3 geldt dan $d[b + c, b + c] = 0$. Pas dan D1 tweemaal toe. Dit levert

$$\begin{aligned} 0 = d[b + c, b + c] &= d[b, b + c] + d[c, b + c] \\ &= d[b, b] + d[b, c] + d[c, b] + d[c, c] \\ &= d[b, c] + d[c, b] \quad (\text{vanwege D3}). \end{aligned}$$

en hieruit volgt D4.

D5. Dit kunnen we zonder inductie rechtstreeks uit de reeds bewezen eigenschappen afleiden. Volgens 2.4.12 geldt $E_{ij}(c)A = A_i[a_i + ca_j]$. Dus met D1, D2 en D3 volgt

$$\begin{aligned} \det E_{ij}(c)A &= \det(A_i[a_i]) + \det(A_i[ca_j]) = \det A + c \det(A_i[a_j]) \\ &= \det A + c \cdot 0 = \det A. \end{aligned}$$

□

Opgave 3.2.6 Leid uit D2, D4 en D5 de volgende eigenschappen af:

- i) $\det E_{ij}(a) = 1$, $\det P_{ij} = -1$ en $\det D_i(a) = a$.
- ii) $\det EB = \det E \det B$ voor iedere $n \times n$ matrix B en iedere elementaire matrix E .
- iii) Leid uit ii) af met inductie naar s dat

$$\det E_1 \dots E_s B = \det E_1 \dots \det E_s \cdot \det B = \det(E_1 \dots E_s) \cdot \det B$$

voor ieder stel elementaire matrices $E_1, \dots, E_s$.

De eigenschappen D2, D4 en D5 beschrijven hoe een determinant zich gedraagt onder vegen. Het uitrekenen van een determinant gaat in de praktijk vaak als volgt: men “veegt” m.b.v. de Gauss eliminatie methode een matrix op *bovendriehoeks vorm* d.w.z. $a_{ij} = 0$ voor alle $i > j$ en gebruikt dan het volgende resultaat

Stelling 3.2.7 Zij $A = (a_{ij})$ een bovendriehoeks matrix. Dan geldt $\det A = a_{11} \cdot a_{22} \dots a_{nn}$ m.a.w. $\det A$ is het product der diagonaal elementen (een zelfde resultaat geldt voor een *onderdriehoeks matrix* d.w.z. $a_{ij} = 0$ voor alle $i < j$).

Bewijs. Met inductie naar n . Juist voor $n = 1$. Laat nu $n \geq 2$ en neem aan dat de stelling al bewezen is voor iedere $(n - 1) \times (n - 1)$ bovendriehoeks matrix. Omdat A een bovendriehoeks matrix is geldt $a_{21} = a_{31} = \dots = a_{n1} = 0$. Dus is volgens definitie 3.2.1

$\det A = a_{11}\tilde{a}_{11} = a_{11}(-1)^{1+1} \det A_{(11)}$.

Maar $A_{(11)}$ is een bovendriehoeks matrix met op de diagonaal de elementen $a_{22}, \dots, a_{nn}$. Dus geldt volgens de inductieaanname dat $\det A_{(11)} = a_{22} \dots a_{nn}$ en dus $\det A = a_{11} \det A_{(11)} = a_{11}a_{22} \dots a_{nn}$. $\square$

Voorbeeld 3.2.8 Zij $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \\ 3 & 0 & 4 \end{pmatrix}$. Bereken $\det A$.

Oplossing. Door tweemaal D5 te gebruiken vinden we $\det A = \det \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & -2 \end{pmatrix}$. Dus

$\det A = -2$ vanwege 3.2.7.

Een voor theoretische doeleinden belangrijke stelling maar ook een van de motivaties om naar zoiets als de determinant te kijken is:

Stelling 3.2.9 Zij A een $n \times n$ matrix. Dan geldt A is inverteerbaar d.e.s.d.a. $\det A \neq 0$.

Bewijs. i) Als A inverteerbaar is bestaan er volgens 2.4.11 elementaire matrices $E_1, \dots, E_s$ zodanig dat $A = E_1 E_2 \dots E_s$. Dan volgt uit opgave 3.2.6 iii), met $B = I_n$, dat $\det A = \det E_1 \dots \det E_s$ en dus $\det A \neq 0$ vanwege 3.2.6 i).

ii) Neem nu aan dat A niet inverteerbaar is. We gebruiken inductie over n . Voor $n = 1$ is $A = (0)$ (anders zou A wel inverteerbaar zijn), dus is $\det A = 0$. Zij nu $n \geq 2$ en neem aan de bewering is voor $n - 1$ bewezen. Als de eerste kolom van A de nul kolom is, dan volgt meteen uit definitie 3.2.1 dat $\det A = 0$. Neem dus aan dat de eerste kolom van A niet de nul kolom is. Zoals aangetoond in het bewijs van 2.4.11 kunnen we dan elementaire matrices $E_1, \dots, E_s$ vinden zodat $B := E_1 \dots E_s A$ als eerste kolom $(1 \ 0 \dots 0)^t$ heeft en een rechtsonder matrix B_{n-1} . Uit definitie 3.2.1 volgt dan dat $\det B = \det B_{n-1}$.

Merk op dat B niet inverteerbaar is (want als B inverteerbaar is dan is ook $A = E_s^{-1} \dots E_1^{-1} B$ inverteerbaar, volgens lemma 2.4.10 herhaald toegepast, maar A is niet inverteerbaar). Maar dan volgt uit opgave 3.2.10 hieronder dat B_{n-1} ook niet inverteerbaar is. Dan volgt uit de inductieaanname dat $\det B_{n-1} = 0$ en dus dat $\det B = 0$. Maar dan volgt uit 3.2.6 dat $\det A = 0$, immers $A = E_s^{-1} \dots E_1^{-1} B$ en ieder E_i^{-1} is weer een elementaire matrix. $\square$

Opgave 3.2.10 Zij B een $n \times n$ matrix met als eerste kolom $(1 \ 0 \dots 0)^t$ en een rechtsonder $(n - 1) \times (n - 1)$ matrix B_{n-1} , dus van de vorm

$$B = \left(\begin{array}{c|ccc} 1 & b_{12} & \dots & b_{1n} \\ \hline 0 & & & \\ \vdots & & B_{n-1} & \\ 0 & & & \end{array} \right)$$

Bewijs dat als B_{n-1} inverteerbaar is, B het ook is. (Aanw.: vermenigvuldig B met $\begin{pmatrix} 1 & 0 \\ 0 & B_{n-1}^{-1} \end{pmatrix}$ en gebruik 2.4.10.)

Een verdere cruciale eigenschap van de determinant is dat ze *multiplicatief* is, d.w.z. dat $\det(AB) = \det A \cdot \det B$. Om dit te bewijzen hebben we echter de uitspraak van stelling 2.4.6 nodig die we nu in een equivalente formulering gaan bewijzen.

Stelling 3.2.11 Zij A een $n \times n$ matrix. Als A niet inverteerbaar is, dan bestaat er een $x \neq 0$ met $Ax = 0$.

Omgekeerd, als zo'n $x \neq 0$ met $Ax = 0$ niet bestaat, moet A dus inverteerbaar zijn.

Bewijs. Als A niet inverteerbaar is, is ook A^t niet inverteerbaar (want voor een inverse B van A^t is B^t een inverse van A). We brengen nu A^t d.m.v. elementaire transformaties op bovendriehoeksvorm: $E_s \dots E_1 A^t = D$ met $D_{ij} = 0$ voor $i > j$. Omdat A^t niet inverteerbaar is, zijn niet alle $D_{ii} \neq 0$, want anders zou $\det A^t \neq 0$ zijn. Zij nu k de grootste index met $D_{kk} = 0$, dan is $D_{jj} \neq 0$ voor $j > k$. Door nu voor $j = k+1, \dots, n$ het passende veelvoud van de j -de rij bij de k -de op te tellen, kunnen we stapsgewijs de hele k -de rij op 0 brengen. Met de bijhorende elementaire matrices $E'_{k+1}, \dots, E'_n$ geldt dus dat

$$\underbrace{E'_n \dots E'_{k+1} E_s \dots E_1}_{E} A^t = EA^t = \begin{pmatrix} d_{11} & * & \dots & & * \\ 0 & d_{22} & & & \vdots \\ \vdots & & \ddots & & \\ & & & 0 & \dots & 0 \\ & & & & d_{k+1,k+1} & * \\ & & & & \ddots & \\ 0 & \dots & & & & d_{nn} \end{pmatrix}$$

Omdat EA^t in de k -de rij een 0-rij heeft, heeft $(EA^t)^t = AE^t$ in de k -de kolom een 0-kolom. Met e_k noteren we de kolom met 1 in de k -de plaats en 0 elders, dan is $A(E^t e_k) = 0$, maar $E^t e_k \neq 0$ omdat E^t een inverteerbare matrix is. $\square$

Stelling 3.2.12 Zij A en B $n \times n$ matrices. Dan geldt

$$\det AB = \det A \cdot \det B.$$

Bewijs. i) Als A inverteerbaar is bestaan er volgens 2.4.11 elementaire matrices $E_1, \dots, E_s$ met $A = E_1 \dots E_s$. Uit opgave 3.2.6 iii) volgt dan $\det AB = \det A \cdot \det B$.

ii) Als A niet inverteerbaar is, dan is AB het ook niet. Stel namelijk dat AB wel inverteerbaar is, dan volgt uit 3.2.11 dat B inverteerbaar is (want als B niet inverteerbaar is, is er $x \neq 0$ met $Bx = 0$, maar dan is ook $ABx = 0$ en dus AB niet inverteerbaar, tegenspraak). Maar dan is B^{-1} inverteerbaar en dus is met 2.4.10 $A = (AB)B^{-1}$ ook inverteerbaar, tegenspraak. Dus is inderdaad AB niet inverteerbaar en dus volgens 3.2.9 $\det A = 0$ en $\det AB = 0$. Dus $\det AB = 0 = 0 \cdot \det B = \det A \cdot \det B$. $\square$

Stelling 3.2.13 Zij A een $n \times n$ matrix. Dan geldt $\det A^t = \det A$.

Bewijs. i) Voor elementaire matrices volgt de stelling uit 3.2.6 i).

Met inductie naar s volgt dan, gebruik makend van 3.2.12, dat $\det A = \det A^t$ als A een product van s elementaire matrices is. M.a.w. vanwege 2.4.11 geldt de stelling als A inverteerbaar is.

ii) Als A niet inverteerbaar is is A^t het ook niet (immers als $A^t C = C A^t = I_n$ voor een zekere $n \times n$ matrix C , dan volgt uit 2.2.17 dat $C^t A = A C^t = I_n$ m.a.w. A is inverteerbaar, een tegenspraak). Dus met 3.2.9 $\det A = 0 = \det A^t$. $\square$

Kijkend naar definitie 3.2.1 zien we dat de eerste kolom van A een speciale rol speelt. Dit is maar schijn, immers zo'n zelfde formule geldt ook voor iedere kolom. Precieser

Stelling 3.2.14 Voor iedere $n \times n$ matrix geldt

- i) $\sum_{i=1}^n a_{ij} \tilde{a}_{ij} = \det A$, voor alle $1 \leq j \leq n$.
- ii) $\sum_{i=1}^n a_{ih} \tilde{a}_{ij} = 0$, voor alle $1 \leq j, h \leq n$ met $j \neq h$.

Bewijs. i) Bekijk de matrix A' ontstaan uit A door de j -de kolom te schrappen en hem vervolgens helemaal naar links te brengen zodat hij de eerste kolom van A' wordt.

Hiervoor zijn $j - 1$ verwisselingen nodig (van plaats j naar plaats 1). Uit 3.2.13 en D4 volgt dan $\det A' = (-1)^{j-1} \det A$. Merk nu op dat $a'_{i1} = a_{ij}$ en $\tilde{a}'_{i1} = (-1)^{i+1} \det A_{(ij)}$. Dus

$$\det A' = \sum_{i=1}^n a'_{i1} \tilde{a}'_{i1} = \sum_{i=1}^n a_{ij} (-1)^{i+1} \det A_{(ij)}$$

waaruit volgt dat

$$\det A = (-1)^{j-1} \sum_{i=1}^n a_{ij} (-1)^{i+1} \det A_{(ij)} = \sum_{i=1}^n a_{ij} (-1)^{i+j} \det A_{(ij)} = \sum_{i=1}^n a_{ij} \tilde{a}_{ij}.$$

ii) Zij A' de matrix ontstaan uit A door de j -de kolom te vervangen door de h -de kolom van A . Dan is A' een matrix met twee gelijke kolommen en dus volgens D3 en 3.2.13 geldt $\det A' = 0$. Pas nu de formule uit i) toe op A' i.p.v. A en merk op dat $a'_{ij} = a_{ih}$ en $\tilde{a}'_{ij} = \tilde{a}_{ij}$. Hieruit volgt dan meteen de gevraagde formule in ii). $\square$

Opmerking 3.2.15 Formule i) uit 3.2.14 heet de *Laplace² ontwikkeling van det A* volgens de j -de kolom. Ook kunnen we de determinant van een matrix uitrekenen via zijn *Laplace ontwikkeling volgens de i -de rij*: dat zien we in de volgende opgave.

Opgave 3.2.16 Zij A een $n \times n$ matrix en $1 \leq i, h \leq n$ met $i \neq h$. Bewijs dat geldt

$$\det A = \sum_{j=1}^n a_{ij} \tilde{a}_{ij} \quad \text{en} \quad \sum_{j=1}^n a_{hj} \tilde{a}_{ij} = 0.$$

(Aanw.: gebruik 3.2.13 en 3.2.14.)

De resultaten uit 3.2.14 en 3.2.16 kunnen we als volgt samenvatten: bij een $n \times n$ matrix A definiëren we een nieuwe matrix, de zgn. *geadjungeerde matrix van A*, genoteerd $\text{adj } A$ als volgt:

$$(\text{adj } A)_{ij} := \tilde{a}_{ji} \quad (\text{let op volgorde } i \text{ en } j!)$$

De relaties uit 3.2.13 en 3.2.14 geven dan onmiddellijk

Stelling 3.2.17 $A \cdot \text{adj } A = \text{adj } A \cdot A = \det A \cdot I_n$.

Gevol 3.2.18 Als A inverteerbaar is dan geldt

$$A^{-1} = (\det A)^{-1} \text{adj } A.$$

De formule uit 3.2.18 wordt nauwelijks toegepast om de inverse matrix A^{-1} echt te berekenen, hiervoor is het algoritme aan het eind van hoofdstuk 2 veel geschikter. Maar ze is voor bepaalde uitspraken over A^{-1} wel handig. Bijvoorbeeld ziet men rechtstreeks dat voor een matrix A met alleen maar gehele getallen als elementen a_{ij} de inverse A^{-1} uit breuken bestaat die hoogstens $\det A$ als noemer hebben.

²Pierre-Simon Laplace, 1749-1827, Franse wiskundige

3.3 Toepassingen

In deze paragraaf behandelen we een aantal toepassingen van de theorie der determinanten.

a) De regel van Cramer

Stelling 3.3.1 (Regel van Cramer) Zij A een $n \times n$ matrix en $b \in \mathbb{R}^n$. Dan geldt:

$$\text{als } Ax = b, \text{ dan } \det A \cdot x_i = \det A_i(b) \text{ voor alle } i.$$

Bewijs. Laten $A_1, \dots, A_n$ de kolommen van A aanduiden. Dan zijn $a_1 := A_1^t, \dots, a_n := A_n^t$ de rijen van $\mathcal{A} := A^t$. Uit $Ax = b$ volgt dan $b = x_1 A_1 + \dots + x_n A_n$ en dus $b^t = x_1 a_1 + \dots + x_n a_n$. Volgens 3.2.13 geldt dan

$$\det A_i(b) = \det \mathcal{A}_i[b^t] = \det \mathcal{A}_i[x_1 a_1 + \dots + x_n a_n].$$

Volgens D1 en D2 is deze laatste determinant gelijk aan

$$x_1 \det \mathcal{A}_i[a_1] + \dots + x_i \det \mathcal{A}_i[a_i] + \dots + x_n \det \mathcal{A}_i[a_n].$$

Omdat volgens D3 $\det \mathcal{A}_i[a_j] = 0$ als $j \neq i$ vinden we

$$\det A_i(b) = x_i \det \mathcal{A}_i[a_i] = x_i \det \mathcal{A} = x_i \det A.$$

□

Merk op dat voor het geval $\det A \neq 0$, dus voor een inverteerbare matrix A de regel van Cramer juist de algemene versie van de oplossingsformule voor stelsels lineaire vergelijkingen is, die we in sectie 3.1 voor de gevallen $n = 1, 2, 3$ hadden gevonden. De aanvullende uitspraak voor niet inverteerbare A is dat het stelsel alleen maar oplosbaar kan zijn als voor iedere kolom i geldt dat $\det A_i(b) = 0$.

Als toepassing van de regel van Cramer geven we een bewijs van de cosinusregel uit de vlakke meetkunde.

Voorbeeld 3.3.2 (Cosinusregel) Gegeven een driehoek ABC met hoeken α , β en γ . De zijden tegenover de hoeken A , B en C hebben lengte respectievelijk a , b en c . Leidt de cosinusregel af, d.w.z. bewijs dat geldt

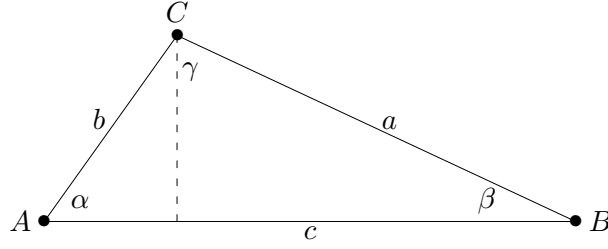
$$a^2 = b^2 + c^2 - 2bc \cos \alpha.$$

Oplossing.

Men ziet onmiddellijk uit de tekening in Figuur 3.1 dat $b \cos \alpha + a \cos \beta = c$ en analoog geldt dat $c \cos \beta + b \cos \gamma = a$ en $c \cos \alpha + a \cos \gamma = b$. We vatten deze uitspraken op als een stelsel van drie lineaire vergelijkingen in de onbekenden $\cos \alpha$, $\cos \beta$ en $\cos \gamma$ d.w.z.

$$\begin{pmatrix} 0 & c & b \\ c & 0 & a \\ b & a & 0 \end{pmatrix} \begin{pmatrix} \cos \alpha \\ \cos \beta \\ \cos \gamma \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

met de regel van Cramer volgt dan $2abc \cos \alpha = ab^2 + ac^2 - a^3$ waaruit $a^2 = b^2 + c^2 - 2bc \cos \alpha$ volgt.



Figuur 3.1: Illustratie voor de cosinusregel

b) De Vandermonde determinant

We zullen nu de determinant van een heel speciale matrix, de zgn. Vandermonde³ matrix, uitrekenen. Deze matrix blijkt in veel problemen op te duiken. We zullen hem o.a. gebruiken om stelling 1.3.1 uit hoofdstuk 1 te bewijzen.

Beschouw $n \geq 2$ reële getallen $a_1, \dots, a_n$. De *Vandermonde* matrix $V(a_1, \dots, a_n)$ is de $n \times n$ matrix met als i -de rij $(a_1^{i-1}, \dots, a_n^{i-1})$, waarbij we met a^0 altijd 1 bedoelen.

Stelling 3.3.3 Voor de Vandermonde matrix $V(a_1, \dots, a_n)$ geldt

$$\det V(a_1, \dots, a_n) = \prod_{1 \leq j < i \leq n} (a_i - a_j).$$

Bewijs. Met inductie naar n : om het schrijfwerk te beperken doen we de redenering voor $n = 4$, maar het zal duidelijk zijn dat deze redenering algemeen werkt. Zij dus

$$V = \begin{pmatrix} 1 & 1 & 1 & 1 \\ a_1 & a_2 & a_3 & a_4 \\ a_1^2 & a_2^2 & a_3^2 & a_4^2 \\ a_1^3 & a_2^3 & a_3^3 & a_4^3 \end{pmatrix}.$$

Trek a_1 keer de derde rij van de vierde af, vervolgens a_1 keer de tweede van de derde en vervolgens a_1 keer de eerste van de tweede. Omdat volgens D5 daardoor de determinant niet verandert vinden we

$$\det V = \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 & a_4 - a_1 \\ 0 & a_2^2 - a_1 a_2 & a_3^2 - a_1 a_3 & a_4^2 - a_1 a_4 \\ 0 & a_2^3 - a_1 a_2^2 & a_3^3 - a_1 a_3^2 & a_4^3 - a_1 a_4^2 \end{pmatrix}.$$

Uit definitie 3.2.1 volgt dan

$$\det V = \det \begin{pmatrix} a_2 - a_1 & a_3 - a_1 & a_4 - a_1 \\ a_2^2 - a_1 a_2 & a_3^2 - a_1 a_3 & a_4^2 - a_1 a_4 \\ a_2^3 - a_1 a_2^2 & a_3^3 - a_1 a_3^2 & a_4^3 - a_1 a_4^2 \end{pmatrix} = \det \begin{pmatrix} a_2 - a_1 & a_3 - a_1 & a_4 - a_1 \\ a_2(a_2 - a_1) & a_3(a_3 - a_1) & a_4(a_4 - a_1) \\ a_2^2(a_2 - a_1) & a_3^2(a_3 - a_1) & a_4^2(a_4 - a_1) \end{pmatrix}.$$

Nu bevat de eerste kolom een factor $a_2 - a_1$, de tweede kolom een factor $a_3 - a_1$ en de derde kolom een factor $a_4 - a_1$. Vanwege D2 en 3.2.13 volgt dan

$$\det V = (a_2 - a_1)(a_3 - a_1)(a_4 - a_1) \det \begin{pmatrix} 1 & 1 & 1 \\ a_2 & a_3 & a_4 \\ a_2^2 & a_3^2 & a_4^2 \end{pmatrix}.$$

³Alexandre-Théophile Vandermonde, 1735-1796, Franse wiskundige

De matrix $\begin{pmatrix} 1 & 1 & 1 \\ a_2 & a_3 & a_4 \\ a_2^2 & a_3^2 & a_4^2 \end{pmatrix}$ is een 3×3 Vandermonde matrix en dus is zijn determinant vanwege de inductieaanname gelijk aan $(a_3 - a_2)(a_4 - a_2)(a_4 - a_3)$. Totaal zien we dan dat

$$\det V = (a_2 - a_1)(a_3 - a_1)(a_4 - a_1)(a_3 - a_2)(a_4 - a_2)(a_4 - a_3).$$

□

Bewijs van stelling 1.3.1 We moeten laten zien dat er precies één kromme van de vorm $y = a_0 + a_1x + \dots + a_{n-1}x^{n-1}$ bestaat die door de punten $(c_1, d_1), \dots, (c_n, d_n)$ gaat. Een punt (c_i, d_i) ligt op de kromme $y = a_0 + a_1x + \dots + a_{n-1}x^{n-1}$ d.e.s.d.a.

$$(3.9) \quad a_0 + a_1c_i + \dots + a_{n-1}c_i^{n-1} = d_i.$$

We moeten dus $a_0, a_1, \dots, a_{n-1}$ zo bepalen dat ze voor ieder i aan de vergelijking 3.9 voldoen m.a.w. we moeten $a_0, a_1, \dots, a_{n-1}$ oplossen uit

$$\begin{pmatrix} 1 & c_1 & \dots & c_1^{n-1} \\ 1 & c_2 & \dots & c_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & c_n & \dots & c_n^{n-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} d_1 \\ \vdots \\ \vdots \\ d_n \end{pmatrix}$$

Omdat gegeven is dat alle c_i verschillend zijn is volgens 3.2.3 $\det C^t \neq 0$, waarbij C de matrix in het linkerlid is. Dus ook $\det C \neq 0$. Maar dan is C vanwege 3.2.9 inverteerbaar en dus bestaat er volgens 2.4.4 precies een oplossing die aan $C(a_0, \dots, a_{n-1})^t = (d_1, \dots, d_n)^t$ voldoet. □

Opgave 3.3.4 Laten $c_1, \dots, c_n$ verschillende reële getallen zijn.

Bewijs dat de functies $e^{c_1x}, \dots, e^{c_nx}$ lineair onafhankelijk zijn d.w.z. laat zien dat geldt: als $\lambda_1, \dots, \lambda_n$ reële getallen zijn zodanig dat $\lambda_1 e^{c_1x} + \dots + \lambda_n e^{c_nx} = 0$ voor alle $x \in \mathbb{R}$, dan is $\lambda_1 = \dots = \lambda_n = 0$.

(Aanw.: vul $x = 0, 1, \dots, n-1$ in de relatie in.)

c) Kortste paden in een graaf

We hadden in 2.3.4 beloofd dat er een slimme methode bestaat om voor een graaf met adjacency matrix A de lengte van het kortste pad tussen twee punten P_i en P_j te bepalen zonder de machten A, A^2 , enz. te hoeven berekenen. Deze methode is een toepassing van de geadjungeerde matrix.

We weten al dat de lengte d van het kortste pad van P_i en P_j het kleinste getal d is zo dat het element $(A^d)_{ij} \neq 0$ is. In een eerste stap passen we een trucje toe dat in heel verschillende situaties (met name in de discrete wiskunde) erg nuttig blijkt, we vermenigvuldigen de adjacency matrix A met een variabele t . Nu kijken we naar de matrix

$$M := I_n + (tA) + (tA)^2 + (tA)^3 + \dots = I_n + tA + t^2A^2 + t^3A^3 + \dots$$

waarbij we de som oneindig door laten gaan. Ieder element van de matrix M is dus een *machtrees* in t . We maken ons hier geen zorgen over convergentie, we identificeren een *machtrees* $\sum_{i=0}^{\infty} a_i t^i$ gewoon met de rij $(a_0, a_1, a_2, \dots)$.

Omdat voor $k < d$ het element $(A^k)_{ij} = 0$ is, weten we dat M_{ij} een machtreeks is die een term $a_d t^d$ als term van laagste graad heeft.

Nu is een machtreeks van de vorm $1 + x + x^2 + \dots$ een *meetkundige reeks* en er geldt

$$1 + x + x^2 + \dots = \frac{1}{1-x} = (1-x)^{-1}.$$

Als we deze formule voor $x = tA$ toepassen, zien we dat

$$M = (I_n - tA)^{-1}$$

Dat $I_n - tA$ een inverteerbare matrix is volgt uit het feit dat $\det(I_n - tA)$ een veelterm $p(t)$ is met $p(0) = 1$, want $\det(I_n - 0.A) = \det I_n = 1$. Er geldt dus i.h.b. dat

$$\det(I_n - tA) = p(t) = 1 + c_1 t + c_2 t^2 + \dots + c_n t^n.$$

Nu passen we 3.2.17 op $I_n - tA$ toe, dit geeft

$$\det(I_n - tA) \cdot M = p(t) \cdot (I_n - tA)^{-1} = \text{adj}(I_n - tA)$$

Omdat we alleen maar in de laagste graad van het element M_{ij} geïnteresseerd zijn, kunnen we de vermenigvuldiging met $p(t)$ negeren, want die verandert de graad niet. We bereiken zo het volgende resultaat:

Stelling 3.3.5 De lengte d van het kortste pad van P_i naar P_j is de graad van de laagste term in $\text{adj}(I_n - tA)_{ij}$, d.w.z. de graad van de laagste term in $\det(I_n - tA)_{(ji)}$ (let op de volgorde van de indices).

Voorbeeld 3.3.6 Bepaal de lengte van het kortste pad van P_1 naar P_3 in de graaf met

$$\text{adjacency matrix } A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} \quad (\text{de graaf uit Figuur 2.2}).$$

Oplossing. Volgens 3.3.5 moeten we de term van laagste graad in $\det(I_n - tA)_{(31)}$ bepalen, dus in de determinant van de matrix

$$\begin{pmatrix} 1 & -t & 0 & 0 & 0 \\ -t & 1 & 0 & 0 & -t \\ 0 & -t & 1 & 0 & -t \\ 0 & 0 & -t & 1 & -t \\ 0 & -t & 0 & -t & 1 \end{pmatrix}_{(31)} = \begin{pmatrix} -t & 0 & 0 & 0 \\ 1 & 0 & 0 & -t \\ 0 & -t & 1 & -t \\ -t & 0 & -t & 1 \end{pmatrix} =: B$$

Ontwikkeling volgens de eerste rij geeft

$$\det B = -t \det \begin{pmatrix} 0 & 0 & -t \\ -t & 1 & -t \\ 0 & -t & 1 \end{pmatrix} = -t \cdot (-t)^3 = t^4$$

De term van laagste graad is dus t^4 en het kortste pad van P_1 naar P_3 heeft dus lengte 4.

Historische opmerkingen

1. Het begrip *determinant* verscheen voor het eerst in 1683 in een manuscript van de Japanse wiskundige Seki Takakazu (1642-1708). Seki's manuscript bevat gedetailleerde berekeningen van determinanten van 2×2 , 3×3 en 4×4 matrices. Zijn determinanten hadden een tegengesteld teken van de onze! Hij gebruikte de determinanten om zekere typen niet lineaire vergelijkingen op te lossen. Seki werkte het grootste deel van zijn leven als accountant voor twee leenheren in Kofu, ten westen van Tokio. In 1693 verscheen de determinant in een brief van Gottfried von Leibniz (1646-1716) aan de Markies van l'Hôpital (1661-1704). Hij bekeek het stelsel

$$10 + 11x + 12y = 0$$

$$20 + 21x + 22y = 0$$

$$30 + 31x + 32y = 0$$

en vond een criterium om te besluiten of dit stelsel een oplossing heeft nl.

$$10.21.32 + 11.22.30 + 12.20.31 = 10.22.31 + 11.20.32 + 12.21.30$$

m.a.w. de determinant van de coëfficiënten matrix moet nul zijn.

De theorie der determinanten ontstond door werk van verschillende wiskundigen aan het eind van de 18-de en het begin van de 19-de eeuw. Met name Gabriël Cramer (1704-1752), Etienne Bézout (1730-1783) in 1764 en Alexandre-Théophile Vandermonde (1735-1796) in 1771 geven verschillende methoden om determinanten uit te rekenen.

Pierre-Simon Laplace (1749-1827) formuleerde en bewees de regel dat door verwisseling van twee naast elkaar gelegen kolommen de determinant met een minteken verandert en dat de determinant van een matrix met twee gelijke kolommen nul is.

Het meest volledige werk in die tijd was dat van Augustin Louis Cauchy (1789-1857) in 1812. In dit werk voerde hij de naam determinant en de dubbele index notatie in en bewees dat men de determinant kan uitrekenen door ontwikkeling via iedere rij en kolom.

2. De *regel van Cramer* verscheen voor het eerst in zijn algemeenheid in "Introduction to the Analysis of Algebraic Curves", 1750 van Gabriël Cramer. Hij was geïnteresseerd in het probleem om de vergelijking van een vlakke kromme van gegeven graad te vinden die gaat door een gegeven aantal punten. Bijvoorbeeld, de algemene tweede-graads kromme, wiens vergelijking gegeven wordt door

$$a + by + cx + dy^2 + exy + x^2 = 0$$

is bepaald door vijf punten. Om, gegeven vijf punten, a, b, c, d en e te bepalen, substitueerde Cramer de coördinaten van de punten in de vergelijking en vond vijf lineaire vergelijkingen in vijf onbekenden. Hij verwees dan naar de appendix van zijn werk, waarin hij de algemene regel als volgt beschreef:

"Men vindt de waarde van iedere onbekende door n breuken te vormen die een gemeenschappelijke noemer hebben die uit net zoveel termen bestaat als er permutaties bestaan van n dingen".

Vervolgens beschreef hij precies hoe die termen en hun juiste teken te vinden zijn. Cramer beschreef echter niet waarom zijn regel werkte! Het bewijs voor de gevallen

$n = 2$ en $n = 3$ verscheen in “A Treatise of Algebra” by Colin Maclaurin (1698-1746), en werd pas in 1748 na zijn dood gepubliceerd. Over het algemene geval werd door Maclaurin niets gezegd.

3. *Alexandre-Théophile Vandermonde* (1735-1796) werd in Parijs geboren als zoon van een dokter. Hij studeerde muziek en verkreeg zijn licentiaat in 1757. Zijn instrument was de viool. Hij volgde een carrière in de muziek totdat hij op 35-jarige leeftijd overstapte op Wiskunde. Een jaar later werd hij gekozen tot lid van de Akademie van Wetenschappen. In de volgende twee jaar leverde hij in vier artikelen een belangrijke bijdrage aan de Wiskunde. De jaren daarna publiceerde hij meerdere artikelen over wetenschap en muziek: hij werkte samen met Bézout en de scheikundige Lavoisier. Ook werkte hij samen met Monge over hoe je staal moet bereiden. In 1778 schreef hij twee artikelen over muziek. In z'n tweede artikel formuleerde hij het idee dat musici zich niets moesten aantrekken van alle (wiskundige) theoriën over muziek, maar alleen moesten afgaan op hun getrainde oren om muziek te beoordelen. Dit leidde er uiteindelijk toe dat de Akademie van Wetenschappen muziek verplaatste van het gebied der Wetenschappen naar het gebied der Kunsten. Het is opmerkelijk dat juist een groot wiskundige als Vandermonde tegen muziek als wiskundige kunst schreef, een positie die de muziek innam sinds de Grieken.

Vandermonde is het meest bekend vanwege de *Vandermonde determinant*. Het is zeker dat hij een zeer belangrijke bijdrage heeft geleverd aan de theorie der determinanten, echter in zijn vier wiskundige artikelen komt deze determinant niet voor. Het is dus nogal vreemd dat deze determinant zijn naam draagt. Een aannemelijk vermoeden, afkomstig van Lebesgue, is dat iemand Vandermondes notatie verkeerd gelezen heeft en daardoor dacht dat de determinant in zijn werk voorkwam!

Hoofdstuk 4

Eigenwaarden en eigenvectoren

4.1 Inleiding

Tot nu toe zijn al onze vectoren en matrices reëel geweest d.w.z. de theorie voor stelsels lineaire vergelijkingen en de theorie der matrices en determinanten zijn behandeld voor de reële getallen. Echter het zal de lezer geen enkele moeite kosten om overal waar “reëel getal”, $\mathbb{R}$ of $\mathbb{R}^n$ staat, dit te vervangen door “complex getal”, $\mathbb{C}$ of $\mathbb{C}^n$; alle stellingen en bewijzen gaan gewoon door.

Bij het nu volgende onderwerp, eigenwaarden en eigenvectoren, is het van belang over de complexe getallen te kunnen beschikken, ook al start je met een matrix bestaande uit reële (of zelfs gehele) getallen. Je kunt het vergelijken met zoiets als de vergelijking $x^2 + 1 = 0$; deze heeft reële coëfficiënten maar geen reële nulpunten, echter wel complexe, namelijk i en $-i$.

Het kunnen vinden van nulpunten in de complexe getallen lukt altijd: dit is de zgn. *Hoofdstelling van de Algebra* die door Gauss in 1799 bewezen werd. Het precieze resultaat is:

Stelling 4.1.1 (Hoofdstelling van de Algebra)

Laten $a_0, a_1, \dots, a_{n-1}$ willekeurige complexe getallen zijn. Dan bestaan er complexe getallen $\lambda_1, \dots, \lambda_n$ zodanig dat

$$x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 = (x - \lambda_1)(x - \lambda_2) \dots (x - \lambda_n),$$

m.a.w. de vergelijking $x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 = 0$ heeft precies n complexe nulpunten, namelijk $\lambda_1, \dots, \lambda_n$ (de λ_i 's hoeven niet allen verschillend te zijn).

In dit hoofdstuk zal A steeds een complexe $n \times n$ matrix aanduiden en alle berekeningen gaan over $\mathbb{C}$ i.p.v. over $\mathbb{R}$ zoals tot nu toe. I.h.b. zal de determinant $\det A$ een complex getal zijn.

4.2 Definitie van eigenwaarden en eigenvectoren

In sectie 3.2 zagen we hoe we aan iedere $n \times n$ matrix een getal, $\det A$, toegevoegd hebben. We zullen nu een verfijning hiervan geven: aan iedere $n \times n$ matrix A voegen we n complexe getallen toe, de zgn. eigenwaarden van A . Het getal $\det A$ zal blijken, op een minteken na,

het product van alle eigenwaarden te zijn. De eigenwaarden en eigenvectoren (zie 4.2.1 hieronder) blijken een fundamentele rol te spelen in allerlei problemen zowel binnen als buiten de wiskunde. We zullen hier niet de hele theorie behandelen (we verwijzen daarvoor naar de colleges LA3 en LA4) maar wel een voor de praktijk zeer belangrijk geval, namelijk wanneer alle eigenwaarden verschillend zijn. Als illustratie van het gebruik van eigenwaarden en eigenvectoren geven we een aantal toepassingen.

Definitie 4.2.1 Zij A een $n \times n$ matrix. Een getal c heet een *eigenwaarde* van A als er een niet-nul vector x in $\mathbb{C}^n$ bestaat met $Ax = cx$. De vector x heet dan een *eigenvector* van A bij de eigenwaarde c .

Voorbeeld 4.2.2 Zij $A = \begin{pmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{pmatrix}$, het magisch vierkant uit hoofdstuk 1 welke 15 als magische som heeft.

Men rekent dan eenvoudig na dat 15 een eigenwaarde van A is met als eigenvector $e := (1 \ 1 \ 1)^t$ m.a.w. $Ae = 15e$.

De volgende stelling geeft aan hoe we alle eigenwaarden van een matrix kunnen bepalen.

Stelling 4.2.3 Zij A een $n \times n$ matrix en c een complex getal. Dan is c een eigenwaarde van A d.e.s.d.a. $\det(cI_n - A) = 0$.

Bewijs. Er geldt dat c een eigenwaarde van A is d.e.s.d.a. er een $x \in \mathbb{C}^n$, $x \neq 0$ bestaat met $Ax = cx$ d.e.s.d.a. er $x \in \mathbb{C}^n$, $x \neq 0$ bestaat met $(cI_n - A)x = 0$. Uit 2.4.5 en 3.2.11 zien we dat dit het geval is d.e.s.d.a. $cI_n - A$ niet inverteerbaar is m.a.w. d.e.s.d.a. $\det(cI_n - A) = 0$, volgens 3.2.9. $\square$

Voorbeeld 4.2.4 Zij $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$. Bepaal de eigenwaarden van A en bij iedere eigenwaarde al zijn eigenvectoren.

Oplossing. Volgens 4.2.3 is c een eigenwaarde van A d.e.s.d.a. $\det(cI_2 - A) = 0$ d.e.s.d.a. $\det \begin{pmatrix} c-1 & -1 \\ 2 & c-4 \end{pmatrix} = 0$ d.e.s.d.a. $(c-1)(c-4) + 2 = 0$ d.e.s.d.a. $c^2 - 5c + 6 = 0$ d.e.s.d.a. $c = 2$ of $c = 3$. M.a.w. $c = 2$ en $c = 3$ zijn alle eigenwaarden van A .

Om de eigenvectoren bij $c = 2$ te vinden, moeten we dus alle $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{C}^2$ bepalen met $Ax = 2x$ m.a.w. met

$$\begin{array}{ll} x_1 + x_2 = 2x_1 & \text{ofwel:} \quad -x_1 + x_2 = 0 \\ -2x_1 + 4x_2 = 2x_2 & \quad -2x_1 + 2x_2 = 0 \end{array}$$

We zien dus dat alle oplossingen van de vorm $\begin{pmatrix} x_1 \\ x_1 \end{pmatrix}$ zijn met $x_1 \in \mathbb{C}$. Dus iedere eigenvector van A bij de eigenwaarde $c = 2$ is van de vorm $x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ met $x_1 \neq 0$. Net zo vinden we dat alle eigenvectoren bij de eigenwaarde $c = 3$ van de vorm $x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ met $x_1 \neq 0$ zijn.

Een cruciale stelling waardoor eigenwaarden en eigenvectoren zowel binnen als buiten de wiskunde gebruikt worden is de volgende:

Stelling 4.2.5 Zij A een $n \times n$ matrix. Neem aan dat de n eigenwaarden $c_1, \dots, c_n$ van A allen verschillend zijn. Laat $v_1, \dots, v_n$ een stel corresponderende eigenvectoren zijn d.w.z. $Av_i = cv_i$ voor iedere i . Zij T de $n \times n$ matrix waarvan voor iedere i de i -de kolom gelijk is aan v_i . Dan geldt:

- 1) T is een inverteerbare matrix.
- 2) $T^{-1}AT = D$, waarbij D de diagonaalmatrix is met $D_{ii} = c_i$ voor alle i , m.a.w. de eigenwaarden van A staan op de diagonaal van D .

Het bewijs van deze stelling is gebaseerd op het volgende lemma.

Lemma 4.2.6 Laat $p \geq 1$ en $v_1, \dots, v_p$ eigenvectoren zijn bij verschillende eigenwaarden $c_1, \dots, c_p$ van A . Als $x_1v_1 + \dots + x_pv_p = 0$ voor zekere $x_i \in \mathbb{C}$, dan is $x_1 = \dots = x_p = 0$.

Bewijs. Met inductie naar p . Omdat $v_1 \neq 0$ (want een eigenvector is niet nul) is het geval $p = 1$ duidelijk. Laat nu $p \geq 2$. Door de relatie van links met A te vermenigvuldigen en te gebruiken dat $Av_i = c_iv_i$ voor iedere i volgt

$$(*) \quad x_1c_1v_1 + \dots + x_pc_pv_p = 0.$$

Door aan de andere kant $x_1v_1 + \dots + x_pv_p = 0$ met c_p te vermenigvuldigen volgt

$$(**) \quad x_1c_pv_1 + \dots + x_pc_pv_p = 0.$$

Uit (*) en (**) volgt dan $x_1(c_1 - c_p)v_1 + \dots + x_{p-1}(c_{p-1} - c_p)v_{p-1} = 0$. Uit de inductieaanname volgt dan dat $x_i(c_i - c_p) = 0$ voor alle $1 \leq i \leq p-1$. Omdat $c_i - c_p \neq 0$ voor alle $1 \leq i \leq p-1$ volgt dat $x_1 = \dots = x_{p-1} = 0$. Dit invullend in de oorspronkelijke relatie, gebruikend dat $v_p \neq 0$, levert dat ook $x_p = 0$. $\square$

Bewijs van stelling 4.2.5

1) Om te bewijzen dat T inverteerbaar is, is het volgens stelling 3.2.11 voldoende te laten zien dat $Tx = 0$ alleen maar kan voor $x = 0$. Neem dus aan dat $Tx = 0$ m.a.w. (zie sectie 2.2 in hoofdstuk 2)

$$x_1v_1 + \dots + x_nv_n = 0.$$

Dan volgt uit 4.2.6 dat $x_1 = \dots = x_n = 0$, dus $x = 0$.

2) $AT = A(v_1, \dots, v_n) = (Av_1, \dots, Av_n) = (c_1v_1, \dots, c_nv_n) = TD$. Dus $T^{-1}AT = D$. $\square$

Definitie 4.2.7 Een matrix A heet *diagonaliseerbaar* als er een inverteerbare $n \times n$ matrix T bestaat en een diagonaalmatrix D zodanig dat $T^{-1}AT = D$.

Voorgaande stelling zegt dus dat iedere $n \times n$ matrix die n verschillende complexe eigenwaarden heeft diagonaliseerbaar is. Voordat we wat toepassingen van deze stelling bekijken, geven we eerst een eenvoudig, maar nuttig lemma

Lemma 4.2.8 i) Voor een diagonaalmatrix $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$ is $D^m = \begin{pmatrix} d_1^m & & 0 \\ & \ddots & \\ 0 & & d_n^m \end{pmatrix}$

voor alle $m \geq 1$.

ii) Zij D een willekeurige $n \times n$ matrix en T een inverteerbare $n \times n$ matrix. Dan is $(TDT^{-1})^m = TD^mT^{-1}$ voor alle $m \geq 1$.

Bewijs. Met inductie naar m . De details worden aan de lezer overgelaten (merk op dat $(TAT^{-1})(TBT^{-1}) = TA(T^{-1}T)BT^{-1} = TABT^{-1}$). $\square$

4.3 Toepassingen

a) Dynamische systemen

We laten nu zien hoe eigenwaarden en eigenvectoren een rol spelen bij de bestudering van zgn. dynamische systemen. Grofweg gesproken is een dynamisch systeem een “systeem” dat op ieder moment in een bepaalde toestand verkeert, maar dat met de tijd verandert. Zo heb je chemische systemen, economische systemen enz. Deze kunnen allen als dynamische systemen gezien worden.

Men onderscheidt continue en discrete dynamische systemen; in het laatste geval vindt de verandering meer stapsgewijs plaats bijvoorbeeld iedere dag, of iedere week enz. Dit alles klinkt nogal vaag! Het volgende voorbeeld uit de biologie zal het hopelijk allemaal wat duidelijker maken.

De gevlekte uil

We bekijken een model dat de populatie dynamica (d.w.z. de verandering van de populatie met de tijd) beschrijft van de gevlekte uil in de bossen rond Willow Creek in California. De populatie in het k -de jaar wordt beschreven door een vector $p_k := (j_k, h_k, v_k)$ in $\mathbb{R}^3$, welke de *toestandsvector* heet. Hierbij duiden j_k , h_k resp. v_k aan het aantal jonge vrouwtjes (in eerste levensjaar), het aantal halfvolwassen vrouwtjes (in tweede levensjaar) en het aantal volwassen vrouwtjes in het k -de jaar. Jaarlijkse metingen hebben geleid tot de volgende informatie: het aantal jonge vrouwtjes in het $(k+1)$ -de jaar is 33% van het aantal volwassen vrouwtjes in het k -de jaar d.w.z. $j_{k+1} = 0.33v_k$. Ook overleefden 18% van de jonge vrouwtjes om het volgende jaar halfvolwassen vrouwtjes te worden d.w.z. $h_{k+1} = 0.18j_k$. Tenslotte overleefde 71% van de halfvolwassen vrouwtjes en 94% van de volwassen vrouwtjes het volgend jaar om volwassen vrouwtjes te worden resp. blijven m.a.w.

$$v_{k+1} = 0.71h_k + 0.94v_k.$$

De cruciale vraag die de onderzoekers zich stelden was: zal de gevlekte uil overleven?

Bovenstaande informatie kan geschreven worden in de vorm van een zgn. discreet lineair dynamisch systeem d.w.z. $p_{k+1} = Ap_k$, waarbij

$$A = \begin{pmatrix} 0 & 0 & 0.33 \\ 0.18 & 0 & 0 \\ 0 & 0.71 & 0.94 \end{pmatrix} \quad \text{en} \quad p_k = \begin{pmatrix} j_k \\ h_k \\ v_k \end{pmatrix}.$$

De vraag is dus: wat gebeurt er met p_k als $k \rightarrow \infty$.

Merk op dat $p_{k+1} = Ap_k = A^2_{p_{k-1}} = \dots = A^k p_1$. De vraag die we dus moeten onderzoeken is: wat gebeurt er met A^k als $k \rightarrow \infty$? Daartoe berekenen we de eigenwaarden van A : deze zijn (afgerond op vier decimalen) $c_1 = 0.9836$, $c_2 = -0.0218 + 0.2059i$ en $c_3 = -0.0218 - 0.2059i$. Omdat ze allen verschillend zijn, bestaat er volgens stelling 4.2.5 een inverteerbare complexe matrix T met $T^{-1}AT = D$, waarbij D de diagonaalmatrix is met c_1, c_2, c_3 op de diagonaal. Dus $A = TDT^{-1}$ en dus met 4.2.6 $A^k = TD^kT^{-1}$ voor alle $k \geq 1$. Omdat $|c_i| < 1$ voor iedere i gaat $|c_i^k| \rightarrow 0$ als $k \rightarrow \infty$ en dus $D^k \rightarrow 0$ als $k \rightarrow \infty$. We zien dus dat

$$v_{k+1} = A^k v_1 = TD^kT^{-1}v_1 \rightarrow 0 \text{ als } l \rightarrow \infty$$

m.a.w. de gevlekte uil sterft uit!

De Fibonacci rij

De Fibonacci rij is de volgende rij getallen

$$1, 1, 2, 3, 5, 8, 13, 21, 34, \dots$$

De rij wordt beschreven door de regels

$$a_0 = 1, a_1 = 1 \text{ en } a_{n+2} = a_{n+1} + a_n \text{ voor alle } n \geq 0.$$

Vraag: Kun je een formule geven voor a_n ?

Omdat een term uit de rij bepaald is door de twee voorafgaande termen kunnen we dit probleem als volgt opvatten als een (discreet) dynamisch systeem:

Definieer $v_n := \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$ als de toestandsvector van het systeem. Dan geldt

$$v_{n+1} = \begin{pmatrix} a_{n+2} \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} a_{n+1} + a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$$

m.a.w.

$$v_{n+1} = Av_n, \text{ waarbij } A := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Net zoals in het vorige voorbeeld krijgen we dan $v_n = A^n v_0$, waarbij $v_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. De eigenwaarden van A zijn

$$c_1 = \frac{1}{2}(1 + \sqrt{5}) \text{ (de gulden snede)} \quad \text{en} \quad c_2 = \frac{1}{2}(1 - \sqrt{5}).$$

De bijbehorende eigenvectoren zijn

$$\begin{pmatrix} \frac{1}{2}(1 + \sqrt{5}) \\ 1 \end{pmatrix} \quad \text{en} \quad \begin{pmatrix} \frac{1}{2}(1 - \sqrt{5}) \\ 1 \end{pmatrix}$$

Dus als we nemen $T = \begin{pmatrix} \frac{1}{2}(1 + \sqrt{5}) & \frac{1}{2}(1 - \sqrt{5}) \\ 1 & 1 \end{pmatrix}$, dan geldt volgens stelling 4.2.5 dat

$$T^{-1}AT = D, \text{ waarbij } D = \begin{pmatrix} c_1 & \\ & c_2 \end{pmatrix}.$$

Dan volgt uit 4.2.8 dat $A^n = T \begin{pmatrix} c_1^n & \\ & c_2^n \end{pmatrix} T^{-1}$ en dus

$$v_n = A^n v_0 = T \begin{pmatrix} c_1^n & \\ & c_2^n \end{pmatrix} T^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

De waarden voor T , c_1 en c_2 invullen geeft dan dat a_n , welke de tweede component van v_n is, gelijk is aan

$$a_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^{n+1} - \left(\frac{1 - \sqrt{5}}{2} \right)^{n+1} \right) \text{ voor alle } n \geq 0!$$

b) Stochastische matrices

Bij veel dynamische systemen heeft de matrix A de eigenschap dat alle elementen van A niet-negatief zijn en dat in iedere kolom de som der elementen gelijk is aan 1. In dit geval noemt men A vaak een *stochastische matrix*: de rijen en kolommen van A worden geïdentificeerd met toestanden en het element A_{ij} geeft de kans aan waarmee het systeem van toestand j naar toestand i over gaat (let op de volgorde). Omdat het systeem vanuit toestand j ergens naartoe moet, geldt $\sum_{i=1}^n A_{ij} = 1$, dus de som over de j -de kolom is 1. In dit geval laat zich aantonen dat 1 altijd een eigenwaarde is en dat alle eigenwaarden absolute waarde ≤ 1 hebben.

Onder zekere voorwaarden (waaraan meestal voldaan is) geldt zelfs dat er op schalingen na een unieke eigenvector met eigenwaarde 1 bestaat en dat alle andere eigenwaarden van absolute waarde < 1 zijn. De precieze resultaten staan bekend als *stellingen van Perron*¹ en *Frobenius*² over niet-negatieve matrices.

Als we voor zo'n systeem een basis uit eigenvectoren zo kiezen dat de eerste basisvector de eigenvector $v = (v_1 \ v_2 \ \dots \ v_n)^t$ met eigenwaarde 1 is, volgt dat D^m voor grote m naar de matrix D_∞ gaat waarbij het $(1,1)$ -element 1 is en alle andere elementen 0 zijn. Maar dan gaat A^m naar

$$\begin{aligned} A_\infty &= T \cdot D_\infty \cdot T^{-1} = T \cdot \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & 0 \end{pmatrix} \cdot T^{-1} \\ &= \begin{pmatrix} v_1 & 0 & \dots & 0 \\ v_2 & 0 & \dots & 0 \\ & & \ddots & \\ v_n & 0 & \dots & 0 \end{pmatrix} \cdot T^{-1} = \begin{pmatrix} s_1 v_1 & s_2 v_1 & \dots & s_n v_1 \\ s_1 v_2 & s_2 v_2 & \dots & s_n v_2 \\ & & \ddots & \\ s_1 v_n & s_2 v_n & \dots & s_n v_n \end{pmatrix} \end{aligned}$$

waarbij $(s_1, \dots, s_n)$ de elementen in de eerste rij van T^{-1} zijn.

Men gaat echter eenvoudig na, dat voor een stochastische matrix A ook de machten A^2 , A^3 enz. stochastische matrices zijn, want dit zijn de kansen op overgangen in 2, 3 enz. stappen. I.h.b. is dus ook de limiet A_∞ een stochastische matrix. Dit betekent dat alle kolommen van A_∞ som 1 moeten hebben en dus geldt

$$s_1 = s_2 = \dots = s_n = \left(\sum_{i=1}^n v_i \right)^{-1}.$$

Als we dus de eigenvector v al zo normeren dat $\sum_{i=1}^n v_i = 1$ is, dan geldt $s_1 = s_2 = \dots = s_n = 1$ en we vinden

$$A_\infty = \begin{pmatrix} v_1 & v_1 & \dots & v_1 \\ v_2 & v_2 & \dots & v_2 \\ \vdots & \vdots & & \vdots \\ v_n & v_n & \dots & v_n \end{pmatrix}$$

dus is A_∞ een matrix waarin alle kolommen gelijk zijn.

¹Oskar Perron, 1880-1975, Duitse wiskundige

²Georg Frobenius, 1849-1917, Duitse wiskundige

Bij een dynamisch systeem met zo'n stochastische matrix A zal dus iedere begintoestand $(b_1 \ b_2 \ \dots \ b_n)^t$ uiteindelijk naar dezelfde evenwichtstoestand streven (tot op een schaling na), namelijk

$$\begin{pmatrix} v_1 & v_1 & \dots & v_1 \\ v_2 & v_2 & \dots & v_2 \\ \vdots & \vdots & & \vdots \\ v_n & v_n & \dots & v_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = n \left(\sum_{i=1}^n b_i \right) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Verspreiding van de Euro munten

Begin 2002 zijn de Euro munten geïntroduceerd en toen vroeg men zich af, hoe de verspreiding van de munten uit de verschillende landen gemodelleerd zou kunnen worden. Er waren inderdaad wiskundige onderzoeksgroepen bezig, die regelmatig *Euro scouts* bericht lieten geven hoeveel van de verschillende munten ze op een gegeven moment in hun portemonnee hadden om zo de verdere verspreiding te kunnen voorspellen.

In een vereenvoudigd model kunnen we *Euroland* splitsen in Nederland, België, Luxemburg en de rest. Om het proces van de verspreiding van de munten te beschrijven, moeten we alleen maar afschatten, hoeveel munten in een jaar vanuit een land naar een ander land gaan en hoeveel er in het land blijven.

We beschouwen de verspreiding van de munten nu als een systeem, waarbij de mogelijke toestanden de landen zijn, waar een munt zich bevindt. Als volgorde van de toestanden (en dus van de rijen en kolommen van de matrix A) leggen we vast: NL, B, L, rest. Voor de kansen van de overgangen schatten we nu (enigszins willekeurig) de matrix

$$A = \begin{pmatrix} 0.7 & 0.1 & 0.03 & 0.006 \\ 0.05 & 0.8 & 0.02 & 0.003 \\ 0.05 & 0.05 & 0.9 & 0.001 \\ 0.2 & 0.05 & 0.05 & 0.99 \end{pmatrix}$$

De eerste kolom geeft bijvoorbeeld aan dat 70% van de munten uit Nederland in Nederland blijven, 5% naar België gaan, 5% naar Luxemburg en de resterende 20% naar andere landen. Andersom komen 10% van de Belgische munten naar Nederland (element a_{12}), 3% van de munten uit Luxemburg en 0.6% van de munten uit andere landen. Merk op dat we ter vereenvoudiging aannemen dat geen munten verdwijnen, hierdoor wordt de kolomsom inderdaad 1.

Als we nu willen weten waar de Nederlandse munten na k jaren terecht zijn gekomen, hoeven we alleen maar de vector $A^k e_1$ te bepalen (met $e_1 = (1 \ 0 \ 0 \ 0)^t$), net zo $A^k e_2$ voor de Belgische munten enz. De kolommen van A^k geven dus de verdeling van de verschillende soorten munten na k jaar aan, de rijen de mix van munten in één land.

Voor de eigenwaarden van A vinden we nu (bij benadering)

$$c_1 = 1, \quad c_2 = 0.92, \quad c_3 = 0.81, \quad c_4 = 0.66$$

We zijn dus precies in de situatie van de stelling over stochastische matrices. De eigenvector voor de eigenwaarde $c_1 = 1$ die als som van zijn componenten 1 heeft is (op drie decimalen

afgerond) $(0.030 \ 0.025 \ 0.037 \ 0.908)^t$, en hieruit volgt dat

$$A^k \rightarrow \begin{pmatrix} 0.030 & 0.030 & 0.030 & 0.030 \\ 0.025 & 0.025 & 0.025 & 0.025 \\ 0.037 & 0.037 & 0.037 & 0.037 \\ 0.908 & 0.908 & 0.908 & 0.908 \end{pmatrix}$$

De Euro munten streven dus naar een evenwichtstoestand waarbij in ieder land dezelfde relatieve hoeveelheid munten uit ieder land terecht komen, maar in Nederland zijn uiteindelijk 3.0% van alle munten, in België 2.5%, in Luxemburg 3.7% en in de rest van de landen zitten 90.8% van de munten.

c) Het idee achter GOOGLE

Het probleem bij een zoekmachine voor het internet is niet zo zeer, om überhaupt documenten te vinden, die bij een zoekaanvraag (een *query*) passen, maar om de passende documenten in een volgorde te brengen zo dat de meest interessante documenten bovenaan staan.

Voor documenten die alle woorden uit een query bevatten, werden bij de vroege zoekmachines onder meer de volgende criteria voor de rangschikking gebruikt:

- Een korter document staat hoger in de lijst, want hoe langer een document hoe eerder kunnen de gezochte woorden er toevallig in voorkomen.
- Hoe dichter de woorden uit de query bij elkaar staan, hoe hoger staat het document in de lijst.
- Als de gezochte woorden meerdere keren voorkomen, wordt het document hoger geplaatst.

Met het snel toenemende aantal beschikbare documenten in het internet werden deze technieken steeds minder bevredigend, vaak moest men de interessante links diep in de lijst van resultaten zoeken.

Het succes van GOOGLE berust juist op een nieuwe manier van rangschikking die ertoe leidt dat in veel gevallen de meest interessante documenten ook op de eerste plaatsen in de lijst van gevonden documenten staan. Het idee van *Sergey Brin* en *Lawrence Page* was, de documenten volgens hun *relevantie* te rangschikken, waarbij ze relevantie als volgt beschrijven:

De relevantie van een document is evenredig met de som van de relevanties der documenten die erop verwijzen.

Dit lijkt in eerste instantie op een circulaire redenering, want om de relevantie van een document te bepalen, moeten we de relevanties van de documenten die erop verwijzen al kennen: Als bijvoorbeeld de documenten 2, 4 en 8 op document 1 verwijzen, de documenten 3, 5, 7 en 11 op 2 en de documenten 1, 4, 9 en 16 op 3, dan krijgen we een stelsel lineaire vergelijkingen als het volgende (waarbij we met r_i de relevantie van document i noteren en k de evenredigheidsfactor is):

$$\begin{aligned} r_1 &= k(r_2 + r_4 + r_8) \\ r_2 &= k(r_3 + r_5 + r_7 + r_{11} + r_{13}) \\ r_3 &= k(r_1 + r_4 + r_9 + r_{16}) \end{aligned}$$

Om bijvoorbeeld r_3 te berekenen, moeten we r_1 al kennen, maar hiervoor hebben we r_2 nodig die wederom van r_3 afhangt.

Maar met behulp van een geschikte matrix kunnen we uit deze vicieuze cirkel ontsnappen door het probleem naar een vraag over eigenwaarden te vertalen. Als de zoekmachine over N documenten in het internet kan zoeken, maken we een $N \times N$ matrix A , waarbij

$$a_{i,j} = \begin{cases} 1 & \text{als document } j \text{ naar document } i \text{ verwijst;} \\ 0 & \text{als document } j \text{ niet naar document } i \text{ verwijst.} \end{cases}$$

Voor de vector $r = (r_1 \ r_2 \ \dots \ r_N)^t$ der relevanties geldt dan

$$r = k \cdot Ar \text{ d.w.z. } r \text{ is een eigenwaarde van } A.$$

De matrix A is een matrix waarin alle elementen ≥ 0 zijn, dit noemt men een *niet-negatieve* matrix. Voor dit soort matrices laat zich aantonen, dat een eigenvector voor de grootste eigenwaarde alleen maar positieve getallen bevat (of negatieve, maar dan kunnen we met -1 vermenigvuldigen). De eigenvectoren voor de andere eigenwaarden hebben zowel positieve als negatieve elementen.

Men definieert nu de relevantie van het i -de document als het element r_i in de eigenvector r van A voor de grootste eigenwaarde. Omdat we achteraf alleen maar relevanties gaan vergelijken, speelt een schaling van r met een positieve factor geen rol.

We merken nog op dat er slimme numerieke methoden bestaan om voor gigantisch grote matrices die *dun bezet* zijn (d.w.z. waarvoor in iedere rij en kolom slechts een paar elementen ongelijk aan 0 zijn) eigenwaarden en eigenvectoren te bepalen.

Een 'eerlijke' tabel voor de eredivisie

Het idee om met behulp van de componenten van een eigenvector een rangschikking te bepalen, laat zich ook op andere vraagstukken toepassen. Als voorbeeld kijken we hier naar het bepalen van een tabel voor de eredivisie.

We nemen aan dat de verschillende teams verschillende speelsterktes $r_1, r_2, \dots, r_{18}$ hebben. De grondgedachte is nu dat het moeilijker (en dus waardevoller) is om tegen een sterke ploeg te winnen dan tegen een zwakke club. Daarom is een overwinning van het i -de team niet altijd 3 punten waard, maar $3r_i$ punten en een gelijkspel r_i punten (de gewone punten worden dus met de speelsterkte vermenigvuldigd). De analoge uitspraak met het principe achter GOOGLE is nu:

De sterkte van een team is evenredig met het aantal behaalde punten, telkens gewogen met de sterkte van het team waar de punten gewonnen zijn.

We noteren nu in de matrix A met het element a_{ij} het aantal punten dat team i tegen team j heeft gewonnen. Voor de vector $r = (r_1 \ r_2 \ \dots \ r_{18})^t$ van speelsterktes geldt dan net als boven dat $r = k \cdot Ar$.

Ook in dit geval is A een niet-negatieve matrix, dus heeft de eigenvector voor de grootste eigenwaarde slechts positieve componenten. Als we de eigenvector zo normeren, dat de som der componenten juist het aantal punten is dat van alle teams samen in de gewone tabel behaald werd, kunnen we onze alternatief bepaalde tabel goed met de gewone tabel vergelijken. Dit doen we voor het seizoen 2006/2007. De matrix A ziet er als volgt uit, waarbij de teams in de volgorde van de afsluitende gewone tabel zijn aangegeven:

$$A = \begin{pmatrix} 0 & 3 & 3 & 3 & 4 & 3 & 4 & 6 & 4 & 3 & 4 & 6 & 4 & 6 & 6 & 4 & 6 & 6 \\ 3 & 0 & 2 & 4 & 3 & 3 & 6 & 6 & 6 & 4 & 6 & 3 & 3 & 6 & 6 & 4 & 4 & 6 \\ 3 & 2 & 0 & 1 & 6 & 4 & 1 & 4 & 6 & 4 & 6 & 6 & 6 & 4 & 6 & 3 & 6 & 4 \\ 3 & 1 & 4 & 0 & 6 & 1 & 3 & 4 & 4 & 6 & 4 & 4 & 3 & 1 & 4 & 6 & 6 & 6 \\ 1 & 3 & 0 & 0 & 0 & 3 & 3 & 4 & 1 & 6 & 4 & 4 & 4 & 3 & 6 & 3 & 6 & 4 \\ 3 & 3 & 1 & 4 & 3 & 0 & 1 & 0 & 1 & 4 & 6 & 1 & 4 & 4 & 6 & 6 & 1 & 6 \\ 1 & 0 & 4 & 3 & 3 & 4 & 0 & 0 & 3 & 1 & 3 & 6 & 6 & 1 & 4 & 6 & 4 & 4 \\ 0 & 0 & 1 & 1 & 1 & 6 & 6 & 0 & 0 & 4 & 4 & 3 & 3 & 6 & 3 & 6 & 4 & 3 \\ 1 & 0 & 0 & 1 & 4 & 4 & 3 & 6 & 0 & 3 & 3 & 3 & 2 & 1 & 3 & 6 & 4 & 4 \\ 3 & 1 & 1 & 0 & 0 & 1 & 4 & 1 & 3 & 0 & 6 & 4 & 3 & 4 & 0 & 1 & 6 & 6 \\ 1 & 0 & 0 & 1 & 1 & 0 & 3 & 1 & 3 & 0 & 0 & 6 & 6 & 1 & 4 & 6 & 6 & 4 \\ 0 & 3 & 0 & 1 & 1 & 4 & 0 & 3 & 3 & 1 & 0 & 0 & 6 & 4 & 4 & 1 & 3 & 4 \\ 1 & 3 & 0 & 3 & 1 & 1 & 0 & 3 & 2 & 3 & 0 & 0 & 0 & 4 & 4 & 3 & 3 & 6 \\ 0 & 0 & 1 & 4 & 3 & 1 & 4 & 0 & 4 & 1 & 4 & 1 & 1 & 0 & 1 & 3 & 1 & 3 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 3 & 3 & 6 & 1 & 1 & 1 & 4 & 0 & 3 & 4 & 3 \\ 1 & 1 & 3 & 0 & 3 & 0 & 0 & 0 & 0 & 4 & 0 & 4 & 3 & 3 & 3 & 0 & 1 & 4 \\ 0 & 1 & 0 & 0 & 0 & 4 & 1 & 1 & 1 & 0 & 0 & 3 & 3 & 4 & 1 & 4 & 0 & 4 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 3 & 1 & 0 & 1 & 1 & 0 & 3 & 3 & 1 & 1 & 0 \end{pmatrix}$$

De grootste eigenwaarde van A is (afgerond) 40.998 en als we de bijhorende eigenvector zo normeren dat de som der componenten hetzelfde is als de som der gewoon behaalde punten, namelijk 848, dan vinden we de volgende tabel (met de plaats in de gewone tabel tussen haakjes):

plaats	team	speelsterkte	punten
1. (2.)	Ajax Amsterdam	78.28	75
2. (1.)	PSV Eindhoven	76.90	75
3. (3.)	AZ Alkmaar	73.28	72
4. (4.)	FC Twente	67.91	66
5. (6.)	Roda JC Kerkrade	55.56	54
6. (7.)	Feyenoord Rotterdam	53.82	53
7. (5.)	sc Heerenveen	53.21	55
8. (8.)	FC Groningen	49.51	51
9. (9.)	FC Utrecht	47.37	48
10. (10.)	NEC Nijmegen	42.92	44
11. (12.)	Vitesse Arnhem	38.10	38
12. (11.)	NAC Breda	37.90	43
13. (13.)	Sparta Rotterdam	37.06	37
14. (14.)	Heracles Almelo	34.90	32
15. (16.)	Excelsior Rotterdam	30.57	30
16. (15.)	Willem II Tilburg	28.45	31
17. (17.)	RKC Waalwijk	25.25	27
18. (18.)	ADO Den Haag	17.03	17

Het valt op dat de m.b.v. speelsterkte berekende punten niet sterk van de daadwerkelijk behaalde punten afwijken, dit geeft aan dat de zwakkere ploegen inderdaad vooral tegen zwakke teams punten halen, terwijl de sterke teams ook van andere sterke teams winnen.

Historische opmerkingen

1. Het begrip *eigenwaarde* is niet afkomstig uit de matrixtheorie maar komt voor het eerst voor in werk van Jean Le Rond d'Alembert (1717-1783) tussen 1740 en 1750 waarin hij stelsels lineaire differentiaalvergelijkingen bestudeert (hij onderzocht de beweging van een snaar waaraan een aantal gewichten hingen). Om zo'n stelsel

$$\frac{d^2 y_i}{dt^2} + \sum_{k=1}^3 a_{ik} y_k = 0, \quad i = 1, 2, 3$$

te bestuderen vermenigvuldigde d'Alembert de i -de vergelijking met een (nog te bepalen) constante v_i , voor iedere i , en telde de vergelijkingen op. Dit geeft

$$\sum_{i=1}^3 v_i \frac{d^2 y_i}{dt^2} + \sum_{i,k=1}^3 v_i a_{ik} y_k = 0.$$

Als de v_i 's zo gekozen worden dat $\sum_{i=1}^3 v_i a_{ki} + \lambda v_k = 0$ voor zekere λ , voor $k = 1, 2, 3$ m.a.w. de vector (v_1, v_2, v_3) is een eigenvector behorende bij de eigenwaarde $-\lambda$ voor de matrix $A = (a_{ki})$, dan voert de substitutie $u = v_1 y_1 + v_2 y_2 + v_3 y_3$ het stelsel over in de volgende vergelijking

$$\frac{d^2 u}{dt^2} + \lambda u = 0.$$

Deze vergelijking kan eenvoudig worden opgelost voor ieder voor de λ 's die aan de derdegraads vergelijking $\det(\lambda I_3 - A) = 0$ voldoen. Bij ieder van de drie λ_i 's vinden we zo een u_i en met deze drie u_i 's kunnen dan y_1, y_2 en y_3 bepaald worden.

Eigenwaarden speelden ook een fundamentele rol in de quantummechanica rond 1925: de elektronen in een atoom kunnen niet alle mogelijke energietoestanden hebben, maar kunnen slechts bepaalde discrete waarden aannemen (de bekende "schillen" rond een proton). Deze discrete waarden bleken de eigenwaarden van een zekere matrix te zijn, de zgn. Hamiltoniaan (in de colleges *quantummechanica* wordt hierop uitvoerig ingegaan).

2. Het idee van Brin en Page om de documenten die aan een query voldoen met behulp van de relevanties te rangschikken die de componenten van een zekere eigenvector zijn, was al veel eerder door T.H. Wei (1952) en M.G. Kendall (1955) voorgesteld om rankings te bepalen. Brin en Page hebben dit idee echter voor het eerst op zoekmachines toegepast.