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 $x_1, \ldots, x_n$ te moeten opschrijven. Immers, het enige wat er bij die bewerkingen verandert zijn de coëfficiënten die voor de $x_1, \ldots, x_n$ 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.

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

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 $\mathrm{Mat_{m,n}}(\mathbb{R})$ of $M_{mn}(\mathbb{R})$. Als $m = n$ schrijven we $\mathrm{Mat_n}(\mathbb{R})$ 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 $c.a_{ij}$. 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 $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$, 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

$$x = x_1e_1 + x_2e_2 + \cdots + x_ne_n.$$

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

$$\begin{array}{c} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \vdots \qquad\qquad\qquad \vdots \\ a_{m1}x_n + \cdots + a_{mn}x_n = b_m \end{array} \tag{2.1}$$

Zij $A := (a_{ij})$ de matrix der coëffiënten 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 $A.x$ van de $m \times n$ matrix A en de $n \times 1$ matrix x door

$$A.x = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \qquad\qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} \tag{2.2}$$

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

$$Ae_j = \text{ de } j\text{-de kolom van } A. \tag{2.3}$$

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

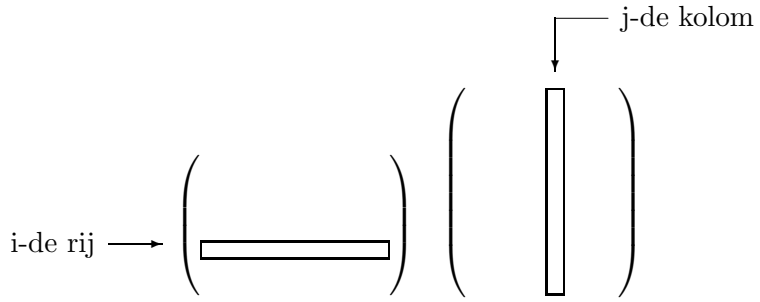
We zullen nu meer algemeen $A.B$ definiëren waarbij $A \in M_{m,n}(\mathbb{R})$ en $B \in M_{n,p}(\mathbb{R})$. Zij daartoe B_j de j-de kolom van B. Dan is volgens (2.2) AB_j een welgedefinieerde kolom uit $\mathbb{R}^n$.

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. $AB = (AB_1\ AB_2\ \ldots\ AB_p)$.

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 $a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$.

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 $x_1 = b_{1j}, \ldots, x_n = b_{nj}$. □

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.

$$(AB)_{11} = (1\ 2\ 3) \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = 1.2 + 2.(-1) + 3.3 = 9$$

$$(AB)_{21} = (4\ 5\ 6) \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = 4.2 + 5.(-1) + 6.3 = 21$$

Net zo vinden we $(AB)_{12} = 1.0 + 2.1 + 3.(-2) = -4$ en $(AB)_{22} = 4.0 + 5.1 + 6.(-2) = -7$. 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

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

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. □

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 *diagonaal matrix* als $a_{ij} = 0$ voor alle (i, j) met $i \neq j$.

Een handige notatie voor diagonale matrices is

$$A = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix} = \mathrm{diag}(\mathrm{a}_{11}, \mathrm{a}_{22}, \ldots, \mathrm{a}_{\mathrm{nn}}).$$

Speciale diagonale matrices zijn
i) de *eenheidsmatrix* $I_n = \mathrm{diag}(1, 1, \ldots, 1)$ die op de diagonaal alleen maar enen heeft;
ii) de *scalaire matrices* $c.I_n = \mathrm{diag}(\mathrm{c}, \mathrm{c}, \ldots, \mathrm{c})$. I.h.b. is $0.I_n$ de nulmatrix in $M_n(\mathbb{R})$.

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

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 (*indcutieaanname)*, 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 $U(2)$, $U(3)$, $\ldots$, $U(m-1)$, $U(m)$ 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 $U(n-1)$, $U(n-2)$, $\ldots$, $U(2)$, $U(1)$ 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 $U(1)$, $U(2)$, $\ldots$, $U(n)$ 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

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

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

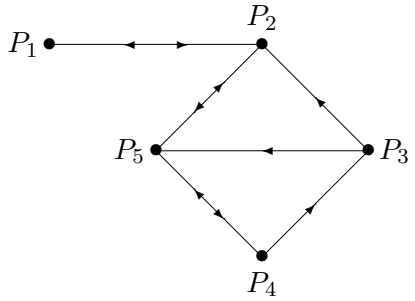
$$\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n.$$

2.3 Toepassingen

a) Grafentheorie en een voorbeeld uit de Sociologie

Stel je hebt een communicatie netwerk tussen 5 stations $P_1, \ldots, P_5$. 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 $P_1, \ldots, P_n$ 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 $P_3 \to P_2 \to P_1$, maar ook via het 3-pad $P_3 \to P_5 \to P_2 \to P_1$. 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 $P_1, \ldots, P_n$. 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 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 $P_1, \ldots, P_n$. 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 $a_{i1}a_{1j} + a_{i2}a_{2j} + \cdots + a_{in}a_{nj}$ 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 $(A^m)_{i1}a_{1j} + (A^m)_{i2}a_{2j} + \cdots + (A^m)_{in}a_{nj}$ 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. □

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 blijkbaar is de lengte van het kortste pad van P_1 naar P_4 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 $P_1 \to P_2 \to P_5 \to P_4 \to P_3$.

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 $A, A^2, \ldots$ 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_i gelijk is aan het element c_{ij} in de matrix $C = A + A^2 + \cdots A^m$.

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

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

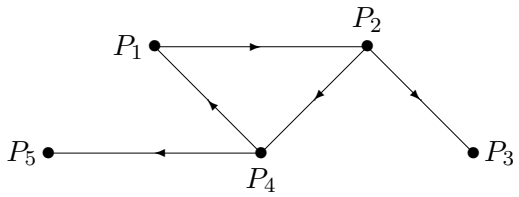
d.w.z. x symboliseert een oneindig grote afstand.

In bovenstaand voorbeeld hebben we dus berekent 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 $P_1, \ldots, P_5$. 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 formuleer moet je je naam invullen en de naam van díe 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 rijsommen:

- 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 dichter 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 $1, 2, 3, \ldots$ en de verschillende soorten aardewerk met $1, 2, 3, \ldots$ 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.

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

Iedere term is 0 of 1. Bijvoorbeeld de term $a_{i2}a_{j2} = 1$ d.e.s.d.a. beide $a_{i2} = 1$ èn $a_{j2} = 1$ 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. □

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} = aantal type aardewerk dat i en j gemeen hebben = aantal type aardewerk dat j en i 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

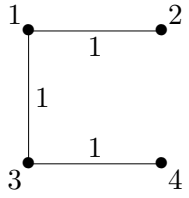
$$g_{12} = 1,\ g_{13} = 1,\ g_{14} = 0,\ g_{23} = 0,\ g_{24} = 0,\ g_{34} = 1.$$

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 $4 \to 3 \to 1 \to 2$ 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 eenvoudig 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. $BA\binom{3}{1} = \binom{3}{1}$. Echter $A\binom{3}{1} = \binom{0}{0}$ en dus $BA\binom{3}{1} = \binom{0}{0}$, 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 $B = C$. □

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 $Ax = b$. □

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 $a \in \mathbb{R}\backslash\{0\}$ 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 $M_3(\mathbb{R})$:

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

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.

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

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 $C = B^{-1}A^{-1}$. □

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. i) 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.

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

iii) We geven nu een bewijs van ($*$) met inductie naar n. Het geval $n = 1$ is duidelijk. Neem dus aan dat $n \geq 2$ en dat we ($*$) al bewezen hebben voor het geval $n - 1$. Merk op dat de eerste kolom van $A = (a_{ij})$ niet nul kan zijn (anders zou $Ae_1 = 0$, een tegenspraak met 2.4.5). Neem aan $a_{i1} \neq 0$. Door, indien $i \neq 1$, de eerste en de i-de rij te verwisselen en vervolgens in de ontstane matrix de eerste rij met $\frac{1}{a_{i1}}$ te vermenigvuldigen, krijgen we een matrix met 1 op de plaats $(1,1)$. Door bij de 2-de rij het $-a_{21}$-voud van de eerste op te tellen, krijgen we op de plaats $(2,1)$ een nul. Net zo kun je op alle plaatsen $\neq (1,1)$ uit de eerste kolom d.m.v. elementaire rij operaties een nul krijgen. Noem de ontstane matrix B.

iv) Bekijk nu de $(n-1) \times (n-1)$ rechtsonder matrix van B d.w.z. de matrix ontstaan uit B door de eerste rij en kolom van B te schrappen en noem die B_{n-1}. Dan is B_{n-1} een inverteerbare matrix: immers als B_{n-1} niet inverteerbaar is dan is ook B_{n-1}^t niet inverteerbaar (zie opgave 2.4.14 hieronder) en dus bestaat er volgens 2.4.6 een $x' \in \mathbb{R}^{n-1}$, $x' \neq 0$ met

$B_{n-1}^t x' = 0$. Maar dan voldoet $x := \binom{0}{x'} \in \mathbb{R}^n$ aan $B^t x = 0$ (ga na!) en dit is in tegenspraak met de inverteerbaarheid van B^t (opgave 2.4.14) vanwege 2.4.6. Dus is blijkbaar B_{n-1} inverteerbaar. Vanwege de inductie veronderstelling kunnen we B_{n-1} d.m.v. elementaire rij operaties herleiden tot I_{n-1}. Door deze rij operaties op B toe te passen vinden we een matrix, zeg C, waarvan de laatste $n-1$ rijen gelijk zijn aan die van I_n (herinner dat de eerste kolom van B e_1 is!). De eerste rij van C is van de vorm $(1, c_2, \ldots, c_n)$.
v) Tel nu $-c_2$ keer de tweede rij van C, welke $(0, 1, 0, \ldots, 0)$ is, op bij de eerste rij van C. We vinden dan nul op de plaats $(1, 2)$. Herhaal dit nu voor de plaatsen $(1, 3)$ t/m $(1, n)$. Het eindresultaat is de matrix I_n. M.a.w. d.m.v. elementaire rij operaties is A te herleiden tot I_n, waarmee $(*)$ en dus de stelling bewezen is. □

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 $E_1, \ldots E_s$ geldt dan

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

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 $E_s \ldots E_1 A = I_n$. Dan

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

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 $\binom{d}{-c}$ en $\binom{-b}{a}$ 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 $ad - bc \neq$ 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. □

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 *coderings matrix*. De ontvanger decodeert de boodschap m.b.v. de inverse van de matrix, welke de *decoderings 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

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

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

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	$*$

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 getallenrechthoek 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 vermelde 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 vermenigvuldiging (zie het college *Ringen 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.