

# Lineaire Algebra



# Inhoudsopgave

<b>1 Stelsels lineaire vergelijkingen</b>	<b>1</b>
1 Enkele inleidende voorbeelden . . . . .	1
2 Algemeenheden . . . . .	4
3 Gauss-Jordan eliminatie . . . . .	8
<b>2 Matrixbewerkingen</b>	<b>15</b>
1 Optellen, vermenigvuldigen en transponeren . . . . .	15
2 Informatie verwerken met matrices . . . . .	18
3 Elementaire matrices en inverteerbaarheid . . . . .	21
4 Error-correcting coding . . . . .	25
<b>3 Vectoren en deelruimten</b>	<b>31</b>
1 $\mathbb{R}^n$ als verzameling van vectoren . . . . .	31
2 Deelruimten . . . . .	36
3 Orthogonaliteit . . . . .	40
4 Lights out . . . . .	42
5 De kleinste-kwadratenmethode . . . . .	49
6 De Gram-Schmidt-procedure . . . . .	51
<b>4 Determinanten</b>	<b>57</b>
1 Definitie en eigenschappen . . . . .	57
2 Toepassingen . . . . .	64

<b>5 Eigenwaarden en eigenvectoren</b>	<b>67</b>
1 Voorbeeld en definities . . . . .	67
2 Eigenschappen . . . . .	70
3 Enkele toepassingen . . . . .	78
4 De spectraalstelling . . . . .	80
5 De Jordan-normaalvorm . . . . .	85
6 Markovmatrices en het PageRank-algoritme . . . . .	87
<b>6 Vectorruimten en lineaire afbeeldingen</b>	<b>95</b>

**Dankwoord.** Deze cursus is voor het overgrote deel op nota's van Prof. Lieven Le Bruyn gebaseerd. Ik ben hem zeer erkentelijk voor het vrij beschikbaar stellen van zijn cursusmateriaal. Veel van de afbeeldingen en foto's werden, zonder expliciete toestemming van de auteurs, van het internet geplukt. Mijn excuses aan deze anonieme bijdragers mochten ze zich hiervoor tekort gedaan voelen.



# 1 Stelsels lineaire vergelijkingen

## 1 Enkele inleidende voorbeelden

De kerntaak van lineaire algebra is het oplossen van stelsels van lineaire vergelijkingen. We beginnen met enkele eenvoudige voorbeelden:

**Raadsels.** (1) *Een klas van 15 kleuters gaat op uitstap. De kleuters zullen ofwel op een fiets, ofwel op een driewieler rijden. De juf pompt eerst alle banden op en ze telt 37 wielen. Hoeveel fietsen zijn er?*

Laat  $x$  het aantal fietsen zijn, en  $y$  het aantal driewielers. De informatie kan dan vertaald worden in het volgend stelsel vergelijkingen

$$\begin{cases} x + y = 15, \\ 2x + 3y = 37. \end{cases}$$

De onbekende  $y$  halen uit de eerste en substitueren in de tweede levert als oplossing  $x = 8$  en  $y = 7$ .

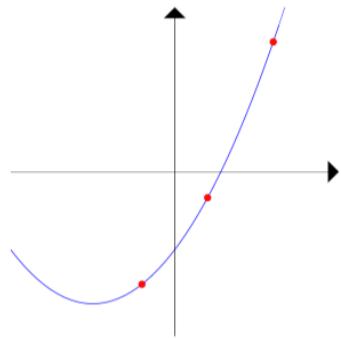
(2) *Bob, de broer van Alice, heeft twee keer zoveel zussen als broers. Alice daarentegen heeft evenveel broers als zussen. Hoeveel kinderen telt deze familie?*

Met  $x$  het aantal jongens en  $y$  het aantal meisjes in de familie, vinden we het stelsel

$$\begin{cases} y = 2(x - 1), \\ x = y - 1. \end{cases}$$

Je vindt makkelijk dat  $x = 3$  en  $y = 4$ .

**Interpolatie.** *Vind de vergelijking van de parabool  $y = ax^2 + bx + c$  die door volgende punten gaat:  $P = (-1, -3)$ ,  $Q = (1, -1)$  en  $R = (3, 4)$ .*



De drie punten substitueren in de vergelijking van de parabool levert het volgende stelsel lineaire vergelijkingen in de onbekenden  $a$ ,  $b$  en  $c$ :

$$\begin{cases} -3 = a - b + c, \\ -1 = a + b + c, \\ 4 = 9a + 3b + c. \end{cases}$$

Uit de eerste vergelijking halen we  $c = b - a - 3$ . Wanneer we dit in de volgende twee vergelijkingen substitueren, wordt dit stelsel

$$\begin{cases} 2 = 2b, \\ 7 = 8a + 4b. \end{cases}$$

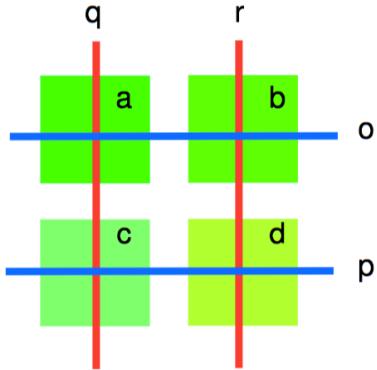
Hieruit volgt dat  $b = 1$ ,  $a = \frac{3}{8}$  en dus  $c = -\frac{19}{8}$ .

**Tomografie.** MRI-scans maken gebruik van absorptie en emissie van energie in de radiofrequenties van het electromagnetische spectrum.



*Stel dat we vier waterstof atomen hebben waarvan de kernen een exitatie hebben met intensiteiten  $a, b, c, d$  (de onbekenden). We kunnen de echointensiteit meten in 4 verschillende*

richtingen  $o, p, q, r$



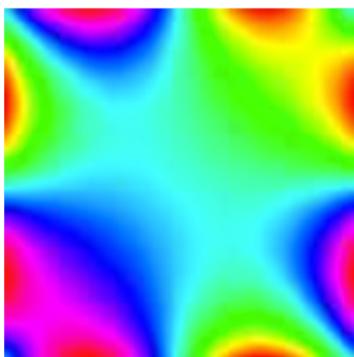
en stel dat we als waarden  $3, 7, 5, 5$  krijgen. Wat zijn dan  $a, b, c, d$ ?

Onze informatie leidt tot het stelsel lineaire vergelijkingen

$$\begin{cases} a + b = 3, \\ c + d = 7, \\ a + c = 5, \\ b + d = 5. \end{cases}$$

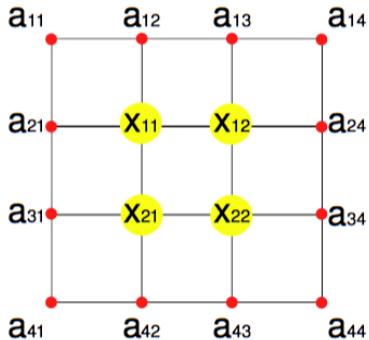
Ga na dat bvb.  $a = 2, b = 1, c = 3$  en  $d = 4$  een oplossing is, maar ook  $a = 0, b = 3, c = 5$  en  $d = 2$ . Dus, we zien dat een stelsel meerdere oplossingen kan hebben. Wat is de algemene oplossing van dit stelsel? (Antwoord: we kunnen  $b, c$  en  $d$  in functie van  $a$  schrijven:  $b = 3 - a, c = 5 - a$  en  $d = 7 - c = 2 + a.$ )

**Evenwichttoestanden.** Sommige problemen leiden tot stelsels met heel veel variabelen.



Neem bvb. een model van een opgespannen vel voorgesteld door een fijnmazig net. De hoogte van elk inwendig roosterpunt (de onbekenden) is het gemiddelde van de hoogten van de 4 directe buren. Gekend zijn de hoogten van de randpunten. Als we  $n^2$  inwendige punten hebben moeten we een stelsel van  $n^2$  lineaire vergelijkingen in  $n^2$  onbekenden oplossen.

Neem  $n = 2$  voorgesteld door de 4 inwendige punten  $x_{ij}$  en de 12 randpunten  $a_{ij}$ .



Wat is het bijhorende stelsel lineaire vergelijkingen?

De vergelijkingen worden:

$$\begin{cases} 4x_{11} = a_{21} + a_{12} + x_{12} + x_{21}, \\ 4x_{12} = x_{11} + a_{13} + a_{24} + x_{22}, \\ 4x_{21} = a_{31} + x_{11} + x_{22} + a_{42}, \\ 4x_{22} = x_{21} + x_{12} + a_{34} + a_{43}. \end{cases}$$

Dit stelsel is al moeilijker op te lossen naar de onbekenden  $x_{ij}$  (Probeer!). In een realistische situatie is  $n = 300$ , bijvoorbeeld, en krijgen we een stelsel met 90.000 onbekenden. Het is evident dat men in dergelijke gevallen behoeft aan een algemeen algoritme.

## 2 Algemeenheden

**Lineaire vergelijkingen.** In elk van de voorgaande voorbeelden kon het probleem geformuleerd worden aan de hand van vergelijkingen (en niet, bijvoorbeeld, ongelijkheden) in een aantal onbekenden  $x_1, \dots, x_n$ . Merk ook op dat in geen enkel van de bovenstaande voorbeelden gebruik werd gemaakt van bv. vierkantsvergelijkingen, of vergelijkingen waarin op de onbekenden goniometrische of andere elementaire functies toegepast werden. Het lineair zijn van de vergelijkingen is essentieel voor de methoden die we in deze cursus wensen te bespreken. Het bepaalt tevens het soort toepassingen die we met die methoden kunnen behandelen.

Eén *lineaire vergelijking* in  $n$  onbekenden  $x_1, \dots, x_n$  is van de vorm

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b,$$

waarbij de reële getallen  $a_1, \dots, a_n \in \mathbb{R}$  de *coëfficiënten* van de vergelijking heten en  $b \in \mathbb{R}$  de *constante term* van de vergelijking is. Een punt  $(p_1, p_2, \dots, p_n) \in \mathbb{R}^n$  noemen we een *oplossing* indien

$$a_1p_1 + a_2p_2 + \dots + a_np_n = b.$$

Een *stelsel van  $m$  lineaire vergelijkingen in  $n$  onbekenden* wordt gegeven door vergelijkingen van de vorm

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases}$$

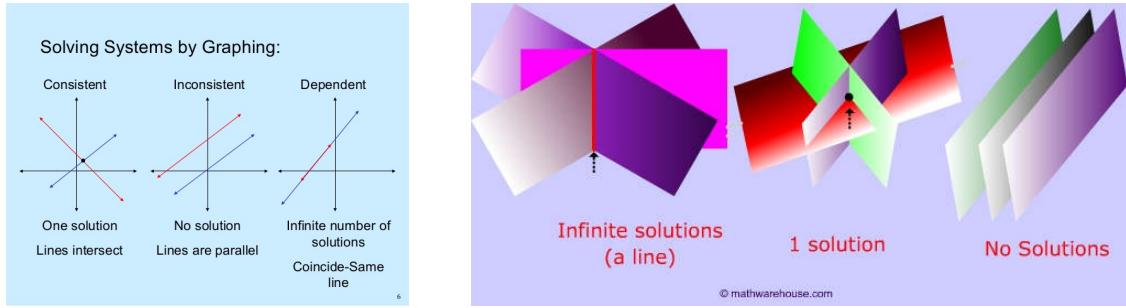
Een *oplossing* van het stelsel is een punt  $p = (p_1, p_2, \dots, p_n) \in \mathbb{R}^n$  dat een oplossing is van elk van de  $m$  lineaire vergelijkingen in het stelsel, dat is, we hebben dat

$$\begin{cases} a_{11}p_1 + a_{12}p_2 + \dots + a_{1n}p_n = b_1, \\ a_{21}p_1 + a_{22}p_2 + \dots + a_{2n}p_n = b_2, \\ \vdots \\ a_{m1}p_1 + a_{m2}p_2 + \dots + a_{mn}p_n = b_m. \end{cases}$$

We kunnen alle oplossingen in een verzameling groeperen. Als deze verzameling leeg is spreken we van een *strijdig* stelsel.

**Meetkundige interpretatie.** Beschouw het geval van een lineaire vergelijking  $ax + by = c$  in twee onbekenden  $(x, y)$ . De oplossingen van deze vergelijking vormen samen een *rechte* in het  $(x, y)$ -vlak  $\mathbb{R}^2$ . Analoog stellen de oplossingen van de vergelijking  $ax + by + cz = d$  een *vlak* in  $\mathbb{R}^3$  voor, met coördinaten  $x, y$  en  $z$ . De verzameling van alle oplossingen van een lineaire vergelijking in  $n$  onbekenden noemt men een *hypervlak* in  $\mathbb{R}^n$ .

Deze meetkundige interpretatie kan gebruikt worden om een idee te krijgen van het aantal oplossingen dat een stelsel van vergelijkingen kan hebben. Bijvoorbeeld, bij een stelsel in twee onbekenden  $(x, y)$  stelt elke vergelijking van het stelsel een rechte in  $\mathbb{R}^2$  voor. Een oplossing van het stelsel is dus een punt dat terzelfdertijd op alle rechten ligt. In de onderstaande figuur werden enkele mogelijke gevallen afgebeeld. Analoog liggen oplossingen van stelsels in 3 onbekenden op de snijpunten en snijlijnen van vlakken in  $\mathbb{R}^3$ . De oplossingsruimte van stelsels van  $m$  vergelijkingen in  $n$  onbekenden is de doorsnede in  $\mathbb{R}^n$  van  $m$  hypervlakken.



**Voorbeeld.** Bekijk het volgende systeem van lineaire vergelijkingen in de onbekenden  $x, y$  en  $z$  (m.a.w. we zoeken oplossingen in  $\mathbb{R}^3$ )

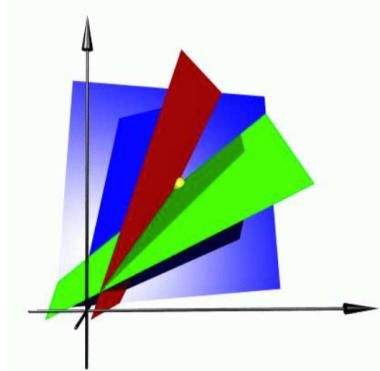
$$\begin{cases} 3x - y - z = 0, \\ -x + 2y - z = 0, \\ -x - y + 3z = 9. \end{cases}$$

Uit de eerste vergelijking halen we dat  $z = 3x - y$ . Als we dat substitueren in de twee volgende vergelijkingen krijgen we het stelsel

$$\begin{cases} -4x + 3y = 0, \\ 8x - 4y = 9. \end{cases}$$

Hiervan geeft de eerste vergelijking  $y = \frac{4}{3}x$  en dit in de tweede substitueren levert  $8x = 27$ . Dus we krijgen  $x = \frac{27}{8}$  en hieruit volgt dat  $y = \frac{9}{2}$  en  $z = \frac{45}{8}$ . We zien dus dat er een unieke oplossing is van het stelsel.

Meetkundig kunnen we de vlakken in  $\mathbb{R}^3$  bekijken die horen bij elk van de lineaire vergelijkingen.



En we merken op dat inderdaad deze drie vlakken in slechts 1 punt snijden, namelijk in  $p = (\frac{27}{8}, \frac{9}{2}, \frac{45}{8})$ .

**Voorstelling aan de hand van matrices.** Stelsels van lineaire vergelijkingen zullen we vaak voorstellen aan de hand van hun bijhorende matrices. Een *matrix* is een rechthoekig schema van symbolen (de *componenten* van de matrix). Voor het overgrote deel van deze cursus zullen deze symbolen reële getallen zijn. Een  $m \times n$  *matrix* bestaat uit  $m$  *rijen* en  $n$  *kolommen*,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Het koppel  $(m, n)$  noemt men de *orde* van de matrix. Wanneer  $m = n$  dan spreken we over een *vierkante matrix*. Een matrix met  $m = 1$  noemt men een *rijmatrix*. Een matrix met  $n = 1$  is een *kolommatrix*. In de rest van de cursus wordt voor dit geval ook vaak de benaming  *$m$ -dimensionale vector* gebruikt. Een matrix met beide  $m = n = 1$  kunnen we identificeren met haar enige component, dus met een reëel getal.

Met het bovenstaande stelsel lineaire vergelijkingen kunnen we nu drie matrices associëren. De eerste matrix,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix},$$

noemen we de *coëfficiëntenmatrix* van het stelsel. De constante termen vormen samen de vector

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

De *uitgebreide matrix* van het stelsel is de  $m \times (n + 1)$  matrix  $B = (A, \vec{b})$  bestaande uit  $A$  en  $\vec{b}$ ,

$$B = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix}.$$

**Belangrijke opmerking.** Zoals eerder vermeld zullen we, voor de eenvoud, bijna altijd werken met vergelijkingen met coefficiënten in  $\mathbb{R}$ , en met matrices met componenten in  $\mathbb{R}$ . De resultaten die we bewijzen zullen echter meestal ook gelden voor coefficiënten in een willekeurig *lichaam* (waarin we kunnen optellen en vermenigvuldigen en waarin elk niet nul element een inverse heeft). Informatici werken vaak over het lichaam  $\mathbb{F}_2 = \{0, 1\}$  van de 2-bits en we zullen hiervan enkele toepassingen geven. In de enkele gevallen waarin we echt moeten werken over  $\mathbb{R}$  (bvb. bij inproducten) zullen we dit explicet vermelden.

### 3 Gauss-Jordan eliminatie

**Equivalent stelsels.** In deze paragraaf zullen we een algemeen algoritme geven om de oplossing(en) te vinden van een lineair stelsel van  $m$  vergelijkingen in  $n$  onbekenden

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases}$$

Twee verschillende stelsels van lineaire vergelijkingen noemen we *equivalent* wanneer ze dezelfde oplossingsruimte hebben. Merk op dat de oplossingen van een stelsel niet zal veranderen onder (een combinatie van) volgende operaties:

- Het verwisselen van twee vergelijkingen in de lijst.
- Het vermenigvuldigen van een vergelijking met een niet nul getal.
- Het aftrekken van een niet-nul veelvoud van een vergelijking van een andere.

Enkel de laatste operatie vraagt misschien wat uitleg. Als  $p = (p_1, \dots, p_n)$  een oplossing is van het stelsel, dan is  $p$  ook een oplossing van het stelsel waarin we de  $i$ -de vergelijking vervangen hebben door

$$(a_{i1}x_1 + \dots + a_{in}x_n) - c(a_{j1}x_1 + \dots + a_{jn}x_n) = b_i - cb_j.$$

Omgekeerd als  $p$  een oplossing is van dit nieuwe stelsel, dus met  $i$ -de vergelijking die hierboven en alle andere vergelijkingen (waaronder de  $j$ -de) van het originele stelsel, dan is ook  $p$  een oplossing van

$$((a_{i1}x_1 + \dots + a_{in}x_n) - c.(a_{j1}x_1 + \dots + a_{jn}x_n)) + c.(a_{j1}x_1 + \dots + a_{jn}x_n) = b_i - c.b_j + c.b_j,$$

maar dit is gewoon de  $i$ -de oorspronkelijke vergelijking.

Naar elk van de operaties uit de lijst zullen we refereren als *elementaire rijoperaties*.

**Rij-echelonvorm.** Sommige stelsels zijn in een zodanige vorm dat men bijna onmiddellijk de oplossing kan aflezen. Beschouw, bijvoorbeeld, het stelsel van 3 vergelijkingen in 4 onbekenden  $x, y, z, v$  gegeven door

$$\begin{cases} x + 2z = 3, \\ y + 3z = 0, \\ v = 2. \end{cases}$$

Via achterwaartse substitutie vinden we dat de oplossingsverzameling gegeven wordt door  $\{(3 - 2z, -3z, z, 2) | z \in \mathbb{R}\}$ . Het 'eenvoudig zijn' van dit stelsel komt omdat haar bijhorende uitgebreide matrix

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 3 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

reeds in rij-echelonvorm staat.

**Definitie 1.** Een matrix staat in rij-echelonvorm als er voldaan is aan volgende eisen.

- Als een rij niet-nul getallen bevat dan is de eerste niet-nul component een 1 (die we de *leidende 1 van de rij* noemen).
- Als een kolom een leidende 1 bevat van een rij, dan zijn alle andere componenten in die kolom gelijk aan 0.
- Als een rij een leidende 1 heeft, dan hebben alle voorgaande rijen ook een leidende 1 die meer naar links staat.

Hieronder vind je een voorbeeld van een rij-echelon matrix, waarin een \* een willekeurig getal voorstelt:

$$\begin{bmatrix} 0 & 1 & * & 0 & 0 & * & * \\ 0 & 0 & 0 & 1 & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hierin hebben de rijen 1, 2 en 3 een leidende 1 en ook de kolommen 2, 4 en 5. Het is duidelijk dat, als bovenstaande matrix de uitgebreide matrix is van een stelsel vergelijkingen, we makkelijk de oplossingen van het stelsel kunnen bepalen via achterwaartse substitutie.

Een  $(m, n + 1)$  matrix in rij-echelonvorm kan hoogstens  $\min(m, n + 1)$  aantal leidende 1-en bevatten.

**Gauss-Jordan algoritme.** Beschouw opnieuw een algemeen stelsel van  $m$  vergelijkingen in  $n$  onbekenden, met uitgebreide matrix

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix}.$$

Het doel is om dit stelsel om te zetten in een 'eenvoudig' equivalent stelsel door enkel gebruik te maken van de bovenvermelde toegestane elementaire rijoperaties. Het doel wordt bereikt wanneer de uitgebreide matrix van het equivalente stelsel in rij-echelonvorm staat.

**Stap 1:** Als alle componenten in de matrix nul zijn, dan zijn we klaar.

**Stap 2:** Zoek de eerste kolom  $j$  die niet-nul getallen bevat en hierin de eerste rij  $i$  zodat  $a_{ij} \neq 0$ . Vermenigvuldig rij  $i$  met  $a_{ij}^{-1}$  zodat ze een leidende 1 krijgt.

**Stap 3:** Verwissel rij  $i$  met de eerste rij in de matrix, zodat de bovenste rij een leidende 1 heeft.

**Stap 4:** Maak alle andere componenten in de  $j$ -de kolom nul door een veelvoud van de bovenste rij af te trekken van de andere rijen. Hierdoor verkrijgen we een matrix van de vorm

$$\begin{bmatrix} 0 & \dots & 0 & 1 & c_{1,j+1} & \dots & c_{1,n+1} \\ 0 & \dots & 0 & 0 & c_{2,j+1} & \dots & c_{2,n+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & c_{m,j+1} & \dots & c_{m,n+1} \end{bmatrix}$$

**Stap 5:** Herhaal voorgaande stappen op de deelmatrix

$$\begin{bmatrix} c_{2,j+1} & \dots & c_{2,n+1} \\ \vdots & & \vdots \\ c_{m,j+1} & \dots & c_{m,n+1} \end{bmatrix}$$

waarna we een matrix krijgen van de vorm

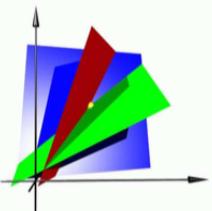
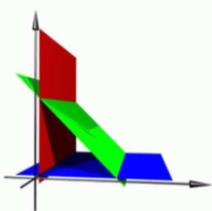
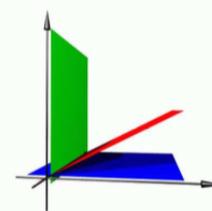
$$\begin{bmatrix} \dots & 0 & 1 & * & \dots & * & d_{1,k} & d_{1,k+1} & \dots & d_{1,n+1} \\ \dots & 0 & 0 & 0 & \dots & 0 & 1 & d_{2,k+2} & \dots & d_{2,n+1} \\ \dots & 0 & 0 & 0 & \dots & 0 & 0 & d_{3,k+2} & \dots & d_{3,n+1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ \dots & 0 & 0 & 0 & \dots & 0 & 0 & d_{m,k+1} & \dots & d_{m,n+1} \end{bmatrix}$$

Hierna kunnen we ook de  $(1, k)$ -component nul krijgen door van de eerste rij  $d_{1k}$ -keer de tweede af te trekken. Merk op dat dit geen invloed heeft op de leidende 1 in de eerste rij en de  $j$ -de kolom.

**Stap 6:** Herhaal het voorgaande tot je overblijft met een nul-deelmatrix of een leidende 1 hebt op positie  $(m, n + 1)$ .

De matrix in rij-echelonvorm verkregen uit een matrix  $X$  noteren we met  $\text{ref}(X)$ .

Hieronder vind je enkele uitwerkingen voor een specifiek stelsel vergelijkingen:

THE GOOD (1 solution)	THE BAD (0 solution)	THE UGLY ( $\infty$ solutions)
		
$\left[ \begin{array}{ccc c} 0 & 1 & 2 & 2 \\ 1 & -1 & 1 & 5 \\ 2 & 1 & -1 & -2 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 2 & 1 & -1 & -2 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 3 & -3 & -12 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & -1 & -4 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & -3 & -6 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 2 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{array} \right]$	$\left[ \begin{array}{ccc c} 0 & 1 & 2 & 2 \\ 1 & -1 & 1 & 5 \\ 1 & 0 & 3 & -2 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & 2 & -7 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & -9 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & -9 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 1 \end{array} \right]$	$\left[ \begin{array}{ccc c} 0 & 1 & 2 & 2 \\ 1 & -1 & 1 & 5 \\ 1 & 0 & 3 & 7 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 1 & 0 & 3 & 7 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & 2 & 2 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & -1 & 1 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & 0 & 3 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right]$ $\left[ \begin{array}{ccc c} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right]$
Rank(A) = 3, Rank(B) = 3.	Rank(A) = 2, Rank(B) = 3.	Rank(A) = 2, Rank(B) = 2. Rank(A) = 2, Rank(B) = 2.

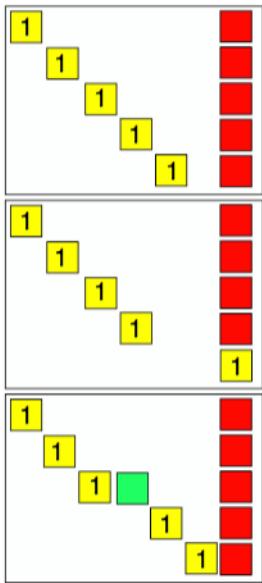
Uit het voorbeeld kunnen we de volgende stelling afleiden, steunend op het idee van achterwaartse substitutie.

**Stelling 1** (Oplossingen van lineaire stelsels). *Een stelsel van  $m$  lineaire vergelijkingen in  $n$  onbekenden met coefficiënt-matrix  $A$  en uitgebreide matrix  $B = (A, \vec{b})$  heeft*

- exact één oplossing indien de eerste  $n$  kolommen van  $\text{ref}(B)$  allen een leidende 1 hebben.
- geen oplossingen indien er een leidende 1 is in de laatste kolom van  $\text{ref}(B)$ .

- oneindig veel oplossingen *indien er kolommen zijn zonder leidende 1 en er ook geen leidende 1 is in de laatste kolom.*

Onderstaande figuur geeft een grafische voorstelling van de gevallen van de stelling.



**Definitie 2.** *Het aantal leidende 1-en in  $\text{ref}(X)$  noemen we de rang van  $X$ .*

We noteren de rang van  $X$  met  $\text{rank}(X)$ . Als  $X$  een  $m \times n$ -matrix is, dan is  $\text{rank}(X) \leq \min(m, n)$ . De rang van een matrix  $X$  kan men ook opvatten als het aantal niet-nulrijen van de matrix  $\text{ref}(X)$ .

De rang van de coefficientenmatrix  $A$  en van de uitgebreide matrix  $B = (A, \vec{b})$  bepalen dus of en hoeveel oplossingen het stelsel heeft. Merk eerst op dat er steeds geldt

$$\text{rank}(A) \leq \text{rank}(B) \leq \min(m, n + 1)$$

( $A$  is een deelmatrix van  $B$ , en dus ook  $\text{ref}(A)$  van  $\text{ref}(B)$ , en kan dus hoogstens evenveel nulrijen bevatten). Tevens geldt er dat de rang van de uitgebreide matrices van twee verschillende maar equivalente stelsels steeds dezelfde zal zijn (ze hebben immers dezelfde rij-echelonvorm).

De voorgaande stelling kan nu uitgedrukt worden in termen van de rangen van de matrices  $A$  en  $B$ .

**Stelling 2** (Oplossingen van lineaire stelsels). *Het stelsel met uitgebreide matrix  $B = (A, \vec{b})$  heeft*

- exact één oplossing *indien  $\text{rank}(A) = \text{rank}(B) = n$ .*

- geen oplossingen *indien*  $\text{rank}(B) = \text{rank}(A) + 1$ .
- oneindig veel oplossingen *indien*  $\text{rank}(A) = \text{rank}(B) < n$ .

**Voorbeeld.** Bepaal de oplossingen van het stelsel

$$\begin{cases} x + y + 3z = -1, \\ 2x - 2y - z = 1, \\ 5x + y + 8z = -2. \end{cases}$$

We gaan eerst na of het stelsel oplosbaar, dan wel strijdig is. Hiervoor bepalen we de rang van  $A$  en  $B$ . Na enkele operaties is de rij-echelonvorm van de uitgebreide matrix

$$\begin{bmatrix} 1 & 0 & \frac{5}{4} & -\frac{1}{4} \\ 0 & 1 & \frac{7}{4} & -\frac{3}{4} \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

waaruit we duidelijk kunnen besluiten dat  $\text{rank}(A) = \text{rank}(B) = 2$ . Het stelsel is dus oplosbaar. Aangezien de rang kleiner is dan het aantal onbekenden zullen er oneindig veel oplossingen zijn. Het equivalente stelsel dat hoort bij  $\text{ref}(B)$  is

$$\begin{cases} x + \frac{5}{4}z = -\frac{1}{4}, \\ y + \frac{7}{4}z = \frac{3}{4}. \end{cases}$$

We kunnen  $z$  dus als een vrije parameter opvatten. De oplossingenverzameling van het stelsel wordt dus gegeven door

$$\left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\frac{1}{4} \\ -\frac{3}{4} \\ 0 \end{bmatrix} + z \begin{bmatrix} -\frac{5}{4} \\ -\frac{7}{4} \\ 1 \end{bmatrix} \mid z \in \mathbb{R} \right\}.$$

**Historische noot.** Johann Carl Friedrich Gauss (1777-1855) (links) ontwikkelde de eliminatiemethode rond 1800 en paste ze toe op sterrenkundige problemen. Op 1 januari 1801 had de Italiaanse astronoom Giuseppe Piazzi (midden) de asteroide Ceres ontdekt (een rotsblok met een diameter van 914km). Gauss was in staat om de baan van Ceres te bepalen uit maar enkele observaties (met een methode analoog aan dat van het interpolatieprobleem uit de eerste paragraaf). Wilhelm Jordan (1842-1899) (rechts) was een Duitse toegepaste

wiskundige die toepassingen gaf van de eliminatiemethode op de [geodesie](#).



# 2 Matrixbewerkingen

## 1 Optellen, vermenigvuldigen en transponeren

We vermelden eerst kort enkele bijzondere matrices.

- Matrices waarin alle componenten nul zijn noemen we *nulmatrices*.
- De *eenheidsmatrix*  $I_n$  van orde  $n$  is de vierkante  $n \times n$ -matrix met enkel 1-en op de diagonalelementen, en 0-en op de niet-diagonalelementen. We kunnen dit kort noteren met  $I_n = (\delta_{ij})$ . Het symbool  $\delta_{ij}$  hierin staat voor de Kronecker-delta. Het is gedefinieerd als  $\delta_{ij} = 1$  als  $i = j$  en 0 als  $i \neq j$ . Wanneer de orde van  $I_n$  uit de context gekend is, noteren we vaak gewoon  $I$ .
- Bij een *diagonaalmatrix*  $D = (a_{ij})$  is  $a_{ij} = 0$  als  $i \neq j$ .

Twee matrices  $A = (a_{ij})$  en  $B = (b_{ij})$  zijn *gelijk* als ze evenveel rijen en evenveel kolommen hebben en als  $a_{ij} = b_{ij}$  voor alle  $i$  en  $j$ . De optelling is enkel gedefinieerd voor twee matrices  $A = (a_{ij})$  en  $B = (b_{ij})$  met evenveel rijen  $m$  en met evenveel kolommen  $n$ . De *som*  $A + B$  is dan de  $m \times n$ -matrix waarvan de component op plaats  $(i, j)$  de som is van de componenten van  $A$  en  $B$  op die plaats, dat is  $(A + B)_{ij} = a_{ij} + b_{ij}$ , of

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}.$$

Als  $A = (a_{ij})$  een  $m$  bij  $n$  matrix is en  $\lambda \in \mathbb{R}$  een getal, dan is het *scalair veelvoud*  $\lambda A$  de  $m$  bij  $n$  matrix waarvan elke component met  $\lambda$  vermenigvuldigd wordt, dat is  $(\lambda A)_{ij} = \lambda a_{ij}$ ,

$$\lambda \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{bmatrix}.$$

In sommige gevallen kunnen twee matrices met elkaar vermenigvuldigd worden. Als  $A$  een  $m$  bij  $n$  matrix is en  $B$  een  $n$  bij  $p$  matrix, dus als het aantal kolommen van  $A$  gelijk is aan het aantal rijen van  $B$ , dan is het product  $A \cdot B$  van  $A$  en  $B$  de  $m$  bij  $p$  matrix (dus zoveel rijen als  $A$  en zoveel kolommen als  $B$ ) die we krijgen door middel van de *rij-kolom regel*:

$$(A \cdot B)_{i,j} = \sum_{k=1}^n a_{ik} b_{kj}.$$

De component van  $A \cdot B$  op plaats  $i, j$  krijgen we dus door de componenten van de  $i$ -de rij van  $A$  paarsgewijs te vermenigvuldigen met de entries van de  $j$ -de kolom van  $B$  (merk op dat er evenveel componenten in beide zijn omdat het aantal kolommen van  $A$  gelijk is aan het aantal rijen van  $B$ ).

### Voorbeeld.

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} 38 & 44 & 40 & 56 \\ 83 & 98 & 113 & 128 \end{bmatrix}.$$

Merk op dat we matrix-vermenigvuldiging kunnen gebruiken om een lineair stelsel in matrix-vorm te schrijven. Als we een stelsel van  $m$  lineaire vergelijkingen in  $n$  onbekenden hebben met coefficiënt matrix  $A$  en constante vector  $\vec{b}$  en  $n$ -dimensionale vector

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

dan kunnen we het stelsel in matrix-vorm schrijven als

$$A \cdot \vec{x} = \vec{b}$$

Ga zelf volgende beweringen na:

**Stelling 3.** *Voor alle matrices  $A, B$  en  $C$  waarvoor de bewerkingen goed gedefinieerd zijn geldt*

1.  $A \cdot I_n = A$  en  $I_m \cdot A = A$ ,
2.  $A \cdot (B + C) = A \cdot B + A \cdot C$  en  $(A + B) \cdot C = A \cdot C + B \cdot C$ ,
3.  $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ ,
4.  $(\lambda A) \cdot B = \lambda (A \cdot B)$ ,  $\forall \lambda \in \mathbb{R}$ ,

5. In het algemeen geldt echter  $A \cdot B \neq B \cdot A$ .

**Complexiteit van een matrixproduct.** De complexiteit van een berekening hangt af van het aantal bewerkingen dat nodig is. Bijvoorbeeld, een matrixvermenigvuldiging van twee  $2 \times 2$ -matrices bevat essentieel 8 vermenigvuldigingen en 4 optellingen/aftrekkingen:

$$\begin{bmatrix} a & b \\ -c & -d \end{bmatrix} \cdot \begin{bmatrix} A & B \\ -C & -D \end{bmatrix} = \begin{bmatrix} aA - bC & aB - bD \\ -cA - dC & -cB + db \end{bmatrix}.$$

Algemeen heeft men  $n^2(n-1)$  optellingen en  $n^3$  vermenigvuldigingen nodig om twee vierkante matrices van orde  $n$  met elkaar te vermenigvuldigen.

In het algemeen gaat, voor een computer, optellen echter veel sneller dan vermenigvuldigen. De *Strassenmethode* reduceert het aantal vermenigvuldigingen in een matrixproduct, ten koste van extra optellingen. Bijvoorbeeld, het product hierboven kunnen we ook schrijven als

$$\begin{bmatrix} a & b \\ -c & -d \end{bmatrix} \cdot \begin{bmatrix} A & B \\ -C & -D \end{bmatrix} = \begin{bmatrix} R & (a-b)D - a(D-B) \\ (d-c)A - d(A-C) & S \end{bmatrix}$$

met  $R = (a+d)(A+D) - (b+d)(C+D) - d(A-C) - (a-b)D$  en  $S = (a+d)(A+D) - (a+c)(A+B) - a(D-B) - (d-c)A$ . Het verschil met de voorgaande uitdrukking is dat er nu slechts 7 (verschillende) vermenigvuldigingen moeten uitgevoerd worden, maar ten koste van 18 optellingen. De Strassenmethode kan uitgebreid worden naar matrices van hogere orde. Voor matrices van zeer grote orde is de methode efficiënter dan de gewone matrixvermenigvuldiging.

**Blokvermenigvuldiging.** Matrices kunnen we vaak in blokken verdelen, met in elk blok een deelmatrix van mogelijk verschillende orde. Twee matrices kunnen ook blok per blok vermenigvuldigd worden als de verdelingen overeenkomen. Bijvoorbeeld de matrices  $A$  en  $B$  werden hieronder in blokken opgesplitst:

$$A = \begin{bmatrix} 1 & 0 & | & 0 & 0 & 0 \\ 0 & 1 & | & 0 & 0 & 0 \\ - & - & | & - & - & - \\ 2 & -1 & | & 4 & 2 & 1 \\ 3 & 1 & | & -1 & 7 & 5 \end{bmatrix} = \begin{bmatrix} I_2 & 0 \\ P & Q \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -2 \\ 5 & 6 \\ - & - \\ 7 & 3 \\ -1 & 0 \\ 1 & 6 \end{bmatrix} = \begin{bmatrix} X \\ Y \end{bmatrix}.$$

Er geldt dan

$$A \cdot B = \begin{bmatrix} I_2 & 0 \\ P & Q \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} X \\ P \cdot X + Q \cdot Y \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ 5 & 6 \\ - & - \\ 30 & 8 \\ 8 & 27 \end{bmatrix}.$$

De laatste bewerking in deze paragraaf is de transponering. Als  $A = (a_{ij})$  een  $m$  bij  $n$  matrix is dan noemen we de *getransponeerde matrix*  $A^{tr}$  de  $n$  bij  $m$  matrix waarvan de kolommen van  $A^{tr}$  de rijen zijn van  $A$  en de rijen van  $A^{tr}$  de kolommen van  $A$ , dat is  $A^{tr} = (b_{ij})$  waarbij voor alle  $i$  en  $j$  geldt dat  $b_{ij} = a_{ji}$ , i.e.

$$\text{als } A = \begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \vdots & \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & \dots & a_{mn} \end{bmatrix} \quad \text{dan is } A^{tr} = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m,2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

Ga opnieuw zelf de volgende beweringen na:

**Stelling 4.** Voor alle matrices  $A$  en  $B$  waarvoor de bewerkingen goed gedefinieerd zijn geldt

1.  $(A^{tr})^{tr} = A$ ,
2.  $(A + B)^{tr} = A^{tr} + B^{tr}$ ,
3.  $(\lambda A)^{tr} = \lambda \cdot A^{tr}$ ,  $\lambda \in \mathbb{R}$ ,
4.  $(A \cdot B)^{tr} = B^{tr} \cdot A^{tr}$ .

## 2 Informatie verwerken met matrices

**Een tewerkstellingsmodel.** In een gemeenschap verliest 30% van de actieve bevolking elk jaar zijn werk en 20% van de werklozen vindt elk jaar werk. De actieve bevolking bestaat uit 10 000 mensen, waarvan 2000 werkzoekend zijn. Hoeveel werkenden en hoeveel werklozen zijn er in deze gemeenschap na 13 jaar?

We behandelen eerst een eenvoudiger probleem: Hoeveel werkenden en hoeveel werklozen zijn er in deze gemeenschap na 1 jaar? De verdeling na 1 jaar wordt gegeven door:

$$\text{werkenden na 1 jaar} = 0.70 \times 8000 + 0.20 \times 2000 = 6000,$$

$$\text{werklozen na 1 jaar} = 0.30 \times 8000 + 0.80 \times 2000 = 4000.$$

In het midden herkennen we het matrixproduct van de matrices

$$A = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \quad \text{en} \quad \vec{x}_0 = \begin{bmatrix} 8000 \\ 2000 \end{bmatrix}.$$

In de eerste rij van  $A$  vind je het percentage van de werkende bevolking dat na 1 jaar nog aan de slag is en het percentage van de werkzoekenden die na een jaar werk gevonden hebben. Analoog staat, in rij 2, het percentage van diegenen met werk die na een jaar hun job verliezen en het percentage van de werkzoekenden die na een jaar nog steeds werkloos zijn. De kolom  $\vec{x}_0$  stelt de verdeling van de actieve bevolking over werkenden en werkzoekenden voor op het begintijdstip. De verdeling na 1 jaar wordt dus inderdaad gegeven door:

$$\begin{aligned} \vec{x}_1 &= A \cdot \vec{x}_0 = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \cdot \begin{bmatrix} 8000 \\ 2000 \end{bmatrix} \\ &= \begin{bmatrix} 0.70 \times 8000 + 0.20 \times 2000 \\ 0.30 \times 8000 + 0.80 \times 2000 \end{bmatrix} = \begin{bmatrix} 6000 \\ 4000 \end{bmatrix}. \end{aligned}$$

De verdeling na 2 jaar vinden we door het voorgaande resultaat als begintoestand te nemen:

$$\begin{aligned} \vec{x}_2 &= A \cdot \vec{x}_1 = A \cdot (A \cdot \vec{x}_0) = (A \cdot A) \cdot \vec{x}_0 = A^2 \cdot \vec{x}_0 \\ &= \begin{bmatrix} 0.55 & 0.3 \\ 0.45 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 8000 \\ 2000 \end{bmatrix} = \begin{bmatrix} 5000 \\ 5000 \end{bmatrix}. \end{aligned}$$

Zo kunnen we natuurlijk nog een tijdje doorgaan. De verdeling na 12 jaar is

$$\vec{x}_{12} = A^{12} \cdot \vec{x}_0 = \begin{bmatrix} 4000, 9765 \\ 5999, 0234 \end{bmatrix},$$

en na 13 jaar vinden we uiteindelijk

$$\vec{x}_{13} = A^{13} \cdot \vec{x}_0 = \begin{bmatrix} 4000, 4882 \\ 5999, 5117 \end{bmatrix}.$$

Na 13 jaar zijn er dus (ongeveer) 4000 werkenden en 6000 werkzoekenden. Als we nog verder zouden rekenen, dan zien we dat er eigenlijk niet veel meer aan die verhouding verandert. Na een lange periode blijft de verdeling van de actieve bevolking over werkende en werkzoekenden (ongeveer) constant. In de volgende hoofdstukken zullen we ondermeer de volgende vragen beantwoorden:

- Zijn er snellere methodes om deze conclusies te bereiken?

- Blijven die conclusies geldig als we de beginconditie  $X_0$  wijzigen?

**Koekjesfabriek 'Sweet Cookie'.** Het bedrijf Sweet Cookie heeft 3 vestigingen. In elk van de vestigingen produceren en verkopen ze 3 soorten koekjes: chocoladekoekjes, kokosnootkoekjes en gewone koekjes. De 3 basisbestanddelen van elk koekje zijn boter, bloem en suiker. Voor de chocoladekoekjes en de kokosnootkoekjes zijn er echter nog extra ingrediënten nodig. In de eerste tabel staan de verkochte hoeveelheden per vestiging en in de tweede tabel vinden we de verhoudingen van de 3 basisbestanddelen per kg koekjes weer.

	chocolade	kokos	gewone		bloem	boter	suiker
Vestiging 1	100 kg	80 kg	100 kg	chocolade	0,5 kg	0,2 kg	0,1 kg
Vestiging 2	80 kg	60 kg	120 kg	kokos	0,4 kg	0,2 kg	0,2 kg
Vestiging 3	120 kg	40 kg	80 kg	gewone	0,6 kg	0,3 kg	0,2 kg

De gegevens in deze twee tabellen verzamelen we in de matrices  $A$  en  $B$ , respectievelijk. Ten slotte gebruiken we  $\vec{p}$  voor de kolom met de prijzen van de bestanddelen, en de kolom  $\vec{e}$  voor de extra kosten per kg koeken:

	prijs per kg		prijs per kg
bloem	2 euro	chocolade	3 euro
boter	5 euro	kokos	2 euro
suiker	1 euro	gewone	0 euro

Matrixrekenen geeft makkelijk de totale kostprijs per vestiging voor de productie van de koekjes. De matrix  $B \cdot \vec{p}$  geeft immers de basisbestanddelenprijs in euro, per kg koekjes, terwijl  $B \cdot \vec{p} + \vec{e}$  de totale kostprijs voor de productie in euro, per kg koekjes bevat. De kolommatrix

$$\vec{k} = A \cdot (B \cdot \vec{p} + \vec{e}) = \begin{bmatrix} 1120 \\ 996 \\ 1004 \end{bmatrix}$$

geeft dan de totale kostprijs voor de productie van de koekjes, in euro, per vestiging.

De directeur van Sweet Cookie wil graag de prijs per kilogram koekjes bepalen om in Vestiging 1 een winst van 260 euro te realizeren, in Vestiging 2 een winst van 244 euro en in Vestiging 3 een winst van 156 euro.

Stel hiertoe de prijs per kg van de 3 soorten koekjes, in matrixvorm, vorm door  $\vec{x} = \begin{bmatrix} x_c \\ x_k \\ x_g \end{bmatrix}$ .

De winst van de twee vestigingen voldoet dan aan:

$$\vec{w} = A \cdot \vec{x} - \vec{k}$$

en dus is  $\vec{x} = A^{-1} \cdot (\vec{w} + \vec{k})$ . Om de vraag van de directeur te kunnen beantwoorden moeten we, in de volgende paragraaf, betekenis geven aan het concept  $A^{-1}$ , en uitleggen hoe we dit kunnen berekenen.

### 3 Elementaire matrices en inverteerbaarheid

**Elementaire matrices.** We noteren de  $m$ -dimensionale vectoren (i.e. kolommen) met enkel een 1 op plaats  $i$  als  $\vec{e}_i$ ,

$$\vec{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i\text{-de rij.}$$

Er geldt dan  $\vec{e}_i^{tr} \cdot \vec{e}_j = \delta_{ij}$ . In het geval  $i \neq j$  is de vierkante  $m \times m$ -matrix  $E_{ij} = \vec{e}_i \cdot \vec{e}_j^{tr}$  een matrix met enkel een 1 op plaats  $(i, j)$ , en verder allemaal nullen,

$$E_{ij} = \begin{bmatrix} 0 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix} \leftarrow i\text{-de rij.}$$

$\uparrow j\text{-de kolom}$

Elke matrix  $A = (a_{ij})$  kan dan geschreven worden als een som van scalaire veelvouden van dergelijke matrices,

$$A = a_{11}E_{11} + \dots + a_{1n}E_{1n} + a_{21}E_{21} + \dots + a_{2n}E_{2n} + \dots + a_{m1}E_{m1} + \dots + a_{mn}E_{mn}.$$

**Lemma 1.** *Er geldt:  $E_{ij} \cdot E_{kl} = \delta_{jk} E_{il}$ .*

*Bewijs.* We moeten aantonen dat, als  $j \neq k$ ,  $E_{ij} \cdot E_{kl} = 0_m$  (met  $0_m$  de nulmatrix van orde  $m$ ) en dat, als  $j = k$ ,  $E_{ij} \cdot E_{jl} = E_{il}$ .

Beide beweringen volgen uit

$$E_{ij} \cdot E_{kl} = (\vec{e}_i \cdot \vec{e}_j^{tr}) \cdot (\vec{e}_k \cdot \vec{e}_l^{tr}) = \vec{e}_i \cdot (\vec{e}_j^{tr} \cdot \vec{e}_k) \cdot \vec{e}_l^{tr}.$$

In het geval  $j \neq k$  is de (1,1)-matrix (= het getal)  $\vec{e}_j^{tr} \cdot \vec{e}_k$  net 0, en dus is  $E_{ij} \cdot E_{kl}$  de nulmatrix. In het geval waarbij  $j = k$  is  $\vec{e}_j^{tr} \cdot \vec{e}_j = 1$  en dus is  $E_{ij} \cdot E_{jl} = \vec{e}_i \cdot \vec{e}_l^{tr} = E_{il}$ .  $\square$

Met de matrices  $E_{ij}$  kunnen we de *elementaire matrices* construeren. Er bestaan drie types van dergelijke matrices.

- **Type I.**  $U_{ij}(a) = I + aE_{ij}, \quad i \neq j$

Dit zijn boven- of onderdriehoeksmatrices met alle diagonalelementen gelijk aan 1 en hoogstens één niet-diagonaalcomponent verschillend van nul (en gelijk aan  $a$ ).

- **Type II.**  $V_{ij} = I + E_{ij} + E_{ji} - E_{ii} - E_{jj}$ .

Hier werden de  $i$ -de en  $j$ -de diagonaalcomponenten van de eenheidsmatrix vervangen door een 0, en er werden twee 1-en toegevoegd (namelijk op plaats plaats  $(i, j)$  en op plaats  $(j, i)$ ).

- **Type III.**  $W_i(c) = I + (c - 1)E_{ii}, \quad c \neq 0$ .

Eén diagonalelement van de eenheidsmatrix werd vervangen door een niet-nul element.

De elementaire  $2 \times 2$  matrices zijn

$$U_{12}(a) = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}, \quad U_{21}(a) = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix} \quad V_{12} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad W_1(c) = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \quad W_2(c) = \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}.$$

Voorbeelden met  $m = 3$  zijn

$$U_{23}(a) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & a \\ 0 & 0 & 1 \end{bmatrix}, \quad V_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad W_3(c) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{bmatrix}.$$

Zij  $A$  een  $m \times n$  matrix waarvan we de  $i$ -de rij noteren met  $X_i$  ( $i = 1, \dots, m$ ). De matrix  $A$  voorvermenigvuldigen met een elementaire matrix heeft volgend effect:

- De matrix  $U_{ij}(a).A$  heeft dezelfde rijen als  $A$ , behalve de rij  $X_i$ , die vervangen werd door de rij  $X_i + aX_j$ .
- De matrix  $V_{ij}.A$  heeft dezelfde rijen als  $A$ , behalve de  $i$ -de rij  $X_i$  en  $j$ -de rij  $X_j$  die van plaats omgewisseld werden.
- De matrix  $W_i(c).A$  heeft dezelfde rijen als  $A$ , behalve de  $i$ -de rij  $X_i$ , die vervangen werd door de rij  $cX_i$ .

Het is duidelijk dat we elk van de elementaire rijoperaties die we kennen uit het Gauss-Jordan eliminatiealgoritme kunnen opvatten als een voorvermenigvuldiging van de uitgebreide matrix met een elementaire matrix. Analoog, als  $A$  een  $m \times n$  matrix is dan hebben we dat

$$\text{ref}(A) = L_1 \dots L_k \cdot A,$$

voor zekere elementaire  $m \times m$  matrices  $L_i$ .

### De inverse van een matrix.

**Definitie 3.** Een vierkante  $n \times n$  matrix  $A$  noemen we *inverteerbaar* of *omkeerbaar* indien er een  $n \times n$  matrix  $B$  bestaat zodat

$$B \cdot A = I_n \quad \text{en} \quad A \cdot B = I_n.$$

**Lemma 2.** Als  $A$  inverteerbaar is dan is de matrix  $B$  met bovenstaande eigenschappen uniek.

*Bewijs.* Als ook  $B_2 \cdot A = I_n$  en  $A \cdot B_2 = I_n$ , dan is  $B_2 = B_2 \cdot I_n = B_2 \cdot A \cdot B = I_n \cdot B = B$ . □

We noteren daarom  $B = A^{-1}$ .

**Stelling 5.** Voor alle inverteerbare  $n \times n$  matrices  $A$  en  $B$  geldt:

$$1. \quad (A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$$

$$2. \quad (A^{tr})^{-1} = (A^{-1})^{tr}$$

*Bewijs.* Het bewijs volgt uit het feit dat de matrices  $B^{-1} \cdot A^{-1}$  en  $(A^{-1})^{tr}$  aan de voorwaarden van een inverse voor  $A \cdot B$  en  $A^{tr}$  voldoen, en uit het feit dat de inverse uniek is. □

De elementaire matrices zijn allen inverteerbaar en hun inverse is een elementaire matrix van hetzelfde type.

**Stelling 6.** Er geldt

$$U_{ij}(a) \cdot U_{ij}(-a) = I, \quad V_{ij} \cdot V_{ij} = I, \quad W_i(c) \cdot W_i(c^{-1}) = I.$$

*Bewijs.* We tonen enkel de eerste en de derde uitspraken aan. Het aantonen van de tweede bewering is analoog, maar vergt meer rekenwerk.

Er geldt, wegens Lemma 1,

$$(I + aE_{ij}) \cdot (I - aE_{ij}) = I - aE_{ij} + aE_{ij} - a^2 E_{ij}E_{ij} = I.$$

Tevens is

$$\begin{aligned}(I + (c - 1)E_{ii}).(I + (\frac{1}{c} - 1)E_{ii}) &= I + (\frac{1}{c} - 1)E_{ii} + (c - 1)E_{ii} + (c - 1)(\frac{1}{c} - 1)E_{ii}E_{ii} \\ &= I + (\frac{1}{c} - 1 + c - 1 + 1 - C - \frac{1}{c} + 1)E_{ii} = I.\end{aligned}$$

□

**Stelling 7.** Voor een vierkante  $n \times n$  matrix  $A$  zijn volgende uitspraken equivalent:

1.  $\text{rank}(A) = n,$
2.  $\text{ref}(A) = I_n,$
3.  $A$  is een product van elementaire matrices,
4.  $A$  is inverteerbaar.
5. Het stelsel  $A.\vec{x} = \vec{b}$  van  $n$  vergelijkingen in  $n$  onbekenden heeft een unieke oplossing (en die is dan noodzakelijk  $\vec{x} = A^{-1}.\vec{b}$ ).

*Bewijs.* (1)  $\Rightarrow$  (2): Als  $\text{rank}(A) = n$  dan is het aantal leidende 1-en net gelijk aan  $n$ . Maar  $n$  is voor een vierkante matrix ook het aantal rijen, waaruit  $\text{ref}(A) = I_n$ .

(2)  $\Rightarrow$  (3): Wanneer  $\text{ref}(A) = I_n$ , dan bestaat er een product van elementaire matrices  $C = L_1 \dots L_k$  zodat  $C.A = I_n$  (via het Gauss-Jordan-algoritme). Elk van die elementaire matrices is inverteerbaar met elementair inverse  $L_i^{-1}$ . Als gevolg is ook  $C$  inverteerbaar, met inverse  $C^{-1} = L_k^{-1} \dots L_1^{-1}$ . Indien we de gelijkheid  $C.A = I_n$  voorvermenigvuldigen met  $C^{-1}$  dan vinden we  $A = C^{-1}$ , wat een product is van elementaire matrices.

(3)  $\Rightarrow$  (4): Dit volgt opnieuw uit het feit dat alle elementaire matrices inverteerbaar zijn.

(4)  $\Rightarrow$  (5): Wanneer  $A$  inverteerbaar is, dan volgt uit  $A.\vec{x} = \vec{b}$  dat  $A^{-1}.A.\vec{x} = A^{-1}.\vec{b}$  of  $\vec{x} = A^{-1}.\vec{b}$ .

(5)  $\Rightarrow$  (1): Uit Stelling 1 volgt dat bij een unieke oplossing de rang van  $A$  (een vierkante matrix) net  $n$  moet zijn. □

In de definitie van een inverteerbare matrix staan eigenlijk twee voorwaarden,  $A.B = I_n$  en  $B.A = I_n$ . De volgende stelling toont aan dat één van beide voorwaarden reeds voldoende is om inverteerbaarheid te besluiten.

**Stelling 8.** Zij  $A$  een  $n \times n$ -matrix. Als er  $n \times n$ -matrix  $B$  bestaat met  $A.B = I_n$ , dan is  $A$  inverteerbaar met  $A^{-1} = B$ .

*Bewijs.* Beschouw  $\text{ref}(A) = L_1 \dots L_k \cdot A$ . We tonen eerst aan dat  $\text{ref}(A)$  geen nulrij heeft, uit het ongerijmde. Indien  $\text{ref}(A)$  wel een nulrij zou hebben, dan krijgt ook  $\text{ref}(A) \cdot B$  een nulrij, via het matrixproduct. Uit  $A \cdot B = I_n$  volgt echter ook dat  $\text{ref}(A) \cdot B = L_1 \dots L_k \cdot B = I_n$ , wat een product van elementaire matrices is. Een dergelijk product is inverteerbaar, uit de vorige stelling. We kunnen dus besluiten dat  $\text{ref}(A)$  geen nulrij heeft.

Omdat  $\text{ref}(A)$  geen nulrij heeft is, per definitie,  $\text{rank}(A) = n$ , wat op zijn beurt betekent dat  $A$  inverteerbaar is. Bovendien geldt dan  $B = A^{-1} \cdot A \cdot B = A^{-1} \cdot I_n = A^{-1}$ .  $\square$

We kunnen natuurlijk een analoge eigenschap aantonen, startend van  $B \cdot A = I_n$ .

Om de inverse  $A^{-1}$  uit te rekenen bekijk je de  $n \times 2n$  matrix  $\begin{bmatrix} A & I_n \end{bmatrix}$ . Uit het bewijs van Stelling 7 kunnen we immers besluiten dat dat

$$\text{ref} \begin{bmatrix} A & I_n \end{bmatrix} = \begin{bmatrix} I_n & A^{-1} \end{bmatrix}.$$

Het idee is dus om net die elementaire matrices (= rijoperaties) uit te voeren die ervoor zorgen dat  $A$  in  $I_n$  omgezet wordt. Diezelfde rijoperaties zorgen er terzelfdertijd voor dat  $I_n$  in  $A^{-1}$  zal omgezet worden.

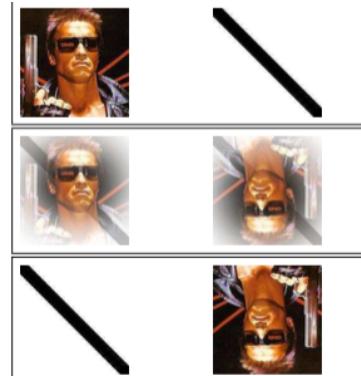
Hieronder staat de uitwerking van deze methode op een voorbeeld.

$$\left[ \begin{array}{cc|cc} 2 & 6 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{array} \right] \quad \left[ \begin{array}{c|c} A & I_n \end{array} \right]$$

$$\left[ \begin{array}{cc|cc} 1 & 3 & 1/2 & 0 \\ 1 & 4 & 0 & 1 \end{array} \right] \quad \left[ \begin{array}{c|c} \dots & \dots \end{array} \right]$$

$$\left[ \begin{array}{cc|cc} 1 & 3 & 1/2 & 0 \\ 0 & 1 & -1/2 & 1 \end{array} \right] \quad \left[ \begin{array}{c|c} \dots & \dots \end{array} \right]$$

$$\left[ \begin{array}{cc|cc} 1 & 0 & 2 & -3 \\ 0 & 1 & -1/2 & 1 \end{array} \right] \quad \left[ \begin{array}{c|c} I_n & A^{-1} \end{array} \right]$$



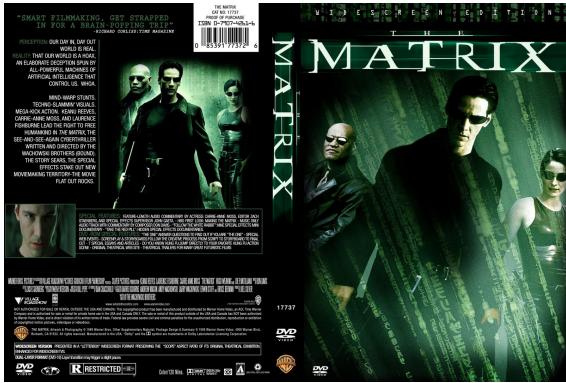
We kunnen besluiten dat

$$\begin{bmatrix} 2 & 6 \\ 1 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -3 \\ -\frac{1}{2} & 1 \end{bmatrix}.$$

Later zullen we een meer directe methode zien om de inverse van een inverteerbare matrix te berekenen met behulp van determinanten.

## 4 Error-correcting coding

De totale opslagruimte van een DVD is ongeveer 5.64GB. Toch zal je via je computer er nooit meer dan ongeveer 4.7GB data op kunnen zetten, dus ongeveer 20% minder. Hoe komt dit?



De reden is dat deze extra ruimte gebruikt wordt voor 'error correcting code' (foutverbeterende code) die op je data wordt losgelaten vooraleer ze op de DVD geschreven wordt. Zonder deze codes zou zelfs het kleinste krasje op je DVD hem totaal onleesbaar maken.

Het idee van foutverbeterende codes is dat je aan je data extra informatie toevoegt die toelaat na te gaan of er in het transitieproces kleine fouten zich hebben voorgedaan en die dan te verbeteren. Hieronder bespreken we het voorbeeld van de *Hamming (7,4) code*. Ze werd in 1950 door Richard Hamming bedacht, toen hij voor Bell Telephone Laboratories werkte, en hij zich ergerde aan de vele fouten die eigen waren aan het inlezen van ponskaarten (waarmee toenmalige computers werkten).

We veronderstellen dat je data een bit-stream is. We splitsen deze stream op in stukken van 4-bits. Bijvoorbeeld, de stream 101101101010... splitsen we op als 1011 en 0110, enzovoort. Het *encoding proces* zal aan elk van deze 4-bits nog 3 extra cijfers toevoegen.

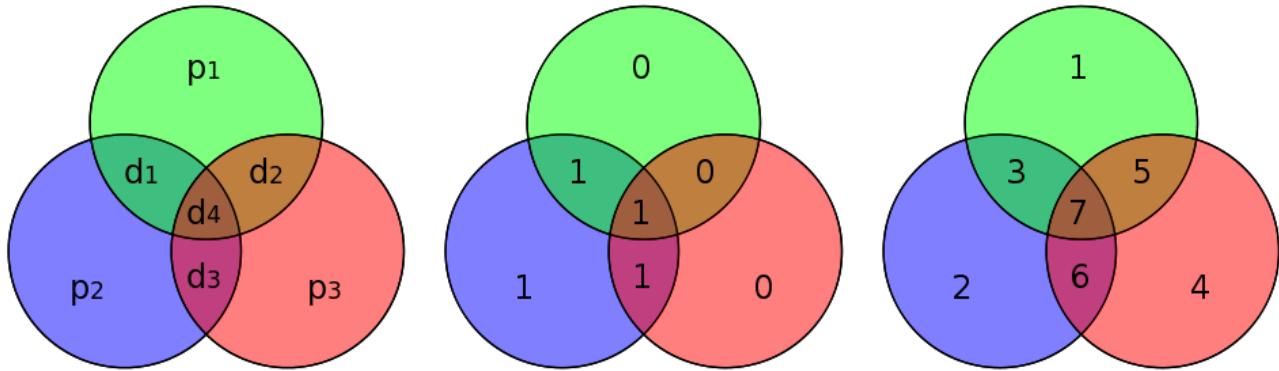
De cijfers die we toevoegen zullen natuurlijk ook enkel 0 of 1 zijn. De verzameling  $\mathbb{F}_2 = \{0, 1\}$  kan voorzien worden van een (inwendige) optelling en vermenigvuldiging, als volgt gedefinieerd:

$$\begin{array}{c} + \quad | \quad 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad \times \quad | \quad 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

Merk op dat je deze bewerkingen ook uit de standaardbewerkingen kan afleiden als je 1 vervangt door 'even' en 0 door 'oneven'. Met deze bewerkingen krijgt de 2-bit verzameling  $\mathbb{F}_2$  de wiskundige structuur van een *lichaam*.

De 3 extra bits die de Hammingcode toevoegt kunnen op volgende wijze bepaald worden.

Beschouw eerst de eerste figuur hieronder.



In de 4 doorsnedes plaatsen we de gekende 4 cijfers  $d_1d_2d_3d_4$  van de bitstream. De getallen  $p_1, p_2$  en  $p_3$  worden nu op zo'n wijze bepaald dat de som van alle getallen in elk van de 3 cirkels steeds 0 is, voor de boven gedefinieerde bewerking. Bijvoorbeeld, voor de 4-bit 1011 krijgen we de tweede figuur hierboven. Indien men  $d_1d_2d_3d_4$ , bijvoorbeeld, codeert als  $p_1p_2d_1p_3d_2d_3d_4$  (zie ook de derde figuur), dan wordt de corresponderende 7-bit dus 0110011.

Een tweede manier om dit encodingproces te beschrijven maakt gebruik van de zogenaamde *generator matrix*

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

Wanneer we de 4-bit opvatten als een kolomvector, dan kunnen we haar links vermenigvuldigen met  $G$  om een kolomvector met 7 componenten te bekomen. Bijvoorbeeld,

$$G \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad G \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \quad G \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Merk op dat dit hetzelfde resultaat geeft als procedure via de doorsnedes. De boodschap 101101101010 wordt dus doorgestuurd als 011001111001101011010.

Tijdens het doorsturen (of opslag) kan het gebeuren dat sommige bits verkeerd worden doorgegeven. We veronderstellen echter dat het kanaal betrouwbaar genoeg is om hoogstens 1 fout te maken voor elke 7 doorgestuurde bits. Aan de andere kant van de lijn wordt de inkomende bitstream opgesplitst in 7-dimensionale vectoren. Voor elk van deze vectoren  $\vec{v}$  wordt eerst het *syndroom*  $\vec{s}$  berekend,

$$\vec{s} = H\vec{v},$$

waar  $H$  staat voor de *parity check matrix* gegeven door

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

Merk op dat de kolommen van  $H$  alle binaire voorstellingen geven van de getallen 1 tot en met 7 (van onder naar boven, en van links naar rechts bekijken).

Het syndroom is een 3-dimensionale vector. Bijvoorbeeld voor de 'goeddoorgestuurde' code 0110011 is het syndroom

$$\vec{s} = H \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Als we echter de 0 van het 5de bit vervangen door een 1 krijgen we een 'foutdoorgestuurde' code. Haar syndroom is dan

$$\vec{s} = H \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

De waarden van het syndroom hadden we kunnen voorspellen. Elke goeddoorgestuurde 7-bit is van de vorm  $G.\vec{w}$ , met  $\vec{w}$  een 4-bits vector. Men rekent makkelijk na dat  $H.G$  net de (3,4)-nulmatrix is, waaruit volgt dat  $H.G.\vec{w}$  noodzakelijk de 3-dimensionale nulvector is.

Maar wat gebeurt er als er een fout is opgetreden? Wat opvalt bij de foutdoor gestuurde code is dat het syndroom net de binaire voorstelling geeft van het getal 5, i.e. net de plaats in de code waar de fout zich bevindt. Ook dat kunnen we algemeen verklaren. Gebruik  $e_i$  voor de 7-vectoren met enkel een 1 op de  $i$ -de plaats. Een stuk code met een fout op plaats  $i$  kan dan voorgesteld worden als

$$G.\vec{w} + \vec{e}_i,$$

en het bijhorende syndroom wordt

$$s = H.(G\vec{w} + \vec{e}_i) = H.\vec{e}_i$$

Het product van een matrix met  $\vec{e}_i$  levert altijd de  $i$ -de kolom van die matrix op. In het geval van  $H$  is de  $i$ -de kolom net de binaire voorstelling van het getal  $i$ . Dit laat ons toe te ontdekken dat er een fout was opgetreden en wel degelijk op de  $i$ -de bit. Als we dit bit aanpassen (van 1 naar 0, of van 0 naar 1) vinden we het correcte codewoord terug.

De Hammingcode geeft dus een manier om 1 enkele fout in de gecodeerde boodschap te verbeteren. Eens deze vrij van fouten verklaard is, kan men de oorspronkelijke boodschap decoderen via de matrix

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

Bijvoorbeeld, de 7-bit 0110011 is afkomstig van de 4-bit gegeven door

$$R. \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Het nadeel van de Hammingcode is echter dat men er steeds vanuit moet gaan dat er slechts 1 fout gemaakt kan worden bij het doorsturen van 7 bits. Stel je immers voor dat een boodschap twee fouten bevat, op plaats 1 en 2. De ontvangen boodschap is dan van het type  $G.\vec{w} + \vec{e}_1 + \vec{e}_2$ . Het syndroom van deze boodschap is dan  $H.\vec{e}_1 + H.\vec{e}_2 = [1\ 0\ 0]^{tr} + [0\ 1\ 0]^{tr} = [1\ 1\ 0]^{tr}$ . Maar, ook  $H.\vec{e}_3 = [1\ 1\ 0]^{tr}$ . De ontvanger kan dus niet onderscheiden tussen een

boodschap met twee fouten (op plaats 1 en 2), en een boodschap met slechts 1 fout (op plaats 3).

De [Hamming-code](#) is een zeer eenvoudige code. Er bestaan betere codes waarin men minder extra bits moet toevoegen of die meerdere fouten kunnen verbeteren. Bij DVDs worden de meer gecompliceerde [Reed-Solomon codes](#) gebruikt, zie bvb. de cursus 'Codetheorie en cryptografie'.

# 3 Vectoren en deelruimten

## 1 $\mathbb{R}^n$ als verzameling van vectoren

Binnen de Lineaire Algebra is een belangrijke rol weggelegd voor de kolommatrices. In de voorgaande hoofdstukken hebben we een kolommatrix vaak een *vector* genoemd. De reden hiertoe is dat de ruimte  $\mathbb{R}^n$  (opgevat als de verzameling van alle  $n$ -dimensionale kolommen) de wiskundige structuur bezit van een *vectorruimte*. Een precieze definitie van dit begrip moet wachten tot Hoofdstuk 6, maar de benaming vectorruimte verwijst naar de 2 bewerkingen die we op kolommen kunnen uitvoeren (net als op alle matrices van gepaste orde): we kunnen kolommen optellen en een scalair veelvoud van een kolom nemen:

$$\vec{v} + \vec{w} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{bmatrix}, \quad \lambda \cdot \vec{v} = \lambda \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{bmatrix}.$$

Verder bestaat er een neutraal element voor de optelling,

$$\vec{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

en heeft elke kolom  $\vec{v}$  een tegengestelde, namelijk  $-\vec{v}$ .

Een veel gebruikte grafische voorstelling voor vectoren is die waarbij de vectoren voorgesteld worden aan de hand van pijlen. De som van de pijlen  $\vec{v}_1$  en  $\vec{v}_2$  correspondeert dan met de diagonale pijl binnen het parallellogram opgespannen door de twee termen (zie bijvoorbeeld de som van de vectoren  $\vec{f}_1$  en  $\vec{f}_2$  in de eerstvolgende figuur hieronder). De scalaire vermenigvuldiging stemt overeen met het op schaal verlengen van de pijl.

De verzameling van de vectoren

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

noemen we de *standaardbasis* van  $\mathbb{R}^n$ . Elke kolomvector  $\vec{v}$  kunnen we dan interpreteren als een *lineaire combinatie* van de speciale vectoren  $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ ,

$$\vec{v} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} = a_1 \vec{e}_1 + a_2 \vec{e}_2 + \dots + a_n \vec{e}_n.$$

Het woord 'basis' verwijst naar twee eigenschappen van de verzameling  $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ .

**Definitie 4.** Een stel vectoren  $\{\vec{v}_1, \dots, \vec{v}_k\}$  in  $\mathbb{R}^n$  noemen we

1. lineair onafhankelijk indien uit  $\lambda_1 \vec{v}_1 + \dots + \lambda_k \vec{v}_k = \vec{0}$  volgt dat  $\lambda_i = 0$  voor alle  $1 \leq i \leq k$ .
2. een voortbrengend stel indien elke vector  $\vec{v} \in \mathbb{R}^n$  geschreven kan worden als een (niet noodzakelijk unieke) lineaire combinatie

$$\vec{v} = a_1 \vec{v}_1 + \dots + a_k \vec{v}_k.$$

3. een basis indien  $\{\vec{v}_1, \dots, \vec{v}_k\}$  zowel lineair onafhankelijk als voortbrengend is.

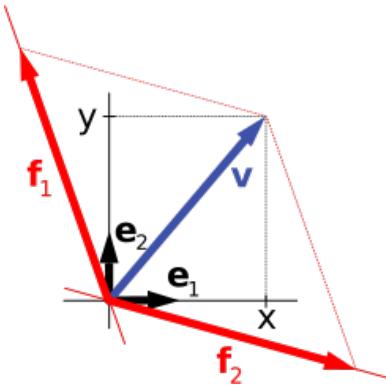
Men kan makkelijk nagaan dat het standaardstel  $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$  aan deze eigenschappen voldoet. Immers, indien

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \vec{0} = \lambda_1 \vec{e}_1 + \dots + \lambda_n \vec{e}_n = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \end{bmatrix}$$

kunnen we enkel concluderen dat  $\lambda_i = 0$  voor alle  $1 \leq i \leq n$ . Het stel is dus lineair onafhankelijk. Bovendien hebben we al gesteld dat elke kolomvector  $\vec{v}$  te schrijven is als een lineaire combinatie van vectoren uit het stel. De verzameling  $\{\vec{e}_1, \dots, \vec{e}_n\}$  is dus inderdaad een basis voor  $\mathbb{R}^n$ .

De standaardbasis is echter niet de enige mogelijke basis voor  $\mathbb{R}^n$ . We illustreren dit met

een voorbeeld in  $\mathbb{R}^2$ .



We gaan na of de vectoren

$$\vec{f}_1 = \begin{bmatrix} -1 \\ 4 \end{bmatrix} \quad \text{en} \quad \vec{f}_2 = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

een basis vormen. Om te beginnen is  $\{\vec{f}_1, \vec{f}_2\}$  lineair onafhankelijk want als

$$\vec{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = a_1 \vec{f}_1 + a_2 \vec{f}_2 = a_1 \begin{bmatrix} -1 \\ 4 \end{bmatrix} + a_2 \begin{bmatrix} 4 \\ -1 \end{bmatrix} = \begin{bmatrix} -a_1 + 4a_2 \\ 4a_1 - a_2 \end{bmatrix}$$

dan moeten we hebben dat  $a_1 = 4a_2$  en  $a_2 = 4a_1$  dus  $a_1 = 16a_1$  wat enkel kan als  $a_1 = 0$  en dus ook  $a_2 = 0$ . Verder is  $\{\vec{f}_1, \vec{f}_2\}$  ook voortbrengend want elke vector  $\vec{v} \in \mathbb{R}^2$  kunnen we schrijven als

$$\vec{v} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \frac{a_1 + 4a_2}{15} \begin{bmatrix} -1 \\ 4 \end{bmatrix} + \frac{4a_1 + a_2}{15} \begin{bmatrix} 4 \\ -1 \end{bmatrix}.$$

Aangezien we nu over twee basissen beschikken kunnen we eenzelfde vector  $\vec{v}$  ontbinden op twee manieren. Bijvoorbeeld, de vector  $\vec{v} = \vec{f}_1 + \vec{f}_2$ , uit de figuur, is de kolom  $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$  t.o.v. de standaardbasis, maar kan ook geschreven worden als de kolom  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$  t.o.v. de nieuwe basis.

De conclusie is dat de vectoren van de vectorruimte  $\mathbb{R}^2$  geïdentificeerd kunnen worden met de 2-dimensionale kolomvectoren over  $\mathbb{R}$ , maar deze identificatie eigenlijk afhangt van de keuze van de basisvectoren.

De volgende stelling geeft een belangrijke karakterisatie van een basis. We beginnen echter met een aantal hulplemma's.

Beschouw een stel vectoren  $\{\vec{v}_1, \dots, \vec{v}_k\}$  in  $\mathbb{R}^n$  die ten opzichte van de standaardbasis volgende kolomvoorstellingen hebben

$$\vec{v}_j = \begin{bmatrix} v_{1j} \\ \vdots \\ v_{nj} \end{bmatrix}.$$

De  $n \times k$ -matrix waarvan de kolommen de vectoren  $\vec{v}_j$  zijn noteren we met

$$V = \begin{bmatrix} v_{11} & \dots & v_{1k} \\ \vdots & & \vdots \\ v_{n1} & \dots & v_{nk} \end{bmatrix} = \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_k \end{bmatrix}.$$

**Lemma 3.** *Als de vectoren  $\{\vec{v}_1, \dots, \vec{v}_k\}$  lineair onafhankelijk zijn in  $\mathbb{R}^n$ , dan is  $k \leq n$ .*

*Bewijs.* Noteer  $\vec{\lambda}$  voor de  $(1, k)$ -kolommatrix met componenten  $\lambda_i$ ,  $i = 1, \dots, k$ . Lineair onafhankelijkheid kunnen we ook uitdrukken als de eigenschap die zegt dat de matrixvergelijking  $V.\vec{\lambda} = \vec{0}$  een unieke oplossing heeft, namelijk  $\vec{\lambda} = \vec{0}$ . Deze matrixvergelijking is een stelsel van  $n$  vergelijkingen in  $k$  onbekenden. Uit Stelling 1 volgt dan dat  $\text{rank}(V)$  net gelijk is aan het aantal onbekenden, in dit geval dus  $k$ . Bij definitie is de rang van een  $n \times k$ -matrix steeds kleiner of gelijk aan  $\min(k, n)$ . We kunnen dus concluderen dat  $k \leq n$ .  $\square$

**Lemma 4.** *Elke basis van  $\mathbb{R}^n$  telt  $n$  vectoren.*

*Bewijs.* We kennen voor  $\mathbb{R}^n$  reeds de standaardbasis, en die bestaat uit  $n$  vectoren. Stel dat  $\{\vec{v}_1, \dots, \vec{v}_k\}$  een tweede basis is. Uit het voorgaande Lemma weten we al dat  $k \leq n$ . We kunnen de redenering echter ook omdraaien. Aangezien  $\{\vec{v}_1, \dots, \vec{v}_k\}$  een basis is kunnen we alle vectoren  $\vec{e}_i$  schrijven als lineaire combinaties van de vectoren  $\vec{v}_j$ ,  $\vec{e}_i = \sum_{j=1}^k w_{ij} \vec{v}_j$ . De matrix  $W = (w_{ij})$  is dan een  $n \times k$ -matrix. Noteer nu  $\vec{\mu}$  voor de  $(1, n)$ -kolommatrix met componenten  $\mu_i$ ,  $i = 1, \dots, n$ . Uitdrukken dat de standaardbasis lineair onafhankelijk is betekent dat de matrixvergelijking (nu in  $n$  onbekenden)  $W.\vec{\mu} = \vec{0}$  een unieke oplossing heeft, namelijk  $\vec{\mu} = \vec{0}$ . Zoals in het vorige lemma leiden we hieruit af dat  $n \leq k$ , en dus is  $k = n$ .  $\square$

Het voorgaande resultaat laat ons toe te zeggen dat de *dimensie* van  $\mathbb{R}^n$  (= het aantal elementen in een basis) gelijk is aan  $n$ . We kunnen ook besluiten dat de matrix  $V$  in dat geval steeds een vierkante  $n \times n$ -matrix is.

**Stelling 9.** *Beschouw een stel vectoren  $\{\vec{v}_1, \dots, \vec{v}_n\}$  van  $n$  vectoren in  $\mathbb{R}^n$ , en de bijhorende matrix  $V$  ten opzichte van de standaardbasis. De volgende uitspraken zijn equivalent:*

1. Het stel  $\{\vec{v}_1, \dots, \vec{v}_n\}$  is lineair onafhankelijk.
2.  $V$  is een inverteerbare  $n \times n$  matrix.
3. Het stel  $\{\vec{v}_1, \dots, \vec{v}_n\}$  is een basis voor  $\mathbb{R}^n$

In dit geval noemen we  $V$  de basisovergangsmatrix voor  $\{\vec{v}_1, \dots, \vec{v}_n\}$ .

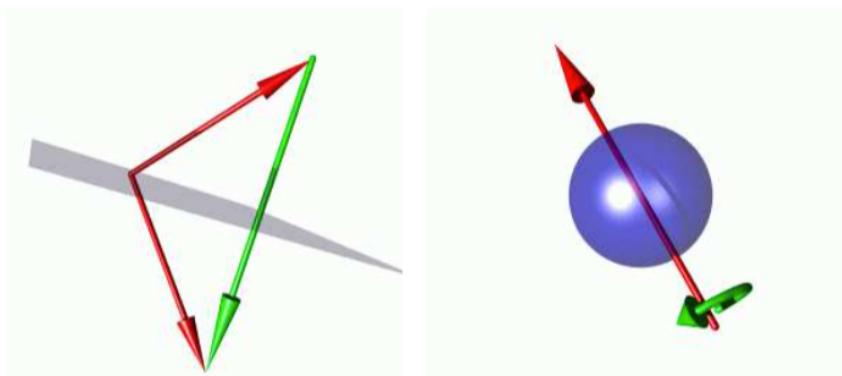
*Bewijs.* (1)  $\Rightarrow$  (2): Uit het bewijs van Lemma 3 weten we dat  $\text{rank}(V) = k = n$ . Uit Stelling 7 volgt dat  $V$  inverteerbaar is.

(2)  $\Rightarrow$  (3): Als  $V$  een inverteerbare  $n \times n$  matrix is dan volgt uit stelling 1 dat elke matrixvergelijking  $V\vec{x} = \vec{b}$  met  $\vec{x}$  en  $\vec{b}$   $n$ -dimensionale kolomvectoren een unieke oplossing heeft, namelijk  $\vec{x} = V^{-1}\vec{b}$ . Als we het geval  $\vec{b} = \vec{0}$  bekijken, dan betekent deze eigenschap dat  $\{\vec{v}_1, \dots, \vec{v}_n\}$  lineair onafhankelijk is.

Anderzijds, voor een willekeurige  $\vec{b}$  kunnen de (uniek bestaande) componenten  $x_i$  van  $\vec{x}$  gebruikt worden om  $\vec{b}$  te schrijven als de lineaire combinatie  $x_1\vec{v}_1 + \dots + x_n\vec{v}_n$ . Hiermee hebben we ook voortbrengendheid aangetoond.

(3)  $\Rightarrow$  (1): Per definitie zijn de vectoren van een basis steeds lineair onafhankelijk.  $\square$

Je kan je natuurlijk afvragen waarom dit alles van belang is. Waarom gebruiken we niet altijd de gemakkelijke standaardbasis?



Wel, afhankelijk van het probleem kan het nuttig zijn een andere basis te nemen om de berekeningen te vergemakkelijken. Bvb. om de reflectie van een straal ten opzichte van een vlak uit te drukken is het makkelijker een basis te kiezen waarvan twee vectoren het vlak opspannen en de derde er orthogonaal op staat. Of, als we rotaties bekijken is het handig één van de basis vectoren te kiezen langs de rotatie-as en de andere twee er loodrecht op.

## 2 Deelruimten

**Definitie 5.** Een deelvectorruimte of kortweg deelruimte van  $\mathbb{R}^n$  is een deelverzameling vectoren  $W \subset \mathbb{R}^n$  met de eigenschappen:

1.  $\vec{0} \in W$ ,
2. als  $\vec{u}, \vec{w} \in W$  dan is ook  $\vec{u} + \vec{w} \in W$ ,
3. als  $\vec{u} \in W$  en  $\lambda \in \mathbb{R}$ , dan is ook  $\lambda\vec{u} \in W$ .

We zeggen dat een stel vectoren  $\{\vec{w}_1, \dots, \vec{w}_d\}$  een *basis is voor de deelruimte*  $W$  indien alle  $\vec{w}_j \in W$ , het stel vectoren  $\{\vec{w}_1, \dots, \vec{w}_d\}$  lineair onafhankelijk is en elke vector  $\vec{v} \in W$  geschreven kan worden als een lineaire combinatie van de vectoren  $\vec{w}_j$ , dat is, er bestaan  $a_j \in \mathbb{R}$  zodat

$$\vec{v} = a_1\vec{w}_1 + \dots + a_d\vec{w}_d.$$

In dit geval zeggen we dat  $W$  *dimensie d heeft*. Ten opzichte van de basis  $\{\vec{w}_1, \dots, \vec{w}_d\}$  kunnen we dan iedere vector  $\vec{v} \in W$  op unieke wijze voorstellen als een  $d$ -dimensionale kolomvector

$$\vec{v} \leftrightarrow \begin{bmatrix} a_1 \\ \vdots \\ a_d \end{bmatrix}$$

en we kunnen bijgevolg de deelruimte  $W$  identificeren met  $\mathbb{R}^d$ .

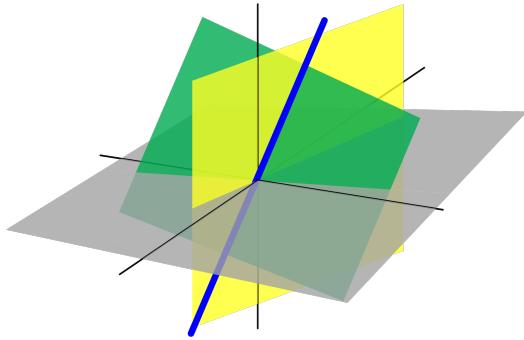
**Het span van een stel vectoren.** Een eenvoudige manier om een dergelijke deelruimte te construeren bestaat uit het verzamelen van alle mogelijke lineaire combinaties van een stel vectoren  $\{\vec{v}_1, \dots, \vec{v}_k\}$  in  $\mathbb{R}^n$ . Een dergelijke deelruimte noemt men het *span* van de vectoren  $\{\vec{v}_1, \dots, \vec{v}_k\}$  en we noteren

$$\text{spa}(\vec{v}_1, \dots, \vec{v}_k) = \{a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_k\vec{v}_k \mid a_i \in \mathbb{R}\}.$$

Bijvoorbeeld, in het span van de vectoren  $\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^{tr}$  en  $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^{tr}$  zitten, bijvoorbeeld, de vectoren  $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^{tr}$ ,  $\begin{bmatrix} 1 & -2 & 1 \end{bmatrix}^{tr}$  en  $\begin{bmatrix} 0 & 2 & 0 \end{bmatrix}^{tr}$ , maar bijvoorbeeld niet de vectoren  $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^{tr}$  of  $\begin{bmatrix} 1 & 1 & 2 \end{bmatrix}^{tr}$ .

Merk op dat de vectoren  $\{\vec{v}_1, \dots, \vec{v}_k\}$  niet noodzakelijk een basis vormen voor  $\text{span}(\vec{v}_1, \dots, \vec{v}_k)$ . Indien, bijvoorbeeld  $\vec{v}_1 = 2\vec{v}_2$ , dan is het duidelijk dat  $\text{span}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k) = \text{span}(\vec{v}_2, \dots, \vec{v}_k)$ .

Hier zijn enkele voorbeelden in  $\mathbb{R}^3$ :



De deelruimte  $\text{span}(\vec{v})$  opgespannen door 1 niet-nul vector in  $\mathbb{R}^3$  is een rechte door de oorsprong. De deelruimte  $\text{span}(\vec{v}, \vec{w})$  opgespannen door 2 niet-nul vectoren in  $\mathbb{R}^3$  is meestal een vlak door de oorsprong *tenzij  $\vec{v}$  en  $\vec{w}$  lineair afhankelijk zijn*. De deelruimte  $\text{span}(\vec{u}, \vec{v}, \vec{w})$  opgespannen door 3 niet-nul vectoren in  $\mathbb{R}^3$  is meestal heel  $\mathbb{R}^3$  *tenzij  $\{\vec{u}, \vec{v}, \vec{w}\}$  lineair afhankelijk zijn*, in dat geval zal  $\text{span}(\vec{u}, \vec{v}, \vec{w})$  een vlak door de oorsprong zijn of een lijn door de oorsprong.

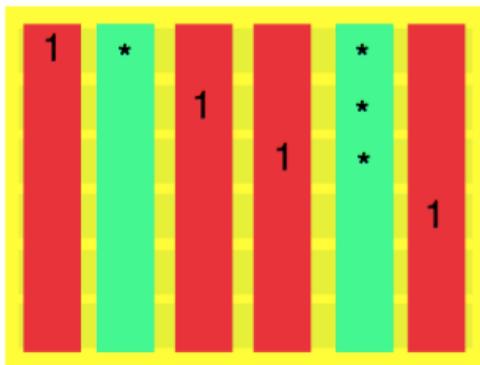
**Stelling 10.** *Beschouw een stel vectoren  $\{\vec{v}_1, \dots, \vec{v}_k\}$  in  $\mathbb{R}^n$  die ten opzichte van de standaardbasis volgende kolomvoorstellingen hebben*

$$\vec{v}_j = \begin{bmatrix} v_{1j} \\ \vdots \\ v_{nj} \end{bmatrix}. \quad \text{Noteer} \quad V = \begin{bmatrix} v_{11} & \dots & v_{1k} \\ \vdots & & \vdots \\ v_{n1} & \dots & v_{nk} \end{bmatrix}$$

voor de  $n \times k$ -matrix waarvan de kolommen de vectoren  $\vec{v}_j$  zijn.

Als  $\{i_1, \dots, i_d\}$  de kolommen in  $\text{ref}(V)$  zijn met een leidende 1 (m.a.w. als de rang van  $V$  gelijk is aan  $d$ ), dan is  $\{\vec{v}_{i_1}, \dots, \vec{v}_{i_d}\}$  een basis voor  $\text{span}(\vec{v}_1, \dots, \vec{v}_k)$ . De dimensie van  $\text{span}(\vec{v}_1, \dots, \vec{v}_k)$  is dus gelijk aan  $d = \text{rank}(V)$ .

*Bewijs.* Laat  $C$  het product zijn van de elementaire matrices die de stappen voorstellen in het Gauss-Jordan algoritme. Dan is  $C.V = \text{ref}(V)$ , waarbij  $\text{ref}(V)$  de volgende vorm heeft



Veronderstel dat er  $d$  kolommen zijn met een leidende 1. Deze kolommen zien eruit als de vectoren  $\{\vec{e}_1, \dots, \vec{e}_d\}$  in  $\mathbb{R}^n$  van de standaardbasis. De deelruimte opgespannen door de kolommen van  $C.V$  is dan gelijk aan de  $d$ -dimensionale deelruimte opgespannen door de vectoren  $\vec{e}_1, \dots, \vec{e}_d$ , dat is

$$\text{span}(C.\vec{v}_1, \dots, C.\vec{v}_k) = \text{span}(C.\vec{v}_{i_1}, \dots, C.\vec{v}_{i_d}) = \text{span}(\vec{e}_1, \dots, \vec{e}_d).$$

We tonen nu aan dat  $\{\vec{v}_{i_1}, \dots, \vec{v}_{i_d}\}$  een basis is voor  $\text{span}(\vec{v}_1, \dots, \vec{v}_k)$ . De vectoren zijn lineair onafhankelijk want als

$$\lambda_{i_1} \vec{v}_{i_1} + \dots + \lambda_{i_d} \vec{v}_{i_d} = \vec{0},$$

dan geldt na voorvermenigvuldigen met  $C$  dat

$$\lambda_{i_1} \vec{e}_1 + \dots + \lambda_{i_d} \vec{e}_d = \vec{0},$$

en wegens de lineaire onafhankelijkheid van (delen) van de standaardbasis volgt onmiddellijk dat  $\lambda_{i_1} = \dots = \lambda_{i_d} = 0$ . De verzameling is tevens voortbrengend. Elke vector  $\vec{v}_j$  (en dus ook elke lineaire combinatie van dergelijke vectoren) kan geschreven worden als een lineaire combinatie van de vectoren  $\{\vec{v}_{i_1}, \dots, \vec{v}_{i_d}\}$ . Uit de rij-echelonvorm van  $C.V = \text{ref}(V)$  is het immers duidelijk dat  $C.\vec{v}_j$  te schrijven is als

$$C.\vec{v}_j = \mu_{j_1} \vec{e}_1 + \dots + \mu_{j_d} \vec{e}_d,$$

Aangewien  $C$  inverteerbaar is kunnen we deze relatie voorvermenigvuldigen met de inverse matrix  $C^{-1}$ , en we krijgen

$$\vec{v}_j = \mu_{j_1} \vec{v}_{i_1} + \dots + \mu_{j_d} \vec{v}_{i_d},$$

wat net voortbrengendheid aantoon. Het aantal elementen van de basis is net  $d = \text{rank}(V)$ , wat het gestelde bewijst.  $\square$

**Deelruimten geassocieerd aan een matrix.** In deze paragraaf veronderstellen we dat  $A$  een  $m \times n$ -matrix is.

**Definitie 6.** *De kern van  $A$  is de deelruimte van  $\mathbb{R}^n$ , gegeven door*

$$\ker(A) = \{\vec{x} \in \mathbb{R}^n \mid A.\vec{x} = \vec{0}\}.$$

*Het beeld van  $A$  is de deelruimte van  $\mathbb{R}^m$ , met*

$$\text{im}(A) = \{\vec{b} \in \mathbb{R}^m \mid \exists \vec{x} \in \mathbb{R}^n \text{ zodat } A.\vec{x} = \vec{b}\}.$$

Ga zelf na dat zowel  $\ker(A)$  als  $\text{im}(A)$  aan de definiërende kenmerken van een deelruimte voldoen. Merk op dat de eerste in  $\mathbb{R}^n$  ligt, en de tweede in  $\mathbb{R}^m$ .

We bespreken enkele interessante eigenschappen van deze deelruimten.

(1) Als  $\vec{a}_1, \dots, \vec{a}_n$  de kolommen zijn van  $A$  dan is een andere karakterisatie voor  $\text{im}(A)$ ,

$$\text{im}(A) = \text{span}(\vec{a}_1, \dots, \vec{a}_n),$$

want als  $\vec{b} \in \text{span}(\vec{a}_1, \dots, \vec{a}_n)$ , dan is

$$\vec{b} = x_1 \vec{a}_1 + \dots + x_n \vec{a}_n,$$

voor zekere  $x_i$ . Als we die in een vector  $\vec{x}$  groeperen, en aan blokvermenigvuldiging denken, dan staat er inderdaad  $\vec{b} = A.\vec{x}$ . We mogen, via Stelling 10, besluiten dat

$$\dim(\text{im}(A)) = \text{rank}(A).$$

(2) Beschouw de rij-echelonvorm  $\text{ref}(A) = C.A$  van  $A$ . Als  $A.\vec{x} = \vec{0}$ , dan is, na vermenigvuldigen met  $C$ , ook  $\text{ref}(A).\vec{x} = \vec{0}$ . Aangezien  $C$  een inverteerbare matrix volgt uit  $\text{ref}(A).\vec{x} = \vec{0}$  ook dat  $A.\vec{x} = \vec{0}$ . We concluderen dus dat

$$\ker(\text{ref}(A)) = \ker(A).$$

**Stelling 11** (Dimensiestelling). *Voor elke  $m \times n$ -matrix geldt*

$$\dim(\ker(A)) + \dim(\text{im}(A)) = n.$$

*Bewijs.* Noem  $s$  de dimensie van  $\ker(A) \subset \mathbb{R}^n$  en beschouw een basis  $\{\vec{v}_1, \dots, \vec{v}_s\}$  van  $\ker(A)$ . Noem  $r$  de dimensie van  $\text{im}(A) \subset \mathbb{R}^m$  en beschouw een basis  $\{\vec{w}_1, \dots, \vec{w}_r\}$  van  $\text{im}(A)$ . Kies nu vectoren  $\vec{u}_j$  in  $\mathbb{R}^n$ , met de eigenschap dat  $A.\vec{u}_j = \vec{w}_j$ ,  $j = 1, \dots, r$ . We zullen aantonen dat de set  $\{\vec{v}_1, \dots, \vec{v}_s, \vec{u}_1, \dots, \vec{u}_r\}$  een basis is voor  $\mathbb{R}^n$ . Omdat alle basissen van  $\mathbb{R}^n$  net  $n$  vectoren bevatten is dan noodzakelijkerwijs  $s + r = n$ .

(1) Voortbrengend? Kunnen we elke  $\vec{x}$  schrijven als een lineaire combinatie van elementen uit de kandidaatbasis? Er geldt natuurlijk dat  $A.\vec{x} \in \text{im}(A)$ , en dus is  $A.\vec{x} = \sum_{j=1}^r \lambda_j \vec{w}_j$ . Hieruit volgt dat  $A.(\vec{x} - \sum_{j=1}^r \lambda_j \vec{u}_j) = \vec{0}$ , wat betekent dat deze vector in  $\ker(A)$ , zit. Er bestaan dus coëfficiënten  $\mu_\alpha$  zodat  $\vec{x} - \sum_{j=1}^r \lambda_j \vec{u}_j = \sum_{\alpha=1}^s \mu_\alpha \vec{v}_\alpha$ . We concluderen dus dat, inderdaad,

$$\vec{x} = \sum_{j=1}^r \lambda_j \vec{u}_j + \sum_{\alpha=1}^s \mu_\alpha \vec{v}_\alpha.$$

(2) Lineair onafhankelijk? Kunnen we uit de voorwaarde  $\vec{0} = \sum_{j=1}^r \lambda_j \vec{u}_j + \sum_{\alpha=1}^s \mu_\alpha \vec{v}_\alpha$  concluderen dat alle  $\lambda_j = 0$  ( $j = 1, \dots, r$ ) en  $\mu_\alpha = 0$  ( $\alpha = 1, \dots, s$ )?

Wanneer we de voorwaarde voorvermenigvuldigen met  $A$ , dan vinden we dat  $\vec{0} = \sum_{j=1}^r \lambda_j A \cdot \vec{u}_j = \sum_{j=1}^r \lambda_j \vec{w}_j$ . Aangezien alle  $\vec{w}_j$  lineair onafhankelijk zijn mogen we concluderen dat  $\lambda_j = 0$  ( $j = 1, \dots, r$ ).

Als we hiermee rekening houden dan blijft er van de voorwaarde enkel  $\vec{0} = \sum_{\alpha=1}^s \mu_\alpha \vec{v}_\alpha$  over. Opnieuw zijn alle  $\vec{v}_\alpha$  lineair onafhankelijk, en dus is ook  $\mu_\alpha = 0$  ( $\alpha = 1, \dots, s$ ).  $\square$

Een andere schrijfwijze voor deze stelling is  $\dim(\ker(A)) = n - \text{rank}(A)$ .

### 3 Orthogonaliteit

**Belangrijke opmerking:** In het eerste hoofdstuk vermeldden we dat de meeste eigenschappen in deze cursus voor vectoren en matrices met componenten in een willekeurig lichaam (zoals  $\mathbb{R}$ ,  $\mathbb{C}$ ,  $\mathbb{F}_2$ , etc) gelden. In deze paragraaf werken we uitzonderlijk uitsluitend met vectoren over  $\mathbb{R}$ .

**Definitie 7.** Het inwendig product of inproduct van twee  $n$ -dimensionale kolomvectoren  $\vec{v}, \vec{w} \in \mathbb{R}^n$  met

$$\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}, \quad \vec{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \quad \text{is} \quad \vec{v} \cdot \vec{w} = v_1 w_1 + \dots + v_n w_n = (\vec{v})^{tr} \cdot \vec{w}.$$

We zeggen dat  $\vec{v}$  en  $\vec{w}$  orthogonaal zijn indien  $\vec{v} \cdot \vec{w} = 0$ . We noteren dan  $\vec{v} \perp \vec{w}$ .

Over de reële getallen  $\mathbb{R}$  is het scalaire product positief definit, dwz. dat als  $\vec{v} \neq \vec{0}$  dat is

$$\vec{v} \cdot \vec{v} = v_1^2 + \dots + v_n^2 > 0.$$

Merk op dat dit niet geldt over het bit-lichaam  $\mathbb{F}_2$  of over de complexe getallen  $\mathbb{C}$ . De *lengte* (of *norm*) van een vector  $\vec{v} \in \mathbb{R}^n$  noteren we met

$$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}.$$

We noemen  $\vec{v}$  een *eenheidsvector* indien  $\|\vec{v}\| = 1$ . Een basis  $\{\vec{v}_1, \dots, \vec{v}_n\}$  van  $\mathbb{R}^n$  noemen we *orthogonaal* indien voor alle  $i \neq j$  geldt dat  $\vec{v}_i \cdot \vec{v}_j = 0$  en we noemen de basis *orthonormaal* indien verder iedere  $\vec{v}_i$  een eenheidsvector is.

**Stelling 12.** (Over de reële getallen  $\mathbb{R}$ )

**1. Pythagoras:** Als  $\vec{x}$  en  $\vec{y}$  orthogonaal zijn dan geldt

$$\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2.$$

**2. Cauchy-Schwartz ongelijkheid:**  $|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \|\vec{y}\|$ .

**3. Driehoeksongelijkheid:**  $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ .

*Bewijs.* (1): Wegens de lineariteit van het inproduct krijgen we:

$$\|\vec{x} + \vec{y}\|^2 = (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) = \vec{x} \cdot \vec{x} + 2\vec{x} \cdot \vec{y} + \vec{y} \cdot \vec{y} = \|\vec{x}\|^2 + \|\vec{y}\|^2,$$

omdat  $\vec{x} \cdot \vec{y} = 0$ .

(2): Merk op dat de eigenschap al zeker geldt voor ofwel  $\vec{x} = \vec{0}$ , of voor  $\vec{y} = \vec{0}$ . Veronderstel nu dat  $\vec{x} \neq \vec{0}$  en  $\vec{y} \neq \vec{0}$ . Er geldt voor alle  $\lambda \in \mathbb{R}$ :

$$0 \leq \|\vec{x} - \lambda\vec{y}\|^2 = (\vec{x} - \lambda\vec{y}) \cdot (\vec{x} - \lambda\vec{y}) = \|\vec{x}\|^2 - 2\lambda(\vec{x} \cdot \vec{y}) + \lambda^2\|\vec{y}\|^2.$$

Beschouw nu het bijzonder geval met  $\lambda = \frac{\vec{x} \cdot \vec{y}}{\|\vec{y}\|^2}$ . Invullen geeft:

$$0 \leq \|\vec{x}\|^2 - \frac{(\vec{x} \cdot \vec{y})^2}{\|\vec{y}\|^2}.$$

Hieruit volgt dat  $(\vec{x} \cdot \vec{y})^2 \leq \|\vec{x}\|^2 \|\vec{y}\|^2$ , wat leidt tot de tweede eigenschap (als je opmerkt dat  $|\dots|$  voor absolute waarde staat).

(3): Door gebruik te maken van de voorgaande eigenschap vinden we

$$\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2 + 2\vec{x} \cdot \vec{y} \leq \|\vec{x}\|^2 + \|\vec{y}\|^2 + 2\|\vec{x}\| \|\vec{y}\| = (\|\vec{x}\| + \|\vec{y}\|)^2.$$

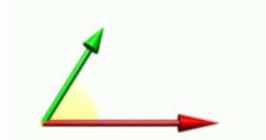
□

Uit de Cauchy-Schwartz ongelijkheid zien we dat

$$-1 \leq \frac{\vec{x}}{\|\vec{x}\|} \cdot \frac{\vec{y}}{\|\vec{y}\|} \leq 1.$$

We kunnen daarom de *hoek*  $\alpha \in [0, \pi]$  tussen de vectoren  $\vec{x}$  en  $\vec{y}$  definiëren aan de hand van

$$\cos(\alpha) = \frac{\vec{x}}{\|\vec{x}\|} \cdot \frac{\vec{y}}{\|\vec{y}\|}.$$



**Definitie 8.** Als  $W$  een deelruimte is van  $\mathbb{R}^n$  dan noemen we het orthogonaal complement van  $W$  de deelruimte van  $\mathbb{R}^n$  gegeven door

$$W^\perp = \{\vec{u} \in \mathbb{R}^n \mid \vec{u} \cdot \vec{w} = 0, \forall \vec{w} \in W\}.$$

(Ga na dat dit inderdaad een deelruimte definieert.)

Indien  $W = \text{span}(\vec{v}_1, \dots, \vec{v}_k)$  dan is  $\vec{u} \in W^\perp$  als  $\vec{u} \cdot \vec{v}_i = 0$ , voor alle  $i = 1 \dots k$ .

**Lemma 5.** Er geldt  $(W^\perp)^\perp = W$  en  $W \cap W^\perp = \{\vec{0}\}$ .

*Bewijs.* We bewijzen enkel de tweede eigenschap en laten de eerste als oefening. Als  $\vec{v}$  in beide deelruimten zit, dan is  $\vec{v} \cdot \vec{v} = 0$ , en dus is  $v_1^2 + v_2^2 + \dots + v_n^2 = 0$ , waaruit  $\vec{v} = \vec{0}$ .  $\square$

Beschouw nu een  $m \times n$ -matrix  $A$ . Dan is  $A^{tr}$  een  $n \times m$ -matrix, waarvoor  $\ker(A^{tr})$  een deelruimte is van  $\mathbb{R}^m$ . Ook  $\text{im}(A)$  is een deelruimte van  $\mathbb{R}^m$ . Wat is hun verband?

**Stelling 13.**  $\text{im}(A)^\perp = \ker(A^{tr})$ .

*Bewijs.* Druk de matrix  $A$  opnieuw uit aan de hand van haar kolommen als  $A = [\vec{a}_1 \dots \vec{a}_n]$ . Een vector  $\vec{y} \in \mathbb{R}^m$  is in  $\ker(A^{tr})$  als en slechts als

$$\vec{0} = A^{tr} \cdot \vec{y} = \begin{bmatrix} (\vec{a}_1)^{tr} \\ \vdots \\ (\vec{a}_n)^{tr} \end{bmatrix} \cdot \vec{y} = \begin{bmatrix} \vec{a}_1 \cdot \vec{y} \\ \vdots \\ \vec{a}_n \cdot \vec{y} \end{bmatrix},$$

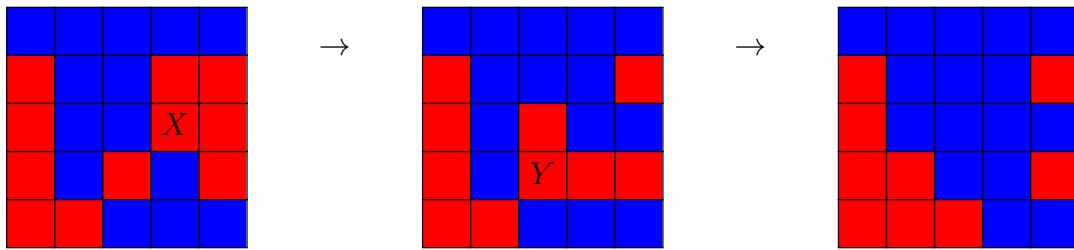
dus als en slechts als  $\vec{y}$  loodrecht staat op alle vectoren uit  $\text{span}(\vec{a}_1, \dots, \vec{a}_n) = \text{im}(A)$ .  $\square$

Een matrix is *symmetrisch* als  $A^{tr} = A$ . Voor een symmetrische matrix geldt dus  $\text{im}(A)^\perp = \ker(A)$ .

## 4 Lights out



”Lights Out”(zie bvb. deze [Lights Out Ad](#)) was een populair spel in de jaren negentig. Het bestaat uit een  $5 \times 5$  grid van al dan niet verlichte toetsen. Als je op een toets drukt verandert niet enkel deze toets van aan in uit (of omgekeerd) maar ook de aan/uit status van de 4 directe buren van de toets (boven, onder, links en rechts). Dus, als we vanuit de bovenstaande situatie drukken op de met een  $X$  aangegeven toets krijgen we de middelste situatie en als we vervolgens de  $Y$  toets indrukken krijgen we de situatie aan de rechterkant:



Het doel van het spel is vanuit een gegeven begin-positie uiteindelijk alle lichten uit te krijgen.

*Bepaal hoeveel legale begin-posities er zijn in Lights Out, dwz. posities die na eindig veel stappen alle lichten uit hebben.*

Eerst willen we een positie in 'Lights Out' op een handige manier coderen. Elke toets heeft twee mogelijke staten: aan (1) en uit (0). We hebben 25 verschillende toetsen, dus de staat van een positie kunnen we volledig beschrijven door een *vector* bestaande uit 25 *coordinaten* die een element zijn uit het *lichaam*  $\mathbb{F}_2$ . Als we de verschillende toetsen als volgt nummeren

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 \\ 21 & 22 & 23 & 24 & 25 \end{bmatrix}$$

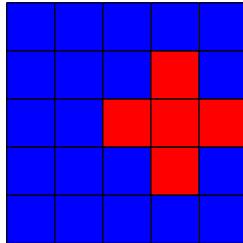
dan worden de drie posities boven weergegeven door de volgende (getransponeerde) *vectoren*:

$$\left\{ \begin{array}{ll} \text{links} & (0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0)^{tr} \\ \text{midden} & (0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0)^{tr} \\ \text{rechts} & (0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0)^{tr}. \end{array} \right.$$

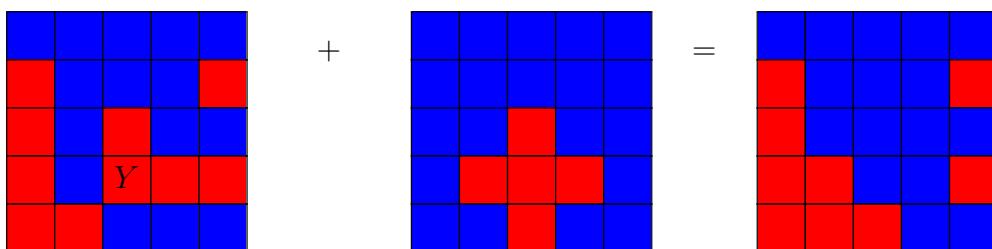
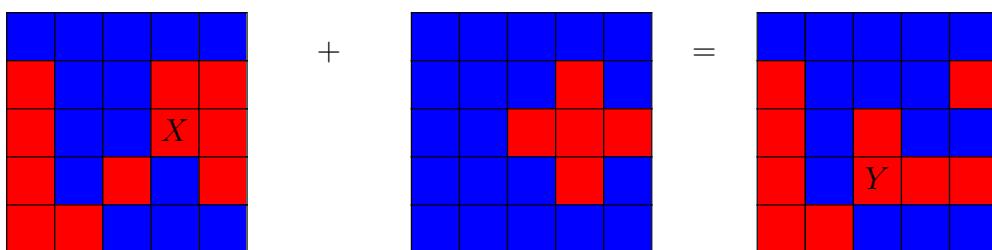
Twee vectoren kunnen we *optellen* door de coordinaten die op dezelfde plaats staan met elkaar op te tellen in  $\mathbb{F}_2$ . Dus, bvb.

$$\begin{aligned} & (0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0) \\ + & (0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0) \\ \hline & (0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0) \end{aligned}$$

Hier is de eerste vector de positie van de linkse positie. De tweede vector codeert de positie



Dit zijn juist de toetsen die van status veranderen als we drukken op de middelste rode knop. De som-vector codeert de middelste positie. Dus we zien dat, indien we een positie-vector optellen bij de vector die de status-verandering voor een bepaalde toets codeert, het resultaat de verkregen nieuwe positie codeert.



Omdat we 25 verschillende toetsen hebben, hebben we ook 25 van deze *statusverandering-*

vectoren:  $\{\vec{s}_1, \dots, \vec{s}_{25}\}$ . De  $25 \times 25$  matrix  $S$  met deze vectoren als kolommen is

$$S = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Zoals eerder vermeld noemen we een beginpositie, gecodeerd als de vector  $\vec{v}$  over  $\mathbb{F}_2$ , legaal indien we een opeenvolging van toetsindrukken kunnen vinden zodat alle lichten uit zijn. Merk op dat het einddoel, namelijk de 'alle lichten uit'-positie, als code de nulvector  $\vec{0}$  heeft. Dus als we na het indrukken van de toetsen  $i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_k$  alle lichten uitkrijgen, dan moet gelden dat

$$\vec{v} + \vec{s}_{i_1} + \vec{s}_{i_2} + \dots + \vec{s}_{i_k} = \vec{0}.$$

Omdat we werken over  $\mathbb{F}_2$ , en omdat daarin de optelling voldoet aan de regels  $0 + 0 = 0$  en  $1 + 1 = 0$ , weten we dat voor elke vector  $\vec{w}$  geldt dat  $\vec{w} + \vec{w} = \vec{0}$ . Bijgevolg kunnen we beide leden van bovenstaande gelijkheid optellen met  $\vec{v}$ . Het linkerlid wordt dan  $\vec{v} + \vec{v} + \vec{s}_{i_1} + \vec{s}_{i_2} + \dots + \vec{s}_{i_k} = \vec{0} + \vec{0} = \vec{0}$ .

$\dots + \vec{s}_{i_k} = \vec{s}_{i_1} + \vec{s}_{i_2} + \dots + \vec{s}_{i_k}$ . Het rechterlid wordt  $0 + \vec{v} = \vec{v}$ . We concluderen dus dat

$$\vec{s}_{i_1} + \vec{s}_{i_2} + \dots + \vec{s}_{i_k} = \vec{v}.$$

Een vector  $\vec{v}$  is dus legaal als en slechts als  $\vec{v}$  een vector is van de deelruimte opgespannen door de statusveranderingsvectoren  $\vec{s}_1, \dots, \vec{s}_{25}$ . Wederom omdat een vector bij zichzelf optellen de nul-vector levert mogen we dus ook veronderstellen dat alle  $\vec{s}_{i,j}$  in deze som verschillend zijn.

Om het aantal legale posities te bepalen moeten we dus de dimensie kennen van de deelruimte  $\text{span}(\vec{s}_1, \vec{s}_2, \dots, \vec{s}_{25})$  opgespannen door de vectoren  $\vec{s}_i$ . Deze deelruimte is echter  $\text{im}(S)$ , en haar dimensie wordt gegeven door de rang van  $S$ . Een kleine berekening leert dat de matrix  $\text{ref}(S)$  er uit ziet als:

Uit de twee nulrijen onderaan concluderen we dat  $\text{rank}(S) = 23$ .

In totaal zijn er  $2^{25}$  verschillende startposities. Niet allen leiden tot succes. Aangezien de deelruimte van legale startposities 23 is, zijn er  $2^{23}$  mogelijke legale startposities. Bijgevolg is slechts  $1/4$  van alle mogelijke posities een legale positie. Verder volgt uit onze redenering dat je elke legale positie kunt uitdoven in ten hoogste 23 zetten.

Kunnen we een methode ontwikkelen om die 23 zetten te weten te komen?

Hiervoor beschouwen we eerst  $\ker(S)$ . Aangezien  $S$  symmetrisch is weten we uit Stelling 13 dat  $\text{im}(S) = (\ker(S))^\perp$ . Noem  $R$  de inverteerbare matrix waarvoor  $\text{ref}(S) = R.S$ . Er geldt  $\ker(S) = \ker(\text{ref}(S))$  (zie net voor de Stelling 11). Via de dimensiestelling weten we dan:

$$\begin{aligned}\dim(\ker(S)) &= \dim(\ker(\text{ref}(S))) = 25 - \dim(\text{im}(\text{ref}(S))) \\ &= 25 - \text{rank}(\text{im}(\text{ref}(S))) = 25 - 23 = 2.\end{aligned}$$

We kunnen hieruit concluderen dat een basis voor  $\ker(S)$  steeds uit 2 vectoren zal bestaan.

De  $(25 \times 25)$  matrix  $\text{ref}(S)$  is van de vorm  $\begin{bmatrix} I_{23} & \vec{k}_1 & \vec{k}_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ . Hierin zijn  $\vec{k}_1$  en  $\vec{k}_2$  twee kolommen met 23 elementen. Als  $\vec{x} = (\vec{y} \ a \ b)^{tr} \in \ker(\text{ref}(S))$  (met opnieuw  $\vec{y}$  een kolom met 23 elementen), dan is

$$\begin{bmatrix} I_{23} & \vec{k}_1 & \vec{k}_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \vec{y} \\ a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

en dus  $\vec{y} + a\vec{k}_1 + b\vec{k}_2 = 0$ , waaruit  $\vec{y} = a\vec{k}_1 + b\vec{k}_2$ .

Stel nu

$$\vec{n}_1 = \begin{bmatrix} \vec{y} \\ a \\ b \end{bmatrix} = \begin{bmatrix} \vec{k}_1 \\ 1 \\ 0 \end{bmatrix} \quad \text{en} \quad \vec{n}_2 = \begin{bmatrix} \vec{y} \\ a \\ b \end{bmatrix} = \begin{bmatrix} \vec{k}_2 \\ 0 \\ 1 \end{bmatrix},$$

dan is  $\{\vec{n}_1, \vec{n}_2\}$  een basis voor  $\ker(S)$ . Deze basis is orthogonaal, want  $\vec{n}_1 \cdot \vec{n}_2 = 0$ .

We zagen dat een beginpositie  $\vec{v}$  legaal was als  $\vec{v} \in \text{span}(\{\vec{s}_1, \vec{s}_2, \dots, \vec{s}_{25}\}) = \text{im}(S)$ . Maar,  $\text{im}(S) = (\ker(S))^\perp$ . Samen met het voorgaande kunnen we dus concluderen:

**Stelling 14.** *Een positie  $\vec{v}$  is legaal/winnable indien  $\vec{v} \cdot \vec{n}_1 = 0$  en  $\vec{v} \cdot \vec{n}_2 = 0$ .*

Een tweede karakterisatie van legale posities gebruikt de definitie van  $\text{im}(S)$ . Een vector  $\vec{v}$  ligt in  $\text{im}(S)$  als er een corresponderende vector  $\vec{x}$  bestaat waarvoor  $S.\vec{x} = \vec{v}$ . Een dergelijke  $\vec{x}$  noemen we een *strategie*. Zoals alle vectoren in  $\mathbb{F}_2$  bestaat  $\vec{x}$  uit 1-en en 0-en. De positie

van de 1-en in een strategie  $\vec{x}$  geven aan op welke toetsen je dient te drukken opdat het spel zou eindigen.

Een voorbeeld maakt dit duidelijk. Voor  $\vec{x} = (1 \ 0 \ 0 \ \dots \ 0)^{tr}$  geldt

$$S \cdot \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \vec{s}_1 & \vec{s}_2 & \dots & \vec{s}_{25} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \vec{s}_1.$$

Analoog zal men voor  $\vec{x} = (1 \ 1 \ 0 \ \dots \ 0)^{tr}$  de legale positie  $\vec{s}_1 + \vec{s}_2$  krijgen, etc.

Wat zijn de strategieën die bij een gegeven  $\vec{v}$  horen? De strategieën zijn niet uniek bepaald. Als  $\vec{x}$  een strategie is voor  $\vec{v}$ , dan ook  $\vec{x} + \vec{n}_1$ ,  $\vec{x} + \vec{n}_2$  en  $\vec{x} + \vec{n}_1 + \vec{n}_2$ , want, bijvoorbeeld:

$$S(\vec{x} + \vec{n}_1 + \vec{n}_2) = S\vec{x} + S\vec{n}_1 + S\vec{n}_2 = \vec{v}.$$

Wat  $\vec{x}$  ook is, de vier vectoren  $\vec{x}$ ,  $\vec{x} + \vec{n}_1$ ,  $\vec{x} + \vec{n}_2$  en  $\vec{x} + \vec{n}_1 + \vec{n}_2$  bestaan elk uit 25 componenten. Merk op dat de getallen op de laatste twee rijen steeds 0 of 1 moeten zijn. Aangezien er slechts vier mogelijkheden zijn heeft één van de vier vectoren  $\vec{x}$ ,  $\vec{x} + \vec{n}_1$ ,  $\vec{x} + \vec{n}_2$  of  $\vec{x} + \vec{n}_1 + \vec{n}_2$  als laatste twee cijfers 0 en 0. Noem die vector  $\vec{x}_0$ .

**Stelling 15.** *Zij  $\vec{v}$  een legale positie. Dan worden de enige strategieën voor  $\vec{v}$  gegeven door:*

$$R.\vec{v}, \quad R.\vec{v} + \vec{n}_1, \quad R.\vec{v} + \vec{n}_2, \quad \text{en} \quad R.\vec{v} + \vec{n}_1 + \vec{n}_2.$$

*Bewijs.* Wegens de vorm van  $\text{ref}(S) = \begin{bmatrix} I_{23} & \vec{k}_1 & \vec{k}_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$  en de laatste twee nullen in  $\vec{x}_0$  geldt:

$$\text{ref}(S).\vec{x}_0 = \vec{x}_0,$$

en dus is

$$R.S.\vec{x}_0 = \vec{x}_0.$$

Aangezien  $S.\vec{x}_0 = \vec{v}$  vinden we dat  $R.\vec{v} = \vec{x}_0$ .

We kunnen dus concluderen dat één van de vier strategieën gelijk is aan  $\vec{x}_0 = R.\vec{v}$ . De andere drie zijn dan  $R.\vec{v} + \vec{n}_1$ ,  $R.\vec{v} + \vec{n}_2$  en  $R.\vec{v} + \vec{n}_1 + \vec{n}_2$ .  $\square$

Hoe spelen we Lights Out? Veronderstel dat de beginpositie van het spel gegeven wordt door  $\vec{v}$ . Bereken dan  $\vec{x} = R\vec{v}$ , of één van de andere drie strategieën. Dit is een vector met

minder dan 23 enen. Druk de corresponderende toetsen in. Indien hierna de lichten uit zijn, dan was  $\vec{v}$  legaal. Indien niet, dan was  $\vec{v}$  geen legale positie en heeft verder spelen geen zin. Naast deze methode, met ten hoogste 23 zetten, bestaan er nog snellere speelwijzen. Je kan wat meer info hierover vinden in het artikel "M. Anderson and T. Feil, Turning lights out with linear algebra, *Mathematics Magazine* 71 (1998), 300-303".

## 5 De kleinste-kwadratenmethode

Beschouw een deelruimte  $V$  van  $\mathbb{R}^n$ . Als  $\{\vec{v}_1, \dots, \vec{v}_k\}$  een orthonormale basis is van  $V$ , dan is de *orthogonale projectie map* de afbeelding van  $\mathbb{R}^n$  naar  $V$  gegeven door

$$\vec{x} \in \mathbb{R}^n \mapsto \text{proj}_V(\vec{x}) = (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \dots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k \in V.$$

Omdat het inproduct lineair is in beide termen volgt dat, voor elke  $\lambda, \mu \in \mathbb{R}$ ,

$$\text{proj}_V(\lambda\vec{v} + \mu\vec{w}) = \lambda\text{proj}_V(\vec{v}) + \mu\text{proj}_V(\vec{w}).$$

**Lemma 6.** Voor elke  $\vec{x} \in \mathbb{R}^n$  geldt  $\vec{x} - \text{proj}_V(\vec{x}) \in V^\perp$ .

*Bewijs.* Neem een  $\vec{w} = w_1\vec{v}_1 + \dots + w_k\vec{v}_k \in V$ , dan

$$\begin{aligned} \vec{w} \cdot (\vec{x} - \text{proj}_V(\vec{x})) &= (w_1\vec{v}_1 + \dots + w_k\vec{v}_k) \cdot \vec{x} - ((\vec{x} \cdot \vec{v}_1)(\vec{w} \cdot \vec{v}_1) + \dots + (\vec{x} \cdot \vec{v}_k)(\vec{w} \cdot \vec{v}_k)) \\ &= w_1\vec{v}_1 \cdot \vec{x} + \dots + w_k\vec{v}_k \cdot \vec{x} - (w_1\vec{x} \cdot \vec{v}_1 + \dots + w_k\vec{x} \cdot \vec{v}_k) = 0. \end{aligned}$$

In de laatste stap hebben we gebruikt dat, voor een orthonormale basis,  $\vec{w} \cdot \vec{v}_i = w_i$ . □

We geven een toepassing waarin een dergelijke projectie aan bod komt.

Beschouw een lineair stelsel van  $m$  vergelijkingen in  $n$  onbekenden,  $A.\vec{x} = \vec{b}$ . Soms is het niet mogelijk om een oplossing te vinden, bijvoorbeeld als  $\vec{b} \notin \text{im}(A)$ . De *kleinste-kwadraatoplossing* van het stelsel is die vector  $\vec{x}^* \in \mathbb{R}^n$  zodat de vector  $A\vec{x}^* \in \mathbb{R}^m$  op minimale afstand ligt van  $\vec{b}$ , tussen alle mogelijke vectoren van de vorm  $A\vec{x}$ .

Omdat  $\text{proj}_{\text{im}(A)}(\vec{b})$  de vector in  $\text{im}(A)$  is op minimale afstand van  $\vec{b}$  zoeken we dus een oplossing  $\vec{x}^*$  van het stelsel van  $m$  lineaire vergelijkingen in  $n$  onbekenden

$$A.\vec{x}^* = \text{proj}_{\text{Im}(A)}(\vec{b}).$$

Via Lemma 6 weten we dat  $\vec{b} - \text{proj}_{\text{im}(A)}(\vec{b}) \in \text{im}(A)^\perp$ . Het is dus via Stelling 13 ook een vector van  $\ker(A^{tr})$ . De kleinste-kwadraatoplossing  $\vec{x}^*$  voldoet hierdoor aan

$$A^{tr}.(\vec{b} - A.\vec{x}^*) = \vec{0}$$

en is dus een oplossing van het stelsel

$$A^{tr}.A.\vec{x}^* = A^{tr}.\vec{b}.$$

Dit laatste stelsel noemen we het *normale stelsel* voor het oorspronkelijke stelsel  $A.\vec{x} = \vec{b}$ . Merk op dat  $A^{tr}.A$  een vierkante matrix van orde  $n$  is (ook al is  $A$  zelf niet vierkant). We gaan na wanneer hij inverteerbaar is.

**Stelling 16.**  $\ker(A) = \ker(A^{tr}.A)$ .

*Bewijs.* Als  $\vec{x} \in \ker(A)$  dan  $A.\vec{x} = \vec{0}$  en dus ook  $A^{tr}.A.\vec{x} = \vec{0}$  dus  $\ker(A) \subset \ker(A^{tr}.A)$ . Omgekeerd, als  $\vec{x} \in \ker(A^{tr}.A)$  dan is  $A^{tr}.(A.\vec{x}) = \vec{0}$  en dus is  $A.\vec{x} \in \ker(A^{tr}) = \text{im}(A)^\perp$  (zie Stelling 13). Dus is

$$A.\vec{x} \in \text{im}(A) \cap \text{im}(A)^\perp,$$

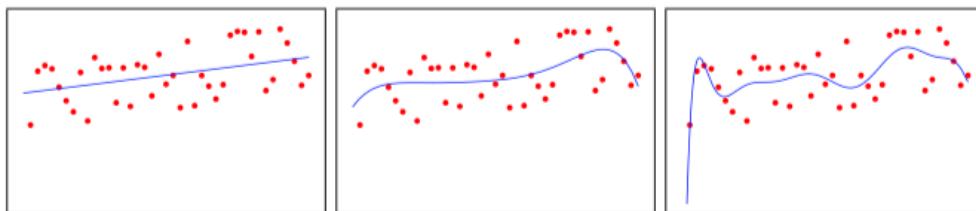
en dus is (wegens Lemma 5)  $A.\vec{x} = \vec{0}$ . Bijgevolg is  $\vec{x} \in \ker(A)$ .  $\square$

**Stelling 17.** Als  $\ker(A) = \vec{0}$  dan wordt de kleinste-kwadraat oplossing voor  $A.\vec{x} = \vec{b}$  gegeven door

$$\vec{x}^* = (A^{tr}.A)^{-1}.A^{tr}.\vec{b}.$$

*Bewijs.* In dit geval is  $\ker(A^{tr}.A) = \ker(A) = \vec{0}$ , met  $A^{tr}.A$  een  $n \times n$  matrix is. Dit betekent dat het lineaire stelsel  $(A^{tr}.A).\vec{x} = \vec{0}$  de unieke oplossing  $\vec{0}$  heeft. Wegens Stelling 7 (Eigenschap 5, met  $\vec{b} = \vec{0}$ ) is  $A^{tr}.A$  dan een inverteerbare  $n \times n$  matrix. Bijgevolg kunnen we beide leden van de normale vergelijking linksvermenigvuldigen met  $(A^{tr}.A)^{-1}$ . We vinden dan de unieke oplossing  $\vec{x}^* = (A^{tr}.A)^{-1}.A^{tr}.\vec{b}$ .  $\square$

Deze kleinste-kwadratenmethode wordt, bijvoorbeeld, gebruikt in de context van *data fitting*. Hieronder staat een voorbeeld met 30 datapunten. De blauwe krommen zijn de polynomen van graad respectievelijk 1, 6 en 11 die de beste benadering van de datapunten geven.



We werken een eenvoudiger versie in detail uit. *Vind het quadratische polynoom  $y = ax^2 + bx + c$  dat de beste benadering geeft voor de data-punten  $(-1, 8), (0, 8), (1, 4)$  en  $(2, 16)$ .*

Laat  $A$  de matrix zijn waarvan de rijen de waarden  $(x^2, x^1, x^0)$  geven van de 4  $x$ -waarden, dan zoeken we de kleinste-kwadraat oplossing voor het lineaire stelsel

$$A \cdot \vec{x} = \vec{b} \quad \text{met} \quad A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \quad \text{en} \quad \vec{b} = \begin{bmatrix} 8 \\ 8 \\ 4 \\ 16 \end{bmatrix}.$$

We berekenen dat

$$A^{tr} \cdot A = \begin{bmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{bmatrix} \quad \text{en dus} \quad \vec{x}^* = (A^{tr} \cdot A)^{-1} \cdot A \cdot \vec{b} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}.$$

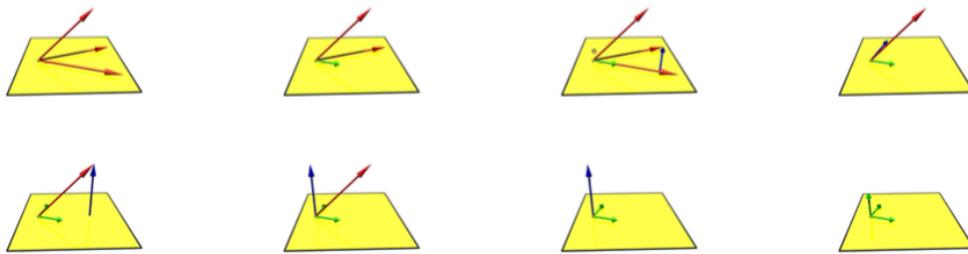
Dit is een inverteerbare matrix en men vindt makkelijk dat  $y = 3x^2 - x + 5$  de beste data-fit geeft.

## 6 De Gram-Schmidt-procedure

In deze paragraaf geven we een manier om een orthonormale basis te construeren, startend van een gegeven basis.

**Stelling 18** (Gram-Schmidt). (*Over de reële getallen  $\mathbb{R}$* ) Er is een inductieve procedure om uit een gegeven basis  $\{\vec{v}_1, \dots, \vec{v}_n\}$  van  $\mathbb{R}^n$  een orthonormale basis  $\{\vec{w}_1, \dots, \vec{w}_n\}$  te construeren waarbij  $\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}$ .

*Bewijs.* Hieronder een grafische voorstelling van het proces in  $\mathbb{R}^3$



We geven een bewijs dat werkt op inductie. Voor de eerste stap in de inductie beschouwen we  $\vec{u}_1 = \vec{v}_1$ , en dan  $\vec{w}_1 = \vec{u}_1 / \|\vec{u}_1\|$ .

Veronderstel nu dat we al een orthonormaal stel vectoren  $\{\vec{w}_1, \dots, \vec{w}_{i-1}\}$  geconstrueerd hebben, voor zekere  $i$ , en laat  $V_{i-1} = \text{span}(\vec{w}_1, \dots, \vec{w}_{i-1})$  de deelruimte zijn, opgespannen door deze vectoren. Definiëer nu de 'volgende' vector

$$\vec{u}_i = \vec{v}_i - \text{proj}_{V_{i-1}}(\vec{v}_i).$$

Via Lemma 6 weten we dat  $\vec{u}_i \in V_{i-1}^\perp$ , en via Lemma 5 ligt deze vector zeker niet in  $V_{i-1}$ .

Normaliseer nu deze vector als

$$\vec{w}_i = \frac{\vec{u}_i}{\|\vec{u}_i\|},$$

dan is nog steeds  $\vec{w}_i \in V_{i-1}^\perp$  en vormen  $\{\vec{w}_1, \dots, \vec{w}_i\}$  een orthonormaal stel vectoren in  $\mathbb{R}^n$ .

Na eindig veel stappen hebben we dus de gevraagde basis  $\{\vec{w}_1, \dots, \vec{w}_n\}$ .  $\square$

We kunnen het Gram-Schmidt proces ook uitdrukken in matrix-vorm.

**Stelling 19** (QR-factorizatie). (*Over de reële getallen*) Iedere inverteerbare  $n \times n$  matrix  $A$  kan geschreven worden als

$$A = Q.R$$

waarbij de kolommen van  $Q$  een orthonormaal stel vectoren voor  $\mathbb{R}^n$  geven en  $R$  een boven-driehoeksmatrix is.

*Bewijs.* Uit het vorige bewijs halen we

$$\left\{ \begin{array}{l} \vec{v}_1 = \|\vec{v}_1\| \vec{w}_1 = r_{11} \vec{w}_1, \\ \vec{v}_2 = (\vec{v}_2 \cdot \vec{w}_1) \vec{w}_1 + \|\vec{v}_2\| \vec{w}_2 = r_{12} \vec{w}_1 + r_{22} \vec{w}_2, \\ \vdots \\ \vec{v}_i = (\vec{v}_i \cdot \vec{w}_1) \vec{w}_1 + \dots + (\vec{v}_i \cdot \vec{w}_{i-1}) \vec{w}_{i-1} + \|\vec{v}_i\| \vec{w}_i = r_{1i} \vec{w}_1 + \dots + r_{ii} \vec{w}_i, \\ \vdots \\ \vec{v}_n = (\vec{v}_n \cdot \vec{w}_1) \vec{w}_1 + \dots + (\vec{v}_n \cdot \vec{w}_{n-1}) \vec{w}_{n-1} + \|\vec{v}_n\| \vec{w}_n = r_{1n} \vec{w}_1 + \dots + r_{nn} \vec{w}_n, \end{array} \right.$$

en dit kunnen we ook in matrix-vorm schrijven

$$A = [\vec{v}_1 \ \dots \ \vec{v}_n] = [\vec{w}_1 \ \dots \ \vec{w}_n] \cdot \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{bmatrix} = Q.R.$$

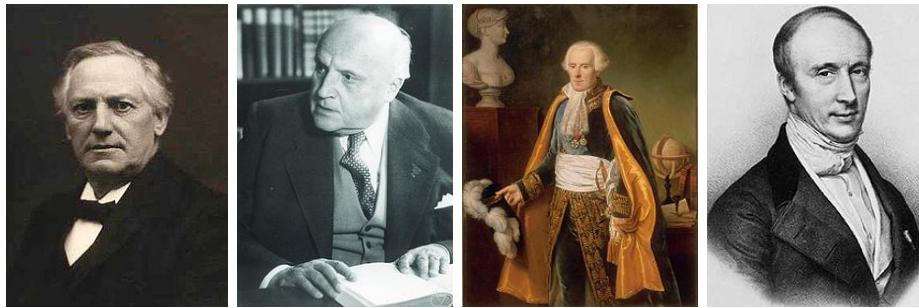
$\square$

**Definitie 9.** Een  $n \times n$  matrix  $Q$  noemen we orthogonaal indien  $Q^{tr} \cdot Q = I_n$ .

Deze eigenschap is equivalent met zeggen dat, net zoals in de stelling, de kolommen van  $Q$  een orthonormaal stel vectoren vormen. Inderdaad,

$$Q^{tr} \cdot Q = \begin{bmatrix} (\vec{w}_1)^{tr} \\ \vdots \\ (\vec{w}_n)^{tr} \end{bmatrix} \cdot \begin{bmatrix} \vec{w}_1 & \dots & \vec{w}_n \end{bmatrix} = \begin{bmatrix} \vec{w}_1 \cdot \vec{w}_1 & \dots & \vec{w}_1 \cdot \vec{w}_n \\ \vdots & & \vdots \\ \vec{w}_n \cdot \vec{w}_1 & \dots & \vec{w}_n \cdot \vec{w}_n \end{bmatrix} = I_n.$$

**Historische noot.** De recursieve formule werd geformuleerd door [Erhard Schmidt](#) (1876-1959) (tweede van links) in 1907. De essentie van de formule stond al in een paper uit 1883 van [Joergen Pedersen Gram](#) (1850-1916) (links) dat Schmid in een voetnoot vermeldt. Eigenlijk werd de procedure al gebruikt door [Pierre Simon Laplace](#) (1749-1827) (tweede van rechts) en door [Augustin Louis Cauchy](#) (rechts) in 1836.



**Een uitgewerkt voorbeeld.** Beschouw de matrix

$$A = \begin{bmatrix} 2 & 1 & 3 & 3 \\ 2 & 1 & -1 & 1 \\ 2 & -1 & 3 & -3 \\ 2 & -1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 \end{bmatrix}.$$

Aangezien  $\text{rank}(A) = 4$  is  $A$  inverteerbaar (Stelling 7) en dus is  $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$  een basis voor  $\mathbb{R}^4$  (Stelling 9). We zoeken nu een orthonormale basis voor  $\mathbb{R}^4$ .

$$\text{Stel } \vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix}. \text{ Dan is } \|\vec{u}_1\| = 4 \text{ en dus } \vec{w}_1 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Stel  $V_1 = \text{span}(\vec{v}_1) = \text{span}(\vec{w}_1)$ . Er geldt:

$$\vec{v}_2 \cdot \vec{w}_1 = \begin{bmatrix} 1 & 1 & -1 & -1 \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 0$$

en dus

$$\begin{aligned}\vec{u}_2 &= \vec{v}_2 - \text{proj}_{V_1}(\vec{v}_2) \\ &= \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} - (\vec{v}_2 \cdot \vec{w}_1) \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}.\end{aligned}$$

Aangezien  $\|\vec{u}_2\| = 2$ , is  $\vec{w}_2 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$ .

Stel nu  $V_2 = \text{span}(\vec{v}_1, \vec{v}_2) = \text{span}(\vec{w}_1, \vec{w}_2)$ . Er geldt:

$$\begin{aligned}\vec{v}_3 \cdot \vec{w}_1 &= \begin{bmatrix} 3 & -1 & 3 & -1 \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 2, \\ \vec{v}_3 \cdot \vec{w}_2 &= \begin{bmatrix} 3 & -1 & 3 & -1 \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} = 0,\end{aligned}$$

en dus

$$\begin{aligned}\vec{u}_3 &= \vec{v}_3 - \text{proj}_{V_2}(\vec{v}_3) \\ &= \begin{bmatrix} 3 \\ -1 \\ 3 \\ -1 \end{bmatrix} - (\vec{v}_3 \cdot \vec{w}_1) \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - (\vec{v}_3 \cdot \vec{w}_2) \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 2 \\ -2 \end{bmatrix}.\end{aligned}$$

Aangezien  $\|\vec{u}_3\| = 4$ , is  $\vec{w}_3 = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$ .

Stel nu  $V_3 = \text{span}(\vec{v}_1, \vec{v}_2, \vec{v}_3) = \text{span}(\vec{w}_1, \vec{w}_2, \vec{w}_3)$ . Er geldt:

$$\vec{v}_4 \cdot \vec{w}_1 = 0, \quad \vec{v}_4 \cdot \vec{w}_2 = 4, \quad \vec{v}_4 \cdot \vec{w}_3 = 0$$

en dus

$$\begin{aligned}
 \vec{u}_4 &= \vec{v}_4 - \text{proj}_{V_3}(\vec{v}_4) \\
 &= \begin{bmatrix} 3 \\ 1 \\ -3 \\ -1 \end{bmatrix} - (\vec{v}_4 \cdot \vec{w}_1) \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - (\vec{v}_4 \cdot \vec{w}_2) \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} - (\vec{v}_4 \cdot \vec{w}_3) \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \\
 &= \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.
 \end{aligned}$$

Aangezien  $\|\vec{u}_4\| = 2$ , is  $\vec{w}_4 = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$ . De 4 vectoren  $\vec{w}_1, \vec{w}_2, \vec{w}_3$  en  $\vec{w}_4$  vormen de gezochte orthonormale basis.

We gaan nog even de QR-decompositie na. Men gaat makkelijk na dat  $A = QR$  met

$$Q = [\vec{w}_1 \quad \vec{w}_2 \quad \vec{w}_3 \quad \vec{w}_4] = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix},$$

en (zie de uitdrukkingen in de stelling)

$$R = \begin{bmatrix} 4 & 0 & 2 & 0 \\ 0 & 2 & 0 & 4 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Wat is nu eigenlijk het voordeel van een QR-decompositie bij het oplossen van een stelsel

lineaire vergelijkingen? Stel je voor dat we de oplossingen  $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$  van de vergelijkingen

$$A.\vec{x} = \vec{b}, \text{ met } \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}.$$

Omdat we weten dat  $Q^{-1} = Q^{tr}$  is dit ook de oplossing van  $R.\vec{x} = Q^{tr}.\vec{b}$  of

$$\begin{bmatrix} 4 & 0 & 2 & 0 \\ 0 & 2 & 0 & 4 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -1 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} 4x_1 + 2x_3 = 2, \\ 2x_2 + 4x_4 = 1, \\ 4x_3 = -1, \\ 2x_4 = 0, \end{cases}$$

waaruit, via achterwaartse substitutie, makkelijk de oplossing volgt,

$$x_4 = 0, \quad x_3 = -\frac{1}{4}, \quad x_2 = \frac{1}{2}, \quad x_1 = \frac{5}{8}.$$

We hebben via deze methode dus nergens de inverse van een matrix moeten berekenen.

# 4 Determinanten

## 1 Definitie en eigenschappen

Beschouw een vierkante matrix  $A$ ,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

**Definitie 10.** De  $(k, l)$ -minor van  $A$  is de  $(n - 1) \times (n - 1)$  matrix  $A_{kl}$  die men bekomt door uit  $A$  de  $k$ -de rij en de  $l$ -de kolom te schrappen.

De determinant wordt inductief gedefinieerd, als volgt:

- De determinant  $\det(A)$  van een  $1 \times 1$  matrix  $A = (a)$  is  $a$  zelf.
- De determinant  $\det(A)$  van een  $n \times n$  matrix  $A$  is de som

$$\begin{aligned} \det(A) &= a_{11} \det(A_{11}) - a_{21} \det(A_{21}) + \dots + (-1)^{i+n} a_{ni} \det(A_{ni}) \\ &= \sum_{i=1}^n (-1)^{1+i} a_{i1} \det(A_{i1}). \end{aligned}$$

Het rechterlid in de definitie heet *de ontwikkeling naar de eerste kolom*. We verduidelijken die benaming later nog.

Voor kleine waarden van  $n$  kan men makkelijk nagaan dat:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc, \quad \det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = aei + bfg + cdh - ceg - fha - bdi.$$

Het is uit de definitie ook onmiddellijk duidelijk dat *de determinant van een diagonalmatrix gelijk is aan het product van de diagonalelementen*. In het bijzonder is  $\det(I_n) = 1$ .

**Stelling 20.** *De determinant is multi-lineair en alternerend in de rijen. Dit wil zeggen dat*

$$\det \begin{bmatrix} R_1 \\ \vdots \\ \lambda R_k \\ \vdots \\ R_n \end{bmatrix} = \lambda \det \begin{bmatrix} R_1 \\ \vdots \\ R_k \\ \vdots \\ R_n \end{bmatrix}, \quad \det \begin{bmatrix} R_1 \\ \vdots \\ R_k + S \\ \vdots \\ R_n \end{bmatrix} = \det \begin{bmatrix} R_1 \\ \vdots \\ R_k \\ \vdots \\ R_n \end{bmatrix} + \det \begin{bmatrix} R_1 \\ \vdots \\ S \\ \vdots \\ R_n \end{bmatrix},$$

$$\det \begin{bmatrix} R_1 \\ \vdots \\ R_l \\ \vdots \\ R_k \\ \vdots \\ R_n \end{bmatrix} = - \det \begin{bmatrix} R_1 \\ \vdots \\ R_k \\ \vdots \\ R_l \\ \vdots \\ R_n \end{bmatrix},$$

waarbij we in de laatste lijn enkel de rijen  $l$  en  $k$  verwisseld hebben.

*Bewijs.* We zullen deze eigenschappen bewijzen door inductie op  $n$ . Voor het geval waarbij  $n = 1$  zijn de eigenschappen triviaal voldaan: Bijvoorbeeld,  $\det(\lambda a) = \lambda a = \lambda \det(a)$ , etc. We nemen dus aan dat de eigenschappen voldaan zijn voor  $(n - 1) \times (n - 1)$ -matrices en tonen aan dat hieruit ook de eigenschap volgt voor  $n \times n$ -matrices.

Voor de eerste eigenschap, noem  $B$  de matrix met  $k$ -de rij  $\lambda R_k$  en  $A$  de matrix met  $k$ -de rij  $R_k$ . We weten dat

$$\det(A) = \sum_{i=1}^n (-1)^{1+i} a_{i1} \det(A_{i1}) \quad \text{en} \quad \det(B) = \sum_{i=1}^n (-1)^{1+i} b_{i1} \det(B_{i1}).$$

We beschouwen nu de  $i$ -de term in beide sommen. We maken een onderscheid tussen de term met net  $i = k$ , en alle andere termen met dus  $i \neq k$ . Als  $i = k$ , dan is  $b_{k1} = \lambda a_{k1}$ , maar, aangezien  $A$  en  $B$  enkel in de  $k$ -de rij verschillen, en we die net schrappen bij de minor is  $A_{k1} = B_{k1}$ . Samen hebben we dat  $(-1)^{1+k} b_{1k} \det(B_{1k}) = \lambda (-1)^{1+k} b_{1k} \det(A_{1k})$ .

Beschouw nu een  $i$ -de term in beide sommen, met nu  $i \neq k$ . Dan is  $b_{i1} = a_{i1}$ . Maar,  $B_{i1}$  is nu een matrix wiens  $k$ -de rij een factor  $\lambda$  verschilt met de  $k$ -de rij van die van  $A_{i1}$ . Aangezien het nu gaat over een matrix van orde  $n - 1$  volgt uit de inductiehypothese dat  $\det(B_{i1}) = \lambda \det(A_{i1})$ . Samen hebben we opnieuw  $(-1)^{1+i} b_{1i} \det(B_{1i}) = \lambda (-1)^{1+i} b_{1i} \det(A_{1i})$ .

Bijgevolg kunnen we in *alle* termen van het rechterlid van  $\det B$  een factor  $\lambda$  voorop plaatsen. Dit bewijst de eerste eigenschap.

Voor de tweede eigenschap noemen we  $C$  de matrix met  $k$ -de rij  $R_k + S$ ,  $A$  degene met rij  $R_k$  en  $B$  degene met rij  $S$ . Uit de definitie van de determinant zien we dat de eigenschap zal volgen als we voor elke term van de verschillende sommen kunnen aantonen dat

$$c_{i1} \det C_{i1} = a_{i1} \det A_{i1} + b_{i1} \det B_{i1}, \quad \forall i = 1, \dots, n.$$

Veronderstel eerst dat  $i = k$ . Dan is duidelijk  $c_{k1} = a_{k1} + b_{k1}$ . Tevens geldt (aangezien we telkens de  $k$ -de rij schrapten):  $C_{k1} = A_{k1} = B_{k1}$ , waaruit de bovenstaande eigenschap volgt.

Veronderstel nu dat  $i \neq k$ . In dit geval is  $c_{i1} = a_{i1} = b_{i1}$ . Maar, via de inductiehypothese weten we dat voor de corresponderende  $(n-1) \times (n-1)$ -matrices geldt:  $\det C_{i1} = \det A_{i1} + \det B_{i1}$  (want ze zijn gelijk op één rij na, en net die rij heeft de gepaste eigenschap). Opnieuw volgt de gezochte eigenschap.

Het bewijs van de derde eigenschap volgt een analoge (maar langere) redenering als die van de eerste twee eigenschappen. We laten het bewijs hier weg.  $\square$

Uit de eerste eigenschap volgt, met  $\lambda = 0$ , onmiddellijk dat *een vierkante matrix met een nulrij determinant nul heeft*.

Merk op dat bij de eerste eigenschappen de twee matrices in de som op één rij na dezelfde rijen moeten hebben opdat de eigenschap geldig zou zijn. Je kan hieruit geen eigenschap afleiden voor  $\det(A + B)$  van twee willekeurige matrices  $A$  en  $B$ : in het algemeen is  $\det(A + B) \neq \det(A) + \det(B)$ .

Uit de laatste eigenschap volgt ook onmiddellijk dat *een matrix met twee gelijke rijen determinant nul heeft*.

Als we de rij echelon vorm  $\text{ref}(A)$  van  $A$  kennen kunnen we ook hun determinanten vergelijken.

**Stelling 21.** *Als  $A$  een  $n \times n$  matrix is en als in het Gauss-Jordan algoritme we  $s$  rijverwisselingen nodig hadden en  $\lambda_1, \dots, \lambda_k$  de factoren zijn waarmee we verschillende rijen moesten vermenigvuldigen om een leidende 1 te krijgen, dan geldt*

$$\det(\text{ref}(A)) = (-1)^s \lambda_1 \lambda_2 \dots \lambda_k \det(A).$$

*Een direct gevolg is dat  $A$  inverteerbaar is als en slechts als  $\det(A) \neq 0$ .*

*Bewijs.* In het Gauss-Jordan algoritme zijn er drie elementaire bewerkingen. Ofwel wisselen we twee rijen en dit verandert de determinant met een factor  $-1$ , ofwel vermenigvuldigen

we een rij met een factor  $\lambda$  en dan verandert ook de determinant met die factor, ofwel tellen we een veelvoud van een rij op bij een andere rij. In dit geval gebruiken we stelling 20:

$$\det \begin{bmatrix} R_1 \\ \vdots \\ R_i + \lambda R_j \\ \vdots \\ R_j \\ \vdots \\ R_n \end{bmatrix} = \det \begin{bmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_j \\ \vdots \\ R_n \end{bmatrix} + \lambda \det \begin{bmatrix} R_1 \\ \vdots \\ R_j \\ \vdots \\ R_n \end{bmatrix} = \det \begin{bmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_j \\ \vdots \\ R_n \end{bmatrix},$$

omdat de determinant van een matrix met twee gelijke rijen nul is.

Wanneer  $A$  inverteerbaar is dan is  $\text{rank}(A) = n$  en dus  $\text{ref}(A) = I_n$ . Aangezien  $\det(I_n) = 1$ , kunnen de factoren in het rechterlid van bovenstaande uitdrukking niet nul zijn, dus  $\det(A) \neq 0$ .

Omgekeerd, veronderstel dat  $\det(A) \neq 0$ . We bewijzen uit het ongerijmde en veronderstellen eerst dat ook  $A$  niet inverteerbaar zou zijn. Maar dan is ook  $\text{rank}(A) \neq n$ , wat betekent dat  $\text{ref}(A)$  een nulrij heeft. Dat zou betekenen dat  $\det(\text{ref}(A)) = 0$ , wat op zijn beurt strijdig is met  $\det(A) \neq 0$ . We concluderen dus dat  $A$  wel degelijk inverteerbaar is in het geval dat  $\det(A) \neq 0$ .  $\square$

**Stelling 22.** *Het stelsel  $A.\vec{x} = \vec{b}$  heeft een unieke oplossing als en slechts als  $\det A \neq 0$ .*

*Bewijs.* De voorwaarde in de stelling betekent dat  $A$  inverteerbaar is, wat betekent dat noodzakelijk  $\vec{x} = A^{-1}.\vec{b}$  de enige oplossing van het stelsel is.  $\square$

In de volgende paragraaf geven we in Stelling 27 nog een bijkomende methode om deze oplossing effectief te berekenen.

**Stelling 23.** *Als  $A$  en  $B$  twee  $n \times n$  matrices zijn dan geldt*

$$\det(A \cdot B) = \det(A) \det(B).$$

*Bewijs.* Als  $A$  of  $B$  niet inverteerbaar is dan geldt hetzelfde voor  $A \cdot B$  en volgt de gelijkheid dus uit vorige stelling. We mogen dus veronderstellen dat zowel  $A$  als  $B$  inverteerbaar zijn (beide  $\det A \neq 0 \neq \det B$ ). Bekijk nu de  $n \times 2n$  matrix

$$\begin{bmatrix} A & A \cdot B \end{bmatrix}$$

en breng deze met het Gauss-Jordan algoritme in rij echelon vorm

$$\text{ref}(\begin{bmatrix} A & A \cdot B \end{bmatrix}) = \begin{bmatrix} I_n & B' \end{bmatrix}.$$

Uit de opmerking na stelling 7 volgt dat  $B' = A^{-1} \cdot (A \cdot B) = B$ . Als we nu kijken naar de verandering in determinant van de twee  $n \times n$  blokken gedurende het proces

$$\begin{bmatrix} A & A \cdot B \end{bmatrix} \longrightarrow \begin{bmatrix} I_n & B \end{bmatrix}$$

dan halen we uit Stelling 21 dat

$$(-1)^s \lambda_1 \dots \lambda_k \det(A) = \det(I_n) = 1 \quad \text{en} \quad (-1)^s \lambda_1 \dots \lambda_k \det(A \cdot B) = \det(B),$$

waarbij  $s$  het aantal rij-wisselingen is en  $\lambda_j$  de factoren zijn waarmee rijen vermenigvuldigd moeten worden om  $A$  in rij-echelonvorm te krijgen. We concluderen dus ook in dit geval  $\det(A \cdot B) = \det(A) \det(B)$ .  $\square$

**Voorbeeld.** Als  $A = \begin{bmatrix} 1 & 2 \\ 3 & 16 \end{bmatrix}$  wat is  $\det(A^{100})$ ? Aangezien  $\det(A) = 10$  is  $\det(A^{100}) = \det(A)^{100} = 10^{100} = 1.\text{gogool}$ .

**Historische noot.** De naam Google is afgeleid uit gogool. Er bestaat ook het getal 1.gogoolplex =  $10^{10^{100}}$ .

Directe gevolgen van deze *multiplicativiteit* van de determinant zijn

$$\det(S^{-1} \cdot A \cdot S) = \det(A) \quad \text{en} \quad \det(A^{-1}) = \det(A)^{-1}.$$

**Lemma 7.** Voor een bovendriehoeksmatrix of benedendriehoeksmatrix is de determinant gelijk aan het product van de diagonalelementen.

*Bewijs.* We bewijzen de eigenschap opnieuw per inductie. In het geval  $n = 1$  is de eigenschap triviaal voldaan. Veronderstel nu dat de eigenschap geldig is voor alle dergelijke  $(n - 1) \times (n - 1)$ -matrices.

Beschouw een boven- of benedendriehoeksmatrix  $A$ . Dan is  $A_{11}$  er één van hetzelfde type, maar een orde lager, en dus met determinant gelijk aan alle diagonalelementen van  $A$ , behalve  $a_{11}$ . Dus is het product  $a_{11} \det A_{11}$  het product van alle diagonalelementen van  $A$ . Maar dit geeft slechts de eerste term in de som van  $\det A$ . We moeten dus aantonen dat alle andere termen in dit geval nul zijn.

In het geval dat  $A$  een bovendriehoeksmatrix is weten we dat  $a_{i1} = 0$  voor alle  $i > 1$ , en dus vallen alle overige termen in  $\det A$  weg.

In het geval dat  $A$  een benedendriehoeksmatrix is, bevat de eerste rij van  $A_{i1}$  (met  $i > 1$ ) allemaal nullen. Uit het bovenstaande weten we dat hieruit volgt dat  $\det A_{i1} = 0$ , en opnieuw vallen alle overige termen in  $\det A$  weg.  $\square$

**Lemma 8.** Voor elke elementaire matrix  $L = U, V, W$  geldt  $\det(L) = \det(L^{tr})$ .

*Bewijs.* Matrices van het type  $V$  en  $W$  zijn symmetrisch, en zijn dus gelijk aan hun getransponeerde. Matrices van het type  $U$  zijn boven-of benedendriehoeksmatrices, en blijven dat ook na het transponeren. Aangezien de diagonalelementen niet veranderen na het transponeren zullen de determinanten van  $U$  en  $U^{tr}$  gelijk zijn, in het licht van de vorige stelling.  $\square$

**Stelling 24.** Voor elke vierkante matrix  $A$  geldt  $\det(A^{tr}) = \det(A)$ .

*Bewijs.* We beschouwen eerst het geval waarbij  $\det(A) = 0$ . Dan is ook  $\det(A^{tr}) = 0$ . We bewijzen dit uit het ongerijmde: indien niet dan zou  $A^{tr}$  een inverse  $B$  hebben met  $A^{tr} \cdot B = I$ . Maar dan is na transponeren  $B^{tr} \cdot A = I$ , wat op zijn beurt zou betekenen dat  $A$  een inverse heeft en dus  $\det A \neq 0$ .

In het geval dat  $\det A \neq 0$  is  $A$  inverteerbaar, en dus het product van elementaire matrices:  $A = L_1 \dots L_k$  (zie Stelling 7). Maar dan is  $A^{tr} = L_k^{tr} \dots L_1^{tr}$  en

$$\det(A^{tr}) = \det(L_k^{tr}) \dots \det(L_1^{tr}) = \det(L_k) \dots \det(L_1) = \det(A),$$

wegens voorgaand Lemma.  $\square$

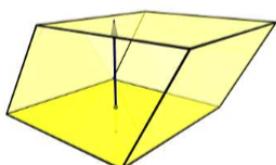
Het gevolg is dat ook als we twee kolommen van plaats wisselen, de determinant met een factor  $-1$  zal veranderen. Als we een kolom met een factor  $\lambda$  veranderen zal ook de determinant met factor  $\lambda$  veranderen.

De determinant van een orthogonale matrix  $Q$  is  $\pm 1$ . Immers  $Q^{tr} \cdot Q = I_n$  en dus

$$1 = \det(Q^{tr} \cdot Q) = \det(Q^{tr}) \det(Q) = \det(Q)^2$$

Dit laat ons toe een meetkundige interpretatie te geven van de determinant in het geval we werken over de reële getallen.

**Stelling 25.** (Over de reële getallen  $\mathbb{R}$ ) De absolute waarde  $|\det(A)|$  van de determinant van een  $n \times n$  matrix  $A$  gelijk aan het volume van de  $n$ -dimensionale parallellepipedum  $E_n$  opgespannen door de kolom vectors  $\vec{v}_i$  van  $A$ .



*Bewijs.* Herinner de QR-decompositie van  $A$  (zie stelling 19) dat is  $A = Q.R$  waarbij  $Q$  orthogonaal is en  $R$  een bovendriehoeksmatrix met diagonaal-entries  $r_{ii} = \|\vec{u}_i\| = \|\vec{v}_i - \text{proj}_{V_{i-1}}(\vec{v}_i)\|$  dat de afstand is van  $\vec{v}_i$  tot de deelruimte  $V_{i-1}$ . Bijgevolg is

$$\det(A) = \det(Q) \det(R) = \pm \prod_{i=1}^n \|\vec{u}_i\|.$$

Nu is het volume van de  $j$ -dimensionale parallellepipedum  $E_j$  met basis  $E_{j-1}$  in  $V_{j-1}$  en hoogte  $\|\vec{u}_i\|$  gelijk aan  $\text{vol}(E_{j-1})\|\vec{u}_i\|$ . Bijgevolg hebben we inductief

$$\text{vol}(E_n) = \prod_{i=1}^n \|\vec{u}_i\|.$$

□

Herinner dat de  $(i,j)$ -minor van  $A$  de  $(n-1) \times (n-1)$  matrix  $A_{ij}$  is die men bekomt door uit  $A$  de  $i$ -de rij en de  $j$ -de kolom te schrappen.

**Stelling 26** (Laplaceontwikkeling naar de  $k$ -de kolom). *Gegeven een  $n \times n$  matrix  $A = (a_{ij})$ . Kies een willekeurige kolom  $k$ . Dan geldt*

$$\begin{aligned} \det(A) &= (-1)^{k+1}a_{1k}\det(A_{1k}) + (-1)^{k+2}a_{2k}\det(A_{2k}) + \dots + (-1)^{k+n}a_{nk}\det(A_{nk}) \\ &= \sum_{i=1}^n (-1)^{k+i}a_{ik}\det(A_{ik}). \end{aligned}$$

*Bewijs.* Stel  $A = [\vec{a}_1 \ \dots \ \vec{a}_{k-1} \ \vec{a}_k \ \vec{a}_{k+1} \ \dots \ \vec{a}_n]$ . Dan is de matrix die we bekomen door op  $A$   $k-1$  kolomwisselingen toe te passen gegeven door  $B = [\vec{a}_k \ \vec{a}_1 \ \dots \ \vec{a}_{k-1} \ \vec{a}_{k+1} \ \dots \ \vec{a}_n]$ . Er geldt dan

$$\det A = (-1)^{k-1} \det B = (-1)^{k-1} \sum_{i=1}^n (-1)^{1+i}b_{i1}\det(B_{i1}) = (-1)^{k-1} \sum_{i=1}^n (-1)^{1+i}a_{ik}\det(A_{ik})$$

omdat  $b_{i1} = a_{ik}$  en  $B_{ik} = A_{ik}$ .

□

Merk ook op dat we, wegens  $\det(A^{tr}) = \det(A)$ , ook mogen ontwikkelen naar rijen in plaats van kolommen. In de praktijk kies je als ontwikkelingskolom of -rij best die kolom of rij die het meest nullen bevat.

**Voorbeeld.** De volgende berekening maakt gebruik van de Laplaceontwikkeling.

$$\det \begin{bmatrix} -2 & 0 & 0 & 1 \\ 2 & 3 & 0 & 1 \\ 7 & -5 & 2 & -2 \\ 0 & -1 & 0 & 1 \end{bmatrix} = 2 \det \begin{bmatrix} -2 & 0 & 1 \\ 2 & 3 & 1 \\ 0 & -1 & 1 \end{bmatrix} = 2 \left( -2 \det \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix} + \det \begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix} \right) = -20.$$

## 2 Toepassingen

Determinanten zijn ook nuttig om de oplossing van een lineair stelsel van  $n$  vergelijkingen in  $n$  onbekenden te bepalen.

**Stelling 27** (Regel van Cramer). *Zij  $A$  een vierkante matrix met  $\det(A) \neq 0$ . Laat  $A_i$  de matrix zijn waarvan de  $i$ -de kolom  $\vec{v}_i$  van  $A$  vervangen is door  $\vec{b}$ . De unieke oplossing  $\vec{x} = [x_1 \dots x_n]^{\text{tr}}$  van het stelsel  $A.\vec{x} = \vec{b}$  wordt gegeven door*

$$x_i = \frac{\det(A_i)}{\det(A)}, \quad \forall i = 1, \dots, n.$$

*Bewijs.* Stel  $A$  voor door  $\begin{bmatrix} \vec{a}_1 & \dots & \vec{a}_n \end{bmatrix}$ , dan is  $A_i = \begin{bmatrix} \vec{a}_1 & \dots & \vec{a}_{i-1} & \vec{b} & \vec{a}_{i+1} & \dots & \vec{a}_n \end{bmatrix}$ . Beschouw een vaste  $i$ . We weten uit voorgaande stellingen dat er een unieke oplossing  $\vec{x}$  besaat, en ze voldoet aan  $\vec{b} = A.\vec{x}$ . Bij definitie is  $\det(A_i)$  gelijk aan

$$\begin{aligned} \det \begin{bmatrix} \vec{a}_1 & \dots & \vec{b} & \dots & \vec{a}_n \end{bmatrix} &= \det \begin{bmatrix} \vec{a}_1 & \dots & A.\vec{x} & \dots & \vec{a}_n \end{bmatrix} = \\ \det \begin{bmatrix} \vec{a}_1 & \dots & \sum_{j=1}^n x_j \vec{a}_j & \dots & \vec{a}_n \end{bmatrix} &= x_i \det \begin{bmatrix} \vec{a}_1 & \dots & \vec{a}_i & \dots & \vec{a}_n \end{bmatrix} = x_i \det(A), \end{aligned}$$

waar de voorlaatste gelijkheid volgt uit het feit dat voor alle  $j \neq i$ -de term in de som

$$x_j \det \begin{bmatrix} \vec{a}_1 & \dots & \vec{a}_j & \dots & \vec{a}_n \end{bmatrix} = 0,$$

omdat de matrix dan twee gelijke kolommen heeft.  $\square$

We kunnen de regel van Cramer en de Laplaceontwikkeling combineren teneinde een expli- ciete beschrijving te geven van de inverse van een inverteerbare  $n \times n$  matrix.

**Stelling 28.** *Als  $A$  een inverteerbare  $n \times n$  matrix is dan geldt*

$$(A^{-1})_{i,j} = (-1)^{i+j} \frac{\det(A_{ji})}{\det(A)}.$$

*Bewijs.* Stel  $A^{-1}$  voor door  $\begin{bmatrix} \vec{x}_1 & \dots & \vec{x}_n \end{bmatrix}$ . Omdat  $A.A^{-1} = I_n = \begin{bmatrix} \vec{e}_1 & \dots & \vec{e}_n \end{bmatrix}$  geldt dat de kolommen  $\vec{x}_j$  van  $A^{-1}$  voldoen aan  $A.\vec{x}_j = \vec{e}_j$ ,  $j = 1, \dots, n$ . Het element op de  $i$ -de rij van  $\vec{x}_j$  is  $(A^{-1})_{ij}$ . Via de regel van Cramer is

$$(A^{-1})_{ij} = (\vec{x}_j)_i = \frac{\det A_i}{\det(A)},$$

met hier

$$\begin{aligned}
 A_i &= \left[ \vec{a}_1 \ \dots \ \vec{a}_{i-1} \ \vec{e}_j \ \vec{a}_{i+1} \ \dots \ \vec{a}_n \right] \\
 &= \begin{bmatrix} a_{11} & \dots & a_{1 i-1} & 0 & a_{1 i+1} & \dots & a_{1 n} \\ a_{21} & \dots & a_{2 i-1} & 0 & a_{2 i+1} & \dots & a_{2 n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{j1} & \dots & a_{j i-1} & 1 & a_{j i+1} & \dots & a_{j n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n i-1} & 0 & a_{n i+1} & \dots & a_{n n} \end{bmatrix}.
 \end{aligned}$$

Wanneer ze nu deze matrix ontbinden naar de  $i$ -de kolom vinden we dat  $\det(A_i) = (-1)^{i+j} \det(A_{ji})$ .

□

Let op de volgorde van de indices  $A_{ji}$  in bovenstaande eigenschap! Voor een inverteerbare  $2 \times 2$  matrix hebben we dus

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

**Historische noot.** [Gabriel Cramer](#) (1704-1752) was een Zwitserse wiskundige die de naar hem genoemde regel gebruikte om stelsels in 5 onbekenden op te lossen. Wederom was de regel al bekend bij wiskundigen voor hem.





# 5 Eigenwaarden en eigenvectoren

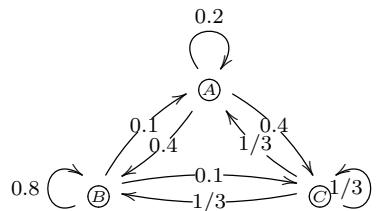
## 1 Voorbeeld en definities

We beginnen met een voorbeeld. Stel dat drie spelers  $A$ ,  $B$  en  $C$  op training de bal rondspelen. Uit vorige trainingen hebben we een patroon kunnen vaststellen: als  $A$  de bal heeft, geeft hij in 40% van de tijd een pas naar  $B$ , 40% een pas naar  $C$  en houdt 20% van de tijd de bal bij.  $B$  is iets egoïstischer en houdt 80% van de tijd de bal bij en verdeelt de overige 20% gelijk over  $A$  en  $C$ .  $C$  daarentegen verdeelt gelijk de tijd tussen de bal bijnouden of een pas naar  $A$  en een pas naar  $C$  te geven.



Stel dat de drie spelers een uurtje trainen en al honderden passen gegeven hebben. Wat is de kans dat je op een gegeven moment de bal zal aantreffen bij  $A$ , bij  $B$  en bij  $C$ ?

De kansen waarmee de bal rondgespeeld wordt kunnen grafisch voorgesteld worden als



of in matrixvorm als

$$P = \begin{bmatrix} \frac{2}{10} & \frac{1}{10} & \frac{1}{3} \\ \frac{4}{10} & \frac{8}{10} & \frac{1}{3} \\ \frac{4}{10} & \frac{1}{10} & \frac{1}{3} \end{bmatrix}.$$

Stel dat bij het begin van de training de bal bij speler  $A$  is, dan is na 1 tijdsinterval de kansverdeling voor de bal

$$P \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix},$$

en na twee tijdsintervallen

$$P^2 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = P \cdot \begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 0.21333\dots \\ 0.5333\dots \\ 0.25333\dots \end{bmatrix}.$$

Na 10 en 20 tijdsintervallen is de kans, respectievelijk,

$$\begin{bmatrix} 0.16111\dots \\ 0.64555\dots \\ 0.19333\dots \end{bmatrix} \quad \text{en} \quad \begin{bmatrix} 0.161290333\dots \\ 0.645161222\dots \\ 0.19354888\dots \end{bmatrix}$$

We zien dus dat de kansverdeling convergeert en zouden graag de limietverdeling kennen.

Stel dat we formeel de  $\infty$ -macht van  $P$  zouden kunnen uitrekenen,

$$\vec{v} = P^\infty \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

dan geldt hiervoor

$$P \cdot \vec{v} = P^{1+\infty} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = P^\infty \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \vec{v}$$

Dit betekent dat  $\vec{v}$  een *eigenvector* is met *eigenwaarde* 1 voor de matrix  $P$ .

**Definitie 11.** Voor een  $n \times n$  matrix  $A$  noemen we een vector  $\vec{v} \in \mathbb{R}^n$  een eigenvector met eigenwaarde  $\lambda$  indien geldt dat

$$A \cdot \vec{v} = \lambda \vec{v}.$$

De verzameling

$$E_\lambda(A) = \{\vec{v} \in \mathbb{R}^n \mid A \cdot \vec{v} = \lambda \vec{v}\}$$

is een deelruimte van  $\mathbb{R}^n$  die we de  $\lambda$ -eigenruimte van  $A$  noemen.

**Opmerking.** (1) Men kan de voorgaande definities ook bekijken vanuit het standpunt van een vaste waarde  $\lambda \in \mathbb{R}$ . Indien de de corresponderende deelruimte  $E_\lambda(A)$  slechts de nulvector  $\vec{0}$  bevat, m.a.w. als de enige eigenvector voor  $\lambda$  slechts de nulvector is, dan spreken we niet van een eigenwaarde  $\lambda$ . Als we spreken over een eigenvector  $\vec{v}$  zullen we ook steeds veronderstellen worden dat  $\vec{v} \neq \vec{0}$ .

(2) Eén van de eigenschappen van een deelruimte is dat als  $\vec{v} \in E_\lambda(A)$ , dan ook  $a\vec{v} \in E_\lambda(A)$ , voor elke  $a \in \mathbb{R}$ . Dat betekent natuurlijk niet dat de eigenruimtes noodzakelijk ééndimensionaal zouden zijn. Het volgende voorbeeld toont dit aan.

**Voorbeeld.** Bepaal  $E_1(A)$ ,  $E_{-1}(A)$  en  $E_{-3}(A)$  voor  $A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ 4 & 4 & 11 \end{bmatrix}$ .

Een vector  $\vec{v} = (v_1, v_2, v_3)$  behoort tot  $E_1(A)$  als  $(A - I_3).\vec{v} = \vec{0}$ , of als

$$\begin{cases} 4v_1 + 8v_2 + 16v_3 = 0, \\ 4v_1 + 8v_3 = 0, \\ 4v_1 4v_2 12v_3 = 0. \end{cases}$$

of als  $v_1 = -2v_3$  en  $v_2 = -v_3$ .

Dus is  $E_1(A) = \left\{ k \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \mid k \in \mathbb{R} \right\}$  ééndimensionaal.

Een vector  $\vec{v} = (v_1, v_2, v_3)$  behoort tot  $E_{-1}(A)$  als  $(A + I_3).\vec{v} = \vec{0}$ , of als

$$\begin{cases} 6v_1 + 8v_2 + 16v_3 = 0, \\ 4v_1 + 2v_2 + 8v_3 = 0, \\ 4v_1 4v_2 10v_3 = 0. \end{cases}$$

Dit stelsel heeft een unieke oplossing:  $(v_1 = 0, v_2 = 0, v_3 = 0)$ . Dus is  $E_{-1}(A) = \{\vec{0}\}$ . Om die reden is  $\lambda = -1$  *geen* eigenwaarde van  $A$ , en is  $E_{-1}(A)$  *geen* eigenruimte voor  $A$ .

Een vector  $\vec{v} = (v_1, v_2, v_3)$  behoort tot  $E_{-3}(A)$  als  $(A + 3I_3).\vec{v} = \vec{0}$ , of

$$\begin{cases} 8v_1 + 8v_2 + 16v_3 = 0, \\ 4v_1 + 4v_2 + 8v_3 = 0, \\ 4v_1 4v_2 8v_3 = 0, \end{cases}$$

of als  $v_1 = -v_2 + -2v_3$ . Dus is  $E_{-3}(A) = \left\{ l \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + m \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \mid l, m \in \mathbb{R} \right\}$  een tweedimensionale deelruimte van  $\mathbb{R}^3$ .

Laat ons nu terugkeren naar het voetbalvoorbeeld. We hadden gezien dat de limietverdeling

een *eigenvector*  $\vec{v}$  met *eigenwaarde* 1 was voor de matrix  $P = \begin{bmatrix} 0.2 & 0.1 & \frac{1}{3} \\ 0.4 & 0.8 & \frac{1}{3} \\ 0.4 & 0.1 & \frac{1}{3} \end{bmatrix}$ . Het stelsel  $(P - I_3).\vec{v} = \vec{0}$  is equivalent met

$$\begin{cases} -0.8v_1 + 0.1v_2 + \frac{1}{3}v_3 = 0, \\ 0.4v_1 - 0.2v_2 + \frac{1}{3}v_3 = 0, \\ 0.4v_1 + 0.1v_2 - \frac{2}{3}v_3 = 0. \end{cases}$$

De oplossingverzameling van dit stelsel worden gegeven door

$$E_1(P) = \left\{ k \begin{bmatrix} 5 \\ 20 \\ 6 \end{bmatrix} \mid k \in \mathbb{R} \right\}.$$

De enige kansvector (met som van de componenten gelijk aan 1) in  $E_1(P)$  is dus

$$\vec{v} = \frac{1}{31} \begin{bmatrix} 5 \\ 20 \\ 6 \end{bmatrix}.$$

Na lange tijd voetballen is de kans het grootst dat  $B$  de bal heeft.

## 2 Eigenschappen

**Stelling 29.** *Als  $A$  een  $n \times n$  matrix is en  $\vec{v}$  een eigenvector is met eigenwaarde  $\lambda$  en  $\vec{w}$  een eigenvector met eigenwaarde  $\mu$ , dan zijn  $\vec{v}$  en  $\vec{w}$  lineair onafhankelijk als  $\lambda \neq \mu$ . In het bijzonder geldt voor  $\lambda \neq \mu$ ,*

$$E_\lambda(A) \cap E_\mu(A) = \{\vec{0}\}.$$

*Bewijs.* Voor het bewijzen van de eerste uitspraak, stel dat  $\vec{w} = \alpha\vec{v}$ . Er geldt dan

$$A.\vec{w} = \mu\vec{w} \quad \text{en} \quad A.\vec{w} = A.\alpha\vec{v} = \alpha A.\vec{v} = \alpha\lambda\vec{v} = \lambda\vec{w}.$$

Dit is een contradictie als  $\lambda \neq \mu$ . De tweede uitspraak volgt uit dezelfde redenering met  $\alpha = 1$ .  $\square$

Als gevolg heeft de matrix  $A$  slechts eindig veel verschillende eigenwaarden  $\lambda$  (waarvoor  $E_\lambda(A) \neq \{\vec{0}\}$ ). Dit komt omdat in  $\mathbb{R}^n$  hoogstens  $n$  verschillende vectoren lineair onafhankelijk kunnen zijn (zie Lemma 3).

**Stelling 30.** *Laat  $\vec{v} \neq \vec{0}$  een eigenvector zijn van  $A$  met eigenwaarde  $\lambda$ , dan is*

$$\det(\lambda I_n - A) = 0.$$

*De eigenwaarden  $\lambda$  van  $A$ , met eigenruimte  $E_\lambda(A) \neq \{\vec{0}\}$ , zijn de wortels van de functie*

$$x \mapsto \chi_A(x) = \det(xI_n - A).$$

*Deze functie is een monische veelterm in  $x$  van graad  $n$ , die we de karakteristieke veelterm noemen van de  $n \times n$  matrix  $A$ .*

*Bewijs.* Als  $\vec{v}$  een niet-nul eigenvector is van  $A$  met eigenwaarde  $\lambda$  dan is  $\vec{v}$  een niet-nul oplossing van het lineaire stelsel

$$(\lambda I_n - A).\vec{v} = \vec{0}.$$

In het geval dat  $\det(\lambda I_n - A) \neq 0$ , dan zou het stelsel enkel de unieke oplossing  $\vec{v} = \vec{0}$  kunnen hebben (zie Stelling 22). We kunnen dus besluiten dat  $\det(\lambda I_n - A) = 0$ .

Het feit dat de functie  $\chi_A(x)$  een veelterm is van graad  $n$  volgt makkelijk uit de Laplaceontwikkeling voor de determinant. De veelterm is 'monisch', wat betekent dat de coëfficiënt die hoort bij  $x^n$  net 1 is.  $\square$

Een belangrijk gevolg van deze stelling is dat er ten hoogste  $n$  verschillende eigenwaarden met niet-nul eigenruimte kunnen bestaan voor een matrix  $A$ . Dit volgt uit het feit dat elk graad  $n$  polynoom ten hoogste  $n$  wortels heeft.

**Stelling 31.** *De getransponeerde matrix  $A^{tr}$  en een geconjugeerde matrix  $S^{-1}.A.S$ , met  $S$  inverteerbaar, hebben dezelfde eigenwaarden als  $A$ .*

*Bewijs.* Deze beweringen volgen uit:

$$\det(xI_n - A^{tr}) = \det((xI_n - A)^{tr}) = \det(xI_n - A)$$

en

$$\det(xI_n - S^{-1}.A.S) = \det(S^{-1}.(xI_n - A).S) = \det(S)^{-1} \det(xI_n - A) \det(S) = \det(xI_n - A)$$

$\square$

Let op: uit het bewijs volgt niet dat deze matrices ook allemaal dezelfde eigenvectoren zouden hebben. Bij de inverse is dat wel het geval.

**Stelling 32.** *Als  $\lambda \neq 0$  een eigenwaarde en  $\vec{v}$  een eigenvector is van een inverteerbare matrix  $A$  dan is  $1/\lambda$  een eigenwaarde en  $\vec{v}$  een eigenvector van de inverse matrix  $A^{-1}$ .*

*Bewijs.* Dit volgt eenvoudig uit  $A^{-1} \cdot A = I$ . □

De karakteristieke veelterm van een  $2 \times 2$  matrix  $A$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is gelijk aan

$$\det \begin{bmatrix} x - a & b \\ c & x - d \end{bmatrix} = (x - a)(x - d) - bc = x^2 - (a + d)x + (ad - bc) = x^2 - sp(A)x + \det(A)$$

waarbij  $sp(A)$  het *spoor* van  $A$  is, dat wil zeggen de som van de diagonalelementen. Bijgevolg heeft  $A$  als eigenwaarden

$$\lambda_{1,2} = \frac{sp(A) \pm \sqrt{sp(A)^2 - 4 \det(A)}}{2}.$$

**Definitie 12.** *Een  $n \times n$ -matrix  $A$  is diagonaliseerbaar indien er een inverteerbare  $n \times n$ -matrix  $S$  en een diagonalmatrix  $D$  bestaan, waarvoor  $A = S.D.S^{-1}$ .*

**Stelling 33.** *Een  $n \times n$  matrix  $A$  is diagonaliseerbaar als en slechts als de matrix  $A$   $n$  lineair onafhankelijke eigenvectoren heeft. In dat geval zijn de kolommen van  $S$  eigenvectoren en de diagonalelementen van  $D$  eigenwaarden van  $A$ .*

*Bewijs.* Veronderstel eerst dat  $A$  diagonaliseerbaar is, met  $A = S.D.S^{-1}$ . Noem  $d_i$  de diagonalelementen van  $D$ . Via Stelling 32 weten we al dat de matrices  $A$  en  $D$  dezelfde eigenwaarden zullen hebben. In het geval van een diagonalmatrix  $D$  zijn de eigenwaarden net de diagonalelementen  $d_i$  (denk aan de karakteristieke vergelijking van  $D$ ). We tonen nu aan dat elke kolom van  $S$  een eigenvector is van  $A$ . De  $i$ -de kolom van  $S$  wordt gegeven door  $S.\vec{e}_i$ . Er geldt voor deze vector dat

$$A.(S.\vec{e}_i) = S.D.S^{-1}.S.\vec{e}_i = S.D.\vec{e}_i = S.(d_i\vec{e}_i) = d_i(S.\vec{e}_i).$$

Dit toont aan dat  $S.\vec{e}_i$  een eigenvector is van  $A$  met eigenwaarde  $d_i$ . We weten echter ook dat  $S$  inverteerbaar is. Volgens Stelling 9 zijn alle  $n$  kolommen van een inverteerbare matrix lineair onafhankelijk.

Beschouw nu de omgekeerde richting. We veronderstellen nu dat  $A$   $n$  onafhankelijke eigenvectoren  $\{\vec{v}_1, \dots, \vec{v}_n\}$  heeft met eigenwaarden  $\lambda_i$ . We kunnen er een matrix  $S$  meeformen:

$$S = \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_n \end{bmatrix}.$$

Deze matrix  $S$  is inverteerbaar, opnieuw wegens Stelling 9. We bepalen nu de  $i$ -de kolom van het product  $S^{-1} \cdot A \cdot S$ :

$$S^{-1} \cdot A \cdot S \cdot \vec{e}_i = S^{-1} \cdot A \cdot \vec{v}_i = S^{-1} \lambda_i \vec{v}_i = \lambda_i S^{-1} \vec{v}_i = \lambda_i \vec{e}_i.$$

Bijgevolg is  $S^{-1} \cdot A \cdot S$  een diagonalmatrix  $D$  met als diagonalelementen eigenwaarden  $\lambda_i$  van  $A$ . Bijgevolg is  $A$  inderdaad te schrijven als  $S \cdot D \cdot S^{-1}$ .  $\square$

**Definitie 13.** Een basis  $\{\vec{v}_1, \dots, \vec{v}_n\}$  van  $\mathbb{R}^n$  waarvan alle vectoren  $\vec{v}_i$  eigenvectoren zijn van een matrix  $A$  noemt men een eigenbasis.

Herinner dat  $n$  lineair onafhankelijke vectoren in  $\mathbb{R}^n$  steeds een basis vormen (zie Stelling 9). De vorige stelling zegt dus dat *diagonaliseerbaarheid van  $A$  equivalent is met het bestaan van een eigenbasis*, corresponderend met  $A$ .

In het bijzonder geval dat alle eigenwaarden verschillend zijn, kunnen we nog een sterkere eigenschap bewijzen.

**Stelling 34.** Als een  $n \times n$  matrix  $A$   $n$  verschillende eigenwaarden heeft, dan is  $A$  steeds diagonaliseerbaar.

*Bewijs.* Wegens de vorige stelling moeten we enkel aantonen dat  $A$ , in dit geval,  $n$  lineair onafhankelijke eigenvectoren heeft. Laat  $\lambda_1, \dots, \lambda_n$  de verschillende eigenwaarden van  $A$  zijn. Omdat  $\det(\lambda_i I_n - A) = 0$  bestaat er voor elke  $1 \leq i \leq n$  een niet-nul eigenvector  $\vec{v}_i$  met eigenwaarde  $\lambda_i$ . We weten al dat deze eigenvectoren voor verschillende eigenwaarden, twee aan twee, lineair onafhankelijk zijn. We beweren nu dat het volledige stel  $\{\vec{v}_1, \dots, \vec{v}_n\}$  ook lineair onafhankelijk is, en dus (wegens Stelling 9) een basis van  $\mathbb{R}^n$ .

Het bewijs gaat op inductie. We weten dat  $\vec{v}_1$  en  $\vec{v}_2$  lineair onafhankelijk zijn. Veronderstel nu dat het stel  $\{\vec{v}_1, \dots, \vec{v}_{i-1}\}$  lineair onafhankelijk is (met  $i \geq 3$ ). We moeten aantonen dat hier uit volgt dat het stel  $\{\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_i\}$  ook lineair onafhankelijk is. Dat zullen we doen aan de hand van een redenering uit het ongerijmde. Veronderstel hiertoe dat er een niet-nul lineaire relatie zou bestaan

$$\vec{v}_i = \alpha_1 \vec{v}_1 + \dots + \alpha_{i-1} \vec{v}_{i-1},$$

met minstens één  $\lambda_j$  niet-nul (met  $j$  een getal in  $1, \dots, i-1$ ). Enerzijds geldt er dan

$$A \cdot \vec{v}_i = \lambda_i \vec{v}_i = \alpha_1 \lambda_i \vec{v}_1 + \dots + \alpha_{i-1} \lambda_i \vec{v}_{i-1},$$

maar anderzijds ook

$$A \cdot \vec{v}_i = \alpha_1 A \cdot \vec{v}_1 + \dots + \alpha_n A \cdot \vec{v}_{i-1} = \alpha_1 \lambda_1 \vec{v}_1 + \dots + \alpha_{i-1} \lambda_{i-1} \vec{v}_{i-1}.$$

Als we beide uitdrukkingen aan elkaar gelijk stellen,

$$\alpha_1(\lambda_1 - \lambda_i) \vec{v}_1 + \dots + \alpha_{i-1}(\lambda_{i-1} - \lambda_i) \vec{v}_{i-1} = \vec{0},$$

en gebruik maken van de lineair onafhankelijkheid van het stel, dan is  $\alpha_1(\lambda_1 - \lambda_i) = \dots \alpha_{i-1}(\lambda_1 - \lambda_i) = 0$ . Aangezien zeker één van de  $\alpha_j \neq 0$  geldt dus  $\lambda_j = \lambda_i$ , met  $j \neq i$ . Maar dat is tegenstrijdig met de aannname dat alle eigenwaarden tegenstrijdig zijn. We besluiten dus dat ook  $\{\vec{v}_1, \dots, \vec{v}_i\}$  lineair onafhankelijk is, wat de redenering op inductie sluit.  $\square$

**Voorbeelden.** Wanneer  $A$  geen  $n$  verschillende eigenwaarden heeft, dan kan het soms zijn dat  $A$  diagonalizeerbaar is, en soms niet. We beschouwen drie matrices, ter verduidelijking,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

De karakteristieke vergelijking van  $A$  is

$$\lambda(\lambda - 1)(\lambda + 1) = 0,$$

zodat er drie verschillende eigenwaarden zijn,  $\lambda = 0$ ,  $\lambda = 1$  en  $\lambda = -1$ . Uit de stelling weten we al dat we een eigenbasis zullen kunnen opstellen.

We berekenen eerst de eigenvectoren  $\vec{v} = [v_1 \ v_2 \ v_3]^T$  die horen bij  $\lambda = 0$ . De vectoren voldoen aan  $A \cdot \vec{v} = \vec{0}$  en dus aan

$$\begin{cases} v_1 = 0, \\ 2v_1 - v_3 = 0. \end{cases}$$

Dit betekent dat de eigenvectoren gegeven worden door  $\vec{v} = [0 \ k \ 0]^T$ , waarbij  $k$  elk willekeurig niet-nul reëel getal kan zijn. Analoog gaat men na dat de eigenvectoren van  $\lambda = 1$  en  $\lambda = -1$  gegeven worden door  $\vec{v} = [l \ 0 \ l]^T$  en  $\vec{v} = [0 \ 0 \ m]^T$ , respectievelijk. Een voorbeeld van een eigenbasis is dan  $\left\{ [0 \ 1 \ 0]^T, [1 \ 0 \ 1]^T, [0 \ 0 \ 1]^T \right\}$ .

Voor de matrix  $B$  is de karakteristieke vergelijking

$$\lambda(\lambda - 1)^2 = 0.$$

Nu zijn er slechts twee verschillende eigenwaarden:  $\lambda = 0$  en  $\lambda = 1$ . De laatste waarde is echter een dubbele wortel van de vergelijking, vandaar dat we zeggen dat de eigenwaarde  $\lambda = 1$  *multipliciteit* 2 heeft.

De eigenvectoren die horen bij  $\lambda = 0$  voldoen aan  $B \cdot \vec{v} = \vec{0}$  en dus aan

$$\begin{cases} v_3 = 0, \\ v_2 + v_3 = 0, \\ v_3 = 0. \end{cases}$$

Dus is  $\vec{v} = \begin{bmatrix} l & 0 & 0 \end{bmatrix}^{tr}$ , met  $l$  een willekeurig niet-nul reëel getal. De eigenruimte  $E_0(B)$  is in dit geval dus 1-dimensionaal.

Analoog voldoen de eigenvectoren bij  $\lambda = 1$  aan  $(B - I) \cdot \vec{v} = \vec{0}$ , of aan

$$\begin{cases} v_1 - v_3 = 0, \\ -2v_3 = 0. \end{cases}$$

Nu zijn de eigenvectoren  $\vec{v} = \begin{bmatrix} 0 & k & 0 \end{bmatrix}^{tr}$ , met  $k$  willekeurig. De eigenruimte  $E_1(B)$  is opnieuw slechts 1-dimensionaal. We kunnen dus nooit 3 lineair onafhankelijke eigenvectoren vinden en de matrix  $B$  is niet diagonalizeerbaar.

Voor de matrix  $C$  is de situatie anders. Opnieuw is de karakteristieke vergelijking

$$\lambda(\lambda - 1)^2 = 0,$$

zodat de eigenwaarden en hun multipliciteit dezelfde zijn als voor  $B$ . De eigenruimte bij  $\lambda = 0$  is opnieuw 1-dimensionaal,

$$E_0(C) = \left\{ \begin{bmatrix} k & 0 & -k \end{bmatrix}^{tr} \mid k \in \mathbb{R} \right\},$$

maar de eigenruimte  $E_1(C)$  is nu 2-dimensionaal. De eigenruimte bij  $\lambda = 1$  is immers van de vorm

$$E_1(C) = \left\{ \begin{bmatrix} 0 & l & m \end{bmatrix}^{tr} \mid l, m \in \mathbb{R} \right\}.$$

In deze ruimte kunnen we, onafhankelijk van elkaar, zowel  $l$  als  $m$  willekeurig kiezen. We kunnen dus wel degelijk 3 lineair onafhankelijke eigenvectoren vinden, bijvoorbeeld  $\begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^{tr}$  (bij  $\lambda = 0$ ),  $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^{tr}$  (bij  $\lambda = 1$ ) en  $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^{tr}$  (bij  $\lambda = 1$ ).

Het verschil tussen de twee laatste gevallen heeft dus te maken met het feit dat bij  $B$  de multipliciteit van de eigenwaarde  $\lambda = 1$  niet overeenstemt met de dimensie van de eigenruimte  $E_1(B)$ . Dit is eigenlijk een algemeen resultaat is, maar het bewijs van onderstaande stelling valt buiten het bestek van deze cursus.

**Stelling 35.** Een  $n \times n$ -matrix  $A$  is diagonaliseerbaar als en slechts als voor alle eigenwaarden  $\lambda$  van  $A$  de multipelititeit gelijk is aan de dimensie van de eigenruimte  $E_\lambda(A)$ .

**Berekening van machten en inverse.** Beschouw opnieuw het geval van een diagonaliseerbare matrix  $A$ . Stel

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

De eigenschap dat een matrix diagonaliseerbaar is,  $A = S.D.S^{-1}$ , is handig in tal van toepassingen. Stel je voor dat je de  $k$ -de macht van een matrix  $A$  wenst te berekenen. Dan is  $A^k = A \cdot \dots \cdot A = S.L^k.S^{-1}$ . Maar,  $S, S^{-1}$  en

$$L^k = \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix},$$

zijn eenvoudig te berekenen, eens je de eigenwaarden en eigenvectoren kent. Voor grote  $k$  is  $S.L^k.S^{-1}$  uitrekenen makkelijker dan  $k$ -keer het matrixproduct van  $A$  met zichzelf.

Men kan op dezelfde wijze ook nagaan dat  $A^{-1} = S.L^{-1}.S^{-1}$ , met

$$L^{-1} = \begin{bmatrix} \frac{1}{\lambda_1} & & & \\ & \frac{1}{\lambda_2} & & \\ & & \ddots & \\ & & & \frac{1}{\lambda_n} \end{bmatrix}.$$

**Definitie 14.** Het spoor van een vierkante matrix is gelijk aan de som van haar diagonalelementen.

We noteren  $\text{sp } A = \sum_{i=1}^n a_{ii}$ . Een onmiddellijk eigenschap is dat, voor twee vierkante matrices van dezelfde orde

$$\text{sp}(A \cdot B) = \text{sp}(B \cdot A).$$

Dit komt omdat  $(A \cdot B)_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$  en dus is

$$\text{sp}(A \cdot B) = \sum_{i=1}^n (A \cdot B)_{ii} = \sum_{i=1}^n \sum_{k=1}^n a_{ik}b_{ki} = \sum_{i=1}^n \sum_{k=1}^n b_{ki}a_{ik} = \sum_{i=1}^n (B \cdot A)_{ii} = \text{sp}(B \cdot A).$$

**Stelling 36.** *Indien  $\lambda_i$  de eigenwaarden van een (diagonalizeerbare) matrix  $A$  zijn, dan geldt  $\det A = \lambda_1 \dots \lambda_n$  en  $\text{sp } A = \lambda_1 + \dots + \lambda_n$ .*

*Bewijs.* Het is evident dat  $\det L = \lambda_1 \dots \lambda_n$  en  $\text{sp } L = \lambda_1 + \dots + \lambda_n$ . Beide eigenschappen volgen dan uit  $A = S.L.S^{-1}$ , uit  $\det(A.B) = \det(A)\det(B)$  en uit  $\text{sp}(A.B) = \text{sp}(B.A)$ .  $\square$

**Stelling 37** (Cayley-Hamilton). *Als  $\chi_A(x) = x^n + \alpha_1 x^{n-1} + \dots + \alpha_n$  de karakteristieke veelterm is van een  $n \times n$  matrix  $A$  dan geldt dat*

$$A^n + \alpha_1 A^{n-1} + \dots + \alpha_{n-1} A + \alpha_n I_n = 0,$$

*met in het rechterlid de nulmatrix.*

*Bewijs.* We bewijzen dit eerst in het geval  $A$   $n$  verschillende eigenwaarden heeft  $\lambda_1, \dots, \lambda_n$  en dus diagonalizeerbaar is met  $S^{-1}.A.S = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Maar dan is

$$(S^{-1}.A.S)^i = \begin{bmatrix} \lambda_1^i & & 0 \\ & \ddots & \\ 0 & & \lambda_n^i \end{bmatrix}$$

en dus is

$$(S^{-1}.A.S)^n + \alpha_1(S^{-1}.A.S)^{n-1} + \dots + \alpha_n I_n = \begin{bmatrix} \chi_A(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & \chi_A(\lambda_n) \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix}$$

Maar we hebben ook

$$(S^{-1}.A.S)^n + \alpha_1(S^{-1}.A.S)^{n-1} + \dots + \alpha_n I_n = S^{-1}.(A^n + \alpha_1 A^{n-1} + \dots + \alpha_n I_n).S.$$

Het algemene geval volgt uit continuiteitsoverwegingen. Als  $A$  eigenwaarden heeft waarvan sommigen met hogere multipliciteit voorkomen, dan kunnen we arbitraire kleine getallen  $\epsilon_{ij}$  vinden zodat de matrix

$$A' = A + \begin{bmatrix} \epsilon_{11} & \dots & \epsilon_{1n} \\ \vdots & & \vdots \\ \epsilon_{n1} & \dots & \epsilon_{nn} \end{bmatrix}$$

$n$  verschillende eigenwaarden heeft en dus is  $\chi_{A'}(A') = (0)_{ij}$ . Omdat de entries van  $\chi_{A'}(A')$  continue functies zijn, volgt dat ook  $\chi_A(A) = 0$ .  $\square$

### 3 Enkele toepassingen

**Fibonacci en zijn konijnen.** De *Fibonaccitallenreeks* geeft de groei aan van een konijnenpopulatie na  $n$  generaties. De reeks begint met

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

Als  $u_n$  het  $n$ -de getal is in de reeks, dan is de recursie-betrekking

$$u_{n+1} = u_n + u_{n-1},$$

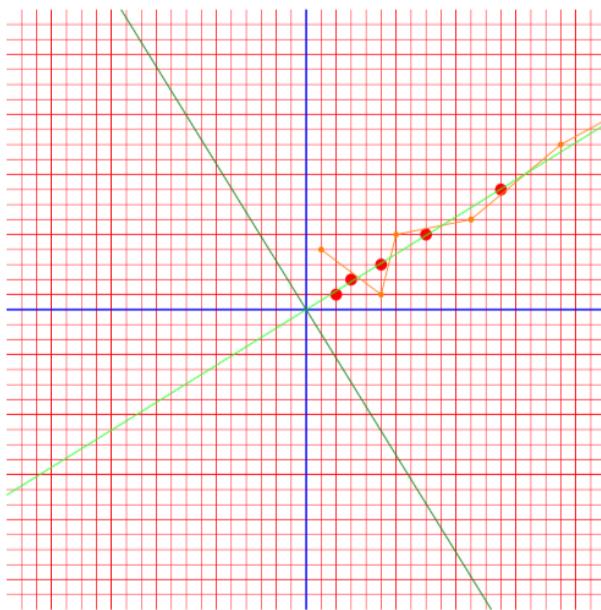
wat betekent dat elk volgend getal de som is van de twee vorige.

Bestaat er een gesloten formule voor het  $n$ -de Fibonaccigetal  $u_n$ ?

Als we opeenvolgende Fibonaccitallen een 2-vector laten vormen, dan is

$$\begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_n \\ u_{n-1} \end{bmatrix}.$$

Deze getallenparen kunnen we grafisch voorstellen als



Voor de matrix  $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$  geldt:  $\begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix} = A^n \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ . We berekenen de eigenwaarden en eigenvectoren van  $A$ . Hier is

$$\chi_A(x) = x^2 - x - 1 = \left( x - \frac{1 + \sqrt{5}}{2} \right) \left( x - \frac{1 - \sqrt{5}}{2} \right).$$

De twee eigenwaarden zijn dus<sup>1</sup>

$$\lambda_{\pm} = \frac{1 \pm \sqrt{5}}{2}.$$

Een stel bijhorende eigenvectoren worden gegeven door:

$$\vec{v}_+ = \begin{bmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{bmatrix} \quad \text{en} \quad \vec{v}_- = \begin{bmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{bmatrix}.$$

We weten dat twee eigenvectoren van twee verschillende eigenwaarden een basis vormen van  $\mathbb{R}^2$ . Elke vector kan dus ontbonden worden t.o.v. deze basis. We vinden hier:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{5}}\vec{v}_+ - \frac{1}{\sqrt{5}}\vec{v}_-.$$

Bijgevolg hebben we

$$\begin{aligned} \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix} &= A^n \cdot \left( \frac{1}{\sqrt{5}}\vec{v}_+ - \frac{1}{\sqrt{5}}\vec{v}_- \right) \\ &= \frac{1}{\sqrt{5}} \left( (\lambda_+)^n \vec{v}_+ - (\lambda_-)^n \vec{v}_- \right) = \begin{bmatrix} \dots \\ \frac{1}{\sqrt{5}} \left( (\lambda_+)^n - (\lambda_-)^n \right) \end{bmatrix}. \end{aligned}$$

Een gesloten formule voor het  $n$ -de Fibonacci getal is dus

$$u_n = \frac{1}{\sqrt{5}} \left( \left( \frac{1+\sqrt{5}}{2} \right)^n - \left( \frac{1-\sqrt{5}}{2} \right)^n \right).$$

Bijvoorbeeld, het 36e getal in de reeks is  $u_{36} = 14\,930\,352$ .

**Een tewerkstellingsmodel.** Beschouw opnieuw het tewerkstellingsmodel dat we in paragraaf 2 van Hoofdstuk 2 bespraken. De matrices die ons interesseerden waren

$$A = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \quad \text{en} \quad \vec{x}_0 = \begin{bmatrix} 8000 \\ 2000 \end{bmatrix}.$$

We zijn nu in staat om de vragen op te lossen die we ons op het eind van die paragraaf stelden. De eigenwaarden van  $A$  zijn nulpunten van

$$\chi_A(x) = x^2 - \frac{3}{2}x + \frac{1}{2} = (x-1) \left( x - \frac{1}{2} \right),$$

---

<sup>1</sup> $\lambda_+$  is de *gulden snede!*

dus  $\lambda_1 = 1$  en  $\lambda_2 = \frac{1}{2}$ . De bijhorende eigenruimten zijn:

$$E_1(A) = \left\{ k \begin{bmatrix} 2 \\ 3 \end{bmatrix} \mid k \in \mathbb{R} \right\} \quad \text{en} \quad E_{\frac{1}{2}}(A) = \left\{ l \begin{bmatrix} 1 \\ -1 \end{bmatrix} \mid l \in \mathbb{R} \right\}.$$

De totale bevolking bestaat uit 10 000 mensen. De verdeling onder werkenden en werkzoekenden is dus steeds van het type

$$\begin{aligned} \vec{x}_0 &= \begin{bmatrix} a \\ 10000 - a \end{bmatrix} \\ &= 2000 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + (a - 4000) \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \end{aligned}$$

voor zekere  $a$ . In de laatste gelijkheid maakten we gebruik van de eigenschap dat 2 eigenvectoren van 2 verschillende eigenwaarden een basis voor  $\mathbb{R}^2$  vormen.

Na  $n$  jaren vinden we

$$\begin{aligned} \vec{x}_n &= A^n \cdot \begin{bmatrix} a \\ 10000 - a \end{bmatrix} = A^n \cdot \left( 2000 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + (a - 4000) \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right) \\ &= 2000 (1)^n \begin{bmatrix} 2 \\ 3 \end{bmatrix} + (a - 4000) \left( \frac{1}{2} \right)^n \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \end{aligned}$$

en dus is

$$\vec{x}_\infty = \lim_{n \rightarrow \infty} (\vec{x}_n) = 2000 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4000 \\ 6000 \end{bmatrix}.$$

Klaarblijkelijk is deze verdeling onafhankelijk van de beginverdeling  $a$ , iets dat we ten tijde van Hoofdstuk 2 reeds vermoedden.

## 4 De spectraalstelling

In deze sectie zullen we even werken met complexe getallen  $a + bi$ , met natuurlijk  $i^2 = -1$ . Herinner dat de *complex toegevoegde* van het complexe getal  $a + ib$  het complexe getal

$$\overline{a + ib} = a - ib$$

is. Er geldt steeds dat het product  $(a + ib)(\overline{a + ib}) = a^2 + b^2$  reëel en positief is. We kunnen ook matrices over  $\mathbb{C}$  beschouwen. Bijvoorbeeld,

$$\begin{bmatrix} 1 \\ 2i \\ 0 \end{bmatrix} \quad \text{en} \quad \begin{bmatrix} 4 & 2i+1 \\ 3i & i-1 \end{bmatrix}$$

zijn voorbeelden van een kolomvector in  $\mathbb{C}^3$  en een matrix over  $\mathbb{C}$ .

Merk op dat eigenwaarden niet noodzakelijk reëel hoeven te zijn, zelfs niet voor reële matrices. Beschouw bijvoorbeeld de matrix  $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ . De eigenwaarden zijn de oplossingen van de karakteristieke vergelijking  $\lambda^2 + 1 = 0$ , en dus gelijk aan  $\lambda = \pm i$ , wat natuurlijk complexe getallen zijn.

In de volgende definitie beschouwen we twee kolomvectoren

$$\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}, \quad \vec{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix},$$

met coëfficiënten  $v_i, w_i$  in  $\mathbb{C}$ .

**Definitie 15.** Het hermitisch product van twee kolomvectoren  $\vec{v}, \vec{w} \in \mathbb{C}^n$  wordt gegeven door

$$(\vec{v}, \vec{w}) = \bar{v}_1 w_1 + \bar{v}_2 w_2 + \dots + \bar{v}_n w_n \in \mathbb{C}.$$

**Eigenschappen.** (1) Voor elke  $\vec{v} \neq \vec{0} \in \mathbb{C}^n$  geldt dat

$$(\vec{v}, \vec{v}) = \bar{v}_1 v_1 + \dots + \bar{v}_n v_n > 0.$$

Merk op dat het gewone scalaire product voor  $\mathbb{C}^n$  deze eigenschap niet heeft want bvb. met

$$\vec{v} = \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{dan} \quad \vec{v} \cdot \vec{v} = 1 \cdot 1 + i \cdot i = 0 \quad \text{maar} \quad (\vec{v}, \vec{v}) = \bar{1}1 + \bar{i}i = 1 - i^2 = 2.$$

(2) Indien  $\vec{v}, \vec{w} \in \mathbb{R}^n$  reële vectoren zijn dan is  $\vec{v} \cdot \vec{w} = (\vec{v}, \vec{w})$ .

(3) Er geldt voor alle  $\vec{v}, \vec{w} \in \mathbb{C}^n$  en  $\lambda \in \mathbb{C}$ :

$$(\lambda \vec{v}, \vec{w}) = \bar{\lambda}(\vec{v}, \vec{w}) \quad \text{en} \quad (\vec{v}, \lambda \vec{w}) = \lambda(\vec{v}, \vec{w}).$$

(4) Voor een reële vierkante matrix  $B$  geldt

$$(B \cdot \vec{v}, \vec{w}) = (\vec{v}, B^{tr} \cdot \vec{w}),$$

want  $(B.\vec{v}, \vec{w}) = \sum_{i=1}^n (\sum_{j=1}^n \overline{B_{ij}v_j}) w_i = \sum_{j=1}^n \bar{v}_j (\sum_{i=1}^n B_{ij}w_i)$ , en dit is het rechterlid.

Voor de rest van de paragraaf beschouwen we *symmetrische reële matrices*. Dit zijn vierkante matrices  $A$ , met de eigenschap  $A^{tr} = A$ . Symmetrische matrices treden heel natuurlijk op in vele toepassingen: in *meetkunde* als veralgemeende inproducten, in *statistiek* als *correlatiematrizes*, in *grafentheorie* als *adjacency matrices*, etc.

**Stelling 38.** *Als  $A$  een symmetrische reële matrix is dan zijn alle eigenwaarden van  $A$  reële getallen.*

*Bewijs.* Voor een eigenwaarde  $\lambda$  van  $A$ , met eigenvector  $\vec{v} \neq \vec{0}$ , geldt:

$$\bar{\lambda}(\vec{v}, \vec{v}) = (A.\vec{v}, \vec{v}) = (\vec{v}, A^{tr}.\vec{v}) = (\vec{v}, A.\vec{v}) = (\vec{v}, \lambda\vec{v}) = \lambda(\vec{v}, \vec{v}).$$

Wegens  $(\vec{v}, \vec{v}) > 0$  moet dus  $\lambda = \bar{\lambda}$ . Dit betekent dat  $\lambda$  een reëel getal is.  $\square$

**Stelling 39.** *Voor een symmetrische reële matrix  $A$  staan eigenvectoren  $\vec{v}$  en  $\vec{w}$  horende bij verschillende eigenwaarden loodrecht op elkaar.*

*Bewijs.* Laat  $A.\vec{v} = \lambda\vec{v}$  en  $A.\vec{w} = \mu\vec{w}$  met  $\lambda \neq \mu \in \mathbb{R}$ . Dan is

$$\lambda\vec{v}.\vec{w} = \lambda(\vec{v}, \vec{w}) = (A.\vec{v}, \vec{w}) = (\vec{v}, A.\vec{w}) = (\vec{v}, \mu\vec{w}) = \mu\vec{v}.\vec{w}$$

en, omdat  $\lambda \neq \mu$  moet dus  $\vec{v}.\vec{w} = 0$ .  $\square$

Herinner dat een orthogonale  $n \times n$  matrix  $S$  voldoet aan  $S^{tr}.S = I_n$ , en dus aan  $S^{-1} = S^{tr}$ .

**Stelling 40** (Spectraalstelling). *Een reële symmetrische matrix  $A$ , met eigenwaarden  $\lambda_i$ , kan gediagonaaliseerd worden met een orthogonale basisverandering, dat is, er bestaat een orthogonale matrix  $S$  zodat*

$$S^{-1}.A.S = S^{tr}.A.S = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

*Bewijs.* We bewijzen enkel het geval waarbij  $A$  **n verschillende eigenwaarden heeft**. Het bewijs voor het algemene geval hoef je niet te kennen.

In dat geval volgt uit Stelling 34 dat  $A$  diagonaliseerbaar is. Veronderstel dat de eigenvectoren  $\vec{v}_1, \dots, \vec{v}_n$  allen lengte 1 hebben (indien niet, vervang ze door  $\frac{1}{\|\vec{v}_i\|}\vec{v}_i$ ). Dan is, wegens de

vorige stelling,  $\{\vec{v}_1, \dots, \vec{v}_n\}$  een orthonormaal stel in  $\mathbb{R}^n$  (en een basis wegens Stelling 33). Hiermee is  $S = [\vec{v}_1 \ \dots \ \vec{v}_n]$  een orthogonale matrix, want

$$S^{tr}.S = \begin{bmatrix} (\vec{v}_1)^{tr} \\ \vdots \\ (\vec{v}_n)^{tr} \end{bmatrix} \cdot \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_n \end{bmatrix} = \begin{bmatrix} \vec{v}_1 \cdot \vec{v}_1 & \dots & \vec{v}_1 \cdot \vec{v}_n \\ \vdots & & \vdots \\ \vec{v}_n \cdot \vec{v}_1 & \dots & \vec{v}_n \cdot \vec{v}_n \end{bmatrix} = I_n.$$

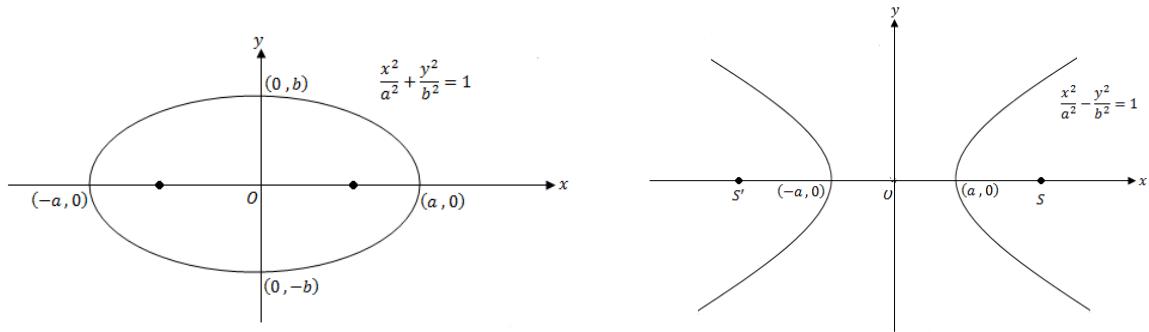
□

De benaming van deze stelling komt uit de quantummechanica waar het *spectrum* van een stralingsbron (de absorptie en emissie lijnen) geassocieerd wordt aan de eigenwaarden van een matrix.

**Een toepassing op ellipsen en hyperbolen.** De standaardvergelijkingen van een ellips en een hyperbool zijn, respectievelijk,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad \text{en} \quad \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

Bij deze uitdrukkingen zijn de  $x$ - en de  $y$ -as net de (orthogonale) symmetrieassen van de krommen.



Voor een willekeurige kwadratische vergelijking, kan men zich dus de vraag stellen:

*Stelt de vergelijking*

$$5x^2 + 8xy + 5y^2 = 1$$

*ook een ellips of een hyperbool voor?*

Noteer  $\vec{x} = [x \ y]^tr$ . Het linkerlid van de vergelijking is dan van de vorm:

$$5x^2 + 8xy + 5y^2 = [x \ y] \cdot \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \vec{x}^{tr} \cdot A \cdot \vec{x}.$$

De matrix  $A$  is symmetrisch. Wegens de spectraalstelling heeft ze reële eigenwaarden en is ze diagonaliseerbaar via een orthogonale matrix  $S$  van eigenvectoren. De eigenwaarden van  $A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$  zijn de oplossingen van  $(5 - \lambda)^2 - 16 = 0$ , en zijn dus hier

$$\lambda_1 = 9, \quad \lambda_2 = 1.$$

Corresponderende eigenvectoren (met lengte 1) zijn:

$$\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Dus kunnen we  $A$  ontbinden als:

$$A = S.D.S^{tr} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Het linkerlid van onze kwadratische vergelijking wordt hiermee:

$$\begin{aligned} 5x^2 + 8xy + 5y^2 &= \vec{x}^{tr} \cdot (S.D.S^{tr}) \cdot \vec{x} = (S^{tr} \cdot \vec{x})^{tr} \cdot D \cdot (S^{tr} \cdot \vec{x}) \\ &= 9 \left( \frac{x+y}{\sqrt{2}} \right)^2 + 1 \left( \frac{y-x}{\sqrt{2}} \right)^2. \end{aligned}$$

Stel nu

$$S^{tr} \cdot \vec{x} = \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix}, \quad \text{met } \tilde{x} = \frac{x+y}{\sqrt{2}}, \quad \tilde{y} = \frac{y-x}{\sqrt{2}},$$

dan is:

$$5x^2 + 8xy + 5y^2 = 1 \quad \Leftrightarrow \quad \frac{\tilde{x}^2}{(1/3)^2} + \frac{\tilde{y}^2}{1^2} = 1.$$

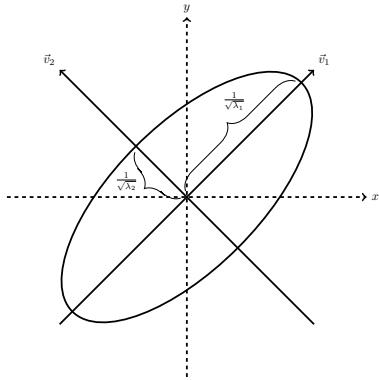
Onze vergelijking stelde dus een ellips voor met kleine as  $\frac{1}{\sqrt{\lambda_1}} = \frac{1}{3}$  en grote as  $\frac{1}{\sqrt{\lambda_2}} = 1$ . Essentieel gaat het hier dus om een ellips omdat beide eigenwaarden positief waren.

We kunnen ons resultaat nog verder interpreteren. Merk op dat

$$\begin{aligned} \vec{x} &= \begin{bmatrix} x \\ y \end{bmatrix} = x\vec{e}_1 + y\vec{e}_2 \\ &= \left( \frac{x+y}{\sqrt{2}} \right) \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \left( \frac{y-x}{\sqrt{2}} \right) \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= \tilde{x}\vec{v}_1 + \tilde{y}\vec{v}_2. \end{aligned}$$

De vector  $\vec{x}$  heeft dus componenten  $(\tilde{x}, \tilde{y})$  t.o.v. de (orthonormale) basis van eigenvectoren  $\{\vec{v}_1, \vec{v}_2\}$ .

De symmetrie-assen van een ellips corresponderen met die punten van de ellips met ofwel  $\tilde{x} = 0$ , ofwel  $\tilde{y} = 0$ . Dus: met de (loodrechte) richtingen  $\vec{v}_1$  en  $\vec{v}_2$ .



We bereiken de volgende conclusies:

- De orthogonale eigenvectoren  $\vec{v}_1$  en  $\vec{v}_2$  geven de richting van de symmetrieassen aan.
- De eigenwaarden  $\lambda_1$  en  $\lambda_2$  bepalen de waarden van de grote en de kleine as.

## 5 De Jordan-normaalvorm

In vorige lessen hebben we gezien dat een  $n \times n$  matrix  $A$  met  $n$  verschillende eigenwaarden gediagonaaliseerd kan worden. Als dit het geval is weten we dat de diagonalmatrix als componenten de eigenwaarden van  $A$  heeft. Echter, niet elke matrix kan worden gediagonaaliseerd. Bekijk, bijvoorbeeld, de  $2 \times 2$  matrix

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

Haar karakteristieke veelterm is  $\chi_A(x) = x^2$ , en dus heeft  $A$  als enige eigenwaarde 0 die voorkomt met multipliciteit twee. Indien  $A$  gediagonaaliseerd kan worden dan zou er een inverteerbare  $2 \times 2$  matrix  $S$  bestaan waarvoor

$$S^{-1} \cdot \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot S = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{met} \quad S = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Dat zou betekenen dat ook  $A.S = 0$ , maar

$$A.S = \begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix},$$

en dus zou  $c = d = 0$  moeten zijn. Maar, dan kan  $S$  niet inverteerbaar zijn.

De stelling hieronder zegt dat elke  $n \times n$  matrix  $A$  over de complexe getallen  $\mathbb{C}$  in een speciale vorm gebracht kan worden, de zogenaamde *Jordan-normaalvorm*.

**Definitie 16.** Voor  $\lambda \in \mathbb{C}$  is een *Jordanblok* met eigenwaarde  $\lambda$  en size  $d$  de  $d \times d$  matrix

$$J_d(\lambda) = \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix},$$

dus met alle diagonaalcomponenten gelijk aan  $\lambda$  en de eerste boven-diagonaalcomponenten allen gelijk aan 1, en de overige componenten gelijk aan 0.

**Stelling 41** (Jordan-normaalvorm). (Over de complexe getallen  $\mathbb{C}$ ) Iedere  $n \times n$  matrix  $A$  met karakteristieke veelterm

$$\chi_A(x) = \prod_{i=1}^k (x - \lambda_i)^{m_i} \quad \text{met} \quad \lambda_i \neq \lambda_j \text{ als } i \neq j$$

kan met een inverteerbare matrix  $S$  in volgende vorm gebracht worden

$$S^{-1}.A.S = \begin{bmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & J_l \end{bmatrix}$$

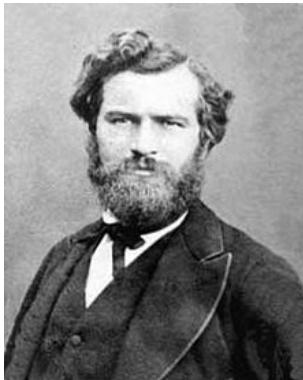
waar elke  $J_j$  een *Jordanblok* is horende bij een eigenwaarde van  $A$  en met size  $d_j$  zodat dat als  $\{j_1, \dots, j_{u_i}\}$  de *Jordan blokken* zijn die horen bij eigenwaarde  $\lambda_i$  we hebben dat

$$m_i = d_{j_1} + \dots + d_{j_{u_i}}$$

Het bewijs van deze stelling is technisch en wordt om die reden weggelaten.

**Historische noot.** De Jordan uit de Jordan-normaalvorm is *niet* dezelfde Jordan van de Gauss-Jordaneliminatie. De normaalvorm werd bewezen door de Franse wiskundige [Camille](#)

Jordan (1838-1922).



## 6 Markovmatrices en het PageRank-algoritme

In de beginlagen van het internet kon men de basisprincipes van elke search engine samenvatten als:

1. Lokaliseer alle webpagina's die publiek toegankelijk zijn.
2. Indexeer de data uit stap 1, op zo'n wijze dat ze efficiënt onderzocht kan worden naar relevante keywords and strings.

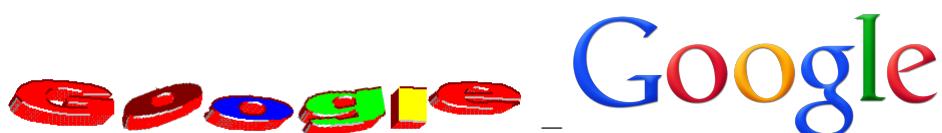
In de late jaren 90 constateerde men echter het volgende probleem:

*“.... Anyone who has used a search engine recently, can readily testify that **the completeness of the index is not the only factor** in the quality of search results. ‘Junk results’ often wash out any results that a user is interested in. In fact, as of November 1997, only one of the top four commercial search engines finds itself (returns its own search page in response to its name in the top ten results). ... People are still only willing to look at the first few tens of results. Because of this, as the collection size grows, we need **tools that have very high precision** (number of relevant documents returned, say in the top tens of results).... ”*

Deze zin komt uit de paragraaf 'Design Goals' van:

S. Brin and L. Page, *The Anatomy of a Large-Scale Hypertextual Web Search Engine*. In: 7th Int. WWW Conf., 1998, Brisbane.

Brin en Page zijn natuurlijk twee van de grondleggers van



Vanaf haar lancering was Google duidelijk beter dan haar concurrenten (Lycos, Yahoo, Altavista, etc): je scheen je niet te moeten worstelen doorheen lange lijsten van irrelevante webpagina's die toevallig ook de gezochte string bevatten.

De basisprincipes van elke search engine kunnen daarom beter als volgt aangepast worden:

1. Lokaliseer alle webpagina's die publiek toegankelijk zijn.
2. Indexeer de data uit stap 1, op zo'n wijze dat ze efficiënt onderzocht kan worden naar relevante keywords and strings.
3. Geef een gewicht aan elke pagina in je database, zodat, wanneer de deelverzameling van webpagina's met de gevraagde informatie gevonden zijn, de belangrijkste pagina's eerst gepresenteerd kunnen worden.

In het geval van Google heet het algoritme dat het gewicht uit stap 3 bepaalt *PageRank*. In de volgende paragrafen gaan we dieper in op de ideeën die aan de basis van PageRank liggen. PageRank is natuurlijk niet het enige beschikbare algoritme voor stap 3, en, voor Google, is het (tegenwoordig) ook niet de enige factor in hun beslissing om bepaalde pagina's eerst te tonen. Bijkomende info vind je in:

- Het oorspronkelijke artikel:

L. Page, S. Brin, M. Rajeev en W. Terry, *The PageRank Citation Ranking: Bringing Order to the Web*, Technical report, Stanford University (1998). Download [hier](#).

- Een leesbaar artikel met meer info:

K. Bryan and T. Leise, *The \$25,000,000,000 Eigenvector, The linear Algebra behind Google*, SIAM Review 48 (2006) 569–581. Bijvoorbeeld [hier](#) te downloaden.

**De vectorruimte van het World Wide Web en de Googlematrix  $G$ .** Beschouw alle webpagina's op het internet, en catalogeer ze in een database. Noem de pagina's  $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_n$ . (Volg [deze link](#) om een idee van  $n$  te krijgen.) We identificeren  $\vec{w}_1$  met de vector  $\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \end{bmatrix}^{tr}$ ,  $\vec{w}_2$  met de vector  $\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \end{bmatrix}^{tr}$ , etc.

Beschouw nu  $V = \text{span}(\vec{w}_1, \vec{w}_2, \dots, \vec{w}_n) = \mathbb{R}^n$ , de ruimte met als basis “de” webpagina's. Een vector in  $V$  is dus een “lineaire combinatie van webpagina's”:

$$\vec{v} = \sum_{i=1}^n \alpha_i \vec{w}_i.$$

Hoe moeten we een dergelijk vector interpreteren?

- Als één  $\alpha_j = 1$  en alle andere  $\alpha_i = 0$  ( $i \neq j$ ), dan is de vector  $\vec{v}$  een webpagina, namelijk de webpagina  $\vec{w}_j$ .
- In de andere gevallen kan je  $\vec{v}$  niet interpreteren als een “echte” webpagina.
- Als alle  $\alpha_i \geq 0$  en  $\sum_{i=1}^n \alpha_i = 1$  is er toch een interpretatie voor  $\vec{v}$  mogelijk, namelijk als “kansvector”: *de coëfficiënt  $\alpha_i$  geeft de kans aan dat je (bv. door een link te volgen) op de webpagina  $\vec{w}_i$  terecht komt.*

Met andere woorden: *Op de “virtuele” webpagina  $\vec{v} = \sum_{i=1}^n \alpha_i \vec{w}_i$  ben je met kans  $\alpha_i$  op webpagina  $\vec{w}_i$ .* Google’s algoritme gebruikt dit idee om een rangorde op de webpagina’s te definiëren.

Hiertoe nemen we, vanaf nu, eerst een hypothese aan. Stel je voor dat iemands browser zich op de webpagina  $\vec{w}_j$  bevindt en dat er links op die pagina staan. Ofwel zal die persoon één van die links volgen, ofwel niet. De hypothese is dat we de kansverdeling hiervoor op voorhand kennen: we nemen aan dat de persoon

- *met een zekere kans  $p$  (bv.  $p = 85\%$ ) een link op die pagina zal volgen.*
- *in het andere geval, met kans  $1 - p$  (= 15%), “random” naar een andere webpagina op het internet zal surfen (door een nieuw webadres in te tikken).*

Het volgende voorbeeld maakt dit idee expliciet. Stel je voor dat je je bevindt op de webpagina  $\vec{w}_j$ , en dat er op die pagina 5 links zijn, namelijk naar de webpagina’s  $\vec{w}_{i_1}, \vec{w}_{i_2}, \dots, \vec{w}_{i_5}$ .

- De kans dat je hierna (na één klik of na het intikken van een nieuw adres) op pagina  $\vec{w}_i$  terecht komt, met  $i \in \{i_1, \dots, i_5\}$  is

$$g_{ij} = 0.85 \cdot \frac{1}{5} + 0.15 \cdot \frac{1}{n}.$$

- De kans dat je op een webpagina  $\vec{w}_l$  terecht komt, met  $l \notin \{i_1, \dots, i_5\}$  is

$$g_{lj} = 0.15 \frac{1}{n}.$$

Dit leidt tot volgend algemeen principe. Je bevindt je op de webpagina  $\vec{w}_j$ , en op die pagina zijn er  $k_j$  links. De kans dat je na een klik op pagina  $\vec{w}_i$  terecht komt is

$$g_{ij} = pm_{ij} + (1 - p) \frac{1}{n},$$

met

$$m_{ij} = \begin{cases} \frac{1}{k_j} & \text{als } k_j \neq 0 \text{ en er een link is van } \vec{w}_j \text{ naar } \vec{w}_i, \\ 0 & \text{als } k_j \neq 0 \text{ en er geen link is van } \vec{w}_j \text{ naar } \vec{w}_i, \\ \frac{1}{n} & \text{als } k_j = 0. \end{cases}$$

Merk op dat, in het laatste geval waarbij  $k_j = 0$ , de kans dat je van  $\vec{w}_j$  naar  $\vec{w}_i$  gaat inderdaad gelijk is aan  $p\frac{1}{n} + (1-p)\frac{1}{n} = \frac{1}{n}$ , zoals verwacht.

Alle getallen  $g_{ij}$ , voor alle  $i$  en voor alle  $j$ , verzamelen we nu samen in een matrix, de “Googlematrix”  $G = (g_{ij})$ . De kans dat je van webpagina  $\vec{w}_j = [0 \dots 1 \dots 0]^{tr}$  op pagina  $\vec{w}_i$  terecht is dan de  $i$ -de component  $g_{ij}$  van de vector

$$G.\vec{w}_j = \begin{bmatrix} g_{11} & \dots & g_{1n} \\ \vdots & \vdots & \vdots \\ g_{n1} & \dots & g_{nn} \end{bmatrix} \cdot \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} g_{1j} \\ \vdots \\ g_{nj} \end{bmatrix}.$$

De betekenis van de matrix  $G$  is dus dat  $G$  aangeeft hoe we “door een link te volgen/door een nieuw adres in te tikken” van de ene webpagina op de andere komen. Als we op de webpagina  $\vec{w}$  zijn dan gaan we door een link te volgen of door een nieuw adres in te tikken naar de “virtuele webpagina”  $G.\vec{w}$ . Dit is een vector wiens componenten de kans aangeven dat we in de volgende stap op een bepaalde webpagina terechtkomen.

Het proces waarbij we in verschillende stappen surfen op het web, startend van webpagina  $\vec{w}$ , kunnen we dan als volgt opvatten:

- Na één klik op een link of na het intikken van een nieuw webadres wordt de kans om op eender welke pagina terecht te komen gegeven door  $G.\vec{w}$ ,
- na twee stappen door  $G.G.\vec{w}$ ,
- na  $k$  stappen door  $G^k.\vec{w}$ .

Om de toestand te kennen na “lang” surfen moet je dus  $\lim_{k \rightarrow +\infty}(G^k.\vec{w})$  berekenen. Dat is natuurlijk een typisch vraagstuk uit de Lineaire Algebra! We moeten eerst enkele bijkomende stellingen bewijzen.

### Markovmatrices en kansvectoren.

**Definitie 17.** Een vector  $\vec{v} = [\alpha_1 \dots \alpha_n]^{tr}$  is een kansvector indien al haar componenten getallen  $\alpha_i$  uit  $[0, 1]$  zijn en de som van alle componenten gelijk is aan 1, i.e.  $\sum_{i=1}^n \alpha_i = 1$ .

Een vierkante matrix  $A$  is een Markovmatrix indien al haar kolommen kansvectoren zijn.

**Stelling 42.** Zij  $A$  een Markovmatrix en  $\vec{v}$  een kansvector. Dan zijn  $A.\vec{v}$  en  $A^k.\vec{v}$  ook kansvectoren.

*Bewijs.* De  $i$ -de component van  $A.\vec{v}$  is  $\sum_{k=1}^n A_{ik}\alpha_k$ . De som van al die componenten is dus

$$\sum_{i=1}^n \sum_{k=1}^n A_{ik}\alpha_k = \sum_{k=1}^n \alpha_k \sum_{i=1}^n A_{ik} = \sum_{k=1}^n \alpha_k = 1.$$

De 2de eigenschap volgt uit de vorige.  $\square$

De “Google-matrix”  $G$  is een dergelijke Markov-matrix. Er geldt immers voor een kolom  $j$ :

- Als  $k_j = 0$ , dan is  $\sum_{i=1}^n g_{ij} = \sum_{i=1}^n \frac{1}{n} = n \frac{1}{n} = 1$ ,
- Als  $k_j \neq 0$ , dan is

$$\sum_{i=1}^n a_{ij} = p \sum_{i=1}^n m_{ij} + (1-p) \sum_{i=1}^n \frac{1}{n} = p \left( \sum_{i=1}^{k_j} \frac{1}{k_j} \right) + 1 - p = 1.$$

Uit de vorige stelling halen we dus dat als we starten van een webpagina  $\vec{w}_i$  (= een kansvector), ook  $G.\vec{w}_i$  een kansvector zal zijn.

We bewijzen enkele verdere eigenschappen van Markovmatrices.

**Stelling 43.** Een Markovmatrix  $A$  heeft een niet-nul eigenvector met eigenwaarde 1.

*Bewijs.* Voor  $A^{tr}$  is de totale som van alle rijen 1 en dus is

$$A^{tr} \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} + a_{21} + \dots + a_{n1} \\ \vdots \\ a_{1n} + a_{2n} + \dots + a_{nn} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Er volgt dat  $A^{tr}$  een niet-nul eigenvector heeft, met eigenwaarde 1, en dus ook  $A$  (De eigenwaarden van  $A$  en  $A^{tr}$  zijn immers steeds dezelfde).  $\square$

De matrix  $P$  van het voetbalprobleem, waarmee we dit hoofdstuk startten, is een voorbeeld van een Markovmatrix. We hadden toen al opgemerkt dat die matrix een eigenwaarde 1 had, en dat de eigenvectoren bij die eigenwaarde een belangrijke rol spelen bij het verloop van het vraagstuk op lange termijn. We zullen straks zien dat een dergelijke rol ook weggelegd is voor de corresponderende eigenvectoren van de Googlematrix.

**Stelling 44.** Voor alle eigenwaarden van een Markovmatrix geldt  $|\lambda_i| \leq 1$  (over  $\mathbb{C}$ ).

*Bewijs.* Zij  $\lambda$  een eigenwaarde van  $A$ , dan is het ook een eigenwaarde van  $A^{tr}$ . Zij  $\vec{z} = [\beta_1 \dots \beta_n]^{tr}$  een corresponderende eigenvector voor  $A^{tr}$ .

Beschouw de component  $\beta_k$  waarvoor  $|\beta_j| \leq |\beta_k|$ , voor alle  $j$ .

Voor die component geldt:  $\sum_{j=1}^n a_{jk}\beta_j = \lambda\beta_k$  (Let op de volgorde van  $j$  en  $k$ !).

Dus is

$$\begin{aligned} |\lambda||\beta_k| &= |\lambda\beta_k| = \left| \sum_{j=1}^n a_{jk}\beta_j \right| \\ &\leq \sum_{j=1}^n a_{jk}|\beta_j| \leq \left( \sum_{j=1}^n a_{jk} \right) |\beta_k| = |\beta_k|. \end{aligned}$$

Vandaar is  $|\lambda| \leq 1$ . □

Men kan dit nog verder verfijnen tot

**Stelling 45.** Een Markovmatrix  $A$  heeft een eigenwaarde  $\lambda_1 = 1$ . Voor alle andere eigenwaarden geldt  $|\lambda_i| < 1$ .

In de volgende paragraaf hebben we dit resultaat nodig, maar het bewijs van deze stelling valt buiten het bestek van deze cursus.

De Googlematrix  $G$  heeft echter een bijkomende eigenschap: alle elementen  $g_{ij}$  zijn *strikt* groter dan nul. Tot slot bewijzen we nog enkele eigenschappen voor Markovmatrices met  $a_{ij} > 0$ .

**Stelling 46.** Zij  $A$  een Markovmatrix met alle  $a_{ij} > 0$ . Elke eigenvector  $\vec{z} = [\beta_1 \dots \beta_n]^{tr} \neq \vec{0}$  van  $E_1(A)$  heeft ofwel allemaal strikt positieve ofwel allemaal strikt negatieve componenten  $\beta_i$ .

*Bewijs.* We nemen aan dat  $|\sum_{i=1}^n \beta_i| < \sum_{i=1}^n |\beta_i|$  en tonen een strijdigheid aan. Omdat  $\vec{z} \in E_1(A)$  is  $\beta_i = \sum_{j=1}^n a_{ij}\beta_j$  en dus:

$$|\beta_i| = \left| \sum_{j=1}^n a_{ij}\beta_j \right| < \sum_{j=1}^n a_{ij}|\beta_j|.$$

Na sommeren over  $i$  vinden we een strijdigheid:

$$\sum_{i=1}^n |\beta_i| < \sum_{j=1}^n |\beta_j| \left( \sum_{i=1}^n a_{ij} \right) = \sum_{j=1}^n |\beta_j|.$$

Dus zijn ofwel alle  $\beta_i \geq 0$  ofwel alle  $\beta_i \leq 0$ . Maar omdat  $\beta_i = \sum_{j=1}^n a_{ij}\beta_j$  en  $a_{ij} > 0$  is  $\beta_i > 0$  of  $\beta_i < 0$  voor alle  $i$ .  $\square$

Merk op dat, als alle  $\beta_i > 0$ , dan is ook  $\sum_{i=1}^n \beta_i > 0$ . Als alle  $\beta_i < 0$ , dan is ook  $\sum_{i=1}^n \beta_i < 0$ . In beide gevallen is  $\sum_{i=1}^n \beta_i \neq 0$ .

**Stelling 47.** *Zij  $A$  een Markovmatrix met alle  $a_{ij} > 0$ . De eigenruimte  $E_1(A)$  is één-dimensionaal.*

*Bewijs.* Veronderstel dat het niet zo is, en dat er dus minstens twee linear onafhankelijke eigenvectoren  $\vec{z}_1 \neq \vec{0}$  en  $\vec{z}_2 \neq \vec{0}$  bestaan. Dan is, voor alle keuzes van  $s$ , ook  $\vec{z}_3 = \vec{z}_1 + s\vec{z}_2$  een eigenvector, verschillend van  $\vec{0}$ . We weten dat  $b_1 = \sum_{i=1}^n \beta_{1i} \neq 0$  en  $b_2 = \sum_{i=1}^n \beta_{2i} \neq 0$ . Neem nu  $s = -b_1/b_2$ , dan is  $\sum_{i=1}^n \beta_{3i} = 0$ , wat een strijdigheid is.  $\square$

**Terug naar PageRank.** We hebben nu alle eigenschappen verzameld om het PageRank-algoritme te begrijpen. Stel dat  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$  een eigenbasis is voor  $G$ , en ontbind een webpagina  $\vec{w}$  als  $\vec{w} = \sum_{i=1}^n \mu_i \vec{v}_i$ . Dan is, wegens  $\lambda_1 = 1$  en  $|\lambda_i| < 1$  voor de andere,

$$\lim_{k \rightarrow +\infty} (G^k \cdot \vec{w}) = \lim_{k \rightarrow +\infty} \left( \sum_{i=1}^n \mu_i (\lambda_i)^k \vec{v}_i \right) = \mu_1 \vec{v}_1.$$

Dus ligt  $\lim_{k \rightarrow +\infty} (G^k \cdot \vec{w}) \in E_1(G)$ .

Hierbij kunnen we twee opmerkingen formuleren. Een webpagina  $\vec{w}$  is een kansvector, en dus is ook  $\lim_{k \rightarrow +\infty} (G^k \cdot \vec{w})$  een kansvector. Ten tweede, aangezien alle eigenvectoren van  $E_1(G)$  allemaal positieve of allemaal negatieve componenten hebben, ligt er binnen de één-dimensionale ruimte  $E_1(G)$  net één kansvector. We concluderen:

*Ongeacht de keuze van de startpagina  $\vec{w}$ , komen we na lang surfen steeds terecht op dezelfde virtuele pagina  $\vec{v}_\infty := \lim_{k \rightarrow +\infty} (G^k \cdot \vec{w}) \in E_1(G)$ .*

Het algoritme werkt daarom als volgt:

- Bepaal de componenten van  $\vec{v}_\infty = (\alpha_1, \dots, \alpha_n)$ . Als je lang aan het surfen bent, kom je met kans  $\alpha_i$  op webpagina  $\vec{w}_i$  terecht!
- De component  $\alpha_i$  van de kanseigenvector  $\vec{v}_\infty$  noemt men de *Pagerank* van de webpagina  $\vec{w}_i$ .
- Stel  $\alpha_j = \max_{i=1, \dots, n} \{\alpha_i\}$ , dan is de kans het grootst dat je na lang surfen op webpagina  $\vec{w}_j$  terechtkomt.
- We kunnen nu de componenten  $\alpha_i$  van de eigenvector  $\vec{v}_\infty$  ordenen volgens grootte. Dit is de rangorde die Google gebruikt.



# 6 Vectorruimten en lineaire afbeeldingen

Tot nu toe hebben we gewerkt met vectoren in  $\mathbb{R}^n$  en met deelruimten van  $\mathbb{R}^n$ . Dit zijn allemaal typische voorbeelden van *vectorruimten*.

**Definitie 18.** Een reële vectorruimte is een verzameling  $V$ , waarvan de elementen vectoren genoemd worden, samen met twee afbeeldingen:  $+ : V \times V \rightarrow V$  en  $\cdot : \mathbb{R} \times V \rightarrow V$ , respectievelijk de optelling en de scalaire vermenigvuldiging, die voldoen aan de volgende axioma's voor elke  $\vec{u}, \vec{v}, \vec{w} \in V$  en voor elke  $a, b \in \mathbb{R}$ :

- (1)  $\vec{u} + \vec{v} = \vec{v} + \vec{u}$  (de optelling is commutatief),
- (2)  $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$  (de optelling is associatief),
- (3) er bestaat een vector  $\vec{0}$  in  $V$  zodat  $\vec{0} + \vec{u} = \vec{u}$  (deze  $\vec{0}$  noemen we de nulvector),
- (4) voor elke  $\vec{v} \in V$  bestaat een vector  $-\vec{v}$  zodat  $\vec{v} + (-\vec{v}) = \vec{0}$ ,
- (5)  $(ab).\vec{u} = a.(b.\vec{u})$ ,
- (6)  $(a + b).\vec{u} = a.\vec{u} + b.\vec{u}$ ,
- (7)  $a.(\vec{u} + \vec{v}) = a.\vec{u} + a.\vec{v}$ ,
- (8)  $0.\vec{u} = \vec{0}$
- (9)  $1.\vec{u} = \vec{u}$ .

In het vervolg schrijven we simpelweg  $a\vec{v}$  i.p.v.  $a.\vec{v}$ .

In een vectorruimte kunnen we ook de notie van *lineair onafhankelijk* en *voortbrengend* (en dus ook van *basis*) definiëren voor een stel vectoren  $\{\vec{v}_i \mid i \in I\} \subset V$ , hier is  $I$  een index-verzameling die eindig of oneindig mag zijn. Indien  $\{\vec{v}_i \mid i \in I\}$  een basis is voor  $V$  en  $\#I = n$ , dan is  $V$  een  $n$ -dimensionale vectorruimte. Er bestaan echter ook oneindig dimensionale vectorruimten.

**Voorbeeld.** Een oneindig dimensionale vectorruimte. Neem  $V = \mathbb{R}[x]$  de verzameling van alle veeltermen in  $\mathbb{R}$ , dat is, alle veeltermen

$$\vec{f} = \lambda_0 + \lambda_1 x + \lambda_2 x^2 + \dots + \lambda_k x^k$$

voor alle mogelijke graden  $k \in \mathbb{N}$  en alle coefficiënten  $\lambda_i \in \mathbb{R}$ . De nul-vector komt overeen met de nul-veelterm 0 ∈  $\mathbb{R}[x]$  en men gaat na dat het stel veeltermen

$$\{\vec{e}_0 = 1, \vec{e}_i = x^i, \forall i \geq 1\}$$

een (oneindige) basis is voor  $V$ . Merk op dat na keuze van deze basis we  $\mathbb{R}[x]$  kunnen identificeren met de oneindige kolom-vectoren  $\mathbb{R}^\infty$  waarvan alle behalve eindig veel entries nul zijn, via

$$\vec{f} \leftrightarrow \begin{bmatrix} \lambda_0 & \lambda_1 & \lambda_2 & \dots \end{bmatrix}^{tr}.$$

Veronderstel nu dat  $V$  eindigdimensionaal is, met dimensie  $n$ . Alle echte berekeningen kunnen pas gebeuren nadat we in de  $n$ -dimensionale vectorruimte  $V$  een basis  $\{\vec{v}_1, \dots, \vec{v}_n\}$  gekozen hebben. In dat geval kunnen we  $V$  identificeren met  $\mathbb{R}^n$ . Elke vector  $\vec{v} \in V$  kan immers geschreven worden als

$$\vec{v} = x_1 \vec{v}_1 + \dots + x_n \vec{v}_n,$$

en dus kunnen we  $\vec{v}$  identificeren met de kolom

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n.$$

**Definitie 19.** Zij  $V$  en  $W$  twee vectorruimten. Een afbeelding  $f : V \rightarrow W$  is *lineair* als voor alle vectoren  $\vec{v}, \vec{v}' \in V$  en alle getallen  $\lambda, \mu \in \mathbb{R}$  geldt dat

$$f(\lambda \vec{v} + \mu \vec{v}') = \lambda f(\vec{v}) + \mu f(\vec{v}').$$

**Definitie 20.** Een deel-vectorruimte of kort deelruimte  $Z$  van een vectorruimte  $V$  is een deelverzameling van  $V$  die met de geïnduceerde operaties zelf een vectorruimte is, d.w.z. dat de optelling en scalaire vermenigvuldiging van vectoren in  $Z$  binnen  $Z$  blijven.

Als  $f : V \rightarrow W$  een lineaire afbeelding is, dan is de *kern van  $f$* ,

$$\ker(f) = \{\vec{v} \in V : f(\vec{v}) = \vec{0}_W\},$$

een deelruimte van  $V$  en het *beeld* van  $f$ ,

$$\text{im}(f) = \{f(\vec{v}) : \vec{v} \in V\},$$

een deelruimte van  $W$ .

**Matrixvoorstellingen.** Stel dat  $\{\vec{v}_1, \dots, \vec{v}_n\}$  een basis is voor  $V$ , en  $\{\vec{w}_1, \dots, \vec{w}_m\}$  een basis is voor  $W$ . Dan kunnen we een lineaire afbeelding  $f : V \rightarrow W$  voorstellen door een  $m \times n$  matrix  $A$  waarvan de  $i$ -de kolom de kolomvector is (t.o.v. de basis  $\{\vec{w}_1, \dots, \vec{w}_m\}$ ) van het beeld  $f(\vec{v}_i) \in W$ .

Als  $f : V \rightarrow W$  en  $g : W \rightarrow X$  beiden lineaire afbeeldingen zijn, dan is ook de *samenstelling*  $g \circ f : V \rightarrow X$  een lineaire afbeelding. Indien we (ten opzichte van gekozen basissen in  $V, W$  en  $Z$ ) de afbeelding  $f$  voorstellen door een  $m \times n$  matrix  $A$  en de afbeelding  $g$  door een  $p \times m$  matrix  $B$ , dan is de voorstelling van  $g \circ f$  ten opzichte van de gekozen basissen gelijk aan het matrix-product  $B.A$ .

**Veranderen van basis.** Merk op dat iemand anders een andere keuze van basissen kan gemaakt hebben en dat dezelfde vector en dezelfde lineaire afbeelding dus voorgesteld worden door andere kolomvectoren of andere  $m$  bij  $n$  matrices. We moeten dus nagaan hoe deze voorstellingen afhangen van de keuze van basissen.

Zij  $\{\vec{v}_1, \dots, \vec{v}_n\}$  en  $\{\vec{u}_1, \dots, \vec{u}_n\}$  twee basissen van dezelfde  $n$ -dimensionale vectorruimte  $V$ , dan noemen we de *basisveranderingsmatrix*  $B$  de  $n \times n$  matrix waarvan de  $j$ -de kolom de kolomvoorstelling geeft van de vector  $\vec{u}_j$  ten opzichte van de basis  $\{\vec{v}_1, \dots, \vec{v}_n\}$ , dus

$$B = (b_{ij}) \quad \text{met} \quad \vec{u}_j = b_{1j}\vec{v}_1 + b_{2j}\vec{v}_2 + \dots + b_{nj}\vec{v}_n.$$

Neem  $\vec{v} \in V$ , en stel dat deze vector een kolomvoorstelling  $v = [y_1 \ \dots \ y_n]^{tr}$  heeft ten opzichte van de basis  $\{\vec{u}_1, \dots, \vec{u}_n\}$ , d.w.z.

$$\vec{v} = y_1\vec{u}_1 + \dots + y_n\vec{u}_n.$$

Uit de constructie van de matrix  $B$  volgt dat het matrix-product

$$x = B.y$$

de kolomvoorstelling is van  $\vec{v}$  ten opzichte van de basis  $\{\vec{v}_1, \dots, \vec{v}_n\}$ . Omgekeerd hebben we

$$B^{-1}.x = y,$$

waarbij  $x$  de voorstelling is van  $\vec{v}$  ten opzichte van  $\{\vec{v}_1, \dots, \vec{v}_n\}$  en  $y$  de voorstelling is ten opzichte van  $\{\vec{u}_1, \dots, \vec{u}_n\}$ .

Op analoge wijze kunnen we de matrixvoorstellingen van dezelfde lineaire afbeelding tussen verschillende basisparen vergelijken. We beschouwen enkel het geval met  $W = V$ , en een lineaire afbeelding  $f : V \rightarrow V$ . Zij  $A$  de matrixvoorstelling van  $f$  ten opzichte van de basis  $\{\vec{u}_1, \dots, \vec{u}_n\}$  in  $V$ . Dit wil zeggen dat de  $j$ -de kolom in de matrix  $A$  bestaat uit de kolomvoorstelling van het beeld  $f(\vec{u}_j) \in V$  ten opzichte van de basis  $\{\vec{u}_1, \dots, \vec{u}_n\}$ . We kunnen de lineaire afbeelding  $f$  echter ook voorstellen ten opzichte van de basis  $\{\vec{v}_1, \dots, \vec{v}_n\}$ , met matrixvoorstelling  $A'$ . Het verband tussen beide matrices is

$$A' = B^{-1} \cdot A \cdot B.$$

**Definitie 21.** *Het spoor en de determinant van een lineaire afbeelding  $f : V \rightarrow V$  zijn het spoor en de determinant van eenderwelke matrixvoorstelling  $A$ .*

Deze definities houden wel degelijk steek omdat  $\det(A') = \det(B^{-1}) \det(A) \det(B) = \det(A)$ , en analoog voor het spoor.

Uit het voorgaande is het duidelijk dat resultaten over reële vectorruimten en lineaire afbeeldingen tussen vectorruimten essentieel zullen steunen op de resultaten die we in de voorgaande hoofdstukken besproken, voor vectoren in  $\mathbb{R}^n$  en matrices over  $\mathbb{R}^n$ .