Deel C

Lineaire Algebra

Aanbevolen achtergrondliteratuur met veel opgaven (en oplossingen):

- Seymour Lipschutz, Marc L. Lipson: (Schaum's Outline of Theory and Problems of) Linear Algebra. McGraw-Hill Companies, 2000, 368 p., ISBN: 0071362002.
- Frank Ayres: (Schaum's Outline of Theory and Problems of) Matrices. McGraw-Hill Companies, 1974, 256 p., ISBN: 0070843791.

Les 11 Stelsels lineaire vergelijkingen

Voorbeeld 1: We beginnen deze les met een heel eenvoudig puzzeltje:

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

Hoe oud zijn Annie, Ben en Cees?

Dit kunnen we makkelijk met een stelsel lineaire vergelijkingen beschrijven, namelijk:

$$\begin{aligned} A &= B + C - 2 \\ A + 5 &= 2(B + 5) \\ B - 2 &= \frac{1}{2}(C - 2). \end{aligned}$$

We kunnen dit stelsel op verschillende manieren oplossen, bijvoorbeeld door vervangen van één van de onbekenden met behulp van één van de vergelijkingen: Uit de eerste vergelijking zien we meteen dat $A = B + C - 2$ en als we dit in de tweede vergelijking invullen, wordt deze

$$B + C - 2 + 5 = 2(B + 5) \text{ of te wel } C + 3 = B + 10$$

Uit de derde vergelijking zien we dat $B = \frac{1}{2}C + 1$ en als we B nu in de nieuwe tweede vergelijking vervangen krijgen we

$$C + 3 = \frac{1}{2}C + 1 + 10 \text{ of } \frac{1}{2}C = 8$$

Dus vinden we als oplossing $C = 16$ en door achteruit door de stappen te gaan volgt dat $B = 9$ en $A = 23$.

Voorbeeld 2: Een iets dramatischer voorbeeld: Twee vliegtuigen vliegen in dezelfde hoogte op koersen die een snijpunt hebben. Vliegtuig A heeft een afstand van 45 mijl van het snijpunt en vliegt met een snelheid van 350 mijl per uur, vliegtuig B heeft een afstand van 75 mijl en een snelheid van 600 mijl per uur. Is er gevaar van een crash?

Een botsing zal gebeuren als de twee vliegtuigen het snijpunt op hetzelfde tijdstip bereiken, dus als er een tijdstip x bestaat met

$$350x = 45 \text{ en } 600x = 75.$$

Maar uit $350x = 45$ volgt $x = 9/70$ en $600 \cdot 9/70 = 540/7 \approx 77.14$ dus is er geen gevaar van een crash. Aan de andere kant kunnen we ook berekenen dat het wel mis kan gaan als vliegtuig B iets te langzaam is, namelijk als hij met snelheid $75 \cdot 70/9 = 583.\overline{3}$ vliegt. In dit geval is inderdaad $x = 9/70$ een oplossing voor beide vergelijkingen.

Deze twee voorbeelden van stelsels lineaire vergelijkingen laten zich natuurlijk bijna in het hoofd oplossen. Maar omdat we in de praktijk vaak soortgelijke problemen tegen komen die dan meestal iets meer vergelijkingen en onbekenden hebben, zal het nuttig zijn hier een algemene oplossingsmethode voor te behandelen.

11.1 Gauss-eliminatie

Lineaire vergelijkingen heten *lineair* omdat er veelvouden en sommen van onbekenden in voorkomen, maar geen producten of machten van onbekenden, zoals A^3 of AB. Dit betekent dat we alle informatie van een vergelijking in de *coëfficiënten* van de onbekenden en in de rechterzijde terug vinden (als we alle onbekenden naar de linkerzijde brengen).

Daarom kunnen we een stelsel lineaire vergelijkingen in een schema schrijven dat we een *matrix* noemen. Dit is een rechthoekig schema van getallen waarbij de rijen met vergelijkingen corresponderen en de kolommen met de onbekenden. Een extra kolom wordt eraan toegevoegd om de rechterzijde van een vergelijking op te slaan.

In het puzzeltje over de leeftijd hebben we het stelsel vergelijkingen

$$\begin{aligned} A - B - C &= -2 \\ A - 2B &= 5 \\ B - \frac{1}{2}C &= 1. \end{aligned}$$

De hierbij horende matrix (het hierbij horende schema) is

$$\left(\begin{array}{ccc|c} 1 & -1 & -1 & -2 \\ 1 & -2 & 0 & 5 \\ 0 & 1 & -\frac{1}{2} & 1 \end{array}\right),$$

de eerste kolom bevat de coëfficiënten van A, de tweede die van B, de derde die van C en de vierde kolom geeft de rechterzijden aan.

Wat hebben we hier nu aan? We kunnen op zo'n matrix een aantal operaties toepassen, de zogeheten *elementaire operaties*, waaronder de oplossingen van het stelsel vergelijkingen ongedeerd blijven. Deze operaties zijn:

- het verwisselen van twee rijen;
- het vermenigvuldigen van een rij met een getal $c \neq 0$;
- het optellen van een veelvoud van een rij bij een andere rij.

Het zal duidelijk zijn dat een oplossing van het oorspronkelijke stelsel ook een oplossing van het zo veranderde stelsel is. Voor de eerste twee typen van operaties is dit bijna vanzelfsprekend, bij de derde komt het erop neer dat

$$X_1 = X_2 \text{ en } Y_1 = Y_2 \Rightarrow c \cdot X_1 + Y_1 = c \cdot X_2 + Y_2.$$

Het belangrijke punt is nu, dat we er door het toepassen van de elementaire operaties ook geen (kunstmatige) oplossingen bij krijgen, want we kunnen alle deze operaties weer ongedaan maken (omkeren): bij de eerste type verwisselen we de rijen gewoon terug, bij de tweede delen we de rij door het getal c (hiervoor is het belangrijk dat $c \neq 0$) en bij de derde trekken we hetzelfde veelvoud weer af.

Op deze manier zien we in dat een oplossing van het veranderde stelsel ook een oplossing van het oorspronkelijke stelsel is.

C.1 Stelling *Het toepassen van elementaire operaties veranderd de verzameling van oplossingen van een stelsel lineaire vergelijkingen niet.*

De strategie is nu, de matrix van het stelsel door elementaire operaties op een vorm te transformeren waaraan we de oplossingen makkelijk af kunnen lezen. Zo'n vorm is de *rijtrapvorm*.

C.2 Definitie Een matrix heeft *rijtrapvorm* als elke rij met meer nullen begint dan de voorafgaande rijen.

Als we niet-nul elementen met $\bullet$ en willekeurige elementen met $*$ noteren, zien typische rijtrapvormen er als volgt uit

$$\begin{pmatrix} \bullet & * & * \\ 0 & \bullet & * \\ 0 & 0 & \bullet \end{pmatrix}, \quad \begin{pmatrix} \bullet & * & * & * \\ 0 & 0 & \bullet & * \end{pmatrix}, \quad \begin{pmatrix} 0 & \bullet & * & * \\ 0 & 0 & \bullet & * \\ 0 & 0 & 0 & * \end{pmatrix}, \quad \begin{pmatrix} \bullet & * & * & * & * \\ 0 & \bullet & * & * & * \\ 0 & 0 & 0 & \bullet & * \\ 0 & 0 & 0 & 0 & * \end{pmatrix},$$

terwijl de volgende matrices niet in rijtrapvorm zijn

$$\begin{pmatrix} \bullet & * & * \\ 0 & \bullet & * \\ \bullet & 0 & * \end{pmatrix}, \quad \begin{pmatrix} 0 & \bullet & * & * \\ 0 & \bullet & * & * \\ 0 & 0 & 0 & * \end{pmatrix} \text{ of } \begin{pmatrix} \bullet & * & * & * & * \\ 0 & \bullet & * & * & * \\ 0 & 0 & 0 & \bullet & * \\ 0 & 0 & 0 & \bullet & * \end{pmatrix}.$$

Om het transformeren van een stelsel lineaire vergelijkingen op rijtrapvorm beter te kunnen bestuderen (en controleren) is het handig om afkortingen voor de elementaire operaties af te spreken:

- Het ver<u>w</u>isselen van de i-de en de j-de rij noemen we $W_{i,j}$.
- Het <u>v</u>ermenigvuldigen van de i-de rij met het getal c noteren we met $V_i(c)$.
- Het <u>o</u>ptellen van het c-voud van de i-de rij bij de j-de rij geven we met $O_{i,j}(c)$ aan.

In het voorbeeld van het puzzeltje moeten we voor de rijtrapvorm in de eerste kolom overal 0en produceren (behalve in de eerste rij). De derde rij is al klaar en van de tweede rij moeten we de eerste aftrekken, dus passen we de operatie $O_{1,2}(-1)$ toe. Dit levert de volgende nieuwe matrix (met dezelfde oplossingen) op:

$$\left(\begin{array}{ccc|c} 1 & -1 & -1 & -2 \\ 0 & -1 & 1 & 7 \\ 0 & 1 & -\frac{1}{2} & 1 \end{array}\right)$$

Nu moeten we alleen maar de tweede bij de derde rij optellen, dus $O_{2,3}(1)$ toepassen, om de rijtrapvorm te bereiken. Dit geeft de matrix:

$$\left(\begin{array}{ccc|c} 1 & -1 & -1 & -2 \\ 0 & -1 & 1 & 7 \\ 0 & 0 & \frac{1}{2} & 8 \end{array}\right)$$

Dit stelsel kunnen we nu eenvoudig van onder naar boven oplossen.

We kunnen de oplossingen ook door verdere toepassingen van elementaire operaties vinden. Hiervoor berekenen we de *gereduceerde rijtrapvorm*. Deze heeft de eigenschap dat het eerste niet-nul element in iedere rij 1 is en dat verder alle elementen in de kolom van zo'n 1 allemaal 0 zijn.

In het voorbeeld passen we dus eerst de operaties $V_2(-1)$ en $V_3(2)$ toe, dit geeft:

$$\left(\begin{array}{ccc|c} 1 & -1 & -1 & -2 \\ 0 & 1 & -1 & -7 \\ 0 & 0 & 1 & 16 \end{array}\right)$$

De tweede kolom is klaar als we $O_{2,1}(1)$ uitvoeren:

$$\left(\begin{array}{ccc|c} 1 & 0 & -2 & -9 \\ 0 & 1 & -1 & -7 \\ 0 & 0 & 1 & 16 \end{array}\right)$$

Tenslotte moeten we in de derde kolom nog $O_{3,1}(2)$ en $O_{3,2}(1)$ toepassen en we eindigen dus met de matrix

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 23 \\ 0 & 1 & 0 & 9 \\ 0 & 0 & 1 & 16 \end{array}\right)$$

die de oplossing expliciet aan geeft.

Het transformeren van een stelsel lineaire vergelijkingen op rijtrapvorm wordt ook wel *vegen* genoemd, maar de meer officiële naam is *Gauss-eliminatie*. De Gauss-eliminatie is een algoritme die zowel theoretisch maar ook praktisch de basis voor bijna alle methoden in het kader van stelsels lineaire vergelijkingen vormt. Hieronder wordt het basis-algoritme expliciet aangegeven.

C.3 Algoritme: Gauss-eliminatie (Vegen)

- Begin met de eerste kolom die niet alleen maar 0en bevat, zeg maar de j-de kolom (meestal zal dit de eerste kolom zijn);

- zorg (als nodig) door verwisseling van rijen ervoor dat in de eerste rij van deze kolom een getal c staat dat niet 0 is, dit getal noemen we een *pivot*;
- pas voor de rijen i van 2 t/m n (het aantal van rijen) de elementaire operatie $O_{1,i}(-d/c)$ toe, waarbij d het getal in de j-de kolom van rij i is, op die manier wordt de j-de kolom geveegd;
- herhaal deze procedure op de submatrix die wordt verkregen door de eerste rij te schrappen.

Als we de *gereduceerde rijtrapvorm* willen bereiken, hebben we nog de volgende stappen nodig:

- Pas voor i van 1 t/m n de operatie $V_i(c_i^{-1})$ toe, waarbij c_i het eerste element $\neq 0$ in de i-de rij is;
- veeg de kolom j waarin de laatste rij zijn eerste 1 heeft, door $O_{n,i}(-d_i)$ voor i van 1 t/m $n-1$ toe te passen, waarbij d_i het element in de j-de kolom van rij i is;
- herhaal deze procedure op de submatrix die wordt verkregen door de laatste rij te schrappen.

11.2 Oplosbaarheid

In het voorbeeld van het puzzel met de leeftijden hebben we een eenduidige oplossing gevonden. In het probleem van de vliegtuigen hadden we de matrix

$$\left(\begin{array}{c|c} 350 & 45 \\ 600 & 75 \end{array}\right)$$

die we door toepassen van $O_{1,2}(-\frac{12}{7})$ naar

$$\left(\begin{array}{c|c} 350 & 45 \\ 0 & 75 - \frac{12}{7}45 \end{array}\right)$$

brengen. Maar $75-\frac{12}{7}45 = \frac{-15}{7}$, dus vertegenwoordigd de laatste rij de vergelijking $0 \cdot x = \frac{-15}{7}$ en deze heeft natuurlijk geen oplossing. We kunnen nagaan dat dit al het algemeen geval van een stelsel zonder oplossing is, er geldt namelijk de volgende stelling:

C.4 Stelling *Een stelsel lineaire vergelijkingen is dan en slechts dan niet oplosbaar, als er in de rijtrapvorm en rij met louter nullen als coëfficiënten voor de onbekenden maar een getal $\neq 0$ voor de rechterzijde voorkomt.*

Als geen zo'n rij bestaat, is het stelsel dus oplosbaar.

De volgende vraag is nu, of voor een oplosbaar stelsel vergelijkingen de oplossing eenduidig is.

Het is duidelijk dat de oplossing niet eenduidig is als we een kolom met alleen maar nullen vinden, want dan mag de bijhorende onbekende elke willekeurige waarde hebben. Maar hetzelfde geldt ook als we trappen van meer dan één kolom breedte vinden.

Voor de kolommen die geen *pivot* (het eerste niet-nul element van een rij) bevatten, mogen we immers ook willekeurige waarden kiezen, want bij het van onder naar boven oplossen zijn alleen maar de pivots van belang.

Voorbeeld 3: Dit zal het volgende voorbeeld illustreren. Stel we vinden voor een stelsel vergelijkingen in x, y, z de rijtrapvorm

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 1 & -3 & 4 \end{array}\right)$$

dan kunnen we het stelsel voor een willekeurig gekozen waarde van z oplossen door $y = 4 + 3z$ en $x = 9 - 2z - y = 5 - 5z$. We noemen z dan ook een *vrije parameter*.

C.5 Stelling *Een oplosbaar stelsel lineaire vergelijkingen heeft zoveel vrije parameters als het verschil tussen het aantal onbekenden en het aantal pivots in de rijtrapvorm aangeeft.*
In het bijzonder is een stelsel lineaire vergelijkingen eenduidig oplosbaar dan en slechts dan als de rijtrapvorm even veel pivots bevat als er onbekenden zijn.

Voorbeeld 4: Als voorbeeld kijken we eens naar een iets groter stelsel vergelijkingen dat we al op rijtrapvorm hebben gebracht. We nemen aan dat de rechterzijden alle 0 zijn:

$$\left(\begin{array}{ccccc|c} \boxed{1} & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & \boxed{1} & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} & 0 \end{array}\right).$$

Dit stelsel heeft 5 onbekenden, 3 pivots (door de hokjes gemarkeerd) en dus 2 vrije parameters, die we in de tweede en vierde kolom kunnen kiezen. Als we de onbekenden x, y, z, v, w noemen, krijgen we bij het van beneden naar boven oplossen:

$$\begin{aligned} w &= 0 \\ z + 2v &= 0 \Rightarrow z = -2v \\ x + y - 2v + v &= 0 \Rightarrow x = -y + v \end{aligned}$$

De algemene oplossing is dus van de vorm $(-y + v, y, -2v, v, 0)$. Merk op dat we hier y en v als vrije parameters hebben *gekozen*. We hadden in principe voor de vrije parameters ook een andere keuze kunnen maken, maar voor de manier hoe we uit de rijtrapvorm de oplossingen door achteruit oplossen bepalen, is onze keuze de meest voor de hand liggende.

11.3 Homogeen/inhomogeen stelsel

Bij het van onder naar boven oplossen van een lineair stelsel vergelijkingen met vrije parameters is het soms omslachtig en lastig om met de parameters te rekenen. Echter kunnen we dit ook op een andere manier aanpakken. Hiervoor is de volgende eenvoudige opmerking belangrijk:

Merk op: *Het verschil van twee oplossingen van een stelsel lineaire vergelijkingen is een oplossing van hetzelfde stelsel vergelijkingen, waarbij weliswaar alle rechterzijden door* 0 *zijn vervangen.*

Dit is in een voorbeeld makkelijk in te zien: Stel dat (X, Y, Z) en (X', Y', Z') oplossingen zijn van de vergelijking

$$ax + by + cz = d,$$

dan is $aX + bY + cZ = d = aX' + bY' + cZ'$ en dus geldt

$$a(X - X') + b(Y - Y') + c(Z - Z') = 0.$$

C.6 Definitie Een stelsel vergelijkingen waarbij alle rechterzijden 0 zijn, heet een *homogeen stelsel*, een algemeen stelsel heet *inhomogeen*.

Het stelsel vergelijkingen dat uit een inhomogeen stelsel wordt verkregen door de rechterzijden door 0 te vervangen heet het *bijhorende homogene stelsel*.

Met behulp van deze begrippen kunnen we nu de oplossingen van een stelsel lineaire vergelijkingen abstract als volgt beschrijven:

C.7 Stelling *De oplossingen van een stelsel lineaire vergelijkingen zijn van de vorm* $P + H$*, waarbij* P *een vaste oplossing van het inhomogene stelsel is en* H *alle oplossingen van het bijhorende homogene stelsel doorloopt.*

Als namelijk P en P' oplossingen van het stelsel vergelijkingen zijn, dan is $P' - P = H$ een oplossing van het bijhorende homogene stelsel en dus is P' van de vorm $P' = P + H$.

Een vast gekozen oplossing van een (inhomogeen) stelsel vergelijkingen noemt men vaak een *particuliere oplossing*. Dit betekent niet dat er iets speciaals met de oplossing aan de hand is, maar onderscheid de oplossing van een *algemene oplossing* die met betrekking tot vrije parameters uitgedrukt is.

Het aardige aan een homogeen stelsel lineaire vergelijkingen is, dat we oplossingen van een homogeen stelsel bij elkaar kunnen optellen en zo weer een oplossing van hetzelfde stelsel krijgen. Als namelijk $aX + bY + cZ = 0$ en $aX' + bY' + cZ' = 0$, dan is ook $a(X + X') + b(Y + Y') + c(Z + Z') = 0$. Net zo goed kunnen we een oplossing ook met een factor vermenigvuldigen, dan blijft het ook een oplossing.

We vinden *alle* oplossingen van een homogeen stelsel vergelijkingen als lineaire combinaties van een aantal *basis-oplossingen*, namelijk van juist zo veel basis-oplossingen als er vrije parameters zijn. Hierbij noemen we een oplossing een basis-oplossing als één van de vrije parameters de waarde 1 heeft en alle andere vrije parameters de waarde 0 hebben. Omdat we voor elke keuze van de vrije parameters een oplossing van het homogene stelsel krijgen, geldt dit natuurlijk ook voor onze speciale keuze van één parameter 1 en de anderen 0.

Voorbeeld 5: Als we naar de enkele vergelijking

$$x + y + z = 0$$

kijken, dan zijn y en z vrije parameters en zijn $(x, y, z) = (-1, 1, 0)$ en $(x, y, z) = (-1, 0, 1)$ de basis-oplossingen.

We kunnen nu de basis-oplossingen met een willekeurige factor vermenigvuldigen en bij elkaar optellen. Zo krijgen we bijvoorbeeld de oplossing $(x, y, z) = (-7, 3, 4)$ door de eerste basis-oplossing met 3 en de tweede met 4 te vermenigvuldigen en deze twee oplossingen bij elkaar op te tellen.

Het is duidelijk dat lineaire combinaties van de basis-oplossingen verschillend zijn als de factoren van de basis-oplossingen verschillend zijn. Andersom kunnen we een willekeurige oplossing van het homogene stelsel steeds schrijven als een combinatie van de basis-oplossingen door de waarden van de vrije parameters als factoren te kiezen.

Tenslotte kunnen we ook voor het vinden van een particuliere oplossing van het inhomogene stelsel nog een trucje toepassen om het makkelijker te maken. Bij een oplossing P van het inhomogene stelsel hebben de vrije parameters zekere waarden en als we de oplossing H van het homogene stelsel waarvoor de vrije parameters dezelfde waarden hebben van P aftrekken, krijgen we een nieuwe oplossing $P' = P - H$ van het inhomogene stelsel. Maar P' is zo gemaakt dat alle vrije parameters de waarde 0 hebben. We vinden dus een particuliere oplossing van het inhomogene stelsel door alle vrije parameters 0 te kiezen en het stelsel voor de resterende onbekenden (die bij de pivots horen) op te lossen.

Bij elkaar genomen hebben we dus de volgende stelling ingezien:

C.8 Stelling *Laat P een particuliere oplossing van een stelsel lineaire vergelijkingen zijn en noem de basis-oplossingen van het bijhorende homogene stelsel $h_1, \ldots h_s$, dan zijn alle oplossingen van het stelsel eenduidig te schrijven als $P + c_1 h_1 + c_2 h_2 + \ldots + c_s h_s$, waarbij c_i willekeurige getallen zijn.*

Let op: Als je de oplossingen van een inhomogeen stelsel bepaalt moet je *eerst* het originele stelsel *met* rechte zijden op rijtrapvorm brengen en pas vervolgens de rechte zijden door 0en vervangen. De rijtrapvorm heeft namelijk in het algemeen andere rechterzijden dan het oorspronkelijke stelsel.

In Voorbeeld 3 met rijtrapvorm

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 0 & 1 & -3 & 4 \end{array}\right)$$

hebben we één vrije parameter, namelijk de derde onbekende. Als we $z = 0$ kiezen vinden we $y = 4$ en $x = 5$ als particuliere oplossing van het inhomogene stelsel. Voor het homogene stelsel kiezen we nu $z = 1$, dit geeft $y = 3$ en $x = -5$ en dus vinden we de algemene oplossing als

$$(x, y, z) = (5, 4, 0) + t \cdot (-5, 3, 1)$$

en dit is precies wat we ook eerder als oplossing hadden gevonden (met z in plaats van t).

Het voordeel van de nu toegepaste methode met een particuliere oplossing en basis-oplossingen van het bijhorende homogene stelsel is, dat we niet meer met de vrije parameters als variabelen hoeven te rekenen, maar alleen maar met getallen.

In het grotere Voorbeeld 4 met 5 onbekenden veranderen we even de rechterzijden en bekijken het stelsel in de onbekenden x, y, z, v, w met rijtrapvorm

$$\left(\begin{array}{ccccc|c} \boxed{1} & 1 & 1 & 1 & 1 & 3 \\ 0 & 0 & \boxed{1} & 2 & 3 & 1 \\ 0 & 0 & 0 & 0 & \boxed{1} & 4 \end{array}\right).$$

De vrije parameters horen bij de tweede en vierde kolom, dus bij de onbekenden y en v. Als we $y = 0$ en $v = 0$ kiezen, vinden we de particuliere oplossing $(10, 0, -11, 0, 4)$.

Voor de basis-oplossingen van het bijhorende homogene stelsel kiezen we eerst $y = 0$ en $v = 1$, dan krijgen we als eerste basis-oplossing $(1, 0, -2, 1, 0)$. Vervolgens kiezen we $y = 1$ en $v = 0$, dit geeft de tweede basis-oplossing $(-1, 1, 0, 0, 0)$.

We vinden nu alle oplossingen van het homogene stelsel door de twee basis-oplossingen met willekeurige factoren te vermenigvuldigen en bij elkaar op te tellen (dit geeft precies de oplossingen die we boven al voor het homogene stelsel hebben bepaald).

Om alle oplossingen van het inhomogene stelsel aan te geven, moeten we bij deze oplossingen alleen maar nog de particuliere oplossing optellen, en we krijgen op deze manier als oplossingen:

$$\begin{aligned}(x, y, z, v, w) &= (10, 0, -11, 0, 4) + r \cdot (1, 0, -2, 1, 0) + s \cdot (-1, 1, 0, 0, 0) \\ &= (10 + r - s, s, -11 - 2r, r, 4).\end{aligned}$$

11.4 Toepassingen

Ter afsluiting van deze les laten we nog twee verschillende typen van toepassingen voor het oplossen van lineaire stelsels vergelijkingen zien.

Grafiek door gegeven punten

De eerste toepassing is het vinden van een grafiek door voorgeschreven punten, bijvoorbeeld door de punten van een meting. Hierbij maken we gebruik van de volgende stelling die we hier niet gaan bewijzen.

C.9 Stelling *Laten $x_1, \ldots, x_n$ verschillende (reële) getallen zijn en $y_1, \ldots, y_n$ willekeurig gekozen waarden. Dan is er precies één veelterm*

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1}$$

van graad $\leq n-1$ zodat de grafiek van $y = f(x)$ door de punten $(x_1, y_1), \ldots,$ (x_n, y_n) loopt, d.w.z. zo dat $f(x_i) = y_i$ voor alle $i = 1, 2, \ldots, n$.

De toepassing van deze stelling ziet er als volgt uit: Neem aan we hebben bij een meting voor de x-waarden 1, 2 en 3 de resultaten 6, 3 en 2 verkregen. We zijn dus op zoek naar een grafiek door de punten $(1, 6)$, $(2, 3)$ en $(3, 2)$. Volgens de stelling kunnen we hiervoor een kwadratische veelterm vinden. We schrijven zo'n veelterm in algemene vorm neer, dan zijn de coëfficiënten de onbekenden voor een stelsel lineaire vergelijkingen. In ons geval hebben we dus de veelterm

$$a_0 + a_1 x + a_2 x^2$$

en door invullen van de gewenste punten vinden we het voglende stelsel vergelijkingen in de onbekenden a_0, a_1 en a_2:

$$\begin{aligned} a_0 + a_1 \cdot 1 + a_2 \cdot 1^2 &= 6 \\ a_0 + a_1 \cdot 2 + a_2 \cdot 2^2 &= 3 \\ a_0 + a_1 \cdot 3 + a_2 \cdot 3^2 &= 2. \end{aligned}$$

De matrix van dit stelsel is

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 1 & 2 & 4 & 3 \\ 1 & 3 & 9 & 2 \end{array}\right)$$

en de rijtrapvorm hiervan wordt

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 3 & -3 \\ 0 & 0 & 2 & 2 \end{array}\right).$$

Door van onder naar boven op te lossen vinden we $a_2 = 1$, $a_1 = -6$ en $a_0 = 11$, dus is de gezochte functie gegeven door $y = x^2 - 6x + 11$.

Stoichiometrische vergelijkingen

Een voorbeeld uit de scheikunde laat zien dat soms ook de oplossingen van een homogeen stelsel interessant kunnen zijn. Bij een chemische reactie zijn de uitgangs- en de eindproducten bekend, maar om de hoeveelheden van de stoffen te bepalen moet een chemische vergelijking in evenwicht gebracht worden (de zogeheten *stoichiometrische vergelijking*). In de reactie

$$Ca + H_3PO_4 \to Ca_3P_2O_8 + H_2$$

hebben we bijvoorbeeld coëfficiënten a_1, a_2 en b_1, b_2 nodig, zodat in de chemische reactie

$$a_1 \cdot Ca + a_2 \cdot H_3PO_4 \to b_1 \cdot Ca_3P_2O_8 + b_2 \cdot H_2$$

op beide zijden voor iedere soort van atomen hetzelfde aantal staat.

In dit voorbeeld levert dit de vergelijkingen $a_1 = 3b_1$, $3a_2 = 2b_2$, $a_2 = 2b_1$ en $4a_2 = 8b_1$ op. Het bijhorende stelsel lineaire vergelijkingen is dus

$$\left(\begin{array}{cccc|c} 1 & 0 & -3 & 0 & 0 \\ 0 & 3 & 0 & -2 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 4 & -8 & 0 & 0 \end{array}\right)$$

Door toepassen van $O_{3,4}(-4)$, $O_{3,2}(-3)$ en $W_{2,3}$ vinden we de rijtrapvorm

$$\left(\begin{array}{cccc|c} 1 & 0 & -3 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 6 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

We zien dat we b_2 als vrije parameter kunnen kiezen, als we die t noemen wordt de oplossing dus $b_2 = t$, $b_1 = \frac{1}{3}t$, $a_2 = \frac{2}{3}t$ en $a_1 = t$. Omdat het om een chemische reactie gaat, willen we graag de kleinste oplossing in de natuurlijke getallen vinden. Hiervoor laten we t door $1, 2, \ldots$ lopen en zien dat we voor $t = 3$ inderdaad zo'n oplossing vinden, namelijk

$$3Ca + 2H_3PO_4 \to Ca_3P_2O_8 + 3H_2$$

Belangrijke begrippen in deze les

- stelsel lineaire vergelijkingen
- matrix van een stelsel vergelijkingen
- elementaire operaties
- (gereduceerde) rijtrapvorm, pivot
- Gauss-eliminatie
- vrije parameter
- homogeen/inhomogeen stelsel
- basis-oplossing, particuliere oplossing

Opgaven

80. Vind alle oplossingen van het volgende stelsel lineaire vergelijkingen:

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

81. Breng de volgende stelsels lineaire vergelijkingen op rijtrapvorm, concludeer of ze oplosbaar zijn en geef, zo ja, alle oplossingen aan.

(i) $$\begin{aligned} 2x_1 - x_2 + 3x_3 &= 2 \\ x_1 + 2x_2 + x_3 &= 1 \\ 3x_1 - 4x_2 + 5x_3 &= 3 \end{aligned}$$

(ii) $$\begin{aligned} x_1 - x_2 + 2x_3 - 3x_4 &= 7 \\ 4x_1 + 3x_3 + x_4 &= 9 \\ 2x_1 - 5x_2 + x_3 &= -2 \\ 3x_1 - x_2 - x_3 + 2x_4 &= -2 \end{aligned}$$

(iii) $$\begin{aligned} x_1 + 2x_2 - x_3 &= 0 \\ 2x_1 + 5x_2 + 2x_3 &= 0 \\ x_1 + 4x_2 + 7x_3 &= 0 \\ x_1 + 3x_2 + 3x_3 &= 0 \end{aligned}$$

(iv) $$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 4x_1 + 5x_2 + 6x_3 &= 2 \\ 7x_1 + 8x_2 + 9x_3 &= 3 \\ 5x_1 + 7x_2 + 9x_3 &= 4 \end{aligned}$$

(v) $$\begin{aligned} x_1 + 2x_2 - 3x_3 + 2x_4 &= 2 \\ 2x_1 + 5x_2 - 8x_3 + 6x_4 &= 5 \\ 3x_1 + 4x_2 - 5x_3 + 2x_4 &= 4 \end{aligned}$$

(vi) $$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 + x_5 &= 0 \\ -x_1 - 2x_2 - 2x_3 + 2x_4 + x_5 &= 0 \\ 2x_1 + 4x_2 + 3x_3 - x_4 &= 0 \\ x_1 + 2x_2 + 2x_3 - 2x_4 - x_5 &= 0 \end{aligned}$$

82. Een stelsel vergelijkingen heeft de rijtrapvorm

$$\left(\begin{array}{cccccc|c} 0 & 1 & 2 & 0 & 1 & 1 & 8 \\ 0 & 0 & 2 & 1 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

(i) Bepaal alle oplossingen van het bijhorende homogene stelsel.

(ii) Vind een oplossing van het inhomogene stelsel.

(iii) Beschrijf alle oplossingen van het inhomogene stelsel.

83. Ga na dat een stelsel lineaire vergelijkingen met meer onbekenden dan vergelijkingen nooit een eenduidige oplossing heeft. Bestaat er in dit geval altijd een oplossing?

84. Bepaal de waarden van de parameter a zo dat de aangegeven stelsels lineaire vergelijkingen: (a) een eenduidige oplossing hebben, (b) meerdere oplossingen hebben, (c) niet oplosbaar zijn:

(i) $$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ 2x_1 + 3x_2 + ax_3 &= 3 \\ x_1 + ax_2 + 3x_3 &= 2 \end{aligned}$$

(ii) $$\begin{aligned} x_1 + x_2 + ax_3 &= 2 \\ 3x_1 + 4x_2 + 2x_3 &= a \\ 2x_1 + 3x_2 - x_3 &= 1 \end{aligned}$$

85. Aan welke voorwaarden moeten a, b en c voldoen opdat de gegeven stelsels lineaire vergelijkingen oplosbaar zijn?

(i) $$\begin{aligned} x_1 + 2x_2 - 3x_3 &= a \\ 2x_1 + 6x_2 - 11x_3 &= b \\ x_1 - 2x_2 + 7x_3 &= c \end{aligned}$$

(ii) $$\begin{aligned} x_1 + 2x_2 - 3x_3 &= a \\ 3x_1 - x_2 + 2x_3 &= b \\ x_1 - 5x_2 + 8x_3 &= c \end{aligned}$$

86. Zij $c_1, c_2, c_3, b \in \mathbb{R}$ en $c_1 \neq 0$. Vind alle oplossingen van $c_1x_1 + c_2x_2 + c_3x_3 = b$. Hoeveel vrije parameters heeft dit 'stelsel'?

87. Vind een veelterm $a_3x^3 + a_2x^2 + a_1x + a_0$ van graad 3, waarvan de grafiek door de punten $(-1, 10)$, $(0, 4)$, $(1, 2)$ en $(2, -2)$ gaat.

88. In een experiment zijn de resultaten afhankelijk van een parameter t. Gemeten worden de waarden $y = 0$ voor $x = 0$, $y = t$ voor $x = 1$ en $y = 3t$ voor $x = 2$. Geef een veelterm $a_2x^2 + a_1x + a_0$ van graad 2 aan die door deze drie punten gaat (de coëfficiënten zijn natuurlijk van t afhankelijk).

89. Breng de chemische vergelijking

$$As_2S_3 + H_2O + HNO_3 \to NO + H_3AsO_4 + H_2SO_4$$

in evenwicht.

Les 12 Vectorruimten en lineaire afbeeldingen

Als een robot bij de *robocup* (het voetbaltoernooi voor robots) een doelpunt wil maken, moet hij eerst in de goede positie komen, d.w.z. geschikt achter de bal staan, voordat hij trapt. Hiervoor moet hij naar de positie van de bal lopen en hij moet zo om zijn as draaien, dat hij in de richting van het doel mikt. Als we dit in wiskundige woorden beschrijven, is eerst een translatie van zijn momentele positie naar de positie van de bal nodig, en vervolgens een rotatie om zijn as.

In de berekeningen die voor zo'n probleem nodig zijn, komen we de volgende vraag tegen: Hoe kunnen we de nieuwe coördinaten van een punt (x, y) berekenen, als we rond een as in het punt (x_0, y_0) om een hoek van φ (bijvoorbeeld van 90 graden) draaien?

Voordat we met dit soort vragen aan de slag kunnen, hebben we een paar basisbegrippen nodig.

12.1 Vectorruimten

Meestal worden *vectoren* als een soort pijltjes in het vlak of in de ruimte ingevoerd. Maar vaak is het handig om een iets abstractere toegang te hanteren, hierbij is een vector gewoon een element van een *vectorruimte*. In principe worden vectorruimten door twee cruciale eigenschappen van hun elementen gekarakteriseerd:

(i) De som van twee vectoren in een vectorruimte is ook weer een vector.

(ii) De vermenigvuldiging van een vector met een factor is ook weer een element van de vectorruimte.

Hierbij wordt verondersteld dat voor het optellen van vectoren dezelfde regels gelden als voor het optellen van getallen, d.w.z. het optellen is commutatief ($x+y = y+x$) en associatief ($x+(y+z) = (x+y)+z$) en dat het vermenigvuldigen met factoren distributief over de optelling van de vectoren is ($a(x+y) = ax+ay$).

C.10 Definitie Een (niet lege) verzameling V van vectoren heet een $\mathbb{R}$-vectorruimte als geldt:

(i) Voor alle $v, w \in V$ is $v + w \in V$, d.w.z. met twee vectoren bevat V ook hun som.

(ii) Voor alle $v \in V$ en $\lambda \in \mathbb{R}$ is $\lambda v \in V$, d.w.z. met een vector v bevat V ook alle *schalingen* van v. De factor λ wordt vaak een *scalair* genoemd.

Een eindige som van scalaire veelvouden van vectoren heet een *lineaire combinatie* van de vectoren. Een vectorruimte is dus een verzameling van vectoren, die afgesloten is onder het vormen van *lineaire combinaties*.

Soms wordt $\mathbb{R}$ ook door de complexe getallen $\mathbb{C}$ vervangen, dan spreekt men van een $\mathbb{C}$-*vectorruimte*. In deze cursus zullen we ons tot $\mathbb{R}$-vectorruimten beperken die we van nu af gewoon *vectorruimten* noemen.

Men ziet onmiddellijk in dat een vectorruimte steeds de *nulvector* $0 = 0 \cdot v$ bevat en voor elke vector $v \in V$ een tegengestelde vector $-v = (-1) \cdot v$ met $v + (-v) = 0$.

De belangrijkste vectorruimten zijn het 2-dimensionale vlak

$$\mathbb{R}^2 = \{ \begin{pmatrix} x \\ y \end{pmatrix} \mid x, y \in \mathbb{R}\}$$

en de 3-dimensionale ruimte

$$\mathbb{R}^3 = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x, y, z \in \mathbb{R}\}.$$

Vaak is het echter ook handig met hoger-dimensionale vectorruimten te werken, bijvoorbeeld als bij een meting n onafhankelijke parameters een rol spelen. De n-dimensionale vectorruimte $\mathbb{R}^n$ bestaat dan gewoon uit de kolommen met n componenten.

In veel situaties is het interessant naar een deelverzameling van een vectorruimte te kijken die zelfs ook weer een vectorruimte is. Dit noemen we een *deelvectorruimte* of *lineaire deelruimte*, maar meestal gewoon kort een *deelruimte*. Een voorbeeld is een lijn door de oorsprong in $\mathbb{R}^2$, bijvoorbeeld de lijn $L = \{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x = 2y\}$. Alle elementen van L zijn van de vorm $\lambda \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ en men gaat eenvoudig na dat dit inderdaad een vectorruimte is.

12.2 Lineaire onafhankelijkheid, basis

We hebben gesteld dat een vectorruimte afgesloten is onder het vormen van linearie combinaties. Om een goed overzicht over de elementen van een vectorruimte te krijgen, zal het handig blijken om slechts lineaire combinaties van een vast gekozen stelsel B van vectoren te bekijken dat aan de volgende eisen voldoet:

(i) Het stelsel B is groot genoeg zo dat zich alle vectoren uit V als lineaire combinaties van vectoren uit het stelsel laten schrijven.

(ii) Het stelsel B is zo klein dat zich iedere vector uit V slechts op een enkele manier als lineaire combinatie van vectoren uit B laat schrijven.

Een stelsel $(v_1, \ldots, v_n)$ van vectoren in een vectorruimte V heet *lineair onafhankelijk* of kort *onafhankelijk* als we geen van de v_i als lineaire combinatie van de andere vectoren v_j kunnen schrijven. Een equivalente beschrijving is dat $(v_1, \ldots, v_n)$ onafhankelijk zijn als de enige mogelijkheid voor een vergelijking $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$ gegeven is door alle $\lambda_i = 0$ te kiezen, dus hebben we de definitie:

C.11 Definitie Een stelsel $(v_1, \ldots, v_n)$ van vectoren heet *lineair onafhankelijk* (of kort *onafhankelijk*) als

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0 \Rightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0.$$

Hieruit laat zich snel een handig criterium voor lineaire onafhankelijkheid afleiden: Een stelsel $(v_1, \ldots, v_n)$ van vectoren is dan en slechts dan onafhankelijk is als

(i) $v_1 \neq 0$,

(ii) voor $2 \leq i \leq n$ is de vector v_i geen lineaire combinatie van $v_1, \ldots, v_{i-1}$, d.w.z. $v_i \neq \lambda_1 v_1 + \ldots + \lambda_{i-1} v_{i-1}$ voor alle keuzes van $\lambda_1, \ldots, \lambda_{i-1}$.

Bijvoorbeeld is $\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right)$ onafhankelijk, want $\begin{pmatrix}1\\1\end{pmatrix}$ is geen veelvoud van $\begin{pmatrix}1\\0\end{pmatrix}$ (kijk naar de tweede component).

Aan de andere kant is $\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\2\end{pmatrix}\right)$ wel afhankelijk, omdat $\begin{pmatrix}1\\2\end{pmatrix} = -1\begin{pmatrix}1\\0\end{pmatrix} + 2\begin{pmatrix}1\\1\end{pmatrix}$ geldt.

C.12 Definitie Zij V een vectorruimte:

(i) Een onafhankelijk stelsel $B = (v_1, \ldots, v_n)$ heet een *basis* van V als zich iedere vector $v \in V$ als lineaire combinatie $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ van de vectoren in B laat schrijven.

(ii) Het aantal van vectoren in een basis voor een vectorruimte V is voor alle bases hetzelfde en heet de *dimensie* van V.

> **Merk op:** Achter het feit dat elke basis van een vectorruimte even veel vectoren bevat, zit een serieuze stelling die we hier echter niet gaan behandelen.

We noemen de λ_i in de uitdrukking $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ de *coördinaten* van v met betrekking tot de basis B. Uit de onafhankelijkheid van B volgt dat de coördinaten eenduidig zijn.

Voorbeeld: Het stelsel $B = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right)$ is een basis van $\mathbb{R}^2$, want iedere vector $\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2$ laat zich schrijven als $\begin{pmatrix}x\\y\end{pmatrix} = (x-y)\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}1\\1\end{pmatrix}$ en $\mathbb{R}^2$ is dus 2-dimensionaal (wat geen verrassing is).

We zullen later in deze les zien hoe we met behulp van een homogeen stelsel lineaire vergelijkingen makkelijk na kunnen gaan of een stelsel vectoren onafhankelijk en een basis is.

Door voor de vectorruimte $\mathbb{R}^n$ van vectoren met n componenten de *standaardbasis*

$$B := \left(\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \ldots, \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}\right)$$

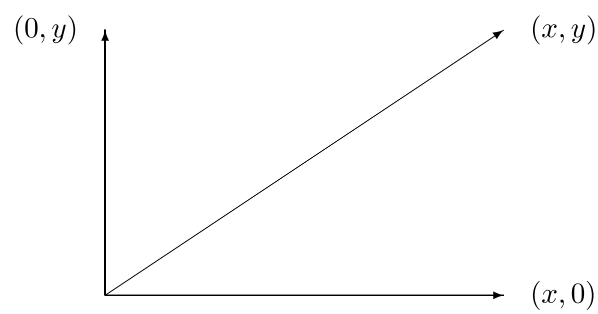
te kiezen, waarbij de i-de vector een 1 in de i-de component heeft en 0 elders, zien we dat $\mathbb{R}^n$ dimensie n heeft. De standaardbasis heeft de mooie eigenschap dat de coördinaten van een vector precies de componenten van de kolomvector zijn.

Let op: Niet elke vectorruimte die uit vectoren met n componenten bestaat (dus een deelverzameling van $\mathbb{R}^n$ is) heeft dimensie n, hij kan ook een lagere dimensie hebben. Het boven genoemde voorbeeld van de lijn L in $\mathbb{R}^2$ is een 1-dimensionale vectorruimte.

12.3 Lineariteit

We bekijken een eenvoudig voorbeeld van een draaiing in het vlak:

Stel, de as waar we om draaien is in de oorsprong, het punt $(0,0)$. Als we om 90 graden (linksom) draaien, komt het punt $(1,0)$ op $(0,1)$ terecht en het punt $(0,1)$ op $(-1,0)$. Een schaling (rekken of krimpen) levert geen enkel probleem op, het punt $(x,0)$ gaat onder de draaiing naar $(0,x)$ en het punt $(0,y)$ gaat naar $(-y,0)$. Een algemeen punt (x,y) krijgen we als som van $(x,0)$ en $(0,y)$, zo als in het volgende plaatje te zien is:



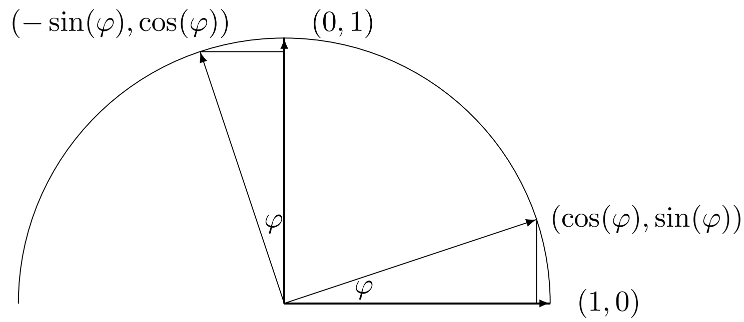
Omdat bij een rotatie de hele rechthoek gedraaid wordt, is het beeld van de som $(x,0)+(0,y)$ hetzelfde als de som van de beelden $(0,x)$ en $(-y,0)$ en dus gelijk aan $(-y,x)$.

We kunnen op deze manier gemakkelijk berekenen dat het punt $(47,11)$ bij een rotatie om 90 graden op $(-11,47)$ terecht komt.

Als we nu in plaats van 90 graden om 60 graden draaien, kunnen we met dezelfde methode weer het beeld van een algemeen punt bepalen, als we weten waar de punten $(1,0)$ en $(0,1)$ naar toe gaan. Met een beetje meetkunde uit de middelbare school zien we in dat $(1,0)$ naar $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ en $(0,1)$ naar $(-\frac{\sqrt{3}}{2}, \frac{1}{2})$ gaat. Een punt (x,y) gaat dus naar het beeldpunt $(\frac{1}{2}x - \frac{\sqrt{3}}{2}y, \frac{\sqrt{3}}{2}x + \frac{1}{2}y)$.

In het algemeen kunnen we een draaiing (rond de oorsprong) om een hoek φ in het vlak met behulp van de cosinus en de sinus beschrijven, namelijk het punt $(1,0)$ gaat naar $(\cos(\varphi), \sin(\varphi))$, het punt $(0,1)$ naar $(-\sin(\varphi), \cos(\varphi))$ en dus gaat het algemene punt (x,y) naar het punt (x',y') met

$$(x',y') = (x\cos(\varphi) - y\sin(\varphi), x\sin(\varphi) + y\cos(\varphi)).$$



We hebben in dit voorbeeld gebruik gemaakt van een belangrijk principe, namelijk dat een draaiing een *lineaire afbeelding* is. Het enige wat we inderdaad gebruikt hebben is, dat we een schaling of een optelling voor en na het afbeelden kunnen uitvoeren zonder het resultaat te veranderen. Dit geeft de volgende definitie:

C.13 Definitie Een afbeelding f op een vectorruimte V heet een *lineaire afbeelding* als geldt:

(i) $f(v + w) = f(v) + f(w)$ voor alle vectoren $v, w \in V$;

(i) $f(\lambda v) = \lambda f(v)$ voor alle vectoren $v \in V$ en alle $\lambda \in \mathbb{R}$.

We hebben boven in feite al gezien dat een rotatie een lineaire afbeelding is, we hoeven namelijk alleen maar de punten (x, y) door vectoren $\begin{pmatrix} x \\ y \end{pmatrix}$ te vervangen. Het cruciale argument is dat de rechthoek met als zijden de vectoren $\begin{pmatrix} x \\ 0 \end{pmatrix}$ en $\begin{pmatrix} 0 \\ y \end{pmatrix}$ onder een rotatie als geheel intact blijft.

Maar: Niet elke eenvoudige afbeelding is lineair. Bijvoorbeeld is een translatie geen lineaire afbeelding, want bij een translatie f over de vector t geldt voor twee vectoren v en w dat $f(v + w) = v + w + t$, maar $f(v) + f(w) = v + t + w + t = v + w + 2t$ en dit is alleen maar hetzelfde als $t = 0$ is.

Een noodzakelijke voorwaarde dat een afbeelding f lineair is, bestaat erin dat $f(0) = 0$ moet gelden, want $f(0) = f(v-v) = f(v)+f(-v) = f(v)-f(v) = 0$. Hiermee kunnen we al zien dat translaties niet lineair zijn.

Omdat translaties zo belangrijk en eenvoudig zijn, is er wel een eigen begrip voor de combinaties van lineaire afbeeldingen en translaties: *affiene afbeeldingen*. Een affiene afbeelding φ op een vectorruimte V is een afbeelding van de vorm $\varphi(v) = f(v) + t$ waarbij f een lineaire afbeelding op V is en $t \in V$ de translatie vector.

We kunnen omgekeerd heel eenvoudig testen of een gegeven afbeelding φ affien is. Omdat elke lineaire afbeelding de nulvector vast laat, moet $\varphi(0)$ de translatie vector zijn. Dus is φ een affiene afbeelding dan en slechts dan als de afbeelding $f(v) := \varphi(v) - \varphi(0)$ een lineaire afbeelding is.

12.4 Matrix van een lineaire afbeelding

We kunnen nu de manier hoe we in het voorbeeld de beeldvectoren van een rotatie bepaald hebben nader toelichten. Het idee is, voor een basis van de vectorruimte de beelden te berekenen en vervolgens de lineariteit te gebruiken. Immers, als $B = (v_1, \ldots, v_n)$ een basis van V is, dan vinden we voor $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ het beeld $f(v)$ als

$$f(v) = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n).$$

Om een lineaire afbeelding handig te beschrijven, gebruiken we (net als bij de stelsels lineaire vergelijkingen) een matrix. Als de afbeelding f van een n-dimensionale vectorruimte V naar een m-dimensionale vectorruimte W gaat, dan hebben we een $m \times n$-matrix A nodig:

We kiezen een basis $(v_1, \ldots, v_n)$ voor V en een basis $(w_1, \ldots, w_m)$ voor W.

In de eerste kolom van de matrix A schrijven we nu het beeld $f(v_1)$ van de eerste basisvector, in de tweede het beeld $f(v_2)$ van de tweede basisvector, enzovoorts. Hierbij zetten we voor de beelden alleen maar de coördinaten met betrekking tot de basis $(w_1, \ldots, w_m)$ neer, dus:

$$f(v_j) = a_1 w_1 + a_2 w_2 + \ldots + a_m w_m \Rightarrow \text{ j-de kolom van A is } \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}.$$

Op deze manier krijgen we de $m \times n$-matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

waarin het element a_{ij} in de i-de rij en de j-de kolom de coördinaat van w_i in het beeld $f(v_j)$ van v_j in de lineaire combinatie

$$f(v_j) = a_{1j} w_1 + a_{2j} w_2 + \ldots + a_{ij} w_i + \ldots + a_{mj} w_m$$

is.

Merk op: Als V en W gewoon de vectorruimten $\mathbb{R}^n$ en $\mathbb{R}^m$ zijn, en we hiervoor de standaardbases $\left(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right)$ van n- resp. m-dimensionale vectoren kiezen, bevat de matrix van f precies de kolomvectoren $f(v_i)$.

In het voorbeeld van boven beschrijven we een rotatie om de hoek φ met betrekking tot de standaardbases door de matrix

$$\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

In sommige situaties is het echter handig een andere basis dan de standaardbasis te kiezen, en dan hebben we de coördinaten met betrekking tot zo'n basis nodig.

Ook *projecties* zijn lineaire afbeeldingen. Als we bijvoorbeeld de loodrechte projectie van de 3-dimensionale ruimte in het $x-y$-vlak bekijken, dan wordt $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ op $\begin{pmatrix} x \\ y \end{pmatrix}$ afgebeeld. De matrix die hierbij hoort is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

Beeld van een vector onder een lineaire afbeelding

Het beeld van een vector v onder een lineaire afbeelding f bepalen we nu, door de matrix van de afbeelding met de (coördinaten van de) vector te *vermenigvuldigen*. Deze vermenigvuldiging moet voor een vector $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ de lineaire combinatie $f(v) = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n)$ weerspiegelen. Dit bereiken we met de definitie:

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ & \ddots & \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} := \begin{pmatrix} a_{11}\lambda_1 + \ldots + a_{1n}\lambda_n \\ \vdots \\ a_{m1}\lambda_1 + \ldots + a_{mn}\lambda_n \end{pmatrix}$$

Dus, de i-de coördinaat van de beeldvector vinden we door tegelijkertijd over de i-de rij van de matrix van de afbeelding en over de λ-vector (de vector met de coördinaten) te lopen en de producten van de paren waar we overheen lopen op te tellen.

Voorbeeld: Het beeld van de vector $\begin{pmatrix} 5 \\ 6 \end{pmatrix}$ onder de lineaire afbeelding met de matrix $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ berekenen we door

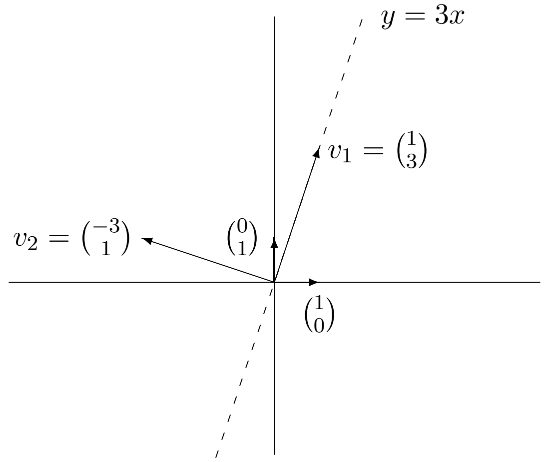
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \cdot 5 + 2 \cdot 6 \\ 3 \cdot 5 + 4 \cdot 6 \end{pmatrix} = \begin{pmatrix} 17 \\ 39 \end{pmatrix}$$

Voorbeelden voor lineaire afbeeldingen in de vlakte zijn naast rotaties ook spiegelingen. De spiegeling in de diagonaal $x = y$ in de vlakte beschrijven we met betrekking tot de standaardbasis door de matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ en het beeld van een algemene vector is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$.

Als we in plaats van de standaardbasis de basis $\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$ kiezen, dan wordt de matrix van de spiegeling $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, want de eerste basisvector ligt in de spiegelingsas en blijft dus ongedeerd en de tweede basisvector staat loodrecht op de spiegelingsas en wordt dus omgedraaid.

Iets ingewikkelder wordt het voorbeeld, als we de spiegeling in de lijn $y = 3x$ bekijken. Hierbij is het handig om een basis zo te kiezen dat één vector op de spiegelingsas ligt, en de andere er loodrecht op staat. Zo'n basis is bijvoorbeeld

$$B := \left(v_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} -3 \\ 1 \end{pmatrix} \right).$$



De matrix van de spiegeling met betrekking tot deze basis is (net als boven) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, want v_1 wordt afgebeeld op $1 \cdot v_1 + 0 \cdot v_2$ en v_2 gaat naar $0 \cdot v_1 + (-1) \cdot v_2$.

Om beelden van de vectoren in de standaardbasis te vinden, moeten we de vectoren $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ en $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ als lineaire combinaties van de gekozen basis schrijven. Er geldt

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 1 \\ 3 \end{pmatrix} - \frac{3}{10} \begin{pmatrix} -3 \\ 1 \end{pmatrix} \text{ en } \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{3}{10} \begin{pmatrix} 1 \\ 3 \end{pmatrix} + \frac{1}{10} \begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

De coördinaten van het beeld van $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ zijn dus $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{10} \\ \frac{-3}{10} \end{pmatrix} = \begin{pmatrix} \frac{1}{10} \\ \frac{3}{10} \end{pmatrix}$ en dus is het beeld van $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ gelijk aan

$$\frac{1}{10} v_1 + \frac{3}{10} v_2 = \begin{pmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{pmatrix}.$$

Op dezelfde manier volgt dat het beeld van $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ gelijk aan $\begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix}$ is. Met betrekking tot de standaardbasis heeft de spiegeling in de as $y = 3x$ dus de matrix $A = \frac{1}{5} \begin{pmatrix} -4 & 3 \\ 3 & 4 \end{pmatrix}$.

12.5 Kern en beeld

Om een lineaire afbeelding goed te kunnen beschrijven, zijn twee verdere begrippen nuttig: de *kern* en het *beeld* van een lineaire afbeelding.

C.14 Definitie De *kern* van een lineaire afbeelding $f : V \to W$ bestaat uit de vectoren, die door f naar 0 worden afgebeeld, dus

$$\ker(f) := \{v \in V \mid f(v) = 0\}.$$

Het *beeld* van een lineaire afbeelding $f : V \to W$ is de verzameling van vectoren, waarop een vector uit V afgebeeld wordt, dus

$$\mathrm{im}(f) := \{f(v) \in W \mid v \in V\}.$$

Merk op, dat met $v \in \ker(f)$ ook alle scalaire veelvouden $\lambda v \in \ker(f)$ zijn, en dat uit $v, w \in \ker(f)$ volgt, dat $v + w \in \ker(f)$ is. De kern is dus afgesloten onder het vormen van lineaire combinaties en is dus een deelruimte van V.

Analoog geldt ook bij het beeld dat lineaire combinaties van elementen in het beeld weer in het beeld liggen, dus is het beeld een deelruimte van W.

Het bepalen van de de kern van een lineaire afbeelding komt erop neer, het homogene stelsel lineaire vergelijkingen op te lossen, dat door de matrix van de afbeelding gegeven is. Voor de kern zijn we namelijk op zoek naar coördinaten $\lambda_1, \ldots, \lambda_n$ zo dat $f(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) = 0$. Met andere woorden zoeken we een vector $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ met $A \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = 0$. Als we dit uitschrijven is het niets anders dan het stelsel lineaire vergelijkingen met matrix A en de λ_i als onbekende.

Met hetzelfde argument zien we nu ook hoe we kunnen testen of een stelsel vectoren onafhankelijk of een basis is: We schrijven de vectoren v_i als kolommen in een matrix A en lossen het bijhorende homogene stelsel lineaire vergelijkingen op. Een vector in de kern van de lineaire afbeelding met matrix A levert dan een combinatie $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$ en als niet alle $\lambda_i = 0$ zijn, is het stelsel vectoren afhankelijk. Maar omdat we het hier over een homogeen stelsel hebben is dit het geval dan en slechts dan als de rijtrapvorm van A een vrije parameter laat zien (minder pivots dan kolommen heeft). We hebben dus ingezien:

C.15 Stelling *Een stelsel $(v_1, \ldots, v_n)$ van vectoren is lineair onafhankelijk als de kern van de matrix A die de v_i als kolommen heeft alleen maar uit de nulvector bestaat.*

Dit is het geval dan en slechts dan als de rijtrapvorm van A even veel pivots als kolommen heeft.

Omdat het aantal elementen in elke basis hetzelfde is, is een onafhankelijk stelsel van n vectoren automatisch een basis van een vectorruimte V van dimensie n, namelijk van de vectorruimte van alle lineaire combinaties van de vectoren in het onafhankelijke stelsel.

Waarschuwing: Let op dat je niet de dimensie van V met het aantal componenten van de vectoren uit V verwart, dit kan namelijk echt groter dan de dimensie zijn.

Het beeld van een lineaire afbeelding vinden we door de afbeelding op een basis los te laten. Dit levert vectoren op die het beeld voortbrengen maar deze zijn niet noodzakelijk onafhankelijk. Als we een basis $(v_1, v_2, \ldots, v_n)$ afbeelden, dan is $f(v_i)$ voor het voortbrengen van het beeld alleen maar nodig als $f(v_i)$ geen lineaire combinatie van $f(v_1), \ldots, f(v_{i-1})$ is.

Er bestaat een belangrijke samenhang tussen de dimensies van de kern en het beeld van een lineaire afbeelding. Stel dat f een lineaire afbeelding van een n-dimensionale vectorruimte V naar een m-dimensionale vectorruimte W is, dan is de matrix van f een $m \times n$-matrix. Om de kern van de afbeelding te bepalen, brengen we de matrix op (gereduceerde) rijtrapvorm. Omdat we voor de kern het homogene stelsel oplossen, is de dimensie van de kern het aantal vrije parameters en dit is het verschil van het aantal n van kolommen en het aantal pivot-kolommen.

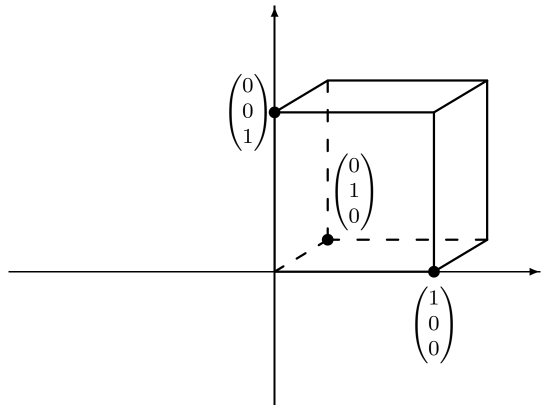
Aan de andere kant kunnen we van de gereduceerde rijtrapvorm heel makkelijk de dimensie van het beeld aflezen, dat is namelijk precies het aantal pivot-kolommen. Verder is het niet moeilijk om in te zien, dat de elementaire operaties waarmee we naar de rijtrapvorm zijn gekomen de dimensie van het beeld niet hebben veranderd (terwijl het beeld zelf wel is veranderd). Dit ligt eraan dat we de elementaire operaties terug kunnen draaien en dat een lineaire relatie tussen kolommen ook na het toepassen van een elementaire operatie nog bestaat. We hebben dus de volgende stelling ingezien:

C.16 Stelling *Als f een lineaire afbeelding van een n-dimensionale vectorruimte is, dan is de som van de dimensies van $\ker(f)$ en $\mathrm{im}(f)$ gelijk aan n.*

Het voordeel van deze stelling ligt voor de hand: we hoeven alleen maar één van de dimensies van $\ker(f)$ of $\mathrm{im}(f)$ te bepalen en kunnen de andere daaruit afleiden.

Als we bijvoorbeeld een lineaire afbeelding met matrix $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}$ hebben, dan berekenen we dat de rijtrapvorm hiervan $\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$ is en concluderen dus dat de kern 1-dimensionaal en het beeld 2-dimensionaal is. We zien dat de eerste twee kolommen $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ en $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ in de (originele) matrix van de afbeelding onafhankelijk zijn, dus zijn deze vectoren een basis voor het beeld en daarom zijn de vectoren in het beeld van de vorm $\begin{pmatrix} x+y \\ x \\ x+2y \end{pmatrix}$.

Voorbeeld: We bekijken de begrippen kern en beeld van een lineaire afbeelding aan de hand van een projectie die een kubus op de volgende manier in het 2-dimensionale vlak afbeeldt:



We kiezen de kubus zo dat de drie aangegeven hoekpunten met de vectoren uit de standaardbasis in $\mathbb{R}^3$ overeen komen. Bij de projectie beelden we nu de vector $\begin{pmatrix}1\\0\\0\end{pmatrix}$ af op de vector $\begin{pmatrix}1\\0\end{pmatrix}$ in het 2-dimensionale vlak en de vector $\begin{pmatrix}0\\0\\1\end{pmatrix}$ op $\begin{pmatrix}0\\1\end{pmatrix}$ in het vlak. Dit heeft het effect dat de voorkant van de kubus een niet vertekend vierkant wordt. Door de keuze van het beeld van de derde basisvector bepalen we nu of we meer van voren of van boven op de kubus kijken. In het plaatje is het beeld van $\begin{pmatrix}0\\1\\0\end{pmatrix}$ de vector $\begin{pmatrix}\frac{1}{3}\\\frac{1}{5}\end{pmatrix}$.

De matrix van deze projectie (met betrekking tot de standaardbases) is $\begin{pmatrix}1 & \frac{1}{3} & 0\\0 & \frac{1}{5} & 1\end{pmatrix}$ want de kolommen van deze matrix zijn de beelden van de basisvectoren in de standaardbasis van $\mathbb{R}^3$.

Het beeld van een algemene vector berekenen we nu door

$$\begin{pmatrix}1 & \frac{1}{3} & 0\\0 & \frac{1}{5} & 1\end{pmatrix}\cdot\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x+\frac{1}{3}y\\\frac{1}{5}y+z\end{pmatrix}.$$

De kern van deze projectie is precies de lijn langs welke we projecteren want dit zijn de punten die op het nulpunt in de vlakte afgebeeld worden. Merk op dat alle lijnen die parallel met deze projectielijn zijn ook op één punt in het 2-dimensionale vlak afgebeeld worden. We berekenen de kern door het homogene stelsel vergelijkingen met als matrix de matrix van de afbeelding op te lossen. In ons geval heeft deze matrix al rijtrapvorm, we kunnen de z-coördinaat als vrije parameter kiezen en krijgen zo als kern de lijn $\{\begin{pmatrix}5t\\-15t\\3t\end{pmatrix}\in\mathbb{R}^3 \mid t\in\mathbb{R}\}$.

Omdat de vectoren $\begin{pmatrix}1\\0\end{pmatrix}$ en $\begin{pmatrix}0\\1\end{pmatrix}$ in het beeld van de projectie liggen is het duidelijk dat het beeld van deze projectie het hele vlak $\mathbb{R}^2$ is. Dit hadden we ook uit het feit kunnen concluderen dat de dimensie van de kern 1 is, want dan moet die dimensie van het beeld volgens stelling C.16 gelijk aan $3-1=2$ zijn.

12.6 Stelsels lineaire vergelijkingen als lineaire afbeeldingen

We hebben in de vorige les stelsels lineaire vergelijkingen in een schema geschreven en vervolgens met behulp van eenvoudige operaties op dit schema de oplossingen van het schema bepaald. We kunnen nu met de nieuwe terminologie van lineaire afbeeldingen de resultaten herformuleren.

Laat A de $m \times n$-matrix van een stelsel lineaire vergelijkingen met m vergelijkingen en n onbekenden zijn. Verder zij b de rechterzijde van het stelsel. De lineaire afbeelding met matrix A (met betrekking tot de standaardbases) noemen we f. Dan geldt:

- Het stelsel is oplosbaar dan en slechts dan als b in $\mathrm{im}(f)$ ligt.
- Het aantal vrije parameters voor de oplossingen is de dimensie van $\ker(f)$.
- De oplossingen zijn van de vorm $P+h$, waarbij P een particuliere oplossing is en $h \in \ker(f)$.

Belangrijke begrippen in deze les

- vectorruimte, deelruimte
- onafhankelijke vectoren, basis, dimensie
- coördinaten van een vector
- lineaire afbeelding
- matrix van een lineaire afbeelding
- kern en beeld van een lineaire afbeelding

Opgaven

90. Een (stomme) robot bij de robocup mikt altijd in de richting waarin hij het laatst gelopen is, behalve als hij expres nog een extra rotatie uitvoert. Onze robot is van het punt $(-10, 10)$ naar de bal toe gelopen die op het punt $(20, -10)$ ligt. Nu draait hij nog om 30 graden (linksom) en haalt vervolgens uit. Het doel ligt tussen de punten $(100, -5)$ en $(100, 5)$.

    (i) Wordt dit een doelpunt (als we van tegenstanders afzien)?

    (ii) Hoe zit het met een rotatie om 45 graden in plaats van 30?

    (iii) Kun je een hoek gokken, zo dat het wel een doelpunt wordt?

91. Ga na of de volgende stelsels van vectoren lineair afhankelijk of onafhankelijk zijn:

$$\text{(i)}\ \left(\begin{pmatrix}1\\3\\-1\end{pmatrix}, \begin{pmatrix}2\\0\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\1\end{pmatrix}\right) \qquad \text{(ii)}\ \left(\begin{pmatrix}1\\1\\-1\end{pmatrix}, \begin{pmatrix}2\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\1\\2\end{pmatrix}\right)$$

(iii) $\left(\begin{pmatrix} 1 \\ -2 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 6 \\ -5 \\ -4 \end{pmatrix} \right)$ (iv) $\left(\begin{pmatrix} 0 \\ 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \right)$

92. Ga na of de volgende afbeeldingen lineair zijn en geef, zo ja, de matrix van de afbeelding met betrekking tot de standaardbases aan:

 (i) $f : \mathbb{R}^2 \to \mathbb{R} : f(x, y) := xy$
 (ii) $f : \mathbb{R} \to \mathbb{R}^2 : f(x) := (2x, 3x)$
 (iii) $f : \mathbb{R}^2 \to \mathbb{R}^2 : f(x, y) := (2x - y, x)$
 (iv) $f : \mathbb{R}^2 \to \mathbb{R}^3 : f(x, y) := (x + 1, 2y, x + y)$
 (v) $f : \mathbb{R}^2 \to \mathbb{R}^3 : f(x, y) := (x, 2y, x + y)$
 (vi) $f : \mathbb{R}^2 \to \mathbb{R}^3 : f(x, y) := (x^2, 2y, x + y)$
 (vii) $f : \mathbb{R}^3 \to \mathbb{R}^2 : f(x, y, z) := (z, x + y)$
 (viii) $f : \mathbb{R}^3 \to \mathbb{R}^2 : f(x, y, z) := (x + 1, y + z)$

93. De veeltermen van graad $\leq n$ vormen een $(n + 1)$-dimensionale vectorruimte met basis $B = (1, x, x^2, \ldots, x^n)$. De (algebraïsche) afgeleide $f'(x)$ van een veelterm $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ is gedefinieerd als

$$f'(x) = na_n x^{n-1} + (n - 1)a_{n-1}x^{n-2} + \ldots + a_1.$$

 (i) Ga na dat de afbeelding $f(x) \to f'(x)$ een lineaire afbeelding is, d.w.z. toon aan dat $(f + g)'(x) = f'(x) + g'(x)$ en $(\lambda f(x))' = \lambda f'(x)$.
 (ii) Bepaal voor $n = 3$ de matrix A van deze lineaire afbeelding (m.b.t. de basis B).
 (iii) Bepaal kern en beeld van A.

94. Ook de afbeelding die een veelterm $f(x)$ naar $g(x) := f(x+1)$ afbeeldt is een lineaire afbeelding (bijvoorbeeld gaat x^2 naar $(x + 1)^2 = x^2 + 2x + 1$).

 (i) Bepaal voor $n = 4$ de matrix A van deze lineaire afbeelding (m.b.t. de basis B uit de vorige opgave).
 (ii) Bepaal (met behulp van de matrix A) het beeld van de veelterm $f(x) = x^4 - 3x^3 + 5x^2 - 7x + 9$.

95. Bepaal alle veeltermen van graad ≤ 3 die een oplossing voor de differentiaalvergelijking $xf'(x) - f(x + 1) = 0$ zijn.
(Aanwijzing: Ga na dat de afbeelding $f(x) \to xf'(x) - f(x + 1)$ lineair is. De kern van deze lineaire afbeelding is dan de verzameling van oplossingen van de differentiaalvergelijking.)

96. Bij een vaak gebruikte projectie van $\mathbb{R}^3$ naar $\mathbb{R}^2$ gaat de vector $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ naar $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, de vector $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ naar $\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ en de vector $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ naar $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

 (i) Bepaal de matrix van deze projectie (m.b.t. de standaardbases).
 (ii) Bepaal het beeld van een kubus met hoekpunten $(\pm 1, \pm 1, \pm 1)$ bij deze projectie. Bereken de beelden van de hoekpunten en maak een plaatje.

(iii) Bereken de kern van deze projectie. Dit zijn de punten die onder de projectie naar de oorsprong in de vlakte gaan.

97. De kubus met hoekpunten $(\pm 1, \pm 1, \pm 1)$ wordt zo om de as door $(-1, -1, -1)$ en $(1, 1, 1)$ gedraaid dat hij weer hetzelfde uitziet. Hiervoor zijn er twee mogelijkheden, waarvan we degene kiezen waarbij het punt $(1, 1, -1)$ naar $(1, -1, 1)$ gaat.

(i) Toon aan dat de vectoren naar de punten $(1, 1, -1)$, $(1, -1, 1)$ en $(-1, 1, 1)$ onafhankelijk en dus een basis voor $\mathbb{R}^3$ zijn. Dit is een geschikte basis omdat deze drie punten de punten zijn die in de kubus met $(1, 1, 1)$ verbonden zijn en daarom door de rotatie onder elkaar verwisseld worden.

(ii) Bepaal de matrix van deze rotatie met betrekking tot de in (i) gekozen basis.

(iii) Bepaal het beeld van de punten $(1, 0, 0)$ (middelpunt van een zijvlak) en $(\frac{1}{2}, \frac{1}{3}, \frac{1}{4})$ (punt ergens binnen de kubus).

Les 13 Matrix product

We hebben gezien hoe we matrices kunnen gebruiken om lineaire afbeeldingen te beschrijven. Om het beeld van een vector onder een afbeelding te bepalen, hebben we al een soort product van een matrix met een vector gedefinieerd. In deze les zullen we nu naar producten van matrices kijken en zien wat we daarmee kunnen.

13.1 Vermenigvuldiging van matrices

Om de manier, hoe we matrices vermenigvuldigen, te motiveren, kijken we eerst naar een toepassing.

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

De achtergrond van dit vraagstuk is natuurlijk van economische aard. Het is op grond van enquêtes wel bekend hoeveel mensen uit een gegeven land naar een ander land reizen, maar het is veel moeilijker om er achter te komen, hoeveel geld de mensen in het buitenland uitgeven. Hierover laat zich met behulp van de verspreiding van de munten wel een uitspraak doen.

In een heel eenvoudig model kunnen we *Euroland* splitsen in Nederland en niet-Nederland. Om het proces van de verspreiding van de munten te beschrijven, moeten we alleen maar afschatten, hoeveel munten in een jaar van Nederland naar het buitenland gaan en hoeveel munten vanuit het buitenland Nederland binnen komen.

Stel dat in een jaar een tiende van de Nederlandse munten naar het buitenland gaan. Verder zullen misschien nog 2% van de munten verdwijnen, dus blijven 88% van de munten in eigen land. Omgekeerd nemen we aan dat van de munten in het buitenland elk jaar 1% naar Nederland komen en 5% verdwijnen. Wat kunnen we bij deze gegevens na 1, 2, 3 jaar voor een mix van munten in Nederland verwachten?

We kunnen het proces van de verspreiding van de munten door een *overgangsmatrix* beschrijven, waar we in de eerste kolom schrijven waar de munten uit Nederland naar toe gaan, in de tweede, waar de munten uit het buitenland naar toe gaan en in de derde, waar de verdwenen munten terecht komen (deze blijven natuurlijk verdwenen). Voor de boven aangegeven schattingen is deze matrix dus

$$A = \begin{pmatrix} 0.88 & 0.01 & 0 \\ 0.10 & 0.94 & 0 \\ 0.02 & 0.05 & 1 \end{pmatrix}$$

Als we nu willen weten hoeveel Nederlandse en buitenlandse munten na 1 jaar in Nederland terecht komen, hoeven we alleen maar A met de vector $v = \begin{pmatrix} N \\ B \\ 0 \end{pmatrix}$ te vermenigvuldigen, waarbij N het aantal geslagen Nederlandse munten en B het aantal geslagen buitenlandse munten zijn. Het resultaat is de vector

$$\begin{pmatrix} 0.88 \cdot N + 0.01 \cdot B \\ 0.10 \cdot N + 0.94 \cdot B \\ 0.02 \cdot N + 0.05 \cdot B \end{pmatrix}$$

die precies weer geeft, dat na een jaar in Nederland 88% van de Nederlandse en 1% van de buitenlandse munten te vinden zijn. De tweede component geeft de mix voor het buitenland aan en de derde de mix van verdwenen munten.

Hoe zit het nu na het tweede en na het derde jaar? In principe doen we hetzelfde opnieuw, we vermenigvuldigen de overgangsmatrix met de vector waarin de hoeveelheden munten staan, dus met het resultaat na het eerste jaar. Dit geeft dus

$$A = \begin{pmatrix} 0.88 & 0.01 & 0 \\ 0.10 & 0.94 & 0 \\ 0.02 & 0.05 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0.88N + 0.01B \\ 0.10N + 0.94B \\ 0.02N + 0.05B \end{pmatrix}$$

$$= \begin{pmatrix} 0.88 \cdot (0.88N + 0.01B) + 0.01 \cdot (0.10N + 0.94B) \\ 0.10 \cdot (0.88N + 0.01B) + 0.94 \cdot (0.10N + 0.94B) \\ 0.02 \cdot (0.88N + 0.01B) + 0.05 \cdot (0.10N + 0.94B) + 1 \cdot (0.02N + 0.05B) \end{pmatrix}$$

$$= \begin{pmatrix} 0.7754N + 0.0182B \\ 0.182N + 0.8846B \\ 0.0426N + 0.0972B \end{pmatrix}$$

Dit zegt dat na twee jaar 18.2% van de Nederlandse munten in het buitenland zitten en nog 77.54% in Nederland (de andere 4.26% zijn inmiddels verdwenen).

Om de mix na drie jaar te bepalen kunnen we nu in principe weer de matrix A met de resultaatvector na twee jaar vermenigvuldigen. Maar we kunnen dit ook op een andere manier beschrijven.

We kunnen namelijk niet alleen maar een matrix met een vector vermenigvuldigen, maar ook een matrix met een matrix als het aantal kolommen van de eerste matrix hetzelfde is als het aantal rijen van de tweede matrix. Om het product $A \cdot B$ voor een $m \times r$-matrix A en een $r \times n$-matrix B te bepalen, beschouwen we B als een collectie van n kolommen. Hoe we een matrix met een kolom vermenigvuldigen hebben we al eerder gezien. Als we A dus met de n kolommen van B vermenigvuldigen, geeft dit n kolommen met telkens m componenten, en deze schrijven we (in de goede volgorde) weer als een $m \times n$-matrix.

C.17 Definitie Zij A de $m \times r$-matrix (a_{ij}) met a_{ij} het element in de i-de rij en j-de kolom, en B de $r \times n$-matrix (b_{ij}), dan is het *matrix product* (of kort *product*) van A en B gedefinieerd als de $m \times n$-matrix $C = A \cdot B$ met element c_{ij} in de i-de rij en j-de kolom waarvoor geldt:

$$c_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj}.$$

Merk op: Twee matrices kunnen alleen maar vermenigvuldigd worden als het aantal kolommen van de eerste overeenkomt met het aantal rijen van de tweede matrix. Het resultaat is dan een matrix met net zo veel rijen als de eerste matrix en net zo veel kolommen als de tweede.

De definitie van het matrix product ziet er erger uit als het is. Een goede manier om de vermenigvuldiging van matrices te onthouden is het volgende schema:

$$\begin{array}{cccc} & & & \text{kolom } j \\ & & & \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & b_{kj} & \vdots \\ b_{r1} & \dots & b_{rn} \end{pmatrix} \\ \text{rij } i & \begin{pmatrix} a_{11} & \dots & a_{1r} \\ \vdots & a_{ik} & \vdots \\ a_{m1} & \dots & a_{mr} \end{pmatrix} & & \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & c_{ij} & \vdots \\ c_{m1} & \dots & c_{mn} \end{pmatrix} \end{array}$$

Om het element c_{ij} in de i-de rij en j-de kolom van C te bepalen lopen we gelijktijdig over de i-de rij van A en de j-de kolom van B. De elementen van A en B waar we tegelijkertijd op staan worden vermenigvuldigd en deze producten worden bij het doorlopen opgeteld.

In het voorbeeld van de verspreiding van de Euro-munten hebben we de mix na één jaar bepaald door $A \cdot v$ te berekenen, de mix na twee jaar was gegeven door de vector $(A \cdot (A \cdot v))$ en dus wordt de mix na drie jaar de vector $(A \cdot (A \cdot (A \cdot v)))$. Maar dit kunnen we ook berekenen als $(A \cdot A \cdot A) \cdot v$ waar we ook $A^3 \cdot v$ voor schrijven. Met de boven aangegeven methode vinden we dat

$$A^2 = A \cdot A = \begin{pmatrix} 0.88 & 0.01 & 0 \\ 0.10 & 0.94 & 0 \\ 0.02 & 0.05 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0.88 & 0.01 & 0 \\ 0.10 & 0.94 & 0 \\ 0.02 & 0.05 & 1 \end{pmatrix} = \begin{pmatrix} 0.7754 & 0.0182 & 0 \\ 0.1820 & 0.8846 & 0 \\ 0.0426 & 0.0972 & 1 \end{pmatrix}$$

en

$$\begin{aligned} A^3 = A^2 \cdot A &= \begin{pmatrix} 0.7754 & 0.0182 & 0 \\ 0.1820 & 0.8846 & 0 \\ 0.0426 & 0.0972 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0.88 & 0.01 & 0 \\ 0.10 & 0.94 & 0 \\ 0.02 & 0.05 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0.684172 & 0.024862 & 0 \\ 0.248620 & 0.833344 & 0 \\ 0.067208 & 0.141794 & 1 \end{pmatrix}. \end{aligned}$$

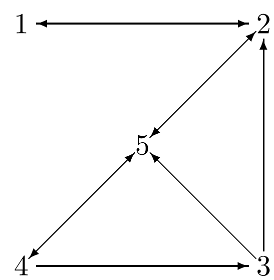
Als voorbeeld gaan we even na hoe we het eerste element in tweede rij van A^3 vinden, dus het getal 0.248620: Omdat $A^3 = A^2 \cdot A$ is, moeten we hiervoor gelijktijdig over de *tweede rij* van A^2 en over de *eerste kolom* van A lopen, op deze manier krijgen we de som van producten:

$$0.1820 \cdot 0.88 + 0.8846 \cdot 0.10 + 0 \cdot 0.02 = 0.24862.$$

De mix na afloop van drie jaar vinden we door de matrix A^3 met de vector $v = \begin{pmatrix} N \\ B \\ 0 \end{pmatrix}$ te vermenigvuldigen, en hieruit kunnen we zien dat er na drie jaar van de in Nederland geslagen munten 68.4% nog (of weer) in Nederland, 24.9% in het buitenland en 6.7% verdwenen zijn.

Toepassing op netwerken

Een verdere toepassing van het machtsverheffen van matrices (naast de ontwikkeling van populaties met behulp van overgangsmatrices) is de analyse van netwerken. Neem aan we hebben het volgende netwerk van (geörienteerde) paden:



Als we de vraag willen beantwoorden hoe lang het kortste pad van 1 naar 3 is, kunnen we dit aan de hand van de machten van een geschikte matrix aflezen. Voor een netwerk met n punten gebruiken we een $n \times n$-matrix $A = (a_{ij})$, waarbij $a_{ij} = 1$ is als er een directe verbinding van punt i naar punt j bestaat en $a_{ij} = 0$ anders. Voor het voorbeeld is dus

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

We kunnen deze matrix nu ook anders interpreteren: Het element a_{ij} zegt, hoeveel *directe paden* er van het punt i naar het punt j zijn (één of geen).

Als we nu $C = A^2$ berekenen, dan is het element c_{ij} hiervan het aantal van *paden van lengte* 2 die van i naar j gaan. Immers, elk pad van lengte 2 van i naar j is van de vorm $i \to k \to j$ en het aantal van deze paden is $a_{ik}a_{kj}$. Als we deze aantallen voor alle k optellen, krijgen we aan de ene kant het aantal paden van lengte 2 van i naar j, maar aan de andere kant is $a_{i1} \cdot a_{1j} + a_{i2} \cdot a_{2j} + \ldots + a_{in} \cdot a_{nj}$ juist het element op de plaats (i, j) in het matrix product $A \cdot A$.

Als we dit argument herhalen, komen we erachter dat het element c_{ij} in de matrix $C = A^m$ het aantal van paden van lengte m is die van punt i naar punt j gaan.

In het voorbeeld is

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

dus zien we dat het kortste pad van punt 1 naar punt 3 lengte 4 heeft, omdat A^4 de kleinste macht van A is waarbij het element op de plaats $(1, 3)$ niet 0 is. Voor alle andere paren van punten geldt, dat een verbinding hooguit lengte 3 heeft.

We kunnen ook eenvoudig het aantal paden van lengte $\leq m$ bepalen, dit is namelijk het element c_{ij} in $C = A + A^2 + \ldots + A^m$. Hierbij tellen we matrices natuurlijk componentsgewijs bij elkaar op.

13.2 Samenstelling van lineaire afbeeldingen

In de vorige les hebben we gezien hoe we lineaire afbeeldingen zoals rotaties en spiegelingen in het 2-dimensionale vlak of in de 3-dimensionale ruimte met behulp van matrices kunnen beschrijven. Nu we weten hoe we matrices vermenigvuldigen, kunnen we ook aangeven hoe zich de samenstelling van twee (of meer) lineaire afbeeldingen laat beschrijven. Hierbij betekent *samenstelling* (net als bij functies zoals we die in de Calculus hebben behandeld) dat we eerst bijvoorbeeld een rotatie en dan een spiegeling op een vector toepassen.

Stel dat A de matrix van een lineaire afbeelding f en B de matrix van een lineaire afbeelding g is. We willen op een (coördinaten-)vector v eerst f en op het resultaat dan g toepassen. Hiervoor zeggen we ook dat we de *samengestelde* of *gecomponeerde afbeelding* $g \circ f$ toepassen. Het beeld onder f vinden we als $u = A \cdot v$ en het beeld hiervan onder g dan als

$$w = B \cdot u = B \cdot (A \cdot v) = (B \cdot A) \cdot v.$$

Dus, in plaats van de twee afbeeldingen achter elkaar toe te passen, kunnen we ook hun matrices (in de goede volgorde) vermenigvuldigen en vervolgens het product met de vector v vermenigvuldigen.

C.18 Stelling *Zij $f : V \to U$ een lineaire afbeelding van de vectorruimte V naar de vectorruimte U en $g : U \to W$ een lineaire afbeelding van U naar W. Ten opzichte van vast gekozen bases voor V, U en W zij A de matrix van f en B de matrix van g.*

Dan heeft de samengestelde afbeelding $g \circ f : V \to W$ (ten opzichte van dezelfde bases voor V en W) de matrix $B \cdot A$.

Voorbeeld: Als we eerst om een hoek φ willen draaien en vervolgens in de as $y = x$ willen spiegelen is de matrix van de samenstelling van deze afbeeldingen het product

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} = \begin{pmatrix} \sin(\varphi) & \cos(\varphi) \\ \cos(\varphi) & -\sin(\varphi) \end{pmatrix}$$

Waarschuwing: In het algemeen is $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$!

Let op dat de vermenigvuldiging van matrices bijna nooit commutatief is, d.w.z. in het algemeen is $A \cdot B \neq B \cdot A$. Bijvoorbeeld heeft de spiegeling in de x-as de matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ en de rotatie om 90 graden heeft de matrix $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Als we eerst de rotatie en dan de spiegeling toepassen, wordt dit beschreven door de matrix

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

Het is eenvoudig om na te gaan dat dit de spiegeling in de as $y = x$ is.

Als we omgekeerd eerst de spiegeling en dan de rotatie toepassen, heeft de samengestelde afbeelding de matrix

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

en dit is de spiegeling in de as $y = -x$.

Het zou eigenlijk geen verrassing zijn, dat het samenstellen van (lineaire) afbeeldingen niet commutatief is, want bij gewone functies zouden we bijvoorbeeld ook nooit vermoeden dat de functies $\sin(e^x)$ en $e^{\sin(x)}$ hetzelfde zijn.

Maar omdat we het samenstellen van lineaire afbeeldingen die we door matrices beschrijven op dezelfde manier noteren als het product van gewone getallen, zou men hier eerder het idee kunnen krijgen dat de twee matrices in het product $A \cdot B$ verwisseld kunnen worden. Dit is echter niet het geval!

13.3 Inverse van een matrix

Bij de stelsels lineaire vergelijkingen hebben we vaker het argument gebruikt dat we door het toepassen van elementaire operaties op de matrix van het stelsel geen oplossingen verliezen of erbij krijgen, omdat we de elementaire operaties terug kunnen draaien.

In de taal van lineaire afbeeldingen betekent het *terugdraaien* van een afbeelding f (die van een vectorruimte V naar zichzelf gaat), dat we een lineaire afbeelding g moeten vinden zo dat de samenstelling van f en g de *identieke afbeelding* id_V is, die elke vector ongedeerd laat, dus waarvoor $id_V(v) = v$ voor iedere vector $v \in V$ geldt.

Omdat de identieke afbeelding in het bijzonder elke basis vector op zich zelf afbeeldt, is de matrix van de identieke afbeelding de matrix

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

met 1 op de diagonaal en 0 elders, die we de *identiteitsmatrix* noemen en met I_n noteren (als V de dimensie n heeft).

Als A de $n \times n$-matrix van de afbeelding f is, hebben we dus een $n \times n$-matrix B nodig zo dat $B \cdot A = I_n$ is. De matrix B noemen we dan de *inverse matrix van A* en noteren deze met A^{-1} (zo als we dat bij gewone getallen gewend zijn).

We kunnen ook omgekeerd naar een matrix C zoeken zo dat $A \cdot C = I_n$ is, maar als we zo'n matrix gevonden hebben, kunnen we deze vergelijking van links met A^{-1} vermenigvuldigen, dit geeft $A^{-1} \cdot A \cdot C = A^{-1}$ en dus $C = A^{-1}$ omdat $A^{-1} \cdot A = I_n$. De samenstelling van een afbeelding en zijn inverse geeft dus onafhankelijk van de volgorde de identieke afbeelding.

C.19 Definitie Voor een $n \times n$-matrix A heet de matrix A^{-1} met

$$A \cdot A^{-1} = A^{-1} \cdot A = I_n$$

de *inverse matrix van A* of kort *inverse van A* als zo'n matrix A^{-1} bestaat.

We zullen nu nagaan, onder welke voorwaarden de inverse matrix A^{-1} bestaat en hoe we in dit geval A^{-1} kunnen bepalen.

Het berekenen van de inverse matrix A^{-1}

De eigenschap die we voor de inverse matrix eisen, is dat $A^{-1} \cdot A = I_n$ of, wat equivalent is, dat $A \cdot A^{-1} = I_n$ is. Maar het laatste kunnen we als stelsel lineaire vergelijkingen opvatten, waarbij A de matrix van het stelsel vergelijkingen is en de kolommen van A^{-1} vectoren met onbekenden zijn. Voor de i-de kolom v_i van A^{-1} geldt, dat hij voldoet aan $A \cdot v_i = e_i$ is, waarbij e_i de i-de standaardbasis vector $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ (met 1 in de i-de component en 0 elders) is.

Het inverteren van een $n \times n$-matrix komt dus neer op het oplossen van een stelsel lineaire vergelijkingen met n verschillende rechterzijden. Om dit stelsel vergelijkingen op te lossen, brengen we de matrix A van het stelsel op rijtrapvorm, waarbij we de verschillende rechterzijden simultaan behandelen.

Uit dit gezichtspunt zien we nu snel in, dat een $n \times n$-matrix A dan en slechts dan een inverse heeft (we zeggen ook: *inverteerbaar* is), als de rijtrapvorm geen 0-rij heeft, of, met andere woorden, als de rijtrapvorm van A precies n pivots en dus geen vrije parameters heeft:

In dit geval weten we namelijk dat het stelsel vergelijkingen met elke rechterzijde eenduidig oplosbaar is.

C.20 Stelling *Een $n \times n$-matrix A is dan en slechts dan inverteerbaar als het bij A horende stelsel lineaire vergelijkingen met elke rechterzijde oplosbaar is. Dit is dan en slechts dan het geval als de rijtrapvorm van A in elke kolom een pivot heeft en dus geen 0-rij bevat.*

Voorbeeld: We willen de inverse van de matrix $A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ bepalen. Hiervoor brengen we het volgende stelsel lineaire vergelijkingen met drie rechterzijden op gereduceerde rijtrapvorm:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

De stappen hiervoor zijn:

$$\underset{\substack{O_{1,2}(-2) \\ O_{1,3}(-1)}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -2 & -2 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \end{array}\right) \underset{\substack{O_{3,2}(1) \\ W_{2,3}}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & -2 & -3 & 1 & 1 \end{array}\right)$$

$$\underset{V_3(-\frac{1}{2})}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & \frac{3}{2} & \frac{-1}{2} & \frac{-1}{2} \end{array}\right) \underset{O_{3,1}(-1)}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{-1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & \frac{3}{2} & \frac{-1}{2} & \frac{-1}{2} \end{array}\right)$$

Dus is $A^{-1} = \frac{1}{2} \begin{pmatrix} -1 & 1 & 1 \\ -2 & 0 & 2 \\ 3 & -1 & -1 \end{pmatrix}$ de inverse van A.

De inverse matrix in het geval $n = 2$

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

Voor een matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ bepalen we net als boven de rijtrapvorm van het bijhorende stelsel lineaire vergelijkingen. We nemen eerst aan, dat $a \neq 0$ is. De stappen zijn dan:

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right) \underset{O_{1,2}(-\frac{d}{a})}{\longrightarrow} \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & d - \frac{cb}{a} & -\frac{c}{a} & 1 \end{array}\right) \underset{V_2(a)}{\longrightarrow} \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & ad - bc & -c & a \end{array}\right).$$

De matrix A is dus inverteerbaar dan en slechts dan als $ad - bc \neq 0$, en in dit geval kunnen we verder doorgaan naar de gereduceerde rijtrapvorm:

$$\underset{O_{2,1}(-\frac{b}{ad-bc})}{\longrightarrow} \left(\begin{array}{cc|cc} a & 0 & \frac{ad}{ad-bc} & -\frac{ba}{ad-bc} \\ 0 & ad - bc & -c & a \end{array}\right) \rightarrow \left(\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right).$$

Dus bestaat de inverse als $ad - bc \neq 0$ is en is voor het geval $a \neq 0$ gelijk aan

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

Voor het geval $a = 0$ vinden we als rijtrapvorm de matrix $\left(\begin{array}{cc|cc} c & 0 & -\frac{d}{b} & 1 \\ 0 & b & 1 & 0 \end{array}\right)$ als $b \neq 0$ is. Als $b = 0$ is, is het duidelijk dat de matrix niet inverteerbaar is, evenzo moet $c \neq 0$ zijn. We komen dan verder naar de gereduceerde rijtrapvorm

$$\left(\begin{array}{cc|cc} 1 & 0 & -\frac{d}{bc} & \frac{1}{c} \\ 0 & 1 & \frac{1}{b} & 0 \end{array}\right).$$

Dit is inderdaad hetzelfde als de boven gevonden inverse A^{-1}, als we daar $a = 0$ invullen, dus geldt de uitspraak

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ met } ad - bc = 0 \Rightarrow A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

voor alle 2×2-matrices.

13.4 Basistransformatie

In de vorige les hebben we gezien dat het soms handig is om de matrix van een lineaire afbeelding met betrekking tot een andere basis dan de standaardbasis te schrijven. De overgang van de matrix van een lineaire afbeelding met betrekking tot één basis naar de matrix met betrekking tot een andere basis noemen we een *basistransformatie*. Met behulp van de inverse matrix kunnen we een basistransformatie nu heel simpel beschrijven.

Het enige wat hiervoor nodig is, is de matrix van de lineaire afbeelding die een basis naar de andere vertaalt. Zij A de matrix van de afbeelding f met betrekking tot de basis $B = (v_1, \ldots, v_n)$, en laat T de matrix zijn die de basisvectoren van een nieuwe basis $C = (w_1, \ldots, w_n)$ met betrekking tot de basis B uitdrukt. Hiermee wordt bedoeld dat de i-de kolom van T de coördinaten van de (nieuwe) basisvector w_i met betrekking tot de (oude) basis B bevat.

Als B speciaal de standaardbasis is, dan zijn de kolommen van T gewoon de vectoren $w_1, \ldots, w_n$.

C.21 Stelling *Met de aangegeven notaties is de matrix van de lineaire afbeelding f met betrekking tot de nieuwe basis C gegeven door*

$$A' = T^{-1} \cdot A \cdot T.$$

We gaan nu even toelichten waarom deze stelling geldt.

De matrix T geeft de coördinaten van de vectoren uit de basis C met betrekking tot de basis B aan. Omdat T^{-1} de inverse matrix van T is, drukt deze omgekeerd de vectoren uit de basis B in de nieuwe basis C uit, dus zijn de kolommen van T^{-1} de coördinaten van de vectoren in de basis B met betrekking tot de basis C.

Als we de samenstelling $A' = T^{-1} \cdot A \cdot T$ nu van rechts naar links lezen en op de i-de basisvector w_i toe passen, dan geldt het volgende:

- De vector $T \cdot w_i$ geeft de coördinaten van de vector w_i met betrekking tot de basis B aan.
- $A \cdot T \cdot w_i$ bevat dus de coördinaten van het beeld $f(w_i)$ met betrekking tot de basis B.
- Door nu nog met T^{-1} te vermenigvuldigen drukken we het beeld $f(w_i)$ weer met betrekking tot de basis C uit. Op deze manier vinden we de gewenste coördinaten van het beeld $f(w_i)$ met betrekking tot de nieuwe basis $(w_1, \ldots, w_n)$.

Voorbeeld: In een voorbeeld in de vorige les hebben we naar de spiegeling in de lijn $y = 3x$ gekeken. We waren toen (met iets moeite) erachter gekomen dat deze spiegeling met betrekking tot de standaardbasis $B = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$ de matrix $A = \frac{1}{5} \begin{pmatrix} -4 & 3 \\ 3 & 4 \end{pmatrix}$ heeft.

Als we als nieuwe basis de basis $C = \left(\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \end{pmatrix} \right)$ met één vector in de spiegelingslijn en één vector loodrecht erop kiezen, kunnen we de transformatiematrix T meteen opschrijven (omdat B de standaardbasis is), er geldt:

$$T = \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix}.$$

Met de formule voor de inverse van een 2×2-matrix berekenen we T^{-1} en krijgen:

$$T^{-1} = \frac{1}{10} \begin{pmatrix} 1 & 3 \\ -3 & 1 \end{pmatrix}.$$

De matrix van de spiegeling in de lijn $y = 3x$ heeft dus met betrekking tot de nieuwe basis C de matrix

$$A' = T^{-1} \cdot A \cdot T = \frac{1}{10} \begin{pmatrix} 1 & 3 \\ -3 & 1 \end{pmatrix} \cdot \frac{1}{5} \begin{pmatrix} -4 & 3 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We hadden overigens net zo goed met de eenvoudige matrix A' kunnen beginnen en hieruit de matrix met betrekking tot de standaardbasis kunnen afleiden. Uit $A' = T^{-1} \cdot A \cdot T$ volgt namelijk (door van links met T en van rechts met T^{-1} te vermenigvuldigen) dat $A = T \cdot A' \cdot T^{-1}$ en dit is in principe de berekening die we in de afgelopen les stapsgewijs hebben uitgevoerd.

Belangrijke begrippen in deze les

- vermenigvuldiging en machtsverheffing van matrices
- overgangsmatrix
- matrix van een netwerk
- samenstelling van lineaire afbeeldingen
- inverse van een matrix
- basistransformatie

Opgaven

98. Bij een zeldzame soort van vogels geven observaties de volgende informatie: In het eerste jaar zijn de vogels in hun jeugd, in het tweede zijn ze pubers en in het derde zijn ze volwassen. Van de volwassen vrouwtjes hebben elk jaar 33% een vrouwelijk kuiken. Slechts 18% van de vrouwelijke kuikens overleven het eerste jaar, 71% van de vrouwelijke pubers overleven het tweede jaar en van de volwassen vrouwtjes overlijden per jaar 6%.

    (i) Beschrijf de ontwikkeling van deze populatie door een (3×3) overgangsmatrix, waarbij het element in de i-de kolom en j-de rij van de matrix voor de overgang van het i-de naar het j-de jaar staat.

    (ii) Bij een telling van de vrouwelijke vogels worden op een gegeven moment 300 jeugdige vogels, 50 pubers en 1000 volwassen vogels geteld. Hoeveel vogels van de verschillende leeftijden verwachten we na afloop van 1, 2, 3 jaar.

    (iii) Heb je een verwachting over de ontwikkeling van deze populatie? Kun je er redenen voor aangeven?

99. In het 2-dimensionale vlak zij f de draaiing om een hoek van 45 graden en g de spiegeling in de lijn $y = x$. Bepaal de matrices van de samengestelde afbeeldingen

$$\text{(i) } f \circ g \qquad \text{(ii) } g \circ f \qquad \text{(iii) } g \circ f \circ g.$$

100. Ga na of de volgende matrices inverteerbaar zijn en bepaal, zo ja, de inverse matrix:

$$\text{(i) } \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} \qquad \text{(ii) } \begin{pmatrix} 2 & 1 & 0 \\ -1 & 3 & 2 \\ 1 & 2 & 4 \end{pmatrix} \qquad \text{(iii) } \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 1 & 2 & 2 & 2 \\ 1 & 3 & 2 & 4 \end{pmatrix}.$$

101. Ga na of de volgende matrices inverteerbaar zijn en bepaal, zo ja, de inverse matrix:

$$A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 2 \\ 7 & 5 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}, \quad D = \begin{pmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{pmatrix},$$

$$E = \begin{pmatrix} 1 & 3 & 4 \\ 3 & -1 & 6 \\ -1 & 5 & 1 \end{pmatrix}, \qquad F = \begin{pmatrix} 1 & -2 & -3 \\ 2 & -1 & -4 \\ 3 & -3 & -5 \end{pmatrix}, \qquad G = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 2 \end{pmatrix}.$$

102. Van een lineaire afbeelding is bekend, dat de vectoren op de lijn $y = x$ onveranderd blijven en dat de vectoren op de lijn $y = -x$ om een factor 2 verlengd worden. Geef de matrix van deze lineaire afbeelding met betrekking tot de standaardbasis aan.

103. Een netwerk met 5 punten wordt beschreven door de matrix

$$A = (a_{ij}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

waarbij $a_{ij} = 1$ betekent dat de punten i en j door een ribbe verbonden zijn en $a_{ij} = 0$ dat ze niet (direct) verbonden zijn.

(i) Teken een plaatje van het netwerk.

(ii) Welke paren van punten zijn door een eenduidig pad van lengte hoogstens 2 verbonden?

(iii) Bepaal het aantal paden van lengte (precies) 3 die van punt 2 naar punt 3 lopen.

Les 14 Eigenwaarden en eigenvectoren

In het voorbeeld van de verspreiding van de Euro-munten hebben we gezien hoe we de mix van munten na afloop van n jaar uit de n-de macht A^n van de overgangsmatrix af kunnen lezen. Een voor de hand liggende vraag is, of er uiteindelijk een stabiel evenwicht wordt bereikt waarin de mix van munten niet meer verandert.

Nu zijn we in het voorbeeld ervan uit gegaan dat munten verdwijnen en deze munten ook verdwenen blijven. Dan is duidelijk dat uiteindelijk alle munten verdwijnen, en de toestand waarin in elk land helemaal geen munten meer zijn is natuurlijk stabiel, maar ook flauw. Een betere aanpak is dat we de verdwenen munten door nieuw geslagen munten weer opvullen.

In het voorbeeld krijgen we zo de nieuwe overgangsmatrix

$$A = \begin{pmatrix} 0.90 & 0.01 \\ 0.10 & 0.99 \end{pmatrix}.$$

Het totale aantal munten blijft altijd hetzelfde omdat alle kolommen de som 1 hebben.

Als er een stabiele evenwichtstoestand bereikt wordt, dan voldoet de vector v met de hoeveelheden munten in Nederland en het buitenland aan

$$A \cdot v = v.$$

Anders gezegd is $A \cdot v - v = 0$ ofwel v ligt in de kern van $A - I$ (waarbij I de identiteitsmatrix is). Dit is een makkelijk probleem, want we hoeven alleen maar een stelsel lineaire vergelijkingen op te lossen. Dit stelsel heeft de matrix

$$A - I = \begin{pmatrix} -0.10 & 0.01 \\ 0.10 & -0.01 \end{pmatrix}$$

en de kern hiervan is $t \cdot \begin{pmatrix} 1 \\ 10 \end{pmatrix}$. Het evenwicht ziet er dus zo uit, dat tien keer zo veel munten in het buitenland dan in Nederland zijn.

Verder kunnen we ook bepalen, hoe de mix er uitziet, die zal namelijk in Nederland en in het buitenland hetzelfde zijn. Waarom dat zo is zullen we later in deze les zien. Als we hogere en hogere machten A^n van A berekenen zullen deze machten voor $n \to \infty$ naar

$$A_\infty = \frac{1}{11} \begin{pmatrix} 1 & 1 \\ 10 & 10 \end{pmatrix}$$

convergeren. Merk op dat de kolommen van A_∞ veelvouden van de evenwichtsvector zijn die zo geschaald zijn dat de som van de componenten 1 is.

14.1 Eigenvectoren

Om een evenwichtstoestand te vinden, hebben we een vector v gezocht waarvoor geldt dat $A \cdot v = v$ is. Iets algemener is het een belangrijk concept om vectoren te vinden die door een lineaire afbeelding alleen maar met een factor vermenigvuldigd worden, maar op dezelfde lijn blijven liggen. Dit zijn dus vectoren die aan de vergelijking

$$A \cdot v = \lambda v$$

voldoen, waarbij λ een scalaire factor is.

C.22 Definitie Voor een $n \times n$-matrix A heet een vector $v \neq 0$ die aan de vergelijking $A \cdot v = \lambda v$ voldoet een *eigenvector* voor de *eigenwaarde* λ (soms ook wel eigenvector *met* eigenwaarde λ) van A.

De evenwichtsvector v van boven die aan de vergelijking $A \cdot v = v$ voldoet, is dus een eigenvector voor de eigenwaarde 1.

We hebben al eerder met eigenvectoren te maken gehad, zonder ze expliciet zo te noemen:

- Bij een spiegeling zijn vectoren op de spiegelingsas eigenvectoren voor de eigenwaarde 1.
- Vectoren op een lijn loodrecht op de spiegelingsas zijn eigenvectoren voor de eigenwaarde -1.
- Vectoren die in de kern van een lineaire afbeelding liggen, zijn eigenvectoren voor de eigenwaarde 0.

Merk op: De nulvector noemen we geen eigenvector, want hij is een eigenvector voor elke eigenwaarde, en we willen graag dat een eigenvector bij een unieke eigenwaarde hoort.

Het is duidelijk dat voor een eigenvector v met eigenwaarde λ ook alle veelvouden cv (met $c \neq 0$) eigenvectoren met eigenwaarde λ zijn, want uit $A \cdot v = \lambda v$ volgt wegens de lineariteit van A dat $A \cdot (cv) = c(A \cdot v) = \lambda(cv)$.

De vraag is nu natuurlijk hoe we eigenwaarden een eigenvectoren van een matrix kunnen vinden. Het belangrijke punt is, dat een eigenvector v met eigenwaarde λ voldoet aan $A \cdot v = \lambda v$, ofwel $(A - \lambda \cdot I) \cdot v = 0$. De vector v ligt dus in de kern van de matrix $A - \lambda \cdot I$ en dit kan alleen maar voor een vector $v \neq 0$ gebeuren als het stelsel lineaire vergelijkingen met matrix $A - \lambda \cdot I$ een vrije parameter heeft. Omdat A even veel rijen en kolommen heeft, kan dit alleen maar gebeuren als we in de rijtrapvorm van de matrix een 0-rij vinden. In het bijzonder is de matrix $A - \lambda \cdot I$ niet inverteerbaar als λ een eigenwaarde van A is. We hebben dus ingezien:

C.23 Stelling *Een eigenwaarde van een matrix A is een getal λ zo dat $A - \lambda \cdot I$ niet inverteerbaar is. Dit is het geval dan en slechts dan als de rijtrapvorm van $A - \lambda \cdot I$ een vrije parameter en dus een 0-rij heeft.*

14.2 Determinanten

Sommigen hebben zeker eens van de determinant van een matrix gehoord. De determinant is een getal, die zich voor een matrix met even veel rijen als kolommen laat berekenen. We zullen nu kort bespreken hoe de determinant van een algemene $n \times n$-matrix berekend wordt en hierbij een aantal eigenschappen van de determinant toelichten. Voor ons cruciaal is de volgende eigenschap van de determinant:

C.24 Stelling *Een $n \times n$-matrix is dan en slechts dan inverteerbaar als zijn determinant niet 0 is.*

De samenhang van inverteerbaarheid en de determinant van een matrix volgt meteen uit de manier hoe we de determinant kunnen berekenen. Hier is een methode, die misschien iets van achter door de borst lijkt, maar die in feite de manier is, hoe in software pakketten als Maple of Mathematica de determinant wordt berekend. Het belangrijkste ingrediënt in deze methode is weer de rijtrapvorm.

C.25 Karakterisering van de determinant

- Voor een matrix in rijtrapvorm is de determinant het product van de elementen op de diagonaal.
- Als we twee rijen (of kolommen) van een matrix verwisselen, wordt de determinant met -1 vermenigvuldigd.
- Als we een rij (of kolom) met een getal c vermenigvuldigen, wordt ook de determinant met c vermenigvuldigd.
- Als we een veelvoud van een rij (of kolom) bij een andere optellen, verandert de determinant niet.

Een belangrijke stelling zegt nu, dat de determinant door deze karakterisering (tot op een constante factor na) eenduidig bepaald is. Deze stelling gaan we in dit college gewoon als feit accepteren.

> Er is ook een meetkundige interpretatie van de determinant: De absolute waarde van de determinant is het volume van het parallelepipedium dat door de kolommen van de matrix opgespannen wordt.

Het feit dat in de karakterisering van de determinant aangegeven wordt, hoe de determinant bij de verschillende *elementaire operaties* verandert, is tegelijkertijd de aanwijzing hoe de determinant berekent wordt.

C.26 Berekening van de determinant

- Definieer een hulpgetal $d := 1$.
- Breng de matrix door elementaire operaties op rijtrapvorm.

- Vervang hierbij voor iedere vermenigvuldigingen van een rij (of kolom) met een factor c het hulpgetal d door $c \cdot d$.
- Voor iedere verwisselingen van rijen (of kolommen), vervang d door $-d$.
- De determinant van de matrix is dan het product van de diagonaalelementen in de rijtrapvorm, gedeeld door d.

Met behulp van deze procedure kunnen we nu rechtstreeks inzien dat een $n \times n$-matrix dan en slechts dan inverteerbaar is als de determinant niet 0 is:

$\rightarrow$ De matrix is inverteerbaar als de rijtrapvorm n pivots heeft die niet 0 zijn, maar dan is ook de determinant niet nul.

$\leftarrow$ Omgekeerd is een matrix niet inverteerbaar als er in de rijtrapvorm een 0-rij is, maar dan is ook de determinant 0 omdat er een 0 op de diagonaal staat.

Determinanten van 2×2-matrices

Voor 2×2-matrices kunnen we een formule voor de determinant bepalen door de determinant van een algemene matrix te berekenen: De matrix

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

heeft voor $a \neq 0$ de rijtrapvorm

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

en omdat we geen rijen hebben verwisseld en ook geen rij met een factor hebben vermenigvuldigd is de determinant

$$\det(A) = a \cdot (d - \frac{c}{a}b) = ad - bc.$$

Dit klopt ook voor $a = 0$, want dan verwisselen we de rijen en hebben als rijtrapvorm $\begin{pmatrix} c & d \\ 0 & b \end{pmatrix}$ en omdat we een keer rijen hebben verwisseld is de determinant gelijk aan $-bc = 0 \cdot d - bc$. De formule van het geval $a \neq 0$ klopt dus ook voor $a = 0$ en geldt dus algemeen.

De determinant van een 2×2-matrix is dus het product van de diagonaalelementen min het product van de dwarsdiagonaal-elementen.

Determinanten van 3×3-matrices

Ook voor 3×3-matrices kunnen we nog een algemene formule uit de rijtrapvorm berekenen. We nemen weer een algemene matrix

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

en negeren voor het moment de gevallen waar een noemer 0 kan zijn. De stappen naar de rijtrapvorm zijn nu:

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \to \begin{pmatrix} a & b & c \\ 0 & e - \frac{d}{a}b & f - \frac{d}{a}c \\ 0 & h - \frac{g}{a}b & i - \frac{g}{a}c \end{pmatrix} = \begin{pmatrix} a & b & c \\ 0 & \frac{1}{a}(ae - bd) & \frac{1}{a}(af - cd) \\ 0 & \frac{1}{a}(ah - bg) & \frac{1}{a}(ai - cg) \end{pmatrix}$$

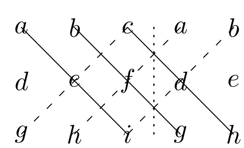
$$\to \begin{pmatrix} a & b & c \\ 0 & \frac{1}{a}(ae - bd) & \frac{1}{a}(af - cd) \\ 0 & 0 & \frac{1}{a}(ai - cg) - \frac{ah-gb}{ae-bd}\frac{af-cd}{a} \end{pmatrix}.$$

Als product van de diagonaalelementen vinden we

$$\begin{aligned} det(A) &= \frac{1}{a}(ae - bd)(ai - cg) - \frac{1}{a}(ah - gb)(af - cd) \\ &= \frac{1}{a}(a^2ei - aceg - abdi + bcdg - a^2fh + acdh + abfg - bcdg) \\ &= aei + bfg + cdh - ceg - afh - bdi. \end{aligned}$$

Ook in dit geval kunnen we nagaan dat deze formule ook voor de exceptionele gevallen $a = 0$ en $ae - bd = 0$ geldt en dus het algemene geval weergeeft.

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



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

Waarschuwing: Het schema van de 3×3-matrices geldt niet voor grotere matrices. Vanaf grootte 4×4 is het inderdaad nodig de matrix op rijtrapvorm te brengen.

14.3 Het bepalen van eigenwaarden en eigenvectoren

Om eigenwaarden en eigenvectoren te vinden hoeven we nu alleen nog de resultaten van boven te combineren: Een eigenvector voor de eigenwaarde λ is een vector in de kern van $A - \lambda \cdot I$ en de kern bevat dan en slechts dan andere elementen dan de nulvector als de determinant $\det(A - \lambda \cdot I) = 0$ is.

Hierbij interpreteren we λ als een variabel, dan wordt de determinant van $A - \lambda \cdot I$ een veelterm in de variabel λ. Deze veelterm heet het *karakteristieke* *polynoom* van de matrix A en de eigenwaarden van A zijn juist de nulpunten van het karakteristieke polynoom.

Voorbeeld: We bepalen de eigenwaarden en eigenvectoren van de matrix

$$A := \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}.$$

We hebben

$$A - \lambda \cdot I = \begin{pmatrix} 1-\lambda & 2 \\ -1 & 4-\lambda \end{pmatrix} \text{ en}$$
$$\det(A - \lambda \cdot I) = (1-\lambda)(4-\lambda) - (-2) = \lambda^2 - 5\lambda + 4 + 2 = \lambda^2 - 5\lambda + 6.$$

Om de eigenwaarden van A te vinden moeten we dus de nulpunten van $\lambda^2 - 5\lambda + 6$ bepalen. Dit kan altijd met de *abc*-formule, maar in dit geval zien we makkelijk dat $\lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3)$ en we hebben dus

$$\lambda^2 - 5\lambda + 6 = 0 \Leftrightarrow \lambda = 2 \text{ of } \lambda = 3.$$

De eigenwaarden van A zijn dus 2 en 3.

Het bepalen van de eigenvectoren is nu eenvoudig. We weten dat de matrix $A - \lambda \cdot I$ voor de gevonden eigenwaarden λ een vector $\neq 0$ in de kern heeft, en elke van deze vectoren is een eigenvector. Merk op dat met een eigenvector v ook elke veelvoud $c \cdot v$ een eigenvector voor dezelfde eigenwaarde is. We vinden dus de eigenvectoren van A als volgt:

$\lambda = 2$: $A - \lambda \cdot I = \begin{pmatrix} -1 & 2 \\ -1 & 2 \end{pmatrix}$ heeft kern $t \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, dus is $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ een eigenvector van A voor de eigenwaarde 2.

Men gaat eenvoudig na dat inderdaad $A \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$.

$\lambda = 3$: $A - \lambda \cdot I = \begin{pmatrix} -2 & 2 \\ -1 & 1 \end{pmatrix}$ heeft kern $t \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, dus is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ een eigenvector van A voor de eigenwaarde 3.

Men gaat weer na dat $A \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$.

Voor een algemene 2×2-matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is de determinant van $A - \lambda \cdot I$ de determinant van $\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}$ en dus gelijk aan

$$(a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + (ad - bc).$$

Dit is een kwadratische veelterm waarvan we de nulpunten altijd met behulp van de *abc*-formule kunnen bepalen.

De matrix in het voorbeeld van de Euro-munten was $A = \begin{pmatrix} 0.90 & 0.01 \\ 0.10 & 0.99 \end{pmatrix}$ en hiervoor hebben we

$$\det(A - \lambda \cdot I) = \lambda^2 - 1.89\lambda + 0.89 = (\lambda - 1)(\lambda - 0.89).$$

De eigenwaarden zijn dus 1 en 0.89 en we kunnen voor deze eigenwaarden nu ook de eigenvectoren bepalen:

De eigenvectoren voor de eigenwaarde 1 zijn de vectoren in de kern van $A - 1 \cdot I = \begin{pmatrix} -0.10 & 0.01 \\ 0.10 & -0.01 \end{pmatrix}$ en die is gelijk aan $t \cdot \begin{pmatrix} 1 \\ 10 \end{pmatrix}$ (wat we al wisten).

De eigenvectoren voor de eigenwaarde 0.89 zijn de vectoren in de kern van $A - 0.89 \cdot I = \begin{pmatrix} 0.01 & 0.01 \\ 0.10 & 0.10 \end{pmatrix}$ en die is gelijk aan $t \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Als we de twee eigenvectoren als basis voor $\mathbb{R}^2$ kiezen, wordt de overgangsmatrix met betrekking tot deze nieuwe basis heel eenvoudig, namelijk de diagonaalmatrix $D = \begin{pmatrix} 1 & 0 \\ 0 & 0.89 \end{pmatrix}$. We kennen ook de transformatiematrix van de nieuwe basis naar de standaardbasis, want die is gewoon $T = \begin{pmatrix} 1 & 1 \\ 10 & -1 \end{pmatrix}$. We weten dus dat $D = T^{-1} \cdot A \cdot T$ geldt. Wat we hieraan hebben zullen we straks zien.

14.4 Toepassingen van eigenwaarden en eigenvectoren

Limieten van overgangsmatrices

Bij processen die we door een overgangsmatrix A beschrijven, zijn we vaak geïnteresseerd in de ontwikkeling op langere termijn. Hiervoor hebben we de machten A^m voor grotere waarden van m nodig en soms zelfs een limiet voor $m \to \infty$. Dit is natuurlijk een lastig probleem, maar als we de eigenwaarden van A kennen en een basis uit eigenvectoren kunnen vinden, valt het inderdaad mee.

Stel dat we voor een $n \times n$-matrix A de eigenwaarden hebben bepaald en n lineair onafhankelijke basisvectoren hebben gevonden. Als $(v_1, \ldots, v_n)$ de eigenvectoren zijn en $(\lambda_1, \ldots, \lambda_n)$ de bijhorende eigenwaarden, dan zij T de matrix met v_i als i-de kolom. Door kolomsgewijs te vergelijken zien we dat $A \cdot T = T \cdot D$, waarbij D een diagonaalmatrix met de λ_i op de diagonaal is:

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Hieruit volgt dat

$$A = T \cdot D \cdot T^{-1} \quad \text{en} \quad D = T^{-1} \cdot A \cdot T.$$

Dit resultaat hadden we met behulp van de in de vorige les behandelde *basistranformatie* ook direct kunnen bereiken: We interpreteren A als matrix van een lineaire afbeelding met betrekking tot de standaardbasis, dan is de diagonaalmatrix D de matrix van dezelfde lineaire afbeelding, maar nu met betrekking tot de basis uit eigenvectoren geschreven. De matrix T met als kolommen de eigenvectoren is juist de matrix van de coördinaten van de eigenvectoren met betrekking tot de standaardbasis.

Voor deze situatie hadden we in de vorige les aangetoond dat

$$D = T^{-1} \cdot A \cdot T.$$

Het aardige is dat we de machten A^m van A nu veel makkelijker kunnen berekenen, want

$$A^m = (T \cdot D \cdot T^{-1})^m = T \cdot D^m \cdot T^{-1}$$

omdat

$$\begin{aligned} &(T \cdot D \cdot T^{-1}) \cdot (T \cdot D \cdot T^{-1}) \cdot \ldots \cdot (T \cdot D \cdot T^{-1}) \\ = {}& T \cdot D \cdot (T^{-1}T) \cdot D \cdot \ldots \cdot (T^{-1}T) D T^{-1} = T \cdot D \cdot D \cdot \ldots \cdot D \cdot T^{-1}. \end{aligned}$$

Maar de machten van D zijn gewoon

$$D^m = \begin{pmatrix} \lambda_1^m & 0 & \ldots & 0 \\ 0 & \lambda_2^m & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & \lambda_n^m \end{pmatrix}$$

en hiervoor kunnen we voor $m \to \infty$ makkelijk zeggen wat er gaat gebeuren: Als $|\lambda_i| < 1$ dan gaat λ_i^m naar 0, als $|\lambda_i| > 1$ dan gaat λ_i^m naar ∞, en als $|\lambda_i| = 1$ hangt het ervan af of $\lambda_i = 1$, $\lambda_i = -1$ of λ_i een complex getal is.

Bij veel processen heeft de overgangsmatrix A de eigenschap dat de som van alle kolommen gelijk is aan 1. In dit geval laat zich aantonen dat 1 altijd een eigenwaarde is en dat alle eigenwaarden absolute waarde ≤ 1 hebben.

Onder zekere voorwaarden (waaraan meestal voldaan is) geldt zelfs dat de kern van $A - I$ dimensie 1 heeft (dus dat er op scalaire na een unieke eigenvector met eigenwaarde 1 bestaat) en dat alle andere eigenwaarden van absolute waarde < 1 zijn.

Als we voor zo'n proces een basis uit eigenvectoren zo kiezen dat de eerste basisvector de eigenvector $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ met eigenwaarde 1 is, volgt dat D^m voor grote m naar de matrix D_∞ gaat waarbij het $(1,1)$-element 1 is en alle andere elementen 0 zijn. Maar dan gaat A^m naar

$$A_\infty = T \cdot D_\infty \cdot T^{-1} = T \cdot \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & 0 \end{pmatrix} \cdot T^{-1}$$

$$= \begin{pmatrix} v_1 & 0 & \dots & 0 \\ v_2 & 0 & \dots & 0 \\ & & \ddots & \\ v_n & 0 & \dots & 0 \end{pmatrix} \cdot T^{-1} = \begin{pmatrix} s_1v_1 & s_2v_1 & \dots & s_nv_1 \\ s_1v_2 & s_2v_2 & \dots & s_nv_2 \\ & & \ddots & \\ s_1v_n & s_2v_n & \dots & s_nv_n \end{pmatrix}$$

waarbij $(s_1, \dots, s_n)$ de elementen in de eerste rij van T^{-1} zijn. In dit geval is dus A_∞ een matrix waarin elke kolom een veelvoud van de eigenvector v met eigenwaarde 1 is.

In het voorbeeld van de Euro-munten vinden we dat de eerste rij van T^{-1} gelijk is aan $(\frac{1}{11}, \frac{1}{11})$, dus gaan de machten A^m voor de matrix $A = \begin{pmatrix} 0.90 & 0.01 \\ 0.10 & 0.99 \end{pmatrix}$ naar $A_\infty = \begin{pmatrix} \frac{1}{11} & \frac{1}{11} \\ \frac{10}{11} & \frac{10}{11} \end{pmatrix}$. Dit zegt dat er uiteindelijk zowel in Nederland als ook in het buitenland relatief hetzelfde aandeel van Nederlandse en van buitenlandse munten terecht komt, maar dat er in het buitenland tien keer zo veel munten zullen zijn dan in Nederland (onafhankelijk van de hoeveelheden munten in het begin).

In het voorbeeld van opgave 98 is de overgangsmatrix

$$A = \begin{pmatrix} 0 & 0 & 0.33 \\ 0.18 & 0 & 0 \\ 0 & 0.71 & 0.94 \end{pmatrix}.$$

De eigenwaarden hiervan zijn de wortels van de veelterm $\lambda^2(\lambda - 0.94) - 0.042174$. Deze zijn niet zo makkelijk met de hand te berekenen, maar bijvoorbeeld met een GRM of met een computeralgebra pakket zoals MAPLE vinden we als numerieke benaderingen: $\lambda_1 = 0.9835927398$, $\lambda_2 = -0.02179636988 + 0.2059184805i$, $\lambda_3 = -0.02179636988 - 0.2059184805i$. De eigenwaarden λ_2 en λ_3 zijn complexe getallen met absolute waarde 0.2070688348, dus zijn alle eigenwaarden van absolute waarde < 1, en dus gaan de machten van A tegen de nulmatrix. De vogels sterven dus uit, als er niets verandert!

Hoofdcomponenten analyse

Bij statistische evaluaties kijkt men vaak naar het gemiddelde van de metingen en naar de afwijkingen van de metingen van het gemiddelde. Als de metingen $x_1, \dots, x_n$ zijn, dan is het gemiddelde (verwachtingswaarde) $\mu = \frac{1}{n} \sum_{i=1}^n x_i$ en de variantie is $\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$.

Als we nu een experiment hebben met verschillende parameters, dan zijn de metingen vectoren met een zeker aantal componenten. Het gemiddelde vinden we weer door de metingen op te tellen en door het aantal metingen te delen, dit geeft voor drie parameters bijvoorbeeld een vector $\mu = \begin{pmatrix} \mu_x \\ \mu_y \\ \mu_z \end{pmatrix}$.

Het is nu niet meer vanzelfsprekend dat de grootste spreiding van de metingen in één van de richtingen van de parameters ligt, want misschien is er een combinatie van parameters, die het grootste effect op het experiment heeft. We willen dus graag de richting bepalen waarin de spreiding maximaal wordt.

De *afstand* in de richting van de vector $v = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ die een vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ van de gemiddelde vector μ heeft, bepalen we door

$$(a(x - \mu_x) + b(y - \mu_y) + c(z - \mu_z))^2$$

(dit zullen we in de volgende les nog verder toelichten) en voor de spreiding in deze richting krijgen we analoog als boven:

$$\sigma_v^2 = \sum_{i=1}^{n} (a(x_i - \mu_x) + b(y_i - \mu_y) + c(z_i - \mu_z))^2.$$

Dit kunnen we (een beetje kunstmatig) ook als een product van matrices schrijven, dan wordt het

$$\sigma_v^2 = \sum_{i=1}^{n} \begin{pmatrix} a & b & c \end{pmatrix} \cdot \begin{pmatrix} x_i - \mu_x \\ y_i - \mu_y \\ z_i - \mu_z \end{pmatrix} \cdot \begin{pmatrix} x_i - \mu_x & y_i - \mu_y & z_i - \mu_z \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$$= \begin{pmatrix} a & b & c \end{pmatrix} \cdot \left(\sum_{i=1}^{n} \begin{pmatrix} x_i - \mu_x \\ y_i - \mu_y \\ z_i - \mu_z \end{pmatrix} \cdot \begin{pmatrix} x_i - \mu_x & y_i - \mu_y & z_i - \mu_z \end{pmatrix} \right) \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

De matrix $V = \sum_{i=1}^{n} \begin{pmatrix} x_i - \mu_x \\ y_i - \mu_y \\ z_i - \mu_z \end{pmatrix} \cdot \begin{pmatrix} x_i - \mu_x & y_i - \mu_y & z_i - \mu_z \end{pmatrix}$ noemt men de *covariantiematrix* van de metingen $\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}, \ldots, \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}$.

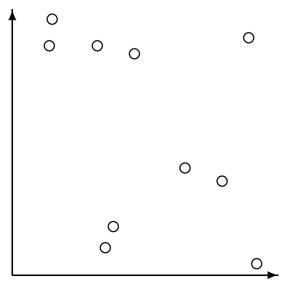
We zijn nu geïnteresseerd in de richting zo dat de spreiding σ_v maximaal wordt. Maar dat gebeurt precies als we voor v de eigenvector voor de grootste eigenwaarde van V kiezen.

In het algemeen wordt deze methode gebruikt om uit een groot aantal parameters de belangrijkste combinaties eruit te vissen. Deze combinaties heten de *hoofdcomponenten.* Men neemt dan bijvoorbeeld in een experiment met honderd parameters de eigenvectoren die bij de grootste tien eigenwaarden horen en heeft hiermee vaak de belangrijkste eigenschappen voor het experiment beschreven. Hier zijn twee voorbeelden van dit soort toepassingen:

(1) Het eerste voorbeeld vindt men in de automatische spraakherkenning, waarbij de signalen door een Fourier-analyse met betrekking tot een groot aantal frequenties worden beschreven. Om nu de verschillende klinkers en medeklinkers te kunnen onderscheiden, wordt gekeken welke combinaties van frequenties een grote spreiding hebben en op die manier wordt het signaal met veel minder parameters beschreven. Een voordeel hierbij is ook nog, dat de gereduceerde parameters robuuster tegen ruis en andere storingen zijn.

(2) Een andere toepassing zit in het comprimeren van informatie. Bij een kleur-plaatje wordt vaak de kleur in het RGB-systeem aangegeven, dat wil zeggen er wordt een waarde voor de intensiteiten van de kleuren rood, groen en blauw aangegeven. Meestal zijn de kleuren niet onafhankelijk, dan laat het plaatje zich ook met de intensiteiten van twee combinaties van rood, groen en blauw beschrijven, de spreiding in de richting van de derde eigenvector is dan heel klein.

Voorbeeld: We gaan een 2-dimensionaal voorbeeld bekijken, om te zien hoe de hoofdcomponenten analyse werkt.

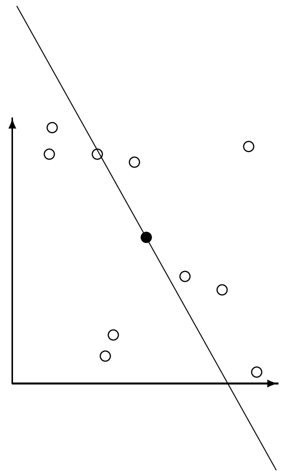


De punten in het plaatje hierboven zijn de tien punten met de coördinaten:

$$\begin{aligned}
(x_1, y_1) &= (14, 86);\\
(x_2, y_2) &= (15, 96);\\
(x_3, y_3) &= (92, 4);\\
(x_4, y_4) &= (65, 40);\\
(x_5, y_5) &= (35, 10);\\
(x_6, y_6) &= (89, 89);\\
(x_7, y_7) &= (79, 35);\\
(x_8, y_8) &= (32, 86);\\
(x_9, y_9) &= (38, 18);\\
(x_{10}, y_{10}) &= (46, 83).
\end{aligned}$$

De x-waarden zijn de eerste 10 paren van cijfers achter de komma in π, de y-waarden de eerste 10 paren van cijfers achter de komma in π^2.

Het middelpunt heeft de coördinaten $(50.5, 54.7)$ en voor de covariantie-matrix vinden we $C = \begin{pmatrix} 7578.5 & -3721.5 \\ -3721.5 & 12162.1 \end{pmatrix}$. De eigenwaarden van C zijn $\lambda_1 = 14240.87313$ en $\lambda_2 = 5499.726869$ en de eigenvector voor de eigenwaarde λ_1 is $\begin{pmatrix} -0.4876624923 \\ 0.8730322409 \end{pmatrix}$. Dit is dus de richting van de grootste spreiding. die in het volgende plaatje als lijn door het middelpunt te zien is.



Belangrijke begrippen in deze les

- eigenwaarden, eigenvectoren
- determinant
- karakteristieke polynoom
- hoofdcomponenten analyse

Opgaven

104. Bereken de determinanten van de volgende matrices:

$$A := \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \qquad B := \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 0 & -1 & 1 \\ 2 & 1 & 2 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}.$$

105. Bereken de determinant van de volgende matrices:

$$A := \begin{pmatrix} x-5 & 7 \\ -1 & x+3 \end{pmatrix}, \qquad B := \begin{pmatrix} 1+x & 1 & 1 \\ 1 & 1+x & 1 \\ 1 & 1 & 1+x \end{pmatrix}.$$

106. Bereken de volgende determinanten van 3×3-matrices:

$$\text{(i) } \det \begin{pmatrix} 2 & 1 & 1 \\ 0 & 5 & -2 \\ 1 & -3 & 4 \end{pmatrix}, \quad \text{(ii) } \det \begin{pmatrix} 2 & 0 & -1 \\ 3 & 0 & 2 \\ 4 & -3 & 7 \end{pmatrix}, \quad \text{(iii) } \det \begin{pmatrix} 3 & 2 & -4 \\ 1 & 0 & -2 \\ -2 & 3 & 3 \end{pmatrix},$$

$$\text{(iv) } \det \begin{pmatrix} 7 & 6 & 5 \\ 1 & 2 & 1 \\ 3 & -2 & 1 \end{pmatrix}, \quad \text{(v) } \det \begin{pmatrix} t+3 & -1 & 1 \\ 5 & t-3 & 1 \\ 6 & -6 & t+4 \end{pmatrix}, \quad \text{(vi) } \det \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}.$$

107. Bereken de volgende determinanten van 4×4-matrices:

$$\det\begin{pmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 2 \\ -1 & -6 & 4 & 3 \end{pmatrix}, \quad \det\begin{pmatrix} 3 & -2 & -5 & 4 \\ -5 & 2 & 8 & -5 \\ -2 & 4 & 7 & -3 \\ 2 & -3 & -5 & 8 \end{pmatrix}, \quad \det\begin{pmatrix} 1 & 2 & 2 & 3 \\ 1 & 0 & -2 & 0 \\ 3 & -1 & 1 & -2 \\ 4 & -3 & 0 & 2 \end{pmatrix}.$$

108. Bepaal voor de matrix $A := \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}$ de eigenwaarden en eigenvectoren. Beschrijf de lineaire afbeelding met matrix A meetkundig.

109. Bepaal de eigenwaarden en eigenvectoren van de volgende matrices. Geef in ieder geval aan of er een basis van $\mathbb{R}^2$ c.q. $\mathbb{R}^3$ bestaat die alleen maar eigenvectoren van de betreffende matrix bevat.

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}, \quad D = \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix},$$

$$E = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{pmatrix}, \qquad F = \begin{pmatrix} 0 & 7 & -6 \\ -1 & 4 & 0 \\ 0 & 2 & -2 \end{pmatrix}, \qquad G = \begin{pmatrix} 2 & -2 & 3 \\ 10 & -4 & 5 \\ 5 & -4 & 6 \end{pmatrix}.$$

110. In de 3-dimensionale ruimte $\mathbb{R}^3$ heeft een rotatie om een hoek φ rond de z-as de matrix $\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) & 0 \\ \sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

De matrix $A := \begin{pmatrix} -\frac{1}{6} & \frac{2}{3} & -\frac{1}{6} \\ -\frac{1}{2} & 1 & \frac{1}{2} \\ \frac{13}{6} & -\frac{2}{3} & \frac{1}{6} \end{pmatrix}$ is ook de matrix van een rotatie, maar niet rond de z-as.

Wat is de rotatie-as van de rotatie A?

111. Zij A de matrix $A := \begin{pmatrix} -4 & 1 \\ -6 & 1 \end{pmatrix}$.

(i) Bereken het karakteristieke polynoom $\det(A - \lambda \cdot I)$, de eigenwaarden en de eigenvectoren van A.

(ii) Bepaal voor willekeurige $m \in \mathbb{N}$ de matrix A^m.

112. Bepaal voor de matrix

$$A = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{pmatrix}$$

(i) het karakteristieke polynoom $\det(A - \lambda \cdot I)$,

(ii) de eigenwaarden van A,

(iii) de eigenvectoren van A,

(iv) de matrix A_∞ waar A^m voor grote m naar toe gaat.

(Hint: $\lambda = 1$ is een eigenwaarde.)

Les 15 Inproducten

Als we het in de meetkunde (of elders) over afstanden en hoeken hebben, dan hebben we daar intuïtief wel een idee van. Maar wat is eigenlijk de afstand van twee punten in het 2-dimensionale vlak?

Als we twee punten (x_1, y_1) en (x_2, y_2) hebben dan is hun afstand de *lengte van de vector tussen de twee punten*. We kunnen dus het meten van afstanden terug brengen op het meten van de lengte van vectoren.

Maar de lengte van de vector $\begin{pmatrix} x \\ y \end{pmatrix}$ is natuurlijk $\sqrt{x^2 + y^2}$, hiervoor hoeven we alleen maar de stelling van Pythagoras op de driehoek met hoekpunten $(0, 0)$, $(x, 0)$ en (x, y) toe te passen: De schuine zijde is de vector waarvan we de lengte l willen weten en de rechthoekzijden hebben lengtes x en y, dus geldt $l^2 = x^2 + y^2$.

In de 3-dimensionale ruimte berekenen we de lengte van een vector op een soortgelijke manier, de lengte van $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is $\sqrt{x^2 + y^2 + z^2}$. Ook hier wordt de stelling van Pythagoras toegepast, in dit geval twee keer: eerst in de $x - y$-vlakte en vervolgens in de vlakte loodrecht op de vector $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$.

Het is nu voor de hand liggend, dat we in een n-dimensionale vectorruimte de lengte van de vector $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ aangeven met $\sqrt{x_1^2 + \ldots + x_n^2}$. Dit komt er op neer, de stelling van Pythagoras herhaald $n - 1$ keer toe te passen.

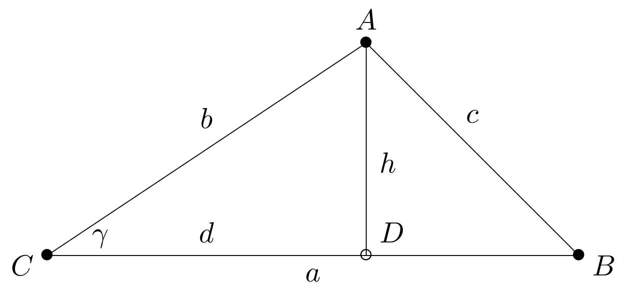
> Men noemt $\sqrt{x_1^2 + \ldots + x_n^2}$ ook de *Euclidische lengte* van de vector $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ om duidelijk te maken dat de lengte in de gewone *Euclidische meetkunde* bedoeld is. Er zijn namelijk ook andere mogelijkheden om de lengte van een vector te definiëren, bijvoorbeeld in de zogeheten *hyperbolische meetkunde*. Terwijl punten met dezelfde afstand van een vast punt in de Euclidische meetkunde op een cirkel of bol liggen, liggen ze in de hyperbolische meetkunde op een hyperbool of hyperboloïde.

Hoe zit het nu met hoeken? We zijn al klaar als we weten wat een hoek in het 2-dimensionale vlak is, want ook in een n-dimensionale vectorruimte liggen twee vectoren altijd in een 2-dimensionale deelruimte en we definiëren de hoek tussen de vectoren gewoon als de hoek in dit vlak.

Het aardige is nu, dat we de definitie van hoeken tussen vectoren uit de definitie voor lengtes van vectoren af kunnen leiden. Dit zal verder geen verrassing zijn, want bij een driehoek waarvan we de lengtes van de zijden kennen, liggen ook de hoeken vast. Het cruciale punt voor de samenhang tussen zijden en hoeken is de *cosinusregel*.

C.27 Cosinusregel *In een driehoek met zijden a, b en c en met de hoek γ tussen de zijden a en b geldt:*

$$c^2 = a^2 + b^2 - 2ab\cos(\gamma).$$



In het geval van een rechthoekige driehoek vinden we de stelling van Pythagoras terug, want dan is γ en hoek van 90 graden en dus $\cos(\gamma) = 0$. Als γ een scherpe hoek is (kleiner dan 90 graden) is $\cos(\gamma) > 0$, dus wordt de tegenover γ liggende zijde c in dit geval korter dan in het rechthoekige driehoek, voor een stompe hoek γ wordt c groter.

Het bewijs van de cosinusregel kunnen we eenvoudig uit het plaatje hierboven aflezen:

De driehoeken CDA en ADB zijn rechthoekig, dus geldt $b^2 = h^2 + d^2$ en $c^2 = h^2 + (a-d)^2$. Als we deze vergelijkingen van elkaar aftrekken krijgen we $b^2 - c^2 = d^2 - a^2 + 2ad - d^2 = -a^2 + 2ad$ en dus $c^2 = a^2 + b^2 - 2ad$. Verder geldt $d = b\cos(\gamma)$ en als we dit voor d invullen, hebben we precies de bewering van de cosinusregel. Voor een stompe hoek γ werkt het bewijs op een analoge manier.

15.1 Het standaardinproduct

We vertalen het berekenen van een hoek in een driehoek nu naar vectoren. Als we de hoek tussen twee vectoren v en w willen weten, verbinden we de punten van de vectoren, en de vector die daar bij hoort is de verschilvector $v - w$. De lengte van een vector v schrijven we als $\|v\|$ en noemen dit vaak ook de *norm* van de vector v. De cosinusregel zegt dat voor de hoek γ tussen de vectoren v en w geldt dat

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\,\|v\|\|w\|\cos(\gamma).$$

Als we deze vergelijking nu voor de twee vectoren $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ en $w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ uitschrijven, krijgen we

$$(v_1 - w_1)^2 + (v_2 - w_2)^2 = v_1^2 + v_2^2 + w_1^2 + w_2^2 - 2\,\|v\| \cdot \|w\|\cos(\gamma)$$

dus $-2v_1w_2 - 2v_2w_2 = -2\,\|v\| \cdot \|w\| \cos(\gamma)$. Voor de hoek γ tussen v en w geldt dus:

$$\cos(\gamma) = \frac{v_1w_1 + v_2w_2}{\|v\|\|w\|}.$$

De teller van deze formule kunnen we beschouwen als een functie van de vectoren v en w, namelijk $\Phi(v,w) = v_1w_1 + v_2w_2$ als v en w zo als boven gegeven zijn.

De functie $\Phi(v,w)$ noemen we een *inproduct* en gebruiken hiervoor meestal de notatie

$$\Phi(v,w) = \langle v, w\rangle.$$

Met deze notatie is de hoek tussen de vectoren v en w gedefinieerd door

$$\cos(\gamma) = \frac{\langle v, w\rangle}{\|v\|\|w\|}.$$

De berekening waarmee we geconcludeerd hebben dat $\|v\|\|w\|\cos(\gamma) = v_1w_1 + v_2w_2$ gaat precies hetzelfde door als v en w niet meer 2-dimensionaal maar n-dimensionale vectoren zijn. Voor $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ en $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ krijgen we dan

$$\|v\|\|w\|\cos(\gamma) = v_1w_1 + v_2w_2 + \ldots + v_nw_n.$$

C.28 Definitie Het *standaardinproduct* van twee vectoren $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ en $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ in de n-dimensionale vectorruimte $\mathbb{R}^n$ is gedefinieerd door

$$\langle v, w\rangle := v_1w_1 + v_2w_2 + \ldots + v_nw_n = \sum_{i=1}^{n} v_iw_i.$$

In het bijzonder geldt voor iedere vector $v \in \mathbb{R}^n$ dat

$$\|v\|^2 = v_1^2 + v_2^2 + \ldots + v_n^2 = \sum_{i=1}^{n} v_i^2 = \langle v, v\rangle,$$

d.w.z. het standaardinproduct van een vector met zich zelf is het kwadraat van zijn (Euclidische) norm.

We kunnen het standaardinproduct ook als een matrix product opvatten, hiervoor schrijven we de eerste vector als een matrix met 1 rij en n kolommen en de tweede als een matrix met n rijen en 1 kolom, dan krijgen we

$$\langle v, w\rangle = \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = v^{tr} \cdot w.$$

Hierbij noemen we v^{tr} de *getransponeerde* vector van v, d.w.z. de vector v geschreven als een rij in plaats van een kolom.

15.2 Inproducten

Als we het standaardinproduct als matrix product interpreteren, zien we meteen in dat het zekere eigenschappen heeft:

$$\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle \quad \text{en} \quad \langle \lambda v, w\rangle = \lambda\langle v, w\rangle$$
$$\langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_2\rangle \quad \text{en} \quad \langle v, \lambda w\rangle = \lambda\langle v, w\rangle$$

Als we de tweede vector in het inproduct vast laten, hebben we dus een lineaire afbeelding van het eerste argument, net zo krijgen we een lineaire afbeelding van het tweede argument als we de eerste vector vast laten. Dit geeft aanleiding om de inproducten in een algemener context in te bedden. Op een soortgelijke manier kunnen we namelijk een heel algemene klasse van afbeeldingen beschrijven, de *bilineaire afbeeldingen*.

C.29 Definitie Voor een vectorruimte V heet een afbeelding Φ, die aan een paar (v, w) van vectoren een getal $\Phi(v, w) \in \mathbb{R}$ toewijst een *bilineaire afbeelding* als geldt:

(i) $\Phi(v_1 + v_2, w) = \Phi(v_1, w) + \Phi(v_2, w)$ voor alle $v_1, v_2, w \in V$

(ii) $\Phi(\lambda v, w) = \lambda\Phi(v, w)$ voor alle $v, w \in V$, $\lambda \in \mathbb{R}$

(iii) $\Phi(v, w_1 + w_2) = \Phi(v, w_1) + \Phi(v, w_2)$ voor alle $v, w_1, w_2 \in V$

(iv) $\Phi(v, \lambda w) = \lambda\Phi(v, w)$ voor alle $v, w \in V$, $\lambda \in \mathbb{R}$.

Als we voor een bilineaire afbeelding één van de argumenten vast laten, krijgen we dus een lineaire afbeelding in het andere argument, daarom de naam.

Voorbeeld: We kijken naar een voorbeeld van een bilineaire afbeelding dat niets meer met de gewone meetkundige aanschouwing te maken heeft, maar het algemenere concept van bilineaire afbeeldingen illustreert.

Neem als vectorruimte de veeltermfuncties van graad ≤ 2, dan kunnen we een bilineaire afbeelding definiëren door

$$\Psi(f(x), g(x)) := \int_0^1 f(x)g(x)dx.$$

Dat dit inderdaad een bilineaire afbeelding is, volgt uit de rekenregels voor integralen. We kiezen de drie eenvoudige veeltermen $(1, x, x^2)$ als basis, dan geldt $\Psi(1, 1) = 1$, $\Psi(1, x) = \Psi(x, 1) = \frac{1}{2}$, $\Psi(1, x^2) = \Psi(x, x) = \Psi(x^2, 1) = \frac{1}{3}$, $\Psi(x, x^2) = \Psi(x^2, x) = \frac{1}{4}$ en $\Psi(x^2, x^2) = \frac{1}{5}$.

We hebben gezien dat het standaardinproduct een speciaal geval van een bilineaire afbeelding is. Maar niet alle bilineaire afbeeldingen zijn geschikt als we het over lengtes en hoeken willen hebben.

In het geval van het standaardinproduct zien we, dat de lengte $\|v\|$ van een vector v gelijk is aan $\sqrt{\langle v, v\rangle}$. Als we nu op een abstracte vectorruimte een bilineaire afbeelding Φ hebben waarmee we een lengte willen definiëren, dan zullen we precies deze samenhang gebruiken.

Omdat het schrijven van wortels soms onhandig is, wordt meestal het kwadraat van de lengte gedefinieerd. Maar hier zit al een belangrijk punt in: Als $\Phi(v,v) < 0$ is, is het kwadraat van de lengte negatief, maar dan is de lengte geen zinvol reëel getal meer.

Ook vectoren van lengte 0 zijn onhandig, we willen dat alleen maar de nulvector lengte 0 heeft. Hieruit volgen de volgende eisen aan een 'nuttige' bilineaire afbeelding Φ:

(i) $\Phi(v,v) \geq 0$ voor alle $v \in V$

(ii) $\Phi(v,v) = 0$ alleen maar voor $v = 0$.

Een bilineaire afbeelding die aan de eisen (i) en (ii) voldoet noemen we *positief definiet*.

Verder hadden we voor het standaardinproduct gezien dat we de hoek tussen twee vectoren kunnen berekenen uit de formule $\cos(\gamma) = \frac{\Phi(v,w)}{\|v\|\|w\|}$. Maar de hoek is natuurlijk onafhankelijk van de volgorde van de vectoren, dus willen we nog graag hebben dat Φ *symmetrisch* is, dus

(iii) $\Phi(v,w) = \Phi(w,v)$ voor alle $v, w \in V$.

C.30 Definitie Een positief definiete, symmetrische bilineaire afbeelding Φ op een vectorruimte V, d.w.z. een bilineaire afbeelding met

(i) $\Phi(v,v) \geq 0$ voor alle $v \in V$;

(ii) $\Phi(v,v) = 0$ alleen maar voor $v = 0$;

(iii) $\Phi(v,w) = \Phi(w,v)$ voor alle $v, w \in V$;

heet een *inproduct* op V.

We zien onmiddellijk in dat het standaardinproduct $\langle v, w\rangle = \sum_{i=1}^n v_i w_i$ een inproduct op $\mathbb{R}^n$ is, de naam is dus terecht gekozen.

Dat de bilineaire afbeelding $\Psi(f(x), g(x)) = \int_0^1 f(x)g(x)dx$ symmetrisch is, is duidelijk en Ψ voldoet ook aan eigenschap (i), want $f(x)^2 \geq 0$ en een integraal over waarden ≥ 0 is zelfs ook ≥ 0. Er zit wel een klein probleempje met eigenschap (ii). De functie $f(x)$ die alleen maar in het punt 1 de waarde 1 heeft, maar elders 0 is, heeft inderdaad integraal 0 zonder zelf de nulvector te zijn. Het probleem zit erin dat dit een functie met een sprong is, dus een niet-continue functie. Als we ons tot continue functies beperken (of nog sterker tot veeltermfuncties), is Ψ wel positief definiet: Als $f(x_0)^2 > 0$ dan vinden we een klein stukje om x_0 heen, waar ook $f(x)^2 > 0$ geldt, omdat $f(x)$ geen sprongen heeft. Maar dan hebben we een positieve vlakte onder de grafiek van $f(x)^2$ gevonden en is de integraal dus > 0.

Waarschuwing: In de toepassingen zijn niet alleen maar inproducten belangrijk. Het meest prominente voorbeeld van een bilineaire afbeelding die wel symmetrisch maar niet positief definiet is, is de afstand in de 4-dimensionale ruimtetijd van de relativiteitstheorie. Hier wordt de zogeheten *Minkowski-afstand* gebruikt die voor een punt met ruimtecoördinaten (x, y, z) en tijd t de afstand $x^2 + y^2 + z^2 - (ct)^2$ geeft. Hierbij is c de snelheid van het licht. Het idee is dat twee punten afstand nul hebben als hun posities in de ruimtetijd door een lichtsignaal verbonden kunnen worden.

15.3 Orthogonale projecties

In de 2-dimensionale vlakte en de 3-dimensionale ruimte is het duidelijk dat twee vectoren *loodrecht* op elkaar staan als hun standaardinproduct 0 is. We gaan daarom ook in het algemeen zeggen, dat twee vectoren v, w ten opzichte van een inproduct loodrecht op elkaar staan, als $\langle v, w\rangle = 0$ is en we noteren dit met $v \perp w$. In plaats van te zeggen dat twee vectoren v en w met inproduct $\langle v, w\rangle = 0$ loodrecht op elkaar staan, zegt men vaak dat v en w *orthogonaal* zijn.

C.31 Definitie Zij $\langle v, w\rangle$ een inproduct op de vectorruimte V.

(i) Twee vectoren v en w heten *orthogonaal* als $\langle v, w\rangle = 0$. Dit wordt genoteerd met $v \perp w$.

(ii) Een stelsel $(v_1, \ldots, v_n)$ van vectoren $v_i \neq 0$ (bijvoorbeeld een basis) heet *orthogonaal*, als $\langle v_i, v_j\rangle = 0$ voor alle $i \neq j$ is.

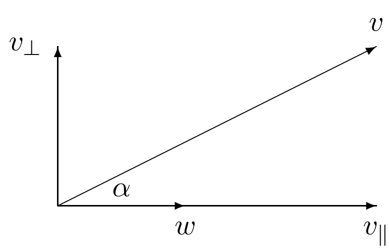
(iii) Een orthogonale basis $(v_1, \ldots, v_n)$ van V waarin alle vectoren lengte 1 hebben, dus met $\|v_i\|^2 = \langle v_i, v_i\rangle = 1$ voor alle i heet een *orthonormale basis* van V.

Men ziet makkelijk in dat een orthogonaal stelsel noodzakelijk lineair onafhankelijk is. Stel namelijk dat $v_n = c_1 v_1 + \ldots + c_{n-1} v_{n-1}$ een lineaire combinatie van de andere vectoren is. Dan is $\langle v_n, v_n\rangle = c_1\langle v_1, v_n\rangle + \ldots + c_{n-1}\langle v_{n-1}, v_n\rangle = 0$ omdat $\langle v_i, v_n\rangle = 0$ voor $1 \leq i \leq n-1$. Maar $\langle v_n, v_n\rangle = 0$ geldt alleen maar voor de nulvector, en die zit (volgens de definitie) nooit in een orthogonaal stelsel.

Een belangrijke eigenschap van het inproduct is, dat zich hiermee de orthogonale projectie van een vector op de lijn door een andere vector handig laat beschrijven. Er geldt namelijk, dat de lineaire afbeelding

$$\pi_w(v) := \langle v, w\rangle \frac{w}{\|w\|^2} = \frac{\langle v, w\rangle}{\langle w, w\rangle}\, w$$

de orthogonale projectie op de lijn door w weergeeft. Dit kunnen we goed aan het volgende plaatje aflezen:



We schrijven v als $v = v_{\|} + v_{\perp}$, waarbij $v_{\|}$ parallel met w is en $v_{\perp}$ loodrecht op w staat, dan is $v_{\|}$ de orthogonale projectie van v langs de lijn door w en er geldt $v_{\|} = v - v_{\perp}$.

Omdat $v_{\perp}$ loodrecht op w staat, geldt $\langle v_{\perp}, w\rangle = 0$ en dus $\langle v - v_{\|}, w\rangle = 0$. Maar $v_{\|}$ ligt in de richting van w en is dus een veelvoud λw van w. Er geldt dus $\langle v - \lambda w, w\rangle = \langle v, w\rangle - \langle \lambda w, w\rangle = 0$ en dus $\langle v, w\rangle = \lambda \langle w, w\rangle$.

C.32 Stelling *Voor de orthogonale projectie $v_{\|}$ van v langs de lijn door w geldt*

$$v_{\|} = \lambda w = \frac{\langle v, w\rangle}{\langle w, w\rangle}\, w.$$

Maar hoe kunnen we nu de orthogonale projectie op een 2-dimensionaal vlak of zelfs een hoger-dimensionale deelruimte berekenen? We kunnen voor een basis natuurlijk de orthogonale projecties op de lijnen door de basisvectoren berekenen. Maar let op: Als de basisvectoren niet loodrecht op elkaar staan is er een probleem. De projectie in de richting van de tweede basisvector heeft dan misschien ook een component in de richting van de eerste basisvector, en dan is de orthogonale projectie in het vlak niet de som van de twee projecties op de lijnen.

Waarschuwend voorbeeld: We willen de vector $v = \begin{pmatrix}1\\2\\3\end{pmatrix}$ in het $x-y$-vlak projecteren. Het is duidelijk dat de projectie $\begin{pmatrix}1\\2\\0\end{pmatrix}$ is, want we hoeven alleen maar de z-component te vergeten. Maar stel we nemen als basis voor het $x-y$-vlak de vectoren $v_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}$ en $v_2 = \begin{pmatrix}1\\1\\0\end{pmatrix}$. Dan is de projectie van v in de richting van v_1 gegeven door $\frac{\langle v, v_1\rangle}{\langle v_1, v_1\rangle}\, v_1 = 1 \cdot v_1$ en de projectie van v in de richting van v_2 is $\frac{\langle v, v_2\rangle}{\langle v_2, v_2\rangle}\, v_2 = \frac{3}{2} \cdot v_2$. De som van deze twee projecties is $\begin{pmatrix}\frac{5}{2}\\\frac{3}{2}\\0\end{pmatrix}$ en dit is niet de juiste projectie in het $x-y$-vlak.

Dit voorbeeld is een illustratie van de volgende stelling:

C.33 Stelling *De orthogonale projectie vanuit een vectorruimte V op een deelruimte U met basis $(v_1, v_2, \ldots, v_r)$ is de som van de projecties in de richtingen van de basisvectoren v_i als de v_i een orthogonale basis van U vormen. Dit geldt* niet *als de v_i niet orthogonaal zijn.*

Voordat we met een projectie aan de slag kunnen, moeten we volgens deze stelling dus eerst ervoor zorgen dat de ruimte waar we op willen projecteren een orthogonale basis heeft. Maar dit kunnen we bereiken, door stapsgewijs van elke basisvector de projecties op de eerder behandelde basisvectoren af te trekken. Dit geeft aanleiding tot een procedure die *Gram-Schmidt orthogonalisatie* heet.

Gram-Schmidt orthogonalisatie

We gaan uit van een basis $(v_1, \ldots, v_n)$ en willen een nieuwe basis $(v'_1, \ldots, v'_n)$ vinden, die orthogonaal is. We construeren de nieuwe orthogonale basis vector voor vector.

- De eerste vector v_1 kunnen we ongedeerd laten, we nemen dus
$$v'_1 := v_1.$$
- Nu hebben we nodig dat v'_2 loodrecht op v'_1 staat. Hiervoor trekken we de orthogonale projectie van v_2 op de lijn door v'_1 van v_2 af, de rest staat dan loodrecht op v'_1. We krijgen dus
$$v'_2 := v_2 - \frac{\langle v_2, v'_1 \rangle}{\langle v'_1, v'_1 \rangle}\, v'_1.$$
- Op deze manier gaan we nu door. Van de vector v_i trekken we de projecties op alle eerder bepaalde basisvectoren in de orthogonale basis af, dus krijgen we algemeen:
$$\begin{aligned} v'_i &:= v_i - \left(\sum_{j=1}^{i-1} \frac{\langle v_i, v'_j \rangle}{\langle v'_j, v'_j \rangle}\, v'_j \right) \\ &= v_i - \frac{\langle v_i, v'_1 \rangle}{\langle v'_1, v'_1 \rangle}\, v'_1 - \frac{\langle v_i, v'_2 \rangle}{\langle v'_2, v'_2 \rangle}\, v'_2 - \ldots - \frac{\langle v_i, v'_{i-1} \rangle}{\langle v'_{i-1}, v'_{i-1} \rangle}\, v'_{i-1}. \end{aligned}$$

Als we de basis nog mooier willen maken, kunnen we ieder van de vectoren v'_i nog op lengte 1 brengen door v'_i door $\|v'_i\|$ te delen. Dit geeft dan een orthonormale basis.

Voorbeeld: We bepalen een orthogonale basis voor het voorbeeld van de veeltermfuncties met inproduct $\Psi(f(x), g(x)) = \int_0^1 f(x)g(x)dx$. We gaan uit van de basis $(v_1, v_2, v_3) = (1, x, x^2)$.

- We beginnen met $v'_1 = v_1 = 1$.
- Vervolgens hebben we $v'_2 = v_2 - \frac{\Psi(v_2, v'_1)}{\Psi(v'_1, v'_1)} v'_1 = x - \frac{\frac{1}{2}}{1} \cdot 1 = x - \frac{1}{2}$.
- Tenslotte is $v'_3 = v_3 - \frac{\Psi(v_3, v'_1)}{\Psi(v'_1, v'_1)} v'_1 - \frac{\Psi(v_3, v'_2)}{\Psi(v'_2, v'_2)} v'_2 = x^2 - \frac{\frac{1}{3}}{1} \cdot 1 - \frac{\frac{1}{12}}{\frac{1}{12}}(x - \frac{1}{2}) = x^2 - \frac{1}{3} - (x - \frac{1}{2}) = x^2 - x + \frac{1}{6}$.

De orthogonale basis die we met behulp van de Gram-Schmidt orthogonalisatie vinden, is dus
$$(1, x - \frac{1}{2}, x^2 - x + \frac{1}{6}).$$

We kunnen vanaf nu ervan uitgaan dat we voor een orthogonale projectie over een orthogonale basis van de deelruimte beschikken, waarop we willen projecteren. Desnoods moeten we zo'n orthogonale basis met behulp van de Gram-Schmidt orthogonalisatie eerst construeren.

Voorbeeld: We berekenen de projectie op een zeker 2-dimensionaal vlak in de 3-dimensionale ruimte. Stel we willen de projectie op het vlak U waarin de vectoren $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ en $v_2 = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}$ liggen.

Eerst hebben we een orthogonale basis voor U nodig. Dit wordt (met behulp van de Gram-Schmidt orthogonalisatie) $v_1' = v_1$ en $v_2' = v_2 - \frac{1}{3}v_1' = \frac{1}{3}\begin{pmatrix} -1 \\ 5 \\ -4 \end{pmatrix}$. De factor $\frac{1}{3}$ kunnen we ook weglaten, daarom kiezen we als orthogonale basis van U de basis

$$\left(w_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, w_2 = \begin{pmatrix} -1 \\ 5 \\ -4 \end{pmatrix} \right).$$

De kwadraten van de lengtes van deze vectoren zijn $\|w_1\|^2 = \langle w_1, w_1 \rangle = 3$ en $\|w_2\|^2 = \langle w_2, w_2 \rangle = 42$.

Voor de inproducten van deze basis met de eenheidsvectoren $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ geldt $\langle e_1, w_1 \rangle = 1$, $\langle e_2, w_1 \rangle = 1$, $\langle e_3, w_1 \rangle = 1$, $\langle e_1, w_2 \rangle = -1$, $\langle e_2, w_2 \rangle = 5$, $\langle e_3, w_2 \rangle = -4$ De projecties van de eenheidsvectoren in het vlak U dat door w_1 en w_2 (en dus ook door v_1 en v_2) opgespannen is, zijn dus:

$$e_1 \mapsto \frac{1}{3}w_1 + \frac{-1}{42}w_2 = \begin{pmatrix} \frac{1}{3} + \frac{1}{42} \\ \frac{1}{3} - \frac{5}{42} \\ \frac{1}{3} + \frac{4}{42} \end{pmatrix} = \begin{pmatrix} \frac{5}{14} \\ \frac{3}{14} \\ \frac{6}{14} \end{pmatrix},$$

$$e_2 \mapsto \frac{1}{3}w_1 + \frac{5}{42}w_2 = \begin{pmatrix} \frac{1}{3} - \frac{5}{42} \\ \frac{1}{3} + \frac{25}{42} \\ \frac{1}{3} - \frac{20}{42} \end{pmatrix} = \begin{pmatrix} \frac{3}{14} \\ \frac{13}{14} \\ \frac{-2}{14} \end{pmatrix},$$

$$e_3 \mapsto \frac{1}{3}w_1 + \frac{-4}{42}w_2 = \begin{pmatrix} \frac{1}{3} + \frac{4}{42} \\ \frac{1}{3} - \frac{20}{42} \\ \frac{1}{3} + \frac{16}{42} \end{pmatrix} = \begin{pmatrix} \frac{6}{14} \\ \frac{-2}{14} \\ \frac{10}{14} \end{pmatrix}.$$

Maar meestal zijn we inderdaad niet geïnteresseerd in de projecties zelf maar in hun coëfficiënten met betrekking tot een orthonormaalbasis van de deelruimte, bijvoorbeeld als we de projecties in een vlak willen tekenen. Als orthonormaalbasis van U krijgen we $(w_1' = \frac{1}{\sqrt{3}}w_1, w_2' = \frac{1}{\sqrt{42}}w_2)$ en beschouwen deze als de standaardbasis van U. Dit levert de projecties

$$e_1 \mapsto \frac{1}{\sqrt{3}}w_1' + \frac{-1}{\sqrt{42}}w_2' \to \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{42}} \end{pmatrix},$$

$$e_2 \mapsto \frac{1}{\sqrt{3}}w_1' + \frac{5}{\sqrt{42}}w_2' \to \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{5}{\sqrt{42}} \end{pmatrix},$$

$$e_3 \mapsto \frac{1}{\sqrt{3}} w_1' + \frac{-4}{\sqrt{42}} w_2' \to \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{-4}{\sqrt{42}} \end{pmatrix}.$$

Het principe van een orthogonale projectie op een deelruimte wordt ook in de Fourieranalyse gebruikt die bijvoorbeeld in de analyse en verwerking van beeld- of geluidssignalen een rol speelt. Hierbij worden periodieke functies (trillingen) beschreven door hun projectie in de deelruimte die door de functies $\sin(kx)$ en $\cos(kx)$ wordt opgespannen. De functies $\sin(kx)$ en $\cos(kx)$ vormen namelijk een orthogonaal stelsel en worden door een geschikte schaling een orthonormale basis.

Bij de hoofdcomponenten analyse hebben we de covariantiematrix $A = \sum_{i=1}^{n} v_i \cdot v_i^{tr}$ berekend, waarbij elke vector v_i het verschil tussen een gemeten vectoren en het gemiddelde van de gemeten vectoren is. De eigenvector van de matrix A voor de grootste eigenwaarde geeft de richting van de grootste spreiding van de resultaten. Het aardige is nu, dat de eigenvectoren voor verschillende eigenwaarden met betrekking tot het standaardinproduct automatisch orthogonaal zijn omdat A een symmetrische matrix is, d.w.z. omdat A gelijk aan zijn *getransponeerde* (of *gespiegelde*) matrix A^{tr} is, waarbij de kolommen als rijen worden geschreven. Voor symmetrische matrices geldt namelijk de volgende stelling:

C.34 Stelling *Voor een symmetrische matrix $A = A^{tr}$ zijn eigenvectoren met verschillende eigenwaarden orthogonaal met betrekking tot het standaardinproduct. Dit betekent: Is $A \cdot v = \lambda v$ en $A \cdot w = \mu w$ met $\lambda \neq \mu$ dan is $\langle v, w \rangle = 0$.*

Deze stelling kunnen we makkelijk inzien: Neem aan dat v een eigenvector met eigenwaarde λ en w een eigenvector met eigenwaarde μ is en dat $\lambda \neq \mu$. Er geldt:

$$\begin{aligned} \lambda\langle v, w\rangle &= \langle \lambda v, w\rangle = \langle Av, w\rangle = (Av)^{tr} w = (v^{tr} A^{tr}) w \\ &= v^{tr}(A^{tr} w) = v^{tr}(Aw) = \langle v, Aw\rangle = \langle v, \mu w\rangle = \mu\langle v, w\rangle \end{aligned}$$

dus is $(\lambda - \mu)\langle v, w\rangle = 0$, en dus $\langle v, w\rangle = 0$ omdat $\lambda \neq \mu$.

In de hoofdcomponenten analyse hoeven we dus alleen maar de eigenvectoren van de m grootste eigenwaarden te kiezen en deze op lengte 1 te normeren, dan hebben we een orthonormaalbasis voor de m-dimensionale deelruimte met de grootste spreiding van metingen en we kunnen ook meteen de orthogonale projectie van de hoger-dimensionale metingen in deze deelruimte aangeven.

Beste approximatie

We hebben gezien dat we de orthogonale projectie v_0 van een vector v in een deelruimte U met orthogonale basis $(v_1, \ldots, v_r)$ kunnen berekenen door

$$v_0 := \frac{\langle v, v_1\rangle}{\langle v_1, v_1\rangle} v_1 + \ldots + \frac{\langle v, v_r\rangle}{\langle v_r, v_r\rangle} v_r.$$

Het aardige aan deze formule is dat we op deze manier de *beste approximatie* van v in de deelruimte U vinden. Hiermee bedoelen we dat v_0 de vector in U is waarvoor het verschil $v - v_0$ de minimale lengte heeft.

C.35 Definitie We noemen een vector v_0 in een deelruimte U van V de *beste approximatie* van de vector $v \in V$ als de lengte van $v - v_0$ minimaal onder alle vectoren uit U is, d.w.z. als $\|v - v_0\| \leq \|v - u\|$ voor alle $u \in U$.

Meetkundig is het redelijk voor de hand liggend, dat de orthogonale projectie juist de beste approximatie aangeeft. Maar omdat we het inproduct in een algemener kader gedefinieerd hebben, is het handig dit even voor een willekeurig inproduct na te gaan:

We schrijven $v = v_0 + w$, dan is $\langle w, u \rangle = 0$ voor alle $u \in U$ omdat v_0 de orthogonale projectie van v in U is. Maar als we nu de lengte van het verschil van v met een willekeurige vector $u \in U$ berekenen, dan is

$$\begin{aligned}
\|v - u\|^2 &= \|(v - v_0) + (v_0 - u)\|^2 \\
&= \langle (v - v_0) + (v_0 - u), (v - v_0) + (v_0 - u) \rangle \\
&= \langle v - v_0, v - v_0 \rangle + \langle v_0 - u, v_0 - u \rangle + 2\langle v - v_0, v_0 - u \rangle \\
&= \|v - v_0\|^2 + \|v_0 - u\|^2 + 2\langle w, v_0 - u \rangle \\
&= \|v - v_0\|^2 + \|v_0 - u\|^2.
\end{aligned}$$

Maar het laatste is minimaal als $\|v_0 - u\| = 0$ is, dus als $u = v_0$ is.

15.4 Toepassing: Zoekmachines

Een eenvoudige maar wel redelijk efficiënte manier om een zoekmachine in elkaar te zetten is gebaseerd op een variatie van het standaardinproduct. We nemen aan dat we een aantal documenten $D_1, D_2, \ldots, D_n$ hebben en dat deze de woorden $t_1, t_2, \ldots, t_m$ (t voor *term*) bevatten. We beschrijven een document of een aanvraag Q (Q voor *query*) door een vector

$$f(Q) = (f_1, f_2, \ldots, f_m)$$

waarbij f_i aangeeft hoe vaak het woord t_i in Q voorkomt.

We zeggen nu dat een aanvraag Q sterk op een document D_i lijkt als de hoek tussen de vectoren $f(Q)$ en $f(D_i)$ klein is, dus als het inproduct $\langle f(Q), f(D_i) \rangle$ tussen de vectoren groot is.

Het inproduct hangt hierbij van zekere *gewichten* af, die rekening ermee houden dat sommige woorden belangrijker zijn dan andere: Een woord dat in alle documenten voorkomt geeft geen informatie over de overeenkomst met een aanvraag, terwijl een woord dat alleen in één document voorkomt een heel sterke indicatie is. Daarom wordt het inproduct $\langle f(Q), f(D_i) \rangle$ niet met het standaardinproduct berekend maar door

$$\langle f(Q), f(D_i) \rangle := \sum_{j=1}^{m} w_j f(Q)_j f(D_i)_j.$$

Er zijn verschillende manieren om de gewichten w_j te bepalen, vaak wordt de *inverse document frequency* gekozen, die gedefinieerd is door

$$w_j := \log(\frac{n}{n_j})$$

waarbij n_j het aantal documenten in de collectie is, dat het woord t_j bevat.

Het maat, waarmee een aanvraag Q op een document D_i lijkt is dus

$$l(Q, D_i) := \frac{\langle f(Q), f(D_i) \rangle}{\|f(Q)\| \|f(D_i)\|}$$

waarbij de lengte $\|f(Q)\|$ weer als $\sqrt{\langle f(Q), f(Q) \rangle}$ gedefinieerd is. De documenten worden nu zo gerangschikt dat de waarde $l(Q, D_i)$ afneemt.

In de praktijk zijn er natuurlijk een aantal trucjes bij de implementatie van zo'n zoekmachine. Een van de belangrijkste is de *inverse index*, die voor een woord t_i zegt in welke documenten hij überhaupt voorkomt. Dan hoeft bij een aanvraag alleen maar naar de documenten te worden gekeken, waarin een van de woorden in de aanvraag voorkomt en dat zijn er meestal niet zo erg veel omdat zinvolle aanvragen alleen maar informatie dragende woorden bevatten. Een lijst van heel onbelangrijke woorden (zogenoemde *stopwords*) wordt meestal expliciet uit de lijst van woorden uitgesloten.

Merk op: De bekende zoekmachine GOOGLE werkt niet op deze manier. Het was juist een reden voor het grote succes van GOOGLE dat de volgorde van de matchende documenten hier anders wordt bepaald, zo dat de meest interessante documenten vaker op de eerste plaatsen terecht komen.

Het idee achter GOOGLE is in feite een redelijk simpel eigenvector probleem: De *relevantie* van een document wordt beschouwd als evenredig met de som van de relevanties van de documenten die erna verwijzen. De relevanties van de documenten zijn dus een eigenvector voor de matrix die de verwijzingen tussen de documenten bevat, en wel de eigenvector voor de grootste eigenwaarde (dan zijn alle relevanties positief). De matchende documenten voor een query worden nu volgens hun op deze manier bepaalde relevanties gerangschikt.

BELANGRIJKE BEGRIPPEN IN DEZE LES

- standaardinproduct
- bilineaire afbeelding
- inproduct
- orthogonaliteit, Gram-Schmidt orthogonalisatie
- orthogonale projectie
- beste approximatie

Opgaven

113. Bepaal in een gewone kubus de volgende hoeken tussen vectoren die van het middelpunt van de kubus uit gaan:

    (i) Tussen de vectoren naar twee hoekpunten die door een ribbe verbonden zijn.

    (ii) Tussen de vectoren naar een hoekpunt en het middelpunt van een zijvlak die het hoekpunt bevat.

    (iii) Tussen de vectoren naar de middelpunten van twee ribben die zich op een zijvlak tegenover liggen.

114. Bepaal een basis van de deelruimten die (met betrekking tot het standaardinproduct) orthogonaal op de volgende vectoren staan:

$$\text{(i)} \begin{pmatrix} 1 \\ -2 \\ 3 \\ 4 \end{pmatrix} \text{ en } \begin{pmatrix} 3 \\ -5 \\ 7 \\ 8 \end{pmatrix} \in \mathbb{R}^4 \qquad \text{(ii)} \begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \\ 2 \end{pmatrix} \text{ en } \begin{pmatrix} 2 \\ 4 \\ 7 \\ 2 \\ -1 \end{pmatrix} \in \mathbb{R}^5$$

115. Laat zien dat de afbeelding $f(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}) := x_1x_2 - x_1y_2 - y_1x_2 + 3y_1y_2$ een inproduct op $\mathbb{R}^2$ definieert (d.w.z. laat zien dat f een positief definiete, symmetrische bilineaire afbeelding is).

116. Laten $v_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ en $v_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ vectoren in $\mathbb{R}^2$ zijn.

    (i) Voor welke $a \in \mathbb{R}$ definieert de afbeelding

$$f(v_1, v_2) := x_1x_2 - 3x_1y_2 - 3y_1x_2 + ay_1y_2$$

    een inproduct op $\mathbb{R}^2$?

    (ii) Voor welke $a, b, c, d \in \mathbb{R}$ definieert de afbeelding

$$f(v_1, v_2) := ax_1x_2 + bx_1y_2 + cy_1x_2 + dy_1y_2$$

    een inproduct op $\mathbb{R}^2$?

117. Op de 2×2-matrices definiëren we een afbeelding $\Phi(A, B) := sp(A \cdot B^{tr})$. Hierbij betekent $sp(X) := X_{11} + X_{22}$ de *spoor* van de matrix X, dus de som van de elementen op de diagonaal.

    (i) Laat zien dat Φ een bilineaire afbeelding is.

    (ii) Ga na dat Φ een inproduct is.

    (iii) Bepaal (met behulp van de Gram-Schmidt orthogonalisatie) een orthonormaalbasis van $\mathbb{R}^{2\times 2}$ met betrekking tot Φ.

118. Bepaal een orthonormaalbasis (met betrekking tot het standaardinproduct) voor de deelvectorruimte $U \subseteq \mathbb{R}^4$ die de vectoren $\begin{pmatrix} -3 \\ -3 \\ 3 \\ 3 \end{pmatrix}, \begin{pmatrix} -5 \\ -5 \\ 7 \\ 7 \end{pmatrix}$ en $\begin{pmatrix} 4 \\ -2 \\ 0 \\ 6 \end{pmatrix}$ bevat.

119. Geef de orthogonale projectie (met betrekking tot het standaardinproduct) van $v = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ in de richting van $w = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ aan.

120. Bereken de orthogonale projectie (met betrekking tot het standaardinproduct) van de vector $v = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$ in het vlak dat de vectoren $v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ en $v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ bevat.

121. Zij $V = \mathbb{R}^4$ met standaardinproduct. De deelruimte $U \subseteq V$ zij opgespannen door de vectoren $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ en $v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. Bereken de beste approximatie van $v = \begin{pmatrix} 2 \\ 2 \\ 3 \\ 1 \end{pmatrix}$ in U, d.w.z. de vector $v_0 \in U$ zo dat $\|v - v_0\|$ minimaal is. (Let op dat (v_1, v_2, v_3) nog geen orthogonaal stelsel is.)