Inhoudsopgave

Voorwoord

Het is de bedoeling dat het onderdeel lineaire algebra van het vak ???? goed aansluit bij het dictaat lineaire algebra, deel 1. Net als in het dictaat van deel 1 zullen er in ook dit deel nog vele taal en wiskundige foutjes (zo niet fouten of blunders) staan. Daarom blijft ook voor dit dictaat de oproep uit deel 1 geldig aan de studenten. Ziet u fouten, geef deze dan aan mij door, opdat toekomstige studenten een beter dictaat verkrijgen.

De volgende onderwerpen zullen aan bod komen.

1. Ophalen kennis lineaire algebra via $\mathbb{C}^n$.
2. De algemene vectorruimten (reeel en complex) worden geïntroduceerd.
3. De lineaire structuur op algemene vectorruimten wordt bestudeerd, via:
   - Deelruimten, opspanning en basis.
   - Coördinatentransformaties en overgangsmatrix.
4. Lineaire afbeeldingen tussen vectorruimtes worden bestudeerd, via:
   - Lineaire afbeeldingen en transformatiematrix.
   - Eigenwaarden, eigenruimtes van een lineaire operator.

Bovenstaande zijn de onderwerpen van het eerste deeeltentamen (tesamen met de onderwerpen differentiaalvergelijkingen).

Het tweede deeltentamen betreft inwendigproductstructeren op vectorruimten (tesamen met de onderwerpen differentiaalvergelijkingen).

1. Het standaard inwendigproduct op $\mathbb{C}^n$ en op andere vectorruimten wordt geïntroduceerd.
2. Inwendig product ruimten worden geïntroduceerd (reeel en complex).
3. Basisrepresentatie inwendigproduct.
4. Projecteren op eindig dimensionale deelruimten.
5. De spectraalstellingen voor reële en complexe matrices.
6. Operatoren theorie, geadjungeerde van een operator, zelfgeadjungeerde operatoren, normale operatoren.

Hoofdstuk 1

Complexe lineaire algebra

1.1 De lineaire structuur van $\mathbb{C}^n$

We worden allen geacht vertrouwd te zijn met de complexe getallen en met de elementaire operaties op $\mathbb{C}$ zoals: optellen, aftrekken, vermenigvuldigen, delen en conjugatie (complex toegevoegde). Ook de polaire vorm (notatie) van een complex getal (d.w.z. $z = |z|e^{i\varphi}$) wordt u geacht te kennen en ermee te kunnen rekenen. Zo niet, bekijk dan de appendix die op blackboard staat over complexe getallen [1].

In deel 1 hebben we de complexe tegenhanger van $\mathbb{R}^n$ (die uit alle reële $n \times 1$ matrices bestaat) al ontmoet, namelijk $\mathbb{C}^n$. Deze bestaat uit alle complexe $n \times 1$ matrices.

We schrijven $\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{C}^n$. Op $\mathbb{C}^n$ definiëren we een optelling en een scalaire vermenigvuldiging (een lineaire structuur) via: als: $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$ en $c \in \mathbb{C}$, zeg:

$$\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \quad \text{en} \quad \mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \quad \text{dan} \quad \mathbf{z} + \mathbf{w} = \begin{bmatrix} z_1 + w_1 \\ \vdots \\ z_n + w_n \end{bmatrix} \quad \text{en} \quad c\mathbf{z} = \begin{bmatrix} cz_1 \\ \vdots \\ cz_n \end{bmatrix}$$

Merk op dat op $\mathbb{C}$ één operatie bestaat die we reëel niet kennen (of liever, die reëel weinig zinvols oplevert): de conjugatie $\overline{z}$ (als $z = a + bi$ dan $\overline{z} = a - bi$). We definiëren de conjugatie op $\mathbb{C}^n$ alsvolgt: als $\mathbf{z} \in \mathbb{C}^n$, zeg:

$$\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \quad \text{dan} \quad \overline{\mathbf{z}} = \begin{bmatrix} \overline{z_1} \\ \vdots \\ \overline{z_n} \end{bmatrix}$$

Met deze optelling en (complexe)scalaire vermenigvuldiging blijkt $\mathbb{C}^n$ een zogeheten (complexe) vectorruimte te zijn, waarmee we bedoelen dat aan de 8 rekenregels uit deel 1 voldaan is.

[1] Naar een handout van Dr. H.A.W.M. Kneppers

STELLING 1.1. *De optelling + voegt aan ieder tweetal* $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$ *en aan iedere scalair* $c \in \mathbb{C}$ *een vector* $\mathbf{z} + \mathbf{w} \in \mathbb{C}^n$ *en* $c\mathbf{z} \in \mathbb{C}^n$ *toe, die voldoet aan de volgende 8 rekenregels.*

1. $\mathbf{z} + \mathbf{w} = \mathbf{w} + \mathbf{z}$. *(*commutatieve eigenschap*)*
2. $(\mathbf{z} + \mathbf{w}) + \mathbf{v} = \mathbf{z} + (\mathbf{w} + \mathbf{v})$. *(*associatieve eigenschap*)*
3. *Er is een nulvector* $\mathbf{0}$ *met de eigenschap* $\mathbf{z} + \mathbf{0} = \mathbf{0} + \mathbf{z} = \mathbf{z}$.
4. *Voor elke* $\mathbf{z}$ *bestaat een tegengestelde* $-\mathbf{z}$ *zo dat* $\mathbf{z} + (-\mathbf{z}) = \mathbf{0}$.
5. $c(\mathbf{z} + \mathbf{w}) = c\mathbf{z} + c\mathbf{w}$. *(*distributieve eigenschap 1*)*
6. $(c + d)\mathbf{z} = c\mathbf{z} + d\mathbf{z}$. *(*distributieve eigenschap 2*)*
7. $c(d\mathbf{z}) = (cd)\mathbf{z}$.
8. $1\mathbf{z} = \mathbf{z}$.

De hele theorie zoals eerder behandeld is voor $\mathbb{R}^n$ kan herhaald worden voor $\mathbb{C}^n$ met letterlijk dezelfde stellingen en dezelfde bewijzen. Want aan puur symbolisch rekenen kan men niet zien of in $\mathbb{R}$ of in $\mathbb{C}$ gerekend wordt (zolang geen ongelijkheden worden gebruikt; de uitspraak $x^2 \geq 0$ is waar in $\mathbb{R}$, niet in $\mathbb{C}$). Bijvoorbeeld, het oplossen van complexe stelsels $[A|\mathbf{b}]$ (of $A\mathbf{z} = \mathbf{b}$) gaat hetzelfde als bij reële stelsels. Dit hebben we bij deel 1 al gezien en daar hebben we ook ingezien dat het oplossen van complexe stelsels, hoewel methodisch hetzelfde, rekentechnisch toch een extra inspanning vraagt. We geven een lijst met de opbouw van de complexe lineaire algebra.

- Als $(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k)$ een stelsel vectoren in $\mathbb{C}^n$ is en als $c_1, c_2, \ldots, c_k$ scalairen in $\mathbb{C}$ zijn, dan wordt de vector van de vorm
$$c_1\mathbf{z}_1 + c_2\mathbf{z}_2 + \cdots + c_k\mathbf{z}_k$$
een *lineaire combinatie* van het stelsel genoemd. Een lineaire combinatie wordt ook genoteerd met het matrixvectorproduct: $[\mathbf{z}_1 \cdots \mathbf{z}_k] \begin{bmatrix} c_1 \\ | \\ c_k \end{bmatrix}$.
De verzameling van *alle* lineaire combinaties van dit stelsel heet de "*(lineaire) opspanning* van $(\mathbf{z}_1, \ldots, \mathbf{z}_k)$") en wordt genoteerd met: $\mathrm{Span}\{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$.
- Als $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k$ een stelsel vectoren in $\mathbb{C}^n$ is dan zeggen we dat dit dit stelsel lineair onafhankelijk is als de eis: $c_1\mathbf{z}_1 + c_2\mathbf{z}_2 + \cdots + c_k\mathbf{z}_k$ noodzakelijk impliceert dat $c_1 = c_2 = \cdots = c_k = 0$. Als dit niet zo is dan heet dit een afhankelijk stel vectoren.
- Een deelverzameling $W \subset \mathbb{C}^n$ heet een lineaire deelruimte van $\mathbb{C}^n$ als:
  1. $\mathbf{0} \in W$,
  2. Als $\mathbf{z}_1, \mathbf{z_2} \in W$, dan ook $\mathbf{z}_1 + \mathbf{z_2} \in W$, voor alle $\mathbf{z}_1, \mathbf{z_2} \in W$,
  3. Als $\mathbf{z} \in W$ en $c \in \mathbb{C}$, dan ook $c\mathbf{z} \in W$, voor alle $\mathbf{z} \in W$ en $c \in \mathbb{C}$.
- De lineaire deelruimtes van $\mathbb{C}^n$ zijn precies die deelverzamelingen die te schrijven zijn als $\mathrm{Span}\{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$.
- Als W een deelruimte is van $\mathbb{C}^n$ dan heet de verzameling $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ een basis voor W als:
  1. $\mathrm{Span}\{\mathbf{b}_1, \ldots, \mathbf{b}_k\} = W$.

2. De verzameling $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ is lineair onafhankelijk.

- Stel dat W een deelruimte is van $\mathbb{C}^n$. Dan bestaat er een basis voor W en twee verschillende bases van W zullen hetzelfde aantal vectoren hebben. Dit aantal wordt de dimensie van W genoemd en wordt genoteerd met $\dim(W)$.

- We merken nog op dat de standaardeenheids vectoren vectoren $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ een basis voor $\mathbb{C}^n$ vormen, ook hier de standaardbasis geheten, waaruit volgt dat $\dim(\mathbb{C}^n) = n$. (Hoewel $\mathbb{C} = \mathbb{C}^1$ meetkundig wordt voorgesteld met een vlak, is de (algebraïsche !) dimensie van $\mathbb{C}^1$ dus 1.)

- Ook in de complexe lineaire algebra geldt de uitbreidingsstelling, de uitdunningsstelling en de dimensiestelling voor deelruimtes.

- Aan een complexe $m \times n$ matrix $A = [\mathbf{a}_1 \ldots \mathbf{a}_k]$ kunnen we de deelruimtes $\mathrm{Col}(A) = \mathrm{Span}\{\mathbf{a}_1, \ldots, \mathbf{a}_k\} \subset \mathbb{C}^m$ en $\mathrm{Nul}(A) = \{\mathbf{z} \in \mathbb{C}^n | A\mathbf{z} = \mathbf{0}\} \subset \mathbb{C}^n$ toevoegen. Op de ons bekende wijze is een basis te bepalen voor deze ruimtes. De pivotkolommen van A vormen een basis van $\mathrm{Col}(A)$, de richtingsvectoren van de oplossing van het stelsel $A\mathbf{z} = \mathbf{0}$ vormen een basis van $\mathrm{Nul}(A)$.

- Voor een complexe $m \times n$ matrix A geldt ook: $\dim(\mathrm{Col}(A)) + \dim(\mathrm{Nul}(A)) = n$ en $\dim(\mathrm{Col}(A)) = \dim(\mathrm{Col}(A^T))$. Dit laatste getal wordt ook de rang van A genoemd.

1.2 Complexe Matrixalgebra

De complexe matrixalgebra is het evenbeeld van de reële matrixalgebra. Letterlijk dezelfde definitie van de som en het product van twee matrices geldt. De inverse van een matrix en de inverse matrix stelling kan letterlijk complex gebruikt worden. Echter, we hebben complex een extra operatie die we kunnen overhevelen naar de matrices, de operatie conjugeren.
Dit levert bij de complexe matrixalgebra de nieuwe operatie de conjugatie van een matrix. $\overline{A}$ is de matrix A met geconjugeerde kentallen, dus $\left(\overline{A}\right)_{i,j} = \overline{a_{i,j}}$.

Voorbeeld 1.2. Als $A = \begin{bmatrix} 1 & 1-i \\ i & 5+2i \\ 4-5i & 4+5i \end{bmatrix}$ dan $\overline{A} = \begin{bmatrix} 1 & 1+i \\ -i & 5-2i \\ 4+5i & 4-5i \end{bmatrix}$.

Bij het bewijs van het volgende lemma zijn de rekenregels voor voor de geconjugeerde $\overline{z}$ nodig. (Dit zijn: $\overline{\overline{z}} = z$, $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$ en $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$.)

LEMMA 1.3. *Laat A, B matrices zijn. Dan geldt:*

1. $\overline{A+B} = \overline{A} + \overline{B}$,

2. $\overline{AB} = \overline{A}\,\overline{B}$,

3. *A is een reële matrix als en slechts als $\overline{A} = A$.*

Het transponeren van complexe matrices gaat zonder enig probleem. Echter, transponeren blijkt voor de complexe algebra een minder belangrijke operatie te zijn. Wat daarvoor in de plaats komt is het zogeheten "Hermitisch[2] transponeren".

[2]Charles Hermite 1822-1901. Hij had moeite met zijn studie en heeft 5 jaar over zijn kandidaatsexamen gedaan. Hij heeft een belangrijke bijdrage geleverd op het gebied van de elliptische functies. Verder heeft hij als eerste het transcendent zijn van e aangetoond.

Definitie: Laat A een (complexe) $m \times n$-matrix zijn. De **Hermites getransponeerde** of **geconjugeerd getransponeerde** A^* van A is de matrix die uit A verkregen wordt door A te transponeren en van elk kental de complex geconjugeerde te nemen, dus

$$A^* = (\overline{A})^T = \overline{(A^T)}.$$

Voorbeeld 1.4. De Hermites getransponeerde van de matrix

$$A = \begin{bmatrix} 1 & 2+i & 0 \\ 3 & i & 1-i \end{bmatrix}$$

wordt gegeven door de matrix

$$A^* = \begin{bmatrix} 1 & 3 \\ 2-i & -i \\ 0 & 1+i \end{bmatrix}.$$

De volgende stelling bevat de rekenregels voor Hermites transponeren. Deze rekenregels (eigenschappen) zijn niet moeilijk te bewijzen.

STELLING 1.5 (Eigenschappen van het Hermites transponeren)**.** *Als A en B twee (complexe) matrices zijn, dan geldt*

1. $(A^*)^* = A$,
2. $(A+B)^* = A^* + B^*$,
3. $(cA)^* = \overline{c}A^*$ *voor ieder complex getal* c,
4. $(AB)^* = B^*A^*$.
5. *A is een reële matrix als en slechts als* $A^* = A^T$.

Met behulp van transponeren zijn in deel 1 op diverse plaatsen speciale typen reële matrices gedefinieerd. We herhalen deze definities nog even.

Definitie: Laat A een reële matrix zijn. A heet

- **symmetrisch** als $A^T = A$,
- **scheefsymmetrisch** als $A^T = -A$.

Afspraak Als we praten over een symmetrische of scheefsymmetrische matrix A, dan zal daarmee zonder verdere uitleg aangenomen worden dat de matrix A een reële matrix is!
Vervolgens gaan we de (scheef)symmetrische matrices generaliseren met behulp van de definitie van de Hermites getransponeerde.

Definitie: Laat A een matrix zijn. A heet

- **Hermites** als $A^* = A$,
- **scheefhermites** als $A^* = -A$.

Opmerking: We sommen een aantal eigenschappen op:

- A is Hermites als en slechts als $a_{ij} = \overline{a}_{ji}$, (voor alle i, j).
- Als A Hermites is dan staan op de hoofddiagonaal reële getallen.
- A is scheefhermites als en slechts als $a_{ij} = -\overline{a}_{ji}$, (voor alle i, j).
- Als A scheefhermites is dan staan op de hoofddiagonaal zuiver imaginaire getallen.
- A is scheefhermites als en slechts als iA Hermites is.
- Er geldt: een (noodzakelijk reële) symmetrische matrix is Hermites en een (noodzakelijk reële) scheefsymmetrische matrix is scheefhermites.

Voorbeeld 1.6. Bekijk de matrices

$$A = \begin{bmatrix} 3 & 1+i & -2 \\ 1-i & -5 & i \\ -2 & -i & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} i & 1-i \\ -1-i & 2i \end{bmatrix} \qquad D = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1+i \\ 1-i & -1 \end{bmatrix} \qquad E = \begin{bmatrix} i & i \\ i & i \end{bmatrix}.$$

Er geldt: A is Hermites en niet scheefsymmetrisch. B is zowel scheefhermites als scheefsymmetrisch, maar niet symmetrisch of Hermites. C is scheefhermites, maar niet Hermites. D is Hermites, maar niet scheefhermites. E is scheefhermites, niet Hermites, niet symmetrisch, niet scheefsymmetrisch.

1.3 Complexe eigenwaarden en eigenvectoren

We willen nog even speciale aandacht vestigen op de theorie van eigenwaarden, eigenvectoren en diagonaliseerbaarheid. Deel 1 heeft zich beperkt tot de (reële) eigenwaarden van reële matrices. Als we dit complex gaan doen moeten we eerst de definities geven en/of aanpassen in het geval ook complexe eigenwaarden worden toegelaten.

Definitie: Een **eigenwaarde** van een vierkante (complexe) $n \times n$ matrix A is een scalar $\lambda \in \mathbb{C}$ waarvoor geldt dat de vergelijking $A\mathbf{z} = \lambda\mathbf{z}$ een niet-triviale oplossing heeft (met andere woorden, er bestaat een niet-nulvector $\mathbf{z} \in \mathbb{C}^n$ met de eigenschap dat $A\mathbf{z} = \lambda\mathbf{z}$).
Zo'n niet-triviale oplossing $\mathbf{z}$ heet een bij λ behorende **eigenvector** van A.
Als λ een eigenwaarde van A is, dan heet de verzameling oplossingen van $A\mathbf{z} = \lambda\mathbf{z}$ met $\mathbf{z} \in \mathbb{C}^n$ de **eigenruimte** van A bij eigenwaarde λ, en wordt genoteerd met E_λ.
De verzameling eigenwaarden heet het spectrum $SP(A)$ van A.

Voor een gegeven $n \times n$ matrix A hebben we weer:

- Het karakteristiek polynoom $p_A(\lambda)$ (in de variabele λ) van A is het n^{de}–graads polynoom $p_A(\lambda) = \det(A - \lambda I)$.
- De eigenwaarden van de matrix A zijn precies de oplossingen van de karakteristieke vergelijking $p_A(\lambda) = \det(A - \lambda I) = 0$.
- De algebraïsche multipliciteit van een eigenwaarde λ_0 van A is de multipliciteit van λ_0 als nulpunt van het karakteristiek polynoom. We schrijven hiervoor: $a.m.(\lambda_0) = k$.
- De sommatie van de algebraïsche multipliciteiten van de verschillende eigenwaarden van A is dus de graad van het polynoom, is dus n. Dus:
$$\sum_{\lambda \in SP(A)} a.m.(\lambda) = n$$
- Als λ_0 een eigenwaarde is van A dan geldt:
$$E_{\lambda_0} = \text{Nul}(A - \lambda_0 I)$$
- De dimensie van de eigenruimte E_{λ_0} heet de meetkundige multipliciteit van A. We schrijven hiervoor: $m.m.(\lambda_0) = \dim(E_{\lambda_0})$. (In deel 1 is het bewijs hiervan niet gegeven, en het bewijs is gewoon lastig.)
- Altijd geldt: als λ_0 een eigenwaarde van A dan $1 \leq m.m.(\lambda_0) \leq a.m.(\lambda_0)$.
- Daarom zal altijd gelden:
$$\sum_{\lambda \in SP(A)} m.m.(\lambda) \leq n$$

Voorbeeld 1.7. Beschouw de complexe matrix: $A = \begin{bmatrix} 0 & 0 & -1 \\ 0 & i & 0 \\ 1 & 0 & 0 \end{bmatrix}$. We doorlopen voor deze matrix bovenstaande stappen. We bepalen eerst het karakteristiek polynoom.
$p_A(\lambda) = \begin{vmatrix} -\lambda & 0 & -1 \\ 0 & i-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix} = (i-\lambda) \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = -(\lambda - i)(\lambda^2+1) = -(\lambda-i)^2(\lambda+i)$. We zien: $\lambda = i$ is eigenwaarde van A met algebraïsche multipliciteit $a.m.(\lambda = i) = 2$ en $\lambda = -i$ is een eigenwaarde van A met algebraïsche multipliciteit $a.m.(\lambda = -i) = 1$.
Inderdaad, de som van de algebraïsche multipliciteiten is 3.
We bepalen E_i en E_{-i} door een basis van beide ruimtes te bepalen.
$E_i = \text{Nul}(A - iI)$ levert dat de vectoren in E_i precies de oplossing is van het stelsel:
$\left[\begin{array}{ccc|c} -i & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & -i & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$. Dit geeft $\mathbf{z} = \begin{bmatrix} iz_3 \\ z_2 \\ z_3 \end{bmatrix} = z_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + z_3 \begin{bmatrix} i \\ 0 \\ 1 \end{bmatrix}$.
Dus: $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} i \\ 0 \\ 1 \end{bmatrix} \right\}$ is een basis E_i en we zien $m.m.(\lambda = i) = 2$.
Evenzo: $E_{-i} = \text{Nul}(A + iI)$ levert dat de vectoren in E_{-i} precies de oplossing is van het stelsel:
$\left[\begin{array}{ccc|c} i & 0 & -1 & 0 \\ 0 & 2i & 0 & 0 \\ 1 & 0 & i & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & i & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$. Dit geeft $\mathbf{z} = \begin{bmatrix} -iz_3 \\ 0 \\ z_3 \end{bmatrix} = z_3 \begin{bmatrix} -i \\ 0 \\ 1 \end{bmatrix}$. Dus: $\left\{ \begin{bmatrix} -i \\ 0 \\ 1 \end{bmatrix} \right\}$ is een basis E_{-i} en we zien $m.m.(\lambda = -i) = 1$.

We vragen ons af of de matrix uit voorafgaand voorbeeld diagonaliseerbaar is. We herinneren ons:

Definitie: Een (eventueel complexe) $n \times n$-matrix A heet **(complex) diagonaliseerbaar** als er een (eventueel complexe) inverteerbare matrix P en een (eventueel complexe) diagonaalmatrix D bestaan zo dat $A = PDP^{-1}$.

De stellingen uit deel 1 (zie college 13) wat betreft het diagonaliseren van reële matrices hebben alle een vertaling naar de nieuwe situtie van het diagonaliseren van complexe matrices. We vermelden expliciet:

LEMMA 1.8. *Stel dat A een complexe $n \times n$ matrix is.*
Als $P = [\,\mathbf{p}_1 \ldots \mathbf{p}_n\,]$ een (niet noodzakelijk inverteerbare) $n \times n$ matrix is en $D = diag(d_1, \ldots, d_n)$ een diagonaalmatrix, dan zijn de volgende uitspraken equivalent:

1. *$AP = PD$,*
2. *Ieder diagonaalkental d_i van de matrix D heeft de eigenschap: $A\mathbf{p}_i = d_i\mathbf{p}_i$ $(i = 1, \ldots, n)$ en als $\mathbf{p}_i \neq \mathbf{0}$ dan is d_i dus een eigenwaarde van A.*

En dus:

STELLING 1.9. *Zij A een (eventueel complexe) $n \times n$ matrix. Dan geldt: A is complex diagonaliseerbaar als en slechts als A n (eventueel complexe) lineair onafhankelijke eigenvectoren heeft.*

Weer geldt:

STELLING 1.10. *Zij A een (eventueel complexe) $n \times n$ matrix. Dan geldt: eigenvectoren bij verschillende (eventueel complexe) eigenwaarden zijn lineair onafhankelijk.*

En dus:

STELLING 1.11. *Zij A een (eventueel complexe) $n \times n$ matrix. Dan geldt: als A n verschillende eigenwaarden heeft dan is A complex diagonaliseerbaar.*

Weer geldt:

STELLING 1.12. *Zij A een (eventueel complexe) $n \times n$ matrix. Er geldt: A is diagonaliseerbaar als en slechts als voor iedere eigenwaarde van A geldt dat de meetkundige multipliciteit gelijk is aan de algebraïsche multipliciteit. Dit is equivalent met de uitspraak:*

$$\sum_{\lambda \in SP(A)} m.m.(\lambda) = n$$

Voorbeeld 1.13. We concluderen dat de matrix A uit voorbeeld 1.7 diagonaliseerbaar is. Inderdaad, daar is de som van de meetkundige multipliciteiten gelijk aan $2 + 1 = 3$ en A is een 3×3 matrix.

Een diagonalisering is te vinden op de gebruikelijke manier: $A = PDP^{-1}$ met $P = \begin{bmatrix} 0 & i & -i \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ (de eigenvectoren in P) en $D = \begin{bmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix}$ (de overeenkomstige eigenwaarden in de diagonaalmatrix D).

Een diagonalisering is handig om te hebben, bijvoorbeeld om A^n te bepalen. Hier zien we: $A^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1}$ en omdat $D^2 = \begin{bmatrix} (i)^2 & 0 & 0 \\ 0 & (i)^2 & 0 \\ 0 & 0 & (-i)^2 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -I_3$ volgt: $A^2 = PD^2P^{-1} = P(-I)P^{-1} = -PP^{-1} = -I$.

Wat de terminologie van deel 1 over diagonaliseren van reële matrices moeten we iets voorzichtiger zijn. We geven eerst een voorbeeld:

Voorbeeld 1.14. Bekijk de matrix

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

De eigenwaarden van deze matrix vinden we uit de vergelijking

$$\det(A - \lambda I) = (1 - \lambda)^2(\lambda^2 + 1) = 0.$$

In deel 1 zouden we zeggen: $\lambda = 1$ is de enige eigenwaarde van A met meetkundige multipliciteit 2 (ga na) en dus is de matrix A niet diagonaliseerbaar.
Nu kunnen we zeggen: de eigenwaarden zijn dus i, $-i$ en 1 (met algebraïsche multipliciteit 2).
Ga na dat de vectoren

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{en} \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

een basis vormen voor de eigenruimte bij eigenwaarde 1. De vectoren

$$\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ i \end{bmatrix} \quad \text{en} \quad \mathbf{u}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -i \end{bmatrix}$$

vormen bases voor de eigenruimten bij respectievelijk de eigenwaarden $-i$ en i. Als we nu definieren

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & i & -i \end{bmatrix} \quad \text{en} \quad D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{bmatrix}$$

dan geldt $A = PDP^{-1}$ en de matrix A is opeens wel diagonaliseerbaar geworden, maar complexe getallen zijn nodig.

Kennelijk moeten we enige voorzichtigheid betrachten met wat we bedoelen “A is diagonaliseerbaar”, als de matrix A reëel is. Daarom:

Definitie: Laat A een reële $n \times n$ matrix zijn.
De matrix A heet **reëel diagonaliseerbaar** als er een reële inverteerbare matrix P en een reële diagonaalmatrix D bestaan zo dat $A = PDP^{-1}$.
De matrix A heet **complex diagonaliseerbaar** als er een (eventueel complexe) inverteerbare matrix P en een (eventueel complexe) diagonaalmatrix D bestaan zo dat $A = PDP^{-1}$.

Opmerking: Een reëel diagonaliseerbare matrix is ook complex diagonaliseerbaar.

De reële matrix A uit bovenstaand voorbeeld 1.14 is niet reëel diagonaliseerbaar, maar wel complex diagonaliseerbaar.

Opmerking: Een reële matrix A kan om *twee redenen* niet reëel-diagonaliseerbaar zijn, namelijk: A heeft een niet-reële eigenwaarde of er bestaat een eigenwaarde λ waarvan de dimensie van de eigenruimte kleiner is dan de algebraïsche multipliciteit van λ.
Echter, een matrix A kan slechts om één reden niet complex-diagonaliseerbaar zijn, namelijk: er bestaat een eigenwaarde λ waarvan de dimensie van de eigenruimte kleiner is dan de algebraïsche multipliciteit van λ.
Herinner: Een matrix die een eigenwaarde λ heeft waarvan de meetkundige multipliciteit strikt kleiner is dan de algebraïsche multipliciteit, heet **defect** (in λ).

Hoofdstuk 2

De lineaire structuur op een vectorruimte

2.1 Vectorruimten

Tijdens voorafgaande colleges in de lineaire algebra zijn de vectorruimten $\mathbb{R}^n$ en de $\mathbb{C}^n$ geïntroduceerd tezamen met hun deelruimten. De denkwereld van de *lineaire structuur* die daar gecreëerd is, blijkt ook in andere "ruimten" dan alleen $\mathbb{R}^n$ of $\mathbb{C}^n$ toepasbaar te zijn.

Bij het rekenen en redeneren in de lineaire structuur van $\mathbb{R}^n$ of $\mathbb{C}^n$ spelen de rekenoperaties optelling en scalaire vermenigvuldiging (vermenigvuldiging met een reëel of complex getal) een rol. Waarom zou dit alleen met kolommen kunnen? Inderdaad, er zijn veel meer objecten waarvoor een optelling en een scalaire vermenigvuldiging gedefinieërd zijn: matrices van gelijke afmeting, getalrijen (eindig of oneindig lang) en reëel- of complexwaardige functies (zie ook de voorbeelden verderop). Het blijkt dat allerlei theorie, die reeds ontwikkeld is voor $\mathbb{R}^n$ of $\mathbb{C}^n$, ook geldt in deze gevallen.

Om iets concreter te worden: bij differentiaalvergelijkingen bestaat de oplossing vaak uit lineaire combinaties van bepaalde basisoplossingen. Als we bijvoorbeeld kijken naar de vergelijking $y''+y = 0$, dan vinden we als oplossing $y = A\cos t + B\sin t$. Hier zijn $\cos t$ en $\sin t$ de basisoplossingen en A en B zijn vrij te kiezen coëfficiënten. Naar analogie van de lineaire combinatie in $\mathbb{R}^n$ zouden we deze basisoplossingen met hun lineaire combinaties ook graag als vectoren willen zien en praten over $\text{Span}\{\cos t, \sin t\}$.

Om precies vast te leggen in welke situaties we kunnen spreken over vectoren, wordt het begrip vectorruimte geïntroduceerd. De eerste die in een boek een *axiomatische* definitie gaf van een vectorruimte (wat ook wel lineaire ruimte genoemd wordt) was Peano.[1] Dit boek werd gepubliceerd in Turijn in 1888. Er is in de definitie sprake van twee verzamelingen. De vectorruimte is een verzameling en de elementen hieruit zullen **vectoren** heten, De vectoren worden in dit dictaat weer aangeduid met een dik gedrukte letter. De tweede verzameling zijn elementen uit een zogeheten **scalairen–lichaam**, waarvan de elementen **scalairen** [2] worden genoemd, en de verzameling $\mathbb{L}$. Meestal zal dit $\mathbb{R}$ of $\mathbb{C}$ zijn. In de informatica is dit vaak $\{0, 1\}$.

[1] Giuseppe Peano (1858-1932). In het boek "Calcolo geometrico secondo l'Ausdehnungslehre di H. Grassmann preceduto dalle operazioni della logica deduttiva", waarin hij de definitie van een vectorruimte gaf, introduceerde hij tevens de notaties $\cap$, $\cup$ en $\in$ voor doorsnede, vereniging en 'is element van'.

[2] Het uit het Frans overgenomen woord scalair (wordt ook op zijn Duits skalar genoemd) is verwant met het Nederlandse woord (peil)schaal. Een scalair is dan ook een schalingsfactor.

Definitie: Een **reële vectorruimte** (of een **reële lineaire ruimte**) is een *niet-lege* verzameling V waarop twee operaties gedefinieërdd zijn:

- Een *optelling* die bij elk tweetal $\mathbf{v}, \mathbf{w} \in V$ een element $\mathbf{v} + \mathbf{w} \in V$ maakt;
- Een *scalaire vermenigvuldiging* die bij elk tweetal $c \in \mathbb{R}$ en $\mathbf{v} \in V$ een element $c\mathbf{v} \in V$ maakt.

Deze optelling en scalaire vermenigvuldiging moeten voldoen aan de volgende acht axioma's (eisen), die moeten gelden voor iedere $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in V en voor iedere c, d in $\mathbb{R}$:

(1) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (*commutatieve eigenschap*)

(2) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (*associatieve eigenschap*)

(3) Er is een element $\mathbf{e}$ in V met de eigenschap dat $\mathbf{e} + \mathbf{v} = \mathbf{v}$ voor alle $\mathbf{v} \in V$

(4) Bij ieder element $\mathbf{v} \in V$ bestaat een element $\mathbf{w}$ in V zo dat $\mathbf{v} + \mathbf{w} = \mathbf{e}$

(5) $c(d\mathbf{v}) = (cd)\mathbf{v}$

(6) $1\mathbf{v} = \mathbf{v}$

(7) $(c + d)\mathbf{v} = c\mathbf{v} + d\mathbf{v}$ (*distributieve eigenschap 1*)

(8) $c(\mathbf{v} + \mathbf{w}) = c\mathbf{v} + c\mathbf{w}$ (*distributieve eigenschap 2*)

Een **complexe vectorruimte** (of een **complexe lineaire ruimte**) wordt op dezelfde manier gedefinieërd, met dit verschil dat de scalairen nu uit $\mathbb{C}$ gekozen worden. In bovenstaande moet elke $\mathbb{R}$ dus door $\mathbb{C}$ vervangen worden.
Een reële vectorruimte wordt ook wel een **vectorruimte over** $\mathbb{R}$ genoemd en een complexe vectorruimte een **vectorruimte over** $\mathbb{C}$.
Elementen van een vectorruimte worden **vectoren** genoemd. De vector $\mathbf{e}$ uit de derde rekenregel wordt **de nulvector** van V genoemd en genoteerd met $\mathbf{0}$. De vector $\mathbf{w}$ uit de vierde rekenregel wordt de **tegengestelde** van $\mathbf{v}$ genoemd en genoteerd met $-\mathbf{v}$.

Opmerking: (Voor hen die een theoretische interesse hebben.) Het is belangrijk dat bij de optelling het resultaat $(\mathbf{v} + \mathbf{w})$ weer in de verzameling V ligt. Men zegt wel: V is **gesloten** ten opzichte van de optelling. Dit moet je controleren. Hetzelfde geldt voor de scalaire vermenigvuldiging: V moet ook gesloten zijn t.o.v. de scalaire vermenigvuldiging, d.w.z het resultaat $c\mathbf{v}$ moet weer tot V behoren (voor alle c uit $\mathbb{R}$ of $\mathbb{C}$ en $\mathbf{v}$ uit V).
Om na te gaan of een gegeven verzameling met twee operaties een vectorruimte is, moeten uiteindelijk elf zaken worden nagegaan. Voordat de acht "rekenaxioma's" gecontroleerd worden, moet je namelijk laten zien dat V niet leeg is en dat V gesloten is onder de optelling en de scalaire vermenigvuldiging.
Tenslotte merken we nog op dat we inderdaad mogen praten over *de* nulvector en *de* tegengestelde van een vector, omdat deze uniek zijn. Zie Stelling 2.1.

Rekenregels

Wij hebben in dit dictaat gekozen voor scalairen uit $\mathbb{R}$ of $\mathbb{C}$. Afhankelijk van de toepassing zijn ook andere keuzen mogelijk. Zo kiest men in de digitale wereld (informatica, coderingstheorie,

cryptografie) de scalairen vaak gelijk aan 0 of 1 en wordt in plaats van $\mathbb{R}$ of $\mathbb{C}$ dus $\{0,1\}$ genomen. Niet elke keuze van scalairen levert een vectorruimte op. De eis die op de verzameling van scalairen gelegd moet worden is dat het een zogenaamd lichaam moet zijn. Dat is een verzameling waarop een "nette" optelling en vermenigvuldiging zijn gedefinieërd. Wat precies een scalairen–lichaam is zullen we niet zeggen. Wel zeggen we dat zo'n lichaam tenminste een 0 en een 1 moet bevatten, en aan de eigenschap: "voor iedere $c \neq 0$ moet een c' in het lichaam bestaan zo dat $cc' = 1$" moet voldoen. De verzameling gehele getallen $\mathbb{Z}$ is dan ook **GEEN** lichaam. Er bestaan dus geen vectorruimtes over $\mathbb{Z}$. Wij beperken ons in dit dictaat tot $\mathbb{R}$ of $\mathbb{C}$. Deze keuze is in het algemeen geschikt voor natuurkundige en elektrotechnische toepassingen.
In het vervolg zullen we regelmatig spreken over **vectorruimten over** $\mathbb{L}$ waarbij voor $\mathbb{L}$ in principe elk lichaam ingevuld kan worden. Voor ons volstaat het dan om slechts te denken aan $\mathbb{R}$ of $\mathbb{C}$. De volgende stelling, waarin enkele rekenregels geformuleerd worden die in elke vectorruimte gelden, is daar een voorbeeld van.

STELLING 2.1. *Zij V een vectorruimte over $\mathbb{L}$. Dan gelden de volgende regels voor willekeurige $\mathbf{v} \in V$ en $c \in \mathbb{L}$.*

(a) De nulvector is uniek.

(b) Elke vector heeft precies één tegengestelde

(c) Voor elke vector $\mathbf{v}$ uit V geldt $-(-\mathbf{v}) = \mathbf{v}$

(d) Voor elk drietal vectoren $\mathbf{v}$, $\mathbf{x}$ en $\mathbf{y}$ uit V geldt: uit $\mathbf{v} + \mathbf{x} = \mathbf{v} + \mathbf{y}$ volgt $\mathbf{x} = \mathbf{y}$

(e) $0\mathbf{v} = \mathbf{0}$,

(f) $c\mathbf{0} = \mathbf{0}$,

(g) $(-1)\mathbf{v} = -\mathbf{v}$,

(h) $c\mathbf{v} = \mathbf{0} \Longrightarrow c = 0$ of $\mathbf{v} = \mathbf{0}$.

Bewijs: Omdat we willen aantonen dat bovenstaande regels in *elke* vectorruimte gelden, zullen we ons zeer strak aan de definiërende rekenregels (1) t/m (8) van vectorruimten moeten houden. Daarom wordt bij elke stap vermeld waarom deze goed is.
(a): Als $\mathbf{0}$ en $\mathbf{0}'$ nulvectoren zijn, dan volgt: $\mathbf{0}' \overset{(3)}{=} \mathbf{0} + \mathbf{0}' \overset{(1)}{=} \mathbf{0}' + \mathbf{0} \overset{(3)}{=} \mathbf{0}$. Bij het eerste =-teken hebben we (3) gebruikt voor $\mathbf{0}$ en bij het derde =-teken voor $\mathbf{0}'$.
(b): Stel dat $\mathbf{b}$ en $\mathbf{c}$ tegengestelden zijn van $\mathbf{a}$, dus $\mathbf{a} + \mathbf{b} = \mathbf{0}$ en $\mathbf{a} + \mathbf{c} = \mathbf{0}$. Dan volgt:

$$\mathbf{b} \overset{(3)}{=} \mathbf{0} + \mathbf{b} = (\mathbf{a} + \mathbf{c}) + \mathbf{b} \overset{(1)}{=} (\mathbf{c} + \mathbf{a}) + \mathbf{b} \overset{(2)}{=} \mathbf{c} + (\mathbf{a} + \mathbf{b}) = \mathbf{c} + \mathbf{0} \overset{(1)}{=} \mathbf{0} + \mathbf{c} \overset{(3)}{=} \mathbf{c}.$$

(c): Er geldt $(-\mathbf{v}) + \mathbf{v} \overset{(1)}{=} \mathbf{v} + (-\mathbf{v}) \overset{(4)}{=} \mathbf{0}$. Hiermee volgt uit (4) en (b) dat $\mathbf{v}$ de tegengestelde is van $-\mathbf{v}$. Anders gezegd: $\mathbf{v} = -(-\mathbf{v})$.
(d): Als geldt $\mathbf{v} + \mathbf{x} = \mathbf{v} + \mathbf{y}$, dan tellen we aan beide zijden van de vergelijking $(-\mathbf{v})$ op. Links komt er dan:

$$(-\mathbf{v}) + (\mathbf{v} + \mathbf{x}) \overset{(2)}{=} (-\mathbf{v} + \mathbf{v}) + \mathbf{x} \overset{(1)}{=} (\mathbf{v} + (-\mathbf{v})) + \mathbf{x} \overset{(4)}{=} \mathbf{0} + \mathbf{x} \overset{(3)}{=} \mathbf{x},$$

en net zo komt er rechts: $(-\mathbf{v}) + (\mathbf{v} + \mathbf{y}) = \mathbf{y}$. Dus volgt nu:

$$\mathbf{x} = (-\mathbf{v}) + (\mathbf{v} + \mathbf{x}) = (-\mathbf{v}) + (\mathbf{v} + \mathbf{y}) = \mathbf{y}.$$

(e): $0\mathbf{v} = (0+0)\mathbf{v} \overset{(7)}{=} 0\mathbf{v} + 0\mathbf{v}$. Tel nu aan beide zijden $(-0\mathbf{v})$ op. Links komt er $0\mathbf{v} + (-0\mathbf{v}) \overset{(4)}{=} \mathbf{0}$ en rechts: $(0\mathbf{v} + 0\mathbf{v}) + (-0\mathbf{v}) \overset{(2)}{=} 0\mathbf{v} + (0\mathbf{v} + (-0\mathbf{v})) \overset{(4)}{=} 0\mathbf{v} + \mathbf{0} \overset{(3)}{=} 0\mathbf{v}$. Dus $0\mathbf{v} = \mathbf{0}$.
(f): $c\mathbf{0} \overset{(3)}{=} c(\mathbf{0}+\mathbf{0}) \overset{(8)}{=} c\mathbf{0} + c\mathbf{0}$. Door weer aan beide zijden $(-c\mathbf{0})$ op te tellen, volgt nu, net als bij het bewijs van (e), $\mathbf{0} = c\mathbf{0}$.
(g): $\mathbf{v} + (-1)\cdot\mathbf{v} \overset{(6)}{=} 1\cdot\mathbf{v} + (-1)\cdot\mathbf{v} \overset{(7)}{=} (1-1)\cdot\mathbf{v} = 0\mathbf{v} \overset{(e)}{=} \mathbf{0}$. Vanwege de uniciteit van de tegengestelde (uitspraak (b)) volgt nu het gestelde.
(h): Stel dat $c\mathbf{v} = \mathbf{0}$. Als $c = 0$, dan zijn we klaar. Daarom nemen we aan dat $c \neq 0$ (en dus bestaat $\frac{1}{c}$) en we tonen aan dat volgt: $\mathbf{v} = \mathbf{0}$. Als volgt: $c\mathbf{v} = \mathbf{0} \Rightarrow \frac{1}{c}(c\mathbf{v}) = \frac{1}{c}\mathbf{0}$ en uit rekenregel (5) en uit (f) volgt: $(\frac{1}{c}c)\mathbf{v} = \mathbf{0} \Rightarrow (1)\mathbf{v} = \mathbf{0} \overset{(6)}{\Rightarrow} \mathbf{v} = \mathbf{0}$. De conclusie volgt. □

Standaard Reële Vectorruimten

We geven nu een tamelijk uitputtende lijst van standaardvoorbeelden van reële vectorruimten. Deze voorbeelden zullen we in de rest van het dictaat regelmatig gebruiken.

Voorbeeld 2.2. Het eerste voorbeeld is natuurlijk $\mathbb{R}^n$. Dit voorbeeld is de afgelopen tijd uitputtend besproken.

Voorbeeld 2.3. We noteren met $\mathcal{M}_{m\times n}(\mathbb{R})$ de verzameling van $m \times n$ matrices met kentallen in $\mathbb{R}$. Twee matrices uit $M_{m\times n}(\mathbb{R})$ kunnen we bij elkaar optellen door de kentallen op de corresponderende plaatsen telkens bij elkaar op te tellen. Een matrix kan met een reëel getal worden vermenigvuldigd door elk van de kentallen met dat reële getal te vermenigvuldigen. Het is eenvoudig in te zien dat $M_{m\times n}(\mathbb{R})$ met deze bewerkingen een vectorruimte is. De vectoren uit deze vectorruimte zijn hier dus matrices. Als illustratie controleren we voor $M_{2\times 2}(\mathbb{R})$ de zevende rekenregel:

$$c_1 \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + c_2 \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \overset{\text{def.sc.verm.}}{=} \begin{bmatrix} c_1a_{11} & c_1a_{12} \\ c_1a_{21} & c_1a_{22} \end{bmatrix} + \begin{bmatrix} c_2a_{11} & c_2a_{12} \\ c_2a_{21} & c_2a_{22} \end{bmatrix} \overset{\text{def.opt.}}{=}$$

$$= \begin{bmatrix} c_1a_{11} + c_2a_{11} & c_1a_{12} + c_2a_{12} \\ c_1a_{21} + c_2a_{21} & c_1a_{22} + c_2a_{22} \end{bmatrix} \overset{\text{eig.van } \mathbb{R}}{=} \begin{bmatrix} (c_1+c_2)a_{11} & (c_1+c_2)a_{12} \\ (c_1+c_2)a_{21} & (c_1+c_2)a_{22} \end{bmatrix} \overset{\text{def.sc.verm.}}{=}$$

$$= (c_1 + c_2) \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

De uitspraak: "kies een $m \times n$ matrix A met reële coëfficienten" kan met de ingevoerde notatie worden afgekort tot: "kies $A \in \mathcal{M}_{m\times n}(\mathbb{R})$".

Voorbeeld 2.4. Bekijk de verzameling van alle oneindige rijtjes (voor het gemak horizontaal geschreven) van reële getallen gedefinieerd door

$$\{(a_0, a_1, a_2, a_3, \cdots) \mid a_0, a_1, a_2, a_3, \cdots \in \mathbb{R}\}.$$

In de Elektro heten zulke rijtjes vaak "tijd discrete signalen".
De optelling en scalaire vermenigvuldiging definieëren we coördinaatsgewijs:

$$(a_0, a_1, a_2, \cdots) + (b_0, b_1, b_2, \cdots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \cdots)$$
$$c(a_0, a_1, a_2, \cdots = (ca_0, ca_1, ca_2, \cdots).$$

De verzameling is gesloten ten opzichte van deze optelling en scalaire vermenigvuldiging. Verder voldoet deze verzameling aan alle vectorruimte-axioma's. Het is daarmee een vectorruimte over $\mathbb{R}$. In feite zijn vectoren nu kolommen (of rijen) met oneindig veel reële kentallen en in die zin is het een uitbreiding van $\mathbb{R}^n$. Deze vectorruimte wordt dan ook als $\mathbb{R}^\infty$ genoteerd.

Door de indices van $-\infty$ tot ∞ te laten lopen, kan nog een rijtjesruimte gemaakt worden. De verzameling is dan

$$\{(\cdots, a_{-2}, a_{-1}, a_0, a_1, a_2, \cdots) \mid \cdots, a_{-2}, a_{-1}, a_0, a_1, a_2, \cdots \in \mathbb{R}\}.$$

De optelling en scalaire vermenigvuldiging worden weer coördinaatsgewijs gedefinieërd:

$$\begin{aligned}(\cdots, a_{-1}, a_0, a_1, \cdots) + (\cdots, b_{-1}, b_0, b_1, \cdots) &= (\cdots, a_{-1} + b_{-1}, a_0 + b_0, a_1 + b_1, \cdots)\\ c(\cdots, a_{-1}, a_0, a_1, \cdots) &= (\cdots, ca_{-1}, ca_0, ca_1, \cdots).\end{aligned}$$

Deze vectorruimte van de reële dubbel oneindige rijtjes wordt genoteerd met $\mathbb{R}^{\pm\infty}$.

Voorbeeld 2.5. Als X een formele variabele is, dan heet

$$p(X) = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n \qquad \text{met } a_i \in \mathbb{R}$$

een reëel polynoom van de graad tenhoogste n. De verzameling van alle reële polynomen van graad tenhoogste n wordt genoteerd met

$$\mathbb{P}_n(\mathbb{R})$$

Pas op, men wordt niet geacht waarden aan de X toe te kennen, reëel polynoom betekent dat de coëfficienten a_k reëel zijn, niet de X.
De graad van $p(X) \in \mathbb{P}_n(\mathbb{R})$ is de grootste $k \leq n$ met $a_k \neq 0$. Mochten alle a_k gelijk zijn aan 0, dan zeggen we dat het polynoom van de graad -1 is.
Op $\mathbb{P}_n(\mathbb{R})$ definiëren we alsvolgt een optelling en een scalaire vermenigvuldiging: als $p(X) = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$ en $q(X) = b_0 + b_1X + b_2X^2 + \cdots + b_nX^n$ dan is

$$\begin{aligned}p(X) + q(X) &= (a_0 + b_0) + (a_1 + b_1)X + (a_2 + b_2)X^2 + \cdots + (a_n + b_n)X^n\\ cp(X) &= (ca_0) + (ca_1)X + (ca_2)X^2 + \cdots + (ca_n)X^n.\end{aligned}$$

De verzameling $\mathbb{P}_n(\mathbb{R})$ is gesloten ten opzichte van deze optelling en scalaire vermenigvuldiging. Verder voldoet deze verzameling aan alle vectorruimte-axioma's. Het is daarmee een vectorruimte over $\mathbb{R}$.
Als we de afspraak maken dat in de notatie van $p(X)$ die termen a_kX^k van $p(X)$ met $a_k = 0$ weggelaten mogen worden (behalve als alle a_k nul zijn, dan schrijven we $p(X) = 0$), dan geldt:

$$\mathbb{P}_n(\mathbb{R}) \subset \mathbb{P}_{n+1}(\mathbb{R})$$

(dit in tegenstelling tot $\mathbb{R}^n$: er geldt NIET: $\mathbb{R}^n \subset \mathbb{R}^{n+1}$).

Het volgende voorbeeld beschrijft de belangrijkste en MEEST voorkomende vectorruimten, namelijk de functieruimten. Juist deze ruimten zijn de reden van de uitbreiding van $\mathbb{R}^n$ tot algemene vectorruimten en ze zijn voor fysische toepassingen van groot belang. Voordat we overgaan tot de precieze formulering eerst even wat constateringen.

Voorbeeld 2.6. We bekijken straks de verzameling V van alle functies naar $\mathbb{R}$ met een vast domein, bijvoorbeeld $\mathbb{R}$, dus $V = \{f | f : \mathbb{R} \to \mathbb{R}\}$. De functies f_1, f_2 en f_3 met de voorschriften

$$f_1(x) = 5x^2 - 2x + 8, \quad f_2(x) = 4\sqrt{3|x|}, \quad f_3(x) = |x - \frac{1}{2}|$$

behoren bijvoorbeeld tot deze verzameling V. Ook de functie f_4 met voorschrift

$$f_4(x) = \begin{cases} 0 & \text{als } x \in \mathbb{Q} \\ 1 & \text{als } x \notin \mathbb{Q} \end{cases}$$

behoort tot V. De functie f_5 met voorschrift $f_5(x) = 2\ln(-5x)$ behoort niet tot V.
Twee functies van $[0,1]$ naar $\mathbb{R}$ kunnen we optellen. Dit levert een nieuwe functie van $[0,1]$ naar $\mathbb{R}$. Zo geldt bijvoorbeeld dat $f_1 + f_2$ de functie g is met voorschrift

$$g(x) = 5x^2 - 2x + 8 + 4\sqrt{3x}.$$

Overigens is het onnodig om de letter g te gebruikten als naam voor de somfunctie, $f_1 + f_2$ is eigenlijk veel duidelijker:

$$(f_1 + f_2)(x) = 5x^2 - 2x + 8 + 4\sqrt{3|x|}.$$

Op dezelfde manier is $f_2 + f_4$ de functie met voorschrift

$$(f_2 + f_4)(x) = \begin{cases} 4\sqrt{3|x|} & \text{als } x \in \mathbb{Q} \\ 4\sqrt{3|x|} + 1 & \text{als } x \notin \mathbb{Q}. \end{cases}$$

Een functie kan ook met een reëel getal vermenigvuldigd worden: $3f_1$ is de functie met voorschrift $(3f_1)(x) = 15x^2 - 6x + 24$ en $-\sqrt{2}f_2$ heeft voorschrift $-4\sqrt{6x}$.

Bovenstaand voorbeeld geeft aanleiding tot een som en scalaire vermenigvuldiging op een verzameling functies met vast domein en dus ook tot een vectorruimtestructuur.

Voorbeeld 2.7. We bekijken hier de verzameling van alle functies van een vast domein D naar $\mathbb{R}$. Deze wordt genoteerd als

$$\mathcal{F}(D, \mathbb{R})$$

Twee functies f_1 en f_2 van D naar $\mathbb{R}$ kunnen we bij elkaar optellen tot een functie $f_1 + f_2$ van D naar $\mathbb{R}$, door in ieder punt van het domein D de twee functiewaarden bij elkaar op te tellen, dus

$$(f_1 + f_2)(x) = f_1(x) + f_2(x) \qquad \text{voor alle } x \in D.$$

Verder kunnen we een functie f in zijn geheel met een scalair $c \in \mathbb{R}$ vermenigvuldigen zodat we de functie cf krijgen via

$$(cf)(x) = c \cdot f(x) \qquad \text{voor alle } x \in D.$$

Met deze optelling en scalaire vermenigvuldiging voldoet de *functieruimte* $\mathcal{F}(D, \mathbb{R})$ aan alle axioma's van een vectorruimte over $\mathbb{R}$. De nulfunctie die elk punt van het domein op $0 \in \mathbb{R}$ afbeeldt, fungeert daarbij als de nulvector. Deze functie wordt vaak (verwarrend) genoteerd met 0. Deze 0 staat dus niet voor het getal nul maar voor een functie, waar je dus waarden voor kunt invullen. (Dus: $0(\mathbf{x}) = 0$, voor alle $\mathbf{x} \in D$) Na invullen van zo'n waarde krijg je wel het *getal* nul. [In het vervolg zullen we vaak een berekening in de trend van "bepaal alle scalairen c en d zo dat $c\sin x + d\cos x = 0$" tegenkomen. Bedoeld wordt dan dat de functie f met voorschrift $f(x) = c\sin x + d\cos x$ gelijk moet zijn aan de nulfunctie. De gelijkheid moet dus gelden voor alle $x \in D$. Vaak kunnen c en d dan bepaald worden door enkele verstandige waarden voor x in te vullen. Hier werken $x = 0$ en $x = \frac{1}{2}\pi$ bijvoorbeeld goed. Die waarden leveren respectievelijk $d = 0$ en $c = 0$.]
De tegengestelde van een functie f is gelijk aan de functie $(-1)f$ want

$$f(x) + ((-1)f)(x) = f(x) + (-f(x)) = 0 = 0(x) \qquad \text{voor alle } x.$$

Merk op dat dit eigenlijk onderdeel (g) van Stelling 2.1 is, dus eigenlijk hadden we dit al bewezen!

Standaard Complexe Vectorruimten

We geven eveneens een lijst van standaardvoorbeelden van complexe vectorruimten. Deze voorbeelden zullen we in de rest van het dictaat regelmatig gebruiken.

Voorbeeld 2.8. Het eerste voorbeeld is natuurlijk $\mathbb{C}^n$. Dit voorbeeld is de afgelopen tijd uitputtend besproken, en we worden geacht experts te zijn op dit gebied.

Voorbeeld 2.9. We noteren met $\mathcal{M}_{m\times n}(\mathbb{C})$ de verzameling van $m \times n$ matrices met kentallen in $\mathbb{C}$. Twee matrices uit $M_{m\times n}(\mathbb{C})$ kunnen we bij elkaar optellen door de kentallen op de corresponderende plaatsen telkens bij elkaar op te tellen. Een matrix kan met een complex getal worden vermenigvuldigd door elk van de kentallen met dat complexe getal te vermenigvuldigen. Het is eenvoudig in te zien dat $M_{m\times n}(\mathbb{C})$ met deze bewerkingen een vectorruimte over $\mathbb{C}$ is. De uitspraak: "kies een $m \times n$ matrix A met complexe coëfficienten" kan met de ingevoerde notatie worden afgekort tot: "kies $A \in \mathcal{M}_{m\times n}(\mathbb{C})$".

Voorbeeld 2.10. Bekijk de verzameling van alle oneindige rijtjes (voor het gemak horizontaal geschreven) van complexe getallen gedefinieërd door

$$\{(a_0, a_1, a_2, a_3, \cdots) \mid a_0, a_1, a_2, a_3, \cdots \in \mathbb{C}\}.$$

De optelling en scalaire vermenigvuldiging definieëren we weer coördinaatsgewijs:

$$\begin{aligned}(a_0, a_1, a_2, \cdots) + (b_0, b_1, b_2, \cdots) &= (a_0 + b_0, a_1 + b_1, a_2 + b_2, \cdots)\\ c(a_0, a_1, a_2, \cdots) &= (ca_0, ca_1, ca_2, \cdots).\end{aligned}$$

De verzameling is gesloten ten opzichte van deze optelling en scalaire (uit $\mathbb{C}$) vermenigvuldiging. Verder voldoet deze verzameling aan alle vectorruimte-axioma's. Het is daarmee een vectorruimte over $\mathbb{C}$. Deze vectorruimte wordt als $\mathbb{C}^\infty$ genoteerd.

Door de indices van $-\infty$ tot ∞ te laten lopen, kan nog een rijtjesruimte gemaakt worden. De verzameling is dan

$$\{(\cdots, a_{-2}, a_{-1}, a_0, a_1, a_2, \cdots) \mid \cdots, a_{-2}, a_{-1}, a_0, a_1, a_2, \cdots \in \mathbb{C}\}.$$

De optelling en scalaire vermenigvuldiging worden weer coördinaatsgewijs gedefinieërd:

$$\begin{aligned}(\cdots, a_{-1}, a_0, a_1, \cdots) + (\cdots, b_{-1}, b_0, b_1, \cdots) &= (\cdots, a_{-1} + b_{-1}, a_0 + b_0, a_1 + b_1, \cdots)\\ c(\cdots, a_{-1}, a_0, a_1, \cdots) &= (\cdots, ca_{-1}, ca_0, ca_1, \cdots).\end{aligned}$$

Deze vectorruimte van de reële dubbel oneindige rijtjes wordt genoteerd met $\mathbb{C}^{\pm\infty}$.

Voorbeeld 2.11. Als X een formele variabele is, dan heet

$$p(X) = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n \quad \text{met } a_i \in \mathbb{C}$$

een complex polynoom van de graad tenhoogste n. De verzameling van alle complexe polynomen van graad tenhoogste n wordt genoteerd met

$$\mathbb{P}_n(\mathbb{C})$$

Op $\mathbb{P}_n(\mathbb{C})$ wordt op dezelfde manier als op $\mathbb{P}_n(\mathbb{R})$ een optelling en een scalaire vermenigvuldiging gedefiniëerd en wordt daarmee een vectorruimte over $\mathbb{C}$.

De voor ons weer belangrijkste vectorruimten zijn de complexe functieruimten.

Voorbeeld 2.12. We bekijken hier de verzameling van alle functies van een vast domein D naar $\mathbb{C}$. Deze wordt genoteerd als

$$\mathcal{F}(D, \mathbb{C})$$

Twee functies f_1 en f_2 van D naar $\mathbb{C}$ kunnen we bij elkaar optellen tot een functie $f_1 + f_2$ van D naar $\mathbb{C}$, door in ieder punt van het domein D de twee functiewaarden bij elkaar op te tellen, dus

$$(f_1 + f_2)(x) = f_1(x) + f_2(x) \qquad \text{voor alle } x \in D.$$

Verder kunnen we een functie f in zijn geheel met een scalair $c \in \mathbb{C}$ vermenigvuldigen zodat we de functie cf krijgen via

$$(cf)(x) = c \cdot f(x) \qquad \text{voor alle } x \in D.$$

Met deze optelling en scalaire vermenigvuldiging voldoet de *functieruimte* $\mathcal{F}(D, \mathbb{C})$ aan alle axioma's van een vectorruimte over $\mathbb{C}$. De nulfunctie die elk punt van het domein op $0 \in \mathbb{C}$ afbeeldt, fungeert daarbij als de nulvector.
Merk op dat bij bijvoorbeeld $\mathcal{F}([a, b], \mathbb{C})$ functies bekeken worden van een reële variabele x, echter de functiewaarde $f(x)$ zit in $\mathbb{C}$!

De Grote L_2 ruimten

In de wiskunde (en in de quantum-mechanica) neemt de Grote L_2 ruimte een bijzonder belangrijke plaats in. Echter, om deze goed te behandelen zouden we eigenlijk eerst een college echte analyse moeten geven, over maattheorie en over de Lebesque–integraal.
Een functie $f : [a, b] \to \mathbb{R}$ (of $\mathbb{C}$) heet **kwadratisch integreerbaar** over het interval $[a, b]$ als

$$\int_a^b |f(x)|^2 dx < \infty.$$

Hier staat $\int_a^b f(x)\,dx$ voor de Lebesque–integraal van f. Voor alle functies waarmee natuurkundigen werken is dit gelijk aan de ons bekende Riemann integraal.
De verzameling van alle reële (resp. complexe) kwadratisch integreerbare functies op $[a, b]$ wordt genoteerd met

$$L_2([a, b], \mathbb{R}) \;\; \text{respectievelijk} \;\; L_2([a, b], \mathbb{C})$$

Dat dit inderdaad een vectorruimte is, valt niet mee om aan te tonen. (Waarom, als f, g kwadratisch integreerbaar zijn, is ook $f + g$ kwadratisch integreerbaar?)
De reden dat $L_2([a, b], \mathbb{R})$ (en $L_2([a, b], \mathbb{C})$) verschillend is van $\mathcal{F}([a, b], \mathbb{R})$ (respectievelijk $\mathcal{F}([a, b], \mathbb{C})$) (of liever verschillend is van de verzameling kwadratisch integreerbare functies in $\mathcal{F}([a, b], \mathbb{R})$) is de volgende aanname die we zullen doen.
"Functies in $L_2([a, b], \mathbb{R})$ die slechts in eindig veel punten van het domein verschillen zullen we identificeren."
Dit is U niet vreemd, dit bent U ook al gewend van het college voortgezette analyse. Daar identificeerde u twee "piecewise–continuous functions" als ze slechts in eindig veel punten verschilde. Wiskundig correct zou zijn functies te identificeren die slechts op een verzameling van maat 0 verschillend waren. Maar dan is kennis van de maattheorie nodig.

2.2 De afhankelijkheid van de vectorruimte axioma's

Opmerking: Deze paragraaf betreft enkele opmerkingen aangaande het verschil tussen axioma's en stelling. Het is zeker geen tentamenstof.

Wat is nu het verschil tussen de 8 axioma's (eisen) uit de definitie van een vectorruimte en de 8 stellingen (uitspraken) uit Stelling 2.1?

Welnu, de uitspraak

$$(-1)\mathbf{v} = -\mathbf{v}$$

is een bijzonder gewenste identiteit. Deze hoeft echter niet als een negende axioma aan de definitie worden toegevoegd, want deze is al een *GEVOLG* van de eerste 8 axioma's (zie stelling 2.1).
Zo kan men zich ook afvragen: zou men axioma 6:

$$1\mathbf{v} = \mathbf{v}$$

niet mogen weglaten uit de axioma's en deze als negende resultaat aan Stelling 2.1 kunnen toevoegen?
Het volgende voorbeeld laat zien dat niet mag. Dit is een voorbeeld van een "soort lineaire structuur", een optelling en een scalaire vermenigvuldiging, die aan de axioma's 1 tot en met 5, 7 en 8 voldoet, maar NIET aan axioma 6. En dit voorbeeld laat zien dat axioma 6 GEEN gevolg kan zijn van de overigen axioma's en dat dit axioma dus echt nodig is (tenminste, als men echt altijd wilt dat $1\mathbf{v} = \mathbf{v}$).

Voorbeeld 2.13. Beschouw $V = \mathbb{R}^2$ met de standaard optelling $\mathbf{u} + \mathbf{v}$ maar de niet standaard scalaire vermenigvuldiging

$$c\mathbf{v} = \mathbf{0} \quad \text{voor alle } \mathbf{v} \in \mathbb{R}^2 \text{ en } c \in \mathbb{R}.$$

Het is eenvoudig in te zien dat deze optelling en scalaire vermenigvuldiging aan alle axioma's voldoet, behalve aan axioma 6. Dit is dus GEEN voorbeeld van een vectorruimte.

Opmerking: Het is wellicht een ziekte onder de wiskundigen, maar die willen er van overtuigd zijn dat iedere axioma van de lijst echt nodig is. Bij de opgaven kunt u nog twee voorbeelden vinden van een "soort lineaire structuur", een optelling en een scalaire vermenigvuldiging, die aan precies 7 van de 8 axioma's voldoen.

In ongeveer 1910 werd ontdekt dat de axioma's van een vectorruimte niet onafhankelijk van elkaar zijn. Axioma 1 ($x + y = y + x$) is af te leiden uit de overige axioma's. En wel op de volgende manier.

Bewijs: Kies $\mathbf{x}$ en $\mathbf{y}$ uit V en fixeer deze.
Merk op: $2\mathbf{x} = (1+1)\mathbf{x} = 1\mathbf{x} + 1\mathbf{x} = \mathbf{x} + \mathbf{x}$. (Gebruikt: axoima 7 en 6.)
In het bijzonder geldt:
$2(\mathbf{x} + \mathbf{y}) = \mathbf{x} + \mathbf{y} + \mathbf{x} + \mathbf{y}$. (We gebruiken axioma 2 door geen haakjes te gebruiken!)
Ook geldt volgens axioma 8:
$2(\mathbf{x} + \mathbf{y}) = 2\mathbf{x} + 2\mathbf{y} = (1+1)\mathbf{x} + (1+1)\mathbf{y} = \mathbf{x} + \mathbf{x} + \mathbf{y} + \mathbf{y}$
We concluderen: $\mathbf{x} + \mathbf{x} + \mathbf{y} + \mathbf{y} = \mathbf{x} + \mathbf{y} + \mathbf{x} + \mathbf{y} \quad (**)$
Vervolgens gebruiken we axioma 3 en 4 door bij deze identiteit linska een $\mathbf{x}$ en rechts een $\mathbf{y}$ af te snoepen, door:
$-\mathbf{x} + \mathbf{x} + \mathbf{x} + \mathbf{y} + \mathbf{y} - \mathbf{y} = -\mathbf{x} + \mathbf{x} + \mathbf{y} + \mathbf{x} + \mathbf{y} - \mathbf{y}$
Via de associatieve wet (axioma 2) krijgen we:
$(-\mathbf{x} + \mathbf{x}) + (\mathbf{x} + \mathbf{y}) + (\mathbf{y} - \mathbf{y}) = (-\mathbf{x} + \mathbf{x}) + (\mathbf{y} + \mathbf{x}) + (\mathbf{y} - \mathbf{y})$
en dus:
$\mathbf{0} + (\mathbf{x} + \mathbf{y}) + \mathbf{0} = \mathbf{0} + (\mathbf{y} + \mathbf{x}) + \mathbf{0}$
d.w.z.
$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
Toch wel een slim en leuk bewijs. $\square$

2.3 Lineaire Opspanning en Deelruimten

In deze paragraaf komen we tot de bestudering van de fundamentele begrippen van de lineaire algebra.

Lineaire Combinaties en Lineaire Omhulsels

Voordat we de betreffende definities geven, eerst de volgende terminologie. We noemen een geordende verzameling van een *eindig* stel niet noodzakelijk verschillende vectoren $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ uit een vectorruimte V een *stelsel vectoren* Als de volgorde van de vectoren van belang is, wordt dit stelsel vectoren genoteerd als

$$(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n),$$

Als de volgorde niet van belang is, kan het stelsel ook genoteerd worden als

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}.$$

en spreken we van een *verzameling vectoren*. Een stelsel vectoren kan eventueel leeg zijn.
We benadrukken nog eens dat in een stelsel vectoren eventueel twee of meer gelijke vectoren kunnen voorkomen.

Definitie: Als $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ (of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$) een niet leeg stelsel (of verzameling) vectoren uit een vectorruimte V over $\mathbb{R}$ (respectievelijk. $\mathbb{C}$) is en als $c_1, c_2, \ldots, c_n$ getallen in $\mathbb{R}$ (respectievelijk $\mathbb{C}$) zijn, dan wordt een vector van de vorm

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

een *lineaire combinatie* van het stelsel genoemd en de getallen $c_1, c_2, \ldots, c_n$ de *gewichten* van de lineaire combinatie.
Als het stelsel leeg is, kunnen we een dergelijke uitdrukking niet opschrijven. Desondanks noemen we in dat geval de nulvector $\mathbf{0}$ een lineaire combinatie van dit lege stelsel.

De verzameling U van *alle* lineaire combinaties van het stelsel heet het *lineaire omhulsel* (of het *lineaire opspansel* of kortweg de span van $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$, of de span van $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$).
Notatie:

$$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) \quad \text{of} \quad \mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}.$$

We zeggen vaak dat het stelsel de verzameling U *voortbrengt* of *opspant*.

Het is je waarschijnlijk al opgevallen dat bovenstaande definitie al eerder gegeven is voor vectoren uit $\mathbb{R}^n$ en $\mathbb{C}^n$. Realiseer je verder dat deze definitie zeer belangrijk is en dat de begrippen lineaire combinatie en lineair omhulsel de sleutel zijn tot het begrijpen van de lineaire algebra.

Opmerking: Uit bovenstaande definitie volgt dat het lineair omhulsel van het lege stelsel uit één vector bestaat, namelijk de nulvector: $\mathrm{Span}() = \mathrm{Span}\{\} = \{\mathbf{0}\}$.
Een lineaire combinatie van een stelsel dat uit slechts één vector $\mathbf{v}$ bestaat, is een scalair veelvoud van die vector. Met andere woorden, $\mathrm{Span}\{\mathbf{v}\}$ (de verzameling van alle lineaire combinaties van dit stelsel dus) bestaat precies uit de vectoren van de vorm $c\mathbf{v}$ met $c \in \mathbb{L}$. In $\mathbb{R}^n$ is dat de rechte lijn door $\mathbf{0}$, in de richting van $\mathbf{v}$.

Bekijk nu een verzameling $\{\mathbf{u},\mathbf{v}\}$ van twee vectoren in $\mathbb{R}^n$, waarbij de ene vector geen veelvoud is van de andere. De verzameling $\text{Span}\{\mathbf{u},\mathbf{v}\}$ van alle lineaire combinaties van dit stelsel is dan een vlak

Opmerking: Soms zal het niet à priori duidelijk zijn "welke span" we bedoelen. Als we zonder verdere informatie praten over

$$\text{Span}\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$$

dan is het mogelijk dat we de span in de reële vectorruimte $\mathbb{R}^2$ bedoelen, maar ook is het mogelijk dat we de span in de complexe vectorruimte $\mathbb{C}^2$ bedoelen. Om deze te onderscheiden schrijven we soms

$$\text{Span}_{\mathbb{R}}\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}\quad\text{dan wel}\quad\text{Span}_{\mathbb{C}}\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$$

om aan te geven dat we lineaire combinaties van $\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}$ bekijken met reële gewichten dan wel met complexe gewichten. Dus:

$$\text{Span}_{\mathbb{R}}\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}=\mathbb{R}^2\quad\text{en}\quad\text{Span}_{\mathbb{C}}\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}=\mathbb{C}^2$$

Meestal is het uit de context duidelijk welke van de twee bedoeld wordt.

De afgelopen tijd zijn al vele voorbeelden gegeven die betrekking hadden op $\mathbb{R}^n$ en $\mathbb{C}^n$. Wij geven nu een voorbeeld dat betrekking heeft op functieruimten.

Voorbeeld 2.14. In de vectorruimte $V=\mathcal{F}([0,2\pi],\mathbb{C})$ beschouwen we de vier functies f_1, f_2, f_3 en f_4 gedefinieërd door

$$f_1(x)=\sin x,\quad f_2(x)=\cos x,\quad f_3(x)=\sin^2 x,\quad\text{en}\quad f_4(x)=\cos^2 x.$$

We stellen de volgende vraag: "behoren de functies

$$f(x)=\sin 2(x+\pi/4)\quad\text{en}\quad g(x)=\cos 2(x+\pi/4)$$

dan tot het lineair omhulsel $U=\text{Span}\{f_1,f_2,f_3,f_4\}$ van de verzameling $\{f_1,\ldots,f_4\}$?"
Uit de definitie van lineair omhulsel volgt dat we moeten nagaan of f en g lineaire combinaties zijn van f_1 t/m f_4. We kunnen eerst proberen met de bekende sinus- en cosinus-regels f en g als lineaire combinaties van de opspannende functies te schrijven:

$$\begin{aligned}f(x)&=\sin 2(x+\pi/4)=\sin(2x+\pi/2)=\cos 2x=\cos^2 x-\sin^2 x\\ g(x)&=\cos 2(x+\pi/4)=\cos(2x+\pi/2)=-\sin 2x=-2\sin x\cos x.\end{aligned}$$

Hieruit volgt dat $f=-f_3+f_4$ en dus een lineaire combinatie is van het stelsel $\{f_1,\ldots,f_4\}$. Bijgevolg behoort f tot U. Voor g weten we dat nog niet!
Stel nu dat g wel in U zit, dus stel dat

$$g=a_1f_1+a_2f_2+a_3f_3+a_4f_4$$

waarbij de a_i scalairen zijn, hier dus complexe getallen. (Want V is een complexe vectorruimte.) Dit is een gelijkheid van functies, dus zou dit voor elke x in $[0,2\pi]$ een gelijkheid moeten opleveren:

$$\cos 2(x+\pi/4)=a_1\sin x+a_2\cos x+a_3\sin^2 x+a_4\cos^2 x$$

We vullen de punten 0, $\frac{1}{2}\pi$, π, $\frac{3}{2}\pi$ in. Dit geeft het stelsel vergelijkingen:

$$\begin{cases} 0 &= & & +a_2 & & +a_4 \\ 0 &= & +a_1 & & +a_3 & \\ 0 &= & & -a_2 & & +a_4 \\ 0 &= & -a_1 & & +a_3 & \end{cases}$$

Lossen we dit stelsel op, dan vinden we $a_1 = a_2 = a_3 = a_4 = 0$. Maar dan zou g gelijk zijn aan de nulfunctie. Dat is echter niet zo, omdat bijvoorbeeld $g(\frac{1}{4}\pi) = -1$. Blijkbaar is g geen element van U.
Merk nog op dat de relatie $f = -f_3 + f_4$ voor f ook gevonden had kunnen worden door — net als bij g gedaan is — punten in te vullen en de coëfficiënten van de lineaire combinatie uit te rekenen. Dit kan handig zijn als we de eigenschappen van de functie niet zouden kennen. In dat geval weten we echter nog niet of de zo gevonden relatie voor alle x geldt. Daarvoor is dan nog een bewijs nodig zoals hierboven voor f gegeven is.

Opmerking: In het vorige voorbeeld hebben we het stelsel nogal omslachtig gedefinieërd via de functies genaamd $f_1, \ldots, f_4$. Daarmee is nadruk gelegd op het feit dat de vectoren hier functies zijn en dat gelijkheid van twee functies betekent dat deze dan ook in ieder punt aan elkaar gelijk zijn. In het vervolg zullen we om een functie aan te duiden voor het gemak vaak alleen het voorschrift gebruiken. In bovenstaand voorbeeld wordt het stelsel dan $\{\sin x, \cos x, \sin^2 x, \cos^2 x\}$. Ofschoon hier eigenlijk *waarden* staan, bedoelen we dan (tenzij anders vermeld) *functies.*

We beëindigen deze paragraaf met een generalisatie van het begrip span. In de oorspronkelijke definitie nemen we namelijk de span van een eindig stel vectoren. We kunnen ook spreken over de span van een oneindige verzameling vectoren.

Definitie: Zij S een verzameling vectoren in een vectorruimte V over $\mathbb{R}$ (of $\mathbb{C}$) (S kan dus oneindig veel vectoren bevatten). Een vector $v \in V$ is een *lineaire combinatie* van S als

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

voor zekere vectoren $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_n$ uit S en zekere scalairen c_1, c_2, …, c_n uit $\mathbb{R}$ (of $\mathbb{C}$).
De verzameling $\operatorname{Span} S$ van alle lineaire combinaties van S heet het *lineaire omhulsel* (of het *lineaire opspansel* of kortweg de span) van S.

Let op, ondanks het feit dat S oneindig veel vectoren kan bevatten, blijven we toch werken met lineaire combinaties van eindig veel vectoren, zogenaamde *eindige lineaire combinaties*

Voorbeeld 2.15. Als in $\mathcal{F}([0, 2\pi], \mathbb{R})$ de oneindige verzameling

$$S = \{1, \cos x, \sin x, \cos^2 x, \sin^2 x, \cos^3 x, \sin^3 x, \ldots\}$$

genomen wordt, dan bestaat $\operatorname{Span} S$ uit alle functies van de vorm

$$f(x) = a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos^2 x + b_2 \sin^2 x + \cdots + a_n \cos^n x + b_n \sin^n x$$

waarbij n een natuurlijk getal is en de a_i en b_i reële getallen zijn.

Lineaire Deelruimten

We hebben in de vorige paragraaf vectorruimten ingevoerd. Nu blijkt dat een heleboel deelverzamelingen van deze vectorruimten zelf ook vectorruimten zijn (waarbij we de optelling en de scalaire vermenigvuldiging hetzelfde nemen als de "moedervectorruimte"). Het vervelende is dat telkens alle eisen van vectorruimte nagegaan moeten worden om te bepalen of zo'n deelverzameling werkelijk een vectorruimte is. We zullen zo zien dat dit helemaal niet nodig is, omdat de axioma's (1) t/m (8) kado krijgen, want deze gelden in de "moedervectorruimte". De volgende definitie representeert dit.

Definitie: Een deelverzameling U van een vectorruimte V over een lichaam $\mathbb{L}$ heet een **deelruimte** van V, als:

(1) U bevat de nulvector van V.

(2) $\mathbf{x} + \mathbf{y} \in U$ voor elke keuze van $\mathbf{x}$ en $\mathbf{y}$ uit U

(3) $c\mathbf{x} \in U$ voor elke keuze van $\mathbf{x}$ uit U en c uit $\mathbb{L}$

Merk op dat deze definitie inderdaad overeenkomt met de definitie van vectorruimte zonder de acht rekenregels. Dat komt omdat de eerste eis (dus U bevat de nulvector van V) equivalent is met de eis "U is niet leeg". Met de derde eis volgt dan namelijk vanzelf dat U de nulvector van V moet bevatten.

Zoals beloofd hebben deelruimten de volgende eigenschap.

STELLING 2.16. *Een deelruimte van een vectorruimte over $\mathbb{L}$ is zelf ook een vectorruimte over $\mathbb{L}$.*

Bewijs: Neem een deelruimte U van een vectorruimte V over $\mathbb{L}$. Allereerst is U dan door eigenschap (1) (uit de definitie van deelruimte) niet leeg en door de eigenschappen (2) en (3) (opnieuw uit de definitie van deelruimte) gesloten ten opzichte van de optelling en de scalaire vermenigvuldiging. Nu moeten nog de acht rekenregels uit de definitie van vectorruimte gecontroleerd worden. De eigenschappen (1), (2), (5), (6), (7) en (8) erft U al gelijk van de grotere ruimte V. Verder zit de nulvector van V op grond van de eerste eis van een deelruimte in U en er is dus aan (3) voldaan. Tenslotte bestaat er, vanwege de vectorruimte-eigenschappen van V, voor elk element $\mathbf{v}$ in U een tegengestelde $-\mathbf{v}$ in V. Maar ligt $-\mathbf{v}$ dan ook in U? Met stelling 2.1(g) weten we dat $-\mathbf{v} = (-1) \cdot \mathbf{v}$ en dit zit in U op grond van eis (3) voor deelruimten. Daarmee is aan alle eisen voor een vectorruimte voldaan. $\square$

Het moge nu duidelijk zijn dat het zinvol is om, bij het aantonen dat een verzameling een vectorruimte is, eerst na te gaan of deze misschien een deelruimte is van een bekende vectorruimte. Het is namelijk veel sneller om de drie eigenschappen van een deelruimte te controleren dan de elf van een vectorruimte. Bovendien zijn deze drie toch eigenschappen die we ook voor een vectorruimte hadden moeten nagaan!

Er zullen nu vele voorbeelden van, soms belangrijke, deelruimten volgen.

Voorbeeld 2.17. In iedere vectorruimte V zit een nulelement $\mathbf{0}$. De deelverzameling $\{\mathbf{0}\}$ van V is een deelruimte van V, de **nulvectorruimte**[3] $\{\mathbf{0}\}$.

De volgende stelling, die we voor eindige opspannende verzamelingen ook bij het eerste deel van lineaire algebra al zijn tegengekomen, kan handig zijn om snel te concluderen dat bepaalde verzamelingen, namelijk lineaire omhulsels, deelruimten zijn.

STELLING 2.18. *Als S een (eventueel oneindig grote) verzameling vectoren is in een vectorruimte V dan is het lineair omhulsel* $\operatorname{Span} S$ *een deelruimte van V.*

Bewijs: De nulvector behoort tot het lineair omhulsel. Neem maar een eindig stel vectoren $\mathbf{u}_1$, $\mathbf{u}_2, \ldots, \mathbf{u}_m$ in S. Dan is $\mathbf{0} = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_n$ een lineaire combinatie van deze vectoren en behoort dus tot $\operatorname{Span} S$.
Verder geldt voor elk tweetal vectoren $\mathbf{u}$ en $\mathbf{v}$ uit het lineair omhulsel dat er vectoren $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in S bestaan en getallen c_i en d_i $(i = 1, \ldots, n)$ zo dat $\mathbf{u} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$ en $\mathbf{v} = d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n$. Ga dit na! De som is dan $\mathbf{u} + \mathbf{v} = (c_1 + d_1)\mathbf{v}_1 + \cdots + (c_n + d_n)\mathbf{v}_n$ en behoort dus ook tot het lineair omhulsel.
Tenslotte geldt voor een vector $\mathbf{u}$ uit $\operatorname{Span} S$ en een scalar c dat ook $c\mathbf{u}$ tot $\operatorname{Span} S$ behoort, omdat er in die situatie eindig veel vectoren $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k$ in S bestaan en scalairen c_i $(i = 1, \ldots, k)$ zo dat $\mathbf{u} = c_1\mathbf{w}_1 + \cdots + c_k\mathbf{w}_k$, waaruit volgt dat $c\mathbf{u} = (c \cdot c_1)\mathbf{w}_1 + \cdots + (c \cdot c_k)\mathbf{w}_k$. Er is dus aan de drie eisen voor deelruimte voldaan. $\square$

Als het stelsel vectoren leeg is, klopt bovenstaande stelling ook: het lineaire opspansel van het lege stelsel is immers gelijk aan de nulvectorruimte (op grond van definitie 2.3) en dus ook een deelruimte van V.

Deelruimten van $\mathbb{R}^n$ en $\mathbb{C}^n$

Als het goed is, is men hiermee vertrouwd. We weten dat iedere deelruimte W in $\mathbb{R}^n$ op te vatten is als kolomruimte van een matrix $A \in \mathcal{M}_{n\times m}(\mathbb{R})$ (met $m \geq \dim(W)$), of als nulruimte van een matrix $B \in \mathcal{M}_{k\times n}(\mathbb{R})$ (met $k \geq n - \dim(W)$).
Evenzo, iedere deelruimte W in $\mathbb{C}^n$ is op te vatten als kolomruimte van een matrix $A \in \mathcal{M}_{n\times m}(\mathbb{C})$ (met $m \geq \dim(W)$), of als nulruimte van een matrix $B \in \mathcal{M}_{k\times n}(\mathbb{C})$ (met $k \geq n - \dim(W)$).

Opmerking: Laat $A = [\mathbf{a}_1 \ldots \mathbf{a}_m] \in \mathcal{M}_{n\times m}(\mathbb{R})$ een reële matrix zijn. Wees dan voorzichtig met het begrip "de kolomruimte van A". Men kan bedoelen

$$\operatorname{Col}(A) = \operatorname{Span}\{\mathbf{a}_1, \ldots \mathbf{a}_m\} \subset \mathbb{R}^n$$

maar ook

$$\operatorname{Col}(A) = \operatorname{Span}\{\mathbf{a}_1, \ldots \mathbf{a}_m\} \subset \mathbb{C}^n$$

Vaak is uit de context duidelijk welke van de twee kolomruimtes bedoeld wordt.

Deelruimten van $\mathbb{R}^\infty$ en $\mathbb{C}^\infty$

We zullen later zien dat dat de ruimten $\mathbb{R}^\infty$ en $\mathbb{C}^\infty$ geen "mooie vectorruimten" zijn omdat er een natuurlijk (standaard) inwendig product mist. Er zijn twee deelruimten waar dit wel mogelijk is, zullen we zien, de zogeheten "**kleine ℓ_2-ruimten**".

[3] In het Engels: zero space

Definitie: $\ell_2(\mathbb{R})$ bestaat uit alle reële rijtjes $\mathbf{x} = (x_1, x_2, \ldots, x_n, \ldots) \in \mathbb{R}^\infty$ met de eigenschap dat

$$\sum_{n=1}^{\infty} x_n^2 < \infty$$

(De zogeheten "kwadratisch sommeerbare rijtjes".) Evenzo: $\ell_2(\mathbb{C})$ bestaat uit alle complexe rijtjes $\mathbf{x} = (x_1, x_2, \ldots, x_n, \ldots) \in \mathbb{C}^\infty$ met de eigenschap dat

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty$$

En inderdaad, $\ell_2(\mathbb{R})$ is een lineaire deelruimte van $\mathbb{R}^\infty$ en $\ell_2(\mathbb{C})$ is een lineaire deelruimte van $\mathbb{C}^\infty$.

Deelruimten van functieruimten

Binnen de vectorruimte $\mathcal{F}(D, \mathbb{L})$ (waarbij $\mathbb{L}$ is $\mathbb{R}$ of $\mathbb{C}$) van functies kunnen we veel deelruimten maken. Zo zijn er bij de functies van het interval $[a, b]$ naar $\mathbb{L}$ in ieder geval de volgende deelruimten:

- de ruimte van alle reële/complexe polynomen van de graad tenhoogste n op $[a, b]$, notatie $\mathbb{P}ol_n([a, b], \mathbb{R})$ of/en $\mathbb{P}ol_n([a, b], \mathbb{C})$, (Pas op: als $p \in \mathbb{P}ol_n([a, b], \mathbb{C})$, zeg $p(x) = a_0 + a_1x + \cdots a_nx^n$, dan is x een reële variabele, want $x \in [a, b]$. Echter, de coëfficienten a_i (en dus ook $p(x)$) zijn uit $\mathbb{C}$.)
- De ruimte van alle reële polynomen $\mathbb{P}ol([a, b], \mathbb{R})$ en de ruimte van alle complexe polynomen $\mathbb{P}ol([a, b], \mathbb{C})$.
- de ruimte van alle continue functies op $[a, b]$ naar $\mathbb{L}$, notatie: $C([a, b], \mathbb{R})$ of/en $C([a, b], \mathbb{C})$.
- de ruimte van alle differentieerbare functies op $[a, b]$, notatie: $D([a, b], \mathbb{R})$ of/en $D([a, b], \mathbb{C})$.
- de ruimte van alle continu-differentieerbare functies op $[a, b]$, dat wil zeggen, differentieerbare functies met continue afgeleide, notatie: $C^1([a, b], \mathbb{R})$ en/of $C^1([a, b], \mathbb{C})$.
- de ruimte van n-maal continu-differentieerbare functies op $[a, b]$, dat wil zeggen, die functies f die n-maal differentieerbaar zijn en waarvan de n^{de} afgeleide $f^{(n)}$ continu is, notatie: $C^n([a, b], \mathbb{R})$ en/of $C^n([a, b], \mathbb{C})$.

In het bijzonder geldt dat

$$\mathbb{P}ol_n([a, b], \mathbb{R}) \subset \mathbb{P}ol([a, b], \mathbb{R}) \subset C^n([a, b], \mathbb{R}) \subset C^1([a, b], \mathbb{R}) \subset D([a, b], \mathbb{R}) \subset C([a, b], \mathbb{R}) \subset \mathcal{F}([a, b], \mathbb{R})$$
$$\mathbb{P}ol_n([a, b], \mathbb{C}) \subset \mathbb{P}ol([a, b], \mathbb{C}) \subset C^n([a, b], \mathbb{C}) \subset C^1([a, b], \mathbb{C}) \subset D([a, b], \mathbb{C}) \subset C([a, b], \mathbb{C}) \subset \mathcal{F}([a, b], \mathbb{C}).$$

Dat het werkelijk deelruimten van $\mathcal{F}([a, b], \mathbb{C})$ zijn, volgt uit eigenschappen van continue, differentieerbare en continu-differentieerbare functies, die bekend zijn uit de analyse. De nulfunctie van $[a, b]$ naar $\mathbb{C}$ is continu, differentieerbaar, enzovoorts en behoort dus tot alle genoemde verzamelingen. Verder geldt voor elk tweetal continue functies dat de som ook weer continu is en dat het product van een continue functie met een (complex) getal ook weer continu is. Voor $C([a, b], \mathbb{C})$ is dus voldaan aan de tweede eis en de derde eis van deelruimte. Hetzelfde geldt voor de andere verzamelingen.

Deelruimten van $L_2([a,b],\mathbb{R})$ en/of $L_2([a,b],\mathbb{C})$

De ruimte die bestaat uit alle stuksgewijs continue functies op $[a,b]$,
notatie: $PC([a,b],\mathbb{R})$ (resp. $PC([a,b],\mathbb{C})$) is een lineaire deelruimte van $L_2([a,b],\mathbb{R})$ (resp. $L_2([a,b],\mathbb{C})$).

2.4 Lineaire (On)Afhankelijkheid

De begrippen lineaire afhankelijkheid en lineaire onafhankelijkheid zijn "de absolute basisbegrippen" uit de lineaire algebra. Dit willen we definiëren in een willekeurige vectorruimte.

Definitie: Laat V een vectorruimte zijn. Dan:

1. De lege verzameling $S = \emptyset$ zullen we een onafhankelijke deelverzameling van V noemen.
2. Een niet leeg eindig stelsel $S = \{\mathbf{v}_1, \dots, \mathbf{v}_p\} \subset V$ heet onafhankelijk als de vectorvergelijking
$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$
alleen de triviale oplossing $c_1 = 0, \dots, c_p = 0$ heeft.
3. Een oneindige verzameling $S = \{\mathbf{v}_i\}_{i\in I} \subset V$ heet onafhankelijk als iedere eindige deelverzameling van S onafhankelijk is.

Tenslotte, een deelverzameling die niet onafhankelijk is, zullen we afhankelijk noemen.

Er volgt direct dat een niet leeg eindig stelsel $S = \{\mathbf{v}_1, \dots, \mathbf{v}_p\} \subset V$ afhankelijk is, als de vectorvergelijking
$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$
een niet–triviale oplossing $c_1, \dots, c_p$ heeft. In het geval van afhankelijkheid wordt een relatie $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$ met niet alle c_i gelijk aan nul een **afhankelijkheidsrelatie** genoemd. Vóórdat we overgaan tot het geven van voorbeelden presenteren we eerst een zeer handig criterium waarmee lineaire (on)afhankelijkheid in één keer nagegaan kan worden. In de gangbare literatuur wordt deze stelling meestal als definitie van afhankelijkheid en onafhankelijkheid gebruikt, terwijl bovenstaande definitie dan een stelling is.

STELLING 2.19 (of definitie)**.** *Laat V een vectorruimte zijn. Dan geldt: Een stelsel vectoren $S = \{\mathbf{v}_i\}_{i\in I} \subset V$ is afhankelijk als tenminste één van de vectoren uit S een lineaire combinatie is van de overigen.*

Bewijs: Als S leeg is, dan is de verzameling onafhankelijk, dus de stelling is niet van toepassing. We bekijken nu het geval dat S eindig maar niet leeg is.
Als S uit één vector bestaat, zeg $S = \{\mathbf{v}\}$ dan is S afhankelijk als $c\mathbf{v} = \mathbf{0}$ een niet triviale oplossing heeft, dus als en slechts als $\mathbf{v} = \mathbf{0}$. Aan de andere kant, de uitspraak "tenminste één van de vectoren uit S een lineaire combinatie is van de overigen" voor deze S is equivalent met "$\mathbf{v}$ is een lineaire combinatie van de lege verzameling" dus evenzo equivalent met $\mathbf{v} = \mathbf{0}$. De conclusie volgt.

Nu het geval dat S eindig is maar tenminste uit twee vectoren bestaat. Zeg $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ met $p \geq 2$. Als tenminste één een lineaire combinatie van de andere vectoren, zeg $\mathbf{v}_k$. Dit betekent dat er getallen c_1, c_2,..., c_{k-1}, c_{k+1},..., c_p bestaan, zo dat

$$\mathbf{v}_k = c_1\mathbf{v}_1 + \cdots + c_{k-1}\mathbf{v}_{k-1} + c_{k+1}\mathbf{v}_{k+1} + \cdots + c_p\mathbf{v}_p,$$

hetgeen herschreven kan worden tot

$$c_1\mathbf{v}_1 + \cdots + c_{k-1}\mathbf{v}_{k-1} + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} + \cdots + c_p\mathbf{v}_p = \mathbf{0},$$

met $c_k = -1$. Blijkbaar heeft de vectorvergelijking

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

meer dan alleen de triviale oplossing ($x_1 = c_1$, $x_2 = c_2$, etc is immers een oplossing waarbij $c_k \neq 0$), en dus is S afhankelijk.
Nu andersom: neem aan dat de vectorvergelijking

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

een niet-triviale oplossing heeft. Als $c_k \neq 0$ dan halen we de andere vectoren (die er zijn) naar de andere kant en we krijgen

$$c_k\mathbf{v}_k = -c_1\mathbf{v}_1 - \cdots - c_{k-1}\mathbf{v}_{k-1} - c_{k+1}\mathbf{v}_{k+1} - \cdots - c_p\mathbf{v}_p,$$

dus (c_k is immers ongelijk aan nul)

$$\mathbf{v}_k = (-\frac{c_1}{c_k})\mathbf{v}_1 + \cdots + (-\frac{c_{k-1}}{c_k})\mathbf{v}_{k-1} + (-\frac{c_{k+1}}{c_k})\mathbf{v}_{k+1} + \cdots + (-\frac{c_p}{c_k})\mathbf{v}_p$$

en hier staat dat $\mathbf{v}_k$ een lineaire combinatie is van de overigen in S.
Tenslotte bekijken we het geval dat S oneindig is. S is afhankelijk als en slechts als een eindige deelverzameling van S afhankelijk is, en hiermee volgt alles uit het eindige geval. $\square$

Hoe gaan we na of een stel vectoren $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ (on)afhankelijk zijn? We moeten zien of de vectorvergelijking

$$c_1\mathbf{a}_1 + \cdots + c_k\mathbf{a}_k = \mathbf{0}$$

wel of niet een niet-trivale oplossing heeft.
In $\mathbb{R}^n$ en $\mathbb{C}^n$ reduceert dit, zoals ons bekend is, tot het oplossen van het stelsel vergelijkingen $[\mathbf{a}_1 \ldots \mathbf{a}_k|\mathbf{0}]$.
Ook in $\mathcal{M}_{n\times m}(\mathbb{R})$ of in $\mathcal{M}_{n\times m}(\mathbb{C})$ wordt dit probleem gereduceerd tot het oplossen van een een stelsel vergelijkingen.

Voorbeeld 2.20. Ga na of de matrices (vectoren) $A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ en $A_3 = \begin{bmatrix} 2 & 1 \\ -3 & -1 \end{bmatrix}$ in $\mathcal{M}_{2\times 2}(\mathbb{R})$ lineair onafhaneklijk zijn.
Antwoord. Stel:

$$c_1A_1 + c_2A_2 + c_3A_3 = \mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

En dus:

$$c_1\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + c_2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + c_3\begin{bmatrix} 2 & 1 \\ -3 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

wat leidt tot het stelsel: $\begin{cases} c_1 & + & c_2 & + & 2c_3 & = & 0 \\ c_1 & + & 2c_2 & + & c_3 & = & 0 \\ & & 3c_2 & - & 3c_3 & = & 0 \\ c_1 & + & 4c_2 & - & c_3 & = & 0 \end{cases} \sim \left[\begin{array}{ccc|c} 1 & 0 & 3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$ wat tot de niet triviale oplossing $c_1 = 3$, $c_2 = -1$ en $c_3 = -1$ leidt. Of te wel, A_1, A_2 en A_3 zijn afhankelijk en $3A_1 - A_2 - A_3 = \mathbf{0}$ is een afhankelijkheidsrelatie tussen A_1, A_2 en A_3.

Echter, voor functieruimten zullen we toch wel andere technieken moeten gebruiken. Veel lineaire afhankelijkheidsrelaties kunnen van formules bekend van de middelbare school ontdekt worden.

Voorbeeld 2.21. Is de verzameling functies (vectoren) $\{f, g, h\}$ met $f(x) = \sin(x)$, $g(x) = \cos(x)$ en $h(x) = \sin(1 + x)$ lineair afhankelijk of niet in de vectorruimte $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$.
Antwoord. Van de middelbare school kennen we de formule:

$$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)$$

In het bijzonder geldt:

$$h(x) = \sin(1 + x) = \sin(1)\cos(x) + \cos(1)\sin(x) = \sin(1)g(x) + \cos(1)f(x)$$

Of te wel, f, g en h zijn afhankelijk en $\cos(1)f + \sin(1)g - h = \mathbf{0}$ =nulfunctie is een afhankelijkheidsrelatie tussen f, g en h. Het zal duidelijk zijn dat als de formule voor $\sin(\alpha + \beta)$ niet bekend is, het zogoed als onbegonnen werk is om de afhankelijkheidsrelatie (en dus de afhankelijkheid) te ontdekken. Hoe zou men ooit de gewichten $\sin(1), \cos(1), -1$ kunnen vinden?

Voorbeeld 2.22. Is de verzameling functies (vectoren) $\{f, g, h\}$ met $f(x) = \sin(x)$, $g(x) = \cos(x)$ en $h(x) = \sin(2x)$ lineair afhankelijk of niet in de vectorruimte $V = C([0, 2\pi], \mathbb{R})$?
Antwoord. We kennen de regel $\sin(2x) = 2\sin(x)\cos(x)$, echter, dit is GEEN afhankelijkheidsrelatie tussen de functies f, g en h. Daarom: stel $c_1 f + c_2 g + c_3 h = \mathbf{0}$ =nulfunctie. Dan:

$$c_1 \sin(x) + c_2 \cos(x) + c_3 \sin(2x) = 0 \quad \text{voor iedere } x \in [0, 2\pi]$$

Als we achteréén de waarden $x = 0, \pi/4$ en $\pi/2$ uit $[0, 2\pi]$ invullen, verkrijgen we de vergelijkingen:

$$\begin{cases} 0c_1 & + & 1c_2 & + & 0c_3 & = & 0 \\ \frac{1}{2}\sqrt{2}c_1 & + & \frac{1}{2}\sqrt{2}c_2 & + & 1c_3 & = & 0 \\ 1c_1 & + & 0c_2 & + & 0c_3 & = & 0 \end{cases}$$

We concluderen dat $c_1 = c_2 = c_3 = 0$ en dus zijn de functies f, g en h lineair onafhankelijk.

Opmerking: Men kan de techniek van voorbeeld 2.22 kunnen toepassen op de functies van voorbeeld 2.21. Stel $c_1 f + c_2 g + c_3 h = \mathbf{0}$ =nulfunctie. Dan:

$$c_1 \sin(x) + c_2 \cos(x) + c_3 \sin(x + 1) = 0 \quad \text{voor iedere } x \in \mathbb{R}$$

We kunnen nu honderden x-waarden gaan invullen, en zo honderden vergelijkingen met slechts 3 onbekenden verkrijgen. Echter, altijd zal de in voorbeeld 2.21 gevonden niet triviale oplossing blijven bestaan. Bovendien zal de aldaar gevonden niet triviale oplossing bijna zeker niet gevonden worden.

Een ander probleem bij functieruimtes die we vanuit $\mathbb{R}^n$ niet kennen, is dat lineaire afhankelijkheid voor een stel gegeven functies kan afhangen van het domein waarop de functies bekeken worden.

Voorbeeld 2.23. Beschouw de functies f en g met $f(x) = x$ en $g(x) = x^2$. Oftewel, bekijk de functies x en x^2, en wel op het twee–punts domein $D = \{0, 1\}$. In de vectorruimte $\mathcal{F}(D, \mathbb{R})$ zijn de twee functies x en x^2 hetzelfde (want ze hebben dezelfde functiewaarde in ieder punt van het domein. Maar dan is $\{x, x^2\}$ een afhankelijke verzameling in $\mathcal{F}(\{0, 1\}, \mathbb{R})$.

Opmerking: In verband met het vorige voorbeeld het volgende. In opgave 5 wordt aangetoond dat als het domein D tenminste $n+1$ punten bevat, dat dan de eindige verzameling functies $\{1, x, \ldots, x^n\}$ onafhankelijk is in de vectorruimte $\mathcal{F}(D, \mathbb{R})$. Merk op dat dit impliceert dat de oneindige verzameling

$$\{1, x, \ldots, x^n, \ldots\ldots\}$$

onafhankelijk is in $\mathcal{F}([a,b], \mathbb{R})$, voor ieder interval $[a,b]$, met $a < b$.

Er bestaat nog een methode om de lineaire onafhankelijkheid van functies te bepalen, een methode gevonden door Wronski[4].
Als $f_1, \ldots, f_k$ een k-tal functies is dat $k-1$ maal gedifferentiëerd kan worden op een domein $[a,b]$ (of heel $\mathbb{R}$) dan heet de determinant

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \ldots & f_k(x) \\ f_1'(x) & f_2'(x) & \ldots & f_k'(x) \\ \vdots & \vdots & \ldots & \vdots \\ f_1^{(k-1)}(x) & f_2^{(k-1)}(x) & \ldots & f_k^{(k-1)}(x) \end{vmatrix}$$

de **Wronskiaan** van de functies $f_1, \ldots, f_k$, en wordt soms ook genoteerd met $W[f_1, \ldots, f_k](x)$.
Merk op dat als de functies $f_1, \ldots, f_k$ afhankelijk zijn, zeg $c_1 f_1 + c_2 f_2 + \ldots + c_k f_k = 0$ (niet alle c_i nul) dan kan dit $(k-1)$ maal gedifferentieerd worden, en we verkrijgen:

$$\begin{array}{lclclclcl} c_1 f_1(x) & + & c_2 f_2(x) & + & \ldots & + & c_k f_k(x) & = & 0 \\ c_1 f_1'(x) & + & c_2 f_2'(x) & + & \ldots & + & c_k f_k'(x) & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ c_1 f_1^{(k-1)}(x) & + & c_2 f_2^{(k-1)}(x) & + & \ldots & + & c_k f_k^{(k-1)}(x) & = & 0 \end{array}$$

en dus

$$c_1 \begin{bmatrix} f_1(x) \\ f_1'(x) \\ \vdots \\ f_1^{(k-1)}(x) \end{bmatrix} + c_2 \begin{bmatrix} f_2(x) \\ f_2'(x) \\ \vdots \\ f_2^{(k-1)}(x) \end{bmatrix} + \cdots + c_k \begin{bmatrix} f_k(x) \\ f_k'(x) \\ \vdots \\ f_k^{(k-1)}(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

(met niet alle $c_i = 0$) en we concluderen dat $W[f_1, \ldots, f_k](x) = 0$, voor iedere x. In contrapositief gesteld hebben we de volgende stelling bewezen.

STELLING 2.24. *(**Wronski's test**) Als er een x bestaat met Wronskiaan $W[f_1, \ldots, f_k](x) \neq 0$, dan zijn de functies $f_1, \ldots, f_k$ lineair onafhankelijk.*

Voorbeeld 2.25. Ga na of de drie functies f, g, h met $f(x) = 1$, $g(x) = e^x$ en $h(x) = e^{2x}$ onafhankelijk zijn in $\mathcal{F}(\mathbb{R}, \mathbb{R})$.
Antwoord. Omdat

$$W[f,g,h](x) = \begin{vmatrix} 1 & e^x & e^{2x} \\ 0 & e^x & 2e^{2x} \\ 0 & e^x & 4e^{2x} \end{vmatrix} = 2e^{3x} \neq 0 \text{ (ga na.)}$$

volgt dat deze drie functies onafhankelijk zijn.

[4]Jozef Hoëné de Wronski (1778-1853) Deze Pool/Fransman heeft een leven vol conflicten en controversies uitgevochten over zijn werk. Veel van zijn werk is totaal vergeten en hij is overleden in grote armoede.

Jammer genoeg geldt de omkering van Stelling 2.24 NIET. Als $W[f_1, \ldots, f_k](x) = 0$, voor alle x, dan zijn de functies $f_1, \ldots, f_k$ niet noodzakelijk lineair afhankelijk. Bij de opgaven kan men een voorbeeld tegenkomen. Dit voorbeeld heeft een ietwat "geforceerd karakter", en dat is geen toeval want men kan aantonen dat voor "een k-tal natuurlijke functies" de omkering wel geldt. Wat "natuurlijke functies" zijn vertellen we niet, dit zou te ver voeren.

2.5 Basis, Coördinaten en dimensie

Basis

Definitie: Een verzameling vectoren B in een vectorruimte V heet een **basis** voor V als het stelsel V opspant (dus $\operatorname{Span} B = V$) en het stelsel B bovendien onafhankelijk is.

Let op! Deelruimten van vectorruimten zijn ook vectorruimten. In bovenstaande definitie hoeft V dus geen standaardvectorruimte te zijn, maar kan het ook een deelruimte van zo'n vectorruimte zijn.

Opmerking: We zullen een eindige basis algemeen noteren met $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$. Merk op dat deze notatie een **ordening** geeft op de basisvectoren uit B. Zo'n basis zullen we "geordende basis" noemen. Aangezien we meestal geïnteresseerd zijn in een geordende basis, zullen we meestal bij iedere eindige basis de ordening meteen vastleggen en dit noteren met $B = (\mathbf{b}_1, \ldots, \mathbf{b}_k)$. Als u deze opmerking niet begrijpt, bekijk dan goed Voorbeeld 2.30 waar een natuurlijke basis op $\mathcal{M}_{m\times n}(\mathbb{R})$ gedefiniëerd wordt zonder natuurlijke ordening. De reden voor de ordenig zal blijken bij de coördinatisering, daarvoor is een ordenig nodig.

Definitie: Een vectorruimte V heet **eindig dimensionaal** als V een **eindige** basis $B = (\mathbf{v}_1, \ldots, \mathbf{v}_k)$ heeft.

We zullen onze aandacht eerst beperken tot de eindig dimensionale vectorruimten.

Voorbeelden van bases

Voorbeeld 2.26. Beschouw de kleinst mogelijke vectorruimte $V = \{\mathbf{0}\}$. Merk op dat de verzameling $\{\mathbf{0}\}$ geen basis is voor V, want deze verzameling is afhankelijk. Echter, de lege verzameling $S = \emptyset$ is wel een basis, want we hebben afgesproken dat de lege verzameling onafhankelijk is en dat $\operatorname{Span} \emptyset = \{\mathbf{0}\} = V$.

Voorbeeld 2.27. Net als in $\mathbb{R}^n$ is ook in $\mathbb{C}^n$ het stelsel

$$\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \right) = (\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n)$$

een (geordende) basis, want het is onafhankelijk en elke vector in $\mathbb{C}^n$ kan als lineaire combinatie van deze vectoren worden geschreven. Deze basis heet de **natuurlijke basis** voor $\mathbb{C}^n$, of ook de **standaardbasis** of **kanonieke basis**. Dus: $\mathbb{R}^n = \mathrm{Span}_{\mathbb{R}}\{\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n\}$ en $\mathbb{C}^n = \mathrm{Span}_{\mathbb{C}}\{\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n\}$.

Voorbeeld 2.28. Zowel in $\mathbb{P}ol_n([a,b], \mathbb{R})$ als in $\mathbb{P}ol_n([a,b], \mathbb{C})$ is het stelsel

$$(1,\ x,\ x^2\ , \ldots ,\ x^n)$$

een (geordende) basis. We hebben al in de opgaves gezien dat dit stelsel onafhankelijk is. Verder is het duidelijk dat elke veelterm met graad $\leq n$ te schrijven is als een lineaire combinatie van deze vectoren, zodat $\mathbb{P}ol_n([a,b], \mathbb{R}) = \mathrm{Span}_{\mathbb{R}}(1,\ x,\ x^2\ , \ldots ,\ x^n)$ en $\mathbb{P}ol_n([a,b], \mathbb{C}) = \mathrm{Span}_{\mathbb{C}}(1,\ x,\ x^2\ , \ldots ,\ x^n)$. Deze basis wordt de **natuurlijke basis** voor $\mathbb{P}ol_n([a,b], \mathbb{R})$ (of $\mathbb{P}ol_n([a,b], \mathbb{C})$) genoemd. Ook de termen **standaardbasis** en **kanonieke basis** worden gebruikt.

Voorbeeld 2.29. In voorbeeld 2.22 hebben we gezien dat het stelsel functies

$$(\sin x, \cos x, \sin 2x)$$

in $\mathcal{F}(\mathbb{R}, \mathbb{C})$ onafhankelijk is.
Daaruit volgt dan direct dat het een basis is voor $V = \mathrm{Span}(\sin x, \cos x, \sin 2x)$. Dat het stelsel de ruimte V opspant is dan namelijk door de definitie van V waar.

Voorbeeld 2.30. Een basis voor $M_{m\times n}(\mathbb{R})$ (of $M_{m\times n}(\mathbb{C})$) is bijvoorbeeld het stelsel $\mathcal{E}$ dat bestaat uit de matrices waarvan één kental gelijk is aan 1, terwijl de andere kentallen nul zijn. Ga na dat deze verzameling vectoren onafhankelijk is en $M_{m\times n}(\mathbb{R})$ (of $M_{m\times n}(\mathbb{C})$) opspant.
Echter, om een geordende basis te zijn is ook een ordening op deze basis nodig. Als $E_{i,j}$ die matrix in $M_{m\times n}(\mathbb{R})$ (of $M_{m\times n}(\mathbb{C})$) is met $(i,j)^{\mathrm{de}}$ kental 1 en andere kentallen 0, dan spreken we de volgende ordening af:

$$\mathcal{E} = (E_{1,1}, E_{1,2}, \ldots, E_{1,n}, E_{2,1}, E_{2,2}, \ldots, E_{2,n}, E_{3,1} \ldots, E_{m,1}, \ldots, E_{m,n})$$

Ook de basis $\mathcal{E}$ wordt de natuurlijke basis, de standaardbasis en/of de kanonieke basis genoemd.
De deelruimte $\mathrm{Symm}_n(\mathbb{R})$ (of $\mathrm{Symm}_n(\mathbb{C})$) van *symmetrische* $n \times n$ matrices (d.w.z. $A^T = A$) heeft als basis

$$\left(\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \right)$$

(dit zijn de matrices waarbij de kentallen a_{ij} en a_{ji} voor één paar (i,j) gelijk zijn aan 1, terwijl de rest gelijk is aan 0). U mag zelf een ordening bepalen.

Bij bovenstaande voorbeelden lag het voor de hand hoe een basis kon worden gekozen. Het ligt niet altijd zo voor de hand. Met name niet als de vectorruimte geen eindige basis heeft. Verder is

het zo dat als er een basis is, deze niet uniek is. Sterker nog: als een vectorruimte een (niet lege) basis heeft dan zijn er zelfs oneindig veel.
Voor deelruimten die als lineair omhulsel van een eindig stel vectoren gegeven zijn hebben we een recept om bases te vinden: verwijder net zo lang vectoren die combinaties zijn van de overigen, tot een onafhankelijk stelsel overblijft. Dit kan echter een vervelend karwei zijn. In het eerste deel van lineaire algebra hebben we gezien hoe het vinden van een basis handig kan voor deelruimten van $\mathbb{R}^n$, als deze als lineair omhulsel (kolomruimte, rijruimte) of als nulruimte gegeven zijn. Deze methoden gelden ook als we in $\mathbb{C}^n$ werken.

Coördinaten t.o.v. een basis

Eén van de redenen waarom bases zo handig zijn is de volgende stelling

STELLING 2.31. *Een stelsel vectoren* $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ *(mag oneindig groot zijn) is een basis voor een vectorruimte* V *(reëel of complex) als en slechts als elke vector uit* V *op precies één manier te schrijven is als een lineaire combinatie van de vectoren uit het stelsel.*

Dit leidt tot:

Definitie: Stel dat V een eindig dimensionale vectorruimte is (over $\mathbb{R}$ of $\mathbb{C}$), zeg met basis $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$. Uit bovenstaande stelling volgt dat bij iedere vector $\mathbf{v} \in V$ een uniek stel scalairen c_1, c_2,..., c_n in $\mathbb{R}$ of $\mathbb{C}$ bestaat zo dat

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n.$$

Deze getallen worden de **coördinaten van v ten opzichte van de basis** B genoemd. Bij $\mathbf{v}$ maken we dan een vector in $\mathbb{R}^n$ of $\mathbb{C}^n$ door de getallen c_1 tot en met c_n in een kolom te zetten. Deze kolomvector noteren we als $[\mathbf{v}]_B$, dus

$$[\mathbf{v}]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

en noemen we de **coördinatisering** van $\mathbf{v}$ ten opzichte van de basis B.

Bewijs: van Stelling 2.31 Ga er eerst van uit dat $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ een basis voor V is. Omdat de basisvectoren V opspannen is elke vector uit V op *tenminste* één manier als lineaire combinatie van de basisvectoren te schrijven.
Stel nu dat voor een vector $\mathbf{v}$ uit V geldt dat

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_p\mathbf{v}_p$$

en ook

$$\mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_p\mathbf{v}_p.$$

Dan volgt

$$\mathbf{0} = \mathbf{v} - \mathbf{v} = (a_1 - b_1)\mathbf{v}_1 + \cdots + (a_p - b_p)\mathbf{v}_p.$$

Uit de onafhankelijkheid van $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ volgt dan direct dat $a_i - b_i = 0$ en dus dat $a_i = b_i$, voor alle i. Maar dat wil zeggen dat de twee lineaire combinaties voor $\mathbf{v}$ hetzelfde zijn en $\mathbf{v}$ dus op één manier te schrijven is als lineaire combinatie van de basisvectoren.
Nu andersom, neem dus aan dat elke vector uit V op precies één manier te schrijven is als een lineaire combinatie van de vectoren uit het stelsel. Duidelijk is dat het stelsel dan V opspant! Nu moeten we nog inzien dat het ook onafhankelijk is. Neem daartoe een combinatie van $\mathbf{v}_1, \ldots, \mathbf{v}_p$ die $\mathbf{0}$ oplevert. Volgens het gegeven is er maar één zo'n combinatie mogelijk. Dat moet dus $0\mathbf{v}_1 + \cdots + 0\mathbf{v}_p$ zijn, waaruit onafhankelijkheid volgt. □

Voorbeeld 2.32. In $\mathbb{P}ol_3([0, 1], \mathbb{R})$ met de standaardbasis $B = (1, x, x^2, x^3)$ zijn de coördinaten van het polynoom $p(x) = -3 + x - x^2 + 3x^3$ de getallen -3, 1, -1 en 3 (in deze volgorde!), wat genoteerd wordt met

$$[p]_B = \begin{bmatrix} -3 \\ 1 \\ -1 \\ 3 \end{bmatrix}.$$

Voor de geordende basis $C = (x^3, x^2, x, 1)$ geldt

$$[p]_C = \begin{bmatrix} 3 \\ -1 \\ 1 \\ -3 \end{bmatrix}.$$

Opmerking: Om de coördinatisering op een unieke manier te bepalen, moet de volgorde van het basisvectoren vastliggen. Dit is de reden geweest voor het invoeren van een stelsel vectoren (en daarmee een basis) als een *geordende* verzameling.

Voorbeeld 2.33. De coördinaten van $A = \begin{bmatrix} 1-i & 1+i \\ 2 & 4 \end{bmatrix} \in \mathcal{M}_{n\times n}(\mathbb{C})$ t.o.v. de standaardbasis $B = (E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$ (zie Voorbeeld 2.30) is: $[A]_B = \begin{bmatrix} 1-i \\ 1+i \\ 2 \\ 4 \end{bmatrix}$.

Het zal nu duidelijk zijn dat het gebruik van coördinaten ten op zichte van een gegeven basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ op de ruimte V, bij een vectorruimte over $\mathbb{R}$ een afbeelding $V \rightsquigarrow \mathbb{R}^n$ en bij een vectorruimte over $\mathbb{C}$ een afbeelding $V \rightsquigarrow \mathbb{C}^n$ geeft, in beide gevallen gedefinieerd door

$$\mathbf{x} \rightsquigarrow [\mathbf{x}]_B$$

(Deze afbeelding heeft, afgezien van de "coördinatisering" geen naam.) De volgende stellingen gaan over het gedrag van deze afbeelding t.o.v. de basisbegrippen.

STELLING 2.34. *V is een eindig dimensionale vectorruimte met basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Dan geldt voor iedere $\mathbf{u}, \mathbf{v} \in V$ en iedere scalar c:*

1. $[\mathbf{u} + \mathbf{v}]_B = [\mathbf{u}]_B + [\mathbf{v}]_B$
2. $[c\mathbf{u}]_B = c[\mathbf{u}]_B$

Bewijs: Merk op dat in de eerste te bewijzen formule, de eerste + staat voor de vectorieële optelling in de vectorrruimte V, terwijl de tweede + staat voor die in $\mathbb{R}^n$ (of $\mathbb{C}^n$).
Stel dat $[\mathbf{u}]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ en $[\mathbf{v}]_B = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$, dus stel dat $\mathbf{u} = c_1\mathbf{b}_1 + \ldots + c_n\mathbf{b}_n$ en $\mathbf{v} = d_1\mathbf{b}_1 + \ldots + d_n\mathbf{b}_n$. Er volgt dat

$$\mathbf{u} + \mathbf{v} = (c_1 + d_1)\mathbf{b}_1 + \ldots + (c_n + d_n)\mathbf{b}_n \quad \text{en} \quad c\mathbf{u} = cc_1\mathbf{b}_1 + \ldots + cc_n\mathbf{b}_n$$

en dus dat

$$[\mathbf{u} + \mathbf{v}]_B = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = [\mathbf{u}]_B + [\mathbf{v}]_B \quad \text{en} \quad [c\mathbf{u}]_B = \begin{bmatrix} cc_1 \\ \vdots \\ cc_n \end{bmatrix} = c[\mathbf{u}]_B$$

De conclusie volgt. □

STELLING 2.35. *V is een eindig dimensionale vectorruimte met basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Dan geldt voor een stel vectoren: $(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ in V :*

1. *$(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ zijn onafhankelijk in V als en slechts als de coördinaten $([\mathbf{u}_1]_B, \ldots, [\mathbf{u}_k]_B)$ onafhankelijk zijn in $\mathbb{R}^n$ (respectievelijk in $\mathbb{C}^n$).*
2. *De vectoren $(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ spannen V op als en slechts als de coördinaten $([\mathbf{u}_1]_B, \ldots, [\mathbf{u}_k]_B)$ de ruimte $\mathbb{R}^n$ (respectievelijk $\mathbb{C}^n$) opspannen.*
3. *$(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ is een basis van V als en slechts als de coördinaten $([\mathbf{u}_1]_B, \ldots, [\mathbf{u}_k]_B)$ een basis van $\mathbb{R}^n$ (respectievelijk $\mathbb{C}^n$) is.*

Bewijs: Onderdeel 1. is een opgave. We bewijzen 2. en dan is 3. een rechtstreeks gevolg van 1. en 2.
Stel dat de vectoren $(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ V opspannen. Kies $\mathbf{x} \in \mathbb{R}^n$ (respectievelijk, kies $\mathbf{x} \in \mathbb{C}^n$.) willekeurig. We tonen aan dat $\mathbf{x} \in \mathrm{Span}\{[\mathbf{u}_1]_B, \ldots, [\mathbf{u}_k]_B\}$.
Alsvolgt, beschouw de vector $\mathbf{v} = x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n \in V$. Merk op dat $[\mathbf{v}]_B = \mathbf{x}$. Omdat $(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ V opspant, bestaan er scalairen $c_1, \ldots, c_n$ zodat $\mathbf{v} = c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k$ en daarom

$$\mathbf{x} = [\mathbf{v}]_B = [c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k]_B = c_1[\mathbf{u}_1]_B + \cdots + c_k[\mathbf{u}_k]_B$$

en dus: $\mathbf{x} \in \mathrm{Span}\{[\mathbf{u}_1]_B, \ldots, [\mathbf{u}_k]_B\}$, voor iedere $\mathbf{x}$.
Omgekeerd: stel dat de coördinaatvectoren $([\mathbf{u}_1]_B, \ldots, [\mathbf{u}_k]_B)$ de ruimte $\mathbb{R}^n$ (respectievelijk $\mathbb{C}^n$) opspannen. Kies $\mathbf{v} \in V$ willekeurig en beschouw $[\mathbf{v}]_B$. Er bestaan $c_1, \ldots, c_k$ zodat

$$[\mathbf{v}]_B = c_1[\mathbf{u}_1]_B + \cdots + c_k[\mathbf{u}_k]_B = [c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k]$$

en dus

$$\mathbf{v} = c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k \in \mathrm{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\},$$

voor iedere $\mathbf{v} \in V$, d.w.z. $\mathrm{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\} = V$. □

Deze stellingen maken het mogelijk om onze kennis over $\mathbb{R}^n$ (respectievelijk, over $\mathbb{C}^n$) in de reële (respectievelijk, complexe) vectorruimte V te stoppen. We krijgen als eerste resultaat:

GEVOLG 2.36. *Stel dat V een eindig dimensionale vectorruimte is met basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Dan geldt: iedere basis van V bestaat precies uit n vectoren.*

Deze stelling maakt het mogelijk om het begrip dimensie voor algemene vectorruimten in te voeren.

Definitie: Stel dat V een eindig dimensionale vectorruimte is met basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Dan zeggen we: de dimensie van V is n en noteren we dit met $\dim(V) = n$.
Als V geen eindige basis dan heet V oneindig dimensionaal en dit noteren we met: $\dim(V) = \infty$

Voorbeeld 2.37. Als we het aantal elementen in de standaardbasis tellen van de vectorruimten $\mathbb{P}ol_n([a, b], \mathbb{R})$, $\mathbb{P}ol_n([a, b], \mathbb{C})$, $\mathcal{M}_{m\times n}(\mathbb{R})$ en $\mathcal{M}_{m\times n}(\mathbb{C})$ dan verkrijgen we:

$$\boxed{\dim(\mathbb{P}ol_n([a, b], \mathbb{R})) = n + 1} \quad \text{en} \quad \boxed{\dim(\mathbb{P}ol_n([a, b], \mathbb{C})) = n + 1}$$

$$\boxed{\dim(\mathcal{M}_{m\times n}(\mathbb{R})) = mn} \quad \text{en} \quad \boxed{\dim(\mathcal{M}_{m\times n}(\mathbb{C})) = mn}.$$

Natuurlijk bestaat er veel meer kennis over $\mathbb{R}^n$ en $\mathbb{C}^n$ die we kunnen terughalen naar V, zoals de dimensiestelling, de uitbreidingsstelling en de uitduningsstelling.

STELLING 2.38. *Stel dat V een eindig dimensionale vectorruimte is met* $\dim(V) = n$. *Dan geldt:*

1. *Ieder n-tal onafhankelijke vectoren in V is een basis van V.*
2. *Ieder n-tal opspannende vectoren in V is een basis van V.*

□

De volgende stelling is ook waar in een oneindig dimensionale vectorruimte, het bewijs is dan niet zo simpel meer. In de opgaven mag U het eindig dimensionale geval bewijzen.

STELLING 2.39. *Stel dat V een eindig dimensionale vectorruimte is met* $\dim(V) = n$. *Dan geldt:*

1. *Ieder onafhankelijk aantal vectoren van V is uit te breiden tot een basis.*
2. *Ieder opspannend aantal vectoren van V is uit te dunnen tot een basis.*

□

Opmerking: Het bewijs laat ook zien hoe je een stel opspannende vectoren kan reduceren tot een basis. Stap over op de coördinaten en reduceer die via vegen tot een basis. Ga nu weer terug naar V.

STELLING 2.40. *Stel dat V een eindig dimensionale vectorruimte is met* $\dim(V) = n$. *Dan geldt:*

1. *Ieder $n + 1$ tal vectoren in V is afhankelijk.*
2. *Ieder onafhankelijk stelsel bestaat uit tenhoogste n vectoren.*

$\square$

Opmerking: Uit bovenstaande stelling volgt dat als een vectorruimte V oneindig veel onafhankelijke vectoren bevat, dan moet deze ruimte oneindig dimensionaal zijn.

Oneindige Dimensionale vectorruimten

We herhalen: de vectorruimte V is oneindig dimensionaal als er geen eindige basis is.

Voorbeeld 2.41. De ruimten $\mathcal{F}([a,b],\mathbb{R})$, $C([a,b],\mathbb{R})$ en $C^n([a,b],\mathbb{R})$ zijn voorbeelden van oneindig dimensionale vectorruimten over $\mathbb{R}$.
Evenzo zijn $\mathcal{F}([a,b],\mathbb{C})$, $C([a,b],\mathbb{C})$ en $C^n([a,b],\mathbb{C})$ voorbeelden van oneindig dimensionale complexe vectorruimten. Dat dit inderdaad zo is volgt uit het feit dat de verzameling

$$\{1, x, x^2, \ldots\ldots, x^n, \ldots\ldots\}$$

een oneindige lineair onafhankelijke verzameling is in ieder van deze zes vectorruimten.

Dat $C([a,b],\mathbb{R})$ oneindig dimensionaal is merken we dagelijks bij het vak Analyse. Niemand heeft daar een eindig lijstje met functies waarmee alle functies te maken zijn. Zelfs een oneindige lijst is niet te maken.
Hoewel oneindige dimensionale ruimten wel een basis hebben, deze worden dan vaak "**Hamelbasis**" genoemd, spelen deze in het oneindig dimensionale geval een minder prominente rol dan in het eindig dimensionale geval.

2.6 Coördinatentransformaties

Het gebeurt regelmatig dat voor verschillende doeleinden verschillende bases interessant zijn en op dat moment is het handig als we snel van de coördinaten ten opzichte van de ene basis kunnen overstappen naar de coördinaten ten opzichte van de andere basis.
Wanneer in een n-dimensionale vectorruimte V zowel een basis B als een basis C gegeven is, hebben we voor een vector $\mathbf{v}$ uit V twee coördinatiseringen: $[\mathbf{v}]_B$ en $[\mathbf{v}]_C$. Overgang van de ene naar de andere schrijfwijze wordt een **coördinatentransformatie** genoemd en blijkt weer te maken te hebben met een matrixvermenigvuldiging. De volgende stelling maakt dit precies.

STELLING 2.42. *Voor een n dimensionale vectorruimte V met bases B en C geldt dat de omzetting van $[\mathbf{v}]_B$ naar $[\mathbf{v}]_C$ (met $\mathbf{v} \in V$) berekend kan worden via de matrixvermenigvuldiging*

$$P\ [\mathbf{v}]_B = [\mathbf{v}]_C,$$

waarbij P de $n \times n$ matrix over is die als kolommen de coördinatisering van de basisvectoren uit B ten opzichte van de basis C heeft. Dus als $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ dan

$$P = \begin{bmatrix} [\mathbf{b}_1]_C & [\mathbf{b}_2]_C & \cdots & [\mathbf{b}_n]_C \end{bmatrix}.$$

Deze matrix wordt de **overgangsmatrix** *(of* **transformatiematrix***) van B naar C genoemd en wordt meestal met $\underset{C \leftarrow B}{P}$ genoteerd.*
Bovendien geldt: als A een $n \times n$ met de eigenschap dat

$$[\mathbf{v}]_C = A[\mathbf{v}]_B \quad \textit{voor alle} \quad \mathbf{v} \in V$$

dan geldt: $A = \underset{C \leftarrow B}{P}$.

Bewijs: Kies maar $\mathbf{v} \in V$ en schrijf deze als lineaire combinatie van de vectoren uit B:

$$\mathbf{v} = c_1\mathbf{b}_1 + c_1\mathbf{b}_2 + \cdots + c_n\mathbf{b}_n.$$

Uit de lineariteit van de coördinatisering volgt dan

$$[\mathbf{v}]_C = c_1[\mathbf{b}_1]_C + c_2[\mathbf{b}_2]_C + \cdots + c_n[\mathbf{b}_n]_C =$$

$$= \begin{bmatrix} [\mathbf{b}_1]_C & [\mathbf{b}_2]_C & \cdots & [\mathbf{b}_n]_C \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \underset{C \leftarrow B}{P}\ [\mathbf{v}]_B.$$

De uniciteit van de overgangsmatrix matrix volgt ook. Stel dat A een matrix is met $[\mathbf{v}]_C = A[\mathbf{v}]_B$, voor alle $\mathbf{v} \in V$. Kies $\mathbf{v} = \mathbf{b}_i$, voor $i = 1, \ldots, n$. Omdat $[\mathbf{b}_i]_B = \mathbf{e}_i$, voor $i = 1, \ldots, n$, volgt: $[\mathbf{b}_i]_C = A[\mathbf{b}_i]_B = A\mathbf{e}_i = i^{\mathrm{e}}$ kolom van A, voor $i = 1, \ldots, n$. Dus zien we dat A en $\underset{C \leftarrow B}{P}$ dezelfde i^{e} kolom heeft (voor $i = 1, \ldots, n$). Kortom: $A = \underset{C \leftarrow B}{P}$. □

De eigenschappen om te onthouden zijn dus:

$$\underset{C \leftarrow B}{P} = \begin{bmatrix} [\mathbf{b}_1]_C & [\mathbf{b}_2]_C & \cdots & [\mathbf{b}_n]_C \end{bmatrix} \quad \text{en} \quad [\mathbf{v}]_C = \underset{C \leftarrow B}{P}[\mathbf{v}]_B \quad (\text{voor alle } \mathbf{v} \in V).$$

Voorbeeld 2.43. Beschouw $\mathbb{C}^2$ met als eerste basis $B = \left(\begin{bmatrix} 1 \\ i \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right)$ en tweede basis $C = \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ i \end{bmatrix} \right)$. Om de overgangsmatrix $\underset{C \leftarrow B}{P}$ van de basis B naar de basis C te bepalen moeten de vectoren van B gecoördinatiseerd worden ten opzichte van C. De eerste kolom van $\underset{C \leftarrow B}{P}$ is immers de coördinaatvector van $\mathbf{b}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ ten opzichte van de basis C, welke gevonden wordt door de vectorvergelijking

$$\begin{bmatrix} 1 \\ i \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ i \end{bmatrix}$$

op te lossen. Vegen van de aangevulde matrix

$$\left[\begin{array}{rr|r} 1 & 0 & 1 \\ -1 & i & i \end{array}\right]$$

tot gereduceerde echelonvorm levert dan $x_1 = 1$ en $x_2 = 1 - i$, de kentallen van de eerste kolom van $\underset{C\leftarrow B}{P}$. Op dezelfde manier moeten we voor de tweede kolom de aangevulde matrix

$$\left[\begin{array}{rr|r} 1 & 0 & -1 \\ -1 & i & 0 \end{array}\right]$$

vegen, hetgeen de kentallen -1 en i oplevert. Blijkbaar is

$$\underset{C\leftarrow B}{P} = \begin{bmatrix} 1 & -1 \\ 1-i & i \end{bmatrix}.$$

We merken nog op dat de twee stelsels in één keer opgelost hadden kunnen worden:

$$\left[\begin{array}{rr|rr} 1 & 0 & 1 & -1 \\ -1 & i & i & 0 \end{array}\right] \sim \left[\begin{array}{rr|rr} 1 & 0 & 1 & -1 \\ 0 & 1 & 1-i & i \end{array}\right]$$

Er is geveegd totdat links de eenheidsmatrix staat. Rechts staan dan de twee kolommen van de gevraagde matrix $\underset{C\leftarrow B}{P}$. Als nu $\mathbf{v}$ een vector in $\mathbb{C}^2$ is met $[\mathbf{v}]_B = \begin{bmatrix} -2 \\ i \end{bmatrix}$ (dus $\mathbf{v} = \begin{bmatrix} -2-i \\ -2i \end{bmatrix}$), dan is $[\mathbf{v}]_C$ snel te vinden, immers

$$[\mathbf{v}]_C = \underset{C\leftarrow B}{P}[\mathbf{v}]_B = \begin{bmatrix} 1 & -1 \\ 1-i & i \end{bmatrix}\begin{bmatrix} -2 \\ i \end{bmatrix} = \begin{bmatrix} -2-i \\ -3+2i \end{bmatrix}.$$

De volgende stelling is handig.

STELLING 2.44. *Als B, C en D bases zijn voor een eindig dimensionale vectorruimte V, dan geldt*

1. $\underset{C\leftarrow B}{P}$ *is inverteerbaar en* $\underset{B\leftarrow C}{P} = \left(\underset{C\leftarrow B}{P}\right)^{-1}$.
2. $\underset{D\leftarrow B}{P} = \underset{D\leftarrow C}{P} \cdot \underset{C\leftarrow B}{P}$

Bewijs: 1. Omdat de vectoren $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ onafhankelijk zijn, zijn de coördinaten $\{[\mathbf{b}_1]_C, \ldots, [\mathbf{b}_n]_C\}$ dat ook en dus is de vierkante matrix $\underset{C\leftarrow B}{P}$ inverteerbaar.
Omdat $[\mathbf{v}]_C = \underset{C\leftarrow B}{P}[\mathbf{v}]_B$ voor alle $\mathbf{v} \in V$ geldt ook:

$$\left(\underset{C\leftarrow B}{P}\right)^{-1}[\mathbf{v}]_C = [\mathbf{v}]_B \quad \text{voor alle } \mathbf{v} \in V$$

en dus: $\underset{B\leftarrow C}{P} = \left(\underset{C\leftarrow B}{P}\right)^{-1}$.
2. We controleren dat de tweede matrix ook de eigenschap:

$$[\mathbf{v}]_D = \underset{D\leftarrow B}{P}[\mathbf{v}]_B \quad \text{heeft, voor alle } \mathbf{v} \in V,$$

de eigenschap die $\underset{D\leftarrow B}{P}$ karakteriseert.
Welnu: $(\underset{D\leftarrow C}{P}\underset{C\leftarrow B}{P})[\mathbf{v}]_B = \underset{D\leftarrow C}{P}(\underset{C\leftarrow B}{P}[\mathbf{v}]_B) = \underset{D\leftarrow C}{P}[\mathbf{v}]_C = [\mathbf{v}]_D$, voor alle $\mathbf{v} \in V$ en de conclusie volgt. □

Voorbeeld 2.45. We bepalen $\underset{C\leftarrow B}{P}$ nog een keer voor de situatie van Voorbeeld 2.43, maar nu door deze matrix uit te rekenen via de standaardbasis basis $E = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$. Het aardige hiervan is dat de overgangsmatrices $\underset{E\leftarrow B}{P}$ en $\underset{E\leftarrow C}{P}$ direct zijn op te schrijven omdat de vectoren uit B en C al gecoördinatiseerd gegeven zijn ten opzichte van de standaardbasis! Er geldt

$$\underset{E\leftarrow B}{P} = \begin{bmatrix} 1 & -1 \\ i & 0 \end{bmatrix} \quad \text{en} \quad \underset{E\leftarrow C}{P} = \begin{bmatrix} 1 & 0 \\ -1 & i \end{bmatrix}.$$

Met bovenstaande stelling volgt dan

$$\underset{C\leftarrow B}{P} = \underset{C\leftarrow E}{P}\underset{E\leftarrow B}{P} = \left(\underset{E\leftarrow C}{P}\right)^{-1}\underset{E\leftarrow B}{P} = \begin{bmatrix} 1 & 0 \\ -1 & i \end{bmatrix}^{-1}\begin{bmatrix} 1 & -1 \\ i & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1-i & i \end{bmatrix}.$$

Een handige truc die toepasbaar is met ruimtes die een standaardbasis E hebben. Dan kan $\underset{C\leftarrow B}{P}$ berekend worden als $\underset{C\leftarrow B}{P} = \left(\underset{E\leftarrow C}{P}\right)^{-1}\underset{E\leftarrow B}{P}$.
Nu horen we te weten dat als we het stelsel $A|B$ (met A een $n \times n$ inverteerbare matrix) vegen naar $I_n|?$, dat in de rechterkolom de matrix $A^{-1}B$ verschijnt, de oplossing van $AX = B$. Dus:

$$[A|B] \sim [I_n|A^{-1}B]$$

GEVOLG 2.46. *Als B, C geordende bases zijn voor een eindig dimensionale vectorruimte V met een geordende standaardbasis E, dan geldt: $\underset{C\leftarrow B}{P}$ is te berekenen door $[\underset{E\leftarrow C}{P}|\underset{E\leftarrow B}{P}]$ te vegen naar $[I_n|\underset{C\leftarrow B}{P}]$.*

Voorbeeld 2.47. We bepalen $\underset{C\leftarrow B}{P}$ nog een keer voor de situatie van Voorbeeld 2.43, maar nu door $[\underset{E\leftarrow C}{P}|\underset{E\leftarrow B}{P}]$ te vegen naar $[I_2|\underset{C\leftarrow B}{P}]$. Welnu: even vegen en we zien:

$$[\underset{E\leftarrow C}{P}|\underset{E\leftarrow B}{P}] = \left[\begin{array}{rr|rr} 1 & 0 & 1 & -1 \\ -1 & i & i & 0 \end{array}\right] \sim \left[\begin{array}{rr|rr} 1 & 0 & 1 & -1 \\ 0 & 1 & 1-i & i \end{array}\right]$$

en dus $\underset{C\leftarrow B}{P} = \begin{bmatrix} 1 & -1 \\ 1-i & i \end{bmatrix}$.

Voorbeeld 2.48. Voor $\mathbb{P}_2([0,1],\mathbb{R})$ nemen we de twee geordende bases $C = (1, x-1, 1+x-x^2)$ en $B = (x-1, x^2+1, 2x^2-x)$. Om $\underset{C\leftarrow B}{P}$ te bepalen beschouwen we de standaardbasis $E = (1, x, x^2)$ en kennelijk

$$\underset{E\leftarrow C}{P} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{en} \quad \underset{E\leftarrow B}{P} = \begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

Om $\underset{C\leftarrow B}{P}$ te bepalen, vegen we

$$[\underset{E\leftarrow C}{P} | \underset{E\leftarrow B}{P}] = \begin{bmatrix} 1 & -1 & 1 & | & -1 & 1 & 0 \\ 0 & 1 & 1 & | & 1 & 0 & -1 \\ 0 & 0 & -1 & | & 0 & 1 & 2 \end{bmatrix}$$

Dit levert: $\begin{bmatrix} 1 & 0 & 0 & | & 0 & 3 & 3 \\ 0 & 1 & 0 & | & 1 & 1 & 1 \\ 0 & 0 & 1 & | & 0 & -1 & -2 \end{bmatrix}$ en dus $\underset{C\leftarrow B}{P} = \begin{bmatrix} 0 & 3 & 3 \\ 1 & 1 & 1 \\ 0 & -1 & -2 \end{bmatrix}$.

Hoofdstuk 3

Lineaire afbeeldingen

3.1 Afbeeldingen van $\mathbb{R}^n$ naar $\mathbb{R}^m$

We gaan er van uit dat de student bekend is met het **functiebegrip**, dus "een functie $f : A \to B$". Een functie wordt bepaald door een **domein** A, een zogeheten **codomein** B en een **functievoorschrift** $f(a)$, een voorschrift die aan iedere $a \in A$ een element $f(a) \in B$ toevoegt. We schrijven "de functie $f : A \to B$ met functievoorschrift $f(a)$". In het algemeen: f is de naam van de functie, $f(x)$ het functievoorschrift.

Zoals in de Analyse functies $f : \mathbb{R} \to \mathbb{R}$ bestudeerd worden (bijvoorbeeld met voorschrift $f(x) = \sin(2\ln(x^2 + 1))$), worden in de lineaire algebra functies (of ook **transformaties**) $T : \mathbb{R}^n \to \mathbb{R}^m$ bekeken, functies dus met domein $\mathbb{R}^n$ en codomein $\mathbb{R}^m$. Maar welke functies? Laten we ons realiseren dat we al functies bekeken hebben. Projecteren op een lineaire deelruimte W geeft een afbeelding $\text{proj}_W : \mathbb{R}^n \to \mathbb{R}^n$ en evenzo, spiegelen t.o.v. een lineaire deelrumte W geeft een afbeelding $S_W : \mathbb{R}^n \to \mathbb{R}^n$. beide konden beschreven worden met een matrix: $\text{proj}_W(\mathbf{x}) = [\text{proj}_W]\mathbf{x}$ en $S_W(\mathbf{x}) = [S_W]\mathbf{x}$. Hierbij staan $[\text{proj}_W]$ en $[S_W]$ voor de bijbehorende projectiematrix en spiegelingsmatrix.
Laten we ons nu realiseren dat iedere $m \times n$ matrix A een afbeelding $T_A : \mathbb{R}^n \to \mathbb{R}^m$ oplevert via het functievoorschrift:

$$T_A(\mathbf{x}) = A\mathbf{x}.$$

Dit soort afbeeldingen worden ook wel matrixtransformaties genoemd en dit blijken precies die functies te zijn (tussen $\mathbb{R}^n$ en $\mathbb{R}^m$) waarin men in de lineaire algebra geïnteresseerd is.

Voorbeeld 3.1. De matrix $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 2 & -1 \end{bmatrix}$ geeft een matrixtransformatie $T_A : \mathbb{R}^3 \to \mathbb{R}^2$. Deze heeft als functievoorschrift:

$$T_A\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 - 3x_3 \\ 2x_1 + 2x_2 - x_3 \end{bmatrix}$$

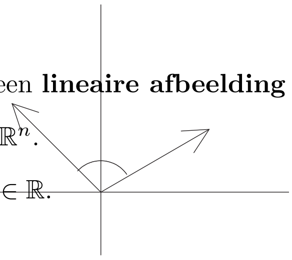
Aan het functievoorschrift van een functie is vaak in één opslag te zien of de betreffende functie een matrixtransformatie is.

Voorbeeld 3.2. Bekijk de functie $T : \mathbb{R}^2 \to \mathbb{R}^3$ met functievoorschrift $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 - 5x_2 \\ 3x_1 + 4x_2 \\ -x_1 + x_2 \end{bmatrix}$.
Aan dit functievoorschrift is inderdaad meteen te zien dat dit een matrixtransformatie betreft en

dat de bijbehorende matrix $A = \begin{bmatrix} 2 & -5 \\ 3 & 4 \\ -1 & 1 \end{bmatrix}$ is.

Daarentegen, de functie $T : \mathbb{R}^2 \to \mathbb{R}^2$ met functievoorschrift $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1^2 - 5x_2^3 \\ 3x_1 + x_2^2 \end{bmatrix}$ is gevoelsmatig toch geen matrixtransformatie. Is dat voldoende bewijs?

Als geen functievoorschrift gegeven is, maar een meetkundige beschrijving van de afbeelding wordt het nog lastiger.
Beschouw de afbeelding $R : \mathbb{R}^2 \to \mathbb{R}^2$ die iedere vector $\mathbf{x}$ over dezelfde hoek θ (linksom) in het vlak draait naar $R(\mathbf{x})$.
Is deze afbeelding een matrixtransformatie?
We zoeken een criterium om aan een afbeelding te kunnen zien of deze deze afbeelding een matrixtransformatie is.



Definitie: Een afbeelding $T : \mathbb{R}^n \to \mathbb{R}^m$ heet een **lineaire afbeelding** als:

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$, voor alle $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.
2. $T(c\mathbf{u}) = cT(\mathbf{u})$, voor alle $\mathbf{u} \in \mathbb{R}^n$ en alle $c \in \mathbb{R}$.

STELLING 3.3. *(i) Iedere matrixtransformatie $T : \mathbb{R}^n \to \mathbb{R}^m$ is een lineaire afbeelding.*
(ii) Stel dat $T : \mathbb{R}^n \to \mathbb{R}^m$ een lineaire afbeelding is. Dan is er een **unieke** *$m \times n$ matrix A zodat*

$$T(\mathbf{x}) = A\mathbf{x}, \quad \textit{voor alle } \mathbf{x} \in \mathbb{R}^n.$$

We noteren $A = [T]$ en er geldt:

$$A = [T] = [\,T(\mathbf{e_1}) \ldots T(\mathbf{e}_n)\,].$$

Bewijs: Beschouw de standaardbasis $(\,\mathbf{e}_1, \ldots, \mathbf{e}_n\,)$ van $\mathbb{R}^n$. Als $\mathbf{x} \in \mathbb{R}^n$ dan

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$$

en dus

$$T(\mathbf{x}) = T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) = x_1T(\mathbf{e}_1) + \cdots + x_nT(\mathbf{e}_n) = [\,T(\mathbf{e_1}) \ldots T(\mathbf{e}_n)\,]\mathbf{x}$$

De uniciteit: stel dat C een $m \times n$ matrix is zodat $T(\mathbf{x}) = C\mathbf{x}$ voor alle $\mathbf{x} \in \mathbb{R}^n$. Dan geldt:

$$\text{de } i^{\text{de}} \text{ kolom van } C = C\mathbf{e}_i = T(\mathbf{e}_i) = \text{de } i^{\text{de}} \text{ kolom van } [T]$$

voor $i = 1 \ldots n$ en dus $C = [T]$. $\square$

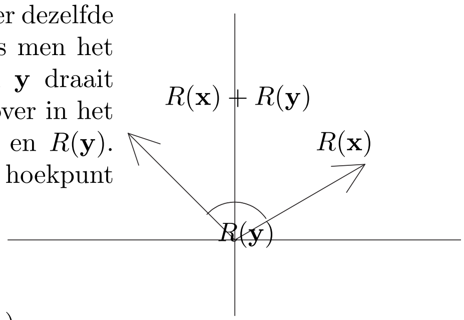
De matrix $A = [T]$ in Stelling 3.3 heet de **standaardmatrix** van T en de basiseigenschap van de standaardmatrix $[T]$ is:

$$T(\mathbf{x}) = [T]\mathbf{x}.$$

(Omdat de formule van $[T]$ gerelateerd is aan de standaardbasis, heet $[T]$ de standaardmatrix van T.) Deze stelling levert ook het antwoord op de vraag gesteld aan het begin van deze paragraaf. De functies $T : \mathbb{R}^n \to \mathbb{R}^m$ waar wij in geïnteresseerd zijn zien er uit als $T(\mathbf{x}) = A\mathbf{x}$.
We kunnen nu de vragen beantwoorden uit Voorbeeld 8.9.
Antwoord 1. De functie $T : \mathbb{R}^2 \to \mathbb{R}^2$ met functievoorschrift $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1^2 - 5x_2^3 \\ 3x_1 + x_2^2 \end{bmatrix}$ is geen matrixtransformatie, want het is geen lineaire afbeelding. Enig simpel rekenwerk laat zien dat $T(2\mathbf{x}) \neq 2T(\mathbf{x})$ als bijvoorbeeld $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
Antwoord 2. De afbeelding $R : \mathbb{R}^2 \to \mathbb{R}^2$ draait iedere vector over dezelfde hoek θ (linksom) in het vlak. Om in te zien dat R lineair is moeten we een meetkundig argument gebruiken. Neem twee vectoren $\mathbf{x}$ en $\mathbf{y}$ en we laten zien dat

$$R(\mathbf{x} + \mathbf{y}) = R(\mathbf{x}) + R(\mathbf{y}).$$

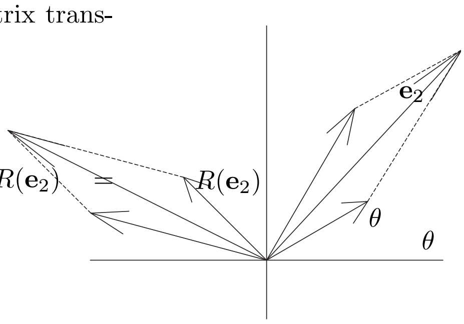
Beschouw in de schets de vectoren $\mathbf{x}$ en $\mathbf{y}$ en hun beelden $R(\mathbf{x})$ en $R(\mathbf{y})$ verkregen door over dezelfde hoek θ te draaien. Maar dan geldt: als men het parallellogram opgespannen door $\mathbf{x}$ en $\mathbf{y}$ draait over de hoek θ (linksom) dan gaat dit over in het parallellogram opgespannen door $R(\mathbf{x})$ en $R(\mathbf{y})$. Dus draait het hoekpunt $\mathbf{x} + \mathbf{y}$ naar het hoekpunt $R(\mathbf{x}) + R(\mathbf{y})$. Dus:



$$R(\mathbf{x} + \mathbf{y}) = R(\mathbf{x}) + R(\mathbf{y}).$$

Evenzo laat men zien dat $R(c\mathbf{x}) = cR(\mathbf{x})$.
We concluderen dat de roteren een matrix transformatie is. De standaardmatrix wordt:

$$[R] = [\, R(\mathbf{e}_1)\; R(\mathbf{e}_2)\,]$$

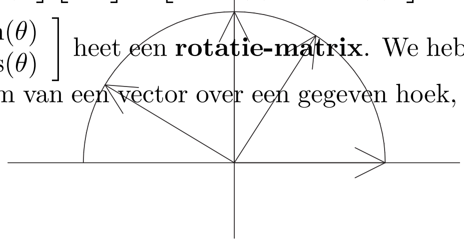


Omdat $R(\mathbf{e}_1) = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$ en $R(\mathbf{e}_2) = \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix}$ geldt:

$$[R] = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

En dus: $R\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos(\theta) - y\sin(\theta) \\ x\sin(\theta) + y\cos(\theta) \end{bmatrix}$.

Een matrix van de vorm $\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$ heet een **rotatie-matrix**. We hebben een expliciete formule gevonden voor de rotatie linksom van een vector over een gegeven hoek, en dus kunnen we



nu iedere vector roteren over welke hoek dan ook. Dit is de kracht van de standaardmatrix. Bijvoorbeeld: willen we de vector $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ over een hoek van 60 graden linksom roteren dan wordt het resultaat:

$$R(\mathbf{x}) = \begin{bmatrix} \cos(60) & -\sin(60) \\ \sin(60) & \cos(60) \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} - \sqrt{3} \\ \frac{1}{2}\sqrt{3} + 1 \end{bmatrix}.$$

De lineaire afbeeldingen waarvan de standaardmatrix een orthogonale matrix is vertonen een heel mooi gedrag, zoa;ls de volgende stelling laat zien.

STELLING 3.4. *Stel $T : \mathbb{R}^n \to \mathbb{R}^n$ is een lineaire afbeelding met standaardmatrix A, d.w.z. $T(\mathbf{x}) = A\mathbf{x}$, voor alle $\mathbf{x} \in \mathbb{R}^n$. Equivalent zijn:*

1. *A is een orthogonale matrix.*
2. *$\|T(\mathbf{x})\| = \|\mathbf{x}\|$, voor alle $\mathbf{x} \in \mathbb{R}^n$.*
3. *$\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$, voor alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.*

Bewijs: *1.⇒3.* Stel dat A orthogonaal is. Dan geldt $A^TA = I$ en dus: $\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \langle A\mathbf{x}, A\mathbf{y}\rangle = (A\mathbf{x})^T(A\mathbf{y}) = \mathbf{x}^TA^TA\mathbf{y} = \mathbf{x}^T\mathbf{y} = \langle \mathbf{x}, \mathbf{y}\rangle$.
3.⇒2. Als geldt: $\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$, voor alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ dan kunnen we deze uitspraak beperken tot de situatie $\mathbf{y} = \mathbf{x}$. We verkrijgen zo: $\langle T(\mathbf{x}), T(\mathbf{x})\rangle = \langle \mathbf{x}, \mathbf{x}\rangle$, voor alle $\mathbf{x} \in \mathbb{R}^n$ en dus $\|T(\mathbf{x})\|^2 = \|\mathbf{x}\|^2$, voor alle $\mathbf{x} \in \mathbb{R}^n$. De conclusie volgt.
2.⇒1. (Dit is het lastige deel van dit bewijs!) Stel dat $\|T(\mathbf{x})\| = \|\mathbf{x}\|$, voor alle $\mathbf{x} \in \mathbb{R}^n$ en bekijk de standaardmatrix $A = [\mathbf{a}_1 \dots \mathbf{a}_n]$ met $\mathbf{a}_i = T(\mathbf{e}_i)$. Dan: $\|\mathbf{a}_i\| = \|T(\mathbf{e}_i)\| = \|\mathbf{e}_i\| = 1$ voor $i = 1, \dots, n$. Fixeer nu $i \neq j$. Omdat $\|\mathbf{e}_i + \mathbf{e}_j\| = \sqrt{2}$ geldt: $\|\mathbf{a}_i + \mathbf{a}_j\| = \|T(\mathbf{e}_i + \mathbf{e}_j)\| = \|\mathbf{e}_i + \mathbf{e}_j\| = \sqrt{2}$ en dus: $2 = \|\mathbf{a}_i + \mathbf{a}_j\|^2 = \langle \mathbf{a}_i + \mathbf{a}_j, \mathbf{a}_i + \mathbf{a}_j\rangle = \|\mathbf{a}_i\|^2 + \|\mathbf{a}_j\|^2 + 2\langle \mathbf{a}_i, \mathbf{a}_j\rangle = 2 + 2\langle \mathbf{a}_i, \mathbf{a}_j\rangle$. En dus: $\langle \mathbf{a}_i, \mathbf{a}_j\rangle = 0$. Kortom, A is vierkant en de kolommen zijn orthonormaal, d.w.z. A is een orthogonale matrix. □

Hetzelfde verhaal kan men ophouden over functies $T : \mathbb{C}^n \to \mathbb{C}^m$. Om precies te zijn:

- Een (complexe) $m \times n$ matrix A geeft een afbeelding (matrixtransformatie) $T_A : \mathbb{C}^n \to \mathbb{C}^m$met het functievoorschrift:
$$T_A(\mathbf{x}) = A\mathbf{x}.$$
- Een afbeelding $T : \mathbb{C}^n \to \mathbb{C}^m$ heet een **lineaire afbeelding** als:
    1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$, voor alle $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$.
    2. $T(c\mathbf{u}) = cT(\mathbf{u})$, voor alle $\mathbf{u} \in \mathbb{C}^n$ en alle $c \in \mathbb{C}$.
- Stelling. De matrixtransformaties $T : \mathbb{C}^n \to \mathbb{C}^m$ zijn precies de lineaire afbeeldingen $L : \mathbb{C}^n \to \mathbb{C}^m$. Ook hier heet de unieke $m \times n$ matrix A zodat
$$T(\mathbf{x}) = A\mathbf{x}, \quad \text{voor alle } \mathbf{x} \in \mathbb{C}^n.$$
de standaardmatrix van T en wordt genoteerd met $[T]$. Er geldt weer:
$$A = [T] = [\,T(\mathbf{e_1}) \dots T(\mathbf{e}_n)\,].$$
Het bewijs is hetzelfde als in het reële geval.

- De volgende stelling geldt weer: als $T : \mathbb{C}^n \to \mathbb{C}^m$ een lineaire afbeelding is, dan zijn equivalent:

  1. $[T]$ is unitair.
  2. $\|T(\mathbf{x})\| = \|\mathbf{x}\|$, voor alle $\mathbf{x} \in \mathbb{C}^n$.
  3. $\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$, voor alle $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$.

  Hiervan is het bewijs van de stappen *1.⇒3.* en *3.⇒2.* letterlijk hetzelfde als in het reële geval. De stap *2.⇒1.* is weer de lastige stap.

  We bekijken weer de standaardmatrix $A = [\mathbf{a}_1 \ldots \mathbf{a}_n]$ met $\mathbf{a}_i = T(\mathbf{e}_i)$. Weer geldt: $\|\mathbf{a}_i\| = \|T(\mathbf{e}_i)\| = \|\mathbf{e}_i\| = 1$ voor $i = 1, \ldots, n$.

  Fixeer nu $i \neq j$. We bekijken weer $\mathbf{e}_i + \mathbf{e}_j$. Omdat $\|\mathbf{e}_i + \mathbf{e}_j\| = \sqrt{2}$ geldt: $\|\mathbf{a}_i + \mathbf{a}_j\| = \|T(\mathbf{e}_i + \mathbf{e}_j)\| = \|\mathbf{e}_i + \mathbf{e}_j\| = \sqrt{2}$ en dus: $2 = \|\mathbf{a}_i + \mathbf{a}_j\|^2 = \langle \mathbf{a}_i + \mathbf{a}_j, \mathbf{a}_i + \mathbf{a}_j\rangle = \|\mathbf{a}_i\|^2 + \|\mathbf{a}_j\|^2 + \langle \mathbf{a}_i, \mathbf{a}_j\rangle + \langle \mathbf{a}_j, \mathbf{a}_i\rangle = 2 + \langle \mathbf{a}_i, \mathbf{a}_j\rangle + \langle \mathbf{a}_j, \mathbf{a}_i\rangle$. Omdat we nu geen commutativiteit van het inwendig product hebben, kunnen we slechts concluderen dat $\langle \mathbf{a}_i, \mathbf{a}_j\rangle + \langle \mathbf{a}_j, \mathbf{a}_i\rangle = 0$. (a)
  Oftewel: $\langle \mathbf{a}_i, \mathbf{a}_j\rangle$ is zuiver imaginair.

  In het complexe geval bekijken we ook de vector $\mathbf{e}_i + i\mathbf{e}_j$. Omdat $\|\mathbf{e}_i + i\mathbf{e}_j\| = \sqrt{2}$ geldt: $\|\mathbf{a}_i + i\mathbf{a}_j\| = \|T(\mathbf{e}_i + i\mathbf{e}_j)\| = \|\mathbf{e}_i + i\mathbf{e}_j\| = \sqrt{2}$ en dus: $2 = \|\mathbf{a}_i + i\mathbf{a}_j\|^2 = \langle \mathbf{a}_i + i\mathbf{a}_j, \mathbf{a}_i + i\mathbf{a}_j\rangle = \|\mathbf{a}_i\|^2 + \|\mathbf{a}_j\|^2 + i\,\langle \mathbf{a}_i, \mathbf{a}_j\rangle - i\,\langle \mathbf{a}_j, \mathbf{a}_i\rangle = 2 + i\,\langle \mathbf{a}_i, \mathbf{a}_j\rangle - i\,\langle \mathbf{a}_j, \mathbf{a}_i\rangle$.
  We zien nu: $\langle \mathbf{a}_i, \mathbf{a}_j\rangle - \langle \mathbf{a}_j, \mathbf{a}_i\rangle = 0$. (b)

  Uit de beide gelijkheden (a) en (b) volgt nu ook dat $\langle \mathbf{a}_i, \mathbf{a}_j\rangle = 0$.

  Kortom, A is vierkant en de kolommen zijn orthonormaal, d.w.z. A is een unitaire matrix.

3.2 Lineaire afbeeldingen tussen vectorruimten

In het vorige college zijn functies tussen $\mathbb{R}^n$ en $\mathbb{R}^m$ bekeken, maar ook functies tussen $\mathbb{C}^n$ en $\mathbb{C}^m$. De voor de lineaire algebra geschikte afbeeldingen waren de matrix transformaties en deze afbeeldingen konden precies gekarakteriseerd worden als de de lineaire afbeeldingen (zie college 9). Nu willen we afbeeldingen tussen vectorruimten gaan bekijken, maar welke? Matrixtransformaties tussen functieruimten bestaan niet. Het ligt voor de hand om de karakterisering te gebruiken van deze afbeeldingen uit college 9.

Definitie: Een **lineaire afbeelding** $\mathcal{L} : V \to W$ tussen twee vectorruimten V en W over $\mathbb{R}$ (of **beide** over $\mathbb{C}$) is een afbeelding $\mathcal{L}$ van V naar W met de eigenschappen:

- $\mathcal{L}(\mathbf{u} + \mathbf{v}) = \mathcal{L}(\mathbf{u}) + \mathcal{L}(\mathbf{v})$, voor elke $\mathbf{u}$ en $\mathbf{v}$ uit V
- $\mathcal{L}(c\mathbf{v}) = c\mathcal{L}(\mathbf{v})$, voor elke $\mathbf{v}$ uit V en elke c uit $\mathbb{R}$ (dan wel $c \in \mathbb{C}$).

Opmerking: Een lineaire afbeelding kan bestaan tussen twee reële vectorruimten of twee complexe vectorruimten. Echter NIET van een reële naar een complexe vectorruimte en ook NIET van een complexe naar een reële vectorruimte.

Opmerking: Merk op dat het begrip matrixtransformatie in eerste instantie niet bestaat tussen twee vectorruimten.

Opmerking: Een *afbeelding* (of *functie* of *transformatie*) $\mathcal{L}$ van V naar W is een voorschrift dat aan elk element $\mathbf{v}$ van V precies één element van W koppelt. Dit element wordt genoteerd met $\mathcal{L}(\mathbf{v})$ en wordt het *beeld* van $\mathbf{v}$ onder $\mathcal{L}$ genoemd. Andersom, als $\mathbf{w}$ uit W het beeld is van een vector $\mathbf{v}$ uit V (dus $\mathbf{w} = \mathcal{L}(\mathbf{v})$), dan wordt $\mathbf{v}$ een *origineel* van $\mathbf{w}$ genoemd. Merk op dat $\mathbf{w}$ meerdere originelen kan hebben (of niet één).

Nu volgt een waslijst aan voorbeelden. Bestudeer deze intensief. Kijk of je overal begrijpt hoe de afbeelding gedefinieërd is (dat wil zeggen: snap van welke vectorruimte V naar welke vectorruimte W de afbeelding gaat en zie in dat het beeld van een vector uit V inderdaad een vector in W is) en waarom deze lineair is.

Voorbeeld 3.5. De **nulafbeelding** van een vectorruimte V naar een vectorruimte W is de afbeelding

$$O : V \to W \text{ gedefinieërd door } O(\mathbf{v}) = \mathbf{0}\ , \text{ voor alle } \mathbf{v} \in V.$$

Dit is een lineaire afbeelding, want

$$O(\mathbf{u} + \mathbf{v}) = \mathbf{0} = \mathbf{0} + \mathbf{0} = O(\mathbf{u}) + O(\mathbf{v})$$

en

$$O(c\mathbf{v}) = \mathbf{0} = c\mathbf{0} = cO(\mathbf{v}).$$

Ook de **identieke afbeelding** op V of **de identiteit** op V

$$\mathrm{Id} : V \to V \text{ gedefinieërd door } \mathrm{Id}(\mathbf{v}) = \mathbf{v}\ , \text{ voor alle } \mathbf{v} \in V$$

is een lineaire afbeelding, want $\mathrm{Id}(\mathbf{u} + \mathbf{v}) = \mathbf{u} + \mathbf{v} = \mathrm{Id}(\mathbf{u}) + \mathrm{Id}(\mathbf{v})$ en $\mathrm{Id}(c\mathbf{v}) = c\mathbf{v} = c\mathrm{Id}(\mathbf{v})$.

Voorbeeld 3.6. De afbeelding

$$\mathcal{L} : \mathcal{F}([-1, 1], \mathbb{R}) \to \mathbb{R} \text{ gedefinieërd door } \ \mathcal{L}(f) = f(0)$$

is een lineaire afbeelding. Immers:

$$\begin{aligned} \mathcal{L}(f + g) &= (f + g)(0) = f(0) + g(0) = \mathcal{L}(f) + \mathcal{L}(g) \\ \mathcal{L}(cf) &= (cf)(0) = cf(0) = c\mathcal{L}(f). \end{aligned}$$

voor alle f en g uit $\mathcal{F}([-1, 1], \mathbb{R})$ en c uit $\mathbb{R}$.

Vraag: Is de afbeelding $\mathcal{L} : \mathcal{F}([-1, 1], \mathbb{R}) \to \mathbb{R}$ gedefinieërd door $\mathcal{L}(f) = f(1)$ een lineaire afbeelding?

Voorbeeld 3.7. We definiëren nu twee schuif-afbeeldingen op $\mathbb{R}^\infty$. Eerst het "terugschuiven": dit is de afbeelding

$$\mathcal{L}_1 : \mathbb{R}^\infty \to \mathbb{R}^\infty \text{ gedefinieërd door } \mathcal{L}_1(x_0, x_1, \dots) = (x_1, x_2, \dots).$$

Deze afbeelding wordt "de shift" genoemd. Dat dit een lineaire afbeelding is, volgt uit

$$\begin{aligned} \mathcal{L}_1\big((x_0, x_1, \dots) + (y_0, y_1, \dots)\big) &= \mathcal{L}_1(x_0 + y_0, x_1 + y_1, \dots) \\ &= (x_1 + y_1, x_2 + y_2, \dots) \\ &= (x_1, x_2, \dots) + (y_1, y_2, \dots) \\ &= \mathcal{L}_1(x_0, x_1, \dots) + \mathcal{L}_1(y_0, y_1, \dots) \end{aligned}$$

en

$$\mathcal{L}_1\big(c(x_0, x_1, \dots)\big) = \mathcal{L}_1(cx_0, cx_1, \dots) = (cx_1, cx_2, \dots) = c(x_1, x_2, \dots) = c\mathcal{L}_1(x_0, x_1, \dots).$$

Het “heenschuiven” is de afbeelding

$$\mathcal{L}_2 : \mathbb{R}^\infty \to \mathbb{R}^\infty \text{ gedefinieërd door } \mathcal{L}_2(x_0, x_1, \dots) = (0, x_0, x_1, \dots).$$

Dit is eveneens een lineaire afbeelding (zie de opgaven).

Vraag: Is de afbeelding $\mathcal{L}_3 : \mathbb{R}^\infty \to \mathbb{R}^\infty$ gedefinieërd door $\mathcal{L}_3(x_0, x_1, \dots) = (1, x_0, x_1, \dots)$ een lineaire afbeelding?

De volgende afbeeldingen spelen een grote rol bij de theorie over differentiaal– en integraalvergelijkingen.

Voorbeeld 3.8. De eerste afbeelding is de afbeelding $\mathcal{D}$ die elke continu differentieerbare functie afbeeldt op zijn afgeleide:

$$\mathcal{D} : C^1([a,b],\mathbb{R}) \to C([a,b],\mathbb{R}) \text{ gedefinieërd door } \mathcal{D}(f) = f'.$$

Dat deze afbeelding lineair is volgt uit de bekende regels $(f+g)' = f' + g'$ en $(cf)' = cf'$.

Voorbeeld 3.9. De tweede afbeelding is de afbeelding $\mathcal{I}$ die elke continue functie f afbeeldt op die primitieve van f die nul is in a:

$$\mathcal{I} : C([a,b],\mathbb{R}) \to C^1([a,b],\mathbb{R}) \text{ gedefinieërd door } \mathcal{I}(f)(x) = \int_a^x f(t)dt.$$

Merk allereerst op dat deze afbeelding goed gedefinieërd is. Immers, de primitieve van een continue functie bestaat en is continu-differentieerbaar (dit is de hoofdstelling van de integraalrekening, bekend van het vak analyse). Verder is $\mathcal{I}$ een lineaire afbeelding, hetgeen volgt uit de rekenregels voor het primitiveren, bekend uit de analyse.

Voorbeeld 3.10. De eerste afbeelding is de afbeelding $\mathcal{D}$ die elke continu differentieerbare functie afbeeldt op de afgeleide in een vast punt $\mathbf{x}_0 \in [a,b]$:

$$\mathcal{D} : C^1([a,b],\mathbb{R}) \to \mathbb{R} \text{ gedefinieërd door } \mathcal{D}(f) = f'(\mathbf{x}_0).$$

Dat deze afbeelding lineair is volgt uit de bekende regels $(f+g)'(\mathbf{x}_0) = f'(\mathbf{x}_0)+g'(\mathbf{x}_0)$ en $(cf)'(\mathbf{x}_0) = cf'(\mathbf{x}_0)$.

Voorbeeld 3.11. De vierde afbeelding is de afbeelding $\mathcal{J}$ die elke continue functie f afbeeldt op het getal $\int_a^b f(t)dt$, de bepaalde integraal van f over $[a,b]$.

$$\mathcal{J} : C([a,b],\mathbb{R}) \to \mathbb{R} \text{ gedefinieërd door } \mathcal{I}(f) = \int_a^b f(x)dx.$$

Ook is $\mathcal{J}$ een lineaire afbeelding, hetgeen volgt uit de rekenregels voor de onbepaalde integraal, bekend uit de analyse.

Het volgende voorbeeld zullen we in het vervolg veelvuldig gebruiken en kan gedefinieërd bij elke eindig dimensionale vectorruimte.

Voorbeeld 3.12. In §2.5 hebben we bij elke vector uit een eindig dimensionale vectorruimte V over $\mathbb{R}$ (dan wel $\mathbb{C}$) met basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ de coördinatisering $[\mathbf{v}]_B$ gedefinieërd. Dit geeft aanleiding tot een afbeelding van V naar $\mathbb{R}^n$ (dan wel $\mathbb{C}^n$), die $\mathbf{v}$ koppelt aan $[\mathbf{v}]_B$, de zogeheten "coördinaatsafbeelding". Notatie

$$\mathbf{v} \mapsto [\mathbf{v}]_B.$$

We kunnen dit de afbeelding $[\,]_B : V \to \mathbb{R}^n$ ($[\,]_B : V \to \mathbb{C}^n$) noemen.
Uit Stelling 2.34 volgt dat deze coördinatisering een lineaire afbeelding is, want: $[\mathbf{u}+\mathbf{v}]_B = [\mathbf{u}]_B + [\mathbf{v}]_B$ en $[c\mathbf{u}]_B = c[\mathbf{u}]_B$.

3.3 Eigenschappen Lineaire afbeeldingen

Het wordt nu tijd om te laten zien dat er ook niet-lineaire afbeeldingen zijn.

Voorbeeld 3.13. De afbeelding $f : \mathbb{R} \to \mathbb{R}$ gedefinieërd door $f(x) = x + 1$ is niet lineair, want $f(2 \cdot 1) = f(2) = 3$ en $2f(1) = 2 \cdot 2 = 4 \neq f(2 \cdot 1)$.
De afbeelding $g : \mathbb{C} \to \mathbb{C}$ met $g(x) = ix$ is wel lineair: $g(x + y) = i(x + y) = ix + iy = g(x) + g(y)$ en $g(cx) = icx = cix = cg(x)$.
De afbeelding $h : \mathbb{R}^2 \to \mathbb{R}$ gedefinieërd door

$$h(x, y) = \begin{cases} x^2/y & \text{als } y \neq 0 \\ 0 & \text{als } y = 0 \end{cases}$$

is ook niet lineair. Er geldt wel $h(c(x, y)) = h(cx, cy) = (cx)^2/(cy) = c \cdot x^2/y = c \cdot h(x, y)$ (als $y \neq 0$), en $h(c(x, 0)) = h(cx, 0) = 0 = c \cdot 0 = c \cdot h(x, 0)$. Aan de andere eigenschap is echter niet voldaan: $h((1, 1) + (1, 0)) = h(2, 1) = 4$ en $h(1, 1) + h(1, 0) = 1 + 0 = 1 \neq 4$.

Het nagaan of een afbeelding lineair is, kan natuurlijk door de beide definieërende voorwaarden te controleren. Een andere mogelijkheid is inzien dat het een reeds bekende lineaire afbeelding is (bijvoorbeeld door een afbeelding van $\mathbb{R}^n$ naar $\mathbb{R}^m$ te schrijven als matrixtransformatie). Aantonen dat een afbeelding niet lineair is betekent in het algemeen een voorbeeld geven waaruit blijkt dat aan tenminste één van de twee definieërende voorwaarden niet voldaan is. Het volgende criterium kan in sommige gevallen ook uitkomst bieden.

STELLING 3.14. *Als voor een afbeelding $T : V \to W$ geldt dat $T(\mathbf{0}) \neq \mathbf{0}$ dan is T geen lineaire afbeelding (ofwel, als T lineair is, dan $T(\mathbf{0}) = \mathbf{0}$).*

Bewijs: Als T wel lineair zou zijn, dan

$$T(\mathbf{0}) = T(0\mathbf{0}) = 0T(\mathbf{0}) = \mathbf{0}$$

hetgeen in tegenspraak is met het gegeven. Overigens is bij dit bewijs alleen de tweede eigenschap van lineaire afbeeldingen gebruikt. □

Bij de afbeelding f uit het vorige voorbeeld hadden we deze stelling kunnen gebruiken, er geldt immers $f(0) = 1$ en f is dus niet lineair. Let op: draai de stelling niet om! Als het beeld van de nulvector wél de nulvector is, hoeft de afbeelding niet lineair te zijn! Dit geldt bijvoorbeeld voor de afbeelding h uit bovenstaand voorbeeld.

Voorbeeld 3.15. Een translatie $T : \mathbb{R}^2 \to \mathbb{R}^2$ over een niet-nulvector $\mathbf{a}$, dat wil zeggen $T(\mathbf{x}) = \mathbf{x} + \mathbf{a}$, is niet lineair, want $T(\mathbf{0}) = \mathbf{a} \neq \mathbf{0}$.

Andere eigenschappen van een lineaire afbeelding $\mathcal{L} : V \to W$ die we zullen gebruiken, zijn de volgende. Je mag ze zelf bewijzen in de opgaven.

- $\mathcal{L}(-\mathbf{x}) = -\mathcal{L}(\mathbf{x})$, voor alle $\mathbf{x} \in V$.
- Als $\mathcal{L} : V \to W$ lineair is, dan geldt $\mathcal{L}(\mathbf{x} - \mathbf{y}) = \mathcal{L}(\mathbf{x}) - \mathcal{L}(\mathbf{y})$, voor alle $\mathbf{x}, \mathbf{y} \in V$.
- $\mathcal{L} : V \to W$ is lineair als en slechts als

  $$\mathcal{L}(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k) = c_1\mathcal{L}(\mathbf{x}_1) + c_2\mathcal{L}(\mathbf{x}_2) + \cdots + c_k\mathcal{L}(\mathbf{x}_k)$$

  voor alle $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in V$ en $c_1, c_2, \ldots, c_k \in \mathbb{L}$.

- Als $\mathcal{L} : V \to W$ lineair is en de verzameling vectoren $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subset V$ is afhankelijk in V, dan is de verzameling vectoren $\{\mathcal{L}(\mathbf{x}_1), \mathcal{L}(\mathbf{x}_2), \ldots, \mathcal{L}(\mathbf{x}_k)\} \subset W$ afhankelijk in W.
- Als $\mathcal{L} : V \to W$ lineair is en de verzameling vectoren $\{\mathcal{L}(\mathbf{x}_1), \mathcal{L}(\mathbf{x}_2), \ldots, \mathcal{L}(\mathbf{x}_k)\} \subset W$ is onafhankelijk in W, dan is de verzameling vectoren $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subset V$ onafhankelijk in V.

We komen nu tot twee belangrijke verzamelingen die bij een lineaire afbeelding gedefinieërd kunnen worden, namelijk het beeld (te vergelijken met het bereik van een functie) en de kern (te vergelijken met de nulpunten van een functie). Deze twee verzamelingen leggen veel eigenschappen van de bijbehorende lineaire afbeelding vast.

Definitie: De **kern** van een lineaire afbeelding $\mathcal{L} : V \to W$ tussen vectorruimten V en W is de verzameling vectoren uit V die door $\mathcal{L}$ op $\mathbf{0}$ worden afgebeeld. De kern wordt genoteerd als $\mathrm{Ker}(\mathcal{L})$ en er geldt dus:

$$\mathrm{Ker}(\mathcal{L}) = \{\mathbf{v} \in V \mid \mathcal{L}(\mathbf{v}) = \mathbf{0}\}.$$

Het **beeld** of **bereik** van $\mathcal{L}$ is de verzameling van alle beelden onder $\mathcal{L}$ en wordt genoteerd als $\mathrm{Im}(\mathcal{L})$, als $\mathcal{L}[V]$ of als $\mathcal{L}(V)$. Er geldt dus

$$\mathrm{Im}(\mathcal{L}) = \{\mathcal{L}(\mathbf{v}) \mid \mathbf{v} \in V\} = \{\mathbf{w} \in W \mid \text{er is een } \mathbf{v} \in V \text{ met } \mathcal{L}(\mathbf{v}) = \mathbf{w}\}$$

Voorbeeld 3.16. De kern van de nulafbeelding $O : V \to W$ is heel V en het beeld is $\{\mathbf{0}\}$, de nulvector van W (en niet meer). Voor de identieke afbeelding $\mathrm{Id} : V \to V$ geldt $\mathrm{Ker}(\mathrm{Id}) = \{\mathbf{0}\}$ en $\mathrm{Im}(\mathrm{Id}) = V$.

Voorbeeld 3.17. De kern van de matrixtransformatie $\mathcal{L}_A : \mathbb{R}^m \to \mathbb{R}^n$ met $A \in M_{m\times n}(\mathbb{R})$ (dus $\mathcal{L}_A(\mathbf{x}) = A\mathbf{x}$ met $\mathbf{x} \in \mathbb{R}^m$) is precies de nulruimte van de matrix A en het beeld is precies de kolomruimte van de matrix A, dus

$$\mathrm{Ker}(\mathcal{L}_A) = \mathrm{Nul}(A) \text{ en } \mathrm{Im}(\mathcal{L}_A) = \mathrm{Col}(A).$$

Zo geldt voor de matrixtransformatie $\mathcal{L} : \mathbb{R}^3 \to \mathbb{R}^2$ uit voorbeeld 3.1 dat de kern de nulruimte van de matrix

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 2 & -1 \end{bmatrix}$$

is (hier $\mathrm{Span}\{(-16, 5, 2)^T\}$) en het beeld de kolomruimte (in dit geval heel $\mathbb{R}^2$, dus $\mathcal{L}$ is surjectief).

Voorbeeld 3.18. De kern van de schuif-afbeelding $\mathcal{L}_1$ uit voorbeeld 3.7 bestaat uit alle vectoren van de vorm $(x_0, 0, \ldots)$ (en is dus $\mathrm{Span}\{(1, 0, 0, \ldots)\}$) en het beeld is heel $\mathbb{R}^\infty$. Voor de schuif-afbeelding $\mathcal{L}_2$ geldt $\mathrm{Ker}(\mathcal{L}_2) = \{\mathbf{0}\}$ en het beeld bestaat uit alle vectoren van de vorm $(0, x_0, x_1, \ldots)$.
Voor de differentieer-afbeelding $\mathcal{D} : C^1([a,b], \mathbb{R}) \to C([a,b], \mathbb{R})$ uit voorbeeld 3.8 geldt dat de kern uit de functies bestaat waarvan de afgeleide de nulfunctie is, en dat zijn precies de constante functies (in dit geval van $[a,b]$ naar $\mathbb{R}$). Het beeld van $\mathcal{D}$ is heel $C([a,b], \mathbb{R})$. De hoofdstelling van de integraalrekening zegt immers dat elke continue functie $f : [a,b] \to \mathbb{R}$ een primitieve F heeft die zelfs continu-differentieerbaar is. Blijkbaar geldt $\mathcal{D}(F) = f$. Elke f is dus een beeld en behoort dus tot $\mathrm{Im}(\mathcal{D})$.
Dat de kern van de integraal-afbeelding $\mathcal{I} : C([a,b], \mathbb{R}) \to C^1([a,b], \mathbb{R})$ (uit voorbeeld 3.9) slechts uit de nulfunctie bestaat, volgt ook uit de hoofdstelling van de integraalrekening. Als f namelijk

tot de kern van $\mathcal{I}$ behoort, is $\mathcal{I}(f)$ de nulfunctie en dus is

$$\int_a^x f(t)\,dt = 0 \text{ , voor alle } x.$$

Als we het linkerlid differentiëren naar x vinden we met de genoemde hoofdstelling $f(x)$, terwijl het rechterlid 0 oplevert. Blijkbaar is f de nulfunctie. Voor het beeld van $\mathcal{I}$ zie de opgaven.

De kern en het beeld van een lineaire afbeelding $\mathcal{L} : V \to W$ zijn niet zomaar deelverzamelingen van de vectorruimten V en W, het zijn zelfs deelruimten. Dit is je misschien bij bovenstaande voorbeelden al opgevallen. Het wordt bewezen in de volgende stelling.

STELLING 3.19. *De kern* $\mathrm{Ker}(\mathcal{L})$ *van een lineaire afbeelding* $\mathcal{L} : V \to W$ *is een deelruimte van* V *en het beeld* $\mathrm{Im}(\mathcal{L})$ *is een deelruimte van* W*.*

Bewijs: De kern en het beeld bevatten beide de nulvector: $\mathbf{0}_V \in \mathrm{Ker}(\mathcal{L})$ en $\mathbf{0}_W \in \mathrm{Im}(\mathcal{L})$, want uit de lineariteit van $\mathcal{L}$ volgt dat $\mathcal{L}(\mathbf{0}_V) = \mathbf{0}_W$
Verder geldt voor de kern: als $\mathbf{v}_1$ en $\mathbf{v}_2$ in de kern zitten, dan

$$\mathcal{L}(\mathbf{v}_1 + \mathbf{v}_2) = \mathcal{L}(\mathbf{v}_1) + \mathcal{L}(\mathbf{v}_2) = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W.$$

en zit de som van $\mathbf{v}_1$ en $\mathbf{v}_2$ dus ook in de kern. Als verder $c \in \mathbb{R}$ (dan wel $c \in \mathbb{C}$) en $\mathbf{v} \in \mathrm{Ker}(\mathcal{L})$, dan volgt

$$\mathcal{L}(c\mathbf{v}) = c\mathcal{L}(\mathbf{v}) = c\mathbf{0}_W = \mathbf{0}_W.$$

en dus is ook $c\mathbf{v}$ element van de kern. Blijkbaar is de kern een deelruimte van V.
Voor het beeld geldt: als $\mathbf{w}_1$ en $\mathbf{w}_2$ in het beeld zitten, dan zijn er vectoren $\mathbf{v}_1$ en $\mathbf{v}_2$ in V zo dat $\mathbf{w}_1 = \mathcal{L}(\mathbf{v}_1)$ en $\mathbf{w}_2 = \mathcal{L}(\mathbf{v}_2)$. Er volgt:

$$\mathbf{w}_1 + \mathbf{w}_2 = \mathcal{L}(\mathbf{v}_1) + \mathcal{L}(\mathbf{v}_2) = \mathcal{L}(\mathbf{v}_1 + \mathbf{v}_2) \in \mathrm{Im}(\mathcal{L}).$$

Neem tenslotte $c \in \mathbb{R}$ (dan wel $c \in \mathbb{C}$) en $\mathbf{w} \in \mathrm{Im}(\mathcal{L})$. Dan is er weer een vector $\mathbf{v}$ in V zo dat $\mathbf{w} = \mathcal{L}(\mathbf{v})$ en dus

$$c\mathbf{w} = c\mathcal{L}(\mathbf{v}) = \mathcal{L}(c\mathbf{v}) \in \mathrm{Im}(\mathcal{L}),$$

waarmee ook $\mathrm{Im}(\mathcal{L})$ een deelruimte is, nu van W. □

De volgende stelling, die het verband tussen de dimensies van de kern en het beeld geeft, staat bekend als de *dimensiestelling*. Een gevolg van deze stelling is: hoe groter de kern, hoe kleiner het beeld.

STELLING 3.20 (dimensiestelling)**.** *Als* $\mathcal{L} : V \to W$ *een lineaire afbeelding van de* <u>*eindig dimensionale*</u> *vectorruimte* V *naar de vectorruimte* W *is, dan geldt*

$$\dim \mathrm{Ker}(\mathcal{L}) + \dim \mathrm{Im}(\mathcal{L}) = \dim V.$$

Bewijs: Als deelruimte van V is ook $\mathrm{Ker}(\mathcal{L})$ eindig-dimensionaal. Er is dus een basis $(\mathbf{u}_1, \ldots, \mathbf{u}_s)$ voor de kern van $\mathcal{L}$ (waarvan de dimensie dus s is). Deze is aan te vullen tot een basis

$$(\mathbf{u}_1, \ldots, \mathbf{u}_s, \mathbf{v}_1, \ldots, \mathbf{v}_r)$$

voor V (waarvan de dimensie dus $r+s$ is). Eventueel mag het aantal r of s nul zijn. De bewering is nu dat het stelsel $(\mathcal{L}(\mathbf{v}_1),\ldots,\mathcal{L}(\mathbf{v}_r))$ een basis vormt voor $\mathrm{Im}(\mathcal{L})$. In dat geval is bovenstaande dimensieformule duidelijk, want dan geldt $\dim V = r+s = \dim \mathrm{Im}(\mathcal{L}) + \dim \mathrm{Ker}(\mathcal{L})$. Om te laten zien dat het inderdaad een basis is voor $\mathrm{Im}(\mathcal{L})$, moet aangetoond worden dat het stelsel onafhankelijk is en dat het $\mathrm{Im}(\mathcal{L})$ opspant.
Onafhankelijkheid: Stel $c_1\mathcal{L}(\mathbf{v}_1)+\cdots+c_r\mathcal{L}(\mathbf{v}_r)=\mathbf{0}$. Uit de lineariteit volgt voor deze lineaire combinatie dat $\mathcal{L}(c_1\mathbf{v}_1+\cdots+c_r\mathbf{v}_r)=\mathbf{0}$ en is het dus een element van de kern van $\mathcal{L}$. Deze is dan te schrijven als lineaire combinatie van de vectoren uit de basis voor de kern:

$$c_1\mathbf{v}_1+\cdots+c_r\mathbf{v}_r = d_1\mathbf{u}_1+\cdots+d_s\mathbf{u}_s.$$

Dit kan herschreven worden tot

$$d_1\mathbf{u}_1+\cdots+d_s\mathbf{u}_s - c_1\mathbf{v}_1-\cdots-c_r\mathbf{v}_r=\mathbf{0}.$$

De vectoren uit deze relatie vormen echter een basis voor V en zijn dus onafhankelijk. Alle coëfficiënten moeten daarom nul zijn en er volgt dat $c_1=\cdots=c_r=0$. De vectoren $\mathcal{L}(\mathbf{v}_1),\ldots,\mathcal{L}(\mathbf{v}_r)$ zijn blijkbaar onafhankelijk.
Opspannen: Kies een $\mathbf{w}$ uit $\mathrm{Im}(\mathcal{L})$. Dan is $\mathbf{w}$ te schrijven als $\mathbf{w}=\mathcal{L}(\mathbf{v})$. De vector $\mathbf{v}$ kunnen we schrijven als lineaire combinatie van de basisvectoren van V:

$$\mathbf{v}=a_1\mathbf{u}_1+\cdots+a_s\mathbf{u}_s+b_1\mathbf{v}_1+\cdots+b_r\mathbf{v}_r.$$

De afbeelding $\mathcal{L}$ toepassen geeft:

$$\mathbf{w}=\mathcal{L}(\mathbf{v})=a_1\mathcal{L}(\mathbf{u}_1)+\cdots+a_s\mathcal{L}(\mathbf{u}_s)+b_1\mathcal{L}(\mathbf{v}_1)+\cdots+b_r\mathcal{L}(\mathbf{v}_r)=b_1\mathcal{L}(\mathbf{v}_1)+\cdots+b_r\mathcal{L}(\mathbf{v}_r).$$

omdat de $\mathbf{u}_i$ tot de kern van $\mathcal{L}$ behoren. Daarmee is w geschreven als lineaire combinatie van de vectoren $\mathcal{L}(\mathbf{v}_1),\ldots,\mathcal{L}(\mathbf{v}_r)$. □

Het komt nogal eens voor dat men geïnteresseerd is in vergelijkingen van het type $\mathcal{L}(\mathbf{x})=\mathbf{b}$ voor een lineaire afbeelding $\mathcal{L}$ van een vectorruimte V naar een vectorruimte W. Hierbij is $\mathbf{b}$ een vector uit W en wordt gezocht naar alle $\mathbf{x}\in V$ die op $\mathbf{b}$ afgebeeld worden. Een voorbeeld hiervan is de bekende matrixvergelijking $A\mathbf{x}=\mathbf{b}$.

STELLING 3.21. *Voor een lineaire afbeelding $\mathcal{L}: V\to W$ tussen vectorruimten V en W en een vector $\mathbf{b}$ uit W geldt dat de oplossingsverzameling van de vergelijking $\mathcal{L}(\mathbf{x})=\mathbf{b}$ gelijk is aan*

$$\mathbf{a}+\mathrm{Ker}(\mathcal{L})=\{\mathbf{a}+\mathbf{v}\mid \mathbf{v}\in\mathrm{Ker}(\mathcal{L})\}.$$

Hierbij is $\mathbf{a}$ <u>een</u> vector waarvan het beeld $\mathbf{b}$ is, met andere woorden, $\mathbf{a}$ is <u>een</u> (vaste) oplossing van de vergelijking.

Bewijs: We moeten aantonen dat twee verzamelingen gelijk zijn: de verzameling S van alle $\mathbf{x}$ waarvoor $\mathcal{L}(\mathbf{x})=\mathbf{b}$ en de verzameling $\mathbf{a}+\mathrm{Ker}(\mathcal{L})$. Neem daartoe eerst $\mathbf{x}$ in S, wat betekent dat $\mathcal{L}(\mathbf{x})=\mathbf{b}$. Merk op dat

$$\mathcal{L}(\mathbf{x}-\mathbf{a})=\mathcal{L}(\mathbf{x})-\mathcal{L}(\mathbf{a})=\mathbf{b}-\mathbf{b}=\mathbf{0}$$

zodat $\mathbf{x}-\mathbf{a}$ een element van $\mathrm{Ker}(\mathcal{L})$ is. Maar dan is $\mathbf{x}\in\mathbf{a}+\mathrm{Ker}(\mathcal{L})$.
Nu andersom: neem een vector $\mathbf{x}$ uit $\mathbf{a}+\mathrm{Ker}(\mathcal{L})$, zeg $\mathbf{x}=\mathbf{a}+\mathbf{v}$. Dan

$$\mathcal{L}(\mathbf{x})=\mathcal{L}(\mathbf{a}+\mathbf{v})=\mathcal{L}(\mathbf{a})+\mathcal{L}(\mathbf{v})=\mathbf{b}+\mathbf{0}=\mathbf{b}.$$

Blijkbaar behoort $\mathbf{x}$ tot S. $\square$

Het mooie van bovenstaande stelling is het volgende: in veel situaties is de kern van een afbeelding $\mathcal{L}$ relatief snel te vinden. Als dan ook nog één oplossing van de vergelijking $\mathcal{L}(\mathbf{x}) = \mathbf{b}$ bekend is, is meteen ook de volledige oplossing van die vergelijking bekend. Tel die ene oplossing maar op bij de vectoren uit de kern!

Voorbeeld 3.22. Beschouw de vectorruimte $V = C^\infty(\mathbb{R}, \mathbb{R})$ en de lineaire afbeelding $L : V \to V$ met:

$$L(y) = y'' - 2y' - 3y$$

Dit is een lineaire afbeelding en als we willen oplossen: $L(y) = x$ (dit geeft de differentiaalvergelijking $y'' - 2y' - 3y = x$) dan zegt bovenstaande stelling dat de oplossing eruit zal zien als: één echte oplossing + de kern. Bij analyse noemde we dit: een particuliere oplossing plus de homogene oplossing!!

3.4 Operaties op lineaire afbeeldingen en inverteerbaarheid

In het algemeen kunnen afbeeldingen $\mathcal{L}_1 : U \to V$ en $\mathcal{L}_2 : V \to W$, zoals bekend, na elkaar uitgevoerd worden:

$$\begin{array}{ccccc} U & \xrightarrow{\mathcal{L}_1} & V & \xrightarrow{\mathcal{L}_2} & W \\ \mathbf{u} & \mapsto & \mathcal{L}_1(\mathbf{u}) & \mapsto & \mathcal{L}_2\big(\mathcal{L}_1(\mathbf{u})\big) \end{array}$$

$\mathbf{u}$ uit U wordt via $\mathcal{L}_1$ naar $\mathcal{L}_1(\mathbf{u})$ in V gestuurd, die vervolgens via $\mathcal{L}_2$ doorgestuurd wordt naar de vector $\mathcal{L}_2\big(\mathcal{L}_1(\mathbf{u})\big)$ in W. Dit levert een nieuwe afbeelding van U naar W, die de *samenstelling* of *compositie* van $\mathcal{L}_1$ en $\mathcal{L}_2$ wordt genoemd en genoteerd wordt als $\mathcal{L}_2 \circ \mathcal{L}_1$ (spreek uit: $\mathcal{L}_2$ *na* $\mathcal{L}_1$). De vraag die nu rijst is of de samenstelling van twee lineaire afbeeldingen weer lineair is. Het antwoord wordt gegeven in de volgende stelling.

STELLING 3.23. *Als $\mathcal{L}_1 : U \to V$ en $\mathcal{L}_2 : V \to W$ lineaire afbeeldingen zijn tussen vectorruimten over $\mathbb{R}$ (of $\mathbb{C}$), dan is ook de compositie $\mathcal{L}_2 \circ \mathcal{L}_1 : U \to W$ een lineaire afbeelding.*

Bewijs: De lineaire afbeelding $\mathcal{L}_2 \circ \mathcal{L}_1$ is gedefinieërd door $(\mathcal{L}_2 \circ \mathcal{L}_1)(\mathbf{u}) = \mathcal{L}_2(\mathcal{L}_1(\mathbf{u}))$. Het bewijs van de stelling volgt direct door uitschrijven van de eigenschappen die moeten gelden voor lineaire afbeeldingen. Daarbij maken we gebruik van de lineariteit van $\mathcal{L}_1$ en $\mathcal{L}_2$:

$$\begin{aligned}
(\mathcal{L}_2 \circ \mathcal{L}_1)(\mathbf{u}_1 + \mathbf{u}_2) &= \mathcal{L}_2\big(\mathcal{L}_1(\mathbf{u}_1 + \mathbf{u}_2)\big) \\
&= \mathcal{L}_2\big(\mathcal{L}_1(\mathbf{u}_1) + \mathcal{L}_1(\mathbf{u}_2)\big) \\
&= \mathcal{L}_2\big(\mathcal{L}_1(\mathbf{u}_1)\big) + \mathcal{L}_2\big(\mathcal{L}_1(\mathbf{u}_2)\big) \\
&= (\mathcal{L}_2 \circ \mathcal{L}_1)(\mathbf{u}_1) + (\mathcal{L}_2 \circ \mathcal{L}_1)(\mathbf{u}_2)
\end{aligned}$$

en

$$(\mathcal{L}_2 \circ \mathcal{L}_1)(c\mathbf{u}) = \mathcal{L}_2\big(\mathcal{L}_1(c\mathbf{u})\big) = \mathcal{L}_2\big(c\mathcal{L}_1(\mathbf{u})\big) = c\mathcal{L}_2\big(\mathcal{L}_1(\mathbf{u})\big) = c(\mathcal{L}_2 \circ \mathcal{L}_1)(\mathbf{u}).$$

In beide gevallen hebben we bij het tweede =-teken de lineariteit van $\mathcal{L}_1$ gebruikte en bij het derde =-teken die van $\mathcal{L}_2$. $\square$

Voorbeeld 3.24. De twee schuif-afbeeldingen $\mathcal{L}_1$ en $\mathcal{L}_2$ uit voorbeeld 3.7 kunnen we op twee manieren samenstellen. Het moge duidelijk zijn dat $(\mathcal{L}_1 \circ \mathcal{L}_2)(\mathbf{v}) = \big(\mathcal{L}_1(\mathcal{L}_2(\mathbf{v})\big) = \mathbf{v}$ voor alle

$\mathbf{v} \in \mathbb{R}^\infty$ hetgeen we beknopt kunnen formuleren als $\mathcal{L}_1 \circ \mathcal{L}_2 = \mathrm{Id}$ (waarbij Id de identieke afbeelding is, gedefinieërd in voorbeeld 3.5). Omgekeerd geldt

$$(\mathcal{L}_2 \circ \mathcal{L}_{11})(x_0, x_1, \dots) = \mathcal{L}_2\big(\mathcal{L}_1(x_0, x_1, \dots)\big) = \mathcal{L}_2(x_1, x_2 \dots) = (0, x_1, x_2 \dots)$$

en dus is $\mathcal{L}_2 \circ \mathcal{L}_1$ niet de identieke afbeelding. Ga nog na dat beide samenstellingen inderdaad lineaire afbeeldingen zijn.

Voorbeeld 3.25. Ook de differentieer-afbeelding $\mathcal{D} : C^1([a,b], \mathbb{R}) \to C([a,b], \mathbb{R})$ en de integreer-afbeelding $\mathcal{I} : C([a,b], \mathbb{R}) \to C^1([a,b], \mathbb{R})$ uit voorbeeld 3.8 kunnen we op twee manieren samenstellen. De compositie $\mathcal{D} \circ \mathcal{I} = \mathrm{Id}$, want toepassing van deze compositie op een functie uit $C([a,b], \mathbb{R})$ betekent dat eerst een primitieve bepaald wordt, die daarna gedifferentieerd wordt. Volgens de hoofdstelling van de integraalrekening geeft dit weer de oorspronkelijke functie.
Verder: $\mathcal{I} \circ \mathcal{D} \neq \mathrm{Id}$. Neem maar de constante functie $f(x) = 3$ uit $C^1([a,b], \mathbb{R})$. Dan geldt $\mathcal{D}(f) = 0$, dus ook $\mathcal{I}(\mathcal{D}(f)) = 0 \neq f$.
Uit de relatie $\mathcal{D} \circ \mathcal{I} = \mathrm{Id}$ kunnen $\mathrm{Ker}(\mathcal{I})$ en $\mathrm{Im}(\mathcal{D})$ nog op een elegante manier bepaald worden (vergelijk met voorbeeld 3.18): als f namelijk element is van $\mathrm{Ker}(\mathcal{I})$, dan moet $\mathcal{I}(f) = 0$ (de nulfunctie!) en dus is $\mathcal{D}\big(\mathcal{I}(f)\big) = 0$. Maar dit laatste is $(\mathcal{D} \circ \mathcal{I})(f) = \mathrm{Id}(f) = f$ en dus is $f = 0$. Conclusie de kern van $\mathcal{I}$ bevat alleen de nulfunctie. Neem nu een functie g in $C([a,b], \mathbb{R})$. Omdat $\mathcal{D} \circ \mathcal{I} = \mathrm{Id}$, volgt $g = \mathcal{D}\big(\mathcal{I}(g)\big)$ en dus behoort g, als $\mathcal{D}$-beeld van $\mathcal{I}(g)$, tot $\mathrm{Im}(\mathcal{D})$. Blijkbaar geldt $\mathrm{Im}(\mathcal{D}) = C([a,b], \mathbb{R})$.

Herinner dat we een afbeelding $\mathcal{L} : U \to V$ inverteerbaar noemen als er een "omkeerafbeelding" bestaat. Dat is een afbeelding $\mathcal{T} : V \to U$ met de volgende eigenschap: Als $\mathcal{L}$ de vector $\mathbf{u}$ naar $\mathbf{v}$ stuurt, dan stuurt $\mathcal{T}$ de vector $\mathbf{v}$ naar $\mathbf{u}$ en andersom. Anders gezegd:

Definitie: Een afbeelding (niet per se lineair!) $\mathcal{L} : U \to V$ is **inverteerbaar** als er een afbeelding $\mathcal{T} : V \to U$ bestaat zodat

$\mathcal{L} \circ \mathcal{T} = \mathrm{Id}_V$ (de identieke afbeelding op V) en
$\mathcal{T} \circ \mathcal{L} = \mathrm{Id}_U$ (de identieke afbeelding op U).

In dat geval wordt $\mathcal{T}$ een **inverse** van $\mathcal{L}$ genoemd en $\mathcal{L}$ een inverse van $\mathcal{T}$.

Herinner verder dat een inverteerbare afbeelding $\mathcal{L} : U \to V$ precies één inverse heeft. Dit noemen we dan <u>de</u> inverse van $\mathcal{L}$, en deze afbeelding wordt dan genoteerd met $\mathcal{L}^{-1} : V \to U$. Merk op dat de inverse afbeelding een vector $\mathcal{L}(\mathbf{u})$ terugstuurt naar $\mathbf{u}$.
Inverteerbare lineaire afbeeldingen hebben de prettige eigenschap dat hun inverse ook weer lineair is:

STELLING 3.26. *De inverse van een inverteerbare lineaire afbeelding is ook weer lineair.*

Bewijs: Neem aan dat $\mathcal{L} : U \to V$ een inverteerbare lineaire afbeelding tussen vectorruimten is en laat $\mathcal{T} : V \to U$ de inverse zijn. We laten zien dat ook $\mathcal{T}$ een lineaire afbeelding is.
Met de lineariteit van $\mathcal{L}$ volgt

$$\mathcal{L}\big(\mathcal{T}(\mathbf{u}) + \mathcal{T}(\mathbf{v})\big) = \mathcal{L}\big(\mathcal{T}(\mathbf{u})\big) + \mathcal{L}\big(\mathcal{T}(\mathbf{v})\big) = \mathbf{u} + \mathbf{v}$$

en

$$\mathcal{L}\big(c\mathcal{T}(\mathbf{u})\big) = c\mathcal{L}\big(\mathcal{T}(\mathbf{u})\big) = c\mathbf{u}.$$

Pas hier $\mathcal{T}$ op toe en gebruik dat $\mathcal{T} \circ \mathcal{L} = \mathrm{Id}_U$. Dan volgt direct:

$$\begin{aligned} \mathcal{T}(\mathbf{u}) + \mathcal{T}(\mathbf{v}) &= \mathcal{T}(\mathbf{u} + \mathbf{v}) \\ c\mathcal{T}(\mathbf{u}) &= \mathcal{T}(c\mathbf{u}), \end{aligned}$$

en $\mathcal{T}$ is dus lineair. $\square$

Hoe is na te gaan of een gegeven afbeelding $T : U \to V$ inverteerbaar is? Als U en V oneindig dimensionaal zijn en zit er waarschijnlijk niets anders op dan terug te vallen op de volgende stelling.

> Een afbeelding is inverteerbaar als en slechts als deze bijectief (dat wil zeggen injectief én surjectief) is.

(Herinner: een afbeelding $\mathcal{L} : U \to V$ is *injectief* (of één-één) als $\mathbf{u} \neq \mathbf{v}$ impliceert dat ook $\mathcal{L}(\mathbf{u}) \neq \mathcal{L}(\mathbf{v})$ (met andere woorden: verschillende originelen hebben verschillende beelden). De afbeelding is *surjectief* als het bereik (beeld) heel V is, dus als $\mathrm{Im}(\mathcal{L}) = V$.)

STELLING 3.27. *Een lineaire afbeelding $\mathcal{L} : V \to W$ is injectief als en slechts als* $\mathrm{Ker}(\mathcal{L}) = \{\mathbf{0}\}$.

Bewijs: Als $\mathcal{L}$ injectief is, dan moet $\mathrm{Ker}(\mathcal{L}) = \{\mathbf{0}\}$. Zo niet, dan zouden twee verschillende vectoren uit $\mathrm{Ker}(\mathcal{L})$ op $\mathbf{0}$ afgebeeld worden.
Omgekeerd, stel dat $\mathrm{Ker}(\mathcal{L}) = \{\mathbf{0}\}$. Als zou gelden: $\mathcal{L}(\mathbf{u}) = \mathcal{L}(\mathbf{v})$, dan $\mathcal{L}(\mathbf{u}-\mathbf{v}) = \mathcal{L}(\mathbf{u}) - \mathcal{L}(\mathbf{v}) = \mathbf{0}$. Dus $\mathbf{u} - \mathbf{v} \in \mathrm{Ker}(\mathcal{L})$, en daarom bij aanname $\mathbf{u} - \mathbf{v} = \mathbf{0}$, d.w.z. $\mathbf{u} = \mathbf{v}$. $\square$

Voorbeeld 3.28. Ondanks het feit dat de differentieer-afbeelding $\mathcal{D}$ en de integreer-afbeelding $\mathcal{I}$ uit voorbeeld 3.8 de eigenschap hebben dat $\mathcal{D} \circ \mathcal{I} = \mathrm{Id}$ (vb 3.25), zijn ze niet inverteerbaar: $\mathcal{D}$ is niet injectief, kijk maar naar de kern (zie vb 3.18 of 3.25) en $\mathcal{I}$ is niet surjectief, kijk maar naar het beeld (zie de opgaven).
Iets dergelijks geldt ook voor de schuifafbeeldingen op $\mathbb{R}^\infty$.

Als U en V eindig dimensionaal zijn zullen we een simpelere methode ontwikkelen om in te zien if $T : U \to V$ inverteerbaar is. We zien nu:

Voorbeeld 3.29. Zij V een vectorruimte met basis $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$. De coördinatisering $\mathbf{v} \mapsto [\mathbf{v}]_B$ is inverteerbaar. We kunnen namelijk direct zijn inverse opschrijven:

$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \mapsto c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n.$$

Definitie: We zeggen dat de vectorruimten U en V, beide over $\mathbb{R}$ (of beide over $\mathbb{C}$) **isomorf** zijn als er een inverteerbare lineaire afbeelding $\mathcal{L} : U \to V$ bestaat. Zo'n afbeelding wordt een **isomorfisme** genoemd.

Opmerking: Merk op dat in de situatie van de definitie ook de inverse van $\mathcal{L}$ een isomorfisme is en dat V en U (verwisselde volgorde) dan ook isomorf zijn.

Isomorfe vectorruimten zijn qua vectorruimte-structuur hetzelfde. Eigenschappen die in de ene ruimte gelden worden via het isomorfisme overgeheveld naar de andere. Zo geldt bijvoorbeeld dat een basis voor U door het isomorfisme overgaat naar een basis voor V Het gevolg is dat berekeningen binnen U eventueel gedaan kunnen worden in V door eerst alles "naar V over te hevelen". Vergeet in zo'n situatie niet het eindantwoord weer terug te sturen naar U.
Uit voorbeeld 3.29 volgt dat elke eindig dimensionale vectorruimte V over $\mathbb{R}$ (of $\mathbb{C}$) door middel van de coördinatisering isomorf is met $\mathbb{R}^n$ (of $\mathbb{C}^n$) waarbij n de dimensie van V is. Wat vectorruimte-eigenschappen betreft, zijn ieder tweetal n-dimensionale vectorruimte over $\mathbb{R}$ (of $\mathbb{C}$) aan elkaar gelijk. Het algemene begrip vectorruimte lijkt voor eindig-dimensionale vectorruimten dus niets nieuws opgeleverd te hebben in vergelijking met de $\mathbb{R}^n$ en $\mathbb{C}^n$ die we al hadden!
Een voorbeeld uit de differentiaal vergelijkingen.
Van het college differentiaalvergelijkingen zijn we bekend met de stelling:

STELLING 3.30. *Beschouw de lineaire homogene differentiaal vergelijking*

$$y^{(n)} + c_{n-1}(x)y^{(n-1)} + \cdots + c_1(x)y' + c_0(x)y = 0$$

met beginvoorwaarden: $y^{(n-1)}(x_0) = r_{n-1}, \ldots, y'(x_0) = r_1, y(x_0) = r_0$*, waarbij* $c_0(x), \ldots, c_{n-1}(x)$ *continue functies zijn op het interval* $[a, b]$ *en* $x_0 \in [a, b]$*.*
Dan bestaat er een unieke oplossing $y \in \mathcal{C}^n([a, b], \mathbb{R})$ *van dit beginwaardeprobleem.*

Met behulp van deze stelling kunnen we nu bewijzen.

STELLING 3.31. *Beschouw de lineaire homogene differentiaal vergelijking*

$$y^{(n)} + c_{n-1}(x)y^{(n-1)} + \cdots + c_1(x)y' + c_0(x)y = 0$$

waarbij $c_0(x), \ldots, c_{n-1}(x)$ *continue functies zijn op het interval* $[a, b]$*. Dan is de oplossingsverzameling* $\mathcal{O}$ *een n-dimensionale deelruimte van* $\mathcal{C}^n([a, b], \mathbb{R})$*.*

Bewijs: Dat $\mathcal{O}$ een lineaire deelruimte is, volgt rechtstreeks en is een opgave. Maar waarom is deze n-dimensionaal? Daartoe fixeren we een punt $x_0 \in [a, b]$ en beschouwen we de afbeelding

$$\mathcal{L} : \mathcal{O} \to \mathbb{R}^n$$

gedefinieerd door: $\mathcal{L}(y) = \begin{bmatrix} y(x_0) \\ y'(x_0) \\ \vdots \\ y^{(n-1)}(x_0) \end{bmatrix}$. De afbeelding is lineair (ga na). Uit Stelling 3.30 volgt dat de afbeelding surjectief (omdat er een oplossing is) en injectief (vanwege de uniciteit van een oplossing) en daarom zal $\mathcal{O}$ isomorf met $\mathbb{R}^n$ zijn en daarom dimensie n hebben. Dus inderdaad, er volgt dat er n lineair onafhankelijke oplossingen zullen bestaan, oplossingen die bij het college differentiaal vergelijkingen gezocht werden. $\square$

Tenslotte kunnen we $y^{(n)} + c_{n-1}(x)y^{(n-1)} + \cdots + c_1(x)y' + c_0(x)y$ beschouwen als een lineaire afbeelding $\mathcal{L} : \mathcal{C}^n([a,b],\mathbb{R}) \to \mathcal{C}([a,b],\mathbb{R})$ via:

$$\mathcal{L}(y) = y^{(n)} + c_{n-1}(x)y^{(n-1)} + \cdots + c_1(x)y' + c_0(x)y$$

Merk op dat de Kern van deze afbeelding eigenlijk de algemene oplossing is van de bijbehorende homogene vergelijking $y^{(n)} + c_{n-1}(x)y^{(n-1)} + \cdots + c_1(x)y' + c_0(x)y = 0$!
Stelling 3.21 levert ons nu dat de oplossing van de differentiaalvergelijking

$$y^{(n)} + c_{n-1}(x)y^{(n-1)} + \cdots + c_1(x)y' + c_0(x)y = p(x)$$

(met $p(x)$ continu) gelijk is aan een particuliere oplossing + de homogene oplossing

3.5 Matrixrepresentaties van lineaire afbeeldingen

We weten dat iedere lineaire afbeelding $T : \mathbb{R}^n \to \mathbb{R}^m$ bepaald wordt door een matrix, de standaardmatrix $[T]$ via $T(\mathbf{x} = [T]\mathbf{x}$. Er geldt: $[T] = [T(\mathbf{e}_1) \cdots T(\mathbf{e}_n)]$.
Er is al eerder gezegd dat lineaire afbeeldingen tussen willekeurige vectorruimtes niet op deze manier door een matrix bepaald kunnen worden. Tenslotte, wat is Af als $f \in \mathcal{F}([0,1],\mathbb{R})$ en A een matrix? Echter, als B een basis is van de vectorruimte, dan bestaat $A[\mathbf{v}]_B$ wel. Zouden die coördinaten gebruikt kunnen worden om een matrix met een lineaire algebra te associeren?
We zullen zien dat, na keuze van bases in domein en codomein, een lineaire afbeelding tussen eindig dimensionale vectorruimten met behulp van coördinatiseringen altijd gezien kan worden als een matrixtransformatie.

STELLING 3.32. *Als $\mathcal{L} : U \to V$ een lineaire afbeelding is tussen de eindig dimensionale vectorruimten U en V (over $\mathbb{L}$) en als $B = (\mathbf{u}_1 \ldots \mathbf{u}_n)$ een basis voor U is en $C = (\mathbf{v}_1 \ldots \mathbf{v}_m)$ een basis voor V, dan geldt voor elke $\mathbf{u}$ uit U dat de coördinatisering van het beeld $\mathcal{L}(\mathbf{u})$ berekend kan worden via een matrixvermenigvuldiging:*

$$[\mathcal{L}(\mathbf{u})]_C = A[\mathbf{u}]_B$$

Hierbij is A de $m \times n$ matrix waarvan de i^{e} kolom gelijk is aan de coördinatisering van het beeld van de i^{e} basisvector uit B ten opzichte van de basis C, dus

$$A = \begin{bmatrix} [\mathcal{L}(\mathbf{u}_1)]_C & [\mathcal{L}(\mathbf{u}_2)]_C & \cdots & [\mathcal{L}(\mathbf{u}_n)]_C \end{bmatrix}.$$

Door de eigenschap $[\mathcal{L}(\mathbf{u})]_C = A[\mathbf{u}]_B$ wordt de matrix A volledig bepaald en A wordt genoteerd als:

$$\underset{C \leftarrow B}{[\mathcal{L}]}$$

en wordt de **(transformatie)matrix van $\mathcal{L}$ ten opzichte van de bases B en (naar) C** *genoemd.*

Bewijs: Neem een vector $\mathbf{u} \in U$. Omdat B een basis voor U is, is $\mathbf{u}$ op precies één manier te schrijven als lineaire combinatie van $\mathbf{u}_1, \ldots, \mathbf{u}_n$, zeg

$$\mathbf{u} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n.$$

Uit de lineariteit van $\mathcal{L}$ volgt dan

$$\mathcal{L}(\mathbf{u}) = c_1\mathcal{L}(\mathbf{u}_1) + c_2\mathcal{L}(\mathbf{u}_2) + \cdots + c_n\mathcal{L}(\mathbf{u}_n).$$

Dit is een vector uit V die we dus kunnen coördinatiseren ten opzichte van de basis C. We krijgen dan

$$\begin{aligned}
[\mathcal{L}(\mathbf{u})]_C &= [c_1\mathcal{L}(\mathbf{u}_1) + c_2\mathcal{L}(\mathbf{u}_2) + \cdots + c_n\mathcal{L}(\mathbf{u}_n)]_C \\
&= c_1[\mathcal{L}(\mathbf{u}_1)]_C + c_2[\mathcal{L}(\mathbf{u}_2)]_C + \cdots + c_n[\mathcal{L}(\mathbf{u}_n)]_C \\
&= \begin{bmatrix} [\mathcal{L}(\mathbf{u}_1)]_C & [\mathcal{L}(\mathbf{u}_2)]_C & \cdots & [\mathcal{L}(\mathbf{u}_n)]_C \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = A[\mathbf{u}]_B.
\end{aligned}$$

Bij het tweede =-teken hebben we de lineariteit van de coördinaatafbeelding gebruikt en bij het derde =-teken de definitie van de matrix-vector vermenigvuldiging.
Tenslotte, zij D een matrix met de eigenschap: $[\mathcal{L}(\mathbf{u})]_C = D[\mathbf{u}]_B$, voor alle $\mathbf{u} \in U$. We tonen aan dat $D = \underset{C\leftarrow B}{[\mathcal{L}]}$. Merk eerst op dat D een $m \times n$ matrix **moet** zijn. Als we voor $\mathbf{u}$ de vector $\mathbf{u}_i \in B$ invullen, dan lezen we: $[\mathcal{L}(\mathbf{u}_i)]_C = D[\mathbf{u}_i]_B$, voor $i = 1 \ldots n$. Omdat $[\mathbf{u}_i]_B = \mathbf{e}_i$ en omdat $D\mathbf{e}_i = i^{\text{de}}$ kolom van D volgt:

$$\text{de } i^{\text{de}} \text{ kolom van } D \text{ is } [\mathcal{L}(\mathbf{u}_i)]_C, \text{ voor } i = 1 \ldots n$$

en dus $D = \underset{C\leftarrow B}{[\mathcal{L}]}$. □

Opmerking: We moeten kunnen inzien dat de standaardmatrix $[\mathcal{T}]$ van een lineaire afbeelding $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^m$ eigenlijk een transformatiematrix van type $\underset{C\leftarrow B}{[\mathcal{T}]}$ is. Maar wat zijn de bases B en C dan? Merk op: als $E_n = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ de standaardbasis is van $\mathbb{R}^n$ dan geldt: $\mathbf{x} = [\mathbf{x}]_{E_n}$.
En dus vertaald de eigenschap: $\mathcal{T}(\mathbf{x}) = [\mathcal{T}]\mathbf{x}$ naar:

$$[\mathcal{T}(\mathbf{x})]_{E_m} = [\mathcal{T}][\mathbf{x}]_{E_n}$$

Kortom:

$$[T] = \underset{E_m\leftarrow E_n}{[\mathcal{T}]}$$

De standaardmatrix is de transformatiematrix t.o.v. de standaardbases!

Voordat we een voorbeeld geven nog het volgende: de stelling zegt eigenlijk twee dingen: ten eerste hoe de matrix $\underset{C\leftarrow B}{[\mathcal{L}]}$ gemaakt moet worden en ten tweede hoe hij gebruikt moet worden!

Voorbeeld 3.33. Zij $\mathcal{L} : \text{Pol}_2(\mathbb{C}) \to \text{Pol}_3(\mathbb{C})$ de afbeelding gedefinieërd door

$$\mathcal{L}(p)(x) = xp'(x) + \int_1^x p(t)\, dt.$$

Ga zelf na dat $\mathcal{L}$ goed gedefiniëerd en lineair is. Voordat we de matrix van $\mathcal{L}$ ten opzichte van de standaardbases gaan bepalen, berekenen we eerst even het beeld van het polynoom p gegeven door $p(x) = 1 - 2ix + 3x^2$. Dat is (volgens het voorschrift)

$$\mathcal{L}(p)(x) = x(-2i + 6x) + \int_1^x (1 - 2it + 3t^2)\, dt = (-2 + i) + (1 - 2i)x + (6 - i)x^2 + x^3.$$

Laat nu $B = (1, x, x^2)$ de standaardbasis voor $\text{Pol}_2(\mathbb{C})$ zijn en $C = (1, x, x^2, x^3)$ de standaardbasis voor $\text{Pol}_3(\mathbb{C})$. De matrix $\underset{C\leftarrow B}{[\mathcal{L}]}$ is volgens stelling 3.32

$$\underset{C\leftarrow B}{[\mathcal{L}]} = \begin{bmatrix} [\mathcal{L}(1)]_C & [\mathcal{L}(x)]_C & [\mathcal{L}(x^2)]_C \end{bmatrix}.$$

Eerst moet kennelijk de functiewaarden van de basisvectoren 1, x en x^2 berekend worden:

$$\begin{aligned}\mathcal{L}(1) &= -1 + x\\ \mathcal{L}(x) &= -\frac{1}{2} + x + \frac{1}{2}x^2\\ \mathcal{L}(x^2) &= -\frac{1}{3} + 2x^2 + \frac{1}{3}x^3\end{aligned}$$

na het bepalen van de coördinaten van deze beelden ten opzichte van de basis C vinden we

$$\underset{C\leftarrow B}{[\mathcal{L}]} = \begin{bmatrix} -1 & -1/2 & -1/3\\ 1 & 1 & 0\\ 0 & 1/2 & 2\\ 0 & 0 & 1/3\end{bmatrix}.$$

Dat deze matrix gebruikt kan worden voor het berekenen van $\mathcal{L}$-beelden illustreren we door het net berekende beeld van $p(x) = 1 - 2ix + 3x^2$ nog een keer te bepalen, nu met $[\mathcal{L}]_{C\leftarrow B}$. Hiervoor zijn de coördinaten van p ten opzichte van B nodig en die zijn $\begin{bmatrix} 1\\ -2i\\ 3\end{bmatrix}$. Met de formule voor de bepaling van de C-coördinaten van het $\mathcal{L}$-beeld uit stelling 3.32 volgt nu

$$[\mathcal{L}(p)]_C = \underset{C\leftarrow B}{[\mathcal{L}]}\,[p]_B = \begin{bmatrix} -1 & -1/2 & -1/3\\ 1 & 1 & 0\\ 0 & 1/2 & 2\\ 0 & 0 & 1/3\end{bmatrix}\begin{bmatrix} 1\\ -2i\\ 3\end{bmatrix} = \begin{bmatrix} -2+i\\ 1-2i\\ 6-i\\ 1\end{bmatrix}.$$

Dit zijn dus de coördinaten van $\mathcal{L}(p)$ ten opzichte van de basis C. Daaruit kan geconcludeerd worden dat

$$\mathcal{L}(p)(x) = (-2+i) + (1-2i)x + (6-i)x^2 + x^3.$$

Dit komt, zoals te verwachten, overeen met het eerder gevonden antwoord. Het voordeel van deze laatste berekening moge duidelijk zijn! $\square$

Opmerking: De matrix van een lineaire afbeelding $\mathcal{L} : U \to V$ hangt natuurlijk af van de bases die voor U en V gekozen worden.

Net als bij lineaire afbeeldingen van $\mathbb{R}^n$ naar $\mathbb{R}^m$ (of van $\mathbb{C}^n$ naar $\mathbb{C}^m$) geldt ook nu dat de matrix van de compositie van lineaire afbeeldingen het product van de matrices van de afzonderlijke afbeeldingen is:

STELLING 3.34. *Als $\mathcal{L}_1 : U \to V$ en $\mathcal{L}_2 : V \to W$ lineaire afbeeldingen tussen eindig dimensionale vectorruimten over $\mathbb{R}$ (of $\mathbb{C}$) zijn en als B, C en D bases zijn voor respectievelijk U, V en W, dan geldt*

$$\underset{D\leftarrow B}{[\mathcal{L}_2 \circ \mathcal{L}_1]} = \underset{D\leftarrow C}{[\mathcal{L}_2]}\,\underset{C\leftarrow B}{[\mathcal{L}_1]}.$$

Bewijs: Beschouw de formules $[\mathcal{L}_1(\mathbf{u})]_C = \underset{C\leftarrow B}{[\mathcal{L}_1]}\,[\mathbf{u}]_B$, voor alle $\mathbf{u} \in U$ en $[\mathcal{L}_2(\mathbf{v})]_D = \underset{D\leftarrow C}{[\mathcal{L}_2]}\,[\mathbf{v}]_C$, voor alle $\mathbf{v} \in V$. Omdat $\mathcal{L}_1(\mathbf{u})$ een element van V is, mag in de laatste formule $\mathbf{v}$ vervangen worden door $\mathcal{L}_1(\mathbf{u})$. Dit geeft aanleiding tot

$$[(\mathcal{L}_2 \circ \mathcal{L}_1)(\mathbf{u})]_D = [\mathcal{L}_2\big(\mathcal{L}_1(\mathbf{u})\big)]_D = \underset{D\leftarrow C}{[\mathcal{L}_2]}\,[\mathcal{L}_1(\mathbf{u})]_C = \underset{D\leftarrow C}{[\mathcal{L}_2]}\,\underset{C\leftarrow B}{[\mathcal{L}_1]}\,[\mathbf{u}]_B.$$

Anderzijds geldt

$$[(\mathcal{L}_2 \circ \mathcal{L}_1)(\mathbf{u})]_D = \underset{D\leftarrow B}{[\mathcal{L}_2 \circ \mathcal{L}_1]}\,[\mathbf{u}]_B.$$

Omdat de matrix $\underset{D\leftarrow B}{[\mathcal{L}_2 \circ \mathcal{L}_1]}$ uniek is en $\underset{D\leftarrow C}{[\mathcal{L}_2]}\ \underset{C\leftarrow B}{[\mathcal{L}_1]}$ "hetzelfde doet", moet nu gelden dat deze twee matrices gelijk zijn. Hiermee is de stelling bewezen. $\square$

We kunnen nu de inverteerbaarheid van de lineaire afbeelding koppelen aan de inverteerbaarheid van de bijbehorende matrixrepresentatie van de afbeelding.

STELLING 3.35. *Als $\mathcal{L} : U \to V$ een lineaire afbeelding is tussen eindig dimensionale vectorruimten U en V over $\mathbb{R}$ (of over $\mathbb{C}$) en als $B = (\mathbf{u}_1 \ldots \mathbf{u}_n)$ een basis voor U is en $C = (\mathbf{v}_1 \ldots \mathbf{v}_m)$ een basis voor V, dan geldt:*

1. *Als $\mathbf{u} \in U$ dan: $\mathbf{u} \in \mathrm{Ker}(\mathcal{L}) \Leftrightarrow [\mathbf{u}]_B \in \mathrm{Nul}(\underset{C\leftarrow B}{[\mathcal{L}]})$.*
2. *Als $\mathbf{v} \in V$ dan: $\mathbf{v} \in \mathrm{Im}(\mathcal{L}) \Leftrightarrow [\mathbf{v}]_C \in \mathrm{Col}(\underset{C\leftarrow B}{[\mathcal{L}]})$.*
3. *Merk op dat de identiteit uit Stelling 3.20: $\dim(U) = \dim(\mathrm{Ker}(\mathcal{L})) + \dim(Im(\mathcal{L})$ nu volgt uit de identiteit: als A een $m \times n$ matrix, dan $n = \dim(\mathrm{Col}(A)) + \dim(\mathrm{Nul}(A))$.*

Bewijs:

1. $\mathbf{u} \in \mathrm{Ker}(\mathcal{L}) \Leftrightarrow \mathcal{L}(\mathbf{u}) = \mathbf{0} \Leftrightarrow [\mathcal{L}(\mathbf{u})]_C = [\mathbf{0}]_C = \mathbf{0} \Leftrightarrow \underset{C\leftarrow B}{[\mathcal{L}]}[\mathbf{u}]_B = \mathbf{0} \Leftrightarrow [\mathbf{u}]_B \in \mathrm{Nul}(\underset{C\leftarrow B}{[\mathcal{L}]})$.

2. $\mathbf{v} \in \mathrm{Im}(\mathcal{L}) \Leftrightarrow$ er bestaat $\mathbf{u} \in U$ zo $\mathcal{L}(\mathbf{u}) = \mathbf{v} \Leftrightarrow$ er bestaat $\mathbf{u} \in U$ zo $[\mathcal{L}(\mathbf{u})]_C = [\mathbf{v}]_C$ $\Leftrightarrow$ er bestaat $\mathbf{u} \in U$ zo $\underset{C\leftarrow B}{[\mathcal{L}]}[\mathbf{u}]_B = [\mathbf{v}]_C \Leftrightarrow [\mathbf{v}]_C \in \mathrm{Col}(\underset{C\leftarrow B}{[\mathcal{L}]})$.

3. Als $\dim(U) = n$ en $\dim(V) = m$ dan is $\underset{C\leftarrow B}{[\mathcal{L}]}$ een $m \times n$ matrix.
   Volgens 1. is de coördinatisering een isomorfisme tussen $\mathrm{Ker}(\mathcal{L})$ en $\mathrm{Nul}(\underset{C\leftarrow B}{[\mathcal{L}]})$ en dus: $\dim(\mathrm{Ker}(\mathcal{L})) = \dim(\mathrm{Nul}(\underset{C\leftarrow B}{[\mathcal{L}]}))$.
   Volgens 2. is de coördinatisering een isomorfisme tussen $\mathrm{Im}(\mathcal{L})$ en $\mathrm{Col}(\underset{C\leftarrow B}{[\mathcal{L}]})$ en dus: $\dim(\mathrm{Im}(\mathcal{L})) = \dim(\mathrm{Col}(\underset{C\leftarrow B}{[\mathcal{L}]}))$.
   En dus volgt de identiteit $n = \dim(U) = \dim(\mathrm{Ker}(\mathcal{L})) + \dim(\mathrm{Im}(\mathcal{L})$ inderdaad uit de identiteit $n = \dim(\mathrm{Nul}(\underset{C\leftarrow B}{[\mathcal{L}]})) + \dim(\mathrm{Col}(\underset{C\leftarrow B}{[\mathcal{L}]}))$.

$\square$

STELLING 3.36. *Als $\mathcal{L} : U \to V$ een lineaire afbeelding is tussen eindig dimensionale vectorruimten U en V over $\mathbb{R}$ (of over $\mathbb{C}$) en als $B = (\mathbf{u}_1 \dots \mathbf{u}_n)$ een basis voor U is en $C = (\mathbf{v}_1 \dots \mathbf{v}_m)$ een basis voor V, dan geldt:*
De afbeelding $\mathcal{L} : U \to V$ is inverteerbaar (als afbeelding) als en slechts als de matrix $\underset{C \leftarrow B}{[\mathcal{L}]}$ inverteerbaar is (als matrix). Bovendien geldt: als $\underset{C \leftarrow B}{[\mathcal{L}]}$ inverteerbaar is dan:

$$\underset{B \leftarrow C}{[\mathcal{L}^{-1}]} = \left(\underset{C \leftarrow B}{[\mathcal{L}]}\right)^{-1}$$

Bewijs: Stel dat de afbeelding $\mathcal{L} : U \to V$ inverteerbaar is. Dan bestaat $T = \mathcal{L}^{-1} : V \to U$ zodat $T \circ L = Id_V$ en $L \circ T = Id_W$. (Merk op dat $\underset{B \leftarrow B}{[Id_V]} = I$ en $\underset{C \leftarrow C}{[Id_W]} = I$. Daarom geldt: $I = \underset{B \leftarrow B}{[Id_V]} = \underset{B \leftarrow B}{[T \circ L]} = \underset{B \leftarrow C}{[T]}\underset{C \leftarrow B}{[L]}$ en $I = \underset{C \leftarrow C}{[Id_W]} = \underset{C \leftarrow C}{[L \circ T]} = \underset{C \leftarrow B}{[L]}\underset{B \leftarrow C}{[T]}$. Er volgt dat de matrix $\underset{C \leftarrow B}{[L]}$ vierkant is, inverteerbaar is en dat $\underset{B \leftarrow C}{[T]} = \underset{B \leftarrow C}{[\mathcal{L}^{-1}]} = \left(\underset{C \leftarrow B}{[L]}\right)^{-1}$.
Omgekeerd, stel dat de matrix $\underset{C \leftarrow B}{[L]}$ inverteerbaar is. We definieëren $T : V \to U$ via de coordinaten door:

$$[T(\mathbf{v})]_B = \left(\underset{C \leftarrow B}{[L]}\right)^{-1} [\mathbf{v}]_C.$$

We claimen dat T een lineaire afbeelding is.
Om in te zien dat $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$, is het voldoende om in te zien dat $[T(\mathbf{v}_1 + \mathbf{v}_2)]_B = [T(\mathbf{v}_1) + T(\mathbf{v}_2)]_B$. Omdat $[T(\mathbf{v}_1) + T(\mathbf{v}_2)]_B = [T(\mathbf{v}_1)]_B + [T(\mathbf{v}_2)]_B$ moeten we laten zien:

$$[T(\mathbf{v}_1 + \mathbf{v}_2)]_B = [T(\mathbf{v}_1)]_B + [T(\mathbf{v}_2)]_B.$$

Welnu: $[T(\mathbf{v}_1 + \mathbf{v}_2)]_B = \left(\underset{C \leftarrow B}{[L]}\right)^{-1} [\mathbf{v}_1 + \mathbf{v}_2]_C = \left(\underset{C \leftarrow B}{[L]}\right)^{-1} [\mathbf{v}_1]_C + \left(\underset{C \leftarrow B}{[L]}\right)^{-1} [\mathbf{v}_2]_C = [T(\mathbf{v}_1)]_B + [T(\mathbf{v}_2)]_B$. Evenzo wordt de identiteit $T(c\mathbf{v}) = cT(\mathbf{v})$ aangetoond.
Tenslotte, omdat $\underset{B \leftarrow C}{[T]} = \left(\underset{C \leftarrow B}{[L]}\right)^{-1}$ zal gelden: $T \circ L = Id_V$ en $L \circ T = Id_W$. Daarom zal L inverteerbaar zijn met inverse $L^{-1} = T$. □

Voorbeeld 3.37. Beschouw de functieruimte $C(\mathbb{R}, \mathbb{R})$ en de deelruimte V opgespannen door de drie functies $e^{3x}, xe^{3x}, x^2e^{3x}$, i.e. $V = \text{Span}\{e^{3x}, xe^{3x}, x^2e^{3x}\}$. Beschouw de lineaire afbeelding $\mathcal{D} : V \to V$ gedefinieerd door:

$$\mathcal{D}(p)(x) = p'(x)$$

(Merk op: dit is inderdaad een afbeelding naar V, want de afgeleide van de functies $e^{3x}, xe^{3x}, x^2e^{3x}$ zitten in V en de rekenregels voor afgeleide geeft dat dit inderdaad een lineaire afbeelding is.)
Vragen:

1. Laat zien dat $B = \{e^{3x}, xe^{3x}, x^2e^{3x}\}$ een basis voor V is.
2. Bepaal $\underset{B \leftarrow B}{[\mathcal{D}]}$.
3. Bepaal een expliciete formule voor de afbeelding $\mathcal{D}$. (Hiermee bedoelen we: als $p \in V$ met $[p]_B = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ dan $[\mathcal{D}(p)]_B = \dots$.) Bepaal hiermee de afgeleide van $p(x) = e^{3x} + xe^{3x} + x^2e^{3x}$.

4. Is de afbeelding $\mathcal{D} : V \to V$ inverteerbaar? Zo ja, bepaal een expliciete formule voor de inverse afbeelding.

5. Bepaal een primitieve van $p(x) = e^{3x} + xe^{3x} + x^2e^{3x}$ in V.

Antwoorden 1. Per definitie van V spant B de ruimte V op en onafhankelijkheid van de drie functies $e^{3x}, xe^{3x}, x^2e^{3x}$ in B volgt bijvoorbeeld via de Wronskiaan (ga na). Dan is B dus een basis.

2. Omdat $\mathcal{D}(e^{3x}) = 3e^{3x}$ volgt $[\mathcal{D}(e^{3x})]_B = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$.

Omdat $\mathcal{D}(xe^{3x}) = e^{3x} + 3xe^{3x}$ volgt $[\mathcal{D}(xe^{3x})]_B = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$.

Omdat $\mathcal{D}(x^2e^{3x}) = 2xe^{3x} + 3x^2e^{3x}$ volgt $[\mathcal{D}(x^2e^{3x})]_B = \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}$.

En dus:

$$\underset{B\leftarrow B}{[\mathcal{D}]} = [\,[\mathcal{D}(e^{3x})]_B\,[\mathcal{D}(xe^{3x})]_B\,[\mathcal{D}(x^2e^{3x})]_B\,] = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

3. $[\mathcal{D}(p)]_B = \underset{B\leftarrow B}{[\mathcal{D}]}[p]_B = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3x_1 + x_2 \\ 3x_2 + 2x_3 \\ 3x_3 \end{bmatrix}$. De afgeleide van de gegeven $p \in V$ is $\mathcal{D}(p)$. Omdat $[p]_B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ volgt $[\mathcal{D}(p)]_B = \begin{bmatrix} 4 \\ 5 \\ 3 \end{bmatrix}$ en dus is de afgeleide:

$$4e^{3x} + 5xe^{3x} + 3x^2e^{3x}.$$

4. Volgens stelling 3.36 is de afbeelding $\mathcal{D} : V \to V$ inverteerbaar (als afbeelding) als en slechts als de matrix $\underset{B\leftarrow B}{[\mathcal{D}]}$ inverteerbaar is (als matrix). Dit laatste is het geval, want $\det(\underset{B\leftarrow B}{[\mathcal{D}]}) = 3^3 \neq 0$.

Omdat $\underset{B\leftarrow B}{[\mathcal{D}^{-1}]} = \left(\underset{B\leftarrow B}{[\mathcal{D}]}\right)^{-1}$ volgt:

$$\underset{B\leftarrow B}{[\mathcal{D}^{-1}]} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 1/3 & -1/9 & 2/27 \\ 0 & 1/3 & -2/9 \\ 0 & 0 & 1/3 \end{bmatrix} \quad \text{(ga na).}$$

Voor de expliciete formule van de inverse krijgen we: als $p \in V$ met $[p]_B = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ dan

$$[\mathcal{D}^{-1}(p)]_B = \underset{B\leftarrow B}{[\mathcal{D}^{-1}]}[p]_B = \begin{bmatrix} 1/3 & -1/9 & 2/27 \\ 0 & 1/3 & -2/9 \\ 0 & 0 & 1/3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{3}x_1 - \frac{1}{9}x_2 + \frac{2}{27}x_3 \\ \frac{1}{3}x_2 - \frac{2}{9}x_3 \\ \frac{1}{3}x_3 \end{bmatrix}.$$

5. Merk op: als $p \in V$ dan is $q = \mathcal{D}^{-1}(p) \in V$ is kennelijk een functie in V met $p = \mathcal{D}(q) = q'$, d.w.z. q is een primitieve van V. Omdat (als $p(x) = e^{3x} + xe^{3x} + x^2e^{3x}$) $[p]_B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ geldt dat de B–coördinaten van een primitieve van p zijn:

$$[\int e^{3x} + xe^{3x} + x^2e^{3x}\,dx]_B = \begin{bmatrix} 1/3 & -1/9 & 2/27 \\ 0 & 1/3 & -2/9 \\ 0 & 0 & 1/3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{8}{27} \\ \frac{1}{9} \\ \frac{1}{3} \end{bmatrix}.$$

en dus:

$$\int e^{3x} + xe^{3x} + x^2e^{3x}\,dx = \frac{12}{27}e^{3x} + \frac{1}{9}xe^{3x} + \frac{1}{3}x^2e^{3x}. \qquad \square$$

We laten vervolgens zien dat de overgangsmatrix $\underset{B\leftarrow C}{P}$ uit college 8 eigenlijk een matrixrepresentatie van een afbeelding is.

STELLING 3.38. *Laat V een vectorruimte zijn met basis B voor V en met een andere basis C voor V. Beschouw de identeit op V, d.w.z. de lineaire afbeelding $Id : V \to V$ met $Id(\mathbf{v}) = \mathbf{v}$, voor alle $\mathbf{v} \in V$. Dan geldt:*

$$\underset{C\leftarrow B}{[Id]} = \underset{C\leftarrow B}{P}.$$

Bewijs: Dit moet in ieder geval duidelijk zijn uit de formules van beide matrices. Als $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ dan:

$$\underset{C\leftarrow B}{P} = [[\mathbf{b}_1]_C \ldots [\mathbf{b}_k]_C]$$

en

$$\underset{C\leftarrow B}{[Id]} = [[Id(\mathbf{b}_1)]_C \ldots [Id(\mathbf{b}_k)]_C].$$

Omdat $Id(\mathbf{b}_i) = \mathbf{b}_1$ volgt de gelijkheid van de twee matrices. $\square$

Tenslotte presenteren we de relatie tussen matrixrepresentaties van dezelfde afbeelding t.o.v. verschillende bases.

STELLING 3.39. *Als $\mathcal{L} : U \to V$ een lineaire afbeelding tussen eindig dimensionale vectorruimten over $\mathbb{L}$ is, en als B_1 een basis voor U en C_1 een basis voor V is, dan is*

$$\underset{C1\leftarrow B1}{[\mathcal{L}]}$$

de matrix van $\mathcal{L}$ ten opzichte van deze bases. Als we voor U en voor V een andere basis kiezen, zeg B_2 en C_2, dan geldt voor de matrixrepresentatie van $\mathcal{L}$ ten opzichte van deze nieuwe bases dat

$$\underset{C2\leftarrow B2}{[\mathcal{L}]} = \underset{C2\leftarrow C1}{P}\ \underset{C1\leftarrow B1}{[\mathcal{L}]}\ \underset{B1\leftarrow B2}{P}.$$

Bewijs: Een formele manier om dit te zien is als samenvoeging van stelling 3.38 en stelling 3.34. Omdat $\mathcal{L} = \mathrm{Id} \circ \mathcal{L} \circ Id$ volgt via de tweede en dan de eerste stelling dat

$$\underset{C2\leftarrow B2}{[\mathcal{L}]} = \underset{C_2\leftarrow C_1}{[Id]}\ \underset{C1\leftarrow B1}{[\mathcal{L}]}\ \underset{B_1\leftarrow B_2}{[Id]} = \underset{C2\leftarrow C1}{P}\ \underset{C1\leftarrow B1}{[\mathcal{L}]}\ \underset{B1\leftarrow B2}{P}.$$

Een andere manier wordt verkregen door na te gaan dat de matrix $D = \underset{C2\leftarrow C1}{P}\underset{C1\leftarrow B1}{[\mathcal{L}]}\underset{B1\leftarrow B2}{P}$ voldoet aan de eigenschap $[\mathcal{L}(\mathbf{u})]_{C_2} = D[\mathbf{u}]_{B_2}$, de eigenschap die $\underset{C2\leftarrow B2}{[\mathcal{L}]}$ volledig karakteriseert. Inderdaad:

$$\underset{C2\leftarrow C1}{P}\underset{C1\leftarrow B1}{[\mathcal{L}]}\underset{B1\leftarrow B2}{P}[\mathbf{u}]_{B_2} = \underset{C2\leftarrow C1}{P}\underset{C1\leftarrow B1}{[\mathcal{L}]}(\underset{B1\leftarrow B2}{P}[\mathbf{u}]_{B_2}) = \underset{C2\leftarrow C1}{P}\underset{C1\leftarrow B1}{[\mathcal{L}]}[\mathbf{u}]_{B_1} =$$

$$\underset{C2\leftarrow C1}{P}(\underset{C1\leftarrow B1}{[\mathcal{L}]}[\mathbf{u}]_{B_1}) = \underset{C2\leftarrow C1}{P}[\mathcal{L}(\mathbf{u})]_{C_1} = [\mathcal{L}(\mathbf{u})]_{C_2}.$$

□

Voorbeeld 3.40. Beschouw $\mathbb{R}^3$ met basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ waarbij

$$\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \ \mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \ \text{ en } \ \mathbf{b}_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

Vraag: Bepaal de standaardmatrix van de lineaire afbeelding $\mathcal{L} : \mathbb{R}^3 \to \mathbb{R}^3$ met de eigenschap $\mathcal{L}(\mathbf{b}_1) = \mathbf{b}_1 + \mathbf{b}_2$, $\mathcal{L}(\mathbf{b}_2) = \mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3$ en $\mathcal{L}(\mathbf{b}_3) = \mathbf{b}_1 - \mathbf{b}_2$.
Antwoord: Laat $E = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ de standaardbasis van $\mathbb{R}^3$ zijn. De vraag is $[\mathcal{L}] = \underset{E\leftarrow E}{[\mathcal{L}]}$ te bepalen. In eerste instantie denk je wellicht dat hiervoor $\mathcal{L}(\mathbf{e}_i)$ berekend moet worden en rechtstreeks uit de gegevens is dat heel wat rekenwerk.
Veel slimmer is je te realiseren dat eigenlijk $\underset{B\leftarrow B}{[\mathcal{L}]}$ eigenlijk gegeven is, want er staat $[\mathcal{L}(\mathbf{b}_1)]_B = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $[\mathcal{L}(\mathbf{b}_2)]_B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ en $[\mathcal{L}(\mathbf{b}_3)]_B = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ en dus:

$$\underset{B\leftarrow B}{[\mathcal{L}]} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

Merk op dat $\underset{E\leftarrow B}{P} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ en daarom:

$$[\mathcal{L}] = \underset{E\leftarrow E}{[\mathcal{L}]} = \underset{E\leftarrow B}{P}\underset{B\leftarrow B}{[\mathcal{L}]}\underset{B\leftarrow E}{P} = \underset{E\leftarrow B}{P}\underset{B\leftarrow B}{[\mathcal{L}]}(\underset{E\leftarrow B}{P})^{-1} = \begin{bmatrix} 1 & 2 & -1 \\ 3/2 & 5/2 & -5/2 \\ 1/2 & 5/2 & -3/2 \end{bmatrix} \quad \text{(ga na).}$$

3.6 Lineaire operatoren, eigenvectoren en eigenwaarden

Definitie: Een lineaire afbeelding $\mathcal{L}$ met hetzelfde domein als codomein wordt een lineaire operator genoemd.

Dus $\mathcal{L} : V \to V$ is een lineaire operator. Voor lineaire operatoren kunnen we het begrip eigenvector en eigenwaarden definiëren.

Definitie: Een vector $\mathbf{v} \neq \mathbf{0}$ uit een vectorruimte V over $\mathbb{L}$ heet een *eigenvector* van de lineaire operator $\mathcal{L} : V \to V$ als er een scalar λ in $\mathbb{L}$ bestaat zo dat

$$\mathcal{L}(\mathbf{v}) = \lambda\mathbf{v}.$$

De scalar $\lambda \in \mathbb{L}$ wordt de bijbehorende *eigenwaarde* genoemd.

Opmerking: Merk op dat de eigenwaarden en eigenvectoren van een $n\times n$ matrix A ook bestaan. Zijn dat adere objecten? Eigenlijk niet. Als A een reële $n \times n$ matrix dan zijn de eigenvectoren en eigenwaarden van de matrix A zoals gedefinieerd in Lineare Algebra deel 1, precies de eigenwaarden en eigenvectoren van de operator $\mathcal{L}_A : \mathbb{R}^n \to \mathbb{R}^n$ gedefinieerd door $\mathcal{L}_A(\mathbf{v}) = A\mathbf{v}$ (de operator die de matrix A als standaardmatrix heeft).
Als A een complexe $n\times n$ matrix dan zijn de eigenvectoren en eigenwaarden van de matrix A, zoals gedefinieerd in college 1, precies de eigenwaarden en eigenvectoren van de operator $\mathcal{L}_A : \mathbb{C}^n \to \mathbb{C}^n$ gedefinieerd door $\mathcal{L}_A(\mathbf{v}) = A\mathbf{v}$.
De complexe eigenwaarden en eigenvectoren van een reële $n \times n$ matrix A, zoals gedefinieerd in college 1, zijn precies de eigenwaarden en eigenvectoren van de operator $\mathcal{L}_A : \mathbb{C}^n \to \mathbb{C}^n$ gedefinieerd door $\mathcal{L}_A(\mathbf{v}) = A\mathbf{v}$ (een operator op $\mathbb{C}^n$ die een reële matrix als standaardmatrix heeft).

Opmerking: Merk op dat volgens onze definitie een reële operator $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^n$ alleen reële eigenwaarden en eigenvectoren kan hebben.

Eigenwaarden en eigenvectoren spelen in zeer vele toepassingen een centrale rol. U zult dit tegenkomen bij bij het oplossen van stelsels differentiaalvergelijkingen, maar ook bij een vak als kwantummechanica.
Een andere context is het volgende: het komt vaak voor dat een vector een bepaalde toestand voorstelt, bijvoorbeeld als de kentallen van een vector in $\mathbb{R}^n$ de aantallen aanhangers voorstellen van de verschillende politieke partijen, of, in de kwantummechanica, waar een toestand wordt aangegeven door een (vector)functie $\Psi(x,t)$ (wat betekent dat $\|\Psi(x,t)\|^2$ de kans is dat een deeltje zich op tijdstip t in positie x bevindt). Een lineaire operator $\mathcal{L}$ representeert dan een toestandsverandering (bijvoorbeeld een bepaalde gebeurtenis die de aanhangers van een partij naar een andere partij doet overlopen, of in het kwantummechanische voorbeeld de impuls, waarbij de operator gelijk is aan $-i\frac{d}{dx}$ en $\Psi(x,t)$ dus veranderd wordt in $-i\frac{d}{dx}\Psi(x,t)$). Men kan dan geïnteresseerd zijn in een stabiele toestand $\mathbf{v}$, waarin de operator geen effect heeft, dus waarvoor $\mathcal{L}(\mathbf{v}) = \mathbf{v}$, d.w.z. $\mathbf{v}$ is een eigenvector bij de eigenwaarde $\lambda = 1$. [1].

Voorbeeld 3.41. Beschouw de operator $\mathcal{L} : \mathrm{Pol}(\mathbb{R},\mathbb{R}) \to \mathrm{Pol}(\mathbb{R},\mathbb{R})$ die gedefinieerd is door

$$\mathcal{L}(p)(x) = x\frac{dp}{dx} = xp'(x).$$

Het is niet al te moeilijk in te zien dat alle polynomen van de vorm $p(x) = x^k$ eigenvectoren zijn van deze operator $\mathcal{L}$, en wel bij de eigenwaarde k. Er geldt immers dat:

$$\mathcal{L}(p)(x) = x\,k\,x^{k-1} = k\,x^k = \, k\,(p)(x).$$

Inderdaad, $\mathcal{L}(p) = k\,(p)$ en dus is p een eigenvector van de operator $\mathcal{L}$ bij de eigenwaarde k.
De eerste les uit dit voorbeeld is: nagaan of een gegeven vector een eigenvector is, is in het algemeen vrij eenvoudig.
Hoe op het idee te komen dat $p(x) = x^k$ wel eens een eigenvector zou kunnen zijn is natuurlijk een tweede. Deze vraag is natuurlijk een stuk lastiger en we zullen dit beantwoorden via de vraag: "zijn de vectoren $p(x) = x^k$ resp. de getallen $k = 0, 1, 2, \ldots$, alle eigenvectoren, resp alle eigenwaarden, van deze opertor"?
Stel dat p een polynoom is bij de eigenwaarde λ. Dan: $\mathcal{L}(p) = \lambda\,(p)$, dus $\mathcal{L}(p)(x) = \lambda\,(p)(x)$ en dus

$$xp'(x) = \lambda\, p(x).$$

We herkennen dit als een separabele differentiaalvergelijking, en hiermee kunnen we niet alleen de polynomen (wat de vraag is) maar zels alle functies $y(x)$ bepalen met $xy'(x) = \lambda\, y(x)$.

[1] De termen eigenvector, eigenwaarde en eigenfunctie zijn afkomstig uit het Duits en zijn overgenomen uit artikelen over kwantummechanica van omstreeks 1920.

Aldus: $xy'(x) = \lambda\, y(x) \Rightarrow \dfrac{y'(x)}{y(x)} = \dfrac{\lambda}{x} \Rightarrow \displaystyle\int \frac{1}{y}\, dy = \int \frac{\lambda}{x}\, dx \Rightarrow \ln|y| = \lambda \ln|x| + C' = \ln|Cx^{\lambda}|$ en dus:

$$y(x) = Cx^{\lambda}.$$

Conclusie: de differentiaalvergelijking heeft veel meer oplossingen (bijvoorbeeld $y(x) = x^{4/3}$), maar de **polynoomoplossingen** van de differentiaalvergelijking, de eigenvectoren van $\mathcal{L}$, zijn precies de polynomen van de vorm $p(x) = Cx^k$. (Dus ëssentieel"zijn $p(x) = x^k$ precies de eigenwaarden) en $k = 0, 1, 2, \ldots$ zijn precies de eigenwaarden. □

Voor het volgende voorbeeld introduceren we de vectorruimte $\mathbb{L}[[X]]$, de verzameling van alle "formele machtreeksen" van de vorm

$$\sum_{k=0}^{\infty} a_k X^k$$

in de formele variabele X met $a_k \in \mathbb{L}$.
We noemen dit een formele machtreeks omdat we de machtreeks niet als functie zullen beschouwen (dus een vraag over het convergentiegebied zullen we niet stellen) maar puur als een formele oneindige sommatie. We zeggen wel: twee machtreeksen $\displaystyle\sum_{k=0}^{\infty} a_k X^k$ en $\displaystyle\sum_{k=0}^{\infty} b_k X^k$ zijn gelijk als $a_k = b_k$, voor $k = 0, \ldots, \infty$. Met de operaties:

- $\displaystyle\sum_{k=0}^{\infty} a_k X^k + \sum_{k=0}^{\infty} b_k X^k = \sum_{k=0}^{\infty} (a_k + b_k) X^k$,
- $\displaystyle c\sum_{k=0}^{\infty} a_k X^k = \sum_{k=0}^{\infty} (ca_k) X^k$

maken we $\mathbb{L}[[X]]$ inderdaad tot een vectorruimte over $\mathbb{L}$.

Voorbeeld 3.42. Om alle eigenwaarden en eigenvectoren van de differentiatie operator $\mathcal{D} : \mathbb{L}[[X]] \to \mathbb{L}[[X]]$, die gedefinieerd is door

$$\mathcal{D}(\sum_{k=0}^{\infty} a_k X^k) = \sum_{k=1}^{\infty} k a_k X^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} X^k$$

op te zoeken, stellen we dat de niet-nulreeks $\sum_{k=0}^{\infty} a_k X^k$ een eigenvector bij eigenwaarde λ is. Dat betekent dat er geldt

$$\sum_{k=0}^{\infty} (k+1) a_{k+1} X^k = \lambda \sum_{k=0}^{\infty} a_k X^k = \sum_{k=0}^{\infty} \lambda a_k X^k.$$

Hieruit volgt dat de coëfficiënten voor $k \geq 0$ moeten voldoen aan:

$$(k+1) a_{k+1} = \lambda a_k.$$

We herschreven dit tot:

$$a_{k+1} = \frac{\lambda}{k+1} a_k.$$

Hier staat dus dat als $a_1 = \frac{\lambda}{1} a_0$, dat dan: $a_2 = \frac{\lambda}{2} a_1 = \frac{\lambda^2}{2\cdot 1} a_0$, $a_3 = \frac{\lambda}{3} a_2 = \frac{\lambda^3}{3\cdot 2\cdot 1} a_0$, enzovoorts. Bij gegeven a_0 liggen dus alle coëfficiënten van de eigenvector–machtreeks bij λ vast, er geldt namelijk dat

$$a_k = a_0 \frac{\lambda^k}{k!}.$$

Voor elke $a_0 \neq 0$ uit $\mathbb{L}$ is de reeks

$$\sum_{k=0}^{\infty} a_0 \frac{\lambda^k}{k!} X^k = a_0 \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} X^k.$$

dus een eigenvector van $\mathcal{D}$ bij eigenwaarde λ.
De conclusie is dat elke $\lambda \in \mathbb{L}$ een eigenwaarde van $\mathcal{D}$ is en dat de bijbehorende eigenvectoren de scalaire veelvouden van de machtreeks $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} X^k$ zijn, op de nul–machtreeks na. $\square$

Opmerking: We zouden bovenstaande operator ook kunnen bekijken als een operator $\mathcal{D}$: $\mathrm{Pol}(\mathbb{R},\mathbb{R}) \to \mathrm{Pol}(\mathbb{R},\mathbb{R})$ via $\mathcal{D}(p)(x) = p'(x)$. Willen we eigenwaarden en eigenvectoren bekijken dan leidt dit tot de differentiaalvergelijking:

$$y' = \lambda y.$$

Deze heeft geen polynoom als oplossing, dus de operator $\mathcal{D} : \mathrm{Pol}(\mathbb{R},\mathbb{R}) \to \mathrm{Pol}(\mathbb{R},\mathbb{R})$ heeft geen eigenwaarden/eigenvectoren.
De oplossing van de differentiaalvergelijking zijn welbekend, $y(x) = C\, e^{\lambda\, x}$. Merk op dat de machtreeksontwikkeling van deze functies wordt:

$$y(x) = C \sum_{k=0}^{\infty} \frac{(\lambda\, x)^k}{k!} = C \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} x^k.$$

Zie de overeenkomst met de formele machtreeks gevonden in het vorige voorbeeld. $\square$

Vorige voorbeelden wekken de suggestie dat de verzameling van eigenvectoren bij gelijke eigenwaarde een deelruimte is als de nulvector toegevoegd wordt. Dat is ook zo, vandaar de volgende definitie.

Definitie: Als λ een eigenwaarde is van een operator $\mathcal{L}$ op een vectorruimte V over $\mathbb{L}$, dan heet de verzameling E_λ, die bestaat uit alle vectoren $\mathbf{v}$ waarvoor $\mathcal{L}(\mathbf{v}) = \lambda\mathbf{v}$ de *eigenruimte* van $\mathcal{L}$ bij de eigenwaarde λ.

Merk op dat de eigenruimte E_λ bij een eigenwaarde λ uit de bijbehorende eigenvectoren én de nulvector (die geen eigenvector is!) bestaat. Verder geldt dat een vector $\mathbf{v}$ tot E_λ behoort als en slechts als $\mathcal{L}(\mathbf{v}) = \lambda\mathbf{v}$. Dit kan herschreven worden tot $\mathcal{L}(\mathbf{v}) - \lambda\mathbf{v} = \mathbf{0}$. Merk op dat $\mathcal{L}(\mathbf{v}) - \lambda\mathbf{v}$ het beeld is van $\mathbf{v}$ onder de operator $\mathcal{L} - \lambda\,\mathrm{Id}$. Blijkbaar behoort $\mathbf{v}$ tot E_λ als en slechts als $(\mathcal{L} - \lambda\,\mathrm{Id})(\mathbf{v}) = \mathbf{0}$, wat wil zeggen dat

$$E_\lambda = \mathrm{Ker}(\mathcal{L} - \lambda\,\mathrm{Id}). \tag{3.1}$$

Hieruit volgt dat de scalar λ een eigenwaarde is zodra de kern van de operator $\mathcal{L} - \lambda\,\mathrm{Id}$ uit meer dan alleen de nulvector bestaat én er volgt dat eigenruimten van een operator op een vectorruimte V deelruimten van V zijn. Dat betekent dat elke lineaire combinatie van eigenvectoren van een operator $\mathcal{L}$ bij dezelfde eigenwaarde ook weer een eigenvector bij die eigenwaarde is (tenzij het de nulvector is). We formuleren dit alles in een stelling:

STELLING 3.43. *De eigenruimte E_λ bij een eigenwaarde λ van een lineaire operator $\mathcal{L}$ op een vectorruimte V is de kern van de operator $\mathcal{L} - \lambda\,\mathrm{Id}$ en is dus een deelruimte van V. Deze deelruimte bevat precies de nulvector van V én alle eigenvectoren met eigenwaarde λ. Er geldt blijkbaar dat elke lineaire combinatie van eigenvectoren bij eigenwaarde λ ook weer een eigenvector bij eigenwaarde λ is (tenzij het de nulvector is).*

3.7 Bepalen eigenwaarden als V oneindig–dimensionaal is

In feite geeft de definitie of (3.1) aan hoe eigenvectoren bij een *gegeven* eigenwaarde berekend kunnen worden. Zo'n berekening hebben we al gezien in voorbeeld 3.41 en voorbeeld 3.42. De vraag is hoe eigenwaarden gevonden kunnen worden. Voor oneindig dimensionale vectorruimten is geen speciale methode beschikbaar en zal de definitie van de operator gebruikt moeten worden, zoals al gedaan is in voorbeeld 3.41 en in voorbeeld 3.42. Nog een laatste wat lastiger voorbeeld.

Voorbeeld 3.44. Op de vectorruimte $\mathrm{Pol}(\mathbb{R}, \mathbb{C})$ wordt de lineaire operator D gedefinieerd door

$$D(p)(x) = (1 - x^2)\frac{d^2p}{dx^2} - 2x\frac{dp}{dx} = (1 - x^2)p''(x) - 2xp'(x).$$

Vaak wordt dit kortweg genoteerd als

$$D = (1 - x^2)\frac{d^2}{dx^2} - 2x\frac{d}{dx}.$$

De bijbehorende differentiaalvergelijking die de eigenwaarde bepaalt luidt:

$$(1 - x^2)y'' - 2xy' = \lambda\, y$$

en kennelijk zoeken we de polynoom–oplossingen van deze differentiaalvergelijking. Echter, we zijn (nog?) niet in staat deze differentiaalvergelijking op te lossen. Daarom zullen we een ad-hoc methode toepassen gebaseerd op de definitie van de operaotor.
We gaan de eigenwaarden en eigenvectoren berekenen. Neem daarvoor aan dat het niet-nulpolynoom p een eigenvector is van de graad n bij eigenwaarde λ, dus $D(p) = \lambda p$.
Dan zijn er scalairen $a_0, \ldots, a_n$ in $\mathbb{C}$, met $a_n \neq 0$, zo dat

$$p(x) = \sum_{k=0}^{n} a_k x^k \quad \text{met} \quad a_n \neq 0.$$

Ga na dat

$$\begin{aligned}
D(p)(x) &= p''(x) - x^2p''(x) - 2xp'(x)\\
&= \sum_{k=0}^{n-2}(k+2)(k+1)a_{k+2}x^k - \sum_{k=2}^{n} k(k-1)a_kx^k - \sum_{k=1}^{n} 2ka_kx^k\\
&= \sum_{k=0}^{n-2}(k+2)(k+1)a_{k+2}x^k - \sum_{k=0}^{n} k(k-1)a_kx^k - \sum_{k=0}^{n} 2ka_kx^k\\
&= \sum_{k=0}^{n-2}(k+2)(k+1)a_{k+2}x^k - \sum_{k=0}^{n}(k^2+k)a_kx^k
\end{aligned}$$

(hierbij moet aangetekend worden dat de term met $\sum_{k=0}^{n-2}$ alleen aanwezig zijn als $n \geq 2$). Dit moet gelijk zijn aan

$$\lambda p(x) = \sum_{k=0}^{n} \lambda a_k x^k$$

en dat is het geval als de coëfficiënten van x^n tot en met x^0 in $D(p)(x)$ en $\lambda p(x)$ twee aan twee gelijk zijn. Gelijkheid van de coëfficiënten van x^n levert

$$-(n^2 + n)a_n = \lambda a_n$$

en omdat $a_n \neq 0$ volgt hieruit dat $\lambda = -n^2 - n$. Hiermee zijn de eigenwaarden van D gevonden, het zijn precies de getallen:

$$\lambda = -(n^2 + n), \quad \text{voor} \quad n = 0, 1, 2, \ldots.$$

Ook hebben we al gezien dat een eigenvector bij $\lambda = -(n^2 + n)$ een n^{e} graads polynoom moet zijn. Nu kunnen ook de bijbehorende eigenvectoren nog bepaald worden. Daarvoor moeten we verder gaan met het gelijkstellen van de diverse coëfficiënten. Gelijkheid van die van x^{n-1} (die overigens alleen voorkomt als $n > 0$) levert

$$-\big((n-1)^2 + (n-1)\big)a_{n-1} = \lambda a_{n-1}.$$

en dus: $-(n^2 - 2n)a_{n-1} = \lambda a_{n-1}$. Omdat $\lambda = -(n^2 + n)$ volgt hieruit dat $na_{n-1} = 0$ en omdat $n > 0$ geldt dus $a_{n-1} = 0$. Voor de coëfficiënten van alle andere machten x^k met $k = 0, \ldots, n-2$ moet vervolgens gelden dat

$$(k+2)(k+1)a_{k+2} - (k^2 + k)a_k = \lambda a_k$$

dus

$$a_k = \frac{(k+2)(k+1)}{(k^2 + k) + \lambda} a_{k+2}.$$

Dit is een *recurrente betrekking* (een relatie die een verband legt tussen een willekeurige term van de rij en één of meer voorgaande termen). Met zo'n betrekking is de rij te construeren vanuit één of meer gegeven termen van de rij.
Hier kan met gegeven a_n achtereenvolgens a_{n-2}, a_{n-4}, enzovoorts berekend worden.
Omdat bekend is dat $a_{n-1} = 0$, kunnen tevens a_{n-3}, a_{n-5}, enzovoorts berekend worden en er volgt dat deze laatste allemaal nul zijn. De eigenvectoren zijn dus polynomen met óf alleen even machten (als het polynoom even graad heeft), óf alleen oneven machten (als het polynoom oneven graad heeft). Hieronder zijn de eigenwaarden en bijbehorende "eigenpolynomen" uitgerekend voor $n = 0$, 1, 2, 3 en 4 (door telkens een willekeurige a_n te nemen, waarmee vervolgens a_{n-2}, a_{n-4}, enzovoorts, uitgerekend kunnen worden; de andere coëfficiënten zijn, zoals hierboven vermeld, gelijk aan nul):

$$\begin{array}{lll}
n = 0: & \lambda = 0; & p(x) = a_0 \\
n = 1: & \lambda = -2; & p(x) = a_1 x \\
n = 2: & \lambda = -6; & p(x) = a_2(x^2 - \frac{1}{3}) = \frac{a_2}{3}(3x^2 - 1) \\
n = 3: & \lambda = -12; & p(x) = a_3(x^3 - \frac{3}{5}x) = \frac{a_3}{5}(5x^3 - 3x) \\
n = 4: & \lambda = -20; & p(x) = a_4(x^4 - \frac{6}{7}x^2 + \frac{6}{70}) = \frac{a_4}{35}(35x^4 - 30x^2 + 3)
\end{array}$$

We eindigen dit voorbeeld met de opmerking dat deze "eigenpolynomen" bij geschikte keuze van a_0, a_1, a_2, …, de *Legendre polynomen* P_n genoemd worden. De genoemde keuze moet zo zijn dat voor alle gevonden polynomen geldt dat $p(1) = 1$. Voor bijvoorbeeld a_4 moet dan $35/8$ genomen worden en P_4 is dan $\frac{1}{8}(35x^4 - 30x^2 + 3)$. Later komen we terug op Legendre polynomen.

3.8 Bepalen van eigenwaarden: het eindig dimensionale geval

In het eindig dimensionale geval kunnen we het begrip diagonaliseerbaar introduceren.

Definitie: Een lineaire operator $\mathcal{L} : V \to V$ heet diagonaliseerbaar als V een basis B heeft bestaande uit eigenvectoren van $\mathcal{L}$.

De volgende stelling verklaart de naam diagonaliseerbaar.

STELLING 3.45. *Beschouw de lineaire operator $\mathcal{L} : V \to V$. Equivalent zijn:*

1. *De operator $\mathcal{L} : V \to V$ is diagonaliseerbaar.*
2. *Er is een basis B op V zodat $\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$ een diagonaalmatrix is.*

Bewijs: Stel (1) is gegeven. Dat wil zeggen dat er een basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ van V bestaat met $\mathbf{b}_i$ een eigenvector van $\mathcal{L}$, zeg $\mathcal{L}(\mathbf{b}_i) = \lambda_i \mathbf{b}_i$. Dan geldt kennelijk $[\mathcal{L}(\mathbf{b}_i)]_B = [\lambda_i \mathbf{b}_i]_B = \lambda_i \mathbf{e}_i$. Maar dan is kennelijk $\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]} = [[\mathcal{L}(\mathbf{b}_1)]_B \ldots [\mathcal{L}(\mathbf{b}_n)]_B]$ een diagonaalmatrix met op de diagonaal de getallen $\lambda_1, \ldots, \lambda_n$, precies de eigenwaarden!
Omgekeerd, stel dat (2) gegeven is, zeg dat $\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$ een diagonaalmatrix is met op de diagonaal de getallen $d_1, \ldots, d_n$. Dan geldt kennelijk voor de i^{de} kolom $[\mathcal{L}(\mathbf{b}_i)]_B$ van deze matrix: $[\mathcal{L}(\mathbf{b}_i)]_B = d_i\, \mathbf{e}_i$ en dus: $\mathcal{L}(\mathbf{b}_i) = d_i\, \mathbf{b}_i$ (voor $i = 1, \ldots n$). Inderdaad, de basisvectoren uit de basis B zijn eigenvectoren. $\square$

Het bepalen van de eigenwaarden en eigenvectoren van een operator $\mathcal{L}$ op een *eindig dimensionale* vectorruimte V gaat als volgt: kies namelijk een basis $\mathcal{B}$ voor V en maak de matrix $A = \underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$. De eigenwaarden van $\mathcal{L}$ zijn dan volgens de volgende stelling precies de eigenwaarden van de matrix A.

STELLING 3.46. *Als $\mathcal{L}$ een lineaire operator op een eindig dimensionale vectorruimte V is en als $\mathcal{B}$ een basis is voor V, dan geldt:*

- *λ is een eigenwaarde van de operator $\mathcal{L}$ als en slechts als λ een eigenwaarde is van de matrix $A = \underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$ en*
- *$\mathbf{v}$ is een eigenvector van de operator $\mathcal{L}$ als en slechts als $[\mathbf{v}]_{\mathcal{B}}$ een eigenvector is van de matrix $A = \underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$*

Bewijs: Er geldt

$$\mathcal{L}(\mathbf{v}) = \lambda\mathbf{v} \iff [\mathcal{L}(\mathbf{v})]_{\mathcal{B}} = [\lambda\mathbf{v}]_{\mathcal{B}} \iff \underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}[\mathbf{v}]_{\mathcal{B}} = \lambda[\mathbf{v}]_{\mathcal{B}} \iff A[\mathbf{v}]_{\mathcal{B}} = \lambda[\mathbf{v}]_{\mathcal{B}}$$

Omdat ook nog geldt dat $\mathbf{v} \neq \mathbf{0}$ als en slechts als $[\mathbf{v}]_{\mathcal{B}} \neq \mathbf{0}$, volgt het gestelde. $\square$

Tenslotte verkrijgen we:

STELLING 3.47. *Stel dat $\mathcal{L} : V \to V$ een lineaire operator op een vectorruimte V met $\dim(V) = n < \infty$. Als $\mathcal{B}$ een basis is voor V, dan geldt dat equivalent zijn:*

1. *De operator $\mathcal{L} : V \to V$ is diagonaliseerbaar.*
2. *De matrix $\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$ is diagonaliseerbaar.*
3. $\displaystyle\sum_{\lambda} \dim E_\lambda(\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}) = n$
4. $\displaystyle\sum_{\lambda} \dim E_\lambda(\mathcal{L}) = n$

Bewijs: Stel dat (1) geldt. Laat C een basis zijn voor V bestaande uit eigenvectoren. Dan is $\underset{\mathcal{C}\leftarrow\mathcal{C}}{[\mathcal{L}]}$ kennelijk een diagonaalmatrix, zeg dat $D = \underset{\mathcal{C}\leftarrow\mathcal{C}}{[\mathcal{L}]}$. Volgens stelling 3.39 geldt

$$\underset{B\leftarrow B}{[\mathcal{L}]} = \underset{B\leftarrow C}{P}\ \underset{C\leftarrow C}{[\mathcal{L}]}\ \underset{C\leftarrow B}{P}.$$

en dus: $\underset{B\leftarrow B}{[\mathcal{L}]} = PDP^{-1}$, d.w.z. de matrix $\underset{B\leftarrow B}{[\mathcal{L}]}$ is diagonaliseerbaar.
Omgekeerd, stel (2), dus stel dat de matrix $\underset{B\leftarrow B}{[\mathcal{L}]}$ diagonaliseerbaar is, d.w.z. er is een basis $\{\mathbf{p}_1, \ldots, \mathbf{p}_n\}$ van $\mathbb{R}^n$ bestaande uit eigenvectoren van de matrix $\underset{B\leftarrow B}{[\mathcal{L}]}$. Als $\underset{B\leftarrow B}{[\mathcal{L}]}\mathbf{p}_i = d_i\mathbf{p}_i$, $P = [\mathbf{p}_1 \ldots \mathbf{p}_n]$ en $D = \text{diag}(d_1, \ldots, d_n)$ dan geldt: $\underset{B\leftarrow B}{[\mathcal{L}]} = PDP^{-1}$. Kies $\mathbf{v}_i \in V$ met $[\mathbf{v}_i]_B = \mathbf{p}_i$ en stel dat $C = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Leg zelf uit dat C een basis van V is (gebruik het inverteerbaar zijn van de matrix P). Merk op dat $\underset{B\leftarrow C}{P} = [[\mathbf{v}_1]_B \ldots [\mathbf{v}_n]_B] = [\mathbf{p}_1 \ldots = \mathbf{p}_n] = P$ (!). Maar dan via stelling 3.39:

$$\underset{C\leftarrow C}{[\mathcal{L}]} = \underset{C\leftarrow B}{P}\ \underset{B\leftarrow B}{[\mathcal{L}]}\ \underset{B\leftarrow C}{P} = P^{-1}\underset{B\leftarrow B}{[\mathcal{L}]}P = D.$$

Kortom, C is een basis van V met $\underset{C\leftarrow C}{[\mathcal{L}]}$ een diagonaalmatrix en dus is de operator $\mathcal{L} : V \to V$ diagonaliseerbaar.
De equivalentie van (2) en (3) komt van de ons bekende uitspraak dat een $n \times n$ matrix (in dit geval $\underset{B\leftarrow B}{[\mathcal{L}]}$) diagonaliseerbaar is als en slechts als de som van de meetkundige multipliciteiten gelijk is aan n.
De equivalentie van (3) en (4) is de gelijkheid van dimensie van de eigenruimte E_λ van de matrix $\underset{B\leftarrow B}{[\mathcal{L}]}$ (die we hier aanduiden met $E_\lambda(\underset{B\leftarrow B}{[\mathcal{L}]})$) en de dimensie van de eigenruimte E_λ van de operator $\mathcal{L}$ (die we hier aanduiden met $\dim E_\lambda(\mathcal{L})$). $\square$

Voorbeeld 3.48. Definieer de lineaire operator $\mathcal{L} : \mathrm{Pol}_2(\mathbb{R}) \to \mathrm{Pol}_2(\mathbb{R})$ door

$$\mathcal{L}(p)(x) = p(1) + p(x) + (1+x)p'(x).$$

De eigenwaarden van deze operator zijn precies de eigenwaarden van de matrix A van $\mathcal{L}$ op een basis $\mathcal{B}$ voor $\mathrm{Pol}_2(\mathbb{R})$. We kunnen voor $\mathcal{B}$ de standaardbasis $\{1, x, x^2\}$ nemen. Ga na dat dan

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

en omdat dit een bovendriehoeksmatrix is, staan de eigenwaarden van deze matrix op de diagonaal (zie lineaire algebra deel 1). De eigenwaarden van $\mathcal{L}$ zijn dus 2 en 3, dezelfde als die van A.
Nu de eigenwaarden bekend zijn kunnen de bijbehorende eigenruimten E_2 en E_3 bepaald worden. We hebben al eerder gezien dat dit de oplossingsverzamelingen zijn van de vergelijkingen $\mathcal{L}(p) = 2p$ en $\mathcal{L}(p) = 3p$, ofwel $E_2 = \mathrm{Ker}(\mathcal{L} - 2\,\mathrm{Id})$ en $E_3 = \mathrm{Ker}(\mathcal{L} - 3\,\mathrm{Id})$. Het bepalen hiervan kan handig gedaan worden via coördinatisering ten opzichte van de basis $\mathcal{B}$: bereken de eigenruimten $E_2 = \mathrm{Nul}(A - 2I)$ en $E_3 = \mathrm{Nul}(A - 3I)$ van A. Deze bestaan uit de eigenvectoren van A en dit zijn volgens de zojuist behandelde stelling coördinatiseringen van de eigenvectoren van $\mathcal{L}$. Als de gevonden vectoren uit $\mathbb{R}^3$ dus omgezet worden naar polynomen, dan zijn de eigenvectoren van $\mathcal{L}$ gevonden.
Ga na dat

$$\mathrm{Nul}(A - 2I) = \mathrm{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\} \quad \text{en} \quad \mathrm{Nul}(A - 3I) = \mathrm{Span}\left\{\begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}\right\}.$$

Terugvertalen naar polynomen levert de eigenruimten

$$E_2 = \mathrm{Span}\{1\} \quad \text{en} \quad E_3 = \mathrm{Span}\{5 + 2X + X^2\}.$$

van $\mathcal{L}$. Merk op dat A en dus ook $\mathcal{L}$ niet diagonaliseerbaar zijn. Inderdaad, de meetkundige multipliciteit van $\lambda = 2$ is kleiner dan de algebraïsche multipliciteit.

Tenslotte nog de volgende opmerking. Laat $\mathcal{L} : V \to V$ een lineaire operator zijn op V en neem aan dat op V de bases B_1 en B_2 voorhande zijn. Volgens Steliing 3.39 wordt het verband tussen de verschillende matrices voor $\mathcal{L}$, de matrices $\underset{B_1 \leftarrow B_1}{[\mathcal{L}]}$ en $\underset{B_2 \leftarrow B_2}{[\mathcal{L}]}$ dan:

$$\underset{B_2 \leftarrow B_2}{[\mathcal{L}]} = \underset{B_2 \leftarrow B_1}{P} \underset{B_1 \leftarrow B_1}{[\mathcal{L}]} \underset{B_1 \leftarrow B_2}{P}.$$

Schrijven we $P = \underset{B_2 \leftarrow B_1}{P}$, dan geldt $\underset{B_1 \leftarrow B_2}{P} = P^{-1}$ en dus

$$\underset{B_2 \leftarrow B_2}{[\mathcal{L}]} = P \underset{B_1 \leftarrow B_1}{[\mathcal{L}]} P^{-1}.$$

De matrices $\underset{B_2 \leftarrow B_2}{[\mathcal{L}]}$ en $\underset{B_1 \leftarrow B_1}{[\mathcal{L}]}$ voldoen aan een speciale relatie, namelijk die van de vorige formule. Matrices die aan zo'n relatie voldoen werden in deel 1 gelijkvormig genoemd.
We kunnen nu concluderen: bij een lineaire operator zijn alle matrices, die we kunnen maken door een basis te kiezen, gelijkvormig. Omgekeerd, als een matrix A gelijkvormig is met $\underset{B_1 \leftarrow B_1}{[\mathcal{L}]}$, zeg $A = P \underset{B_1 \leftarrow B_1}{[\mathcal{L}]} P^{-1}$ voor een zekere matrix P, dan kunnen we P beschouwen als een overgangsmatrix, waarbij in de kolommen de coördinaten staan van de oude basis B_1 ten opzichte van een nieuwe basis B_2 (en dus staan in P^{-1} de coördinaten van de nieuwe basis B_2 ten opzichte van de oude basis B_1). De matrix A is dan de matrix van $\mathcal{L}$ ten opzichte van deze nieuwe basis B_2. We vinden daarmee:

$$A \text{ is de matrix van } \mathcal{L} \text{ ten opzichte van een basis } B_2 \iff A \text{ is gelijkvormig met } \underset{B_1 \leftarrow B_1}{[\mathcal{L}]}.$$

Hoofdstuk 4

$\mathbb{C}^n$ met zijn inwendigproduct structuur

4.1 Het standaard Inwendig Product op $\mathbb{C}^n$

De behandelde theorie wat betreft het standaard inwendig product op de Euclidische ruimte $\mathbb{R}^n$ gaat niet geheel door voor de unitaire ruimte $\mathbb{C}^n$. We willen namelijk dat een inwendig product op $\mathbb{C}^n$ de lengte van een vector bepaald door de wortel van het inwendig product van die vector met zichzelf, dus net als in $\mathbb{R}^n$:

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

Hieruit volgt al dat het standaard inwendige product uit $\mathbb{R}^n$ niet zomaar overgenomen kan worden naar $\mathbb{C}^n$, want we zien als we dit zouden doen dat bij de vector

$$\mathbf{z} = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

zou gelden dat:

$$\langle \mathbf{z}, \mathbf{z} \rangle = 1 \cdot 1 + i \cdot i = 0.$$

Merk op dat in $\mathbb{C}^1 = \mathbb{C}$ het begrip lengte ons bekend via de formule:

$$|z|^2 = \overline{z}z.$$

Dit motiveert de volgende definitie (waarbij we opmerken dat het standaard inwendige product op $\mathbb{C}^n$ genoteerd zal worden met $\langle , \rangle$):

Definitie: Laat

$$\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \quad \text{en} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

twee vectoren in $\mathbb{C}^n$ zijn. Het **(standaard) inwendig product** van $\mathbf{u}$ en $\mathbf{v}$ is het complexe getal

$$\langle \mathbf{u}, \mathbf{v} \rangle = \overline{u}_1 v_1 + \overline{u}_2 v_2 + \cdots + \overline{u}_n v_n.$$

Opmerking: Jammer genoeg wordt in de literatuur niet eenduidig het standaard inwendig product gekozen. Sommige kiezen voor conjugatie op het tweede kental, dus $\langle \mathbf{u}, \mathbf{v} \rangle = u_1\overline{v}_1 + u_2\overline{v}_2 + \cdots + u_n\overline{v}_n$. Anderen, zoals wij, kiezen voor de conjugatie op het EERSTE kental, nogmaals:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \overline{u}_1 v_1 + \overline{u}_2 v_2 + \cdots + \overline{u}_n v_n.$$

Voorbeeld 4.1. Als $\mathbf{z} = \begin{bmatrix} 1 \\ i \end{bmatrix}$ en $\mathbf{w} = \begin{bmatrix} 1+i \\ 3i \end{bmatrix}$ dan geldt

$$\langle \mathbf{z}, \mathbf{z} \rangle = 1 \cdot 1 + (-i) \cdot i = 2, \qquad \langle \mathbf{w}, \mathbf{w} \rangle = (1-i) \cdot (1+i) + (-3i) \cdot 3i = 11,$$

en

$$\langle \mathbf{z}, \mathbf{w} \rangle = 1 \cdot (1+i) + (-i) \cdot 3i = 4 + i, \qquad \langle \mathbf{w}, \mathbf{z} \rangle = (1-i) \cdot 1 + (-3i) \cdot i = 4 - i.$$

Het lijkt erop dat het inwendig product van een vector met zichzelf inderdaad reëel en niet-negatief is.

Herinner: in $\mathbb{R}^n$ kon het standaardinwendig product geschreven kan worden als een matrixproduct via $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y}$. Hetzelfde geldt in $\mathbb{C}^n$, want:

$$\boxed{\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^*\mathbf{v}}$$

Dat wil zeggen dat ook dit inwendig product een matrix product is, en sommige van de rekenregels volgen rechtstreeks uit de rekenregels voor producten van matrices.

Het valt in bovenstaand voorbeeld op dat het standaard inwendig product op $\mathbb{C}^n$ niet commutatief is, dat wil zeggen er geldt niet $\langle \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle$. Dit in tegenstelling tot het standaard inwendig product op $\mathbb{R}^n$. Blijkbaar heeft dit inwendig product niet dezelfde eigenschappen (=rekenregels) als het inwendig product op $\mathbb{R}^n$. We formuleren de eigenschappen in de volgende stelling. Vergelijk ze met de eigenschapppen voor het standaard inwendig product op $\mathbb{R}^n$.

STELLING 4.2 (Rekenregels van het standaard inwendig product op $\mathbb{C}^n$)**.** *Als* $\mathbf{u}$*,* $\mathbf{v}$ *en* $\mathbf{w}$ *vectoren in* $\mathbb{C}^n$ *zijn en als* c *een complexe scalar is, dan geldt:*

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$,
2. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$,
3. $\langle \mathbf{u}, c\mathbf{v} \rangle = c \langle \mathbf{u}, \mathbf{v} \rangle$,
4. $\langle \mathbf{u}, \mathbf{u} \rangle \in \mathbb{R}$ *en* $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$. *Bovendien:* $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ *als en slechts als* $\mathbf{u} = \mathbf{0}$.

Voor het bewijs van deze stelling verwijzen we naar de opgaven.

Merk op dat het inwendig product inderdaad aan de eerder genoemde twee eisen voldoet. Verder geldt dat het inwendig product "lineair in de tweede factor" is (zie rekenregels 2. en 3.). Het inwendig product op $\mathbb{R}^n$ is ook lineair in de eerste factor. Het volgende gevolg laat zien dat dit op $\mathbb{C}^n$ *niet* het geval is:

GEVOLG 4.3. *Laat* $\mathbf{u}$, $\mathbf{v}$ *en* $\mathbf{w}$ *vectoren in* $\mathbb{C}^n$ *zijn en laat* c *een complexe scalar zijn. Dan geldt:*

1. $\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle$,
2. $\langle c\mathbf{u}, \mathbf{v}\rangle = \overline{c}\,\langle \mathbf{u}, \mathbf{v}\rangle$,

Ook voor dit bewijs verwijzen we naar de opgaven.

We definieren nu de **lengte** of **norm** $\|\mathbf{v}\|$ van een vector $\mathbf{v}$ in $\mathbb{C}^n$ door

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}.$$

Vanwege eigenschap 4 is de wortel goed gedefinieerd. Toon zelf aan dat:

$$\|c\mathbf{v}\| = |c|\,\|\mathbf{v}\|$$

Vectoren van lengte 1 heten weer **eenheidsvectoren**. Vectoren kunnen genormeerd worden, d.w.z. bij iedere vector $\mathbf{z}$ bestaat een scalair veelvoud van die vector met lengte 1, en wel:

$$\frac{1}{\|\mathbf{z}\|}z$$

In het reële geval bestonden er *twee* scalaire veelvouden van $\mathbf{z}$ met lengte 1, namelijk $\dfrac{\pm 1}{\|\mathbf{z}\|}z$. In het complexe geval zijn dit er veel meer, namelijke alle vectoren van de vorm

$$\alpha\frac{1}{\|\mathbf{z}\|}z \quad (\alpha \in \mathbb{C} \ \text{ met } \ |\alpha| = 1).$$

Hoeken tussen vectoren kunnen in $\mathbb{C}^n$ niet op een zinvolle manier gedefinieërd worden. Alleen de hoek van 90 graden blijkt van belang te zijn.

Definitie: De vectoren $\mathbf{u}$ en $\mathbf{v}$ in $\mathbb{C}^n$ heten **orthogonaal** (of ook loodrecht) als hun inwendig product nul is, dus als $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

Notatie: $\mathbf{u}\perp\mathbf{v}$ als $\mathbf{u}$ en $\mathbf{v}$ orthogonaal zijn.

Merk op dat als geldt dat $\mathbf{u}$ en $\mathbf{v}$ orthogonaal zijn, dus $\langle \mathbf{u}, \mathbf{v}\rangle = 0$, dat dan ook geldt dat $\langle \mathbf{v}, \mathbf{u}\rangle = \overline{\langle \mathbf{u}, \mathbf{v}\rangle} = 0$. Dus als $\mathbf{u}$ en $\mathbf{v}$ orthogonaal zijn, dan zijn $\mathbf{v}$ en $\mathbf{u}$ ook orthogonaal (gelukkig wel).

Tenslotte: de drie basisstellingen voor het inwendig product gelden ook in $\mathbb{C}^n$. Om precies te zijn:

STELLING 4.4. (Cauchy-Schwarz)
Als $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ *dan:*

$$|\,\langle \mathbf{u}, \mathbf{v}\rangle\,| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$$

(De ongelijkheid van Cauchy-Schwarz zal later bewezen worden.)

Tenslotte wordt op $\mathbb{C}^n$ een afstandsbegrip tussen vectoren ingevoerd evenals op $\mathbb{R}^n$. We zeggen **afstand** tussen de vectoren $\mathbf{u}, \mathbf{v}$ is het getal:

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

De belangrijkste ongelijkheid bij het afstandsbegrip luidt alsvolgt:

STELLING 4.5. (Driehoeks-ongelijkheid)
Als $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ *dan:*

$$\|\mathbf{u} - \mathbf{v}\| \leq \|\mathbf{u} - \mathbf{w}\| + \|\mathbf{w} - \mathbf{v}\|$$

Ook van de driehoeksongelijkheid kan in college 13 een bewijs gevonden worden. Tenslotte formulieren we nog de stelling van Pythagoras.

STELLING 4.6. (De stelling van Pythagoras)
Als $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ *met* $\mathbf{u} \perp \mathbf{v}$ *dan:*

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

4.2 Orthogonaliteit in $\mathbb{C}^n$

Als in deel 1 definiëren we:

Definitie: Als $W \subset \mathbb{C}^n$ een deelruimte is van $\mathbb{C}^n$, dan is het **orthogonaal complement** van W:

$$\{\mathbf{z} \in \mathbb{C}^n : \langle \mathbf{z}, \mathbf{w} \rangle = 0, \text{ voor alle } \mathbf{w} \in W\}$$

en deze verzameling wordt genoteerd met $W^{\perp}$.

In de unitaire ruimte $\mathbb{C}^n$ gelden dezelfde resultaten als in $\mathbb{R}^n$ wat betreft $W^{\perp}$ (maar een bewijs is natuurlijk wel nodig).

STELLING 4.7. *Laat W een lineaire deelruimte zijn van $\mathbb{C}^n$. Dan geldt:*

1. *Als $W = \text{Span}\{\mathbf{a}_1 \dots \mathbf{a}_k\}$, dan:*

$$\mathbf{z} \in W^\perp \quad \textit{als en slechts als} \quad \langle \mathbf{z}, \mathbf{a}_i \rangle = 0, \quad \textit{voor} \quad i = 1 \dots k.$$

2. *$W^\perp$ is ook een lineaire deelruimte van $\mathbb{C}^n$.*
3. *$(W^\perp)^\perp = W$.*
4. *$W \cap W^\perp = \{\mathbf{0}\}$.*

Herinner dat in $\mathbb{R}^n$ geldt: $\text{Col}(A)^\perp = \text{Nul}(A^T)$. Dit geldt niet in $\mathbb{C}^n$! De volgende stelling demonstreert een fenomeen wat we nog vaak zullen tegenkomen en wel:

"als in een algemene reële formule A^T staat, dan wordt in de algemene complexe formule transponeren vervangen door Hermitisch transponeren!"

Inderdaad, dit zagen we al bij het standaard inwendigproduct. Op $\mathbb{R}^n$ geldt: $\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T\mathbf{b}$ maar op $\mathbb{C}^n$ wordt dit: $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^*\mathbf{b}$.

STELLING 4.8. *Laat A een complexe $n \times m$ matrix zijn. dan geldt:*

1. *$\text{Col}(A)^\perp = \text{Nul}(A^*)$,*
2. *$\text{Nul}(A)^\perp = \text{Col}(A^*)$.*

Met behulp van deze stelling kunnen we van iedere gegeven lineaire deelruimte W van $\mathbb{C}^n$ een basis bepalen voor zijn orthogonaal complement. Tevens volgt:

$$\dim(W) + \dim(W^\perp) = n.$$

Vervolgens presenteren we §9.2 (die voor $\mathbb{R}^n$ was) voor $\mathbb{C}^n$.

Definitie: Een verzameling (of stelsel) vectoren $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ in $\mathbb{C}^n$ heet **orthogonaal** als

$$\mathbf{u}_i \perp \mathbf{u}_j, \quad \text{voor alle } i, j = 1, \dots, k \text{ met } i \neq j.$$

Die verzameling vectoren heet **orthonormaal** als deze orthogonaal is en bovendien geldt: $\|\mathbf{u}_i\| = 1$, voor $i = 1, \dots, k$.

Net als in $\mathbb{R}^n$ geldt:

STELLING 4.9. *Een orthogonaal stelsel* $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ *met* $\mathbf{u}_i \neq \mathbf{0}$ *(*$i = 1, \ldots, k$*) in* $\mathbb{C}^n$ *is lineair onafhankelijk.*

Ook geldt dat ieder orthogonaal stelsel niet nulvectoren te normeren is tot een orthonormaal stelsel!

Definitie: Laat W een lineaire deelruimte van $\mathbb{C}^n$ zijn met basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$. De basis $\mathcal{B}$ heet een **orthogonale basis** als $\mathcal{B}$ bovendien (behalve het basis zijn) een orthogonale verzameling is. Evenzo: de basis $\mathcal{B}$ heet een **orthonormale basis** (of **unitaire basis**) als $\mathcal{B}$ bovendien een orthonormale verzameling is.

Definitie: Laat $\mathbf{a}_1 \cdots \mathbf{a}_k$ enkele vectoren uit $\mathbb{C}^n$ zijn en $A = [\,\mathbf{a}_1 \cdots \mathbf{a}_k\,]$. De $k \times k$ matrix

$$A^*A$$

heet **de Grammatrix** van de vectoren $\mathbf{a}_1 \cdots \mathbf{a}_k$.

Opmerking: Wellicht onnodig om op te merken, maar mochten $\mathbf{a}_1 \cdots \mathbf{a}_k$ vectoren uit $\mathbb{R}^n$ zijn, dan is de Grammatrix (=namelijk A^TA), zoals gedefiniëerd in §9.2, gelijk aan de Grammatrix zoals gedefiniëerd in deze paragraaf.

STELLING 4.10. *Als* $\mathbf{a}_1 \cdots \mathbf{a}_k$ *enkele vectoren uit* $\mathbb{C}^n$ *zijn en* $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_k \end{bmatrix}$ *dan:*

$$A^*A = \begin{bmatrix} \langle \mathbf{a}_1, \mathbf{a}_1 \rangle & \langle \mathbf{a}_1, \mathbf{a}_2 \rangle & \cdots & \langle \mathbf{a}_1, \mathbf{a}_k \rangle \\ \langle \mathbf{a}_2, \mathbf{a}_1 \rangle & \langle \mathbf{a}_2, \mathbf{a}_2 \rangle & \cdots & \langle \mathbf{a}_2, \mathbf{a}_k \rangle \\ \vdots & \vdots & & \vdots \\ \langle \mathbf{a}_k, \mathbf{a}_1 \rangle & \langle \mathbf{a}_k, \mathbf{a}_2 \rangle & \cdots & \langle \mathbf{a}_k, \mathbf{a}_k \rangle \end{bmatrix}$$

Via deze stelling verkrijgen we:

GEVOLG 4.11. *Als A een matrix is, dan:*

1. *Een (complexe) $m \times n$-matrix A heeft orthogonale kolommen als en slechts als A^*A een diagonaalmatrix is.*
2. *Een (complexe) $m \times n$-matrix A heeft orthonormale kolommen als en slechts als $A^*A = I$.*

Dit leidt tot het begrip unitaire matrix, het complexe equivalent van "orthogonale matrix".

Definitie: Laat A een (complexe) matrix zijn. A heet **unitair** (een unitaire matrix) als A **vierkant** is en als $A^*A = I$.

Unitaire matrices hebben de volgende eigenschappen (vergelijk de situatie met reële orthogonale matrices).

STELLING 4.12. *Voor een* ***vierkante*** *(complexe) matrix U zijn de volgende uitspraken equivalent:*

1. *U is unitair,*
2. *$U^*U = I$,*
3. *De kolommen van U zijn orthonormaal,*
4. *De kolommen van U vormen een orthonormale basis voor $\mathbb{C}^n$,*
5. *U is inverteerbaar en $U^{-1} = U^*$,*
6. *$UU^* = I$,*
7. *De rijen van U zijn orthonormaal,*
8. *De rijen van U vormen een orthonormale basis voor $\mathbb{C}^n$,*
9. *U^* is unitair.*

Voor het bewijs verwijzen we weer naar de opgaven.

Afspraak: als we over orthogonale matrices spreken dan zullen we het altijd over reële matrices hebben. Een unitaire matrix is complex en mag dus ook reëel zijn.

De volgende stellingen mag u in de opgaven over dit college bewijzen.

STELLING 4.13. *Het product van twee unitaire matrices is weer unitair.*

STELLING 4.14. *Laat U een unitaire $n \times n$ matrix zijn en $\mathbf{z}, \mathbf{w}$ vectoren in $\mathbb{C}^n$. Dan geldt:*

1. $\|\mathbf{z}\| = \|U\mathbf{z}\|$.
2. $\langle \mathbf{z}, \mathbf{w} \rangle = \langle U\mathbf{z}, U\mathbf{w} \rangle$.

4.3 Spectraal eigenschappen Speciale Matrices

Zoals we (horen te) weten is *het spectrum* van een matrix gelijk aan de verzameling eigenwaarden van die matrix. Spectrale eigenschappen van een matrix zijn dus eigenschappen die de eigenwaarden van die matrix betreffen. Ook worden hiermee vaak eigenschappen van de eigenruimten van die matrix bedoeld.
We gaan eerst spectrale eigenschappen van complexe matrices bekijken en die toepassen in het reële geval. We gaan eerst Hermitese en/of scheef Hermitese matrices bestuderen, en het zal blijken dat eigenwaarden en eigenvectoren van deze speciale matrices mooie eigenschappen hebben. Al de komende bewijzen zijn gebaseerd op de identiteit

$$\langle Ax, y \rangle = \langle x, A^*y \rangle$$

(Deze identiteit is bijna triviaal: $\langle Ax, y \rangle = (Ax)^*y = (x^*A^*)y = x^*(A^*y) = \langle x, A^*y \rangle$)
We beginnen met de Hermitese matrices:

STELLING 4.15. *Laat A een Hermitese $n \times n$-matrix zijn. Dan geldt*

1. *A heeft slechts reële eigenwaarden,*
2. *eigenvectoren bij verschillende eigenwaarden zijn orthogonaal.*

Bewijs: 1. We merken eerst op dat A Hermitisch is, d.w.z. $A^* = A$ en dit levert nu dat geldt:

$$\langle \mathbf{z}, A\mathbf{w} \rangle = \langle A\mathbf{z}, \mathbf{w} \rangle, \quad \text{voor alle } \ \mathbf{z}, \mathbf{w} \in \mathbb{C}^n.$$

Laat $\lambda \in \mathbb{C}$ een eigenwaarde zijn en $\mathbf{x} \in \mathbb{C}^n$ met $\mathbf{x} \neq \mathbf{0}$ een bijbehorende eigenvector, dus $A\mathbf{x} = \lambda\mathbf{x}$. Dan geldt:

$$\langle A\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{x}, A\mathbf{x} \rangle \Rightarrow \langle \lambda\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{x}, \lambda\mathbf{x} \rangle \Rightarrow \overline{\lambda} \langle \mathbf{x}, \mathbf{x} \rangle = \lambda \langle \mathbf{x}, \mathbf{x} \rangle \Rightarrow \overline{\lambda}\|\mathbf{x}\|^2 = \lambda\|\mathbf{x}\|^2$$

Dus $\lambda\|\mathbf{x}\|^2 = \overline{\lambda}\|\mathbf{x}\|^2$ en omdat $\mathbf{x} \neq \mathbf{0}$ (dus $\|\mathbf{x}\| \neq 0$) volgt $\lambda = \overline{\lambda}$, hetgeen betekent dat λ reëel is.
2. Laat $\mathbf{x}_1 \neq \mathbf{0}$ een eigenvector bij eigenwaarde λ_1 zijn en $\mathbf{x}_2 \neq \mathbf{0}$ een eigenvector bij eigenwaarde λ_2 zijn, waarbij $\lambda_1 \neq \lambda_2$. Dan geldt (op dezelfde manier als bij het bewijs van 1.):

$$\langle A\mathbf{x}_1, \mathbf{x}_2 \rangle = \langle \mathbf{x}_1, A\mathbf{x}_2 \rangle \Rightarrow \langle \lambda_1\mathbf{x}_1, \mathbf{x}_2 \rangle = \langle \mathbf{x}_1, \lambda_2\mathbf{x}_2 \rangle \Rightarrow \overline{\lambda}_1 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \lambda_2 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle \Rightarrow \lambda_1 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \lambda_2 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle$$

(De laatste stap is correct omdat we net bewezen hebben dat de eigenwaarden van een Hermitese matrix reëel zijn). Uit $\lambda_1 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \lambda_2 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle$, ofwel $(\lambda_1 - \lambda_2) \langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 0$, volgt, omdat $\lambda_1 \neq \lambda_2$, dat $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 0$, d.w.z. dat $\mathbf{x}_1$ en $\mathbf{x}_2$ orthogonaal zijn. □

Voor scheefhermitese matrices geldt iets dergelijks:

STELLING 4.16. *Laat A een scheefhermitese $n \times n$-matrix zijn. Dan geldt*

1. *A heeft slechts zuiver imaginaire eigenwaarden,*
2. *eigenvectoren bij verschillende eigenwaarden zijn orthogonaal.*

Bewijs: De uitspraken 1 en 2 kunnen op dezelfde manier bewezen worden als in de vorige stelling. We kunnen echter ook als volgt te werk gaan: Definieer $B = iA$. Dan is B een Hermitese matrix. Laat λ een eigenwaarde zijn van A met eigenvector $\mathbf{x} \neq \mathbf{0}$. Dan geldt

$$B\mathbf{x} = iA\mathbf{x} = i\lambda\mathbf{x}.$$

We zien dus: Als λ een eigenwaarde van A is, dan is $i\lambda$ een eigenwaarde van B (en andersom), en als $\mathbf{x}$ een eigenvector van A is, dan is $\mathbf{x}$ ook een eigenvector van B (en andersom). Hieruit volgt het gestelde. □

Zelfs voor unitaire matrices kunnen we iets dergelijks zeggen:

STELLING 4.17. *Laat U een unitaire $n \times n$-matrix zijn. Dan geldt*

1. *Elke eigenwaarde van U heeft modulus 1,*
2. *eigenvectoren bij verschillende eigenwaarden zijn orthogonaal.*

Bewijs: Voor de eerste uitspraak realiseren we ons eerst dat stelling 2.14 impliceert dat uit $U\mathbf{x} = \lambda\mathbf{x}$ (met $\mathbf{x} \neq \mathbf{0}$) volgt dat $\|\mathbf{x}\| = \|U\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$. Maar dat kan alleen als $|\lambda| = 1$ (want $\mathbf{x} \neq \mathbf{0}$).
Voor uitspraak 2 nemen we twee eigenvectoren $\mathbf{x} \neq \mathbf{0}$ en $\mathbf{y} \neq \mathbf{0}$ bij twee verschillende eigenwaarden λ en μ. Voordat we het eigenlijke bewijs geven, merken we even op dat $|\lambda|^2 = \overline{\lambda}\lambda = 1$ en dus dat

$$\overline{\lambda} = \frac{1}{\lambda}.$$

Hieruit volgt namelijk

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle U\mathbf{x}, U\mathbf{y} \rangle = \langle \lambda\mathbf{x}, \mu\mathbf{y} \rangle = \overline{\lambda}\mu \langle \mathbf{x}, \mathbf{y} \rangle = \frac{\mu}{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle .$$

Omdat

$$\frac{\mu}{\lambda} \neq 1$$

(want $\mu \neq \lambda$) moet gelden $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. □

Voorbeeld 4.18. De eigenwaarden van de orthogonale matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ zijn 1 en -1. Ze hebben dus inderdaad modulus 1 en zijn in dit geval reëel.
De eigenwaarden van de orthogonale matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ zijn i en $-i$. Ook deze hebben modulus 1, maar deze zijn nu niet-reëel.

STELLING 4.19. *Laat A een hermitese, scheef–hermitese of een unitaire matrix zijn. Dan geldt dat: "eigenruimten E_{λ_1} en E_{λ_2} van A met $\lambda_1 \neq \lambda_2$ onderling orthogonaal zijn".*

4.4 Symmetrische Matrices en de reële Spectraaldecompositie

Herinner dat een matrix A symmetrisch is als $A = A^T$ èn als A een reële matrix is. (Dit laatste hebben wij afgesproken, maar is absoluut geen universele afspraak.) We zullen om vergissingen te voorkomen steeds spreken over "reële symmetrische matrices". Deze komen veel voor in de praktijk en we zullen zien dat er rondom deze symmetrische reële matrices een erg mooie theorie ontwikkeld is.

We beginnen met de opmerking dat een reëel symmetrische matrix Hermitesch is, en daarom levert Stelling 4.15 ons:

STELLING 4.20. *Zij A een reële symmetrische matrix. Dan geldt: het karakteristiek polynoom $p_A(\lambda)$ van A heeft alleen reële nulpunten.*

en

STELLING 4.21. *Zij A een reële symmetrische matrix. Dan geldt: eigenvectoren bij verschillende eigenwaarden zijn orthogonaal.*

De volgende stelling, waarvan het bewijs lastig is, maar opgeschreven is voor hen die geïnteresseerd zijn, vertelt ons dat iedere reële symmetrische matrix diagonaliseerbaar is.

STELLING 4.22. *Zij A een reële symmetrische matrix. Dan geldt: voor iedere eigenwaarde van A zijn de algebraïsche multipliciteit en de meetkundige multipliciteit gelijk.*

Bewijs: (Dit bewijs is lastig, maar is toch maar toegevoegd.) We bewijzen de uitspraak met inductie naar afmetingen van A. De uitspraak is zeker waar voor alle (symmetrische) 1×1 matrices. Stel nu dat de uitspraak waar is voor alle symmetrische $k \times k$ matrices.
Laat A een symmetrische $(k+1) \times (k+1)$ matrix zijn en stel dat $\lambda = \lambda_0$ een eigenwaarde is van A met algebraïsche multipliciteit n_0. Als $n_0 = 1$ dan zijn zeker de algebraïsche multipliciteit en de meetkundige multipliciteit van λ_0 gelijk.
Er bestaat zeker één eigenvector $\mathbf{q}_1 \in \mathbb{R}^{k+1}$ bij λ_0 (genormeerd op lengte 1, dus $|\mathbf{q}_1| = 1$) van

A. Breidt $\{\mathbf{q}_1\}$ uit tot een orthonormale basis $\{\mathbf{q}_1, \ldots, \mathbf{q}_{k+1}\}$ van $\mathbb{R}^{k+1}$. (Waarom kan dat?) en beschouw de orthogonale $(k+1) \times (k+1)$ matrix $Q = [\mathbf{q}_1 \ldots \mathbf{q}_{k+1}]$. Dan:

$$B = Q^T A Q = \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_{k+1}^T \end{bmatrix} A[\mathbf{q}_1 \ldots \mathbf{q}_{k+1}] = \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_{k+1}^T \end{bmatrix} [A\mathbf{q}_1 \ldots A\mathbf{q}_{k+1}] =$$

$$= \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_{k+1}^T \end{bmatrix} [\lambda_0 \mathbf{q}_1\, A\mathbf{q}_2 \ldots A\mathbf{q}_{k+1}] = \begin{bmatrix} \lambda_0 & \mathbf{0} & \ldots \\ \mathbf{0} & B_1 & \\ \vdots & & \end{bmatrix}$$

We merken op:

- B is een symmetrische matrix, want $B^T = (Q^T A Q)^T = Q^T A^T (Q^T)^T = Q^T A Q = B$; en dus is B_1 ook een symmetrische $(k \times k)$ matrix.
- B en A zijn gelijkvormig (nodig: Q is orthogonale matrix), want $B = Q^T A Q = Q^{-1} A Q$ en dus hebben A en B hetzelfde karakteristiek polynoom en dezelfde eigenwaarden met dezelfde (algebraïsche en meetkundige) multipliciteiten.
- In het bijzonder, λ_0 is een eigenwaarde van B met algebraïsche multipliciteit $n_0 > 1$. Merk op dat $\mathbf{e}_1 \in \mathbb{R}^{k+1}$ een eigenvector is van B bij λ_0.
- Omdat $B - \lambda I_{k+1} = \begin{bmatrix} \lambda_0 - \lambda & \mathbf{0} & \ldots \\ \mathbf{0} & B_1 - \lambda I_k & \\ \vdots & & \end{bmatrix}$ geldt:

$$p_A(\lambda) = p_B(\lambda) = \det(B - \lambda I_{k+1}) = (\text{ontwikkelen naar eerste kolom}) =$$

$$(\lambda_0 - \lambda) \det(B_1 - \lambda I_k) = (\lambda_0 - \lambda) p_{B_1}(\lambda)$$

- Omdat λ_0 een nulpunt is van $p_B(\lambda)$ met multipliciteit $n_0 > 1$, zal λ_0 een nulpunt zijn van $p_{B_1}(\lambda)$ met multipliciteit $n_0 - 1$.

Op de symmetrische $k \times k$ matrix B_1 kunnen we de inductie verondersrelling los laten: de meetkundige multipliciteit van λ_0 t.o.v. B_1 zal ook $n_0 - 1$ zijn. Laat $\mathbf{v}_1, \ldots, \mathbf{v}_{n_0-1}$ onafhankelijke eigenvectoren in $\mathbb{R}^k$ zijn van B_1 t.o.v. λ_0. Als we definieren $\mathbf{w}_i = \begin{bmatrix} 0 \\ \mathbf{v}_i \end{bmatrix} \in \mathbb{R}^{k+1}$ voor $(i = 1, \ldots, n_0 - 1)$ dan zijn

$$\mathbf{e}_1, \mathbf{w}_1, \ldots, \mathbf{w}_{n_0-1}$$

n_0 lineair onafhankelijke eigenvectoren van B t.o.v. de eigenwaarde λ_0 (ga na). We concluderen dat de meetkundige multipliciteit van λ_0 ook t.o.v. B inderdaad n_0 is, en dit zal dan ook gelden t.o.v. A. □

Merk op dat uit de voorafgaande Stelling 4.20 en Stelling 4.22 volgt dat iedere reële symmetrische matrix reëel diagonaliseerbaar is.

Orthogonale Diagonalisering van reële matrices

Definitie: Een vierkante matrix A heet **orthogonaal diagonaliseerbaar** als A reeël is en als er een (reële) orthogonale matrix Q bestaat en een (reële) diagonaal matrix D zodat $A = QDQ^{-1}$.

STELLING 4.23. *Stel dat A een reële $n \times n$ matrix is. Dan geldt:*
A is symmetrisch als en slechts als A orthogonaal diagonaliseerbaar is.

Bewijs: Als A orthogonaal diagonaliseerbaar is, dan bestaat een orthogonale matrix Q en een diagonaal matrix D zodat $A = QDQ^{-1}$. Omdat $Q^{-1} = Q^T$ (want Q is een orthogonale matrix) geldt: $A = QDQ^{-1} = QDQ^T$ en dus:

$$A^T = (QDQ^T)^T = (Q^T)^T D^T Q^T = QDQ^T = A$$

dat wil zeggen: A is een symmetrische matrix.
Omgekeerd: stel dat A symmetrisch is. Beschouw het karakteristiek polynoom $p_A(\lambda)$. Volgens Stelling 4.20 zijn alle nulpunten reëel, zeg $\lambda_1, \ldots, \lambda_k$. Zeg ook dat λ_i algebraïsche multipliciteit n_i $(i = 1, \ldots, k)$. Kennelijk geldt:

$$n_1 + \cdots + n_k = n$$

Volgens Stelling 4.22 is de meetkundige multipliciteit van λ_i ook gelijk aan n_i $(i = 1 \ldots k)$ en dus kunnen we voor de eigenruimte E_{λ_i} een orthonormale basis, zeg $\mathcal{B}_i = \{\mathbf{q}_{(i,1)}, \ldots, \mathbf{q}_{(i,n_i)}\}$ bepalen, (voor $i - 1, \ldots, k$). Beschouw

$$\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_k$$

$\mathcal{B}$ bestaat uit eenheidsvectoren en daarom zegt Stelling 4.21 dat $\mathcal{B}$ een orthonormale verzameling is van n eigenvectoren. Daarom geldt: als we vectoren uit $\mathcal{B}$ in een matrix stoppen verkrijgen we een orthogonale matrix Q en als D de diagonaal matrix wordt met op de diagonaal de overeenkomstige eigenwaarden, dan geldt:

$$A = QDQ^{-1} = QDQ^T$$

en dus is A orthogonaal diagonaliseerbaar. □

Merk op dat het bewijs ons precies vertelt hoe een symmetrische matrix orthogonaal gediagonaliseerd moet worden. Bepaal voor iedere eigenwaarde λ van de symmetrische matrix A een orthonormale basis van de eigenruimte E_λ en ga met deze bases dan verder met het standaard diagonalisering proces!

Voorbeeld 4.24. Bepaal een orthogonale diagonalisering van de symmetrische matrix

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

als gegeven is dat het karakteristiek polynoom $p_A(\lambda) = -(\lambda - 4)(\lambda - 1)^2$ is.
Antwoord:
Uit $p_A(\lambda) = -(\lambda - 4)(\lambda - 1)^2 = 0$ zien we dat $\lambda = 4$ en $\lambda = 1$ de eigenwaarden zijn met a.m.$(\lambda = 4) = 1$ en met a.m.$(\lambda = 1) = 2$.
De volgende stap bestaat uit het bepalen van een basis van de eigenruimten E_4 en E_1.
Basis voor $E_4 = \text{Nul}(A - 4I)$ wordt bepaald door het oplossen van het stelsel $[A - 4I|\mathbf{0}]$ en dit levert $\mathcal{C}_1 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} = \{\mathbf{b}_1\}$ als basis.
Evenzo: een basis voor $E_1 = \text{Nul}(A - I)$ wordt bepaald door het oplossen van het stelsel $[A - I|\mathbf{0}]$ en dit levert $\mathcal{C}_2 = \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\} = \{\mathbf{b}_2, \mathbf{b}_3\}$ als basis.

Vroeger waren we nu klaar, we zouden zeggen: als $P = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ en $D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
dan is $A = PDP^{-1}$ een diagonalisering van A. Maar dit is GEEN orthogonale diagonalisering, want de matrix P is GEEN orthogonale matrix.
Merk op dat $\mathbf{b}_1 \perp \mathbf{b}_2$ en $\mathbf{b}_1 \perp \mathbf{b}_3$, want eigenvectoren bij verschillende eigenwaarden zijn orthogonaal bij een symmetrische matrix. Echter, $\mathbf{b}_2$ en $\mathbf{b}_3$ zijn niet orthogonaal (dit zijn eigenvectoren bij dezelfde eigenwaarde!).
Een orthonormale basis voor E_4 is makkelijk, we hoeven alleen maar te normeren. Dit levert:
$\mathcal{B}_1 = \left\{ \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \right\} = \{\mathbf{q}_1\}$

Bij E_1 is het Gram-Schmidt proces nodig. $\mathbf{w}_2 = \mathbf{b}_2$ en $\mathbf{w}_3' = \mathbf{b}_3 - \frac{\mathbf{b}_3 \cdot \mathbf{w}_2}{\mathbf{w}_2 \cdot \mathbf{w}_2}\mathbf{w}_2 = \begin{bmatrix} -1/2 \\ 1 \\ -1/2 \end{bmatrix}$, wat na schaling $\mathbf{w}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ en $\mathbf{w}_3 = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$ oplevert. Dus $\{\mathbf{w}_2, \mathbf{w}_3\}$ is een orthogonale basis voor E_1 en na normeren verkrijgen we de orthonormale basis $\left\{ \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix} \right\} = \{\mathbf{q}_2, \mathbf{q}_3\}$.
Tenslotte: als we definiëren

$$Q = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \end{bmatrix} \text{ en } D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

dan is $A = QDQ^{-1} = QDQ^T$ de gezochte orthogonale diagonalisering van A. △

De Spektraaldecompositie stelling

Laat A een symmetrische $n \times n$ matrix zijn. Stel dat $\lambda_1, \ldots, \lambda_n$ de eigenwaarden zijn van A en laat $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ een orthonormale basis zijn van $\mathbb{R}^n$, waarbij $\mathbf{q}_i$ een eigenvector is van λ_i (voor $i = 1, \ldots, n$). Schrijven we $Q = [\mathbf{q}_1 \ldots \mathbf{q}_n]$ en $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$ dan geldt:

$$A = QDQ^T = [\mathbf{q}_1 \ldots \mathbf{q}_n] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} = [\lambda_1 \mathbf{q}_1 \ldots \lambda_n \mathbf{q}_n] \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix}$$

Passen we op dit laatste product kolom-rij expansie toe dan krijgen we vorm 1 van de spectraal decompositie:

STELLING 4.25. (De Spektraaldecompositie Stelling, vorm 1)
Laat A een reële symmetrische $n \times n$ matrix zijn. Dan bestaat een orthonormale basis $B = \{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ van $\mathbb{R}^n$ bestaande uit eigenvectoren van A, zeg $A\mathbf{q}_i = \lambda_i \mathbf{q}_i$, voor $i = 1, \ldots, n$. Er geldt:

$$\boxed{A = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T}$$

Met andere woorden: A is een lineaire combinatie van projectiematrices op één dimensionale eigenruimtes. Bovenstaande decompositie van A heet "een spektraaldecompositie van A".

Voorbeeld 4.26. (Vervolg Voorbeeld 4.24) Beschouw de matrix A uit Voorbeeld 4.24. Een spektraaldecompositie van deze matrix is:

$$A = \lambda_1 \mathbf{q}_1\mathbf{q}_1^T + \lambda_2 \mathbf{q}_2\mathbf{q}_2^T + \lambda_3 \mathbf{q}_3\mathbf{q}_3^T = 4\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} +$$

$$1\begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}\begin{bmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} + 1\begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}\begin{bmatrix} -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \end{bmatrix} =$$

$$4\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} + 1\begin{bmatrix} 1/2 & 0 & -1/2 \\ 0 & 0 & 0 \\ -1/2 & 0 & 1/2 \end{bmatrix} + 1\begin{bmatrix} 1/6 & -2/6 & 1/6 \\ -2/6 & 4/6 & -2/6 \\ 1/6 & -2/6 & 1/6 \end{bmatrix}$$

Merk op dat deze decompositie NIET uniek is. Een andere orthonormale bases $\{\mathbf{q}_2', \mathbf{q}_3'\}$ voor de eigenruimte E_1 leidt tot een andere spektraal decompositie. $\triangle$

Voorbeeld 4.27. Het gebrek aan uniciteit van de spektraaldecompositie stelling vorm 1 wordt met name duidelijk geïllustreerd door de eenheidsmatrix $A = I$. Omdat bij $A = I$ iedere vector (ongelijk nulvector) een eigenvector is bij de eigenwaarde $\lambda = 1$. Dus geldt voor iedere orthonormale basis $B = \{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ van $\mathbb{R}^n$:

$$\boxed{I = \mathbf{q}_1\mathbf{q}_1^T + \cdots + \mathbf{q}_n\mathbf{q}_n^T}$$

Er bestaat een tweede vorm van de spektraaldecompositie die als voordeel heeft dat deze wel uniek is. Om deze te vinden gebruiken we de notatie uit het bewijs van Stelling 4.23. Dus:
A is een symmetrische matrix, $\lambda_1, \ldots, \lambda_k$ zijn de eigenwaarden van A, λ_i heeft multipliciteit n_i $(i = 1, \ldots, k)$ en $\mathcal{B}_i = \{\mathbf{q}_{(i,1)}, \ldots, \mathbf{q}_{(i,n_i)}\}$ is een orthonormale basis van de eigenruimte E_{λ_i}, voor $i = 1, \ldots, k$. Als we schrijven $Q_i = [\mathbf{q}_{(i,1)} \cdots \mathbf{q}_{(i,n_i)}]$ dan weten we uit deel 1 dat de projectie op de eigenruimte E_{λ_i} de standaardmatrix:

$$[\mathrm{proj}_{E_{\lambda_i}}] = \mathbf{q}_{(i,1)}\mathbf{q}_{(i,1)}^T + \cdots + \mathbf{q}_{(i,n_i)}\mathbf{q}_{(i,n_i)}^T = Q_i Q_i^T$$

heeft. Als we in vorm 1 van de spektrale decompositie van A de projectiematrices bij dezelfde eigenwaarden bij elkaar nemen krijgen we:

$$A = \lambda_1[\mathrm{proj}_{E_{\lambda_1}}] + \cdots + \lambda_k[\mathrm{proj}_{E_{\lambda_k}}]$$

Dit heet de spectraal decompositie van A, vorm 2. Met andere woorden:
"de matrix A is de lineaire combinatie van de projectiematrices op de (onderling loodrechte) eigenruimtes met als gewichten: de bijbehorende eigenwaarden!"

STELLING 4.28. (De Spektraaldecompositie Stelling, vorm 2)
Laat A een reële symmetrische $n \times n$ matrix zijn met eigenwaarden $\lambda_1, \ldots, \lambda_k$. Beschouw de orthogonale projecties $\mathrm{proj}_{E_{\lambda_i}} : \mathbb{R}^n \to \mathbb{R}^n$ op de eigenruimtes E_{λ_i}. Dan geldt:

$$\boxed{A = \lambda_1[\mathrm{proj}_{E_{\lambda_1}}] + \cdots + \lambda_k[\mathrm{proj}_{E_{\lambda_k}}]}$$

Voorbeeld 4.29. (Vervolg Voorbeeld 4.24 en 4.26) Beschouw de matrix A uit Voorbeeld 4.24. Vorm 2 van de spektrale decompositie van A wordt:

$$A = 4[\text{proj}_{E_4}] + 1[\text{proj}_{E_1}] = 4[\mathbf{q}_1][\mathbf{q}_1]^T + 1[\mathbf{q}_2\mathbf{q}_3][\mathbf{q}_2\mathbf{q}_3]^T =$$

$$4\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} + 1\begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} \\ 0 & 2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{6} \end{bmatrix}\begin{bmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \end{bmatrix} =$$

$$4\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} + 1\begin{bmatrix} 2/3 & -2/6 & -2/6 \\ -2/6 & 2/3 & -2/6 \\ -2/6 & -2/6 & 2/3 \end{bmatrix}$$

en dit laatste is de gewenste decompositie van A. △

4.5 Unitaire Diagonaliseerbaarheid

In het vorige college is voor reële matrices we het begrip “orthogonale diagonaliseerbaarheid”. Herinner:

Definitie: Een *reële* matrix A heet **orthogonaal diagonaliseerbaar** als er een *reële* diagonaalmatrix D en een orthogonale matrix Q bestaan zo dat

$$A = QDQ^{-1}(= QDQ^T).$$

Het ligt nu voor de hand om te definiëren:

Definitie: Een (eventueel complexe) matrix A heet **unitair diagonaliseerbaar** als er een (eventueel complexe) diagonaalmatrix D en een unitaire matrix U bestaan zo dat

$$A = UDU^{-1}(= UDU^*).$$

In deze context spreekt men ook vaak over unitaire gelijkvormigheid: Twee matrices A en B heten **unitair gelijkvormig** als er een unitaire matrix U bestaat zo dat $A = UBU^{-1}(= UBU^*)$. Merk op dat een matrix unitair diagonaliseerbaar is als en slechts als deze matrix unitair gelijkvormig is met een diagonaalmatrix.

We weten precies welke reële matrices orthogonaal diagonaliseerbaar zijn, want we herinneren ons uit het vorige college:

“Een reële matrix A is orthogonaal diagonaliseerbaar als en slechts als A symmetrisch is”.

De vraag is welke matrices unitair diagonaliseerbaar zijn.
In tegenstelling tot “gewone” diagonaliseerbaarheid bestaat hiervoor een eenvoudig criterium. Het vergt wel wat werk om dit te bewijzen! De stelling die hierbij een centrale rol speelt is het zogenaamde lemma van Schur. Dit lemma zegt dat *elke vierkante matrix unitair gelijkvormig met een driehoeksmatrix* is.

LEMMA 4.30. *[Lemma van Schur] Elke vierkante matrix is unitair gelijkvormig met een bovendriehoeksmatrix, m.a.w. bij een (eventueel) complexe $n \times n$-matrix A bestaat een unitaire matrix U en een bovendriehoeksmatrix R zo dat $A = URU^*$.*

Bewijs: We bewijzen dit met inductie naar de afmetingen van A. Voor $n = 1$ is de stelling duidelijk (een 1×1-matrix is namelijk een bovendriehoeksmatrix, en dus kan voor U de 1×1-eenheidsmatrix genomen worden en voor R de matrix A zelf). We veronderstellen nu dat de stelling waar is voor alle $(n-1)\times(n-1)$-matrices. We bewijzen dat het dan ook geldt voor een $n \times n$-matrix.
We weten dat A tenminste één eigenwaarde λ_1 heeft. Laat $\mathbf{v}_1$ een bijbehorende eigenvector zijn met lengte 1. Met de methode van Gram-Schmidt kunnen we vectoren $\mathbf{v}_2, \ldots, \mathbf{v}_n$ vinden, zo dat $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ een orthonormale basis voor $\mathbb{C}^n$ is. Definieer de unitaire matrix U_1 nu door

$$U_1 = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}.$$

Dan geldt

$$AU_1 = \begin{bmatrix} \lambda_1\mathbf{v}_1 & A\mathbf{v}_2 & \cdots & A\mathbf{v}_n \end{bmatrix}.$$

Hiermee kunnen we $U_1^* A U_1$ als volgt als blokmatrix schrijven

$$U_1^* A U_1 = \begin{bmatrix} \mathbf{v}_1^* \\ \mathbf{v}_2^* \\ \vdots \\ \mathbf{v}_n^* \end{bmatrix} \begin{bmatrix} \lambda_1\mathbf{v}_1 & A\mathbf{v}_2 & \cdots & A\mathbf{v}_n \end{bmatrix} = \left[\begin{array}{c|c} \lambda_1 & \cdots \\ \hline 0 & \\ \vdots & A_1 \\ 0 & \end{array}\right].$$

Hierbij is A_1 een $(n-1)\times(n-1)$-matrix, waarop we de inductieveronderstelling kunnen toepassen. Deze zegt dat er een unitaire $(n-1)\times(n-1)$-matrix C bestaat, en een bovendriehoeksmatrix B zo dat

$$A_1 = CBC^* \quad \text{ofwel} \quad B = C^* A_1 C.$$

Maak nu de matrix

$$U_2 = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & C & \\ 0 & & & \end{array}\right].$$

Omdat C unitair is, is U_2 dit ook. Omdat het product van twee unitaire matrices weer unitair is (zie de opgaven over unitaire matrices, §8.3), volgt dat de matrix $U = U_1U_2$ unitair is. Er geldt

$$U^*AU = U_2^*(U_1^* A U_1)U_2 = U_2^* \left[\begin{array}{c|c} \lambda_1 & \cdots \\ \hline 0 & \\ \vdots & A_1 \\ 0 & \end{array}\right] U_2 = \left[\begin{array}{c|c} \lambda_1 & \cdots \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right].$$

De laatste gelijkheid is te verifieren door U_2^* en U_2 ook als blokmatrices te schrijven en vervolgens het product uit te schrijven. Omdat B een bovendriehoeksmatrix is, is $R = U^*AU$ ook een bovendriehoeksmatrix. Hiermee is de inductie voltooid. □

We voeren nu een speciale klasse van vierkante matrices in, namelijk de zogenaamde normale matrices. We zullen laten zien dat dit precies de unitair diagonaliseerbare matrices zijn.

Definitie: Een (eventueel complexe) vierkante matrix A heet **normaal** als $A^*A = AA^*$ (dus als A commuteert met zijn hermites getransponeerde).

Voorbeeld 4.31. Alle Hermitese, scheefhermitese, unitaire, symmetrische, scheefsymmetrische en orthogonale matrices zijn normaal. Een voorbeeld van een normale matrix die niet van één van deze typen is, is

$$\begin{bmatrix} 0 & 0 & \sqrt{2} \\ 1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix}.$$

Bedenk zelf een niet-normale vierkante matrix.

Opmerking: Het lijkt wellicht op het eerste gezicht niet zo een zware eis voor een matrix om normaal te zijn. Echter, de meeste matrices zijn niet normaal. In de eerste plaats blijkt dit uit de komende toegift: commuterende (diagonaliseerbare) matrices A en B zijn gemeenschappelijk diagonaliseerbaar, d.w.z. dat er een matrix P bestaat zo dat dat $A = PD_1P^{-1}$ en $B = PD_2P^{-1}$. Dus de kolommen van P vormen een basis van $\mathbb{C}^n$ die zowel bestaat uit eigenvectoren van A als uit eigenvectoren van B. De laatste stelling van deze paragraaf laat zien dat voor een normale matrix A geldt dat A en A^* precies dezelfde eigenvectoren hebben.

LEMMA 4.32. *We hebben de volgende eigenschappen:*

1. *Als de matrix A normaal is en unitair gelijkvormig met een matrix B, dan is B ook normaal,*
2. *Een normale bovendriehoeksmatrix is een diagonaalmatrix.*

Bewijs: We bewijzen uitspraak 1: Schrijf $B = UAU^*$ met U unitair. Dan geldt

$$BB^* = (UAU^*)(UAU^*)^* = (UAU^*)(UA^*U^*) = UA(U^*U)A^*U^* = UAA^*U^*$$

en

$$B^*B = (UAU^*)^*(UAU^*) = (UA^*U^*)(UAU^*) = UA^*(U^*U)AU^* = UA^*AU^*.$$

Omdat A normaal is geldt dat $AA^* = A^*A$. Hieruit volgt direct dat $BB^* = B^*B$ en dus dat B normaal is.
Voor uitspraak 2 nemen we een normale $n \times n$ bovendriehoeksmatrix A. Schrijf $C = AA^* = A^*A$, dus

$$\begin{aligned} C &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{bmatrix} \begin{bmatrix} \overline{a}_{11} & & & \\ \overline{a}_{12} & \overline{a}_{22} & & \\ \vdots & \vdots & \ddots & \\ \overline{a}_{1n} & \overline{a}_{2n} & \cdots & \overline{a}_{nn} \end{bmatrix} \\ &= \begin{bmatrix} \overline{a}_{11} & & & \\ \overline{a}_{12} & \overline{a}_{22} & & \\ \vdots & \vdots & \ddots & \\ \overline{a}_{1n} & \overline{a}_{2n} & \cdots & \overline{a}_{nn} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{bmatrix}. \end{aligned}$$

We kunnen de kentallen van C dus op twee manieren uitrekenen. We doen het voor de kentallen op de hoofddiagonaal:

$$c_{11} = a_{11}\overline{a}_{11} + a_{12}\overline{a}_{12} + \cdots + a_{1n}\overline{a}_{1n} = a_{11}\overline{a}_{11}$$

Hieruit volgt dat

$$a_{12}\overline{a}_{12} + \cdots + a_{1n}\overline{a}_{1n} = |a_{12}|^2 + \cdots + |a_{1n}|^2 = 0.$$

Dit kan alleen als $a_{12} = \cdots = a_{1n} = 0$. Op dezelfde manier vinden we

$$c_{22} = a_{22}\overline{a}_{22} + a_{23}\overline{a}_{23} + \cdots + a_{2n}\overline{a}_{2n} = a_{12}\overline{a}_{12} + a_{22}\overline{a}_{22}$$

Omdat $a_{12} = 0$ volgt hieruit dat

$$a_{23}\overline{a}_{23} + \cdots + a_{2n}\overline{a}_{2n} = |a_{23}|^2 + \cdots + |a_{2n}|^2 = 0.$$

Dit kan alleen als $a_{23} = \cdots = a_{2n} = 0$. Op dezelfde manier kunnen we, via de berekening van c_{33} aantonen dat $a_{34} = \cdots = a_{3n} = 0$, enzovoorts. Zo blijkt dat alle kentallen van A boven de diagonaal nul moeten zijn. A is dus een diagonaalmatrix. □

Met dit lemma kunnen we het eerder genoemde criterium bewijzen:

STELLING 4.33. *Een matrix A is unitair diagonaliseerbaar als en slechts als A normaal is.*

Bewijs: Neem eerst aan dat A unitair diagonaliseerbaar is. Dat betekent dat A unitair gelijkvormig is met een diagonaalmatrix. Omdat een diagonaalmatrix normaal is, volgt uit de eerste uitspraak van Lemma 4.32 dat A ook normaal is.
Neem nu aan dat A normaal is. Volgens het Lemma van Schur (Lemma 4.30) is A unitair gelijkvormig met een bovendriehoeksmatrix B. Volgens de eerste uitspraak van het Lemma 4.32 is B dus ook normaal. Uit de tweede uitspraak van hetzelfde lemma volgt dan dat B een diagonaalmatrix is. We concluderen dat A unitair diagonaliseerbaar is. □

GEVOLG 4.34. *Alle Hermitese, scheefhermitese, unitaire, symmetrische, scheefsymmetrische en orthogonale matrices zijn unitair diagonaliseerbaar.*

In het complexe geval zijn de unitair diagonaliseerbare matrices dus precies de normale matrices. Maar pas op: een reëel matrix A kan unitair diagonaliseerbaar zijn zonder (reëel) orthogonaal diagonaliseerbaar te zijn. Neem maar een niet symmetrische orthogonale (reële) matrix A. Dan: A is reëel maar niet symmetrisch, dus NIET orthogonaal diagonaliseerbaar. Maar: A is orthogonaal, dus normaal dus unitair diagonaliseerbaar!

STELLING 4.35. *Laat A een normale matrix zijn. Dan geldt dat A en A^* precies dezelfde eigenvectoren hebben.*

Bewijs: Laat A een normale matrix zijn. Stel dat $\mathbf{u}$ een eigenvector van A is, zeg $A\mathbf{u} = \lambda\mathbf{u}$. Beschouw de vector $\mathbf{u}_1 = \frac{\mathbf{u}}{\|\mathbf{u}\|}$. Breidt deze ene vector uit tot een orthonormale basis van de eigenruimte E_λ en daarna tot een orthonormale basis $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ van $\mathbb{C}^n$ van eigenvectoren van A, door orthonormale bases van andere eigenruimtes toe te voegen. Als $U = [\,\mathbf{u}_1 \ldots \mathbf{u}_n\,]$, dan is U een unitaire matrix en er bestaat een diagonaalmatrix D zodat $A = UDU^{-1} = UDU^{-1} = UDU^*$. Dan $A^* = (UDU^*)^* = (U^*)^*\overline{D}U^* = U\overline{D}U^{-1}$. maar dan zegt lemma 13.4 dat U bestaat uit eigenvectoren van A^*. Dus $\mathbf{u}_1 = \frac{\mathbf{u}}{\|\mathbf{u}\|}$ is eigenvector van A^* en dus $\mathbf{u}$ ook.
Omgekeerd, ook A^* is normaal en hetzelfde bewijs laat daarom zien dat een eigenvector van A^* ook eigenvector is van $(A^*)^* = A$. De conclusie volgt. □

4.6 Normale matrices en de complexe Spektraaldecompositie

Laat A een normale $n \times n$ matrix zijn. A heeft een unitaire diagonalisering, daarom stel dat $\lambda_1, \ldots, \lambda_n$ de eigenwaarden zijn van A en laat $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ een orthonormale basis zijn van $\mathbb{C}^n$, waarbij $\mathbf{u}_i$ een eigenvector is van λ_i (voor $i = 1, \ldots, n$). Schrijven we $U = [\mathbf{u}_1 \ldots \mathbf{u}_n]$ en $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ dan geldt:

$$A = UDU^* = [\mathbf{u}_1 \ldots \mathbf{u}_n] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1{}^* \\ \vdots \\ \mathbf{u}_n{}^* \end{bmatrix} = [\lambda_1\mathbf{u}_1 \ldots \lambda_n\mathbf{u}_n] \begin{bmatrix} \mathbf{u}_1^* \\ \vdots \\ \mathbf{u}_n^* \end{bmatrix}$$

Passen we op dit laatste product kolom-rij expansie toe dan krijgen we: $A = \lambda_1\mathbf{u}_1\mathbf{u}_1^* + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^*$. We hebben bewezen:

STELLING 4.36. (De (complexe) Spektraledecompositie Stelling, vorm 1)
Laat A een normale $n \times n$ matrix zijn. Dan bestaat een orthonormale basis $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ van $\mathbb{C}^n$ bestaande uit eigenvectoren van A, zeg $A\mathbf{u}_i = \lambda_i\mathbf{u}_i$, voor $i = 1, \ldots, n$. Er geldt:

$$\boxed{A = \lambda_1\mathbf{u}_1\mathbf{u}_1^* + \cdots + \lambda_n\mathbf{u}_n\mathbf{q}_n^*}$$

Echter, de conclusie: "A is een lineaire combinatie van projectiematrices op één dimensionale eigenruimtes" kunnen we nog niet trekken want projecteren in $\mathbb{C}^n$ is nog niet besproken, dus projectiematrices ook nog niet. We zullen zien dat de conclusie wel juist is. Lees verder in sectie 5.8

Hoofdstuk 5

Inwendig productruimten

5.1 Inwendig producten

Ooit zijn de standaard inwendig producten op $\mathbb{R}^n$ en $\mathbb{C}^n$ ingevoerd. Dit doet wellicht vermoeden dat er ook zoiets zal bestaan als een "niet–standaard inwendig product op deze ruimten". En inderdaad, dit is het geval. We gaan zelfs inwendig producten op een vectorruimte bekijken.
De reden waarom we geïnteresseerd zijn in inwendig producten op vectorruimten is dat er veel meetkunde gekoppeld is aan het inwendig product. Kijk maar naar de inhoud van college 2 en 3, alle geïntroceerde meetkunde in $\mathbb{C}^n$ is gebeurd aan de hand van het standaard inwendig product. (Evenzo in deel 1 voor $\mathbb{R}^n$.) We willen dit gaan zien, en wel in iedere vectorruimte met een zogeheten " inwendig product–structuur" er op. (Niet alle vectorruimtes hebben dat.)
Dan volgt nu de definitie van een inwendig product. (Vergelijk de definitie overigens ook met het tijdens voortgezette analyse ingevoerde inwendig product.)

Definitie: Een *inwendig product* op een vectorruimte V over $\mathbb{L}$ (waarbij $\mathbb{L} = \mathbb{R}$ of $\mathbb{L} = \mathbb{C}$) is een voorschrift dat aan elk tweetal vectoren $\mathbf{u}$ en $\mathbf{v}$ een getal uit $\mathbb{L}$ toekent. Dit getal wordt genoteerd[1] met $\langle \mathbf{u}, \mathbf{v} \rangle$en moet de volgende eigenschappen hebben:

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$ voor alle $\mathbf{u}$ en $\mathbf{v}$ in V
2. $\langle \mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2 \rangle = \langle \mathbf{u}, \mathbf{v}_1 \rangle + \langle \mathbf{u}, \mathbf{v}_2 \rangle$ voor alle $\mathbf{u}$, $\mathbf{v}_1$ en $\mathbf{v}_2$ in V
3. $\langle \mathbf{u}, c\mathbf{v} \rangle = c \langle \mathbf{u}, \mathbf{v} \rangle$ voor alle $\mathbf{u}$, $\mathbf{v}$ in V en $c \in \mathbb{L}$
4. $\langle \mathbf{u}, \mathbf{u} \rangle$ is reëel en $\langle \mathbf{u}, \mathbf{u} \rangle > 0$ voor elke $\mathbf{u} \neq \mathbf{0}$ uit V. (Positief definiet)

Een vectorruimte met een inwendig product heet een *inwendig productruimte.* Een inwendig product op een complexe vectorruimte wordt ook wel een *hermitisch*[2] *product* genoemd en de vectorruimte in dat geval een *hermitisch productruimte.*

[0] In de natuurkunde wordt vaak de notatie $\langle \mathbf{u} | \mathbf{v} \rangle$ gebruikt voor een inwendig product.
[1] Er zijn boeken waarbij de eisen 2 en 3 niet voor de tweede factor, maar voor de eerste geformuleerd worden. Hier moet je altijd alert op zijn.
[2] Charles Hermite 1822-1901. Hij had moeite met zijn studie en heeft 5 jaar over zijn kandidaatsexamen gedaan. Hij heeft een belangrijke bijdrage geleverd op het gebied van de elliptische functies. Verder heeft hij als eerste het transcendent zijn van e aangetoond.

Opmerking: Als we een inwendig product op een reële vectorruimte V hebben, is $\langle \mathbf{u}, \mathbf{v}\rangle$ voor elke $\mathbf{u}$ en $\mathbf{v}$ uit V een reëel getal en dus is $\overline{\langle \mathbf{u}, \mathbf{v}\rangle} = \langle \mathbf{u}, \mathbf{v}\rangle$. De eerste eigenschap van het inwendig product kan dan dus vervangen worden door

$$\langle \mathbf{v}, \mathbf{u}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle .$$

Met andere woorden: een inwendig product op een *reële vectorruimte* V is een voorschrift dat aan elk tweetal vectoren $\mathbf{u}$ en $\mathbf{v}$ een getal uit $\mathbb{R}$ toekent. Dit getal wordt genoteerd met $\langle \mathbf{u}, \mathbf{v}\rangle$ en moet de volgende eigenschappen hebben:

1. $\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle$ voor alle $\mathbf{u}$ en $\mathbf{v}$ in V
2. $\langle \mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2\rangle = \langle \mathbf{u}, \mathbf{v}_1\rangle + \langle \mathbf{u}, \mathbf{v}_2\rangle$ voor alle $\mathbf{u}$, $\mathbf{v}_1$ en $\mathbf{v}_2$ in V
3. $\langle \mathbf{u}, c\mathbf{v}\rangle = c\,\langle \mathbf{u}, \mathbf{v}\rangle$ voor alle $\mathbf{u}$, $\mathbf{v}$ in V en $c \in \mathbb{L}$
4. $\langle \mathbf{u}, \mathbf{u}\rangle$ is reëel en $\langle \mathbf{u}, \mathbf{u}\rangle > 0$ voor elke $\mathbf{u} \neq \mathbf{0}$ uit V. (Positief definiet)

Rekenregels voor een inwendig product

De eigenschappen 2 en 3 zeggen dat een inwendig product *lineair* is *in de tweede variabele*. Hoe zit dat met de eerste variabele? De volgende stelling, waarin nog een aantal rekenregels verwoord worden, laat zien dat het in het reële geval wel zo is, maar in het complexe geval niet. Scalaire vermenigvuldiging in de eerste variabele is in het complexe geval namelijk *niet* altijd gelijk aan vermenigvuldiging van het inwendig product met die scalar.

STELLING 5.1 (Rekenregels voor een inwendig product). *Voor een inwendig product op een inwendig productruimte V over $\mathbb{L}$ geldt*

1. $\langle \mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}\rangle = \langle \mathbf{u}_1, \mathbf{v}\rangle + \langle \mathbf{u}_2, \mathbf{v}\rangle$ *voor alle* $\mathbf{u}_1$, $\mathbf{u}_2$ *en* $\mathbf{v}$ *in* V.
2. *Als* $\mathbb{L} = \mathbb{C}$ *dan* $\langle c\mathbf{u}, \mathbf{v}\rangle = \overline{c}\,\langle \mathbf{u}, \mathbf{v}\rangle$ *voor alle* $\mathbf{u}$, $\mathbf{v}$ *in* V *en* $c \in \mathbb{C}$.
   Als $\mathbb{L} = \mathbb{R}$ *dan* $\langle c\mathbf{u}, \mathbf{v}\rangle = c\,\langle \mathbf{u}, \mathbf{v}\rangle$ *voor alle* $\mathbf{u}$, $\mathbf{v}$ *in* V *en* $c \in \mathbb{R}$.
3. $\langle \mathbf{u}, \mathbf{0}\rangle = \langle \mathbf{0}, \mathbf{u}\rangle = 0$ *voor alle* $\mathbf{u}$ *in* V.
4. $\langle \mathbf{u}, \mathbf{u}\rangle = 0$ *als en slechts als* $\mathbf{u} = \mathbf{0}$.
5. *Als voor alle* $\mathbf{u}$ *in* V *geldt dat* $\langle \mathbf{x}, \mathbf{u}\rangle = \langle \mathbf{y}, \mathbf{u}\rangle$, *dan* $\mathbf{x} = \mathbf{y}$

Bewijs:

1. $\langle \mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}\rangle = \overline{\langle \mathbf{v}, \mathbf{u}_1 + \mathbf{u}_2\rangle} = \overline{\langle \mathbf{v}, \mathbf{u}_1\rangle + \langle \mathbf{v}, \mathbf{u}_2\rangle} = \overline{\langle \mathbf{v}, \mathbf{u}_1\rangle} + \overline{\langle \mathbf{v}, \mathbf{u}_2\rangle} = \langle \mathbf{u}_1, \mathbf{v}\rangle + \langle \mathbf{u}_2, \mathbf{v}\rangle$.
2. Als $c \in \mathbb{L}$ dan volgt: $\langle c\mathbf{u}, \mathbf{v}\rangle \overset{(1)}{=} \overline{\langle \mathbf{v}, c\mathbf{u}\rangle} \overset{(3)}{=} \overline{c\,\langle \mathbf{v}, \mathbf{u}\rangle} = \overline{c} \cdot \overline{\langle \mathbf{v}, \mathbf{u}\rangle} \overset{(1)}{=} \overline{c}\,\langle \mathbf{u}, \mathbf{v}\rangle$.
   De reële situatie volgt hier direct uit omdat voor reële c geldt dat $\overline{c} = c$.
3. $\langle \mathbf{u}, \mathbf{0}\rangle = \langle \mathbf{u}, 0\mathbf{0}\rangle = 0\,\langle \mathbf{u}, \mathbf{0}\rangle = 0$. Evenzo:
   $\langle \mathbf{0}, \mathbf{u}\rangle = \langle 0\mathbf{0}, \mathbf{u}\rangle = \overline{0}\,\langle \mathbf{0}, \mathbf{u}\rangle = 0$.
4. Als $\mathbf{u} \neq \mathbf{0}$ dan per definitie $\langle \mathbf{u}, \mathbf{u}\rangle > 0$.

5. Neem voor het bewijs van de vijfde rekenregel aan dat $\langle \mathbf{x}, \mathbf{u}\rangle = \langle \mathbf{y}, \mathbf{u}\rangle$ voor alle $\mathbf{u}$ in V. Daaruit volgt dat

$$0 = \langle \mathbf{x}, \mathbf{u}\rangle - \langle \mathbf{y}, \mathbf{u}\rangle = \langle \mathbf{x}, \mathbf{u}\rangle + \langle -\mathbf{y}, \mathbf{u}\rangle = \langle \mathbf{x} - \mathbf{y}, \mathbf{u}\rangle$$

voor alle $\mathbf{u}$ in V (hier zijn respectievelijk de tweede rekenregel van deze stelling en de tweede eis uit de definitie van een inwendig product gebruikt). In het bijzonder geldt dus dat $\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y}\rangle = 0$. Uit de vierde rekenregel van deze stelling volgt dan dat $\mathbf{x} - \mathbf{y} = \mathbf{0}$ en dus dat $\mathbf{x} = \mathbf{y}$

□

5.2 Standaard inwendig producten

Voorbeeld 5.2. Het standaard inwendig product op $\mathbb{R}^n$ is gedefinieerd door

$$\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^T\mathbf{y}.$$

Vaak wordt dit ook wel het *dot product* van $\mathbf{x}$ en $\mathbf{y}$ genoemd en genoteerd met $\mathbf{x} \cdot \mathbf{y}$. Voor $\mathbf{x} = (x_1, \dots, x_n)$ en $\mathbf{y} = (y_1, \dots, y_n)$ geldt dus dat

$$\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y} = x_1y_1 + \cdots + x_ny_n.$$

Ga zelf na dat dit inderdaad een inwendig product op $\mathbb{R}^n$ definieert. (Dat wil zeggen, ga na dat aan de voorwaarden van de definitie “reëel inwendig product” voldoet.)

Voorbeeld 5.3. Het standaard inwendig product op $\mathbb{C}^n$ is gedefinieerd door

$$\langle \mathbf{u}, \mathbf{v}\rangle = \overline{\mathbf{u}}^T\mathbf{v} = \mathbf{u}^*\mathbf{v} = \overline{u_1}v_1 + \overline{u_2}v_2 + \cdots + \overline{u_n}v_n.$$

Ga zelf nog eens na dat dit inderdaad een inwendig (hermitisch) product op $\mathbb{C}^n$ definieert.

Voorbeeld 5.4. Op de ruimte $\mathbb{R}^\infty$ is geen ‘natuurlijk inwendig product te definiëren, dit in tegenstelling tot $\ell_2(\mathbb{R})$ en $\ell_2(\mathbb{C})$. Dit is het grote voordeel van de ruimte $\ell_2(\mathbb{R})$ boven $\mathbb{R}^\infty$ (resp. $\ell_2(\mathbb{C})$ boven $\mathbb{C}^\infty$). De volgende uitdrukking definieert een inwendig product op $\ell_2(\mathbb{R})$ die het *standaard inwendig product* op $\ell_2(\mathbb{R})$genoemd wordt:

$$\langle (x_1, x_2, \dots), (y_1, y_2, \dots)\rangle = x_1y_1 + x_2y_2 + \cdots = \sum_{n=1}^{\infty} x_ny_n.$$

Dat het aan de vier eisen van een inwendig product voldoet is geheel analoog aan het standaard inwendig product op $\mathbb{R}^n$. Omdat we hier te maken hebben met een oneindige som (een reeks dus), moeten we alleen eerst nog nagaan of dit inwendig product wel altijd een eindige uitkomst heeft (met andere woorden, we moeten nagaan of de reeks convergent is).

Merk daartoe ten eerste op dat als $\mathbf{x} = \{x_n\}_{n=1}^\infty$ en $\mathbf{y} = \{y_n\}_{n=1}^\infty$ uit $\ell_2(\mathbb{R})$ zijn, dat dan $\displaystyle\sum_{n=0}^{\infty} |x_n|^2 < \infty$ en $\displaystyle\sum_{n=0}^{\infty} |y_n|^2 < \infty$ en omdat $|x_ny_n| = |x_n||y_n| \leq \frac{1}{2}|x_n|^2 + \frac{1}{2}|y_n|^2$ (want: $(|x_n| - |y_n|)^2 \geq 0$), volgt dat de reeks $\displaystyle\sum_{n\geq 0} x_ny_n$ absoluut convergent, en dus convergent is.

De volgende uitdrukking wordt het *standaard inwendig (hermitisch) product* op $\ell_2(\mathbb{C})$ genoemd:

$$\langle (x_1, x_2, \dots), (y_1, y_2, \dots)\rangle = \overline{x_1}y_1 + \overline{x_2}y_2 + \cdots = \sum_{n=1}^{\infty} \overline{x_n}y_n.$$

De convergentie van de reeks $\sum_{n=1}^{\infty} \overline{x_n} y_n$ volgt ook uit de ongelijkheid $|\overline{x_n} y_n| = |x_n||y_n| \leq \frac{1}{2}|x_n|^2 + \frac{1}{2}|y_n|^2$ en het feit dat ook voor complexe reeksen geldt: absoluut convergente reeksen zijn convergent.

Voorbeeld 5.5. Bij het vak voortgezette analyse uit het eerste jaar zijn we al het standaard inwendig product op $C([a, b], \mathbb{R})$ en $C([a, b], \mathbb{C})$ tegengekomen. Op $C([a, b], \mathbb{R})$ is dit:

$$\langle f, g\rangle = \int_a^b f(x)g(x)\, dx.$$

En op $C([a, b], \mathbb{C})$ wordt dit:

$$\langle f, g\rangle = \int_a^b \overline{f(x)} g(x)\, dx.$$

Dat dit inderdaad een inwendig product is, is niet zo lastig aan te tonen. De eerste drie eisen zijn rechtstreeks te verifiëren. De vierde eis: $\langle f, f\rangle = \int_a^b \overline{f(x)} f(x)\, dx = \int_a^b |f(x)|^2\, dx \geq 0$ en tenslotte: stel dat $\langle f, f\rangle = 0$. Als f niet de nulfunctie zou zijn, zeg $f(x_0) \neq 0$ met $x_0 \in [a, b]$, dan, vanwege de continuïteit van f, is er een heel deelinterval $[p, q]$ met $x_0 \in [p, q] \subset [a, b]$ zodat $|f(x)|^2 \geq \frac{1}{2}|f(x_0)|^2$ als $x \in [p, q]$. Maar dan zou gelden dat $\langle f, f\rangle = \int_a^b |f(x)|^2\, dx > 0$, dit in tegenspraak met de aanname dat $\langle f, f\rangle = 0$. Kortom, f moet wel de nulfunctie zijn.

Opmerking: Een voor de hand liggende vraag is: waarom kan de formule $\langle f, g\rangle = \int_a^b f(x)g(x)\, dx$ niet als inwendig product gebruikt worden op de vectorruimte $F([a, b], \mathbb{R})$, de vectorruimte van alle functies op $[a, b]$?
Welnu, het eerste probleem is dat er veel functies bestaan waarvoor de integraal $\langle f, g\rangle = \int_a^b f(x)g(x)\, dx$ domweg niet bestaat. Dit probleem zou nog te vermijden zijn door niet alle functies te nemen, maar je te beperken tot de vectorruimte van de integreerbare funcies op $[a, b]$.
Maar dan doemt een tweede probleem op: aan de eis "$\langle f, f\rangle = 0 \Rightarrow f = 0$" moet voldaan worden. Beschouw de functie

$$f(x) = \begin{cases} 0 & x \neq \frac{a+b}{2} \\ 1 & x = \frac{a+b}{2} \end{cases}$$

Deze functie f is integreerbaar op $[a, b]$ en heeft de eigenschap: $\langle f, f\rangle = \int_a^b (f(x))^2\, dx = 0$. Maar f is zeker niet de nulfunctie. Er is maar één manier om aan deze eis te voldoen, functies die in eindig veel punten uit het domein verschillen hetzelfde verklaren. Dit gebeurt in de vectorruimten $PC([a, b], \mathbb{R})$ en $L_2([a, b], \mathbb{R})$, en inderdaad, op deze vectorruimten vormt $\langle f, g\rangle = \int_a^b f(x)g(x)\, dx$ een inwendig product.

5.3 Niet standaard inwendig producten

We bekijken eerst niet standaard inwendig producten op $\mathbb{R}^n$. Laten we maar beginnen met een voorbeeld.

Voorbeeld 5.6 (Andere inwendig producten op $\mathbb{R}^n$)**.** De uitdrukking

$$\langle \mathbf{x}, \mathbf{y}\rangle = 3x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$$

definieërt een inwendig product op $\mathbb{R}^2$. Uit het feit dat deze uitdrukking te schrijven is in de vorm

$$\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^T A\mathbf{y}$$

met A de symmetrische matrix

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix},$$

(ga na of je dit inziet!)(maar zie ook de volgende stelling). Via deze schrijfwijze volgt direct de eerste, tweede en derde eis voor een inwendig product, en wel volgen deze uit de rekenregels voor de matrixvermenigvuldiging en transponeren. De eerste eis bewijs je bijvoorbeeld als volgt:

$$\langle \mathbf{y}, \mathbf{x} \rangle = (\langle \mathbf{y}, \mathbf{x} \rangle)^T = (\mathbf{y}^T A \mathbf{x})^T = \mathbf{x}^T A^T (\mathbf{y}^T)^T = \mathbf{x}^T A \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle .$$

Bij het eerste =-teken werd gebruikt dat het inwendig product een (in dit geval) reëel getal is, wat dus gelijk is aan zijn getransponeerde en bij het vierde =-teken dat A symmetrisch is (dus $A^T = A$). Rest aan te tonen dat de uitdrukking positief definiet is. Schrijf daartoe

$$\langle \mathbf{x}, \mathbf{x} \rangle = 3x_1^2 + 2x_1x_2 + x_2^2 = (x_1 + x_2)^2 + 2x_1^2.$$

Het is duidelijk dat dit altijd ≥ 0 is en alleen gelijk aan 0 als $x_1 + x_2 = 0$ én $x_1 = 0$, dus alleen als $\mathbf{x} = \mathbf{0}$. Kortom, dit is een niet–standaard inwendig product op $\mathbb{R}^2$. Merk op dat op dit nieuwe inwendig product niet geldt dat $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = 0$ of $\langle \mathbf{e}_1, \mathbf{e}_1 \rangle = 1$. $\square$

Bovenstaande aanpak is geen toeval, want:

STELLING 5.7. *Ieder inwendig product $\langle \mathbf{x}, \mathbf{y} \rangle$ op $\mathbb{R}^n$ is te schrijven als $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$, met een unieke symmetrische matrix A. Bovendien geldt dat $\langle \mathbf{x}, \mathbf{y} \rangle$ te schrijven is*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle$$

Bewijs: Om dit in te zien, schrijven we:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n, \mathbf{y} \rangle = x_1 \langle \mathbf{e}_1, \mathbf{y} \rangle + \cdots + x_n \langle \mathbf{e}_n, \mathbf{y} \rangle =$$

$$= \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{y} \rangle \\ \vdots \\ \langle \mathbf{e}_n, \mathbf{y} \rangle \end{bmatrix} = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} \begin{bmatrix} \langle \mathbf{e}_1, y_1 \mathbf{e}_1 + \cdots + y_n \mathbf{e}_n \rangle \\ \vdots \\ \langle \mathbf{e}_n, y_1 \mathbf{e}_1 + \cdots + y_n \mathbf{e}_n \rangle \end{bmatrix}$$

$$= [x_1 \ldots x_n] \begin{bmatrix} y_1 \langle \mathbf{e}_1, \mathbf{e}_1 \rangle + \cdots + y_n \langle \mathbf{e}_1, \mathbf{e}_n \rangle \\ \vdots \\ y_1 \langle \mathbf{e}_n, \mathbf{e}_1 + \cdots + y_n \langle \mathbf{e}_n, \mathbf{e}_n \rangle \rangle \end{bmatrix} = [x_1 \ldots x_n] \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{e}_1 \rangle & \cdots & \langle \mathbf{e}_1, \mathbf{e}_n \rangle \\ \vdots & & \vdots \\ \langle \mathbf{e}_n, \mathbf{e}_1 \rangle & \cdots & \langle \mathbf{e}_1, \mathbf{e}_n \rangle \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

De uniciteit van de matrix A volgt uit het feit dat $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \mathbf{e}_i^T A \mathbf{e}_j = [A]_{i,j}$. Tenslotte, als we het product uitwerken volgt de laatste claim. $\square$

Als we moeten nagaan of een gegeven formule een inwendigproduct op $\mathbb{R}^n$ is, dan moeten we ons realiseren dat de formule een sommatie van termen $a_{i,j} x_i y_j$ is. De aanpak zal zijn: schrijf deze formule als $\mathbf{x}^T A \mathbf{y}$, ga na of A symmetrisch is (zo niet, dan niet aan eis (1) voldaan). Vervolgens moeten we eis 4 controleren

Definitie: Laat A een reële symmetrische $n \times n$ matrix zijn. De matrix heet positief definiet, als $\mathbf{x}^T A \mathbf{x} > 0$, voor alle $\mathbf{x} \neq \mathbf{0}$.

Merk op dat de definiet positive symmetrische matrices precies die matrices zijn waarvoor de formule $\mathbf{x}^T A \mathbf{y}$ een inwendig product op $\mathbb{R}^n$ geeft.
Om voor een gegeven symmetrische matrix A na te gaan of de formule

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$$

inderdaad een inwendig product op $\mathbb{R}^n$ geeft (dus om eis 4 te controleren) volstaat vaak kwadraat-afsplitsen in $\mathbf{x}^T A \mathbf{x}$. Toch kan dit soms vrij lastig zijn. We geven eerst nog twee voorbeelden.

Voorbeeld 5.8. De uitdrukking

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_1 y_2 + x_2 y_1 + x_2 y_2$$

definieert geen inwendig product op $\mathbb{R}^2$.
Om dit in te zien schrijven we eerst:

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_1 y_2 + x_2 y_1 + x_2 y_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

Er volgt direct dat aan de eerste, tweede en derde eis voor een inwendig product voldaan is.
Tenslotte zien we in dat niet aan eis vier voldaan is.

$$\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + 2x_1 x_2 + x_2^2 = (x_1 + x_2)^2.$$

Het is duidelijk dat dit altijd ≥ 0 is. Maar we zien ook dat als $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ dat dan ook $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ terwijl $\mathbf{x} \neq 0$. □

Tenslotte nog een voorbeeld op $\mathbb{R}^3$.

Voorbeeld 5.9. Laat zien dat de uitdrukking

$$\langle \mathbf{x}, \mathbf{y} \rangle = 2x_1 y_1 + 2x_2 y_2 + 5x_3 y_3 - x_1 y_2 - x_2 y_1 - 2x_1 y_3 - 2x_3 y_1 + x_2 y_3 + x_3 y_2$$

een inwendig product op $\mathbb{R}^3$ definieert.
Om dit in te zien schrijven we eerst:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 2 & -1 & -2 \\ -1 & 2 & 1 \\ -2 & 1 & 5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Er volgt direct dat aan de eerste, tweede en derde eis voor een inwendig product voldaan is.
Tenslotte controleren we of aan eis vier voldaan is.

$$\langle \mathbf{x}, \mathbf{x} \rangle = 2x_1^2 + 2x_2^2 + 5x_3^2 - 2x_1 x_2 - 4x_1 x_3 + 2x_2 x_3 =$$

Ga dit schrijven als sommatie kwadraten, hierbij uitgaande van de dubbelproducten:

$$= (x_1 - x_2)^2 + (x_1 - 2x_3)^2 + (x_2 + x_3)^2$$

Het is duidelijk dat dit altijd ≥ 0 is. We zien ook dat als $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ dat dan volgt: $x_1 - x_2 = 0$, $x_1 - 2x_3 = 0$ en $x_2 + x_3 = 0$. Uit oplossen van het stelsel volgt snel dat $\mathbf{x} = \mathbf{0}$. Dus de conclusie volgt. $\square$

Wil het via kwadraat afsplistsen niet lukken dan willen de volgende stellingen nog wel eens helpen. We presenteren geen bewijs.

STELLING 5.10. *Laat A een reële symmetrische $n \times n$ matrix zijn. Equivalent zijn:*

1. *De matrix A is definiet positief.*
2. *Alle eigenwaarden van A zijn (strikt) positief.*

STELLING 5.11. *Laat A een reële symmetrische $n \times n$ matrix zijn. Equivalent zijn:*

1. *De matrix A is definiet positief.*
2. *Alle determinanten van de linksboven $k \times k$ ondermatrices van A zijn (strikt) positief.*

Loop zelf de drie bovenstaande voorbeelden opnieuw door met de twee criteria en beslis zelf hoe u het wilt doen.
Hoe ziet een niet–standaard inwendig product op $\mathbb{C}^n$ eruit?

STELLING 5.12. *Ieder inwendig product $\langle \mathbf{x}, \mathbf{y} \rangle$ op $\mathbb{C}^n$ is te schrijven als $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* A\mathbf{y}$, met een unieke Hermitische matrix A. Bovendien geldt dat $\langle \mathbf{x}, \mathbf{y} \rangle$ te schrijven is*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \overline{x_i} y_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle$$

Bewijs:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n, \mathbf{y} \rangle = \overline{x_1} \langle \mathbf{e}_1 \mathbf{y}, + \rangle \cdots + \overline{x_n} \langle \mathbf{e}_n \mathbf{y}, = \rangle$$

$$= \begin{bmatrix} \overline{x_1} & \dots & \overline{x_n} \end{bmatrix} \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{y} \rangle \\ \vdots \\ \langle \mathbf{e}_n, \mathbf{y} \rangle \end{bmatrix} = \begin{bmatrix} \overline{x_1} & \dots & \overline{x_n} \end{bmatrix} \begin{bmatrix} \langle \mathbf{e}_1, y_1 \mathbf{e}_1 + \cdots + y_n \mathbf{e}_n \rangle \\ \vdots \\ \langle \mathbf{e}_n, y_1 \mathbf{e}_1 + \cdots + y_n \mathbf{e}_n \rangle \end{bmatrix}$$

$$= [\overline{x_1} \dots \overline{x_n}] \begin{bmatrix} y_1 \langle \mathbf{e}_1, \mathbf{e}_1 \rangle + \cdots + y_n \langle \mathbf{e}_1, \mathbf{e}_n \rangle \\ \vdots \\ y_1 \langle \mathbf{e}_n, \mathbf{e}_1 + \cdots + y_n \langle \mathbf{e}_n, \mathbf{e}_n \rangle \rangle \end{bmatrix} = [\overline{x_1} \dots \overline{x_n}] \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{e}_1 \rangle & \cdots & \langle \mathbf{e}_1, \mathbf{e}_n \rangle \\ \vdots & & \vdots \\ \langle \mathbf{e}_n, \mathbf{e}_1 \rangle & \cdots & \langle \mathbf{e}_1, \mathbf{e}_n \rangle \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

De uniciteit van de matrix A volgt uit het feit dat $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \mathbf{e}_i^T A \mathbf{e}_j = [A]_{i,j}$ en dit laat ook zien dat de matrix A Hermitisch is. Tenslotte, als we het product uitwerken volgt de laatste claim.

$\square$

Definitie: Laat A een Hermitische $n \times n$ matrix zijn. De matrix heet positief definiet, als $\mathbf{x}^* A\mathbf{x} > 0$, voor alle $\mathbf{x} \neq \mathbf{0}$.

STELLING 5.13. *Laat A een Hermitische $n \times n$ matrix zijn. Equivalent zijn:*

1. *De matrix A is definiet positief.*
2. *Alle eigenwaarden van A zijn (strikt) positief.*
3. *Alle determinanten van de linksboven $k \times k$ ondermatrices van A zijn (strikt) positief.*

Om bij een gegeven formule eis 4 na te gaan in het complexe geval via kwadraat afsplitsen is in het algemeen bijzonder lastig. We zullen één voorbeeld geven.

Voorbeeld 5.14. Beschouw de uitdrukking op $\mathbb{C}^2$, gedefinieerd door

$$\langle \mathbf{x}, \mathbf{y} \rangle = 2\overline{x_1}y_1 + (1+i)\overline{x_1}y_2 + (1-i)\overline{x_2}y_1 + \overline{x_2}y_2 = \overline{\mathbf{x}}^T A\mathbf{y} = \mathbf{x}^* A\mathbf{y}$$

met $A = \begin{bmatrix} 2 & 1+i \\ 1-i & 1 \end{bmatrix}$. Ga na of dit een inwendig product is op $\mathbb{C}^2$.
Merk op dat A Hermitisch is (dus $A^* = A$). De notatie $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* A\mathbf{y}$ geeft dat aan de eisen (2) en (3) voldaan is. Ook aan eis (1) is voldaan, want:

$$\overline{\langle \mathbf{y}, \mathbf{x} \rangle} = \langle \mathbf{y}, \mathbf{x} \rangle^* = (\mathbf{y}^* A\mathbf{x})^* = \mathbf{x}^* A^* \mathbf{y} = \mathbf{x}^* A\mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle .$$

(overigens kan dit natuurlijk ook ingezien worden door $\overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ uit te schrijven en te constateren dat het resultaat hetzelfde is als $\langle \mathbf{x}, \mathbf{y} \rangle$).
We moeten nu alleen nog na gaan of het product positief definiet is. Daartoe schrijven we $\langle \mathbf{x}, \mathbf{x} \rangle$ uit:

$$\left\langle \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\rangle = 2|x_1|^2 + (1+i)\overline{x_1}x_2 + (1-i)x_1\overline{x_2} + |x_2|^2 = |(1-i)x_1 + x_2|^2.$$

(Hierbij vind ik de laatste stap lastig!) Aan deze laatste vorm zien we dat $\overline{\langle \mathbf{x}, \mathbf{x} \rangle}$ altijd ≥ 0 is. Maar we zien ook dat bijvoorbeeld voor de vector $\mathbf{x} = \begin{bmatrix} -1 \\ 1-i \end{bmatrix}$ geldt: $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ terwijl $\mathbf{x} \neq \mathbf{0}$. Deze constatering toont aan dat de uitdrukking geen inwendig product is. □

Merk op dat criterium 3 van Stelling 5.13 ons wel makkelijk doet inzien dat in bovenstaand voorbeeld geen inwendig product gegeven wordt door de formule. Inderdaad, de matrix $A = \begin{bmatrix} 2 & 1+i \\ 1-i & 1 \end{bmatrix}$ heeft twee linkerboven ondermatrices, de matrix $[2]$, die een positieve determinant en de matrix A zelf, die determinant nul heeft.

Het komt niet zo heel erg vaak voor dat op $\mathbb{R}^n$ en $\mathbb{C}^n$ andere inwendig producten dan de standaard inwendig producten gebruikt worden. Als met functieruimten gewerkt wordt gebeurt dit wel veel vaker. Het volgende voorbeeld geeft verschillende inwendig producten op functieruimten. Zie overigens ook voortgezette analyse.

Voorbeeld 5.15. Als ρ een positieve (dus $\rho(x) > 0$ voor alle x uit het domein) continue functie op $[a, b]$ is, dan is elke uitdrukking van de vorm

$$\langle f, g \rangle = \int_a^b \rho(x)\overline{f(x)}g(x)\, dx$$

een inwendig product op $C([a,b],\mathbb{L})$. In één van de opgsven mag u dit controleren. (Maar het is ook een inwendig product op $PC([a,b],\mathbb{L})$ of op $L_2([a,b],\mathbb{L})$.) Een speciaal geval hiervan is het veel voorkomende inwendig product op $C([0,2\pi],\mathbb{L})$ gedefinieerd door

$$\langle f,g\rangle = \frac{1}{2\pi}\int_0^{2\pi} \overline{f(x)}g(x)\;dx$$

(voor ρ is dus een constante functie genomen).
Soms werken we bij functieruimten niet met functies waarvan het domein een interval van de vorm $[a,b]$ is, maar met functies waarvan het domein onbegrensd is. Je moet dan net even iets kritischer zijn omdat de integraal die het inwendig product definieert dan moet convergeren. Zo is

$$\langle f,g\rangle = \int_0^{\infty} \overline{f(x)}g(x)e^{-x}\;dx$$

een inwendig product op deelruimten van functies uit $\mathcal{F}([0,\infty),\mathbb{L})$ die de eigenschap hebben dat bovenstaande integraal convergeert voor elke tweetal functies uit die deelruimte. In dit geval is de verzameling $Pol([0,\infty),\mathbb{L})$ van alle polynomen een voorbeeld van zo'n deelruimte.
Evenzo is

$$\langle f,g\rangle = \int_{-\infty}^{\infty} \overline{f(x)}g(x)e^{-x^2}\;dx$$

een inwendig product op $Pol((-\infty,\infty),\mathbb{L})$

5.4 Meetkunde met behulp van een inwendig product

In de ruimten $\mathbb{R}^n$ en $\mathbb{C}^n$ is het standaard inwendig product gebruikt om allerlei meetkundige aspecten als lengte van een vector, afstand tussen twee vectoren, loodrechte stand van vectoren en orthogonale projecties op deelruimten te definiëren. In deze en de volgende paragraaf zullen we dit allemaal nog eens over doen, maar nu voor een willekeurige inwendig productruimte en er zal blijken dat vrijwel alles wat in $\mathbb{R}^n$ en $\mathbb{C}^n$ is gedaan, zogoed als letterlijk is over te nemen, tot aan de bewijzen toe.

Als eerste zal de lengte of norm van een vector gedefinieerd worden.

Lengte van een vector

Definitie: In een inwendig productruimte V wordt het getal

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v},\mathbf{v}\rangle}.$$

de *lengte* of *norm* van een vector $\mathbf{v}$ genoemd.

Omdat een inwendig product per definitie positief definiet is kan de wortel van $\langle \mathbf{v},\mathbf{v}\rangle$ altijd genomen worden en is de lengte blijkbaar goed gedefinieerd.

Voorbeeld 5.16. Om de lengte van de functie f, met voorschrift $f(x) = \sin x$, uit $C([0,2\pi],\mathbb{R})$ voorzien van het standaard inwendig product te berekenen, moet eerst

$$\langle f,f\rangle = \int_0^{2\pi} \sin^2 x\;dx$$

bepaald worden. Ga na dat dit π is. De lengte van $\sin x$ is in deze inwendig productruimte dus $\sqrt{\pi}$. Realiseer je overigens dat deze lengte NIETS met de lengte van de grafiek van f te maken heeft!
Voor de lengte van $f_n(x) = e^{inx}$ (n een geheel getal) als functie uit $C([0, 2\pi], \mathbb{C})$ voorzien van het inwendig product

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} \overline{f(x)}g(x)\, dx$$

geldt

$$\|f_n\|^2 = \langle f_n, f_n\rangle = \frac{1}{2\pi}\int_0^{2\pi} \overline{f_n(x)}f_n(x)\, dx = \frac{1}{2\pi}\int_0^{2\pi} e^{-inx}e^{inx}\, dx = \frac{1}{2\pi}\int_0^{2\pi} 1\, dx = 1$$

en dus $\|f_n\| = 1$.

We merken nog op dat de lengte van een vector niet alleen van de vector afhangt, maar ook van het gekozen inwendig product! Als we op $\mathbb{R}^3$ dus een ander inwendig product kiezen dan het standaard inwendig product, is de lengte van een vector in het algemeen niet de natuurlijke lengte.

De volgende stelling geeft een aantal eigenschappen van de norm. Ga na dat dit allemaal eigenschappen zijn, die we, op basis van onze intuïtie voor het begrip lengte, ook zouden willen hebben!

STELLING 5.17. *(Rekenregels voor lengte)*
Voor alle vectoren $\mathbf{v}$ *en* $\mathbf{w}$ *en voor elke scalar* c *uit een inwendig productruimte* V *over* $\mathbb{L}$ *geldt:*

- $\|\mathbf{v}\| \geq 0$;
- $\|\mathbf{v}\| = 0 \Longleftrightarrow \mathbf{v} = \mathbf{0}$;
- $\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$.

Eenheidsvector en normaliseren van een vector

Bij een vector $\mathbf{v}$ ongelijk aan nul willen we vaak een vector $\widehat{\mathbf{v}}$ met lengte 1 maken die dezelfde richting heeft als $\mathbf{v}$. De manier waarop deze gevonden wordt volgt direct uit de eigenschappen van norm: zoek een positieve reële scalar c zo dat $1 = \|c\mathbf{v}\| = |c|\|\mathbf{v}\| = c\|\mathbf{v}\|$. Blijkbaar voldoet $c = \frac{1}{\|\mathbf{v}\|}$ en we concluderen

$$\widehat{\mathbf{v}} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}.$$

Dit proces om van een vector een vector met lengte 1 te maken die dezelfde richting heeft, noemen we *normaliseren* en $\widehat{\mathbf{v}}$ wordt wel de bij $\mathbf{v}$ horende *genormaliseerde vector* genoemd. Een vector met lengte 1 heet ook wel een *eenheidsvector*.

Voorbeeld 5.18. De functie $\widehat{f}(x) = \frac{1}{\sqrt{\pi}}\sin x$ is de bij $f(x) = \sin x$ horende functie van lengte 1 die dezelfde "richting" heeft als $f(x) = \sin x$, tenminste, als we werken in $C([0, 2\pi], \mathbb{R})$ voorzien van het standaard inwendig product.

Afstand tussen vectoren

Nu we de lengte van een vector hebben, kunnen we ook de afstand tussen twee vectoren definiëren.

Definitie: In een inwendig productruimte V over $\mathbb{L}$ wordt de *afstand* $d(\mathbf{v}, \mathbf{w})$ tussen de vectoren $\mathbf{v}$ en $\mathbf{w}$ uit V gedefinieerd door

$$d(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\|.$$

Voorbeeld 5.19. De afstand tussen $f(x) = \sin x$ en $g(x) = \cos x$ in $C([0, 2\pi], \mathbb{R})$ voorzien van het standaard inwendig product is de wortel van

$$\|f - g\|^2 = \int_0^{2\pi} (f(x) - g(x))^2 \, dx = \int_0^{2\pi} (\sin^2 x + \cos^2 x - 2 \sin x \cos x) \, dx = 2\pi,$$

dus $\sqrt{2\pi}$.

Opmerking: De afstand tussen f en g is dus $\sqrt{\int_0^{2\pi} (f(x) - g(x))^2 \, dx}$. Omdat de wortel en het kwadraat niet tegen elkaar wegvallen, zal dit niet gelijk zijn aan $\int_0^{2\pi} |f(x) - g(x)| \, dx$, de oppervlakte tussen de grafieken. Toch is dit wel een goede intuïtie: "de afstand tussen twee functies is klein, als de oppervlakte tussen de grafieken klein is".

Loodrechte stand van vectoren
We definiëren weer:

Definitie: We zeggen dat vectoren $\mathbf{v}$ en $\mathbf{w}$ uit een inwendig productruimte V *loodrecht op elkaar* staan of *orthogonaal* zijn als

$$\langle \mathbf{v}, \mathbf{w} \rangle = 0.$$

We noteren dit als $\mathbf{v} \perp \mathbf{w}$.

Voorbeeld 5.20. In $C([0, 2\pi], \mathbb{C})$ voorzien van het inwendig product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(x)} g(x) \, dx$$

geldt dat voor verschillende gehele getallen n en m de functies $f_n(x) = e^{inx}$ en $f_m(x) = e^{imx}$ orthogonaal zijn:

$$\begin{aligned} \langle f_n, f_m \rangle &= \frac{1}{2\pi} \int_0^{2\pi} \overline{e^{inx}} e^{imx} \, dx = \frac{1}{2\pi} \int_0^{2\pi} e^{i(m-n)x} \, dx \\ &= \frac{1}{2\pi i(m-n)} e^{i(m-n)x} \Big|_0^{2\pi} = 0 \end{aligned}$$

Eén van de meest bekende stellingen die met loodrechte stand te maken heeft is de stelling van Pythagoras. Deze blijkt in elke inwendig productruimte te gelden!

STELLING 5.21 (Stelling van Pythagoras)**.** *Als in een inwendig productruimte V de vectoren $\mathbf{u}$ en $\mathbf{v}$ orthogonaal zijn, dan geldt*

$$\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{u}+\mathbf{v}\|^2.$$

In een reële inwendig productruimte geldt de omkering ook, dus als daar bovenstaande gelijkheid geldt, dan volgt daaruit dat $\mathbf{u}$ en $\mathbf{v}$ orthogonaal zijn.

Het bewijs mag je zelf geven in een opgave.

5.5 Meetkunde met behulp van een inwendig product (vervolg)

Als op een vectorruimte V een inwendig product $\langle , \rangle$ aanwezig is, hebben we tot nu toe gedefinieerd:

- De lengte of norm van een vector: $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$.
- Eenheidsvectoren
- Normaliseren van een vector $\mathbf{v} \neq \mathbf{0}$.
- Afstand $d(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\|$ tussen de vectoren $\mathbf{v}, \mathbf{w}$.
- Loodrechte stand van vectoren.
- De stelling van Pythagoras.

De orthogonale projectie

Ook in inwendig product ruimten willen we de orthogonale projectie van de vector $\mathbf{x}$ op de vector $\mathbf{u}$ definiëren. Deze zullen we noteren met:

$$\mathrm{proj}_{\mathbf{u}}(\mathbf{x}).$$

Het is duidelijk dat $\mathrm{proj}_{\mathbf{u}}(\mathbf{x})$ een scalair veelvoud zal zijn van $\mathbf{u}$. We stellen daarom:

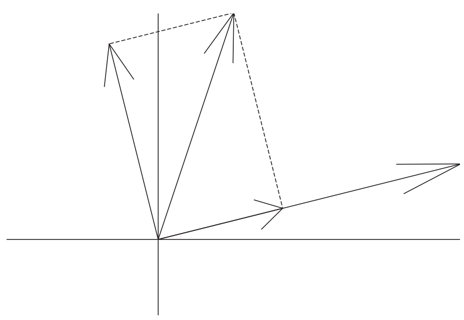
$$\mathrm{proj}_{\mathbf{u}}(\mathbf{x}) = \lambda\mathbf{u}$$

en het zal dat scalaire veelvoud zijn van $\mathbf{u}$ zo dat

$$\mathbf{x} - \lambda\mathbf{u} \perp \mathbf{u}.$$

Hieruit is λ te bepalen. Inderdaad:
$\mathbf{u} \perp \mathbf{x} - \lambda\mathbf{u} \Rightarrow \langle \mathbf{u}, \mathbf{x} - \lambda\mathbf{u}\rangle = 0 \Rightarrow \langle \mathbf{u}, \mathbf{x}\rangle - \lambda \langle \mathbf{u}, \mathbf{u}\rangle = 0$ en dus

$$\lambda = \frac{\langle \mathbf{u}, \mathbf{x}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}$$



Definitie: Laat $\mathbf{x}$ en $\mathbf{u}$ vectoren in de inwendig productruimte V zijn met $\mathbf{u} \neq \mathbf{0}$. De **orthogonale projectie** van $\mathbf{x}$ op $\mathbf{u}$ is gelijk aan de vector:

$$\mathrm{proj}_{\mathbf{u}}(\mathbf{x}) = \left(\frac{\langle \mathbf{u}, \mathbf{x} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\right) \mathbf{u}.$$

De coëfficiënt $\frac{\langle \mathbf{u}, \mathbf{x} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}$ heet de Fourier-coëfficiënt bij de projectie.

Beroemde ongelijkheden

We beginnen met de beroemde onglijkheid van Cauchy[3]-Schwarz[4].

STELLING 5.22 (De *ongelijkheid van Cauchy-Schwarz*)**.** *Voor alle vectoren* $\mathbf{v}$ *en* $\mathbf{w}$ *uit een inwendig productruimte* V *over* $\mathbb{L}$ *geldt:*

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \cdot \|\mathbf{w}\|.$$

Verder geldt: er treedt gelijkheid op (dus $|\langle \mathbf{v}, \mathbf{w} \rangle| = \|\mathbf{v}\| \cdot \|\mathbf{w}\|$*) als en slechts als* $\mathbf{v}$ *en* $\mathbf{w}$ *afhankelijk zijn.*

Bewijs: Als $\mathbf{w} = \mathbf{0}$ dan is de ongelijkheid duidelijk. Neem daarom aan dat $\mathbf{w} \neq \mathbf{0}$. Beschouw nu de projectie van $\mathbf{v}$ op $\mathbf{w}$: $\mathrm{proj}_{\mathbf{w}}(\mathbf{v}) = \left(\frac{\langle \mathbf{w}, \mathbf{v} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle}\right) \mathbf{w}$. Daarom: $\mathbf{v} - \mathrm{proj}_{\mathbf{w}}(\mathbf{v}) \perp \mathbf{w}$. Pythagoras geeft:

$$\|\mathbf{v} - \mathrm{proj}_{\mathbf{w}}(\mathbf{v})\|^2 + \|\mathrm{proj}_{\mathbf{w}}(\mathbf{v})\|^2 = \|\mathbf{v}\|^2$$

en dus: $\|\mathrm{proj}_{\mathbf{w}}(\mathbf{v})\|^2 \leq \|\mathbf{v}\|^2$ en dus $\|\mathrm{proj}_{\mathbf{w}}(\mathbf{v})\| \leq \|\mathbf{v}\|$. Dus:

$$\left\| \left(\frac{\langle \mathbf{w}, \mathbf{v} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle}\right) \mathbf{w} \right\| \leq \|\mathbf{v}\|$$

en dus:

$$\left(\frac{|\langle \mathbf{w}, \mathbf{v} \rangle|}{\|\mathbf{w}\|^2}\right) \|\mathbf{w}\| \leq \|\mathbf{v}\|$$

en de ongelijkheid volgt.
Stel nu dat $|\langle \mathbf{v}, \mathbf{w} \rangle| = \|\mathbf{v}\| \cdot \|\mathbf{w}\|$. Bekijk bovenstaand bewijs. Als $\mathbf{w} = \mathbf{0}$ (dan is er gelijkheid) dan zijn $\mathbf{v}, \mathbf{w}$ duidelijk afhankelijk. Als $\mathbf{w} \neq \mathbf{0}$, dan laat het bewijs zien dat gelijkheid impliceert dat $\|\mathbf{v} - \mathrm{proj}_{\mathbf{w}}(\mathbf{v})\| = 0$ en dus $\mathbf{v} - \mathrm{proj}_{\mathbf{w}}(\mathbf{v}) = \mathbf{0}$. Omdat dan $\mathrm{proj}_{\mathbf{w}}(\mathbf{v}) = \mathbf{v}$ en omdat $\mathrm{proj}_{\mathbf{w}}(\mathbf{v})$ zeker een scalair veelvoud van $\mathbf{w}$ is, volgt: $\mathbf{v}$ is een scalair veelvoud van $\mathbf{w}$. Dus $\mathbf{v}, \mathbf{w}$ zijn afhankelijk. $\square$

[3]Augustin Louis Cauchy (1789-1857). Abel schreef in 1826 over hem: "Cauchy is gek en daar is niets aan te doen, maar hij is op dit moment de enige die weet hoe wiskunde gedaan moet worden." Cauchy heeft zich met veel beziggehouden. Hij heeft onder andere de theorie ontwikkeld voor Fourier-getransformeerden en diagonalisatie van matrices.

[4]Karl Herman Amandus Schwarz (1843-1921). Hij schreef ter gelegenheid van de 70-ste verjaardag van Weierstrass een Festschrift, waarin onder andere de ongelijkheid van Cauchy-Schwarz voorkomt en waar Picard zijn existentie-bewijs van oplossingen van differentiaalvergelijkingen op gebaseerd heeft.

STELLING 5.23 (De *driehoeksongelijkheid*)**.** *Voor alle vectoren* $\mathbf{v}$ *en* $\mathbf{w}$ *uit een inwendig productruimte* V *over* $\mathbb{L}$ *geldt:*

$$\|\mathbf{v}+\mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$$

Bewijs: Voor het bewijs van de driehoeksongelijkheid wordt $\|\mathbf{v}+\mathbf{w}\|^2$ uitgewerkt. Hierbij passen we bij de eerste ongelijkheid toe dat voor complexe getallen z geldt dat $\operatorname{Re} z \leq |z|$ en bij de tweede de ongelijkheid van Cauchy-Schwarz.

$$\begin{aligned}
\|\mathbf{v}+\mathbf{w}\|^2 &= \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w} \rangle \\
&= \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle \\
&= \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle + \overline{\langle \mathbf{v}, \mathbf{w} \rangle} + \langle \mathbf{w}, \mathbf{w} \rangle \\
&= \langle \mathbf{v}, \mathbf{v} \rangle + 2\operatorname{Re}(\langle \mathbf{v}, \mathbf{w} \rangle) + \langle \mathbf{w}, \mathbf{w} \rangle \\
&\leq \langle \mathbf{v}, \mathbf{v} \rangle + 2|\langle \mathbf{v}, \mathbf{w} \rangle| + \langle \mathbf{w}, \mathbf{w} \rangle \\
&= \|\mathbf{v}\|^2 + 2|\langle \mathbf{v}, \mathbf{w} \rangle| + \|\mathbf{w}\|^2 \\
&\leq \|\mathbf{v}\|^2 + 2\|\mathbf{v}\| \cdot \|\mathbf{w}\| + \|\mathbf{w}\|^2 \\
&= (\|\mathbf{v}\| + \|\mathbf{w}\|)^2
\end{aligned}$$

Door nu aan beide zijden de wortel te nemen vinden we de gezochte ongelijkheid. □

Voorbeeld 5.24. Als we $\mathbb{R}^n$ voorzien van het standaard inwendig product, dan geldt voor de vectoren $\mathbf{v} = (1, 1, \ldots, 1)$ en $\mathbf{w} = (\frac{1}{1}, \frac{1}{2}, \ldots, \frac{1}{n})$ dat

$$\|\mathbf{v}\| = \sqrt{n}, \quad \|\mathbf{w}\| = \sqrt{\frac{1}{1} + \frac{1}{4} + \cdots + \frac{1}{n^2}} \quad \text{en} \quad \langle \mathbf{v}, \mathbf{w} \rangle = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n}.$$

Uit de ongelijkheid van Cauchy-Schwarz volgt dan

$$\frac{1}{\sqrt{n}} \left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} \right) \leq \sqrt{\frac{1}{1} + \frac{1}{4} + \cdots + \frac{1}{n^2}}.$$

Juist voor dit type ongelijkheden wordt Cauchy-Schwarz vaak gebruikt.

Er is ook nog een zogenaamde *omgekeerde driehoeksongelijkheid* Zie hiervoor opgave **??**.

Orthogonaal complement

Loodrechte stand van vectoren kunnen we uitbreiden tot loodrechte stand van een vector op een deelruimte.

Definitie: We zeggen dat een vector $\mathbf{x}$ uit een inwendig productruimte V *loodrecht op de deelruimte* W van V staat als $\mathbf{x}$ loodrecht staat op iedere vector uit W. Notatie: $\mathbf{x} \perp W$.
Het *orthogonale complement* van W is de verzameling die bestaat uit alle vectoren die loodrecht staan op W. We noteren deze verzameling als $W^{\perp}$.

Voorbeeld 5.25. Het is duidelijk dat in een inwendig productruimte V geldt dat elke vector loodrecht staat op de nulvector. Dat betekent dat het orthogonale complement van de deelruimte $\{\mathbf{0}\}$ heel V is, dus $\{\mathbf{0}\}^{\perp} = V$. Ga zelf na dat $V^{\perp} = \{\mathbf{0}\}$.

Voorbeeld 5.26. Ondanks het feit dat in Pol($\mathbb{R}$) met inwendig product

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$$

geldt dat het polynoom x^2 loodrecht staat op het polynoom x (ga dit na), behoort x^2 niet tot het orthogonale complement van $W = \text{Span}\{1, x\}$. Dat komt omdat x^2 niet loodrecht staat op elke vector uit W. Zo geldt bijvoorbeeld dat x^2 niet loodrecht staat op het constante polynoom 1 (ga dit ook na).
De vector $p(x) = 3x^2 - 1$ behoort wel tot het orthogonale complement van W: snel is in te zien dat deze vector loodrecht staat op beide opspannende vectoren 1 en x van W, waaruit geconcludeerd kan worden dat $3x^2 - 1$ dan ook loodrecht staat op *elke* vector uit W. Deze zijn namelijk van de vorm $c_1 \cdot 1 + c_2 \cdot x$ en

$$\langle p(x), c_1 \cdot 1 + c_2 \cdot x\rangle = c_1\,\langle p(x), 1\rangle + c_2\,\langle p(x), x\rangle = 0$$

waaruit het volgt.

Bovenstaand voorbeeld kan gegeneraliseerd worden tot de volgende stelling, die een handige methode geeft om voor een eindig dimensionale deelruimte W het orthogonale complement $W^{\perp}$ te berekenen of te bepalen of een vector loodrecht op W staat (met andere woorden, bij het bepalen of die vector tot het orthogonale complement $W^{\perp}$ behoort).

STELLING 5.27. *Het orthogonale complement van een deelruimte W van een inwendig productruimte V is een deelruimte van V.*

Het bewijs van deze stelling mag je zelf geven.
Ook geldt weer:

STELLING 5.28. *Als W een deelruimte van een inwendig productruimte V is, zeg $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p, \ldots\}$, dan behoort $\mathbf{w}$ tot $W^{\perp}$ als en slechts als $\mathbf{w} \perp \mathbf{v}_i$ voor alle i.*

Ook deze stelling mag je zelf bewijzen.
Veel eigenschappen die we kennen uit het eindig dimensionale geval, gelden nog steeds. Bijvoorbeeld:

$$W \cap W^{\perp} = \emptyset$$

Andere eigenschappen niet, zoals het volgende voorbeeld laat zien.

Voorbeeld 5.29. Beschouw $V = \ell_2(\mathbb{R})$ met het standaard inwendig product. Beschouw de deelruimte $W = \text{Span}\{\mathbf{e}_1, \ldots, \mathbf{e}_n, \ldots\}$. Er geldt: $W \neq V$ maar er geldt ook: $W^{\perp} = \{\mathbf{0}\}$.
Want: als $\mathbf{x} = (x_1, \ldots, x_n, \ldots) \in W^{\perp}$, dan $\mathbf{x} \perp \mathbf{e}_i$ voor alle i, dus $0 = \langle \mathbf{x}, \mathbf{e}_i\rangle = x_i$ voor alle i. Als $x_i = 0$ voor alle i, dan $\mathbf{x} = \mathbf{0}$.

Opmerking: Merk op dat Voorbeeld 5.29 geldt:

$$\left(W^{\perp}\right)^{\perp} \neq W.$$

Inderdaad, $\left(W^{\perp}\right)^{\perp} = \{\mathbf{0}\}^{\perp} = V \neq W$.

5.6 Orthogonale verzamelingen en de Grammatrix

We bouwen door op orthogonaliteit, op dezelfde manier zoals we in $\mathbb{R}^n$ en $\mathbb{C}^n$ gedaan hebben met het standaard inwendig product. We "herhalen" de volgende definitie.

Definitie: Een (eventueel oneindige) verzameling vectoren S in V heet een *orthogonale verzameling vectoren* als elk tweetal vectoren uit S orthogonaal is. Als ook nog geldt dat elke vector een eenheidsvector is (dus lengte 1 heeft), dan wordt het ook wel een *orthonormale verzameling* genoemd.

Voorbeeld 5.30. Uit voorbeeld 5.20 blijkt dat de functies f_0, f_1, f_{-1}, f_2, f_{-2},... met voorschrift $f_n(x) = e^{inx}$ een (oneindige) orthogonale verzameling vormen in $C([0, 2\pi], \mathbb{C})$ voorzien van het inwendig product

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} \overline{f(x)}g(x)\, dx.$$

Uit voorbeeld 5.16 blijkt bovendien dat elke f_n een eenheidsvector is, en dus vormen de functies zelfs een orthonormale verzameling[5].

De eerste, van $\mathbb{R}^n$ en $\mathbb{C}^n$ bekende,prettige eigenschap van orthogonale verzamelingen is dat deze onafhankelijk zijn (tenminste..., als ze de nulvector niet bevatten).

STELLING 5.31. *Een (eventueel oneindige) orthogonale verzameling niet-nulvectoren is lineair onafhankelijk.*

Het bewijs van deze stelling mag U zelf geven.

Net als in $\mathbb{R}^n$ en $\mathbb{C}^n$ introduceren we het begrip orthogonale basis.

Definitie: Een (eventueel oneindige) basis B van een vectorruimte V heet een *orthogonale basis* (respectievelijk een *orthonormale basis*) als deze, behalve basis, als verzameling orthogonaal (respectievelijk orthonormaal) is

Eén van de prettige eigenschappen is dan dat de coördinaten van vectoren ten opzichte van zo'n orthogonale basis relatief snel te vinden zijn. Net als in $\mathbb{R}^n$ en $\mathbb{C}^n$ geldt:

[5] Mede om deze reden zijn deze functies bruikbaar voor het bepalen van Fourierreeksen, zie voortgezette analyse.

STELLING 5.32. *Als $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ een eindige orthogonale basis is van de vectorruimte V, dan geldt voor iedere $\mathbf{x} \in V$:*

$$\mathbf{x} = \frac{\langle \mathbf{b}_1, \mathbf{x} \rangle}{\langle \mathbf{b}_1, \mathbf{b}_1 \rangle}\mathbf{b}_1 + \cdots + \frac{\langle \mathbf{b}_n, \mathbf{x} \rangle}{\langle \mathbf{b}_n, \mathbf{b}_n \rangle}\mathbf{b}_n,$$

met andere woorden, $[\mathbf{x}]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ *met* $c_i = \dfrac{\langle \mathbf{b}_i, \mathbf{x} \rangle}{\langle \mathbf{b}_i, \mathbf{b}_i \rangle}$, *voor* $i = 1, \dots, n$.

Het bewijs van deze stelling mag U zelf geven.
Merk op dat de vorige stelling zegt de i-de kental van $[\mathbf{x}]_B$ de Fourier–coëfficiënt is van de projectie van $\mathbf{x}$ op de vector $\mathbf{b}_i$. Als de basis bovendien orthonormaal is dan krijgen we:

$$[\mathbf{x}]_B = \begin{bmatrix} \langle \mathbf{b}_1, \mathbf{x} \rangle \\ \langle \mathbf{b}_2, \mathbf{x} \rangle \\ \vdots \\ \langle \mathbf{b}_n, \mathbf{x} \rangle \end{bmatrix}$$

Ook het bepalen van de transformatiematrix matrix $\underset{B \leftarrow B}{[\mathcal{L}]}$ van een lineaire afbeelding $\mathcal{L} : V \to W$ gaat snel als W een inwendig productruimte is met een eindige orthonormale basis B. Dat komt omdat de kolommen van deze matrix coördinatiseringen van vectoren uit W ten opzichte van de basis B zijn en deze zijn snel te berekenen.

Voorbeeld 5.33. Beschouw $\mathbb{R}^3$ (met standaard inwendig product) met orthonormale basis $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$, waarbij:

$$\mathbf{b}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \text{en} \quad \mathbf{b}_3 = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$$

Definieer de operator $\mathcal{L} : \mathbb{R}^3 \to \mathbb{R}^3$ door

$$\mathcal{L}(\mathbf{x}) = \mathcal{L}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 - x_2 \\ x_1 + x_2 + x_3 \\ x_1 - x_3 \end{bmatrix}.$$

We bepalen de matrix van $\mathcal{L}$ ten opzichte van de basis B, dus de matrix $\underset{B \leftarrow B}{[\mathcal{L}]}$. (Merk op dat de matrix $\underset{E \leftarrow E}{[\mathcal{L}]}$ (hier is E de standaardbasis), de standaardmatrix dus, zo is af te lezen: $\underset{E \leftarrow E}{[\mathcal{L}]} = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$. Vroeger zouden we $\underset{B \leftarrow B}{[\mathcal{L}]}$ berekend hebben via: $\underset{B \leftarrow B}{[\mathcal{L}]} = \underset{B \leftarrow E}{P}\underset{E \leftarrow E}{[\mathcal{L}]}\underset{E \leftarrow B}{P}$.)

De eerste kolom is de coördinatisering van $\mathcal{L}(\mathbf{b}_1) = \frac{1}{\sqrt{2}}\begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$ ten opzichte van de basis B en dat is, net als bij bovenstaand voorbeeld

$$[\mathcal{L}(\mathbf{b}_1)]_B = \begin{bmatrix} \langle \mathbf{b}_1, \mathcal{L}(\mathbf{b}_1) \rangle \\ \langle \mathbf{b}_2, \mathcal{L}(\mathbf{b}_1) \rangle \\ \langle \mathbf{b}_3, \mathcal{L}(\mathbf{b}_1) \rangle \end{bmatrix} = \begin{bmatrix} 1 \\ 4/\sqrt{6} \\ 2/\sqrt{12} \end{bmatrix}.$$

De tweede kolom is de coördinatisering van $\mathcal{L}(\mathbf{b}_2) = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ ten opzichte van de basis B, dus

$$[\mathcal{L}(\mathbf{b}_2)]_B = \begin{bmatrix} \langle \mathbf{b}_1, \mathcal{L}(\mathbf{b}_2)\rangle \\ \langle \mathbf{b}_2, \mathcal{L}(\mathbf{b}_2)\rangle \\ \langle \mathbf{b}_3, \mathcal{L}(\mathbf{b}_2)\rangle \end{bmatrix} = \begin{bmatrix} 3/\sqrt{6} \\ 0 \\ 3/\sqrt{18} \end{bmatrix},$$

en, tenslotte, de derde kolom is de coördinatisering van $\mathcal{L}(\mathbf{b}_3) = \frac{1}{\sqrt{6}} \begin{bmatrix} -4 \\ 2 \\ -2 \end{bmatrix}$ ten opzichte van de basis B, dus

$$[\mathcal{L}(\mathbf{b}_3)]_B = \begin{bmatrix} \langle \mathbf{b}_1, \mathcal{L}(\mathbf{b}_3)\rangle \\ \langle \mathbf{b}_2, \mathcal{L}(\mathbf{b}_3)\rangle \\ \langle \mathbf{b}_3, \mathcal{L}(\mathbf{b}_3)\rangle \end{bmatrix} = \begin{bmatrix} -6/\sqrt{12} \\ 0 \\ 1 \end{bmatrix}.$$

Blijkbaar

$$\underset{B\leftarrow B}{[\mathcal{L}]} = \begin{bmatrix} \langle \mathbf{b}_1, \mathcal{L}(\mathbf{b}_1)\rangle & \langle \mathbf{b}_1, \mathcal{L}(\mathbf{b}_2)\rangle & \langle \mathbf{b}_1, \mathcal{L}(\mathbf{b}_3)\rangle \\ \langle \mathbf{b}_2, \mathcal{L}(\mathbf{b}_1)\rangle & \langle \mathbf{b}_2, \mathcal{L}(\mathbf{b}_2)\rangle & \langle \mathbf{b}_2, \mathcal{L}(\mathbf{b}_3)\rangle \\ \langle \mathbf{b}_3, \mathcal{L}(\mathbf{b}_1)\rangle & \langle \mathbf{b}_3, \mathcal{L}(\mathbf{b}_2)\rangle & \langle \mathbf{b}_3, \mathcal{L}(\mathbf{b}_3)\rangle \end{bmatrix} = \begin{bmatrix} 1 & 3/\sqrt{6} & -6/\sqrt{12} \\ 4/\sqrt{6} & 0 & 0 \\ 2/\sqrt{12} & 3/\sqrt{18} & 1 \end{bmatrix}$$

Opmerking: In bovenstaand voorbeeld gold voor de matrix van $\mathcal{L}$ ten opzichte van de basis B dat het kental van de i-de rij j-de kolom gelijk is aan $\langle \mathbf{b}_i, \mathcal{L}(\mathbf{b}_j)\rangle$. Dit is geen toeval: als B een eindige ortho*normale* basis voor de inwendig productruimte V is, dan geldt dit voor elke operator $\mathcal{L} : V \to V$.

Gram-matrices

In een *eindig dimensionale* inwendig productruimte bestaat een eindige basis B. In dit geval kan de (voor de hand liggende) vraag gesteld worden of het inwendig product van twee vectoren $\mathbf{x}$ en $\mathbf{y}$ berekend kan worden via de B-coördinaten $[\mathbf{x}]_B$ en $[\mathbf{y}]_B$. Tenslotte, dit doen we ook in $\mathbb{R}^n$ (en $\mathbb{C}^n$), maar dan via de coördinaten t.o.v. de standaardbasis. Dit blijkt inderdaad te kunnen, het inwendig product kan geassocieerd worden via een matrix met de coördinaten. Die matrix wordt de Grammatrix van het inwendig product bij die basis genoemd. In het voorgaande is deze matrix al een paar keer gebruikt. We geven nu eerst de nette definitie.

Definitie: In een eindig dimensionale inwendig productruimte V van dimensie n met basis $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ wordt de *Gram-matrix van het inwendig product ten opzichte van de basis* B gedefinieerd door

$$G_B = \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{v}_1\rangle & \cdots & \langle \mathbf{v}_1, \mathbf{v}_n\rangle \\ \vdots & & \vdots \\ \langle \mathbf{v}_n, \mathbf{v}_1\rangle & \cdots & \langle \mathbf{v}_n, \mathbf{v}_n\rangle \end{bmatrix}.$$

Voorbeeld 5.34. Ga na dat op $\mathbb{R}^n$ of $\mathbb{C}^n$ (met het standaard inwendig product) de Gram-matrix van de standaardbasis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ de $n \times n$ eenheidsmatrix is.
Ga na dat de matrices uit de voorbeelden 5.6 en 5.14 de Gram-matrices van de betreffende inwendig producten ten opzichte van de standaardbases zijn.

De volgende stelling geeft aan hoe het inwendig product van vectoren te berekenen via de coördinaten t.o.v. een of andere basis B.

STELLING 5.35. *In een eindig dimensionale inwendig productruimte V met basis $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ geldt voor alle vectoren $\mathbf{u}$ en $\mathbf{w}$ dat*

$$\langle \mathbf{u}, \mathbf{w} \rangle = ([\mathbf{u}]_B)^* \; G_B \; [\mathbf{w}]_B = \overline{([\mathbf{u}]_B)}^T \; G_B \; [\mathbf{w}]_B.$$

Bewijs: Beschouw de vectoren $\mathbf{u}$ en $\mathbf{w}$. Er zijn scalairen $a_1, \ldots, a_n$ en $b_1, \ldots, b_n$ zo dat $\mathbf{u} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ en $\mathbf{w} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$. Er volgt:

$$\langle \mathbf{u}, \mathbf{w} \rangle = \langle a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n, \mathbf{w} \rangle = \overline{a_1} \langle \mathbf{v}_1, \mathbf{w} \rangle + \cdots + \overline{a_n} \langle \mathbf{v}_n, \mathbf{w} \rangle =$$

$$= \begin{bmatrix} \overline{a_1} & \cdots & \overline{a_n} \end{bmatrix} \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{w} \rangle \\ \vdots \\ \langle \mathbf{v}_n, \mathbf{w} \rangle \end{bmatrix} = \begin{bmatrix} \overline{a_1} & \cdots & \overline{a_n} \end{bmatrix} \begin{bmatrix} \langle \mathbf{v}_1, b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n \rangle \\ \vdots \\ \langle \mathbf{v}_n, b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n \rangle \end{bmatrix} =$$

$$= \begin{bmatrix} \overline{a_1} & \cdots & \overline{a_n} \end{bmatrix} \begin{bmatrix} b_1 \langle \mathbf{v}_1, \mathbf{v}_1 \rangle + \cdots + b_n \langle \mathbf{v}_1, \mathbf{v}_n \rangle \\ \vdots \\ b_1 \langle \mathbf{v}_n, \mathbf{v}_1 \rangle + \cdots + b_n \langle \mathbf{v}_n, \mathbf{v}_n \rangle \end{bmatrix} =$$

$$= \begin{bmatrix} \overline{a_1} & \cdots & \overline{a_n} \end{bmatrix} \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{v}_1, \mathbf{v}_n \rangle \\ \vdots & & \vdots \\ \langle \mathbf{v}_n, \mathbf{v}_1 \rangle & \cdots & \langle \mathbf{v}_n, \mathbf{v}_n \rangle \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \overline{([\mathbf{u}]_B)}^T \; G_B \; [\mathbf{w}]_B,$$

hetgeen te bewijzen was. □

Nu wat eigenschappen wat de Grammatrix.

STELLING 5.36 (Eigenschappen van de Gram-matrix)**.** *In het volgende is G_B de Gram-matrix van een inwendig product op een eindig dimensionale inwendig productruimte ten opzichte van een basis B voor die ruimte. Dan geldt*

- *G_B is inverteerbaar*
- *G_B is Hermitisch, dus $G_B^* = G_B$. In het reële geval is G symmetrisch.*
- *G_B is definiet positief.*
- *De getallen op de diagonaal van G_B zijn reëel en positief.*
- *B is een orthogonale basis als en slechts als G_B een diagonaalmatrix is.*
- *B is een orthonormale basis als en slechts als G_B de eenheidsmatrix is.*

Bewijs: We bewijzen de eerste uitspraak. We tonen aan dat de Gram–matrix G_B inverteerbaar is door te laten zien dat $\mathrm{Nul}(G_B) = \{\mathbf{0}\}$. Alsvolgt: stel dat $G_B\mathbf{x} = \mathbf{0}$. Dan geldt zeker $\mathbf{x}^* G_B \mathbf{x} = 0$. Kies $\mathbf{v} \in V$ met $[\mathbf{v}]_B = \mathbf{x}$. Volgens Stelling 5.35 geldt:

$$\langle \mathbf{v}, \mathbf{v} \rangle = [\mathbf{v}]_B{}^* A [\mathbf{v}]_B = \mathbf{x}^* A \mathbf{x} = 0$$

Dus er geldt zeker: $\mathbf{v} = \mathbf{0}$ en dus ook $\mathbf{x} = [\mathbf{v}]_B = [\mathbf{0}]_B = \mathbf{0}$.
De inverteerbaarheid van G_B volgt.
De derde uitspraak. Kies $\mathbf{z} \in \mathbb{C}^n$ en beschouw $\mathbf{w} = z_1\mathbf{b}_1 + \cdots + z_n\mathbf{b}_n \in V$. Dan $[\mathbf{w}]_B = \mathbf{z}$. Daarom: $\mathbf{z}^* G_B \mathbf{z} = ([\mathbf{w}]_B)^* \ G_B \ [\mathbf{w}]_B = \langle \mathbf{w}, \mathbf{w} \rangle \geq 0$ en $0 = \mathbf{z}^* G_B \mathbf{z}$ impliceert dat $([\mathbf{w}]_B)^* \ G_B \ [\mathbf{w}]_B = \langle \mathbf{w}, \mathbf{w} \rangle = 0$ dus $\mathbf{w} = \mathbf{0}$ en dus $\mathbf{z} = [\mathbf{w}]_B = \mathbf{0}$.
De rest van de uitspraken is bijna triviaal. $\square$

Nu komt een mijns inziens bijzonder verrassend resultaat. Stel dat we het inwendig product willen gaan berekenen via de coördinaten van een orthonormale basis B van een eindig dimensionale ruimte V. Dat ligt voor de hand want de voorafgaande stelling zegt dat dan $G_B = I$, de eenheidsmatrix, en dan is er altijds iets moois te verwachten.
Als $\mathbf{v}$ en $\mathbf{w}$ vectoren in V zijn, zeg

$$[\mathbf{v}]_B = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \quad \text{en} \quad [\mathbf{w}]_B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix},$$

dan volgt dat

$$\langle \mathbf{v}, \mathbf{w} \rangle = \overline{([\mathbf{v}]_B)}^T \ I \ [\mathbf{w}]_B = \overline{a_1} b_1 + \cdots + \overline{a_n} b_n.$$

Ten opzichte van deze basis is het product op de coördinatisering dus gelijk aan het standaard inwendig product op $\mathbb{L}^n$! Deze constatering is belangrijk genoeg om in een aparte stelling te vermelden.

STELLING 5.37. *Het inwendig product op een n-dimensionale reële (complexe) vectorruimte V komt in de coördinatisering ten opzichte van een orthonormale basis B overeen met het standaard inwendig product op $\mathbb{R}^n$ (dan wel $\mathbb{C}^n$).*

We hebben eerder gezien bij een eindigdimensionale vectorruimte V met basis B, dat de coördinatisering afbeelding

$$\mathbf{x} \rightsquigarrow [\mathbf{x}]_B$$

de lineaire structuur van V op de lineaire structuur van $\mathbb{R}^n$ (of $\mathbb{C}^n$) wordt gelegd. En dit gebeurt zodanig dat de lineaire structuren niet te onderscheiden (isomorf) zijn.
Nu zien we dat als V een eindig dimensionaal inwendig product ruimte is en als voor B een orthonormale basis gekozen wordt, de coördinatisering niet alleen de lineaire structuur bewaart, maar ook de inwendig product structuur van V op die van $\mathbb{R}^n$ (of $\mathbb{C}^n$) gelegd wordt.. Onder de coördinatisering wordt V gewoon als $\mathbb{R}^n$ (of $\mathbb{C}^n$) gezien met het standaard inwendig product. Alle resultaten van $\mathbb{R}^n$ (of $\mathbb{C}^n$) aangaande het standaard inwendig product kunnen naar V teruggehaald worden. En eigenlijk hebben we dat in de vorige paragrafen al zitten doen; lengte, Pythagoras, Cauchy-Schwarz etc. Eindig dimensionaal is niets nieuws te verwachten.
Een ander voorbeeld hiervan:

GEVOLG 5.38. *Als V een eindig dimensionaal inwendig productruimte is met* $\dim(V) = n$, *en als W een deelruimte van V is, dan geldt:*

$$\dim V = \dim W + \dim W^{\perp}.$$

(Maar we moeten nog wel aantonen dat iedere eindig dimensionale inwendig product ruimte V ook echt een orthonormale basis heeft.) Nieuwe lineaire algebra is te verwachten in de oneindig dimensionale inwendig product ruimten.

5.7 De Orthogonale projectie

Evenals in $\mathbb{R}^n$ en $\mathbb{C}^n$ zouden we de orthogonale projectie van een vector op een deelruimte W willen definiëren. We gebruiken dezelfde definitie als in deze vectorruimtes

Definitie: Gegeven zijn een deelruimte W van een inwendig productruimte V en een vector $\mathbf{x}$ uit V. Een *loodrechte* of *orthogonale projectie* van $\mathbf{x}$ op W is een vector $\mathbf{w}$ in W met de eigenschap dat $\mathbf{x} - \mathbf{w} \perp W$ (ofwel $\mathbf{x} - \mathbf{w} \in W^\perp$). We noteren $\mathbf{w} = \text{proj}_W(\mathbf{x})$.

Merk op dat op dit moment niet duidelijk is of projecties altijd zullen bestaan. Eenvoudige voorbeelden laten zien dat dit niet het geval is.

Voorbeeld 5.39. Beschouw nogmaals voorbeeld 5.29. Hier werd de ruimte $V = \ell_2(\mathbb{R})$ (met het standaard inwendig product) bekeken met de deelruimte $W = \text{Span}\{\mathbf{e}_1, \ldots, \mathbf{e}_n, \ldots\}$. We hebben daar laten zien dat $W \neq \ell_2(\mathbb{R})$ en $W^\perp = \{\mathbf{0}\}$. Dit laatste zal tot gevolg hebben dat als $\mathbf{x} \notin W$, dat er geen $\mathbf{w} \in W$ zal bestaan met $\mathbf{x} - \mathbf{w} \in W^\perp$. (Inderdaad, als $\mathbf{x} - \mathbf{w} \in W^\perp$ dan $\mathbf{x} - \mathbf{w} = \mathbf{0}$ en dus $\mathbf{w} = \mathbf{x} \notin W$. maar dan bestaat er dus geen orthogonale projectie van $\mathbf{x}$ op W!

Ook is het in eerste instantie niet duidelijk of de projectie, als die bestaat, uniek zal zijn. Dat geldt gelukkig wel.

STELLING 5.40. *Als de orthogonale projectie van* $\mathbf{x}$ *op* W *bestaat, dan is deze uniek.*

Het bewijs van deze stelling mag U zelf geven in opgave **??** van deze paragraaf. (Het is letterlijk hetzelfde bewijs als bij $\mathbb{C}^n$.

Nu de eigenschap die het nut van projecties verklaart:

STELLING 5.41. *Als de orthogonale projectie van een vector* $\mathbf{x}$ *uit een inwendig productruimte* V *op een deelruimte* W *bestaat, dan is het precies die vector uit* W *waarvoor de afstand tot* $\mathbf{x}$ *zo klein mogelijk is.*

Het bewijs van deze stelling mag U zelf geven in opgave **??** van deze paragraaf. Ook hier geldt dat het bewijs bekend van $\mathbb{C}^n$ gecopiëerd kan worden.

Voorbeeld 5.42. Beschouw nogmaals het voorbeeld 5.39. Bovenstaande stelling maakt ook duidelijk dat een vector $\mathbf{x} = (x_1, \ldots, x_n, \ldots) \notin W$ niet op W geprojecteerd kan worden. Bekijk namelijk $\mathbf{w}_n = (x_1, \ldots, x_n, 0, 0, 0, \ldots) \in W$. Het zal duidelijk zijn dat $\|\mathbf{x} - \mathbf{w}_n\|$, de afstand van $\mathbf{x}$ tot $\mathbf{w}_n$, gelijk is aan $\sqrt{\sum_{k=n+1}^{\infty} x_k^2}$ en dus $\lim_{n\to\infty} \|\mathbf{x} - \mathbf{w}_n\| = 0$. Dit laatste zal tot gevolg hebben dat er geen punt in W zal bestaan met kleinste afstand tot $\mathbf{x}$, en dus bestaat er geen projectie.

Voorbeeld 5.43. In Pol($\mathbb{R}$) met inwendig product

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\, dx$$

is de orthogonale projectie van $f(x) = x^3$ op de deelruimte $W = \mathrm{Pol}_2(\mathbb{R})$ gelijk aan $p(x) = \frac{3}{5}x$. Er geldt namelijk dat $p(x)$ tot W behoort en dat $f(x) - p(x) = x^3 - \frac{3}{5}x$ tot $W^\perp$ behoort. Het eerste is duidelijk en het tweede volgt uit het feit dat $W = \mathrm{Span}\{1, x, x^2\}$ en dat de inwendig producten $\left\langle 1, x^3 - \frac{3}{5}x\right\rangle$, $\left\langle x, x^3 - \frac{3}{5}x\right\rangle$ en $\left\langle x^2, x^3 - \frac{3}{5}x\right\rangle$ allemaal nul zijn, zodat $f(x) - p(x)$ loodrecht op de opspannende vectoren van W staat en dus (volgens stelling 5.28) inderdaad tot $W^\perp$ behoort. Uit bovenstaande stelling volgt dat het polynoom $p(x) = \frac{3}{5}x$ het polynoom van graad 2 of minder is, dat het dichtst bij het polynoom $f(x) = x^3$ ligt, tenminste, als we werken met het afstandsbegrip dat bepaald wordt door het inwendig product uit het voorbeeld. In dit geval is p dus het polynoom van graad 2 of minder waarbij

$$\|f - p\|^2 = \int_{-1}^{1} \big(f(x) - p(x)\big)^2\, dx \tag{5.1}$$

zo klein mogelijk is. Meetkundig zou je dus kunnen zeggen dat $p(x) = \frac{3}{5}x$ het polynoom van graad twee of minder is, waarvoor de oppervlakte tussen de grafieken van $p(x)$ en $f(x) = x^3$ op het interval $[-1, 1]$ zo klein mogelijk is. Dit laatste is natuurlijk niet helemaal waar omdat het om de integraal van het kwadraat van de verschilfunctie gaat, maar als de oppervlakte tussen grafieken klein is, is ook de integraal van het kwadraat van de verschilfunctie klein, en andersom.
Overigens is met formule (5.1) ook te begrijpen dat een ander inwendig product leidt tot een andere projectie.

Dit voorbeeld laat zien dat het relatief eenvoudig is om te controleren of een gegeven vector de orthogonale projectie is van een andere gegeven vector op een deelruimte W. Het voorbeeld laat niet zien hoe de projectie berekend kan worden.
Tenslotte:

STELLING 5.44. *(Decompositiestelling eindig dimensionale deelruimtes) Als V een inwendig productruimte is met een eindig dimensionale deelruimte W, dan geldt: Iedere $\mathbf{x}$ uit V is te schrijven als:*

$$\mathbf{x} = \mathbf{w} + \mathbf{w}^\sharp$$

(met $\mathbf{w} \in W$ en $\mathbf{w}^\sharp \in W^\perp$) op unieke wijze. En dus bestaat de orthogonale projectie $\mathrm{proj}_W : V \to V$ *op iedere eindig dimensionale deelruimte.*

Bewijs: Het bewijs is een copie van het bewijs voor $\mathbb{C}^n$. Dus neem een orthogonale basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ van W en beschouw

$$\mathbf{w} = \frac{\langle \mathbf{b}_1, \mathbf{x}\rangle}{\langle \mathbf{b}_1, \mathbf{b}_1\rangle}\mathbf{b}_1 + \cdots + \frac{\langle \mathbf{b}_n, \mathbf{x}\rangle}{\langle \mathbf{b}_n, \mathbf{b}_n\rangle}\mathbf{b}_n.$$

Laat zien dat $\mathbf{x} - \mathbf{w} \in W^\perp$ door aan te tonen dat $\mathbf{x} - \mathbf{w} \perp \mathbf{b}_1$, voor $i = 1, \ldots, n$. $\square$

Tenslotte:

STELLING 5.45. *De afbeelding* $\mathrm{proj}_W : V \to V$ *is voor een eindig dimensionale deelruimte* W *van* V *een goed gedefinieerde lineaire afbeelding.*

Bewijs: Dat de afbeelding goed gedefinieerd is, is hierboven al vermeld. Rest het aantonen van lineariteit. Neem daarvoor $\mathbf{x}$ en $\mathbf{y}$ in V en schrijf $\mathbf{p} = \mathrm{proj}_W(\mathbf{x}) + \mathrm{proj}_W(\mathbf{y})$. We moeten laten zien dat $\mathbf{p} = \mathrm{proj}_W(\mathbf{x} + \mathbf{y})$ dus dat $\mathbf{p} \in W$ en $(\mathbf{x} + \mathbf{y}) - \mathbf{p} \in W^\perp$. Het eerste is triviaal omdat W een deelruimte is en omdat $\mathrm{proj}_W(\mathbf{x})$ en $\mathrm{proj}_W(\mathbf{y})$ tot W behoren. Het tweede volgt uit het feit dat $W^\perp$ een deelruimte is en het feit dat zowel $\mathbf{x} - \mathrm{proj}_W(\mathbf{x})$ als $\mathbf{y} - \mathrm{proj}_W(\mathbf{y})$ tot $W^\perp$ behoren. Dat geldt dan namelijk ook voor de som van deze twee vectoren en dat is $(\mathbf{x} + \mathbf{y}) - \mathbf{p}$.
Neem nu $\mathbf{x}$ in V en $c \in \mathbb{L}$. Schrijf $\mathbf{p} = c\,\mathrm{proj}_W(\mathbf{x})$. We moeten laten zien dat $\mathbf{p} = \mathrm{proj}_W(c\mathbf{x})$ dus dat $\mathbf{p} \in W$ en $c\mathbf{x} - \mathbf{p} \in W^\perp$. Het eerste is duidelijk omdat W een deelruimte is en het tweede is waar omdat $W^\perp$ ook een deelruimte is en $c\mathbf{x} - \mathbf{p} = c(\mathbf{x} - \mathrm{proj}_W(\mathbf{x}))$. $\square$

Tenslotte de vraag: "Wat zijn nu precies de deelruimtes waarop we kunnen projecteren". Zonder bewijs vermelden we:

STELLING 5.46. *(Decompositiestelling eindig dimensionale deelruimtes) Als* V *een inwendig productruimte is en* W *een deelruimte, dan geldt: de orthogonale projectie* $\mathrm{proj}_W : V \to V$ *op* W *bestaat als en slechts als:* $W = (W^\perp)^\perp$.

5.8 Het bepalen van Orthogonale projecties

In het vorig college is de projectie van een vector op een (eindig dimensionale) deelruimte geïntroduceerd. Nu willen we de projectie effectief berekenen.

Methode 1

Als we bij de deelruimte W een orthogonale basis tot onze beschikking hebben, dan kunnen we de projectie van $\mathbf{x}$ verkrijgen door op iedere basis vector orthogonaal te projecteren en de projecties op te tellen. Het probleem van deze methode is wel dat een orthogonale basis voor W nodig is. Als die er in eerste instantie niet is, dan zal deze bepaald moeten worden. Een probleem wat we ook al in $\mathbb{C}^n$ kende.

STELLING 5.47. *Als* $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ *een orthogonale basis voor een deelruimte* W *van de inwendig productruimte* V *is dan geldt voor* $\mathbf{x} \in V$ *dat*

$$\mathrm{proj}_W(\mathbf{x}) = \frac{\langle \mathbf{b}_1, \mathbf{x} \rangle}{\langle \mathbf{b}_1, \mathbf{b}_1 \rangle}\mathbf{b}_1 + \cdots + \frac{\langle \mathbf{b}_n, \mathbf{x} \rangle}{\langle \mathbf{b}_n, \mathbf{b}_n \rangle}\mathbf{b}_n.$$

Bewijs: Zie het bewijs van de decompositiestelling uit college 13. $\square$

Voorbeeld 5.48. We gaan de projectie van $f(x) = x^3$ uit Pol($\mathbb{R}$) op $W = \text{Pol}_2(\mathbb{R})$ bepalen. Hier is Pol($\mathbb{R}$) voorzien van het inwendigproduct $\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)dx$.
Echter, de standaard basis $B = \{1, x, x^2\}$ is niet orthogonaal. Die moeten we dan eerst maken. Straks volgt een structurele methode om er één te vinden, maar voor W hebben we er eigenlijk al één gezien in voorbeeld 5.26. Daar is namelijk geconstateerd dat $3x^2 - 1$ loodrecht staat op de polynomen 1 en x. Omdat 1 en x ook loodrecht op elkaar staan, volgt dat $\{1, x, 3x^2 - 1\}$ een orthogonale verzameling niet nulvectoren is. Volgens stelling 5.31 zijn deze drie polynomen onafhankelijk en omdat ze duidelijk in de drie-dimensionale ruimte $\text{Pol}_2(\mathbb{R})$ liggen vormen ze een basis voor W. We hebben dus een orthogonale basis voor W te pakken en kunnen de orthogonale projectie uitrekenen met de formule van de vorige stelling:

$$\begin{aligned}\text{proj}_W(f) &= \frac{\langle 1, f(x)\rangle}{\langle 1,1\rangle}\cdot 1 + \frac{\langle x, f(x)\rangle}{\langle x,x\rangle}\cdot x + \frac{\langle 3x^2-1, f(x)\rangle}{\langle 3x^2-1, 3x^2-1\rangle}\cdot(3x^2-1)\\ &= 0\cdot 1 + \frac{2/5}{2/3}\cdot x + 0\cdot(3x^2-1) = \frac{3}{5}x\end{aligned}$$

hetgeen op basis van voorbeeld 5.26 te verwachten was.

Methode 2

We beschouwen het vorige voorbeeld nogmaals met de nieuwe methode.

Voorbeeld 5.49. We laten zien hoe in Pol($\mathbb{R}$) de orthogonale projectie $p(x)$ van $f(x) = x^3$ op $W = \text{Pol}_2(\mathbb{R})$ berekend kan worden als het inwendig product gedefinieerd is door

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\, dx.$$

Voor deze berekening hebben we een basis voor W nodig, bijvoorbeeld de standaardbasis $B = \{1, x, x^2\}$. Omdat $p(x)$ tot W moet behoren, is $p(x)$ een lineaire combinatie van de drie basisvectoren, dus

$$p(x) = c_1 + c_2x + c_3x^2$$

voor zekere getallen c_1, c_2 en c_3.
Verder moet $f(x) - p(x)$ loodrecht op W staan, wat betekent dat de inwendig producten van $f(x) - p(x)$ met de basisvectoren uit B nul moeten zijn, dus

$$\begin{aligned}\langle 1, f(x)-p(x)\rangle &= 0 \quad \text{ofwel} \quad \langle 1, p(x)\rangle = \langle 1, f(x)\rangle\\ \langle x, f(x)-p(x)\rangle &= 0 \quad \text{ofwel} \quad \langle x, p(x)\rangle = \langle x, f(x)\rangle\\ \left\langle x^2, f(x)-p(x)\right\rangle &= 0 \quad \text{ofwel} \quad \left\langle x^2, p(x)\right\rangle = \left\langle x^2, f(x)\right\rangle\end{aligned}$$

Invullen van $p(x) = c_1 + c_2x + c_3x^2$ leidt tot de volgende drie vergelijkingen met de onbekenden c_1, c_2 en c_3:

$$\begin{array}{llll} c_1\langle 1,1\rangle & +c_2\langle 1,x\rangle & +c_3\left\langle 1,x^2\right\rangle & =\langle 1, f(x)\rangle\\ c_1\langle x,1\rangle & +c_2\langle x,x\rangle & +c_3\left\langle x,x^2\right\rangle & =\langle x, f(x)\rangle\\ c_1\left\langle x^2,1\right\rangle & +c_2\left\langle x^2,x\right\rangle & +c_3\left\langle x^2,x^2\right\rangle & =\left\langle x^2, f(x)\right\rangle\end{array}$$

Al de inwendig producten in deze vergelijkingen kunnen berekend worden ($f(x)$ is immers x^3), wat het volgende stelsel lineaire vergelijkingen oplevert:

$$\begin{array}{rrrl} 2c_1 & & +\frac{2}{3}c_3 & = 0\\ & +\frac{2}{3}c_2 & & = \frac{2}{5}\\ \frac{2}{3}c_1 + & & +\frac{2}{5}c_3 & = 0\end{array}$$

met oplossing $c_1 = 0$, $c_2 = \frac{3}{5}$ en $c_3 = 0$. Daaruit volgt dat $p(x) = \frac{3}{5}x$, hetgeen in overeenstemming is met voorbeeld 5.43.

STELLING 5.50. *Stel $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ een basis is voor een eindig dimensionale deelruimte W van een inwendig productruimte V. Als $\mathbf{x}$ een vector is uit V, dan bestaat zijn orthogonale projectie*

$$\mathrm{proj}_W(\mathbf{x}) = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$$

en kunnen we de c_i vinden door het volgende lineaire stelsel op te lossen:

$$\begin{bmatrix} \langle \mathbf{b}_1, \mathbf{b}_1 \rangle & \cdots & \langle \mathbf{b}_1, \mathbf{b}_n \rangle \\ \vdots & & \vdots \\ \langle \mathbf{b}_n, \mathbf{b}_1 \rangle & \cdots & \langle \mathbf{b}_n, \mathbf{b}_n \rangle \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \langle \mathbf{b}_1, \mathbf{x} \rangle \\ \vdots \\ \langle \mathbf{b}_n, \mathbf{x} \rangle \end{bmatrix}.$$

Bewijs: We merken eerst op dat de coëfficiënten matrix van dit stelsel de Gram–matrix G_B is. Volgens Stelling 5.36 is deze matrix inverteerbaar, en dus is er een unieke oplossing van dit stelsel. Beschouw de unieke oplossing $c_1, \ldots, c_n$ en de vector:

$$\mathbf{w} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$$

Als we kunnen aantonen dat $\mathbf{x} - \mathbf{w} \in W^\perp$, dan geldt per definitie van de projectie:

$$\mathrm{proj}_W(\mathbf{x}) = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$$

Om $\mathbf{x} - \mathbf{w} \in W^\perp$ aan te tonen is voldoende te laten zien dat

$$\langle \mathbf{b}_i, \mathbf{x} - \mathbf{w} \rangle = 0, \quad \text{voor} \quad i = 1, \ldots, n.$$

Welnu, $\langle \mathbf{b}_i, \mathbf{x} - \mathbf{w} \rangle = \langle \mathbf{b}_i, \mathbf{x} \rangle - \langle \mathbf{b}_i, \mathbf{w} \rangle = \langle \mathbf{b}_i, \mathbf{x} \rangle - \langle \mathbf{b}_i, c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n \rangle = \langle \mathbf{b}_i, \mathbf{x} \rangle - (c_1 \langle \mathbf{b}_i, c_1\mathbf{b}_1 \rangle + \cdots + c_n \langle \mathbf{b}_i, \mathbf{b}_n \rangle) = \langle \mathbf{b}_i, \mathbf{x} \rangle - \langle \mathbf{b}_i, \mathbf{x} \rangle = 0$, voor $i = 1, \ldots, n$. □

Projecteren in $\mathbb{C}^n$ en projectiematrices

Omdat $\mathbb{C}^n$ een complexe inwendigproduct ruimte is, kunnen we ook op deelruimtes W van $\mathbb{C}^n$ projecteren. De bijbehorende projectie $\mathrm{proj}_W : \mathbb{C}^n \to \mathbb{C}^n$ is een lineaire afbeelding en de standaardmatrix van deze afbeelding heet weer een projectiematrix en wordt genoteerd met $[\mathrm{proj}_W]$. Kortom: de formule $\mathrm{proj}_W(\mathbf{z}) = [\mathrm{proj}_W]\mathbf{z}$ geldt ook complex. Hoe deze projectiematrix te verkrijgen?

STELLING 5.51. *Stel W een deelruimte is van $\mathbb{C}^n$ en stel dat $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ een <u>orthonormale basis</u> is voor W en schrijf $U = [\mathbf{u}_1 \ldots \mathbf{u}_k]$. Dan geldt:*

1. *Als S de ééndimensionale deelruimte $\mathrm{Span}\{\mathbf{u}_1\}$ is, dan: $[\mathrm{proj}_S] = \mathbf{u}_1\mathbf{u}_1^*$.*
2. *$[\mathrm{proj}_W] = \mathbf{u}_1\mathbf{u}_1^* + \cdots + \mathbf{u}_k\mathbf{u}_k^* = UU^*$.*

Bewijs: 1. Volgens stelling 5.47 geldt dat:

$$\mathrm{proj}_S(\mathbf{x}) = \frac{\langle \mathbf{u}_1, \mathbf{x} \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 = \frac{\langle \mathbf{u}_1, \mathbf{x} \rangle}{\|\mathbf{u}_1\|^2}\mathbf{u}_1 = \langle \mathbf{u}_1, \mathbf{x} \rangle \mathbf{u}_1 = (\mathbf{u}_1^*\mathbf{x})\mathbf{u}_1 = \mathbf{u}_1(\mathbf{u}_1^*\mathbf{x}) = (\mathbf{u}_1\mathbf{u}_1^*)\mathbf{x}$$

Merk op dat de scalaire vermenigvuldiging $c\mathbf{x}$ weer als matrixproduct $\mathbf{x}c$ geschreven is.
2. Volgens stelling 5.47 en eerste deel dit bewijs geldt dat:

$$\text{proj}_W(\mathbf{x}) = \frac{\langle \mathbf{u}_1, \mathbf{x}\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 + \cdots + \frac{\langle \mathbf{u}_k, \mathbf{x}\rangle}{\langle \mathbf{u}_k, \mathbf{u}_k\rangle}\mathbf{u}_n = (\mathbf{u}_1\mathbf{u}_1^*)\mathbf{x} + \cdots + (\mathbf{u}_k\mathbf{u}_k^*)\mathbf{x} =$$
$$= (\mathbf{u}_1\mathbf{u}_1^* + \cdots + \mathbf{u}_k\mathbf{u}_k^*)\mathbf{x} = UU^*\mathbf{x}$$

Merk op dat kolom-rij expansie is toegepast. □

Een bestudering van het bewijs van stelling 5.50 levert de volgende formule die de projectiematrix via een willekeurige basis geeft.

STELLING 5.52. *Stel dat W een deelruimte is van $\mathbb{C}^n$ en stel dat $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ een basis is voor W en schrijf $A = [\mathbf{a}_1 \ldots \mathbf{a}_k]$. Dan geldt:*

$$\text{proj}_W(\mathbf{x}) = A(A^*A)^{-1}A^*\mathbf{x}.$$

We zijn uiteindelijk in staat de spectraalstelling voor normale bases te geven. Herinner:

STELLING 5.53. (De (complexe) Spektraledecompositie Stelling, vorm 1)
Laat A een normale $n \times n$ matrix zijn. Dan bestaat een orthonormale basis $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ van $\mathbb{C}^n$ bestaande uit eigenvectoren van A, zeg $A\mathbf{u}_i = \lambda_i\mathbf{u}_i$, voor $i = 1, \ldots, n$. Er geldt:

$$\boxed{A = \lambda_1\mathbf{u}_1\mathbf{u}_1^* + \cdots + \lambda_n\mathbf{u}_n\mathbf{q}_n^*}$$

Met andere woorden: A is een lineaire combinatie van projectiematrices (van projecties op onderling loodrechte eigenvectoren)! Dit heet een spektrale decompositie van A Maar als een eigenwaarde een hogere multipliciteit heeft is deze spectraal decompositie niet uniek. Het extreme voorbeeld hiervan is weer de eenheidsmatrix.

Voorbeeld 5.54. Iedere orthonormale basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ van $\mathbb{C}^n$ is een orthonormale basis van $\mathbb{C}^n$ bestaande uit eigenvectoren (bij $\lambda = 1$) van de eenheidsmatrix I en dus geldt:

$$\boxed{I = \mathbf{u}_1\mathbf{u}_1^* + \cdots + \mathbf{u}_n\mathbf{u}_n^*}$$

Ook complex bestaat een vorm van de spektraaldecompositie stelling die als voordeel heeft dat deze wel uniek is. A is een normale matrix, $\lambda_1, \ldots, \lambda_k$ zijn de eigenwaarden van A, λ_i heeft multipliciteit n_i $(i = 1, \ldots, k)$ en $\mathcal{B}_i = \{\mathbf{u}_{(i,1)}, \ldots, \mathbf{u}_{(i,n_i)}\}$ is een orthonormale basis van de eigenruimte E_{λ_i}, voor $i = 1, \ldots, k$. Als we schrijven $U_i = [\mathbf{u}_{(i,1)} \ldots \mathbf{u}_{(i,n_i)}]$ dan weten we van Stelling 5.51 dat de projectie op de eigenruimte E_{λ_i} de standaardmatrix:

$$[\text{proj}_{E_{\lambda_i}}] = \mathbf{u}_{(i,1)}\mathbf{u}_{(i,1)}^* + \cdots + \mathbf{u}_{(i,n_i)}\mathbf{u}_{(i,n_i)}^* = U_iU_i^*$$

heeft. Als we in vorm 1 van de spektrale decompositie van A de projectiematrices bij dezelfde eigenwaarden bij elkaar nemen krijgen we:

$$A = \lambda_1[\text{proj}_{E_{\lambda_1}}] + \cdots + \lambda_k[\text{proj}_{E_{\lambda_k}}]$$

Met andere woorden: A is de lineaire combinatie van de projectiematrices (van projecties op al de onderling loodrechte eigenruimtes) met als gewichten: de bijbehorende eigenwaarden!

STELLING 5.55. *(***De Spektraaldecompositie Stelling, vorm 2***)*
Laat A een normale $n \times n$ matrix zijn met eigenwaarden $\lambda_1, \ldots, \lambda_k$. Beschouw de projecties $\mathrm{proj}_{E_{\lambda_i}} : \mathbb{C}^n \to \mathbb{C}^n$ op de eigenruimtes E_{λ_i}. Dan geldt:

$$\boxed{A = \lambda_1[\mathrm{proj}_{E_{\lambda_1}}] + \cdots + \lambda_k[\mathrm{proj}_{E_{\lambda_k}}]}$$

Het Gram-Schmidt orthogonaliseringsproces

Op dezelfde manier als in $\mathbb{R}^n$ en $\mathbb{C}^n$ kunnen we bij een gegeven eindige basis van een inwendig productruimte V om zetten naar een orthogonale basis voor W. Eerst een voorbeeld.

Voorbeeld 5.56. Voor de constructie van een orthogonale basis voor $V = \mathrm{Pol}_2([-1,1], \mathbb{R})$, voorzien van het inwendig product

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)\, dx.$$

wordt begonnen met de standaardbasis $B = \{1, x, x^2\}$ voor V. Het geluk is dat de polynomen 1 en x al orthogonaal zijn. Definieer daarom de deelruimte $W_2 = \mathrm{Span}\{1, x\}$ van V. Als aan $\{1, x\}$ een niet-nulvector uit V wordt toegevoegd die loodrecht staat op W_2, dan is een orthogonale verzameling van drie niet-nulvectoren in de drie-dimensionale V gevonden. Deze drie vectoren vormen onmiddellijk een orthogonale basis voor V. En ... deze derde vector is snel gevonden: volgens het voorgaande is $x^2 - \mathrm{proj}_{W_2}(x^2)$ namelijk zo'n vector! [Waarom behoort deze vector tot $\mathrm{Pol}_2([-1,1], \mathbb{R})$?]
Omdat $\{1, x\}$ een orthogonale basis voor W_2 is, geldt volgens stelling 5.47

$$\mathrm{proj}_{W_2}(x^2) = \frac{\langle 1, x^2 \rangle}{\langle 1, 1 \rangle} \cdot 1 + \frac{\langle x, x^2 \rangle}{\langle x, x \rangle} \cdot x.$$

Als derde basisvector kan dus

$$x^2 - \frac{\langle 1, x^2 \rangle}{\langle 1, 1 \rangle} \cdot 1 - \frac{\langle x, x^2 \rangle}{\langle x, x \rangle} \cdot x = x^2 - \frac{2/3}{2} \cdot 1 - 0 \cdot x = x^2 - \frac{1}{3}$$

genomen worden. Dus bijvoorbeeld $\{1, x, x^2 - \frac{1}{3}\}$ is een orthogonale basis voor $\mathrm{Pol}_2([-1,1], \mathbb{R})$, maar natuurlijk ook $\{1, x, 3x^2 - 1\}$ of $\{\frac{1}{2}\sqrt{2}, \frac{1}{2}x\sqrt{6}, \sqrt{10}(\frac{3}{4}x^2 - \frac{1}{4}\}$. De laatste is zelfs een orthonormale basis voor $\mathrm{Pol}_2([-1,1], \mathbb{R})$.

Dan nu het Gram-Schmidt orthogonaliseringsproces.

STELLING 5.57 (*Gram-Schmidt orthogonaliseringsproces*). *Als* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *een basis is voor een deelruimte* W *van een inwendig productruimte* V, *dan vormen de vectoren*

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 \\
\mathbf{w}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{w}_1, \mathbf{v}_2 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 \\
\mathbf{w}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{w}_1, \mathbf{v}_3 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 - \frac{\langle \mathbf{w}_2, \mathbf{v}_3 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle}\mathbf{w}_2 \\
&\vdots \\
\mathbf{w}_n &= \mathbf{v}_n - \frac{\langle \mathbf{w}_1, \mathbf{v}_n \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 - \cdots - \frac{\langle \mathbf{w}_{n-1}, \mathbf{v}_n \rangle}{\langle \mathbf{w}_{n-1}, \mathbf{w}_{n-1} \rangle}\mathbf{w}_{n-1}
\end{aligned}$$

een orthogonale basis voor W.

Bewijs: Net als in voorbeeld 5.56 wordt de orthogonale basis stapje voor stapje opgebouwd via orthogonale projecties op verschillende deelruimten van W, namelijk de deelruimten

$$W_1 = \mathrm{Span}\{\mathbf{v}_1\}, \quad W_2 = \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\}, \quad W_3 = \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}, \quad \text{etc.}$$

Merk op dat $W_n = W$. We laten zien dat voor $k = 2, \ldots, n$ het in de stelling geconstrueerde stelsel vectoren $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ een orthogonaal stelsel is dat W_k opspant. Voor het geval $k = n$ hebben we dan de geldigheid van de stelling bewezen, want de $\mathbf{w}_k$'s spannen W op en hun aantal is precies $\dim W$. Ze vormen dus een basis voor W.

We beginnen met $k = 2$. Merk op dat uit het feit dat $\mathbf{w}_1 = \mathbf{v}_1$ volgt dat $\frac{\langle \mathbf{w}_1, \mathbf{v}_2 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1$ volgens stelling 5.47 precies de loodrechte projectie $\mathrm{proj}_{W_1}(\mathbf{v}_2)$ van $\mathbf{v}_2$ op W_1 is. Daarom staat $\mathbf{v}_2 - \mathrm{proj}_{W_1}(\mathbf{v}_2)$ (en dat is precies $\mathbf{w}_2$!) loodrecht op W_1 en daarmee op $\mathbf{w}_1$. Het stelsel $\{\mathbf{w}_1, \mathbf{w}_2\}$ is dus orthogonaal. Nu moet nog aangetoond worden dat $\mathrm{Span}\{\mathbf{w}_1, \mathbf{w}_2\} = W_2$. Het is duidelijk dat $\mathbf{w}_1 = \mathbf{v}_1$ tot W_2 behoort. Verder zijn $\mathbf{v}_2$ en $\mathrm{proj}_{W_1}(\mathbf{v}_2)$ elementen van W_2 en $\mathbf{w}_2$ als lineaire combinatie van deze twee dus ook. Blijkbaar $\mathrm{Span}\{\mathbf{w}_1, \mathbf{w}_2\} \subset W_2$. Andersom is het duidelijk dat $\mathbf{v_1}$ tot $\mathrm{Span}\{\mathbf{w}_1, \mathbf{w}_2\}$ behoort en omdat $\mathbf{v}_2 = \mathbf{w}_2 + \frac{\langle \mathbf{w}_1, \mathbf{v}_2 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1$ een lineaire combinatie van $\mathbf{w}_1$ en $\mathbf{w}_2$ is, volgt dit ook voor $\mathbf{v}_2$. Blijkbaar $W_2 = \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\} \subset \mathrm{Span}\{\mathbf{w}_1, \mathbf{w}_2\}$.

Zo kunnen we doorgaan: nu we bewezen hebben dat $\{\mathbf{w}_1, \mathbf{w}_2\}$ een orthogonaal stelsel is dat W_2 opspant, kunnen we laten zien dat hetzelfde geldt voor $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ en W_3. Daaruit kunnen we dan weer laten zien dat $\{\mathbf{w}_1, \ldots, \mathbf{w}_4\}$ een orthogonaal stelsel is dat W_4 opspant, enzovoorts.

Gelukkig hoeven we niet elke stap apart op te schrijven, maar kunnen we dat in één keer doen. Als we namelijk bewezen hebben dat $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ een orthogonaal stelsel is dat W_k opspant is de vector $\mathbf{w}_{k+1}$ volgens stelling 5.47 gelijk aan de vector $\mathbf{v}_{k+1} - \mathrm{proj}_{W_k} \mathbf{v}_{k+1}$ en die staat daarom loodrecht op W_k en dus loodrecht op $\mathbf{w}_1, \ldots, \mathbf{w}_k$. Omdat de laatst genoemde vectoren al een orthogonaal stelsel vormden, is $\{\mathbf{w}_1, \ldots, \mathbf{w}_{k+1}\}$ ook een orthogonaal stelsel.

Dat $\mathrm{Span}\{\mathbf{w}_1, \ldots, \mathbf{w}_{k+1}\} = W_{k+1}$ volgt op dezelfde manier als de situatie $k = 2$. Het is namelijk duidelijk dat $\mathbf{w}_1$ tot en met $\mathbf{w}_k$ tot W_{k+1} behoren omdat $W_k \subset W_{k+1}$. Verder is $\mathbf{w}_{k+1}$ het verschil van $\mathbf{v}_{k+1}$, die in W_{k+1} ligt, en $\mathrm{proj}_{W_k} \mathbf{v}_{k+1}$, die in W_k en dus ook in W_{k+1} ligt. Dus $\mathbf{w}_{k+1}$ behoort ook tot W_{k+1}, waarmee $\mathrm{Span}\{\mathbf{w}_1, \ldots, \mathbf{w}_{k+1}\} \subset W_{k+1}$. Andersom geldt dat $\mathbf{v}_1$ tot en met $\mathbf{v}_k$ tot W_k en dus tot $\mathrm{Span}\{\mathbf{w}_1, \ldots, \mathbf{w}_{k+1}\}$ behoren. Ook is duidelijk dat $\mathbf{v}_{k+1} = \mathbf{w}_{k+1} + \frac{\langle \mathbf{w}_1, \mathbf{v}_{k+1} \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 + \cdots + \frac{\langle \mathbf{w}_k, \mathbf{v}_{k+1} \rangle}{\langle \mathbf{w}_k, \mathbf{w}_k \rangle}\mathbf{w}_k$ tot dit lineair omhulsel behoort, zodat $\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}\} = W_{k+1} \subset \mathrm{Span}\{\mathbf{w}_1, \ldots, \mathbf{w}_{k+1}\}$. Blijkbaar is $\{\mathbf{w}_1, \ldots, \mathbf{w}_{k+1}\}$ een orthogonaal stelsel dat W_{k+1} opspant! $\square$

Het volgende voorbeeld is als illustratie toegevoegd vanwege het feit dat er iets verrassends kan gebeuren. Laat je niet afleiden door het alternatieve inwendig product dat gekozen is, omdat het fenomeen dat optreedt niet veroorzaakt wordt door de keuze van het inwendig product.

Voorbeeld 5.58. De bedoeling is dat een orthogonale basis gevonden gaat worden voor de deelruimte van $\mathbb{C}^3$ opgespannen door de vectoren

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ i \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1+5i \\ 0 \\ 4i \end{bmatrix} \quad \text{en} \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 5-i \\ -1-i \end{bmatrix}$$

waarbij $\mathbb{C}^3$ voorzien is van het inwendig product

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* A \mathbf{y}$$

met

$$A = \begin{bmatrix} 1 & i & 0 \\ -i & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix}.$$

De orthogonale basis wordt uiteraard gevonden door toepassen van Gram-Schmidt en voor de eerste vector wordt dus $\mathbf{w}_1 = \mathbf{v}_1 = (1, i, 1)$ genomen. Om de tweede vector te bepalen zijn de inwendig producten

$$\langle \mathbf{w}_1, \mathbf{v}_2 \rangle = \begin{bmatrix} 1 & -i & 1 \end{bmatrix} \begin{bmatrix} 1 & i & 0 \\ -i & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1+5i \\ 0 \\ 4i \end{bmatrix} = 4 + 12i$$

en

$$\langle \mathbf{w}_1, \mathbf{w}_1 \rangle = \begin{bmatrix} 1 & -i & 1 \end{bmatrix} \begin{bmatrix} 1 & i & 0 \\ -i & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ i \\ 1 \end{bmatrix} = 4.$$

nodig. Er volgt

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{w}_1, \mathbf{v}_2 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 = \begin{bmatrix} 1+5i \\ 0 \\ 4i \end{bmatrix} - \frac{4+12i}{4} \begin{bmatrix} 1 \\ i \\ 1 \end{bmatrix} = \begin{bmatrix} 2i \\ 3-i \\ -1+i \end{bmatrix}.$$

Voor de derde vector worden eerst de inwendig producten

$$\langle \mathbf{w}_1, \mathbf{v}_3 \rangle = -8i, \quad \langle \mathbf{w}_2, \mathbf{v}_3 \rangle = 34 \quad \text{en} \quad \langle \mathbf{w}_2, \mathbf{w}_2 \rangle = 34$$

berekend, zodat

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{w}_1, \mathbf{v}_3 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{w}_2, \mathbf{v}_3 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 = \mathbf{v}_3 - \frac{-8i}{4} \mathbf{w}_1 - \frac{34}{34} \mathbf{w}_2 = \mathbf{0},$$

en dat is verrassend! Het gevonden stelsel $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ is weliswaar orthogonaal, maar het is geen basis! Hoe kan dat? Daarvoor moet je je realiseren dat $\mathbf{w}_3 = \mathbf{v}_3 - \text{proj}_{W_2}(\mathbf{v}_3)$ waarbij $W_2 = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\} = \text{Span}\{\mathbf{w}_1, \mathbf{w}_2\}$ (dit staat in het bewijs van Gram-Schmidt). Omdat $\mathbf{w}_3 = \mathbf{0}$ is $\mathbf{v}_3$ blijkbaar gelijk aan zijn orthogonale projectie op W_2. Dat betekent dat $\mathbf{v}_3$ een element van $W_2 = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ is, waaruit volgt dat het stelsel $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ dat W opspant afhankelijk is en dus zelf al geen basis was voor W. Omdat $\mathbf{v}_1$ en $\mathbf{v}_2$ wel onafhankelijk zijn vormen deze twee vectoren een basis voor W. Daaruit kan geconcludeerd worden dat $\{\mathbf{w}_1, \mathbf{w}_2\}$ een orthogonale basis voor W is. Tenslotte: $\{\frac{1}{2}\mathbf{w}_1, \frac{1}{\sqrt{34}}\mathbf{w}_2\}$ is een orthonormale basis voor W.

Een generalisatie van het geconstateerde in bovenstaand voorbeeld leidt tot het volgende: als het Gram-Schmidt proces begonnen wordt met een afhankelijk stel vectoren, dan wordt op zeker moment een nulvector $\mathbf{w}_k$ gevonden, wat betekent dat $\mathbf{v}_k$ een lineaire combinatie is van zijn voorgangers. Als we $\mathbf{v}_k$ en de nulvector $\mathbf{w}_k$ in het vervolg van het proces weglaten, kunnen we doorgaan met Gram-Schmidt en vinden we zo doorgaand vanzelf een orthogonale basis.

5.9 Maximaal Orthogonale verzamelingen

Laat V een inwendig product ruimte zijn.

Definitie: Een verzameling S van niet–nulvectoren in V heet *maximaal orthogonaal*, als deze orthogonaal is, en er is geen niet nulvector $\mathbf{v} \in V$ zodat de verzameling $S \cup \{\mathbf{v}\}$ ook orthogonaal is.

Evenzo definiëren we het begrip maximaal orthonormaal.

Opmerking: Maximaal orthogonale (orthonormale) verzamelingen worden ook wel volledige (complete) orthogonale verzamelingen genoemd. Ook worden ze wel orthogonale basis genoemd, hoewel dit een onjuiste benaming. Het volgende voorbeeld laat al zien dat in een oneindig dimensionale ruimte een maximaal orthogonale verzameling geen basis hoeft te zijn.

We zullen eerst het meest simpele voorbeeld van dit verschijnsel bestuderen.

Voorbeeld 5.59. Beschouw $V = \ell_2(\mathbb{R})$ met standaard inwendigproduct. Bekijk ook de orthonormale verzameling $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n, \ldots\}$ in V. We hebben ooit aangetoond dat $\text{Span}\{E\}^{\perp} = \{\mathbf{0}\}$. Dat betekent dat aan de verzameling E geen enkele vector $\mathbf{x} \neq \mathbf{0}$ toegevoegd kan worden zonder de orthogonaliteit te verpesten. Merk dus op dat E een maximaal orthogonale verzameling is in $\ell_2(\mathbb{R})$, maar merk ook op dat E geen basis is voor $\ell_2(\mathbb{R})$.

Voorbeeld 5.60. (Vervolg) Beschouw nogmaals $V = \ell_2(\mathbb{R})$ met de maximaal orthonormale verzameling $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n, \ldots\}$. Stel $\mathbf{x} = (x_1, x_2, \ldots, x_n, \ldots) \in \ell_2(\mathbb{R})$. Hoewel $\mathbf{x}$ niet in $\text{Span}(E)$ hoeft te zitten, kunnnen we $\mathbf{x}$ wel krijgen als een limiet van een rij van elementen uit E, en wel op de volgende manier.
Beschouw $W_n = \text{Span}\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$. W_n is eindig dimensionaal, en dus kunnen we projecteren op W_n. Beschouw

$$\mathbf{w}_n = \text{proj}_{W_n}(\mathbf{x}) = (x_1, x_2, \ldots, x_n, 0, 0, 0, \ldots)$$

We verkrijgen op deze manier een rij $\{\mathbf{w}_n\}_{n=1}^{\infty}$ waarvoor geldt:

$$\lim_{n\to\infty} \mathbf{w}_n = \mathbf{x}.$$

Hoewel dit in dit geval intuïtief duidelijk zal zijn, merken we op dat $\lim\limits_{n\to\infty} \mathbf{w}_n = \mathbf{x}$ een limiet is in de ruimte $\ell_2(\mathbb{R})$. We moeten zeggen wat we hiermee bedoelen! Hiermee bedoelen we dat voor de (reële) rij $\{\|\mathbf{w}_n - \mathbf{x}\|\}_{n=1}^{\infty}$ geldt:

$$\lim_{n\to\infty} \|\mathbf{w}_n - \mathbf{x}\| = 0$$

Inderdaad, de vectoren $\mathbf{w}_n$ in $\ell_2(\mathbb{R})$ komen op den duur op voorgeschreven afstand $\varepsilon > 0$ van $\mathbf{x}$, en dit voor iedere $\varepsilon > 0$.

Als V eindig dimensionaal is, dan is het niet moeilijk in te zien dat de maximaal orthogonale verzamelingen precies de orthogonale bases van V zijn.
Oneindig dimensionaal geldt dat niet meer. Een beroemd voorbeeld van een maximaal orthogonale verzameling is de volgende:

Voorbeeld 5.61. De verzameling

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos nx, \sin nx, \ldots\ldots$$

is een maximaal orthogonale verzameling in $C([-\pi,\pi],\mathbb{R})$ (met standaard inwendig product). Zelfs in de inwendig product ruimte $L_2([0,2\pi],\mathbb{R})$ is deze verzameling maximaal orthogonaal. Dat deze verzameling orthogonaal is, is niet moeilijk aan te tonen via de product formules voor $\cos px \cos qx$, etc. Dat deze verzameling maximaal orthogonaal is, is echt lastig en zullen we U besparen.
Na normeren verkrijgen we de volgende maximaal orthonormale verzameling:

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \ldots\ldots$$

In de Fourier analyse wordt deze maximaal orthonormale verzameling veel gebruikt, en wel op de volgende manier. Beschouw de eindig dimensionale deelruimte:

$$W_n = \mathrm{Span}\{1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos nx, \sin nx\}.$$

Neem een functie $f \in C([0,2\pi],\mathbb{R})$. Omdat W_n eindig dimensionaal is kunnen we $F_n = \mathrm{proj}_{W_n}(f)$ bekijken. Omdat de gegeven opspannende vectoren van W_n orthogonaal zijn, geldt:

$$F_n(x) = \mathrm{proj}_{W_n}(f)(x) = a_0 + \sum_{k=1}^{n} a_k \cos(kx) + \sum_{k=1}^{n} b_k \sin(kx)$$

met: $a_0 = \dfrac{\langle 1, f\rangle}{\langle 1, 1\rangle} = \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} f(x)dx,$

en: $a_k = \dfrac{\langle \cos(kx), f\rangle}{\langle \cos(kx), \cos(kx)\rangle} = \dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} f(x)\cos(kx)dx$, voor $k = 1, \ldots, n$

en: $b_k = \dfrac{\langle \sin(kx), f\rangle}{\langle \sin(kx), \sin(kx)\rangle} = \dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} f(x)\sin(kx)dx$, voor $k = 1, \ldots, n$.

De functie F_n wordt wel de n^{de} orde Fourier benadering van f genoemd. Als in het vorige voorbeeld zal weer gelden:

$$\lim_{n\to\infty} F_n = f \quad \text{wat betekent:} \quad \lim_{n\to\infty} \|F_n - f\| = 0$$

(Het zij opgemerkt dat dit NIET hoeft te betekenen dat $\lim_{n\to\infty} F_n(x) = f(x)$, voor alle $x \in [-\pi,\pi]$.)
Het is wel gebruikelijk om de Fourierreeks $F(x)$ van de functie f

$$F(x) = \lim_{n\to\infty} F_n(x) = a_0 + \sum_{k=1}^{\infty} a_k \cos(kx) + \sum_{k=1}^{n} b_k \sin(kx)$$

te beschouwen.

Voorbeeld 5.62. De verzameling

$$1, \cos x, \cos 2x, \ldots, \cos nx, \ldots\ldots$$

is een maximaal orthogonale verzameling in $C([0,\pi],\mathbb{R})$ (met standaard inwendig product). Zelfs in de inwendig product ruimte $L_2([0,\pi],\mathbb{R})$ is deze verzameling maximaal orthogonaal. Dit is vrij makkelijk te bewijzen als de familie uit Voorbeeld 5.61 verondersteld wordt maximaal orthogonaal te zijn.
Na normeren verkrijgen we de volgende maximaal orthonormale verzameling:

$$\frac{1}{\sqrt{\pi}}, \frac{\sqrt{2}\cos x}{\sqrt{\pi}}, \frac{\sqrt{2}\cos 2x}{\sqrt{\pi}}, \ldots, \frac{\sqrt{2}\cos nx}{\sqrt{\pi}}, \ldots\ldots$$

Zoals beschreven in de vorige voorbeelden kan men een gegeven functie $f \in C([0,\pi],\mathbb{R})$ projecteren op $W_n = \text{Span}\{1, \cos x, \cos 2x, \ldots, \cos nx\}$. De functie $C_n = \text{proj}_{W_n}(f)$ heet de n^{de} Fourier cosinus benadering van f en in de ruimte $C([0,\pi],\mathbb{R})$ geldt weer: $\lim_{n\to\infty} C_n = f$ (dus $\lim_{n\to\infty} \|C_n - f\| = 0$).
Merk op:

$$C_n(x) = a_0 + \sum_{k=1}^{n} a_k \cos(kx) \quad \text{met} \quad a_0 = \frac{\langle 1, f\rangle}{\langle 1, 1\rangle} = \frac{1}{\pi}\int_0^\pi f(x)dx$$

$$\text{en} \quad a_k = \frac{\langle \cos(kx), f\rangle}{\langle \cos(kx), \cos(kx)\rangle} = \frac{1}{2\pi}\int_0^\pi f(x)\cos(kx)dx$$

De reeks $C(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(nx)$ heet de cosinusreeks van f op het interval $[0,\pi]$

Voorbeeld 5.63. De verzameling

$$1, \sin x, \sin 2x, \ldots, \sin nx, \ldots\ldots$$

is een maximaal orthogonale verzameling in $C([0,\pi],\mathbb{R})$. Zelfs in de inwendig product ruimte $L_2([0,\pi],\mathbb{R})$ is deze verzameling maximaal orthogonaal. Dit is vrij makkelijk te bewijzen als de familie uit Voorbeeld 5.61 verondersteld wordt maximaal orthogonaal te zijn.
Na normeren verkrijgen we de maximaal orthonormale verzameling in $C([0,\pi],\mathbb{R})$:

$$\frac{1}{\sqrt{\pi}}, \frac{\sqrt{2}\sin x}{\sqrt{\pi}}, \frac{\sqrt{2}\sin 2x}{\sqrt{\pi}}, \ldots, \frac{\sqrt{2}\sin nx}{\sqrt{\pi}}, \ldots\ldots$$

Als in het vorige voorbeeld kan de n^{de} Fourier sinus benadering S_n van $f \in C([0,\pi],\mathbb{R})$ geïntroduceerd worden en de sinusreeks.

Voorbeeld 5.64. De verzameling

$$1, e^{ix}, e^{-ix}, e^{2ix}, e^{-2ix}, \ldots, e^{inx}, e^{-inx}, \ldots\ldots$$

is een maximaal orthogonale verzameling in $C([-\pi,\pi],\mathbb{C})$. Zelfs in de inwendig product ruimte $L_2([-\pi,\pi],\mathbb{C})$ is deze verzameling maximaal orthogonaal. Dit is vrij makkelijk te bewijzen als de familie uit Voorbeeld 5.61 verondersteld wordt maximaal orthogonaal te zijn.
Na normeren verkrijgen we de maximaal orthonormale verzameling in $C([-\pi,\pi],\mathbb{C})$:

$$\frac{1}{\sqrt{2\pi}}, \frac{e^{ix}}{\sqrt{2\pi}}, \frac{e^{-ix}}{\sqrt{2\pi}}, \frac{e^{2ix}}{\sqrt{2\pi}}, \frac{e^{-2ix}}{\sqrt{2\pi}}, \ldots, \frac{e^{nix}}{\sqrt{2\pi}}, \frac{e^{-nix}}{\sqrt{2\pi}}, \ldots\ldots$$

Als in voorbeeld 5.61 kan de n^{de} exponentiëele benadering van $f \in C([-\pi,\pi],\mathbb{C})$ geïntroduceerd worden en de exponentiëele reeks.

Hoofdstuk 6

De geadjungeerde van een lineaire operator

6.1 De geadjungeerde van een lineaire afbeelding

Op eindig dimensionale ruimten $\mathbb{R}^n$ (en $\mathbb{C}^n$) hebben we aan de ene kant $m \times n$ matrices bekeken. Ook de lineaire afbeeldingen $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$ zijn aan bod gekomen. De begrippen zijn aan elkaar gekoppeld via de standaardmatrix. Er zijn wiskundigen die zeggen: “matrices, daar draait het om”maar er is ook een groep die zegt: “nee hoor, lineaire afbeeldingen, die moeten centraal staan”.
Merk op dat voor de oneindig dimensionale lineaire algebra er eigenlijk geen keus is. Lineaire afbeeldingen, daar draait het om. Ook voor het oneindig dimensionale geval willen we alle theorie die voor de matrices voor hande is op de lineaire afbeeldingen kunnen loslaten.
Veel begrippen zijn zowel voor vierkante matrices als voor operatoren bekeken, zoals: eigenwaarden, eigenvectoren en diagonaliseren voor operatoren besproken.
Maar andere begrippen, zoals symmetrisch, Hermitisch, orthogonaal, unitair, normaal $\cdots$ hebben we alleen voor matrices, maar in ieder geval in het eindig dimensionale geval zouden die ook voor de afbeeldingen moeten bestaan.
Er is één operatie die we voor matrices kennen, maar niet voor afbeeldingen. Dat is het (Hermitisch) transponeren. Een $m \times n$ matrix A induceert een afbeelding $L_A : \mathbb{C}^n \to \mathbb{C}^m$ (via $L_A(\mathbf{x}) = A\mathbf{x}$) terwijl de $n \times m$ matrix A^* een afbeelding $L_{A^*} : \mathbb{C}^m \to \mathbb{C}^n$ oplevert (via $L_{A^*}(\mathbf{y}) = A^*\mathbf{y}$).
Probleem: hoe de afbeelding L_{A^*} uit de afbeelding L_A te herkennen?
Welnu, in college 4 zijn veel eigenschappen uit de volgende eigenschap verkregen. Voor elke tweetal vectoren $\mathbf{x} \in \mathbb{C}^n$ en $\mathbf{y} \in \mathbb{C}^m$ geldt namelijk

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^*\mathbf{y} \rangle$$

Dit volgt uit het feit dat het standaard inwendig product op $\mathbb{C}^n$ als matrixproduct geschreven kan worden via $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\mathbf{u}}^T\mathbf{v} = \mathbf{u}^*\mathbf{v}$. Hieruit en uit de eigenschappen voor Hermitisch transponeren blijkt dat

$$\langle A\mathbf{x}, \mathbf{y} \rangle = (A\mathbf{x})^*\mathbf{y} = (\mathbf{x}^*A^*)\mathbf{y} = \mathbf{x}^*(A^*\mathbf{y}) = \langle \mathbf{x}, A^*\mathbf{y} \rangle \, .$$

Maar hebben we een relatie tussen L_A en L_{A^*} verkregen: voor elke tweetal vectoren $\mathbf{x} \in \mathbb{C}^n$ en $\mathbf{y} \in \mathbb{C}^m$ geldt namelijk

$$\langle L_A(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, L_{A^*}(\mathbf{y}) \rangle$$

Dit alles geeft aanleiding tot de volgende definitie.

Definitie: Zij $\mathcal{L} : V \to W$ een lineaire afbeelding tussen de inwendig productruimtes V en W. Een lineaire afbeelding $\mathcal{L}^* : W \to V$ met de eigenschap

$$\langle \mathcal{L}(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, \mathcal{L}^*(\mathbf{w}) \rangle \text{ voor alle } \mathbf{v} \in V \text{ en } \mathbf{w} \in W, \tag{6.1}$$

wordt de (een?) *geadjungeerde afbeelding* $\mathcal{L}^*$ van $\mathcal{L}$ genoemd.

Opmerking: (i) Merk op dat $\langle \mathcal{L}(\mathbf{v}), \mathbf{w} \rangle$ het inwendig product in W betreft, terwijl $\langle \mathbf{v}, \mathcal{L}^*(\mathbf{w}) \rangle$ het inwendig product in V betreft. We noteren dit soms met:

$$\langle \mathcal{L}(\mathbf{v}), \mathbf{w} \rangle_W = \langle \mathbf{v}, \mathcal{L}^*(\mathbf{w}) \rangle_V .$$

(ii) We kunnen inderdaad spreken over *de* geadjungeerde afbeelding, want als deze bestaat, is hij uniek. Stel namelijk dat zowel $\mathcal{L}_1$ als $\mathcal{L}_2$ een geadjungeerde van $\mathcal{L}$ is. Neem een vector $\mathbf{w} \in W$. Volgens vergelijking (6.2) moet dan voor alle $\mathbf{v}$ uit V gelden:

$$\langle \mathcal{L}(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, \mathcal{L}_1(\mathbf{w}) \rangle \quad \text{en} \quad \langle \mathcal{L}(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, \mathcal{L}_2(\mathbf{w}) \rangle$$

en daarom zal voor alle $\mathbf{v}$ uit V gelden:

$$\langle \mathbf{v}, \mathcal{L}_1(\mathbf{w}) - \mathcal{L}_2(\mathbf{w}) \rangle = \langle \mathbf{v}, \mathcal{L}_1(\mathbf{w}) \rangle - \langle \mathbf{v}, \mathcal{L}_2(\mathbf{w}) \rangle = \langle \mathcal{L}(\mathbf{v}), \mathbf{w} \rangle - \langle \mathcal{L}(\mathbf{v}), \mathbf{w} \rangle = 0.$$

Dit kan alleen als $\mathcal{L}_1(\mathbf{w}) - \mathcal{L}_2(\mathbf{w}) = \mathbf{0}$ (neem maar $\mathbf{v} = \mathcal{L}_1(\mathbf{w}) - \mathcal{L}_2(\mathbf{w})$) waaruit volgt dat $\mathcal{L}_1(\mathbf{w}) = \mathcal{L}_2(\mathbf{w})$, en wel voor alle $\mathbf{w}$ uit W. Blijkbaar zijn $\mathcal{L}_1$ en $\mathcal{L}_2$ gelijke afbeeldingen.
(iii) Wat niet duidelijk is of iedere lineaire afbeelding $\mathcal{L} : V \to W$ ook inderdaad een geadjungeerde heeft. En inderdaad, dit hoeft niet zo te zijn in het oneindig dimensionale geval. Zelfs niet als $W = V$ (dus als $\mathcal{L}$ een operator is). We zullen hiervan geen voorbeeld geven. In het algemeen geldt dat "mooi gedefinieerde operatoren een geadjungeerde hebben".
(iv) Merk op dat de geadjungeerde afbeelding van het inwendig product afhangt. Als op de inwendig product ruimte V of W het inwendig product veranderd wordt, dan verandert ook de geadjungeerde afbeelding $\mathcal{L}^*$ van een gegeven lineaire afbeelding $\mathcal{L} : V \to W$.
(v) Merk op dat dat de uitdrukking $\langle \mathcal{L}(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, \mathcal{L}^*(\mathbf{w}) \rangle$ (voor alle $\mathbf{v} \in V$ en $\mathbf{w} \in W$) via de rekenregels van het inwendig product (eerst $\overline{\langle \mathcal{L}(\mathbf{v}), \mathbf{w} \rangle} = \overline{\langle \mathbf{v}, \mathcal{L}^*(\mathbf{w}) \rangle}$) geeft dat: $\langle \mathbf{w}, \mathcal{L}(\mathbf{v}) \rangle = \langle \mathcal{L}^*(\mathbf{w}), \mathbf{v} \rangle$, voor alle $\mathbf{v} \in V$ en $\mathbf{w} \in W$.
We zeggen:

$\mathcal{L}$ mag door het inwendig product gehaald worden en wordt $\mathcal{L}^*$.

STELLING 6.1. *Stel dat V, W een inwendig product ruimtes zijn. Als $\mathcal{L} : V \to W$ een geadjungeerde heeft dan:*

1. *heeft $\mathcal{L}^* : W \to V$ dat ook en $(\mathcal{L}^*)^* = \mathcal{L}$.*
2. *heeft $c\mathcal{L} : V \to W$ dat ook en $(c\mathcal{L})^* = \bar{c}\mathcal{L}^*$.*

Bewijs: De lineaire afbeelding $\mathcal{L} : V \to W$ heeft een geadjungeerde, d.w.z. er is een (unieke) lineaire afbeelding $\mathcal{L}^* : W \to V$ met

$$\langle \mathcal{L}(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, \mathcal{L}^*(\mathbf{w}) \rangle \quad \text{voor alle } \mathbf{v} \in V \text{ en } \mathbf{w} \in W.$$

1. Beschouw nu $\mathcal{L}^* : W \to V$. Omdat bovenstaande identiteit geeft dat:

$$\langle \mathcal{L}^*(\mathbf{w}), \mathbf{v} \rangle = \langle \mathbf{w}, \mathcal{L}(\mathbf{v}) \rangle \quad \text{voor alle } \mathbf{w} \in W \text{ en } \mathbf{v} \in V.$$

bestaat er kennelijk een afbeelding, namelijk $\mathcal{L}$, die voldoet aan de eis van de geadjungeerde van $\mathcal{L}^*$. Dus: $\mathcal{L} = \mathcal{L}^{**}$.
2. Omdat $\langle (c\mathcal{L})(\mathbf{v}), \mathbf{w} \rangle = \langle c(\mathcal{L}(\mathbf{v})), \mathbf{w} \rangle = \overline{c} \langle \mathcal{L}(\mathbf{v}), \mathbf{w} \rangle = \overline{c} \langle \mathbf{v}, \mathcal{L}^*(\mathbf{w}) \rangle = \langle \mathbf{v}, \overline{c}(\mathcal{L}^*(\mathbf{w})) \rangle = \langle \mathbf{v}, (\overline{c}\mathcal{L}^*)(\mathbf{w}) \rangle$, voor alle $\mathbf{v} \in V$ en $\mathbf{w} \in W$, volgt het gestelde. □

STELLING 6.2. *Stel dat V, W, Z een inwendig product ruimtes zijn.*

1. *Als de afbeeldingen $\mathcal{L}_i : V \to W$ $(i = 1, 2)$ geadjungeerde hebben, dan heeft de afbeelding $\mathcal{L}_1 + \mathcal{L}_2 : V \to W$ dat ook en $(\mathcal{L}_1 + \mathcal{L}_2)^* = (\mathcal{L}_1)^* + (\mathcal{L}_2)^*$.*
2. *Als de afbeeldingen $\mathcal{L}_1 : V \to W$ en $\mathcal{L}_2 : W \to Z$ geadjungeerde hebben, dan heeft de afbeelding $\mathcal{L}_2 \circ \mathcal{L}_1 : V \to Z$ dat ook en $(\mathcal{L}_2 \circ \mathcal{L}_1)^* = (\mathcal{L}_1)^* \circ (\mathcal{L}_2)^*$.*
3. $\mathrm{Id} : V \to V$ *heeft een geadjungeerde en* $\mathrm{Id}^* = \mathrm{Id}$.

U mag de bewijzen zelf leveren in de opgaven.
Tenslotte zullen we nog geadjungeerden van structureel simpele afbeeldingen bepalen.

Voorbeeld 6.3. Laat V een rëel inwendig product ruimte zijn. We bekijken lineaire afbeeldingen $L : \mathbb{R} \to V$. Merk op dat $\mathbb{R}$ een inwendig product ruimte is met $\langle r, s \rangle_{\mathbb{R}} = rs$.
Fixeer $\mathbf{v} \in V$ en definieer de afbeelding $\varphi_{\mathbf{v}} : \mathbb{R} \to V$ bekijken gedefiniëerd door: $\varphi_{\mathbf{v}}(r) = r\mathbf{v}$. Merk op dat de afbeelding $\varphi_{\mathbf{v}} : \mathbb{R} \to V$ inderdaad een lineaire afbeelding is. (Ga na!)
We willen de geadjungeerde afbeelding bepalen. Kennelijk heeft deze geadjungeerde afbeelding $\varphi_{\mathbf{v}}^* : V \to \mathbb{R}$ de eigenschap:

$$\langle \varphi_{\mathbf{v}}(r), \mathbf{u} \rangle_V = \langle r, \varphi_{\mathbf{v}}^*(\mathbf{u}) \rangle_{\mathbb{R}}, \quad \text{for all } \; r \in \mathbb{R}, \; \mathbf{u} \in V$$

Hieruit kunnen we de formule van $\varphi_{\mathbf{v}}^*(\mathbf{u})$ halen.
Links lezen we: $\langle \varphi_{\mathbf{v}}(r), \mathbf{u} \rangle = \langle r\mathbf{v}, \mathbf{u} \rangle = r \langle \mathbf{v}, \mathbf{u} \rangle$.
Rechts lezen we: $\langle r, \varphi_{\mathbf{v}}^*(\mathbf{u}) \rangle_{\mathbb{R}} = r\varphi_{\mathbf{v}}^*(\mathbf{u})$, dus kennelijk geldt:

$$r \langle \mathbf{v}, \mathbf{u} \rangle = r\varphi_{\mathbf{v}}^*(\mathbf{u}), \quad \text{for all } \; r \in \mathbb{R}, \; \mathbf{u} \in V$$

en dus:

$$\varphi_{\mathbf{v}}^*(\mathbf{u}) = \langle \mathbf{v}, \mathbf{u} \rangle$$

Herinner: als $\mathbf{q} \in \mathbb{R}^n$ met $\|\mathbf{q}\| = 1$ dan is $\mathbf{q}\mathbf{q}^T$ de projectiematrix van de projectie op $\mathrm{Span}\{\mathbf{q}\}$ d.w.z.

$$\mathrm{proj}_{\mathrm{Span}\{\mathbf{q}\}}(\mathbf{x}) = \mathbf{q}\mathbf{q}^T(\mathbf{x})$$

(Dus: $\mathbf{q}\mathbf{q}^T$ is de standaardmatrix van de projectie op $\mathrm{Span}\{\mathbf{q}\}$. Een zelfde soort uitspraak hebben we nu gekregen in een oneindig dimensionaal inwendig product ruimte!

STELLING 6.4. *Laat V een rëel inwendig product ruimte zijn. Als $\mathbf{q} \in V$ met $\|\mathbf{q}\| = 1$ dan geldt voor $W = \mathrm{Span}\{\mathbf{q}\}$ dat:*

$$\mathrm{proj}_W = \varphi_{\mathbf{q}} \circ \varphi_{\mathbf{q}}^*$$

Bewijs: We kennen de formule voor de projectie op W, omdat $\{\mathbf{q}\}$ een orthonormale basis is voor W. We concluderen:

$$\mathrm{proj}_W(\mathbf{x}) = \left(\frac{\langle \mathbf{q}, \mathbf{x}\rangle}{\langle \mathbf{q}, \mathbf{q}\rangle}\right)\mathbf{q} = \left(\frac{\langle \mathbf{q}, \mathbf{x}\rangle}{\|\mathbf{q}\|^2}\right)\mathbf{q} = (\langle \mathbf{q}, \mathbf{x}\rangle)\mathbf{q}$$

Aan de andere kant: we kunnen $\varphi_{\mathbf{q}} \circ \varphi^*_{\mathbf{q}}(\mathbf{x})$ uitwerken met de formules uit Voorbeeld 14.15. Dit zijn: $\varphi_{\mathbf{q}}(r) = r\mathbf{q}$ en $\varphi^*_{\mathbf{q}}(\mathbf{x}) = \langle \mathbf{q}, \mathbf{x}\rangle$. Dan verkrijgen we:

$$(\varphi_{\mathbf{q}} \circ \varphi^*_{\mathbf{q}})(\mathbf{x}) = \varphi_{\mathbf{q}}(\varphi^*_{\mathbf{q}}(\mathbf{x}) = \varphi_{\mathbf{q}}(\langle \mathbf{q}, \mathbf{x}\rangle) = (\langle \mathbf{q}, \mathbf{x}\rangle)\mathbf{q} = \mathrm{proj}_W(\mathbf{x})$$

De conclusie volgt. □

Tenslotte bepalen we de geadjungeerde van de projectie proj_W zelf.

STELLING 6.5. *Laat V een (rëel of complex) inwendig product ruimte zijn. Als W een eindig dimensionale deelruimte is en de projectie* proj_W *wordt beschouwd als* $\mathrm{proj}_W : V \to V$ *(dus niet als* $\mathrm{proj}_W : V \to W$*), dan geldt:*

$$\mathrm{proj}_W{}^* = \mathrm{proj}_W$$

Bewijs: Het is voldoende aan te tonen:

$$\langle \mathrm{proj}_W(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, \mathrm{proj}_W(\mathbf{y})\rangle\,, \quad \text{voor alle } \ \mathbf{x}, \mathbf{y} \in V.$$

Om dit in te zien schrijven we: $\mathbf{x} = \mathbf{w}_1 + \mathbf{w}_1^\sharp$ en $\mathbf{y} = \mathbf{w}_2 + \mathbf{w}_2^\sharp$ met $\mathbf{w}_i \in W$ en $\mathbf{w}_i^\sharp \in W^\perp$. We zien:
$\langle \mathrm{proj}_W(\mathbf{x}), \mathbf{y}\rangle = \left\langle \mathbf{w}_1, \mathbf{w}_2 + \mathbf{w}_2^\sharp \right\rangle = \langle \mathbf{w}_1, \mathbf{w}_2\rangle + \left\langle \mathbf{w}_1, \mathbf{w}_2^\sharp \right\rangle = \langle \mathbf{w}_1, \mathbf{w}_2\rangle$.
Evenzo: $\langle \mathbf{x}, \mathrm{proj}_W(\mathbf{y})\rangle = \left\langle \mathbf{w}_1 + \mathbf{w}_1^\sharp, \mathbf{w}_2 \right\rangle = \langle \mathbf{w}_1, \mathbf{w}_2\rangle$. De conclusie volgt. □

6.2 De geadjungeerde van een lineaire operator

De algemene definitie van de geadjungeerde van een afbeelding $\mathcal{L} : V \to W$ is gepresenteerd. Wij zijn echter het meest geïnteresseerd in de situatie $V = W$, dus als $\mathcal{L} : V \to V$ een operator is. We herhalen de definitie voor dit speciale geval.

Definitie: Stel $\mathcal{L} : V \to V$ een lineaire operator op de inwendig productruimtes V. Een lineaire operator $\mathcal{L}^* : V \to V$ met de eigenschap

$$\langle \mathcal{L}(\mathbf{v}), \mathbf{u}\rangle = \langle \mathbf{v}, \mathcal{L}^*(\mathbf{u})\rangle \ \text{ voor alle } \mathbf{v}, \mathbf{u} \in V \tag{6.2}$$

wordt de *geadjungeerde operator* $\mathcal{L}^*$ van $\mathcal{L}$ genoemd.

Als $\mathcal{L} : V \to V$ een geadjungeerde operator heeft, dan loopt de geadjungeerde operator $\mathcal{L}^*$ ook van V naar V. Toch is het soms goed voor het inzicht dat men zich realiseert dat $\mathcal{L}^* : V \to V$ de omgekeerde richting heeft van $\mathcal{L} : V \to V$. We geven enkele voorbeelden.

Voorbeeld 6.6. In college 14 hebben we gezien dat de geadjungeerde $\mathcal{L}^*$ van de operator $\mathcal{L}(\mathbf{x}) = A\mathbf{x}$ op $\mathbb{C}^n$ met standaard inwendig product bestaat. Er geldt $\mathcal{L}^*(\mathbf{x}) = A^*\mathbf{x}$. Hierbij is A een (eventueel complexe) $n \times n$ matrix.

De geadjungeerde van de afbeelding $\mathcal{L}$ op $\mathbb{R}^n$ (met standaard inwendig product) gedefinieerd door $\mathcal{L}(\mathbf{x}) = A\mathbf{x}$ is de afbeelding $\mathcal{L}^*$ op $\mathbb{R}^n$ met $\mathcal{L}^*(\mathbf{x}) = A^T\mathbf{x}$. □

Voorbeeld 6.7. Bekijk de operator $\mathcal{L}$ op Beschouw de vectorruimte $C([-1,1],\mathbb{C})$ voorzien van het standaard inwendig product $\langle f, g\rangle = \int_{-1}^{1} \overline{f(t)}g(t)\, dt$. Bekijk de operator $\mathcal{L}$ op $C([-1,1],\mathbb{C})$ die gedefinieerd wordt door

$$\mathcal{L}(f)(x) = \int_{-1}^{1} e^{x-t} f(t)\, dt = e^x \int_{-1}^{1} e^{-t} f(t)\, dt.$$

Dus $\mathcal{L}$ voegt aan iedere functie f een veelvoud van de exponentiële functie e^x toe. We laten zien dat de operator $\mathcal{L}^*$ gedefinieerd door

$$\mathcal{L}^*(f)(x) = \int_{-1}^{1} e^{t-x} f(t)\, dt = e^{-x} \int_{-1}^{1} e^{t} f(t)\, dt.$$

de geadjungeerde van $\mathcal{L}$ is. Dit doen we door $\langle \mathcal{L}(f), g\rangle$ en $\langle f, \mathcal{L}^*(g)\rangle$ te berekenen en vervolgens te beredeneren dat ze gelijk zijn.

$$\begin{aligned}
\langle \mathcal{L}(f), g\rangle &= \int_{-1}^{1} \overline{\mathcal{L}(f)(x)}g(x)\, dx = \int_{-1}^{1} \overline{\left(e^x \int_{-1}^{1} e^{-t} f(t)\, dt\right)}\, g(x)\, dx = \\
&= \int_{-1}^{1} e^x \overline{\left(\int_{-1}^{1} e^{-t} f(t)\, dt\right)}\, g(x)\, dx = \overline{\left(\int_{-1}^{1} e^{-t} f(t)\, dt\right)} \int_{-1}^{1} e^x g(x)\, dx = \\
&= \int_{-1}^{1} e^{-t}\overline{f(t)}\, dt \int_{-1}^{1} e^x g(x)\, dx.
\end{aligned}$$

en

$$\begin{aligned}
\langle f, \mathcal{L}^*(g)\rangle &= \int_{-1}^{1} \overline{f(x)}\mathcal{L}^*(g)(x)\, dx = \int_{-1}^{1} \overline{f(x)}\; e^{-x} \left(\int_{-1}^{1} e^t g(t) dt\right)\, dx \\
&= \int_{-1}^{1} e^t g(t) dt \int_{-1}^{1} e^{-x}\overline{f(x)}\, dx
\end{aligned}$$

Het is wellicht niet duidelijk dat deze integralen hetzelfde opleveren. Maar als we in de eerste integraal de variabele x de naam t en de variabele t de naam x geven, dan verschijnt de tweede intergraal (met omgekeerde integratie volgorde). En dus hebben beide integralen dezelfde waarde. Dus volgt dat $\langle \mathcal{L}(f), g\rangle = \langle f, \mathcal{L}^*(g)\rangle$, hetgeen we wilden aantonen. □

In bovenstaand voorbeeld bestaat de geadjungeerde van de gegeven operator. Het zal iedereen duidelijk zijn dat het (bij dit voorbeeld) verdraaid lastig zal zijn om de geadjungeerde operator zelf te verzinnen (in het voorbeeld werd de formule van de geadjungeerde gegeven). Met name als je niet van te voren zou weten of deze wel bestaat. De vraag is natuurlijk: bestaan ze altijd en als de geadjungeerde bestaat, hoe vind je die dan?

Eindig dimensionaal zal de geadjungeerde altijd bestaan (zie stelling 6.9) en eigenlijk geeft deze stelling ook een methode hoe de geadjungeerde dan te vinden.

In de opgaven vindt u een voorbeeld van een operator op een oneindig-dimensionale inwendig productruimte die geen geadjungeerde heeft. Ook bestaat in het oneindig-dimensionale geval geen standaardprocedure die de geadjungeerde oplevert. Je bent dan aangewezen op de definitie en op je puzzelend vermogen. We geven een voorbeeld.

Voorbeeld 6.8. Laat V een complex inwendig productruimte zijn. Fixeer $\mathbf{a} \in V$. Definieer de operator $\mathcal{L} : V \to V$ via

$$\mathcal{L}(\mathbf{x}) = \langle \mathbf{a}, \mathbf{x} \rangle \, \mathbf{a}$$

Bepaal een formule voor de geadjungeerde operator $\mathcal{L}^*$. (U mag zelf nagaan dat de afbeelding lineair is.) Er moet gelden:

$$\langle \mathcal{L}(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, \mathcal{L}^*(\mathbf{y}) \rangle$$

De eerste kunnen we bepalen. $\langle \mathcal{L}(\mathbf{x}), \mathbf{y} \rangle = \langle \langle \mathbf{a}, \mathbf{x} \rangle \, \mathbf{a}, \mathbf{y} \rangle$ en kennelijk moeten we dit omschrijven naar $\langle \mathbf{x}, ?? \rangle$, waarbij ?? gelijk is aan $\mathcal{L}^*(\mathbf{y})$. Welnu:

$$\langle \mathcal{L}(\mathbf{x}), \mathbf{y} \rangle = \langle \langle \mathbf{a}, \mathbf{x} \rangle \, \mathbf{a}, \mathbf{y} \rangle = \overline{\langle \mathbf{a}, \mathbf{x} \rangle} \, \langle \mathbf{a}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{a} \rangle \, \langle \mathbf{a}, \mathbf{y} \rangle = \langle \mathbf{x}, \langle \mathbf{a}, \mathbf{y} \rangle \, \mathbf{a} \rangle$$

Kennelijk $\mathcal{L}^*(\mathbf{y}) = \langle \mathbf{a}, \mathbf{y} \rangle \, \mathbf{a}$. Opmerkelijk, kennelijk geldt: $\mathcal{L}$ en $\mathcal{L}^*$ zijn dezelfde afbeelding! $\square$

Dan nu de stelling voor het eindig-dimensionale geval:

STELLING 6.9. *De geadjungeerde van een lineaire operator $\mathcal{L}$ op een eindig-dimensionale inwendig productruimte V over $\mathbb{L}$ bestaat (en is uniek).*
Als voor V een orthonormale basis $\mathcal{B}$ gekozen wordt en als A de matrix van $\mathcal{L}$ ten opzichte van deze basis is, dan geldt dat A^* de matrix van $\mathcal{L}^*$ is ten opzichte van $\mathcal{B}$. Met andere woorden:

$$\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}^*]} = \underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}^{\,*}.$$

mits de basis $\mathcal{B}$ orthonormaal is.

Bewijs: Definieer $\mathcal{L}^*$ via coördinatisering ten opzichte van $\mathcal{B}$ door $[\mathcal{L}^*(\mathbf{x})]_{\mathcal{B}} = A^*[\mathbf{x}]_{\mathcal{B}}$. We zullen laten zien dat deze operator op V de geadjungeerde van $\mathcal{L}$ is. Daarmee is dan aangetoond dat deze bestaat (uniciteit was al bekend).
Omdat $\mathcal{B}$ een orthonormale basis is, hebben we $\langle \mathbf{v}, \mathbf{w} \rangle = ([\mathbf{v}]_{\mathcal{B}})^*[\mathbf{w}]_{\mathcal{B}}$ (zie stelling 5.37). Dus voor alle $\mathbf{x}$ en $\mathbf{y}$ uit V geldt dat

$$\langle \mathcal{L}(\mathbf{x}), \mathbf{y} \rangle = ([\mathcal{L}(\mathbf{x})]_{\mathcal{B}})^*[\mathbf{y}]_B = (A[\mathbf{x}]_{\mathcal{B}})^*[\mathbf{y}]_{\mathcal{B}} = ([\mathbf{x}]_{\mathcal{B}})^*(A^*[\mathbf{y}]_{\mathcal{B}}) = ([\mathbf{x}]_{\mathcal{B}})^*[\mathcal{L}^*(\mathbf{y})]_{\mathcal{B}} = \langle \mathbf{x}, \mathcal{L}^*(\mathbf{y}) \rangle$$

waarmee het gestelde volgt. $\square$

Opmerking: Als gewerkt wordt op een reële vectorruimte komt hermitisch transponeren overeen met transponeren. Dus kan in bovenstaande de "*" overal vervangen worden door "T". Behalve dan bij $\mathcal{L}^*$. Ook in het reële geval wordt de geadjungeerde van een operator $\mathcal{L}$, $\mathcal{L}^*$ genoemd en dus niet $\mathcal{L}^T$.

De stelling zegt niet alleen dat de geadjungeerde van operatoren op eindig-dimensionale vectorruimten altijd bestaat, maar geeft zelfs een methode om de geadjungeerde in die situatie te bepalen. Deze methode werkt als je een orthonormale basis voor de vectorruimte hebt. In het volgende voorbeeld laten we zien hoe dat gaat.

Voorbeeld 6.10. Beschouw $\mathbb{R}^2$ met het standaardinwendig product en beschouw de operator $\mathcal{L} : \mathbb{R}^2 \to \mathbb{R}^2$ gegeven door $\mathcal{L}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$. We bepalen $\mathcal{L}^* : \mathbb{R}^2 \to \mathbb{R}^2$.

Beschouw $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$, de standaardbasis. Kennelijk geldt dat $\underset{\mathcal{E}\leftarrow\mathcal{E}}{[\mathcal{L}]} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Omdat $\mathcal{E} =$

$\{\mathbf{e}_1, \mathbf{e}_2\}$ een orthonormale basis is, geldt: $\underset{\mathcal{E}\leftarrow\mathcal{E}}{[\mathcal{L}^*]} = \left(\underset{\mathcal{E}\leftarrow\mathcal{E}}{[\mathcal{L}]}\right)^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. En dus: $\mathcal{L}^*\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$. U ziet: $\mathcal{L}^* = \mathcal{L}$. $\square$

Voorbeeld 6.11. Voorzie nu $\mathbb{R}^2$ van het inwendig product $\langle \mathbf{x}, \mathbf{y}\rangle = x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2$. Bekijk weer de operator $\mathcal{L}$ op $\mathbb{R}^2$ gegeven door $\mathcal{L}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$. We bepalen opnieuw $\mathcal{L}^* : \mathbb{R}^2 \to \mathbb{R}^2$ (en zullen inzien dat inderdaad $\mathcal{L}^*$ afhangt van het i.p.)
Er geldt nog steeds dat $\underset{\mathcal{E}\leftarrow\mathcal{E}}{[\mathcal{L}]} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, maar nu is $\mathcal{E}$ *geen* orthonormale basis. Om de geadjungeerde afbeelding te vinden, voeren we de volgende stappen uit:
1. Bepaal een orthonormale basis: via Gram-Schmidt vinden we, vanuit de standaardbasis $\mathcal{E}$, de orthonormale basis (ga na!):

$$\mathcal{B} = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}\right).$$

2. Bepaal ten opzichte van $\mathcal{B}$ de matrix van $\mathcal{L}$:

$$\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]} = \underset{\mathcal{B}\leftarrow\mathcal{E}}{P}\,\underset{\mathcal{E}\leftarrow\mathcal{E}}{[\mathcal{L}]}\,\underset{\mathcal{E}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}.$$

3. Bepaal de hermitisch getransponeerde: dat is dan de matrix van $\mathcal{L}^*$ ten opzichte van $\mathcal{B}$. Dus:

$$\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}^*]} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}.$$

4. Ten opzichte van de standaardbasis is de matrix van $\mathcal{L}^*$ dan

$$\underset{\mathcal{E}\leftarrow\mathcal{E}}{[\mathcal{L}^*]} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{P}\,\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}^*]}\,\underset{\mathcal{B}\leftarrow\mathcal{E}}{P} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 3 \\ 0 & -1 \end{bmatrix}.$$

Er geldt dus $\mathcal{L}^*\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 3x_2 \\ -x_2 \end{bmatrix}$. $\square$

Er is nog een tweede methode om de geadjungeerde van een operator te bepalen. Deze werkt met een willekeurige basis en maakt gebruik van de Grammatrix van het inwendig product ten opzichte van die basis.

STELLING 6.12. *Als $\mathcal{L}$ een lineaire operator op een eindig-dimensionale inwendig productruimte V over $\mathbb{L}$ is en als $\mathcal{B}$ een basis is voor V, dan geldt*

$$\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}^*]} = \left(G_{\mathcal{B}}\,\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}\,G_{\mathcal{B}}{}^{-1}\right)^*.$$

Bewijs: We gebruiken Stelling 5.35 om de inwendige producten te berekenen.

$$\langle \mathcal{L}(\mathbf{x}), \mathbf{y}\rangle = [\mathcal{L}(\mathbf{x})]^*_{\mathcal{B}} G_{\mathcal{B}}[\mathbf{y}]_{\mathcal{B}} = \left(\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}[\mathbf{x}]_{\mathcal{B}}\right)^* G_{\mathcal{B}}\,[\mathbf{y}]_{\mathcal{B}} = ([\mathbf{x}]_{\mathcal{B}})^*\left(\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}\right)^* G_{\mathcal{B}}\,[\mathbf{y}]_{\mathcal{B}}$$

en

$$\langle \mathbf{x}, \mathcal{L}^*(\mathbf{y})\rangle = ([\mathbf{x}]_{\mathcal{B}})^* G_{\mathcal{B}}\,[\mathcal{L}^*(\mathbf{y})]_{\mathcal{B}} = ([\mathbf{x}]_{\mathcal{B}})^* G_{\mathcal{B}}\,\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}^*]}[\mathbf{y}]_{\mathcal{B}}.$$

Omdat deze twee voor elke $\mathbf{x}$ en $\mathbf{y}$ gelijk zijn, volgt: $\left(\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}\right)^* G_{\mathcal{B}} = G_{\mathcal{B}} \underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}^*]}$, zodat

$$\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}^*]} = G_{\mathcal{B}}^{-1} \left(\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}\right)^* G_{\mathcal{B}} = \left(G_{\mathcal{B}} \underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]} {G_{\mathcal{B}}}^{-1}\right)^*,$$

(we gebruiken hier dat de matrix $G_{\mathcal{B}}$ Hermitisch is) hetgeen bewezen moest worden. □

Voorbeeld 6.13. Als bovenstaande stelling toegepast wordt op de operator en inwendig productruimte uit Voorbeeld 6.11 t.o.v. de standaardbasis $\mathcal{E}$, dan hebben we

$$G_{\mathcal{E}} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{en} \quad M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

zodat

$$\underset{\mathcal{E}\leftarrow\mathcal{E}}{[\mathcal{L}^*]} = \left(G_{\mathcal{E}} \underset{\mathcal{E}\leftarrow\mathcal{E}}{[\mathcal{L}]} {G_{\mathcal{E}}}^{-1}\right)^* = \left(\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}\right)^* = \begin{bmatrix} 1 & 3 \\ 0 & -1 \end{bmatrix}.$$

Dit is in overeenstemming met het resultaat van Voorbeeld 6.11. □

6.3 Zelfgeadjungeerde (hermitische) operatoren

We zullen in deze paragraaf de klasse van zogenaamde zelfgeadjungeerde operatoren definiëren. Er zal in deze paragraaf blijken dat bij dit soort operatoren een *orthogonale eigenbasis* (dat is een dus een volledig uit eigenvectoren bestaande orthogonale basis) bestaat voor de vectorruimte waarop de operator werkt, mits deze eindig dimensionale is.

Definitie: Een lineaire operator $\mathcal{L}$ op een inwendig productruimte V heet *zelfgeadjungeerd* als $\mathcal{L}$ gelijk is aan zijn geadjungeerde $\mathcal{L}^*$, wat wil zeggen dat

$$\langle \mathcal{L}(\mathbf{v}), \mathbf{w}\rangle = \langle \mathbf{v}, \mathcal{L}(\mathbf{w})\rangle\,, \text{ voor alle } \mathbf{v} \text{ en } \mathbf{w} \text{ in } V.$$

Als V een reële vectorruimte is, noemen we zo'n operator ook wel *symmetrisch* en als V een complexe vectorruimte zeggen we wel dat de operator *hermitisch* is.

Voorbeeld 6.14. Uit de inleiding van §6.1 (of Stelling 6.9) volgt dat een matrixtransformaties $T : \mathbb{C}^n \to \mathbb{C}^n$ (met standaard inwendig product) met $T(\mathbf{x}) = A\mathbf{x}$ een zelfgeadjungeerde lineaire operator is als en slechts als A een hermitische matrix is.
Evenzo is een matrixtransformatie $T : \mathbb{R}^n \to \mathbb{R}^n$ (met standaard inwendig product) met $T(\mathbf{x}) = A\mathbf{x}$ zelfgeadjungeerd als en slechts als A een symmetrische matrix is. □

De volgende stelling, die bij *eindig dimensionale* inwendige productruimten gebruikt kan worden om na te gaan of een operator zelfgeadjungeerd is, volgt direct uit stelling 6.9.

STELLING 6.15. *Een lineaire operator op een eindig-dimensionale inwendig productruimte is zelfgeadjungeerd als en slechts als de matrix ten opzichte van een orthonomale basis hermitisch is (in het complexe geval) of symmetrisch (in het reële geval).*

Herinner hierbij dat symmetrische matrices per definitie reëel zijn.

Nagaan of een operator op een oneindig-dimensionale vectorruimte zelfgeadjungeerd is, zal met de definitie moeten gebeuren, zie het volgende voorbeeld.

Voorbeeld 6.16. We laten zien dat de lineaire operator $D : \mathrm{Pol}(\mathbb{R}) \to \mathrm{Pol}(\mathbb{R})$ uit voorbeeld 3.44, die gedefinieerd werd door

$$D(p)(X) = (1 - X^2)\frac{d^2p}{dX^2} - 2X\frac{dp}{dX} = (1 - X^2)p''(X) - 2Xp'(X),$$

een symmetrische operator is, mits $\mathrm{Pol}(\mathbb{R})$ voorzien wordt van het inwendig product

$$\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)\,dx.$$

Er geldt namelijk

$$\begin{aligned}
\langle D(p), q\rangle &= \int_{-1}^{1} D(p)(x)q(x)\,dx \\
&= \int_{-1}^{1} \big((1 - x^2)p''(x) - 2xp'(x)\big)\,q(x)\,dx \\
&= \int_{-1}^{1} \frac{d}{dx}\big((1 - x^2)p'(x)\big)\,q(x)\,dx \\
&= \big[(1 - x^2)p'(x)q(x)\big]_{-1}^{1} - \int_{-1}^{1}(1 - x^2)p'(x)q'(x)\,dx \\
&= -\int_{-1}^{1}(1 - x^2)p'(x)q'(x)\,dx.
\end{aligned}$$

Bij het één na laatste =-teken is partieel geïntegreerd. Als in deze uitdrukking p en q omgewisseld worden verandert er niets!! Dat betekent dat $\langle D(p), q\rangle = \langle D(q), p\rangle$. Verder werken we hier met een inwendig product op een reële vectorruimte zodat $\langle D(q), p\rangle = \langle p, D(q)\rangle$. Dus $\langle D(p), q\rangle = \langle p, D(q)\rangle$ en daaraan zien je dat D een symmetrische operator is[1].

Eigenwaarden en eigenvectoren van zelf geadjungeerde operatoren

We herinneren ons de stelling dat iedere Hermitische (symmetrische) matrix A reële eigenwaarden heeft en (unitair) diagonaliseerbaar is, en dat de ruimte $\mathbb{C}^n$ (of $\mathbb{R}^n$) een *orthogonale* basis van eigenvectoren van A heeft(zie Stelling 4.33). Dan vertelt Stelling 6.15 ons:

STELLING 6.17. *Stel $\mathcal{L} : V \to V$ is een zelfgeadjungeerde lineaire operator op een eindig-dimensionale inwendig productruimte V. Dan zijn alle eigenwaarden van $\mathcal{L}$ reëel en V heeft een orthogonale basis van eigenvectoren van $\mathcal{L}$.*

Bewijs: Laat B een orthonormale basis van V zijn en we beschouwen de coördinatisering via deze basis B. Volgens Stelling 5.37 geldt:

$$\langle \mathbf{v}, \mathbf{u}\rangle_V = \langle [\mathbf{v}]_B, [\mathbf{u}]_B\rangle_{\mathbb{C}^n}$$

Beschouw de matrixtransformatie $\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$ $*$. We weten:

[1]De hier geschetste methode wordt in het bij kwantummechanica gebruikte boek toegepast.

1. λ is eigenwaarde van de matrix $\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$ als en slechts als λ eigenwaarde is van de operator $\mathcal{L}$, en

2. $\mathbf{v}$ is eigenvector van $\mathcal{L}$ bij λ als en slechts als $[\mathbf{v}]_B$ eigenvector is van de matrix $\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$ bij λ.

Tenslotte geeft stelling 6.15 dat de matrix $\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$ Hermitisch is. Dan zijn de eigenwaarden van $\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$ reëel, dus geeft 1. dat die van $\mathcal{L}$ dat ook zijn. Tenslotte, een orthogonale basis van $\mathbb{C}^n$ van eigenvectoren van $\underset{\mathcal{B}\leftarrow\mathcal{B}}{[\mathcal{L}]}$ levert volgens 2. (een basis van V van) eigenvectoren van $\mathcal{L}$ op. Volgens $*$ is dat zelfs een orthogonale basis van V. $\square$

We willen soortgelijke resultaten voor het oneindig dimensionale geval. Het bewijs van de volgende reultaten is hetzelfde als dat Stelling 4.15 in § 4.3.)

STELLING 6.18. *Een zelfgeadjungeerde (hermitische, symmetrische) operator $\mathcal{L}$ op een inwendig product ruimte V heeft de volgende eigenschappen:*

(i) de eigenwaarden zijn allemaal reëel;

(ii) eigenvectoren bij verschillende eigenwaarden staan loodrecht.

Bewijs: Neem een eigenwaarde λ en een bijbehorende eigenvector $\mathbf{v}$. Dan hebben we:

$$\lambda \langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{v}, \lambda\mathbf{v} \rangle = \langle \mathbf{v}, \mathcal{L}(\mathbf{v}) \rangle = \langle \mathcal{L}(\mathbf{v}), \mathbf{v} \rangle = \langle \lambda\mathbf{v}, \mathbf{v} \rangle = \overline{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle .$$

Omdat een eigenvector per definitie niet nul is, is $\langle \mathbf{v}, \mathbf{v} \rangle = \|\mathbf{v}\|^2$ ook ongelijk aan nul. De gevonden vergelijking kan dus door $\langle \mathbf{v}, \mathbf{v} \rangle$ gedeeld worden, zodat overblijft $\lambda = \overline{\lambda}$. Dat betekent dat λ reëel is, waarmee de eerste uitspraak bewezen is.
Voor de tweede uitspraak nemen we eigenvectoren $\mathbf{v}_1$ bij eigenwaarde λ_1 en $\mathbf{v}_2$ bij eigenwaarde λ_2 met $\lambda_1 \neq \lambda_2$. Dan volgt

$$\lambda_2 \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \lambda_2\mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathcal{L}(\mathbf{v}_2) \rangle = \langle \mathcal{L}(\mathbf{v}_1), \mathbf{v}_2 \rangle = \langle \lambda_1\mathbf{v}_1, \mathbf{v}_2 \rangle = \overline{\lambda}_1 \langle \mathbf{v}_1, \mathbf{v}_2 \rangle .$$

en dus geldt

$$(\lambda_2 - \overline{\lambda_1}) \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0.$$

Omdat λ_1 vanwege de eerste uitspraak reëel is, hebben we $\overline{\lambda}_1 = \lambda_1$ en omdat $\lambda_1 \neq \lambda_2$ volgt dat $(\lambda_2 - \overline{\lambda_1}) \neq 0$. Maar dan $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$, wat zegt dat $\mathbf{v}_1$ en $\mathbf{v}_2$ loodrecht op elkaar staan. $\square$

Tenslotte, is de eigenschap dat iedere Hermitische (symmetrische) matrix A unitair diagonaliseeerbaar is (equivalent: $\mathbb{C}^n$ (of $\mathbb{R}^n$) heeft een orthonormale basis van eigenvectoren van A) te vertalen naar de oneindig dimensionale situatie? Het antwoord is ja. Echter, we hebben niet voldoende kennis ontwikkeld om de volgende stelling te bewijzen.

STELLING 6.19. *Stel dat $\mathcal{L} : V \to V$ een zelfgeadjungeerde (hermitische, symmetrische) operator op een inwendig product ruimte V is. Dan bestaat er in V een maximaal orthonormale verzameling bestaande uit eigenvectoren van $\mathcal{L}$.*

Voorbeeld 6.20. Kijk nog eens naar de operator D uit de voorbeelden 3.44 en 6.16. In het tweede voorbeeld hebben we gezien dat D een symmetrische operator is. We weten dus dat de eigenwaarden reëel zijn en dat eigenvectoren bij verschillende eigenwaarden loodrecht op elkaar staan. In voorbeeld 3.44 hebben we berekend dat de getallen van de vorm $-n^2-n$ voor $n = 0,1,2,\ldots$ de eigenwaarden van D zijn. Deze zijn inderdaad reëel. Ook hebben we gezien dat het n^{e}-graads (Legendre–)polynoom P_n de eigenruimte bij eigenwaarde $-n^2-n$ opspant.
Met bovenstaande stelling volgt nu dat de polynomen P_0, P_1, P_2, …een oneindige orthogonale verzameling vormt. Het zijn immers eigenvectoren bij verschillende eigenwaarden van een symmetrische operator.
In college 11 werd beweerd dat we deze Legendre–polynomen ook vinden als het Gram-Schmidt proces toegepast wordt op $1, X, X^2, X^3, \ldots$. We zullen nu laten zien dat dit (op een factor na) inderdaad zo is.
Immers: als Gram-Schmidt wordt toegepast op $1, X, X^2, X^3, \ldots$ vind je achtereenvolgens polynomen

p_0 van graad 0,
$p_1 \in \mathrm{Span}\{p_0\}^\perp = \mathrm{Span}\{1\}^\perp$ van graad 1,
$p_2 \in \mathrm{Span}\{p_0,p_1\}^\perp = \mathrm{Span}\{1,X\}^\perp$ van graad 2,
$p_3 \in \mathrm{Span}\{p_0,p_1,p_2\}^\perp = \mathrm{Span}\{1,X,X^2\}^\perp$ van graad 3,
enzovoorts.

Er geldt dus dat p_n een polynoom is van graad n uit $\mathrm{Span}\{1,X,\ldots,X^{n-1}\}^\perp$. Kijk nu naar het polynoom P_n: deze heeft graad n en staat loodrecht op de ruimte opgespannen door $P_0,\ldots,P_{n-1}$. Omdat dit $n-1$ onafhankelijke polynomen zijn in het lineair omhulsel van 1, X, X^2, …, X^{n-1} en omdat de dimensie hiervan $n-1$ is, vormen $P_0,\ldots,P_{n-1}$ een basis voor $\mathrm{Span}\{1,X,X^2,\ldots,X^{n-1}\}$. Blijkbaar is de ruimte opgespannen door $P_0,\ldots,P_{n-1}$ precies gelijk aan $\mathrm{Span}\{1,X,X^2,\ldots,X^{n-1}\}$ en geldt dus dat P_n en p_n polynomen van graad n uit $\mathrm{Span}\{1,X,\ldots,X^{n-1}\}^\perp$ zijn. Omdat $\dim(\mathrm{Span}\{1,X,\ldots,X^{n-1}\}) = n-1$ vormen de polynomen van graad n die loodrecht op deze span staan een 1-dimensionale ruimte. Dat betekent dat p_n en P_n scalaire veelvouden van elkaar zijn en dus op een factor na precies hetzelfde zijn.
Beschouw dit als winst: toepassen van Gram-Schmidt betekent relatief veel rekenwerk terwijl het produceren van de eigenvectoren (en dus van de Legendre polynomen) met de methode van voorbeeld 3.44 veel makkelijker gaat. Tenslotte merken we op dat uit stelling 6.19 volgt dat de verzameling Legendre-polynomen $\{p_0,p_1,\ldots,p_n,\ldots\}$ een maximale orthogonale verzameling in $\mathrm{Pol}(\mathbb{R})$ is.

Toepassing in de kwantummechanica

De zelfgeadjungeerde operatoren spelen een grote rol in de kwantummechanica. De meetbare grootheden (positie, impuls, energie van een deeltje) worden namelijk door zelfgeadjungeerde operatoren beschreven. De eigenwaarden van deze operatoren zijn de waarden die deze meetbare grootheden kunnen aannemen (bij energie zijn het de energieniveaus). Deze operatoren zijn dus van groot belang in de fysica en een voorname reden om zo ver te gaan met lineaire algebra. In voorbeeld 6.21 gaan we hier iets verder op in (dit voorbeeld is niet van belang voor het vak, maar slechts bedoeld om een link te leggen tussen lineaire algebra en kwantummechanica).

Voorbeeld 6.21. We kunnen er niet aan ontkomen om één van de belangrijkste toepassingen van de theorie van hermitische operatoren aan te stippen, namelijk de kwantummechanica. In de kwantummechanica worden toestanden van een deeltje aangegeven door functies die genoteerd worden als $\Psi(\mathbf{x})$. Hierbij is $\mathbf{x}$ de plaats (dus element van $\mathbb{R}^n$) en $\Psi(\mathbf{x})$ een element van $\mathbb{C}^n$. Met zo'n toestandsfunctie kan de kans om een deeltje binnen een bepaald volume V aan te treffen berekend worden. Deze kans is gelijk aan

$$\int_V \|\Psi(\mathbf{x})\|^2 \, d\mathbf{x} = \int_V \overline{\Psi(\mathbf{x})}\Psi(\mathbf{x}) \, d\mathbf{x}.$$

De kans om het deeltje ergens in de hele ruimte aan te treffen is natuurlijk 1. De vectorruimte U opgespannen door alle toestandsfuncties is een hermitisch-productruimte als we het inwendig product definiëren door[2]

$$\langle f, g \rangle = \int_{\mathbb{R}^3} \overline{f(\mathbf{x})} g(\mathbf{x})\, d\mathbf{x}.$$

De toestandsfuncties zijn dan genormeerde vectoren uit U.

Alle te meten waarden van grootheden als impuls en plaats blijken eigenwaarden te zijn van lineaire operatoren die werken op de toestandsfuncties. Als $\Psi(x_1, x_2, x_3) = (\psi_1, \psi_2, \psi_3)$ dan wordt de lineaire operator $\hat{\mathbf{p}} : U \to U$ die de impuls "representeert" gedefinieerd door

$$\hat{\mathbf{p}}(\Psi) = -i \begin{bmatrix} \frac{d\psi_1}{dx_1} \\ \frac{d\psi_2}{dx_2} \\ \frac{d\psi_3}{dx_3} \end{bmatrix}$$

(kort genoteerd: $\hat{\mathbf{p}} = -i\nabla = (-i\frac{d}{dx_1}, -i\frac{d}{dx_2}, -i\frac{d}{dx_3})$) en de operator $\hat{\mathbf{x}} : U \to U$ die de plaats "representeert" gedefinieerd door

$$\hat{\mathbf{x}}(\Psi) = (x_1\psi_1, x_2\psi_2, x_3\psi_3)$$

(kort genoteerd: $\hat{\mathbf{x}} = (x_1, x_2, x_3)$).

Dit alles is zo geconstrueerd dat de verwachtingswaarde van de impuls in een toestand Ψ gelijk is aan $\langle \Psi, \hat{\mathbf{p}}(\Psi) \rangle$ en die van de plaats aan $\langle \Psi, \hat{\mathbf{x}}(\Psi) \rangle$. Als Ψ een eigenvector is (van $\hat{\mathbf{p}}$ respectievelijk $\hat{\mathbf{x}}$) is deze verwachtingswaarde precies de eigenwaarde.

Bij kwantummechanica zal (met stellingen uit de lineaire algebra) blijken dat de eerder genoemde operatoren zelfgeadjungeerd zijn en dus reële eigenwaarden hebben, hetgeen te verwachten is omdat verwachtingen ook reëel zijn. Ook bestaat er (zoals we in §6.7 zullen zien) bij de operatoren een basis voor U die volledig uit eigenvectoren bestaat. Het gevolg daarvan is dat elke toestand een lineaire combinatie is van eigenvectoren en de verwachtingswaarden dezelfde combinatie, maar dan van de bijbehorende eigenwaarden.

Om deze reden wordt in de kwantummechanica van alle meetbare grootheden (=operatoren) geëist dat ze hermitisch zijn.

6.4 Hilbertruimten

De vertaling van de eindigdimensionale resultaten naar het oneindig dimensionale geval is vrij succesvol. Om deze nog verder te verbeteren moeten we ook inzien wat er fout kan gaan. Waar we ons in het eindig dimensionale geval nooit zorgen hoefde te maken was continuïteit van afbeeldingen. Een lineaire afbeelding $L : \mathbb{R}^n \to \mathbb{R}^m$ is gewoon continu. Tuusen oneindig dimensionale ruimtes is dat echter niet het geval. Maar waarschijnlijk zien we op dit moment nog helemaal niet in wat dat zal betekenen.

Omdat we op een inwendig product ruimte V een afstandsbegrip hebben via $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$, kunnen we op een V het begrip: "limiet van een rij" introduceren.

[2]Het product wordt ook vaak geschreven als $\langle f|g \rangle$ en een verwachtingswaarde van $\mathbf{p}$ als $\langle \Psi|\mathbf{p}|\Psi \rangle$. De (toestands-)functies worden vaak geschreven als $|\Psi\rangle$ en deze worden *kets* genoemd. De complex geconjugeerde (toestands-)functies worden dan geschreven als $\langle \Psi|$ en worden *bra*'s genoemd. Aan elkaar geplakt geven de bra $\langle f|$ en de ket $|g\rangle$ het product $\langle f|g\rangle$, net als de woorden aan elkaar geplakt het woord *bra(c)ket* opleveren.

Definitie: Laat V een inwendig product ruimte zijn en stel dat $\{\mathbf{v}_n\}_{n\in\mathbb{N}}$ een rij van vectoren in V is. We zeggen dat deze rij in V convergent is (met limiet $\ell \in V$), als geldt: voor iedere $\varepsilon > 0$ bestaat een $n_0 \in \mathbb{N}$ met de eigenschap: $d(\mathbf{v}_n, \ell) < \varepsilon$, voor iedere $n > n_0$.

We schrijven weer: $\lim\limits_{n\to\infty} \mathbf{v}_n = \ell$. Men kan weer aantonen dat als een rij een limiet heeft, dan is die limiet uniek bepaald. (Merk op dat dit precies de definitie van limiet is die we kennen in $\mathbb{R}$.) Omdat het afstandsbegrip in V gekoppeld is aan de norm op V, heet de rij $\{\mathbf{v}_n\}$ ook wel: "convergent in de norm".
We zijn dit al eerder tegengekomen op $C([-\pi,\pi],\mathbb{R})$. Aan een continue functie $f : [-\pi,\pi] \to \mathbb{R}$ kunnen we de Fourier rij $\{F_n\}_{n\geq 1}$ in $C([-\pi,\pi],\mathbb{R})$ toevoegen.
(Herinner: $F_n(x) = a_0 + \sum\limits_{k=1}^{n}(a_k\cos(kx) + b_k\sin(kx))$.) In de ruimte $V = C([-\pi,\pi],\mathbb{R})$ zal gelden:

$$\lim_{n\to\infty} F_n = f$$

Nu we het limiet begrip hebben kunnen we continuïteit definiëren.

Definitie: Laat $L : V \to W$ een lineaire afbeelding zijn tussen de inwendig product ruimtes V en W. De afbeelding L heet continu als voor iedere rij $\{\mathbf{v}_n\}$ in V met $\lim\limits_{n\to\infty} \mathbf{v}_n = \ell$ geldt:

$$\lim_{n\to\infty} L(\mathbf{v}_n) = L(\ell).$$

De continue lineaire afbeeldingen blijken de belangrijke lineaire afbeeldingen te zijn. Projecties ijn altijd continu. Men kan aantonen:

STELLING 6.22. *Laat $L : V \to W$ een continue lineaire afbeelding zijn tussen de inwendig product ruimtes V en W. De afbeelding L heeft een geadjungeerde afbeelding $L^* : W \to V$.*

Vervolgens zal men ook kritischer zijn op de te bestuderen inwendig productruimtes. Daarom:

Definitie: Laat V een inwendig product ruimte zijn en stel dat $\{\mathbf{v}_n\}_{n\in\mathbb{N}}$ een rij van vectoren in V is. We zeggen dat deze rij een Cauchyrij is als geldt: voor iedere $\varepsilon > 0$ bestaat een $n_0 \in \mathbb{N}$ met de eigenschap: $d(\mathbf{v}_n, \mathbf{v}_m) < \varepsilon$, voor alle $n, m > n_0$.

Intuïtief: een rij is een Cauchyrij als de termen op den duur willekeurig dicht bij elkaar komen. In het bijzonder zal gelden dat convergente rijen Cauchyrijen zijn.

Men kan aantonen dat in sommige ruimten ook het tegenovergestelde geldt:

STELLING 6.23. *Een Cauchyrij in $\mathbb{R}$ zal convergent zijn. Dit geldt ook in $\mathbb{R}^n$ en $\mathbb{C}^n$.*

Inwendig productruimtes met deze eigenschap worden volledig genoemd en blijken de interessante ruimtes te zijn.

Definitie: Een Hilbertruimte V is een inwendig productruimte met de eigenschap dat iedere Cauchyrij convergent is.

Welke ruimtes te we zijn tegengekomen zijn Hilbertruimtes? Niet al te veel!

STELLING 6.24. *Eindig dimensionale inwendigproduct ruimtes zijn Hilbertruimtes. Van de oneindig domensionale ruimtes die we zijn tegengekomen zijn alleen $\ell_2(\mathbb{R})$, $\ell_2(\mathbb{C})$, $L_2([a,b],\mathbb{R})$ en $L_2([a,b],\mathbb{C})$ Hilbertruimtes.*

6.5 De spectraalstelling voor zelfgeadjungeerde operatoren

Herriner de spectraalstelling voor Hermitische matrices, en wel vorm 1 van deze stelling.

STELLING 6.25. *(Spectraalstelling, vorm 1) Laat A een hermitische $n \times n$ matrix zijn. Laat $B = \{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ een orthonormale basis zijn van $\mathbb{C}^n$ bestaande uit eigenvectoren van A, zeg $A\mathbf{q}_i = \lambda_i \mathbf{q}_i$ met $\lambda_i \in \mathbb{R}$ $(i = 1, \ldots, n)$. Dan geldt:*

$$A = \sum_{i=1}^{n} \lambda_i \mathbf{q}_i \mathbf{q}_i{}^* = \sum_{i=1}^{n} \lambda_i [\mathrm{proj}_{\mathrm{Span}\{\mathbf{q}_i\}}]$$

De Hermitische matrix is een lineaire combinatie van projectiematrices $\mathbf{q}_i \mathbf{q}_i{}^* = [\mathrm{proj}_{\mathrm{Span}\{\mathbf{q}_i\}}]$ van projecties op onderling loodrechte $\mathrm{Span}\{\mathbf{q}_i\} \subset E_{\lambda_i}$ met gewichten: de (reële) eigenwaarden.
En nu de zelfgeadjungeerde operatoren $\mathcal{L} : V \to V$. We weten uit stelling 6.19 dat V een maximaal orthonormale familie eigenvectoren heeft. In stelling 6.4 hebben we geleerd de projectie op de 1 dimensionale deelruimte $W = \mathrm{Span}\{\mathbf{q}\}$ (met $\|\mathbf{q}\| = 1$) te beschrijven als $\mathrm{proj}_W = \varphi_{\mathbf{q}} \circ \varphi_{\mathbf{q}}^*$.
En ja, de stelling is nog steeds waar in het oneindig dimensionale geval. Echter, er zijn wat restricyies nodig (die in het eindig dimensionale geval geen extra eisen zijn). De operatoren moeten continu zijn en de onderliggende inwendig product ruimte moet een Hilbertrumte zijn. Voor bewijzen verwijzen we naar de literatuur.

STELLING 6.26. *(Spectraalstelling, vorm 1) Stel dat $\mathcal{L} : V \to V$ een continue zelfgeadjungeerde operator is op een Hilbertruimten V. Laat $\{\mathbf{q}_i | i \in I\}$ een maximaal orthonormale verzameling in V zijn, bestaande uit eigenvectoren van $\mathcal{L}$, zeg $\mathcal{L}(\mathbf{q}_i) = \lambda_i \mathbf{q}_i$ met $\lambda_i \in \mathbb{R}$ ($i \in I$). Dan geldt:*

$$\mathcal{L} = \sum_{i \in I} \lambda_i (\varphi_{\mathbf{q}_i} \circ \varphi^*_{\mathbf{q}_i}) = \sum_{i \in I} \lambda_i \operatorname{proj}_{\operatorname{Span}\{\mathbf{q}_i\}}$$

De zelfgeadjungeerde operator is een oneindige som van van projecties op onderling loodrechte $\operatorname{Span}\{\mathbf{q}_i\} \subset E_{\lambda_i}$ met gewichten: de (reële) eigenwaarden.
Merk op dat er, afgezien van de afwezigheid van het bewijs, nog een probleem aan deze stelling kleeft. Er wordt gesproken over een oneindige som van operatoren en eigenlijk weten we op dit moment niet wat daarmee bedoelt wordt (omdat geen afstandsbegrip op de vectorruimte van alle operatoren van V naar V geintroduceerd is).
We weten dat een nadeel van vorm 1 van de spectraalstelling is dat deze decompositie i.h.a. niet uniek is. Daarvoor hadden we vorm 2. Herriner voor Hermitische matrices:

STELLING 6.27. *(Spectraalstelling, vorm 2) Laat A een hermitische $n \times n$ matrix zijn, zeg met (reëele) eigenwaarden $\lambda_1, \ldots, \lambda_k$ en eigenruimten E_{λ_i} $(i = 1, \ldots, k)$. Dan geldt:*

$$A = \sum_{i=1}^{k} \lambda_i [\operatorname{proj}_{E_{\lambda_i}}]$$

Voor operatoren luidt vorm 2 als volgt.

STELLING 6.28. *(Spectraalstelling, vorm 2) Stel dat $\mathcal{L} : V \to V$ een continue zelfgeadjungeerde operator is op een Hilbertruimte V. Zeg met eigenwaarden $\{\lambda_i : i \in I\}$. Dan geldt:*

$$\mathcal{L} = \sum_{i \in I} \lambda_i \operatorname{proj}_{E_{\lambda_i}}$$

Tenslotte merken we op dat ook hier een probleem is. Natuurlijk, afwezigheid van het bewijs en aanwezigheid van een oneindige sommatie. Maar hier staat ook de projectie op de deelruimte E_λ. Die bestaat kennelijk, ook als E_λ oneindig dimensionaal is. Met onze opgebouwde kennis kunnen wij op dit moment niet inzien waarom die projectie zou bestaan. Daarvoor moet je de literatuur inzien.

6.6 Unitaire en orthogonale operatoren

We behandelen nu een tweede type operator waarvoor ook geldt dat eigenvectoren bij verschillende eigenwaarden loodrecht staan. Dit wordt de vertaling van het begrip unitaire matrix (of orthogonale matrix, in het reële geval) naar de operatoren toe. Stelling 3.4 en de laatste stelling van college 8 kunnen ons op het idee brengen van de volgende definitie.

Definitie: Een lineaire operator $\mathcal{L}$ op een inwendig of productruimte V is een *unitaire operator* (als V een complexe vectorruimte is) of *orthogonale operator* (als V een reële vectorruimte is) als

$$\langle \mathcal{L}(\mathbf{v}), \mathcal{L}(\mathbf{w}) \rangle = \langle \mathbf{v}, \mathbf{w} \rangle \text{ voor alle } \mathbf{v} \text{ en } \mathbf{w} \text{ in } V.$$

Voordat we voorbeelden van unitaire/orthogonale operatoren geven, zullen we een stelling bewijzen die lijkt op Stelling 4.17 voor unitaire matrices. Merk op dat het bewijs hetzelfde is als in Stelling 4.17

STELLING 6.29. *Een unitaire of orthogonale operator $\mathcal{L}$ op een inwendig product ruimte V heeft de volgende eigenschappen:*

(i) de eigenwaarden hebben allemaal modulus 1 (dus als λ een eigenwaarde is, dan $|\lambda| = 1$);

(ii) eigenvectoren bij verschillende eigenwaarden staan loodrecht.

Bewijs: Neem een eigenwaarde λ en een bijbehorende eigenvector $\mathbf{v}$. Er geldt dan

$$\langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathcal{L}(\mathbf{v}), \mathcal{L}(\mathbf{v}) \rangle = \langle \lambda \mathbf{v}, \lambda \mathbf{v} \rangle = \overline{\lambda} \lambda \langle \mathbf{v}, \mathbf{v} \rangle = |\lambda|^2 \langle \mathbf{v}, \mathbf{v} \rangle$$

Omdat een eigenvector per definitie niet nul is, is $\langle \mathbf{v}, \mathbf{v} \rangle = \|\mathbf{v}\|^2$ ook ongelijk aan nul. De gevonden vergelijking kan dus door $\langle \mathbf{v}, \mathbf{v} \rangle$ gedeeld worden, zodat overblijft $|\lambda|^2 = 1$ en dus $|\lambda| = 1$. Hiermee is de eerste uitspraak bewezen.
Voor de tweede uitspraak nemen we eigenvectoren $\mathbf{v}_1$ bij eigenwaarde λ_1 en $\mathbf{v}_2$ bij eigenwaarde λ_2 met $\lambda_1 \neq \lambda_2$. Merk op dat $\overline{\lambda_1} \lambda_1 = |\lambda_1|^2 = 1$ en dus $\overline{\lambda_1} = 1/\lambda_1$. Er volgt

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathcal{L}(\mathbf{v}_1), \mathcal{L}(\mathbf{v}_2) \rangle = \langle \lambda_1 \mathbf{v}_1, \lambda_2 \mathbf{v}_2 \rangle = \overline{\lambda_1} \lambda_2 \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \frac{\lambda_2}{\lambda_1} \langle \mathbf{v}_1, \mathbf{v}_2 \rangle$$

en dus

$$(1 - \frac{\lambda_2}{\lambda_1}) \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0.$$

Omdat $\lambda_1 \neq \lambda_2$ volgt dat $(1 - \frac{\lambda_2}{\lambda_1}) \neq 0$. Maar dan $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$, wat zegt dat $\mathbf{v}_1$ en $\mathbf{v}_2$ loodrecht op elkaar staan. □

Merk op dat uit de definitie volgt een een unitaire operator $\mathcal{L} : V \to V$ injectief is. Inderdaad, als $\mathbf{L}(\mathbf{v}) = \mathbf{0}$, dan geeft de identiteit $0 = \langle \mathbf{0}, \mathbf{0} \rangle = \langle \mathcal{L}(\mathbf{v}), \mathcal{L}(\mathbf{v}) \rangle = \langle \mathbf{v}, \mathbf{v} \rangle$ dat $\|\mathbf{v}\| = 0$ en dus $\mathbf{v} = \mathbf{0}$. We hebben aangetoond dat $\mathrm{Ker}(\mathcal{L}) = \{\mathbf{0}\}$, d.w.z. $\mathcal{L}$ is injectief. Als V eindig dimensionaal is, zal een unitaire operator $\mathcal{L} : V \to V$ injectief zijn en dus ook surjectief en dus inverteerbaar. Oneindig dimensionaal hoeft dat niet.

Voorbeeld 6.30. De shift $\Sigma_2 : \ell_2(\mathbb{R}) \to \ell_2(\mathbb{R})$ met

$$\Sigma_2(x_1, \ldots, x_n, \ldots) = (0, x_1, \ldots, x_n, \ldots)$$

is orthogonaal (ga na), maar niet surjectief (ga na) dus niet inverteerbaar. Evenzo, de shift $\Sigma_2 : \ell_2(\mathbb{C}) \to \ell_2(\mathbb{C})$ is unitair, maar niet inverteerbaar. Omdat deze shift niet inverteerbaar is, kan dus ook niet gelden: $\mathcal{L}^* = \mathcal{L}^{-1}$!

Wel geldt het volgende:

STELLING 6.31. *Laat $\mathcal{L} : V \to V$ een operator zijn. Equivalent zijn:*

1. *$\mathcal{L} : V \to V$ is een surjectieve unitaire operator.*
2. *$\mathcal{L}$ heeft een geadjungeerde en er geldt: $\mathcal{L}^* = \mathcal{L}^{-1}$*

Bewijs: Stel dat $\mathcal{L} : V \to V$ is een surjectieve unitaire operator. Omdat $\mathcal{L}$ ook injectief is, en surjectief, zal $\mathcal{L}$ zeker inverteerbaar zijn. Als we in de identiteit

$$\langle \mathcal{L}(\mathbf{v}), \mathcal{L}(\mathbf{w}) \rangle = \langle \mathbf{v}, \mathbf{w} \rangle, \quad \text{voor alle } \ \mathbf{v}, \mathbf{w}$$

de variable $\mathcal{L}(\mathbf{v})$ even $\mathbf{p}$ noemen, dan kan $\mathbf{p}$ iedere waarde in V aannemen, want $\mathcal{L}$ is surjectief, en er geldt: $\mathbf{v} = \mathcal{L}^{-1}(\mathbf{p})$. De identiteit leest dus als:

$$\langle \mathbf{p}, \mathcal{L}(\mathbf{w}) \rangle = \left\langle \mathcal{L}^{-1}(\mathbf{p}), \mathbf{w} \right\rangle, \quad \text{voor alle } \ \mathbf{p}, \mathbf{w} \in V$$

Dit was de $\mathcal{L}^*$ definieërende karakteristieke eigenschap uit § 6.1 en dus: $\mathcal{L}^* = \mathcal{L}^{-1}$.
Omgekeerd, stel dat $\mathcal{L}^* = \mathcal{L}^{-1}$. Dan is $\mathcal{L}$ inverteerbaar en dus surjectief. Als we in de eigenschap:

$$\langle \mathbf{v}, \mathcal{L}(\mathbf{w}) \rangle = \langle \mathcal{L}^*(\mathbf{v}), \mathbf{w} \rangle, \quad \text{voor alle } \ \mathbf{v}, \mathbf{w} \in V$$

van $\mathcal{L}^*$, voor $\mathcal{L}^*$ $\mathcal{L}^{-1}$ lezen, krijgen we:

$$\langle \mathbf{v}, \mathcal{L}(\mathbf{w}) \rangle = \left\langle \mathcal{L}^{-1}(\mathbf{v}), \mathbf{w} \right\rangle, \quad \text{voor alle } \ \mathbf{v}, \mathbf{w} \in V.$$

Lezen we nu voor $\mathcal{L}^{-1}(\mathbf{v})$, zeg $\mathbf{z}$, dan kan ook $\mathbf{z}$ iedere waarde in V aannemen (want ook $\mathcal{L}^{-1}$ is surjectief) dan $\mathbf{v} = \mathcal{L}(\mathbf{z})$ en we lezen:

$$\langle \mathcal{L}(\mathbf{z}), \mathcal{L}(\mathbf{w}) \rangle = \langle \mathbf{z}, \mathbf{w} \rangle, \quad \text{voor alle } \ \mathbf{z}, \mathbf{w} \in V.$$

Hier staat dat $\mathcal{L}$ unitair is. □

Unitaire en orthogonale afbeeldingen hebben allerlei eigenschappen. Zo geldt bijvoorbeeld dat ze *lengtebewarend* zijn (en orthogonale afbeeldingen ook nog *hoekbewarend* (zie de opgaven). Dat wil zeggen dat het beeld van een vector dezelfde lengte heeft als de oorspronkelijke vector (en dat de hoek tussen twee beelden hetzelfde is als de hoek tussen de oorspronkelijke vectoren). In feite is dat niet zo verrassend: lengtes en hoeken hebben met inwendig producten te maken en unitaire (orthogonale) afbeeldingen zijn per definitie "inwendig-product-bewarend". Genoemde eigenschappen verklaren ook de naamgeving voor deze afbeeldingen: "orthogonale afbeelding" omdat twee onderling loodrechte vectoren door een orthogonale afbeelding weer worden afgebeeld op onderling loodrechte vectoren en "unitair" omdat eenheidsvectoren weer op eenheidsvectoren worden afgebeeld.
De volgende stelling laat zien dat zelfs geldt dat elke lengtebewarende operator unitair (orthogonaal) is.

STELLING 6.32. *Voor een lineaire operator $\mathcal{L}$ op een inwendig productruimte V zijn de volgende voorwaarden equivalent.*

(i) *$\mathcal{L}$ is unitair;*

(ii) *$\|\mathcal{L}(\mathbf{v})\| = \|\mathbf{v}\|$ voor alle $\mathbf{v}$ in V;*

(iii) *$\mathcal{L}$ beeldt eenheidsvectoren af op eenheidsvectoren.*

Bewijs: De bewijzen van "(i)⇒(ii)" en van "(ii)⇒(iii)" mag je zelf geven (zie opgave **??**). We bewijzen nu eerst dat "(iii)⇒(ii)". Neem daartoe een vector $\mathbf{v}$ uit V. Als $\mathbf{v} \neq \mathbf{0}$ kunnen we de eenheidsvector $\hat{\mathbf{v}} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$ maken, waarvoor op grond van (iii) geldt dat $\|\mathcal{L}(\hat{\mathbf{v}})\| = 1$. Omdat $\mathcal{L}$ lineair is volgt

$$\mathcal{L}(\hat{\mathbf{v}}) = \frac{1}{\|\mathbf{v}\|}\mathcal{L}(\mathbf{v}),$$

dus $\|\mathcal{L}(\mathbf{v})\| = \|\mathbf{v}\|$. Natuurlijk geldt (ii) ook als $\mathbf{v} = \mathbf{0}$: omdat $\mathcal{L}$ lineair is weten we dat $\mathcal{L}(\mathbf{0}) = \mathbf{0}$ en dus ook $\|\mathcal{L}(\mathbf{0})\| = \|\mathbf{0}\|$
Nu het lastigste deel van het bewijs: (ii)⇒(i). We bewijzen het voor de situatie dat V een complexe vectorruimte is. Neem $\mathbf{v}$ en $\mathbf{w}$ uit V. Uit (ii) volgt dat $\|\mathcal{L}(\mathbf{v}+\mathbf{w})\|^2 = \|\mathbf{v}+\mathbf{w}\|^2$. Dit uitwerken met behulp van inwendig producten levert

$$\|\mathcal{L}(\mathbf{v})\|^2 + \|\mathcal{L}(\mathbf{w})\|^2 + 2\mathrm{Re}\,\langle\mathcal{L}(\mathbf{v}), \mathcal{L}(\mathbf{w})\rangle = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\mathrm{Re}\,\langle\mathbf{v}, \mathbf{w}\rangle\,.$$

Omdat, wederom op grond van (ii), $\|\mathcal{L}(\mathbf{v})\| = \|\mathbf{v}\|$ en $\|\mathcal{L}(\mathbf{w})\| = \|\mathbf{w}\|$, volgt hieruit

$$\mathrm{Re}\,\langle\mathcal{L}(\mathbf{v}), \mathcal{L}(\mathbf{w})\rangle = \mathrm{Re}\,\langle\mathbf{v}, \mathbf{w}\rangle\,.$$

Nu moet nog aangetoond worden dat ook $\mathrm{Im}\,\langle\mathcal{L}(\mathbf{v}), \mathcal{L}(\mathbf{w})\rangle = \mathrm{Im}\,\langle\mathbf{v}, \mathbf{w}\rangle$. Dat gaat op dezelfde manier door in plaats van $\mathbf{w}$ de vector $i\mathbf{w}$ in te vullen. Je vindt dan namelijk

$$\mathrm{Re}\,\langle\mathcal{L}(\mathbf{v}), \mathcal{L}(i\mathbf{w})\rangle = \mathrm{Re}\,\langle\mathbf{v}, i\mathbf{w}\rangle$$

dus

$$\mathrm{Re}(i\,\langle\mathcal{L}(\mathbf{v}), \mathcal{L}(\mathbf{w})\rangle) = \mathrm{Re}(i\,\langle\mathbf{v}, \mathbf{w}\rangle).$$

Gebruik nu dat voor complexe getallen z geldt dat $\mathrm{Re}(iz) = -\,\mathrm{Im}(z)$. □

Voorbeeld 6.33. Neem een functie h in $C([0, 2\pi], \mathbb{C})$ (met standaard inwendig product) die de eigenschap heeft dat $|h(x)| = 1$ voor alle $x \in [0, 2\pi]$ en definieer de operator $\mathcal{L}$ op $C([0, 2\pi], \mathbb{C})$ door $\mathcal{L}(f) = hf$ (dus $\mathcal{L}(f)(x) = h(x)f(x)$). Dan geldt

$$\begin{aligned}
\|\mathcal{L}(f)\|^2 = \langle\mathcal{L}(f), \mathcal{L}(f)\rangle &= \int_0^{2\pi} \overline{h(x)f(x)}h(x)f(x)\;dx \\
&= \int_0^{2\pi} \overline{h(x)}h(x)\overline{f(x)}f(x)\;dx \\
&= \int_0^{2\pi} |h(x)|^2\overline{f(x)}f(x)\;dx \\
&= \int_0^{2\pi} \overline{f(x)}f(x)\;dx = \|f\|^2.
\end{aligned}$$

De operator is blijkbaar lengtebewarend en dus unitair.

GEVOLG 6.34. *De samenstelling van twee unitaire (orthogonale) operatoren is weer unitair (orthogonaal). (Dit geldt niet voor zelfgeadjungeerde operatoren.)*

Bewijs: Dit is direct duidelijk als we het derde criterium van stelling 6.32 bekijken: als $\mathbf{v}$ een eenheidsvector is, is $\mathcal{L}_2(\mathbf{v})$ ook een eenheidsvector en $\mathcal{L}_1(\mathcal{L}_2(\mathbf{v}))$ dus ook. De samenstelling $\mathcal{L}_1 \circ \mathcal{L}_2$ is dus lengtebewarend en daarom unitair. Dat dit niet geldt voor zelfgeadjungeerde operatoren blijkt onder andere uit het feit dat het product van symmetrische matrices niet weer symmetrisch hoeft te zijn. □

Uit de definitie volgt nog een karakterisering die straks handig blijkt te zijn:
We hebben nog niet gekeken naar matrixrepresentaties van unitaire (orthogonale) operatoren in de situatie dat ze werken op eindig-dimensionale inwendig productruimten. Dat gaan we nu doen. Wel is het dan handig om ten opzichte van een orthonormale basis te werken.

STELLING 6.35. *Stel dat V een eindig-dimensionale inwendig productruimte is met orthonormale basis B en laat $L : V \to V$ een operator zijn. Dan geldt: de operator $L : V \to V$ is een unitaire (orthogonale) operator als en slechts als de transformatiematrix $\underset{B \leftarrow B}{[L]}$ ten opzichte van de* orthonormale *basis B een unitaire (orthogonale) matrix.*

6.7 Normale operatoren en de spectraalstelling

In deze paragraaf gaan we het begrip normale matrix op operator niveau bekijken. en we gaan het begrip normale operator introduceren.

Definitie: Een lineaire operator $\mathcal{L}$ op een inwendig productruimte V is *normaal* als $\mathcal{L}$ en $\mathcal{L}^*$ commuteren, dus als

$$\mathcal{L} \circ \mathcal{L}^* = \mathcal{L}^* \circ \mathcal{L}.$$

Voorbeeld 6.36. Alle zelfgeadjungeerde operatoren en alle surjectieve unitaire (en orthogonale) operatoren zijn normaal.

Een niet surjectieve unitaire operator hoeft niet normaal te zijn.

Voorbeeld 6.37. De shift $\Sigma_2 : \ell_2(\mathbb{C}) \to \ell_2(\mathbb{C})$ met

$$\Sigma_2(x_1, \ldots, x_n, \ldots) = (0, x_1, \ldots, x_n, \ldots)$$

is unitair maar niet surjectief. We weten ook dat ${\Sigma_2}^* = \Sigma_1$ (met $\Sigma_1(x_1, \ldots, x_n, \ldots) = (x_2, \ldots, x_n, \ldots)$. Maar dan geldt: ${\Sigma_2}^* \circ \Sigma_2 = \Sigma_1 \circ \Sigma_2 = Id \neq \Sigma_2 \circ \Sigma_1 = \Sigma_2 \circ {\Sigma_2}^*$.

De volgende stelling geeft het verband tussen normale operatoren op een eindig dimensionale ruimte en normale matrices:

STELLING 6.38. *Een operator $\mathcal{L}V \to V$ op een eindig-dimensionale complexe inwen- dig productruimte V is normaal als en slechts als de transformatiematrix $\underset{B\leftarrow B}{[L]}$ van $\mathcal{L}$ ten opzichte van een* orthonormale *basis B een normale matrix is.*

Bewijs: Dit volgt direct uit het feit dat A^* de matrix van $\mathcal{L}^*$ is ten opzichte van de basis (het is namelijk een orthonormale basis) en het feit dat de matrix van de samenstelling van twee lineaire afbeeldingen ten opzichte van de basis het product van de matrices is. □

We weten:

STELLING 6.39. *Een vierkante (eventueel complexe) matrix A is normaal als en slechts als in $\mathbb{C}^n$ een orthonormale basis van eigenvectoren van A bestaat. Dat wil zeggen, als en slechts als er een unitaire matrix U en een diagonaalmatrix D bestaan zo dat $A = UDU^{-1} = UDU^*$.*

In operatortaal wordt dat:

STELLING 6.40 (De spectraalstelling, eenvoudige formulering)**.** *Een lineaire operator $\mathcal{L}$: $V \to V$ op een eindig-dimensionale complexe inwendig productruimte V is een normale operator als en slechts als in V een orthonormale basis van eigenvectoren van $\mathcal{L}$ bestaat.*

Tenslotte de spectraalstellingen. Eerst vorm 1.

STELLING 6.41. *(Spectraalstelling, vorm 1) Stel dat $\mathcal{L} : V \to V$ een continue normale operator is op een Hilbertruimten V. Dan bestaat een maximaal orthonormale verzameling $\{\mathbf{q}_i | i \in I\}$ in V zijn, bestaande uit eigenvectoren van $\mathcal{L}$, zeg $\mathcal{L}(\mathbf{q}_i) = \lambda_i \mathbf{q}_i$ met $\lambda_i \in \mathbb{C}$ $(i \in I)$. Dan geldt:*

$$\mathcal{L} = \sum_{i\in I} \lambda_i(\varphi_{\mathbf{q}_i} \circ \varphi^*_{\mathbf{q}_i}) = \sum_{i\in I} \lambda_i \operatorname{proj}_{\mathrm{Span}\{\mathbf{q}_i\}}$$

Merk op dat stelling 6.26 een onderdeel van stelling 6.41 is. Echter bij stelling 6.26 is er sprake van reele eigenwaarden. Bij stelling 6.41 kunnen de eigenwaarden complex zijn.

We weten dat een nadeel van vorm 1 van de spectraalstelling is dat deze decompositie i.h.a. niet uniek is. Daarvoor hadden we vorm 2.

Voor operatoren luidt vorm 2 als volgt.

STELLING 6.42. *(Spectraalstelling, vorm 2) Stel dat $\mathcal{L} : V \to V$ een continue normale operator is op een Hilbertruimte V. Zeg met eigenwaarden $\{\lambda_i : i \in I\}$. Dan geldt:*

$$\mathcal{L} = \sum_{i \in I} \lambda_i \operatorname{proj}_{E_{\lambda_i}}$$

Toegift 1: Een voorbeeld van een unitaire matrix, de discrete Fourrier transformatie

De discrete Fourrier transformatie (de DFT) is een lineaire afbeelding $F : \mathbb{C}^N \to \mathbb{C}^N$. Deze voert een rij complexe getallen $(z_0, z_1, \cdots, z_{N-1})$ ter lengte N om in een rij $(Z_0, Z_1, \cdots, Z_{N-1})$ van complexe getallen ter lengte N, via de formule:

$$Z_k = \sum_{n=0}^{N-1} z_n e^{-\frac{2\pi i}{N}kn}, \qquad k = 0, 1, \cdots, N-1$$

Dit is niet de enige formule die men tegenkomt in de literatuur, ook de formules:

$$Z_k = \frac{1}{N}\sum_{n=0}^{N-1} z_n e^{-\frac{2\pi i}{N}kn}, \qquad k = 0, 1, \cdots, N-1$$

en

$$Z_k = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} z_n e^{-\frac{2\pi i}{N}kn}, \qquad k = 0, 1, \cdots, N-1$$

zijn gebruikelijke “DFT’s”. Wij kiezen voor de laatste formule omdat dit wiskundig gezien de mooiste is. Ook zullen we trouw blijven aan de lineaire algebra notatie, de vector wordt vertikaal geschreven en genummerd van 1 tot N (i.p.v. 0 tot $N-1$).

Dus als $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix}$, dan $F(\mathbf{z}) = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_N \end{bmatrix} = \mathbf{Z}$ met $Z_k = \dfrac{1}{\sqrt{N}}\displaystyle\sum_{n=1}^{N} z_n e^{-\frac{2\pi i}{N}(k-1)(n-1)}$, $\quad k = 1, 2, \cdots, N$.

Als we schrijven $\zeta = e^{-\frac{2\pi i}{N}}$ (en omdat $\zeta^k = e^{-\frac{2\pi i}{N}k}$ (en $\zeta^0 = 1$)) verkrijgen we zo:

$$F(\mathbf{z}) = \mathbf{Z} = \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \\ \vdots \\ Z_N \end{bmatrix} = \frac{1}{\sqrt{N}}\begin{bmatrix} z_1 & + & z_2 & + & z_3 & + & \cdots & + & z_N \\ z_1 & + & z_2\zeta & + & z_3\zeta^2 & + & \cdots & + & z_N\zeta^{N-1} \\ z_1 & + & z_2\zeta^2 & + & z_3\zeta^4 & + & \cdots & + & z_N\zeta^{2(N-1)} \\ \vdots & + & \vdots & + & \vdots & + & \cdots & + & \vdots \\ z_1 & + & z_2\zeta^{N-1} & + & z_3\zeta^{2(N-1)} & + & \cdots & + & z_N\zeta^{(N-1)^2} \end{bmatrix} =$$

$$= \frac{1}{\sqrt{N}}\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \zeta & \zeta^2 & \cdots & \zeta^{N-1} \\ 1 & \zeta^2 & \zeta^4 & \cdots & \zeta^{2(N-1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \zeta^{N-1} & \zeta^{2(N-1)} & \cdots & \zeta^{(N-1)^2} \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_N \end{bmatrix} = U\mathbf{z}$$

We hebben laten zien dat de discrete Fourrier transformatie inderdaad een lineaire transformatie is met standaardmatrix bovenstaande matrix U. Om te laten zien dat $U = [\,\mathbf{u}_1 \cdots, \mathbf{u}_N\,]$ een unitaire matrix is volstaat het om te laten zien dat de verzameling $\{\mathbf{u}_1 \cdots, \mathbf{u}_N\}$ orthonormaal is. Merk op dat

$$U_{i,j} = \frac{1}{\sqrt{N}} e^{-\frac{2\pi i}{N}(i-1)(j-1)}$$

Daarom

$$\langle \mathbf{u}_\ell, \mathbf{u}_k \rangle = \frac{1}{N} \sum_{j=1}^{N} \overline{e^{-\frac{2\pi i}{N}(\ell-1)(j-1)}} e^{-\frac{2\pi i}{N}(k-1)(j-1)} = \frac{1}{N} \sum_{j=1}^{N} e^{+\frac{2\pi i}{N}(\ell-1)(j-1)} e^{-\frac{2\pi i}{N}(k-1)(j-1)} =$$

$$\frac{1}{N} \sum_{j=1}^{N} e^{-\frac{2\pi i}{N}(k-\ell)(j-1)} = \frac{1}{N} \sum_{j=1}^{N} \left(e^{-\frac{2\pi i}{N}(k-\ell)}\right)^{(j-1)}$$

Hier staat de som van een meetkundige reeks en omdat $1 + a + \cdots + a^{N-1} = \dfrac{1-a^N}{1-a}$ volgt:

$$\langle \mathbf{u}_\ell, \mathbf{u}_k \rangle = \frac{1}{N} \sum_{j=1}^{N} \left(e^{-\frac{2\pi i}{N}(k-\ell)}\right)^{(j-1)} = \frac{1}{N} \frac{1-(e^{-\frac{2\pi i}{N}(k-\ell)})^N}{1-e^{-\frac{2\pi i}{N}(k-\ell)}} = \frac{1}{N} \frac{1-e^{-2\pi i(k-\ell)}}{1-e^{-\frac{2\pi i}{N}(k-\ell)}} = \frac{1}{N} \frac{1-1}{1-e^{-\frac{2\pi i}{N}(k-\ell)}} = 0$$

Dus inderdaad, $\mathbf{u}_\ell \perp \mathbf{u}_k$, als $k \neq \ell$. Vervolgens:

$$\|\mathbf{u}_\ell\|^2 = \langle \mathbf{u}_\ell, \mathbf{u}_\ell \rangle = \frac{1}{N} \sum_{j=1}^{N} \overline{e^{-\frac{2\pi i}{N}(\ell-1)(j-1)}} e^{-\frac{2\pi i}{N}(\ell-1)(j-1)} = \frac{1}{N} \sum_{j=1}^{N} e^{+\frac{2\pi i}{N}(\ell-1)(j-1)} e^{-\frac{2\pi i}{N}(k-1)(j-1)} = \frac{1}{N} \sum_{j=1}^{N} 1 = 1$$

Dus inderdaad, $\|\mathbf{u}_\ell\| = 1$, voor $\ell = 1 \cdots N$. We hebben aangetoond dat $\{\mathbf{u}_1, \ldots, \mathbf{u}_N\}$ een orthonormale verzameling in $\mathbb{C}^N$ is en we concluderen dat U een unitaire matrix is.

Opmerking: $\{\mathbf{u}_1, \ldots, \mathbf{u}_N\}$ is een veel gebruikte orthonormale basis voor $\mathbb{C}^N$. □

Nu is de inverse operator ook te bepalen. Merk op dat $U^T = U$ en we zien dat voor de unitaire U geldt: $U^{-1} = U^* = \overline{U}$.
Omdat $\mathbf{Z} = F(\mathbf{z}) = U\mathbf{z}$ volgt: $\mathbf{z} = U^{-1}\mathbf{Z} = \overline{U}\mathbf{Z}$, m.a.w.:

$$z_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} Z_n e^{+\frac{2\pi i}{N}kn}, \quad k = 1, \cdots, N$$

Opmerking: Als de DFT formule

$$Z_k = \sum_{n=1}^{N} z_n e^{-\frac{2\pi i}{N}kn}, \quad k = 1, \cdots, N$$

gebruikt wordt, dan ziet de inverse formules er als volgt uit. Kennelijk geldt nu: $\mathbf{Z} = B\mathbf{z}$ met $B = \sqrt{N}U$. Daarom $B^{-1} = (\sqrt{N}U)^{-1} = \dfrac{1}{\sqrt{N}}U^{-1} = \dfrac{1}{\sqrt{N}}\overline{U}$ en dus:

$$z_k = \frac{1}{N} \sum_{n=1}^{N} Z_n e^{+\frac{2\pi i}{N}kn}, \quad k = 1, \cdots, N \qquad \square$$

Omdat de DFT matrix unitair is geldt: $\langle F(\mathbf{v}), F(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$ (zie opgave ?? van college 2) en dus:

$$\sum_{n=1}^{N} \overline{V_n} W_n = \sum_{n=1}^{N} \overline{v_n} w_n \qquad \text{(formule Plancherel)}$$

en

$$\|\mathbf{V}\|^2 = \sum_{n=1}^{N} \overline{V_n} V_n = \sum_{n=1}^{N} \overline{v_n} v_n = \|\mathbf{v}\|^2 \qquad \text{(formule Parseval)}$$

Opmerking: Als de DFT formule

$$Z_k = \sum_{n=1}^{N} z_n e^{-\frac{2\pi i}{N} kn}, \qquad k = 1, \cdots, N$$

gebruikt wordt, dan zien de formules van Plancherel en Parseval er als volgt uit.

$$\frac{1}{N} \sum_{n=1}^{N} \overline{V_n} W_n = \sum_{n=1}^{N} \overline{v_n} w_n \qquad \text{(formule Plancherel)}$$

en

$$\frac{1}{N} \|\mathbf{V}\|^2 = \frac{1}{N} \sum_{n=1}^{N} \overline{V_n} V_n = \sum_{n=1}^{N} \overline{v_n} v_n = \|\mathbf{v}\|^2 \qquad \text{(formule Parseval)}$$

Toegift 2: Commuterende matrices en simultaan diagonaliseerbaarheid

Definitie: De reële matrices A en B heten simultaan reëel diagonaliseerbaar als een reële inverteerbare matrix P bestaat en twee reële diagonaalmatrices D_1 en D_2 zodat $A = PD_1P^{-1}$ en $B = PD_2P^{-1}$.

Simultane reële diagonaliseerbaarheid van A en B is equivalent met het bestaan van een basis van $\mathbb{R}^n$ die zowel bestaat uit eigenvectoren van A als uit eigenvectoren van B. Evenzo:

Definitie: De matrices A en B heten simultaan (complex) diagonaliseerbaar als een (complexe) inverteerbare matrix P bestaat en twee (complexe) diagonaalmatrices D_1 en D_2 zodat $A = PD_1P^{-1}$ en $B = PD_2P^{-1}$.

Simultane (complexe) diagonaliseerbaarheid van A en B is equivalent met het bestaan van een basis van $\mathbb{C}^n$ die zowel bestaat uit eigenvectoren van A als uit eigenvectoren van B

LEMMA 8.1. *Laat A en B twee simultaan diagonaliseerbare matrices zijn. Dan geldt: $AB = BA$.*

Bewijs: Stel D_1 en D_2 zijn diagonaalmatrices en P een matrix zodat geldt: $A = PD_1P^{-1}$ en $B = PD_2P^{-1}$. Dan geldt:

$$AB = PD_1P^{-1}PD_2P^{-1} = PD_1D_2P^{-1} \quad \text{en} \quad BA = PD_2P^{-1}PD_1P^{-1} = PD_2D_1P^{-1}.$$

Omdat voor diagonaalmatrices geldt $D_1D_2 = D_2D_1$ concluderen we: $D_1D_2 = D_2D_1$. □

Verrassend voor mij geldt ook het omgekeerde. Hieraan is al af te lezen dat commutativiteit van het product van twee matrices een strenge eis is op A en B.

STELLING 8.2. *1. Laat A en B twee reëel diagonaliseerbare matrices zijn met: $AB = BA$. Dan geldt: A en B zijn simultaan reëel diagonaliseerbaar.*
2. Laat A en B twee diagonaliseerbare matrices zijn met: $AB = BA$. Dan geldt: A en B zijn simultaan diagonaliseerbaar.

We zullen alleen het bewijs leveren voor het reële geval. Het complexe geval gaat letterlijk hetzelfde, vervang $\mathbb{R}^n$ steeds door $\mathbb{C}^n$. Het bewijs is echter niet geheel triviaal en vergt enkelijke lemma's.

Definitie: Laat A een $n \times n$ matrix zijn. We zeggen dat een deelruimte W van $\mathbb{R}^n$ invariant is onder A (of: A-invariant is) als geldt: als $\mathbf{w} \in W$ dan ook $A\mathbf{w} \in W$, voor alle $\mathbf{w} \in W$.

LEMMA 8.3. *Laat A een reële $n \times n$ matrix zijn en W een A-invariante deelruimte van $\mathbb{R}^n$. Dan geldt: als $\mathbf{v}_1, \dots \mathbf{v}_k$ eigenvectoren van A zijn bij verschillende eigenwaarden $\lambda_1, \dots, \lambda_k$ met $\mathbf{v}_1 + \dots + \mathbf{v}_k \in W$, dan $\mathbf{v}_i \in W$, voor $i = 1, \dots, k$.*

Bewijs: We tonen dit aan met inductie naar k.
Als $k = 1$ dan is de uitspraak duidelijk correct.
Stel de uitspraak is juist bij ieder k-tal vectoren van A bij verschillende eigenwaarden.
Neem nu een $k+1$-tal eigenvectoren bij verschillende eigenwaarden, zeg de eigenvectoren $\mathbf{v}_1, \dots \mathbf{v}_k, \mathbf{v}_{k+1}$ bij de onderling verschillende eigenwaarden $\lambda_1, \dots, \lambda_k, \lambda_{k+1}$ met

$$\mathbf{v}_1 + \dots + \mathbf{v}_k + \mathbf{v}_{k+1} \in W. \quad (1)$$

Onze aannames impliceren dat $A(\mathbf{v}_1 + \dots + \mathbf{v}_k + \mathbf{v}_{k+1}) \in W \Rightarrow A(\mathbf{v}_1) + \dots + A(\mathbf{v}_k) + A(\mathbf{v}_{k+1}) \in W$ en dus:

$$\lambda_1 \mathbf{v}_1 + \dots + \lambda_k \mathbf{v}_k + \lambda_{k+1} \mathbf{v}_{k+1} \in W. \quad (2)$$

Vermenigvuldig de eerste formule met λ_{k+1} en trek de tweede formule hiervan af. We verkrijgen:

$$(\lambda_{k+1} - \lambda_1)\mathbf{v}_1 + \dots + (\lambda_{k+1} - \lambda_k)\mathbf{v}_k \in W.$$

Omdat $\lambda_{k+1} - \lambda_i \neq 0$ zijn $(\lambda_{k+1} - \lambda_i)\mathbf{v}_i$ $(i = 1, \dots, k)$ k-veel eigenvectoren bij verschillende eigenwaarden λ_i $(i = 1, \dots, k)$.
De inductieaanname geeft dat $(\lambda_{k+1} - \lambda_i)\mathbf{v}_i \in W$ voor $(i = 1, \dots, k)$. En dus (omdat $\lambda_{k+1} - \lambda_i \neq 0$) volgt ook dat $\mathbf{v}_i \in W$ voor $(i = 1, \dots, k)$. Tenslotte hebben we dan ook $\mathbf{v}_{k+1} \in W$, als verschil van de twee vectoren $\mathbf{v}_1 + \dots + \mathbf{v}_k + \mathbf{v}_{k+1}$ en $\mathbf{v}_1 + \dots + \mathbf{v}_k$ uit W.
Hiermee is ook de uitspraak voor $k + 1$ vectoren bewezen. □

LEMMA 8.4. *Laat A een reëel diagonaliseerbare $n \times n$ matrix zijn en W een A-invariante deelruimte van $\mathbb{R}^n$. Dan geldt: W heeft een basis bestaande uit eigenvectoren van de matrix A.*

Bewijs: De matrix A is reëel diagonaliseerbaar, dus bestaat er een basis $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ van $\mathbb{R}^n$ bestaande uit eigenvectoren van A.
De deelruimte W heeft ook een basis, zeg $\mathbf{w}_1, \dots, \mathbf{w}_k$.
Iedere van deze vectoren $\mathbf{w}_i$ is een sommatie van de basis van $\mathbb{R}^n$ bestaande uit eigenvectoren van A. Door in deze sommatie eigenvectoren bij dezelfde eigenwaarden bij elkaar te nemen en te zien als één eigenvector, zien we: "iedere vector $\mathbf{w}_i$ is te schrijven als:

$$\mathbf{w}_i = \mathbf{V}_1^i + \dots + \mathbf{V}_{\ell_i}^i$$

waarbij $\mathbf{V}_1^i, \dots, \mathbf{V}_{\ell_i}^i$ eigenvectoren van A zijn bij verschillende eigenwaarden.
Via het vorige lemma (hier gebruiken we dat W A-invariant is) kunnen we concluderen dat de eigenvectoren $\mathbf{V}_1^i, \dots, \mathbf{V}_{\ell_i}^i$ in W moeten zitten.
We kunnen concluderen dat W wordt opgespannen door eindig veel eigenvectoren van A. Via de uitdunningsstelling kunnen we hieruit een basis (van eigenvectoren van A) voor W kiezen. □

We hebben genoeg bagage ingeladen om stelling 8.2 te bewijzen.

Bewijs: Laat A en B commuterende reëel diagonaliserende matrices zijn.
Beschouw de eigenruimten van B, zeg $E_{\lambda_1}, \dots, E_{\lambda_k}$.
Claim: ieder van deze eigenruimtes is A-invariant.
Om deze claim te bewijzen is het commuteren $AB = BA$ nodig. Kies E_λ, een eigenruimte van B, en kies $\mathbf{x} \in E_\lambda$. Dan geldt zeker $B\mathbf{x} = \lambda\mathbf{x}$.

Uit $AB = BA$ volgt: $AB\mathbf{x} = BA\mathbf{x}$, dus $A\lambda\mathbf{x} = BA\mathbf{x}$ en dus $\lambda(A\mathbf{x}) = B(A\mathbf{x})$. We lezen hieruit: $A\mathbf{x}$ is een eigenvector van B bij de eigenwaarde λ, d.w.z. $A\mathbf{x} \in E_\lambda$.
Hiermee is de claim bewezen.
Nu passen we lemma 8.4 toe en concluderen dat iedere eigenruimte E_λ van B een basis heeft bestaande uit eigenvectoren van A. Deze basis bestaat natuurlijk ook uit eigenvectoren van B.
Als we op deze manier alle eigenruimtes van B doorlopen vinden we een basis van $\mathbb{R}^n$ bestaande uit zowel eigenvectoren van A als van B, en daarmee is de simultane diagonalisering van A en B gevonden. $\square$

Opmerking: Pas op. In bovenstaande situatie kan men niet concluderen dat A en B dezelfde eigenruimtes hebben. Neem maar $B = I$ en A een matrix met meerdere eigenruimtes.

Bovenstaande stelling gaat ook op als meerdere matrices commuteren. Om precies te zijn:

STELLING 8.5. *Laat $A_1, \ldots, A_k$ diagonaliseerbare matrices zijn met: $A_iA_j = A_jA_i$, voor alle i, j. Dan geldt: $A_1, \ldots, A_k$ zijn simultaan diagonaliseerbaar (d.w.z. er is één matrix P zodat $A_i = PD_iP^{-1}$, voor alle i).*

Toegift 3: Gegeneraliseerde eigenvectoren, de Jordan normaalvorm

Herinner: een vector $\mathbf{v} \in \mathbb{C}^n$ heet een eigenvector van de matrix A bij de eigenwaarde λ als $(A - \lambda I)\mathbf{v} = \mathbf{0}$. We weten: een matrix A is (complex) diagonaliseerbaar als en slechts als $\mathbb{C}^n$ een basis heeft bestaande uit eigenvectoren van A.
Wat als zo'n basis niet bestaat? Kennelijk is A dan niet te diagonaliseren. Maar bestaat er dan wellicht wel een andere vorm/type standaard matrix waarmee A gelijkvormig is?

Definitie: Laat $\lambda \in \mathbb{C}$ een getal zijn. Een $k \times k$ matrix van de vorm:

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

heet een **Jordanblok** en wordt genoteerd met J_λ^k of kortweg J_λ of zelfs J.

Merk op: bij het Jordanblok J_λ bestaat de diagonaal uit louter lambda's en daarboven een "bovendiagonaal" van énen. De rest van de kentallen zijn 0.

Definitie: Een matrix A heet in **Jordannormaalvorm** als deze bestaat uit Jordanblokken over de diagonaal en de rest louter uit nullen

Voorbeeld 9.1. De matrices A_1 en A_2 met:

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix} \quad A_2 = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

zijn matrices in Jordannormaalvorm. Inderdaad, de matrix A_1 bestaat op de diagonaal uit de Jordanblokken:

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

Evenzo: A_2 bestaat uit de Jordanblokken:

$$\begin{bmatrix} 2 \end{bmatrix} \quad \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

Men kan bewijzen:

STELLING 9.2. *Iedere matrix A is gelijkvormig met een matrix in Jordannormaalvorm*

Het bewijs van bovenstaande stelling is verdraaid lastig en presenteren we niet. We zullen wel enige stappen aangegeven hoe een matrix op Jordannormaalvorm te brengen is. Als eerste stap geven we de volgende stelling.

STELLING 9.3. *Als de matrix A gelijkvormig is met de matrix J in Jordannormaalvorm bestaande uit de Jordanblokken $J_{\lambda_1}^{\ell_1}, \ldots, J_{\lambda_k}^{\ell_k}$, dan zijn $\lambda_1, \ldots, \lambda_k$ precies de eigenwaarden van A.*

Bewijs: Omdat A en J gelijkvormig zijn hebben ze hetzelfde karakteristiek polynoom. H Het karakteristiek polynoom van J is simpel te bepalen:

$$p_A(\lambda) = p_J(\lambda) = (\lambda_1 - \lambda)^{l_1} \cdots (\lambda_k - \lambda)^{\ell_k} \quad \text{(ga na)}$$

en de conclusie volgt. □

Om tot de volgende stappen te komen beschouwen we de matrices A_1 en A_2 uit Voorbeeld 9.1 nogmaals.

Voorbeeld 9.4. (Vervolg) Bekijk nogmaals de matrix A_1. Dan:

1. De vector $\mathbf{e}_1$ is eigenvector van A_1 bij $\lambda = 1$, dus: $A\mathbf{e}_1 = 1\mathbf{e}_1$, equivalent: $(A_1 - 1I_6)\mathbf{e}_1 = \mathbf{0}$.
2. De vector $\mathbf{e}_2$ is ook een eigenvector van A_1, echter bij de eigenwaarde $\lambda = 2$.
   Dus: $(A_1 - 2I_6)\mathbf{e}_2 = \mathbf{0}$.
3. De vector $\mathbf{e}_3$ is geen eigenvector van A_1, wel geldt: $(A_1 - 2I_6)\mathbf{e}_3 = \mathbf{e}_2$. $(A_1 - 2I_6)^2\mathbf{e}_3 = \mathbf{0}$.
4. De vector $\mathbf{e}_4$ is weer een eigenvector van A_1 en wel bij de eigenwaarde $\lambda = 3$.
   Dus: $(A_1 - 3I_6)\mathbf{e}_4 = \mathbf{0}$.
5. De vector $\mathbf{e}_5$ is geen eigenvector van A_1, wel geldt: $(A_1 - 3I_6)\mathbf{e}_4 = \mathbf{e}_3$. $(A_1 - 3I_6)^2\mathbf{e}_5 = \mathbf{0}$.
6. De vector $\mathbf{e}_6$ is geen eigenvector van A_1, wel geldt: $(A_1 - 3I_6)\mathbf{e}_6 = \mathbf{e}_5$. $(A_1 - 3I_6)^3\mathbf{e}_6 = \mathbf{0}$.

Min of meer dezelfde observaties zijn geldig voor A_2, met dat verschil dat al de eigenwaarden 2 zijn. Er geldt bijvoorbeeld: $(A_2 - 2I_6)^3\mathbf{e}_6 = \mathbf{0}$.

We stellen vervolgens de vraag: bij een gegeven lineaire afbeelding $T : \mathbb{C}^n \to \mathbb{C}^n$: wat moet gelden voor een basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ opdat de transformatiematrix $\underset{B \leftarrow B}{[T]} = [\, [T(\mathbf{b}_1)]_B \cdots [T(\mathbf{b}_n)]_B \,]$ een Jordanblok J_λ^n is?
De volgende stelling is simpel te verifiëren.

STELLING 9.5. *Gegeven een afbeelding $T : \mathbb{C}^n \to \mathbb{C}^n$ en een basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$. Equivalent zijn:*

1. *De transformatiematrix $\underset{B \leftarrow B}{[T]} =$ is een Jordanblok J^n_λ.*
2. *Er geldt: $(T - \lambda I)(\mathbf{b}_1) = \mathbf{0}$ en $(T - \lambda I)(\mathbf{b}_k) = \mathbf{b}_{k-1}$ voor $k = 2, \ldots, n$.*

Hier staat dus: $\mathbf{b}_1$ is eigenvector van T bij λ en $\mathbf{b}_2$ is een oerbeeld $\mathbf{b}_1$ onder $T - \lambda I$ etc. In het bijzonder zal gelden voor de basis B uit Stelling 9.5: $(T - \lambda I)^k(\mathbf{b}_k) = \mathbf{0}$.

Definitie: Laat A een $n \times n$ matrix zijn. Een vector $\mathbf{b} \in \mathbb{C}^n$ heet een **gegeneraliseerde eigenvector** van A bij de eigenwaarde λ als $k \in \mathbb{N}$ bestaat zodat $(A - \lambda I)^k(\mathbf{b}) = \mathbf{0}$.
Een **cykel** van gegeneraliseerde eigenvectoren van A is een rij $\mathbf{b}_1, \ldots, \mathbf{b}_\ell$ van gegeneraliseerde eigenvectoren met: $(A - \lambda I)(\mathbf{b}_1) = \mathbf{0}$ en $(A - \lambda I)(\mathbf{b}_k) = \mathbf{b}_{k-1}$ voor $k = 2, \ldots, \ell$.

De cykel wordt genoteerd met

$$\mathbf{b}_\ell \to \cdots \to \mathbf{b}_2 \to \mathbf{b}_1 \to \mathbf{0}$$

Hier staat de pijl voor de operatie (afbeelding) $A - \lambda I$. Merk op dat in deze notatie van de cykel alleen de laatste vector $\mathbf{b}_1$ een eigenvector is.
Zonder bewijs vermelden we:

STELLING 9.6. *Laat A een $n \times n$ matrix zijn.*

1. *Als $\mathbf{b}_\ell \to \cdots \to \mathbf{b}_2 \to \mathbf{b}_1 \to \mathbf{0}$ een cykel in A is dan zijn de vectoren $\mathbf{b}_1, \ldots, \mathbf{b}_\ell$ lineair onafhankelijk.*
2. *Als $\mathbf{b}^i_{\ell_i} \to \cdots \to \mathbf{b^i}_1 \to \mathbf{0}$ $(i = 1, \ldots, k)$ cykels in A zijn waarvan de eigenvectoren $\{\mathbf{b}^i_1 | i = 1..k\}$ onafhankelijk zijn, dan is de hele verzameling gegeneraliseerde eigenvectoren $\{\mathbf{b}^i_j | j = 1..\ell_i,\ i = 1..k\}$ onafhankelijk.*

STELLING 9.7. *Laat A een $n \times n$ matrix zijn. Dan heeft $\mathbb{C}^n$ een basis B die puur bestaat uit cykels van gegeneraliseerde eigenvectoren van A. Zeg dit zijn de cykels $\mathbf{b}^i_{\ell_i} \to \cdots \to \mathbf{b^i}_1 \to \mathbf{0}$ bij λ_i $(i = 1, \ldots, k)$. Als*

$$P = [\,\mathbf{b}^1_1 \cdots \mathbf{b}^1_{\ell_1}\, \mathbf{b}^2_1 \cdots \mathbf{b}^2_{\ell_2} \cdots\cdots ,]$$

dan is $A = PJP^{-1}$ met J een matrix in Jordannormaalvorm bestaande uit de Jordanblokken $J^{\ell_1}_{\lambda_1}, \ldots, J^{\ell_k}_{\lambda_k}$. Op de volgorde van de Jordanblokken na, is de matrix J uniek. Die heet dan ook de Jordaannormaalvorm van A.

Een matrix op Jordannormaalvorm brengen is dus een zoektocht naar de cykels van de matrix.

Voorbeeld 9.8. (Zie § 7.8, opgave 17 uit Boyce, duPrima.)
Beschouw de matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ -3 & 2 & 4 \end{bmatrix}$. Enig rekenwerk laat zien dat $p_A(\lambda) = (2 - \lambda)^3$. De enige eigenwaarden van A is dus $\lambda = 2$ en enig rekenwerk laat zien dat $\dim(E_2) = 1$ en dat $\left\{ \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$ een basis voor de eigenruimte is. Tot zover niets nieuws.
Nu kunnen we met onze nieuwe kennis concluderen dat dit moet betekenen dat de Jordannormaalvorm van A moet bestaan uit één Jordanblok, en wel

$$J^3_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Vervolgens zoeken we een bijbehorende cykel.

Het eindpunt van de cykel moet (een veelvoud van) $\mathbf{b}_1 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ zijn. Omdat $(A-2I)(\mathbf{b}_2) = \mathbf{b}_1$ lossen we op $(A-2I)\mathbf{x} = \mathbf{b}_1$ wat levert: $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$. Kies $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. (We moeten een $\mathbf{b}_2$ nemen die onder $A-2I$ een oerbeeld heeft. Gelukkig doet dat hier er niet toe, iedere keuze van $\mathbf{b}_2$ zal zo'n oerbeeld hebben!)

Oplossen van $(A-2I)\mathbf{x} = \mathbf{b}_2$ levert $\mathbf{x} = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} + c \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$. Neem dus bijvoorbeeld $\mathbf{b}_3 = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}$.

Dan is $\mathbf{b}_3 \to \mathbf{b}_1 \to \mathbf{b}_1 \to \mathbf{0}$ de gezochte cykel.

Conclusie: als $P = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 0 \\ -1 & 0 & 3 \end{bmatrix}$ dan $A = PJ_2^3P^{-1}$.

Voorbeeld 9.9. (Zie § 7.8, opgave 18 uit Boyce, duPrima.)

Beschouw de matrix $A = \begin{bmatrix} 5 & -3 & -2 \\ 8 & -5 & -4 \\ -4 & 3 & 3 \end{bmatrix}$. Enig rekenwerk laat zien dat $p_A(\lambda) = (1-\lambda)^3$. De enige eigenwaarden van A is dus $\lambda = 1$ en enig rekenwerk laat zien dat $\dim(E_1) = 2$ en dat $\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -3 \end{bmatrix} \right\}$ een basis voor de eigenruimte is. Tot zover niets nieuws.

Nu onze nieuwe kennis. Omdat $\dim(E_1) = 2$ zullen er dus twee cykels zijn. Op dot moment zien we dat de Jordannormaalvorm uniek bepaald moet zijn, namelijk

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

Voor de cykel ter lengte twee hebben we dus een eigenvector met oerbeeld onder $A - I$ nodig. Echter, noch $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, noch $\begin{bmatrix} 0 \\ 2 \\ -3 \end{bmatrix}$ blijken zo'n oerbeeld te hebben. (Deze kunnen dus allebei als cykel ter lengte één gebruikt worden.) Een eigenvector met oerbeeld onder $A-I$ zal er uit moeten zien als $\mathbf{v} = c_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 2 \\ -3 \end{bmatrix}$. We moeten dus nagaan voor welke c_1, c_2 de vergelijking $(A-I)\mathbf{x} = \mathbf{v}$ consistent is. Omdat:

$$\left[\begin{array}{rrr|l} 4 & -3 & -2 & c_1 \\ 8 & -6 & -4 & 2c_2 \\ -4 & 3 & 2 & 2c_1 - 3c_2 \end{array}\right] \sim \left[\begin{array}{rrr|l} 4 & -3 & -2 & c_1 \\ 0 & 0 & 0 & 2c_2 - 2c_1 \\ 0 & 0 & 0 & 3c_1 - 3c_2 \end{array}\right]$$

zien we dat dit het geval is als $c_1 = c_2$. Nemen we $c_1 = c_2 = 2$ dan krijgen we $\mathbf{b}_1^2 = \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix}$.

Oplossen $(A-I)\mathbf{x} = \mathbf{b}_1^2$ levert $\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} + d_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + d_2 \begin{bmatrix} 0 \\ 2 \\ -3 \end{bmatrix}$.

Neem bijvoorbeeld $\mathbf{b}_2^2 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$.

Conclusie: als $P = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 4 & 0 \\ 2 & -2 & -1 \end{bmatrix}$ dan: $A = PJP^{-1}$.

In het algemeen kunnen we als in bovenstaande voorbeelden een matrix naar Jordannormaalvorm brengen. Dus:
Stap 1: Bepaal alle eigenwaarden λ met hun algebraïsche multpliciteit $a.m.(\lambda)$.
Stap 2: Bepaal alle meetkundige multipliciteiten $m.m.(\lambda)$.
Stap 3.

1. Als $m.m.(\lambda) = a.m.(\lambda)$: geen probleem.

2. Als $m.m.(\lambda) = 1$ geen probleem, er is 1 cykel te vinden als in voorbeeld 9.8.

3. Als $1 < m.m.(\lambda) < a.m.(\lambda)$ is i.h.a. het lastige geval. Als bijvoorbeeld $m.m.(\lambda) = 3$ en $a.m.(\lambda) = 5$ dan zijn er 3 cykels op te speuren die tesamen een 5×5 matrix vullen. Het is altijd goed om je eerst voor te stellen waar je naar moet gaan zoeken. Hier hebben we de keuze tussen:
   of twee cykels ter lengte één en één cykel ter lengte drie
   of één cykel ter lengte één en twee cykels ter lengte twee.
   Voorzichtig speuren als in voorbeeld 9.9 volgt.