SYLLABUS

LINEAIRE ALGEBRA 2

****

R.J.Kooman

Universiteit Leiden

najaar 2007

INHOUDSOPGAVE

I. Algemene begrippen.

Vectorruimten 1
Lineaire deelruimte, lineaire onafhankelijkheid, basis 2
Lineaire afbeeldingen 4
Lineaire afbeeldingen en matrices 7
Basistransformaties 8
Directe som en projectie 9
Quotiëntverzamelingen en quotiëntruimte 11
Restrictieafbeelding en quotiëntafbeelding 12
Het tensorproduct van vectorruimten 13

II. Determinant en spoor.

De determinant van een matrix 15
Permutaties 15
Eigenschappen van determinanten 16
De Wronskiaan 17
De determinant van Vandermonde 18
Ontwikkeling van de determinant naar een kolom 19
Het spoor van een matrix 20
Het volume van een k-blok in $\mathbf{R}^n$ 20
De afstand van een punt tot een k-blok in $\mathbf{R}^n$ 21

III. Spectraaltheorie van endomorfismen in eindig-dimensionale complexe vectorruimten.

Eigenwaarden en eigenvectoren 23
Diagonaliseerbare en nilpotente afbeeldingen 24
Gegeneraliseerde eigenvectoren. Een meetkundige interpretatie van de algebraïsche multipliciteit. 25
De Jordan-normaalvorm. 27
Gelijkvormige matrices. 30
Minimumpolynoom. De stelling van Cayley-Hamilton 31
Gemeenschappelijke eigenvectoren van commuterende endomorfismen 32
De cirkels van Gershgorin 32
Appendix 33

IV. Inwendige producten op vectorruimten.
Inproducten op reële vectorruimten 35
Inproducten op complexe vectorruimten 36
Norm en afstand. De ongelijkheid van Schwarz 37
De methode van Gram-Schmidt en QR-decompositie van een matrix 38
Representatie t.o.v. een orthonormale basis 39
De geadjungeerde van een lineaire afbeelding 40
Orthogonaal complement en orthogonale projectie 41
De matrix van een orthogonale projectie 43
Toepassing: de methode van kleinste kwadraten 44
Unitaire en orthogonale afbeeldingen. 44

V. De duale van een vectorruimte. 49
De getransponeerde afbeelding 50
De annihilator van een lineaire deelruimte 50
Duale vectorruimte en tensorproduct 50

VI. Genormeerde vectorruimten.
De norm van een lineaire afbeelding. 51
Banach- en Hilbertruimten. Convergente rijen van lineaire afbeeldingen. 53
De e-macht van een matrix. 54
Vector- en matrixwaardige functies 55
Toepassing: stelsels van lineaire differentiaalvergelijkingen 56

VII. Spectraaltheorie van normale afbeeldingen.
Normale afbeeldingen 58
Symmetrische matrices 60
Kwadratische vormen op $\mathbf{R}^n$ 61
Rayleighquotiënt en minimaxprincipe 62

VIII. Positief-definiete matrices. 64
De polaire decompositie 66
De singuliere-waardendecompositie van een matrix. 66

Index. 69

I. ALGEMENE BEGRIPPEN.

Vectorruimten.

Definitie: Een *vectorruimte* over K (met $K = \mathbf{R}$ of $\mathbf{C}$) is een niet-lege verzameling V met twee bewerkingen, een optelling en een *scalaire vermenigvuldiging*, zodanig dat de volgende eigenschappen gelden (in het onderstaande zijn $u, v, w \in V$ en $\lambda, \mu \in K$).
1. $v + w = w + v$ (commutativiteit van de optelling).
2. $(v + w) + u = v + (w + u)$ (associativiteit van de optelling).
3. er is een nulelement 0 (ook wel genoteerd als 0_V) zodanig dat $0 + v = v + 0 = v$ voor alle $v \in V$.
4. Elke $v \in V$ heeft een *inverse* $-v$, zodanig dat $v + (-v) = 0$.
5. $\lambda(v + w) = \lambda v + \lambda w$ (distributieve eigenschap 1).
6. $(\lambda + \mu)v = \lambda v + \mu v$ (distributieve eigenschap 2).
7. $(\lambda\mu)v = \lambda(\mu v)$ (associativiteit van de scalaire vermenigvuldiging).
8. $1 \cdot v = v$

K heet het *lichaam* van de scalairen. Als $K = \mathbf{R}$ dan noemen we V een reële vectorruimte, als $K = \mathbf{C}$, dan heet V een complexe vectorruimte. (Ook andere lichamen K kunnen optreden als lichaam van scalairen. We beperken ons in dit college tot $K = \mathbf{R}$ resp. $\mathbf{C}$ en gaan niet verder op het begrip lichaam in.)

Voorbeelden van vectorruimten: 1. $V = \mathbf{R}^n$ bestaande uit geordende rijtjes $(x_1, x_2, \ldots, x_n)$, *vectoren* genoemd, met $x_i \in \mathbf{R}$, en de componentsgewijze optelling en scalaire vermenigvuldiging, is een vectorruimte over $\mathbf{R}$. In plaats van rijtjes schrijven we vectoren in $\mathbf{R}^n$ in de regel als kolomvectoren, dus $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Voor deze kolomvector schrijven we ook wel $(x_1, x_2, \ldots, x_n)^T$, waarbij ^{T} staat voor *transponeren*, d.w.z. rijen en kolommen in een matrix omwisselen.
2. Geheel analoog is $V = \mathbf{C}^n$ een vectorruimte over $\mathbf{C}$.
3. De vectorruimte van $m \times n$-matrices met elementen in $K = \mathbf{R}$ of $\mathbf{C}$ met de componentsgewijze optelling en scalaire vermenigvuldiging is een vectorruimte over K. We noteren deze als
$\mathcal{M}(m \times n, K)$.
4. De verzameling van de complexe getallen $a + bi$ (met $a, b \in \mathbf{R}$) vormt een reële vectorruimte.
5. De vectorruimte $P(K)$ van polynomen $a_0 + a_1X + \ldots a_nX^n$ met coëfficiënten $a_i \in K$ met de termsgewijze optelling en scalaire vermenigvuldiging is een vectorruimte over K (d.w.z. $(a_0 + a_1X + \ldots + a_nX^n) + (b_0 + b_1X + \ldots + b_nX^n) = (a_0 + b_0) + (a_1 + b_1)X + \ldots + (a_n + b_n)X^n$ en $\lambda(a_0 + a_1X + \ldots + a_nX^n) = (\lambda a_0 + \lambda a_1X + \ldots + \lambda a_nX^n)$).
6. De vectorruimte van reëel- resp. complexwaardige functies van een niet-lege verzameling X naar $\mathbf{R}$ resp. $\mathbf{C}$ (bijvoorbeeld $X = [a, b] \in \mathbf{R}$ met $a < b$). De optelling en scalaire vermenigvuldiging zijn gedefinieerd als $(f + g)(x) = f(x) + g(x)$ en $(\lambda f)(x) = \lambda f(x)$ voor $\lambda \in K$.

We zullen het woord *vector* ook in een algemene zin voor een element van een vectorruimte gebruiken. Verder geven we scalairen meestal aan met Griekse letters $\lambda, \mu, \nu, \ldots$, maar ook wel als $x_1, x_2 \ldots$, (bijvoorbeeld wanneer het de componenten van een vector $\mathbf{x}$ zijn).

De eigenschappen 1-8 definiëren een vectorruimte. Andere eigenschappen moeten we uit deze acht afleiden. Een voorbeeld is de eigenschap dat $0 \cdot v = 0_V$ voor alle $v \in V$. Dit volgt uit:

$$0 \cdot v = (0+0) \cdot v = 0 \cdot v + 0 \cdot v,$$

vanwege eigenschap 6. De inverse links en rechts optellen geeft (eigenschap 4)

$$0_V = -0 \cdot v + 0 \cdot v = -0 \cdot v + 0 \cdot v + 0 \cdot v = 0 \cdot v.$$

Opmerking: Net zo is aan te tonen: 1. Voor elke $v \in V$ geldt dat $(-1) \cdot v = -v$. 2. Zowel het nulelement als de inverse van een vector v zijn uniek bepaald.

Lineaire deelruimte, lineaire onafhankelijkheid, basis.

Zij V een vectorruimte over K. Een niet-lege deelverzameling $W \subset V$ is een *lineaire deelruimte* van V als voor elke $v, w \in W$ en $\lambda \in K$ geldt dat $v + w \in W$ en $\lambda v \in W$. Een lineaire deelruimte van een vectorruimte V is zelf dus weer een vectorruimte t.a.v. de optelling en scalaire vermenigvuldiging in V. In het bijzonder is elke vectorruimte een lineaire deelruimte van zichzelf. Elke lineaire deelruimte is het *opspansel* $\text{span}\{a_1, a_2, \ldots\}$ van een verzameling vectoren (bestaande uit alle eindige (!) *lineaire combinaties* van deze vectoren). Een stelsel vectoren $\{b_1, b_2, \ldots\}$ in een vectorruimte V heet *lineair onafhankelijk* als er geen niet-triviale eindige lineaire afhankelijkheidsrelaties zijn, m.a.w. uit $\lambda_1 b_1 + \ldots + \lambda_k b_k = 0$ volgt dat $\lambda_1 = \ldots = \lambda_k = 0$ voor elke eindige k. Als een opspannend stelsel vectoren van een lineaire deelruimte tevens lineair onafhankelijk is dan heet het stelsel een *basis* van de lineaire deelruimte. Iedere vector in de lineaire deelruimte is dan op een unieke manier te schrijven als een eindige lineaire combinatie van de opspannende vectoren.

Voorbeelden. 1. Als V een vectorruimte is met nulelement 0_V, dan zijn zowel $\{0_V\}$ als V zelf lineaire deelruimten.

2. $V = K^n$ ($K = \mathbf{R}$ of $\mathbf{C}$). Laat $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$. Iedere vector $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ is op een unieke manier te schrijven als een lineaire combinatie $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$. $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ heet de *standaardbasis* van K^n en $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ noemen we de *standaardrepresentatie* van de vector $\mathbf{x}$.

Laat $f_1, \ldots, f_k$ lineair onafhankelijke vectoren zijn. Dan bestaat het opspansel $\text{span}_K\{f_1, \ldots, f_k\}$ uit alle lineaire combinaties van de vectoren $f_1, \ldots, f_k$ met coëfficiënten in K. Dit opspansel is een lineaire deelruimte van K^n met basis $f_1, \ldots, f_k$.

3. $\mathbf{R}^n$ is geen lineaire deelruimte van $\mathbf{C}^n$.
4. $V = \mathcal{M}(n \times n, K)$. Een matrix A heet *symmetrisch*, resp. *antisymmetrisch* als $A^T = A$ resp. $A^T = -A$. De symmetrische $n \times n$-matrices vormen een lineaire deelruimte $\mathrm{Sym}^n(K)$. Evenzo vormen de antisymmetrische $n \times n$-matrices een lineaire deelruimte $\mathrm{Ant}^n(K)$. Andere voorbeelden van lineaire deelruimten zijn de verzameling diagonaalmatrices, de verzameling bovendriehoeksmatrices (of onderdriehoeksmatrices). De verzameling van inverteerbare matrices is daarentegen geen lineaire deelruimte van V.
5. Laat $V = P(K)$, de vectorruimte van polynomen met coëfficiënten in K (K is uiteraard weer $\mathbf{R}$ of $\mathbf{C}$). $P(K)$ is het opspansel van $1, X, X^2, \ldots$. Een lineaire deelruimte wordt gevormd door de verzameling polynomen $P_n(K)$ van graad hoogstens n. Dit is $\mathrm{span}\{1, X, \ldots, X^n\}$. Een ander voorbeeld van een lineaire deelruimte wordt gevormd door de verzameling polynomen in $P(K)$ zodat $P(2) = 0$. Deze laatste lineaire deelruimte heeft als basis $\{X - 2, (X - 2)^2, (X - 2)^3, \ldots\}$. De polynomen van graad precies n vormen geen lineaire deelruimte.
6. V is de vectorruimte van K-waardige functies van $[a, b]$ naar K. Een lineaire deelruimte wordt gevormd door de continue functies van $[a, b]$ naar K, notatie $C([a, b], K)$. De (K-waardige) differentieerbare functies op $[a, b]$ vormen een lineaire deelruimte van de laatste en dus ook van V. Bases van deze lineaire deelruimte zijn niet echter aan te geven.

Merk op dat in een vectorruimte alleen eindige lineaire combinaties gedefinieerd zijn, ook als er oneindig veel elementen in een basis zitten. Beschouw als voorbeeld de vectorruimte $P(K)$ van polynomen over K, die als basis $1, X, X^2, X^3 \ldots$ heeft. Oneindige lineaire combinaties, van $1, X, X^2, \ldots$, de *formele machtreeksen*, liggen niet in $P(K)$.

In een vectorruimte V zijn er oneindig veel mogelijkheden om een basis te kiezen (behalve voor $V = \{0\}$, dat geen basis heeft). Elk lineair onafhankelijk stelsel dat de vectorruimte opspant voldoet. Wat wel onafhankelijk van de keuze van de basis is, is het aantal vectoren waaruit een basis bestaat. Dit zullen we nu aantonen.

Propositie 1.1: Een lineair onafhankelijk stelsel in V heeft nooit meer vectoren dan een basis van V.

Bewijs: We geven het bewijs voor het geval dat V een basis heeft bestaande uit eindig veel vectoren. Laat $\{v_1, \ldots, v_n\}$ een basis van V zijn en $\{w_1, \ldots, w_m\}$ een lineair onafhankelijk stelsel. We laten zien dat elke vector w_i uit het lineaire onafhankelijke stelsel kan worden verwisseld met een basisvector v_j zo, dat het stelsel lineair onafhankelijk blijft. Stel nl. dat dit voor (zeg) w_1 niet het geval is. Dan is het stelsel $\{v_j, w_2, \ldots, w_m\}$ lineair afhankelijk voor elke basisvector v_j. Omdat $w_2, \ldots, w_m$ volgens de aanname lineair onafhankelijk zijn, is elke v_j een lineaire combinatie van $w_2, \ldots, w_m$. Daar de v_j's een basis vormen is w_1 een lineaire combinatie van de v_j's en dus van $w_2, \ldots, w_m$. Maar dan zijn $w_1, w_2, \ldots, w_m$ lineair afhankelijk, in tegenspraak met de aanname. Conclusie: er is een v_j, noem deze v_{j_1}, zodat $\{v_{j_1}, w_2, \ldots, w_m\}$ lineair onafhankelijk is. Op dezelfde wijze kunnen we w_2 uitwisselen tegen, zeg v_{j_2}, zodat $\{v_{j_1}, v_{j_2}, w_3, \ldots, w_m\}$ lineair onafhankelijk is.

Zo verdergaand vinden we tenslotte m lineair onafhankelijke vectoren $v_{j_1}, \ldots, v_{j_m}$. In het bijzonder zijn deze verschillend, dus $m \leq n$. $\diamond$

Gevolgen 1.2: a. Als een vectorruimte V hoogstens eindig veel lineair onafhankeijke vectoren bevat dan heeft elke basis van V evenveel vectoren. Dit aantal noemen we de *dimensie* van V (notatie: $\dim(V)$). Als er oneindig veel lineair onafhankelijke vectoren zijn, dan is de dimensie van V oneindig. (Ook in dit geval kunnen we verschillende soorten van oneindig onderscheiden. Als er een oneindige rij $\{f_1, f_2, \ldots\}$ lineair onafhankelijke vectoren bestaat die V opspannen dan zeggen we dat de dimensie van V *aftelbaar oneindig* is.)
b. Zij V een vectorruimte van eindige dimensie n. Een lineair onafhankelijk stelsel in V dat n vectoren bevat is een basis. M.a.w., een basis is een maximaal lineair onafhankelijk stelsel.

Bewijs: Stel nl. dat $\{a_1, \ldots, a_n\}$ een lineair onafhankelijk stelsel is dat V niet opspant. Dan is er een vector $a_{n+1} \in V$ die lineair onafhankelijk is van $a_1, \ldots, a_n$. Het lineair onafhankelijke stelsel $\{a_1, \ldots, a_{n+1}\}$ bevat dan meer vectoren dan een basis. Dit is in tegenspraak met Propositie 1.1. $\diamond$
c. Zij V een vectorruimte van eindige dimensie n. Een opspannend stelsel dat uit precies n vectoren bestaat is een basis.
d. Zij V een vectorruimte en W een lineaire deelruimte. Dan is $\dim(W) \leq \dim(V)$. Als $\dim(W) = \dim(V)$ en V heeft eindige dimensie, dan is $W = V$.
e. De dimensie van de reële, resp. complexe vectorruimten $\mathbf{R}^n$, resp. $\mathbf{C}^n$ is n. De vectorruimte van polynomen $P(K)$ heeft dimensie ∞.
f. $\mathbf{C}^n = \mathrm{span}_{\mathbf{C}}\{e_1, e_2, \ldots, e_n\}$ is op te vatten als een reële vectorruimte van dimensie $2n$, nl. als $\mathrm{span}_{\mathbf{R}}\{e_1, ie_1, \ldots, e_n, ie_n\}$.

Opgave: Laat V een vectorruimte zijn van eindige dimensie n en laat $\{a_1, \ldots, a_k\}$ een lineair onafhankelijk stelsel zijn. Laat zien dat er vectoren $a_{k+1}, \ldots, a_n \in V$ zijn zodanig dat $\{a_1, \ldots, a_n\}$ een basis is van V, m.a.w. elk lineair onafhankelijk stelsel is aan te vullen tot een basis.

Opgave: Toon aan dat de kleinste lineaire deelruimte van de vectorruimte $\mathbf{C}^n$ die $\mathbf{R}^n = \mathrm{span}_{\mathbf{R}}\{e_1, e_2, \ldots, e_n\}$ bevat, $\mathbf{C}^n$ zelf is.

Lineaire afbeeldingen.

We beschouwen nu afbeeldingen tussen vectorruimten. Een centrale rol wordt gespeeld door afbeeldingen die de vectorruimtestructuur behouden:

Definitie: Laat V, W vectorruimten zijn over K ($K = \mathbf{R}$ of $\mathbf{C}$). De afbeelding $T : V \to W$ heet *lineair* als voor elke $v, v' \in V$ en $\lambda \in K$ geldt dat: 1. $T(v+v') = T(v)+T(v')$ en 2. $T(\lambda v) = \lambda T(v)$.

Voorbeelden:

1. $T : K^n \to K^m$ gegeven door $T(\mathbf{x}) = A\mathbf{x}$ met A een $m \times n$-matrix is een lineaire afbeelding. Omgekeerd zullen we zien dat elke lineaire afbeelding van K^n naar K^m van de vorm $T(\mathbf{x}) = A\mathbf{x}$ is met A een $m \times n$-matrix.

2. Zij $C = C([a,b], K)$ de vectorruimte van K-waardige continue functies op het interval $[a,b] \subset \mathbf{R}$. De afbeelding $T : C \to K$ gegeven door $T(f) = f(c)$ (met $c \in [a,b]$) is een lineaire afbeelding.
3. Laat C als in (2) gedefinieerd zijn. De afbeelding $T : C \to C$ gegeven door $T(f) = fg$ voor een vaste $g \in C$ is lineair.
4. $M = M(m \times n, K)$ de vectorruimte van $m \times n$-matrices met elementen in K. $T : M \to K$ gegeven door $T(A) = A_{ij}$ is lineair (A_{ij} is het element in de i-e rij en j-e kolom van A. We schrijven ook wel $A = (A_{ij})$).
5. $T : M \to M$ gegeven door $T(A) = A^T$ (transponeren) is lineair.
6. Laat V een vectorruimte zijn. De *identieke afbeelding* $id_V : V \to V$, gedefinieerd als de afbeelding die ieder element op zichzelf afbeeldt, m.a.w. $id_V(v) = v$ voor alle $v \in V$, is lineair.

Laat V en W vectorruimten over hetzelfde lichaam K zijn. Op de verzameling lineaire afbeeldingen van V naar W kunnen we de structuur van een vectorruimte leggen: als $T, U : V \to W$ lineaire afbeeldingen zijn, dan definiëren we de som $T + U$ en het scalair product λT d.m.v. $(T + U)(v) = T(v) + U(v)$ en $(\lambda T)(v) = \lambda T(v)$ waarbij $v \in V$. Het is duidelijk dat $T + U$ en λT ook lineaire afbeeldingen van V naar W zijn. De vectorruimte van lineaire afbeeldingen van V naar W geven we aan als $\mathcal{L}(V, W)$.

Opgave: Toon aan: als V en W eindige dimensie hebben dan is $\dim(\mathcal{L}(V,W)) = \dim(V)\dim(W)$. Geef ook een basis van $\mathcal{L}(V,W)$ aan.

Definitie: Een lineaire afbeelding $T : V \to V$ van een vectorruimte in zichzelf noemen we ook een *(lineair) endomorfisme*. De vectorruimte $\mathcal{L}(V,V)$ van endomorfismen van een vectorruimte V noteren we korter als als $\mathcal{L}(V)$. $\mathcal{L}(V)$ heeft een rijkere algebraïsche structuur dan een vectorruimte, doordat er naast de optelling en scalaire vermenigvuldiging ook nog d.m.v. de compositie van twee endomorfismen een vermenigvuldiging is gedefinieerd. Laat immers $T, U \in \mathcal{L}(V)$, en zij $TU : V \to V$ gedefinieerd d.m.v. $TU(v) = T(U(v))$ voor $v \in V$. Het is eenvoudig om na te gaan dat TU weer lineair is (doe dit!). Voor de vermenigvuldiging gelden de extra eigenschappen:

i. $S(T + U) = ST + SU$, $(T + U)S = TS + US$ (distributiviteit).
ii. $S(TU) = (ST)U$ (associativiteit).
iii. $\lambda TU = (\lambda T)U = T(\lambda U)$ voor $\lambda \in K$.

Een vectorruimte met een vermenigvuldiging zodanig dat eigenschappen (i-iii) gelden, noemen we een *algebra.* Merk op dat de vermenigvuldiging niet commutatief hoeft te zijn. De algebra $\mathcal{L}(V)$ heeft ook een eenheidselement $I = id_V$, zodat $IT = TI = T$ voor alle $t \in V$. Niet elke algebra heeft een eenheidselement. Een ander voorbeeld van een algebra is de vectorruimte $\mathcal{M}(n \times n, K)$ van $n \times n$-matrices met elementen in K, met de gewone matrixoptelling en -vermenigvuldiging. Deelalgebra's hiervan zijn de algebra's bestaande uit de $n \times n$-bovendriehoeksmatrices en de $n \times n$-stricte bovendriehoeksmatrices (met nullen op de hoofddiagonaal). De laatste algebra heeft geen eenheidselement.

We voeren nu een aantal begrippen in die van belang zijn bij de bestudering van afbeeldingen.

Definities:

i. Zij $T : V \to W$ een lineaire afbeelding: Als $U \subset V$ dan heet $T(U) = \{v \in W : v = T(u)$ voor zekere $u \in U\}$ het *beeld* van U; $T(U)$ is dus de verzameling van beelden van elementen uit U onder T. Als U een lineaire deelruimte van V is, dan is $T(U)$ een lineaire deelruimte van W. Als $U = V$, dan schrijven we i.p.v. $T(V)$ ook wel $\mathrm{im}(T)$. $\mathrm{im}(T)$ heet het *bereik* van T. De dimensie van $\mathrm{im}(T)$ heet de *rang* van T; als V, W eindige dimensie hebben, dan is de rang van T gelijk aan de rang van een matrix van T. T heet *surjectief* (of: *op*) als $T(V) = W$.

ii. Als $Z \subset W$ dan heet de verzameling $T^{-1}(Z) = \{v \in V : T(v) \in Z\}$ van T het *inverse beeld* van Z. $T^{-1}(Z)$ is de verzameling originelen van elementen van Z onder T. Als Z een lineaire deelruimte is van W, dan is het inverse beeld een lineaire deelruimte van V. Voor het inverse beeld van $\{0_W\}$ (de *kern* of *nulruimte* van T) schrijven we ook wel $\ker(T)$ of $T^{-1}(0)$. T heet *injectief* (of *1-1*) als elke $w \in W$ hoogstens één origineel heeft. Een lineaire afbeelding $T : V \to W$ is injectief dan en slechts dan als de nulruimte alleen uit het nulelement van V bestaat.

iii. Als T injectief en surjectief is, dan heet T *bijectief* of *inverteerbaar*. In dit geval kunnen we een *inverse afbeelding* $T^{-1} : W \to V$ definiëren zodanig dat $T \circ T^{-1} = id_W$ en $T^{-1} \circ T = id_V$, m.a.w. $T^{-1}(w) = v$ dan en slechts dan als $T(v) = w$. Een inverteerbare lineaire afbeelding $T : V \to W$ noemen we ook wel een *vectorruimte-isomorfisme*. V en W heten in dat geval *isomorfe* vectorruimten.

Voorbeeld: De afbeelding $T : \mathbf{R}^3 \to \mathbf{R}^2$ wordt gegeven door $T(\mathbf{x}) = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 3 \end{pmatrix} \mathbf{x}$. T heeft rang 2 en is dus surjectief: $T(\mathbf{R}^3) = \mathbf{R}^2$.

De afbeelding $U : \mathbf{R}^2 \to \mathbf{R}^3$ wordt gegeven door $U(\mathbf{x}) = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 3 \end{pmatrix} \mathbf{x}$. $\mathrm{Ker}(U)$ bestaat alleen uit de nulvector $\mathbf{0}$. U is dus injectief.

Voor inverteerbare lineaire afbeeldingen geldt:

Propositie 1.3: De inverse van een inverteerbare lineaire afbeelding is lineair.

Bewijs: Laat $T : V \to W$ een lineaire afbeelding zijn tussen vectorruimten V en W. Laat $w_1, w_2 \in W$. Daar T inverteerbaar is, zijn er $v_1, v_2 \in V$ zodanig dat $T(v_1) = w_1$ en $T(v_2) = w_2$. Dan is

$$T^{-1}(w_1 + w_2) = T^{-1}(T(v_1) + T(v_2)) = T^{-1}(T(v_1 + v_2)) = v_1 + v_2 = T^{-1}(w_1) + T^{-1}(w_2).$$

Verder is

$$T^{-1}(\lambda w_1) = T^{-1}(\lambda T(v_1)) = T^{-1}(T(\lambda v_1)) = \lambda v_1 = \lambda T^{-1}(w_1). \quad \diamond$$

Opgave: Toon aan: als $S : V \to W$ en $T : W \to Z$ inverteerbare lineaire afbeeldingen zijn, dan is $TS : V \to Z$ inverteerbaar en $(TS)^{-1} = S^{-1}T^{-1}$.

Voorbeeld: De vectorruimte $P_n(K)$ van polynomen van graad hoogstens n is isomorf met de vectorruimte K^{n+1}. Een vectorruimte-isomorfisme is de afbeelding $T : P_n(K) \to K^{n+1}$ gedefinieerd door

$$T(x_0 + x_1 X + \ldots + x_n X^n) = (x_0, x_1, \ldots, x_n)^T.$$

Het voorafgaande voorbeeld kunnen we direct generaliseren naar het geval van twee vectorruimten van dezelfde dimensie: laat V, W vectorruimten over K van dezelfde dimensie n zijn. Dan zijn V en W isomorf. Immers laat $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ een basis van V en $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ een basis van W zijn. Laat de lineaire afbeelding $T : V \to W$ gegeven zijn door $T(\mathbf{a}_j) = \mathbf{b}_j$ (zoals we weten is T geheel bepaald door de beelden van de basisvectoren). Het is nu duidelijk dat T een vectorruimte-isomorfisme is.

Gevolg 1.4: Twee vectorruimten met hetzelfde lichaam K van scalairen en dezelfde eindige dimensie zijn isomorf.

Een speciaal geval van het bovenstaande krijgen we door $W = K^n$ en $\mathcal{B}$ de standaardbasis te nemen. Zij $v = v_1\mathbf{a}_1 + \ldots + v_n\mathbf{a}_n$; dan wordt het vectorruimte-isomorfisme $T : V \to K^n$ gegeven door $T(v) = (v_1, \ldots, v_n)^T$. T heet de *coördinaatafbeelding* (m.b.t. de basis $\mathcal{A}$). We noteren hiervoor $T = B_{\mathcal{A}}$. Voor de *coördinaatvector* $B_{\mathcal{A}}(v)$ noteren we ook $v_{\mathcal{A}}$.

Voor een lineaire afbeelding bestaat het volgende verband tussen zijn rang en de dimensie van de kern:

Propositie 1.5 (*dimensiestelling*): Zij $T : V \to W$ lineair. Dan geldt:

$$\dim\ \ker(T) \ +\ \mathrm{rang}(T) \ =\ \dim(V). \tag{1.1}$$

Bewijs: zij $\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$ een basis van $\ker(T)$ en vul deze aan tot een basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ van V. Dan wordt $\mathrm{im}(T)$ opgespannen door $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$ en dus door $T(\mathbf{b}_{m+1}), \ldots, T(\mathbf{b}_n)$ omdat $T(\mathbf{b}_1) = \ldots = T(\mathbf{b}_m) = 0$. We tonen nu aan dat het stelsel $\{T(\mathbf{b}_{m+1}, \ldots, T(\mathbf{b}_n)\}$ lineair onafhankelijk en dus een basis van $\mathrm{im}(T)$ is. Neem aan dat $\lambda_{m+1}T(\mathbf{b}_{m+1}) + \ldots + \lambda_n T(\mathbf{b}_n) = 0$. Dan is $T(\lambda_{m+1}\mathbf{b}_{m+1} + \ldots + \lambda_n\mathbf{b}_n) = 0$ dus de vector $\lambda_{m+1}\mathbf{b}_{m+1} + \ldots + \lambda_n\mathbf{b}_n$ ligt in $\ker(T)$. Maar omdat $\mathbf{b}_{m+1}, \ldots, \mathbf{b}_n$ lineair onafhankelijk van de vectoren $\mathbf{b}_1, \ldots, \mathbf{b}_m$ gekozen zijn, is $\lambda_{m+1} = \ldots = \lambda_n = 0$. Conclusie: $\dim\ \mathrm{im}(T) = n - m = \dim(V) - \dim\ \ker(T)$. $\diamond$

Lineaire afbeeldingen en matrices.

Een lineaire afbeelding $T : V \to W$ wordt geheel bepaald door de beelden van de basisvectoren. Immers als $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ een basis is van V (n kan eventueel oneindig zijn), dan is elke $v \in V$ een eindige lineaire combinatie $\lambda_1\mathbf{b}_1 + \ldots \lambda_k\mathbf{b}_k$ en dus is $T(v) = \lambda_1 T(\mathbf{b}_1) + \ldots + \lambda_k T(\mathbf{b}_k)$. Als V, W eindig-dimensionaal zijn, dan kan T d.m.v. een matrix worden gerepresenteerd. Dit gaat op de volgende manier: laat $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ een basis van V zijn en $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_m\}$ een basis van W. Dan zijn er getallen $A_{ij} \in K$ zodanig dat $T(\mathbf{b}_j) = \sum_{i=1}^m A_{ij}\mathbf{c}_i$. Laat A de matrix (A_{ij}) zijn, m.a.w. A_{ij} is het element in de i-e rij en j-e kolom van A. Dan bevat de j-e kolomvector de coëfficiënten van $T(\mathbf{b}_j)$ t.o.v. de basis $\mathcal{C}$. Laat nu $v = v_1\mathbf{b}_1 + \dots v_k\mathbf{b}_k$ zijn. De kolomvector $T(v)_\mathcal{C}$ van coëfficiënten van $T(v)$ t.o.v. de basis $\mathcal{C}$ is nu het matrixproduct van de matrix A met de kolomvector $v_\mathcal{B} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$. De matrix A noemen we de *matrix van* T *t.o.v. de bases* $\mathcal{B}$ en $\mathcal{C}$. We noteren deze matrix ook wel als $T_\mathcal{C}^\mathcal{B}$. We kunnen dit resultaat kort schrijven als

$$T(v)_\mathcal{C} = T_\mathcal{C}^\mathcal{B}(v_\mathcal{B}).$$

Als V eindig-dimensionaal is met dimensie n, dan is voor elke basis $\mathcal{B}$ van V de matrix $(id_V)_\mathcal{B}^\mathcal{B}$ de eenheidsmatrix I_n.

In het geval dat $V = K^n$ en $W = K^m$, kunnen we voor $\mathcal{B}$ en $\mathcal{C}$ de standaardbases van V en W kiezen. In dit geval is $v_\mathcal{B}$, resp $v_\mathcal{C}$ de standaardrepresentatie van v in V resp. W en we schrijven dan v voor $v_\mathcal{B}$ resp. $v_\mathcal{C}$. Dan is $T(v) = Av$ met $A = T_\mathcal{C}^\mathcal{B}$. Een lineaire afbeelding van K^n naar K^m is dus (in de standaardrepresentatie) van de vorm $v \to Av$ voor een $m \times n$-matrix A. A heet de *standaardmatrix* van de afbeelding T.

Opmerking: Als V en W eindig-dimensionaal zijn, en $A = T_\mathcal{C}^\mathcal{B}$ is de matrix van $T \in \mathcal{L}(V, W)$ t.o.v. zekere bases $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ en $\mathcal{C}$ van V resp. W, dan is $n = \dim(V)$ het aantal kolommen van A; verder is $v = \lambda_1\mathbf{b}_1 + \dots + \lambda_n\mathbf{b}_n \in \ker(T)$ dan en slechts dan als $A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = 0$ dus is $\dim(\ker(T)) = \dim(\ker(A))$. Dan geldt, m.b.v. Propositie 1.3,

$$\operatorname{rang}(A) := n - \dim(\ker(A)) = \dim(V) - \dim(\ker(T)) = \operatorname{rang}(T).$$

Hierbij is de rang van de matrix A gelijk aan het aantal lineair onafhankelijke kolomvectoren van de matrix A.

Basistransformaties.

Laat V een eindig-dimensionale vectorruimte zijn en $T : V \to V$ een lineaire afbeelding. We onderzoeken hoe de matrix van de afbeelding T verandert als we overgaan op een andere basis van V. Laat $\mathcal{A} := \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ en $\mathcal{C} := \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ twee bases zijn van V. Voor $v \in V$ zijn de coördinaatafbeeldingen $B_\mathcal{A}, B_\mathcal{C} : V \to K^n$ gedefinieerd door $B_\mathcal{A}(v) = v_\mathcal{A}$, $B_\mathcal{C}(v) = v_\mathcal{C}$ met $v_\mathcal{A} = (v_1, v_2, \dots, v_n)^T$ waarbij $v_1, \dots, v_n$ de coördinaten van v t.o.v. de basis $\mathcal{A}$ zijn. De afbeelding $B_\mathcal{C}^\mathcal{A} : K^n \to K^n$ gegeven door $B_\mathcal{C}^\mathcal{A}(v_\mathcal{A}) = v_\mathcal{C}$ is lineair en inverteerbaar, nl. $B_\mathcal{C}^\mathcal{A} = (B_\mathcal{A}^\mathcal{C})^{-1}$. De matrix van deze afbeelding noemen we ook $B_\mathcal{C}^\mathcal{A}$. Merk op dat de kolomvectoren van $B_\mathcal{C}^\mathcal{A}$ gelijk zijn aan $(\mathbf{a}_1)_\mathcal{C}, \dots, (\mathbf{a}_n)_\mathcal{C}$. Laat nu $T : V \to V$ een lineaire afbeelding zijn. De matrix $T_\mathcal{A}^\mathcal{A}$ van T t.o.v. de basis $\mathcal{A}$ is gedefinieerd als $T_\mathcal{A}^\mathcal{A}(v_\mathcal{A}) = (T(v))_\mathcal{A}$ (en analoog voor $T_\mathcal{C}^\mathcal{C}$). Nu geldt:

$$T_\mathcal{C}^\mathcal{C} = B_\mathcal{C}^\mathcal{A} T_\mathcal{A}^\mathcal{A} B_\mathcal{A}^\mathcal{C} = B_\mathcal{C}^\mathcal{A} T_\mathcal{A}^\mathcal{A} (B_\mathcal{C}^\mathcal{A})^{-1}. \tag{1.2}$$

In het bijzonder zien we dat twee matrices van dezelfde lineaire afbeelding t.o.v. verschillende basis gelijkvormig zijn. (Matrices A, B heten *gelijkvormig* als $B = U^{-1}AU$ voor zekere inverteerbare matrix U.)

Voorbeeld. Laat $V = P_1(\mathbf{R}) = \text{span}\{1, X\}$, de vectorruimte van reële polynomen van graad hoogstens 1. Laat $\mathcal{A} = \{1, X\}$ en $\mathcal{C} = \{1 + X, X\}$ twee bases van V zijn. Dan zijn de basistransformatiematrices

$$B_{\mathcal{A}}^{\mathcal{C}} = \big((1+X)_{\mathcal{A}}, X_{\mathcal{A}}\big) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{en } B_{\mathcal{C}}^{\mathcal{A}} = (B_{\mathcal{A}}^{\mathcal{C}})^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Het polynoom $p(X) = 2X + 3$ heeft coëfficiënten $p_{\mathcal{A}} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ en

$$p_{\mathcal{C}} = B_{\mathcal{C}}^{\mathcal{A}} p_{\mathcal{A}} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Inderdaad is $p(X) = 3(1 + X) - X$.

Laat verder $T : V \to V$ de lineaire afbeelding zijn die wordt gegeven door $T(p) = -2p(x) + xp'(x)$. De matrix van T t.o.v. de basis $\mathcal{A} = \{1, X\}$ is $T_{\mathcal{A}}^{\mathcal{A}} = \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}$. De matrix van T t.o.v. de basis $\mathcal{C}$ is nu

$$T_{\mathcal{C}}^{\mathcal{C}} = B_{\mathcal{C}}^{\mathcal{A}} T_{\mathcal{A}}^{\mathcal{A}} B_{\mathcal{A}}^{\mathcal{C}} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 1 & -1 \end{pmatrix}.$$

Inderdaad zien we dat $T(1 + X) = -2 - X = -2(1 + X) + X$ en $T(X) = -X = 0(1 + X) - X$.

Directe som en projectie.

Laat V een vectoruimte en U, W lineaire deelruimten van V zijn. De *som* $U + W$ van U en W bestaat uit alle lineare combinaties $u + w$ met $u \in U$ en $w \in W$. Het is de kleinste lineaire deelruimte van V die U en W omvat. Als $U \cap W \neq \{0_V\}$ dan is de schrijfwijze niet uniek (vergelijk het geval dat $V = \mathbf{R}^3$ en U, W twee snijdende vlakken door de oorsprong zijn). Als $U \cap W = \{0_V\}$ dan is elke $v \in U + W$ op precies één manier te schrijven als som $u + w$ met $u \in U$ en $w \in W$. Immers neem aan dat $u + w = u' + w'$ voor $u, u' \in U$ en $w, w' \in W$. Dan is $u - u' = w' - w \in U \cap W$, dus $u = u'$ en $w = w'$. De som heet dan de *directe som*: $U \oplus W$.

Analoog is een vectorruimte V de directe som $V = U_1 \oplus \ldots \oplus U_k$ van lineaire deelruimten $U_1, \ldots, U_k$ als iedere $v \in V$ op een unieke manier als een lineaire combinatie $v = u_1 + \ldots + u_k$ met $u_j \in U_j$ is te schrijven. Merk op dat $U \oplus V \oplus W = (U \oplus V) \oplus W = U \oplus (V \oplus W)$.

Definitie: Zij $V = U \oplus W$ voor zekere lineaire deelruimten U en W. De afbeeldingen $\pi_U : V \to V$ en $\pi_W : V \to V$ zijn als volgt gedefinieerd: voor $v \in V$ zijn er unieke $u \in U$ en $w \in W$ zodanig dat $v = u + w$. Dan is

$$\pi_U(v) = u, \qquad \pi_W(v) = w.$$

π_U heet de *projectie op U langs W*; π_W heet de projectie op W langs U.

De projecties π_U en π_W zijn lineaire afbeeldingen en verder geldt

$$\pi_U^2 = \pi_U, \quad \pi_W = \pi_W^2$$

terwijl

$$\pi_U \pi_W = \pi_W \pi_U = 0, \qquad \pi_U + \pi_W = id_V.$$

Analoog zijn voor $V = U_1 \oplus \ldots \oplus U_k$ de projecties $\pi_{U_1}, \ldots, \pi_{U_k}$ gedefinieerd, waarbij $\pi_{U_j}(v) = u_j$ waarbij $v = u_1 + \ldots + u_k$ en $u_j \in U_j$ $(j = 1, \ldots, k)$. Ook hier geldt:

$$\pi_{U_1} + \ldots + \pi_{U_k} = id_V, \quad \pi_{U_i} = \pi_{U_i}^2, \quad \pi_{U_i}\pi_{U_j} = 0 \quad (i \neq j).$$

In het algemeen noemen we een afbeelding $P : V \to V$ een *projectie* als P lineair is en $P^2 = P$. In dit geval is $V = \ker(P) \oplus \mathrm{im}(P)$ en $P = \pi_{\mathrm{im}(P)}$. Immers voor elke $v \in V$ geldt dan $v = P(v) + (v - P(v))$ en $P(v) \in \mathrm{im}(P)$, $v - P(v) \in \ker(P)$. Dus V is de som van $\ker(P)$ en $\mathrm{im}(P)$. Om te laten zien dat de som een directe som is, nemen we $v \in \mathrm{im}(P) \cap \ker(P)$. Omdat $v \in \mathrm{im}(P)$, is $v = P(w)$ voor zekere $w \in W$. Dan is $P(v) = P^2(w) = P(w) = v$. Anderzijds is $P(v) = 0$ omdat $v \in \ker(P)$, en dus is $v = 0$. Uit het bovenstaande volgt ook dat $\mathrm{im}(P)$ bestaat uit de vectoren $v \in V$ zodat $P(v) = v$. (In termen van eigenvectoren kunnen we zeggen dat een projectie P eigenwaarden 0 en 1 heeft en $\mathrm{im}(P)$ is de eigenruimte bij eigenwaarde 1).

Opmerking: Als $P : V \to V$ een projectie is, dan is $id_V - P : V \to V$ eveneens een projectie en

$$\mathrm{im}(id_V - P) = \ker(P), \qquad \ker(id_V - P) = \mathrm{im}(P).$$

Voorbeelden: 1. Zij $V = \mathbf{R}^2$, $U = \mathrm{span}\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$ en $W = \mathrm{span}\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}\}$. $\pi_U : V \to V$ noemen we de projectie op U langs W. Er geldt dus

$$\pi_U \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \pi_U \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

We bepalen de matrix P van π_U t.o.v. de standaardbasis $\{\mathbf{e}_1, \mathbf{e}_2\}$. Laat $\mathcal{B}$ de basis $\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\}$ zijn. De matrix van π_U t.o.v. de basis $\mathcal{B}$ is dus $(\pi_U)_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, en de matrix P is dan

$$P = (\pi_U)_{\mathcal{E}}^{\mathcal{E}} = B_{\mathcal{E}}^{\mathcal{B}} (\pi_U)_{\mathcal{B}}^{\mathcal{B}} B_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

2. Zij $V = P_n(\mathbf{C})$, de vectorruimte van complexe polynomen van graad hoogstens n. Laat $U = \mathrm{span}(1, X)$ en $W = \mathrm{span}(X^2, \ldots, X^n)$. Dan is $V = U \oplus W$ en voor $p \in P_n$ is $\pi_U(p) = p(0) + p'(0)X$.

3. Laat $V = \mathcal{M}(n \times n, K)$ de vectorruimte van $n \times n$-matrices zijn met coëfficiënten in het lichaam K. U is de lineaire deelruimte van symmetrische matrices, W is de lineaire deelruimte

van antisymmetrische matrices. Dan is $V = U \oplus W$ en voor $A \in V$ is $\pi_U(A) = (A + A^T)/2$, $\pi_W(A) = (A - A^T)/2$.

Quotiëntverzamelingen en quotiëntruimten.

Een *equivalentierelatie* $\sim$ op een verzameling U is een relatie waarvoor de volgende drie eigenschappen gelden:

1. $v \sim v$ voor alle $v \in V$ (reflexiviteit)
2. Als $v \sim w$ dan $w \sim v$ (symmetrie)
3. Als $u \sim v$ en $v \sim w$ dan is $u \sim w$ (transitiviteit)

Voorbeelden van een equivalentierelaties zijn:

1. Laat V een niet-lege verzameling zijn. Voor $a, b \in V$ laat $a \sim b$ dan en slechts dan als $a = b$. Dit is een equivalentierelatie.
2. Laat $U = \mathbf{Z}$ en N een positief geheel getal. $a \sim b$ voor $a, b \in U$ als $a - b$ deelbaar is door N. We schrijven dit als $a \equiv b \bmod N$.
3. $U = \mathcal{M}(n \times n, K)$, de verzameling van $n \times n$-matrices. $A \sim B$ (voor $A, B \in U$) als A en B gelijkvormige matrices zijn, d.w.z. $B = C^{-1}AC$ voor een zekere matrix $C \in U$.
4. Laat V een vectorruimte zijn. Voor $a, b \in V$ laat $a \sim b$ als a, b lineair afhankelijk zijn en $a, b \neq 0_V$ en verder $0_V \sim 0_V$. $\sim$ is een equivalentierelatie.

Als U een verzameling is met een equivalentierelatie $\sim$ dan kunnen we U verdelen in *equivalentieklassen*, zodanig dat in een equivalentieklasse alle elementen van U zitten die equivalent aan elkaar zijn. Zo'n equivalentieklasse noteren we als $\bar{a}$: in de klasse $\bar{a}$ zitten alle elementen van U die equivalent zijn met a. Als $a \sim b$ dan is dus $\bar{a} = \bar{b}$. Het element a heet een *representant* van de equivalentieklasse $\bar{a}$.

In het geval van voorbeeld 2 zijn er N equivalentieklassen $\bar{0}, \bar{1}, \ldots, \overline{N-1}$. De equivalentieklasse $\bar{k}$ bestaat uit alle getallen die gelijk zijn aan k mod N, d.w.z. de getallen van de vorm $k + mN$ voor $m \in \mathbf{Z}$.
In het geval van voorbeeld 4 zijn de equivalentieklassen de lijnen door 0_V m.u.v. 0_V zelf en verder is er de klasse die alleen uit 0_V bestaat.

De verzameling van equivalentieklassen heet een *quotiëntverzameling*. We noteren deze als $U/\sim$.
We bekijken nu het geval dat U een vectorruimte is. Er bestaat dan een equivalentierelatie $\sim$ zodanig dat de quotiëntverzameling zelf weer een vectorruimte is.
Laat V een vectorruimte over K zijn en W een lineaire deelruimte van V; de relatie "$v \sim v'$ dan en slechts dan als $v - v' \in W$" is een equivalentierelatie op V. De quotiëntverzameling noteren we in dit geval als V/W. Merk op dat W gelijk is aan de klasse $\bar{0}$. We tonen aan dat op V/W de structuur van een vectorruimte gelegd kan worden. De optelling en scalaire vermenigvuldiging van twee klassen zijn als volgt gedefinieerd: voor $v, w \in V$ en $\lambda \in K$ laat

$$\bar{v} + \bar{w} = \overline{v + w}, \quad \lambda\bar{v} = \overline{\lambda v}.$$

Er moet nog worden nagegaan dat deze optelling en scalaire vermenigvuldiging goed zijn gedefinieerd: laat v' en w' twee willekeurige representanten van de klassen $\bar{v}$, resp. $\bar{w}$. We moeten nu aantonen dat $\overline{v'+w'} = \overline{v+w}$ en dat $\overline{\lambda v'} = \overline{\lambda v}$. Maar als $v \sim v'$ en $w \sim w'$, dan is $v' - v \in W$ en $w' - w \in W$ en dus is $(v' + w') - (v + w) \in W$ en ook $\lambda v' - \lambda v \in W$. De bewerkingen zijn dus inderdaad goed gedefinieerd. Met deze optelling en scalaire vermenigvuldiging vormt V/W een vectorruimte, de *quotiëntvectorruimte* van V en W.

Propositie 1.6. Als V en W eindig-dimensionale vectorruimten zijn, dan geldt voor de dimensie van V/W de volgende identiteit:

$$\dim(V/W) = \dim(V) - \dim(W). \tag{1.3}$$

Bewijs: Beschouw de *kanonieke afbeelding* $T : V \to V/W$ gedefinieerd door $T(v) = \bar{v}$ voor V. T is een lineaire afbeelding (ga dit na); verder is T surjectief en $\ker(T) = W$. Volgens de dimensiestelling is dus

$$\dim(V) = \dim(\ker(T)) + \operatorname{rang}(T) = \dim(W) + \dim(V/W). \ \diamond$$

We kunnen quotiëntruimten gebruiken om andere gelijkheden tussen dimensies van verschillende vectorruimten af te leiden. Een voorbeeld is het volgende: laat V, W lineaire deelruimten zijn van een vectorruimte U. Dan is de som $V + W$ en ook de doorsnede $V \cap W$ een lineaire deelruimte. De volgende relatie geldt tussen de dimensies:

Propositie 1.7.

$$\dim(V + W) + \dim(V \cap W) = \dim(V) + \dim(W). \tag{1.4}$$

In het bijzonder volgt in het geval dat $V + W$ een directe som is (d.w.z. als $V \cap W = \{0\}$)

$$\dim(V \oplus W) = \dim(V) + \dim(W). \tag{1.5}$$

Bewijs: De afbeelding $T : V \to V+W/W$ die v op $(v \bmod W)$ afbeeldt, is goed gedefinieerd, lineair en surjectief. De kern van T bestaat uit alle elementen uit V die in W liggen dus $\ker(T) = V \cap W$. Uit de dimensieformules (1.2) en (1.3) volgt dus

$$\dim(V) = \dim(V + W/W) + \dim(V \cap W) = \dim(V + W) - \dim(W) + \dim(V \cap W).\diamond$$

Restrictieafbeelding en quotiëntafbeelding.

Laat V een vectorruimte zijn en W een lineaire deelruimte van V. Zij verder $T : V \to V$ een lineair endomorfisme van V dat W in zichzelf afbeeldt, d.w.z. $T(W) \subset W$. We kunnen de afbeelding T dan opvatten als een afbeelding in $\mathcal{L}(W)$. We noemen deze afbeelding de *restrictie* $T|_W$ van T op W. Er geldt dus $T|_W(w) = T(w)$ voor $w \in W$ (voor $v \notin W$ is $T|_W$ niet gedefinieerd). Verder kunnen we de *quotiëntafbeelding* $\bar{T} : V/W \to V/W$ definiëren d.m.v. $\bar{T}(\bar{v}) = \overline{T(v)}$. Ga zelf na dat $\bar{T}$ goed gedefinieerd en lineair is. Neem nu aan dat V eindige dimensie n heeft. Laat

$\mathcal{W} = \{w_1, \ldots, w_m\}$ een basis van W zijn. Vul deze aan tot een basis $\mathcal{V} = \{w_1, \ldots, w_n\}$ van V. Het stelsel vectoren $\mathcal{W}' = \{\overline{w_{m+1}}, \ldots, \overline{w_n}\}$ is nu lineair onafhankelijk in V/W en dus een basis (waarom?). We zeggen ook dat het stelsel $\{w_{m+1}, \ldots, w_n\}$ *lineair onafhankelijk modulo* W is. Merk op dat een stelsel vectoren $\{y_1, \ldots, y_k\}$ lineair onafhankelijk modulo de lineaire deelruimte W is als uit $\lambda_1 y_1 + \ldots + \lambda_k y_k \in W$ volgt dat alle λ_i nul zijn.

Opgave: Ga na dat de matrix $T_{\mathcal{V}}^{\mathcal{V}}$ van de afbeelding T t.o.v. de basis $\mathcal{V}$ van de vorm $\begin{pmatrix} A & B \\ O & C \end{pmatrix}$ is, waarbij A de matrix van $T|_W$ is t.o.v. de basis $\mathcal{W}$ van W en C de matrix van $\bar{T}$ t.o.v. de basis $\mathcal{W}'$ van V/W. In het bijzonder geldt de volgende uitdrukking voor de determinanten:

$$\det(T) = \det(T|_W) \det(\bar{T}). \tag{1.6}$$

(Determinanten vormen het onderwerp van hoofdstuk II.)

Het tensorproduct van vectorruimten.

We beginnen met een voorbeeld. Beschouw de vectorruimte $V = P_n(K)$ van polynomen in X van graad hoogstens n en coëfficiënten in K. Een element van V kunnen we dus schrijven als $a_0 + a_1 X + \ldots + a_n X^n$. Laat $W = P_m(K)$ de vectorruimte zijn van polynomen in Y van graad hoogstens m. W bestaat dus uit polynomen van de vorm $b_0 + b_1 Y + \ldots + b_m Y^m$ met $b_j \in K$. Een polynoom in de twee variabelen X en Y van graad hoogstens n in X en graad hoogstens m in Y is een lineaire combinatie van de vorm $\displaystyle\sum_{i=0}^{n} \sum_{j=0}^{m} c_{ij} X^i Y^j$. Zo'n polynoom is een element van de vectorruimte opgespannen door de $(n+1)(m+1)$ basiselementen $X^i Y^j$. Deze vectorruimte noemen we het *tensorproduct* van de vectorruimten V en W, notatie $V \otimes W$.

In het algemeen laat V een vectorruimte zijn met basis $\{e_1, \ldots, e_n\}$ en W een vectorruimte (over hetzelfde lichaam van scalairen K) met basis $\{f_1, \ldots, f_m\}$ (n en m mogen ∞ zijn). Het tensorproduct $V \otimes W$ (of $V \otimes_K W$) is de vectorruimte van dimensie mn opgespannen door de basisvectoren $e_i \otimes f_j$ en met coëfficiënten in K. Het tensorproduct $V \otimes W$ is onafhankelijk van de keuze van de bases van V en W. In hoofdstuk IV zullen we een definitie van het tensorproduct geven die onafhankelijk is van een basis in V en W. Merk nog op dat $V \otimes W$ en $W \otimes V$ verschillende (maar wel isomorfe) vectorruimten zijn.

Als $v = \sum_{i=1}^{n} v_i e_i \in V$ en $w = \sum_{j=1}^{m} w_j f_j \in W$ dan is het tensorproduct $v \otimes w$ gedefinieerd als $\sum_{i=1}^{n} \sum_{j=1}^{m} v_i w_j e_i \otimes f_j$. $v \otimes w$ is dus een element van $V \otimes W$. In de fysische literatuur wordt vaak vw i.p.v. $v \otimes w$ geschreven.
Voor het tensorproduct van twee vectoren geldt (ga na):

$$\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w) \quad (\lambda \in K). \tag{1.7}$$

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w; \qquad v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2. \tag{1.8}$$

Opmerking: Een willekeurige vector in $V \otimes W$ is altijd een eindige lineaire combinatie van tensorproducten $v_i \otimes w_i$ met $v_i \in V$ en $w_i \in W$. Niet elke vector in $V \otimes W$ is echter zelf zo'n tensorproduct. Zie opgave I.34.

We kunnen herhaald tensorproducten van vectorruimten nemen. In zo'n geval geldt de associatieve eigenschap $U \otimes (V \otimes W) = (U \otimes V) \otimes W$. De uitdrukking $V_1 \otimes V_2 \otimes \ldots \otimes V_n$ is dus goed gedefinieerd als de V_i alle vectorruimten over K zijn. Ga na dat de dimensie van deze tensorproductruimte is gelijk aan het product van de dimensies van de afzonderlijke factoren. Tenslotte nog een opmerking over de notatie: i.p.v. het herhaald n-voudig tensorproduct $V \otimes \ldots \otimes V$ schrijven we ook wel $V^{\otimes n}$.

Voorbeeld: Laat V de vectorruimte zijn van reëel-of complexwaardige functies op een verzameling X. Dan is $V \otimes_K K^n$ (met $K = \mathbf{R}$ resp. $\mathbf{C}$) de vectorruimte van functies op een verzameling X met waarden in K^n. Een element van deze vectorruimte is dus te schrijven als een rijtje $f = (f_1, \ldots, f_n)$ met $f_j \in V$. Er geldt:

$$(f_1, \ldots, f_n) + (g_1, \ldots, g_n) = (f_1 + g_1, \ldots, f_n + g_n), \quad \lambda(f_1, \ldots, f_n) = (\lambda f_1, \ldots, \lambda f_n).$$

II. DETERMINANT EN SPOOR.

De determinant van een matrix.

Definitie: Een (n-de orde) *determinant* is een n-lineaire alternerende vorm op K^n (waarbij $K = \mathbf{R}$ of $\mathbf{C}$) die de waarde 1 aanneemt op de standaardbasis, d.w.z.

1. $\det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) \in K$ voor $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in K^n$.
2. $\det(\lambda\mathbf{a}_1 + \mu\mathbf{b}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \lambda \det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) + \mu \det(\mathbf{b}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ voor $\mathbf{a}_1, \mathbf{b}_1, \ldots, \mathbf{a}_n \in K^n$ en $\lambda, \mu \in K$ (*lineariteit in de eerste component*).
3. $\det(\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n) = -\det(\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_n)$ (de determinant is een *alternerende* vorm.) In het bijzonder is de determinant nul als twee van de $\mathbf{a}_j$'s gelijk zijn.
4. $\det(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) = 1$ (de determinant is 1 op de standaardbasis van K^n).

Lineariteit in de andere $n-1$ componenten volgt uit lineariteit in de eerste component samen met de alternerendheid. De determinant $\det(A)$ van een $n \times n$-matrix A met kolomvectoren $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is gedefinieerd als $\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. Uit eigenschap 3 volgt meteen dat $\det(\mathbf{a}_1, \ldots, \mathbf{a}_n) = 0$ als minstens twee van de $\mathbf{a}_j$'s gelijk zijn.

Uit de definitie volgt een unieke vorm voor de determinant: voor $i = 1, \ldots, n$ is $\mathbf{a}_i = \sum_{i_j=1}^{n} a_{i_j i}\mathbf{e}_{i_j}$; vul deze uitdrukkingen in voor $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in de determinant. Wegens de multilineariteit (lineariteit in elke component) kunnen we de n sommen samen met de coëfficiënten $a_{i_j j}$ buiten de determinant halen en dan vinden we

$$\det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \ldots \sum_{i_n=1}^{n} \epsilon_{i_1 i_2 \ldots i_n} a_{i_1 1} a_{i_2 2} \ldots a_{i_n n}. \tag{2.1}$$

Hierbij is het *Levi-Civitasymbool* $\epsilon_{i_1 i_2 \ldots i_n}$ gedefinieerd als $\det(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_n})$. Als twee van de indices i_j gelijk zijn, dan is $\epsilon_{i_1 i_2 \ldots i_n} = 0$. In het geval dat de indices alle verschillend zijn, passen we de alternerendheidseigenschap toe om de waarde van $\epsilon_{i_1 i_2 \ldots i_n}$ te bepalen: door herhaald omwisselen van twee e_{i_j}'s in de determinant kunnen we bereiken dat de e_{i_j}'s in de juiste volgorde staan. Dus $\det(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_n}) = (-1)^N \det(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ waarbij N het aantal benodigde verwisselingen is. Dus als alle indices verschillend zijn, is $\epsilon_{i_1 i_2 \ldots i_n} = 1$ of -1 als het aantal van deze paarverwisselingen even resp. oneven is. Dat dit goed gedefinieerd is, m.a.w. het even of oneven zijn van het aantal paarverwisselingen hangt niet af van de wijze waarop we verwisselen, volgt uit de theorie van de permutaties.

Voorbeeld. $\epsilon_{2413} = (-1)^3 = -1$: $(2413) \to (1423) \to (1243) \to (1234)$ (3 paarverwisselingen).

Permutaties. Een *permutatie* op n objecten, zeg de getallen $1, 2, \ldots, n$, is een bijectie van de verzameling $\{1, \ldots, n\}$ op zichzelf. Elke permutatie is een compositie van paarverwisselingen; dit zijn permutaties p zodat $p(i) = j$, $p(j) = i$ voor twee verschillende getallen $i, j \in \{1, \ldots, n\}$ en zodat $p(k) = k$ als $k \neq i$ of j. Een permutatie is op meer manieren als een compositie van paarverwisselingen te schrijven, maar altijd is het aantal paarverwisselingen hetzij oneven, hetzij even. Het is niet moeilijk om na te gaan dat het aantal even (resp. oneven) is als het aantal paren (i, j) zodat $i < j$ en $p(i) > p(j)$ even (resp. oneven) is: immers bij elke paarverwisseling neemt het aantal van dergelijke paren met 1 toe- of af. Het aantal is nul dan en slechts dan als p de identieke permutatie is. Als p de compositie van een even (resp. oneven) aantal paarverwisselingen is, dan is het teken $\sigma(p)$ van p gelijk aan 1 (resp. -1). Voor permutaties p en p' op $\{1, \ldots, n\}$ geldt dan dat $\sigma(p \circ p') = \sigma(p)\sigma(p')$ en $\sigma(p^{-1}) = \sigma(p)$. Als voor de permutatie p geldt dat $p(1) = i_1, p(2) = i_2, \ldots, p(n) = i_n$, dan is $\sigma(p) = \epsilon_{i_1 i_2 \ldots i_n}$.

We kunnen de determinant nu ook schrijven als $\det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \sum_p \sigma(p) a_{p(1)1} a_{p(2)2} \ldots a_{p(n)n}$ waarbij de som wordt genomen over alle permutaties van $\{1, 2, \ldots, n\}$. Verder geldt:

$$\sum_{i_1, i_2, \ldots, i_n} \epsilon_{i_1 i_2 \ldots i_n} a_{j_1 i_1} a_{j_2 i_2} \ldots a_{j_n i_n} = \det(\mathbf{a}_{j_1}, \mathbf{a}_{j_2}, \ldots, \mathbf{a}_{j_n}) = \epsilon_{j_1 j_2 \ldots j_n} \det(A). \tag{2.2}$$

Eigenschappen van determinanten. Een gevolg van (2.2) is

Propositie 2.1. Laat A, B twee $n \times n$-matrices zijn. Dan geldt

i. $\det(A) = \det(A^T)$.

ii. $\det(AB) = \det(A)\det(B)$.

Bewijs: (i.)

$$\det(A^T) = \sum_p \sigma(p) a_{1p(1)} \ldots a_{np(n)} = \sum_p \sigma(p) a_{p^{-1}(1)1} \ldots a_{p^{-1}(n)n}$$

waarbij p loopt over alle permutaties van $\{1, \ldots, n\}$. In de tweede gelijkheid hebben we gebruikt dat als $j = p(i)$, dan $i = p^{-1}(j)$, dus a_{ij} is dan te schrijven als zowel $a_{ip(i)}$ als $a_{p^{-1}(j)j}$. M.b.v. $\sigma(p) = \sigma(p^{-1})$ volgt dan

$$\det(A^T) = \sum_p \sigma(p^{-1}) a_{p^{-1}(1)1} \ldots a_{p^{-1}(n)n} = \det(A)$$

daar sommeren over p hetzelfde is als sommeren over p^{-1}.

(ii.) We geven de elementen van AB aan d.m.v. $(AB)_{ij}$. Dan geldt

$$\det(AB) = \sum_{i_1, \ldots, i_n} \epsilon_{i_1 \ldots i_n} (AB)_{1i_1} \ldots (AB)_{ni_n} = \sum_{i_1, \ldots, i_n} \sum_{j_1, \ldots, j_n} \epsilon_{i_1 \ldots i_n} a_{1j_1} b_{j_1 i_1} \ldots a_{nj_n} b_{j_n i_n}.$$

M.b.v. (2.2) kunnen we de laatste term schrijven als

$$\det(B) \sum_{j_1, \ldots, j_n} \epsilon_{j_1 \ldots j_n} a_{1j_1} \ldots a_{nj_n} = \det(B)\det(A). \quad \diamond$$

Uit de laatste eigenschap volgt, m.b.v. $UU^{-1} = I_n$ dat als U een inverteerbare matrix is, dan $\det(U^{-1}) = (\det(U))^{-1}$. I.h.b. is $\det(U) \neq 0$ als U inverteerbaar is. Verder volgt uit $\det(U^{-1}AU) = \det(U^{-1})\det(A)\det(U) = \det(A)$ dat twee gelijkvormige matrices A en $B = U^{-1}AU$ dezelfde determinant hebben. Deze eigenschap kunnen we gebruiken om de determinant van een lineair endomorfisme te definiëren: laat V een vectorruimte van eindige dimensie zijn en $T : V \to V$ een lineaire afbeelding. De determinant $\det(T)$ van T is nu als volgt gedefinieerd: kies een basis $\mathcal{B}$ van V. De matrix van T t.o.v. deze basis is $T_{\mathcal{B}}$. Dan is $\det(T) := \det(T_{\mathcal{B}})$. Deze definitie is onafhankelijk van de gekozen basis.

Als A inverteerbaar is, dan is $\det(A) \neq 0$. Omgekeerd, als de matrix $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ singulier (d.w.z. niet-inverteerbaar) is, dan is de rang van A kleiner dan n en de dimensie van de kern van A is groter dan 0. Kies een basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ van $\ker(A)$ en vul deze aan tot een basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ van K^n. De matrix $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ heeft dan rang n en is inverteerbaar, dus $\det(B) \neq 0$. De eerste k kolommen van de matrix AB zijn dan nul en dus is $\det(AB) = 0$ volgens eigenschap 2. Maar dan is $\det(A) = 0$. De determinant van de $n \times n$-matrix A is dus nul dan en slechts dan als A singulier is. Anders geformuleerd: $\det(\mathbf{a}_1, \ldots, \mathbf{a}_n) = 0$ dan en slechts dan als het stelsel $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ lineair afhankelijk is.

Zij A een $m \times n$-matrix. Kies k rijen en k kolommen. De determinant van de $k \times k$-deelmatrix van A die ontstaat uit A door alleen de elementen die in de gekozen rijen en kolommen voorkomen te nemen heet een *minor* of *onderdeterminant* van orde k.

Propositie 2.2: De rang van een $m \times n$-matrix A is k dan en slechts dan als er minoren van orde k zijn die ongelijk zijn aan nul en alle minoren van orde groter dan k nul zijn.

Bewijs: Zij k de rang van A. Dan zijn er k lineair onafhankelijke kolommen. Door zo nodig de kolommen te permuteren kunnen we ervoor zorgen dat de eerste k kolommen lineair onafhankelijk zijn. Merk op dat de rang van de matrix niet verandert bij het verwisselen van kolommen of rijen. Beschouw nu de $m \times k$-deelmatrix van A die ontstaat door alleen de eerste k kolommen te nemen. Omdat de rang van de deelmatrix k is, zijn er k lineair onafhankelijke rijen. Door weer zo nodig te permuteren kunnen we aannemen dat de eerste k rijen lineair onafhankelijk zijn. De $k \times k$-deelmatrix die nu ontstaat door alleen de eerste k rijen te nemen heeft rang k en dus een determinant die ongelijk aan nul is. Maar deze determinant is een minor van A van orde k.

Omgekeerd, als A een $\ell \times \ell$-deelmatrix heeft met determinant ongelijk aan nul, dan zijn de rijen en kolommen in A waaruit deze deelmatrix is samengesteld, lineair onafhankelijk. De rang van A is dus minstens ℓ. $\diamond$

De Wronskiaan.

We gebruiken het bovenstaande resultaat om een criterium af te leiden voor de lineaire onafhankelijkheid van een stelsel differentieerbare functies. Laat $f_1, \ldots, f_n$ $n-1$ keer differentieerbare (reële of complexe) functies op een interval $[a, b] \subset \mathbf{R}$ zijn. Het stelsel functies $f_1, \ldots, f_n$ is lineair afhankelijk indien er $\lambda_1, \ldots, \lambda_n \in K(= \mathbf{R}$ of $\mathbf{C})$ bestaan, niet alle gelijk aan nul, zodanig dat

$$\lambda_1 f_1(x) + \ldots + \lambda_n f_n(x) = 0 \quad \text{voor } x \in [a, b].$$

Dit is het geval dan en slechts dan als voor zekere $(\lambda_1, \ldots, \lambda_n) \neq (0, \ldots, 0)$

$$\begin{cases} \lambda_1 f_1(x) + \ldots + \lambda_n f_n(x) & = & 0 \\ \lambda_1 f_1'(x) + \ldots + \lambda_n f_n'(x) & = & 0 \\ \vdots & & \vdots \\ \lambda_1 f_1^{(n-1)}(x) + \ldots + \lambda_n f_n^{(n-1)}(x) & = & 0 \end{cases}$$

en dit is het geval indien de determinant $W(f_1, \ldots, f_n)(x) = \begin{vmatrix} f_1^{(n-1)}(x) & \ldots & f_n^{(n-1)}(x) \\ \vdots & \ddots & \vdots \\ f_1'(x) & \ldots & f_n'(x) \\ f_1(x) & \ldots & f_n(x) \end{vmatrix}$ gelijk is aan nul. $W(f_1, \ldots, f_n)$ heet de *Wronskiaan* van het stelsel $\{f_1, \ldots, f_n\}$. Omgekeerd geldt: als $W(f_1, \ldots, f_n)(x) = 0$ op $[a, b]$ en voor elke $x \in [a, b]$ is er een i zodanig dat
$W(f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_n)(x) \neq 0$, dan is het stelsel $\{f_1, \ldots, f_n\}$ lineair afhankelijk. (Voor een voorbeeld neem $f_1(x) = x^3$ en $f_2(x) = |x|^3$. Er geldt dat $W(f_1, f_2)(x) = 0$ voor $x \in \mathbf{R}$ maar f_1, f_2 zijn op $\mathbf{R}$ lineair onafhankelijk.)

Voorbeeld: Laat $f_1(x) = \sin x$ en $f_2(x) = \cos x$ op $[0, \pi]$. De Wronskiaan $W(\sin x, \cos x) = \begin{vmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{vmatrix}$ is identiek gelijk aan 1 en dus is het stelsel $\{\sin x, \cos x\}$ lineair onafhankelijk op $[0, \pi]$.

Toepassing: Beschouw de homogene lineaire 2-de orde differentiaalvergelijking

$$y''(x) + c_1(x)y'(x) + c_0(x)y(x) = 0, \quad x \in [a, b] \tag{2.3}$$

waarbij de coëfficiënten c_0, c_1 continue (reële of complexe) functies op het interval $[a, b] \subset \mathbf{R}$ zijn. Het is eenvoudig in te zien dat als y_1 en y_2 oplossingen zijn, dan is $\alpha y_1 + \beta y_2$ voor $\alpha, \beta \in \mathbf{R}$ of $\mathbf{C}$ ook een oplossing. De oplossingen van de differentiaalvergelijking (2.3) vormen dus een vectorruimte (over $\mathbf{R}$ of $\mathbf{C}$). In feite kan worden aangetoond dat de dimensie van deze vectorruimte gelijk is aan 2. Voor twee oplossingen y_1, y_2 geldt dat $W = W(y_1, y_2) = y_1'y_2 - y_1y_2'$. Uit de d.v. volgt dat $W' = -c_1 W$ dus $W(y_1, y_2)(x) = \exp(-\int_d^x c_1(t)dt)W(y_1, y_2)(d)$ voor $d \in \mathbf{R}$. y_1, y_2 zijn dus lineair afhankelijk dan en slechts dan als $W(y_1, y_2)(d) = 0$ voor een enkel punt $d \in [a, b]$. In het bijzonder zijn er niet meer dan twee lineair onafhankelijke oplossingen.

De determinant van Vandermonde. De volgende determinant is vaak nuttig bij kwesties over lineaire (on)afhankelijkheid:

Propositie 2.3: Laat $x_1, \ldots, x_n$ complexe getallen zijn. Dan is

$$V_n(x_1, x_2, \ldots, x_n) := \begin{vmatrix} x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{(n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \ldots & x_n \\ 1 & 1 & \ldots & 1 \end{vmatrix} = \prod_{1 \leq i < j \leq n} (x_i - x_j).$$

Bewijs: We passen inducie naar n toe. Voor $n = 1$ staat links en rechts van het =-teken 1 (een leeg product is gelijk aan 1). Stel de bewering is waar voor V_k met $k < n$. Vat in $V_n(x_1, x_2, \ldots, x_n)$

het element x_1 op als een variabele en laat $x_2, \dots, x_n$ vaste getallen zijn. We kunnen aannemen dat deze verschillend zijn, omdat anders de determinant zeker nul is. Dan is $V_n(x_1, x_2, \dots, x_n)$ een polynoom in x_1 van graad $n-1$ met kopcoëfficiënt (d.w.z. de coëfficiënt van de hoogste macht x_1^{n-1}) gelijk aan $V_{n-1}(x_2, \dots, x_n)(\neq 0)$ en nulpunten $x_1 = x_2, \dots, x_n$. Het polynoom is dus te ontbinden in (lineaire) factoren:

$$V_n(x_1, x_2, \dots, x_n) = V_{n-1}(x_2, \dots, x_n) \prod_{j=2}^{n} (x_1 - x_j) = \prod_{2 \leq i < j \leq n} (x_i - x_j) \prod_{j=2}^{n} (x_1 - x_j)$$

volgens de inductieveronderstelling. De uitdrukking in het rechterlid is precies $\prod_{1 \leq i < j \leq n} (x_i - x_j)$ $\diamond$.

Toepassing: Laat $x_0, \dots, x_n$ verschillende reële getallen zijn en laat $y_0, \dots, y_n$ willekeurige reële getallen zijn. Dan is er precies één polynoom P van graad hoogstens n zodanig dat $P(x_j) = y_j$ voor $j = 0, \dots, n$.

Bewijs: Laat $P(X) = a_n X^n + \dots + a_1 X + a_0$ en stel $P(x_j) = y_j$. Dan is

$$a_n x_j^{n-1} + \dots + a_1 x_j + a_0 = y_j \quad \text{voor } j = 0, \dots n.$$

Dit levert een stelsel van $n+1$ vergelijkingen in de $n+1$ onbekenden $a_n, \dots, a_0$. De coëfficiënten-determinant van het stelsel is $V_n(x_0, x_1, \dots, x_n)$ en deze is ongelijk aan nul omdat $x_0, \dots, x_n$ verschillend zijn. Maar dan heeft het stelsel precies één oplossing (immers als A een inverteerbare $N \times N$-matrix is dan heeft de vergelijking $A\mathbf{x} = \mathbf{b}$ voor elke $\mathbf{b} \in \mathbf{C}^N$ de unieke oplossing $\mathbf{x} = A^{-1}\mathbf{b}$). $\diamond$

Ontwikkeling van een determinant naar een kolom.

Uit (2.1) zien we dat in elk van de $n!$ termen in de uitdrukking van $\det(A)$ er precies één element uit elke rij en één element uit elke kolom voorkomt. We kunnen (2.1) voor vaste k schrijven als $\det(A) = \sum_{i=1}^{n} a_{ik} A_{ik}$, waarbij

$$A_{ik} = \sum_{i_1 \dots \hat{i_k} \dots i_n \neq i} \epsilon_{i_1 \dots i_n} a_{i_1 1} \dots \hat{a}_{i_k k} \dots a_{i_n n}$$

de *cofactor* van a_{ik} is. Het dakje boven de factor $a_{i_k k}$ betekent dat deze wordt weggelaten. In het bijzonder is

$$A_{11} = \sum_{i_2 \dots i_n = 2}^{n} \epsilon_{1 i_2 \dots i_n} a_{i_2 2} \dots a_{i_n n} = \sum_{i_2 \dots i_n = 2}^{n} \epsilon_{i_2 \dots i_n} a_{i_2 2} \dots a_{i_n n},$$

m.a.w. A_{11} is de determinant van de $(n-1) \times (n-1)$-matrix die uit A ontstaat door de 1e rij en kolom weg te laten. Door eerst in de matrix A de i-e rij naar de eerste rij te verplaatsen in door de rij $i-1$ keer te verwisselen met de rij die er boven staat en vervolgens de j-e kolom naar de eerste kolom te verplaatsen door de kolom $j-1$ keer te verwisselen met de kolom die er net voor staat, kunnen we ervoor zorgen dat het element a_{ij} in de 1e rij en kolom komt te staan. De cofactor van dit element in de zo ontstane matrix is dus gelijk aan de determinant van de matrix die uit A ontstaat door de i-e rij en j-e kolom weg te laten (merk op dat de volgorde van de andere rijen en kolommen niet is veranderd). Anderzijds is de determinant van de zo ontstane matrix gelijk aan $(-1)^{i+j-2}\det(A) = (-1)^{i+j}\det(A)$. M.a.w. de cofactor A_{ij} van a_{ij} is gelijk aan $(-1)^{i+j}$ maal de determinant van de matrix die uit A ontstaat door de i-e rij en j-e kolom weg te laten.

Als voorbeeld ontwikkelen we de volgende determinant naar de 2e kolom:

$$\begin{vmatrix} 2 & 1 & -1 \\ 3 & 1 & -2 \\ 1 & 2 & 2 \end{vmatrix} = -\begin{vmatrix} 3 & -2 \\ 1 & 2 \end{vmatrix} + \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} - 2\begin{vmatrix} 2 & -1 \\ 3 & -2 \end{vmatrix} = -8 + 5 - 2 \cdot -1 = -1.$$

Omdat $\det(A) = \det(A^T)$ kunnen we de determinant ook berekenen door te ontwikkelen naar een rij i.p.v. een kolom.

In de vorige paragraaf hebben we gezien dat $\sum_{i=1}^n a_{ik}A_{ik} = \det(A)$ voor $k = 1, \ldots n$. Beschouw nu de matrix A' die uit A ontstaat door de k-e kolom te vervangen door een andere, zeg de ℓ-e kolom. A' heeft dan twee gelijke kolommen en ontwikkelen naar de k-e kolom levert dan $0 = \det(A') = \sum_{i=1}^n a_{i\ell}A_{ik}$ (omdat de cofactoren A_{ik} onafhankelijk zijn van $a_{1k}, \ldots, a_{nk}$ zijn de cofactoren van de elementen in de k-e kolom van A' gelijk aan de cofactoren in de k-e kolom van A). We hebben dus aangetoond:

$$\sum_{i=1}^n a_{i\ell}A_{ik} = \delta_{k\ell}\det(A). \tag{2.4}$$

We kunnen (2.4) schrijven als een matrixproduct: definieer de *geadjungeerde matrix* $\mathrm{adj}(A)$ van A als de getransponeerde van de matrix van cofactoren van A, m.a.w. $(\mathrm{adj}(A))_{ij} = A_{ji}$. Dan zegt (2.3) dat

$$A \cdot \mathrm{adj}(A) = \det(A) \cdot I. \tag{2.5}$$

Gevolg: als $\det(A) \neq 0$ dan is A inverteerbaar met inverse $A^{-1} = \dfrac{1}{\det(A)}\mathrm{adj}(A)$.

Het spoor van een matrix.

Zij A een $n \times n$-matrix. Het *spoor* $\mathrm{tr}(A)$ van A is de som van de elementen op de hoofddiagonaal: $\mathrm{tr}(A) = \sum_{i=1}^n A_{ii}$. Als B een $n \times n$-matrix is, dan geldt $\mathrm{tr}(AB) = \sum_{i,j} A_{ij}B_{ji} = \mathrm{tr}(BA)$. In het bijzonder hebben twee gelijkvormige matrices hetzelfde spoor: $\mathrm{tr}(U^{-1}AU) = \mathrm{tr}(UU^{-1}A) = \mathrm{tr}(A)$. Dit biedt net als in het geval van de determinant de mogelijkheid om het spoor $\mathrm{tr}(T)$ van een lineaire afbeelding $T : V \to V$ voor V eindig-dimensionaal op basis-onafhankelijke wijze te definiëren als het spoor van een willekeurige matrix van T.

Het volume van een k-blok in $\mathbf{R}^n$.

Het k-blok opgespannen door $\mathbf{a}_1, \ldots, \mathbf{a}_k$ (met $\mathbf{a}_1, \ldots, \mathbf{a}_k$ lineair onafhankelijke vectoren in $\mathbf{R}^n$) is de verzameling $\{\mathbf{x} = t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + \ldots + t_k\mathbf{a}_k : 0 \leq t_i \leq 1,\ i = 1, \ldots, k\}$. Het volume $V(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$

wordt berekend als volgt: Laat $A = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$ de matrix met kolomvectoren $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ zijn. Dan is

$$V(\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_k) = \sqrt{\det(A^T A)}. \tag{2.6}$$

Voor het geval dat $k = n$ is $V(\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_n) = |\det(\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_n)|$. Een schets van een bewijs voor $k = n$: Het volume verandert niet als we een van de zijvlakken van het k-blok evenwijdig aan zichzelf verschuiven, m.a.w. $V(\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_n) = V(\mathbf{a}'_1, \mathbf{a}_2 \ldots, \mathbf{a}_n)$ waarbij $\mathbf{a}'_1 - \mathbf{a}_1$ een lineaire combinatie is van $\mathbf{a}_2, \ldots, \mathbf{a}_n$. We kunnen nu een zodanig lineaire combinatie kiezen dat $\mathbf{a}'_1$ orthogonaal is met $\mathbf{a}_2, \ldots, \mathbf{a}_n$. Op soortgelijke wijze kunnen we bij $\mathbf{a}_2$ een lineaire combinatie van $\mathbf{a}_3, \ldots, \mathbf{a}_k$ optellen zo, dat de somvector a'_2 orthogonaal is met $\mathbf{a}_3, \ldots, \mathbf{a}_n$ en uiteraard ook met $\mathbf{a}'_1$. Zo verdergaand vinden we een orthogonaal stelsel vectoren $\{\mathbf{a}'_1, \ldots, \mathbf{a}'_n\}$ zodat $V(\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_n) = V(\mathbf{a}'_1, \mathbf{a}'_2 \ldots, \mathbf{a}'_n)$. Evenzo geldt, wegens multilineariteit, dat $\det(\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_n) = \det(\mathbf{a}'_1, \mathbf{a}_2 \ldots, \mathbf{a}_n) = \ldots =$
$= \det(\mathbf{a}'_1, \mathbf{a}'_2 \ldots, \mathbf{a}'_n) =: \det(A')$. Tenslotte is wegens orthogonaliteit

$$V(\mathbf{a}'_1, \mathbf{a}'_2 \ldots, \mathbf{a}'_n)^2 = \|\mathbf{a}'_1\| \cdot \|\mathbf{a}'_2\| \ldots \|\mathbf{a}'_n\|^2 = \det(A'^T A') = \det(A')^2.$$

Voor het algemene geval ($k \leq n$) gebruiken we dat $V(\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_k) = V(\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_k, \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n)$ waarbij $\{\mathbf{b}_{k+1}, \ldots, \mathbf{b}_n\}$ een orthonormaal stelsel is zodat $(\mathbf{a}_i, \mathbf{b}_j) = 0$ en dat
$\det(A^T A) = \det(\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n)^2$. (De begrippen orthogonaliteit en orthonormaliteit worden in hoofdstuk IV behandeld.)

Aan de determinant van een lineaire afbeelding $T : \mathbf{R}^n \to \mathbf{R}^n$ kunnen we nu een meetkundige interpretatie geven: laat K een n-blok in $\mathbf{R}^n$ zijn, opgespannen door vectoren $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Zoals we in de vorige paragraaf gezien hebben, is het volume van K gelijk aan $V(K) = |\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)|$. Het beeld $T(K)$ is een n-blok (eventueel gedegenereerd), opgespannen door $T(\mathbf{a}_1), \ldots, T(\mathbf{a}_n)$. Dan geldt, m.b.v. Prop.2.1,

$$V(T(K)) = |\det(T(\mathbf{a}_1), \ldots, T(\mathbf{a}_n))| = |\det(T)||\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)| = |\det(T)|V(K).$$

De absolute waarde van de determinant geeft dus de vergrotingsfactor aan voor het volume van een n-blok in $\mathbf{R}^n$ onder de afbeelding T. Voor andere "nette" gebieden G in R^n (zoals het inwendige van een n-bol) kunnen we het volume benaderen door G zo efficiënt mogelijk te overdekken met n-blokken. Onder de afbeelding T wordt het volume van al deze blokken met dezelfde factor $|\det(T)|$ vergroot; ditzelfde geldt dan ook voor het volume van G zelf.
Zonder bewijs merken we nog op dat het teken van de determinant te maken heeft met de oriëntatie: afbeeldingen met negatieve determinant (zoals spiegelingen) keren de oriëntatie om: linksdraaiend wordt rechtsdraaiend.

De afstand van een punt tot een k-vlak in $\mathbf{R}^n$. M.b.v. de uitdrukking voor het volume van een k-blok leiden we eenvoudig de volgende uitdrukking af voor de afstand $d(B, V)$ van een punt B tot het k-vlak $V = \text{span}(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ af: zij $\mathbf{b}$ de vector $\overrightarrow{OB}$. Dan is

$$d(B, V) = \frac{V(\mathbf{b}, \mathbf{a}_1, \ldots, \mathbf{a}_k)}{V(\mathbf{a}_1, \ldots, \mathbf{a}_k)}. \tag{2.7}$$

Voor de afstand van een punt B tot een lijn $\ell = \text{span}\{\mathbf{a}\}$ in $\mathbf{R}^3$ geeft dit

$$d(B, \ell) = \frac{V(\mathbf{b}, \mathbf{a})}{V(\mathbf{a})} = \frac{\|\mathbf{b} \times \mathbf{a}\|}{\|\mathbf{a}\|}$$

waarbij $\mathbf{b}$ de vector OB voorstelt.

III. SPECTRAALTHEORIE VAN COMPLEXE ENDOMORFISMEN

Eigenwaarden en eigenvectoren.

Zij V een vectorruimte over $K = \mathbf{R}$ of $\mathbf{C}$ van eindige dimensie n en $T : V \to V$ een lineaire afbeelding. $v \in V$ heet een *eigenvector* van T als er een $\lambda \in K$ bestaat zodat $T(v) = \lambda v$ en $v \neq 0$. λ heet een *eigenwaarde* van T. De verzameling eigenvectoren met eigenwaarde λ samen met de nulvector noemen we de *eigenruimte* $\mathrm{Eig}_T(\lambda)$ bij eigenwaarde λ. $\mathrm{Eig}_T(\lambda) = \ker(T - \lambda \cdot id)$ is een lineaire deelruimte van V. De eigenwaarden zijn de nulpunten (in K) van het *karakteristieke polynoom* $\chi_T(x) = \det(T - x \cdot id)$. χ_T is een polynoom van graad n en T heeft dus hoogstens n eigenwaarden. (Het karakteristieke polynoom wordt ook wel gedefinieerd als $\det(x \cdot id - T)$. Dit scheelt een factor ± 1.) De *algebraïsche multipliciteit* van de eigenwaarde λ is de multipliciteit van λ als nulpunt van het karakteristieke polynoom en de *meetkundige multipliciteit* is de dimensie van de bijbehorende eigenruimte. ($a \in K$ is een nulpunt van multipliciteit k van het polynoom P als $P(a) = 0, \ldots, P^{(k-1)}(a) = 0$ maar $P^{(k)}(a) \neq 0$.) Zoals we nog zullen zien is de meetkundige multipliciteit nooit groter dan de algebraïsche.

De bovenstaande begrippen worden overgedragen op een $n \times n$-matrix A: een vector $\mathbf{x} \in K^n$, $\mathbf{x} \neq 0$, heet een eigenvector van A met eigenwaarde λ indien $A\mathbf{x} = \lambda\mathbf{x}$. De eigenwaarden van A zijn de nulpunten van het karakteristieke polynoom $\chi_A(x) = \det(A - x \cdot I_n)$. Gelijkvormige matrices hebben hetzelfde karakteristiek polynoom: laat U een matrix zijn zodanig dat $B = U^{-1}AU$. Dan is

$$\begin{aligned}\chi_B(x) = \det(B - x \cdot id) = \det(U^{-1}AU - x \cdot id) = \det U^{-1}(A - x \cdot id)U) = \\ = \det(U^{-1}) \det(A - x \cdot id) \det(U) = \det(A - x \cdot id) = \chi_A(x).\end{aligned}$$

Verder geldt dat $\mathbf{y}$ een eigenvector van B is met eigenwaarde λ dan en slechts dan als $U\mathbf{y}$ een eigenvector van A is met eigenwaarde λ. Er geldt dus $U(\mathrm{Eig}_B(\lambda)) = \mathrm{Eig}_A(\lambda)$. Omdat U inverteerbaar is geldt dat $\dim \mathrm{Eig}_B(\lambda) = \dim \mathrm{Eig}_A(\lambda)$.

Voorbeeld: Laat $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Dan is $\chi_A(x) = \begin{vmatrix} 1-x & 1 \\ 0 & 1-x \end{vmatrix} = (1-x)^2$. A heeft een eigenwaarde 1 met algebraïsche multipliciteit 2. De meetkundige multipliciteit is gelijk aan de dimensie van de kern van de matrix $A - I = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Deze matrix heeft rang 1 en dus is dim $\ker(A - I) = 1$. De meetkundige multipliciteit van de eigenwaarde 1 is dus gelijk aan 1.

De verzameling eigenwaarden van de afbeelding T heet het *spectrum* van T. In oneindig-dimensionale vectorruimten is de theorie van eigenwaarden en eigenvectoren wat ingewikkelder. Zo kan het spectrum van een lineaire afbeelding bestaan uit een discreet en een continu deel. In dit college bekijken we alleen het geval dat de dimensie eindig is.

Propositie 3.1: Zij A een complexe $n \times n$-matrix met eigenwaarden $\lambda_1, \ldots, \lambda_n$ (met algebraïsche multipliciteit geteld). Dan is

$$\det(A) = \lambda_1 \cdot \ldots \cdot \lambda_n, \qquad \mathrm{tr}(A) = \lambda_1 + \ldots + \lambda_n. \tag{3.1}$$

Bewijs: Het karakteristieke polynoom $\chi_A(x) = \det(A - x \cdot I)$ splitst in n lineaire complexe factoren: $\chi_A(x) = \prod_{j=1}^{n}(\lambda_j - x)$ waarbij $\lambda_1, \ldots, \lambda_n$ de eigenwaarden zijn. Nu volgt meteen dat $\det(A) = \chi_A(0) = \prod_{j=1}^{n} \lambda_j$. De coëfficiënt van x^{n-1} is enerzijds gelijk aan $(-1)^{n-1} \sum_{j=1}^{n} \lambda_j$ en anderzijds is het gelijk aan $(-1)^{n-1}\mathrm{tr}(A)$. $\diamond$

Diagonaliseerbare en nilpotente afbeeldingen.

Propositie 3.2: Eigenvectoren behorende bij verschillende eigenwaarden van een lineaire afbeelding $T : V \to V$ zijn lineair onafhankelijk.

Bewijs: Laat $v_1, \ldots, v_k$ lineair eigenvectoren zijn bij verschillende eigenwaarden $\lambda_1, \ldots, \lambda_k$ van T. Stel dat er getallen $a_1, \ldots, a_k \in K$ zijn, zo, dat

$$a_1 v_1 + \ldots + a_k v_k = 0. \tag{3.2}$$

Dan is ook

$$0 = T(a_1 v_1 + \ldots + a_k v_k) = a_1 T(v_1) + \ldots + a_k T(v_k) = a_1 \lambda_1 v_1 + \ldots + a_k \lambda_k v_k. \tag{3.3}$$

Door (3.2) met λ_k te vermenigvuldigen en van (3.3) af te trekken, vinden we

$$a_1(\lambda_1 - \lambda_k)v_1 + \ldots + a_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0. \tag{3.4}$$

Neem nu aan dat $v_1, \ldots, v_k$ lineair afhankelijk zijn. Er zijn nu twee mogelijkheden: (i) $v_1, \ldots, v_{k-1}$ zijn lineair afhankelijk. (ii) $v_1, \ldots, v_{k-1}$ zijn lineair onafhankelijk. In dit geval is $a_1 = \ldots = a_{k-1} = 0$. Uit (3.2) volgt dan $a_k v_k = 0$ en dus $a_k = 0$. In dit geval zijn $v_1, \ldots, v_k$ lineair onafhankelijk, wat in tegenspraak is met de aanname. Conclusie: $v_1, \ldots, v_{k-1}$ zijn lineair afhankelijk. Door dezelfde redenering op $v_1, \ldots, v_{k-1}$ toe te passen, vinden we dat $v_1, \ldots, v_{k-2}$ lineair afhankelijk zijn. Zo verdergaand, vinden we uiteindelijk dat v_1 lineair afhankelijk is, m.a.w. $v_1 = 0$. Maar dit is een tegenspraak met de aanname dat v_1 een eigenvector is. $\diamond$

Gevolg: Als $T : V \to V$ $n = \dim(V)$ verschillende eigenwaarden heeft, dan heeft V een basis bestaande uit eigenvectoren van T. In dit geval noemen we T *diagonaliseerbaar.*

Ten opzichte van zo'n basis $\mathcal{B}$ van eigenvectoren is de matrix $T_{\mathcal{B}}^{\mathcal{B}}$ van T een diagonaalmatrix. Analoog noemen we een $n \times n$-matrix A diagonaliseerbaar (over K) als K^n een basis van eigenvectoren van A heeft. Dit is precies het geval als er een inverteerbare matrix U met elementen in K bestaat zodat $D = U^{-1}AU$ een diagonaalmatrix is. De kolommen van U bestaan dan uit eigenvectoren van A. Een afbeelding $T : V \to V$ is dus diagonaliseerbaar als de matrix van de afbeelding (t.o.v. een willekeurige basis) diagonaliseerbaar is. Omdat de meetkundige multipliciteit van een eigenwaarde nooit groter is dan de algebraïsche, geldt dat een afbeelding $T : V \to V$ diagonaliseerbaar is dan en slechts dan als alle nulpunten van het karakteristiek polynoom in K liggen én de meetkundige multipliciteit van elke eigenwaarde gelijk is aan de algebraïsche. Door

$K = \mathbf{C}$ te nemen kunnen we ervoor zorgen dat aan de eerste voorwaarde is voldaan (maar niet aan de tweede).

Voorbeelden:

1. De matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ heeft eigenwaarden i en $-i$. De matrix is dus niet diagonaliseerbaar over $\mathbf{R}$ (er zijn geen reële eigenwaarden en eigenvectoren), maar wel over $\mathbf{C}$.
2. De matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is niet-diagonaliseerbaar: de eigenwaarde 1 heeft meetkundige multipliciteit 1 en algebraïsche multipliciteit 2.
3. Een afbeelding $T : V \to V$ (resp. een $n \times n$-matrix N) heet *nilpotent* als $T^k = 0$ (resp. $N^k = O$) voor zekere gehele $k > 0$. Een nilpotente afbeelding (resp. matrix) heeft 0 als enige eigenwaarde. Als er een basis van eigenvectoren bestaat, dan is $T(v) = 0$ voor alle $v \in V$ (resp. $N\mathbf{v} = \mathbf{0}$ voor alle $\mathbf{v} \in K^n$), en dus is $T = 0$ resp. $N = O$. De enige nilpotente diagonaliseerbare afbeelding (resp. matrix) is dus de nulafbeelding (resp. nulmatrix).

In de rest van dit hoofdstuk laten we V een complexe vectorruimte zijn. We gaan het volgende resultaat aantonen:

Stelling 3.3: Zij V een complexe eindig-dimensionale vectorruimte en $T : V \to V$ een lineaire afbeelding. Dan is er een unieke diagonaliseerbare afbeelding D en een unieke nilpotente afbeelding N zodanig dat $T = D + N$ en $DN = ND$. D en N zijn polynomen in T.

Opmerkingen: (1) Een analoog resultaat geldt uiteraard voor matrices: voor een $n \times n$-matrix A kunnen we de stelling immers toepassen op de afbeelding van $\mathbf{C}^n \to \mathbf{C}^n$ gegeven door $\mathbf{x} \to A\mathbf{x}$.
(2) De eis dat D en N commuteren is essentieel voor de uniciteit: als $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, dan is $A = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 3 & 0 \end{pmatrix}$, waarbij steeds de linkermatrix diagonaliseerbaar is en de rechtermatrix nilpotent (waarom?). De linker- en rechtermatrix in het 2e en 3e lid commuteren echter niet. In feite is A zelf diagonaliseerbaar, dus het diagonaliseerbare deel D van A is A zelf en het nilpotente deel N is O.

Gegeneraliseerde eigenvectoren. Een meetkundige interpretatie van de algebraïsche multipliciteit.

Laat dus V een eindig-dimensionale complexe vectorruimte en $T \in \mathcal{L}(V)$ zijn, m.a.w. T is een lineaire afbeelding van $V \to V$. Zij $\lambda \in \mathbf{C}$ een eigenwaarde van T. Dan is

$$\ker(T - \lambda \cdot id) \subset \ker(T - \lambda \cdot id)^2 \subset \ldots \ker(T - \lambda \cdot id)^{m(\lambda)-1} \subset \ker(T - \lambda \cdot id)^{m(\lambda)} = \ldots$$

een rij van inclusies. Immers als $(T - \lambda \cdot id)^m(v) = 0$, dan is zeker $(T - \lambda \cdot id)^{m+1}(v) = 0$.
Zoals we weten, geldt voor lineaire deelruimten U, W van een eindig-dimensionale vectorruimte V: als $U \subset W$ dan is $\dim(U) \leq \dim(W)$ en gelijkheid geldt alleen als $U = W$. Hieruit volgt dat in de rij van inclusies vanaf zekere index $k = m(\lambda)$ gelijkheid optreedt. Als eenmaal gelijkheid optreedt

dan zijn de volgende termen in de rij inclusies ook gelijk. Neem immers aan dat $\ker(T-\lambda\cdot id)^{k-1} = \ker(T-\lambda\cdot id)^k$ voor zekere k. Als nu geldt dat $v \in \ker(T-\lambda\cdot id)^{k+1}$ dan $(T-\lambda\cdot id)v \in \ker(T-\lambda\cdot id)^k = \ker(T-\lambda\cdot id)^{k-1}$ en dus is $v \in \ker(T-\lambda\cdot id)^k$. Dus $m(\lambda)$ (de index waarbij gelijkheid optreedt) is hoogstens gelijk aan $n = \dim(V)$. De lineaire deelruimte $E_\lambda := \ker(T-\lambda\cdot id)^{m(\lambda)}$ heet de *gegeneraliseerde eigenruimte* van T bij (de eigenwaarde) λ. Elementen van E_λ (op de nulvector na) heten *gegeneraliseerde eigenvectoren.* Evenals voor gewone eigenvectoren geldt:

Propositie 3.4: Gegeneraliseerde eigenvectoren van een endomorfisme $T : V \to V$ behorende bij verschillende eigenwaarden zijn lineair onafhankelijk.

Bewijs: Laat $\lambda_1, \ldots, \lambda_k$ de eigenwaarden van T zijn, en $v_1+\ldots+v_k = 0$ voor $v_1 \in E_{\lambda_1}, \ldots, v_k \in E_{\lambda_k}$. We moeten aantonen: $v_1 = \ldots = v_k = 0$. Noem $N_j = T - \lambda_j \cdot id$. Merk op dat de N_j's onderling commuteren, d.w.z. $N_iN_j = N_jN_i$. Laat verder $k_j \geq 0$ het kleinste gehele getal zijn zodanig dat $N_j^{k_j}v_j = 0$. Stel dat $v_1 \neq 0$, dus $k_1 > 0$. Beschouw de afbeelding $N_1' := N_1^{k_1-1}N_2^{k_2}\ldots N_k^{k_k}$. Omdat $N_1'(v_j) = 0$ voor $j = 2,\ldots,k$ is ook $N_1'(v_1) = 0$. Maar
$N_1'(v_1) = \prod_{j=2}^k(\lambda_1-\lambda_j)^{k_j}\cdot(T-\lambda_1)^{k_1-1}(v_1)$. Daar alle λ_j's verschillend zijn is $(T-\lambda_1)^{k_1-1}(v_1) = 0$, wat in tegenspraak is met de definitie van k_1. Dus moet $v_1 = 0$ zijn. Op analoge wijze volgt dat $v_2 = \ldots = v_k = 0$.) $\diamond$

Uit Propositie 3.4 volgt dat de som E van gegeneraliseerde eigenruimten zelfs een directe som is:

$$E = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_k}.$$

We tonen aan dat in feite geldt $E = V$. Merk eerst op dat T de gegeneraliseerde eigenruimten invariant laat: $T(E_\lambda) \subset E_\lambda$. Dus de quotiëntafbeelding $\bar{T} : V/E_\lambda \to V/E_\lambda$ is goed gedefinieerd voor elke eigenwaarde λ van T. Uit (1.6) volgt dat $\chi_T = \chi_{T|E_\lambda}\cdot\chi_{\bar{T}}$. $\bar{T}$ heeft dus geen andere eigenwaarden dan T. Verder heeft $T|E_\lambda$ geen andere eigenwaarden dan λ, d.w.z. $\chi_{T|E_\lambda}(x) = (x-\lambda)^m$ voor $m = \dim(E_\lambda)$. We laten zien dat λ geen eigenwaarde is van $\bar{T}$. Stel nl. $\bar{T}(\bar{v}) = \lambda\bar{v}$. Dan is $(T-\lambda)v \in E_\lambda$, dus $(T-\lambda)v$ is een gegeneraliseerde eigenvector van T met eigenwaarde λ, Maar dan is v zelf een gegeneraliseerde eigenvector met eigenwaarde λ (waarom?) dus $v \in E_\lambda$ en $\bar{v} = \bar{0}$. I.h.b. heeft $\chi_{\bar{T}}$ dus geen nulpunten λ meer; m.a.w., $m = \dim(E_\lambda)$ is precies de algebraïsche multipliciteit van λ. Daar de som van de algebraïsche multipliciteiten precies gelijk is aan $n = \dim(V)$ is $\dim(E) = \sum_{i=1}^{k}\dim(E_{\lambda_i}) = \dim(V)$ en dus is $E = V$. We hebben dus aangetoond:

Stelling 3.5. Zij V een eindig-dimensionale complexe vectorruimte en $T \in \mathcal{L}(V)$ een endomorfisme. Dan is V de directe som van de gegeneraliseerde eigenruimten van T en de algebraïsche multipliciteit van een eigenwaarde λ van T is gelijk aan de dimensie van de bijbehorende gegeneraliseerde eigenruimte. In het bijzonder volgt dat de meetkundige multipliciteit van een eigenwaarde - de dimensie van de gewone eigenruimte - nooit groter kan zijn dan de algebraïsche multipliciteit.

We zijn nu toe aan het bewijs van stelling 3.3. (Het gedeelte dat D en N polynomen zijn in T stellen we uit tot het eind van het hoofdstuk.)

Bewijs van stelling 3.3: Laat P_j de projectie zijn op E_{λ_j}. Dan geldt $P_j^2 = P_j$, $\sum_{j=1}^k P_j = id_V$ en $P_iP_j = 0$ als $i \neq j$. Dan is de restrictie $T|_{E_j} =: T_j = \lambda_j id_{E_j} + N_j$ met N_j nilpotent en

$$T = \sum_{j=1}^{k} T_j P_j = \sum_{j=1}^{k} \lambda_j P_j + \sum_{j=1}^{k} N_j P_j =: D + N, \tag{3.5}$$

waarbij $D = \sum_{j=1}^k \lambda_j P_j$ diagonaliseerbaar is en N nilpotent. $[N, D] = ND - DN = O$ volgt uit de eigenschappen van de projecties. Om de uniciteit aan te tonen nemen we aan dat $T = D' + N'$ met D' diagonaliseerbaar en N' nilpotent zodanig dat $[D', N'] = O$. Laat v een gegeneraliseerde eigenvector van D' zijn bij eigenwaarde λ. Daar ook $[T, D'] = O$, volgt dat $D'T(v) = TD'(v) = \lambda T(v)$. T beeldt dus de eigenruimte E'_λ van D' bij de eigenwaarde λ af in zichzelf. Laat T_λ de restrictie zijn van T tot E'_λ en P'_λ de projectie op E'_λ. Dan is $D' = \sum_\lambda \lambda P'_\lambda$, waarbij over de eigenwaarden van D' wordt gesommeerd, en $T = \sum_\lambda T_\lambda P'_\lambda$. Maar dan is $N' = \sum_\lambda (T_\lambda - \lambda \cdot id) P'_\lambda$ en omdat N' nilpotent is, is ook $N'_\lambda = T_\lambda - \lambda \cdot id$ voor elke eigenwaarde λ een nilpotent endomorfisme op E'_λ. Op E'_λ is dan $(T_\lambda - \lambda \cdot id)^m = O$ voor zekere $m > 0$ en dus is λ een eigenwaarde van T en de gegeneraliseerde eigenruimte E'_λ bij λ is dan bevat in E_λ. Maar omdat $V = \bigoplus_\lambda E_\lambda = \bigoplus_\lambda E'_\lambda$, is $E'_\lambda = E_\lambda$ en dus is $D' = \sum_\lambda \lambda P_\lambda = D$ en dus ook $N' = N$. $\diamond$

Opmerkingen: (i.) Uit het bewijs volgt dat het diagonaliseerbare deel D van de afbeelding T gelijk is aan $D = \sum_{j=1}^k \lambda_j P_j$ waarbij $\lambda_1, \ldots, \lambda_k$ de eigenwaarden van T zijn en P_j de projectie op de gegeneraliseerde eigenruimte van λ_j.

(ii.) Ten opzichte van een geordende basis $\mathcal{B}$ van gegeneraliseerde eigenvectoren (zo geordend dat eigenvectoren bij dezelfde eigenwaarden bij elkaar staan) is de matrix van T een blokdiagonaalmatrix:

$$T_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \lambda_1 I_{a_1} & O & \ldots & O \\ O & \lambda_2 I_{a_2} & \ldots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \ldots & \lambda_k I_{a_k} \end{pmatrix} + \begin{pmatrix} N_1 & O & \ldots & O \\ O & N_2 & \ldots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \ldots & N_k \end{pmatrix}. \tag{3.6}$$

Hierbij is de eerste matrix in het rechterlid een diagonaalmatrix met de eigenwaarden op de hoofddiagonaal (hierbij is a_j de algebraïsche multipliciteit van λ_j). Dit is de matrix $D_{\mathcal{B}}^{\mathcal{B}}$ van het diagonaliseerbare deel van D. De rechtermatrix is de matrix $N_{\mathcal{B}}^{\mathcal{B}}$ van het nilpotente deel N; de matrices N_j zijn nilpotente $a_j \times a_j$-matrices. Dat de linker- en rechtermatrix commuteren is evident.

De Jordan-normaalvorm.

De matrices N_j in (3.6) liggen op gelijkvormigheid na vast. Indien we andere bases van de gegeneraliseerde eigenruimten kiezen, blijft de matrix van het diagonaliseerbare deel hetzelfde - deze hangt alleen af van de eigenwaarden en hun algebraïsche multipliciteit - maar de matrices N_j gaan over in gelijkvormige matrices. Door geschikte bases van de gegeneraliseerde eigenruimten te kiezen, kunnen we een zo eenvoudig mogelijke vorm voor de matrices N_j proberen te vinden. Het volgende resultaat geldt:

Stelling 3.6. Laat V een eindig-dimensionale complexe vectorruimte zijn en $T : V \to V$ een lineaire afbeelding. Er bestaat een basis van V, bestaande uit gegeneraliseerde eigenvectoren van T, zodanig dat de matrix van T t.o.v. deze basis een blokdiagonaalmatrix $\text{diag}(J_1, J_2, \dots, J_m) = \begin{pmatrix} J_1 & O & \dots & O \\ O & J_2 & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & J_m \end{pmatrix}$ en waarbij de *Jordanblokken* J_j vierkante matrices van de vorm $J_k = \lambda_k I + N_k$ met $\lambda_k \in \mathbf{C}$ een eigenwaarde van T, I een eenheidsmatrix, en N_k een $i_k \times i_k$-matrix met $(N_k)_{ij} = 1$ als $j - i = 1$ en $(N_k)_{ij} = 0$ als $j - i \neq 1$ $(i, j = 1, \dots, i_k)$, dus $N_k = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$ (†).
De Jordanblokken $J_1, \dots, J_m$ zijn uniek bepaald op volgorde van de blokken na.

De in stelling 3.6 genoemde matrix heet een *Jordan-normaalvorm* van de afbeelding T. Elke matrix is dus gelijkvormig met een matrix in Jordan-normaalvorm. Een basis $\mathcal{B}$ van V met de eigenschap dat de matrix $T_{\mathcal{B}}^{\mathcal{B}}$ van T t.o.v. $\mathcal{B}$ een Jordan-normaalvorm is heet een *Jordanbasis*. Een Jordanbasis bestaat dus altijd uit gegeneraliseerde eigenvectoren van T.

Bewijs van Stelling 3.6: Omdat T de gegeneraliseerde eigenruimten E_{λ_i} invariant laat, d.w.z. $T(E_{\lambda_i}) \subset E_{\lambda_i}$, is het voldoende om de restricties N_{λ_i} van $T - \lambda_i \cdot id$ tot E_{λ_i} te bekijken. Neem dus $\lambda = \lambda_i$ vast, en laat $N = N_\lambda : E_\lambda \to E_\lambda$. De afbeelding N is nilpotent. Laat m het kleinste positieve gehele getal zijn zodanig dat $N^m = 0$, en laat $F_j = \ker(N^j)$. Dan is $F_0 = \{0\} \subset F_1 \subset \dots \subset F_m = E_\lambda$ en de inclusies zijn echte inclusies.
We construeren nu een basis van F_m op de volgende wijze: kies eerst een maximaal lineair onafhankelijk stelsel modulo F_{m-1} $\{c_1^1, \dots, c_{k_1}^1\}$ in F_m, m.a.w. $\{\overline{c_1^1}, \dots, \overline{c_{k_1}^1}\}$ is een basis van de quotiëntruimte F_m/F_{m-1}. Dan is $\{c_1^1, \dots, c_{k_1}^1, Nc_1^1, \dots, Nc_{k_1}^1\}$ lineair onafhankelijk modulo F_{m-2}. Vul dit stelsel aan met $\{c_2^2, \dots, c_{k_2}^2\}$ tot een maximaal lineair onafhankelijk stelsel modulo F_{m-2}. Noem dit nieuwe stelsel $\{b_1^2, \dots, b_{\ell_2}^2\}$. Nu gaan we op dezelfde manier verder: zij $\{b_1^p, \dots, b_{\ell_p}^p\}$ een maximaal lineair onafhankelijk stelsel modulo F_{m-p}, dan is $\{b_1^p, \dots, b_{\ell_p}^p, Nb_1^p, \dots, Nb_{\ell_p}^p\}$ lineair onafhankelijk modulo F_{m-p-1} en we vullen dit aan met $\{c_1^{p+1}, \dots, c_{k_{p+1}}^{p+1}\}$ tot een maximaal lineair onafhankelijk stelsel modulo F_{m-p-1}, dat we $\{b_1^{p+1}, \dots, b_{\ell_{p+1}}^{p+1}\}$ noemen. Uiteindelijk krijgen we zo een maximaal lineair onafhankelijk stelsel $\{b_1^m, \dots, b_{\ell_m}^m\}$ modulo F_0. Dit is een basis van $F_0 = E_\lambda$. We herordenen deze basiselementen in groepjes $N^{m-p}c_j^p, N^{m-p-1}c_j^p, \dots, c_j^p$ (merk op dat $N^{m-p+1}c_j^p = 0$). De matrix van N t.o.v. deze geordende basis heeft dan de vorm van een blokdiagonaalmatrix $\text{diag}(N_1, \dots, N_\ell)$ met N_j een matrix van de vorm (†), m.a.w. N_j is een Jordanblok. Merk op dat elke c_j^p aanleiding geeft tot een Jordanblok van grootte $m - p + 1$. $\diamond$

Elke complexe $n \times n$-matrix is dus gelijkvormig met een matrix in Jordan-normaalvorm. Omdat matrices waarbij tegelijkertijd de rijen en kolommen zijn gepermuteerd gelijkvormig zijn (zie opgave III.14) is de Jordan-normaalvorm van het endomorfisme T geheel bepaald op permutatie van de Jordanblokken na. Anderzijds geldt dat twee Jordan-normaalvormen die niet door een permutatie van de Jordanblokken in elkaar over gaan, niet gelijkvormig zijn; de Jordan-normaalvorm is dus (op permutatie van de blokken na) bepaald door het aantal Jordanblokken bij een gegeven eigenwaarde van een gegeven afmeting (de afmeting van een $\ell \times \ell$-Jordanblok is ℓ). In feite geldt het volgende resultaat:

Propositie 3.7. Zij J een Jordan-normaalvorm van de lineaire afbeelding $T : V \to V$. Het aantal Jordanblokken $J_i = aI + N_i$ bij eigenwaarde a van afmeting minstens k is gelijk aan rang$(T - a \cdot id)^{k-1}$ - rang$(T - a \cdot id)^k$.

Bewijs: J is de matrix van T t.o.v. een zekere basis van V. Laat $J_i = bI + N_i$ een Jordanblok van J zijn van afmeting ℓ_i (het is dus een $\ell_i \times \ell_i$-matrix). N_i heeft de vorm (†). Dan is $(J_i - aI)^k$ een Jordanblok van $(T - a \cdot id)^k$. Als $a \neq b$, dan is de rang van $(J_i - aI)^k$ gelijk aan ℓ_i: het is immers een bovendriehoeksmatrix met $(b - a)^k$ op de hoofddiagonaal. Als $a = b$, dan is $(J_i - aI)^k = N_i^k$. Ga na dat de rang van N_i^k gelijk is aan $\ell_i - k$ als $k \leq \ell_i$ en 0 als $k > \ell$. Het verschil tussen de rang van $(J_i - aI)^{k-1}$ en de rang van $(J_i - aI)^k$ is dus 1 als $a = b$ en $k \leq \ell_i$, en 0 in de andere gevallen. De rang van de matrix $(T - a \cdot id)^k$ is verder gelijk aan de som van de rangen van zijn Jordanblokken $(J_i - aI)^k$, en het verschil tussen de rang van $(T - a \cdot id)^k$ en de rang van $(T - a \cdot id)^{k-1}$ is dus gelijk aan het aantal Jordanblokken $J_i = bI + N_i$ met $b = a$ en grootte $\ell_i \geq k$. ⋄

Gevolg: Het aantal Jordanblokken van afmeting precies k bij een eigenwaarde a van T hangt alleen af van de rangen van $(T - a \cdot id)^m$ voor $m = 1, \ldots, n$. In het bijzonder is het aantal Jordanblokken bij een gegeven eigenwaarde a gelijk aan n - rang$(T - a \cdot id) = \dim(\ker(T - a \cdot id))$ en dit is weer gelijk aan de meetkundige multipliciteit van de eigenwaarde a. Merk nog op dat de algebraïsche multipliciteit van a gelijk is aan de som van de afmetingen van de Jordanblokken bij eigenwaarde a.

Voorbeeld: Laat $A = \begin{pmatrix} 3 & 1 & -1 & 1 & 0 \\ 0 & 3 & 2 & 2 & 1 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$. Het karakteristieke polynoom van A is $\chi_A(X) = -(X-3)^4(X-1)$. De eigenwaarde 1 heeft algebraïsche en meetkundige multipliciteit 1; voor een eigenvector geldt $(A - I)\mathbf{v} = \mathbf{0}$. Een eigenvector is dus $\mathbf{v}_5 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -2 \\ 4 \end{pmatrix}$. Er is uiteraard maar 1

Jordanblok met afmeting 1. De eigenwaarde 3 heeft algebraïsche multipliciteit 4. Er geldt

$$A-3I = \begin{pmatrix} 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 2 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}, (A-3I)^2 = \begin{pmatrix} 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, (A-3I)^3 = \begin{pmatrix} 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & -8 \end{pmatrix}.$$

De rangen van $(A-3I), (A-3I)^2$ en $(A-3I)^3$ zijn resp. 3, 2 en 1. Volgens Propositie 3.6 is het aantal Jordanblokken (van afmeting minstens 1) dus gelijk aan $5 - \text{rang}(A-3I) = 2$; dit is ook de meetkundige multipliciteit van de eigenwaarde 3. Het aantal Jordanblokken met afmeting minstens 2 is $\text{rang}(A-3I) - \text{rang}(A-3I)^2 = 1$, en het aantal Jordanblokken van afmeting minstens 3 is $\text{rang}(A-3I)^2 - \text{rang}(A-3I)^3 = 1$. Daar de som van de afmetingen van de Jordanblokken 4 is (de algebraïsche multipliciteit), is er één Jordanblok bij eigenwaarde 3 met afmeting 3 en één met afmeting 1. Een Jordan-normaalvorm is dus $\mathcal{J} = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$.

Tenslotte bepalen we een Jordanbasis. Merk op dat $\ker(A-3I)^3 = \{\mathbf{x} \in \mathbf{C}^5 : x_5 = 0\}$. Dit is de gegeneraliseerde eigenruimte E_3 bij eigenwaarde 3. Verder is $\ker(A-3I)^2 = \{\mathbf{x} \in \mathbf{C} : x_5 = 0, x_3 + x_4 = 0\}$ en $\ker(A-3I)^3 = \{\mathbf{x} \in \mathbf{C} : x_5 = x_3 + x_4 = x_2 - x_3 + x_4 = 0\}$. Laat $\mathbf{v}_3$ een vector zijn in $\ker(A-3I)^3$ die niet in $\ker(A-3I)^2$ ligt: neem $\mathbf{v}_3 = \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$. Dan is $\mathbf{v}_2 = (A-3I)\mathbf{v}_3 = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ en $\mathbf{v}_1 = (A-3I)\mathbf{v}_2 = (A-3I)\mathbf{v}_3 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. $\mathbf{v}_1$ ligt in $\ker(A-3I)$. Vul $\{\mathbf{v}_1\}$ aan tot een basis $\{\mathbf{v}_1, \mathbf{v}_4\}$ van $\ker(A-3I)$ door (bijvoorbeeld) $\mathbf{v}_4 = \begin{pmatrix} 0 \\ 2 \\ 1 \\ -1 \\ 0 \end{pmatrix}$ te kiezen (elke vector in $\ker(A-3I)$ die lineair onafhankelijk is van $\mathbf{v}_1$ voldoet). Nu is $E_3 = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ en $E_1 = \text{span}\{\mathbf{v}_5\}$. $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\}$ vormt een Jordanbasis bij de matrix A, m.a.w. $\mathcal{J}$ is de matrix van de afbeelding $\mathbf{x} \to A\mathbf{x}$ t.o.v. de basis $\mathcal{V}$.

Gelijkvormige matrices.

We hebben gezien dat elke complexe matrix gelijkvormig is met een matrix in Jordan-normaalvorm. Twee complexe $n \times n$-matrices zijn gelijkvormig dan en slechts dan als zij dezelfde Jordan-normaalvorm hebben. Twee matrices A en B zijn dus gelijkvormig als de rang van $(A - aI)^k$ gelijk is aan de rang van $(B - aI)^k$ voor alle $a \in \mathbf{C}$ en $k \in \mathbf{N}$. In het bijzonder hebben gelijkvormige matrices dezelfde eigenwaarden met dezelfde algebraïsche en meetkundige multipliciteiten, maar het omgekeerde is i.h.a. niet het geval.

Propositie 3.8: Laat A, B twee reële $n \times n$-matrices zijn die (complex) gelijkvormig zijn, m.a.w. er bestaat een complexe $n \times n$-matrix U zodanig dat $A = UBU^{-1}$. Dan bestaat er ook een reële matrix Z zodanig dat $A = ZBZ^{-1}$.

Bewijs: Schrijf $U = V + iW$ met reële matrices V, W. Als $W = O$ dan zijn we klaar. Neem dus aan dat $W \neq O$. Dan is $AV = VB$ en $AW = WB$ en dus ook $A(V + aW) = (V + aW)B$ voor $a \in \mathbf{C}$. We zoeken een reële a zodanig dat $V + aW$ inverteerbaar is, dus $\det(V + aW) \neq 0$. Nu is $\det(V + aW)$ hetzij een polynoom van graad hoogstens n in a, hetzij identiek gelijk aan 0. Het laatste geval doet zich echter niet voor omdat $\det U = \det(V + iW) \neq 0$. Dus zijn er hoogstens n reële getallen a met $\det(V + aW) = 0$. Conclusie: matrices die complex gelijkvormig zijn, zijn ook reëel gelijkvormig.

Voorbeelden: (1.) Omdat $\mathrm{rang}(A^T - aI)^k = \mathrm{rang}((A - aI)^k)^T = \mathrm{rang}(A - aI)^k$ voor alle a, b, k, is elke matrix A gelijkvormig met zijn getransponeerde A^T.

(2.) Zij A een reële 2×2-matrix met een niet-reële eigenwaarde $\lambda = re^{i\phi}$. Dan is $\overline{\lambda} = re^{-i\phi}$ ook een eigenwaarde. Beide eigenwaarden hebben algebraïsche multipliciteit 1. De matrix $B = r\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$ heeft dezelfde eigenwaarden als A en is (reëel) gelijkvormig met A (een Jordan-normaalvorm van beide is $\mathrm{diag}(re^{i\phi}, re^{-i\phi})$).

Minimumpolynoom, de stelling van Cayley-Hamilton.

Laat V en $T \in \mathcal{L}(V)$ als boven zijn. Als λ een eigenwaarde is van T dan laat $m(\lambda)$ het kleinste gehele getal m zijn zodat de restrictie van $(T - \lambda \cdot id)^m$ tot de gegeneraliseerde eigenruimte E_λ de nulafbeelding is. $m(\lambda)$ is tevens het kleinste getal n zodat $\ker(T - \lambda \cdot id)^n$ gelijk is aan $\ker(T - \lambda \cdot id)^{n+1}$. Merk op dat $m(\lambda)$ gelijk is aan de grootte van het grootste Jordanblok in de Jordan-normaalvorm van de matrix. In het bijzonder is dan $m(\lambda)$ kleiner of gelijk aan de algebraïsche multipliciteit van λ. Beschouw nu het polynoom $M(X) = \prod_\lambda (X - \lambda \cdot id)^{m(\lambda)}$ waarbij het product wordt genomen over de eigenwaarden van T (we kunnen overigens evengoed het product nemen over alle $\lambda \in \mathbf{C}$ omdat $m(\lambda) = 0$ als λ geen eigenwaarde is). Omdat alle factoren in $M(T)$ met elkaar commuteren is $M(T)v = 0$ als v in een gegeneraliseerde eigenruimte E_ν ligt (door de factoren commuteren kunnen we er dan immers voor zorgen dat de factor $(T - \nu \cdot id)^{m(\nu)}$ geheel aan de rechterkant staat in het product en deze factor annihileert elke $v \in E_\nu$. Maar omdat V de directe som is van de gegeneraliseerde eigenruimten E_λ, is $M(T)v = 0$ voor alle $v \in V$ en dus is $M(T)$ de nulafbeelding. $M(X)$ is het polynoom van laagste graad zodanig dat $M(T) = O$ (in het bijzonder is $M(X)$ op een voorfactor $a \in \mathbf{C}$ het unieke polynoom met deze eigenschap). $M(X)$ heet het *minimumpolynoom* van T.

Omdat $m(\lambda)$ kleiner of gelijk is aan de algebraïsche multipliciteit van λ is $M(X)$ een deler van het karakteristieke polynoom $\chi_T(X)$. Hieruit volgt onmiddellijk

Stelling 3.9 (Cayley-Hamilton). Laat V een eindig-dimensionale vectorruimte zijn en $T : V \to V$ een endomorfisme. Dan is $\chi_T(T) = O$.

Voorbeeld: Beschouw de matrix $A = \begin{pmatrix} 3 & 1 & -1 & 1 & 0 \\ 0 & 3 & 2 & 2 & 1 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ in het hierboven behandelde voorbeeld. De grootte van het grootste Jordanblok bij eigenwaarde 1 is 1, bij eigenwaarde 3 is het 3. Het minimumpolynoom bij A is dus $m_A(X) = (X-3)^3(X-1)$. Dit is inderdaad een deler van het karakteristieke polynoom $\chi_A(X) = -(X-3)^4(X-1)$.

Gemeenschappelijke eigenvectoren van commuterende endomorfismen.

Propositie 3.10. Zij V een eindig-dimensionale complexe vectorruimte. Laat $T, U \in \mathcal{L}(V)$ twee commuterende endomorfismen zijn. Dan is er een basis van V bestaande uit gemeenschappelijke gegeneraliseerde eigenvectoren van T en U. (m.a.w. er is een basis $\{f_1, \dots, f_n\}$ zodanig dat f_j een gegeneralizeerde eigenvector van zowel T als U is).

Bewijs: Zij λ een gegeneraliseerde eigenwaarde van T en E_λ de bijbehorende gegeneralizeerde eigenruimte. Voor $v \in E_\lambda$ geldt dat $(T-\lambda)^m(v) = 0$ voor zekere $m > 0$. Uit $(T-\lambda)^m U(v) = U(T-\lambda)^m(v) = 0$ volgt dat $U(E_\lambda) \subset E_\lambda$. De restricties $U|_{E_\lambda}$ zijn dus goed gedefinieerd. Kies een basis van E_λ, bestaande uit gegeneralizeerde eigenvectoren van U (deze bestaat uiteraard ook uit gegeneralizeerde eigenvectoren van T). Daar V de directe som van de gegeneralizeerde eigenruimten E_λ is, is de vereniging van deze bases een basis van V. Deze heeft de gewenste eigenschap. $\diamond$

Opmerking 1. In het geval dat T en U beide diagonaliseerbaar zijn zegt Prop.3.10 dat er dan een gemeenschappelijke basis van eigenvectoren van T en U bestaat.

Opmerking 2. Propositie 3.10 is te generalizeren naar het geval van meer dan twee commuterende afbeeldingen. De bewering luidt dan: als $T_1, \dots, T_k : V \to V$ paarsgewijs commuterende lineaire afbeeldingen zijn, dan is er een basis van V, bestaande uit gemeenschappelijke gegeneraliseerde eigenvectoren van $T_1, T_2, \dots T_k$.

De cirkels van Gershgorin.

In deze paragraaf leiden we een resultaat af over de ligging van eigenwaarden in het complexe vlak:

Propositie 3.11 (de cirkels van Gershgorin). Zij A een complexe $n \times n$-matrix met elementen a_{ij}, $i, j = 1, \dots, n$. Dan liggen alle eigenwaarden van A in de vereniging van de cirkelschijven $C_i : |z - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$.

Bewijs: Laat $\lambda \in \mathbf{C}$ een eigenwaarde van A zijn en $\mathbf{x}$ een bijbehorende eigenvector. Laat i de index zijn zodat x_i de grootste modulus heeft van alle componenten van $\mathbf{x}$, m.a.w. $|x_i| \geq |x_j|$ voor $j = 1, \dots, n$. Dan volgt uit $\sum_{j=1}^n a_{ij} x_j = \lambda x_i$ dat

$$(\lambda - a_{ii}) x_i = \sum_{j \neq i} a_{ij} x_j.$$

Delen door x_i en de modulus nemen geeft dan

$$|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|. \quad \diamond$$

De cirkels $|z - a_{ii}| = \sum_{j \neq i} |a_{ij}|$ heten de *cirkels van Gershgorin* bij A.

Appendix.

We bewijzen in deze appendix het resultaat (ook reeds genoemd in stelling 3.3) dat als $T : V \to V$ een lineaire afbeelding is op een eindig-dimensionale complexe vectorruimte V, dan zijn het diagonaliseerbare deel D en het nilpotente deel N van T polynomen in T. De polynomen zijn uniek als we eisen dat de graad hoogstens M is, waarbij M de graad is van het minimumpolynoom van T.

Bewijs: Laat A de matrix van T zijn t.o.v. een Jordanbasis. We geven het bewijs voor A. A is dus een matrix in Jordan-normaalvorm. Laat D het diagonaliseerbare deel van A zijn - D is dus een diagonaalmatrix met de eigenwaarden van A op de diagonaal - en N het nilpotente deel.

Neem aan dat A (en dus T) m verschillende eigenwaarden $a_1, \dots, a_m \in \mathbf{C}$ heeft. We bekijken eerst het geval dat A geen eigenwaarde 0 heeft. Laat m_ℓ het aantal eigenwaarden met (minstens) een Jordanblok van afmeting minstens ℓ (dus $m_1 = m$) en laat L de afmeting van het grootste Jordanblok zijn (dus $m_\ell = 0$ voor $\ell > L$).

We tonen aan dat de matrices $I, D, D^2, \dots, D^{n-1}$ lineair onafhankelijk zijn in de vectorruimte $\mathcal{M}(n \times n, \mathbf{C})$ van complexe $n \times n$-matrices voor $n = m$ maar niet voor $n > m$. Als

$$\lambda_1 I + \lambda_2 D + \ldots + \lambda_n D^{n-1} = O$$

voor zekere complexe getallen $\lambda_1, \dots, \lambda_n$, dan is

$$\lambda_1 + \lambda_2 a_j + \ldots + \lambda_n a_j^{n-1} = 0$$

voor $j = 1, \dots, m$ en omgekeerd. Dit zijn m vergelijkingen met n onbekenden. Als $n > m$ dan is er zeker een oplossing met niet alle $\lambda_j = 0$. Als $n = m$ dan is de coëfficiëntendeterminant een determinant van Vandermonde $V(a_1, \dots, a_m) \neq 0$; de enige oplossing is dus die met alle $\lambda_j = 0$. Dus $I, D, \dots, D^{m-1}$ zijn lineair onafhankelijk en D^m is een polynoom in $I, \dots, D^{m-1}$.

Op dezelfde wijze geldt dat de matrices $N^\ell, N^\ell D, \dots, N^\ell D^{n-1}$ lineair onafhankelijk zijn dan en slechts dan als $n \leq m_\ell$; $N^\ell D^{m_\ell}$ is dan een polynoom in $N^\ell D^j$ voor $0 \leq j < m_\ell$. Omdat $N^\ell D^j$ en $N^k D^i$ niet-nulelementen hebben op verschillende nevendiagonalen indien $k \neq \ell$ zijn de M matrices $N^\ell D^j$ met $0 \leq \ell < L$ en $0 \leq j < m_\ell$ lineair onafhankelijk, waarbij $M = \displaystyle\sum_{\ell=1}^{L} m_\ell$.

Laat n_j de afmeting van het grootste Jordanblok bij de eigenwaarde a_j zijn. Het minimumpolynoom van A heeft dan graad $M' = \sum_{j=1}^{m} n_j$. Dus zijn de matrices $I, A, \ldots, A^{M'-1}$ lineair onafhankelijk.

Nu geldt dat $M = M'$: beschouw de $m \times L$-matrix B met $B_{ij} = 1$ als er een Jordanblok van afmeting minstens j is bij de eigenwaarde a_i en $B_{ij} = 0$ anders. In de i-e rij staan nu precies n_i enen; in totaal zijn er M' enen. In de j-e kolom staan precies m_j enen, in totaal zijn dit M enen. Maar dan is $M = M'$.

Door $A = D + N$, $A^2 = (D + N)^2$ etc. uit te schrijven en termen $N^\ell D^j$ te herschrijven als sommen van termen $N^\ell D^k$ met $0 \leq k < m_\ell$, kunnen we $I, A, \ldots, A^{M-1}$ schrijven als lineaire combinaties van de M matrices $N^\ell D^j$ met $0 \leq \ell < L$ en $0 \leq j < m_\ell$. Dit geeft een lineaire afbeelding U van de vectorruimte van polynomen van graad hoogstens $M - 1$ in A (dus $\text{span}\{I, A, \ldots, A^{M-1}\}$) naar de vectorruimte van polynomen in N en D die wordt opgespannen door $I, D, \ldots, N^\ell D^j, \ldots, N^{L-1} D^{m_L - 1}$. Omdat $I, A, \ldots, A^{M-1}$ lineair onafhankelijk zijn, is U injectief en dus, omdat beide vectorruimten dimensie M hebben, is U inverteerbaar. Maar dan zijn $U^{-1}(D)$ en $U^{-1}(N)$ polynomen in A van graad hoogstens $M - 1$. ($U^{-1}(N)$ kan O zijn; $U^{-1}(D)$ niet).

Als A een eigenwaarde 0 heeft, dan beschouw $A - \mu I$ waarbij μ geen eigenwaarde van A is. $A - \mu I$ heeft diagonaliseerbaar deel $D - \mu I$ en nilpotent deel N. Uit het bovenstaande volgt dat $D - \mu I$ en N polynomen in $A - \mu I$ zijn; dus zijn D en N polynomen in A.

We sluiten af met een voorbeeld. Laat $A = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$. Dan is $m_1 = m = 2, m_2 = 2$ en $M = 4$. De matrices I, D, N, ND zijn lineair onafhankelijk, evenals I, A, A^2, A^3. Merk op dat $(D - I)(D - 2I) = O$ dus $D^2 = 3D - 2I$. Dan is

$$I = I, \; A = D + N, \; A^2 = (D + N)^2 = 3D + 2ND - 2I, \; A^3 = (D + N)^3 = 7D + 9ND - 6N - 6I.$$

Inverteren geeft dan

$$D = -2A^3 + 9A^2 - 12A + 6I, \quad N = 2A^3 - 9A^2 + 13A - 6I.$$

IV. INWENDIGE PRODUCTEN.

Inproducten op reële vectorruimten.

Definitie. Zij V een reële vectorruimte. Een *inwendig product* (of *scalair product*) $(\, , \,)$ op V is een afbeelding van $V \times V$ naar $\mathbf{R}$ (d.w.z. voor $v, w \in V$ is $(v, w) \in \mathbf{R}$) zodanig dat

a. (bilineariteit): voor elke $w \in V$ zijn de afbeeldingen $V \to \mathbf{R}$ gegeven door $v \to (v, w)$ en $v \to (w, v)$ lineair. M.a.w. $(v + v', w) = (v, w) + (v', w)$ en $(\lambda v, w) = \lambda(v, w)$ voor $v, w \in V$ en $\lambda \in \mathbf{R}$ en analoog voor de tweede component.

b. (symmetrie): $(v, w) = (w, v)$ voor $v, w \in V$

c. (positieve definietheid): $(v, v) > 0$ als $v \in V$ en $v \neq 0$.

Een vectorruimte met een inwendig product heet ook wel een *Euclidische* vectorruimte.

Voorbeelden: 1. $V = \mathbf{R}^n$. Voor $\mathbf{x} = (x_1, \ldots, x_n)^T$ en $\mathbf{y} = (y_1, \ldots, y_n)^T$ is $(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y} = x_1 y_1 + \ldots + x_n y_n$ een inwendig product, het standaard-inwendig product. Dit inproduct noteren we ook wel als $\mathbf{x} \cdot \mathbf{y}$.

2. $V = \mathbf{R}^n$. Laat $w_1, \ldots, w_n$ positieve getallen zijn. Dan is $(\mathbf{x}, \mathbf{y}) = w_1 x_1 y_1 + \ldots + w_n x_n y_n$ een inwendig product.

3. Zij $V = \mathbf{R}^n$. We zoeken de meest algemene vorm voor een inwendig product $(\, , \,)$ op V. Laat $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$ en $\mathbf{y} = \sum_{i=1}^n y_j \mathbf{e}_j$ vectoren in V zijn. Dan is wegens bilineariteit

$$(\mathbf{x}, \mathbf{y}) = (\sum_{i=1}^n x_i \mathbf{e}_i, \sum_{j=1}^n y_j \mathbf{e}_j) = \sum_{i,j=1}^n x_i y_j (\mathbf{e}_i, \mathbf{e}_j) = \sum_{i,j=1}^n a_{ij} x_i y_j$$

met $a_{ij} = (\mathbf{e}_i, \mathbf{e}_j)$. Laat A de matrix met elementen a_{ij} zijn. Dan is $(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$. Elke bilineaire vorm $(\, , \,)$ op $\mathbf{R}^n$ is dus van deze vorm. Als we tevens symmetrie eisen, dan is $\mathbf{x}^T A \mathbf{y} = (\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x}) = \mathbf{y}^T A \mathbf{x} = \mathbf{x}^T A^T \mathbf{y}$ en dus is A symmetrisch ($A = A^T$). Verder volgt uit positieve definietheid dat $\mathbf{x}^T A \mathbf{x} > 0$ voor alle $\mathbf{x} \neq \mathbf{0}$. Een symmetrische matrix met deze eigenschap noemen we een *positief-definiete* matrix. Positief-definiete matrices bestuderen we in hoofdstuk VII.

4. Laat $\langle \, , \, \rangle$ op $\mathbf{R}^2$ gegeven zijn door $\langle \mathbf{x}, \mathbf{y} \rangle = 2x_1 y_1 + 5x_2 y_2 + x_1 y_2 + x_2 y_1$. We kunnen deze uitdrukking schrijven in de vorm $\mathbf{x}^* B \mathbf{y}$ met $B = \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}$ en $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. De matrix B is symmetrisch. Omdat

$$\langle \mathbf{x}, \mathbf{x} \rangle = (x_1 - x_2)^2 + (x_1 + 2x_2)^2$$

is $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ en gelijkheid geldt als $x_1 - x_2 = 0, x_1 + 2x_2 = 0$ dus voor $x_1 = x_2 = 0$ (en dus $\mathbf{x} = \mathbf{0}$). De vorm is dus een inwendig product op $\mathbf{R}^2$.

Opmerking: In hoofdstuk 8 zullen we zien dat een symmetrische matrix positief definiet is dan en slechts dan als alle eigenwaarden positief zijn. Het is eenvoudig na te gaan dat dit voor de matrix B het geval is.

5. V is de vectorruimte van reële continue functies op het interval $[a, b]$ ($a < b$). Dan is $(f, g) = \int_a^b f(x) g(x) dx$ een inwendig product op V.

6. $V = \mathcal{M}(m \times n, \mathbf{R})$, de vectorruimte van reële $m \times n$-matrices. Een inwendig product op V wordt gegeven door $(A, B) = \operatorname{tr}(A^T B)$.

Niet elke vectorruimte heeft een inwendig product. Op eindig dimensionale vectorruimten kunnen we echter altijd een inwendig product definiëren, bijvoorbeeld zoals in voorbeeld 1, waarbij x_i, y_i de componenten t.o.v. een gegeven basis zijn.

Inproducten op complexe vectorruimten.

We bekijken nu het geval dat V een complexe vectorruimte is.

Definitie. Zij V een complexe vectorruimte. Een *(hermites) inwendig product* (of scalair product) $(\ ,\ )$ op V is een afbeelding van $V \times V$ naar $\mathbf{C}$ (d.w.z. voor $v, w \in V$ is $(v, w) \in \mathbf{C}$) zodanig dat

a. (sesquilineariteit): voor $w \in V$ is de afbeelding $V \to \mathbf{C}$ gegeven door $v \to (w, v)$ lineair, m.a.w. $(w, v + v') = (w, v) + (w, v')$ en $(w, \lambda v) = \lambda(w, v)$ voor $v, w \in V$ en $\lambda \in \mathbf{C}$. Voor elke $w \in V$ is de afbeelding gegeven door $v \to (v, w)$ *antilineair*, d.w.z. $(v + v', w) = (v, w) + (v', w)$ en $(\lambda v, w) = \overline{\lambda}(v, w)$ voor $v, w \in V$ en $\lambda \in \mathbf{C}$.

b. $(w, v) = \overline{(v, w)}$ voor $v, w \in V$

c. (positieve definietheid): $(v, v) > 0$ als $v \in V$ en $v \neq 0$.

Merk op dat de antilineariteit in de tweede component wordt geïmpliceerd door (gewone) lineariteit in de eerste component samen met eigenschap b. Verder volgt uit eigenschap b dat $(v, v) \in \mathbf{R}$ voor alle $v \in V$. Een vorm $(\ ,\ )$ met eigenschappen a en b heet *hermites*. Een eindig-dimensionale complexe vectorruimte met een hermites inwendig product heet ook wel een *unitaire* vectorruimte.

Voorbeelden: 1. $V = \mathbf{C}^n$. Voor $\mathbf{x} = (x_1, \ldots, x_n)^T$ en $\mathbf{y} = (y_1, \ldots, y_n)^T$ is $(\mathbf{x}, \mathbf{y}) = \overline{x_1} y_1 + \ldots + \overline{x_n} y_n$ een (hermites) inwendig product, het *standaard-hermites inproduct.* (Merk op: in de literatuur geldt antilineariteit soms juist voor de tweede component i.p.v. de eerste; overeenkomstig wordt dan $(\mathbf{x}, \mathbf{y}) = x_1 \overline{y_1} + \ldots + x_n \overline{y_n}$).

2. V is de vectorruimte van complexe continue functies op het interval $[a, b] \in \mathbf{R}$ $(a < b)$. Dan is $(f, g) = \int_a^b \overline{f(x)} g(x) dx$ een inwendig product op V.

3. Als in het geval van reële vectorruimten tonen we aan: De meest algemene sesquilineaire vorm op $\mathbf{C}^n$ is $(\mathbf{x}, \mathbf{y}) = \overline{\mathbf{x}}^T A \mathbf{y}$. Deze vorm is hermites als $A = A^*$ waarbij de *hermites geadjungeerde* A^* van A gedefinieerd is als de complex gecongujeerde van A^T. Een matrix A waarvoor $A = A^*$ heet een *hermitese* matrix. Merk op dat we $(\mathbf{x}, \mathbf{y})$ kunnen schrijven als $\mathbf{x}^* A \mathbf{y}$.

4. Laat $\langle\ ,\ \rangle$ op $\mathbf{C}^2$ gegeven zijn door

$$\langle \mathbf{x}, \mathbf{y} \rangle = 3\overline{x_1} y_1 + 2\overline{x_1} y_2 + 2\overline{x_2} y_1 + 6\overline{x_2} y_2.$$

We kunnen deze uitdrukking schrijven in de vorm $\mathbf{x}^* B \mathbf{y}$ met $B = \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix}$ en $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Als in het reële geval geldt dat een hermitese matrix positief definiet is precies indien alle eigenwaarden positief zijn. De matrix B is hermites en heeft positieve eigenwaarden 2 en 7. De vorm $\langle\ ,\ \rangle$ is dus een hermites inwendig product op $\mathbf{C}^n$.

5. $V = \mathcal{M}(m \times n, \mathbf{C})$. Een inproduct op V wordt gegeven door $(A, B) = \operatorname{tr}(A^* B)$.

Norm en afstand. De ongelijkheid van Schwarz.

Met behulp van een inwendig product kunnen we een norm en een afstand definiëren. Eerst geven we de definities van een norm, resp. een afstand.

Definitie: Een *norm* op een vectorruimte V over een lichaam K is een afbeelding $\|\ \| : V \to \mathbf{R}_{\geq 0}$ zodanig dat

i. $\|v\| > 0$ voor alle $v \in V$ en $v \neq 0$.

ii. $\|\lambda v\| = |\lambda| \|v\|$ voor $v \in V$ en $\lambda \in \mathbf{R}$.

iii. (driehoeksongelijkheid): Voor $v, w \in V$ geldt: $\|v + w\| \leq \|v\| + \|w\|$.

Merk op dat uit (ii) volgt dat $\|0\| = 0$. Als eigenschappen (ii) en (iii) gelden maar eigenschap (i) vervangen is door de eis dat $\|v\| \geq 0$ voor alle $v \in V$, dan spreken we van een *seminorm.* Een vectorruimte met een norm heet een *genormeerde* vectorruimte.

M.b.v. een norm kunnen we een afstand definiëren:

In het algemeen moet een afstandsfunctie $d(\ ,\ )$ aan de volgende eigenschappen voldoen:

a. $d(v, w) \geq 0$ en gelijkheid geldt slechts als $v = w$.

b. $d(v, w) = d(w, v)$.

c. $d(v, w) \leq d(v, u) + d(u, w)$ (driehoeksongelijkheid) voor $u, v, w \in V$.

Ga na dat $d(v, w) = \|v - w\|$ inderdaad aan deze eigenschappen voldoet, m.a.w. op een genormeerde vectorruimte kan een afstand worden gedefinieerd.

Als een inproduct $(\ ,\ )$ op V gegeven is dan definiëren we de (Euclidische) *norm* van een vector $v \in V$ d.m.v. $\|v\| = \sqrt{(v, v)}$. Deze norm, die door het inwendig product wordt geïnduceerd, noemen we wel de Euclidische norm. Om te laten zien dat dit inderdaad een norm is, moeten we nagaan dat aan de eigenschappen (i-iii) van normen voldaan is. Om dit aan te tonen gebruiken we de volgende cruciale ongelijkheid:

Propositie 4.1. (ongelijkheid van Schwarz) Zij V een Euclidische vectorruimte. Dan geldt de volgende ongelijkheid:

$$|(v, w)|^2 \leq (v, v)(w, w) \quad (v, w \in V). \tag{4.1}$$

Gelijkheid geldt dan en slechts dan als v en w lineair afhankelijk zijn.

Bewijs: Als $w = 0$ dan volgt de ongelijkheid direct. Neem dus aan dat $w \neq 0$. Dan geldt, wegens positieve definietheid,

$$0 \leq (v + \lambda w, v + \lambda w) = (v, v) + \lambda(v, w) + \bar{\lambda}(w, v) + |\lambda|^2 (w, w)$$

voor alle $\lambda \in \mathbf{C}$. Kies nu $\lambda = -\dfrac{(w,v)}{(w,w)}$. Dan volgt meteen dat $(w,v)(v,w) \leq (v,v)(w,w)$. Gelijkheid geldt slechts als $v + \lambda w = 0$, m.a.w. als v, w lineair afhankelijk zijn. $\diamond$

Gevolg 4.2: De afbeelding $\| \ \| : V \to \mathbf{R}$ definieert inderdaad een norm op V

Bewijs: Eigenschap (i) volgt uit de positieve definietheid van het inproduct: $(v,v) > 0$ als $v \neq 0_V$. Eigenschap (ii) volgt uit $(\lambda v, \lambda v) = \overline{\lambda}\lambda(v,v) = |\lambda|^2(v,v)$. Eigenschap (iii) is een gevolg van de ongelijkheid van Schwarz:

$$\|v+w\|^2 = (v+w, v+w) = (v,v) + (v,w) + (w,v) + (w,w) = (v,v) + 2\mathrm{Re}(v,w) + (w,w) \leq$$

$$\leq \|v\|^2 + 2|(v,w)| + \|w\|^2 \leq (\|v\| + \|w\|)^2. \quad \diamond$$

In een reële vectorruimte kunnen we nu de hoek θ tussen twee vectoren $v, w \in V$ (mits $v, w \neq 0$) op teken na definiëren door $\cos\theta = \dfrac{(v,w)}{\|v\|\|w\|}$. De ongelijkheid van Schwarz garandeert dat $-1 \leq \cos\theta \leq 1$ en dat $\theta = 0$ of π precies dan indien v en w lineair afhankelijk zijn. (Bij het college lineaire algebra 1 hebben we gezien dat dit voor $V = \mathbf{R}^2$ en $\mathbf{R}^3$ met het standaard-inproduct precies overeen komt met de gebruikelijke hoek.) Verder is $\theta = \pi/2$ precies in het geval dat $(v,w) = 0$ (voor $v, w \neq 0$).

Definitie: (i.) Zij V een vectorruimte met een inproduct. Twee vectoren $v, w \in V$ heten *orthogonaal* als $(v,w) = 0$.
(ii.) Een stelsel vectoren $v_1, \ldots, v_n \in V$ heet een *orthogonaal* stelsel als elk tweetal vectoren uit het stelsel orthogonaal is en geen van de vectoren de nulvector is. Als bovendien geldt dat $\|v_i\| = 1$ voor $i = 1, \ldots, n$ dan heet het stelsel *orthonormaal*.

Voorbeeld. De standaardbasis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ in K^n is een orthonormaal stelsel.

De methode van Gram-Schmidt en QR-decompositie.

De volgende procedure (de *methode van Gram-Schmidt*) levert een methode om uitgaande van een gegeven lineair onafhankelijk stelsel $\{v_1, \ldots, v_n\}$ een orthonormaal stelsel te maken zodanig dat beide stelsels dezelfde lineaire deelruimte opspannen: definieer $v_1', \ldots, v_n'$ als volgt:

$$v_1' = v_1,\ v_2' = v_2 - \frac{(v_1', v_2)}{(v_1', v_1')}v_1',\ v_3' = v_3 - \frac{(v_1', v_3)}{(v_1', v_1')}v_1' - \frac{(v_2', v_3)}{(v_2', v_2')}v_2', \ldots$$

Nu geldt (ga na): $(v_2', v_1') = 0, (v_3', v_1') = (v_3', v_2') = 0$ etc. Het stelsel $\{v_1', \ldots, v_n'\}$ is dus een orthogonaal stelsel en $\mathrm{span}\{v_1', \ldots, v_n'\} = \mathrm{span}\{v_1, \ldots, v_n\}$. Om een orthonormaal stelsel te verkrijgen moeten we door de normen delen: als $w_i = v_i'/\|v_i'\|$ $(i = 1, \ldots, n)$, dan is $\{w_1, \ldots, w_n\}$ een orthonormaal stelsel.

Deze orthonormalisatieprocedure geeft aanleiding tot een ontbinding van een willekeurige $m \times n$-matrix waarvan de kolomvectoren lineair onafhankelijk zijn: laat $V = (v_1 \ldots v_n)$ een $m \times n$-matrix zijn met kolomvectoren $v_1, \ldots, v_n$. Pas de methode van Gram-Schmidt toe op de kolomvectoren $v_1, \ldots, v_n$. Dit levert, met dezelfde notatie als hierboven, een orthogonaal stelsel $\{v_1', \ldots, v_n'\}$, waarbij voor $i = 1, \ldots, n$ de vector $v_i = r_{1i}v_1' + r_{i-1,i}v_{i-1}' + v_i'$ een lineaire combinatie is van $v_1', \ldots, v_i'$. In matrixvorm kunnen we dit schrijven als $V = V'R'$ waarbij $V' = (v_1' \ldots v_n')$ en R' de bovendriehoeksmatrix is met elementen $R_{ij}' = r_{ij}$ voor $i < j$ en $R_{ii}' = 1$. Laat vervolgens $w_i = v_i'/\|v_i'\|$ en $Q = (w_1, \ldots, w_n)$. Dan is $V = QR$ waarbij R de bovendriehoeksmatrix is met elementen $R_{ij} = R_{ij}'\|v_i'\|$. Dit heet de *QR-decompositie* van de matrix V. De kolomvectoren van de matrix Q vormen een orthonormaal stelsel.

Voorbeeld: Laat $V = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = (v_1, v_2)$. Dan vormen de vectoren

$$v_1' = v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_2' = v_2 - \frac{v_2 \cdot v_1'}{v_1' \cdot v_1'} v_1' = \begin{pmatrix} -1/2 \\ 1/2 \end{pmatrix}$$

een orthogonaal stelsel. Normaliseren geeft

$$w_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad w_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

en er geldt dat

$$v_1 = v_1' = \sqrt{2}w_1, \qquad v_2 = v_2' + \frac{3}{2}v_1' = \frac{3}{2}\sqrt{2}w_1 + \frac{1}{2}\sqrt{2}w_2$$

zodat

$$V = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{pmatrix} \begin{pmatrix} \sqrt{2} & \frac{3}{2}\sqrt{2} \\ 0 & \frac{1}{2}\sqrt{2} \end{pmatrix} = QR.$$

Representatie t.o.v. een orthonormale basis.

Zij V een eindig-dimensionale reële of complexe vectorruimte met een (hermites) inproduct $(\ ,\ )$. We tonen het volgende aan:

Propositie 4.3: Zij $\mathcal{N} = \{v_1, \ldots, v_n\}$ een orthonormale basis van V. Dan geldt

a. Voor $x \in V$ geldt $x = (v_1, x)v_1 + \ldots + (v_n, x)v_n$ m.a.w. de coördinaten van een vector x t.o.v. de basis $\mathcal{N}$ zijn gegeven door de inproducten met de basisvectoren van $\mathcal{N}$.

b. Voor $x, y \in V$ is

$$(x, y) = \overline{x_1}y_1 + \ldots + \overline{x_n}y_n = \overline{(v_1, x)}(v_1, y) + \ldots + \overline{(v_n, x)}(v_n, y).$$

(In het geval dat V een reële vectorruimte is, is uiteraard $x_i = \overline{x_i}$ voor alle i.)

c. Zij $T : V \to V$ een lineaire afbeelding en $A = T_{\mathcal{N}}^{\mathcal{N}}$ de matrix van T t.o.v. de basis $\mathcal{N}$. Dan is $A_{ij} = (v_i, T(v_j))$.

Bewijs:

a. Schrijf $x = x_1 v_1 + \ldots + x_n v_n$. Dan is

$$(v_j, x) = x_1(v_j, v_1) + \ldots + x_n(v_j, v_n) = x_j$$

daar $(v_j, v_i) = 1$ als $i = j$ en 0 als $i \neq j$.

b.

$$(x, y) = \Big(\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j \Big) = \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{x_i} y_j (v_i, v_j) = \sum_{i=1}^{n} \overline{x_i} y_i.$$

c. Daar $A = (T(v_1)_{\mathcal{N}} \ldots T(v_n)_{\mathcal{N}})$, is het element A_{ij} gelijk aan de i-e component (ten opzichte van de basis $\{v_1, \ldots, v_n\}$) van $T(v_j)$. Volgens (a) is dit gelijk aan $(v_i, T(v_j))$. $\diamond$

De geadjungeerde van een lineaire afbeelding.

Laat V, W vectorruimten over hetzelfde lichaam K met (hermites) inwendige producten $(\ ,\ )_V$ en $(\ ,\ )_W$ zijn. Zij $T : V \to W$ een lineaire afbeelding. De geadjungeerde T^* is een lineaire afbeelding van W naar V zodanig dat $(w, Tv)_W = (T^* w, v)_V$ voor alle $v \in V$ en $w \in W$. (In complexe vectorruimten wordt wel de benaming *hermites geadjungeerde* gebruikt om te benadrukken dat het hier gaat om de geadjungeerde t.a.v. een hermites (i.p.v. een Euclidisch) inproduct.) Merk op dat T^*, als hij bestaat, uniek is. Een afbeelding $T : V \to V$ waarvoor $T = T^*$ heet *zelfgeadjungeerd* of ook wel *hermites.*

Voorbeelden: 1. Laat $V = \mathbf{R}^n$ en $W = \mathbf{R}^m$ met het standaardinproduct en $T(\mathbf{x}) = A\mathbf{x}$ voor $\mathbf{x} \in \mathbf{R}^n$ waarbij A een zekere $m \times n$-matrix is (A is de standaardmatrix van T). Dan is $T^*(\mathbf{y}) = A^T \mathbf{y}$, m.a.w. de standaardmatrix van de geadjungeerde afbeelding is de getransponeerde van de matrix van de afbeelding zelf. Immers voor $\mathbf{x} \in \mathbf{R}^n$ en $\mathbf{y} \in \mathbf{R}^m$ is

$$(\mathbf{y}, T(\mathbf{x})) = (\mathbf{y}, A\mathbf{x}) = \mathbf{y}^T A \mathbf{x} = (A^T \mathbf{y})^T \mathbf{x} = (A^T \mathbf{y}, \mathbf{x}) = (T^*(\mathbf{y}), \mathbf{x}).$$

T is zelfgedadjungeerd dan en slechts dan als $A = A^T$, dus als A symmetrisch. Om deze reden heet T ook wel een *symmetrische afbeelding.*
2. Laat $V = \mathbf{C}^n$ en $W = \mathbf{C}^m$ met het standaard-hermites inproduct en $T(\mathbf{x}) = A\mathbf{x}$ voor $\mathbf{x} \in \mathbf{C}^n$ waarbij A een zekere $m \times n$-matrix is (A is weer de standaardmatrix van T). Dan is $T^*(\mathbf{y}) = A^* \mathbf{y}$, waarbij $A^* = \overline{A^T}$. Immers voor $\mathbf{x} \in \mathbf{R}^n$ en $\mathbf{y} \in \mathbf{R}^m$ is

$$(\mathbf{y}, T(\mathbf{x})) = (\mathbf{y}, A\mathbf{x}) = \mathbf{y}^* A \mathbf{x} = (A^* \mathbf{y})^* \mathbf{x} = (A^* \mathbf{y}, \mathbf{x}) = (T^*(\mathbf{y}), \mathbf{x}).$$

3. Beschouw een n-dimensionale complexe vectorruimte V met een inproduct $(\ ,\ )$. Zij $a \in V$ een gegeven vector en de lineaire afbeelding $T : V \to V$ wordt gegeven door $T(x) = x - i(a, x)a$. Dan is voor $x, y \in V$

$$(y, T(x)) = (y, x - i(a, x)a) = (y, x) - i(a, x)(y, a) = (y + i(a, y)a, x) = (T^*(y), x)$$

en de geadjungeerde afbeelding wordt dus gegeven door $T^*(y) = y + i(a, y)a$.

4. Laat V de vectorruimte zijn van oneindig vaak differentieerbare (reële) functies op een interval $[a, b]$ (met $a < b$) in $\mathbf{R}$, met de eigenschap dat voor $f \in V$ geldt dat $f(a) = f(b) = 0$. Ga na dat dit een vectorruimte is (het is een lineaire deelruimte van de vectorruimte van alle reële continue functies op $[a, b]$). Op V is een inproduct gegeven door $(f, g) = \int_a^b f(x)g(x)dx$. Laat $T : V \to V$ gegeven zijn door $T(f) = f'$. Dan is voor $f, g \in V$:

$$(g, T(f)) = \int_a^b g(x)f'(x)dx = -\int_a^b g'(x)f(x) = (T^*(g), f)$$

en de geadjungeerde afbeelding is dus $T^*(g) = -g'$.
De voorbeelden 1 en 2 zijn in zekere zin te generaliseren naar algemene vectorruimten:

Propositie 4.4 (de matrix van de geadjungeerde). Laat V een eindig-dimensionale reële (resp. complexe) vectorruimte met inwendig product (,) zijn. Zij $\mathcal{N} = \{v_1, \ldots, v_m\}$ een orthonormale basis van V en zij verder $T : V \to V$ een lineaire afbeelding. De matrix van de geadjungeerde T^* t.o.v. de basis $\mathcal{N}$ is de getransponeerde (resp. de hermites geadjungeerde) van de matrix $T^{\mathcal{N}}_{\mathcal{N}}$ van T t.o.v. de basis $\mathcal{N}$.

Bewijs: Laat $A = (A_{ij})_{i,j=1\ldots n}$ de matrix van T t.o.v. $\mathcal{N}$ zijn en $B = (B_{ij})$ de matrix van T^* t.o.v. $\mathcal{N}$. Uit Propositie 4.3 volgt dat $A_{ij} = (v_i, T(v_j))$ en $B_{ij} = (v_i, T^*(v_j))$. Dan

$$B_{ij} = (v_i, T^*(v_j)) = (T(v_i), v_j) = \overline{(v_j, T(v_i))} = \overline{A_{ji}} = A^*_{ij}$$

en dus is $B = A^*$. In het geval dat V een reële vectorruimte is, is uiteraard $A^* = A^T$. $\diamond$

Opmerking: Zij V een eindig-dimensionale complexe vectorruimte met een (hermites) inproduct. Uit het bovenstaande volgt: een afbeelding $T : V \to V$ is hermites dan en slechts dan als voor de matrix A van T t.o.v. een (willekeurige) orthonormale basis geldt dat $A = A^*$. Een matrix met deze eigenschap noemen we een *hermitese* matrix.

Opmerking: In de fysische literatuur wordt de notatie A^* vaak gebruikt voor de complex geconjugeerde van een matrix of een operator en voor de (hermites) geadjungeerde wordt de notatie $A^\dagger$ gebruikt.

Orthogonaal complement en orthogonale projectie.

Laat V een eindig-dimensionale vectorruimte met inwendig product zijn en W een lineaire deelruimte van V. Het *orthogonaal complement* $W^\perp$ van W is de lineaire deelruimte van V bestaande uit de vectoren $x \in V$ zodanig dat $(x, w) = 0$ voor alle $w \in W$. Ga na dat dit inderdaad een lineaire deelruimte is.

Propositie 4.5. Voor de dimensies geldt:

$$\dim(W) + \dim(W^\perp) = \dim(V). \tag{4.2}$$

Bewijs: Laat $\{w_1, \ldots, w_m\}$ een orthonormale basis van W zijn (m.b.v. de methode van Gram-Schmidt kunnen we altijd een orthonormale basis van W uit een gewone basis construeren). Beschouw de afbeelding $P_W : V \to V$ gegeven door

$$P_W(v) = (w_1, v)w_1 + \ldots + (w_m, v)w_m. \tag{4.3}$$

P_W is lineair, verder geldt $P_W(v) \in W$ en $P_W(w) = w$ als $w \in W$ dus i.h.b. is $\mathrm{im}(P_W) = W$. Verder is $\ker(P_W) = W^\perp$. De bewering volgt nu meteen uit (1.2). $\diamond$

Aangezien $W \cap W^\perp = \{0_V\}$ is V gelijk aan de directe som $W \oplus W^\perp$. De afbeelding P_W is een projectie op W (ga na). Omdat $W^\perp$ de projectierichting is, noemen we P_W de *orthogonale projectie* op W. Een orthogonale projectie is altijd zelfgeadjungeerd: $P_W = P_W^*$. Zelfs geldt

Stelling 4.6. De lineaire afbeelding $P : V \to V$ is een orthogonale projectie dan en slechts dan als $P^2 = P$ en $P^* = P$.

Voordat we het bewijs geven, bewijzen we eerst een ander resultaat, dat een verband geeft tussen geadjungeerde en orthogonaal complement:

Propositie 4.7. Zij $T : V \to W$ een lineaire afbeelding tussen vectorruimten met een inproduct. Dan geldt:

$$\ker(T^*) = (\mathrm{im}(T))^\perp \quad \text{en } \mathrm{im}(T^*) \subset (\ker(T))^\perp. \tag{4.4}$$

Als de dimensies van V en W eindig zijn, geldt zelfs $\mathrm{im}(T^*) = (\ker(T))^\perp$. Verder is $\mathrm{rang}(T) = \mathrm{rang}(T^*)$.

Bewijs: (i) De eerste identiteit volgt uit de volgende rij van equivalenties:

$$w \in (\mathrm{im}(T))^\perp \iff (w, Tv) = 0 \text{ voor alle } v \in V \iff (T^*w, v) = 0\ \forall v \in V \iff T^*(w) = 0.$$

In het bijzonder volgt uit (1.2) en (4.2) (in het geval dat de dimensies eindig zijn):

$$\dim(W) - \ \mathrm{rang}(T^*) = \dim\ \ker(T^*) = \dim(W) - \ \mathrm{rang}(T).$$

(ii) Laat $v \in\ \mathrm{im}(T^*)$ en $z \in\ \ker(T)$. Er bestaat een $w \in W$ zodanig dat $T^*(w) = v$. Dan is

$$(z, v) = (z, T^*(w)) = (T(z), w) = 0$$

en dus is $v \in (\ker(T))^\perp$. In het eindig-dimensionale geval geldt gelijkheid omdat de dimensies van $\mathrm{im}(T^*)$ en $(\ker(T))^\perp$ gelijk zijn (vergelijk gevolg 4 op blz.4). $\diamond$

Opgave: Laat V een e.d. vectorruimte zijn met een inproduct. Toon aan dat, als W een lineaire deelruimte is van W: $(W^\perp)^\perp = W$. Laat verder zien dat voor de geadjungeerde van een lineaire afbeelding geldt: $(T^*)^* = T$

Bewijs van Stelling 4.6: We hebben al eerder gezien dat een lineaire afbeelding P een projectie is precies wanneer $P^2 = P$. Het enige wat nog moet worden aangetoond is dat $\mathrm{im}(P) = (\ker(P))^\perp$ dan en slechts dan als $P = P^*$. Als $P = P^*$ dan volgt $\mathrm{im}(P) = (\ker(P))^\perp$ direct uit Propositie 4.7. Omgekeerd, als $\mathrm{im}(P) = (\ker(P))^\perp$ dan volgt uit Propositie 4.7 dat $\mathrm{im}(P) = \mathrm{im}(P^*)$ en $\ker(P) = \ker(P^*)$. Omdat P en P^* beide projecties zijn, is $P = P^*$ (een projectie ligt geheel vast door zijn kern en zijn beeld). $\diamond$

Opgave: Waarom is $P^* : V \to V$ een projectie als $P : V \to V$ een projectie is?

De matrix van een orthogonale projectie.

Zij V een eindig-dimensionale vectorruimte met een inproduct en een orthonormale basis $\mathcal{E} = \{e_1, \ldots, e_n\}$. en $P_W : V \to V$ de orthogonale projectie op de k-dimensionale lineaire deelruimte W. We bepalen de matrix $P_{\mathcal{E}}^{\mathcal{E}}$ van P_W t.o.v. de basis $\mathcal{E}$. Laat $\{f_1, \ldots, f_k\}$ een orthonormale basis van W zijn. Dan is voor $x \in V$

$$P_W(x) = (f_1, x)f_1 + \ldots + (f_k, x)f_k$$

en dus is volgens Prop.4.3b

$$P_W(x)_{\mathcal{E}} = (f_1, x)(f_1)_{\mathcal{E}} + \ldots + (f_k, x)(f_k)_{\mathcal{E}} =$$

$$= (f_1)_{\mathcal{E}}(f_1)^*_{\mathcal{E}} x_{\mathcal{E}} + \ldots + (f_k)_{\mathcal{E}}(f_k)^*_{\mathcal{E}} x_{\mathcal{E}}$$

dus

$$P_{\mathcal{E}}^{\mathcal{E}} = (f_1)_{\mathcal{E}}(f_1)^*_{\mathcal{E}} + \ldots + (f_k)_{\mathcal{E}}(f_k)^*_{\mathcal{E}} = FF^*$$

waarbij F de $n \times k$-matrix $((f_1)_{\mathcal{E}} \ldots (f_k)_{\mathcal{E}})$ is. Merk op dat, omdat $\{f_1, \ldots, f_k\}$ een orthonormaal stelsel is, $F^*F = I_n$.

Voorbeeld: Zij W de lineaire deelruimte van $\mathbf{R}^3$ die wordt opgespannen door de vectoren $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ en $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Een orthonormale basis van W wordt gegeven door $\{f_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, f_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}\}$.

Laat $F = (f_1, f_2) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$. De matrix van orthogonale projectie op W is dus $FF^* = FF^T = \frac{1}{3} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$.

Er bestaat ook een uitdrukking voor de matrix van P_W in termen van de matrix $A = ((a_1)_{\mathcal{E}} \ldots (a_k)_{\mathcal{E}})$, waarbij $\{a_1, \ldots, a_k\}$ een willekeurige basis is van W. Laat $\{f_1, \ldots, f_k\}$ een orthonormale basis van

W zijn, en F als boven gedefinieerd. Dan is $A = FR$ een QR-decompositie, waarbij R een inverteerbare rechterbovendriehoeksmatrix is. Dan is $I = F^*F = (AR^{-1})^*AR^{-1}$, dus $R^*R = A^*A$. Nu is

$$P_{\mathcal{E}}^{\mathcal{E}} = AR^{-1}(AR^{-1})^* = A(R^*R)^{-1}A^* = A^*(A^*A)^{-1}A.$$

Opmerking: Als $A = (\mathbf{a}_1, \ldots, \mathbf{a}_k)$ en $B = (\mathbf{b}_1, \ldots, \mathbf{b}_k)$ $n \times k$-matrices zijn, dan is $(A^*B)_{ij} = \mathbf{a}_i^*\mathbf{b}_j = (\mathbf{a}_i, \mathbf{b}_j)$, m.a.w. de elementen van de matrix A^*B zijn precies de standaard-inwendige producten van de kolomvectoren van A. De $k \times k$-matrix A^*A de *Gram-matrix* van A. De matrix A en zijn Gram-matrix hebben dezelfde rang (waarom?) I.h.b. is de Gram-matrix A^*A inverteerbaar precies indien $k \leq n$ en A rang k heeft.

Toepassing; De methode van kleinste kwadraten. In het geval dat A een $n \times k$-matrix is met $n > k$, heeft de vergelijking $A\mathbf{x} = \mathbf{b}$ niet voor elke $\mathbf{b} \in \mathbf{R}^n$ een oplossing $\mathbf{x} \in \mathbf{R}^k$. In het geval dat de rang van A maximaal, dus k, is, is er wel een unieke "beste benadering". $\mathbf{x} \in \mathbf{R}^k$ beschouwen we als de beste benadering van de oplossing indien $\|A\mathbf{x} - \mathbf{b}\|$ minimaal is. $A\mathbf{x}$ is dus de orthogonale projectie van $\mathbf{b}$ op de kolomruimte van A en uit het voorafgaande volgt meteen dat dan $\mathbf{x} = (A^*A)^{-1}A^*\mathbf{b}$. De matrix $(A^*A)^{-1}A^*$ heet de *pseudoinverse* van A. Een voorbeeld hiervan is het volgende: laat $(x_1, y_1), \ldots, (x_n, y_n)$ een rij (reële) meetpunten zijn, waarbij $x_1, \ldots, x_n$ verschillende getallen zijn, de "invoerwaarden" en $y_1, \ldots, y_n$ de meetwaarden. Gezocht wordt een lijn $y = cx + d$ die "zo goed mogelijk" bij de meetpunten past. We bepalen c, d zo, dat de som van de kwadraten $\sum_{j=1}^n (y_j - cx_j - d)^2$ minimaal is. Dit komt neer op het vinden van een vector $\mathbf{x} = \begin{pmatrix} c \\ d \end{pmatrix}$ zodanig dat $\|A\mathbf{x} - \mathbf{b}\|$ minimaal is, waarbij

$$A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \quad \text{en} \quad \mathbf{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Deze procedure heet de *methode van kleinste kwadraten.*

Orthogonale en unitaire afbeeldingen.

Definitie: Een complexe $n \times n$-matrix U waarvoor geldt dat $U^*U = I_n$, heet een *unitaire* matrix. Een reële $n \times n$-matrix Q waarvoor geldt dat $Q^TQ = I_n$ heet een *orthogonale* matrix.

Propositie 4.8: De volgende beweringen zijn equivalent:

i. U is een unitaire (resp. orthogonale) matrix.
ii. De kolomvectoren van U vormen een orthonormaal stelsel in $\mathbf{C}^n$ (resp. $\mathbf{R}^n$).
iii. De rijvectoren van U vormen een orthonormaal stelsel in $\mathbf{C}^n$ (resp. $\mathbf{R}^n$).

Bewijs: De equivalentie van (i) en (ii) volgt direct omdat U^*U de Gram-matrix is van de kolomvectoren van A. De equivalentie van (i) en (iii) volgt uit het feit dat U^T unitair (resp. orthogonaal) is als U unitair (resp. orthogonaal) is. Dit laatste volgt uit $(U^T)^* = (U^*)^T = (U^{-1})^T = (U^T)^{-1}$.$\diamond$

Definitie: Laat V een complexe (resp. reële) vectorruimte zijn met een inwendig product. Een lineaire afbeelding $T \in \mathcal{L}(V)$ heet *unitair* (resp. *orthogonaal*) als $T^*T = id_V$.

Opmerking: Uit Propositie 4.4 volgt onmiddellijk voor het geval dat V een eindig-dimensionale vectorruimte is, dat T unitair resp. orthogonaal is dan en slechts dan als de matrix van T t.o.v. een orthonormale basis van V een unitaire resp. een orthogonale matrix is.

Propositie 4.9: Laat V een complexe (resp. reële) eindig-dimensionale vectorruimte met inwendig product zijn. De volgende beweringen zijn equivalent:

i. $T \in \mathcal{V}$ is unitair (resp. orthogonaal).
ii. $\|Tx\| = \|x\|$ voor alle $x \in V$.
iii. $(Tx, Ty) = (x, y)$ voor alle $x, y \in V$.

Bewijs: (i.$\Rightarrow$ ii.) Laat $x \in V$. Dan

$$\|Tx\|^2 = (Tx, Tx) = (x, T^*Tx) = (x, x) = \|x\|^2.$$

(ii.$\Rightarrow$ iii.) Laat $x, y \in V$ en $\lambda \in \mathbf{C}$. Dan volgt uit

$$(T(x + \lambda y), T(x + \lambda y)) = (x + \lambda y, x + \lambda y)$$

dat

$$(Tx, Tx) + 2\text{Re } \lambda(Tx, Ty) + |\lambda|^2(Ty, Ty) = (x, x) + 2\text{Re } \lambda(x, y) + |\lambda|^2(y, y)$$

dat

$$\text{Re } \lambda(Tx, Ty) = \text{Re } \lambda(x, y)$$

voor $\lambda \in \mathbf{C}$. Door $\lambda = 1$, resp. $\lambda = -i$ te kiezen, vinden we dat

$$\text{Re } (Tx, Ty) = \text{Re } (x, y) \qquad \text{en Im } (Tx, Ty) = \text{Im } (x, y).$$

(iii.$\Rightarrow$ i.) Uit $(Tx, Ty) = (x, y)$ volgt dat $(x, T^*Ty) = (x, y)$, en dus $(x, T^*Ty - y) = 0$ voor alle $x, y \in V$. Maar dan is $T^*T = id_V$. $\diamond$

Voorbeeld. Laat V een eindig-dimensionale vectorruimte zijn met een inproduct en W een lineaire deelruimte. $P_W : V \to V$ is de orthogonale projectie op W. Laat $S_W = 2P_w - id_V$. Dan is $S_W = S_W^*$ en $S_W^2 = id_V$. S_W is dus hermites en unitair (resp. orthogonaal). S_W heet een orthogonale spiegeling in W.

We leiden een spectraalstelling af voor unitaire en orthogonale afbeeldingen.

Stelling 4.10: Zij V een complexe vectorruimte van eindige dimensie met inproduct en $T : V \to V$ een unitaire afbeelding. Dan geldt:

a. Elke eigenwaarde van T heeft modulus 1.

b. Eigenvectoren behorende bij verschillende eigenwaarden van T zijn orthogonaal.

c. Zij W een lineaire deelruimte van W die invariant is onder T, d.w.z. $T(W) \subset W$. Dan is $W^\perp$ invariant onder T.

d. T heeft een orthonormale basis van eigenvectoren.

Bewijs:

a. Zij $Tx = \lambda x$ voor $x \neq 0$. Dan is $\|x\| = \|Tx\| = |\lambda|\|x\|$, dus $|\lambda| = 1$.

b. Zij $Tx = \lambda x$ en $Ty = \mu y$. Dan is

$$(x, y) = (Tx, Ty) = \overline{\lambda}\mu(x, y).$$

Als λ en μ verschillende eigenwaarden van T zijn, dan is $\overline{\lambda}\mu \neq 1$ en dus $(x, y) = 0$.

c. Merk op dat de restrictie $T|_W : W \to W$ unitair is en dus inverteerbaar. Laat nu $x \in W^\perp$. Dan is $(Tx, w) = (x, T^{-1}w) = 0$ voor $w \in W$ en dus $Tx \in W^\perp$.

d. We bewijzen dit met inductie naar de dimensie n van V. Voor $n = 1$ is de bewering waar. Neem aan dat de bewering waar is als $n < N$. Laat nu $\dim(V) = N$ zijn. Zij x een eigenvector van T met $\|x\|$ en W is de lineaire deelruimte $\mathrm{span}\{x\}$. Volgens (c) is $T(W^\perp) = W^\perp$ en de restrictie $T|_{W^\perp}$ heeft volgens de inductieveronderstelling een orthonormale basis van eigenvectoren. Deze basis, aangevuld met x vormt een orthonormale basis van eigenvectoren van T. $\diamond$

Gevolg 4.11: Zij U een complexe $n \times n$-matrix. Dan is er een unitaire matrix V en een diagonaalmatrix D zodanig dat $U = VDV^*$. We zeggen dat U *unitair gelijkvormig* is met een diagonaalmatrix.

Opmerking: Een lineaire afbeelding $T \in \mathcal{L}(V)$ voor V een complexe resp. reële vectorruimte die een orthonormale basis van eigenvectoren heeft heet *unitair* (resp. *orthogonaal*) *diagonaliseerbaar.*

Uit de spectraalstelling voor unitaire afbeeldingen leiden we een spectraalstelling voor orthogonale afbeeldingen op een reële vectorruimte af.

Stelling 4.12: Zij V een reële eindig-dimensionale vectorruimte met een inproduct en $T \in \mathcal{L}(V)$ een orthogonale afbeelding. Dan is er een orthonormale basis van V zodanig dat de matrix van T t.o.v. deze basis de vorm

$$\begin{pmatrix} I_k & O & O & \dots & O \\ O & -I_\ell & O & \dots & O \\ O & O & R(\phi_1) & \dots & O \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ O & O & O & \dots & R(\phi_m) \end{pmatrix} \tag{4.5}$$

heeft, waarbij $R(\phi)$ de 2×2-matrix $\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$ is.($\phi \in \mathbf{R}$)

Opmerking: 1. Omdat $R(0) = I_2$ en $R(\pi) = -I_2$ kunnen we in bovenstaande uitdrukking aannemen dat $k, \ell \in \{0, 1\}$.

Stelling 4.12 is equivalent met de bewering dat elke (reële) orthogonale matrix orthogonaal gelijkvormig is met een matrix van de vorm (4.5). Zonder beperking der algemeenheid kunnen we dus veronderstellen dat $V = K^n$ en $T : V \to V$ de afbeelding $T : \mathbf{x} \to Q\mathbf{x}$ is, waarbij Q een orthogonale matrix is. Laat $V = \mathbf{C}^n$ met het standaard-hermites inproduct zijn. Voor elke lineaire deelruimte W van V definiëren we $W_R = \{\mathbf{x} \in W : (\mathbf{x}, \mathbf{e}_j) \in \mathbf{R} \text{ voor } 1 \le j \le n\}$. I.h.b. is V_R=span$_{\mathbf{R}}\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ een reële vectorruimte isomorf met $\mathbf{R}^n$ en W_R een reële lineaire deelruimte van V_R. De restrictie van het hermites standaard-inproduct tot V_R geeft het standaard-inproduct op V_R (m.b.t. de basis $\{e_1, \ldots, e_n\}$). I.h.b. is $(W_R)^\perp = (W^\perp)_R$ waarbij het orthogonaal complement wordt genomen t.a.v. het inproduct op V_R, resp. op V.

Lemma 4.13: Zij V een complexe vectorruimte en W een lineaire deelruimte van V. Dan is de dimensie van W gelijk aan de (reële) dimensie van W_R dan en slechts dan als $W = \overline{W}$, m.a.w. voor elke $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n \in W$ is ook $\overline{\mathbf{x}} = \overline{x_1}\mathbf{e}_1 + \ldots + \overline{x_n}\mathbf{e}_n \in W$.

Bewijs: Zij P_W de orthogonale projectie op W. Als $\{f_1, \ldots, f_k\}$ een orthonormale basis is van W, dan is de matrix van P_W gelijk aan $\sum_{i=1}^k f_j f_j^*$. Als W invariant is onder complexe conjugatie, dan is $\{\overline{f_1}, \ldots, \overline{f_k}\}$ eveneens een orthonormale basis van W, en dus is de matrix van P_W gelijk aan $\sum_{i=1}^k \overline{f_j f_j}^*$. De matrix van P_W is dus een reële matrix. Nu is de dimensie van W gelijk aan de rang van P_W. Als P_W reëel is, dan is W_R het beeld van P_W en de dimensie van W_R is dus gelijk aan de rang van P_W. Omgekeerd, als de dimensies van W en W_R gelijk zijn dan is elke basis van W_R ook een basis van W. Hieruit volgt meteen dat als $\mathbf{x} \in W$, dan $\overline{\mathbf{x}} \in W$. $\diamond$

Bewijs van Stelling 4.12: Laat $\hat{Q} : V \to V$ de afbeelding $\mathbf{x} \to Q\mathbf{x}$ zijn en $\hat{Q}_R : V_R \to V_R$ de afbeelding $\mathbf{x} \to Q\mathbf{x}$. $\hat{Q}$ is unitair en we passen Stelling 4.10 toe. We passen inductie toe naar de dimensie n van V. Voor $n = 1$ valt er niets te bewijzen. Neem aan dat de bewering waar is voor vectrorruimten van dimensie kleiner dan n. We onderscheiden twee gevallen:

1. $\hat{Q}$ heeft alleen reële eigenwaarden. Dan is $V = V_1 \oplus V_{-1}$ de directe som van de eigenruimten bij eigenwaarden 1 resp. -1 en de beide eigenruimten zijn onderling orthogonaal. Omdat Q een reële matrix is, hebben $(V_1)_R$ en $(V_{-1})_R$ dezelfde dimensies als V_1 resp. V_{-1} en dus is $V_R = (V_R)_1 \oplus (V_R)_{-1}$. Beide eigenruimten zijn ook onderling orthogonaal. Er is dus een orthonormale basis van eigenvectoren van $\hat{Q}_R$ en t.o.v. deze basis heeft de matrix van $\hat{Q}_R$ de vorm (4.5) met $k + \ell = n$ en $m = 0$.

(2). $\hat{Q}$ heeft een niet-reële eigenwaarde $e^{i\phi}$. Laat $\mathbf{x}$ een eigenvector zijn. We kunnen $\mathbf{x} = \mathbf{y} + i\mathbf{z}$ schrijven met $\mathbf{y}, \mathbf{z} \in V_R$. Merk op dat de complex geconjugeerde $\overline{\mathbf{x}} = \mathbf{y} - i\mathbf{z}$ eigenvector is bij de eigenwaarde $e^{-i\phi}$. Uit

$$Q(\mathbf{y} + i\mathbf{z}) = (\cos\phi + i\sin\phi)(\mathbf{y} + i\mathbf{z})$$

volgt dat

$$Q\mathbf{y} = \cos\phi\mathbf{y} - \sin\phi\mathbf{z}, \quad Q\mathbf{z} = \sin\phi\mathbf{y} + \cos\phi\mathbf{z}.$$

Verder volgt uit

$$0 = (\overline{\mathbf{x}}, \mathbf{x}) = (\mathbf{y} + i\mathbf{z})^T(\mathbf{y} + i\mathbf{z})$$

dat $\mathbf{y}^T\mathbf{z} = 0$ en $\mathbf{y}^T\mathbf{y} = \mathbf{z}^T\mathbf{z}$, dus door $\mathbf{y}$ en $\mathbf{z}$ zo nodig met eenzelfde factor te vermenigvuldigen, kunnen we aannemen dat $\{\mathbf{y}, \mathbf{z}\}$ een orthonormaal stelsel is. De lineaire deelruimte W' van V_R opgespannen door $\mathbf{y}$ en $\mathbf{z}$ is invariant onder $\hat{Q}_R$ en de restrictie van $\hat{Q}_R$ tot W' heeft t.o.v. de basis $\{\mathbf{y}, \mathbf{z}\}$ de matrix $R(\phi)$. Verder laat $\hat{Q}_R$ het orthogonaal complement $(W'_R)^\perp = ((W')^\perp)_R$ invariant. We kunnen nu de inductie-onderstelling toepassen op de restrictie van $\hat{Q}_R$ tot $(W'_R)^\perp$. Een orthonormale basis van $(W'_R)^\perp$ verenigd met $\{\mathbf{y}, \mathbf{z}\}$ geeft een orthonormale basis van V_R. $\diamond$

We passen Stelling 4.12 toe op het geval dat $V = \mathbf{R}^n$ met $n = 2$ of $n = 3$:

$n = 2$: Voor een orthogonale afbeelding $T : V \to V$ zijn er twee mogelijkheden: (1) T heeft t.o.v. een orthonormale basis de matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$; T is in dit geval een orthogonale spiegeling. (2) De matrix van T t.o.v. een orthonormale basis is $R(\phi)$ voor zekere ϕ. T is in dit geval een rotatie om 0_V over een hoek ϕ. Merk op dat de matrix dezelfde vorm heeft voor elke willekeurige orthonormale basis. Gevallen 1 en 2 zijn verder te onderscheiden doordat in geval 1 de determinant van T gelijk is aan -1, in geval 2 is de determinant gelijk aan +1.

$n = 3$. In dit geval heeft T t.o.v. zekere orthonormale basis $\{f_1, f_2, f_3\}$ de vorm $\begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}$. T stelt een rotatie voor over een hoek ϕ om de as $\text{span}\{f_1\}$, in het geval dat $\pm 1 = -1$ wordt de rotatie nog gevolgd door een loodrechte spiegeling in het vlak opgespannen door f_2, f_3. In het eerste geval heet T een *draaiing* of *rotatie*, in het tweede geval heet T een *draaispiegeling*. Beide gevallen worden onderscheiden door het teken van de determinant van T ($\det(T) = \pm 1$). Verder is de rotatiehoek ϕ eenvoudig te bepalen via $\text{tr}(T) = \pm 1 + 2\cos\phi$.

V. DE DUALE VAN EEN VECTORRUIMTE

Laat V een eindig-dimensionale reële of complexe vectorruimte zijn met een (hermites) inwendig product $(\ ,\ )$. Voor een vaste $v \in V$ is de afbeelding $i_v : V \to K$ (waarbij $K = \mathbf{R}$ resp. $\mathbf{C}$) gegeven door $i_v(w) = (v, w)$ een lineaire afbeelding en dus een element van de vectoruimte $\mathcal{L}(V, K)$ (ga dit na). Omgekeerd, als $f \in \mathcal{L}(V, K)$, dan is er een $v \in V$ zodanig dat $f = i_v$. Immers, laat $\{v_1, \ldots, v_n\}$ een orthonormale basis zijn van V (zo'n basis bestaat altijd). f wordt geheel bepaald door de beelden $f(v_j) = a_j$ van de basisvectoren. Laat nu $v = \overline{a_1}v_1 + \ldots + \overline{a_n}v_n$. Dan is $i_v(v_j) = (v, v_j) = a_j = f(v_j)$ en dus is $f = i_v$.

Gevolg: de afbeelding $v \to i_v$ (voor $v \in V$) is een vectorruimte-isomorfisme tussen V en $\mathcal{L}(V, K)$. De vectorruimte $\mathcal{L}(V, K)$ heet de *duale vectorruimte* van V. Een gebruikelijke notatie voor $\mathcal{L}(V, K)$ is V^*. In het bijzonder geldt als $\dim(V) < \infty$:

$$\dim(V) = \dim(V^*). \tag{5.1}$$

Opmerking: Ook voor willekeurige vectorruimten V is de duale gedefinieerd als $V^* = \mathcal{L}(V, K)$. Als $\dim(V) = \infty$ levert de afbeelding $v \to i_v$ echter geen isomorfisme tussen V en V^*. Merk op dat i_v zelfs niet gedefinieerd is als V niet een vectorruimte met inproduct is. Zie opgave V.8 voor een voorbeeld.

Voor een willekeurige vectorruimte V bestaat er een natuurlijk (of *kanoniek*) isomorfisme tussen V en een lineaire deelruimte van $V^{**} = \mathcal{L}(V^*, K)$: zij immers $v \in K$. Dan is de afbeelding $v^\sharp : V^* \to K$ gegeven door $v^\sharp(f) = f(v)$ een lineaire afbeelding (ga na) en dus een element van V^{**}. De afbeelding die v afbeeldt op $v^\sharp$ is een injectieve lineaire afbeelding van V naar V^{**}, dus een vectorruimte-isomorfisme van V op een lineaire deelruimte van V^{**} die we dan met V kunnen identificeren. Deze afbeelding heet natuurlijk of kanoniek omdat deze niet afhangt van speciale keuzen (van een basis of een inproduct). Merk op dat dit niet het geval is voor de afbeelding $v \to i_v$ van V naar V^* omdat i_v afhangt van het inproduct in V. Merk nog op dat als $\dim(V)$ eindig is, dan is $\dim(V) = \dim(V^*) = \dim(V^{**})$ en dan is de afbeelding $v \to v^\sharp$ zelfs surjectief en dus een vectorruimte-isomorfisme. In dit geval kunnen we V^{**} en V identificeren.

Propositie 5.1 (duale basis). Zij $\{v_1, \ldots, v_n\}$ een basis van de vectorruimte V. Laat $f_1, \ldots, f_n \in V^*$ de lineaire afbeeldingen van V naar K zijn zodanig dat $f_j(v_i) = \delta_{ij}$ (d.w.z. $f_j(v_i) = 1$ als $i = j$ en $f_j(v_i) = 0$ als $i \neq j$). Dan is $\{f_1, \ldots, f_n\}$ een basis van V^*.

Bewijs: Omdat $\dim(V) = \dim(V^*)$, is het voldoende om aan te tonen dat $f_1, \ldots, f_n$ lineair onafhankelijk zijn. Neem hiertoe aan dat $0 = \lambda_1 f_1 + \ldots + \lambda_n f_n$ voor zekere $\lambda_1, \ldots, \lambda_n \in K$. Dan is

$$0 = (\lambda_1 f_1 + \ldots + \lambda_n f_n)(v_j) = \lambda_j \quad (j = 1, \ldots, n). \quad \diamond$$

$\{f_1, \ldots, f_n\}$ heet de *duale basis van* $\{v_1, \ldots, v_n\}$. Als V een inproduct heeft, dan zijn er vectoren $v^1, \ldots, v^n \in V$ zodanig dat $f_j = i_{v^j}$, m.a.w. $f_j(v) = (v^j, v)$. I.h.b. is $(v^j, v_i) = \delta_{ij}$. We noemen $\{v^1, \ldots, v^n\}$ eveneens de duale basis van $\{v_1, \ldots, v_n\}$. (In tegenstelling tot de duale basis $\{f_1, \ldots, f_n\}$ is de duale basis $\{v^1, \ldots, v^n\}$ dus een basis van V.)

De getransponeerde van een afbeelding. Laat V en W vectorruimten zijn en $T : V \to W$ een lineaire afbeelding. We definiëren de afbeelding $T' : W^* \to V^*$ d.m.v. $(T'f)(v) = f(Tv)$ waarbij $f \in W^*$ en $v \in V$. T' is weer een lineaire afbeelding en heet de *getransponeerde* of *pull-back* van T. Als V en W eindig-dimensionale vectorruimten zijn met een inproduct, dan bestaat er een nauw verband tussen de getransponeerde T' en de geadjungeerde T^*. Zoals we hebben gezien bestaat er voor elke $f \in W^*$ een $w \in W$ zodanig dat $f = i_w$. Dan is voor $v \in V$

$$(T'i_w)(v) = i_w(Tv) = (w, Tv) = (T^*w, v) = i_{T^*w}(v).$$

We hebben daarmee aangetoond:

Propositie 5.2. Laat V, W eindig-dimensionale vectorruimten zijn met inwendig product en $T \in \mathcal{L}(V, W)$. Dan is $T'i_w = i_{T^*w}$.

De annihilator van een lineaire deelruimte. Zij V een vectorruimte en U een lineaire deelruimte van V. De *annihilator* $U^\perp$ van U is de lineaire deelruimte van V' bestaande uit de lineaire afbeeldingen $f : V \to K$ zodanig dat $f(u) = 0$ voor alle $u \in U$. Als V een inwendig product heeft, dan is het orthogonaal complement $U^\perp$ van U een lineaire deelruimte van V. De notatie $U^\perp$ is dus dubbelzinnig. Wel geldt het volgende verband:

Opgave: Laat zien dat voor $w \in V$ geldt: $w \in U^\perp$ (lin. deelruimte van V) dan en slechts dan als $i_w \in U^\perp$ (lin. deelruimte van V').

Duale vectorruimte en tensorproduct. Laat V en W eindig-dimensionale vectorruimten over het lichaam K zijn. In opgave I.30 is aangetoond dat er een isomorfisme is tussen het tensorproduct $V \otimes W$ en de vectorruimte $\mathcal{L}(V, W)$ van lineaire afbeeldingen van V naar W. Zo'n isomorfisme is niet kanoniek, omdat de vorm afhangt van de basis en niet behouden blijft onder basistransformaties (verg. opgave V.9). Er bestaat echter wel een kanoniek isomorfisme $\phi : V^* \otimes W \to \mathcal{L}(V, W)$: daar $V^* \otimes W$ wordt opgespannen door tensorproducten $v' \otimes w$ met $v' \in V$ en $w \in W$ is het voldoende om ϕ vast te leggen op tensorproducten en vervolgens lineair tot $V^* \otimes W$ voort te zetten. Laat nu $\phi(v' \otimes w)(v) = v'(v)w$ voor $v \in V$. $\phi(v' \otimes w)$ is dan een lineaire afbeelding van V naar W. Merk op dat deze definitie onafhankelijk is van een basiskeuze. Het is niet moeilijk om aan te tonen dat ϕ goed gedefinieerd is en een vectorruimte-isomorfisme.

VI. GENORMEERDE VECTORRUIMTEN

De norm van een lineaire afbeelding.

Zij V een genormeerde vectorruimte en $T \in \mathcal{L}(V)$. De norm $\|T\|$ van T is gedefinieerd als

$$\|T\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|T(\mathbf{x})\|}{\|\mathbf{x}\|} = \sup_{\|\mathbf{x}\|=1} \|T(\mathbf{x})\|.$$

Als $\dim(V)$ eindig is, dan bestaat het supremum. I.h.a. noemen we T *begrensd* als het supremum bestaat. Merk op dat als T begrensd is, voor alle $\mathbf{x} \in V$ geldt dat $\|T(\mathbf{x})\| \leq \|T\|\|\mathbf{x}\|$. In het bijzonder is een begrensde lineaire afbeelding $T : V \to V$ (t.o.v. de norm) een continue functie op V: voor elke $\epsilon > 0$ is er een $\delta > 0$ zodat als $\|\mathbf{x} - \mathbf{y}\| < \delta$, dan is $\|T(\mathbf{x}) - T(\mathbf{y})\| < \epsilon$ voor $\delta = \dfrac{\epsilon}{\|T\|}$. Omdat de eenheidsbol $\|\mathbf{x}\| = 1$ in een eindig-dimensionale vectorruimte compact (d.w.z. begrensd en gesloten) is, neemt de $T(\mathbf{x})$ hier dan een maximum aan (immers een continue functie neemt op een compacte verzameling een maximum en een minimum aan). Als $\dim(V) < \infty$, dan kunnen we in de definitie van de norm dus het supremum vervangen door het maximum.

Propositie 6.1. De norm heeft de volgende eigenschappen ($T, S \in \mathcal{L}(V)$, begrensd.):

i. $\|T\| \geq 0$ en gelijkheid geldt slechts als T de nulafbeelding is.
ii. $\|\lambda T\| = |\lambda|\|T\|$ voor $\lambda \in K$.
iii. $\|T + S\| \leq \|T\| + \|S\|$
iv. $\|TS\| \leq \|T\|\|S\|$.

Bewijs: Eigenschappen (i)-(iii) volgen uit de definitie en de eigenschappen van de norm in V. Eigenschap (iv) bewijzen we als volgt: laat $\mathbf{x} \in V$, $\mathbf{x} \neq \mathbf{0}$. Dan is

$$\|TS(\mathbf{x})\| \leq \|T\|\|S(\mathbf{x})\| \leq \|T\|\|S\|\|\mathbf{x}\|,$$

en dus is $\|TS\| = \max\limits_{\|\mathbf{x}\|=1} \|TS(\mathbf{x})\| \leq \|T\|\|S\|$. $\diamond$

Opmerking: Als V een inwendig product heeft en de norm is geïnduceerd door het inproduct, dan is voor $T \in \mathcal{L}(V)$:
$\|T\| = \sup_{\|\mathbf{x}\|=\|\mathbf{y}\|=1} |(T(\mathbf{x}), \mathbf{y})|$. Hieruit volgt voor de geadjungeerde afbeelding: $\|T\| = \|T^*\|$ (zie opgave VI.3).

De norm van een $n \times n$-matrix A kunnen we op geheel analoge wijze definiëren als $\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$ waarbij $\mathbf{x} \in K^n$. Voor een reële matrix A is het maximum hetzelfde indien we $\mathbf{x} \in \mathbf{R}^n$ als indien we $\mathbf{x} \in \mathbf{C}^n$ nemen (ga na).

Voorbeelden: 1. Laat $P : V \to V$ een orthogonale projectie zijn, $P \neq O$. Dan is $\|P\|$=1 (zie opgave VI.3).
2. De norm van een orthogonale (resp. unitaire) afbeelding is 1.

Voor het berekenen van de norm van een lineaire afbeelding resp. een matrix is het volgende resultaat nuttig:

Propositie 6.2. Zij V een eindig-dimensionale vectorruimte met inproduct en $T : V \to V$ een lineaire afbeelding. Dan geldt: $\|T\|^2 = \|T^*T\|$.

Bewijs: Laat $x \in V$ met $\|x\| = 1$. Dan is

$$\|Tx\|^2 = (Tx, Tx) = (T^*Tx, x) \leq \|T^*Tx\| \leq \|T^*T\|.$$

Neem nu in het linkerlid het supremum over alle $x \in V$ met $\|x\| = 1$. Dan volgt $\|T\|^2 \leq \|T^*T\|$. Omgekeerd is $\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2$. $\diamond$

Opmerking: In hoofdstuk 7 zullen we zien dat voor een hermitese matrix (zoals T^*T) de norm gelijk is aan de modulus van de grootste eigenwaarde (zoals we zullen zien zijn de eigenwaarden van T^*T niet-negatieve reële getallen). De (niet-negatieve) wortels van de eigenwaarden van T^*T heten de *singuliere waarden* van T.
Conclusie: de norm van T is gelijk aan de grootste singuliere waarde van T.
Voorbeeld: Laat $J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Dan is $J^*J = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. De singuliere waarden van J zijn dus 0 en 1 en $\|J\| = 1$.

Het volgende verband geldt tussen de norm van een matrix en zijn elementen:

Propositie 6.3. Zij A een $m \times n$-matrix met elementen A_{ij}. Dan geldt

$$\max_{i,j} |A_{ij}|^2 \leq \|A\|^2 \leq \sum_{i,j} |A_{ij}|^2. \tag{6.1}$$

Bewijs: $A\mathbf{e}_j$ is gelijk aan de j-e kolomvector van A. I.h.b. is dus $\max_{i=1}^n |A_{ij}| \leq \|A\mathbf{e}_j\| \leq \|A\|$. Neem nu het maximum over $j = 1, \ldots, n$. Hieruit volgt de eerste ongelijkheid. Voor de tweede ongelijkheid laat $\|\mathbf{x}\| = 1$. Dan is

$$\|A\mathbf{x}\|^2 \leq \sum_{i=1}^n \left(\sum_{j=1}^n |A_{ij}x_j| \right)^2 \leq \sum_{i=1}^n \sum_{j=1}^n |A_{ij}|^2 \sum_{j=1}^n |x_j|^2 = \sum_{i,j=1}^n |A_{ij}|^2.$$

De tweede ongelijkheid in bovenstaande uitdrukking is precies de ongelijkheid van Schwarz. $\diamond$

Gevolg 6.4. Laat $\{A^{(n)}\}_{n=1}^\infty$ een rij $m \times n$- matrices zijn met elementen $A_{ij}^{(n)}$. Dan geldt:

$$\lim_{n\to\infty} A_{ij}^{(n)} = 0 \text{ voor alle } i, j \Leftrightarrow \lim_{n\to\infty} \|A^{(n)}\| = 0.$$

We hebben in het vorige hoofdstuk gezien dat een inwendig product op een vectorruimte een norm induceert d.m.v. $\|v\| = \sqrt{(v, v)}$. Deze norm heet de *Euclidische* norm of de 2-norm. Er zijn ook vectorruimten met een norm die niet van een inwendig product afkomstig zijn. We noemen twee voorbeelden van normen op K^n:
1. De 1-norm op K^n: $\|\mathbf{x}\| = \sum_{i=1}^n |x_i|$ waarbij x_i de i-de component van de vector $\mathbf{x}$ is.

2. De sup-norm (of maximum-norm) op K^n: $\|\mathbf{x}\| = \max_{i=1,\dots,n} |x_i|$.

In het geval van een eindig-dimensionale vectorruimte V zijn alle normen *equivalent*, d.w.z. als $\| \|_1$ en $\| \|_2$ twee normen zijn, dan bestaan er positieve getallen a en b zodanig dat voor elke $x \in V$ geldt dat $a\|x\|_1 \leq \|x\|_2 \leq b\|x\|_1$. In het geval van oneindig-dimensionale vectorruimten geldt dit i.h.a. niet. Een vectorruimte waarop een norm gedefinieerd is, heet een *genormeerde* vectorruimte. Voor een lineaire afbeelding $T : V \to W$ tussen genormeerde vectorruimten V en W kunnen we de norm van T analoog aan de Euclidische norm definiëren als $\sup_{\|x\|=1} \|T(x)\|$. In het geval dat de dimensies oneindig zijn, is het mogelijk dat $\|T\| = \infty$.

Banach- en Hilbertruimten. Convergente rijen van lineaire afbeeldingen.

Zij V een genormeerde vectorruimte met norm $\| \|$. Een rij $\{x_j\}_{j=1}^{\infty}$ in V heet een *Cauchyrij* of *fundamentaalrij* als er voor elke $\epsilon > 0$ een N bestaat zodanig dat $\|x_j - x_k\| < \epsilon$ voor $j, k \geq N$. We zeggen dat een rij $\{x_j\}_{j=1}^{\infty}$ in V convergeert indien er een $x \in V$ bestaat zodanig dat $\lim_{j\to\infty} \|x - x_j\| = 0$. Een vectorruimte V heet *volledig* of *compleet* indien elke Cauchyrij in V convergeert. Een volledige genormeerde vectorruimte noemen we ook wel een *Banachruimte*. Alle eindig-dimensionale genormeerde vectorruimten zijn Banachruimten.

Een vectorruimte V met een inwendig product $(\, , \,)$ is i.h.b. een genormeerde vectorruimte. Indien V volledig is t.a.v. de door het inproduct geïnduceerde norm, dan heet V een *Hilbertruimte*. Voorbeelden van Hilbertruimten zijn:

1. $V = \ell_2(K)$, de vectorruimte van oneindige rijtjes $(x_1, x_2, \dots)$ (met $x_i \in K$) met de componentsgewijze optelling en scalaire vermenigvuldiging, zodanig dat $\sum_{j=1}^{\infty} |x_j|^2$ convergeert. Het inproduct is gedefinieerd als $(x, y) = \sum_{i=1}^{\infty} \overline{x_i} y_i$. Merk op dat uit de ongelijkheid van Schwarz (voor het eindig-dimensionale geval) volgt dat de reeks convergeert.

2. De vectorruimte $\mathcal{L}_2([a, b], K)$ van kwadratisch integreerbare reële resp. complexe functies op het interval $[a, b]$ met inproduct $(f, g) = \displaystyle\int_a^b \overline{f(x)} g(x) dx$.

Andere voorbeelden van genormeerde vectorruimten:

3. $V = \ell_p(K)$ (waarbij $p \geq 1$)), dit is zoals in (1) de vectorruimte van oneindige rijtjes $(x_1, x_2, \dots)$ maar nu zodanig dat $\sum_{j=1}^{\infty} |x_j|^p$ convergeert. De norm van het rijtje $\mathbf{x} = (x_1, x_2, \dots)$ is gedefinieerd als $\|\mathbf{x}\|_p = (\sum_{j=1}^{\infty} |x_j|^p)^{1/p}$. Als $p \geq 1$ is dit een norm op V, d.w.z. de driehoeksongelijkheid $\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$ geldt. Deze ongelijkheid voor p-normen wordt de *ongelijkheid van Minkowski* genoemd.

4. Analoog is de vectorruimte $\mathcal{L}_p([a, b], K)$ van K-waardige functies op $[a, b]$ zodanig dat $\int_a^b |f(x)|^p dx$ bestaat voor $p \geq 1$ een genormeerde vectorruimte met norm $\|f\| = (\int_a^b |f(x)|^p dx)^{1/p}$

5. De vectorruimte $C([a, b], K)$ van K-waardige continue functies op $[a, b]$ is een genormeerde vectorruimte met de *sup*-norm $\|f\| = \sup_{[a,b]} |f(x)|$. Convergentie t.a.v. de sup-norm wordt ook wel *uniforme convergentie* genoemd. Omdat een uniform convergente Cauchyrij van continue functies convergeert naar een continue functie, is de vectorruimte $C([a, b], K)$ volledig.

Laat nu V, W genormeerde vectorruimten zijn, en W bovendien volledig. Zij $\{T_j : V \to W\}_{j=1}^{\infty}$ een Cauchy-rij van begrensde lineaire afbeeldingen, m.a.w. voor elke $\epsilon > 0$ is $\|T_j - T_k\| < \epsilon$ zodra $j, k \geq N$. Daar voor $x \in V$ geldt dat $\|T_j(x) - T_k(x)\| \leq \|T_j - T_k\|\|x\|$, is de rij $\{T_j(x)\}_{j=1}^{\infty}$ voor elke $x \in V$ een Cauchyrij, en daar W volledig is, is de rij convergent. De limiet noemen we $T(x)$. Het is niet moeilijk om aan te tonen dat $T : V \to V$ een lineaire afbeelding is. T heet de limiet van de rij $\{T_j\}_{j=1}^{\infty}$. Op dezelfde manier kan ook de som van een convergente reeks $\sum_{j=1}^{\infty} T_j$ worden gedefinieerd als de limiet van de rij van partiële sommen $\{\sum_{j=1}^{n} T_j\}_{n=1}^{\infty}$.

Voorbeelden: 1. Zij V een Banachruimte en $T : V \to V$ lineair met $\|T\| < 1$, dan is de afbeelding $id_V - T$ injectief: immers uit $T(v) = v$ en $v \neq 0$ volgt dat $\|v\| = \|T(v)\| \leq \|T\|\|v\| < \|v\|$. De reeks $id_V + T + T^2 + \ldots$ is een Cauchyreeks (aangezien $\|T^m + T^{m+1} \ldots + T^n\| \leq \dfrac{\|T^m\|}{1 - \|T\|}$ en het rechterlid willekeurig klein wordt als $m \to \infty$) en convergeert dus. De limiet is gelijk aan $(id_V - T)^{-1}$. 2. Zij V een eindig-dimensionale vectorruimte en $T : V \to V$ willekeurig. De reeks $\sum_{j=0}^{\infty} T^j/j!$ convergeert (op V kunnen we altijd een norm definiëren). De limiet noemen we e^T of $\exp(T)$.

De e-macht van een matrix. Op dezelfde manier als in het laatste voorbeeld is voor een willekeurige $n \times n$-matrix A de e-macht e^A gedefinieerd als $I_n + A + A^2/2! + \ldots$. Verder zijn $\sin(A)$ en $\cos(A)$ gedefinieerd d.m.v. $\sin(A) = \dfrac{e^{iA} - e^{-iA}}{2i}$ en $\cos(A) = \dfrac{e^{iA} + e^{-iA}}{2}$. Voor $\|A\| < 1$ is $\log(I_n + A) = A - A^2/2 + A^3/3 + \ldots$ gedefinieerd en er geldt: $e^{\log(A)} = A$.

Merk op dat i.h.a. niet geldt: $e^{A+B} = e^A \cdot e^B$; als A en B commuteren (dus $AB = BA$), dan geldt $e^{A+B} = e^A \cdot e^B$ echter wel. Het bewijs verloopt op dezelfde manier als voor reële of complexe getallen. Als A een willekeurige $n \times n$-matrix is, dan is A te schrijven als een som $B + N$ van een diagonalizeerbare matrix B en een nilpotente matrix N en $BN = NB$ (verg. hoofdstuk III). In dit geval geldt $e^A = e^B \cdot e^N$. Daar $N^n = O$, is $e^N = I + N + \ldots + \frac{1}{(n-1)!} N^{n-1}$ een polynoom van hoogstens graad $n - 1$; e^B is als volgt te berekenen: omdat B diagonaliseerbaar is, is er een inverteerbare matrix U zodat $B = UDU^{-1}$ waarbij D een diagonaalmatrix is (U is een matrix van eigenvectoren van B). Dan geldt: $e^B = Ue^D U^{-1}$ (vergelijk opgave VI.7). Als $D = \mathrm{diag}(a_1, \ldots, a_n)$, dan is $e^D = \mathrm{diag}(e^{a_1}, \ldots, e^{a_n})$ (ga dit na).
Voorbeeld 1: Zij $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. A heeft eigenwaarden 1 en 3 met eigenvectoren $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ resp. $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. A is dus diagonaliseerbaar (dus $B = A$ en $N = O$). Verder is $A = UDU^{-1}$ met $U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ en $D = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$, dus (merk op dat U orthogonaal is)

$$A = Ue^D U^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^3 & 0 \\ 0 & e \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e^3 + e & e^3 - e \\ e^3 - e & e^3 + e \end{pmatrix}.$$

Voorbeeld 2: Zij $A = \begin{pmatrix} 0 & -1 \\ 4 & -4 \end{pmatrix}$. A heeft karakteristiek polynoom $(X + 2)^2$, dus $B = -2I =$

$\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$ en $N = \begin{pmatrix} 2 & -1 \\ 4 & -2 \end{pmatrix}$. Nu is $e^N = I + N$ dus

$$e^A = e^D \cdot e^N = \begin{pmatrix} e^{-2} & 0 \\ 0 & e^{-2} \end{pmatrix} \begin{pmatrix} 3 & -1 \\ 4 & -1 \end{pmatrix} = e^{-2} \begin{pmatrix} 3 & -1 \\ 4 & -1 \end{pmatrix}.$$

Vector- en matrixwaardige functies.

We beschouwen afbeeldingen $A : I \to \mathcal{M}(m \times n, K)$ van een interval $I = [a, b] \subset \mathbf{R}$ naar de vectorruimte van reële of complexe $m \times n$-matrices. De componenten $A_{ij}(t)$ van de matrixfunctie $A(t)$ zijn dus (reële of complexe) functies van $[a, b]$. $A(t)$ is continu, resp. differentieerbaar als de componenten $A_{ij}(t)$ dat zijn. Continuïteit en differentieerbaarheid kunnen ook in termen van de norm worden gedefinieerd. Zoals we gezien hebben wordt de norm van een $m \times n$-matrix gedefinieerd als $\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$. Dan geldt: de op het interval I gedefinieerde matrixwaardige functie $A(t)$ is continu in $c \in I$ dan en slechts dan als voor elke $\epsilon > 0$ er een $\delta > 0$ bestaat zodanig dat indien $|t - c| < \delta$ en $t \in I$, dan $\|A(t) - A(c)\| < \epsilon$, m.a.w. $\lim_{t \to c} \|A(t) - A(c)\| = 0$ (waarbij $t \in I$ wordt genomen). M.b.v. Gevolg 6.3 is onmiddellijk in te zien dat beide definities van continuïteit gelijkwaardig zijn.

$A(t)$ is differentieerbaar in een inwendig punt c van I dan en slechts dan als $A(t) - A(c) = (t - c)B + (t - c)R(t)$ voor $t \in I$ waarbij $\lim_{t \to c} \|R(t)\| = 0$. De matrix B is de afgeleide $A'(c)$ van $A(t)$ in $t = c$ en heeft als elementen precies $A'_{ij}(c)$. Merk op dat $B = \lim\limits_{t \to c} \dfrac{A(t) - A(c)}{t - c}$. $A(t)$ heet continu, resp. differentieerbaar op het interval I als $A(t)$ continu, resp. differentieerbaar is in elk punt van I. Nu gelden de gebruikelijke regels voor het differentiëren: $(A + B)'(t) = A'(t) + B'(t)$ en $(AB)'(t) = A'(t)B(t) + A(t)B'(t)$ als de som en het product goed gedefinieerd zijn en $A(t), B(t)$ differentieerbaar zijn.

Voorbeelden: 1. Laat $A(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$. Dan is $A(t)$ differentieerbaar op $I = \mathbf{R}$ en $A'(t) = \begin{pmatrix} -\sin t & -\cos t \\ \cos t & -\sin t \end{pmatrix}$.

2. Laat $A(t) = e^{tB}$ voor B een $n \times n$-matrix. Dan is $A(t)$ differentieerbaar op $\mathbf{R}$ en $A'(t) = Be^{tB}$. We laten eerst zien dat $A(t)$ differentieerbaar is in $t = 0$: Voor $t \in \mathbf{R}$ is $e^{tB} = I + tB + tR(t)$ met $R(t) = \frac{1}{2}B^2 t + \frac{1}{6}B^3 t^2 + \ldots$ dus

$$\|R(t)\| \leq \frac{1}{2}\|B\|^2 |t| + \frac{1}{6}\|B\|^3 |t|^2 + \ldots$$

en in het rechterlid staat een convergente machtreeks die naar 0 convergeert als $t \to 0$ dus $\lim_{t \to 0} \|R(t)\| = 0$ en $A'(0) = B$. Voor differentieerbaarheid in $t = c$ schrijven we $A(t) = e^{cB} e^{uB}$ met $u = t - c$. Dan is $A(t) = e^{cB}(I + uB + uR(u))$ met $\|R(u)\| \to 0$ als $u \to 0$, dus is $A(t)$ differentieerbaar in c en $A'(c) = e^{cB}B = Be^{cB}$.

Propositie 6.5: Zij $A(t)$ voor $t \in I$ een $n \times n$-matrix. Dan geldt:

1. Als $A(t)$ continu is in $t = c \in I$ en $A(c)$ is inverteerbaar, dan is $A(t)$ inverteerbaar in een omgeving $(c - \epsilon, c + \epsilon)$ van c en $A^{-1}(t)$ is continu in $t = c$.

2. Als $A(t)$ inverteerbaar en differentieerbaar in $t = c$, dan is $A^{-1}(t)$ differentieerbaar in c en

$$\frac{d}{dt} A^{-1}(t)\Big|_{t=c} = -A^{-1}(c)A'(c)A^{-1}(c).$$

Voor het bewijs gebruiken we het volgende lemma:

Lemma 6.6: Als A een inverteerbare $n \times n$-matrix is en B is een matrix zodanig dat $\|A - B\| < \|A^{-1}\|^{-1}$, dan is B inverteerbaar en

$$\|B^{-1}\| \leq \frac{1}{\|A^{-1}\|^{-1} - \|A - B\|}. \tag{6.1}$$

Bewijs: Neem aan dat $B\mathbf{x} = \mathbf{0}$ voor $\mathbf{x} \in \mathbf{C}^n$ en $\mathbf{x} \neq \mathbf{0}$. Dan volgt uit $\|\mathbf{x}\| \leq \|A^{-1}\|\|A\mathbf{x}\|$ dat

$$\|A^{-1}\|^{-1}\|\mathbf{x}\| \leq \|A\mathbf{x}\| = \|(A - B)\mathbf{x}\| \leq \|A - B\|\|\mathbf{x}\| < \|A^{-1}\|^{-1}\|\mathbf{x}\|.$$

Dit is een tegenspraak, m.a.w. uit $B\mathbf{x} = \mathbf{0}$ volgt dat $\mathbf{x} = \mathbf{0}$ dus B is inverteerbaar. Verder is, voor $\mathbf{x} \in \mathbf{C}^n$ en $B\mathbf{y} = \mathbf{x}$,

$$\frac{\|B^{-1}\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|\mathbf{y}\|}{\|B\mathbf{y}\|} = \frac{\|\mathbf{y}\|}{\|(B - A)\mathbf{y} + A\mathbf{y}\|} \leq$$

$$\leq \frac{\|\mathbf{y}\|}{|\|A\mathbf{y}\| - \|(A - B)\mathbf{y}\||} \leq \frac{\|\mathbf{y}\|}{\|A^{-1}\|^{-1}\|\mathbf{y}\| - \|A - B\|\|\mathbf{y}\|} = \frac{1}{\|A^{-1}\|^{-1} - \|A - B\|}.$$

Door het maximum te nemen over alle $\mathbf{x} \neq \mathbf{0}$ wordt het linkerlid in de bovenstaande rij ongelijkheden gelijk aan $\|B^{-1}\|$ en (6.1) volgt onmiddellijk. $\diamond$

Bewijs van Propositie 6.5: (1.) Uit het lemma volgt onmiddellijk dat

$$\|A^{-1} - B^{-1}\| = \|A^{-1}(B - A)B^{-1}\| \leq \|A^{-1}\|\|A - B\|\|B^{-1}\| \leq \frac{\|A^{-1}\|\|A - B\|}{\|A^{-1}\|^{-1} - \|A - B\|}.$$

Laat nu in bovenstaande uitdrukking $A = A(c)$ en $B = A(t)$ voor $|t - c|$ klein. Dan volgt de bewering uit $\|A(t) - A(c)\| \to 0$ $(t \to c)$. (Inverteerbaarheid van $A(t)$ in een omgeving van c volgt ook direct uit $\det A(c) \neq 0$ en uit de continuïteit van de functie $t \to \det A(t)$.
(2.) Differentieerbaarheid volgt uit continuïteit van de inverse m.b.v.

$$A^{-1}(t) - A^{-1}(c) = A^{-1}(t)\big(A(c) - A(t)\big)A^{-1}(c).$$

Als we links en rechts delen door $t - c$ en de limiet voor $t \to c$ nemen zien we dat $A^{-1}(t)$ differentieerbaar is in $t = c$ en $(A^{-1})'(c) = -A^{-1}(c)A'(c)A^{-1}(c)$. $\diamond$

Toepassing: stelsels van lineaire differentiaalvergelijkingen. De e-macht speelt een rol bij de oplossing van stelsels lineaire differentiaalvergelijkingen: evenals een stelsel algebraïsche lineaire vergelijkingen kunnen we een stelsel

$$\begin{cases} x_1'(t) &= & a_{11}x_1(t) + a_{12}x_2(t) + \ldots + a_{1n}x_n(t) \\ \quad\vdots & & \vdots \\ x_n'(t) &= & a_{n1}x_1(t) + a_{n2}x_2(t) + \ldots + a_{nn}x_n(t) \end{cases} \tag{6.2}$$

schrijven als

$$\mathbf{x}'(t) = A\mathbf{x}(t) \tag{6.3}$$

waarbij $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))^T$ en A de matrix $(a_{ij})_{i,j=1\ldots n}$ is. Dan geldt: de oplossing $\mathbf{x}(t)$ is voor alle $t \in \mathbf{R}$ gedefinieerd en gelijk aan: $\mathbf{x}(t) = e^{tA}\mathbf{x}(0)$.

Bewijs: Laat $\mathbf{y}(t) = e^{-tA}\mathbf{x}(t)$. Dan is $\mathbf{y}'(t) = -e^{-tA}A\mathbf{x}(t) + e^{tA}\mathbf{x}'(t) = \mathbf{0}$. Maar dan is $\mathbf{y}(t)$ constant, dus $\mathbf{y}(t) = \mathbf{y}(0) = \mathbf{x}(0)$.

Toepassing: Los het stelsel $\begin{cases} x'(t) = & -y(t) \\ y'(t) = & x(t) + 2y(t) \end{cases}$ op. Dit stelsel is te schrijven als $\mathbf{x}'(t) = A\mathbf{x}(t)$ met $A = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$ en $\mathbf{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$. We berekenen e^{tA}: hiertoe schrijven we A als de som $A = D + N$ van een diagonalizeerbare matrix D en een nilpotente matrix N zodanig dat D en N commuteren (vergelijk hfd.III). A heeft eigenwaarde 1 met algebraïsche multipliciteit 2, dus $D = I$ (de eenheidsmatrix) en $N = \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}$. Daar $N^2 = O$, is $e^{tN} = I + tN$ en $e^{tA} = e^{tD} \cdot e^{tN} = e^t(I + tN)$. De oplossing van het stelsel is dus $z(t) = e^t((1-t)x(0) - ty(0))$ en $y(t) = e^t(tx(0) + (1+t)y(0))$.

Een lineaire differentiaalvergelijking van orde n

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \ldots + a_0 y(t) = 0 \tag{6.4}$$

kunnen we schrijven als een stelsel van n lineaire eerste-orde differentiaalvergelijkingen: laat $x_1(t) = y^{(n-1)}(t), \ldots, x_{n-1}(t) = y'(t)$, $x_n(t) = y(t)$. Dan is de d.v. (6.4) gelijkwaardig met het stelsel

$$\begin{cases} x_1'(t) & = & -a_{n-1}x_1(t) + \ldots - a_1 x_{n-1}(t) - a_0 x_n(t) \\ x_2'(t) & = & -x_1(t) \\ \vdots & & \vdots \\ x_n'(t) & = & -x_{n-1}(t) \end{cases}$$

Voorbeeld: Beschouw de d.v. $y'' + 2y' + y = 0$. Laat $y' = z$, $A = \begin{pmatrix} -2 & -1 \\ 1 & 0 \end{pmatrix}$ en $\mathbf{x}(t) = \begin{pmatrix} z(t) \\ y(t) \end{pmatrix}$. Dan is $\mathbf{x}'(t) = A\mathbf{x}(t)$. Ga na dat A te ontbinden is als $A = -I + N$ met $N^2 = O$ en dus is $\mathbf{x}(t) = e^{-t}(I + Nt)\mathbf{x}(0)$ en $\begin{cases} z(t) = e^{-t}(1-t)z_0 - e^{-t}ty_0 \\ y(t) = e^{-t}tz_0 + e^{-t}(1+t)y_0 \end{cases}$ waarbij $y(0) = y_0$ en $y'(0) = z_0$. De oplossing $y(t)$ is dus een lineaire combinatie van de oplossingen $y_1(t) = e^{-t}$ en $y_2(t) = te^{-t}$. De oplossing vormt dus een lineaire deelruimte van de vectorruimte $C_\infty(\mathbf{R})$ van oneindig vaak differentieerbare functies op $\mathbf{R}$. In het algemeen vormen de oplossingen van de d.v. (6.4) een n-dimensionale lineaire deelruimte van $C_\infty(\mathbf{R})$.

Als de coëfficiënten van (6.4) niet constant zijn, maar wel continue functies zijn op een interval $[a, b] \subset \mathbf{R}$ (met $a < b$), dan heeft de verzameling oplossingen van (6.4) eveneens de structuur van een vectorruimte van dimensie n over $\mathbf{C}$ (onder de gewone optelling en scalaire vermenigvuldiging van functies).

VII. SPECTRAALTHEORIE VAN NORMALE AFBEELDINGEN

Normale afbeeldingen.

In het voorgaande zijn we een aantal voorbeelden tegengekomen van lineaire afbeeldingen die een orthonormale basis van eigenvectoren hebben, zoals symmetrische en unitaire afbeeldingen. We zullen zien dat de klasse van lineaire afbeeldingen $V \to V$ voor een eindig-dimensionale (reële resp. complexe) vectorruimte V die *orthogonaal*, resp. *unitair diagonaliseerbaar* (d.w.z. een orthonormale basis van eigenvectoren hebben) zijn, eenvoudig gekarakteriseerd kan worden. In het vervolg nemen we steeds aan dat V een eindig-dimensionale reële of complexe vectorruimte is met (hermites) inproduct en dimensie n.

Zij $T : V \to V$ lineair. Kies een orthonormale basis $\mathcal{B}$ van V en laat A de matrix van T zijn t.o.v. deze basis. Zoals we hebben gezien heeft T een basis van eigenvectoren precies dan als er een inverteerbare matrix U bestaat zodanig dat $U^{-1}AU$ een diagonaalmatrix is (het is eenvoudig in te zien dat U dan een matrix van eigenvectoren is). T is dus orthogonaal, resp. unitair diagonalizeerbaar als $U^{-1}AU$ een diagonaalmatrix is, waarbij U een orthogonale resp. unitaire matrix is. Als V een reële vectorruimte is, en T is orthogonaal diagonaliseerbaar, dan is de diagonaalmatrix $D := U^{-1}AU = U^TAU$ symmetrisch en dus is A ook symmetrisch: $A^T = (UDU^T)^T = UDU^T = A$ en de afbeelding T is dan een symmetrische afbeelding (zie hfd.IV). Omgekeerd zullen we zien dat een symmetrische afbeelding T orthogonaal diagonaliseerbaar is. Dit karakteriseert de orthogonaal diagonalizeerbare afbeeldingen in reële eindig-dimensionale vectorruimten volledig.

Het complexe geval verloopt iets minder direct. Als T unitair diagonaliseerbaar is, dan is de diagonaalmatrix $D = U^{-1}AU = U^*AU$ alleen dan hermites indien T louter reële eigenwaarden heeft: in dit laatste geval zien we, analoog aan het symmetrische geval, dat dan A en dus T hermites is. Omgekeerd geldt dat elke hermitese afbeelding unitair diagonaliseerbaar is en louter reële eigenwaarden heeft. Als T ook niet-reële eigenwaarden heeft, is $D \neq D^*$ maar wel geldt dat $DD^* = D^*D$ en dus ook $AA^* = (UDU^*)(UDU^*)^* = UDD^*U^* = A^*A$. Aangezien A de matrix t.o.v. een orthonormale basis is, is dan ook $TT^* = T^*T$.

Definitie: Een lineaire afbeelding $T : V \to V$ (resp. een $n \times n$-matrix A) heet *normaal* als $T^*T = TT^*$ (resp. $AA^* = A^*A$).

Voorbeelden: 1. Hermitese afbeeldingen en matrices zijn normaal.
2. Unitaire afbeeldingen en matrices zijn normaal.
3. Antihermitese afbeeldingen en matrices zijn normaal. (Een afbeelding heet antisymmetrisch resp. antihermites als $T^* = -T$). Merk op dat in complexe vectorruimten geldt: T is antihermites dan en slechts dan als iT hermites is. In reële vectorruimten geldt niet een dergelijke 1-1 relatie tussen antisymmetrische en symmetrische afbeeldingen.

Voor een complexe e.d. vectorruimte V geldt het volgende verband tussen unitaire diagonalizeerbaarheid en normaliteit:

Propositie 7.1 (*spectraalstelling voor normale afbeeldingen:*) Een lineaire afbeelding $T : V \to V$ is unitair diagonaliseerbaar dan en slechts dan als T normaal is.

Bewijs: We hoeven alleen nog te bewijzen dat een normale afbeelding unitair diagonaliseerbaar is. Laat dus $T : V \to V$ een normale afbeelding zijn. Het bewijs bestaat uit drie stappen:
1. Voor $x \in V$ is $\|Tx\| = \|T^*x\|$. Immers

$$\|Tx\|^2 = (Tx, Tx) = (T^*Tx, x) = (TT^*x, x) = (T^*x, T^*x) = \|T^*x\|^2.$$

2. Laat λ een eigenwaarde van T zijn en x een eigenvector. Dan is x een eigenvector van T^* met eigenwaarde $\overline{\lambda}$. Immers volgens (1) is

$$0 = \|(T - \lambda \cdot id)x\| = \|(T^* - \overline{\lambda} \cdot id)x\|.$$

3. We passen inductie naar de dimensie toe. Voor dimensie 1 valt er niets te bewijzen. Stel nu dat de bewering waar is voor vectorruimten van dimensie kleiner dan $n = \dim(V)$. Laat x een eigenvector van T zijn. We tonen aan dat zowel T als T^* het orthogonaal complement X van span$\{x\}$ in zichzelf afbeelden. Laat dus $y \in X$. Dan is $(Ty, x) = (y, T^*x) = \bar{\lambda}(y, x) = 0$ en analoog is $(T^*y, x) = 0$. Bekijk nu de restrictie $T|_X$ van T tot X. Aangezien de geadjungeerde van de restrictie de restrictie van de geadjungeerde is, m.a.w. $(T|_X)^* = T^*|_X$ (waarom?) is $T|_X : X \to X$ weer een normale afbeelding en volgens de inductieveronderstelling is er een orthonormale basis van eigenvectoren van $T|_X$ (en dus van T) in X. Samen met $x/\|x\|$ vormen deze vectoren een orthonormale basis van eigenvectoren van T. $\diamond$

Gevolg 7.2 (*spectraaldecompositie van een normale matrix*):
Zij A een normale matrix. Dan is $A = VDV^*$ met V unitair en D een diagonaalmatrix. Dus we kunnen A schrijven als

$$A = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^* + \lambda_2 \mathbf{v}_2 \mathbf{v}_2^* + \ldots + \lambda_n \mathbf{v}_n \mathbf{v}_n^*$$

waarbij $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ een orthonormale basis van eigenvectoren van A en $\lambda_1, \lambda_2, \ldots, \lambda_n$ de bijbehorende eigenwaarden zijn. Immers

$$A = VDV^* = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \begin{pmatrix} \mathbf{v}_1^* \\ \mathbf{v}_2^* \\ \vdots \\ \mathbf{v}_n^* \end{pmatrix} = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^* + \ldots + \lambda_n \mathbf{v}_n \mathbf{v}_n^*.$$

Voor een expliciet voorbeeld zie onder Propositie 7.4, waar het analoge geval van een (reële) symmetrische matrix wordt behandeld.

Gevolg 7.3. Een normale afbeelding $T : V \to V$ is hermites dan en slechts dan als alle eigenwaarden van T reëel zijn.

Bewijs: Laat $\{v_1, \ldots, v_n\}$ een orthonormale basis van eigenvectoren van T zijn. Dan is $T(v_i) = \lambda_i v_i$ en $T^*(v_i) = \overline{\lambda_i} v_i$ voor $i = 1, \ldots, n$. Dus $T = T^*$ precies dan als $\lambda_i = \overline{\lambda_i}$ voor alle i.

Merk op dat voor een orthogonale projectie P_W op de lineaire deelruimte $W = \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ de eigenwaarden $\lambda_1 = \ldots = \lambda_k = 1$ en $\lambda_{k+1} = \ldots = \lambda_n = 0$ zijn en de matrix t.o.v. een orthonormale basis is dan $\mathbf{v}_1\mathbf{v}_1^* + \ldots + \mathbf{v}_k\mathbf{v}_k^*$, een resultaat dat we al eerder hebben afgeleid. Een willekeurige normale afbeelding T is dus een som $\lambda_1 P_1 + \ldots + \lambda_\ell P_\ell$ waarbij nu $\lambda_1, \ldots, \lambda_k$ de verschillende eigenwaarden zijn en P_i de orthogonale projectie op de eigenruimte bij λ_i. $\diamond$

Opmerking: Zij $T \in \mathcal{L}(V)$. Dan zijn er hermitese afbeeldingen T_1, T_2 zodanig dat $T = T_1 + iT_2$. Laat immers $T_1 = (T + T^\dagger)/2$ en $T_2 = (T - T^\dagger)/2i$. Nu is T normaal dan en slechts dan als T_1, T_2 commuteren. M.b.v. Propositie 3.10 (over gemeenschappelijke eigenwaarden van commuterende lineaire afbeeldingen) volgt nu de spectraalstelling voor normale afbeeldingen eenvoudig uit de spectraalstelling voor hermitese afbeeldingen.

Symmetrische matrices. We bestuderen nu het geval van een (eindig-dimensionale) reële vectorruimte V. We hebben reeds gezien dat de orthogonaal diagonalizeerbare afbeeldingen in elk geval symmetrisch moeten zijn. Ook geldt hier het omgekeerde: symmetrische afbeeldingen zijn orthogonaal diagonaliseerbaar. We bekijken hiertoe de matrix A van een symmetrische afbeelding T t.o.v. een orthonormale basis. Deze is symmetrisch en dus i.h.b. hermites. De eigenwaarden zijn dus alle reëel. Nu volgt echter niet onmiddellijk uit het complexe geval dat er een orthonormale basis van eigenvectoren is omdat de eigenvectoren $\mathbf{v}_1, \ldots, \mathbf{v}_n$ die we vonden in het complexe geval complexe coördinaten kunnen hebben. We kunnen A echter schrijven op unieke wijze als een som van orthogonale projecties: $A = \lambda_1 P_1 + \ldots + \lambda_\ell P_\ell$ met verschillende reële eigenwaarden $\lambda_1, \ldots, \lambda_\ell$ en met P_i een som van termen van de vorm $\mathbf{v}_j\mathbf{v}_j^*$. Daar A reëel is, is $A = \overline{A}$, dus

$$\lambda_1\mathbf{v}_1\mathbf{v}_1^* + \ldots + \lambda_n\mathbf{v}_n\mathbf{v}_n^* = A = \overline{A} = \lambda_1\overline{\mathbf{v}_1}\overline{\mathbf{v}_1}^* + \ldots + \lambda_n\overline{\mathbf{v}_n}\overline{\mathbf{v}_n}^*$$

en daar ook $\{\overline{\mathbf{v}_1}, \ldots, \overline{\mathbf{v}_n}\}$ een orthonormale basis van V is, is $P_j = \overline{P_j}$ voor $j = 1, \ldots, \ell$. Dus de projectiematrices zijn reëel. Maar zoals we weten heeft een reële orthogonale projectie een orthogonale basis van (reële) eigenvectoren. De eigenvectoren van P_j zijn de eigenvectoren van A met eigenwaarde λ_j. De symmetrische matrix A heeft dus een orthonormale basis van eigenvectoren in de reële vectorruimte V. We hebben dus aangetoond dat een symmetrische afbeelding $T : V \to V$ in een reële vectorruimte orthogonaal diagonaliseerbaar is.

Propositie 7.4 (*spectraaldecompositie van een symmetrische matrix*):
Zij A een reële symmetrische $n \times n$-matrix. Dan zijn er een orthonormale basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ van $\mathbf{R}^n$ en reële getallen $\lambda_1, \ldots, \lambda_n$ zodanig dat $A = \lambda_1\mathbf{v}_1\mathbf{v}_1^T + \lambda_2\mathbf{v}_2\mathbf{v}_2^T \ldots + \lambda_n\mathbf{v}_n\mathbf{v}_n^T$.

Voorbeeld: Laat $A = \begin{pmatrix} 2 & 4 & -1 \\ 4 & 2 & 1 \\ -1 & 1 & -1 \end{pmatrix}$. A is symmetrisch en heeft eigenwaarden $\lambda_1 = 6, \lambda_2 = -3$ en $\lambda_3 = 0$ met (genormaliseerde) eigenvectoren

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{v}_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \text{ resp. } \mathbf{v}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}.$$

De spectraaldecompositie is dus

$$A = 6u_1u_1^T - 3u_2u_2^T = 6 \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix} - 3 \begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}.$$

In termen van matrices wordt dit

$$A = UDU^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{3} & -1/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{3} & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & 2/\sqrt{6} \end{pmatrix} \begin{pmatrix} 6 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{3} & -1/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{3} & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & 2/\sqrt{6} \end{pmatrix}^T.$$

Het begrip spectraaldecompositie geldt voor beide schrijfwijzen.

Kwadratische vormen op $\mathbf{R}^n$.

De uitdrukking $q(\mathbf{x}) = \sum_{i,j=1}^{n} a_{ij}x_ix_j$ waarbij $\mathbf{x} \in \mathbf{R}^n$ en $a_{ij} \in \mathbf{R}$ heet een *kwadratische vorm*. Omdat we zonder beperking van de algemeenheid mogen aannemen dat $a_{ij} = a_{ji}$ kan $q(\mathbf{x})$ worden uitgedrukt als $\mathbf{x}^T A\mathbf{x}$ waarbij $A = (a_{ij})$ een symmetrische matrix is. De kwadratische vorm kunnen we d.m.v. een orthogonale coördinatentransformatie in diagonaalvorm schrijven: laat U een orthogonale matrix zijn zodat $U^TAU = D$ een diagonaalmatrix is en $\mathbf{y} = U^T\mathbf{x}$, dan is

$$q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \mathbf{x}^T UDU^T\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2.$$

Hierbij zijn $\lambda_1, \ldots, \lambda_n$ de eigenwaarden van A. In principe zijn er ook andere mogelijkheden om de kwadratische vorm in diagonaalvorm te schrijven: $q(\mathbf{x}) = \mathbf{z}^T D'\mathbf{z} = \mu_1 z_1^2 + \ldots + \mu_n z_n^2$ voor zekere $\mathbf{z} = V^T\mathbf{x}$ zodat $D' = V^TAV$ een diagonaalmatrix is. Merk op dat V geen orthogonale matrix hoeft te zijn. Hierbij kunnen de coëfficiënten μ_i i.h.a. verschillen van de coëfficiënten λ_j. Het aantal positieve resp. negatieve) coëfficiënten in de gediagonaliseerde kwadratische vorm is echter altijd constant. De reden is dat het aantal positieve (resp. negatieve) coëfficiënten in de vorm $q(\mathbf{x}) = \mu_1 z_1^2 + \ldots + \mu_n z_n^2$ gelijk is aan de maximale dimensie van een lineaire deelruimte W van $\mathbf{R}^n$ zodanig dat $q(\mathbf{x}) > 0$ (resp. $q(\mathbf{x}) < 0$) is voor alle $\mathbf{x} \in W$ met $\mathbf{x} \neq \mathbf{0}$. Dit feit staat bekend als de *traagheidsstelling van Sylvester*. Als p het aantal positieve eigenwaarden van A is en q het aantal negatieve eigenwaarden. dan heet $p + q$ de *rang* van de kwadratische vorm $\mathbf{x}^T A\mathbf{x}$ en $p - q$ heet de *signatuur*.

Toepassing: Zij $f : \mathbf{R}^n \to \mathbf{R}$ een minstens tweemaal continu differentieerbare functie in het punt $\mathbf{0} \in \mathbf{R}^n$. Dan is voor $\|\mathbf{x}\|$ voldoende klein $f(\mathbf{x}) = f(\mathbf{0}) + \nabla f(\mathbf{0})^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T H\mathbf{x} + o(\|\mathbf{x}\|^2)$ waarbij $\nabla f(\mathbf{0}) \in \mathbf{R}^n$ de gradiënt van f in $\mathbf{x} = \mathbf{0}$ voorstelt en de Hessematrix H de symmetrische matrix $(\frac{\partial^2 f}{\partial x_i \partial x_j})_{ij}$ voorstelt. $\mathbf{0}$ is een stationair punt van f als $\nabla f(\mathbf{0}) = \mathbf{0}$. In dit geval wordt het gedrag van f in de omgeving van $\mathbf{0}$ in hoogste orde door H bepaald. Nu geldt voor $\|\mathbf{x}\|$ klein en $\mathbf{y} = U^T\mathbf{x} \in \mathbf{R}^n$ (met U als boven):

$$f(\mathbf{x}) = f(\mathbf{0}) + \frac{1}{2}\sum_{i=1}^{n} \lambda_i y_i^2 + o(\|x\|^2)$$

waarbij $\lambda_1, \ldots, \lambda_n$ de eigenwaarden van H zijn.
Gevolg:

1. Als alle eigenwaarden van H positief (resp. negatief) zijn, dan neemt f in $\mathbf{0}$ een minimum (resp. maximum) aan. (H is dan een *positief*, resp. *negatief definiete* matrix, zie hfd.VIII)
2. Als H zowel positieve als negatieve eigenwaarden heeft (H heet dan *indefiniet*), dan neemt f geen maximum en ook geen minimum aan in $\mathbf{x} = \mathbf{0}$. Het punt $\mathbf{x} = \mathbf{0}$ heet dan een *zadelpunt* van f.

Het Rayleighquotiënt en het minimax-principe.

Zij A een normale matrix. Het quotiënt $R_A(\mathbf{x}) = \dfrac{(\mathbf{x}, A\mathbf{x})}{(\mathbf{x}, \mathbf{x})}$ dat voor $\mathbf{x} \neq \mathbf{0}$ gedefinieerd is, heet het *Rayleighquotiënt* van A. M.b.v. de spectraaldecompositie kunnen we dit schrijven als

$$R_A(\mathbf{x}) = \frac{\lambda_1 \mathbf{x}^*\mathbf{v}_1\mathbf{v}_1^*\mathbf{x} + \lambda_2 \mathbf{x}^*\mathbf{v}_2\mathbf{v}_2^*\mathbf{x} + \ldots + \lambda_n \mathbf{x}^*\mathbf{v}_n\mathbf{v}_n^*\mathbf{x}}{\mathbf{x}^*\mathbf{x}} = \frac{\lambda_1|(\mathbf{x}, \mathbf{v}_1)|^2 + \lambda_2|(\mathbf{x}, \mathbf{v}_2)|^2 + \ldots + \lambda_n|(\mathbf{x}, \mathbf{v}_n)|^2}{|(\mathbf{x}, \mathbf{v}_1)|^2 + |(\mathbf{x}, \mathbf{v}_2)|^2 + \ldots + |(\mathbf{x}, \mathbf{v}_n)|^2}$$

waarbij $\lambda_1, \ldots, \lambda_n$ de eigenwaarden van A zijn en $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ een orthonormale basis van eigenvectoren. Als A hermites is dan zijn de eigenwaarden reëel en we kunnen ze dan naar aflopende grootte ordenen: $\lambda_1 \geq \ldots \geq \lambda_n$. Nu zien we onmiddellijk uit de bovenstaande uitdrukking dat het bereik van het Rayleighquotiënt precies het reële interval $[\lambda_n, \lambda_1]$ is.

De eigenwaarden van een hermitese matrix kunnen op eenvoudige wijze d.m.v. het Rayleighquotiënt worden gekarakteriseerd: zo is duidelijk dat

$$\lambda_1 = \max_{\mathbf{x} \neq \mathbf{0}} R_A(\mathbf{x}), \qquad \lambda_n = \min_{\mathbf{x} \neq \mathbf{0}} R_A(\mathbf{x}).$$

Als λ_1 en een eigenvector $\mathbf{v}_1$ bekend zijn, dan is $\lambda_2 = \max_{(\mathbf{x}, \mathbf{v}_1)=0, \mathbf{x} \neq \mathbf{0}} R_A(\mathbf{x})$, m.a.w. de een-na grootste eigenwaarde is het maximum van het Rayleighquotiënt over het orthogonaal complement van span$\{\mathbf{v}_1\}$. De eigenwaarde λ_3 is dan het maximum van $R_A(\mathbf{x})$ op het orthogonaal complement van span$\{\mathbf{v}_1, \mathbf{v}_2\}$, en zo verder. Het nadeel van deze methode is dat kleinere eigenwaarden worden uitgedrukt in termen van de eigenruimten van de grotere eigenwaarden en dit is een bezwaar als we eigenwaarden van verschillende hermitese matrices of afbeeldingen willen vergelijken. Een karakterisering van de eigenwaarden die onafhankelijk is van de eigenruimten van andere eigenwaarden is echter wel mogelijk: zo is λ_2 het minimum van $R_A(\mathbf{x})$ op de tweedimensionale lineaire deelruimte opgespannen door $\mathbf{v}_1$ en $\mathbf{v}_2$, en als we een willekeurige lineaire deelruimte W van dimensie 2 bekijken dan is het duidelijk uit de bovenstaande uitdrukking voor $R_A(\mathbf{x})$ dat het minimum van het Rayleighquotiënt daar nooit groter dan λ_2 kan zijn. Dus $\lambda_2 = \max_{W:\dim(W)=2} \min_{\mathbf{x}\in W,\mathbf{x}\neq\mathbf{0}} R_A(\mathbf{x})$. Op analoge manier zien we in dat

$$\lambda_k = \max_{W:\dim(W)=k} \min_{\mathbf{x}\in W,\mathbf{x}\neq\mathbf{0}} R_A(\mathbf{x}) \qquad (k = 1, 2, \ldots, n).$$

Dit noemen we het *minimax-principe.*

Beschouw de kwadratische vorm op $\mathbf{R}^3$ gegeven door

$$q(\mathbf{x}) = 2x_1^2 + 2x_2^2 - x_3^2 + 8x_1x_2 - 2x_1x_3 + 2x_2x_3.$$

$q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ met A de symmetrische matrix $\begin{pmatrix} 2 & 4 & -1 \\ 4 & 2 & 1 \\ -1 & 1 & -1 \end{pmatrix}$. A heeft eigenwaarden 6,-3 en 0, dus q heeft rang 2 en signatuur 1-1=0. T.o.v. een orthonormale basis van eigenvectoren van A is $q(\mathbf{x}) = 6y_1^2 - 3y_2^2$, waarbij $\mathbf{y} = U^T\mathbf{x}$ voor U een orthogonale matrix van eigenvectoren van A. Omdat $\|\mathbf{x}\| = \|\mathbf{y}\|$, neemt $q(\mathbf{x})$ op $\|\mathbf{x}\| = 1$ alle waarden tussen -3 en 6 aan. Merk op dat $q(\mathbf{x}) = R_A(\mathbf{x}) \cdot \|\mathbf{x}\|$.

Tenslotte leiden we een uitdrukking af voor de norm van een normale afbeelding:

Propositie 7.5: Zij V een eindig-dimensionale vectorruimte met een inwendig product en $T : V \to V$ een normale afbeelding. Laat $\lambda_1, \ldots, \lambda_n$ de eigenwaarden van T zijn zodanig geordend dat $|\lambda_1| \geq \ldots \geq |\lambda_n|$. Dan is $\|T\| = |\lambda_1|$. Een soortgelijk resultaat geldt voor normale matrices.

Bewijs: Laat $\{u_1, \ldots, u_n\}$ een orthonormale basis van eigenvectoren van T zijn. Voor $x \in V$ is $x = \sum_{j=1}^n (u_j, x)u_j$ en $T(x) = \sum_{j=1}^n \lambda_j(u_j, x)u_j$. Dan is, als $\|x\|^2 = \sum_{j=1}^n |(x, u_j)|^2 = 1$,

$$\|T(x)\|^2 = \sum_{j=1}^n |\lambda_j|^2 |(u_j, x)|^2 \leq |\lambda_1|^2$$

en gelijkheid geldt als $x = u_1$. Dus is

$$\|T\| = \max_{\|x\|=1} \|T(x)\| = |\lambda_1|. \quad \diamond$$

Voorbeeld: De matrix $A = \begin{pmatrix} 2 & 4 & -1 \\ 4 & 2 & 1 \\ -1 & 1 & -1 \end{pmatrix}$ is normaal en heeft eigenwaarden 6,-3 en 0. Dus is $\|A\| = 6$.

VIII. POSITIEF-DEFINIETE MATRICES

Zij A een normale $n \times n$-matrix. A heet *positief-definiet* als $\mathbf{x}^* A\mathbf{x} > 0$ voor alle $\mathbf{x} \in \mathbf{C}^n$, $\mathbf{x} \neq \mathbf{0}$ (in het complexe geval) resp. $\mathbf{x}^T A\mathbf{x} > 0$ voor alle $\mathbf{x} \in \mathbf{R}^n$, $\mathbf{x} \neq \mathbf{0}$ (in het reële geval). A heet *positief semi-definiet* als $\mathbf{x}^* A\mathbf{x} \geq 0$ voor alle $\mathbf{x} \in \mathbf{C}^n$ (in het complexe geval) resp. $\mathbf{x}^T A\mathbf{x} \geq 0$ voor alle $\mathbf{x} \in \mathbf{R}^n$ (in het reële geval). Zoals we in hoofdstuk 4 al gezien hebben, zijn alle inproducten op $\mathbf{C}^n$ (resp. $\mathbf{R}^n$) van de vorm $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* A\mathbf{y}$ (resp. $\mathbf{x}^T A\mathbf{y}$) waarbij A een positief-definiete $n \times n$-matrix is. We noteren kort $A > O$ als A positief-definiet is. (resp. $A \geq O$ als A positief-semidefiniet). We gaan in het volgende steeds uit van het complexe geval. Voor reële matrices gelden dezelfde resultaten als we de hermites geadjungeerden A^* vervangen door de getransponeerden A^T.

Propositie 8.1. Een normale matrix A is positief-definiet (resp. positief-semidefiniet) als alle eigenwaarden van A positief (resp. niet-negatief) zijn. In het bijzonder zijn alle positief (semi-)definiete matrices hermites.

Bewijs: Uit de spectrale decompositie voor A volgt:

$$\mathbf{x}^* A\mathbf{x} = \lambda_1 \mathbf{x}^* \mathbf{v}_1 \mathbf{v}_1^* \mathbf{x} + \lambda_2 \mathbf{x}^* \mathbf{v}_2 \mathbf{v}_2^* \mathbf{x} + \ldots + \lambda_n \mathbf{x}^* \mathbf{v}_n \mathbf{v}_n^* \mathbf{x} = \lambda_1 |(\mathbf{x}, \mathbf{v}_1)|^2 + \lambda_2 |(\mathbf{x}, \mathbf{v}_2)|^2 + \ldots + \lambda_n |(\mathbf{x}, \mathbf{v}_n)|^2$$

voor $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ een orthonormale basis van eigenvectoren van A en $\lambda_1, \ldots, \lambda_n$ de eigenwaarden. Hieruit volgt de bewering direct. $\diamond$

Propositie 8.2. 1. Een $n \times n$-matrix A is positief-semidefiniet dan en slechts dan als $A = B^* B$ voor zekere matrix B. A is positief-definiet als bovendien B inverteerbaar is.
2. Als A een positief-definiete $n \times n$-matrix is en U is een inverteerbare $n \times n$-matrix, dan is $U^* AU$ positief-definiet. Als A positief-semidefiniet is, dan is $U^* AU$ positief-semidefiniet.

Bewijs: 1. Voor $\mathbf{x} \in \mathbf{C}^n$ is $\mathbf{x}^* B^* B\mathbf{x} = (B\mathbf{x}, B\mathbf{x}) = \|B\mathbf{x}\|^2$. $\|B\mathbf{x}\| \geq 0$ voor alle $\mathbf{x}$ en als B inverteerbaar is, dan is $\|B\mathbf{x}\| > 0$ voor $\mathbf{x} \neq \mathbf{0}$. Omgekeerd, zij A positief-(semi)definiet. Dan is volgens de spectraaldecompositie $A = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^* + \ldots + \lambda_n \mathbf{v}_n \mathbf{v}_n^*$ met $\lambda_i > 0$ (resp. ≥ 0) en $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ orthonormaal. Laat nu $B = \sqrt{\lambda_1} \mathbf{v}_1 \mathbf{v}_1^* + \ldots + \sqrt{\lambda_n} \mathbf{v}_n \mathbf{v}_n^*$. Dan is $B = B^*$ en $A = B^2 = B^* B$. 2. Doe dit zelf. $\diamond$

Opmerking: De matrix $B = \sqrt{\lambda_1} \mathbf{v}_1 \mathbf{v}_1^* + \ldots + \sqrt{\lambda_n} \mathbf{v}_n \mathbf{v}_n^*$ is de enige positief-(semi)definiete matrix B zodat $A = B^2$. B heet de *vierkantswortel* uit A. We noteren $B = \sqrt{A}$. Op dezelfde manier kunnen we andere functies van normale (resp. hermitese of positief-(semi)definiete) matrices definiëren, zoals A^p voor $p > 0$ (en ook voor $p < 0$ indien A hermites en inverteerbaar is), en $\log(A)$ voor A positief-definiet.

Voorbeeld: Laat $A = \frac{1}{5} \begin{pmatrix} 8 & 6 \\ 6 & 17 \end{pmatrix}$. A is positief-definiet, heeft eigenwaarden 1 en 4 en heeft spectraaldecompositie

$$A = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}.$$

Dan is

$$\sqrt{A} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 6 & 2 \\ 2 & 9 \end{pmatrix}.$$

Gevolg 8.3: Laat A, B hermitese reële of complexe $n \times n$-matrices zijn, zodanig dat $A > 0$. Dan bestaat er een (reële resp. complexe) inverteerbare matrix Q zodanig dat $Q^*AQ = I_n$ en $Q^*BQ = D$ een diagonaalmatrix is.

Bewijs: Volgens Propositie 8.2 is $A = C^*C$ voor zekere positief-definiete matrix C. Dan is $(C^{-1})^*AC^{-1} = I_n$ en $(C^{-1})^*BC^{-1}$ is hermites. Volgens stelling 7.4 (resp. 7.2) bestaat er een orthogonale resp. unitaire matrix U zodanig dat $U^*(C^{-1})^*BC^{-1}U = D$ een diagonaalmatrix is, en uiteraard is $U^*U = I_n$. Laat nu $Q = C^{-1}U$. $\diamond$

M.b.v. het minimax-principe kunnen we de eigenwaarden vergelijken van twee hermitese matrices waarvan het verschil een positief-semidefiniete matrix is:

Propositie 8.4. Laat A en B twee hermitese $n \times n$-matrices zijn zodanig dat $A - B \geq O$. Laat $\lambda_1 \geq \ldots \geq \lambda_n$ de eigenwaarden van A zijn en $\mu_1 \geq \ldots \geq \mu_n$ de eigenwaarden van B. Dan is

$$\lambda_1 \geq \mu_1, \lambda_2 \geq \mu_2, \ldots, \lambda_n \geq \mu_n.$$

Bewijs: Omdat $A - B$ positief-semidefiniet is, is het Rayleighquotiënt $R_{A-B}(\mathbf{x}) \geq 0$ voor alle $\mathbf{x} \neq \mathbf{0}$. Maar $R_{A-B}(\mathbf{x}) = R_A(\mathbf{x}) - R_B(\mathbf{x})$. De bewering volgt nu onmiddellijk uit het minimax-principe. $\diamond$

M.b.v. het volgende criterium is na te gaan of een matrix A positief-definiet is zonder de eigenwaarden en eigenvectoren te bepalen.

Propositie 8.5: (a.) Zij A een $n \times n$-matrix en zij A_k de $k \times k$-deelmatrix die uit A ontstaat door de laatste $n-k$ rijen en kolommen van A weg te laten (i.h.b. is $A_n = A$). Dan is A positief-definiet dan en slechts dan als A hermites is en $\det(A_k) > 0$ voor $k = 1, 2, \ldots, n$.

Bewijs: Voor $n = 1$ is de bewering zeker waar. We passen inductie naar n toe. Uit Propositie 3.1 volgt dat $\det(A) > 0$ als A positief-definiet is. Laat nu $1 \leq k \leq n$ en laat $\mathbf{x} = (\mathbf{x}_k, \mathbf{0})^T = (x_1, \ldots, x_k, 0, \ldots, 0)^T$, $\mathbf{x}_k \in K^k$. Daar $\mathbf{x}^T A \mathbf{x} = \mathbf{x}_k^T A_k \mathbf{x}_k$, volgt dat A_k positief-definiet is als A positief definiet is. I.h.b. is $\det(A_k) > 0$ voor $k = 1, \ldots, n$. Omgekeerd, neem aan dat $\det(A_k) > 0$ voor $k = 1, \ldots, n$. Volgens de inductieveronderstelling is A_k positief-definiet voor $1 \leq k < n$. Verder is $A = \begin{pmatrix} A_{n-1} & \mathbf{b} \\ \mathbf{b}^T & \alpha \end{pmatrix}$, waarbij dus $\alpha = a_{nn}$. Omdat A_{n-1} positief-definiet en dus inverteerbaar is, is er een unieke vector $\mathbf{c} \in K^{n-1}$ zodanig dat $A_{n-1}\mathbf{c} + \mathbf{b} = \mathbf{0}$. Laat $C = \begin{pmatrix} I_{n-1} & \mathbf{c} \\ \mathbf{0}^T & 1 \end{pmatrix}$. Dan is $\det(C) = 1$ en is de matrix $C^*AC = \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0}^T & \alpha' \end{pmatrix}$ met $\alpha' = \mathbf{c}^* A_{n-1}\mathbf{c} +$

$\mathbf{b}^*\mathbf{c} + \mathbf{c}^*\mathbf{b} + \alpha = -\mathbf{c}^* A_{n-1}\mathbf{c} + \alpha$ positief-definiet dan en slechts dan als A positief-definiet is (volgens Propositie 8.2). Tevens is $\det(A) = \det(C^*AC) = \det(A_{n-1})\alpha'$. Omdat $\det(A) > 0$, is $\alpha' > 0$ en verder is, voor $\mathbf{x} = (\mathbf{x}_{n-1}, x_n)^T \neq \mathbf{0}^T$,

$$\mathbf{x}^* C^* AC\mathbf{x} = \mathbf{x}^*_{n-1} A_{n-1}\mathbf{x}_{n-1} + \alpha'|x_n|^2 > 0.$$

Maar dan is C^*AC, en dus ook A, positief-definiet. Merk tenslotte nog op dat als $A = A_n$ positief-definiet is, dan

$$\alpha' = \frac{\det(A_n)}{\det(A_{n-1})} = -\mathbf{c}^* A_{n-1}\mathbf{c} + a_{nn} \leq a_{nn}. \quad \diamond \tag{8.1}$$

Opmerking: Door het in het bewijs gegeven argument herhaald toe te passen (d.w.z. ook voor $A_{n-1}, A_{n-2}, \ldots$) zien we: als A een positief-definiete matrix is, dan is er een rechterbovendriehoeks-matrix B (met enen op de hoofddiagonaal) zodanig dat $B^*AB = \mathrm{diag}(d_1, \ldots, d_n)$ een diagonaal-matrix is. Verder is $d_i = \dfrac{\det(A_i)}{\det(A_{i-1})}$.

Toepassing: De hermitese 2×2-matrix $A = \begin{pmatrix} a & b \\ \bar{b} & c \end{pmatrix}$ is positief definiet dan en slechts dan indien $a > 0$ en $ac - |b|^2 > 0$.

Voorbeeld: De matrix $A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 2 \\ -1 & 2 & a \end{pmatrix}$. De matrices $A_1 = (1)$ en $A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ hebben beide positieve determinant. De matrix $A_3 = A$ heeft determinant $a - 10$. A is dus positief-definiet als $a > 10$.

Door (8.1) herhaald toe te passen volgt onmiddellijk het volgende resultaat:

Propositie 8.6: Zij $A = (a_{ij})$ een positief-definiete $n \times n$-matrix. Dan is

$$0 < \det(A) \leq \prod_{j=1}^{n} a_{jj}. \tag{8.2}$$

Een gevolg van Propositie 8.6 is de volgende ongelijkheid voor de determinant van een willekeurige matrix:

Stelling 8.7: (*de ongelijkheid van Hadamard*) Zij B een $n \times n$-matrix met kolomvectoren $\mathbf{b}_1, \ldots, \mathbf{b}_n$. Dan is

$$|\det(B)| \leq \prod_{j=1}^{n} \|\mathbf{b}_j\|. \tag{8.3}$$

Bewijs: Als B niet-inverteerbaar is dan is $\det(B) = 0$ en valt er niets te bewijzen. Als B inverteerbaar is, dan is $B^T B$ positief-definiet. Er geldt dat $(B^T B)_{jj} = \|\mathbf{b}_j\|^2$ (vergelijk de discussie over Gram-matrices in hoofdstuk IV). Volgens Propositie 2.1 en 8.6 geldt nu

$$|\det B|^2 = \det B^T \det B = \det B^T B \leq \prod_{j=1}^{n} \|\mathbf{b}_j\|^2. \quad \diamond$$

Opmerking: Meetkundig geïnterpreteerd zegt de ongelijkheid van Hadamard dat de inhoud van een parallellopipedum met gegeven ribben het grootst is als de ribben onderling loodrecht staan.

De polaire decompositie.

Zij A een inverteerbare $n \times n$-matrix. Dan is A^*A positief definiet en heeft een unieke wortel $S = \sqrt{A^*A}$. S is positief-definiet en dus volgt uit $S^2 = A^*A$ dat $(S^{-1}A^*)(AS^{-1}) = I_n$. De matrix $U = AS^{-1}$ is dus een unitaire matrix. We hebben aangetoond dat een inverteerbare matrix het product is van een unitaire en een positief-definiete matrix: $A = US$. Deze decompositie $A = US$ is uniek en wordt de *polaire decompositie* van A genoemd. Merk op dat $A = US = (USU^*)U = S'U$ waarbij S' opnieuw positief-definiet is. Voor een niet-inverteerbare matrix A bestaat er overigens ook een decompositie $A = SU$ met S positief-semidefiniet, maar deze is niet meer uniek.

Voorbeeld: Laat $A = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}$. Dan is $A^*A = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix} = S^2$ en $S = \sqrt{A^*A} = \begin{pmatrix} \sqrt{8} & 0 \\ 0 & \sqrt{2} \end{pmatrix}$.
Tenslotte is $U = AS^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ dus de polaire decompositie van A is

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{8} & 0 \\ 0 & \sqrt{2} \end{pmatrix}.$$

De singuliere waardendecompositie van een matrix.

Laat A een willekeurige $m \times n$-matrix zijn. De matrix A^*A is een positief-semidefiniete $n \times n$-matrix en heeft n niet noodzakelijk verschillende niet-negatieve eigenwaarden die we in aflopende volgorde noteren als $\sigma_1^2 \geq \sigma_2^2 \geq \ldots \geq \sigma_n^2$ waarbij alle $\sigma_i \geq 0$ zijn. De getallen $\sigma_1, \ldots, \sigma_n$ heten de *singuliere waarden* van A. We tonen het volgende opmerkelijke resultaat aan:

Stelling 8.8. Zij A een willekeurige complexe $m \times n$-matrix. Dan zijn er orthonormale stelsels $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ in $\mathbf{C}^m$ en $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in $\mathbf{C}^n$ zodat

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^* + \ldots + \sigma_k \mathbf{u}_k \mathbf{v}_k^*$$

waarbij $\sigma_1, \ldots, \sigma_k$ de positieve singuliere waarden van A zijn. k is gelijk aan de rang van A. Voor een reële matrix geldt een analoog resultaat.

Bewijs: Zonder beperking der algemeenheid nemen we aan dat $m \geq n$ (het geval $n > m$ volgt dan door de hermites geadjungeerde van de matrix te nemen). A^*A is een positief-semidefiniete $n \times n$-matrix met eigenwaarden $\sigma_1^2 \geq \ldots \geq \sigma_n^2$, waarbij $\sigma_j = 0$ voor $j > k$ en $\sigma_j > 0$ voor $j \leq k$. Laat $\mathbf{v}_1, \ldots, \mathbf{v}_n$ een corresponderende orthonormale basis van eigenvectoren zijn. Definieer nu voor $i = 1, \ldots, k$ de vectoren $\mathbf{u}_i \in \mathbf{C}^m$ d.m.v. $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$. Dan is

$$\sigma_i \sigma_j (\mathbf{u}_i, \mathbf{u}_j) = (\sigma_i \mathbf{u}_i, \sigma_j \mathbf{u}_j) = (A\mathbf{v}_i, A\mathbf{v}_j) = (A^*A\mathbf{v}_i, \mathbf{v}_j) = \sigma_i^2 (\mathbf{v}_i, \mathbf{v}_j)$$

en omdat $(\mathbf{v}_i, \mathbf{v}_j) = \delta_{ij}$, is ook $(\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}$, m.a.w. $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is een orthonormaal stelsel. Vul dit aan tot een orthonormale basis van $\mathbf{C}^m$. Nu geldt

$$A(\mathbf{v}_1, \ldots, \mathbf{v}_n) = (A\mathbf{v}_1, \ldots, A\mathbf{v}_n) = (\sigma_1\mathbf{u}_1, \ldots, \sigma_k\mathbf{u}_k, 0, \ldots, 0)$$

ofwel $AV = U\Sigma$ waarbij U en V de unitaire matrices zijn met kolomvectoren $\mathbf{u}_1, \ldots, \mathbf{u}_m$ resp. $\mathbf{v}_1, \ldots, \mathbf{v}_n$ en $\Sigma = \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_n \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$. Maar dan is $A = U\Sigma V^*$ en dit kunnen we uitwerken tot

$$A = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m) \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_n \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \mathbf{v}_1^* \\ \mathbf{v}_2^* \\ \vdots \\ \mathbf{v}_n^* \end{pmatrix} = \sigma_1\mathbf{u}_1\mathbf{v}_1^* + \ldots + \sigma_n\mathbf{u}_n\mathbf{v}_n^*$$

dus $A = \sigma_1\mathbf{u}_1\mathbf{v}_1^* + \ldots + \sigma_k\mathbf{u}_k\mathbf{v}_k^*$ en de rang van A is inderdaad precies k. $\diamond$

Opmerking: Met de notatie van de stelling geldt:

1. $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is een basis van de kolomruimte van A.
2. $\{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ is een basis van de nulruimte van A.
3. $\{\overline{\mathbf{v}_1}, \ldots, \overline{\mathbf{v}_k}\}$ is een basis van de rijruimte van A.
4. Daar $A^* = \sigma_1\mathbf{v}_1\mathbf{u}_1^* + \ldots + \sigma_k\mathbf{v}_k\mathbf{u}_k^*$ is $\{u_{k+1}, \ldots, u_m\}$ een basis van de nulruimte van A^*.

De ontbinding $A = U\Sigma V^*$ heet een *singuliere waardendecompositie* (SVD) van A.

Nogmaals de polaire decompositie van een $n \times n$-matrix.

Zij A een $n \times n$-matrix. Laat $U\Sigma V^*$ een SVD van A zijn waarbij Σ een positief-semidefiniete diagonaalmatrix is (Σ is positief definiet als A inverteerbaar is). Daar U en V unitaire matrices zijn, is

$$A = (UV^*)(V\Sigma V^*) = (U\Sigma U^*)(UV^*)$$

dus $A = OS = S'O$ waarbij $O = UV^*$ een unitaire matrix is en S,S' positief-(semi)definiete matrices zijn. Daar $O = e^{iH}$ voor een hermitese matrix H, is de ontbinding ook te schrijven als $A = e^{iH}S = S'e^{iH}$ met H hermites en S, S' positief-(semi)definiet. Dit is precies de polaire decompositie van A.

INDEX

afstand	37
afstand (tot lineaire deelruimte)	21
algebra	5
alternerende vorm	15
annihilator	50
antilineair	36
antihermites	58
antisymmetrische matrix	3
associatief	1
Banachruimte	53
basis	2
basistransformatie	8
beeld	6
bereik	6
bijectief	6
bilineair	35
Cauchyrij, fundamentaalrij	53
Cayley-Hamilton	32
cofactor	19
commutatief	1
continu	55
coördinaatafbeelding	7
coördinaatvector	7
determinant	15
diagonaliseerbaar	24, 46
differentieerbaar	55
dimensie	4
dimensiestelling	7
directe som	9
distributief	1
draaiing	48
draaispiegeling	48
driehoeksongelijkheid	37
duale (vectorruimte, basis)	49
eigenruimte	23

eigenvector	23
eigenwaarde	23
e-macht van een matrix	54
endomorfisme	5
equivalentieklasse	11
equivalentierelatie	11
geadjungeerde matrix	20
gegeneraliseerde eigenruimte, eigenvector	26
gelijkvormig (matrix)	9
genormeerde vectorruimte	37
Gershgorin, cirkels van	33
getransponeerde afbeelding	50
Gram-matrix	44
Gram-Schmidt	38
Hadamard, ongelijkheid van	66
hermites (afbeelding, matrix)	40, 41
hermites (inproduct)	36
hermites geadjungeerde	37
Hilbertruimte	53
identieke afbeelding	5
indefiniet	62
injectief	6
inverse (afbeelding)	6
inverse beeld	6
inverteerbaar	6
inwendig product, inproduct	35
isomorf	6
Jordanbasis	28
Jordanblok	28
Jordan-normaalvorm	28
kanonieke afbeelding	12
karakteristiek polynoom	23
kern, nulruimte	6
kleinste kwadraten	44
kwadratische vorm	61
Levi-Civitasymbool	15
lichaam (van scalairen)	1
lineair onafhankelijk	2
lineair onafhankelijk (modulo een lin.deelruimte)	13

lineaire afbeelding	4
lineaire deelruimte	2
matrix van een afbeelding	8
minimax-principe	63
minimumpolynoom	31
minor, onderdeterminant	17
(algebraïsche, meetkundige) multipliciteit	23
negatief-definiet	62
nilpotent	24
norm (van een vector)	37
norm (van een afbeelding, matrix)	51, 52, 63
normaal (matrix, afbeelding)	58
opspansel	2
oriëntatie	21
orthogonaal (vector, stelsel)	38
orthogonaal (matrix, afbeelding,..)	44,45
orthogonaal complement	41
orthogonale projectie	42
orthonormaal	38
permutatie	15
polaire decompositie	67
positief-definiet	35, 64
projectie	9
pseudoinverse	44
quotiëntafbeelding	12
quotiëntvectorruimte	12
quotiëntverzameling	11
QR-decompositie	39
rang	6, 61
Rayleighquotiënt	62
representant	11
restrictie (afbeelding)	12
scalaire vermenigvuldiging	1
Schwarz, ongelijkheid van	37
semidefiniet	64
seminorm	37
sesquilineair	36
signatuur	61
singuliere waarden	52, 67

singuliere waardendecompositie	67
som	9
spectraaldecompositie (normale, symmetrische matrix)	59, 60
spectraalstelling (unitaire, orthogonale afbeelding)	45
spectraalstelling (normale afbeelding)	59
spectrum	23
spoor	20
standaardbasis	2
standaardmatrix	8
standaard-hermites (inproduct)	36
standaardrepresentatie	2
surjectief	6
symmetrisch (afbeelding)	40, 60
symmetrisch (matrix)	3, 60
symmetrisch (vorm)	35
tensorproduct	13
traagheidsstelling van Sylvester	61
unitair (vectorruimte)	36
unitair (matrix, afbeelding)	44, 45
unitair diagonaliseerbaar	46
unitair gelijkvormig	46
Vandermonde, determinant van	18
vector	1
vectorruimte	1
vectorruimte-isomorfisme	6
vierkantswortel (van een matrix)	64
volledig (vectorruimte)	53
volume	21
Wronskiaan	18
zelfgeadjungeerd	40