

Hoofdstuk 7

Jordan normaalvorm

Zoals we zagen hangt de matrix die behoort bij een lineaire transformatie af van de keuze van een basis voor de ruimte. In dit hoofdstuk buigen we ons over de vraag of er altijd een basiskeuze bestaat waarvoor de matrix diagonaalvorm heeft; het antwoord blijkt negatief te zijn, en de Jordan normaalvorm is een antwoord op de vraag hoe mooi de matrix wel altijd gemaakt kan worden.

Definities 7.1 Een matrix $D \in M_{n,n}(L)$ heet *diagonaalmatrix* als $D_{ij} = 0$ voor alle $i \neq j$, (met $1 \leq i, j \leq n$). Een matrix $M \in M_{n,n}(L)$ heet *diagonaliseerbaar* als er een met M geconjugeerde matrix bestaat die diagonaalmatrix is. Een lineaire transformatie T van een eindig-dimensionale vectorruimte V heet *diagonaliseerbaar* als er een basis $\mathcal{B}$ bestaat zodat $M_T^{\mathcal{B}}$ een diagonaalmatrix is.

Stelling 7.2 *Een lineaire transformatie T van een eindig-dimensionale vectorruimte V is diagonaliseerbaar dan en slechts dan als $M_T^{\mathcal{B}}$ diagonaliseerbaar is voor elke basis $\mathcal{B}$ van V .*

Bewijs. Dat volgt onmiddellijk uit de definities, omdat twee matrices behorende bij dezelfde transformatie maar ten opzichte van verschillende bases met elkaar geconjugeerd zijn.

Laat nu T een diagonaliseerbare lineaire transformatie zijn, en $\mathcal{B} = b_1, b_2, \dots, b_n$ een basis ten opzichte waarvan $M_T^{\mathcal{B}} = D$ diagonaalmatrix is. Dan is $Db_i = D_{ii}b_i$, met $D_{ii} \in L$, met andere woorden: de vectoren b_i worden door D met een factor D_{ii} vermenigvuldigd, en zijn dus allen eigenvector van T .

Omgekeerd, als je in V een basis $\mathcal{B} = b_1, b_2, \dots, b_n$ kunt vinden bestaande uit eigenvectoren, zodat $Tb_i = \lambda_i b_i$, voor zekere $\lambda_i \in L$, dan is $M_T^{\mathcal{B}}$ de diagonaalmatrix D met op de diagonaal $D_{ii} = \lambda_i$.

We hebben zo, samenvattend, de volgende stelling bewezen.

Stelling 7.3 *Een lineaire transformatie $T: V \rightarrow V$ is diagonaliseerbaar dan en slechts dan als er een basis van V is die bestaat uit eigenvectoren voor T .*

De grootste nog resterende complicatie is gelegen in de mogelijkheid dat eigenwaarden met multipliciteiten voorkomen. Eerst laten we zien dat de situatie overzichtelijk is wanneer alle multipliciteiten 1 zijn.

Stelling 7.4 *Als $x_1, \dots, x_k$ eigenvectoren van T zijn met eigenwaarden $\lambda_1, \dots, \lambda_k$, en er geldt $\lambda_i \neq \lambda_j$ voor $1 \leq i < j \leq k$, dan zijn $x_1, \dots, x_k$ lineair onafhankelijk.*

Bewijs. Inductie naar k . Voor $k = 1$ is de stelling zeker waar. Veronderstel eens dat

$$\mu_1 x_1 + \cdots + \mu_k x_k = 0,$$

met $k > 1$, en x_i als in de stelling. Er geldt dan

$$(T - \lambda_k)(\mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_k x_k) = (\lambda_1 - \lambda_k)\mu_1 x_1 + \cdots + (\lambda_{k-1} - \lambda_k)\mu_{k-1} x_{k-1} = 0.$$

Op grond van de inductiehypothese is $x_1, x_2, \dots, x_{k-1}$ lineair onafhankelijk, dus

$$(\lambda_j - \lambda_k)\mu_j = 0,$$

voor $j = 1, 2, \dots, k-1$, en omdat $\lambda_j \neq \lambda_k$ moet $\mu_j = 0$, voor $j = 1, 2, \dots, k-1$. Maar dan ook $\mu_k = 0$, en we zijn klaar!

Gevolg 7.5 *Als T een transformatie van een n -dimensionale L -vectorruimte is, en T heeft over L precies n verschillende eigenwaarden, dan is T diagonaliseerbaar over L .*

Bewijs. De $n = \dim V$ verschillende eigenvectoren x_i zijn op grond van voorgaande stelling onafhankelijk en vormen dus een basis voor V , ten opzichte waarvan T diagonaalvorm heeft.

Let wel dat er in feite twee voorwaarden staan in dit gevolg: alle nulpunten van het karakteristieke polynoom moeten in L liggen en verschillend zijn!

Stelling 7.6 *Als T een diagonaliseerbare transformatie van de eindig-dimensionale L -vectorruimte V is, dan splitst het karakteristieke polynoom $p_T(x) \in L[x]$ in $L[x]$ volledig in lineaire factoren.*

Bewijs. Ten opzichte van zekere basis is T een diagonaalmatrix over L ; op de diagonaal staan dan de eigenwaarden λ_i van T . Die moeten in het lichaam L zitten, en dus valt p_T over L uiteen in de lineaire factoren $x - \lambda_i$.

Het belangrijkste onderdeel van de laatste stelling is dat p_T al over het lichaam L in lineaire factoren uiteen valt. Dat garandeert natuurlijk nog niet dat alle eigenwaarden verschillend zijn. De omkering van Stelling 7.6 blijkt ook niet te gelden!

Voorbeeld 7.7 Beschouw de lineaire transformatie van $\mathbb{Q}^3$ gegeven door

$$A = \begin{pmatrix} 3 & 1 & -2 \\ -1 & 0 & 5 \\ -1 & -1 & 4 \end{pmatrix}.$$

Het karakteristieke polynoom van A is

$$\det |\lambda I_3 - A| = (\lambda - 2)^2(\lambda - 3).$$

Om E_2 te bepalen kijk je naar de kern van

$$\begin{pmatrix} -1 & -1 & 2 \\ 1 & 2 & -5 \\ 1 & 1 & -2 \end{pmatrix};$$

omdat de rang van deze matrix 2 is, zal de kern dimensie 1 hebben. Daarom is E_2 ook 1-dimensionaal, net als E_3 . Andere eigenwaarden zijn er niet, en dus kan er geen basis voor $\mathbb{Q}^3$ zijn bestaande uit eigenvectoren: A is niet diagonaliseerbaar!

Het niet-diagonaliseerbaar zijn van A in het voorbeeld werd veroorzaakt doordat voor een zekere eigenwaarde ($\lambda = 2$) de multipliciteit van die eigenwaarde (2) groter was dan de dimensie (1) van de eigenruimte E_λ .

De dimensie van E_λ wordt ook wel de *meetkundige multipliciteit* van λ genoemd; de volgende stelling drukt dan uit dat de meetkundige multipliciteit ten hoogste gelijk is aan de algebraïsche multipliciteit.

Stelling 7.8 *Als T eigenwaarde λ met multipliciteit m heeft dan geldt*

$$1 \leq \dim E_\lambda \leq m.$$

Bewijs. Omdat λ eigenwaarde is, is $\det(\lambda I_n - M_T) = 0$, dus geldt $\dim E_\lambda = \dim \text{Ker}(\lambda I_n - M_T) \geq 1$.

Voor de andere ongelijkheid vullen we een basis van de deelruimte E_λ van dimensie ℓ aan tot een basis $\mathcal{B}$ voor V . Ten opzichte van die basis is de matrix

$$M_T^{\mathcal{B}} = \begin{pmatrix} D & A \\ 0 & C \end{pmatrix},$$

waar

$$D = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda \end{pmatrix},$$

een $\ell \times \ell$ matrix is, en dus (vgl. Lemma 6.33):

$$p_T(x) = (x - \lambda)^\ell \cdot \det(xI_{n-\ell} - C).$$

De multipliciteit m van het nulpunt λ is dus minstens $\ell = \dim E_\lambda$.

Definitie 7.9 Als W een L -vectorruimte is en $V_1, V_2, \dots, V_k$ zijn lineaire deelruimten van W , dan is de *som* van die deelruimten de lineaire deelruimte gedefinieerd door

$$V = \sum_{i=1}^k V_i = V_1 + V_2 + \cdots + V_k = \{v_1 + v_2 + \cdots + v_k : v_i \in V_i\}.$$

Een som heet een *directe som*, en we schrijven $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$, als geldt dat elke $v \in V$ op precies één manier te schrijven is als $v = v_1 + v_2 + \cdots + v_k$ met $v_i \in V_i$ voor $1 \leq i \leq k$.

Stelling 7.10 *Zij V een vectorruimte met lineaire deelruimten $V_1, V_2, \dots, V_k$. Dan zijn de volgende beweringen equivalent:*

- (i) $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$;
- (ii) $V = V_1 + V_2 + \cdots + V_k$ en $V_i \cap \sum_{j \neq i} V_j = \{0\}$ voor $i = 1, 2, \dots, k$;
- (iii) $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_k$ vormt een basis voor V als $\mathcal{B}_i$ een basis voor V_i is ($1 \leq i \leq k$).

Bewijs. (i) $\Rightarrow$ (ii) Als $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ dan zeker $V = V_1 + V_2 + \cdots + V_k$, en veronderstel dat $v \in V_i \cap \sum_{j \neq i} V_j$. Dan is $0 = 0 + 0 = v + (-v)$ en dat geeft twee verschillende manieren om $0 \in V$ als som van elementen van V_i en $\sum_{j \neq i} V_j$ te schrijven, in tegenspraak met de aanname dat de som direct is, tenzij $v = 0$. Dus is 0 de enige vector in de doorsnede.

(ii) $\Rightarrow$ (iii) Laat $\mathcal{B}_i = b_1^{(i)}, \dots, b_{n_i}^{(i)}$ een basis voor V_i zijn. Het is eenvoudig in te zien dat $\mathcal{B} = \mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_k$ dan een opspannend stelsel voor $V_1 + V_2 + \dots + V_k$ is. We moeten laten zien dat de voorwaarde $V_i \cap \sum_{j \neq i} V_j = \{0\}$ voor $i = 1, 2, \dots, k$ impliceert dat $\mathcal{B}$ een basis is. Veronderstel eens dat $\mathcal{B}$ geen basis is, dan is er kennelijk een afhankelijkheid van de vorm

$$\sum_{j=1}^k (\lambda_1^{(j)} b_1^{(j)} + \dots + \lambda_{n_j}^{(j)} b_{n_j}^{(j)}) = 0.$$

Veronderstel dat $\lambda_r^{(i)} \neq 0$ voor zekere $1 \leq r \leq n_i$; als voor elke $j \neq i$ en elke index m met $1 \leq m \leq n_j$ geldt dat $\lambda_m^{(j)} = 0$ dan is $w = \lambda_1^{(i)} b_1^{(i)} + \dots + \lambda_{n_i}^{(i)} b_{n_i}^{(i)} = 0$ met tenminste 1 coëfficiënt ongelijk aan 0, in tegenspraak met het feit dat $\mathcal{B}_i$ een basis voor V_i is. Maar dan moet $w \neq 0$, en is $w \in V_i$ en $-w \in \sum_{j \neq i} V_j$, en dus zit w in de doorsnede van beide, in tegenspraak met onze aanname. Kennelijk kan er geen $\lambda_r^{(i)} \neq 0$ bestaan: het stelsel $\mathcal{B}$ is onafhankelijk.

(iii) $\Rightarrow$ (i) Laat $\mathcal{B}_i$ een basis voor V_i zijn, en veronderstel dat $\mathcal{B} = \mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_k$ een basis voor V is; het is direct duidelijk dat $V = V_1 + V_2 + \dots + V_k$, we moeten aantonen dat deze som *direct* is. Veronderstel eens dat dit niet zo is, dat wil zeggen, er is een $v \in V$ die twee schrijfwijzen $v = v'_1 + \dots + v'_k = v''_1 + \dots + v''_k$ heeft met $v'_i, v''_i \in V_i$ voor $1 \leq i \leq k$. Voor tenminste één i is $v'_i \neq v''_i$; schrijf $w' = v'_i$ en $w'' = v''_i$, en ook $x' = \sum_{j \neq i} v'_j$ en $x'' = \sum_{j \neq i} v''_j$. Dan is $w' + x' = w'' + x'' = v$, en dus $w' - w'' = x'' - x'$. Schrijf elk van de v'_r en v''_r nu op de basis $\mathcal{B}_r$ voor $1 \leq r \leq k$; dan hebben we een niet-triviale lineaire combinatie van elementen van $\mathcal{B}$ met als som 0, in tegenspraak met de veronderstelling dat $\mathcal{B}$ een basis voor V is.

Gevolg 7.11 Als $V = V_1 \oplus \dots \oplus V_k$ dan is $\dim V = \dim V_1 + \dots + \dim V_k$.

Voorbeelden 7.12 Elke vectorruimte is de directe som van de 1-dimensionale deelruimten $\langle b_i \rangle$ opgespannen door een basisvector. Zo is L^3 natuurlijk te schrijven als $L^3 = L \oplus L \oplus L = L \cdot e_1 \oplus L \cdot e_2 \oplus L \cdot e_3$, maar ook is $L^3 = L^2 \oplus L$.

Als $U \subset V$ een lineaire deelruimte is, bestaat er altijd een deelruimte $T \subset V$ zodat $T \oplus U = V$. Dat is eenvoudig in te zien door een basis van U aan te vullen met vectoren $t_1, \dots, t_k$ tot een basis voor V . De ruimte opgespannen door de t_i kan dan voor T genomen worden; T heet in zo'n geval een *complement* van U .

Definitie 7.13 De *directe som van vierkante matrices* $A^{(t)} \in M_{n_t \times n_t}(L)$, for $t = 1, 2, \dots, k$, is de vierkante matrix $A \in M_{n \times n}(L)$, met $n = n_1 + n_2 + \dots + n_k$, gedefinieerd door:

$$A_{ij} = \begin{cases} A_{ij}^{(1)} & 0 < i, j \leq n_1 \\ A_{ij}^{(2)} & n_1 < i, j \leq n_2 \\ \vdots & \vdots \\ A_{ij}^{(k)} & n - n_k = n_1 + n_2 + \dots + n_{k-1} < i, j \leq n_k \\ 0 & \text{elders.} \end{cases}$$

Dit schrijven we wel als

$$A = A^{(1)} \oplus A^{(2)} \oplus \dots \oplus A^{(k)} = \begin{pmatrix} A^{(1)} & 0 & \dots & 0 \\ 0 & A^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A^{(k)} \end{pmatrix}.$$

Opmerkingen 7.14 Een matrix als in de directe som $A^{(1)} \oplus \dots \oplus A^{(k)}$ heet wel een *blok-diagonaal matrix* bestaande uit k blokken. Elke matrix is zo op flauwe manier een blok-diagonaal matrix bestaand uit 1 blok. Een $n \times n$ diagonaalmatrix is een blok-diagonaal matrix bestaande uit n blokken.

Het verband tussen directe sommen van deelruimten en van matrices wordt gegeven in de stelling beneden; verderop in het hoofdstuk zullen we ons bezighouden met het vinden van geschikte blokken.

Stelling 7.15 *Laat een eindig-dimensionale vectorruimte V met een lineaire transformatie T gegeven zijn, en veronderstel dat er deelruimten $U_1, U_2, \dots, U_k$ van V bestaan zodanig dat $V = U_1 \oplus U_2 \oplus \dots \oplus U_k$ en U_i is invariant ten opzichte van T . Dan geldt:*

$$M_T = M_{T_1} \oplus M_{T_2} \oplus \dots \oplus M_{T_k},$$

wanneer $T_i = T|_{U_i}$ de beperking tot U_i is, $M_T = M_T^{\mathcal{B}}$ de matrix van T ten opzichte van $\mathcal{B} = \mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_k$ is, en $M_{T_i} = M_{T_i}^{\mathcal{B}_i}$ de matrix van T_i ten opzichte van $\mathcal{B}_i$, voor bases $\mathcal{B}_i$ van U_i .

Bewijs. Bekijk het beeld van een basis vector uit $\mathcal{B}_j$: dat is een lineaire combinatie van de vectoren uit $\mathcal{B}_j$ zelf, gegeven door de juiste kolom uit M_{T_j} .

Stelling 7.16 *Laat T een lineaire transformatie zijn van de L -vectorruimte V , van eindige dimensie n . Dan zijn de volgende beweringen equivalent:*

- (i) T is diagonaliseerbaar;
- (ii) $p_T(x)$ splitst in $L[x]$ in lineaire factoren:

$$p_T = \prod_{\lambda} (x - \lambda)^{m_{\lambda}}$$

en $\dim E_{\lambda} = m_{\lambda}$, waar λ de verschillende eigenwaarden van T doorloopt.

- (iii) $V = \oplus_{\lambda} E_{\lambda}$.

Bewijs. Als (i) geldt, splitst p_T volledig in lineaire factoren over L volgens Stelling 7.6. Kies nu een basis $\mathcal{B}$ ten opzichte waarvan T door een diagonaalmatrix wordt gegeven; dan is elke $b_i \in \mathcal{B}$ eigenvector bij één van de eigenwaarden λ van T , en bovendien is het duidelijk dat $B_{\lambda} = \{b \in \mathcal{B} : Tb = \lambda b\}$ een basis is voor E_{λ} , voor elke eigenwaarde λ . Maar dan is $\#B_{\lambda} = \dim E_{\lambda}$ en

$$n = \#\mathcal{B} = \sum_{\lambda} \#B_{\lambda} = \sum_{\lambda} \dim E_{\lambda} \leq \sum_{\lambda} m_{\lambda} = \deg p_T = n,$$

op grond van Stelling 7.8, dus moet $\dim E_{\lambda} = m_{\lambda}$, voor elke eigenwaarde λ .

Als $\lambda' \neq \lambda$ dan is het duidelijk dat $E_{\lambda'} \cap E_{\lambda} = \{0\}$, want $\lambda'v = Tv = \lambda v$ voor een v in de doorsnijding. Omdat u' en u zelfs lineair onafhankelijk zijn als $u' \in E_{\lambda'}$ en $u \in E_{\lambda}$, op grond van Stelling 7.4, vormt $\cup_{\lambda} B_{\lambda}$ dan een onafhankelijk stelsel in V als B_{λ} een basis voor E_{λ} is. Geldt nu (ii), dan is $n = \#\cup_{\lambda} B_{\lambda}$, en dus is $\cup_{\lambda} B_{\lambda}$ een basis voor V , terwijl $E_{\lambda'} \cap E_{\lambda} = \{0\}$. Daaruit volgt (iii).

Tenslotte volgt (i) direct uit (iii) omdat een basis voor E_{λ} bestaat uit eigenvectoren van T . Ten opzichte van een basis voor V die bestaat uit de vereniging van bases voor de E_{λ} heeft T dus een diagonaalmatrixvoorstelling.

De stelling zegt dus, onder andere, dat T diagonaliseerbaar is dan en slechts dan als V de directe som is van zijn eigenruimten.

We hebben reeds gezien dat lineaire transformaties niet altijd diagonaliseerbaar zijn. Zelfs als het karakteristieke polynoom geheel in lineaire factoren splitst over het lichaam waarover we werken zien we uit Stelling 7.16 dat het voor diagonaliseerbaarheid nog noodzakelijk is dat de dimensies van alle eigenruimten gelijk zijn aan de multipliciteiten van de bijbehorende eigenwaarden.

Voorbeeld 7.17 Het is eenvoudig om voorbeelden van lineaire transformaties te maken (in vectorruimten van eindige dimensie over een willekeurig lichaam!) die niet diagonaliseerbaar zijn terwijl het karakteristieke polynoom wel volledig in lineaire factoren uiteenvalt. Zorg er bijvoorbeeld voor dat 0 de enige eigenwaarde is, door een bovendriehoeksmatrix voor de transformatie te kiezen met bovendien *op* de diagonaal allemaal nullen. Dan is het karakteristieke polynoom gelijk aan x^n , en 0 de enige eigenwaarde, met multipliciteit n . De eigenruimte bij $\lambda = 0$ is de kern van de matrix, waarvan de dimensie gelijk is aan $n - r$, met r de rang van de matrix. Zodra er niet-nul elementen in de bovendriehoek staan is de rang groter dan nul en dus de dimensie van de kern kleiner dan n .

Wat we in dit hoofdstuk verder willen laten zien is dat een lineaire transformatie T van een eindig-dimensionale vectorruimte V , mits het karakteristieke polynoom ervan geheel in lineaire factoren uiteenvalt, ‘bijna diagonaliseerbaar’ is, in een heel precieze zin. Namelijk, dat er dan een basis voor V bestaat ten opzichte waarvan de transformatie een matrix heeft die bestaat uit ‘bijna-diagonaal’ blokken. Preciezer, ten opzichte van die basis is de matrix van T de directe som van Jordan blokken, waar een $k \times k$ *Jordan blok* $J_k(\lambda)$ een $k \times k$ matrix van de vorm

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

is. Een directe som van Jordan blokken heet een *Jordan matrix*.

Definitie 7.18 Zij T een transformatie van de eindig-dimensionale vectorruimte V . Met I geven we de identieke afbeelding id_V aan. Een vector $v \in V$ met $v \neq 0$ is *gegeneraliseerde eigenvector bij* λ als er een $k \in \mathbb{Z}_{\geq 1}$ bestaat zodat $(\lambda I - T)^k(v) = 0$. De *gegeneraliseerde eigenruimte bij* λ van T is de deelruimte

$$G_\lambda = \{v \in V \mid \exists k \in \mathbb{Z}_{\geq 1} : (\lambda I - T)^k(v) = 0\}.$$

Voorbeeld 7.19 Laat de 6×6 Jordanmatrix J gegeven zijn als directe som van $J_1(2) \oplus J_2(5) \oplus J_3(2)$:

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

ten opzichte van de basis $b_1, \dots, b_6$. De vectoren b_1, b_2, b_4 zijn dan eigenvectoren (bij eigenwaarden 2, 5, 2), terwijl b_3, b_5, b_6 gegeneraliseerde eigenvector zijn, bij

5, 2, 2. Immers, $(5I_6 - J)(b_3) = b_2$, dus $(5I_6 - J)^2(b_3) = 0$, en net zo $(2I_6 - J)(b_6) = b_5$, en $(2I_6 - J)(b_5) = b_4$, zodat $(2I_6 - J)^3(b_6) = (2I_6 - J)^2(b_5) = (2I_6 - J)(b_4) = 0$. De ruimte G_2 is de directe som van een 1-dimensionale en een 3-dimensionale deelruimte, terwijl E_2 dimensie 2 heeft.

Definitie 7.20 Als v een gegeneraliseerde eigenvector van T is bij λ , en m minimaal zodat $(\lambda I - T)^m(v) = 0$, dan noemen we $(\lambda I - T)^{m-1}(v)$, $(\lambda I - T)^{m-2}(v)$, $\dots$, $(\lambda I - T)(v)$, v , een *cykel van gegeneraliseerde eigenvectoren* voor T van lengte m . Merk op dat de eerste vector in zo'n cykel de enige eigenvector is.

Lemma 7.21 *De vectoren in een cykel van gegeneraliseerde eigenvectoren voor T zijn lineair onafhankelijk.*

Bewijs. Met inductie naar de lengte m . Voor $m = 1$ is het duidelijk. Laat nu $v_{m-1}, \dots, v_0$ zo'n cykel zijn bij λ , en veronderstel dat $\mu_{m-1}v_{m-1} + \dots + \mu_0v_0 = 0$. Pas $(\lambda I - T)$ toe: omdat $(\lambda I - T)v_i = v_{i+1}$ voor $i = m-2, \dots, 0$ en $(\lambda I - T)v_{m-1} = 0$, geeft dat $\mu_{m-2}v_{m-1} + \dots + \mu_0v_1 = 0$. Volgens de inductiehypothese zijn dan $\mu_i = 0$ voor $i = m-2, \dots, 0$. Dan ook $\mu_{m-1}v_{m-1} = 0$ en dus $\mu_{m-1} = 0$ (want $v_{m-1} \neq 0$ is eigenvector).

Stelling 7.22 *De matrix van de transformatie T is ten opzichte van een basis $\mathcal{B}$ een Jordanmatrix dan en slechts dan als $\mathcal{B}$ een vereniging van cyclen van gegeneraliseerde eigenvectoren is.*

Bewijs. Bij elk $k \times k$ Jordanblok hoort een cykel van lengte k , dat is duidelijk. Omgekeerd geeft zo'n cykel $v_{k-1}, \dots, v_0$ ook een Jordanblok, precies omdat steeds $(\lambda I - T)v_i = v_{i+1}$ voor $i = k-2, \dots, 0$ en $(\lambda I - T)v_{k-1} = 0$.

Het idee is nu om bij elke eigenvector van T een cykel van gegeneraliseerde eigenvectoren van maximale lengte te maken, en als basis voor de ruimte de vereniging van die cyclen te nemen. We moeten nog laten zien dat zo'n vereniging van cyclen een lineair onafhankelijk stelsel geeft, en dat we zo een aantal vectoren krijgen dat de dimensie van de hele ruimte is.

We zagen al dat de cyclen zelf louter onafhankelijke vectoren bevatten. De volgende stelling zegt dat cyclen bij *verschillende* eigenwaarden ook onafhankelijk zijn. Uit het voorbeeld blijkt dat dat niet genoeg is: er kunnen ook verschillende cyclen bij dezelfde λ zijn; merk ook op dat dit stelling 7.4 generaliseert.

Stelling 7.23 *Als $x_1, \dots, x_k$ gegeneraliseerde eigenvectoren van T zijn bij eigenwaarden $\lambda_1, \dots, \lambda_k$, en er geldt $\lambda_i \neq \lambda_j$ voor $1 \leq i < j \leq k$, dan zijn $x_1, \dots, x_k$ lineair onafhankelijk.*

Bewijs. Inductie naar k . Voor $k = 1$ is de stelling zeker waar. Zij $k > 1$, en x_i als in de stelling. Laat m_i minimaal zijn met de eigenschap dat $(\lambda_i I - T)^{m_i}(x_i) = 0$. Dan is, zoals we al zagen, $w_i = (\lambda_i I - T)^{m_i-1}(x_i)$ eigenvector voor T bij de eigenwaarde λ_i . Veronderstel nu eens dat

$$\mu_1 x_1 + \mu_2 x_2 + \dots + \mu_k x_k = 0,$$

dan is enerzijds

$$\left(\prod_{i=1}^{k-1} (\lambda_i I - T)^{m_i} \right) \cdot (\lambda_k I - T)^{m_k-1} (\mu_1 x_1 + \mu_2 x_2 + \dots + \mu_k x_k)$$

gelijk aan 0, maar anderzijds gelijk aan

$$\mu_k \left(\prod_{i=i}^{k-1} (\lambda_i I - T)^{m_i} \right) \cdot w_k = \mu_k \prod_{i=i}^{k-1} (\lambda_i - \lambda_k)^{m_i} \cdot w_i.$$

Maar w_k is eigenvector bij λ_k , dus niet nul, en ook is $\lambda_i \neq \lambda_k$. Maar dan moet $\mu_k = 0$. Op grond van de inductiehypothese is $x_1, \dots, x_{k-1}$ lineair onafhankelijk, dus zijn alle μ_i gelijk aan 0: alle x_i zijn lineair onafhankelijk.

Gevolg 7.24 Als $\mathcal{B}_i$ een basis is voor de gegeneraliseerde eigenruimte G_{λ_i} , en $\lambda_i \neq \lambda_j$ voor $i \neq j$ dan is $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_k$ een onafhankelijk stelsel.

Lemma 7.25 Als $C_1, \dots, C_k$ cycli van gegeneraliseerde eigenvectoren van T zijn bij dezelfde eigenwaarde λ , en de eigenvectoren $w_i \in C_i$ vormen een lineair onafhankelijk stelsel, dan vormt $C = C_1 \cup C_2 \cup \dots \cup C_k$ een lineair onafhankelijk stelsel vectoren.

Bewijs. Het bewijs gaat weer met inductie naar het aantal elementen n van C . Voor de inductiestap kijken we naar de deelruimte W opgespannen door n de elementen van C_i ; die ruimte is invariant onder $\lambda I - T$, en het beeld wordt opgespannen door de $n - k$ vectoren uit $C \setminus \{w_1, \dots, w_k\}$ (de laatste vector uit elke cykel weggelaten). Onder de inductiehypothese heeft dat beeld dan dimensie precies $n - k$. De kern van $\lambda I - T$ bevat de k onafhankelijke eigenvectoren w_i , en de dimensiestelling zegt dan dat W dimensie n heeft.

Gevolg 7.26 Als het karakteristieke polynoom van de transformatie T op V volledig in lineaire factoren splitst, dan bestaat er een basis voor V ten opzichte waarvan de matrix van T Jordanmatrix is.

Bewijs. Kies voor elke λ een basis voor E_λ , en bij elke basisvector een cykel van gegeneraliseerde eigenvectoren.

Stelling 7.27 Laat T een transformatie van de L -vectorruimte V zijn waarvoor het karakteristieke polynoom $p_T(x)$ over L splitst: $p_T(x) = (x - \lambda_1)^{m_1} \dots (x - \lambda_k)^{m_k}$ (met $\lambda_i \neq \lambda_j$ als $i \neq j$). Dan is voor $i = 1, \dots, k$:

$$G_{\lambda_i} = \text{Ker}((\lambda_i I - T)^{m_i}), \quad \dim G_{\lambda_i} = m_i.$$

Bewijs. Kies een basis $\mathcal{B}$ voor V ten opzichte waarvan T een Jordanmatrix heeft. Op de diagonaal komen precies m_i eigenwaarden λ_i voor, dus $\dim G_{\lambda_i} \geq m_i$. Anderzijds vormt een vereniging van bases voor G_{λ_i} een basis voor V , dus

$$n = \sum_{i=1}^k \dim G_{\lambda_i} \geq \sum_{i=1}^k m_i = n.$$

Omdat per definitie

$$G_{\lambda_i} \supset \text{Ker}((\lambda_i I - T)^{m_i}),$$

resteert slechts de bewering

$$G_{\lambda_i} \subset \text{Ker}((\lambda_i I - T)^{m_i}),$$

te bewijzen. Omdat $\dim G_{\lambda_i} = m_i$ kan geen cykel van gegeneraliseerde eigenvectoren in G_{λ_i} langer zijn dan m_i . Dus elke vector in G_{λ_i} is bevat in $\text{Ker}((\lambda_i I - T)^{m_i})$.

Gevolg 7.28 Onder dezelfde aannamen als in de voorgaande stelling geldt:

$$T \text{ is diagonaliseerbaar} \iff \forall \lambda : G_\lambda = E_\lambda.$$