Technische Hogeschool Eindhoven

Onderafdeling der Wiskunde

Matrixtheorie

door prof. dr. G.W. Veltkamp

Dictaatnr. 2.287 Prijs f 12,-

TECHNISCHE HOGESCHOOL EINDHOVEN
Afdeling Algemene Wetenschappen
Onderafdeling der Wiskunde

MATRIXTHEORIE

door

Prof. Dr. G.W. Veltkamp

Najaarssemester 1979

TECHNISCHE HOGESCHOOL EINDHOVEN

Onderafdeling der Wiskunde

MATRIXTHEORIE

door

Prof. dr. G.W. Veltkamp

Najaarssemester 1979

OKT. 80
VB

Inhoudsbeschrijving

MATRIXTHEORIE

Najaarssemester 1979

JdG, 29 Juli 2005

I N H O U D

Hoofdstuk I. Inleiding matrixalgebra

1. Definities en rekenregels

1.1. Definitie van een matrix

Definitie. (1.1.1)

Zij $\mathcal{L}$ een verzameling en $m,n \in \mathbb{N}$. Een $m \times n$ matrix A over $\mathcal{L}$ is een afbeelding van $\{1,\ldots,m\} \times \{1,\ldots,n\}$ in $\mathcal{L}$:

$$(i,j) \mapsto \alpha_{ij} \in \mathcal{L}\ , \qquad 1 \le i \le m, \quad 1 \le j \le n\ .$$

Deze beeldverzameling wordt genoteerd als

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \alpha_{21} & \cdots & \alpha_{2n} \\ & & \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} .$$

De α_{ij} heten de elementen van de matrix.
We schrijven ook $A = (\alpha_{ij})$. En in plaats van α_{ij} schrijven we vaak $(A)_{ij}$.

De verzameling van alle $m \times n$ matrices over $\mathcal{L}$ noemen we $M_{m,n}(\mathcal{L})$ of, als duidelijk is wat $\mathcal{L}$ is, $M_{m,n}$. Een andere veel gebruikte notatie is $\mathcal{L}^{m\times n}$.

De nu volgende definities van de elementaire bewerkingen op matrices hebben zin indien $\mathcal{L}$ tenminste een commutatieve ring met eenheidselement is (bv. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{C}$, polynomen over een commutatieve ring met eenheidselement). Vanaf §2 veronderstellen we dat $\mathcal{L}$ een lichaam is (bv. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$). Dat is niet overal strikt nodig, maar het vereenvoudigt de zaak wel. Het belangrijke geval dat $\mathcal{L}$ een polynoomring is valt dus buiten onze beschouwingen, zie hiervoor bv. P. Lancaster, Lambda-matrices and vibrating systems. Oxford, 1966.

1.2. Elementaire bewerkingen op matrices

Definitie. (1.2.1)

De matrix $0 \in M_{m,n}$, met $(0)_{ij} = 0$ (het nulelement uit $\mathcal{L}$) voor alle i en j, heet nulmatrix uit $M_{m,n}$. De matrix $I \in M_{n,n}$, met $(I)_{ij} = 0$ als $i \neq j$ en $(I)_{ij} = 1$ (het eenheidselement uit $\mathcal{L}$) voor $i = j$, heet eenheidsmatrix uit $M_{n,n}$ (alleen als $m = n$!).

Definitie.

Als $A \in M_{m,n}$ en $B \in M_{m,n}$ dan is $A + B$ de matrix uit $M_{m,n}$ waarvoor geldt

$$(A + B)_{ij} := (A)_{ij} + (B)_{ij}\ , \quad \text{alle i en j .}$$

Eigenschappen. (1.2.3)

$$A + B = B + A\ ,$$
$$(A + B) + C = A + (B + C)\ ,$$
$$A + 0 = 0 + A = A\ ,$$
$$\forall_{A \in M_{m,n}}\ \exists_{(-A) \in M_{m,n}}\ :\ A + (-A) = 0\ .$$

Hieruit volgt: $M_{m,n}$ vormt een commutatieve groep t.o.v. de optelling met 0 als nulelement.

Definitie. (1.2.4)

Als $A \in M_{m,n}$ en $\alpha \in \mathcal{L}$ dan is αA de matrix uit $M_{m,n}$ waarvoor geldt

$$(\alpha A)_{ij} := \alpha (A)_{ij}\ , \quad \text{alle i en j .}$$

Eigenschappen. (1.2.5)

$$(\alpha\beta)A = \alpha(\beta A)\ ,$$
$$(\alpha + \beta)A = \alpha A + \beta A\ ,$$
$$\alpha(A + B) = \alpha A + \alpha B\ ,$$
$$1.A = A\ ,$$
$$(-1).A = -A\ ,$$
$$0.A = 0\ ,$$
$$\alpha.0 = 0\ .$$

Hieruit volgt: als $\mathcal{L}$ een lichaam is dan is $M_{m,n}(\mathcal{L})$ een vectorruimte over $\mathcal{L}$ t.o.v. optelling en vermenigvuldiging met scalairen.

Definitie. (1.2.6)

Als $A \in M_{m,k}$ en $B \in M_{k,n}$ dan is AB de matrix uit $M_{m,n}$ waarvoor geldt

$$(AB)_{ij} := \sum_{\ell=1}^{k} (A)_{i\ell}(B)_{\ell j} , \quad 1 \le i \le m, \quad 1 \le j \le n .$$

Eigenschappen. (1.2.7)

Als de afmetingen van de optredende matrices bij elkaar passen dan geldt

$$\begin{aligned}
&(AB)C = A(BC) , \\
&(A + B)C = AC + BC , \\
&A(B + C) = AB + AC , \\
&(\alpha A)B = A(\alpha B) = \alpha(AB) , \\
&A0 = 0 , \\
&0A = 0 , \\
&AI = A , \\
&IA = A .
\end{aligned}$$

Daar in het algemeen het product andere afmetingen heeft dan de factoren kunnen we in het algemeen niet zeggen dat $M_{m,n}$ een ring is. Wel geldt: $M_{n,n}$ is een niet-commutatieve ring met eenheidselement t.o.v. optelling en vermenigvuldiging; I fungeert als eenheidselement.

Opmerkingen. (1.2.8)

1) Als $A \in M_{m,n}$, $B \in M_{n,m}$ dan zijn zowel AB als BA gedefinieerd. Zij hebben echter afmetingen $m \times m$, resp. $n \times n$. Ook als $m = n$ is in het algemeen $AB \neq BA$.
   Voorbeeld:

$$\begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix}\begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix} = \begin{pmatrix} 0 & ca+db \\ 0 & 0 \end{pmatrix} .$$

2) Ook als $\mathcal{L}$ een lichaam is, is $M_{n,n}(\mathcal{L})$ $(n \ge 2)$ geen lichaam. Want het is mogelijk dat $A \neq 0$, $B \neq 0$ en $AB = 0$ (zie het voorbeeld bij 1)), zodat $M_{n,n}$ een ring met nuldelers en dus zeker geen lichaam is (in een lichaam

geldt: als $A \neq 0$ en $AB = 0$ dan is $B = 0$, want $A \neq 0 \Rightarrow A^{-1}$ bestaat en dus $B = A^{-1}.AB = A^{-1}.0 = 0$).

1.3. Transponeren van een matrix

Definitie. (1.3.1)

Aan $A \in M_{m,n}$ wordt toegevoegd de getransponeerde matrix $A^T \in M_{n,m}$ waarbij

$$(A^T)_{ij} := (A)_{ji}\ , \qquad 1 \leq i \leq n, \quad 1 \leq j \leq m\ .$$

Eigenschappen. (1.3.2)

$$(A^T)^T = A\ ,$$
$$(\alpha A)^T = \alpha A^T\ ,$$
$$(A + B)^T = A^T + B^T\ ,$$
$$(AB)^T = B^T A^T\ .$$

Definitie. (1.3.3)

Aan $A \in M_{m,n}(\mathbb{C})$ wordt toegevoegd de hermitisch geconjugeerde matrix $A^H \in M_{n,m}(\mathbb{C})$ waarbij

$$(A^H)_{ij} := \overline{(A)_{ji}}\ , \qquad 1 \leq i \leq n, \quad 1 \leq j \leq m\ .$$

Eigenschappen. (1.3.4)

$$(A^H)^H = A\ ,$$
$$(\alpha A)^H = \bar{\alpha} A^H\ ,$$
$$(A + B)^H = A^H + B^H\ ,$$
$$(AB)^H = B^H A^H\ .$$

Opmerking. (1.3.5)

In plaats van A^T en A^H ziet men veel andere notaties, b.v. A' i.p.v. A^T, A^* i.p.v. A^H.

1.4. Partitioneren van matrices

Een uiterst belangrijk hulpmiddel in de matrixtheorie is het zgn. partitioneren van matrices. Daarbij wordt een matrix beschouwd als te zijn opgebouwd uit deelmatrices volgens een rechthoekig schema.

Definitie. (1.4.1)

Zij voor $p \in \{1,\ldots,M\}$ en $q \in \{1,\ldots,N\}$

$$A_{pq} \in M_{m_p,n_q} ,$$

zij $m := \sum_1^M m_p$ en $n := \sum_1^N n_q$.

Indien de elementen van $A \in M_{m,n}$ gedefinieerd zijn door

$$(A)_{ij} = (A_{pq})_{k\ell}$$

als

$$i = \sum_{s=1}^{p-1} m_s + k , \quad 1 \le k \le m_p , \quad 1 \le p \le M ,$$

$$j = \sum_{s=1}^{q-1} n_s + \ell , \quad 1 \le \ell \le n_q , \quad 1 \le q \le N ,$$

dan heet

$$\left(\begin{array}{c|c|c} A_{11} & \cdots & A_{1N} \\ \hline \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot \\ \hline A_{M1} & \cdots & A_{MN} \end{array}\right)$$

een partitie van A in blokken A_{pq} en we schrijven

$$A = (A_{pq}) .$$

Eigenschappen. (1.4.2)

1) Als $A = (A_{pq})$ dan is $\alpha A = (\alpha A_{pq})$.
2) Als A en B dezelfde afmetingen hebben en "op dezelfde wijze" gepartitioneerd zijn dan volgt uit $A = (A_{pq})$, $B = (B_{pq})$ dat

$$A + B = (A_{pq} + B_{pq}) .$$

3) Zij $A = (A_{pq})$, $p = 1,\dots,M$, $q = 1,\dots,K$ met

$$A_{pq} \in M_{m_p,k_q} ,$$

en $B = (B_{pq})$, $p = 1,\dots,K$, $q = 1,\dots,N$ met

$$B_{pq} \in M_{k_p,n_q} .$$

Dan is $C := AB$ gedefinieerd en $C = (C_{pq})$, $p = 1,\dots,M$, $q = 1,\dots,N$ met

$$C_{pq} = \sum_{r=1}^{K} A_{pr}B_{rq} , \quad 1 \le p \le M , \quad 1 \le q \le N . \tag{1.4.3}$$

N.B. De verticale partitie van A moet passen bij de horizontale partitie van B; de horizontale indeling van C is dezelfde als die van A en de verticale indeling is dezelfde als die van B.

De juistheid van (1.4.3) is intuitief erg duidelijk, een formeel bewijs, gebaseerd op (1.4.1), is administratief een beetje vervelend.

4) Als $A = (A_{pq})$, $p = 1,\dots,M$, $q = 1,\dots,N$ dan is $A^T = (B_{pq})$, $p = 1,\dots,N$, $q = 1,\dots,M$ met

$$B_{pq} = A^T_{qp} .$$

Analoog met A^H.

<u>Bijzondere gevallen.</u> (1.4.4)

a) Alle blokken 1×1: $A_{pq} = (\alpha_{pq})$.

b) Alle blokken $m \times 1$:

$$A = (A_1 \mid \dots \mid A_n) . \tag{1.4.5}$$

$A_q \in M_{m,1}$ heet de q-de kolom van A.

c) Alle blokken $1 \times n$:

$$A = \begin{pmatrix} A_1 \\ \hline \cdot\cdot \\ \hline A_m \end{pmatrix} = (A_1^T \mid \dots \mid A_m^T)^T . \tag{1.4.6}$$

$A_p \in M_{1,n}$ heet de p-de rij van A.

d) Zij $A = (A_1 \mid \dots \mid A_N)$ met $A_q \in M_{m,n_q}$, $q = 1,\dots,N$, en

$B = (B_1^T \mid \dots \mid B_M^T)^T$ met $B_p \in M_{m_p,n}$, $p = 1,\dots,M$.

Dan is AB gedefinieerd indien $\sum_1^M m_p = \sum_1^N n_q$.

Als ook de blokafmetingen passen, d.w.z. als $M = N$ en $m_r = n_r$, $1 \le r \le M$, dan is

$$AB = \sum_{r=1}^{M} A_r B_r \in M_{m,n} . \tag{1.4.7}$$

Anderzijds is het product BA gedefinieerd indien $m = n$ en er geldt dan

$$BA = \left(\begin{array}{c|c|c} B_1A_1 & \cdots & B_1A_N \\ \hline \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \hline B_MA_1 & \cdots & B_MA_N \end{array}\right) = (\{BA\}_{pq}) \tag{1.4.8}$$

met $\{BA\}_{pq} = B_pA_q \in M_{m_p,n_q}$.

Als $X \in M_{n,m}$ dan is BXA gedefinieerd en er geldt

$$BXA = (\{BXA\}_{pq}) \tag{1.4.9}$$

met $\{BXA\}_{pq} = B_pXA_q \in M_{m_p,n_q}$.

1.5. Het spoor van een matrix

Definitie. (1.5.1)

Het spoor (engels: trace) van een (vierkante!) matrix $A \in M_{n,n}$ is

$$tr(A) := \sum_{i=1}^{n} (A)_{ii} .$$

Eigenschappen. (1.5.2)

1. $tr(\alpha A + \beta B) = \alpha tr(A) + \beta tr(B)$,
2. $tr(A^T) = tr(A)$; $tr(A^H) = \overline{tr(A)}$,
3. als $A \in M_{m,n}$, $B \in M_{n,m}$ dan is

$$\mathrm{tr}(AB) = \mathrm{tr}(BA) = \sum_{i=1}^{m} \sum_{j=1}^{n} (A)_{ij}(B)_{ji} .$$

4. als $A = (A_1 \mid \dots \mid A_n)$, $B = (B_1 \mid \dots \mid B_n)$ met $A_j \in M_{m,1}$, $B_j \in M_{m,1}$ dan is

$$\mathrm{tr}(A^T B) = \sum_{j=1}^{n} (A_j^T B_j)_{11}$$

(merk op dat $A_j^T B_j$ een 1×1-matrix is; we laten in dit geval de indices $(\ )_{11}$ meestal weg).

1.6. Permutatiematrices

Definitie. (1.6.1)

Een permutatie van de orde m is een afbeelding $p : i \mapsto p(i)$ van $\mathbb{N}_m := \{1,\dots,m\}$ op zichzelf.

Stelling. (1.6.2)

Bij iedere permutatie p van de orde m bestaat een inverse permutatie p^{-1} van de orde m, zodanig dat

$$p^{-1}(p(i)) = i , \quad i \in \mathbb{N}_m ,$$

$$p(p^{-1}(j)) = j , \quad j \in \mathbb{N}_m .$$

Bewijs.

Daar $\mathbb{N}_m$ eindig is en p $\mathbb{N}_m$ afbeeldt op zichzelf treedt ieder element van $\mathbb{N}_m$ precies eenmaal als beeld op. □

Definitie.

Aan een permutatie p van de orde m wordt toegevoegd een permutatiematrix $P \in M_{m,m}$ waarvoor geldt

$$(P)_{ij} = \delta_{i,p(j)} = \delta_{p^{-1}(i),j} = \begin{cases} 1 & \text{als } i = p(j) \\ 0 & \text{als } i \neq p(j) \end{cases} .$$

<u>Gevolg</u> (ga na). (1.6.4)

De i-de rij van P is de $p^{-1}(i)$-de rij van I,
de j-de kolom van P is de $p(j)$-de kolom van I,
de i-de rij van P^T is de $p(i)$-de rij van I,
de j-de kolom van P^T is de $p^{-1}(j)$-de kolom van I.

Voorbeeld:

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}, \quad p^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix},$$

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad P^T = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

<u>Stelling</u>. (1.6.5)

Als de permutatiematrix P hoort bij de permutatie p dan is P^T de permutatiematrix die hoort bij p^{-1}. Er geldt $PP^T = P^TP = I$.

<u>Bewijs</u>.

De eerste bewering volgt direct uit (1.6.4). Ook volgt uit (1.4.8) en (1.6.4)

$$\begin{aligned}(P^TP)_{ij} &= (\text{i-de rij van } P^T) \times (\text{j-de kolom van } P) \\ &= (p(i)\text{-de rij van } I) \times (p(j)\text{-de kolom van } I) \\ &= \delta_{p(i),p(j)} = \delta_{ij}\ .\end{aligned}$$

Analoog voor PP^T. □

<u>Opmerking</u>. (1.6.6)

Na invoering van het begrip inverse matrix kunnen we (1.6.5) uitspreken als: de bij p horende permutatiematrix P is inverteerbaar met als inverse de bij p^{-1} horende permutatiematrix P^T.

<u>Stelling</u>. (1.6.7)

Zij $P \in M_{m,m}$ en $Q \in M_{n,n}$ permutatiematrices behorende bij permutaties p, resp. q. Zij $A \in M_{m,n}$. Dan geldt:

de i-de rij van $P^T A$ is de p(i)-de rij van A,
de j-de kolom van AQ is de q(j)-de kolom van A,
het i,j-de element van $P^T AQ$ is het p(i),q(j)-de element van A:

$$(P^T AQ)_{ij} = (A)_{p(i),q(j)} \ . \tag{1.6.8}$$

Bewijs.

Zelf, gebruik (1.4.9).

Opgave. (1.6.9)

Als P_1 en $P_2 \in M_{m,m}$ permutatiematrices zijn, behorende bij permutaties p_1 en p_2, dan is $P_2 P_1$ de permutatiematrix horend bij de permutatie $p_2 \circ p_1$, gedefinieerd als $(p_2 \circ p_1)(i) = p_2(p_1(i))$. Leid hieruit (1.6.5) af.

2. Regulariteit en inverteerbaarheid

2.1. Definities en eenvoudige stellingen

Definitie. (2.1.1)

$A \in M_{m,n}$ heet rechtsregulier (notatie: $A \in RR$) als

$$\forall_{k\in\mathbb{N}}\ \forall_{X\in M_{n,k}} : AX = 0 \Rightarrow X = 0 .$$

A heet linksregulier ($A \in LR$) als

$$\forall_{k\in\mathbb{N}}\ \forall_{Y\in M_{k,m}} : YA = 0 \Rightarrow Y = 0 .$$

Definitie. (2.1.2)

$A \in M_{m,n}$ heet rechtsinverteerbaar ($A \in RI$) als er een $X \in M_{n,m}$ bestaat zo dat

$$AX = I \quad (\text{met } I \in M_{m,m}) .$$

X heet dan rechterinverse van A.

A heet linksinverteerbaar ($A \in LI$) als er een $Y \in M_{n,m}$ bestaat zo dat

$$YA = I \quad (\text{met } I \in M_{n,n}) .$$

Y heet dan linkerinverse van A.

Definitie. (2.1.3)

A heet regulier als $A \in RR$ en $A \in LR$.

A heet inverteerbaar ($A \in I$) als $A \in RI$ en $A \in LI$.

Stelling. (2.1.4)

$$A \in RR \iff \forall_{X\in M_{n,1}} : AX = 0 \Rightarrow X = 0 ,$$

$$A \in LR \iff \forall_{Y\in M_{1,m}} : YA = 0 \Rightarrow Y = 0 .$$

Bewijs.

$\Rightarrow$: triviaal uit definitie van RR.

$\Leftarrow$: Stel $X \in M_{n,k}$ $AX = 0$. Partitioneer X in kolommen

$$X = (X_1 \mid \dots \mid X_k) .$$

Dan is $AX_j = 0$, $j = 1,\dots,k$. Uit het rechterlid van (2.1.3) volgt $X_j = 0$, $j = 1,\dots,k$, dus $X = 0$. □

Stelling. (2.1.5)

$A \in LI \Rightarrow A \in RR$,

$A \in RI \Rightarrow A \in LR$,

$A \in I \Rightarrow A \in RR$ en $A \in LR$.

Bewijs.

Zij $YA = I$. Dan volgt uit $AX = 0$ dat $0 = YAX = X$, dus $A \in RR$. □

Stelling. (2.1.6)

$A \in RR \Rightarrow AX = Z$ heeft hoogstens één oplossing X,

$A \in LR \Rightarrow YA = Z$ heeft hoogstens één oplossing Y,

$A \in RI \Rightarrow AX = Z$ heeft minstens één oplossing X,

$A \in LI \Rightarrow YA = Z$ heeft minstens één oplossing Y.

Bewijs.

Als $AX_1 = Z$ en $AX_2 = Z$, dan is $A(X_1 - X_2) = 0$, dus volgens (2.1.1) $X_1 - X_2 = 0$. Als $AX_0 = I$ dan voldoet $X = X_0Z$ aan $AX = Z$. □

Stelling. (2.1.7)

$A \in RI$ en $A \in RR \Rightarrow A \in I$,

$A \in LI$ en $A \in LR \Rightarrow A \in I$.

Bewijs.

Zij X rechterinverse van A. Dan is $AX = I$, $AXA = A$, dus $A(XA - I) = 0$. Uit $A \in RR$ volgt nu $XA = I$, dus $A \in LI$, dus $A \in I$. □

Stelling. (2.1.8)

Als A inverteerbaar is, dan heeft A precies één linker- en één rechterinverse; deze zijn gelijk.

Bewijs.

Als $A \in LI$ dan heeft A volgens (2.1.5) en (2.1.6) hoogstens één rechterinverse. Analoog als $A \in RI$. Als $A \in LI$ en $\in RI$, dan heeft A dus precies één

rechterinverse X en één linkerinverse Y.
En er geldt $X = YA.X = Y.AX = Y$. □

Notatie. (2.1.9)

Als A inverteerbaar is dan wordt de matrix die de unieke linkerinverse en rechterinverse van A is aangeduid met A^{-1}: *de inverse van A*.

Stelling. (2.1.10)

$$A \in I \Rightarrow A^{-1} \in I \text{ en } (A^{-1})^{-1} = A \,.$$

Stelling. (2.1.11)

Zij $A \in M_{m,n}$, $B \in M_{m,m}$, $C \in M_{n,n}$, B en $C \in I$ en

$$A' := BAC \,.$$

Dan geldt $A \in RR \leftrightarrow A' \in RR$ en analoog met LR, RI, LI en I.
Als X rechterinverse van A is, dan is $C^{-1}XB^{-1}$ rechterinverse van A'. Analoog met linkerinverse en inverse.

Stelling. (2.1.12)

Als A_1 en A_2 zo zijn dat A_1A_2 bestaat dan geldt:

- A_1 en $A_2 \in RR \Rightarrow A_1A_2 \in RR$,
- $A_1A_2 \in RR \Rightarrow A_2 \in RR$,
- A_1 en $A_2 \in RI \Rightarrow A_1A_2 \in RI$,
- $A_1A_2 \in RI \Rightarrow A_1 \in RI$.

Analoog met LR, LI en I.

Stelling. (2.1.13)

Zij *)

$$A = \left(\begin{array}{c|c} A_{11} & \\ \hline & A_{22} \end{array}\right) \,.$$

Dan geldt $A \in RR \Leftrightarrow A_{11}$ en $A_{22} \in RR$.
Analoog met LR, RI, LI en I.

---

*) In gepartitioneerde matrices schrijft men op een plaats waar een blok nullen staat vaak niets (in plaats van 0_{12}, 0_{21}, etc.).

Stelling. (2.1.14)

Zij $A = (A_1 \mid A_2)$.
Dan geldt

$$A \in RR \Rightarrow A_1 \text{ en } A_2 \in RR\ ,$$
$$A \in LI \Rightarrow A_1 \text{ en } A_2 \in LI\ ,$$
$$A_1 \in LR \text{ of } A_2 \in LR \Rightarrow A \in LR\ ,$$
$$A_1 \in RI \text{ of } A_2 \in RI \Rightarrow A \in RI\ .$$

Bewijs.

Zij $A_1 \in M_{m,n_1}$, $A_2 \in M_{m,n_2}$, $A \in M_{m,n}$, $n = n_1 + n_2$.

a) Zij $A \in RR$. Stel $A_1X_1 = 0$ met $X_1 \in M_{n_1,k}$. Vul X_1 aan met n_2 rijen nullen tot $X = \left(\frac{X_1}{0}\right) \in M_{n,k}$. Dan is $AX = 0$, dus $X = 0$ (want $A \in RR$), dus $X_1 = 0$. Dus $A_1 \in RR$.

b) Stel A heeft een linkerinverse $Y \in M_{n,m}$.
Partitioneer Y: $Y = \left(\frac{Y_1}{Y_2}\right)$ met $Y_1 \in M_{n_1,m}$, $Y_2 \in M_{n_2,m}$. Uit $YA = I$ $(n \times n)$ volgt dan $Y_1A_1 = I_1$ $(n_1 \times n_1)$, $Y_1A_2 = 0$ $(n_1 \times n_2)$, $Y_2A_1 = 0$ $(n_2 \times n_1)$, $Y_2A_2 = I_2$ $(n_2 \times n_2)$. Dit impliceert dat A_1 en A_2 linkerinversen hebben.

c) Zij $A_1 \in LR$. Uit $YA = 0$ volgt $YA_1 = 0$, dus $Y = 0$, zodat $A \in LR$.

d) Stel $A_1 \in RI$ met rechterinverse $X_1 \in M_{n_1,m}$. Dan is $A_1X_1 = I$ $(m \times m)$. Vul X_1 aan met n_2 rijen nul tot $X = \left(\frac{X_1}{0}\right) \in M_{n,m}$. Dan geldt $AX = A_1X_1 = I$, dus $A \in RI$.

Stelling. (2.1.15)

Zij

$$A = \left(\begin{array}{c|c} A_{11} & \\ \hline A_{21} & A_{22} \end{array}\right) .$$

Dan geldt

- $A_{11} \in RR$ en $A_{22} \in RR \Rightarrow A \in RR \Rightarrow A_{22} \in RR$,
- $A_{11} \in LR$ en $A_{22} \in LR \Rightarrow A \in LR \Rightarrow A_{11} \in LR$,
- $A_{11} \in RI$ en $A_{22} \in RI \Rightarrow A \in RI \Rightarrow A_{11} \in RI$,
- $A_{11} \in LI$ en $A_{22} \in LI \Rightarrow A \in LI \Rightarrow A_{22} \in LI$.

<u>Bewijs</u>.

a) Stel $A_{11} \in RR$ en $A_{22} \in RR$. Stel $AX = 0$. Partitioneer X: $X = \left(\frac{X_1}{X_2}\right)$. Dan is

$$A_{11}X_1 = 0\ , \qquad (*)$$

$$A_{21}X_1 + A_{22}X_2 = 0\ . \qquad (**)$$

Daar $A_{11} \in RR$ volgt uit (*) $X_1 = 0$. En dan volgt uit (**), daar $A_{22} \in RR$, dat $X_2 = 0$. Dus $X = 0$, dus $A \in RR$.
Als, omgekeerd, $A \in RR$ en $A_{22}X_2 = 0$ dan is ook $AX = 0$ met $X = (0 \mid X_2^T)^T$.
Uit $A \in RR$ volgt $X = 0$, dus $X_2 = 0$. Dus $A_{22} \in RR$.
Analoog voor LR.

b) Zij A_{11} en $A_{22} \in RI$ met rechterinversen X_{11} en X_{22}.
Dan is

$$X = \left(\begin{array}{c|c} X_{11} & \\ \hline -X_{22}A_{21}X_{11} & X_{22} \end{array}\right)$$

rechterinverse van A, want $AX = I$ (ga na).
Als, omgekeerd, $A \in RI$ met rechterinverse X, dan is (met voor de hand liggende partitie van X) $A_{11}X_{11} = I_{11}$, dus $A_{11} \in RI$.
Analoog voor LI. □

<u>Gevolg</u>. (2.1.16)

Als

$$A = \left(\begin{array}{c|c} A_{11} & \\ \hline A_{21} & A_{22} \end{array}\right) .$$

met A_{11} en $A_{22} \in I$ dan is $A \in I$ en

$$A^{-1} = \left(\begin{array}{c|c} A_{11}^{-1} & \\ \hline -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} \end{array}\right) .$$

Opmerking. (2.1.17)

Dat uit $A \in RR$ niet volgt dat $A_{11} \in RR$ blijkt uit het voorbeeld

$$A = \left(\begin{array}{cc|c} 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) .$$

In (2.4.5) zal blijken dat de volledige equivalentie wel geldt als A_{11} vierkant is.

Stelling. (2.1.18)

Als $A \in RR$ dan is $A^T \in LR$. Analoog voor LR, RI, LI.

Opgave. (2.1.19)

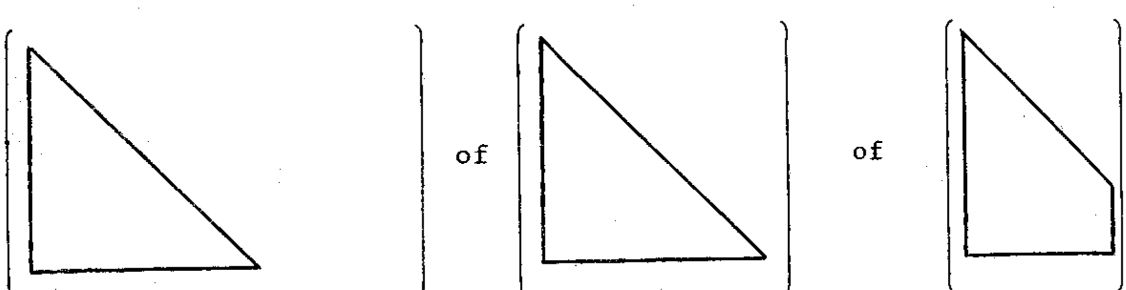
Gebruik (2.1.18) om de "tweede helft" van stellingen als (2.1.7) en (2.1.12) te bewijzen en om "spiegelbeelden" aan (2.1.14) en (2.1.15) toe te voegen.

2.2. Driehoeksmatrices

Definitie. (2.2.1)

$A \in M_{m,n}$ heet linkerdriehoeksmatrix als $(A)_{ij} = 0$ voor $j > i$. Analoog rechterdriehoeksmatrix. We zeggen dat A reguliere diagonaal heeft indien $(A)_{ii} \neq 0$ voor $1 \leq i \leq \min(m,n)$.

We eisen niet dat $n = m$. Een linkerdriehoeksmatrix kan er dus uitzien als



$n > m$ $\qquad$ $n = m$ $\qquad$ $n < m$

Definitie. (2.2.2)

Een linkerblokdriehoeksmatrix is een gepartitioneerde matrix waarvan de blokken A_{ij} met $j > i$ alle 0 zijn.

Stelling. (2.2.3)

Als $A \in M_{m,k}$ en $B \in M_{k,n}$ linkerdriehoeksmatrices zijn, dan is AB een linkerdriehoekmatrix.
Analoog voor rechterdriehoeksmatrices.

Bewijs.

Daar

$$(AB)_{ij} = \sum_{\ell=1}^{k} (A)_{i\ell}(B)_{\ell j}$$

en $(A)_{i\ell} = 0$ voor $\ell > i$, $(B)_{\ell j} = 0$ voor $\ell < j$ is $(AB)_{ij} = 0$ voor $i < j$. □

Opmerking. (2.2.4)

Een analoge stelling geldt voor blokdriehoeksmatrices (ga na).

Stelling. (2.2.5)

Zij $A \in M_{m,n}$ een linkerdriehoeksmatrix. Dan geldt

. $A \in I \Leftrightarrow m = n$ en A heeft reguliere diagonaal ,
. als $A \in I$ dan is A^{-1} linkerdriehoeksmatrix .

Analoog voor rechterdriehoeksmatrices.

Bewijs.

We passen inductie naar $k := \min(m,n)$ toe.
Als $k = m = 1$ en $n > 1$ dan geldt (ga na) $A \notin RR$, als $k = n = 1$ en $m > 1$ dan geldt $A \notin LR$. Hieruit volgt (ga na) dat de bewering waar is voor $k = 1$.

Zij nu $k > 1$ en zij

$$A = \left(\begin{array}{c|c} A_{11} & \\ \hline A_{21} & A_{22} \end{array}\right)$$

met $A_{11} \in M_{1,1}$.

Stel dat $A \in I$ met inverse

$$X = \left(\begin{array}{c|c} X_{11} & X_{12} \\ \hline X_{21} & X_{22} \end{array}\right)$$

(met $X_{11} \in M_{1,1}$). Dan is $A_{11}X_{11} = I_{11}$ dus $A_{11} \neq 0$ en $A_{11} \in I$. Uit $A_{11}X_{12} = 0$ volgt $X_{12} = 0$ en daarna blijkt dat $A_{22}X_{22} = X_{22}A_{22} = I_{22}$, dus $A_{22} \in I$. Volgens de inductieveronderstelling is dan A_{22} vierkant met **reguliere diagonaal.** Derhalve is ook A vierkant met reguliere diagonaal.
Stel nu dat A vierkant is met **reguliere** diagonaal. Dan is $A_{11} \neq 0$, dus $A_{11} \in I$ en A_{22} is vierkant met **reguliere** diagonaal, dus (inductieveronderstelling) $A_{22} \in I$ en A_{22}^{-1} is linkerdriehoeksmatrix. Uit (2.1.16) volgt nu dat $A \in I$ en dat A^{-1} linkerdriehoeksmatrix is. □

Gevolg. (2.2.6)

Zij $A \in M_{m,n}$ een linker- of rechterdriehoeksmatrix met reguliere diagonaal. Dan geldt

$$m \geq n \Leftrightarrow A \in LI \Leftrightarrow A \in RR\ ,$$
$$m \leq n \Leftrightarrow A \in RI \Leftrightarrow A \in LR\ .$$

Bewijs dit zelf met behulp van het feit dat, als $k := \min(m,n)$, het linksboven $k{\times}k$-blok van A inverteerbaar is.

Opmerking. (2.2.7)

Voor blokdriehoeksmatrices geldt een stelling analoog aan (2.2.5). Een bewijs kan pas gegeven worden nadat (2.4.2) of iets equivalents bewezen is.

2.3. Een fundamentele decompositie

De volgende decompositie is fundamenteel voor een aantal bewijzen van stellingen over regulariteit, inverteerbaarheid, rang, enz.

Stelling. (2.3.1)

Zij $A \in M_{m,n}$,

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

met $A_{11} \in M_{k,k}$ $(1 \leq k < \min(m,n))$ en A_{11} inverteerbaar *).
Zij

$$A_{22.1} := A_{22} - A_{21}A_{11}^{-1}A_{12} \tag{2.3.2}$$

(ga na dat de afmetingen passen!).
Dan geldt

$$A = \left(\begin{array}{c|c} I_{11} & \\ \hline A_{21}A_{11}^{-1} & I_{22} \end{array}\right)\left(\begin{array}{c|c} A_{11} & \\ \hline & A_{22.1} \end{array}\right)\left(\begin{array}{c|c} I_{11} & A_{11}^{-1}A_{12} \\ \hline & I'_{22} \end{array}\right), \tag{2.3.3}$$

met $I_{11} \in M_{k,k}$, $I_{22} \in M_{m-k,m-k}$, $I'_{22} \in M_{n-k,n-k}$. De eerste en de derde factor van het rechterlid zijn inverteerbaar.

Bewijs.

De geldigheid van (2.3.3) volgt direct door uitwerken van de producten.
De inverteerbaarheid van de eerste en derde factor volgt uit (2.1.16).

Opmerkingen. (2.3.4)

1. de door (2.3.2) gedefinieerde matrix $A_{22.1}$ wordt Schur-complement van A_{11} in A genoemd.
2. Het nut van (2.3.3) zien we als volgt in. Daar de eerste en derde factor inverteerbaar zijn volgt uit (2.1.11) dat

$$A \in RR \leftrightarrow \left(\begin{array}{c|c} A_{11} & \\ \hline & A_{22.1} \end{array}\right) \in RR$$

   en daar A_{11} inverteerbaar is, is dit volgens (2.1.13) weer equivalent met $A_{22.1} \in RR$. Analoog voor LR, RI, LI. $A_{22.1}$ is echter kleiner dan A. Hier ligt dus een mogelijkheid voor bewijzen door middel van volledige inductie.
3. Natuurlijk geldt niet voor iedere A dat er een k is zodanig dat A_{11} inverteerbaar is. Maar wel geldt: als $A \neq 0$ dan zijn niet alle elementen van A nul en dus zijn er permutatiematrices P en Q zodanig dat

---

*) We zullen straks zien dat alleen vierkante matrices inverteerbaar zijn, daarom veronderstellen we A_{11} vierkant.

$$P^T A Q = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

met $A_{11} \in M_{1,1}$, $A_{11} \neq 0$. Daar P en Q inverteerbaar zijn, geldt dus bv.

$$A = P \left(\begin{array}{c|c} I_{11} & \\ \hline A_{21}A_{11}^{-1} & I_{22} \end{array}\right) \left(\begin{array}{c|c} A_{11} & \\ \hline & A_{22.1} \end{array}\right) \left(\begin{array}{c|c} I_{11} & A_{11}^{-1}A_{12} \\ \hline & I'_{22} \end{array}\right) Q^T .$$

$$= B \left(\begin{array}{c|c} A_{11} & \\ \hline & A_{22.1} \end{array}\right) C$$

met inverteerbare B en C. Pas nu (2.1.11) weer toe, etc.

Opgave. (2.3.5)

Stel dat

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

met A_{11} inverteerbaar.

Dan geldt: A inverteerbaar $\Leftrightarrow$ $A_{22.1}$ inverteerbaar. En als $X := A^{-1}$, $X_{22} := A_{22.1}^{-1}$, dan is

$$X = \left(\begin{array}{c|c} A_{11}^{-1} + A_{11}^{-1}A_{12}X_{22}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}X_{22} \\ \hline -X_{22}A_{21}A_{11}^{-1} & X_{22} \end{array}\right) .$$

Bewijs dit uit (2.3.1) met behulp van (2.1.11), (2.1.13) en (2.1.16). Definieer op voor de hand liggende wijze $X_{11.2}$ en laat zien dat $X_{11.2} = A_{11}^{-1}$.

Opgave. (2.3.6)

Zij A als in (2.3.5) met $A_{22} \in I$. Definieer, analoog aan (2.3.2), een matrix $A_{11.2}$ en bepaal een decompositie van A, analoog aan (2.3.3). Formuleer en bewijs nu een analogon van (2.3.5).

2.4. Verband tussen regulariteit, inverteerbaarheid en afmetingen

Stelling. (2.4.1)

Zij $A \in M_{m,n}$. Dan geldt

$$A \in RR \leftrightarrow A \in LI \Rightarrow m \geq n ,$$
$$A \in LR \leftrightarrow A \in RI \Rightarrow m \leq n ,$$
$$A \in LR \text{ en } A \in RR \leftrightarrow A \in I \Rightarrow m = n .$$

Bewijs.

Zij $A \in RR$. We moeten bewijzen dat dan $A \in LI$ en $m \geq n$ (uit (2.1.5) volgt dat $A \in LI \Rightarrow A \in RR$).
Pas volledige inductie naar n toe. Zij $n = 1$. Als $A = 0$ dan is $A \notin RR$. Dus $A \neq 0$. Als $m = 1$ dan bestaat zelfs A^{-1}. Als $m > 1$ dan is er een permutatiematrix P zodat

$$P^T A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$$

met $A_1 \in M_{1,1}$ en $A_1 \neq 0$. Dus A_1^{-1} bestaat en bv. $(A_1^{-1} \mid 0)P^T$ is linkerinverse van A.
Zij $n > 1$. Als $m = 1$ dan is $A \notin RR$ (ga na). Dus $m > 1$. Als de hele eerste kolom van A nul is dan $A \notin RR$. Er is dus een permutatiematrix P zodanig dat

$$P^T A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

met $A_{11} \in M_{1,1}$, $A_{11} \neq 0$. Dus bestaat A_{11}^{-1} en uit (2.3.1) volgt: er zijn inverteerbare matrices B en C zodanig dat

$$A = B \left(\begin{array}{c|c} A_{11} & \\ \hline & A_{22.1} \end{array}\right) C .$$

Uit (2.1.10) en (2.1.12) volgt dat $A_{22.1} \in RR$. Uit de inductieveronderstelling volgt nu dat $A_{22.1} \in LI$ en $m-1 \geq n-1$, dus $m \geq n$.
Uit (2.1.10) en (2.1.12) volgt dan dat $A \in LI$.
Analoog als $A \in LR$. $\square$

<u>Stelling.</u> (2.4.2)

Zij $A \in M_{n,n}$ (dus vierkant).
Dan geldt

$$A \in RR \Leftrightarrow A \in LR \Leftrightarrow A \in I .$$

<u>Bewijs.</u>

Inductie naar n. Triviaal als $n = 1$. Voor $n > 1$ net als in het bewijs van (2.4.1). Zelf! □

<u>Opmerkingen.</u> (2.4.3)

1. Linksregulier + rechtsregulier is dus identiek met inverteerbaar. Daarom wordt een inverteerbare (vierkante) matrix voortaan ook wel regulier genoemd.

2. Uit (2.4.1) en (2.4.2) volgt dat er voor $A \in M_{m,n}$ slechts de volgende mogelijkheden zijn:
   . als $m = n$ dan $A \in I$ of A irregulier
   . als $m > n$ dan ($A \in LI$, $A \notin LR$) of A irregulier
   . als $m < n$ dan ($A \in RI$, $A \notin RR$) of A irregulier
   (hierbij betekent irregulier: $A \notin RR \cup LR$).

3. Van de drie eigenschappen:

   $$A \in RR,\ A \in LR,\ A \text{ vierkant}$$

   impliceert ieder tweetal de derde en dus ook dat $A \in I$.

<u>Opgave.</u> (2.4.4)

Zij $A = A_1A_2$.
Als A_2 vierkant is dan geldt $A \in RR \Leftrightarrow A_1 \in RR$ en $A_2 \in I$.
Als A_1 vierkant is dan geldt $A \in LR \Leftrightarrow A_1 \in I$ en $A_2 \in LR$.
Analoog met LI en RI.
Dit is een aanvulling op (2.1.12).

<u>Opgave.</u> (2.4.5)

Zij

$$A = \left(\begin{array}{c|c} A_{11} & \\ \hline A_{21} & A_{22} \end{array}\right) .$$

Dan geldt, in aanvulling op (2.1.15),

- als A_{11} vierkant is, dan $A \in RI \leftrightarrow A_{11} \in I$ en $A_{22} \in RI$.
- als A_{22} vierkant is, dan $A \in LI \leftrightarrow A_{11} \in LI$ en $A_{22} \in I$.
- als A_{11} of A_{22} vierkant is dan $A \in I \leftrightarrow A_{11}$ en $A_{22} \in I$.

<u>Opgave</u>. (2.4.6)

Formuleer en bewijs een stelling voor blokdriehoeksmatrices analoog aan (2.2.5).

3. Decompositiestellingen. Equivalentie. Rang.

We willen een matrix $A \in M_{m,n}$ schrijven als product van **eenvoudiger** matrices. We zullen (voorlopig) bewijzen dat dat op de volgende manieren kan.

a) $A = BJC$ met B en C inverteerbaar en

$$J = \left(\begin{array}{c|c} I_{11} & 0 \\ \hline 0 & 0 \end{array}\right)$$

met I_{11} de $r \times r$ eenheidsmatrix, $0 \le r \le \min(m,n)$ (het is toegestaan dat $r = 0$ of $r = \min(m,n)$, J heeft dan slechts één blok-rij en/of -kolom).

b) $A = B_1C_1$ met B_1 $m \times r$ en rechtsregulier, C_1 $r \times n$ en linksregulier, $1 \le r \le \min(m,n)$, (voor $A = 0$ is deze representatie niet mogelijk).

We zullen bewijzen dat het bij een matrix A horende getal r, en dus ook de matrix J eenduidig bepaald is (B en C, cq. B_1 en C_1 zijn dit geenszins). We noemen deze r de rang van A en J de met A equivalente normaalvorm. Regulariteits- en inverteerbaarheidseigenschappen van A zijn dezelfde als van J en uit J eenvoudig af te lezen.

3.1. De algemene decompsitiestelling

Definitie. (3.1.1)

$J \in M_{m,n}$ heet normaalvorm van rang r uit $M_{m,n}$ als

$$0 \le r \le \min(m,n) \ ,$$

$$(J)_{ij} = \begin{cases} 1 & \text{als } 1 \le i = j \le r \\ 0 & \text{anders} \ . \end{cases}$$

Opmerking. (3.1.2)

- Als $r = m = n$ dan is $J = I$, dus J inverteerbaar;
- als $r = m < n$ dan is $J = (I_1 \mid 0)$, dus $J \in RI$ (met rechterinverse J^T), $J \in LR$, $J \notin RR$;
- als $r = n < m$ dan is $J = \left(\frac{I_1}{0}\right)$, dus $J \in LI$ (met linkerinverse J^T), $J \in RR$, $J \notin LR$;
- als $1 \le r < \min(m,n)$ dan is

$$J = \left(\begin{array}{c|c} I_{11} & 0 \\ \hline 0 & 0 \end{array}\right) \quad \text{met } I_{11} \in M_{r,r} ,$$

dus $J \notin LR$, $J \notin RR$;

. als $r = 0$ dan is $J = 0$.

Hiermee zijn alle mogelijkheden opgesomd.

Stelling **(algemene decompositiestelling)** (3.1.3)

Zij $A \in M_{m,n}$. Dan zijn er

. reguliere matrices $B \in M_{m,m}$ en $C \in M_{n,n}$,

. een normaalvorm $J \in M_{m,n}$

zodat

$$A = BJC . \tag{3.1.4}$$

Bewijs.

Als $A = 0$ dan triviaal met $J = 0$, $B = I$ ($m \times m$), $C = I'$ ($n \times n$).

Zij nu $A \neq 0$. We bewijzen (3.1.4) met inductie naar $k := \min(m,n)$.

Als $k = m = n = 1$ dan triviaal.

Zij $k = m = 1 < n$. Daar $A \neq 0$ is er een permutatiematrix Q zo dat $AQ = (A_{11} \mid A_{12})$ met $A_{11} \in M_{1,1}$ en $\neq 0$. Dan is

$$A = (I_{11} \mid 0) \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline & I_{22} \end{array}\right) Q^T = BJC$$

met $I_{11} \in M_{1,1}$, $I_{22} \in M_{n-1,n-1}$, $B = I_{11}$, $J = (I_{11} \mid 0)$. Uit (2.1.16) volgt dat C regulier is.

Analoog als $k = n = 1 < m$.

Zij nu $k > 1$. Daar $A \neq 0$ zijn er permutatiematrices P_1 en Q_1 zo dat

$$P_1^T A Q_1 = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

met $A_{11} \in M_{1,1}$ en $\neq 0$. Kies B_{11} en $C_{11} \in M_{1,1}$ zo dat

$$A_{11} = B_{11} C_{11} .$$

Analoog aan (2.3.3) geldt dan

$$P_1^T A Q_1 = \left(\begin{array}{c|c} B_{11} & \\ \hline B_{21} & I_{22} \end{array}\right) \left(\begin{array}{c|c} I_{11} & \\ \hline & A_{22.1} \end{array}\right) \left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline & I_{22} \end{array}\right) \qquad (*)$$

met

$$B_{21} := A_{21}C_{11}^{-1}, \ C_{12} := B_{11}^{-1}A_{12} \text{ en}$$

$$A_{22.1} := A_{22} - A_{21}A_{11}^{-1}A_{12} = A_{22} - B_{21}C_{21} \ .$$

$A_{22.1}$ heeft afmetingen $(m-1) \times (n-1)$, dus volgens de inductieveronderstelling zijn er reguliere $B_{22} \in M_{m-1,m-1}$ en $C_{22} \in M_{n-1,n-1}$ en een normaalvorm $J_{22} \in M_{m-1,n-1}$ zo dat

$$A_{22.1} = B_{22}J_{22}C_{22} \ .$$

Daarmee volgt uit (*)

$$A = P_1 \left(\begin{array}{c|c} B_{11} & \\ \hline B_{21} & B_{22} \end{array}\right) \left(\begin{array}{c|c} I_{11} & \\ \hline & J_{22} \end{array}\right) \left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline & C_{22} \end{array}\right) Q_1^T = BJC \ .$$

Uit (2.1.16) volgt dat B en C regulier zijn; J is duidelijk een normaalvorm uit $M_{m,n}$ en (3.1.4) geldt. □

<u>Corollarium</u> (driehoeksdecompositie). (3.1.5)

Als $A \in M_{m,n}$ dan zijn er

- permutatiematrices $P \in M_{m,m}$ en $Q \in M_{n,n}$
- een reguliere linksondermatrix $L \in M_{m,m}$
- een reguliere rechtsbovenmatrix $R \in M_{n,n}$
- een normaalvorm $J \in M_{m,n}$

zo dat

$$P^T AQ = LJR \ . \qquad (3.1.6)$$

<u>Bewijs</u>.

Dit is grotendeels hetzelfde als dat van (3.1.3). Alleen volgt nu uit de inductieveronderstelling het bestaan van P_{22}, Q_{22}, L_{22}, R_{22} en J_{22} zo dat

$$P_{22}^T A_{22.1} Q_{22} = L_{22} J_{22} R_{22}$$

en daarmee vinden we

$$\left(\begin{array}{c|c} I_{11} & \\ \hline & P_{22}^T \end{array}\right) P_1^T A Q_1 \left(\begin{array}{c|c} I_{11} & \\ \hline & Q_{22} \end{array}\right) = \left(\begin{array}{c|c} B_{11} & \\ \hline P_{22}^T B_{21} & L_{22} \end{array}\right) \left(\begin{array}{c|c} I_{11} & \\ \hline & J_{22} \end{array}\right) \left(\begin{array}{c|c} C_{11} & C_{12} Q_{22} \\ \hline & R_{22} \end{array}\right)$$

waaruit (3.1.6) direct volgt. □

Opgave. (3.1.7)

Bewijs (2.4.1) uit (3.1.3), (2.1.11) en (3.1.2).

3.2. Equivalentie. Rang

Definitie. (3.2.1)

Matrices A en A' heten equivalent als er inverteerbare B en C bestaan zo dat

$$A = BA'C \ .$$

Notatie: $A \sim A'$.

Opmerking.

Ga na dat dit een equivalentierelatie is:

$$A \sim A, \ A \sim A' \Rightarrow A' \sim A, \ A \sim A' \text{ en } A' \sim A'' \Rightarrow A \sim A'' \ .$$

De algemene decompositiestelling zegt dat iedere $A \in M_{m,n}$ equivalent is met een normaalvorm J uit $M_{m,n}$. We bewijzen nu dat A met precies één J equivalent is, m.a.w. de rang van de J uit de algemene decompositiestelling is eenduidig door A bepaald (B en C zijn dat geenszins, hangen af van de gekozen permutatiematrices).

Stelling. (3.2.2)

Bij iedere $A \in M_{m,n}$ is er precies één normaalvorm J uit $M_{m,n}$ waarmee A equivalent is.

Bewijs.

Dat er een J is volgt uit de algemene decompositiestelling.
Stel dat A equivalent is met twee verschillende normaalvormen J en J' uit

$M_{m,n}$ met rangen r en r'. Dan zijn ook J en J' equivalent, dus

$$J' = BJC$$

met reguliere B en C. Partitioneer (we veronderstellen dat $1 \leq r < \min(m,n)$, $1 \leq r' < \min(m,n)$, behandel de andere gevallen zelf):

$$\left(\begin{array}{c|c} I'_{11} & \\ \hline & \end{array}\right) = \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right)\left(\begin{array}{c|c} I_{11} & \\ \hline & \end{array}\right)\left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right),$$

met $I'_{11} \in M_{r',r'}$, $I_{11} \in M_{r,r}$, $B_{11} \in M_{r',r}$, $C_{11} \in M_{r,r'}$. Hieruit volgt

$$B_{11}C_{11} = I'_{11} ,$$

dus $B_{11} \in RI$, dus volgens (2.4.1) $r' \leq r$.
Analoog $r \leq r'$ (ga uit van $J = B'J'C'$).
Dus $r = r'$ en $J = J'$. □

<u>Definitie.</u> (3.2.3)

De rang van een matrix $A \in M_{m,n}$ is de rang van de normaalvorm uit J waarmee hij equivalent is.
<u>Notatie</u>: rang(A) of r(A).
Als $A \in M_{m,n}$ rang r heeft dan schrijven we $A \in M^r_{m,n}$.

<u>Stelling.</u> (3.2.4)

Matrices A en A' zijn equivalent dan en slechts dan als zij dezelfde afmetingen en rang hebben.

<u>Bewijs.</u>

A en A' kunnen slechts equivalent zijn als zij gelijke afmetingen hebben, want inverteerbare matrices zijn vierkant (2.4.1).
De rest volgt uit (3.2.1) en (3.2.2). □

<u>Stelling.</u> (3.2.5)

Voor $A \in M_{m,n}$ geldt

$$0 \leq r(A) \leq \min(m,n) ,$$
$$r(A) = 0 \Leftrightarrow A = 0 ,$$

$$r(A) = m\ (\le n) \Leftrightarrow A \in LR \Leftrightarrow A \in RI\ ,$$
$$r(A) = n\ (\le m) \Leftrightarrow A \in RR \Leftrightarrow A \in LI\ ,$$
$$r(A) = m = n \Leftrightarrow A \in LR \text{ en } A \in RR \Leftrightarrow A \in I\ .$$

Bewijs.

Zij A equivalent met een normaalvorm J. Volgens (2.1.10) geldt $A \in RR \Leftrightarrow J \in RR$ en analoog voor LR, RI en LI. De diverse beweringen volgen nu uit (3.1.2). □

Opmerking. (3.2.6)

Een groot deel van deze resultaten is ook al in §2.4 bewezen, alleen het verband met de rang is nieuw. Ga na dat het mogelijk is om §2.4 geheel achterwege te laten en alles te baseren op de algemene decompositiestelling. Wel moeten we dan het feit dat $A \in RI \Rightarrow m \le n$ apart uit (3.1.3) en (3.1.2) bewijzen omdat dit nodig is in het bewijs van (3.2.2).

Opgave. (3.2.7)

Zij $A \in M^r_{m,n}(\mathcal{L})$. Zij $\mathcal{L}$ deellichaam van een groter lichaam $\mathcal{L}'$ (bv. $\mathcal{L} = \mathbb{R}$, $\mathcal{L}' = \mathbb{C}$). Dan kunnen we A ook opvatten als element van $M_{m,n}(\mathcal{L}')$. Bewijs dat $A \in M^r_{m,n}(\mathcal{L}')$.

Opgave. (3.2.8)

Zij $A = BJC \in M^r_{m,n}$ met J de normaalvorm met rang r uit $M_{m,n}$, B en C regulier. Dan geldt $A = B'JC'$ met reguliere B' en C' dan en slechts dan als

$$B' = B\left(\begin{array}{c|c} X_{11} & X_{12} \\ \hline & X_{22} \end{array}\right)\ , \quad C' = \left(\begin{array}{c|c} Y_{11} & \\ \hline Y_{21} & Y_{22} \end{array}\right)C$$

met X_{11}, X_{22}, Y_{11} en $Y_{22} \in I$ en $X_{11}Y_{11} = I_{11} \in M_{r,r}$. Dit resultaat geeft een overzicht van de mate van niet-eenduidigheid van B en C. (Als $r = m$ of/en $r = n$ dan moeten de formules passend gemodificeerd worden.)

3.3. Bijzondere gevallen van de algemene decompositiestelling. De alternatieve decompositiestelling

Stelling. (speciale decompositiestelling). (3.3.1)

Als $A \in M^n_{m,n}$ (dus $A \in RR$) dan is er een $B \in M^m_{m,m}$ (dus $B \in I$) zo dat

$$A = BJ ,$$

waarin J de normaalvorm uit $M^n_{m,n}$ is.
Als $A \in M^m_{m,n}$ (dus $A \in LR$) dan is er een $C \in M^n_{n,n}$ (dus $C \in I$) zo dat

$$A = JC ,$$

waarin J de normaalvorm uit $M^m_{m,n}$ is.

Bewijs.

Als $m = n$ dan is de zaak triviaal.
Stel $A \in M^n_{m,n}$ met $m > n$. Volgens (3.1.3) en (3.1.2) zijn er dan reguliere B en C zo dat

$$A = BJC = B\left(\frac{I_1}{0}\right)C = B\left(\frac{C}{0}\right) = B\left(\begin{array}{c|c} C & \\ \hline & I_2 \end{array}\right)\left(\frac{I_1}{0}\right) = B'J ,$$

waarin

$$B' := (B_1 | B_2)\left(\begin{array}{c|c} C & \\ \hline & I_2 \end{array}\right) = (B_1C | B_2)$$

volgens (2.1.12) en (2.1.13) regulier is.
Analoog als $A \in M^m_{m,n}$ met $n > m$. □

Corollarium (speciale uitbreidingsstelling) (3.3.2)

Als $A \in M^n_{m,n}$ met $m > n$ (dus $A \in RR$, $A \notin I$), dan is er een $A' \in M_{m,m-n}$ zo dat $(A \mid A') \in M^m_{m,m}$ (dus $\in I$). Immers, volgens (3.3.1) is er een $B \in M^m_{m,m}$ zo dat

$$A = BJ = (B_1 \mid B_2)\left(\frac{I_1}{0}\right) = B_1 .$$

We kunnen dus nemen $A' := B_2$ (en dus $(A \mid A') = B$).
Analoog: als $A \in M^n_{m,n}$ met $m < n$ (dus $A \in LR$, $A \notin I$) dan is er een

$A'' \in M_{m-n,n}$ zo dat $(\frac{A}{A''}) \in M^n_{n,n}$ (dus $\in$ I).

Stelling (alternatieve decompositiestelling). (3.3.3)

Zij $A \in M_{m,n}$, $A \neq 0$.
De volgende uitspraken zijn equivalent:

i) $r(A) = r$;

ii) er is een $B_1 \in M^r_{m,r}$ (dus $\in$ RR)

en een $C_1 \in M^r_{r,n}$ (dus $\in$ LR)

zodat

$$A = B_1C_1 .$$

Bewijs.

i) $\Rightarrow$ ii): De algemene decompositiestelling (3.1.3), samen met de definitie van rang (3.2.3) zegt: er zijn reguliere B en C en een normaalvorm van rang r zodanig dat

$$A = BJC = (B_1 \mid B_2)\left(\begin{array}{c|c} I_1 & \\ \hline & \end{array}\right)\left(\begin{array}{c} C_1 \\ \hline C_2 \end{array}\right) = B_1C_1$$

waarin $B_1 \in M_{m,r}$, $C_1 \in M_{r,n}$. Uit (2.1.14) volgt dat $B_1 \in$ RR, $C_1 \in$ LR.

ii) $\Rightarrow$ i): Volgens (3.3.2) zijn er bij B_1 en C_1 reguliere B en C zo dat

$$B_1 = B\left(\begin{array}{c} I_1 \\ \hline 0 \end{array}\right) , \quad C_1 = (I_1 \mid 0)C$$

met $I_1 \in M_{r,r}$. Hieruit volgt

$$A = B\left(\begin{array}{c} I_1 \\ \hline 0 \end{array}\right)(I_1 \mid 0)C = BJC$$

met J de normaalvorm uit $M^r_{m,n}$. Uit (3.2.3) volgt nu dat $r(A) = r$. □

Corollarium (alternatieve driehoeksdecompositie) (3.3.4)

Als $A \in M^r_{m,n}$, $A \neq 0$, dan zijn er

- permutatiematrices $P \in M_{m,m}$ en $Q \in M_{n,n}$,
- een linkerondermatrix $L_1 \in M^r_{m,r}$ met reguliere diagonaal,

. een rechtsbovenmatrix $R_1 \in M^r_{r,n}$ met reguliere diagonaal
zo dat

$$P^TAQ = L_1R_1 \; .$$

Bewijs.

Volgt uit de driehoeksdecompositie (3.1.5), (2.2.5) en (2.2.6). □

Opmerking. (3.3.5)

We hebben nu twee karakteriseringen van de rang van een matrix A:

i) als $A = BJC$ met B en C regulier, J een normaalvorm, dan is $r(A) = r(J)$,

ii) als $A = B_1C_1$, $B_1 \in RR$, $C_1 \in LR$, dan is $r(A) = \text{breedte}(B_1) = \text{hoogte}(C_1)$.

Beide karakteriseringen zijn onder omstandigheden nuttig. In (4.2.6) wordt nog een andere nuttige karakterisering gegeven.

Opgave. (3.3.6)

Zij $A = B_1C_1 = B_1'C_1'$ met B_1 en $B_1' \in RR$, C_1 en $C_1' \in LR$.
Bewijs met behulp van (2.4.1) rechtstreeks dat $r(B_1) = r(B_1')$ en dat er een reguliere Z is zo dat $B_1' = B_1Z$, $C_1' = Z^{-1}C_1$.
Breng het resultaat in verband met (3.2.2) en (3.2.8).

Opgave. (3.3.7)

Geef een rechtstreeks bewijs van de alternatieve driehoeksdecompositie.
Leid hieruit de andere decompositiestellingen af zonder de resultaten van §2.4 te gebruiken.

Opgave. (3.3.8)

Zij $B_1 \in RR$, $B_1 \notin I$. Zij U_1 linkerinverse van B_1. Volgens de alternatieve decompositiestelling zijn er $B_2 \in RR$ en $U_2 \in LR$ zo dat $I - B_1U_1 = B_2U_2$.

Bewijs dat $\left(\dfrac{U_1}{U_2}\right)$ inverse is van $(B_1 \mid B_2)$ en dat dus $(B_1 \mid B_2) \in I$.

Merk op dat hieruit (3.3.1) volgt, zodat de algemene decompositiestelling (in de vorm $A = BJC$) ook uit de alternatieve decompositiestelling afgeleid kan worden.

<u>Opgave.</u> (3.3.9)

Zij $B = (B_1 \mid B_2) \in I$ met $B_1 \in M_{m,r}$. Bewijs dat $B' := (B_1 \mid B_2') \in I$ dan en slechts dan als

$$B_2' = B_1X_{12} + B_2X_{22}$$

met $X_{12} \in M_{r,m-r}$ en $X_{22} \in M^{m-r}_{m-r,m-r}$ (dus $\in I$).

Als $B^{-1} = U = \left(\frac{U_1}{U_2}\right)$ met $U_1 \in M_{r,m}$, bepaal dan $(B')^{-1}$.

Analoog voor $C = \left(\frac{C_1}{C_2}\right)$, $C' = \left(\frac{C_1}{C_2'}\right)$, $C^{-1} = W = (W_1 \mid W_2)$ met $C_1 \in M_{r,n}$, $W_1 \in M_{n,r}$.

Vergelijk het resultaat met (3.2.8) en met (3.3.2).

3.4. <u>Uitbreidings- en slectiestellingen</u>

In (3.3.2) bleek dat als $A \in RR$ (resp. LR) doch $A \notin I$, A met een aantal kolommen (rijen) uitgebreid kan worden tot een reguliere matrix. Kunnen we, algemener, een matrix $A \in M^r_{m,n}$ met $r < \min(m,n)$ door toevoeging van kolommen en/of rijen uitbreiden tot een matrix die regulier of tenminste links- of rechtsregulier is?

Omgekeerd, kunnen we uit A door schrapping van een aantal kolommen en/of rijen een deelmatrix selecteren die regulier of tenminste links- of rechtsregulier is?

Met behulp van de decompositiestellingen is het antwoord op deze vragen eenvoudig.

<u>Stelling</u> (uitbreidingsstelling). (3.4.1)

Zij $A \in M^r_{m,n}$.

Als $r < m$ dan is er een $A' \in M^{m-r}_{m,m-r}$ zo dat

$$(A \mid A') \in M^m_{m,m+n-r} \qquad \text{(dus } \in LR) .$$

Als $r < n$ dan is er een $A'' \in M^{n-r}_{n-r,n}$ zo dat

$$\left(\frac{A}{A''}\right) \in M^n_{n+m-r,n} \qquad \text{(dus } \in RR) .$$

Bewijs.

Zij $r < m$. Volgens de algemene decompositiestelling is

$$A = BJC = (B_1 \mid B_2)\left(\begin{array}{c|c} I_{11} & \\ \hline & \end{array}\right)\left(\begin{array}{c} C_1 \\ \hline C_2 \end{array}\right) = B_1C_1$$

met $B \in I$, $C_1 \in LR$.

Neem nu $A' := B_2$. Volgens (2.1.14) en (3.2.5) is $B_2 \in M^{m-r}_{m,m-r}$ en er geldt

$$(A \mid A') = (B_1C_1 \mid B_2) = B\left(\begin{array}{c|c} C_1 & \\ \hline & I_{22} \end{array}\right) = BC' .$$

Uit (2.1.12) volgt dat $C' \in LR$, zodat volgens (2.1.10) ook $(A \mid A') \in LR$, dus $\in M^m_{m,m+n-r}$.

Analoog in het andere geval. $\square$

Opmerking. (3.4.2)

Als $r < \min(m,n)$ dan blijkt door twee maal toepassen van de stelling dat A door toevoeging van $n-r$ rijen en $m-r$ kolommen uitgebreid kan worden tot een reguliere $(m+n-r)\times(m+n-r)$ matrix (waarvan A het linksbovenblok is). Ook geldt: als A' en A'' de in het bewijs van (3.4.1) geconstrueerde uitbreidingen van A zijn dan is

$$\left(\begin{array}{c|c} A & A' \\ \hline A'' & 0 \end{array}\right) = \left(\begin{array}{c|c|c} B_1 & B_2 & \\ \hline & & I_{23} \end{array}\right)\left(\begin{array}{c|c|c} I_{11} & & \\ \hline & & I_{23} \\ \hline & I_{32} & \end{array}\right)\left(\begin{array}{c|c} C_1 & \\ \hline C_2 & \\ \hline & I_{32} \end{array}\right)$$

met $I_{23} \in M_{n-r,n-r}$, $I_{32} \in M_{m-r,m-r}$; daar de drie factoren van het rechterlid regulier zijn is het linkerlid dat ook!

Stelling (selectiestelling). (3.4.3)

Zij $A \in M^r_{m,n}$, $A \neq 0$.

a) Als $r < n$ dan is er een permutatiematrix Q zo dat

$$AQ = (A_1 \mid A_2) \text{ met } A_1 \in M^r_{m,r} \text{ en } A_2 = A_1G_{12} ;$$

hieruit volgt

$$A = A_1G_1 \text{ met } A_1 \in RR, G_1 \in LR .$$

b) Als $r < m$ dan is er een permutatiematrix P zo dat

$$P^TA = \left(\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right) \text{ met } A_1 \in M^r_{r,n} \text{ en } A_2 = F_{21}A_1 \ ;$$

hieruit volgt

$$A = F_1A_1 \text{ met } F_1 \in RR,\ A_1 \in LR .$$

c) Als $r < \min(m,n)$ dan zijn er permutatiematrices P en Q zo dat

$$P^TAQ = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

met $A_{11} \in M^r_{r,r}$, $A_{12} = A_{11}G_{12}$, $A_{21} = F_{21}A_{11}$, $A_{22} = F_{21}A_{11}G_{12}$; (*)

hieruit volgt

$$A = F_1A_{11}G_1 \text{ met } F_1 \in RR,\ A_{11} \in I,\ G_1 \in LR . \qquad (**)$$

<u>Bewijs</u>.

We bewijzen alleen het geval c). Zij $r < \min(m,n)$.
Volgens de driehoeksdecompositiestelling zijn er P, Q, L en R zo dat

$$P^TAQ = LJR = \left(\begin{array}{c|c} L_{11} & \\ \hline L_{21} & L_{22} \end{array}\right)\left(\begin{array}{c|c} I_{11} & \\ \hline & \end{array}\right)\left(\begin{array}{c|c} R_{11} & R_{12} \\ \hline & R_{22} \end{array}\right) =$$

$$= \left(\begin{array}{c} L_{11} \\ \hline L_{21} \end{array}\right)(R_{11} \mid R_{12}) := \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

met $A_{ij} = L_{i1}R_{1j}$. Daar L en dus ook L_{11} reguliere linkerdriehoeksmatrix is en R en dus ook R_{11} een reguliere rechterdriehoeksmatrix is, gelden (*) en (**) met

$$F_{21} = L_{21}L_{11}^{-1} , \quad G_{12} = R_{11}^{-1}R_{12} , \quad F_1 = P\left(\begin{array}{c} I_{11} \\ \hline F_{12} \end{array}\right) , \quad G_1 = (I_{11} \mid G_{12})Q^T .$$

Analoog in de andere gevallen. □

Definitie. (3.4.4)

Zij $A \in M_{m,n}$, $1 \leq m' \leq m$, $1 \leq n' \leq n$. $A' \in M_{m',n'}$ heet submatrix van A indien A' verkregen wordt uit A door $m-m'$ rijen en $n-n'$ kolommen te schrappen. Of meer formeel: indien er permutaties p van $\{1,\ldots,m\}$ en q van $\{1,\ldots,n\}$ met bijbehorende permutatiematrices P en Q zijn zo dat

i) $1 \leq i_1 < i_2 \leq m' \Rightarrow p(i_1) < p(i_2)$, $1 \leq j_1 < j_2 \leq n' \Rightarrow q(j_1) < q(j_2)$.

ii) $$P^T AQ = \left(\begin{array}{c|c} A' & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right), \text{ dus } (A')_{ij} = (A)_{p(i),q(j)} \text{ voor } 1 \leq i \leq m',\ 1 \leq j \leq n'.$$

Opgave. (3.4.5)

Ga na dat de formele definitie klopt met de definitie door middel van schrapping.

Bewijs dat als wel aan ii) maar niet aan i) voldaan is, er permutatiematrices P' en Q' zijn zo dat $P'^T A'Q'$ submatrix van A is.

Opgave. (3.4.6)

Bewijs dat een $m \times n$ matrix $(2^m-1)(2^n-1)$ submatrices heeft, waaronder $\binom{m+n}{m} - 1$ vierkante.

Met behulp van het begrip submatrix kunnen we de inhoud van (3.4.4) herformuleren.

Stelling. (3.4.7)

Zij $A \in M^r_{m,n}$, $A \neq 0$.

a) A heeft een submatrix $A' \in M^r_{m,r}$ (dus hoogte m, breedte r en $\in RR$), er is een $G_1 \in LR$ zo dat $A = A'G_1$.

b) A heeft een submatrix $A' \in M^r_{r,n}$ (dus breedte n, hoogte r en $\in LR$), er is een $F_1 \in RR$ zo dat $A = F_1A'$.

c) A heeft een submatrix $A' \in M^r_{r,r}$ (dus $r \times r$ en $\in I$), er zijn $F_1 \in RR$ en $G_1 \in LR$ zo dat $A = F_1A'G_1$.

Bewijs.

Zelf, gebruik (3.4.5). □

Opmerking. (3.4.8)

We zullen in (4.1) bewijzen dat

. voor iedere submatrix A' van A geldt $r(A') \leq r(A)$; r is dus de afmeting van de grootste reguliere submatrix van A;

. voor iedere submatrix A' van A met $r(A') = r(A)$ (maar niet noodzakelijk $A' \in I$) geldt: er zijn $F_1 \in RR$ en $G_1 \in LR$ zo dat $A = F_1A'G_1$.

4. Toepassingen van de decompositiestellingen

4.1. Eigenschappen van de rang

Stelling. (4.1.1)

Zij $A \in M^r_{m,n}$. Dan geldt

- $A^T \in M^r_{n,m}$,
- als $\alpha \neq 0$ dan $\alpha A \in M^r_{m,n}$,
- als
$$A' = \left(\begin{array}{c|c} A & \\ \hline & \end{array}\right) \in M_{m+p,n+q} ,$$
dan $A' \in M^r_{m+p,n+q}$.

Bewijs.

Stel, conform (3.2.3), $A = BJC$ met B en $C \in I$ en J de normaalvorm uit $M^r_{m,n}$. Dan geldt

- $A^T = C^T J^T B^T$
- $\alpha A = (\alpha B)JC$
- $$\left(\begin{array}{c|c} A & \\ \hline & \end{array}\right) = \left(\begin{array}{c|c} B & \\ \hline & I_1 \end{array}\right)\left(\begin{array}{c|c} J & \\ \hline & \end{array}\right)\left(\begin{array}{c|c} C & \\ \hline & I_2 \end{array}\right) ,$$
met $I_1 \in M_{p,p}$, $I_2 \in M_{q,q}$.

Ga nu na dat in de drie rechterleden de eerste en derde factor regulier zijn en dat de tweede factor een passende normaalvorm is. □

Stelling. (4.1.2)

Als $A' = BAC$ met $B \in RR$, $C \in LR$ dan is $r(A') = r(A)$.

Bewijs.

Stel, conform (3.3.3), $A = B_1C_1$, $B_1 \in RR$, $C_1 \in LR$, dan $r(B_1) = r(C_1) = r(A) = r$ (zeg). Nu is $A' = BB_1 . C_1C$. Uit (2.1.12) volgt dat $BB_1 \in RR$ en $C_1C \in LR$. Dus is volgens (3.3.3) $r(A') = r$, want BB_1 heeft breedte r. □

<u>Stelling</u>. (4.1.3)

Als $A = (A_1 \mid A_2)$, $A \in M^r_{m,n}$, $A_j \in M^{r_j}_{m,n_j}$ $(j = 1,2)$, dan is

$$\max(r_1,r_2) \le r \le r_1 + r_2 ,$$

$$r - n_2 \le r_1 \le r,\ r - n_1 \le r_2 \le r .$$

<u>Bewijs</u>.

a. Stel $A = B_1C_1$, conform (3.3.3), met $B_1 \in RR$, $C_1 \in LR$.
   Zij $C_1 = (C_{11} \mid C_{12})$ zodanig dat $A_j = B_1C_{1j}$, $(j = 1,2)$. Daar $B_1 \in RR$ is volgens (4.1.2) $r_j = r(C_{1j})$, dus $r_j \le r$, want C_{1j} heeft hoogte r.
b. Stel $A_j = B_{1j}C_{j1}$ met $B_{1j} \in RR$, $C_{j1} \in LR$ $(j = 1,2)$. Dan is

$$A = (B_{11}C_{11} \mid B_{12}C_{12}) = (B_{11} \mid B_{12})\left(\begin{array}{c|c} C_{11} & \\ \hline & C_{21} \end{array}\right) .$$

   Daar volgens (2.1.13) de tweede factor $\in$ LR, is volgens (4.1.2) $r = r((B_{11} \mid B_{12})) \le \text{breedte}\ ((B_{11} \mid B_{12})) = r_1 + r_2$. Tenslotte volgt uit $r \le r_1 + r_2$ dat $r_1 \ge r - r_2 \ge r - n_2$ en analoog voor r_2.

<u>Gevolg</u>. (4.1.4)

Als $A' \in M^{r'}_{m',n'}$ een submatrix is van $A \in M^r_{m,n}$, dan is

$$r - (m-m') - (n-n') \le r' \le r .$$

<u>Stelling</u> (Sylvester). (4.1.5)

Als $A = A_1A_2$, $A \in M^r_{m,n}$, $A_1 \in M^{r_1}_{m,k}$, $A_2 \in M^{r_2}_{k,n}$, dan is

$$r_1 + r_2 - k \le r \le \min(r_1,r_2) .$$

<u>Bewijs</u>.

Zij $A_1 = BJC$, $A_2 = B'J'C'$ met B $(m \times m)$, C $(k \times k)$, B' $(k \times k)$ en C' $(n \times n)$ regulier, $J \in M^{r_1}_{m,k}$ en $J' \in M^{r_2}_{k,n}$ normaalvormen.

Dan is $A = B \cdot JDJ' \cdot C'$ met $D = CB'$ $(k \times k$, regulier). Zij

$$D = \left(\begin{array}{c|c} D_{11} & D_{12} \\ \hline D_{21} & D_{22} \end{array}\right), \text{ met } D_{11} \in M_{r_1,r_2} .$$

Dan is

$$JDJ' = \left(\begin{array}{c|c} D_{11} & 0 \\ \hline 0 & 0 \end{array}\right) .$$

Volgens (4.1.2) en (4.1.1) is $r(A) = r(JDJ') = r(D_{11})$. Daar D $k \times k$ en regulier is volgt uit (4.1.4)

$$r_1 + r_2 - k = k - (k - r_1) - (k - r_2) \leq r(D_{11}) \leq \min(r_1,r_2). \qquad \square$$

<u>Opmerkingen</u>. (4.1.6)

1. $A_1A_2 = 0$ impliceert dus $r_1 + r_2 \leq k$.
2. Met $m_1 := m$, $n_1 := m_2 := k$, $n_2 := n$ (dus $A_1 \in M_{m_1,n_1}$, $A_2 \in M_{m_2,n_2}$) kunnen we de linkerongelijkheid ook formuleren als

$$m - r \leq (m_1 - r_1) + (m_2 - r_2)$$

of, equivalent, als

$$n - r \leq (n_1 - r_1) + (n_2 - r_2) .$$

<u>Stelling</u> (Frobenius). (4.1.7)

Als $A = A_1A_2A_3$ dan is

$$r(A) \geq r(A_1A_2) + r(A_2A_3) - r(A_2) .$$

<u>Bewijs</u>.

Zij, conform (3.3.3), $A_2 = B_2C_2$ met $B_2 \in RR$, $C_2 \in LR$, $r(B_2) = r(C_2) = r(A_2) = r_2$ (zeg). Dan is $A = A_1B_2 \cdot C_2A_3$ en volgens (4.1.5) is $r(A) \geq r(A_1B_2) + r(C_2A_3) - r_2$. Daar $C_2 \in LR$ is volgens (4.1.2) $r(A_1B_2) = r(A_1B_2C_2) = r(A_1A_2)$, analoog $r(C_2A_3) = r(A_2A_3)$. $\square$

<u>Stelling</u>. (4.1.8)

Als $A = A_1 + A_2$, $r(A) = r$, $r(A_j) = r_j$ $(j = 1,2)$, dan is

$$|r_1 - r_2| \le r \le r_1 + r_2 .$$

Bewijs.

$A = (A_1 \mid A_2)\binom{I}{I}$ met I de $n \times n$ eenheidsmatrix. Dus volgens (4.1.6) en (4.1.3) $r \le r((A_1 \mid A_2)) \le r_1 + r_2$. Daar $A_1 = A - A_2$ geldt nu ook $r_1 \le r + r_2$ en analoog $r_2 \le r + r_1$. □

Stelling. (4.1.9)

Als

$$A = \left(\begin{array}{c|c} A_{11} & \\ \hline & A_{22} \end{array}\right)$$

met $A_{11} \in M^{r_1}_{m_1,n_1}$, $A_{22} \in M^{r_2}_{m_2,n_2}$, dan is $r(A) = r_1 + r_2$.

Gevolg. (4.1.10)

Zij

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

met $A_{11} \in I$. Dan is, als $A_{22.1} := A_{22} - A_{21}A_{11}^{-1}A_{12}$,

$$r(A) = r(A_{11}) + r(A_{22.1}) .$$

(Dit volgt uit de decompositie (2.3.1)).

Opgave. (4.1.11)

Als

$$A = \left(\begin{array}{c|c} A_{11} & \\ \hline A_{21} & A_{22} \end{array}\right)$$

met $A_{11} \in M^{r_1}_{m_1,n_1}$, $A_{22} \in M^{r_2}_{m_2,n_2}$, dan is

$$r_1 + r_2 \le r(A) \le r_1 + r_2 + \min(n_1-r_1,m_2-r_2) = \min(r_1+m_2,n_1+r_2) .$$

Analoog gespiegeld.

Hint: merk op dat

$$A = \left(\begin{array}{c|c} A_{11} & \\ \hline & I_{11} \end{array}\right)\left(\begin{array}{c|c} I_{11} & \\ \hline A_{21} & I_{22} \end{array}\right)\left(\begin{array}{c|c} I_{11} & \\ \hline & A_{22} \end{array}\right) ;$$

en pas nu (4.1.5) toe.

Opmerking. (4.1.12)

Als $A_{11} \in RR$ of $A_{22} \in LR$ dan is $r(A) = r_1 + r_2$ (bewijs dit ook rechtstreeks).

Opgave. (4.1.13)

$$\text{Zij } A = BJC = (B_1 \mid B_2)\left(\begin{array}{c|c} I_{11} & \\ \hline & \end{array}\right)\left(\begin{array}{c} C_1 \\ \hline C_2 \end{array}\right) \in M^r_{m,n}.$$

a) Bij iedere $A' \in M_{m,k}$ zijn er eenduidig bepaalde $C_1' \in M_{r,k}$ en $C_2' \in M_{m-r,k}$ zo dat $A' = B_1C_1' + B_2C_2'$.
Bewijs dat $r((A \mid A') = r + k \Leftrightarrow r(C_2') = k \Leftrightarrow C_2' \in RR$.

b) Karakteriseer analoog de matrices $A'' \in M_{\ell,n}$ waarvoor $r((\frac{A}{A''})) = r + \ell$.

c) Zij $A' \in M_{m,k}$ en $A'' \in M_{\ell,n}$ zo dat $r((A \mid A')) = r + k$, $r((\frac{A}{A''})) = r + \ell$.

Dan is $r\left(\left(\begin{array}{c|c} A & A' \\ \hline A'' & 0 \end{array}\right)\right) = r + k + \ell$.

Vergelijk de resultaten met (3.4.1), (3.3.9) en (3.4.2).

Stelling. (4.1.14)

Zij $A = A_1A_2$. Dan geldt

- $r(A) = r(A_1) \Leftrightarrow \exists_X : A_1 = AX = A_1A_2X$,
- $r(A) = r(A_2) \Leftrightarrow \exists_Y : A_2 = YA = YA_1A_2$.

Bewijs.

We bewijzen alleen de eerste uitspraak.
$\Rightarrow$: Zij $r(A) = r(A_1) = r$. Zij, conform (3.3.3), $A_1 = B_1C_1$ met $B_1 \in RR$, $C_1 \in LR$. Dan is $A = B_1 \cdot C_1A_2$, dus volgens (4.1.2) $r(C_1A_2) = r(A) = r$. Nu is de hoogte van C_1, en dus ook die van C_1A_2, gelijk aan $r(A_1) = r = r(C_1A_2)$. Dus volgens (3.2.5) $C_1A_2 \in RI$, dus $\exists_Z : C_1A_2Z = I$, dus $AZ = B_1$, dus

$A_1 = B_1C_1 = AZC_1 = AX$.
$\Leftarrow$: triviaal met (4.1.6). □

Opmerking. (4.1.15)

Een nuttige variant op bovenstaande stelling is:
Stel dat A_1 en A_2 dezelfde hoogte hebben. Van de drie uitspraken

i) $r(A_1) = r(A_2)$,

ii) $\exists_{X_{12}} : A_2 = A_1X_{12}$,

iii) $\exists_{X_{21}} : A_1 = A_2X_{21}$

impliceert ieder tweetal de derde.
En analoog: stel dat A_1 en A_2 dezelfde breedte hebben. Van de drie uitspraken

i) $r(A_1) = r(A_2)$,

ii) $\exists_{Y_{21}} : A_2 = Y_{21}A_1$,

iii) $\exists_{Y_{12}} : A_1 = Y_{12}A_2$

impliceert ieder tweetal de derde.

Opgave. (4.1.16)

Zij $A \in M^r_{m,n}$, $U \in M_{p,m}$, $W \in M_{n,q}$. Dan geldt

$$r(UA) = r(AW) = r \Leftrightarrow r(UAW) = r \; .$$

Bewijs dit zowel met (4.1.15), met (4.1.7) als met (3.3.3).

Stelling. (4.1.17)

Als $A = (A_1 \mid A_2)$ dan geldt

$$r(A) = r(A_1) \Leftrightarrow \exists_{X_{12}} : A_2 = A_1X_{12} \; .$$

Als $A = \left(\frac{A_1}{A_2}\right)$ dan geldt

$$r(A) = r(A_1) \Leftrightarrow \exists_{Y_{21}} : A_2 = Y_{21}A_1 \; .$$

Bewijs.

We bewijzen alleen de eerste uitspraak. Er geldt

$$A_1 = A\left(\frac{I_{11}}{0}\right) ,$$

waarin I_{11} een eenheidsmatrix is met afmetingen gelijk aan de breedte van A_1.
Uit (4.1.14) volgt nu dat $r(A_1) = r(A) \Leftrightarrow \exists_{X_1} : A = A_1X_1$, of wel
$(A_1 \mid A_2) = A_1(X_{11} \mid X_{12})$, dus $\exists_{X_{12}} : A_2 = A_1X_{12}$. □

Opgave. (4.1.18)

Bewijs (4.1.17) ook rechtstreeks, analoog met het bewijs van (4.1.14).
Bewijs ook (4.1.14) uit (4.1.17).

Opgave. (4.1.19)

Als

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

en $r(A_{11}) = r(A)$, dan zijn er $X_1 \in LR$ en $Y_1 \in RR$ zo dat

$$A = Y_1A_{11}X_1 .$$

Vergelijk dit met (3.4.8).

Opgave. (4.1.20)

Zij $A \in M_{n,n}$ en

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & \end{array}\right)$$

met $A_{11} \in M_{n-1,n-1}$.
Als A_{11} regulier is dan geldt: A regulier $\Leftrightarrow A_{21}A_{11}^{-1}A_{12} \neq 0$.
Als A_{11} singulier is dangeldt:

$$A \text{ regulier} \Leftrightarrow r((A_{11} \mid A_{12})) = r\left(\left(\frac{A_{11}}{A_{21}}\right)\right) = n - 1 .$$

Opgave. (4.1.21)

Zij A regulier, $A = (A_1 \mid A_2)$, $A^{-1} = \binom{X_1}{X_2}$ met $A_1 \in M_{n,k}$, $X_1 \in M_{k,n}$.
Zij $E_1 \in M_{n,k}$. Bewijs dat

$$(A_1 + E_1 \mid A_2) \in I \Leftrightarrow I_{11} + X_1E_1 \in I \quad (k\times k)$$

en zo ja druk dan $(A_1 + E_1 \mid A_2)^{-1}$ uit in X_1, X_2 en $(I_{11} + X_1E_1)^{-1}$.

Opgave. (4.1.22)

Zij $A \in M_{n,n}$ regulier, $B \in M_{n,1}$, $C \in M_{1,n}$. Bewijs dat

$$A + BC \text{ regulier} \Leftrightarrow CA^{-1}B \neq -1$$

en zo ja dat dan

$$(A + BC)^{-1} = A^{-1} - (1 + CA^{-1}B)^{-1}A^{-1}BCA^{-1}$$

(formule van Sherman en Morrison, merk op dat $A^{-1}B \in M_{n,1}$, $CA^{-1} \in M_{1,n}$ en $CA^{-1}B \in M_{1,1}$).

Opgave. (4.1.23)

Zij $A \in M_{n,n}$ regulier, $D \in M_{k,k}$ regulier, $B \in M_{n,k}$, $C \in M_{k,n}$. Dan geldt:

$$n - r(A-BD^{-1}C) = k - r(D-CA^{-1}B) .$$

Als $A - BD^{-1}C$ en $D - CA^{-1}B$ regulier zijn dan geldt

$$(A-BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D-CA^{-1}B)^{-1}CA^{-1} .$$

(formule van Woodbury).
Zie in dat (4.1.21) en (4.1.22) speciale gevallen hiervan zijn. Bekijk ook de gevallen $A = I$, $D = I$.

4.2. De vergelijkingen AX = Z en YA = Z

In deze paragraaf is $A \in M^r_{m,n}$. Gevraagd wordt bij gegeven $Z \in M_{m,q}$ een $X \in M_{n,q}$ zo dat

$$AX = Z . \tag{4.2.1}$$

En analoog vragen we bij gegeven $Z \in M_{p,n}$ een $Y \in M_{p,m}$ zo dat

$$YA = Z \ . \tag{4.2.2}$$

(Omdat we X en Z in kolommen kunnen partitioneren is het eigenlijk voldoende om het geval $q = 1$ te beschouwen; omdat we Y en Z in rijen kunnen partitioneren is het geval $p = 1$ voldoende.)

<u>Stelling</u>. (4.2.3)

Zij $A = BJC$ met B en C regulier en J de normaalvorm uit $M^r_{m,n}$.
Zij $B^{-1} = U = \left(\frac{U_1}{U_2}\right)$ met $U_1 \in M_{r,m}$, $C^{-1} = W = (W_1 \mid W_2)$ met $W_1 \in M_{n,r}$.
Dan geldt:

1) (4.2.1) is oplosbaar dan en slechts dan als $U_2Z = 0$;
   als (4.2.1) oplosbaar is dan is de verzameling der oplossingen

$$\{X \mid X = W_1U_1Z + W_2S_2, S_2 \in M_{n-r,q} \text{ willekeurig}\} \ .$$

2) (4.2.2) is oplosbaar dan en slechts dan als $ZW_2 = 0$;
   als (4.2.2) oplosbaar is dan is de verzameling der oplossingen

$$\{Y \mid Y = ZW_1U_1 + T_2U_2, \ T_2 \in M_{p,m-r} \text{ willekeurig}\} \ .$$

<u>Bewijs</u>.

Beschouw de vergelijking $BJCX = Z$. Dit is equivalent met $JS = UZ$, $X = WS$.
Partitioneer de vergelijking $JS = UZ$:

$$\left(\begin{array}{c|c} I_{11} & \\ \hline & \end{array}\right)\left(\begin{array}{c} S_1 \\ \hline S_2 \end{array}\right) = \left(\begin{array}{c} U_1 \\ \hline U_2 \end{array}\right) Z \ ,$$

dus
$$S_1 = U_1Z \ ,$$
$$0 = U_2Z \ .$$

Hieruit volgt de oplosbaarheidsvoorwaarde $U_2Z = 0$. En als hieraan voldaan is dan voldoet $S_1 = U_1Z$, S_2 willekeurig en dus

$$X = W_1S_1 + W_2S_2 = W_1U_1Z + W_2S_2 \ .$$

Analoog in het andere geval. □

Speciale gevallen. (4.2.4)

1. Als $r = m = n$ dan bestaat A^{-1} (= WU). De vergelijking (4.2.1) en (4.2.2) hebben dan voor iedere Z een eenduidige oplossing, nl.

$$X = A^{-1}Z, \text{ resp. } Y = ZA^{-1} .$$

2. Als $r = m < n$ dan is $A \in RI$ zodat bij (4.2.1) de oplosbaarheidsvoorwaarde vervalt; ook is $A \in LR$ zo dat de oplossing van (4.2.2), als hij bestaat, eenduidig is.

   We kunnen (zie (3.3.1)) in dit geval C zo kiezen dat

$$A = JC = (I_1 \mid 0)\left(\frac{C_1}{C_2}\right) = C_1 .$$

   Als dan weer $C^{-1} = W = (W_1 \mid W_2)$, dan geldt
   - W_1 is rechterinverse van A,
   - (4.2.1) is oplosbaar voor iedere Z en de oplossingsverzameling is

$$\{X \mid X = W_1Z + W_2S_2,\ S_2 \text{ willekeurig}\} ,$$

   - (4.2.2) is oplosbaar dan en slechts dan als $ZW_2 = 0$,
   - als (4.2.2) oplosbaar is dan is de oplossing

$$Y = ZW_1 .$$

3. Als $r = n \leq m$ dan is $A \in RR$ zo dat de oplossing van (4.2.1), als hij bestaat, eenduidig is; ook is $A \in LI$ zo dat bij (4.2.2) de oplosbaarheidsvoorwaarde vervalt.

   Als we B zo kiezen dat

$$A = BJ = (B_1 \mid B_2)\left(\frac{I_1}{0}\right) = B_1$$

   en $B^{-1} = U = \left(\frac{U_1}{U_2}\right)$, dan geldt
   - U_1 is linkerinverse van A,
   - (4.2.1) is oplosbaar dan en slechts dan als $U_2Z = 0$,
   - als (4.2.1) oplosbaar is dan is de oplossing

$$X = U_1Z ,$$

   - (4.2.2) is oplosbaar voor iedere Z, de oplossingsverzameling is

$$\{Y \mid Y = ZU_1 + T_2U_2,\ T_2 \text{ willekeurig}\}\ .$$

Opgave. (4.2.5)

Ga met behulp van (3.3.9) na dat de oplosbaarheidsvoorwaarden en de oplossingen onafhankelijk zijn van de gekozen splitsingen van A.

De inhoud van (4.2.3) beschrijft de oplosbaarheid en oplossingsverzameling volledig. We geven nu echter nog enige formuleringen die geen beroep doen op een splitsing van A.

Stelling. (4.2.6)

Zij $A \in M^r_{m,n}$. Dan geldt

$$n - r = \max\{r(X) \mid AX = 0\} = \max\{k \mid X \in M^k_{n,k},\ AX = 0\}\ ,$$

$$m - r = \max\{r(Y) \mid YA = 0\} = \max\{k \mid Y \in M^k_{k,m},\ YA = 0\}\ .$$

Bewijs.

Uit (4.2.3) volgt (in de notatie van (4.2.3)) dat $AX = 0 \Leftrightarrow X = W_2S_2$ met S_2 willekeurig. Hieruit volgt $r(X) \le r(W_2) = n-r$. En daar we $X = W_2$ kunnen nemen is deze bovengrens voor $r(X)$ ook bereikbaar, zelfs met een $X \in RR$. Analoog voor de tweede uitspraak.

Stelling. (4.2.7)

Zij $A \in M^r_{m,n}$.
Als $AX_0 = 0$ en $r(X_0) = n-r$ dan geldt
i) $AX = 0 \Leftrightarrow \exists_S : X = X_0S$,
ii) $YA = Z$ oplosbaar $\Leftrightarrow ZX_0 = 0$.
Als $Y_0A = 0$ en $r(Y_0) = m-r$ dan geldt
i) $YA = 0 \Leftrightarrow \exists_T : Y = TY_0$,
ii) $AX = Z$ oplosbaar $\Leftrightarrow Y_0Z = 0$.

Bewijs.

Uit (4.2.6) volgt dat de genoemde matrices X_0 en Y_0 bestaan. Uit (4.2.3) volgt (in de notatie van (4.2.3)) dat $AX_0 = 0$ en $r(X_0) = n-r$ impliceert $X_0 = W_2S_{20}$ met (daar $W_2 \in RR$ en rang $n-r$ heeft) $S_{20} \in LR$, dus $\in RI$. Maar dan is de uitspraak $\exists_S : X = X_0S$ equivalent met $\exists_{S_2} : X = W_2S_2$ en de uitspraak $ZX_0 = 0$ is equivalent met $ZW_2 = 0$. De eerste twee beweringen volgen nu uit (4.2.3). □

Stelling. (4.2.8)

$AX = Z$ is oplosbaar $\Leftrightarrow \forall_Y : (YA = 0 \Rightarrow YZ = 0)$,
$YA = Z$ is oplosbaar $\Leftrightarrow \forall_X : (AX = 0 \Rightarrow ZX = 0)$.

Bewijs.

Dat de voorwaarden nodig zijn is triviaal. Dat ze voldoende zijn volgt uit (4.2.7): neem $Y = Y_0$ met Y_0 zo dat $Y_0A = 0$ en $r(Y_0) = m-r$, resp. $X = X_0$ met $AX_0 = 0$ en $r(X_0) = n-r$. □

Stelling. (4.2.9)

$$AX = Z \text{ is oplosbaar } \Leftrightarrow r((A \mid Z)) = r(A) .$$

$$YA = Z \text{ is oplosbaar } \Leftrightarrow r\left(\left(\frac{A}{Z}\right)\right) = r(A) .$$

Bewijs.

Als $AX = Z$ dan is $(A \mid Z) = A(I \mid X)$, dus $r((A \mid Z)) = r(A)$, want $(I \mid X) \in LR$. Als $r((A \mid Z)) = r(A) = r$, dan is er volgens (4.2.6) een Y_0 met $r(Y_0) = m-r$ en $Y_0(A \mid Z) = 0$, dus $Y_0A = 0$ en $Y_0Z = 0$. Volgens (4.2.7) is dit voldoende voor de oplosbaarheid van $AX = Z$.

Opgave. (4.2.10)

De uitspraken van (4.2.9) zijn geheel equivalent met die van (4.1.17)! Bewijs (4.2.7) met behulp van (4.2.9).

Opgave. (4.2.11)

Zij $A \in M^r_{m,n}$, $Z \in M_{m,p}$,

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right), \quad Z = \left(\begin{array}{c} Z_1 \\ \hline Z_2 \end{array}\right),$$

met $A_{11} \in M^r_{r,r}$, $Z_1 \in M_{r,p}$.

Bewijs dat het stelsel $AX = Z$ oplosbaar is dan en slechts dan als

$$Z_2 = A_{21}A_{11}^{-1}Z_1$$

en dat de oplossingen, als zij bestaan, zijn

$$X = \begin{pmatrix} A_{11}^{-1}Z_1 \\ \\ 0 \end{pmatrix} + \begin{pmatrix} -A_{11}^{-1}A_{12} \\ \\ I_{22} \end{pmatrix} X_2$$

met willekeurige $X_2 \in M_{n-r,p}$.

Breng deze uitspraken in verband met bovenstaande stellingen, zowel als met de elementaire oplossingsprocedure via eliminatie van X_1 (gebruik dat $A_{22.1} = 0$).

Doe hetzelfde voor het stelsel $YA = Z$.

Bewijs nu de stellingen uit deze paragraaf opnieuw, uitgaande van de selectie stelling (3.4.3).

4.3. Gegeneraliseerde inversen

We hebben op één van de resultaten uit (4.2.3) nog onvoldoende nadruk gelegd, namelijk op het feit dat voor iedere Z waarvoor $AX = Z$ oplosbaar is, $X = W_1U_1.Z$ een oplossing is en analoog dat voor iedere Z waarvoor $YA = Z$ oplosbaar is, $Y = Z.W_1U_1$ een oplossing is. Met andere woorden, de (van Z onafhankelijke!) matrix

$$A^- := W_1U_1 \; (\in M_{n,m}) \tag{4.3.1}$$

heeft de eigenschappen

i) als $AX = Z$ oplosbaar is, dan is $X = A^-Z$ een oplossing ,

ii) als $YA = Z$ oplosbaar is, dan is $Y = ZA^-$ een oplossing .

Op grond hiervan noemen we A^- een gegeneraliseerde inverse van A.

Ook de oplosbaarheidsvoorwaarden en op de oplossingsverzamelingen uit (4.2.3) kunnen met behulp van A^- uitgedrukt worden. Hiertoe merken we eerst op dat als (in de notatie van (4.2.3)) $B = (B_1 \mid B_2)$, $C = \left(\frac{C_1}{C_2}\right)$, met $B_1 \in M_{m,r}$, $C_1 \in M_{r,n}$, er geldt

$$AA^- = B_1U_1, \quad A^-A = W_1C_1 ,$$

en (daar $B_1U_1 + B_2U_2 = I$, $W_1C_1 + W_2C_2 = I$)

$$I - AA^- = B_2U_2, \quad I - A^-A = W_2C_2 \; .$$

Daar $B_2 \in RR$ en $C_2 \in LR$ kunnen we de oplosbaarheidsvoorwaarde $U_2Z = 0$ en $ZW_2 = 0$ dus vertalen in

iii) $AX = Z$ is oplosbaar dan en slechts dan als $(I - AA^-)Z = 0$,

iv) $YA = Z$ is oplosbaar dan en slechts dan als $Z(I - A^-A) = 0$.

Daar $C_2 \in RI$ is er bij iedere S_2 een S_2' zo dat $W_2S_2 = (I - A^-A)S_2'$ en omgekeerd; analoog is er bij iedere T_2 een T_2' zo dat $T_2U_2 = T_2'(I - AA^-)$ en omgekeerd. Hieruit volgt

v) de oplossingsverzameling van de vergelijking $AX = 0$ is $\{X \mid X = (I - A^-A)S,\ S \text{ willekeurig}\}$,

vi) de oplossingsverzameling van de vergelijking $YA = 0$ is $\{Y \mid Y = T(I - AA^-),\ T \text{ willekeurig}\}$.

Door combinatie van i) en v), resp. ii) en vi) kunnen ook de oplossingsverzamelingen van $AX = Z$ en $YA = Z$ beschreven worden.

Natuurlijk hangt de door (4.3.1) bepaalde gegeneraliseerde inverse A^- af van de gekozen splitsing $A = BJC$. Kan men een karakterisering van gegeneraliseerde inversen vinden die onafhankelijk is van een splitsing?

Stelling. (4.3.2)

Zij $A \in M_{m,n}$, $A^- \in M_{n,m}$.
Als tenminste één van de uitspraken i) t/m vi) geldt, dan voldoet A^- aan

$$AA^-A = A \; . \tag{4.3.3}$$

Als A^- aan (4.3.3) voldoet dan gelden de uitspraken i) t/m vi).

Bewijs.

Als i) geldt, kies dan daarin $Z = A$. De vergelijking $AX = A$ is zeker oplosbaar; volgens i) is A^-A oplossing, derhalve is $AA^-A = A$.
Als iii) geldt, neem dan daarin $Z = A$.
Als v) geldt, neem dan daarin $S = I$.
Analoog als ii), iv) of vi) gelden.
Stel nu dat A^- aan (4.3.3) voldoet. Als $AX = Z$ oplosbaar is dan is $(I - AA^-)Z = (I - AA^-)AX = 0$. Omgekeerd, als $(I - AA^-)Z = 0$, dan is $X = A^-Z$ kennelijk oplossing van $AX = Z$. Dit bewijst iii) en i).

Tenslotte: voor ieder S is $A(I - A^-A)S = 0$ en als $AX = 0$ dan is $X = (I - A^-A)X$. Dit bewijst v).
Analoog voor ii), iv) en vi). □

De matrix A^- uit (4.3.1) voldoet uiteraard aan (4.3.3). Hij voldoet ook aan (ga na)

$$A^-AA^- = A^- \text{ en } r(A^-) = r(A) . \tag{4.3.4}$$

Stelling. (4.3.5)

Zij $A \in M_{m,n}$, $A^- \in M_{n,m}$. Van de drie eigenschappen

i) $AA^-A = A$

ii) $A^-AA^- = A^-$

iii) $r(A) = r(A^-)$

impliceert ieder tweetal de derde.

Bewijs.

Uit i) volgt met (4.1.6) dat $r(A) \leq r(A^-)$, uit ii) volgt dat $r(A^-) \leq r(A)$, i) en ii) impliceren dus iii).
Stel nu dat i) en iii) gelden. Uit i) volgt met (4.1.5) dat $r(A) = r(AA^-A) \leq r(A^-A) \leq r(A)$, dus $r(A^-A) = r(A)$. Met iii) volgt hieruit dat ook $r(A^-A) = r(A^-)$. Uit (4.1.14) volgt dat er een X is zo dat $A^- = A^-AX$ en dus $A^-AA^- = A^-AA^-AX =$ (met i)) $A^-AX = A^-$.
Analoog als ii) en iii) gelden. □

Definitie. (4.3.6)

Zij $A \in M_{m,n}$, $A^- \in M_{n,m}$.
A^- heet gegeneraliseerde inverse (g-inverse) van A als

$$AA^-A = A .$$

A^- heet reflexieve gegeneraliseerde inverse (rg-inverse) van A als

$$AA^-A = A \text{ en } A^-AA^- = A^- .$$

Eenvoudige eigenschappen. (4.3.7)

1. Iedere matrix A heeft tenminste één rg-inverse.
2. Als A^- g-inverse van A is dan is $r(A^-) \geq r(A)$; het gelijkteken geldt dan en slechts dan als A^- rg-inverse van A is (dit volgt uit (4.1.5) en (4.3.5)).

3. Als A^- rg-inverse van A is dan is A rg-inverse van A^- (de reflexiviteit).
4. Als A regulier is dan is A^{-1} de enige g-inverse (en dus ook de enige rg-inverse) van A.
5. Als $A \in RR$ dan vallen de klassen der g-inversen, der rg-inversen en der linker-inversen van A samen (als $A \in RR$ dan is $AA^-A = A$ equivalent met $A^-A = I$ en dit impliceert $A^-AA^- = A^-$).
6. Als $A \in LR$ dan vallen de klassen der g-inversen, der rg-inversen en der rechter-inversen van A samen.
7. Als $A = 0$ ($\in M_{m,n}$) dan is 0 ($\in M_{n,m}$) de enige rg-inverse van A; iedere $A^- \in M_{n,m}$ is g-inverse van A.
8. Als $A \in M^r_{m,n}$ dan geldt voor alle g-inversen A^- van A

$$r(AA^-) = r(A^-A) = r \ ,$$

$$r(I - AA^-) = m-r, \ r(I - A^-A) = n-r \ .$$

(De laatste regel volgt uit (4.3.2.v) en (4.2.6) of uit (4.1.6.1), (4.1.8) en de voorlaatste regel.)

<u>Opgave.</u> (4.3.8)

Als X en Y g-inversen van A zijn dan is $A^- := YAX$ een rg-inverse van A. A^- is de enige rg-inverse van A waarvoor geldt

$$AA^- = AX \ , \quad A^-A = YA \ .$$

<u>Opgave.</u> (4.3.9)

De vergelijking $A_1XA_2 = Z$ heeft een oplossing dan en slechts dan als ieder der vergelijkingen $A_1Y_1 = Z$ en $Y_2A_2 = Z$ oplosbaar is. Als de vergelijking oplosbaar is en A_1^- en A_2^- zijn g-inversen van A_1, resp. A_2 dan is de verzameling der oplossingen

$$\{X \mid X = A_1^-ZA_2^- + W - A_1^-A_1WA_2A_2^- \ , \ W \text{ willekeurig}\} \ .$$

<u>Opgave.</u> (4.3.10)

Zij A^- een g-inverse van A. Dan is de verzameling van alle g-inversen van A

$$\{X \mid X = A^-AA^- + W - A^-AWAA^-, \ W \text{ willekeurig}\} \ .$$

En de verzameling der rg-inversen van A is

$$\{X \mid X = (A^- + (I - A^-A)W)A(A^- + W(I - AA^-)), \ W \text{ willekeurig}\}.$$

<u>Opgave</u>. (4.3.11)

Zij X_1 en X_2 g-inversen van A. Dan zijn er reguliere matrices Z_{12} en Z_{21} zo dat

$$AX_2 = AX_1Z_{12}, \quad X_2A = Z_{21}X_1A \ ,$$

$$I - AX_1 = Z_{12}(I - AX_2), \quad I - X_1A = (I - X_2A)Z_{21} \ .$$

Hieruit volgt dat noch de oplosbaarheidsvoorwaarden, noch de oplossingsverzamelingen uit (4.3.2) van de speciale keuze van A^- afhangen.

Uit het volgende resultaat blijkt de omvang van de verzameling der g-inversen van een algemene matrix.

<u>Stelling</u>. (4.3.12)

Zij $A = BJC$ met B en C regulier en J de normaalvorm uit $M^r_{m,n}$. Zij $0 < r < \min(m,n)$. A^- is g-inverse van A dan en slechts dan als

$$A^- = C^{-1} \left(\begin{array}{c|c} I_{11} & W_{12} \\ \hline W_{21} & W_{22} \end{array}\right) B^{-1}$$

met $I_{11} \in M_{r,r}$, W_{12}, W_{21} en W_{22} willekeurig. Er geldt

$$r(A^-) = r(A) + r(W_{22} - W_{21}W_{12}) \ .$$

A^- is rg-inverse van A dan en slechts dan als

$$A^- = C^{-1}\left(\frac{I_{11}}{W_{21}}\right)(I_{11} \mid W_{12})B^{-1}$$

met $I_{11} \in M_{r,r}$, W_{12} en W_{21} willekeurig.

<u>Bewijs</u>.

A^- is g-inverse, resp. rg-inverse van A dan en slechts dan als $W := CA^-B$ g-inverse, resp. rg-inverse van J is. Werk dit verder uit. De uitspraak over $r(A^-)$ volgt bv. uit (4.1.10). □

<u>Opgave</u>. (4.3.13)

Leid (4.3.10) af uit (4.3.12).

Opgave. (4.3.14)

A en A^- zijn rg-inverse van elkaar dan en slechts dan als er reguliere B en C en een normaalvorm J bestaan zo dat

$$A = BJC,\ A^- = C^{-1}J^TB^{-1} .$$

Opgave. (4.3.15)

A en A^- zijn rg-inversen van elkaar dan en slechts dan als er reguliere $B = (B_1 \mid B_2)$ en $W = (W_1 \mid W_2)$ bestaan zo dat

$$A(W_1 \mid W_2) = (B_1 \mid 0),\ A^-(B_1 \mid B_2) = (W_1 \mid 0) .$$

Opgave. (4.3.16)

A en A^- zijn rg-inverse van elkaar dan en slechts dan als er B_1 en $W_1 \in RR$ en C_1 en $U_1 \in LR$ bestaan zo dat

$$A = B_1C_1,\ A^- = W_1U_1,\ U_1B_1 = C_1W_1 = I_{11} .$$

Bewijs dit, behalve uit (4.3.14), ook rechtstreeks uit de definitie (4.3.6) en de decompositiestelling (3.4.1). Bewijs op analoge wijze (4.3.5).

Opgave. (4.3.17)

Zij $A = B_1C_1$ met $B_1 \in RR$, $C_1 \in LR$. Zij A^- een g-inverse van A. Dan zijn er een eenduidig bepaalde linkerinverse B_1^- van B_1 en een eenduidig bepaalde rechterinverse C_1^- van C_1 zo dat

$$AA^- = B_1B_1^-,\ \ A^-A = C_1^-C_1 .$$

Bovendien geldt

$$A^-AA^- = C_1^-B_1^-$$

(dus ook: als A^- rg-inverse is dan is $A^- = C_1^-B_1^-$).

Opgave. (4.3.18)

Zij

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right),\ X = \left(\begin{array}{c|c} A_{11}^- & 0_{12} \\ \hline 0_{21} & 0_{22} \end{array}\right) ,$$

waarin $A \in M_{m,n}$, $X \in M_{n,m}$ en A_{11}^- een g-inverse van A_{11} is.
Bewijs dat de volgende uitspraken equivalent zijn:

i) $r(A_{11}) = r(A)$,

ii) $(I - A_{11}A_{11}^-)A_{12} = 0$, $A_{21}(I - A_{11}^-A_{11}) = 0$, $A_{21}A_{11}^-A_{12} = A_{22}$,

iii) X is g-inverse van A.

Wanneer is X rg-inverse van A?

Opgave. (4.3.19)

Zij $A = A_1 + A_2$ en A^- een g-inverse van A. Dan zijn de volgende uitspraken equivalent:

i) $r(A_1) + r(A_2) = r(A)$,

ii) $(I - AA^-)A_1 = 0$, $A_1(I - A^-A) = 0$, $A_1A^-A_1 = A_1$,

iii) $(I - AA^-)A_2 = 0$, $A_2(I - A^-A) = 0$, $A_2A^-A_2 = A_2$,

iv) $A_iA^-A_j = \delta_{ij}A_j$, $i,j = 1,2$.

Opgave. (4.3.20)

Zij A, U en W zo dat $Z := UAW$ bestaat. Zij $X = WZ^-U$, waarin Z^- een g-inverse van Z is. Bewijs dat X g-inverse van A is dan en slechts dan als $r(Z) = r(A)$.
Hint: gebruik (4.1.16) en (4.1.14) of leid het resultaat af uit (4.3.18) door eerst aan te nemen dat U en W normaalvormen zijn.

Opgave. (4.3.21)

Zij $A \in M_{m,n}$ en $X \in M_{n,m}$. Dan geldt

$$r(A) - r(A - AXA) = r(X) - r(X - XAX) .$$

Dit is een generalisatie van (4.3.5)!

Opgave. (4.3.22)

Zij $A = (A_1 \mid A_2)$, A_1^- een g-inverse van A_1 en $A_2' := (I - A_1A_1^-)A_2$. Dan is

$$r(A) = r(A_1) + r(A_2') . \qquad (*)$$

Analoog: als $A = \left(\frac{A_1}{A_2}\right)$ en $A_2' = A_2(I - A_1^-A_1)$ dan geldt (*) eveneens.
Hint: bewijs dat A equivalent is met $A' := (A_1 \mid A_2')$ en dat $r(A') = r(A_1) + r(A_2')$.

Opgave. (4.3.23)

Zij

$$A = \left(\begin{array}{c|c} A_{11} & \\ \hline A_{21} & A_{22} \end{array}\right) .$$

Zij A_{11}^- en A_{22}^- g-inversen van A_{11}, resp. A_{22}.

Bewijs dat

$$r(A) = r(A_{11}) + r(A_{22}) + r((I_{22} - A_{22}A_{22}^-)A_{21}(I_{11} - A_{11}^-A_{11})) .$$

Hint: gebruik (4.3.22) tweemaal.

Bewijs ook dat

$$\left(\begin{array}{c|c} A_{11}^- & \\ \hline -A_{22}^-A_{21}A_{11}^- & A_{22}^- \end{array}\right)$$

een g-inverse van A is dan en slechts dan als $(I_{22} - A_{22}A_{22}^-)A_{21}(I_{11} - A_{11}^-A_{11})=0$.

Hoofdstuk II. Vectorruimten en matrices

5. Matrixtheorie en lineaire algebra

5.1. Vectorruimten

Definitie. (5.1.1)

Zij $\mathcal{L}$ een lichaam met elementen α, β, λ, μ, ξ, η, ...
Een vectorruimte V over $\mathcal{L}$ is een verzameling met elementen $\underline{u},\underline{v},\ldots$ met daarop twee operaties:

a) optelling: $V \times V \to V : \underline{u},\underline{v} \mapsto \underline{u} + \underline{v}$, met eigenschappen

$$\underline{u} + \underline{v} = \underline{v} + \underline{u} \ ,$$

$$(\underline{u} + \underline{v}) + \underline{w} = \underline{u} + (\underline{v} + \underline{w}) \ ,$$

$$\exists_{\underline{0}\in V} : \underline{u} + \underline{0} = \underline{0} + \underline{u} = \underline{u} \ ,$$

$$\forall_{\underline{u}\in V} \ \exists_{(-\underline{u})\in V} : \underline{u} + (-\underline{u}) = \underline{0} \ .$$

b) scalaire vermenigvuldiging: $\mathcal{L} \times V \to V$: $\alpha,\underline{u} \mapsto \alpha\underline{u}$, met eigenschappen

$$(\alpha\beta)\underline{u} = \alpha(\beta\underline{u}) \ ,$$

$$(\alpha + \beta)\underline{u} = \alpha\underline{u} + \beta\underline{u} \ ,$$

$$\alpha(\underline{u} + \underline{v}) = \alpha\underline{u} + \alpha\underline{v} \ ,$$

$1.\underline{u} = \underline{u}$ (waarbij 1 het eenheidselement uit $\mathcal{L}$ is) .

Definitie. (5.1.2)

$V' \subset V$ heet lineaire deelruimte van V als V', met de operaties uit V, weer een vectorruimte is.

Nodig en voldoende hiervoor is: als $\underline{u} \in V'$ en $\underline{v} \in V'$ dan $\underline{u} + \underline{v} \in V'$ en als $\underline{u} \in V'$, $\alpha \in \mathcal{L}$, dan $\alpha\underline{u} \in V'$.

Definitie. (5.1.3)

$\underline{u}_1,\ldots,\underline{u}_n \in V$ heten afhankelijk als er $(\xi_1,\ldots,\xi_n) \in \mathcal{L}^n$ *) is zodat

---

*) $\mathcal{L}^n$ is - voorlopig - het cartesisch product $\mathcal{L} \times \mathcal{L} \times \ldots \times \mathcal{L}$, dus de verzameling der rijen $(\xi_1,\ldots,\xi_n)$ met $\forall_i \xi_i \in \mathcal{L}$.

$$(\xi_1,\ldots,\xi_n) \neq (0,\ldots,0) \text{ en } \sum_1^n \xi_i \underline{u}_i = \underline{0} \ .$$

$\underline{u}_1,\ldots,\underline{u}_n \in V$ heten onafhankelijk als ze niet afhankelijk zijn, dus als

$$\sum_1^n \xi_i \underline{u}_i = \underline{0} \Rightarrow (\xi_1,\ldots,\xi_n) = (0,\ldots,0) \ .$$

Definitie. (5.1.4)

Als $\underline{u}_1,\ldots,\underline{u}_n \in V$ dan heet

$$[\underline{u}_1,\ldots,\underline{u}_n] := \{\underline{u} \mid \exists_{(\xi_1,\ldots,\xi_n)\in\mathcal{L}^n} : \underline{u} = \sum_1^n \xi_i \underline{u}_i\}$$

het lineair opspansel van $\underline{u}_1,\ldots,\underline{u}_n$.

Stelling. (5.1.5)

$[\underline{u}_1,\ldots,\underline{u}_n]$ is een lineaire deelruimte van V.

Definitie. (5.1.6)

Zij $\underline{u}_1,\ldots,\underline{u}_n \in V$. $(\underline{u}_1,\ldots,\underline{u}_n)$ heet basis voor V indien

i) $\underline{u}_1,\ldots,\underline{u}_n$ onafhankelijk,
ii) $[\underline{u}_1,\ldots,\underline{u}_n] = V$.

Stelling. (5.1.7)

$(\underline{u}_1,\ldots,\underline{u}_n)$ basis voor V $\Leftrightarrow \forall_{\underline{u}\in V} \ \exists!_{(\xi_1,\ldots,\xi_n)\in\mathcal{L}^n} : \underline{u} = \sum_1^n \xi_i \underline{u}_i$.

Definitie. (5.1.8)

V heet eindig dimensionaal als er $\underline{u}_1,\ldots,\underline{u}_n$ zijn zo dat $V = [\underline{u}_1,\ldots,\underline{u}_n]$.
V heet oneindig dimensionaal als V niet eindig dimensionaal is.

Stelling. (5.1.9)

V is oneindig dimensionaal $\Leftrightarrow$

$$\forall_{n\in\mathbb{N}} \ \forall_{\underline{u}_1,\ldots,\underline{u}_n\in V} \ \exists_{\underline{u}_{n+1}\in V} : \underline{u}_{n+1} \notin [\underline{u}_1,\ldots,\underline{u}_n] \ .$$

Stelling. (5.1.10)

Als V eindig dimensionaal is dan is iedere lineaire deelruimte van V eindig dimensionaal.

Stelling. (5.1.11)

Ieder eindig dimensionale vectorruimte heeft een basis.

Opmerking. (5.1.12)

De definitie (5.1.1) laat toe dat een vectorruimte slechts uit het element $\underline{0}$ bestaat. En iedere vectorruimte heeft de verzameling $\{\underline{0}\}$ als - zg. triviale - deelruimte. Stellingen als (5.1.11) moeten voor de triviale vectorruimten met enige voorzichtigheid geïnterpreteerd worden: $\{\underline{0}\}$ heeft een basis van 0 elementen (deze interpretatie is in overeenstemming met het gebruik bij lege sommen e.d.). In het vervolg zullen we het triviale uitzonderingsgeval als regel niet apart noemen.

Definitie. (5.1.13)

Zij V_1 en V_2 deelruimten van V. De doorsnede $V_1 \cap V_2$ en de som $V_1 + V_2$ van V_1 en V_2 zijn gedefinieerd door

$$V_1 \cap V_2 := \{\underline{u} \in V \mid \underline{u} \in V_1 \text{ en } \underline{u} \in V_2\} ,$$
$$V_1 + V_2 := \{\underline{u} \in V \mid \exists_{\underline{u}_1 \in V_1} \exists_{\underline{u}_2 \in V_2} : \underline{u} = \underline{u}_1 + \underline{u}_2\} .$$

De schrijfwijze $V_3 = V_1 \oplus V_2$ duidt aan:

$$V_1 \cap V_2 = \{\underline{0}\} \text{ en } V_3 = V_1 + V_2 .$$

Men spreekt in dit geval van de directe som van V_1 en V_2 (ga na dat in dit geval de bij een $\underline{u} \in V_3$ behorende $\underline{u}_1 \in V_1$ en $\underline{u}_2 \in V_2$ eenduidig bepaald zijn). Als $V_1 \oplus V_2 = V$ dan heten V_1 en V_2 complementair t.o.v. elkaar (in V).

Stelling. (5.1.14)

Doorsnede en som van lineaire deelruimten van V zijn lineaire deelruimten van V.

Definitie. (5.1.15)

Deelruimten $V_1, \ldots, V_k$ van V heten onafhankelijk als $\underline{u}_i \in V_i$ $(i = 1, \ldots, k)$ en $\Sigma\, \underline{u}_i = \underline{0}$ impliceert $\underline{u}_i = \underline{0}$ $(i = 1, \ldots, k)$.

Stelling. (5.1.16)

Zij $V_1,\dots,V_k$ deelruimten van V. De volgende beweringen zijn equivalent.

i) $V_1,\dots,V_k$ zijn onafhankelijk.

ii) bij iedere $\underline{u} \in V_1 + \dots + V_k$ zijn er eenduidig bepaalde $\underline{u}_i \in V_i$ $(i = 1,\dots,k)$ zo dat $\underline{u} = \Sigma \underline{u}_i$.

iii) voor alle niet-lege deelverzamelingen I_1 en I_2 van $\{1,\dots,k\}$ geldt

$$I_1 \cap I_2 = \emptyset \Rightarrow \sum_{i\in I_1} V_i \cap \sum_{i\in I_2} V_i = \{\underline{0}\} .$$

Gevolg. (5.1.17)

Als $V_1,\dots,V_k$ onafhankelijke deelruimten van V zijn dan is

$$V_1 + \dots + V_k = ((V_1 \oplus V_2) \oplus \dots) \oplus V_k$$

en hierin mogen de indices willekeurig gepermuteerd worden. We kunnen daarom voor het rechterlid zonder bezwaar $V_1 \oplus V_2 \oplus \dots \oplus V_k$ schrijven.

Opgave. (5.1.18)

Bewijs dat de volgende beweringen ook equivalent zijn met de beweringen uit (5.1.16):

iv) voor $j = 1,\dots,k$ geldt $\left(\sum_{\substack{i=1\\ i\neq j}}^{k} V_i\right) \cap V_j = \{\underline{0}\}$.

v) voor $j = 2,\dots,k$ geldt $\left(\sum_{i=1}^{j-1} V_i\right) \cap V_j = \{\underline{0}\}$.

Opgave. (5.1.19)

Zij V_1,V_2,V_3 deelruimten van V. Bewijs rechtstreeks:
als $(V_1 \oplus V_2) \cap V_3 = \{\underline{0}\}$ dan is ook $V_1 \cap (V_2 \oplus V_3) = \{\underline{0}\}$.

5.2. Lineaire afbeeldingen

Definitie. (5.2.1)

Zij V_1 en V_2 vectorruimten over $\mathcal{L}$.

Een afbeelding A: $V_2 \to V_1$ heet lineair als

$$\forall_{\underline{u},\underline{v}\in V_2}\ \forall_{\alpha,\beta\in\mathcal{L}}: A(\alpha\underline{u} + \beta\underline{v}) = \alpha A\underline{u} + \beta A\underline{v} \ .$$

De verzameling van alle lineaire afbeeldingen van V_2 in V_1 noemen we $LA(V_1,V_2)$. Als A en $B \in LA(V_1,V_2)$, α en $\beta \in \mathcal{L}$ dan is $\alpha A + \beta B \in LA(V_1,V_2)$ gedefinieerd door

$$\forall_{\underline{u}\in V_2} : (\alpha A + \beta B)\underline{u} = \alpha A\underline{u} + \beta B\underline{u} \ .$$

Als $A \in LA(V_1,V_2)$, $B \in LA(V_2,V_3)$ dan is $AB \in LA(V_1,V_3)$ gedefinieerd door

$$\forall_{\underline{u}\in V_3} : (AB)\underline{u} = A(B\underline{u}) \ .$$

<u>Definitie</u>. (5.2.2)

Zij $A \in LA(V_1,V_2)$. Dan is

$N(A) := \{\underline{u} \in V_2 \mid A\underline{u} = \underline{0}\}$, de nulruimte van A ,

$R(A) := \{\underline{v} \in V_1 \mid \exists_{\underline{u}\in V_2} : A\underline{u} = \underline{v}\}$, de range van A .

<u>Stelling</u>. (5.2.3)

Als $A \in LA(V_1,V_2)$ dan is $N(A)$ lineaire deelruimte van V_2 en $R(A)$ lineaire deelruimte van V_1.

<u>Definitie</u>. (5.2.4)

$A \in LA(V_1,V_2)$ heet

injectief (eeneenduidig) als $N(A) = \{\underline{0}\}$,
surjectief (afbeelding op) als $R(A) = V_1$,
bijectief (eeneenduidig op) als $N(A) = \{\underline{0}\}$ en $R(A) = V_1$.

<u>Stelling</u>. (5.2.5)

Als $A \in LA(V_1,V_2)$ injectief is dan heeft iedere $\underline{v} \in R(A)$ precies één origineel in V_2; als A bijectief is dan heeft A een eenduidig bepaalde inverse $A^{-1} \in LA(V_2,V_1)$ zodanig dat

$$\forall_{\underline{u}\in V_2}\ A^{-1}A\underline{u} = \underline{u},\ \forall_{\underline{v}\in V_1}\ AA^{-1}\underline{v} = \underline{v} \ .$$

A^{-1} is weer bijectief en $(A^{-1})^{-1} = A$.

Stelling. (5.2.6)

$A \in LA(V_1,V_2)$ en $B \in LA(V_2,V_1)$ zijn bijectief en elkaars inversen dan en slechts dan als

$$\forall_{\underline{u}\in V_2}\ \forall_{\underline{v}\in V_1} \ : \ (\underline{v} = A\underline{u} \leftrightarrow \underline{u} = B\underline{v}) \ .$$

Definitie. (5.2.7)

Vectorruimte V_1 en V_2 heten isomorf als V_1 en V_2 lineair en bijectief op elkaar afgebeeld kunnen worden.

5.3. De vectorruimte $\mathcal{L}^m$ en de lineaire afbeeldingen $LA(\mathcal{L}^m,\mathcal{L}^n)$

Definitie. (5.3.1)

De ruimte $\mathcal{L}^m$ bestaat uit de geordende rijen van m elementen van $\mathcal{L}$. Optelling en vermenigvuldiging zijn gedefinieerd door: als $\underline{x} = (\xi_1,\ldots,\xi_m)$ en $\underline{y} = (\eta_1,\ldots,\eta_m) \in \mathcal{L}^m$, α en $\beta \in \mathcal{L}$, dan is

$$\alpha\underline{x} + \beta\underline{y} = (\alpha\xi_1 + \beta\eta_1,\ldots,\alpha\xi_m + \beta\eta_m) \ .$$

ξ_i noemen we i-de component of kental of coördinaat van $\underline{x}$. De vectoren $\underline{e}_k = (\delta_{1k},\ldots,\delta_{mk})$, $k = 1,\ldots,m$ (met $\delta_{ij} = 1$ als $i = j$, anders 0) heten coördinaatvectoren in $\mathcal{L}^m$.

Stelling. (5.3.2)

$\mathcal{L}^m$ is een eindig dimensionale vectorruimte over $\mathcal{L}$; $\underline{e}_1,\ldots,\underline{e}_m$ vormen een basis voor $\mathcal{L}^m$ (de zogenaamde natuurlijke basis).

Stelling. (5.3.3)

$M_{m,1}(\mathcal{L})$ en $M_{1,m}(\mathcal{L})$ zijn, met de optelling en vermenigvuldiging met scalairen van matrices, vectorruimten over $\mathcal{L}$ en isomorf met $\mathcal{L}^m$.

Bewijs.

Voeg aan $\underline{x} = (\xi_1,\ldots,\xi_m) \in \mathcal{L}^m$ toe

$$X = \begin{pmatrix} \xi_1 \\ \hline \vdots \\ \hline \xi_m \end{pmatrix} \in M_{m,1}(\mathcal{L}) \text{ en } X^T = (\xi_1 \mid \ldots \mid \xi_m) \in M_{1,m}(\mathcal{L}) \ .$$

Deze toevoegingen zijn lineair en bijectief (ga na). □

Notatie. (5.3.4)

Als $\underline{x} \in \mathcal{L}^m$ dan noemen we het toegevoegde element uit $M_{m,1}$ meestal x (een kolomvector) en het toegevoegde element uit $M_{1,m}$ x^T (een rijvector). Vanaf §5.5 zullen we het verschil tussen $M_{m,1}$ en $\mathcal{L}^m$ vaak geheel negeren en met x zowel een element van $\mathcal{L}^m$ als de bijbehorende kolomvector uit $M_{m,1}$ bedoelen.

We beschouwen nu het verband tussen de lineaire afbeeldingen $\mathcal{A}$ uit $LA(\mathcal{L}^m,\mathcal{L}^n)$ en de matrices A uit $M_{m,n}(\mathcal{L})$.

Definitie. (5.3.5)

Zij $\mathcal{A} \in LA(\mathcal{L}^m,\mathcal{L}^n)$. Zij $(\underline{e}_1,\ldots,\underline{e}_m)$ de natuurlijke basis in $\mathcal{L}^m$ en $(\tilde{\underline{e}}_1,\ldots,\tilde{\underline{e}}_n)$ de natuurlijke basis in $\mathcal{L}^n$. Zij

$$\mathcal{A}\tilde{\underline{e}}_j = \sum_{i=1}^{m} \alpha_{ij}\underline{e}_i \ , \qquad j = 1,\ldots,n \ .$$

Dan heet

$$A = (\alpha_{ij}) \in M_{m,n}(\mathcal{L})$$

de matrix van $\mathcal{A}$ t.o.v. de natuurlijke bases in $\mathcal{L}^m$ en $\mathcal{L}^n$. (De toevoeging "t.o.v. ..." laten we weg als geen verwarring dreigt).

Stelling. (5.3.6)

Als $\mathcal{A}$ en $\mathcal{B} \in LA(\mathcal{L}^m,\mathcal{L}^n)$ met matrices A en B $\in M_{m,n}$, en α en $\beta \in \mathcal{L}$ dan is $\alpha A + \beta B$ de matrix van $\alpha\mathcal{A} + \beta\mathcal{B}$.
Als $\mathcal{A} \in LA(\mathcal{L}^m,\mathcal{L}^n)$ en $\mathcal{B} \in LA(\mathcal{L}^n,\mathcal{L}^k)$ met matrices $A \in M_{m,n}$ en $B \in M_{n,k}$, dan is $AB \in M_{m,k}$ de matrix van $\mathcal{AB} \in LA(\mathcal{L}^m,\mathcal{L}^k)$.

Bewijs.

Volgt uit de definities (5.2.1), (1.2.2), (1.24) en (1.2.6). Ga na. □

Opmerking. (5.3.7)

De definitie van de vermenigvuldiging van twee matrices is geheel gemotiveerd door de wens dat AB de matrix van $\mathcal{AB}$ is. □

Stelling. (5.3.8)

$LA(\mathcal{L}^m,\mathcal{L}^n)$ en $M_{m,n}(\mathcal{L})$ zijn isomorfe vectorruimten.

Bewijs.

Dat $LA(\mathcal{L}^m,\mathcal{L}^n)$ en $M_{m,n}(\mathcal{L})$ vectorruimten over $\mathcal{L}$ zijn volgt uit de definities van optelling en vermenigvuldiging met scalairen voor afbeeldingen, resp. matrices.
De eerste bewering van (5.3.6) zegt dat de door (5.3.5) gegeven afbeelding van $LA(\mathcal{L}^m,\mathcal{L}^n)$ in $M_{m,n}(\mathcal{L})$ lineair is. Dat deze afbeelding ook bijectief is volgt uit het feit dat $\mathcal{A}$ eenduidig door $\mathcal{A}\tilde{\underline{e}}_1,\ldots,\mathcal{A}\tilde{\underline{e}}_n$ (en dus door de kolommen van A) bepaald is. □

Stelling. (5.3.9)

Zij $\mathcal{A} \in LA(\mathcal{L}^m,\mathcal{L}^n)$ en A de bijbehorende matrix. Dan geldt voor iedere $\underline{x} \in \mathcal{L}^n$ en $\underline{y} \in \mathcal{L}^m$ met bijbehorende $x \in M_{n,1}$ en $y \in M_{m,1}$

$$\underline{y} = \mathcal{A}\underline{x} \Leftrightarrow y = Ax\ .$$

Definitie. (5.3.10)

Zij $A \in M_{m,n}(\mathcal{L})$. Dan is

$$N(A) := \{\underline{x} \in \mathcal{L}^n \mid Ax = 0\},\quad \text{de nulruimte van A}$$

$$R(A) := \{\underline{y} \in \mathcal{L}^m \mid \exists_{\underline{x}\in\mathcal{L}^n} : y = Ax\},\quad \text{de range van A}\ .$$

Stelling. (5.3.11)

Als A de matrix is van een afbeelding $\mathcal{A} \in LA(\mathcal{L}^m,\mathcal{L}^n)$ dan is $N(A) = N(\mathcal{A})$, $R(A) = R(\mathcal{A})$.

5.4. Matrices en deelruimten van $\mathcal{L}^m$

Als $A = (a_1 \mid \ldots \mid a_n) \in M_{m,n}$ (met $a_i \in M_{m,1}$) dan heten $\underline{a}_1,\ldots,\underline{a}_n$ de kolomvectoren van A. Uit (5.3.10) en (5.1.4) volgt dan

$$R(A) = [\underline{a}_1,\ldots,\underline{a}_n]\ , \tag{5.4.1}$$

daarom noemen we R(A) ook wel de kolomruimte van A.
Met behulp van deze relatie tussen matrices uit $M_{m,n}$ en deelruimten van $\mathcal{L}^m$ kunnen we nu eigenschappen van deelruimten van $\mathcal{L}^m$ beschrijven en bewijzen via matrixtheorie (het omgekeerde kan ook, zie §5.6).

Stelling. (5.4.2)

Als $A = (a_1 \mid \dots \mid a_n)$ dan geldt: $(\underline{a}_1,\dots,\underline{a}_n)$ is basis voor $R(A) \Leftrightarrow A \in RR$.

Bewijs.

Uit (5.4.1) en (5.1.6) volgt dat de linker uitspraak geldt dan en slechts dan als $\underline{a}_1,\dots,\underline{a}_n$ lineair onafhankelijk zijn. Maar volgens (5.1.3) is dit equivalent met

$$\forall_{(\xi_1,\dots,\xi_n)\in\mathcal{L}^n} \left(\sum_i \xi_i \underline{a}_i = \underline{0} \Rightarrow (\xi_1,\dots,\xi_n) = (0,\dots,0)\right) ,$$

en dus met

$$\forall_{x\in M_{n,1}} (Ax = 0 \Rightarrow x = 0) ,$$

dat wil zeggen: met $A \in RR$. □

Stelling. (5.4.3)

$$R(A) = \mathcal{L}^m \Leftrightarrow A \in RI .$$

Bewijs.

$$R(A) = \mathcal{L}^m \Leftrightarrow \forall_{\underline{y}\in\mathcal{L}^m} \exists_{\underline{x}\in\mathcal{L}^n} : y = Ax$$

$$\Leftrightarrow \exists_{X\in M_{m,n}} : AX = I \Leftrightarrow A \in RI .$$ □

Gevolgen. (5.4.4)

1. $(\underline{a}_1,\dots,\underline{a}_n)$ is basis voor $\mathcal{L}^m \Leftrightarrow n = m$ en $A \in I$.
2. iedere basis voor $\mathcal{L}^m$ heeft m elementen.

Stelling. (5.4.5)

Als A_1 en A_2 beide hoogte m hebben dan geldt

$$R(A_2) \subset R(A_1) \Leftrightarrow \exists_X : A_2 = A_1 X .$$

Gevolgen. (5.4.6)

1. $R(A_2) \subset R(A_1) \Rightarrow r(A_2) \le r(A_1)$, het gelijkteken geldt dan en slechts dan als $R(A_2) = R(A_1)$ (dit laatste volgt uit (4.1.15)).

2. Als $A_1 \in RR$ en $A_2 = A_1X$ dan geldt (ga na)

$$R(A_2) = R(A_1) \Leftrightarrow X \in RI .$$

3. Als $A = B_1C_1$ met $B_1 \in RR$, $C_1 \in LR$ dan is $R(A) = R(B_1)$ en de kolommen van B_1 vormen een basis voor $R(A)$.
4. Ieder basis voor een deelruimte $V \subset \mathcal{L}^m$ heeft evenveel elementen (dit volgt uit (5.4.2) en (5.4.6.1)).

Definitie. (5.4.7)

De dimensie van een deelruimte V van $\mathcal{L}^m$ is het aantal elementen van een basis voor V.

Gevolgen. (5.4.8)

1. $\dim(\mathcal{L}^m) = m$.
2. $\dim(R(A)) = r(A)$ (volgt uit (5.4.6.3)).
3. als $V \subset \mathcal{L}^m$ en $\dim(V) = r$ dan is er een $A \in M^r_{m,r}$ (dus $A \in RR$) zo dat $V = R(A)$.
4. als $V_2 \subset V_1 \subset \mathcal{L}^m$ dan is $\dim(V_2) \leq \dim(V_1)$ met gelijkteken dan en slechts dan als $V_2 = V_1$ (dit is een vertaling van (5.4.6.1)).

Stelling. (5.4.9)

Een deelruimte $V \subset \mathcal{L}^m$ heeft dimensie n dan en slechts dan als V isomorf is met $\mathcal{L}^n$.

Bewijs.

$\Rightarrow$: Als de kolommen $\underline{a}_1,\ldots,\underline{a}_n$ van A een basis voor V vormen dan is de afbeelding $\mathcal{A} \in LA(V,\mathcal{L}^n)$, gegeven door

$$\underline{y} = \mathcal{A}\underline{x} \Leftrightarrow y = Ax,\ \underline{y} \in V,\ x \in \mathcal{L}^n,$$

een bijectie (ga na).

$\Leftarrow$: Als $\mathcal{A} \in LA(V,\mathcal{L}^n)$ een bijectie is dan vormen $(\mathcal{A}\underline{e}_1,\ldots,\mathcal{A}\underline{e}_n)$ (met $(\underline{e}_1,\ldots\underline{e}_n)$ de natuurlijke basis voor $\mathcal{L}^n$) een basis voor V (ga na). □

Gevolg. (5.4.10)

Deelruimten $V_1 \subset \mathcal{L}^{m_1}$ en $V_2 \subset \mathcal{L}^{m_2}$ zijn isomorf dan en slechts dan als ze dezelfde dimensie hebben.

Stelling. (5.4.11)

Zij $A = (A_1 \mid \dots \mid A_k)$ met $A_i \in RR$ $(i = 1,\dots,k)$. Dan zijn de volgende beweringen equivalent:

i) $A \in RR$,

ii) $r(A) = \Sigma r(A_i)$,

iii) $\dim(R(A)) = \Sigma \dim(R(A_i))$,

iv) $R(A_1),\dots,R(A_k)$ onafhankelijk.

Bewijs.

De equivalentie van i), ii) en iii) is triviaal. Uitspraak iv) betekent: $\Sigma A_i x_i = 0 \leftrightarrow \forall_i : A_i x_i = 0 \leftrightarrow \forall_i x_i = 0$ (daar alle $A_i \in RR$) en dit is equivalent met uitspraak i). □

Gevolg. (5.4.12)

Zij $V_1,\dots,V_k$ deelruimten van $\mathcal{L}^m$ en $V := V_1 + \dots + V_k$. Dan zijn de volgende beweringen equivalent:

i) $\dim(V) = \Sigma \dim(V_i)$,

ii) $V_1,\dots,V_k$ onafhankelijk.

Als deze beweringen gelden dan is $V = V_1 \oplus \dots \oplus V_k$ en er is een $A = (A_1 \mid \dots \mid A_k) \in RR$ zodat $R(A_i) = V_i$, $R(A) = V$.

Opgave. (5.4.13)

Zij V_1 en V_2 deelruimten van $V \subset \mathcal{L}^m$. Bewijs dat van de volgende drie uitspraken ieder tweetal de derde impliceert:

i) $V_1 \cap V_2 = \{\underline{0}\}$,

ii) $V_1 + V_2 = V$,

iii) $\dim(V_1) + \dim(V_2) = \dim(V)$.

We noemen en bewijzen nu enige uitbreidingsstellingen.

Stelling. (5.4.14)

Als V_1 een echte deelruimte van $V \subset \mathcal{L}^m$ is en $(\underline{a}_1,\dots,\underline{a}_k)$ een basis voor V_1 is dan zijn er $\underline{a}_{k+1},\dots,\underline{a}_n$ zo dat $(\underline{a}_1,\dots,\underline{a}_n)$ een basis voor V is.

Bewijs.

Zij $(\underline{b}_1,\dots,\underline{b}_n)$ basis voor V. Zij $A_1 = (a_1 \mid \dots \mid a_k)$, $B = (b_1 \mid \dots \mid b_n)$.

Volgens (5.4.5) is dan $A_1 = BX_1$ met $X_1 \in RR$ (omdat $A_1 \in RR$). Volgens (3.3.2) is er een X_2 zo dat $X = (X_1 \mid X_2)$ regulier is. Neem voor $\underline{a}_{k+1},\ldots,\underline{a}_n$ de kolommen van BX_2, dan is $(a_1 \mid \ldots \mid a_n) = BX$. Pas nu (5.4.6.2) toe. □

Gevolg. (5.4.15)

Bij iedere $V_1 \subset V \subset \mathcal{L}^m$ is er een V_2 die t.o.v. V_1 complementair is in V, d.w.z.

$$V_1 \oplus V_2 = V \ .$$

Er geldt: $\dim(V_1) + \dim(V_2) = \dim(V)$.

Opmerking. (5.4.16)

Als in (5.4.14) $(\underline{b}_1,\ldots,\underline{b}_n)$ een gegeven basis voor V is, dan kunnen we voor $\underline{a}_{k+1},\ldots,\underline{a}_n$ een geschikt (n-k)-tal uit $\underline{b}_1,\ldots,\underline{b}_n$ kiezen. Daar nl. de matrix X_1 uit het bewijs $\in RR$ is is er volgens de selectiestelling (3.4.3) een permutatiematrix P zo dat

$$P^T X_1 = \left(\frac{X_{11}}{X_{22}}\right), \text{ dus } A_1 = BP\left(\frac{X_{11}}{X_{21}}\right)$$

met X_{11} regulier. Neem nu voor $\underline{a}_{k+1},\ldots,\underline{a}_n$ de kolommen van $BP\left(\frac{0}{I_{22}}\right)$, dus $\underline{b}_{p(k+1)},\ldots,\underline{b}_{p(n)}$, als p de bij P horende permutatie is.

Stelling. (5.4.17)

Zij V_1 en V_2 onafhankelijke deelruimten van $\mathcal{L}^m$ en $V := V_1 \oplus V_2$.
Zij $(A_1 \mid A_2) \in RR$ zo dat $R(A_1) = V_1$, $R(A_2) = V_2$.
Zij $V_2' = R(A_2')$, met $A_2' \in RR$, ook deelruimte van $\mathcal{L}^m$. Dan geldt:

$$V_1 \text{ en } V_2' \text{ zijn complementair t.o.v. } V, \text{ d.w.z. } V_1 \oplus V_2' = V \ ,$$

dan en slechts dan als er X_{12} en X_{22} zijn zo dat $X_{22} \in I$ en

$$A_2' = A_1 X_{12} + A_2 X_{22} \ .$$

Bewijs.

Daar $V_1 + V_2' = R((A_1 \mid A_2'))$ is volgens (5.4.5) en (5.4.6.2) $V_1 + V_2' = V$ dan en slechts dan als

$$(A_1 \mid A_2') = (A_1 \mid A_2)X$$

met $X \in RI$. En volgens (5.4.11) is $V_1 \cap V_2' = \{\underline{0}\}$ dan en slechts dan als $(A_1 \mid A_2') \in RR$, dus als $X \in RR$.
Derhalve is $V_1 \oplus V_2' = V$ dan en slechts dan als $X \in I$. De matrix X heeft echter noodzakelijk de vorm

$$X = \left(\begin{array}{c|c} I_{11} & X_{12} \\ \hline & X_{22} \end{array}\right)$$

en volgens (2.4.5) geldt dat $X \in I \leftrightarrow X_{22} \in I$. □

Een deelruimte V van $\mathcal{L}^m$ kunnen we, behalve als kolomruimte van een matrix $A \in M_{m,p}$ ook geven als nulruimte van een matrix $A \subset M_{q,m}$.

Stelling. (5.4.18)

Zij $A_1 \in M_{m,p}$, $A_2 \in M_{q,m}$. Dan zijn de volgende uitspraken equivalent:

i) $R(A_1) = N(A_2)$,

ii) $A_2A_1 = 0$ en $r(A_1) + r(A_2) = m$.

Bovendien geldt:

. bij iedere $A_1 \in M^r_{m,p}$ is er een $A_2 \in M^{m-r}_{m-r,m}$ zo dat $N(A_2) = R(A_1)$,

. bij iedere $A_2 \in M^r_{q,m}$ is er een $A_1 \in M^{m-r}_{m,m-r}$ zo dat $R(A_1) = N(A_2)$.

Bewijs.

Vertaling van (4.2.6) en (4.2.7). □

Gevolgen. (5.4.19)

1. Als $A \in M^r_{m,n}$ dan is $\dim(N(A)) = n-r$.
2. Als $V \subset \mathcal{L}^m$ met dimensie m-r dan is er een $A \in M^r_{r,m}$ (dus $A \in LI$) zo dat $V = N(A)$.
3. Als B en U elkaars inverse zijn en

$$B = (B_1 \mid B_2) \;,\; U = \left(\begin{array}{c} U_1 \\ \hline U_2 \end{array}\right)$$

met $B_1 \in M_{m,r}$, $U_1 \in M_{r,m}$ $(1 \le r \le m)$ dan is $R(B_1) = N(U_2)$, $R(B_2) = N(U_1)$.
Hieruit volgt dat $N(U_1) \oplus N(U_2) = \mathcal{L}^m$.

Stelling. (5.4.20)

Als A_1 en A_2 beide breedte m hebben dan geldt

$$N(A_1) \subset N(A_2) \leftrightarrow \exists_Y : A_2 = YA_1 .$$

(Vergelijk dit met (5.4.5)!)

Bewijs.

$N(A_1) \subset N(A_2)$ is equivalent met $\forall_X : A_1X = 0 \Rightarrow A_2X = 0$.
Volgens (4.2.8) is dit equivalent met $\exists_Y : A_2 = YA_1$.

Gevolgen. (vgl. (5.4.6)) (5.4.21)

1. $N(A_1) \subset N(A_2) \Rightarrow r(A_1) \ge r(A_2)$ met gelijkteken dan en slechts dan als $N(A_1) = N(A_2)$.
2. Als $A_1 \in LR$ en $A_2 = YA_1$ dan geldt $N(A_2) = N(A_1) \leftrightarrow Y \in LI$.
3. Als $A = B_1C_1$ met $B_1 \in RR$, $C_1 \in LR$ dan is $N(A) = N(C_1)$.

Opgave. (5.4.22)

De vergelijking $A_1XA_2 = Z$ is oplosbaar dan en slechts dan als $R(Z) \subset R(A_1)$, $N(Z) \supset N(A_2)$. Bewijs dit met behulp van (5.4.5) en (5.4.20). Vergelijk formulering en bewijs met (4.3.9).

Stelling. (5.4.23)

Als A_1 hoogte m en A_2 breedte m heeft dan geldt
1. $R(A_1) \cap N(A_2) = \{\underline{0}\} \leftrightarrow N(A_2A_1) = N(A_1) \leftrightarrow r(A_2A_1) = r(A_1)$.
2. $R(A_1) + N(A_2) = \mathcal{L}^m \leftrightarrow R(A_2A_1) = R(A_2) \leftrightarrow r(A_2A_1) = r(A_2)$.
3. $R(A_1) \oplus N(A_2) = \mathcal{L}^m \leftrightarrow r(A_2A_1) = r(A_2) = r(A_1)$.

Bewijs.

1. $R(A_1) \cap N(A_2) = \{\underline{0}\} \leftrightarrow \forall_x : (A_2A_1x = 0 \Rightarrow A_1x = 0) \leftrightarrow N(A_2A_1) = N(A_1) \leftrightarrow$
$\leftrightarrow r(A_2A_1) = r(A_1)$ (volgens (5.4.21.1)).
2. $R(A_1) + N(A_2) = \mathcal{L}^m \leftrightarrow \forall_y \exists_x : A_2(y - A_1x) = 0 \leftrightarrow R(A_2A_1) = R(A_2) \leftrightarrow$
$\leftrightarrow r(A_2A_1) = r(A_2)$ (volgens (5.4.6.1)).
3. Volgt uit 1. en 2. □

Opmerking. (5.4.24)

Als $A_1 \in RR$ dan is 1. ook equivalent met $A_2A_1 \in RR$.
Als $A_2 \in LR$ dan is 2. ook equivalent met $A_2A_1 \in LR$.
Als $A_1 \in RR$ en $A_2 \in LR$ dan is 3. ook equivalent met $A_2A_1 \in I$.

Zij nu V_1 en $V_2 \subset \mathcal{L}^m$, voorgesteld als kolomruimte of als nulruimte van matrices A_1, resp. A_2. We gaan $V_1 + V_2$ en $V_1 \cap V_2$ beschrijven en hun dimensies bepalen.

Stelling. (5.4.25)

Als A_1 en A_2 beide hoogte m hebben dan geldt:
1. $R(A_1) + R(A_2) = R((A_1 \mid A_2))$.
2. als A_1 en $A_2 \in RR$ dan is $R(A_1) \cap R(A_2)$ isomorf met $N((A_1 \mid A_2))$.
Als A_1 en A_2 beide breedte m hebben dan geldt:
3. $N(A_1) \cap N(A_2) = N((\frac{A_1}{A_2}))$.

Als A_1 hoogte m en A_2 breedte m heeft dan geldt:
4. als $A_1 \in RR$ dan is $R(A_1) \cap N(A_2)$ isomorf met $N(A_2A_1)$.

Bewijs.

De uitspraken 1. en 3. zijn triviaal.
De isomorfie in uitspraak 2. construeren we als volgt. Aan $\underline{x} \in N((A_1 \mid A_2))$ voegen we toe $\underline{y} \in R(A_1) \cap R(A_2)$, bepaald door $y = (A_1 \mid 0)x = A_1x_1 = -A_2x_2$. Deze toevoeging is lineair, injectief (want $\underline{y} = \underline{0} \Rightarrow x_1 = 0$ en $x_2 = 0$, daar A_1 en $A_2 \in RR$) en surjectief (want $\underline{y} \in R(A_1) \cap R(A_2)$ impliceert existentie van x_1 en x_2 zo dat $y = A_1x_1 = -A_2x_2 = (A_1 \mid 0)x$ met $\underline{x} \in N((A_1 \mid A_2))$.
De isomorfie in uitspraak 4. construeren we door aan $\underline{x} \in N(A_2A_1)$ toe te voegen $\underline{y} \in R(A_1) \cap N(A_2)$, bepaald door $y = A_1x$. Deze toevoeging is lineair, injectief (omdat $A_1 \in RR$) en surjectief (want $\underline{y} \in R(A_1) \cap N(A_2)$ impliceert existentie van een $\underline{x} \in N(A_2A_1)$ zo dat $y = A_1x$). □

Stelling. (5.4.26)

Als V_1 en $V_2 \subset \mathcal{L}^m$ dan is

$$\dim(V_1 + V_2) + \dim(V_1 \cap V_2) = \dim(V_1) + \dim(V_2). \qquad (*)$$

Bewijs.

Neem A_1 en $A_2 \in RR$ zo dat $R(A_1) = V_1$, $R(A_2) = V_2$. Dan is $V_1 + V_2 = R((A_1 \mid A_2))$, $V_1 \cap V_2$ isomorf met $N((A_1 \mid A_2))$ zodat volgens (5.4.25) 1. en 2., het linkerlid van (*) gelijk is aan de breedte van $(A_1 \mid A_2)$. □

Stelling. (5.4.27)

1. Als A_1 en A_2 hoogte m hebben dan is

$$\dim(R(A_1) + R(A_2)) = r((A_1 \mid A_2)) ,$$

$$\dim(R(A_1) \cap R(A_2)) = r(A_1) + r(A_2) - r((A_1 \mid A_2)) .$$

2. Als A_1 en A_2 breedte m hebben dan is

$$\dim(N(A_1) \cap N(A_2)) = m - r(\left(\frac{A_1}{A_2}\right)) ,$$

$$\dim(N(A_1) + N(A_2)) = m - r(A_1) - r(A_2) + r(\left(\frac{A_1}{A_2}\right)) .$$

3. Als A_1 hoogte m en A_2 breedte m heeft dan is

$$\dim(R(A_1) \cap N(A_2)) = r(A_1) - r(A_2A_1) ,$$

$$\dim(R(A_1) + N(A_2)) = m - r(A_2) + r(A_2A_1) .$$

Bewijs.

In 1. en 2. volgt de eerste relatie rechtstreeks uit (5.4.25). In 3. volgt de eerste relatie ook uit (5.4.25) want als $A_1 = B_1C_1$ met $B_1 \in RR$, $C_1 \in LR$ dan is

$$\dim(R(A_1) \cap N(A_2)) = \dim(R(B_1) \cap N(A_2)) =$$

$$= \dim(N(A_2B_1)) = \text{breedte}(A_2B_1) - r(A_2B_1) = r(A_1) - r(A_2A_1) .$$

In alle gevallen volgt de tweede relatie uit de eerste met behulp van (5.4.26). □

Opgave. (5.4.28)

Bewijs (5.4.26) door met (5.4.15) deelruimten V_1' en V_2' te construeren zo dat

$$(V_1 \cap V_2) \oplus V_1' = V_1, \quad (V_1 \cap V_2) \oplus V_2' = V_2 \quad .$$

en te bewijzen dat

$$(V_1 \cap V_2) \oplus V_1' \oplus V_2' = V_1 + V_2 \ .$$

Opgave. (5.4.29)

Als V_1 en $V_2 \subset \mathcal{L}^m$, met dimensies d_1 en d_2 dan volgt uit (5.4.26) en triviale ongelijkheden dat

$$\max(d_1,d_2) \le \dim(V_1 + V_2) \le \min(d_1 + d_2,m) \ ,$$

$$\max(0,d_1+d_2-m) \le \dim(V_1 \cap V_2) \le \min(d_1,d_2) \ .$$

Bewijs nu (4.1.5) uit (5.4.27.3).

Opgave. (5.4.30)

Bewijs dat, als A_1 en A_2 hoogte m hebben en A_1^- een g-inverse van A_1 is,

$$R(A_1) \cap R((I - A_1A_1^-)A_2) = \{\underline{0}\} \ .$$

Analoog, als A_1 en A_2 breedte m hebben

$$N(A_1) + N(A_2(\ I - A_1^-A_1)) = \mathcal{L}^m \ .$$

Leid hieruit (4.3.22) af.

Opgave. (5.4.31)

Zij $A = (A_1 \mid A_2) \in RR$ en $A_2' = A_1X_{12} + A_2X_{22}$. Dan geldt

. $R(A_1) \cap R(A_2') = \{\underline{0}\} \Leftrightarrow N(X_{22}) \subset N(X_{12}) \Leftrightarrow X_{12} = W_{12}X_{22}$.

. $R(A_1) + R(A_2') = R(A) \Leftrightarrow X_{22} \in RI$.

Vergelijk dit met (5.4.17).

Opgave. (5.4.32)

Zij A_1 en $A_2 \in M_{m,n}$, $A := A_1 + A_2$ met rangen r_1,r_2,r. Dan zijn de volgende uitspraken equivalent:

i) $r_1 + r_2 = r$,

ii) $R(A_1) + R(A_2) = R(A)$, $R(A_1) \cap R(A_2) = \{\underline{0}\}$,

iii) $N(A_1) \cap N(A_2) = N(A)$, $N(A_1) + N(A_2) = \mathcal{L}^n$.

Dit is een aanvulling op (4.3.19).

5.5. Matrices en lineaire afbeeldingen

We weten al dat $M_{m,n}$ isomorf is met $LA(\mathcal{L}^m,\mathcal{L}^n)$. We beschrijven nu eigenschappen van matrices in termen die horen bij de bijbehorende lineaire afbeeldingen, zoals beeldruimte, nulruimte. Belangrijke objecten zijn de zg. projectoren en de gegeneraliseerde inversen.

Definitie. (5.5.1)

$E \in M_{m,m}$ heet projector als $E^2 = E$.

Stelling. (5.5.2)

Zij $E \in M_{m,m}$. De volgende beweringen zijn equivalent:

i) E is projector,

ii) $I - E$ is projector,

iii) $R(E) = N(I - E)$,

iv) $N(E) = R(I - E)$,

v) $r(E) + r(I - E) = m$,

vi) $\forall_X : R(X) \subset R(E) \Leftrightarrow X = EX$,

vii) $\forall_Y : N(Y) \supset N(E) \Leftrightarrow Y = YE$,

viii) $E = 0$ of $E = B_1C_1$ met $C_1B_1 = I_{11}$ (dus $B_1 \in RR$, $C_1 \in LR$, C_1 linkerinverse van B_1, B_1 rechterinverse van C_1, $R(E) = R(B_1)$, $N(E) = N(C_1)$).

Bewijs.

De meeste van deze beweringen volgen eenvoudig uit elkaar (ga na).
Dat v) $\Rightarrow$ iv) zien we als volgt. Voor iedere $E \in M_{m,m}$ geldt $N(E) \subset R(I - E)$.
Uit v) volgt echter $\dim(N(E)) = m - r(E) = r(I - E) = \dim(R(I - E))$. Dus (met (5.4.8.4)) $N(E) = R(I - E)$.
Dat i) $\Leftrightarrow$ viii) zien we door substitutie van $E = B_1C_1$ met $B_1 \in RR$, $C_1 \in LR$ in $E^2 = E$. □

Stelling. (5.5.3)

Als E en $E' \in M_{m,m}$ projectoren zijn dan geldt

1) $R(E) \cap N(E) = \{\underline{0}\}$, $R(E) + N(E) = \mathcal{L}^m$.

2) $R(E') \subset R(E)$ en $N(E') \subset N(E) \Rightarrow E' = E$.

Bewijs.

De uitspraken 1) volgen uit de relaties $E^2 = E$ en $E + (I - E) = I$ (ga na).
De uitspraak 2) volgt uit (5.5.2.vi) en (5.5.2.vii):
$R(E') \subset R(E) \Rightarrow E' = EE'$ en $N(E') \subset N(E) \Rightarrow E = EE'$.

Gevolg. (5.5.4)

Als $E \in M_{m,m}$ projector is dan is

$$R(E) \oplus N(E) = \mathcal{L}^m$$

en E is ook de enige projector met deze range en nulruimte.
Daar $x \in R(E) \leftrightarrow Ex = x$ en voor $x \notin R(E)$ geldt $x - Ex \in N(E)$, noemt men E de projector op R(E) langs N(E).
Notatie:

$$E = E_{V_1,V_2}$$

is de projector met $R(E) = V_1$, $N(E) = V_2$.
Uit de volgende stelling volgt dat bij iedere V_1 en $V_2 \in \mathcal{L}^m$ de voorwaarde $V_1 \oplus V_2 = \mathcal{L}^m$ ook voldoende is voor het bestaan van de projector E_{V_1,V_2}.

Stelling. (5.5.5)

Als $V_1 \oplus V_2 = \mathcal{L}^m$ dan is er een eenduidig bepaalde projector E zo dat $R(E) = V_1$, $N(E) = V_2$.

Bewijs.

We geven een meetkundig en een meer algebraïsch bewijs van deze stelling.

a) Bij iedere $\underline{x} \in \mathcal{L}^m$ horen volgens (5.1.13) eenduidige $\underline{x}_1 \in V_1$ en $\underline{x}_2 \in V_2$ zo dat $\underline{x} = \underline{x}_1 + \underline{x}_2$. De afbeelding $\mathcal{E} : \mathcal{L}^m \to V_1$ gegeven door $\mathcal{E}\underline{x} = \underline{x}_1$ is lineair (ga na), voldoet aan $R(\mathcal{E}) = V_1$, $N(\mathcal{E}) = V_2$ en aan $\mathcal{E}^2 = \mathcal{E}$. De bij $\mathcal{E}$ horende matrix E is dus de gevraagde projector.

b) Uit (5.4.12) volgt dat er een $B = (B_1 \mid B_2) \in I$ is zo dat $R(B_i) = V_i$ $(i = 1,2)$. Zij E de gevraagde projector. De eis $V_1 \subset R(E)$ is volgens (5.5.2.vi) equivalent met $EB_1 = B_1$. De eis $V_2 \subset N(E)$ is equivalent met $EB_2 = 0$. E moet dus voldoen aan $E(B_1 \mid B_2) = (B_1 \mid 0) = BJ$ als J de normaalvorm uit $M_{m,m}$ met rang gelijk aan $\dim(V_1)$ is. Dus moet $E = BJB^{-1}$ en dit is inderdaad een projector met $R(E) = R(B_1)$, $N(E) = R(B_2)$ (ga na). De eenduidigheid volgt uit (5.5.3). □

<u>Gevolgen.</u> (5.5.6)

1. $E \in M_{m,m}$ is projector dan en slechts dan als

$$E = BJB^{-1}$$

met B regulier en J een normaalvorm uit $M_{m,m}$.

2. Als $B = (B_1 \mid B_2) \in I$, $B^{-1} = U = \left(\frac{U_1}{U_2}\right)$ met $B_1 \in M_{m,r}$, $U_1 \in M_{r,m}$, dan is

$$B_1U_1 = E_{R(B_1),N(U_1)} = E_{R(B_1),R(B_2)} = E_{N(U_2),N(U_1)} .$$

3. Voor het spoor van een projector geldt

$$tr(E) = r(E) \cdot 1$$

waarin 1 het eenheidselement van $\mathcal{L}$ is (dit volgt met (1.5.2.3) uit (5.5.6.1) of (5.5.2.viii)).

<u>Opgave.</u> (5.5.7)

Bewijs de volgende generalisatie van (5.5.5).
Zij $V_1,\dots,V_k \subset \mathcal{L}^m$ onafhankelijk, $V := V_1 \oplus \dots \oplus V_k$. Dan zijn er bij iedere V' zo dat $V \oplus V' = \mathcal{L}^m$ eenduidig bepaalde $E_1,\dots,E_k$ zo dat
1. E_i is projector op V_i $(i = 1,\dots,k)$,
2. $E_iE_j = 0$, $i \neq j$,
3. $\Sigma\, E_i$ is projector op V langs V'.
Wat is $N(E_i)$? Bekijk ook het geval dat $V = \mathcal{L}^m$.

<u>Opgave.</u> (5.5.8)

Zij $E_1,\dots,E_k \in M_{m,m}$. Beschouw de volgende uitspraken:
i) E_i is projector $(i = 1,\dots,k)$,
ii) $E_iE_j = 0$, $i \neq j$,
iii) $\Sigma\, E_i$ is projector,
iv) $\Sigma\, r(E_i) = r(\Sigma\, E_i)$.
Bewijs:
a) als i) en ii) gelden dan zijn $R(E_1),\dots,R(E_k)$ onafhankelijk en $\Sigma\, E_i$ is projector op $R(E_1) \oplus \dots \oplus R(E_k)$,
b) i) $\wedge$ ii) $\leftrightarrow$ iii) $\wedge$ iv) (hint: stel $E_i = B_iC_i$, $B_i \in RR$, $C_i \in LR$, $B = (B_1 \mid \dots \mid B_k)$, $C = (C_1^T \mid \dots \mid C_k^T)^T$, druk i) $\wedge$ ii), iii) en iv) uit in eigenschappen van B en C),

c) als de karakteristiek van $\mathcal{L}$ nul is (dit betekent dat $n \in \mathbf{Z}$, $\alpha \in \mathcal{L}$, $n\alpha = 0 \Rightarrow n = 0$ of $\alpha = 0$) dan i) $\wedge$ iii) $\Rightarrow$ ii) (hint: gebruik (5.5.6.2) en (1.5.2.1)),

d) als $k = 2$ en de karakteristiek van $\mathcal{L} \neq 2$ (dit betekent dat $\alpha \in \mathcal{L}$, $\alpha + \alpha = 0 \Rightarrow \alpha = 0$) dan i) $\wedge$ iii) $\Rightarrow$ ii),

e) als $\Sigma\, E_i = I$ dan ii) $\Rightarrow$ i).

We gaan nu het verband tussen een projector en representaties van zijn range en nulruimte verder na.

Stelling. (5.5.9)

Zij $A \in M_{m,n}$. Dan geldt:

$E_1 \in M_{m,m}$ is projector op $R(A) \Leftrightarrow \exists_{A^-} : A^-$ is g-inverse van A en $AA^- = E_1$;
$E_2 \in M_{n,n}$ is projector langs $N(A) \Leftrightarrow \exists_{A^-} : A^-$ is g-inverse van A en $A^-A = E_2$;
als A^- een rg-inverse van A is dan is

$$AA^- = E_{R(A),N(A^-)}\ , \quad A^-A = E_{R(A^-),N(A)}\ .$$

Bewijs.

Als E_1 projector is en $R(A) \subset R(E_1)$ dan is volgens (5.5.2.vi) $A = E_1A$; als tevens $R(E_1) \subset R(A)$ dan is er volgens (5.4.5) een X zo dat $E_1 = AX$ en dus $A = AXA$. Omgekeerd, als $AXA = A$ en $E_1 = AX$, dan is $E_1A = A$ en daaruit volgt $E_1^2 = E_1$ en $R(E_1) = R(A)$.
Analoog voor E_2 (gebruik (5.5.2.vii) en (5.4.20)).
De laatste bewering volgt uit de eerste twee met behulp van de reflexiviteit. □

Opmerkingen. (5.5.10)

1. Als A^- g-inverse van A is dan geldt

$$R(B) \subset R(A) \Leftrightarrow AA^-B = B$$
$$N(C) \supset N(A) \Leftrightarrow CA^-A = C\ .$$

2. Ga na welke bijzonderheden optreden als $r(A) = m$ of $r(A) = n$.
3. Als E_1 projector op $R(A)$ is dan is er zelfs een rg-inverse zo dat $AA^- = E_1$, want als A^- g-inverse is zodat $AA^- = E_1$, dan is ook $AA^-AA^- = E_1$ en volgens (4.3.8) is A^-AA^- rg-inverse van A. Analoog voor E_2.
4. Uit de volgende stelling blijkt dat we bij gegeven E_1 en E_2 zelfs met dezelfde A^- kunnen werken.

Stelling. (5.5.11)

Zij $A \in M_{m,n}$, E_1 een projector op $R(A)$ en E_2 een projector langs $N(A)$. Dan is er een eenduidig bepaalde rg-inverse A^- van A zo dat

$$AA^- = E_1 \ , \quad A^-A = E_2 \ .$$

Deze A^- heeft verder als eigenschappen:

a) als X en Y zo zijn dat $AX = E_1$, $YA = E_2$, dan is

$$YAX = A^- \ ,$$

b) voor iedere g-inverse Z van A geldt

$$E_2ZE_1 = A^- \ ,$$

c) als voor een g-inverse Z van A geldt

$$R(Z) \subset R(E_2) \ , \quad N(Z) \supset N(E_1) \ ,$$

dan is $Z = A^-$.

Bewijs.

Volgens (5.5.9) zijn er X en Y zo dat $AXA = A$, $AYA = A$ en $AX = E_1$, $YA = E_2$. Als A^- voldoet dan moet dus $AA^- = AX$, $A^-A = YA$ en $A^-AA^- = A^-$. Daaruit volgt $A^- = A^-AA^- = YAA^- = YAX$. Omgekeerd volgt door verificatie dat $A^- := YAX$ aan de eisen voldoet. De eenduidigheid volgt uit het feit dat $A^- = YAX = E_2X = YE_1$, zodat A^- alleen van E_1 en E_2 en niet van de specifieke keus van X en Y afhangt.
Als Z voldoet aan $AZA = A$ dan is $E_2ZE_1 = A^-AZAA^- = A^-AA^- = A^-$.
Als Z voldoet aan $AZA = A$, $R(Z) \subset R(E_2)$, $N(Z) \supset N(E_1)$ dan is $Z = E_2Z = ZE_1$, dus $Z = E_2ZE_1 = A^-$. □

Zij nu deelruimten V_1 en $V_2 \subset \mathcal{C}^m$ gegeven door $V_1 = R(A_1)$, $V_2 = N(A_2)$ met $A_1 \in M_{m,q}$, $A_2 \in M_{p,n}$. Zij $V_1 \oplus V_2 = \mathcal{C}^m$. Volgens (5.5.5) is dan $E := E_{V_1,V_2}$ eenduidig bepaald. Kunnen we E in A_1 en A_2 uitdrukken?
Zeker is $E = A_1A_1^- = A_2^-A_2$ met passende g-inversen A_2^- en A_1^-. Hieruit volgt dat $E = E^2 = A_1A_1^-A_2^-A_2 = A_1ZA_2$ en $A_2A_1 = A_2EA_1 = A_2A_1ZA_2A_1$, zo dat Z een g-inverse van A_2A_1 is. E heeft dus de vorm

$$E = A_1(A_2A_1)^-A_2 \qquad (*)$$

met $(A_2A_1)^-$ een g-inverse van A_2A_1. We bewijzen nu dát, voor iedere g-inverse $(A_2A_1)^-$, het rechterlid van (*) de (unieke) projector op $R(A_1)$ langs $N(A_2)$ is.

<u>Stelling</u>. (5.5.12)

Zij $A_1 \in M_{m,q}$, $A_2 \in M_{p,m}$ en

$$R(A_1) \oplus N(A_2) = \mathcal{C}^m .$$

Zij $(A_2A_1)^-$ een g-inverse van A_2A_1 en

$$X_1 := (A_2A_1)^- A_2, \; X_2 := A_1(A_2A_1)^- ,$$

$$E := A_1(A_2A_1)^- A_2 = A_1X_1 = X_2A_2 .$$

Dan geldt

$$E = E_{R(A_1),N(A_2)} ,$$

X_1 is rg-inverse van A_1 met $N(X_1) = N(A_2)$,

X_2 is rg-inverse van A_2 met $R(X_2) = R(A_1)$.

<u>Bewijs</u>.

Uit $A_2A_1(A_2A_1)^-A_2A_1 = A_2A_1$ volgt $A_2(I - E)A_1 = 0$. Hieruit volgt
a) $R((I - E)A_1) \subset N(A_2)$ en dus, omdat $R(E) \subset R(A_1)$, ook
$R((I - E)A_1) \subset R(A_1) \cap N(A_2) = \{\underline{0}\}$, dus $(I - E)A_1 = 0$;
b) $N(A_2(I - E)) \supset R(A_1)$ en dus, omdat $N(E) \supset N(A_2)$, ook
$N(A_2(I - E)) \supset R(A_1) + N(A_2) = \mathcal{C}^m$, dus $A_2(I - E) = 0$.
Deze resultaten impliceren

$$A_1 = EA_1 = A_1X_1A_1 , \quad A_2 = A_2E = A_2X_2A_2 . \qquad (*)$$

Hieruit volgt al dat X_1 en X_2 g-inversen van A_1, resp. A_2 zijn en (met (5.5.9)) dat E de projector is op $R(A_1)$ langs $N(A_2)$.
Ook volgt uit (*) en de definities van E, X_1 en X_2 dat
$R(A_1) = R(E) \subset R(X_2) \subset R(A_1)$ dus $R(X_2) = R(A_1) = R(E)$,
$N(A_2) = N(E) \supset N(X_1) \supset N(A_2)$ dus $N(X_1) = N(A_2) = N(E)$.
Hieruit volgt met (5.5.2) $X_1 = X_1E = X_1A_1X_1$, $X_2 = EX_2 = X_2EX_2$, zodat X_1 en X_2 rg-inversen van A_1, resp. A_2 zijn. □

Opmerkingen. (5.5.13)

1. Dit bewijs is tevens een derde bewijs van (5.5.5).
2. Als $A_1 \in RR$ en $A_2 \in LR$ dan is volgens (5.4.24) A_2A_1 regulier en dus geldt dan

$$E_{R(A_1),N(A_2)} = A_1(A_2A_1)^{-1}A_2 \ , \quad X_1 = (A_2A_1)^{-1}A_2 \ , \ X_2 = A_1(A_2A_1)^{-1} \ .$$

Verifieer rechtstreeks dat deze E, X_1 en X_2 de genoemde eigenschappen hebben. Merk op dat (5.5.6.2) een speciaal geval hiervan is.

Opgave. (5.5.14)

Bewijs de equivalentie van de volgende beweringen:

i) $R(A_1) \cap N(A_2) = \{\underline{0}\}$,
ii) $N(A_2A_1) = N(A_1)$,
iii) $r(A_2A_1) = r(A_1)$,
iv) $\exists_Y : A_1 = YA_2A_1$,
v) $(A_2A_1)^-A_2$ is g-inverse van A_1,
vi) $A_1(A_2A_1)^-A_2$ is projector op $R(A_1)$,

en ook van de beweringen

i) $R(A_1 + N(A_2) = \mathcal{L}^m$,
ii) $R(A_2A_1) = R(A_2)$,
iii) $r(A_2A_1) = r(A_2)$,
iv) $\exists_X : A_2 = A_2A_1X$,
v) $A_1(A_2A_1)^-$ is g-inverse van A_2,
vi) $A_1(A_2A_1)^-A_2$ is projector langs $N(A_2)$,

(zie ook (5.4.23)). Leid hieruit (5.5.12) af.

Zij nu A en A^- rg-inversen van elkaar. We bekijken de relaties tussen de ranges en de nulruimten van A en A^-.

Stelling. (5.5.15)

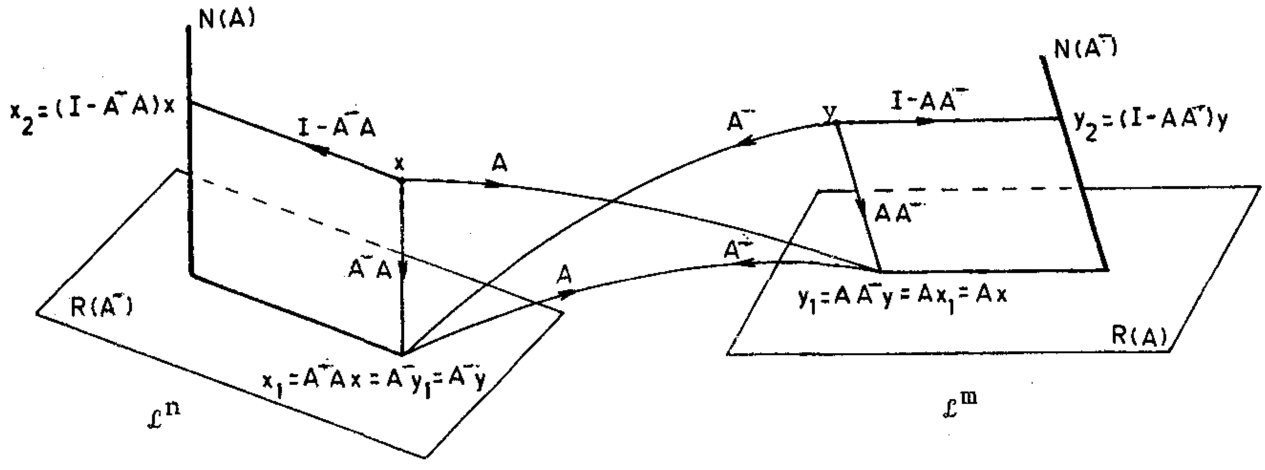
Zij $A \in M_{m,n}$ en $A^- \in M_{n,m}$ rg-inversen van elkaar. Dan geldt

1. $R(A) \oplus N(A^-) = \mathcal{L}^m$, $R(A^-) \oplus N(A) = \mathcal{L}^n$.
2. $AA^- = E_{R(A),N(A^-)}$, $A^-A = E_{R(A^-),N(A)}$.
3. De afbeeldingen $A : R(A^-) \to R(A)$ en $A^- : R(A) \to R(A^-)$ zijn bijecties en elkaars inversen.

Bewijs.

De beweringen 2, zijn al in (5.5.9) bewezen. De beweringen 1. volgen dááruit met (5.5.4).
Tenslotte: als $x = A^-u \in R(A^-)$, $y = Av \in R(A)$ dan geldt $y = Ax \leftrightarrow Av = AA^-u \leftrightarrow A^-Av = A^-u \leftrightarrow A^-y = x$ en daaruit volgt met (5.2.6) dat de door A voortgebrachte afbeelding van $R(A^-)$ in $R(A)$ een bijectie is met als inverse de door A^- voortgebrachte afbeelding van $R(A)$ in $R(A^-)$. □

In onderstaande tekening is gepoogd de diverse verzamelingen en de afbeeldingen daartussen weer te geven. De punten $\underline{x} \in \mathcal{L}^n$ en $\underline{y} \in \mathcal{L}^m$ zijn zo gekozen dat voor $x_1 := A^-Ax \in R(A^-)$ en $y_1 := AA^-y \in R(A)$ geldt $y_1 = Ax_1$ en $x_1 = A^-y_1$.



$\dim(R(A^-)) = r,$	$\dim(R(A)) = r,$
$\dim(N(A)) = n - r,$	$\dim(N(A^-)) = m - r,$
$R(A^-) \oplus N(A) = \mathcal{L}^n,$	$R(A) \oplus N(A^-) = \mathcal{L}^m.$

Opgave. (5.5.16)

Zij $A \in M_{m,n}$ en A^- een g-inverse van A. Dan geldt

$$R(A) \cap N(A^-) = \{\underline{0}\}\ , \quad R(A) + N(A^-) = N(A^- - A^-AA^-)\ ,$$
$$R(A^-) \cap N(A) = R(A^- - A^-AA^-)\ , \quad R(A^-) + N(A) = \mathcal{L}^n\ .$$

Vergelijk dit met (4.3.21).

Het blijkt dat de eerste bewering van (5.5.15) de verzameling der rg-inversen van A ook karakteriseert in die zin dat $N(A^-)$ en $R(A^-)$ vrij gekozen mogen worden, mits de genoemde relaties met $R(A)$ en $N(A)$ gelden en dat daardoor A^- eenduidig bepaald is.

Stelling. (5.5.17)

Zij $A \in M_{m,n}$. Bij iedere $V_1 \subset \mathcal{L}^m$ en $V_2 \subset \mathcal{L}^n$ zo dat

$$R(A) \oplus V_1 = \mathcal{L}^m , \quad V_2 \oplus N(A) = \mathcal{L}^n ,$$

is er een eenduidig bepaalde g-inverse A^- van A zo dat

$$R(A^-) = V_2 , \quad N(A^-) = V_1 .$$

Deze A^- is rg-inverse van A en voldoet aan

$$AA^- = E_{R(A),V_1} , \quad A^-A = E_{V_2,N(A)} .$$

Bewijs.

Volgens (5.5.5) bestaan eenduidig bepaalde projectoren E_1 en E_2 zo dat $E_1 = E_{R(A),V_1}$ en $E_2 = E_{V_2,N(A)}$. Volgens (5.5.11) is er daarbij een eenduidig bepaalde rg-inverse A^- van A zo dat $AA^- = E_1$, $A^-A = E_2$. Volgens (5.5.9) is $R(A^-) = R(E_2) = V_2$, $N(A^-) = N(E_1) = V_1$ en volgens (5.5.11), deel c), is A^- ook de enige g-inverse van A waarvoor deze relaties gelden. $\square$

In de opgaven (5.5.18) t/m (5.5.20) worden enige min of meer expliciete methoden aangegeven om bij bepaalde representaties van V_1 en V_2 de A^- uit (5.5.17) te bepalen.

Opgave. (5.5.18)

Zij $A \in M_{m,n}$, $U \in M_{p,m}$, $W \in M_{n,q}$ zo dat

$$R(A) \oplus N(U) = \mathcal{L}^m , \quad R(W) \oplus N(A) = \mathcal{L}^n .$$

Dan is (voor iedere g-inverse $(UAW)^-$ van UAW)

$$A^- = W(UAW)^-U$$

de rg-inverse van A met $R(A^-) = R(W)$, $N(A^-) = N(U)$.

Opgave. (5.5.19)

Zij $A \in M_{m,n}$. Zij $B_2 \in RR$ zo dat $R(A) \oplus R(B_2) = \mathcal{L}^m$ en $C_2 \in LR$ zo dat $N(C_2) \oplus N(A) = \mathcal{L}^n$.
Als $A = B_1C_1$ met $B_1 \in RR$, $C_1 \in LR$ dan zijn $B := (B_1 \mid B_2)$ en $C := (C_1^T \mid C_2^T)^T$ regulier. Zij $B^{-1} = U = (U_1^T \mid U_2^T)^T$ en $C^{-1} = W = (W_1 \mid W_2)$.
Dan is

. $A^- := W_1U_1$ de rg-inverse van A met $R(A^-) = N(C_2)$, $N(A^-) = R(B_2)$,
. $B_2^- := U_2$ de linker-inverse van B_2 met $N(B_2^-) = R(A)$,
. $C_2^- := W_2$ de rechter-inverse van C_2 met $R(C_2^-) = N(A)$.

Bewijs ook dat B_2 en C_2 aan de gegeven voorwaarden voldoen dan en slechts dan als

$$\left(\begin{array}{c|c} A & B_2 \\ \hline C_2 & 0 \end{array}\right)$$

regulier is en dat de inverse van deze matrix dan is

$$\left(\begin{array}{c|c} A^- & C_2^- \\ \hline B_2^- & 0 \end{array}\right).$$

Opgave. (5.5.20)

Ga na dat existentie en eenduidigheid van de A^- uit (5.5.17) ook door de volgende rechtstreekse constructie bewezen kunnen worden.
Zij W_1 en $W_2 \in RR$ zo dat $R(W_1) = V_2$, $R(W_2) = N(A)$. Dan is $W := (W_1 \mid W_2)$ regulier.
Zij $B_1 := AW_1$, dan is $B_1 \in RR$ en $R(B_1) = R(A)$. Zij $B_2 \in RR$ zo dat $R(B_2) = V_1$. Dan is $B := (B_1 \mid B_2)$ regulier en $A = BJW^{-1}$.
Uit $N(A^-) \supset V_1 = R(B_2)$ en $R(A^-) \subset V_2 = R(W_1)$ volgt nu dat $A^-B = (W_1Z_{11} \mid 0)$.
Als A^- g-inverse is dan moet $B_1 = AA^-B_1 = AW_1Z_{11} = B_1Z_{11}$, dus $Z_{11} = I_{11}$.
Als de gevraagde A^- bestaat dan is dus $A^- = WJ^TB^{-1}$. Maar deze voldoet ook.
Vergelijk het resultaat met (4.3.15).

We vragen nu nog wat de bij een $A \in M_{m,n}$ behorende afbeelding van $\mathcal{L}^n$ in $\mathcal{L}^m$ doet met een deelruimte $V \subset \mathcal{L}^n$. Het beeld van V noemen we A(V):

$$A(V) := \{\underline{y} \in \mathcal{L}^m \mid \exists_{\underline{x}\in V} : y = Ax\} .$$

Stelling. (5.5.21)

Zij $A \in M_{m,n}$, $V \subset \mathcal{L}^n$. Dan is

$$\dim(A(V)) = \dim(V) - \dim(V \cap N(A)) .$$

Bewijs.

Zij A_1 zo dat $V = R(A_1)$. Dan is $A(V) = R(AA_1)$ en de bewering is equivalent met

$$r(AA_1) = r(A_1) - \dim(R(A_1) \cap N(A)) ,$$

d.w.z., met (5.4.27.3). □

Gevolg. (5.5.22)

De afbeelding $A : V \to A(V)$ is een bijectie dan en slechts dan als $V \cap N(A) = \{\underline{0}\}$.

Opgave. (5.5.23)

Als A en A_1 breedte n hebben dan is

$$\dim(A(N(A_1))) = r(\left(\frac{A}{A_1}\right)) - r(A_1) .$$

Bewijs dit zowel met (5.5.21) als met (4.3.22).

Opgave. (5.5.24)

Bewijs (5.5.21) door voor $V \cap N(A)$ een basis te kiezen en deze aan te vullen tot basis voor V.

Opgave. (5.5.25)

Uit (5.5.21) volgt dat, als $V_1 \subset V$,

$$\dim(V_1) - \dim(A(V_1)) \le \dim(V) - \dim(A(V)) .$$

Leid hiermee de ongelijkheid van Frobenius (4.1.7) af.

Opgave. (5.5.26)

Uit het resultaat van (4.3.23) blijkt dat $r(A) = r(A_{11}) + r(A_{22})$ dan en slechts dan als

$$A_{21}(N(A_{11})) \subset R(A_{22}) .$$

Bewijs dit ook rechtstreeks door de nulruimte van A te bepalen.

Opgave. (5.5.27)

Zij $A \in M_{m,m}$. Bewijs dat de volgende uitspraken equivalent zijn:

i) $R(A) = R(A^2)$,

ii) $N(A) = N(A^2)$,

iii) $R(A) \cap N(A) = \{\underline{0}\}$,

iv) $R(A) + N(A) = \mathcal{C}^m$,

v) A heeft een rg-inverse A^- met $R(A^-) = R(A)$, $N(A^-) = N(A)$,

vi) A heeft een rg-inverse A^- waarvoor $AA^- = A^-A$.

Men noemt matrices met deze eigenschappen matrices met index 1.

Bewijs dat in dit geval A op de volgende manieren kan worden voorgesteld:

vii) er zijn B_1, B_1^- en Z_{11} zo dat $B_1^-B_1 = I_{11}$, $Z_{11} \in I$ en $A = B_1Z_{11}B_1^-$,

er geldt dan $A^- = B_1Z_{11}^{-1}B_1^-$.

viii) er zijn reguliere B en Z_{11} zo dat

$$A = B\left(\begin{array}{c|c} Z_{11} & \\ \hline & \end{array}\right)B^{-1}\ , \text{ er geldt dan } A^- = B\left(\begin{array}{c|c} Z_{11}^{-1} & \\ \hline & \end{array}\right)B^{-1}\ .$$

De aldus voor een matrix met index 1 door v) eenduidig bepaalde rg-inverse A^- wordt de groepsinverse van A genoemd om de volgende reden.

Definieer voor $k \in \mathbb{Z}$

$$\varphi_k := \begin{cases} A^k & \text{als } k > 0 \\ AA^- = A^-A & \text{als } k = 0 \\ (A^-)^{|k|} & \text{als } k < 0\ . \end{cases}$$

Dan geldt (ga na) voor alle k en $m \in \mathbb{Z}$ dat $\varphi_k\varphi_m = \varphi_{k+m}$. Hieruit volgt dat de verzameling $\{\varphi_k\}$ een commutatieve groep is (met de matrix-vermenigvuldiging als groepsoperatie) met φ_0 als neutraal element.

Ga na dat uit vii) volgt $\varphi_k = B_1Z_{11}^kB_1^-$ (alle $k \in \mathbb{Z}$).

Bewijs tenslotte met een 2×2-voorbeeld dat niet alle vierkante matrices index 1 hebben.

Opgave. (5.5.28)

Zij $A \in M_{m,m}$ met rg-inverse A^-, $E_1 := AA^-$, $E_2 := A^-A$.

Dan is er een reguliere Z zo dat

$$A = E_1ZE_2\ ,\quad A^- = E_2Z^{-1}E_1\ .$$

5.6. Een alternatieve opzet

We hebben in het voorgaande het begin van de theorie van de eindig-dimensionale lineaire ruimten afgeleid met behulp van matrixtheorie en met name met het enige niet-triviale stuk gereedschap dat we daar ontwikkeld hebben: de splitsings- en uitbreidingsstellingen.
Kan het ook omgekeerd? Het antwoord op deze vraag is bevestigend.
Het essentiële gereedschap, waarmee men bij de opbouw van de theorie der eindig dimensionale lineaire ruimten de matrixtheorie kan omzeilen is de zogenaamde Austauschsatz van Steinitz (of een variant daarvan):

Stelling. (5.6.1)

Zij $1 \leq k \leq n$, $(\underline{a}_1,\ldots,\underline{a}_{k-1}, \underline{b}_k,\ldots,\underline{b}_n)$ een basis voor een vectorruimte V, $\underline{a}_k \in V$ en onafhankelijk van $\underline{a}_1,\ldots,\underline{a}_{k-1}$. Dan is er een permutatie p van $\{k,\ldots,n\}$ zodanig dat $(\underline{a}_1,\ldots,\underline{a}_k, \underline{b}_{p(k+1)},\ldots,\underline{b}_{p(n)})$ een basis voor V is.

Bewijs.

$\underline{a}_k \in V$, dus $\underline{a}_k = \sum\limits_1^{k-1} \alpha_i\underline{a}_i + \sum\limits_k^n \beta_i\underline{b}_i$. Omdat $\underline{a}_k \notin [\underline{a}_1,\ldots,\underline{a}_{k-1}]$ zijn niet alle β_i nul. Er is dus een permutatie p van $\{k,\ldots,n\}$ zodanig dat $\beta_{p(k)} \neq 0$ en dus

$$\underline{b}_{p(k)} = \sum_1^k \gamma_i\underline{a}_i + \sum_{k+1}^n \gamma_i\underline{b}_{p(i)} .$$

Hieruit volgt (ga na!) dat $\underline{a}_1,\ldots,\underline{a}_k, \underline{b}_{p(k+1)},\ldots,\underline{b}_{p(n)}$ V opspannen, maar ze zijn ook onafhankelijk omdat $\underline{a}_1,\ldots,\underline{a}_{k-1}, \underline{b}_{p(k+1)},\ldots,\underline{b}_{p(n)}$ onafhankelijk zijn en $\underline{a}_k$ niet in de door deze vectoren opgespannen deelruimte ligt. □

Uit deze stelling volgt (ga na) dat iedere basis van V evenveel elementen heeft, zodat de definitie van dimensie zinvol wordt. Vervolgens kan men met de Austauschsatz uitbreidingsstellingen zoals bijv. (5.4.14) bewijzen. Met deze uitbreidingsstelling volgt voor een lineaire afbeelding A van een n-dimensionale V_2 in een m-dimensionale V_1 (na de definitie van $N(A)$ en $R(A)$ en het bewijs dat dit lineaire deelruimten van V_2 resp. V_1 zijn) dat als $A \neq 0$ er een r met $1 \leq r \leq \min(m,n)$ en bases $(\underline{w}_1,\ldots,\underline{w}_n)$ voor V_2 en $(\underline{b}_1,\ldots,\underline{b}_m)$ voor V_1 zijn zodanig dat $(\underline{w}_{r+1},\ldots,\underline{w}_n)$ een basis voor $N(A)$ is, $(\underline{b}_1,\ldots,\underline{b}_r)$ een basis voor $R(A)$ is en

$$A\underline{w}_j = \begin{cases} \underline{b}_j & \text{als } j \le r\,, \\ \underline{0} & \text{als } j > r\,. \end{cases} \tag{5.6.2}$$

We kunnen nu hieruit de hoofdzaken van de matrixtheorie halen. Zij $A = (a_1 | \ldots | a_n) \in M_{m,n}$. Door vergelijking van definities volgt: $(\underline{a}_1,\ldots,\underline{a}_n)$ is een basis $\Leftrightarrow A \in RR$, $\underline{a}_1,\ldots,\underline{a}_n$ spannen $\mathcal{L}^m$ op $\Leftrightarrow A \in LI$. Met de Austausch-satz en de uitbreidingsstelling volgt nu (ga na): Als $A \in M_{m,n}$ en RR is dan is $n \le m$; als $n = m$ dan is $A \in RI$ en dus inverteerbaar en regulier (zie (2.1.7)); als $n < m$ is dan kan A worden uitgebreid tot een reguliere matrix $(A \mid A')$, deze is inverteerbaar, dus $A \in LI$ (ga na). Men name geldt dus: $(\underline{a}_1,\ldots,\underline{a}_n)$ is een basis voor $\mathcal{L}^m \Leftrightarrow m = n$ en A is inverteerbaar. Door naar de rijvectoren van A te kijken volgen analoog de "gespiegelde" resultaten.
Om de algemene splitingsstelling te krijgen moeten we A identificeren met een lineaire afbeelding van $\mathcal{L}^n$ in $\mathcal{L}^m$. Stelling (5.6.2) zegt dan: er is een $r \le \min(m,n)$, en er zijn inverteerbare matrices $V \in M_{m,n}$ en $B \in M_{m,m}$ zodanig dat

$$AW = A(W_1 \mid W_2) = (B_1 \mid 0) = (B_1 \mid B_2)\left(\begin{array}{c|c} I_{11} & 0 \\ \hline 0 & 0 \end{array}\right) = BJ \text{ met } I_{11} \in M_{r,r}\,.$$

Met $C := W^{-1}$ volgt nu $A = BJC$.

6. Algemene eindig dimensionale vectorruimten

6.1. Coördinatisering

Definitie. (6.1.1)

Zij V een eindig dimensionale vectorruimte. Zij $(\underline{u}_1,\dots,\underline{u}_m)$ een basis voor V. De bij een $\underline{u} \in V$ eenduidig behorende vector $\underline{x} = (\xi_1,\dots,\xi_m) \in \mathcal{L}^m$ zo dat

$$\underline{u} = \Sigma \xi_i u_i$$

heet coördinatenvector van $\underline{u}$ t.o.v. de basis $(\underline{u}_1,\dots,\underline{u}_m)$. De afbeelding $C \in LA(\mathcal{L}^m, V)$, gedefinieerd door $C\underline{u} = \underline{x}$ heet coördinatiserende afbeelding t.o.v. de basis $(\underline{u}_1,\dots,\underline{u}_m)$.

Stelling. (6.1.2)

De coördinatiserende afbeelding C t.o.v. een basis $(\underline{u}_1,\dots,\underline{u}_m)$ beeldt V lineair en bijectief af op $\mathcal{L}^m$, heeft dus een inverse.
C is geheel bepaald door de relaties

$$C\underline{u}_j = \underline{e}_j\ , \quad j = 1,\dots,m\ .$$

Stelling. (6.1.3)

Als V een basis van m elementen heeft dan is V isomorf met $\mathcal{L}^m$.

Stelling. (6.1.4)

Als V eindig dimensionaal is dan heeft iedere basis voor V evenveel elementen.

Definitie. (6.1.5)

De dimensie van eén eindig dimensionale vectorruimte is het aantal elementen in een basis.

Stelling. (6.1.6)

Eindig dimensionale vectorruimten zijn isomorf dan en slechts dan als ze dezelfde dimensie hebben.

<u>Definitie.</u> (6.1.7)

Stel V heeft bases $(\underline{u}_1,\ldots,\underline{u}_m)$ en $(\underline{u}'_1,\ldots,\underline{u}'_m)$. Dan zijn er getallen σ_{ij} zo dat

$$\underline{u}'_j = \sum_{i=1}^{m} \sigma_{ij}\underline{u}_i \ , \qquad j = 1,\ldots,m \ .$$

De matrix $S := (\sigma_{ij})$ heet overgangsmatrix van de basis $(\underline{u}_1,\ldots,\underline{u}_m)$ naar de basis $(\underline{u}'_1,\ldots,\underline{u}'_m)$.
<u>N.B.</u> De kolommen van S zijn de coördinaten van de basisvectoren $\underline{u}'_j$ t.o.v. de basis $(\underline{u}_1,\ldots,\underline{u}_m)$.

<u>Stelling.</u> (6.1.8)

S is regulier, S^{-1} is de overgangsmatrix van $(\underline{u}'_1,\ldots,\underline{u}'_m)$ naar $(\underline{u}_1,\ldots,\underline{u}_m)$.
Als C en C' de coördinatiserende afbeeldingen t.o.v. de genoemde bases zijn, dan is S de matrix van CC'^{-1} t.o.v. de natuurlijke basis in $\mathcal{L}^m$.
Als $\underline{u} \in V$ coördinatenvectoren $\underline{x}$, resp. $\underline{x}'$ t.o.v. de genoemde bases heeft dan geldt

$$x = Sx' \ . \tag{6.1.9}$$

<u>Bewijs.</u>

De eerste bewering is vrijwel triviaal.
De tweede bewering volgt met (5.3.5) uit

$$CC'^{-1}\underline{e}_j = C\underline{u}'_j = C(\sum_i \sigma_{ij}\underline{u}_i) = \sum_i \sigma_{ij}\underline{e}_i \ .$$

De derde bewering volgt met (5.3.9) uit

$$\underline{x} = C\underline{u} = CC'^{-1}\underline{x}'$$

of, alternatief, uit

$$\underline{u} = \sum_j \xi'_j\underline{u}'_j = \sum_{ij} \xi'_j\sigma_{ij}\underline{u}_i = \sum_i \xi_i\underline{u}_i$$

met $\xi_i = \sum_j \sigma_{ij}\xi'_j$. $\square$

Via de isomorfie en, zo nodig, coördinatisering zijn nu alle resultaten over $\mathcal{L}^m$ over te brengen naar m-dimensionale vectorruimten en hun deelruimten, b.v. (5.4.26) en (5.4.16) (dit is equivalent met de Austauschsatz).

Voorbeeld.

Zij V de vectorruimte van alle polynomen met graad < m. We beschouwen twee verschillende bases.
Basis 1 heeft als basiselementen de interpolatiepolynomen van Lagrange t.o.v. interpolatiepunten $x_1,\dots,x_m$ (met $x_i \neq x_j$ voor $i \neq j$):

$$\underline{\ell}_j : \ell_j(x) := \prod_{\substack{i=1\\ i\neq j}}^{m} \frac{x - x_i}{x_j - x_i}, \quad j = 1,\dots,m .$$

Deze vormen een basis voor V en voor $\underline{p} \in V$ geldt

$$p(x) = \sum_{j=1}^{m} p(x_j)\ell_j(x) ,$$

of

$$\underline{p} = \Sigma\beta_j\underline{\ell}_j , \quad \text{met } \beta_j = p(x_j) .$$

De afbeelding $C \in LA(\mathcal{L}^m,V)$ is dus gedefinieerd door

$$C\underline{p} = \underline{b} = (\beta_1,\dots,\beta_m)' .$$

De tweede basis heeft als elementen de machten van x:

$$\underline{p}_j : p_j(x) := x^{j-1} , \quad j = 1,\dots,m .$$

Voor $\underline{p} \in V$ geldt dan

$$p(x) = \sum_{j=1}^{m} \frac{p^{(j-1)}(0)}{(j-1)!}\, x^{j-1}$$

of

$$\underline{p} = \Sigma\gamma_j\underline{p}_j \quad \text{met } \gamma_j = \frac{p^{(j-1)}(0)}{(j-1)!} .$$

De afbeelding $C' \in LA(\mathcal{L}^m, V)$ wordt gedefinieerd door $C'\underline{p} = \underline{c} = (\gamma_1, \ldots, \gamma_m)$.
Wat is nu de overgangsmatrix S van de eerste naar de tweede basis? Volgens (6.1.9) geldt b = Sc en uit

$$\beta_i = p(x_i) = \Sigma \frac{p^{(j-1)}(0)}{(j-1)!} x_i^{j-1} = \Sigma\gamma_j x_i^{j-1}$$

volgt dus

$$S = \begin{pmatrix} 1 & x_1 & \ldots & x_1^{m-1} \\ 1 & x_2 & \ldots & x_2^{m-1} \\ . & .. & \ldots & \ldots. \\ 1 & x_m & \ldots & x_m^{m-1} \end{pmatrix} . \qquad (6.1.10)$$

Verifieer dit ook rechtstreeks aan de definitie (6.1.7)!

<u>Opgave.</u> (6.1.11)

Partitioneer en splits S uit (6.1.10) als

$$S = \left(\begin{array}{c|c} S_1 & b_1 \\ \hline d_1^T & \delta_1 \end{array}\right) = \left(\begin{array}{c|c} S_1 & \\ \hline d_1^T & \beta_1 \end{array}\right) \left(\begin{array}{c|c} I_1 & c_1 \\ \hline & 1 \end{array}\right)$$

met $S_1 \in M_{m-1,m-1}$, $c_1 := S_1^{-1} b_1$, $\beta_1 = \delta_1 - d_1^T c_1$.

Bewijs dat de elementen van c_1 de Taylor-coëfficiënten zijn van het interpolatiepolynoom $q(x)$ met graad $\le m-2$ dat de functie x^{m-1} interpoleert in de punten $x_1, \ldots, x_{m-1}$ en dat $d_1^T c_1 = q(x_m)$. Daar

$$q(x) = x^{m-1} - \prod_{j=1}^{m-1} (x - x_j)$$

volgt hieruit

$$\beta_1 = \prod_{j=1}^{m-1} (x_m - x_j) .$$

Met inductie naar m en eenvoudige regels voor determinanten volgt hieruit

$$\det(S) = \beta_1 \det(S_1) = \prod_{1 \le j < k \le m} (x_k - x_j)$$

(van der Monde).

Opgave. (6.1.12)

Stel V heeft bases $(\underline{u}_1,\dots,\underline{u}_m)$, $(\underline{u}_1',\dots,\underline{u}_m')$ en $(\underline{u}_1'',\dots,\underline{u}_m'')$. Noem de overgangsmatrix van de eerste naar de tweede basis S_{12}, analoog S_{23} en S_{13}. Bewijs dat

$$S_{13} = S_{12}S_{23} .$$

Opgave. (6.1.13)

Zij V een n-dimensionale deelruimte van $\mathcal{L}^m$ met basis $(\underline{b}_1,\dots,\underline{b}_n)$. Dan is er bij iedere $\underline{a} \in V$ een $\underline{x} = (\xi_1,\dots,\xi_n) \in \mathcal{L}^n$ zo dat $\underline{a} = \Sigma\xi_j\underline{b}_j$, of, in matrixtaal, $a = Bx$, met $B = (b_1 | \dots | b_n)$.
Bewijs dat $B \in RR$ en dat voor iedere $\underline{a} \in V$

$$x = B^- a ,$$

waarin B^- een willekeurige linkerinverse van B is.
(N.b. Omdat V ingebed is in $\mathcal{L}^m$ wordt de coördinatiserende afbeelding - die een afbeelding is van een n-dimensionale deelruimte van $\mathcal{L}^m$ op $\mathcal{L}^n$ - hier gerepresenteerd door een $n \times m$ matrix B^- met rang n).

6.2. Lineaire afbeeldingen van V_2 in V_1

Definitie. (6.2.1)

Zij V_1 en V_2 vectorruimten met bases $(\underline{v}_1,\dots,\underline{v}_m)$, resp. $(\underline{u}_1,\dots,\underline{u}_n)$. Zij $\mathcal{A} \in LA(V_1,V_2)$ lineair. Dan heet de matrix $A = (\alpha_{ij})$, gedefinieerd door

$$\mathcal{A}\underline{u}_j = \sum_{i=1}^{m} \alpha_{ij}\underline{v}_i , \qquad j = 1,\dots,n ;$$

de matrix van $\mathcal{A}$ t.o.v. de genoemde bases voor V_1 en V_2.
N.B. De kolommen van A zijn de coördinaten van $\mathcal{A}\underline{u}_j$ t.o.v. $(\underline{v}_1,\dots,\underline{v}_m)$.

Stelling. (6.2.2)

Zij V_1, V_2, $\mathcal{A}$ en A als in (6.2.1). Zij $C_1 : V_1 \to \mathcal{L}^m$ en $C_2 : V_2 \to \mathcal{L}^n$ de bijbehorende coördinatiserende afbeeldingen. Dan is A de matrix van de afbeelding $C_1\mathcal{A}C_2^{-1} \in LA(\mathcal{L}^m,\mathcal{L}^n)$ t.o.v. de natuurlijke bases in $\mathcal{L}^n$ en $\mathcal{L}^m$.

Bewijs.

Zij $(\tilde{\underline{e}}_1,\ldots,\tilde{\underline{e}}_n)$, resp. $(\underline{e}_1,\ldots,\underline{e}_m)$ de natuurlijke bases voor $\mathcal{L}^n$ en $\mathcal{L}^m$.
Uit (6.1.2) en (6.2.1)

$$C_1AC_2^{-1}\tilde{\underline{e}}_j = C_1A\underline{u}_j = \sum_i \alpha_{ij}C_1\underline{v}_i = \Sigma\alpha_{ij}\underline{e}_i \ .$$

De bewering volgt nu uit (5.3.5). □

Gevolg. (6.2.3)

Als $\underline{u} = \Sigma\xi_j\underline{u}_j \in V_2$ en $\underline{v} = \Sigma\eta_i\underline{v}_i \in V_1$, dan geldt

$$\underline{v} = A\underline{u} \iff y = Ax \ ,$$

met $\underline{x} = (\xi_1,\ldots,\xi_n)$, $\underline{y} = (\eta_1,\ldots,\eta_m)$.

Stelling. (6.2.4)

Stel

V_1 heeft bases $(\underline{v}_1,\ldots,\underline{v}_m)$ en $(\underline{v}'_1,\ldots,\underline{v}'_m)$,
V_2 heeft bases $(\underline{u}_1,\ldots,\underline{u}_n)$ en $(\underline{u}'_1,\ldots,\underline{u}'_n)$.

Stel $A \in LA(V_1,V_2)$ heeft matrices A en A' t.o.v. genoemde basisparen. Dan geldt

$$A' = S_1^{-1}AS_2 \ ,$$

waarin S_1 de overgangsmatrix is van de basis $(\underline{v}_1,\ldots,\underline{v}_m)$, naar de basis $(\underline{v}'_1,\ldots,\underline{v}'_m)$ en S_2 de overgangsmatrix van de basis $(\underline{u}_1,\ldots,\underline{u}_n)$ naar de basis $(\underline{u}'_1,\ldots,\underline{u}'_n)$.

Bewijs.

Volgt uit (6.2.2) en (6.1.8), daar

$$C'_1AC'^{-1}_2 = C'_1C_1^{-1} \cdot C_1AC_2^{-1} \cdot C_2C'^{-1}_2 \ .$$ □

Opmerking.

We kunnen dit als volgt verifiëren:

als $\underline{u} = \Sigma\xi_j\underline{u}_j = \Sigma\xi_j'\underline{u}_j' \in V_1$ en $\underline{v} = \Sigma\eta_i\underline{v}_i = \Sigma\eta_i'\underline{v}_i' \in V_2$ dan geldt $\underline{v} = A\underline{u} \Leftrightarrow$ $\Leftrightarrow y = Ax \Leftrightarrow y' = A'x'$. Volgens (6.1.9) is echter $x = S_2x'$ en $y = S_1y'$, dus $y' = S_1^{-1}AS_2x'$, dus $A' = S_1^{-1}AS_2$.

Gevolgen. (6.2.5)

1. A en A' zijn equivalent dan en slechts dan als zij matrices zijn van een afbeelding $A \in LA(V_1,V_2)$ t.o.v. verschillende paren bases in V_2 en V_1.

2. Zij A de matrix van een afbeelding $A \in LA(\mathcal{L}^m,\mathcal{L}^n)$ t.o.v. de natuurlijke bases in $\mathcal{L}^n$ en $\mathcal{L}^m$. Zij $(\underline{b}_1,\dots,\underline{b}_m)$ en $(\underline{v}_1,\dots,\underline{v}_n)$ andere bases voor $\mathcal{L}^m$ resp. $\mathcal{L}^n$.
   Dan is de matrix van A t.o.v. deze bases $B^{-1}AV$, waarin $B = (b_1| \dots |b_m)$, $V = (v_1| \dots |v_n)$.

Opgave. (6.2.6)

Definieer de rang $r(A)$ van een afbeelding $A \in LA(V_1,V_2)$; de range $R(A)$ en de nulruimte $N(A)$. Generaliseer nu (5.4.8.2), (5.4.14).

Opgave. (6.2.7)

Zij $A \in LA(V_1,V_2)$ met rang r. Bewijs dat er bases in V_2 en V_1 zijn zodanig dat de matrix van A t.o.v. deze bases de normaalvorm uit $M^r_{m,n}$ is:

$$A = \left(\begin{array}{c|c} I_1 & 0 \\ \hline 0 & 0 \end{array}\right),$$

met I_1 de $r \times r$ eenheidsmatrix.

Opgave. (6.2.8)

Zij $A \in LA(V,V)$. Zij $V_1 \subset V$ zo dat $AV_1 \subset V_1$ (d.w.z. $\forall_{\underline{u}\in V_1}$: $A\underline{u} \in V_1$, V_1 heet dan een t.o.v. A invariante deelruimte van V).
Zij $(\underline{u}_1,\ldots,\underline{u}_m)$ een basis voor V zo dat $(\underline{u}_1,\ldots,\underline{u}_n)$ een basis voor V_1 is.
Dan ziet de matrix van A t.o.v. deze basis er uit als

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right), \text{ met } A_{11} \in M_{n,n}.$$

(A_{11} is de matrix van de zg. restrictie van A tot V_1, dwz. het door A geïnduceerde element van $LA(V_1,V_1)$). Als er ook een $V_2 \subset V$ is zo dat $V_1 \oplus V_2 = V$ en $AV_2 \subset V_2$, dan kunnen we de basis zo kiezen dat $A_{12} = 0$.

Opgave. (6.2.9)

Zij V de ruimte der polynomen $\underline{p}$ met graad < m. Zij $A \in LA(V,V)$ de afbeelding die aan een polynoom zijn afgeleide toevoegt (dit is een lineaire afbeelding van V in V!). Kiezen we als basis voor V de polynomen

$$\underline{p}_j : p_j(x) = \frac{x^{j-1}}{(j-1)!}, \qquad j = 1,\ldots,m,$$

dan heeft A t.o.v. deze basis de matrix

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

(de niet ingevulde elementen zijn ook nul).
Wat zijn A^2, A^3, ... en de matrices hiervan?
Bepaal ook de rg-inversen van A en onderzoek de hierbij horende afbeelding uit LA(V,V) aan de hand van (5.5.15) en (5.5.16).

7. Vectorruimten met inproduct (unitaire vectorruimten).

7.1. Inproduct en orthogonaliteit

Definitie. (7.1.1)

Een inproduct in een vectorruimte V over $\mathbb{R}$ is een afbeelding $V \times V \to \mathbb{R}$: $\underline{u} \in V$, $\underline{v} \in V \mapsto (\underline{u},\underline{v}) \in \mathbb{R}$, waarvoor geldt

i) $(\underline{u},\underline{v}) = (\underline{v},\underline{u})$,

ii) $(\underline{u},\underline{v} + \underline{w}) = (\underline{u},\underline{v}) + (\underline{u},\underline{w})$,

iii) $(\underline{u},\alpha\underline{v}) = \alpha(\underline{u},\underline{v})$,

iv) $\underline{u} \neq \underline{0} \Rightarrow (\underline{u},\underline{u}) > 0$.

Een inproduct in een vectorruimte V over $\mathbb{C}$ is een afbeelding $V \times V \to \mathbb{C}$ waarvoor ii), iii) en iv) gelden en in plaats van i)

ia) $(\underline{u},\underline{v}) = \overline{(\underline{v},\underline{u})}$.

Opmerkingen. (7.1.2)

1. Eis iv) veronderstelt een ordening in het grondlichaam $\mathcal{L}$, daarom zijn alleen $\mathbb{R}$ en $\mathbb{C}$ interessant ($\mathbb{C}$ ook, omdat uit ia) volgt dat $(\underline{u},\underline{u})$ reëel is).
2. In het reële geval volgt uit i) en iii) dat

   $$(\alpha\underline{u},\underline{v}) = \alpha(\underline{u},\underline{v}) ,$$

   in het complexe geval geldt echter

   $$(\alpha\underline{u},\underline{v}) = \bar{\alpha}(\underline{u},\underline{v}) .$$

3. Vectorruimten met inproduct over $\mathbb{R}$ noemt men wel Euclidische vectorruimten, die over $\mathbb{C}$ heten unitaire vectorruimten.

In het vervolg is V steeds een eindig dimensionale unitaire vectorruimte. Bijna alle stellingen gelden overigens ook voor Euclidische vectorruimten.

Kunnen we $\mathbb{C}^m$ zelf voorzien van een inproduct en kunnen we het inproduct in een m-dimensionale V in verband brengen met een inproduct in $\mathbb{C}^m$?

Definitie. (7.1.3)

Het zg. natuurlijke inproduct van $x = (\xi_1, \ldots, \xi_m) \in \mathbb{C}^m$ en $y = (\eta_1, \ldots, \eta_m) \in \mathbb{C}^m$ is

$$(x,y) := \Sigma \bar{\xi}_i \eta_i = x^H y \ .$$

Ga na dat met dit inproduct $\mathbb{C}^m$ een unitaire ruimte wordt.

Definitie. (7.1.4)

$\underline{u}$ en $\underline{v} \in V$ heten orthogonaal als $(\underline{u},\underline{v}) = 0$ (deze relatie is symmetrisch!).
Deelruimten V' en $V'' \in V$ heten orthogonaal als $\forall_{\underline{u} \in V'} \ \forall_{\underline{v} \in V''} \ (\underline{u},\underline{v}) = 0$.
$\underline{u}_1, \ldots, \underline{u}_n \in V$ heten orthonormaal als

$$(\underline{u}_i, \underline{u}_j) = \delta_{ij}, \quad i,j = 1,\ldots,n \ .$$

Stelling. (7.1.5)

Als $\underline{u}_1, \ldots, \underline{u}_m \in V$ orthonormaal zijn dan zijn ze onafhankelijk.
Zij vormen een basis voor V dan en slechts dan als $m = \dim(V)$.

Stelling. (7.1.6)

Zij $(\underline{u}_1, \ldots, \underline{u}_m)$ een basis voor V. Zij $\mathbf{x}$ en $\mathbf{y} \in \mathbb{C}^m$ de coördinatenvectoren t.o.v. deze basis van $\underline{u}$ en $\underline{v} \in V$ (dus $\underline{u} = \Sigma \xi_i \underline{u}_i$, etc.).
Dan zijn de volgende uitspraken equivalent:

i) $(\underline{u}_1, \ldots, \underline{u}_m)$ is een orthonormale basis,
ii) voor alle $\underline{u}$ en $\underline{v} \in V$ geldt $(\underline{u},\underline{v}) = x^H y$.

Bewijs.

i) $\Rightarrow$ ii): $(\underline{u},\underline{v}) = (\Sigma \xi_i \underline{u}_i, \Sigma \eta_j \underline{u}_j) = \Sigma \bar{\xi}_i \eta_j (\underline{u}_i, \underline{u}_j) = \Sigma \bar{\xi}_i \eta_i = x^H y$.
ii) $\Rightarrow$ i): Neem $\underline{u} = \underline{u}_i$, $\underline{v} = \underline{u}_j$, dan is $x = e_i$, $y = e_j$ (natuurlijke basisvectoren uit $\mathbb{C}^m$), dus

$$(\underline{u}_i, \underline{u}_j) = e_i^H e_j = \delta_{ij} \ . \qquad \square$$

Opmerking. (7.1.7)

In het algemeen geldt: als $(\underline{u}_1,\ldots,\underline{u}_m)$ een basis voor V is, dan is er een matrix $G \in M_{mm}$ zo dat voor alle $\underline{u}$ en $\underline{v} \in V$ met coördinaatvectoren x en $y \in \mathbb{C}^m$ geldt

$$(*) \qquad (\underline{u},\underline{v}) = x^H G y .$$

Neem nl. $(G)_{ij} = (\underline{u}_i,\underline{u}_j)$. De matrix G is positief definiet (zie § 10.2), hetgeen direct uit de eisen iia) en iv) voor het inproduct volgt. Omgekeerd zij V een vectorruimte over $\mathbb{C}$ met basis $(\underline{u}_1,\ldots,\underline{u}_m)$. Dan wordt bij iedere positief definite G door (*) een inproduct gedefinieerd (ga na). $(\underline{u}_1,\ldots,\underline{u}_m)$ is t.o.v. dit inproduct een orthonormale basis dan en slechts dan als G = I.

Stelling. (Gram-Schmidt orthogonalisatie). (7.1.8)

Zij $\underline{u}_1,\ldots,\underline{u}_n \in V$ onafhankelijk. Dan zijn er $\underline{q}_1,\ldots,\underline{q}_n \in V$ zo dat voor $1 \le k \le n$ $(\underline{q}_1,\ldots,\underline{q}_k)$ een orthonormale basis voor $[\underline{u}_1,\ldots,\underline{u}_k]$ is.

Bewijs.

Inductie naar n. Het geval n = 1 is triviaal (neem $\underline{q}_1 = (\underline{u}_1,\underline{u}_1)^{-\frac{1}{2}}\underline{u}_1$). Stel $\underline{q}_1,\ldots,\underline{q}_{n-1}$ zo dat voor $1 \le k \le n-1$ $(\underline{q}_1,\ldots,\underline{q}_k)$ orthonormale basis is voor $[\underline{u}_1,\ldots,\underline{u}_k]$. Daar $\underline{u}_n \notin [\underline{u}_1,\ldots,\underline{u}_{n-1}]$ is dan

$$\underline{v} := \underline{u}_n - \sum_{1}^{n-1} (\underline{q}_i,\underline{u}_n)\underline{q}_i$$

$\perp \underline{q}_1,\ldots,\underline{q}_{n-1}$ en $\neq \underline{0}$. Neem nu $\underline{q}_n = (\underline{v},\underline{v})^{-\frac{1}{2}}\underline{v}$, dan is $(\underline{q}_1,\ldots,\underline{q}_n)$ orthonormale basis voor $[\underline{u}_1,\ldots,\underline{u}_n]$. □

Opmerkingen. (7.1.9)

1) Uit deze stelling volgt dat in een unitaire vectorruimte iedere eindig dimensionale deelruimte een orthonormale basis heeft.

2) Er is een rechtsboven matrix $R = (\rho_{ij})$ zo dat

$$(*) \qquad \underline{u}_j = \sum_{i=1}^{i} \rho_{ij}\underline{q}_i$$

3) Zij $(\underline{u}_1,\ldots,\underline{u}_m)$ basis voor V, $R = (\rho_{ij})$ een reguliere matrix en $\underline{g}_1,\ldots,\underline{g}_m$ bepaald door (*). Dan is $(\underline{g}_1,\ldots,\underline{g}_m)$ orthonormale basis voor V dan en slechts dan als $R^H R = G$ met G als in (7.1.7). In (10.2.7) zal bewezen worden dat er bij iedere positief definiete G een R (en zelfs een rechtsboven R) is die voldoet aan deze relatie. Dit levert een tweede methode om orthonormale bases te construeren.

<u>Stelling</u>. (uitbreidingsstelling). (7.1.10)

Zij V m-dimensionaal en $V_1 \subset V$ n-dimensionaal met orthonormale basis $(\underline{g}_1,\ldots,\underline{g}_n)$. Dan zijn er $\underline{g}_{n+1},\ldots,\underline{g}_m \in V$ zo dat $(\underline{g}_1,\ldots,\underline{g}_m)$ orthonormale basis is voor V.

<u>Bewijs</u>.

Volgens (5.4.14) zijn er $\underline{u}_{n+1},\ldots,\underline{u}_m \in V$ zo dat $(\underline{g}_1,\ldots,\underline{g}_n, \underline{u}_{n+1},\ldots,\underline{u}_m)$ basis is voor V. Orthogonaliseer deze basis; de eerste n vectoren blijven dan onveranderd. □

<u>Stelling</u>. (projectiestelling). (7.1.11)

Zij V een vectorruimte met inproduct en $V_1 \subset V$ n-dimensionaal. Dan is de verzameling

$$V_1^\perp := \{\underline{w} \in V \mid \forall_{\underline{v} \in V_1} : (\underline{v},\underline{w}) = 0\}$$

een deelruimte van V (het zgn. orthogonaal complement van V_1 in V).
Er geldt

$$V_1 \oplus V_1 = V\ , \quad (V_1^\perp)^\perp = V_1\ .$$

<u>Bewijs</u>.

Dat $V_1^\perp$ een lineaire deelruimte van V is is duidelijk.
Als $\underline{u} \in V_1 \cap V_1^\perp$ dan is $(\underline{u},\underline{u}) = 0$, dus $\underline{u} = \underline{0}$. Dus $V_1 \cap V_1^\perp = \{\underline{0}\}$.
Zij $(\underline{g}_1,\ldots,\underline{g}_n)$ een orthonormale basis voor V_1. Zij $\underline{u} \in V$, $\underline{v} = \Sigma(\underline{g}_i,\underline{u})\underline{g}_i$ en $\underline{w} = \underline{u} - \underline{v}$. Dan is $\underline{v} \in V_1$, $(\underline{g}_i,\underline{w}) = 0$ voor $i = 1,\ldots,n$, dus $\underline{w} \in V_1^\perp$. Hieruit volgt dat $V_1 + V_1^\perp = V$.

Uit $V_1 \perp V_1^\perp$ volgt dat $V_1 \subset (V_1^\perp)^\perp$. Zij nu $\underline{w} \in (V_1^\perp)^\perp$. Volgens het bovenstaande zijn er $\underline{u} \in V_1$ en $\underline{v} \in V_1^\perp$ zo dat $\underline{w} = \underline{u} + \underline{v}$. Dan is echter $(\underline{v},\underline{u}) = 0$ en $(\underline{v},\underline{w}) = 0$, dus $(\underline{v},\underline{v}) = 0$, dus $\underline{v} = \underline{0}$ en $\underline{w} \in V_1$. Derhalve $(V_1^\perp)^\perp \subset V_1$. □

Opmerking. (7.1.12)

In het bovenstaande bewijs is niet gebruikt dat V eindig dimensionaal is. Als $\dim(V) < \infty$ dan volgt uit $V_1 \oplus V_1^\perp = V$ dat

$$\dim((V_1^\perp)^\perp) = \dim(V) - \dim(V_1^\perp) = \dim(V_1)$$

en dan impliceert $V_1 \subset (V_1^\perp)^\perp$ al dat $V_1 = (V_1^\perp)^\perp$.

Opgave. (7.1.13)

Zij V_1 en $V_2 \subset V$. Dan geldt

$$V_1 \subset V_2 \Leftrightarrow V_1^\perp \supset V_2^\perp \,,$$
$$(V_1 \cap V_2)^\perp = V_1^\perp + V_2^\perp \,, \quad (V_1 + V_2)^\perp = V_1^\perp \cap V_2^\perp \,,$$
$$V_1 \cap V_2 = \{\underline{0}\} \Leftrightarrow V_1^\perp + V_2^\perp = V \,, \quad V_1 + V_2 = V \Leftrightarrow V_1^\perp \cap V_2^\perp = \{\underline{0}\} \,,$$
$$V_1 \oplus V_2 = V \Leftrightarrow V_1^\perp \oplus V_2^\perp = V \,.$$

Stelling. (7.1.14)

Zij $(\underline{q}_1,\ldots,\underline{q}_m)$ een orthonormale basis voor V en $(\underline{v}_1,\ldots,\underline{v}_n)$ een basis voor een deelruimte $V' \subset V$. Zij $S = (\sigma_{ij}) \in M_{m,n}$ gedefinieerd door

$$\underline{v}_j = \sum_{i=1}^{m} \sigma_{ij}\underline{q}_i \,, \qquad j = 1,\ldots,n$$

(een gegeneraliseerde overgangsmatrix, vgl. (6.1.7)).
Dan geldt: $(\underline{v}_1,\ldots,\underline{v}_n)$ is orthonormaal dan en slechts dan als S voldoet aan $S^H S = I$ (S heet dan kolom-unitair, zie verder § 8.1).

Bewijs.

Er geldt:

$$(\underline{v}_k,\underline{v}_\ell) = \sum_{i=1}^{m} \bar{\sigma}_{ik}\sigma_{i\ell} = (S^H S)_{k\ell} \,.$$ □

Gevolg. (7.1.15)

Als $(\underline{q}_1,\ldots,\underline{q}_m)$ en $(\underline{q}_1',\ldots,\underline{q}_m')$ orthonormale bases voor V zijn dan voldoet de overgangsmatrix S aan $S^H S = I$ (dus, daar S vierkant is, S inverteerbaar en $S^{-1} = S^H$, S heet dan unitair, zie § 8.2).

Stelling. (7.1.16)

Zij V_1 en V_2 eindig dimensionale vectorruimten met inproduct en $A \in LA(V_1,V_2)$ lineair. Dan is er een eenduidig bepaalde lineaire afbeelding $A^* \in LA(V_2,V_1)$ zo dat

$$\forall_{\underline{u}\in V_1}\ \forall_{\underline{v}\in V_2} : (\underline{u},A\underline{v}) = (A^*\underline{u},\underline{v}) .$$

Als A de matrix van A t.o.v. orthonormale bases $(\underline{u}_1,\ldots,\underline{u}_m)$ voor V_1 en $(\underline{v}_1,\ldots,\underline{v}_n)$ voor V_2 is dan is A^H de matrix van A^* t.o.v. deze bases.

Bewijs.

Dat er hoogstens één lineaire A^* is die voldoet volgt uit het feit dat $(A^*\underline{u},\underline{v}) = 0$ voor alle $\underline{u}$ en $\underline{v}$ impliceert dat $A^*\underline{u} = \underline{0}$ voor alle $\underline{u}$ (neem nl. $\underline{v} = A^*\underline{u}$) en dus $A^* = 0$.
Kies nu orthonormale bases. Zij $A = (\alpha_{ij})$ de matrix van A t.o.v. deze bases, d.w.z.

$$A\underline{v}_j = \sum_i \alpha_{ij}\underline{u}_i, \quad j = 1,\ldots,n .$$

Definieer A^* door

$$A^*\underline{u}_i = \sum_j \bar{\alpha}_{ij}\underline{v}_j, \quad i = 1,\ldots,m .$$

Dan is $(\bar{\alpha}_{ji}) = A^H$ de matrix van A^* en er geldt voor alle i en j

$$(A^*\underline{u}_i,\underline{v}_j) = \alpha_{ij} = (\underline{u}_i,A\underline{v}_j) .$$

Hieruit volgt eenvoudig dat voor alle $\underline{u} \in V_1$ en $\underline{v} \in V_2$

$$(A^*\underline{u},\underline{v}) = \Sigma\alpha_{ij}(\underline{u},\underline{u}_i)(\underline{v}_j,\underline{v}) = (\underline{u},A\underline{v}) . \qquad \square$$

Opmerking. (7.1.17)

Men noemt A^* de aan A geadjungeerde afbeelding. Als $V_1 = V_2 = V$, dan zijn A en A^* afbeeldingen van V in V; is in dit geval $A^* = A$ dan heet A zelfgeadjungeerd, voor de matrix t.o.v. een orthonormale basis voor V geldt dan $A^H = A$, d.w.z., A is hermitisch (en omgekeerd).

Opgave (7.1.18)

Zij V een unitaire vectorruimte met inproduct (,). Stel dat de afbeelding van V×V in $\mathbb{C}$: $\underline{u} \in V$, $\underline{v} \in V \mapsto \langle\underline{u},\underline{v}\rangle \in \mathbb{C}$ ook voldoet aan de eisen ia), ii), iii), iv) uit (7.1.1). Dan is er een afbeelding $G \in LA(V,V)$ zodat voor alle $\underline{u}$ en $\underline{v} \in V$

$$\langle\underline{u},\underline{v}\rangle = (G\underline{u},\underline{v}) \ .$$

G is zelfgeadjungeerd en positief definiet.

Opgave (7.1.19)

Zij $A \in LA(V,V)$ zelfgeadjungeerd. Als $V_1 \subset V$ invariant t.o.v. A (d.w.z., $A(V_1) \subset V_1$) dan is ook $V_1^{\perp}$ invariant t.o.v. A.

Hoofdstuk III. Vervolg matrixalgebra.

8. Orthogonale decompositie

8.1. Matrices over $\mathbb{R}$ of $\mathbb{C}$

In dit hoofdstuk kijken we naar matrices met elementen uit $\mathbb{R}$ of $\mathbb{C}$ en met name naar eigenschappen die met andere grondlichamen niet altijd gelden. Dat er met $\mathbb{R}$ of $\mathbb{C}$ als grondlichaam speciale eigenschappen gelden wordt in hoofdzaak veroorzaakt door het feit dat

$$\text{als } x \in \mathbb{R}^m \text{ dan } x^T x \geq 0 \text{ en } x^T x = 0 \Leftrightarrow x = 0 , \tag{8.1.1}$$

$$\text{als } x \in \mathbb{C}^m \text{ dan } x^H x \geq 0 \text{ en } x^H x = 0 \Leftrightarrow x = 0 . \tag{8.1.2}$$

Dat dit niet voor ieder lichaam geldt zien we aan voorbeelden. Als bv. $\mathcal{L}$ bestaat uit de gehele getallen modulo een priemgetal p (deze vormen een (eindig) lichaam) dan geldt voor $x = (1,\ldots,1)^T \in \mathcal{L}^p$ dat $x^T x = 0$. Een ander voorbeeld: in $\mathbb{C}^m$ geldt niet dat $x^T x = 0 \Rightarrow x = 0$.
Uit (8.1.1) of (8.1.2) volgt dat $(x,x) := x^T x$, resp. $(x,x) := x^H x$ een inproduct is in $\mathbb{R}^m$, resp. $\mathbb{C}^m$; daaruit volgen zaken als het bestaan van orthonormale bases (via het feit dat iedere vector $\neq 0$ genormeerd kan worden), de projectiestelling, etc.

In het vervolg van dit hoofdstuk werken we steeds met matrices over $\mathbb{C}$ (voor matrices over $\mathbb{R}$ gelden grotendeels dezelfde resultaten, enkele afwijkingen geven we aan).

Gevolgen van (8.1.2). (8.1.3)

1. $A^H AX = 0 \Leftrightarrow AX = 0$.
2. $A \in RR \Leftrightarrow A^H A \in I$; $A \in LR \Leftrightarrow AA^H \in I$.
3. $N(A^H AB) = N(AB)$.
4. $R(BAA^H) = R(BA)$.
5. $r(A) = r(A^H) = r(AA^H) = r(A^H A) = r(AA^H A) = r(A^H AA^H) = \ldots$.
6. $(\forall_y : \mathrm{Re}(y^H x) = 0) \Leftrightarrow x = 0$.

Bewijs zelf!

Definitie. (8.1.4)

$A \in M_{m,m}(\mathbb{C})$ (dus vierkant) heet hermitisch als $A^H = A$.
$A \in M_{m,m}(\mathbb{R})$ heet symmetrisch als $A^T = A$.

Stelling. (8.1.5)

Zij $A \in M^r_{m,m}$ hermitisch. Dan zijn er

$B_1 \in M^r_{m,r}$ (dus $B_1 \in RR$)

$Z_{11} \in M^r_{r,r}$ (dus $Z_{11} \in I$) en hermitisch

zo dat

$$A = B_1 Z_{11} B_1^H .$$

Bewijs.

Zij $A = B_1 C_1^H$ met B_1 en $C_1 \in RR$ (dit is een nu meer passende vorm van de decompositiestelling). Zij U_1^H linker inverse van B_1. Dan volgt uit $A = B_1 C_1^H = C_1 B_1^H$ dat $A = B_1 U_1^H A U_1 B_1^H = B_1 Z_{11} B_1^H$ met $Z_{11} = U_1^H A U_1 \in M_{r,r}$ en hermitisch. Uit deze relaties volgt tevens $r(A) \le r(Z_{11}) \le r(A)$. □

Opgave. (8.1.6)

Onderzoek de niet-eenduidigheid van de splitsing (8.1.5).

Opgave. (8.1.7)

Zij $A \in M_{m,m}$ en $A^2 = A^H A$. Dan is A hermitisch.
Hint: Stel $A = B_1 C_1^H$ met B_1 en $C_1 \in RR$.

Stelling. (8.1.8)

Zij $A \in M_{m,n}$. Dan geldt

$$(R(A))^\perp = N(A^H), \quad (N(A))^\perp = R(A^H) .$$

Hint: Stel $A = B_1 C_1^H$ met B_1 en $C_1 \in RR$.

Bewijs.

$y \in (R(A))^\perp \Leftrightarrow \forall_x : y^H A x = 0 \Leftrightarrow A^H y = 0 \Leftrightarrow y \in N(A^H)$. En analoog. □

Gevolgen. (8.1.9)

1) $$R(A) \cap N(A^H) = \{\underline{0}\},\ R(A) \oplus N(A^H) = \mathbb{C}^m ,$$
$$N(A) \cap R(A^H) = \{\underline{0}\},\ N(A) \oplus R(A^H) = \mathbb{C}^n .$$

2) Volgens (5.4.5) geldt $R(A_2) \subset R(A_1) \Leftrightarrow \exists_X: A_2 = A_1X$,
volgens (5.4.20) geldt $N(A_2) \supset N(A_1) \Leftrightarrow \exists_Y: A_2 = YA_1$.
Het bewijs van de eerste bewering is vrijwel triviaal, dat van de tweede bewering wat minder doorzichtig. Met behulp van (8.1.4) en (7.1.13) is echter (5.4.20) eenvoudig uit (5.4.5) af te leiden:

$$N(A_2) \supset N(A_1) \Leftrightarrow (R(A_2^H))^\perp \supset (R(A_1^H))^\perp \Leftrightarrow R(A_2^H) \subset R(A_1^H)$$

$$\Leftrightarrow \exists_Y: A_2^H = A_1^HY^H \Leftrightarrow \exists_Y: A_2 = YA_1 .$$

3) Als A hermitisch is dan is $(R(A))^\perp = N(A)$.

4) De vergelijking $Ax = z$ is oplosbaar dan en slechts dan als $z \in R(A) = (N(A^H))^\perp$, dus als $z \perp N(A^H)$, dus als $\forall y : y^HA = 0^H \Rightarrow y^Hz = 0$ (vgl. (4.2.8)!).

5) Als A_1 en A_2 hoogte m hebben dan geldt (vgl. (15.4.18))

$$R(A_1) = (R(A_2))^\perp \Leftrightarrow A_2^HA_1 = 0 \text{ en } r(A_1) + r(A_2) = m .$$

Opgave. (8.1.10)

Zij $A \in M_{m,n}$. Dan is er bij iedere $y \in \mathcal{L}^m$ een $x \in \mathcal{L}^n$ zo dat $A^H(y - Ax) = 0$ en daarbij is Ax eenduidig bepaald. Bewijs dit met behulp van (8.1.3) en leid hieruit af dat $R(A) \oplus N(A^H) = \mathbb{C}^m$.

Opgave. (8.1.11)

Zij $A \in M_{m,m}$. Bewijs dat de volgende uitspraken equivalent zijn:

i) $(R(A))^\perp = N(A)$,

ii) $R(A^H) = R(A)$,

iii) $N(A^H) = N(A)$.

Matrices met deze eigenschappen heten range-hermitian.
Bewijs dat hermitische matrices en ook alle matrices waarvoor $AA^H = A^HA$ (de zg. normale matrices) range-hermitian zijn.
Bewijs dat een matrix die range-hermitian is, index 1 (zie (5.5.27)) heeft.
Bewijs dat A range-hermitian is dan en slechts dan als

$$A = B_1Z_{11}B_1^H$$

met $B_1 \in RR$, $Z_{11} \in I$. Wat is dan de groepsinverse (zie (5.5.27)) van A?

Opgave. (8.1.12)

Zij $A = (A_1 \mid A_2)$ regulier met inverse $X = \left(\frac{X_1^H}{X_2^H}\right)$.

Dan zijn de volgende beweringen equivalent:

i) $A_1^H A_2 = 0$,

ii) $X_1 = A_1(A_1^H A_1)^{-1}$,

iii) $A_1 X_1^H = X_1 A_1^H$,

iv) $X_1^H X_2 = 0$.

8.2. Unitaire matrices, Gram-Schmidt.

Definitie. (8.2.1)

$U \in M_{m,n}$ heet links-unitair (ook wel: kolom-unitair) als $U^H U = I$.
Als ook $UU^H = I$ dan heet U unitair.

Eenvoudige eigenschappen. (8.2.2)

1) Als U_1 links-unitair is dan $U_1 \in RR$, $n \leq m$ en U_1^H is linkerinverse van U_1.
2) De kolommen van U_1 zijn orthonormaal (d.w.z. $u_i^H u_j := \delta_{ij}$) dan en slechts dan als U_1 links-unitair is.
3) U is unitair dan en slechts dan als U regulier is en $U^{-1} = U^H$.
4) Een U is unitair dan en slechts dan als zijn kolommen een orthonormale basis voor $\mathbb{C}^m$ vormen.
5) Als U_1 links-unitair is met $n < m$ dan is er een links-unitaire U_2 zo dat $U = (U_1 \mid U_2)$ unitair is (dit is een vertaling van (7.1.10)).
6) Als $U = (U_1 \mid U_2)$ dan geldt

$$U \text{ is links-unitair} \Leftrightarrow (U_1 \text{ en } U_2 \text{ links-unitair en } U_1^H U_2 = 0) ,$$

$$U \text{ is unitair} \Leftrightarrow (U_1 \text{ en } U_2 \text{ links-unitair en } U_1 U_1^H + U_2 U_2^H = I) .$$

7) Als U_1 en U_2 links-unitair zijn dan geldt voor $U = (U_1 \mid U_2)$

$$U \text{ is unitair} \Leftrightarrow R(U_2) = (R(U_1))^{\perp} .$$

8) Als U_1 en V_1 links-unitair zijn dan geldt

$$R(V_1) \subset R(U_1) \Leftrightarrow V_1 = U_1 Z_{11} \text{ met } Z_{11} \text{ links-unitair} ,$$
$$R(V_1) = R(U_1) \Leftrightarrow V_1 = U_1 Z_{11} \text{ met } Z_{11} \text{ unitair.}$$

9) Als U_1 en V_1 links-unitair zijn en U_1V_1 bestaat dan is U_1V_1 links-unitair.
10) De unitaire $m \times m$-matrices vormen een groep: met U en V zijn ook UV en U^{-1} unitaire $m \times m$-matrices.
11) Als U unitair en rechtsboven is (dus met name als U unitair en diagonaal is), dan is U van de vorm $U = \text{diag}(e^{i\theta_1},\dots,e^{i\theta_m})$, met reële θ's (als U rechtsboven is dan is volgens (2.2.5) U^{-1} ook rechtsboven, dus $U = U^{-H}$ is zowel linksonder als rechtsboven).
12) Een permutatiematrix P is unitair.

Opgave. (8.2.3)

Bewijs (8.2.2.7) ook met behulp van het feit dat

$$N(U_1^H) = R(I - U_1U_1^H)) \ .$$

Stelling. (Gram-Schmidt). (8.2.4)

Zij $A \in M^n_{m,n}$ (dus $A \in RR$). Dan is er een links-unitaire $U_1 \in M_{m,n}$ en een reguliere rechter driehoeksmatrix R_1 zo dat $A = U_1R_1$.

Bewijs.

Volgt uit (7.1.8) en (7.1.9.2). □

Opmerking. (8.2.5)

De constructie in het bewijs van (7.1.8) levert U_1 en R_1 zo, dat de diagonaalelementen van R_1 positief zijn en uit de constructie blijkt (ga na) dat U_1 en R_1 hierdoor eenduidig bepaald zijn. We kunnen dit ook rechtstreeks inzien.
Iets algemener geldt

als $U_1R_1 = U_2R_2$, met U_1 en U_2 links-unitair, R_1 en R_2 rechtsboven en regulier, dan is er een unitaire diagonaalmatrix D zo dat $U_2 = U_1D$, $R_2 = D^{-1}R_1$.

Uit het gegeven volgt nl. $U_1^HU_2 = R_1R_2^{-1}$ en $U_2^HU_1 = R_2R_1^{-1}$. Daaruit volgt: als $Z := R_1R_2^{-1}$ dan is $Z^H = Z^{-1}$, dus Z unitair. Maar Z is ook rechtsboven, volgens (8.2.2.11) moet Z dus diagonaal zijn.

Stelling. (8.2.6)

Zij $A \in M^r_{m,n}$. Dan is er een links-unitaire $U_1 \in M_{m,r}$ en een rechtsreguliere $C_1 \in M_{n,r}$ zo dat $A = U_1C_1^H$.

Bewijs.

Zij $A = B_1C_1^H$ met B_1 en $C_1 \in RR$. Schrijf met Gram-Schmidt $B_1 = U_1R_{11}$ met U_1 links-unitair en R_{11} regulier. Neem $C_1' = C_1R_{11}^H$. □

Opgave. (8.2.7)

Onderzoek de niet-eenduidigheid van U_1 en C_1.

Opgave. (8.2.8)

Zij $A \in M_{m,n}^r$. Bewijs dat er een permutatiematrix P, een links-unitaire matrix U_1 en een rechtsbovenmatrix R_1 met positieve diagonaalelementen is zo, dat

$$AP = U_1R_1 \ .$$

Laat ook zien dat P, U_1 en R_1 gevonden kunnen worden met een Gram-Schmidt constructie met kolomverwisselingen en dat hiermee ook een constructie voor de splitsing uit (8.2.6) aangegeven is.

Opgave. (8.2.9)

Bewijs (8.2.2.5) door toepassing van (8.2.6) op de matrix $I - U_1U_1^H$. Onderzoek de niet-eenduidigheid van U_2.

Stelling. (8.2.10)

Zij $A \in M_{m,n}^r$, $A \neq 0$. Dan zijn er

$$\text{unitaire } U = (U_1 \mid U_2) \text{ met } U_1 \in M_{m,r} \text{ (en links-unitair)} \ ,$$

$$\text{unitaire } V = (V_1 \mid V_2) \text{ met } V_1 \in M_{m,r} \text{ (en links-unitair)} \ ,$$

$$Z = \left(\begin{array}{c|c} Z_{11} & 0 \\ \hline 0 & 0 \end{array}\right) \in M_{m,n} \text{ met } Z_{11} \in M_{r,r}^r \text{ (dus regulier)}$$

zo dat

$$A = UZV^H = U_1Z_{11}V_1^H \ .$$

Als A hermitisch is dan kan Z zo gekozen worden dat V = U en Z hermitisch is:

$$A = UZU^H = U_1Z_{11}U_1^H \ .$$

Bewijs.

Zij $A = B_1C_1^H$ met $B_1 \in M^r_{m,r}$, $C_1 \in M^r_{n,r}$. Volgens (8.2.4) zijn er links-unitaire U_1 en V_1 zo dat $B_1 = U_1X_{11}$, $C_1 = V_1Y_{11}$ met X_{11} en Y_{11} regulier. Hieruit volgt $A = U_1Z_{11}V_1^H$ met $Z_{11} = X_{11}Y_{11}^H$, regulier. Breid nu U_1 en V_1 volgens (8.2.2.5) uit tot unitaire U en V en vul Z_{11} aan met nullen.
Als A hermitisch is dan gaan we uit van de splitsing (8.1.5): $A = B_1Z_{11}B_1^H$ met $B_1 \in RR$ en Z_{11} hermitisch en regulier (zelf verder). □

Opgave. (8.2.11)

Als $A = U_1Z_{11}V_1^H = U_1'Z_{11}'V_1'^H$, zo als in (8.2.10), dan zijn er unitaire X_{11} en Y_{11} zo dat $U_1' = U_1X_{11}$, $V_1' = V_1Y_{11}$, $Z_{11}' = X_{11}^HZ_{11}Y_{11}$.

Opgave. (8.2.12)

A is range-hermitian (zie (8.1.11)) dan en slechts dan als

$$A = U_1Z_{11}U_1^H$$

met U_1 links-unitair en Z_{11} regulier.

Stelling. (8.2.13)

Zij A gesplitst als in (8.2.10). Dan geldt:

de kolommen van U_1 vormen een orthonormale basis voor $R(A)$,
de kolommen van U_2 vormen een orthonormale basis voor $N(A^H)$,
de kolommen van V_1 vormen een orthonormale basis voor $R(A^H)$,
de kolommen van V_2 vormen een orthonormale basis voor $N(A)$.

Bewijs.

Uit $A = U_1Z_{11}V_1^H$ volgt $R(A) = R(U_1)$ want $Z_{11}V_1^H \in RI$. En eveneens $N(A) = N(V_1^H)$, want $U_1Z_{11} \in LI$. Volgens (8.1.8) en (8.2.2.7) geldt $N(V_1^H) = (R(V_1))^\perp = R(V_2)$. Analoog de andere relaties. □

8.3. Orthogonale projectoren

In (5.5.1) werden reeds projectoren gedefinieerd en een aantal eenvoudige eigenschappen ervan genoemd. We voegen nu een extra eis, nl. hermiticiteit, toe. We krijgen dan de zgn. orthogonale projectoren (i.t.t. de algemene projectoren uit (5.5.1) die we scheef kunnen noemen).

Definitie. (8.3.1)

$E \in M_{m,m}$ heet orthogonale projector indien

i) $E^2 = E$,

ii) E is hermitisch (dus $E^H = E$).

Eenvoudige eigenschappen. (8.3.2)

Behalve de in (5.5.2) genoemde eigenschappen i) t/m vii) die gelden voor alle projectoren en deze ook karakteriseren, geldt voor de orthogonale projectoren E:

ix) E is projector en $\forall_x$: $Ex \perp (I - E)x$,

x) $(R(E))^\perp = R(I - E)$ (= $N(E)$ als E projector is),

xi) $E^H(I - E) = 0$.

Ieder van deze eigenschappen karakteriseert ook de orthogonale projectoren: als $E \in M_{m,m}$ en ix), x) of xi) geldt, dan is E orthogonale projector. Uit ix) volgt nl. door te nemen $x = y + z$ met $y \in R(E)$, $z \in R(I - E)$ dat $R(E) \perp R(I - E)$. Deze bewering, die ook direct uit x) volgt, is equivalent met xi). Maar xi) impliceert $E^H = E^HE$, dus E is hermitisch en $E = E^2$.

Opmerkingen. (8.3.3)

1. Uit (8.3.2.x) en (5.5.3.2) volgt: als E en E' orthogonale projectoren zijn en $R(E') = R(E)$ of $N(E') = N(E)$ dan is $E' = E$.
2. Eveneens volgt uit (8.3.2.x) dat een projector orthogonaal is dan en slechts dan als hij range-hermitian is (d.w.z. $R(E) = R(E^H)$, zie (8.1.11)). Men name geldt: als E projector is en normaal is (d.w.z. $EE^H = E^HE$) dan is E orthogonale projector (bewijs dit ook rechtstreeks!).
3. Als E_1 en E_2 orthogonale projectoren zijn dan geldt $E_1E_2 = 0 \Leftrightarrow E_2E_1 = 0 \Leftrightarrow R(E_1) \perp R(E_2) \Leftrightarrow E_1 + E_2$ is orthogonale projector (op $R(E_1) \oplus R(E_2)$ langs $N(E_1) \cap N(E_2)$).
   Bewijs dat dit rechtstreeks uit de definitie van orthogonale projector volgt. Vergelijk het resultaat met (5.5.7) en met (7.1.13).

Stelling. (8.3.4)

$E \neq 0$ is orthogonale projector dan en slechts dan als er een links-unitaire U_1 is zo dat $E = U_1U_1^H$.

E is orthogonale projector dan en slechts dan als er een unitaire U is en een normaalvorm J zodat $E = UJU^H$.

Bewijs.

Dat de voorwaarden voldoende zijn, is duidelijk. Als E orthogonale projector is, dan is E hermitisch en dus kan E volgens (8.2.10) geschreven worden als

$$E = U_1 Z_{11} U_1^H$$

met U_1 links-unitair en Z_{11} hermitisch en regulier. Uit $E^2 = E$ volgt direct $Z_{11}^2 = Z_{11}$ en dus $Z_{11} = I_{11}$. Uit de uitbreidingsstelling (8.2.2.5) volgt dat er een U_2 is zo dat $U = (U_1/U_2)$ unitair is en dan is $E = UJU^H$ met J de normaalvorm met $r(J) = r(E)$ (vgl. (3.1.1). Deze voorstelling geldt ook nog als $E = 0$.

Gevolgen. (8.3.5)

1) Als de kolommen van U_1 een orthonormale basis voor een deelruimte V van $\mathbb{C}^m$ vormen dan is $E = U_1 U_1^H$ de orthogonale projector op V (langs $V^\perp$). Als $(U_1 \mid U_2)$ unitair is (vgl. (8.2.2.5)) dan is $I - E = U_2 U_2^H$ de orthogonale projector op $V^\perp$ (langs V).

2) Als, zoals in (8.2.10), $A = U_1 Z_{11} V_1^H$ met U_1 en V_1 links-unitair en Z_{11} regulier, dan is, als $U = (U_1 \mid U_2)$ en $V = (V_1 \mid V_2)$ unitair zijn,

$E_1 := U_1 U_1^H$ de orthogonale projector op $R(A)$,

$I - E_1 = U_2 U_2^H$ de orthogonale projector op $N(A^H)$,

$E_2 := V_1 V_1^H$ de orthogonale projector op $R(A^H)$,

$I - E_2 = V_2 V_2^H$ de orthogonale projector op $N(A)$.

Dit volgt uit (8.2.13)).

Stelling. (8.3.6)

Zij $A \in M_{m,n}$, $E \in M_{m,m}$. De volgende vier uitspraken zijn equivalent:

i) E is de orthogonale projector op $R(A)$ langs $N(A^H)$,

ii) E is projector, $R(E) \subset R(A)$, $N(E) \subset N(A^H)$,

iii) $\exists_X$: $E = AX$, $A^H AX = A^H$,

iv) $\exists_X$: $E = AX$, $AXA = A$, AX hermitisch.

Bewijs.

i) $\Rightarrow$ ii) : triviaal.

ii) $\Rightarrow$ iii) : $R(E) \subset R(A) \Rightarrow \exists_X : E = AX$; E projector met $N(E) \subset N(A^H) \Rightarrow A^H(I - E) = 0$; combinatie levert iii).

iii) $\Rightarrow$ iv) : $A^HAX = A^H \Rightarrow X^HA^HAX = X^HA^H \Rightarrow AX$ hermitisch en $A = (A^HAX)^H = AXA$.

iv) $\Rightarrow$ i) : volgt uit (5.5.9), (8.3.1), (8.3.2.x) en (8.1.8). □

Opmerkingen. (8.3.7)

1. Iedere X die aan iii) voldoet, voldoet ook aan iv) en omgekeerd. Als voor X een rg-inverse van A gekozen wordt (dat kan volgens (5.5.10.3) dan is volgens (5.5.9) $N(X) = N(E) = N(A^H)$.
2. Uit (8.1.3.4) volgt dat de vergelijking $A^HAX = A^H$ een oplossing X heeft, uit (8.1.3.3) volgt dat AX onafhankelijk is van de gekozen oplossing; dit bewijst opnieuw existentie en eenduidigheid van E. Uit (4.3.2) volgt dat een oplossing van $A^HAX = A^H$ is

$$X = (A^HA)^-A^H$$

met willekeurige g-inverse $(A^HA)^-$. Deze X is g-inverse van A en, daar $r(X) \leq r(A^H) = r(A)$, zelfs rg-inverse, zodat volgens (8.3.7.1) $N(X) = N(A^H)$. Voor de projector E vinden we hiermee de uitdrukking

$$E = A(A^HA)^-A^H$$

die onafhankelijk van de gekozen g-inverse is. Vergelijk deze resultaten met (5.5.12)!

Opgave. (8.3.8)

Zij $A \in M_{m,n}$, $E \in M_{n,n}$. De volgende uitspraken zijn equivalent:

i) E is de orthogonale projector langs N(A) op $R(A^H)$,

ii) E is projector, $N(A) \subset N(E)$, $R(A^H) \subset R(E)$,

iii) $\exists_Y : E = YA$, $YAA^H = A^H$,

iv) $\exists_Y : E = YA$, $AYA = A$, YA hermitisch.

Voor iedere g-inverse $(AA^H)^-$ is

$$Y := A^H(AA^H)^-$$

een rg-inverse van A die voldoet aan iii) en iv) en waarvoor $R(Y) = R(A^H)$.

Voor E geldt (onafhankelijk van de keuze van $(AA^H)^-$

$$E = A^H(AA^H)^-A \ .$$

Vergelijk dit met (5.5.12).

8.4. De euclidische norm voor vectoren. Kleinste kwadraten-problemen.

Definitie. (8.4.1)

De euclidische norm van $x \in \mathbb{C}^m$ is

$$\|x\| := (x^Hx)^{\frac{1}{2}} \ .$$

Eenvoudige eigenschappen. (8.4.2)

1) $\|x\| \geq 0$, $\|x\| = 0 \Leftrightarrow x = 0$.
2) $\|\alpha x\| = |\alpha| \ \|x\|$.
3) $\|x + y\|^2 = \|x\|^2 + 2 \ \mathrm{Re}(x^Hy) + \|y\|^2$.
4) $|x^Hy| \leq \|x\| \ \|y\|$ (ongelijkheid van Cauchy, volgt uit 3)).
5) $\|x + y\| \leq \|x\| + \|y\|$ (driehoeksongelijkheid, volgt uit 3) en 4)).
6) $\mathrm{Re}(x^Hy) = 0 \Leftrightarrow \|x + y\|^2 = \|x\|^2 + \|y\|^2$ (Pythagoras).
7) $(\forall_{\alpha\in\mathbb{C}} : \|x + \alpha y\| \geq \|x\|) \Leftrightarrow x^Hy = 0$.
8) Als U_1 links-unitair is dan is $\|U_1x\| = \|x\|$.

Stelling. (8.4.3)

Zij $E \in M_{m,m}$ orthogonale projector. Dan geldt

$$\forall_{x\in\mathbb{C}^m} : \|Ex\| \leq \|x\|$$

met gelijkteken dan en slechts dan als $x \in R(E)$.

Opgave. (8.4.4)

Als $E \in M_{m,m}$ projector is en

$$\forall_{x\in\mathbb{C}^m} : \|Ex\| \leq \|x\|$$

dan is E orthogonale projector.

We bespreken nu een aantal onderling samenhangende minimaliseringsopgaven van het zg. kleinste-kwadraten'type.

Stelling. (8.4.5)

Zij $V_1 \subset \mathbb{C}^m$, E_1 de orthogonale projector op V_1. Dan geldt voor iedere $z \in \mathbb{C}^m$

$$\min\{\|z - y\| \mid y \in V_1\} = \|(I - E_1)z\|$$

en het minimum wordt bereikt dan en slechts dan als $y = E_1 z$.

Bewijs.

Voor alle toegelaten y is $y = E_1 y$ en dus volgt met Pythagoras

$$\|z - y\|^2 = \|(I - E_1)(z - y)\|^2 + \|E_1(z - y)\|^2 =$$

$$= \|(I - E_1)z\|^2 + \|E_1 z - y\|^2 \geq \|(I - E_1)z\|^2 ,$$

met gelijkteken dan en slechts dan als $y = E_1 z$ (deze y is toegelaten). □

Stelling. (8.4.6)

Zij $V_2 \subset \mathbb{C}^n$, E_2 de orthogonale projector langs V_2. Dan geldt voor iedere $z \in \mathbb{C}^n$

$$\min\{\|x\| \mid x - z \in V_2\} = \|E_2 z\|$$

en het minimum wordt bereikt dan en slechts dan als $x = E_2 z$.

Bewijs.

Voor de toegelaten x geldt $E_2(x - z) = 0$ en dus is

$$\|x\|^2 = \|E_2 x\|^2 + \|(I - E_2)x\|^2 = \|E_2 z\|^2 + \|x - E_2 z\|^2 \geq \|E_2 z\|^2$$

met gelijkteken dan en slechts dan als $x = E_2 z$ (deze x is toegelaten). □

Stelling. (8.4.7)

Zij $A \in M_{m,n}$, E_1 de orthogonale projector op $R(A)$ en A^- een g-inverse van A. Dan geldt voor iedere $z \in \mathbb{C}^m$

$$\min\{\|z - Ax\| \mid x \in \mathbb{C}^n\} = \|(I - E_1)z\|$$

en het minimum wordt bereikt dan en slechts dan als $Ax = E_1 z$, of wel: als $x - A^- E_1 z \in N(A)$.

Bewijs.

Pas (8.4.5) toe met $V_1 = R(A)$. Voor de minimaliserende x moet dan $Ax = E_1 z$. Daar $AA^- E_1 = E_1$ (zie (5.5.10.1) is dit equivalent met $x - A^- E_1 z \in N(A)$. □

Stelling. (8.4.8)

Zij $A \in M_{m,n}$, $E_2 = YA$ de orthogonale projector langs N(A). Dan geldt voor iedere $z \in \mathbb{C}^n$

$$\min\{\|x\| \mid x - z \in N(A)\} = \|E_2 z\|$$

en het minimum wordt bereikt dan en slechts dan als $x = E_2 z$.

Bewijs.

Pas (8.4.6) toe met $V_2 = N(A)$. Voor de minimaliserende x geldt dan $x = E_2 z$. □

Stelling. (8.4.9)

Zij $A \in M_{m,n}$, E_1 de orthogonale projector op R(A), E_2 de orthogonale projector langs N(A), A^- een g-inverse van A en

$$A^+ := E_2 A^- E_1 \ . \tag{8.4.10}$$

Dan geldt voor iedere $z \in \mathbb{C}^m$

$$\min\{\|\hat{x}\| \mid \hat{x} \in \mathbb{C}^n, \|z - A\hat{x}\| = \min\{\|z - Ax\| \mid x \in \mathbb{C}^n\}\} = \|A^+ z\|$$

en het minimum wordt bereikt dan en slechts dan als $\hat{x} = A^+ z$.

Dat wil zeggen: onder alle $\hat{x}$ die $\|z - Ax\|$ minimaliseren heeft $A^+ z$ de kleinste norm.

Bewijs.

Volgens (8.4.7) minimaliseert $\hat{x}$ de uitdrukking $\|z - Ax\|$ dan en slechts dan als $\hat{x} - A^- E_1 z \in N(A)$. En onder deze nevenconditie is volgens (8.4.8) $\|\hat{x}\|$ minimaal dan en slechts dan als $\hat{x} = E_2 A^- E_1 z = A^+ z$. □

Opmerkingen. (8.4.11)

1. Volgens (5.5.11) is de door (8.4.10) gedefinieerde $A^+ \in M_{n,m}$ de eenduidig bepaalde matrix waarvoor geldt

$$\begin{aligned} &A^+ \text{ is rg-inverse van } A, \\ &AA^+ = E_1 \text{ (= de orthogonale projector op } R(A)), \\ &A^+A = E_2 \text{ (= de orthogonale projector langs } N(A)) \;. \end{aligned} \tag{8.4.12}$$

Volgens (5.5.9) is (ga na) aan (8.4.12) voldaan dan en slechts dan als

$$\begin{aligned} &A^+ \text{ is rg-inverse van } A, \\ &AA^+ \text{ is hermitisch}, \\ &A^+A \text{ is hermitisch} \;. \end{aligned} \tag{8.4.13}$$

Anderzijds volgt uit (5.5.11) dat aan (8.4.12) voldaan is zodra A^+ g-inverse van A is met $R(A^+) = R(E_2)$, $N(A^+) = N(E_1)$. Daar

$$R(E_2) = (N(E_2))^\perp = (N(A))^\perp = R(A^H) \;,$$

$$N(E_1) = (R(E_1))^\perp = (R(A))^\perp = N(A^H) \;,$$

is dus (8.4.12) ook equivalent met

$$\begin{aligned} &A^+ \text{ is g-inverse van } A, \\ &R(A^+) = R(A^H) \;,\; N(A^+) = N(A^H) \;. \end{aligned} \tag{8.4.14}$$

De door (8.4.12), (8.4.13) of (8.4.14) eenduidig bepaalde A^+ heet pseudo-inverse van A en wordt in §9 nader besproken.

2. Daar volgens (8.4.14) $R(A^+) = R(A^H)$, $N(A^+) = N(A^H)$ is er volgens (5.4.22) een Z zo dat $A^+ = A^H Z A^H$. Daar A^+ g-inverse van A is geldt dan

$$A^HAA^H \cdot Z \cdot A^HAA^H = A^HA \cdot A^+ \cdot AA^H = A^HAA^H \;,$$

zodat Z g-inverse van A^HAA^H is en

$$A^+ = A^H(A^HAA^H)^-A^H \;. \tag{8.4.15}$$

Omgekeerd geldt, als we het rechterlid van (8.4.15) X noemen, dat $A^HAXAA^H = A^HAA^H$ en hieruit volgt met (8.1.3) en het feit dat $R(X) \subset R(A^H)$, $N(X) \supset N(A^H)$ eenvoudig dat X g-inverse van A is met $R(X) = R(A^H)$, $N(X) = N(A^H)$. Voor iedere keuze van $(A^HAA^H)^-$ levert het resultaat van (8.4.15) dus de door (8.4.14) eenduidig bepaalde A^+. Vergelijk dit resultaat met (5.5.18)!

8.5. Inproduct en norm voor matrices.

We kunnen $M_{m,n}(\mathbb{C})$ opvatten als een vectorruimte die isomorf is met $\mathbb{C}^{mn}$ en deze ruimte unitair maken door een voor de hand liggen inproduct te definiëren. Tevens vinden we zo een euclidische norm voor matrices. Het blijkt dat inproduct en norm ook nuttige eigenschappen hebben t.a.v. de matrixvermenigvuldiging.

Definitie. (8.5.1)

Het inproduct van $A = (\alpha_{ij})$ en $B = (\beta_{ij}) \in M_{m,n}(\mathbb{C})$ is

$$\langle A,B\rangle := \sum_{i,j} \overline{\alpha_{ij}}\, \beta_{ij} .$$

Eenvoudige eigenschappen. (8.5.2)

1. $\langle A,B\rangle = \mathrm{tr}(A^H B) = \mathrm{tr}(BA^H)$ (zie §1.5) .
2. $\langle A,B\rangle = \langle B^H, A^H\rangle = \overline{\langle B,A\rangle}$.
3. $\langle A,A\rangle$ is reëel en niet negatief; $\langle A,A\rangle = 0 \Leftrightarrow A = 0$.
4. $\langle C,\alpha A + \beta B\rangle = \alpha\langle C,A\rangle + \beta\langle C,B\rangle$.
5. $\langle A,BC\rangle = \langle AC^H,B\rangle = \langle B^H A,C\rangle$ (dit volgt uit 1) .
6. Als U_1 en V_1 linksunitair zijn dan is

$$\langle U_1 A V_1^H,\ U_1 B V_1^H\rangle = \langle A,B\rangle .$$

7. Als $A = (a_1 \mid \dots \mid a_n)$, $B = (b_1 \mid \dots \mid b_n)$ dan is

$$\langle A,B\rangle = \sum_j a_j^H b_j .$$

(gevolg: de isomorfie tussen $M_{m,1}(\mathbb{C})$ en $\mathbb{C}^m$ geldt ook voor het inproduct).

Definitie. (8.5.3)

De euclidische norm van $A \in M_{m,n}(\mathbb{C})$ is

$$\|A\| := \langle A,A\rangle^{\frac{1}{2}} = \Big(\sum_{i,j} |\alpha_{ij}|^2\Big)^{\frac{1}{2}} .$$

Eenvoudige eigenschappen. (8.5.4)

1. $\|A\| \geq 0$, $\|A\| = 0 \Leftrightarrow A = 0$.
2. $\|\alpha A\| = |\alpha|\ \|A\|$.
3. $\|A + B\|^2 = \|A\|^2 + 2\,\mathrm{Re}(\langle A,B\rangle) + \|B\|^2$.
4. $|\langle A,B\rangle| \leq \|A\|\,\|B\|$.
5. $\|A + B\| \leq \|A\| + \|B\|$.
6. $A^H B = 0 \Rightarrow \|A + B\|^2 = \|A\|^2 + \|B\|^2$.
7. $\|A\|^2 = \|A^H\|^2 = \mathrm{tr}(A^H A) = \mathrm{tr}(AA^H)$.
8. Voor een normaalvorm J is $\|J\|^2 = r(J)$; voor de m×m eenheidsmatrix I is $\|I\|^2 =$

9. Als U_1 links-unitair is dan is $\|U_1\|^2 = r(U_1)$.
10. Als E orthogonale projector is dan is $\|E\|^2 = tr(E) = r(E)$.
11. Als U_1 en V_1 links-unitair zijn dan is $\|U_1AV_1^H\| = \|A\|$.
12. Als $A = (a_1 \mid \dots \mid a_n)$ dan is $\|A\|^2 = \Sigma\|a_j\|^2$ (met name geldt dit als n = 1!).
13. Als $D = \mathrm{diag}(\delta_1,\dots,\delta_n)$ diagonaalmatrix is dan is

    $\|AD\| \le \|A\| \max (|\delta_1|,\dots,|\delta_n|)$.
14. $\|AB\| \le \|A\|\|B\|$.

Uit 1., 2. en 5. volgt dat de afbeelding $\|\ \| : M_{m,n}(\mathbb{C}) \to \mathbb{R}^+$ een norm is in $M_{m,n}(\mathbb{C})$; op grond van 14. noemen we deze norm submultiplicatief.

Opgave (8.5.5)

Zij

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

normaal (d.w.z. $A^HA = AA^H$) met vierkante A_{11} en A_{22}. Dan geldt

$$\|A_{21}\| = \|A_{12}\| .$$

Met name geldt dit als A unitair is.

Opgave (8.5.6)

Zij E_1 en $E_2 \in M_{m,m}$ orthogonale projectoren met dezelfde rang. Dan is

$$\|E_1(I - E_2)\| = \|(I - E_1)E_2\| .$$

Hint: Stel $E_1 = UJU^H$, $E_2 = VJV^H$ met U en V unitair en J een normaalvorm uit $M_{m,m}$.

Opgave (8.5.7)

Zij $A \in M_{m,n}$. Zij $U_1 \in M_{m,p}$ en $V_1 \in M_{n,q}$ links-unitair. Dan is

$$\|U_1^HAV_1\| \le \|A\|$$

met gelijkteken dan en slechts dan als $R(A) \subset R(U_1)$ en $R(A^H) \subset R(V_1)$.

<u>Opgave</u> (8.5.8)

Bewijs (8.1.7) door $\| A - A^H \|^2$ te bepalen.

9. De pseudo-inverse.

9.1. Inleiding.

In 4.3 en 5.5 bespraken we reflexieve gegeneraliseerde inversen van een matrix $A \in M_{m,n}$. Dat zijn matrices $X \in M_{n,m}$ die voldoen aan

$$AXA = A,\ XAX = X\ . \tag{9.1.1}$$

Het bleek dat $E_1 := AX$ en $E_2 := XA$ dan de (scheve) projectoren zijn die bepaald zijn door de eisen

E_1 is projector op $R(A)$ langs $N(X)$,

E_2 is projector op $R(X)$ langs $N(A)$.

X, E_1 en E_2 zijn pas eenduidig bepaald door de keuze van $N(X)$ (die moet voldoen aan $R(A) \oplus N(X) = \mathcal{L}^m$) en de keuze van $R(X)$ (die moet voldoen aan $N(A) \oplus R(X) = \mathcal{L}^n$).

Hieruit volgt direct dat, als $\mathcal{L} = \mathbb{C}$, er een eenduidig bepaalde X is die aan (9.1.1) voldoet en waarvoor E_1 en E_2 orthogonale projectoren zijn. Volgens (8.3.2) en (8.1.8) zijn deze extra-eisen namelijk equivalent met de eisen dat

$$N(X) = N(E_1) = (R(E_1))^{\perp} = (R(A))^{\perp} = N(A^H)\ ,$$

$$R(X) = R(E_2) = (N(E_2))^{\perp} = (N(A))^{\perp} = R(A^H)\ .$$

En er geldt inderdaad $R(A) \oplus N(A^H) = \mathbb{C}^m$, $N(A) \oplus R(A^H) = \mathbb{C}^n$. Deze aldus eenduidig bepaalde gegeneraliseerde inverse X heet pseudo-inverse van A (ook wel: Moore-Penrose inverse van A) en wordt aangeduid met A^+.

In §8.4 bleek dat deze pseudo-inverse belangrijk is bij de oplossing van zg. kleinste kwadraten problemen.

We bespreken in het volgende de pseudo-inverse zonder de resultaten uit 4.3 en 5.5 te gebruiken.

9.2. Definitie en eenvoudige eigenschappen

Stelling. (9.2.1)

Bij iedere $A \in M_{m,n}$ is er precies één $X \in M_{n,m}$ die voldoet aan de zg. Penrose-condities

$$AXA = A\ , \tag{9.2.2}$$

$$XAX = X\ , \tag{9.2.3}$$

$$AX \text{ is hermitisch}\ , \tag{9.2.4}$$

$$XA \text{ is hermitisch}\ . \tag{9.2.5}$$

Bewijs.

Als $A = 0$ dan volgt uit (9.2.3) dat $X = 0$ moet zijn, deze voldoet ook aan de andere eisen.
Zij nu $A \neq 0$ en stel, conform (8.2.10),

$$A = U_1Z_{11}V_1^H$$

met U_1 en V_1 links-unitair en Z_{11} regulier.
Veronderstel eerst dat er een X bestaat die aan (9.2.2) t/m (9.2.5) voldoet.
Uit de laatste drie eisen volgt dat

$$X = XA.X.AX = A^HX^HXX^HA^H\ ,$$

dus X is van de vorm $X = V_1W_{11}U_1^H$. Substitutie in (9.2.2) levert
$U_1(Z_{11}W_{11}Z_{11} - Z_{11})V_1^H = 0$ en dus, daar U_1 en $V_1 \in RR$, $Z_{11} \in I$, moet $W_{11} = Z_{11}^{-1}$.
Conclusie: als X bestaat dan geldt noodzakelijk

$$X = V_1Z_{11}^{-1}U_1^H\ . \tag{9.2.6}$$

Omgekeerd geldt echter: als X door (9.2.6) gedefinieerd wordt dan is

$$AX = U_1U_1^H,\ XA = V_1V_1^H$$

en daaruit volgt eenvoudig dat X aan de eisen (9.2.2) t/m (9.2.5) voldoet.
Hiermee zijn eenduidigheid en existentie aangetoond (waarom volgt uit het bovenstaande dat X onafhankelijk is van de gekozen splitsing van A?). □

Opmerking. (9.2.7)

De eenduidigheid van X kunnen we ook als volgt bewijzen.
Stel dat X en X' voldoen aan (9.2.2) t/m (9.2.5). Volgens (8.3.6) zijn (9.2.2) en (9.2.4) equivalent met de eis $A^H AX = A^H$. Daaruit volgt $A^H A(X - X') = 0$, dus $A(X - X') = 0$ (waarom?). Analoog $(X - X')A = 0$. Maar dan volgt uit (9.2.3) dat $X' = X'AX' = XAX = X$.

Opgave. (9.2.8)

Stel, conform (8.2.10), $A = UZV^H$ met U en V unitair en

$$Z = \left(\begin{array}{c|c} Z_{11} & 0 \\ \hline 0 & 0 \end{array}\right),$$

waarin $Z_{11} \in M^r_{r,r}$. Stel $X = VWU^H$ met

$$W = \left(\begin{array}{c|c} W_{11} & W_{12} \\ \hline W_{21} & W_{22} \end{array}\right),$$

waarin $W_{11} \in M_{r,r}$.
Vertaal nu de eisen (9.2.2) t/m (9.2.5) in eisen voor W en bepaal daaruit W. Ga ook na welke vrijheid voor W er over blijft als we één of meer der eisen laten vallen.

Definitie. (9.2.9)

De bij een $A \in M_{m,n}$ door de eisen (9.2.2) t/m (9.2.5) eenduidig bepaalde matrix $A^+ := X \in M_{n,m}$ heet pseudo-inverse van A (ook wel: Moore-Penrose inverse).

De relaties die A^+ definiëren kunnen op vele wijzen gecombineerd worden tot nieuwe identiteiten. Nuttig zijn b.v.

$$A = AA^+A = AA^HA^{+H} = A^{+H}A^HA, \tag{9.2.10}$$

$$A^+ = A^+AA^+ = A^+A^{+H}A^H = A^HA^{+H}A^+, \tag{9.2.11}$$

$$A^H = A^+AA^H = A^HAA^+, \tag{9.2.12}$$

$$A^{+H} = AA^+A^{+H} = A^{+H}A^+A. \tag{9.2.13}$$

Eenvoudige eigenschappen (9.2.14)

1) $r(A^+) = r(A)$,
2) $(A^+)^+ = A$,
3) $(A^H)^+ = (A^+)^H$ (vandaar de notatie A^{+H}),
4) $(A^HA)^+ = A^+A^{+H}$ (ga na),
5) $R(A^+) = R(A^H) = (N(A))^\perp$ (volgt uit (9.2.11) en (9.2.12)),
6) $N(A^+) = N(A^H) = (R(A))^\perp$ (idem),
7) AA^+ is de orthogonale projector op $R(A)$ langs $N(A^H) = N(A^+)$,
8) A^+A is de orthogonale projector op $R(A^H) = R(A^+)$ langs $N(A)$,
9) De afbeelding A: $R(A^H) \to R(A)$ is een bijectie met inverse A^+: $R(A) \to R(A^H)$.
10) Als A hermitisch is dan is A^+ hermitisch en $AA^+ = A^+A$ is de orthogonale projector op $R(A) = R(A^+)$ langs $N(A) = N(A^+)$.

Vergelijk de eigenschappen 5) t/m 9) met de resultaten uit (5.5.15) en met de daarbij behorende tekening.

9.3. Uitdrukkingen voor A^+ in eenvoudige gevallen

Er geldt: (9.3.1)

1) A regulier: $A^+ = A^{-1}$,
2) A of A^H links-unitair: $A^+ = A^H$,
3) A orthogonale projector: $A^+ = A$,
4) $A \in RR$: $A^+ = (A^HA)^{-1}A^H$,
5) $A \in LR$: $A^+ = A^H(AA^H)^{-1}$,
6) $A = B_1C_1$, $B_1 \in RR$, $C_1 \in LR$: $A^+ = C_1^+B_1^+$,
7) $A = U_1Z_1V_1^H$, U_1 en V_1 links-unitair: $A^+ = V_1Z_1^+U_1^H$.

Verifieer deze uitspraken m.b.v. de definities of de daaruit afgeleide relaties (uitspraak 4 volgt b.v. onmiddellijk uit (9.2.12)).

In het algemeen geldt niet $(AB)^+ = B^+A^+$. Nodige en voldoende voorwaarden opdat dat geldt zijn aangegeven door Greville:

Stelling. (9.3.2)

Er geldt $(AB)^+ = B^+A^+$ dan en slechts dan als

$$(I - A^+A)BB^HA^+A = 0 \ , \tag{9.3.3}$$

$$(I - BB^+)A^HABB^+ = 0 \ , \tag{9.3.4}$$

of equivalent: $A((R(B))^\perp) \perp A(R(B))$, $B^H((R(A^H))^\perp) \perp B^H(R(A^H))$.

Bewijs.

Zij $C := AB$, $X := B^+A^+$.

Uit (9.2.12) en (9.2.5) volgt dat (9.3.3) equivalent is met

$$0 = (I - A^+A)BB^HA^H = (I - A^+A)BC^H \ . \tag{9.3.5}$$

We bewijzen dat dit equivalent is met

$$(I - XC)C^H = 0 \ . \tag{9.3.6}$$

Immers, uit (9.3.5) volgt dat $A^+ABC^H = BC^H$, zodat

$$(I - XC)C^H = (I - B^+A^+AB)C^H = (I - B^+B)C^H = 0 \ ,$$

omdat volgens (9.2.12) $(I - B^+B)B^H = 0$.

En uit (9.3.6) volgt door vermenigvuldiging van links met B^HB

$$0 = B^HB(I - B^+A^+AB)C^H = B^H(I - A^+A)BC^H \ .$$

Daar $I - A^+A = (I - A^+A)^H(I - A^+A)$ volgt hieruit met (8.1.3.4) dat (9.3.5) geldt.

Geheel analoog blijkt dat (9.3.4) equivalent is met

$$C^H(I - CX) = 0 \ . \tag{9.3.7}$$

Volgens (8.3.8) en (8.3.6) zijn (9.3.6) en (9.3.7) samen equivalent met drie van de vier Penrose-condities, nl.

$$CXC = C, \ CX \text{ en } XC \text{ hermitisch} \ .$$

Hieruit volgt al dat de voorwaarden van de stelling nodig zijn. Om te bewijzen dat ze ook voldoende zijn moeten we nog aantonen dat ze implicieren dat $XCX = X$.

Nu volgde reeds uit de voorwaarden dat $CXC = C$ en dus

$$\begin{aligned} 0 &= (C - CXC)^H = (A(I - BB^+A^+A)B)^H \\ &= B^H(I - A^+ABB^+)A^H \ . \end{aligned}$$

Met (9.2.11) volgt hieruit dat dan ook

$$0 = B^+(I - A^+ABB^+)A^+ = X - XCX \ . \qquad \square$$

Opgave. (9.3.8)

Zij A en B zo dat AB gedefinieerd is. Zij

$$B_1 := A^+AB,\ A_1 := AB_1B_1^+ .$$

Dan is (Cline)

$$AB = A_1B_1,\ (AB)^+ = B_1^+A_1^+ .$$

Hint: bewijs eerst dat, als $AB = C$, $R(C) = R(A_1)$, $N(C) = N(B_1)$ en dus $CC^+ = A_1A_1^+$, $C^+C = B_1^+B_1$.

Opgave. (9.3.9)

Bewijs rechtstreeks dat $(A^HA)^+ = A^+A^{+H}$ en laat zien dat dit ook uit (9.3.2) of uit (9.3.8) volgt.

9.4. Andere karakteriseringen van A^+

We geven nu enkele, min of meer meetkundig geformuleerde, karakteriseringen van A^+.

Stelling. (9.4.1)

Zij AX de orthogonale projector op R(A) en XA de orthogonale projector op R(X). Dan is $X = A^+$.

Bewijs.

Uit (8.3.6) volgt dat deze eisen geheel equivalent zijn met de definiërende eisen (9.2.2) t/m (9.2.5). □

Stelling. (9.4.2)

Zij AX de orthogonale projector op R(A) en XA de orthogonale projector langs N(A). Dan geldt: $r(X) = r(A) \Leftrightarrow X = A^+$.

Bewijs.

Stel $r(X) = r(A)$. Uit $R(A^H) = R(XA) \subset R(X)$ volgt dan $R(A^H) = R(X)$ en hiermee is de bewering teruggevoerd tot (9.4.1). De omkering staat in (9.2.14.1). □

Stelling. (9.4.3)

Zij AX de orthogonale projector op R(A) en $R(X) \subset R(A^H)$. Dan is $X = A^+$.

Bewijs.

Als AX projector is op $R(A)$ dan is $AXA = A$ en dus $R(I - XA) = N(A)$ (ga nog eens na). Uit $R(X) \subset R(A^H)$ volgt $N(X^H) \supset N(A)$ en dus $X^H(I - XA) = 0$ of $X^H XA = X^H$. Uit (8.3.6) volgt dan dat XA orthogonale projector is op $R(X)$ en hiermee is de bewering teruggevoerd tot (9.4.1). □

Opmerkingen. (9.4.4)

1) Een variant op deze stelling is: Zij X oplossing van $A^H AX = A^H$ en $R(X) \subset R(A^H)$ (of, equivalent, $R(X) \perp N(A)$). Dan is $X = A^+$. Dit volgt uit (8.3.6).
2) Een variant op het bewijs van (9.4.3), die gebruik maakt van existentie en eigenschappen van A^+, is: Als AX de orthogonale projector op $R(A)$ is dan is $AX = AA^+$. Uit $R(X) \subset R(A^H)$ volgt (daar A^+A projector is op $R(A^H)$) $X = A^+AX$, zodat $X = A^+AA^+ = A^+$.

   Bedenk dergelijke varianten ook voor de andere bewijzen uit deze paragraaf.

Stelling. (9.4.5)

Zij AX orthogonale projector en $R(X) = R(A^H)$. Dan is $X = A^+$.

Bewijs.

Als AX projector is dan is $(I - AX)AX = 0$; $R(A^H) \subset R(X)$ impliceert dan $(I - AX)AA^H = 0$ dus (met (8.1.3)) $(I - AX)A = 0$, dus $R(AX) = R(A)$.
De rest volgt nu uit (9.4.3). □

Stelling. (9.4.6)

Zij AX projector op $R(A)$ en zij $R(X) \subset R(A^H)$ en $N(X) \supset N(A^H)$. Dan is $X = A^+$.

Bewijs.

Als AX projector is op $R(A)$ dan is $AXA = A$ en dus (ga na) $N(A) = R(I - XA)$, $R(A) = N(I - AX)$. Uit $R(X) \subset R(A^H)$ volgt $N(X^H) \supset N(A)$ en dus $X^H(I - XA) = 0$. Uit $N(X) \supset N(A^H)$ volgt $R(A^H) \subset R(A)$ en dus $(I - AX)X^H = 0$. Hermiticiteit van XA en AX volgen nu direct, evenals $XAX = X$. □

Stelling. (9.4.7)

Zij AX de orthogonale projector op R(A) en YA de orthogonale projector langs N(A). Dan is $YAX = A^+$ en $AX = AA^+$, $YA = A^+A$.

Bewijs.

Uit AXA = A en AYA = A volgt voor Z := YAX dat AZ = AX, dus AZ hermitisch; eveneens ZA = YA dus ZA hermitisch. Verder is AZA = AYAXA = A en ZAZ = YAXAYAX = YAX = Z. □

Opmerking. (9.4.8)

Ga na dat, hoewel X en Y i.h.a. niet eenduidig bepaald zijn, AX, YA en Z dat wel zijn. Leid zo de eenduidigheid van A^+ af. Leid ook de existentie van A^+ af uit (8.3.6), (8.3.8) en (9.4.7).

Opgave. (9.4.9)

Zij $X := (A^HA)^-A^H$, $Y := A^H(AA^H)^-$ met willekeurige g-inversen. Dan is X rg-inverse van A met $N(X) = N(A^H)$ en Y is rg-inverse van A met $R(Y) = R(A^H)$. Hieruit volgt dat AX de orthogonale projector op R(A) en YA de orthogonale projector op $R(A^H)$ is en dat $YAX = A^+$. Ga na dat dit alles volgt uit (5.5.12) maar geef ook directe bewijzen door toepassing van (8.1.3).
Bewijs ook dat

$$(A^HA)^+A^H = A^H(AA^H)^+ = A^+ .$$

Opgave. (9.4.10)

Bewijs (bv. met (4.2.9)) dat de vergelijking $A^HAA^HZ = A^H$ een oplossing heeft en dat voor iedere oplossing Z geldt $A^HZ = A^+$.
Bewijs ook dat voor iedere g-inverse $(A^HAA^H)^-$ geldt

$$A^H(A^HAA^H)^-A^H = A^+ .$$

Opgave. (9.4.11)

Zij $A = (A_1 \mid A_2)$. Dan geldt

$$A_1^HA_2 = 0 \Leftrightarrow A^+ = \begin{pmatrix} A_1^+ \\ \hline A_2^+ \end{pmatrix} .$$

Opgave. (9.4.12)

Zij $A = (A_1 \mid a)$ met $a \in M_{m,1}$. Dan is

$$A^+ = \left(\begin{array}{c} A_1^+ - xy^H \\ \hline y^H \end{array}\right),$$

waarin $x = A_1^+ a$ en

$$y^H = \begin{cases} (1 + x^H x)^{-1} x^H A_1^+ & \text{als } a \in R(A_1) \\ r^+ := (r^H r)^{-1} r^H, \text{ met } r = (I - A_1 A_1^+)a & \text{als } a \notin R(A_1). \end{cases}$$

Druk $E := AA^+$ uit met behulp van $E_1 := A_1 A_1^+$ en a. Bepaal ook A^+A.

Opgave. (9.4.13)

Zij

$$A = \left(\begin{array}{c|c} A_{11} & \\ \hline A_{21} & A_{22} \end{array}\right).$$

Bewijs dat A^+ linker blokdriehoeksvorm heeft dan en slechts dan als

$$A_{21}(I - A_{11}^+ A_{11}) = 0, \quad (I - A_{22}A_{22}^+)A_{21} = 0$$

en zo ja dat dan

$$A^+ = \left(\begin{array}{c|c} A_{11}^+ & \\ \hline -A_{22}^+ A_{21} A_{11}^+ & A_{22}^+ \end{array}\right).$$

Bepaal, in het geval dat $A_{22} \in M_{1,1}$, A^+ ook als niet aan de condities voldaan is.

Opgave. (9.4.14)

Als A vierkant is dan geldt $AA^+ = A^+A$ dan en slechts dan als A "range-hermitian" (dwz. $R(A) = R(A^H) = (N(A))^{\perp}$, zie (8.1.12)) is.

9.5. Verband met kleinste kwadraten.

Stelling. (9.5.1)

Zij $z \in \mathbb{C}^m$ en $A \in M_{m,n}$.
De functie $f: \mathbb{C}^n \to \mathbb{R}$, gedefinieerd door

$$f(x) = \| z - Ax \|^2 ,$$

is minimaal dan en slechts dan als $x - A^+z \in N(A)$. Voor alle minimaliserende $x \neq A^+z$ geldt $\|x\| > \|A^+z\|$ (dus onder alle minimaliserende x heeft A^+z de kleinste norm).

Bewijs.

Daar $z - Ax = (I - AA^+)z + A(A^+z - x)$ en $(I - AA^+)^H A = 0$ volgt uit (8.4.2.6) dat

$$f(x) = \| (I - AA^+)z \|^2 + \| A(A^+z - x) \|^2 .$$

Dit is minimaal dan en slechts dan als $A(A^+z - x) = 0$, dus als $x - A^+z \in N(A)$.
Hieruit volgt $x = A^+z + (I - A^+A)x$ en dus (daar $(I - A^+A)^H A^+ = 0$)

$$\|x\|^2 = \|A^+z\|^2 + \|(I - A^+A)x\|^2 .$$

Dit is minimaal dan en slechts dan als $(I - A^+A)x$, dus als $x = A^+z$. □

Opmerking. (9.5.2)

In feite herhalen we hier het bewijs van (8.4.10). Daar we nu over de pseudo-inverse A^+ beschikken kan dit zeer kort geformuleerd worden.
Merk op dat we in het bewijs gebruikt hebben dat A^+ aan alle vier de Penrose-condities (9.2.2) t/m (9.2.5) voldoet. Bij beperkter minimaliseringsopgaven, zoals (8.4.8) en (8.4.9), kan gewerkt worden met g-inversen die slechts aan een deel van de Penrose-condities voldoen.

Opgave. (9.5.3)

$$A^+ = \lim_{\alpha \downarrow 0} (\alpha I + A^H A)^{-1} A^H = \lim_{\alpha \downarrow 0} A^H (\alpha I + AA^H)^{-1} ,$$

$$I - A^+A = \lim_{\alpha \downarrow 0} (I + \alpha^{-1} A^H A)^{-1}, \quad I - AA^+ = \lim_{\alpha \downarrow 0} (I + \alpha^{-1} AA^H)^{-1}$$

(met elementsgewijze limiet).
Hint: merk op dat

$$(\alpha I + A^H A)A^+ = A^H (I + \alpha A^{+H} A^+)$$

en dat $\alpha I + A^H A$ en $I + \alpha A^{+H} A^+ \in I$ voor $\alpha > 0$ (waarom?).

<u>Opgave.</u> (9.5.4)

Zij $A \in M_{m,n}$, $z \in \mathbb{C}^m$.

a) Zij $\alpha > 0$. Het minimum van

$$\|z - Ax\|^2 + \alpha^2\|x\|^2$$

wordt bereikt dan en slechts dan als

$$x = x_\alpha := (A^HA + \alpha^2 I)^{-1}A^Hz = A^H(AA^H + \alpha^2 I)^{-1}z .$$

Hint: minimaliseer

$$\| \binom{z}{0} - \binom{A}{\alpha I}x\|^2 .$$

b) Met x_α als boven geldt

$$\|z - Ax_\alpha\| = \min\{\|z - Ax\| \mid \|x\| \le \|x_\alpha\|\} ,$$

$$\|x_\alpha\| = \min\{\|x\| \mid \|z - Ax\| \le \|z - Ax_\alpha\|\} .$$

c) Met x_α als boven geldt: $\|x_\alpha\|$ gaat als functie van α monotoon van $\|A^+z\|$ naar 0 als α van 0 naar ∞ gaat, $\|z - Ax_\alpha\|$ gaat monotoon van $\|(I - AA^+)z\|$ naar $\|z\|$.

d) Zij $\rho > 0$. Definieer

$$\hat{x}_\rho := \begin{cases} x_\alpha \text{ met } \alpha \text{ zo dat } \|x_\alpha\| = \rho \text{ als } \|A^+z\| > \rho \\ A^+z \text{ als } \|A^+z\| \le \rho . \end{cases}$$

Dan geldt

$$\|z - A\hat{x}_\rho\| = \min\{\|z - Ax\| \mid \|x\| \le \rho\} .$$

e) Zij $\varepsilon > 0$. Definieer

$$\hat{x}_\varepsilon := \begin{cases} x_\alpha \text{ met } \alpha \text{ zo dat } \|z - Ax_\alpha\| = \varepsilon \text{ als } \|(I - AA^+)z\| < \varepsilon \\ A^+z \text{ als } \|(I - AA^+)z\| \ge \varepsilon . \end{cases}$$

Dan geldt

$$\|\hat{x}_\varepsilon\| = \min\{\|x\| \mid \|z - Ax\| \le \max(\varepsilon, \|(I - AA^+)z\|)\} .$$

Opgave. (9.5.5)

Zij $A \in M_{m,n}$ en $f : M_{n,m} \to \mathbb{R}$ gedefinieerd door

$$f(X) = \| I - AX \|^2$$

met de in (8.5.3) gedefinieerde euclidische norm voor matrices.

Bewijs dat:

1) X minimaliseert $f \Leftrightarrow X = A^+ + Y$ met $AY = 0$,

2) voor alle minimaliserende $X \neq A^+$ geldt $\| X \| > \| A^+ \|$.

10. Hermitische matrices

10.1. Splitsingen

We herinneren aan de definitie (8.1.4) van hermitische matrices en aan de voor hermitische matrices geldende splitsingen

$$A = B_1 Z_{11} B_1^H , \qquad (8.1.5)$$

$$A = U_1 Z'_{11} U_1^H , \qquad (8.2.10)$$

met $B_1 \in RR$, U_1 links-unitair, Z_{11} en Z'_{11} regulier en hermitisch.
In dit hoofdstuk zal o.a. blijken dat we de splitsing (8.1.5) nog kunnen vereenvoudigen: we kunnen B_1 zo kiezen dat Z_{11} een diagonaalmatrix is waarvan de diagonaalelementen +1 of -1 zijn. In het belangrijke deelgeval dat A positief definiet is kunnen we B_1 zelfs zo kiezen dat Z_{11} de eenheidsmatrix is.
Later (in 15.2) zullen we bewijzen dat ook (8.2.10) vereenvoudigd kan worden. U_1 kan zo gekozen worden dat Z'_{11} een diagonaalmatrix is.

10.2. Positief (semi-)definiete matrices

Definitie. (10.2.1)

A heet positief definiet ($A \in PD$) als A hermitisch is en

$$\forall_{x \neq 0} \; x^H A x > 0 .$$

A heet positief semi-definiet ($A \in PSD$) als A hermitisch is en

$$\forall_x \; x^H A x \geq 0 .$$

Stelling. (10.2.2)

Als $A \in PSD$ dan geldt voor alle x en y

$$|y^H A x|^2 \leq x^H A x \cdot y^H A y .$$

Bewijs.

De vorm $(y,x) := y^H Ax$ is een zg. semi-definiet inproduct in $\mathbb{C}^n$. Dat wil zeggen, de eigenschappen ia), ii), iii) uit (7.1.1) gelden maar iv) moet vervangen worden door $\forall_x : (x,x) \geq 0$. Deze eigenschappen zijn voldoende voor het bewijs van de ongelijkheid van Cauchy-Schwarz:

$$|(y,x)|^2 \leq (x,x)(y,y) \quad ,$$

waarmee de bewering van de stelling equivalent is. □

Gevolgen. (10.2.3)

1. Als $A \in PSD$ dan $x^H Ax = 0 \leftrightarrow Ax = 0$ (neem $y = Ax$).
2. $A \in PSD$ en A regulier $\Rightarrow A \in PD$.

Opgave. (10.2.4)

A heet indefiniet als A hermitisch is en A noch $-A$ is PSD.
Bewijs dat voor hermitische A geldt

$$A \in PSD \leftrightarrow \forall_x : Ax \neq 0 \Rightarrow x^H A > 0 \, ,$$
$$-A \in PDS \leftrightarrow \forall_x : Ax \neq 0 \Rightarrow x^H Ax < 0 \, ,$$
$$A \text{ indefiniet } \leftrightarrow \exists_x : Ax \neq 0, \ x^H Ax = 0 \, .$$

Eenvoudige eigenschappen van PD en PSD matrices. (10.2.5)

1) $A \in PD \Rightarrow \forall_k: (A)_{kk} > 0$.

2) $A \in PSD \Rightarrow \forall_k: (A)_{kk} \geq 0$; $(A)_{kk} = 0 \Rightarrow \forall_\ell (A)_{k\ell} = (A)_{\ell k} = 0$, want volgens (10.2.2) impliceert $e_k^T Ae_k = 0$ dat $Ae_k = 0$.

3) $A \in PSD \Rightarrow \forall_{k,\ell}: |(A)_{k\ell}|^2 \leq (A)_{kk}(A)_{\ell\ell}$.

4) Als $A = B_1 Z_{11} B_1^H$ met $B_1 \in RR$, Z_{11} regulier, dan $A \in PSD \Leftrightarrow Z_{11} \in PD$.

5) $A \in PSD \Rightarrow C^H AC \in PSD$.

6) Als $A \in PD$ dan $C \in RR \Rightarrow C^H AC \in PD$.

7) Als $C \in RI$ dan $C^H AC \in PSD$ (resp. PD) $\Rightarrow A \in PSD$ (resp. PD).

8) Als $A \in PSD$ dan $\langle X, AX \rangle \geq 0$ met gelijkteken dan en slechts dan als $AX = 0$.

9) Als $A \in PSD$ dan is $I + A$ regulier en voor alle x is $\|(I+A)^{-1} x\| \leq \|x\|$.

10) $A \in PD \leftrightarrow A^{-1} \in PD$; $A \in PSD \leftrightarrow A^+ \in PSD$.

<u>Stelling.</u> (10.2.6)

Zij

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

hermitisch en A_{11} vierkant.

Zij $A_{22.1} := A_{22} - A_{21}A_{11}^{+}A_{12}$ (dit is een generalisatie van (2.3.2)).

Er geldt dan

$$A \in PD \leftrightarrow \begin{cases} \text{i)} \ A_{11} \in PD \\ \text{ii)} \ A_{22.1} \in PD \end{cases}$$

$$A \in PSD \leftrightarrow \begin{cases} \text{i)} \ A_{11} \in PSD \\ \text{ii)} \ A_{22.1} \in PSD \\ \text{iii)} \ R(A_{12}) \subset R(A_{11}),\ N(A_{21}) \supset N(A_{11}) \ . \end{cases}$$

<u>Bewijs.</u>

Als $A \in PD$ dan is volgens (10.2.5.5) ook $A_{11} \in PD$, dus $A_{11} \in I$. Derhalve geldt in dit geval (als $A_{11} \in M_{k,k}$) dat $R(A_{11}) = \mathbb{C}^k \supset R(A_{21})$ en $N(A_{11}$ $N(A_{11})) = \{\underline{0}\} \subset N(A_{21})$. In het PD-geval is de eigenschap iii) dus een gevolg van i), hoeft dus niet apart vermeld te worden.

We bewijzen nu dat iii) geldt als $A \in PSD$. Stel $x_1 \in N(A_{11})$. Dan geldt (ga na) voor $x := (x_1^H \mid 0)^H$ dat $x^H Ax = 0$, dus $Ax = 0$, dus $A_{21}x_1 = 0$, dus $x_1 \in N(A_{21})$. Derhalve is $N(A_{11}) \subset N(A_{21})$; deze relatie is echter equivalent met $R(A_{11}) \supset R(A_{21})$, want $R(A_{11}) = (N(A_{11}))^{\perp}$ en $R(A_{21}) = (N(A_{21}^H))^{\perp} = (N(A_{12}))^{\perp}$.

Stel nu dat iii) geldt. Dan is $A_{11}A_{11}^{+}A_{12} = A_{12}$, $A_{21}A_{11}^{+}A_{11} = A_{21}$ en door narekenen zien we dat

$$A = \left(\begin{array}{c|c} I_{11} & \\ \hline A_{21}A_{11}^{+} & I_{22} \end{array}\right)\left(\begin{array}{c|c} A_{11} & \\ \hline & A_{22.1} \end{array}\right)\left(\begin{array}{c|c} I_{11} & A_{11}^{+}A_{12} \\ \hline & I_{22} \end{array}\right) . \qquad (*)$$

Daar de eerste en de derde factor van het rechterlid regulier zijn volgt uit (10.2.5) dat $A \in PSD$ (resp. PD) $\leftrightarrow A_{11}$ en $A_{22.1} \in PSD$ (resp. PD).

Hiermee is alles bewezen (ga na). □

<u>Gevolg.</u> (10.2.7)

Als $A \in PSD$ en

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

met vierkante A_{11} dan geldt

$$A = \left(\begin{array}{c|c} I_{11} & \\ \hline A_{21}A_{11}^{+} & I_{22} \end{array}\right)\left(\begin{array}{c|c} A_{11} & \\ \hline & A_{22.1} \end{array}\right)\left(\begin{array}{c|c} I_{11} & A_{11}^{+}A_{12} \\ \hline & I_{22} \end{array}\right)$$

met $A_{22.1} := A_{22} - A_{21}A_{11}^{+}A_{12}$.

<u>Stelling.</u> (10.2.8)

Als $A \in PSD$ dan is er een reguliere linksondermatrix L met diagonaalelementen 1 en een diagonaalmatrix D zo dat

$$A = LDL^H .$$

<u>Bewijs.</u>

Inductie naar n, waarbij de inductiestap uitgevoerd wordt met behulp van (10.2.7) waarin $A_{11} \in M_{1,1}$. □

<u>Gevolgen.</u> (10.2.9)

1. Als $A \in PD$ dan is er een reguliere linksondermatrix $\tilde{L}$ zo dat

   $$A = \tilde{L}\tilde{L}^H$$

   (Cholesky).
   Immers, in dit geval geldt voor de D uit (10.2.8) dat $D \in PD$; er is dus een $\tilde{D} \in PD$ zo dat $D = \tilde{D}^2$; neem nu $\tilde{L} := L\tilde{D}$.

2. Als $A \in PSD$ met rang r dan is er een linksondermatrix $\tilde{L}_1 \in M^r_{n,r}$ zo dat

   $$A = \tilde{L}_1\tilde{L}_1^H .$$

   Immers, van de matrix D uit (10.2.8) zijn in dit geval precies r diagonaalelementen positief en de andere zijn 0, er geldt dus ook $A = L_1D_1L_1^H$ met $L_1 \in M^r_{n,r}$ linksonder en $D_1 \in M^r_{r,r}$, diagonaal en $\in PD$.

3. De matrix $\tilde{L}$ uit (10.2.9.1) heeft positieve diagonaalelementen. Dit geldt niet noodzakelijk voor de diagonaalelementen van de $\tilde{L}_1$ uit (10.2.9.2), deze kunnen ook 0 zijn. Wel geldt (ga na): als $A \in PSD$ met rang r dan is er een permutatiematrix P en een linksondermatrix $\hat{L}_1 \in M^r_{n,r}$ met reguliere diagonaal zo dat

$$P^T AP = \hat{L}_1 \hat{L}_1^H .$$

4. Voor veel doeleinden is de volgende formulering voldoende:

$A \in PSD$ met rang r $\leftrightarrow A = B_1 B_1^H$ met $B_1 \in M^r_{n,r}$.

Opgave. (10.2.10)

Zij $A = LL^H$ met $L \in M_{n,r}$, linksonder en $(L)_{kk} \neq 0$, $1 \le k \le r$. Als ook $A = L_1 L_1^H$ met L_1 linksonder en $\in$ RR dan is $L_1 = LD$ met D diagonaal en unitair (dus $D = \mathrm{diag}(e^{i\theta_1},\dots,e^{i\theta_r})$ met reële θ's).

Opgave. (10.2.11)

Zij $A = B_1 B_1^H = CC^H$ met $B_1 \in M^r_{n,r}$ en $C \in M_{n,k}$. Dan is er een links-unitaire matrix $U_1 \in M_{k,r}$ zo dat $C = B_1 U_1^H$.

Opgave. (10.2.12)

Zij A_1 en $A_2 \in PSD$ met dezelfde afmetingen. Dan geldt

$$r(A_1 + A_2) \ge \max(r(A_1), r(A_2)).$$

Wanneer geldt het gelijkteken?

Opgave. (10.2.13)

Bewijs (8.2.1) (Gram-Schmidt) uit (10.2.9.1).
Bewijs (8.2.5) uit (10.2.9.3).

Opgave. (10.2.14)

Zij E orthogonale projector. Dan zijn E en $I - E \in PSD$ (ga na). Bewijs dat Cholesky, toegepast op E, een links-unitaire U_1 levert zo dat $E = U_1U_1^H$ (8.3.3). Leid hieruit voorts een constructie af om bij gegeven links-unitaire U_1 een U_2 te vinden zo dat $(U_1 \mid U_2)$ unitair is (8.2.2.5).

Opgave. (10.2.15)

Zij

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & 0 \end{array}\right)$$

met $A_{11} \in PSD$, $A_{21} = A_{12}^H$ (dus A hermitisch).
Dan geldt

1. $\left(\begin{array}{c} x_1 \\ \hline x_2 \end{array}\right) \in N(A) \Leftrightarrow x_1 \in N(A_{11}) \cap N(A_{21}),\ x_2 \in N(A_{12})$.

2. $r(A) = r\left(\left(\begin{array}{c} A_{11} \\ \hline A_{21} \end{array}\right)\right) + r(A_{12}) = r((A_{11} \mid A_{12})) + r(A_{21})$.

3. $\left(\begin{array}{c} y_1 \\ \hline y_2 \end{array}\right) \in R(A) \Leftrightarrow y_1 \in R(A_{11}) + R(A_{12}),\ y_2 \in R(A_{21})$.

Opgave. (10.2.16)

Zij

$$A = \left(\begin{array}{c|c} V & X \\ \hline X^H & 0 \end{array}\right), \quad C = \left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right)$$

met $V \in PSD$.
Dan geldt (Rao)

1. C is g-inverse van A dan en slechts dan als
   i) $VC_{11}X = 0,\ X^HC_{11}V = 0,\ X^HC_{11}X = 0,$
   ii) $XC_{21}X = X,\ X^HC_{12}X^H = X^H,$
   iii) $VC_{12}X^H = XC_{21}V = -XC_{22}X^H = V - VC_{11}V.$

2. Als C g-inverse van A is, dan geldt ook

iv) $VC_{11}^H VC_{11}V = VC_{11}V$,

$VC_{11}V$ is hermitisch, $\in$ PSD en onafhankelijk van de keuze van de g-inverse C,

v) C_{21} en C_{12}^H zijn g-inversen van X, $\begin{pmatrix} C_{11} \\ \hline C_{21} \end{pmatrix}$ is g-inverse van $(V \mid X)$.

vi) $tr(VC_{11}) = r((V \mid X)) - r(X)$.

N.b. Deze resultaten zijn van belang bij de behandeling van het zg. algemene Gauss-Markov model in de multivariate statistiek.

Hint: Pas (10.2.15) toe op de relaties $A(I - CA) = 0$, $A(I - C^HA) = 0$. vi) volgt uit v) door toepassing van (5.5.6.3).

Opgave. (10.2.17)

Zij $A = (a_1 \mid \ldots \mid a_n)$ vierkant en regulier. Zij, volgens Cholesky, $A^HA = LL^H$. Door gebruikmaking van een beetje determinantentheorie volgt hieruit

$$|\det(A)|^2 = \det(A^HA) = |\det(L)|^2 = \prod_{k=1}^{n} |(L)_{kk}|^2 .$$

Bewijs nu de ongelijkheid van Hadamard:

$$|\det(A)| \leq \prod_{k=1}^{n} \| a_k \| ,$$

met gelijkteken dan en slechts dan als $a_k^Ha_\ell = 0$, $k \neq \ell$.
(Hint: druk $\| a_k \|^2 = a_k^Ha_k$ uit in de elementen van L).

10.3. Algemene hermitische matrices. Congruentie.

Voor $A \in PSD$ is er volgens (10.2.9.4) een $B_1 \in RR$ zo dat

$$A = B_1 B_1^H .$$

Een dergelijke splitsing bestaat natuurlijk ook alleen als $A \in PSD$ omdat altijd $B_1 B_1^H \in PSD$. We zullen nu bewijzen dat er voor een algemene hermitische matrix een splitsing

$$A = B_1 B_1^H - B_2 B_2^H \tag{10.3.1}$$

met $(B_1 \mid B_2) \in RR$ bestaat. Hierin zijn B_1 en B_2 in het algemeen geenszins eenduidig bepaald, het getal

$$s := r(B_1) - r(B_2) = \textbf{breedte } (B_1) - \textbf{breedte } (B_2)$$

echter wel (stelling van Sylvester); s wordt de signatuur van A genoemd.
Een variant op (10.3.1) is

$$A = BJB^H \tag{10.3.2}$$

met $B \in I$ en

$$J = \left(\begin{array}{c|c|c} I_1 & & \\ \hline & -I_2 & \\ \hline & & 0 \end{array}\right) \tag{10.3.3}$$

met (als $r(B_1) = p$, $r(B_2) = q$) $I_1 \in M_{p,p}$, $I_2 \in M_{q,q}$. We zeggen dat A en J congruent zijn en dat J een normaalvorm is. De stelling van Sylvester luidt dan:
A is met precies één normaalvorm (gekarakteriseerd door z'n rang $r = p + q$ en z'n signatuur $s = p - q$) congruent.

Stelling. (10.3.4)

Zij $A \in M_{n,n}$ hermitisch. Dan is er een reguliere matrix B en een hermitische diagonaalmatrix D zo dat

$$A = BDB^H.$$

Bewijs.

We bewijzen eerst dat er bij A een permutatiematrix P en een reguliere linksondermatrix L is zo dat

$$P^T AP = LTL^H \, , \qquad (*)$$

waarin T een blok-diagonaalmatrix is met uitsluitend 1×1- en 2×2-blokken.
Inductie naar n. Voor n = 1 en ook als A = 0 is de bewering triviaal. Zij n > 1 en A ≠ 0. Als voor tenminste één diagonaalelement $(A)_{kk}$ geldt $(A)_{kk} = \alpha \neq 0$ dan is er een permutatiematrix P (horend bij verwisseling van de indices 1 en k) zo dat het 1×1-linksbovenblok van P^TAP regulier is. Als alle diagonaalelementen van A nul zijn dan is er (omdat A ≠ 0) een buitendiagonaalelement $(A)_{k\ell}$ ($k \neq \ell$!) met $(A)_{k\ell} = \beta \neq 0$. Er is dan een permutatiematrix P (horend bij verwisseling van de indices 1 en k, 2 en ℓ) zo dat het 2×2-linksbovenblok van P^TAP is

$$\begin{pmatrix} 0 & \beta \\ \bar{\beta} & 0 \end{pmatrix} \qquad (**)$$

en dit is regulier.
Zij nu

$$P^TAP = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

met A_{11} 1×1 of 2×2 en regulier. Dan geldt, als in (2.3.1),

$$P^TAP = \left(\begin{array}{c|c} I_{11} & \\ \hline L_{21} & I_{22} \end{array}\right) \left(\begin{array}{c|c} T_{11} & \\ \hline & A_{22.1} \end{array}\right) \left(\begin{array}{c|c} I_{11} & L_{21}^H \\ \hline & I_{22} \end{array}\right)$$

met $T_{11} = A_{11}$, $L_{21} = A_{21}A_{11}^{-1}$ en $A_{22.1} = A_{22} - A_{21}A_{11}^{-1}A_{12}$.
Daar $A_{22.1}$ weer hermitisch is kunnen we het inductiebewijs afmaken analoog aan het bewijs van (3.1.4).

Hiermee is de existentie van de splitsing (*) bewezen. Als T een diagonaalmatrix is dan zijn we klaar. Als T ook 2×2 blokken bevat dan merken we op dat er bij iedere hermitische 2×2 matrix T_1 van de vorm (**) een reguliere matrix U en een diagonaalmatrix D is zo dat

$$T_1 = U_1D_1U_1^H \, ,$$

bv.

$$U_1 = \begin{pmatrix} 1 & -e^{i\theta} \\ e^{-i\theta} & 1 \end{pmatrix}, \quad D_1 = \begin{pmatrix} \frac{1}{2}|\beta| & \\ & -\frac{1}{2}|\beta| \end{pmatrix} \qquad (***)$$

met $\theta = \arg(\beta)$.

Maak nu het bewijs af. □

Opmerkingen. (10.3.5)

1. Voor de matrix B geldt

$$B = PLU \;;$$

hierin is P een permutatiematrix, L een reguliere linksondermatrix en U een blok-diagonaalmatrix waarvan ieder diagonaalblok of een 1×1 eenheidsmatrix of een 2×2 matrix van de vorm (***) is. De diagonaalelementen van L zijn 1 en overal waar U een 2×2 blok heeft, heeft L een eenheidsmatrix als 2×2 diagonaalblok. LU is dus een blok - linksondermatrix waarvan de diagonaalblokken 1×1 - eenheidsmatrices of 2×2 matrices U_1 van de vorm (***) zijn.
2. Als A een symmetrische matrix over een willekeurig lichaam $\mathcal{L}$ is dan gelden stelling en bewijs ook (met B^T in plaats van B^H).

Gevolgen. (10.3.6)

1. Bij iedere hermitische matrix A is er een reguliere matrix $\tilde{B}$ en een normaalvorm J van de vorm (10.3.3) zo dat

$$A = \tilde{B}J\tilde{B}^H . \qquad (*)$$

Er is nl. een permutatiematrix P zo dat voor de matrix D uit (10.3.4) geldt

$$P^T DP = \mathrm{diag}(D_1^2,\ -D_2^2,\ 0)$$

met D_1 en $D_2 \in PD$ (en diagonaal). Met $\tilde{D} := \mathrm{diag}(D_1, D_2, I_3)$ geldt dus

$$D = P\tilde{D}J\tilde{D}P^T$$

zodat (*) uit (10.3.4) volgt met $\tilde{B} = BP\tilde{D}$.

2. Bij iedere hermitische matrix A is er een matrix $(B_1 \mid B_2) \in RR$ zo dat

$$A = B_1 B_1^H - B_2 B_2^H . \qquad (**)$$

Partitioneer nl. de matrix $\tilde{B}$ uit (10.3.6.1) als $\tilde{B} = (B_1 \mid B_2 \mid B_3)$ met passende afmetingen.

Omgekeerd volgt uit (**) natuurlijk weer de existentie van de splitsing (*) met $\tilde{B} = (B_1 \mid B_2 \mid B_3)$ met geschikte B_3. Merk op dat de voorwaarde $(B_1 \mid B_2) \in RR$ hierbij nodig is!

3. Bij iedere hermitische matrix A zijn er matrices A_1 en $A_2 \in PSD$ zodanig dat

$$A = A_1 - A_2 \ , \ r(A) = r(A_1) + r(A_2) \ .$$

Neem nl. $A_1 = B_1B_1^H$, $A_2 = B_2B_2^H$ met B_1 en B_2 als in (**).

Stelling. (10.3.7)

Zij A_1 en $A_2 \in PSD$, $A := A_1 - A_2$ en

$$r(A) = r(A_1) + r(A_2) \ .$$

Zij $A_1' \in PSD$ zo dat

$$\forall_x : x^H A_1' x \geq x^H A x \ .$$

Dan geldt $r(A_1') \geq r(A_1)$.

Bewijs.

Voor $x \in N(A_1') \cap N(A_2)$ geldt

$$0 = x^H A_1' x \geq x^H A x = x^H A_1 x \geq 0 \ ,$$

dus (volgens (10.2.3.1)) $x \in N(A_1)$. Hiermee is bewezen dat

$$N(A_1') \cap N(A_2) \subset N(A_1) \cap N(A_2) \subset N(A) \ ,$$

waaruit (bv. met (5.4.26) en (5.4.19.1) volgt

$$n - r(A_1) - r(A_2) = n - r(A) = \dim(N(A)) \geq$$

$$\geq \dim(N(A_1') \cap N(A_2)) \geq \dim(N(A_1')) + \dim(N(A_2)) - n = n - r(A_1') - r(A_2) \ . \ \square$$

Gevolgen. (10.3.8)

1. Als

$$A = A_1 - A_2 = A_1' - A_2'$$

met A_1, A_2, A_1' en $A_2' \in PSD$ en $r(A_1) + r(A_2) = r(A)$ dan is

$$r(A_1') \geq r(A_1), \ r(A_2') \geq r(A_2) \ .$$

Als ook $r(A_1') + r(A_2') = r(A)$ dan is $r(A_1') = r(A_1)$, $r(A_2') = r(A_2)$.

2. Bij iedere hermitische A met rang r is er een geheel getal s (de signatuur van A) zodanig dat voor iedere splitsing conform (10.3.6.3):

$$A = A_1 - A_2 \text{ , met } A_1 \text{ en } A_2 \in \text{PSD}, \ r(A_1) + r(A_2) = r(A) \ ,$$

geldt

$$r(A_1) - r(A_2) = s$$

(stelling van Sylvester).
De existentie van dergelijke splitsingen volgt uit (10.3.6.2).

3. Als A hermitisch is met rang r en signatuur s dan volgt uit $A = A_1 - A_2$ met A_1 en $A_2 \in$ PSD dat $r(A_1) \geq (r + s)/2$, $r(A_2) \geq (r - s)/2$.

4. Als $A = B_1B_1^H - B_2B_2^H$ met $(B_1 \mid B_2) \in$ RR dan geldt

$$r(A) = r(B_1) + r(B_2), \ s(A) = r(B_1) - r(B_2) \ .$$

Opgave. (10.3.9)

Zij $A = A_1 - A_2$ met A_1 en $A_2 \in$ PSD en $R(A_1) \perp R(A_2)$. Dan is $r(A_1) + r(A_2) = r(A)$.
Als ook $A = A_1' - A_2'$ met A_1' en $A_2' \in$ PSD en $R(A_1') \perp R(A_2')$ dan is $A_1' = A_1$, $A_2' = A_2$.
Nb. De splitsing $A = A_1 - A_2$ is onder de gegeven condities voor A_1 en A_2 dus uniek. De existentie van een dergelijke splitsing wordt pas in §15 bewezen.
Hint: uit het gegeven volgt dat $A_1^2A_2' + A_1A_2'^2 = 0$, leid daaruit met (10.2.5.8) af dat $A_1A_2' = A_2'A_1 = 0$, analoog geldt $A_2A_1' = A_1'A_2 = 0$. Pas nu (8.5.4.3) toe.

Definitie. (10.3.10)

Hermitische matrices A en A' heten congruent als er een reguliere matrix S bestaat zo dat

$$A' = S^HAS$$

(dit is een equivalentierelatie!).

Stelling. (10.3.11)

Bij een hermitische matrix A is er precies één normaalvorm

$$J = \left(\begin{array}{c|c|c} I_1 & & \\ \hline & -I_2 & \\ \hline & & 0 \end{array}\right)$$

waarmee A congruent is.
Als $I_1 \in M_{p,p}$, $I_2 \in M_{q,q}$ dan geldt

$$p + q = r(A),\ p - q = s(A)\ .$$

Stelling. (10.3.12)

Hermitische matrices A en A' zijn congruent dan en slechts dan als ze dezelfde afmetingen, rang en signatuur hebben.

Opgave. (10.3.13)

Zij A hermitisch met rang r en signatuur s.
Bewijs dat

$$\max\{r(X_1)\ |\ X_1^H A X_1 \in PD\} = (r + s)/2,\ \max\{r(X_2)\ |-X_2^H A X_2 \in PD\} = (r-s)/2\ .$$

Hint: Als $X_1^H A X_1 \in PD$ dan is er een X_2 zo dat $(X_1 \mid X_2) \in I$ en $X_1^H A X_2 = 0$. Pas (10.3.11) toe op $X_2^H A X_2$.

10.4. Hermitische bilineaire en kwadratische vormen

Definitie. (10.4.1)

Een hermitische bilineaire vorm over $\mathbb{C}^n$ is een afbeelding $\varphi : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ waarvoor geldt:

i) $\varphi(y,x_1 + x_2) = \varphi(y,x_1) + \varphi(y,x_2)$,

ii) $\varphi(y,\alpha x) = \alpha\varphi(y,x)$,

iii) $\varphi(x,y) = \overline{\varphi(y,x)}$.

Gevolgen:

iv) $\varphi(\alpha y,x) = \bar{\alpha}\varphi(y,x)$,

v) $\varphi(x,x)$ is reëel .

Stelling. (10.4.2)

Voor iedere hermitische bilineaire vorm φ is er precies één hermitische matrix $A \in M_{n,n}$ zo dat

$$\varphi(y,x) = y^H Ax .$$

Bewijs.

Neem $(A)_{ij} = \varphi(e_i,e_j)$. □

Definitie. (10.4.3)

Een hermitische kwadratische vorm over $\mathbb{C}^n$ is een afbeelding $\psi : \mathbb{C}^n \to \mathbb{R}$ die is afgeleid van een hermitische bilineaire vorm φ:

$$\psi(x) = \varphi(x,x) .$$

Natuurlijk kan men op geheel analoge wijze symmetrische bilineaire en kwadratische vormen over $\mathbb{R}^n$ definiëren, eis iii) wordt dan vervangen door $\varphi(x,y) = \varphi(y,x)$. Er is echter een opmerkelijk verschil tussen vormen over $\mathbb{C}^n$ en vormen over $\mathbb{R}^n$, dat blijkt uit de volgende drie stellingen.

Stelling. (10.4.4)

Zij $A \in M_{n,n}(\mathbb{C})$ en $\forall_{x\in\mathbb{C}^n} : x^H Ax = 0$. Dan is $A = 0$.
Zij $A \in M_{n,n}(\mathbb{R})$ en $\forall_{x\in\mathbb{R}^n} : x^T Ax = 0$. Dan is $A + A^T = 0$
(en als bovendien gegeven is dat A symmetrisch is, dan is $A = 0$).

Bewijs.

Zij $A \in M_{n,n}(\mathbb{C})$. Uit het gegeven volgt dat voor alle x en $y \in \mathbb{C}^n$ en $\alpha \in \mathbb{C}$ geldt $(x + \alpha y)^H A(x + \alpha y) = 0$, waaruit volgt (ga na) dat

$$\mathrm{Re}(\alpha x^H Ay) = 0 .$$

Daar α willekeurig is volgt hieruit dat $y^H Ax = 0$, dus $A = 0$. Uit dezelfde redenering blijkt dat we in het reële geval niet verder komen dan $A + A^T = 0$. □

Stelling. (10.4.5)

Zij $A \in M_{n,n}(\mathbb{C})$ en $\forall_{x \in \mathbb{C}^n} : x^H Ax \in \mathbb{R}$. Dan is A hermitisch.

Bewijs.

Als $x^H Ax$ reëel is dan is $x^H Ax = \overline{x^H Ax} = x^H A^H x$. Dus geldt $\forall_x : x^H(A - A^H)x = 0$. Dus met (10.4.4): $A - A^H = 0$. □

Opmerking.

In het reële geval is het gegeven dat $x^H Ax \in \mathbb{R}$ natuurlijk triviaal. Er kan dus geen analogon zijn van (10.4.5) in het reële geval. Wel geldt hier (en niet in het complexe geval)

Stelling. (10.4.6)

Zij $A \in M_{n,n}(\mathbb{R})$. Dan is er een symmetrische matrix $A' \in M_{n,n}(\mathbb{R})$ zodanig dat

$$\forall_{x \in \mathbb{R}^n} : x^T Ax = x^T A' x \ .$$

Bewijs.

Daar voor alle x $\quad x^T Ax = x^T A^T x$ kunnen we nemen $A' = \frac{1}{2}(A + A^T)$, A' is dan symmetrisch. □

De resultaten over hermitische matrices kunnen natuurlijk vertaald worden in resultaten over hermitische kwadratische vormen. We noemen alleen het equivalent van (10.3.6.2) en de stelling van Sylvester (10.3.8.2).

Stelling. (10.4.7)

Zij ψ een kwadratische vorm over $\mathbb{C}^n$. Dan zijn er

een getal r, $0 \le r \le n$: de rang van ψ,
een getal s, $-r \le s \le r$: de signatuur van ψ,
onafhankelijke $b_1,\ldots,b_r \in \mathbb{C}^n$,

zodat

$$\psi(x) = \sum_{1}^{p} |b_i^H x|^2 - \sum_{p+1}^{r} |b_i^H x|^2 ,$$

waarin $p = (r + s)/2$.
De getallen r en s zijn eenduidig door ψ bepaald, de vectoren $b_1,\ldots,b_r$ niet.
De vormen $\chi_i(x) := b_i^H x$ heten lineaire vormen.

De definitie van congruentie van hermitische matrices krijgt een meetkundige zin als we deze matrices beschouwen als matrices van hermitische bilineaire vormen. Zij $\varphi(y,x) = y^H Ax$ met A hermitisch. A is de matrix van φ t.o.v. de basis $(e_1,\ldots,e_n)$ in $\mathbb{C}^n$. Kies nu een andere (niet noodzakelijk orthonormale) basis $(s_1,\ldots,s_n)$ in $\mathbb{C}^n$. Dan is $S := (s_1| \ldots |s_n)$ de overgangsmatrix (zie(6.1.7)) van de basis $(e_1,\ldots,e_n)$ naar $(s_1,\ldots,s_n)$. En als x' de coördinatenvector van x t.o.v. de nieuwe basis is dan is (zie 6.1.9) $x = Sx'$. Derhalve is ook

$$\varphi(y,x) = y'^H S^H ASx' = y'^H A'x'$$

met $A' = S^H AS$. A' is dus de matrix van φ t.o.v. de basis $(s_1,\ldots,s_n)$ (ga na dat $(A')_{ij} = \varphi(s_i,s_j)$). We zien dat A' hermitisch is (ook als de nieuwe basis niet orthonormaal is!) *en congruent met* A. *En uit* (10.3.12) volgt dat rang en signatuur meetkundige eigenschappen van φ zijn, onafhankelijk van de keuze van een basis in $\mathbb{C}^n$ (en ook dat er een basis - als regel niet orthogonaal - in $\mathbb{C}^n$ is ten opzichte waarvan de matrix van φ de normaalvorm (10.3.3) heeft).

Opgave. (10.4.8)

A en $A' \in M_{n,n}(\mathbb{C})$ zijn dan en slechts dan congruent als er een hermitische bilineaire vorm φ is waarvan A en A' de matrices zijn t.o.v. zekere bases in $\mathbb{C}^n$.

Hoofdstuk IV. Gelijkvormigheid

11. Inleiding.

Matrices A en B $\in M_{n,n}(\mathcal{L})$ heten gelijkvormig als er een reguliere matrix $S \in M_{n,n}(\mathcal{L})$ is zo dat

$$B = S^{-1}AS \ . \tag{11.1}$$

Dit is **duidelijk een** equivalentie relatie. De meetkundige betekenis zien we door A te beschouwen als matrix van een afbeelding $\mathcal{A}: \mathcal{L}^n \to \mathcal{L}^n$ t.o.v. de natuurlijke basis $(e_1,\ldots,e_n)$ van $\mathcal{L}^n$. Ga nu over op een andere basis $(s_1,\ldots,s_n)$ van $\mathcal{L}^n$. Uit (6.2.4) volgt dan dat de matrix van $\mathcal{A}$ t.o.v. deze basis gegeven wordt door B volgens (11.1), waarin S de overgangsmatrix van de basis $(e_1,\ldots,e_n)$ naar de basis $(s_1,\ldots,s_n)$ is. Uit (6.1.7) volgt dat $S = (s_1 \mid \ldots \mid s_n)$.

Het is natuurlijk van belang om bij een matrix A te zoeken naar een gelijkvormigheidstransformatie $A \to B := S^{-1}AS$ zo dat B een eenvoudige vorm heeft: we hebben dan een basis voor $\mathcal{L}^n$ gevonden t.o.v. waarvan de met A corresponderende afbeelding $\mathcal{A}$ een eenvoudige structuur heeft. En ideaal daarbij kan zijn: zorg dat B een diagonaalmatrix is, dus

$$S^{-1}AS = \Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_n) \ . \tag{11.2}$$

Als dat lukt, dan volgt uit $AS = S\Lambda$ dat

$$As_j = \lambda_j s_j \ , \qquad j = 1,\ldots,n \ .$$

De basisvector s_j is dan eigenvector van A bij eigenwaarde λ_j.

Helaas blijkt dat slechts een beperkte (zij het grote) klasse van matrices diagonaliseerbaar is in de zin van (11.2). Voor de overige matrices moeten we de wens dat B diagonaal is verzwakken tot de eis dat B blokdiagonaal is:

$$S^{-1}AS = \mathrm{diag}(B_{11},\ldots,B_{pp}) \ , \tag{11.3}$$

met "zo klein mogelijke" (d.w.z. niet verder te "blokdiagonaliseren") vierkante diagonaalblokken.
Als we de kolommen van S partitioneren passend bij de partitionering van B dan volgt uit (11.3)

$$AS_j = S_jB_{jj} ,$$

d.w.z., A beeldt de door de kolommen van S_j opgespannen deelruimte V_j van $\mathcal{L}^n$ af in zichzelf. V_j heet dan een invariante deelruimte van $\mathcal{L}^n$ t.o.v. A. En B_{jj} beschrijft hoe A V_j in zichzelf afbeeldt. We hebben dus de bij A horende afbeelding $\mathcal{A}$ opgesplitst in p afbeeldingen met matrices $B_{11},\ldots,B_{pp}$.

Een bijzonder belang van reductie naar eenvoudiger vorm via gelijkvormigheidstransformaties ligt in het feit dat uit (11.3) volgt dat voor $k \in \mathbb{N}$ geldt

$$A^k = SB^kS^{-1} = S\ \mathrm{diag}(B_{11}^k,\ldots,B_{pp}^k)S^{-1} .$$

Dit maakt het mogelijk om functies van een matrix A (bv. polynomen, maar ook machtreeksen) te berekenen door ze te berekenen voor de eenvoudiger matrices B_{jj}.

Het zal blijken dat de existentie van eigenwaarden, eigenvectoren en invariante deelruimten van A afhangt van de ontbindbaarheid van het bij A behorende zgn. minimaalpolynoom, een polynoom met graad $\leq n$ met coëfficiënten uit het lichaam $\mathcal{L}$. Nu is in een willekeurig lichaam $\mathcal{L}$ een polynoom φ met graad > 1 en coëfficiënten uit $\mathcal{L}$ niet altijd te ontbinden in factoren φ_1 en φ_2 met graad ≥ 1 en coëfficiënten uit $\mathcal{L}$. (Voorbeeld: het polynoom $x^2 + 1$, beschouwd als polynoom over $\mathbb{R}$, is niet te schrijven als product van twee eerstegraadspolynomen met coëfficiënten uit $\mathbb{R}$.) Voor het lichaam $\mathbb{C}$ geldt dit wel (hoofdstelling van de algebra). De algemene theorie van de reductie van matrices naar eenvoudiger vorm door middel van gelijkvormigheidstransformaties is daarom voor matrices over een lichaam waarin de hoofdstelling van de algebra niet geldt gecompliceerder dan voor matrices over $\mathbb{C}$.
We zullen deze algemene theorie alleen volledig ontwikkelen voor matrices over $\mathbb{C}$ (of een ander lichaam waarin de hoofdstelling van de algebra geldt), maar alle resultaten die gelden in alle lichamen $\mathcal{L}$ formuleren voor matrices over $\mathcal{L}$. Voor de ontbrekende delen van de theorie voor matrices over een algemener lichaam zij verwezen naar boeken*).

*) bv. Gantmacher, F.R., The theory of matrices, Vol. I, Ch. VII, New York, 1959 (Chelsea) of Hoffmann, K. and R. Kunze, Linear algebra, Englewood Cliffs (NJ), 1971 (Prentice-Hall).

12. Eigenwaarden, eigenvectoren, minimaalpolynomen, diagonaliseerbare matrices.

In deze paragraaf is $A \in M_{n,n}(\mathcal{L})$ met willekeurig lichaam $\mathcal{L}$.

Opmerking. (12.0.1)

We zullen in het vervolg werken met polynomen $\varphi(z)$ over $\mathcal{L}$ waarin we voor de variabele z een matrix A substitueren. Daarom hier enige opmerkingen over polynomen.

Een polynoom φ over $\mathcal{L}$ is een vorm

$$\varphi(z) = \sum_{j=0}^{m} \gamma_j z^j$$

met coëfficiënten γ_j uit $\mathcal{L}$ en een ongespecificeerde variabele z. φ heet triviaal als alle γ_j nul zijn. Als $\gamma_m \neq 0$ dan heeft $\varphi(z)$ graad m.

Polynomen φ_1 en φ_2 heten gelijk als hun coëfficiënten overeenstemmen. En som, product, vermenigvuldiging met scalairen vatten we op als relaties tussen coëfficiëntenrijen (die zo gedefinieerd zijn dat bv. de vermenigvuldiging commutatief is). De polynomen vormen zo een commutatieve ring over $\mathcal{L}$. Pas als we in een punt $z = \lambda$ een waarde aan een polynoom willen toekennen hoeven we te zeggen wat λ is. De enige eisen die daarbij gesteld worden zijn dat λ behoort tot een verzameling waarbinnen machtsverheffen en lineair combineren met coëfficiënten uit $\mathcal{L}$ mogelijk is. Als ook vermenigvuldiging mogelijk is en voor alle p en q (geheel, ≥ 0) geldt $\lambda^p.\lambda^q = \lambda^{p+q}$ dan impliceert $\varphi(z) = \varphi_1(z)\varphi_2(z)$ dat ook $\varphi(\lambda) = \varphi_1(\lambda)\varphi_2(\lambda) = \varphi_2(\lambda)\varphi_1(\lambda)$.

Merk nu op dat aan de genoemde eisen voldaan is als we als "actuele parameter" λ hetzij een element uit $\mathcal{L}$, hetzij een matrix $A \in M_{n,n}(\mathcal{L})$ nemen (met $A^0 := I$). N.B. Er geldt voor $\alpha \in \mathcal{L}$: $\varphi(\alpha I) = \varphi(\alpha)I$.

12.1. Eigenwaarden en eigenruimten

Definities. (12.1.1)

1. $\lambda \in \mathcal{L}$ heet eigenwaarde van A als $A - \lambda I$ singulier is.
2. Het spectrum van A, $\sigma(A)$, is de verzameling van alle eigenwaarden van A.
3. Als $\lambda \in \sigma(A)$ dan heet $N(A - \lambda I)$ de eigenruimte van A bij λ. De vectoren

$x \neq 0$ uit $N(A - \lambda I)$ heten eigenvectoren van A bij λ, $\dim(N(A - \lambda I))$ heet geometrische multipliciteit van λ.

Stelling. (12.1.2)

Zij A en B gelijkvormig. Dan hebben A en B dezelfde eigenwaarden met dezelfde geometrische multipliciteiten. Als $B = S^{-1}AS$ en x is eigenvector van A dan is $S^{-1}x$ eigenvector van B.

De volgende stelling geeft een uitspraak over de onafhankelijkheid van de eigenruimten en daarmee een bovengrens voor het aantal eigenwaarden.

Stelling. (12.1.3)

Stel dat $\lambda_1,\ldots,\lambda_p$ onderling verschillende eigenwaarden van A zijn met geometrische multipliciteiten $n_1,\ldots,n_p$. Zij

$$S = (S_1 \mid \ldots \mid S_p)$$

zo dat de kolommen van S_j een basis vormen voor $N(A - \lambda_j I)$. Dan is $S \in RR$.

Bewijs.

Zij $L_1(z),\ldots,L_p(z)$ de interpolatiepolynomen van Lagrange met basispunten $\lambda_1,\ldots,\lambda_p$:

$$L_k(z) = \prod_{\substack{i=1 \\ i\neq k}}^{p} \frac{z - \lambda_i}{\lambda_k - \lambda_i} .$$

Uit $AS_j = \lambda_j S_j$ volgt dan (ga na)

$$L_k(A)S_j = L_k(\lambda_j)S_j = \begin{cases} 0 & \text{als } j \neq k , \\ S_k & \text{als } j = k . \end{cases}$$

Stel nu dat

$$Sx = \sum_{j=1}^{p} S_j x_j = 0 .$$

Dan geldt voor $1 \leq k \leq p$

$$0 = L_k(A)Sx = \Sigma L_k(A)S_j x_j = S_k x_k ,$$

dus $x_k = 0$ (want $S_k \in RR$). Derhalve is $x = 0$ en dus $S \in RR$. $\square$

Gevolgen. (12.1.4)

1. $N(A - \lambda_1 I),\dots,N(A - \lambda_p I)$ zijn onafhankelijk, $\sum_1^p n_j \le n$ (zie (5.4.19)).
2. A heeft hoogstens n verschillende eigenwaarden.
3. De door de kolommen van S opgespannen deelruimte van $\mathcal{L}^n$ is een t.o.v. A invariante deelruimte en er geldt

$$AS = S\Lambda \text{ met } \Lambda = \mathrm{diag}(\lambda_1 I_1,\dots,\lambda_p I_p) ,$$

waarin I_j de $n_j \times n_j$ eenheidsmatrix is.
4. Als $\Sigma n_j = n$ dan is S regulier, A gelijkvormig met Λ, $\sigma(A) = \{\lambda_1,\dots,\lambda_p\}$ en

$$N(A - \lambda_1 I) \oplus \dots \oplus N(A - \lambda_p I) = \mathcal{L}^n.$$

Opgave. (12.1.5)

Zij $A = \mathrm{diag}(A_1,\dots,A_k)$ met vierkante A_i. Dan is

$$\sigma(A) = \bigcup_{i=1}^{k} \sigma(A_i),\ \dim(N(A - \lambda I)) = \sum_{i=1}^{k} \dim(N(A_i - \lambda I_i)) .$$

12.2. Het minimaalpolynoom

Stelling. (12.2.1)

Bij iedere $A \in M_{n,n}(\mathcal{L})$ zijn er niet-triviale polynomen φ over $\mathcal{L}$ zo dat $\varphi(A) = 0$ (de nulmatrix).

Bewijs.

Beschouw bij zekere $N \ge 0$ de n^2 homogene lineaire vergelijkingen

$$\gamma_N A^N + \gamma_{N-1} A^{N-1} + \dots + \gamma_0 A^0 = 0 \tag{12.2.2}$$

met de $N + 1$ onbekenden $\gamma_N,\dots,\gamma_0$. Voor voldoende grote N (zeker als $N \ge n^2$ want dan zijn er meer onbekenden dan vergelijkingen) hebben deze vergelijkingen een niet-triviale oplossing. □

Definities. (12.2.3)

1. Een niet-triviaal polynoom φ waarvoor geldt $\varphi(A) = 0$ heet annihilerend polynoom van A.

2. Een polynoom φ_0 heet minimaalpolynoom van A als
   a) φ_0 is annihilerend polynoom,
   b) als ook φ annhilerend polynoom is dan is graad(φ) $\geq$ graad(φ_0),
   c) hoogste coëfficiënt van φ_0 is 1.

Stelling. (12.2.4)

Bij iedere matrix A hoort een eenduidig bepaald minimaalpolynoom φ_0 met graad ≥ 1.
Een polynoom φ is annihilerend polynoom dan en slechts dan als φ_0 deler is van φ.

Bewijs.

De existentie van φ_0 volgt uit (12.2.1): neem voor N de laagste waarde waarvoor het stelsel (12.2.2) een niet-triviale oplossing heeft, voor zo'n oplossing is stellig $\gamma_N \neq 0$ en $N > 0$; deel zo nodig door γ_N.
Als φ een annihilerend polynoom is en φ_0 geen deler van φ zou zijn, dan zouden we door deling met rest ($\varphi = \varphi_0\psi + \varphi_1$) een annihilerend polynoom φ_1 kunnen vinden met graad lager dan φ_0. Tegenspraak, dus φ_0 deler van φ. De omkering van deze uitspraak is triviaal.
Uit de tweede bewering van de stelling volgt ook de eenduidigheid van het minimaalpolynoom (ga na!). $\square$

Eenvoudige eigenschappen. (12.2.5)

1. Als A en B gelijkvormig zijn dan hebben zij hetzelfde minimaalpolynoom.
2. Als φ_0 minimaalpolynoom van A is dan is φ_0^*, gedefinieerd door $\varphi_0^*(z) = \varphi_0(z + \alpha)$ minimaalpolynoom van $A - \alpha I$.
3. Als φ_1 en φ_2 annihilerend polynoom van A zijn dan is ggd(φ_1,φ_2) ook annihilerend polynoom van A (volgt uit (12.2.4)).
4. Zij $A = \mathrm{diag}(A_1,\dots,A_p)$ met vierkante diagonaalblokken. Zij $\varphi_1,\dots,\varphi_p$ de minimaalpolynomia van $A_1,\dots,A_p$. Dan is

$$\varphi_0 := \mathrm{kgv}(\varphi_1,\dots,\varphi_p)$$

   het minimaalpolynoom van A (merk op dat voor ieder polynoom φ

$$\varphi(A) = \mathrm{diag}(\varphi(A_1),\dots,\varphi(A_p)) .$$

5. A en A^T hebben hetzelfde minimaalpolynoom.

Stelling. (12.2.6)

Als φ annihilerend polynoom is van A dan geldt

$$\lambda \in \sigma(A) \Rightarrow \varphi(\lambda) = 0 .$$

Als φ_0 minimaalpolynoom is van A dan geldt

$$\lambda \in \sigma(A) \Leftrightarrow \varphi_0(\lambda) = 0 .$$

Bewijs.

Zij φ annihilerend polynoom van A. Dan is graad$(\varphi) \geq 1$. Volgens de reststelling is er bij iedere $\lambda \in \mathcal{L}$ een niet-triviaal polynoom ψ zo dat

$$\varphi(z) = (z - \lambda)\psi(z) + \varphi(\lambda)$$

en dus (zie (12.0.1))

$$(A - \lambda I)\psi(A) = -\varphi(\lambda)I .$$

Hieruit volgt

- $\varphi(\lambda) \neq 0 \Rightarrow A - \lambda I$ regulier $\Rightarrow \lambda \notin \sigma(A)$,
- $\varphi(\lambda) = 0$ en $\lambda \notin \sigma(A) \Rightarrow \psi(A) = 0$, dus φ is niet het minimaalpolynoom van A.

Uit deze uitspraken volgen de beweringen van de stelling. □

Gevolgen. (12.2.7)

1. $\sigma(A) = \{\lambda \in \mathcal{L} \mid \varphi_0(\lambda) = 0\}$.
2. Iedere $A \in M_{n,n}(\mathbb{C})$ heeft minstens één eigenwaarde (volgt uit de hoofdstelling van de algebra).
3. Een $A \in M_{n,n}(\mathbb{R})$ hoeft in $\mathbb{R}$ geen enkele eigenwaarde te hebben. Voorbeeld:

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

Als niet $\alpha = \delta$ en $\beta = \gamma = 0$ dan is het minimaalpolynoom (ga na) $\varphi_0(z) = z^2 - (\alpha + \delta)z + \alpha\delta - \beta\gamma$. Dit heeft geen nulpunten in $\mathbb{R}$ als $(\alpha - \delta)^2 + 4\beta\gamma < 0$.

Opmerkingen. (12.2.8)

Het zogenaamde karakteristieke polynoom Δ van A is gedefinieerd door

$$\Delta(z) := \det(zI - A) .$$

Uit de rekenregels voor determinanten volgt dat Δ een polynoom is over $\mathcal{L}$ met graad n.
Daar uit de determinantentheorie volgt dat $\lambda I - A$ singulier is dan en slechts dan als $\Delta(\lambda) = 0$, heeft Δ dezelfde nulpunten in $\mathcal{L}$ als φ_0. Er geldt zelfs meer, nl.

> het karakteristieke polynoom Δ van A is een annihilerend polynoom van A d.w.z.: $\Delta(A) = 0$ (stelling van Hamilton-Cayley).

Hieruit volgt

> het minimaalpolynoom φ_0 van A is een deler van het karakteristieke polynoom Δ van A.

Een bewijs van de stelling van Hamilton-Cayley met determinantentheorie is eenvoudig. Uit de regel van Cramer volgt nl. dat voor $\lambda \notin \sigma(A)$

$$(\lambda I - A)^{-1} = Q(\lambda)/\Delta(\lambda)$$

waarbij $Q(\lambda)$ een polynoom (van graad $< n$) is met matrices als coëfficiënten. Dit is equivalent met

$$\Delta(\lambda)I = (\lambda I - A)Q(\lambda) .$$

Door te bedenken dat deze relatie een relatie is tussen coëfficiëntenrijen zien we in dat we λ door A mogen vervangen, waaruit volgt dat $\Delta(A) = 0$.

Een gevolg van het bovenstaande is dat de graad van het minimaalpolynoom van een $n \times n$ matrix altijd $\leq n$ is.
Daar wij geen gebruik van determinantentheorie willen maken gebruiken we de stelling van Hamilton-Cayley in het vervolg niet en we zullen pas later bewijzen dat graad$(\varphi_0) \leq n$ (nl. in (13.2.18)).

Opgave. (12.2.9)

Zij $E \in M_{n,n}$ een projector. Bepaal het minimaalpolynoom, de eigenwaarden en de eigenruimten van E. Pas (12.1.4.4) toe op E en breng het resultaat in verband met (5.5.6.1).

12.3. Diagonaliseerbare matrices.

Definitie. (12.3.1)

$A \in M_{n,n}(\mathcal{L})$ heet diagonaliseerbaar (in $\mathcal{L}$) als A gelijkvormig is met een diagonaalmatrix:

$$A = S\Lambda S^{-1} ,$$

met $S \in M_{n,n}(\mathcal{L})$ regulier en $\Lambda \in M_{n,n}(\mathcal{L})$ diagonaal.

Eigenschappen. (12.3.2)

In het volgende bespreken we een aantal eigenschappen van een diagonaliseerbare matrix A, de bijbehorende projectoren E_j en het minimaalpolynoom φ_0 van A. Tevens laten we zien dat verschillende van deze eigenschappen de diagonaliseerbare matrices ook karakteriseren.

A) Zij $A \in M_{n,n}$ diagonaliseerbaar. Zonder beperking der algemeenheid mogen we dan veronderstellen dat

$$A = S\Lambda S^{-1} = S\ \mathrm{diag}(\lambda_1 I_1,\dots,\lambda_p I_p)S^{-1} \qquad (12.3.3)$$

met $\lambda_1,\dots,\lambda_p \in \mathcal{L}$, $p \geq 1$ en $\lambda_i \neq \lambda_j$ voor $i \neq j$.
Stel dat I_j afmetingen $n_j \times n_j$ heeft en dat

$$S = (S_1 \mid \dots \mid S_p) ,$$

waarin S_j breedte n_j heeft.
Dan geldt (ga na):

a) $\sigma(A) = \{\lambda_1,\dots,\lambda_p\}$, dat wil zeggen, $A - \lambda I$ is singulier dan en slechts dan als $\lambda \in \{\lambda_1,\dots,\lambda_p\}$,

b) de kolommen van S_j vormen een basis voor de eigenruimte $N(A - \lambda_j I)$, de geometrische multipliciteit van λ_j is dus n_j,

c) $\quad \Sigma \dim(N(A - \lambda_i I)) = \Sigma n_i = n$,

d) $\quad N(A - \lambda_1 I) \oplus \dots \oplus N(A - \lambda_p I) = \mathcal{L}^n$,

(dit volgt uit het feit dat S regulier is, zie (5.4.12)),

e) voor ieder polynoom φ geldt

$$\varphi(A) = S\ \mathrm{diag}(\varphi(\lambda_1)I_1,\dots,\varphi(\lambda_p)I_p)S^{-1} ;$$

φ is dus annihilerend polynoom van A dan en slechts dan als $\forall_i \varphi(\lambda_i) = 0$,

f) het minimaalpolynoom van A is (waarom?)

$$\varphi_0(z) = \prod_i (z - \lambda_i) .$$

Definieer nu voor $j = 1,\dots,p$

$$E_j := S\ \mathrm{diag}(0,\dots,I_j,\dots,0)S^{-1} . \qquad (12.3.4)$$

Dan geldt (ga na):

g) $r(E_j) = n_j > 0$,

h) $E_iE_j = 0$ als $i \neq j$, $E_j^2 = E_j$, $\Sigma E_i = I$,

i) $E_jA = AE_j = \lambda_jE_j$, $\Sigma\lambda_iE_i = A$.

Zij

$$L_j(z) := \prod_{i\neq j} \frac{z-\lambda_i}{\lambda_j-\lambda_i} = \frac{\varphi_0(z)}{(z-\lambda_j)\varphi_0'(\lambda_j)}$$

het j-de interpolatiepolynoom van Lagrange op de punten $\lambda_1,\dots,\lambda_p$. Dan geldt $L_j(\lambda_i) = \delta_{ij}$ en met e) volgt hieruit

j) $L_j(A) = E_j$.

De E_j zijn dus polynomen in A met coefficienten die uitgedrukt kunnen worden in de eigenwaarden van A.

B) We beginnen nu opnieuw en veronderstellen dat $E_1,\dots,E_p \in M_{n,n}$ voldoen aan

$$E_j \neq 0,\ E_iE_j = 0 \text{ voor } i \neq j\ ,\ \Sigma E_i = I . \qquad (12.3.5)$$

Dan geldt (ga na, vgl. (5.5.7) en (5.5.8))

a) $E_1,\dots,E_p$ zijn projectoren op onderling onafhankelijke deelruimten van $\mathcal{L}^n$ en

$$R(E_1) \oplus \dots \oplus R(E_p) = \mathcal{L}^n ,$$

b) als $n_j := r(E_j)$ dan is $\Sigma n_j = n$.

Stel voorts dat $\lambda_1,\dots,\lambda_p$ onderling verschillende elementen van $\mathcal{L}$ zijn en definieer

$$A := \sum_i \lambda_iE_i . \qquad (12.3.6)$$

Dan geldt (ga na):

c) $E_jA = AE_j = \lambda_jE_j$,

d) daar $x \in N(A - \lambda_j I) \leftrightarrow \sum_i (\lambda_i - \lambda_j)E_i x = 0 \leftrightarrow E_i x = 0$ voor $i \neq j \leftrightarrow x \in R(E_j)$, geldt

$$N(A - \lambda_j I) = R(E_j) \ ;$$

λ_j is dus eigenwaarde van A met geometrische multipliciteit n_j,

e) uit $E_j(A - \lambda_j I) = 0$ volgt $R(A - \lambda_j I) \subset N(E_j)$ en daar $x \in N(E_j) \Rightarrow$ $x = \sum_{i \neq j} E_i x = (A - \lambda_j I) \sum_{i \neq j} (\lambda_i - \lambda_j)^{-1} E_i x \in R(A - \lambda_j I)$, geldt

$$R(A - \lambda_j I) = N(E_j)$$

en dus ook

$$R(A - \lambda_j I) \oplus N(A - \lambda_j I) = \mathcal{L}^n \ ,$$

f) voor $\lambda \notin \{\lambda_1, \ldots, \lambda_p\}$ is $A - \lambda I$ regulier en

$$(A - \lambda I)^{-1} = \sum_i (\lambda_i - \lambda)^{-1} E_i \ ;$$

met d) volgt hieruit dat $\sigma(A) = \{\lambda_1, \ldots, \lambda_p\}$, en met b) dat de som van de geometrische multipliciteiten n is,

g) $$(A - \lambda_j I)^- := \sum_{i \neq j} (\lambda_i - \lambda_j)^{-1} E_i$$

is de rg-inverse van $A - \lambda_j I$ waarvan de range $R(A - \lambda_j I)$ en de nulruimte $N(A - \lambda_j I)$ is (vergelijk dit met (5.5.24)),

h) voor ieder polynoom φ geldt

$$\varphi(A) = \sum_i \varphi(\lambda_i) E_i \ ;$$

φ is dus annihilerend polynoom dan en slechts dan als $\forall_i \varphi(\lambda_i) = 0$; het minimaalpolynoom van A is

$$\varphi_0(z) = \prod_i (z - \lambda_i) \ ,$$

i) als L_j het j-de interpolatiepolynoom van Lagrange is op de punten $\lambda_1, \ldots, \lambda_p$ dan volgt uit h) dat

$$L_j(A) = \sum_i L_j(\lambda_i) E_i = E_j \ .$$

De formule (12.3.6), in samenhang met (12.3.5), heet de spectraalvoorstelling van A. Uit de resultaten in f) t/m i) blijkt hoe met behulp van de spectraalvoorstelling functies van A eenvoudig gedefinieerd of beschreven kunnen worden.

C) We beginnen nog eens opnieuw. Stel dat $A \in M_{n,n}$ als minimaalpolynoom heeft

$$\varphi_0(z) = \prod_{i=1}^{p} (z - \lambda_i) , \qquad (12.3.7)$$

waarin $\lambda_1,\dots,\lambda_p$ onderling verschillend zijn. Zij weer $L_j(z)$ het j-de interpolatiepolynoom van Lagrange op $\lambda_1,\dots,\lambda_p$. Dan heeft L_j graad p-1 en als $p \geq 2$ is dan geldt (waarom?)

$$\sum_i L_i(z) = 1 , \quad \sum_i \lambda_i L_i(z) = z . \qquad (*)$$

Definieer nu

$$E_j := L_j(A) , \quad j = 1,\dots,p . \qquad (12.3.8)$$

Dan geldt (ga na)

a) daar de graad van L_j lager is dan die van φ_0 is $E_j \neq 0$,

b) daar voor $i \neq j$ $\varphi_0(z)$ deler is van $L_i(z)L_j(z)$ volgt uit $\varphi_0(A) = 0$ dat $E_iE_j = 0$ voor $i \neq j$,

c) uit (*) volgt dat, als $p \geq 2$,

$$\Sigma E_i = I , \quad \Sigma \lambda_i E_i = A$$

en deze relaties gelden ook als p = 1 omdat dan $A = \lambda_1 I$ en $E_1 = I$.

Hiermee zijn we terug bij het uitgangspunt (12.3.5) en (12.3.6) van B).

<u>Stelling</u>. (12.3.9)

Zij $A \in M_{n,n}(\mathcal{C})$, $p \geq 1$ en $\lambda_1,\dots,\lambda_p \in \mathcal{C}$ onderling verschillend.
Dan zijn de volgende beweringen equivalent:

i) A is diagonaliseerbaar en $\sigma(A) = \{\lambda_1,\dots,\lambda_p\}$,

ii) er zijn $E_1,\dots,E_p \in M_{n,n}$ zo dat

$$E_j \neq 0, \; E_iE_j = 0 \text{ voor } i \neq j, \; \Sigma E_i = I, \; \Sigma\lambda_iE_i = A ,$$

iii) het minimaalpolynoom van A is $\varphi_0(z) := \Pi(z - \lambda_i)$,

iv) $\{\lambda_1,\dots,\lambda_p\} \subset \sigma(A)$ en $\Sigma \dim(N(A - \lambda_i I)) = n$.

De projectoren $E_1,\dots,E_p$ uit ii) zijn eenduidig door A bepaald.

<u>Bewijs</u>.

In de beschouwingen van (12.3.2) bleek onder meer:
in A) dat i) ⇒ ii) , iii) en iv),

in B) dat ii) $\Rightarrow$ iii) en iv),
in C) dat iii) $\Rightarrow$ ii).
De bewering dat iv) $\Rightarrow$ i) staat in (12.1.4.4).
De eenduidigheid van de projectoren E_j uit ii) volgt uit het feit dat in B) bleek dat uit ii) volgt dat $E_j = L_j(A)$, waarbij de Lagrange-polynomen L_j geheel door de eigenwaarden van A bepaald zijn. □

Stelling. (12.3.10)

Zij $A = \mathrm{diag}(A_1,\dots,A_k)$ met vierkante A_j. Dan geldt:
A diagonaliseerbaar $\leftrightarrow$ alle A_j diagonaliseerbaar.

Bewijs.

Volgt uit (12.2.5.4) en het feit dat i) en iii) uit (12.3.9) equivalent zijn.

□

Stelling. (12.3.11)

Zij $A \in M_{n,n}$ diagonaliseerbaar en $V \subset \mathcal{L}^n$ een t.o.v. A invariante deelruimte (dus $A(V) \subset V$). Dan is er een basis voor V bestaande uit eigenvectoren van A.

Bewijs.

Stel dat de kolommen van $U_1 \in M^k_{n,k}$ een basis voor V vormen. Dan is er een $B_{11} \in M_{k,k}$ zo dat

$$AU_1 = U_1B_{11} .$$

Voor ieder polynoom φ geldt dan $\varphi(A)U_1 = U_1\varphi(B_{11})$ en dus is het minimaalpolynoom φ_0 van A annihilerend polynoom van B_{11} (want $U_1 \in RR$). Hieruit volgt dat het minimaalpolynoom van B_{11} ontbindbaar is (in $\mathcal{L}$) in eerstegraadsfactoren, waaruit met (12.3.9) volgt dat B_{11} diagonaliseerbaar is. Zij $B_{11} = T_1\Lambda_1T_1^{-1}$ met T_1 regulier en Λ_1 diagonaal. Dan vormen de kolommen van $S_1 := U_1T_1$ ook een basis voor V en er geldt $AS_1 = AU_1T_1 = U_1B_{11}T_1 = U_1T_1\Lambda_1 = S_1\Lambda_1$ zo dat de kolommen van S_1 eigenvectoren van A zijn. □

Opmerking. (12.3.12)

B_{11} is de matrix van de restrictie van de afbeelding A tot de deelruimte V t.o.v. de basis gevormd door de kolommen van U_1. Uit bovenstaand bewijs volgt dat de diagonaliseerbaarheid van A die van B_{11} impliceert (dit is een generalisatie van (12.3.10)) en dat $\sigma(B_{11}) \subset \sigma(A)$.

Opgave. (12.3.13)

Zij $A \in M_{m,m}$, $B \in M_{n,n}$ en $X \in M^r_{m,n}$ zo dat

$$AX = XB .$$

Als A of B diagonaliseerbaar is dan zijn er $Y_1 \in M^r_{m,r}$, $Z_1 \in M^r_{r,n}$ en $\Lambda_1 \in M_{r,r}$ (diagonaal) zo dat

$$X = Y_1Z_1,\ AY_1 = Y_1\Lambda_1,\ Z_1B = \Lambda_1Z_1 .$$

Opgave. (12.3.14)

Zij A en V als in (12.3.11). Dan is er een t.o.v. A invariante complementaire deelruimte V' (d.w.z., $V \oplus V' = \mathcal{L}^n$).
Hint: zij $\sigma(A) = \{\lambda_1,\ldots,\lambda_p\}$ en $V_j := V \cap N(A - \lambda_j I)$.
Dan is (12.3.11) equivalent met $V = V_1 \oplus \ldots \oplus V_p$.

Stelling. (12.3.15)

Zij $A_1,\ldots,A_m \in M_{n,n}$ diagonaliseerbaar en zij

$$A_kA_\ell = A_\ell A_k,\ k,\ell = 1,\ldots,m .$$

Dan zijn er een reguliere U en diagonaalmatrices $\Lambda_1,\ldots,\Lambda_m$ zo dat

$$U^{-1}A_kU = \Lambda_k,\ k = 1,\ldots,m$$

(d.w.z., $A_1,\ldots,A_m$ worden door dezelfde gelijkvormigheidstransformatie gediagonaliseerd).

Bewijs.

Inductie naar m. Het geval m = 1 is triviaal.
Zij m > 1. Daar A_1 diagonaliseerbaar is, is er een reguliere S zo dat

$$S^{-1}A_1S = \mathrm{diag}(\lambda_{11}I_1,\ldots,\lambda_{1p}I_p) ,$$

waarin $\lambda_{11},\ldots,\lambda_{1p}$ onderling verschillend zijn en I_j afmetingen $n_j \times n_j$ heeft.
Definieer voor later gebruik

$$B_{1j} := \Lambda_{1j} := \lambda_{1j}I_j ,\quad j = 1,\ldots,p ,$$

$$B_1 := S^{-1}A_1S = \mathrm{diag}(B_{11},\ldots,B_{1p}) = \mathrm{diag}(\Lambda_{11},\ldots,\Lambda_{1p}) . \qquad (*)$$

Zij nu

$$B_k := S^{-1}A_kS\ , \quad k = 2,\dots,m\ .$$

Uit het gegeven volgt dat deze matrices diagonaliseerbaar zijn en ook dat

$$B_kB_\ell = B_\ell B_k\ , \quad k,\ell = 1,\dots,m\ . \tag{**}$$

Neem hierin $\ell = 1$. Uit de vorm (*) van B_1 en het feit dat $\lambda_{11},\dots,\lambda_{1p}$ onderling verschillend zijn volgt (ga na) dat voor $k = 2,\dots,p$

$$B_k = \mathrm{diag}(B_{k1},\dots,B_{kp})\ ,$$

waarin B_{kj} afmetingen $n_j{\times}n_j$ heeft.
Uit (12.3.10) volgt dat de B_{kj} diagonaliseerbaar zijn en uit (**) dat ook

$$B_{kj}B_{\ell j} = B_{\ell j}B_{kj}\ , \quad k,\ell = 2,\dots,m\ , \quad j = 1,\dots,p\ .$$

Uit de inductieveronderstelling volgt nu dat er reguliere matrices $U_1,\dots,U_p$ zijn zo dat

$$U_j^{-1}B_{kj}U_j = \Lambda_{kj}\ , \quad k = 2,\dots,m\ , \quad j = 1,\dots,p\ , \tag{***}$$

met diagonale Λ_{kj} (afmetingen $n_j{\times}n_j$).
Omdat $B_{1j} = \Lambda_{1j}$ een veelvoud van de eenheidsmatrix I_j is geldt (***) echter ook voor $k = 1$.
Hieruit volgt dat, als

$$U = S\ \mathrm{diag}(U_1,\dots,U_p)\ ,$$

voor $k = 1,\dots,m$ geldt

$$U^{-1}A_kU = \mathrm{diag}(\Lambda_{k1},\dots,\Lambda_{kp}) =: \Lambda_k\ . \qquad \square$$

<u>Opgave</u>. (12.3.16)

Zij $A \in M_{n,n}$ en $\lambda \in \sigma(A)$. Bewijs dat de volgende beweringen equivalent zijn:
i) $N(A - \lambda I) \oplus R(A - \lambda I) = \mathcal{L}^n$ (d.w.z., A heeft index 1, zie (5.5.27)),
ii) λ is enkelvoudig nulpunt van het minimaal polynoom φ_0 van A,
iii) er is een reguliere matrix U zo dat

$$U^{-1}AU = \mathrm{diag}(\lambda I_1, B_2)\ ,$$

met B_2 zo dat $\lambda \notin \sigma(B_2)$.

Hint voor i) $\Rightarrow$ ii): $\lambda \in \sigma(A) \Rightarrow \varphi_0(z) = (z - \lambda)^h \psi(z)$. $h > 1$ zou impliceren dat $N(A - \lambda I) \cap R(A - \lambda I) \neq \{\underline{0}\}$.
Hint voor ii) $\Rightarrow$ i): Als $\varphi_0(\lambda) = 0$, $\varphi_0'(\lambda) \neq 0$ dan is $\varphi_0(z) = \Sigma\alpha_j(z - \lambda)^j$ met $\alpha_0 = 0$, $\alpha_1 \neq 0$ en dus is er een polynoom χ zo dat $A - \lambda I = (A - \lambda I)^2\chi(A)$.

Opgave. (12.3.17)

Zij $A \in M_{n,n}$ en λ enkelvoudig nulpunt van het minimaalpolynoom φ_0 van A. Dan is de matrix E gedefinieerd door

$$L(z) := \frac{\varphi_0(z)}{(z - \lambda)\varphi_0'(\lambda)}\ , \quad E = L(A)$$

de projector op $N(A - \lambda I)$ langs $R(A - \lambda I)$. Bewijs hiermee dat in (12.3.16) i) uit ii) volgt.

Opgave. (12.3.18)

Als de karakteristiek van $\mathcal{L}$ nul is dan is bewering ii) van (12.3.9) equivalent met:
er zijn projectoren $E_1,\ldots,E_p \in M_{n,n}$ zo dat $E_i \neq 0$, $\Sigma E_i = I$, $\Sigma\lambda_i E_i = A$.
Zie (5.5.8).

Opgave. (12.3.19)

Zij $A = S\Lambda S^{-1}$ met S en Λ als in (12.3.3). Zij ook $A = S'\Lambda'S'^{-1}$ met S' regulier en Λ'diagonaal.
Bewijs dat er een permutatiematrix P is zo dat

$$P^T\Lambda'P = \Lambda\ , \quad S'P = S\ \mathrm{diag}(Z_{11},\ldots,Z_{pp})\ ,$$

waarin de diagonaalblokken Z_{ii} afmetingen $n_i \times n_i$ hebben en regulier zijn. Deze uitspraak beschrijft de mate van niet-eenduidigheid van de diagonaliseerbaarheid. Geef ook een meetkundige interpretatie.

13. Reductie van een algemene matrix

In deze paragraaf is $A \in M_{n,n}(\mathcal{L})$ met willekeurig lichaam $\mathcal{L}$, behalve in 13.4, daar is $\mathcal{L} = \mathbb{C}$.

13.1. Index en multipliciteitenrij van een eigenwaarde. Nilpotente matrices

Stelling. (13.1.1)

Zij $A \in M_{n,n}(\mathcal{L})$ en $\lambda \in \sigma(A)$.
Zij $N_j := N((A - \lambda I)^j)$ en $\mu_j := \dim(N_j)$.
Dan is er een h met $1 \le h \le n$ zo dat

$$\{\underline{0}\} = N_0 \subset N_1 \subset \ldots \subset N_h = N_{h+1} = \ldots$$

met echte inclusies en (dus)

$$0 = \mu_0 < \mu_1 < \ldots < \mu_h = \mu_{h+1} = \ldots$$

Bewijs.

Zeker geldt $N_{j-1} \subset N_j$. Daaruit volgt met (5.4.8.1) dat $\mu_{j-1} \le \mu_j$ met gelijkteken dan en slechts dan als $N_{j-1} = N_j$.
Stel nu $N_{j-1} = N_j$. We bewijzen dat dan ook $N_j = N_{j+1}$. Stel nl. $x \in N_{j+1}$. Dan is $(A - \lambda I)x \in N_j = N_{j-1}$, dus $x \in N_j$.
De rij $\mu_0, \mu_1, \ldots$ is dus strict stijgend totdat hij constant wordt (en dit moet vanaf een $h \le n$ omdat $\mu_j \le n$ voor alle j). □

Gevolg. (13.1.2)

Als $R_j := R((A - \lambda I)^j)$ dan geldt, met echte inclusies,

$$\mathcal{L}^n = R_0 \supset R_1 \supset \ldots \supset R_h = R_{h+1} = \ldots .$$

Definitie. (13.1.3)

Het in (13.1.1) optredende getal h heet index van de eigenwaarde λ en de rij getallen

$$\Sigma(\lambda) := (\mu_1, \ldots, \mu_h)$$

heet multipliciteitenrij van λ.
Het getal μ_h heet algebraische multipliciteit van λ (en μ_1 is de geometrische multipliciteit).

Stelling. (13.1.4)

Index en multipliciteitenrij van een eigenwaarde van A zijn invariant onder gelijkvormigheidstransformaties van A en ook bij overgang van A naar A^T.

Stelling. (13.1.5)

De rij getallen $\mu_0, \mu_1, \ldots$ is concaaf, voor $j \geq 1$ geldt

$$\mu_{j+1} - \mu_j \leq \mu_j - \mu_{j-1} .$$

Bewijs.

Pas de ongelijkheid van Frobenius (4.1.7) toe op het product

$$(A - \lambda I)^{j+1} = (A - \lambda I).(A - \lambda I)^{j-1}.(A - \lambda I) . \qquad \square$$

Opgave. (13.1.6)

Schets de grafiek van μ_j als functie van j, rekening houdend met (13.1.1) en (13.1.5). Bewijs dat, voor $1 \leq j \leq h$, $\mu_1 + j - 1 \leq \mu_j \leq j\mu_1$. Bewijs dat $\mu_1 = 1$ impliceert $\mu_j = j$. Bewijs (13.1.1) uit (13.1.5). Bepaal de multipliciteitenrij van de eigenwaarde 0 van

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

(alle niet ingevulde elementen zijn nul).

Opgave. (13.1.7)

Bepaal minimaalpolynoom, eigenwaarden en multipliciteitenrijen van de matrix

$$A = \begin{pmatrix} 0 & \alpha & \beta \\ & 1 & \gamma \\ & & 0 \end{pmatrix} .$$

Opgave. (13.1.8)

Zij

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline & A_{22} \end{array}\right)$$

met A_{11} en A_{22} vierkant. Zij $\lambda \in \sigma(A_{ii})$ met index h_i en algebraïsche multipliciteit m_i $(i = 1,2)$. Dan is $\lambda \in \sigma(A)$ met index h en algebraïsche multipliciteit m, waarbij

$$\max(h_1,h_2) \le h \le h_1 + h_2, \qquad m = m_1 + m_2 .$$

Hint: Stel $\lambda = 0$. Merk op dat

$$A^k = \left[\begin{array}{c|c} A_{11}^k & A_{12}^{(k)} \\ \hline & A_{22}^k \end{array}\right]$$

met $A_{12}^{(k)} = \sum_{\ell=1}^{k} A_{11}^{k-\ell} A_{12} A_{22}^{\ell-1}$.

Laat zien dat voor $k \ge h_1 + h_2$ geldt $A_{12}^{(k)}(N(A_{22}^k)) \subset R(A_{11}^k)$, waaruit volgt dat $\dim(N(A^k)) = \dim(N(A_{11}^k)) + \dim(N(A_{22}^k))$ (zie ook (5.5.26)).

Opgave. (13.1.9)

Bewijs dat de algebraïsche multipliciteit van een eigenwaarde λ van een driehoeksmatrix A gelijk is aan het aantal der diagonaalelementen van A die gelijk zijn aan λ.

Hint: inductie naar n met behulp van (13.1.8).

Stelling. (13.1.10)

Zij $A \in M_{m,n}$, $B \in M_{n,m}$. Dan hebben AB en BA dezelfde eigenwaarden $\neq 0$ en hierbij ook dezelfde multipliciteitenrijen.

Bewijs.

Stel $\lambda \in \sigma(AB)$, $\lambda \neq 0$. Zij $U_j \in RR$, $R(U_j) = N((AB - \lambda I)^j)$, $j \ge 1$.
Uit $(AB - \lambda I)^j U_j = 0$ volgt $0 = B(AB - \lambda I)^j U_j = (BA - \lambda I)^j BU_j$, dus $R(BU_j) \subset N((BA - \lambda I)^j)$. Uit $(AB - \lambda I)^j U_j = 0$ en $\lambda \neq 0$ volgt echter dat $N(BU_j) \subset N(U_j) = \{0\}$, zodat $BU_j \in RR$ en dus

$$\dim(N((BA - \lambda I)^j)) \ge r(BU_j) = r(U_j) = \dim(N((AB - \lambda I)^j)) .$$

Analoog omgekeerd. □

Opmerking. (13.1.11)

Voor de eigenwaarde $\lambda = 0$ geldt de bewering in het algemeen niet (wel als A of B regulier, want dan zijn AB en BA gelijkvormig). Voorbeeld:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Stelling. (13.1.12)

De volgende uitspraken zijn equivalent:

i) λ is eigenwaarde van A met index h,

ii) λ is h-voudig nulpunt van het minimaalpolynoom φ_0 van A.

Bewijs.

Zonder beperking der algemeenheid (ga na) veronderstellen we dat $\lambda = 0$. Omdat volgens (12.2.6) $0 \in \sigma(A) \Leftrightarrow \varphi_0(0) = 0$ zijn we klaar als we bewijzen dat

$$(*) \qquad \varphi_0(z) = z^h\psi(z) \text{ met } \psi(0) \neq 0$$

impliceert dat $\mu_{h-1} < \mu_h = \mu_{h+1}$ (met $\mu = \dim(N(A^j))$).
Daar φ_0 minimaalpolynoom van A is volgt uit (*) dat $A^h\psi(A) = 0$, $A^{h-1}\psi(A) \neq 0$, dus $R(\psi(A))$ is deelruimte van $N(A^h)$ en niet van $N(A^{h-1})$, dus $N(A^{h-1})$ is echte deelruimte van $N(A^h)$, dus $\mu_{h-1} < \mu_h$.
Uit (*) volgt ook dat er een polynoom χ is zo dat

$$\psi(0)z^h = \varphi_0(z) + z^{h+1}\chi(z) .$$

Daar $\psi(0) \neq 0$ impliceert dit dat $N(A^{h+1}) \subset N(A^h)$, dus $N(A^{h+1}) = N(A^h)$, dus $\mu_{h+1} = \mu_h$. □

Definitie. (13.1.13)

$A \in M_{n,n}(\mathcal{L})$ heet nilpotent als er een $k \in \mathbb{N}$ is zo dat $A^k = 0$.
Als $A^h = 0$ en $A^{h-1} \neq 0$ dan heet h index van nilpotentie. De rij getallen $\Sigma := (\mu_1,\ldots,\mu_h)$ met $\mu_j := \dim(N(A^j))$ heet multipliciteitenrij.

Stelling. (13.1.14)

De volgende beweringen zijn equivalent:

i) A nilpotent met index h,

ii) $\sigma(A) = \{0\}$, de eigenwaarde 0 heeft index h en algebraïsche multipliciteit n,

iii) het minimaalpolynoom van A is $\varphi_0(z) = z^h$.

Bewijs.

Bewering i) is equivalent met

$$A^{h-1} \neq A^h = 0 \ , \qquad (*)$$

bewering ii) is equivalent met

$$A - \lambda I \text{ regulier voor } \lambda \neq 0 \text{ en } \dim(N(A^{h-1})) < \dim(N(A^h)) = n \ .$$

Hieruit blijkt direct de equivalentie van i) en ii), aangezien een nilpotente matrix geen andere eigenwaarden dan 0 kan hebben (als $Ax = \lambda x$ met $\lambda \neq 0$ en $x \neq 0$ dan is $A^j x \neq 0$ voor alle j).
Verder volgt uit i) dat z^h annihilerend polynoom is en z^{h-1} niet. Het minimaalpolynoom is dus deler van z^h en niet van z^{h-1}. Dus $\varphi_0(z) = z^h$. Uit iii) volgt direct (*) en dus i). □

Opmerkingen. (13.1.15)

1. Als $\mathcal{L} = \mathbb{C}$ dan geldt zelfs (ga na):

$$A \text{ nilpotent} \Leftrightarrow \sigma(A) = \{0\} \ .$$

2. De index van nilpotentie is gelijk aan de index van de eigenwaarde 0. Analoog de multipliciteitenrij. Hieruit volgt met (13.1.1) dat $h \leq n$.

Opgave. (13.1.16)

Een driehoeksmatrix is dan en slechts dan nilpotent als alle diagonaalelementen nul zijn.

Opgave. (13.1.17)

Zij A driehoeksmatrix. Bewijs met behulp van (13.1.16) dat de algebraïsche multipliciteit van een eigenwaarde λ gelijk is aan het aantal der diagonaalelementen van A die gelijk zijn aan λ.
Hint: stel $\lambda = 0$. Als de bovenste m diagonaalelementen van A nul en de andere ongelijk nul zijn dan is volgens (13.1.16) het bovenste $m \times m$ blok van A nilpotent en het onderste $(m - n) \times (m - n)$ blok is regulier. Het algemene geval kan hiertoe teruggebracht worden door op te merken dat er bij iedere A van de vorm

$$A = \begin{pmatrix} a & c \\ 0 & b \end{pmatrix}$$

een reguliere S is (ga na!) zo dat

$$S^{-1}AS = \begin{pmatrix} b & c \\ 0 & a \end{pmatrix} .$$

Opgave. (13.1.18)

Zij $A_2 - A_1 = H$ met H nilpotent en $A_1H = HA_1$ (dus ook $A_2H = HA_2$). Dan is $\sigma(A_2) = \sigma(A_1)$ en iedere eigenwaarde heeft bij A_2 dezelfde algebraïsche multipliciteit als bij A_1.
Hint: als H index h heeft dan is $N(A_1^k) \subset N(A_2^{k+h-1})$.

13.2. Grofreductie

Zij $0 \in \sigma(A)$ met index h. Zij $N_j := N(A^j)$, $R_j := R(A^j)$.
We weten uit (13.1.1) en (13.1.2) dat

$$\{0\} = N_0 \subset N_1 \subset \ldots \subset N_h = N_{h+1} = \ldots \qquad (*)$$

$$\mathcal{L}^n = R_0 \supset R_1 \supset \ldots \supset R_h = R_{h+1} = \ldots \qquad (**)$$

met allemaal echte inclusies. Verder is voor iedere j $\dim(R_j) + \dim(N_j) = n$.
Wanneer is $R_j \oplus N_j = \mathcal{L}^n$?

Stelling. (13.2.1)

Zij $0 \in \sigma(A)$ met index h. Met N_j en R_j als boven geldt voor $i \geq 1$, $j \geq 1$

1) $N_i \cap R_j = \{0\} \Leftrightarrow j \geq h$,

2) $N_i + R_j = \mathcal{L}^n \Leftrightarrow i \geq h$.

Bewijs.

De eerste bewering is equivalent met $\forall_y$: $A^{i+j}y = 0 \Rightarrow A^iy = 0$, dus met $N_{i+j} \subset N_j$. Uit $(*)$ blijkt dat hiervoor nodig en voldoende is dat $j \geq h$. De tweede bewering is equivalent met $\forall_x \exists_y$: $A^i(x - A^jy) = 0$, dus met $R_i \subset R_{i+j}$. Uit $(**)$ blijkt dat hiervoor nodig en voldoende is dat $i \geq h$. □

Gevolgen. (13.2.2)

1. Als $0 \in \sigma(A)$ met index h dan geldt

$$N(A^h) \oplus R(A^h) = \mathcal{L}^n .$$

2. Daar voor iedere i en j $A(N(A^j)) = N(A^{j-1}) \subset N(A^j)$ en $A(R(A^i)) = R(A^{i+1}) \subset R(A^i)$ kunnen we zeggen: $N(A^h)$ en $R(A^h)$ zijn complementaire, t.o.v. A invariante deelruimten van $\mathcal{L}^n$.

Opgave. (13.2.3)

Bewijs dat, als $\mu_j := \dim(N_j)$, voor $i,j \geq 0$ geldt

$$\dim(N_i \cap R_j) = \mu_{i+j} - \mu_j ,$$
$$\dim(N_i + R_j) = n - \mu_{i+j} + \mu_i .$$

Stelling. (13.2.4)

De volgende uitspraken zijn equivalent:

i) 0 is h-voudig nulpunt van het minimaalpolynoom φ_0 van A,

ii) 0 is eigenwaarde van A met index h,

iii) er is een reguliere matrix U zo dat

$$U^{-1}AU = \mathrm{diag}(B_{11}, B_{22}) ,$$

waarin B_{11} nilpotent is met index h en B_{22} regulier is.

Bovendien geldt in deze situatie:

iv) de afmetingen van B_{11} zijn gelijk aan de algebraïsche multipliciteit van de eigenwaarde 0 van A en de multipliciteitenrij van B_{11} is dezelfde als die van de eigenwaarde 0 van A,

v) als φ_2 het minimaalpolynoom van B_{22} is dan is

$$\varphi_0(z) = z^h \varphi_2(z) .$$

Bewijs.

De equivalentie van i) en ii) is reeds in (13.1.12) bewezen. We bewijzen hieronder dat ii) $\Rightarrow$ iii), iv) en v). De implicatie iii) $\Rightarrow$ ii) is een direct gevolg van (13.1.4), (13.1.13) en het feit dat B_{22} regulier is.
Stel nu dat ii) geldt. Volgens (13.2.2) zijn dan $N(A^h)$ en $R(A^h)$ complementaire, t.o.v. A invariante deelruimten van $\mathcal{C}^n$. Zij $U = (U_1 \mid U_2) \in M_{n,n}$ regulier en zo dat $R(U_1) = N(A^h)$, $R(U_2) = R(A^h)$. U_1 en U_2 hebben dan breedte m (de algebraïsche multipliciteit van de eigenwaarde 0 van A), resp. $n-m$ en er zijn matrices $B_{11} \in M_{m,m}$ en $B_{22} \in M_{n-m,n-m}$ zo dat

$$AU_1 = U_1 B_{11},\ AU_2 = U_2 B_{22},\ AU = U\,\mathrm{diag}(B_{11}, B_{22}) .$$

Er geldt dan ook voor iedere k $A^k U_1 = U_1 B_{11}^k$, $A^k U_2 = U_2 B_{22}^k$. Uit $R(U_1) = N(A^h)$ volgt (daar $U_1 \in RR$) dat $B_{11}^h = 0$ en (daar $U_2 \in RR$) dat B_{22}^h en dus ook B_{22} regulier is. B_{11} is dus nilpotent en daar B_{22} regulier is, is $\dim(N(A^k)) =$

$= \dim(N(B_{11}^k))$, zodat B_{11} dezelfde index en multipliciteitenrij heeft als de eigenwaarde 0 van A. Tenslotte volgt uit (13.1.14), (12.2.5.4) en (13.1.4) direct dat iii) $\Rightarrow$ v). $\square$

Opmerking. (13.2.5)

De bewering v) omvat i). Derhalve is (omdat $0 \in \sigma(A) \Leftrightarrow \varphi_0(0) = 0$) de equivalentie van i) en ii) hierboven opnieuw aangetoond.

Opgave. (13.2.6)

Als $U^{-1}AU = \mathrm{diag}(B_{11}, B_{22})$, $V^{-1}AV = \mathrm{diag}(C_{11}, C_{22})$ met B_{11} en C_{11} nilpotent, B_{22} en C_{22} regulier, dan zijn B_{11} en C_{11} gelijkvormig, evenals B_{22} en C_{22}; $U^{-1}V$ is van de vorm $\mathrm{diag}(W_{11}, W_{22})$.
(Hier wordt dus de essentiële eenduidigheid van de reductie (13.2.4) aangetoond.)

Opgave. (13.2.7)

Zij φ_2 als in (13.2.4). Dan is

$$R(\varphi_2(A)) = N(A^h) .$$

Opgave. (13.2.8)

Zij A als in (13.2.4). Zij $U = (U_1 \mid U_2)$ regulier en zo dat $R(U_1) = N(A^h)$. Dan geldt

$$U^{-1}AU = \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline & B_{22} \end{array}\right)$$

met B_{11} en B_{22} als in (13.2.4). Bewijs dit zonder (13.2.4) te gebruiken.

Opgave. (13.2.9)

Als

$$B = \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline & B_{22} \end{array}\right)$$

met B_{11} nilpotent en B_{22} regulier dan is er precies één V van de vorm

$$V = \left(\begin{array}{c|c} I_{11} & V_{12} \\ \hline & I_{22} \end{array}\right)$$

zo dat $V^{-1}BV = \mathrm{diag}(B_{11},B_{22})$.
Hint: Bewijs dat $V_{12}B_{22} - B_{11}V_{12} = B_{12}$ een stelsel lineaire vergelijkingen voor de elementen van V_{12} is dat (dank zij het feit dat B_{11} nilpotent en B_{22} regulier is) voor iedere B_{12} een eenduidige oplossing heeft.

Stelling (grofreductie). (13.2.10)

Zij $A \in M_{n,n}(\mathcal{L})$ met minimaalpolynoom φ_0, spectrum $\{\lambda_1,\ldots,\lambda_p\}$, indices $h_1,\ldots,h_p$ en multipliciteitenrijen $\Sigma_1,\ldots,\Sigma_p$. Dan zijn de volgende beweringen equivalent.

i) Er is een reguliere S zo dat

$$S^{-1}AS = \mathrm{diag}(\lambda_1 I_1 + H_1,\ldots,\lambda_p I_p + H_p)$$

met $H_1,\ldots,H_p$ nilpotent, H_j heeft index h_j en multipliciteitenrij Σ_j,

ii) de som van de algebraïsche multipliciteiten van de eigenwaarden van A is n,

iii) het minimaalpolynoom van A is ontbindbaar in lineaire factoren:

$$\varphi_0(z) = \prod_1^p (z - \lambda_j)^{h_j},$$

iv) er is een diagonaliseerbare A_0 en een nilpotente H zo dat

$$A = A_0 + H,\ A_0H = HA_0$$

(er geldt dan bovendien dat $\sigma(A_0) = \sigma(A)$ en dat een eigenwaarde λ bij beide matrices dezelfde algebraïsche multipliciteit heeft).

Bewijs.

Uit i) volgen ii) en iii) met behulp van (13.1.4), (12.2.5.2), (12.2.5.4) en (13.1.14).
Dat uit ii) of iii) weer i) volgt wordt bewezen door volledige inductie naar p. Voor p = 1 volgt de bewering uit (13.1.14). De inductiestap volgt zowel in het geval ii) ⇒ i) als in het geval iii) ⇒ i) uit (13.2.4) en (13.1.4) (ga na).
iv) volgt uit i) door te nemen $A_0 = S\,\mathrm{diag}(\lambda_1 I_1,\ldots,\lambda_p I_p)S^{-1}$ en op te merken dat voor een diagonaliseerbare matrix algebraïsche en geometrische multipliciteit samenvallen.
We bewijzen nu dat iv) ⇒ iii). Stel dat H index h heeft. Dan geldt, dankzij de verwisselbaarheid van A_0 en H dat voor iedere $\mu \in \mathcal{L}$

$$(A - \mu I)^h = \sum_{j=0}^{h-1} \binom{h}{j} H^j (A_0 - \mu I)^{h-j} = C(\mu)(A_0 - \mu I) ,$$

waarin $C(\mu)$ verwisselbaar is met A_0. Uit (12.3.9) volgt dat het minimaalpolynoom van A_0 van de vorm $\Pi(z - \mu_j)$ is. Dit impliceert $\Pi(A_0 - \mu_j I) = 0$, dus ook $\Pi(A - \mu_j I)^h = 0$, dus $\Pi(z - \mu_j)^h$ is annihilerend polynoom van A. Het minimaalpolynoom van A is daar deler van, is dus ontbindbaar in lineaire factoren.

□

Opmerking. (13.2.11)

Zij A als in de aanhef van (13.2.10) en stel dat de beweringen van (13.2.10) niet gelden. Dan volgt uit (13.2.4) en (13.1.4) door inductie naar p (ga na) dat er een reguliere S is zodat

$$S^{-1}AS = \mathrm{diag}(\lambda_1 I_1 + H_1, \ldots, \lambda_p I_p + H_p, B_{p+1}) ,$$

met voor $j \le p$ H_j als in (13.2.10) terwijl B_{p+1} geen eigenwaarden in $\mathcal{L}$ heeft. Het minimaalpolynoom van A is nu

$$\varphi_0(z) = \prod_1^p (z - \lambda_j)^{h_j} \psi(z) ,$$

waarin ψ het minimaalpolynoom van B_{p+1} is (dat (in $\mathcal{L}$) geen nulpunten heeft).

In 13.3 gaan we verder in het geval dat $A \in M_{n,n}(\mathbb{C})$. Dan is het minimaalpolynoom van A ontbindbaar in lineaire factoren en dus gelden de uitspraken van (13.2.10). Voor het geval dat het minimaalpolynoom niet volledig ontbindbaar is, noemen we nu nog een paar resultaten. Met name bewijzen we dat de graad van het minimaalpolynoom van een $A \in M_{n,n}$ altijd $\le n$ is (ga na dat dit voor het geval van een in lineaire factoren ontbindbaar minimaalpolynoom uit (13.2.10) en (13.1.15.2) volgt!).

Stelling. (13.2.12)

Zij φ_1 en φ_2 niet-triviale polynomen. Als $\omega = \mathrm{ggd}(\varphi_1, \varphi_2)$ dan zijn er polynomen ψ_1 en ψ_2 zo dat

$$\omega = \varphi_1\psi_1 + \varphi_2\psi_2 .$$

Opgave. (13.2.13)

Bewijs (13.2.12).
Hint: inductie naar min(graad(φ_1),graad(φ_2)) met behulp van de algorithme van Euclides: als graad(φ_2) $\le$ graad(φ_1) dan is of φ_2 een deler van φ_1 en dus ggd(φ_1,φ_2) = φ_2 of er is een polynoom χ en een niet-triviaal polynoom φ_3 met graad(φ_3) < graad(φ_2) zodat $\varphi_1 = \varphi_2\chi + \varphi_3$; in het laatste geval is ggd(φ_1,φ_2) = ggd(φ_2,φ_3).

Stelling. (13.2.14)

Zij $A \in M_{n,n}$ en φ annihilerend polynoom van A. Zij $\varphi = \varphi_1\varphi_2$ waarin φ_1 en φ_2 relatief priem (d.w.z. ggd(φ_1,φ_2) = 1) zijn en hoogstegraadscoëfficiënt 1 hebben. Dan is er een reguliere U zo dat

$$U^{-1}AU = \mathrm{diag}(B_{11},B_{22}) ,$$

waarbij φ_i annihilerend polynoom van B_{ii} is (i = 1,2).
φ_1 en φ_2 zijn minimaalpolynoom van B_{11}, resp. B_{22} dan en slechts dan als φ minimaalpolynoom van A is.

Bewijs.

Daar ggd(φ_1,φ_2) = 1 zijn er volgens (13.2.12) polynomen ψ_1 en ψ_2 zo dat

$$\varphi_1\psi_1 + \varphi_2\psi_2 = 1 .$$

Derhalve geldt:

a) $\qquad \varphi_1(A)\varphi_2(A) = 0 ,$

b) $\qquad \varphi_1(A)\psi_1(A) + \varphi_2(A)\psi_2(A) = I .$

Uit a) volgt

$$R(\varphi_2(A)) \subset N(\varphi_1(A)),\ R(\varphi_1(A)) \subset N(\varphi_2(A))$$

en uit b) volgt (ga na)

$$N(\varphi_1(A)) \subset R(\varphi_2(A)),\ N(\varphi_2(A)) \subset R(\varphi_1(A)) .$$

Derhalve geldt

$$N(\varphi_1(A)) = R(\varphi_2(A)) =: V_1,\ N(\varphi_2(A)) = R(\varphi_1(A)) =: V_2 .$$

Uit b) volgt ook (ga na)

$$N(\varphi_1(A)) \cap N(\varphi_2(A)) = \{\underline{0}\}, \ R(\varphi_1(A)) + R(\varphi_2(A)) = \mathcal{L}^n ,$$

dus

$$V_1 \oplus V_2 = \mathcal{L}^n .$$

Daar $A(V_1) = R(A\varphi_2(A)) \subset R(\varphi_2(A)) = V_1$ en analoog voor V_2 zijn V_1 en V_2 complementaire, t.o.v. A invariante deelruimten van $\mathcal{L}^n$. Als $U = (U_1 \mid U_2)$ regulier is en zo dat $R(U_i) = V_i$ $(i = 1,2)$ dan zijn er dus B_{11} en B_{22} zo dat $AU_i = U_iB_{ii}$, dus $AU = U \operatorname{diag}(B_{11},B_{22})$.
Uit $\varphi_i(A)U_i = U_i\varphi_i(B_{ii})$ volgt (daar $R(U_i) = V_i = N(\varphi_i(A))$ en $U_i \in RR$) dat φ_i annihilerend polynoom van B_{ii} is.
Zij φ_{i0} het minimaalpolynoom van B_{ii}. Dan is φ_{i0} deler van φ_i, dus φ_{10} en φ_{20} zijn relatief priem. Volgens (12.2.5.1) is het minimaalpolynoom van A dan gelijk aan $\varphi_{10}\varphi_{20}$. Hieruit volgt dat $\varphi = \varphi_1\varphi_2$ het minimaalpolynoom van A is dan en slechts dan als $\varphi_{10} = \varphi_1$, $\varphi_{20} = \varphi_2$. □

<u>Opgave</u>. (13.2.15)

Zij $E_1 := \varphi_2(A)\psi_2(A)$, $E_2 := \varphi_1(A)\psi_1(A)$. Bewijs dat E_1 de projector is op V_1 langs V_2 en E_2 de projector op V_2 langs V_1.

<u>Opgave</u>. (13.2.16)

Zij φ, φ_1 en φ_2 als in (13.2.14). Zij $\tilde{\varphi}_1$ zo dat φ_1 deler is van $\tilde{\varphi}_1$ en dat $\operatorname{ggd}(\tilde{\varphi}_1,\varphi_2) = 1$. Dan geldt

$$R(\tilde{\varphi}_1(A)) = R(\varphi_1(A)), \ N(\tilde{\varphi}_1(A)) = N(\varphi_1(A)) .$$

Merk op dat (13.1.1) en (13.1.2) hier zeer speciale gevallen zijn!

<u>Stelling</u> (priemdecompositiestelling). (13.2.17)

Zij $A \in M_{n,n}(\mathcal{L})$ met minimaalpolynoom φ_0. Zij

$$\varphi_0(z) = \prod_1^p (\varphi_j(z))^{h_j} ,$$

waarin $\varphi_1,\ldots,\varphi_p$ onontbindbaar (in $\mathcal{L}$), onderling verschillend, met graad ≤ 1 en met hoogste graadscoëfficiënt 1 zijn. Dan is er een reguliere S zo dat

$$S^{-1}AS = \operatorname{diag}(B_{11},\ldots,B_{pp}) ,$$

waarin B_{jj} minimaalpolynoom $(\varphi_j)^{h_j}$ heeft $(j = 1,\ldots,p)$.

Bewijs.

Volledige inductie naar p op basis van (13.2.14). □

Stelling. (13.2.18)

De graad van het minimaalpolynoom van een $A \in M_{n,n}(\mathcal{L})$ is niet groter dan n.

Bewijs.

Uit (13.2.17) volgt dat we de stelling alleen hoeven te bewijzen voor het geval dat het minimaalpolynoom φ_0 van A de vorm $\varphi_0 = \varphi^h$ heeft met φ niet ontbindbaar in $\mathcal{L}$.

Zij $(u_1,\ldots,u_n)$ een basis voor $\mathcal{L}^n$. Definieer de polynomen ψ_j $(j = 1,\ldots,n)$ door

$$\psi_j(A)u_j = 0,\ \tilde{\psi}_j(A)u_j = 0 \Rightarrow \text{graad}(\tilde{\psi}_j) \geq \text{graad}(\psi_j)\ ,$$

$$\text{hoogstegraadscoëfficiënt van } \psi_j = 1\ .$$

Stellig is graad$(\psi_j) \leq n$. Want graad$(\psi_j) = m$ betekent dat de vectoren $u_j, Au_j, \ldots, A^{m-1}u_j$ onafhankelijk zijn.

Anderzijds zijn $\psi_1,\ldots,\psi_n$ delers van φ_0 en is zelfs (ga na)

$$\varphi_0 = \text{kgv}(\psi_1,\ldots,\psi_n)\ ;$$

gezien de speciale vorm van φ_0 volgt hieruit dat minstens één der ψ_j gelijk moet zijn aan φ_0. Dus graad$(\varphi_0) \leq n$. □

13.3. Reductie van matrices over $\mathbb{C}$

Afspraak. (13.3.1)

In deze paragraaf is $A \in M_{n,n}(\mathbb{C})$ met spectrum $\sigma(A) = \{\lambda_1,\ldots,\lambda_p\}$, indices $h_1,\ldots,h_p$, multipliciteitenrijen $\Sigma_1,\ldots,\Sigma_p$, geometrische multipliciteiten $n_1,\ldots,n_p$ en algebraïsche multipliciteiten $m_1,\ldots,m_p$.

Stelling. (13.3.2)

Zij A als in (13.3.1). Dan is er een reguliere S zo dat

$$S^{-1}AS = \text{diag}(\lambda_1 I_{11} + H_{11},\ldots,\lambda_p I_{pp} + H_{pp})\ ,$$

waarin I_{jj} en H_{jj} afmetingen $m_j \times m_j$ hebben en H_{jj} nilpotent is met index h_j en multipliciteitenrij Σ_j.

Het minimaalpolynoom van A is

$$\prod_j (z - \lambda_j)^{h_j} .$$

Bewijs.

Zij φ_0 het minimaalpolynoom van A. Daar in $\mathbb{C}$ de hoofdstelling van de algebra geldt is φ_0 volledig factoriseerbaar:

$$\varphi_0(z) = \prod_{j=1}^{p} (z - \lambda_j)^{h_j} ,$$

waarin volgens (12.2.6) de getallen $\lambda_1,\ldots,\lambda_p$ het volledige spectrum van A vormen en volgens (13.2.4) $h_1,\ldots,h_p$ de bijbehorende indices zijn.
De rest volgt uit (13.2.10). □

Opmerkingen. (13.3.3)

1) Men kan ook een rechtstreeks bewijs van (13.3.2) geven, uitgaande van (12.2.7.2) (existentie van tenminste één eigenwaarde) en (13.2.4) (reductiestap). Voer dit uit en merk op dat op deze manier het begrip minimaalpolynoom nauwelijks nodig is (alleen voor het existentiebewijs in (12.2.7.2) maar dat zou ook via het karakteristieke polynoom kunnen).
2) Het karakteristieke polynoom van A is

$$\Delta(z) = \prod_j (z - \lambda_j)^{m_j} ,$$

d.w.z., de algebraïsche multipliciteit van een eigenwaarde λ_j is de macht waarmee de factor $(z - \lambda_j)$ voorkomt in $\Delta(z)$.
Immers, met de determinantentheorie volgt uit (13.3.2) dat

$$\Delta(z) = \prod_j \Delta_j(z) ,$$

waarin Δ_j het karakteristieke polynoom van $\lambda_j I_{jj} + H_{jj}$ is. Daar Δ_j graad m_j heeft en als enig nulpunt de enige eigenwaarde λ_j van $\lambda_j I_{jj} + H_{jj}$ heeft, moet $\Delta_j(z) = (z - \lambda_j)^{m_j}$ (merk op dat deze uitspraak berust op de factoriseerbaarheid van Δ_j in lineaire factoren, dus op het feit dat we over $\mathbb{C}$ werken).
Gevolg: de stelling van Hamilton-Cayley (zie (12.2.8)) geldt in $M_{n,n}(\mathbb{C})$.

3) De volgende relaties voor h_j, n_j en m_j volgen uit (13.1.1), (13.1.5) en (13.3.2)

$$1 \le h_j \le m_j, \quad p \le \Sigma h_j \le n \,,$$

$$1 \le n_j \le m_j, \quad p \le \Sigma n_j \le n \,,$$

$$n_j + h_j - 1 \le m_j \le h_j n_j \,.$$

Uit de laatste relatie (of rechtstreekser) volgt dat

$$h_j = 1 \Leftrightarrow n_j = m_j, \; n_j = 1 \Leftrightarrow h_j = m_j \,.$$

4) Uit bovenstaande ongelijkheden volgt ook (ga na)

$$\Sigma h_j = p \Leftrightarrow \forall_j : h_j = 1 \Leftrightarrow \forall_j : n_j = m_j \Leftrightarrow \Sigma n_j = n \,,$$

$$\Sigma n_j = p \Leftrightarrow \forall_j : n_j = 1 \Leftrightarrow \forall_j : h_j = m_j \Leftrightarrow \Sigma h_j = n \,,$$

$$p = n \Leftrightarrow \Sigma h_j = \Sigma n_j = n \Leftrightarrow \forall_j : h_j = n_j = 1 \,.$$

<u>Definitie</u>. (13.3.4)

Zij A als in (13.3.1).

1) A heet gedegenereerd als $p < n$, d.w.z. A heeft minder dan n verschillende eigenwaarden.
2) A heet defectief als $\Sigma n_j < n$, d.w.z. er bestaat niet een n-tal onafhankelijke eigenvectoren van A.
3) A heet derogatoir*) als $\Sigma h_j < n$; d.w.z. de graad van het minimaalpolynoom van A is kleiner dan n (of: het minimaalpolynoom is een echte deler van het karakteristiek polynoom).

<u>Eigenschappen</u>. (13.3.5)

Uit (13.3.4) volgt

A niet gedegenereerd $\Leftrightarrow$ A heeft n verschillende eigenwaarden ($p = n$)
$\Leftrightarrow \forall_j : m_j = 1$
$\Leftrightarrow \forall_j : h_j = 1, \; n_j = 1$
$\Leftrightarrow$ A niet defectief en niet derogatoir ($\Sigma n_j = \Sigma h_j = n$),

A niet defectief $\Leftrightarrow$ A is diagonaliseerbaar ($\Sigma n_j = n$)
$\Leftrightarrow \forall_j : n_j = m_j$
$\Leftrightarrow \forall_j : h_j = 1$
$\Leftrightarrow$ graad minimaalpolynoom is p ($\Sigma h_j = p$),

*) to derogate = schenden, verkleinen

A niet derogatoir $\Leftrightarrow$ graad minimaalpolynoom is n ($\Sigma h_j = n$)

$\Leftrightarrow \forall_j : h_j = m_j$

$\Leftrightarrow \forall_j : n_j = 1$

$\Leftrightarrow$ A heeft slechts p onafhankelijke eigenvectoren ($\Sigma n_j = p$).

Opgave. (13.3.6)

A heet irreducibele linker-Hessenberg matrix als

$$(A)_{ij} = 0 \text{ voor } j > i+1, \; (A)_{i,i+1} \neq 0 \quad (1 \le i < n) \; .$$

Bewijs dat zo'n A niet derogatoir is.
Hint: Voor iedere λ heeft $A - \lambda I$ minstens rang $n-1$.

Opgave. (13.3.7)

Zij

$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{pmatrix}.$$

a) Bewijs dat het minimaalpolynoom van A is

$$\varphi_0(z) = z^n + a_1 z^{n-1} + \ldots + a_n \; .$$

A heet de bij het polynoom φ_0 behorende companion-matrix.
Hint: Uit $Ae_{j+1} = e_j - a_{n-j} e_n$ $(1 \le j < n)$ en $Ae_1 = -a_n e_n$ volgt dat $(A^n + a_1 A^{n-1} + \ldots + a_n I)e_n = 0$; uit (13.3.6) volgt dat graad$(\varphi_0) = n$.

b) Zij $x(\lambda) := (1, \lambda, \ldots, \lambda^{n-1})^T$. Bewijs dat

$$(\lambda I - A)x(\lambda) = \varphi_0(\lambda) e_n \; . \qquad (*)$$

c) Zij λ h-voudig nulpunt van φ_0. Definieer voor $j = 1, \ldots, h$

$$x_j = \frac{1}{(j-1)!} \frac{d^{j-1}}{d\lambda^{j-1}} x(\lambda), \; X = (x_1 | \ldots | x_h) \; .$$

Bewijs dat $X \in RR$ en dat $AX = X(\lambda I + J)$ met

$$J = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \quad (h \times h) \ .$$

Hint: $j - 1$ keer differentiëren van $(*)$ levert $Ax_j = \lambda x_j + x_{j-1}$ $(2 \le j \le h)$.

Opgave. (13.3.8)

Een matrix $A \in M_{n,n}(\mathbb{C})$ is diagonaliseerbaar dan en slechts dan als er bij iedere t.o.v. A invariante deelruimte $V \subset \mathcal{L}^n$ een invariant complement V' is (dit is een omkering van (12.3.14)).
Hint: gebruik (12.2.5.4).

Als toepassing van de in $\mathbb{C}$ geldende priemdecompositiestelling (13.3.2) beschouwen we het volgende probleem.
Zij $A \in M_{m,m}(\mathbb{C})$ en $B \in M_{n,n}(\mathbb{C})$. Beschouw de afbeelding Φ: $M_{m,n}(\mathbb{C}) \to M_{m,n}(\mathbb{C})$ gegeven door

$$X \in M_{m,n}(\mathbb{C}) \mapsto Y := AX - XB \in M_{m,n}(\mathbb{C}) \ .$$

Is er bij iedere $Y \in M_{m,n}(\mathbb{C})$ een eenduidige $X \in M_{m,n}(\mathbb{C})$ met $\Phi(X) = Y$? Of anders: is Φ bijectief?

Stelling. (13.3.9)

De hierboven gedefinieerd afbeelding Φ is bijectief dan en slechts dan als $\sigma(A)$ en $\sigma(B)$ disjunct zijn.

Bewijs.

Φ is een lineaire afbeelding van de mn-dimensionale vectorruimte $M_{mn}(\mathbb{C})$ in zichzelf, Φ is dus bijectief dan en slechts dan als $N(\Phi) = \{0\}$.
Stel dat $\lambda \in \sigma(A) \cap \sigma(B)$. Dan is er een $u \in \mathbb{C}^m$ zodat $u \neq 0$ en $Au = \lambda u$ en (omdat $\sigma(B) = \sigma(B^T)$) een $v \in \mathbb{C}^n$ zo dat $v \neq 0$ en $v^TB = \lambda v^T$. Hieruit volgt dat $uv^T \in M_{m,n}$, $uv^T \neq 0$ en $Auv^T = \lambda uv^T = uv^TB$, dus $uv^T \in N(\Phi)$, dus $N(\Phi) \neq \{0\}$.
Stel nu dat $\sigma(A) \cap \sigma(B) = \emptyset$. Volgens (13.3.2) zijn er reguliere S en T zo dat

$$S^{-1}AS = \mathrm{diag}(\lambda_1 I_{11} + H_{11}, \ldots, \lambda_p I_{pp} + H_{pp}) \ ,$$

$$T^{-1}BT = \mathrm{diag}(\mu_1 I'_{11} + H'_{11}, \ldots, \mu_q I'_{qq} + H'_{qq})$$

waarin alle H_{ii} en H'_{jj} nilpotent zijn. Zij $X \in M_{m,n}$, $Z := S^{-1}XT$ en partitioneer de rijen van Z net als de diagonaal van $S^{-1}AS$ en de kolommen van Z net als de diagonaal van $T^{-1}BT$. Dan is $AX = XB$ equivalent met

$$(\lambda_i I_{ii} + H_{ii})Z_{ij} = Z_{ij}(\mu_j I'_{jj} + H'_{jj}),\ i = 1,\dots,p,\ j = 1,\dots,q\ .$$

We kunnen hiervoor ook schrijven

$$C_{ij}Z_{ij} = Z_{ij}H'_{jj} \text{ met } C_{ij} := (\lambda_i - \mu_j)I_{ii} + H_{ii}\ .$$

Hieruit volgt dat voor alle $k \in \mathbb{N}$t $C_{ij}^k Z_{ij} = Z_{ij}(H'_{jj})^k$. Daar C_{ij} regulier is (omdat $\lambda_i \neq \mu_j$) en H'_{jj} nilpotent is volgt nu eenvoudig dat $Z_{ij} = 0$. Dit geldt voor alle i en j, zodat uit $AX = XB$ volgt dat $X = 0$. □

Opgave. **(13.3.10)**

Zij $H \in M_{m,m}$, I' en $H' \in M_{n,n}$ nilpotent met indices h, resp. h'.
Dan heeft voor $\alpha \neq 0$ de vergelijking

$$\alpha X - HX - XH' = Y$$

voor iedere $Y \in M_{m,n}$ een unieke oplossing $X \in M_{m,n}$. Bepaal voor X een uitdrukking van de vorm

$$X = \sum_{i=0}^{h-1}\sum_{j=0}^{h'-1} \gamma_{ij} H^i Y H'^j\ .$$

Geef aan hoe met behulp hiervan en van (13.3.2) de oplossing van $AX - XB = Y$ in het geval dat $\sigma(A) \cap \sigma(B) = \emptyset$ bepaald kan worden.

14. Fijnreductie van een nilpotente matrix. Stelling van Jordan

14.1. Geometrische aanpak

Afspraak. (14.1.1)

In deze paragraaf is $A \in M_{m,m}(\mathcal{L})$, nilpotent met index h en multipliciteitenrij $\Sigma = (\mu_1,\dots,\mu_h)$ met $\mu_h = m$.
Zij verder $N_j := N(A^j)$ met dimensie μ_j en

$$\nu_j := \mu_j - \mu_{j-1},\ 1 \le j \le h,\ \nu_{h+1} := 0\ .$$

We weten al (13.1.1) dat $\{\underline{0}\} = N_0 \subset N_1 \subset \dots \subset N_h = \mathcal{L}^m$ met allemaal echte inclusies. Ook geldt $A(N_j) \subset N_{j-1}$. We willen nu deelruimten $V_1,\dots,V_h$ van resp. $N_1,\dots,N_h$ construeren zo, dat

$$N_j = N_{j-1} \oplus V_j,\ \text{dus } N_j = V_1 \oplus\dots\oplus V_j,\ j = 1,\dots,h\ .$$

Natuurlijk is dan $A^i(V_j) \subset N_{j-i}$. We kunnen de V_j zo kiezen dat zelfs $A^i(V_j) \subset V_{j-i}$ en $\dim(A^i(V_j)) = \dim(V_j)$ als $i < j$.

Na de constructie van de V_j kiezen we bases in de V_j en wel zo dat als x basisvector in V_j is, $A^i x$ basisvector in V_{j-i} is $(1 \le i < j)$. De matrix van de met A corresponderende afbeelding ten opzichte van de aldus gevonden basis voor $\mathcal{L}^m$ krijgt dan een zeer eenvoudige vorm.

Stelling. (14.1.2)

Als A voldoet aan (14.1.1) dan zijn er deelruimten $V_1,\dots,V_h$ van $\mathcal{L}^m$ zo dat

i) $N_j = N_{j-1} \oplus V_j = V_1 \oplus \dots \oplus V_j, \ 1 \le j \le h$,

ii) $A(V_j) \subset V_{j-1}$ voor $2 \le j \le h$, $A(V_1) = \{\underline{0}\}$.

Voor deze V_j geldt bovendien

iii) $A^i(V_j) \subset V_{j-i}$ en $\dim(A^i(V_j)) = \nu_j$ voor $0 \le i < j \le h$.

Bewijs.

1) We construeren de V_j in de volgorde $V_h,\dots,V_1$.
2) Aangezien $N_{h-1} \subset N_h$ is er volgens (5.4.15) een V_h zo dat $N_{h-1} \oplus V_h = N_h$.
3) Zij $1 \le j < h$. Veronderstel dat V_{j+1} zo is dat

$$N_j \oplus V_{j+1} = N_{j+1}$$

(voor $j = h - 1$ is op grond van 2) aan deze veronderstelling voldaan). Natuurlijk is dan $A(V_{j+1}) \subset A(N_{j+1}) \subset N_j$. Anderzijds is $A(V_{j+1}) \cap N_{j-1} = \{\underline{0}\}$, want als $x \in V_{j+1}$ en $Ax \in N_{j-1}$ dan is $x \in N_j$, dus $x \in N_j \cap V_{j+1} = \{\underline{0}\}$.
4) Daar volgens 3) $A(V_{j+1})$ en N_{j-1} disjuncte deelruimten zijn van N_j bestaat $N_{j-1} \oplus A(V_{j+1})$ en is er volgens (5.4.15) een deelruimte $V_j' \subset N_j$ (mogelijk met dimensie 0) zo dat

$$(N_{j-1} \oplus A(V_{j+1})) \oplus V_j' = N_j \ .$$

Dan is stellig $A(V_{j+1}) \cap V_j' = \{\underline{0}\}$, zodat $V_j := A(V_{j+1}) \oplus V_j'$ bestaat. En hiervoor geldt (omdat de direct som associatief is (zie (5.1.17))),

$$N_{j-1} \oplus V_j = N_j \ .$$

Hiermee is de constructie van $V_h,\dots,V_1$, die aan i) en ii) voldoen, geheel aangegeven.

5) Uit (5.4.15) volgt dat

$$\dim(V_j) = \dim(N_j) - \dim(N_{j-1}) = \mu_j - \mu_{j-1} = \nu_j .$$

Uit ii) volgt door inductie naar i dat $A^i(V_j) \subset V_{j-i}$ voor $i < j$. En uit (5.5.21) volgt

$$\dim(A^i(V_j)) = \dim(V_j) - \dim(V_j \cap N_i) .$$

Maar voor $0 < i < j$ is $V_j \cap N_i = \{\underline{0}\}$, omdat $N_i \subset N_{j-1}$ en $V_j \cap N_{j-1} = \{\underline{0}\}$. Dit bewijst iii). □

<u>Gevolg.</u> (14.1.3)

Daar $A(V_{j+1}) \subset V_j$ volgt uit iii)

$$\nu_{j+1} = \dim(A(V_{j+1})) \le \dim(V_j) = \nu_j .$$

Dit is een rechtstreeks bewijs van (13.1.5) (voor nilpotente matrices, maar het is eenvoudig te generaliseren voor matrices die 0 als eigenwaarde hebben met index h - ga na).

<u>Opgave.</u> (14.1.4)

Als, in de notatie van (14.1.1), $V'_h := V_h$ en $V'_{h-1},\ldots,V'_1$ zo zijn dat $A(V_{j+1}) \oplus V'_j = V_j$, dan geldt

$$V_j = A^{h-j}(V'_h) \oplus A^{h-j-1}(V'_{h-1}) \oplus \ldots \oplus V'_j .$$

Hint: bewijs eerst dat voor $i < j$

$$A^{i+1}(V_{j+1}) \oplus A^i(V'_j) = A^i(V_j) .$$

14.2. <u>Basiskeuze</u>

We gaan nu bases kiezen in $V_h,\ldots,V_1$.
Zij $(x_{h,1},\ldots,x_{h,\nu_h})$ een basis voor V_h (dimensie $\nu_h \ge 1$).
Daar $A(V_h) \subset V_{h-1}$ en $\dim(A(V_h)) = \dim(V_h) = \nu_h$ zijn de vectoren $x_{h-1,\ell} := Ax_{h,\ell}$, $1 \le \ell \le \nu_h$, onafhankelijke vectoren uit V_{h-1}. Als $\nu_{h-1} = \nu_h$ (d.w.z., $\dim(V_{h-1}) = \dim(V_h)$) dan vormen ze een basis voor V_{h-1}. Als $\nu_{n-1} > \nu_h$ dan kunnen we er nog $\nu_{h-1} - \nu_h$ vectoren $x_{h-1,\ell}$ bij kiezen zo dat de vectoren $x_{h-1,\ell}$, $1 \le \ell \le \nu_{h-1}$, een basis voor V_{h-1} vormen (de vectoren $x_{h-1,\ell}$ met $\nu_h < \ell \le \nu_{h-1}$ vormen een basis voor V'_h). Nu naar V_{h-2}, enz.

Het volgende pseudo-Algol programma beschrijft deze basiskeuze nog eens (bedenk dat $\nu_{h+1} = 0$):

$\underline{\text{for}}$ $j := h$ $\underline{\text{step}}$ -1 $\underline{\text{until}}$ 1 $\underline{\text{do}}$

$\underline{\text{begin}}$ $\underline{\text{for}}$ $\ell := 1$ $\underline{\text{step}}$ 1 $\underline{\text{until}}$ ν_{j+1} $\underline{\text{do}}$ $x_{j,\ell} := Ax_{j+1,\ell}$;

$\underline{\text{for}}$ $\ell := \nu_{j+1} + 1$ $\underline{\text{step}}$ 1 $\underline{\text{until}}$ ν_j $\underline{\text{do}}$

<kies $x_{j,\ell} \in V_j$ zo dat $x_{j,1},\ldots,x_{j,\ell}$ onafhankelijk zijn>

$\underline{\text{end}}$

Stelling. (14.2.1)

Zij A als in (14.1.1). De met de bovenstaande constructie gevonden vectoren $x_{j,\ell}$, $1 \le j \le h$, $1 \le \ell \le \nu_j$, vormen een basis voor $\mathcal{L}^m$. Daarbij vormen de vectoren $x_{j,\ell}$, $1 \le \ell \le \nu_j$, een basis voor V_j en er geldt

$$A^i x_{j,\ell} = \begin{cases} x_{j-i,\ell} & \text{als } 0 \le i < j\,, \\ 0 & \text{als } i \ge j\,. \end{cases}$$

Bewijs.

Dit volgt direct uit de constructie en het feit dat $V_1 \oplus \ldots \oplus V_h = \mathcal{L}^m$. De laatste bewering volgt uit de constructie en het feit dat $A^j(V_j) = \{\underline{0}\}$. □

Opmerking. (14.2.2)

Als $x \in N_j$ maar $x \notin N_{j-1}$ dan heet x hoofdvector (principal vector) van graad j van A bij de eigenwaarde 0 (de hoofdvectoren van graad 1 zijn dus de eigenvectoren). De verzameling van hoofdvectoren van graad j is geen lineaire deelruimte van $\mathcal{L}^m$. x is hoofdvector van graad j dan en slechts dan als er een $y \neq 0$ uit V_j is zo dat $x - y \in N_{j-1}$. In deze zin representeren de elementen van $V_j \backslash \{\underline{0}\}$ dus de verzameling van alle hoofdvectoren van graad j. Ook geldt: $N_j \backslash \{\underline{0}\}$ is de verzameling van alle hoofdvectoren met graad $\le$ j.

14.3. Normaalvormen voor nilpotente matrices

We willen nu van de met een nilpotente A corresponderende afbeelding $A: \mathcal{L}^m \to \mathcal{L}^m$ de matrix t.o.v. de in 14.2 geconstrueerde basis $(x_{j,\ell})$ bepalen. Deze basisvectoren zijn op dit moment echter tweedimensionaal (d.w.z., met twee indices) geordend:

$$V_1:\ \underbrace{x_{1,1},\dots,x_{1,\nu_h}}_{A^{h-1}(V_h)},\underbrace{\dots,x_{1,\nu_{h-1}}}_{A^{h-2}(V'_{h-1})},\dots,x_{1,\nu_2},\underbrace{\dots,x_{1,\nu_1}}_{V'_1},$$

$$\dots \qquad \dots \qquad \dots \qquad \dots$$

$$V_{h-1}:\ \underbrace{x_{h-1,1},\dots,x_{h-1,\nu_h}}_{A(V_h)},\underbrace{\dots,x_{h-1,\nu_{h-1}}}_{V'_{h-1}},$$

$$V_h:\ \underbrace{x_{h,1},\dots,x_{h,\nu_h}}_{V_h}. \tag{14.3.1}$$

We moeten ze nu in een eendimensionale rij ordenen.

A) Eerste ordening (horizontaal lexicografisch):

Neem

$$X = (X_1|\dots|X_h)$$

met

$$X_j = (x_{j,1}|\dots|x_{j,\nu_j}) .$$

De kolommen van X_j bevatten dus de gekozen basis voor V_j.
Uit (14.2.1) volgt direct (ga na)

$$AX_1 = 0,\ AX_j = X_{j-1}H_{j-1,j},\ 1 < j \le h ,$$

waarin

$$H_{j-1,j} = \begin{pmatrix} I_{jj} \\ \hline 0 \end{pmatrix} , \tag{14.3.2}$$

afmetingen $\nu_{j-1} \times \nu_j$, met I_{jj} de $\nu_j \times \nu_j$ eenheidsmatrix.
Hiermee is bewezen:

<u>Stelling</u>. (14.3.3)

Zij A als in (14.1.1). Dan is er een reguliere matrix X zo dat

$$X^{-1}AX = H$$

met

$$H = \begin{pmatrix} 0 & H_{12} & & & \\ & 0 & H_{23} & & \\ & & \ddots & \ddots & \\ & & & 0 & H_{h-1,h} \\ & & & & 0 \end{pmatrix},$$

waarin het j-de diagonaalblok afmetingen $\nu_j \times \nu_j$ heeft $(j = 1,\dots,h)$ en waarin de matrices $H_{j-1,j}$ $(2 \le j \le h)$ er uit zien als in (14.3.2).

<u>Opmerkingen</u>. 14.3.4)

1. Als A niet defectief is dan is $h = 1$, $\nu_1 = m$, dus $X = X_1$, $A = H = 0$ $(m \times m)$. Als A niet derogatoir is dan is $h = m$ en alle ν_j zijn 1; er geldt dan $H = J^{(m)}$, met

$$J^{(m)} := \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}, \qquad (m \times m) .$$

2. De bij A horende matrix H is eenduidig bepaald door de getallen $\nu_1,\dots,\nu_h$, dus door de multipliciteitenrij Σ van A. Dit geldt ook omgekeerd: als H bekend is dan is de multipliciteitenrij bekend. H wordt daarom een normaalvorm voor de nilpotente matrix A genoemd; de zgn. blok-Jordan normaalvorm. De matrix X is niet eenduidig bepaald omdat hij afhangt van de keuze van de bases voor $V_h, V'_{h-1},\dots,V'_1$.

<u>Opgave</u>. (14.3.5)

Bewijs dat voor $1 \le k < h$

$$X^{-1}A^kX = H^k = \begin{pmatrix} 0 & \cdots & H_{1,k+1} & & \\ & \ddots & & \ddots & \\ & & \ddots & & H_{h-k,h} \\ & & & \ddots & \vdots \\ & & & & 0 \end{pmatrix}$$

met

$$H_{j-k,j} = \begin{pmatrix} I_{jj} \\ \hline 0 \end{pmatrix},$$

afmetingen $\nu_{j-k} \times \nu_j$ $(j = k+1,\dots,h)$.

Bepaal hieruit rechtstreeks de nulruimten van H^k en A^k.
(Gebruik dat $H_{j-k,j} \in RR$!)

B) Tweede ordening (verticaal lexicografisch).

Neem:

$$Y = (Y_1|Y_2|\dots|Y_{\nu_1})$$

met

$$Y_\ell = (x_{1,\ell}|\dots|x_{j,\ell}) ,$$

waarbij j door ℓ bepaald wordt uit de eis

$$\nu_{j+1} < \ell \le \nu_j \quad (1 \le j \le h), \text{ of: } j = \max\{i \mid \nu_i \ge \ell\} .$$

Daar $0 = \nu_{h+1} < \nu_h \le \dots \le \nu_1$ is voor iedere ℓ met $1 \le \ell \le \nu_1$ hierdoor eenduidig een j bepaald.
De kolommen van Y_ℓ zijn de vectoren die in de ℓ-de kolom van het schema (14.3.1) onder elkaar staan. Als $0 < \ell \le \nu_h$ dan bevat Y_ℓ h kolommen. Als $\nu_h < \ell \le \nu_{h-1}$ dan bevat Y_ℓ $h-1$ kolommen, maar als $\nu_h = \nu_{h-1}$ dan komen Y_ℓ van dit type niet voor in Y. Enz., enz. De Y_ℓ met $\nu_2 < \ell \le \nu_1$ bevatten (als ze bestaan, d.w.z. als $\nu_2 < \nu_1$) ieder precies één kolom.

Er volgt nu, als $\nu_{j+1} < \ell \le \nu_j$, uit (14.2.1) dat

$$AY_\ell = Y_\ell J^{(j)} ,$$

waarin

$$J^{(j)} = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix} , \qquad (14.3.6)$$

met afmetingen $j \times j$ (een zg. Jordankast).
Hieruit volgt:

<u>Stelling.</u> (14.3.7)

Zij A als in (14.1.1). Er is een reguliere matrix Y zo dat

$$Y^{-1}AY = J = \mathrm{diag}(\underbrace{J^{(h)},\dots,J^{(h)}}_{\nu_h \text{ stuks}},\underbrace{J^{(h-1)},\dots,J^{(h-1)}}_{\nu_{h-1}-\nu_h \text{ stuks}},\dots,\underbrace{J^{(1)},\dots,J^{(1)}}_{\nu_1-\nu_2 \text{ stuks}}).$$

Opmerkingen. (14.3.8)

1. Als A niet defectief is dan is $h = 1$ en $J = 0$ (merk op dat $J^{(1)}$ de 1×1 nulmatrix is). Als A niet derogatoir is dan is $h = m$ en $\nu_j = 1$ voor $1 \leq j \leq h$, derhalve is nu $J = J^{(m)}$.
2. De bij A horende matrix J is ook eenduidig bepaald door de multipliciteitenrij Σ van A en omgekeerd, de matrix Y uiteraard niet. J heet de zgn. Jordan normaalvorm van de nilpotente matrix A.
3. De door de kolommen van een Y_ℓ met breedte j opgespannen deelruimte wordt door A in zichzelf afgebeeld. Men noemt deze deelruimte de t.o.v. A cyclische deelruimte horend bij de vector $x_{1,\ell} \in N(A)$. De door de kolommen van Y_ℓ gegeven basis van deze cyclische deelruimte bevat één eigenvector, nl. $x_{1,\ell}$, een hoofdvector van graad 2, nl. $x_{2,\ell},\ldots$, een hoofdvector van graad h. $\mathcal{L}^n$ is aldus opgesplitst als directe som van

ν_h	cyclische	deelruimten	met	dimensie	h
$\nu_{h-1} - \nu_h$	"	"	"	"	$h-1$
-----	-----	-----	-----	-----	-----
$\nu_1 - \nu_2$	"	"	"	"	1.

Ga na dat analoog bij iedere eigenvector y_0 een A-cyclische deelruimte $\{y_0,\ldots,y_{j-1}\}$ hoort zo dat $Ay_i = y_{i-1}$, $1 \leq i \leq j$, waarbij de dimensie j van deze deelruimte bepaald wordt door de eis dat $y_0 \in R(A^{j-1})$ en $y_0 \notin R(A^j)$. Het minimaalpolynoom van de matrix $J^{(j)}$ die de afbeelding van een j-dimensionale cyclische deelruimte in zichzelf beschrijft is z^j (ga na). Cyclische deelruimten spelen ook een belangrijke rol bij de beschrijving van de fijnstructuur van matrices waarvan het minimaalpolynoom een macht is van een (in $\mathcal{L}$) niet ontbindbaar polynoom met graad > 1.

Opgave. (14.3.9)

Zij J een $m \times m$ Jordankast. Zij $1 \leq k \leq m$. Bepaal de (unieke!) t.o.v. J invariante deelruimte V van $\mathcal{L}^m$ met dimensie k. Merk op dat V een t.o.v. J cyclische deelruimte is.

Opgave. (14.3.10)

Zij $A \in M_{m,m}$ nilpotent. Bewijs dat iedere t.o.v. A invariante deelruimte V van $\mathcal{L}^m$ directe som is van t.o.v. A cyclische deelruimten.
Hint: generaliseer het bewijs van (14.1.2) en kies daarna bases zoals in (14.3.7).

Door hetzij z'n blok-Jordan normaalvorm, hetzij z'n Jordan normaalvorm wordt een nilpotente matrix volledig beschreven behoudens een gelijkvormigheids-transformatie. Uit (14.3.4.2) en (14.3.8.2) volgt direct

Stelling. (14.3.11)

Zij A en A' nilpotent. Dan zijn de volgende uitspraken equivalent:

i) A en A' zijn gelijkvormig.

ii) A en A' hebben dezelfde multipliciteitenrij Σ.

iii) A en A' hebben dezelfde blok-Jordan normaalvorm.

iv) A en A' hebben dezelfde Jordan normaalvorm.

Opgave. (14.3.12)

Zij A nilpotent met blok-Jordan normaalvorm H en Jordan normaalvorm J. Dan is er een permutatiematrix P zo dat $J = P^T HP$. Construeer H, J en P voor het geval dat $h = 3$, $\nu_3 = 1$, $\nu_2 = 2$, $\nu_1 = 2$.

14.4. Algebraïsche aanpak van de reductie naar blok-Jordan vorm

Het is mogelijk de fijnreductie van een nilpotente matrix A wat meer algebraisch te beschrijven dan in 14.1. We kiezen dan eerst op eenvoudige manier een basis in $\mathcal{L}^m$, zo dat de matrix C van de bij A horende afbeelding $\mathcal{A}$ t.o.v. deze basis nog niet erg mooi is, maar wel al een deel van het karakteristieke van de blok-Jordan matrix H heeft: C is blok-bovendiagonaal met nul-blokken op de diagonaal, terwijl de blokken $C_{j-1,j} \in RR$ zijn. Daarna bewijzen we de gelijkvormigheid van C met H door constructie van een matrix S zo dat $CS = SH$.

Stelling. (14.4.1)

Zij A nilpotent en verder als in (14.1.1). Dan is er een reguliere matrix U zo dat

$$U^{-1}AU = C$$

met

$$C = \begin{pmatrix} 0 & C_{12} & \cdots & C_{1,h} \\ & 0 & & C_{2,h} \\ & & \ddots & \vdots \\ & & & C_{h-1,h} \\ & & & 0 \end{pmatrix}$$

waarin het j-de diagonaalblok afmetingen $\nu_j \times \nu_j$ heeft $(j = 1,\dots,h)$ en de matrices $C_{j-1,j} \in RR$ met afmetingen $\nu_{j-1} \times \nu_j$.

Bewijs.

Kies $U = (U_1|\dots|U_h)$ zo dat $U_j \in RR$ en $N_{j-1} \oplus R(U_j) = N_j$, $j = 1,\dots,h$.
Dit kan, daar $N_{j-1} \subset N_j$. U_j heeft afmetingen $m \times \nu_j$. Dan is $AU_1 = 0$ en, daar $A(N_j) \subset N_{j-1}$, zijn er C_{ij} zo dat

$$AU_j = \sum_{i=1}^{j-1} U_i C_{ij}, \qquad 2 < j \leq h .$$

Daar $R(U) = R(U_1) \oplus\dots\oplus R(U_h) = R(N_h) = \mathcal{L}^m$ is U regulier.
We hoeven dus nog slechts te bewijzen dat $C_{j-1,j} \in RR$.
Stel dat $C_{j-1,j}y_j = 0$. Dan is

$$AU_j y_j = \sum_{i=1}^{j-2} U_i C_{ij} y_j \in N_{j-2} ,$$

dus $U_j y_j \in N_{j-1}$. Maar dat betekent $U_j y_j = 0$, want $R(U_j) \cap N_{j-1} = \{\underline{0}\}$. Daar $U_j \in RR$ is $y_j = 0$ en dus $C_{j-1,j} \in RR$. □

Opmerkingen. (14.4.2)

1. Uit $C_{j-1,j} \in RR$ volgt weer dat $\nu_{j-1} \geq \nu_j$ (13.1.5).
2. In (15.1.3) zal aangegeven worden hoe men $U_1, U_2, \dots$ kan construeren met een methode die numeriek goed te implementeren is.
3. Uit de structuur van C is de multipliciteitenrij van A direct af te lezen.
4. Uit (14.4.1) volgt dat voor een nilpotente matrix A geldt $tr(A) = 0$.

Stelling. (14.4.3)

Zij C als in (14.4.1). Dan is er een reguliere S zo dat

$$S^{-1}CS = H$$

met H (de blok-Jordan normaalvorm van C) als in (14.3.3).

Bewijs.

We proberen met een blok-bovendriehoeksmatrix S van de vorm

$$S = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1h} \\ & S_{22} & \cdots & S_{2h} \\ & & & \\ & & & S_{hh} \end{pmatrix},$$

waarbij S_{jj} afmetingen $\nu_j \times \nu_j$ heeft, aan de eisen te voldoen. S is regulier dan en slechts dan als de diagonaalblokken S_{jj} regulier zijn.
Van de gepartitioneerde vorm van de vergelijking CS = SH blijft als niet-triviaal over (ga na)

$$\sum_{\ell=i+1}^{j} C_{i\ell}S_{\ell j} = S_{i,j-1}H_{j-1,j}, \qquad 1 \le i < j \le h . \tag{*}$$

We partitioneren $S_{i,j-1}$ nog verder

$$S_{i,j-1} = (S'_{i,j-1} \mid S''_{i,j-1})$$

met $S'_{i,j-1}$ $\nu_i \times \nu_j$ en $S''_{i,j-1}$ $\nu_i \times (\nu_{j-1} - \nu_j)$. Dan is, gezien de vorm (14.3.2) van $H_{j-1,j}$, het rechterlid van (*) gelijk aan $S'_{i,j-1}$ en daarvoor moet dus gelden

$$S'_{i,j-1} = \sum_{\ell=i+1}^{j} C_{i\ell}S_{\ell j}, \qquad 1 \le i < j \le h . \tag{**}$$

Daar dit de enige vergelijkingen zijn waaraan voldaan moet worden, kunnen we S nu als volgt construeren. Kies de h-de blokkolom van S (dus de blokken $S_{1,h},\ldots,S_{h,h}$) en zo dat $S_{h,h}$ regulier is. Bepaal nu met (**) (en j = h) de eerste ν_h kolommen van de h-1ste blokkolom van S (dus de blokken $S'_{1,h-1},\ldots,S'_{h-1,h-1},0$). Vul (als $\nu_{h-1} > \nu_h$) de h-1ste blokkolom van S nog aan met de blokken $S''_{1,h-1},\ldots,S''_{h-1,h-1},0$ zo, dat $S_{h-1,h-1}$ regulier wordt. Kan dit laatste? Ja, als $S'_{h-1,h-1} \in RR$ is. Maar uit (**) volgt

$$S'_{h-1,h-1} = C_{h-1,h}S_{hh}$$

en daarin is $C_{h-1,h} \in RR$ en S_{hh} regulier, dus $S'_{h-1,h-1} \in RR$.
Analoog kunnen we de overige blokkolommen van S construeren. □

Opgave. (14.4.4)

Ga na dat de keuzevrijheid in S geheel overeen komt met de keuzevrijheid in de basisvectoren van de ruimten $V_h,\dots,V_1$ uit (14.1). Wat is de nulruimte van C^k?

Opgave. (14.4.5)

Bepaal bij een blok-Jordan normaalvorm H alle reguliere S zo dat SH = HS. Interpreteer het resultaat meetkundig. Welke vereenvoudigingen treden op als H niet derogatoir is?

Opgave. (14.4.6)

Bewijs dat er een transformatie $C = V^{-1}CV$ met $V = \text{diag}(V_1,\dots,V_h)$ van de matrix C uit (14.4.1) is zo dat de blokken $C_{j-1,j}$ de vorm $H_{j-1,j}$ uit (14.3.2) hebben. Bij de transformatie van C naar H, zoals in (14.4.3), kunnen de diagonaalblokken S_{jj} dan eenheidsmatrices zijn. Ga ook na hoe men de transformatie kan bepalen die C overvoert in de Jordan normaalvorm J.

14.5. De stelling van Jordan

Afspraak. (14.5.1)

In deze paragraaf is $A \in M_{n,n}(\mathbb{C})$ met spectrum $\sigma(A) = \{\lambda_1,\dots,\lambda_p\}$, geometrische multipliciteiten $n_1,\dots,n_p$, algebraische multipliciteiten $m_1,\dots,m_p$ en multipliciteitenrijen $\Sigma_1,\dots,\Sigma_p$.

Opmerking. (14.5.2)

Alle stellingen gelden ook voor $A \in M_{n,n}(\mathcal{L})$ mits A een in lineaire factoren ontbindbaar minimaalpolynoom heeft (zie (13.2.10)).

Stelling (Jordan). (14.5.3)

Als A aan (14.5.1) voldoet dan is er een reguliere S zo dat

$$S^{-1}AS = \text{diag}(\lambda_1 I_1 + J_1,\dots,\lambda_p I_p + J_p)$$

waarin

I_k de $m_k \times m_k$ eenheidsmatrix

J_k de volgens (14.3.7) bij een nilpotente matrix met multipliciteitenrij Σ_k horende Jordan normaalvorm.

Bewijs.

Combinatie van (13.3.2) (grofreductie naar diagonaalblokken $\lambda_k I_k + H_k$ met H_k nilpotent met multipliciteitenrij Σ_k) en (14.3.7) (fijnreductie van een nilpotente matrix, dus ook van matrices van de vorm $\lambda_k I_k + H_k$!). □

Opmerkingen. (14.5.4)

1. Uit (14.3.8.2) volgt dat het rechterlid van (14.5.3) door A eenduidig bepaald is behoudens permutatie van de diagonaalblokken $\lambda_k I_k + J_k$. Het rechterlid van (14.5.3) heet de bij A horende Jordan normaalvorm.
2. Voor een aantal toepassingen is het handiger, de fijnreductie van de J_k uit (14.5.3) uit te schrijven en in plaats van (14.5.3) te schrijven

$$S^{-1}AS = \mathrm{diag}(\lambda_1 I_{11} + J_{11}, \lambda_1 I_{12} + J_{12}, \ldots, \lambda_p I_{p,n_p} + J_{p,n_p}) ,$$

waarbij de $J_{k\ell}$ de vorm (14.3.6) hebben en iedere λ_k precies n_k maal voorkomt. In de terminologie van (14.3.8.2): als we S overeenkomstig partitioneren dan vormen de kolommen van de $S_{k\ell}$ bases voor cyclische deelruimten t.o.v. A die samen $\mathbb{C}^n$ opspannen.
3. Voor het spoor van A geldt (ga na)

$$\mathrm{tr}(A) = \Sigma m_j \lambda_j .$$

Stelling. (14.5.5)

Zij A en A' $\in M_{n,n}(\mathbb{C})$. De volgende beweringen zijn equivalent:

i) A en A' zijn gelijkvormig.
ii) A en A' hebben hetzelfde spectrum met dezelfde multipliciteitenrijen.
iii) A en A' hebben Jordan normaalvormen die behoudens permutaties van blokken gelijk zijn.

14.6. Enkele toepassingen van de stelling van Jordan

We geven in enigszins informele stijl enkele toepassingen van de stelling van Jordan.

1. Zij J een $h \times h$ Jordankast (van de vorm (14.3.6)) en $A = \lambda I + J$. Dan is, omdat I en J commuteren, voor $s \in \mathbb{N}t$

$$A^s = \sum_{\ell=0}^{s} \binom{s}{\ell} \lambda^{s-\ell} J^\ell = \sum_{\ell=0}^{\min(s,h-1)} \binom{s}{\ell} \lambda^{s-\ell} J^\ell .$$

Hieruit volgt, omdat $(J^\ell)_{ij} = \delta_{j,i+\ell} = \delta_{\ell,j-i}$ (ga na),

$$(A^s)_{ij} = \begin{cases} \binom{s}{j-i}\lambda^{s-j+i} & \text{als } 0 \le j - i \le \min(s,h-1) \\ 0 & \text{anders} \end{cases} .$$

Een gevolg hiervan is (bedenk dat $\binom{s}{\ell}$ een polynoom in s is met graad ℓ)

a) $\lim_{s\to\infty} A^s = 0 \Leftrightarrow |\lambda| < 1$,

b) A^s begrensd voor $s \to \infty$ $\Leftrightarrow |\lambda| < 1$ of $|\lambda| = 1$ en $h = 1$,

c) $$\lim_{s\to\infty} \sqrt[s]{|(A^s)_{ij}|} = \begin{cases} |\lambda| & \text{als } i \le j \\ 0 & \text{anders} \end{cases} ,$$

d) $\lim_{s\to\infty} \sqrt[s]{\|A^s\|} = |\lambda|$.

2. Voor een algemene matrix $A \in M_{n,n}(\mathbb{C})$ gelden de beweringen a), b) en d) evenzeer, mits de criteria rechts gelden voor iedere $\lambda \in \sigma(A)$. Het bewijs loopt via de stelling van Jordan in de vorm (14.5.4.2), waarbij alleen nog opgemerkt hoeft te worden dat natuurlijk $\lim S^{-1}A^sS = S^{-1}(\lim A^s)S$, enz. Met name geldt

$$\lim \sqrt[s]{\|A^s\|} = \max_{\lambda\in\sigma(A)} |\lambda| .$$

3. Analoge relaties kan men, als A regulier is, bewijzen voor A^{-s} $(s \in \mathbb{N}t)$, op grond van het feit dat

$$(\lambda I + J)^{-s} = \sum_{\ell=0}^{h-1} \binom{-s}{\ell}\lambda^{-s-\ell}J^\ell .$$

4. Beschouw de differentiaalvergelijking

$$\frac{dY}{dt} = AY , \qquad (14.6.1)$$

waarin A een $n \times n$ matrix is (onafhankelijk van t) en Y een $n \times n$ matrix is waarvan de elementen van t afhangen (de afgeleide van Y kunnen we door elementsgewijze differentiatie definiëren). Y heet fundamentele oplossing van (14.6.1) als Y voldoet aan de beginvoorwaarde

$$Y(0) = I . \qquad (14.6.2)$$

We zullen straks existentie (voor $-\infty < t < \infty$) en eenduidigheid van de fundamentele oplossing aantonen en tevens dat hij voldoet aan

$$\frac{dY}{dt} = AY = YA \ . \tag{14.6.3}$$

Eerst echter enkele gevolgtrekkingen en toepassingen.

a) Uit (14.6.3) volgt dat

$$\frac{d}{dt}\, Y(t)Y(t_0 - t) = Y(t)A.Y(t_0 - t) - Y(t).AY(t_0 - t) = 0 \ ,$$

dus

$$Y(t)Y(t_0 - t) = Y(0)Y(t_0) = Y(t_0) \ . \tag{*}$$

Door $t_0 = 0$ te nemen zien we hieruit dat $Y(t)$ regulier is en dat

$$(Y(t))^{-1} = Y(-t) \ . \tag{14.6.4}$$

b) Eveneens volgt uit (*) dat

$$Y(t_1 + t_2) = Y(t_1)Y(t_2) \ .$$

c) Zij $Z(t)$ een $n \times k$ matrix die voldoet aan

$$\frac{dZ}{dt} = AZ \ . \tag{14.6.5}$$

Dan geldt

$$\frac{d}{dt}\, Y(-t)Z(t) = -Y(-t)A.Z(t) + Y(-t).AZ(t) = 0 \ ,$$

waaruit met (14.6.4) volgt

$$Z(t) = Y(t)Z(0) \ . \tag{14.6.6}$$

Omgekeerd is voor iedere $C \in M_{n,k}$ $Y(t)C$ oplossing van (14.6.5). Met behulp van de fundamentele oplossing $Y(t)$ kunnen we dus alle oplossingen van (14.6.5) bepalen. Tevens volgt hieruit dat de fundamentele oplossing eenduidig is.

d) Het is eveneens niet moeilijk om te bewijzen dat de oplossingen van

$$\frac{dZ}{dt} = AZ + F$$

(met $F \in M_{n,k}$ en afhankelijk van t) gegeven worden door

$$\begin{aligned} Z(t) &= \int_0^t Y(t-\tau)F(\tau)\,d\tau + Y(t)Z(0) = \\ &= Y(t)\Big[\int_0^t Y(-\tau)F(\tau)\,d\tau + Z(0)\Big] \ . \end{aligned}$$

5. We construeren nu de fundamentele oplossing Y die aan (14.6.1) en (14.6.2) voldoet. Zij eerst weer $A = \lambda I + J$ met J een Jordankast. Dan luidt (14.6.1)

$$\frac{dY}{dt} = \lambda Y + JY \; . \tag{14.6.7}$$

Stel $Y(t) = e^{\lambda t}V(t)$. Dan gaat (14.6.7) over in

$$\frac{dV}{dt} = JV \tag{14.6.8}$$

met beginvoorwaarde $V(0) = I$. Uit (14.6.8) volgt echter

$$\frac{d^{\ell}V}{dt^{\ell}} = \begin{cases} J^{\ell}V & 0 \le \ell < h \\ 0 & \ell \ge h \; . \end{cases}$$

Dus $V(t)$ is een polynoom in t met graad $h-1$ en omdat

$$\frac{d^{\ell}V}{dt^{\ell}}(0) = J^{\ell}V(0) = J^{\ell}$$

geldt

$$V(t) = \sum_{\ell=0}^{h-1} \frac{t^{\ell}}{\ell!} J^{\ell} \; ,$$

dus

$$Y(t) = e^{\lambda t} \sum_{=0}^{h-1} \frac{t^{\ell}}{\ell!} J^{\ell} \; .$$

Een algemene matrix A kunnen we met (14.5.3) (of (14.5.4.2)) uitdrukken in z'n Jordan normaalvorm

$$A = S \operatorname{diag}(\lambda_1 I_1 + J_1, \ldots, \lambda_p I_p + J_p) S^{-1} \tag{14.6.9}$$

en voor de bijbehorende fundamentele oplossing Y geldt dan (ga na)

$$Y(t) = S \operatorname{diag}(e^{\lambda_1 t} \sum_{\ell=0}^{h_1-1} \frac{t^{\ell}}{\ell!} J_1^{\ell}, \ldots) S^{-1} \; . \tag{14.6.10}$$

Ga na dat hieruit volgt dat $Y(t)A = AY(t)$.

6. We definiëren bij een $A \in M_{n,n}$

$$\exp(tA) = \sum_{\ell=0}^{\infty} \frac{t^{\ell}}{!} A^{\ell} \; .$$

De reeks rechts convergeert (elementsgewijs) voor iedere t.

Want als

$$M_\ell := n \max_{i,j} |(A^\ell)_{ij}|$$

dan is (ga na)

$$M_\ell \le M_1 M_{\ell-1} \le M_1^\ell .$$

Met behulp van een standaard redenering volgt nu dat

$$\exp((t_1 + t_2)A) = \exp(t_1 A)\exp(t_2 A) .$$

Hieruit volgt met name dat exp(tA) regulier is en

$$(\exp(tA))^{-1} = \exp(-tA) .$$

Tenslotte geldt

$$\exp((t + h)A) - \exp(tA) = \exp(tA).hA. \sum_1^\infty \frac{h^{\ell-1}}{\ell!} A^{\ell-1}$$

waaruit volgt

$$\frac{d}{dt} \exp(tA) = \exp(tA).A = A.\exp(tA) .$$

Daar de fundamentele oplossing Y(t) van (14.6.1) eenduidig is en exp(0) = I, zien we

$$Y(t) = \exp(tA) .$$

7. Het is eenvoudig in te zien dat

$$\exp(t(A + B)) = \exp(tA)\exp(tB)$$

dan en slechts dan als A en B commuteren.
Daarmee vinden we (ga na)

$$\exp(t(\lambda I + J)) = e^{\lambda t} \sum_{\ell=0}^{h-1} \frac{t^\ell}{\ell!} J^\ell$$

en nu volgt eenvoudig dat, als A geschreven is in de vorm (14.6.9), het rechterlid van (14.6.10) gelijk is aan exp(tA).
Bewijs nu zelf dat

a) $\lim_{t\to\infty} \exp(tA) = 0 \Leftrightarrow \forall_k\colon \mathrm{Re}(\lambda_k) < 0$.
b) $\exp(tA)$ begrensd voor $t \to \infty \Leftrightarrow \forall_k\colon (\mathrm{Re}(\lambda_k) < 0$ of $\mathrm{Re}(\lambda_k) = 0$ en $h_k = 1)$.

Opgave. (14.6.11)

Beschouw de n-de orde differentiaalvergelijking

$$\frac{d^n z}{dt^n} + a_1 \frac{d^{n-1} z}{dt^{n-1}} + \ldots + a_n z = 0 .$$

Bewijs dat iedere oplossing $z(t)$ begrensd is voor $t \to \infty$ dan en slechts dan als voor alle nulpunten λ van het zgn. karakteristieke polynoom

$$\lambda^n + a_1 \lambda^{n-1} + \ldots + a_n = 0$$

geldt: $\mathrm{Re}(\lambda) < 0$ of $\mathrm{Re}(\lambda) = 0$ en λ enkelvoudig nulpunt.
(Hint: Maak gebruik van de bij het karakteristieke polynoom behorende companion-matrix, gedefinieerd in (13.4.7)).

Opgave. (14.6.12)

Het is mogelijk het limietgedrag van A^s en $\exp(tA)$ te onderzoeken op basis van de grofreductiestelling (13.3.2). Ga dit na.

15. Unitaire gelijkvormigheidstransformaties

We beschouwen in deze paragraaf matrices $A \in M_{n,n}(\mathbb{C})$ en vragen ons af in hoeverre we met een unitaire gelijkvormigheidstransformatie

$$B = U^{-1}AU = U^H AU, \quad U \text{ unitair} ,$$

A op eenvoudiger vorm kunnen brengen.
De afleidingen in deze paragraaf zullen grotendeels onafhankelijk zijn van de resultaten uit § 13 en § 14.

15.1. Lemma van Schur

Stelling (het zgn. lemma van Schur). (15.1.1)

Bij iedere $A \in M_{n,n}(\mathbb{C})$ is er een unitaire U zo dat

$$U^H AU = R ,$$

waarin R een rechter driehoeksmatrix is.

Bewijs.

Volledige inductie naar n. Triviaal als $n = 1$.
Zij $n > 1$. Daar A een matrix over $\mathbb{C}$ is heeft A volgens (12.2.7.2) minstens één eigenwaarde λ_1 met geometrische multipliciteit n_1.

Volgens (8.2.2.5) is er een unitaire matrix

$$U^{(1)} = (U_1 \mid U_2)$$

zo dat de kolommen van U_1 een orthonormale basis vormen voor $N(A - \lambda_1 I)$.

Uit $AU_1 = \lambda_1 U_1$, $U_1^H U_1 = I_{11}$, $U_2^H U_1 = 0$ volgt dat

$$U^{(1)H} A U^{(1)} = \left(\begin{array}{c|c} \lambda_1 I_{11} & A_{12}^{(1)} \\ \hline & A_{22}^{(1)} \end{array}\right),$$

met

$$I_{11} \in M_{n_1,n_1},\ A_{12}^{(1)} = U_1^H A U_2,\ A_{22}^{(1)} = U_2^H A U_2 .$$

Volgens de inductieveronderstelling is er nu een unitaire $U_{22}^{(2)}$ zo dat $U_{22}^{(2)H} A_{22}^{(1)} U_{22}^{(2)} = R_{22}$ (rechtsboven).

Met $U := U^{(1)} \operatorname{diag}(I_{11}, U_{22}^{(2)})$ geldt derhalve

$$U^H A U = \left(\begin{array}{c|c} \lambda_1 I_{11} & R_{12} \\ \hline & R_{22} \end{array}\right) = R ,$$

waarin $R_{12} = A_{12}^{(1)} U_{22}^{(2)}$. □

Een nadere analyse van de inductiestap in het bewijs van (15.1.1) met behulp van de begrippen index en multipliciteitenrij uit § 13 levert de volgende reductiestelling.

<u>Stelling.</u> (15.1.2)

Als A een eigenwaarde λ met index h en multipliciteitenrij $\Sigma = (\mu_1, \ldots, \mu_h)$ heeft, dan is er een unitaire U zo dat

$$U^H A U = \left(\begin{array}{c|c} \lambda I_{11} & B_{12} \\ \hline & B_{22} \end{array}\right)$$

met I_{11} de $\mu_1 \times \mu_1$ eenheidsmatrix. De matrix $(B_{12}^T \mid B_{22}^T - \lambda I_{22})^T$ is $\in$ RR. Als $h = 1$ dan is $B_{22} - \lambda I_{22}$ regulier, als $h > 1$ dan is λ eigenwaarde van B_{22} met index $h-1$ en multipliciteitenrij $(\mu_2 - \mu_1, \ldots, \mu_h - \mu_1)$.

Bewijs.

Stel voor het gemak $\lambda = 0$. Zij $U = (U_1 \mid U_2)$ unitair met $U_1 \in M_{n,\mu_1}$ zo dat $R(U_1) = N(A)$. Dan geldt

$$U^H AU = \left(\begin{array}{c|c} 0 & B_{12} \\ \hline & B_{22} \end{array}\right) =: (0 \mid B_2) , \tag{$*$}$$

waarin B_2 afmetingen $n \times (n-\mu_1)$ heeft. Daar $\dim(N(A)) = \mu_1$ moet $B_2 \in RR$ zijn.

Uit $(*)$ volgt

$$U^H A^{j+1} U = (0 \mid B_2 B_{22}^j) .$$

Daar $B_2 \in RR$ is $\dim(N(A^{j+1})) = \mu_1 + \dim(N(B_{22}^j))$ (ga na) en dus

$$\dim(N(B_{22}^j)) = \mu_{j+1} - \mu_1 .$$

Hieruit volgt de bewering. □

Opmerkingen. (15.1.3)

1. Constructie van U en B zou als volgt kunnen geschieden. Als $0 \in \sigma(A)$ dan is er volgens (8.2.6), toegepast op A^H, een links-unitaire U_2 en een rechts-reguliere C_2 zo dat $A^H = U_2 C_2^H$, of wel, $A = C_2 U_2^H$. Volgens (8.2.2.5) is er een U_1 zo dat $U = (U_1 \mid U_2)$ unitair is en hiervoor geldt $AU = (0 \mid C_2)$, dus $U^H AU = (0 \mid B_2)$ met $B_2 = U^H C_2 \in RR$.
   In de numerieke practijk bepaalt men liever niet eerst U_2 en C_2 met Gram-Schmidt en daarna U_1, maar rechtstreeks met behulp van zg. Householder-transformaties een unitaire U en een permutatiematrix P zo dat $P^T AU = (0 \mid C_2)$, waarin C_2 een trapeziumvorm heeft. Daarmee bepaalt men $B_2 = U^H PC_2$.
2. Met (15.1.2) ter uitvoering van de inductiestap kan door volledige inductie de volgende generalisatie van (14.4.1) bewezen worden (ga na):
   Als $\lambda \in \sigma(A)$ met index h en algebraische multipliciteit $m < n$ dan is er een unitaire U zo dat

$$U^H AU = \begin{pmatrix} \lambda I_{11} & C_{12} & \cdots & C_{1,h} & C_{1,h+1} \\ & \lambda I_{22} & \cdots & C_{2,h} & C_{2,h+1} \\ & & & & \\ & & & C_{h-1,h} & C_{h-1,h+1} \\ & & & \lambda I_{h,h} & C_{h,h+1} \\ & & & & C_{h+1,h+1} \end{pmatrix} \tag{$*$}$$

waarin de matrices $C_{j-1,j} \in RR$ (voor $1 < j \leq h$) en $C_{h+1,h+1} - \lambda I_{h+1,h+1}$ regulier is met afmetingen $(n-m) \times (n-m)$.

3. Zelfs is het mogelijk om te zorgen dat de blokken $C_{j-1,j}$ $(1 < j \leq h)$ rechtsboven matrices zijn: ga van $U^H AU$ over op $V^H U^H AUV$, waarin $V = \mathrm{diag}(V_1,\dots,V_h,V_{h+1})$ unitair is en $V_{h-1},\dots,V_1$ na keuze van V_{h+1} en V_h successievelijk bepaald worden uit de eis dat

$$C_{j-1,j}V_j = V_{j-1}R_{j-1,j}$$

met $R_{j-1,j}$ rechtsboven en V_{j-1} unitair.

4. Uit het rechterlid van (*) zijn de multipliciteitenrij, de nulruimten en de cyclische invariante deelruimten (zie (14.3.8.3)) behorend bij de eigenwaarde λ eenvoudig af te lezen (ga na). Bovendien blijkt door inductie naar het aantal eigenwaarden:
Als $A \in M_{n,n}(\mathbb{C})$ eigenwaarden $\lambda_1,\dots,\lambda_p$ met algebraische multipliciteiten $m_1,\dots,m_p$ heeft, dan geldt

$$\mathrm{tr}(A) = \sum_1^p m_j\lambda_j, \quad \|A\|^2 \geq \sum_1^p m_j|\lambda_j|^2 .$$

5. Daar unitaire transformaties numeriek stabiel zijn kan de matrix C ook numeriek met goede nauwkeurigheid bepaald worden (zie voor verdere numerieke details A. Ruhe, An algorithm for numerical determination of the structure of a general matrix, BIT 10 (1970), 196-210 en Golub-Wilkinson, SIAM Review 18 (1976), 578-619). De verdere, niet-unitaire, gelijkvormigheidstransformaties waarmee C behandeld moet worden (zie bv. (13.2.10) voor de grofreductie en (14.4.4) voor de fijnreductie) om op de Jordan normaalvorm te komen, zijn in het algemeen niet numeriek stabiel.

Opgave. (15.1.4)

Laat zien dat men mag eisen dat de matrix U uit opgave **(13.2.8) unitair is.**

Opgave. (15.1.5)

Als $A_1,\dots,A_m \in M_{n,n}(\mathbb{C})$ commuteren dan zijn er een unitaire U en driehoeksmatrices $R_1,\dots,R_m$ zo dat

$$U^H A_k U = R_k, \qquad k = 1,\dots,m$$

(vergelijk (12.3.14)). Hint: gebruik (15.1.2) en het feit dat, als $\lambda \in \sigma(A_1)$, $A_k(N(A_1 - \lambda I)) \subset N(A_1 - \lambda I)$ voor $k = 2,\dots,m$.

15.2. Unitaire diagonalisatie. Normale matrices

We vragen nu welke matrices A door een unitaire gelijkvormigheidstransformatie op diagonaalvorm gebracht kunnen worden. A heet dan unitair diagonaliseerbaar.

Als A unitair diagonaliseerbaar is:

$$U^H AU = \Lambda,\ U \text{ unitair},\ \Lambda \text{ diagonaal}\ ,$$

dan is

$$AA^H = U\Lambda U^H . U\Lambda^H U^H = U.\Lambda\Lambda^H.U^H = A^H A$$

omdat voor iedere diagonaalmatrix $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$ geldt dat

$$\Lambda\Lambda^H = \Lambda^H\Lambda = \mathrm{diag}(|\lambda_1|^2,\ldots,|\lambda_n|^2)\ .$$

De voorwaarde $AA^H = A^H A$ is dus nodig voor unitaire diagonaliseerbaarheid. We zullen zien dat hij ook voldoende is.

<u>Definitie.</u> (15.2.1)

$A \in M_{n,n}(\mathbb{C})$ heet normaal als

$$AA^H = A^H A\ .$$

<u>Stelling.</u> (15.2.2)

Als A normaal is dan geldt

. A^H is normaal,

. voor ieder polynoom φ over $\mathbb{C}$ is $\varphi(A)$ normaal,

. $U^H AU$ is normaal (U unitair).

<u>Stelling.</u> (15.2.3)

Een rechterdriehoeksmatrix is normaal dan en slechts dan als hij een diagonaalmatrix is.

<u>Bewijs.</u>

Dat de voorwaarde voldoende is, is duidelijk. Dat hij nodig is, bewijzen we met inductie naar n.

Het geval n = 1 is triviaal. Zij

$$R = \left(\begin{array}{c|c} R_{11} & R_{12} \\ \hline & R_{22} \end{array}\right)$$

een normale driehoeksmatrix met $R_{11} \in M_{1,1}$. Dan is het 1,1-blok van R^HR $R_{11}^HR_{11}$ en dat van RR^H is $R_{11}R_{11}^H + R_{12}R_{12}^H$. Daar R_{11} normaal is impliceert dit $R_{12}R_{12}^H = 0$, dus $R_{12} = 0$. Uit de gelijkheid van de 2,2-blokken van R^HR en RR^H volgt nu dat R_{22} normaal is, dus volgens de inductieveronderstelling een diagonaalmatrix. □

Opgave. (15.2.4)

Zij

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

met vierkante diagonaalblokken. Als A normaal is geldt: $\|A_{12}\| = \|A_{21}\|$.
Als $A_{21} = 0$ geldt: A is normaal dan en slechts dan als $A_{12} = 0$ en A_{11} en A_{22} normaal zijn.
Hint: bedenk dat $\|A_{12}\|^2 = tr(A_{12}^HA_{12}) = tr(A_{12}A_{12}^H)$.

Stelling. (15.2.5)

Een matrix $A \in M_{n,n}(\mathbb{C})$ is unitair diagonaliseerbaar dan en slechts dan als hij normaal is.

Bewijs.

Dat de voorwaarde nodig is, is in de inleiding van deze paragraaf al aangetoond.
Dat de voorwaarde voldoende is volgt uit (15.1.1) (lemma van Schur), (15.2.2) en (15.2.3). □

Opmerking. (15.2.6)

Zij $U^HAU = \Lambda$ = diagonaal. Zonder beperking der algemeenheid mogen we aannemen dat $\Lambda = diag(\lambda_1 I_{11},\ldots,\lambda_p I_{pp})$, waarin $\lambda_1,\ldots,\lambda_p$ de onderling verschillende eigenwaarden van A zijn. Als we U overeenkomstig partitioneren: $U = (U_1|\ldots|U_p)$, dan geldt

$$A = \sum_{1}^{p} \lambda_j U_j U_j^H = \sum_{1}^{p} \lambda_j E_j \ , \qquad (15.2.7)$$

waarin de E_j orthogonale projectoren zijn zo dat

$$E_iE_j = 0 \text{ voor } i \neq j,\ \Sigma E_j = I \ . \qquad (15.2.8)$$

Dit is de spectraalvoorstelling van A (vgl. § 12.3) waarin voor een normale matrix de projectoren dus orthogonaal zijn. Dit geldt ook omgekeerd: als in (15.2.7) en (15.2.8) de E_j orthogonale projectoren zijn (en de λ's willekeurig uit $\mathbb{C}$) dan is A normaal (ga na).

<u>Opgave.</u> (15.2.9)

Bewijs rechtstreeks dat als E projector en normaal is, E een orthogonale projector is.

<u>Opgave.</u> (15.2.10)

Als A normaal is dan is $N(A^2) = N(A)$. Bewijs dit en leid hieruit rechtstreeks af dat een normale matrix diagonaliseerbaar is.

<u>Opgave.</u> (15.2.11)

Als $A = \Sigma\lambda_j E_j$, $E_i E_j = \delta_{ij} E_j$, $\Sigma E_j = I$, dan geldt: A normaal $\Leftrightarrow$ alle E_j orthogonale projectoren.
Hint: gebruik (12.3.12).

<u>Opgave.</u> (15.2.12)

A is normaal dan en slechts dan als

i) $A = H + iK$ met H en K hermitisch, $HK = KH$.
ii) $A = UH$, met U unitair, H hermitisch en p.s.d., $UH = HU$.
iii) A^H is polynoom in A.
iv) Voor iedere $x \in \mathbb{C}^n$ is $\|Ax\| = \|A^H x\|$.
v) Voor iedere $\alpha \in \mathbb{C}$ is $N(A - \alpha I) = (R(A - \alpha I))^{\perp}$ (d.w.z. $A - \alpha I$ is range-hermitian).
vi) Voor iedere deelruimte $V \subset \mathbb{C}^n$ geldt: als V invariant is t.o.v. A dan is $V^{\perp}$ ook invariant t.o.v. A.

<u>Stelling.</u> (15.2.13)

Stel A heeft eigenwaarden $\lambda_1,\ldots,\lambda_p$, met algebraische multipliciteiten $m_1,\ldots,m_p$. Dan geldt

$$A \text{ normaal} \Leftrightarrow \|A\|^2 = \sum_1^n m_j |\lambda_j|^2 .$$

Bewijs.

Volgt uit (15.1.3.4). □

Definitie. (15.2.14)

A heet hermitisch indien $A^H = A$,
A heet anti-hermitisch indien $A^H = -A$,
A heet unitair als $A^HA = AA^H = I$.

Stelling. (15.2.15)

Hermitische, anti-hermitische en unitaire matrices zijn normaal en dus unitair diagonaliseerbaar:

$$U^HAU = \Lambda \ .$$

Het spectrum (de diagonaalelementen van Λ) heeft de volgende eigenschappen

1) Als A hermitisch is dan is het spectrum reëel ($\Lambda = \Lambda^H$).
2) Als A anti-hermitisch is dan is het spectrum zuiver imaginair ($\Lambda = -\Lambda^H$).
3) Als A unitair is dan ligt het spectrum op de eenheidscirkel ($\Lambda^H\Lambda = I$).

Opmerking. (15.2.16)

Het omgekeerde geldt natuurlijk ook: als A unitair diagonaliseerbaar is en een reëel spectrum heeft, dan is A hermitisch, enz.

Opgave. (15.2.17)

Zij

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

met $\alpha,\beta,\gamma,\delta \in \mathbb{C}$. Dan geldt

$$\text{A is normaal} \Leftrightarrow |\beta| = |\gamma| \wedge \beta(\bar{\alpha} - \bar{\delta}) = \bar{\gamma}(\alpha - \delta)$$

$$\Leftrightarrow (\beta = \gamma = 0) \vee (|\beta| = |\gamma| \neq 0 \wedge (\alpha - \delta)^2/(\beta\gamma) \geq 0)).$$

Zij nu A normaal. Definieer, als $|\beta| = |\gamma| \neq 0, \theta$, $(\beta\gamma)^{\frac{1}{2}}$ en φ door

$$(\beta\gamma)^{\frac{1}{2}} = \beta e^{-i\theta} = \gamma e^{i\theta}, \ -\pi/2 < \arg((\beta\gamma)^{\frac{1}{2}}) \leq \pi/2 \ ;$$

$$\text{ctn } 2\varphi = \tfrac{1}{2}(\alpha - \delta)/(\beta\gamma)^{\frac{1}{2}}, \ -\pi/4 < \varphi \leq \pi/4 \ .$$

Als $\beta = \gamma = 0$ dan is $\theta = \varphi = (\beta\gamma)^{\frac{1}{2}} = 0$.

Er geldt dan

$$AU = U \operatorname{diag}(\lambda_1,\lambda_2) ,$$

met

$$U = \begin{pmatrix} \cos\varphi & -e^{i\theta}\sin\varphi \\ e^{-i\theta}\sin\varphi & \cos\varphi \end{pmatrix},$$

$$\lambda_1 = \alpha + (\beta\gamma)^{\frac{1}{2}}\tan\varphi, \quad \lambda_2 = \delta - (\beta\gamma)^{\frac{1}{2}}\tan\varphi.$$

Welke vereenvoudiging treedt op als A hermitisch is?

15.3. Hermitische matrices

Stelling. (15.3.1)

Zij $A \in M_{n,n}$ hermitisch met rang r en signatuur s (zie (10.3.1)). Dan is de som der multipliciteiten van de positieve eigenwaarden $\frac{1}{2}(r + s)$ en die van de negatieve eigenwaarden $\frac{1}{2}(r - s)$.

Bewijs.

De eigenwaarden van A zijn reëel. We kunnen ze zo ordenen dat er een unitaire U is zo dat

$$A = U \operatorname{diag}(D_1^2,-D_2^2,0)U^H = V \operatorname{diag}(I_1,-I_2,0)V^H ,$$

waarin D_1 en D_2 diagonaalmatrices met positieve diagonaalelementen zijn en $V = U \operatorname{diag}(D_1,D_2,I_3)$. A is dus congruent met $\operatorname{diag}(I_1,-I_2,0)$. De bewering volgt nu uit (10.3.8.2).

Gevolg. (15.3.2)

Een hermitische A is positief definiet dan en slechts dan als z'n eigenwaarden positief zijn en positief semi-definiet dan en slechts dan als z'n eigenwaarden niet negatief zijn (bewijs dit ook rechtstreeks!).

Opgave. (15.3.3)

Zij $A \in M_{n,n}$ hermitisch en V een t.o.v. A invariante deelruimte van $\mathbb{C}^n$. Dan is er in V een orthonormale basis, bestaande uit eigenvectoren van A (n.b. dit volgt uit (12.3.13), maar kan hier wat eenvoudiger bewezen worden). Bewijs ook dat het orthogonaal complement van V ook invariant is t.o.v. A. Bewijs dezelfde uitspraken ook voor een normale A (gebruik (15.2.4)).

We behandelen nu de zgn. minimax eigenschappen van de eigenwaarden van een hermitische matrix en enkele daarmee samenhangende zaken.

Stelling. (15.3.4)

Zij $A = U\Lambda U^H \in M_{n,n}$ hermitisch met U unitair en $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$ met $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Dan geldt voor $1 \leq k \leq n$

$$\lambda_k = \max\{\min_{\substack{x\in V\\ x\neq 0}} \frac{x^H Ax}{x^H x} \mid V \subset \mathbb{C}^n,\ \dim(V) \geq k\} =$$

$$= \min\{\max_{\substack{x\in V\\ x\neq 0}} \frac{x^H Ax}{x^H x} \mid V \subset \mathbb{C}^n,\ \dim(V) \geq n-k+1\}\ .$$

Bewijs.

Zij $U = (u_1|\ldots|u_n)$. Dan is voor $x = \Sigma\xi_j u_j \neq 0$

$$\frac{x^H Ax}{x^H x} = \frac{\sum_1^n \lambda_j|\xi_j|^2}{\sum_1^n |\xi_j|^2}\ .$$

Daar $x \in \{u_1,\ldots,u_k\}$ equivalent is met $\xi_j = 0$ voor $k+1 \leq j \leq n$ vinden we eenvoudig dat

$$\lambda_k = \min_{\substack{x\in\{u_1,\ldots,u_k\}\\ x\neq 0}} \frac{x^H Ax}{x^H x}$$

en analoog

$$\lambda_k = \max_{\substack{x\in\{u_k,\ldots,u_n\}\\ x\neq 0}} \frac{x^H Ax}{x^H x}\ .$$

Stel nu dat $V \subset \mathbb{C}^n$ met $\dim(V) \geq k$. Uit (5.4.26) volgt dan dat $\dim(V \cap \{u_k,\ldots,u_n\}) \geq 1$ en dus is

$$\min_{\substack{x\in V\\ x\neq 0}} \frac{x^H Ax}{x^H x} \le \max_{\substack{x\in\{u_k,\dots,u_n\}\\ x\neq 0}} \frac{x^H Ax}{x^H x} = \lambda_k \ .$$

Voor $V = \{u_1,\dots,u_k\}$ geldt echter het gelijkteken en daarmee is de max-min eigenschap van λ_k bewezen. Analoog de min-max eigenschap. □

Stelling. (15.3.5)

Zij $A \in M_{n,n}$ hermitisch met eigenwaarden $\lambda_1 \ge \dots \ge \lambda_n$. Zij $U_1 \in M_{n,k}$ links-unitair. Zij $B := U_1^H AU_1$ met eigenwaarden $\mu_1 \ge \dots \ge \mu_k$. Dan geldt voor $1 \le j \le k$

$$\lambda_{n+1-j} \le \mu_{k+1-j},\ \mu_j \le \lambda_j \ .$$

Bewijs.

Stel dat bij de eigenwaarden $\mu_1,\dots,\mu_k$ van B orthonormale eigenvectoren $y_1,\dots,y_k$ horen. Dan geldt (ga na)

$$\mu_j = \min_{y\in\{y_1,\dots,y_j\}} \frac{y^H By}{y^H y} = \min_{x\in\{U_1y_1,\dots,U_1y_j\}} \frac{x^H Ax}{x^H x} \le \lambda_j \ ,$$

omdat de ruimte $\{U_1y_1,\dots,U_1y_j\}$ j-dimensionaal is.

Analoog

$$\mu_{k+1-j} = \max_{y\in\{y_{k+1-j},\dots,y_k\}} \frac{y^H By}{y^H y} = \max_{x\in\{U_1y_{k+1-j},\dots,U_1y_k\}} \frac{x^H Ax}{x^H x} \ge \lambda_{n+1-j} \ .$$

□

Gevolg. (15.3.6)

Door voor U_1 de matrix bestaande uit de eerste k kolommen van de eenheidsmatrix te nemen, vinden we het volgende resultaat.

Zij

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)$$

$\in M_{n,n}$ en hermitisch en $A_{11} \in M_{k,k}$. Als A eigenwaarden $\lambda_1 \ge \dots \ge \lambda_n$ en A_{11} eigenwaarden $\mu_1 \ge \dots \ge \mu_k$ heeft dan geldt voor $1 \le j \le k$

$$\lambda_{n+1-j} \le \mu_{k+1-j},\ \mu_j \le \lambda_j \ .$$

In het geval dat $k = n - 1$ volgt hieruit

$$\lambda_n \le \mu_{n-1} \le \lambda_{n-1} \le \ldots \le \lambda_2 \le \mu_1 \le \lambda_1$$

(de zgn. scheiding der eigenwaarden van A en A_{11}).

Opgave. (15.3.7)

Zij A, B en C hermitisch, $A = B + C$ met eigenwaarden resp. $\alpha_1 \ge \ldots \ge \alpha_n$, $\beta_1 \ge \ldots \ge \beta_n$ en $\gamma_1 \ge \ldots \ge \gamma_n$. Dan geldt voor $k \ge 0$, $\ell \ge 0$, $k + \ell \le n - 1$

$$\beta_{n-k} + \gamma_{n-\ell} \le \alpha_{n-k-\ell},\ \alpha_{1+k+\ell} \le \beta_{1+k} + \gamma_{1+\ell}\ .$$

Door $\ell = 0$ te nemen blijkt dat voor $j = 1,\ldots,n$

$$\gamma_n \le \alpha_j - \beta_j \le \gamma_1,\ |\alpha_j - \beta_j| \le \max_i |\gamma_i|\ .$$

Hint: voor V_1 en $V_2 \subset \mathbb{C}^n$ geldt

$$\max_{x\in V_1\cap V_2} \frac{x^HAx}{x^Hx} \le \max_{x\in V_1} \frac{x^HBx}{x^Hx} + \max_{x\in V_2} \frac{x^HCx}{x^Hx}$$

en $\dim(V_1 \cap V_2) \ge \dim(V_1) + \dim(V_2) - n$.

Opgave. (15.3.8)

Zij $A = B + \sigma cc^H$ met $B \in M_{n,n}$ hermitisch, $\sigma \in \mathbb{R}$, $c \in \mathbb{C}^n$ en $c^Hc = 1$.
Stel dat A en B eigenwaarden $\alpha_1 \ge \ldots \ge \alpha_n$ resp. $\beta_1 \ge \ldots \ge \beta_n$ hebben. Dan geldt voor $j = 1,\ldots,n$

$$\alpha_j = \beta_j + \sigma\theta_j$$

met $0 \le \theta_j \le 1$ en $\Sigma\theta_j = 1$.

Stelling. (15.3.9)

Zij A en $B \in M_{n,n}$ hermitisch met eigenwaarden $\alpha_1 \ge \ldots \ge \alpha_n$, resp. $\beta_1 \ge \ldots \ge \beta_n$. Dan is $\langle A,B\rangle$ reëel en

$$\sum_1^n \alpha_j\beta_{n+1-j} \le \langle A,B\rangle \le \sum_1^n \alpha_j\beta_j\ .$$

Bewijs.

Uit (8.5.2) volgt voor hermitische A en B

$$\overline{\langle A,B\rangle} = \langle B,A\rangle = \langle A^H,B^H\rangle = \langle A,B\rangle \ .$$

We bewijzen nu de rechterongelijkheid door inductie naar n. De linkerongelijkheid volgt analoog of door vervanging van B door -B en β_j door $-\beta_{n+1-j}$. Het geval n = 1 is triviaal. Zij n > 1. Zonder beperking der algemeenheid mogen we veronderstellen dat $\alpha_n = 0$. Immers, voor iedere $\alpha \in \mathbb{R}$ geldt

$$\langle A - \alpha I,B\rangle = \langle A,B\rangle - \alpha\langle I,B\rangle = \langle A,B\rangle - \alpha\,\mathrm{tr}(B)$$

en

$$\sum_1^n (\alpha_j - \alpha)\beta_j = \Sigma\alpha_j\beta_j - \alpha\Sigma\beta_j = \Sigma\alpha_j\beta_j - \alpha\,\mathrm{tr}(B) \ .$$

Stel dus $\alpha_n = 0$. Dan is er een unitaire U (waarvan de laatste kolom eigenvector van A is bij α_n) zodat

$$U^HAU = A' = \left(\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & 0 \end{array}\right), \ U^HBU=B'=\left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right)$$

met A_{11} en $B_{11} \in M_{n-1,n-1}$. Uit (8.5.2) volgt dat

$$\langle A,B\rangle = \langle A',B'\rangle = \langle A_{11},B_{11}\rangle \ .$$

Nu kunnen we de inductieveronderstelling toepassen. De eigenwaarden van A_{11} zijn $\alpha_1 \geq \dots \geq \alpha_{n-1}$ en die van B_{11} $\beta_1' \geq \dots \geq \beta_{n-1}'$, waarbij volgens (15.3.6) $\beta_j' \leq \beta_j$ $(j = 1,\dots,n-1)$. Derhalve is

$$\langle A,B\rangle = \langle A_{11},B_{11}\rangle \leq \sum_1^{n-1} \alpha_j\beta_j' \leq \sum_1^n \alpha_j\beta_j \ ,$$

omdat $\alpha_j \geq 0$ en $\alpha_n = 0$. □

Gevolg (stelling van Wielandt-Hoffman). (15.3.10)

Zij A en B als in (15.3.9). Dan geldt

$$\sum_1^n (\alpha_j - \beta_j)^2 \leq \|A - B\|^2 \ , \quad \|A + B\|^2 \leq \sum_1^n (\alpha_j + \beta_j)^2 \ .$$

Immers,

$$\|A \pm B\|^2 = \langle A \pm B,\ A \pm B\rangle = \|A\|^2 + \|B\|^2 \pm 2\langle A,B\rangle$$

en $\|A\|^2 = \Sigma\alpha_j^2$, $\|B\|^2 = \Sigma\beta_j^2$.

16. Unitaire equivalentie

16.1. Singuliere waarden decompositie

Stelling. (16.1.1)

Zij $A \in M^r_{m,n}(\mathbb{C})$. Dan zijn er
unitaire $U \in M_{m,m}$ en $V \in M_{n,n}$,
een diagonaalmatrix $\Sigma \in M_{m,n}$ van de vorm

$$\Sigma = \left(\begin{array}{c|c} \Sigma_1 & 0 \\ \hline 0 & 0 \end{array}\right) ,$$

waarin $\Sigma_1 = \mathrm{diag}(\sigma_1,\ldots,\sigma_r)$ met $\sigma_1 \geq \ldots \geq \sigma_r > 0$,
zo dat

$$A = U\Sigma V^H . \qquad (16.1.2)$$

Ook zijn er links-unitaire $U_1 \in M_{m,r}$ en $V_1 \in M_{n,r}$, zo dat

$$A = U_1\Sigma_1V_1^H . \qquad (16.1.3)$$

Bewijs.

A^HA is hermitisch, $n \times n$, PSD en heeft rang r (want $A^HAx = 0 \Leftrightarrow Ax = 0$). Dus heeft A^HA precies r positieve eigenwaarden $\sigma_1^2,\ldots,\sigma_r^2$ en een spectraalvoorstelling van de vorm

$$A^HA = V_1\Sigma_1^2V_1^H ,$$

waarin $\Sigma_1 = \mathrm{diag}(\sigma_1,\ldots,\sigma_r)$ en V_1 links-unitair is. Zonder beperking der algemeenheid veronderstellen we dat $\sigma_1 \geq \ldots \geq \sigma_r > 0$. Definieer nu $U_1 := AV_1\Sigma_1^{-1}$, dan is $U_1^HU_1 = \Sigma_1^{-1}V_1^HA^HAV_1\Sigma_1^{-1} = I_{11}$, dus U_1 is links-unitair. En uit $A^HA(I - V_1V_1^H) = 0$ volgt $A(I - V_1V_1^H) = 0$, dus $A = AV_1V_1^H = U_1\Sigma_1V_1^H$. Hiermee is (16.1.3) bewezen. (16.1.2) volgt daaruit door aanvulling van U_1 en V_1 tot unitaire matrices en van Σ_1 tot $m \times n$ matrix. □

Definitie. (16.1.4)

De getallen $\sigma_1,\ldots,\sigma_r$ heten singuliere waarden van A, (16.1.2), zowel als (16.1.3) worden singuliere waarden decompositie van A genoemd.
Soms is het handig om $\sigma_j := 0$ voor $r < j \le \max(m,n)$ te definiëren.

Gevolgen. (16.1.5)

1) $$A^HA = V\Sigma^H\Sigma V^H = V_1\Sigma_1^2V_1^H ,$$

$$AA^H = U\Sigma\Sigma^HU^H = U_1\Sigma_1^2U_1^H .$$

De diagonaalelementen van Σ_1^2 zijn dus de eigenwaarden $\neq 0$ zowel van A^HA als van AA^H (inclusief van multipliciteiten), de kolommen van V_1, resp. U_1 zijn de bijbehorende eigenvectoren; $\sigma_1,\ldots,\sigma_r$ zijn dus eenduidig bepaald, U_1 en V_1 niet geheel (ga na).

2) $$A = U_1\Sigma_1V_1^H , \quad A^H = V_1\Sigma_1U_1^H ,$$

$$A^+ = V_1\Sigma_1^{-1}U_1^H , \quad A^{+H} = U_1\Sigma_1^{-1}V_1^H .$$

Dit volgt uit (9.3.1.7). De singuliere waarden van A^H zijn dus ook $\sigma_1,\ldots,\sigma_r$, die van A^+ zijn $\sigma_r^{-1},\ldots,\sigma_1^{-1}$.

3) Zij $U_1 = (u_1|\ldots|u_r)$, $V_1 = (v_1|\ldots|v_r)$. Dan vormen $(u_1,\ldots,u_r)$ een orthonormale basis voor R(A) en $(v_1,\ldots,v_r)$ een orthonormale basis voor $R(A^H) = R(A^+)$. Er geldt

$$Av_j = \sigma_ju_j , \quad A^Hu_j = \sigma_jv_j , \quad A^+u_j = \sigma_j^{-1}v_j , \quad A^{+H}v_j = \sigma_j^{-1}u_j .$$

Vergelijk dit met (8.2.10) en (9.2.14.9).

4) $$\|A\|^2 = \sum_{j=1}^{r} \sigma_j^2 , \quad \|A^+\|^2 = \sum_{j=1}^{r} \sigma_j^{-2} .$$

Dit volgt uit (16.1.5.2) en (8.5.4.11).

5) Als A normaal is dan is $A = U\Lambda U^H$ met U unitair en $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$. Zij $\Sigma := \mathrm{diag}(|\lambda_1|,\ldots,|\lambda_n|)$ en $D = \mathrm{diag}(e^{i\theta_1},\ldots,e^{i\theta_n})$ zo dat $\Lambda = \Sigma D$. Dan is $A = U\Sigma V^H$, met $V := UD^H$, een singuliere waarden ontbinding van A. De singuliere waarden van een normale matrix zijn dus de absolute waarden van de eigenwaarden (met passende multipliciteiten).

Opgave. (16.1.6)

Bewijs dat $A \in M_{m,n}$ links-unitair is dan en slechts dan als $\sigma_1(A) = \ldots = \sigma_n(A) = 1$.

Opgave (16.1.7)

Zij $A = uv^H$ met u en $v \in \mathbb{C}^n$, $\|u\| = \|v\| = 1$. Bepaal de singuliere waarden en de eigenwaarden (met indices en multipliciteiten) van A.

We kunnen de singuliere waarden decompositie $A = U\Sigma V^H$ interpreteren als een equivalentierelatie (zie (3.2.1)) tussen A en Σ met unitaire factoren U en V^H: een zgn. unitaire equivalentie. Er geldt dan

Stelling. (16.1.8)

Bij iedere $A \in M_{m,n}(\mathbb{C})$ is er één Σ zo als in (16.1.1) waarmee A unitair equivalent is.

Stelling. (16.1.9)

Matrices A en A' zijn unitair equivalent dan en slechts dan als ze dezelfde afmetingen, rang en singuliere waarden hebben (vergelijk dit met (3.2.4)).

Opgave. (16.1.10)

Bewijs (9.5.3) door voor A z'n singuliere waarden decompositie te substitueren.

Ook voor singuliere waarden gelden minimax stellingen.

Stelling. (16.1.11)

Als $A \in M^r_{m,n}$ met singuliere waarden $\sigma_1 \geq \ldots \geq \sigma_r > 0$ en $\sigma_{r+1} := \ldots \sigma_n := 0$

dan geldt voor $1 \le k \le n$

$$\sigma_k = \max\{\min_{\substack{x \in V \\ x \neq 0}} \frac{\|Ax\|}{\|x\|} \mid V \subset \mathbb{C}^n,\ \dim(V) \ge k\} =$$

$$= \min\{\max_{\substack{x \in V \\ x \neq 0}} \frac{\|Ax\|}{\|x\|} \mid V \subset \mathbb{C}^n,\ \dim(V) \ge n-k+1\} \quad .$$

Bewijs.

Voor $x \in \mathbb{C}^n$ geldt

$$\frac{\|Ax\|^2}{\|x\|^2} = \frac{x^H A^H Ax}{x^H x} \quad .$$

Daar $A^H A$ eigenwaarden $\sigma_1^2 \ge \ldots \ge \sigma_r^2 > \sigma_{r+1}^2 = \ldots = \sigma_n^2 = 0$ heeft (zie (16.1.5.1)) volgen de beweringen direct uit (15.3.4). □

Gevolgen. (16.1.12)

1. $$\sigma_1(A) = \max_{\substack{x \in \mathbb{C}^n \\ x \neq 0}} \frac{\|Ax\|}{\|x\|} , \qquad \sigma_n(A) = \min_{\substack{x \in \mathbb{C}^n \\ x \neq 0}} \frac{\|Ax\|}{\|x\|} \quad .$$

2. Als $A \in M_{n,n}$ dan geldt voor iedere eigenwaarde λ van A

$$\sigma_n(A) \le |\lambda| \le \sigma_1(A) \quad .$$

Immers, er is een eigenvector $x \neq 0$ waarvoor $\|Ax\|/\|x\| = |\lambda|$.

3. Als A en $B \in M_{m,n}$ dan geldt (ga na)

$$|\sigma_k(A) - \sigma_k(B)| \le \sigma_1(A - B) , \qquad k = 1,\ldots,\min(m,n) \quad .$$

Opgave. (16.1.13)

Zij $E \in M^r_{m,m}$ een projector. Bewijs dat

$$\sigma_1(E) \geq \ldots \geq \sigma_r(E) \geq 1$$

en dat alle gelijktekens gelden dan en slechts dan als E een orthogonale projector is.

Opgave. (16.1.14)

Zij $A \in M_{n,n}$ regulier, met eigenwaarden $\lambda_1,\ldots,\lambda_n$, zo gerangschikt dat $0 < |\lambda_n| \leq \ldots \leq |\lambda_1|$. Uit (16.1.12.2) volgt dat $0 < \sigma_n \leq |\lambda_n| \leq |\lambda_1| \leq \sigma_1$. Met behulp van determinanten vinden we

$$\prod_1^n \sigma_j = (\det(A^H A))^{\frac{1}{2}} = |\det(A)| = \prod_1^n |\lambda_j| .$$

Leid hieruit af dat

$$1 \leq |\lambda_n|/\sigma_n \leq (\sigma_1/\sigma_n)^{(n-1)/n} , \quad (\sigma_n/\sigma_1)^{(n-1)/n} \leq |\lambda_1|/\sigma_1 \leq 1 ,$$

zodat $|\lambda_1|$ en $|\lambda_n|$ veel dichter bij elkaar kunnen liggen dan σ_1 en σ_n.

Verifieer dat de nieuwe ongelijkheden "best possible" zijn door de eigenwaarden en singuliere waarden van

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

te bepalen. Vergelijk dit met het resultaat (16.1.5.5) voor normale matrices.

Opgave. (16.1.15)

Als A en $B \in M_{m,n}$ dan geldt voor $k \geq 0$, $\ell \geq 0$, $k + \ell \leq \min(m,n) - 1$

$$\sigma_{1+k+\ell}(A + B) \leq \sigma_{1+k}(A) + \sigma_{1+\ell}(B) .$$

Als $A \in M_{m,p}$ en $B \in M_{p,n}$ dan geldt voor $k \geq 0$, $\ell \geq 0$, $k + \ell \leq p - 1$

$$\sigma_{1+k+\ell}(AB) \leq \sigma_{1+k}(A)\sigma_{1+\ell}(B) ,$$

$$\sigma_{p-k-\ell}(AB) \geq \sigma_{p-k}(A)\sigma_{p-\ell}(B) .$$

Opgave. (16.1.16)

Zij $A = (A_1 \mid A_2)$, $r_i := r(A_i)$ $(i = 1,2)$.
Dan geldt

$$\sigma_k(A) \geq \max(\sigma_k(A_1),\sigma_k(A_2)) ,$$

$$\sigma_{r_1+r_2-k}(A) \leq \min(\sigma_{r_1-k}(A_1),\sigma_{r_2-k}(A_2)) .$$

In het geval dat $r_2 = 1$ volgt hieruit

$$\sigma_{k+1}(A) \leq \sigma_k(A_1) \leq \sigma_k(A) .$$

Stelling. (16.1.17)

Zij $A \in M^r_{m,n}$ met singuliere waarden $\sigma_1 \geq \ldots \geq \sigma_r > 0$. Dan zijn de eigenwaarden van de hermitische matrix

$$C := \left(\begin{array}{c|c} 0 & A \\ \hline A^H & 0 \end{array}\right)$$

$\sigma_1,\ldots,\sigma_r,0,-\sigma_r,\ldots,-\sigma_r$, waarbij de eigenwaarde 0 de multipliciteit m+n-2r heeft.

Bewijs.

Zij

$$A = U\Sigma V^H = (U_1 \mid U_2)\left(\begin{array}{c|c} \Sigma_1 & \\ \hline & \end{array}\right)\left(\begin{array}{c} V_1^H \\ \hline V_2^H \end{array}\right)$$

met U en V unitair en $\Sigma_1 = \mathrm{diag}(\sigma_1,\ldots,\sigma_r)$. Zij

$$W := \begin{pmatrix} 2^{-\frac{1}{2}}U_1 & U_2 & 0 & -2^{-\frac{1}{2}}U_1 \\ 2^{-\frac{1}{2}}V_1 & 0 & V_2 & 2^{-\frac{1}{2}}V_1 \end{pmatrix} .$$

Dan is (ga na) W unitair en uit $A(V_1 \mid V_2) = (U_1\Sigma_1|0)$, $A^H(U_1 \mid U_2) = (V_1\Sigma_1|0)$ volgt

$$CW = W\,\mathrm{diag}(\Sigma_1,0,0,-\Sigma_1) ,$$

waarbij de nul-blokken afmetingen $(m - r) \times (m - r)$, resp. $(n - r) \times (n - r)$ hebben. Hiermee is de spectraaldecompositie van C gevonden. □

Gevolgen. (16.1.18)

Als A en B $\in M_{m,n}$ singuliere waarden $\sigma_j(A)$ en $\sigma_j(B)$ hebben dan volgt uit (16.1.17) met (15.3.9) en (15.3.10) dat

$$\Sigma\sigma_k(A)\sigma_k(B) \geq |\mathrm{Re}(\langle A,B\rangle)| ,$$

$$\Sigma(\sigma_k(A) - \sigma_k(B))^2 \leq \|A \pm B\|^2 \leq \Sigma(\sigma_k(A) + \sigma_k(B))^2 .$$

16.2. Benadering van een matrix door een matrix met lagere rang

Beschouw het volgende probleem. Zij $A \in M^r_{m,n}$ en zij $0 \leq k \leq r$. Gevraagd wordt een matrix $B \in M_{m,n}$ met $r(B) = k$ zo dat $\|A - B\|$ minimaal is.
Met behulp van (16.1.18) is het minimum eenvoudig te bepalen.

<u>Stelling</u> (Eckart-Young). (16.2.1)

Zij $A \in M^r_{m,n}$ met singuliere waarden $\sigma_1 \geq \ldots \geq \sigma_r > 0$. Zij $0 \leq k \leq r$. Dan is

$$\min\{\|A - B\|^2 \mid B \in M^k_{m,n}\} = \sum_{k+1}^{r} \sigma_j^2 .$$

B minimaliseert dan en slechts dan als er links-unitaire $U_1 = (u_1|\ldots|u_r)$ en $V_1 = (v_1|\ldots|v_r)$ zijn zo dat

$$A = \sum_1^r \sigma_j u_j v_j^H, \quad B = \sum_1^k \sigma_j u_j v_j^H .$$

<u>Bewijs</u>.

Uit (16.1.18) volgt dat voor iedere $B \in M^k_{m,n}$ geldt

$$\|A - B\|^2 \geq \sum_1^{\min(m,n)} (\sigma_j(A) - \sigma_j(B))^2 \geq \sum_{k+1}^{r} \sigma_j^2 .$$

En deze ondergrens is ook bereikbaar want als A de singuliere waarden decompositie

$$A = U_1 \Sigma_1 V_1^H = \sum_1^r \sigma_j u_j v_j^H$$

heeft dan geldt voor

$$B := \sum_1^k \sigma_j u_j v_j^H$$

dat

$$\|A - B\|^2 = \sum_{k+1}^{r} \sigma_j^2 .$$

Hiermee is de bewering over het minimum bewezen.
Stel nu dat B minimaliseert. Zij

$$B = \tilde{U}\tilde{\Sigma}\tilde{V}^H = (\tilde{U}_1 \mid \tilde{U}_2)\left(\begin{array}{c|c} \tilde{\Sigma}_1 & \\ \hline & \end{array}\right)\begin{pmatrix} \tilde{V}_1^H \\ \\ \tilde{V}_2^H \end{pmatrix}$$

een singuliere waarden decompositie van B en zij

$$\tilde{U}^H A \tilde{V} = \left(\begin{array}{c|c} Z_{11} & Z_{12} \\ \hline Z_{21} & Z_{22} \end{array}\right) .$$

Dan is (ga na)

$$\| A - B \|^2 = \| \tilde{\Sigma}_1 - Z_{11} \|^2 + \| Z_{12} \|^2 + \| Z_{21} \|^2 + \| Z_{22} \|^2 .$$

Als $Z_{11} \neq \tilde{\Sigma}_1$ dan zou

$$B' := \tilde{U} \left(\begin{array}{c|c} Z_{11} & \\ \hline & \end{array}\right) \tilde{V}^H$$

een matrix met rang $\leq k$ zijn waarvoor $\| A - B' \| < \| A - B \|$. Dus moet $Z_{11} = \tilde{\Sigma}_1$. Analoog bewijst men dat $Z_{12} = 0$, $Z_{21} = 0$.
Zij nu $Z_{22} = U_{22}\tilde{\Sigma}_2 V_{22}^H$ een singuliere waarden decompositie van Z_{22} met unitaire U_{22} en V_{22}. Zij $U := \tilde{U}\,\mathrm{diag}(I_{11}, U_{22})$, $V = \tilde{V}\,\mathrm{diag}(I_{11}, V_{22})$.
Dan geldt

$$A = U\,\mathrm{diag}(\tilde{\Sigma}_1, \tilde{\Sigma}_2)V^H, \quad B = U\,\mathrm{diag}(\tilde{\Sigma}_1, 0)V^H .$$

Laat $\tilde{\sigma}_1 \geq \ldots \geq \tilde{\sigma}_k$, resp. $\tilde{\sigma}_{k+1} \geq \ldots \geq \tilde{\sigma}_{\min(m,n)}$ de diagonaalelementen van $\tilde{\Sigma}_1$ en $\tilde{\Sigma}_2$ zijn. Dan is de rij $\{\tilde{\sigma}_1, \ldots, \tilde{\sigma}_{\min(m,n)}\}$ een permutatie van de rij $\{\sigma_1, \ldots, \sigma_r, 0, \ldots, 0\}$. En daar

$$\sum_{k+1}^{\min(m,n)} \tilde{\sigma}_j^2 = \| \tilde{\Sigma}_2 \|^2 = \| A - B \|^2 = \sum_{k+1}^{r} \sigma_j^2$$

moet (ga na) $\tilde{\sigma}_j = \sigma_j$ voor $1 \leq j \leq r$ en $\tilde{\sigma}_j = 0$ voor $j > r$ zijn. De bewering over de minimaliserende matrices B volgt hier direct uit. $\square$

Opgave. (16.2.2)

De minimaliserende matrix B in (16.2.1) is dan en slechts dan eenduidig bepaald als $\sigma_{k+1} < \sigma_k$.

Opgave. (16.2.3)

Als $A \in M^r_{m,n}$, $B \in M^k_{m,n}$ met $0 \le k \le r$ en

$$\|A - B\| = \min\{\|A - B'\| \mid B' \in M^k_{m,n}\} ,$$

dan geldt $B^H(A - B) = 0$, $(A - B)B^H = 0$, of wel

$$BB^+A = B = AB^+B , \qquad (*)$$

zodat de kolommen van B de orthogonale projecties zijn van de kolommen van A op R(B) en analoog voor de rijen. Bewijs dit zonder singuliere waarden decompositie te gebruiken (gebruik (8.5.4.6)!).
Bewijs ook dat als, omgekeerd, $A \in M^r_{m,n}$ en $B \in M^k_{m,n}$ aan (*) voldoen, er een permutatie p van $\{1,\ldots,r\}$ is zo dat

$$\|A - B\|^2 = \sum_{k+1}^{r} \sigma^2_{p(j)}(A) .$$

Opgave. (16.2.4)

Zij $A \in M_{mn}$. Bepaal $Q \in M_{m,n}$, zo dat

Q is links-unitair, $\|A - Q\|$ is minimaal.

Bewijs dat voor de optimale Q geldt $A = QZ$ met $Z \in PSD$ en dat

$$\|A - Q\| \le \|A^HA - I\| .$$

Hint: Beschouw eerst het geval dat $A = \left(\frac{\Sigma_1}{0}\right)$ met $\Sigma_1 = \mathrm{diag}(\sigma_1,\ldots,\sigma_n)$.

Opgave. (16.2.5)

Zij Q en W links-unitair en van de vorm

$$Q = \left(\frac{Q_1}{Q_2}\right) , \quad W = \left(\frac{W_1}{0}\right) ,$$

waarin Q_1 en W_1 vierkant zijn. Bewijs dat er bij gegeven Q een W van de gegeven vorm is zo dat

$$\|Q_1 - W_1\| \le \|Q_2\|^2 , \quad \|Q - W\| \le \|Q_2\|(1 + \|Q_2\|^2)^{\frac{1}{2}} .$$

6.3. De operator-norm (of 2-norm) van een matrix

Uit (16.1.12.1) :

$$\sigma_1(A) = \max_{x \in \mathbb{C}^n, x \neq 0} \frac{\|Ax\|}{\|x\|} \tag{16.3.1}$$

volgt dat de afbeelding $\sigma_1 : M_{m,n}(\mathbb{C}) \to \mathbb{R}^+$ de volgende eigenschappen heeft (ga na):

i) $\sigma_1(A) \ge 0,\ \sigma_1(A) = 0 \Leftrightarrow A = 0$,

ii) $\sigma_1(\alpha A) = |\alpha|\sigma_1(A)$,

iii) $\sigma_1(A + B) \le \sigma_1(A) + \sigma_1(B)$,

iv) $\sigma_1(AB) \le \sigma_1(A)\sigma_1(B)$,

v) $\sigma_1(A) = \min\{\mu \mid \forall_x : \|Ax\| \le \mu\|x\|\}$.

Op grond van de eerste vier eigenschappen is $\sigma_1(\ )$ een submultiplicatieve matrix-norm. Op grond van de vijfde eigenschap wordt $\sigma_1(A)$ de operatornorm of 2-norm van A genoemd, horende bij de Euclidische of 2-norm $\|x\|$ voor vectoren x uit $\mathbb{C}^n$. We schrijven voortaan

$$\|A\|_2 := \sigma_1(A) . \tag{16.3.2}$$

N.B. De in (8.5) ingevoerde norm $\|A\|$ wordt wel de Frobenius- of Schurnorm van A genoemd en geschreven als $\|A\|_F$.

Eigenschappen (16.3.3)

1. $\|A\|_2 = \max\{\lambda^{\frac{1}{2}} | \lambda \in \sigma(A^H A)\} = \max\{\lambda^{\frac{1}{2}} | \lambda \in \sigma(AA^H)\}$.

   Als A normaal is dan is $\|A\|_2 = \max\{|\lambda| \,|\, \lambda \in \sigma(A)\}$.

2. $\|A\|_2 = \|A^H\|_2$.

3. Als $A \in M^r_{m,n}$ dan is $\|A^+\|_2 = 1/\sigma_r(A)$.

4. Als $A = \operatorname{diag}(A_{11},\dots,A_{kk})$ dan is

   $$\|A\|_2 = \max \|A_{jj}\|_2 \ .$$

   Dit volgt eenvoudig uit (16.3.1).

5. Als U en V links-unitair zijn dan is

   $$\|UAV^H\|_2 = \|A\|_2 \ .$$

6. Uit (16.1.5.4) volgt dat voor $A \in M^r_{m,n}$

   $$r^{-\frac{1}{2}}\|A\| \le \|A\|_2 \le \|A\| \le r^{\frac{1}{2}}\|A\|_2 \ .$$

7. Er geldt

   $$\|AB\| \le \|A\|_2\|B\| \ , \quad \|AB\| \le \|A\|\|B\|_2 \ .$$

   De eerste bewering volgt uit (8.5.4.12) door $B = (b_1 \mid \dots \mid b_n)$ te schrijven.

8. Uit (16.1.12.3) volgt dat, als A en $B \in M_{m,n}$,

   $$|\sigma_k(B) - \sigma_k(A)| \le \|B - A\|_2 \ , \qquad k = 1,\dots,\min(m,n) \ .$$

9. Als $A \in M^r_{m,n}$, $B \in M_{m,n}$ en $\|B - A\|_2 < \sigma_r(A)$ dan volgt uit (16.3.3.8) dat $\sigma_r(B) \ge \sigma_r(A) - \|B - A\|_2 > 0$ en dus $r(B) \ge r(A)$. Dit bewijst de semi-continuïteit van de rang; in een voldoend kleine omgeving van een matrix met rang r ligt geen matrix met lagere rang.

Opgave. (16.3.7)

Zij $A \in M_{m,n}$, $B \in M_{k,n}$. Als $N(B) \subset N(A)$ dan geldt

$$\max_{\substack{x \in \mathbb{C}^n \\ Bx \neq 0}} \frac{\|Ax\|}{\|Bx\|} = \|AB^+\|_2 .$$

Opgave. (16.3.8)

Als $E \neq 0$ projector is dan is $\|E\|_2 \geq 1$; het gelijkteken geldt dan en slechts dan als E orthogonale projector is (vgl. (8.4.4) en (16.1.13)).

Opgave. (16.3.9)

(16.3.3.9) kan als volgt verscherpt worden. Als A en $B \in M_{m,n}$ en $\|A^+(B-A)\|_2 < 1$ of $\|(B-A)A^+\|_2 < 1$ dan is $r(B) \geq r(A)$.
Hint : ga uit van $AA^+B = A(I + A^+(B-A))$.

Opgave. (16.3.10)

Als E_1 en E_2 projectoren zijn dan geldt

$$\|E_1 - E_2\|_2 < 1 \Rightarrow r(E_1) = r(E_2) .$$

Bewijs dat zowel met (16.1.13) en (16.3.3.8) als door op te maken dat $r(E_1) > r(E_2) \Rightarrow R(E_1) \cap N(E_2) \neq \{\underline{0}\}$.

Opgave. (16.3.11)

Zij $A \in M^r_{m,n}$, $B \in M_{m,n}$, $\|B - A\|_2 < \sigma_r(A)$. Dan geldt
- als $r(B) = r(A)$ dan is

$$\frac{\|A^+\|_2}{1 + \|A^+\|_2\|B-A\|_2} \leq \|B^+\|_2 \leq \frac{\|A^+\|_2}{1 - \|A^+\|_2\|B-A\|_2}$$

- als $r(B) > r(A)$ dan is

$$\|B^+\|_2 \geq 1/\|B - A\|_2 .$$

Opgave. (16.3.12)

Zij A en B $\in M_{m,n}$. Bewijs dat

$$B^+ - A^+ = -B^+(B-A)A^+ + B^+B^{+H}(B-A)^H(I-AA^+) + (I-B^+B)(B-A)^HA^{+H}A^+$$
$$= -A^+(B-A)B^+ + A^+A^{+H}(B-A)^H(I-BB^+) + (I-A^+A)(B-A)^HB^{+H}B^+ \ .$$

Hint: gebruik (9.2.10) t/m (9.2.13).
Geef ook uitdrukkingen voor $BB^+ - AA^+$ en $B^+B - A^+A$. In het geval dat $r(B) = r(A)$ volgen uit deze uitdrukkingen met (16.3.11) schattingen voor $\|B^+ - A^+\|_2$, etc., in termen van $\|B - A\|_2$ en $\|A^+\|_2$.

Opgave. (16.3.13)

Zij $A \in M^r_{m,n}$. Dan geldt

$$\min\{\|A - B\|_2 \mid B \in M^k_{m,n}\} = \sigma_{k+1}(A)$$

en het minimum wordt bereikt voor de zelfde B die $\|A - B\|$ minimaliseren volgens (16.3.1).

Stelling. (16.3.14)

Zij

$$W = \left(\begin{array}{c|c} W_{11} & W_{12} \\ \hline W_{21} & W_{22} \end{array}\right)$$

unitair en $W_{11} \in M_{k,k}$. Dan geldt

$$\|W_{21}\|_2 = \|W_{12}\|_2 = (1 - \sigma_k^2(W_{11}))^{\frac{1}{2}} \ .$$

Bewijs

Uit $W_{11}^HW_{11} + W_{21}^HW_{21} = I_{11}$ volgt met (16.1.12.1)

$$\|W_{21}\|_2^2 = \max_{x_1\in\mathbb{C}^k, x_1\neq 0} \frac{\|W_{21}x_1\|^2}{\|x_1\|^2} = 1 - \min_{x_1\in\mathbb{C}^k, x_1\neq 0} \frac{\|W_{11}x_1\|^2}{\|x_1\|^2} = 1 - \sigma_k^2(W_{11}) .$$

Analoog volgt uit $W_{11}W_{11}^H + W_{12}W_{12}^H = I_{11}$ dat

$$\|W_{12}\|_2^2 = \|W_{12}^H\|_2 = 1 - \sigma_k^2(W_{11}^H) = 1 - \sigma_k^2(W_{11}) . \qquad \square$$

Stelling. (16.3.15)

Zij E_1 en $E_2 \in M_{m,m}$ orthogonale projectoren met dezelfde rang. Dan geldt

$$\|(I - E_1)E_2\|_2 = \|E_1(I - E_2)\|_2 = \|E_1 - E_2\|_2 .$$

Bewijs.

Zij $E_1 = UJU^H$, $E_2 = VJV^H$ met U en V unitair en J de normaalvorm uit $M_{m,m}$ met $r(J) = r(E_1) = r(E_2) =: r$ (vgl. (8.3.4)).

Zij

$$U^HV = W = \left(\begin{array}{c|c} W_{11} & W_{12} \\ \hline W_{21} & W_{22} \end{array}\right)$$

met $W_{11} \in M_{r,r}$. Dan is W unitair en

$$(I - E_1)E_2 = U(I - J)WJV^H = U\left(\begin{array}{c|c} & \\ \hline W_{21} & \end{array}\right)V^H ,$$

$$E_1(I - E_2) = UJW(I - J)V^H = U\left(\begin{array}{c|c} & W_{12} \\ \hline & \end{array}\right)V^H ,$$

$$E_1 - E_2 = U(JW - WJ)V^H = U\left(\begin{array}{c|c} & W_{12} \\ \hline -W_{21} & \end{array}\right)V^H .$$

Uit (16.3.3.5) volgt dat $\|(I - E_1)E_2\|_2 = \|W_{21}\|_2$ en $\|E_1(I - E_2)\|_2 = \|W_{12}\|_2$. En met (16.3.3.4) zien we eenvoudig dat $\|E_1 - E_2\|_2 = \max(\|W_{12}\|_2, \|W_{21}\|_2)$. De bewering van de stelling volgt nu uit (16.3.14).

Opmerkingen. (16.3.16)

1. Als E_1 en E_2 rangen r_1 en r_2 hebben (eventueel verschillend) dan volgt door een kleine aanpassing van bovenstaand bewijs dat

$$\|E_1 - E_2\|_2 = \max(\|(I - E_1)E_2\|_2 \ , \ \|E_1(I - E_2)\|_2) \ .$$

2. Ongeacht de rangen van E_1 en E_2 geldt dat $\|(I - E_1)E_2\| \le 1 \; \|E_1(I - E_2)\|_2 \le 1$ en dus $\|E_1 - E_2\| \le 1$. De laatste bewering volgt overigens ook al uit het feit dat voor iedere x geldt (ga na) $-x^H x \le -x^H E_2 x \le x^H(E_1 - E_2)x \le x^H E_1 x \le x^H x$, zodat de eigenwaarden van $E_1 - E_2$ in $[-1,1]$ liggen.

3. Als $r(E_2) > r(E_1)$ dan is $R(E_2) \cap N(E_1) \ne \{\underline{0}\}$ (waarom?) en daaruit volgt dat er een x is *zo dat* $E_2(I - E_1)x = (E_2 - E_1)x = E_2 x = x$, zodat nu $\|(I - E_1)E_2\|_2 = \|E_2(I - E_1)\|_2 = \|E_1 - E_2\|_2 = 1$. De bewering uit (16.3.16.1) is dus weinig interessant als $r(E_1) \ne r(E_2)$ en zwakker dan de inhoud van (16.3.15) als $r(E_1) = r(E_2)$.

16.4. De hoek tussen twee deelruimten

Zij V_1 en V_2 deelruimten van $\mathbb{C}^m$ met dimensies n_1, resp. n_2. Zij E_1 en E_2 de orthogonale projectoren op V_1, resp. V_2.

Definitie. (16.4.1)

De hoek $\theta_{21} = \theta(V_2, V_1)$ van V_2 met V_1 is bepaald door

$$\sin \theta_{21} = \max_{\substack{y \in V_2 \\ y \ne 0}} \frac{\|(I - E_1)y\|}{\|y\|} \ , \qquad 0 \le \theta_{21} \le \pi/2 \ .$$

Eenvoudige eigenschappen (16.4.2)

1. In plaats van $\| (I - E_1)y \|$ kunnen we volgens (8.4.5) ook schrijven

$$\min_{z \in V_1} \| y - z \| ,$$

zodat ook

$$\sin \theta_{21} = \max_{y \in V_2} \min_{z \in V_1} \frac{\| y - z \|}{\| y \|} .$$

2. $\theta_{21} = 0 \Leftrightarrow \forall_{y \in V_2} : E_1 y = y \Leftrightarrow V_2 \subset V_1$.

3. $\theta_{21} = \pi/2 \Leftrightarrow \exists_{y \in V_2} : y \neq 0 \wedge E_1 y = 0 \Leftrightarrow V_2 \cap V_1^{\perp} \neq \{\underline{0}\}$.

4. $\dim(V_2) > \dim(V_1) \Rightarrow \theta_{21} = \pi/2$. Immers, in dit geval is $\dim(V_2) + \dim(V_1^{\perp}) > n$ en dus $V_2 \cap V_1^{\perp} \neq \{\underline{0}\}$.

5. Daar voor iedere y geldt $\| (I - E_1)y \|^2 + \| E_1 y \|^2 = \| y \|^2$, is

$$\cos \theta_{21} = \min_{y \in V_2} \frac{\| E_1 y \|}{\| y \|}$$

en, als $\theta_{21} < \pi/2$,

$$\tan \theta_{21} = \max_{y \in V_2} \frac{\| (I - E_1)y \|}{\| E_1 y \|} .$$

Stelling. (16.4.3)

Zij V_1, V_2, E_1 en E_2 als in de aanhef van 16.4.
Dan geldt

$$\sin \theta_{21} = \| (I - E_1)E_2 \|_2 , \quad \sin \theta_{12} = \| E_1 (I - E_2) \|_2 .$$

Als $\dim(V_1) = \dim(V_2)$ dan geldt

$$\theta(V_2,V_1) = \theta(V_1,V_2) = \theta$$

en

$$\sin\theta = \|(I - E_1)E_2\|_2 = \|E_1(I - E_2)\|_2 = \|E_1 - E_2\|_2 .$$

Bewijs.

Voor $y \in V_2$ geldt dat $y = E_2 y$ en dus

$$\|(I - E_1)y\| = \|(I - E_1)E_2 y\| \le \|(I - E_1)E_2\|_2 \|y\| \quad ,$$

zodat $\sin\theta_{21} \le \|(I - E_1)E_2\|_2$.
Anderzijds geldt voor $x \in \mathbb{C}^m$ dat $E_2 x \in V_2$ en dus met (16.4.1) en (16.3.3.11).

$$\|(I - E_1)E_2 x\| \le \sin\theta_{21}\|E_2 x\| \le \sin\theta_{21}\|x\| \quad ,$$

zodat $\|(I - E_1)E_2\|_2 \le \sin\theta_{21}$. Derhalve is $\sin\theta_{21} = \|(I - E_1)E_2\|_2$.
Door verwisseling van de indices 1 en 2 vinden we
$\sin\theta_{12} = \|(I - E_2)E_1\|_2 = \|E_1(I - E_2)\|_2$(dank zij het feit dat $I - E_2$ en E_1 hermitisch zijn).
De beweringen voor het geval dat $\dim(V_1) = \dim(V_2)$ volgen nu rechtstreeks uit (16.3.15). □

We bekijken nu nog enkele situaties waarbij V_1 en V_2 op een speciale manier gegeven zijn.

Stelling. (16.4.4)

Zij $U = (U_1 \mid U_2)$ en $U' = (U_1' \mid U_2') \in M_{m,m}$ unitair en $R(U_1) = V_1 := R(M)$, $V_2 := R(U_1')$ met dimensies n_1, resp. n_2.
Dan geldt

$$\sin\theta_{21} = \|U_2^H U_1'\|_2 \;, \quad \sin\theta_{12} = \|U_1^H U_2'\|_2 \;;$$

als $n_2 \le n_1$ dan is

$$\cos\theta_{21} = \sigma_{n_2}(U_1^H U_1') \;,$$

als $n_1 \le n_2$ dan is

$$\cos\theta_{12} = \sigma_{n_1}(U_1^H U_1') \;.$$

Bewijs.

Volgt uit (16.4.1) en (16.4.2.5) omdat $E_1 = U_1U_1^H$, $I - E_1 = U_2U_2^H$ en
$y \in V_2 \quad y = U_1'x$. □

Opmerking. (16.4.5)

Zij $n: = \min(n_1,n_2)$. Zonder beperking der algemeenheid mogen we veronderstellen (ga na) dat U en U' zo gekozen zijn dat

$$U_1^HU_1' = \mathrm{diag}(\cos\varphi_1,\ldots,\cos\varphi_n)$$

(afmetingen $n_1 \times n_2$) met

$$0 \le \varphi_1 \le \varphi_2 \le \quad \varphi_n \le \pi/2 \ .$$

Er geldt dan

$$\varphi_n = \min(\theta_{21},\theta_{12}) \ .$$

Als $U_1 = (u_1|\ldots|u_{n_1})$, $U_1' = (u_1'|\ldots|u_{n_2}')$ dan is

$$u_i^Hu_j = \delta_{ij} \ , \quad i,j = 1,\ldots,n_1 \ ,$$

$$u_i'^Hu_j = \delta_{ij} \ , \quad i,j = 1,\ldots,n_2$$

$$u_i'^Hu_j' = \delta_{ij}\cos\varphi_i, \ i = 1,\ldots,n_1, \ j = 1,\ldots,n_2$$

Op grond hiervan heten $\{u_1,\ldots,u_{n_1}\}$en $\{u_1',\ldots,u_{n_2}'\}$ kanonieke bases van V_1, resp. V_2 en $\varphi_1,\ldots,\varphi_n$ de kanonieke hoeken tussen V_1 en V_2. Merk op dat u_i loodrecht staat op u_j' voor $j \neq i$ en een hoek op φ_i vormt met u_i' (als $i \le n$).
Merk ook op dat

$$\dim(V_1 \cap V_2) = k \Leftrightarrow 0 = \varphi_1 = \ldots = \varphi_k < \varphi_{k+1} \ .$$

Stelling. (16.4.6)

Zij $U = (U_1 \mid U_2)$ unitair met $U_1 \in M_{m,n}$. Zij $V_1 = R(U_1)$, $V_2 = R(U_1W_{11} + U_2W_{21}) = R(UW_1)$ met zekere

$$W_1 = \left(\frac{W_{11}}{W_{21}}\right) \in M_{m,k} \ .$$

Dan geldt: $\theta_{21} < \pi/2 \Leftrightarrow N(W_{11}) \subset N(W_{21}) \Leftrightarrow W_{21} = T_{21}W_{11}$ met $T_{21} = W_{21}W_{11}^+$ en zo ja dan is $\tan\theta_{21} = \|T_{21}\|_2$;

Als $W_1 \in M_{m,n}^n$ dan is $\dim(V_2) = \dim(V_1)$ en nu geldt: $\theta_{21} (= \theta) < \pi/2 \Leftrightarrow W_{11} \in I$ en zo ja dan is

$$V_2 = R(U_1 + U_2T_{21}), \quad \tan\theta = \|T_{21}\|_2$$

met $T_{21} = W_{21}W_{11}^{-1}$. □

Bewijs.

$\theta_{21} < \pi/2 \Leftrightarrow V_2 \cap V_1^{\perp} = \{\underline{0}\} \Leftrightarrow \forall_x : (W_{11}x = 0 \Rightarrow W_{21}x = 0) \Leftrightarrow N(W_{11}) \subset N(W_{21}) \Leftrightarrow$
$W_{21} = W_{21}W_{11}^{+}W_{11} = T_{21}W_{11}$.
Uit (16.4.2.5) volgt (ga na) dat, als $\theta_{21} < \pi/2$,

$$\tan \theta_{21} = \max_{W_{11}x \neq 0} \frac{\| W_{21}x \|}{\| W_{11}x \|} = \max_{W_{11}x \neq 0} \frac{\| T_{21}W_{11}x \|}{\| W_{11}x \|} .$$

Hieruit volgt direct dat $\tan \theta_{21} \leq \| T_{21} \|_2$. Maar daar voor iedere y geldt $\| T_{21}y \| = \| W_{21}W_{11}^{+}y \| \leq \tan \theta_{21} \| W_{11}W_{11}^{+}y \| \leq \tan \theta_{21} \| y \|$, geldt ook $\| T_{21} \|_2 \leq \tan \theta_{21}$.

Als $W_1 \in RR$ dan is $N(W_{11}) \cap N(W_{21}) = \{\underline{0}\}$ en dus $N(W_{11}) \subset N(W_{21}) \Leftrightarrow N(W_{11}) = \{\underline{0}\} \Leftrightarrow W_{11} \in RR$. De rest is nu duidelijk. □

Opgave. (16.4.7)

Zij X_1 en $X_2 \in M^n_{m,n}$, $V_i := R(X_{ii})$, $i = 1,2$ en $\theta := \theta(V_2,V_1)$. Dan geldt

$$\sin \theta = \| (I - X_1X_1^{+})(X_2 - X_1)X_2^{+} \|_2 .$$

Voorts geldt

$$\theta < \pi/2 \Leftrightarrow X_1^H X_2 \in I \Leftrightarrow I + (X_2 - X_1)X_1^{+} \in I$$

en zo ja, dan is

$$\tan \theta = \| (I - X_1X_1^{+})(X_2 - X_1)(X_1^H X_2)^{-1} X_1^H \|_2$$

$$= \| (I - X_1X_1^{+})(X_2 - X_1)X_1^{+}(I + (X_2 - X_1)X_1^{+})^{-1} \|_2 .$$

Als $\| (X_2 - X_1)X_1^{+} \|_2 < 1$ dan is zeker $\theta < \pi/2$ en

$$\tan \theta \leq \frac{\| (X_2 - X_1)X_1^{+} \|_2}{1 - \| (X_2 - X_1)X_1^{+} \|_2} .$$

Literatuur

[1] Bellman, R.; Introduction to matrix analysis. New York, McGraw-Hill, 1960.

[2] Ben-Israel, A. en Th.N.E. Greville; Generalized inverses: theory and applications. New York, Wiley, 1974.

[3] Faddeev, D.K. en V.N. Faddeeva; Computational methods of linear algebra, vertaald uit het Russisch. San Francisco, Freemann, 1963.

[4] Gantmacher, F.R.; Matrizenrechnung, 2 delen, vertaald uit het Russisch, Berlin, Deutscher Verlag der Wissensch., 1958 (ook engelse editie).

[5] Halmos, P.R.; Finite dimensional vector spaces. New York, Springer, 1970.

[6] Hoffman, K. en R. Kunze; Linear Algebra. Englewood Cliffs, Prentice-Hall, 1961.

[7] Householder, A.S.; The theory of matrices in numerical analysis. New York, Blaisdell, 1964.

[8] Marcus, M. en H. Minck; A survey of matrix theory and matrix inequalities. Boston, Allyn and Bacon, 1964.

[9] Noble, B. en J.W. Daniel, Applied linear algebra, Prentice-Hall Inc., Englewood-Cliff, 1977.

[10] Wilkinson, J.H.; The algebraic eigenvalue problem. Oxford, Clarendon, 1965.

[11] Zurmühl, R.; Matrizen und ihre technische Anwendungen. 4. Auflage. Berlin, Springer, 1964.

I N D E X