Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 1994-86

Linear Transformations of Higher-Order Spaces: Geometrical Concepts and Decompositions [1]

Lieven De Lathauwer[2] Bart De Moor[3] Joos Vandewalle

December 1994

Extract from:
L. De Lathauwer
Signaalverwerking m.b.v. Hogere-Orde Tensoren
Jaarverslag 1e mandaat specialisatiebeurs I.W.O.N.L.
Sept. 1993

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/delathau/reports/ldl-94-86.ps.Z*

[2]ESAT - Katholieke Universiteit Leuven, Kardinaal Mercierlaan 94, 3001 Leuven (Heverlee), Belgium, tel 32/16/32 11 11, fax 32/16/32 19 86, email: *Lieven.DeLathauwer@esat.kuleuven.ac.be*. Lieven De Lathauwer is a research assistant with the I.W.T. (Flemish Institute for Support of Scientific-Technological Research in Industry). The research reported in this paper was partially supported by the Belgian Institute for Scientific Research in Industry and Agriculture (I.W.O.N.L.) and the Belgian Programme on Interuniversity Poles of Attraction (IUAP-17, IUAP-50), initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. It was also supported by the European Community Research program ESPRIT, Basic Research Working Group nr. 6620 (ATHOS). The scientific responsibility rests with its authors.

[3]Bart De Moor is a research associate with the N.F.W.O. (Belgian National Fund for Scientific Research).

1. Inleiding

In de Aanvraag van vorig jaar hebben we ons voorgenomen de theorie van matrices als lineaire afbeeldingen tussen vectorruimten uit te breiden tot een theorie van "lineaire afbeeldingen tussen hogere-orde ruimten", waarvoor tensoren het formalisme bij uitstek bieden. In het eerste mandaat hebben we deze uitbreiding gerealiseerd.

De veralgemeningen die we hebben uitgewerkt, en waarvan we hier de voornaamste opsommen, konden in vele gevallen worden bekomen door de betrokken hogere-orde grootheden voor te stellen als (eerste-orde) vectoren in een hoogdimensionele vectorruimte, vervolgens een matrixtheorema uit te werken voor deze pseudo-vectorstructuren, en tenslotte opnieuw het tensorformalisme in te voeren.

We hebben ons eerst en vooral afgevraagd hoe (hogere-orde) lineaire afbeeldingen via een tensorformalisme konden worden aangepakt. We hebben kunnen bewijzen dat een samenstelling van hogere-orde transformaties leidt tot een inwendig produkt van de betrokken tensoren, hetgeen een ruimere visie geeft op de klassieke definitie van een matrixprodukt. Door ons af te vragen welke hogerdimensionele objecten door een hogere-orde transformatie in hun eigen (veralgemeende) "richting" worden afgebeeld, zijn we gekomen we tot een veralgemening van de eigenwaardenontbinding. We zijn er ook in geslaagd de belangrijke singuliere waardenontbinding uit te breiden. Ook geometrische concepten zoals orthogonaliteit en norm werden veralgemeend.

De rechtstreekse aanpak van hogere-orde afbeeldingen met tensoren biedt, naast een compacte en overzichtelijke notatie, het voordeel dat symmetrieën (in b.v. hogere-orde cumulanten) en positiegebonden informatie (in b.v. beelden) in het formalisme behouden blijven. We illustreren de praktische toepasbaarheid van het concept aan de hand van de vierde-orde signaaldeelruimte technieken van Cardoso [1].

2. Tensoren als lineaire afbeelding

Door $p' \leq p$ subscript-indices en $q' \leq q$ superscript-indices te selecteren in een (p, q)-tensor, kan een lineaire afbeelding worden geformaliseerd van de lineaire

vectorruimte van de (p,q)-tensoren naar de lineaire vectorruimte van de $(p-p',q-q')$-tensoren. Op die manier bekomen we een algemeen formalisme voor lineaire transformaties van hogere-orde vectorruimten, met als speciaal geval voor orde 1, de klassieke lineaire afbeelding van vectoren door matrices.

We geven het voorbeeld van een (2,2)-tensor $\mathbb{A}$ met $p'=1$ en $q'=1$:

$$Y_i^j = \mathbb{A}_{il}^{jk} X_k^l$$

Het *beeld* van deze lineaire afbeelding kan gedefinieerd worden als de lineaire vectorruimte van alle (1,1)-tensoren die het beeld zijn door de (2,2)-tensor van een (1,1)-tensor X_k^l. De *rang* wordt gedefinieerd als de dimensie van het beeld. De *kern* van de lineaire afbeelding is de lineaire vectorruimte van alle (1,1)-tensoren die door de (2,2)-tensoren op de 0-tensor worden afgebeeld. Dimensie van kern en beeld zijn met elkaar verbonden door de dimensiestelling

$$\dim(\text{Kern}) + \dim(\text{Beeld}) = \dim(\text{Brongebied}) = k + l$$

Het is belangrijk op te merken dat eenzelfde tensor meerdere lineaire afbeeldingen kan voorstellen, afhankelijk van welke indices men voor de sommatie selecteert.

Stelling 0.1 *De samenstelling van lineaire afbeeldingen wordt voorgesteld door het inwendig product van hun tensoren over hun gemeenschappelijke indices.*

Bewijs: We stellen bij wijze van voorbeeld de lineaire afbeeldingen

$$Y_i^j = \mathbb{A}_{il}^{jk} X_k^l \text{ en } Z_m^n = \mathbb{B}_{mj}^{ni} Y_i^j$$

samen. De totale afbeelding, van X naar Z, nl.

$$Z_m^n = \mathbb{C}_{ml}^{nk} X_k^l$$

komt overeen met het inwendig product van $\mathbb{A}$ en $\mathbb{B}$ over de indices i en j:

$$\mathbb{C}_{ml}^{nk} = \langle \mathbb{A}_{il}^{jk}, \mathbb{B}_{mj}^{ni} \rangle_{ij} = \mathbb{B}_{mj}^{ni} \mathbb{A}_{il}^{jk}$$

□

Deze stelling breidt het klassieke resultaat voor matrices uit, waar de lineaire afbeelding $f_C = f_B \circ f_A$ wordt voorgesteld door het matrixprodukt $C = BA$.

4.3. Eenheidstensor

Een belangrijk type tensor is uiteraard de eenheidstensor $\mathbb{I} = (\mathbb{I}^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p})$, die o.a. de identieke afbeelding voorstelt. Hij wordt geconstrueerd uit het uitwendig produkt van klassieke eenheidsmatrices $I = (\delta^j_i)$. (Het deltasymbool is 1 bij gelijkheid van zijn indices, en 0 in de andere gevallen.)

Voorbeeld 1:

$$\mathbb{I}^{j_1 j_2}_{i_1 i_2} \triangleq \delta^{j_1}_{i_1} \delta^{j_2}_{i_2}$$

$$\begin{aligned} Y_{i_1 i_2} &= \mathbb{I}^{j_1 j_2}_{i_1 i_2} X_{j_1 j_2} \quad \Rightarrow Y_{i_1 i_2} = X_{i_1 i_2} \\ M^{j_1}_{i_2} &= \mathbb{I}^{j_1 j_2}_{i_1 i_2} N^{i_1}_{j_2} \quad \Rightarrow M^{j_1}_{i_2} = N^{j_1}_{i_2} \end{aligned}$$

□

Het is belangrijk op te merken dat de eenheidstensor slechts in het tweede-orde geval gelijk is aan $(\delta^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p})$:

$$\mathbb{I}^{j_1}_{i_1} = \delta^{j_1}_{i_1}$$

Omdat de eenheidsmatrix I invariant is onder basisveranderingen $X = A\tilde{X}$, $Y^* = A^{-1}\tilde{Y}^*$ in de vectorruimte $\mathbf{E}$ en de duale ruimte van lineaire vormen $\mathbf{F}$, zijn de componenten van de eenheidstensor $\mathbb{I}$ ook basisonafhankelijk. Dit is niet algemeen het geval voor $(\delta^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p})$, zodat de bedoelde tensor slechts goed gedefinieerd is indien ook de particuliere basis werd gespecifieerd.

Voorbeeld 2:
We illustreren de afhankelijkheid van basis voor $(\delta^{j_1 j_2}_{i_1 i_2})$. In de basissen B_E van $\mathbf{E}$ en B_F van $\mathbf{F}$ definieert deze tensor volgende multilineaire vorm:

$$\alpha = \delta^{j_1 j_2}_{i_1 i_2} w_{j_1} x_{j_2} y^{i_1} z^{i_2}$$

Beschouw nu de basistransformatie $u_j = A^s_j \tilde{u}_s$ in $\mathbf{E}$, die overeenkomt met de transformatie $v^i = B^i_r \tilde{v}^r$ in $\mathbf{F}$, met het verband

$$A^s_j B^i_s = A^i_s B^s_j = \delta^i_j$$

dan geldt

$$\begin{aligned} \alpha &= \tilde{\delta}^{s_1 s_2}_{r_1 r_2} \tilde{w}_{s_1} \tilde{x}_{s_2} \tilde{y}^{r_1} \tilde{z}^{r_2} \\ &= \delta^{j_1 j_2}_{i_1 i_2} A^{s_1}_{j_1} \tilde{w}_{s_1} A^{s_2}_{j_2} \tilde{x}_{s_2} B^{i_1}_{r_1} \tilde{y}^{r_1} B^{i_2}_{r_2} \tilde{z}^{r_2} \end{aligned}$$

zodat

$$\tilde{\delta}^{s_1 s_2}_{r_1 r_2} = \delta^{j_1 j_2}_{i_1 i_2} A^{s_1}_{j_1} A^{s_2}_{j_2} B^{i_1}_{r_1} B^{i_2}_{r_2}$$

De componenten $\tilde{\delta}^{s_1 s_1}_{r_1 r_1}$ b.v., met $s_1 \neq r_1$ worden gegeven door

$$\tilde{\delta}^{s_1 s_1}_{r_1 r_1} = (A^{s_1}_{j_1})^2 (B^{j_1}_{r_1})^2$$

hetgeen niet noodzakelijk gelijk is aan 0.

□

3. Scalair produkt, orthogonaliteit, norm

Definitie 0.1 *Het scalair product van twee tensoren* $\mathbb{A}^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$ *en* $\mathbb{B}^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$ *wordt gedefinieerd als het inwendig product over alle indices, van de duale tensor* $\overline{\mathbb{A}}$ *met* $\mathbb{B}$*:*

$$\langle \mathbb{A}, \mathbb{B} \rangle \triangleq \overline{\mathbb{A}}^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} \mathbb{B}^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p}$$

Dit tensor scalair produkt is een uitbreiding van het gekende scalair produkt van vectoren X en Y:

$$\langle X, Y \rangle = X^* Y$$

Een ander voorbeeld is het scalair produkt van matrices. Beschouwen we de matrices $A = (a^j_i)$ en $B = (b^j_i)$, dan komt hun scalair produkt overeen met:

$$\langle A, B \rangle = a^j_i \overline{b}^i_j = \text{spoor}(AB^*)$$

Definitie 0.2 *Tensoren waarvan het scalair produkt 0 is, zijn orthogonaal.*

Deze definitie veralgemeent het orthogonaliteitsbegrip voor vectoren. Door de veralgemening van het scalair produkt wordt het dus mogelijk ook tensoren geometrisch te benaderen.

Definitie 0.3 *De Frobenius-norm van een tensor* $\mathbb{A}$ *wordt gegeven door*

$$\|\mathbb{A}\|^2_F = \overline{\mathbb{A}}^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} \mathbb{A}^{j_1 j_2 \ldots j_q}_{i_1 i_2 \ldots i_p} \tag{1}$$

Deze definitie is een uitbreiding van de Frobenius-norm van vectoren en matrices. Ook de Frobenius-norm geeft geometrische informatie: hij kan geïnterpreteerd worden als een veralgemeende "lengte" van de tensor. Het kwadraat van deze norm kan bovendien worden gezien als de "energie" in de tensor.

5. Eigenwaardenontbinding

Gegeven is de $(p+q, p+q)$-tensor $(\mathbb{A}_{i_1 i_2 \ldots i_p u_1 u_2 \ldots u_q}^{j_1 j_2 \ldots j_q v_1 v_2 \ldots v_p})$, geassocieerd met een lineaire afbeelding door sommatie over de indices $u_1 u_2 \ldots u_q; v_1 v_2 \ldots v_p$ met bereik $U_1 U_2 \ldots U_q; V_1 V_2 \ldots V_p$.

Stelling 0.2 *Als*

$$\mathbb{A}_{i_1 i_2 \ldots i_p u_1 u_2 \ldots u_q}^{j_1 j_2 \ldots j_q v_1 v_2 \ldots v_p} = \overline{\mathbb{A}}_{u_1 u_2 \ldots u_q i_1 i_2 \ldots i_p}^{v_1 v_2 \ldots v_p j_1 j_2 \ldots j_q}$$

dan heeft $\mathbb{A}$ *reële eigenwaarden* λ_e *en* $N = \prod_{k=1}^{q} U_k \prod_{l=1}^{p} V_l$ *orthonormale eigentensoren. De veralgemeende eigenwaardenontbinding wordt gegeven door*

$$\mathbb{A}_{i_1 i_2 \ldots i_p u_1 u_2 \ldots u_q}^{j_1 j_2 \ldots j_q v_1 v_2 \ldots v_p} = \sum_{\nu=1}^{N} \lambda_\nu (\mathbb{E}_\nu)_{i_1 i_2 \ldots i_p}^{j_1 j_2 \ldots j_q} (\overline{\mathbb{E}}_\nu)_{u_1 u_2 \ldots u_q}^{v_1 v_2 \ldots v_p} \tag{2}$$

Bewijs:

Onderstel

$$\mathbb{Y}_{i_1 i_2 \ldots i_p}^{j_1 j_2 \ldots j_q} = \mathbb{A}_{i_1 i_2 \ldots i_p u_1 u_2 \ldots u_q}^{j_1 j_2 \ldots j_q v_1 v_2 \ldots v_p} \mathbb{X}_{v_1 v_2 \ldots v_p}^{u_1 u_2 \ldots u_q}$$

dus ook

$$\mathbb{Y}_{v_1 v_2 \ldots v_p}^{u_1 u_2 \ldots u_q} = \mathbb{A}_{v_1 v_2 \ldots v_p j_1 j_2 \ldots j_q}^{u_1 u_2 \ldots u_q i_1 i_2 \ldots i_p} \mathbb{X}_{i_1 i_2 \ldots i_p}^{j_1 j_2 \ldots j_q}$$

Naar analogie met matrices kan men inzien dat $\mathbb{A}$ een Hermitiaanse structuur vertoont indien de bovenstaande coëfficienten complex toegevoegd zijn:

$$\mathbb{A}_{i_1 i_2 \ldots i_p u_1 u_2 \ldots u_q}^{j_1 j_2 \ldots j_q v_1 v_2 \ldots v_p} = (\mathbb{A}_{v_1 v_2 \ldots v_p j_1 j_2 \ldots j_q}^{u_1 u_2 \ldots u_q i_1 i_2 \ldots i_p})^\star \equiv \overline{\mathbb{A}}_{u_1 u_2 \ldots u_q i_1 i_2 \ldots i_p}^{v_1 v_2 \ldots v_p j_1 j_2 \ldots j_q}$$

Bemerk de veranderde volgorde van de coëfficienten.

□

De veralgemeende eigenwaardenontbinding kan als volgt worden geïnterpreteerd. De veralgemeende eigentensoren vormen een orthonormale basis voor de (p,q)-tensoren. Deze basistensoren hebben de eigenschap dat zij bij lineaire afbeelding door $\mathbb{A}$ eenvoudigweg gescaleerd worden. Zij worden dus a.h.w. in hun veralgemeende "richting" afgebeeld. De mate waarin zo'n eigentensor $\mathbb{E}_\nu$ wordt versterkt/verzwakt, wordt gegeven door de veralgemeende eigenwaarde λ_ν.

6. Singuliere Waardenontbinding

Gegeven is de $(p+q, r+s)$-tensor $(\mathbb{A}_{i_1 i_2 \ldots i_p u_1 u_2 \ldots u_q}^{j_1 j_2 \ldots j_r v_1 v_2 l dots v_s})$ waarvan de indices i_k $(1 \leq k \leq p)$, u_l $(1 \leq l \leq q)$, j_m $(1 \leq m \leq r)$, v_n $(1 \leq n \leq s)$ een bereik hebben van resp. I_k, U_l, J_m en V_n. Definieer $N = \min\{\prod_{k=1}^{q} U_k \prod_{l=1}^{p} V_l, \prod_{m=1}^{p} I_m \prod_{n=1}^{r} J_r\}$

Stelling 0.3 *De veralgemeende Singuliere Waardenontbinding, geassocieerd met de lineaire afbeelding door sommatie over $u_1 u_2 \ldots u_q; v_1 v_2 \ldots v_s$, wordt gegeven door*

$$\mathbb{A}_{i_1 i_2 \ldots i_p u_1 u_2 \ldots u_q}^{j_1 j_2 \ldots j_r v_1 v_2 \ldots v_s} = \sum_{\nu=1}^{N} \sigma_\nu (\mathbb{S}_\nu)_{i_1 i_2 \ldots i_p}^{j_1 j_2 \ldots j_r} (\overline{\mathbb{T}}_\nu)_{u_1 u_2 \ldots u_q}^{v_1 v_2 \ldots v_s} \tag{3}$$

met

$$\sigma_\nu \geqslant 0$$

$$\langle \mathbb{S}_\nu, \mathbb{S}_{\nu'} \rangle = \begin{cases} 0 & (\nu \neq \nu') \\ 1 & (\nu = \nu') \end{cases}$$

$$\langle \mathbb{T}_\nu, \mathbb{T}_{\nu'} \rangle = \begin{cases} 0 & (\nu \neq \nu') \\ 1 & (\nu = \nu') \end{cases}$$

Bewijs:

De veralgemening kan worden aangetoond door de tensor $\mathbb{A}$ te "ont-wikkelen" tot het matrixniveau, van deze matrix de Singuliere Waardenontbinding te bepalen en tenslotte de bekomen pseudo-matrices weer voor te stellen als tensoren.

□

De veralgemeende Singuliere Waardenontbinding geeft een belangrijk inzicht in het mechanisme van een willekeurige hogere-orde lineaire afbeelding. De rechter singuliere tensoren $\mathbb{T}_\nu$ vormen een orthonormale basis voor de (s,q)-tensoren van de bronruimte. Zij worden bij de lineaire afbeelding door $\mathbb{A}$ afgebeeld op de overeenkomende linker singuliere tensoren $\mathbb{S}_\nu$, gescaleerd met de veralgemeende singuliere waarden σ_ν. Deze linker singuliere tensoren vormen een orthonormale basis voor de (p,r)-tensoren van het beeldgebied.

7. Illustratie: vierde-orde signaaldeelruimte

Ofschoon de meeste principes die opduiken wanneer men tensoren beschouwt als lineaire afbeeldingen, vrij goed te voorspellen zijn vanuit de analogie met het tweede-orde matrixgeval, zijn ze zeer zinvol en bieden ze een kader waarin talrijke problemen van signaalverwerking kunnen worden uitgewerkt.

Het belang van de voorgestelde zienswijze, wordt mooi geïllustreerd door het werk van Cardoso[1] aangaande hogere-orde signaalverwerking voor smalbandige meerkanaalssystemen (b.v. [1]). In de artikels over dit onderwerp werden reeds een beperkt aantal van bovenstaande ideeën uitgewerkt.

Men beschouwt het lineair verband

$$y_i = a_i^j x_j$$

tussen I sensoruitgangen y_i en J smalbandige bronsignalen x_j.

Uit de lineariteitseigenschap van cumulanten volgt dat de spreiding (tweede orde) en $(2,2)$-cumulant of "quadricovariantie" (vierde orde) van Y gegeven worden door

$$C_j^k(Y) = a_j^r C_r^s(X) \overline{a}_s^k$$

$$\mathbb{C}_{jl}^{ik}(Y) = \overline{a}_r^i a_j^s \mathbb{C}_{su}^{rt} \overline{a}_t^k a_l^u$$

Door sommatie over de indices k (en l) definieert men 2 lineaire afbeeldinge n waarvan het beeld voor spreiding en quadricovariantie respectievelijk "signaal-deelruimte" ("signal subspace") en "4-de orde signaal deelruimte" ("4th-order signal subspace - FOSS") genoemd wordt.

[1] Eveneens participant in ATHOS

Bij onafhankelijke bronsignalen zijn de cumulanten van X diagonaal. Definiëren we σ_ν en c_ν als de scalaire $(1,1)$- en $(2,2)$- cumulant van bron x_ν en a_ν als de ν-de stuurvector, dan geldt:

$$C_j^k(Y) = \sum_\nu \sigma_\nu (a_\nu)^k (a_\nu)_j$$

$$\mathbb{C}_{jl}^{ik}(Y) = \sum_\nu c_\nu (a_\nu)^i (a_\nu)_j (a_\nu)^k (a_\nu)_l$$

Het is onmiddellijk duidelijk dat bij lineair onafhankelijke stuurvectoren de FOSS dimensie J heeft. In het geval van gecorreleerde bronnen, is de FOSS echter van dimensie J^2. De signaal-deelruimte daarentegen, is in beide gevallen J-dimensionaal.

De voordelen van de benadering van hogere-orde tensoren als lineaire afbeeldinge n, in vergelijking met de klassieke tweede-orde benadering, kunnen dan worden samengevat als volgt:

- Men kan het onderscheid maken tussen bronsignalen die al dan niet volledig onafhankelijk zijn. De benadering maakt het dus mogelijk de betrokken signalen grondiger te analyseren.
- Door de rang-1 matrices van de FOSS te bepalen, kan men de onbekende modelparameters a_i^j identificeren [2].
- Men kan aantonen dat, in tegenstelling tot de tweede-orde analyse, het aantal te identificeren stuurvectoren het aantal sensoren mag overtreffen [3].
- De benadering van hogere-orde tensoren als lineaire afbeeldingen maakt het mogelijk de hardware-kost van het aantal sensoren te verlagen.
- Door de hogere-orde benadering maakt men de analyse blind voor Gaussiaanse meetruis.

Referenties

[1] J.-F. Cardoso, "Higher-Order Narrow-Band Array Processing," in J.L. Lacoume (Ed.), *Higher Order Statistics.* Proceedings of the International Signal Processing Workshop on Higher Order Statistics, Chamrousse, France, July 10-12, 1991, pp. 39-48.

[2] J.-F. Cardoso, "Fourth-Order Cumulant Structure Forcing. Application to Blind Array Processing." *Proceedings IEEE Workshop on SSAP-92*, pp. 136-139.

[3] J.-F. Cardoso, "Super-Symmetric Decomposition of the Fourth-Order Cumulant Tensor. Blind Identification of More Sources than Sensors," *Proceedings ICASSP-91,* Vol. 5, pp. 3109-3112.