Tensor Decomposition

16.01.20
Karolina Spiel,
Korbinian Schmidhuber

Vertiefung der Grundlagen der Computerlinguistik WiSe 19/20
Prof. Dr. Klaus Schulz, Luisa März

Inhalt

Was ist ein Tensor?

Was ist ein Tensor?

Skalar

Wie viel beträgt die Temperatur in München?

Nur **1** Komponente / Einheit / Zahl A nötig:
z.B. **A** = 0°C, 273 K, …

→ **0** Basisvektoren pro Komponente

München

Was ist ein Tensor?

Skalar

Wie viel beträgt die Temperatur in München?

Nur **1** Komponente / Einheit / Zahl A nötig:
z.B. **A** = 0°C, 273 K, …

→ **0** Basisvektoren pro Komponente

→ **Rang-0 Tensor**

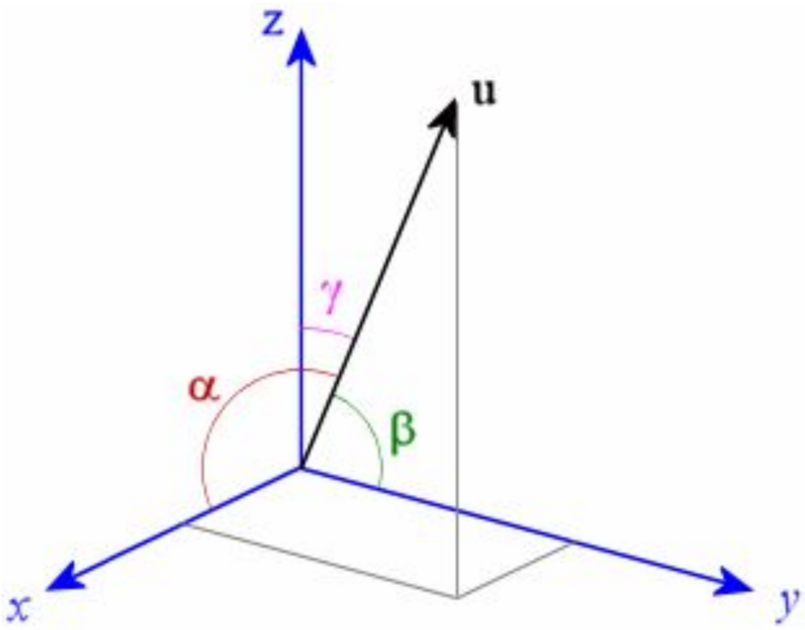
Was ist ein Tensor?

Vektor

Wie weit ist die Spitze des Olympiaturms entfernt?

3 Komponenten nötig: z.B.
19 km Richtung Osten ($\mathbf{A_x}$),
12 km Richtung Süden ($\mathbf{A_y}$),
291 m in die Höhe ($\mathbf{A_z}$)

$$\vec{u} = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} = \begin{pmatrix} 19 \\ 12 \\ 0.291 \end{pmatrix} = 21.2 \text{ km}$$



→ **1** Basisvektor pro Komponente

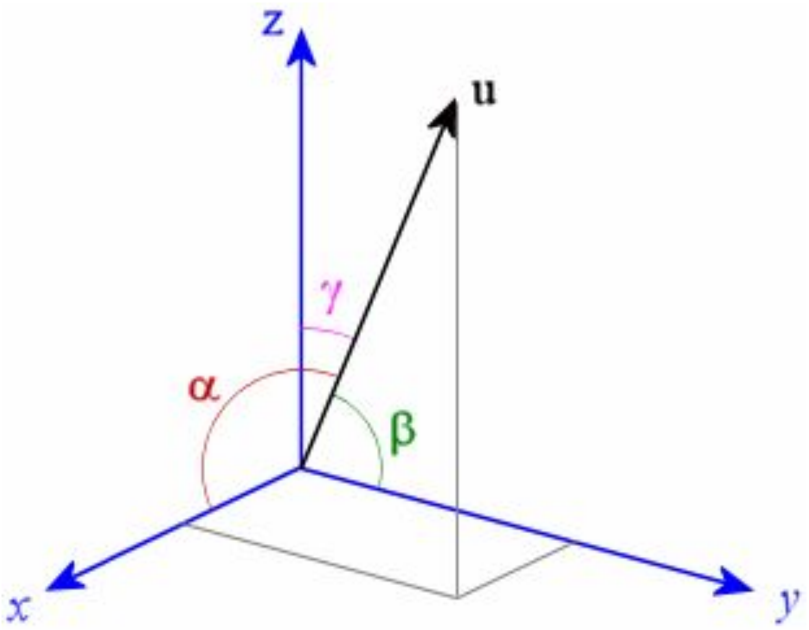
Was ist ein Tensor?

Vektor

Wie weit ist die Spitze des Olympiaturms entfernt?

3 Komponenten nötig: z.B.
19 km Richtung Osten (**A_x**),
12 km Richtung Süden (**A_y**),
291 m in die Höhe (**A_z**)

$$\vec{u} = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} = \begin{pmatrix} 19 \\ 12 \\ 0.291 \end{pmatrix} = 21.2 \text{ km}$$



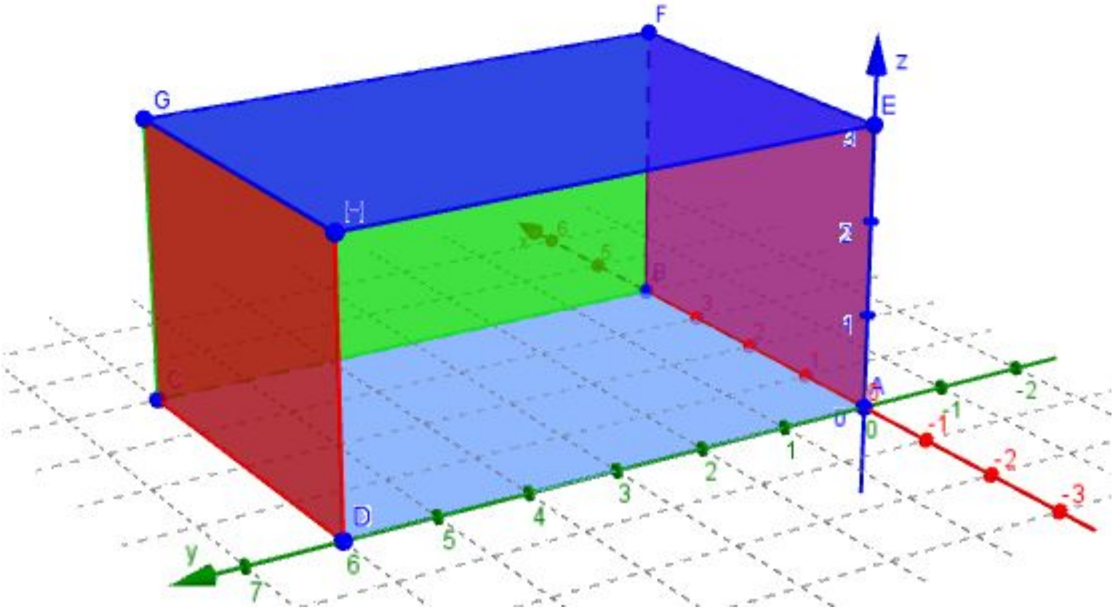
→ **1** Basisvektor pro Komponente

→ **Rang-1 Tensor**

Was ist ein Tensor?

Matrix

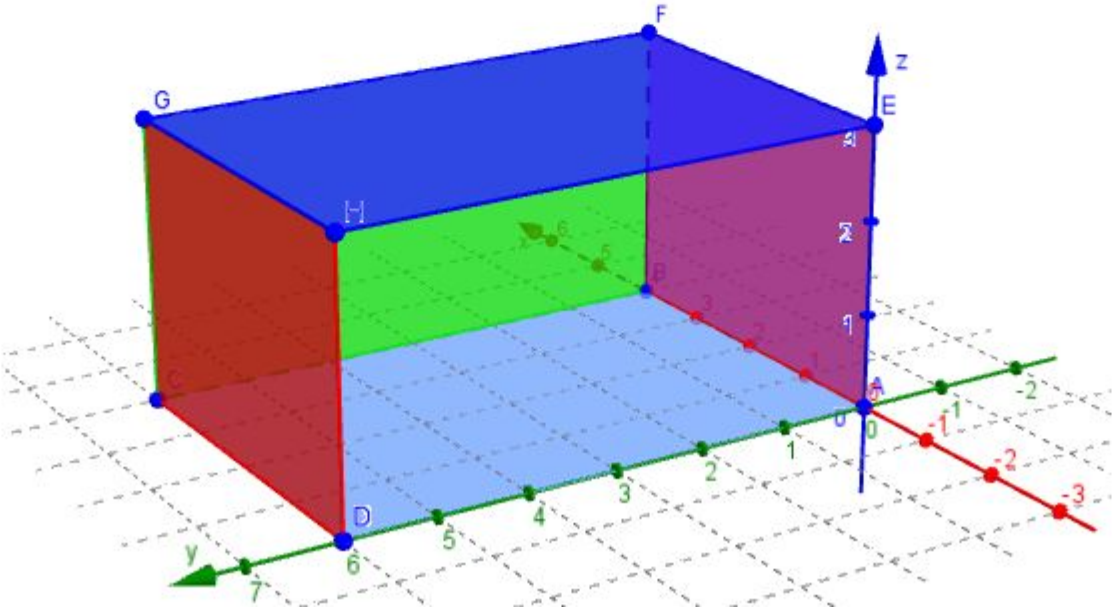
Welche Kräfte wirken an einem Punkt in einem Stahlträger im Olympiaturm?



Was ist ein Tensor?

Matrix

Welche Kräfte wirken an einem Punkt in einem Stahlträger im Olympiaturm?



9 Komponenten nötig:

Für die Fläche mit Flächenvektor entlang x-Achse und Kraft entlang x-Achse: $\mathbf{A}_{xx}$, für die Fläche mit Flächenvektor entlang x-Achse und Kraft entlang y-Achse: $\mathbf{A}_{xy}$, für die Fläche mit Flächenvektor entlang x-Achse und Kraft entlang z-Achse: $\mathbf{A}_{xz}$, …

Was ist ein Tensor?

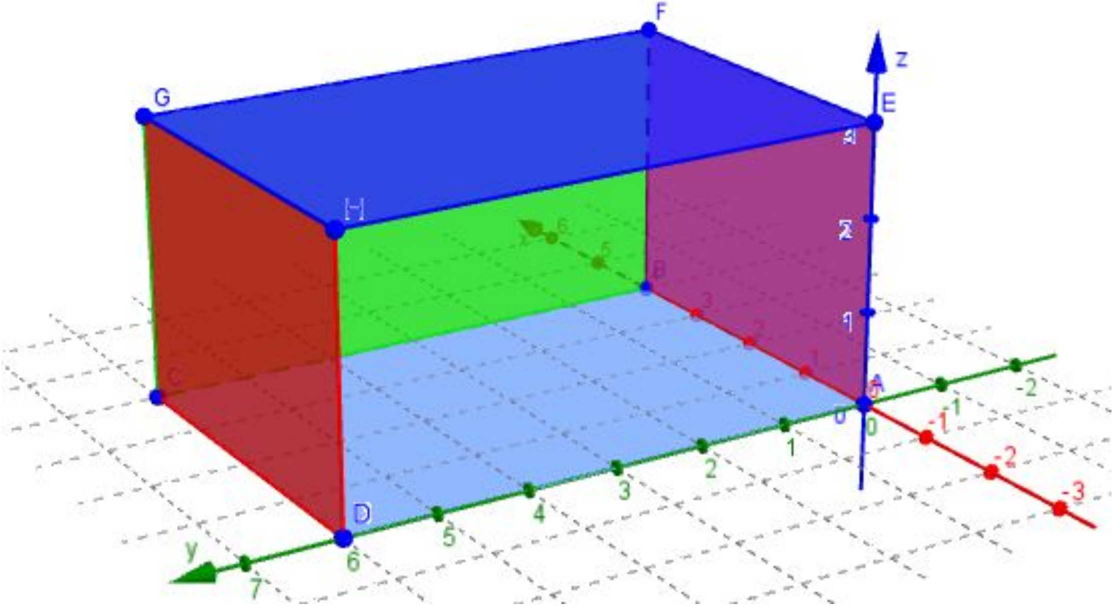
Matrix

Welche Kräfte wirken an einem Punkt in einem Stahlträger im Olympiaturm?

9 Komponenten nötig:

$$\begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix}$$

→ **2** Basisvektoren pro Komponente



Was ist ein Tensor?

Matrix

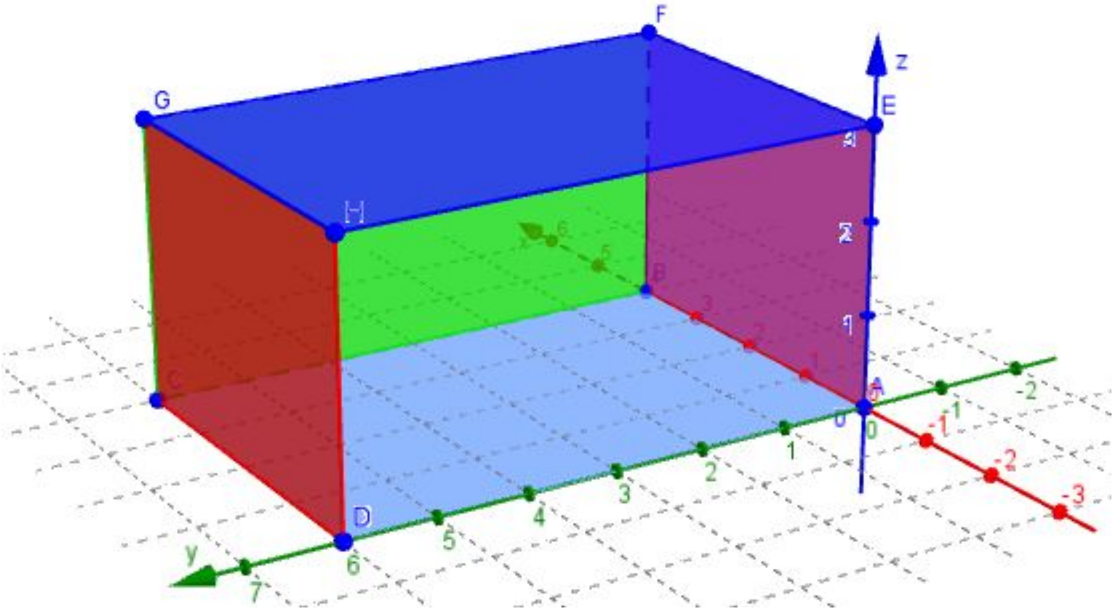
Welche Kräfte wirken an einem Punkt in einem Stahlträger im Olympiaturm?

9 Komponenten nötig:

$$\begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix}$$

→ **2** Basisvektoren pro Komponente

→ **Rang-2 Tensor**



Was ist ein Tensor?

"In an **n-dimensional space**, a tensor of **rank k** is specified by $\mathbf{n^k}$ parameters, known as the components of the tensor." (Deshmukh 2019: 37)

Rang-0 Tensor in 3d: 1 Komponente (A),

Rang-1 Tensor in 3d: 3 Komponenten (A_x, A_y, A_z),

Rang-2 Tensor in 3d: 9 Komponenten (A_{xx}, A_{xy}, A_{xz}, …),

Was ist ein Tensor?

"In an **n-dimensional space**, a tensor of **rank k** is specified by $\mathbf{n^k}$ parameters, known as the components of the tensor." (Deshmukh 2019: 37)

Rang-0 Tensor in 3d: 1 Komponente (A),

Rang-1 Tensor in 3d: 3 Komponenten (A_x, A_y, A_z),

Rang-2 Tensor in 3d: 9 Komponenten (A_{xx}, A_{xy}, A_{xz}, …),

Rang-3 Tensor in 3d: 27 Komponenten (A_{xxx}, A_{xyx}, A_{xxz}, …),

Rang-4 Tensor in 3d: 81 Komponenten, (A_{xxxx}, A_{xyxx}, A_{xyyx}, …)

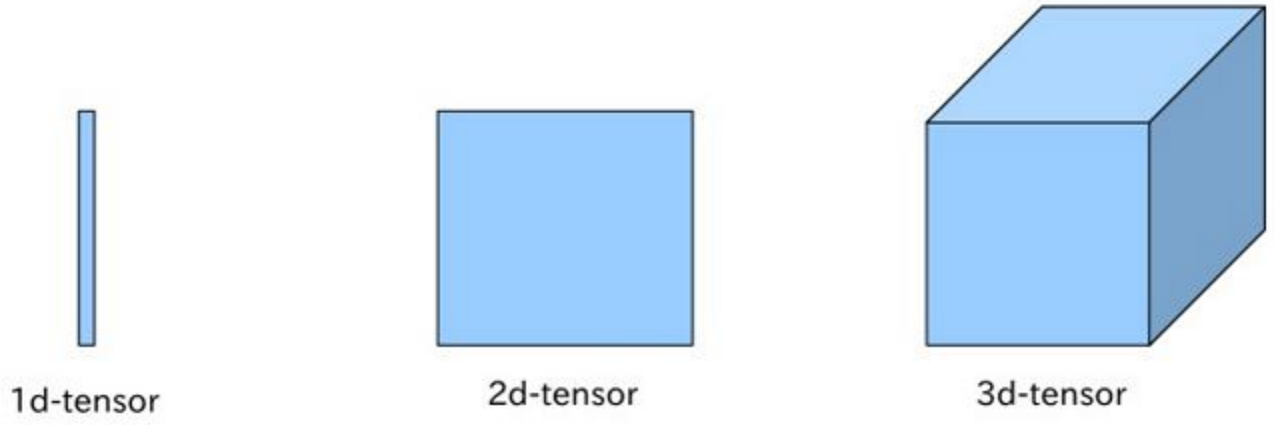
usw.

Was ist ein Tensor?

"In an **n-dimensional space**, a tensor of **rank k** is specified by $\mathbf{n^k}$ parameters, known as the components of the tensor." (Deshmukh 2019: 37)

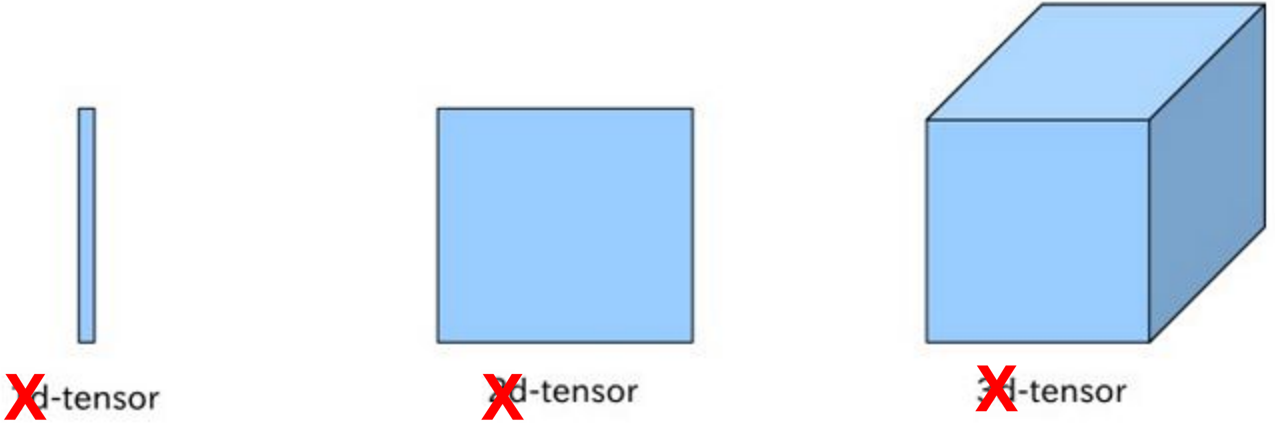
Was ist ein Tensor?

"In an **n-dimensional space**, a tensor of **rank k** is specified by $\mathbf{n^k}$ parameters, known as the components of the tensor." (Deshmukh 2019: 37)



Was ist ein Tensor?

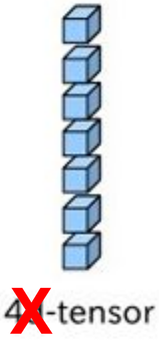
"In an **n-dimensional space**, a tensor of **rank k** is specified by **n^k** parameters, known as the components of the tensor." (Deshmukh 2019: 37)



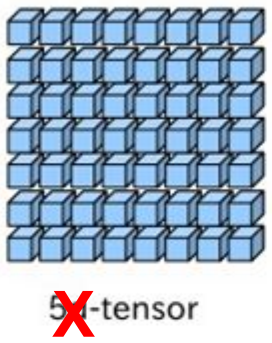
RANG-1 Tensor | RANG-2 Tensor | RANG-3 Tensor

Was ist ein Tensor?

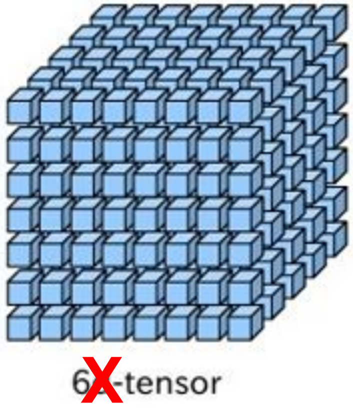
"In an **n-dimensional space**, a tensor of **rank k** is specified by **n**^{k} parameters, known as the components of the tensor." (Deshmukh 2019: 37)



RANG-4 Tensor



RANG-5 Tensor



RANG-6 Tensor

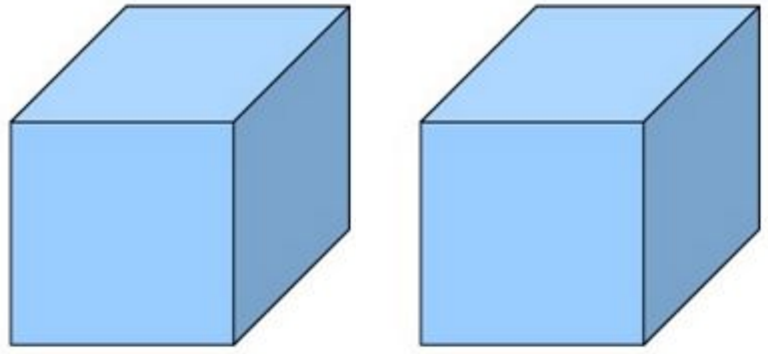
Grundlegende Tensor-Konzepte

Skalarprodukt zweier Tensoren

$$\langle A, B \rangle = \sum_{i,j,k} a_{ijk} b_{ijk}$$

Skalarprodukt zweier Tensoren

$$\langle A, B \rangle = \sum_{i,j,k} a_{ijk} b_{ijk}$$



Frobenius-Norm eines Tensors

$$\|A\|_F = \langle \mathrm{A}, \mathrm{A} \rangle^{1/2} = \left(\sum_{i,j,k} a_{ijk}^2 \right)^{1/2}$$

→ Entspricht der Wurzel des Skalarprodukts des Tensors mit sich selbst

Tensor-Matrix Multiplikation

- Für einen Rang 3 Tensor und eine Matrix gibt es 3 Multiplikationen (Für jeden Rang eine)

- Für einen Tensor $\mathcal{A} \in \mathbb{R}^{l \times m \times n}$ und eine Matrix $U \in \mathbb{R}^{l_0 \times l}$ ist die sogenannte 1-mode Multiplikation wie folgt definiert

Tensor-Matrix Multiplikation

Tensor: $\mathcal{A} \in \mathbb{R}^{l \times m \times n}$, **Matrix:** $U \in \mathbb{R}^{l_0 \times l}$

1-Mode Multiplikation:

$$(\mathcal{A} \times_1 U)(j, i_2, i_3) = \sum_{k=1}^{l} u_{j,k} * a_{k,i_2,i_3}$$

Tensor-Matrix Multiplikation

Zum Vergleich:

1-Mode Tensor-Matrix Multiplikation:

$$(\mathcal{A} \times_1 U)(j, i_2, i_3) = \sum_{k=1}^{l} u_{j,k} * a_{k,i_2,i_3}$$

Matrix-Matrix Multiplikation:

$$(A \times U)(j, i) = \sum_{k=1}^{l} u_{j,k} * a_{k,i}$$

Tensor-Matrix Multiplikation

1-Mode Tensor-Matrix Multiplikation: $(\mathcal{A} \times_1 U)(j, i_2, i_3) = \sum_{k=1}^{l} u_{j,k} * a_{k,i_2,i_3}$

2-Mode Tensor-Matrix Multiplikation: $(\mathcal{A} \times_2 U)(i_1, j, i_3) = \sum_{k=1}^{m} u_{j,k} * a_{i_1,k,i_3}$

3-Mode Tensor-Matrix Multiplikation: $(\mathcal{A} \times_3 U)(i_1, i_2, j) = \sum_{k=1}^{n} u_{j,k} * a_{i_1,i_2,k}$

Tensor-Decomposition-Algorithmen

Tensor-Decomposition-Algorithmen

Bekannteste Algorithmen:

1. **Canonical Polyadic Decomposition (CPD)**
2. **Tucker Decomposition (TD)**

- Beide: Tensor-Zerlegung in ein dyadisches Produkt, aber mit unterschiedlichen Struktureigenschaften
- CDP i.d.R. zur Schätzung latenter Parameter, TD i.d.R. u.a. zur Komprimierung und Dimensionsreduktion
- Basis für viele weitere Zerlegungen

Tensor-Decomposition-Algorithmen

Bekannteste Algorithmen:

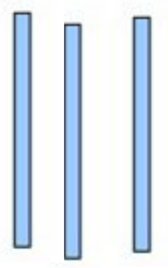
1. **Canonical Polyadic Decomposition (CPD)**

   = Rang-Zerlegung:

   Tensor wird als **Summe einer (endlichen) Anzahl an Rang-1 Tensoren** ausgedrückt

   → z.B. Jennrich Algorithm, Alternating Square Algorithm, Tensor Power Method, …

   (siehe hierfür Rabanser et al. 2017)



Tensor-Decomposition-Algorithmen

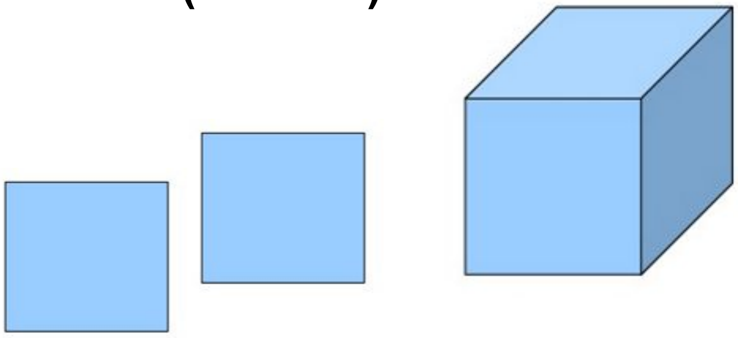
Bekannteste Algorithmen:

2. **Tucker Decomposition (TD)**

   Tensor wird zerlegt in **einen Kern-Tensor und mehrere Matrizen**

   z.B. Zerlegung in 1 Kern-Tensor und 3 Matrizen

   → **Higher Order Singular Value Decomposition (HOSVD)**



Higher Order Singular Value Decomposition (HOSVD)

Bei der HOSVD wird ein Tensor in einen Kern-Tensor und 3 Matrizen zerlegt:

Higher Order Singular Value Decomposition (HOSVD)

Bei der HOSVD wird ein Tensor in einen Kern-Tensor und 3 Matrizen zerlegt:

$$\mathcal{A} \in \mathbb{R}^{l \times m \times n}$$

Higher Order Singular Value Decomposition (HOSVD)

Bei der HOSVD wird ein Tensor in einen Kern-Tensor und 3 Matrizen zerlegt:

$$\mathcal{A} \in \mathbb{R}^{l \times m \times n}$$
$$\mathcal{S} \in \mathbb{R}^{l \times m \times n}$$

Higher Order Singular Value Decomposition (HOSVD)

Bei der HOSVD wird ein Tensor in einen Kern-Tensor und 3 Matrizen zerlegt:

$$\mathcal{A} \in \mathbb{R}^{l \times m \times n}$$

$$S \in \mathbb{R}^{l \times m \times n}$$

$$U^{(1)} \in \mathbb{R}^{l \times l}, U^{(2)} \in \mathbb{R}^{m \times m}, U^{(3)} \in \mathbb{R}^{n \times n}$$

Higher Order Singular Value Decomposition (HOSVD)

Bei der HOSVD wird ein Tensor in einen Kern-Tensor und 3 Matrizen zerlegt:

$$\mathcal{A} \in \mathbb{R}^{l \times m \times n}$$

$$\mathcal{S} \in \mathbb{R}^{l \times m \times n}$$

$$U^{(1)} \in \mathbb{R}^{l \times l}, U^{(2)} \in \mathbb{R}^{m \times m}, U^{(3)} \in \mathbb{R}^{n \times n}$$

HOSVD: $\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}$

Higher Order Singular Value Decomposition (HOSVD)

Die Bestandteile der Zerlegung haben dabei folgende Eigenschaften:

Higher Order Singular Value Decomposition (HOSVD)

Die Bestandteile der Zerlegung haben dabei folgende Eigenschaften:

- U(1), U(2), U(3) sind orthogonal Matrizen

Higher Order Singular Value Decomposition (HOSVD)

Die Bestandteile der Zerlegung haben dabei folgende Eigenschaften:

- U(1), U(2), U(3) sind orthogonal Matrizen
- Zwei jeweils unterschiedliche Slices von S sind orthogonal in Bezug auf das Skalarprodukt (<S(i,:,:), S(j,:,:)> = <S(:,i,:), S(:,j,:)> = <S(:,:,i), S(:,:,j)> = 0, mit i ungleich j)

Higher Order Singular Value Decomposition (HOSVD)

Die 1-Mode Singulärwerte sind definiert als:

$$\sigma_i^{(1)} = ||S(i,:,:)||_F, \quad i = 1, ..., l$$

Higher Order Singular Value Decomposition (HOSVD)

Sie besitzen die Eigenschaft, dass Singulärwerte mit zunehmendem Index kleiner werden:

$$\sigma_1^{(1)} \geq \sigma_2^{(1)} \geq \dots \geq \sigma_I^{(1)}$$

Higher Order Singular Value Decomposition (HOSVD)

- 2-Mode Singulärwerte und 3-Mode Singulärwerte sind analog dazu definiert und besitzen dieselbe Eigenschaft

Anwendungen von Tensor Decomposition

- Kompression von Tensoren
- Dimensionsreduktion
- Inferenz in Latent Variable Models
- Entdecken neuer Relationen in Multilayer Networks
- uvm.

Quellen und Empfehlungen

Literatur:

Deshmukh, P.C. (2019): **Foundations of Classical Mechanics.** Cambridge: University Press.

Elden, L. (2007): **Matrix Methods in Data Mining and Pattern Recognition**. Philadelphia: SIAM.

Rabanser, S., Shchur, O., Günnemann, S. (2017): **Introduction to Tensor Decomposition and their Applications in Machine Learning.** München: Technische Universität.

Abbildungen:

S. 6 f.: http://geogebra-rlp.zum.de/images/b/ba/Quader_auf_KOS.PNG [Letzter Zugriff: 15.01.20]

S. 8 ff.:https://www.intmath.com/vectors/img/cosines.gif [Letzter Zugriff: 15.01.20]

S. 15 ff.: https://www.cc.gatech.edu/~san37/img/dl/tensor.png [Letzter Zugriff: 15.01.20]

YouTube:

What's a Tensor: https://youtu.be/f5liqUk0ZTw [Letzter Zugriff: 15.01.20]

Introduction to Tensors: https://youtu.be/uaQeXi4E7gA [Letzter Zugriff: 15.01.20]