

# Basen, Unterräume, Gram-Schmidt

Katja Markert

Institut für Computerlinguistik  
Uni Heidelberg  
markert@cl.uni-heidelberg.de

May 21, 2019

- 1 Bisher: Sparse Embeddings
- 2 Probleme: Zuviel Rauschen, versteckte Ähnlichkeiten
- 3 Eine Möglichkeit zum Erreichen von dichten Vektoren ist mittels Matrizenzerlegungen (Singular Value Decomposition)
- 4 Jetzt: Hintergrund zu den mathematischen Begriffen und Methoden, die für Matrizenzerlegung notwendig ist
- 5 Insbesondere: Orthonormalität und Orthonormalbasen

- 1 Basen und Unterräume
- 2 Gram-Schmidt Orthonormalisierungsverfahren

- 1 Basen und Unterräume
- 2 Gram-Schmidt Orthonormalisierungsverfahren

Gegeben sei ein (reeller) Vektorraum  $V$ . Gilt für eine nicht-leere Teilmenge  $U \subseteq V$ , dass

- 1  $\vec{u}_1 + \vec{u}_2 \in U$  für alle  $\vec{u}_1, \vec{u}_2 \in U$
- 2  $\lambda \cdot \vec{u} \in U$  für alle  $\vec{u} \in U, \lambda \in \mathbb{R}$

so nennt man  $U$  einen **Unterraum** von  $V$ . Man sagt,  $U$  ist bezüglich der Vektoraddition und Skalarmultiplikation **abgeschlossen**.

Welche der folgenden Mengen ist ein Untervektorraum von  $\mathbb{R}^3$  ?

- 1 Sei  $U$  die Menge der Vektoren im  $\mathbb{R}^3$  mit der ersten Komponente  $=0$ , also  $U = \{(0, y, z) | y, z \in \mathbb{R}\}$
- 2 Sei  $U$  die Menge der Vektoren im  $\mathbb{R}^3$  mit der ersten Komponente  $=1$ , also  $U = \{(1, y, z) | y, z \in \mathbb{R}\}$
- 3 Sei  $U$  die Menge der Vektoren im  $\mathbb{R}^3$  so dass gilt  $U = \{(x, y, z) | x + y = z; x, y, z \in \mathbb{R}\}$

- Jeder Schnitt zweier Untervektorräume ist wieder ein Untervektorraum
- Jeder Unterraum muss den Nullvektor enthalten.

Gegeben sei ein reeller Vektorraum  $V$  und eine endliche Teilmenge  $M := \{\vec{v}_1, \dots, \vec{v}_n\}$  von  $V$ . Dann definieren wir die Menge ihrer **Linearkombinationen** als **lineare Hülle oder span** von  $M$ .

$$\text{span}(M) := \{\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 \dots + \lambda_n \vec{v}_n \mid \lambda_i \in \mathbb{R} \text{ fuer alle } i\}$$

Die lineare Hülle der leeren Menge wird als  $\{\vec{0}\}$  definiert.



- Die lineare Hülle bildet immer einen Untervektorraum. Man nennt ihn den von  $M$  **aufgespannten** Unterraum.
- Was ist die lineare Hülle eines einzelnen Vektors im  $\mathbb{R}^2$ ?
- Was ist im  $\mathbb{R}^3$  die lineare Hülle der beiden Vektoren  $(4, 0, 0)$  und  $(0, 3, 0)$ ?

Vektoren  $\{\vec{v}_1, \dots, \vec{v}_n\} \in V$  heißen **linear unabhängig** wenn für alle  $\lambda_i \in \mathbb{R}$  gilt

$$\lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n = \vec{0} \iff \lambda_i = 0 \text{ fuer alle } i$$

Beispiele:

- 1 Im  $\mathbb{R}^2$  sind die Vektoren  $(2, 2)$  und  $(-6, 4)$  linear unabhängig
- 2 Im  $\mathbb{R}^3$  sind die Vektoren  $(1, 0, 0)$ ,  $(0, 1, 0)$  und  $(0, 0, 1)$  linear unabhängig
- 3 Im  $\mathbb{R}^3$  sind die Vektoren  $(2, 3, 4)$ ,  $(1, 1, 5)$  und  $(3, 4, 9)$  linear abhängig

- 1 Enthält unsere Vektormenge den Nullvektor, so ist sie nie linear unabhängig
- 2 Bei linear unabhängigen Vektoren lässt sich kein Vektor als Linearkombination der anderen darstellen.
- 3  $n$  linear unabhängige Vektoren spannen einen Untervektorraum der Dimension  $n$  auf
- 4 Der  $\mathbb{R}^n$  hat die Dimension  $n$ , d.h.  $n + 1$  Vektoren sind immer linear abhängig.
- 5 Eine Menge linear unabhängiger Vektoren, die einen Vektorraum aufspannen, nennen wir eine **Basis** eines Vektorraums.

- 1  $(1, 0)$  und  $(0, 1)$  ist eine Basis des  $\mathbb{R}^2$ . Die Vektoren sind unabhängig und alle Vektoren lassen sich als Linearkombination der beiden darstellen. Die Basis ist auch orthonormal, d.h. die Vektoren sind Einheitsvektoren und stehen aufeinander senkrecht. (Wie waren noch mal die mathematischen Definitionen von Länge und Orthogonalität?). Dies nennt man auch die **Standardbasis**.
- 2 Im  $\mathbb{R}^2$  sind die Vektoren  $(2, 2)$  und  $(-6, 4)$  linear unabhängig und bilden eine Basis. Alle Vektoren lassen sich als Linearkombination der beiden darstellen. Die Basis besteht nicht aus Einheitsvektoren und die Vektoren sind nicht orthogonal zueinander.

- 1 Ich habe im  $\mathbb{R}^2$  die Basis  $B_1 = (2, 2)$  und  $(-6, 4)$  gewählt. Der Vektor  $(1, 2)_{B_1}$  unter der Basis  $B_1$  ist dann  $1 \cdot (2, 2) + 2 \cdot (-6, 4) = (-10, 10)$  in der Standardbasis.
- 2 Basistransformation im allgemeinen (nicht zur Standardbasis) ist komplizierter. Ich verweise auf entsprechende Textbücher
- 3 Wir sind wegen SVD vor allem an Mengen von Orthonormalvektoren und Orthonormalbasen interessiert.

- 1 Es sei der  $\mathbb{R}^n$  gegeben sowie die Abbildung  $f: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $f(\vec{v}, \vec{w}) := 0 \iff \vec{v} = \vec{w}$  und  $f(\vec{v}, \vec{w}) := 1$  sonst.
  - 1 Zeigen Sie, dass  $f$  eine Metrik auf dem  $\mathbb{R}^n$  definiert
  - 2 Zeigen Sie, dass diese Metrik nicht von einer Norm induziert ist. (Hinweis: Wäre die Metrik von einer Norm induziert, müsste für alle Vektoren gelten  $f(\vec{v}, 0) = \|\vec{v}\|$ . Was folgt daraus? Kann das alle Normaxiome erfüllen?)

- 1 Gegeben seien die folgenden drei Vektoren:  
 $\vec{a} = (1, 2, 3)$ ,  $\vec{b} = (3, 2, 1)$ ,  $\vec{c} = (5, 4, 5)$  und  $\vec{d} = (1, 4, 3)$ 
  - 1 Sind  $\vec{a}, \vec{b}, \vec{c}, \vec{d}$  linear unabhängig?
  - 2 Sind  $\vec{a}, \vec{b}, \vec{c}$  linear unabhängig?
  - 3 Sind  $\vec{a}, \vec{b}, \vec{d}$  linear unabhängig?
  - 4 Berechnen Sie den Winkel zwischen  $\vec{c}$  und  $\vec{d}$
- 2 Zeigen Sie, dass jeder Unterraum eines Vektorraums  $V$  den Nullvektor enthalten muss
- 3 Gegeben sei der  $\mathbb{R}^2$ . Sind alle Geraden Unterräume?
- 4 Welche Unterräume gibt es im  $\mathbb{R}$ ?

- 1 Basen und Unterräume
- 2 Gram-Schmidt Orthonormalisierungsverfahren**



Gegeben sei eine endliche, linear unabhängige Menge von Vektoren im  $\mathbb{R}^n$ , d.h.  $S = \{\vec{w}_1, \dots, \vec{w}_k\}$  mit  $k \leq n$ . Diese spannen einen Untervektorraum  $U$  auf. Wie finde ich eine Orthonormalbasis  $S' = \{\vec{v}_1, \dots, \vec{v}_k\}$  von  $U$ ? Gesucht also:

- 1 Alle  $\vec{v}_i$  haben Länge 1
- 2  $S' = \{\vec{v}_1, \dots, \vec{v}_k\}$  ist eine Basis von  $U$
- 3  $\langle \vec{v}_i, \vec{v}_j \rangle = 0$  wenn  $i \neq j$

Lösung: Das Gram-Schmidtsche Orthonormalisierungsverfahren

- 1 Normalisierung des ersten Vektors

$$v_1 = \frac{w_1}{\|w_1\|}$$

- 2 Orthogonalisierung des zweiten Vektors  $w_2$  in Bezug auf  $v_1$

$$v'_2 = w_2 - \langle v_1, w_2 \rangle \cdot v_1$$

- 3 Normalisieren des  $v'_2$

$$v_2 = \frac{v'_2}{\|v'_2\|}$$

- 4 Orthogonalisieren des dritten Vektors  $w_3$  in Bezug auf  $v_1$  und  $v_2$

$$v'_3 = w_3 - \langle v_1, w_3 \rangle \cdot v_1 - \langle v_2, w_3 \rangle v_2$$

- 5 Normalisieren des  $v'_3$

$$v_3 = \frac{v'_3}{\|v'_3\|}$$

$$v'_i = w_i - \sum_{j=1}^{i-1} \langle v_j, w_i \rangle \cdot v_j$$

$$v_i = \frac{v'_i}{\|v'_i\|}$$

Gegeben seien die drei Vektoren

$$w_1 = (1, 0, 2, 1), w_2 = (2, 2, 3, 1), w_3 = (1, 0, 1, 0).$$

$$v_1 = \frac{w_1}{\|w_1\|} = \left( \frac{1}{\sqrt{6}}, 0, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right)$$

$$v'_2 = w_2 - \langle v_1, w_2 \rangle \cdot v_1 = (2, 2, 3, 1) - \frac{9}{\sqrt{6}} \cdot \left( \frac{1}{\sqrt{6}}, 0, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right) = (0.5, 2, 0, -0.5)$$

$$v_2 = \frac{v'_2}{\|v'_2\|} = \left( \frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}, 0, -\frac{\sqrt{2}}{6} \right)$$

$$v'_3 = w_3 - \langle v_1, w_3 \rangle \cdot v_1 - \langle v_2, w_3 \rangle v_2 = \left( \frac{4}{9}, -\frac{2}{9}, 0, -\frac{4}{9} \right)$$

$$v_3 = \frac{v'_3}{\|v'_3\|} = \left( \frac{2}{3}, \frac{-1}{3}, 0, -\frac{2}{3} \right)$$

# Warum funktioniert der Algorithmus?

Vollständiger Beweis unter

<http://mathonline.wikidot.com/the-gram-schmidt-process>

- Die generierten Vektoren sind alle normal, da wir explizit normalisieren
- Die Vektoren spannen den gleichen Raum auf, da wir jeweils nur Linearkombinationen zur Konstruktion verwenden
- Orthogonalität am Beispiel:

$$v_1 = \left(\frac{1}{\sqrt{6}}, 0, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right), v_2 = \left(\frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}, 0, -\frac{\sqrt{2}}{6}\right), v_3 = \left(\frac{2}{3}, \frac{-1}{3}, 0, -\frac{2}{3}\right)$$

Und es gilt zum Beispiel

$$\langle v_1, v_3 \rangle = \frac{1}{\sqrt{6}} \cdot \frac{2}{3} + 0 \cdot \frac{-1}{3} + \frac{2}{\sqrt{6}} \cdot 0 + \frac{1}{\sqrt{6}} \cdot \left(-\frac{2}{3}\right) = 0$$

Wie wende ich Gram-Schmidt auf eventuell linear abhängige Vektoren an?

- 1 Hier bekomme ich manchmal  $v'_i$ , die der Nullvektor sind
- 2 Diese lösche ich einfach und überspringe diesen Vektor und gehe gleich zum nächsten über.

Leider gibt es bei der automatischen Ausführung des Gram-Schmidt-Verfahrens irgendwann ziemlich viele Rundungsfehler, so dass die resultierenden Vektoren nicht mehr orthonormal sind.

- 1 Modified Gram-Schmidt löst dies
- 2 Es gibt auch noch andere Orthogonalisierungsverfahren



Wende das Gram-Schmidt-Orthogonalisierungsverfahren auf die folgenden Vektoren im  $\mathbb{R}^2$  an  $w_1 = (3, 2)$  und  $w_2 = (2, 2)$ .

**Auflösung: Wikipedia**

[https://de.wikipedia.org/wiki/Gram-Schmidtsches\\_Orthogonalisierungsverfahren](https://de.wikipedia.org/wiki/Gram-Schmidtsches_Orthogonalisierungsverfahren)

- Linearkombinationen und lineare Unabhängigkeit erlaubt es uns, Unterräume und Basen zu definieren
- Ein endlicher Raum mit einem Skalarprodukt hat immer eine orthonormale Basis. Wir können diese Basis mit dem Gram-Schmidt-Orthonormalisierungsverfahren finden.

- Fischer: Lineare Algebra
- Ein sehr schönes Video zu Gram-Schmidt:  
<https://www.youtube.com/watch?v=sqwVvJL3WTw>