Euklidische Vektorräume

Irene Winkler & Nico Görnitz

Arbeitsgruppe Maschinelles Lernen

Organisatorisches

- 1. Tag: Mengen, Körper, (euklidische) Vektorräume
- 2. Tag: Lineare Abbildungen, Matrizen & Determinante
- 3. Tag: Polynome, Eigenwerte & Eigenvektoren
- 4. Tag: Zufallsvariablen, Verteilungen & Erwartungswerte

- 10.15-11.45: VL (und Abgabe der Hausaufgaben)
- 11.45-15.00: Eigenständige Bearbeitung der Übungsaufgaben
- 15.00-16.30: Besprechung der Übungsaufgaben

Grundlage für den benoteten Leistungsnachweis (2 SWS bzw. 3 SP) ist eine Klausur (90 Minuten), auf Wunsch stellen wir bei bestandener Klausur auch einen unbenoteten Leistungsnachweis aus. Voraussetzung für die Teilnahme an der Klausur ist das Erreichen von mindestens der Hälfte aller möglichen Punkte in den Hausaufgaben, die Ergebnisse in den Übungsaufgaben gehen nicht in die Note ein.

Maschinelles Lernen = Lernen aus Daten

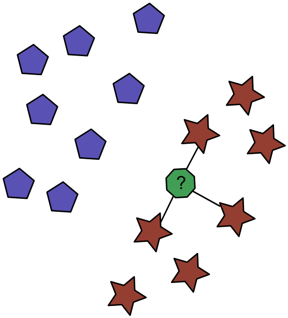
z.B. Bildklassifikation, Spamerkennung, DNA-Segmentierung, ...
Vorgehensweise:

1. Extraction von Merkmalen und Repräsentation als Vektoren
2. Anwendung von Methoden aus Linearer Algebra, Stochastik, ... um aus den Datenpunkten zu lernen
3. Anwendung des Gerlernten

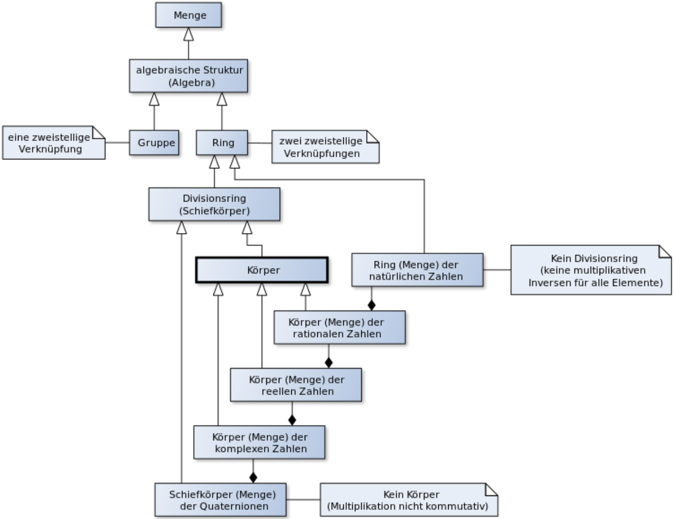
Bsp: k-Nearest-Neighbor-Klassifikation

- Gegeben sind n Datenpunkte $X = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ und dazugehörige Labels $\mathbf{y} = (y_1, y_2, \dots, y_n)$ mit $y_i \in \{-1, +1\}$
- Ein (nicht gesehener) Testdatenpunkt $\mathbf{v}$ erhält dann sein vorhergesagtes Label nach: $y = f(\mathbf{v}) = \operatorname{sign} \frac{1}{k} \sum_{i \in \mathcal{I}} y_i$
- Die Kardinalität der Indexmenge $\mathcal{I} \subset \{1, 2, \dots, n\}$ ist k und enthält genau die Indizes (aus der Trainingsmenge), für die gilt:

$$||\mathbf{v}-\mathbf{x}_i||_2^2 = ||\mathbf{v}||_2^2 - 2\langle \mathbf{v}, \mathbf{x}_i \rangle + ||\mathbf{x}_i||_2^2 \leq ||\mathbf{v}-\mathbf{x}_j||_2^2 \quad \forall i \in \mathcal{I}, j \in \{1, 2, \dots, n\} \backslash \mathcal{I}$$



Algebraische Strukturen



Gruppe

Definition. Eine *Gruppe* ist eine Menge G zusammen mit einer Verknüpfungen $+ : G \times G \mapsto G$ (Addition) in der gilt:

- Es gibt ein neutrales und ein inverses Element:
  - $e + a = a \quad \forall a \in G$
  - Zu jedem $a \in G$ gibt es ein inverses Element $-a \in G$
- Es gilt das Assoziativgesetz:
  $(a + b) + c = a + (b + c) \quad \forall a, b, c \in G$

Eine Gruppe G heisst *abelsch* (kommutativ), sollte ausserdem noch gelten: $a + b = b + a \quad \forall a, b \in G$

Körper

Definition. Ein *Körper* ist eine Menge K zusammen mit zwei Verknüpfungen $+ : K \times K \mapsto K$ (Addition) und $\cdot : K \times K \mapsto K$ (Multiplikation) in der gilt:

- $(K,+)$ ist eine abelsche Gruppe:
  - Die Addition ist assoziativ und kommutativ
  - Es existiert ein Element neutrales Element $0 \in K$
  - Zu jedem $a \in K$ gibt es ein inverses Element $-a \in K$
- $(K \backslash \{0\},\cdot)$ ist eine abelsche Gruppe:
  - Die Addition ist assoziativ und kommutativ
  - Es existiert ein Element neutrales Element $1 \in K$
  - Zu jedem $a \in K$ gibt es ein inverses Element $a^{-1} \in K$
- Es gelten die Distributivgesetze: $a \cdot (b + c) = a \cdot b + a \cdot c$

Vektorräume

Vektoren sind ...

- die Elemente eines Vektorraums
- können miteinander addiert werden
- können mit einem Skalar multipliziert werden

Beispiele: $\mathbb{R}^n$, Raum der Matrizen, Raum der reellwertigen Funktionen auf $[-1, 1]$, ...

Vektorräume

Definition. Ein *Vektorraum* ist eine Menge V zusammen mit einer Abbildung $+ : V \times V \mapsto V$ (Addition) und einer Abbildung $\cdot : K \times V \mapsto V$ (skalare Multiplikation) mit

- (V,+) ist eine abelsche Gruppe ist, d.h.
  - Die Addition ist assoziativ und kommutativ
  - Es existiert ein Element $0 \in V$ ("Nullvektor")
  - Zu jedem $\mathbf{v} \in V$ gibt es ein inverses Element
- Für alle $\mathbf{v}, \mathbf{w} \in V$ und $\lambda, \mu \in K$ gilt:
  - $\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$
  - $1\mathbf{v} = \mathbf{v}$
  - $\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$, $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$

Untervektorräume

Definition. Sei V ein reeller Vektorraum. Eine Teilmenge $U \subset V$ heisst *Untervektorraum von V* wenn $U \neq \emptyset$ und U abgeschlossen bzgl. der Addition und der Skalarmultiplikation ist, d.h.

$$\forall \mathbf{v}, \mathbf{w} \in U, \lambda \in \mathbb{R} : \mathbf{v} + \mathbf{w} \in U \text{ und } \lambda \mathbf{v} \in U.$$

Ist U ein Untervektorraum von V, so ist U ein Vektorraum.

Beispiele: alle n-Tupel, die ein bestimmtes Gleichungssystem lösen; Raum der differenzierbaren reelwertigen Funktionen auf $[-1, 1]$, ...

Lineare Hülle

Definition und Satz. Seien $\mathbf{v}_1 \ldots \mathbf{v}_n \in V$ Elemente eines Vektorraums V. Die Menge

$$\{\lambda_1 \mathbf{v_1} + \ldots + \lambda_n \mathbf{v_n} | \lambda_i \in \mathbb{R}\} \subset V$$

ist ein Unterverktorraum von V und heißt *lineare Hülle* bzw. der von $\mathbf{v_1}, \ldots, \mathbf{v_n}$ *aufgespannte / erzeugte Raum*.

Affine & Konvexe Hülle

Definition Seien $\mathbf{v}_1 \ldots \mathbf{v}_n \in V$ Elemente eines Vektorraums V. Die Menge

$$\{\lambda_1 \mathbf{v_1} + \ldots + \lambda_n \mathbf{v_n} | \lambda_i \in \mathbb{R} \wedge \sum_{i=1}^{n} \lambda_i = 1\} \subset V$$

heißt *affine Hülle*.

Definition Seien $\mathbf{v}_1 \ldots \mathbf{v}_n \in V$ Elemente eines Vektorraums V. Die Menge

$$\{\lambda_1 \mathbf{v_1} + \ldots + \lambda_n \mathbf{v_n} | 0 \leq \lambda_i \leq 1 \wedge \sum_{i=1}^{n} \lambda_i = 1\} \subset V$$

heißt *konvexe Hülle*.

Lineare Unabhängigkeit

Definition. Sie V ein reeller Vektorraum. Die Vektoren $\mathbf{v_1}, \ldots, \mathbf{v_n} \in V$ heißen *linear unabhängig*, wenn

$$\lambda_1 \mathbf{v_1} + \ldots + \lambda_n \mathbf{v_n} = 0 \Rightarrow \lambda_1 = \ldots = \lambda_n = 0$$

Äquivalent dazu: wenn sich keiner der Vektoren als Linearkombination der anderen darstellen lässt.

Basis & Dimension

Definition. Sei V ein reeller Vektorraum. Vektoren $\mathbf{v_1}, \ldots, \mathbf{v_n} \in V$ bilden eine *Basis von V*, wenn sie linear unabhängig sind und wenn jedes $\mathbf{v} \in V$ Linearkombination der $\mathbf{v_i}$ ist.

Satz und Definition. Je zwei endliche Basen eines reellen Vektorraumes V haben gleiche Länge. Diese Länge ist die *Dimension* des Vektorraums.

Satz. Ist $\mathbf{v_1}, \ldots, \mathbf{v_n}$ eine Basis von V, dann gibt es zu jedem $\mathbf{v} \in V$ *genau eine* Linearkombination.

Euklidische Vektorräume

- Zusätzliche Struktur auf Vektorräumen: Länge, Abstand, Winkel
- $\Rightarrow$ Norm, Skalarprodukt

Norm - Länge eines Vektors

Definition. Sei V ein reeller Vektorraum. Eine *Norm* ist eine Abbildung $\|\cdot\| : V \mapsto \mathbb{R}^+$ mit folgende Eigenschaften:

1. $\|\mathbf{v}\| \geq 0$ und $\|\mathbf{v}\| = 0 \Leftrightarrow \mathbf{v} = 0$
2. $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ (Dreiecksungleichung)
3. $\|\lambda \mathbf{v}\| = |\lambda| \|\mathbf{v}\|$

für alle $\mathbf{v}, \mathbf{w} \in V, \lambda \in \mathbb{R}$

Beispiele auf $\mathbb{R}^n$:

- 2-Norm $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$
- 1-Norm $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$
- max-Norm $\|\mathbf{x}\|_\infty = \max_{i=1}^n |x_i|$

Skalarprodukt

Standardskalarprodukt im $\mathbb{R}^2$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = x_1 y_1 + x_2 y_2$$

Was bedeutet das?

- Es gilt $\langle \mathbf{x}, \mathbf{y} \rangle = \cos(\angle \mathbf{x}, \mathbf{y}) \cdot \|\mathbf{x}\| \cdot \|\mathbf{y}\|$
- Für $\|\mathbf{x}\| = 1$: Länge der Projektion von $\mathbf{y}$ auf $\mathbf{x}$

Skalarprodukt

Definition. Sei V ein reeller Vektorraum. Unter einem *Skalar - oder innerem Produkt auf V* versteht man eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \mapsto \mathbb{R}$ mit folgende Eigenschaften:

1. bilinear: $\langle \lambda \mathbf{v}, \mathbf{w} \rangle = \lambda \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \lambda \mathbf{w} \rangle$, $\langle \mathbf{v} + \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$, $\langle \mathbf{v}, \mathbf{w} + \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle$
2. symmetrisch: $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$
3. positiv definit: $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ und $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Leftrightarrow \mathbf{v} = 0$

für alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ und $\lambda \in \mathbb{R}$

Satz (Cauchy-Schwarzsche Ungleichung). Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum. Dann gilt

$$\forall \mathbf{v}, \mathbf{w} \in V : | \langle \mathbf{v}, \mathbf{w} \rangle | \leq \|\mathbf{v}\| \|\mathbf{w}\|$$

Skalarprodukt: Beispiele

- Standardskalarprodukt auf dem $\mathbb{R}^n$:

$$\langle \mathbf{x}, \mathbf{y} \rangle := x_1 y_1 + \ldots + x_n y_n = \mathbf{x}^T \mathbf{y} \qquad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^n)$$

- Auf $\mathbb{R}^n$ mit $\lambda_1, \ldots, \lambda_n \geq 0$:

$$\langle \mathbf{x}, \mathbf{y} \rangle := \lambda_1 x_1 y_1 + \ldots + \lambda_n x_n y_n \qquad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^n)$$

- Auf $\mathcal{F} = L_2(X) = \{f : X \mapsto \mathbb{R} \mid \int_X f(x)^2 dx < \infty\}$, dem Raum der quadratisch integrierbaren Funktionen auf einem kompakten $X \subset \mathbb{R}^n$

$$\langle f, g \rangle := \int_X f(x) g(x) dx \qquad (f, g \in \mathcal{F})$$

Skalarprodukt

Definition. Unter einem *euklidischen Vektorraum* versteht man ein Paar $(V, \langle\cdot,\cdot\rangle)$ bestehend aus einem reellen Vektorraum V und einem Skalarprodukt $\langle\cdot,\cdot\rangle$ auf V.

Hilbertraum: Ein reeller oder komplexer Vektorraum zusammen mit einem Skalarprodukt, der vollständig bezüglich der durch das Skalarprodukt induzierten Metrik ist (d.h. in dem jede Cauchy-Folge konvergiert).

Euklidische Norm

Satz und Definition. Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum. Das Skalarprodukt induziert eine Norm $\|\mathbf{v}\| := \sqrt{\langle\mathbf{v},\mathbf{v}\rangle}$ ($\mathbf{v} \in V$), die *euklidische Norm* genannt wird.

Orthogonale Vektoren

Definition. Zwei Vektoren $\mathbf{v}, \mathbf{w}$ eines euklidischen Vektorraumes heißen *orthogonal zueinander* $(\mathbf{v} \perp \mathbf{w})$ wenn $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.

Definition. Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in einem euklidischen Vektorraum heißen *orthonormal* oder *Orthonormalsystem* wenn $\|\mathbf{v}_i\| = 1$ für alle i und $\mathbf{v}_i \perp \mathbf{v}_j$ für $i \neq j$.

Orthonomalbasis

Definition. Eine Basis eines euklidischen Vektorraums aus orthonormalen Vektoren wird *Orthonormalbasis* genannt.

Satz. Ist $\mathbf{v}_1, \ldots, \mathbf{v}_n$ eine Orthonomalbasis eines euklidischen Vektorraums $(V, \langle \cdot, \cdot \rangle)$, so gilt für jedes $\mathbf{x} \in V$:

$$\mathbf{x} = \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{v}_i \rangle \, \mathbf{v}_i$$

Geg: Eine Basis. Ziel: Orthogonale Basis. Methode: Gram-Schmidt Orthogonalisierungsverfahren

Orthogonale Projektion

Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum und U ein r -dimensionaler Untervektorraum mit einer Orthonormalbasis $\mathbf{u}_1, \ldots, \mathbf{u}_r$.
Die orthogonale Projektion eines Vektors $\mathbf{v} \in V$ auf U ist gegeben durch

$$p(\mathbf{v}) = \sum_{i=1}^{r} \langle \mathbf{v}, \mathbf{u}_i \rangle \, \mathbf{u}_i$$

.