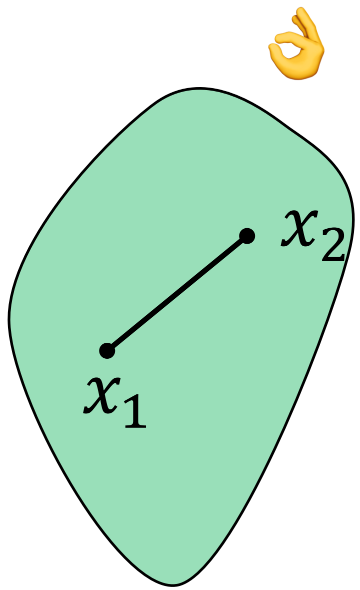
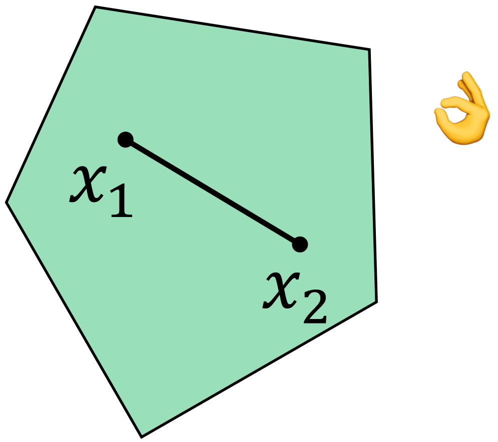
2b. Grundlagen

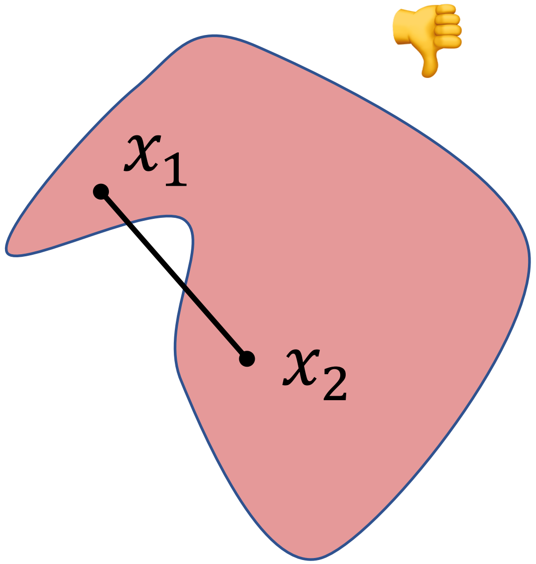
Konvexe Mengen

Optimierung SoSe 2020

Dr. Alexey Agaltsov







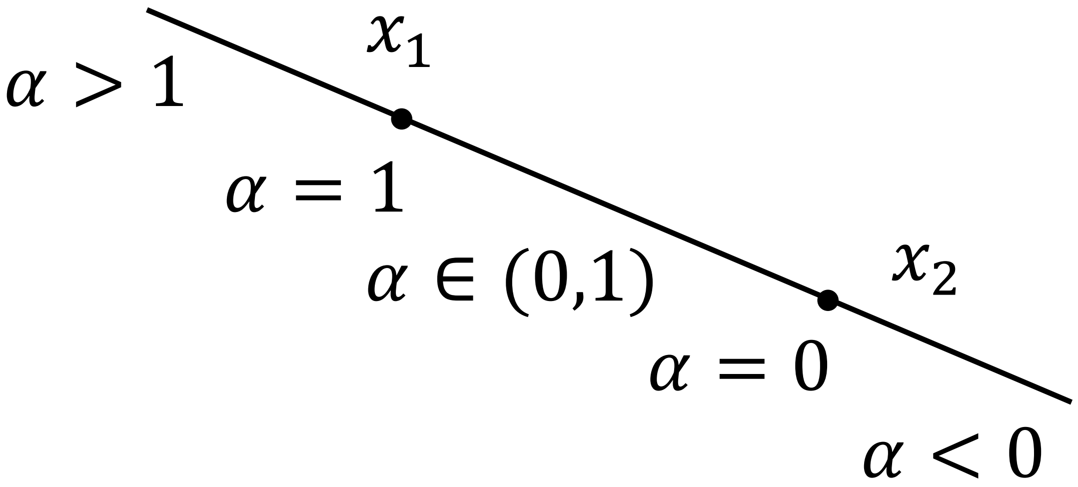
Plan

- Affine Mengen
- Konvexe Mengen
- Projektion

Geraden und Strecken

- Die Gerade durch die Punkte $x_1 \neq x_2$ aus $\mathbb{R}^n$ enthält die Punkten

$$y = \alpha x_1 + (1-\alpha)x_2, \quad \alpha \in \mathbb{R}$$



- Die Strecke zwischen x_1 und x_2 entspricht den Punkten mit $\alpha \in [0,1]$ $\alpha = 1$

Affine Mengen

- Eine Menge $K \subseteq \mathbb{R}^n$ heißt affin, falls die Gerade durch jede $x_1 \neq x_2$ in K liegt ebenfalls in K:

$$\alpha x_1 + (1-\alpha) x_2 \in K \quad \forall (x_1, x_2 \in K, \alpha \in \mathbb{R})$$

- Als Affinkombination von $x_1, \ldots, x_k \in K$ bezeichnen wir einen Punkt:

$$y = \alpha_1 x_1 + \cdots + \alpha_k x_k$$
$$\alpha_1 + \cdots + \alpha_k = 1$$

- Mit einer Induktion kann man zeigen, dass K affin ist ist gdw. K alle affine Kombinationen seiner Punkte enthält

Beispiel: Lineare Gleichungssysteme

$$K = \{x: Ax = b\}$$
$$A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$$

Seien x_1, x_2 so, dass $Ax_1 = b,\ Ax_2 = b$

$$A(\alpha x_1 + (1-\alpha)x_2) = \alpha \underbrace{Ax_1}_{b} + (1-\alpha)\underbrace{Ax_2}_{b} = b$$

- Die Lösungsmenge eines Systems der Lineargleichungen ist affin

Affine Menge und lineare Unterräume

- Sei K eine affine Menge und $x_0 \in K$. Setze

$$V = K - x_0 := \{x - x_0 : x \in K\}$$

V ist ein linearer Unterraum:

Seien $v_1, v_2 \in V$ und $\alpha, \beta \in \mathbb{R}$

$$\alpha v_1 + \beta v_2 + x_0 = \alpha\underbrace{(v_1 + x_0)}_{\in K} + \beta\underbrace{(v_2 + x_0)}_{\in K} + (1 - \alpha - \beta)\underbrace{x_0}_{\in K}$$

$$\alpha v_1 + \beta v_2 + x_0 \in K$$

Also $\alpha v_1 + \beta v_2 \in K - x_0 = V$

Affine Menge und lineare Unterräume

K ist affin $\longleftrightarrow$ $K = V + x_0, \quad x_0 \in K$

V linearer Unterraum (unabh. von x_0)

- Als Dimension $\dim K$ von K bezeichnen wir die Dimension von V

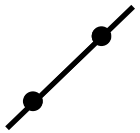
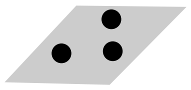
- Affine Mengen entsprechen den Folgenden geometrischen Objekten:

$$\dim K = 0: \text{ Punkt}$$
$$\dim K = 1: \text{ Gerade}$$
$$\dim K = 2: \text{ Ebene}$$
$$\dim K \geq 3: \text{ Hyperebene}$$

Affine Hülle

- Als affine Hülle von $K \subseteq \mathbb{R}^n$ bezeichnen wir die Menge:

$$\mathbf{aff}(K) \coloneqq \{\alpha_1 x_1 + \cdots + \alpha_N x_N \mid x_1, \dots, x_N \in K,\ \alpha_1 + \cdots + \alpha_N = 1\}$$

$\mathbf{aff}(\bullet) = \bullet$ $\qquad$ $\mathbf{aff}(\bullet\ \bullet) =$  $\qquad$ $\mathbf{aff}(\bullet\ \bullet\ \bullet) =$ 

- Äquivalent, $\mathbf{aff}(K)$ ist die minimale affine Menge die K enthält:

$$\forall (A \text{ affin}, K \subseteq A)\ \ \mathbf{aff}(K) \subseteq A$$

- Also ist K affin gdw. $K = \mathbf{aff}(K)$

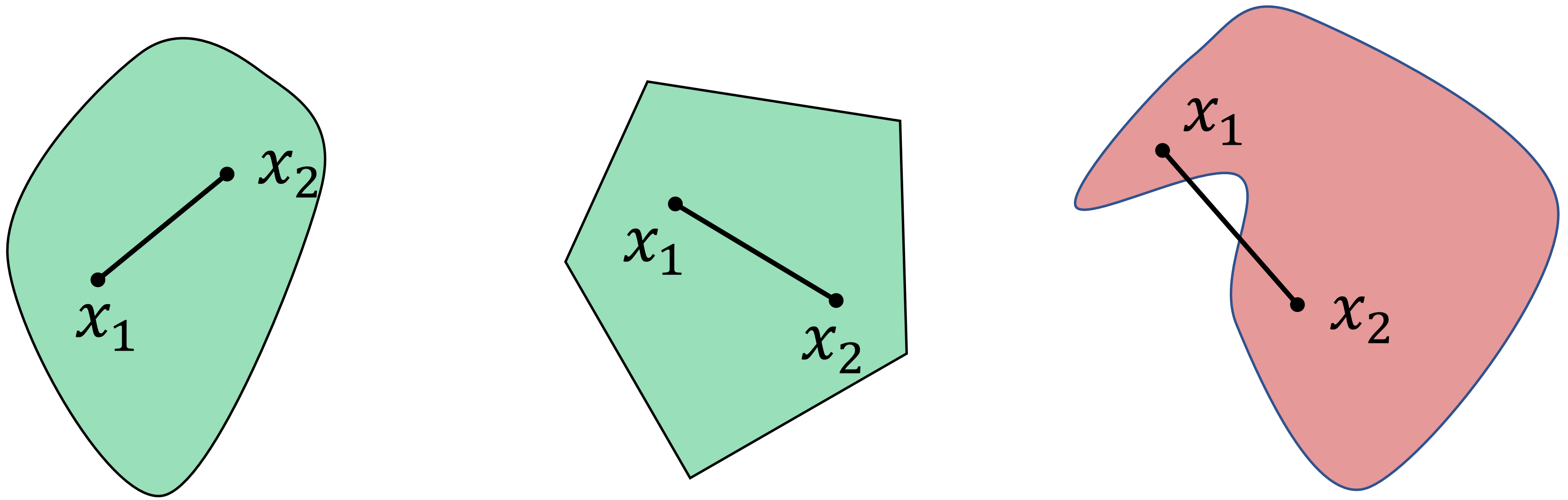
Plan

- Affine Mengen
- Konvexe Mengen
- Projektionssatz

Konvexe Mengen

- Eine Menge $K \subseteq \mathbb{R}^n$ heißt konvex falls für alle $x_1 \neq x_2$ in K die Strecke zwischen x_1 und x_2 ebenfalls K gehört:

$$\alpha x_1 + (1-\alpha)x_2 \in K \quad \forall x_1, x_2 \in K, \alpha \in [0,1]$$



Konvexkombinationen

- Als Konvexkombination von Punkten $x_1, \dots, x_k$ bezeichnen wir jeden Punkt y mit Darstellung:

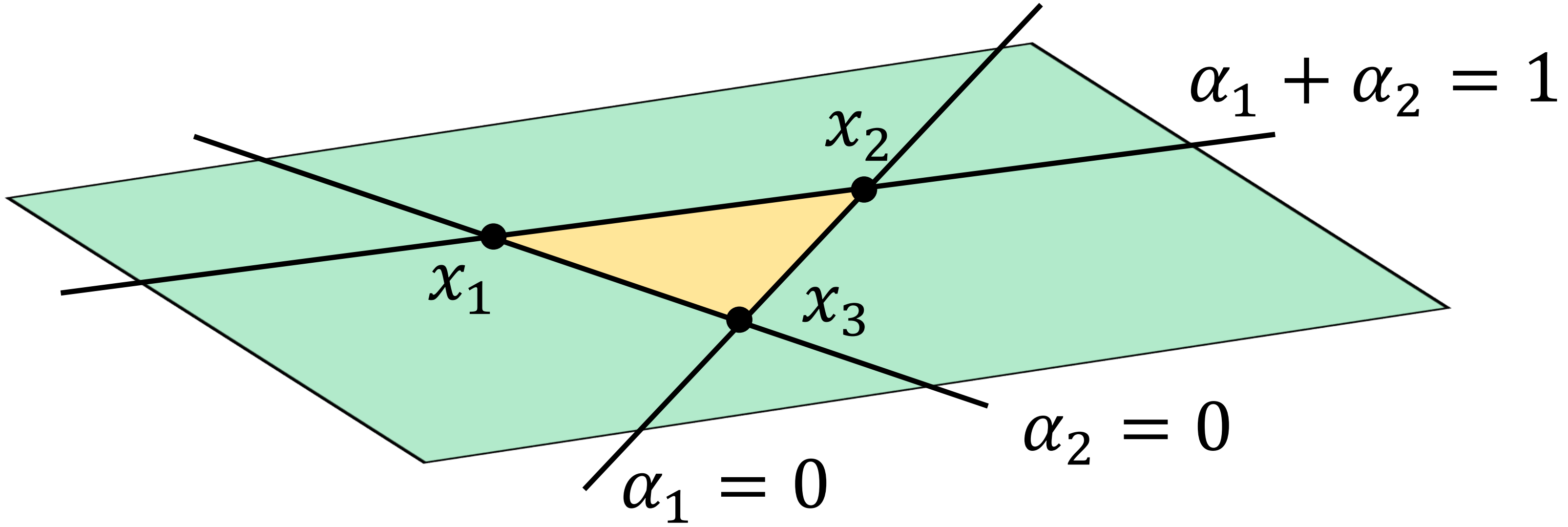
$$y = \alpha_1 x_1 + \cdots + \alpha_k x_k$$

$$\alpha_1 + \cdots + \alpha_k = 1, \qquad \alpha_i \geq 0, i = 1, \dots, k$$

- Mit einer Induktion kann man zeigen, dass eine Menge K konvex ist gdw. K alle Konvexkombinationen ihre Punkte enthält

Beispiel

$$y = \alpha_1 x_1 + \alpha_2 x_2 + (1 - \alpha_1 - \alpha_2) x_3$$



Affinkombinationen von x_1, x_2, x_3 entsprechen der Ebene durch x_1, x_2, x_3

Konvexkombinationen von x_1, x_2, x_3 liegen im Dreieck

Konvexe Hülle

- Als konvexe Hülle von $K \subseteq \mathbb{R}^n$ bezeichnen wir die Menge:

$$\mathbf{conv}(K) := \{\alpha_1 x_1 + \cdots + \alpha_k x_k \mid x_i \in K, \alpha_i \geq 0, i = 1, \ldots, k, \alpha_1 + \cdots + \alpha_k = 1\}$$

- Äquivalent, $\mathbf{conv}(K)$ ist die minimale konvexe Menge, die K enthält

$$\forall (C \text{ konvex}, K \subseteq C) \;\; \mathbf{conv}(K) \subseteq C$$

- Also ist K konvex gdw. $K = \mathbf{conv}(K)$

Beispiel: Konvexe Matrizenräume

$$\mathbb{S}^n := \{X \in \mathbb{R}^{n\times n} : X^T = X\}$$

$$\mathbb{S}^n_{\succcurlyeq} := \{X \in \mathbb{R}^{n\times n} : X^T = X, X \succcurlyeq 0\}$$

$$\mathbb{S}^n_{\succ} := \{X \in \mathbb{R}^{n\times n} : X^T = X, X \succ 0\}$$

Wir werden zeigen, dass $\mathbb{S}^n$, $\mathbb{S}^n_{\succcurlyeq}$, $\mathbb{S}^n_{\succ}$ konvexe Mengen sind

Beweis

Behauptung: $\mathbb{S}^n := \{X \in \mathbb{R}^{n\times n} : X^T = X\}$ ist konvex

Seien $A, B \in \mathbb{S}^n$, $\alpha \in (0,1)$

$$Z := \alpha A + (1-\alpha)B$$

$$Z^T = \alpha \underbrace{A^T}_{A} + (1-\alpha)\underbrace{B^T}_{B} = Z$$

$$Z \in \mathbb{S}^n$$

Beweis

Behauptung: $\mathbb{S}^n_{\succ} := \{X \in \mathbb{S}^n : X \succ 0\}$ ist konvex

Seien $A, B \in \mathbb{S}^n_{\succ}$, $\alpha \in (0,1)$

$$Z = \alpha A + (1-\alpha)B$$

$$\forall (x \in \mathbb{R}^n : x \neq 0)$$

$$x^T Z x = \alpha \underbrace{x^T A x}_{>0} + (1-\alpha)\underbrace{x^T B x}_{>0} > 0$$

$$Z \in \mathbb{S}^n_{\succ}$$

Ähnlich für $\mathbb{S}^n_{\succeq}$

Lemma 2.6. Durchschnitt konvexer Mengen

Seien $K_i, i \in I$ konvex

$$K := \cap_{i \in I} K_i$$

Dann ist K konvex

Beweis

Seien $x_1, x_2 \in K, \alpha \in (0,1)$

$y = \alpha x_1 + (1-\alpha)x_2$

$x_1, x_2 \in K_i \quad \forall i \in I$

K_i ist konvex

$y \in K_i \quad \forall i \in I$

$y \in K$

K ist konvex

$\square$

Plan

- Affine Mengen
- Konvexe Mengen
- Projektionssatz

Projektion auf eine Menge

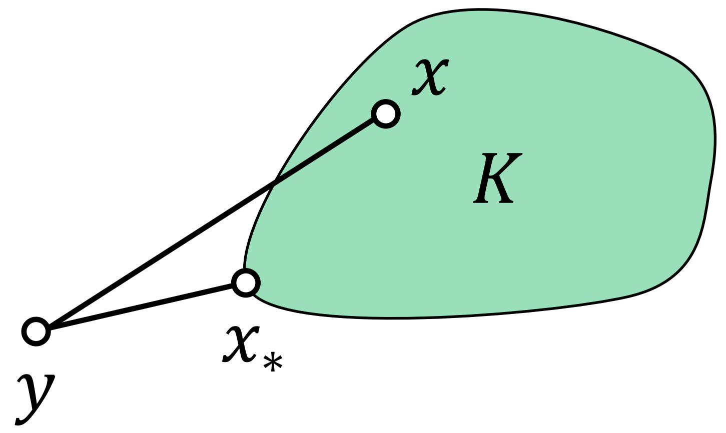
- Sei $\langle\cdot,\cdot\rangle$ ein beliebiges Skalarprodukt auf $\mathbb{R}^n$ und $\|\cdot\|$ die zugehörige Norm

- Als Abstand zwischen $K \subseteq \mathbb{R}^n$ und $y \in \mathbb{R}^n$ bezeichnen wir:

$$\mathbf{dist}(y, K) = \inf\{\|x - y\| : x \in K\}$$

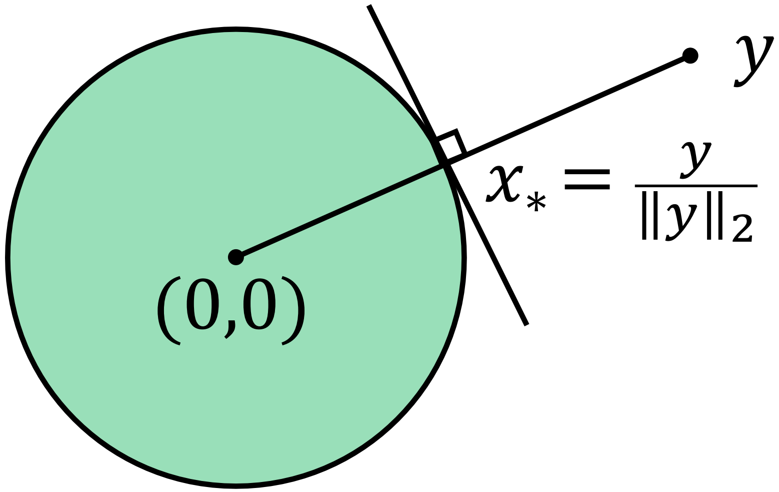
- Wir bezeichnen einen Punkt $x_* \in K$ als Projektion von y auf K falls:

$$\|y - x_*\| = \mathbf{dist}(y, K)$$



Eindeutige Projektion

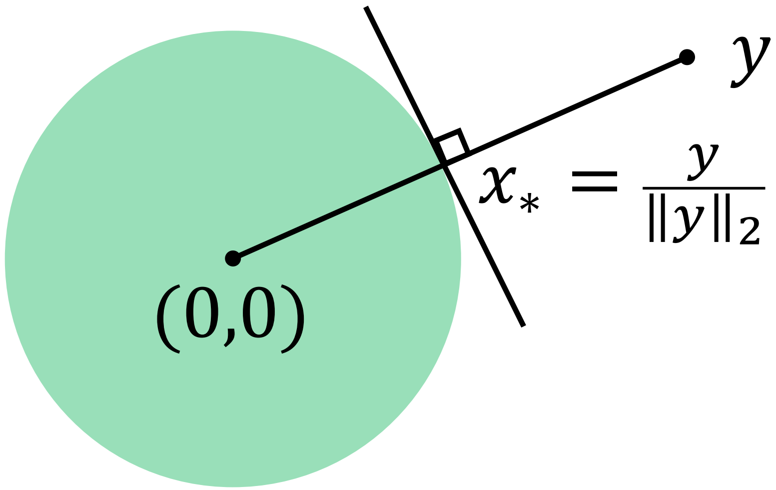
$$\|x\| = \|x\|_2 := \sqrt{x_1^2 + x_2^2} \text{ für } x = (x_1, x_2) \in \mathbb{R}^2$$



$$K = \{x \in \mathbb{R}^2 : \|x\|_2 \le 1\}$$

eindeutige Projektion $x_* = \frac{y}{\|y\|_2}$

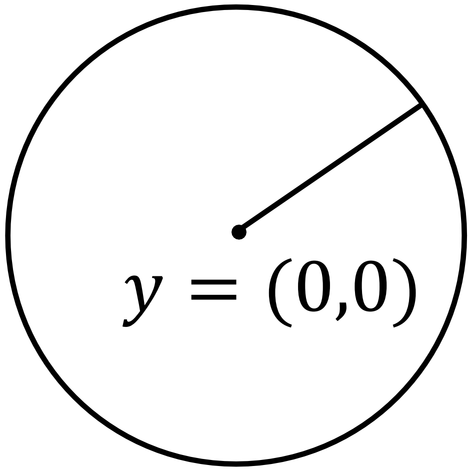
Keine Projektion



$$K = \{x \in \mathbb{R}^2 : \|x\|_2 < 1\}$$

keine Projektion

Mehrere Projektionen



$$K = \{x : \|x\|_2 = 1\}$$

Jeder Punkt $x \in K$ ist eine Projektion

Lemma 2.7. Existenz der Projektion

- Sei $K \subseteq \mathbb{R}^n$ nichtleer und abgeschlossen und sei $y \in \mathbb{R}^n$
- Dann existiert zumindest eine Projektion $x_* \in K$ von y auf K

Beweis

Setze $f(x) = \begin{cases} \|y - x\| & x \in K \\ \infty & sonst \end{cases}$

Satz 2.4

$\mathbf{dom}(f) = K$ ist abgeschlossen und $f \in C(K)$

f ist unterhalbstetig

Entweder ist K beschränkt oder $f(x) \to \infty$ für $\|x\| \to \infty$

Satz 2.5

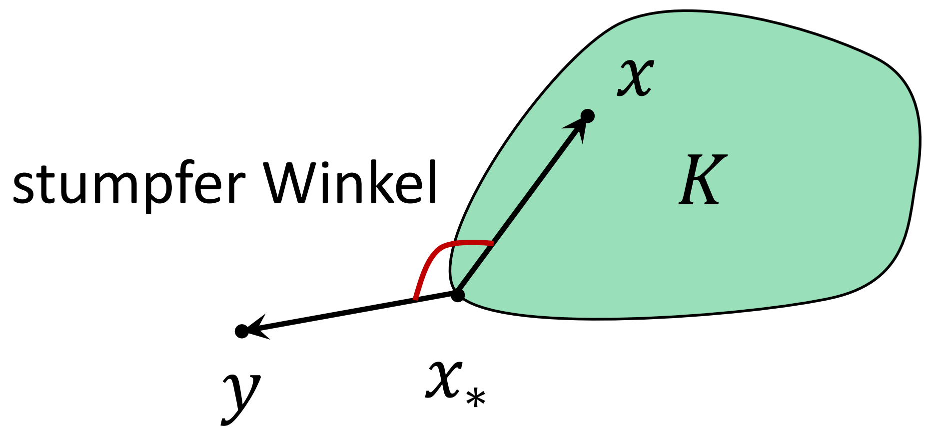
$$\exists x_* \in K\colon \underbrace{f(x_*)}_{\|y - x_*\|} = \underbrace{\min_{x \in K} f(x)}_{\mathbf{dist}(y, K)}$$

$\square$

Satz 2.8. Projektionssatz

- Sei $K \subseteq \mathbb{R}^n$ nichtleer, abgeschlossen und konvex und sei $y \in \mathbb{R}^n$
- Dann existiert eine eindeutige Projektion $x_* \in K$ von y auf K
- Außerdem ist $x_* \in K$ genau dann die Projektion von y, wenn:

$$\langle y - x_*, x - x_* \rangle \leq 0 \quad \forall x \in K$$



Beweis: Charakterisierung

Lemma 2.7 ⇒ Existenz

Seien x_* ist eine Projektion von y auf K

$x \in K, t \in [0,1]$

$$\|y - x_*\|^2 \leq \Big\|y - \overbrace{\big(x_* + t(x - x_*)\big)}^{z \in K}\Big\|^2$$

$$= \|y - x_*\|^2 - 2t\langle y - x_*, x - x_*\rangle + t^2\|x - x_*\|^2$$

$$-2t\langle y - x_*, x - x_*\rangle + t^2\|x - x_*\|^2 \geq 0$$

dividiere durch t

$t \to 0$

$$\langle y - x_*, x - x_*\rangle \leq 0$$

Beweis: Eindeutigkeit

Seien x_1^*, x_2^* zwei Projektionen von y auf K

Charakterisierung

$$\langle y - x_2^*, x_1^* - x_2^* \rangle \leq 0$$

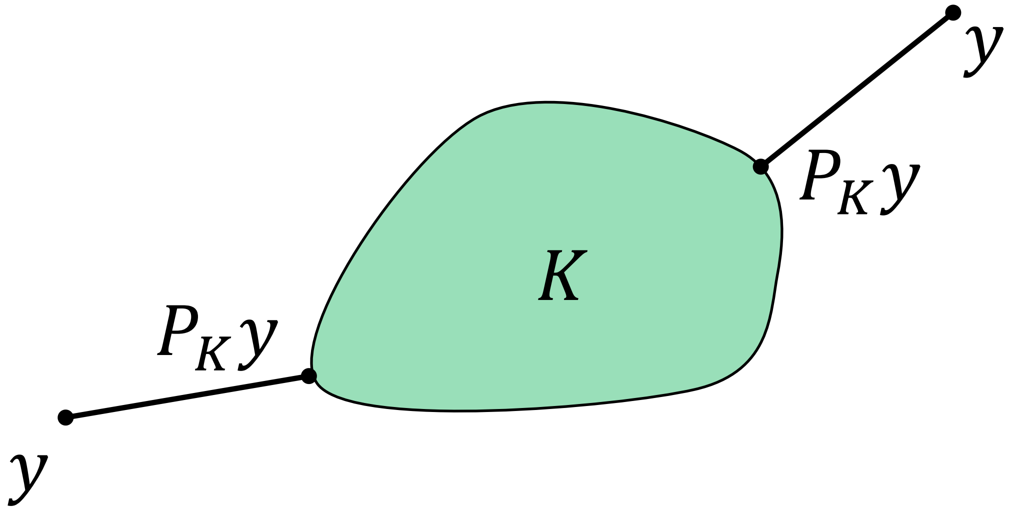
$$\underbrace{\langle y - x_1^*, x_2^* - x_1^* \rangle}_{\langle x_1^* - y,\, x_1^* - x_2^* \rangle} \leq 0$$

addiere

$\|x_2^* - x_1^*\|^2 \leq 0$ also $x_1^* = x_2^*$

Projektor

Sei $K \subseteq \mathbb{R}^n$ nichtleer und konvex, so definieren wir den Projektor auf K als:



$$P_K : \mathbb{R}^n \to K$$

$$P_K(y) = \operatorname{argmin}_{x \in K} \|y - x\|$$

Satz 2.9. Projektor auf einen Unterraum

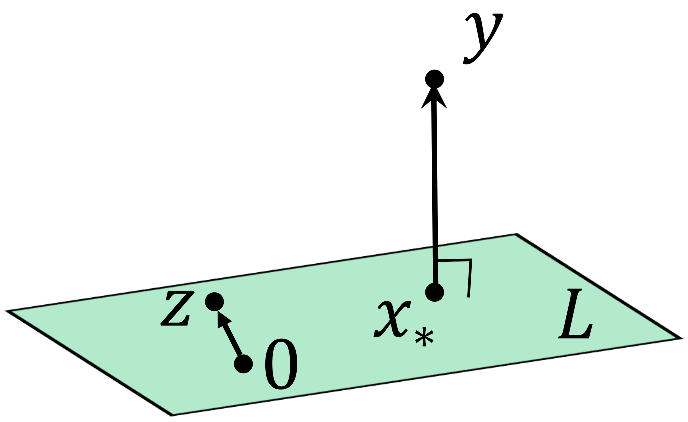
Sei $L \subseteq \mathbb{R}^n$ ein linearer Unterraum. Dann gilt:

- P_L ist eine lineare Abbildung $\quad P_L(\alpha x + \beta y) = \alpha P_L(x) + \beta P_L(y)$
- P_L ist symmetrisch:

$$\langle x, P_L y\rangle = \langle P_L x, y\rangle \quad \forall (x, y \in L)$$

- Sei $x_* \in L$. Dann gilt $x_* = P_L y$ gdw.

$$\langle y - x_*, z\rangle = 0 \quad \forall z \in L$$



$y - x_* \in L^\perp$

Beweis: Charakterisierung

Behauptung: Sei $x_* \in L$. Dann $x_* = P_L y$ gdw. $\langle y - x_*, z \rangle = 0 \;\forall z \in L$

Sei $x_* \in L$ Satz 2.8

$$x_* = P_L y \quad \text{gdw. } \langle y - x_*, \underbrace{x - x_*}_{z} \rangle \le 0 \quad \forall x \in L$$

$x \in L \Leftrightarrow z \in L$

$$\langle y - x_*, z \rangle \le 0 \quad \forall z \in L$$

$z \mapsto -z$

$$\langle y - x_*, z \rangle \ge 0 \quad \forall z \in L$$

$$\langle y - x_*, z \rangle = 0 \quad \forall z \in L$$

Beweis: Linearität

Seien $y_1, y_2 \in L, \alpha_1, \alpha_2 \in \mathbb{R}$

$$x_1^* = P_L y_1, x_2^* = P_L y_2$$

$$\langle y_1 - x_1^*, z \rangle = 0 \quad \forall z \in L$$

$$\langle y_2 - x_2^*, z \rangle = 0 \quad \forall z \in L$$

$(1)\times\alpha_1 + (2)\times\alpha_2$

$$\langle y_\alpha - x_\alpha^*, z \rangle = 0 \quad \forall z \in L$$

$$x_\alpha^* = \alpha_1 x_1^* + \alpha_2 x_2^*$$

$$y_\alpha = \alpha_1 y_1 + \alpha_2 y_2$$

Charakterisierung

$$x_\alpha^* = P_L y_\alpha$$

Beweis: Symmetrie

Seien $y_1, y_2 \in L$, $\alpha_1, \alpha_2 \in \mathbb{R}$

$$x_1^* = P_L y_1, x_2^* = P_L y_2$$

Charakterisierung

$$\langle y_1 - x_1^*, x_2^* \rangle = 0$$

$$\langle y_2 - x_2^*, -x_1^* \rangle = 0$$

addiere

$$\langle y_1, \underbrace{x_2^*}_{P_L y_2} \rangle = \langle y_2, \underbrace{x_1^*}_{P_L y_1} \rangle$$

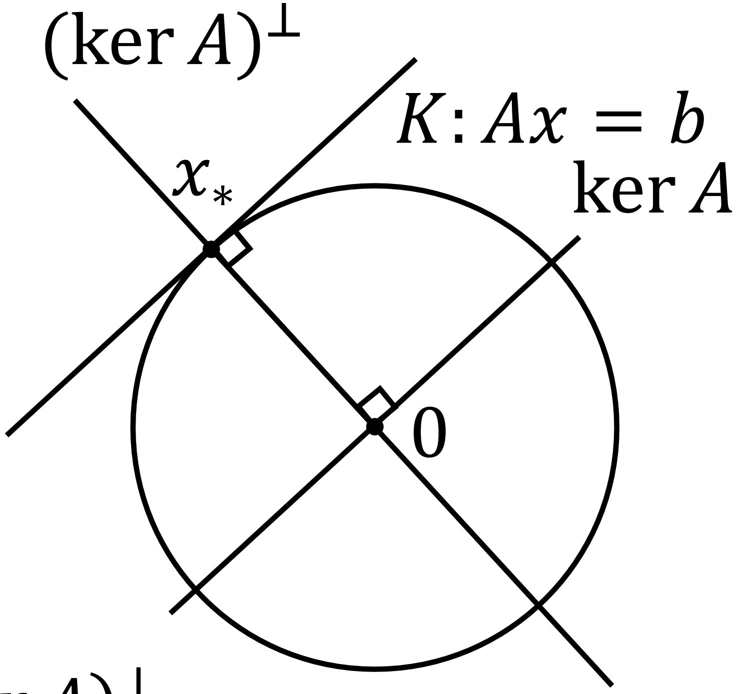
$\square$

Beispiel 2.10

$$\text{Minimiere } \|x\|_2$$

$$\text{u.d.N. } Ax = b \qquad A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$$

$$\mathrm{rang}(A) = m \leq n$$



- Eine optimale Lösung x_* ist die Projektion von 0 auf $K = \{x \in \mathbb{R}^n : Ax = b\}$
- Da K abgeschlossen und konvex ist, ist x_* eindeutig bestimmt
- Aufgabe. Beweisen Sie die Formel $\{x_*\} = K \cap (\ker A)^\perp$

Zusammenfassung

- Affine Mengen
- Konvexe Mengen
- Projektionssatz

Nächstes Video

- 2c. Grundlagen: Konvexe Funktionen