不動点理論と最適化

明治大学理工学部情報科学科
飯塚秀明

連続最適化および関連分野に関する
夏季学校

2021年8月23日, 24日

明治大学
MEIJI UNIVERSITY

連続最適化および関連分野に関する夏季学校

- ○統計数理研究所　田中未来 先生
- ○慶應義塾大学　成島康史 先生
- ○関西大学　檀寛成 先生
- ○京都大学　佐藤寛之 先生

この場を借りて感謝申し上げます。

本講義の目的

明治大学 MEIJI UNIVERSITY

- □不動点とは何か？
- □不動点を見つけるとなぜ嬉しいのか？

という素朴な問に対する回答例の提示

- ■講義 1: ヒルベルト空間、便利な不等式
- ■講義 2: 不動点, 非拡大写像, 最適化
- ■講義 3: 不動点近似法

参考文献

□高橋渉著 非線形･凸解析学入門

高橋渉先生　東京工業大学名誉教授
1944.01.22～2020.11.19
ご専門：非線形解析
論文数：500以上　引用数：10000以上
博士学生：43名

不動点理論を学んで良かったこと

- 高橋先生の教え：数学の問題を**高い立場**で捉える
- 最適化や工学（機械学習）の研究者と異なる視点

均衡問題

↓

不動点問題

↓

最適化

⇒

深層学習に
現れる
非凸最適化
や均衡問題

- 有限次元ユークリッド空間とヒルベルト空間
- よく使う不等式・等式

3次元ユークリッド空間

$$\mathbb{R}^3$$

(I) $x, y \in \mathbb{R}^3, \alpha \in \mathbb{R}$,

和:

$$x + y \in \mathbb{R}^3$$

スカラー倍:

$$\alpha x \in \mathbb{R}^3$$

を定義すると、性質①から⑧を満たす。

$\mathbb{R}^3$: ベクトル空間

$\cup$

x : ベクトル

ヒルベルト空間

H : 集合

(I) $x, y \in H, \alpha \in \mathbb{R}$,

和:

$$x + y \in H$$

スカラー倍:

$$\alpha x \in H$$

を定義して、性質①から⑧を満たすとする。

H : 線形空間

$\cup$

x : ベクトル

$$x = (a_1, a_2, a_3), y = (b_1, b_2, b_3)$$

$$x + y := (a_1 + b_1, a_2 + b_2, a_3 + b_3),$$

$$\alpha x := (\alpha a_1, \alpha a_2, \alpha a_3)$$

$x, y, z \in \mathbb{R}^3, \alpha, \beta \in \mathbb{R} \qquad x = (a_1, a_2, a_3)$

① $(x + y) + z = x + (y + z)$

② $x + y = y + x$

③ $x + 0 = x \quad 0 = (0, 0, 0)$

④ $x + (-x) = 0 \qquad -x = (-a_1, -a_2, -a_3)$

⑤ $(\alpha\beta)x = \alpha(\beta x)$

⑥ $1x = x$

⑦ $\alpha(x + y) = \alpha x + \alpha y$

⑧ $(\alpha + \beta)x = \alpha x + \beta x$

3次元ユークリッド空間

$\mathbb{R}^3$

(I) 和・スカラー倍を定義すると、
性質①から⑧を満たす。

(II) 内積

$$\langle x, y\rangle := a_1 b_1 + a_2 b_2 + a_3 b_3$$

を定義すると、性質①から④を満たす。

① $\langle x, x\rangle \geq 0; \langle x, x\rangle = 0 \iff x = 0$
② $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
③ $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$
④ $\langle x, y\rangle = \langle y, x\rangle$

(III) ノルム $\|x\| := \sqrt{\langle x, x\rangle}$
を定義すると、性質①から③を満たす。
① $\|x\| \geq 0; \|x\| = 0 \iff x = 0$
② $\|\alpha x\| = |\alpha|\|x\|$
③ $\|x+y\| \leq \|x\| + \|y\|$

ヒルベルト空間

H : 集合

(I) 和・スカラー倍を定義して、
性質①から⑧を満たすとする。

H: 線形空間

(II) 内積

$\langle x, y\rangle \in \mathbb{R}$

を定義して、性質①から④を満たすとする。

H: 内積空間

Cauchy-Schwarz の不等式

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

(III) ノルム $\|x\| := \sqrt{\langle x, x\rangle}$
を定義すると性質①から③を満たす。

H : ノルム空間

Cauchy-Schwarz の不等式が成り立つことの証明

内積空間 *H* 上の任意の *x, y* に対して

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

$$\boxed{\|x\|^2 := \langle x, x\rangle}$$

よく使う等式：２乗ノルム展開

$$\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$$

$$\because \|x+y\|^2 = \langle x+y, x+y\rangle$$

$$\overset{②}{=} \langle x, x+y\rangle + \langle y, x+y\rangle$$

$$\overset{④, ②}{=} \langle x, x\rangle + \langle y, x\rangle + \langle x, y\rangle + \langle y, y\rangle$$

$$\overset{④}{=} \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$$

②
$\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$

④
$\langle x, y\rangle = \langle y, x\rangle$

Cauchy-Schwarz の不等式が成り立つことの証明

内積空間 *H* 上の任意の *x*, *y* に対して

$$|\langle x, y\rangle| \leq \|x\|\|y\| \qquad \boxed{\|x\|^2 := \langle x, x\rangle}$$

$\because$ $t \in \mathbb{R}$

$$0 \underset{①}{\leq} \|tx + y\|^2 = \|tx\|^2 + 2\langle tx, y\rangle + \|y\|^2$$

$$\overset{③, ④}{=} t^2\|x\|^2 + 2t\langle x, y\rangle + \|y\|^2$$

③: $\boxed{\langle \alpha x, y\rangle = \alpha\langle x, y\rangle}$

$$\|tx\|^2 = \langle tx, tx\rangle \overset{③}{=} t\langle x, tx\rangle$$

④: $\boxed{\langle x, y\rangle = \langle y, x\rangle}$

$$\overset{④}{=} t\langle tx, x\rangle \overset{③}{=} t^2\langle x, x\rangle = t^2\|x\|^2$$

Cauchy-Schwarz の不等式が成り立つことの証明

明治大学 MEIJI UNIVERSITY

内積空間 *H* 上の任意の *x*, *y* に対して

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

$$\boxed{\|x\|^2 := \langle x, x\rangle}$$

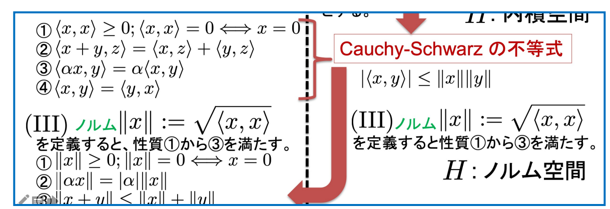
$\because$ $t \in \mathbb{R}$

$$0 \leq t^2\|x\|^2 + 2t\langle x, y\rangle + \|y\|^2$$

$$\mathrm{D} = \langle x, y\rangle^2 - \|x\|^2\|y\|^2 \leq 0$$

$$\therefore \langle x, y\rangle^2 \leq \|x\|^2\|y\|^2$$

$$\therefore |\langle x, y\rangle| \leq \|x\|\|y\|$$



演習問題：Cauchy-Schwarz ⇒ 三角不等式

3次元ユークリッド空間

$$\mathbb{R}^3$$

(I) 和・スカラー倍を定義すると、性質①から⑧を満たす。

(II) 内積

$$\langle x, y \rangle := a_1 b_1 + a_2 b_2 + a_3 b_3$$

を定義すると、性質①から④を満たす。

(III) ノルム

$$\|x\| := \sqrt{\langle x, x \rangle}$$

を定義すると、性質①から③を満たす。

定理(参考文献57頁; (C1)):

$(x_n) \subset \mathbb{R}^3 \ : \ \text{bounded}$

$$\Longrightarrow \exists (x_{n_i}) \subset (x_n) : x_{n_i} \to x \in \mathbb{R}^3$$

ヒルベルト空間

H : 集合

(I) 和・スカラー倍を定義して、性質①から⑧を満たすとする。

(II) 内積

$$\langle x, y \rangle \in \mathbb{R}$$

を定義して、性質①から④を満たすとする。

(III) ノルム

$$\|x\| := \sqrt{\langle x, x \rangle}$$

を定義すると性質①から③を満たす。

(IV) 以下を満たすとする:

$(x_n) \subset H \ : \ \text{bounded}$

$$\Longrightarrow \exists (x_{n_i}) \subset (x_n) : x_{n_i} \to x \in H$$

H : ヒルベルト空間

H: 内積空間【※ (I), (II) を満たす空間】

H: ヒルベルト空間

$\underset{\text{def}}{\Longleftrightarrow}$ **(IV) H 上の有界な点列が弱収束する部分列をもつ。**

$$\exists (x_{n_i}) \subset (x_n) : x_{n_i} \rightharpoonup x \in H$$

$$\Big\Updownarrow \text{def}$$

内積の意味で強収束する

$$\forall y \in H, \langle x_{n_i}, y \rangle \to \langle x, y \rangle$$

よく見るヒルベルト空間の定義との関係

明治大学 MEIJI UNIVERSITY

H: 内積空間【※ (I), (II) を満たす空間】

H: ヒルベルト空間

$$\underset{\text{def}}{\iff}$$

H 上の有界な点列が弱収束する部分列をもつ

$$\iff$$

H: 完備（complete）※ 参考文献99頁

□強収束（ノルムの意味で収束する）

$$x_n \to x \underset{\text{def}}{\iff} \lim_{n\to\infty} \|x_n - x\| = 0$$

□弱収束（内積の意味で収束する）

$$x_n \rightharpoonup x \underset{\text{def}}{\iff} \forall y \in H,\ \lim_{n\to\infty} \langle x_n, y\rangle = \langle x, y\rangle$$

$$\iff \forall y \in H,\ \lim_{n\to\infty} |\langle x_n, y\rangle - \langle x, y\rangle| = 0$$

強収束 $\Rightarrow$ 弱収束

$$\boxed{x_n \to x \Longrightarrow x_n \rightharpoonup x} \quad \text{i.e.,}$$

$$\lim_{n\to\infty} \|x_n - x\| = 0 \Longrightarrow \lim_{n\to\infty} |\langle x_n, y\rangle - \langle x, y\rangle| = 0$$

$\because \quad \forall y \in H,$

$$\begin{aligned}
&|\langle x_n, y\rangle - \langle x, y\rangle| \\
\overset{③,②}{=}\ &|\langle x_n - x, y\rangle| \\
\overset{\text{Cauchy-Schwarz}}{\leq}\ &\|x_n - x\|\|y\| \to 0
\end{aligned}$$

$\blacksquare$

②
$\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$

③
$\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$

3次元ユークリッド空間

$$\mathbb{R}^3$$

(I) 和・スカラー倍を定義すると、
性質①から⑧を満たす。

(II) 内積

$$\langle x, y \rangle := a_1 b_1 + a_2 b_2 + a_3 b_3$$

を定義すると、性質①から④を満たす。

(III) ノルム

$$\|x\| := \sqrt{\langle x, x \rangle}$$

を定義すると、性質①から③を満たす。

定理:

$$(x_n) \subset \mathbb{R}^3 \ : \ \text{bounded}$$

$$\Longrightarrow \exists (x_{n_i}) \subset (x_n) : x_{n_i} \to x \in \mathbb{R}^3$$

ヒルベルト空間

H : 集合

(I) 和・スカラー倍を定義して、
性質①から⑧を満たすとする。

(II) 内積

$$\langle x, y \rangle \in \mathbb{R}$$

を定義して、性質①から④を満たすとする。

(III) ノルム

$$\|x\| := \sqrt{\langle x, x \rangle}$$

を定義すると性質①から③を満たす。

(IV) 以下を満たすとする:

$$(x_n) \subset H \ : \ \text{bounded}$$

$$\Longrightarrow \exists (x_{n_i}) \subset (x_n) : x_{n_i} \to x \in H$$

H : ヒルベルト空間

ヒルベルト空間の例（有限次元ユークリッド空間）

- 有限次元ユークリッド空間
- 有限次元ユークリッド空間では

強収束 ⇔ 弱収束

- 強収束 ⇒ 弱収束（証明済）
- 弱収束 ⇒ 強収束（演習問題）

ヒルベルト空間の例（l^2 空間）

明治大学 MEIJI UNIVERSITY

□ 2 乗総和可能な数列全体からなる空間

$$\ell^2 := \left\{ x = (x_i)_{i=1}^{\infty} : \sum_{i=1}^{\infty} |x_i|^2 < \infty \right\}$$

$x, y \in \ell^2,$

$$\text{(I)}\ \ x + y := (x_1 + y_1, x_2 + y_2, \ldots)$$

$$\alpha x := (\alpha x_1, \alpha x_2, \ldots)$$

$$\text{(II)}\ \ \langle x, y \rangle := \sum_{i=1}^{\infty} x_i y_i \qquad \text{(III)}\ \ \|x\| := \sqrt{\langle x, x \rangle}$$

例題2.3

不動点近似法の証明でよく使う不等式・等式

- **Cauchy-Schwarz の不等式（証明済）**

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

- **三角不等式（演習問題）**

$$\|x+y\| \leq \|x\| + \|y\|$$

- **２乗ノルム展開（証明済）**

$$\|x \pm y\|^2 = \|x\|^2 \pm 2\langle x, y\rangle + \|y\|^2$$

- **平行四辺形公式（２乗ノルム展開から）**

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

不動点近似法の証明でよく使う不等式・等式

- **Cauchy-Schwarz の不等式（証明済）**

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

- **三角不等式（演習問題）**

$$\|x + y\| \leq \|x\| + \|y\|$$

- **２乗ノルム展開（証明済）**

$$\|x \pm y\|^2 = \|x\|^2 \pm 2\langle x, y\rangle + \|y\|^2$$

- **本当によく使う等式** $\alpha \in \mathbb{R}$

$$\|\alpha x + (1-\alpha)y\|^2$$

$$= \alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x - y\|^2$$

本当によく使う等式の証明

$$
\begin{aligned}
&\|\alpha x + (1-\alpha)y\|^2 \\
&\quad = \|\alpha x\|^2 + \|(1-\alpha)y\|^2 + 2\langle \alpha x, (1-\alpha)y\rangle \\
&\quad = \alpha^2\|x\|^2 + (1-\alpha)^2\|y\|^2 + \boxed{2}\alpha(1-\alpha)\boxed{\langle x, y\rangle}
\end{aligned}
$$

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

$$
\boxed{2\langle x, y\rangle} = \|x\|^2 + \|y\|^2 - \|x-y\|^2
$$

$$
\begin{aligned}
&\|\alpha x + (1-\alpha)y\|^2 \\
&\quad = \alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x-y\|^2
\end{aligned}
$$

[redacted]

ヒルベルト空間上の近似法の収束解析の流れ

明治大学 MEIJI UNIVERSITY

① 近似法で生成される点列が有界であることを示す（よく使う不等式・等式）

② 弱収束する部分列が存在する（ヒルベルト空間の定義）

③ 弱収束先が解になることを示す（オピアルの定理）

④ 全体点列の弱収束先も解になることを示す（オピアルの定理）

オピアル (Opial) の定理

H: ヒルベルト空間

$$x_n \rightharpoonup x, \quad x \neq y$$

$$\Longrightarrow \varliminf_{n\to\infty} \|x_n - x\| < \varliminf_{n\to\infty} \|x_n - y\|$$

※証明で使用する事実:

①弱収束する点列は有界 (参考文献: 定理5.4.3)

②ε-n 論法

③下極限の性質 (参考文献: 1.4 章)

参考文献: 定理5.4.4（111頁）※付録

ここまでのまとめ

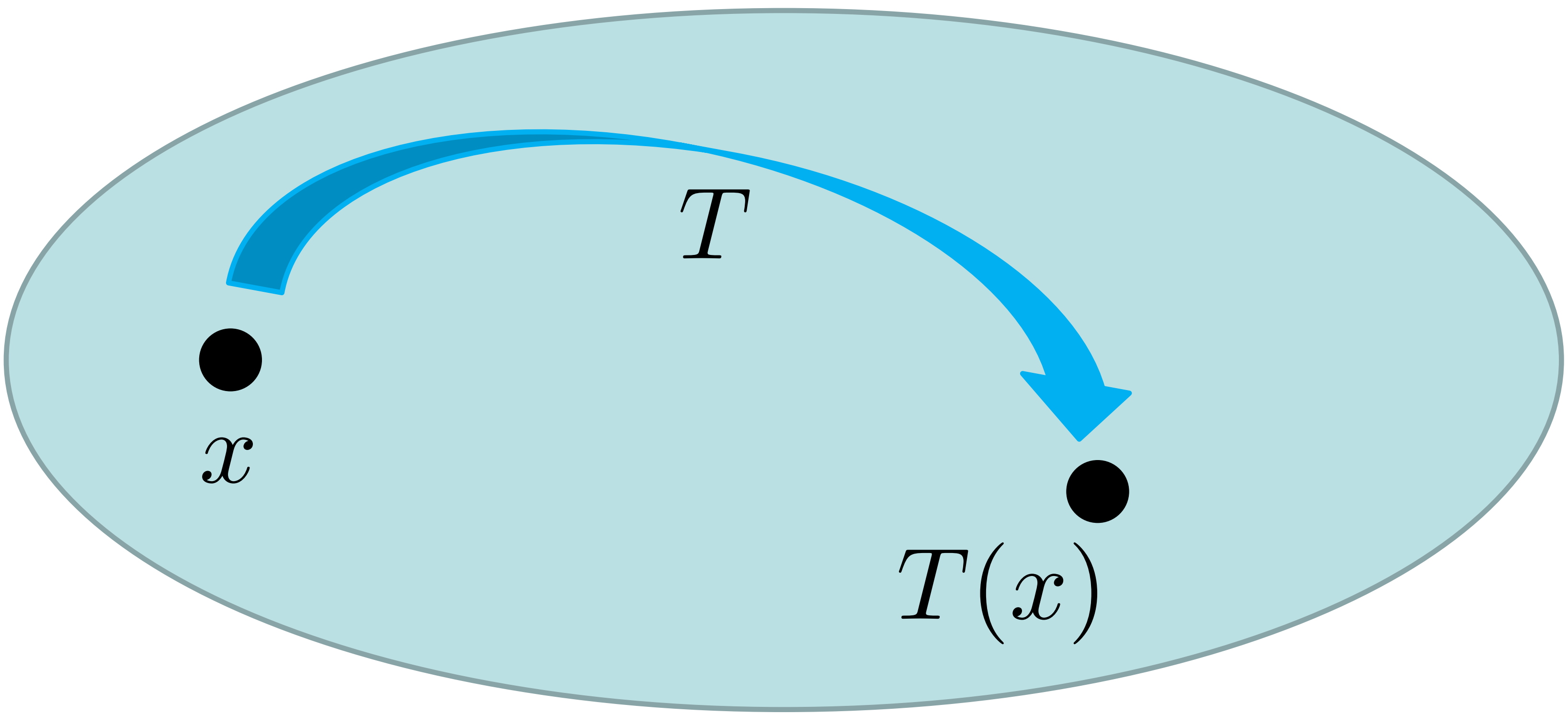
- ヒルベルト空間
  - 有界な点列が弱収束する部分列をもつ内積空間（完備な内積空間）
  - 有限次元ユークリッド空間、l^2 空間
- ヒルベルト空間上で成り立つ便利な不等式・等式

明治大学 MEIJI UNIVERSITY

- 不動点
- 非拡大写像
- 非拡大写像の不動点
- 例

不動点とは？

$$\begin{array}{rccc} T: & H & \to & H \\ & \rotatebox{90}{$\in$} & & \rotatebox{90}{$\in$} \\ & x & \mapsto & T(x) \end{array}$$



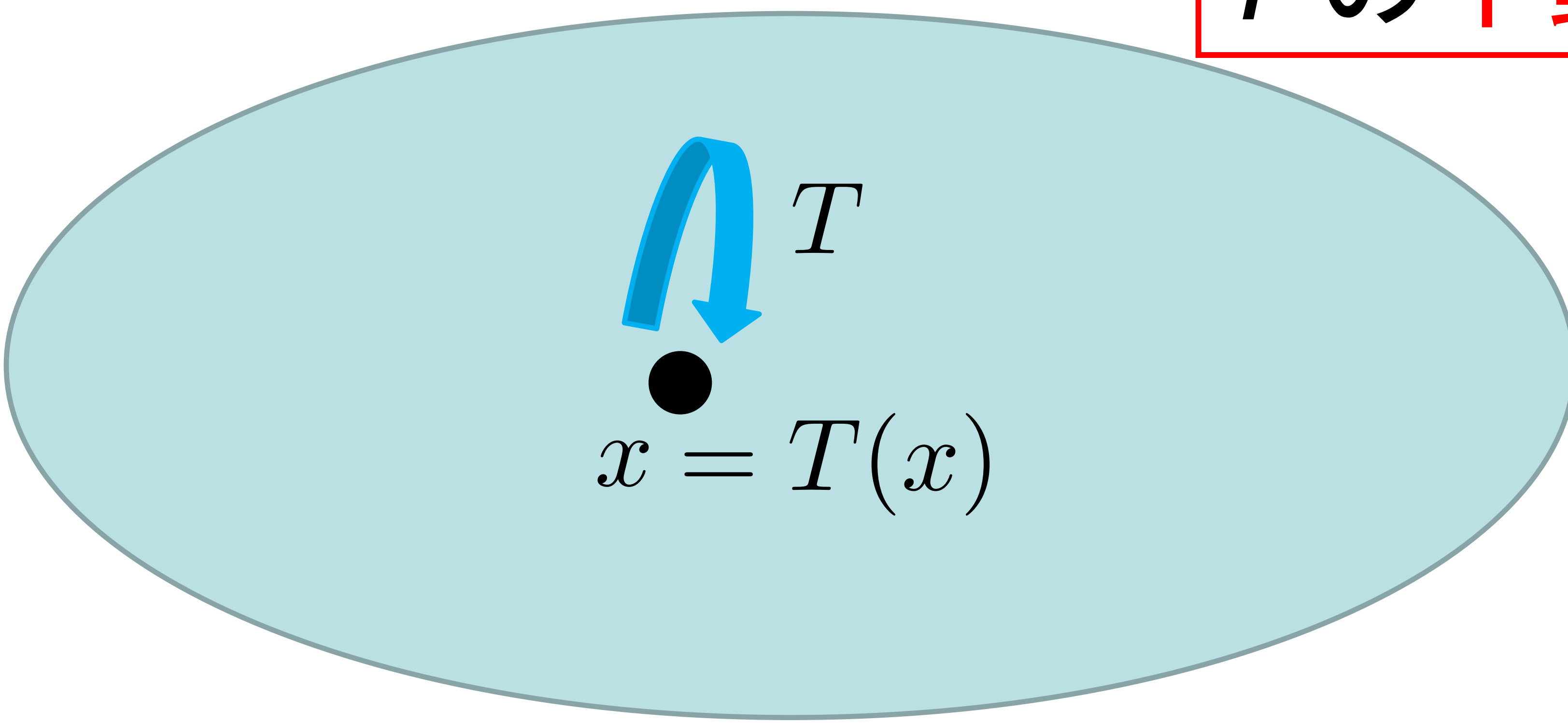
不動点とは？

$$T: H \to H$$

$$T(x) = x$$

T を作用させても**動かない点**のこと

$\boxed{T \text{ の不動点}}$



不動点とは？

$$T : H \to H$$

Tを作用させても

$$T(x) = x$$

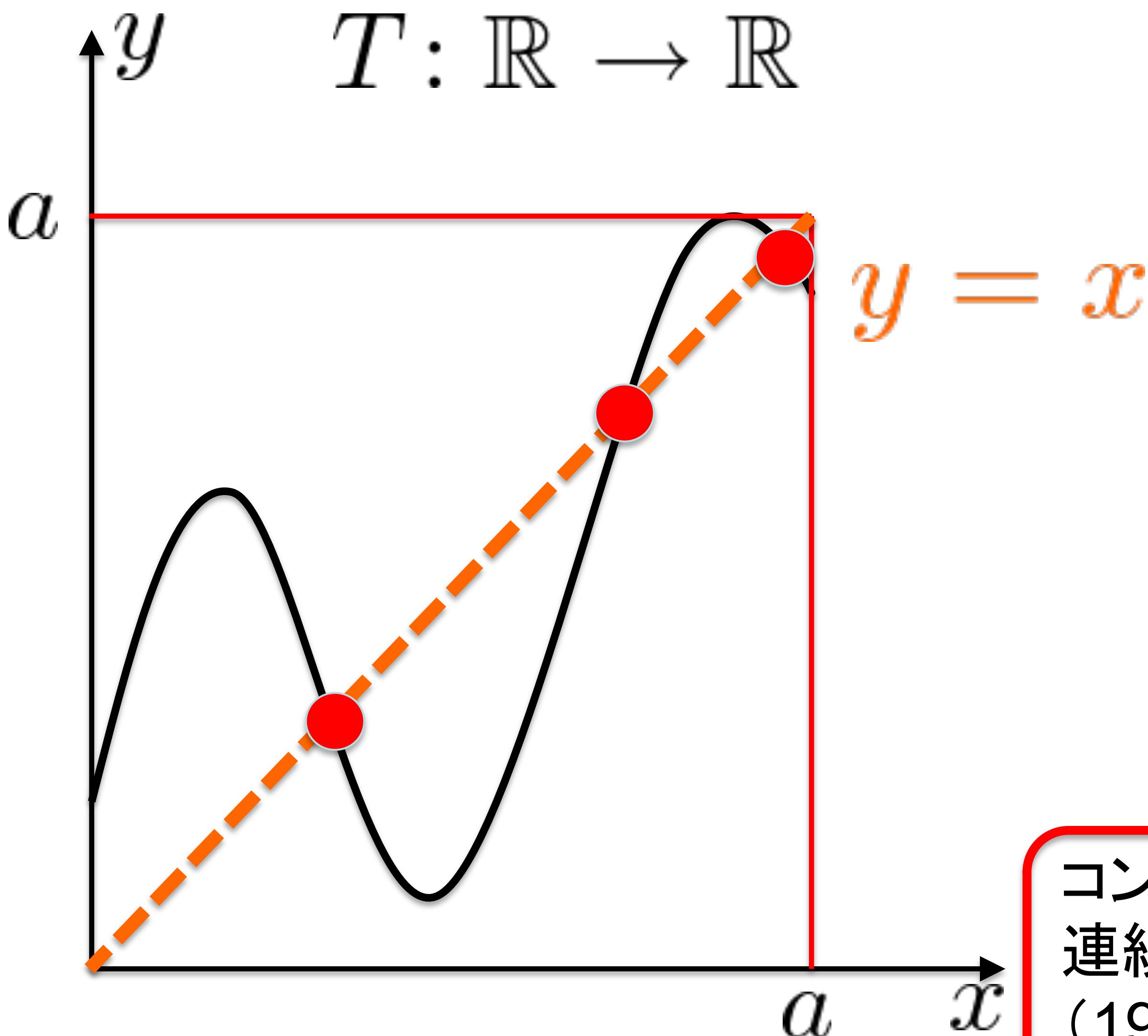
動かない点のこと

$\boxed{T\text{の不動点}}$

Tの不動点集合

$$\mathrm{Fix}(T) := \{x \in H : T(x) = x\}$$

色々な不動点の存在（不動点定理）



Dr. L.E.J. Brouwer
（1881--1966）

コンパクト凸集合上の
連続写像は不動点をもつ
（1909）

色々な不動点の存在（不動点定理）

Dr. Stefan Banach
（1892--1945）

完備距離空間上の縮小写像の不動点は一意に存在する(1922)

$$x_{n+1} = T(x_n)$$

$$x_n \to x^\star$$

縮小写像 T の不動点に収束する

色々な不動点の存在（不動点定理）

角谷 静夫 先生
[Dr. Shizuo Kakutani]
（1911--2004）

位相空間上のある集合値写像には不動点が存在する（Nash均衡の存在が言える）（1941）

色々な不動点の存在（不動点定理）

非拡大写像（半群）の不動点の存在性を明らかにした

高橋 渉 先生
[Dr. Wataru Takahashi]
（1944--2020）

非拡大写像

***H*: ヒルベルト空間**

***T*: *H* → *H*, 非拡大（nonexpansive）**

$$\underset{\mathrm{def}}{\iff} \forall x, y \in H,$$

$$\|T(x) - T(y)\| \leq \|x - y\|$$

非拡大写像と例

H: ヒルベルト空間

T: ***H*** → ***H***, 非拡大（nonexpansive）

$$\underset{\text{def}}{\Longleftrightarrow} \forall x, y \in H,$$

$$\|T(x) - T(y)\| \leq \|x - y\|$$

- 非拡大写像ー不動点集合
- 例1: 射影ー射影先
- 例2: 射影の積ー射影先の共通部分
- 例3: 最急降下ー凸最適解の集合

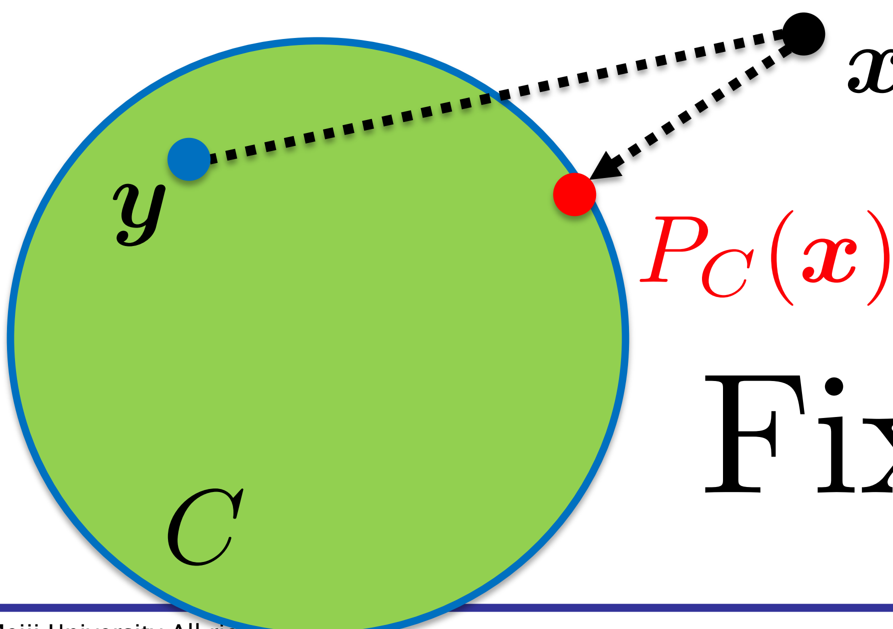
例1: 射影（Projection, Projector）

H: ヒルベルト空間

C: H の空でない閉凸集合

P_C: $H \to C$, 射影

$$\overset{\text{def}}{\iff} P_C(x) \in C, \ \|P_C(x) - x\| = \min_{y \in C} \|y - x\|$$



$$\mathrm{Fix}(P_C) = C$$

例1: 射影（Projection, Projector）の性質

C: *H* の空でない閉凸集合

P_C: $H \to C$, 射影

1. 射影の不動点集合＝射影先

$$\mathrm{Fix}(P_C) = C$$

2. 射影と内積の関係

$$\forall y \in C, \quad \langle x - P_C(x), y - P_C(x) \rangle \leq 0$$

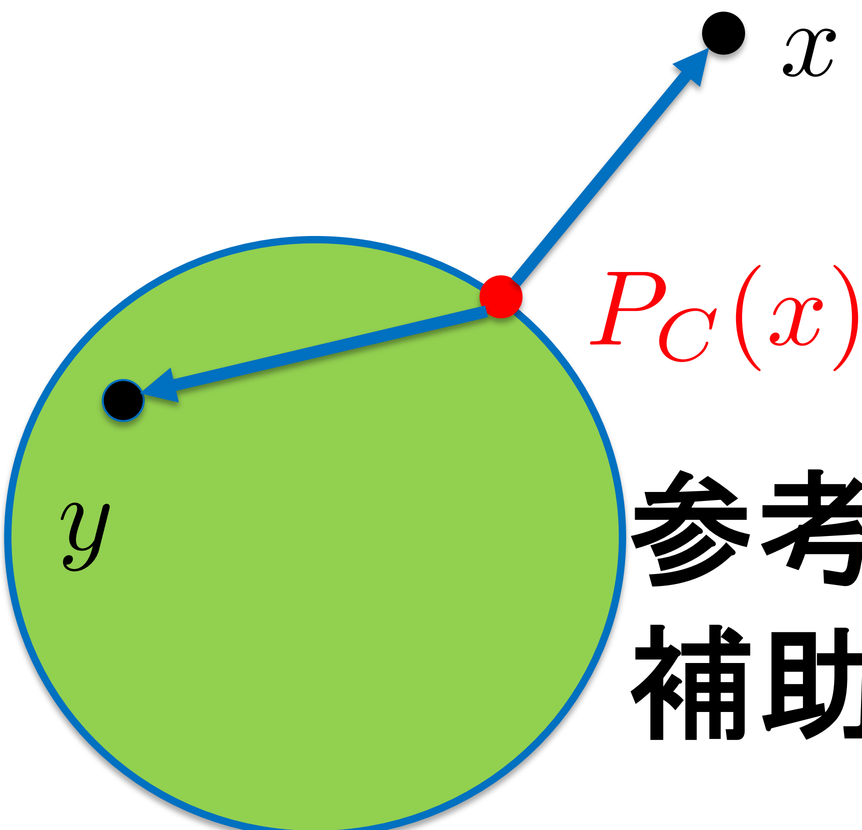
3. 射影は（堅）非拡大

$$\forall x, y \in H,$$

$$\|P_C(x) - P_C(y)\|^2 \leq \langle x - y, P_C(x) - P_C(y) \rangle$$

例1 性質 2.

$$\langle x - P_C(x), y - P_C(x) \rangle \leq 0$$



参考文献
補助定理5.2.2（102頁）

例1 性質 3. の証明

明治大学 MEIJI UNIVERSITY

$$\forall x, y \in H,$$

$$\|P_C(x) - P_C(y)\|^2 \leq \langle x - y, P_C(x) - P_C(y) \rangle$$

$\because$ $\forall z \in C,$ **例1 性質 2.**

$$\langle x - P_C(x), \boxed{z} - P_C(x) \rangle \leq 0$$

$$\langle y - P_C(y), \boxed{z} - P_C(y) \rangle \leq 0$$

$$\langle x - P_C(x), \boxed{P_C(y)} - P_C(x) \rangle \leq 0$$

$$\iff \langle P_C(x) - x, P_C(x) - P_C(y) \rangle \leq 0$$

$$\langle y - P_C(y), \boxed{P_C(x)} - P_C(y) \rangle \leq 0$$

例1 性質 3. の証明

$$\forall x, y \in H, \qquad \boxed{\langle x, z\rangle + \langle y, z\rangle = \langle x + y, z\rangle}$$

$$\|P_C(x) - P_C(y)\|^2 \leq \langle x - y, P_C(x) - P_C(y)\rangle$$

$$\because \quad \langle P_C(x) - x, P_C(x) - P_C(y)\rangle \leq 0$$

$$\langle y - P_C(y), P_C(x) - P_C(y)\rangle \leq 0$$

$$\langle (P_C(x) - x) + (y - P_C(y)), P_C(x) - P_C(y)\rangle \leq 0$$

$$\langle (P_C(x) - P_C(y)) - (x - y), P_C(x) - P_C(y)\rangle \leq 0$$

$$\langle P_C(x) - P_C(y), P_C(x) - P_C(y)\rangle - \langle x - y, P_C(x) - P_C(y)\rangle \leq 0$$

$$\|P_C(x) - P_C(y)\|^2 - \langle x - y, P_C(x) - P_C(y)\rangle \leq 0$$

[redacted]

例1 性質3. 射影は堅非拡大 ⇒ 非拡大

❑ 射影は堅非拡大

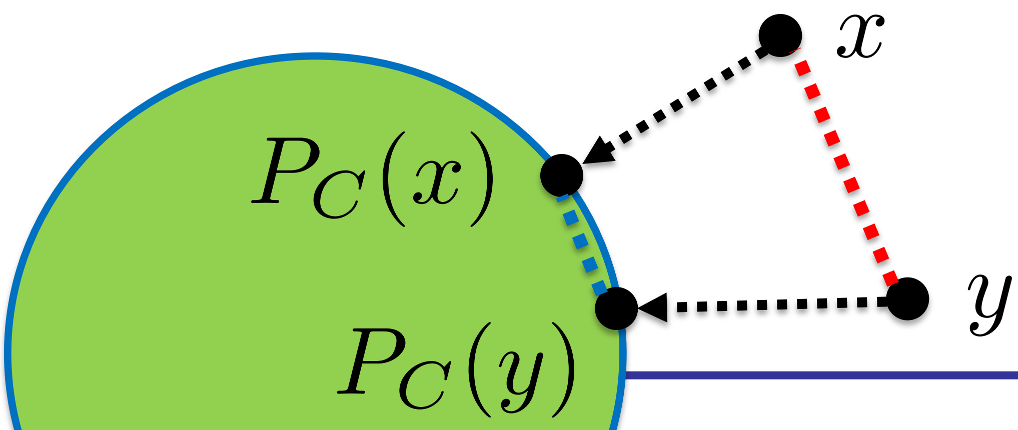
$$\|P_C(x) - P_C(y)\|^2 \leq \langle x - y, P_C(x) - P_C(y)\rangle$$

Cauchy-Schwarz

$$\leq \|x - y\| \|P_C(x) - P_C(y)\|$$

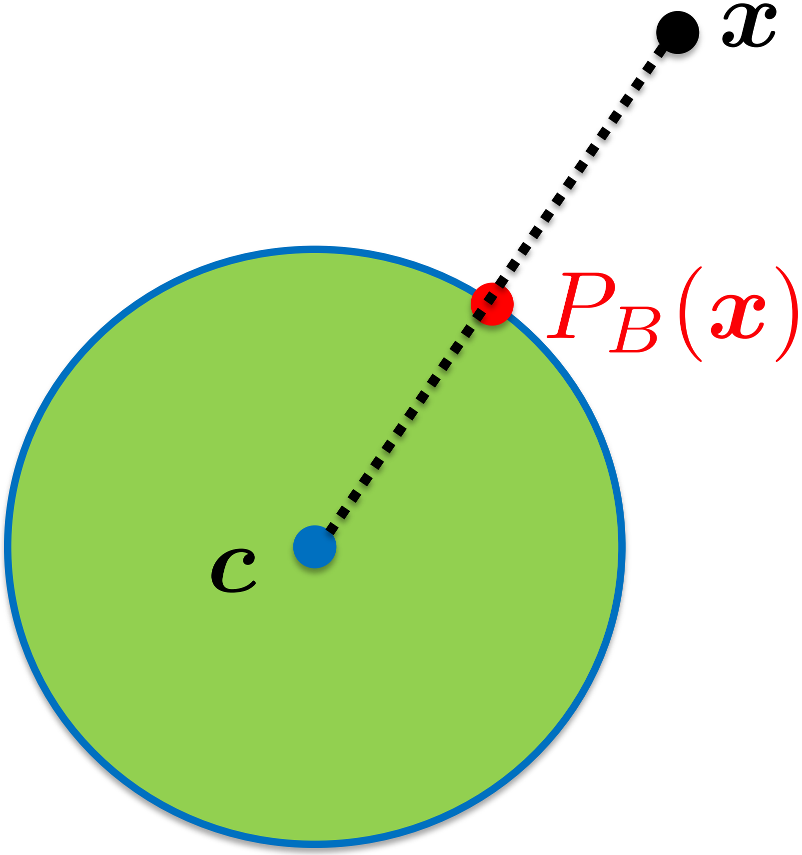
❑ 射影は非拡大でもある

$$\|P_C(x) - P_C(y)\| \leq \|x - y\|$$



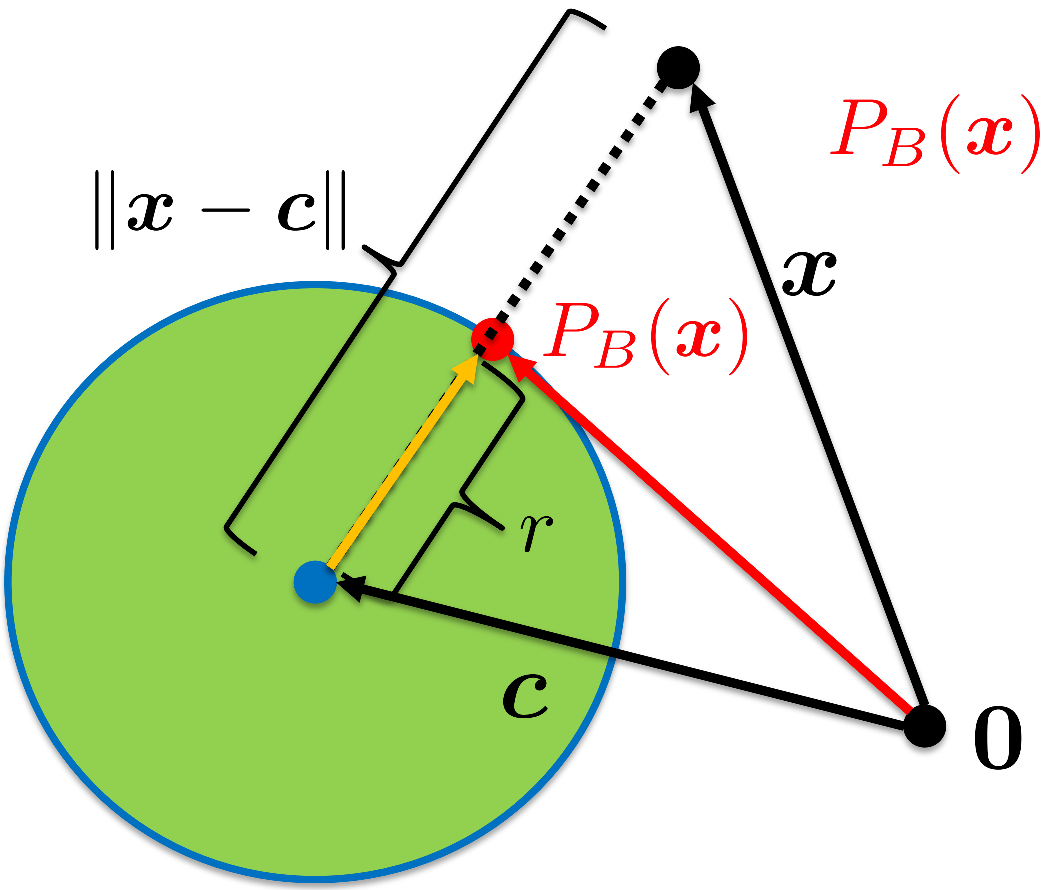
閉球への射影

$$B(\boldsymbol{c}, r) := \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{c}\| \leq r\}$$



閉球への射影

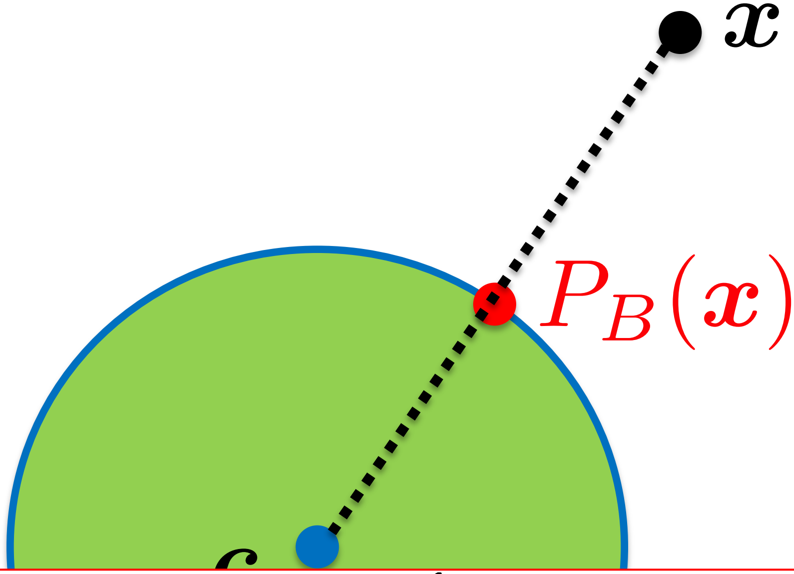
$$B(\boldsymbol{c}, r) := \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{c}\| \leq r\}$$



$$P_B(\boldsymbol{x}) = \boldsymbol{c} + \frac{r}{\|\boldsymbol{x} - \boldsymbol{c}\|}(\boldsymbol{x} - \boldsymbol{c})$$

閉球への射影

$$B(\boldsymbol{c}, r) := \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{c}\| \leq r\}$$



$$P_B(\boldsymbol{x}) = \begin{cases} \boldsymbol{c} + \dfrac{r}{\|\boldsymbol{x} - \boldsymbol{c}\|}(\boldsymbol{x} - \boldsymbol{c}) & (\boldsymbol{x} \notin B) \\ \boldsymbol{x} & (\boldsymbol{x} \in B) \end{cases}$$

半空間への射影

$$\boldsymbol{a} \neq \boldsymbol{0}, r \in \mathbb{R}$$

$$H(\boldsymbol{a}, r) := \{\boldsymbol{x} \in \mathbb{R}^n : \langle \boldsymbol{a}, \boldsymbol{x} \rangle \leq r\}$$

$$P_H(\boldsymbol{x}) = ??$$

演習問題

非拡大写像と例

H: ヒルベルト空間

***T*: *H* → *H*, 非拡大（nonexpansive）**

$$\underset{\text{def}}{\Longleftrightarrow} \forall x, y \in H,$$

$$\|T(x) - T(y)\| \leq \|x - y\|$$

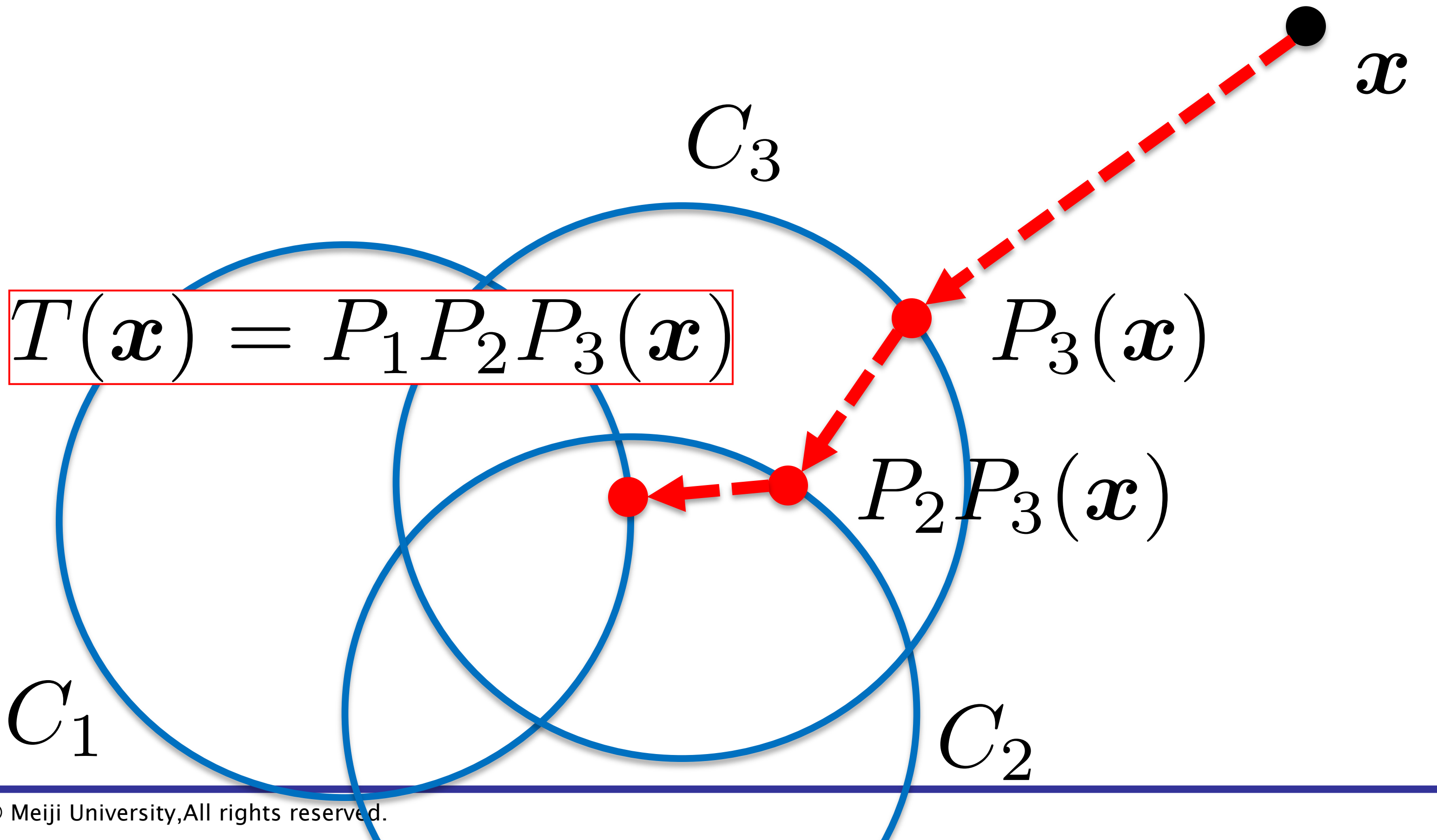
- **非拡大写像ー不動点集合**
- **例1: 射影ー射影先**
- **例2: 射影の積ー射影先の共通部分**
- **例3: 最急降下ー凸最適解の集合**

例2: 射影の積

***H*: ヒルベルト空間**

C_i: H の空でない閉凸集合 ($i = 1,2,\cdots,m$)

$P_i = P_{C_i}$ $T = P_1 P_2 \cdots P_m$



例2: 射影の積の性質

H: ヒルベルト空間

C_i: *H* の空でない閉凸集合 (i = 1,2,⋯,m)

$P_i = P_{C_i}$ $T = P_1 P_2 \cdots P_m$

1.　射影積の不動点集合＝射影先の共通部分

$$\mathrm{Fix}(T) = \bigcap_{i=1}^{m} C_i \neq \emptyset$$

凸実行可能問題

2.　射影積は非拡大

$\forall x, y \in H,$

$$\|T(x) - T(y)\| \leq \|x - y\|$$

例2 性質 2. の証明

$$T = P_1 P_2 \cdots P_m$$

$$\|P_i(x) - P_i(y)\| \leq \|x - y\|$$

$$\boxed{\text{例1 性質 3.}}$$

$$\begin{aligned}
\|T(x) - T(y)\| &= \|P_1((P_2 \cdots P_m)(x)) - P_1((P_2 \cdots P_m)(y))\| \\
&\leq \|(P_2 \cdots P_m)(x) - (P_2 \cdots P_m)(y)\| \\
&\leq \|(P_3 \cdots P_m)(x) - (P_3 \cdots P_m)(y)\| \\
&\vdots \\
&\leq \|x - y\|
\end{aligned}$$

[redacted]

例2 性質 1. の証明

$$T = P_1 P_2 \cdots P_m$$

$$\boxed{\mathrm{Fix}(T) \subset \bigcap_{i=1}^{m} C_i \neq \emptyset}$$

演習問題

$$\mathrm{Fix}(T) = \bigcap_{i=1}^{m} C_i \neq \emptyset$$

$$\mathrm{Fix}(T) \supset \bigcap_{i=1}^{m} C_i \neq \emptyset$$

$\boxed{\text{例1 性質 1.}}$ $\boxed{x = P_i(x)}$

$$\because \quad x \in \bigcap_{i=1}^{m} C_i \iff \forall i, x \in C_i = \mathrm{Fix}(P_i)$$

$$\Longrightarrow T(x) = P_1 P_2 \cdots P_{m-1} P_m(x)$$

$$= P_1 P_2 \cdots P_{m-1}(x)$$

$$= \cdots = x$$

例2 の応用例（ネットワーク資源割当）

C_i: H の空でない閉凸集合 ($i = 1,2,\cdots,m$)

$$T = P_1 P_2 \cdots P_m$$

$$\mathrm{Fix}(T) = \bigcap_{i=1}^{m} C_i \neq \emptyset$$

- ネットワーク上のユーザ i が所望する条件 C_i
- ネットワーク全体の条件 $\bigcap_{i=1}^{m} C_i$

H.I.: Distributed Optimization for Network Resource Allocation With Nonsmooth Utility Functions, IEEE Transactions on Control of Network Systems 6 (4): 1354-1365 (2019)

例2 の応用例（機械学習）

C_i: H の空でない閉凸集合 ($i = 1,2,\cdots,m$)

$$T = P_1 P_2 \cdots P_m$$

$$\mathrm{Fix}(T) = \bigcap_{i=1}^{m} C_i \neq \emptyset$$

- **スパース学習:** $C_{\mathrm{s}} = \left\{ \boldsymbol{x} \colon \|x\|_1 = \sum_{i=1}^{n} |x_i| \leq s \right\}$
- **ダイバーシティ学習:** $C_{\mathrm{d}} = \{ \boldsymbol{x} \colon f_{\mathrm{d}}(\boldsymbol{x}) \leq d \}$
- **スパース&ダイバーシティ学習:** $C_{\mathrm{s}} \cap C_{\mathrm{d}}$

H.I.: Stochastic Fixed Point Optimization Algorithm for Classifier Ensemble, IEEE Transactions on Cybernetics 50 (10): 4370–4380 (2020)

例2 の応用例（機械学習; もう少し丁寧に）

- 制約集合: $\mathbb{R}^n_+ \cap C_\mathrm{s} \cap C_\mathrm{d}$ ← とても複雑

  計算可能な非拡大写像

$$T = P_+ P_\mathrm{s} P_\mathrm{d}$$
$$\Longrightarrow \mathrm{Fix}(T) = \mathbb{R}^n_+ \cap C_\mathrm{s} \cap C_\mathrm{d}$$

- 目的関数: データ *i* に関する損失関数 f_i の総和

$$f(\boldsymbol{x}) = \sum_{i=1}^{N} f_i(\boldsymbol{x})$$

- 自然に思いつくアルゴリズム ← 計算困難

$$\boldsymbol{x}_{k+1} = \boxed{P_{\mathbb{R}^n_+ \cap C_\mathrm{s} \cap C_\mathrm{d}}}(\boldsymbol{x}_k - \alpha_k \nabla f_i(\boldsymbol{x}_k))$$

例2 の応用例（機械学習; もう少し丁寧に）

明治大学 MEIJI UNIVERSITY

- 制約集合: $\mathbb{R}_+^n \cap C_{\mathrm{s}} \cap C_{\mathrm{d}}$ ← とても複雑

  計算可能な非拡大写像

$$T = P_+ P_{\mathrm{s}} P_{\mathrm{d}}$$

$$\Longrightarrow \mathrm{Fix}(T) = \mathbb{R}_+^n \cap C_{\mathrm{s}} \cap C_{\mathrm{d}}$$

- 目的関数: データ i に関する損失関数 f_i の総和

$$f(\boldsymbol{x}) = \sum_{i=1}^{N} f_i(\boldsymbol{x})$$

- 不動点最適化アルゴリズム ← 計算可能

$$\boldsymbol{x}_{k+1} = T(\boldsymbol{x}_k) - \alpha_k \nabla f_i(T(\boldsymbol{x}_k))$$

H.I.: Stochastic Fixed Point Optimization Algorithm for Classifier Ensemble, IEEE Transactions on Cybernetics 50 (10): 4370–4380 (2020)

例3: 最急降下

H: ヒルベルト空間

C: H の空でない閉凸集合

$f: H \to \mathbb{R}$, **凸**

$\nabla f: H \to H$, **リプシッツ連続**

$$\underset{\text{def}}{\Longleftrightarrow} \exists L > 0 : \forall x, y \in H,$$

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$$

例: $A \in \mathcal{S}^n; 0 \leq \lambda_{\min} \leq \lambda_{\max} \quad \boldsymbol{b} \in \mathbb{R}^n$

$$f(\boldsymbol{x}) = \frac{1}{2}\langle \boldsymbol{x}, A\boldsymbol{x}\rangle - \langle \boldsymbol{b}, \boldsymbol{x}\rangle$$

$\Longrightarrow \nabla f(\boldsymbol{x}) = A\boldsymbol{x} - \boldsymbol{b} : \lambda_{\max}$ ー リプシッツ連続

例3: 最急降下

H: ヒルベルト空間

C: H の空でない閉凸集合

$f : H \to \mathbb{R}$, **凸**

$\nabla f : H \to H$, **L-リプシッツ連続**

$T : H \to H$,

$$T(x) = P_C(x - \lambda \nabla f(x))$$

最急降下法

$$\boxed{x_{k+1} = x_k - \lambda \nabla f(x_k)}$$

$$\lambda \in \left(0, \frac{2}{L}\right]$$

例3: 最急降下の性質

H: ヒルベルト空間

$$\lambda \in \left(0, \frac{2}{L}\right]$$

C: *H* の空でない閉凸集合

$f : H \to \mathbb{R}$, 凸, $\nabla f : H \to H$, *L*-リプシッツ連続

$$T(x) = P_C(x - \lambda \nabla f(x))$$

1. *T* の不動点集合＝*C* 上の *f* の最小点集合

$$\mathrm{Fix}(T) = \operatorname*{argmin}_{x \in C} f(x)$$

2. *T* は非拡大

$\forall x, y \in H,$

$$\|T(x) - T(y)\| \leq \|x - y\|$$

例3 性質 2. の証明

$$T(x) = P_C(x - \lambda \nabla f(x)) \qquad \lambda \in \left(0, \frac{2}{L}\right]$$

$$\Longrightarrow \|T(x) - T(y)\| \leq \|x - y\|$$

$\because$

$$\|T(x) - T(y)\|^2$$

$$= \|P_C(x - \lambda \nabla f(x)) - P_C(y - \lambda \nabla f(y))\|^2$$

$$\leq \|(x - \lambda \nabla f(x)) - (y - \lambda \nabla f(y))\|^2$$

例1 性質 3.
P_Cは非拡大

$$= \|(x - y) - \lambda(\nabla f(x) - \nabla f(y))\|^2$$

例3 性質 2. の証明

$$T(x) = P_C(x - \lambda \nabla f(x)) \qquad \lambda \in \left(0, \frac{2}{L}\right]$$

$$\Longrightarrow \|T(x) - T(y)\| \leq \|x - y\|$$

∵

Theorem18.15
(i), (v) 322頁

$$\|T(x) - T(y)\|^2$$

$$\leq \|(x - y) - \lambda(\nabla f(x) - \nabla f(y))\|^2$$

$$= \|x - y\|^2 - 2\lambda \langle x - y, \nabla f(x) - \nabla f(y) \rangle$$

$$+\lambda^2 \|\nabla f(x) - \nabla f(y)\|^2$$

$$\geq \frac{1}{L} \|\nabla f(x) - \nabla f(y)\|^2$$

参考文献: H.H. Bauschke, P.L. Combettes: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer International Publishing (2017)

例3 性質 2. の証明

$$T(x) = P_C(x - \lambda \nabla f(x)) \qquad \lambda \in \left(0, \frac{2}{L}\right]$$

$$\Longrightarrow \|T(x) - T(y)\| \leq \|x - y\|$$

$\because$

$$\|T(x) - T(y)\|^2$$

$$\leq \|(x - y) - \lambda(\nabla f(x) - \nabla f(y))\|^2$$

$$= \|x - y\|^2 - 2\lambda \boxed{\langle x - y, \nabla f(x) - \nabla f(y)\rangle}$$

$$+\lambda^2 \|\nabla f(x) - \nabla f(y)\|^2 \qquad \geq \frac{1}{L}\|\nabla f(x) - \nabla f(y)\|^2$$

$$\leq \|x - y\|^2 - \lambda\left(\frac{2}{L} - \lambda\right)\|\nabla f(x) - \nabla f(y)\|^2$$

例3 性質 2. の証明

$$T(x) = P_C(x - \lambda \nabla f(x)) \qquad \lambda \in \left(0, \frac{2}{L}\right]$$

$$\Longrightarrow \|T(x) - T(y)\| \leq \|x - y\|$$

$\because$

$$\|T(x) - T(y)\|^2$$

$$\leq \|x - y\|^2 - \lambda \left(\frac{2}{L} - \lambda \right) \|\nabla f(x) - \nabla f(y)\|^2$$

$$\leq \|x - y\|^2 \qquad \blacksquare$$

H. I. : Iterative Algorithm for Solving Triple-hierarchical Constrained Optimization Problem, Journal of Optimization Theory and Applications, 148 (3): 580-592 (2011)

例3 性質 1. の証明（付録）

$$T(x) = P_C(x - \lambda \nabla f(x))$$

$$\mathrm{Fix}(T) = \operatorname*{argmin}_{x \in C} f(x)$$

ここまでのまとめ（非拡大写像と例）

明治大学 MEIJI UNIVERSITY

***H*: ヒルベルト空間**

***T*: *H* → *H*, 非拡大（nonexpansive）**

$$\underset{\text{def}}{\Longleftrightarrow} \forall x, y \in H,$$

$$\|T(x) - T(y)\| \leq \|x - y\|$$

- **例1: 射影ー射影先** $\boxed{P_C - \mathrm{Fix}(P_C) = C}$
- **例2: 射影の積ー射影先の共通部分**

$$\boxed{T = P_1 \cdots P_m - \mathrm{Fix}(T) = \bigcap C_i}$$

- **例3: 最急降下ー凸最適解の集合**

$$\boxed{T = P_C(I - \lambda \nabla f) - \mathrm{Fix}(T) = \underset{x \in C}{\mathrm{argmin}}\, f(x)}$$

□不動点近似法

- ○不動点を見つけるための手法
- ○Krasnosel’skiĭ-Mann

□最近の研究動向

Krasnosel’skiĭ-Mann 不動点近似法

Dr. Mark Krasnosel’skiĭ
(1920--1997)

Dr. William Mann
(1920--2006)

Krasnosel’skiĭ-Mann 不動点近似法
(1955)　　　(1953)

Krasnosel'skiĭ-Mann 不動点近似法

H: ヒルベルト空間

T: $H \to H$, 非拡大写像, Fix(T)≠∅

$$\alpha \in (0,1)$$

$$\boxed{\begin{aligned} &x_0 \in H, \\ &x_{k+1} = \alpha x_k + (1-\alpha)T(x_k) \end{aligned}}$$

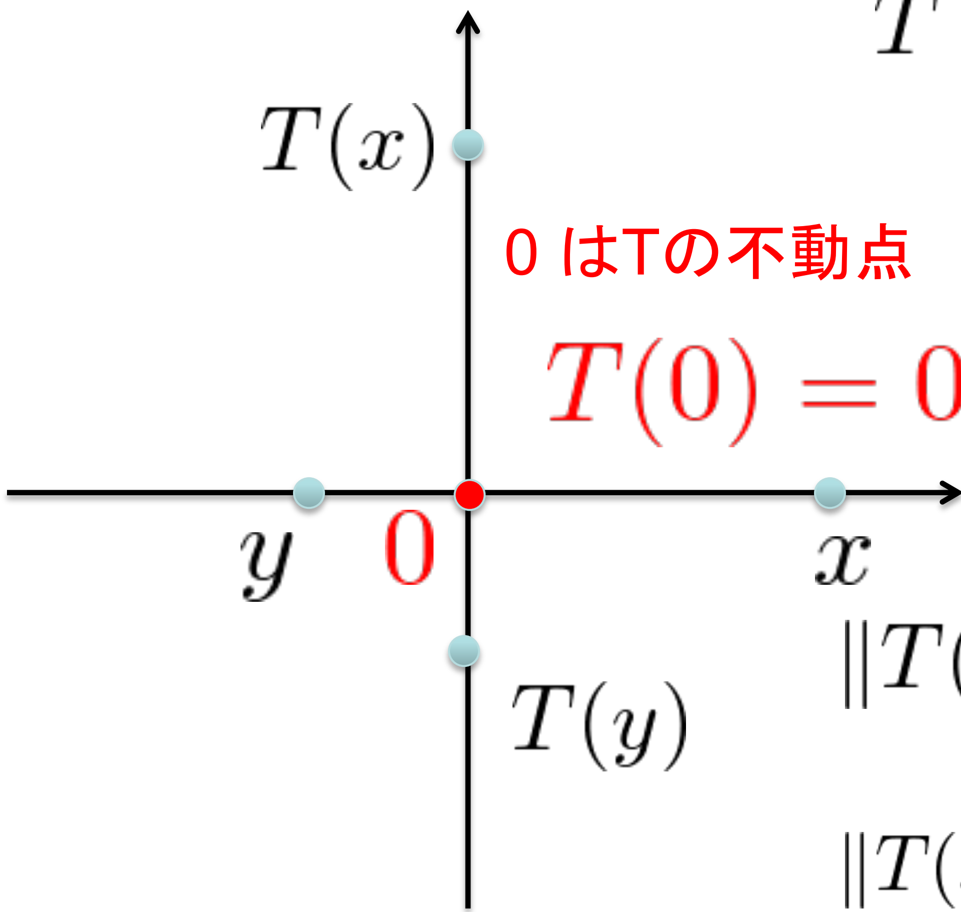
$$(k = 0, 1, \ldots)$$

$$\Longrightarrow x_k \rightharpoonup x^\star \in \mathrm{Fix}(T)$$

回転写像

$$T := \begin{bmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{bmatrix}$$

$$\|T(x)\| = \|x\|$$



0 はTの不動点

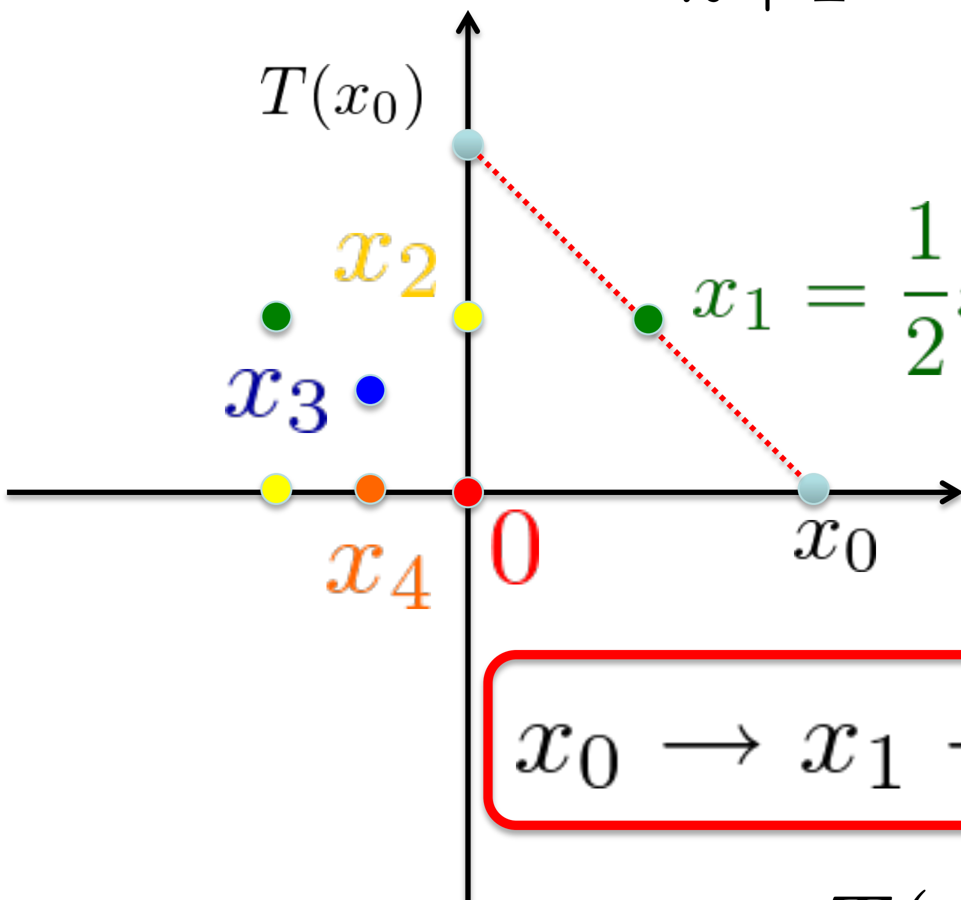
$$T(0) = 0$$

$$\|T(x-y)\| = \|x-y\|$$

$$\|T(x) - T(y)\| = \|x-y\|$$

Krasnosel'skiĭ-Mann 不動点近似法の例

$$x_{k+1} = \frac{1}{2}x_k + \frac{1}{2}T(x_k)$$

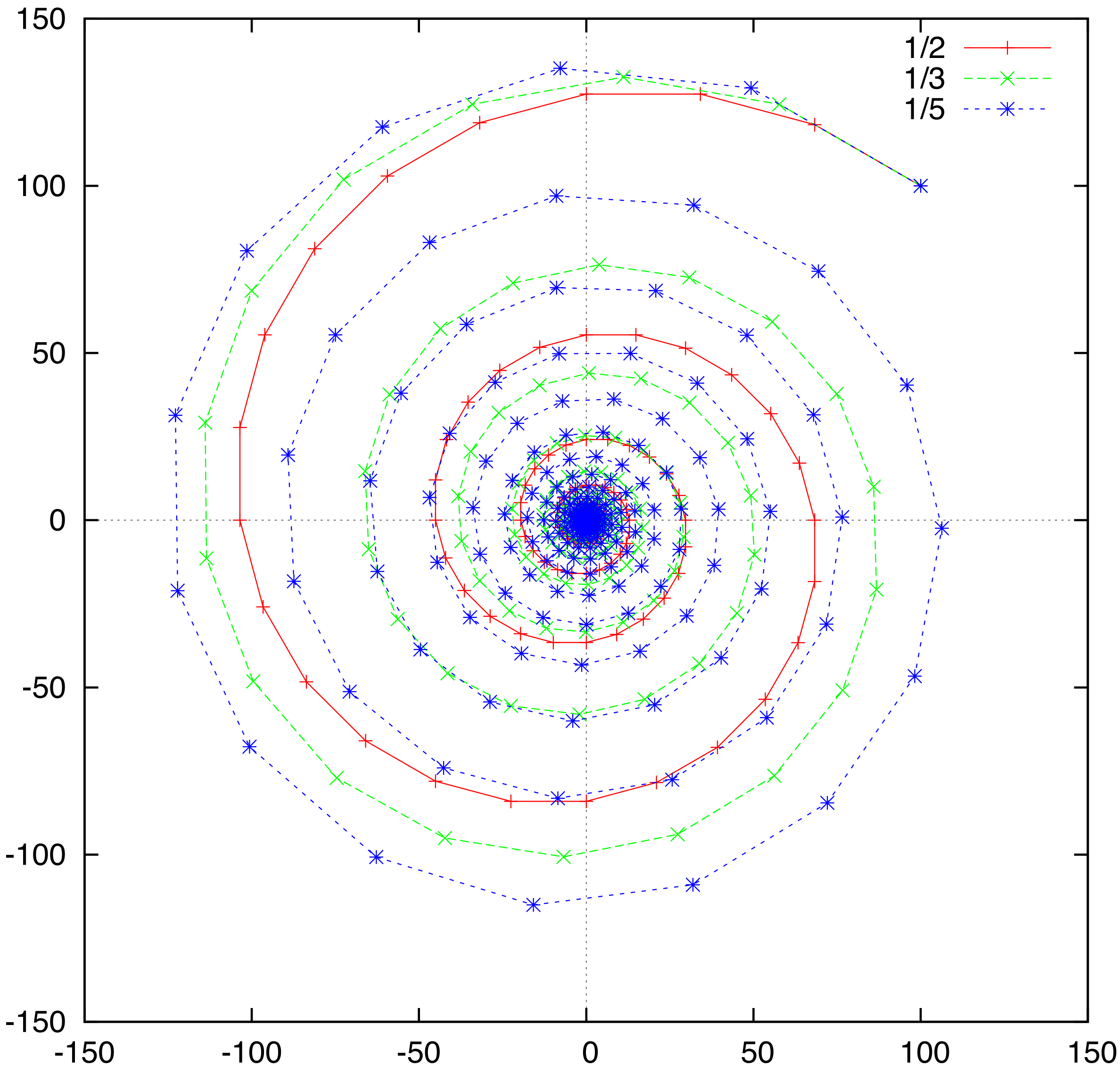


$$x_1 = \frac{1}{2}x_0 + \frac{1}{2}T(x_0)$$

$$\boxed{x_0 \longrightarrow x_1 \longrightarrow x_2 \longrightarrow \cdots \longrightarrow 0}$$

$x_{k+1} = T(x_k)$ は 0 に収束しない

数値例（30°回転）



ヒルベルト空間上の近似法の収束解析の流れ

① 近似法で生成される点列が有界であることを示す（よく使う不等式・等式）

② 弱収束する部分列が存在する（ヒルベルト空間の定義）

③ 弱収束先が解になることを示す（オピアルの定理）

④ 全体点列の弱収束先も解になることを示す（オピアルの定理）

証明

① 近似法で生成される点列が**有界**であることを示す
（よく使う**不等式・等式**）

$$x_0 \in H,$$

$$x_{k+1} = \alpha x_k + (1-\alpha)T(x_k)$$

$$x \in \mathrm{Fix}(T),$$

$$\|x_{k+1} - x\|^2 \leq \cdots \leq \|x_k - x\|^2$$

本当によく使う等式

$$\|\alpha x + (1-\alpha)y\|^2$$
$$= \alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x-y\|^2$$

証明

明治大学 MEIJI UNIVERSITY

① 近似法で生成される点列が**有界**であることを示す（よく使う**不等式・等式**）

$$\boxed{\begin{aligned} &x_0 \in H, \\ &x_{k+1} = \alpha x_k + (1-\alpha)T(x_k) \end{aligned}}$$

$$x \in \mathrm{Fix}(T),$$

$$\begin{aligned} \|x_{k+1} - x\|^2 &= \|\alpha x_k + (1-\alpha)T(x_k) - x\|^2 \\ &= \|\alpha(x_k - x) + (1-\alpha)(T(x_k) - x)\|^2 \end{aligned}$$

$$\boxed{\begin{aligned} &\|\alpha x + (1-\alpha)y\|^2 \\ &\quad = \alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x-y\|^2 \end{aligned}}$$

証明

① 近似法で生成される点列が有界であることを示す（よく使う不等式・等式）

$$x \in \mathrm{Fix}(T),$$

$$\begin{aligned}
\|x_{k+1} - x\|^2 &= \|\alpha x_k + (1-\alpha)T(x_k) - x\|^2 \\
&= \|\alpha(x_k - x) + (1-\alpha)(T(x_k) - x)\|^2 \\
&= \alpha\|x_k - x\|^2 + (1-\alpha)\|T(x_k) - x\|^2 \\
&\quad - \alpha(1-\alpha)\|x_k - T(x_k)\|^2
\end{aligned}$$

$$\boxed{\begin{aligned}
&\|\alpha x + (1-\alpha)y\|^2 \\
&\quad = \alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x-y\|^2
\end{aligned}}$$

証明

① 近似法で生成される点列が有界であることを示す（よく使う不等式・等式）

$$x \in \mathrm{Fix}(T),$$

$$\|x_{k+1} - x\|^2 = \alpha\|x_k - x\|^2 + (1-\alpha)\boxed{\|T(x_k) - x\|^2}$$

$$= \|T(x_k) - T(x)\|^2 \leq \|x_k - x\|^2$$

$$-\alpha(1-\alpha)\|x_k - T(x_k)\|^2$$

$$\leq \|x_k - x\|^2 - \alpha(1-\alpha)\|x_k - T(x_k)\|^2$$

$$\leq \|x_k - x\|^2$$

$$\therefore \exists \lim_{k\to\infty} \|x_k - x\| \qquad \therefore (x_k) : \text{bounded}$$

証明

① 近似法で生成される点列が有界であることを示す（よく使う不等式・等式）

$$\|x_{k+1} - x\|^2 \leq \|x_k - x\|^2 - \alpha(1-\alpha)\|x_k - T(x_k)\|^2$$

$$\alpha(1-\alpha)\|x_k - T(x_k)\|^2 \leq \|x_k - x\|^2 - \|x_{k+1} - x\|^2$$

$$\boxed{\exists \lim_{k\to\infty} \|x_k - x\|}$$

$$\therefore \lim_{k\to\infty} \|x_k - T(x_k)\|^2 = 0$$

$$\therefore \lim_{k\to\infty} \|x_k - T(x_k)\| = 0$$

証明

① 近似法で生成される点列が**有界**であることを示す（よく使う**不等式・等式**）

② **弱収束**する部分列が存在する（**ヒルベルト空間**の定義）

$$\exists (x_{k_i}) \subset (x_k) : x_{k_i} \rightharpoonup x \in H$$

③ 収束先が**不動点**になることを示す（**オピアルの定理**）：**背理法** $\boxed{x \neq T(x)}$

$$\lim_{i\to\infty} \|x_{k_i} - x\| < \lim_{i\to\infty} \|x_{k_i} - T(x)\|$$

$$\lim_{k\to\infty} \|x_k - T(x_k)\| = 0$$

証明 ③

$$\exists (x_{k_i}) \subset (x_k) : x_{k_i} \rightharpoonup x \in H$$

$$\boxed{x \neq T(x)}$$

$$\varliminf_{i\to\infty} \|x_{k_i} - x\| < \varliminf_{i\to\infty} \|x_{k_i} - T(x)\|$$

$$= \varliminf_{i\to\infty} \|x_{k_i} - T(x_{k_i}) + T(x_{k_i}) - T(x)\|$$

$$\leq \varliminf_{i\to\infty} \{\|x_{k_i} - T(x_{k_i})\| + \|T(x_{k_i}) - T(x)\|\}$$

三角不等式

証明 ③

$$\exists (x_{k_i}) \subset (x_k) : x_{k_i} \rightharpoonup x \in H$$

$$\boxed{x \neq T(x)}$$

$$\varliminf_{i\to\infty} \|x_{k_i} - x\| < \varliminf_{i\to\infty} \|x_{k_i} - T(x)\|$$

$$\leq \varliminf_{i\to\infty} \{\|x_{k_i} - T(x_{k_i})\| + \|T(x_{k_i}) - T(x)\|\}$$

$$\boxed{\lim_{k\to\infty} \|x_k - T(x_k)\| = 0}$$

$$= \lim_{i\to\infty} \|x_{k_i} - T(x_{k_i})\| + \varliminf_{i\to\infty} \|T(x_{k_i}) - T(x)\|$$

$$\boxed{\varliminf_{n\to\infty} (a_n + b_n) = \lim_{n\to\infty} a_n + \varliminf_{n\to\infty} b_n}$$

参考文献
問題1.4.4 (2)

証明 ③

$$\exists (x_{k_i}) \subset (x_k) : x_{k_i} \rightharpoonup x \in H$$

$$\boxed{x \neq T(x)}$$

$$\varliminf_{i\to\infty} \|x_{k_i} - x\| < \varliminf_{i\to\infty} \|x_{k_i} - T(x)\|$$

$$\leq \varliminf_{i\to\infty} \{\|x_{k_i} - T(x_{k_i})\| + \|T(x_{k_i}) - T(x)\|\}$$

$$= \lim_{i\to\infty} \|x_{k_i} - T(x_{k_i})\| + \varliminf_{i\to\infty} \|T(x_{k_i}) - T(x)\|$$

$$= \varliminf_{i\to\infty} \|T(x_{k_i}) - T(x)\| \leq \varliminf_{i\to\infty} \|x_{k_i} - x\|$$

T : 非拡大

矛盾

証明

明治大学 MEIJI UNIVERSITY

① 近似法で生成される点列が有界であることを示す（よく使う不等式・等式）

② 弱収束する部分列が存在する（ヒルベルト空間の定義）

$$\exists (x_{k_i}) \subset (x_k) : x_{k_i} \rightharpoonup x \in H$$

③ 弱収束先が不動点になることを示す（オピアルの定理）背理法 $\boxed{x \neq T(x)} \Rightarrow \boxed{x = T(x)}$

④ 全体点列の弱収束先も不動点になることを示す（オピアルの定理）

全体点列が x に弱収束する

⇔ すべての弱収束する部分列の収束先が x

（参考文献 定理 5.4.1）

$$\exists (x_{k_i}) \subset (x_k) : x_{k_i} \rightharpoonup x \in \mathrm{Fix}(T)$$

証明③の議論 $\Longrightarrow x = z$

$$\exists (x_{k_j}) \subset (x_k) : x_{k_j} \rightharpoonup z \in \mathrm{Fix}(T)$$

背理法 $\boxed{x \neq z}$ $\exists \lim_{k\to\infty} \|x_k - x\| \ (x \in \mathrm{Fix}(T))$

オピアルの定理

$$\varliminf_{i\to\infty} \|x_{k_i} - x\| < \varliminf_{i\to\infty} \|x_{k_i} - z\|$$

証明 ④

全体点列が x に弱収束する

⇔ すべての弱収束する部分列の収束先が x

（参考文献 定理 5.4.1）

$$\left.\begin{array}{l}\exists (x_{k_i}) \subset (x_k) : x_{k_i} \rightharpoonup x \in \mathrm{Fix}(T) \\ \exists (x_{k_j}) \subset (x_k) : x_{k_j} \rightharpoonup z \in \mathrm{Fix}(T)\end{array}\right. \Longrightarrow x = z$$

背理法 $\boxed{x \neq z}$ $\exists \lim_{k\to\infty} \|x_k - x\| \ (x \in \mathrm{Fix}(T))$

オピアルの定理

$$\lim_{k\to\infty} \|x_k - x\| < \lim_{k\to\infty} \|x_k - z\|$$

証明 ④

$$\exists (x_{k_i}) \subset (x_k) : x_{k_i} \rightharpoonup x \in \mathrm{Fix}(T)$$

$$\exists (x_{k_j}) \subset (x_k) : x_{k_j} \rightharpoonup z \in \mathrm{Fix}(T)$$

$\boxed{x \neq z}$ $\boxed{\exists \lim_{k\to\infty} \|x_k - z\|}$

$$\lim_{k\to\infty} \|x_k - x\| < \lim_{k\to\infty} \|x_k - z\| = \underline{\lim}_{j\to\infty} \|x_{k_j} - z\|$$

$\boxed{x \neq z}$: オピアルの定理

$$< \underline{\lim}_{j\to\infty} \|x_{k_j} - x\| = \lim_{k\to\infty} \|x_k - x\|$$

矛盾

$$\therefore x = z$$

$$\therefore x_k \rightharpoonup x \in \mathrm{Fix}(T) \quad \blacksquare$$

Krasnosel'skiĭ-Mann 不動点近似法の収束率

H: ヒルベルト空間

T: $H \to H$, 非拡大写像, Fix(T)≠∅

$$\alpha \in (0,1)$$

$$\boxed{\begin{aligned} &x_0 \in H, \\ &x_{k+1} = \alpha x_k + (1-\alpha)T(x_k) \end{aligned}}$$

$$\Longrightarrow \|x_k - T(x_k)\| = o\left(\frac{1}{\sqrt{k}}\right)$$

証明

$$\forall k \in \mathbb{N}, \|x_{k+1} - T(x_{k+1})\| \leq \|x_k - T(x_k)\|$$

$\because$

$$\|x_{k+1} - T(x_{k+1})\|$$

$$= \|\alpha x_k + (1-\alpha)T(x_k) - T(x_{k+1})\|$$

$$= \|\alpha(x_k - T(x_{k+1})) + (1-\alpha)(T(x_k) - T(x_{k+1}))\|$$

$$\leq \alpha\|x_k - T(x_{k+1})\| + (1-\alpha)\|T(x_k) - T(x_{k+1})\|$$

三角不等式

$$\leq \alpha\|x_k - T(x_{k+1})\| + (1-\alpha)\|x_k - x_{k+1}\|$$

T: 非拡大

証明

$$\forall k \in \mathbb{N}, \|x_{k+1} - T(x_{k+1})\| \leq \|x_k - T(x_k)\|$$

$\because$

$$\|x_{k+1} - T(x_{k+1})\|$$

$$\leq \alpha\|x_k - T(x_{k+1})\| + (1-\alpha)\|x_k - x_{k+1}\|$$

$$\leq \alpha\boxed{\|x_k - x_{k+1}\|} + \alpha\|x_{k+1} - T(x_{k+1})\|$$

$$+ (1-\alpha)\|x_k - x_{k+1}\|$$

$$\boxed{\|x_k - x_{k+1}\|} = \|x_k - \alpha x_k - (1-\alpha)T(x_k)\|$$

$$= (1-\alpha)\|x_k - T(x_k)\|$$

証明

$$\forall k \in \mathbb{N}, \|x_{k+1} - T(x_{k+1})\| \leq \|x_k - T(x_k)\|$$

$\because$

$$
\begin{aligned}
&\|x_{k+1} - T(x_{k+1})\| \\
&\leq \boxed{\|x_k - x_{k+1}\|} + \alpha\|x_{k+1} - T(x_{k+1})\| \\
&= (1-\alpha)\|x_k - T(x_k)\| + \alpha\|x_{k+1} - T(x_{k+1})\|
\end{aligned}
$$

$$
\begin{aligned}
\boxed{\|x_k - x_{k+1}\|} &= \|x_k - \alpha x_k - (1-\alpha)T(x_k)\| \\
&= (1-\alpha)\|x_k - T(x_k)\|
\end{aligned}
$$

証明

$$\forall k \in \mathbb{N}, \|x_{k+1} - T(x_{k+1})\| \leq \|x_k - T(x_k)\|$$

$\because$

$$\begin{aligned}
&\|x_{k+1} - T(x_{k+1})\| \\
&\leq \|x_k - x_{k+1}\| + \alpha\|x_{k+1} - T(x_{k+1})\| \\
&= (1-\alpha)\|x_k - T(x_k)\| + \alpha\|x_{k+1} - T(x_{k+1})\|
\end{aligned}$$

$$\therefore \|x_{k+1} - T(x_{k+1})\| \leq \|x_k - T(x_k)\|$$

証明

$$\forall k \in \mathbb{N}, \|x_{k+1} - T(x_{k+1})\| \leq \|x_k - T(x_k)\|$$

①

$x \in \mathrm{Fix}(T)$,

$$\|x_{k+1} - x\|^2 \leq \|x_k - x\|^2 - \alpha(1-\alpha)\|T(x_k) - x_k\|^2$$

$$\alpha(1-\alpha)\|T(x_k) - x_k\|^2 \leq \|x_k - x\|^2 - \|x_{k+1} - x\|^2$$

$$\begin{aligned}\alpha(1-\alpha)\sum_{k=1}^{n}\|T(x_k) - x_k\|^2 &\leq \sum_{k=1}^{n}\left\{\|x_k - x\|^2 - \|x_{k+1} - x\|^2\right\} \\ &= \|x_1 - x\|^2 - \|x_{n+1} - x\|^2 \\ &\leq \|x_1 - x\|^2\end{aligned}$$

証明

明治大学 MEIJI UNIVERSITY

$$\forall k \in \mathbb{N}, \|x_{k+1} - T(x_{k+1})\| \leq \|x_k - T(x_k)\|$$

$$\alpha(1-\alpha)\sum_{k=1}^{n} \underbrace{\boxed{\|T(x_k) - x_k\|^2}}_{\geq \|T(x_n) - x_n\|^2} \leq \|x_1 - x\|^2$$

$$\alpha(1-\alpha)n\|T(x_n) - x_n\|^2 \leq \|x_1 - x\|^2$$

$$\|T(x_n) - x_n\|^2 \leq \frac{\|x_1 - x\|^2}{\alpha(1-\alpha)n}$$

$$\therefore \|T(x_n) - x_n\| \leq \frac{\|x_1 - x\|}{\sqrt{\alpha(1-\alpha)n}} \quad \blacksquare$$

【参考】適応手法の収束率

- AMSGrad (Adaptive Mean Square Gradient)
  - 凸最適化 (Reddi et al.; ICLR2018):

$$\frac{1}{T}\left(\sum_{t\in\mathcal{T}} f_t(\boldsymbol{x}_t) - f^\star\right) = \mathcal{O}\left(\sqrt{\frac{1+\log T}{T}}\right) \Rightarrow \mathcal{O}\left(\frac{1}{\sqrt{T}}\right)$$

  - 非凸最適化 (Chen et al.; ICLR2019):

$$\min_{k=1,2,\ldots,n} \mathbb{E}\left[\|\nabla f(\boldsymbol{x}_k)\|^2\right] = \mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right) \Rightarrow \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$$

- Krasnosel’skiĭ-Mann

$$\mathcal{O}\left(\frac{1}{\sqrt{k}}\right)$$

H.I.: Appropriate Learning Rates of Adaptive Learning Rate Optimization Algorithms for Training Deep Neural Networks, IEEE Transactions on Cybernetics (2021)

ここまでのまとめ（Krasnosel’skiĭ-Mann）

- **Krasnosel’skiĭ-Mann 不動点近似法**

$$x_{k+1} = \alpha x_k + (1-\alpha) T(x_k)$$

- **非拡大写像のある不動点に弱収束することを証明した**

$$x_k \rightharpoonup x^\star \in \mathrm{Fix}(T)$$

- **収束率は**

$$\|x_k - T(x_k)\| = \mathcal{O}\left(\frac{1}{\sqrt{k}}\right)$$

プログラミングの演習問題

$$B_i = B(\boldsymbol{c}_i, r_i) \ (i = 1, 2, \ldots, m) \qquad B = \bigcap_{i=1}^{m} B_i$$

$$T_1 = P_1 P_2 \cdots P_m$$

$$T_2 = P_1 \left(\sum_{i=2}^{m} w_i P_i \right) \qquad (w_i) \subset (0, 1) : \sum_{i=2}^{m} w_i = 1$$

$$\boldsymbol{x}_{k+1} = \alpha \boldsymbol{x}_k + (1 - \alpha) T_1(\boldsymbol{x}_k)$$

$$\boldsymbol{x}_{k+1} = \alpha \boldsymbol{x}_k + (1 - \alpha) T_2(\boldsymbol{x}_k)$$

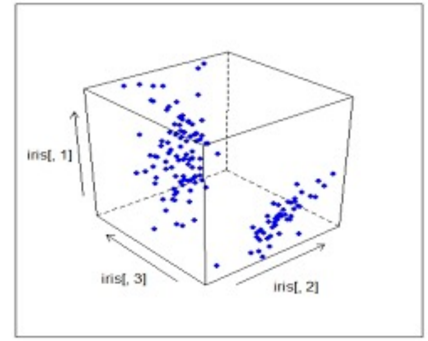
B ≠ ∅ ⇒ 挙動を調べてみましょう。

B = ∅ ⇒ 挙動を調べてみましょう。

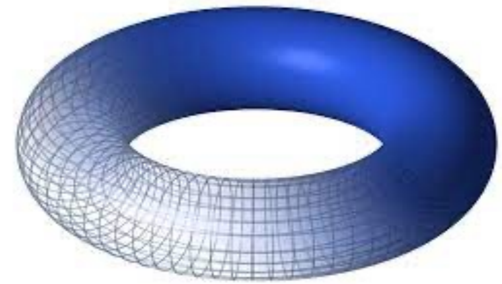
最近の研究動向（多様体へ）

明治大学
MEIJI UNIVERSITY

- 深層ニューラルネットワークは通常
  ユークリッド空間 (まっすぐな空間)



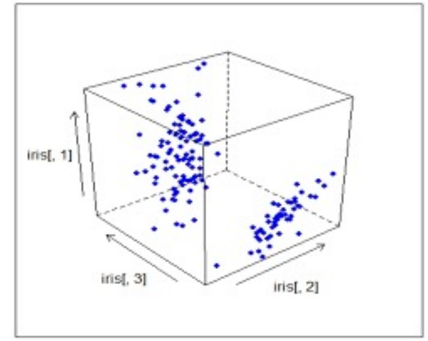
- **自然言語処理**へ応用する時
  多様体 (**曲がった空間**)
  での最適化が必要



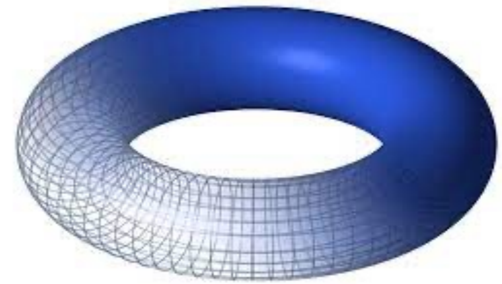
- **8/24 13:00－ リーマン多様体上の最適化理論とその周辺 佐藤寛之先生**

最近の研究動向（多様体へ）

明治大学 MEIJI UNIVERSITY



- 深層ニューラルネットワークは通常
  ユークリッド空間 (まっすぐな空間)
- 自然言語処理へ応用する時
  多様体 (曲がった空間)
  での最適化が必要



- リーマンAdam・リーマンAMSGrad
  の開発に成功（既存手法よりも高性能）

Hiroyuki SAKAI, H.I.: Riemannian Adaptive Optimization Algorithms and Its Applications to Natural Language Processing, IEEE Transactions on Cybernetics (2021)

最近の研究動向（多様体上の不動点問題）

明治大学
MEIJI UNIVERSITY

- 2015年 6/27: 成島先生、佐藤先生との研究打ち合わせ
- 多様体上での非拡大写像の不動点問題
  - 射影ー射影先
  - 射影積ー射影先の共通部分
  - リゾルベントーゼロ点集合
  - 自然な拡張ができる
- 不動点最適化アルゴリズムを多様体へ拡張

H.I., Hiroyuki SAKAI : Riemannian Stochastic Fixed Point Optimization Algorithm, submitted

これで講義はお終いです

- オピアルの定理の証明
- 例3 性質 1. の証明

オピアルの定理の証明 $\boxed{\varliminf_{n\to\infty} \|x_n - x\| < \varliminf_{n\to\infty} \|x_n - y\|}$

$$\|x_n - y\|^2 = \|(x_n - x) + (x - y)\|^2$$
$$= \|x_n - x\|^2 + 2\langle x_n - x, x - y\rangle + \|x - y\|^2$$

$$2\langle x_n - x, x - y\rangle + \|x - y\|^2$$
$$= \|x_n - y\|^2 - \|x_n - x\|^2$$
$$= (\|x_n - y\| + \|x_n - x\|)(\|x_n - y\| - \|x_n - x\|)$$
$$\leq M\ (\|x_n - y\| - \|x_n - x\|)$$

?

$$\because\ x_n \rightharpoonup x \Longrightarrow (x_n)\ :\ \text{bounded}$$

オピアルの定理の証明 $\boxed{\varliminf_{n\to\infty} \|x_n - x\| < \varliminf_{n\to\infty} \|x_n - y\|}$

$$x_n \rightharpoonup x \Longrightarrow \varepsilon = \frac{\|x-y\|^2}{4} > 0$$

$$\exists n_0 \in \mathbb{N} : \forall n \geq n_0$$

$$|\langle x_n - x, x - y\rangle| \leq \varepsilon = \frac{\|x-y\|^2}{4}$$

$$\forall n \geq n_0,$$

$$2\langle x_n - x, x - y\rangle + \|x-y\|^2 \geq -2\varepsilon + \|x-y\|^2$$

$$= \frac{\|x-y\|^2}{2} > 0$$

$$\because x \neq y$$

オピアルの定理の証明 $\varliminf_{n\to\infty} \|x_n - x\| < \varliminf_{n\to\infty} \|x_n - y\|$

$$\forall n \geq n_0$$

$$0 < 2\langle x_n - x, x - y\rangle + \|x - y\|^2$$

$$= (\|x_n - y\| + \|x_n - x\|)(\|x_n - y\| - \|x_n - x\|)$$

$$\leq M(\|x_n - y\| - \|x_n - x\|)$$

$$2\langle x_n - x, x - y\rangle + \|x - y\|^2 + M\|x_n - x\|$$

$$\leq M\|x_n - y\|$$

$$\varliminf_{n\to\infty} \left\{2\langle x_n - x, x - y\rangle + \|x - y\|^2 + M\|x_n - x\|\right\}$$

$$\leq \varliminf_{n\to\infty} M\|x_n - y\|$$

オピアルの定理の証明 $\displaystyle\lim_{n\to\infty}\|x_n - x\| < \lim_{n\to\infty}\|x_n - y\|$

明治大学 MEIJI UNIVERSITY

$$\underline{\lim_{n\to\infty}}\left\{2\langle x_n - x, x - y\rangle + \|x - y\|^2 + M\|x_n - x\|\right\}$$

$$\leq \underline{\lim_{n\to\infty}}\, M\|x_n - y\|$$

$$\underline{\lim_{n\to\infty}}\left\{2\langle x_n - x, x - y\rangle + \|x - y\|^2 + M\|x_n - x\|\right\}$$

$$\geq \underline{\lim_{n\to\infty}}\, 2\langle x_n - x, x - y\rangle + \underline{\lim_{n\to\infty}}\,\|x - y\|^2 + \underline{\lim_{n\to\infty}}\, M\|x_n - x\|$$

$$= \|x - y\|^2 + M\underline{\lim_{n\to\infty}}\,\|x_n - x\|$$

$$\|x - y\|^2 + M\underline{\lim_{n\to\infty}}\,\|x_n - x\| \leq M\underline{\lim_{n\to\infty}}\,\|x_n - y\|$$

オピアルの定理の証明 $\boxed{\varliminf_{n\to\infty} \|x_n - x\| < \varliminf_{n\to\infty} \|x_n - y\|}$

$$\|x-y\|^2 + M \varliminf_{n\to\infty} \|x_n - x\| \leq M \varliminf_{n\to\infty} \|x_n - y\|$$

$$0 < \|x-y\|^2 \leq M \left\{ \varliminf_{n\to\infty} \|x_n - y\| - \varliminf_{n\to\infty} \|x_n - x\| \right\}$$

$$\therefore \varliminf_{n\to\infty} \|x_n - x\| < \varliminf_{n\to\infty} \|x_n - y\| \qquad \blacksquare$$

例3 性質 1. の証明

$$\boxed{\mathrm{Fix}(T) = \operatorname*{argmin}_{x \in C} f(x)}$$

$$T(x) = P_C(x - \lambda \nabla f(x))$$

$$\mathrm{Fix}(T) \subset \operatorname*{argmin}_{x \in C} f(x)$$

$$\because\ x \in \mathrm{Fix}(T) \Longleftrightarrow x = P_C(x - \lambda \nabla f(x))$$

例1 性質 2.

$$\Longrightarrow \forall y \in C,$$

$$\langle x - \lambda \nabla f(x) - x, y - x \rangle \leq 0$$

$$\Longleftrightarrow \forall y \in C, \langle \nabla f(x), y - x \rangle \geq 0$$

例3 性質 1. の証明

$$\boxed{\mathrm{Fix}(T) = \operatorname*{argmin}_{x \in C} f(x)}$$

明治大学 MEIJI UNIVERSITY

$$T(x) = P_C(x - \lambda \nabla f(x))$$

$$\mathrm{Fix}(T) \subset \operatorname*{argmin}_{x \in C} f(x)$$

$$\because \quad x \in \mathrm{Fix}(T) \Longrightarrow \forall y \in C, \langle \nabla f(x), y - x \rangle \geq 0$$

$$f \ : \mathrm{convex} \Longleftrightarrow \forall y \in H, f(y) \geq f(x) + \langle y - x, \nabla f(x) \rangle$$

$$\begin{aligned} \Longrightarrow & \forall y \in C, \\ & f(y) \geq f(x) + \langle y - x, \nabla f(x) \rangle \\ & \qquad \geq f(x) \\ \Longleftrightarrow & \ x \in \operatorname*{argmin}_{y \in C} f(y) \end{aligned}$$

例3 性質 1. の証明

$$\boxed{\mathrm{Fix}(T) = \operatorname*{argmin}_{x\in C} f(x)}$$

$$T(x) = P_C(x - \lambda\nabla f(x))$$

$$\mathrm{Fix}(T) \supset \operatorname*{argmin}_{x\in C} f(x)$$

f: 凸

$$\because\ x \in \operatorname*{argmin}_{y\in C} f(y) \Longrightarrow \forall y \in C,$$

$$f(x) \geq f(y) + \langle x - y, \nabla f(y)\rangle$$

$$\Longrightarrow \forall y \in C,$$

$$\langle y - x, \nabla f(y)\rangle \geq 0$$

補助定理6.1.7

$$\Longrightarrow \forall y \in C,$$

$$\langle y - x, \nabla f(x)\rangle \geq 0$$

例3 性質 1. の証明

$$\boxed{\mathrm{Fix}(T) = \operatorname*{argmin}_{x \in C} f(x)}$$

$$T(x) = P_C(x - \lambda \nabla f(x))$$

$$\mathrm{Fix}(T) \supset \operatorname*{argmin}_{x \in C} f(x)$$

f: 凸

$$\because \quad x \in \operatorname*{argmin}_{y \in C} f(y) \Longrightarrow \forall y \in C, \quad \langle y - x, \nabla f(x) \rangle \geq 0$$

$$\Longrightarrow \forall y \in C, \quad \langle x - \lambda \nabla f(x) - x, y - x \rangle \leq 0$$

例1 性質 2.

$$\Longleftrightarrow x = P_C(x - \lambda \nabla f(x))$$

$$\Longleftrightarrow x = T(x)$$

[redacted]