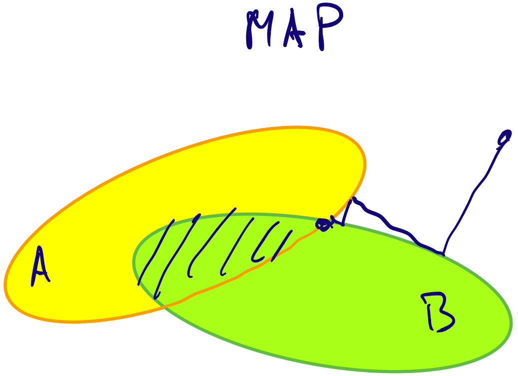
A Lyapunov perspective to projection algorithms

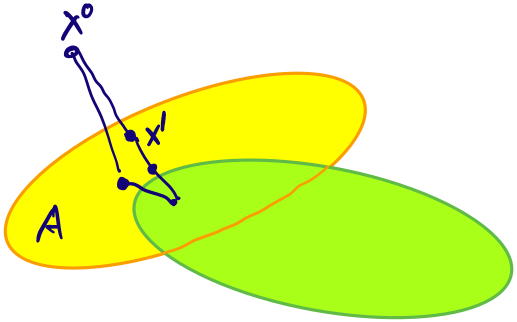
Björn S. Rüffer, The University of Newcastle, Australia

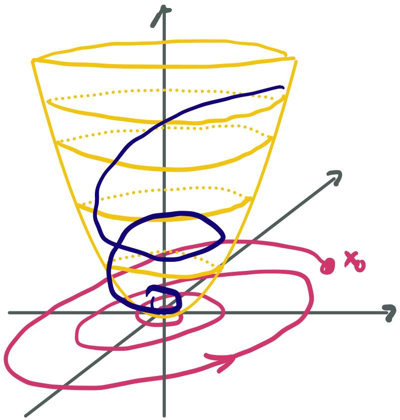
Variational Analysis and Optimisation Webinar | Mathematics of Computation and Optimisation (MoCAO) | AustMS Special Interest Group

2020/10/14

Overview







attraction ⟷ convergence

\+

stability ⟷ start close, stay close

=

asymptotic stability

The setting

- Notation to (mostly) follow [BC17].
- H a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$, and norm-induced topology
- For $D \subset H$ be nonempty and $T\colon D \to H$ denote
  $\operatorname{Fix} T := \{x \in D\colon Tx = x\}$

$T : D \to H$ is called

- *firmly nonexpansive* if for all $x, y \in D$,

$$\|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leq \|x - y\|^2; \tag{1}$$

- *nonexpansive* if it is Lipschitz continuous with constant 1, i.e., for all $x, y \in D$,

$$\|Tx - Ty\| \leq \|x - y\|; \tag{2}$$

- *strictly nonexpansive* if for all $x, y \in D$,

$$x \neq y \implies \|Tx - Ty\| < \|x - y\|; \tag{3}$$

- *firmly quasinonexpansive* if for all $x \in D$, $y \in \operatorname{Fix} T$,

$$\|Tx - y\|^2 + \|Tx - x\|^2 \leq \|x - y\|^2; \tag{4}$$

- *quasinonexpansive* if for all $x \in D$, $y \in \operatorname{Fix} T$,

$$\|Tx - y\| \leq \|x - y\|; \tag{5}$$

  and
- *strictly quasinonexpansive* if for all $x \in D \setminus \operatorname{Fix} T$, $y \in \operatorname{Fix} T$,

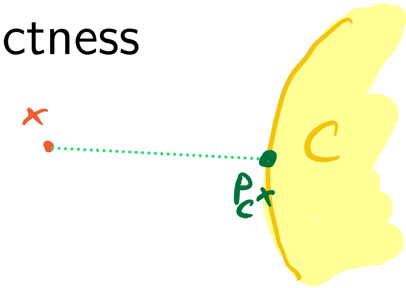
$$\|Tx - y\| < \|x - y\|. \tag{6}$$

Clearly (1) $\implies$ (2) $\implies$ (5); (1) $\implies$ (4) $\implies$ (6) $\implies$ (5); and (3) $\implies$ (6).

Nonexpansiveness under compositions, structure of Fix T

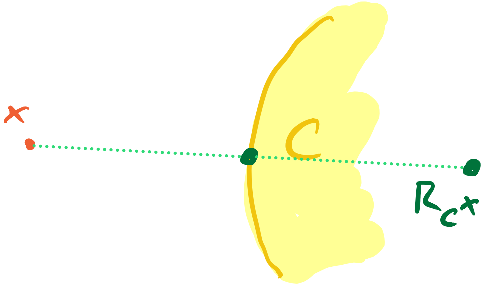
- Let $T_1, T_2 \colon D \to D$ be quasinonexpansive, one of them strictly quasinonexansive, with $\operatorname{Fix} T_1 \cap \operatorname{Fix} T_2 \neq \emptyset$. Then the composition $T_1 T_2$ is quasinonexpansive and $\operatorname{Fix} T_1 T_2 = \operatorname{Fix} T_1 \cap \operatorname{Fix} T_2$.
- If both T_1 and T_2 are strictly quasinonexansive then so is $T_1 T_2$.
- If $T_1, T_2 \colon D \to H$ are nonexpansive, then so is $T_1 T_2$, and $\frac{1}{2}(\operatorname{Id} + T_1)$ is firmly nonexpansive.
- If $T \colon D \to H$ is quasinonexpansive, D nonempty, closed and convex, then $\operatorname{Fix} T$ is closed and convex.

Projectors onto convex sets, precompactness



- Let $C \subset H$ be nonempty, closed, and convex. The *projector onto* C, denoted by $P_C \colon H \to C$, maps each $x \in H$ to the unique $y \in C$ that attains $\operatorname{dist}(x, C) := \inf_{y \in C} \|x - y\|$.
- The projector P_C is firmly nonexpansive and hence continuous, and $\operatorname{Fix} P_C = C$.
- The map $R_C := 2P_C - \operatorname{Id}$ is the *reflector across* C and is nonexpansive.

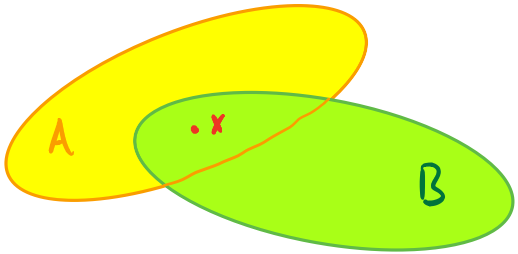
- A set in H is *precompact* if it is contained in a compact set, or, equivalently, if its closure is compact.



The main problem

Problem

Given two nonempty, closed, convex subsets $A, B \subset H$*, find* $x \in A \cap B$*.*



Different but same problem

Given: n nonempty, closed, convex sets C_i, $i = 1, \ldots, n$

Aim: find $x \in \bigcap_{i=1}^{n} C_i$

"Solution": Let

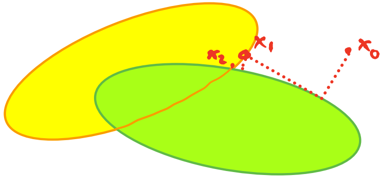
$$A := C_1 \times C_2 \times \ldots \times C_n \subset H^n$$

with projector $P_A = (P_{C_1}, \ldots, P_{C_n})$ and

$$B := \{z = (x, x, \ldots, x) \in H^n \colon x \in H\}$$

and projector $P_B z = (\frac{1}{n} \sum_i z_i, \ldots, \frac{1}{n} \sum_i z_i)$.

MAP/POCS



The first algorithm is known as the *Method of Alternating Projections* (MAP) or *Projections onto Convex Sets* (POCS), and it is simply to iterate the system

$$x^+ = P_A P_B x \tag{7}$$

MAP/POCS

The first algorithm is known as the *Method of Alternating Projections* (MAP) or *Projections onto Convex Sets* (POCS), and it is simply to iterate the system

$$x^+ = P_A P_B x \tag{7}$$

The problem is *consistent* if $A \cap B \neq \emptyset$.

MAP/POCS

The first algorithm is known as the *Method of Alternating Projections* (MAP) or *Projections onto Convex Sets* (POCS), and it is simply to iterate the system

$$x^+ = P_A P_B x \tag{7}$$

The problem is *consistent* if $A \cap B \neq \emptyset$.

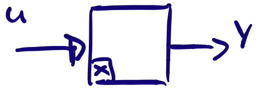
Theorem (Brègman)

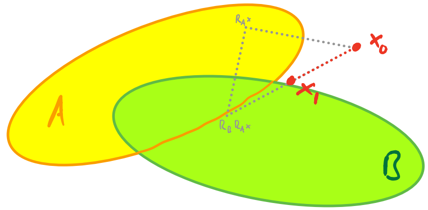
If Problem 1 is consistent then algorithm (7) *converges weakly to a point in* $A \cap B$*.*

Weak vs strong convergence

- A sequence $\{x_n\}_{n\in\mathbb{N}}$ in H converges (strongly, or "in norm") to x, in symbols $x_n \to x$ as $n \to \infty$, if $\|x_n - x\| \to 0$ as $n \to \infty$.
- The sequence convergences weakly to x, in symbols $x_n \rightharpoonup x$, if instead for all $y \in H$, $\langle x_n - x, y\rangle \to 0$ as $n \to \infty$.
- More generally speaking, the weak topology is coarser than the strong topology, and of course strong convergence implies weak convergence.
- Both notions coincide if H is finite dimensional.

Douglas–Rachford Algorithm





$$x^+ = \frac{1}{2}(x + R_B R_A x) := T_{A,B} x \qquad (8)$$
$$y = P_A x.$$

The dual of the Douglas–Rachford Algorithm is the Alternating Direction Method of Multipliers (ADMM), and in this version the algorithm has seen widespread applications in control theory.

Theorem (Lions and Mercier)

If Problem 1 is consistent, then the sequence x_n generated by (8) converges weakly to a point $x \in \mathrm{Fix}\, T_{A,B}$ with $y = P_A x \in A \cap B$.

Generalisations

The feasibility problem 1 is merely a special case of

$$\min_{x \in H} f(x) + g(x) \tag{9}$$

where $f, g \colon H \to [-\infty, \infty]$. This contains Problem 1 when $f = i_A$ and $g = i_B$, the indicator functions of the sets A and B, where

$$i_C(x) = \begin{cases} 0 & \text{for } x \in C, \\ \infty & \text{otherwise.} \end{cases}$$

Generalisations

The feasibility problem 1 is merely a special case of

$$\min_{x \in H} f(x) + g(x) \tag{9}$$

where $f, g \colon H \to [-\infty, \infty]$. This contains Problem 1 when $f = i_A$ and $g = i_B$, the indicator functions of the sets A and B, where

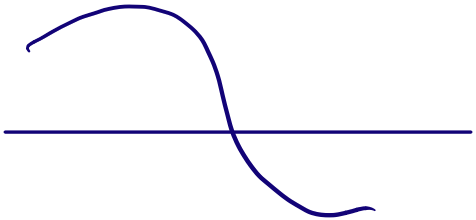
$$i_C(x) = \begin{cases} 0 & \text{for } x \in C, \\ \infty & \text{otherwise.} \end{cases}$$

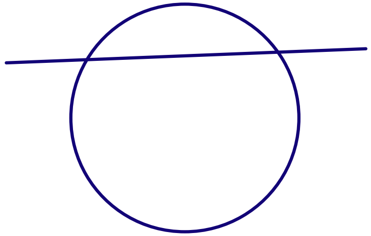
Problem (9) in turn is a special case of the inclusion problem,

$$\text{find } x \in H \text{ such that } 0 \in \mathbb{A}(x) + \mathbb{B}(x)$$

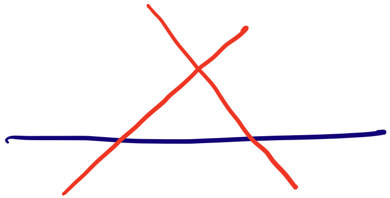
for set-valued, maximally monotone operators $\mathbb{A}, \mathbb{B} \colon H \rightrightarrows H$.

Lyapunov perspective





- weak convergence and Lyapynov functions?
- Borwein–Sims: circle and line problem (non-convex)
- Benoist 2015: convergence proof for DRA of circle and line using Lyapunov function
- Dao–Tam 2019: convergence proof for DRA of line and graph of a function (non-convex) using Lyapunov function
- Gildadi–R 2019: convergence proof for DRA of line and two lines (non-convex) using Lyapunov function



Abstract dynamical system

A map $\phi\colon \mathbb{N} \times H \to H$ is a *dynamical system* if

1. $\phi(0, x) = x$ for all $x \in H$;
2. $\phi(n, \phi(k, x)) = \phi(n + k, x)$ for all $n, k \in \mathbb{N}$ and $x \in H$;
3. ϕ is continuous in the sense that $x_k \to x$ and $n \in \mathbb{N}$ implies $\phi(n, x_k) \to \phi(n, x)$.

Associated with a dynamical system is the difference equation

$$x^+ = Tx \tag{10}$$

with $Tx := \phi(1, x)$, and conversely, (10) with continuous T gives rise to the dynamical system $\phi(n, x) = T^n x$.

Notation

- The orbit of x is denoted by $\mathcal{O}(x) := \{\phi(n,x) \colon n \in \mathbb{N}\}$ and it is forward invariant, i.e. $T\mathcal{O}(x) \subset \mathcal{O}(x)$.
- The limit set $\omega(x) := \bigcap_{n\geq 0} \overline{\mathcal{O}\big(\phi(n,x)\big)} = \big\{z \in H \colon \exists\{n_j\}_{j\in\mathbb{N}} \subset \mathbb{N}, n_j \to \infty \text{ as } j \to \infty, \phi(n_j,x) \to z\big\}$ is closed and forward invariant.
- If $\mathcal{O}(x)$ is precompact, then $\omega(x)$ is nonempty, precompact, and invariant, i.e., $T\omega(x) = \omega(x)$.

Lyapunov functions in the sense of LaSalle

Let $D \subset H$. A function $V \colon D \to \mathbb{R}$ is a *Lyapunov function* of (10) if

1. V is continuous; and
2. $\dot{V}(x) := V(Tx) - V(x) \leq 0$ for all $x \in D$.

$V(x) = -1$

Lyapunov functions in the sense of LaSalle

Let $D \subset H$. A function $V\colon D \to \mathbb{R}$ is a *Lyapunov function* of (10) if

1. V is continuous; and
2. $\dot{V}(x) := V(Tx) - V(x) \leq 0$ for all $x \in D$.

We define

$$E := \{x \in \overline{D}\colon \dot{V} = 0\}.$$

Lyapunov functions in the sense of LaSalle

Let $D \subset H$. A function $V\colon D \to \mathbb{R}$ is a *Lyapunov function* of (10) if

1. V is continuous; and
2. $\dot{V}(x) := V(Tx) - V(x) \leq 0$ for all $x \in D$.

We define

$$E := \{x \in \overline{D} \colon \dot{V} = 0\}.$$

Denote by M the largest invariant set in E,

Lyapunov functions in the sense of LaSalle

Let $D \subset H$. A function $V \colon D \to \mathbb{R}$ is a *Lyapunov function* of (10) if

1. V is continuous; and
2. $\dot{V}(x) := V(Tx) - V(x) \leq 0$ for all $x \in D$.

We define

$$E := \{x \in \overline{D} \colon \dot{V} = 0\}.$$

Denote by M the largest invariant set in E, and by

$$V^{-1}(c) = \{x \in D \colon V(x) = c\}.$$

LaSalle's invariance principle

Theorem ([LaS76])

$x^t = \Gamma x$

Let $V : D \to \mathbb{R}$ be a Lyapunov function for (10). If $\mathcal{O}(x)$ is precompact and $\mathcal{O}(x) \subset D$, then $\omega(x) \subset M \cap V^{-1}(c)$ for some $c = c(x)$.

Back to the projection algorithms

- Both algorithms are of the form (10) with T continuous as the composition of continuous operators.
- For the method of alternating projections, both P_A and P_B are firmly nonexpansive, so $T = P_A P_B$ is nonexpansive as well as strictly quasinonexpansive and $\operatorname{Fix} T = A \cap B$.
- For the Douglas–Rachford Algorithm, $T = T_{A,B}$ is firmly nonexpansive, as $R_B R_A$ is nonexpansive as composition of nonexpansive operators, so in particular, T is strictly quasinonexpansive. The fixed point set of T can be characterised quite precisely, however, here we simply note that in general it is **not equal** to $A \cap B$.

Theorem

Let T be strictly quasinonexpansive and Fix T *be nonempty. The function* $V\colon H \to \mathbb{R}$ *given by* $V(x) := \operatorname{dist}(x, \operatorname{Fix} T)$ *is a Lyapunov function for* (10), *positive definite with respect to* Fix T, *and* $E = M = V^{-1}(0) = \operatorname{Fix} T$.

$x^+ = Tx$

Theorem

Let T be strictly quasinonexpansive and $\operatorname{Fix} T$ *be nonempty. The function* $V\colon H \to \mathbb{R}$ *given by* $V(x) := \operatorname{dist}(x, \operatorname{Fix} T)$ *is a Lyapunov function for* (10), *positive definite with respect to* $\operatorname{Fix} T$, *and* $E = M = V^{-1}(0) = \operatorname{Fix} T$.

Proof.

- V is defined on $D = H$, which is closed and convex, so that $\operatorname{Fix} T$ is also closed and convex.

Theorem

Let T be strictly quasinonexpansive and $\operatorname{Fix} T$ be nonempty. The function $V\colon H \to \mathbb{R}$ given by $V(x) := \operatorname{dist}(x, \operatorname{Fix} T)$ is a Lyapunov function for (10), positive definite with respect to $\operatorname{Fix} T$, and $E = M = V^{-1}(0) = \operatorname{Fix} T$.

Proof.

- V is defined on $D = H$, which is closed and convex, so that $\operatorname{Fix} T$ is also closed and convex.
- Both the norm and $P_{\operatorname{Fix} T}$ are continuous, so $V(x) = \|x - P_{\operatorname{Fix} T} x\|$ is continuous as well.

Theorem

Let T be strictly quasinonexpansive and $\operatorname{Fix} T$ *be nonempty. The function* $V\colon H \to \mathbb{R}$ *given by* $V(x) := \operatorname{dist}(x, \operatorname{Fix} T)$ *is a Lyapunov function for* (10), *positive definite with respect to* $\operatorname{Fix} T$, *and* $E = M = V^{-1}(0) = \operatorname{Fix} T$.

Proof.

- V is defined on $D = H$, which is closed and convex, so that $\operatorname{Fix} T$ is also closed and convex.
- Both the norm and $P_{\operatorname{Fix} T}$ are continuous, so $V(x) = \|x - P_{\operatorname{Fix} T}x\|$ is continuous as well.
- Clearly, $V(x) = 0$ if and only if $x \in \operatorname{Fix} T$, and $V(x) > 0$ for $x \notin \operatorname{Fix} T$.

Theorem

Let T be strictly quasinonexpansive and $\operatorname{Fix} T$ *be nonempty. The function* $V\colon H \to \mathbb{R}$ *given by* $V(x) := \operatorname{dist}(x, \operatorname{Fix} T)$ *is a Lyapunov function for* (10)*, positive definite with respect to* $\operatorname{Fix} T$*, and* $E = M = V^{-1}(0) = \operatorname{Fix} T$*.*

Proof.

- V is defined on $D = H$, which is closed and convex, so that $\operatorname{Fix} T$ is also closed and convex.
- Both the norm and $P_{\operatorname{Fix} T}$ are continuous, so $V(x) = \|x - P_{\operatorname{Fix} T} x\|$ is continuous as well.
- Clearly, $V(x) = 0$ if and only if $x \in \operatorname{Fix} T$, and $V(x) > 0$ for $x \notin \operatorname{Fix} T$.
- Let $x \notin \operatorname{Fix} T$ and $y \in \operatorname{Fix} T$, then $V(Tx) = \operatorname{dist}(Tx, \operatorname{Fix} T) = \|Tx - P_{\operatorname{Fix} T} Tx\| \leq \|Tx - y\| = \|Tx - Ty\| < \|x - y\|$, where inequality comes from (6).

Theorem

Let T be strictly quasinonexpansive and $\operatorname{Fix} T$ be nonempty. The function $V\colon H \to \mathbb{R}$ given by $V(x) := \operatorname{dist}(x, \operatorname{Fix} T)$ is a Lyapunov function for (10), positive definite with respect to $\operatorname{Fix} T$, and $E = M = V^{-1}(0) = \operatorname{Fix} T$.

Proof.

- V is defined on $D = H$, which is closed and convex, so that $\operatorname{Fix} T$ is also closed and convex.
- Both the norm and $P_{\operatorname{Fix} T}$ are continuous, so $V(x) = \|x - P_{\operatorname{Fix} T} x\|$ is continuous as well.
- Clearly, $V(x) = 0$ if and only if $x \in \operatorname{Fix} T$, and $V(x) > 0$ for $x \notin \operatorname{Fix} T$.
- Let $x \notin \operatorname{Fix} T$ and $y \in \operatorname{Fix} T$, then $V(Tx) = \operatorname{dist}(Tx, \operatorname{Fix} T) = \|Tx - P_{\operatorname{Fix} T} Tx\| \leq \|Tx - y\| = \|Tx - Ty\| < \|x - y\|$, where inequality comes from (6).
  In particular, for $y = P_{\operatorname{Fix} T} x \in \operatorname{Fix} T$, we have $V(Tx) < \|x - P_{\operatorname{Fix} T} x\| = \operatorname{dist}(x, \operatorname{Fix} T) = V(x)$. This establishes that $E = V^{-1}(0) = \operatorname{Fix} T$.

Theorem

Let T be strictly quasinonexpansive and $\operatorname{Fix} T$ *be nonempty. The function* $V\colon H \to \mathbb{R}$ *given by* $V(x) := \operatorname{dist}(x, \operatorname{Fix} T)$ *is a Lyapunov function for* (10)*, positive definite with respect to* $\operatorname{Fix} T$*, and* $E = M = V^{-1}(0) = \operatorname{Fix} T$*.*

Proof.

- V is defined on $D = H$, which is closed and convex, so that $\operatorname{Fix} T$ is also closed and convex.
- Both the norm and $P_{\operatorname{Fix} T}$ are continuous, so $V(x) = \|x - P_{\operatorname{Fix} T}x\|$ is continuous as well.
- Clearly, $V(x) = 0$ if and only if $x \in \operatorname{Fix} T$, and $V(x) > 0$ for $x \notin \operatorname{Fix} T$.
- Let $x \notin \operatorname{Fix} T$ and $y \in \operatorname{Fix} T$, then $V(Tx) = \operatorname{dist}(Tx, \operatorname{Fix} T) = \|Tx - P_{\operatorname{Fix} T}Tx\| \leq \|Tx - y\| = \|Tx - Ty\| < \|x - y\|$, where inequality comes from (6).
  In particular, for $y = P_{\operatorname{Fix} T}x \in \operatorname{Fix} T$, we have $V(Tx) < \|x - P_{\operatorname{Fix} T}x\| = \operatorname{dist}(x, \operatorname{Fix} T) = V(x)$. This establishes that $E = V^{-1}(0) = \operatorname{Fix} T$.
- As $\operatorname{Fix} T$ is invariant under T, we clearly have $M = E$. $

Corollary

Let T be strictly quasinonexpansive and $\operatorname{Fix} T \neq \emptyset$. *Then every orbit $\mathcal{O}(x)$ of (10) is bounded. If, in addition, $\mathcal{O}(x)$ is confined to a finite dimensional subspace of H, then it is precompact, so that* $\omega(x) = \operatorname{Fix} T$, *and in fact,* $T^n x \to \operatorname{Fix} T$ *as* $n \to \infty$.

Theorem

Let T be nonexpansive and $\operatorname{Fix} T \neq \emptyset$*. The function* $V \colon H \to \mathbb{R}$ *given by* $V(x) := \|x - Tx\|$ *is a Lyapunov function for* (10) *and positive definite with respecect to* $\operatorname{Fix} T$.

Theorem

Let T be nonexpansive and $\operatorname{Fix} T \neq \emptyset$*. The function* $V\colon H \to \mathbb{R}$ *given by* $V(x) := \|x - Tx\|$ *is a Lyapunov function for* (10) *and positive definite with respecet to* $\operatorname{Fix} T$*.*

Proof.

- Again V is continuous as a composition of continuous functions and operators.
- Clearly $V(x) = 0$ if $x = Tx$ and positive otherwise.
- For $x \notin \operatorname{Fix} T$ we compute $V(Tx) = \|Tx - T^2x\| \leq \|x - Tx\| = V(x)$, using (2). $\$$

$E = \{ x : V(Tx) - V(x) = 0 \}$

What about Hundal's example of MAP failing strong convergence?

- H infinite dimensional Hilbert space with orthonormal basis $\{e_i\}_{i\in\mathbb{N}}$
- Define

$$\begin{aligned} A &:= \{x \in H \colon \langle x, e_1\rangle \leq 0\} \text{ and} \\ B &:= \overline{\text{cone}}\{p(t) \colon t \geq 0\}, \text{ with} \end{aligned}$$

$$\begin{aligned} p(t) &:= e_{\lfloor t\rfloor+2} \cos f(t) + e_{\lfloor t\rfloor+3} \sin f(t) + e_1 e^{-100t^3} \text{ and} \\ f(t) &:= \frac{\pi}{2}(x - \lfloor x\rfloor). \end{aligned}$$

- A is a nonempty closed halfspace, hence convex.
- B is the closure of the conical hull of a spiraling sequence of points $p(t)$, so convex as well.
- With starting point $x_0 = p(1)$, Hundal then demonstrates through a sequence of technical lemmas that the orbit $\mathcal{O}(x_0)$ contains a subsequence that remains close to the sequence of points $p(n)$, $n \geq 1$, which all have norm close to 1. However, the intersection $A \cap B = \{0\}$, so clearly $T^n x_0 \not\to 0$ in norm as $n \to \infty$. On the other hand, $T^n x_0 \rightharpoonup 0$ by Theorem 2.

Details: [Hun04]

What about weak convergence?

Lemma

Let the sequence $\{x_n\}_{n\in\mathbb{N}}$ be weakly convergent to $x \in H$, but not strongly. Then the sequence is not precompact.

Proof. As $\{x_n\}_{n\in\mathbb{N}}$ does not converge strongly to x, there must exist and $\epsilon > 0$ and a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ so that $\|x_{n_k} - x\| > \epsilon$ for all $k \geq 0$. Now we argue by contradiction and assume that $\{x_n\}_{n\in\mathbb{N}}$ is precompact. Hence $\{x_{n_k}\}_{k\in\mathbb{N}}$ is precompact as well and thus admits a convergent subsequence $\{x_{n_{k_l}}\}_{l\in\mathbb{N}}$, say $x_{n_{k_l}} \to y$ as $l \to \infty$. By uniqueness of weak limits, we must have $y = x$. However, this contradicts that the subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ stays away by an ϵ from x, a contradiction. Hence $\{x_n\}_{n\in\mathbb{N}}$ is not precompact. $\$$

Summary

- apparent discrepancy between weak convergence results and strong convergence implied by Lyapunov theory
- Lyapunov theory only applies to precompact trajectories, ruling out counterexamples
- Lyapunov theory applies in finite dimensional settings

Main references

Heinz H. Bauschke and Patrick L. Combettes.
Convex analysis and monotone operator theory in Hilbert spaces.
Springer, Cham, 2nd edition, 2017.

Hein S. Hundal.
An alternating projection that does not converge in norm.
Nonlin. Anal. Th. Meth. Appl., 57(1):35–61, 2004.

J. P. LaSalle.
The stability of dynamical systems.
Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1976.
With an appendix: "Limiting equations and stability of nonautonomous ordinary differential equations" by Z. Artstein, Regional Conference Series in Applied Mathematics.