A General Approach for Convergence Analysis of Adaptive Sampling-Based Signal Processing

Holger Boche[1] and Ullrich J. Mönich[2]

[1]Technische Universität München
Lehrstuhl für Theoretische Informationstechnik

[2]Massachusetts Institute of Technology
Research Laboratory of Electronics

May 26, 2015
11th International Conference on Sampling Theory and Applications
(SampTA 2015)

Outline

Adaptive vs. Non-Adaptive Approximation

Non-Adaptive Approximation:

Shannon sampling series

$$(U_N f)(t) = \sum_{k=-N}^{N} f(k) \frac{\sin(\pi(t-k))}{\pi(t-k)}$$

Approximation $U_N f$ uses all samples $\{f(k)\}_{k=-N}^{N}$.

Adaptive Approximation:

Selection of specific samples $f(k)$ for the approximation.

Consequences:

- Non-Adaptive Approximation: sequence of linear operators. Convergence analysis with Banach–Steinhaus theorem (uniform boundedness principle)
- Adaptive Approximation: leads to non-linear operators.

Signal Spaces

- $\mathcal{B}_\sigma$ is the set of all entire functions f with the property that for all $\epsilon > 0$ there exists a constant $C(\epsilon)$ with $|f(z)| \leqslant C(\epsilon)\exp\big((\sigma+\epsilon)|z|\big)$ for all $z \in \mathbb{C}$.

Definition (Bernstein Space)

The Bernstein space $\mathcal{B}_\sigma^p$ consists of all signals in $\mathcal{B}_\sigma$, whose restriction to the real line is in $L^p(\mathbb{R})$, $1 \leqslant p \leqslant \infty$.

Definition (Paley-Wiener Space)

For $1 \leqslant p \leqslant \infty$ we denote by $\mathcal{PW}_\sigma^p$ the Paley-Wiener space of signals f with a representation $f(z) = \frac{1}{2\pi}\int_{-\sigma}^{\sigma} g(\omega)\, e^{iz\omega}\, d\omega$, $z \in \mathbb{C}$, for some $g \in L^p[-\sigma,\sigma]$.

The norm for $\mathcal{PW}_\sigma^p$ is given by $\|f\|_{\mathcal{PW}_\sigma^p} = \left(\frac{1}{2\pi}\int_{-\sigma}^{\sigma} |\hat{f}(\omega)|^p\, d\omega\right)^{1/p}$.

Residual Set

- A subset M of a metric space X is said to be nowhere dense in X if the closure $[M]$ does not contain a non-empty open set of X. M is said to be of the first category (or meager) if M is the countable union of sets each of which is nowhere dense in X.
- A set that is not of the first category is called a set of the second category.
- The complement of a set of the first category is called a residual set.

Baire's Theory:

- Residual sets are "large".
- In a complete metric space, a residual set is dense and a set of the second category.
- The countable intersection of residual sets is always a residual set.

Divergence of the Shannon Sampling Series

- For signals $f \in \mathcal{PW}_{\pi}^{p}$, $1 < p < \infty$, the Shannon sampling series converges absolutely and uniformly on all of $\mathbb{R}$.

Divergence of the Shannon Sampling Series

- For signals $f \in \mathcal{PW}_\pi^p$, $1 < p < \infty$, the Shannon sampling series converges absolutely and uniformly on all of $\mathbb{R}$.
- However, for $p = 1$, i.e., for $f \in \mathcal{PW}_\pi^1$ we have

$$\limsup_{N \to \infty} \left(\max_{t \in \mathbb{R}} \sum_{k=-N}^{N} f(k) \frac{\sin(\pi(t-k))}{\pi(t-k)} \right) = \infty, \tag{1}$$

  that is the peak value diverges as N tends to infinity.

Divergence of the Shannon Sampling Series

- For signals $f \in \mathcal{PW}_\pi^p$, $1 < p < \infty$, the Shannon sampling series converges absolutely and uniformly on all of $\mathbb{R}$.
- However, for $p = 1$, i.e., for $f \in \mathcal{PW}_\pi^1$ we have

$$\limsup_{N\to\infty} \left(\max_{t\in\mathbb{R}} \sum_{k=-N}^{N} f(k) \frac{\sin(\pi(t-k))}{\pi(t-k)} \right) = \infty, \qquad (1)$$

  that is the peak value diverges as N tends to infinity.

- Recently strengthened in [BF14] : There exists a signal $f \in \mathcal{PW}_\pi^1$ such that

$$\lim_{N\to\infty} \left(\max_{t\in\mathbb{R}} \sum_{k=-N}^{N} f(k) \frac{\sin(\pi(t-k))}{\pi(t-k)} \right) = \infty. \qquad (2)$$

Important difference in the divergence behavior of (1) and (2)

[BF14] H. Boche and B. Farrell, "Strong divergence of reconstruction procedures for the Paley-Wiener space PW_π^1 and the Hardy space H^1," *Journal of Approximation Theory*, vol. 183, pp. 98–117, Jul. 2014

Weak Divergence vs. Strong Divergence

We say a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$

- diverges weakly if $\limsup_{n\to\infty}|x_n| = \infty$.
- diverges strongly if $\lim_{n\to\infty}|x_n| = \infty$.

Weak lim sup divergence:

Merely guarantees the existence of a subsequence $\{N_n\}_{n\in\mathbb{N}}$ for which we have $\lim_{n\to\infty} x_{N_n} = \infty$.
Leaves the possibility that there exist a different subsequences $\{N_n^*\}_{n\in\mathbb{N}}$ such that $\limsup_{n\to\infty} x_{N_n^*} < \infty$.

Strong lim divergence:

Divergence for all subsequences $\{N_n\}_{n\in\mathbb{N}}$.

Weak Divergence vs. Strong Divergence

We say a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$

- diverges weakly if $\limsup_{n\to\infty}|x_n| = \infty$.
- diverges strongly if $\lim_{n\to\infty}|x_n| = \infty$.

Weak lim sup divergence:
Merely guarantees the existence of a subsequence $\{N_n\}_{n\in\mathbb{N}}$ for which we have $\lim_{n\to\infty} x_{N_n} = \infty$.
Leaves the possibility that there exist a different subsequences $\{N_n^*\}_{n\in\mathbb{N}}$ such that $\limsup_{n\to\infty} x_{N_n^*} < \infty$.

Strong lim divergence:
Divergence for all subsequences $\{N_n\}_{n\in\mathbb{N}}$.

adaptive techniques not convergent $\Leftrightarrow$ strong divergence

Banach–Steinhaus Theorem, Weak Divergence, and Residual Sets

Divergence results as in

$$\limsup_{N\to\infty}\left(\max_{t\in\mathbb{R}}\sum_{k=-N}^{N} f(k)\frac{\sin(\pi(t-k))}{\pi(t-k)}\right)=\infty$$

are usually proved by using the Banach–Steinhaus theorem (uniform boundedness principle).

- → The obtained divergence is in terms of the $\limsup$ (weak divergence) and not a statement about strong divergence.
- → Strength of the Banach–Steinhaus theorem: the divergence statement holds for all functions from a residual set.

Approximation Processes in Banach Spaces

General and abstract setting: two Banach spaces, B_1 and B_2, and a bounded linear operator $T\colon B_1 \to B_2$.

Goal: approximate T by a sequence of bounded linear operators $\{U_N\}_{N\in\mathbb{N}}$ (with a simpler structure, e.g., finite dimensional range, ...), mapping from B_1 in B_2.

Assumptions

1. There exists a dense subset S_1 of B_1 such that

$$\lim_{N\to\infty} \|U_N f - Tf\|_{B_2} = 0$$

for all $f \in S_1$.

2. We have $\lim_{N\to\infty} \|U_N\| = \infty$.

Discussion of Assumptions

❶ $\lim_{N\to\infty} \|U_N f - Tf\|_{B_2} = 0$ for all $f \in \mathcal{S}_1$.

❷ $\lim_{N\to\infty} \|U_N\| = \infty$.

Remarks:

- For the Shannon sampling series and $\mathcal{PW}_\pi^1$ assumption ❶ is fulfilled. ($\mathcal{PW}_\pi^2$ is a dense subset of $\mathcal{PW}_\pi^1$.)

Discussion of Assumptions

1. $\lim_{N\to\infty} \|U_N f - Tf\|_{B_2} = 0$ for all $f \in S_1$.
2. $\lim_{N\to\infty} \|U_N\| = \infty$.

Remarks:

- $\limsup_{N\to\infty} \|U_N\| = \infty$ is necessary.
  If $\limsup_{N\to\infty} \|U_N\| < \infty \Rightarrow$ convergence for all $f \in B_1$, nothing needs to be analyzed.
- If $\limsup_{N\to\infty} \|U_N\| = \infty$ then the set

$$D_{BS} = \left\{ f \in B_1 \colon \limsup_{N\to\infty} \|U_N f\|_{B_2} = \infty \right\}$$

is a residual set in B_1.

Discussion of Assumptions

❶ $\lim_{N\to\infty} \|U_N f - Tf\|_{B_2} = 0$ for all $f \in \mathcal{S}_1$. ❷ $\lim_{N\to\infty} \|U_N\| = \infty$.

Remarks:

- If $\liminf_{N\to\infty} \|U_N\| < \infty$ then there exists a universal subsequence $\{N_l\}_{l\in\mathbb{N}}$ of the natural numbers such that

  $$\lim_{l\to\infty} \|U_{N_l} f - Tf\|_{B_2} = 0$$

  for all $f \in B_1$.
- Adaptive signal processing would lead to convergence for the whole space B_1.
- Hence, interesting case is ❷.

Set of Signals with Strong Divergence

We analyze the set

$$\mathrm{D}_{\mathsf{strong}} = \left\{ f \in \mathcal{B}_1 \colon \lim_{N\to\infty} \|U_N f\|_{\mathcal{B}_2} = \infty \right\}.$$

$\mathrm{D}_{\mathsf{strong}}$ contains all the signals for which adaptive signal processing is useless.

No Universal Subsequence

- ❷ $\Rightarrow$ there cannot exist a universal subsequence $\{N_l\}_{l \in \mathbb{N}}$ such that $\lim_{l \to \infty} \|U_{N_l} f - Tf\|_{B_2} = 0$ is true for all $f \in B_1$.

Relevant question:

- Is possible to find for every signal $f \in B_1$ a subsequence $\{N_l(f)\}_{l \in \mathbb{N}}$ of the natural numbers such that $\lim_{l \to \infty} \|U_{N_l} f - Tf\|_{B_2} = 0$ is true?
- In this case the subsequence $\{N_l(f)\}_{l \in \mathbb{N}}$ is adapted to the signal f.

answer negative $\quad \Leftrightarrow \quad D_{\text{strong}} \neq \emptyset$

Observation

Observation

Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two Banach spaces and $T\colon \mathcal{B}_1 \to \mathcal{B}_2$ a bounded linear operator. Further, let $\{U_N\}_{N\in\mathbb{N}}$ be a sequence of bounded linear operators mapping from $\mathcal{B}_1$ to $\mathcal{B}_2$ such that ❶ and ❷ are fulfilled. If $\mathcal{D}_{\text{strong}} \neq \emptyset$ then $\mathcal{D}_{\text{strong}}$ is dense in $\mathcal{B}_1$.

Set of Signals without Strong Divergence

Set of signals for which the approximation process diverges and adaptive signal processing leads to a bounded approximation process:

$$D_{\mathsf{sb}} = D_{\mathsf{BS}} \setminus D_{\mathsf{strong}}$$

$$D_{\mathsf{sb}} = \left\{ f \in B_1 \colon \limsup_{N \to \infty} \|U_N f\|_{B_2} = \infty \text{ and } \liminf_{N \to \infty} \|U_N f\|_{B_2} < \infty \right\}$$

Set of Signals without Strong Divergence

Set of signals for which the approximation process diverges and adaptive signal processing leads to a bounded approximation process:

$$\mathrm{D}_{\mathsf{sb}} = \mathrm{D}_{\mathsf{BS}} \setminus \mathrm{D}_{\mathsf{strong}}$$

$$\mathrm{D}_{\mathsf{sb}} = \left\{ f \in \mathrm{B}_1 \colon \limsup_{N\to\infty} \|U_N f\|_{\mathrm{B}_2} = \infty \text{ and } \liminf_{N\to\infty} \|U_N f\|_{\mathrm{B}_2} < \infty \right\}$$

Theorem

Let B_1 and B_2 be two Banach spaces and $T\colon \mathrm{B}_1 \to \mathrm{B}_2$ a bounded linear operator. Further, let $\{U_N\}_{N\in\mathbb{N}}$ be a sequence of bounded linear operators mapping from B_1 to B_2 such that ① and ② are fulfilled. Then D_{sb} is a residual set in B_1.

Set of Signals with Strong Divergence

Corollary

Let B_1 and B_2 be two Banach spaces and $\mathrm{T}\colon \mathrm{B}_1 \to \mathrm{B}_2$ a bounded linear operator. Further, let $\{\mathrm{U}_\mathrm{N}\}_{\mathrm{N}\in\mathbb{N}}$ be a sequence of bounded linear operators mapping from B_1 to B_2 such that ❶ and ❷ are fulfilled. Then the set D_{strong} is either empty or a meager set.

Examples that Fit into the Framework

1. Shannon sampling series (global convergence):
   $B_1 = \mathcal{PW}_\pi^1$, $B_2 = \mathcal{B}_\pi^\infty$, $T = \mathrm{Id}$,

$$(U_N f)(t) = \sum_{k=-N}^{N} f(k) \frac{\sin(\pi(t-k))}{\pi(t-k)}.$$

2. System approximation process (global convergence):
   $B_1 = \mathcal{PW}_\pi^1$, $B_2 = \mathcal{B}_\pi^\infty$, $T\colon \mathcal{PW}_\pi^1 \to \mathcal{PW}_\pi^1$ a stable LTI system,

$$(U_N f)(t) = \sum_{k=-N}^{N} f(k) h_T(t-k),$$

   where $h_T(t) = (T\,\mathrm{sinc})(t)$.

3. System approximation process (pointwise convergence at $t \in \mathbb{R}$):
   $B_1 = \mathcal{PW}_\pi^1$, $B_2 = \mathbb{C}$, $T\colon \mathcal{PW}_\pi^1 \to \mathbb{C}$, $f \mapsto (\tilde{T} f)(t)$, where $\tilde{T}\colon \mathcal{PW}_\pi^1 \to \mathcal{PW}_\pi^1$ is a stable LTI system.

4. Non-equidistant sampling series and system approximation processes.

Examples

Strong divergence for Shannon sampling series and peak value of the system approximation process for the Hilbert transform.

In both cases, the set of signals with strong divergence can be at most a meager set.

Stable Linear Time Invariant Systems

A linear system $T : \mathcal{PW}_\pi^1 \to \mathcal{PW}_\pi^1$ is called stable linear time invariant (LTI) system if:

- T is bounded, i.e., $\|T\| = \sup_{\|f\|_{\mathcal{PW}_\pi^1} \leqslant 1} \|Tf\|_{\mathcal{PW}_\pi^1} < \infty$ and
- T is time invariant, i.e., $\big(Tf(\cdot - a)\big)(t) = (Tf)(t - a)$ for all $f \in \mathcal{PW}_\pi^1$ and $t, a \in \mathbb{R}$.

The Hilbert transform H and the low-pass filter are stable LTI systems.

Example (Hilbert transform)

The Hilbert transform Hf of a signal $f \in \mathcal{PW}_\pi^1$ is defined by

$$(Hf)(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} -i \operatorname{sgn}(\omega) \hat{f}(\omega) \, e^{i\omega t} \, d\omega,$$

where sgn denotes the signum function.

Representation of Stable LTI Systems

- For every stable LTI system $T : \mathcal{PW}_\pi^1 \to \mathcal{PW}_\pi^1$ there is exactly one function $\hat{h}_T \in L^\infty[-\pi, \pi]$ such that

$$(Tf)(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{h}_T(\omega) \hat{f}(\omega) \, e^{i\omega t} \, d\omega$$

  for all $f \in \mathcal{PW}_\pi^1$, and the integral is absolutely convergent.
- Every $\hat{h}_T \in L^\infty[-\pi, \pi]$ defines a stable LTI system $T : \mathcal{PW}_\pi^1 \to \mathcal{PW}_\pi^1$.

The operator norm $\|T\| := \sup_{\|f\|_{\mathcal{PW}_\pi^1} \leqslant 1} \|Tf\|_{\mathcal{PW}_\pi^1}$ is given by $\|T\| = \|\hat{h}_T\|_\infty$.

System Approximation

Approximate the system output Tf from the samples of f.

$$\sum_{k=-\infty}^{\infty} f(k) h_T(t-k).$$

Convergence is more problematic than the convergence behavior of the Shannon sampling series.

Abbreviation: $$(T_N f)(t) = \sum_{k=-N}^{N} f(k) h_T(t-k).$$

Theorem (BM10)

For all $t \in \mathbb{R}$ *there exists stable LTI system* $T \colon \mathcal{PW}_\pi^1 \to \mathcal{PW}_\pi^1$ *and a signal* $f \in \mathcal{PW}_\pi^1$ *such that*

$$\limsup_{N\to\infty} |(T_N f)(t)| = \infty.$$

Even true for oversampling and arbitrary choice of the reconstruction kernel.

[BM10] H. Boche and U. J. Mönich, "Sampling-type representations of signals and systems," *Sampling Theory in Signal and Image Processing*, vol. 9, no. 1–3, pp. 119–153, Jan., May, Sep. 2010

System Approximation

Divergence is only weak $\limsup$ divergence (no strong divergence).

Adaptive choice of a subsequence $\{N_n\}_{n\in\mathbb{N}}$ may lead to a convergent approximation process.

The subsequence will in general depend on the signal f.
$\rightarrow T_{N_n(f)}$ would be adapted to the signal f.

Task of adaptive signal processing:
Find an index sequence (depending on the signal f) that leads to convergence.

System Approximation

Theorem

Let $T\colon \mathcal{PW}_\pi^1 \to \mathcal{PW}_\pi^1$ *be a stable LTI system,* $t \in \mathbb{R}$*, and* $f \in \mathcal{PW}_\pi^1$*. There exists a monotonically increasing subsequence* $\{N_k = N_k(t, f, T)\}_{k\in\mathbb{N}}$ *of the natural numbers such that* $\lim_{k\to\infty}(T_{N_k} f)(t) = (Tf)(t)$*.*

- In certain cases, adaptive signal processing leads to an approximation process that is convergent for all $f \in \mathcal{PW}_\pi^1$.
- Shows that strong divergence is a stronger statement than assumption ❷ ($\lim_{N\to\infty}\|U_N\| = \infty$).

[BM10] H. Boche and U. J. Mönich, “Strong divergence for system approximations,” *Problems of Information Transmission*, 2015, to be published

Set of Signals with a Convergent Subsequence

Set of signals for which the approximation process diverges and adaptive signal processing leads to a convergent approximation process:

$$\mathcal{D}_{\mathsf{sc}} = \left\{ f \in \mathcal{B}_1 \colon \limsup_{N\to\infty} \|U_N f\|_{\mathcal{B}_2} = \infty \text{ and } \liminf_{N\to\infty} \|U_N f - Tf\|_{\mathcal{B}_2} = 0 \right\}.$$

- $\mathcal{D}_{\mathsf{sc}} \subset \mathcal{D}_{\mathsf{sb}}$
- We already know that $\mathcal{D}_{\mathsf{sb}}$ is a residual set.

Set of Signals with a Convergent Subsequence

Set of signals for which the approximation process diverges and adaptive signal processing leads to a convergent approximation process:

$$D_{sc} = \left\{ f \in B_1 \colon \limsup_{N\to\infty} \|U_N f\|_{B_2} = \infty \text{ and } \liminf_{N\to\infty} \|U_N f - Tf\|_{B_2} = 0 \right\}.$$

- $D_{sc} \subset D_{sb}$
- We already know that D_{sb} is a residual set.

The next theorem shows that even D_{sc} is a residual set.

Theorem

Let B_1 and B_2 be two Banach spaces and $T\colon B_1 \to B_2$ a bounded linear operator. Further, let $\{U_N\}_{N\in\mathbb{N}}$ be a sequence of bounded linear operators mapping from B_1 to B_2 such that ❶ and ❷ are fulfilled. Then D_{sc} is a residual set in B_1.

D_{sb} is a residual set: Main Steps of the Proof I

$$D_{sb} = \left\{ f \in B_1 \colon \limsup_{N\to\infty} \|U_N f\|_{B_2} = \infty \text{ and } \liminf_{N\to\infty} \|U_N f\|_{B_2} < \infty \right\}$$

$$D_2 = \left\{ f \in B_1 \colon \limsup_{N\to\infty} \|U_N f\|_{B_2} = \infty \text{ and } \liminf_{N\to\infty} \|U_N f\|_{B_2} \leqslant \|Tf\|_{B_2} + 1 \right\}$$

We will show that D_2 is a residual set. Since $D_2 \subset D_{sb}$, this implies that D_{sb} is also a residual set.

For $M, N, K \in \mathbb{N}$ we consider the set

$$D(M,N,K) = \left\{ f \in B_1 \colon \|U_N f\|_{B_2} > K \text{ and } \|U_M f\|_{B_2} < \|Tf\|_{B_2} + 1 \right\}.$$

We first prove that $D(M,N,K)$ is an open set. Then, we prove that for all $N_0, K \in \mathbb{N}$ the set

$$\widetilde{D}(N_0, K) = \bigcup_{M,N \geqslant N_0} D(M,N,K) \tag{3}$$

is dense in B_1.

D_{sb} is a residual set: Main Steps of the Proof II

Since $\widetilde{D}(N_0, K)$ is the union of open sets, it follows that $\widetilde{D}(N_0, K)$ is open.
Thus, we have established that $\widetilde{D}(N_0, K)$ is an open set that is dense in B_1.
This is true for all N_0 and K in $\mathbb{N}$. It follows that

$$D_3 = \bigcap_{K=1}^{\infty} \bigcap_{N_0=1}^{\infty} \widetilde{D}(N_0, K) \tag{4}$$

is a residual set in B_1.
Let $f \in D_3$ be arbitrary but fixed. Form (3) and (4) we see that for every $N_0, K \in \mathbb{N}$ there exist natural numbers $N_{N_0,K}$ and $M_{N_0,K}$ satisfying $\min\{N_{N_0,K}, M_{N_0,K}\} \geqslant N_0$, $\|U_{N_{N_0,K}} f\|_{B_2} > K$, and $\|U_{M_{N_0,K}} f\|_{B_2} < \|Tf\|_{B_2} + 1$.
It follows that

$$\limsup_{N\to\infty} \|U_N f\|_{B_2} = \infty$$

and

$$\liminf_{N\to\infty} \|U_N f\|_{B_2} \leqslant \|Tf\|_{B_2} + 1,$$

that is, we have $f \in D_2$. This shows that $D_3 \subset D_2$, and since D_3 is a residual set, it follows that D_2 and consequently D_{sb} is a residual set.

Conclusion

- Weak divergence $\rightarrow$ full theory given by Banach–Steinhaus (residual set)
- Examples for strong divergence: Shannon sampling series, Hilbert transform approximation
- Strong divergence $\Leftrightarrow$ adaptive techniques not convergent
- Strong divergence at most for a meager set
- Showed examples for approximation processes that are weakly divergent but not strongly
- Strong divergence stronger statement than $\lim_{N\to\infty}\|U_N\| = \infty$
- With adaptivity, reduction from a residual set to a meager set possible

Thank you!