The Kolmogorov Superposition Theorem:

A framework for multivariate computation

Jonas Actor

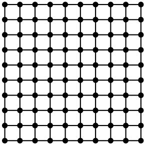
Rice University

7 May 2018

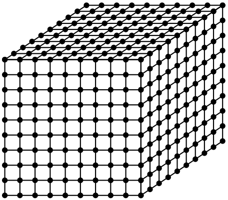
Contact: jonasactor@rice.edu

Curse of Dimensionality

10 nodes



10^2 nodes



10^3 nodes

Cost of simulation / computation grows **exponentially**

A Question [1]

Do functions of three variables exist at all?

[1] Pólya and Szegö, Problems and Theorems of Analysis, 1925 (German), transl. 1945, reprinted 1978

A Question, Restated

Can functions of three variables be expressed using functions of only two variables?

Example:
Cardano's Formula for roots of a cubic equation

A Better Question

Hilbert's 13th Problem (1900)

Can the solution x to the 7th degree polynomial equation

$$x^7 + ax^3 + bx^2 + cx + 1 = 0$$

be represented by a finite number of compositions of bivariate continuous (algebraic) functions using the three variables a, b, c?

Answer: Arnol'd 1957 [2]

Any **continuous** function of three variables can be expressed using **continuous** functions of only two variables.

[2] Arnol'd, Dokl. Akad. Nauk SSSR 114:5, 1957

The Next Question

Can continuous functions of three variables be expressed using functions of only **one** variable and addition?

Answer: Kolmogorov 1957[3]

Any continuous $f : [0,1]^n \to \mathbb{R}$ can be written as

$$f(x_1, \ldots, x_n) = \sum_{q=0}^{2n} \chi \left(\sum_{p=1}^{n} \psi_{p,q}(x_p) \right)$$

[3] Kolmogorov, Dokl. Akad. Nauk SSSR 114:5, 1957

What's going on here?

$$f(x_1, \ldots, x_n) = \sum_{q=0}^{2n} \chi\left(\sum_{p=1}^{n} \psi_{p,q}(x_p) \right)$$

What's going on here?

$$f(x_1, \ldots, x_n) = \sum_{q=0}^{2n} \chi\left(\sum_{p=1}^{n} \psi_{p,q}(x_p) \right)$$

Original function $f : [0,1]^n \to \mathbb{R}$

What's going on here?

$$f(x_1, \ldots, x_n) = \sum_{q=0}^{2n} \chi\left(\sum_{p=1}^{n} \psi_{p,q}(x_p) \right)$$

Inner function $\psi_{p,q} : [0, 1] \to \mathbb{R}$

What's going on here?

$$f(x_1, \ldots, x_n) = \sum_{q=0}^{2n} \chi\left(\sum_{p=1}^{n} \psi_{p,q}(x_p) \right)$$

Addition

What's going on here?

$$f(x_1, \dots, x_n) = \sum_{q=0}^{2n} \chi\left(\sum_{p=1}^{n} \psi_{p,q}(x_p) \right)$$

Outer function $\chi : \mathbb{R} \to \mathbb{R}$

What's going on here?

$$f(x_1, \ldots, x_n) = \sum_{q=0}^{2n} \chi\left(\sum_{p=1}^{n} \psi_{p,q}(x_p) \right)$$

Function composition

Implication

$$\Psi^q(x_1, \ldots, x_n) = \sum_{p=1}^{n} \psi_{p,q}(x_p)$$

(independent of f)

KST: $f \longmapsto \chi$

No Free Lunch...

Traded smoothness for variables:

- No KST[4] for $\psi_{p,q} \in C^1([0,1])$
- Current $\psi_{p,q} \in$ Hölder$([0,1]) \to$ bad for computation
- Possible $\psi_{p,q} \in$ Lip$([0,1])$

[4] Vituskin, DAN, 95:701–704, 1954.

Smoothness

Goal: Construct Lipschitz functions $\psi_{\mathbf{p},\mathbf{q}}$

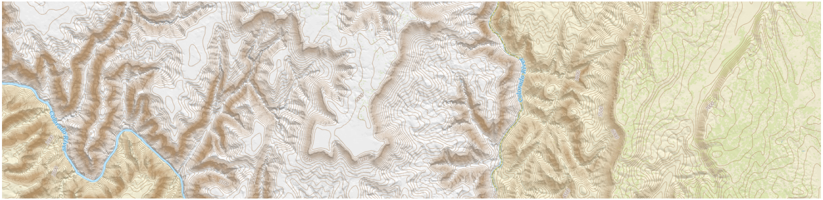
Rest of the Talk

1. Constructive KST
2. The Fridman Strategy
3. Reparameterization Approach
4. Conclusions and Outstanding Tasks

Table of Contents

An Analogy

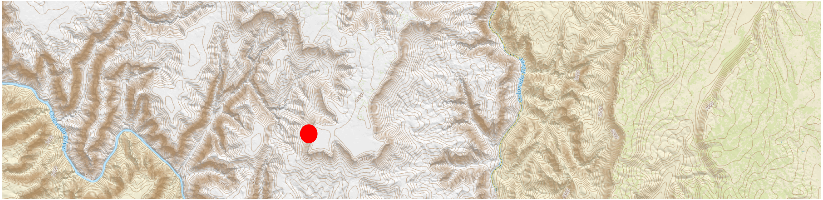
$$f(x, y) = \text{ elevation of Grand Canyon at } (x, y)$$



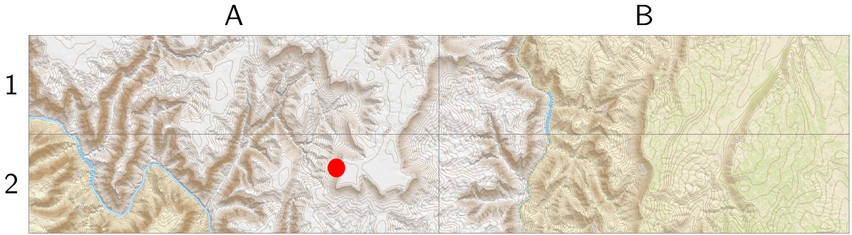
https://viewer.nationalmap.gov/basic/

An Analogy

$$f(\text{Thor's Temple}) = \ ?$$



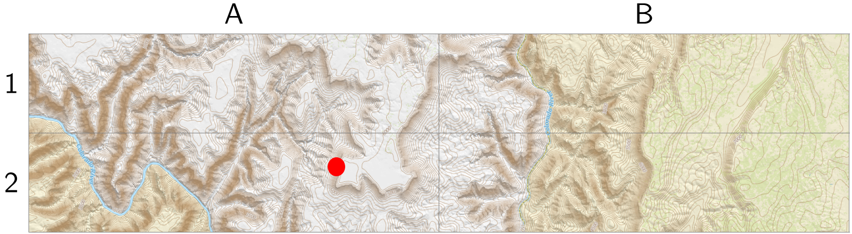
An Analogy



Bright Angel Point	A1	Kaibab Nat'l Forest	B1
Desert View	B2	Thor's Temple	A2
Grand Canyon Lodge	A1	Walhalla's Overlook	A2

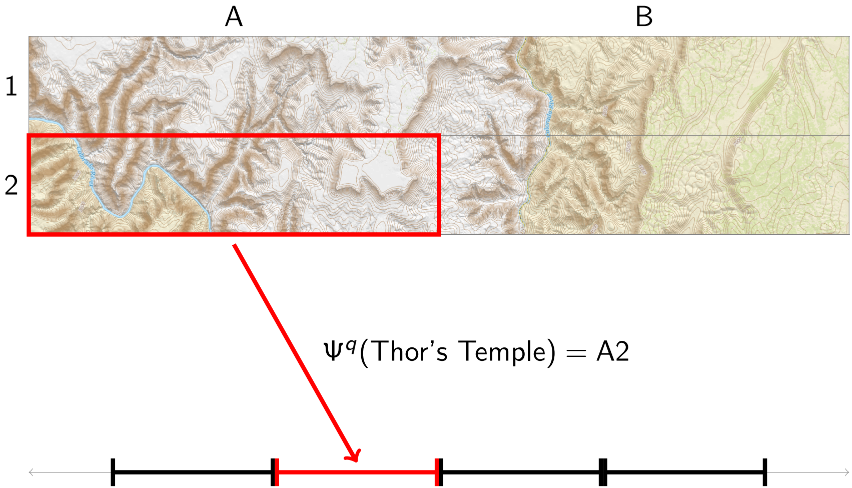
Locations real; elevations are not.

An Analogy



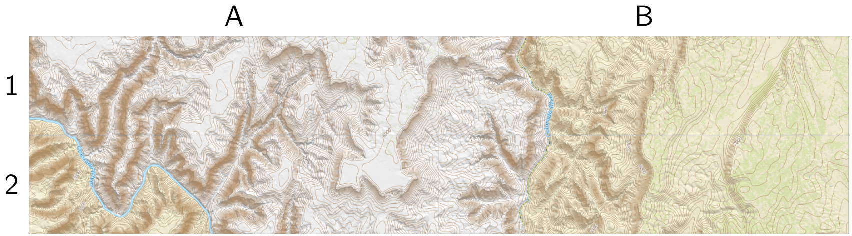
Elevation (ft)	A1	8200	B1	8000
	A2	9000	B2	7600

An Analogy

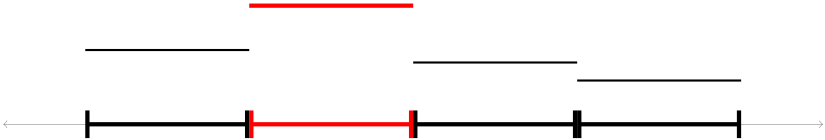


$\Psi^q(\text{Thor's Temple}) = \text{A2}$

An Analogy



$$f(\text{Thor's Temple}) = \chi(\text{A2})$$



An Analogy

A B

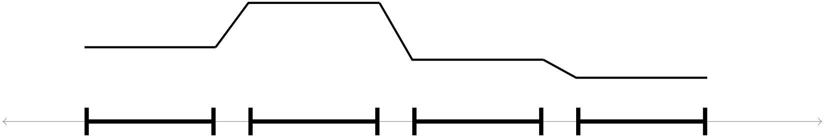
1

2

An Analogy

A B

1
2

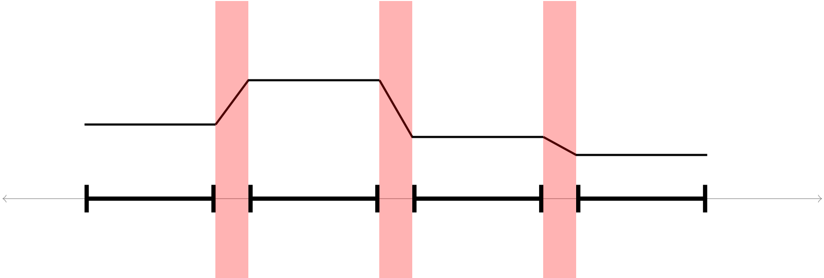


An Analogy

A B

1

2



Decomposition Process

Strategy

- Leave gaps between squares $\rightarrow$ gives room to vary continuously
- Duplicate and slightly shift the squares to cover the gaps
- Define functions $\psi_{p,q}$ as we refine squares

Decisions

- How big to make the gaps
- How to assign values to functions

Sprecher's Reduction[5]

Define 1 function instead of $2n^2 + n$ functions:

$$\psi_{p,q}(x_p) = \lambda_p\,\psi(x_p + q\varepsilon)$$

$\lambda_1, \ldots, \lambda_n$ integrally independent

[5] Sprecher, Trans. AMS, 115:340–355, 1965

Requirements (1/2)

Refinement:

Squares $\mathcal{S}^k$ get smaller as $k \to \infty$

More Than Half:

Each point in at least $\frac{n+1}{2n+1}$ sets of shifted squares

Requirements (2/2)

Disjoint Image:

Image of squares under Ψ^q disjoint

Monotonicity:

Function ψ strictly monotonic increasing

Bounded Slope:

Function ψ Lipschitz

Table of Contents

Background

First proof of Lipschitz inner KST functions[6]

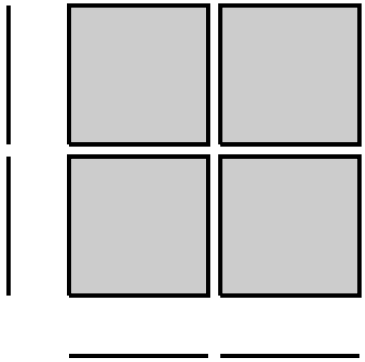
Predominant approach for Lipschitz KST functions

Never successfully implemented

[6] Dokl. Akad. Nauk SSR, 177:5:1019–1022, 1967.

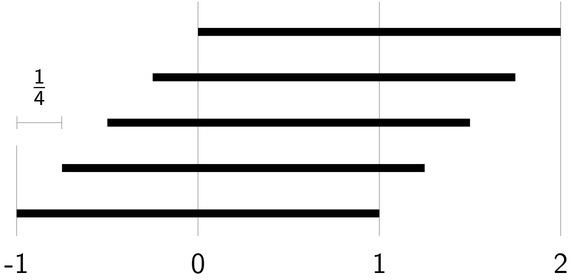
Defining our Squares

Squares: Cartesian product of 1D family of intervals

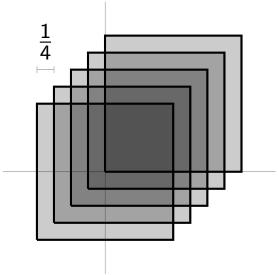


Space Decomposition

5 copies of the same family of intervals, shifted by $\frac{1}{4}$

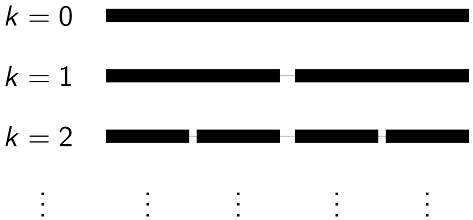


Cartesian product for each shifted family of intervals



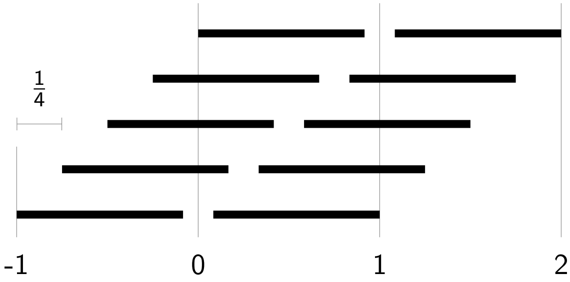
Space Decomposition

Each *refinement level* $k \in \mathbb{N}$,
break largest intervals roughly in half



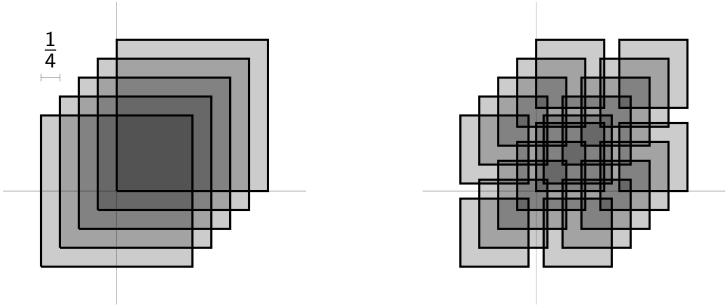
Space Decomposition

Breaks are copied in **each** shifted family of intervals . . .



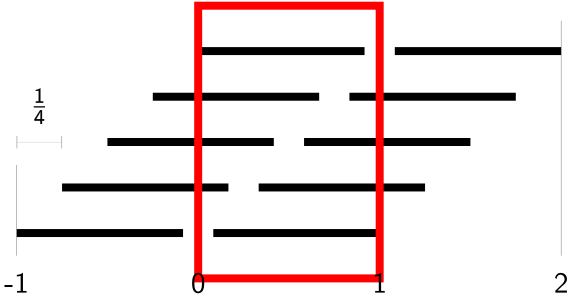
Space Decomposition

. . . and each shifted family of squares.



Requirements on Breaks

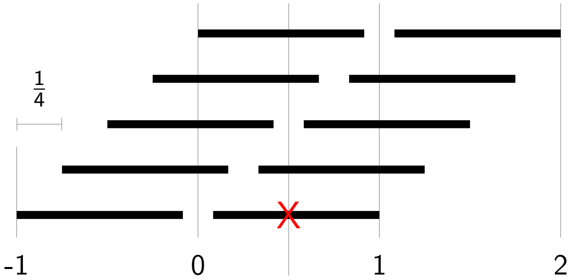
Every $x \in [0, 1]$ in at least **All But One** of shifted families



All But One on intervals $\iff$ More Than Half on squares

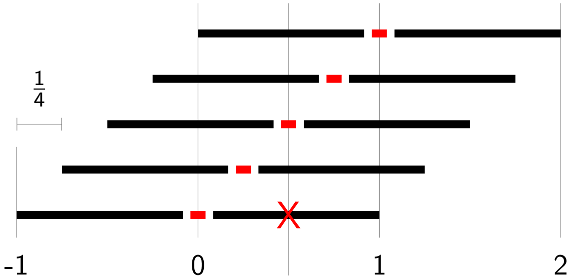
Choosing Breaks

Keep inserting breaks $\longrightarrow$ lose All But One Condition



Choosing Breaks

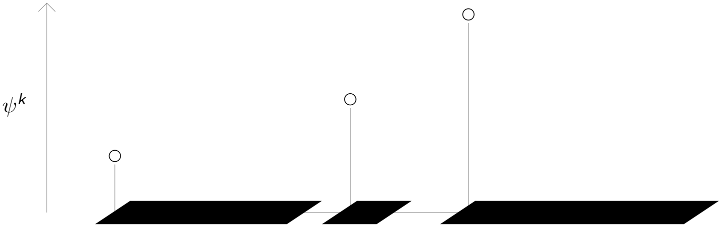
Solution: plug the gaps that cause problems



Inner Function ψ

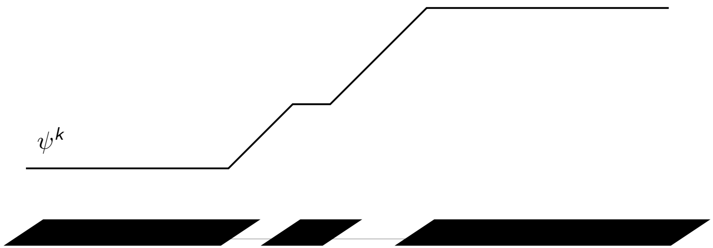
At each refinement level k:

- Assign a value of ψ^k on left endpoint of each interval
- Value is fixed for all future k



Inner Function ψ

ψ^k constant on intervals, linear on gaps



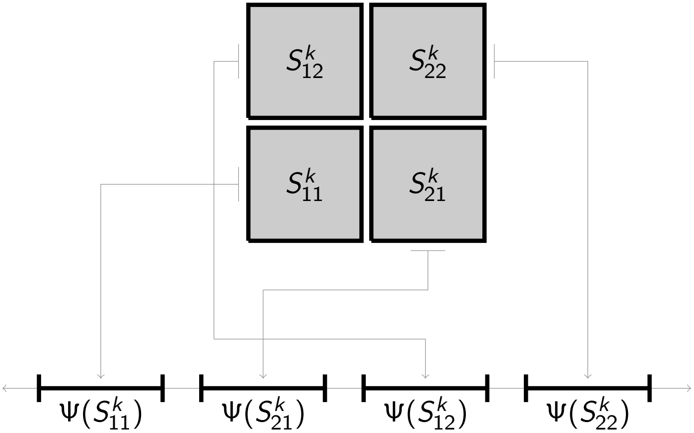
large plugs $\rightarrow$ small gaps $\rightarrow$ steep ψ^k

Fixing Values: Disjoint Image Condition

For any two squares S, $S' \in \mathcal{S}^k$,

$$\Psi(S) \cap \Psi(S') = \emptyset.$$

Disjoint Image Condition



Disjoint Image Condition

Don't know ψ at level k

Can't enforce Disjoint Image Condition without ψ

Two remedies:

- Fridman's Disjoint Image Condition
- Conservative Disjoint Image Condition

Fridman's Disjoint Image Condition

For any two squares S, $S' \in \mathcal{S}^k$,

$$\Psi^k(S) \cap \Psi^k(S') = \emptyset.$$

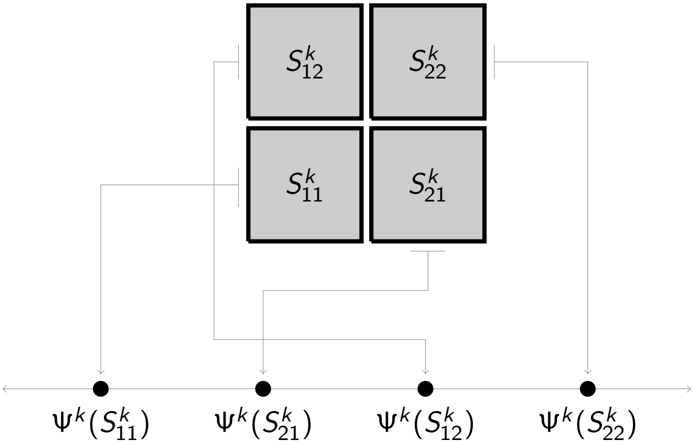
Fridman's Disjoint Image Condition

For any two squares S, $S' \in \mathcal{S}^k$,

$$\Psi^k(S) \cap \Psi^k(S') = \emptyset.$$

Not $\Psi(S)$, $\Psi(S')$

Fridman's Disjoint Image Condition



Fridman's Disjoint Image Condition

$$\Psi(x) = \sum_{p=1}^{n} \lambda_p \, \psi(x_p + q\varepsilon)$$

$\lambda_1, \ldots, \lambda_n$ integrally independent

ψ^k rational values at left endpoints

Fridman's Condition is always met!
Does not enforce disjointness in limit

Conservative Disjoint Image Condition

Use Lipschitz bound to define intervals $\Delta_{\mathbf{i}}^{k}$ that are guaranteed to contain $\Psi(S_{\mathbf{i}}^{k})$

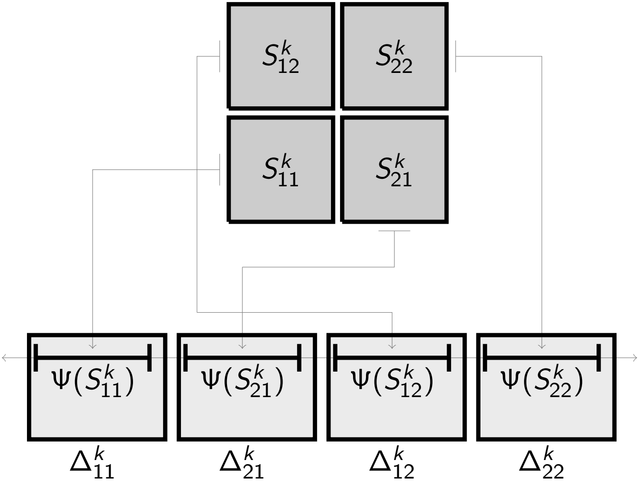
$$\Psi^{k}(S_{\mathbf{i}}^{k}) \subseteq \Psi(S_{\mathbf{i}}^{k}) \subseteq \Delta_{\mathbf{i}}^{k}.$$

Conservative Disjoint Image Condition

For any two squares $S^k_{\mathbf{i}}$, $S^k_{\mathbf{i}'} \in \mathcal{S}^k$
and corresponding intervals $\Delta^k_{\mathbf{i}}, \Delta^k_{\mathbf{i}'}$,

$$\Delta^k_{\mathbf{i}} \cap \Delta^k_{\mathbf{i}'} = \emptyset.$$

Conservative Disjoint Image Condition



Conservative Disjoint Image Condition

Requires removing **at least half** of each interval

Points not contained in enough shifted copies of $\mathcal{S}^k$ to reconstruct χ

Implications

Choose 2 of 3:

- (Conservative) Disjoint Image Condition
- All But One Condition
- Bounded Slope Condition

Table of Contents

Reasoning

KST inner functions are strictly monotonic increasing

Reasoning

KST inner functions are strictly monotonic increasing

. . . which are of Bounded Variation

Reasoning

KST inner functions are strictly monotonic increasing

. . . which are of Bounded Variation

. . . so they define rectifiable curves

Reasoning

KST inner functions are strictly monotonic increasing

. . . which are of Bounded Variation

. . . so they define rectifiable curves

. . . which have Lipschitz reparameterizations.

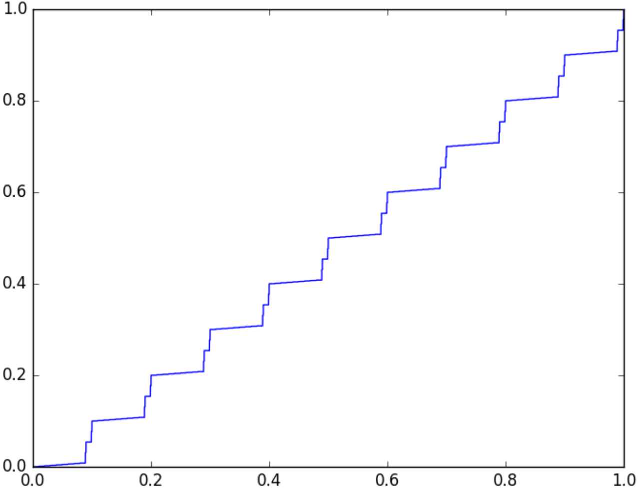
Approach

1. Find (non-Lipschitz) function $\widehat{\psi}$
2. Define reparameterization σ
3. Create Lipschitz ψ from $\widehat{\psi}$ using σ
4. Construct squares $\mathcal{S}^k$ that meet conditions

Hölder Continous $\widehat{\psi}$

- Define iteratively: $\widehat{\psi} = \lim_{k\to} \widehat{\psi}^k$
- Fix values at points with k digits in base γ expansion
    - Small increase for most points
    - Large increase for expansions ending in $\gamma - 1$
- Linearly interpolate between fixed values

Köppen's KST Inner Function $\widehat{\psi}^7$



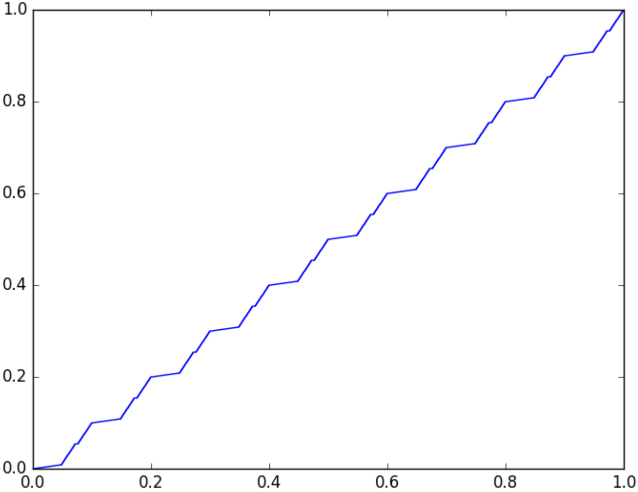
Lipschitz Reparameterization

Reparameterization $\sigma : [0,1] \to [0,1]$

$$\sigma(x) = \frac{\text{arclength of } \widehat{\psi} \text{ from } 0 \text{ to } x}{\text{total arclength of } \widehat{\psi} \text{ from } 0 \text{ to } 1}$$

$$\psi(x) = \widehat{\psi}(\sigma^{-1}(x))$$

Lipschitz Reparameterization ψ^7



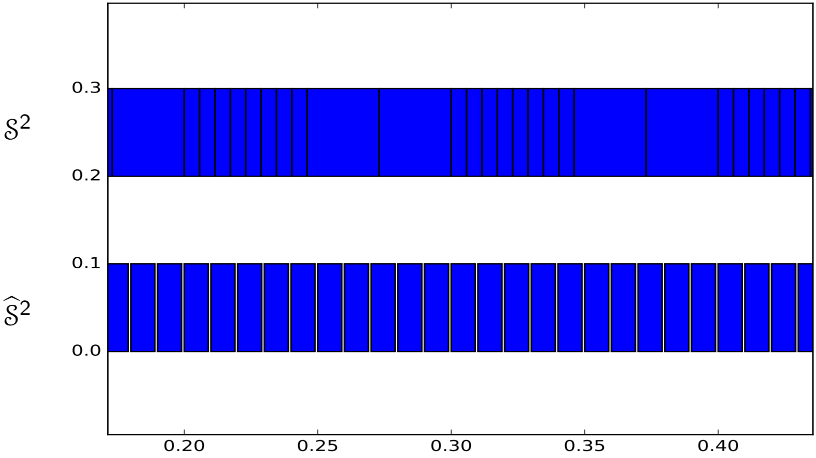
Correct Squares

$\psi = \widehat{\psi} \circ \sigma^{-1}$

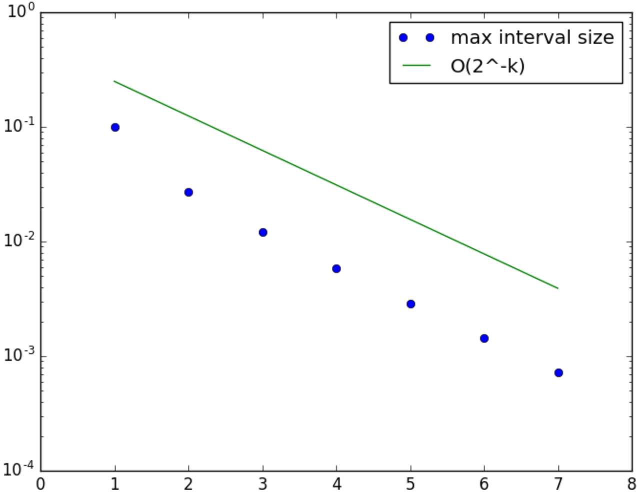
Intervals for $\widehat{\psi}$ are $\widehat{\mathcal{S}}^k$
Intervals for ψ are $\mathcal{S}^k = \sigma(\widehat{\mathcal{S}}^k)$

$\widehat{\psi}$ satisfies Disjoint Image Condition on $\widehat{\mathcal{S}}^k$
ψ satisfies Disjoint Image Condition on $\mathcal{S}^k$

Comparing $\widehat{\mathcal{S}}^k$ and $\mathcal{S}^k$



Largest Interval Size for $\mathcal{S}^k$



Reparameterized Intervals

Refinement Condition:
Intervals shrink at $O(2^{-k})$ not $O(\gamma^{-k})$

All But One Condition:
Gaps between intervals get smaller under σ

Takeaway

Function $\psi = \widehat{\psi} \circ \sigma^{-1}$ is a Lipschitz KST inner function

Squares $\mathcal{S}^k = \sigma(\widehat{\mathcal{S}}^k)$ are KST squares

Table of Contents

Review

- Motivated why KST is interesting, hard
- Illustrated how KST works
- Outlined Fridman Strategy
- Showed where Fridman Strategy fails
- Posed new reparameterization approach
- Justified reparameterization meets requirements

Outstanding Tasks

Complete proofs for reparameterization argument for ψ

Framework for computing outer function χ

- Requires squares at each level k
- Requires final Ψ

First Application: Image compression