Understanding Multivariate Computation using the Kolmogorov Superposition Theorem

Matthew Knepley

Computer Science and Engineering
University at Buffalo

Firedrake 2019, Durham University
Durham, UK   September 27, 2019

Argonne NATIONAL LABORATORY

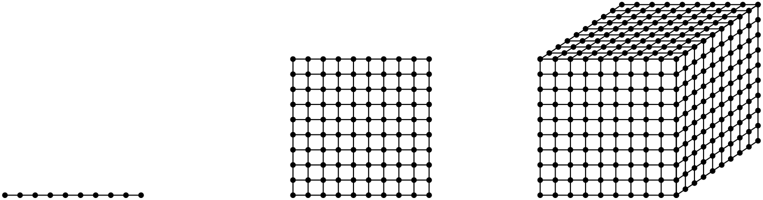
University at Buffalo

Collaborators

KST Representation Collaboration

Jonas Actor

Curse of Dimensionality



10 nodes

10^2 nodes

10^3 nodes

Cost of simulation / computation grows **exponentially**

Main Question

Do functions of three variables exist at all?[1]

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

Q: Can any function of three variables be expressed using functions of only two variables?[2]

[2] Hilbert, Göttinger Nachrichten, 1900

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

Q: Can any function of three variables be expressed using functions of only two variables?[2]

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

[2] Hilbert, Göttinger Nachrichten, 1900

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

Q: Can any function of three variables be expressed using functions of only two variables?[2]

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

A: Any *continuous* function of three variables can be expressed using *continuous* functions of only two variables.[3]

[2] Hilbert, Göttinger Nachrichten, 1900

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

Q: Can any function of three variables be expressed using only univariate functions and addition?

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

Q: Can any function of three variables be expressed using only univariate functions and addition?

A: Yes[4], for any continuous $f : [0,1]^n \to \mathbb{R}$

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

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

What's going on here?

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

What's going on here?

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

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

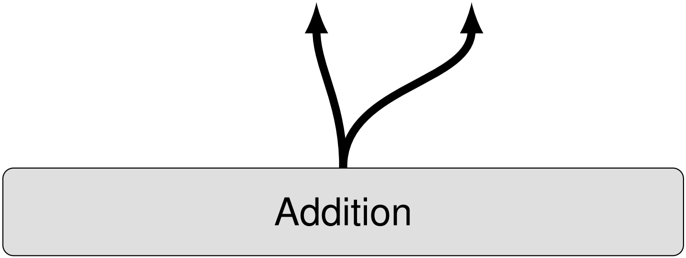
What's going on here?

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

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

What's going on here?

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



What's going on here?

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

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

What's going on here?

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

Function composition

2D: $n = 2$

$$\begin{aligned} f(x, y) = {} & \chi_0\left(\psi_{x,0}(x) + \psi_{y,0}(y)\right) \\ & + \chi_1\left(\psi_{x,1}(x) + \psi_{y,1}(y)\right) \\ & + \chi_2\left(\psi_{x,2}(x) + \psi_{y,2}(y)\right) \\ & + \chi_3\left(\psi_{x,3}(x) + \psi_{y,3}(y)\right) \\ & + \chi_4\left(\psi_{x,4}(x) + \psi_{y,4}(y)\right) \end{aligned}$$

2D: $n = 2$

$$
\begin{aligned}
f(x, y) = \chi_0\big(\lambda_x\psi(x) &+ \lambda_y\psi(y)\big) \\
+ \chi_1\big(\lambda_x\psi(x+\epsilon) &+ \lambda_y\psi(y+\epsilon)\big) \\
+ \chi_2\big(\lambda_x\psi(x+2\epsilon) &+ \lambda_y\psi(y+2\epsilon)\big) \\
+ \chi_3\big(\lambda_x\psi(x+3\epsilon) &+ \lambda_y\psi(y+3\epsilon)\big) \\
+ \chi_4\big(\lambda_x\psi(x+4\epsilon) &+ \lambda_y\psi(y+4\epsilon)\big)
\end{aligned}
$$

Single ψ function (Sprecher)

$$
\begin{aligned}
f(x, y) = &\chi\big(\lambda_x \psi(x) + \lambda_y \psi(y)\big) \\
&+ \chi\big(\lambda_x \psi(x + \epsilon) + \lambda_y \psi(y + \epsilon) + \delta\big) \\
&+ \chi\big(\lambda_x \psi(x + 2\epsilon) + \lambda_y \psi(y + 2\epsilon) + 2\delta\big) \\
&+ \chi\big(\lambda_x \psi(x + 3\epsilon) + \lambda_y \psi(y + 3\epsilon) + 3\delta\big) \\
&+ \chi\big(\lambda_x \psi(x + 4\epsilon) + \lambda_y \psi(y + 4\epsilon) + 4\delta\big)
\end{aligned}
$$

Single χ function (Lorentz) and ψ function (Sprecher)

Universality

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

is independent of f, so that

$$KST : f \to \chi$$

Simple Example

$$xy = \exp^{\ln x + \ln y}$$

Simple Example

$$xy = \exp^{\ln x + \ln y}$$

Modulus of continuity of *exp* compensates for flatness of *ln*

$$xy = \exp^{\ln x + \ln y}$$

Controlling regularity of ϕ is the key to efficiency

Regularity of Inner Functions

(Vitushkin 1964)
Inner functions ϕ cannot be smooth

$\mathcal{O}(n^2)$ cross-derviatives in f

$\mathcal{O}(n)$ derivatives in ϕ and χ

Regularity of Inner Functions

(Diaconis and Shahshahani 1984)
Exactly characterizes the nullspace for projection pursuit

$$f(x_1, \ldots, x_n) \approx \sum_{q=1}^{Q} \chi_q \left(\sum_{p=1}^{n} \lambda_{pq} x_p \right)$$

(Sprecher 1965)
Explicitly constructs a ϕ
that is Hölder$\left(\frac{\ln 2}{\ln 2n+2}\right)$

(Braün and Griebel 2009)
Demonstrates that such a function
cannot be computed efficiently

(Actor 2018)
Constructs a Lipschitz(2) ϕ

A remapping of Sprecher's ϕ
that is efficiently computable

General Approximations

- Need hierarchy
- Need nonlinearity
- Need to control regularity

Function composition is a powerful nonlinear mechanism missing from FEM

Thank You!

http://cse.buffalo.edu/~knepley