Introduction to Radial Basis Function Method

How to Interpolate Scattered Data with Radial Basis

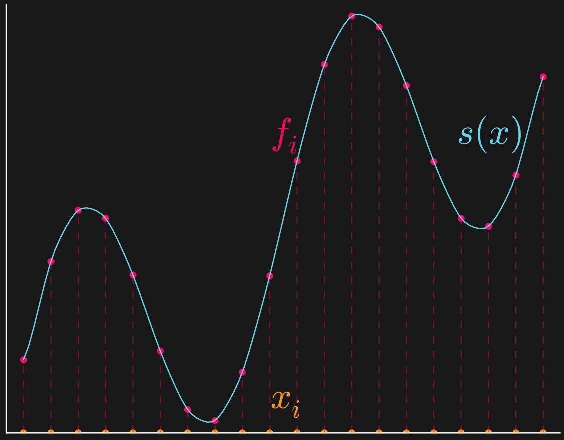
Jesse Bettencourt

McMaster University
Dr. Kevlahan

jessebett@gmail.com
github.com/jessebett

Motivation

Given a set of measurements $\{f_i\}_{i=1}^N$ taken at corresponding data sites $\{x_i\}_{i=1}^N$ we want to find an interpolation function $s(x)$ that informs us on our system at locations different from our data sites.



Motivation

Given a set of measurements $\{f_i\}_{i=1}^N$ taken at corresponding data sites $\{x_i\}_{i=1}^N$ we want to find an interpolation function $s(x)$ that informs us on our system at locations different from our data sites.

Examples of Data Sites and Measurments

1D: A series of temperature measurements over a time period

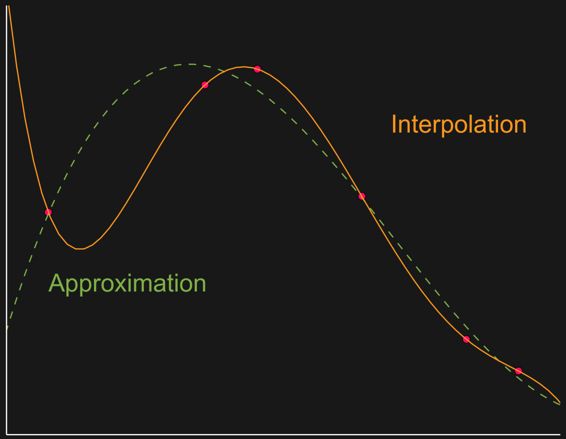
2D: Surface temperature of a lake based on measurements collected at sample surface locations

3D: Distribution of temperature within a lake

n-D: Machine learning, financial models, system optimization

What makes a good fit?

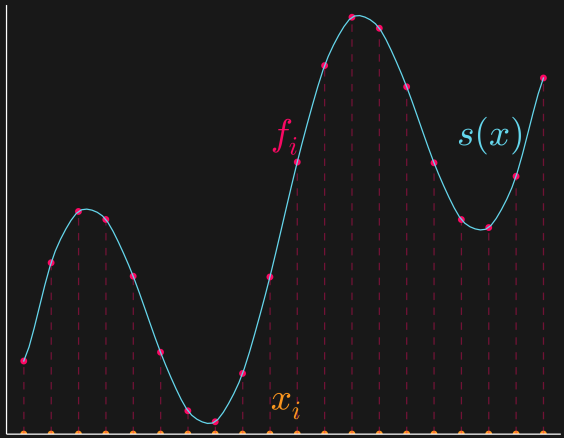
- Interpolation: $s(x)$ exactly matches our measurements at our data sites.
- Approximation: $s(x)$ closely matches our measurements at our data sites, e.g. with Least Squares



For today's purposes...

we will only consider interpolation.

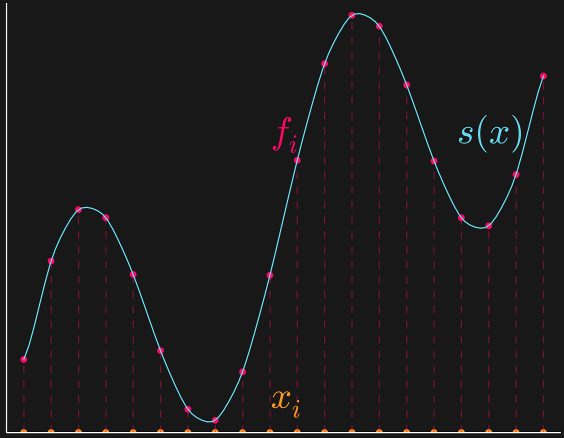
- Interpolation: $s(x_i) = f_i \; \forall i \in \{0 \ldots N\}$



Our Problem, Restated

Interpolation of Scattered Data

Given data (x_i, f_i), $i = 1, \ldots, N$, such that $x_i \in \mathbb{R}^n$, $f_i \in \mathbb{R}$, we want to find a continuous function $s(x)$ such that $s(x_i) = f_i$ $\forall i \in \{0 \ldots N\}$



A Familiar Approach

Convenient Assumtption

Assume $s(x)$ is a linear combination of basis functions ψ_i

$$s(x) = \sum_{i=1}^{N} \lambda_i \psi_i$$

A Familiar Approach

Convenient Assumtption

Assume $s(x)$ is a linear combination of basis functions ψ_i

$$s(x) = \sum_{i=1}^{N} \lambda_i \psi_i$$

Interpolation as a Linear System

Following this assumption we have a system of linear equations

$$A\boldsymbol{\lambda} = \mathbf{f}$$

where

A is called the interpolation matrix whose entries are given by

$$a_{ij} = \psi_j(x_i) \qquad\qquad i, j = 1 \dots N$$

and

$$\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_N]^T$$

$$\mathbf{f} = [f_1, \dots, f_N]^T$$

The Well-Posed Problem

$$A\boldsymbol{\lambda} = \mathbf{f}$$

Solving this linear system, thus finding $s(x)$, is only possible if the problem well-posed, i.e., $\exists$ a unique solution

The Well-Posed Problem

$$A\boldsymbol{\lambda} = \mathbf{f}$$

Solving this linear system, thus finding $s(x)$, is only possible if the problem well-posed, i.e., $\exists$ a unique solution

Result from introductory linear algebra:

The problem will be well-posed if and only if the interpolation matrix A is non-singular, i.e., $\det(A) \neq 0$.

The Well-Posed Problem

$$A\boldsymbol{\lambda} = \mathbf{f}$$

Solving this linear system, thus finding $s(x)$, is only possible if the problem well-posed, i.e., $\exists$ a unique solution

Result from introductory linear algebra:
The problem will be well-posed if and only if the interpolation matrix A is non-singular, i.e., $\det(A) \neq 0$.

Note: The non-singularity of A will depend on our choice of basis functions, $\psi_{i=1}^{N}$

Easily Well-Posed in 1D

In 1D, many choices of basis functions will guarantee a well-posed problem as long as the data-sites are distinct.

Easily Well-Posed in 1D

In 1D, many choices of basis functions will guarantee a well-posed problem as long as the data-sites are distinct.

Example

We are familiar with polynomial interpolation, interpolating from N data sites with a $(N-1)$-degree polynomial.

$$\psi_{i=1}^{N} = \{1, x, x^2, x^3, \ldots, x^{N-1}\}$$

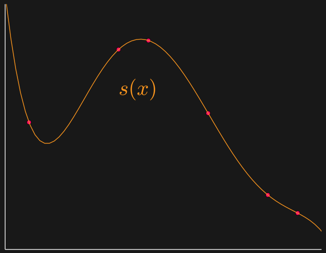
Easily Well-Posed in 1D

In 1D, many choices of basis functions will guarantee a well-posed problem as long as the data-sites are distinct.

Example

We are familiar with polynomial interpolation, interpolating from N data sites with a $(N-1)$-degree polynomial.

$$\psi_{i=1}^{N} = \{1, x, x^2, x^3, \ldots, x^{N-1}\}$$



$s(x) = -0.02988x^5 + 0.417x^4 - 2.018x^3 + 3.694x^2 - 1.722x - 5.511e^{-14}$

A Problem in Higher Dimensions

For n-Dimensions where $n \geq 2$ there is no such guarantee.

A Problem in Higher Dimensions

For n-Dimensions where $n \geq 2$ there is no such guarantee.

For any set of basis functions, $\psi_{i=1}^{N}$ (chosen independently of the data sites) $\exists$ a set of distinct data sites $\{x_i\}_{i=1}^{N}$ such that the interpolation matrix becomes singular.

A Problem in Higher Dimensions

For n-Dimensions where $n \geq 2$ there is no such guarantee.

For any set of basis functions, $\psi_{i=1}^{N}$ (chosen independently of the data sites) $\exists$ a set of distinct data sites $\{x_i\}_{i=1}^{N}$ such that the interpolation matrix becomes singular.

Implication: If we choose our basis functions independently of the data, we are not guaranteed a well-posed problem.

A Problem in Higher Dimensions

For n-Dimensions where $n \geq 2$ there is no such guarantee.

For any set of basis functions, $\psi_{i=1}^{N}$ (chosen independently of the data sites) $\exists$ a set of distinct data sites $\{x_i\}_{i=1}^{N}$ such that the interpolation matrix becomes singular.

Implication: If we choose our basis functions independently of the data, we are not guaranteed a well-posed problem.

Note: This results from the Haar-Mairhuber-Curtis Theorem

A Solution in Higher Dimensions

Implication: If we choose our basis functions independently of the data, we are not guaranteed a well-posed problem.

Solution?

A Solution in Higher Dimensions

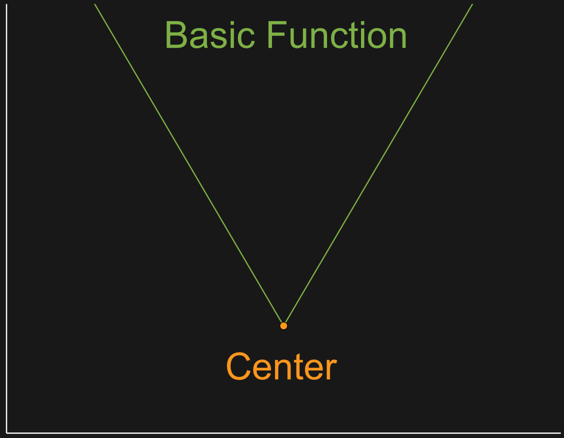
Implication: If we choose our basis functions independently of the data, we are not guaranteed a well-posed problem.

Solution? Choose basis functions depending on the data!

Basis Functions Depending on Data

First, consider what we call the basic function

$$\psi(x) = |x|$$



Basis Functions Depending on Data

First, consider what we call the basic function

$$\psi(x) = |x|$$

To produce our set of basis functions, we take translates of the basic function.

$$\psi_i(x) = ||x - x_i|| \qquad\qquad i = 1, \ldots, N$$

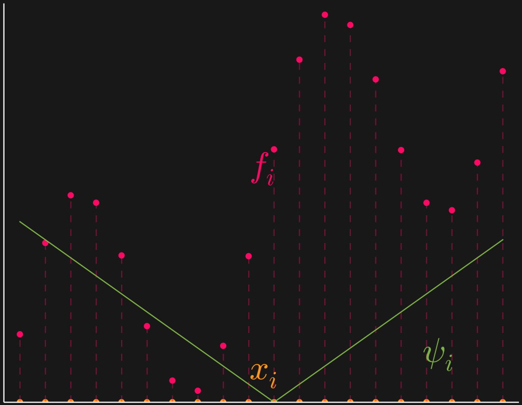
So each basis function, $\psi_i(x)$, is our basic function shifted so that the center or knot is positioned on a data site, x_i.

Note: It's possible to have other choices of centers, but in most implementations the centers coincide with data sites.

Basis Functions Depending on Data

Each basis function, $\psi_i(x)$, is our basic function shifted so that the center is positioned on a data site, x_i.

$$\psi_i(x) = ||x - x_i|| \qquad\qquad i = 1, \ldots, N$$



Radial Basis Functions

$$\psi_i(x) = ||x - x_i|| \qquad\qquad i = 1, \dots, N$$

Notice that $\psi_i(x)$ are radially symmetric about their centers, for this reason we call these functions Radial Basis Functions (RBF).

Since the basis functions only depend on distance, the interpolation matrix becomes

$$A = \begin{bmatrix} ||x_1 - x_1|| & ||x_1 - x_2|| & \cdots & ||x_1 - x_N|| \\ ||x_2 - x_1|| & ||x_2 - x_2|| & \cdots & ||x_2 - x_N|| \\ \vdots & \vdots & \ddots & \vdots \\ ||x_N - x_1|| & ||x_N - x_2|| & \cdots & ||x_N - x_N|| \end{bmatrix}$$

called a distance matrix.

The Distance Matrix

Distance matrices, with Euclidean distances, for distinct points in $\mathbb{R}^s$ are always non-singular.

This means that our interpolation problem

$$\begin{bmatrix} ||x_1 - x_1|| & ||x_1 - x_2|| & \cdots & ||x_1 - x_N|| \\ ||x_2 - x_1|| & ||x_2 - x_2|| & \cdots & ||x_2 - x_N|| \\ \vdots & \vdots & \ddots & \vdots \\ ||x_N - x_1|| & ||x_N - x_2|| & \cdots & ||x_N - x_N|| \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{bmatrix}$$

is well-posed!

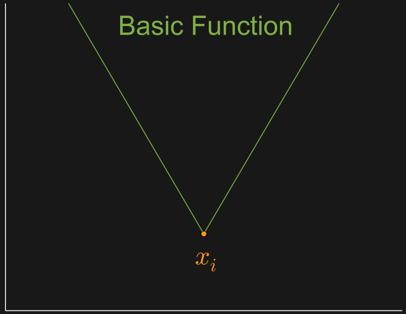
Our interpolant becomes $s(x) = \sum_{i=1}^{N} \lambda_i ||x - x_i||$

Building a Better Basic Function

Basic function

$$\psi_i(x) = ||x - x_i||$$

has a discontinuity in its first derivative at x_i.



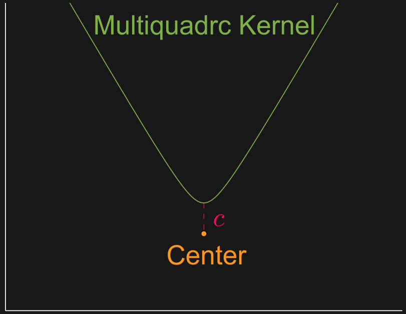
This causes the interpolant to have a discontinuous first derivative at each data site.
Obviously not ideal.

Building a Better Basic Function

In 1968, R.L. Hardy showed that we can remedy this problem by changing our basic function so it's C^∞.

Hardy's Multiquadrc Kernel

$$\psi(x) = \sqrt{c^2 + x^2} \qquad \text{where } c \neq 0.$$

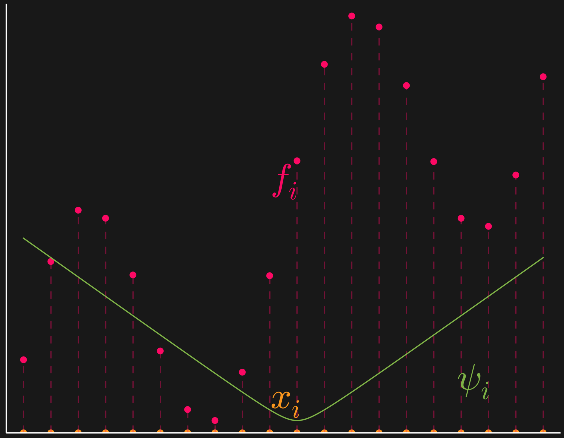


Note: c is called the shape parameter. The case when $c = 0$ is the previous basic function.

Radial Basis Kernels

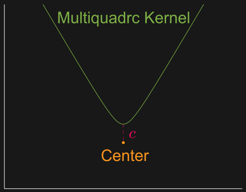
As before, we can generate our basis functions by translating Hardy's basic function to center on our data sites.

$$\psi_i(x) = \sqrt{c^2 + (||x - x_i||)^2}$$



Radial Basis Kernels

Hardy's Multiquadric function is still radially symmetric about its center

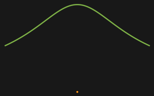
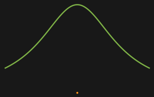
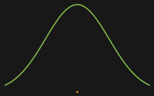


we this function a Kernel.
All Kernels are functions only of distance from center, and can be written generally as $\phi(||x - x_i||)$ or $\phi(r)$

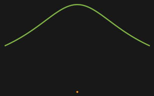
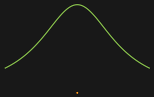
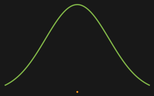
The RBF Method

$$s(x) = \sum_{i=1}^{N} \lambda_i \phi(||x - x_i||) = \sum_{i=1}^{N} \lambda_i \phi(r) \qquad r = ||x - x_i||$$

Radial Basis Kernels

Common RBF Kernels	$\phi(r)$	
Multiquadric	$\sqrt{1+(\epsilon r)^2}$	
Inverse Multiquadric	$\frac{1}{\sqrt{1+(\epsilon r)^2}}$	
Inverse Quadratic	$\frac{1}{1+(\epsilon r)^2}$	
Gaussian	$e^{-(\epsilon r)^2}$	

Radial Basis Kernels

Common RBF Kernels	$\phi(r)$	
Multiquadric	$\sqrt{1+(\epsilon r)^2}$	
Inverse Multiquadric	$\frac{1}{\sqrt{1+(\epsilon r)^2}}$	
Inverse Quadratic	$\frac{1}{1+(\epsilon r)^2}$	
Gaussian	$e^{-(\epsilon r)^2}$	

Note: One of these things is not like the others.

What About Well-Posed?

Our interpolation matrix is no longer the distance matrix. Can we still expect well-posed?

$$A = \begin{bmatrix} \phi_1(x_1) & \phi_1(x_2) & \cdots & \phi_1(x_N) \\ \phi_2(x_1) & \phi_2(x_2) & \cdots & \phi_2(x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(x_1) & \phi_N(x_2) & \cdots & \phi_N(x_N) \end{bmatrix}$$

What About Well-Posed?

Our interpolation matrix is no longer the distance matrix. Can we still expect well-posed?

$$A = \begin{bmatrix} \phi_1(x_1) & \phi_1(x_2) & \cdots & \phi_1(x_N) \\ \phi_2(x_1) & \phi_2(x_2) & \cdots & \phi_2(x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(x_1) & \phi_N(x_2) & \cdots & \phi_N(x_N) \end{bmatrix}$$

A New, Less Sexy Condition:

If interpolation matrix, A, is symetric positive-definite, then A is nonsingular and our system is well-posed.

Positive-Definite

Our matrix, A, is positive-definite if

$$t^T A t > 0 \qquad \forall t = [t_1, t_2, \ldots, t_n] \neq 0 \in \mathbb{R}^n$$

Which results from a positive-definite kernel, $\phi : \mathbb{R}^s \times \mathbb{R}^s \to \mathbb{R}$:

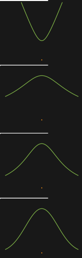
$$\sum_{i=1}^{N} \sum_{j=1}^{N} \phi(||x - x_i||) t_i \bar{t}_j > 0 \quad \forall t = [t_1, t_2, \ldots, t_n] \neq 0 \in \mathbb{C}^n$$

Useful Properties of Positive Definite Matricies

- All positive eigenvalues $\implies$ Non-Singular
- More efficent solving methods, e.g., Cholskey factorization

Radial Basis Kernels

Common RBF Kernels	$\phi(r)$
Multiquadric	$\sqrt{1+(\epsilon r)^2}$
Inverse Multiquadric	$\frac{1}{\sqrt{1+(\epsilon r)^2}}$
Inverse Quadratic	$\frac{1}{1+(\epsilon r)^2}$
Gaussian	$e^{-(\epsilon r)^2}$

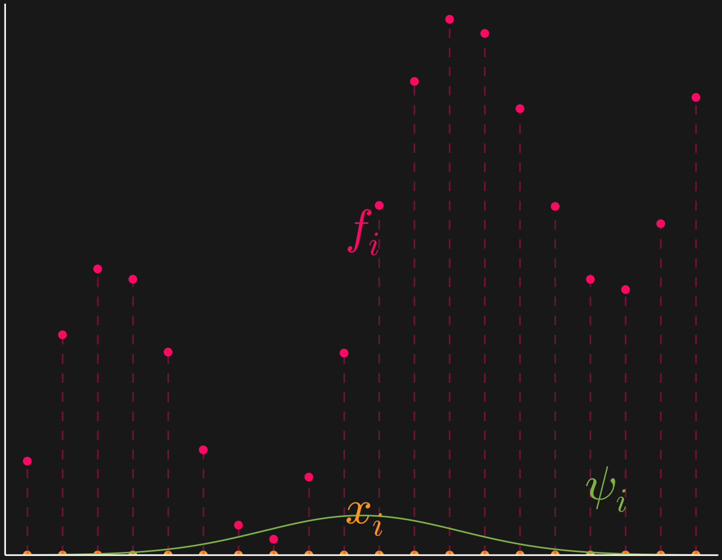


Hardy's Multiquadric Kernel is not positive-definite. However, it is conditionally negative-definite.

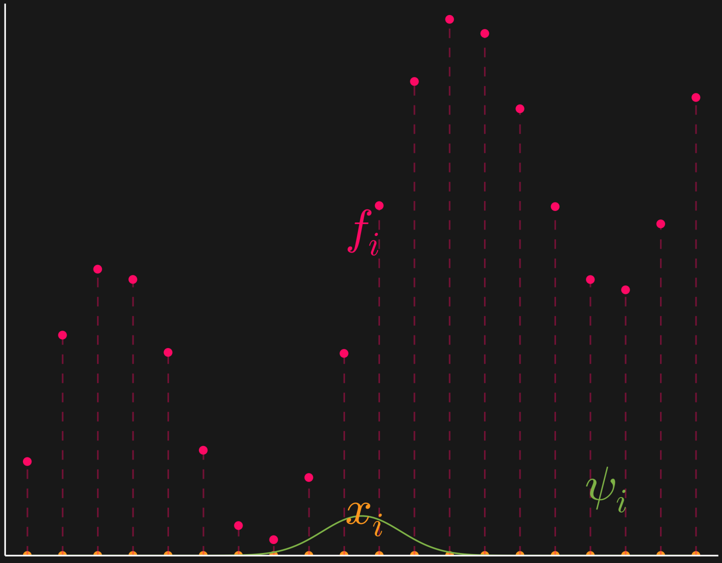
Properties of Multiquadric Matrix

- One positive eigenvalue and $(n-1)$ negative eigenvalues $\implies$ Non-Singular
- Not generally subject to positive-definite solving methods

Visualizing the RBF Method



Visualizing the RBF Method



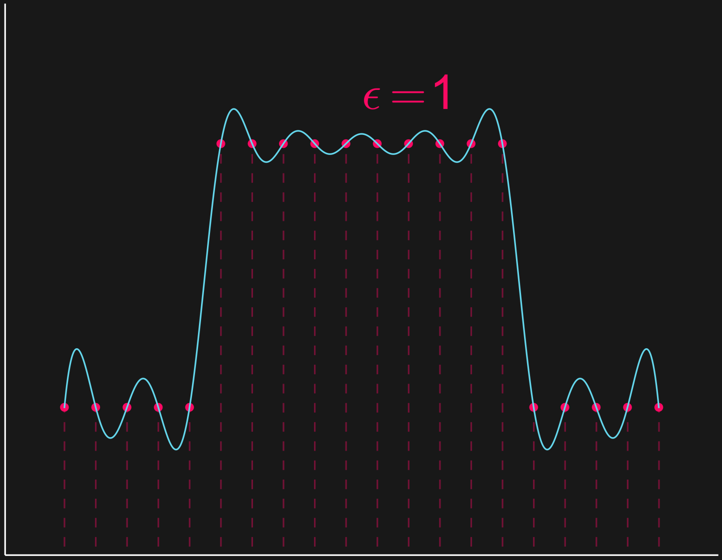
Well-Posed $\neq$ Well-Conditioned

We now know that our system is well-posed, so a unique solution exists.

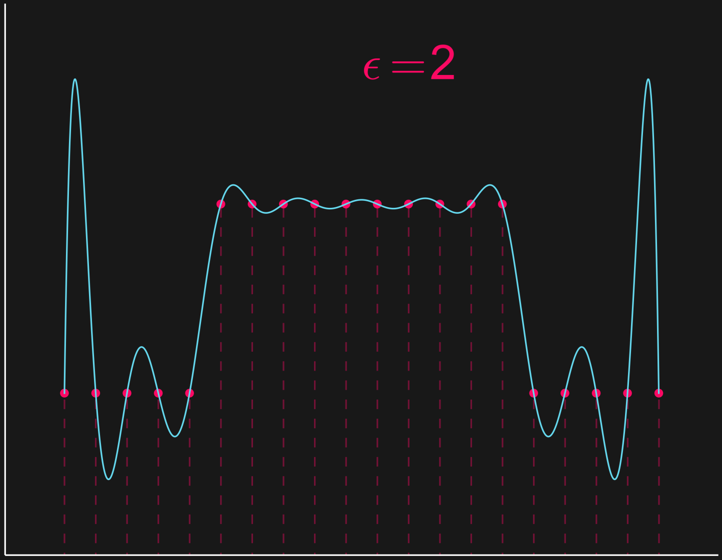
However, this solution isn't always accessible using numerical methods, making it ill-conditioned due to a loss of percision in computationally solving the linear system.

Radial Basis Interpolation has the propensity to be ill-conditioned, especially when choosing shape parameter, ϵ.

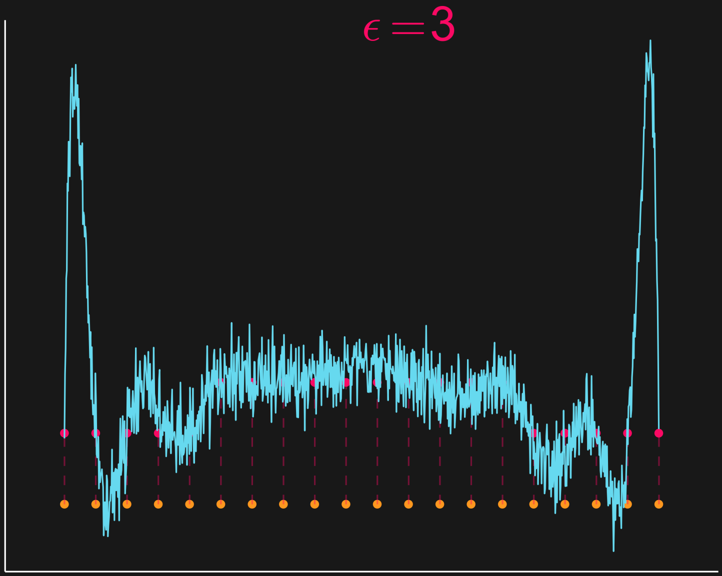
Pure Mathematicians can Leave Now



Pure Mathematicians can Leave Now

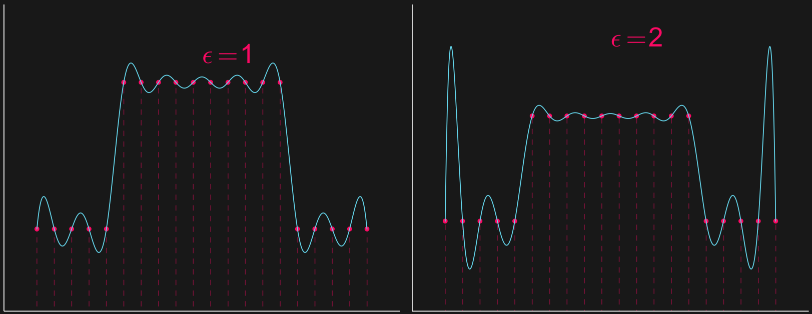


Pure Mathematicians can Laugh Now



Considerations When Using RBFs

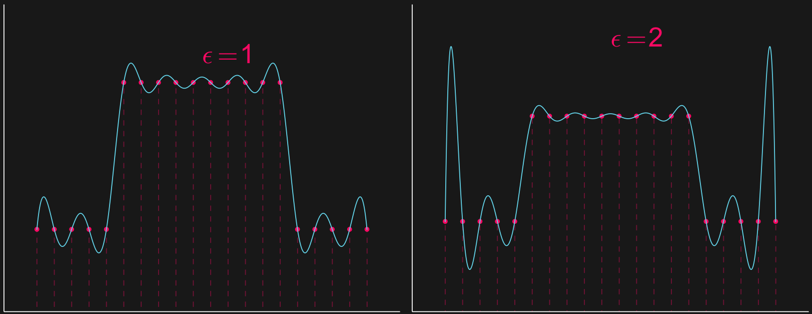
- The interpolation matrix, A, is dense $\Longrightarrow$ Computationally Expensive
- Choosing shape parameter, ϵ, often involves optimization
- RBF interpolation can easily be ill-conditioned
- Interpolation error blows up near bounderies



Considerations When Using RBFs

Interpolation error blows up near bounderies

Why? Because our basis functions are centered at our data and kernels will be asymetric near boundaries.



Implication? RBFs are great for interpolating data with no boundaries! e.g. global bathymetry data

Conclusion

Hopefully with this talk you

- Refreshed your memory on interpolation
- Recognized interpolation through linear combination of basis functions
- Appreciated that RBFs are simply data-dependent basis functions
- Understood how RBFs are proven to be Well-Posed
- Understood how RBFs can still be Ill-Conditioned
- Will feel comfortable approachng RBFs for your appropriate interpolation needs

References

Martin Buhmann.
Radial basis functions: theory and implementations.
Cambridge University Press, 5 edition, 2003.

Greg Fasshauer.
Mesh Free Methods (590), 2012.

Michael Mongillo.
Choosing Basis Functions and Shape Parameters for Radial Basis Function Methods.
Technical report, Illinois Institute of Technology, 2011.

Grady Wright.
Radial Basis Function Interpolation: Numerical and Analytical Developments.
PhD thesis, University of Colorado, 2003.