Introduction to Radial Basis Functions

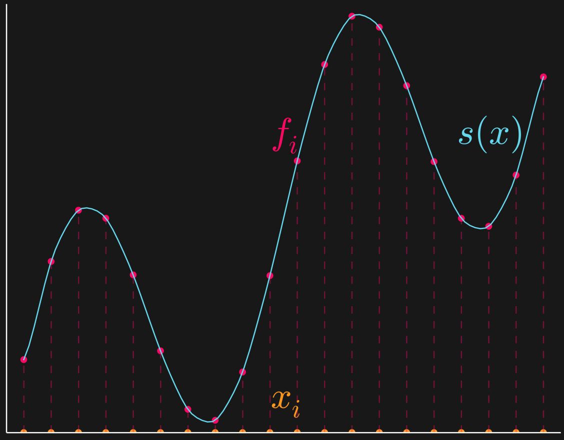
Jesse Bettencourt

McMaster University

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

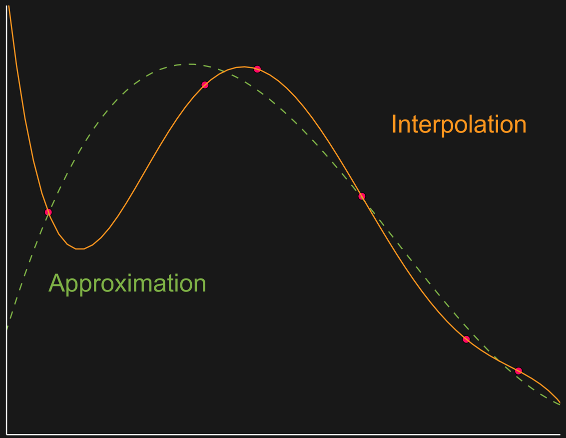
2D: Surface temperature of a lake based on measurements collected at sample surface locations

3D: Distribution of temperature within a lake

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

What makes a good fit?

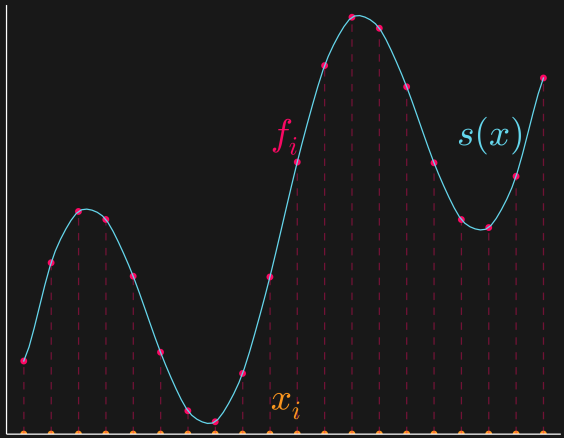
- Interpolation: $s(x)$ exactly matches our measurements at our data sites.
- Approximation: $s(x)$ closely matches our measurements at our data sites, e.g. with Least Squares



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

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

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

$$A_{ii} = \psi_i(x_i),\ i, i = 1 \dots N$$

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

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

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

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, $\psi_{i=1}^{N}$

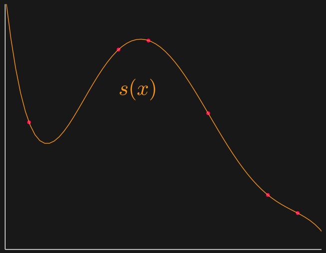
Easily Well-Posed in 1D

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

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.

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

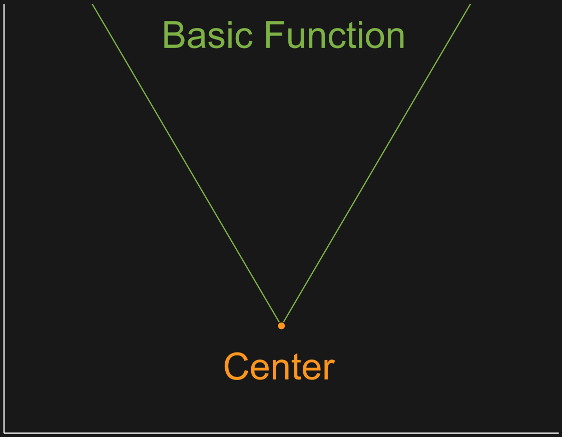
Implication: If we choose our basis functions independently of the data, we are not guaranteed a well-posed problem.

Solution?
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.

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

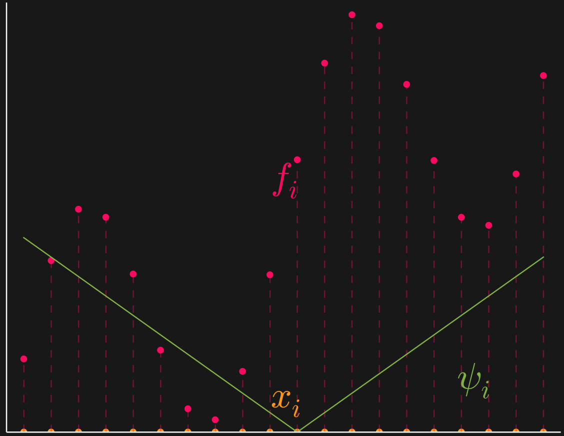
So each basis function, $\psi_i(x)$, is our basic function shifted so that the center or knot is positioned on a data site, x_i.

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.

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



Radial Basis Functions

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

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

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!

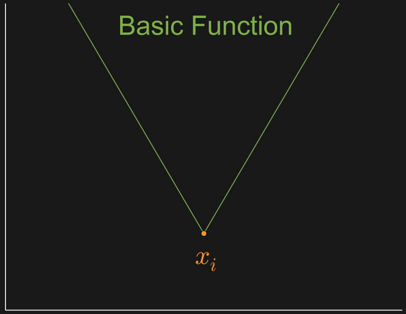
Our interpolant becomes $s(x) = \sum_{i=1}^{N} \lambda_i ||x - x_i||$

Building a Better Basic Function

Basic function

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

has a discontinuity in its first derivative at x_i.



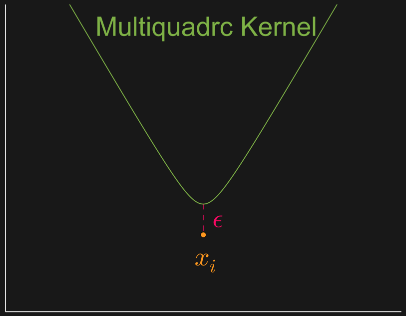
This causes the interpolant to have a discontinuous first derivative at each data site.
Obviously not ideal.

Building a Better Basic Function

In 1968, R.L. Hardy showed that we can remedy this problem by changing our basic function to one with continuous derivatives.

Hardy's Multiquadrc Kernel

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

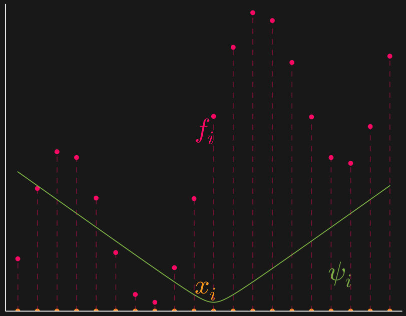


Note: The case where $\epsilon = 0$ is the previous basic function.

Radial Basis Kernels

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

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



Notice that the Hardy's Multiquadric function is still radially symmetric about its center, making it a Radial Basis Function (RBF). All RBFs are functions only of distance from center, and can be written generally as $\phi(||x - x_i||)$.

Radial Basis Kernels

The RBF Method

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

There are a few commonly used RBF Kernels: