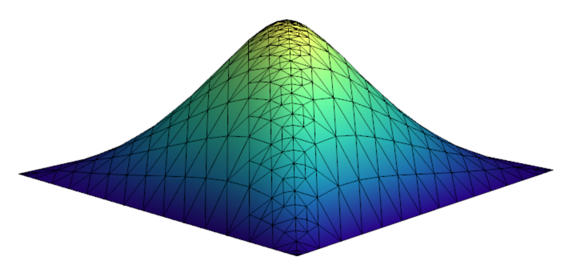
# A 2-Parameter Family of Kernels for Data Interpolation


Casey Bylund[1] Will Mayner[2]

[1]University of San Francisco
[2]Brown University

## Motivation

- In many application areas, one encounters a function $f$ which is difficult or expensive to evaluate, for example the output of a sophisticated computer simulation.
- One desires an approximation of $f$ that is easier to evaluate than $f$ itself, *i.e.*, an interpolant.
- A popular way of constructing these surrogate functions is to use **positive definite kernels**.
- The choice of which kernel to use has a significant effect on the accuracy of the resulting approximation.
- So, we'd like a systematic way of choosing the most suitable kernel for the particular application.

## Introduction

- We study a 2-parameter family of **compact Matérn kernels** arising as **Green's functions** associated with differential equations of the form

$$\left(-\left|\frac{\mathrm{d}^2}{\mathrm{d}x^2}\right| + \varepsilon^2\,\mathrm{Id}\right)^{\beta} K_{\beta,\varepsilon}(x,y) \;=\; \delta(x-y), \qquad (1)$$

subject to the boundary conditions

$$\begin{aligned} K_{\beta,\varepsilon}(0,y) &= K''_{\beta,\varepsilon}(0,y) = \cdots = K^{2(\beta-1)}_{\beta,\varepsilon}(0,y) = 0 \\ K_{\beta,\varepsilon}(L,y) &= K''_{\beta,\varepsilon}(L,y) = \cdots = K^{2(\beta-1)}_{\beta,\varepsilon}(L,y) = 0, \end{aligned} \qquad (2)$$

where $\delta$ is the Dirac delta function and $[0,L]\times[0,L]$ is the domain on which our kernel is defined.
- Current kernel methods optimize $\varepsilon$, the "shape" parameter.

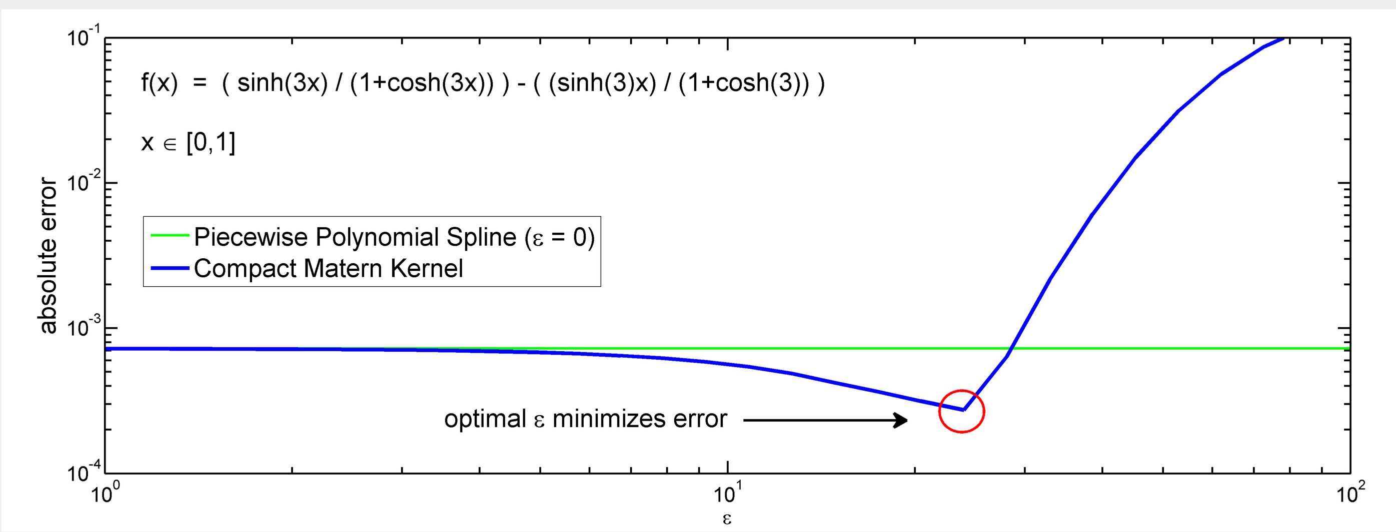



**Figure:** Relative RMS error as $\varepsilon$ varies, using kernels with $\beta = 8$. Note that there is a *nonzero* value of $\varepsilon$ that minimizes error.

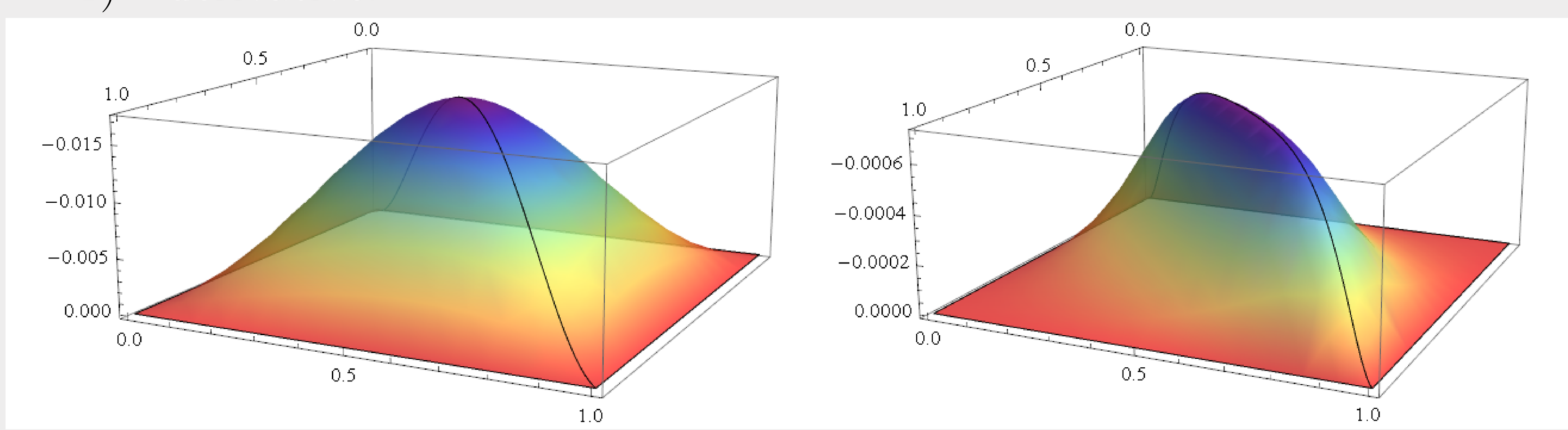
- By introducing the new parameter $\beta$, the number of iterations of the differential operator $\left(-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \varepsilon^2\,\mathrm{Id}\right)$, we can force the kernel $K_{\beta,\varepsilon}$ to be continuously differentiable up to the $(2\beta-2)^{\text{th}}$ derivative. In this sense $\beta$ is a **smoothness parameter**.
- Closed forms of these kernels are known when $\varepsilon = 0$.
- In that special case, we recover certain piecewise polynomial splines, which are well-understood and can be expressed as

$$K_{\beta,0}(x,y) = \frac{(2L)^{2\beta-1}}{(2\beta)!}\left[B_{2\beta}\left(\frac{|x-y|}{2L}\right) - B_{2\beta}\left(\frac{x+y}{2L}\right)\right],$$

with Bernoulli polynomials $B_{2\beta}$.

## Closed Forms when $\varepsilon > 0$ via Green's functions

- A Green's function is a type of function used to solve differential equations subject to boundary conditions.
- The kernels in this family are continuously differentiable up to the $(2\beta-2)^{\text{th}}$ derivative, and have a jump of 1 in the $(2\beta-1)^{\text{th}}$ derivative along the line $y = x$ due to the Dirac delta function on the RHS of (1).
- We split the domain $[0,L]\times[0,L]$ along the interface $x = y$ and find the Green's function on either side.
- We force the two sides to agree along the interface with "gluing" conditions.
- With the boundary conditions in equation (2) and these interface conditions, a Green's function is uniquely determined.
- The Green's function for the original boundary value problem is then given by the piecewise union of those of each half of the domain.
- This domain-splitting approach sidesteps the difficulty of the discontinuity in the $(2\beta-1)^{\text{th}}$ derivative.



(a) $\varepsilon = 1$ (b) $\varepsilon = 7$

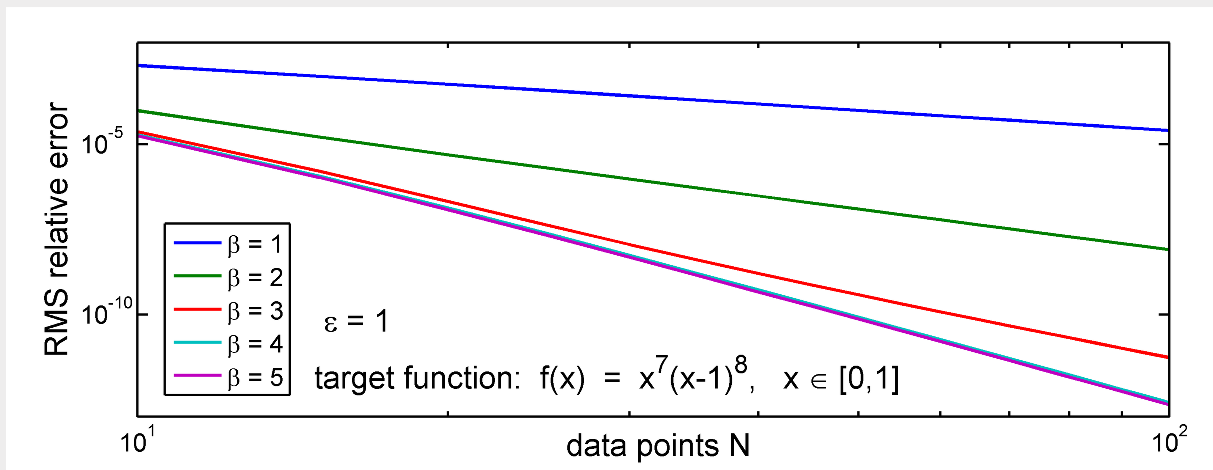
**Figure:** Surface plot of $K_{2,\varepsilon}$. $\beta = 2,\ \varepsilon > 0,\ L = 1$. Note the interface $x = y$ in black, and the effect of the shape parameter.

## Convergence Behavior

- We perform numerical experiments to investigate how fast the interpolant converges to the target function as $N$, the number of data points, increases.
- Here we assume that the error in the interpolant is proportional to $N^{-p}$ (we refer to the exponent $p$ as the *order of convergence*).
- If the function to be approximated satisfies the boundary conditions in (2) for all *even* derivatives up to order $2n$, then we observe that the rate of convergence seems to be $p = 2\beta$ and it increases with $\beta$ until $\beta > n+1$, after which the rate of convergence remains constant:

| $\beta$ | $p$ |
|---|---|
| 1 | 1.93 |
| 2 | 4.11 |
| 3 | 6.21 |
| 4 | 6.24 |
| 5 | 6.24 |

(a) Convergence orders




(b) Error as $N$ increases for different values of $\beta$

**Figure:** The target function $f(x) = x^7(x-1)^8$ satisfies the left boundary conditions for all even derivatives up to the $6^{\text{th}}$, and those on the right for all even derivatives up to the $8^{\text{th}}$. Note how the rate of convergence increases as $\beta$ increases until a boundary condition is violated. This occurs even if only one of the boundary conditions is not satisfied, and even when higher-order boundary conditions are satisfied.

## Existence of Optimal $\beta$

- The smoothness parameter is only useful if there is an optimal value between 1 and $\infty$. Numerical evidence below seems to indicate that this is the case, though there is still work to be done.
- When $\frac{\beta}{2}$ is greater than the target function's highest-order even derivative that satisfies the boundary conditions, most of the error is concentrated at the boundary.
- As $\beta$ increases, the error at the boundary spreads toward the interior, while the interior enjoys better and better accuracy (until the influence of the boundary overtakes this improvement):

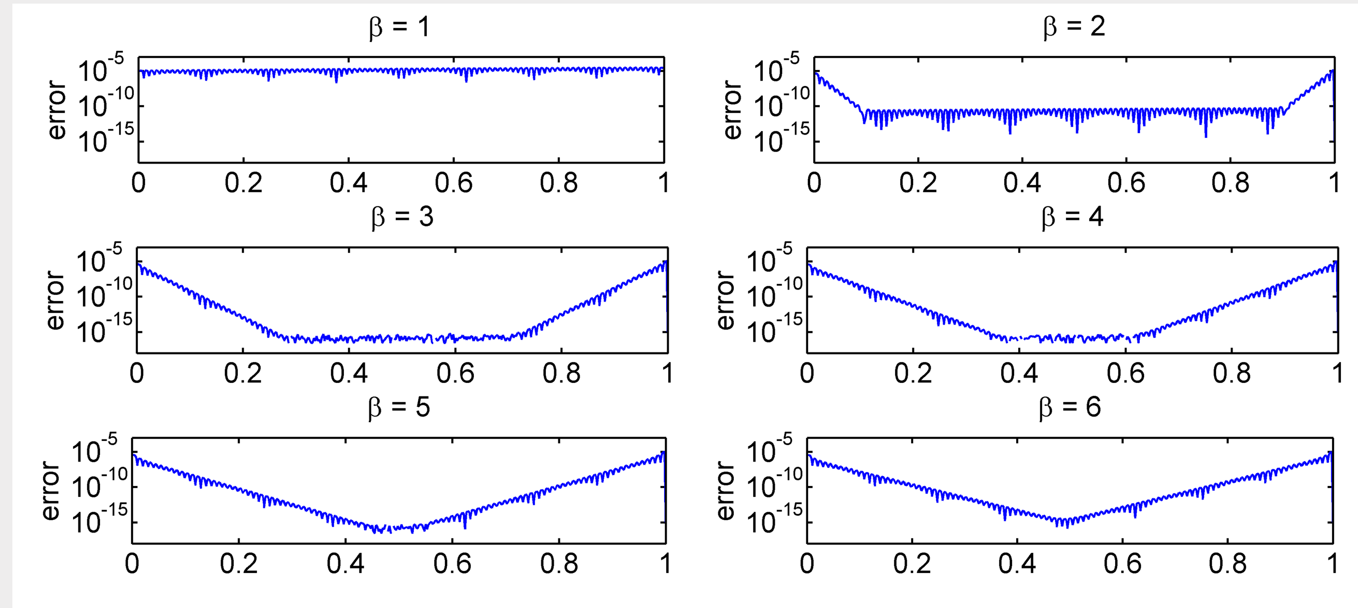



**Figure:** Error profiles over the interval $[0,1]$ for increasing $\beta$, with target function $f(x) = \mathrm{e}^x - (1-x) - \mathrm{e}x$. This function satisfies the boundary conditions only for $\beta = 1$. Note how error begins spreading towards the interior when $\beta = 2$, yet error continues to decrease in the interior. The interplay between these two phenomena indicates that there is an optimal choice of $\beta$, at least for a subregion of the domain.

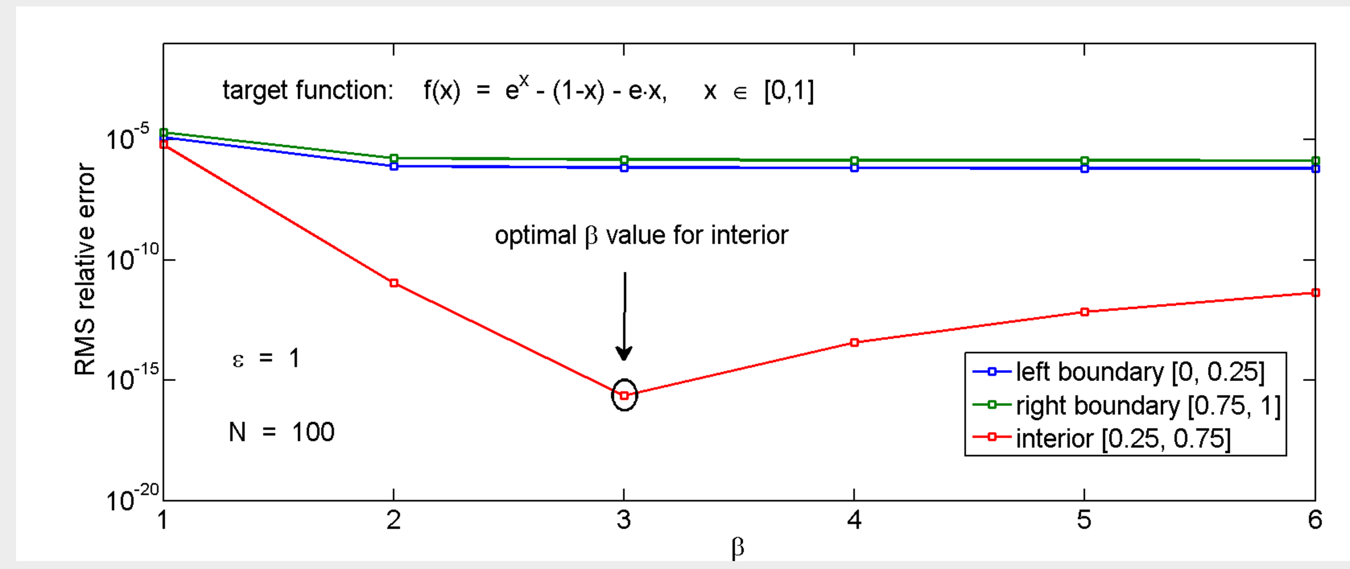



**Figure:** RMS error for interior and boundary regions. When we consider the interior of the interval separately from the boundaries, we observe that there is an optimal $\beta$.

## Conclusions

- The degree to which the target function and its derivatives satisfy the boundary conditions has a significant influence on the convergence behavior of the interpolant as the kernels become smoother.
- Accuracy can be gained in a region of the domain by optimizing $\beta$ for that region. This justifies the introduction of the smoothness parameter.

## Acknowledgments


- Special thanks to our mentor, Greg Fasshauer, and to Mike McCourt.
- This work was supported by NSF grant DMS–1115392, Kernel Methods for Numerical Computation.


## References


[1] Simon Hubbert and Stefan Müller. Thin plate spline interpolation on the unit interval. Numerical Algorithms, 45:167–177, 2007.