Bivariate Splines for De-Convolution

Ming-Jun Lai[1]

based on a joint work with Tianhe Zhou

mjlai@uga.edu

May 20, 2017

[1]This author is associated with Department of Mathematics, The University of Georgia, Athens, GA 30602, U.S.A. mjlai@uga.edu. His research is supported by the National Science Foundation under the grant #DMS 1521537.

Table of Contents

Introduction: the Research Problem

Suppose that we are given a Fredholm integral equation of first kind

$$F(\mathbf{x}) = \int_{\Omega} K(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}, \quad \mathbf{x} \in \Omega \tag{1}$$

over a bounded polygonal domain $\Omega \in \mathbf{R}^2$. We are interested in how to approximate $f(\mathbf{y})$, $\mathbf{y} \in \Omega$, assuming that $K(\mathbf{x}, \mathbf{y})$ is a given kernel function and $F(\mathbf{x})$. When $K(\mathbf{x}, \mathbf{y}) = K(\mathbf{x} - \mathbf{y})$, $F(\mathbf{x})$ is a convolution function. It is an inverse problem.

More precisely, we are going to use bivariate splines ([Lai and Schumaker, 2007[2]] and [Awanou, Lai and Wenston, 2006[3]]) to approximate $f(\mathbf{y})$.

[2] M. J. Lai and L. L. Schumaker, Spline Functions over Triangulations, Cambridge University Press, 2007.

[3] G. Awanou, M. J. Lai, and P. Wenston: The multivariate spline method for numerical solution of partial differential equations and scattered data interpolation, Wavelets and Splines (G. Chen and M. J. Lai, eds.). Nashboro Press, 24–74, (2006).

Introduction: Motivations

This research problem has several motivations. Let me explain some of them as follows.

- Numerical Solution of Integral Equations of Second Kind
- Image De-blurring
- Boundary Element Methods for Numerical Solution of Poisson Equations
- Learning Theory
- Machine Learning
- etc....

Motivation (1)

One often needs to numerically solve a linear integral equation of second kind: find $u \in L^2(\Omega)$ satisfying

$$u(\mathbf{x}) = \int_\Omega \int_\Omega K(\mathbf{x}, \mathbf{y}) u(\mathbf{y}) d\mathbf{y} + F(\mathbf{x}), \quad \mathbf{x} \in \Omega, \tag{2}$$

where $K(\mathbf{x}, \mathbf{y})$ and $F(\mathbf{x})$ are given (cf, [M. Schultz,73[4]] and [Chen, Micchelli, Xu, 2015[5]]). Letting $\delta(\mathbf{x}, \mathbf{y})$ be a Dirac function over Ω such that $f(\mathbf{x}) = \int_\Omega \delta(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}$, we can rewrite (2) in

$$F(\mathbf{x}) = \int_\Omega (\delta(\mathbf{x}, \mathbf{y}) - K(\mathbf{x}, \mathbf{y})) u(\mathbf{y}) d\mathbf{y}, \quad \mathbf{x} \in \Omega. \tag{3}$$

Thus, letting $\widetilde{K}(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x}, \mathbf{y}) - K(\mathbf{x}, \mathbf{y})$, (2) is our research problem.

[4] M. H. Schultz, Spline Analysis, Prentice-Hall, 1973.

[5] Z. Chen, C. A. Micchelli, and Y. Xu, Multiscale Methods for Fredholm Integral Equations, Cambridge University Press, 2015

Kernel Functions

Typically, $K(\mathbf{x}, \mathbf{y}) = K(\mathbf{x} - \mathbf{y})$ is a given kernel function. A weakly singular kernel function or a Mercer kernel are typical examples.

- $\delta(x-y) = \dfrac{1}{2\pi}\int_{-\infty}^{\infty} \exp(ip(x-y))dp$ in the univariate setting. It can be approximated by *nascent delta* function δ_ϵ such that $f * \delta_\epsilon \to f$ as $\epsilon \to 0_+$.
- In $\mathbb{R}^d$ with $d \geq 2$,

$$K(\mathbf{x}-\mathbf{y}) = \begin{cases} -\frac{1}{2\pi}\log|\mathbf{x}-\mathbf{y}|, & \text{if } d = 2 \\ -\frac{1}{(d-2)\omega_{d-1}}\frac{1}{|\mathbf{x}-\mathbf{y}|^{d-1}}, & \text{if } d \geq 3. \end{cases} \tag{4}$$

- $K(\mathbf{x}-\mathbf{y}) = \exp(-\alpha|\mathbf{x}-\mathbf{y}|^2)$ for some $\alpha > 0$.
- box spline kernels, radial basis kernel, etc..

Image De-Convolution

A blurred image is usually modeled as a convolution of a clean image with a blurring kernel with some noises:

$$F(\mathbf{x}) = \int_\Omega K(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y} + N(\mathbf{x}), \quad \mathbf{x} \in \Omega \tag{5}$$

where K is a blurring kernel, $f(\mathbf{y})$ is a clean image, $N(\mathbf{x})$ stands for noises. That is, given $F(\mathbf{x})$ noised blurred image, find a clean image $f(\mathbf{y})$. It is also called non-blind image deconvolution if $K(\mathbf{x}, \mathbf{y})$ is known. Otherwise, it is called blind image deconvolution (cf. e.g. [Chai and Shen, 2007[6]], [Daubechies, Teschke, Vese, 2007[7]], [Cai and Shen, 2010[8]]).

[6] A. Chai and Z. Shen, Deconvolution: A wavelet frame approach, Numer. Math., 106 (2007), pp. 529–587.

[7] I. Daubechies, G. Teschke, and L. Vese, Iteratively solving linear inverse problems under general convex constraints, Inverse Problems and Imaging, 1 (2007), pp. 29–46.

[8] J. Cai and Z. Shen, Framelet based deconvolution, J. Comput. Math. 28 (3) (2010) 289–308.

Geopotential Estimate

More precisely, suppose that the geopotential on an orbital surface S_o, e.g., 500 km above the Earth surface is given. One is interested in estimating the geopotential on the ground level S_e of the Earth. Since the geopotential $G(z)$ is a harmonic function satisfying the Laplace equation with boundary values at Earth surface with radius $R_e = 6378.136$km, one knows

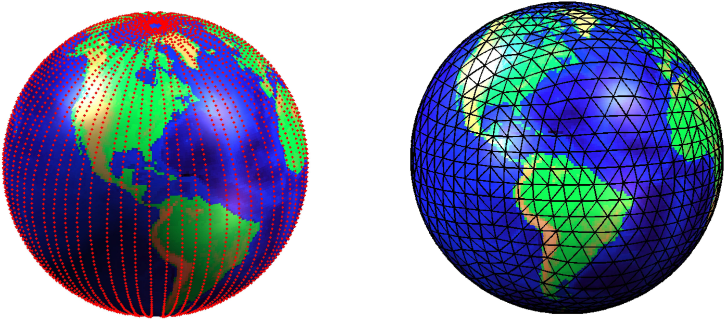
$$G(\mathbf{z}) = \int_{\mathbf{S}_e} K(\mathbf{z}, \mathbf{x}) G(\mathbf{x}) d\mathbf{x} \tag{6}$$

with a known Poisson kernel $K(\mathbf{z}, \mathbf{x})$ (cf. [Heiskanen and Moritz, 67[9]]), where $\mathbf{S}_e$ denotes a spherical surface around the Earth with radius R_e and $G(z)$ is given at $|\mathbf{z}| = R_e + 500$km (cf. e.g. [Lai, Shum, Baramidze and Wenston, 2009[10]])

[9] W. Heiskanen and H. Moritz, Physical Geodesy, Freeman, San Franscico, 1967.

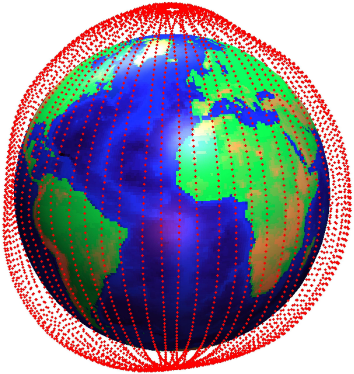
[10] Lai, M. -J., Shum, C. K., Baramidze, V. and Wenston, P., Triangulated Spherical Splines for Geopotential Reconstruction, Journal of Geodesy, vol. 83 (2009) pp. 695–708.

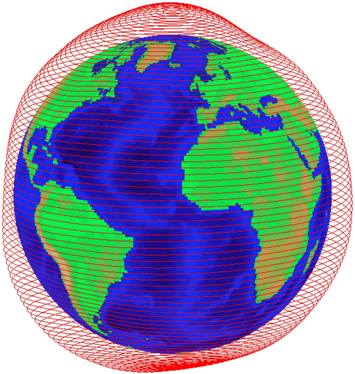
Geopotential Reconstruction [11]



[11] Lai, M. J., Shum, C. K., Baramidze, V. and Wenston, P., Triangulated Spherical Splines for Geopotential Reconstruction, Journal of Geodesy, vol. 83 (2009) pp. 695–708.

Spherical Spline Interpolation





We use spherical splines to reconstruct the geopotential values from the geopotential measurements from a German satellite which was lunched around 2002. Our spline surface interpolates the given data set.

Classic Learning Theory

Another example is the computation in the classic learning theory. To estimate the regression function f_ρ

$$f_\rho(\mathbf{x}) = \int_X y d\rho(y|\mathbf{x}), \quad \mathbf{x} \in X,$$

one would like to approximate the minimizer f_λ^* of the quadratic functional

$$f_\lambda := \arg \min_{f \in \mathcal{H}_K} \int_X (f - f_\rho)^2 d\rho_X + \lambda \|f\|_K^2, \tag{7}$$

where $\lambda > 0$ is a fixed penalty parameter, $\mathcal{H}_K$ is a Reproducing Kernel Hilbert Space (RKHS) and ρ is a conditional probability.
Here K is a continuous, symmetric and positive semidefinite kernel.
$\|f\|_K^2 = \int_X f(\mathbf{x}) \int_X K(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\rho_X(\mathbf{y}) d\rho_X(\mathbf{x})$.
See, e.g. [Cucker and Smale, 2001[12]].

[12] F. Cucker and S. Smale, On the mathematical foundations of learning. Bull. Amer. Math. Soc., 39(2001), 773–795.

Classic Learning Theory

It is known that f_λ exists, is unique, and satisfies the following

$$\lambda f_\lambda(\mathbf{x}) + \int_X K(\mathbf{x}, \mathbf{y}) f_\lambda(\mathbf{y}) d\rho_X(\mathbf{y}) = \int_X K(\mathbf{x}, \mathbf{y}) f_\rho(\mathbf{y}) d\rho_X(\mathbf{y}) \tag{8}$$

We refer to [Smale and Zhou, 2003[13]] and [F. Cucker, D.-X. Zhou, 2007[14]] for details.
Similar to the integral equation of second kind, the above (8) can be formulated as a de-convolution.

[13] S. Smale and D.-X. Zhou, Estimating the approximation error in learning theory. Anal. Appl. 1(2003), 17–41.

[14] F. Cucker and D. -X. Zhou, Learning theory : an approximation theory viewpoint, Cambridge University Press, 2007

Existing Numerical Methods

There are many approaches to tackle the numerical approximation of the de-convolution.

- Framelet Approach (cf. [Cai and Shen, 2010]);
- Galerkin Methods (cf. [Chen, Micchelli, Xu, 2015]);
- Collocation Methods (cf. [Chen, Micchelli, Xu, 2002[15]]);
- Petrov-Galerkin Methods (cf. [Chen and Xu, 1998[16]]),
- Degenerate Kernel Methods (cf. [Chen, Micchelli, Xu, 2015])
- Multiscale Methods (cf. [Chen, Micchelli, Xu, 2015])
- Learning Schemes, (cf. D. X. Zhou and his collaborators, Smale and Y. Yao, 2006, ...)

---

[15] Z. Chen, C. A. Micchelli, Y. Xu, Fast collocation methods for second kind integral equations, SIAM J. Numerical Analysis, 40(2002), 344–375.

[16] Z. Chen and Y. Xu, The Petrov-Galerkin and integral Petrov-Galerkin methods for second kind integral equations, SIAM J. Numerical Analysis, 35(1998), 406–434.

Difficulties and Challenges

Major difficulties in numerical computation for de-convolution are

- 1) the matrix associated with linear systems from the most methods is **dense** and the **size** of the matrix increases quickly in the multi-dimensional setting;
- 2) A given $F(\mathbf{x})$ is noisy and has some errors, how can one recovery $f(\mathbf{y})$ accurately?
- 3) When $K(\mathbf{x}, \mathbf{y})$ is very smooth, even $f(\mathbf{y})$ is not smooth at all, the given $F(\mathbf{x})$ is very reasonable. In this case, can we recovery $f(\mathbf{y})$?
- 4) Another problem is how to deal with this problem over irregular domains, any polygonal domain in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Can we use spline functions to help?

What are Multivariate Splines?

Let $\mathbf{P}_d$ be the space of all polynomials of degree $d \geq 1$. Let Δ be a triangulation of a domain $\Omega \subset \mathbb{R}^2$. For integers $d \geq 1$, $-1 \leq r \leq d$ define by

$$S_d^r(\Delta) = \{s \in C^r(\Omega), s|_t \in \mathbf{P}_d, t \in \Delta\}$$

the spline space of smoothness r and degree d over Δ.

In general, let $\mathbf{r} = (r_1, \cdots, r_n)$ with $r_i \geq 0$ be a vector of integers. Define

$$S_d^{\mathbf{r}}(\Delta) = \{s \in C^{-1}(\Omega), s|_{e_i} \in C^{r_i}, e_i \in E\},$$

where E is the collection of interior edges of $\triangle$. Each spline in $S_d^{\mathbf{r}}(\triangle)$ has variable smoothness.

This can handle the situation of hanging nodes in a triangulation!

Definition of Spline Functions

Let $T = \langle (x_1, y_1), (x_2, y_2), (x_3, y_3) \rangle$. For any point (x, y), let b_1, b_2, b_3 be the solution of

$$\begin{aligned} x &= b_1 x_1 + b_2 x_2 + b_3 x_3 \\ y &= b_1 y_1 + b_2 y_2 + b_3 y_3 \\ 1 &= b_1 + b_2 + b_3. \end{aligned}$$

Fix a degree $d > 0$. For $i + j + k = d$, let

$$B_{ijk}(x, y) = \frac{d!}{i!j!k!} b_1^i b_2^j b_3^k$$

which is called Bernstein-Bézier polynomials.
For each $T \in \Delta$, let

$$S|_T = \sum_{i+j+k=d} c_{ijk}^T B_{ijk}(x, y).$$

We use $\mathbf{s} = (c_{ijk}^T, i + j + k = d, T \in \Delta)$ be the coefficient vector to denote a spline function in $S_d^{-1}(\Delta)$. This setup can include the discontinuous

Evaluation and Derivatives

We use the de Casteljau algorithm to evaluate a Bernstein-Bézier polynomial at any point inside the triangle. It is a simple and stable computation. See [Lai and Schumaker, 2007[17]]

It is also used for computation of derivatives
Let $T = \langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$ and $S|_T = \sum_{i+j+k=d} c_{ijk} B_{ijk}(x,y)$. Then directional derivative

$$D_{\mathbf{v}_2-\mathbf{v}_1} S|_T = d \sum_{i+j+k=d-1} (c_{i,j+1,k} - c_{i+1,j,k}) B_{ijk}(x,y).$$

Similar for $D_{\mathbf{v}_3-\mathbf{v}_1} S|_T$.
D_x and D_y are linearly combinations of these two directional derivatives.

[17] M. -J. Lai and Larry L. Schumaker, Spline Functions over Triangulations, Cambridge University Press, 2007.

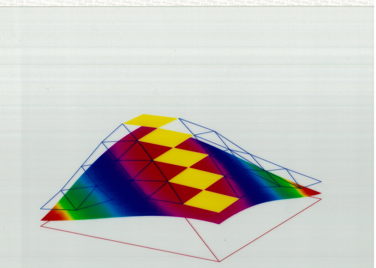
Smoothness Condition between Triangles

Let T_1 and T_2 be two triangles in Δ which share a common edge e. Then $S \in C^r(T_1 \cup T_2)$ if and only if the coefficients of $c_{ijk}^{T_1}$ and $c_{ijk}^{T_2}$ satisfy the following linear conditions. E.g.,

$$S \in C^0(T_1 \cup T_2) \text{ iff } c_{0,j,k}^{T_1} = c_{j,k,0}^{T_2}, j + k = d$$

$$S \in C^1(T_1 \cup T_2) \text{ iff } c_{1,j,k}^{T_1} = b_1 c_{j+1,k,0}^{T_2} + b_2 c_{j,k+1,0}^{T_2} + b_3 c_{j,k,1}^{T_2}$$

for $i + k = d - 1$ and etc. (cf. [Farin'86] and [de Boor'87]). We code them by H**c**=0.



Integration

Let s be a spline in $S_d^r(\triangle)$ with $s|_T = \sum_{i+j+k=d} c_{ijk}^T B_ijk(x,y)$, $T \in \triangle$. Then

$$\int_\Omega s(x,y)dxdy = \sum_{T\in\triangle} \frac{A_T}{\binom{d+2}{2}} \sum_{i+j+k=d} c_{ijk}^T.$$

If $p = \sum_{i+j+k=d} a_{ijk} B_{ijk}(x,y)$ and $q = \sum_{i+j+k=d} b_{ijk} B_{ijk}(x,y)$ over a triangle T, then

$$\int_T p(x,y)q(x,y)dxdy = \mathbf{a}^\top M_d \mathbf{b},$$

where $\mathbf{a} = (a_{ijk}, i+j+k=d)^\top$, $\mathbf{b} = (b_{ijk}, i+j+k=d)^\top$, M_d is a symmetric matrix with known entries (cf. [Chui and Lai, 1990]). Similarly, we have

$$\int_T p(x,y)q(x,y)r(x,y)dxdy = \mathbf{a}^\top A_d \mathbf{b} \odot \mathbf{c}.$$

which can be used for weighted inner products of polynomials.

Spline Approximation Order

Theorem ([Lai and Schumaker'98)

[a]] *Suppose that* $\triangle$ *is a* β*-quasi-uniform triangulation of domain* $\Omega \in \mathbf{R}^2$ *and suppose that* $d \geq 3r+2$. *Fix* $0 \leq m \leq d$. *Then for any* f *in a Sobolev space* $W_p^{m+1}(\Omega)$, *there exists a quasi-interpolatory spline* $Q_f \in S_d^r(\triangle)$ *such that*

$$\|f - Q_f\|_{k,p,\Omega} \leq C|\triangle|^{m+1-k}|f|_{d+1,p,\Omega}, \forall 0 \leq k \leq m+1 \tag{9}$$

for a constant $C > 0$ *independent of* f, *but dependent on* β *and* d.

[a] Lai, M. J. and Schumaker, L. L., Approximation Power of Bivariate Splines, Advances in Computational Mathematics, vol. 9 (1998) pp. 251–279.

When $d \geq 3r+2$, $S_d^r(\triangle)$ has a super-spline subspace which consists of a locally supported basis and achieves the full approximation order (9).

The Galerkin Method

Let $\mathcal{S}$ be a spline space which has a locally supported stable basis $\{B_\xi, \xi \in \mathcal{M}\}$. Let $S_f = \sum_{\xi \in \mathcal{M}} c_\xi B_\xi$ be the approximation of f satisfying

$$\sum_{\xi \in \mathcal{M}} c_\xi \int_\Omega \int_\Omega K(\mathbf{x}, \mathbf{y}) B_\xi(\mathbf{y}) B_\eta(\mathbf{x}) d\mathbf{y} d\mathbf{x} = \int_\Omega F(\mathbf{x}) B_\eta(\mathbf{x}) d\mathbf{x} \tag{10}$$

for all $B_\eta \in \mathcal{S}$. This is called the Galerkin method.
The linear system $A\mathbf{c} = \mathbf{b}$ with

$$A = [A_{\xi,\eta}], A_{\xi,\eta} = \int_\Omega \int_\Omega K(\mathbf{x}, \mathbf{y}) B_\xi(\mathbf{y}) B_\eta(\mathbf{x}) d\mathbf{y} d\mathbf{x}. \tag{11}$$

We say $K(\mathbf{x}, \mathbf{y})$ is positive definite with respect to a spline space $\mathcal{S}$ if $\int_\Omega B_\eta(\mathbf{x}) \int_\Omega K(\mathbf{x}, \mathbf{y}) B_\xi(\mathbf{y}) d\mathbf{y} d\mathbf{x} > 0$ for all $B_\eta, B_\xi \in \mathcal{S}$.

The Galerkin Method (II)

Theorem

Suppose that $K(\mathbf{x}, \mathbf{y})$ is positive definite with respect to a spline space $\mathcal{S}$. There exists a unique solution S_f satisfying the weak equations (10).

Theorem

Suppose that $K(\mathbf{x}, \mathbf{y})$ is positive definite with respect to the standard $L^2(\Omega)$ and bounded from the above. Suppose that $f \in H^{\ell+1}(\Omega)$ for $\ell \geq 0$. Then

$$\|f - S_f\|_{2,\Omega} \leq C|\triangle|^{\ell+1}|f|_{\ell+1,\Omega} \tag{12}$$

for a positive constant C independent of f, where $0 \leq \ell \leq d$.

Similarly, we can extend these results to the setting when $F(\mathbf{x}) = \int_\Omega K(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}$ is a compact operator mapping from $L^2(\Omega)$ to itsself and is a bi-injection.

Computational Consideration

Definition

A kernel function K decays off diagonally if

$$|K(\mathbf{x},\mathbf{y})| \le \frac{C}{d(\mathbf{x},\mathbf{y})^\sigma + 1},$$

for $\sigma > 0$, where $d(\mathbf{x},\mathbf{y})$ is the distance between the $\mathbf{x}$ and $\mathbf{y}$.

For example, $K(\mathbf{x},\mathbf{y}) = \exp(-\sigma\|\mathbf{x}-\mathbf{y}\|^2)$.
When the support Ω_ξ of B_ξ and Ω_η of B_η are far away, our spline method gives

$$|A_{\xi,\eta}| = |\int_\Omega \int_\Omega K(\mathbf{x},\mathbf{y}) B_\xi(\mathbf{y}) B_\eta(\mathbf{x}) d\mathbf{y} d\mathbf{x}| \le \frac{C}{d(\Omega_\eta,\Omega_\xi)^\sigma + 1},$$

where Ω_ξ and Ω_η are the support of B_ξ and B_η, respectively.

Some New Approximation Schemes

We now propose a few new scheme to approximate f from the given information F.

- A Least Squares Method;
- Discrete Least Squares Method;
- Penalized Least Squares Method;
- Some Nonlinear Schemes;

A Least Squares Method

For a given convolution function

$$F(\mathbf{x}) = \int_\Omega K(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y},$$

we approximate f by $S_f \in S^r_d(\triangle)$ which solves the following minimization problem:

$$S_f = \arg \min_{s \in S^r_d(\triangle)} \int_\Omega \left(F(\mathbf{x}) - \int_\Omega K(\mathbf{x}, \mathbf{y}) s(\mathbf{y}) d\mathbf{y} \right)^2 d\mathbf{x}. \qquad (13)$$

This is a standard least squares approach for de-convolution.

- Let us see that the minimizer S_f exists and is unique.
- We also need to determine how well S_f approximates f in terms of size $|\triangle|$ of triangulation of Ω.
- How do we compute the solution and what the computational cost is.

Coerciveness and Uniqueness

Definition

We say $K(\mathbf{x}, \mathbf{y})$ is coercive with respect to spline space $S^r_d(\triangle)$ if

$$\int_\Omega K(\mathbf{x}, \mathbf{y}) s(\mathbf{y}) d\mathbf{y} = 0, \;\; a.e. \;\; on \;\; \Omega,$$

then $s(\mathbf{y}) = 0, \quad$ on $\;\Omega$.

Then, we have the following

Theorem (Existence and Uniqueness)

Suppose that the kernel function $K(\mathbf{x}, \mathbf{y})$ is coercive with respect to a spline space $S^r_d(\triangle)$. Then the minimization problem (13) has a unique solution in $S^r_d(\triangle)$.

Approximation of Least Squares Method

Let $S^r_d(\triangle) = \{\phi_1, \cdots, \phi_m\}$ and let $\{\phi_j, j = m+1, m+2, \cdots, \}$ be a basis of the orthogonal complement space of $S^r_d(\triangle)$ in a Hilbert space, e.g, $H = L_2(\Omega)$ or Sobolev space $H = W^r_2(\Omega)$.
Then we can write

$$f = \sum_{j=1}^{\infty} c_j \phi_j.$$

Note that f is the solution of the following the minimization:

$$f = \arg \min_{c_1, c_2, \cdots} \int_\Omega \left(F(\mathbf{x}) - \int_\Omega K(\mathbf{x}, \mathbf{y}) \sum_{j=1}^{\infty} c_j \phi_j(\mathbf{y}) d\mathbf{y} \right)^2 d\mathbf{x}.$$

Approximation of Least Squares Method

It follows that

$$\int_\Omega \int_\Omega K(\mathbf{x},\mathbf{y})f(\mathbf{y})d\mathbf{y} \int_\Omega K(\mathbf{x},\mathbf{y})\phi_j(y)d\mathbf{y}d\mathbf{x} = \int_\Omega F(\mathbf{x}) \int_\Omega K(\mathbf{x},\mathbf{y})\phi_j(\mathbf{y})d\mathbf{y}d\mathbf{x}$$

for all $j = 1, 2, \cdots, \infty$ while the spline minimization in (13) gives

$$\int_\Omega \int_\Omega K(\mathbf{x},\mathbf{y})S_f(\mathbf{y})d\mathbf{y} \int_\Omega K(\mathbf{x},\mathbf{y})\phi_j(\mathbf{y})d\mathbf{y}d\mathbf{x} = \int_\Omega F(\mathbf{x}) \int_\Omega K(\mathbf{x},\mathbf{y})\phi_j(\mathbf{y})d\mathbf{y}d\mathbf{x}$$

for all $j = 1, 2, \cdots, m$. It thus follows that

$$\int_\Omega \int_\Omega K(\mathbf{x},\mathbf{y})\left(f(\mathbf{y}) - S_f(\mathbf{y})\right)d\mathbf{y} \int_\Omega K(\mathbf{x},\mathbf{y})\phi_j(\mathbf{y})d\mathbf{y}d\mathbf{x} = 0 \qquad (14)$$

for all $j = 1, 2, \cdots, m$.

Approximation of Least Squares Method

Let Q_f be the quasi-interpolatory spline in $S^r_d(\triangle)$ which achieves the optimal order of approximation of f from $S^r_d(\triangle)$. Then (14) implies that

$$\begin{aligned}
& \int_\Omega \left(\int_\Omega K(\mathbf{x},\mathbf{y})(f(\mathbf{y}) - S_f(\mathbf{y}))dy \right)^2 dx \\
= & \int_\Omega \int_\Omega K(\mathbf{x},\mathbf{y})\,(f(\mathbf{y}) - S_f(\mathbf{y}))\,dy \int_\Omega K(\mathbf{x},\mathbf{y})\,(f(\mathbf{y}) - Q_f(\mathbf{y}))\,dydx \\
\leq & \left(\int_\Omega \left(\int_\Omega K(\mathbf{x},\mathbf{y})(f(\mathbf{y}) - S_f(\mathbf{y}))dy \right)^2 dx \right)^{\frac{1}{2}} \times \\
& \left(\int_\Omega \left(\int_\Omega K(\mathbf{x},\mathbf{y})(f(\mathbf{y}) - Q_f(\mathbf{y}))dy \right)^2 dx \right)^{\frac{1}{2}}.
\end{aligned}$$

That is, we have

$$\int_\Omega \left(\int_\Omega K(\mathbf{x},\mathbf{y})(f(\mathbf{y}) - S_f(\mathbf{y}))dy \right)^2 dx \leq \int_\Omega \left(\int_\Omega K(\mathbf{x},\mathbf{y})(f(\mathbf{y}) - Q_f(\mathbf{y}))dy \right)^2$$

Approximation of De-Convolution

When $K(\mathbf{x}, \mathbf{y})$ is coercive with respect to $L^2(\Omega)$, we can show there exists a positive constant C_K such that

$$\int_\Omega (\int_\Omega K(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y})^2 \geq C_K \int_\Omega (f(\mathbf{x}))^2 d\mathbf{x}, \quad \forall f \in L^2(\Omega).$$

By using the quasi-interpolant Q_f of f (cf. Theorem 1 above), we obtain the following

Theorem (Approximation Properties)

Suppose that the kernel function $K(\mathbf{x}, \mathbf{y})$ is bounded above and coercive. Suppose that $f \in C^{\ell+1}(\Omega)$ for $0 \leq \ell \leq d$. Then the solution S_f from the minimization problem (13) approximates f in the following sense:

$$\|f(\mathbf{x}) - S_f(\mathbf{x})\|_{L_2(\Omega)} \leq C|\triangle|^{\ell+1}|f|_{\ell+1,\Omega}$$

for a constant C dependent on d, $\frac{M_K}{C_K}$ and the boundary of Ω, where $|\triangle|$ is the maximal length of the edges of $\triangle$ and $|f|_{\ell,\Omega}$ denotes the L_2 norm of the ℓ^{th} derivatives of f over Ω.

Coercivity

What kind of kernel functions K are coercive with respect to $L^2(\Omega)$?
It is clear when $K(\mathbf{x},\mathbf{y}) = K_1(\mathbf{x})K_2(\mathbf{y})$ is a separate kernel, $K(\mathbf{x},\mathbf{y})$ will not be coercive.

Theorem

Suppose that $K(\mathbf{x},\mathbf{y}) = K(\mathbf{x}-\mathbf{y})$ and $\widehat{F}(\omega) \neq 0$. Then K is coercive.

The following kernels are coercive:

- $K(\mathbf{x},\mathbf{y}) = \exp(-\sigma\|\mathbf{x}-\mathbf{y}|^2)$, for some $\sigma > 0$;
- $K(\mathbf{x},\mathbf{y}) = \exp(-\sigma\|\mathbf{x}-\mathbf{y}|)$, where $\sigma > 0$;
- $K(\mathbf{x},\mathbf{y}) = \sqrt{\sigma + \|\mathbf{x}-\mathbf{y}|^2}$, where $\sigma > 0$;
- $K(\mathbf{x},\mathbf{y}) = 1/\sqrt{\sigma + \|\mathbf{x}-\mathbf{y}|^2}$, where $\sigma > 0$;
- $K(\mathbf{x},\mathbf{y}) = 1/(\sigma + \|\mathbf{x}-\mathbf{y}|^2)$, where $\sigma > 0$;

A Discrete Least Squares Method

Next we consider the discrete least squares approximation of f. Let $\mathbf{x}_i, i = 1, \cdots, n$ be some designed points in Ω which are evenly distributed with respect to $S_d^r(\triangle)$ in the following sense.

Definition

We say that $\mathbf{x}_i \in \Omega, i = 1, \cdots, n$ are evenly distributed over Ω with respect to $S_d^r(\triangle)$ if $\int_\Omega K(x_\ell, y) f(y) dy = 0, \forall \ell = 1, \cdots, n$ for a spline function $f(y) \in S_d^r(\triangle)$, then $f(y) \equiv 0$.

The discrete least squares approximation $\widehat{S_{f,n}} \in S_d^r(\triangle)$ is the solution of

$$\widehat{S_{f,n}} = \arg \min_{s \in S_d^r(\triangle)} \frac{1}{n} \sum_{i=1}^{n} \left(F(\mathbf{x}_i) - \int_\Omega K(\mathbf{x}_i, \mathbf{y}) s(\mathbf{y}) d\mathbf{y} \right)^2 . \qquad (15)$$

A Useful Lemma

To study the minimizer $\widehat{S_{f,n}}$ of (15), we may relate it to the least squares solution S_f. The following well-known lemma is needed.

Lemma

Let A be an invertible matrix and $\tilde{A}$ be a perturbation of A satisfying $\|A^{-1}\|\|A-\tilde{A}\| < 1$. Suppose that x and $\tilde{x}$ are the exact solutions of $Ax = b$ and $\tilde{A}\tilde{x} = \tilde{b}$, respectively. Then

$$\frac{\|x-\tilde{x}\|}{\|x\|} \le \frac{\kappa(A)}{1-\kappa(A)\frac{\|A-\tilde{A}\|}{\|A\|}} \left[\frac{\|A-\tilde{A}\|}{\|A\|} + \frac{\|b-\tilde{b}\|}{\|b\|}\right].$$

Here, $\kappa(A)$ denotes the condition number of matrix A.

Another Useful Lemma

In addition, we need the following

Lemma

Suppose that $\triangle$ is a β-quasi-uniform triangulation. Suppose that $d \geq 3r+2$. Then there exist two positive constants C_1 and C_2 independent of $\triangle$ such that for any spline function $S \in S^r_d(\triangle)$ with coefficient vector $\mathbf{s} = (s_1, \cdots, s_m)^T$ with $S = \sum_{i=1}^m s_i \phi_i$,

$$C_1 |\triangle|^2 \|\mathbf{s}\|^2 \leq \|S\|^2 \leq C_2 |\triangle|^2 \|\mathbf{s}\|^2.$$

A proof of this lemma can be found in [Lai and Schumaker, 2007[18]].

[18] M. J. Lai and L. L. Schumaker, Spline Functions over Triangulations, Cambridge University Press, 2007.

Approximation of Discrete Least Squares Method

Theorem

Suppose that the kernel function $K(\mathbf{x}, \mathbf{y})$ is nontrivial, bounded and coercive on Ω. Suppose that the designed points $x_\ell, \ell = 1, \cdots, n$ are evenly distributed with respect to $S^r_d(\triangle)$. Suppose that $f \in C^\ell(\Omega)$ for $0 \le \ell \le d+1$. Then the solution $\widehat{S_{f,n}}$ from the minimization problem (15) approximates g in the following sense:

$$\|f(x) - \widehat{S_{f,n}}(x)\|_{L_2(\Omega)} \le C_3|\triangle|^\ell |f|_{\ell,\Omega} + \frac{C_4\|f\|_{L_2(\Omega)}}{n}$$

for a constant C_3 dependent on d, $\frac{M_2}{M_1}$ and the smallest angle $\theta_\triangle$, and for a constant C_4 dependent on d, C_1, C_2 and $\kappa(A)$.

Penalized Discrete Least Method

In general, it is not easy to figure out if the designed points are evenly distributed over Ω with respect to $S_d^r(\triangle)$ or not. We thus propose another approach to approximate the function f. Mainly we seek a solution $\widetilde{S_{f,n,\lambda}} \in S_d^r(\triangle)$ which solves the following minimization problem:

$$\widetilde{S_{f,n,\lambda}} = \arg \min_{s \in S_d^r(\triangle)} \frac{1}{n} \sum_{i=1}^{n} (f(x_i) - \int_\Omega K(x_i, y)s(y)dy)^2 + \lambda E_r(s), \quad (16)$$

where $\lambda > 0$ is a parameter and $E_q(s)$ denotes the energy functional of s in the following sense:

$$E_q(s) = \int_\Omega \sum_{k=0}^{q} \sum_{i+j=k} (D_1^i D_2^j s)^2 dx,$$

where q is an integer with $0 \le q \le r$ and D_1 and D_2 stand for the partial derivatives with respect to the first and second variables.

Approximation of Penalized Least Squares Method

Theorem

Suppose that the kernel function $K(\mathbf{x}, \mathbf{y})$ is nontrivial, bounded and coercive on unit square $\Omega = [a, b] \times [c, d]$. Suppose that the designed points $x_\ell, \ell = 1, \cdots, n$ are evenly distributed with respect to the polynomial space $\mathcal{P}_{q-1}$. Suppose that $f \in C^\ell(\Omega)$ for $0 \leq \ell \leq d + 1$. Then the solution $\widetilde{S_{f,n,\lambda}}$ from the minimization problem (16) approximates g in the following sense:

$$\|f(x) - \widetilde{S_{f,n}}(x)\|_{L_2(\Omega)} \leq C_3|\triangle|^\ell |f|_{\ell,\Omega} + C_4(\frac{1}{n} + \lambda)\|f\|_{L_2(\Omega)}.$$

Here C_3 is a constant dependent on d, $\frac{M_2}{M_1}$ and the smallest angle $\theta_\triangle$, C_4 is a constant dependent on d, C_1, C_2, dimension of spline space $S^r_d(\triangle)$, and the condition number of $\kappa(A)$ of A.

Nonlinear Least Squares Methods

When the kernel function k is ill-posed, we can regularize the L^2 minimization by considering:
Find the minimizer $u_{f,\lambda} \in L^2(\Omega)$ of the following

$$\min_{s \in L^2(\Omega)} \frac{1}{2} \int_\Omega \left(F(\mathbf{x}) - \int_\Omega K(\mathbf{x}, \mathbf{y}) s(\mathbf{y}) d\mathbf{y} \right)^2 d\mathbf{x} + \frac{\lambda}{2} \int_\Omega s(\mathbf{y})^2 d\mathbf{y}, \qquad (17)$$

where $\lambda > 0$ is a parameter and the last term is a standard regularizer.
Certainly, in general, we should study the minimization problem:

$$\min_{s \in H^1(\Omega)} \frac{1}{2} \int_\Omega \left(F(\mathbf{x}) - \int_\Omega K(\mathbf{x}, \mathbf{y}) s(\mathbf{y}) d\mathbf{y} \right)^2 d\mathbf{x} + \frac{\lambda}{2} \int_\Omega |\nabla s(\mathbf{y})| d\mathbf{y}, \qquad (18)$$

where ∇s stands for the gradient of s.
Many standard analysis can be carried out. We omit the details.

The difficulty as any conventional method

We have already seen that a linear system associated with least squares method has a problem that

$$A_{ij} \neq 0, \forall i, j = 1, \cdots, n$$

and n is usually large for a problem in $\mathbb{R}^2$ or $\mathbb{R}^3$. Solution of $A\mathbf{c} = \mathbf{b}$ is not easy when n is large. Similar for discrete least squares, penalized least squares methods, etc..
For integral equation of the second kind, A multiscale method was proposed to reduce the computational difficulty. See [Chen, Micchelli, Xu, 2015[19]] which summarizes many research results developed in the last 20 years by these researchers and their collaborators.

[19] Z. Chen, C. A. Micchelli, Y. Xu, Multiscale Methods for Fredholm Integral Equations, Cambridge University Press, 2015

A Multiscale Method

Following the ideas in [Chen, Micchelli, and Xu, 2015], let $S_0 = S^r_d(\triangle)$ be a spline space over a triangulation $\triangle$ of domain Ω and $S_k = S^r_d(\triangle_k)$ for $k \geq 1$, where $\triangle_k$ is the kth uniform refinement of $\triangle$.
It is known that $S_k \subset S_{k+1}$ and $\bigcup_{k\geq 0} S_k = L^2(\Omega)$. Writing

$$S_k = W_k \oplus S_{k-1}$$

for $k \geq 1$ and letting $W_0 = S_0$, we have

$$L^2(\Omega) = \bigcup_{k\geq 0} W_k.$$

We can use these subspaces $W_k, k \geq 0$ to de-convolution.

Polynomial Decay of Kernel Functions

Recall $K(\mathbf{x}, \mathbf{y}) = K(\mathbf{x} - \mathbf{y})$ has a good decay property if

$$|K(\mathbf{x}, \mathbf{y})| \le \frac{C}{|\mathbf{x} - \mathbf{y}|^{\sigma} + 1}, \quad \forall |\mathbf{x} - \mathbf{y}| \to \infty \tag{19}$$

for some $\sigma > 0$.
Furthermore, K is of smooth decay of order $k \ge 1$ if

$$|D_{\mathbf{x}}^{\alpha} D_{\mathbf{y}}^{\beta} K(\mathbf{x}, \mathbf{y})| \le \frac{C}{|\mathbf{x} - \mathbf{y}|^{\sigma + |\alpha| + |\beta|} + 1}, \quad \forall |\mathbf{x} - \mathbf{y}| \to \infty \tag{20}$$

for some $\sigma > 0$ and $\alpha, \beta \in \mathbb{Z}_{+}^{2}$ with $|\alpha| \le k, |\beta| \le k$.
For example, $K(\mathbf{x}, \mathbf{y}) = \exp(-\alpha |\mathbf{x} - \mathbf{y}|^2)$. For another example, $K(\mathbf{x}, \mathbf{y}) = 1/(|\mathbf{x} - \mathbf{y}| + 1)$.

Advantages of Multiscale Methods

According to [Chen, Micchelli and Xu, 2015], a multiscale method enables us to have a linear system whose coefficient matrix with faster decay property:

Lemma (Chen, Micchelli and Xu, 2015, p. 205)

Suppose that the kernel $K(\mathbf{x}, \mathbf{y})$ is of smooth decay of order $k \geq 1$. Then for $\phi_i \in W_i$ and $\phi_j \in W_j$ with $i \neq j$,

$$|A_{ij}| \leq C2^{-k(i+j)}, \tag{21}$$

if W_0 contains all polynomials of degree $\leq k$.

Their construction of $W_j, j \geq 1$ is based on discontinuous piecewise polynomial functions over a rectangular domain based on tensor products. How to construct $W_j, j \geq 1$ for spline space $S_d^r(\triangle)$ over an arbitrary triangulation of any polygonal domain Ω is still a **difficulty**.

Our Construction of $W_j, j \geq 1$

Construction of W_j is not easy, in particular, when $r \geq 1$.
Let me explain how I do it.

- 1) Consider $\tilde{S}_j = S_d^{-1}(\triangle_j)$ first.
  For any $s \in \tilde{S}_j$, we can write s in terms of $\tilde{S}_{j+1}$. As these spline spaces are of finite dimensions, we can express in terms of the coefficient vector of a spline $s \in \tilde{S}_j$. Let $\mathbf{c}_j$ be the coefficient vector of $s \in \tilde{S}_j$.
- 2) Next for $s \in S_d^r(\triangle_j), r \geq 0$, we have a smoothness matrix H such that $H_j \mathbf{c}_j = 0$.
- 3) Since $\tilde{S}_j \subset \tilde{S}_{j+1}$, we can write any spline $s \in \tilde{S}_j$ in term of $\mathbf{c}_{j+1}$. One can find a refinement matrix R_j such that

$$\mathbf{c}_{j+1} = R_j \mathbf{c}_j \tag{22}$$

  with $H_j \mathbf{c} = 0 = H_{j+1} \mathbf{c}_{j+1}$. Note that R_j is a tall matrix.

Our Construction of $W_j, j \geq 1$(II)

We need spline $w \in W_{j+1} \subset S^r_d(\triangle_{j+1})$ with coefficent vector $\mathbf{d}_{j+1} = Q_j \mathbf{a}_{j+1}$ with Q_j being a tall matrix such that

$$0 = \int_\Omega s_j w dxdy = \mathbf{c}_{j+1}^\top M_{j+1} \mathbf{d}_{j+1} = \mathbf{c}^\top R_j^\top M_{j+1} Q_j \mathbf{a}_{j+1}, \qquad (23)$$

where M_{j+1} is a mass matrix introduced before.

For R_j, we can find Q_j such that $R_j^\top M_{j+1} Q_j = 0$. That is, Q_j is matrix whose columns spanning the null space $M_{j+1} R_j$.

In order to make w whose coefficient vector $Q_j \mathbf{a}$ in $S^r_d(\triangle_{j+1})$, $Q_j \mathbf{a}$ must satisfy the smoothness conditions. Thus, we look for

$$Q_j \mathbf{a} \in \text{ span } \{Q_j \mathbf{a} : H_{j+1} Q_j \mathbf{a}_j = 0\}, \qquad (24)$$

where H_{j+1} is the smoothness condition matrix.

In order to find some locally supported orthogonal basis functions in W_{j+1}, we use a compressive sensing technique to solve

$$\min \|Q_j \mathbf{a}\|_0 : H_{j+1} Q_j \mathbf{a} = 0, \mathbf{a} \neq 0\}. \qquad (25)$$

More study on the construction

One is able to find a few locally supported basis functions in W_j. However, there are many globally supported basis functions.

Research Problem: Does there exist a set of locally supported functions in W_j spanning W_j?

Research Problem: if they exist, are they scaleble?

Spline Approximation of Kernels

For a continuous function $F(\mathbf{x})$ defined on $\mathbf{x} \in \Omega$, we can use spline interpolation of F over all domain points of degree d on triangle T for all $T \in \triangle$ to approximate F. That is, let

$$S_F(\mathbf{x}) = \sum_{T \in \triangle} \sum_{i+j+k=d} c^F_{ijk,T} \phi_{ijk,T}(\mathbf{x})$$

be the spline interpolation of F satisfying

$$S_F(\xi^T_{ijk}) = F(\xi^T_{ijk}), \quad \forall i+j+k=d, T \in \triangle.$$

where $\phi_{ijk,T}$ is a Bernstein-Bézier polynomial of degree d supported only on triangle $T \in \triangle$. This can be done easily.

Spline Approximation of Kernels (II)

For any continuous function $K(\mathbf{x}, \mathbf{y})$ with $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ defined on $\Omega \times \Omega$, we let

$$S_K(\mathbf{x}, \mathbf{y}) = \sum_{T_1 \in \triangle} \sum_{T_2 \in \triangle} \sum_{i+j+k=d} \sum_{o+p+q=d} c^{T_1, T_2}_{ijk,opq} \phi_{ijk,T_1}(\mathbf{x}) \phi_{opq,T_2}(\mathbf{y})$$

be a spline interpolant of $K(\mathbf{x}, \mathbf{y})$ satisfying

$$S_K(\xi^{T_1}_{ijk}, \xi^{T_2}_{opq}) = K(\xi^{T_1}_{ijk}, \xi^{T_2}_{opq}), \quad \forall i+j+k=d, o+p+q=d, \tag{26}$$

for all $T_1, T_2 \in \triangle$, where $\xi^T_{ijk}, i+j+k=d$ are domain points on triangle T for all $T \in \triangle$ and $\phi_{ijk,T}, i+j+k=d$ are Bernstein-Bézier polynomials defined on triangle T for all $T \in \triangle$.
There is a way to use $O(N^2_\triangle)$ to compute S_K, where $N_\triangle$ stands for the number of triangles in $\triangle$.

Spline Approximation of Kernels: Numerical Results

Let us show $\|S_K(\mathbf{x},\mathbf{y}) - K(\mathbf{x},\mathbf{y})\|_\infty$ as follows.

functions	8 triangles	32 triangles	128 triangles
$\exp(-10((x_1-y_1)^2+(x_2-y_2)^2))$	0.1889	0.0760	0.0229
$\frac{1}{10+(x_1-y_1)^2+(x_2-y_2)^2}$	0.000940	0.000238	0.0000599
$\log(1+((x_1-y_1)^2+(x_2-y_2)^2))$	0.0876	0.0234	0.0060
$1+(x_1-y_1)^2+(x_2-y_2)^2$	0.0960	0.0240	0.0060

Table : Maximum Errors for various functions

Summary

- De-convolution is an inverse problem and has many applications of potential importance.
- We have discussed how to use bivariate splines for numerical de-convolution. A few new approaches are proposed. Approximation properties are studied. It is interesting to know how well we have

$$\|f - S_f\|_{L^\infty(\Omega)}, \text{ and } \|f - S_f\|_{L^1(\Omega)}. \tag{27}$$

- A coercive concept for kernel functions is introduced. What kind of kernels $K(\mathbf{x}, \mathbf{y})$ are coercive in general?
- De-convolution is computationally expensive. It requires more in-depth study. More study on construction of the orthogonal complements is needed.
- Our study can be easily extended to the 3D setting by using trivariate spline functions.
- We also plan to extend our study to deal with blind de-convolution problem.

The End

Thank you $\ddot{\smile}$

mjlai@uga.edu