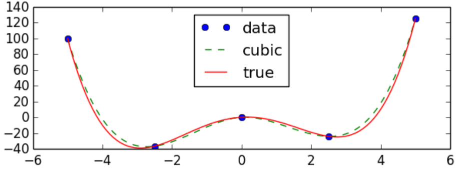
RBF Response Surfaces with Inequality Constraints

David Bindel

Department of Computer Science
Cornell University

14 March 2015

Background: Surrogate-based global optimization



Goal: Optimize

$$f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$$

Assume

- Ω compact (usually a rectangular prism)
- f may be "nice", but is black-box
- Evaluating f is expensive

Idea: Sample, fit a surrogate $\hat{f}$, repeat.

Motivation: Partial information and gray boxes

Costly to compute $f(x)$, but may get bounds fast:

- Trivial bounds (e.g. $0 \leq f(x) \leq 1$)
- Nontrivial-but-cheap bounds (e.g. via Taylor expansion)
- Iterates of a solver (e.g. via bisection)
- Partial sum of a separable function, e.g.

$$f(x) = \sum_{j=1}^{m} \|g(x_j) - g_j\|^2$$

Goal:

- Incorporate bounds into surrogate (today).
- Don't finish unpromising evaluations (another time).

Radial basis function (RBF) approximation

$$s(x) = \sum_{j=1}^{N} c_j \phi(\|x - x_j\|) + p(x)$$

- $X = \{x_i\}_{i=1}^N$ is the set of *centers*
- $\phi : \mathbb{R} \to \mathbb{R}$ is a *radial basis function*
- $p \in \mathcal{P}_{d-1}$ is a *polynomial tail*

Interpolate f at $\{x_j\}_{j=1}^N$ and satisfy discrete orthogonality:

$$\sum_{j=1}^{N} c_j q(x_j) = 0, \quad \forall q \in \mathcal{P}_{d-1}$$

RBF interpolation

Given basis $\{p_j(x)\}$ for $\mathcal{P}_{d-1}$, interpolation system is

$$\begin{bmatrix} \Phi & \Pi \\ \Pi^T & 0 \end{bmatrix} \begin{bmatrix} c \\ a \end{bmatrix} = \begin{bmatrix} f_X \\ 0 \end{bmatrix}$$

where

- $c = \begin{bmatrix} c_1 & \ldots & c_N \end{bmatrix}^T$ is the coefficient vector
- $p(x) = \sum_j a_j p_j(x)$ is the polynomial tail
- $\Pi_{ij} = p_j(x_i)$
- $\Phi_{ij} = \phi(\|x_i - x_j\|)$

When is this well posed? When is there an "energy"?

Conditional positive definite RBFs

[Micchelli, 1986]: ϕ is *conditionally positive definite of order* d if for all $X = \{x_1, \ldots, x_N\}$ distinct and $c \neq 0$ s.t.

$$\sum_{j=1}^{N} c_j q(x_j) = 0, \quad \forall q \in \mathcal{P}_{d-1},$$

we have that

$$\sum_{i,j} c_i c_j \phi(\|x_i - x_j\|) > 0$$

Conditional positive definite RBFs

		$\phi(r)$	Order
Cubic	Schoenberg, 1946	r^3	2
Thin-plate	Duchon, 1976	$r^2 \log r$	2
Multiquadric	Hardy, 1968	$-\sqrt{\gamma^2 + r^2}$	1
Inverse multiquadric		$(\gamma^2 + r^2)^{-1/2}$	0
Gaussian		$\exp(-r^2/\gamma^2)$	0

Conditional positive definite RBFs

For an appropriate degree tail, the interpolation system

$$\begin{bmatrix} \Phi & \Pi \\ \Pi^T & 0 \end{bmatrix} \begin{bmatrix} c \\ a \end{bmatrix} = \begin{bmatrix} f_X \\ 0 \end{bmatrix}$$

is the KKT system for

$$\min \frac{1}{2} c^T \Phi c - c^T f_X \text{ s.t. } \Pi^T c = 0.$$

Optimization well-posed if Π is full rank
$\equiv q \in \mathcal{P}_{d-1}$ uniquely identified by values on X.

Physically: Problem is statically determinate (no rigid-body modes).

Energy interpretation

Two splines with form

$$s(x) = \sum_{j=1}^{N} c_j \phi(\|x - x_j\|) + p(x)$$

Define a semi-definite form ("energy semi-inner product")

$$(s, \tilde{s}) = \sum_{j,k} c_j \tilde{c}_k \phi(\|x_j - x_k\|) = \sum_j c_j \tilde{s}(x_j)$$

Corresponding semi-norm is $|s| = (s, s)^{1/2}$.

Native space $\equiv$ closure of set of splines under semi-norm.
Interpolating spline minimizes $|s|$ under interpolation constraints.

Incorporating bounds

Set $X = E \cup B$ where

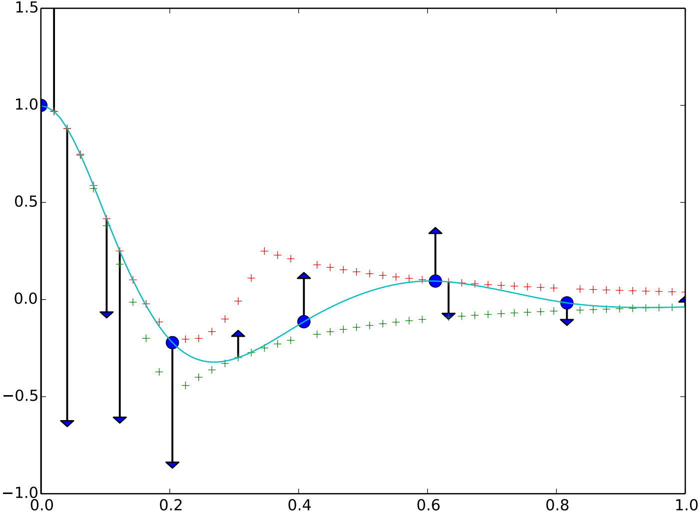
$$E = \{x_j\}_{j=1}^{|E|}, \qquad s(x_i) = f(x_i)$$
$$B = \{x'_j\}_{j=1}^{|B|}, \qquad -\infty \leq \ell_i \leq s(x_i) \leq u_i \leq \infty.$$

and minimize $|s|$ subject to these constraints. KKT conditions:

$$\begin{aligned} s(x_i) &= f(x_i) \\ s(x'_i) &= \ell_i \implies c'_i \geq 0 \\ s(x'_i) &= u_i \implies c'_i \leq 0 \\ \ell_i \leq s(x'_i) &\leq u_i \implies c'_i = 0. \end{aligned}$$

Include “forces” (nonzero coeffs) to push surface within bounds.

Incorporating bounds



Correction formulation

Suppose

$$
\begin{aligned}
s_0(x) &= \text{spline with only interpolation} \\
s(x) &= \text{spline with interpolation + bounds} \\
&= s_0(x) + G(x)
\end{aligned}
$$

Note that

$$|s|^2 = |s_0|^2 + 2(s_0, G) + |G|^2 = |s_0|^2 + |G|^2.$$

Minimizing $|G|$ means minimizing $|s_0|$

Error analysis

Set

$$
\begin{aligned}
f &\in \text{native space for } \phi \\
s(x) &= \text{spline with only interpolation} \\
\tilde{s}(x) &= \text{spline with interpolation + bounds}
\end{aligned}
$$

Standard analysis: bound error semi-norm $|f - s|$ and apply

$$
|f(x) - s(x)| \leq P(x)|f - s| \qquad |f(x) - \tilde{s}(x)| \leq P(x)|f - \tilde{s}|
$$

Error:

$$
\begin{aligned}
|f - s|^2 = |(f - \tilde{s}) + (\tilde{s} - s)|^2 &= |f - \tilde{s} + G|^2 \\
&= |f - \tilde{s}|^2 + 2(f - \tilde{s}, G) + |G|^2.
\end{aligned}
$$

Error analysis

Write inner product term as

$$(f - \tilde{s}, G) = \sum_{j=1}^{|B|} (f(x_i) - \tilde{s}(x_i))c_i'$$

and complementarity yields

$$\begin{aligned} c_i' > 0 &\implies \tilde{s}(x_i) = \ell_i \leq f(x_i) \\ c_i' < 0 &\implies \tilde{s}(x_i) = u_i \geq f(x_i) \end{aligned}$$

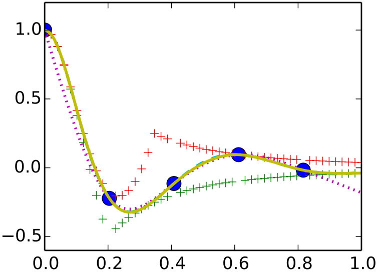
Either way,

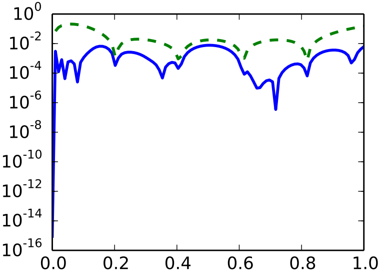
$$(f(x_i) - \tilde{s}(x_i))c_i' \geq 0.$$

Therefore $(f - \tilde{s}, G) \geq 0$, and so

$$\begin{aligned} |f - \tilde{s}|^2 &= |f - s|^2 - 2(f - \tilde{s}, G) - |G|^2 \\ &\leq |f - s|^2 - |G|^2 \end{aligned}$$

Error example





Correction formulation, concretely

Coefficients for G satisfy

$$\begin{bmatrix} \Phi_{BB} & \Phi_{BE} & \Pi_B \\ \Phi_{EB} & \Phi_{EE} & \Pi_E \\ \Pi_B^T & \Pi_E^T & 0 \end{bmatrix} \begin{bmatrix} c' \\ \Delta c \\ \Delta a \end{bmatrix} = \begin{bmatrix} g \\ 0 \\ 0 \end{bmatrix}$$

where $g = G_B$ satisfies $\hat{\ell} \leq g \leq \hat{u}$ for $\hat{\ell} = \ell - s_X$ and $\hat{u} = u - s_X$.

Eliminate Δc and Δa to get Schur complement

$$S = \Phi_{BB} - \begin{bmatrix} \Phi_{BE} & \Pi_B \end{bmatrix} \begin{bmatrix} \Phi_{EE} & \Pi_E \\ \Pi_E^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} \Phi_{EB} \\ \Pi_B^T \end{bmatrix}$$

Now c' minimizes $(c')^T S c'$ subject to $\hat{\ell} \leq g \leq \hat{u}$.

Solver strategy

Active set method to solve

$$c' = \text{argmin}_d \{d^T S d : \hat{\ell} \leq Sd \leq \hat{u}\}$$

- Allows warm-start in context of global optimization.
- Maintain $QSQ^T = R^T R$.
- Permutation Q moves free variables toward front.
- Update Q and R in $O(N^2)$ time per iteration. (Actually $O(Nm)$ where m is distance variable moves.)

Adding points or updating bounds tends to be cheap.

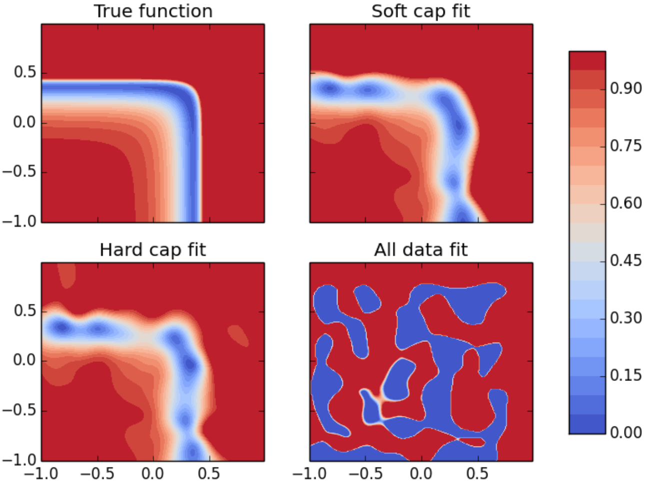
Example: Capped surfaces

Consider

$$f(x,y) = \sum_{j=1}^{10} [2 + 2j - \exp(jx) - \exp(jy)]^2$$

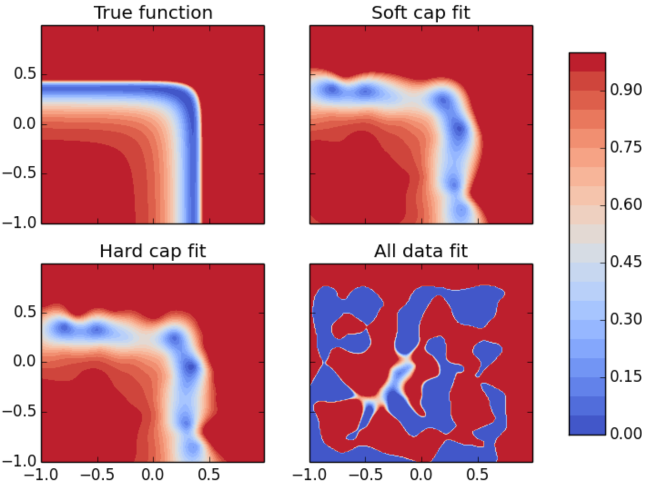
- Large function values can cause interpolant to oscillate
- Idea: replace large function values with a “cap”
- Hard cap: interpolate $\min(f(x,y), M)$
- Soft cap: replace large $f(x_i, y_i)$ with $f(x_i, y_i) \geq M$

Cubic

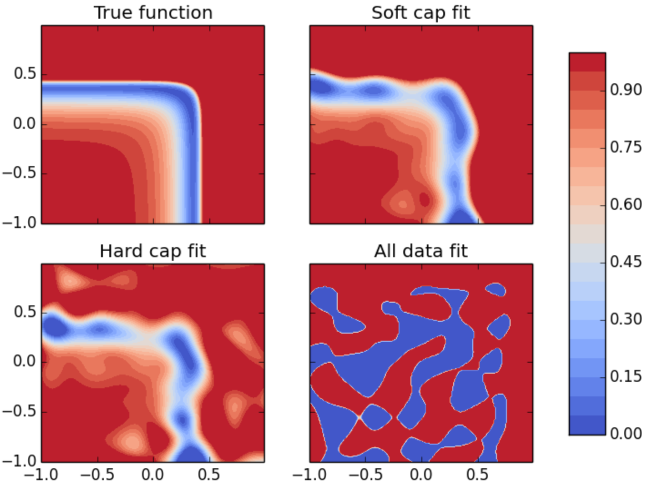


True function
Soft cap fit
Hard cap fit
All data fit
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
0.90
0.75
0.60
0.45
0.30
0.15
0.00

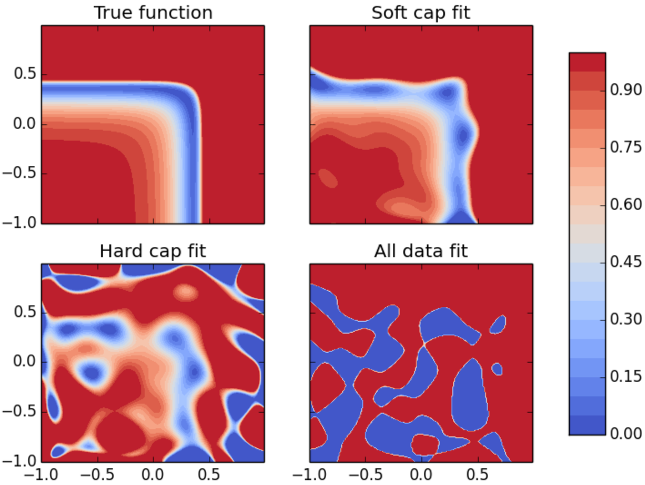
Thin-plate


True function
Soft cap fit
Hard cap fit
All data fit
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
0.90
0.75
0.60
0.45
0.30
0.15
0.00

Multiquadric



Gaussian



True function
Soft cap fit
Hard cap fit
All data fit
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
0.90
0.75
0.60
0.45
0.30
0.15
0.00

Conclusions

- Energy interpretation of RBF interpolant
  $\implies$ Easy to add upper/lower bounds at points
- Python+C code (PyRBFbound) is now public on BitBucket:

  `https://bitbucket.org/dbindel/pyrbfbound/`
- Lots of possible extensions
  - Point bounds + RBF-QR and contour-Padé interpolation
  - Incorporation with scalable solvers (e.g. FMM) for 2D/3D
  - Enforcing continuous lower/upper bounds, integral bounds