On some insight and extensions of the Radial Basis Function method.

Andrea Cassioli[1]

[1]LIX - ECOLE POLYTECHNIQUE (FR)
cassioli@lix.polytechinique.fr

11/07/2013

Outline

1. Brief summary of the RBF method
2. Some more insight on the bumpiness
3. Other variants of the RBF method

The RBF method

Problem definition

We aim to solve

$$\min_{x \in X} f(x)$$

where:

- $X \subset \mathbb{R}^n$ is the (bounded) feasible set of x
- $f(x)$ is a black-box function whose evaluation is "**costly**"

The RBF method

Basic building blocks

For a given (unisolvent) set of samples S:

$$s(x|S) = \sum_{y \in S} \lambda_y \phi(\|x - y\|) + p(x, c) = \Phi(x)\lambda + P(x)^T c$$

Coefficients λ, c are computed solving the linear system

$$\Phi\lambda + Pc = f$$
$$P^T \lambda = 0$$

the degree of the polynomial depends on $\phi()$.

The RBF method

Basic algorithm

determine suitable S_0;
$k \leftarrow 0$;
while *stopping criteria not fulfilled* **do**

- determine $s_k(x|S_k)$;
- $y_k = \arg\min_{x \in X} s_k(x)$;
- select the aspiration level $\hat{f}$;
- determine x_{k+1} based on $\mu(x|y_k, \hat{f})$;
- $S_{k+1} = S_k \cup x_k + 1$;

end

The RBF method
Comments

A lot of freedom:

- which radial basis
- the degree of the polynomial
- how solve the auxiliary problems
- how to select the reference value

The RBF method

Do we really care about convergence?

"In principle, these methods may have convergence guarantees if the point selection strategy is well-chosen; but this is irrelevant in view of the fact that for expensive functions, only few (perhaps up to 1000) function evaluations are admissible"[a]

[a] Arnold Neumaier, "Complete search in continuous global optimization and constraint satisfaction", *Acta numerica* 13.1 (2004): 271–369.

The RBF method

Convergence

Based on the well known theorem of Törn, A., Zilinskas, A.:

Theorem

If an algorithm generates a sequence of points that are dense in in the feasible set X it converges to the optimal solution.

Basically we will get arbitrary close to optimum...

Theorem

If an RBF method is well posed (see usual properties of $s()$*) and*

- S_k *is unisolvent*
- *for the reference value holds that*

$$\hat{f}_{k+1} < \min_x s(x|S_k)$$

- x_{k+1} *is a minimizer(maximizer) for the bumpiness function*

then the point selected at iteration $k+1$ *is distinct for any other points in* S_k*.*

Corollary

If for the designed RBF method the previous theorem holds, then the method converges to the global optimum of $f(x)$.

Proof.

If the previous theorem holds, then the sequence of points $\{x_i\}$ is dense in X and for the Torn and Zilinskas theorem we converge to the global optimum. □

The RBF method

Convergence

- Granted if an infinite subsequence of sampled point is dense in X
- In some cases convergence to first-order stationary points
- In probability if we can sample the feasible set along $\{x_k\}$

On the bumpiness function

A step back on the interpolant function

Let consider 1D cubic splines for a set $\{x_1, \ldots, x_k\}$:

- must fulfill some condition on first/second derivatives
- it's *natural* if $s^{''}(x) = 0$
- minimize

  $$I(s) = \int_{-\infty}^{\infty} s^{''}(x)^2 dx$$

  which a curvature measure

The 1D cubic RBF is a natural spline

$$s(x) = \sum_{i=1}^{k} \lambda_i |x - x_i|^3 + c_1 + c_2 x$$

,

On the bumpiness function

A step back on the interpolant function

Generalize to a general radial basis $\phi()$, we obtain

$$I(s) = \int_{-\infty}^{\infty} s''(x)^2 dx = \ldots = 12\lambda^T \Phi \lambda + 12 P \lambda,$$

but asking for a natural spline we get

$$I(s) = 12\lambda^T \Phi \lambda$$

An 1D RBF is the natural spline for that basis and set of points.

On the bumpiness function

A step back on the interpolant function

Moving to the multidimensional case, we note that the $I(s)$ function comes from the product

$$< u(x,\lambda), v(x,\mu) >= (-1)^m \sum_{i=1}^{k} \lambda_i v(x_i) = (-1)^m \sum_{i=1}^{k} \mu_i u(x_i),$$

yielding

$$< s, s >= (-1)^m \lambda^T \Phi \lambda$$

which is a seminorm once $P^T\lambda = 0$.

On the bumpiness function
Meaning

For the surrogate model, *centers* are fixed, and we look for the λ's.

For the *bumpiness*, one center is not fixed (the next point) and we minimize the seminorm of

$$s(x|S \cup \hat{x}) = s(x|S) + (\hat{f} - s(\hat{x}|S))\hat{L}(x)$$

where L is an RBF of the same family that attains 1 in $\hat{x}$ and 0 everywhere else.

On the bumpiness function

Meaning

How to think about the bumpiness?
Imagine:

- an elastic carpet that has be fixed in points at certain heights
- put your finger at the aspiration level
- move it until you find the point in which the carpet resists less to your pressure
- this is the next point!

On the bumpiness function

Pros...

- ”simple” method
- a meaningful concept
- promote convergence
- allow for balancing exploration/intensification via the aspiration level

On the bumpiness function

...and Cons

- hard to optimize (very bumpy...)
- numerically unstable (log scaling)
- boundary ”effect”
- difficult to relate to the geometry of S
- requires a (good) lower bound on the optimal value of the surrogate model

CORS-RBF[2]

1. use the surrogate model has merit function solving

$$\min_{x \in X} s(x)$$
$$\|x - x_i\| \geq \Delta_i \qquad i = 1 \ldots k$$

2. Parallel version[1]

[1] Rommel Gagalac Regis, "Global optimization of computationally expensive functions using serial and parallel radial basis function algorithms", Diss., 2004.

[2] Rommel G Regis and Christine A Shoemaker, "Constrained global optimization of expensive black box functions using radial basis functions", *Journal of Global Optimization* 31.1 (2005): 153–171.

SRS[3]

1. next point among a pool of perturbations of the best solution so far
2. putative points scored using $s()$ and/or the geometry of the sample set
3. several variants depending on the scoring and globalization strategies
4. convergence in probability

[3] Rommel G Regis and Christine A Shoemaker, “A stochastic radial basis function method for the global optimization of expensive functions”, *INFORMS Journal on Computing* 19.4 (2007): 497–509.

ConstrLMSRBF[4]

1. extends the LMSRBF version of SRS
2. build surrogates models for both objective function and contraints
3. require a first feasible point
4. consider feasibility violation in scoring the putative points
5. tested up to 4000 function evaluations

[4] Rommel G Regis, "Stochastic radial basis function algorithms for large-scale optimization involving expensive black-box objective and constraint functions", *Computers & Operations Research* 38.5 (2011): 837–853.

BOOSTER[5]

A Trust-Region based RBF method (no bumpiness):

1. test the model for *”validity”* and add new points if necessary
2. find a minimizer of the model in the TR
3. compute the improvement ratio
4. update TR

[5] Rodrigue Oeuvray and Michel Bierlaire, “BOOSTERS: A derivative-free algorithm based on radial basis functions”, *International Journal of Modelling & Simulation* 29.1 (2009): 26.

ORBIT[6][7]

Extends BOOSTER with a more complex handling of the TR.

1. use only a subset of samples
2. the surrogate is built enforcing well conditioning (fully linearity)
3. the next point is the (approximate) minimizer of the surrogate on the TR
4. very complex framework

---

[6] Stefan M Wild, Rommel G Regis, and Christine A Shoemaker, "ORBIT: Optimization by radial basis function interpolation in trust-regions", *SIAM Journal on Scientific Computing* 30.6 (2008): 3197–3219.

[7] Stefan M Wild and Christine Shoemaker, "Global convergence of radial basis function trust region derivative-free algorithms", *SIAM Journal on Optimization* 21.3 (2011): 761–781.

ARBF[8]

1. extends the original RBF method
2. consider BB constraints as penalty
3. select next point using bumpiness
4. aspiration level is varied and putative next points clustered
5. in some cases the aspiration level is ignored and the optimum of $s()$ is used

[8] Kenneth Holmström, “An adaptive radial basis algorithm (ARBF) for expensive black-box global optimization”, *Journal of Global Optimization* 41.3 (2008): 447–464.

ARBF

At iteration k, the Grid Mode uses a set $w = \{w_1 \dots w_t\}$ of positive weight and determines

$$x_i = \arg\min_{x \in X} \mu(x, s_k - w_i f_\Delta) \quad i = 1 \dots t,$$

points are then clustered[9] and one is selected using heuristics.

[9] Donald R Jones, "A taxonomy of global optimization methods based on response surfaces", *Journal of global optimization* 21.4 (2001): 345–383.

qualSolve[10]

1. use an alternative merit function
2. consider approximation and interpolation
3. extends to multi-objective optimization

[10] Stefan Jakobsson et al., "A method for simulation based optimization using radial basis functions", *Optimization and Engineering* 11.4 (2010): 501–532.

qualSolve

From interpolation to approximation $\eta[0,1]$:

$$\begin{aligned} \min \quad & \eta\lambda^T\Phi\lambda + (1-\eta)\|\epsilon\|^2 \\ s.t. \quad & \\ & \Phi\lambda + Pc = \epsilon + f \\ & P^T\lambda = 0 \\ & \epsilon \in \mathbb{R}^k \end{aligned}$$

1. $\eta \to 0$ yields original RBF method
2. $\eta \to 1$ yield the smoothest surrogate model

The choice of η can be done using cross-validation.

qualSolve

It maximizes

$$Q(y) = \int_{\Omega} (U_S(x) - U_{S \cup y}(x)) \omega(s(x|S)) dV(x)$$

where $\omega()$ is a suitable weight function and

$$U(x) = \min_{z \in S} \|x - z\|$$

References I

Holmström, Kenneth. "An adaptive radial basis algorithm (ARBF) for expensive black-box global optimization". *Journal of Global Optimization* 41.3 (2008): 447–464. Print.

Jakobsson, Stefan, et al. "A method for simulation based optimization using radial basis functions". *Optimization and Engineering* 11.4 (2010): 501–532. Print.

Jones, Donald R. "A taxonomy of global optimization methods based on response surfaces". *Journal of global optimization* 21.4 (2001): 345–383. Print.

Neumaier, Arnold. "Complete search in continuous global optimization and constraint satisfaction". *Acta numerica* 13.1 (2004): 271–369. Print.

References II

Oeuvray, Rodrigue and Michel Bierlaire. “BOOSTERS: A derivative-free algorithm based on radial basis functions”. *International Journal of Modelling & Simulation* 29.1 (2009): 26. Print.

Regis, Rommel G. “Stochastic radial basis function algorithms for large-scale optimization involving expensive black-box objective and constraint functions”. *Computers & Operations Research* 38.5 (2011): 837–853. Print.

Regis, Rommel G and Christine A Shoemaker. “A stochastic radial basis function method for the global optimization of expensive functions”. *INFORMS Journal on Computing* 19.4 (2007): 497–509. Print.

——. .“Constrained global optimization of expensive black box functions using radial basis functions”. *Journal of Global Optimization* 31.1 (2005): 153–171. Print.

References III

- Regis, Rommel Gagalac. "Global optimization of computationally expensive functions using serial and parallel radial basis function algorithms". Diss. 2004. Print.
- Wild, Stefan M, Rommel G Regis, and Christine A Shoemaker. "ORBIT: Optimization by radial basis function interpolation in trust-regions". *SIAM Journal on Scientific Computing* 30.6 (2008): 3197–3219. Print.
- Wild, Stefan M and Christine Shoemaker. "Global convergence of radial basis function trust region derivative-free algorithms". *SIAM Journal on Optimization* 21.3 (2011): 761–781. Print.