Hindawi
Discrete Dynamics in Nature and Society
Volume 2017, Article ID 4082432, 9 pages
https://doi.org/10.1155/2017/4082432

Research Article

A Frame-Based Conjugate Gradients Direct Search Method with Radial Basis Function Interpolation Model

Xiaowei Fang[1,2] and Qin Ni[1]

[1]*College of Sciences, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*
[2]*Department of Mathematics, Huzhou University, Huzhou 313000, China*

Correspondence should be addressed to Xiaowei Fang; fangxiaowei@163.com

Received 20 July 2016; Accepted 19 December 2016; Published 19 January 2017

Academic Editor: Rigoberto Medina

In this paper, we propose a new hybrid direct search method where a frame-based PRP conjugate gradients direct search algorithm is combined with radial basis function interpolation model. In addition, the rotational minimal positive basis is used to reduce the computation work at each iteration. Numerical results for solving the CUTEr test problems show that the proposed method is promising.

1. Introduction

In this paper, we consider the following problem:

$$\min_{x \in \mathbb{R}^n} \quad f(x), \tag{1}$$

where function f is assumed to be continuously differentiable from $\mathbb{R}^n$ into $\mathbb{R}$, and the derivative information is unavailable or untrustworthy, for example, because of noise and using finite differences. Problem (1) has numerous applications in engineering, such as the helicopter rotor blade design [1, 2], the aeroacoustic shape design [3], groundwater community problems [4], and medical image registration problems [5].

There are two main methods for solving (1). The first class of methods is the model based methods, which are constructed by means of multivariate interpolation, including under and overdetermined. These methods were introduced by Powell [6] and Winfield [7] and were developed by [8–11]. The second class of methods is the direct search methods which are based on the comparison rules of objective function values. These methods were pioneered by Hooke and Jeeves [12]. The convergence theory was established by Torczon [13, 14]. Audet and Dennis [15] proposed a general framework for direct search method. Coope and Price [16] extended the PRP method [17, 18] to solve (1) and presented a frame-based conjugate gradients direct search algorithm (Max-PRP for short). In each iteration, Max-PRP employed the fixed maximal positive basis to estimate the first and second gradients; then the search direction is determined by employing the PRP formula. Numerical tests showed that the Max-PRP was effective on a wide variety of unconstrained optimization problem. In addition, some classical and modern direct search methods were introduced by Kolda et al. [19].

Generally, model based methods are more efficient than direct search methods in that they are able to exploit structure inherently in the problem. But direct search methods are simpler to code and to parallelize. Therefore, it is natural to try to combine both methods. In 2010, Custódio et al. [20] proposed a hybrid method integrating minimum Frobenius norm quadratic interpolation models in a direct search framework and numerical results showed that the addition of quadratic interpolation models improved the performance of the direct search method. In 2013, Conn and Le Digabel [21] showed that the use of quadratic interpolation models can improve the efficiency of the mesh adaptive direct search method.

The above hybrid algorithms were based on the quadratic interpolation models. In 2008, Wild et al. [22] presented a new derivative-free algorithm (ORBIT for short), which employed radial basis function (RBF) interpolation models.

The RBF interpolation models allowed ORBIT to interpolate nonlinear functions using fewer function evaluations than the quadratic interpolation models. In 2013, Wild and Shoemaker [23] proved the global convergence of the ORBIT under some mild assumptions. Numerical results showed that the method using RBF interpolation models outperformed methods using quadratic interpolation models.

Motivated by the efficiency of the ORBIT, we propose a new hybrid direct search method, which combines the frame-based conjugate gradients strategies with the RBF interpolation models. In each iteration, a minimal positive basis is used to construct the frame. In a maximal positive basis, $2n$ function values are computed, while, in a minimal positive basis, $n+1$ function values are evaluated. So the computation work in the new hybrid direct search method can be reduced. In addition, when the trial point of RBF interpolation models cannot satisfy the decrease condition, we employ PRP formula to get the search direction, which is similar to the Max-PRP. Furthermore, we rotate the minimal positive basis according to the local topography of objective function, making our method more effective in practice. The convergence is established under some mild conditions. Some numerical results show that the proposed method is promising.

This paper is organized a s follows. In Section 2, we present some basic notions for positive basis, frame, and describe our method. In Section 3, we prove the convergence of the proposed method. In Section 4, numerical results show the efficiency of method derived in this paper compared to Max-PRP [16]. Concluding remarks are given in Section 5. The default norm used in this paper is Euclidean.

2. The New Hybrid Direct Search Method

We first state the definition about positive basis, which can be found in [24].

Definition 1. Positive basis $\mathscr{V}$ in $\mathbb{R}^n$ is a set of vectors with the following two properties:

(i) Every vector in $\mathbb{R}^n$ is a nonnegative linear combination of the members of $\mathscr{V}$.

(ii) No proper subset of $\mathscr{V}$ satisfies (i).

It is easy to know that cardinality of any positive basis $\mathscr{V}$ satisfies $n+1 \le |\mathscr{V}| \le 2n$. Two famous and simple examples of positive bases are

$$\mathscr{V}_{\min} = \left\{ v_1, \dots, v_n, -\sum_{i=1}^{n} v_i \right\}, \quad \mathscr{V}_{\max} = \{ v_1, \dots, v_n, -v_1, \dots, -v_n \}, \tag{2}$$

where $\{v_1, \dots, v_n\}$ is a basis for $\mathbb{R}^n$, $\mathscr{V}_{\min}$ represents the minimal positive basis, and $\mathscr{V}_{\max}$ represents the maximal positive basis.

In addition, we give some concepts about frames, which were proposed by Coope and Price [25, 26].

Definition 2. A frame can be defined as

$$\Phi = \{ x + hv \colon v \in \mathscr{V} \}, \tag{3}$$

where $x \in \mathbb{R}^n$ is a central point of a frame, $h > 0$ is frame size, and $\mathscr{V}$ is a positive basis in $\mathbb{R}^n$.

Definition 3. A frame Φ is a minimal frame, if and only if

$$f(x) \le f(y), \quad \forall y \in \Phi. \tag{4}$$

Definition 4. A frame Φ is a quasi minimal frame if and only if

$$f(x) \le f(y) + \epsilon, \quad \forall y \in \Phi, \tag{5}$$

where $\epsilon = h^{1+\mu}$, μ is a positive constant, and the corresponding central point x is called a quasi minimal point.

Let x_k be kth iterate. We will discuss the strategy of RBF interpolation model, search direction, and rotation of positive basis in detail below.

2.1. RBF Interpolation Model. Choose a positive basis $\mathscr{V}_k = \{v_1^k, v_2^k, \dots, v_q^k\}$ $(q \ge n+1)$ and obtain a set of interpolate data points $Y = \{y_1, y_2, \dots, y_{n_q}\} \in \mathbb{R}^n$ $(n_q \ge q+1)$, where $y_1 = x_k$, $y_2 = x_k + h_k v_1^k, \dots, y_{q+1} = x_k + h_k v_q^k$, $h_k > 0$ is the frame size, and the other points of set Y are chosen in the subset of previously evaluated points.

RBF interpolation model is a popular model for optimization, and some theory and implementations can be found in [27]. Corresponding to the set of interpolate data points Y, we get the following RBF interpolation model:

$$\widehat{m}_k(x) = \sum_{i=1}^{n_q} \lambda_i \phi(\|x - y_i\|) + \sum_{j=1}^{r} \gamma_j p_j(x), \tag{6}$$

where $\phi : \mathbb{R}_+ \to \mathbb{R}$ is a radial basis function and $\lambda_1, \dots, \lambda_{n_q}, \gamma_1, \dots, \gamma_r \in \mathbb{R}$ are parameters to be determined. $p_1, \dots, p_r$ are polynomial tails used in the context of RBF interpolation models, which most frequently are linear.

In addition, coefficients $\lambda_1, \dots, \lambda_{n_q}$ are required to satisfy

$$\sum_{i=1}^{n_q} \lambda_i p_j(y_i) = 0, \quad j = 1, \dots, r. \tag{7}$$

These, in conjunction with n_q interpolation conditions $\widehat{m}_k(y_l) = f(y_l)$, $l = 1, \dots, n_q$.

We define the linear system:

$$\begin{bmatrix} \Phi & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} \Lambda \\ \Gamma \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix}, \tag{8}$$

where $\Lambda = [\lambda_1, \dots, \lambda_{n_q}]^T$, $\Gamma = [\gamma_1, \dots, \gamma_r]^T$, $\mathbf{f} = [f(y_1), \dots, f(y_{n_q})]^T$, and $\Phi_{li} = \phi(\|y_l - y_i\|)$ for $l, i \in \{1, \dots, n_q\}$, $P_{ij} = p_j(y_i)$ for $i \in \{1, \dots, n_q\}$, $j \in \{1, \dots, r\}$. We employ null-space method to solve system (8), which is similar to the approach of [22].

Then, we minimize the RBF interpolation model by solving the following problem:

$$\min_{x\in B(x_k;\Delta_k)} \widehat{m}_k(x), \tag{9}$$

where $B(x_k;\Delta_k) = \{x \in \mathbb{R}^n: \|x - x_k\| \leq \Delta_k\}$, $\Delta_k = \delta_k h_k \max\{\|v_1^k\|, \dots, \|v_q^k\|\}$, and δ_k is the radius factor parameter.

2.2. PRP Direction. Consider the following linear model:

$$\overline{m}_k(x) = f(x_k) + g_k^T(x - x_k), \tag{10}$$

where $g_k \in \mathbb{R}^n$. The coefficients can be determined by q regression interpolation conditions:

$$\overline{m}_k\left(x_k + h_k v_l^k\right) = f\left(x_k + h_k v_l^k\right), \quad l = 1, \dots, q. \tag{11}$$

Then, we have that

$$f\left(x_k + h_k v_l^k\right) = f(x_k) + h_k g_k^T v_l^k, \quad l = 1, \dots, q. \tag{12}$$

This system can be solved by the method of least squares. For example, if we choose the positive basis $\mathscr{V}_k$ as $\mathscr{V}_{\min}$, and $v_i = e_i$ $(i = 1, \dots, n)$, where e_i is ith unit vector, then ith element of g_k is calculated according to the following formula:

$$(g_k)_i = \frac{1}{h_k}\left(f(x_k + h_k e_i) - \frac{f(x_k - h_k e) + \sum_{j=1}^n f\left(x_k + h_k e_j\right)}{n+1}\right), \tag{13}$$

where $e = \sum_{j=1}^n e_j$. The PRP direction is obtained by

$$d_k = \begin{cases} -g_k & \text{if } k = 0 \\ -g_k + \max\left\{0, \dfrac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2}\right\} d_{k-1} & \text{if } k > 0. \end{cases} \tag{14}$$

2.3. Rotation of the Positive Basis. In order to modify the positive basis such that at least one of the new directions is more closely conformed to the local behavior of the function, we rotate the positive basis at each step. This idea is similar to that in [28].

Suppose that

$$\mathscr{V}_k = \left\{v_1^k, \dots, v_n^k, \dots, v_q^k\right\}, \tag{15}$$

where $\{v_1^k, \dots, v_n^k\}$ is a basis for $\mathbb{R}^n$. Denote

$$\delta x = x_{k+1} - x_k = \left(\delta x_1, \dots, \delta x_n\right)^T, \tag{16}$$

where $\delta x_i \in \mathbb{R}$ $(i = 1, \dots, n)$ describe the movements performed along the vectors v_i^k $(i = 1, \dots, n)$ in previous iterations.

We get positive basis $\mathscr{V}_{k+1}$ by rotating $\mathscr{V}_k$. Firstly, we obtain n linearly independent vectors according to $\mathscr{V}_k$:

$$\overline{v}_i^{k+1} = \begin{cases} v_i^k & \text{if } \delta x_i = 0 \\ \sum_{l=i}^n \delta x_l v_l^k & \text{if } \delta x_i \neq 0, \end{cases} \tag{17}$$

where $\overline{v}_i^{k+1}$ represents the sum of all the movements made in the directions v_l^k for $l = i, \dots, n$. The lemma 8.5.4 of [29] proved that $\{\overline{v}_1^{k+1}, \dots, \overline{v}_n^{k+1}\}$ is linearly independent.

Secondly, we use the Gram-Schmidt orthogonalization method to get a class of standard orthogonal basis:

$$\left\{v_1^{k+1}, \dots, v_n^{k+1}\right\}. \tag{18}$$

Finally, we can get

$$\mathscr{V}_{k+1} = \left\{v_1^{k+1}, \dots, v_n^{k+1}, v_{n+1}^{k+1}, \dots, v_q^{k+1}\right\}, \tag{19}$$

where $\{v_{n+1}^{k+1}, \dots, v_q^{k+1}\}$ is combined with $\{v_1^{k+1}, \dots, v_n^{k+1}\}$ according to the same combination principal as $\mathscr{V}_k$.

For instance, if $q = n + 1$, we obtain

$$\mathscr{V}_{k+1} = \left\{v_1^{k+1}, \dots, v_n^{k+1}, -\sum_{i=1}^n v_i^{k+1}\right\}. \tag{20}$$

Supposing that $\{z_m\}$ is the sequence of quasi minimal iteration points, then the above process can be summarized as the following algorithm.

Algorithm 5.

Step 0 (initializations). Choose initial point $x_0 \in \mathbb{R}^n$, positive basis $\mathscr{V}_0$, step length h_0, and radius factor parameter δ_0. Choose $\lambda > 1$, $\mu > 0$, $\epsilon_k = h_k^{1+\mu}$. Set $k = 0$, $m = 0$.

Step 1 (checking the stopping condition). If the stopping condition is not met, then go to Step 2, otherwise output the lowest known point and stop.

Step 2 (determining the frame). Create a frame Φ_k at iterate x_k according to the positive basis $\mathscr{V}_k$ and step length h_k, and calculate the corresponding function values.

Step 3 (building the RBF interpolation model). Evaluate the RBF model parameters according to formula (8), and get solution $\widehat{x}_{k+1}$ of subproblem (9). If $f(x_k) - f(\widehat{x}_{k+1}) > \epsilon_k$, then set $x_{k+1} = \widehat{x}_{k+1}$ and go to Step 6, otherwise go to Step 4.

Step 4 (obtaining the PRP direction). Obtain the search direction d_k using (14), execute the line search process to find β_k, and set $\overline{x}_{k+1} = x_k + \beta_k h_k d_k/\|d_k\|$. If $f(x_k) - f(\overline{x}_{k+1}) > \epsilon_k$, then set $x_{k+1} = \overline{x}_{k+1}$ and go to Step 6, otherwise go to Step 5.

Step 5 (updating the current iteration point). Let x_{k+1} be defined by the following rule:

$$f(x_{k+1}) = \min_{v_k \in V_k}\left(f(x_k), f(\widehat{x}_{k+1}), f(\overline{x}_{k+1}), f(x_k + h_k v_k)\right). \tag{21}$$

Step 6 (rotating the positive basis and updating the some parameters). Obtain $\mathcal{V}_{k+1}$ according to (15)–(19) and compute δ_{k+1}. If frame Φ_k is a quasi minimal frame, then set $h_{k+1} = h_k/\lambda$, $m = m + 1$, $z_m = x_k$, otherwise set $h_{k+1} = h_k$. In addition, increment k by one and go to Step 1.

Remark 6. In Step 3, we set $p_1, \ldots, p_r$ as linear polynomial tails, and $\phi(\|x - y_i\|) = \|x - y_i\|^3$ in (8).

3. Convergence Analysis

Now we have the following convergent property of Algorithm 5.

Theorem 7. *Supposing that the sequence of function value $\{f(x_k)\}$ is bounded, then the sequence $\{z_m\}$ is infinite.*

Proof. Assume that $\{z_m\}$ is finite; let z_M be the final quasi minimal point and $z_M = x_{\tilde{k}}$.

From Steps 3, 4, and 5 of Algorithm 5, we know that

$$f(x_{\tilde{k}+1}) < f(x_{\tilde{k}}) - \epsilon_{\tilde{k}}, \tag{22}$$

or

$$f(x_{\tilde{k}+1}) \le \min_{v_{\tilde{k}} \in V_{\tilde{k}}} \left(f(x_{\tilde{k}}), f(x_{\tilde{k}} + h_{\tilde{k}} v_{\tilde{k}}) \right), \tag{23}$$

where $\epsilon_{\tilde{k}} = (h_{\tilde{k}})^{1+\mu}$ $(\mu > 0)$ is a positive constant, $h_{\tilde{k}}$, $V_{\tilde{k}}$ are frame size and positive basis corresponding to iterate $x_{\tilde{k}}$, respectively. Supposing that $\Phi_{\tilde{k}}$ is the frame corresponding to quasi minimal iterate $x_{\tilde{k}}$, then the frame $\Phi_{\tilde{k}+1}$ is not quasi minimal. From Definition 4, it follows that there exists at least a vector $v_l^{\tilde{k}}$ $(v_l^{\tilde{k}} \in \mathcal{V}_{\tilde{k}})$, such that

$$f\left(x_{\tilde{k}} + h_{\tilde{k}} v_l^{\tilde{k}}\right) < f(x_{\tilde{k}}) - \epsilon_{\tilde{k}}. \tag{24}$$

By (22), (23), and (24), we have

$$f(x_{\tilde{k}+1}) < f(x_{\tilde{k}}) - \epsilon_{\tilde{k}}. \tag{25}$$

Then, we have

$$\begin{aligned} f(x_{\tilde{k}+r}) &< f(x_{\tilde{k}+r-1}) - \epsilon_{\tilde{k}+r-1} \\ &< f(x_{\tilde{k}+r-2}) - \sum_{i=\tilde{k}+r-2}^{\tilde{k}+r-1} \epsilon_i < \cdots \\ &< f(x_{\tilde{k}+1}) - \sum_{i=\tilde{k}+1}^{\tilde{k}+r-1} \epsilon_i, \end{aligned} \tag{26}$$

where r is a positive integer and $r \ge 3$.

Because frame $\Phi_{\tilde{k}}$ is the final quasi minimal frame, by Step 6 of Algorithm 5, we know that h_k is a positive constant for $k > \tilde{k}$; that is,

$$h_{\tilde{k}+1} = h_{\tilde{k}+2} = \cdots = h_{\tilde{k}+r-1}, \tag{27}$$

$$\epsilon_{\tilde{k}+1} = \epsilon_{\tilde{k}+2} = \cdots = \epsilon_{\tilde{k}+r-1}. \tag{28}$$

By (25), (26), and (28), we have

$$\begin{aligned} f(x_{\tilde{k}+r}) &< f(x_{\tilde{k}+1}) - \sum_{i=\tilde{k}+1}^{\tilde{k}+r-1} \epsilon_i \\ &< f(x_{\tilde{k}}) - (r-1)\epsilon_{\tilde{k}+1} - \epsilon_{\tilde{k}}. \end{aligned} \tag{29}$$

If we ignore the stopping condition and let $r \to +\infty$, then $f(x_{\tilde{k}+r}) \to -\infty$, which contradicts the condition that $\{f(x_k)\}$ is bounded. The proof of this theorem is complete. □

Theorem 8. *Assume the following conditions are satisfied:*

(A1) f is continuously differentiable.

(A2) $\|v_l^k\| \le M$ for $l = 1, \ldots, q$ and $k = 0, 1, \ldots$, where M is a positive constant and v_l^k is the lth vector in $\mathcal{V}_k$.

Then each cluster point of $\{z_m\}$ is a stationary point of f.

Proof. Let z_∞ be an arbitrary cluster point of $\{z_m\}$ and the subsequence $\{z_m\}_K$ converge to z_∞, where K is an infinite subset of natural numbers. Assume $z_{\breve{m}} \in \{z_m\}_K$, and $z_{\breve{m}} = x_{\breve{k}}$. According to Taylor expansion and (A1), we have

$$\begin{aligned} f\left(z_{\breve{m}} + h_{\breve{m}} v_l^{\breve{m}}\right) &= f(z_{\breve{m}}) + h_{\breve{m}} \nabla f(z_{\breve{m}})^T v_l^{\breve{m}} \\ &\quad + o\left(\left\|h_{\breve{m}} v_l^{\breve{m}}\right\|\right), \end{aligned} \tag{30}$$

for all $v_l^{\breve{m}} \in \mathcal{V}_{\breve{m}}$, where $h_{\breve{m}}$, $\mathcal{V}_{\breve{m}}$ are frame size and positive basis corresponding to the iteration point $z_{\breve{m}}$, respectively, and $v_l^{\breve{m}}$ is lth vector of $\mathcal{V}_{\breve{m}}$. From Definition 4, we have

$$f\left(z_{\breve{m}} + h_{\breve{m}} v_l^{\breve{m}}\right) \ge f(z_{\breve{m}}) - (h_{\breve{m}})^{1+\mu}, \quad \forall v_l^{\breve{m}} \in \mathcal{V}_{\breve{m}}. \tag{31}$$

Combining (30) and (31) with (A2), we obtain

$$\nabla f(z_{\breve{m}})^T v_l^{\breve{m}} \ge -\frac{o(h_{\breve{m}})}{h_{\breve{m}}} M - (h_{\breve{m}})^{\mu}, \quad \forall v_l^{\breve{m}} \in \mathcal{V}_{\breve{m}}. \tag{32}$$

Let $\breve{m} \to +\infty$, then $z_{\breve{m}} \to z_\infty$, $v_l^{\breve{m}} \to v_l^\infty$, $h_{\breve{m}} \to h_\infty$. According to Step 6 of Algorithm 5, we have $h_\infty \to 0$. Combining these with (32) and (A1), we have

$$\nabla f(z_\infty)^T v_l^\infty \ge 0, \quad \forall v_l^\infty \in \mathcal{V}_\infty. \tag{33}$$

Let the numbers of $\mathcal{V}_{\breve{m}}$ be $v_1^{\breve{m}}, \ldots, v_q^{\breve{m}}$, then there exist q nonnegative coefficients η_i $(i = 1, \ldots, q)$ such that

$$-\nabla f(z_\infty) = \sum_{i=1}^{q} \eta_i v_i^\infty. \tag{34}$$

Combining (33) and (34), we have

$$0 \ge -\nabla f(z_\infty)^T \nabla f(z_\infty) = \sum_{i=1}^{q} \eta_i (v_i^\infty)^T \nabla f(z_\infty) \ge 0, \tag{35}$$

which yields $\nabla f(z_\infty) = 0$. The proof of this theorem is complete. □

Remark 9. Although Theorem 8 needs the assumed condition (A1), in practice, we do not solve derivative-free problems that accurately. So we only assure that f is continuously differentiable near the stationary point.

Table 1: The information about benchmark problems set P_1.

Problem	n_p	m_p
(1) Linear (full rank)	9	45
(2) Linear (rank 1)	7	35
(3) Linear (rank 1 with 0 columns & rows)	7	35
(4) Rosenbrock	2	2
(5) Helical valley	3	3
(6) Powell singular	4	4
(7) Freudenstein and Roth	2	2
(8) Bard	3	15
(9) Meyer	3	16
(10) Watson	6	31
(11) Waston	9	31
(12) Waston	12	31
(13) Waston	31	31
(14) Box 3-dimensional	3	10
(15) Jennrich and Sampson	2	10
(16) Brown and Dennis	4	20
(17) Chebyquad	6	6
(18) Chebyquad	7	7
(19) Chebyquad	8	8
(20) Chebyquad	9	9
(21) Brown almost-linear	10	10
(22) Bdqrtic	8	8
(23) Bdqrtic	10	12
(24) Bdqrtic	11	14
(25) Bdqrtic	12	16
(26) Cube	5	5
(27) Cube	6	6
(28) Cube	8	8
(29) Mancino	5	5
(30) Mancino	8	8
(31) Mancino	10	10
(32) Mancino	12	12
(33) Penalty II	4	8
(34) Penalty II	6	12
(35) Penalty II	8	16
(36) Penalty II	10	20
(37) Penalty II	12	24
(38) Variably dimensioned	8	10
(39) Variably dimensioned	9	11
(40) Variably dimensioned	10	12
(41) Variably dimensioned	11	13
(42) Variably dimensioned	12	14
(43) Broyden tridiagonal	6	6
(44) Broyden tridiagonal	7	7
(45) Broyden tridiagonal	8	8
(46) Broyden tridiagonal	9	9
(47) Broyden tridiagonal	10	10
(48) Broyden tridiagonal	11	11
(49) Broyden tridiagonal	12	12

Table 1: Continued.

Problem	n_p	m_p
(50) Broyden banded	4	4
(51) Broyden banded	7	7
(52) Broyden banded	9	9
(53) Broyden banded	10	10
(54) Broyden banded	11	11
(55) Linear (full rank)	100	200
(56) Linear (full rank)	200	400
(57) Linear (rank 1)	100	200
(58) Linear (rank 1)	200	400
(59) Linear (rank 1 with 0 columns & rows)	100	200
(60) Linear(rank 1 with 0 columns & rows)	200	400
(61) Chebyquad	100	100
(62) Chebyquad	200	200
(63) Brown almost-linear	100	100
(64) Brown almost-linear	200	200
(65) Bdqrtic	100	200
(66) Bdqrtic	200	400
(67) Cube	100	100
(68) Cube	200	200
(69) Mancino	100	100
(70) Mancino	200	200
(71) Penalty II	100	200
(72) Penalty II	200	400
(73) Variably dimensioned	100	102
(74) Variably dimensioned	200	202
(75) Broyden tridiagonal	100	100
(76) Broyden tridiagonal	200	200
(77) Broyden banded	100	100
(78) Broyden banded	200	200

4. Numerical Experiments

In this section, we discuss numerical test results for Algorithm 5. Our tests are performed on a PC with Intel Core Duo CPU (I5-3470@3.20 GHz, 3.60 GHz) and 8 GB RAM, using MATLAB 7.12.0.

To compare our algorithm to Max-PRP, we choose to work with the performance profiles [30] and data profiles [31] for derivative-free optimization. The performance profile is the following fraction:

$$\rho_s(\alpha) = \frac{1}{|P|}\left|\left\{p \in P: \frac{t_{p,s}}{\min\left\{t_{p,s}: s \in S\right\}} \le \alpha\right\}\right|, \qquad (36)$$

where P is the set of benchmark problems, S is the set of optimization solvers, $t_{p,s}$ is the number of function evaluations required to satisfy the convergence test for problem $p \in P$ on solver $s \in S$.

The data profile is defined that

$$d_s(\kappa) = \frac{1}{|P|}\left|\left\{p \in P: \frac{t_{p,s}}{n_p + 1} \le \kappa\right\}\right|, \qquad (37)$$

where n_p is the number of variables in $p \in P$.

Table 2: The information about benchmark problems set P_2.

Problem	n_p	m_p
(1) Raydan 1	10	10
(2) Raydan 1	20	20
(3) Raydan 1	100	100
(4) Raydan 2	10	10
(5) Raydan 2	20	20
(6) Raydan 2	100	100
(7) Diagonal 1	10	10
(8) Diagonal 1	20	20
(9) Diagonal 1	100	100
(10) Diagonal 2	10	10
(11) Diagonal 2	20	20
(12) Diagonal 2	100	100
(13) Diagonal 3	10	10
(14) Diagonal 3	20	20
(15) Diagonal 3	100	100
(16) Diagonal 4	10	5
(17) Diagonal 4	20	10
(18) Diagonal 4	100	50
(19) Diagonal 5	10	10
(20) Diagonal 5	20	20
(21) Diagonal 5	100	100
(22) Diagonal 6	10	10
(23) Diagonal 6	20	20
(24) Diagonal 6	100	100
(25) Hager	10	10
(26) Hager	20	20
(27) Hager	100	100
(28) Extended TET	10	5
(29) Extended TET	20	10
(30) Extended TET	100	50
(31) Extended Maratos	10	5
(32) Extended Maratos	20	10
(33) Extended Maratos	100	50
(34) Extended Cliff	10	5
(35) Extended Cliff	20	10
(36) Extended Cliff	100	50
(37) ARWHEAD	10	9
(38) ARWHEAD	20	19
(39) ARWHEAD	100	99
(40) ENGVAL1	10	9
(41) ENGVAL1	20	19
(42) ENGVAL1	100	99
(43) GENHUMPS	10	9
(44) GENHUMPS	20	19
(45) GENHUMPS	100	99
(46) MCCORMCK	10	9
(47) MCCORMCK	20	19
(48) MCCORMCK	100	99
(49) COSINE	10	9
(50) COSINE	20	19
(51) COSINE	100	99
(52) SINE	10	9
(53) SINE	20	19
(54) SINE	100	99

Table 2: Continued.

Problem	n_p	m_p
(55) HIMMELBG	10	5
(56) HIMMELBG	20	10
(57) HIMMELBG	100	50
(58) HIMMELH	10	5
(59) HIMMELH	20	10
(60) HIMMELH	100	50

We use the following convergence condition:

$$f(x) \leq f_L + \tau (f(x_0) - f_L), \quad (38)$$

where x_0 is the initial point for the test problem, $\tau > 0$ is tolerance, f_L is the best function value achieved by any solvers within μ_f function evaluations, and μ_f is a positive integer.

The benchmark problems set P in our experiments is proposed in [32, 33] and CUTEr test problem set [34]. The problems set P includes 78 nonlinear least squares problems P_1 and 60 normal nonlinear programming problems P_2. Tables 1 and 2 show some information about test problems, where n_p is the number of variables and m_p is the number of components. The problems of Table 1 are defined by

$$f(x) = \sum_{k=1}^{m_p} (f_k(x))^2. \quad (39)$$

The problems of Table 2 are defined by

$$f(x) = \sum_{k=1}^{m_p} f_k(x). \quad (40)$$

In all problems, we have

$$2 \leq n_p \leq 200, \quad p = 1, \ldots, 138. \quad (41)$$

In addition, we define the maximum computational budget as 150 simplex gradients, where the computational budget of a simplex gradients is equal to $n_p + 1$ function evaluations, so $\mu_f = 30150$.

The parameters of our numerical experiments are listed as follows: $h_0 = 1$, $\mathscr{V}_0 = \mathscr{V}_{\min}$, $v_i^0 = e_i$ $(i = 1, \ldots, n_p)$, $\lambda = 4$, $\mu = 0.5$, $\delta_0 = 1$, and δ_k takes value 2 if the previous iteration was successful, or 1 otherwise.

In the RBF interpolation model of (8), we set $p_1, \ldots, p_r$ as linear polynomial tails and $\phi(\|x - y_i\|) = \|x - y_i\|^3$. In addition, we set $p_{\max} = 3n$ as the maximum number of points considered in the interpolate data points $Y = \{y_1, y_2, \ldots, y_{n_q}\}$. All the previously evaluated points are used to compute the RBF interpolation model when its number is lower than $p_{\max}$. Similar to [20], whenever there are more previously evaluated points than $p_{\max}$ for building the RBF interpolation model, 80% of the desired points are selected as the ones nearest to the current iterate and the last 20% are chosen as the ones further away from the current iterate. This strategy is adopted in order to preserve the geometry and diversify the information used in the RBF interpolation model.

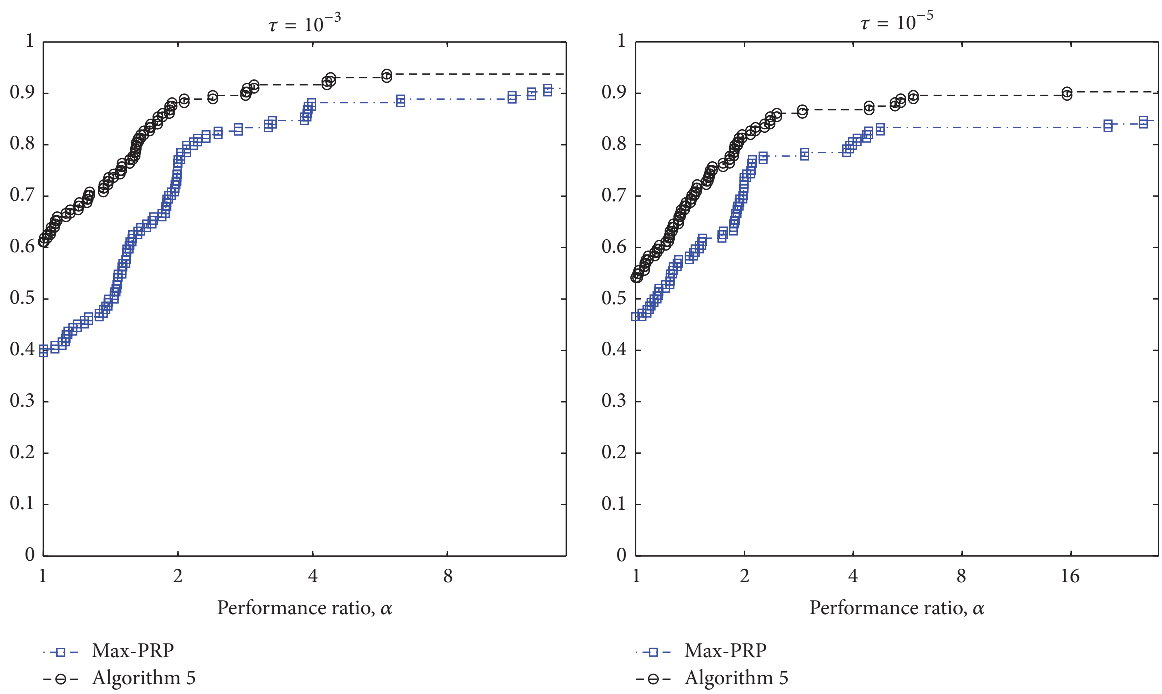

FIGURE 1: Performance profiles $\rho_s(\alpha)$ for problems P.

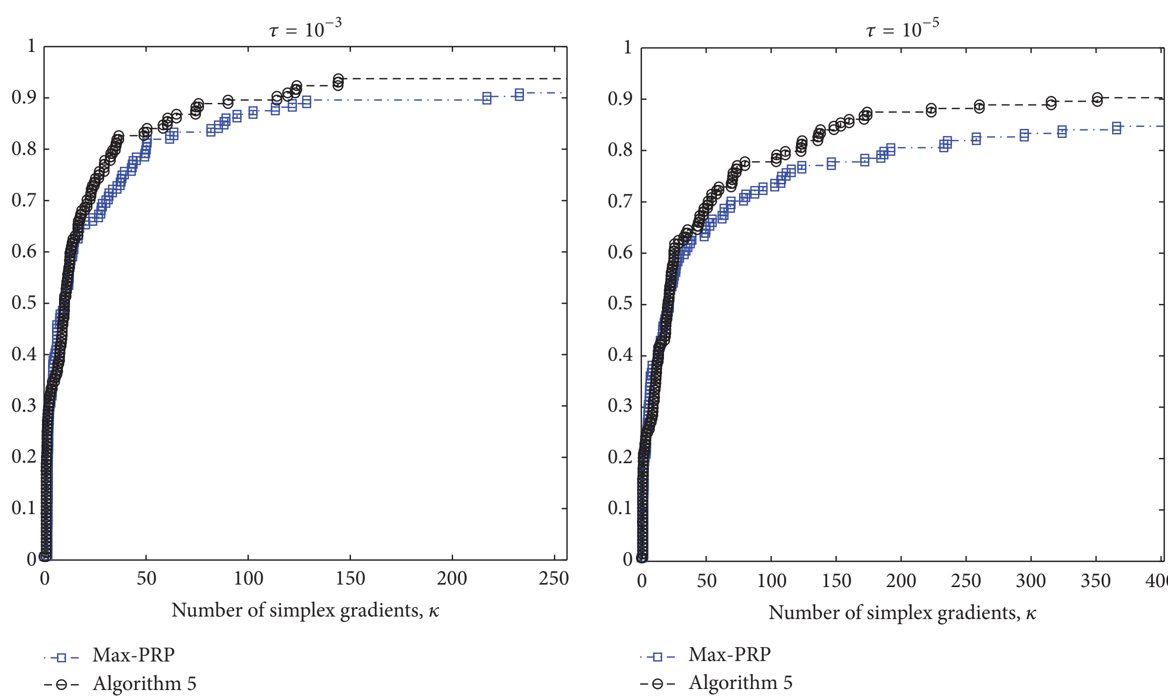

FIGURE 2: Data profiles $d_s(\kappa)$ for problems P.

In Figure 1, we show the performance profiles related to Algorithm 5 and Max-PRP. As we can see, Algorithm 5 outperforms Max-PRP when $\tau = 10^{-3}$, and the difference is significantly large as the performance ratio α decreases. In addition, Algorithm 5 guarantees better results than Max-PRP when $\tau = 10^{-5}$. For example, Algorithm 5 can solve about 90% test problems, while Max-PRP only solves no more than 85%, if performance ratio $\alpha = 16$.

The data profiles of Algorithm 5 and Max-PRP are reported in Figure 2. When the number of simplex gradients κ is larger than 40, Algorithm 5 performs better than Max-PRP as it solves a higher percentage of problems. For example, with a budget of 400 simplex gradients and $\tau = 10^{-5}$, Algorithm 5 solves almost 90% of the problems, while Max-PRP solves roughly 85% of the problems.

5. Conclusion

The computational results which are presented in this paper show that Algorithm 5 appears quite competitive. The performance profiles and the data profiles of numerical results indicate that Algorithm 5 often reduces the number of function evaluations which is required to reach stationary point and is superior to Max-PRP.

Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (11071117 and 11274109), the Natural Science Foundation of Jiangsu Province (BK20141409), and the Natural Science Foundation of Huzhou University (KX21072).

References

[1] A. J. Booker, P. D. Frank, J. E. Dennis Jr. et al., "Managing surrogate objectives to optimize a helicopter rotor design further experiments," AIAA MDO, 98-4717, 1998.

[2] A. J. Booker, J. E. Dennis Jr., P. D. Frank et al., "Optimization using surrogate objectives on a helicopter test example," in *Computational Methods for Optimal Design and Control*, pp. 49–58, Birkhäuser, Boston, Mass, USA, 1998.

[3] A. L. Marsden, M. Wang, J. E. Dennis Jr., and P. Moin, "Optimal aeroacoustic shape design using the surrogate management framework," *Optimization and Engineering*, vol. 5, no. 2, pp. 235–262, 2004.

[4] P. Mugunthan, C. A. Shoemaker, and R. G. Regis, "Comparison of function approximation, heuristic, and derivative-based methods for automatic calibration of computationally expensive groundwater bioremediation models," *Water Resources Research*, vol. 41, no. 11, pp. 1–17, 2005.

[5] R. Oeuvray and M. Bierlaire, "A new derivative-free algorithm for the medical image registration problem," *International Journal of Modelling and Simulation*, vol. 27, no. 2, pp. 115–124, 2007.

[6] M. J. D. Powell, "A direct search optimization method that models the objective and constraint functions by linear interpolation," in *Advances in Optimization and Numerical Analysis*, pp. 51–67, Springer Netherlands, 1994.

[7] D. Winfield, "Function minimization by interpolation in a data table," *IMA Journal of Applied Mathematics*, vol. 12, no. 3, pp. 339–347, 1973.

[8] A. R. Conn, K. Scheinberg, and P. L. Toint, "Recent progress in unconstrained nonlinear optimization without derivatives," *Mathematical Programming*, vol. 79, no. 1–3, pp. 397–414, 1997.

[9] A. R. Conn, K. Scheinberg, and L. N. Vicente, "Global convergence of general derivative-free trust-region algorithms to first- and second-order critical points," *SIAM Journal on Optimization*, vol. 20, no. 1, pp. 387–415, 2009.

[10] M. J. D. Powell, "Least Frobenius norm updating of quadratic models that satisfy interpolation conditions," *Mathematical Programming*, vol. 100, no. 1, pp. 183–215, 2004.

[11] K. Scheinberg and P. L. Toint, "Self-correcting geometry in model-based algorithms for derivative-free unconstrained optimization," *SIAM Journal on Optimization*, vol. 20, no. 6, pp. 3512–3532, 2010.

[12] R. Hooke and T. A. Jeeves, ""Direct search" solution of numerical and statistical problems," *Journal of the ACM*, vol. 8, no. 2, pp. 212–229, 1961.

[13] V. Torczon, "On the convergence of the multidirectional search algorithm," *SIAM Journal on Optimization*, vol. 1, no. 1, pp. 123–145, 1991.

[14] V. Torczon, "On the convergence of pattern search algorithms," *SIAM Journal on Optimization*, vol. 7, no. 1, pp. 1–25, 1997.

[15] C. Audet and J. Dennis, "Analysis of generalized pattern searches," *SIAM Journal on Optimization*, vol. 13, no. 3, pp. 889–903, 2003.

[16] I. D. Coope and C. J. Price, "A direct search frame-based conjugate gradients method," *Journal of Computational Mathematics*, vol. 22, no. 4, pp. 489–500, 2004.

[17] B. T. Polyak, "The conjugate gradient method in extremal problems," *USSR Computational Mathematics and Mathematical Physics*, vol. 9, no. 4, pp. 94–112, 1969.

[18] E. Polak and G. Ribiere, "Note sur la convergence de méthodes de directions conjuguées," *Revue Francaise Informatique Recherche Operationnelle*, vol. 3, no. 1, pp. 35–43, 1969.

[19] T. G. Kolda, R. M. Lewis, and V. Torczon, "Optimization by direct search: new perspectives on some classical and modern methods," *SIAM Review*, vol. 45, no. 3, pp. 385–482, 2003.

[20] A. L. Custódio, H. Rocha, and L. N. Vicente, "Incorporating minimum Frobenius norm models in direct search," *Computational Optimization and Applications*, vol. 46, no. 2, pp. 265–278, 2010.

[21] A. R. Conn and S. Le Digabel, "Use of quadratic models with mesh-adaptive direct search for constrained black box optimization," *Optimization Methods and Software*, vol. 28, no. 1, pp. 139–158, 2013.

[22] S. M. Wild, R. G. Regis, and C. A. Shoemaker, "ORBIT: optimization by radial basis function interpolation in trust-regions," *SIAM Journal on Scientific Computing*, vol. 30, no. 6, pp. 3197–3219, 2008.

[23] S. M. Wild and C. Shoemaker, "Global convergence of radial basis function trust-region algorithms for derivative-free optimization," *SIAM Review*, vol. 55, no. 2, pp. 349–371, 2013.

[24] A. R. Conn, K. Scheinberg, and L. N. Vicente, *Introduction to Derivative-Free Optimization*, vol. 8 of *MPS/SIAM Series on Optimization*, SIAM, Philadelphia, Pa, USA, 2008.

[25] I. D. Coope and C. J. Price, "Frame based methods for unconstrained optimization," *Journal of Optimization Theory and Applications*, vol. 107, no. 2, pp. 261–274, 2000.

[26] I. D. Coope and C. J. Price, "On the convergence of grid-based methods for unconstrained optimization," *SIAM Journal on Optimization*, vol. 11, no. 4, pp. 859–869, 2001.

[27] M. D. Buhmann, *Radial Basis Functions: Theory and Implementations*, vol. 12 of *Cambridge Monographs on Applied and Computational Mathematics*, Cambridge University Press, Cambridge, UK, 2003.

[28] L. Grippo and F. Rinaldi, "A class of derivative-free nonmonotone optimization algorithms employing coordinate rotations and gradient approximations," *Computational Optimization and Applications. An International Journal*, vol. 60, no. 1, pp. 1–33, 2015.

[29] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, John Wiley & Sons, New York, NY, USA, 3rd edition, 2006.

[30] E. D. Dolan and J. J. Moré, "Benchmarking optimization software with performance profiles," *Mathematical Programming*, vol. 91, no. 2, pp. 201–213, 2002.

[31] J. J. Moré and S. M. Wild, "Benchmarking derivative-free optimization algorithms," *SIAM Journal on Optimization*, vol. 20, no. 1, pp. 172–191, 2009.

[32] N. Andrei, "An unconstrained optimization test functions collection," *Advanced Modeling and Optimization*, vol. 10, no. 1, pp. 147–161, 2008.

[33] J. J. Moré, B. S. Garbow, and K. E. Hillstrom, "Testing unconstrained optimization software," *Association for Computing Machinery. Transactions on Mathematical Software*, vol. 7, no. 1, pp. 17–41, 1981.

[34] N. I. M. Gould, D. Orban, and P. L. Toint, "CUTEr and SifDec: a constrained and unconstrained testing environment revisited," *ACM Transactions on Mathematical Software*, vol. 29, no. 4, pp. 373–394, 2003.