

$$\begin{aligned}
\textbf{QP2}(x^k, B^k) \quad & \min \frac{1}{2} (d^k)^T B^k d^k + \sum_{i=1}^m \nabla f_i(x^k)^T d^k \\
\text{s.t.} \quad & \nabla f_i(x^k)^T d^k + f_i(x^k) - f_i(\hat{x}) \leq 0, i \in M = \{1, \dots, m\} \\
& \nabla g_i(x^k)^T d + g_i(x^k) \leq 0, i \in I = \{1, \dots, l\}
\end{aligned}$$

### 3.2 A new MOSQP algorithm

**Algorithm 2 (MOSQP algorithm):**

**Iteration point set initialization:** Initialize the iteration point set  $X_0 = \{x_1, x_2, \dots, x_N\}$ ,  $x_i \in R^n$ , They are all feasible points.

**Iterative point set feasible spread:** Kth spread for  $X_0 = \{x_1, x_2, \dots, x_N\}$ ,  $x_i \in R^n$ .

Step1 Let  $j = 0$ ,  $B = I_n$ ,  $0 < \sigma < 1$ ,  $i = 1$ ,  $M > 1$ ,  $0 < A < 1$ ,  $X = X_0$ ,  $T = \emptyset$ ,  $c > 0$ ;

Step2 Let  $x^j \in X_j$ , Solving the sub-problem  $\textbf{QP}_i(x^j, I)$ , and let  $d^j = d(x^j, I_n)$ , if  $d^j = 0$ ,  $T = T \cup \{x^j\}$ , Otherwise turn Step3;

Step3 Let  $\alpha^j = 1$ , to Step4;

Step4 If  $f_i(x^j + \alpha^j d^j) \leq f_i(x^j) + \sigma \alpha^j [\nabla_x f(x^j)]^T d^j$ , and  $g_i(x^j + \alpha^j d^j) \leq 0, i \in I$ , to

Step6 Otherwise to Step5;

Step5 Let  $\alpha^j = A \alpha^j$ , to Step6;

Step6 Let  $T = T \cup \{x^j + \alpha^j d^j\}$ ,  $i = i + 1$ , if  $i = m + 1$  to Step7, Otherwise, to Step2;

Step7 Let  $j = j + 1$ ,  $X_j = T$ ,  $i = 0$ ,  $X = X \cup X_j$ , if  $j = K$ , to step8, Otherwise, let  $T = \emptyset$ , to Step2;

Step8 Output X for the point where the crowding function value  $C(x_i)$  is greater than  $c$ .

**Note 3.2.1** All the points obtained by diffusion steps are in the feasible domain, ensuring that the iteration point enters the Pareto optimal solution stage. When the initial iteration point is taken as the reference point, the feasible domain of the problem (P2) is not empty. Secondly, Step8 is designed to remove repeated points during the diffusion process and close points, and  $c$  takes different values according to the order of magnitude corresponding to the specific question.

**Pareto Optimal solution of the multi-objective optimization problem:**

Step1  $\Leftrightarrow S = \emptyset$ ,  $\frac{1}{b} < k < 1$ ,  $b > 0$ ,  $\varepsilon > 0$ ,  $j = 0$ ,  $B = I_n$ ,  $0 < \sigma < 1$ ,  $j = 0$ ,  $i = 1$ ,

$M > 1$ ,  $0 < A < 1$ ;

Step2 If the subproblem  $\mathbf{QP2}(x^j, B^j)$  is feasible, Solving the sub-problem  $\mathbf{QP2}(x^j, B^j)$ , let  $d^j = d(x^j, B^j)$ ,  $d^j = 0$ ,  $S = S \cup x^j$ , Otherwise to Step3;

If the problem  $\mathbf{QP2}(x^j, B^j)$  is not feasible, let  $d^j = -\nabla_x P_{\varepsilon^j}^k(x^j, \pi^j)$  to Step8;

Step3 let  $\varepsilon^j = A^k \varepsilon^j$ , st,  $\varepsilon^j \leq \min_{i \in I_- \neq \emptyset} (-g_i(x^j), f_i(\hat{x}) - f_i(x^j))$ ,

if  $(\nabla_x \hat{P}_{\varepsilon^j}^k(x^j, \pi^j))^T d^j \leq -\frac{1}{2}(d^j)^T B^j d^j$ , to Step5, Otherwise to Step4;

Step4  $\pi^j = M\pi^j$ , to Step3;

Step5  $\alpha^j = 1$ , to Step6;

Step6 if  $\hat{P}_{\varepsilon^j}^k(x^j + \alpha^j d^j, \pi^j) \leq \hat{P}_{\varepsilon^j}^k(x^j, \pi) + \sigma \alpha^j [\nabla_x \hat{P}_{\varepsilon^j}^k(x^j, \pi^j)]^T d^j$ , to Step13; Otherwise to Step7;

Step7  $\alpha^j = A\alpha^j$ , to Step6;

Step8 if  $\sum_{g_i(x^j) > 0} d_2(x^j) \nabla_x g_i(x^j) + \sum_{f_i(x^j) > f_i(\hat{x}^j)} d_3(x^j) \nabla_x f_i(x^j) \neq 0$ , to step10, Otherwise to

Step9;

Step9 let  $\varepsilon^j = A\varepsilon^j$ , to Step8;

Step10  $(\nabla_x \hat{P}_{\varepsilon^j}^k(x^j, \pi^j))^T d^j \leq -\max \{g_i(x^j), i \in I_+; f_i(x^j) - f_i(\hat{x}), i \in M_+\} \leq 0$  to Step5, Otherwise to Step12;

Step11 let  $\pi^j = M\pi^j$ , to Step11;

Step12 If the problem is feasible, Correction of  $B^{j+1}$  by using the BFGS formula, Otherwise let  $B^{j+1} = B^j$ , to Step14;

Step13 let  $x^{j+1} = x^j + \alpha^j d^j$ ,  $\varepsilon^{j+1} = \varepsilon^j$ ,  $\pi^{j+1} = \pi^j$ ,  $j = j + 1$ ,  $S = S \cup \{x^j + \alpha^j d^j\}$ , to Step14;

Step14 let  $j = 0$ ,  $i = i + 1$ , if  $i = M + 1$ , to Step15;

Step15  $S = \text{Non-dominated}(S)$ , Output S for the point where the crowding function value  $C(x_i)$  is greater than  $c$ .

**Note3.2.2** The  $d_2(x^j)$  and  $d_3(x^j)$  in the algorithm are respectively:

$$d_2(x^j) = k(g_i(x^j) + \varepsilon b^{-1})^{k-1}, \quad d_3(x^j) = k(f_i(x^j) - f_i(\hat{x}) + \varepsilon b^{-1})^{k-1}$$

Secondly, the set of the 50 Pareto optimal solutions with the largest crowding degree function value in S for the original multi-objective optimization problem is designed to remove the Pareto front repeat solutions obtained by the MOSQP algorithm.

**Theorem3.2.1** If  $x^*$  is the KKT point of the problem (P2), it is the Pareto critical point of the original multi-objective optimization problem.

Proof: If  $x^*$  is the KKT point of the problem (P2), then:

$$\begin{aligned}
& \sum_{i=1}^m \nabla f_i(x^*) + \sum_{i=1}^m \mu_i \nabla f_i(x^*) + \sum_{i=1}^l \lambda_i^* \nabla g_i(x^*) = 0 \\
& g_i(x^*) \leq 0, \lambda_i^* \geq 0, \forall i \in I \\
& \lambda_i^* g_i(x^*) = 0 \\
& f_i(x^*) - f_i(\hat{x}) \leq 0, \mu_i \geq 0, i \in M = \{1, \dots, m\} \\
& \mu_i (f_i(x^*) - f_i(\hat{x})) = 0
\end{aligned}$$

So  $x^*$  satisfied:

$$\begin{aligned}
& \sum_{i=1}^m (1 + \mu_i) \nabla f_i(x^*) + \sum_{i=1}^l \lambda_i^* \nabla g_i(x^*) = 0 \\
& g_i(x^*) \leq 0, \lambda_i^* \geq 0, \forall i \in I \\
& \lambda_i^* g_i(x^*) = 0
\end{aligned}$$

So  $x^*$  is the KKT point of

$$\begin{aligned}
& \min \sum_{i=1}^m (1 + \mu_i) f_i(x) \\
& s.t. g_i(x) \leq 0, i \in I
\end{aligned}$$

And because  $1 + \mu_i > 0$ ,  $x^*$  is the Pareto critical point of the original multi-objective constraint optimization problem (CMOP).

Theorem 3.2.1 also shows that the proposed MOSQP algorithm will converge to the Pareto approximation front of the original multi-objective constraint optimization problem (CMOP).

## 4 Numerical experiments

### 4.1 Algorithm performance measure

First, the measure of the algorithm performance is defined as follows:

**Purity metric:** Purity metric is used to compare the number of non-dominated solutions obtained by different algorithms.

$$\bar{t}_{p,s} = \frac{|F_p|}{|F_{p,s} \cap F_p|}$$

From the definition, the smaller  $\bar{t}_{p,s}$ , the better the quality of the solution. Obviously,  $\bar{t}_{p,s} = \infty$  means that the algorithm cannot generate any non-dominant points in the reference Pareto front of the corresponding problem.