Notes on NEWUOA

Zaikun Zhang *

June 14, 2021 12:23pm

We use $\mathcal{X}_k$ to denote the set of interpolation points at iteration k. In addition, we define $\overline{\mathcal{X}_k}$ to be a point in $\mathcal{X}_k$ such that

$$f(\overline{\mathcal{X}_k}) = \min\{f(x) : x \in \mathcal{X}_k\}. \tag{0.1}$$?eq:xopt?

If multiple points attain the minimum, then we take the earliest one visited by the algorithm. Meanwhile, we define

$$\delta(\mathcal{X}_k) = \max\{\|x - \overline{\mathcal{X}_k}\| : x \in \mathcal{X}_k\}. \tag{0.2}$$?eq:xdist?

Let $\mathcal{Q}$ be the linear space of all the n-variable polynomials with degree at most two. The elements in $\mathcal{Q}$ will be referred to as quadratics, even though they may be linear or constant. Given $\mathcal{X}_k$, we denote

$$\mathcal{Q}_k = \{Q \in \mathcal{Q} : Q(x) = f(x) \text{ for } x \in \mathcal{X}_k\}, \tag{0.3}$$?eq:quak?

which is the set of quadratics that interpolate f at $\mathcal{X}_k$. Note that $\mathcal{Q}_k$ is an affine subset of $\mathcal{Q}$.

$$\|Q\|_{\mathrm{S}} = \|\nabla^2 Q\|_{\mathrm{F}} \tag{0.4}$$?eq:snorm?

$$\mathrm{Proj}_{\mathcal{C}}^{\mathrm{P}} \quad \mathrm{P}_{\mathcal{C}}^{\mathrm{S}}(\Phi) = \operatorname{argmin}\{\|Q - \Phi\|_{\mathrm{S}} : Q \in \mathcal{C}\}$$

*Hong Kong Polytechnic University, zaikun.zhang@polyu.edu.hk

Algorithm 0.1 OPTimization based on Interpolation Models (OPTIM)

?⟨alg:optim⟩? Input $\Delta_0 \in (0, +\infty)$, $\tau > 0$, $m \in \{1, 2, \ldots, (n+1)(n+2)/2\}$, and $\mathcal{X}_0 \subset \mathbb{R}^n$ with $|\mathcal{X}_0| = m$ and $\kappa(\mathcal{X}_0) \le \kappa_0$. Set $k = 0$.

1. **Model construction**. Pick $Q_k \in \mathcal{Q}_k$.
2. **Trust-region step evaluation**. Calculate

$$x_k^{\mathrm{a}} \approx \operatorname{argmin}\{Q_k(x) : \|x - \overline{\mathcal{X}_k}\| \le \Delta_k\}. \tag{0.5}$$?eq:xadd?

   If $\|x_k^{\mathrm{a}} - \overline{\mathcal{X}_k}\| < \alpha\Delta_k$, then set $\Delta_{k+1} = \theta\Delta_k$; otherwise, update Δ_k to Δ_{k+1} according to $r_k = [f(\overline{\mathcal{X}_k}) - f(x_k^{\mathrm{a}})]/[Q_k(\overline{\mathcal{X}_k}) - Q_k(x_k^{\mathrm{a}})]$.
3. **Interpolation set update**. Let

$$x_k^{\mathrm{d}} \approx \operatorname{argmin}\{\kappa(\mathcal{X}_k \cup x_k^{\mathrm{a}} \setminus x) : x \in \mathcal{X}_k\}. \tag{0.6}$$?eq:xdrop?

   If $\|x_k^{\mathrm{a}} - \overline{\mathcal{X}_k}\| \ge \alpha\Delta_k$ and either $r_k > \eta_0$ or $\kappa(\mathcal{X}_k \cup x_k^{\mathrm{a}} \setminus x_k^{\mathrm{d}}) \le \kappa_0$, then set $\mathcal{X}_{k+1} = \mathcal{X}_k \cup x_k^{\mathrm{a}} \setminus x_k^{\mathrm{d}}$.
4. **Geometry improvement**. If $\|x_k^{\mathrm{a}} - \overline{\mathcal{X}_k}\| < \alpha\Delta_k$, or if $r_k \le \eta_0$ and $\kappa(\mathcal{X}_k \cup x_k^{\mathrm{a}} \setminus x_k^{\mathrm{d}}) > \kappa_0$, then set $\mathcal{X}_{k+1} = \mathcal{X}_k \cup y_k^{\mathrm{a}} \setminus y_k^{\mathrm{d}}$ with

$$y_k^{\mathrm{d}} = \operatorname{argmax}\{\|y - \overline{\mathcal{X}_k}\| : y \in \mathcal{X}_k\}, \tag{0.7}$$?eq:ydrop?

$$y_k^{\mathrm{a}} \approx \operatorname{argmin}\{\kappa(\mathcal{X}_k \cup y \setminus y_k^{\mathrm{d}}) : \|y - \overline{\mathcal{X}_k}\| \le \Delta_k\}. \tag{0.8}$$?eq:yadd?

Algorithm 0.2 NEWUOA

?⟨alg:newuoa⟩?

Input $\Delta_0 \in (0, +\infty)$, $\rho_{\text{end}} > 0$, $m \in \{n{+}2, n{+}3, \ldots, (n{+}1)(n{+}2)/2\}$, and $\mathcal{X}_0 \subset \mathbb{R}^n$ with $|\mathcal{X}_0| = m$ and $\kappa(\mathcal{X}_0) \le \kappa_0$. Define $Q_0 = \mathrm{P}^{\mathrm{s}}_{\mathcal{Q}_0}(0)$. Set $k = 0$.

1. **Trust-region step evaluation**. Set

$$x_k^{\mathrm{a}} \approx \operatorname{argmin}\{Q_k(x) : \|x - \overline{\mathcal{X}}_k\| \le \Delta_k\}. \tag{0.9}$$?eq:xaddn?

$$\mathsf{S} = \mathbb{1}(\|x_k^{\mathrm{a}} - \overline{\mathcal{X}}_k\| < \rho_k/2), \tag{0.10}$$?eq:shortd?

$$\mathtt{R} = \mathbb{1}(\mathsf{S} = 1 \text{ and the errors in recent models are small}). \tag{0.11}$$?eq:redrho?

If $\mathtt{R} = 1$, then let $\mathcal{X}_{k+1} = \mathcal{X}_k$, $Q_{k+1} = Q_k$, and go to step 4. If $\mathtt{R} = 0$ and $\mathsf{S} = 1$, then set $\Delta_{k+1} = \max\{\Delta_k/10,\, \rho_k\}$. If $\mathsf{S} = 0$, then evaluate $r_k = [f(\overline{\mathcal{X}}_k) - f(x_k^{\mathrm{a}})]/[Q_k(\overline{\mathcal{X}}_k) - Q_k(x_k^{\mathrm{a}})]$ and update Δ_k to Δ_{k+1} according to r_k.

2. **Interpolation set update**. Let

$$x_k^{\mathrm{d}} \approx \operatorname{argmin}\{\kappa(\mathcal{X}_k \cup x_k^{\mathrm{a}} \setminus x) : x \in \mathcal{X}_k\}. \tag{0.12}$$?eq:xdropn?

If $\mathsf{S} = 0$ and either $r_k > 0$ or $\kappa(\mathcal{X}_k \cup x_k^{\mathrm{a}} \setminus x_k^{\mathrm{d}}) \le \kappa_0$, then set

$$\mathcal{X}_k^+ = \mathcal{X}_k \cup x_k^{\mathrm{a}} \setminus x_k^{\mathrm{d}}, \quad Q_k^+ = \mathrm{P}^{\mathrm{s}}_{\mathcal{Q}_k^+}(Q_k); \tag{0.13}$$?eq:updateq1?

otherwise, $\mathcal{X}_k^+ = \mathcal{X}_k$ and $Q_k^+ = Q_k$.

3. **Geometry improvement**. Let $\Delta_k^+ = \max\{\min\{\delta(\mathcal{X}_k^+)/10,\, \Delta_{k+1}/2\},\, \rho_k\}$, and

$$y_k^{\mathrm{d}} = \operatorname{argmax}\{\|y - \overline{\mathcal{X}_k^+}\| : y \in \mathcal{X}_k^+\}, \tag{0.14}$$?eq:ydropn?

$$y_k^{\mathrm{a}} \approx \operatorname{argmin}\{\kappa(\mathcal{X}_k^+ \cup y \setminus y_k^{\mathrm{d}}) : \|y - \overline{\mathcal{X}_k^+}\| \le \Delta_k^+\}. \tag{0.15}$$?eq:yaddn?

If $\delta(\mathcal{X}_k^+) \ge 2\Delta_{k+1}$ and either $\mathsf{S} = 1$ or $r_k < 1/10$, then set

$$\mathcal{X}_{k+1} = \mathcal{X}_k^+ \cup y_k^{\mathrm{a}} \setminus y_k^{\mathrm{d}}, \quad Q_{k+1} = \mathrm{P}^{\mathrm{s}}_{\mathcal{Q}_{k+1}}(Q_k^+); \tag{0.16}$$?eq:updateq2?

otherwise, $\mathcal{X}_{k+1} = \mathcal{X}_k^+$ and $Q_{k+1} = Q_k^+$. If $\delta(\mathcal{X}_k^+) < 2\Delta_{k+1}$, $\max\{\Delta_{k+1},\, \|x_k^{\mathrm{a}} - \overline{\mathcal{X}}_k\|\} \le \rho_k$, and either $\mathsf{S} = 1$ or $r_k \le 0$, then set $\mathtt{R}$ to 1.

4. **Resolution enhancement**. If $\mathtt{R} = 0$, then set $\rho_{k+1} = \rho_k$. If $\mathtt{R} = 1$ and $\rho_k > \rho_{\text{end}}$, then reduce ρ_k by about a factor of 10 to obtain ρ_{k+1} and set $\Delta_{k+1} = \max\{\rho_k/2,\, \rho_{k+1}\}$. If $\mathtt{R} = 1$ and $\rho_k \le \rho_{\text{end}}$, then exit.