Notes on NEWUOA

Zaikun Zhang *

June 13, 2021 12:37pm

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.

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

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

Algorithm 0.1 OPTimization based on Interpolation Models (OPTIM)

?⟨alg:optim⟩? Input $\Delta_0 \in (0, +\infty)$, $\tau > 0$, $m \in \{n+2, n+3, \ldots, (n+1)(n+2)/2\}$, and $\mathcal{X}_0 \subset \mathbb{R}^n$ with $|\mathcal{X}_0| = m$ and $\kappa(\mathcal{X}_0) \leq \kappa_0$. Set $k = 0$.

1. **Model construction**. Pick $Q_k \in \{Q : Q(x) = f(x) \text{ for all } x \in \mathcal{X}_k\}$.
2. **Trust-region step evaluation**. Calculate

$$x_k^{\mathrm{g}} \approx \operatorname{argmin}\{Q_k(x) : \|x - \overline{\mathcal{X}_k}\| \leq \Delta_k\}. \tag{0.3}$$?eq:xget?

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

$$x_k^{\mathrm{d}} \approx \operatorname{argmin}\{\kappa(\mathcal{X}_k \cup x_k^{\mathrm{g}} \setminus x) : x \in \mathcal{X}_k\}, \tag{0.4}$$?eq:xdrop?

   and set $\mathcal{X}_{k+1} = \mathcal{X}_k \cup x_k^{\mathrm{g}} \setminus x_k^{\mathrm{d}}$ if $r_k > 0$ or $\kappa(\mathcal{X}_k \cup x_k^{\mathrm{g}} \setminus x_k^{\mathrm{d}}) \leq \kappa_0$.
4. **Geometry improvement**. If $\|x_k^{\mathrm{g}} - \overline{\mathcal{X}_k}\| < \alpha\Delta_k$, or $\|x_k^{\mathrm{g}} - \overline{\mathcal{X}_k}\| > \alpha\Delta_k$ but $r_k \leq 0$ and $\kappa(\mathcal{X}_k \cup x_k^{\mathrm{g}} \setminus x_k^{\mathrm{d}}) > \kappa_0$, then calculate

$$y_k^{\mathrm{d}} = \operatorname{argmax}\{\|y - \overline{\mathcal{X}_k}\| : y \in \mathcal{X}_k\}, \tag{0.5}$$?eq:ydrop?

$$y_k^{\mathrm{g}} \approx \operatorname{argmin}\{\kappa(\mathcal{X}_k \cup y \setminus y_k^{\mathrm{d}}) : \|y - \overline{\mathcal{X}_k}\| \leq \Delta_k\}, \tag{0.6}$$?eq:yget?

   and set $\mathcal{X}_{k+1} = \mathcal{X}_k \cup y_k^{\mathrm{g}} \setminus y_k^{\mathrm{d}}$.

Algorithm 0.2 NEWUOA

?⟨alg:newuoa⟩? Input $\Delta_0 \in (0, +\infty)$, $\tau > 0$, $m \in \{n+2, n+3, \ldots, (n+1)(n+2)/2\}$, and $\mathcal{X}_0 \subset \mathbb{R}^n$ with $|\mathcal{X}_0| = m$ and $\kappa(\mathcal{X}_0) \le \kappa_0$. Define $Q_0 = \operatorname{argmin}\{\|\nabla Q\|_{\mathrm{F}} : Q \in \mathcal{Q} \text{ and } Q(x) = f(x) \text{ for all } x \in \mathcal{X}_0\}$. Set $k = 0$.

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

$$x_k^{\mathrm{g}} \approx \operatorname{argmin}\{Q_k(x) : \|x - \overline{\mathcal{X}_k}\| \le \Delta_k\}. \tag{0.7}$$?eq:xgetn?

$$\mathsf{S} = \mathbb{1}(\|x_k^{\mathrm{g}} - \overline{\mathcal{X}_k}\| < \rho_k/2), \tag{0.8}$$?eq:shortd?

$$\mathtt{R} = \mathbb{1}(\mathsf{S} = 1 \text{ and the errors in recent models are small}). \tag{0.9}$$?eq:redrho?

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

2. **Interpolation set and model update**. If $\mathsf{S} = 0$, then calculate

$$x_k^{\mathrm{d}} \approx \operatorname{argmin}\{\kappa(\mathcal{X}_k \cup x_k^{\mathrm{g}} \setminus x) : x \in \mathcal{X}_k\}, \tag{0.10}$$?eq:xdropn?

and let $\mathcal{X}_k^+ = \mathcal{X}_k \cup x_k^{\mathrm{g}} \setminus x_k^{\mathrm{d}}$ if $r_k > 0$ or $\kappa(\mathcal{X}_k \cup x_k^{\mathrm{g}} \setminus x_k^{\mathrm{d}}) \le \kappa_0$. In any other case, $\mathcal{X}_k^+ = \mathcal{X}_k$. Set

$$Q_k^+ = \operatorname{argmin}\{\|Q - Q_k\|_{\mathrm{F}} : Q \in \mathcal{Q} \text{ and } Q(x) = f(x) \text{ for all } x \in \mathcal{X}_k^+\}. \tag{0.11}$$?eq:updateq1?

3. **Geometry improvement**. If $\mathsf{S} = 1$ or $r_k < 1/10$, then set

$$y_k^{\mathrm{d}} = \operatorname{argmax}\{\|y - \overline{\mathcal{X}_k^+}\| : y \in \mathcal{X}_k^+\}. \tag{0.12}$$?eq:ydropn?

If $\|y_k^{\mathrm{d}} - \overline{\mathcal{X}_k^+}\| \ge 2\Delta_{k+1}$, then define $\Delta_k^+ = \max\{\rho_k, \min\{\|y_k^{\mathrm{d}} - \overline{\mathcal{X}_k^+}\|/10, \Delta_{k+1}/2\}\}$, calculate

$$y_k^{\mathrm{g}} \approx \operatorname{argmin}\{\kappa(\mathcal{X}_k^+ \cup y \setminus y_k^{\mathrm{d}}) : \|y - \overline{\mathcal{X}_k^+}\| \le \Delta_k^+\}, \tag{0.13}$$?eq:ygetn?

and set $\mathcal{X}_{k+1} = \mathcal{X}_k^+ \cup y_k^{\mathrm{g}} \setminus y_k^{\mathrm{d}}$.

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