A Penalty Decomposition Approach for Multi-Objective Cardinality-Constrained Optimization Problems

SIAM Conference on Optimization 2021, Virtual Conference, 23th July 2021

Matteo Lapucci

DINFO, Univerisity of Florence

Two Popular Topics in Modern Optimization

Sparse Optimization

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \ & f(x) \\ \text{s.t.} \ & \|x\|_0 \leq s, \\ & x \in X. \end{aligned}$$

Multi-Objective Optimization

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \ & F(x) = (f_1(x), \ldots, f_m(x))^T \\ \text{s.t.} \ & x \in X. \end{aligned}$$

Two Popular Topics in Modern Optimization

Sparse Optimization

$$\min_{x\in\mathbb{R}^n} f(x)$$
$$\text{s.t. } \|x\|_0 \leq s,$$
$$x \in X.$$

Multi-Objective Optimization

$$\min_{x\in\mathbb{R}^n} F(x) = (f_1(x), \ldots, f_m(x))^T$$
$$\text{s.t. } x \in X.$$

- Lots of recent literature...

Two Popular Topics in Modern Optimization

Sparse Optimization

$$\min_{x\in\mathbb{R}^n} f(x) \quad \text{s.t.} \quad \|x\|_0 \le s, \quad x \in X.$$

Multi-Objective Optimization

$$\min_{x\in\mathbb{R}^n} F(x) = (f_1(x), \dots, f_m(x))^T \quad \text{s.t.} \quad x \in X.$$

- Lots of recent literature...
- ...but not concerning the combined setting.

Multi-Objetive Problem with Cardinality Constraint

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \; & F(x) = (f_1(x), \ldots, f_m(x))^T \\ \text{s.t.} \; & \|x\|_0 \leq s, \\ & Ax = b, \quad l \leq x \leq u. \end{aligned} \tag{MOCCOP}$$

Multi-Objetive Problem with Cardinality Constraint

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \ & F(x) = (f_1(x), \dots, f_m(x))^T \\ \text{s.t.} \ & \|x\|_0 \le s, \\ & Ax = b, \quad l \le x \le u. \end{aligned} \qquad (\texttt{MOCCOP})$$

- $F : \mathbb{R}^n \to \mathbb{R}^m$ continuously differentiable;

Multi-Objetive Problem with Cardinality Constraint

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \; & F(x) = (f_1(x), \ldots, f_m(x))^T \\ \text{s.t.} \; & \|x\|_0 \le s, \\ & Ax = b, \quad l \le x \le u. \end{aligned} \qquad \texttt{(MOCCOP)}$$

- $F : \mathbb{R}^n \to \mathbb{R}^m$ continuously differentiable;
- ℓ_0 pseudo-norm $\|x\|_0 = |\{i \mid x_i \neq 0\}|$;

Multi-Objetive Problem with Cardinality Constraint

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n}\ & F(x) = (f_1(x), \ldots, f_m(x))^T \\
\text{s.t. }\ & \|x\|_0 \leq s, \\
& Ax = b, \quad l \leq x \leq u.
\end{aligned}
\tag{MOCCOP}
$$

- $F : \mathbb{R}^n \to \mathbb{R}^m$ continuously differentiable;
- ℓ_0 pseudo-norm $\|x\|_0 = |\{i \mid x_i \neq 0\}|$;
- $s < n$.

Preliminaries

- **Sparse Optimization:**
  - Support set: $I_1(x) = \{i \mid x_i \neq 0\}$.
  - Super support set: $J \subseteq \{1, \ldots, n\}$ s.t.
    - $|J| = s$,
    - $J \supseteq I_1(x)$.

Preliminaries

- **Sparse Optimization:**
  - Support set: $I_1(x) = \{i \mid x_i \neq 0\}$.
  - Super support set: $J \subseteq \{1, \dots, n\}$ s.t.
    - $|J| = s$,
    - $J \supseteq I_1(x)$.
- **Multi-objective Optimization:**
  - Pareto Optimality: $\nexists y \in X : F(y) \leq F(x),\ F(x) \neq F(y)$.
  - Weak Pareto Optimality: $\nexists y \in X : F(y) < F(x)$.
  - Pareto stationarity:
$$\theta(x) = \min_{z \in X, \|z-x\| \leq 1} \max_{j=1,\dots,m} \nabla f_j(x)^T (z - x) = 0.$$
  - Steepest common descent direction:
$$\bar{d} = \bar{z} - x : \bar{z} \in \operatorname*{arg\,min}_{z \in X, \|z-x\| \leq 1} \max_{j=1,\dots,m} \nabla f_j(x)^T (z - x).$$
  - MOPGD algorithm: $x^{k+1} = x^k - \alpha_k \bar{d}^k$.

Optimality Conditions for MOCCOP

- Feasible directions at $\bar{x}$:

Optimality Conditions for MOCCOP

- Feasible directions at $\bar{x}$:

$$d \in \mathcal{D}(\bar{x}) \text{ s.t. } \begin{cases} d \in \mathbb{R}^n, \\ \|d_{I_0(\bar{x})}\|_0 \leq s - \|\bar{x}\|_0, \\ d_i \leq 0 \text{ if } \bar{x}_i = u_i, \\ d_i \geq 0 \text{ if } \bar{x}_i = l_i, \\ Ad = 0. \end{cases}$$

Optimality Conditions for MOCCOP

- Feasible directions at $\bar{x}$:

$$d \in \mathcal{D}(\bar{x}) \text{ s.t. } \begin{cases} d \in \mathbb{R}^n, \\ \|d_{I_0(\bar{x})}\|_0 \leq s - \|\bar{x}\|_0, \\ d_i \leq 0 \text{ if } \bar{x}_i = u_i, \\ d_i \geq 0 \text{ if } \bar{x}_i = l_i, \\ Ad = 0. \end{cases}$$

- Pareto Stationarity:

Optimality Conditions for MOCCOP

- Feasible directions at $\bar{x}$:

$$d \in \mathcal{D}(\bar{x}) \text{ s.t. } \begin{cases} d \in \mathbb{R}^n, \\ \|d_{I_0(\bar{x})}\|_0 \leq s - \|\bar{x}\|_0, \\ d_i \leq 0 \text{ if } \bar{x}_i = u_i, \\ d_i \geq 0 \text{ if } \bar{x}_i = l_i, \\ Ad = 0. \end{cases}$$

- Pareto Stationarity:

$$\theta(\bar{x}) = \min_{d \in \mathcal{D}(\bar{x}), \|d\| \leq 1} \max_{j=1,\dots,m} \nabla f_j(\bar{x})^T d = 0.$$

Optimality Conditions for MOCCOP

- Feasible directions at $\bar{x}$:

$$d \in \mathcal{D}(\bar{x}) \text{ s.t. } \begin{cases} d \in \mathbb{R}^n, \\ \|d_{I_0(\bar{x})}\|_0 \leq s - \|\bar{x}\|_0, \\ d_i \leq 0 \text{ if } \bar{x}_i = u_i, \\ d_i \geq 0 \text{ if } \bar{x}_i = l_i, \\ Ad = 0. \end{cases}$$

- Pareto Stationarity:

$$\theta(\bar{x}) = \min_{d \in \mathcal{D}(\bar{x}), \|d\| \leq 1} \max_{j=1,\dots,m} \nabla f_j(\bar{x})^T d = 0.$$

  - Necessary condition of (local) weak Pareto optimality.

Optimality Conditions for MOCCOP

- Multi-objective Lu-Zhang conditions:

Optimality Conditions for MOCCOP

- Multi-objective Lu-Zhang conditions:
  $\exists$ super support set J at $\bar{x}$ such that:

$$\min_{\substack{d_J \in \mathbb{R}^s \\ \|d_J\| \leq 1 \\ A_J d_J = 0 \\ d_i \leq 0 \text{ if } \bar{x}_i = u_i \\ d_i \geq 0 \text{ if } \bar{x}_i = l_i}} \max_{j=1,\dots,m} \nabla_J f_j(\bar{x})^T d_J = 0$$

Optimality Conditions for MOCCOP

- Multi-objective Lu-Zhang conditions:
  $\exists$ super support set J at $\bar{x}$ such that:

$$\min_{\substack{d_J \in \mathbb{R}^s \\ \|d_J\| \leq 1 \\ A_J d_J = 0 \\ d_i \leq 0 \text{ if } \bar{x}_i = u_i \\ d_i \geq 0 \text{ if } \bar{x}_i = l_i}} \max_{j=1,\dots,m} \nabla_J f_j(\bar{x})^T d_J = 0$$

  - Necessary condition of Pareto-stationarity.

Penalty Decomposition Approach: Basics

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \ & F(x) \\ \text{s.t.} \ & \|x\|_0 \leq s, \\ & Ax = b, \\ & l \leq x \leq u. \end{aligned}$$

Penalty Decomposition Approach: Basics

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \ & F(x) \\ \text{s.t.} \ & \|x\|_0 \le s, \\ & Ax = b, \\ & l \le x \le u. \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} \min_{x,y \in \mathbb{R}^n} \ & F(x) \\ \text{s.t.} \ & \|y\|_0 \le s, \\ & x - y = 0, \\ & Ax = b, \\ & l \le x, y \le u. \end{aligned}$$

Penalty Decomposition Approach: Basics

$$\begin{aligned}\min_{x\in\mathbb{R}^n}\ & F(x)\\ \text{s.t.}\ & \|x\|_0 \le s,\\ & Ax = b,\\ & l \le x \le u.\end{aligned} \quad \Leftrightarrow \quad \begin{aligned}\min_{x,y\in\mathbb{R}^n}\ & F(x)\\ \text{s.t.}\ & \|y\|_0 \le s,\\ & x - y = 0,\\ & Ax = b,\\ & l \le x, y \le u.\end{aligned}$$

- Multi-objective Penalty Function:

$$Q_\tau(x,y) = F(x) + \tau\left(\|x-y\|^2 + \|Ax-b\|^2\right) e$$

Penalty Subproblems

- x **subproblem**:

$$\min_{l \le x \le u} F(x) + \tau \left(\|x - y\|^2 + \|Ax - b\|^2 \right) e$$

Penalty Subproblems

- x **subproblem**:

$$\min_{l \leq x \leq u} F(x) + \tau \left(\|x - y\|^2 + \|Ax - b\|^2 \right) e$$

  - Continuously differentiable bound-constrained multi-objective optimization problem.

Penalty Subproblems

- x **subproblem**:

$$\min_{l \leq x \leq u} F(x) + \tau \left(\|x - y\|^2 + \|Ax - b\|^2 \right) e$$

  - Continuously differentiable bound-constrained multi-objective optimization problem.

- y **subproblem:**

$$\min_{\substack{l \leq y \leq u \\ \|y\|_0 \leq s}} \tau \|x - y\|^2$$

Penalty Subproblems

- x **subproblem**:

$$\min_{l \leq x \leq u} F(x) + \tau \left(\|x - y\|^2 + \|Ax - b\|^2 \right) e$$

  - Continuously differentiable bound-constrained multi-objective optimization problem.

- y **subproblem:**

$$\min_{\substack{l \leq y \leq u \\ \|y\|_0 \leq s}} \tau \|x - y\|^2$$

  - Scalar problem with closed form solution.

Penalty Decomposition Algorithm

Algorithm 1: `MultiObjectiveSparsePenaltyDecomposition`

Input: $\tau_0 > 0$, $\sigma > 1$, $x^0 = y^0 \in \mathbb{R}^n$ s.t. $\|x^0\|_0 \leq s$, $\{\varepsilon_k\}$ s.t. $\varepsilon_k \to 0$.

for $k = 0, 1, \dots$ **do**

- $x_{\text{trial}} = \texttt{MOPGD}(Q_{\tau_k}(\cdot, y^k), x^k, \varepsilon_k)$
- **if** $Q_{\tau_k}(x_{\text{trial}}, y^k) \leq F(x^0)$ **then**
  - $\hat{x}, \hat{y} = x^k, y^k$; // passed boundedness check
- **else**
  - $\hat{x}, \hat{y} = x^0, y^0$; // reset if necessary
- /* Repeat up to approximate Pareto-stat. for Q_{τ_k} w.r.t. x */
- **while** $\theta_{Q_{\tau_k}}(\hat{x}; \hat{y}) < -\varepsilon_k$ **do**
  - $\hat{x} = \texttt{MOPGD}(Q_{\tau_k}(\cdot, \hat{y}), \hat{x}, \varepsilon_k)$; // x update by steepest descent
  - $\hat{y} = \arg\min_y \frac{\tau_k}{2} \|\hat{x} - y\|^2$; // closed form y update (projection)
- $\tau_{k+1} = \sigma \tau_k$; // increase penalty parameter
- $x^{k+1}, y^{k+1} = \hat{x}, \hat{y}$; // new iterate

Output: The sequence $\{x^k, y^k\}$.

Convergence Properties

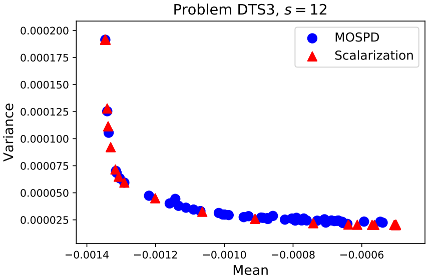
Theorem

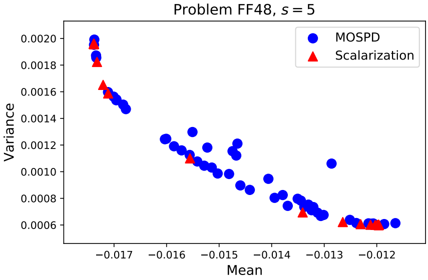
Let $\{x^k, y^k\}$ be the sequence generated by the `MultiObjectiveSparsePenaltyDecomposition` *Algorithm on a (*`MOCCOP`*). Then $\{x^k, y^k\}$ admits cluster points and every cluster point $(\bar{x}, \bar{y})$ is such that*

- $\bar{x} = \bar{y}$,
- $\|\bar{x}\|_0 \leq s$,
- $\bar{x}$ *satisfies the Multi-objective Lu-Zhang conditions for problem (*`MOCCOP`*).*

Preliminary Experiments

- Problem: Multi-objective Sparse Portfolio Selection
- 48 variables
- Multi-start MOSPD strategy vs. Scalarization





Conclusions

- We started a discussion on MOCCOP;
- We theoretically analyzed the problem in terms of optimality conditions;
- We proposed a convergent Penalty Decomposition type algorithm to tackle nonconvex MOCCOP exploiting MO descent methods;
- Promising experimental results support the practical quality of the proposed method.

A Penalty Decomposition Approach for Multi-Objective Cardinality-Constrained Optimization Problems

SIAM Conference on Optimization 2021, Virtual Conference, 23th July 2021

Matteo Lapucci

DINFO, Univerisity of Florence