An augmented Lagrangian method for mixed-integer nonconvex optimization with nonlinear constraints

D.Papadimitriou[1], Bang Vu[1]

3nLab & Belgium Research Center (BeRC), Huawei

dpapadimitriou@3nlab.org,congbang.vu.ext@huawei.com

Mots-clés : *nonlinear optimization, mixed-Integer programming, augmented Lagrangian.*

In this paper, we consider two composite optimization problems with nonlinear constraints. The first involves homogeneous variable spaces which yields a very generic nonlinear mathematical programming model (**P1**). The second where some variables belong to the real Hilbert space, and others to the binary space $\mathbb{Z}_2$ (**P2**). The nonconvexity of the latter is not assumed to be exclusively caused by the integrality constraints.

Problem Assume $\mathcal{H}$ and $\mathcal{G}$ are real finite dimensional Hilbert spaces. Let $f\colon \mathcal{H} \to]-\infty,+\infty[$ and $g\colon \mathcal{G} \to]-\infty,+\infty[$ be continuously differentiable functions (at least $\mathcal{C}^1$) with μ_f and μ_g-Lipschitz continuous gradient, respectively. Let $c_1\colon \mathcal{H}\times\mathcal{G} \to \mathbb{R}^{m_1}$ and $c_2\colon \mathcal{H}\times\mathcal{G} \to \mathbb{R}^{m_2}$ be smooth functions with μ_{c_1} and μ_{c_2}-Lipschitz continuous gradient. The problem is to

$$
\begin{aligned}
&\text{minimize } f(x)+g(y) && (1)\\
&\text{subject to } c_1(x,y)=b_1, c_2(x,y)\le b_2,\\
&\mathbf{P1}:\ (x,y)\in C(\text{closed convex})\subset \mathcal{H}\times\mathcal{G},\ b_1\in\mathbb{R}^{m_1},\ b_2\in\mathbb{R}^{m_2} && (2)\\
&\mathbf{P2}:\ x\in\mathcal{H},\ y\in\{0,1\}^n,\ b_1\in\mathbb{R}^{m_1},\ b_2\in\mathbb{R}^{m_2}. && (3)
\end{aligned}
$$

Finding a global optimizer for such nonconvex optimization problem with both nonlinear inequality and equality constraints is NP-hard in general. Instead, the handling of such problem considers the finding of the so-called Karush-Kuhn-Tucker (KKT) points. The KKT conditions provide a generalization of the Lagrange multiplier theorem to inequality constrained problems. More precisely, under some constraint qualification conditions (to ensure existence of KKT multipliers), the *necessary* conditions to be verified by a KKT point are stationarity, primal/dual feasibility and complementarity slackness. From this perspective, the main objectives of this paper are i) to propose an Augmented Lagrangian method (ALM) for the model described hereabove with both homogeneous and mixed variable spaces that converges to KKT points, and ii) to establish the complexity of finding a KKT point of these problems.

Main Algorithm The proposed ALM-based Algorithm 1 relies on Backtracking Line Search to determine the primal stepsize t_k and in turn to decrease of the (generalized) Augmented Lagrangian $\mathcal{L}_{\rho_k}(u_k,\lambda_{1,k},\lambda_{2,k})$ where $u=(x,y)\in\mathcal{H}\times\mathcal{G}$, ρ_k denotes the smoothing parameter and $\lambda_{1,k}$, $\lambda_{2,k}$ the KKT multipliers. The update rule of the dual stepsize σ_k (Step 3) ensures that $(\sigma_k)_{k\in\mathbb{N}}$ is gradually increasing, bounded and the stepsize t_k obtained by line search is not too small. This algorithm is structured as a single loop, i.e., it does not require calling a first-order subsolver to compute inner iterates (primal subproblems). For **P2**, we first reformulate the constraints $y\in\{0,1\}^n$ into $\overline{y}\ (=2y-1)\in\{-1,1\}^n$. Then, using [1], the constraints $\overline{y}\in\{-1,1\}^n$ can be equivalently reformulated as $\overline{y}\in\{-1,1\}^n \Leftrightarrow \{\overline{y}\mid -1\le\overline{y}\le 1, \|\overline{y}\|_2^2=n\}$. The original problem becomes an instance of Problem **P1** with

$$
\begin{cases}
c_1\colon (x,\overline{y})\mapsto (\overline{h}_1(x,\overline{y}),\|\overline{y}\|_2^2),\\
c_2\colon (x,\overline{y})\mapsto \overline{h}_2(x,\overline{y}),\\
C=\big\{(x,\overline{y})\mid x\in\mathcal{H}, -1\le\overline{y}\le 1\big\}.
\end{cases}
\tag{4}
$$

To guarantee feasibility, certain regularity conditions need to be further imposed.

Assumption 1 Let Z be an nonempty subset of $\mathcal{H}$ and Y_1 a subset of $\mathbb{R}^{m_1}$ and Y_2 a subset of $\mathbb{R}^{m_2}$. Assume i) $\mu_0 = \sup_{u\in Z} \max\{\|J_{c_1}(u)^\star\|, \|J_{c_2}(u)^\star\|\} < +\infty$ (**A1**), where $J_c(x)$ denotes the Jacobian of the function c at x and ii) the uniform regularity conditions of both c_1 and c_2 on Z with constant $\zeta \in \,]0,+\infty[$ with respect to Y_1 and Y_2 ; more precisely, $\exists \zeta \in \,]0,+\infty[$ such that $(\forall (u, v_1) \in Z \times Y_1)\ \|J_{c_1}(u)^\star v_1\| \geq \zeta\|v_1\|$ and $(\forall (u, v_2) \in Z \times Y_2)\ \|J_{c_2}(u)^\star v_2\| \geq \zeta\|v_2\|$ (**A2**).

Algorithm 1 ALM algorithm with backtracking

1: Set $u_0 \in C$, $\lambda_{1,0} \in \mathbb{R}^{m_1}$ $\lambda_{2,0} \in \mathbb{R}^{m_2}$, $u_{-1} \neq u_0$

2: Set $\sigma_{-1} \gg 1$, $\rho_{-1} > 0$, $\varepsilon \in \,]0,(\mu_f+\mu_g)/2[$, $(\theta,\nu,\vartheta) \in \,]0,1[^3$

3: Compute μ_0 and ζ from **A1** and **A2**, respectively.

4: **for** $k \leftarrow 0 : n$ **do**

5: $\quad\triangleright$ **Step 1** : Select $\rho_k \in \,]0,\infty[$ such that

$$\begin{cases} \xi_k := \dfrac{\mu_f+\mu_g}{2} - \dfrac{\rho_k}{2}\Big(\|J_{c_1}(u_k)\|^2 + \|J_{c_1}(u_k)\|^2\Big) > \varepsilon \\ \rho_k\|c_1(u_k) - b_1\| \leq \sigma_{k-1}\|u_k - u_{k-1}\| \\ \rho_k\||c_2(u_k) - b_2 - P_S\big(c_2(u_k) - b_2 + \rho_k^{-1}\lambda_{2,k}\big)\|| \leq \sigma_{k-1}\|u_k - u_{k-1}\| \\ \rho_k - \rho_{k-1} < \varepsilon\sigma_{k-1} \end{cases} \tag{5}$$

6: $\quad\triangleright$ **Step 2** : Compute $v_{1,k}$, $v_{2,k}$ and d_k

$$v_{1,k} = \lambda_{1,k} + \rho_k(c_1(u_k) - b_1),\ v_{2,k} = \lambda_{2,k} + \rho_k\big(c_2(u_k) - b_2 - P_S(c_2(u_k) - b_2 + \rho_k^{-1}\lambda_{2,k})\big) \tag{6}$$

$$d_k = \frac{-\big(\nabla(f+g)(u_k) + J_{c_1}(u_k)^\star v_{1,k} + J_{c_2}(u_k)^\star v_{2,k}\big)}{\mu_g+\mu_f} \tag{7}$$

7: $\quad\triangleright$ **Step 3** : Set $\sigma_k = \sigma_{k-1} + \vartheta/k$ and find $t_k = \theta^j$, $j \in \mathbb{N}$ such that

$$\begin{cases} \varphi_k\big(P_C(u_k + t_k d_k)\big) < \varphi_k(u_k) + t_k\nu\Delta\varphi_k(u_k, d_k) + \mathcal{O}\big(\frac{1}{(k+1)^{1+\varepsilon}}\big) \\ 8(1+\varepsilon)t_k\zeta^{-2}\Big(\big(\mu_0\sigma_k + \mu_{c_1}\|\lambda_{1,k}\|\big)^2 + \big(\mu_0\sigma_k + \mu_{c_2}\|\lambda_{2,k}\|\big)^2\Big) \leq \sigma_k\xi_k \end{cases} \tag{8}$$

8: $\quad\triangleright$ **Step 4** : Update

$$u_{k+1} = P_C\big(u_k + t_k d_k\big) \tag{9}$$

$$\lambda_{1,k+1} = \lambda_{1,k} + \sigma_k\big(c_1(u_{k+1}) - b_1\big),\ \lambda_{2,k+1} = \lambda_{2,k} + \sigma_k\Big(c_2(u_{k+1}) - b_2 - P_S\big(c_2(u_{k+1}) - b_2 + \frac{\lambda_{2,k}}{\rho_k}\big)\Big) \tag{10}$$

9: **end for**

Convergence properties In this paper, we determine the conditions for local convergence to a critical point of $\mathcal{L}$ by considering the sequences $(u_k, \lambda_{1,k}, \lambda_{2,k})_{k\in\mathbb{N}}$ produced by Algorithm 1. For this purpose, the following properties are assumed. Let $(u_k)_{k\in\mathbb{N}} \subset Z \subseteq C$.

- **C0** : the constraints (c_1, c_2) verify condition **A2** ;
- **C1** : the sequence $(\mathcal{L}_{\rho_k}(u_k, \lambda_{1,k}, \lambda_{2,k}))_{k\in\mathbb{N}}$ is lower bounded ;
- **C2** : the sequence $(\rho_k)_{k\in\mathbb{N}}$ is upper bounded ;
- **C3** : the function c_1 is coercive, i.e., $\lim_{\|u\|\to+\infty} \|c_1(u)\| = +\infty$.

Moreover, assuming that the i) limiting continuity, ii) sufficient decrease of the generalized Lagrangian, and iii) gradient boundedness conditions are verified in addition to **C1**, then, our Algorithm 1 provides globally convergent sequences, i.e., for arbitrary starting point, the algorithm generates a sequence that converges to a solution.

Références

[1] G. Yuan and B. Ghanem, An Exact Penalty Method for Binary Optimization Based on MPEC Formulation, Proc. of the Thirty-First AAAI Conf. on Artificial Intelligence (AAAI-17), pp.2867-2875, 2017.

[2] D. Papadimitriou and B. Vu, An augmented Lagrangian method for nonconvex composite optimization problems with nonlinear constraints, submitted (under review), 2022.