DIFFERENTIAL EQUATION METHOD BASED ON APPROXIMATE AUGMENTED LAGRANGIAN FOR NONLINEAR PROGRAMMING

LI JIN AND HONGYING HUANG*

School of Mathematics, Zhejiang Ocean University
Key Laboratory of Oceanographic Big Data Mining & Application of Zhejiang Province
Zhoushan, Zhejiang 316022, China

(Communicated by Jinyan Fan)

Abstract. This paper analyzes the approximate augmented Lagrangian dynamical systems for constrained optimization. We formulate the differential systems based on first derivatives and second derivatives of the approximate augmented Lagrangian. The solution of the original optimization problems can be obtained at the equilibrium point of the differential equation systems, which lead the dynamic trajectory into the feasible region. Under suitable conditions, the asymptotic stability of the differential systems and local convergence properties of their Euler discrete schemes are analyzed, including the locally quadratic convergence rate of the discrete sequence for the second derivatives based differential system. The transient behavior of the differential equation systems is simulated and the validity of the approach is verified with numerical experiments.

1. Introduction. This paper considers the following constrained optimization problems:

\[
\min f(x) \quad \text{s.t. } G(x) \in K, \tag{1}
\]

where \( K = \{0_q \} \times \mathbb{R}^{n-p} \), \( G = (g_1, \ldots, g_p)' \), \( f : \mathbb{R}^n \to \mathbb{R} \) and \( g_i : \mathbb{R}^n \to \mathbb{R} \) (\( i = 1, \cdots, p \)) are twice continuously differentiable functions. The augmented Lagrangian method is one of the most popular algorithms for solving problem \((1)\). Originally, the method was applied to problems with equality constraints \([10, 19]\) and later generalized to problems with inequality constraints \([20, 21]\). The first augmented Lagrangian, namely the proximal Lagrangian, was introduced by Rockafellar \([22]\) and the theory of augmented Lagrangians was developed in, e.g., Ioffe \([13]\), Bertsekas \([2, 3]\) and Rockafellar \([23]\) for constrained optimization problems.

Before going further, we mention that the approximate augmented Lagrangian method was proposed by Huang et al.\([12]\), in which the approximate augmented...
Lagrangian with the parameter \( \tau \) for problem (1) is defined

\[
L_\varepsilon(x, \lambda, \tau) := f(x) + \sum_{i=1}^{q} \lambda_i g_i(x) + \frac{\tau}{2} \sum_{i=1}^{q} g_i^2(x) + \frac{1}{2\tau} \sum_{i=q+1}^{p} \left( \frac{1}{4} p_c^2(x, \lambda_i) - \lambda_i^2 \right),
\]

where \( \varepsilon \geq 0, \tau \geq 1, \) and

\[ p_c(x, \lambda_i) = \lambda_i + \tau g_i(x) + \sqrt{(\lambda_i + \tau g_i(x))^2 + \varepsilon^2}. \]

Especially, the approximate augmented Lagrangian \( L_\varepsilon(x, \lambda, \tau) \) with \( \varepsilon = 0 \) reduces to the augmented Lagrangian of problem (1) in the following form

\[
L_\tau(x, \lambda) := f(x) + \frac{1}{2\tau} \left[ \| \Pi_{\mathbb{R}^q \times \mathbb{R}^p_{\geq 0}} (\lambda + \tau G(x)) \|_2^2 - \| \lambda \|_2^2 \right].
\]

It is well known that the augmented Lagrangian function is not twice continuously differentiable even when both \( f \) and \( g \) are twice continuously differentiable and the convergence rate of the augmented Lagrangian method is proportional to \( 1/\tau \), that is linear (see refs. [2] for more discussions). However, the approximate augmented Lagrangian defined by (2) with \( \varepsilon > 0 \) is different from the augmented Lagrangian in that it is also twice continuously differentiable if both \( f \) and \( g \) are twice continuously differentiable. In some sense, the approximate augmented Lagrangian can be deemed as the smooth function of the augmented Lagrangian. This characteristic may allow us to apply the differential equation method to solve constrained optimization problems. Early differential equation methods were introduced by Arrow and Hurwicz [1], some results have been addressed in the work (see Refs. [9]-[18] for details). Among them, Evtushenko and Zhadan [8]-[7] have studied, by using the so-called space transformation techniques, a family of numerical methods for solving optimization problems with equality and inequality constraints. The proposed algorithms are based on the numerical integration of the systems of ordinary differential equation. Along this line, Zhang [25]-[26] and Jin [14]-[17] studied modified versions of differential equation methods.

In this paper, we analyze differential equation methods for constrained optimization problems. Rather than solving an optimization problem by multiple iterations using a digital computer, one can obtain the solution at an equilibrium point by setting up the associated differential equation systems, which lead the dynamic trajectory into the feasible region. The main idea of the method is to construct the differential systems based on the approximate augmented Lagrangian. The corresponding systems mainly consist of first-order derivatives and second-order derivatives of the approximate augmented Lagrangian. It is demonstrated that the equilibrium point of this system coincides with the solution to the constrained optimization problems. The analysis of the method is made on the basis of the stability theory of the solution of ordinary differential equations. Motivated by finding an approach to the construction of differential systems to solve constrained optimization problems without using space transformations of Evtushenko and Zhadan, we construct systems based on the approximate augmented Lagrangian. The differential systems are carried out without using the space transformation and this feature provides a high rate of convergence. Under a set of suitable conditions, we prove the asymptotical stability of the differential systems and local convergence.
properties of their Euler discrete schemes, including the locally quadratic convergence rate of the discrete sequence for second order derivatives based differential equation systems.

The paper is organized as follows. The approximate augmented Lagrangian and related properties are introduced in Section 2.1. The differential system based on the approximate augmented Lagrangian is derived and the asymptotic stability theorem is established under mild conditions in Section 2.2. In Section 2.3, the Euler discrete schemes for the differential system are presented and the local convergence theorem is demonstrated. In Section 3, we construct a second-order derivative based differential system and prove the asymptotic stability of the system. Euler discrete schemes and their local convergence properties are obtained, including the locally quadratic convergence rate of the discrete sequence for second order derivatives based differential system. The numerical results show that the method has better stability and higher precision in Section 4.

2. Differential systems based on first-order derivatives. In this section, we mention some preliminaries that will be used throughout this paper. Without loss of generality, we assume that \( f(x) \) and \( g_i(x) \), \( i = 1, \cdots, p \), are twice continuously differentiable, then the classical Lagrangian for problem (1) is defined by \( L(x, \lambda) = f(x) + \langle \lambda, G(x) \rangle \), denote \( E = \{ i \mid i = 1, 2, \cdots, q \} \) and the active set of indices \( I(x) = \{ i \mid g_i(x) = 0, i = q + 1, \cdots, p \} \). Let \( x^* \) be a local optimal point to problem (1) and the pair \( (x^*, \lambda^*) \) be the corresponding KKT point, which satisfies the following conditions:

\[
\nabla_x L(x^*, \lambda^*) = \nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla g_i(x^*) = 0, \tag{4}
\]

\[
\lambda_i^* g_i(x^*) = 0, \quad \lambda_i^* \geq 0, \quad g_i(x^*) \leq 0, \quad i = q + 1, \cdots, p.
\]

Let the Jacobian uniqueness conditions, proposed in [11], hold at \( (x^*, \lambda^*) \):

1. The multipliers \( \lambda_i^* > 0 \), \( i \in I(x^*) \).
2. The gradients \( \nabla g_i(x^*) \), \( i \in E \cup I(x^*) \) are linearly independent.
3. \( y^\top \nabla_{xx}^2 L(x^*, \lambda^*) y > 0 \), \( \forall 0 \neq y \in \{ y \mid \nabla g_i(x^*)^\top y = 0, i \in E \cup I(x^*) \} \).

2.1. The properties of the approximate augmented Lagrangian. The following lemma will be used in the proof of the forthcoming theorem.

Lemma 2.1. (See Ref.[3]). Let \( A \) be a \( n \times n \) symmetric matrix, \( B \) be a \( p \times n \) matrix, \( U = \text{diag}(\mu) \), where \( \mu = (\mu_1, \cdots, \mu_p) > 0 \). Assume that \( \langle Ay, y \rangle \geq \lambda \langle y, y \rangle \) holds true for some \( \lambda > 0 \) and all \( y \in \mathbb{R}^n \) satisfying \( By = 0 \). Then there are scalars \( k_0 > 0 \) and \( c \in (0, \lambda) \) such that, for any \( k \geq k_0 \),

\[
\langle (A + kB^T UB)x, x \rangle \geq c(x, x), \quad \forall x \in \mathbb{R}^n.
\]

Now we discuss the properties of the approximate augmented Lagrangian.

Theorem 2.2. Let \( (x^*, \lambda^*) \) be a KKT point of (1), the Jacobian uniqueness conditions hold at \( (x^*, \lambda^*) \), then

(I): \( \lim_{\varepsilon \to 0^+} \nabla_x L_\varepsilon(x^*, \lambda^*, \tau) = 0_n, \quad \tau \geq 0. \)

(II): \( \exists \tilde{\tau} > 0 \) and \( c > 0 \), for any \( \tau \geq \max\{\tilde{\tau}, 1\} \) and \( \varepsilon \) sufficiently close to zero, such that

\[
\langle \nabla_{xx}^2 L_\varepsilon(x^*, \lambda^*, \tau) w, w \rangle \geq c \langle w, w \rangle, \quad \forall w \in \mathbb{R}^n.
\]
Proof. The gradient and Hessian of $L_\varepsilon(x, \lambda, \tau)$ with respect to $x$ are

$$
\nabla_x L_\varepsilon(x, \lambda, \tau) = \nabla f(x) + \sum_{i=1}^{q} \lambda_i \nabla g_i(x) + \tau \sum_{i=1}^{q} g_i(x) \nabla g_i(x)
+ \frac{1}{2\tau} \sum_{i=q+1}^{p} \frac{1}{2} \nabla p_\varepsilon(x, \lambda_i) \nabla x p_\varepsilon(x, \lambda_i),
$$

$$
\nabla^2_{xx} L_\varepsilon(x, \lambda, \tau) = \nabla^2 f(x) + \sum_{i=1}^{q} \lambda_i \nabla^2 g_i(x) + \tau \sum_{i=1}^{q} g_i(x) \nabla^2 g_i(x)
+ \tau \sum_{i=1}^{q} \nabla g_i(x) \nabla g_i(x)^\top + \frac{1}{2\tau} \sum_{i=q+1}^{p} \frac{1}{2} (\nabla x p_\varepsilon(x, \lambda_i) \nabla x p_\varepsilon(x, \lambda_i))^\top
$$

where

$$
\nabla x p_\varepsilon(x, \lambda_i) = \frac{\tau \nabla g_i(x) p_\varepsilon(x, \lambda_i)}{\sqrt{\lambda_i \tau g_i(x))^2 + \varepsilon^2}},
$$

$$
\nabla^2_{xx} p_\varepsilon(x, \lambda_i) = \frac{\tau \nabla g_i(x) \nabla x p_\varepsilon(x, \lambda_i)}{\sqrt{\lambda_i \tau g_i(x))^2 + \varepsilon^2}} + \frac{\tau \nabla^2 g_i(x) p_\varepsilon(x, \lambda_i)}{\sqrt{\lambda_i \tau g_i(x))^2 + \varepsilon^2}} \frac{(\sum_{i=1}^{q} \nabla g_i(x)) \nabla g_i(x)^\top}{\lambda_i \tau g_i(x))^2 + \varepsilon^2}.
$$

In particular, since $x^*$ and $\lambda^*$ satisfy the Jacobian uniqueness conditions, we have

$$
\lim_{\varepsilon \to 0^+} p_\varepsilon(x^*, \lambda_i^*) = \begin{cases} 
2\lambda_i^*, & i \in I(x^*) \\
0, & \text{otherwise}
\end{cases},
$$

$$
\lim_{\varepsilon \to 0^+} \nabla x p_\varepsilon(x^*, \lambda_i^*) = \begin{cases} 
2\tau \nabla g_i(x^*), & i \in I(x^*) \\
0, & \text{otherwise}
\end{cases},
$$

$$
\lim_{\varepsilon \to 0^+} \nabla^2_{xx} p_\varepsilon(x^*, \lambda_i^*) = \begin{cases} 
2\tau \nabla^2 g_i(x^*), & i \in I(x^*) \\
0, & \text{otherwise}
\end{cases},
$$

and in view of Eq. (5), it follows that

$$
\lim_{\varepsilon \to 0^+} \nabla x L_\varepsilon(x^*, \lambda^*, \tau) = \nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla g_i(x^*) = \nabla x L(x^*, \lambda^*) = 0_n,
$$

$$
\lim_{\varepsilon \to 0^+} \nabla^2_{xx} L_\varepsilon(x^*, \lambda^*, \tau) = \nabla^2_{xx} L(x^*, \lambda^*) + \tau \sum_{i \in E \cup I(x^*)} \nabla g_i(x^*) \nabla g_i(x^*)^\top.
$$

By the Jacobian uniqueness conditions, we have that $y^T \nabla^2_{xx} L(x^*, \lambda^*) y > 0, \forall 0 \neq y \in \{y \mid \nabla g_i(x^*)^\top y = 0, i \in E \cup I(x^*)\}$. So by applying Lemma 2.1 with $A = \nabla^2_{xx} L(x^*, \lambda^*)$, it follows that there exists $\hat{\tau} > 0$ and $c > 0$, for any $\tau \geq \max\{\hat{\tau}, 1\}$ and $\varepsilon$ sufficiently close to 0, such that

$$
\langle \nabla^2_{xx} L_\varepsilon(x^*, \lambda^*, \tau) w, w \rangle \geq c(w, w), \forall w \in \mathbb{R}^n.
$$

\qed
2.2. The asymptotic stability of the differential system. To obtain the numerical solution of problem (1), we seek the limit point of the solutions of the system described by the following vector differential equation

\[
\frac{dz}{dt} = -\Lambda(z, \tau) = - \begin{bmatrix} \nabla z L_e(x, \lambda, \tau) \\ -G_q(x) \\ -P_\lambda(x) \end{bmatrix}, \tag{7}
\]

where \( z^T = (x^T, \lambda^T) \), \( G_q(x) = (g_1(x) \ g_2(x) \ \ldots \ g_q(x))^T \) and

\[
P_\lambda(x) = \begin{bmatrix}
\frac{1}{2\tau} \left( \frac{1}{2} p_\tau(x, \lambda_{q+1}) \frac{\partial}{\partial \lambda_{q+1}} p_\tau(x, \lambda_{q+1}) - 2\lambda_{q+1} \right) \\
\frac{1}{2\tau} \left( \frac{1}{2} p_\tau(x, \lambda_{q+2}) \frac{\partial}{\partial \lambda_{q+2}} p_\tau(x, \lambda_{q+2}) - 2\lambda_{q+2} \right) \\
\vdots \\
\frac{1}{2\tau} \left( \frac{1}{2} p_\tau(x, \lambda_p) \frac{\partial}{\partial \lambda_p} p_\tau(x, \lambda_p) - 2\lambda_p \right)
\end{bmatrix}.
\]

We will investigate the local behaviour of trajectories of system (7) in the neighborhood of the point \( z^* \).

**Theorem 2.3.** Assume that \( f \) and \( G \) are twice continuously differentiable, let \((x^*, \lambda^*)\) be a KKT pair, and the Jacobian uniqueness conditions hold at \((x^*, \lambda^*)\).

Then, for any \( \tau \geq \max\{\hat{\tau}, 1\} \) and \( \varepsilon \) sufficiently close to \( 0 \), the system (7) is asymptotically stable at \((x^*, \lambda^*)\).

**Proof.** Without loss of generality, we assume that \( I(x^*) = \{q+1, q+2, \ldots, q+r\} \) with \( q+r \leq p \). Linearizing system (7) in the neighborhood of \( z^* \),

\[
\frac{dz}{dt} = -Q(z - z^*), \tag{8}
\]

where

\[
Q = \begin{bmatrix}
\nabla z L_e(x^*, \lambda^*, \tau) & \hat{G}(x^*, \lambda^*)^T \\
-\hat{G}(x^*, \lambda^*) & Q_{22}
\end{bmatrix}, \quad Q_{22} = \text{diag}(0, \ldots, 0, \psi_{q+1}, \ldots, \psi_p),
\]

\[
\hat{G}(x^*, \lambda^*) = (\nabla g_1(x^*), \nabla g_2(x^*), \ldots, \nabla g_q(x^*), \hat{p}_{q+1}(x^*, \lambda_{q+1}^*) \ldots, \hat{p}_p(x^*, \lambda_p^*))
\]

and

\[
\psi_j = \frac{1}{2\tau} \left(1 - \frac{1}{2} \left( \frac{2\lambda_j^*}{\sqrt{\lambda_j^* \varepsilon^2 + \varepsilon^2}} + \frac{\lambda_j^* \varepsilon^2}{(\lambda_j^* \varepsilon^2 + \varepsilon^2)(\lambda_j^* \varepsilon^2 + \varepsilon^2)} \right) \right),
\]

\[
\psi_l = \frac{1}{2\tau} \left(1 - \frac{1}{2} \left( \frac{2\tau g_l(x^*)}{(\tau g_l(x^*))^2 + \varepsilon^2} + \frac{\tau g_l(x^*) \varepsilon^2}{(\tau g_l(x^*))^2 + \varepsilon^2}(\tau g_l(x^*))^2 + \varepsilon^2) \right) \right),
\]

\[
\hat{p}_i(x^*, \lambda_i^*) = -\frac{1}{4\tau} \left( \frac{\partial}{\partial \lambda_i} p_e(x^*, \lambda_i^*) \nabla_x p_e(x^*, \lambda_i^*) + p_e(x^*, \lambda_i^*) \nabla_x \left( \frac{\partial}{\partial \lambda_i} p_e(x^*, \lambda_i^*) \right) \right),
\]

where \( q+1 \leq i \leq p, q+1 \leq j \leq q+r, q+r+1 \leq l \leq p \).

The stability of system (7) is determined by the properties of the roots of the characteristic equation

\[
\det(Q - \lambda I_{n+p}) = 0. \tag{9}
\]

We will show that the real part of each eigenvalue of \( Q \) is strictly positive, and then the result will follow from Theorem 2.2. For any complex vector \( y \), denote by \( \hat{y} \) its complex conjugate, and for any complex number \( \beta \), denote by \( \text{Re}(\beta) \) its real part.
Let \( \alpha \) be an eigenvalue of \( Q \), and let \( (\omega_1^T, \omega_2^T)^T \neq 0 \) be a corresponding eigenvector, where \( \omega_1 \) and \( \omega_2 \) are complex vectors of dimension \( n \) and \( p \), respectively. We have
\[
\text{Re}\{ \hat{\omega}_1^T, \hat{\omega}_2^T \} Q \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \text{Re}\{ \alpha (\hat{\omega}_1^T, \hat{\omega}_2^T) \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} \} = \text{Re}(\alpha)(\|\omega_1\|^2 + \|\omega_2\|^2). \tag{10}
\]
In view of the definition of \( Q \), we have
\[
\text{Re}\{ \hat{\omega}_1^T, \hat{\omega}_2^T \} Q \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \text{Re}\{ \hat{\omega}_1^T \nabla_{xx}^2 L_c(x^*, \lambda^*, \tau) \omega_1 - \hat{\omega}_2^T \hat{G}(x^*, \lambda^*)^T \omega_1 \\
+ \hat{\omega}_1^T \hat{G}(x^*, \lambda^*) \omega_2 + \hat{\omega}_2^T Q_{22} \omega_2 \}.
\]
Since for any real \( n \times p \) matrix \( M \), it follows that
\[
\text{Re}\{ \hat{\omega}_2^T M^T \omega_1 \} = \text{Re}\{ \hat{\omega}_1^T M \omega_2 \}, \tag{11}
\]
while at the same time, by using Eq. (9),
\[
\hat{\omega}_2^T Q_{22} \omega_2 = \hat{\omega}_2^T \text{diag}(\psi_{q+1}, \cdots, \psi_p) \omega_2, \tag{12}
\]
where \( \hat{\omega}_2^T = (\omega_{21}^T, \omega_{22}^T) \), \( \omega_{21} \) and \( \omega_{22} \) are complex vectors of dimension \( q \) and \( p - q \), respectively. It follows from Eqs. (11) and (12) that
\[
\text{Re}\{ \hat{\omega}_1^T, \hat{\omega}_2^T \} Q \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \text{Re}\{ \hat{\omega}_1^T \nabla_{xx}^2 L_c(x^*, \lambda^*, \tau) \omega_1 \\
+ \hat{\omega}_2^T \text{diag}(\psi_{q+1}, \cdots, \psi_p) \omega_2 \} \tag{13}
\]
The above relation and Eq. (10) imply that
\[
\text{Re}(\alpha)(\|\omega_1\|^2 + \|\omega_2\|^2) = \text{Re}\{ \hat{\omega}_1^T \nabla_{xx}^2 L_c(x^*, \lambda^*, \tau) \omega_1 \\
+ \hat{\omega}_2^T \text{diag}(\psi_{q+1}, \cdots, \psi_p) \omega_2 \}. \tag{14}
\]
Since the matrix \( \nabla_{xx}^2 L_c(x^*, \lambda^*, \tau) \) is positive definite, and when \( \tau > 0 \), we obtain \( \psi_j > 0 \), \( q + 1 < j < p \), then \( \forall \omega_1 \neq 0 \) or \( \forall \omega_2 \neq 0 \),
\[
\text{Re}\{ \hat{\omega}_1^T \nabla_{xx}^2 L_c(x^*, \lambda^*, \tau) \omega_1 + \hat{\omega}_2^T \text{diag}(\psi_{q+1}, \cdots, \psi_p) \omega_2 \} > 0, \tag{15}
\]
it follows from Eqs. (14) and (15) that either \( \text{Re}(\alpha) > 0 \) or else \( \omega_1 = 0, \omega_2 = 0 \). But for \( \omega_1 = 0, \omega_2 = 0 \), the equation
\[
Q \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \alpha \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}, \tag{16}
\]
yields
\[
\hat{G}(x^*, \lambda^*) \omega_2 = 0,
\]
which is also expressed as
\[
[\nabla g_1(x^*) \nabla g_2(x^*) \cdots \nabla g_q(x^*)] \omega_{21} + [\hat{p}_{q+1}(x^*, \lambda^*_{q+1})] \cdots [\hat{p}_p(x^*, \lambda^*_p)] \omega_{22} = 0,
\]
and implies that
\[
[\nabla g_1(x^*) \nabla g_2(x^*) \cdots \nabla g_q(x^*)] \omega_{21} = 0.
\]
Since \( \nabla g_1(x^*), \nabla g_2(x^*), \cdots, \nabla g_q(x^*) \) has rank \( q \), it follows that \( \omega_{21} = 0 \). This contradicts our earlier assumption that \( (\omega_1^T, \omega_2^T) \neq 0 \). Consequently, we must have \( \text{Re}(\alpha) > 0 \).

Therefore, we have that all eigenvalues of \( -Q \) have negative real parts, it follows from Lyapunov’s stability theorem of the first-order approximation that \((x^*, \lambda^*)\) is a local asymptotically stable equilibrium point of Eq. (7).
2.3. Euler discrete schemes. Integrating the system (7) by the Euler method, one obtains the iterate process

\[ z_{k+1} = z_k - h_k \Lambda(z_k, \tau), \quad (17) \]

where \( h_k \) is a stepsize.

The following theorem tells us that the Euler discrete schemes (17) with a constant stepsize is locally convergent.

**Theorem 2.4.** Assume that \( f \) and \( G \) are twice continuously differentiable, let \((x^*, \lambda^*)\) be a KKT pair, and the Jacobian uniqueness conditions hold at \((x^*, \lambda^*)\). Then there exists \( \bar{h} > 0 \) such that for any \( 0 < h_k < \bar{h} \) and \( \tau \geq \max\{r, 1\} \), the iterations defined by (17) converge locally to \((x^*, \lambda^*)\).

**Proof.** Let \( \mathbf{z}^\top = (x^\top, \lambda^\top) \) and

\[ F(z_k) = z_k - h_k \Lambda(z_k, \tau). \quad (18) \]

The convergence of the iterations (17) will be proved if we demonstrate that a \( \bar{h} \) can be chosen such that the iterations defined by

\[ z_{k+1} = F(z_k) \]

converge to \( z^* \) whenever \( z_0 \) is in a neighborhood of \( z^* \) and \( 0 < h_k < \bar{h} \). Let \( \nabla F(z) \) be the transpose of Jacobian of \( F(z) \) and \( \nu_1, \ldots, \nu_{n+p} \) be the eigenvalues of the matrix \( \nabla F(z^*)^\top \). The analysis involves investigating roots of the equation

\[ \det(\nabla F(z^*)^\top - \nu_j I_{n+p}) = 0, \quad 1 \leq j \leq n + p. \]

From the proof of Theorem 2.3, let \((a_j + ib_j), 1 \leq j \leq n + p\) are eigenvalues of the matrix \( Q \), we have \( a_j > 0, 1 \leq j \leq n + p \). Then

\[ \nu_j = (1 - ha_j) - i(hb_j), \]

the condition \(|\nu_j| < 1\) can be written as

\[ h < 2a_j/(a_j^2 + b_j^2), \quad 1 \leq j \leq n + p. \]

Let

\[ \bar{h} = \min\{2a_j/(a_j^2 + b_j^2) \mid 1 \leq j \leq n + p\}, \]

then the spectral radius of \( \nabla F(z^*)^\top \) is strictly less than 1 for \( h < \bar{h} \), and the iterations generated by the scheme (17) is locally convergent to \((x^*, \lambda^*)\) (see Evtushenko [15]). The proof is completed.

\[ \square \]

3. **Differential systems based on second-order derivatives.** The method in Section 2 can be viewed as a gradient approach, through the system (7), for solving Problem (1). Now we discuss the Newton version. The continuous version of Newton method leads to the initial value problem for the following system of ordinary differential equations

\[ \nabla Z(x, \tau) \frac{dz}{dt} = - \text{diag} \gamma_i \begin{bmatrix} \nabla X_L(x, \lambda, \tau) \\ -G_q(x) \\ -P_{q}(x) \end{bmatrix}, \quad (19) \]

where \( \nabla Z(x, \tau) \) is the transpose of Jacobian of \( Z(x, \tau) \) and \( \gamma_i > 0, 1 \leq i \leq n + p \).

**Theorem 3.1.** Assume that \( f \) and \( G \) are twice continuously differentiable, let \((x^*, \lambda^*)\) be a KKT pair, and the Jacobian uniqueness conditions hold at \((x^*, \lambda^*)\). Then for any \( \tau \geq \max\{r, 1\} \), \( \nabla Z(x^*, \tau) \) is a nonsingular matrix.
Proof. Let \( z^T = (x^T, \lambda^T) \), note that if \((x^*, \lambda^*)\) is a local minimum-Lagrange multiplier pair, and satisfies the Jacobian uniqueness conditions, then the matrix \(Q\) is invertible, this was shown as the part of Theorem 2.3 that all eigenvalues of this matrix are strictly positive. Therefore, \(\nabla_z \Lambda(z^*, \tau)\) is nonsingular.

\textbf{Theorem 3.2}. Assume that \(f\) and \(G\) are twice continuously differentiable, let \((x^*, \lambda^*)\) be a KKT pair, and the Jacobian uniqueness conditions hold at \((x^*, \lambda^*)\). Then, for any \(\tau \geq \max\{\tilde{\tau}, 1\}\), the system of Eq. (19) is asymptotically stable at \(z^*\).

Proof. From the second order smoothness of \(L_e(x, \lambda, \tau)\) around \(z^*\), we have

\[
\Lambda(z, \tau) = \Lambda(z^*, \tau) + \nabla_z \Lambda(z^*, \tau) \delta(z) + \Theta(\delta(z)),
\]

where \(\delta(z) = z - z^*\), \(\Theta(\delta(z)) = o(\|\delta(z)\|^2)\).

Linearizing system of Eq. (19) at the point \(z^*\), we obtain

\[
\frac{d\delta(z)}{dt} = -\dot{Q}(z^*) \delta(z), \quad \dot{Q}(z^*) = \left(\nabla_z \Lambda(z^*, \tau)\right)^{-1} \text{diag}_{1 \leq i \leq n+p} \gamma_i \nabla_z \Lambda(z^*, \tau).
\]

Matrix \(\dot{Q}(z^*)\) is similar to matrix \(\text{diag}_{1 \leq i \leq n+p} \gamma_i\), therefore, they have the same eigenvalues \(\lambda_i = \gamma_i > 0\), \(1 \leq i \leq n + p\). According to Lyapunov linearization principle, we have that the equilibrium point \(z^*\) is asymptotically stable.

\textbf{3.1. Euler discrete schemes for system (19)}. Integrating system (19) by the Euler method, one obtains the iterate process

\[
z_{k+1} = z_k - h_k (\nabla_z \Lambda(z_k, \tau))^{-1} \text{diag}_{1 \leq i \leq n+p} \gamma_i \Lambda(z_k, \tau),
\]

(20)

where \(z^T = (x^T, \lambda^T)\), and \(h_k\) is a stepsize. The equation (20) provides a basic implementation of the Newton iteration under the assumption that \(\nabla_z \Lambda(z^*, \tau)\) is invertible.

\textbf{Theorem 3.3}. Assume that \(f\) and \(G\) are twice continuously differentiable, let \((x^*, \lambda^*)\) be a KKT pair, and the Jacobian uniqueness conditions hold at \((x^*, \lambda^*)\). Then, for any \(\tau \geq \max\{\tilde{\tau}, 1\}\), the discrete version \(\{z_k\}\) defined by (20) locally converges with at least linear rate to the point \(z^*\) if the stepsize \(h_k\) is fixed and \(h_k < 2/\max_{1 \leq i \leq n+p} \gamma_i\). If \(h_k = 1\) and \(\gamma_i = 1\), \(1 \leq i \leq n + p\), then the sequence \(\{z_k\}\) converges quadratically to \(z^*\).

Proof. The proof of this theorem is nearly identical to the proof of Theorem 2.4 and to the proof of convergence of Newton’s method.

4. Numerical experiments. In this section, we study the performances of our method through examples to test the theoretical results achieved and the efficiency of the system of Eq. (19). We implement the Runge-Kutta algorithm in Matlab and use Matlab function ode45 to solve the differential system of Eq. (19). All the programs are run on a laptop with Intel(R) Core(TM) i7 CPU M 620 @ 2.67GHz, 2.67GHz and 4 GB of RAM. Now we report the preliminary numerical results for the testing problems in Ref.[11], which are tabulated in Table 1, the comparison is provided between the optimal value \(f(x^*)\) in Eq. (19) and \(F(x^*)\) in Ref.[26].

\[
\min \quad f(x) = x_1 x_4 (x_1 + x_2 + x_3) + x_3
\]

\text{subject to} \quad x_1 x_2 x_3 x_4 - 25 \geq 0 \quad x_1^2 + x_2^2 + x_3^2 + x_4^2 - 40 = 0 \quad 1 \leq x_i \leq 5, \quad i = 1, \ldots, 4.
To construct the experiment, we set $\tau = 20$ and $\varepsilon = 0.01$, the approximate augmented Lagrangian of the problem takes the form (2) which is twice continuously differentiable, then we run the differential system of Eq.(19) for multiple times and these replications always show similar performances. In Problem 71, we choose the differentiable, then we run the differential system of Eq.(19) for multiple times and the optimal value of Problem 71 can be derived, which are $x^* = (1, 4.7430, 3.8212, 1.3794)^T$ and $f(x^*) = 17.014$ respectively. We plot the performances of the solution and the objective value, it respectively displays the local behaviour of the trajectories of $x$ and $z^T = (x^T, \lambda^T)$ in Figure 1, and the objective values and the cost function in Figure 2, where we denotes the cost function $S(z) = ||\nabla_z L_\varepsilon(x, \lambda, \tau)||$. Figure 1 shows that there is a substantially reduced variability in terms of the cost function value as the number of iterations increased, and at same time the objective value of Problem 71 converges to the optimal value very quickly, from the plot we see the local convergence of the differential system of Eq.(19).

$$\min \ f(x) = (x_1 - x_2)^2 + (x_2 + x_3 - 2)^2 + (x_4 - 1)^2 + (x_5 - 1)^2$$

s.t. $\begin{align*}
x_1 + 3x_2 &= 0 \\
x_3 + x_4 - 2x_5 &= 0 \\
x_2 - x_5 &= 0 \\
-10 \leq x_i \leq 10, & \quad i = 1, \ldots, 5.
\end{align*}$

(P53)

We choose the infeasible initial point $x_0 = (2, 2, 2, 2, 2)^T$ to see the performance of the differential system of Eq.(19), the optimal solution and the optimal value of Problem 53 can be obtained, which are $x^* = (-0.7674, 0.2558, 0.6279, -0.1163, 0.2558)^T$ and $f(x^*) = 4.0930$ respectively. Fig.3-4 displays the local behaviour of the trajectories of the differential system of Eq.(19), from the plot we see that the differential equation approach starts from the objective value of the infeasible initial point and converges to the optimal solution. Especially, we again observed that the cost function value by the differential system of Eq.(19) gets closer and closer to zero as the iterations gradually increased.

$$\min \ f(x) = x_1^2 + x_2^2 + x_1x_2 - 14x_1 - 16x_2 + (x_3 - 10)^2 + 4(x_4 - 5)^2 + (x_5 - 3)^2 + 2(x_6 - 1)^2 + 5x_7^2 + 7(x_8 - 11)^2 + 2(x_9 - 10)^2 + (x_{10} - 7)^2 + 45$$

s.t. $\begin{align*}
105 - 4x_1 - 5x_2 + 3x_7 - 9x_8 &\geq 0 \\
-10x_1 + 8x_2 + 17x_7 - 2x_8 &\geq 0 \\
8x_1 - 2x_2 - 5x_9 + 2x_{10} + 12 &\geq 0 \\
-3(x_1 - 2)^2 - 4(x_2 - 3)^2 - 2x_3^2 + 7x_4 + 120 &\geq 0 \\
-5x_2^2 - 8x_2 - (x_3 - 6)^2 + 2x_4 + 40 &\geq 0 \\
-0.5(x_1 - 8)^2 - 2(x_2 - 4)^2 - 3x_3^2 + x_6 + 30 &\geq 0 \\
-x_3^2 - 2(x_2 - 2)^2 + 2x_1x_2 - 14x_5 + 6x_6 &\geq 0 \\
3x_1 - 6x_2 - 12(x_9 - 8)^2 + 7x_{10} &\geq 0
\end{align*}$

(P113)

Similar to the examples, the numerical results of Problem 100 and Problem 113 are reported in Table 1, Figure 6-9 respectively displays the local behaviour of the trajectories of the solution and the objective value. We already reported examples to illustrate the theoretical results achieved and the efficiency of the differential system of Eq.(19), the numerical results given show that the differential equation approach enjoys better stability and higher precision.
\[
\begin{align*}
\min \quad & f(x) = (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2 + 10x_5^6 \\
& + 7x_6^2 + x_3^4 - 4x_6x_7 - 10x_6 - 8x_7 \\
\text{s.t.} \quad & 127 - 2x_1^2 - 3x_2^2 - x_3 - 4x_4^2 - 5x_5 \geq 0 \\
& 282 - 7x_1 - 3x_2 - 10x_3^2 - x_4 + x_5 \geq 0 \\
& 196 - 23x_1 - x_2^2 - 6x_5^2 + 8x_7 \geq 0 \\
& -4x_1^4 - x_2^4 + 3x_1x_2 - 2x_3^3 - 5x_6 + 11x_7 \geq 0
\end{align*}
\]
5. Conclusion. In this paper, we propose a differential equation approach based on the augmented Lagrangian function for constrained optimization problems. We construct the first derivative based and the second derivative based differential equation systems by using the approximate augmented Lagrangian function, which is twice continuously differentiable. Under a set of suitable conditions, we prove the asymptotic stability of the two differential systems and local convergence properties of their Euler discrete schemes, including the locally quadratic convergence rate of...
the discrete algorithm for second order derivatives based differential equation system. The numerical results show the differential equation approach works well on both stability and precision. The study on the dual theory of the nonlinear augmented Lagrangian and the numerical experiments for large-scale testing problems deserve our further research.
Figure 7. Performances of the variable $x$ and $z$ in Problem 113

Figure 8. Performances of the cost function and the objective function in Problem 113

Acknowledgments. The authors would like to thank all three referees and the associate editor for their constructive comments which eventually lead to a significant improvement in the revision.
Table 1. numerical results

| Test | n  | p  | q  | IT  | S(z)               | f(x*)          | F(x*)          |
|------|----|----|----|-----|--------------------|----------------|----------------|
| P.71 | 4  | 10 | 1  | 349 | 8.125604 × 10^{-10} | 17.014         | 17.0140173    |
| P.53 | 5  | 13 | 3  | 127 | 1.175666 × 10^{-11} | 4.0930        | 4.093023      |
| P.100| 7  | 4  | 0  | 967 | 3.829630 × 10^{-12} | 678.6796       | 680.6300573   |
| P.113| 10 | 8  | 0  | 991 | 2.452665 × 10^{-12} | 24.3062        | 24.306291     |

REFERENCES

[1] K. J. Arrow and L. Hurwicz, Reduction of constrained maxima to saddle point problems, in Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability, (eds. J. Neyman), University of California Press, Berkeley, (1956), 1–20.
[2] D. P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods, in Computer Science and Applied Mathematics, Academic Press Inc, New York, 1982.
[3] D. P. Bertsekas, Nonlinear Programming, Athena Scientific, 1999.
[4] Yu. G. Evtushenko, Numerical Optimization Techniques, Optimization Software, New York, 1985.
[5] Yu. G. Evtushenko and V. G. Zhadan, Barrier-projective methods for nonlinear programming, Comp. Math. Math. Phys., 34 (1994), 579–590.
[6] Yu. G. Evtushenko and V. G. Zhadan, Stable barrier-projection and barrier-Newton methods in nonlinear programming, Comput. Optim. Appl., 3 (1994), 289–303.
[7] Yu. G. Evtushenko and V. G. Zhadan, Stable barrier-projection and barrier-Newton methods for linear and nonlinear programming, in Algorithms for Continuous Optimization (eds. E. Spedicato), Kulwer Academic Publishers, 434 (1994), 255–285.
[8] Yu. G. Evtushenko, Two numerical methods of solving nonlinear programming problems, Sov. Math. Dokl., 15 (1974), 420–423.
[9] A. V. Fiacco and G. P. McCormick, Sequential Unconstrained Minimization Techniques, in Nonlinear Programming, John Wiely, New York, 1968.
[10] M. Hestenes, Multiplier and gradient methods, J. Optim. Theor. Appl., 4 (1969), 303–320.
[11] W. Hock and K. Schittkowski, Test Examples for Nonlinear Programming Codes, in Lecture Notes Economics and Mathematical Systems, Springer-Verlag, Berlin Heidelberg-New York, 1981.
[12] X. X. Huang, K. L. Teo and X. Q. Yang, Approximate augmented Lagrangian functions and nonlinear semidefinite programs, Acta Mathematica Sinica (English Series), 22 (2006), 1283–1296.
[13] A. Ioffe, Necessary and sufficient conditions for a local minimum 3: Second-order conditions and augmented duality, SIAM J. Control Optim., 17 (1979), 266–288.
[14] L. Jin, L. W. Zhang and X. T. Xiao, Two differential equation systems for equality-constrained optimization, Appl. Math. Comput., 190 (2007), 1030–1039.
[15] L. Jin, L. W. Zhang and X. T. Xiao, Two differential equation systems for inequality constrained optimization, Appl. Math. Comput., 188 (2007), 1334–1343.
[16] L. Jin, A stable differential equation approach for inequality constrained optimization problems, Appl. Math. Comput., 206 (2008), 186–192.
[17] L. Jin, H. X. Yu and Z. S. Liu, Differential systems for constrained optimization via a nonlinear augmented Lagrangian, Appl. Math. Comput., 235 (2014), 482–491.
[18] P. Q. Pan, New ODE methods for equality constrained optimization (1)-equations, J. Comput. Math., 10 (1992), 77–92.
[19] M. J. D. Powell, A method for non-linear constraints in minimizations problems in optimization, in (eds. R. Fletcher), (1969), 283–298.
[20] R. T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, Math. Oper. Res., 1 (1976), 97–116.
[21] R. T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim., 14 (1976), 877–898.
[22] R. T. Rockafellar, Augmented Lagrange multiplier functions and duality in nonconvex programming, *SIAM J. Control*, 12 (1974), 268–285.

[23] R. T. Rockafellar, Lagrange multipliers and optimality, *SIAM Rev.*, 35 (1993), 183–238.

[24] H. Yamadhiita, A differential equation approach to nonlinear programming, *Math. Prog.*, 18 (1980), 155–168.

[25] L. W. Zhang, A modified version to the differential system of Evtushenko and Zhanda for solving nonlinear programming, in *Numerical Linear Algebra and Optimization* (eds. Ya-xiang Yuan), Science Press, (1999), 161–168.

[26] L. W. Zhang, Q. Li and X. Zhang, Two differential systems for solving nonlinear programming problems, *OR Trans.*, 4 (2000), 33–46.

Received May 2018; 1st revision January 2019; 2nd revision February 2019.

E-mail address: jinli@zjou.edu.cn
E-mail address: huanghy@lsec.cc.ac.cn