Advances in Applied Mathematics 应用数学进展, **2024, 13(12), 5225-5236**
Published Online December 2024 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2024.1312504

一类解决可分离优化问题动力系统的渐进行为分析

董兴港

湖南大学电气与信息工程学院，湖南 长沙

收稿日期：2024年11月24日；录用日期：2024年12月18日；发布日期：2024年12月26日

摘 要

本文提出了一类关于二阶原始变量和一阶对偶变量组合的惯性原对偶动力系统，利用该惯性原对偶动力系统来解决用来解决一类可分离的优化问题。我们的方法与传统的优化算法不同，该方法利用连续性的方程来解决最优化问题，在具体的应用过程当中，能够使最目标函数快速且渐近的收敛到最优解。我们主要得到两个重要理论：一方面在解决凸问题时，得到了该系统解的存在唯一性；另一方面，基于Lyapunov分析方法的应用，在得到目标函数收敛到最优解的同时，还得到了其收敛的速率为$\mathcal{O}(\frac{1}{t^2})$，这也是已知的关于动力系统最好的收敛性结果。

关键词

惯性动力系统，可分离优化，快速收敛性

A Gradual Asymptotic Behavior Analysis of Dynamical Systems for Solving Separable Optimization Problems

Xinggang Dong

College of Electrical and Information Engineering, Hunan University, Changsha Hunan

Nov. 24^{th}, 2024; accepted: Dec. 18^{th}, 2024; published: Dec. 26^{th}, 2024

文章引用: 董兴港. 一类解决可分离优化问题动力系统的渐进行为分析[J]. 应用数学进展, 2024, 13(12): 5225-5236.
DOI: 10.12677/aam.2024.1312504

Abstract

This paper introduces a class of inertial primal-dual dynamical systems involving second-order primal variables and first-order dual variables, which are used to solve a class of separable optimization problems. Our approach differs from traditional optimization algorithms in that it utilizes continuous equations to solve optimization problems, enabling the objective function to converge quickly and asymptotically to the optimal solution in practical applications. We have obtained two important theoretical results: on the one hand, the system has a unique solution; on the other hand, based on the application of the Lyapunov analysis method, we have not only obtained the convergence of the objective function to the optimal solution but also achieved a convergence rate of $\mathcal{O}(\frac{1}{t^2})$, which is also the best known convergence result for dynamical systems.

Keywords

Inertial Dynamical Systems, Separable Optimization, Fast Convergence

1. 引言

在本文中，我们研究了以下二阶原始变量和一阶对偶变量组合的惯性原对偶动力系统：

$$\begin{cases} \ddot{x}(t)+\dfrac{\alpha}{t}\dot{x}(t)=-\nabla g(x)-A^T\lambda(t)-A^T(Ax(t)+By(t)-b), \\ \ddot{y}(t)+\dfrac{\alpha}{t}\dot{y}(t)=-\nabla h(y)-B^T\lambda(t)-B^T(Ax(t)+By(t)-b), \\ \dot{\lambda}(t)=t\big(A(x(t)+By(t)-b)\big)+\dfrac{t^2}{\alpha-1}(A\dot{x}(t)+B\dot{y}(t)). \end{cases} \tag{1.1}$$

其中，$g:\mathcal{X}\to\mathbb{R}$ 和$h:\mathcal{Y}\to\mathbb{R}$ 是两个光滑凸函数，∇g 和∇h 分别是其梯度函数，且满足Lipschitz 连续性，其Lipschitz 常数分别为L_1 和L_2. $A:\mathcal{X}\to\mathcal{H}$、$B:\mathcal{Y}\to\mathcal{H}$ 是两个连续线性算子，$\mathcal{X}$、$\mathcal{Y}$ 和$\mathcal{H}$ 是实Hilbert 空间，$t\geq t_0>0$、$\alpha\geq 3$ 是正实数。令$A^T:\mathcal{H}\to\mathcal{X}$ 和$B^T:\mathcal{H}\to\mathcal{Y}$ 分别是A 和B

的共轭算子。在本文中，始终假设以下条件严格成立：

H : ∇g 和∇h 是Lipschitz 连续的，其Lipschitz 常数分别为L_1 和L_2.

近年来，惯性动力系统方法已广泛应用于解决如下无约束平滑优化问题：

$$\min \Phi(x), \tag{1.2}$$

其中，$\Phi(x)$ 是一个平滑函数。梯度下降法是广泛使用的一阶方法，用于最小化问题(1.2) 中的目标函数。其通过以下方式迭代更新解：

$$x^{k+1} = x^k - \alpha \nabla \Phi(x^k).$$

通过简单的计算，容易得到梯度下降法的梯度系统，其具体形式如下：

$$\dot{x}(t) = -\nabla \Phi(x(t)). \tag{1.3}$$

普遍认为，梯度下降法在凸函数上的收敛速度是$\mathcal{O}(\frac{1}{k})$. 值得注意的是，微分方程(1.3) 的收敛速度也与梯度下降法一致，为$\mathcal{O}(\frac{1}{t})$.

后来，为了更好地解决无约束优化问题(1.2)，一些学者提出了惯性二阶动力系统方法。二阶动力系统方法相比一阶方法具有更好的收敛性。在这些二阶动力系统方法中，Polyak [1,2] 提出的带摩擦的重球系统，是一个著名且经典的方法。重球系统可以表示如下：

$$\ddot{x}(t) + \gamma \dot{x}(t) + \nabla \Phi(x) = 0, \tag{1.4}$$

其中，$\gamma > 0$ 是一个阻尼系数。通过引入阻尼项$\gamma \dot{x}(t)$ ，他证明了相较于一阶梯度系统，目标函数$\Phi(x(t))$ 沿着轨迹$x(t)$ 的收敛速度得到了改善。将γ 替换为函数$\gamma(t)$，我们可以将系统(1.4)重写为：

$$\ddot{x}(t) + \gamma(t)\dot{x}(t) + \nabla \Phi(x) = 0. \qquad (IGS_\gamma).$$

最近，系统(IGS_γ) 引起了许多研究者的关注，因为它与许多快速数值优化方法有着密切的联系。如果选择某些特定的函数作为$\gamma(t)$，则可以得到一些具体的收敛速率和结果。

2016年，Su等学者 [3] 通过选择(IGS_γ) 中的$\gamma(t) = \frac{\alpha}{t}$，研究了以下系统：

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla \Phi(x) + \varepsilon(t)x(t) = 0, \qquad (AVD_{\alpha,\varepsilon}).$$

如果$\Phi(x)$ 是一个连续可微的凸函数，并且其梯度是Lipschitz 连续的，他们证明了系统(AVD_α) 可以看作是Nesterov 加速梯度算法 [4] 的一个连续版本，其中选择$\alpha = 3$. 对于$\alpha \geq 3$，他们还证明了目标函数值的收敛速度为$\mathcal{O}(\frac{1}{t^2})$.

上述惯性动力系统主要应用于求解无约束优化问题，对于线性约束等约束优化问题，我们不能直接使用这些系统。因此，基于原对偶框架的二阶动力系统方法被研究用来解决结构化的约束

优化问题。2023 年，Zeng、Lei 和Chen 提出了以下原对偶动力系统 [5]：

$$\begin{cases} \ddot{x}(t) + \dfrac{\alpha}{t}\dot{x}(t) = -\nabla f(x(t)) - A^T(\lambda(t) + \beta t\dot{\lambda}(t)) - A^T(Ax(t) - b), \\ \ddot{\lambda}(t) + \dfrac{\alpha}{t}\dot{\lambda}(t) = A(x(t) + \beta t\dot{x}(t)) - b, \end{cases}$$

用来求解线性约束问题：

$$\begin{aligned} &\min \quad f(x) \\ &s.t. Ax = b. \end{aligned}$$

其中，f 是一个连续可微的凸函数，A 是一个连续线性算子。他们分析了上述系统的渐近行为，并建立了增广拉格朗日函数$\mathcal{L}$ 的收敛速率。

此外，He、Hu 和Fang 提出了一个“二阶原始”+ “一阶对偶”动力系统 [6]：

$$\begin{cases} \ddot{x}(t) + \gamma\dot{x}(t) = -\beta(t)(\nabla f(x(t)) + A^T\lambda(t) + \sigma A^T(Ax(t) - b)), \\ \dot{\lambda}(t) = \beta(t)(A(x(t) + \delta\dot{x}(t)) - b), \end{cases}$$

其中$\gamma, \delta > 0$，且$\beta : [t_0, +\infty) \to (0, +\infty)$ 是一个时间缩放系数，在分析收敛性质时起着至关重要的作用。类似地，为了建立原对偶动力系统轨迹的强收敛性，Zhu、Hu 和Fang 提出了以下原对偶动力系统 [7]：

$$\begin{cases} \ddot{x}(t) + \gamma\dot{x}(t) = -\nabla f(x(t)) - A^T\lambda(t) - \rho A^T(Ax(t) - b) - \varepsilon(t)x(t), \\ \dot{\lambda}(t) = t\left(A(x(t) + \dfrac{t}{\alpha - 1}\dot{x}(t)) - b\right), \end{cases}$$

他们证明了轨迹$x(t)$ 强收敛到最小范数元素。有关原对偶动力系统的更多详细信息，建议读者参考 [8–10] 及其中的参考文献。

一个简单但重要的可分离约束优化问题是以下线性等式约束问题：

$$\begin{aligned} &\min f(x) + g(y), \\ &s.t. Ax + By = b. \end{aligned} \tag{1.5}$$

其中$f : \mathcal{X} \to \mathbb{R}$ 和$g : \mathcal{Y} \to \mathbb{R}$ 是两个平滑的凸函数，∇g 和∇h 分别是具有模数L_1 和L_2 的Lipschitz 连续函数。$A : \mathcal{X} \to \mathcal{H}$，$B : \mathcal{Y} \to \mathcal{H}$ 是两个连续线性算子，$x \in \mathcal{X}$，$y \in \mathcal{Y}$，$b \in \mathcal{H}$，其中$\mathcal{X}$，$\mathcal{Y}$ 和$\mathcal{H}$ 是实Hilbert 空间。近年来，问题(1.5) 在许多应用领域中出现，包括机器学习 [11]、图像处理 [12,13] 等。

本文研究了一类原对偶动力系统(1.1)，用于求解以下问题：

$$\begin{aligned} &\min g(x) + h(y), \\ &s.t. Ax + By = b. \end{aligned} \tag{1.6}$$

本工作的主要贡献可以总结如下：

- 提出了一种新的“二阶原始”+“一阶对偶”的惯性原对偶动力系统，用于求解双块可分离优化问题，并且证明了解的存在唯一性;
- 通过构造适当的能量函数，不仅建立了沿着轨迹$(x(t), y(t))$ 收敛到目标函数的结果，还证明了增广拉格朗日函数$\mathcal{L}_\sigma$ 和拉格朗日函数$\mathcal{L}$ 的收敛速率。

2. 准备知识

在本节中，我们将介绍本文中使用的一些符号和预备知识。首先，引入与问题(1.6) 相关的一些定义。

由于问题(1.6) 是一个约束优化问题，定义与问题(1.6) 相关的拉格朗日函数$\mathcal{L} : \mathcal{X} \times \mathcal{Y} \times \mathcal{H} \to \mathbb{R}$ 为：

$$\mathcal{L}(x, y, \lambda) = g(x) + h(y) + \langle \lambda, Ax + By - b \rangle, \tag{2.1}$$

其中，$\lambda > 0$ 是拉格朗日乘子。而增广拉格朗日函数$\mathcal{L}_\sigma : \mathcal{X} \times \mathcal{Y} \times \mathcal{H} \to \mathbb{R}$ 定义为：

$$\mathcal{L}_\sigma(x, y, \lambda) = g(x) + h(y) + \langle \lambda, Ax + By - b \rangle + \frac{1}{2}\|Ax + By - b\|^2. \tag{2.2}$$

我们将问题(1.6) 的KKT 点集记为Ω，即当且仅当$(x^*, y^*, \lambda^*) \in \Omega$ 满足以下条件时：

$$\begin{cases} -A^T\lambda^* = \nabla f(x^*), \\ -B^T\lambda^* = \nabla g(y^*), \\ Ax^* + By^* - b = 0. \end{cases} \tag{2.3}$$

显然，$\Omega \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{H}$。在本文中，我们始终假设$\Omega \neq \emptyset$。

根据拉格朗日函数$\mathcal{L}$ 和增广拉格朗日函数$\mathcal{L}_\sigma$的定义，我们可以得到以下等式：

$$\begin{cases} -A^T\lambda^* = \nabla f(x^*), \\ -B^T\lambda^* = \nabla g(y^*), \\ Ax^* + By^* - b = 0. \end{cases} \Leftrightarrow \begin{cases} -A^T\lambda^* = \nabla f(x^*) + A^T(Ax^* + By^* - b), \\ -B^T\lambda^* = \nabla g(y^*) + B^T(Ax^* + By^* - b), \\ Ax^* + By^* - b = 0. \end{cases} \tag{2.4}$$

这等价于：

$$(x^*, y^*, \lambda^*) \in \Omega \Leftrightarrow \begin{cases} \nabla_x\mathcal{L}(x^*, y^*, \lambda^*) = 0, \\ \nabla_y\mathcal{L}(x^*, y^*, \lambda^*) = 0, \\ \nabla_\lambda\mathcal{L}(x^*, y^*, \lambda^*) = 0. \end{cases} \Leftrightarrow \begin{cases} \nabla_x\mathcal{L}_\sigma(x^*, y^*, \lambda^*) = 0, \\ \nabla_y\mathcal{L}_\sigma(x^*, y^*, \lambda^*) = 0, \\ \nabla_\lambda\mathcal{L}_\sigma(x^*, y^*, \lambda^*) = 0. \end{cases}$$

根据Ω 的定义，我们可以立即得出以下结论：若$(x^*, y^*, \lambda^*) \in \Omega$，则它是拉格朗日函数$\mathcal{L}$ 和增

广拉格朗日函数$\mathcal{L}_\sigma$ 的鞍点。

具体而言，对于拉格朗日函数$\mathcal{L}$，满足以下不等式：

$$\mathcal{L}(x^*, y^*, \lambda) \le \mathcal{L}(x^*, y^*, \lambda^*) \le \mathcal{L}(x, y, \lambda^*), \quad \forall (x, y, \lambda) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{H}. \tag{2.5}$$

类似地，对于增广拉格朗日函数$\mathcal{L}_\sigma$，有：

$$\mathcal{L}_\sigma(x^*, y^*, \lambda) \le \mathcal{L}_\sigma(x^*, y^*, \lambda^*) \le \mathcal{L}_\sigma(x, y, \lambda^*), \quad \forall (x, y, \lambda) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{H}. \tag{2.6}$$

从(2.5) 和(2.6) 可以得出，增广拉格朗日函数$\mathcal{L}_\sigma(x, y, \lambda)$ 和拉格朗日函数$\mathcal{L}(x, y, \lambda)$ 具有相同的鞍点。

本文讨论的问题定义在希尔伯特空间$\mathcal{H}$ 中，希尔伯特空间中的内积用$\langle \cdot, \cdot \rangle$ 表示，范数记作$\| \cdot \|$。

对于实值的凸且可微的函数$g : \mathcal{H} \to \mathbb{R}$，若其梯度$\nabla g$ 满足L_g- 利普希茨连续性，则有：

$$\|\nabla g(x) - \nabla g(y)\| \le L_g \|x - y\|, \quad \forall x, y \in \mathcal{H}.$$

如果函数g 是σ-强凸的，意味着$g(\cdot) - \frac{\sigma}{2}\| \cdot \|^2$ 是凸的，其中$\sigma > 0$。此外，如果g 是连续可微的，那么对于任意的$x, y \in \mathcal{H}$，都有：

$$g(y) \ge g(x) + \langle \nabla g(x), y - x \rangle + \frac{\sigma}{2}\|y - x\|^2.$$

由此，我们可以进一步分析两种函数在优化中的角色，尤其是在动力系统轨迹收敛性和优化性能评估中。最后，我们将引入一个重要的引理，它在系统(1.1) 的收敛性分析中起到关键作用。

引理2.1. *[14, 引理A.3] 假设$\delta > 0$且$f \in L^1(\delta, +\infty)$ 是一个非负连续函数。进一步假设$\psi : [\delta, +\infty) \to [0, +\infty)$是一个非递减的函数且$\lim\limits_{t \to +\infty} \psi(t) = +\infty$. 可得$\lim\limits_{t \to +\infty} \frac{1}{\psi(t)} \int_\delta^t \psi(s) f(s) ds = 0$.*

3. 全局解的存在性与唯一性

在本节中，我们将展示惯性原始对偶动力学系统(1.1) 的全局解的存在性和唯一性。从方程(1.1) 中，通过重新排列，可以得到：

$$\begin{cases} \ddot{W}(t) + \dfrac{\alpha}{t}\dot{W}(t) = -\nabla H(W(t)) - C^T\lambda(t) - C^T(CW(t) - b), \\ \dot{\lambda}(t) = t(CW(t) - b) + \dfrac{t^2}{\alpha - 1} C\dot{W}(t), \end{cases} \tag{3.1}$$

其中$W(t) = (x(t), y(t))^T \in \mathcal{W}$，且$\mathcal{W} = \mathcal{X} \times \mathcal{Y}$，$C = (A, B)$，且$H(W(t)) = (f(x(t)), g(y(t)))^T$.

在证明系统(1.1) 解的存在性和唯一性之前，我们首先给出一个全局解的定义。

定义3.1. 称$W:[t_0,+\infty)\to\mathcal{X}\times\mathcal{Y}$，$\lambda:[t_0,+\infty)\to\mathcal{H}_2$ 为系统*(3.1)* 的全局解，当满足以下性质：

(i) $W,\dot{W}:[t_0,+\infty)\to\mathcal{W}$ 且$\lambda:[t_0,+\infty)\to\mathcal{H}$ 在局部是绝对连续的，换句话说，对于每个区间$[t_0,T]$，$T<+\infty$，它们在该区间内是绝对连续的；

(ii)

$$\begin{cases}\ddot{W}(t)+\dfrac{\alpha}{t}\dot{W}(t)=-\nabla H(W(t))-C^T\lambda(t)-C^T(CW(t)-b),\\ \dot{\lambda}(t)=t(CW(t)-b)+\dfrac{t^2}{\alpha-1}C\dot{W}(t),\end{cases}$$

对于几乎所有的$t\geq t_0$；

(iii) $(W(t_0),\dot{W}(t_0),\lambda(t_0))=(W_0,\dot{W}_0,\lambda_0)$.

现在，准备基于柯西-利普希茨-皮卡德定理（Cauchy-Lipschitz-Picard theorem）证明二阶原始变量和一阶对偶变量组合的惯性动力学系统(1.1) 的全局解的存在性和唯一性。

定理3.2. 对于任意初始点$W_0,\dot{W}_0\in\mathcal{W}$ 和$\lambda_0\in\mathcal{H}$，足动力系统*(1.1)* 足动力系统*(1.1)* 存在一个唯一的全局解。

证明 定义$G(t)=(W(t),\dot{W}(t),\lambda(t))$，并且定义$F:[t_0,T]\times\mathcal{W}\times\mathcal{W}\times\mathcal{H}$ 为

$$F(t,u,v,m)=\big(v,\mu_1(t,u,v,m),\mu_2(t,u,v,m)\big) \tag{3.2}$$

其中

$$\begin{cases}\mu_1(t,u,v,m)=-\dfrac{\alpha}{t}v-\nabla H(u)-C^Tm-C^T(Cu-b),\\ \mu_2(t,u,v,m)=t(Cu-b)+\dfrac{t^2}{\alpha-1}Cv.\end{cases} \tag{3.3}$$

显然，从(3.1)、(3.2) 和$G(t)$ 的定义可以看出，(3.1) 可以重写为一个一阶动力系统，具体为

$$\begin{cases}\dot{G}(t)=F(t,G(t))=F(t,W(t),\dot{W}(t),\lambda(t)),\\ G(t_0)=(W_0,\dot{W}_0,\lambda_0).\end{cases} \tag{3.4}$$

我们将首先证明，对于每个$t\geq t_0$，$F(t,\cdot,\cdot,\cdot)$ 是$L(t)$-Lipschitz 连续的。具体地，对于每对$(u_1,v_1,m_1),(u_2,v_2,m_2)\in\mathcal{W}\times\mathcal{W}\times\mathcal{H}$，根据(3.2)，有：

$\| F(t,u_1,v_1,m_1)-F(t,v_2,u_2,m_2)\|$
$=\sqrt{\| v_1-v_2\|^2+\|\mu_1(t,u_1,v_1,m_1)-\mu_1(t,u_2,v_2,m_2)\|^2+\|\mu_2(t,u_1,v_1,m_1)-\mu_2(t,u_2,v_2,m_2)\|^2}$. 通过计算有

$$\begin{aligned}&\mu_1(t,u_1,v_1,m_1)-\mu_1(t,u_2,v_2,m_2)\\ =&\frac{\alpha}{t}(v_2-v_1)+\nabla H(u_2)-\nabla H(u_1)+C^T(m_2-m_1)+C^TC(u_2-u_1),\end{aligned} \tag{3.5}$$

和

$$\begin{aligned}&\mu_2(u_1,v_1,m_1)-\mu_2(u_2,v_2,m_2)\\=&tC(u_1-u_2)+\frac{t^2}{\alpha-1}C(v_1-v_2).\end{aligned}\tag{3.6}$$

根据假设条件∇f 和∇g 是局部Lipschitz 连续的，可以得到：

$$\parallel \nabla H(u_2)-\nabla H(u_1)\| \leq\parallel L(u_2-u_1)\parallel\leq (L_H)\|u_2-u_1\|,\tag{3.7}$$

其中

$$L=\begin{pmatrix}L_1 & 0\\ 0 & L_2\end{pmatrix},\ \ L_H=L_1+L_2.$$

通过基本不等式和上述公式，可以得到：

$$\begin{aligned}&\parallel F(t,u_1,v_1,m_1)-F(t,u_2,v_2,m_2)\parallel\\ \leq&\sqrt{h_1(t)\parallel u_1-u_2\parallel^2+h_2(t)\|v_1-v_2\|^2+h_3(t)\|m_1-m_2\|^2}\end{aligned}\tag{3.8}$$

其中

$$\begin{cases}h_1(t)=5L_H^2+5\|C^TC\|^2+2\|C^TC\|t^2,\\ h_2(t)=1+\dfrac{5\alpha^2}{t^2}+\dfrac{4t^4}{(\alpha-1)^2}\|C^TC\|,\\ h_3(t)=5\|C^TC\|+1.\end{cases}$$

进一步可得

$$\begin{aligned}&\parallel F(t,u_1,v_1,m_1)-F(t,u_2,v_2,m_2)\parallel\\ \leq&\sqrt{h_1(t)\parallel u_1-u_2\parallel^2+h_2(t)\|v_1-v_2\|^2+h_3(t)\|m_1-m_2\|^2}\\ \leq&\sqrt{h_1(t)+h_2(t)+h_3(t)}\sqrt{\|u_1-u_2\|^2+\|v_1-v_2\|^2+\|m_1-m_2\|^2}\\ =&\sqrt{h_1(t)+h_2(t)+h_3(t)}\|(u_1,v_1,m_1)-(u_2,v_2,m_2)\|\end{aligned}\tag{3.9}$$

令$L(t)=\sqrt{h_1(t)+h_2(t)+h_3(t)}$, 然后有

$$\begin{aligned}&\parallel F(t,u_1,v_1,m_1)-F(t,u_2,v_2,m_2)\parallel\\ \leq& L(t)\|(u_1,v_1,m_1)-(u_2,v_2,m_2)\|\end{aligned}\tag{3.10}$$

因此，$F(t,\cdot,\cdot,\cdot)$ 对于每个$t\geq t_0$ 都是$L(t)$-Lipschitz 连续的。因此容易看到，$L(t)$ 在区间$[t_0,T]$ 上是可积的，因此$L(\cdot)\in L^1_{\text{loc}}([t_0,+\infty))$.

接下来，将证明对于所有$u,v\in\mathcal{W}$ 和$m\in\mathcal{H}$，$F(\cdot,u,v,m)\in L^1_{\text{loc}}([t_0,+\infty),\mathcal{W}\times\mathcal{W}\times\mathcal{H})$. 取

任意的$u, v \in \mathcal{W}$ 和$m \in \mathcal{H}$，根据F 的定义，对于$t_0 < T < +\infty$，得到

$$
\begin{aligned}
&\int_{t_0}^{T} \|F(t,u,v,m)\|dt \\
=&\int_{t_0}^{T} \sqrt{\|v\|^2+\|g_1\|^2+\|g_2\|^2}dt \\
\leq&\int_{t_0}^{T} \sqrt{\tilde{h}_1\frac{1}{t^2}+\tilde{h}_2 t^4++\tilde{h}_3 t^2+\tilde{h}_4}dt \\
\leq&\sqrt{\tilde{h}_1+\tilde{h}_2+\tilde{h}_3+\tilde{h}_4}\int_{t_0}^{T} \sqrt{\frac{1}{t^2}+t^4+t^2+1}dt
\end{aligned}
\tag{3.11}
$$

其中$\tilde{h}_1 = 5\alpha^2\|v\|^2$, $\tilde{h}_2 = \frac{2\|C^TC\|\|v\|^2}{(\alpha-1)^2}$, $\tilde{h}_3 = 2\|Cu-b\|^2$, $\tilde{h}_4 = \|v\|^2 + 5(\|\nabla H(u)\|^2 + \|C^TC\|\|m\|^2 + \|c^TC\|\|Cu-b\|^2)$.

因此，基于(3.11) 以及$\frac{1}{t}$、t^4、t^2 和$\varepsilon(t)$ 对于任何$t \geq t_0$ 都是连续的这一事实，可以得出：

$$
F(\cdot,u,v,m,n) \in L^1_{loc}([t_0,+\infty), \mathcal{H}_1 \times \mathcal{W} \times \mathcal{W} \times \mathcal{H})
$$

结合这个关系式与(3.10) 以及结果$L(\cdot) \in L^1_{loc}([t_0,+\infty))$，再利用Cauchy-Lipschitz-Picard 定理，我们可以得出系统(1.1) 存在唯一的全局解，这也意味着由于∇H 的Lipschitz 连续性，系统(1.1) 存在唯一的全局解。

4. 系统的渐近行为分析

在本节中，我们将分析(1.1) 系统轨迹$(x(t), y(t))$ 的收敛行为。为了分析该原对偶动力系统(1.1) 的收敛行为，分为两个定理进行研究。

定理4.1. *假设$(x(t), y(t), \lambda(t))$ 是系统(1.1) 的全局解，且$(x^*, y^*, \lambda^*) \in \Omega$. 令$\varepsilon : [t_0,+\infty) \to (0,+\infty)$ 为一个非增函数，满足$\int_{t_0}^{+\infty}\frac{\varepsilon(t)}{t}dt < +\infty$. 则以下结论成立:*

(i) $\lim\limits_{t\to+\infty} \mathcal{L}_\sigma(x(t), y(t), \lambda^*) - \mathcal{L}_\sigma(x^*, y^*, \lambda^*) = 0$;

(ii) $\lim\limits_{t\to+\infty} \mathcal{L}(x(t), y(t), \lambda^*) - \mathcal{L}(x^*, y^*, \lambda^*) = 0$.

证明 考虑一个能量函数$\tilde{E}(t) : [t_0, T) \to [0,+\infty)$，它由以下公式给出：

$$
\begin{aligned}
\tilde{E}(t) =&\mathcal{L}_\sigma(x(t), y(t), \lambda^*) - \mathcal{L}_\sigma(x^*, y^*, \lambda^*) + \frac{1}{2}\|\frac{\alpha-1}{t}(x(t)-x^*)+\dot{x}(t)\|^2 \\
&+ \frac{1}{2}\|\frac{\alpha-1}{t}(y(t)-y^*)+\dot{y}(t)\|^2 + \frac{\alpha-1}{2t^2}\|\lambda(t)-\lambda^*\|^2.
\end{aligned}
\tag{4.1}
$$

根据动力系统(1.1)的定义，可得

$$
\begin{aligned}
\dot{\tilde{E}}(t) = & -\frac{\alpha-1}{t}\langle x(t)-x^*, \nabla g(x(t))\rangle - \frac{2(\alpha-1)}{t^2}\langle x(t)-x^*, \dot{x}(t)\rangle \\
& -\frac{\alpha-1}{t}\langle y(t)-y^*, \nabla h(y(t))\rangle - \frac{2(\alpha-1)}{t^2}\langle y(t)-y^*, \dot{y}(t)\rangle \\
& -\frac{1}{t}\|\dot{x}(t)\|^2 - \frac{(\alpha-1)^2}{t^3}\|x(t)-x^*\|^2 - \frac{1}{t}\|\dot{y}(t)\|^2 - \frac{(\alpha-1)^2}{t^3}\|y(t)-y^*\|^2 \\
& -\frac{(\alpha-1)^2}{t^3}\|\lambda(t)-\lambda^*\|^2 - \frac{\alpha-1}{t}\|Ax(t)+By(t)-b\|^2 - \frac{\alpha-1}{t}\langle \lambda^*, Ax(t)+By(t)-b\rangle.
\end{aligned} \tag{4.2}
$$

进一步可得

$$
\begin{aligned}
\dot{\tilde{E}}(t) \le & -\frac{\alpha-1}{t}\langle x(t)-x^*, \nabla g(x(t))\rangle - \frac{2(\alpha-1)}{t^2}\langle x(t)-x^*, \dot{x}(t)\rangle \\
& -\frac{\alpha-1}{t}\langle y(t)-y^*, \nabla h(y(t))\rangle - \frac{2(\alpha-1)}{t^2}\langle y(t)-y^*, \dot{y}(t)\rangle \\
& -\frac{1}{t}\|\dot{x}(t)\|^2 - \frac{(\alpha-1)^2}{t^3}\|x(t)-x^*\|^2 - \frac{1}{t}\|\dot{y}(t)\|^2 - \frac{(\alpha-1)^2}{t^3}\|y(t)-y^*\|^2 \\
& -\frac{(\alpha-1)^2}{t^3}\|\lambda(t)-\lambda^*\|^2 - \frac{\alpha-1}{t}\|Ax(t)+By(t)-b\|^2 - \frac{\alpha-1}{t}\langle \lambda^*, Ax(t)+By(t)-b\rangle.
\end{aligned} \tag{4.3}
$$

利用上述不等式和(4.1) 中对$\tilde{E}(t)$ 的定义，结合$\alpha > 3$以及g, h的凸性，可以得到：

$$
\frac{2}{t}\tilde{E}(t) + \dot{\tilde{E}}(t) \le 0. \tag{4.4}
$$

通过求积分，进一步有

$$
t^2\tilde{E}(t) \le t_0^2\tilde{E}(t_0).
$$

然后可得

$$
\tilde{E}(t) \le \frac{t_0^2\tilde{E}(t_0)}{t^2}. \tag{4.5}
$$

结合(4.5),从而可的

$$
\lim_{t\to+\infty}\tilde{E}(t) = 0. \tag{4.6}
$$

利用上述关系和(4.1) 中对$\tilde{E}(t)$ 的定义，再结合关系$\mathcal{L}_\sigma(x(t), y(t), \lambda^*) - \mathcal{L}_\sigma(x^*, y^*, \lambda^*) \ge \mathcal{L}(x(t), y(t), \lambda^*) - \mathcal{L}(x^*, y^*, \lambda^*) \ge 0$，可以得出：

$$
\lim_{t\to+\infty}\mathcal{L}_\sigma(x(t), y(t), \lambda^*) - \mathcal{L}_\sigma(x^*, y^*, \lambda^*) = 0, \quad \text{and} \quad \lim_{t\to+\infty}\mathcal{L}(x(t), y(t), \lambda^*) - \mathcal{L}(x^*, y^*, \lambda^*) = 0.
$$

定理4.2. *设$(x^*, y^*, \lambda^*) \in \Omega$ 且$(x(t), y(t), \lambda(t))$ 是动态系统(1.1) 的全局解。则以下结论成立：*

(i) $\mathcal{L}_\sigma(x(t), y(t), \lambda^*) - \mathcal{L}_\sigma(x^*, y^*, \lambda^*) = \mathcal{O}(\frac{1}{t^2})$;

(ii) $\mathcal{L}(x(t), y(t), \lambda^*) - \mathcal{L}(x^*, y^*, \lambda^*) = \mathcal{O}(\frac{1}{t^2})$.

证明 根据关系(4.5)，易得

$$t^2\tilde{E}(t) \leq t_0^2\tilde{E}(t_0) < +\infty. \tag{4.7}$$

结合(4.1) 中对$\tilde{E}(t)$ 的定义，我们有

$$\mathcal{L}_\sigma(x(t), y(t), \lambda^*) - \mathcal{L}_\sigma(x^*, y^*, \lambda^*) = \mathcal{O}(\frac{1}{t^2}),$$

以及

$$\mathcal{L}(x(t), y(t), \lambda^*) - \mathcal{L}(x^*, y^*, \lambda^*) = \mathcal{O}(\frac{1}{t^2}).$$

完成了证明。

参考文献

[1] Polyak, B.T. (1964) Some Methods of Speeding up the Convergence of Iteration Methods. *USSR Computational Mathematics and Mathematical Physics*, **4**, 1-17. https://doi.org/10.1016/0041-5553(64)90137-5

[2] Polyak, B.T. (1987) Introduction to Optimization, Optimization Software. In: *Translations Series in Mathematics and Engineering*. Publications Division Inc.

[3] Su, W.J., Boyd, S. and Candés, E. (2016) A Differential Equation for Modeling Nesterov's Accelerated Gradient Method: Theory and Insights. *Journal of Machine Learning Research*, **17**, 5312-5354.

[4] Nesterov,Y.E. (1983) A Method for Solving the Convex Programming Problem with Convergence Rate $\mathcal{O}(1/k^2)$. *Doklady Akademii Nauk*, **269**, 543-547.

[5] Zeng, X., Lei, J. and Chen, J. (2023) Dynamical Primal-Dual Nesterov Accelerated Method and Its Application to Network Optimization. *IEEE Transactions on Automatic Control*, **68**, 1760-1767. https://doi.org/10.1109/tac.2022.3152720

[6] He, X., Hu, R. and Fang, Y.P. (2022) "Second-Order Primal" + "First-Order Dual" Dynamical Systems with Time Scaling for Linear Equality Constrained Convex Optimization Problems. *IEEE Transactions on Automatic Control*, **67**, 4377-4383.

[7] Zhu, T., Hu, R. and Fang, Y. (2024) Tikhonov Regularized Second-Order Plus First-Order Primal-Dual Dynamical Systems with Asymptotically Vanishing Damping for Linear Equality Constrained Convex Optimization Problems. *Optimization*, 1-28. https://doi.org/10.1080/02331934.2024.2407515

[8] He, X., Hu, R. and Fang, Y.P. (2021) Convergence Rates of Inertial Primal-Dual Dynamical Methods for Separable Convex Optimization Problems. *SIAM Journal on Control and Optimization*, **59**, 3278-3301. https://doi.org/10.1137/20m1355379

[9] He, X., Tian, F., Li, A. and Fang, Y. (2023) Convergence Rates of Mixed Primal-Dual Dynamical Systems with Hessian Driven Damping. *Optimization*, 1-26. https://doi.org/10.1080/02331934.2023.2253813

[10] Hulett, D.A. and Nguyen, D. (2023) Time Rescaling of a Primal-Dual Dynamical System with Asymptotically Vanishing Damping. *Applied Mathematics Optimization*, **88**, Article No. 27. https://doi.org/10.1007/s00245-023-09999-9

[11] Bach, F. (2013) Learning with Submodular Functions: A Convex Optimization Perspective. *Foundations and Trends in Machine Learning*, **6**, 145-373. https://doi.org/10.1561/2200000039

[12] Wright, J., Ganesh, A. and Rao, S. (2009) Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices via Convex Optimization. https://api.semanticscholar.org/CorpusID:212563318

[13] Condat, L. (2012) A Primal-Dual Splitting Method for Convex Optimization Involving Lipschitzian, Proximable and Linear Composite Terms. *Journal of Optimization Theory and Applications*, **158**, 460-479. https://doi.org/10.1007/s10957-012-0245-9

[14] Attouch, H., Chbani, Z. and Riahi, H. (2018) Combining Fast Inertial Dynamics for Convex Optimization with Tikhonov Regularization. *Journal of Mathematical Analysis and Applications*, **457**, 1065-1094. https://doi.org/10.1016/j.jmaa.2016.12.017