

# 原 - 对偶加速动力系统的稳定性分析

曾杨洋

浙江师范大学数学科学学院, 浙江 金华

收稿日期: 2026年4月9日; 录用日期: 2026年5月2日; 发布日期: 2026年5月12日

## 摘要

本文从连续时间动力系统的角度, 研究一类用于线性约束凸优化问题的扰动原 - 对偶加速动力系统。通过构建一个同时刻画原变量与对偶变量的统一Lyapunov函数, 在凸与强凸两种情形下分别推导了扰动情形下的能量衰减性与渐近收敛性。

## 关键词

凸优化, 动力系统, Lyapunov函数

# Stability Analysis of a Primal-Dual Accelerated Dynamical System

Yangyang Zeng

School of Mathematical Sciences, Zhejiang Normal University, Jinhua Zhejiang

Received: April 9, 2026; accepted: May 2, 2026; published: May 12, 2026

## Abstract

From the perspective of continuous-time dynamical systems, this paper studies a class of perturbed primal-dual accelerated dynamical systems for linearly constrained convex optimization problems. By introducing a unified Lyapunov function that captures both the primal and dual variables, we derive energy decay estimates and establish asymptotic convergence properties in the presence of perturbations for both convex and strongly convex cases.

## Keywords

Convex Optimization, Dynamical Systems, Lyapunov Function



## 1. 引言

在现代科学计算与工程应用中，大量实际问题都可以归结为凸优化问题[1]。这类问题广泛存在于机器学习[2]、信号处理[3]、图像处理等诸多领域。随着数据规模的持续增长与模型复杂度的不断提升，如何在保证算法稳定性与可靠性的前提下进一步加快收敛速度，已成为凸优化研究中的核心问题之一。

近年来，从连续时间动力系统的角度研究优化算法的收敛行为，为理解和设计加速算法提供了一种统一而直观的分析工具。该方法通过构造与离散迭代算法相对应的微分方程模型，将优化算法的迭代过程视为动力系统的时间演化轨迹，从而利用动力系统理论中的稳定性分析与能量方法，揭示算法收敛速率背后的内在机制[4]。在动力系统框架下，梯度下降法、重球法以及 Nesterov 加速梯度法等典型模型可以在同一视角下进行分析，算法收敛性也可以与系统轨道对平衡点的稳定性以及能量衰减性质联系。

在文献[5]中，Su W, Boyd S 和 Candes E J 构建了一个连续微分方程模型，用于描述 Nesterov 加速梯度下降法(Nesterov Accelerated Gradient Descent, NAG) [6]，该模型如下：

$$\ddot{x}(t) + \frac{3}{t}\dot{x}(t) + \nabla f(x(t)) = 0,$$

揭示了其  $O(1/k^2)$  收敛速率的本质。

Attouch H 和 Chbani Z 对上述微分方程模型进行了拓展：将模型中的常数 3 替换为参数  $\alpha$ ，对该动力系统开展了收敛性分析。并围绕带渐近消失粘性项的惯性动力学系统

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla f(x(t)) = g(t),$$

深入分析其轨迹渐近行为，不仅证明了当  $\alpha > 3$  且扰动项满足如下条件时，

$$\int_0^{+\infty} t \|g(t)\| dt < +\infty,$$

系统轨迹弱收敛至最优解，还在强凸或极小值点集内部非空等场景下建立强收敛性。与此同时，这类研究还通过时间离散化导出了相应 Nesterov 型加速算法，揭示了连续动力学与离散快速优化方法之间的深层联系，为算法收敛性分析提供了统一的 Lyapunov 框架[7]。

在上述动力系统研究的基础上，Luo 和 Chen 在文献[8]中进一步从常微分方程数值求解器的稳定性角度解释加速机制出发，并进一步构造了新的动力系统(Nesterov accelerated gradient flow, NAG flow)，其基本形式可写为：

$$\begin{cases} \dot{x}(t) = v(t) - x(t), \\ \gamma(t)\dot{v}(t) = \mu(x(t) - v(t)) - \nabla f(x(t)), \end{cases}$$

其中  $\dot{\gamma}(t) = \mu - \gamma(t)$ 。该系统等价地可以写成如下二阶形式：

$$\gamma(t)\ddot{x}(t) + (\mu + \gamma(t))\dot{x}(t) + \nabla f(x(t)) = 0.$$

作者构造了相应的 Lyapunov 函数并证明该系统沿轨迹具有指数衰减性质。

进一步地，考虑到实际计算中不可避免的外部扰动、数值误差、随机噪声以及梯度近似误差等因素，

文献[9]将扰动作为外部输入引入动力系统,提出了适用于(可能非光滑)凸优化的二阶加速微分包含模型,即

$$\gamma(t)\ddot{x}(t) + (\mu + \gamma(t))\dot{x}(t) + \partial f(x(t)) \ni \xi(t), \quad \dot{\gamma}(t) = \mu - \gamma(t),$$

其中  $f$  为闭凸非光滑函数,  $\partial f$  表示次微分, 而  $\xi(t)$  用于描述噪声、近似梯度或近端误差以及离散化残差等扰动因素。在连续时间框架下, 作者不仅证明了适当意义下解的存在性, 还构造了一种新的 Lyapunov 函数, 建立了与扰动相关的衰减估计和误差界, 从而说明外部扰动对系统收敛行为的影响。该模型清晰地揭示了扰动对加速动力系统能量衰减与收敛行为的影响, 并为分析不精确加速算法提供了连续时间理论基础。

事实上, 在实际优化问题中不可避免存在数值误差、随机扰动或梯度噪声等因素[10]。扰动因素存在时, 优化相关动力系统是否仍能保持收敛、收敛速率会受到怎样的影响, 以及离散化后算法是否保持可证明的收敛性, 是衡量方法鲁棒性与实际应用潜力的重要方面。尤其是在约束优化问题和拉格朗日建模框架中, 原-对偶动力系统是一类极其常见且结构清晰的连续模型。对于这类系统, 研究其在扰动作用下的稳定性、收敛性以及能量演化规律, 不仅有助于深化对原-对偶加速机制的理解, 也能够为后续算法设计、离散格式构造和鲁棒性分析提供连续时间层面的理论支撑。

基于上述背景, 本文关注带线性约束优化问题中的原-对偶加速动力系统, 重点研究在外部扰动作用下系统解的收敛行为。

## 2. 预备知识

设  $\langle \cdot, \cdot \rangle$  为通常的  $\ell_2$  内积, 并定义  $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$  (即由该内积诱导的范数)。对于适当、闭且凸的函数  $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ , 若存在  $\mu \geq 0$  且满足

$$f(y) \geq f(x) + \langle p, y - x \rangle + \frac{\mu}{2} \|y - x\|^2,$$

则称  $f \in \mathcal{S}_\mu^0$ , 其中  $p \in \partial f(x)$ 。记  $\mathcal{S}_\mu^1$  为  $\mathcal{S}_\mu^0$  中所有连续可微函数构成的集合; 进一步地, 若  $f \in \mathcal{S}_\mu^1$  且其梯度满足 Lipschitz 连续性 ( $L$  为 Lipschitz 常数), 即满足:

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \leq L \|x - y\|,$$

则称  $f \in \mathcal{S}_{\mu,L}^{1,1}$ 。

定义带约束凸优化问题的增广拉格朗日函数为:

$$\mathcal{L}(x, \lambda) := f(x) + \langle \lambda, Ax - b \rangle \quad \forall (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m.$$

设  $(x^*, \lambda^*)$  为  $\mathcal{L}(x, \lambda)$  的鞍点, 这意味着

$$\min_x \mathcal{L}(x, \lambda^*) = \mathcal{L}(x^*, \lambda^*) = \max_\lambda \mathcal{L}(x^*, \lambda),$$

且  $(x^*, \lambda^*)$  也满足 Karush-Kuhn-Tucker (KKT) 条件。

## 3. 含扰动原-对偶加速动力系统的稳定性分析

本文研究一类用于线性约束凸优化问题的含扰动原-对偶加速动力系统, 系统具体形式如下

$$\begin{cases} \dot{\lambda}(t) = y(t) - \lambda(t), \\ \beta(t) \dot{y}(t) = \nabla_\lambda \mathcal{L}(v(t), \lambda(t)) + g_2(t), \\ \dot{x}(t) = v(t) - x(t), \\ \gamma(t) \dot{v}(t) = \mu(x(t) - v(t)) - \nabla_x \mathcal{L}(x(t), y(t)) + g_1(t). \end{cases} \quad (3.1)$$

其中,  $g_1(t)$  和  $g_2(t)$  表示外部扰动或计算误差项。参数函数  $\gamma(t)$  与  $\beta(t)$  分别描述原始变量与对偶变量的时间尺度, 其演化满足  $\dot{\gamma}(t) = \mu - \gamma(t), \dot{\beta}(t) = -\beta(t)$ , 初始条件为  $\gamma(0) = \gamma_0 > 0, \beta(0) = \beta_0 > 0$ 。

易证  $\gamma(t) = \gamma_0 e^{-t} + \mu(1 - e^{-t}), \beta(t) = \beta_0 e^{-t}, \forall t \geq 0$ , 并记  $\gamma_{\min} = \min\{\gamma_0, \mu\}$ 。

除特别说明外, 本文中均假设系统(3.1)的解存在且唯一。

本节研究系统(3.1)的渐近最小化性质, 即系统轨道是否能够渐近逼近原问题的最优解并满足约束条件。为此, 下面构造一个同时刻画原始变量与对偶变量偏离鞍点程度的 Lyapunov 函数, 并利用该函数分析系统在扰动存在时的渐近行为及衰减性质。

设  $X(t) = (x(t), \lambda(t), v(t), y(t))$  是系统(3.1)的唯一解, 定义一个 Lyapunov 函数  $\mathcal{E}(t) := L(t) - \Xi(t)$ , 其中

$$\begin{cases} \Xi(t) := \int_0^t e^{s-t} \langle g_1(s), v(s) - x^* \rangle ds + \int_0^t e^{s-t} \langle g_2(s), y(s) - \lambda^* \rangle ds, \\ L(t) := \mathcal{L}(x(t), \lambda^*) - \mathcal{L}(x^*, \lambda(t)) + \frac{\gamma(t)}{2} \|v(t) - x^*\|^2 + \frac{\beta(t)}{2} \|y(t) - \lambda^*\|^2. \end{cases}$$

函数  $L(t)$  由两部分组成, 第一项  $\mathcal{L}(x(t), \lambda^*) - \mathcal{L}(x^*, \lambda(t))$  描述了当前点相对于鞍点  $(x^*, \lambda^*)$  的拉格朗日函数误差。根据拉格朗日函数的鞍点性质, 该项非负。第二部分由能量项构成, 其中

$$\frac{\gamma(t)}{2} \|v(t) - x^*\|^2, \frac{\beta(t)}{2} \|y(t) - \lambda^*\|^2,$$

分别衡量原始变量与对偶变量在加速方向上的能量项。

在存在扰动的情形下, 本文进一步引入修正项  $\Xi(t)$ , 并研究修正后的 Lyapunov 函数  $\mathcal{E}(t)$ 。通过这种构造, 即使在扰动存在时, 仍可建立相应的能量衰减不等式, 并为后续推导指数型或多项式型衰减估计提供基础。

**引理 1** 若  $f \in S_{\mu, L}^{1,1}$ , 其中  $0 \leq \mu \leq L < \infty$ ; 并且  $X(t) = (x(t), \lambda(t), v(t), y(t))$  是系统(3.1)的唯一解, 则对定义的  $\mathcal{E}(t)$ , 以下不等式成立

$$\dot{\mathcal{E}}(t) \leq -\mathcal{E}(t) - \frac{\mu}{2} \|\dot{x}(t)\|^2,$$

并可以推出

$$\mathcal{E}(t) + \frac{\mu}{2} \int_0^t e^{s-t} \|\dot{x}(s)\|^2 ds \leq e^{-t} \mathcal{E}(0), \quad 0 \leq t < \infty,$$

且

$$L(t) \leq e^{-t} (2L(0) + R^2(t)),$$

其中

$$R(t) := \int_0^t e^{s/2} \left( \frac{\|g_1(s)\|}{\sqrt{\gamma(s)}} + \frac{\|g_2(s)\|}{\sqrt{\beta(s)}} \right) ds.$$

**证明:**

$$\begin{aligned} \frac{d}{dt}(L(t)) &= \langle \dot{x}(t), \nabla_x \mathcal{L}(x(t), \lambda^*) \rangle - \langle \dot{\lambda}(t), \nabla_\lambda \mathcal{L}(x^*, \lambda(t)) \rangle + \langle \gamma(t) \dot{v}(t), v(t) - x^* \rangle \\ &\quad + \langle \beta(t) \dot{y}(t), y(t) - \lambda^* \rangle + \frac{\dot{\gamma}(t)}{2} \|v(t) - x^*\|^2 + \frac{\dot{\beta}(t)}{2} \|y(t) - \lambda^*\|^2 \\ &= I_1 + I_2 + I_3, \end{aligned}$$

其中,

$$\begin{aligned} I_1 &= \frac{\mu - \gamma(t)}{2} \|v(t) - x^*\|^2 - \frac{\beta(t)}{2} \|y(t) - \lambda^*\|^2 + \langle \mu(x(t) - v(t)), v(t) - x^* \rangle, \\ I_2 &= \langle v(t) - x(t), \nabla f(x(t)) + A^T \lambda^* \rangle - \langle y(t) - \lambda(t), Ax^* - b \rangle \\ &\quad - \langle \nabla f(x(t)) + A^T y(t), v(t) - x^* \rangle + \langle Av(t) - b, y(t) - \lambda^* \rangle, \\ I_3 &= \langle g_1(t), v(t) - x^* \rangle + \langle g_2(t), y(t) - \lambda^* \rangle, \end{aligned}$$

利用如下恒等式

$$\mu \langle x(t) - v(t), v(t) - x^* \rangle = \frac{\mu}{2} (\|x(t) - x^*\|^2 - \|v(t) - x^*\|^2 - \|v(t) - x(t)\|^2),$$

该恒等式虽形式简单, 却是推导中的一个重要工具。首先, 将  $I_1$  重写为

$$\begin{aligned} I_1 &= \frac{\mu}{2} \|v(t) - x^*\|^2 + \frac{\mu}{2} (\|x(t) - x^*\|^2 - \|v(t) - x^*\|^2 - \|v(t) - x(t)\|^2) \\ &\quad - \frac{\gamma(t)}{2} \|v(t) - x^*\|^2 - \frac{\beta(t)}{2} \|y(t) - \lambda^*\|^2 \\ &= \frac{\mu}{2} \|x(t) - x^*\|^2 - \frac{\mu}{2} \|\dot{x}(t)\|^2 - \frac{\gamma(t)}{2} \|v(t) - x^*\|^2 - \frac{\beta(t)}{2} \|y(t) - \lambda^*\|^2, \end{aligned}$$

其次根据拉格朗日函数定义以及 KKT 条件, 将  $I_2$  重写为

$$\begin{aligned} I_2 &= \langle x^* - x(t), \nabla_x \mathcal{L}(x(t), \lambda^*) \rangle + \langle v(t) - x^*, \nabla_x \mathcal{L}(x(t), \lambda^*) \rangle \\ &\quad - \langle v(t) - x^*, \nabla_x \mathcal{L}(x(t), y(t)) \rangle + \langle Av(t) - b, y(t) - \lambda^* \rangle, \\ &= \langle x^* - x(t), \nabla_x \mathcal{L}(x(t), \lambda^*) \rangle + \langle v(t) - x^*, A^T (\lambda^* - y(t)) \rangle + \langle Av(t) - b, y(t) - \lambda^* \rangle \\ &= \langle x^* - x(t), \nabla_x \mathcal{L}(x(t), \lambda^*) \rangle. \end{aligned}$$

再由  $\mathcal{L}(\cdot, \lambda^*) \in \mathcal{S}_\mu^1$  可推出:

$$I_2 \leq \mathcal{L}(x^*, \lambda^*) - \mathcal{L}(x(t), \lambda^*) - \frac{\mu}{2} \|x(t) - x^*\|^2,$$

因此,

$$\begin{aligned} \frac{d}{dt} L(t) &\leq \mathcal{L}(x^*, \lambda^*) - \mathcal{L}(x(t), \lambda^*) - \frac{\mu}{2} \|\dot{x}(t)\|^2 - \frac{\gamma(t)}{2} \|v(t) - x^*\|^2 - \frac{\beta(t)}{2} \|y(t) - \lambda^*\|^2 \\ &\quad + \langle g_1(t), v(t) - x^* \rangle + \langle g_2(t), y(t) - \lambda^* \rangle \\ &= -L(t) - \frac{\mu}{2} \|\dot{x}(t)\|^2 + \langle g_1(t), v(t) - x^* \rangle + \langle g_2(t), y(t) - \lambda^* \rangle. \end{aligned}$$

再由  $\Xi(t)$  的定义可得

$$\begin{aligned} \frac{d}{dt} \Xi(t) &= \langle g_1(t), v(t) - x^* \rangle + \langle g_2(t), y(t) - \lambda^* \rangle \\ &\quad - \int_0^t e^{s-t} \langle g_1(s), v(s) - x^* \rangle ds - \int_0^t e^{s-t} \langle g_2(s), y(s) - \lambda^* \rangle ds, \end{aligned}$$

即

$$\frac{d}{dt} \Xi(t) = \langle g_1(t), v(t) - x^* \rangle + \langle g_2(t), y(t) - \lambda^* \rangle - \Xi(t).$$

再由上两式相减得到

$$\dot{\mathcal{E}}(t) \leq -\mathcal{E}(t) - \frac{\mu}{2} \|\dot{x}(t)\|^2,$$

将不等式乘以  $e^t$  并对  $[0, t]$  积分即可得到

$$\mathcal{E}(t) + \frac{\mu}{2} \int_0^t e^{s-t} \|\dot{x}(s)\|^2 ds \leq e^{-t} \mathcal{E}(0), \quad 0 \leq t < \infty,$$

接下来对  $\Xi(t)$  行估计, 从而得到  $L(t)$  的显式衰减界。

由  $\Xi(t)$  的定义以及 Cauchy-Schwarz 不等式可得

$$|\Xi(t)| \leq \int_0^t e^{s-t} (\|g_1(s)\| \|v(s) - x^*\| + \|g_2(s)\| \|y(s) - \lambda^*\|) ds.$$

注意到  $L(t)$  的定义中包含如下两项:

$$\frac{\gamma(t)}{2} \|v(t) - x^*\|^2, \quad \frac{\beta(t)}{2} \|y(t) - \lambda^*\|^2,$$

因此对任意  $s \in [0, t]$ , 均有

$$\|v(s) - x^*\| \leq \sqrt{\frac{2L(s)}{\gamma(s)}}, \quad \|y(s) - \lambda^*\| \leq \sqrt{\frac{2L(s)}{\beta(s)}}.$$

代入上式得到

$$|\Xi(t)| \leq \int_0^t e^{s-t} \sqrt{2L(s)} \left( \frac{\|g_1(s)\|}{\sqrt{\gamma(s)}} + \frac{\|g_2(s)\|}{\sqrt{\beta(s)}} \right) ds.$$

再由  $L(t)$  的定义可得

$$\begin{aligned} e^t L(t) &\leq L(0) + e^t \Xi(t) = L(0) + \int_0^t e^s (\langle g_1(s), v(s) - x^* \rangle + \langle g_2(s), y(s) - \lambda^* \rangle) ds \\ &\leq L(0) + \int_0^t e^s (\|g_1(s)\| \|v(s) - x^*\| + \|g_2(s)\| \|y(s) - \lambda^*\|) ds \\ &\leq L(0) + \int_0^t e^s \sqrt{2L(s)} \left( \frac{\|g_1(s)\|}{\sqrt{\gamma(s)}} + \frac{\|g_2(s)\|}{\sqrt{\beta(s)}} \right) ds. \end{aligned}$$

由文献[7]中 Lemma A.1 可得

$$e^{\frac{t}{2}} \sqrt{2L(t)} \leq \sqrt{2L(0)} + R(t).$$

令

$$w(s) = e^{\frac{s}{2}} \cdot \left( \frac{\|g_1(s)\|}{\sqrt{\gamma(s)}} + \frac{\|g_2(s)\|}{\sqrt{\beta(s)}} \right),$$

则

$$\begin{aligned} e^t \Xi(t) &\leq \int_0^t w(s) \cdot e^{\frac{s}{2}} \cdot \sqrt{2L(s)} ds \\ &\leq \int_0^t w(s) \cdot (\sqrt{2L(0)} + R(s)) ds \\ &= \sqrt{2L(0)} \int_0^t w(s) ds + \int_0^t w(s) \cdot R(s) ds \\ &= \sqrt{2L(0)} \cdot R(t) + \frac{1}{2} R^2(t), \end{aligned}$$

即

$$\begin{aligned}\Xi(t) &\leq e^{-t} \left( \sqrt{2L(0)} \cdot R(t) + \frac{1}{2} R^2(t) \right), \\ L(t) &\leq e^{-t} \left( L(0) + \sqrt{2L(0)} \cdot R(t) + \frac{1}{2} R^2(t) \right) \leq e^{-t} (2L(0) + R^2(t)).\end{aligned}$$

引理得证。

**定理 1** 若  $f \in \mathcal{S}_0^1$ , 且扰动项  $g_1(t)$  与  $g_2(t)$  满足

$$e^{-\frac{t}{2}} \int_0^t e^s (\|g_1(s)\| + \|g_2(s)\|) ds \rightarrow 0 \quad (t \rightarrow \infty),$$

则系统(3.1)的解对任意  $t \geq 0$ , 有

$$\mathcal{L}(x(t), \lambda^*) - \mathcal{L}(x^*, \lambda(t)) \leq 2e^{-t} L(0) + \frac{e^{-t}}{\min\{\gamma_0, \beta_0\}} \left( \int_0^t e^s (\|g_1(s)\| + \|g_2(s)\|) ds \right)^2.$$

**证明:** 由引理 1 可得

$$L(t) \leq e^{-t} (2L(0) + R^2(t)).$$

因为  $\mu = 0$ , 可得  $\gamma(t) = \gamma_0 e^{-t}, \beta(t) = \beta_0 e^{-t}$ . 于是

$$\begin{aligned}\mathcal{L}(x(t), \lambda^*) - \mathcal{L}(x^*, \lambda(t)) &\leq 2e^{-t} L(0) + e^{-t} \left( \int_0^t e^{s/2} \left( \frac{\|g_1(s)\|}{\sqrt{\gamma(s)}} + \frac{\|g_2(s)\|}{\sqrt{\beta(s)}} \right) ds \right)^2 \\ &\leq 2e^{-t} L(0) + e^{-t} \left( \int_0^t e^s \left( \frac{\|g_1(s)\|}{\sqrt{\gamma_0}} + \frac{\|g_2(s)\|}{\sqrt{\beta_0}} \right) ds \right)^2 \\ &\leq 2e^{-t} L(0) + \frac{e^{-t}}{\min\{\gamma_0, \beta_0\}} \left( \int_0^t e^s (\|g_1(s)\| + \|g_2(s)\|) ds \right)^2.\end{aligned}$$

定理得证。

**定理 2** 若  $f \in \mathcal{S}_\mu^1, \mu > 0$ , 且扰动项  $g_1(t)$  与  $g_2(t)$  满足

$$e^{-\frac{t}{2}} \int_0^t e^{\frac{s}{2}} (\|g_1(s)\| + e^s \|g_2(s)\|) ds \rightarrow 0 \quad (t \rightarrow \infty),$$

则系统(3.1)的解对任意  $t \geq 0$ , 有

$$\mathcal{L}(x(t), \lambda^*) - \mathcal{L}(x^*, \lambda(t)) \leq 2e^{-t} L(0) + \frac{e^{-t}}{\min\{\gamma_{\min}, \beta_0\}} \left( \int_0^t e^{\frac{s}{2}} (\|g_1(s)\| + e^s \|g_2(s)\|) ds \right)^2,$$

其中  $\gamma_{\min} = \min\{\mu, \gamma_0\}$ 。

定理 1 证明与定理 2 证明类似, 故此处不再赘述。

这里关于扰动项的积分条件, 本质上要求扰动在适当的时间加权意义下是可积的, 即扰动虽然允许在演化过程中存在, 但其累计影响不能随着时间无限放大。

从实际背景看, 这类假设并不苛刻。如在随机优化或不精确梯度计算中, 若噪声方差随时间衰减, 则对应的扰动项通常满足本文所需的加权可积条件。此外, 在原-对偶迭代中若每一步子问题仅被近似求解, 而近似残差能够逐步减小, 则这类残差同样可归入本文讨论的扰动范畴。因此, 本文的扰动条件可以涵盖方差随时间衰减的随机噪声及渐近消失的计算误差等常见情形。

**定理 3** 在定理 1 或定理 2 的条件下, 系统(3.1)的解满足如下结论, 对任意  $t \geq 0$ , 有

$$0 \leq \mathcal{L}(x(t), \lambda^*) - \mathcal{L}(x^*, \lambda(t)) \leq e^{-t} (2L(0) + R(t)^2),$$

并且

$$\|Ax(t) - b\| \leq e^{-t} (\mathcal{H}_0 + \sqrt{2\beta_0} R(t)),$$

以及

$$|f(x(t)) - f(x^*)| \leq e^{-t} (2L(0) + \mathcal{H}_0 \|\lambda^*\| + \sqrt{2\beta_0} \|\lambda^*\| R(t) + R^2(t)),$$

其中

$$\mathcal{H}_0 := 2\sqrt{\beta_0 L(0)} + \beta_0 \|y_0 - \lambda^*\| + \|Ax_0 - b\|.$$

**证明:** 因为  $L(t)$  的前两项就是  $\mathcal{L}(x, \lambda^*) - \mathcal{L}(x^*, \lambda)$ , 且其余两项非负, 故

$$0 \leq \mathcal{L}(x, \lambda^*) - \mathcal{L}(x^*, \lambda) \leq e^{-t} (2L(0) + R(t)^2).$$

令

$$\xi(t) := y(t) - \beta^{-1}(t)(Ax(t) - b) - \int_0^t \beta^{-1}(s) g_2(s) ds,$$

由系统(3.1)及上式可得  $\dot{\xi}(t) = 0$ , 故  $\xi(t) = \xi(0)$ 。故有

$$\|Ax(t) - b\| = \beta(t) \|y(t) - \xi(0)\| \leq \beta(t) (\|y(t) - \lambda^*\| + \|\xi(0) - \lambda^*\|).$$

再由引理 1 中  $L(t)$  的估计以及  $\beta(t) = \beta_0 e^{-t}$ , 可得

$$\beta_0 \|y(t) - \xi(0)\| \leq \sqrt{4\beta_0 L(0) + 2\beta_0 R^2(t)}.$$

因此,

$$\begin{aligned} \|Ax(t) - b\| &\leq e^{-t} (\sqrt{4\beta_0 L(0) + 2\beta_0 R^2(t)} + \beta_0 \|y_0 - \lambda^*\| + \|Ax_0 - b\|) \\ &\leq e^{-t} (\mathcal{H}_0 + \sqrt{2\beta_0} R(t)). \end{aligned}$$

再由  $\mathcal{L}(x(t), \lambda^*) - \mathcal{L}(x^*, \lambda(t)) = f(x(t)) - f(x^*) + \langle \lambda^*, Ax(t) - b \rangle$  可得

$$\begin{aligned} |f(x(t)) - f(x^*)| &\leq \|\lambda^*\| \|Ax(t) - b\| + e^{-t} (2L(0) + R^2(t)) \\ &\leq e^{-t} (2L(0) + \mathcal{H}_0 \|\lambda^*\| + \sqrt{2\beta_0} \|\lambda^*\| R(t) + R^2(t)). \end{aligned}$$

定理得证。

本文所采用的扰动条件可以理解为系统内部的阻尼与校正机制最终能够压制外部误差的累积影响, 从而恢复整体收敛性。应当说明的是, 这里给出的条件主要是便于分析的一组充分条件, 未必是收敛所必需的最弱条件; 但若完全去除这类约束, 则一般无法从理论上保证系统仍然收敛到原问题的最优解及其对应鞍点。

#### 4. 总结与展望

本文从连续时间动力系统的角度出发, 研究了一类适用于带线性约束凸优化问题的原 - 对偶加速动力系统。在连续时间框架下, 通过构造同时刻画原始变量与对偶变量行为的 Lyapunov 函数, 对连续系统

和离散算法的长期行为进行了统一分析。最后分别在凸与强凸两种情形下建立扰动存在时系统的能量衰减性质及渐近收敛性结果。

本文的理论分析主要集中在凸优化问题。未来可考虑将相关分析框架推广到更一般的非线性约束情形，或进一步拓展到变分不等式、鞍点问题等更广泛的模型，并结合近端算子与单调算子理论处理更一般的非光滑目标函数。

## 参考文献

- [1] Boyd, S. and Vandenberghe, L. (2004) *Convex Optimization*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511804441>
- [2] Kervadec, H., Dolz, J., Yuan, J., Desrosiers, C., Granger, E. and Ben Ayed, I. (2022) Constrained Deep Networks: Lagrangian Optimization via Log-Barrier Extensions. *2022 30th European Signal Processing Conference (EUSIPCO)*, Belgrade, 29 August-2 September 2022, 962-966. <https://doi.org/10.23919/eusipco55093.2022.9909927>
- [3] Lei, W., Zhang, Y. and Chen, Z. (2025) Minimum Peak Sidelobe Ratio Filter for MIMO Radar via Convex Optimization. *Signal Processing*, **226**, Article 109626. <https://doi.org/10.1016/j.sigpro.2024.109626>
- [4] Vassilis, A., Jean-François, A. and Charles, D. (2018) The Differential Inclusion Modeling FISTA Algorithm and Optimality of Convergence Rate in the Case  $B \leq 3$ . *SIAM Journal on Optimization*, **28**, 551-574. <https://doi.org/10.1137/17m1128642>
- [5] Su, W., Boyd, S. and Candès, E. (2016) A Differential Equation for Modeling Nesterov's Accelerated Gradient Method: Theory and Insights. *Journal of Machine Learning Research*, **17**, 1-43.
- [6] Nesterov, Y. (1983) A Method for Solving the Convex Programming Problem with Convergence Rate  $O(1/k^2)$ . *Doklady Akademii Nauk SSSR*, **269**, 543-547.
- [7] Attouch, H., Chbani, Z., Peypouquet, J. and Redont, P. (2018) Fast Convergence of Inertial Dynamics and Algorithms with Asymptotic Vanishing Viscosity. *Mathematical Programming*, **168**, 123-175. <https://doi.org/10.1007/s10107-016-0992-8>
- [8] Luo, H. and Chen, L. (2022) From Differential Equation Solvers to Accelerated First-Order Methods for Convex Optimization. *Mathematical Programming*, **195**, 735-781. <https://doi.org/10.1007/s10107-021-01713-3>
- [9] Luo, H. (2023) Accelerated Differential Inclusion for Convex Optimization. *Optimization*, **72**, 1139-1170. <https://doi.org/10.1080/02331934.2021.2002327>
- [10] Ghadimi, S. and Lan, G.H. (2012) Optimal Stochastic Approximation Algorithms for Strongly Convex Stochastic Composite Optimization I: A Generic Algorithmic Framework. *SIAM Journal on Optimization*, **22**, 1469-1492.