A Survey on Finite-, Fixed-, and Prescribed-Time Convergent Optimization Algorithms: Control-Theoretic Perspectives

21st IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA 2025)

Osama F. Abdel Aal
Department of Mechanical Engineering
University of California, Merced
o.abdelaal@ucmerced.edu

Outline

Introduction

- Optimization algorithms lie at the heart of modern artificial intelligence and machine learning techniques. In most applications, fast and efficient algorithms are desired for solving the optimization problem. This is particularly true in machine learning applications where large data sets lead to larger problem instances and potentially larger computational time.
- As a result, stochastic gradient descent (SGD), its variants such as mini-batch SGD, Adam, momentum-based, and accelerated stochastic methods have emerged as popular choices.
- In developing accelerated optimization algorithms, the discrete-time framework often proves non-intuitive and restrictive from an analytical standpoint. In contrast, continuous-time algorithms provide better intuition, and simpler and elegant proofs are often obtained by leveraging the tools of Lyapunov stability theory. Indeed, the connection between ordinary differential equations and optimization has been recognized for several decades

Introduction

- Despite much progress, there remain two main limitations for continuous-time algorithms: (1) most of the analysis has focused on asymptotic and exponential convergence, i.e., convergence as time tends to infinity; and (2) there have been few systematic studies on developing discrete-time implementations such that the accelerated convergence properties of the continuous-time algorithm are preserved.
- This presentation focuses on continuous-time (accelerated) gradient flow dynamics with finite/fixed-time convergence guarantees. The notion of finite-time stability (FTS), which is a precursor to the notion of fixed-time stability, was proposed in the seminal work ([Bhat and Bernstein, 2000]). A system is said to be finite-time stable if the trajectories converge to the equilibrium in a finite amount of time, called the settling time. The settling time may depend on the initial conditions, and can potentially grow unbounded as the initial conditions go farther away from the equilibrium point. Fixed-time stability (FxTS), on the other hand, is a stronger notion, which requires the settling time to be uniformly bounded for all initial conditions, i.e., convergence within a fixed time can be guaranteed ([Polyakov, 2011]).

Introduction

- Recently, [Polyakov et al., 2019]) introduced the notion of consistent discretization for finite, and fixed-time stable dynamical systems. In particular, they proposed an implicit discretization scheme that preserves the convergence behavior of the continuous time system. However, these results are of little use for the optimization community, since, a) the requirement of the dynamics being homogeneous cannot be satisfied unless the equilibrium point, in this case, the optimizer, is known, and b) implicit discretization schemes are not easy to implement, thus, making it difficult to use these schemes for iterative methods.
- The authors in [Benosman et al.,] showed that the FTS flow, re-scaled gradient flow, and signed-gradient flow, all with a finite-time convergence, when discretized using various explicit schemes, such as Euler discretization or Runge-Kutta method, preserve the convergence behavior in the discrete-time, i.e., the minimizer could be computed within a finite number of iterations for a class of convex optimization problems

Preliminaries

Preliminaries

In this section, the required definitions and stability analysis are introduced to highlight the property of the proposed class of gradient flows that satisfies the prescribed or arbitrary time of convergence.

Consider the nonautonomous nonlinear system defined as:

$$\dot{x} = f(t, x), \quad x(t_0) = x_0 \tag{1}$$

where $x \in \mathbb{R}^n$ is the state vector, and $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}^n$ is a nonlinear function satisifes $f(t, 0) = 0$, that defines the origin $x = 0$ as an equilibrium point of the dynamical system. $t_0 \in \mathbb{R}_{\geq 0}$ is the initial time. The following definitions are introduced:

Finite Time Stability

Definition [Polyakov, 2011]

The origin of system (1) is considered globally finite-time stable if it is globally asymptotically stable and every solution $x_{(t_0,x_0)}(t)$ of (1) reaches the origin within a finite time, i.e.,

$$x_{(t_0,x_0)}(t) = 0, \quad \forall t \geq t_0 + \mathcal{T}(t_0, x_0) \tag{2}$$

where $\mathcal{T} : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is a time function that measures the time of convergence to the origin starting from some initial condition (t_0, x_0).

Finite Time Stability- Lyapunov definition

The system is finite-time stable if there exists a continuously differentiable, positive-definite function $V(x)$, known as a Lyapunov function, such that:

$$\dot{V}(x) \leq -cV(x)^p, \quad 0 < p < 1, \quad c > 0. \tag{3}$$

Under this condition, the time required for the system state to reach equilibrium can be estimated as:

$$T_f \leq \frac{V(x_0)^{1-p}}{c(1-p)}, \tag{4}$$

demonstrating explicit dependence on the initial condition x_0.

Fixed Time Stability

Definition [Polyakov, 2011]

The origin of the system (1) is defined to be fixed time stable if it is globally finite time stable and the time function $\mathcal{T}$ has an upper bound, say, $\tau_{max} > 0$ such that:

$$\mathcal{T}(t_0, x_0) \leq \tau_{max}, \quad \forall x_0 \in \mathbb{R}^n, \forall t_0 \in \mathbb{R}_{\geq 0} \tag{5}$$

Fixed Time Stability- Lyapunov definition

FXTS Lyapunov definition

A system achieves fixed-time stability if there exists a Lyapunov function satisfying:

$$\dot{V}(x) \leq -c_1 V(x)^{p_1} - c_2 V(x)^{p_2}, \quad 0 < p_1 < 1, \quad p_2 > 1, \tag{6}$$

where $c_1, c_2 > 0$ are positive constants. The guaranteed upper bound for convergence time is given by:

$$T_f \leq \frac{1}{c_1(1-p_1)} + \frac{1}{c_2(p_2-1)}. \tag{7}$$

This property makes fixed-time approaches particularly useful for control applications requiring robust, uniform performance regardless of initial conditions.

Prescribed Time Stability

Definition [Polyakov, 2011]

The origin of the system (1) is said to be a prescribed/arbitrary time stable if it is fixed time stable, and $\exists \tau_p \in \mathbb{R}_{\geq 0}$, with no dependence on any system parameters or initial conditions and can be predefined or designed in advance. If $\mathcal{T}(t_0, x_0) = \tau_p$, then the origin is strictly prescribed-time convergent, while $\mathcal{T}(t_0, x_0) < \tau_p$ indicates weakly prescribed-time convergence.

Prescribed Time Stability- Lyapunov definition

A control law achieves prescribed-time convergence if the Lyapunov function satisfies:

$$\dot{V}(x,t) \leq -\frac{V(x)}{T_p - t}, \tag{8}$$

ensuring that $V(x)$ reaches zero precisely at $t = T_p$. Unlike fixed-time methods, this approach allows precise timing control, making it particularly suitable for *real-time optimization and scheduling problems*.

From Exponential Stability to Finite/Fixed Time Stability

Exponential Stability

We consider time-invariant dynamical systems,

$$\dot{x} = F(x) \tag{9}$$

with an equilibrium point at the origin, $F(0) = 0$.

Definition

The origin is the globally exponentially stable equilibrium point of 9 if there are positive constants M and ρ that are independent of $x(0)$ such that the solution $x(t)$ to system (9) satisfies,

$$\|x(t)\| \leq Me^{-\rho t}\|x(0)\|, \quad \forall t \geq 0. \tag{10}$$

Exponential Stability

Lemma **1** characterizes a necessary and sufficient condition for the exponential stability.

Lemma 1

The origin is the globally exponentially stable equilibrium point of system (9) if and only if there exists a continuously differentiable Lyapunov function V that satisfies,

$$\begin{aligned} k_1\|x\|^2 \leq V(x) \leq k_2\|x\|^2 \\ \dot{V}(t) \leq -k_3\|x(t)\|^2, \quad \forall t \geq 0 \end{aligned} \tag{11}$$

along the solutions of system 9.

From Exponential to Finite/Fixed-Time Stability: Applications to Optimization [Ozaslan and Jovanović, 2024]

Consider a globally Lipschitz continuous function F, i.e., $\|F(x)-F(y)\| \leq L\|x-y\|, \quad \forall x, y \in \mathbb{R}^n$, then, the equilibrium point of the modified system:

$$\dot{x} = \sigma(x)F(x), \tag{12}$$

is globally finite-time stable, where $\sigma : \mathbb{R}^n \to \mathbb{R}_+ \cup \{0\}$ is a scaling factor defined as:

$$\sigma(x) = \begin{cases} 0, & F(x) = 0, \\ \eta\|F(x)\|^{-\lambda}, & \text{otherwise}, \end{cases} \tag{13}$$

with $\eta > 0$ and $\lambda \in (0, 1)$.

From Exponential to Finite/Fixed-Time Stability: Applications to Optimization [Ozaslan and Jovanović, 2024]

Furthermore, the global fixed-time stability under two additional conditions: in addition to being Lipschitz continuous, the vector field F satisfies

$$\|F(x)\| \geq m\|x\|^{\beta}, \quad \forall x \in \mathbb{R}^n, \tag{14}$$

for some positive parameters m and β. The scaling factor is then modified to:

$$\sigma(x) = \begin{cases} 0, & F(x) = 0, \\ \eta_1\|F(x)\|^{-\lambda_1} + \eta_2\|F(x)\|^{\lambda_2}, & \text{otherwise}, \end{cases} \tag{15}$$

where $\eta_1, \eta_2, \lambda_2$ are positive parameters and $\lambda_1 \in (0, 1)$. When these hold, a uniform upper bound on the settling time of the system can be obtained for all initial conditions.

From Exponential to Finite/Fixed-Time Stability: Applications to Optimization

Finite-time stability.

$$\dot{V} = \langle \nabla_x V(x), \dot{x} \rangle = \sigma(x) \langle \nabla_x V(x), F(x) \rangle.$$
$$\leq -k_3 \sigma(x) \|x\|^2 = -\frac{k_3 \eta \|x\|^2}{\|F(x)\|^\lambda}. \tag{16}$$

By the Lipschitz continuity condition, we get: $\dot{V} \leq -\frac{k_3\eta}{L^\lambda} \|x\|^{2-\lambda}$. Applying the quadratic upper bound on the Lyapunov function, we obtain:

$$\dot{V} \leq -\frac{k_3 \eta}{L^\lambda k_2^{1-\lambda/2}} V^{1-\lambda/2}. \tag{17}$$

Setting $\alpha = 1 - \lambda/2$ and $c = \frac{k_3 \eta}{L^\lambda k_2^\alpha}$ yields $\dot{V} \leq -cV^\alpha$ with $c > 0$ and $\alpha \in (1/2, 1)$. The settling-time upper bound $T(x(0)) \leq \frac{2k_2 L^\lambda}{k_3 \eta \lambda} \|x(0)\|^\lambda$.

From Exponential to Finite/Fixed-Time Stability: Applications to Optimization

Taking the scaling factor: $\sigma(x) = \begin{cases} 0, & F(x) = 0, \\ \eta_1 \|F(x)\|^{-\lambda_1} + \eta_2 \|F(x)\|^{\lambda_2}, & \text{otherwise}, \end{cases}$

Fixed-time stability.

$$\begin{aligned} \dot{V} &= \langle \nabla_x V(x), \dot{x} \rangle = \sigma(x) \langle \nabla_x V(x), F(x) \rangle \\ &\leq -k_3 \sigma(x) \|x\|^2 \\ &= -k_3 \left(\eta_1 \frac{\|x\|^2}{\|F(x)\|^{\lambda_1}} + \eta_2 \|x\|^2 \|F(x)\|^{\lambda_2} \right) \\ &\leq -\frac{k_3 \eta_1}{L^{\lambda_1}} \|x\|^2 - \lambda_1 - k_3 \eta_2 m^{\lambda_2} \|x\|^2 + \beta\lambda_2 \\ &\leq -\frac{k_3 \eta_1}{L^{\lambda_1} k_2^{1-\lambda_1/2}} V^{1-\lambda_1/2} - \frac{k_3 \eta_2 m^{\lambda_2}}{k_2^{1+\beta\lambda_2/2}} V^{1+\beta\lambda_2/2}. \end{aligned}$$

□

Asymptotic Convergence of Classical Gradient Flows

The Proposed Control-theoretic Approach

Consider the unconstrained optimization problem below, aim to minimize a cost function $J : \mathbb{R}^n \to \mathbb{R}$,

$$\min_{\theta \in \mathbb{R}^n} \quad J(\theta). \tag{18}$$

A necessary condition for the convergence of the GF in (18) to an optimal solution is given in Assumption (22) below,

Assumption 1

The function J has its minimum value $J^* = J(\theta^*) = \min J(\theta) > -\infty$ at $\theta^* \in \mathbb{R}^n$, i.e., $J^* > -\infty$.

An important tool in the theory of optimization is the Polyak-Łojasiewicz (PL) inequality describing the gradient dominance property. The following is the formal definition of PL inequality.

Asymptotic Convergence of Gradient Descent

Continuous-time optimization dynamics provide a foundational framework for analyzing and designing gradient-based optimization methods [Boyd and Vandenberghe, 2004]. This section presents a comparative study of classical continuous-time gradient descent (GD) and its accelerated counterpart, focusing on their convergence properties and asymptotic behavior in convex optimization.

The continuous-time gradient flow is given as:

$$\dot{\theta}(t) = -\nabla f(\theta(t)), \tag{19}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable convex function and $\theta(t) \in \mathbb{R}^n$ denotes the optimization variable.

Asymptotic Convergence of Gradient Descent

Theorem (Convergence of Gradient Flow)

Consider the differentiable function f. Let be convex with Lipschitz continuous gradient ∇f. Then the gradient flow dynamics in (19) *has the convergence rate:*

$$f(\theta(t)) - f^\star \leq \frac{m_1}{t}, \tag{20}$$

for some constant $m_1 > 0$, where $f^\star = \min_\theta f(\theta)$.

This result implies that classical gradient flow achieves a sublinear $\mathcal{O}(1/t)$ convergence rate. The dynamics are asymptotically stable, with convergence to the minimizer occurring as $t \to \infty$.

Asymptotic Convergence of Accelerated Methods

The accelerated gradient descent method, inspired by Nesterov's discrete-time algorithm (NAG) [Nesterov, 2013], can be expressed in continuous time by the second-order ODE [Su et al., 2016]:

$$\ddot{\theta}(t) + \frac{r}{t}\dot{\theta}(t) + \nabla f(\theta(t)) = 0, \tag{21}$$

where $r \geq 3$ is a damping term.

Theorem (Convergence of NAG [Su et al., 2016])

assume $f \in \mathcal{C}^1$ be a convex function, and the gradient ∇f is Lipschitz continuous. Then the rate of convergence of the accelerated dynamics given in (21) is characterized by:

$$f(\theta(t)) - f^\star \leq \frac{m_2}{t^2}, \tag{22}$$

for some constant $m_2 > 0$.

Asymptotic Convergence of Accelerated Methods

Proof.

We define the following storage function:

$$\begin{aligned}\mathcal{E}(t) = t^2 \left(f(\theta(t)) - f^*\right) + \\ \frac{1}{2}\left\|(r-1)(\theta(t)-\theta^\star) + t\dot{\theta}(t)\right\|^2 .\end{aligned} \tag{23}$$

Upon computing the derivative of $\mathcal{E}(t)$, one gets:

$$\begin{aligned}\dot{\mathcal{E}}(t) &= -(r-1)t\langle \nabla f(\theta(t)), \theta(t) - \theta^\star\rangle \\ &\quad + 2t(f(\theta(t)) - f^\star) \\ &\leq (3-r)t(f(\theta(t)) - f^\star).\end{aligned} \tag{24}$$

□

Asymptotic Convergence of Accelerated Methods

Proof.

where the convexity implies $\langle \nabla f(\theta(t)), \theta(t) - \theta^{\star} \rangle \geq f(\theta(t)) - f^{*}$.
If $r \geq 3$, then $\mathcal{E}(t)$ is descresing and

$$t^2 (f(\theta(t)) - f^{\star}) \leq \mathcal{E}(t_0), \quad \forall t \geq t_0. \tag{25}$$

and in the case $r > 3$:

$$\int_{t_0}^{\infty} (r-3) t (f(\theta(t)) - f^{\star}) \, dt \leq \mathcal{E}(t_0). \tag{26}$$

Thus, if f is convex and $r \geq 3$, the solution of (21) satisfies:

$$f(\theta(t)) - f^{\star} = \mathcal{O}\left(\frac{1}{t^2}\right). \tag{27}$$

□

Asymptotic Convergence of Accelerated Methods

1. This improved rate $\mathcal{O}(1/t^2)$ is asymptotic, meaning that while the method is faster in theory, it still requires infinite time to exactly reach the optimizer. Moreover, the second-order dynamics can introduce oscillations near the minimum, potentially affecting stability.
2. From a dynamical systems perspective, gradient flow (19) represents a first-order, monotone dissipation process, whereas accelerated dynamics ((21)) mimic inertial systems with vanishing damping, allowing faster energy dissipation [Attouch et al., 2016, Adly and Attouch, 2020].
3. While both systems converge asymptotically, accelerated dynamics exhibit a provably faster convergence rate under convexity assumptions. However, they may also exhibit overshoot or sensitivity to initial conditions, particularly in non-smooth or ill-conditioned settings.
4. Recent works explore modified gradient dynamics that ensure finite-time convergence. These typically involve nonlinear rescaling or discontinuous vector fields and achieve convergence to the optimizer in finite time [Romero and Benosman, 2020], as opposed to the asymptotic nature of (19) and (21), such methods present an

Newtons Gradient Flow

The classical Newtons method is given by:

$$\dot{\theta}(t) = -\left(\nabla^2 J(\theta)\right)^{-1} \nabla J(\theta). \tag{28}$$

It is well established that, under certain conditions on the function J, the method in (28) achieves exponential convergence [Beck, 2014]. We now introduce the following assumption regarding the objective function J:

Assumption

The function $J \in C^2(\mathbb{R}^n, \mathbb{R})$ is strictly convex. Additionally, its Hessian $\nabla^2 J(\theta)$ is invertible for all $\theta \in \mathbb{R}^n$, and the gradient norm $\|\nabla J(\theta)\|$ is radially unbounded.

Newtons Gradient Flow

Theorem

If J satisfies Assumptions (22) and (29), then the trajectories of (28) converge exponantially to the optimal point θ^*.

Proof.

Consider the unbounded Lyapunov function $\mathcal{V} = \frac{1}{2}\|\nabla J(\theta)\|^2$, one can prove the exponential convergence of (28):

$$\begin{aligned} \dot{\mathcal{V}} &= (\nabla J)^T (\nabla^2 J)\dot{\theta} \\ &= -(\nabla J)^T \left(\rho \nabla J(\theta)\right) \\ &= -\rho\|\nabla J(\theta)\|^2 \\ &\leq -2\rho\mathcal{V}(t), \end{aligned} \tag{29}$$

□

Constrained Optimization Problems

We consider the optimization problem given as:

$$\min_{\theta \in \mathbb{R}^n} J(\theta), \quad \text{s.t.} \quad A\theta = b, \tag{30}$$

where $A \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m$, and $\mathbf{J}(\theta)$ is a strongly convex and smooth function.
The existing work in the literature considers the exponential asymptotic convergence of the PDGD problem; the work presented in [Qu and Li, 2019] rigorously investigated the exponential stability of the PDGD dynamics under the assumption of strongly convex and smooth objective functions subject to linear equality constraints. In [Ding and Jovanović, 2019, Ding and Jovanović, 2020], a Lyapunov-based method was employed to demonstrate the global exponential stability of the PDGD dynamics derived from the proximal augmented Lagrangian framework. Furthermore, the work in [Cerone et al., 2025, Centorrino et al., 2025, Cerone et al., 2024], aimed to design a new exponentially convergent PDGD algorithm via a feedback control approach where the Lagrange multipliers that serve as the control inputs are tuned via a PI controller. [Garg and Panagou, 2021] designed a fixed-time convergent PDGD using the classical fractional power scaling approach.

Constrained Optimization Problems

Assumption

The function J is twice differentiable, μ-strongly convex and ℓ-smooth, i.e., for all $\theta, \theta' \in \mathbb{R}^n$,

$$\mu\|\theta - \theta'\|^2 \leq \langle \nabla J(\theta) - \nabla J(\theta'), \theta - \theta' \rangle \leq \ell\|\theta - \theta'\|^2. \tag{31}$$

Assumption

The matrix A is assumed to be of full row rank and
$\gamma_1 I \preceq AA^T \preceq \gamma_2 I \quad$ for some $\gamma_1, \gamma_2 > 0$.

Remark

Assumption (32) is a standard condition in the analysis of constrained optimization problems. The full row rank property of the matrix A ensures that the feasible set is nonempty and closed. Under this condition, the coercivity of the convex objective function J guarantees the existence of a solution to problem (30).

Constrained Optimization Problems

Let $\mathcal{L}(\theta, \beta)$ be the augmented Lagrangian associated with problem (30) defined as:

$$\mathcal{L}(\theta, \beta) = J(\theta) + \beta^\top (A\theta - b), \tag{32}$$

where $\beta \in \mathbb{R}^m$ is the Lagrangian multiplier. The primal-dual GF dynamics can be used to compute the saddle points associated with the Lagrangian $\mathcal{L}(\theta, \beta)$ in (32). The dynamical system is given by:

$$\begin{aligned} \dot{\theta} &= -\nabla_\theta \mathcal{L}(\theta, \beta) \\ \dot{\beta} &= \kappa \nabla_\beta \mathcal{L}(\theta, \beta), \end{aligned} \tag{33}$$

where $\kappa > 0$ is constant. Let's define $\omega = [\theta^\top, \beta^\top]^\top$, and similarly define $\omega^* = [(\theta^*)^\top, (\beta^*)^\top]^\top$ as the equilibrium point of (33), then the adopted PDGD problem can be written as:

$$\dot{\omega} = F(\omega), \tag{34}$$

where

$$F(\omega) := \begin{bmatrix} -\nabla_\theta \mathcal{L}(\theta, \beta) \\ \kappa \nabla_\beta \mathcal{L}(\theta, \beta) \end{bmatrix} = \begin{bmatrix} -\left(\nabla J(\theta) + A^\top \beta\right) \\ \kappa \left(A\theta - b\right) \end{bmatrix} \tag{35}$$

Exponenetial Convergence of PDGD

Theorem

The equality constrained optimization problem in (30) converge to its optimal point (θ^*, λ^*) *exponentially.*

Proof.

See [Qu and Li, 2019, Appendix A]. □

A survey of Existing Methods

Literature Review

Breaking the Convergence Barrier: Optimization via Fixed-Time Convergent Flows [Budhraja et al., 2022]

$$\dot{x} = -c_1 \frac{\nabla f(x)}{\|\nabla f(x)\|^{\frac{p_1-2}{p_1-1}}} - c_2 \frac{\nabla f(x)}{\|\nabla f(x)\|^{\frac{p_2-2}{p_2-1}}} \tag{36}$$

where $c_1; c_2 > 0$, $p_1 > 2$ and $p_2 \in (1, 2)$.the function f satisfies the Polyak-Lojaseiwicz inequality; the Lyapunov method was used to prove FXTC. Robustness is also studied.

Definition: [Polyak-Łojasiewicz inequality]

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function with $\mathcal{S} \neq \emptyset$. Then, f satisfies the Polyak-Łojasiewicz (PŁ) inequality on $\mathcal{S}$, if there exists $r > \min f$ and $\mu > 0$ such that:

$$2\mu[f(x) - \min f] \leq \|\nabla f(x)\|^2 \quad , \forall x \in [\min f < f < r] \tag{37}$$

we then describe the function to be $f \in P Ł_\mu(\mathcal{S})$. This inequality simply requires that the gradient grows faster than a quadratic function as we move away from the optimal function value. Note that this inequality implies that every stationary point is a global minimum. But unlike Strong Convexity, it does not imply that there is a unique solution, so this condition is that it is some weakening of strong convexity. Recall, that when we discussed the linear convergence of GD we highlighted the key property that strong convexity gave us, if we are far away from optimal then strong convexity will ensure that the gradient is large. The PL condition is a more direct statement of that desirable property (but highlights the crucial fact that this is all we need for linear rates, i.e. we do not need convexity itself).

Proof

Consider the Lyapunov candidate function:

$$V(x) = f(x) - f^*. \tag{38}$$

As f^* is the minimum value of f and x^* is the unique minimizer, it holds that $V(x) > 0$ for all $x \neq x^*$. The time derivative of $V(x)$ is given by:

$$\dot{V}(x) = \nabla f(x)^T \dot{x}. \tag{39}$$

Substituting (36) into (39), we get:

$$\dot{V}(x) = -c_1 \|\nabla f(x)\|^{\frac{p_1}{p_1 - 1}} - c_2 \|\nabla f(x)\|^{\frac{p_2}{p_2 - 1}}. \tag{40}$$

Using the Polyak-Łojasiewicz inequality:

$$\frac{1}{2}\|\nabla f(x)\|^2 \geq \mu(f(x) - f^*), \tag{41}$$

we obtain:

$$\dot{V} \leq -c_1(2\mu V)^{\frac{p_1}{2(p_1-1)}} - c_2(2\mu V)^{\frac{p_2}{2(p_2-1)}}. \tag{42}$$

Define $p := c_1(2\mu)^{\frac{p_1}{2(p_1-1)}}$ and $q := c_2(2\mu)^{\frac{p_2}{2(p_2-1)}}$. Since $p_1 > 2$ and $p_2 \in (1, 2)$, we have $\alpha := \frac{p_1}{2(p_1-1)} \in (0, 1)$ and $\beta := \frac{p_2}{2(p_2-1)} > 1$. Thus, the conditions for fixed-time stability are satisfied, and the settling time T is bounded by:

$$T \leq \frac{1}{p(1-\alpha)} + \frac{1}{q(\beta-1)}. \tag{43}$$

This completes the proof. □

Literature Review

Examples

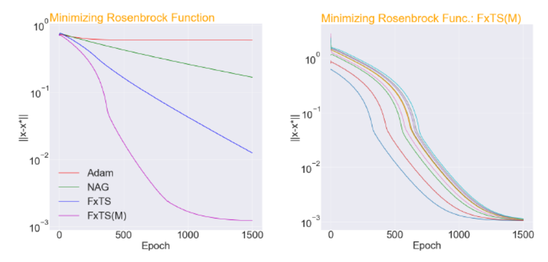

Figure: Minimization of Rosenbrock function, Comparison of various optimization algorithms for the initial condition (0.3, 0.8).

FIRST-ORDER OPTIMIZATION INSPIRED FROM FINITETIME CONVERGENT FLOWS [Zhang et al., 2020]

$$\dot{x} = F_{q-RGF}(x) = -c \frac{\nabla f(x)}{\|\nabla f(x)\|^{\frac{q-2}{q-1}}} \tag{44}$$

$$\dot{x} = F_{SGF}(x) = -c \frac{\nabla f(x)}{\|\nabla f(x)\|^{\frac{1}{q-1}}} sign(\nabla f(x)) \tag{45}$$

where $c > 0, q > 1$. the function f is -gradient dominated of order p, In this paper, the convergence behavior of an Euler discretization for the q-RGF and q-SGF was investigated. Convergence guarantees were provided in terms of closeness of solutions

Fillipov framework

See [Cortes, 2006, Moulay et al., 2019, Li and Wang, 2022]

Fixed-Time Stable Gradient Flows: Applications to Continuous-Time Optimization [Garg and Panagou, 2020]

The paper designed FXTC-GFs for four different optimization problems:

1. **Unconstrained Optimization: FxTS-GF Scheme**

$$\dot{x} = -c_1 \frac{\nabla f(x)}{\|\nabla f(x)\|^{\frac{p_1-2}{p_1-1}}} - c_2 \frac{\nabla f(x)}{\|\nabla f(x)\|^{\frac{p_2-2}{p_2-1}}}, \tag{46}$$

   where $c_1, c_2 > 0$, $p_1 > 2$, and $1 < p_2 < 2$.

2. **Newton's Method: FxTC Scheme**

$$x = -(\nabla^2 f(x))^{-1} \left(c_1 \frac{\nabla f(x)}{\|\nabla f(x)\|^{\frac{p_1-2}{p_1-1}}} + c_2 \frac{\nabla f(x)}{\|\nabla f(x)\|^{\frac{p_2-2}{p_2-1}}} \right), \tag{47}$$

   where $c_1, c_2 > 0$, $p_1 > 2$, and $1 < p_2 < 2$.

Fixed-Time Stable Gradient Flows: Applications to Continuous-Time Optimization [Garg and Panagou, 2020]

1. **Convex Optimization With Linear Equality Constraints**

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad Ax = b, \tag{48}$$

   where $A \in \mathbb{R}^{m \times n}$ is full row rank, and f is a coercive function.

2. **FXTS OF SADDLE-POINT DYNAMICS**

$$\max_{z \in \mathbb{R}^m} \min_{x \in \mathbb{R}^n} F(x, z), \tag{49}$$

   where $F(x, z)$ satisfies local strict convexity–concavity conditions.

Continued

Assumption: The function $f \in C^2(R^n, R)$ is strictly convex. Furthermore, the Hessian $\nabla^2 f$ invertible for all $x \in R^n$ and the norm of the gradient, ∇f is radially unbounded.

Proof: FXT Newton's Method

Consider the Lyapunov function $V(x) = \frac{1}{2}\|\nabla f(x)|^2$:

$$
\begin{aligned}
\dot{V} &= \nabla f^T \nabla^2 f \dot{x} = -\nabla f^T \nabla^2 f \dot{x} \\
&= -\nabla f^T \left[c_1 \frac{\nabla f(x)}{\|\nabla f(x)\|^{\frac{p_1-2}{p_1-1}}} + c_2 \frac{\nabla f(x)}{\|\nabla f(x)\|^{\frac{p_2-2}{p_2-1}}} \right] \\
&= -c_1 \|\nabla f\|^{2-\frac{p_1-2}{p_1-1}} - c_2 \|\nabla f\|^{2-\frac{p_2-2}{p_2-1}} \\
&\leq -c_1 2^{\frac{\alpha_1}{2}} V^{\frac{\alpha_1}{2}} - c_2 2^{\frac{\alpha_2}{2}} V^{\frac{\alpha_2}{2}}.
\end{aligned} \tag{50}
$$

where $\alpha_1 = \frac{p_1-2}{p_1-1}, \alpha_2 = \frac{p_2-2}{p_2-1}$.

Continued

Assumption: The function $f \in C_{1,1}^{loc}(\mathbb{R}^n, \mathbb{R})$ has a unique minimizer $x = x^*$ and satisfies the Polyak-Lojasiewicz (PL) inequality or is dominated by gradients.dients. Under this assumption of gradient dominance it was shown in ([Karimi et al., 2016], Theorem 2) that the function $f(x)$ has a quadratic growth, i.e.:

$$f(x) - f^* \geq \frac{\mu}{2}\|x - x^*\|^2 \quad , \forall x \in \mathbb{R}^n \tag{51}$$

Proof: FXT Unconstrained optimization problem

Consider the Lyapunov function $V(x) = \frac{1}{2}(f - f^*)^2$:

$$\begin{aligned} \dot{V} &= -c_1(f - f^*)\|\nabla f\|^{\alpha_1} - c_2(f - f^*)\|\nabla f\|^{\alpha_2} \\ &\leq -c_1(2\mu_f)^{\frac{\alpha_1}{2}}(f - f^*)^{1+\frac{\alpha_1}{2}} - c_2(2\mu_f)^{\frac{\alpha_2}{2}}(f - f^*)^{1+\frac{\alpha_2}{2}}. \end{aligned} \tag{52}$$

where $\alpha_1 = 2 - \frac{p_1-2}{p_1-1}, \alpha_2 = 2 - \frac{p_2-2}{p_2-1}$.

Literature Review

Finite-time and Fixed-time Convergence in Continuous-time Optimization [Chen et al., 2021]

Definition 1. For a function $f(x)$ whose gradient exists, if there exists a scalar $L > 0$ such that

$$\langle \nabla f(\theta_1) - \nabla f(\theta_2), \theta_1 - \theta_2 \rangle \leq L \|\theta_1 - \theta_2\|^2, \tag{53}$$

for any θ_1 and θ_2 belonging to the definition domain of $f(x)$, then $f(x)$ is said to have an L-continuous gradient.

Definition 2. For a convex function $f(x)$ whose gradient exists, if there exists a scalar $\mu > 0$ such that

$$\langle \nabla f(\theta_1) - \nabla f(\theta_2), \theta_1 - \theta_2 \rangle \geq \mu \|\theta_1 - \theta_2\|^2, \tag{54}$$

for any θ_1 and θ_2 belonging to the definition domain of $f(x)$, then $f(x)$ is said to be μ-strong convex.

Finite-time and Fixed-time Convergence in Continuous-time Optimization [Chen et al., 2021]

Consider the gradient method:

$$\dot{x} = -\rho \frac{\nabla f(x)}{\|\nabla f(x)\|^2}, \tag{55}$$

where $\rho > 0$ is the step size.

Theorem 1. If the convex function $f(x)$ has an L-continuous gradient, then algorithm (6) can reach the minimum point x^* in a finite time.

Proof. Consider the Lyapunov function $V = \|x - x^*\|^2$. Taking the time derivative,

$$\dot{V} = 2(x - x^*)^T \dot{x} = -2\rho \frac{(x - x^*)^T \nabla f(x)}{\|\nabla f(x)\|^2} \leq -\frac{2\rho}{L}. \tag{56}$$

proving the finite-time bound $\square$.

Finite-time and Fixed-time Convergence in Continuous-time Optimization [Chen et al., 2021]

Now consider a more general case:

$$\dot{x} = -\rho \frac{\nabla f(x)}{\|\nabla f(x)\|^{\alpha}}, \quad 0 < \alpha < 2. \tag{57}$$

Theorem 2. If the convex function $f(x)$ has an L-continuous gradient and is μ-strong convex, then algorithm (7) can reach the minimum point x^* in a finite time.
Proof. Consider the Lyapunov function $V = \|x - x^*\|^2$. Taking the time derivative,

$$\dot{V} = -2\rho \frac{(x - x^*)^T \nabla f(x)}{\|\nabla f(x)\|^{\alpha}} \leq -\frac{2\rho\mu^{2-\alpha}}{L} V^{\frac{2-\alpha}{2}}. \tag{58}$$

which gives a finite-time bound $\square$.

Concluding Remarks

1. The advancement from asymptotic convergence to finite-time, fixed-time or prescribed-time shows a systematic improvement in optimization theory.
2. these convergence strategies provide effective frameworks for improving the efficiency of optimization algorithms in ML and AI.
3. Several research directions can be extended: designing discretization schemes of the continuous- time gradient flows while preserving the finite-time stability attracts the attention of some researchers.
4. On the other hand, developing algorithms that is robust against noisy gradient environments while maintaining finite-time convergence guarantees is a challenging aspect.
5. Ongoing research continues to investigate and expand the applications of these convergence concepts for the ML and AI techniques. Furthermore, experimental validations of the aforementioned theoretical advancements for robotics, autonomous systems, and industrial processes involving time-constrained problems will be valuable

References

Adly, S. and Attouch, H. (2020).
Finite convergence of proximal-gradient inertial algorithms combining dry friction with hessian-driven damping.
SIAM Journal on Optimization, 30(3):2134–2162.

Attouch, H., Peypouquet, J., and Redont, P. (2016).
Fast convex optimization via inertial dynamics with hessian driven damping.
Journal of Differential Equations, 261(10):5734–5783.

Beck, A. (2014).
Introduction to nonlinear optimization: Theory, algorithms, and applications with MATLAB.
SIAM.

Benosman, M., Romero, O., and Cherian, A.
First-order optimization algorithms via discretization of finite-time convergent flows.

Bhat, S. P. and Bernstein, D. S. (2000).
Finite-time stability of continuous autonomous systems.
SIAM Journal on Control and optimization, 38(3):751–766.

Boyd, S. P. and Vandenberghe, L. (2004).
Convex optimization.
Cambridge university press.

Budhraja, P., Baranwal, M., Garg, K., and Hota, A. (2022).
Breaking the convergence barrier: Optimization via fixed-time convergent flows.
In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 6115–6122.

Centorrino, V., Rossi, F., Bullo, F., and Russo, G. (2025).
Proximal gradient dynamics and feedback control for equality-constrained composite optimization.
arXiv preprint arXiv:2503.15093.

The End