A Simple Convergence Proof for Stochastic Approximation Using Converse Lyapunov Theory

M. Vidyasagar FRS

SERB National Science Chair, IIT Hyderabad

Theory of Reinforcement Learning Reunion
16 November 2021, Simons Institute, Berkeley

Problem Formulation

Suppose $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^d$. The aim is to find a solution to $\mathbf{f}(\boldsymbol{\theta}) = \mathbf{0}$, when *only noisy measurements* of $\mathbf{f}(\cdot)$ are available.

Start with an initial guess $\boldsymbol{\theta}_0 \in \mathbb{R}^d$. At step $t \geq 0$, let

$$\mathbf{y}_{t+1} = \mathbf{f}(\boldsymbol{\theta}_t) + \boldsymbol{\xi}_{t+1},$$

where $\boldsymbol{\xi}_{t+1}$ is the measurement error. Update via

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \alpha_t \mathbf{y}_{t+1} = \boldsymbol{\theta}_t + \alpha_t(\mathbf{f}(\boldsymbol{\theta}_t) + \boldsymbol{\xi}_{t+1}),$$

where $\{\alpha_t\}_{t\geq 1}$ is a predetermined sequence of step sizes.

Question: When does $\boldsymbol{\theta}_t \to \boldsymbol{\theta}^*$, where $\mathbf{f}(\boldsymbol{\theta}^*) = \mathbf{0}$?

Started by Robbins and Monro (1951).

Some Standard Assumptions

(F). $\boldsymbol{\theta}^*$ is the unique solution of $\mathbf{f}(\boldsymbol{\theta}) = \mathbf{0}$.

(N). Define $\boldsymbol{\theta}_0^t = \{\boldsymbol{\theta}_0, \cdots, \boldsymbol{\theta}_t\}$, and let $\mathcal{F}_t = \sigma(\boldsymbol{\theta}_0^t, \boldsymbol{\xi}_1^t)$. Then (i) the measurements are unbiased, i.e.,

$$E(\boldsymbol{\xi}_{t+1} | \mathcal{F}_t) = \mathbf{0} \text{ a.s.},$$

and (ii) the conditional variance grows quadratically, i.e., $\exists d < \infty$ such that

$$E(\|\boldsymbol{\xi}_{t+1}\|_2^2 | \mathcal{F}_t) \leq d(1 + \|\boldsymbol{\theta}_t\|_2^2).$$

(S). Robbins-Monro (RM) conditions:

$$\sum_{t=0}^{\infty} \alpha_t = \infty, \sum_{t=0}^{\infty} \alpha_t^2 < \infty.$$

IIT Hyderabad

A Typical Theorem

Theorem

Suppose (F), (N), and (S) hold. If $\mathbf{f}(\cdot)$ *satisfies some more conditions, and if the iterates* $\{\boldsymbol{\theta}_t\}$ *are bounded almost surely, then* $\boldsymbol{\theta}_t \to \boldsymbol{\theta}^*$, *a.s. as* $t \to \infty$.

Almost sure boundedness of the iterates ("stability") is a part of the hypothesis, not a conclusion.

Question: Can the stability of the iterates be made a *conclusion, instead of being a part of the hypotheses?*

Borkar-Meyn Theorem (2000)

Assumptions:

- All the standard assumptions (F), (N), (S).
- $\mathbf{f}(\cdot)$ is globally Lipschitz continuous, i.e., $\exists L < \infty$ such that

$$\|\mathbf{f}(\boldsymbol{\theta}) - \mathbf{f}(\boldsymbol{\phi})\|_2 \leq L\|\boldsymbol{\theta} - \boldsymbol{\phi}\|_2, \ \forall \boldsymbol{\theta}, \boldsymbol{\phi} \in \mathbb{R}^d.$$

- There is a "limit function" $\mathbf{f}_\infty$ such that

$$\frac{\mathbf{f}(r\boldsymbol{\theta})}{r} \to \mathbf{f}_\infty(\boldsymbol{\theta}) \text{ as } r \to \infty,$$

  uniformly over compact subsets of $\mathbb{R}^d$.
- $\mathbf{0}$ is a globally exponentially stable equilibrium of

$$\dot{\boldsymbol{\theta}} = \mathbf{f}_\infty(\boldsymbol{\theta}).$$

Borkar-Meyn Theorem (2000) – Cont'd

Theorem

Under the stated assumptions,

1. $\{\boldsymbol{\theta}_t\}$ *is bounded almost surely.*
2. $\boldsymbol{\theta}_t \to \boldsymbol{\theta}^*$ *as* $t \to \infty$.

The a.s. boundedness of $\{\boldsymbol{\theta}_t\}$ is a *conclusion*, not a hypothesis.

Proof is based on the ODE method, which states that the sample paths of the iterates "converge" to the *deterministic* solution trajectories of the ODE $\dot{\boldsymbol{\theta}} = \mathbf{f}_\infty(\boldsymbol{\theta})$.

Method pioneered by Ljung (1974), Deveritskii and Fradkov (1974), Kushner-Clark (1978); see also Métivier-Priouret (1984).

Rather technical – worthwhile to find an easier proof.

Gladyshev's Theorem (1965)

Theorem

Assumptions (F), (N), but not (S). In addition

$$\inf_{\epsilon < \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|_2 < 1/\epsilon} \langle \boldsymbol{\theta} - \boldsymbol{\theta}^*, \mathbf{f}(\boldsymbol{\theta}) \rangle < 0, 0 < \epsilon < 1.$$

Then

1. *If* $\sum_{t=0}^{\infty} \alpha_t^2 < \infty$, *then* $\{\boldsymbol{\theta}_t\}$ *is bounded almost surely.*
2. *If in addition* $\sum_{t=0}^{\infty} \alpha_t = \infty$, *then* $\boldsymbol{\theta}_t \to \boldsymbol{\theta}^*$ *almost surely as* $t \to \infty$.

If $\mathbf{f}(\cdot)$ is continuous, the above is equivalent to

$$\langle \boldsymbol{\theta} - \boldsymbol{\theta}^*, \mathbf{f}(\boldsymbol{\theta}) \rangle < 0, \ \forall \boldsymbol{\theta} \neq \boldsymbol{\theta}^*,$$

or $\mathbf{f}(\cdot)$ is a "passive" function.

Beyond Gladyshev's Theorem

- Very easy proof, based on supermartingale theory.
- Clear "division of labor": Square-summability of step sizes gives stability, and divergence of step sizes gives convergence.

Can this approach be extended *beyond* passive functions?

Yes, by using "converse" Lyapunov theory (topic of this lecture).

Suppose $\boldsymbol{\theta}^*$ is the *only* solution of $\mathbf{f}(\boldsymbol{\theta}) = \mathbf{0}$. Then $\boldsymbol{\theta}^*$ is also the only equilibrium of the ODE $\dot{\boldsymbol{\theta}} = \mathbf{f}(\boldsymbol{\theta})$.

"Forward" Lyapunov theory: If there exists a function V with certain properties, then $\boldsymbol{\theta}^*$ has certain stability properties.

"Converse" Lyapunov theory: If the equilibrium $\boldsymbol{\theta}^*$ has certain stability properties, then there exists a suitable V.

Definition of Global Exponential Stability

Suppose $\mathbf{f}$ is globally Lipschitz continuous, and define $\mathbf{s} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ via: $\mathbf{s}(t, \boldsymbol{\theta})$ is the unique solution of

$$\frac{d\mathbf{s}(t, \boldsymbol{\theta})}{dt} = \mathbf{f}(\mathbf{s}(t, \boldsymbol{\theta})), \mathbf{s}(0, \boldsymbol{\theta}) = \boldsymbol{\theta}.$$

Suppose $\mathbf{f}(\boldsymbol{\theta}^*) = \mathbf{0}$. The equilibrium $\boldsymbol{\theta}^*$ is **globally exponentially stable (GES)** if there exist $\mu < \infty, \gamma > 0$ such that

$$\|\mathbf{s}(t, \boldsymbol{\theta}) - \boldsymbol{\theta}^*\|_2 \leq \mu \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|_2 \exp(-\gamma t), \ \forall t \geq 0, \ \forall \boldsymbol{\theta} \in \mathbb{R}^d.$$

IIT Hyderabad

Standard Converse Lyapunov Theorem for GES

Theorem

Suppose $\mathbf{f}$ *is globally Lipschitz continuous, that* $\boldsymbol{\theta}^*$ *is a GES equilibrium. Then the function* $V : \mathbb{R}^d \to \mathbb{R}_+$ *defined by*

$$V(\boldsymbol{\theta}) := \int_0^\infty \|\mathbf{s}(t, \boldsymbol{\theta})\|_2^2 dt.$$

satisfies the following: There exist $c_1, c_2, c_3 > 0$ *such that*

$$c_1 \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|_2^2 \leq V(\boldsymbol{\theta}) \leq c_2 \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|_2^2,$$

$$\dot{V}(\boldsymbol{\theta}) \leq -c_3 \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|_2^2, \ \forall \boldsymbol{\theta} \in \mathbb{R}^d,$$

where

$$\dot{V}(\boldsymbol{\theta}) = \langle \nabla V(\boldsymbol{\theta}), \mathbf{f}(\boldsymbol{\theta}) \rangle.$$

This is *not good enough* for current application.

IIT Hyderabad

New Converse Lyapunov Theorem for GES

Theorem

Suppose in addition that $\mathbf{f} \in \mathcal{C}^2$*, and that*[a]

$$\sup_{\boldsymbol{\theta} \in \mathbb{R}^d} \|\nabla^2 f_i(\boldsymbol{\theta})\|_S \cdot \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|_2 < \infty, \ \forall i \in [d].$$

Choose

$$0 < \kappa < \gamma, \frac{\ln \mu}{\gamma - \kappa} \leq T < \infty, V(\boldsymbol{\theta}) := \int_0^T e^{\kappa\tau} \|\mathbf{s}(\tau, \boldsymbol{\theta} - \boldsymbol{\theta}^*\|_2^2 \, d\tau$$

Then V *is* $\mathcal{C}^2$*, and also satisfies*

$$\|\nabla^2 V(\boldsymbol{\theta})\|_S \leq 2M, \ \forall \boldsymbol{\theta} \in \mathbb{R}^d.$$

[a] Here $\|\cdot\|_S$ denotes the spectral norm, and $[d] = \{1, \ldots, d\}$.

Builds on earlier work of Corless and Glielmo (1998).

Convergence of Stochastic Approximation

Theorem

Suppose (i) $\boldsymbol{\theta}^*$ is the only zero of $\mathbf{f}(\cdot)$, (ii) $\boldsymbol{\theta}^*$ is a GES equilibrium of $\dot{\boldsymbol{\theta}} = \mathbf{f}(\boldsymbol{\theta})$, (iii) $\mathbf{f}(\cdot)$ is globally Lipschitz continuous, and (iv)

$$\sup_{\boldsymbol{\theta} \in \mathbb{R}^d} \|\nabla^2 f_i(\boldsymbol{\theta})\|_S \cdot \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|_2 < \infty, \ \forall i \in [d].$$

Suppose further that $\{\boldsymbol{\xi}_t\}$ satisfies (N). Then

1. *If $\sum_{t=0}^{\infty} \alpha_t^2 < \infty$, then $\{\boldsymbol{\theta}_t\}$ is bounded almost surely.*
2. *If in addition $\sum_{t=0}^{\infty} \alpha_t = \infty$, then $\boldsymbol{\theta}_t \to \boldsymbol{\theta}^*$ almost surely as $t \to \infty$.*

We don't need

$$\mathbf{f}_\infty := \lim_{r \to \infty} \mathbf{f}(r\boldsymbol{\theta})/r,$$

but Borkar-Meyn (2000) don't need (iv).

Sketch of Proof

Construct a suitable Lyapunov function V *with a globally bounded Hessian*. Since

$$c_1 \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|_2^2 \leq V(\boldsymbol{\theta}) \leq c_2 \|\boldsymbol{\theta} - \boldsymbol{\theta}^*\|_2^2,$$

$\{\boldsymbol{\theta}_t\}$ is bounded if and only if $\{V(\boldsymbol{\theta}_t\}$ is bounded.

Define a new stochastic process

$$Z_t = a_t V(\boldsymbol{\theta}_t) + b_t,$$

and define constants a_t, b_t recursively so that

$$E(Z_{t+1} | \mathcal{F}_t) \leq Z_t \text{ a.s.},$$

By construction, $a_t \downarrow a_\infty \geq 1$ and $b_t \downarrow b_\infty \geq 0$. Hence $\{Z_t\}$ is a nonnegative supermartingale. So $Z_t \to \zeta$, some random variable. So $V(\boldsymbol{\theta}_t)$ is bounded, and so is $\{\boldsymbol{\theta}_t\}$ (almost surely).

Convergence of $\boldsymbol{\theta}_t$ to $\boldsymbol{\theta}^*$ follows via a separate argument.

Batch Stochastic Gradient Descent

- Another application of supermartingale methods (not directly related to converse Lyapunov theory) is "Batch Stochastic Gradient Descent" (BSGD) for convex optimization.
- It is widely used in Deep Learning because the dimension d is huge (though the problems are not convex).
- Suppose we wish to find a global minimum of a convex function $J : R^d \to \mathbb{R}$ using gradient descent:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \alpha_t \mathbf{e}_{S(t)} \circ [-\nabla J(\boldsymbol{\theta}_t) + \boldsymbol{\xi}_{t+1}],$$

  where $S(t) \subseteq [d]$ is the (randomly chosen) set of components to be updated at time t, $\mathbf{e}_{S(t)}$ equals 1 on $S(t)$ and 0 elsewhere, and $\circ$ denotes the Hadamard (componentwise) product.
- We update only $|S(t)|$ components at time t.

Convergence of Batch Stochastic Gradient Descent

- Using the present "supermartingale" approach, the convergence of BSGD can be established *even with noisy measurements*, provided each component of $\boldsymbol{\theta}$ is updated infinitely often.
- Currently available proofs assume noise-free measurements, and don't work with noisy measurements.

A preprint combining both applications will be up on arxiv very soon!

IIT Hyderabad

Future Work – 1: Actor-Critic Algorithms

- Actor-Critic algorithms in RL correspond to two time scale SA.
- The ODE method is *even more intricate* in this case; see e.g., Lakshminarayanan and Bhatnagar (2017).
- Converse Lyapunov theory for two time scale systems is fairly straight-forward.
- However, "off the shelf" theory may not work; we may need to invent new theory (as here).

Future Work – 2: TDL with Function Approximation

- For RL problems with large state space, Temporal Difference Learning (TDL) with function approximation is a popular approach.
- Paper by Tsitsiklis and Van Roy (1997) uses "ODE-like" methods.
- An alternative approach based on converse theory for "partial stability" may work.

Both approaches are under investigation.

IIT Hyderabad

A Shameless Plug

I am preparing a set of notes with the working title *Reinforcement Learning via Stochastic Approximation*. I will keep posting drafts on my website:

https://www.iith.ac.in/∼m_vidyasagar/RL/Gen/RL-Notes.pdf

Note: Current content is badly out of date.

Thank You!

IIT Hyderabad