Gaussian Approximation and Multiplier Bootstrap for Polyak-Ruppert Averaged Linear Stochastic Approximation with Applications to TD Learning

Sergey Samsonov, Eric Moulines, Qi-Man Shao,
Zhuo-Song Zhang, Alexey Naumov

HSE University

November 13, 2024

Linear Stochastic Approximation

- Given $\bar{\mathbf{A}} \in \mathbb{R}^{d \times d}$ and $\bar{\mathbf{b}} \in \mathbb{R}^d$, we aim at finding $\theta^\star \in \mathbb{R}^d$, which is a solution of
$$\bar{\mathbf{A}}\theta^\star = \bar{\mathbf{b}}\,.$$
- Our analysis is based on noisy observations $\{(\mathbf{A}(Z_n), \mathbf{b}(Z_n))\}_{n \in \mathbb{N}}$. Here $\mathbf{A} : \mathsf{Z} \to \mathbb{R}^{d \times d}$, $\mathbf{b} : \mathsf{Z} \to \mathbb{R}^d$ are measurable mappings.

LSA algorithm, Robbins and Monro [1951]

For a sequence of step sizes $\{\alpha_k\}$, and initialization θ_0, consider the sequences of estimates $\{\theta_n\}_{n \in \mathbb{N}}, \{\bar{\theta}_n\}_{n \geq 2}$ given by

$$\theta_k = \theta_{k-1} - \alpha_k \{\mathbf{A}(Z_k)\theta_{k-1} - \mathbf{b}(Z_k)\}\,, \quad k \geq 1,$$
$$\bar{\theta}_n = n^{-1} \sum_{k=n}^{2n-1} \theta_k\,, \quad n \geq 2\,.$$

I.I.D. Noise

Sequence $\{Z_k\}_{k \in \mathbb{N}}$ is an i.i.d. sequence taking values in a state space $(\mathsf{Z}, \mathcal{Z})$ with distribution π satisfying $\mathbb{E}[\mathbf{A}(Z_1)] = \bar{\mathbf{A}}$ and $\mathbb{E}[\mathbf{b}(Z_1)] = \bar{\mathbf{b}}$;

We write $\mathbf{A}_k$ instead of $\mathbf{A}(Z_k)$, and $\mathbf{b}_k$ instead of $\mathbf{b}(Z_k)$, respectively.

Normal approximation

CLT

Under appropriate conditions on the step sizes $\{\alpha_k\}_{k\in\mathbb{N}}$ and noisy observations $\{\mathbf{A}(Z_k)\}_{k\in\mathbb{N}}$, it is known that

$$\sqrt{n}(\bar{\theta}_n - \theta^\star) \xrightarrow{d} \mathcal{N}(0, \Sigma_\infty)\,,$$

where Σ_∞ is the asymptotic covariance matrix, see e.g. Fort [2015].

Berry-Esseen bounds

Our aim is to obtain the non-asymptotic type bounds for

$$\rho_n^{\mathrm{Conv}} := \sup_{B\in\mathrm{Conv}(\mathbb{R}^d)} \left| \mathbb{P}\big(\sqrt{n}(\bar{\theta}_n - \theta^\star) \in B\big) - \mathbb{P}(\Sigma_\infty^{1/2}\eta \in B) \right| \,,$$

where $\eta \sim \mathcal{N}(0, \mathrm{I}_d)$, and $\mathrm{Conv}(\mathbb{R}^d)$ is a set of convex sets in $\mathbb{R}^d$.

Linear Stochastic Approximation

- Let $\{Z_k\}_{k \in \mathbb{N}}$ be an i.i.d.sequence and consider the recurrence

$$\theta_k = \theta_{k-1} - \alpha_k \{\mathbf{A}(Z_k)\theta_{k-1} - \mathbf{b}(Z_k)\} \tag{1}$$

- Set

$$\tilde{\mathbf{A}}(z) = \mathbf{A}(z) - \bar{\mathbf{A}}\,, \quad \tilde{\mathbf{b}}(z) = \mathbf{b}(z) - \bar{\mathbf{b}}\,,$$

  and introduce

$$\varepsilon(z) = \mathbf{A}(z)\theta^\star - \mathbf{b}(z)\,, \quad \Sigma_\varepsilon = \mathbb{E}[\varepsilon(Z)\varepsilon(Z)^\top]\,.$$

Assumptions

Assumption A1

- Sequence $\{Z_n\}_{n\in\mathbb{N}}$ is a sequence of i.i.d. random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution π.
- It holds that

$$\mathsf{C}_A = \sup_{z\in\mathsf{Z}} \|\mathbf{A}(z)\| \vee \sup_{z\in\mathsf{Z}} \|\tilde{\mathbf{A}}(z)\| < \infty$$

and $-\bar{\mathbf{A}}$ is Hurwitz. Moreover,

$$\int_{\mathsf{Z}} \mathbf{A}(z)\mathrm{d}\pi(z) = \bar{\mathbf{A}}, \quad \int_{\mathsf{Z}} \mathbf{b}(z)\mathrm{d}\pi(z) = \bar{\mathbf{b}}, \quad \|\varepsilon\|_\infty = \sup_{z\in\mathsf{Z}} \|\varepsilon(z)\| < +\infty\,.$$

- For the noise covariance matrix

$$\Sigma_\varepsilon = \int_{\mathsf{Z}} \varepsilon(z)\varepsilon(z)^\top \mathrm{d}\pi(z) \tag{2}$$

it holds that its smallest eigenvalue is bounded away from 0, that is,

$$\lambda_{\mathsf{min}} := \lambda_{\mathsf{min}}(\Sigma_\varepsilon) > 0\,. \tag{3}$$

Normal approximation for LSA-PR

Theorem (Rates of normal approximation for LSA)

Assume A1, let $\alpha_k = c_0/k^{\gamma}$, $\gamma \in [1/2; 1)$, and let n be large enough. Then the following bound holds:

$$\rho_n^{\mathrm{Conv}} \lesssim \frac{d^{1/2}\|\varepsilon\|_\infty^3}{\lambda_{\min}^{3/2}\sqrt{n}} + \frac{C_4}{\lambda_{\min}} \exp\left\{-\frac{c_0 a n^{1-\gamma}}{2(1-\gamma)}\right\}\|\theta_0 - \theta^\star\| + \frac{C_1}{\lambda_{\min} n^{(1-\gamma)/2}} + \frac{C_2}{\lambda_{\min} n^{1-\gamma}} + \frac{C_3}{\lambda_{\min} n^{\gamma/2}}\,, \quad (4)$$

where C_1, C_2, C_3, C_4 are problem-specific constants.

Setting here $\alpha_k = c_0/\sqrt{k}$, we obtain that

$$\rho_n^{\mathrm{Conv}} \lesssim n^{-1/4} + \Delta_1 \exp\{-c_0 a \sqrt{n}\}\|\theta_0 - \theta^\star\|\,.$$

Comparison

- Srikant [2024] considers TD learning with Markov noise and obtained $\rho_n^{\mathrm{Conv}} \lesssim n^{-1/8}$;
- Anastasiou et al. [2019] consider smooth Wasserstein distance and obtained $\mathrm{d}_K(\sqrt{n}(\bar{\theta}_n - \theta^\star), Y) \lesssim n^{-1/6}$.

Multiplier bootstrap for LSA

- We aim to provide confidence intervals for $\sqrt{n}(\bar{\theta}_n - \theta^\star)$ without directly estimating the asymptotic covariance matrix Σ_∞, following Fang et al. [2018];
- Let $\mathcal{W}^{2n} = \{W_\ell\}_{1 \le \ell \le 2n}$ - i.i.d. random variables, independent of $\mathcal{Z}^{2n} = \{Z_\ell\}_{1 \le \ell \le 2n}$, where $\mathbb{E}[W_1] = 1$, $\operatorname{Var}[W_1] = 1$;
- We write $\mathbb{P}^{\mathsf{b}} = \mathbb{P}(\cdot|\mathcal{Z}^{2n})$ and $\mathbb{E}^{\mathsf{b}} = \mathbb{E}(\cdot|\mathcal{Z}^{2n})$. Generate M independent samples $(w_n^\ell, \ldots, w_{2n}^\ell)$, $1 \le \ell \le M$ and consider M recursively updated perturbed LSA estimates

$$
\begin{aligned}
\theta_k^{\mathsf{b},\ell} &= \theta_{k-1}^{\mathsf{b},\ell} - \alpha_k w_k^\ell \{\mathbf{A}(Z_k)\theta_{k-1}^{\mathsf{b},\ell} - \mathbf{b}(Z_k)\}\,, \quad k \ge n+1\,, \quad \theta_n^{\mathsf{b},\ell} = \theta_n\,, \\
\bar{\theta}_n^{\mathsf{b},\ell} &= n^{-1} \sum_{k=n}^{2n-1} \theta_k^{\mathsf{b},\ell}\,, \quad n \ge 1\,.
\end{aligned}
$$

We further use a short notation $\bar{\theta}_n^{\mathsf{b}}$ for $\bar{\theta}_n^{\mathsf{b},1}$.

Multiplier Bootstrap validity

Theorem (Bootstrap validity for LSA)

Assume A1, let $\alpha_k = c_0/\sqrt{k}$, and let n be large enough. Then with $\mathbb{P}$ – probability at least $1 - 6/n$ it holds that

$$\sup_{B \in \mathrm{Conv}(\mathbb{R}^d)} |\mathbb{P}^{\mathsf{b}}(\sqrt{n}(\bar{\theta}_n^{\mathsf{b}} - \bar{\theta}_n) \in B) - \mathbb{P}(\sqrt{n}(\bar{\theta}_n - \theta^\star) \in B)| \lesssim \frac{C_5 \log n}{\lambda_{\min} n^{1/4}} + \frac{C_6 \sqrt{d} \log n}{\lambda_{\min} \sqrt{n}} + \frac{C_7 \mathrm{e}^{-(c_0/2)a\sqrt{n}}}{\lambda_{\min}} \|\theta_0 - \theta^\star\|$$

where C_5, C_6, C_7 are problem-specific constants.

Thank you!

References I

Andreas Anastasiou, Krishnakumar Balasubramanian, and Murat A. Erdogdu. Normal approximation for stochastic gradient descent via non-asymptotic rates of martingale CLT. In Alina Beygelzimer and Daniel Hsu, editors, *Proceedings of the Thirty-Second Conference on Learning Theory*, volume 99 of *Proceedings of Machine Learning Research*, pages 115–137. PMLR, 25–28 Jun 2019. URL https://proceedings.mlr.press/v99/anastasiou19a.html.

Yixin Fang, Jinfeng Xu, and Lei Yang. Online bootstrap confidence intervals for the stochastic gradient descent estimator. *Journal of Machine Learning Research*, 19(78):1–21, 2018. URL http://jmlr.org/papers/v19/17-370.html.

G. Fort. Central limit theorems for stochastic approximation with controlled Markov chain dynamics. *ESAIM: PS*, 19:60–80, 2015.

Herbert Robbins and Sutton Monro. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951.

R Srikant. Rates of convergence in the central limit theorem for markov chains, with an application to TD learning. *arXiv preprint arXiv:2401.15719*, 2024.