Lecture-13: Perturbation theory

1 Perturbation theory for proof of convergence

Proposition 1.1. *From the local exponential stability condition, there exist positive constants $k, c_\ell, c_u, c_p, \phi$ and a Lyapunov function $V : \mathbb{R}^d \to \mathbb{R}_+$ such that*

$$c_\ell \|Nh(t)\|^2 \leqslant V(Nh(t)) \leqslant c_u \|Nh(t)\|^2, \tag{1}$$

$$\|\nabla V(Nh(t))\| \leqslant Nc_p \|h(t)\|, \tag{2}$$

$$\dot{V}(Nh) \leqslant -\phi V(Nh), \text{ for } \|\Phi_t(a)\| \leqslant k. \tag{3}$$

Proposition 1.2. *It follows from the global asymptotic stability condition that for any $k > 0$, there exists $t_k \in \mathbb{R}_+$ such that $\|\Phi_t(a)\| \leqslant k$ for all $t \geqslant t_k$.*

Proof. From the global asymptotic stability, we have for any initial point $a \in \mathcal{M}(\mathcal{Z})$ and $\epsilon > 0$, there exists $t_\epsilon > 0$ such that $\|\Phi_t(a) - a^*\| < \epsilon$ for all $t > t_\epsilon$. Since $\|\Phi_t(a)\| \leqslant \|\Phi_t(a) - a^*\| + \|a^*\|$, the result follows. □

Corollary 1.3. *Under local exponential stability and global asymptotic stability conditions on the mean-field model, there exists a Lyapunov function $V : \mathbb{R}^d \to \mathbb{R}_+$ such that $V(Nh(t)) \leqslant V(Nh(t_k))e^{-\phi(t-t_k)}$ for all $t \geqslant t_k$.*

Proof. Recall that at time $t = t_k$, we have $V(Nh(t)) \leqslant V(Nh(t_k))$. Further for times $t \geqslant t_k$, we have $\|\Phi_t(a) - a^*\| \leqslant k$ and hence $(\dot{V}(Nh(t)) + \phi V(Nh(t)))e^{-\phi(t-t_k)} \leqslant 0$. □

Lemma 1.4. *Let $d \triangleq |\mathcal{Z}|$, and consider the Mckean-Vlasov equation $\dot{a} = f(a)$ for distribution process $a : \mathbb{R}_+ \to \mathcal{M}(\mathcal{Z})$, that satisfies the following two conditions.*

Condition 1. **Lipschitz partial derivatives.** *First order partial derivatives $\frac{\partial f_z(a)}{\partial a_w}(a)$ exist and are Lipschitz for all $w, z \in \mathcal{Z}$.*

Condition 2. **Stability.** *The mean-field model is globally asymptotically stable and locally exponentially stable.*

Consider $a, b \in \mathcal{M}(\mathcal{Z})$ such that $\|b - a\| \leqslant \frac{\tilde{c}}{N}$ for some $\tilde{c} > 0$, then there exist positive constants c and σ independent of N such that

$$\|N(b-a)\nabla\Phi_t(a)\| \leqslant ce^{-\sigma t}, \qquad \|\Phi_t(b) - \Phi_t(a) - (b-a)\nabla\Phi_t(a)\|_1 \leqslant \frac{c}{N^2}e^{-\sigma t}.$$

Proof. Recall that $e(t, a, b) = \Phi_t(b) - \Phi_t(a) - (b-a)\nabla\Phi_t(a)$ and $h(t, a, b) = (b-a)\nabla\Phi_t(a)$. We will fix a, b and simplify the notation to $e(t)$ and $h(t)$. Recall that

$$\nabla_a f(\Phi_t(a)) = \sum_{i,j=1}^{d} e_j^T \frac{\partial f_i(\Phi_t(a))}{\partial a_j} e_i = \sum_{i,j=1}^{d} e_j^T \sum_{k=1}^{d} \frac{\partial f_i(\Phi_t(a))}{\partial \Phi_t(a)_k} \frac{\partial \Phi_t(a)_k}{\partial a_j} e_i = \nabla_a \Phi_t(a) \nabla_{\Phi_t(a)} f(\Phi_t(a)).$$

We can write the time derivative of h as

$$\dot{h} = (b-a)\nabla_a f(\Phi_t(a)) = (b-a)\nabla_a \Phi_t(a) \nabla_{\Phi_t(a)} f(\Phi_t(a)).$$

Therefore, we can write the time derivative of error e as

$$\dot{e} = f(\Phi_t(b)) - f(\Phi_t(a)) - h\nabla_{\Phi_t(a)} f(\Phi_t(a)).$$

From the definition of error, we observe that $e(0) = 0$. We observe that

$$\frac{d}{dt}\|h\|^2 = 2\left\langle h, \dot{h} \right\rangle = 2\left\langle h, h\nabla f(\Phi_t(a)) \right\rangle.$$

1. Since h is defined on a bounded set and first-order partial derivatives are Lipschitz, the first-order partial derivatives are bounded, as there exists a constant $\kappa > 0$ such that $\left|\frac{d}{dt}\|h\|^2\right| \leqslant \kappa \|h\|^2$. This together with the fact that $h(0) = b - a$ implies that $\|h(t)\| \leqslant \|b-a\| e^{\frac{\kappa}{2}t}$. From local exponential stability condition, we obtain that $V(Nh(t)) \leqslant e^{-\phi(t-t_k)} V(Nh(t_k))$ for $t \geqslant t_k$. Thus, we can write

$$c_\ell \|Nh(t)\|^2 \leqslant V(Nh(t)) \leqslant c_u e^{-\phi(t-t_k)} \|Nh(t_k)\|^2.$$

Thus, we can write

$$\|Nh(t)\| \leqslant \sqrt{\frac{c_u}{c_\ell}} e^{-\frac{\phi}{2}(t-t_k)} \|Nh(t_k)\| \leqslant \sqrt{\frac{c_u}{c_\ell}} e^{\frac{(\phi+\kappa)}{2}t_k} \|N(b-a)\| e^{-\frac{\phi}{2}t}. \tag{4}$$

2. Fix $z \in \mathcal{Z}$ where $d \triangleq |\mathcal{Z}|$. From the mean-value theorem for continuous function $f : \mathcal{M}(\mathcal{Z}) \to \mathcal{M}(\mathcal{Z})$,

$$f_z(\Phi_t(b)) - f_z(\Phi_t(a)) = (\Phi_t(b) - \Phi_t(a)) \nabla f_z(\Phi_t(a) + (\Phi_t(b) - \Phi_t(a))\xi_z)^T$$

for some $\xi \in [0,1]^{\mathcal{Z}}$. Recall that $e(t) + \frac{1}{N} Nh(t) = \Phi_t(b) - \Phi_t(a)$. Thus, we have

$$\dot{e}_z = e \nabla f_z(\Phi_t(a) + (e+h)\xi_z)^T + \frac{1}{N} Nh(\nabla f_z(\Phi_t(a) + (e+h)\xi_z)^T - \nabla f_z(\Phi_t(a))^T).$$

Since the first-order partial-derivatives of f are Lipschitz, there exists $L_z, M_z > 0$ such that $\|\nabla f_z(b) - \nabla f_z(a)\| \leqslant L_z \|b-a\|$ and $\sup_{z \in \mathcal{Z}, b \in \mathbb{R}^{\mathcal{Z}}} \|\nabla f_z(b)\| < M_z$. Defining a constant $B \triangleq \sup_{z \in \mathcal{Z}, N \in \mathbb{N}} (M_z + \frac{L_z}{N}\|Nh\|) \vee L_z \|Nh\|$ independent of N, we get

$$|\dot{e}_z| \leqslant \|e\| M_z + \frac{L_z}{N}\|Nh\| (\|e\| + \frac{1}{N}\|Nh\|) \leqslant \|e\| (M_z + \frac{L_z}{N}\|Nh\|) + \frac{L_z}{N^2}\|Nh\| \leqslant B\|e\| + \frac{B}{N^2}.$$

It follows that $\|\dot{e}\| \leqslant B\sqrt{d}\|e\| + \frac{B\sqrt{d}}{N^2}$, and since $e(0) = 0$, we obtain $\|e(t)\| \leqslant \frac{1}{N^2}(e^{B\sqrt{d}t} - 1)$. We can write the time derivative of Lyapunov function as

$$\dot{V}(e) = \langle \dot{e}, \nabla V(e) \rangle = \langle f(\Phi_t(b)) - f(\Phi_t(a)) - h \nabla f(\Phi_t(a)), \nabla V(e) \rangle.$$

Consider $t \geqslant t_k$ such that $\|\Phi_t(a)\| \leqslant k$. From (3), we know that $\dot{V}(Nh) = \langle \nabla V(Nh), Nh \nabla f(\Phi_t(a)) \rangle \leqslant -\phi V(Nh)$. Replacing Nh by e in this equation, we get $\langle \nabla V(e), e \nabla f(\Phi_t(a)) \rangle \leqslant -\phi V(e)$. Therefore, we obtain

$$\dot{V}(e) \leqslant -\phi V(e) + \sum_{z \in \mathcal{Z}} (f_z(\Phi_t(b)) - f_z(\Phi_t(a)) - (e+h)\nabla f_z(\Phi_t(a))^T) \frac{\partial V(e)}{\partial e_z}.$$

Substituting $e + h = \Phi_t(b) - \Phi_t(a)$ and $\xi_z \in [0,1]$ from mean-value theorem in the above equation and using the L_z-Lipschitz property for ∇f_z, we obtain

$$\dot{V}(e) \leqslant -\phi V(e) + \sum_{z \in \mathcal{Z}} (e+h) \Big[\nabla f_z(\Phi_t(a) + (e+h)\xi_z)^T - \nabla f_z(\Phi_t(a))^T\Big] \frac{\partial V(e)}{\partial e_z} \leqslant -\phi V(e) + \|e+h\|^2 \langle L, \nabla V(e) \rangle.$$

Replacing Nh by e in (2), we obtain $\|\nabla V(e)\| \leqslant c_p \|e\|$. Taking L as a constant for all $z \in \mathcal{Z}$ and applying Hölder's inequality to $\langle L, \nabla V(e) \rangle$ and from Minkowski inequality that implies $\|e+h\|^2 \leqslant 2(\|e\|^2 + \|h\|^2)$, we get

$$\dot{V}(e) \leqslant -\phi V(e) + 2L\sqrt{d} c_p (\|e\|^3 + \frac{1}{N^2}\|Nh\|^2 \|e\|).$$

Substituting Nh by e in (1), we obtain that $c_\ell \|e\|^2 \leqslant V(e) \leqslant c_u \|e\|^2$, and thus

$$\dot{V}(e) \leqslant -\Big(\phi - \frac{2Lc_p\sqrt{d}}{c_\ell}\|e\|\Big) V(e) + \frac{2L\sqrt{d}c_p}{\sqrt{c_\ell}} \frac{1}{N^2}\|Nh\|^2 \sqrt{V(e)}.$$

Defining a constant $\hat{c} \triangleq L\sqrt{d}\tilde{c}^2 \frac{c_u c_p}{(c_\ell)^{3/2}} e^{(\phi+\kappa)t_k}$ it follows from (4) that

$$\dot{V}(e) \leqslant -\Big(\phi - \frac{2Lc_p\sqrt{d}}{c_\ell}\|e\|\Big)V(e) + \frac{2\hat{c}}{N^2}\sqrt{V(e)}e^{-\phi t}.$$

Defining $W \triangleq \sqrt{V}$ and considering time t_ℓ such that $\|e\| \leqslant \frac{\phi c_\ell}{4Lc_p\sqrt{d}}$ for all $t \leqslant t_\ell$, we obtain

$$\dot{W}(e) \leqslant -\frac{\phi}{4}W(e) + \frac{\hat{c}}{N^2}e^{-\phi t}, \text{ for all } t \geqslant t_\ell \wedge t_k.$$

We observe that $\frac{d}{dt}(W(e(t))e^{\frac{\phi}{4}t}) = (\dot{W}(e) + \frac{\phi}{4}W(e))e^{\frac{\phi}{4}t} \leqslant \frac{\hat{c}}{N^2}e^{\frac{\phi}{4}t-\phi t}$ for all $t \geqslant t_k$. Integrating both sides over $t \geqslant t_k$, we obtain

$$W(e(t)) \leqslant W(e(t_k))e^{-\frac{\phi}{4}(t-t_k)} + \frac{\hat{c}}{N^2}\int_{t_k}^{t} e^{-\frac{\phi}{4}(t-\tau)-\phi\tau}d\tau \leqslant \Big[W(e(t_k))e^{\frac{\phi}{4}t_k} + \frac{4\hat{c}}{3\phi N^2}\Big]e^{-\frac{\phi}{4}t} \leqslant \frac{C}{N^2}e^{-\frac{\phi}{4}t}.$$

Since $W = \sqrt{V}$, substituting Nh by e in (1), we obtain

$$\|e\|_1 \leqslant \sqrt{d}\,\|e\|_2 \leqslant \frac{\sqrt{d}}{\sqrt{c_\ell}}W(e) \leqslant \frac{C\sqrt{d}}{\sqrt{c_\ell}N^2}e^{-\frac{\phi}{4}t}.$$

□

Corollary 1.5. *Under the conditions in Lemma 1.4, we have*

$$\int_{t\in\mathbb{R}_+} \|\langle\nabla\Phi_t(a), N(b-a)\rangle\|^2\,dt \leqslant \frac{c^2}{2\sigma}, \qquad \int_{t\in\mathbb{R}_+} \|e(t,a,b)\|_1 \leqslant \frac{c}{\sigma N^2}.$$

Theorem 1.6. *The empirical distribution processes of a family of CTMCs $((X^N : \Omega \to \mathcal{Z}^N) : N \in \mathbb{N})$ converge to the equilibrium point a^* of the mean-field model in the mean-square sense with rate $O(h_1(N) \vee h_2(N))$, i.e.,*

$$\mathbb{E}_{\pi^{A^N(\infty)}}\left\|A^N(\infty) - a^*\right\|^2 = O\Big(h_1(N) \vee h_2(N)\Big),$$

when the following conditions hold.

Condition 1. **Asymptotically accurate mean-field model**.

$$\mathbb{E}_{\pi^{A^N(\infty)}}\left\|f(A^N(\infty)) - \sum_{b:b\neq A^N(\infty)} Q^{A^N}_{A^N(\infty),b}(b - A^N(\infty))\right\| = O(h_1(N)).$$

Condition 2. **Bounded mean state transitions**. $\mathbb{E}_{\pi^{A^N(\infty)}}\sum_{b:b\neq A^N(\infty)} Q^{A^N}_{A^N(\infty),b}\left\|b - A^N(\infty)\right\|^2 = O(h_2(N))$.

Condition 3. **Bounded state difference**. $\max_{a,b:Q^{A^N}_{a,b}>0}\|b-a\| = o(1)$.

Condition 4. **Lipschitz partial derivatives.** *The first order partial derivatives $\frac{\partial f_w}{\partial a_z}$ exist and are Lipschitz for all $w, z \in \mathcal{Z}$.*

Condition 5. **Stability.** *The mean-field model is globally asymptotically stable and is locally exponentially stable.*

Proof. Recall that $h(t) = (b-a)\nabla\Phi_t(a)$ and $g(a) = -\int_{t\in\mathbb{R}_+}\|\Phi_t(a) - a^*\|^2\,dt$ is the solution to the Poisson equation, where the integral is finite and hence we can exchange integral and derivative to obtain $\langle\nabla g(a), b-a\rangle = -\int_{t\in\mathbb{R}_+} 2\langle(\Phi_t(a) - a^*), h(t)\rangle\,dt$. Taking $N(b-a) = e_w - e_z$ for $z, w \in \mathcal{Z}$, and recalling that $\|Nh(t)\| \leqslant ce^{-\sigma t}$ for some $c, \sigma > 0$, we observe that

$$|\nabla g(a)| \leqslant \int_{t\in\mathbb{R}_+} 2\,\|\Phi_t(a) - a^*\|\,\|Nh(t)\|\,dt < \infty.$$

Thus, we can conclude that $\|\nabla g(a)\| \leqslant K$ for all $a \in \mathcal{M}(\mathcal{Z})$. We further recall that we can write

$$\mathbb{E}_{\pi^{A^N(\infty)}} \left\|A^N(\infty) - a^*\right\|^2 = \mathbb{E}_{\pi^{A^N(\infty)}} \Bigg[\Bigg\langle \nabla g(A^N(\infty)), \Big(f(A^N(\infty)) - \sum_{b: b \neq A^N(\infty)} Q^{A^N}_{A^N(\infty), b}(b - A^N(\infty))\Big) \Bigg\rangle - \sum_{b: b \neq A^N(\infty)} Q^{A^N}_{A^N(\infty), b} \Big(g(b) - g(A^N(\infty)) - \big\langle \nabla g(A^N(\infty)), (b - A^N(\infty)) \big\rangle \Big) \Bigg].$$

From the Hölder's inequality applied to the inner product, triangle inequality, and the fact that $\sup_a \|\nabla g(a)\| < K$, we obtain

$$\mathbb{E}_{\pi^{A^N(\infty)}} \left\|A^N(\infty) - a^*\right\|^2 \leqslant K \mathbb{E}_{\pi^{A^N(\infty)}} \left\| f(A^N(\infty)) - \sum_{b: b \neq A^N(\infty)} Q^{A^N}_{A^N(\infty), b}(b - A^N(\infty)) \right\| + \mathbb{E}_{\pi^{A^N(\infty)}} \Bigg[\sum_{b: b \neq A^N(\infty)} Q^{A^N}_{A^N(\infty), b} \left| g(b) - g(A^N(\infty)) - \big\langle \nabla g(A^N(\infty)), (b - A^N(\infty)) \big\rangle \right| \Bigg].$$

From proof of Lemma 1.4, we observe that

$$\int_{t \in \mathbb{R}_+} \|e(t)\|_1 = O(\|b - a\|^2), \qquad \int_{t \in \mathbb{R}_+} \|Nh(t)\|_1 = O(\|b - a\|).$$

Further, we can show that there exists a constant b independent of N such that

$$|g(b) - g(a) - \langle \nabla g(a), (b - a) \rangle| \leqslant b \int_{t \in \mathbb{R}_+} \|e(t)\|_1 \, dt + \frac{1}{N^2} \int_{t \in \mathbb{R}_+} \|h(t)\|^2 \, dt = O(\|b - a\|^2).$$

□

Remark 1. Relaxing the perfect mean-field model assumption implies that the CTMC A^N is no longer required to be density dependent. Furthermore, relaxing the bounded state transition condition makes the result applicable to CTMCs for which the number of *jumps* during a transition is a function of N instead of a constant. The rate of convergence in these cases depends on the distance between the generator of the CTMC with N particles and the mean-field model, and depends on the mean-square jump size of the CTMC A^N at steady-state.