Fluid limits for Markov chains II

James Norris

Durham Symposium on Markov Processes, Mixing Times and Cutoff 2017
• Martingales
• Martingale inequalities
• Gronwall
• Localization
• Long-time estimates for stable flows
• Averaging over fast variables
Recap

$(\xi_t)_{t \geq 0}$ Markov chain in $E$, $x : E \to V$, $X_t = x(\xi_t)$

$$\beta(\xi) = \lim_{t \to 0} t^{-1} \mathbb{E}(x(\xi_t) - x(\xi_0) | \xi_0 = \xi) = \int_E \{x(\eta) - x(\xi)\} q(\xi, d\eta)$$

$$\alpha(\xi) = \lim_{t \to 0} t^{-1} \mathbb{E}(|x(\xi_t) - x(\xi_0)|^2 | \xi_0 = \xi) = \int_E |x(\eta) - x(\xi)|^2 q(\xi, d\eta)$$

We look for conditions under which $X_t$ is close to the solution of the differential equation $\dot{x}_t = b(x_t)$ with high probability.

We will assume that the transition kernel $q$ is bounded.

We assume for now that $V$ has an inner product. We used this in defining $\alpha$. 
Martingales

For any bounded measurable function \( f : E \to \mathbb{R} \), the following processes are martingales

\[
M_t = f(\xi_t) - f(\xi_0) - \int_0^t Qf(\xi_s)ds, \quad N_t = M_t^2 - \int_0^t \mathcal{E}(f)(\xi_s)ds
\]

where

\[
Qf(\xi) = \int_E \{ f(\eta) - f(\xi) \} q(\xi, d\eta), \quad \mathcal{E}(f)(\xi) = \int_E |f(\eta) - f(\xi)|^2 q(\xi, d\eta).
\]

Write \( x = (x_i) \) in some orthonormal basis of \( E \). Then

\[
\beta = Qx, \quad \alpha = \sum_i \alpha_i, \quad \text{where} \quad \alpha_i = \mathcal{E}(x_i)
\]

so the following processes are martingales

\[
M_t = X_t - X_0 - \int_0^t \beta(\xi_s)ds, \quad N_t = |M_t|^2 - \int_0^t \alpha(\xi_s)ds.
\]
Martingale inequalities

If the diffusivity \( \alpha \) is small, then so is the martingale \( M \).

- **Doob’s \( L^2 \) inequality**

  For all stopping times \( T \),

  \[
  \mathbb{E} \left( \sup_{t \leq T} |M_t|^2 \right) \leq 4 \mathbb{E} \int_0^T \alpha(\xi_s) ds.
  \]

- **Exponential martingale inequality**

  Assume that the jumps of the \( i \)th coordinate process are bounded uniformly by \( \Delta_i \). Then, for all \( i \), all stopping times \( T \) and all \( \delta, \tau \in (0, \infty) \),

  \[
  \mathbb{P} \left( \sup_{t \leq T} |M^i_t| > \delta \text{ and } \int_0^T \alpha_i(\xi_s) ds \leq \tau \right) \leq 2e^{-\delta^2/(2A\tau)}
  \]

  where \( A \in [1, \infty) \) is given by \( A \log A = \delta \Delta_i / \tau \).
Gronwall

Subtract the equations

\[ X_t = X_0 + M_t + \int_0^t \beta(\xi_s) \, ds, \quad x_t = x_0 + \int_0^t b(x_s) \, ds \]

to obtain

\[ |X_t - x_t| \leq |X_0 - x_0| + |M_t| + \left| \int_0^t (\beta(\xi_s) - b(x_s)) \, ds \right|. \]

Fix \( T \geq 0 \) and \( \delta > 0 \). Set \( \varepsilon = e^{-K T} \delta/3 \) where \( K \) is a Lipschitz constant for \( b \). Consider the events \( \Omega_0 = \{ |X_0 - x_0| \leq \varepsilon \} \),

\[ \Omega_1 = \left\{ \int_0^T |\beta(\xi_t) - b(X_t)| \, dt \leq \varepsilon \right\}, \quad \Omega_2 = \left\{ \sup_{t \leq T} |M_t| \leq \varepsilon \right\}. \]

On \( \Omega_0 \cap \Omega_1 \cap \Omega_2 \), we have, for all \( t \leq T \),

\[ |X_t - x_t| \leq 3\varepsilon + K \int_0^t |X_s - x_s| \, ds \]

so \( |X_t - x_t| \leq \delta \) by Gronwall’s lemma.
Localization

The *tube argument* allows to localize these estimates.

Assume that \((x_t : t \in [0, t_0])\) is continuous, with

\[ x_t = x_0 + \int_0^t b(x_s) \, ds, \quad t \in [0, t_0]. \]

Fix an open set \(U\) containing every point at distance at most \(\delta\) from \(\{x_t : t \in [0, t_0]\}\). Assume only that \(|\nabla b| \leq K\) in \(U\).

In the Gronwall argument, take

\[ T = \inf \{ t \geq 0 : |X_t - x_t| \notin U \} \land t_0 \]

to see that, on \(\Omega_0 \cap \Omega_1 \cap \Omega_2\),

\[ \sup_{t \leq T} |X_t - x_t| \leq \delta. \]

In particular \(X_T \in U\), so \(T = t_0\). So, on the same event,

\[ \sup_{t \leq t_0} |X_t - x_t| \leq \delta. \]
Long-time estimates for stable flows

Recall

\((\xi_t)_{t \geq 0}\) is a Markov chain in \(E\), \(x : E \to V\), \(X_t = x(\xi_t)\)

\[\beta(\xi) = \lim_{t \to 0} t^{-1} \mathbb{E}(x(\xi_t) - x(\xi_0) | \xi_0 = \xi) = \int_E (x(\eta) - x(\xi)) q(\xi, d\eta)\].

We look for conditions under which \(X_t\) is close to the solution of the differential equation \(\dot{x}_t = b(x_t)\) with high probability.

Take \(E = V = \mathbb{R}^d\) and \(b = \beta\) and \(X_0 = x_0\). Here we will suppose that the associated flow of diffeomorphisms

\[\dot{\phi}_t(x) = b(\phi_t(x)), \quad \phi_0(x) = x\]

has the following stability properties: for some \(\lambda > 0\) and \(B < \infty\),

\[|\nabla \phi_t(x)y| \leq e^{-\lambda t} |y|, \quad |\nabla^2 \phi_t(x)(y, y)| \leq B e^{-\lambda t} |y|^2.\]

This forces \(b\) to have a stable fixed point. Something close to \(b(x) = Ax\) with \(\langle Ax, x \rangle \leq -\lambda |x|^2\) will work.
Long-time estimates for stable flows

We interpolate from $x_T$ to $X_T$ using $(\phi_{T-t}(X_t) : t \in [0, T])$.

The following process is a martingale

$$M_t = \phi_{T-t}(X_t) - \phi_T(X_0) - \int_0^t \rho(T - s, X_s)ds$$

where

$$\rho(s, x) = \int_E \{\phi_s(y) - \phi_s(x) - \nabla \phi_s(x)(y - x)\} q(x, dy).$$

Moreover

$$\mathbb{E}(|M_t|^2) = \int_0^t \sigma(T - s, X_s)ds$$

where

$$\sigma(s, x) = \int_E \{\phi_s(y) - \phi_s(x)\}^2 q(x, dy).$$
Long-time estimates for stable flows

Now

\[ X_T - \phi_T(x_0) = M_T + \int_0^T \rho(T - t, X_t) \, dt \]

and from our stability assumptions

\[ \sigma(s, x) \leq e^{-2\lambda s} \alpha(x), \quad |\rho(s, x)| \leq B e^{-\lambda s} \alpha(x)/2 \]

so

\[ \mathbb{E}(|M_T|^2) \leq \|\alpha\|_\infty \int_0^T e^{-2\lambda(T-s)} \, ds \leq \frac{\|\alpha\|_\infty}{2\lambda} \]

and

\[ \left| \int_0^T \rho(T - s, X_s) \, ds \right| \leq \frac{B \|\alpha\|_\infty}{2\lambda}. \]

So we get a uniform-in-time estimate

\[ \|X_T - \phi_T(x_0)\|_2 \leq \sqrt{\frac{\|\alpha\|_\infty}{2\lambda}} + \frac{B \|\alpha\|_\infty}{2\lambda}. \]
Averaging over fast variables (joint with M. Luczak)

Recall join-the-shorter-queue with memory:

- $N$ queues, each serves at rate 1
- customers arrive at rate $N\lambda$ for some $\lambda < 1$
- choose a queue at random and compare with memory queue
- join the shorter queue and update the memory

$$Z_t^k = \text{proportion of queues of length at least } k$$
$$Y_t = \text{length of memory queue}.$$

Use fluid coordinate map $x(z, y) = z$. The drift of $Z^k$ is

$$\beta_k(z, y) = \lambda z_{k-1}1_{\{y \geq k-1\}} - \lambda z_k 1_{\{y \geq k\}} - (z_k - z_{k+1}).$$
Averaging over fast variables

In general, for a Markov chain \((\xi_t)_{t \geq 0}\) in \(E\), we may distinguish between fluid and fast coordinates

\[
x : E \to V, \quad y : E \to I
\]

and consider the *drift* and the *local transition rates*

\[
\beta(\xi) = \int_E \{x(\eta) - x(\xi)\} q(\xi, d\eta), \\
\gamma(\xi, y') = q(\xi, \{\eta \in E : y(\eta) = y'\}).
\]

Let us suppose that

\[
\beta(\xi) = b(x(\xi), y(\xi)), \quad \gamma(\xi, y') = g_{x(\xi)}(y(\xi), y')
\]

where \(G_x = (g_x(y, y'))_{y, y' \in I}\) is the generator of a Markov chain.
Averaging over fast variables

We may guess that the fluid coordinates behave approximately as

\[ \dot{x}_t = \bar{b}(x_t) \]

where \( \bar{b} \) is the effective drift

\[ b(x) = \sum_y b(x, y) \pi_x(y) \]

with \( \pi_x \) the invariant distribution of \( G_x \).

- How to build this into quantitative estimates?
- When does it work?
Averaging over fast variables

Fix a reference state \( \bar{y} \in I \) and consider the function

\[
\chi(x, y) = \mathbb{E} \int_0^T \{ b(x, y_t) - b(x, \bar{y}_t) \} dt
\]

where

- \( T = \inf \{ t \geq 0 : y_t = \bar{y}_t \} \)
- \((y_t)_{t \geq 0}\) and \((\bar{y}_t)_{t \geq 0}\) have generator \( G_x \) with \( y_0 = y, \bar{y}_0 = \bar{y} \).

Assume we can couple \((y_t)_{t \geq 0}\) and \((\bar{y}_t)_{t \geq 0}\) so that

\[
\sup_{x \in V, y \in I} \mathbb{E}_{(x,y)}(T) \leq \tau.
\]

Then \(|\chi(x, y)| \leq \tau \| b \|_{\infty} \) and

\[
G \chi(x, y) = \sum_{y' \in I} g_x(y, y') \chi(x, y') = b(x, y) - \bar{b}(x).
\]

The notion that \( Y_t = y(\xi_t) \) converges fast to equilibrium is quantified in treating \( \tau \) as small.
Averaging over fast variables

We make a small correction to the fluid variable

\[
\bar{x}(\xi) = x(\xi) - \chi(x(\xi), y(\xi)), \quad \bar{X}_t = \bar{x}(\xi_t).
\]

Then

\[
\bar{X}_t = \bar{X}_0 + M_t + \int_0^t \bar{b}(\bar{X}_s)ds + \Delta_t
\]

where

\[
\Delta_t = \int_0^t \int_E \left\{ \chi(x(\eta), y(\eta)) - \chi(x(\xi_s), y(\eta)) \right\} q(\xi_s, d\eta)ds.
\]

We can make hypotheses so that \(M\) is small (as above) and also \(\Delta\). Then the Gronwall argument gives an estimate on the deviation from \(x_t\) of \(\bar{X}_t\) and hence of \(X_t\).