Fluid Approximation to Controlled Markov Chains with Local Transitions

A.Piunovskiy

University of Liverpool
piunov@liv.ac.uk

MDP under investigation

Theorem on the standard fluid approximation

Interpretation of the fluid scaling

Example

Refined fluid model

Example

Summary

MDP under investigation

The Markov Decision Process under consideration is defined by the following elements

$\mathbf{X} = \{0, 1, 2, \dots\}$ is the state space;
$\mathbf{A}$ is the action space (Borel);
$p(z|x, a)$ is the transition probability;
$r(x, a)$ is the one-step loss,
$\varphi : \mathbf{X} \to \mathbf{A}$ is a stationary deterministic policy.

We consider absorbing (at 0) model with the total undiscounted loss, up to the absorption.

Suppose real functions $q^+(y,a) > 0$, $q^-(y,a) > 0$, and $\rho(y,a)$ on $\mathbb{R}^+ \times \mathbf{A}$ are given such that
$q^+(y,a) + q^-(y,a) \leq 1$; $q^+(0,a) = q^-(0,a) = \rho(0,a) \stackrel{\triangle}{=} 0$.

For a fixed $n \in \mathbf{N}$ (scaling parameter), we consider a random walk defined by

$$
{}^n p(z|x,a) = \begin{cases} q^+(x/n,a), & \text{if } z = x+1; \\ q^-(x/n,a), & \text{if } z = x-1; \\ 1 - q^+(x/n,a) - q^-(x/n,a), & \text{if } z = x; \\ 0 & \text{otherwise}; \end{cases}
$$

$$
{}^n r(x,a) = \frac{\rho(x/n,a)}{n}.
$$

Let a piece-wise continuous function $\psi(y): \mathbb{R}^+ \to \mathbf{A}$ be fixed and introduce the continuous-time stochastic process

$$^nY(\tau) = I\{\tau \in \left[\frac{t}{n}, \frac{t+1}{n}\right)\}\ ^nX_t/n, \qquad t = 0, 1, 2, \dots, \quad (1)$$

where the discrete-time Markov chain ^nX_t is governed by control policy $\varphi(x) \stackrel{\triangle}{=} \psi(x/n)$. Under rather general conditions, if $\lim_{n\to\infty}\ ^nX_0/n = y_0$ then, for any τ, $\lim_{n\to\infty}\ ^nY(\tau) = y(\tau)$ almost surely, where the deterministic function $y(\tau)$ satisfies

$$y(0) = y_0; \qquad \frac{dy}{d\tau} = q^+(y, \psi(y)) - q^-(y, \psi(y)). \qquad (2)$$

Hence it is not surprising that in the absorbing case (if $q^-(y,a) > q^+(y,a)$) the objective

$$ {}^n v^{\varphi}_{{}^n X_0} = E^{\varphi}_{{}^n X_0}\left[\sum_{t=1}^{\infty} {}^n r(X_{t-1}, A_t)\right] $$

converges to

$$ v^{\psi}(y_0) = \int_0^{\infty} \rho(y(\tau), \psi(y(\tau)) d\tau \qquad (3) $$

as $n \to \infty$.

Theorem

Suppose all functions $q^+(y,\psi(y)),\ q^-(y,\psi(y)),\ \rho(y,\psi(y))$ *are piece-wise continuously differentiable;*

$$q^-(y,\psi(y)) > \underline{q} > 0, \quad \inf_{y>0} \frac{q^-(y,\psi(y))}{q^+(y,\psi(y))} = \tilde{\eta} > 1;$$

$$\sup_{y>0} \frac{|\rho(y,\psi(y))|}{\eta^y} < \infty,$$

where $\eta \in (1,\tilde{\eta})$.
Then, for an arbitrary fixed $\hat{y} \geq 0$

$$\lim_{n\to\infty} \sup_{0\leq x\leq \hat{y}n} |\, {}^n v_x^{\varphi} - v^{\psi}(x/n)| = 0.$$

As a corollary, if one solves a rather simple optimization problem $v^{\psi}(y) \to \inf_{\psi}$ then the control policy $\varphi^*(x) = \psi^*(x/n)$, coming from the optimal (or nearly optimal) feedback policy ψ^*, will be nearly optimal in the underlying MDP, if n is large enough.

As a corollary, if one solves a rather simple optimization problem $v^{\psi}(y) \to \inf_{\psi}$ then the control policy $\varphi^*(x) = \psi^*(x/n)$, coming from the optimal (or nearly optimal) feedback policy ψ^*, will be nearly optimal in the underlying MDP, if n is large enough.

It is possible to give an upper bound on the level of the accuracy of the fluid approximation, in terms of the initial data only. This upper bound on the error approaches zero as $1/n$.

As a corollary, if one solves a rather simple optimization problem $v^{\psi}(y) \to \inf_{\psi}$ then the control policy $\varphi^*(x) = \psi^*(x/n)$, coming from the optimal (or nearly optimal) feedback policy ψ^*, will be nearly optimal in the underlying MDP, if n is large enough.

It is possible to give an upper bound on the level of the accuracy of the fluid approximation, in terms of the initial data only. This upper bound on the error approaches zero as $1/n$.

After discussing the meaning of the fluid scaling, an example will be presented, showing that condition $\sup_{y>0} \frac{|\rho(y,\psi(y))|}{\eta^y} < \infty$ in this Theorem is important.

Interpretation of the fluid scaling

Suppose packets of information, 1 KB each, arrive to a switch at the rate q^+ MB/sec and are served at the rate q^- MB/sec ($q^+ + q^- \leq 1$). We observe the process starting from initial state x packets up to the moment when the queue is empty. Let the holding cost be h per MB per second, so that $\rho(y) = hy$, where y is the amount of information (MB). (For simplicity, we consider the uncontrolled model.)

One can consider batch arrivals and batch services of 1000 packets every second; then $n = 1$ and $r(x) = hx = \rho(x)$.

One can consider batch arrivals and batch services of 1000 packets every second; then $n = 1$ and $r(x) = hx = \rho(x)$.

Perhaps, it would be more accurate to consider particular packets; then probabilities q^+ and q^- will be the same, but the time unit is $\frac{1}{1000}$ sec, so that the arrival and service rates (MB/sec) remain the same. The loss function will obviously change:

$$r(x)\frac{\text{packets}}{1/1000sec} = h\left(\frac{x}{1000}\right)\frac{1}{1000} = \frac{\rho(x/n)}{n},$$

where $n = 1000$.

Now, the total holding cost $V(x)$ up to the absorption at $x = 0$, satisfies equation

$$V(x) =$$

$$\frac{\rho(x/n)}{n} + q^+ V(x+1) + q^- V(x-1) + (1 - q^+ - q^-) V(x),$$

$$x = 1, 2, \ldots$$

Now, the total holding cost $V(x)$ up to the absorption at $x = 0$, satisfies equation

$$V(x) =$$

$$\frac{\rho(x/n)}{n} + q^+ V(x+1) + q^- V(x-1) + (1 - q^+ - q^-) V(x),$$

$$x = 1, 2, \ldots$$

Since we measure information in MB, it is reasonable to introduce function $\tilde{v}(y)$, such that $V(x) = \tilde{v}(x/n)$: $\tilde{v}(y)$ is the total holding cost given the initial queue was y MB.

Now, after the obvious rearrangements, we obtain

$$\rho\left(\frac{x}{n}\right) + \frac{q^{+}\left[\tilde{v}\left(\frac{x}{n}+\frac{1}{n}\right) - \tilde{v}\left(\frac{x}{n}\right)\right]}{\frac{1}{n}} + \frac{q^{-}\left[\tilde{v}\left(\frac{x}{n}-\frac{1}{n}\right) - \tilde{v}\left(\frac{x}{n}\right)\right]}{\frac{1}{n}}$$

$$= 0.$$

This is a version of the Euler method for solving differential equation

$$\rho(y) + (q^{+} - q^{-})\frac{dv(y)}{dy} = 0,$$

and we expect that $V(x) = \tilde{v}(x/n) \approx v(x/n)$.

Example

This example shows that condition $\sup_{y>0} \frac{|\rho(y,\psi(y))|}{\eta^y} < \infty$ in the above theorem is important.

Let $\mathbf{A} = [1, 2]$, $q^+(y, a) = ad^+$, $q^-(y, a) = ad^-$ for $y > 0$, where $d^- > d^+ > 0$ are fixed numbers such that $2(d^+ + d^-) \leq 1$. Put $\rho(y, a) = a^2\gamma^{y^2}$, where $\gamma > 1$ is a constant.

To solve the fluid model $v^\psi(y) \to \inf_\psi$, we use the dynamic programming approach. One can see that the Bellman function $v^*(y) \stackrel{\triangle}{=} \inf_\psi v^\psi(y)$ has the form

$$v^*(y) = \int_0^y \inf_{a\in\mathbf{A}} \left\{ \frac{\rho(u, a)}{q^-(u, a) - q^+(u, a)} \right\} du$$

and satisfies the Bellman equation

$$\inf_{a \in \mathbf{A}} \left\{ \frac{dv^*(y)}{dy} \left[q^+(y,a) - q^-(y,a) \right] + \rho(y,a) \right\} = 0,$$

$$v^*(0) = 0.$$

Hence function

$$v^*(y) = v^{\psi^*}(y) = \int_0^y \frac{\gamma^{u^2}}{d^- - d^+} du$$

is well defined and $\psi^*(y) \equiv 1$ is the optimal policy.

and satisfies the Bellman equation

$$\inf_{a \in \mathbf{A}} \left\{ \frac{dv^*(y)}{dy} \left[q^+(y,a) - q^-(y,a) \right] + \rho(y,a) \right\} = 0,$$

$$v^*(0) = 0.$$

Hence function

$$v^*(y) = v^{\psi^*}(y) = \int_0^y \frac{\gamma^{u^2}}{d^- - d^+} du$$

is well defined and $\psi^*(y) \equiv 1$ is the optimal policy.

On the opposite, for any control policy π in the underlying MDP, ${}^n v_x^\pi = \infty$ for all $x > 0$, $n \in \mathbf{N}$.

Refined fluid model

Assume that all the conditions of the above theorem are satisfied, except for $q^-(y, \psi(y)) > \underline{q} > 0$. Since the control policies ψ and φ are fixed, we omit them in the formulae below.

Since $q^-(y)$ can approach zero, and $q^+(y) < q^-(y)$, the stochastic process ${}^nX_t$ can spend too much time around the (nearly absorbing) state $x > 0$ for which $q^-(x/n) \approx q^+(x/n) \approx 0$, so that ${}^nv_{{}^nX_0}$ becomes big and can even approach infinity as $n \to \infty$.

The situation becomes good again if, instead of inequalities

$$q^-(y) > \underline{q} > 0 \qquad \sup_{y>0} \frac{|\rho(y)|}{\eta^y} < \infty,$$

we impose condition

$$\sup_{y>0} \frac{|\rho(y)|}{[q^+(y) + q^-(y)]\eta^y} < \infty$$

and introduce the transformed functions

$$\hat{q}^+(y) \triangleq \frac{q^+(y)}{q^+(y) + q^-(y)}; \quad \hat{q}^-(y) \triangleq \frac{q^-(y)}{q^+(y) + q^-(y)};$$

$$\hat{\rho}(y) \triangleq \frac{\rho(y)}{q^+(y) + q^-(y)}.$$

We call the "hat" deterministic model

$$y(0) = y_0; \quad \frac{dy}{du} = \hat{q}^+(y) - \hat{q}^-(y); \quad \tilde{\hat{v}}(y_0) = \int_0^\infty \hat{\rho}(y(u))du,$$

similar to the previous one, *refined* fluid model.

We call the "hat" deterministic model

$$y(0) = y_0; \quad \frac{dy}{du} = \hat{q}^+(y) - \hat{q}^-(y); \quad \tilde{\hat{v}}(y_0) = \int_0^\infty \hat{\rho}(y(u))du,$$

similar to the previous one, *refined* fluid model.

Now

$$\lim_{n\to\infty} \sup_{0\le x\le \hat{y}n} |\,{}^n v_x - \tilde{\hat{v}}(x/n)| = 0,$$

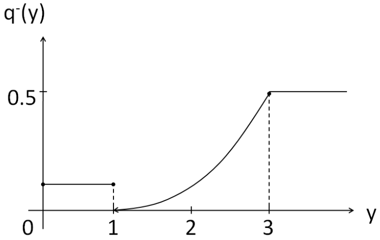
and the convergence is of the rate $1/n$.

Example

As an example, suppose

$$q^-(y) = 0.1\ I\{y \in (0,1]\} + 0.125\ (y-1)^2 I\{y \in (1,3]\} + 0.5\ I\{y > 3\};$$

$$q^+(y) = 0.2\ q^-(y); \qquad \rho(y) = 8\ q^-(y).$$



For the original fluid model, we have $\frac{dy}{d\tau} = -0.1\ (y-1)^2$, and, if initial state $y_0 = 2$ then $y(\tau) = 1 + \frac{10}{\tau+10}$, so that $\lim_{\tau\to\infty} y(\tau) = 1$.

On the opposite, since $q^-, q^+ > 0$ for $y > 0$, and there is a negative trend, process ${}^nY(\tau)$ starting from ${}^nX_0/n = y_0 = 2$ will be absorbed at zero, but the moment of the absorbtion is postponed for later and later as $n \to \infty$ because the process spends more and more time in the neighbourhood of 1.

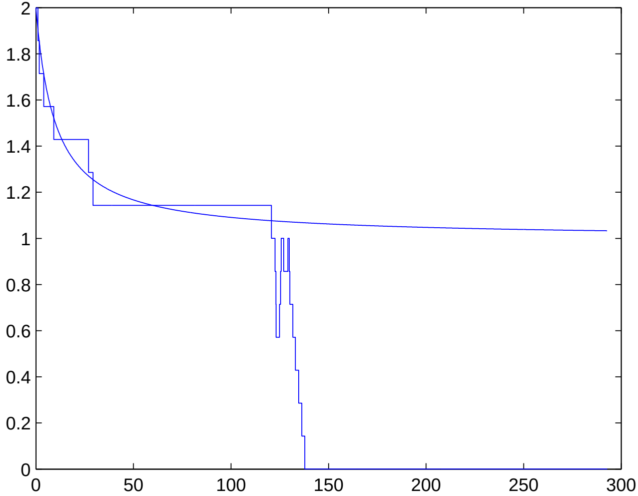


Figure: Stochastic process and its fluid approximation, $n = 7$.

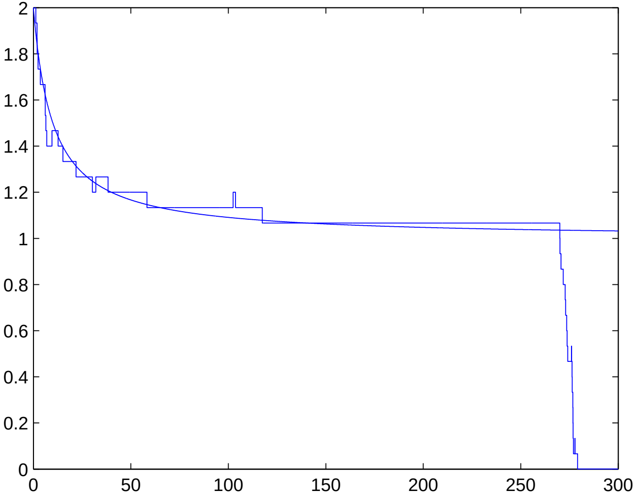


Figure: Stochastic process and its fluid approximation, $n = 15$.

When using the original fluid model, we have

$$\lim_{T\to\infty} \lim_{n\to\infty} E_{2n} \left[\sum_{t=1}^{\infty} I\{t/n \le T\} \, {}^n r(X_{t-1}, A_t) \right]$$

$$= \lim_{T\to\infty} \int_0^T \rho(y(\tau)) d\tau = 10.$$

But we are interested in the following limit:

$$\lim_{n\to\infty} \lim_{T\to\infty} E_{2n} \left[\sum_{t=1}^{\infty} I\{t/n \le T\} \, {}^n r(X_{t-1}, A_t) \right]$$

which in fact equals 20.

When using the refined fluid model, we can calculate $\hat{\rho}(y) = \frac{8}{1.2}$ and $y(u) = 2 - \frac{2}{3}u$, so that the y-process is absorbed at zero at time moment $u = 3$. Therefore,

$$\lim_{n\to\infty} \lim_{T\to\infty} E_{2n}\left[\sum_{t=1}^{\infty} I\{t/n \le T\}\ {}^{n}r(X_{t-1}, A_t)\right]$$

$$= \int_0^{\infty} \hat{\rho}(y(u))du = \int_0^{3} \frac{8}{1.2} du = 20$$

Summary

Summary

- fluid scaling works for processes with local transitions (there are also examples on the multiple-dimentional random walks);

Summary

- fluid scaling works for processes with local transitions (there are also examples on the multiple-dimentional random walks);

- fluid approximation is useful for solving optimal control problems;

Summary

- fluid scaling works for processes with local transitions (there are also examples on the multiple-dimentional random walks);

- fluid approximation is useful for solving optimal control problems;

- one can give an explicit upper bound for the approximation error;

Summary

- fluid scaling works for processes with local transitions (there are also examples on the multiple-dimentional random walks);

- fluid approximation is useful for solving optimal control problems;

- one can give an explicit upper bound for the approximation error;

- the refined fluid model is applicable in some cases when the standard approach fails to work.

Thank you for attention