Modeling controlled Markov chains

Specifying Markov chains

Anderson and Kurtz (2015)

Let X be a continuous time Markov chain where each discontinuity $X(t) - X(t-) \in \{\zeta_1, \ldots, \zeta_m\}$. Then

$$X(t) = X(0) + \sum_{l=1}^{m} R_l(t)\zeta_l$$

where $R_l(t)$ counts the number of ζ_l discontinuities. Then

$$P(X(t+\Delta t)-X(t) = \zeta_l|\mathcal{F}_t) \approx P(R_l(t+\Delta t)-R_l(t) = 1|\mathcal{F}_t) \approx \lambda_l(X(t))\Delta t$$

suggests $R_l(t) = Y_l(\int_0^t \lambda_l(X(s))ds)$, where Y_l is a unit Poisson process. Let $Y_1, \ldots, Y_m$ be independent unit Poisson processes.

$$X(t) = X(0) + \sum_{l=1}^{m} Y_l(\int_0^t \lambda_l(X(s))ds)\zeta_l \qquad (1)$$

$$\tau_i = \inf\{t > 0 : \sum_{l=1}^{m} R_l(t) \geq i\} \qquad \tau_\infty = \lim_{i\to\infty} \tau_i$$

Definition of the generator

Setting

$$\mathbb{A}f(x) = \sum_l \lambda_l(x)(f(x+\zeta_l) - f(x)) \tag{2}$$

For $\widetilde{R}_l(t) = R_l(t) - \int_0^t \lambda_l(X(s))ds$,

$$\begin{aligned} f(X(t)) - f(X(0)) - \int_0^t \mathbb{A}f(X(s))ds \\ = \sum_l \int_0^t (f(X(s-)+\zeta_l) - f(X(s-)))d\widetilde{R}_l(s) \end{aligned}$$

is a martingale. $\mathcal{D}(\mathbb{A}) = \{f : \#\{x : f(x) \neq 0\} < \infty\}$

$\mathbb{A}$ is the *generator* for the Markov chain.

Martingale problem

X is a solution of the martingale problem for $\mathbb{A}$ if

$$M_f(t) = f(X(t)) - f(X(0)) - \int_0^t \mathbb{A}f(X(s))ds$$

is a martingale for each $f \in \mathcal{D}(\mathbb{A})$.

All we need to know for today is that $E[M_f(t)] = 0$.

Forward equation

The martingale property implies

$$E[f(X(t))] = E[f(X(0))] + \int_0^t E[\mathbb{A}f(X(s))]ds$$

and taking $f(x) = \mathbf{1}_{\{y\}}(x)$, we have

$$P\{X(t) = y\} = P\{X(0) = y\} + \int_0^t (\sum_l \lambda_l(y - \zeta_l)P\{X(s) = y - \zeta_l\} - \sum_l \lambda_l(y)P\{X(s) = y\})ds$$

giving the Kolmogorov forward or *master equation* for the distribution of X. In particular, defining $p_y(t) = P\{X(t) = y\}$ and $p_y = P\{X(0) = y\}$, $\{p_y\}$ satisfies the system of differential equations

$$\dot{p}_y(t) = \sum_l \lambda_l(y - \zeta_l)p_{y-\zeta_l}(t) - (\sum_l \lambda_l(y))p_y(t), \quad p_y(0) = \nu_y \tag{3}$$

Equivalence of formulation

The solution of the stochastic equation is a solution of the martingale problem assuming $X(t) = \infty$ for $t \geq \inf\{s : \max_l R_l(s) = \infty\}$ and $f(\infty) = 0$. Each solution of the martingale problem corresponds to a solution of the forward equation.

Each solution for the forward equation corresponds to a solution of the martingale problem. Each solution of the martingale problem corresponds to a solution of the stochastic equation until the first time $\max_l R_l(t) = \infty$.

Specifying controlled Makov chains

$\mathbb{U}$, a nice space giving the collection of possible controls.

Given a control process U with values in $\mathbb{U}$,

$$X(t) = X(0) + \sum_{l=1}^{m} Y_l(\int_0^t \lambda_l(X(s), U(s))ds)\zeta_l.$$

Using the future of the Y_l to select the controls is cheating!

The corresponding martingale problem:

$$M_f(t) = f(X(t)) - f(X(0)) - \int_0^t \mathbb{A}f(X(s), U(s))ds$$

where

$$\mathbb{A}f(x, u) = \sum_l \lambda_l(x, u)(f(x + \zeta_l) - f(x)).$$

The controlled master equation

$E[f(X(t))] = \int_{\mathbb{S}} f(x)\nu_t(dx) = \nu_t f = \sum_{x\in\mathbb{S}} f(x)P\{X(t)=x\}$

$E[h(X(t),U(t))] = \int_{\mathbb{S}\times\mathbb{U}} h(x,u)\mu(dx,du) = \mu_t h.$

The martingale has expectation 0, that is,

$$E[f(X(t))] - E[f(X(0))] - \int_0^t E[\mathbb{A}f(X(s),U(s))]ds = 0$$

so

$$\nu_t f = \nu_0 f + \int_0^t \mu_s \mathbb{A} f dx \tag{4}$$

Theorem 1 *(Almost, but not quite) If $\{\nu_t, \mu_t\}$ satisfy (4), then there is a correspond solution of the controlled martingale problem.*

Relaxed controls

Any measure μ on a reasonable product space $\mathbb{S} \times \mathbb{U}$ can be written $\mu(dx, du) = \eta(x, du)\nu(dx)$. In general, the controlled martingale is

$$M_f(t) = f(X(t)) - f(X(0)) - \int_0^t \int_{\mathbb{U}} \mathbb{A}f(X(s), u)\eta(X(s), du)ds$$

Cost minimization

Cost function $c(x,u)$ depends on both state and control.

Long run average cost

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t E[c(X(s),U(s))]ds = \int_{\mathbb{S}\times\mathbb{U}} c(x,u)\pi(dx,du)$$

Would expect

$$\begin{aligned} 0 &= \lim_{t\to\infty} \frac{1}{t} E[f(X(t)) - f(X(0)) - \int_0^t \mathbb{A}f(X(s),U(s))ds] \\ &= \int_{\mathbb{S}\times\mathbb{U}} \mathbb{A}f(x,u)\pi(dx,du) \end{aligned}$$

Linear programming and relaxed controls

Manne (1960); Stockbridge (1990); Bhatt and Borkar (1996); Kurtz and Stockbridge (1998)

Minimize

$$\int_{\mathbb{S}\times\mathbb{U}} c(x,u)\pi(dx,du) \tag{5}$$

subject to

$$\int_{\mathbb{S}\times\mathbb{U}} \mathbb{A}f(x,u)\pi(dx,du) = 0 \qquad f \in \mathcal{D}(\mathbb{A}). \tag{6}$$

Assuming a minimizing π exists, then setting $\pi(dx,du) = \pi_0(dx)\eta(x,du)$, there exists a stationary solution of the controlled martingale problem, that is,

$$M_f(t) = f(X(t)) - f(X(0)) - \int_0^t \int_{\mathbb{U}} \mathbb{A}f(X(s),u)\eta(X(s),du)ds$$

is a martingale for $f \in \mathcal{D}(\mathbb{A})$.

Existence of minimizing π

Theorem 2 *Assume* $c(x,u) \geq 0$, $\{(x,u) : c(x,u) \leq a\}$ *is compact for every* $a > 0$, *and there exists* $0 < \beta < 1$ *such that for each* $f \in \mathcal{D}(\mathbb{A})$ *there exist* a_f, b_f *such that*

$$|\mathbb{A}f(x,u)| \leq a_f + b_f c(x,u)^\beta. \tag{7}$$

If there exists a π *satisfying (6) such that (5) is finite, then there exists a minimizing* π.

The requirement that $\beta < 1$ in (7) was left out of the statement of the corresponding theorem in Kurtz and Stockbridge (1998) (see Kurtz and Stockbridge (1999)). If $\beta = 1$ is required for (7) to hold, singular controls may arise.

Example

Aoki, Lillacci, Gupta, Baumschlager, Schweingruber, and Khammash (2019); Brait, Gupta, and Khammash (2016)

$$
\begin{aligned}
\emptyset &\overset{\mu}{\rightarrow} Z_1 \\
\emptyset &\overset{\theta X_2}{\rightarrow} Z_2 \\
Z_1 + Z_2 &\overset{\eta}{\rightarrow} \emptyset \\
\emptyset &\overset{\kappa Z_1}{\rightarrow} X_1
\end{aligned}
$$

Why doesn't the big general theorem cover the cases that this crowd cares about

- Delays. The BGT assumes the *current* value of the state is known.
- Partial information. The BGT assumes complete information about the current value of the state.
- Long run average. The cost may be a function of the state at the time the cell splits.

References

David F. Anderson and Thomas G. Kurtz. *Stochastic analysis of biochemical systems*, volume 1 of *Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems*. Springer, Cham; MBI Mathematical Biosciences Institute, Ohio State University, Columbus, OH, 2015. ISBN 978-3-319-16894-4; 978-3-319-16895-1.

Stephanie K. Aoki, Gabriele Lillacci, Ankit Gupta, Armin Baumschlager, David Schweingruber, and Mustafa Khammash. A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature*, 570:533–537, 2019.

Abhay G. Bhatt and Vivek S. Borkar. Occupation measures for controlled Markov processes: characterization and optimality. *Ann. Probab.*, 24(3):1531–1562, 1996. ISSN 0091-1798. doi: 10.1214/aop/1065725192. URL https://doi-org.ezproxy.library.wisc.edu/10.1214/aop/1065725192.

Corentin Brait, Ankit Gupta, and Mustafa Khammash. Antithetic integral feedback ensures robust perfect adaptaion in noisy biomolecular networks. *Cell Systems*, 2:15–26, 2016.

Thomas G. Kurtz and Richard H. Stockbridge. Existence of Markov controls and characterization of optimal Markov controls. *SIAM J. Control Optim.*, 36(2):609–653, 1998. ISSN 0363-0129. doi: 10.1137/S0363012995295516. URL https://doi-org.ezproxy.library.wisc.edu/10.1137/S0363012995295516.

Thomas G. Kurtz and Richard H. Stockbridge. Erratum: "Existence of Markov controls and characterization of optimal Markov controls" [SIAM J. Control Optim. **36** (1998), no. 2, 609–653 (electronic); MR1616514 (99b:93051)]. *SIAM J. Control Optim.*, 37(4):1310–1311, 1999. ISSN 0363-0129. doi: 10.1137/S0363012998343033. URL https://doi-org.ezproxy.library.wisc.edu/10.1137/S0363012998343033.

Alan S. Manne. Linear programming and sequential decisions. *Management Sci.*, 6:259–267, 1960. ISSN 0025-1909. doi: 10.1287/mnsc.6.3.259. URL https://doi-org.ezproxy.library.wisc.edu/10.1287/mnsc.6.3.259.

Richard H. Stockbridge. Time-average control of martingale problems: a linear programming formulation. *Ann. Probab.*, 18(1):206–217, 1990. ISSN 0091-1798. URL http://links.jstor.org.ezproxy.library.wisc.edu/sici?sici=0091-1798(199001)18:1<206:TCOMPA>2.0.CO;2-Z&origin=MSN.

Abstract

Modeling controlled Markov chains

Methods of specifying continuous time Markov chains will be reviewed and then extended to controlled processes. Martingale arguments show that the controlled process can be characterized by a controlled analog of the master equation. Long run average optimal control problems will be introduced indicating how the solution is given by the solution of a linear programming problem, in general infinite dimensional, analogous to the approach of Manne for finite state Markov control problems. Drawing heavily on a recent paper of Aoki, Lillacci, Gupta Baumschlager, Schweingruber, and Khammash, the possibility of implementing the optimal control in a chemical network will be explored.