2022 International Symposium on Nonlinear Theory and Its Applications,
NOLTA2022, Virtual, December 12-15, 2022

Estimation of the Critical Transition Probability Using Quadratic Polynomial Approximation with Skewness Filtering

Makito Oku†

†Institute of Natural Medicine, University of Toyama
2630 Sugitani, Toyama-shi, Toyama 930-0194, Japan
Email: oku@inm.u-toyama.ac.jp

Abstract—Critical transitions are large-scale state transitions that occur occasionally in various complex systems such as ecosystems, climate systems, and financial markets. In order to help predict critical transitions before they occur, several measures for early warning signals have been proposed such as the variance and autocorrelation of a state variable. However, they are related to the linear stability and cannot be used to estimate the critical transition probability. In this study, I propose a nonlinearity-based method for estimating the critical transition probability. It is based on my previous method using quadratic polynomial approximation, and skewness filtering is added as a reject option so that predictions are made only when the observed distribution is sufficiently skewed. The proposed method uses either the least squares method (LSM) or maximum likelihood estimation (MLE). The proposed method is applied to May model, a mathematical model of an ecosystem, as an example case. The results of numerical simulations show that the proposed method has much better precision than the previous method without skewness filtering. In addition, it is found that MLE requires much less data points than LSM if auto-correlation is weak.

1. Introduction

Critical transitions are large-scale state transitions that occur occasionally in various complex systems such as ecosystems, climate systems, and financial markets [1, 2]. In order to help predict critical transitions before they occur, several measures for early warning signals have been proposed such as the variance, autocorrelation, and recovery rate [3–6]. These measures can be calculated even if the equation describing the dynamics of the target system or its parameters are unknown.

However, the abovementioned measures are related to the linear stability of a stable equilibrium or fixed point [7], and nonlinearity is not considered. Because they contain no information on the basin of attraction, they cannot be used to estimate the critical transition probability, i.e., the probability of the occurrence of a critical transition within a given period.

In order to overcome this issue, in a previous study, I proposed a nonlinearity-based approach in which the equation of the target dynamical system was approximated by a quadratic polynomial based on the observed data, and the critical transition probability was estimated from the potential shape and the noise intensity [8]. Unfortunately, the precision of the proposed method was quite low. Even if 10^5 data points were available, the estimated mean escape time $\hat{T}$ was more than twice or less than half of the true value with a probability of 50 % or more.

In this study, I propose an improved method based on the previous method. Skewness filtering is added as a reject option so that predictions are made only when the observed distribution is sufficiently skewed. I also investigate difference between the least squares method and maximum likelihood estimation.

The rest of the paper is organized as follows. In Section 2, theoretical background is explained. In Section 3, the proposed estimation method is explained. In Section 4, simulation settings are explained. In Section 5, the results of numerical simulations are shown. In Section 6, discussion and conclusions are given.

2. Theoretical Background

In this study, we consider continuous-time stochastic dynamical system that can be described as follows:

$$\mathrm{d}x = f(x)\mathrm{d}t + \sigma\,\mathrm{d}W, \tag{1}$$

where $x \in \mathbb{R}$ is a state variable, $f : \mathbb{R} \to \mathbb{R}$ is a time-invariant continuously differentiable nonlinear function, $t \in \mathbb{R}$ is time, $\sigma > 0$ is noise strength, and W is Wiener process. We assume that there exists a potential function U satisfying $f = -U'$, where U' denotes the derivative of U with respect to x. We also assume that U has a local minimum at x^s and a local maximum at x^u, as shown in Fig. 1. They are corresponding to a pair of stable and unstable equilibria that collides at a saddle-node bifurcation for the case of $\sigma \to 0$. We assume $x^u < x^s$ without loss of generality. Initial state x_0 is assumed to be close to x^s.

When x exceeds x^u enough, it is determined that a critical transition has occurred. It has been known that the mean escape time T can be approximated as follows [9]:

$$T = \frac{2\pi}{\sqrt{-f'(x^s)f'(x^u)}} \exp\left(\frac{2}{\sigma^2}(U(x^u) - U(x^s))\right), \tag{2}$$

ORCID iDs Makito Oku: 0000-0002-0282-2361

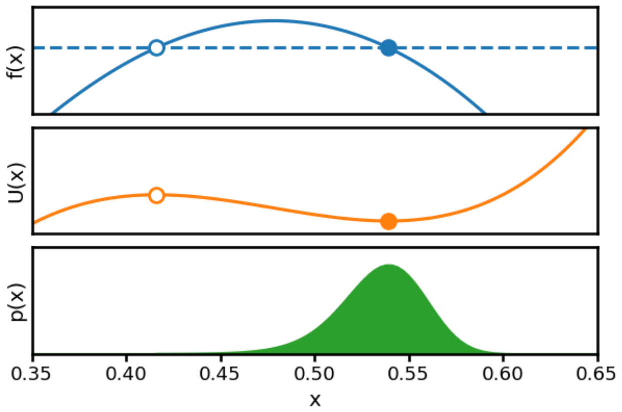

Figure 1: Schematic diagram of $f(x)$, $U(x)$, and $p(x)$. Filled and empty circles indicate x^s and x^u, respectively. Dashed line in the top panel shows the line with $f(x) = 0$.

where f' denotes the derivative of f. The reciprocal of T is called the Kramers' escape rate [10]. If T is sufficiently large, the cumulative transition probability for a given period t can be described as $P(y \leq t) = 1 - \exp(-t/T)$.

The quasi-stationary distribution before a transition can be approximated by the Boltzmann distribution as follows [8]:

$$p(x) = \frac{1}{Z}\exp\left(-\frac{2}{\sigma^2}U(x)\right), \quad x > x^u, \tag{3}$$

where $Z > 0$ is a normalizing constant.

3. Estimation Method

In this section, an estimation method for the critical transition probability is explained. We assume that f is unknown and time series data $\mathcal{D} = \{x_1, \ldots, x_N\}$ is available. The measurement interval is assumed to be constant and denoted as Δt. We use quadratic polynomial approximation of f instead of linear approximation in order to estimate x^u.

Two variants are considered. First one is to directly estimate f by using the least squares method (LSM). The approximated function is described as follows:

$$f(x) \approx \hat{f}(x) = a_0 + a_1 x + a_2 x^2, \tag{4}$$

where $a_0, a_1, a_2 \in \mathbb{R}$ are coefficients. They are estimated by applying LSM to two-dimensional data points $\{(x_n, \Delta x_n)\}$ $(n = 1, \ldots, N-1)$ where $\Delta x_n = (x_{n+1} - x_n)/\Delta t$. U, f', x^s, and x^u are estimated from $\hat{f}$. σ is estimated to be $\hat{\sigma} = \sqrt{\Delta t}\,\mathrm{std}(\Delta x - \hat{f}(x))$, where std denotes standard deviation. T is estimated by using (2).

Another method is to directly estimate $g = (2/\sigma^2)U$ by using the maximum likelihood estimation (MLE). The approximated function is described as follows:

$$\frac{2}{\sigma^2}U(x) \approx \hat{g}(x) = \eta_1 x + \eta_2 x^2 + \eta_3 x^3, \tag{5}$$

where $\eta_1, \eta_2, \eta_3 \in \mathbb{R}$ are coefficients. The likelihood function is $p(x) = \exp(-\hat{g}(x))/Z$, and the coefficients can be obtained by solving the following linear equation [11]:

$$\begin{bmatrix} 1 & 2\overline{x} & 3\overline{x^2} \\ 2\overline{x} & 4\overline{x^2} & 6\overline{x^3} \\ 3\overline{x^2} & 6\overline{x^3} & 9\overline{x^4} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 6\overline{x} \end{bmatrix}. \tag{6}$$

The noise strength σ is estimated in a similar manner to the first method. U, f, f', x^s, and x^u are estimated from $\hat{g}$ and $\hat{\sigma}$. T is estimated by using (2).

When the distribution of the observations $\mathcal{D}$ is not skewed, it is difficult to estimate x^u correctly, which leads to unreliable estimation of the mean escape time T. In order to resolve this issue, skewness filtering is introduced; predictions are made only when the skewness of the observed distribution is below a threshold θ.

4. Simulation Settings

The proposed method was applied to May model [12]. It is a mathematical model of an ecosystem and used in many studies on critical transitions [1, 5, 13–15]. May model has mono- and bi-stable regions, and saddle-node bifurcations occur between them except for the cusp [8], as shown in Fig. 2. May model is described as follows:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x) = r\,x\left(1 - \frac{x}{K}\right) - \frac{c\,x^2}{x^2 + h^2}, \tag{7}$$

where x is the population size, r is the growth rate, K is the maximum population size, c is the consumption rate, and h is the half saturation constant. Because r and K can be set to 1 without loss of generality [8, 12], we assume $r = K = 1$ henceforth.

The stochastic differential equation (1) was solved using Euler-Maruyama method as follows:

$$x_{n+1} = x_n + f(x_n)\Delta t + \sigma\sqrt{\Delta t}\,\xi_n, \tag{8}$$

where ξ_n is a pseudo-random number following the standard normal distribution. Because $x = 0$ corresponds to an unstable equilibrium of the original noise-less May model (7), x_{n+1} was replaced with 0 when it was negative.

The parameter settings were as follows: $h = 0.1$, $c = 0.257$, $\Delta t = 0.1$, $\sigma = 0.01$, and $x_0 = x^s \simeq 0.539$. The closest bifurcation point was $c \simeq 0.260$.

We also consider resampling from the time series data $\mathcal{D} = \{x_1, \ldots, x_N\}$. Every k points $\{x_1, x_{1+k}, x_{1+2k}, \ldots\}$ were extracted.

5. Results

Figure 3 shows cumulative probability of empirical skewness. The probability of empirical skewness being less than -0.5 was more than 20 % for $N = 10^4$ and $N = 10^5$ and approximately 15 % for $N = 10^3$.

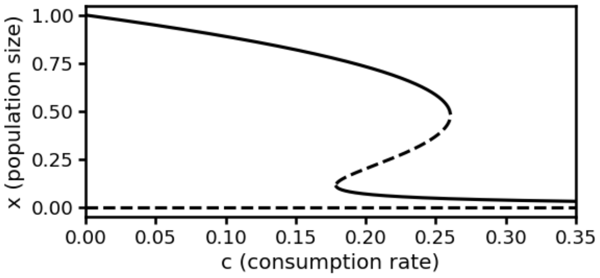

Figure 2: Bifurcation diagram of May model. Solid and dashed lines show stable and unstable equilibria, respectively. $r = K = 1$ and $h = 0.1$.

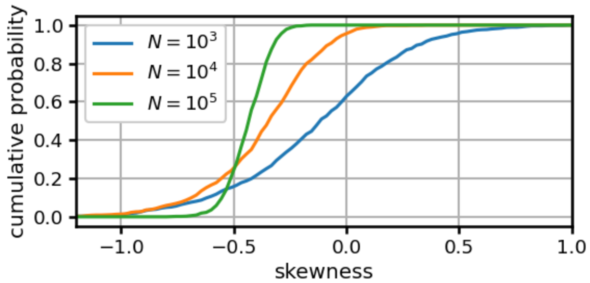

Figure 3: Cumulative probability of empirical skewness. The number of trials was 1 000.

Figure 4 shows relations between the threshold of skewness θ and the absolute error of the mean escape time. The absolute error tended to increase as θ increased. The precision was similar between LSM and MLE for $N = 10^4$ and $N = 10^5$. The relative error $\hat{T}/T$ was approximately ±50 % for both cases when $\theta = -0.5$ and $N = 10^5$.

Figure 5 shows relations between the resampling interval k and the absolute error of the mean escape time. The absolute error increased largely as k increased for the case of LSM. In contrast, the absolute error did not increase much for MLE. Please notice that the scale of the horizontal axis is different between Fig. 5A and Fig. 5B. When $N = 10^5$, the relative error $\hat{T}/T$ for the case of MLE was approximately ±60 % and ±70 % for $k = 10$ and 100, respectively.

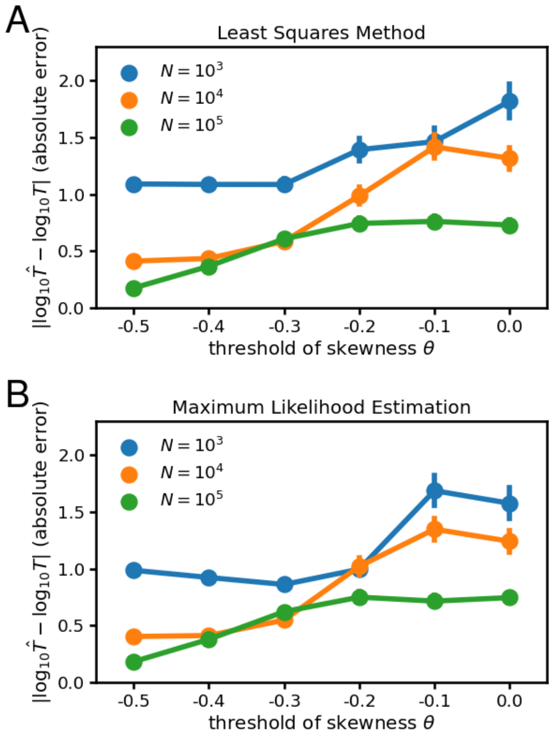

Figure 4: Relations between the threshold of skewness and the absolute error of the mean escape time for the cases of (A) least squares method and (B) maximum likelihood estimation. Error bars show 95 % confidence intervals. The number of trials was 1 000. $k = 1$.

6. Discussion and Conclusions

I discuss how to choose the threshold of skewness θ in practice. When θ is too small, predictions are refrained in most cases as shown in Fig. 3. On the other hand, when θ is too large, the prediction error becomes huge as shown in Fig. 4. The acceptable range of θ depends on the situation, and it should be determined regarding the trade-off. For example, $\theta = -0.5$ would be acceptable if 1 000 uncorrelated data points, which correspond to the case of $N = 10^5$ and $k = 100$, are available, and one requires that the prediction probability is more than 20 % and the relative prediction error $\hat{T}/T$ is within ±70 % on average.

Next, I discuss why MLE was better than LSM with resampling interval $k > 1$ as shown in Fig. 5. Because larger k generally causes weaker auto-correlation of time series data, a possible reason is that the two variants of the proposed method responded differently to auto-correlation. In fact, even if the number of data points were the same, MLE showed different performance with $(N, k) = (10^3, 1)$, $(10^4, 10)$, and $(10^5, 100)$.

Next, I discuss an unresolved issue on the MLE approach. The estimated mean escape time $\hat{T}$ is biased even if independently and identically distributed data points drawn from the quasi-stationary distribution (3) were used. A possible reason is that data points below x^u are missing.

Future works are the correction of the abovementioned bias, application of the proposed method to other cases, dealing with irregular measurement intervals, and so on.

In conclusion, I have proposed a method for estimating critical transition probability using quadratic polynomial approximation with skewness filtering. It was applied to May model. The results of numerical simulations showed that the precision of the proposed method was much better than that of the previous method without skewness filtering [8]. It was also found that MLE required much less data points than LSM if auto-correlation was weak.

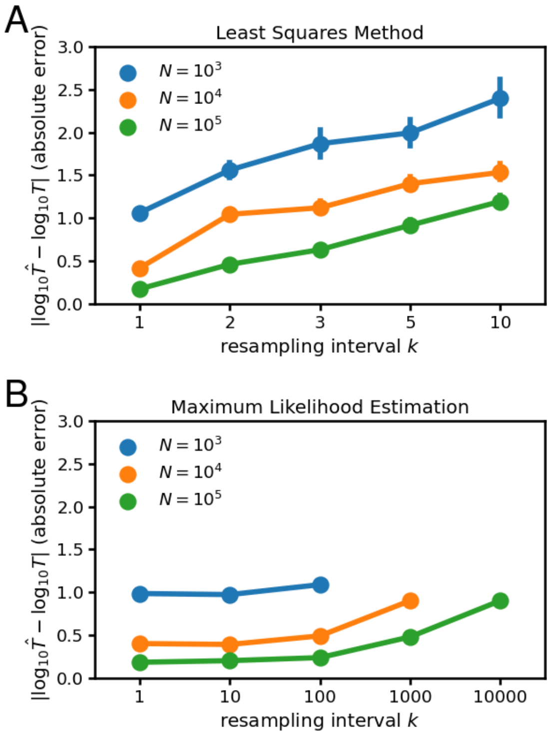

Figure 5: Relations between the resampling interval and the absolute error of the mean escape time for the cases of (A) least squares method and (B) maximum likelihood estimation. Error bars show 95 % confidence intervals. The number of trials was 1 000. $\theta = -0.5$.

Acknowledgments

The author would like to thank NOLTA2022 organizing committee members and the anonymous reviewers. This study was supported by JSPS KAKENHI Grant Number JP15H05707 and JST Moonshot R&D Grant Number JPMJMS2021.

References

[1] Marten Scheffer, Jordi Bascompte, William A. Brock, Victor Brovkin, Stephen R. Carpenter, Vasilis Dakos, Hermann Held, Egbert H. van Nes, Max Rietkerk, and George Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, 2009.

[2] Marten Scheffer, Stephen R. Carpenter, Timothy M. Lenton, Jordi Bascompte, William Brock, Vasilis Dakos, Johan van de Koppel, Ingrid A. van de Leemput, Simon A. Levin, Egbert H. van Nes, Mercedes Pascual, and John Vandermeer. Anticipating critical transitions. *Science*, 338(6105):344–348, 2012.

[3] Stephen R. Carpenter and William A. Brock. Rising variance: A leading indicator of ecological transition. *Ecol. Lett.*, 9(3):311–318, 2006.

[4] Hermann Held and Thomas Kleinen. Detection of climate system bifurcations by degenerate fingerprinting. *Geophys. Res. Lett.*, 31(23):L23207, 2004.

[5] Egbert H. van Nes and Marten Scheffer. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *Am. Nat.*, 169(6):738–747, 2007.

[6] Annelies J. Veraart, Elisabeth J. Faassen, Vasilis Dakos, Egbert H. van Nes, Miquel Lürling, and Marten Scheffer. Recovery rates reflect distance to a tipping point in a living system. *Nature*, 481(7381):357–359, 2012.

[7] Makito Oku and Kazuyuki Aihara. On the covariance matrix of the stationary distribution of a noisy dynamical system. *NOLTA*, 9(2):166–184, 2018.

[8] Makito Oku. Estimation of the critical transition probability using quadratic polynomial approximation. *IEICE Tech. Rep.*, 118(498):57–62, 2019. (in Japanese).

[9] Crispin W. Gardiner. *Handbook of stochastic methods for physics, chemistry and the natural sciences*. Springer-Verlag Berlin Heidelberg, 2nd edition, 1985.

[10] Thomas Kleinen, Hermann Held, and Gerhard Petschel-Held. The potential role of spectral properties in detecting thresholds in the earth system: Application to the thermohaline circulation. *Ocean Dynamics*, 53(2):53–63, 2003.

[11] Aapo Hyvärinen. Estimation of non-normalized statistical models by score matching. *JMLR*, 6(24):695–709, 2005.

[12] Robert M. May. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature*, 269(5628):471–477, 1977.

[13] Vasilis Dakos, Stephen R. Carpenter, William A. Brock, Aaron M. Ellison, Vishwesha Guttal, Anthony R. Ives, Sonia Kéfi, Valerie Livina, David A. Seekell, Egbert H. van Nes, and Marten Scheffer. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLOS ONE*, 7(7):e41010, 2012.

[14] Takuma Nakagawa, Makito Oku, and Kazuyuki Aihara. Early warning signals by dynamical network markers. *SEISAN KENKYU*, 68(3):271–274, 2016. (in Japanese).

[15] Tadayoshi Matsumori, Makito Oku, and Kazuyuki Aihara. Relations between early-warning signals based on dynamical network markers and network structures. In *Proc. NOLTA'17*, pages 229–232, 2018.