Continuous-Time Averaged Models of Discrete-Time Stochastic Systems: Survey and Open Problems

A.L. Fradkov

A survey of a number of papers devoted to continuous-time modeling of discrete-time stochastic systems is given. It is concluded that, although different approaches to averaged (approximate) models justifying are in use, the procedures of building the averaged (approximate) models are similar in different papers. In addition to the deterministic (ODE) model some stochastic continuous-time models described by SDE are introduced. A new result concerning evaluation of the ODE model accuracy over the infinite time interval under partial stability condition is presented. Applications in adaptation, optimization and control are discussed.

1. Introduction

The employment of continuous-time models for analysis and synthesis of discrete-time stochastic systems has started in the 1970s. Hundreds of papers and a number of monographs [7, 20, 43, 44, 51, 11], concerning both application of the machinery and its justification have been published since then. However few authors attempt to compare and unify different lines of research. An additional problem is that quite a number of the results were published in Russian, i.e. they are not well known in the West. An outstanding impact in the area was made by the celebrated paper by L.Ljung [48] that was later listed among 25 seminal papers of the 20th century in control [17]. Currently the paper [48] has got more than 650 citations. What is most impressive its citing rate is about 20 citations per year and it is not going to decrease till now.

In the present paper we survey several avenues of research in continuous-time modeling of discrete-time stochastic systems. Among them the method proposed in [18] and further developed in a few papers and in a monograph [20] has some peculiarities allowing to analyze algorithms with the gain sequence not tending to zero.

2. Continuous-time model building

The method of averaging has a wide applicability in modern control system theory, dynamical systems theory, nonlinear mechanics, etc. [2, ?]. The essence of the method is in separation of slow and fast components of system motion, followed by averaging out the fast motion effects. The formal analysis of the technique for continuous-time systems one can find e.g. in [?, 55] (for deterministic case) and in [55, 63] (for stochastic case).

A specific form of averaging for discrete-time stochastic systems was developed in [18, 20] and, independently in [48] and then applied to various problems in identification and adaptive control. Below the scheme of [18, 20, 48] is described.

Consider a discrete-time stochastic system

\[ x_{k+1} = x_k + \gamma_k F(x_k, f_k), \quad k = 0, 1, 2, \ldots, \]  

where \( x_k \in \mathbb{R}^n \) — state vector, \( f_k \in \mathbb{R}^m \) — random disturbance vector, \( \gamma_k \) — gain parameter. Create the averaged continuous system (continuous model)

\[ \frac{dx}{dt} = A(x), \]  

where \( A(x) = \lim_{k \to \infty} EF(x, f_k) \) (the existence of the limit is assumed). Typical relationships between the discrete-time system and its continuous model are as follows.

1. If the gains \( \gamma_k \) are sufficiently small \( (\gamma_k \leq \gamma) \) then the trajectories \( \{x_k\} \) of (1) are close to the trajectories of (2) \( \{x(t_k)\} \), where \( t_k = \gamma_0 + \cdots + \gamma_{k-1} \).

2. If the gains \( \gamma_k \) tend to zero as \( k \to \infty \) then some asymptotic properties of the solutions of (1) (e.g. stability, ultimate boundedness, etc.) may be similar to those of the solutions of the continuous model (2).

In the case of similarity between (1) and (2) in the above sense one can use simplified model (2) instead of (1) for the purposes of system analysis and design.
Such an approach was called the method of continuous models [18, 20], the ODE approach [48] or the Derevitskii-Fradkov-Ljung (DFL) scheme [31]. Below the term ‘method of continuous models’ will be used since it takes into account two aspects:

— averaging is not the only way of the model generating (in some cases there is a similarity between (1) and (2) even for nonstochastic disturbances $f_k$ investigated in [20]);

— one can use different types of models (e.g., stochastic differential equations).

3. Continuous-time model justifying

A number of rigorous results are known justifying applicability of continuous models for sufficiently small gains $\gamma_k$. Small value of the gains is prerequisite of separation of motions in system. It implies that the disturbance $f_k$ changes faster than the system state $x_k$. The standard condition of averaging is weak dependence of $f_k$ and $f_s$ for large $|k-s|$ (e.g. independence of $f_k$ and $f_s$ when $k \neq s$).

Probably the first results on justifying the averaging for discrete stochastic systems in control theory belong to Meerkov [52], who used discrete averaged model

$$z_{k+1} = z_k + \gamma_k A(z_k)$$

(replacing (3) by (2) creates no extra mathematical problems). The proofs in [52] are based on Krylov–Bogoliubov averaging method [55]. Similarly to the 1st and 2nd Bogoliubov theorems the convergence in probability of solutions of (1) and (3) on finite time interval and, under assumption of asymptotic stability of the model (2), the closeness of the trajectories on infinite interval were established for independent $f_k$.

Significant progress of the method was made by Ljung [47]–[51] who also used Krylov–Bogoliubov approach. In [48] the dependent $f_k$ were treated generated by controlled Markov chain. Moreover, the case $\gamma_k \to 0$ was examined. It was demonstrated that in this case model (2) is responsible for the stability or instability of system (1).

Further development was made by Kul’chitsky [40]–[41] who studied the averaging for some functional of the state vector rather then for the state vector as a whole. It allowed to weaken the restrictive boundedness condition of [48].

if the gain parameter goes to zero at a suitable rate similar in spirit results were obtained [4, 5] without requirements on the dynamics of the model employing a certain set-valued deterministic model.

Another series of results [18]–[22] is based on the machinery developed by S.N. Bernstein who introduced the concept of stochastic differential equation (SDE) as early as in 1934 [9] and established the conditions of the convergence in distribution (weak convergence) of trajectories of (1) either to ODE (2) or to some SDE [10]. In [18] the mean square bounds of the model (2) accuracy were obtained both for finite and for infinite time interval. E.g. it was shown (in [18] for independent $f_k$ and in [2] for $f_k$ satisfying strong mixing conditions) that under Lipschitz and growth conditions

$$\|A(z) - A(z')\| \leq L_1\|z - z'\|, b(z) \leq L_2(1 + \|z\|^2),$$

where $b(z) = E\|F(z, f_k) - A(z)\|^2$ the following inequality holds:

$$E \max_{0 \leq t \leq T} \|x_k - x(t)\|^2 \leq C_1 e^{C_2 T} \gamma,$$

where $\gamma = \max_{1 \leq k \leq N} \gamma_k, T_C \leq T, C_1 > 0, C_2 > 0$.

In the case when the continuous model (2) is exponentially stable it was additionally shown in [18, 20] that the accuracy of approximation over infinite time interval is of order $\gamma^2$ for some $0 < \alpha < 1$. Namely, there exist $\gamma > 0$, such that for $\gamma_k \leq \gamma < \gamma$ the following inequalities hold

$$E \|x_k - x(t_k)\|^2 \leq C_3 \gamma^\alpha, k = 1, 2, \ldots,$$

where numbers $C_3 > 0, \alpha > 0$ do not depend on $\gamma$.

Though the averaging scheme of [18]–[22] is similar to that of Ljung [48], the analytical results are different in that they allow to analyze dynamics of the systems over finite or infinite time intervals rather than convergence as $t \to \infty$. Moreover the results of [18]–[22] are applicable to the cases when the gain $\gamma_k$ does not tend to zero which is important in many applications.

Finally an elegant approach was developed by Kushner [42]–[44] who used weak convergence theory for random functions. This framework however is convenient for the studying of asymptotics when $\gamma \to 0 \ (\gamma_k \equiv \gamma)$ rather than for evaluating mean distance between the trajectories for finite values of $\gamma$.

4. Stochastic continuous model

The inequality (5) shows that the distance between trajectories of (1) and (2) is of order $(\gamma_k)^{1/2}$. This error arises in part due to random fluctuations. Therefore the model taking in account stochasticity potentially may have higher accuracy. Employing the framework of averaging for SDE [29, 63] yields the stochastic continuous model [6, 7, 37]

$$dy = A(x(t))dt + (\gamma(t) B(x(t))^{1/2} dw, \quad (7)$$
where \( \gamma(t) \equiv \gamma_k \) for \( t_k \leq t \leq t_{k+1} \), \( B(x) = \lim_{k \to \infty} \text{M} h(x, f_k) \text{h}(x, f_k)^T \), \( h(x, f_k) = F(x, f_k) - A(x) \), \( w(t) \) — standard Wiener stochastic process, \( x(t) \) — solution of deterministic model (2). In [18] the following stochastic model

\[
dy = A(y(t))dt + (\gamma(t)B(y(t)))^{1/2}dw
\]

was suggested. Model (8) does not use the solutions of deterministic model (2). It was shown that the conditional incremental covariances of solutions of both (7) and (8) coincide with corresponding characteristics of (1) with the accuracy of order \( \gamma_k^5 \). In [7, 20] the family of stochastic models having higher accuracy was introduced. E.g. the accuracy of model

\[
dy = \left[ I_n - \frac{1}{2} \gamma(t) \frac{\partial A(y)}{\partial y} \right] A(y)dt + (\gamma(t)B(y(t)))^{1/2} dw
\]

in terms of the conditional incremental covariances is of order \( \gamma_k^3 \). Note that the model (9) is nothing but the Stratonovich version of the SDE (8).

5. Further results

A number of further results were aimed at extension of the approximation theorems and relaxing their conditions. The approximation bounds were extended to the systems under relaxed Lipschitz and growth conditions [19, 20], to the right hand sides depending on \( \gamma_k \) [19], to the hybrid (discrete-continuous) systems [3]. In the case when the model (2) possesses ultimate boundedness instead of asymptotic stability the ultimate boundedness of initial system (1) was established [19, 20].

New problems such as synchronization and control of networks have become popular during last decade. They demand for new approximation results. One of new demands is to study accuracy of continuous modes over infinite time interval under partial stability assumption for (2) instead of asymptotic stability. For such cases the following theorem can be useful.

**Definition 1** Let \( \Omega, \Omega_0, \Omega \subseteq \Omega_0 \) be closed subsets of \( \mathbb{R}^n \) and \( \Omega \) consists of equilibria of (2). The set \( \Omega \) is called \( \Omega_0 \)–pointwise stable if it is Lyapunov stable and any solution starting from \( \Omega_0 \) tends to a point from \( \Omega \) when \( t \to \infty \).

**Theorem 1** Let Lipschitz and growth conditions (4) hold. Let there exist a smooth mapping \( h : \mathbb{R}^n \to \mathbb{R}^n \) and a bounded set \( \Omega_0 \subseteq \mathbb{R}^n \) such that rank of \( \partial y/\partial z = 1 \) for \( z \in \Omega = \{ z \in \Omega_0 : h(z) = 0 \} \) and the set \( \Omega \) is \( \Omega_0 \)–pointwise stable. Let there exist a twice continuously differentiable function \( V(z) \) and positive numbers \( x_1, x_2, x_3 \) such that

\[
V(z) \leq -x_1 V(z),
\]

\[
|\frac{\partial^2 V(z)}{\partial z_i \partial z_j}| \leq x_3, V(z) \geq x_2 ||h(z)||^2.
\]

Then there exists \( \bar{\gamma} > 0, K_2 > 0, 0 < \alpha < 1 \) such that for \( 0 \leq \gamma_k \leq \bar{\gamma} \) the following inequalities hold

\[
E||y_k - y(t_k)||^2 \leq K_2 \gamma^2, k = 1, 2, ..., \]

where \( y_k = h(z_k), y(t_k) = h(z(t_k)) \).

**Proof.** Let \( \bar{z}(t) \) be the solution of the ODE (2) with the initial condition \( \bar{z}(0) = z_0 \in \Omega_0 \). Pointwise stability implies existence of \( \bar{z} = \lim_{t \to \infty} \bar{z}(t) \in \Omega \). It follows from (10) that \( A(z) = 0 \). Let \( \epsilon > 0 \) be chosen such that rank of \( \partial y/\partial z = 1 \) for \( z \in \Omega_0 \), where \( \Omega_0 = \{ z \in \Omega_0 : \text{dist}(z, \Omega) \leq \epsilon \} \). Then \( E||A(z)|| \leq \epsilon \) (We use notation \( K = \text{const} \), if \( K \) depends only on \( L_1, L_2, \mu, \epsilon, z_0, n, i.e. \) does not depend on \( \gamma_k, t_k \), where \( t_k = \sum_{i=0}^{k-1} \gamma_i \)).

Proof of Lemma 1 is standard and therefore omitted.

**Lemma 1** If the numbers \( \mu_i \geq 0 \) satisfy inequalities \( \mu_k \leq (1 + r_1 \gamma_k)\mu_{k-1} + r_2 \gamma_k, k = 1, 2, ..., \) where \( r_1, r_2 > 0 \), then \( \mu_k \leq (\mu_0 + r_2/r_1)\exp(r_1 t_k) \). If \( r_1 < 0 \) and \( 0 \leq \gamma_k \leq -1/r_1 \), then \( \mu_k \leq -3r_2/r_1 + \mu_0 \exp(r_1 t_k) \).

**Proof of Lemma 1** is standard and therefore omitted.

**Lemma 2** Let the conditions of the theorem hold and the vectors \( d_k \in \mathbb{R}^n \) are defined as follows:

\[
d_{k+1} = d_k + y_k A(d_k), \quad k = 0, 1, 2, ..., \]

Then for \( 0 \leq \gamma_k \leq \gamma < \rho_1 \) the following inequalities hold:

\[
V(z + y_k A(z)) \leq (1 - \rho_2 \gamma) V(z), \]

\[
x_2 ||h(d_k)||^2 \leq V(z_0) \exp(-\rho_2 t_k), \]

\[
x_2 E||z_k||^2 \leq V(z_0) \exp(-\rho_2 t_k) + K_1 \gamma, \]

where \( \rho = x_1 x_2/(\gamma x_3(L_1^2 + L_2)) \), \( \rho_1 = x_1/2 \).

**Proof of Lemma 2** The relations (14) follow from the Taylor expansion and the inequalities (11):

\[
V(z + y_k A(z)) \leq V(z) + y_k \nabla V(z)^T A(z) + \gamma_k^2 x_3 ||A(z)||^2/2 \leq (1 - x_1 \gamma_k) V(z) + y_k \nabla V(z)^T L_2^2 ||z||^2/2 \leq [1 - y_k (x_1 - \gamma n^2 x_3 (L_1^2/(2x_2))] V(z), \]

\[
V(z) \leq -x_1 V(z),
\]

\[
|\frac{\partial^2 V(z)}{\partial z_i \partial z_j}| \leq x_3, V(z) \geq x_2 ||h(z)||^2.
\]
Theorem 2. The theorem provides an upper mean square bound for the distance between the current state and the limit manifold $\Omega = \{z \in \Omega : h(z) = 0\}$. An open problem is relaxation of the pointwise stability condition. Another problem is extending the results to discontinuous models important for economic games, and pattern recognition (some special cases are considered in [33, 46]).

6. Applications of Continuous Models

There are three stages of continuous model using: a) model building; b) model justifying; c) model analyzing (either analytic or numerical). The stage b) including checking the conditions of appropriate theorems sometimes happens to be rather involved. In many cases the theorems serve as "moral support" [48] of the designer's intuition.

Continuous models were used for the analyzing of algorithms of identification [6, 7, 8, 18, 20, 24, 42, 43, 44, 49, 51, 60], optimization [13, 20, 22, 26, 37, 44, 16], filtering [6, 7, 41, ?, 54] and adaptive control [3, 19, 20, 21, 23, 27, 40, 50, 56, 62]; stochastic eigenvalue seeking [58, 65]; games solving [52, 61]; pattern recognition [18, 46, 49]; learning of neural networks [39]. A number of recent works open new networks related application areas: analyzing convergence of learning algorithms for coverage control of mobile sensing agents [14], distributed learning and cooperative control for multi-agent systems [15, 34], distributed topology control of wireless networks [12], etc.

7. Conclusion

Using the continuous models one can simplify the stability and performance analysis of adaptive systems and facilitate discrete-time system design by means of continuous-time design methods. Continuous models provide more detailed information about system behavior than, e.g., Lyapunov function. The main approaches to justifying averaging method for discrete-time stochastic systems are Krylov–Bogoliubov’s approach [48, 52], Bernstein’s approach [18–21] and weak convergence approach [42]. However the procedures of building the averaged (approximate) models are essentially the same. Basic conditions for applicability of averaging are stability of the system and mixing properties of disturbances.

A number of researches are devoted to analysis of the systems with constant or not tending to zero gain (learning rate). Perhaps, the first result of such kind was published in [18]. Unlike many other papers, in [18] approximation bounds were established for nonconstant not tending to zero gain. In this paper it was shown that
the stability restriction of [18] can be relaxed to partial stability. Further relaxation is an avenue of further research.

References


[38] Ioannou P., Kokotovic P.V. Adaptive systems with reduced models. NY: Springer–Verlag, 1983.


