MoViC2020

Parameter Estimation via Fokker–Planck Type Residual: Application to Linear Stationary Random Vibration

Katsutoshi YOSHIDA1*, Yoshikazu YAMANAKA1

¹Department of Mechanical Systems Engineering, Utsunomiya University, 7-1-2 Yoto, Utsunomiya, Tochigi 321-8585, Japan

*yoshidak@cc.utsunomiya-u.ac.jp yyamanaka@cc.utsunomiya-u.ac.jp

Abstract

In this study, we propose a new method that is useful for estimating unknown parameter values of stochastic differential equation (SDE) models, based on probability density function (PDF) data measured from random dynamical systems. As our method does not require explicit description of PDF, it can be applied to the SDE models even when their PDFs are hardly derived in explicit forms due to multiplicative-noise terms, nonlinear terms, and so on. Therefore, our method is expected to provide a versatile tool to dynamically parameterize measured PDF data. In our proposed method, it is assumed that a measured PDF is obtained from a random dynamical system whose structure is described by a known SDE model with unknown parameter values. With the help of Itô calculus, the Fokker–Planck equation (FPE) is derived from the SDE model. The measured PDF and a candidate of parameter values are substituted into the FPE to calculate a FPE residual. Our method is applied to two random vibration systems. Their FPE residuals tend to zero as the parameter values tend to exact values, showing that our proposed FPE residual can be utilized for unknown parameter estimation of SDE models.

Key words : Parameter, Estimation, Fokker-Planck equation, Probability density function

1. Introduction

Efficient and accurate modeling of random dynamical systems is an old and new issue in the fields of physics (Hatjispyros et al., 2007), engineering (Poulimenos and Fassois, 2006; Yoshida et al., 2019), economics (Bhattacharya and Majumdar, 2003), and so on.

In this study, we propose a new method that estimates unknown parameter values of SDE models based on measured PDF data of random dynamical systems. As shown in Fig. 1, our proposed method assumes that a measured PDF is obtained from a random dynamical system whose structure is described by a known SDE model with unknown parameter values. The SDE model is transformed into the corresponding FPE with the help of Itô calculus. The measured PDF and a candidate of parameter values are substituted into the FPE to



calculate a FPE residual. The resulting FPE residual is expected to tend to zero as the parameter values tend to exact values.

Our continuous time domain approach clearly differs from other typical previous methods using discrete time models and/or frequency domain techniques (Hatjispyros et al., 2007; Poulimenos and Fassois, 2006). Moreover, our method has the advantage that it does not require explicit description of PDF; it can be applied to the SDE models whose PDFs are hardly derived in explicit forms due to multiplicative-noise terms, nonlinear terms, and so on.

In this study, our proposed method is tested on two linear random vibration systems. The resulting FPE residuals tend to zero as the parameter values tend to exact values, meaning that our proposed method is capable of estimating unknown parameters of SDE models.

The rest of the paper is organized as follows: Section 2 describes our problem formulation. Section 3 describes

1

our proposed method of parameter estimation. Section 4 demonstrates capability of our proposed method by numerical examples. Section 5 concludes our study.

2. Problem formulation

We consider an *n*-dimensional time-invariant random dynamical system described by a Stratonovich SDE as

$$d\mathbf{x} = \mathbf{f}^{\circ}(\mathbf{x}; \mathbf{q})dt + G(\mathbf{x}; \mathbf{q}) \circ d\mathbf{B}, \quad \mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{R}^n, \tag{1}$$

where *t* is time, *x* is the random state vector, f° is an *n*-dimensional vector-valued function, G(x) is an $n \times m$ matrix-valued function, *B* is an *m*-dimensional standard Brownian motion, and *q* is a *Q*-dimensional parameter vector.

In this study, we impose the following assumptions on our problem:

- The system is stable.
- The system functions f° and G are structurally known.
- The value of the parameter vector \boldsymbol{q} is unknown.
- A stationary PDF data $\bar{p}(\mathbf{x})$ of the state \mathbf{x} is obtained for a certain q.

Our problem is to estimate q that makes the SDE (1) reproduce the measured $\bar{p}(x)$ shape in the sense of least squares.

3. Method of parameter estimation

3.1. Fitting measure based on FPE residual

Although the Stratonovich SDE (1) provides a proper stochastic description of classical mechanics with random fluctuations, it is incompatible with Itô calculus (Gardiner, 2004) that we use in this study. To make it compatible with Itô calculus, we transform (1) to the Itô SDE,

$$d\mathbf{x} = \mathbf{f}(\mathbf{x}; \mathbf{q})dt + G(\mathbf{x}; \mathbf{q})d\mathbf{B}, \quad \mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{R}^n,$$
(2)

which is mathematically equivalent to (1) through

$$f_i(\boldsymbol{x};\boldsymbol{q}) = f_i^{\circ}(\boldsymbol{x};\boldsymbol{q}) + \frac{1}{2} \sum_{k=1}^n \sum_{j=1}^m G_{kj}(\boldsymbol{x};\boldsymbol{q}) \frac{\partial G_{ij}}{\partial x_k}(\boldsymbol{x};\boldsymbol{q}), \quad i = 1, \cdots, n.$$
(3)

Hereafter, $(\cdot)_i$ and $(\cdot)_{ij}$ denote *i*th and (i, j)th components of vector and matrix, respectively.

According to Itô calculus (Gardiner, 2004), the PDF of the solution of (2) satisfies the following FPE:

$$\mathsf{FPE}(p(\boldsymbol{x},t);\boldsymbol{q}) \coloneqq \frac{\partial p(\boldsymbol{x},t)}{\partial t} + \nabla \boldsymbol{J}(p(\boldsymbol{x},t);\boldsymbol{q}) = 0, \quad p(\boldsymbol{x},0) = \delta(\boldsymbol{x}-\boldsymbol{x}_0) \coloneqq \prod_{i=1}^n \delta(x_i - (\boldsymbol{x}_0)_i), \tag{4}$$

where $\delta(\mathbf{x})$ is the *n*-dimensional Dirac's delta function and

$$\nabla \boldsymbol{J}(\boldsymbol{p}(\boldsymbol{x});\boldsymbol{q}) \coloneqq \sum_{i=1}^{n} \frac{\partial J_i(\boldsymbol{p}(\boldsymbol{x});\boldsymbol{q})}{\partial x_i}$$
(5)

is the divergence of the vector field J(p(x); q) of x (called a probability current) whose components are given by

$$J_i(p(\mathbf{x}); \mathbf{q}) := f_i(\mathbf{x}; \mathbf{q}) p(\mathbf{x}) - \frac{1}{2} \sum_{j=1}^n \frac{\partial}{\partial x_j} \left\{ (G(\mathbf{x}; \mathbf{q}) G(\mathbf{x}; \mathbf{q})^T)_{ij} p(\mathbf{x}) \right\}, \quad i = 1, \cdots, n.$$
(6)

Letting $p(\mathbf{x}, t; \mathbf{q})$ be the exact solution of (4) and $\bar{p}(\mathbf{x}, t)$ be an estimated solution, we propose a *PDF-fitness* as

$$E(\bar{p}(\boldsymbol{x},t);\boldsymbol{q}) \coloneqq \int_{\mathcal{D}} \{\mathsf{FPE}(\bar{p}(\boldsymbol{x},t);\boldsymbol{q})\}^2 d\boldsymbol{x},\tag{7}$$

where \mathcal{D} is an appropriate state-space domain and $\mathsf{FPE}(\bar{p}(\boldsymbol{x},t);\boldsymbol{q})$ is a FPE residual. Obviously, $E(\bar{p}(\boldsymbol{x},t);\boldsymbol{q}) = 0$ for $\bar{p}(\boldsymbol{x},t) = p(\boldsymbol{x},t;\boldsymbol{q})$. Furthermore, in this study, we expect the following asymptotic property:

$$E(\bar{p}(\boldsymbol{x},t);\boldsymbol{q}) \to 0 \quad (\bar{p}(\boldsymbol{x},t) \to p(\boldsymbol{x},t;\boldsymbol{q})).$$
(8)

Based on the above, we can solve the optimization problem:

$$\boldsymbol{q}^* = \arg\min_{\boldsymbol{q}} E(\bar{p}(\boldsymbol{x}, t); \boldsymbol{q}), \tag{9}$$

where $\bar{p}(\mathbf{x}, t)$ is substituted by the measured PDF data in the following sections.

In addition, since this study assumes that the system is time-invariant and stable, the FEP has a stationary solution, i.e., $p(\mathbf{x}, t) \rightarrow p(\mathbf{x}) \ (t \rightarrow \infty)$; this yields $\partial p / \partial t = 0$ and simplifies the FPE residual to the stationary version:

$$\mathsf{FPE}(\bar{p}(\boldsymbol{x});\boldsymbol{q}) \coloneqq \nabla \boldsymbol{J}(\bar{p}(\boldsymbol{x});\boldsymbol{q}). \tag{10}$$

In the following sections, we use the stationary (10) to solve the optimization problem (9) and estimate unknown q with the stationary PDF data.

3.2. Numerical implementation

We consider a time-series of the state vector x of length N measured from the random dynamical system (1) as

$$\mathbf{x}(k) := \mathbf{x}(t_k), \quad t_k := t_0 + k\Delta t, \quad k = 0, \cdots, N - 1$$
 (11)

where Δt is a sampling interval. A sufficiently large initial time t_0 is taken to obtain a stationary response. **3.2.1. Quantization of state space** A hyperrectanglar state-space domain \mathcal{D} is taken as

$$\mathcal{D} := \prod_{i=1}^{n} [\underline{x}_i, \overline{x}_i) \subset \mathbb{R}^n, \tag{12}$$

where \prod denotes Cartesian product of sets and $x \in [\underline{x}_i, \overline{x}_i)$ denotes $\underline{x}_i \leq x < \overline{x}_i$. The domain \mathcal{D} is divided into a direct sum of S^n uniform hyperrectangular cells of the form:

$$\mathcal{D}[\boldsymbol{s}] \coloneqq \prod_{i=1}^{n} \left[x_i[s_i] - \frac{\Delta_i}{2}, x_i[s_i] + \frac{\Delta_i}{2} \right), \quad \Delta_i \coloneqq \frac{\overline{x}_i - \underline{x}_i}{S}, \tag{13}$$

where s_i is the *i*th component of the *n*-dimensional index:

$$s \in \mathbb{S} := \{(s_1, \cdots, s_n) \mid 1 \le s_i \le S \text{ for } 1 \le \forall i \le n\}$$

$$(14)$$

and

$$x_i[s_i] := \underline{x}_i + \left(s_i - \frac{1}{2}\right) \Delta_i \tag{15}$$

is the *i*th component of the center point x[s] of the *s*th cell.

In this gridded state-space domain, we designate the value of a function $F(\mathbf{x})$ at a grid-point $\mathbf{x}[\mathbf{s}]$ as

$$F[\mathbf{s}] := F(\mathbf{x}[\mathbf{s}]) \tag{16}$$

and approximate its partial derivative with a numerical derivative in the following form:

$$\frac{\partial F}{\partial x_i}(\boldsymbol{x}[\boldsymbol{s}]) \approx \frac{\Delta F}{\Delta x_i}[\boldsymbol{s}] := \begin{cases} \{F[\boldsymbol{s} + \boldsymbol{e}_i] - F[\boldsymbol{s}]\}(\Delta_i)^{-1} & (s_i = 1), \\ \{F[\boldsymbol{s}] - F[\boldsymbol{s} - \boldsymbol{e}_i]\}(\Delta_i)^{-1} & (s_i = S), \\ \{F[\boldsymbol{s} + \boldsymbol{e}_i] - F[\boldsymbol{s} - \boldsymbol{e}_i]\}(2\Delta_n)^{-1} & (Otherwise), \end{cases}$$
(17)

where e_i denotes the *i*th standard basis vector in \mathbb{S} .

3.2.2. Construction of PDF data We count the frequency N[s] of $x(k) \in \mathcal{D}[s]$ for each *s* and obtain a numerical probability density $\bar{p}[s]$ at each grid-point x[s] as follows.

$$\bar{p}[s] := \frac{N[s]}{N\Delta_1 \cdots \Delta_n}, \quad s \in \mathbb{S},$$
(18)

which we call a PDF data.

3.2.3. Implementation of the PDF-fitness We numerically implement the FPE residual in (10) as

$$\mathsf{FPE}[\bar{p}, s; q] := \sum_{i=1}^{N} \frac{\Delta J_i[\bar{p}, s; q]}{\Delta x_i} \tag{19}$$

with a numerical probability current of the form:

$$J_{i}[\bar{p}, s; q] := f_{i}[s; q]\bar{p}[s] - \frac{1}{2} \sum_{j=1}^{N} \frac{\Delta}{\Delta x_{j}} \left\{ (G[s; q]G[s; q]^{T})_{ij}\bar{p}[s] \right\}.$$
(20)



Fig. 2 PDF data of linear vibration system (23) on the (x_1, x_2) -domain $\mathcal{D} = [-4, 4) \times [-4, 4)$ with the bin number S = 50, which was numerically generated by means of the procedure described in Section 3.2.2.



Fig. 3 PDF-fitness value $E = E[\bar{p}; q]$ with $\nu = \bar{q}_3 = 1$ fixed; the square root value is plotted. On the plane E = 0, the small cross mark indicates the exact parameter value \bar{q} and the curves indicate contours of E.

Therefore, we have a numerical representation of the PDF-fitness in (7) as

$$E[\bar{p};\boldsymbol{q}] := \sum_{\boldsymbol{s}\in S} \{\mathsf{FPE}[\bar{p},\boldsymbol{s};\boldsymbol{q}]\}^2$$
(21)

to solve our parameter estimation problem:

$$\boldsymbol{q}^* = \arg\min_{\boldsymbol{q}} E[\bar{p}; \boldsymbol{q}], \tag{22}$$

based on the measured PDF $\bar{p}[s]$ with the known system structure (1).

4. Application to random linear vibration systems

In this final section, we demonstrate capability of our proposed method described in Section 3.

4.1. Random linear vibration system with additive noise

To provide reliable tests on our method, we start with a linear vibration system whose PDF is explicitly known. Specifically, we consider a linear vibration system with an additive noise only, given by

$$\ddot{x} + c\dot{x} + kx = v w(t), \tag{23}$$

where *c* is a damping, *k* is a stiffness, *v* is a noise strength, and w(t) is a standard Gaussian white noise. Using a state vector $\mathbf{x} := (x, \dot{x})^T$ (*T* denotes transpose) and a parameter vector $\mathbf{q} := (k, c, v)$, we rewrite (23) to the SDE from with

$$\boldsymbol{f}^{\circ}(\boldsymbol{x};\boldsymbol{q}) \coloneqq \begin{bmatrix} 0 & 1\\ -k & -c \end{bmatrix}, \quad \boldsymbol{G}(\boldsymbol{x};\boldsymbol{q}) \coloneqq \begin{bmatrix} 0\\ \nu \end{bmatrix}, \quad \boldsymbol{B} \coloneqq \boldsymbol{B} \text{ (scalar)}.$$
(24)

We obtained a sample of time-series \mathbf{x}_n of length $N = 10^7$ by numerically solving (23) from $x(0) = \dot{x}(0) = 0$ for $\mathbf{q} = \bar{\mathbf{q}} := (1, 0.5, 1)$ with time step $\Delta t = 0.01$. To obtain a stationary data, the same length of initial response was skipped to store. For numerical integration, a fourth-order Runge–Kutta–Gill method was used with the white noise term w(t) simulated by $w(k) \approx W_k / \sqrt{\Delta t}$ ($k = 0, 1, \cdots$) where W_k is normal pseudo-random numbers and $1/\sqrt{\Delta t}$ is the numerical factor required for integrating stochastic differential equations (Higham, 2012). From this stationary \mathbf{x}_n , the PDF data $\bar{p}[s]$ was numerically generated as shown in Fig. 2, by means of the procedure described in Section 3.2.2, on the (x_1, x_2)-domain $\mathcal{D} = [-4, 4) \times [-4, 4)$ with the bin number S = 50.

4.1.1. Estimation of independent parameters Fig. 3 plots the value of our proposed PDF-fitness $E = E[\bar{p}; q]$ on (k, c)-parameter plane with $v = \bar{q}_3 = 1$ fixed; the square root value is plotted for ease of viewing. On the plane E = 0, the small cross mark indicates the exact parameter value \bar{q} and the curves indicate contours of E. Obviously, our proposed E forms a smooth and unimodal concave shape and its minimal point appears to be close to the exact \bar{q} . This implies that our optimization problem in (21) can be solved by gradient methods. In this way, the unknown parameter values were estimated as (k, c) = (0.99980749, 0.49415439) with $E = 4.529 \times 10^{-4}$; their estimation errors were (-0.019%, -1.169%) of the exact (k, c) = (1, 0.5).

Therefore, the results show that our proposed method achieved a reasonable degree of accuracy of the unknown parameter estimation.



Fig. 4 PDF fitness value $E = E[\bar{p}; q]$ plotted on (ν, c) parameter plane with $k = \bar{q}_1 = 1$ fixed. The level set $c = \nu^2/2$ (or $\nu^2/(2c) = 1$) appears as the bottom of
the ravine shape of *E*.



Fig. 5 Numerically extracted level set from the PDF in Fig. 2. The small circles indicate our estimated *c*-value as a function of ν with $k = q_1 = 1$ fixed. The solid curve indicates the exact level set $c = \nu^2/2$.

4.1.2. Extraction of parameter dependency Our method can also be applied to extracting parameter dependency. As the system (23) is linear and additive Gaussian, its stationary PDF becomes Gaussian and can be explicitly derived as (Risken, 1989)

$$p(\mathbf{x}) = \frac{c\,\sqrt{k}}{\pi\nu^2} \exp\left\{-\frac{c}{\nu^2}(kx_1^2 + x_2^2)\right\}.$$
(25)

Such a Gaussian distribution is determined by only the following statistics:

$$E[x_1] = E[x_2] = 0, \quad V[x_1] = \frac{v^2}{2ck}, \quad V[x_2] = \frac{v^2}{2c}, \quad C[x_1, x_2] = 0,$$
 (26)

where E[x] and V[x] denote mean and variance of x, respectively, and C[x, y] denotes covariance between x and y. In these statistics, only the variances depend on the system parameters and they have a common factor $v^2/(2c)$. In other words, the shape of the PDF (25) is parameterized by q = (k, c, v) under the constraint $v^2/(2c) = V[x_2]$. Particularly, in this example, since the PDF data was built for q = (1, 0.5, 1), our PDF-fitness *E* must vanish along the level set $v^2/(2c) = V[x_2] = 1^2/(2 \times 0.5) = 1$.

Fig. 4 shows the corresponding *E*-value on (v, c)-parameter plane with $k = \bar{q}_1 = 1$ fixed. The abovementioned level set $c = v^2/2$ (or $v^2/(2c) = 1$) appears as the bottom of the ravine shape of *E*. In Fig. 5, the small circles indicate our estimated *c*-value for given *k* and *v* values where *k* is fixed to the exact $k = q_1 = 1$ and *v* is taken at 25 uniform grid points within the range $0 \le v \le 1$. The solid curve indicates the exact level set $c = v^2/2$. In these results, the maximal estimation error of *c* was less than -2.81% of the exact value.

In this way, our proposed method also achieved a reasonable degree of accuracy of extracting unknown parameter dependency.

4.2. Random linear vibration system with both additive and multiplicative noises

The second example is of the system whose PDF is not explicitly known. This is simply given by a linear vibration system subjected to both additive and multiplicative noises as follows.

$$\boldsymbol{f}^{\circ}(\boldsymbol{x};\boldsymbol{q}) := \begin{bmatrix} 0 & 1 \\ -k & -c \end{bmatrix}, \quad \boldsymbol{G}(\boldsymbol{x};\boldsymbol{q}) := \begin{bmatrix} 0 & 0 \\ v_1 & v_2 x_1 \end{bmatrix}, \quad \boldsymbol{B} := \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad (27)$$

where B_1 and B_2 are independent standard Brownian motions and v_1 and v_2 are the corresponding noise strengths. This system is linear but its explicit PDF is not yet known (Gitterman, 2005; Nakao, 1998; Zorzano et al., 1999).

Using the same procedure and conditions as those in the first example, we generated the PDF data of (27) for $\boldsymbol{q} = \boldsymbol{\bar{q}} := (k, c, v_1, v_2) = (1, 0.5, 1, 0.6)$. Then, we applied our method to estimating (v_1, v_2) values with $(k, c) = (\boldsymbol{\bar{q}}_1, \boldsymbol{\bar{q}}_2)$ fixed. The resulting estimation was $(v_1, v_2) = (1.00762436, 0.58836647)$ with the PDF-fitness $E = 7.267 \times 10^{-4}$ and thus their estimation errors were (0.762%, -1.939%) of the exact $(v_1, v_2) = (1, 0.6)$.

The results show that our method yielded a reasonable degree of accuracy even when the explicit PDF formulation is not available.

Given the above, it is clearly demonstrated that our proposed method has an advantage of not requiring any explicit PDF expression to estimate parameters. Although this advantage will become distinguished in nonlinear problems, we have restricted ourselves in this study to linear problems to demonstrate primal examples of our method.

5. Conclusion

In this study, we developed a new method that estimates the unknown parameter values of SDE models from PDF data without any explicit PDF formulation. For this purpose, we used FPE residuals and developed a PDF fitness measure that tend to zero as the parameter values tend to exact values. Using the proposed measure, we estimated unknown parameter values of linear random vibration systems and obtained the following results.

As for a random linear vibration system with an additive noise only:

• The independent parameters (i.e, the damping and stiffness) were estimated with estimation errors of under 1.7%.

• The relationship between the dependent parameters (i.e, the damping and noise strength) were extracted with estimation errors of under 2.9%.

As for a random linear vibration system with both additive and multiplicative noises:

• The independent parameters (i.e, the additive and multiplicative noise-strengths) were estimated with estimation errors of under 2.0%.

The above results lead to the conclusion that our method achieved sufficient practical accuracy of the unknown parameter estimation even when the explicit PDF formulation is not available. In future work, we plan to apply our method to nonlinear problems for which the advantage of our method will become distinguished.

Acknowledgment

We would like to thank Prof. Hiroya Nakao for suggesting the topic treated in this paper. This work was funded by JSPS KAKENHI Grant Numbers JP18H01391 and JP17H06552.

References

- Bhattacharya, R. and Majumdar, M., Random dynamical systems: a review, Economic Theory, Vol.23, No.1 (2003), pp.13–38 (2004).
- Gardiner, C. W. Handbook of stochastic methods for physics, chemistry and the natural sciences (2004), Vol.13 of Springer Series in Synergetics. Springer-Verlag, third edition.
- Gitterman, M., Classical harmonic oscillator with multiplicative noise, Physica A: Statistical Mechanics and its Applications, Vol.352, No.2 (2005), pp.309–334.
- Hatjispyros, S., Nicoleris, T., and Walker, S. G., Parameter estimation for random dynamical systems using slice sampling, Physica A: Statistical Mechanics and its Applications, Vol.381 (2007), pp.71–81.
- Higham, D. J., An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations, SIAM Review, Vol.43, No.3 (2012), pp.525–546.
- Nakao, H., Asymptotic power law of moments in a random multiplicative process with weak additive noise, Phys. Rev. E, Vol.58 (1998), pp.1591–1600.
- Poulimenos, A. and Fassois, S., Parametric time-domain methods for non-stationary random vibration modelling and analysis a critical survey and comparison, Mechanical Systems and Signal Processing, Vol.20, No.4 (2006), pp.763–816.
- Risken, H. Fokker-Planck equation: methods of solution and applications (1989). Springer-Verlag Berlin, second edition.
- Yoshida, K., Sato, K., and Yamanaka, Y., Simple degree-of-freedom modeling of the random fluctuation arising in human-bicycle balance, Applied Sciences, Vol.9, No.10 (2019), p.2154.
- Zorzano, M., Mais, H., and Vazquez, L., Numerical solution of two dimensional Fokker–Planck equations, Applied Mathematics and Computation, Vol.98, No.2 (1999), pp.109–117.