Simplified formulas for the mean and variance of linear stochastic differential equations

J.C. Jimenez[a]

[a] *Instituto de Cibernetica, Matematica y Fisica, Calle 15, No. 551, entre C y D, Vedado, La Habana, Cuba.*
e-mail: (see http://sites.google.com/site/locallinearization/)

Abstract

Explicit formulas for the mean and variance of the solutions of linear stochastic differential equations are derived in terms of an exponential matrix. This result improved a previous one by means of which the mean and variance are expressed in terms of a linear combination of higher dimensional exponential matrices. The important role of the new formulas for the system identification as well as numerical algorithms for their practical implementation are pointed out.

Keywords: Stochastic Differential Equations, System Identification, Local Linearization Filter

1. Introduction

Linear Stochastic Differential Equations (SDEs) define one of the more simple class of equations frequently used for modeling a variety of random phenomena. Since long time ago, they have been the focus of intensive researches resulting in a broad and deep knowledge of the properties of their solutions. Among these properties, the mean and variance of the solutions have been well studied. Specifically, the Ordinary Differential Equations (ODEs) that describe the dynamics of the mean and variance are well know (see, e.g., [1]). However, since the explicit solutions of these ODEs were rarely available, numerical solutions were required during some time. Typically, these approximate solutions were computed by means of a numerical integrator for the differential equations or by a numerical quadrature applied to the integral representation of the mean and variance (see, e.g., [17, 15]). Later, in addition to the explicit formulas for the mean and variance of the scalar linear SDEs and for the multidimensional linear SDEs with diagonal drift and diffusion coefficients, explicit formulas could be obtained as well for multidimensional linear SDEs with additive noise. Indeed, by using the main result of [21], the mean and variance of these equations could be expressed in terms of exponential matrices that, nowadays, can be accurate and efficiently computed (see, e.g., [18]). More recently, in [12, 13], explicit formulas for the mean and variance of linear SDEs with multiplicative and/or additive noises were derived in terms of a linear combination of seven exponential matrices. The formulas were obtained as solution of the mentioned ODEs for the mean and variance by using the main result of [21] as well. Apart from being of mathematical interest, these explicit formulas have played a crucial role in the practical implementation of suboptimal linear filters [12], Local Linearization filters [13] and approximate Innovation estimators [14] for the identification of continuous-discrete state space models. In a variety of applications, these methods have shown high effectiveness and efficiency for the estimation of unobserved components and unknown parameters of SDEs given a set of discrete observations. Remarkable is the identification, from actual data, of neurophysiological, financial and molecular models among others (see, e.g., [2, 5, 9, 19, 20]). Therefore, a simplification of the formulas for the mean and variance of linear SDEs might imply a sensible reduction of the computational cost of the mentioned system identification methods and, consequently, a positive impact in applications.

In this paper, simplified explicit formulas for the mean and variance of linear SDEs are obtained in terms of just one exponential matrix of lower dimensionality. The formulas are derived from the solution of the ODEs that describe the evolution of the mean and the second moment of the SDEs. The variance is then obtained from the well-known formula that involves the first two moments. The computational benefits of the simplified formulas are pointed out.

2. Notation and Preliminaries

Let us consider the d-dimensional linear stochastic differential equation

$$d\mathbf{x}(t) = (\mathbf{A}\mathbf{x}(t) + \mathbf{a}(t))dt + \sum_{i=1}^{m} (\mathbf{B}_i\mathbf{x}(t) + \mathbf{b}_i(t))d\mathbf{w}^i(t) \tag{1}$$

for all $t \in [t_0, T]$, where $\mathbf{w} = (\mathbf{w}^1, \dots, \mathbf{w}^m)$ is an m-dimensional standard Wiener process, $\mathbf{A}$ and $\mathbf{B}_i$ are $d \times d$ matrices, and $\mathbf{a}(t) = \mathbf{a}_0 + \mathbf{a}_1 t$ and $\mathbf{b}_i(t) = \mathbf{b}_{i,0} + \mathbf{b}_{i,1} t$ are d-dimensional vectors. Suppose that there exist the first two moments of $\mathbf{x}$ for all $t \in [t_0, T]$.

The ordinary differential equations for the d-dimensional vector mean $\mathbf{m}_t = E(\mathbf{x}(t))$ and the $d \times d$ matrix second moment $\mathbf{P}_t = E(\mathbf{x}(t)\mathbf{x}^\intercal(t))$ of $\mathbf{x}(t)$ are [16]

$$\frac{d\mathbf{m}_t}{dt} = \mathbf{A}\mathbf{m}_t + \mathbf{a}(t)$$

and

$$\frac{d\mathbf{P}_t}{dt} = \mathbf{A}\mathbf{P}_t + \mathbf{P}_t\mathbf{A}^\intercal + \sum_{i=1}^{m} \mathbf{B}_i\mathbf{P}_t\mathbf{B}_i^\intercal + \mathscr{B}(t),$$

where

$$\mathscr{B}(t) = \mathbf{a}(t)\mathbf{m}_t^\intercal + \mathbf{m}_t\mathbf{a}^\intercal(t) + \sum_{i=1}^{m} \mathbf{B}_i\mathbf{m}_t\mathbf{b}_i^\intercal(t) + \mathbf{b}_i\mathbf{m}_t^\intercal\mathbf{B}_i^\intercal(t) + \mathbf{b}_i(t)\mathbf{b}_i^\intercal(t). \tag{2}$$

The solution of these equations can be written as [12, 10]

$$\mathbf{m}_t = \mathbf{m}_0 + \mathbf{L}\mathbf{e}^{\mathbf{C}(t-t_0)}\mathbf{r} \tag{3}$$

and

$$vec(\mathbf{P}_t) = e^{\mathscr{A}(t-t_0)}(vec(\mathbf{P}_0) + \int_0^{t-t_0} e^{-\mathscr{A}s} vec(\mathscr{B}(s+t_0))ds), \tag{4}$$

where $\mathbf{m}_0 = E(\mathbf{x}(t_0))$ and $\mathbf{P}_0 = E(\mathbf{x}(t_0)\mathbf{x}^\intercal(t_0))$ are the first two moments of $\mathbf{x}$ at t_0, and the matrices $\mathbf{C}$, $\mathbf{L}$ and $\mathbf{r}$ are defined as

$$\mathbf{C} = \begin{bmatrix} \mathbf{A} & \mathbf{a}_1 & \mathbf{A}\mathbf{m}_0 + \mathbf{a}(t_0) \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \in \Re^{(d+2)\times(d+2)}, \tag{5}$$

$\mathbf{L} = \begin{bmatrix} \mathbf{I}_d & \mathbf{0}_{d\times 2} \end{bmatrix}$ and $\mathbf{r}^\intercal = \begin{bmatrix} \mathbf{0}_{1\times(d+1)} & 1 \end{bmatrix}$ for non-autonomous equations (i.e., with non zero $\mathbf{a}_1$, $\mathbf{b}_{i,1}$); and as

$$\mathbf{C} = \begin{bmatrix} \mathbf{A} & \mathbf{A}\mathbf{m}_0 + \mathbf{a}(t_0) \\ 0 & 0 \end{bmatrix} \in \Re^{(d+1)\times(d+1)}, \tag{6}$$

$\mathbf{L} = \begin{bmatrix} \mathbf{I}_d & \mathbf{0}_{d\times 1} \end{bmatrix}$ and $\mathbf{r}^\intercal = \begin{bmatrix} \mathbf{0}_{1\times d} & 1 \end{bmatrix}$ for autonomous equations (i.e., with $\mathbf{a}_1 = \mathbf{b}_{i,1} = \mathbf{0}$). Here,

$$\mathscr{A} = \mathbf{A} \oplus \mathbf{A} + \sum_{i=1}^{m} \mathbf{B}_i \otimes \mathbf{B}_i^\intercal \tag{7}$$

is a $d^2 \times d^2$ matrix, and $\mathbf{I}_d$ is the d-dimensional identity matrix. The symbols *vec*, $\oplus$ and $\otimes$ denote the vectorization operator, the Kronecker sum and product, respectively.

The following lemma provides simple expressions to computing multiple integrals involving matrix exponentials such those appearing in (4).

Lemma 1. *([21]) Let $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$ and $\mathbf{A}_4$ be square matrices, n_1, n_2, n_3 and n_4 be positive integers, and set m to be their sum. If the $m \times m$ block triangular matrix $\mathbf{M}$ is defined by*

$$\mathbf{M} = \begin{array}{c} \left[\begin{array}{cccc} \mathbf{A}_1 & \mathbf{B}_1 & \mathbf{C}_1 & \mathbf{D}_1 \\ 0 & \mathbf{A}_2 & \mathbf{B}_2 & \mathbf{C}_2 \\ 0 & 0 & \mathbf{A}_3 & \mathbf{B}_3 \\ 0 & 0 & 0 & \mathbf{A}_4 \end{array} \right] \end{array} \begin{array}{l} \} \, n_1 \\ \} \, n_2 \\ \} \, n_3 \\ \} \, n_4 \end{array},$$

then for $s \geq 0$

$$\left[\begin{array}{cccc} \mathbf{F}_1(s) & \mathbf{G}_1(s) & \mathbf{H}_1(s) & \mathbf{K}_1(s) \\ 0 & \mathbf{F}_2(s) & \mathbf{G}_2(s) & \mathbf{H}_2(s) \\ 0 & 0 & \mathbf{F}_3(s) & \mathbf{G}_3(s) \\ 0 & 0 & 0 & \mathbf{F}_4(s) \end{array} \right] = \exp(s\mathbf{M}),$$

where

$$\mathbf{F}_j(s) \equiv \exp(\mathbf{A}_j s), \textit{ for } j = 1,2,3,4$$

$$\mathbf{G}_j(s) \equiv \int_0^s \exp(\mathbf{A}_j(s-u))\mathbf{B}_j \exp(\mathbf{A}_{j+1}u)du, \textit{ for } j = 1,2,3$$

$$\begin{aligned}\mathbf{H}_j(s) \equiv & \int_0^s \exp(\mathbf{A}_j(s-u))\mathbf{C}_j \exp(\mathbf{A}_{j+2}u)du \\ & + \int_0^s \int_0^u \exp(\mathbf{A}_j(s-u))\mathbf{B}_j \exp(\mathbf{A}_{j+1}(u-r))\mathbf{B}_{j+1} \exp(\mathbf{A}_{j+2}r)drdu\end{aligned}$$

$$\begin{aligned}\mathbf{K}_1(s) \equiv & \int_0^s \exp(\mathbf{A}_1(s-u))\mathbf{D}_1 \exp(\mathbf{A}_4 u)du \\ & + \int_0^s \int_0^u \exp(\mathbf{A}_1(s-u))[\mathbf{C}_1 \exp(\mathbf{A}_3(u-r))\mathbf{B}_3 + \mathbf{B}_1 \exp(\mathbf{A}_2(u-r))\mathbf{C}_2] \exp(\mathbf{A}_4 r)drdu. \\ & + \int_0^s \int_0^u \int_0^r \exp(\mathbf{A}_1(s-u))\mathbf{B}_1 \exp(\mathbf{A}_2(u-r))\mathbf{B}_2 \exp(\mathbf{A}_3(r-w))\mathbf{B}_3 \exp(\mathbf{A}_4 w)dwdrdu.\end{aligned}$$

A generalization of the above lemma for integrals with higher multiplicity is given by the following theorem.

Theorem 2. *([4]) Let $d_1, d_2, \ldots, d_n$, be positive integers. If the $n \times n$ block triangular matrix $\mathbf{A} = [(\mathbf{A}_{lj})]_{l,j=1:n}$ is defined by*

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \ldots & \mathbf{A}_{1n} \\ \mathbf{0} & \mathbf{A}_{22} & \ldots & \mathbf{A}_{2n} \\ \mathbf{0} & \mathbf{0} & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{nn} \end{pmatrix},$$

where $(\mathbf{A}_{lj}), l, j = 1, \ldots, n$ are $d_l \times d_j$ matrices such that $d_l = d_j$ for $l = j$. Then for $t \geqslant 0$

$$e^{\mathbf{A}t} = \begin{pmatrix} \mathbf{B}_{11}(t) & \mathbf{B}_{12}(t) & \ldots & \mathbf{B}_{1n}(t) \\ \mathbf{0} & \mathbf{B}_{22}(t) & \ldots & \mathbf{B}_{2n}(t) \\ \mathbf{0} & \mathbf{0} & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{nn}(t) \end{pmatrix},$$

with

$$\begin{aligned} \mathbf{B}_{ll}(t) &= e^{\mathbf{A}_{ll}t}, l = 1, \ldots, n \\ \mathbf{B}_{lj}(t) &= \int_0^t \mathbf{M}^{(l,j)}(t, s_1) ds_1 \\ &\quad + \sum_{k=1}^{j-l-1} \int_0^t \int_0^{s_1} \ldots \int_0^{s_k} \sum_{l < i_1 < \ldots < i_k < j} \mathbf{M}^{(l,i_1,\ldots,i_k,j)}(t, s_1, \ldots, s_{k+1}) ds_{k+1} \ldots ds_1, \\ l &= 1, \ldots, n-1, \ j = l+1, \ldots, n, \end{aligned}$$

where for any multi-index $(i_1, \ldots, i_k) \in \mathbb{N}^k$ and vector $(s_1, \ldots, s_k) \in \mathbb{R}^k$ the matrices $\mathbf{M}^{(i_1,\ldots,i_k)}(s_1, \ldots, s_k)$ are defined by

$$\mathbf{M}^{(i_1,\ldots,i_k)}(s_1, \ldots, s_k) = \left(\prod_{r=1}^{k-1} e^{\mathbf{A}_{i_r i_r}(s_r - s_{r+1})} \mathbf{A}_{i_r i_{r+1}} \right) e^{\mathbf{A}_{i_k i_k} s_k}.$$

3. Simplified formulas for the first two moments

In this section simplified formulas for the first two moments of the linear SDE (1) and two of their special forms are derived. Equations with multiplicative and additive noises as well autonomous and nonautonomous equations are distinguished.

3.1. Equations with multiplicative and/or additive noises

Lemma 3.

$$vec(\mathscr{B}(s+t_0)) = \mathscr{B}_1 + \mathscr{B}_2 s + \mathscr{B}_3 s^2 + \mathscr{B}_4 \mathbf{e}^{\mathbf{C}s}\mathbf{r} + s\mathscr{B}_5 \mathbf{e}^{\mathbf{C}s}\mathbf{r}, \tag{8}$$

for all $s \geq 0$*, where the vectors* $\mathscr{B}_1 = vec(\beta_1) + \beta_4 \mathbf{m}_0$*,* $\mathscr{B}_2 = vec(\beta_2) + \beta_5 \mathbf{m}_0$ *and* $\mathscr{B}_3 = vec(\beta_3)$*, and the matrices* $\mathscr{B}_4 = \beta_4 \mathbf{L}$ *and* $\mathscr{B}_5 = \beta_5 \mathbf{L}$ *are defined in terms of the matrices*

$$\beta_1 = \sum_{i=1}^{m} (\mathbf{b}_{i,0} + \mathbf{b}_{i,1} t_0)(\mathbf{b}_{i,0} + \mathbf{b}_{i,1} t_0)^{\intercal}$$

$$\beta_2 = \sum_{i=1}^{m} (\mathbf{b}_{i,0} + \mathbf{b}_{i,1} t_0)\mathbf{b}_{i,1}^{\intercal} + \mathbf{b}_{i,1}(\mathbf{b}_{i,0} + \mathbf{b}_{i,1} t_0)^{\intercal}$$

$$\beta_3 = \sum_{i=1}^{m} \mathbf{b}_{i,1}\mathbf{b}_{i,1}^{\intercal}$$

$$\beta_4 = (\mathbf{a}_0 + \mathbf{a}_1 t_0) \oplus (\mathbf{a}_0 + \mathbf{a}_1 t_0) + \sum_{i=1}^{m} (\mathbf{b}_{i,0} + \mathbf{b}_{i,1} t_0) \otimes \mathbf{B}_i + \mathbf{B}_i \otimes (\mathbf{b}_{i,0} + \mathbf{b}_{i,1} t_0)$$

$$\beta_5 = \mathbf{a}_1 \oplus \mathbf{a}_1 + \sum_{i=1}^{m} \mathbf{b}_{i,1} \otimes \mathbf{B}_i + \mathbf{B}_i \otimes \mathbf{b}_{i,1}.$$

Proof. The formula for $vec(\mathscr{B}(s+t_0))$ is directly obtained by substituting (3) in (2) with $t = s + t_0$. □

The main result of this paper is the following.

Theorem 4. *Let* $\mathbf{x}$ *be the solution of the linear SDE (1) with moments* $\mathbf{m}_0 = E(\mathbf{x}(t_0))$ *and* $\mathbf{P}_0 = E(\mathbf{x}(t_0)\mathbf{x}^{\intercal}(t_0))$ *at* t_0*. Then, the first two moments of* $\mathbf{x}$ *can be computed as*

$$\mathbf{m}_t = \mathbf{m}_0 + \mathbf{L}_2 e^{\mathbf{M}(t-t_0)}\mathbf{u}$$

and

$$vec(\mathbf{P}_t) = \mathbf{L}_1 e^{\mathbf{M}(t-t_0)}\mathbf{u}$$

for all $t \in [t_0, T]$*, where the* $(d^2+2d+7)-$*dimensional vector* $\mathbf{u}$ *and the matrices* $\mathbf{M}$*,* $\mathbf{L}_1$*,*

$\mathbf{L}_2$ *are defined as*

$$\mathbf{M} = \begin{bmatrix} \mathscr{A} & \mathscr{B}_5 & \mathscr{B}_4 & \mathscr{B}_3 & \mathscr{B}_2 & \mathscr{B}_1 \\ \mathbf{0} & \mathbf{C} & \mathbf{I}_{d+2} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 2 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & 0 \end{bmatrix}, \ \mathbf{u} = \begin{bmatrix} vec(\mathbf{P}_0) \\ \mathbf{0} \\ \mathbf{r} \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

and

$$\mathbf{L}_2 = \begin{bmatrix} \mathbf{0}_{d\times(d^2+d+2)} & \mathbf{I}_d & \mathbf{0}_{d\times 5} \end{bmatrix}, \mathbf{L}_1 = \begin{bmatrix} \mathbf{I}_{d^2} & \mathbf{0}_{d^2\times(2d+7)} \end{bmatrix},$$

with $\mathscr{B}_i$ *defined as in Lemma 3,* $\mathbf{C}$*,* $\mathbf{r}$ *in (5), and* $\mathscr{A}$ *in (7).*

Proof. From (4) and (8) follows that

$$vec(\mathbf{P}_t) = \mathbf{F}_1 vec(\mathbf{P}_0) + \mathbf{K}_1 + \mathbf{H}_1\mathbf{r}$$

where

$$\mathbf{F}_1 = e^{\mathscr{A}(t-t_0)},$$

$$\mathbf{K}_1 = \int_0^{t-t_0} e^{\mathscr{A}(t-t_0-s)}\mathscr{B}_1 ds + \int_0^{t-t_0}\int_0^{s} e^{\mathscr{A}(t-t_0-s)}\mathscr{B}_2 duds + 2\int_0^{t-t_0}\int_0^{s}\int_0^{u} e^{\mathscr{A}(t-t_0-s)}\mathscr{B}_3 drduds$$

and

$$\mathbf{H}_1 = \int_0^{t-t_0} e^{\mathscr{A}(t-t_0-s)}\mathscr{B}_4 e^{\mathbf{C}s} ds + \int_0^{t-t_0}\int_0^{s} e^{\mathscr{A}(t-t_0-s)}\mathscr{B}_5 e^{\mathbf{C}s} duds.$$

Further, with $\mathbf{F}_3 = \mathbf{e}^{\mathbf{C}(t-t_0)}$, (3) can be written as

$$\mathbf{m}_t = \mathbf{m}_0 + \mathbf{L}\mathbf{F}_3\mathbf{r},$$

where the matrices $\mathbf{L}$, $\mathbf{C}$, $\mathbf{r}$ are defined as in (5). Thus, by a direct application of Theorem

2 (alternatively, Lemma 1 for $\mathbf{F}_1$, $\mathbf{F}_3$ and $\mathbf{H}_1$ can be used by simplicity) follows that

$$\begin{bmatrix} \mathbf{F}_1 & - & \mathbf{H}_1 & - & - & \mathbf{K}_1 \\ \mathbf{0} & - & - & - & - & - \\ \mathbf{0} & \mathbf{0} & \mathbf{F}_3 & - & - & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & - & - & - \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & - & - \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & - \end{bmatrix} = e^{\mathbf{M}(t-t_0)} \text{ with } \mathbf{M} = \begin{bmatrix} \mathscr{A} & \mathscr{B}_5 & \mathscr{B}_4 & \mathscr{B}_3 & \mathscr{B}_2 & \mathscr{B}_1 \\ \mathbf{0} & \mathbf{C} & \mathbf{I}_{d+2} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 2 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & 0 \end{bmatrix},$$

where $\mathscr{B}_i$ are defined as in Lemma 3, and $\mathscr{A}$ in (7). This implies that

$$\mathbf{m}_t = \mathbf{m}_0 + \mathbf{L}\mathbf{v}_2$$

and

$$vec(\mathbf{P}_t) = \mathbf{v}_1,$$

where the d^2-dimensional vector $\mathbf{v}_1$ and the d-dimensional vector $\mathbf{v}_2$ are defined as

$$\begin{bmatrix} \mathbf{v}_1 \\ - \\ \mathbf{v}_2 \\ - \\ - \\ - \end{bmatrix} = e^{\mathbf{M}(t-t_0)}\mathbf{u} \text{ with } \mathbf{u} = \begin{bmatrix} vec(\mathbf{P}_0) \\ \mathbf{0} \\ \mathbf{r} \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Proof concludes by verifying that $\mathbf{L}\mathbf{v}_2 = \mathbf{L}_2\mathbf{v}$ and $\mathbf{v}_1 = \mathbf{L}_1\mathbf{v}$, where $\mathbf{v}^\intercal = [\mathbf{v}_1, -, \mathbf{v}_2, - - -]$, $\mathbf{L}_2 = \begin{bmatrix} \mathbf{0}_{d\times(d^2+d+2)} & \mathbf{L} & \mathbf{0}_{d\times 3} \end{bmatrix}$ and $\mathbf{L}_1 = \begin{bmatrix} \mathbf{I}_{d^2} & \mathbf{0}_{d^2\times(2d+7)} \end{bmatrix}$. $\square$

For autonomous equations with multiplicative and/or additive noises the formulas of the previous theorem can be simplified as follows.

Theorem 5. *Let $\mathbf{x}$ be the solution of the linear SDE (1) with $\mathbf{a}_1 = \mathbf{b}_{i,1} = \mathbf{0}$. Let $\mathbf{m}_0 = E(\mathbf{x}(t_0))$ and $\mathbf{P}_0 = E(\mathbf{x}(t_0)\mathbf{x}^\intercal(t_0))$ be moments of $\mathbf{x}$ at t_0. Then, the first two moments of $\mathbf{x}$ can be computed as*

$$\mathbf{m}_t = \mathbf{m}_0 + \mathbf{L}_2 e^{\mathbf{M}(t-t_0)}\mathbf{u}$$

and

$$vec(\mathbf{P}_t) = \mathbf{L}_1 e^{\mathbf{M}(t-t_0)}\mathbf{u}$$

for all $t \in [t_0, T]$, where the $(d^2+d+2)-$dimensional vector $\mathbf{u}$ and the matrices $\mathbf{M}_1$, $\mathbf{L}_1$

are defined as

$$\mathbf{M} = \begin{bmatrix} \mathscr{A} & \mathscr{B}_1 & \mathscr{B}_4 \\ \mathbf{0} & 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C} \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} vec(\mathbf{P}_0) \\ 1 \\ \mathbf{r} \end{bmatrix}$$

and

$$\mathbf{L}_2 = \begin{bmatrix} \mathbf{0}_{d\times(d^2+1)} & \mathbf{I}_d & \mathbf{0}_{d\times 1} \end{bmatrix}, \mathbf{L}_1 = \begin{bmatrix} \mathbf{I}_{d^2} & \mathbf{0}_{d^2\times(d+2)} \end{bmatrix}$$

with $\mathscr{B}_1$ *and* $\mathscr{B}_4$ *defined as in Lemma 3,* $\mathbf{C}$*,* $\mathbf{r}$ *in (6), and* $\mathscr{A}$ *in (7).*

Proof. Since $\mathbf{a}_1 = \mathbf{b}_{i,1} = \mathbf{0}$, $\mathscr{B}_2 = \mathscr{B}_3 = \mathscr{B}_5 = \mathbf{0}$. Thus, from (4) and (8) follows that

$$vec(\mathbf{P}_t) = \mathbf{F}_1 vec(\mathbf{P}_0) + \mathbf{G}_1 + \mathbf{H}_1\mathbf{r},$$

where

$$\mathbf{F}_1 = e^{\mathscr{A}(t-t_0)},$$

$$\mathbf{G}_1 = \int_0^{t-t_0} e^{\mathscr{A}(t-t_0-s)}\mathscr{B}_1 ds$$

and

$$\mathbf{H}_1 = \int_0^{t-t_0} e^{\mathscr{A}(t-t_0-s)}\mathscr{B}_4 e^{\mathbf{C}s} ds.$$

By a direct application of Lemma 1 follows that

$$\mathbf{m}_t = \mathbf{m}_0 + \mathbf{L}\mathbf{F}_3\mathbf{r}$$

and

$$vec(\mathbf{P}_t) = \mathbf{F}_1 vec(\mathbf{P}_0) + \mathbf{G}_1 + \mathbf{H}_1\mathbf{r},$$

where the matrices $\mathbf{F}_1$, $\mathbf{F}_3$, $\mathbf{K}_1$, and $\mathbf{H}_1$ are defined as

$$\begin{bmatrix} \mathbf{F}_1 & \mathbf{G}_1 & \mathbf{H}_1 \\ \mathbf{0} & - & - \\ \mathbf{0} & \mathbf{0} & \mathbf{F}_3 \end{bmatrix} = e^{\mathbf{M}(t-t_0)} \text{ with } \mathbf{M} = \begin{bmatrix} \mathscr{A} & \mathscr{B}_1 & \mathscr{B}_4 \\ \mathbf{0} & 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C} \end{bmatrix},$$

with $\mathscr{B}_1$ and $\mathscr{B}_4$ defined as in Lemma 3, $\mathbf{C}$, $\mathbf{r}$ in (6), and $\mathscr{A}$ in (7). This implies that

$$\mathbf{m}_t = \mathbf{m}_0 + \mathbf{L}\mathbf{v}_2$$

and

$$vec(\mathbf{P}_t) = \mathbf{v}_1,$$

where the d^2-dimensional vector $\mathbf{v}_1$ and the d-dimensional vector $\mathbf{v}_2$ are defined as

$$\begin{bmatrix} \mathbf{v}_1 \\ - \\ \mathbf{v}_2 \end{bmatrix} = e^{\mathbf{M}(t-t_0)}\mathbf{u} \ \text{ with } \mathbf{u} = \begin{bmatrix} vec(\mathbf{P}_0) \\ 1 \\ \mathbf{r} \end{bmatrix}.$$

Proof concludes by verifying that $\mathbf{L}\mathbf{v}_2 = \mathbf{L}_2\mathbf{v}$ and $\mathbf{v}_1 = \mathbf{L}_1\mathbf{v}$, where $\mathbf{v}^\intercal = [\mathbf{v}_1, -, \mathbf{v}_2]$, $\mathbf{L}_2 = \begin{bmatrix} \mathbf{0}_{d\times(d^2+1)} & \mathbf{L} \end{bmatrix}$ and $\mathbf{L}_1 = \begin{bmatrix} \mathbf{I}_{d^2} & \mathbf{0}_{d^2\times(d+2)} \end{bmatrix}$. □

3.2. Equations with additive noise

For autonomous SDEs with additive noise an additional simplification of the explicit formulas for the first two moments can be archived.

Theorem 6. *Let* $\mathbf{x}$ *be the solution of the linear SDE (1) with* $\mathbf{B}_i = \mathbf{0}$ *and* $\mathbf{a}_1 = \mathbf{b}_{i,1} = \mathbf{0}$. *Let* $\mathbf{m}_0 = E(\mathbf{x}(t_0))$ *and* $\mathbf{P}_0 = E(\mathbf{x}(t_0)\mathbf{x}^\intercal(t_0))$ *be moments of* $\mathbf{x}$ *at* t_0. *Then, the first two moments of* $\mathbf{x}$ *can be computed as*

$$\mathbf{m}_t = \mathbf{m}_0 + \mathbf{k}_1$$

and

$$\mathbf{P}_t = \mathbf{F}_1\mathbf{P}_0\mathbf{F}_1^\intercal + \mathbf{H}_1\mathbf{F}_1^\intercal + \mathbf{F}_1\mathbf{H}_1^\intercal$$

for all $t \in [t_0, T]$, *where the matrices* $\mathbf{F}_1$, $\mathbf{H}_1$, *and* $\mathbf{k}_1$ *are defined as*

$$\begin{bmatrix} \mathbf{F}_1 & - & \mathbf{H}_1 & \mathbf{k}_1 \\ \mathbf{0} & - & - & - \\ \mathbf{0} & \mathbf{0} & - & - \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & - \end{bmatrix} = e^{\mathbf{M}(t-t_0)},$$

being

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{a}_0 & \mathbf{a}_0\mathbf{m}_0^\intercal + \frac{1}{2}\sum_{i=1}^{m}\mathbf{b}_{i,0}\mathbf{b}_{i,0}^\intercal & \mathbf{A}\mathbf{m}_0 + \mathbf{a}_0 \\ \mathbf{0} & 0 & (\mathbf{A}\mathbf{m}_0 + \mathbf{a}_0)^\intercal & 0 \\ \mathbf{0} & \mathbf{0} & -\mathbf{A}^\intercal & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 \end{bmatrix}.$$

a $(2d+2)\times(2d+2)$ *matrix.*

Proof. Thanks to $\mathbf{B}_i = \mathbf{0}$, $\mathbf{a}_1 = \mathbf{b}_{i,1} = \mathbf{0}$, the commutation of the matrices $(\mathbf{I}\otimes\mathbf{A})$ and

$(\mathbf{A}\otimes\mathbf{I})$, and by using expressions (4) and (2) it is obtained that

$$\mathbf{P}_t = e^{\mathbf{A}(t-t_0)}\mathbf{P}_0 e^{\mathbf{A}^\intercal(t-t_0)} + \int\limits_0^{t-t_0} e^{\mathbf{A}(t-t_0-s)}\mathscr{B}(s+t_0)e^{\mathbf{A}^\intercal(t-t_0-s)}ds,$$

where now $\mathscr{B}(s+t_0)$ reduces to

$$\mathscr{B}(s+t_0) = \mathbf{a}_0\mathbf{m}_{s+t_0}^\intercal + \mathbf{m}_{s+t_0}\mathbf{a}_0^\intercal + \sum_{i=1}^{m}\mathbf{b}_{i,0}\mathbf{b}_{i,0}^\intercal.$$

Taking into account that $\mathbf{m}_{s+t_0}$ can be rewritten as [12]

$$\mathbf{m}_{s+t_0} = \mathbf{m}_0 + \mathbf{k}_1,$$

with

$$\mathbf{k}_1 = \int\limits_0^{s} e^{\mathbf{A}(s-u)}(\mathbf{A}\mathbf{m}_0 + \mathbf{a}_0)du$$

we have that

$$\mathbf{a}_0\mathbf{m}_{s+t_0}^\intercal = \mathbf{a}_0\mathbf{m}_0^\intercal + \mathbf{a}_0(\mathbf{A}\mathbf{m}_0 + \mathbf{a}_0)^\intercal\int\limits_0^{s} e^{\mathbf{A}^\intercal(s-u)}du,$$

and so $\mathbf{P}_t$ can be rewritten as

$$\mathbf{P}_t = \mathbf{F}_1\mathbf{P}_0\mathbf{F}_1^\intercal + \mathbf{H}_1\mathbf{F}_1^\intercal + \mathbf{F}_1\mathbf{H}_1^\intercal,$$

where

$$\mathbf{F}_1 = e^{\mathbf{A}(t-t_0)}$$

and

$$\begin{aligned}\mathbf{H}_1 = & \int\limits_0^{t-t_0} e^{\mathbf{A}(t-t_0-s)}(\mathbf{a}_0\mathbf{m}_0^\intercal + \frac{1}{2}\sum_{i=1}^{m}\mathbf{b}_{i,0}\mathbf{b}_{i,0}^\intercal)e^{-\mathbf{A}^\intercal s}ds \\ & + \int\limits_0^{t-t_0} e^{\mathbf{A}(t-t_0-s)}\mathbf{a}_0(\mathbf{A}\mathbf{m}_0 + \mathbf{a}_0)^\intercal\int\limits_0^{s} e^{-\mathbf{A}^\intercal u}duds.\end{aligned}$$

Proof concludes by a direct application of Lemma 1 $\square$

Note that Theorem 6 provides an explicit formula for the first two moments of autonomous linear equations with additive noise in terms of just one exponential matrix. This new result complements both, the well known formulas for the mean and variance of these equations that can be straightforward obtained from Theorem 1 in [21] (Lemma 1 here) in terms of two exponential matrices of dimensions $d+1$ and $2d$, and the formulas for the mean and variance obtained in [3] in terms of just one exponential matrix of dimension $2d+2$.

4. Computational issues and numerical simulations

Theorem 4 provides explicit formulas for the first two moments of the linear SDE (1) in terms of an exponential matrix of dimension d^2+2d+7. By using the well-known expression

$$var(\mathbf{z}) = E(\mathbf{z}\mathbf{z}^\intercal) - E(\mathbf{z})E(\mathbf{z}^\intercal)$$

for the variance $var(\mathbf{z})$ of a random variable $\mathbf{z}$, the variance $var(\mathbf{x})$ of $\mathbf{x}$ solution of (1) can be straightforward computed as

$$var(\mathbf{x}(t)) = \mathbf{P}_t - \mathbf{m}_t\mathbf{m}_t^\intercal$$

for all $t \in [t_0, T]$, where $\mathbf{m}_t$ and $\mathbf{P}_t$ are given as in Theorem 4.

By taking into account that the explicit formulas for the mean and variance of (1) obtained in [12, 13] involve the computation of seven exponential matrices of different dimensions up to a maximum of $3d^2+4d+4$, it is obvious the remarkable benefits of the new simplified formulas. From a computational viewpoint, this includes a considerable reduction of the computer storage capacity and the computational time required for their evaluations through the well known Padé method [18, 11] for exponential matrices. But, in addition, the new formulas allow the efficient computation of the mean and variance of high dimensional systems of the linear SDEs by means of the Krylov subspace method [18, 11] for exponential matrices, which is crucial in many practical situations. Other advantage of the exponential form of these formulas is the flow property the exponential operator, which allows an extra reduction of the computational time when the mean and variance of (1) are required on consecutive time instants with multiplicity. In this case, the firsts two conditional moments at the first time instant after the initial one is computed through the exponential matrix of Theorem 4, whereas the others at the remainder times are computed by simple multiplications of the exponential matrix just mentioned.

Theorems 5 and 6 provide explicit formulas for the first two moments of autonomous linear SDEs, which involve an exponential matrix of lower dimensionality: d^2+d+2 for equation with multiplicative noise, and $2d+2$ for equations with additive noise. This yields extra advantages in a number of important applications.

Type of SDE / dimension	2	8
Non Autonomous, with Multiplicative Noise	0.457	0.036
Autonomous, with Multiplicative Noise	0.322	0.021
Autonomous, with Additive Noise	0.072	0.001

Table I: Relative computational time between the new and old formulas for the moments of linear SDEs.

As illustration, the performance of the new and old formulas are compared for three types of linear equations. In particular, the equations

$$d\mathbf{x}(t) = (-\mathbf{H}\mathbf{x}(t) + \mathbf{1}t)dt + \mathbf{H}\mathbf{x}(t)d\mathbf{w}(t),$$
$$d\mathbf{x}(t) = -\mathbf{H}\mathbf{x}(t)dt + \mathbf{H}\mathbf{x}(t)d\mathbf{w}(t),$$

and

$$d\mathbf{x}(t) = -\mathbf{H}\mathbf{x}(t)dt + \mathbf{1}d\mathbf{w}(t)$$

with $t \in [0,1]$ and initial conditions $E(\mathbf{x}(0)) = \mathbf{1}$, $E(\mathbf{x}(0)\mathbf{x}^{\intercal}(0)) = \mathbf{1}\mathbf{1}^{\intercal}$ were considered, where $\mathbf{H}$ denotes the $d \times d$ Hilbert matrix and $\mathbf{1}$ the d-dimensional unit vector. The formulas of the Theorems 4, 5 and 6 were used to compute the moments of the first, second and third equation, respectively, at $t = 1$. With the same purpose, the formulas of Theorem 3 in [13] were used for the three SDEs. For equations with dimensions $d = 2$ and $d = 8$, Table I presents the relative computational time between the new and old formulas, which is calculated as the ratio of the CPU time consumed for these formulas in each equation. Observe as, in all the cases, the new formulas exhibit a substantial reduction of the computational cost. As it was expected, this reduction clearly increases with the dimensionality and the simplicity of the equation.

Finally, it is worth noting that the simplified formulas derived here have allowed a computationally efficient implementation of the approximate filters and estimators recently proposed in [6], [7] and [8] for the identification of diffusion processes from a reduced number of discrete observations distant in time.

5. Conclusions

In this paper, explicit formulas for the mean and variance of linear stochastic differential equations were derived in terms of an exponential matrix. With respect to the formulas proposed in a previous paper the new ones have a number of clear advantages: 1) they involve the computation of just one exponential matrix of lower dimensionality; 2) for high dimensional SDEs, they can be straightforward computed though the Krylov subspace method; 3) for consecutive time instants with multiplicity, their flow property can be used; and 4) they reduces to simpler forms for autonomous SDEs and for equations with additive noise. From numerical viewpoint, this implies a significant reduction of the computer storage capacity and the computational time.

Acknowledgement: The author thanks to Prof. A. Yoshimoto for his invitation to the Institute of Statistical Mathematics, Japan, where the manuscript was completed.

Bibliography

[1] ARNOLD L., Stochastic Differential Equations: Theory and Applications, Wiley-Interscience Publications, New York, 1974. 1

[2] CALDERON C.P., HARRIS N.C., KIANG C.H., COX D.D. (2009). Analyzing single-molecule manipulation experiments, Journal of Molecular Recognition, 22, 356 - 362. 2

[3] CARBONELL F., JIMENEZ J.C., BISCAY R.J. (2006) Weak local linear discretizations for stochastic differential equations: convergence and numerical schemes, J. Comput. Appl. Math., 197 578 - 596. 12

[4] CARBONELL F., JIMENEZ J.C., PEDROSO L. (2008). Computing multiple integrals involving matrix exponentials, J. Comput. Appl. Math., 213, 300-305. 5

[5] CHIARELLA C., HUNG H., TO T.D. (2009). The volatility structure of the fixed income market under the HJM framework: A nonlinear filtering approach. Comput. Stat. Data Anal. 53, 2075-2088. 2

[6] JIMENEZ J.C. (2012) Approximate linear minimum variance filters for continuous-discrete state space models: convergence and practical algorithms. To appear in Int. J. Control. http://arxiv.org/pdf/1207.6023v2.pdf. 13

[7] JIMENEZ J.C. (2012) Approximate discrete-time schemes for the estimation of diffusion processes from complete observations, submitted. http://arxiv.org/pdf/1212.1788v1.pdf. 13

[8] JIMENEZ J.C. (2012) Approximate continuous-discrete filters for the estimation of diffusion processes from partial and noisy observations, submitted. http://arxiv.org/pdf/1212.3721v1.pdf 13

[9] JIMENEZ J.C., BISCAY R., OZAKI T. (2006). Inference methods for discretely observed continuous-time stochastic volatility models: A commented overview, Asia-Pacific Financial Markets, 12, 109-141. 2

[10] JIMENEZ J.C., CARBONELL F. (2005). Rate of convergence of Local Linearization schemes for initial-value problems, Appl. Math. Comput., 171, 1282-1295. 3

[11] JIMENEZ J.C., DE LA CRUZ H., Convergence rate of strong Local Linearization schemes for stochastic differential equations with additive noise, BIT, 52 (2012) 357-382. 12

[12] JIMENEZ J.C., OZAKI T. (2002). Linear estimation of continuous-discrete linear state space models with multiplicative noise, Systems & Control Letters, 47, 91-101. 2, 3, 11, 12

[13] JIMENEZ J.C., OZAKI T. (2003). Local Linearization filters for nonlinear continuous-discrete state space models with multiplicative noise. Int. J. Control, 76, 1159-1170. 2, 12, 13

[14] JIMENEZ J.C., OZAKI T. (2006). An approximate innovation method for the estimation of diffusion processes from discrete data, J. Time Series Analysis, 27, 77-97. 2

[15] JIMENEZ J.C., VALDES P.A., RODRIGUEZ L.M., RIERA J. AND BISCAY R. (1998) Computing the noise covariance matrix of the local linearization scheme for the numerical solution of stochastic differential equations, Appl. Math. Letters, 11, 19-23. 1

[16] KLOEDEN P.E. AND PLATEN E., Numerical Solution of Stochastic Differential Equations, Springer-Verlag, Berlin, Second Edition, 1995. 2

[17] MORA C. (2005) Weak exponential schemes for stochastic differential equations with noise, IMA J. Numer. Anal., 25, 486-506. 1

[18] MOLER C. AND VAN LOAN C. (2003). Nineteen dubious ways to compute the exponential of a matrix, SIAM Review, 45, 3-49. 2, 12

[19] RIERA, J.J., WATANABE J., IWATA K., MIURA N., AUBERT E., OZAKI T. AND KAWASHIMA R. (2004). A state-space model of the hemodynamic approach: nonlinear filtering of BOLD signals. Neuroimage, 21, 547-567. 2

[20] VALDES P.A., JIMENEZ J.C., RIERA J., BISCAY R. AND OZAKI T. (1999). Nonlinear EEG analysis based on a neural mass model. Biol. Cyb., 81, 415-424. 2

[21] C.F. VAN LOAN (1978). Computing integrals involving the matrix exponential, IEEE Trans. Autom. Control, AC-23, 395-404. 1, 2, 4, 12