1 Introduction

We consider two-dimensional correlated Brownian motion with absorbing boundaries:

$$X(t) = x_0 + \mu_x t + \sigma_x W_x(t) \qquad a_x < X(t) < b_x \tag{1}$$
$$Y(t) = y_0 + \mu_y t + \sigma_y W_y(t) \qquad a_y < Y(t) < b_y \tag{2}$$

where W_i are standard Brownian motions with $\mathrm{Cov}(W_1(t), W_2(t)) = \rho t$ for $0 < t' \leq t$. In particular, we find the joint transition density function for $(X(t), Y(t))$ under the boundary conditions:

$$\Pr\left(X(t) \in dx, Y(t) \in dy | a_x < X(t') < b_x, a_y < Y(t') < b_y, 0 < t' \leq t, X(0) = x_0, Y(0) = y_0, \theta\right), \tag{3}$$

with $\theta := (\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$. This density, which we shorten to $q(x,y,t)$ from now on, is the solution to the Fokker-Planck equation [Oksendal, 2013]:

$$\frac{\partial}{\partial t} q(x,y,t') = -\mu_x \frac{\partial}{\partial x} q(x,y,t') - \mu_y \frac{\partial}{\partial y} q(x,y,t') + \frac{1}{2}\sigma_x^2 \frac{\partial^2}{\partial x^2} q(x,y,t') + \rho\sigma_x\sigma_y \frac{\partial^2}{\partial x \partial y} q(x,y,t') + \frac{1}{2}\sigma_y^2 \frac{\partial^2}{\partial y^2} q(x,y,t'), \tag{4}$$
$$q(a_x,y,t') = q(b_x,y,t') = q(x,a_y,t') = q(x,b_y,t') = 0, \tag{5}$$
$$0 < t' \leq t.$$

Differentiating $q(x,y,t)$ with respect to the boundaries produces the transition density of a particle beginning and ending at the points $(X_1(0), X_2(0))$ and $(X_1(t), X_2(t))$, respectively, while attaining the minima a_x/a_y and maxima b_x/b_y in each coordinate direction:

$$\frac{\partial^4}{\partial a_x \partial b_x \partial a_y \partial b_y} q(x,y,t) =$$

$$\Pr\left(X(t) \in dx, Y(t) \in dy \,\middle|\, \min_{t'} X(t') = a_x, \max_{t'} X(t') = b_x, \min_{t'} Y(t') = a_y, \max_{t'} Y(t') = b_y, 0 < t' \leq t, X(0) = x_0, Y(0) = y_0, \theta\right). \tag{6}$$

The transition density (3) with less than 4 boundaries has been used in computing first passage times [Kou et al., 2016, Sacerdote et al., 2016], with application to structural models in credit risk and default correlations [Haworth et al., 2008, Ching et al., 2014]. He et al. [1998] use variants of (6) with respect to some of the boundaries to price financial derivative instruments whose payoff depends on observed maxima/minima.

Closed-form solutions to (4) - (5) are available for some parameter regimes. When $\rho = 0$, the transition density of the process is the solution to a well-understood Sturm-Liouville problem where the eigenfunctions of the differential operator are sine functions. When $a_1 = -\infty$ and $b_1 = \infty$, the method of images can be used to enforce the remaining boundaries. For either $a_1, a_2 = -\infty$ or $b_1, b_2 = \infty$, eigenfunction of the Fokker-Plank equation can be found in radial coordinates. Both of these techniques are used and detailed by He et al. [1998]. However, to the best of our knowledge, there is no closed-form solution to the general problem in (4) - (5). This also limits the available ways to compute (6), with the most straightforward approach being finite difference with respect to the boundary conditions. This, however, requires one to solve at least 16 eigenvalue problems to evaluate the density function for a single observation, motivating the need for an efficient numerical method to solve (4) - (5).

It is still possible to approach the general problem by proposing a biorthogonal expansion in time and space (Risken [1989], sections 6.2), where the eigenfunctions for the differential operator are approximated as sinusoidal series satisfying the boundary conditions. However, a drawback of this out-of-the-box solution is that the system matrix for the corresponding eigenvalue problem is large and dense. An alternative is to use a finite difference scheme to directly solve the evolution problem after suitable transformations. However, both of these methods need a high degree of numerical resolution to produce practically useful approximations of the transition density. We conjecture that these inefficiencies come from either using a *separable* representation for the differential operator (trigonometric series) or introducing numerical diffusion (finite difference).

In this paper, we propose a solution to the general problem (4) - (5) which is obtained by combining a small-time analytic solution with a finite-element method. Our method directly takes into account the correlation parameter present in the differential operator in order to efficiently represent the analytic small-time solution and propagate it forward in time. We apply our computational method to estimate equation parameters with a maximum likelihood approach in settings where the model assumptions of constant $(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$ and Brownian motion driving stochastic evolution are appropriate.

2 Approximate Numerical Solutions

Before considering any solutions to the full Fokker-Planck equation (4) - (5), we simplify the PDE by proposing an exponential decomposition of the solution and using the fact that parameters $(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$ are constant:

$$q(x,y,t) = \exp(\alpha x + \beta y + \gamma t)p(x,y,t).$$

We can find α, β and γ, as well as a suitable scaling transformation, such that $p(\xi, \eta, \tau)$ satisfies the diffusion equation:

$$\begin{aligned}
\frac{\partial}{\partial\tau}p(\xi,\eta,\tau') &= \frac{1}{2}\sigma_\xi^2\frac{\partial^2}{\partial\xi^2}p(\xi,\eta,\tau') + \rho\sigma_\xi\sigma_\eta\frac{\partial^2}{\partial\xi\partial\eta}p(\xi,\eta,\tau') + \frac{1}{2}\sigma_\eta^2\frac{\partial^2}{\partial\eta^2}p(\xi,\eta,\tau'), && (7)\\
&:= \mathcal{L}p(\xi,\eta,\tau), && (8)\\
p(\xi,\eta,\tau) &= 0 && \text{for } (\xi,\eta) \in [0,1]\times[0,1],\\
p(\xi,\eta,0) &= \delta(\xi-\xi_0)\delta(\eta-\eta_0)
\end{aligned}$$

on the unit square. The transformations $\tilde{\theta} \to \theta$ as well as $(\xi, \eta) \to (x, y)$ allow us to go from $p(\xi, \eta, \tau)$ to $p(x, y, t)$ without trouble. Note here that under this transformation ρ remains the same as in the original coordinate frame. We will call equation (7) the *normalized* problem and will consider its solution without loss of generality.

2.1 Trigonometric Expansion

Following Section 6.2 of Risken [1989], we may use the biorthogonal decomposition of the solution as a sum of eigenfunctions and time-dependent coefficients determined by eigenvalues:

$$p(\xi,\eta,\tau) = \phi_\nu(\xi,\eta)e^{-\lambda_\nu\tau}, \tag{9}$$

where the eigenfunctions $\phi_\nu(\xi, \eta)$ satisfy the boundary conditions. Because the differential operator in the normalized problem (7) is self-adjoint [PROVE], the family of eigenfunctions is complete in the Hilbert space L^2 [CITE]. Moreover, the eigenvalues are bounded below by 0, so that the solution behaves as expected (see section 6.3 of Risken [1989]).

Since we require $\phi_\nu(\xi, \eta)$ to be zero on the boundaries, we may represent the eigenfunction as a linear combination of sines

$$\phi_\nu(\xi,\eta) = \sum_{l=0}^{L}\sum_{m=0}^{M} c_{l,m,\nu}\sin(2\pi l\xi)\sin(2\pi m\eta) := \Psi(\xi,\eta)^T c_\nu,$$

where we have truncated the infinite series for some suitably large L and M and defined

$$\begin{aligned}
\psi_{l,m}(\xi,\eta) &= \sin(2\pi l\xi)\sin(2\pi m\eta),\\
\Psi(\xi,\eta) &= (\psi_{0,0}(\xi,\eta),\ldots,\psi_{L,M}(\xi,\eta))^T,\\
c_\nu &= (c_{0,0,\nu},\ldots,c_{L,M,\nu})^T.
\end{aligned}$$

The biorthogonal representation (9) leads to the eigenvalue problem

$$\mathcal{L}\phi_\nu = -\lambda_\nu \phi_\nu, \tag{10}$$

where $\mathcal{L}$ is the differential operator in the normalized Fokker-Planck equation. Applying $\mathcal{L}$ to ϕ_ν produces the linear system

$$\mathcal{L}\phi_\nu = \mathcal{L}(\Psi(\xi,\eta)^T c_\nu) = \mathcal{L}(\Psi(\xi,\eta)^T)c_\nu = (A\Psi(\xi,\eta))^T c_\nu,$$

where A is a constant matrix dependent on $\tilde{\theta}$. In the case where $\rho = 0$, A is diagonal because $\left\{\psi_{l,m}(\xi,\eta)\right\}_{l,m}$ are the eigenfunctions to $\mathcal{L}$. When $\rho \neq 0$, A is no longer diagonal and is in fact dense. This caused by the mixing terms

$$\frac{\partial^2}{\partial\xi\partial\eta}\sin(2\pi l\xi)\sin(2\pi m\eta) = (2\pi l)(2\pi m)\cos(2\pi l\xi)\cos(2\pi m\eta)$$

being the products of cosine functions, which have an inefficient sine series representation [CITE]. Substituting the linear representation of $\mathcal{L}\phi_\nu$ into the eigenvalue problem (10), we arrive to the system

$$\Psi(\xi,\eta)^T A^T c_\nu = -\lambda_\nu \Psi(\xi,\eta)^T c_\nu \quad \Leftrightarrow \quad A^T c_\nu = -\lambda_\nu c_\nu$$

whose solution gives the family of orthonormal eigenfunctions. As mentioned already, the efficiency of this approach is dependent on the cost of solving the eigenvalue problem $A^T c_\nu = -\lambda_\nu c_\nu$.

2.2 Finite Difference

A finite difference method used to solve the problem (7) defines an approximate solution over some grid of points $\{(\xi_l,\eta_m)\}_{l=0,m=0}^{L,M}$ over $[0,1]\times[0,1]$:

$$q(\xi,\eta,\tau) \approx \sum_l \sum_m c_{l,m}(\tau)\delta(\xi-\xi_l)\delta(\eta-\eta_m) = \Delta(\xi,\eta)^T c(t), \tag{11}$$

where $c(\tau) = (c_{0,0}(\tau),\dots,c_{L,M}(\tau)^T$ and $\Delta(\xi,\eta) = (\delta(\xi-\xi_0)\delta(\eta-\eta_0),\dots,\delta(\xi-\xi_L)\delta(\eta-\eta_M))$, where we have once again separated the spatial and temporal components of the problem as in the previous section. This is a suitable choice, because the differential operator $\mathcal{L}$ is linear and constant and there is therefore no need to perform approximation in time, i.e. we take derivative with respect to time directly:

$$\frac{\partial}{\partial\tau}q(\xi,\eta,\tau) \approx \Delta(\xi,\eta)^T \frac{\partial c(\tau)}{\partial\tau}$$

The differential operator $\mathcal{L}$ for the representation in equation (11) is approximated with a finite difference operator $\mathcal{L}_{FD}$ such that approximate derivatives are defined on the grid:

$$\mathcal{L}q(\xi,\eta,\tau) \approx \mathcal{L}_{FD}q(\xi,\eta,\tau) = \sum_l \sum_m f(c_{l,m}(\tau))\delta(\xi-\xi_l)\delta(\eta-\eta_m) := \Delta(\xi,\eta)^T f(c(t))$$

where $f(c_{l,m}(\tau))$ is a function of some neighboring coefficient values at (ξ_l,η_m).

For a central difference scheme on a regular $N\times N$ grid aligned with the boundaries (with step size $h = 1/(N-1)$),

$$\frac{\partial^2}{\partial\xi^2}q(\xi_l,\eta_m,\tau) \approx \frac{c_{l+1,m}(\tau)-2c_{l,m}(\tau)+c_{l-1,m}(\tau)}{h^2} = \frac{1}{h^2}A_{l,m,\xi^2}c(\tau),$$

where A_{l,m,ξ^2} is some all-zero row vector except for three entries of 1 corresponding to the grid points $(\xi_{l-1},\eta_m),(\xi_l,\eta_m),(\xi_{l+1},\eta_m)$. For the mixing term, the approximation is

$$\frac{\partial^2}{\partial\xi\partial\eta}q(\xi_l,\eta_m,\tau) \approx \frac{c_{l+1,m+1}(\tau)-c_{l+1,m-1}(\tau)-c_{l-1,m+1}(\tau)+c_{l-1,m-1}(\tau)}{4h^2} = \frac{1}{4h^2}A_{l,m,\xi\eta}c(\tau).$$

The finite difference approximation of $\mathcal{L}$ can be written as a linear transformation of $c(\tau)$:

$$\mathcal{L}q(\xi,\eta,\tau) \approx \Delta^T(\xi,\eta) \underbrace{\left(\frac{1}{2}\sigma_\xi^2 \frac{1}{h^2} A_{\xi^2} + \rho\sigma_\xi\sigma_\eta A_{\xi\eta} + \frac{1}{2}\sigma_\xi^2 \frac{1}{h^2} A_{\eta^2} \right)}_{A} c(\tau),$$

where we have composed the row vectors for the different derivative terms as matrices $A_{\xi^2}, A_{\xi\eta}$, and A_{η^2}. The system of differential equations for $c(\tau)$ is therefore completely determined by our choice of step size h, as well as the parameter values $(\sigma_\xi, \sigma_\eta, \rho)$:

$$\begin{aligned} \frac{\partial c(\tau)}{\partial \tau} &= Ac(\tau) \\ \Rightarrow c(\tau) &= \exp(A\tau)\, c(0) \end{aligned} \tag{12}$$

For non-small τ, we must find the eigenvalue decomposition of A in order to solve the evolution problem. It should be noted here that a regular grid approach with a constant h is appealing, because it allows us to fill once and store the matrices $A_{\xi^2}, A_{\eta^2}, A_{\xi\eta}$, which saves valuable computational resources if we are to solve the finite difference eigenproblem (12) repeatedly for different parameter values $\tilde{\theta}$.

Unlike the system matrix for the trigonometric expansion, the system matrix here is sparse, as was demonstrated for A_{ξ^2} explicitly, so that the eigenvalue problem is much cheaper. Further, the system matrix can be made even sparser by considering a 45°

2.3 Finite Element Method

The method we use relies on two parts:

1. a small-time analytic solution $q(x,y,t_\varepsilon)$ for the IC/BC problem,

2. a family of orthonormal basis functions which represent $q(x,y,t_\varepsilon)$ parsimoniously.

By combining 1) and 2), we can efficiently find a weak solution to the PDE (7) via the finite element method [Shaidurov, 2013]. Convergence of our method to the strong solution under the $L^2(\bar{\Omega})$ norm is guaranteed as long as the family we propose is complete in the Banach space of functions induced under $L^2(\bar{\Omega})$ [Salsa, 2016].

The small-time solution is derived by considering the fundamental solution $G(x,y|t,x_0,y_0)$ for the unbounded problem in (7), which is the bivariate Gaussian density with mean and covariance determined by the initial condition and the diffusion parameters [Stakgold and Holst, 2011]. We can then find a small enough t_ε such that $G(x,y|t_\varepsilon,x_0,y_0)$ is numerically zero on three of the four boundaries of $\bar{\Omega}$. The zero-condition on the remaining boundary is enforced by suitably reflecting $G(x,y|t_\varepsilon,x_0,y_0)$ about the boundary. The small-time solution therefore takes on the analytic form

$$q(x,y,t_\varepsilon) = G(x,y|t_\varepsilon,x_0,y_0) - G(x,y|t_\varepsilon,x_0',y_0'),$$

for some known $(x_0', y_0', t_\varepsilon)$.

The construction of the orthonormal basis functions is motivated by the Green's function for the unbounded problem: before performing Gram-Schmidt orthogonalization, the finite family of basis functions are of the form

$$\tilde{\psi}_k(x,y|x_k,y_k,\rho,\sigma) = N\left((x,y)^T \middle| (x_k,y_k)^T, \quad \begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{pmatrix} \right) x(1-x)y(1-y).$$

The advantage of these basis elements is that they better resolve the fundamental for the unbounded problem by taking into account ρ in the covariance of each kernel. By performing Gram-Schmidt orthogonalization under the $L^2(\Omega)$ norm, we arrive at a family of orthonormal functions which can better resolve small-time solutions having a large correlation coefficient.

3 Estimation

As an application of our computational method, we estimate the parameters in (6). However, before we do so, we prove a lemma to show that maximum likelihood estimates based on the approximate solution are asymptotically efficient.

Lemma 1. *The maximum likelihood estimator is consistent as $n \to \infty$ and $k \to \infty$:*

$$\hat{\theta}_{n,k} \to \theta$$

.

Proof. By the definition of weak convergence, given the weak solution q_k and the classical solution q, for any continuous function f,

$$\langle q_k | f \rangle \to \langle q | f \rangle \text{ as } k \to \infty.$$

Because f can be any function in L^2, we can choose f to be $\exp(ilx)$ for any integer l. This means that the characteristic function of X_k converges pointwise to the characteristic function of X. By Levy's continuity theorem, this means that

$$X_k \xrightarrow{d} X \text{ as } k \to \infty.$$

Next, given Theorem 4.1 in Singler [2008], we know that, for each k, q_k satisfies the criteria A1 - A6 in Casella and Berger [2002] to guarantee that, for data $X_k \sim F_k(\theta)$,

$$\hat{\theta}_{n,k}(X_k) \xrightarrow{p} \theta$$

as $n \to \infty$. Moreover, we are guaranteed asymptotic efficiency. In other words, the MLE estimator for $(\sigma_x, \sigma_y, \rho)$ based on the likelihood function under F_k for data sampled from F_k is asymptotically efficient. Now we need to show that the same holds for data sampled from F as $k \to \infty$.

To do this, we will use Chebyshev's inequality:

$$\Pr_X \left(\left| \hat{\theta}_{n,k}(X) - \theta \right| \geq \varepsilon \right) \leq \frac{\mathrm{E}_X \left[(\hat{\theta}_{n,k}(X) - \theta)^2 \right]}{\varepsilon^2}.$$

By the Maximum theorem, $\hat{\theta}_{n,k}(x)$ is a continuous function with respect to x, and further because we have bounded $\hat{\theta}$ from below and above,

$$\mathrm{E}_{X_k} \left[(\hat{\theta}_{n,k}(X_k) - \theta)^2 \right] \to \mathrm{E}_X \left[(\hat{\theta}_{n,k}(X) - \theta)^2 \right] \text{ as } k \to \infty$$

by the portmanteau lemma. Finally, because $\hat{\theta}_{k,n}$ is asymptotically efficient, we can show that

$$\mathrm{E}_{X_k} \left[(\hat{\theta}_{n,k}(X_k) - \theta)^2 \right] \to 0 \text{ as } n \to \infty,$$

since the expected value of the estimator tends to θ and its variance goes to 0 when $n \to \infty$. Therefore, given any $\varepsilon > 0$ and $\delta > 0$, we can find a sufficiently large n and k such that

$$\Pr_X \left(\left| \hat{\theta}_{n,k}(X) - \theta \right| \geq \varepsilon \right) \leq \frac{\mathrm{E}_X \left[(\hat{\theta}_{n,k}(X) - \theta)^2 \right]}{\varepsilon^2} < \delta$$

□

References

George Casella and Roger L Berger. *Statistical inference*, volume 2. Duxbury Pacific Grove, CA, 2002.

Wai-Ki Ching, Jia-Wen Gu, and Harry Zheng. On correlated defaults and incomplete information. *arXiv preprint arXiv:1409.1393*, 2014.

Helen Haworth, Christoph Reisinger, and William Shaw. Modelling bonds and credit default swaps using a structural model with contagion. *Quantitative Finance*, 8(7):669–680, 2008.

Hua He, William P Keirstead, and Joachim Rebholz. Double lookbacks. *Mathematical Finance*, 8(3):201–228, 1998.

Steven Kou, Haowen Zhong, et al. First-passage times of two-dimensional brownian motion. *Advances in Applied Probability*, 48(4):1045–1060, 2016.

Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

Hannes Risken. *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer-Verlag, 1989.

Laura Sacerdote, Massimiliano Tamborrino, and Cristina Zucca. First passage times of two-dimensional correlated processes: Analytical results for the wiener process and a numerical method for diffusion processes. *Journal of Computational and Applied Mathematics*, 296:275–292, 2016.

Sandro Salsa. *Partial differential equations in action: from modelling to theory*, volume 99. Springer, 2016.

Vladimir Viktorovich Shaidurov. *Multigrid methods for finite elements*, volume 318. Springer Science & Business Media, 2013.

John R Singler. Differentiability with respect to parameters of weak solutions of linear parabolic equations. *Mathematical and Computer Modelling*, 47(3):422–430, 2008.

Ivar Stakgold and Michael J Holst. *Green's functions and boundary value problems*, volume 99. John Wiley & Sons, 2011.