Abstract

Keywords: Diffusion equation, regular bounded domain

1 Motivation

2 Solution on $\Omega \subset \mathbb{R}$

In this Section we will demonstrate the method outlined in Section 1 where the solution is defined on a bounded interval on $\mathbb{R}$. In this case, we have the true solution to the diffusion equation. We will compare the asymptotic expansion to the true solutio.

The PDE we will solve is the following BC/IC problem

$$\frac{\partial}{\partial t} q(x,t) = \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2} q(x,t), \tag{1}$$

$$q(x,0) = \delta_{x_0}(x), \tag{2}$$

$$q(a,t) = q(b,t) = 0. \qquad \text{(i.e. } \Omega = [a,b]) \tag{3}$$

Without loss of generality we will assume

$$a = 0, \qquad b = 1.$$

Problem (1) - (3) can be solved in a variety of ways. We will use the method of images, which repeatedly reflects the fundamental solution

$$q_{fundamental}(x,t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left\{ -\frac{1}{2\sigma^2 t}(x-x_0)^2 \right\}$$

about the boundary points a and b. The steps for the full solutions are as follows:

Step 1: Select a kernel $f(x|t)$ for the basis expansion,

Step 2: Perform Gram-Schmidt orthogonalization on the polynomials basis,

Step 3: Compute the weight for each basis element,

Step 4: Profit.

2.1 A suitable kernel for the basis elements

As noted in the motivating Section 1, the kernel we will use must be in $C^\infty(a,b)$, and it must obey the boundary conditions. Moreover, the basis kernel must be chosen such that

i) derivatives $f'(x)$ can be computed easily,

ii) integrals $\int_\Omega x^m f(x|t)^2\, dx$ can be computed easily

Consideration i) suggests that $f(x|t)$ should be of polynomial form. Consideration ii) suggests that $f(x|t)$ should be a known pdf over $[a,b]$, taking on zero at a and b. Given these requirements, the Beta distribution comes to mind:

$$f(x|t,\alpha,\beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1},$$

where $B(\alpha,\beta)$ is the beta function. Our choice for α and β is not very restricted. However, we will outline a few heuristics by which we can choose these parameters. Note that there may exist and optimal choise for (α,β) in terms of the accuracy of the asymptotic expansion with respect to the true solution $q(x,t)$. However, we will not prove anything in this vein here.

First, as long as

$$\alpha, \beta > 1, \tag{4}$$

the mode for the distribution is guaranteed to exist, so that the boundary conditions are met.

Aside from $\alpha > 1$ and $\beta > 1$, we can pick any (α, β) pair for our kernel. However, given that $f(x|t)$ can be thought of as implicitly dependent upon t, and that the variance of the fundamental solution is $\sigma^2 t$, a first, reasonable guess for (α, β) can be given by the solution to the equation:

$$\begin{aligned} \mathrm{Var}[X] &:= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \sigma^2 t, \qquad (5) \\ p(X \in dx) &= f(x|t,\alpha,\beta). \end{aligned}$$

By the same logic, noting the mean for the fundamental solution, we can require

$$\begin{aligned} \mathrm{E}[X] &:= \frac{\alpha}{\alpha+\beta} = x_0, \qquad (6) \\ p(X \in dx) &= f(x|t,\alpha,\beta). \end{aligned}$$

Finally, we may require that $\alpha, \beta \in \mathbb{Z}$, since this will guarantee that

$$\frac{\partial^k}{\partial x^k} f(x|t,\alpha,\beta) = 0$$

for large enough k. [georgid: This may not prove important, but I will keep it here anyway]

Thus, to set α and β, we simultaneously solve (5) and (6), then round α and β to the closest integer greater than or equal to 2. Since α and β are dependent upon t, we will keep t in our notation for f, albeit implicitly. In other words, once we choose α and β, we will not be able to take derivatives of f with respect to t. We will denote the kernel as $f(x|\alpha,\beta;t)$.

2.2 Gram-Schmidt orthogonalization on the polynomials basis

The family of (polynomial) functions $\{x^m f(x|\alpha,\beta;t)\}_{m=0}^{\infty}$ spans the space of $L^2([a,b])$ functions. We generate the basis elements $\{u_m(x|\alpha,\beta;t)\}_{m=0}^{\infty}$ by setting

$$\begin{aligned} v_0(x|\alpha,\beta;t) &= f(x|\alpha,\beta;t), \\ u_0(x|\alpha,\beta;t) &= \frac{f(x|\alpha,\beta;t)}{\|f(x|\alpha,\beta;t\|}, \qquad (7) \\ \|f(x|\alpha,\beta;t)\| &\equiv \left(\int_\Omega f(x|\alpha,\beta;t)^2 dx\right)^{1/2}. \qquad (8) \end{aligned}$$

The integral in (8) is easy to compute because of the form we have chosen for the kernel f :

$$\begin{aligned} \int_\Omega f(x|\alpha,\beta;t)^2 dx &= \int_\Omega \left(\frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}\right)^2 dx, \\ &= \int_\Omega \frac{1}{B(\alpha,\beta)^2} x^{(2\alpha-1)-1}(1-x)^{(2\beta-1)-1} dx, \\ &= \frac{B(2\alpha-1,2\beta-1)}{B(\alpha,\beta)^2}, \\ \left(\int_\Omega f(x|\alpha,\beta;t)^2 dx\right)^{1/2} &= \sqrt{\frac{B(2\alpha-1,2\beta-1)}{B(\alpha,\beta)^2}}. \end{aligned}$$

Next, for $u_1(x;\alpha,\beta;t)$,

$$\begin{aligned} v_1(x|\alpha,\beta;t) &= xf(x|\alpha,\beta;t) - \langle xf(x|\alpha,\beta;t)|u_0(x|\alpha,\beta;t)\rangle\, u_0(x|\alpha,\beta;t) \qquad (9) \\ &= \left(x - \frac{\langle xf(x|\alpha,\beta;t)|u_0(x|\alpha,\beta;t)\rangle}{\|f(x|\alpha,\beta;t)\|}\right) f(x|\alpha,\beta;t) \qquad (10) \\ \langle xf(x|\alpha,\beta;t)|u_0(x|\alpha,\beta;t)\rangle &= \int_\Omega \frac{xf(x|\alpha,\beta;t)^2}{\|f(x|\alpha,\beta;t)\|} = \frac{1}{\|f(x|\alpha,\beta;t)\|}\int_\Omega \frac{1}{B(\alpha,\beta)^2} x^{2\alpha-1}(1-x)^{(2\beta-1)-1} dx \qquad (11) \\ u_1(x|\alpha,\beta;t) &= \frac{v_1(x|\alpha,\beta;t)}{\|v_1(x|\alpha,\beta;t)\|} \qquad (12) \\ \|v_1(x|\alpha,\beta;t)\| &= \int_\Omega \left(x - \frac{\langle xf(x|\alpha,\beta;t)|u_0(x|\alpha,\beta;t)\rangle}{\|f(x|\alpha,\beta;t)\|}\right)^2 f(x|\alpha,\beta;t)^2 dx \qquad (13) \\ u_1(x|\alpha,\beta;t) &= \frac{v_1(x|\alpha,\beta;t)}{\|v_1(x|\alpha,\beta;t)\|} = p_1(x) f(x|\alpha,\beta;t), \qquad (14) \end{aligned}$$

where $p_1(x)$ is a first-order polynomial. In general,

$$v_m(x|\alpha,\beta;t) = x^m f(x|\alpha,\beta;t) - \sum_{m'=0}^{m-1} \langle x^m f(x|\alpha,\beta;t) | u_{m'}(x|\alpha,\beta;t)\rangle\, u_{m'}(x|\alpha,\beta;t)$$
$$u_m(x|\alpha,\beta;t) = \frac{v_m(x|\alpha,\beta;t)}{\|v_m(x|\alpha,\beta;t)\|} \equiv p_m(x|\alpha,\beta;t) f(x|\alpha,\beta;t)$$

In finding the basis, we will have to perform two main types calcluations:

1) polynomial multiplication: $p_m(x|\alpha,\beta;t)p_n(x|\alpha,\beta;t)$
2) integration of the form: $\int_\Omega x^m f(x|\alpha,\beta;t)^2 dx$

In `R`, the package `mpoly` will be used to handle 1). Calculation 2) can be performed relatively easily due to the form of $f(x|\alpha,\beta;t)$, as show in (15).

$$\int_\Omega x^m f(x|\alpha,\beta;t)^2 dx = \int_\Omega x^m \frac{1}{B(\alpha,\beta)^2} x^{2\alpha-2}(1-x)^{2\beta-2} dx = \frac{1}{B(\alpha,\beta)^2}\int_\Omega x^{2\alpha+m-2}(1-x)^{2\beta-2}dx = \frac{B(2\alpha+m-1, 2\beta-1)}{B(\alpha,\beta)^2} \tag{15}$$

2.3 Computing the Weights of the Basis Elements

Given the set of orthonormal functions $\{u_m(x|\alpha,\beta;t)\}_{m=0}^{\infty}$ spanning $L^2([a,b])$, and assuming that $q(x,t) \in L^2([a,b])$, we can write down

$$q(x,t) = \sum_{m=0}^{\infty} c_m u_m(x|\alpha,\beta;t), \tag{16}$$
$$\text{with } c_m = \int_\Omega q(x,t) u_m(x|\alpha,\beta;t) dx. \tag{17}$$

Since each u_m is the product of two polynomials, $u_m(x|\alpha,\beta;t) \in C^\infty([a,b])$ and is square-integrable. Therefore, we can write

$$\int_\Omega \frac{\partial^k \delta_{x_0}(x)}{\partial x^2} u_m(x|\alpha,\beta;t)dx = (-1)^k \int_\Omega \delta_{x_0}(x) \frac{\partial^k u_m(x|\alpha,\beta;t)}{\partial x^k} dx = (-1)^k \left.\frac{\partial^k u_m(x|\alpha,\beta;t)}{\partial x^k}\right|_{x=x_0} \tag{18}$$

Equipped with (18) and that $q(x,0) = \delta_{x_0}(x)$, we can compute the integrals in (17) by using the Taylor expansion of $q(x,t)$ about $t=0$:

$$\begin{aligned}
c_m(t;\alpha,\beta,x_0) = \int_\Omega q(x,t)u_m(x|\alpha,\beta;t)dx &= \int_\Omega \left[\underbrace{q(x,0)}_{\delta_{x_0}(x)} + \sum_{k=1}^{\infty} \frac{t^k}{k!} \left.\frac{\partial^k q(x,t)}{\partial t^k}\right|_{t=0} \right] u_m(x|\alpha,\beta;t)dx \\
&= \int_\Omega \delta_{x_0}(x) u_m(x|\alpha,\beta;t)dx + \sum_{k=1}^{\infty} \frac{t^k}{k!} \int_\Omega \left.\frac{\partial^k q(x,t)}{\partial t^k}\right|_{t=0} u_m(x|\alpha,\beta;t)dx \\
&= \int_\Omega \delta_{x_0}(x) u_m(x|\alpha,\beta;t)dx + \sum_{k=1}^{\infty} \frac{t^k}{k!} \int_\Omega \left(\frac{1}{2}\sigma^2\right)^k \frac{\partial^{2k}\delta_{x_0}(x)}{\partial x^{2k}} u_m(x|\alpha,\beta;t)dx \\
&= u_m(x_0|\alpha,\beta;t) + \sum_{k=1}^{\infty} \frac{t^k}{k!}(-1)^{2k} \left(\frac{1}{2}\sigma^2\right)^k \int_\Omega \delta_{x_0}(x) \frac{\partial^{2k} u_m(x|\alpha,\beta;t)}{\partial x^{2k}} dx \\
&= u_m(x_0|\alpha,\beta;t) + \sum_{k=1}^{\infty} \frac{t^k}{k!}(-1)^{2k} \left(\frac{1}{2}\sigma^2\right)^k \frac{\partial^{2k} u_m(x_0|\alpha,\beta;t)}{\partial x^{2k}} \\
&= u_m(x_0|\alpha,\beta;t) + \sum_{k=1}^{\infty} \frac{t^k}{k!} \left(\frac{1}{2}\sigma^2\right)^k \frac{\partial^{2k} u_m(x_0|\alpha,\beta;t)}{\partial x^{2k}}.
\end{aligned}$$

Note that

$$
\begin{aligned}
u_m^{(k)}(x|\alpha,\beta;t) &= \sum_{j=0}^{k} \begin{pmatrix} k \\ j \end{pmatrix} p_m^{(j)}(x|\alpha,\beta;t) f^{(k-j)}(x|\alpha,\beta;t) \\
f'(x|\alpha,\beta;t) &= \frac{1}{B(\alpha,\beta)} \left[(\alpha-1)x^{\alpha-2}(1-x)^{\beta-1} + (\beta-1)x^{\alpha-1}(1-x)^{\beta-2}\right] \\
&= \frac{1}{B(\alpha,\beta)} \left[(\alpha-1)B(\alpha-1,\beta)f(x|\alpha-1,\beta;t) + (\beta-1)B(\alpha,\beta-1)f(x|\alpha,\beta-1;t)\right]
\end{aligned}
$$

The k^{th} order derivative of the polynomial $p_m(x|\alpha,\beta;t)$ can be computed on the fly with `mpoly`. The recursive definition of

3 Another Kernel Choice

Another suitable kernel choise the bump function

$$
f(x) = \begin{cases} \exp\left\{-\frac{1}{1-x^2}\right\}, & \text{if } |x| < 1 \\ 0, & \text{otherwise} \end{cases} \tag{19}
$$

In thise case, $f \in C^{\infty}(-1,1)$. As outlined perviously, we need to compute

$$
f^{(k)}(x), \int_{-1}^{1} x^m f(x)^2\, dx
$$

For the first calculation, note that

$$
f'(x) = f(x)(1-x^2)^{-2}(-2x) \equiv f(x)P_{1,0}(x)^{-1}P_{1,1}(x)
$$

Assuming that

$$
f^{(k)}(x) = f(x)P_{k,0}(x)^{-1}P_{k,1}(x),
$$

we have that

$$
\begin{aligned}
f^{(k+1)}(x) &= f'(x)P_{k,0}(x)^{-1}P_{k,1}(x) \\
&\qquad + f(x)\left[-P_{k,0}(x)^{-2}P'_{k,0}(x)P_{k,1}(x) + P_{k,0}(x)^{-1}P'_{k,1}(x)\right] \\
&= \left(f(x)P_{1,0}(x)^{-1}P_{1,1}(x)\right)P_{k,0}(x)^{-1}P_{k,1}(x) \\
&\qquad + f(x)P_{k,0}(x)^{-2}\left[-P'_{k,0}(x)P_{k,1}(x) + P_{k,0}(x)P'_{k,1}(x)\right] \\
&= f(x)\left(P_{k,0}(x)P_{1,0}(x)\right)^{-1}P_{1,1}(x)P_{k,1}(x) \\
&\qquad + f(x)P_{k,0}(x)^{-2}\left[-P'_{k,0}(x)P_{k,1}(x) + P_{k,0}(x)P'_{k,1}(x)\right]
\end{aligned}
$$

$$
f^{(k+1)}(x) = f(x)\left[P_{k,0}(x)P_{1,0}(x)P_{k,0}(x)^2\right]^{-1}\left(P_{1,1}(x)P_{k,1}(x)P_{k,0}(x)^2 + P_{k,0}(x)P_{1,0}(x)\left[-P'_{k,0}(x)P_{k,1}(x) + P_{k,0}(x)P'_{k,1}(x)\right]\right)
$$

$$
\boxed{P_{k+1,0} = P_{k,0}(x)P_{1,0}(x)P_{k,0}(x)^2}
$$

$$
\boxed{P_{k+1,1} = P_{1,1}(x)P_{k,1}(x)P_{k,0}(x)^2 + P_{k,0}(x)P_{1,0}(x)\left[-P'_{k,0}(x)P_{k,1}(x) + P_{k,0}(x)P'_{k,1}(x)\right]}
$$

4 Mesh-Free Finite Element Method

Here we take a similar approach, where we seek the *weak solution* of the PDE (1) - (3). A weak solution is any function $q_{weak}(x,t)$ where, for any $\phi(x) \in C_0^{\infty}(\Omega)$,

$$
\int_{\Omega} \partial_t q_{weak}(x,t)\phi(x)\,dx - \frac{1}{2}\sigma^2 \int_{\Omega} \partial_x^2 q_{weak}(x,t)\phi(x)\,dx = 0. \tag{20}
$$

We can relax the definition of a weak solution in the following way. Consider a countable family of functions, or elements, $\{\psi_k(x)\}_{k=1}^{\infty}$ dense in the space of all $L_2(\Omega)$ functions. It can be shown that a function is a weak solution to the problem if (20) holds for each $\psi_k(x)$. In other words,

$$\int_{\Omega} \partial_t q_{weak}(x,t)\psi_k(x)\,dx - \frac{1}{2}\sigma^2 \int_{\Omega} \partial_x^2 q_{weak}(x,t)\psi_k(x)\,dx = 0. \tag{21}$$

It should be noted that each element $\psi_k(x)$ need not be $C_0^{\infty}(\Omega)$. In fact, under the Ritz method, $\psi_k(x) \in C_0^1(\Omega)$ in the weak sense. As long as the basis elements are dense in L_2 and they satisfy the boundary conditions, we can solve the problem in the weak formulation. This can be proved using Friedrichs' mollification (See exercise 2.12 in **?**). Furthermore, we can show that the weak solution converges to the classical solution. In application, we restrict ourselves to a subset of $\{\psi_k(x)\}_{k=1}^{\infty}$:

$$\{\psi_k(x)\}_{k=1}^{K},$$

applying (21) only to elements $\psi_1(x)$ through $\psi_K(x)$. For the purposes of this section, we will work with a finite family of basis elements which are orthonormal (this can be achieved by following the Gram-Schmidt procedure outlined above). Further, we impose the form for the approximate solution $q_K(x,t)$:

$$q_K(x,t) = \sum_{k=1}^{K} \xi_k(t)\psi_k(x).$$

Plugging into (21),

$$\int_{\Omega} \sum_{k=1}^{K} (\partial_t \xi_k(t))\psi_k(x)\psi_l(x)\,dx - \frac{1}{2}\sigma^2 \int_{\Omega} \sum_{k=1}^{K} \xi_k(t)(\partial_x^2 \psi_k(x))\psi_l(x)\,dx = 0,$$
$$\sum_{k=1}^{K} (\partial_t \xi_k(t)) \int_{\Omega} \psi_k(x)\psi_l(x)\,dx - \frac{1}{2}\sigma^2 \sum_{k=1}^{K} \xi_k(t) \int_{\Omega} (\partial_x^2 \psi_k(x))\psi_l(x)\,dx = 0.$$

Using the boundary conditions and integration by parts,

$$\int_{\Omega} (\partial_x^2 \psi_k(x))\psi_l(x)\,dx = -\int_{\Omega} \partial_x \psi_k(x) \partial_x \psi_l(x)\,dx.$$

Using the orthonormality of $\psi_k(x)$,

$$\int_{\Omega} \psi_k(x)\psi_l(x)\,dx = \delta(k-l).$$

The solution condition then becomes

$$\sum_{k=1}^{K} (\partial_t \xi_k(t))\delta(k-l) + \frac{1}{2}\sigma^2 \sum_{k=1}^{K} \xi_k(t) \int_{\Omega} \partial_x \psi_k(x))\partial_x \psi_l(x)\,dx = 0. \tag{22}$$

Defining

$$\xi(t) = (\xi_1(t), \ldots, \xi_K(t))^T, \qquad \psi(x) = (\psi_1(x), \ldots, \psi_K(x))^T,$$

condition (22) sets up the system of equations

$$\partial_t \xi(t) = -\frac{1}{2}\sigma^2 A \xi(t), \qquad [A]_{ij} = \int_{\Omega} \partial_x \psi_i(x))\partial_x \psi_j(x)\,dx. \tag{23}$$

The matrix A is symmetric, so that an eigendecomposition for it exists. Moreover, the solution to the system of ODEs in (23) is given by

$$\xi(t) = e^{-\frac{1}{2}\sigma^2 At}\xi(0), \quad \xi_k(0) = \int_{\Omega} \psi_k(x)\delta(x-x_0)\,dx = \psi_k(x_0), \quad q_K(x,t) = \psi(x)^T \xi(t) = \psi(x)^T \left(e^{-\frac{1}{2}\sigma^2 At}\right)\xi(0). \tag{24}$$

The definition for $\xi(0)$ can be derived from the orthonormality of $\{\psi_k(x)\}_{k=1}^{K}$. Obviously, sparsity in $\{\psi_k(x)\}_{k=1}^{K}$ makes calculating the solution easy.

4.1 Basis Element Choice

It is known that the family of Gaussian pdfs is dense in $L_2(\mathbb{R})$ (See exercise 3.6 in **?**). This motivates the choice for the family of basis elements that we will use for this problem:

$$\left\{(x-a)(b-x)\phi(x|\mu_k,\sigma_k^2)\right\}_{k=1}^K, \qquad \phi(x|\mu_k,\sigma_k^2) = \frac{1}{\sqrt{2\pi\sigma_k^2}}\exp\left\{-\frac{1}{2\sigma_k^2}(x-\mu_k)^2\right\}. \tag{25}$$

Multiplying the Gaussian densities ϕ by $(x-a)(b-x)$ ensures that each basis element obeys the boundary conditions. We have to prove the the above family of basis elements is dense in $L_2(\Omega)$.

The choice of (μ_k,σ_k^2) is up to us. However, for a fixed K, we can choose a (locally) optimum sequences of $(\mu_k,\sigma_k^2)_{k=1}^K$ in the following way. Given $(\mu_k,\sigma_k^2)_{k=1}^K$, we can first follow Gram-Schmidt to make the basis elements orthonormal, then use the solution in (24) to compute the weak solution $q_K(x,t)$ for each $x\in\Omega$. Then we can compute

$$\left\|\partial_t q_K(x,t) - \frac{1}{2}\sigma^2\partial_x^2 q_K(x,t)\right\|_{2,\Omega}. \tag{26}$$

This norm will tell us approximately how closely the weak solution $q_K(x,t)$ solves the original problem in the classical sense. We can then use `R`'s `optim` to search over the space $(\mu_k,\sigma_k^2)_{k=1}^K$ for the best sequence by minimizing (26), given some starting point. In this case, the starting point for μ_k are evenly spaced points on the domain $[a,b]$, including the boundaries. The starting point for the basis variances are all $\sigma_k^2 = 0.05$. We can see in Figure (1) that already for $K=6$ the approximate solution reaches a qualitatively reasonble accuracy, and certainly good enough for effective maximum likelihood estimation.

4.2 Path-Integral Formulation and Differentiation With Respect to Boundaries

The Taylor expansion of the semigroup operator in (24) can be used to obtain the path-integral form for this problem:

$$\begin{aligned}\xi(t+\Delta t) &= e^{-\frac{1}{2}\sigma^2 A\Delta t}\xi(t) \approx \left(I - \frac{1}{2}\sigma^2 A\Delta t\right)\xi(t),\\ \Rightarrow q(x,t+\Delta t) &\approx \psi^T(x)\left(I-\frac{1}{2}\sigma^2 A\Delta t\right)\xi(t),\end{aligned} \tag{27}$$

for a small enought Δt. We will lend (27) a *probabilistic* interpretation. First, if we recall that the elements in $\psi(x)$ are orthonormal, we will notice that the matrix

$$\int_\Omega \psi(x)\psi^T(x)dx = I.$$

Therefore, we can re-write (27) as

$$\begin{aligned}q(x,t+\Delta t) &\approx \psi^T(x)\left(I-\frac{1}{2}\sigma^2 A\Delta t\right)\xi(t),\\ &= \psi^T(x)\left(I-\frac{1}{2}\sigma^2 A\Delta t\right)e^{-\frac{1}{2}\sigma^2 At}\xi(0),\\ &= \int_\Omega \psi^T(x)\left(I-\frac{1}{2}\sigma^2 A\Delta t\right)\psi(y)\psi^T(y)e^{-\frac{1}{2}\sigma^2 At}\xi(0)dy,\end{aligned}$$

The first and second parts in the integral are the probabilities

$$\begin{aligned}P(X_{t+\Delta t} = x, a < X_t < b, \forall t\in[t,t+\Delta t]|X_t = y) &\approx \psi^T(x)\left(I-\frac{1}{2}\sigma^2 A\Delta t\right)\psi(y)\\ P(X_t = y, a<X_t<b,\forall t\in[0,t]|X_0 = x_0) &= \psi^T(y)e^{-\frac{1}{2}\sigma^2 At}\xi(0)\\ \Rightarrow \int_\Omega P(X_{t+\Delta t} = x, a<X_t<b,\forall t\in[t,t+\Delta t]|X_t=y)&P(X_t=y,a<X_t<b,\forall t\in[0,t]|X_0=x_0)dy\\ &\approx \int_\Omega \psi^T(x)\left(I-\frac{1}{2}\sigma^2 A\Delta t\right)\psi(y)\psi^T(y)e^{-\frac{1}{2}\sigma^2 At}\xi(0)dy\end{aligned}$$

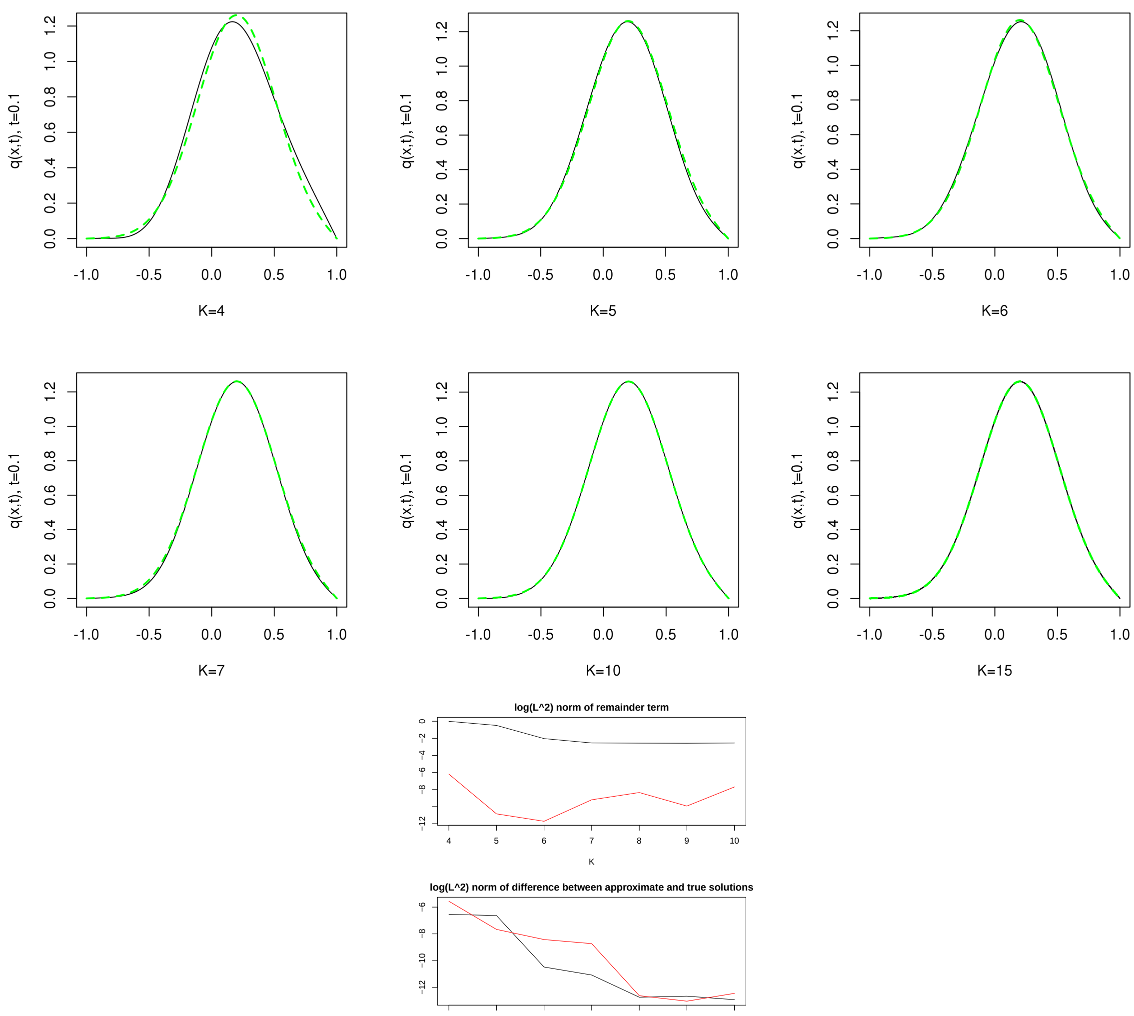

Figure 1: Approximate solution (solid black) and true solution (dashed green) for different number K of optimally chosen basis elements $\left\{(a-x)(b-x)\phi(x|\mu_k,\sigma_k^2)\right\}_{k=1}^{K}$.

In general, again for small enough Δt,

$$P(X_t = x, a < X_t < b, \forall t \in [0,t] | X_0 = x_0) \approx q_K(x,t) \approx$$
$$\approx \int_\Omega \cdots \int_\Omega \left[\psi^T(x)\left(I - \frac{1}{2}\sigma^2 A\Delta t\right)\psi(x_n)\right]\left[\psi^T(x_{n-1})\left(I - \frac{1}{2}\sigma^2 A\Delta t\right)\psi(x_{n-2})\right]\cdots\left[\psi^T(x_1)\left(I - \frac{1}{2}\sigma^2 A\Delta t\right)\psi(x_0)\right] dx_n \cdots dx_1 \tag{28}$$

Equation (28) is the **path integral formulation** for the solution to the initial-BV problem. We can reduce (28) to

$$q_K(x,t) \approx \psi^T(x)\left(I - \frac{1}{2}\sigma^2 A\Delta t\right)^{t/\Delta t}\psi(x_0).$$

Before we take derivatives with respect to the boundaries, first note

$$P(X_t = x, a < X_t < b, \forall t \in [0,t] | X_0 = x_0) = P(X_t = x, \max_{t' \in [0,t]}\{X_{t'}\} < b, \min_{t' \in [0,t]}\{X_{t'}\} > a | X_0 = x_0) = q(x,t).$$

Hence,

$$P\left(X_t = x, \max_{t' \in [0,t]}\{X_{t'}\} = b, \min_{t' \in [0,t]}\{X_{t'}\} = a | X_0 = x_0\right) = -\frac{\partial^2}{\partial a \partial b} q(x,t). \tag{29}$$

This is what we are after. To compute the derivative, we can do

$$-\frac{\partial^2}{\partial a \partial b} q(x,t) \approx -\frac{\partial^2}{\partial a \partial b}\left(\psi(x)^T\left(e^{-\frac{1}{2}\sigma^2 At}\right)\psi(0)\right) \approx -\frac{\partial^2}{\partial a \partial b}\left(\psi^T(x)\left(I - \frac{1}{2}\sigma^2 A\Delta t\right)^{t/\Delta t}\psi(x_0)\right). \tag{30}$$

The above derivatives are expensive to compute. Moreover, at a first glance, they have no immediate probabilistic interpretation. However, under the path-integral formulation, they do. From (28)

$$-\frac{\partial^2}{\partial a \partial b} q_K(x,t) = \int_\Omega \cdots \int_\Omega -\frac{\partial^2}{\partial a \partial b}\left(\left[\psi^T(x)\left(I - \frac{1}{2}\sigma^2 A\Delta t\right)\psi(x_n)\right]\cdots\left[\psi^T(x_1)\left(I - \frac{1}{2}\sigma^2 A\Delta t\right)\psi(x_0)\right]\right) dx_n \cdots dx_1. \tag{31}$$

Note that I have placed the derivatives on the inside of the integral. This follows from the Dominated Convergence Theorem since the derivatives are bounded over Ω. This is not cheaper than the other representations, however, consider one of the elements in the sum of derivatives:

$$-\int_\Omega \cdots \int_\Omega \frac{\partial}{\partial a}\left[\psi^T(x)\left(I - \frac{1}{2}\sigma^2 A\Delta t\right)\psi(x_n)\right]\cdots\frac{\partial}{\partial b}\left[\psi^T(x_1)\left(I - \frac{1}{2}\sigma^2 A\Delta t\right)\psi(x_0)\right] dx_n \cdots dx_1.$$

The above expression is the probability

$$P(X_t = x, \min\{X_t\} = a, \max\{X_t\} = b, \tau_a \in [t_n, t], \tau_b \in [0, t_1] | X_0 = x_0),$$

where $\tau_a(\tau_b)$ is the time when X_t reaches boundary a (b). This motives us to write down

$$q(x,t|\tau_a \in \Delta t_k, \tau_b \in \Delta t_l) = -\int_\Omega \cdots \int_\Omega \frac{\partial}{\partial a}\left[\psi^T(x)\left(I - \frac{1}{2}\sigma^2 A\Delta t\right)\psi(x_n)\right]\cdots\frac{\partial}{\partial b}\left[\psi^T(x_1)\left(I - \frac{1}{2}\sigma^2 A\Delta t\right)\psi(x_0)\right] dx_n \cdots dx_1.$$

Therefore, in the context of MCMC, if we introduce the auxiliary variables τ_a, τ_b, we can make evaluations of the likelihood much faster. The next questions is, therefore, how to sample the hitting times τ_a and τ_b.

4.3 Sampling hitting times

Let τ be the first hitting time for Ω. By definition of our solution $q(x,t)$,

$$P(\tau > t) = \int_\Omega q(x,t)dx \Rightarrow P(\tau \le t) = 1 - \int_\Omega q(x,t)dx$$
$$p(\tau = t) = -\frac{\partial}{\partial t}\int_\Omega q(x,t)dx$$

Appendices

A Calculation

First consider the product

$$
\begin{aligned}
\phi(x|\mu_k,\sigma_k^2)\phi(x|\mu_l,\sigma_l^2) &= \frac{1}{\sqrt{2\pi\sigma_k^2}}\cdot\frac{1}{\sqrt{2\pi\sigma_l^2}}\cdot\exp\left\{-\frac{1}{2}\left[\frac{(x-\mu_k)^2}{\sigma_k^2}+\frac{(x-\mu_l)^2}{\sigma_l^2}\right]\right\} \\
&= \frac{1}{\sqrt{4\pi^2}\sqrt{\sigma_k^2\sigma_l^2}}\cdot\exp\left\{-\frac{1}{2}\left[\frac{x^2-2x\mu_k+\mu_k^2}{\sigma_k^2}+\frac{x^2-2x\mu_l+\mu_l^2}{\sigma_l^2}\right]\right\} \\
&= \frac{1}{\sqrt{4\pi^2}\sqrt{\sigma_k^2\sigma_l^2}}\cdot\exp\left\{-\frac{1}{2}\left[x^2\left(\frac{1}{\sigma_k^2}+\frac{1}{\sigma_l^2}\right)-2x\left(\frac{\mu_k}{\sigma_k^2}+\frac{\mu_l}{\sigma_l^2}\right)+\left(\frac{\mu_k^2}{\sigma_k^2}+\frac{\mu_l^2}{\sigma_l^2}\right)\right]\right\} \\
&= \frac{1}{\sqrt{4\pi^2}\sqrt{\sigma_k^2\sigma_l^2}}\cdot\exp\left\{-\frac{1}{2}\left(\frac{\mu_k^2}{\sigma_k^2}+\frac{\mu_l^2}{\sigma_l^2}\right)\right\}\cdot\exp\left\{-\frac{1}{2}\left[x^2\left(\frac{1}{\sigma_k^2}+\frac{1}{\sigma_l^2}\right)-2x\left(\frac{\mu_k}{\sigma_k^2}+\frac{\mu_l}{\sigma_l^2}\right)\right]\right\} \\
&= \frac{1}{\sqrt{4\pi^2}\sqrt{\sigma_k^2\sigma_l^2}}\cdot\exp\left\{-\frac{1}{2}\left(\frac{\mu_k^2}{\sigma_k^2}+\frac{\mu_l^2}{\sigma_l^2}\right)\right\} \\
&\quad\times\exp\left\{-\frac{1}{2}\left(\frac{1}{\sigma_k^2}+\frac{1}{\sigma_l^2}\right)\left[x-\left(\frac{1}{\sigma_k^2}+\frac{1}{\sigma_l^2}\right)^{-1}\left(\frac{\mu_k}{\sigma_k^2}+\frac{\mu_l}{\sigma_l^2}\right)\right]^2\right\} \\
&\quad\times\exp\left\{\frac{1}{2}\left(\frac{1}{\sigma_k^2}+\frac{1}{\sigma_l^2}\right)^{-1}\left(\frac{\mu_k}{\sigma_k^2}+\frac{\mu_l}{\sigma_l^2}\right)^2\right\} \\
&= \frac{1}{\sqrt{4\pi^2}\sqrt{\sigma_k^2\sigma_l^2}}\cdot\exp\left\{-\frac{1}{2}\left(\frac{\mu_k^2}{\sigma_k^2}+\frac{\mu_l^2}{\sigma_l^2}\right)\right\}\cdot\exp\left\{\frac{1}{2}\left(\frac{1}{\sigma_k^2}+\frac{1}{\sigma_l^2}\right)^{-1}\left(\frac{\mu_k}{\sigma_k^2}+\frac{\mu_l}{\sigma_l^2}\right)^2\right\} \\
&\quad\times\ker\left(x\,\middle|\left(\frac{1}{\sigma_k^2}+\frac{1}{\sigma_l^2}\right)^{-1}\left(\frac{\mu_k}{\sigma_k^2}+\frac{\mu_l}{\sigma_l^2}\right),\left(\frac{1}{\sigma_k^2}+\frac{1}{\sigma_l^2}\right)^{-1}\right) \\
&:= C(\mu_k,\sigma_k^2;\mu_l,\sigma_l^2)\ker\left(x\,\middle|\left(\frac{1}{\sigma_k^2}+\frac{1}{\sigma_l^2}\right)^{-1}\left(\frac{\mu_k}{\sigma_k^2}+\frac{\mu_l}{\sigma_l^2}\right),\left(\frac{1}{\sigma_k^2}+\frac{1}{\sigma_l^2}\right)^{-1}\right) \\
C(\mu_k,\sigma_k^2;\mu_l,\sigma_l^2) &:= \frac{1}{\sqrt{4\pi^2\sigma_k^2\sigma_l^2}}\cdot\exp\left\{-\frac{1}{2}\left(\frac{\mu_k^2}{\sigma_k^2}+\frac{\mu_l^2}{\sigma_l^2}\right)\right\}\cdot\exp\left\{\frac{1}{2}\left(\frac{1}{\sigma_k^2}+\frac{1}{\sigma_l^2}\right)^{-1}\left(\frac{\mu_k}{\sigma_k^2}+\frac{\mu_l}{\sigma_l^2}\right)^2\right\}
\end{aligned}
$$

$$
\begin{aligned}
\tilde{\psi}_k(x) &= (x-a)(b-x)\phi(x|\mu_k,\sigma_k^2) \\
\int_\Omega\tilde{\psi}_k(x)^2dx &= \int_\Omega(x-a)^2(b-x)^2\phi(x|\mu_k,\sigma_k^2)^2\,dx \\
&= \int_\Omega(x-a)^2(b-x)^2C(\mu_k,\sigma_k^2;\mu_k,\sigma_k^2)\phi(x|\mu_k,\sigma_k^2/2)\,dx \\
&= C(\mu_k,\sigma_k^2;\mu_k,\sigma_k^2)\int_\Omega(x^2-2ax+a^2)(b^2-2xb+x^2)\phi(x|\mu_k,\sigma_k^2/2)\,dx \\
&= C(\mu_k,\sigma_k^2;\mu_k,\sigma_k^2)\int_\Omega(x^4-x^3(2b+2a)+x^2(a^2+b^2+4ab)-x(2ab^2+2ba^2)+a^2b^2)\phi(x|\mu_k,\sigma_k^2/2)\,dx \\
&= C(\mu_k,\sigma_k^2;\mu_k,\sigma_k^2)\left(M_4-M_3(2b+2a)+M_2(a^2+b^2+4ab)-M_1(2ab^2+2ba^2)+a^2b^2\right) \\
M_k &:= \int_\Omega x^k\phi(x|\mu_k,\sigma_k^2/2)\,dx
\end{aligned}
$$