Polynomial Chaos Expansion (PCE)

Valerie Tsao

October 30, 2025

Contents

1 Introduction

This section is more meant for me to summarize some of the things I read, as well as put some thought into digesting these concepts on my own. A lot of the ideas I jot down here are because I thought them necessary to know in order to complete the exercises that follow. Please feel free to skip over this section and proceed to the solutions below.

Polynomial chaos expansion (PCE) is a way for us to represent random variables in terms of (orthogonal) polynomials, similar to a Taylor/Maclaurin series or a Fourier series. This series gets truncated in practice to get accurate approximations that converge much faster than typical Monte Carlo simulations. They also allow for closed-form formulas for mean, variance, etc. unlike other forms of decompositions/expansions. Primarily, this has been used in uncertainty quantification and sensitivty analysis applications to easily summarize distributions that may have weird forms. Unlike in Principal Component Analysis (PCA), where the basis is data-dependent (i.e. the eigenvectors of a covariance matrix), the basis for PCE can be chosen a-priori according to the input distribution.

2 Some preliminaries and notes

2.1 Orthogonal polynomials

So, why do we need *orthogonal* polynomials? Let's say we were trying to approximate a function using a basis of vectors, like sines and cosines in Fourier series. If our basis vectors are orthogonal, computing coefficients is easy; we can just project onto each basis element independently. However,

if they're not orthogonal, then we'd need to solve a system of equations, which is very computationally expensive, and doesn't always give the unique best coefficients. The same principle applies here, with the added fact that orthogonality depends on the probability distribution now.

When polynomials are orthogonal, this means that

$$< P_n, P_m >= \int P_n(\xi) P_m(\xi) \rho(\xi) d\xi = 0 \quad \text{for} \quad n \neq m,$$

where $\rho(\xi)$ is the probability density function (PDF). From this, we can see that even with the same functions but different distributions, we can get different orthogonality results.

There are a few classical orthogonal polynomials, like the Hermite polynomials (for Gaussian distribution), the Laguerre polynomials (for Gamma distribution), and the Jacobi polynomials (for Beta distribution).[2] If we are working with a distribution that doesn't match one of the named few, often we need to perform a transformation of random variables (i.e. via an inverse CDF transform) to one that does match (for example, $U[a, b] \mapsto U[-1, 1]$).

Legendre polynomials A special case of the Jacobi is the Legendre polynomials, which are associated with the Uniform distribution on $[-1, 1]$. They are the natural basis for functions defined in an interval with no preference for any particular value.

Speaking of which, our first exercise is defined on the Uniform distribution! The first few Legendre polynomials are defined by:

$$\begin{aligned}
P_0(x) &= 1, \\
P_1(x) &= x, \\
P_2(x) &= \frac{1}{2}(3x^2 - 1), \\
P_3(x) &= \frac{1}{2}(5x^3 - 3x), \\
P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3), \\
&\vdots
\end{aligned}$$

Rigorously, Legendre polynomials are orthogonal with respect to the uniform measure on $[-1, 1]$:

$$\int_{-1}^{1} P_n(x) P_m(x) dx = \frac{2}{2n+1} \delta_{nm}.$$

They can be calculated by the following recurrence relation:

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x),$$

with $P_0(x) = 1$ and $P_1(x) = x$. Alternatively though, they can be defined using Rodrigues' Formula, which gives us:

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}[(x^2 - 1)^n].$$

Hermite polynomials Another set of orthogonal polynomials are the Hermite polynomials, associated with the standard Normal distribution $\mathcal{N}(0, 1)$. They are the natural basis for Gaussian randomness, which is quite prevalent due to the Central Limit Theorem.

For the ensuing notes, we will follow the probabilist's Hermite polynomials, He_n as opposed to the physicist's H_n,[1] taking note that there does indeed exist two conventions. The first few Hermite polynomials are:

$$\begin{aligned}
\mathrm{He}_0(x) &= 1, \\
\mathrm{He}_1(x) &= x, \\
\mathrm{He}_2(x) &= x^2 - 1, \\
\mathrm{He}_3(x) &= x^3 - 3x, \\
\mathrm{He}_4(x) &= x^4 - 6x^2 + 3, \\
\mathrm{He}_5(x) &= x^5 - 10x^3 + 15x, \\
&\vdots
\end{aligned}$$

We see here that each polynomials has degree n and removes the mean of x^n under the Gaussian measure:

$$\mathbb{E}[\mathrm{He}_n(\xi)] = 0 \quad \text{for } \xi \sim \mathcal{N}(0,1) \text{ and } n \geq 1.$$

Rigorously, Hermite polynomials are orthogonal with respect to the standard Gaussian measure:

$$\mathbb{E}[\mathrm{He}_n(\xi)\mathrm{He}_m(\xi)] = \int_{-\infty}^{\infty} \mathrm{He}_n(x)\mathrm{He}_m(x)\frac{e^{-x^2/2}}{\sqrt{2\pi}}dx = n! \cdot \delta_{nm}$$

Note the normalization: $\mathbb{E}[\mathrm{He}_n^2] = n!$ (meaning factorial growth). We calculate them with the following recurrence relation:

$$\mathrm{He}_{n+1}(x) = x \cdot \mathrm{He}_n(x) - n \cdot \mathrm{He}_{n-1}(x),$$

with $\mathrm{He}_0(x) = 1$ and $\mathrm{He}_1(x) = x$. Using Rodrigues' Formula again, we can find

$$\mathrm{He}_n(x) = (-1)^n e^{x^2/2}\frac{d^n}{dx^n}e^{-x^2/2}.$$

There exists deeper connections to stochastic theory as for why Hermite polynomials represent Gaussians, but we won't get into that here. For more information, see the Focker-Plank equation,[3;4] Wiener chaos,[5] and optimal basis.[6]

Key Property: If ξ follows distribution with density $\rho(\xi)$, then the polynomials $P_n(\xi)$ satisfy

$$\int P_n(\xi)P_m(\xi)\rho(\xi)d\xi = h_n\delta_{nm}$$

where δ_{nm} is the Kronecker delta and h_n is a normalization constant.

Main Ideas

⇒ For Legendre polynomials, the normalization constants are:

$$h_n = \int_{-1}^{1} P_n(\xi) P_m(\xi) d\xi = \frac{2}{2n+1}.$$

Note: this assumes $U[0,1]$! If not the case, we need to multiply by the PDF of the uniform distribution.

⇒ For Hermite polynomials, the normalization constants are:

$$h_n = \int_{-\infty}^{\infty} \mathrm{He}_n(\xi) \mathrm{He}_m(\xi) d\xi = \sqrt{2\pi} n!.$$

Note: this assumes $\mathcal{N}(0,1)$. For non-standard Gaussian, we'd need to multiply by another normalization constant.

2.2 Polynomial Chaos Expansion

Definition

The PCE of a random variable Y depending on random input ξ with a known distribution can be defined as

$$Y = \sum_{t=0}^{\infty} c_i \Psi_i(\xi),$$

where $\Psi_i(\xi)$ are orthogonal polynomials with respect to the distribution of ξ, and c_i are deterministic coefficients.

The coefficient formula. To compute the coefficients, we use orthogonality:

$$c_i = \frac{\mathbb{E}[Y \Psi_i(\xi)]}{\mathbb{E}[\Psi_i^2(\xi)]}.$$

In practice, we truncate to order P:

$$Y \approx \sum_{i=0}^{P} c_i \Psi_i(\xi).$$

2.3 Computing statistics using PCEs

This is where PCE shines! Once we have the expansion $Y \approx \sum_{i=0}^{P} c_i \Psi_i(\xi)$, the mean easily follows:

$$\mathbb{E}[Y] = c_0.$$

This is because $\Psi_0 = 1$ and all other polynomials have zero mean. The variance is similar:

$$Var(Y) = \sum_{i=1}^{P} c_i^2 \mathbb{E}[\Psi_i^2].$$

2.4 Computing coefficients in practice

There exist two main approaches to computing our coefficients: the projection method and the Galerkin method.

⇒ The non-intrusive method involves computing the expected values numerically:

$$c_i = \frac{1}{\mathbb{E}[\Psi_i^2]} \int Y(\xi)\, \Psi_i(\xi)\, \rho(\xi)\, d\xi \approx \frac{1}{\mathbb{E}[\Psi_i^2]} \sum_{k=1}^{N_q} Y(\xi_k)\, \Psi_i(\xi_k)\, w_k,$$

where ξ_k are quadrature points and w_k are weights. This process is called non-intrusive because we do not modify the governing equations whatsoever. It does, however, suffer from the curse of dimensionality, so when working in high dimensions this often becomes too expensive.

⇒ The Galerkin method involves substituting the expansion directly into the governing equations and projecting onto each basis function. This is called intrusive because we must modify/approximate the governing equations, and so it no longer strictly enforces them.

For linear ODEs, these two methods are equivalent. For nonlinear ODEs, the truncation errors differ. The Galerkin approach propagates uncertainty more accurately but requires computing expectation of nonlinear terms.

3 Exercises

Problem 1

Let X Uniform[-1,1] and $Y = \sin(\pi X)$.

(a) Expand Y using PCE.

(b) Determine $\mathbb{E}[Y]$ and $\mathrm{Var}(Y)$ from the expansion.

Solution. (a) In this case, since we are working with Uniform[-1, 1] random variables, so we use Legendre polynomials, and therefore $\Psi_i = P_n$. To compute the PCE coefficients, we write out the expansion:

$$Y = \sin(\pi X) = \sum_{n=0}^{\infty} c_n P_n(X).$$

For convenience, let's just write out the first four terms of the PCE of Y. Using the definition of the Legendre polynomials from earlier,

$$\begin{aligned} Y &\approx c_0 P_0(X) + c_1 P_1(X) + c_2 P_2(X) + c_3 P_3(X) + \cdots \\ &= c_0 \cdot 1 + c_1 \cdot X + c_2 \cdot \frac{1}{2}(3X^2 - 1) + c_3 \cdot \frac{1}{2}(5X^3 - 3X) + \cdots \end{aligned}$$

To compute c_n, we can follow the formula:

$$c_n = \frac{\mathbb{E}[Y \Psi_i(\xi)]}{\mathbb{E}[\Psi_i^2(\xi)]} = \frac{1}{h_n} \cdot \frac{1}{2} \int_{-1}^{1} \sin(\pi x) P_n(x) dx,$$

where

$$h_n = \mathbb{E}[P_n^2] = \int_{-1}^{1} P_n(x)p_X(x)dx = \int_{-1}^{1} P_n^2(x)\frac{1}{2}dx = \frac{1}{2}\cdot\frac{2}{2n+1} = \frac{1}{2n+1}.$$

We know that a uniform distribution $U[a,b]$ has PDF $p_X(x) = \frac{1}{b-a}$. Thus

$$c_n = \frac{2n+1}{2}\int_{-1}^{1}\sin(\pi x)P_n(x)dx.$$

Now we can easily compute the coefficients!

$$\begin{aligned}
c_0 &= \int_{-1}^{1}\sin(\pi x)\cdot 1dx = 0,\\
c_1 &= \frac{3}{2}\int_{-1}^{1}\sin(\pi x)\cdot x dx = \frac{3}{\pi},\\
c_2 &= \frac{5}{2}\int_{-1}^{1}\sin(\pi x)\cdot\frac{1}{2}(3x^2-1)dx = 0,\\
c_3 &= \frac{7}{2}\int_{-1}^{1}\sin(\pi x)\cdot\frac{1}{2}(5x^3-3x)dx = \frac{7}{\pi}-\frac{105}{\pi^3}.
\end{aligned}$$

Therefore our truncated expansion can be written as the following:

$$\boxed{Y \approx \frac{3}{\pi}P_1(X) + \left(\frac{7}{\pi}-\frac{105}{\pi^3}\right)P_3(X) + \cdots = \frac{3}{\pi}x + \left(\frac{7}{\pi}-\frac{105}{\pi^3}\right)\cdot\frac{1}{2}(5X^3-3X) + \cdots}$$

(b) We've been asked to find $\mathbb{E}[Y]$ and $\mathrm{Var}(Y)$. To sanity check, lets first find these values using just probability theory so that we may check our answers later.

$$\mathbb{E}[Y] = \mathbb{E}[\sin(\pi X)] = \int_{\Omega}\sin(\pi x)\,p_X(x)\,dx = \int_{-1}^{1}\sin(\pi x)\cdot\frac{1}{2}\,dx = 0.$$

$$\mathrm{Var}(Y) = \mathbb{E}[\sin^2(\pi X)] - \mathbb{E}[\sin(\pi X)]^2 = \mathbb{E}[\sin^2(\pi X)] = \int_{-1}^{1}\sin^2(\pi x)\cdot\frac{1}{2}\,dx = \frac{1}{2}.$$

Now lets try to verify this with PCE. To find the mean, we note that:

$$\boxed{\mathbb{E}[Y] = c_0 = 0.}$$

To find the variance, we use the formula from before:

$$\begin{aligned}\mathrm{Var}(Y) &= \sum_{n=1}^{K} c_n^2 \, \mathbb{E}[P_n^2] \\ &= c_0^2\mathbb{E}[P_0^2] + c_1^2\mathbb{E}[P_1^2] + c_2^2\mathbb{E}[P_2^2] + c_3^2\mathbb{E}[P_3^2] + \cdots\end{aligned}$$

Since only the odd terms contribute,

$$\mathrm{Var}(Y) = c_1^2 h_1 + c_3^2 h_3 + \cdots = \left(\frac{3}{\pi}\right)^2 \cdot \frac{1}{3} + \left(\frac{7}{\pi} - \frac{105}{\pi^3}\right)^2 \cdot \frac{1}{7} + \cdots \approx \boxed{0.49561.}$$

This is obviously not *exactly* equal to the variance we found earlier (which was 0.5), but we can code up a Python simulation to see that it indeed converges.

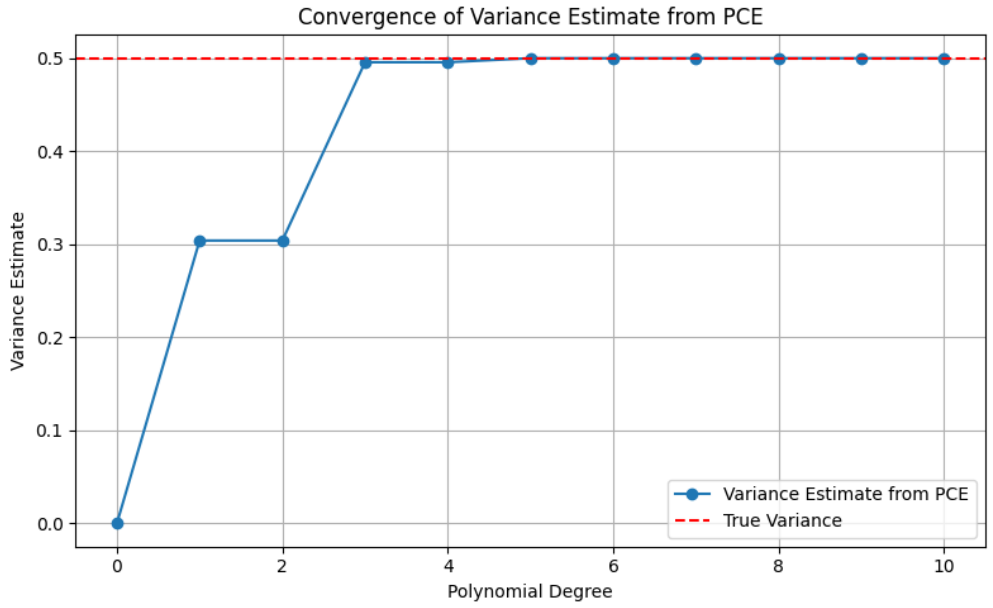

Figure 1: Plot displaying the true and approximated solution of variance using PCE.

Problem 2

Consider the stochastic ODE:

$$\frac{du(t,\xi)}{dt} = \lambda(\xi)u(t,\xi) \quad u(0,\xi) = u_0(\xi)$$

where $\lambda(\xi) = \lambda_0 + \lambda_1\xi$ with $\xi \sim \mathcal{N}(0,1)$, and λ_0 and λ_1 are constants.

(a) Find the solution of this ODE analytically and expand it.

(b) Expand the ODE directly.

(c) Comment on the two approaches: are they equivalent after truncation?

(d) Determine the expectations and variances, and comment on their values with respect to the sign of λ_0 and λ_1.

Solution. (a) Because ξ is Gaussian, we will make use of the Hermite polynomials in this problem. For each realization ξ, our solution will be

$$u(t,\xi) = u_0 e^{\lambda(\xi)t} = u_0 e^{(\lambda_0+\lambda_1\xi)t}.$$

We can let $u(t,\xi) = c_n \sum_{n=0}^{\infty} \mathrm{He}_n(\xi)$, where He_n are the Hermite polyonamils. Their coefficients can be found as follows

$$\begin{aligned} c_n(t) &= \frac{1}{n!}\mathbb{E}[(u(t,\xi)\mathrm{He}_n(\xi))] && (1)\\ &= \frac{u_0}{n!}\mathbb{E}[e^{(\lambda_0+\lambda_1\xi)t}\mathrm{He}_n(\xi)] && (2)\\ &= \frac{u_0}{n!}e^{\lambda_0 t}(\lambda_1 t)^n e^{(\lambda_1 t)^2/2}] && (3)\\ &= u_0 \exp\left(\lambda_0 t + \frac{(\lambda_1 t)^2}{2}\right)\cdot\frac{(\lambda_1 t)^n}{n!}. && (4) \end{aligned}$$

In simplifying from the Eqn. 2 to 3 above, we've used the fact that for $\xi \sim \mathcal{N}(0,1)$, and therefore the moment-generating function is given by

$$\mathbb{E}[e^{a\xi}\mathrm{He}_n(\xi)] = a^n e^{a^2/2}.$$

For our case, we've taken $a = (\lambda_0 + \lambda_1)t$. Then, the truncated expansion of order 2 is given by

$$\begin{aligned} u(t,\xi) &\approx c_0(t) + c_1(t)\mathrm{He}_1(\xi) + c_2(t)\mathrm{He}_2(\xi) && (5)\\ &= u_0 \exp\left(\lambda_0 t + \frac{\lambda_1^2 t^2}{2}\right)\left[1 + (\lambda_1 t)\xi + \frac{(\lambda_1 t)^2}{2}(\xi^2 - 1)\right]. && (6) \end{aligned}$$

Our final solution of the truncated solution is then

$$\boxed{u(t,\xi) \approx u_0 \exp\left(\lambda_0 t + \frac{\lambda_1^2 t^2}{2}\right)\left[1 + (\lambda_1 t)\xi + \frac{(\lambda_1 t)^2}{2}(\xi^2 - 1)\right].}$$

(b) For the Galerkin method, let's try to solve the system for $N = 2$ polynomial terms. We assume the expansion $u(t,\xi) = c_0(t) + c_1(t)\mathrm{He}_1(\xi) + c_2(t)\mathrm{He}_2(\xi)$. With the recurrence relation identities, note that

$$\xi\mathrm{He}_0 = \mathrm{He}_1, \quad \xi\mathrm{He}_1 = \mathrm{He}_2 + \mathrm{He}_0, \quad \xi\mathrm{He}_2 = \mathrm{He}_3 + 2\mathrm{He}_1.$$

We plug into $\frac{du(t,\xi)}{dt} = (\lambda_0 + \lambda_1\xi)u(t,\xi)$. The left-hand side is then

$$\frac{du(t,\xi)}{dt} = \dot{c_0}\mathrm{He}_0 + \dot{c_1}\mathrm{He}_1 + \dot{c_2}\mathrm{He}_2,$$

and the right-hand side is given by:

$$(\lambda_0 + \lambda_1\xi)u = (\lambda_0 c_0 + \lambda_1 c_1)\mathrm{He}_0 + (\lambda_0 c_1 + \lambda_1 c_0 + 2\lambda_1 c_2)\mathrm{He}_1 + (\lambda_0 c_2 + \lambda_1 c_1)\mathrm{He}_2 + (\lambda_1 c_2)\mathrm{He}_3.$$

Now, we can match the coefficients of He_0, He_1, He_2 and drop the He_3 term (for the purposes of our truncation) to get the following:

$$\boxed{\begin{aligned} \dot{c_0} &= \lambda_0 c_0 + \lambda_1 c_1, \\ \dot{c_1} &= \lambda_0 c_1 + \lambda_1 c_0 + 2\lambda_1 c_2, \\ \dot{c_2} &= \lambda_0 c_2 + \lambda_1 c_1 \end{aligned}} \qquad \boxed{c_0(0) = u_0, \quad c_1(0) = 0, \quad c_2(0) = 0}$$

In matrix form, this becomes

$$\begin{bmatrix} \dot{c_0} \\ \dot{c_1} \\ \dot{c_2} \end{bmatrix} = \begin{bmatrix} \lambda_0 & \lambda_1 & 0 \\ \lambda_1 & \lambda_0 & 2\lambda_1 \\ 0 & \lambda_1 & \lambda_0 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}.$$

Solving this 3×3 linear ODE system gives $c_0(t), c_1(t), c_2(t)$; the approximation is

$$u(t,\xi) \approx c_0(t) + c_1(t)\mathrm{He}_1(\xi) + c_2(t)\mathrm{He}_2(\xi).$$

To do this explicitly, we write in matrix form, $\dot{\mathbf{c}} = A\mathbf{c}$ with

$$A = \begin{bmatrix} \lambda_0 & \lambda_1 & 0 \\ \lambda_1 & \lambda_0 & 2\lambda_1 \\ 0 & \lambda_1 & \lambda_0 \end{bmatrix}.$$

We write $A = \lambda_0 I + \lambda_1 B$ with

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}.$$

We check that $B^3 = 3B$ (hence the polynomial is $x(x^2 - 3)$). Therefore

$$e^{tA} = e^{\lambda_0 t} e^{\lambda_1 t B} \left(I + \alpha(t)B + \beta(t)B^2\right),$$

with

$$\alpha(t) = \frac{\sinh(\sqrt{3}\lambda_1 t)}{\sqrt{3}}, \quad \beta(t) = \frac{\cosh(\sqrt{3}\lambda_1 t) - 1}{3}.$$

Applying this to the initial vector $\mathbf{c}(0) = (u_0, 0, 0)^\top$ and using $B(1,0,0)^\top = (0,1,0)^\top$, $B^2(1,0,0)^\top = (1,0,1)^\top$, gives explicit coefficients:

$$\boxed{\begin{aligned} c_0(t) &= u_0 e^{\lambda_0 t}\frac{2+\cosh(\sqrt{3}\lambda_1 t)}{3},\\ c_1(t) &= u_0 e^{-\lambda_0 t}\frac{\sinh(\sqrt{3}\lambda_1 t)}{\sqrt{3}},\\ c_2(t) &= u_0 e^{\lambda_0 t}\frac{\cosh(\sqrt{3}\lambda_1 t)-1}{3}. \end{aligned}}$$

Thus the 3-term approximation is

$$u(t,\xi) \approx c_0(t) + c_1(t)\mathrm{He}_1(\xi) + c_2(t)\mathrm{He}_2(\xi),$$

with c_0, c_1, c_2 as above. As a sanity check: at $t = 0, \cosh(0) = 1, \sinh(0) = 0$, so $c_0(0) = u_0$, $c_1(0) = c_2(0) = 0$.)

To quickly check the two-term case, if we kept only $\mathrm{He}_0, \mathrm{He}_1$, the system $\dot{c_0} = \lambda_0 c_0 + \lambda_1 c_1$, $\dot{c_1} = \lambda_0 c_1 + \lambda_1 c_0$ has a closed form

$$c_0(t) = u_0 e^{\lambda_0 t}\cosh(\lambda_1 t), \quad c_1(t) = u_0 e^{\lambda_0 t}\sinh(\lambda_1 t).$$

(c) Are the two approaches equivalent?

Asymptotically, yes; both solutions from pt.(a) and (b) give the same infinite series. However, for finite orders, the projection method truncates the exponential series while the Galerkin method truncates the polynomial chaos term. This is only true, however, due to the fact that our ODE is linear.

(d) For the mean, we know that $\mathbb{E}[u(t,\xi)] = c_0(t)$. Using the analytical solution, we get

$$\mathbb{E}[u(t,\xi)] = u_0\mathbb{E}[e^{(\lambda_0+\lambda_1\xi)t}] = u_e^{\lambda_0 t}\mathbb{E}[e^{\lambda_1 t\xi}] = u_0 e^{\lambda_0 t}e^{(\lambda_1 t)^2/2},$$

where we used the fact that $\mathbb{E}[e^{a\xi}] = e^{a^2/2}$ for $\xi \sim \mathcal{N}(0,1)$. Thus

$$\boxed{\mathbb{E}[u,(t,\xi)] = u_0 \exp\left(\lambda_0 t + \frac{\lambda_1^2 t^2}{2}\right).}$$

For the variance, using the fact that $\mathrm{Var}(u(t,\xi)) = \sum_{n=1}^{\infty} c_n^2 n!$, we get

$$\begin{aligned} \mathrm{Var}(u(t,\xi)) &= \mathbb{E}[u^2] - \mathbb{E}[u]^2 \\ &= u_0^2 e^{2\lambda_0 t}\mathbb{E}[e^{2\lambda_1 t\xi}] - u_0^2 e^{2\lambda_0 t + \frac{\lambda_1^2 t^2}{2}} \\ &= u_0^2 e^{2\lambda_0 t + 2\lambda_1^2 t^2} - u_0^2 e^{2\lambda_0 t + \lambda_1^2 t^2}. \end{aligned}$$

Thus

$$\boxed{\mathrm{Var}(u(t,\xi)) = u_0^2 e^{2\lambda_0 t + \lambda_1^2 t^2}(e^{\lambda_1^2 t^2} - 1).}$$

Regarding the effect of signs on λ_0 and λ_1 with respect to the mean and variance, we observe:

$\Rightarrow$ when $\lambda_0 > 0$: growth on average; both mean and variance grow exponentially

$\Rightarrow$ when $\lambda_0 < 0$: decay on average but behavior depends on λ_1

- if $|\lambda_0| > \lambda_1^2/t$: mean decays (stable on average)

- if $|\lambda_0| < \lambda_1^2 t/2$: mean can still grow (variance effect dominates)

$\Rightarrow$ when $\lambda_1 = 0$: deterministic case where $\text{Var}(u) = 0$

$\Rightarrow$ because λ_1 only appears as λ_1^2 in mean and variance, the sign of it has no impact

$\Rightarrow$ with high $|\lambda_1|$: there is high uncertainty; variance term $e^{\lambda_1^2 t^2}$ dominates

4 Conceptual Understanding

The coefficients of our polynomials are selected such that their expansion would be the orthogonal projection of u on the span of these basis functions. The choice of the polynomial family is particularly important; they MUST match the input distribution in order for things to converge nicely. Different distributions implies different weights which implies different orthogonal polynomial families. If misaligned, the coefficients would no longer be minimizing the mean square error. It is precisely because of this property of orthogonality that computing moments is nice with PCE.

By formulation, PCE minimizes the root mean square error (RMSE). Mathematically, this is called the L^2 norm and it represents the energy of a random variable (although more commonly called the variance). By minimizing the L^2 error, we are finding the best approximation of the original random variable on average. Physically, this represents the average energy difference between the true and approximated solution. The con to this is that while your solution is meant to have low error on average across various random states, you are not optimizing for any one state in particular.

If we expand this to an application example in climate, we could take the following hypothetical: given a task to model temperature fields across a region, we may (very likely) have variables for which have high uncertainty due to partial information. This could be variables like humidity, surface albedo, etc. If we treat each of these uncertainties a s a random variable, then we'd be able to build a PCE for the temperature field. I'm honestly not sure if this would work, but perhaps this means we wouldn't need explicit detailed physics for our field. The PCE would act similar to a surrogate model and provide us with a way to quantify the uncertainty in our input variables and how those impact temperature behavior.

References

[1] George Arfken. *Mathematical Methods for Physicists*. Academic Press, Inc., San Diego, third edition, 1985.

[2] Matthew Baring. Lecture notes for phys 516,. URL https://www.ruf.rice.edu/~baring/phys516/phys516_2021_lec_041321.pdf.

[3] Johnson C. M. Fok, Benyu Guo, and Tao Tang. Combined hermite spectral-finite difference method for the fokker-planck equation. *Math. Comput.*, 71(240):1497–1528, October 2002. ISSN 0025-5718. doi: 10.1090/S0025-5718-01-01365-5. URL https://doi.org/10.1090/S0025-5718-01-01365-5.

[4] Manoochehr Khasi and Fereshteh Samadi. Numerical solutions of fokker–planck equation based on hermite functions. *Communications in Combinatorics, Cryptography & Computer Science*, 2022(1):1–6, 2022.

[5] Francesco Russo and Pierre Vallois. *Hermite Polynomials and Wiener Chaos Expansion*, pages 309–332. Springer International Publishing, Cham, 2022. ISBN 978-3-031-09446-0. doi: 10.1007/978-3-031-09446-0_9. URL https://doi.org/10.1007/978-3-031-09446-0_9.

[6] J.J. Westerink. Ce 60130 finite element methods lecture notes. URL https://coast.nd.edu/jjwteach/www/www/60130/New%20Lecture%20Notes_PDF/CE60130_Lecture%208%20with_footer-v04.pdf.