Solving the steady-state diffusion equation with uncertainty Midyear report

Virginia Forstall
vhfors@gmail.com

Advisor: Howard Elman
elman@cs.umd.edu
Department of Computer Science

December 16, 2011

Abstract

The goal of this project is to efficiently solve a steady-state diffusion equation with a random diffusion coefficient. Although, such equations can be solved using Monte-Carlo methods, the lengthy computation time can be constraining. Using a Karhunen-Loéve expansion allows the random coefficient to be approximated with a finite sum of random variables. This expansion combined with a Galerkin method or stochastic collocation method should reduce computation time.

1 Introduction

The goal of this project is to explore computational algorithms for the steady-state diffusion equation with a random field as the diffusion coefficient. The formal problem is

$$-\nabla \cdot (k(x,\omega)\nabla u) = f(x)\ . \tag{1}$$

The boundary conditions are chosen to be deterministic. Specifically let $u(x,\omega) = g(x)$ on ∂D_D and $\frac{\partial u}{\partial n} = 0$ on ∂D_n. The diffusion coefficient, $k(x,\omega)$, is a random field of the form, $k = e^{a(x,\omega)}$ where $a(x,\omega)$ is also a random field. This assumption ensures k is positive for all x and thus guarantees existence and uniqueness of the solution, u. Assume the spatial domain $D \subset \mathbb{R}^2$ is bounded.

The solution, $u(x,\omega)$ is a function of the sample space. It will be the goal of this project to calculate the moments of this solution. Another meaningful result would be the cumulative distribution function of the solution.

Applications of this include modeling groundwater flow through a porous medium. To do this the permeability of the medium is needed at every location in the spatial domain. However, knowing this at every point is infeasible. Instead, the permeability can be treated as a random variable at each point in the domain. Since the permeability at one point is related to the points around it, these points will be correlated. Thus the model of a random field is reasonable.

One method for solving these problems is to write $\log(k(x,\omega)) = a(x,\omega)$ as an infinite series expansion of random variables [7]. This expansion, known as the Karhunen-Loéve expansion, can be truncated to provide an approximation of this field and has a linear dependence on the random variables. In this project, the approximation will be found by taking the Karhunen-Loéve expansion of $k(x,\omega)$ instead.

In section 2 the algorithm is outlined. Section 3 discusses some preliminary results of these methods. Section 4 discuss the Monte-Carlo method, which will be used to verify the results. The status of the project is addressed in Section 5 and Section 6 contains and updated schedule.

2 Approach

The algorithm consists of two key parts. The first is to numerically approximate the random field. This will be done using a Karhunen-Loéve expansion which is derived from the covariance matrix of the random field. The eigenpairs of the covariance are used to define the random field and this section discusses methods for obtaining the eigenparis as well descriptions of the random field to be used. The second part of the algorithm is a method for approximate the solution of the partial differential equation.

2.1 Approximating the random field

A random field can be written as an infinite series known as the Karhunen-Loéve expansion. The expansion,

$$a(x,\xi) = \mu(x) + \sum_{s=1}^{\infty} \sqrt{\lambda_s} f_s(x)\xi_s \ , \tag{2}$$

is written in terms of the eigenvalues, λ_s and eigenfunctions, $f_s(x)$, of the covariance matrix of this field and uncorrelated random variables ξ_s. Keeping a finite number of terms with the largest magnitude eigenvalues produces an approximation of the random field.

The eigenpairs of the random field satisfy,

$$(\mathcal{C}f)(x) = \int_D C(x,y)f(y)dy = \lambda f(x) \ , \tag{3}$$

where $C(x,y)$ is the covariance function of the random field,

$$C(x,y) = \int_\Omega (a(x,\omega) - \mu(x))(a(y,\omega) - \mu(y))dP(\omega) \ , \tag{4}$$

where

$$\mu(x) = \int_\Omega a(x,\omega)dP(\omega) \ . \tag{5}$$

When the spatial domain D is discretized, eigenpairs of the operator can be approximated using the covariance matrix. Let C be the covariance matrix such that

$$C_{ij} = C(x_i, x_j) \ . \tag{6}$$

Following the discussion of [1], we determine the relationship between the eigenvalues and eigenfunctions of the operator in Equation 3 and the eigenvalues and eigenvectors of the covariance matrix in Equation 6.

Consider a discretization in space which depends on a parameter h. Denote the basis functions of the space as $h_j(x)$. The eigenfunctions, $f_n(x)$, can be written as a linear combination of the basis functions,

$$f_n(x) = \sum_j s_j^n h_j(x) \ . \tag{7}$$

Requiring that error terms are normal to the space, gives rise to a discrete approximation to the eignevalue problem in Equation 3,

$$GS = \Lambda MS \tag{8}$$

where

$$G_{ij} = \int_D \int_D C(x_1, x_2) h_i(x_2) dx_2 h_j(x_1) dx_1 \tag{9}$$

$$S_{ij} = s_i^j \tag{10}$$

$$\Lambda_{ij} = \delta_{ij} \lambda_i \tag{11}$$

$$M_{ij} = \int_D h_j(x_1) h_i(x_1) dx_1 \ . \tag{12}$$

The matrix G is defined in terms of the covariance function $C(x_1, x_2)$. The eigenvalues (the diagonal of Λ) and the eigenvectors (the columns of S) provide the information needed to determine the KL expansion. Thus for known covariance functions (i.e. where G can be easily computed), we have a method for approximating random field.

Consider a one-dimensional domain with uniform intervals. Let $D = [a, b]$ and h be the width of the uniform intervals. The usual hat functions serve as a basis. Evaluating the integrals on the discrete domain using the Riemann midpoint rule,

$$G_{ij} = h^2 C(x_i, x_j) = h^2 C_{ij} \tag{13}$$

$$M_{ij} = h\delta_{ij} \ . \tag{14}$$

The eigenvalue problem becomes

$$hCS = \Lambda S \ . \tag{15}$$

For the two dimensional domain with uniform intervals, a similar computation can be performed and the eigenvalues of the operator are the eigenvalues of $h_x h_y C$ where h_x is the width of the intervals in the x direction and h_y is the width of the intervals in the y direction. More complicated discretizations can be considered.

Thus, it will be possible to approximate the eigenpairs needed in the KL expansion by solving the eigenvalue problem for the covariance matrix C. There are two methods we consider for determining this matrix. The first is used when an analytic expression for the covariance function is known; the other method takes samples of the random field to generate the sample covariance matrix.

2.1.1 Method 1: Evaluate the covariance function

The first method can be used when the analytic expression for the covariance is known. This function is evaluated at each pair of points on the spatial domain to form the matrix in Equation 6. It is computationally easier and is definitely preferred. However, an analytic expression for the covariance function is not always known so a method using samples of the random field would be necessary.

2.1.2 Method 2: Form the sample covariance matrix

If the covariance function $C(x_1, x_2)$ is unknown for a random field and the random field $a(x, \xi)$ can be sampled, the eigenvalues and eigenvectors can be found by constructing the sample covariance matrix. This is defined as

$$\widehat{C}_{ij} = \frac{1}{n}\sum_{k=1}^{n}(a(x_i, \omega_k) - \hat{\mu}_i)(a(x_j, \omega_k) - \hat{\mu}_j) \tag{16}$$

$$\hat{\mu}_i = \frac{1}{n}\sum_{k=1}^{n} a(x_i, \omega_k) \tag{17}$$

where n is the number of samples and $\hat{\mu}$ is the sample mean. This matrix $\widehat{C}$ is the sampled is known as the sample covariance matrix.

To do this we do not necessarily need the full matrix $\widehat{C}$. Define a matrix,

$$B_{ik} = a(x_i, \omega_k) - \hat{\mu}_i \; . \tag{18}$$

Then the sample covariance matrix can be written as

$$\widehat{C} = \frac{1}{n} BB^T \; . \tag{19}$$

Therefore, the eigenvalues of $\widehat{C}$ can be computed without generating the full sample covariance matrix. Take the singular value decomposition, $B = U\Sigma V^T$. The eigenvalues of $\widehat{C}$ are $\frac{1}{n}\Sigma^2$ and the eigenvectors are the columns of U. The SVD approach is preferred when there are a large number of samples n taken. This will also ensure that the computed eigenvalues are real.

2.1.3 Gaussian random field

A Gaussian random field in one dimension has covariance function

$$C(x_1, x_2) = \sigma^2 \exp(-|x_1 - x_2|/b) \tag{20}$$

where σ^2 is the variance of the stationary random field and b is the correlation length. Note that large values of b indicate that random variables at points that are near each other are highly correlated. For this covariance function, the analytic expressions for the eigenvalues and eigenfunctions can be found by solving Equation 3 analytically. The eigenvalues for a one-dimensional domain $D = [-l, l]$, from [1] are

$$\lambda_n = \sigma^2 \frac{2b}{\omega_n^2 + b^2} \tag{21}$$

$$\lambda_n^* = \sigma^2 \frac{2b}{\omega_n^{*2} + b^2} \tag{22}$$

where ω_n and ω_n^* solve the following:

$$b - \omega \tan(\omega l) = 0 \tag{23}$$

$$\omega^* + b \tan(\omega^* l) = 0 \; . \tag{24}$$

The eigenfunctions are

$$f_n(x) = \frac{\cos(\omega_n x)}{\sqrt{l+\frac{\sin(2\omega_n l)}{2\omega_n}}} \tag{25}$$

$$f_n^*(x) = \frac{\sin(\omega_n^* x)}{\sqrt{l-\frac{\sin(2\omega_n^* l)}{2\omega_n^*}}} \ . \tag{26}$$

Solving Equations 23 and 24 with Newton's method gives the numerical values for the eigenvalues of the random field. The KL expansion of a Gaussian random field a

$$a(x,\vec{\xi}) = \mu(x) + \sum_{n=1}^{\infty} \sqrt{\lambda_n} f_n(x)\xi_n \tag{27}$$

has Gaussian random variables $\xi_n \sim N(0,1)$.

The Gaussian field outline above is one-dimensional. A two-dimensional Gaussian covariance function

$$C((x_1,y_1),(x_2,y_2)) = \sigma^2 \exp\left[\frac{-|x_1-x_2|}{b_1} + \frac{-|y_1-y_2|}{b_2}\right] \tag{28}$$

can be separated into an x-component and y-component. This allows the eigenvalues for the two-dimensional problem to be calculated using the one-dimensional expressions.

2.1.4 Lognormal random field

If $a(x,\xi)$ is a Gaussian random variable at every point in the spatial domain, $c(x,\xi) = \exp(a(x,\xi))$ is lognormal at every point in the spatial domain. The mean, variance, and covariance function for a lognormal random variable, Y, are

$$E[Y] = e^{\sigma^2/2} \tag{29}$$

$$Var[Y] = e^{2\mu+\sigma^2}(e^{\sigma^2}-1) \tag{30}$$

$$LC(x_1,x_2) = e^{2\mu+\sigma^2}(e^{C(x_1,x_2)}-1) \ , \tag{31}$$

where $\ln(Y)$ is a normal random variable with mean μ, standard deviation σ and covariance function, $C(x_1,x_2)$ [5].

2.2 Method for solution of the PDE

The primary portion of the project for next semester will be to write one solution method for the stochastic PDE. There are two possibilities which will be investigated for solving the steady-state diffusion equation: the stochastic collocation method and the stochastic Galerkin method.

2.2.1 Stochastic Galerkin method

First the stochastic weak formulation of the problem is needed. The function space, $L^2(D)$ is defined as

$$L^2(D) = \{v : D \to \mathbb{R} | \int_D v(x)^2 dx < \infty\} \ . \tag{32}$$

Considering $H^1(D) = \{v(x,y) : D \to \mathbb{R} | v, \partial v/\partial x, \partial v/\partial y \in L^2(D)\}$, define the following subspaces:

$$H_0^1(D) = \{u \in H^1(D) : u = 0 \text{ on } \partial D_d\} \tag{33}$$

$$H_g^1(D) = \{u \in H^1(D) : u = g(x) \text{ on } \partial D_d\} \, . \tag{34}$$

Note that for the probability space, Ω, $L^2(\Omega)$ is the set of random variables which have finite variance.The weak formulation is to find $u \in H_g^1(D) \times L^2(\Omega)$ such that

$$b(u,v) = l(v), \quad \forall v \in H_0^1(D) \times L^2(\Omega) \tag{35}$$

where

$$b(u,v) = \int_\Omega \int_D c(x,\omega) \nabla u(x,\omega) \cdot \nabla v(x,\omega) dx \, dP(\omega) \tag{36}$$

$$l(v) = \int_\Omega \int_D f(x) v(x,\omega) dx \, dP(\omega) \, . \tag{37}$$

The stochastic Galerkin method finds the stochastic discretization using a basis formed with polynomials of the M random variables kept in the KL expansion. Increasing the degree of the polynomials improves the approximation. This produces a larger matrix that can be solved using the deterministic Galerkin finite element method. Even though the matrix is larger, computation time can still be managed by taking advantage of structure and/or sparsity. For examples, see [4],[7].

2.2.2 Stochastic Collocation method

The random field has been approximated by a finite number of random variables using the KL expansion. Therefore, the solution, $u(x,\omega)$ can be written as $u(x,\vec{\xi})$ where $\vec{\xi}$ is a vector of length M of the uncorrelated random variables. Define Γ to be the image of the random varialbes. The probability space can be written as $L^2(\Gamma) = \{X : \Gamma \to \mathbb{R} | E[X^2] < \infty\}$.

A set of points, $Z_p \subset \Gamma$, are known as collocation points, where each point is a vector of length M. The number of these points in this set, N_c, is a parameter of the method. It is the maximum degree of the polynomials that form the basis of the finite basis of $L^2(\Gamma)$. A greater number of points increases the accuracy of the results, but naturally adds computation time. Keeping this number significantly below the number of realizations in the Monte-Carlo method will produce savings in computation time.

At each collocation point, there is a single value for the field, k. Denote this $k_i^{(M)}$. At each point a deterministic weak formulation holds. That is find $u \in H_g^1$ s.t.

$$\int_D k_i^{(M)}(x) \nabla u_i(x) \cdot \nabla v_i(x) dx = \int_D f(x) v_i(x) dx \;\; \forall v_i \in H_0^1(D) \, . \tag{38}$$

Thus there are N_c problems that can each be solved deterministically. The last step is to use Lagrange interpolation to find an approximation of u at points in the sample space that are not in the set of collocation points.

3 Results

The implementation of this project will be using Matlab and a desktop computer with 1.9 GB RAM. The code makes use of functions from E. Ullman.

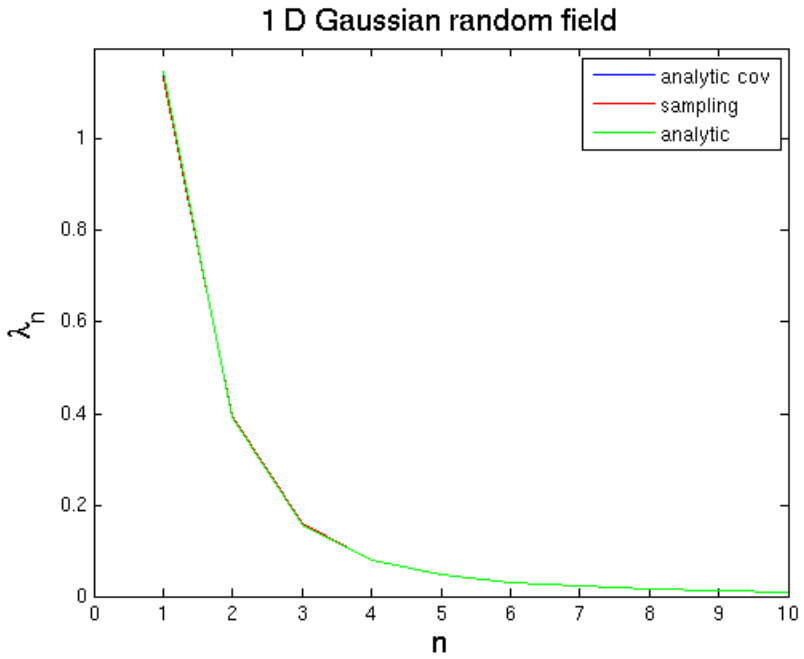


Figure 1: Eigenvalues of Gaussian random field with parameters $b = 1$, $n = 10000$ computed using three different methods. Note that methods 1 and 2 produce nearly identical results.

3.1 Eigenpairs for Gaussian random field $a(x, \xi)$

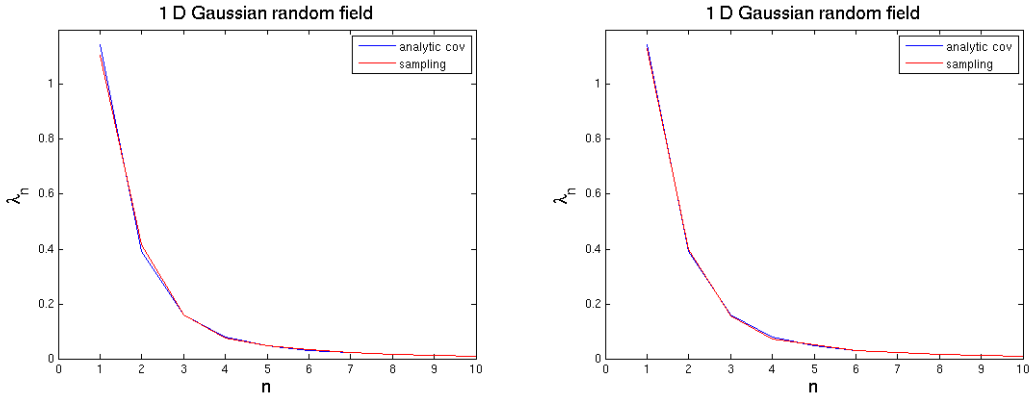
There are three ways in which the eigenvalues of $a(x, \xi)$ for a Gaussian random field with covariance function, $C(x_1, x_2) = \sigma^2 \exp(-|x_1 - x_2|/b)$ were found. Method 1 solves Equations 23 and 24 for the eigenfrequencies using Newton's method. Method 2 uses the analytic expression for the covariance to build the covariance matrix. Method 3 builds the sample covariance matrix.

To compare Method 3 to the others, samples of $a(x, \xi)$ are needed. Consider the KL expansion of $a(x, \xi)$ as defined by Equation 27. Naturally we cannot compute and infinite expansion, so we introduce a parameter M_a. Then the first $M_a + 1$ terms are used with the analytic eigenpairs obtained in Method 1 to compute the KL expansion. This will serve as the samples $a(x_i, \xi_j)$ that are used to build the matrices outlined in Section 2. To maintain an accurate representation of the random field, a condition is introduced

$$\frac{\sum_{i=1}^{M_a} \lambda_i}{\sum_{i=1}^{\infty} \lambda_i} \geq .95 \ . \tag{39}$$

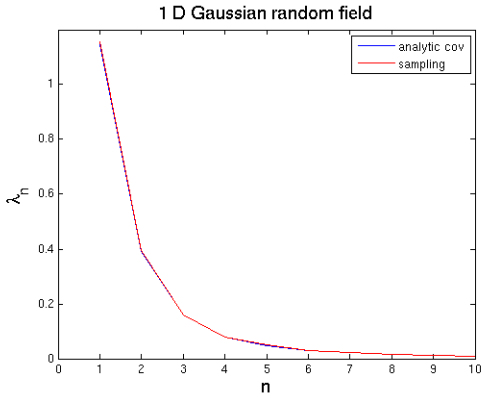
Numerical experiments used a one-dimensional domain $D = [-1, 1]$ with $n_x = 100$ evenly spaced intervals. Figure 1 shows the magnitude of the eigenvalues computed using the three methods sorted in order of decreasing magnitude. The results show that the values for Method 1 and 2 are nearly identical, while the Method 3 (the sampling method) is not as close. Figure 2 the sampling method is shown to converge as the number of samples, n increases.

Consider the effect of the correlation length b on the eigenvalues. Recall that a higher value of b indicates a stronger correlation between random variables that are close. For smaller values of b, Figure 3 shows the magnitude of the eigenvalues decreases. This makes it more difficut to acheive accuracy in the sampling method and requires more terms to satisfy the condition in Equation 39.



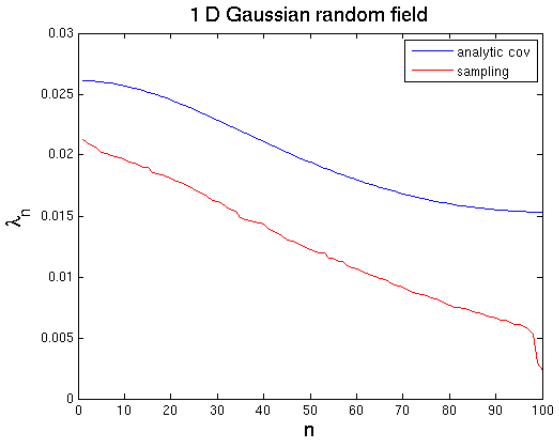
(a) n=100

(b) n=1000

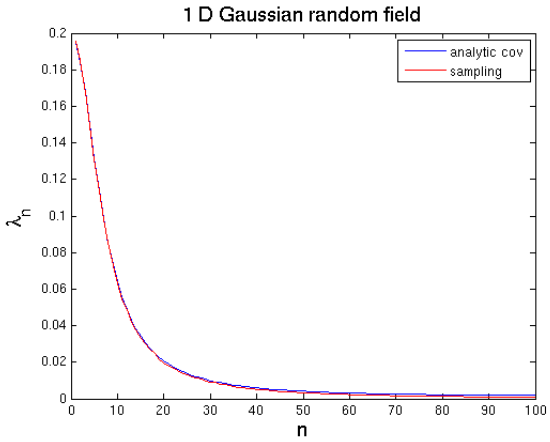


(c) n=10000

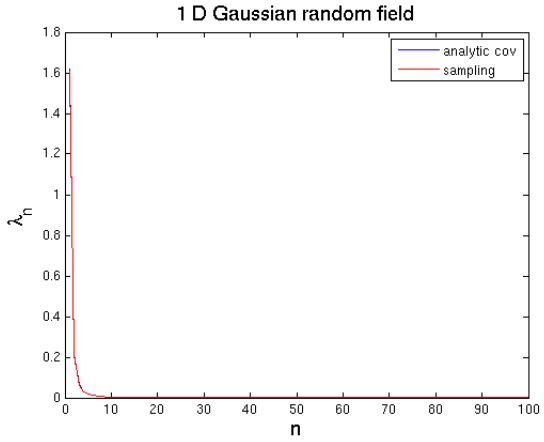
Figure 2: Eigenvalues of $a(x, \xi)$, a Gaussian random field. The sampling method converges to the eigenvalues obtained from analytic covariance function as the number of samples, n is increased. The number of terms kept in the KL expansion for Method 3 was $M_a = 100$ and the correlation length is $b = 1$.



(a) $b = .01, n = 10000, M = 100$



(b) $b = .1, n = 10000, M = 100$



(c) $b = 3, n = 10000, M = 100$

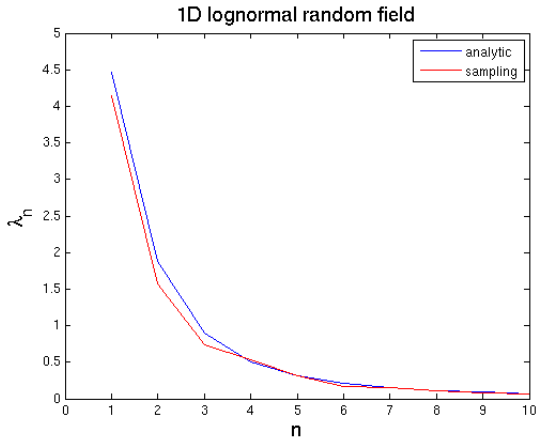
Figure 3: Eigenvalues of $a(x, \xi)$, a Gaussian random field. The effect of correlation length, b, on the eigenvalues obtained using Methods 2 and 3. The magnitude of the largest eigenvalue increases for larger values of b.

Since the two-dimensional Gaussian random field has a separable covariance function (Equation 28), the analytic expressions for the eigenvalues and eigenfunctions are also known. Methods 1-3 were verified for a two dimensional domain, $D = [0, 1] \times [0, 1]$ with 32 grid points in each direction. The eigenvectors of the three methods were also verified.

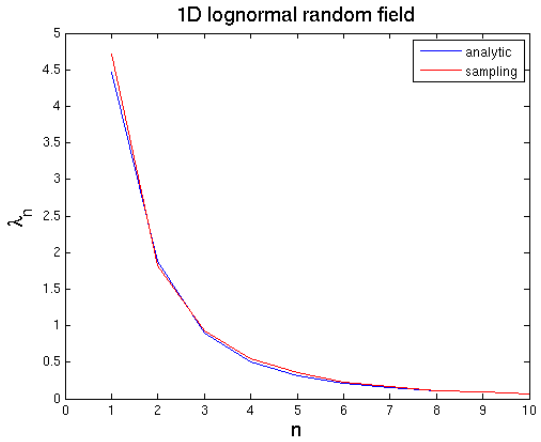
3.2 Eigenpairs for a lognormal random field, $k(x, \xi)$

Although analytic expressions for eigenpairs for a lognormal random field do no exist, Equation 31 is an analytic expression for the covariance matrix. Therefore Methods 2 and 3 can be compared. The sampling covariance matrix uses samples

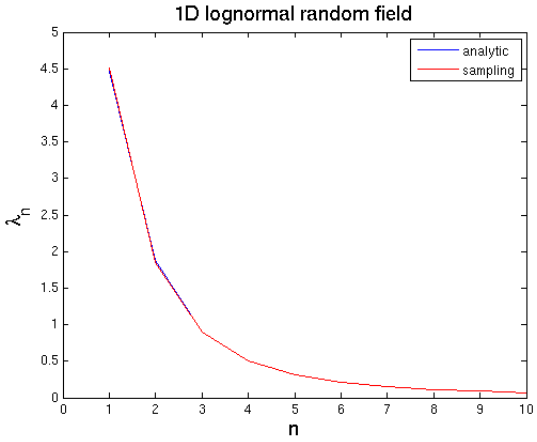
$$k(x_i, \xi_j) = \exp(\mu(x_i) + \sum_{n=1}^{M_a} \sqrt{\lambda_n} f_n(x_i) \xi_j, n) \tag{40}$$



(a) n=1000



(b) n=10000



(c) n=100000

Figure 4: Eigenvalues of $k(x, \xi)$, a lognormal random field. The eigenvalues obtained using the sample covariance matrix, converge to the analytic covariance matrix results as the number of samples is increased. Tests use correlation length $b = 1$.

Comparing these two methods, convergence is obtained (See Figure 3.2). Likewise the eigenvalues were verified for the two-dimensional domain.

4 Validation

Validation of the solution method will be performed using Monte-Carlo simulations. The coefficient, $k(x, \omega)$, will be sampled q times. The number of samples q will be greater than p in the stochastic collocation method. These samples k_i for $i = 1, .., q$ each produce a finite element problem requiring q solves of the discrete problem. Denote the solution of this problem be $u_h^i(x)$. The moments of these solutions can then be found. For example, the mean

$$E_q[u_h] = \frac{1}{q} \sum_{i=1}^{q} u_h^i(x) \ . \tag{41}$$

The error is $e = E[u] - E_q[u_h] = E[u] - E[u_h] + E[u_h] - E_q[u_h]$. The first difference, $E[u] - E[u_h]$ approaches 0 as $h \to 0$ assuming that finite basis of the Galerkin method follows the usual requirements. The second difference, $E[u_h] - E_q[u_h] \to 0$ as $q \to \infty$. We also know that $E[u]$ is the solution to the deterministic problem when $E[c]$ is used as the diffusion coefficient, thus we can compare these results. This approach will likely take many iterations. One example, in [2], took between $q = 10^4$ and $q = 10^6$ to converge.

5 Conclusion

The procedure for determining the eigenvalues from the sample covariance matrix has been verified. At the beginning of the project, this is the method that would be used to find the eigenvalues of the lognormal random field $k(x, \eta)$. Through the process of this verification an analytic expression for the covariance function was found in [5]. Therefore Method 2 will be used instead of the sampling method.

This means we have most of the components of the KL expansion

$$k(x, \eta) = k_0 + \sum_{n=1}^{M_c} \sqrt{\lambda_n} f_n(x) \eta_n \ . \tag{42}$$

The remaining of this approximation is the distribution of the random variables η_n.

A new verification step willl be introduced here which compares the moments of the random field produced with $k(x, \xi)$ as in the sampling method with the moments obtained using this expansion, $k(x, \eta)$.

6 Schedule

The project remains on schedule, but the components have been rearranged a bit. An adjusted schedule is below.

Stage 2: December

- Finish construction of the approximation
- Compare moments of random field and its approximation
- Write code which generates Monte-Carlo solutions

Stage 3: January- February

- Run the Monte-Carlo simulations
- Write solution method

Stage 4: March - April

- Run numerical method
- Analyze accuracy and validity of the method

7 References

References

[1] R. Ghanem, P. Spanos, Stochastic Finite Elements: A spectral approach, Dover Publications, Mineola, New York, 2003.

[2] A. Gordon, Solving stochastic elliptic partial differential equations via stochastic sampling methods , M.S Thesis, University of Manchester, 2008.

[3] C. Moler, Numerical Computing with Matlab, Chapter 10: Eigenvalues and Singular Values, 2004,`http://www.mathworks.com/moler/chapters.html`.

[4] C.E. Powell and H.E. Elman, Block-diagonal preconditioning for spectral stochastic finite-element systems, IMA Journal of Numerical Analysis, 29, (2009), 350-375.

[5] J. Rendu, Normal and lognormal estimation, Mathematical Geology, 11, 4, (1979), 407-422.

[6] C. Schwab and R. Todor, Karhunen-Loéve approximation of random fields by generalized fast multipole methods, Journal of Computational Physics, 217, (2006), 100-122.

[7] E. Ullmann, H. C. Elman, and O. G. Ernst, Efficient iterative solvers for stochastic Galerkin discretization of log-transformed random diffusion problems, UMIACS-TR-2011-04, 2011.

[8] X. Wan and G. Karniadakis, Solving elliptic problems with non-Gaussian spatially-dependent random coefficients, Computational Methods in Applied Mechanical Engineering, 198, (2009), 1985-1995.

[9] D. Xiu, Numerical Methods for Stochastic Computations, Princeton University Press, New Jersey, 2010.

[10] D. Xiu and J. Hesthaven, High-order collocation methods for differential equations with random inputs, SIAM Journal on Scientific Computing, 27, (2005), 1118-1139.