Solving the steady state diffusion equation with uncertainty

Final Presentation

Virginia Forstall
vhfors@gmail.com

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

May 6, 2012
Problem

The equation to be solved is

$$-\nabla \cdot (k(x, \omega) \nabla u(x, \omega)) = f(x),$$

where $k = e^{a(x, \omega)}$ is a lognormal random field.

- Assume a bounded spatial domain $D \subset \mathbb{R}^2$.
- The boundary conditions are deterministic.

$$u(x, \omega) = g(x) \text{ on } \partial D_D$$

$$\frac{\partial u}{\partial n} = 0 \text{ on } \partial D_n.$$

- Models groundwater flow through a porous medium [15].
Outline

1. Approximate the random field using the Karhunen-Loéve expansion.

2. Solve the PDE using stochastic Galerkin method.

3. Compare mean and variance of the solution to those obtained using the Monte-Carlo method.
Karhunen-Loéve expansion

The expansion is

\[ a(x, \xi) = a_0(x) + \sum_{s=1}^{\infty} \sqrt{\lambda_s} a_s(x) \xi_s. \] (2)

- \( a_0(x) \) is the mean of the random field.
- The random variables \( \xi_s \) are uncorrelated with \( E[\xi_s] = 0 \), \( \text{Var}[\xi_s] = 1 \).
- The \( \lambda_s \) and \( a_s(x) \) are eigenpairs which satisfy

\[ (C a_s)(x_1) = \int_D C(x_1, x_2) a_s(x_2) dx_2 = \lambda_s a_s(x_1), \] (3)

where \( C(x_1, x_2) \) is the covariance function of the random field.
Discretization of the eigenvalue problem

- The square domain $D$ is discretized intervals of equal size $h$ in each direction.
- The eigenvalues of the covariance operator satisfy
  \[ h^2 CV = \Lambda V \]  
  where $C_{ij} = C(x_i, x_j)$.
- The approximation of the eigenfunctions are
  \[ a_s(x_i) = \frac{1}{h} V_{is} \]
Discretization of the eigenvalue problem

- Alternatively, $q$ samples of the random field can be used to form the sample covariance matrix:

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

where $\hat{a}_i$ is the sample mean.
Covariance function

- The covariance function for the Gaussian random field with mean \( \mu \) and variance \( \sigma^2 \) is

\[
C_g((x_1, y_1), (x_2, y_2)) = \sigma^2 \exp \left( \frac{-|x_1 - x_2|}{b_x} + \frac{-|y_1 - y_2|}{b_y} \right)
\]  

(7)

where \( b_x, b_y \) are the correlation lengths.

- The eigenvalues and eigenfunctions have analytic expressions for this covariance function.
Validation of eigenpairs

- Two methods for finding eigenpairs verified by comparing with analytic expressions for the Gaussian random field.

**Figure:** Eigenvalues of Gaussian random field with parameters $b = 1$, $q = 10000$ computed using analytic expression and the two covariance matrices.
Covariance function

- The covariance function for the lognormal random field can be written as a function of the Gaussian covariance function.

\[ C_l((x_1, y_1), (x_2, y_2)) = e^{2\mu+\sigma^2} (e^{C_g((x_1, y_1),(x_2, y_2))} - 1) \] [9]. (8)

- This expression is used to build the covariance matrix and find the eigenpairs.

- The sampling method would allow this model to be used when structure of the random field is unknown but can be sampled at various points in space.
Karhunen-Loéve expansion

The lognormal random field can be approximated two ways:

\[
k(x, \xi) = \exp[a_0(x) + \sum_{i=1}^{m_g} \sqrt{\lambda_i} a_i(x) \xi_i]
\] (9)

\[
\hat{k}(x, \eta) = k_0(x) + \sum_{i=1}^{m_l} \sqrt{\mu_i} k_i(x) \eta_i.
\] (10)

- \{\xi_i\} are independent Gaussian random variables, so the joint probability density function, \(\rho(\xi)\), is known.
- The joint density function of the random variables, \(\eta_i\), is needed.
- Let \(m = \max(m_g, m_l)\) and the joint density function, \(\hat{\rho}(\eta)\), is found using a change of variables.
Probability density function

- Define matrices $A = [a_1|a_2|...|a_m]$ and $K = [k_1|k_2|...|k_m]$ where the columns are the eigenfunctions evaluated at the points in the spatial discretization.

- Define the mass matrix $B_{ij} = \int_D \phi_i(x)\phi_j(x)dx$. 

- $A$ was normalized so that $A^TBA = I$.

- Define the diagonal matrices $\Lambda = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_m)$ and $M = \text{diag}(\mu_1, \mu_2, ..., \mu_m)$.

- Define vectors $\xi = [\xi_1, \xi_2, ..., \xi_m]^T$ to be the standard normal random variables in the Gaussian random field and $\eta = [\eta_1, \eta_2, ..., \eta_m]^T$ to be the unknown random variables in the lognormal expansion.
Probability density function

\[ \Lambda A^T B(k(x, \xi) - a_0(x)) = \Lambda A^T B(\hat{k}(x, \eta) - a_0(x)) . \tag{11} \]

\[ \xi = g(\eta) = \Lambda A^T B(\ln(k_0 + KM\eta) - a_0) . \tag{12} \]

- \( \hat{\rho}(\eta) = \rho(g(\eta))|J(\eta)| \) describes the density for \( \eta \) such that \( k_0 + KM\eta > 0 \).
- \( |J(\eta)| \) is the absolute value of the determinant of the Jacobian, which we can find since \( g(\eta) \) is differentiable.
Probability density function

Figure: Probability density function for 1d field $m = 1, b = 10$. 
Probability density function

**Figure:** Probability density function for $m = 2, b = 10$. 
Probability density function

- The joint density function can be used to find the marginal density functions which have mean 0 and variance 1 as expected.
- Samples of $k(x, \xi)$ were generated with $m N(0,1)$ samples for each instance.
- To find samples of $\hat{k}(x, \eta)$, accept/reject sampling is used.
- Uniform samples over the support of density function are generated and the probability density function evaluated at those values.
- In addition another uniform sample on (0,1) is generated, if this value is above the value of the pdf it is kept as a sample of the distribution.
- The sample mean of $k(x, \xi)$ and $\hat{k}(x, \eta)$ can now be compared.
Probability density function

**Figure:** Compare $E[\hat{k}(x, \eta)]$ and $E[k(x, \xi)]$ using Monte-Carlo method. Lognormal samples found using accept/reject technique.
Deterministic diffusion equation

- For the Monte-Carlo method the deterministic diffusion equation needs to be solved.

\[-\nabla \cdot (k(x)\nabla u(x)) = f(x)\,.
\]

- Let $D$ be square and discretize with bilinear elements on quadrilaterals of size $h$ by $h$.
- Let $\phi_j(x)$ denote the basis functions.
Deterministic diffusion equation

Find $u(x) \in H^1_E(D) = \{ u \in H^1(D) : u = g(x) \text{ on } \partial D_d \}$ such that

$$\int_D k(x) \nabla u(x) \cdot \nabla v(x) dx = \int_D f(x)v(x) dx$$

(14)

is satisfied for all $v(x) \in H^1_0(D)$.
Deterministic diffusion equation

- The finite element solution is

\[ u_h(x) = \sum_{j=1}^{n} u_j \phi_j(x) + \sum_{j=n+1}^{n+n_d} u_j \phi_j(x). \]  \hspace{1cm} (15)

where \( n \) is the number of elements on the interior and \( n_d \) is the number of elements on the boundary.

- To find the coefficients \( u_j \), solve \( Au = b \) where

\[ A_{ij} = \int_D k(x) \nabla \phi_j(x) \cdot \nabla \phi_i(x) \, dx \]  \hspace{1cm} (16)

\[ b_i = \int_D \phi_i(x)f(x) \, dx - \sum_{j=n+1}^{n+n_d} u_j \int_D k(x) \nabla \phi_j(x) \cdot \nabla \phi_i(x) \, dx \]  \hspace{1cm} (17)
Deterministic diffusion equation

- Implemented with Matlab package Incompressible Flow & Iterative Solver Software (IFISS) [10].

\[ D = [0, 1] \times [0, 1]. \]

\[ f(x) = 1, \quad g(x) = 0, \quad k(x) = 1, \quad h = 0.0625, \quad n_d = 64, \quad n = 225 \]

- Analytic solution [1]:

\[
\begin{aligned}
  u(x, y) &= \frac{16}{\pi^4} \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \frac{\sin((2k + 1)\pi x) \sin((2l + 1)\pi y)}{(2k + 1)(2l + 1)((2k + 1)^2 + (2l + 1)^2)} \\
  \|u_h(x) - u(x)\|_2^2 &= 3.31 \times 10^{-3} \quad (19)
\end{aligned}
\]
**Deterministic diffusion equation**

![Deterministic diffusion equation](image)

**Figure:** Deterministic solution, $k(x) = 1$
Monte Carlo Method

- Use the deterministic solver for each of $q$ finite element problems, denoting the solution $u^i_h(x)$ for $i = 1, \ldots, q$.

- The sample mean of the solution is

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

$$ \text{Var}_q[u_h] = \frac{1}{q-1} \sum_{i=1}^{q} (u^i_h(x)^2 - E_q[u_h]^2) $$

- The error in the mean is

Monte Carlo Method

- Using the KL expansion from the Gaussian random field

\[ k(x, \xi) = \exp[a_0(x) + \sum_{s=1}^{m_g} \sqrt{\lambda_s} a_s(x) \xi_s] \] (22)

where \( \xi_s \) are independent and \( N(0, 1) \) [13].

- Sample the \( m_g \) standard normal random variables \( q \) times to produce samples \( k_i(x) \) for \( i = 1, \ldots, q \).

- Running method with \( \sigma = 0 \) differs from the deterministic solution \( \sim 10^{-15} \).
Monte Carlo Method

Figure: Monte Carlo solution: $E[k(x, \xi)] = 1$, $f(x) = 1$, $m = 5$, $g(x) = 0$.

(a) $\sigma = 0.001$, $q = 100$  
(b) $\sigma = 0.5$, $q = 100000$
Stochastic weak formulation

- Write the solution as combination of basis functions which can be used to estimate statistical properties of the solution.
- The stochastic basis functions are analogous to the spatial basis functions used in the deterministic method.
- Using the KL expansion, the probability space, $\Omega$, is approximated by $\Gamma$, where $\Gamma$ is the support of the joint density function of the random variables in the expansion.
- The weak formulation of the problem is to find $u \in H^1(D) \otimes L^2(\Gamma)$ such that the following holds for all $v \in H^1_0(D) \otimes L^2(\Gamma)$

$$\int_{\Gamma} \int_{D} \hat{k}(x, \eta) \nabla u \cdot \nabla v \hat{\rho}(\eta) \, dx \, d\eta = \int_{\Gamma} \int_{D} f v \hat{\rho}(\eta) \, dx \, d\eta \quad (23)$$
Chaos polynomials

- The spatial discretization uses the same bilinear elements as the deterministic problem \( \phi_i(x) \).
- The stochastic discretization uses chaos polynomials
  \[
  \psi_j(\eta) = \psi_{j_1}(\eta_1)\psi_{j_2}(\eta_2)\ldots\psi_{j_m}(\eta_m) .
  \] (24)
- The chaos polynomials are chosen to be orthonormal so that
  \[
  E[\psi_i(\eta)\psi_j(\eta)] = \delta_{ij}.
  \]
Chaos polynomials

- The number of basis polynomials is chosen by setting an upper bound \((N)\) on the degree of the polynomials.

\[
\text{deg}(\psi_j) = \text{deg}(\psi_{j1}) + \ldots + \text{deg}(\psi_{jm}) \leq N \ \forall j
\] (25)

- The polynomials can be reindexed \(j = 1, \ldots, n_\eta\) where

\[
n_\eta = \binom{N + m}{m}.
\] (26)
Stochastic Galerkin method

- The solution is written as a combination of products of the two basis functions.

\[ u_h(x, \eta) = \sum_{i=1}^{n_x} \sum_{j=1}^{n_\eta} u_{ij} \phi_i(x) \psi_j(\eta) \]  

(27)

\[ v(x, \eta) = \phi(x) \psi(\eta) \]  

(28)

- The problem is to find the coefficients \( u_{ij} \) which satisfy

\[ \sum_{i=1}^{n_x} \sum_{j=1}^{n_\eta} \int_\Gamma \int_D u_{ij} \hat{k}(x, \eta) \nabla \phi_i(x) \cdot \nabla \phi_k(x) \psi_j(\eta) \psi_l(\eta) \hat{\rho}(\eta) dxd\eta = \int_\Gamma \int_D f(x) \phi_k(x) \psi_l(\eta) \hat{\rho}(\eta) dxd\eta \]  

(29)

for each \( k = 1, \ldots, n_x \) and \( l = 1, \ldots, n_\eta \).
Stochastic Galerkin method

- Define $\mu_0 = 1$ and $\eta_0 = 1$, so the KL expansion can be written

$$\hat{k}(x, \eta) = \sum_{s=0}^{m} \sqrt{\mu_s} k_s(x) \eta_s . \quad (30)$$

- The solution $u$ can be found by solving $\hat{A}u = b$ where

$$\hat{A} = \sum_{p=0}^{m} G_p \otimes A_p \quad (31)$$

$$[A_p]_{ik} = \int_{D} \sqrt{\mu_p} k_p(x) \nabla \phi_i(x) \cdot \nabla \phi_k(x) dx \quad (32)$$

$$[G_p]_{jl} = \int_{\Gamma} \eta_p \psi_j(\eta) \psi_l(\eta) \hat{\rho}(\eta) d\eta \quad (33)$$

$$b = \int_{D} f(x) \phi_k(x) dx \int_{\Gamma} \psi_l(\eta) \hat{\rho}(\eta) d\eta \quad (34)$$
The mean and the variance of the solution are

\[ E[u(x, \eta)] = \sum_{i=1}^{n_x} u_{i1} \phi_i(x) \]  

(35)

\[ \text{Var}[u(x, \eta)] = \sum_{i=1}^{n_x} \sum_{k=1}^{n_x} \sum_{j=2}^{n_\eta} u_{ij} u_{kj} \phi_i(x) \phi_k(x) \]  

(36)
Orthogonal polynomials

- Introduce the assumption $\hat{\rho}(\eta) = \hat{\rho}_1(\eta_1)\hat{\rho}_2(\eta_2)\ldots\hat{\rho}_m(\eta_m)$.

- The integral for $[G_p]_{jl}$ becomes

\[
[G_p]_{jl} = \int_{\Gamma_1} \psi_j(\eta_1)\psi_j(\eta_1)\psi_l(\eta_1)\hat{\rho}_1(\eta_1) d\eta_1 \ldots
\]
\[
\int_{\Gamma_p} \eta_p\psi_j(\eta_p)\psi_l(\eta_p)\hat{\rho}_p(\eta_p) d\eta_p \ldots
\]
\[
\int_{\Gamma_m} \psi_j(\eta_m)\psi_l(\eta_m)\hat{\rho}_m(\eta_m) d\eta_m
\]

- For each of the random variables the $i$th component of the orthogonal polynomials $\psi_j(\eta_i)$ is constructed using the three-term recurrence relation, where the coefficients are found using the Stieljtes procedure.
Orthogonal polynomials

- The \( k + 1 \) degree polynomial is:

\[
\psi_{k+1}(\eta_i) = (\eta_i - \alpha_k)\psi_k(\eta_i) - \beta_k\psi_{k-1}(\eta_i)
\]  

(38)

for \( k = 0, 1, \ldots \), where \( \psi_{-1}(\eta_i) = 0 \) and \( \psi_0(\eta_i) = 1 \).

- The recurrence coefficients are

\[
\alpha_k = \frac{\int \eta_i\psi_k(\eta_i)\psi_k(\eta_i)\rho_i(\eta_i)d\eta_i}{\int \psi_k(\eta_i)\psi_k(\eta_i)\rho_i(\eta_i)d\eta_i}
\]  

(39)

for \( k = 0, 1, 2, \ldots \) and

\[
\beta_k = \frac{\int \psi_k(\eta_i)\psi_{k-1}(\eta_i)\rho_i(\eta_i)d\eta_i}{\int \psi_{k-1}(\eta_i)\psi_{k-1}(\eta_i)\rho_i(\eta_i)d\eta_i}
\]  

(40)

for \( k = 1, 2, \ldots \).
Stieljes procedure

- Let \([a, b]\) be the support of \(\rho_i(\eta_i)\) and discretize with \(R\) points.
- The coefficients are:

\[
\alpha_{k,R} = \frac{\sum_{t=1}^{R} \eta_{it} w_t \psi_{k,R}(\eta_{it}) \psi_{k,R}(\eta_{it}) \rho_i(\eta_{it})}{\sum_{t=1}^{R} w_t \psi_{k,R}(\eta_{it}) \psi_{k,R}(\eta_{it}) \rho(\eta_{it})} \tag{41}
\]

\[
\beta_{k,R} = \frac{\sum_{t=1}^{R} w_t \psi_{k-1,R}(\eta_{it}) \psi_{k-1,R}(\eta_{it}) \rho_i(\eta_{it})}{\sum_{t=1}^{R} w_t \psi_{k-1,R}(\eta_{it}) \rho(\eta_{it})} \tag{42}
\]
The weights and nodes are found using a Fejer quadrature where the nodes are related to the roots of the Chebyshev polynomials.

\[
\eta_v = \frac{1}{2} (b - a) \cos \left( \frac{2v - 1}{2M} \right) + \frac{1}{2} (a + b)
\]

\[
w_v = \frac{1}{M} \left( 1 - 2 \sum_{n=1}^{\left\lfloor M/2 \right\rfloor} \frac{\cos(2n(\frac{2v-1}{2M}))}{4n^2 - 1} \right)
\]

for \( v = 1, \ldots, M \).
Stieljes procedure

- The procedure was implemented for Matlab by Gautschi [4] where the interval is broken up into component intervals.
- This procedure was called to construct the polynomials for each of the $m$ marginal density functions.
Results, $m = 1$

Figure: Stochastic Galerkin solution: $E[k(x, \eta)] = 1$, $f(x) = 1$, $m = 1$, $\sigma = 0.1$, $b_x = b_y = 10$. 

(a) Mean  
(b) Variance
Results, $m = 1$

- Validation of the stochastic Galerkin method is achieved by comparing with the Monte-Carlo solution.
- With standard deviation, $\sigma = 0.1$, and correlation lengths, $b_x = b_y = 10$ the first eigenvalue includes 93.22% of the variance.

\[
\frac{\|E[u]_{MC} - E[u]_{SG}\|_2}{\|E[u]_{MC}\|_2} = 4.61 \times 10^{-4}
\]

\[
\frac{\|sd_{MC} - sd_{SG}\|_2}{\|sd_{MC}\|_2} = 9.00 \times 10^{-3}
\]

- Given the small standard deviation, the solution is not so different from the deterministic result.
Results, $m = 2$

Figure: Stochastic Galerkin solution: $E[k(x, \eta)] = 1$, $f(x) = 1$, $m = 2$, $\sigma = 0.5$, $b_x = b_y = 10$. 

(a) Mean  
(b) Variance
Results, $m = 2$

- Assume $\hat{\rho}(\eta) = \rho_1(\eta_1)\rho_2(\eta_2)$.
- $\sigma = 0.5$ and $b_x = b_y = 10$, which incorporates $94.67\%$ of the variance in the first two eigenvalues.

\[
\frac{\|E[u]_{MC} - E[u]_{SG}\|_2}{\|E[u]_{MC}\|_2} = 3.95 \times 10^{-3}
\]

\[
\frac{\|sd_{MC} - sd_{SG}\|_2}{\|sd_{MC}\|_2} = 1.32 \times 10^{-1}
\]
Results, $m = 2$

- The Monte-Carlo method with $q = 100,000$ takes approximately 3.5 hours.
- The stochastic Galerkin method took approximately 0.5 hours.
- Unlike the Monte-Carlo method, the SG method scales as a function of the number of random variables.
- The majority of the time spent on the stochastic Galerkin method is in computing the two marginal density functions needed.
Deliverables

- Code to compute the moments of the solution using the Monte-Carlo method
  - Verified using $\sigma = 0$ and comparing with deterministic solution.
- Code to compute the moments of the solution using a KL expansion and stochastic Galerkin method.
  - Implemented for expansions of up to two random variables and standard deviation up to $\sigma = 0.5$ and verified using the Monte-Carlo results.
- Comparison of the results for varying number of terms in the KL expansion.
- Comparison of computational cost for the two methods
Conclusion

- The stochastic Galerkin method performs faster than Monte-Carlo methods for $m = 1$ and $m = 2$.
- A different quadrature routine to compute the marginal density functions could improve computation time.
- The assumption about separability of the density function does not hold for standard deviations much higher than $\sigma = 0.5$.
- Having the joint density function illustrates using the direct expansion of the lognormal random field can be used to solve this problem.
- The stochastic collocation method does not require orthogonal polynomials, so no assumption of separability would be needed.
References


References


