A Stochastic Method for Flow in Random Porous Media

Darin Comeau

Oral Comprehensive Exam

December 3, 2010

Problem of Interest

- Modeling subsurface flow accurately requires knowledge of the hydraulic conductivity parameter K.
- In reality, measurements of K are only known at nonuniform locations, contain uncertainty, and can vary over orders of magnitude within an aquifer.
- Solving the problem deterministically is unreasonable; instead K is thought of as being a *random field*.
- Stochastic modeling efforts typically limited by homogeneity in K; unrealistic in real world settings.

Goal

The goal is to develop an efficient method for solving the stochastic flow equation in random porous media.

Problem Statement

The steady state flow equation through a random porous medium D in saturated conditions is governed by:

Continuity Equation & Darcy's Law

$$\nabla \cdot q(\mathbf{x}) = -g(\mathbf{x}) \qquad q(\mathbf{x}) = -K(\mathbf{x})\nabla h(\mathbf{x}) \qquad \mathbf{x} \in D$$

- $q(\mathbf{x})$ is flux, $h(\mathbf{x})$ is hydraulic head, $g(\mathbf{x})$ is source/sink term.
- $K(\mathbf{x})$ is hydraulic conductivity - uncertainty leads to treating the parameter as a random variable, log normally distributed.
- Reparameterize by normally distributed $Y(\mathbf{x}, \theta) = \log K(\mathbf{x}, \theta)$.

Stochastic Flow Equation

$$\nabla \cdot [\exp(Y(\mathbf{x}, \theta))\nabla h(\mathbf{x})] = g(\mathbf{x}) \qquad \mathbf{x} \in D, \theta \in \Theta$$

Solution Approach

We will take the following steps to solve the stochastic flow equation, i.e. solve for the statistics of the random field $h(\mathbf{x})$:

1. Decompose the log conductivity field Y into a Karhunen-Loeve (KL) expansion.
2. Decompose the hydraulic head field h with a polynomial chaos (PC) expansion.
3. Evaluate the coefficients of the PC expansion using the probabilistic collocation method (PCM).
4. Calculate statistical information about h using these PC coefficients.

KL expansion

Decompose Y into its deterministic (mean) and random (fluctuation) components:

$$Y(\mathbf{x},\theta) = \overline{Y}(\mathbf{x}) + Y'(\mathbf{x},\theta)$$

The covariance of the zero-mean fluctuation Y' is given by

$$C_Y(\mathbf{x},\mathbf{y}) = \left\langle Y'(\mathbf{x},\theta)Y'(\mathbf{y},\theta)\right\rangle$$

This is a symmetric, positive definite operator, and thus admits a spectral decomposition:

$$C_Y(\mathbf{x},\mathbf{y}) = \sum_{n=1}^{\infty} \lambda_n f_n(\mathbf{x}) f_n(\mathbf{y})$$

KL expansion

The (deterministic) eigenfunctions and eigenvalues are solutions of the integral equation

$$\int_D C_Y(\mathbf{x},\mathbf{y}) f_n(\mathbf{x}) d\mathbf{x} = \lambda_n f_n(\mathbf{y})$$

and have the following properties:

- The set of eigenfunctions $\{f_n(\mathbf{x})\}_{n=1}^{\infty}$ forms a complete orthonormal set for $L^2(D)$.
- For each eigenvalue λ_n, there corresponds at most a finite number of linearly independent eigenfunctions.
- There are at most a countably infinite set of eigenvalues, all of which are positive real numbers.

KL expansion

Use this eigenfunction basis to expand the random process Y in the Karhunen-Loeve expansion:

$$Y(\mathbf{x}, \theta) = \overline{Y}(\mathbf{x}) + \sum_{n=1}^{\infty} \xi_n(\theta)\sqrt{\lambda_n} f_n(\mathbf{x})$$

where the ξ_i are orthogonal random variables with mean zero, unit variance, and the eigenvalues are arranged to be non-increasing. The variance of Y is distributed among the eigenvalues λ_n:

$$\sigma_Y^2 \mathsf{Vol}(D) = \sum_{n=1}^{\infty} \lambda_n$$

Additionally, when the underlying process Y is Gaussian, the ξ_i are actually standard normal random variables.

KL expansion

Based on the rate of decay of the eigenvalues λ_n, we can truncate the KL expansion N terms to approximate Y.

Error Minimization

The KL expansion of Y is optimal in the sense that the expected squared truncation error is minimized.

The stochastic flow equation we are trying to solve is

$$\nabla \cdot \left\{ \exp\left[\overline{Y}(\mathbf{x}) + \sum_{n=1}^{N} \xi_n(\theta)\sqrt{\lambda_n} f_n(\mathbf{x}) \right] \nabla h(\mathbf{x}) \right\} = g(\mathbf{x})$$

PC expansion

Since the covariance of the hydraulic head field is not known, need an alternate expansion, and we turn to polynomial chaos:

- Let $\{\xi_n(\theta)\}_{n=1}^{\infty}$ be a set of orthogonal normal random variables on a probability space Θ.
- Let $\hat{\Gamma}_d$ be the set of all polynomials on the set of indeterminates $\{\xi_n(\theta)\}_{n=1}^{\infty}$ of degree $\leq d$.
- Let $\Gamma_d \subset \hat{\Gamma}_d$ be the set of polynomials orthogonal to $\hat{\Gamma}_{d-1}$.
- Γ_d is the *polynomial chaos of order* d.
- The space spanned by the union of the sets of polynomials Γ_d will, be dense in $L^2(\Theta)$ (under some general conditions).

Note polynomials in Γ_d have infinitely many terms for finite d.

PC expansion

Thus a general square integrable random variable $\phi \in L^2(\Theta)$ can be expressed as

$$\begin{aligned}\phi(\theta) &= a_0\Gamma_0 \\ &+ \sum_{i_1=1}^{\infty} a_{i_1}\Gamma_1\left(\xi_{i_1}(\theta)\right) \\ &+ \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1} a_{i_1 i_2}\Gamma_2\left(\xi_{i_1}(\theta),\xi_{i_2}(\theta)\right) \\ &+ \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3}\Gamma_3\left(\xi_{i_1}(\theta),\xi_{i_2}(\theta),\xi_{i_3}(\theta)\right) + \cdots\end{aligned}$$

Reindex and concisely write as

$$\phi(\theta) = \sum_{j=0}^{\infty} c_j \Psi_j(\xi(\theta)) \qquad \xi(\theta) = \{\xi_n(\theta)_{n=1}^{\infty}\}$$

PC expansion

To truncate this expansion, there are two controlling factors:

- random dimension N - the number of random variables $\{\xi_n(\theta)\}_{n=1}^N$ that are used as indeterminates.
- order d - the highest order of PC Γ_d used.

The number of terms in the expansion is then

$$p = \binom{N+d}{d} = \frac{(N+d)!}{N!d!}$$

PC expansion of h

$$\hat{h}(\mathbf{x},\theta) = \sum_{j=1}^{p} c_j(\mathbf{x})\Psi_j\left(\xi_1(\theta),\ldots,\xi_N(\theta)\right)$$

We need to determine the random functionals $\Psi_j(\xi(\theta))$ and deterministic coefficients $c_j(\mathbf{x})$.

PC expansion

Aside - Hermite polynomials

The nth Hermite polynomial is given by

$$H_n(x) = (-1)^n \exp(x^2) \frac{d^n}{dx^n} \exp(-x^2)$$

The set of Hermite polynomials form a complete orthogonal basis for the Hilbert space with weighted inner product:

$$(f, g) = \int_{\mathbb{R}} f(x) g(x) \exp\left(-\frac{x^2}{2}\right) dx$$

The first few Hermite polynomials are

$$H_0(x) = 1 \qquad H_1(x) = x$$

$$H_2(x) = x^2 - 1 \qquad H_3(x) = x^3 - 3x$$

PC expansion

If $\{\xi_n(\theta)\}_{n=1}^N$ are Gaussian random variables, then orthogonality is with respect to Gaussian measure. These are exactly multidimensional Hermite polynomials, so we use its generating function to evaluate terms in the dth order polynomial chaos:

Terms of dth PC expansion for h

$$\Gamma_d(\xi) = (-1)^d \exp\left(\frac{1}{2}\xi^T\xi\right) \frac{\partial^d}{\partial\xi_{i_1}\cdots\partial\xi_{i_d}} \exp\left(-\frac{1}{2}\xi^T\xi\right)$$

where

$$\xi = \xi(\theta) = (\xi_{i_1}(\theta), \ldots, \xi_{i_d}(\theta))^T$$

The functionals $\Psi_j(\xi(\theta))$ will be be made of terms of this form. We now need to determine the coefficients $c_j(\mathbf{x})$.

Weighted Residual

For a general differential operator $\mathcal{L}$ and a stochastic differential equation $\mathcal{L}y(\mathbf{x},\theta) = f(\mathbf{x})$, with approximate solution $\hat{y}(\mathbf{x},\theta) = \sum_{n=1}^{p} c_j(\mathbf{x})\Psi_j(\xi(\theta))$, the residual for the differential equation is

$$R\left(\xi(\theta)\right) = \mathcal{L}\hat{y} - f$$

We want to minimize this with respect to the random vector $\xi(\theta)$ in some sense, so we form a weak formulation and integrate against some chosen weight functions $\{w_j(\xi(\theta))\}$:

$$\langle R\left(\xi(\theta)\right) w_j(\xi(\theta))\rangle = 0$$

PCM - Solving Residual

In the probabilistic collocation method, the weighting functions are delta functions at prescribed collocation points ξ^j:

$$w_j(\xi(\theta)) = \delta(\xi(\theta) - \xi^j)$$

For each j, the residual is $R(\xi^j) = 0$, and has solution $\hat{h}^j(\mathbf{x})$ to:

$$\nabla \cdot \left\{ \exp\left[\overline{Y}(\mathbf{x}) + \sum_{n=1}^{N} \xi_n^j \sqrt{\lambda_n} f_n(\mathbf{x}) \right] \nabla \hat{h}^j(\mathbf{x}) \right\} - g(\mathbf{x}) = 0$$

This is a deterministic equation, which we can solve for $\hat{h}^j$, a *representation* of the hydraulic head field.

PCM - Determining Coefficients

Need to use representations $\hat{h}^j$ to determine coefficients c_j in

$$\hat{h}(\mathbf{x}, \theta) = \sum_{j=1}^{p} c_j(\mathbf{x}) \Psi_j(\xi(\theta))$$

From p collocation ξ^j and corresponding representations $\hat{h}^j$, form matrix equation

$$\tilde{h}(\mathbf{x}) = C(\mathbf{x}) Z$$

$\tilde{h}(\mathbf{x}) = [\hat{h}_1(\mathbf{x}), \ldots, \hat{h}_p(\mathbf{x})]^T \quad Z_{ji} = \Psi_i(\xi^j) \quad C(\mathbf{x}) = [c_1(\mathbf{x}), \ldots, c_p(\mathbf{x})]^T$

If Z has full rank, can invert to solve for $C(\mathbf{x})$.

PCM - Choosing Collocation Points

- Since expansion of h is in terms of Hermite polynomials up to H_d, natural choice for values to populate entries of N dimensional ξ^j are roots of H_{d+1}.
- ξ is multivariate standard normal, so choose entries with highest probability (e.g. 0) first.
- Need to ensure $Z_{ji} = \Psi_i(\xi^j)$ is full rank, so reject vectors that lead to redundancy.

PCM - Post-Processing

Calculate moments of h

$$\hat{h}(\mathbf{x},\theta) = \sum_{j=1}^{p} c_j(\mathbf{x})\Psi_j(\xi(\theta))$$

By construction, Ψ_j are mean zero (except for constant term $\Psi_1 = 1$), and are mutually orthogonal. Mean and variance of h approximated by

$$\langle h(\mathbf{x})\rangle = c_1(\mathbf{x}) \qquad \sigma_h^2 = \sum_{j=2}^{p} c_j^2(\mathbf{x})\sigma_{\Psi_j}^2$$

Example - KL Expansion

- 1D example, interval of length $L = 10$ units, 151 uniformly spaced nodes, Dirichlet boundary conditions.
- Covariance $C_Y(x, y) = \sigma_Y^2 \exp(-|x - y|/\eta)$, with variance and correlation length σ_Y^2, η.
- Eigenfunctions/values for KL expansion are:

$$f_n(x) = \frac{1}{\sqrt{(\eta^2\omega_n^2 + 1)\frac{L}{2} + \eta}} \left[\eta\omega_n \cos(\omega_n x) + \sin(\omega_n x)\right]$$

$$\lambda_n = \sigma_Y^2 \frac{2\eta}{\eta^2\omega_n^2 + 1} \qquad \text{where} \quad (\eta^2\omega^2 - 1)\sin(\omega L) = 2\eta\omega\cos(\omega L)$$

Example - KL Expansion

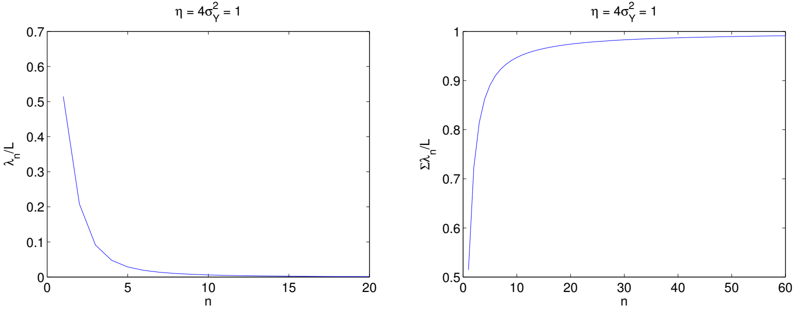

Figure: Eigenvalues for KL expansion with kernel $C_Y = \sigma_Y^2 \exp(-|x-y|/\eta)$. Example truncates at $N = 3$.

Example - PC Expansion

- PC expansion of h up to order $d = 2$:

$$\hat{h}(\mathbf{x}) = a_0(\mathbf{x}) + \sum_{i=1}^{3} a_i(\mathbf{x})\xi_i + \sum_{1=i\leq j=3} a_{i,j}(\mathbf{x})\left(\xi_i\xi_j - \delta_{ij}\right)$$

- Total of $p = 10$ terms in PC expansion.
- Collocation point entries chosen from roots of $H_3(x) = x^3 - 3x^2$, $0, \pm\sqrt{3}$ (i.e., (0,0,0) is the first collocation point).

Monte Carlo Benchmark

- Results compared to benchmark MC simulation with 10,000 realizations.
- Log conductivity field generation used Cholesky decomposition of covariance matrix.
- True variogram

$$\gamma(s) = \frac{1}{2}\left\langle (Y(x) - Y(x+s))^2 \right\rangle$$

$$= \sigma_Y^2 \left(1 - \exp(s/\eta)\right)$$

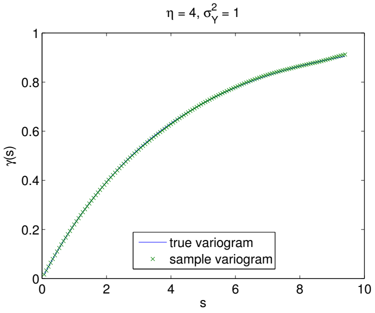

Figure: True variogram vs. Sample variogram of generated Y realizations.

Results - Head Profiles

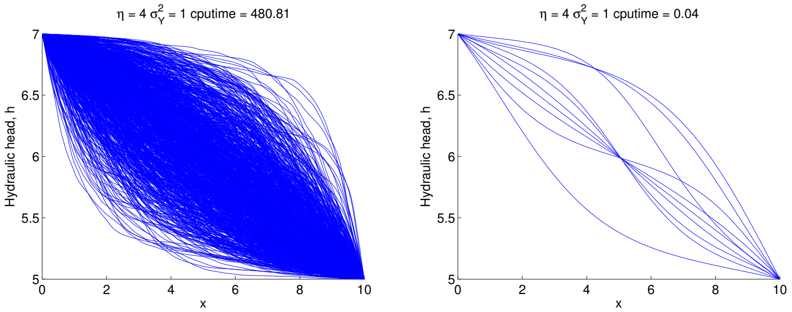

Figure: The first 1000 realizations of hydraulic head from 10,000 MC simulations (left). 10 representations of hydraulic head calculated using second-order PCM (right).

Results - Head Statistics

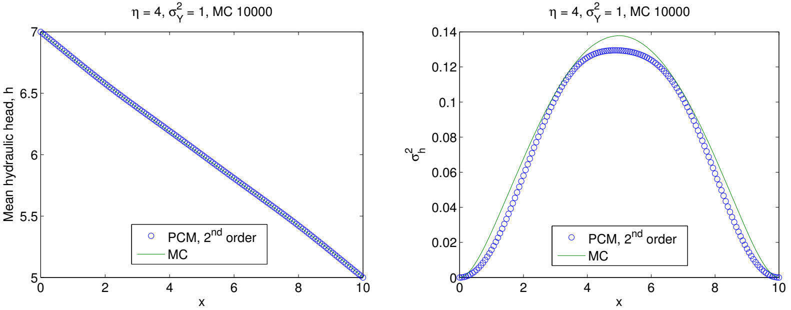

Figure: The mean (left) and variance (right) of the hydraulic head derived from the second order PCM and MC.

Results - Head Statistics

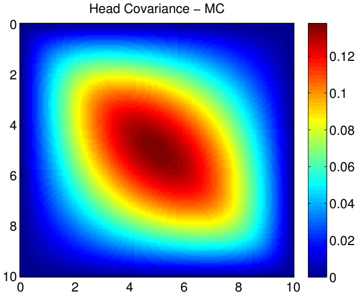

Figure: Covariance of the hydraulic head derived from MC (left) the second order PCM (right), calculated as $C_h(x,y) = \sum_{j=2}^{p} c_j(x)c_j(y)\sigma^2_{\Psi_j}$

Results - Effect of σ_Y^2

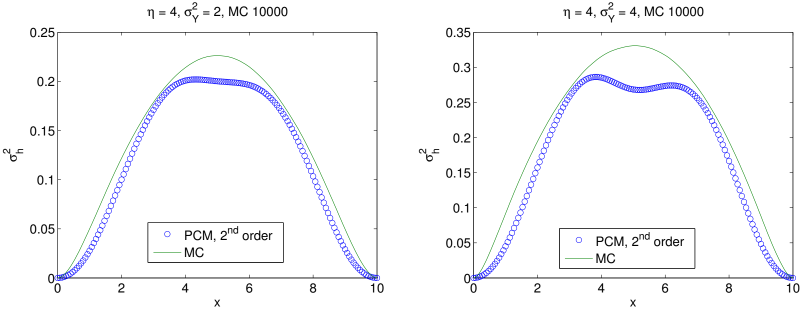

Figure: MC and 2nd order PCM head solutions with increased variance $\sigma_Y = 2, 4$.

Results - Effect of η

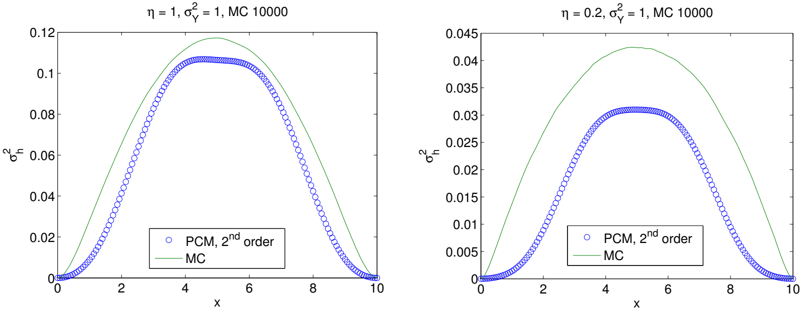

Figure: MC and 2nd order PCM head solutions with decreased correlation length $\eta = 1, .2$, corresponding to 15, 3 nodes per correlation length.

Summary

- We outlined a solution approach to the stochastic flow equation by:
  - Representing the log conductivity field with a KL expansion
  - Representing the hydraulic head field with a PC expansion
  - Evaluating the PC expansion coefficients with PCM
  - Evaluating statistics on the hydraulic head field through the coefficients
- We implemented the method in low random dimension and low order.
- Solutions were compared against benchmark MC simulation, performed well for certain variances and correlation length.
- Solutions began to deviate at large variance or small correlation length.

References

- R. Ghanem and PD Spanos. *Polynomial chaos in stochastic finite elements*. Journal of Applied Mechanics, 57(1):197202, 1990.
- RG Ghanem and PD Spanos. *Stochastic finite elements: a spectral approach*. Springer-Verlag, New York, 1991.
- **H. Li and D. Zhang. Probabilistic collocation method for flow in porous media: Comparisons with other stochastic methods. Water Resour. Res, 43:4448, 2007.**
- Y.K. Zhang. *Stochastic methods for flow in porous media: coping with uncertainties*. Academic Press, 2002.
- D. Zhang and Z. Lu. *An efficient, high-order perturbation approach for flow in random porous media via Karhunen-Loeve and polynomial expansions*. Journal of Computational Physics, 194(2):773794, 2004.

Acknowledgements

Thank you to Dr. Restrepo, Dr. Neuman, Dr. Winter, and Dr. Bhattacharya for forming my committee and academic support.