Generation of a stationary Gaussian random field

Jocelyne Erhel
SAGE team, Inria, Rennes, France

co-authors
Mestapha Oumouni (SAGE team, Inria, Rennes)
Géraldine Pichot (SAGE team, Inria, Rennes)
Anthony Beaudoin (UMR Pprime, university of Poitiers)
Jean-Raynald de Dreuzy (UMR Geosciences, CNRS, Rennes)

MOMAS, Nice, October 2015
1 Introduction
1 Introduction

2 Random field generation
1. Introduction
2. Random field generation
3. Numerical experiments
1 Introduction

2 Random field generation

3 Numerical experiments

4 Conclusion
**Random field** $K(x)$  
Domain $\Omega$ of $\mathbb{R}^d$, $x \in \Omega$  

$$K(x) = \exp(\mu + \sigma Y(x))$$  

where $Y$ is a zero mean normal correlated field,  
$\mu$ is the mean of $K$, $\sigma$ is the variance of $K(x)$  

**Stationary field**  
$\text{cov}[Y](x, y) = \text{cov}[Y](x - y, 0)$  
Covariance function:  

$$\text{cov}[Y](x, y) = C(|x - y|)$$  

Remark that $C(0) = 1$
Covariance matrix
Domain $\Omega = [0, 1]$
Discrete grid of $N + 1$ equally spaced points $x_0 = 0, x_1, \ldots, x_N = 1$
Sampled covariance function $c_i = C(x_i)$ with $i = 0, 1, \ldots, N$
Covariance matrix

$$R_{ij} = C(|x_i - x_j|) = \text{cov}[Y](x_i, x_j)$$

Discrete sampled field

$$Y_i = Y(x_i)$$

random vector of normal variables with zero mean and covariance matrix $R$

$$K_i = \exp(\mu + \sigma Y_i)$$
Covariance matrix $R$

$$R_{ij} = C(|x_i - x_j|) = \text{cov}[Y](x_i, x_j)$$

$R$ is a symmetric positive definite Toeplitz matrix of size $N + 1$

$$R = \begin{pmatrix}
  c_0 & c_1 & \cdots & c_{N-1} & c_N \\
  c_1 & c_0 & c_1 & \cdots & \cdots \\
  \vdots & \vdots & \ddots & \vdots & \vdots \\
  \vdots & \vdots & \ddots & c_1 & c_0 \\
  c_{N-1} & \cdots & c_1 & c_0 & c_1 \\
  c_N & c_{N-1} & \cdots & c_1 & c_0
\end{pmatrix}$$
Generating a sampled Gaussian field

Factorization

\[ R = BB^T \]

Vector \( \theta \) of random normal variables with zero mean and uncorrelated

\[ Y = B\theta \]

is a random vector of normal variables with zero mean and covariance matrix \( R \)

Generation of \( N + 1 \) random variables
Factorizations of a covariance matrix

- Cholesky factorization: $R = LL^T$
- Eigenvalue decomposition: $R = V\Delta V^T = (V\sqrt{\Delta})(V\sqrt{\Delta})^T$
- Spectral decomposition

Cholesky and eigenvalue decompositions are quite expensive for large $N$. The matrix $R$ can be ill-conditioned, implying numerical difficulties.
Circulant embedding of a covariance matrix: 1d case

Circulant embedding of $R$

$a = (c_0, \ldots, c_N, c_{N-1}, \ldots, c_1) \in \mathbb{R}^{2N}$

Circulant matrix $A = \text{circ}(a)$

The matrix $A$ is real symmetric of size $2N$ and is spd if $N$ is large enough

Example with $N = 4$

$$A = \begin{pmatrix} c_0 & c_1 & c_2 & c_3 & c_4 & c_3 & c_2 & c_1 \\ c_1 & c_0 & c_1 & c_2 & c_3 & c_4 & c_3 & c_2 \\ c_2 & c_1 & c_0 & c_1 & c_2 & c_3 & c_4 & c_3 \\ c_3 & c_2 & c_1 & c_0 & c_1 & c_2 & c_3 & c_4 \\ c_4 & c_3 & c_2 & c_1 & c_0 & c_1 & c_2 & c_3 \\ c_3 & c_4 & c_3 & c_2 & c_1 & c_0 & c_1 & c_2 \\ c_2 & c_3 & c_4 & c_3 & c_2 & c_1 & c_0 & c_1 \\ c_1 & c_2 & c_3 & c_4 & c_3 & c_2 & c_1 & c_0 \end{pmatrix}$$
Spectral decomposition of the circulant embedding

**Eigenvalue decomposition of** $A$

Discrete Fourier Transform: $F = (e^{-2i\pi kj/2N})_{0 \leq j, k \leq 2N-1}$

Inverse Discrete Fourier Transform: $F^{-1} = \frac{1}{2N} F^* = \frac{1}{2N} \bar{F}$

$$A = FDF^{-1} = \frac{1}{2N} FDF^*$$

where $D = \text{diag}(s)$ with $s = \bar{F}a = Fa$ (s is real)
$s_{2N-k} = s_k$, $k = 1, \ldots, N$
Generating two random fields with $4N$ random variables

**Factorization of $A$**
If $A$ is spd then $A = (F \sqrt{D/2N})(F \sqrt{D/2N})^*$

$$A = (1/\sqrt{2NF} \text{diag}(\sqrt{s}))(1/\sqrt{2NF} \text{diag}(\sqrt{s}))^*$$

If $A$ is not spd, $a$ can be completed by a padding or $N$ can be increased

**Random complex vector of size $2N$**

$$\theta = \text{Re}(\theta) + i\text{Im}(\theta)$$

$\text{Re}(\theta)$ and $\text{Im}(\theta)$ random normal zero mean uncorrelated real variables

$$Y_1 = \frac{1}{\sqrt{2N}} \text{Re}(F \text{diag}(\sqrt{s}) \theta)(0 : N) \text{ and } Y_2 = 1/\sqrt{2N} \text{Im}(F \text{diag}(\sqrt{s}) \theta)(0 : N)$$

Two vectors of random normal zero mean $N + 1$ variables with correlation $R$
Generating one random field with $2(N+1)$ random variables

Random complex vector $u + iv$ of size $N + 1$

$u$ and $v$ random Gaussian zero mean uncorrelated real variables with

$E[u_k^2] = E[v_k^2] = 1/2$, $k = 1, \ldots, N - 1$

$v_0 = v_N = 0$

$E[u_0^2] = E[v_N^2] = 1$

Random complex vector of size $2N$

$$\theta = \text{Re}(\theta) + i\text{Im}(\theta)$$

with

$\text{Re}(\theta) = (u_0, \ldots, u_N, u_{N-1}, \ldots, u_1)$ and $\text{Im}(\theta) = (v_0, \ldots, v_N, -v_{N-1}, \ldots, -v_1)$

$$Y = \frac{1}{\sqrt{2N}}(F\text{diag}(\sqrt{s})\theta)(0 : N)$$

Vector of real normal zero mean $N + 1$ variables with correlation $R$
Using \((N + 1)\) random variables with uniform law

**Approximation of a Gaussian law by a uniform law**

Real random vector \(\phi\) of size \(N + 1\)
with \(\phi_k \in [0, 2\pi]\) random variable of uniform law

**Random complex vector of size \(2N\)**

\(\theta_k = e^{i\phi_k}, k = 1, \ldots, N - 1\) and \(\theta_0 = \sqrt{2}\cos(\phi_0), \theta_N = \sqrt{2}\cos(\phi_N)\)
and \(\theta = (\theta_0, \ldots, \theta_N, \overline{\theta_{N-1}}, \ldots, \overline{\theta_1})\)

\[Y = \frac{1}{\sqrt{2N}}(F\text{diag}(\sqrt{s})\theta)(0 : N)\]

Vector of real zero mean \(N + 1\) variables with approximate Gaussian law and correlation \(R\)
Algorithm with uniform laws

- Sample the covariance function: vector $c$ and circulant vector $a$
- Apply DFT to compute $s = Fa$
- Generate random phase vector $\phi$ and compute complex vector $\theta$
- Apply DFT to compute $F(diag(\sqrt{s})\theta)$
- Compute random discrete field $Y$

Implementation using FFT and random number generator
Discrete regular grid in the domain $\Omega$ with $N + 1$ points in each direction
The covariance matrix $R$ is block-Toeplitz
The embedding uses symmetry in each direction to get a block-circulant matrix $A$
The random vectors also use symmetry in each direction
Parallel algorithm

Domain decomposition of the regular grid with one process per subdomain
Splitting of the circulant vector \( a \) in FFT with one process per block
Communications between processes to redistribute data
First numerical example: exponential covariance

Domain $\Omega = [0 \, L_x] \times [0 \, L_y]$

Covariance function

$$C(x) = \exp(-|x_1|/l_1 - |x_2|/l_2)$$

Ergodic field

$$\lim_{|D| \to \infty} \frac{1}{|D|} \int_D Y(x) \, dx = E[Y] = 0$$

$$\lim_{|D| \to \infty} \frac{1}{|D|} \int_D Y(x)^2 \, dx = C(0) = 1$$
Domain size: $L_x = L_y = 2^8 = 256$ and correlation lengths: $l_1 = l_2 = 2$

Two discrete random fields
Monte-Carlo convergence: $M$ realizations of $Y^{(m)}$

$$\tilde{C}_i = \frac{1}{M} \sum_{m=1}^{M} Y_i^{(m)} Y_0^{(m)}$$

Error $|C_i - \tilde{C}_i|$ with $M = 10^4$
Approximate mean and variance using ergodicity

\[ \mu = \frac{1}{N + 1} \sum_{i=0}^{N} Y_i \]

\[ \sigma^2 = \frac{1}{N} \sum_{i=0}^{N} Y_i^2 \]

<table>
<thead>
<tr>
<th>N</th>
<th>(2^5 \times 2^5)</th>
<th>(2^8 \times 2^8)</th>
<th>(2^{10} \times 2^{10})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mu)</td>
<td>0.0281</td>
<td>-0.0102</td>
<td>0.004</td>
</tr>
<tr>
<td>(\sigma^2)</td>
<td>0.9579</td>
<td>0.9983</td>
<td>0.9988</td>
</tr>
</tbody>
</table>
Domain $\Omega = [0 \, L_x] \times [0 \, L_y]$

Covariance function

$$C(x) = \exp(-|x_1|^2/l_1^2 - |x_2|^2/l_2^2)$$

The circulant matrix $A$ is spd only for large domains and large $N$
Domain size: $L_x = L_y = 2^8 = 256$ and correlation lengths: $l_1 = l_2 = 3$
One discrete random field $\sigma_Y$ with $\sigma = 2$
Monte-Carlo error on the covariance function with $M = 10^4$
Third numerical example: Gaussian covariance

Domain $\Omega = [0 \ L_x] \times [0 \ L_y]$

Covariance function

$$C(x) = \exp(-|x_1|^2/l_1^2 - |x_2|^2/l_2^2)$$

The circulant matrix $A$ is spd only for large domains and large $N$
Numerical results

Domain size: $L_x = 2$ and $L_y = 1$ and correlation lengths: $l_1 = l_2 = 1/10$

One discrete random field $K = \exp(\sigma Y)$ with $\sigma = 1$

Monte-Carlo error on the covariance function with $M = 10^4$
Conclusion

Summary

- Several variants for generating a correlated Gaussian field
- Parallel software GENFIELD (soon available)
- Use of random fields in hydrogeology
Conclusion

Summary

- Several variants for generating a correlated Gaussian field
- Parallel software GENFIELD (soon available)
- Use of random fields in hydrogeology

Future work

- Comparison of K-L expansion and spectral decomposition
- Generation of $K$ when the discrete density is given
- Use of random fields with multigrid methods