

共分散行列の平方根分解をもとにした 正規確率場および対数正規確率場の発生法

立川康人¹・椎葉充晴²

¹正会員 博士(工学) 京都大学助教授 防災研究所水害研究部門(〒611-0011 宇治市五ヶ庄)

²正会員 工博 京都大学教授 工学研究科土木システム工学専攻(〒606-8501 京都市左京区吉田本町)

Chebyshev 多項式を用いて共分散行列を近似的に平方根分解することを原理とした正規確率場、および対数正規確率場の発生法について述べる。対象領域を同じ大きさの矩形グリッドで覆い、homogeneous な正規確率場、または対数正規確率場に従う $N_x \times N_y$ 個の値を発生させる場合、ここで提案するアルゴリズムを用いれば、実行に必要な計算機メモリは最低 $3N_x N_y$ 個でよい。

Key Words : Gaussian random field, lognormal random field, matrix decomposition, Chebyshev polynomial

1. はじめに

分布型流出モデルにおけるモデルパラメータの空間分布、または降水分布が流出シミュレーション結果にどのような影響を与えるかを調査するための基本的なツールとして、正規確率場、または対数正規確率場を発生させる計算機アルゴリズムを提案する。地下水文学、水文地質学、石油工学などの分野では、透水係数などの流れのパラメータの空間場をある確率場モデルに基づいて発生させ、それが流れにどのような影響を持つかを解析する研究がこれまでにも相当数行われている。表層の雨水の流れ解析においても、モデルパラメータの空間場の模擬発生と分布型流出モデルとを組み合わせた雨水流れの解析は、モデルパラメータの空間分布が流出にどのような影響を及ぼすかを調査するための有効な手法である。また、降雨場の時空間分布が流出に及ぼす影響を解析する上でも、このような確率場の発生アルゴリズムは有効に機能するものと思われる。

正規確率場の発生法としてはこれまでに

- スペクトル法
- Turning Bands 法
- 行列分解法

が提案されてきた(文献^{1),2),3)}などを参照されたい)。スペクトル法、Turning Bands 法は与えられた共分散構造を近似的に満たす確率場を発生させる方法であり、発生させる場の確率変数の個数が多い場合に用いられる。一方、行列分解法は以下で述べる LU 分解法が代表的な方法であり、与えた共分散構造を正しく反映させる確率場の発生法である。シミュレーションの領域が小さい場合は、この方法が最も有効な方法である。

ここで提案する方法は、Davis⁴⁾の手法をもとにした行列分解法のうちの一つであり、Chebyshev 多項式を用いて近似的に共分散行列を平方根分解する原理を用いる。このアルゴリズムは、シミュレーション領域が大きい場合にも適用でき、かつ計算機メモリの使用量も少ない。対象領域を同じ大きさの矩形グリッドで覆い、発生点が $N_x \times N_y$ 個の homogeneous な確率場を発生させる場合、実行に必要な計算機メモリは最低 $3N_x N_y$ 個でよい。また、この方法は直接的であり、スペクトル法や Turning Bands 法に比べて簡単でわかりやすい。

2. 水文量の時空間場の発生とそれを水文流出システムに用いる意義

分布型流出モデルやレーダー雨量計のデータを扱うことができるようになり、水文量の空間分布を直接考慮する水文流出シミュレーションが比較的容易にできるようになった現在、筆者らは、以下のことがこれからの重要な研究課題であると考えている。

- 分布型流出モデルにおけるパラメータの決定法
- 流出現象のモデル化スケールの明確化
- 空間分布する水文量を直接考慮する新たな洪水防御計画の考え方

分布型流出モデルを実用的なモデルとしていくためには、モデルパラメータをどのように決定するかということが重要な課題となる。モデルパラメータは、通常、モデル構造やモデル分解能に依存するため、雨量・流量データから決定せざるを得ない場合が多い。しかし、分布型モデルで必要とする空間的に分布する数種

類のモデルパラメータを水文データのみから決定することは容易ではない。これに対処するためには、モデルパラメータの様々な空間場を発生させて、モデルパラメータの空間分布(流域内部での分布特性や空間相関スケールなど)の仕方が流出に及ぼす影響を解析する必要がある。この解析によって、たとえば、パラメータ分布の分散のみがハイドログラフの形成を支配しており、それ以外の分布情報は重要ではないなどの知見が得られれば、それをもとに分布モデルパラメータの決定に閑して新たな考え方を創出することができる。

流出現象のモデル化スケールを明確にすることも実用上重要な課題である。ある対象流域が与えられた場合、そこでの流出を再現するための適切なモデルの時間・空間分解能(分割基準)は、流域の大きさ、降雨場変動の時空間スケールなどに関連すると予想される。これらの関係を明らかにすることにより、合理的なモデル化の分解能(分割基準)を提示することができる。そのためには、時空間スケールの異なる降雨場を発生させてそれを分布型流出モデルへの入力とし、流域面積の異なる多地点での流出シミュレーション結果を分析する方法が考えられる。正規確率場の理論を用いた降雨場の発生モデルとしては Bras and Rodriguez-Iturbe⁵⁾によるモデルがある。

さらに、分布型流出モデルと降雨場の確率・統計的発生モデルとの組合せは新たな洪水防御計画を考えることを可能とする。たとえば、従来の基本高水の決定法では、空間的に変動する水文量の取扱い方が不十分である。仮に、レーダー雨量計によって降雨場の確率・統計的な特性を抽出することができ、その確率・統計的特性に従う降雨場を発生させることができれば、それを分布型流出モデルへの入力として流域の様々な地点での河川流量を確率的に評価することができるようになる。

これらの研究課題に取り組むための共通の基本的ツールとして、確率場の発生手法が必須となる。

3. 正規確率場の発生法

(1) LU 分解法

正規確率場の伝統的な発生法は、共分散行列を LU 分解(Cholesky 分解)する方法である。 R を $N \times N$ 次元の共分散行列とすると R は対称行列であり、 R が正定値行列ならば R を次のように Cholesky 分解することができる。

$$R = LL^T$$

ここで L は $N \times N$ の下三角行列である。この下三角行列 L に、 N 次元の互いに無相関なランダムベクトル $w \sim N(0, I)$, I は $N \times N$ の単位行列、を乗ずれば、

得られる N 次元ベクトル $y = Lw$ が求めるべき確率場となる。なぜなら

$$\begin{aligned} E[yy^T] &= E[Lww^T L^T] \\ &= LE[ww^T]L^T \\ &= LL^T \\ &= R \end{aligned}$$

であり、 y の共分散行列は R となるからである。ここで $E[\cdot]$ は確率変数の平均を取る操作の意味で用いた。 w は、たとえば Box and Muller 法を用いて連続的に発生させた正規乱数を N 個並べて作成したベクトルとすればよい。

この方法は非常に簡単で直接的な方法であるが、共分散行列のサイズが大きくなると、計算機メモリを大量に取るために実行が不可能となる欠点を持つ。たとえば、二次元の領域を考えて、そこに 100×100 のグリッドを設定し、そのグリッド点での値をある共分散行列に従って発生させるとする。この場合、共分散行列のサイズは $100^2 \times 100^2 = 10^8$ となる。共分散行列をある閲数で与えてメモリを取らないようにするとしても、それを Cholesky 分解した結果を格納しておくためには、約 5×10^7 個の変数を格納するメモリが必要となる。一つの変数が利用する計算機メモリの容量を 4 byte とすると約 200 Mbyte の計算機メモリが必要なことになり、計算の実行が難しいことがわかる。

(2) 平方根分解法

計算機メモリ取得の問題を回避するために、Davis⁴⁾によって提案された手法をもとにした正規確率場の発生法を提案する。Davis は、共分散行列の平方根分解行列を多項式で近似することを原理とした正規確率場の発生法を述べた。ここでは、この多項式として Chebyshev 多項式を導入した手法とアルゴリズムを展開する。

まず、 $N \times N$ 次元の共分散行列 R を $N \times N$ 次元の対称行列 S の積

$$R = SS \quad (1)$$

に分解することを考える。もし、この行列 S を求めることができれば、LU 分解法と同様に S にランダムベクトル w を乗じてできるベクトル $y = Sw$ が求めるべき確率場となる。なぜならば

$$\begin{aligned} E[yy^T] &= E[Sww^TS^T] \\ &= SE[ww^T]S^T \\ &= SS^T \\ &= SSS^T \\ &= SSS^T \\ &= R \end{aligned}$$

となって、 y の共分散行列は R となるからである。

共分散行列 \mathbf{R} は対称行列であり、かつ非負正定値行列なので、次のように固有値分解することができる。

$$\mathbf{R} = \mathbf{Q} \mathbf{A} \mathbf{Q}^T$$

ここで \mathbf{Q} は $N \times N$ 次元の直交行列、 \mathbf{A} は $N \times N$ 次元の対角行列であり、 $\mathbf{A} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$ 、であって、対角成分には \mathbf{R} の固有値が並ぶ。 \mathbf{R} は非負正定値行列なので、すべての固有値は 0 以上の値を取る。したがって、

$$\mathbf{A}^{1/2} = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_N})$$

という行列を設定することができ、 \mathbf{S} を

$$\mathbf{S} = \mathbf{Q} \mathbf{A}^{1/2} \mathbf{Q}^T \quad (2)$$

とすれば、 \mathbf{S} は次式によって式(1)を満たしていることがわかる。

$$\begin{aligned} \mathbf{S} \mathbf{S} &= \mathbf{Q} \mathbf{A}^{1/2} \mathbf{Q}^T \mathbf{Q} \mathbf{A}^{1/2} \mathbf{Q}^T \\ &= \mathbf{Q} \mathbf{A}^{1/2} \mathbf{A}^{1/2} \mathbf{Q}^T \\ &= \mathbf{Q} \mathbf{A} \mathbf{Q}^T \\ &= \mathbf{R} \end{aligned}$$

以上のように、共分散行列 \mathbf{R} に対する平方根行列 \mathbf{S} は常に存在する。問題は \mathbf{S} をどのように求めるか、または近似するかということになる。直接、共分散行列を固有値分解することは、LU 分解すること以上に計算負荷が大きくなるために採用できない。そこで、以下に述べる近似手法を用いる。

式(2)を具体的に書くと

$$\mathbf{S} = \mathbf{Q} \begin{pmatrix} \sqrt{\lambda_1} & & & 0 \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ 0 & & & \sqrt{\lambda_N} \end{pmatrix} \mathbf{Q}^T \quad (3)$$

となる。ここで、固有値の最小値と最大値を含む区間で \sqrt{x} を多項式

$$\sqrt{x} = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \dots + \alpha_p x^p + \delta \quad (4)$$

で近似することを考える。ここで、 $\alpha_0, \dots, \alpha_p$ は定係数であり δ は近似誤差を表す。式(3)中の対角行列の対角成分に式(4)を適用すると

$$\begin{aligned} \mathbf{S} &= \mathbf{Q} \begin{pmatrix} \alpha_0 + \alpha_1 \lambda_1 + \dots + \alpha_p \lambda_1^p + \delta_1 & & & 0 \\ & \alpha_0 + \alpha_1 \lambda_2 + \dots + \alpha_p \lambda_2^p + \delta_2 & & \\ & & \ddots & \\ 0 & & \alpha_0 + \alpha_1 \lambda_N + \dots + \alpha_p \lambda_N^p + \delta_N & \end{pmatrix} \mathbf{Q}^T \\ &= \mathbf{Q} (\alpha_0 \mathbf{I} + \alpha_1 \mathbf{A} + \dots + \alpha_p \mathbf{A}^p + \Delta) \mathbf{Q}^T \\ &= \alpha_0 \mathbf{I} + \alpha_1 \mathbf{R} + \dots + \alpha_p \mathbf{R}^p + \mathbf{Q} \Delta \mathbf{Q}^T \end{aligned} \quad (5)$$

となる。ここで、 $\Delta = \text{diag}(\delta_1, \delta_2, \dots, \delta_N)$ 、 \mathbf{I} は $N \times N$ 次元の単位行列であり、上式の展開には

$$\mathbf{R}^n = (\mathbf{Q} \mathbf{A} \mathbf{Q}^T)^n = \mathbf{Q} \mathbf{A}^n \mathbf{Q}^T$$

を用いた。結局、式(5)に示すように、固有値の最大値と最小値を含む区間で \sqrt{x} を多項式で精度良く近似することができれば、 \mathbf{S} の近似値を得ることができる。近似区間の下端は 0 とすればよい。ここでの固有値は 0 以上の値を取るからである。上端は Gershgorin circle theorem⁶⁾ を用いて、共分散行列 \mathbf{R} の各要素の絶対値の各行和の最大値

$$r_M = \max_i(r_i), \quad r_i = \sum_{j=1}^N |R_{i,j}|, \quad i = 1, \dots, N \quad (6)$$

とする。 r_M の値は固有値の最大値以上の値となり、すべての固有値は区間 $[0, r_M]$ に含まれる。以上が Davis が展開した手法の要点である。

この \sqrt{x} の多項式による近似式として Chebyshev 多項式を用いることにする。Chebyshev 多項式は、minimax 多項式とほとんど同じ多項式であることが知られており、かつ容易に求めることができる。Chebyshev 多項式を用いると、区間 $[-1, 1]$ における任意のスカラー関数 $f(x)$ は

$$f(x) \simeq \left[\sum_{k=0}^{P-1} c_k T_k(x) \right] - \frac{c_0}{2} \quad (7)$$

と近似することができる⁷⁾。ここで

$$T_k(x) = \cos(k \arccos x), \quad k \geq 0 \quad (8)$$

である。 $c_k, k = 0, \dots, P-1$ は

$$\begin{aligned} c_k &= \frac{2}{P} \sum_{i=1}^P f(x_i) T_i(x_i) \\ &= \frac{2}{P} \sum_{i=1}^P f \left(\cos \left(\frac{\pi(i-1/2)}{P} \right) \right) \\ &\quad \times \cos \left(\frac{\pi k(i-1/2)}{P} \right) \end{aligned} \quad (9)$$

によって定められる係数であり、 P は項数を表す。 $x_i, i = 1, \dots, P$ は

$$x_i = \cos \left(\frac{\pi(i-1/2)}{P} \right)$$

として Chebyshev 多項式近似の理論によって定められる補間点である。式(8)を具体的に書くと

$$\left. \begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ &\vdots \\ T_{k+1}(x) &= 2xT_k(x) - T_{k-1}(x), \quad k \geq 1 \end{aligned} \right\} \quad (10)$$

となる。任意の区間 $a \leq x \leq b$ に対して Chebyshev 多項式近似を適用する場合は

$$x = \frac{b-a}{2}y + \frac{a+b}{2}, \quad -1 \leq y \leq 1 \quad (11)$$

として変数変換を施せば良い。

以上の準備をもとに $0 \leq x \leq r_M$ において

$$x = \frac{r_M}{2}(y+1), \quad -1 \leq y \leq 1$$

と変数変換し, $f(x) = \sqrt{x}$ に対して式(7)を適用して

$$\begin{aligned} \sqrt{x} &\simeq \left[\sum_{k=0}^{P-1} c_k T_k(y) \right] - \frac{c_0}{2} \\ &= \left[\sum_{k=0}^{P-1} c_k T_k \left(\frac{2}{r_M} x - 1 \right) \right] - \frac{c_0}{2} \end{aligned} \quad (12)$$

と近似する. この式を式(3)に適用すると

$$\begin{aligned} S &\simeq Q \begin{pmatrix} \left[\sum_{k=0}^{P-1} c_k T_k \left(\frac{2}{r_M} \lambda_1 - 1 \right) \right] - \frac{c_0}{2} & 0 \\ \ddots & \ddots \\ 0 & \left[\sum_{k=0}^{P-1} c_k T_k \left(\frac{2}{r_M} \lambda_N - 1 \right) \right] - \frac{c_0}{2} \end{pmatrix} Q^T \\ &= Q \left[\sum_{k=0}^{P-1} c_k T_k \left(\frac{2}{r_M} A - I \right) \right] Q^T - \frac{c_0}{2} I \\ &= \left[\sum_{k=0}^{P-1} c_k T_k \left(\frac{2}{r_M} Q A Q^T - I \right) \right] - \frac{c_0}{2} I \\ &= \left[\sum_{k=0}^{P-1} c_k T_k \left(\frac{2}{r_M} R - I \right) \right] - \frac{c_0}{2} I \\ &= \left[\sum_{k=0}^{P-1} c_k T_k(R') \right] - \frac{c_0}{2} I \end{aligned} \quad (13)$$

となる. ここで R' は $N \times N$ 次元の行列

$$R' = \frac{2}{r_M} R - I$$

であり, 式(13)中の $T_k(\cdot)$ は行列に対しても式(10)の演算を施す意味で用いた. なお, 式(13)の導出においては, $R = Q D Q^T$, Q を直交行列, $D = \text{diag}(d_1, \dots, d_N)$ とした場合に, $f(\cdot)$ は実数演算を施すオペレータであるとして

$$\begin{aligned} f(R) &= f(Q D Q^T) \\ &= Q f(D) Q^T \\ &= Q \begin{pmatrix} f(d_1) & & 0 \\ & \ddots & \\ 0 & & f(d_N) \end{pmatrix} Q^T \end{aligned}$$

であることを用いた. 求めたい正規確率場は式(13)の右辺に N 次元の互いに無相関のランダムベクトル w を乗じることによって得ることができる.

(3) 近似精度に関する考察

式(13)の Chebyshev 多項式の項数 P を決定するために, ある項数で打ち切った場合の誤差を見積もることを考える.

今, 式(13)において, 非常に大きな項数 L を取って S が S_L で十分に近似できたとする. 次ぎに, それを

項数 P で打ち切ったときの S の近似行列を S_P とする. このとき

$$\begin{aligned} S_L - S_P &= \sum_{k=P}^L c_k T_k(R') \\ &= Q \begin{pmatrix} \left[\sum_{k=P}^L c_k T_k \left(\frac{2}{r_M} \lambda_1 - 1 \right) \right] & 0 \\ \ddots & \ddots \\ 0 & \left[\sum_{k=P}^L c_k T_k \left(\frac{2}{r_M} \lambda_N - 1 \right) \right] \end{pmatrix} Q^T \end{aligned}$$

である. ここで, $S_L - S_P$ のノルム $\|S_L - S_P\|$ を求めることを考える. 任意の N 次元ベクトル x に対して $\|S_L - S_P\|$ は

$$\|S_L - S_P\| = \max_{x \neq 0} \frac{\|(S_L - S_P)x\|}{\|x\|}$$

で定義される.

$$\|(S_L - S_P)x\| \leq \|S_L - S_P\| \|x\|$$

なので, $\|S_L - S_P\|$ が充分小さくなるような P を取ればよいことになる. $\|S_L - S_P\|$ を求めると

$$\begin{aligned} \|S_L - S_P\| &= \|Q\| \left\| \begin{pmatrix} \left[\sum_{k=P}^L c_k T_k \left(\frac{2}{r_M} \lambda_1 - 1 \right) \right] & 0 \\ \ddots & \ddots \\ 0 & \left[\sum_{k=P}^L c_k T_k \left(\frac{2}{r_M} \lambda_N - 1 \right) \right] \end{pmatrix} \right\| \|Q^T\| \\ &= \max_i \left[\sum_{k=P}^L c_k T_k \left(\frac{2}{r_M} \lambda_i - 1 \right) \right] \end{aligned} \quad (14)$$

となる. ここで, 直行行列のノルムが 1 であること, 対角行列のノルムは対角成分の最大値であることを用いた. さらに, Chebyshev 多項式のすべての $T_k()$ は ±1 の間の値を取ることから

$$\|S_L - S_P\| \leq d, \quad d = \sum_{k=P}^L |c_k| \quad (15)$$

が得られる. したがって, 項数 P を決定するためには, 非常に大きな項数 L に対して Chebyshev 多項式の係数を求め, P から L までの係数の絶対値の和が充分に小さくなるように P を選べばよい.

(4) 具体的な計算機アルゴリズム

Chebyshev 多項式は式(10)に示す関係があるため

$$\begin{aligned} T_{k+1}(R')w &= 2R'T_k(R')w - T_{k-1}(R')w, \quad k \geq 1 \end{aligned} \quad (16)$$

が成り立つ. この関係式を逐次用いることにより, 具体的な計算機アルゴリズムは以下のようになる.

a	b	c	d	e	f	g	h	i
b	a	b	e	d	e	h	g	h
c	b	a	f	e	d	i	h	g
d	e	f	a	b	c	d	e	f
e	d	e	b	a	b	e	d	e
f	e	d	c	b	a	f	e	d
g	h	i	d	e	f	a	b	c
h	g	h	e	d	e	b	a	b
i	h	g	f	e	d	c	b	a

図-1 3×3 の矩形領域を考えた場合の等方性を持つ確率場に対する共分散行列の形

a	b	c	d	g	h	i	l	m
b	a	b	e	d	g	j	i	l
c	b	a	f	e	d	k	j	i
d	e	f	a	b	c	d	g	h
g	d	e	b	a	b	e	d	g
h	g	d	c	b	a	f	e	d
i	j	k	d	e	f	a	b	c
l	i	j	g	d	e	b	a	b
m	l	i	h	g	d	c	b	a

図-2 3×3 の矩形領域を考えた場合の異方性を持つ確率場に対する共分散行列の形

- 1) \sqrt{x} を多項式近似する区間を決定する (式(6)).
- 2) 係数 c_k , $k = 0, \dots, P - 1$ を決定する (式(9)).
- 3) $w \sim N(0, I)$ を発生させる.
- 4) 式(13)に w を乗じて正規確率場を発生させる. この計算には、式(13)において $k = 0$ のとき $T_0(\mathbf{R}') = I$, $k = 1$ のとき $T_1(\mathbf{R}') = \mathbf{R}'$, $k \geq 2$ のときには式(16)の関係を逐次用いる. 具体的な計算手順は以下の通り.

```

# k = 0 のとき
z0 = w; rf = c0z0 - c0/2 w
# k = 1 のとき
z1 = R'w; rf = rf + c1z1
# 2 ≤ k < P のとき
for (k = 2; k < P; k++) {
    if (k is even number) {
        z0 = 2R'z1 - z0; rf = rf + ckz0
    }
    else {
        z1 = 2R'z0 - z1; rf = rf + ckz1
    }
}
rf = rf + m

```

z_0 , z_1 は N 次元の作業用ベクトルを意味し、 N 次元ベクトル r_f に最終的に近似的に求めた正規確率場のシミュレーション値が格納される. m は平均値ベクトルである. N は発生させる確率変数の個数であり、=記号はベクトルの代入の意味に用いた.

このアルゴリズムを実行するために必要となる計算機メモリは、 N 次元ベクトル z_0 , z_1 , r_f と $N \times N$ 次元の行列 R' である. 共分散行列を関数形で与えるならば R' の各要素の値は記憶しておく必要がないため、最低 $3N$ 個の計算機メモリがあればよい. たとえば、 100×100 の 2 次元領域の確率場を発生させる場合、 $N = 100 \times 100$ であるため、10,000 個の記憶領域を持つ一次元配列を 3 つ用意すればよい. ただし、 R' の値が参照されるたびにその値を計算すると、行列のサイズが大きい場合には非常に計算時間がかかる. そこで、

Davis が示した方法をもとに、少ないメモリで共分散行列を記憶する方法を採用する.

今、 3×3 の矩形領域を考える. 各グリッドは同じ大きさの矩形であり、各グリッド点間の確率変数の空間相関は距離のみによって決まる等方性を持つ確率場を考える. この場合、共分散行列は図-1 のような形になる. 図-1 からわかるように、この場合、共分散行列は 3 種類の 3×3 の小行列から構成されており、各小行列は斜めの成分の値が等しい対称行列となっている. この性質を利用すれば、小行列を表現するために必要なメモリは 3 個であり、小行列は 3 種類あるので 9 個のメモリでこの共分散行列を表現することができる. 一般に、同じサイズの矩形で領域を覆い、距離のみによって変数間の相関の値を設定するならば、 $N_x \times N_y$ のグリッドを設定した場合、その共分散行列を表現するために必要なメモリの個数は $N_x N_y$ 個でよい. この特性を利用すれば、作業用に用いる $3N_x N_y$ 個のメモリと合わせてこのアルゴリズムでは $4N_x N_y$ 個のメモリを用いて正規確率場を発生させることができる.

次に、各グリッド点間の確率変数の空間相関が距離と方向によって決まる異方性を持つ確率場を考える. この場合、等方性の場合と同様に 3×3 の矩形領域を考えると、共分散行列は図-2 のような形になる. この場合、等方性を持つ場合と同様に、共分散行列は 3 種類の 3×3 の小行列から構成され、各小行列は斜めの成分の値が等しい行列となるが、対角部分以外にある小行列は対称行列とはならない. そのために、共分散行列を表現するために必要なメモリの個数は、等方性を持つ場合よりも多くなる. 同じサイズの矩形で $N_x \times N_y$ のグリッドで領域を覆った場合、共分散行列を表現するために必要なメモリの個数は、対角部分の小行列について N_x 個、対角部分以外の小行列については $2N_x - 1$ 個であり対角部分以外にある小行列は $N_y - 1$ 種類があるので、合計 $2N_x N_y - N_x - N_y + 1$ 個となる. したがって、作業用に用いる $3N_x N_y$ 個のメモリと合わせて $5N_x N_y$ 個のメモリを取ることができれば異方性をもつ正規確率場を発生させることになる.

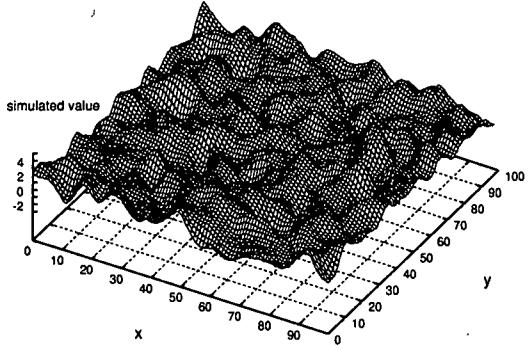


図-3 100×100 のグリッドに対し Gauss 共分散関数を与えて発生させた正規確率場の例。

4. 発生させた確率場の例

(1) 正規確率場

格子間隔が x 方向, y 方向ともに 1.0 であるような 100×100 の矩形領域を設定し, 正規分布に従うように発生させた確率場の一例を 図-3 に示す。この例では, 共分散関数 $C(h)$ を Gauss 関数

$$C(h) = C(0) \exp\left(-\frac{h^2}{a^2}\right)$$

で与えた。 h は距離, $C(0)$ は分散の値, a は相関長さである。ここでは, $C(0) = 1.0$, $a = 5.0$, 発生させる確率場の平均値 $m = 1.0$ とした。また, 式(15)を用いて Chebyshev 多項式の項数の違いによる誤差を見積もったところ, $L = 1000$, $P = 30$ のとき $d = 0.091$, $P = 50$ のとき $d = 0.053$, $P = 100$ のとき $d = 0.024$ であった。そこで, d が 5% 程度のところを取って $P = 50$ とした。

図-4 は 図-3 で示した確率場のシミュレーション値の領域内のヒストグラム, 図-5 は 図-3 で示した確率場のシミュレーション値の空間相關を計算した結果である。図-5 の作成においては, 距離 h を 0.2 ごとの区間に区切り, その区間に相関係数を計算した結果をプロットした。たとえば $9.8 < h \leq 10.0$ の区間で相関係数を計算し, $h = 9.9$ のところにその値をプロットした。このような計算方法を取ったのは, 特定の方向のみでなく, 多方向のサンプル点を含めて空間相関値を得るためにある。発生させた確率場の統計的な特性は理論値とよく一致している。計算時間は DEC Alpha 21164 (500MHz) を搭載した計算機 (OS, Windows NT) を用いた場合, 約 10 分であった。

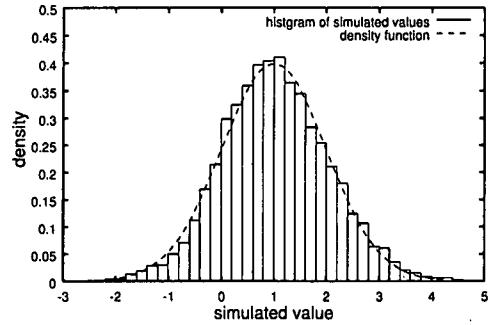


図-4 発生させた正規確率場のヒストグラムと正規分布に従う確率密度関数との比較。

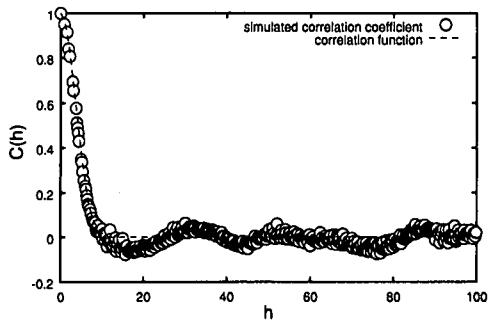


図-5 発生させた正規確率場の空間相関係数と Gauss 相関係数との比較。

(2) 対数正規確率場

次に, 対数正規分布に従う確率場を発生させることを考える。今, 正規分布に従う二つの確率変数を

$$X_1 \sim N(m_{X_1}, \sigma_{X_1}^2), \quad X_2 \sim N(m_{X_2}, \sigma_{X_2}^2)$$

とし, $Y_1 = e^{X_1}$, $Y_2 = e^{X_2}$ の共分散 C_{Y_1, Y_2} を考える。

$$m_{Y_i} = \exp\left(m_{X_i} + \frac{\sigma_{X_i}^2}{2}\right), \quad i = 1, 2 \quad (17)$$

$$\sigma_{Y_i}^2 = m_{Y_i}^2 \{\exp(\sigma_{X_i}^2) - 1\}, \quad i = 1, 2 \quad (18)$$

という関係式と, X_1 と X_2 との共分散を C_{X_1, X_2} として

$$X_1 + X_2$$

$$\sim N(m_{X_1} + m_{X_2}, \sigma_{X_1}^2 + \sigma_{X_2}^2 + 2C_{X_1, X_2})$$

であることを用いると

$$C_{Y_1, Y_2}$$

$$= E[Y_1 Y_2] - E[Y_1]E[Y_2]$$

$$= \exp\{m_{X_1} + m_{X_2} + \frac{1}{2}(\sigma_{X_1}^2 + \sigma_{X_2}^2 + 2C_{X_1, X_2})\}$$

$$- m_{Y_1} m_{Y_2}$$

$$= m_{Y_1} m_{Y_2} \{\exp(C_{X_1, X_2}) - 1\}$$

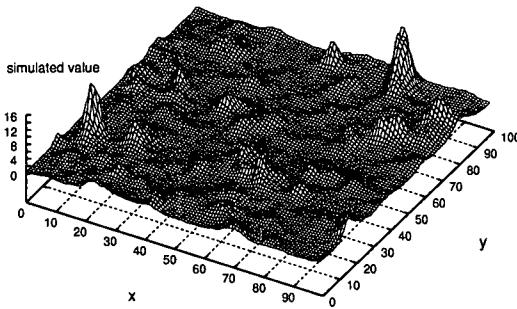


図-6 100×100 のグリッドに対し Gauss 共分散関数を与えて発生させた対数正規確率場の例。

となる。ここで、 X は共分散の値が距離 h のみによって決まる等方性を持つ確率場であるとする。 X の共分散関数を $C_X(h)$ とすると、 Y の共分散関数 $C_Y(h)$ は上式より

$$C_Y(h) = m_Y^2 \{ \exp(C_X(h)) - 1 \} \quad (19)$$

となる。以上をもとに、 m_Y , σ_Y^2 , $C_Y(h)$ を与えて等方性を持つ対数正規分布に従う確率場を発生させる手順は次のようなになる。

- 1) 式(18)より σ_X^2 を求める。
- 2) 式(17)より m_X を求める。
- 3) 式(19)より $C_X(h)$ を決定する。
- 4) m_X , σ_X^2 , $C_X(h)$ に従う正規確率場 X を発生させる。
- 5) $Y = e^X$ によって発生させた X を Y に変換する。

図-6 はこの手順により発生させた対数正規確率場の例である。共分散関数 $C_Y(h)$ は Gauss 関数

$$C_Y(h) = C_Y(0) \exp\left(-\frac{h^2}{a^2}\right)$$

で与えた。この場合、式(19)より

$$C_X(h) = \ln \left\{ \frac{C_Y(0)}{m_Y^2} \exp\left(-\frac{h^2}{a^2}\right) + 1 \right\}$$

となる。ここでは、 $C_Y(0) = 1.0$, $a = 5.0$, $m_Y = 1.0$ とした。また、式(15)を用いて Chebyshev 多項式の項数の違いによる誤差を見積もったところ、 $L = 1000$, $P = 30$ のとき $d = 0.083$, $P = 50$ のとき $d = 0.047$, $P = 100$ のとき $d = 0.021$ であった。そこで、 d が 5% 程度のところを取って $P = 50$ とした。

図-7 は図-6 で示した確率場のシミュレーション値の領域内でのヒストグラム、図-8 は図-6 で示した確率場のシミュレーション値の空間相関を計算した結果である。図-8 における空間相関値の計算の方法は図-5 の場合と同様である。発生させた確率場の統計的な特性は理論値とよく一致している。

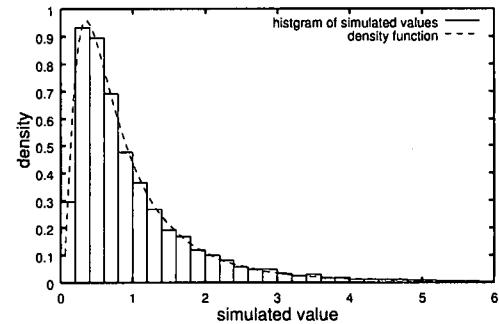


図-7 発生させた対数正規確率場のヒストグラムと対数正規分布に従う確率密度関数との比較。

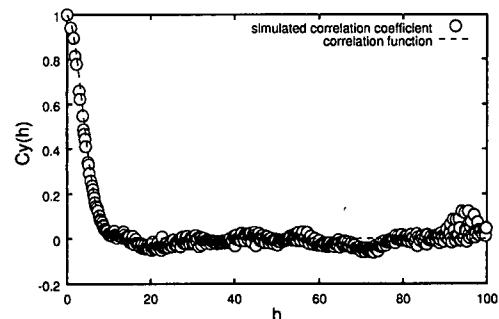


図-8 発生させた対数正規確率場の空間相関係数と Gauss 相関係数との比較。

5. まとめ

分布型流出モデルにおけるモデルパラメータの空間分布、または降水分布が流出シミュレーション結果にどのような影響を与えるかを調査するための基本的なツールとして、正規確率場、対数正規確率場を発生させるアルゴリズム、および計算機プログラムを構築した。この手法は、Davis⁴⁾ によって提案された手法に Chebyshev 多項式近似を導入した方法であり、発生させる確率変数の個数を N とすると、最低 $3N$ 個の計算機メモリを取得することができれば実行が可能となる。計算時間は 100×100 のグリッド領域を対象とした場合、DEC Alpha 21164 (500MHz) を搭載した計算機 (OS: Windows NT) を用いて約 10 分であった。

これにより発生させるモデルパラメータの空間場と分布型流出モデルとを組み合わせて、モデルパラメータの空間分布の仕方が流出にどのような影響を及ぼすかを調査する研究を現在進めている。こうした研究は、分布型流出モデルの分布モデルパラメータを如何に決定するかという実用的に重要な課題に対しても有効な知見を与えるものと思われる。

ここで述べたアルゴリズムを用いて確率場を発生させるプログラムソースコードは <http://fmd.dpri.kyoto-u.ac.jp/~tatikawa/srf.html> から取得することができる。なお、本研究は、筆者が文部省在外研究員としてオクラホマ大学滞在中に行われたものである。

最後に、本論文を査読して頂きました匿名の三名の方からのご意見により、本論が洗練されました。この場を借りてお礼申し上げます。

参考文献

- 1) Bras, R. L., and Rodriguez-Iturbe, I. : Random functions and hydrology, Addison-Wesley, 1985.
- 2) Christakos, G. : Random field models in earth science, Academic Press, 1992.
- 3) Deutsch, C. V. and Journel, A. G. : GSLIB: Geostatistical software library and user's guide, 2nd edition, Oxford University Press, 1998.
- 4) Davis, W.M. : Generating large stochastic simulations - the matrix polynomial approximation method, *Mathematical Geology*, 19(2), pp. 99-107, 1987.
- 5) Bras, R. L. and Rodriguez-Iturbe, I. : Rainfall generation: a nonstationary time-varying multidimensional model, *Water Resources Research*, 12(1), pp. 450-456, 1976.
- 6) Golub, G. H. and Van Loan, C. F. : Matrix computations, 3rd edition, Johns Hopkins University Press, p. 395, 1996.
- 7) Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. : Numerical Recipes in C, 2nd edition, Cambridge University Press, pp. 190-194, 1992.

(1999. 7. 28 受付)

GAUSSIAN AND LOG NORMAL RANDOM FIELD GENERATION BASED ON THE SQUARE ROOT DECOMPOSITION OF A COVARIANCE MATRIX

Yasuto TACHIKAWA and Michiharu SHIIBA

A method to generate a Gaussian and log normal random field is presented. The method is based on a matrix factorization technique of decomposition of a covariance matrix into its square root matrix approximately by using the Chebyshev polynomials. When we generate a Gaussian or lognormal random field with $N_x \times N_y$ points covered by a mesh of a rectangular grid, the algorithm presented here requires only $3N_x N_y$ entries to be stored.