Homogeneity, Isotropy, and Log-normality: Some Thoughts on Kolmogorov's Turbulence Theory

Chunsheng Ma

SAMSI and Wichita State University

Outline

1. The turbulence problem
2. Why is probability and statistics methods involved in this type of non-random systems?
3. Kolmogorov 41 theory
4. Homogeneity (stationarity)
5. Isotropy
6. Log-Gaussian
7. Long range dependence
8. Conclusions

1. The turbulence problem

"*The turbulence problem is an age-old topic of discussion among fluid dynamicists. It is not a problem of physical law; it is a problem of description.*

Turblence is a state of fluid motion, governed by known dynamical laws - the Navier-Stokes equations ..."

P. A. Durbin and B. A. Pettersson Reif (2000),

Statisitical Theory and Modeling for Turbulent Flows. Wiley.

"*The Navier-Stokes equation probably contains all of the turbulence*".

U. Frisch (1995),

Turbulence. Cambridge Univ. Press.

- The Navier-Stokes equation

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla P + \nu \nabla^2 \mathbf{v} + \mathbf{F}$$

$$\nabla \cdot \mathbf{v} = 0$$

where

$P = P(\mathbf{s}, t)$ is the pressure at location $\mathbf{s}$ and time t,

$\mathbf{F}$ an external force,

ν the kinemate

The Navier-Stokes equations are formally deterministic.

2. Why are probability and statistics methods involved in this type of non-random systems?

- No mathematical solution so far.
- Experimental data, measure error $\longrightarrow$ Statistical method
- Similar to that we use Monte Carlo methods to evaluate, for instance,

$$\int_0^1 \frac{\sin x}{x} dx$$

3. Kolmogorov 41 theory

In the statistical theory of turbulence, it is customary to assume that the velocity field of a fluid in turbulent flow may be represented by a random vector field

$$(v_1(\mathbf{s};t), v_2(\mathbf{s};t), v_3(\mathbf{s};t))$$

where

$\mathbf{s} = (s_1, s_2, s_3)$ spatial coordinate

t time

A real-valued random field defined over a space-time domain $\mathcal{S} \times \mathcal{T}$

$$Z(\mathbf{s}; t), \ \mathbf{s} \in \mathcal{S}, \quad t \in \mathcal{T}$$

◦ **Mean function**

$$\mu(s; t) = \mathrm{E}Z(\mathbf{s}; t)$$

◦ **Covariance function**

$$C(\mathbf{s}_1, \mathbf{s}_2; \ t_1, t_2) = \mathrm{E}[\{Z(\mathbf{s}_1; t_1) - \mu(\mathbf{s}_1; t_1)\}\{Z(\mathbf{s}_2; t_2) - \mu(\mathbf{s}_2; t_2)\}]$$

◦ **Variogram**, structure function (variance of the increments)

$$\gamma(\mathbf{s}_1, \mathbf{s}_2; \ t_1, t_2) = \tfrac{1}{2} \operatorname{var}\{Z(\mathbf{s}_1; t_1) - Z(\mathbf{s}_2; t_2)\}, \qquad \mathbf{s}_1, \mathbf{s}_2 \in \mathcal{S}, \ t_1, t_2 \in \mathcal{T}$$

The existence of the covariance function implies that of the variogram, with

$$\gamma(\mathbf{s}_1, \mathbf{s}_2;\ t_1, t_2) = \frac{1}{2}\{C(\mathbf{s}_1, \mathbf{s}_1;\ t_1, t_1) + C(\mathbf{s}_2, \mathbf{s}_2;\ t_2, t_2)\} - C(\mathbf{s}_1, \mathbf{s}_2;\ t_1, t_2),$$

but not vice versa.

A fractional Brownian motion

An intrinsically stationary Gaussian process $\{Z(t), t \in \mathbb{R}\}$ with variogram

$$\gamma(t) = |t|^\alpha, \qquad t \in \mathbb{R},$$

and covariance

$$C(t_1, t_2) = |t_1|^\alpha + |t_2|^\alpha - |t_1 - t_2|^\alpha, \qquad t_1,\ t_2 \in \mathbb{R},$$

where $0 < \alpha \leq 2$.

The Fourier transform of $\gamma(t)$, not $C(t_1, t_2)$,

$$|t|^\alpha \quad \propto \quad \int_0^\infty \omega^{-(\alpha+1)}\{1 - \cos(t\omega)\} d\omega$$

When $\alpha = \frac{2}{3}$, it leads to Kolmogorov $\frac{2}{3}$ law or $-\frac{5}{3}$ law for a temporal margin

4. Homogeneity (stationarity)

Stationary in space and time

if $\mu(\mathbf{s};t)$ is a constant, and

$C(\mathbf{s}_1, \mathbf{s}_2;\ t_1, t_2)$ depends only on $\mathbf{s}_1 - \mathbf{s}_2$ and $t_1 - t_2$

Write $C(\mathbf{s}_1, \mathbf{s}_2;\ t_1, t_2) = C(\mathbf{s}_1 - \mathbf{s}_2;\ t_1 - t_2)$

Stationary in space

if $\mu(\mathbf{s};t)$ does not depend on $\mathbf{s}$, and

$C(\mathbf{s}_1, \mathbf{s}_2;\ t_1, t_2)$ depends only on $\mathbf{s}_1 - \mathbf{s}_2$ and t_1, t_2

Stationary in time

if $\mu(\mathbf{s};t)$ does not depend on t, and

$C(\mathbf{s}_1, \mathbf{s}_2;\ t_1, t_2)$ depends only on $\mathbf{s}_1, \mathbf{s}_2$ and $t_1 - t_2$

Intrinsic stationarity

Intrinsically stationary in space and time

if $\gamma(\mathbf{s}_1, \mathbf{s}_2;\; t_1, t_2)$ depends only on $\mathbf{s}_1 - \mathbf{s}_2$ and $t_1 - t_2$

Write $\gamma(\mathbf{s}_1, \mathbf{s}_2;\; t_1, t_2) = \gamma(\mathbf{s}_1 - \mathbf{s}_2;\; t_1 - t_2)$

Intrinsically stationary in space

if $\gamma(\mathbf{s}_1, \mathbf{s}_2;\; t_1, t_2)$ depends only on $\mathbf{s}_1 - \mathbf{s}_2$ and t_1, t_2

Intrinsically stationary in time

if $\gamma(\mathbf{s}_1, \mathbf{s}_2;\; t_1, t_2)$ depends only on $\mathbf{s}_1, \mathbf{s}_2$ and $t_1 - t_2$

5. Isotropy

Isotropic function: $g(\|\mathbf{s}\|)$,

$$\text{where } \|\mathbf{s}\| = \left(\sum_{k=1}^{d} s_k^2 \right)^{\frac{1}{2}}, \ \mathbf{s} \in \mathbb{R}^d.$$

Geometrically anisotropic function: $g(\|A\mathbf{s}\|)$,

where A is a $d \times d$ matrix.

There are many reasons for the popular use of the isotropic or geometrically anisotropic covariance function and variogram in various areas.

The simplest reason would be just as that the Euclidean distance is the most popularly used distance.

A less known reason is that an isotropic or geometrically anisotropic model would be the only choice in certain circumstances, for instance, when the underlying random field is smooth enough.

A purely spatial version (Ma (2007), PAMS):

Let $C(x), x \in \mathbb{R}$, be an even and twice continuously differentiable function, and let $\gamma(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d$, be a homogeneous function, *i.e.*,

$$\gamma(\alpha \mathbf{s}) = |\alpha| \gamma(\mathbf{s}), \quad \forall \mathbf{s} \in \mathbb{R}^d, \ \ \forall \alpha \in \mathbb{R}.$$

If $C(\gamma(\mathbf{s})), \mathbf{s} \in \mathbb{R}^d$, is a covariance function, then $\gamma(\mathbf{s})$ must be of the form

$$\gamma(\mathbf{s}) = \|A\mathbf{s}\|, \quad \mathbf{s} \in \mathbb{R}^d,$$

where A is a $d \times d$ matrix.

An interpretation:

Let $\{Z(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d\}$ be a random field with covariance $C(\gamma(\mathbf{s}))$. For fixed $s_2^0, \ldots, s_d^0$, the univariate projection $\{Z(s_1, s_2^0, \ldots, s_d^0), s_1 \in \mathbb{R}\}$ has the covariance

$$C(\gamma(s_1, 0, \ldots, 0)) = C(s_1 \gamma(1, 0, \ldots, 0)), \qquad s_1 \in \mathbb{R}.$$

Its derivative exists in the mean square sense,

$$\lim_{h \to 0} \mathrm{E} \left\{ \frac{Z(s_1 + h, s_2^0, \ldots, s_d^0) - Z(s_1, s_2^0, \ldots, s_d^0)}{h} - \frac{\partial}{\partial s_1} Z(s_1, s_2^0, \ldots, s_d^0) \right\}^2 = 0,$$

and possesses the covariance $-C''(s_1 \gamma(1, 0, \ldots, 0)) \gamma^2(1, 0, \ldots, 0)$.

A random field with such a smooth property is nothing but an isotropic (or geometrically anisotropic) one.

Example 1

$$C(x) = \theta \exp(-\alpha_1 x^2) + (1-\theta)\exp(-\alpha_2 x^2), \quad x \in \mathbb{R},$$

where $\alpha_2 > \alpha_1 > 0$.

Since $C''(x)$ exists everywhere on the real line, neither

$$C(\mathbf{s}) = \theta \exp(-\alpha_1 |\mathbf{s}|^2) + (1-\theta)\exp(-\alpha_2 |\mathbf{s}|^2), \quad \mathbf{s} \in \mathbb{R}^d,$$

where $|\mathbf{s}| = \sum_{k=1}^{d} |s_k|$ is the ℓ_1-norm,

nor

$$C(\mathbf{s}) = \theta \exp[-\alpha_1\{\max(|s_1|,|s_2|)\}^2] + (1-\theta)\exp[-\alpha_2\{\max(|s_1|,|s_2|)\}^2], \quad \mathbf{s} \in \mathbb{R}^d,$$

could be a covariance function in $\mathbb{R}^d$.

The isotropic function

$$C(\mathbf{s}) = \theta \exp(-\alpha_1 \|\mathbf{s}\|^2) + (1-\theta)\exp(-\alpha_2 \|\mathbf{s}\|^2), \quad \mathbf{s} \in \mathbb{R}^d,$$

is the covariance function of a Gaussian random field if and only if

$$\left\{1 - \left(\frac{\alpha_2}{\alpha_1}\right)^{\frac{d}{2}}\right\}^{-1} \quad \leq \quad \theta \quad \leq \quad 1.$$

Example 2

$$C(x) = 1 - \{1 - \exp(-\alpha_1|x|)\}\{1 - \exp(-\alpha_2|x|)\}, \quad x \in \mathbb{R},$$

is twice continuously differentiable on $\mathbb{R}$, where $\alpha_1 > 0$, $\alpha_2 > 0$.

When $\gamma(\mathbf{s})$ is homogeneous, the only chance for $C(\gamma(\mathbf{s})), \mathbf{s} \in \mathbb{R}^d$, to be a covariance function is

$$\gamma(\mathbf{s}) = \|A\mathbf{s}\|, \quad \mathbf{s} \in \mathbb{R}^d,$$

where A is a $d \times d$ matrix.

A space-time version (Ma (2007)):

Let an even function $C(x,t), x \in \mathbb{R}, t \in \mathcal{T}$ have continuous second-order derivative with respect to x, where $\mathcal{T} \subset \mathbb{R}$, and

let $\gamma(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d$, be a homogeneous function, *i.e.*,

$$\gamma(\alpha \mathbf{s}) = |\alpha| \gamma(\mathbf{s}), \quad \forall \mathbf{s} \in \mathbb{R}^d, \;\; \forall \alpha \in \mathbb{R}.$$

If $C(\gamma(\mathbf{s}), t), \mathbf{s} \in \mathbb{R}^d, t \in \mathcal{T}$, is a covariance function on $\mathbb{R}^d \times \mathcal{T}$,

then $\gamma(\mathbf{s})$ must be of the form

$$\gamma(\mathbf{s}) = \|A\mathbf{s}\|, \quad \mathbf{s} \in \mathbb{R}^d,$$

where A is a $d \times d$ matrix.

6. Log-Gaussian random field

A positive random field $\{Z(\mathbf{s}), \mathbf{s} \in \mathcal{D}\}$ is said to be a log-Gaussian random field if $\{\ln Z(\mathbf{s}), \mathbf{s} \in \mathcal{D}\}$ is a Gaussian random field.

The finite-dimensional distribution function of $\{Z(\mathbf{s}),\ \mathbf{s} \in \mathcal{D}\}$,

$$
\begin{aligned}
&P(Z(\mathbf{s}_1) \le u_1, \ldots, Z(\mathbf{s}_n) \le u_n) \\
= &\begin{cases} P(\ln Z(\mathbf{s}_1) \le \ln u_1, \ldots, \ln Z(\mathbf{s}_n) \le \ln u_n), & \text{if } \ u_1, \ldots, u_n > 0, \\ 0, & \text{otherwise,} \end{cases}
\end{aligned}
$$

relates closely to a Gaussian distribution function, and a log-Gaussian random field is thus characterized by it mean and covariance functions, just like a Gaussian random field.

But, unlike a Gaussian random field, the mean and covariance function sof a log-Gaussian random field are often tied each other.

(i) If $\{Z(\mathbf{s}), \mathbf{s} \in \mathcal{D}\}$ is a log-Gaussian random field with mean $\mu(\mathbf{s})$ and covariance $C(\mathbf{s}_1, \mathbf{s}_2)$, then $\mu(\mathbf{s})$ is positive, $C(\mathbf{s}_1, \mathbf{s}_2) > -\mu(\mathbf{s}_1)\mu(\mathbf{s}_2)$, and $C(\mathbf{s}_1, \mathbf{s}_2)$ and $\ln\{1 + \mu^{-1}(\mathbf{s}_1)\mu^{-1}(\mathbf{s}_2)C(\mathbf{s}_1, \mathbf{s}_2)\}$ are positive definite on $\mathcal{D}$.

(ii) Conversely, if $\mu(\mathbf{s}), \mathbf{s} \in \mathcal{D}$, is a positive function, and $\ln\{1 + \mu^{-1}(\mathbf{s}_1)\mu^{-1}(\mathbf{s}_2)C(\mathbf{s}_1, \mathbf{s}_2)\}$ is positive definite on $\mathcal{D}$, then there exists a log-Gaussian random field with mean $\mu(\mathbf{s})$ and covariance $C(\mathbf{s}_1, \mathbf{s}_2)$.

The positive definiteness is a necessary but not sufficient condition for a real function to be the covariance function of a log-Gaussian random field.

Example 3

The function

$$C(x_1, x_2) = \min(x_1,\ x_2), \qquad x_1,\ x_2 \geq 0,$$

is known to be the covariance function of the Wiener or Brownian motion process on $[0, \infty)$. It is associated with a log-Gaussian stochastic process on $[0, \infty)$ whose mean could be an arbitrary positive constant, since

$$\ln\{1 + \mu^{-2} C(x_1, x_2)\} = \min\{\ln(1 + \mu^{-2}\sigma^2 x_1),\ \ln(1 + \mu^{-2}\sigma^2 x_2)\}, \qquad x_1,\ x_2 \geq 0,$$

is positive definite on $[0, \infty)$ for every positive constant μ.

Example 4

The function

$$C(x) = \begin{cases} 1, & x = 0, \\ \theta, & x = \pm 1, \\ 0, & x \neq 0, \pm 1, \quad x \in \mathbb{Z}, \end{cases}$$

is the covariance function of a stationary first-order moving average Gaussian process on $\mathbb{Z}$ if and only if $|\theta| \leq \frac{1}{2}$.

It is the covariance function of a stationary log-Gaussian process with a positive mean μ if and only if

$$\{(1 + \mu^{-2})^{-\frac{1}{2}} - 1\}\mu^2 \leq \theta \leq \{(1 + \mu^{-2})^{\frac{1}{2}} - 1\}\mu^2,$$

in such a way μ is tied up with θ or the covariance $C(x)$. The above domain of θ is tighter than $|\theta| \leq \frac{1}{2}$ since

$$-\frac{1}{2} < \{(1 + \mu^{-2})^{-\frac{1}{2}} - 1\}\mu^2, \quad \text{and} \quad \{(1 + \mu^{-2})^{\frac{1}{2}} - 1\}\mu^2 < \frac{1}{2}.$$

Example 5 (Matheron (1989))

The function

$$C(\mathbf{s}) = \exp(-\|\mathbf{s}\|), \qquad \mathbf{s} \in \mathbb{R}^d,$$

is positive definite.

When $d = 1$, it is the covariance function of a stationary log-Gaussian process with an arbitrary mean $\mu > 0$.

When $d \geq 2$, it is the covariance function of a stationary log-Gaussian process with mean $\mu \geq \mu_0$, where the threshold μ_0 is a positive constant, but $\mu_0 =?$.

A conjecture:

Assume that a positive function $C(\mathbf{s}_1, \mathbf{s}_2)$ on $\mathcal{D}$ is the covariance function associated with a log-Gaussian random field whose mean is μ_0. Then, for any positive constant $\mu \geq \mu_0$, $C(\mathbf{s}_1, \mathbf{s}_2)$ is also the covariance function on $\mathcal{D}$ associated with a log-Gaussian random field whose mean is μ.

Mandelbrot, B. B. (1972).

Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence. In: *Statistical Models and Turbulence*, pages 333-351, Edited by J. Ehlers, et al.. Lecture Notes in Physics, #12. Springer.

Too smooth ?

Example 6 (Matheron (1989))

The function

$$C(\mathbf{s}) = \exp(-\|\mathbf{s}\|^2), \qquad \mathbf{s} \in \mathbb{R}^d,$$

is positive definite, and infinitely differentiable in $\mathbb{R}^d$.

But, it is NOT associated with any log-Gaussian random field.

Long-range dependence

Assume that $\gamma(\mathbf{s}_1, \mathbf{s}_2)$ is a variogram for a Gaussian random field on $\mathcal{D}$.

(i) If κ is a positive constant with $0 < \kappa \leq 1$, then each of the following functions is the covariance function for a log-Gaussian random field on $\mathcal{D}$ with any positive constant mean

$$C(\mathbf{s}_1, \mathbf{s}_2) = \{1 + \gamma(\mathbf{s}_1, \mathbf{s}_2)\}^{-\kappa}, \qquad \mathbf{s}_1, \mathbf{s}_2 \in \mathcal{D}, \tag{1}$$

$$C(\mathbf{s}_1, \mathbf{s}_2) = \{1 + \ln(1 + \gamma(\mathbf{s}_1, \mathbf{s}_2))\}^{-\kappa}, \qquad \mathbf{s}_1, \mathbf{s}_2 \in \mathcal{D}, \tag{2}$$

$$C(\mathbf{s}_1, \mathbf{s}_2) = \left\{\frac{\alpha_1 + \alpha_2\gamma(\mathbf{s}_1, \mathbf{s}_2)}{1 + \alpha_1 + \alpha_2\gamma(\mathbf{s}_1, \mathbf{s}_2)}\right\}^{\kappa}, \qquad \mathbf{s}_1, \mathbf{s}_2 \in \mathcal{D}. \tag{3}$$

(ii) If κ is a positive constant, then (1) is a covariance function for a Gaussian random field on $\mathcal{D}$.

The function (1), (2), or (3) is power-law decay or has long-range dependence, for which the reason is based on a known fact that a variogram $\gamma(\mathbf{s}_1, \mathbf{s}_2)$ behaves at most like $\|\mathbf{s}_1 - \mathbf{s}_2\|^2$.

Questions

1. The sum or difference of two independent log-Gaussian random fields may not be a log-Gaussian random field.

Is the sum of two covariance functions of log-Gaussian random fields still a covariance function associated with a log-Gaussian random field?

2. The product or ratio of two independent log-Gaussian random fields is also a log-Gaussian random field.

Is the product of two covariances of log-Gaussian random fields still a covariance function associated with a log-Gaussian random field?

8. Conclusions