Probabilistic Engineering Mechanics 48 (2017) 68–75

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

journal homepage: www.elsevier.com/locate/probengmech

Translation models revisited

M. Grigoriu

Cornell University, Ithaca NY 14853–3501, USA

ARTICLE INFO

Keywords:
Diffusion processes
Gaussian/non-Gaussian random elements
Probability
Random vectors
Stochastic processes
Translation models

ABSTRACT

Translation models have been defined as memoryless mappings of Gaussian elements which match exactly/approximately target marginal distributions/correlations. We extend this class of translation models to include memoryless mappings of non-Gaussian elements. It is shown that quantities of interest inferred from equivalent translation models, i.e., models which share the same marginal distributions and have similar second moments, can differ significantly. It is suggested to construct families of equivalent translation models and select members of these families which are optimal for given quantities of interest.

1. Introduction

Gaussian vectors and random functions have and continue to be used extensively in applications because they are conceptually simple and can be calibrated to relatively small data sets as their laws are completely defined by the first two moments. Moreover, Gaussian models are consistent with physics in some applications, e.g., the seismic ground acceleration process [1], and provide convenient tools to, e.g., construct posterior distributions for unknown parameters [2] and describe two-phase microstructures [3] (Sect. 8.2).

The popularity of the Gaussian models also relates to constraints of most available non-Gaussian models which limit their use in applications [4]. Translation models introduced in [5] constitute a notable exception. Like Gaussian models, they are conceptually simple and computationally efficient. Moreover, they can capture a broad range of features of non-Gaussian vectors and random functions and have been proved useful in a broad range of applications [6–10]. Yet, translation models have limitations which are largely caused by the crude manner in which they capture the dependence between target random elements, i.e., they characterize dependence by correlation. This approach can result in unsatisfactory approximations for some quantities of interest.

Our objectives are to (1) extend the class of translation models in [5] to include memoryless transformations of non-Gaussian elements and (2) compare the performance of translation models defined as images of Gaussian and non-Gaussian elements, referred to as *Gauss-translation* and *NGauss-translation* random vectors and functions.

It is shown that families of equivalent translation models, i.e., models which share the same marginal distributions and have similar second moments, can be constructed for target random elements and that quantities of interest inferred from equivalent translation models can differ significantly. This suggests to describe target random elements by families of translation models and select optimal members of these families for given quantities of interest. Theoretical arguments on Gauss- and NGauss-translation models are illustrated by numerical examples which include random vectors and stochastic processes. Several quantities of interest are used to assess the performance of Gauss- and NGauss-translation models, e.g., extremes and temporal averages for stochastic processes.

2. Translation vectors

Translation models for non-Gaussian vectors have been defined as memoryless mappings of Gaussian vectors and have two properties. They match exactly and, generally, approximately target marginal distributions and correlations, respectively. Let X be a non-Gaussian $\mathbb{R}^d$-valued random variable with marginal distributions $\{F_{0,i}\}$, finite second moments, and scaled covariance matrix $\xi_{ij} = E\left[\left(X_i - E[X_i]\right)\left(X_j - E[X_j]\right)\right]/\sqrt{\mathrm{Std}[X_i]\,\mathrm{Std}[X_j]}$, $i, j = 1, \ldots, d$. The components of the Gauss-translation model $\tilde{X}$ of X are defined by

$$\tilde{X}_i = F_{0,i}^{-1} \circ \left(\Phi(G_i)\right) =: h_i(G_i), \quad i = 1, \ldots, d, \tag{1}$$

where Φ denotes the distribution of the standard Gaussian variable $N(0,1)$ and $\{G_i\}$ are $N(0,1)$ variable with correlations $\{\rho_{ij} = E[G_i G_j]\}$. Properties of this class of translation models including their existence are discussed in, e.g., [11] (Chap. 3) and [12]. We only note that the covariances of $\tilde{X}$ and of the Gaussian vector $G = (G_1, \ldots, G_d)$ are such that $|\xi_{ij}| \le |\rho_{ij}|$, $\xi_{ij} = 1$ for $\rho_{ij} = 1$, and $\xi_{ij} \ge -1$ for $\rho_{ij} = -1$. The latter property shows that translation models may not be able to

E-mail address: mdg12@cornell.edu.

http://dx.doi.org/10.1016/j.probengmech.2017.05.002
Received 9 March 2017; Accepted 17 May 2017
Available online 26 May 2017

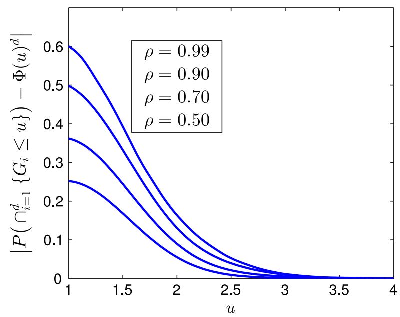

Fig. 1. Difference $\left|P\left(\cap_{i=1}^d \{G_i \le u\}\right) - \Phi(u)^d\right|$ for $\rho = 0.5$, 0.7, 0.9, and 0.99.

capture strong negative correlations between components of target non-Gaussian vectors.

Conceptual simplicity and computational efficiency are the main features of Gauss-translation models. Their implementation only requires to calculate optimal correlations for the Gaussian image G of the target vector X. Generally, these correlation are not far from the target correlations [11] (Sect. 3.1). This observation can be used to initiate optimization algorithms for finding optimal correlations in the Gaussian space. Once the correlations in the Gaussian space have been obtained, the Gauss-translation model $\tilde{X}$ is completely defined. Its samples can be obtained from samples of the Gaussian vector G via the mappings in Eq. (1). We also note that Gauss-translation models are uniquely defined for given metric of the discrepancy between the correlations of X and $\tilde{X}$.

Following are properties of Gauss- and NGauss-translation vectors which are relevant for applications and numerical examples which illustrate similarities and differences between translation models.

Property 1. *Simultaneously large values of components of Gaussian vectors are independent.*

Proof. It can be shown that [13] (Corollary 4.2.4)

$$\begin{aligned}&\left|P\left(G_1 \le u, \dots, G_d \le u\right) - \Phi(u)^d\right|\\&\le c \sum_{1\le i<j\le d} |\rho_{ij}| \exp\left(-\frac{u^2}{1+|\rho_{ij}|}\right),\end{aligned} \tag{2}$$

where $\{G_i\}$ are standard Gaussian variables with correlations $\{\rho_{ij} = E[G_i G_j]\}$ and $c > 0$ is a constant which depends on $\max\{|\rho_{ij}|\}$ assumed to be strictly smaller than unity. This inequality implies $\left|P\left(G_1 \le u, \dots, G_d \le u\right) - \Phi(u)^d\right| \to 0$, $u \to \infty$, i.e., the approximation $P\left(G_1 \le u, \dots, G_d \le u\right) \simeq \Phi(u)^d$ holds for large thresholds u, so that large samples of Gaussian variables are independent regardless their correlation, provided it is not perfect. ▲

For example, suppose $G = (G_1, \dots, G_d)$ is a standard Gaussian vector with equally correlated components, i.e., $G_i = \sqrt{\rho}\,N + \sqrt{1-\rho}\,N_i$, $i = 1, \dots, d$, $\rho > 0$, where N, N_i are independent $N(0,1)$ so that

$$\begin{aligned}P\left(\cap_{i=1}^d \{G_i \le u\}\right) &= E\left[P\left(\cap_{i=1}^d \{N_i \le (u - \sqrt{\rho}\,N)/\sqrt{1-\rho}\}\right)\right]\\&= E\left[\Phi\left((u - \sqrt{\rho}\,N)/\sqrt{1-\rho}\right)^d\right]\end{aligned}$$

which is $\Phi(u)^d$ for $\rho = 0$. Fig. 1 plots the difference $\left|P\left(G_1 \le u, \dots, G_d \le u\right) - \Phi(u)^d\right|$ of probabilities against the threshold u for $d = 10$ and several values of ρ. The difference decreases with u for all values of ρ in agreement with Property 1.

Property 2. *Simultaneously large values of components of Gauss-translation vectors are independent.*

Proof. The random variables $\{G_i\}$, $i = 1, \dots, d$, in Eq. (1) are standard Gaussian but, generally, the vector $G = (G_1, \dots, G_d)$ is not Gaussian. To define $\tilde{X}$, we need to select a covariance matrix $\rho = \{\rho_{ij}\}$ for G which is optimal in the sense that the covariance matrix of $\tilde{X}$ is as close as possible to the target covariance matrix in some metric. Since the events $\{G_i \le u\}$, $i = 1, \dots, d$, are asymptotically ($u \to \infty$) independent, so are their images provided the mappings $\{G_i \mapsto h_i(G_i)\}$ are monotonically increasing, i.e.,

$$P\left(\cap_{i=1}^d \{\tilde{X}_i \le x_i\}\right) = P\left(\cap_{i=1}^d \{G_i \le u_i\}\right),$$

where $u_i = \Phi^{-1} \circ F_i(x_i)$, $i = 1, \dots, d$. The asymptotic independence of the right tails of the components $\{\tilde{X}_i\}$ of translation vectors follows from the normal comparison lemma [13] (Theorem 4.2.1) which shows $\left|P\left(\cap_{i=1}^d \{G_i \le u_i\}\right) - \prod_{i=1}^d \Phi(u_i)\right| \to 0$, $u_i \to \infty$. ▲

Property 3. *Joint distributions of Gauss-translation and target vectors $\tilde{X}$ can be inconsistent in the sense that, if $F_{0,1} = F_{0,2}$, we have $P\left(\tilde{X}_1 \le x_1, \tilde{X}_2 \le x_2\right) = P\left(\tilde{X}_1 \le x_2, \tilde{X}_2 \le x_1\right)$ while the target vector X may not have this property.*

Proof. Consider the special case of a target random vector X with dimension $d = 2$ so that the components of its Gauss-translation model in Eq. (1) are $\tilde{X}_i = F_{0,1}^{-1} \circ \Phi(G_i) = h(G_i)$, $i = 1, 2$, where $G_i \sim N(0,1)$ with correlation $\rho = E[G_1 G_2]$. The joint distribution of $\tilde{X}$ has the expression

$$\begin{aligned}P\left(\tilde{X}_1 \le x_1, \tilde{X}_2 \le x_2\right) &= P\left(G_1 \le h^{-1}(x_1), G_2 \le h^{-1}(x_2)\right)\\&= \Phi\left(h^{-1}(x_1), h^{-1}(x_2); \rho\right),\end{aligned}$$

where $\Phi(\cdot,\cdot;\rho)$ denotes the joint distribution of $G = (G_1, G_2)$. Since $\Phi(u, v; \rho) = \Phi(v, u; \rho)$, $u, v \in \mathbb{R}$, we have $P\left(\tilde{X}_1 \le x_1, \tilde{X}_2 \le x_2\right) = P\left(\tilde{X}_1 \le x_2, \tilde{X}_2 \le x_1\right)$. However, target random vectors may or may not have this property, e.g., the Type C bivariate extreme-value distribution $F(x_1, x_2) = \exp\left[-\max\{\exp(-x) + (1-\phi)\exp(-y), \exp(-y)\}\right]$ has marginal distributions $F_i(x) = \exp\left(-\exp(-x)\right)$, $i = 1, 2$, but is not symmetric [4] (Sect. 2.3). ▲

We also note that Gaussian vectors are distribution-isotropic in the sense that projections of Gaussian vectors on arbitrary directions are Gaussian as linear forms of their components while non-Gaussian vectors do not have this property. For example, suppose X is two-dimensional with components $X_i = N_i^2$, $i = 1, 2$, where $N_i \sim N(0,1)$ with correlation $\rho_N = 0.25$. The solid and dash lines in the left panel of Fig. 2 show that the skewness and kurtosis coefficients of $X(\theta) = X_1 \cos(\theta) + X_2 \sin(\theta)$ depends strongly on $\theta \in [0, \pi/2]$. In contrast, the skewness and kurtosis of the Gaussian image of the translation model of X are invariant with θ and equal to 0 and 3. The middle and right panels in the figure show histograms of $X(\theta)$ for $\theta = 0$ and $\pi/4$. They are consistent with the estimates of skewness and kurtosis in the left panel.

Consider now an extension of Gauss-translation models defined as memoryless mappings of non-Gaussian vectors. Let Y be an $\mathbb{R}^d$-valued random vector with joint distribution F and marginal distributions $\{F_i\}$, $i = 1, \dots, d$, and set

$$\tilde{X}_i = F_{0,i}^{-1} \circ \left(F_i(Y_i)\right) =: h_i(Y_i), \quad i = 1, \dots, d, \tag{3}$$

where $\{Y_i\}$ are the components of Y. We refer to the random vector $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}_d)$ as NGauss-translation model since its image is a non-Gaussian vector. As for the Gauss-translation models, the marginal distributions of $\tilde{X}$ in Eq. (3) match exactly target marginal distributions. The dependence between the components of $\tilde{X}$ is defined by that between the components of Y and the mappings $\{Y_i \mapsto \tilde{X}_i\}$.

The construction of NGauss-translation models requires knowledge of the joint distribution of Y. Although the class of non-Gaussian vectors with known marginal distributions is rather limited [4], it can be used to enrich the class of Gauss-translation models. Let $\mathcal{M}$ denote a family of random vectors Y with known joint distributions, which includes Gaussian vectors. Suppose the members of $\mathcal{M}$ have been mapped into translation models for a target random vector X. We select the model which best describes a particular quantity of interest derived from X. The optimal models may or may not be Gauss-translation vectors.