Probabilistic Engineering Mechanics 24 (2009) 312–320

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

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

An innovating analysis of the Nataf transformation from the copula viewpoint

Régis Lebrun [a], Anne Dutfoy [b,*]

[a] EADS Innovation Works, System Engineering - Applied Mathematics, 12, rue Pasteur BP76, 92152 Suresnes Cedex, France

[b] EDF R&D, Industrial Risk Management, 1, avenue du Général de Gaulle, 92140 Clamart, France

ARTICLE INFO

Article history:
Received 2 July 2007
Received in revised form
4 August 2008
Accepted 4 August 2008
Available online 13 August 2008

Keywords:
The Nataf transformation
Copula
Stochastic modelling

ABSTRACT

This article gives new insight on the Nataf transformation, a widely used tool in reliability analysis. After recalling some basics concerning the copula theory, we explain this transformation in the light of the copula theory and we uncover all the hidden hypothesis made on the dependence structure of the probabilistic model when using this transformation.

Some important results concerning dependence modelling are given, such as the risk related to the use of a linear correlation matrix to describe the dependence structure, and the importance of tail dependence in probabilistic modelling for safety assessment. This contribution should allow the reader to be much more aware of the pitfalls in dependence modelling when relying solely on the Nataf transformation.

1. Introduction

The numerical study of a physical system requires the simulation of a set of equations which model its behaviour. These simulations are mainly computer intensive numerical simulations. We are interested, then, in taking a decision on the basis of a criterion evaluated from some characteristic variables, depending on the values of the input data of the model.

Studies which treat uncertainties aim at evaluating the influence of the uncertainties related to the input data on the characteristic variables of the system. In the framework of probabilistic studies, input data are modelled with probabilistic distributions and propagated through the model to compute the distribution of the characteristic variables, which become a random vector associated to a probability density function.

When the criterion is that one particular characteristic value exceeds a given threshold, the problem of evaluating the probability p of this threshold exceedance can be exposed as follows: let $\underline{X} = (X_1, \dots, X_n)$ be the probabilistic input vector of the n uncertain input data of the model, f its joint probability density function, g: $\mathbb{R}^n \rightarrow \mathbb{R}$ the numerical model (also called the *limit state function*), $Y = g(\underline{X})$ the characteristic variable of interest, and s the given threshold. Then, $p = \mathbb{P}(Y > s) = \int_{\mathcal{D}_s} f(\underline{x})\mathrm{d}\underline{x}$, where $\mathcal{D}_s = \{\underline{X} \in \mathbb{R}^n / g(\underline{X}) > s\}$ is called the *failure domain*.

In the reliability context, authors such as in [3] mention two main difficulties: g and the boundary of $\mathcal{D}_s$ are not analytical expressions but are typically given by a finite element model often requiring high CPU costs, and f is unknown. The first point prevents the use of classical numerical methods to evaluate integrals (Monte Carlo simulations, ...), whereas the second point raises the problem of modelling a joint probability distribution based only on information often reduced to the marginal distributions of $\underline{X}$ and some linear correlation coefficients when one wants to take into account some dependence between the input parameters.

That is why authors recommend the use of the Nataf isoprobabilistic transformation (see [1,10]) to map the *physical space* of the probabilistic input data into the *standard space*, where all the variables are independent and follow the same normal distribution with zero mean and unit variance. Then, within the standard space, we make a first-order or second-order geometrical approximation of the boundary of the failure domain, which allows us to compute an approximation of p thanks to an analytic expression.

This method, widely used in probabilistic uncertainty propagation studies, makes several hypotheses which might not be fulfilled (in that case, the Nataf transformation is not mathematically defined), or which might not agree with the real probability distribution of $\underline{X}$ (in that case, the approximation might be largely wrong).

In this article, we rewrite the Nataf transformation thanks to the copula theory. This innovating point of view explicates the hypotheses of the Nataf transformation in term of probabilistic modelling, which makes it possible to understand plainly the limitations of its use.

In the first part of the article, we detail the Nataf transformation in its usual presentation (as found in e.g. [6]) and how it is usually

* Corresponding author. Tel.: +33 1 47 65 47 27; fax: +33 1 47 65 51 73.
E-mail addresses: regis.lebrun@eads.net (R. Lebrun), anne.dutfoy@edf.fr (A. Dutfoy).

0266-8920/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.probengmech.2008.08.001

used in probabilistic propagation of uncertainties (see e.g. [2]). We recall the interest of such a transformation and the related probabilistic indicators obtained as by-products.

The second part introduces the concept of copula and elliptical distribution and recalls their principal properties. We focus on the key results that will be used in the demonstrations of the following parts.

The concept of copula gives a new insight on the isoprobabilistic Nataf transformation and its hypothesis: in the third part, we demonstrate that the Nataf transformation makes the important hypotheses of a normal dependence structure for the random input vector $\underline{X}$ and maps it into a gaussian vector with independent, zero mean and unit variance components.

Finally, we list all the hypotheses underlying the Nataf transformation and the possible risks associated with its use. In particular, we explain the probabilistic consequences of using a normal dependence structure and the difficulties related to its parameterization with a linear correlation matrix.

2. Traditional use of the Nataf transformation

Very often, probabilistic data available about the random vector $\underline{X}$ are the marginal distributions (which are supposed here to have finite second-order moments) with cumulative distribution functions F_i and, in the particular case of correlated components, the linear correlation matrix $\underline{\underline{R}} = (r_{ij})_{ij}$, with

$$r_{ij} = E\left[\left(\frac{X_i - \mu_i}{\sigma_i}\right)\left(\frac{X_j - \mu_j}{\sigma_j}\right)\right] \tag{1}$$

where μ_i and σ_i are the mean and standard deviation of X_i.

In order to perform reliability analysis such as the computation of a probability of failure, the probabilistic modelling is completed thanks to the Nataf transformation:

Definition 1 (*The Nataf Transformation, [10]*). Let $\underline{X}$ be a random vector with marginal cumulative distribution functions F_i and a linear correlation matrix $\underline{\underline{R}}$.

The Nataf transformation T is the composition of two functions $T = T_2 \circ T_1$ such that

$$T_1 : \underline{X} \mapsto \underline{Y} = \begin{pmatrix} \Phi^{-1}(F_1(X_1)) \\ \Phi^{-1}(F_2(X_2)) \\ \vdots \\ \Phi^{-1}(F_n(X_n)) \end{pmatrix} \tag{2}$$

and

$$T_2 : \underline{Y} \mapsto \underline{U} = \underline{\underline{\Gamma}}\,\underline{Y}. \tag{3}$$

Here, $\underline{Y}$ is supposed to be a gaussian vector with a correlation matrix $\underline{\underline{R}}_0$ (supposed to be positive definite) and with standard normal marginal distributions $\mathcal{N}(0, 1)$.

The matrix $\underline{\underline{\Gamma}}$ is any square-root of $\underline{\underline{R}}_0^{-1}$ and Φ is the cumulative distribution function of the standard normal distribution. The vector $\underline{U}$ is thus a gaussian vector with the same marginal distributions as $\underline{Y}$ but with independent components.

Remark 2. A common choice for $\underline{\underline{\Gamma}}$ is the Cholesky factor of $\underline{\underline{R}}_0^{-1}$, with some extra precautions if $\underline{\underline{R}}^{-1}$ is ill-conditioned. See [7] for the definition of the Cholesky factor and how to manage the stability issue.

The correlation matrix $\underline{\underline{R}}_0$ is called the *fictive correlation matrix*. In general $\underline{\underline{R}}_0 \neq \underline{\underline{R}}$. Indeed, we have the following relation between $\underline{\underline{R}}$ and $\underline{\underline{R}}_0$:

$$\begin{aligned} r_{ij} &= E\left[\left(\frac{F_i^{-1}(\Phi(Y_i)) - \mu_i}{\sigma_i}\right)\left(\frac{F_j^{-1}(\Phi(Y_j)) - \mu_j}{\sigma_j}\right)\right] \\ &= \frac{1}{\sigma_i\sigma_j}\iint_{\mathbb{R}^2}\left\{(F_i^{-1}(\Phi(y_i)) - \mu_i)(F_j^{-1}(\Phi(y_j)) - \mu_j)\right. \\ &\quad \left.\times\, \varphi_2(y_i, y_j, r_{0ij})\right\}\,\mathrm{d}y_i\mathrm{d}y_j \end{aligned} \tag{4}$$

where φ_2 is the bivariate standard normal probability density function with correlation r_{0ij}:

$$\varphi_2(y_i, y_j) = \frac{1}{2\pi\sqrt{1 - r_{0ij}^2}}\exp\left(-\frac{y_i^2 - 2r_{0ij}y_iy_j + y_j^2}{2(1 - r_{0ij}^2)}\right). \tag{5}$$

Remark 3. The computation of the coefficients r_{0ij} might be difficult for two reasons. The first one is that it involves the resolution of the integral equation (4), which is not guaranteed to have a solution, in particular if r_{ij} is too close to 1 or −1. The second one is that even if each coefficient r_{0ij} can be computed, there is no guarantee that the resulting matrix $\underline{\underline{R}}_0$ will be symmetric definite positive.

The Nataf transformation is said to map the *physical space* where $\underline{X}$ takes its values into the *standard space* where $\underline{U}$ takes its values. The interest of the standard space is that we can rewrite the expression of the probability of failure as

$$\begin{aligned} p = \mathbb{P}(Y > s) &= \int_{\mathcal{D}_s} f(\underline{x})\,\mathrm{d}\underline{x} \\ &= \int_{\mathcal{D}_s^U} \varphi_n(\underline{u})\,\mathrm{d}\underline{u} \end{aligned} \tag{6}$$

where the limit state function g has been transformed by T into $G = g \circ T^{-1}$ and the failure domain $\mathcal{D}_s$ into $\mathcal{D}_s^U = \{\underline{U} \in \mathbb{R}^n / G(\underline{U}) > s\}$, where φ_n is the probability density function of the standard n-dimensional normal distribution:

$$\varphi_n(\underline{u}) = \frac{1}{(2\pi)^{n/2}}\exp\left(-\frac{1}{2}\|\underline{u}\|^2\right). \tag{7}$$

The first expression involves the integral of the unknown function f over a complex domain $\mathcal{D}_s$, whereas the second expression involves the integral of the known function φ_n over the complex domain $\mathcal{D}_s^U$.

The main interest of the Nataf transformation is that φ_n is a rapidly decreasing function of $\|\underline{u}\|$, which leads us to suppose that most of the contribution of $\varphi_n(\underline{u})$ to the integral (6) is concentrated in the vicinity of the point of $\mathcal{D}_s^U$ that is the nearest to the origin of the standard space. This point, called the *design point* and denoted P^*, is located on the hypersphere of minimal radius that is tangent to the boundary of the failure domain. It enables us to make a geometrical simplification of the failure domain $\mathcal{D}_s^U$, by modifying its boundary. The so-called FORM method is obtained by a linearization of this boundary at the design point.

We will not go further in the description of the FORM method and the other various extensions such as the SORM method, as we are mainly focused on the reinterpretation of the Nataf transformation as a tool for modelling stochastic dependence.

3. Introduction to dependence modelling and copula

In this section, we recall some basic results on dependence concepts and on the copula theory, in order to enable the reader to reinterpret the Nataf transformation. For the first point, a much more detailed introduction can be found in [8], whereas [11] gives a detailed introduction to the copula theory and the demonstration of all the results presented in this section.

From elementary textbooks, one knows what the stochastic independence between two random variables is: informally, these variables are said to be independent if any information gathered about one of them gives no information about the other. More formally, they are independent if and only if their joint distribution takes the form of a product of the marginal distributions. So, expressed this way, the modelling of the dependence between random variable is, in the most general case, the determination of the *joint distribution* of these variables.

The modelling of stochastic dependence appears to be the determination of a multidimensional function which looks like a complex mathematical object. Several propositions can be found in the statistical literature to synthesize the dependence between two random variables through the determination of a scalar value associated to the two variables, leading to the concept of *measure of association*: it is a general concept that encompasses more specific ones, such as the measure of concordance and the measure of dependence (see [11]).

In this section, we review the three most usual candidates as a measure of association, namely the linear correlation, Spearman's rho and Kendall's tau, and see whether they are proper measures of association.

Let us start with the definition of the general concept of measure of association:

Definition 4 (*Measure of Association*). A *measure of association* r between the components X_1 and X_2 of a random vector $(X_1, X_2)^t$ is a scalar-valued function of X_1 and X_2 with the following properties:

(1) $-1 \leq r(X_1, X_2) \leq 1$
(2) If X_1 and X_2 are independent, $r(X_1, X_2) = 0$
(3) If g_1 and g_2 are strictly increasing functions, we have:

$$r(X_1, X_2) = r(g_1(X_1), g_2(X_2)). \tag{8}$$

A measure of association is then a normalized scalar which quantifies the way two random variables X_1 and X_2 are linked together, being positively associated if $r > 0$ and negatively associated if $r < 0$. The case $r = 0$ is an indication of a possible independence between X_1 and X_2, but not more. The key point is the invariance of r by any change of scale for X_1 and X_2, even a nonlinear one.

Let us see if the most widely used coefficients for tracking dependence are proper measures of association.

The *linear correlation* is often used as a measure of association. This is mainly because, in the context of gaussian vectors, the part of the joint distribution function which is related to the dependence structure is exactly parameterized by the linear correlation matrix.

Definition 5 (*Linear Correlation*). Let $(X_1, X_2)^t$ be a random vector with finite second moments. The *linear correlation* $\rho(X_1, X_2)$ between X_1 and X_2 is given by

$$\rho = E\left[\left(\frac{X_1 - \mu_1}{\sigma_1}\right)\left(\frac{X_2 - \mu_2}{\sigma_2}\right)\right] \tag{9}$$

where μ_i and σ_i are the mean and standard deviation of X_i.

It is well known that the linear correlation coefficient does *not* fulfil point (3) of the definition of a measure of association: it is readily seen with the relation (9).

The sampling definition of this coefficient is:

Definition 6 (*Linear Correlation, Sampling*). Let $((x_1^k, x_2^k)^t)_k$ be a sample of size N of the random vector $(X_1, X_2)^t$. The sampling linear correlation coefficient $\hat{\rho}(X_1, X_2)$ is given by

$$\hat{\rho}(X_1, X_2) = \frac{N \sum_{k=1}^{N} x_1^k x_2^k - \sum_{k=1}^{N} x_1^k \sum_{k=1}^{N} x_2^k}{\sqrt{N \sum_{k=1}^{N} (x_1^k)^2 - \left(\sum_{k=1}^{N} x_1^k\right)^2} \sqrt{N \sum_{k=1}^{N} (x_2^k)^2 - \left(\sum_{k=1}^{N} x_2^k\right)^2}}. \tag{10}$$

In order to fix the non-invariance of the linear correlation by a nonlinear marginal transformation, the computation of the linear correlation can be made on the *rank* of the variables instead of their values. This leads to the definition of *Spearman's rho*:

Definition 7 (*Spearman's Rho*). Let $(X_1, X_2)^t$ be a random vector with marginal cumulative distribution functions F_1 and F_2. Spearman's rho $\rho_S(X_1, X_2)$ is defined by

$$\rho_S(X_1, X_2) = \rho(F_1(X_1), F_2(X_2)). \tag{11}$$

The only drawback of the linear correlation as a measure of association has been removed, so Spearman's rho is a proper measure of association.

The sampling definition of this coefficient is:

Definition 8 (*Spearman's Rho, Sampling*). Let $((x_1^k, x_2^k)^t)_k$ be a sample of size N of the random vector $(X_1, X_2)^t$. If there is no tie, i.e. $\forall i, j, (i \neq j) \Rightarrow (x_1^i \neq x_1^j \text{ or } x_2^i \neq x_2^j)$, the sampling Spearman's rho $\hat{\rho}_S(X_1, X_2)$ is given by

$$\hat{\rho}_S(X_1, X_2) = 1 - \frac{6 \sum_{k=1}^{N} d_k^2}{N(N^2 - 1)} \tag{12}$$

where $d_k = \text{rank}(x_1^k) - \text{rank}(x_2^k)$; otherwise it is the sampling definition of the linear correlation coefficient applied to the ranks of the sample.

Instead of using a measure based on correlation, one can use another approach based on the notion of *concordance*. If $(X_1', X_2')^t$ is an independent copy of the random vector $(X_1, X_2)^t$, X_1 and X_2 are said to be *positively concordant* or *concordant* for short if $\mathbb{P}\left((X_1 - X_1')(X_2 - X_2') > 0\right) > 1/2$, and *negatively concordant* or *discordant* for short if $\mathbb{P}\left((X_1 - X_1')(X_2 - X_2') < 0\right) > 1/2$. *Kendall's tau* is a measure of association based on this concept:

Definition 9 (*Kendall's Tau*). Let $(X_1, X_2)^t$ be a random vector, and let $(X_1', X_2')^t$ be an independent copy of $(X_1, X_2)^t$. Kendall's tau $\tau(X_1, X_2)$ is the difference between the probability of concordance and the probability of discordance between X_1 and X_2:

$$\tau(X_1, X_2) = \mathbb{P}((X_1 - X_1')(X_2 - X_2') > 0) - \mathbb{P}((X_1 - X_1')(X_2 - X_2') < 0). \tag{13}$$

As the notion of concordance is clearly invariant by strictly increasing marginal transformation, Kendall's tau is a proper measure of association.

The sampling definition of this coefficient is:

Definition 10 (*Kendall's Tau, Sampling*). Let $((x_1^k, x_2^k)^t)_k$ be a sample of size N of the random vector $(X_1, X_2)^t$. The sampling Kendall's tau $\hat{\tau}_S(X_1, X_2)$ is given by

$$\hat{\tau}_S(X_1, X_2) = \frac{4 \sum_{i=1}^{N-1} \sum_{j=1+1}^{N} \left(\mathbf{1}_{(x_1^i - x_1^j)(x_2^i - x_2^j) > 0} - \mathbf{1}_{(x_1^i - x_1^j)(x_2^i - x_2^j) < 0}\right)}{N(N-1)} - 1 \tag{14}$$

where $\mathbf{1}_A$ is equal to 1 if A is true, and zero otherwise.

Despite the simplicity of these measures of association, they are not able to fully describe the dependence structure of a random vector. At best, they can identify that the components are *not* independent, but no more. Going back to the real task, the definition of the joint probability density function of the random vector, we introduce the cornerstone of the *full* modelling of stochastic dependence, namely the concept of *copula*:

Definition 11. A copula is a function C defined on $[0,1]^n$ verifying:

(1) for all $\underline{u} \in [0,1]^n$ with at least one component equal to 0, $C(\underline{u}) = 0$ (C is *grounded*);
(2) C is n-increasing:

$$\sum_{i_1=1}^{2} \cdots \sum_{i_n=1}^{2} (-1)^{i_1+\cdots+i_n} C(u_{1i_1}, \ldots, u_{ni_n}) \geq 0 \tag{15}$$

with $u_{j1} = a_j$ and $u_{j2} = b_j \forall j \in \{1,\ldots,n\}$ and $\underline{a}, \underline{b} \in [0,1]^n$, $\underline{a} \leq \underline{b}$
(3) For all $\underline{u} \in [0,1]^n$ with $u_i = 1 \forall i \in \{1,\ldots,n\}, i \neq k$:

$$C(\underline{u}) = u_k. \tag{16}$$

A copula can be seen as the restriction to $[0,1]^n$ of the cumulative distribution function of a distribution whose support is $[0,1]^n$ and with uniform marginal distributions on $[0,1]$. It is not evident from the definition that this concept is the best suited for the modelling of stochastic dependence. It is made clear thanks to the following theorem:

Theorem 12 (*Sklar, 1959*). *Let F be a cumulative density function of dimension n whose marginal distributions are F_i. There exists a copula C of dimension n such that for $\underline{x} \in \overline{\mathbb{R}}^n$, we have*

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)). \tag{17}$$

If the marginal distributions F_i are continuous, the copula C is unique; otherwise, it is uniquely determined on $\mathrm{Range}(F_1) \times \cdots \times \mathrm{Range}(F_n)$.

In the case of continuous marginal distributions, for all $\underline{u} \in [0,1]^n$, we have

$$C(\underline{u}) = F(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)) \tag{18}$$

and

$$p(\underline{x}) = c(F_1(x_1), \ldots, F_n(x_n)) \prod_{i=1}^{n} p_i(x_i) \tag{19}$$

where p_i is the probability density function of the i-th marginal distribution of $\underline{X}$ and c is defined by

$$c(u_1, \ldots, u_n) = \frac{\partial^n C}{\partial u_1 \ldots \partial u_n}(u_1, \ldots, u_n). \tag{20}$$

From this theorem, the role of copulas as a dependence modelling tool is clear: the value taken by any joint cumulative distribution function is the value taken by a copula, once the effect of the marginal cumulative distribution functions have been taken into account. Conversely, a copula is what remains of a joint cumulative distribution once the action of the marginal cumulative distribution functions has been removed.

From a dependence modelling point of view, the first result of Sklar's theorem explains how to build a full probabilistic model based on a set of 1D marginal distributions and a dependence structure given by a copula. The second result explains how to build a catalogue of reusable dependence structures from a given set of multidimensional distributions.

Some examples of bidimensional copulas are given in Table 1.

In order to apprehend the effect of switching from one copula to another in dependence modelling, and to see that this task is much more involved than the determination of a coefficient of correlation, we draw the joint probability distribution function of several bidimensional distributions with the same standard normal marginal distributions, a linear correlation of 0.8 and different copulas; see Fig. 1.

Table 1
Examples of usual bidimensional copulas

Name	$C(u_1, u_2)$
Independent	$u_1 u_2$
Normal	$\int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{s^2-2\rho st+t^2}{2(1-\rho^2)}\right) \mathrm{d}s\,\mathrm{d}t$
Student	$\int_{-\infty}^{T_\nu^{-1}(u_1)} \int_{-\infty}^{T_\nu^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \left(1+\frac{s^2-2\rho st+t^2}{\nu(1-\rho^2)}\right)^{-(\nu+2)/2} \mathrm{d}s\,\mathrm{d}t$
Frank	$-\frac{1}{\theta}\log\left(1+\frac{(\mathrm{e}^{-\theta u_1}-1)(\mathrm{e}^{-\theta u_2}-1)}{\mathrm{e}^{-\theta}-1}\right)$
Clayton	$\left(u_1^{-\theta}+u_2^{-\theta}-1\right)^{-1/\theta}$
Gumbel	$\exp\left(-\left((-\log(u_1))^\theta+(-\log(u_2))^\theta\right)^{1/\theta}\right)$

In addition to Sklar's theorem, we give the most useful properties of copulas for our purpose:

Proposition 13. *If $\underline{X}$ has as a copula C and if $(\alpha_1, \ldots, \alpha_n)$ are n strictly increasing functions defined respectively on the supports of the X_i, then C is also the copula of $(\alpha_1(X_1), \ldots, \alpha_n(X_n))$.*

This property shows that any measure of association between X_1 and X_2 must be a function only of the copula that links X_1 and X_2 and not of their marginal distributions.

If we are interested in the bidimensional marginal distributions of a random vector $\underline{X}$, we have the following property that links the copula of a bidimensional extracted random vector (X_i, X_j) and the copula of the distribution of $\underline{X}$:

Proposition 14 (*Bidimensional Marginals*). *Let $\underline{X}$ be a random vector with a distribution defined by its copula C and its marginal distributions F_i. The cumulative distribution function F_{ij} of the bidimensional random vector (X_i, X_j) with $i < j$ is defined by its marginal distributions (F_i, F_j) and the copula C_{ij} through the relation*

$$F_{ij}(x_i, x_j) = C_{ij}(F_i(x_i), F_j(x_j)) \tag{21}$$

with $C_{ij}(u_i, u_j) = C(1, \ldots, 1, u_i, 1, \ldots, 1, u_j, 1, \ldots, 1)$, where u_i and u_j are respectively at position i and j.

This result will be used to reformulate the relation (4) in terms of extracted bidimensional copulas.

4. New interpretation of the Nataf transformation through the copula theory

The Nataf transformation is the composition of two transformations T_1 and T_2, with an additional hypothesis that upon the action of T_1, the initial random vector $\underline{X}$ is mapped into a gaussian vector $\underline{Y} = T_1(\underline{X})$.

Formalizing the hypothesis underlying the Nataf transformation leads to:

Proposition 15 (*Normal Copula Through the Nataf Transformation*). *Let $\underline{X}$ be a random vector with unknown copula C_X, known marginal cumulative distribution functions F_i and known linear correlation matrix $\underline{\underline{R}}$. Assuming that this vector is mapped into a gaussian vector $\underline{Y} = T_1(\underline{X})$ with distribution $\mathcal{N}(\underline{0}, \underline{\underline{R_0}})$ upon the action of T_1 as defined in (2) is equivalent to the assumption that C_X is the normal copula parameterized by the correlation matrix $\underline{\underline{R}}_0$.*

The demonstration is a direct application of the invariance of the copula by strictly increasing transformation of the components of a random vector. By definition of the normal copula, the copula C_Y of $\underline{Y}$ is exactly the normal copula $C^N_{\underline{\underline{R}}_0}$ parameterized by $\underline{\underline{\mathrm{cor}}}[\underline{Y}] = \underline{\underline{R}}_0$. Then, the transformation T_1 is bijective, and its inverse is

Fig. 1. Iso-density contours of three copulas (left column) and three distributions (right column) built upon these copulas with standard normal marginal distributions and a linear correlation $\rho = 0.8$. It is worth noticing the differences between these probability density functions, even if they share the same marginal distributions and the same linear correlation matrix.

$$T_1^{-1} : \underline{Y} \mapsto \underline{X} = T_1^{-1}(\underline{Y}) = \begin{pmatrix} F_1^{-1} \circ \Phi(Y_1) \\ F_2^{-1} \circ \Phi(Y_2) \\ \vdots \\ F_n^{-1} \circ \Phi(Y_n) \end{pmatrix}. \quad (22)$$

This transformation only acts on the marginal distributions of $\underline{Y}$, and is a strictly increasing transformation which preserves the copula of the transformed random vector (see Proposition 13). We conclude that $C_X = C_Y = C^N_{\underline{\underline{R}}_0}$.

From the definition (4) of the correlation matrix $\underline{\underline{R}}_0$ and the expression of a bidimensional marginal probability density function using (19) and (21), we have

$$r_{ij} = \frac{1}{\sigma_i \sigma_j} \iint_{\mathbb{R}^2} (x_i - \mu_i)(x_j - \mu_j) c_{ij}(F_i(x_i), F_j(x_j))$$