MONTE CARLO STATISTICS; SIMULATING CORRELATED MULTI-VARIATE STOCHASTIC VARIABLES

The following mathematical description is used in a simulation algorithm aimed at simulating wells, i.e. 1D-stratigraphic profiles with attached physical properties. In the algorithm, wells are constructed from so-called integration framework entities. These entities are grouped at three different scale levels. It is considered important that geological knowledge controls the selection of framework entities and that unrealistic realisations of variables can be redrawn. This implies that wells must be constructed one-by-one, entity-by-entity and variable-by-variable.

Variables in a computer are simulated using a (pseudo-) random number generator. When random variables are correlated, it is not simple, however, to simulate random draws using such a (pseudo-) random number generator. This is especially true when the variables must be drawn one-by-one, as in our application. The realisations of already drawn variables will in that case influence the realisation of the variable to be drawn. For example, let us assume that a positive correlation exist between the thicknesses of two layers. When for the first layer a small thickness is drawn, then also for the second layer a small thickness must be drawn. In the case of normally distributed random variables, it is possible to draw the variables consecutively from the marginal distributions. Each time a variable is to be drawn, its marginal distribution must first be updated for the already drawn variables to which it is correlated.

In the following discussion $\underline{X}$ is a stochastic vector. In our algorithm, $\underline{X}$ comprises all stochastic variables required for the simulation. A component of $\underline{X}$ is denoted by X_i. Examples of components are sonic, density, thickness and user-defined variables attached to framework entities. Each component X_i is assumed to be normally distributed with expectation μ_i and variance σ_i^2, symbolically written as: $X_i \sim N(\mu_i, \sigma_i^2)$. The vector of expectation will be denoted $\underline{\mu}$. The components are assumed to be correlated. The covariance between components i and j is indicated by σ_{ij}. Note, that the covariance between component i and itself, σ_{ii} equals σ_i^2. The matrix of covariances will be denoted as Σ. When the covariance σ_{ij} is normalised with the standard deviations σ_i and σ_j, we obtain the correlation coefficient ρ_{ij}, symbolically written as: $\rho_{ij} = \sigma_{ij} \Big/ (\sigma_i * \sigma_j)$. The matrix of correlation coefficients will be denoted by C. Sets of components can be grouped into subvectors of $\underline{X}$ denoted by $\underline{X}^{(i)}$, An example of a subvector $\underline{X}^{(i)}$ is that part of stochastic vector $\underline{X}$ comprising correlated thicknesses of a set of layers. The theorems given hereafter apply to the general case of drawing entire subvectors. However, for design reasons, the variables are, drawn one-by-one, in the final implementation of the algorithm. In other words the subvector $\underline{X}^{(i)}$ to be drawn has only one component. This is illustrated by the example at the end of this Appendix.

We require two theorems for our algorithm to work. Theorem 1.1 is used for updating the expectation and covariance matrix of a variable to be drawn, given some already drawn correlated variables (Mardia, 1979). This theorem requires the covariance matrix to be specified completely. In general, the user will not be in a position to specify all coefficients. Therefore, the unspecified correlation coefficients must be approximated first. This is accomplished with Theorem 1.2 (Meeuwissen et.al., 1994).

In the following discussion, first the two theorems are given, followed by an illustration of their use with an example.

Theorem 1.1

First we introduce some notation. Let $\underline{X}$ be a n-dimensional stochastic vector which is partitioned as follows:

$$\underline{X} = \begin{pmatrix} \underline{X}^{(1)} \\ \underline{X}^{(2)} \end{pmatrix}, \tag{1.1}$$

with expectation $\mathrm{E}[\underline{X}]$ equal to $\underline{\mu}$:

$$\underline{\mu} = \mathrm{E}[\underline{X}] = \begin{pmatrix} \underline{\mu}^{(1)} \\ \underline{\mu}^{(2)} \end{pmatrix}, \tag{1.2}$$

and a positive definite covariance matrix $Cov(\underline{X})$ given by:

$$\Sigma = Cov(\underline{X}) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}. \tag{1.3}$$

Suppose $\underline{X}$ is multivariate normally distributed with expectation $\underline{\mu}$ and covariance matrix Σ, which can be symbolically written as:

$$\underline{X} \sim MVN(\underline{\mu}, \Sigma). \tag{1.4}$$

Here $\sim$ denotes 'is distributed as' and MVN indicates multivariate normally distributed. Then the conditional distribution of $\underline{X}^{(1)}$ given a realisation $\underline{x}^{(2)}$ of $\underline{X}^{(2)}$ is multivariate normally distributed with expectation:

$$\hat{\underline{\mu}}^{(1)} = \underline{\mu}^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}(\underline{x}^{(2)} - \underline{\mu}^{(2)}), \tag{1.5}$$

where $\hat{\underline{\mu}}^{(1)}$ is the updated expectation. The updated covariance matrix $\hat{\Sigma}_{11}$ is given by:

$$\hat{\Sigma}_{11} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}. \tag{1.6}$$

Theorem 1.2

Suppose X_1, X_2 and X_3 are correlated random variables which satisfy:

$$E[X_1 | X_2 = x_2] \quad \text{is linear in } x_2, \tag{1.7}$$

and

$E[X_1|X_3 = x_3]$ is linear in x_3. (1.8)

Then, given the correlation coefficients ρ_{12} between the pairs X_1 and X_2 and ρ_{13} between X_1 and X_3, the correlation coefficient ρ_{23} is given by:

$$\rho_{23} = \rho_{12}\rho_{13}. \tag{1.9}$$

The conditions in the theorem imply, say for X_1, X_2, that given a realisation x_2 of variable X_2, the expectation of X_1 shifts linearly towards x_2. For normal distributions this is always satisfied, as can be seen from theorem 1.1, equation (1.5).

Although this theorem applies to three variables with one missing correlation coefficient only, we are going to use it also, without strict theoretical justification, for more than three variables where several correlation coefficients may be missing. We must note here, that, for more then three correlated variables, the positive definiteness of the covariance matrix may be violated by this procedure. In practice we have seen this happen only in some rare cases.

We will illustrate the use of these theorems with the following example. Suppose the correlation matrix has been specified for five variables as follows:

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & * & 0.8 & * \\ 0 & * & 1 & * & 0.6 \\ 0 & 0.8 & * & 1 & 0.4 \\ 0 & * & 0.6 & 0.4 & 1 \end{bmatrix}. \tag{1.10}$$

In this particular example, ρ_{24} , ρ_{35} and ρ_{45} are known coefficients and ρ_{34} , ρ_{25} and ρ_{23} are unknown, which is indicated in the matrix by the * symbol. Using 1.9 we can determine two of the unspecified correlation coefficients.

$$\rho_{34} = \rho_{35}\rho_{54} = 0.24, \tag{1.11}$$

and

$$\rho_{25} = \rho_{24}\rho_{45} = 0.32. \tag{1.12}$$

However, ρ_{23} cannot be determined by combination of two of the given correlation coefficients. In a second step, we can approximate it using the previously determined correlation coefficients:

$$\rho_{23} = \rho_{24}\rho_{43}, \tag{1.13}$$

which can be expanded using 1.11 to:

$$\rho_{23} = \rho_{24}\rho_{35}\rho_{54} = 0.192. \tag{1.14}$$

Note, that we could also have used:

$$\rho_{23} = \rho_{25}\rho_{53} = \rho_{24}\rho_{45}\rho_{53}. \tag{1.15}$$

In this particular case, the same value for ρ_{23} will be obtained for (1.14) and (1.15). In general, however the approximation is not unique. If several combinations are possible, in which the number of initially specified correlation coefficients differs, then a selection is made from the combinations with the least number of initial coefficients. From these we, arbitrarily choose one of the possible combinations. Thus, if in a different example, ρ_{23}, ρ_{34}, ρ_{35}, ρ_{45} would have been specified, then we can obtain ρ_{25}, either from:

$$\rho_{25} = \rho_{23}\rho_{35}, \tag{1.16}$$

or, from:

$$\rho_{25} = \rho_{23}\rho_{34}\rho_{45}. \tag{1.17}$$

The former expression is favoured because it contains less specified correlation coefficients.

With respect to the approximate nature of the procedure, we emphasise that after multiplying correlation coefficients, the resulting number comes closer and closer to zero. Therefore, the effect of the resulting approximation of the correlation coefficient decreases rapidly. Hence, we argue that making an error in the approximation has little effect when many terms are involved.

After application of the above procedure, the correlation matrix of (1.10) can be approximated by:

$$\tilde{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0.192 & 0.8 & 0.32 \\ 0 & 0.192 & 1 & 0.24 & 0.6 \\ 0 & 0.8 & 0.24 & 1 & 0.4 \\ 0 & 0.32 & 0.6 & 0.4 & 1 \end{bmatrix}. \tag{1.18}$$

We can now draw samples for all variables. Suppose we would like to draw them in the order X_3, X_5, X_1, X_2, X_4. When selecting X_3, no other has been drawn, so we can just draw it from its marginal probability density function $X_3 \sim N(\mu_3, \sigma_3^2)$. Now X_5 must be drawn, conditioned on the x_3 value. Using theorem 1.1, we find:

$$\hat{\mu}_5 = \mu_5 + \sigma_{35}(\sigma_3^2)^{-1}(x_3 - \mu_3), \tag{1.19}$$

and

$$\hat{\sigma}_5^2 = \sigma_5^2 - \sigma_{35}(\sigma_3^2)^{-1}\sigma_{53}, \tag{1.20}$$

where

$$\sigma_{35} = \rho_{35}\sigma_3\sigma_5, \tag{1.21}$$

is the covariance between X_3 and X_5. Now X_5 can be drawn from $N(\hat{\mu}_5, \hat{\sigma}_5^2)$.

Now X_1 is to be drawn. Since it is independent of X_2, X_3, X_4 and X_5 it can be drawn from its marginal distribution $N(\mu_1, \sigma_1^2)$. Finally, for X_2 and X_4 we use:

$$\hat{\mu}_2 = \mu_2 + \begin{bmatrix}\sigma_{23}\sigma_{25}\end{bmatrix}\begin{bmatrix}\sigma_3^2 & \sigma_{35}\\ \sigma_{35} & \sigma_5^2\end{bmatrix}^{-1}\begin{bmatrix}x_3-\mu_3\\ x_5-\mu_5\end{bmatrix}, \tag{1.22}$$

$$\hat{\sigma}_2^2 = \sigma_2^2 - \begin{bmatrix}\sigma_{23}\sigma_{25}\end{bmatrix}\begin{bmatrix}\sigma_3^2 & \sigma_{35}\\ \sigma_{35} & \sigma_5^2\end{bmatrix}^{-1}\begin{bmatrix}\sigma_{23}\\ \sigma_{25}\end{bmatrix}, \tag{1.23}$$

and

$$\hat{\mu}_4 = \mu_4 + \begin{bmatrix}\sigma_{24}\sigma_{34}\sigma_{54}\end{bmatrix}\begin{bmatrix}\sigma_2^2 & \sigma_{23} & \sigma_{25}\\ \sigma_{23} & \sigma_3^2 & \sigma_{35}\\ \sigma_{25} & \sigma_{35} & \sigma_5^2\end{bmatrix}^{-1}\begin{bmatrix}x_2-\mu_2\\ x_3-\mu_3\\ x_5-\mu_5\end{bmatrix}, \tag{1.24}$$

$$\hat{\sigma}_4^2 = \sigma_4^2 - \begin{bmatrix}\sigma_{24}\sigma_{34}\sigma_{54}\end{bmatrix}\begin{bmatrix}\sigma_2^2 & \sigma_{23} & \sigma_{25}\\ \sigma_{23} & \sigma_3^2 & \sigma_{35}\\ \sigma_{25} & \sigma_{35} & \sigma_5^2\end{bmatrix}^{-1}\begin{bmatrix}\sigma_{24}\\ \sigma_{34}\\ \sigma_{54}\end{bmatrix}, \tag{1.25}$$

respectively.

This allows us to draw the variables one by one in any order. Also, we can redraw any one of the variables when needed, and condition on the latest drawn value for each of the correlated variables.