Application of Disjunctive Kriging in Sequential Simulation[1]

David Casson (3drc1@queensu.ca)
Julian M. Ortiz (julian.ortiz@queensu.ca)

Abstract

Quantifying uncertainty in ore body modelling improves economic outcomes from mine development and operation. Simulation is a frequently utilized technique in geostatistics providing information on the level of uncertainty associated with estimated ore body properties at unsampled locations, both individually and in aggregate. Available simulation techniques do not allow for informed application of a random function model relying instead on: i) the assumption that the random function is MultiGaussian (as in sequential Gaussian simulation), ii) no assumption on the random function model (as in sequential indicator simulation), or iii) brute force inference of higher order statistics (multipoint simulation with machine learning). Disjunctive Kriging utilizes polynomial expansions of random function families to estimate parameters at unsampled locations. An algorithm has been developed to apply Disjunctive Kriging in sequential simulation using the Hermitian expansion of a Gaussian random function. Using the same conditioning information, simulation results closely match those from sequential Gaussian simulation. This suggests applying the algorithm with non Gaussian random function families and associated polynomial expansions will provide a valuable sequential simulation tool.

1. Introduction

1.1. Context in Mineral Resource Estimation

Traditionally, the qualified person completing a Mineral Resource estimate applied kriging to arrive at a best estimate for mineral resource grade and tonnage. Kriging is a deterministic method unable to provide a relevant measure of uncertainty associated with its deterministic estimates. The qualified person would classify the mineral resource as measured, indicated, or inferred based on personal experience and industry association guidance related to drill hole spacing and deposit type. For example, in a copper porphyry setting drills holes of a certain spacing might result in indicated resources, while drill holes at slightly greater spacing would be inferred. These specific spacings differ depending on the type of ore body and geological setting.

Simulation provides a quantification of uncertainty in resource modelling. This allows an estimator to arrive at a quantitative view of how probable it is that a given block has a given value (e.g., probability that grade of block is above cut off). This probabilistic estimate of uncertainty can be carried through the

[1] Cite as: Casson D., Ortiz J. M. (2021) Application of Disjunctive Kriging in Sequential Simulation, Predictive Geometallurgy and Geostatistics Lab, Queen's University, Annual Report 2021, paper 2021-05, 46-60.

mine design and operation to make better decisions, understanding the range of possible operational and financial outcomes and associated likelihood.

Simulation infers the properties of the regionalized variable under study based on the available data samples. The accuracy of the resulting simulations as a predictor of outcomes and probabilities depends entirely on how well the inferred model matches the true properties of the regionalized variable. Said simply, the appropriateness of the chosen model for the variable under study will dictate the accuracy and usefulness of simulation results.

1.2. Traditional Simulation Techniques

Commonly applied simulation techniques include Sequential Gaussian Simulation and Sequential Indicator Simulation. As is described further below, both these sequential simulation methods require a random draw from an estimated Conditional Cumulative Distribution Function ("CCDF") at unsampled locations. The CCDF, which is based on information in a neighborhood deemed to be relevant, reflects the probability that the unknown value at the location in question is below any value in its possible range. Simulation will not be successful if the CCDF is not accurate. Sequential Gaussian Simulation relies on the assumption that the regionalized variable under study is Multigaussian in nature. This is often not true. Simulation will also not be successful if it is not practical to implement. Sequential Indicator Simulation makes no assumption on the model of the variable under study but requires computation and modelling of many indicator variograms which makes it a cumbersome approach and challenging to implement in practice.

The assumption of MultiGaussianity provides the theoretical basis for many methods in geostatistics. The specific indicator properties of a multipoint (multivariate) Gaussian variable (i.e., indicator variogram) can be used to define a Gaussian function for the local CCDF. These Gaussian local distribution functions determine the probabilities of potential values at the unsampled locations, providing a measure of "local" uncertainty. Sequential Gaussian Simulation ("SGS") extends this approach to estimate global uncertainty amongst a population of unsampled locations, applying Monte-Carlo simulation along with the inferred local Gaussian CCDF's at unsampled locations to arrive at a simulated point value. Sequential simulation takes a random path through unsampled locations and once a simulated value is determined at a given location, it is assumed fixed and incorporated into the CCDF's for subsequently simulated unsampled locations. Repeating this sequential simulation provides multiple potential realizations of the unsampled locations creating, in aggregate, a measure of uncertainty of the entire unsampled set of locations. Such an approach is applied to create a measure of uncertainty related to grade and tonnage estimates, e.g., the probability across 1000 simulations that a given block is ore or waste.

The Gaussian approach described above is in contrast to the indicator approach applied in Sequential Indicator Simulation ("SIS") which makes no assumption regarding model function or parameters. While the indicator approach can account for non-symmetric distributions and/or non-diffusive distributions, it has several limitations described below (Machuca-Mory et al 2008):

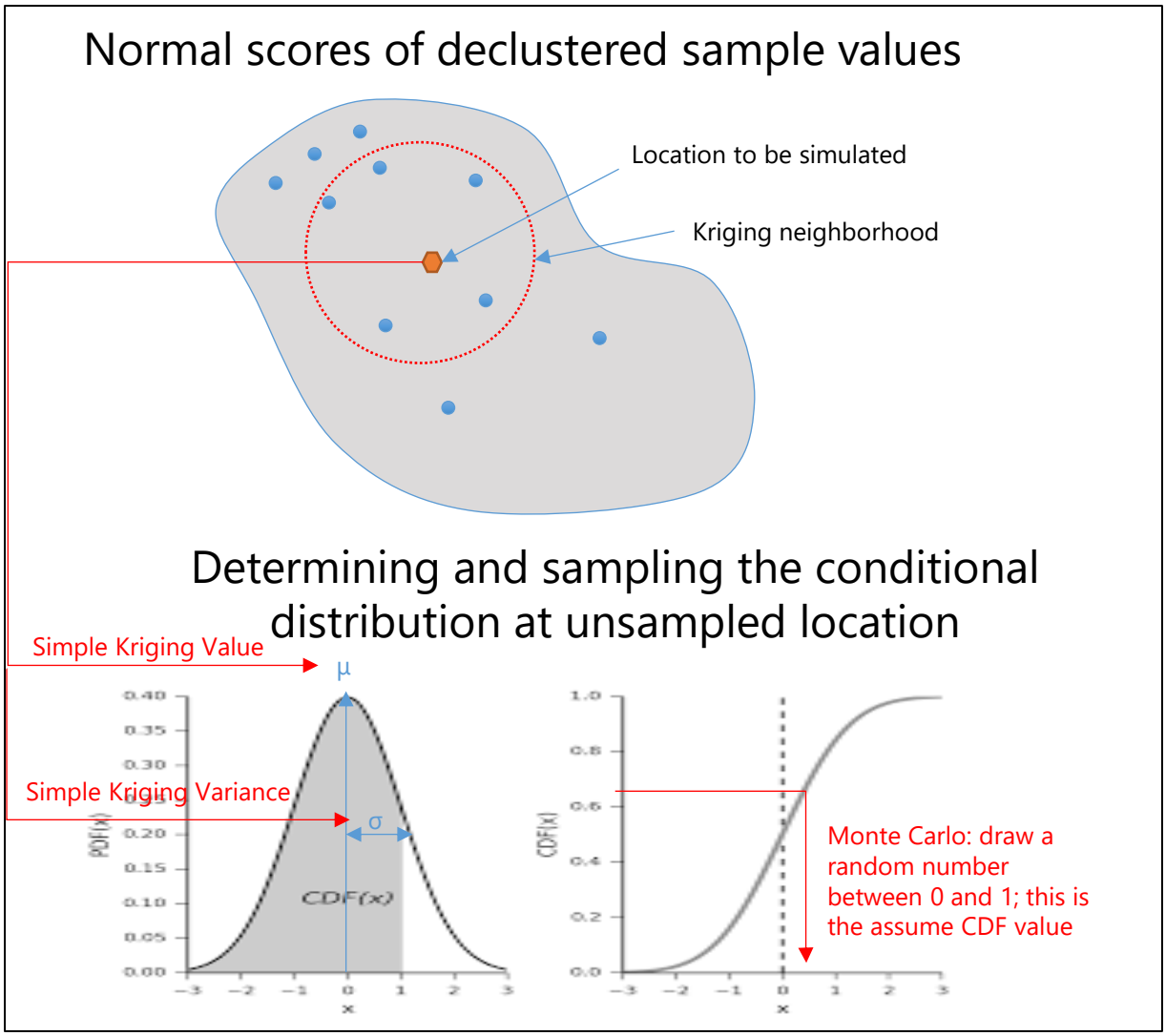
- Discretization and interpolation of the conditional CDF creates error.

- Cokriging of indicator thresholds (which is theoretically required) is not practical with a linear model of co-regionalization (challenging to limit cross and direct indicator variograms to linear combinations of specified families of models).
- Tail behavior is difficult to model given sparse relevant data points.

The steps to sequential Gaussian simulation are described below.

1. Decluster data to get representative distribution;
2. Transform data to normal scores, based on the representative distribution;
3. Visit nodes in a random path;
4. At every node, search for nearby samples or previously simulated nodes;
5. Krige the normal scores using samples and nodes (this requires the variogram model of the normal score transform of the original variable);
6. Draw a value from the conditional distribution;
7. Back transform the simulated value.

The figure below adapted from Ortiz (2019) illustrates the key steps in SGS simulation.

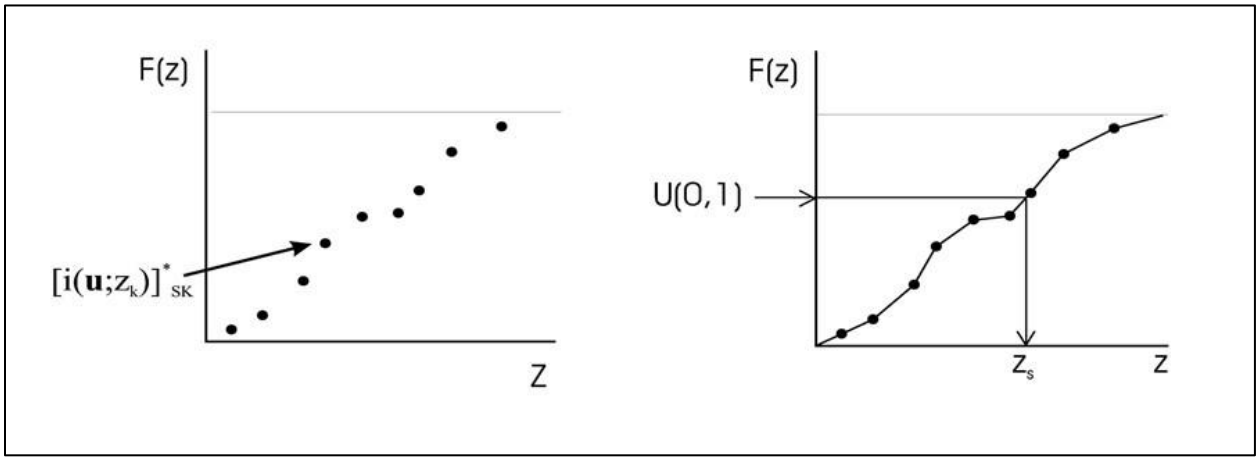


Sequential indicator simulation is non-parametric in nature. No assumptions are made regarding the bi-variate distribution.

The steps to sequential indicator simulation are described below.

1. Visit nodes in a random path;
2. At every node, search for nearby samples or previously simulated nodes;
3. Select a number of threshold values for the variable in question;
4. For each threshold:
    a. Code nearby known or simulated data points as "1" or "0" if they are below of above the threshold respectively;
    b. Apply simple kriging to these values (requires a variogram fit to the indicator data points);
    c. Result is probability that value at unsampled location is below the threshold;
5. Interpolate the threshold-probability data points into a CDF curve;
6. Draw a value from the conditional distribution.

The figure below adapted from Ortiz (2019) illustrates the interpolation approach to modelling the CCDF using indicator kriging.



1.3. Polynomial Expansions for Modelling Probability Density Functions and Cumulative Density Functions

Determination of local conditional Probability Density Functions ("PDF's") is an important component of modern geostatistics. Conditional PDF's (and corresponding Cumulative Distribution Functions or "CDF's") allow for simulation of an attribute at an unsampled location conditioned to the available and relevant data. The sequential simulation technique relies on a random draw from the CCDF at the unsampled location.

Asymptotic expansions are a mathematical technique for approximating a function with finite variance by a weighted average of a family of functions related to a "developing distribution function" or random function family. This type of mathematical formulation can be applied to model CDF's in situations where suitable developing distribution functions are selected.

In the first part of the twentieth century, significant research was completed on approximating empirical distributions with theoretical functions. In many cases, this work focused on the use of asymptotic expansions, in which the error of the approximation approaches zero as a parameter (the order) of the expansion approaches infinite (Wallace, 1958). The premise upon which these expansions are based is found in the Charlier Differential Series.

A random function is completely described by its moments. For any random function, the moments can be defined with the moment generating function shown below.

$$M_X(t) = E[e^{tX}] \quad (1)$$

In a similar manner to the Moment Generating Function, a random variable and its distribution can be defined by its Characteristic Function *"f(t)"* or sometimes *"φ(t)"*. The Characteristic Function is similar to the Moment Generating Function. It resembles a Fourier Transform (but in complex conjugate) of the underlying random variable's PDF. The expression is shown below.

$$\varphi_X(t) = E[ei^{tX}] \quad (2)$$

Cumulants are similar but not identical to moments. A random variable can be completely described by its cumulants or its moments. The cumulants are defined by the Cumulant Generating Function, which is the natural logarithm of the characteristic function. As a result, the characteristic functions can be written as the natural exponent of the cumulant generating function. This is shown below for a characteristic function "*f(t)*".

$$f(t) = e^{\left[\sum_{r=1}^{\infty} \kappa_r \frac{(it)^r}{r!}\right]} \quad (3)$$

Using the above expression for both an empirical characteristic function *f(t)* and a chosen theoretical characteristic function *ψ(t)*, an expression may be defined to relate the theoretical and empirical characteristic functions in terms of the difference between cumulants of each (κ_r and γ_r respectively) as shown below.

$$f(t) = e^{\left[\sum_{r=1}^{\infty} (\kappa_r - \gamma_r) \frac{(it)^r}{r!}\right]} \psi(t) \quad (4)$$

If the PDF of the chosen theoretical distribution (random function family) referred to here as *Ψ(x)* and all of its derivatives are continuous and have finite variance (i.e., vanish in the extremes but are continuous otherwise), integration by parts is possible and therefore expression of the two PDF's (theoretical and empirical) based on the difference between the cumulants is possible. This is shown below where "D" is the differential operator and comes from the integral relationship between probability distribution functions and associated characteristic functions.

$$F(x) = e^{\left[\sum_{r=1}^{\infty} (\kappa_r - \gamma_r) \frac{(-D)^r}{r!}\right]} \Psi(x) \quad (5)$$

By selecting a family of theoretical distributions with finite variance, we can create an expansion of the probability distribution function to be approximated based on the derivatives of the theoretical PDF (and associated CDF) selected and the differences between observed and theoretical cumulants. As noted by Mustapha and Dimitrakopolous (2010) "it is relatively easy in many statistical situations to determine moments, but it is extremely hard or impossible to determine the distributions themselves".

1.4. Review of Hermitian Polynomial Expansion

The Edgeworth Approximation is a specific case of an asymptotic expansion, where the theoretical distribution function is selected to be the normal distribution (which satisfies the criteria of finite variance). In choosing the normal distribution we are not assuming that the empirical distribution to be modelled is truly Gaussian, we are simply choosing the normal distribution to base our expansion on. If the empirical were perfectly Gaussian then no expansion would be needed to write the empirical PDF in terms of the theoretical Gaussian PDF. Intuitively the closer the empirical distribution is to the theoretical the "better" the expansion should be (i.e., how close the expansion approximates the theoretical distribution for any given truncation of the expansion series). Once the choice of the Gaussian distribution is made for the theoretical developing function, the relationship between the cumulants can be further defined based on the theoretical cumulants of the normal distribution. The mean and variance of the empirical and theoretical distribution functions are set equal resulting in an expansion based on the difference of the higher order cumulants. The higher order cumulants of the theoretical normal distribution are zero which further simplifies the equation to depend only on higher order cumulants of the empirical distribution. The empirical cumulative distribution function to be approximated can be written as an expansion of a polynomial and the derivatives of the normal distribution as shown in the equation below based on the observed cumulants of the empirical distribution. Here the λ values are derived from the observed cumulant values of each order in the empirical data set.

$$F_n(x) = \phi(x) - \frac{\lambda_3 \phi^3(x)}{-6\sqrt{n}} + \frac{1}{n}\left[\frac{\lambda_4 \phi^4(x)}{24} + \frac{\lambda_3^2 \phi^6(x)}{72}\right] + \cdots \qquad (6)$$

Given the unique relationship between the normal distribution and its derivatives, and its well-known representation in Hermite Polynomials, the Edgeworth Approximation can be rewritten as a function of Hermite Polynomials, $h_{(n)}(x)$, which is shown below for an expansion up to order 6.

$$e_4(x) = \phi(x)\left[1 + \frac{\kappa_3 h^3(x)}{6\sqrt{n}} + \frac{\kappa_4 h^4(x)}{24n} + \frac{\kappa_3^2 h^6(x)}{72n}\right] \qquad (7)$$

The Hermite Polynomial of order n, $h_{(n)}(y)$, is defined in the below equation where $g(y)$ is the standard gaussian pdf function.

$$H_n(y) = \frac{1}{\sqrt{n!}\, g(y)} \frac{d^n g(y)}{dy^n} \qquad \forall n \geq 0 \qquad (8)$$

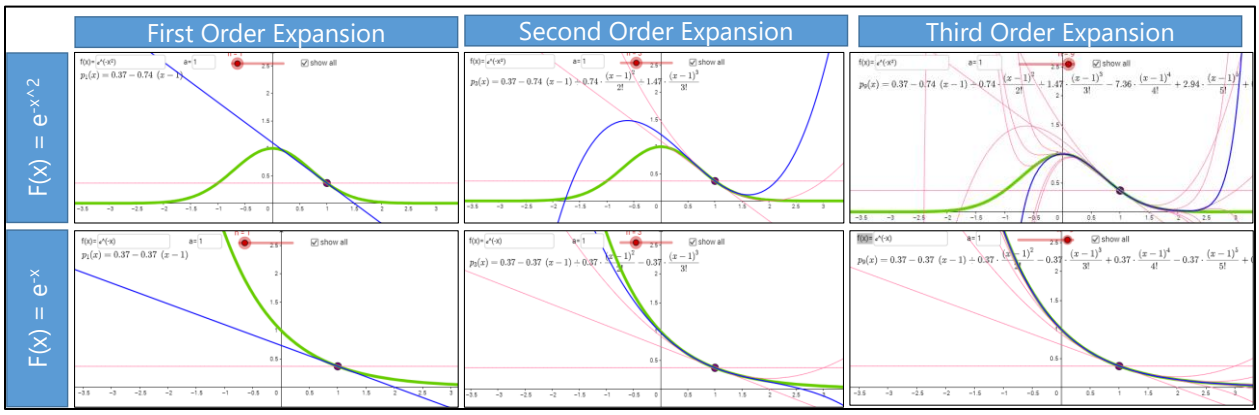
The expression for calculating Hermite Polynomials can be simplified to a recursive formula as shown below, called Rodrigues Formula, where $H_0 = 1$ and $H_1 = -y$.

$$H_{n+1}(y) = -\frac{1}{\sqrt{n+1}} y H_n(y) - \sqrt{\frac{n}{n+1}} H_{n-1}(y) \qquad \forall n \geq 0 \qquad (9)$$

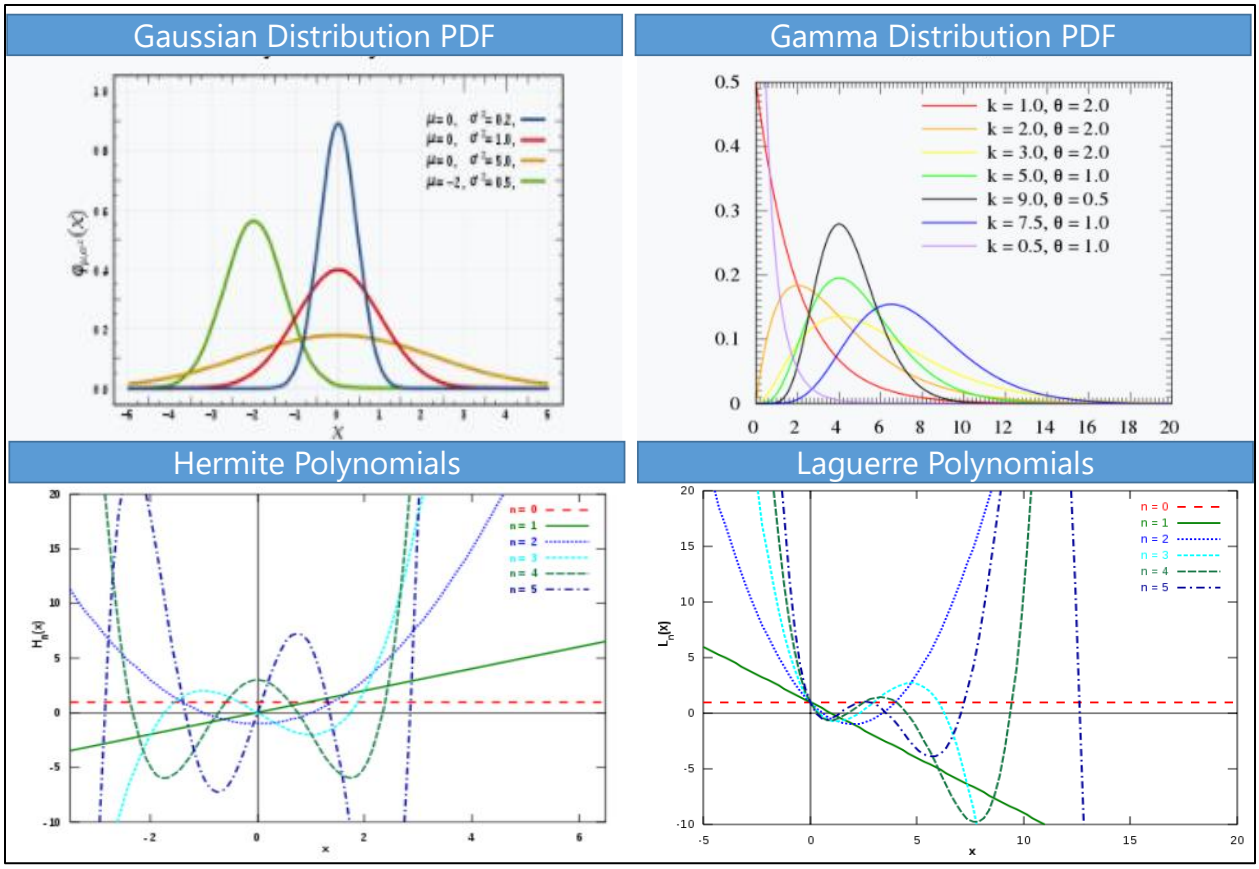
$$f(y(u)) = \sum_{n=0}^{\infty} f_n H_n(y(u)) \qquad (10)$$

1.5. Practical Illustration of Modelling a Finite Function with an Asymptotic Polynomial Expansion

An image of a Taylor polynomial approximation for two functions is shown in the figure below to illustrate the concept of polynomial approximation for a finite function.



Extending the concept to Gaussian and Gamma functions, the images below illustrate how Hermitian and Laguerrian polynomials are suited to approximating Gaussian and Gamma random functions.



1.6. Disjunctive Kriging

Disjunctive Kriging involves the use of polynomial asymptotic expansions to estimate values at unsampled locations. Asymptotic polynomial expansions are used to define the global anamorphosis function based on the sample data and respective transformed values. This can be a Gaussian transform / Hermitian expansion or a Gamma transform / Laguerrian expansion depending on which distribution is deemed most appropriate for the attribute being measured. The polynomial expansion values calculated for sample values/locations can be used in conjunction with simple kriging to determine the transformed (e.g., normal score or "gamma score") value at an unsampled location. The global anamorphosis function is used to determine the "raw" back transformed value of the attribute at the unsampled location.

Disjunctive kriging is simple co-kriging of the polynomial expansion of all orders (so informed by relative location and polynomial values at neighboring sampled locations). The polynomial expansions considered are orthogonal basis; the covariance between polynomials of different orders is zero and the covariance of polynomials of different orders is linked to a single variogram model; as a result, disjunctive kriging becomes simple kriging of the polynomial values for the transformed data values and then a linear sum of the resulting polynomial values across orders. This approach allows for the calculation of an expected value at an unsampled location based on statistical distance of nearby data points and reflecting the local CCDF as embedded in the polynomial values across orders at nearby sampled locations.

The polynomial expansion can be chosen to correspond to a given random function selection. While disjunctive kriging using Gaussian / Hermitian expansions is well documented as a substitute for simple MultiGaussian kriging, its application with other families of bivariate function families (such as MultiGamma) has not been as thoroughly described. Work has been done testing the relative appropriateness of a given expansion for modeling conditional PDF's by comparing a Hermitian expansion for the PDF and a Laguerrian expansion for the PDF, with the actual PDF from the exhaustive data set.

Application in Sequential Simulation

The polynomial approximation technique combined with disjunctive kriging can be used to define an expression for the local CCDF. Such polynomial expansions reflect the choice of bi-variate random function family. Ortiz (2004) provides a good overview of fitting a finite function (any finite function is acceptable) with an expansion of Hermite Polynomials. The approach is to set the function equal to a weighted sum of the Hermite Polynomial values as shown below. The coefficient value for a given order n is solved by calculation of the expected value of the function and the Hermite Polynomial of a given order n. The same author provides a method of determining the CDF function (indicator function $I_Y(u;y_c) = prob\ y \leq y_c$) of a given distribution based on Hermite Polynomial values as shown in the equations below. (note: $G(y)$ is the standard normal CDF)

$$I_Y(u; y_c) = \sum_{p=0}^{P} \psi_p H_p\big(Y(u)\big) \qquad (11)$$

$$\psi_0 = G(y_c) \qquad (12)$$

$$\psi_p = \frac{1}{\sqrt{p}} H_{p-1}(y_c) g(y_c) \qquad (13)$$

These equations let us choose (or observe) our experimental moments (mean and variance of Gaussian PDF $g(y_c)$ above) which inform the coefficients of the Hermitian expansion. The approximation of the CDF is a sum of products at each order of a coefficient (based on experimental moments) and the Hermitian polynomial. At each unsampled location we apply disjunctive kriging to determine the polynomial values based on the polynomial values calculated at the sample locations. We then apply the above equations (derived from the observed moments) to model the local CCDF function. A random draw can then be taken from this CCDF to arrive at a simulated value for the unsampled location.

2. Simulation Methodology

2.1. Prior Work

Emery (2006) also completed work on the use of various bivariate random function families to estimate ore body attributes at unsampled locations with Disjunctive Kriging. Work has also been completed (Emery 2002) using isofactorial representation of the bivariate random function family and disjunctive kriging to simulate ore body parameters ("sequential isofactorial simulation"). In this instance, while the theory is well described, the practical implementation including the approach sampling the local conditional cumulative distribution function is not.

2.2. Simulation Approach

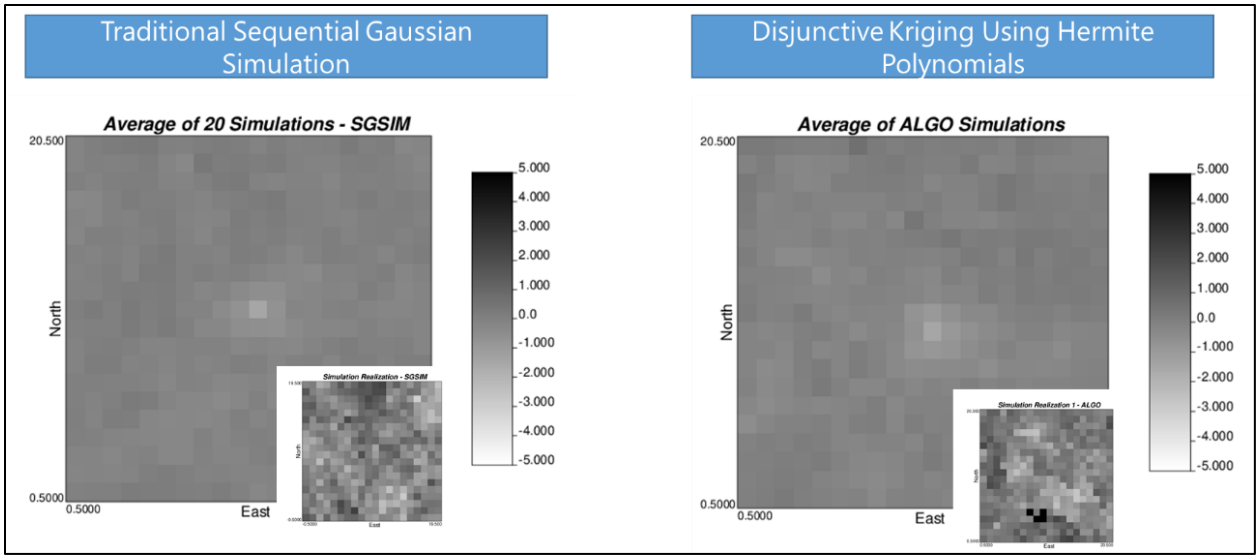
A sequential simulation algorithm was developed for sequential simulation using disjunctive kriging of Hermitian polynomials. The Hermitian polynomials up to a selected order "N" are calculated at all sampled locations based on normal score values. A variogram model is fitted to normal score values. Simple kriging is completed to solve for the Hermitian values of each order at the unsampled location. The covariance between polynomials of an order N is based on the normal score value variogram with the resulting correlogram value raised to the power "N".

Once the Hermitian values at the unsampled location are determined, the CCDF function at that location, (as a weighted sum of the Hermitian values) is used for the random draw. The weightings for each order Hermitian polynomial in the CCDF are derived based on equations 11, 12 and 13 shown in section 1.6 above. The result is an equation for the CCDF value at an unsampled location as a function of the actual (unknown) variable value at that location.

Sequential simulation draws a random number between zero and one and assumes this to be the CCDF value. Because the CCDF (by definition) is a monotonic function, a guess and check bounding algorithm was designed to iteratively determine the corresponding actual variable value that the CCDF random draw corresponds to (within a specified tolerance). The algorithm evaluates the CCDF value for the endpoints and midpoint of a range of the underlying variable; this range is then halved based on the relative location of the randomly drawn simulated CCDF value. The process is repeated until the drawn CCDF value is deemed close enough to one of the calculated CCDF values.

3. Results

It is possible to compare the results of the Hermitian disjunctive kriging simulation algorithm against frequently utilized traditional sequential Gaussian simulation programs. For such a comparison a 200x200 empty array was initialized and a seed value of -1.5 is placed in the center of the array. The seeded array is run through both the Hermitian disjunctive kriging algorithm, and GSLIB's Sequential Gaussian Simulation ("SGS") program. The results of both are shown below and indicate that simulation using Hermitian polynomials appears to accurately simulate a multi-Gaussian distribution.



4. Extensions of Methodology

4.1. MultiGamma Random Function

Wilson and Wragg (1973) provide three methods "…for the reconstruction of a continuous probability density function f(x) from given values of the moments of the distribution." One of these methods involves use of an asymptotic expansion, specifically an expansion of Laguerre Polynomials. This approach is suited to distributions that are expected or observed to be "Gamma like". This contrasts with the Hermite Polynomial expansions described by Edgeworth that are appropriate for distributions observed or expected to be "Gaussian like".

When dealing with the Hermitian expansion, the Gaussian distribution function is characterized by its mean and variance, which may be determined in a straightforward manner from the values of the sample data. In the Laguerrian case, the Gamma distribution function is characterized by a shape parameter α and a rate parameter β that are not as easily determined from the sample data. The shape and rate parameter are key inputs to the PDF and CDF functions for Gamma distributions. In order to utilize the Laguerrian expansion, we require a value for the shape and rate parameters. Using certain integral conditions, these parameters can be determined according to the below equations based on the first (mean) and second (variance) moments of the transformed sample data (Mustapha and Dimitrakopoulos 2010).

$$\alpha = \frac{2m_1^2 - m_2}{m_2 - m_1^2} \qquad (14)$$

$$\beta = \frac{m_1}{m_2 - m_1^2} \qquad (15)$$

With these parameters calculated, the equations of the Laguerrian expansion and polynomials are set out below.

$$P(z) \approx P_{\infty}^{\alpha}(z) = \sum_{n=0}^{\infty} r_n L_n^{(\alpha)}(z) \phi_{GAM}(z) \quad (16)$$

$$\phi_{GAM}(z) = \frac{\beta}{\Gamma(\alpha+1)} z^{\alpha} e^{-z} \quad (17)$$

$$L_n^{(\alpha)}(z) = \sum_{i=0}^{n} \frac{(-1)^i}{i!} \binom{n+\alpha}{n-i} z^i \quad (18)$$

$$r_n = \frac{n!\,\Gamma(\alpha+1)}{\Gamma(n+\alpha+1)} \sum_{i=0}^{n} \frac{(-1)^i}{i!} \binom{n+\alpha}{n-i} \beta^i m_i \quad (19)$$

It can be seen above that the PDF value as determined by the Laguerrian expansion is fully defined by the observed moments of order 1, 2, …, i of the available (and relevant) sample data. In the Hermitian case, as previously described, the PDF is determined based on the observed cumulants of the empirical distribution where such cumulants can be directly related to observed moments. These equations let us choose (or observe) our experimental moments which inform the coefficients and thus define the global PDF distribution function. The approximation of the local PDF (or CDF) is a sum of products across orders of a coefficient (based on experimental moments) and a Laguerre or Hermite polynomial. The Laguerrian case equations are more complicated than the Hermitian case; requiring both higher orders of cumulants and also incorporation of theoretical moments.

Mustapha and Dimitrakopoulos (2010) noted that a Laguerrian expansion around a Gamma distribution "…is suited for simulating high complex natural phenomena that deviate from Gaussianity". They examined certain data sets comparing a Hermitian expansion for the PDF and a Laguerrian expansion for the PDF, with the actual PDF from the exhaustive data set. When applying the Hermitian expansion, they utilized certain practical corrections to the Hermitian expansion such as the Saddle Point Approximation. The results showed better performance of a Laguerrian expansion relative to a Hermitian expansion of the same order (which had negative probabilities) or a Hermitian expansion with Saddlepoint Approximation (which was undefined in certain areas).

The choice of theoretical distribution used to develop the PDF function to approximate the empirical CCDF is very important under truncation. While in the case of an infinite expansion, any choice of finite developing distribution function will allow for convergence of the approximation, selection of a theoretical distribution that more closely fits the empirical data should reduce the order of the expansion required to achieve a certain quality of fit. This in turn allows for more efficient computational implementation. Application of our mapping tool for determination of the "best" suited random function family (and associated orthogonal polynomial expansion) has potential to improve computational efficiency and accuracy of ore body parametric simulation.

4.2. Destructuration of Grade

Emery (2008) identified that the destructuration of grade can be modelled by "randomizing" the correlation between the orders of the polynomials (i.e., the correlations itself becomes a random function). This correlation random function can range between a regular BiGaussian (or BiGgamma) model with no destructuration or a full mosaic model representing complete destructuration. This approach takes the pure BiGaussian (and BiGamma) model and extends them to more generalized "Hermitian" and "Laguerrian" models respectively. Without this adjustment, the BiGaussian and BiGamma models would be described as diffusive. The diffusive property mathematically requires that the correlogram of the polynomials of order "p" are equal to the correlogram of the variable raised to the power p. Practically, the diffusive property means that non-uniform connectivity of values (i.e. connectivity of extreme values) cannot be reflected in the model. This is due to the higher order correlograms trending to zero as the power p increases (pure nugget).

In order to create non-diffusive more generalized models, the correlation coefficient is randomized. The correlation coefficient of higher orders is similarly the randomized correlation coefficient raised to that specific order p. A Beta distribution for the randomized variable is often selected with parameters βρ(h) and β(1-ρ(h)) where the scalar parameter β takes a value between zero and one. Under this condition, it can be shown that the correlation coefficient for order "p" is described by the equation below rather than the diffusive case where it is $\rho(h)^p$ (Chiles and Delfiner 1999)

$$\forall p \in N^* \quad T_p(h) = \frac{\Gamma(\beta)\Gamma(\beta\rho(h)+p)}{\Gamma\big(\beta\rho(h)\big)\Gamma(\beta+p)} \qquad (20)$$

The above equation allows us to calculate the correlation coefficient for our Hermitian and Laguerrian polynomial expansions to create non-diffusive models. The question of how to choose the best β for a given data set is not readily apparent. Emery (2005) incorporates the randomizing distribution as a function of β into the equation for variograms of order ω, such that using the equations previously outlining a selection criterion can be calculated; the clear drawback of this approach is the inability to separate the choice of bi-variate random function family and choice of destructuration (β). An alternate approach described by Emery in 2002 involves the use of the relation between the observed variogram and madogram as shown in the equation below for the Laguerrian case. A similar expression can be derived for the Hermitian case.

$$\gamma_1(h) = \frac{\Gamma(\alpha+\frac{1}{2})\Gamma(\beta)}{\sqrt{\pi}\Gamma(\alpha)\Gamma(\beta+\frac{1}{2})}\frac{\Gamma(\frac{\beta\gamma(h)}{\alpha}+\frac{1}{2})}{\Gamma(\frac{\beta\gamma(h)}{\alpha})} \qquad (21)$$

These equations will be used to create a tool that optimizes the choice of β to the available data. In practical terms the equation above can be thought of as a method to assess the amount of destructuration present in a data set.

5. Conclusions

A sequential simulation algorithm has been developed with the ability to model non-Gaussian random function based simulations. Polynomial expansions are used to encode information content in the available sample data. Disjunctive kriging is used to model the CCDF function at unsampled locations. Unsampled locations are visited in a random order and at each location a random draw is made from the local CCDF function. The simulation algorithm is applied with a Gaussian random function model and associated Hermitian polynomial expansion. The results under specific conditioning data are in line with the results from a traditional sequential Gaussian simulation verifying the accuracy of the approach. Incorporation of a Gamma random function model (and associated Laguerre Expansion) is outlined. Consideration on using a coefficient of destructuration to allow simulation of non-diffusive random functions is also described.

6. Acknowledgments

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN-2017-04200 and RGPAS-2017-507956.

7. References

Bhattacharya R., (2010) Normal Approximation and Asymptotic Expansions Society for Industrial and Applied Mathematics

Chiles J.P., Delfiner P., (1999) Geostatistics Modelling Spatial Uncertainty Wiley, New York NY

Dimitrakopoulos, M. (2009). High-order Statistics of Spatial Random Fields: Exploring Spatial Cumulants for Modeling Complex Non-Gaussian and Non-linear Phenomena. Mathematical Geosciences, 42(1), 65–99

Edgeworth F., (1907) On the Representation of Statistical Frequency by a Series. Journal of the Royal Statistical Society, 70 (1), 102–106

Emery X., (2002) Conditional simulation of nonGaussian random functions, Mathematical Geology, 34, p 79-100

Emery X., (2005a) Variograms of Order ω: a tool to validate a bivariate distribution model, Mathematical Geology, 37, p. 163-181

Emery X., (2005b) Geostatistical Simulation of Random Fields with Bivariate Isofactorial Distributions by Adding Mosaic Models. Stochastic environmental research and risk assessment, 19 (5), 348–360

Emery X., (2005c) Conditional Simulation of Random Fields with Bivariate Gamma Isofactorial Distributions. Mathematical geology, 37 (4), 419–445

Emery X., (2006) A Disjunctive Kriging Program for Assessment Point-Support Conditional Distributions Computers and Geosciences 32 965-983

Emery X., (2007) Using the Gibbs Sampler for Conditional Simulation of Gaussian-Based Random Fields. Computers & geosciences, 33 (4), 522–537

Emery X., (2008) Substitution Random Fields with Gaussian and Gamma Distributions: Theory and Application to a Pollution Data Set. Mathematical geosciences, 40 (1), 83–99

Emery X., Kremer F., (2008) A Survey of Random Field Models for Simulating Mineral Grades, in VII International Geostatistics Congress.

Goovaerts, P. (1997) Geostatistics for Natural Resources Evaluation, Oxford University Press, New York

Machuca-Mory D. F., Ortiz J. M., and Deutsch, C. V., (2008) On the challenge of using sequential indicator simulation for the estimation of recoverable reserves, International Journal of Mining, Reclamation and Environment, 22(4): 285-299

Mustapha D., Dimitrakopoulos R., (2010) Generalized Laguerre Expansions of Multivariate Probability Densities with Moments. Computers & mathematics with applications 60 (7), 2178–2189

Mustapha D., Dimitrakopoulos R., (2010b) High-order Stochastic Simulation of Complex Spatially Distributed Natural Phenomena. Mathematical geosciences. [Online] 42 (5), 457–485.

Ortiz J., (2004) A Step by Step Guide to Bi-Gaussian Disjunctive Kriging. In Geostatistics Banff; Springer Netherlands: Dordrecht; Vol. 14, pp 1097–1102

Ortiz J., (2019) MINE 834 Lectures. Queens University

Skovgaard M., (1986) On Multivariate Edgeworth Expansions International Statistical Review / Revue Internationale De Statistique, vol. 54, no. 2, pp. 169–186

Taylor Polynomial Expansions Website https://mathinsight.org/applet/taylor_polynomial accessed January 2021

Wallace D.L., (1958) Asymptotic Approximations to Distributions, The Annals of Mathematical Statistics, vol. 29, no. 3, pp. 635–654

Wilson G. A., Wragg A., (1973) Numerical methods for approximating continuous probability density functions using moments IMA J. Appl Math. 12 165173

Withers N., (2014) The Dual Multivariate Charlier and Edgeworth Expansions Statistics & probability letters 87, 76–85

Yao D., Dimitrakopoulos R., (2018) A New Computational Model of High-Order Stochastic Simulation Based on Spatial Legendre Moments, Mathematical geosciences 50 (8), 929–960

.