

## **Méthode de déformation graduelle**

### **1- Algorithme de base (Hu 2000)**

Soit 2 réalisations gaussiennes non-conditionnelles et de moyenne 0 de  $Z_0$  (pour  $Z_0(x)$ ) et  $Z_1$ .

On forme  
 $Z = \cos(t)Z_0 + \sin(t)Z_1$ .

On notera que  $\cos^2(t) + \sin^2(t) = 1$ . Cette condition est nécessaire pour que le nouveau  $Z$  ait même covariance que  $Z_0$  et  $Z_1$  et même moyenne 0. On peut choisir le paramètre «  $t$  » (scalaire) de façon à optimiser une fonction objectif. On peut ensuite itérer en générant  $Z_2$  et en le combinant à  $Z$ , pour former un nouveau  $Z$  et ainsi de suite. Au fur et à mesure des itérations, la fonction objectif ne peut que demeurer stable ou décroître.

Problème : après quelques itérations la fonction décroît très lentement car toute nouvelle réalisation a une très faible probabilité de permettre de se rapprocher de l'objectif. Une des raisons est que ce qui s'améliore en certains points se détériore ailleurs.

Idées possibles pour contrer le problème:

#### **i- Plusieurs réalisations simultanées**

Combiner d'un seul coup  $p > 2$  réalisations avec  $p-1$  paramètres à déterminer. Ainsi, avec  $p=3$ , on écrit :

$$Z = \cos(t_1)\cos(t_2) Z_0 + \sin(t_1) \cos(t_2) Z_1 + \sin(t_2) Z_2$$

La somme des 3 coefficients au carré donne 1.

Avec  $p+1$  réalisations on aura :

$$Z = a_0 Z_0 + a_1 Z_1 + \dots + a_p Z_p$$

où :

$$a_0 = \prod_{i=1}^p \cos(t_i), a_i = \sin(t_i) \prod_{j=i+1}^p \cos(t_j) \text{ et } a_p = \sin(t_p)$$

À nouveau, la somme des coefficients au carré donne 1.

#### **ii. Découper le domaine en plusieurs zones et déformer différemment chaque zone**

Ici, il faut se placer dans le cas non corrélé pour pouvoir modifier les valeurs gaussiennes séparément sans affecter la structure spatiale théorique. Bref, on ne va pas déformer  $Z$  directement mais le  $Y$  ( $N(0, I)$ ) à l'origine du  $Z$  (cas FFT-Ma ou LU) ou même la v.a. uniforme à l'origine de  $Z$  pour SGS.

Soit  $Z$  obtenu par FFT-Ma. On exprime  $Z = \text{FFT}^{-1} [\text{FFT}(Y) \text{FFT}(C(h))^{0.5}]$ .

Soit 2 champs de v.a.  $N(0,1)$  indépendantes  $Y_0$  et  $Y_1$ . On forme :

$$Y = \cos(t)Y_0 + \sin(t)Y_1.$$

On choisit la valeur de «  $t$  » telle que le  $Z$  obtenu permet de décroître la fonction objectif.

Pour plusieurs zones en  $Z$ , ( $Z_i$ ,  $i=1\dots p$ )  $Y_{0,i}$  et  $Y_{1,i}$ , on forme :

$$[Y_i] = [\cos(t_i)Y_{0,i} + \sin(t_i)Y_{1,i}]$$

On a donc  $p$  paramètres à optimiser (nombre de zones) mais on ne combine que 2 réalisations de nombres aléatoires indépendants à la fois. Pour chaque  $Y(t)$  obtenu, on calcule  $Z$  (par FFT-Ma) et la fonction objectif. Comme le FFT-Ma n'agit que localement lorsqu'on modifie une valeur, si l'on impose que tous les  $t_i$  soient 0 sauf  $t_k$  alors on modifie seulement la zone  $k$  et sa frontière immédiate.

Pour LU (non conditionnel ou conditionnel), le même raisonnement s'applique que pour FFT-Ma. À noter que les  $y_i$  correspondant à des données observées  $z_i$  doivent demeurer inchangés au cours de la fusion des divers champs.

Pour le SGS (voir Hu, 2001), on tire indépendamment à chaque étape d'une  $U(0,1)$ . On peut donc voir une réalisation donnée comme une transformation d'un champ  $U(0,1)$ . On peut aussi voir le champ  $U(0,1)$  comme étant  $F(N(0,1))$  ( $F$  la fonction de répartition). On peut donc associer un champ gaussien  $N(0,1)$  à chaque réalisation SGS, ce qui permet d'appliquer la procédure précédente.

Toute modification de  $Y$  à un point donné demande de recalculer la moyenne de toutes les distributions conditionnelles du SGS situées en aval du point perturbé dans la visite aléatoire des points à simuler. Il est donc avantageux de sauvegarder les poids du krigeage pour éviter d'avoir à tout refaire. Sinon, l'algorithme peut devenir lent.

## 2- Cas de données conditionnantes

Pour LU et SGS, le conditionnement est immédiat si l'on effectue la déformation graduelle sur le champ indépendant  $Y$  à la source de  $Z$  plutôt que sur le champ  $Z$  directement. Dans le cas LU, on n'a qu'à imposer les  $y_i$  correspondant aux  $z_i$  observés. Dans le cas du SGS, les points connus  $z_i$  ne font pas partie du trajet aléatoire, il n'y a donc pas lieu d'y associer de  $u_i$  et des  $y_i$ . On procède donc comme dans le cas non-conditionnel (la déformation se fait sur les  $Y$ , lesquels donnent  $U$  et finalement  $Z$ ). Pour toutes les méthodes (LU, SGS, FFT-MA, bandes tournantes,...) on peut aussi appliquer une des deux approches suivantes.

### i. Par post-conditionnement par krigeage (Hu, 2000)

On applique l'algorithme suivant une des variantes précédentes aux réalisations non-conditionnelles. À chaque itération, et pour chaque valeur du ou des paramètres ( $t$ ) que l'on considère, on post-conditionne les simulations avant d'évaluer la fonction objectif. Comme le post-conditionnement est en général très rapide (surtout sous forme duale), l'algorithme demeure réaliste.

### ii. Déformation graduelle directement sur des simulations conditionnelles (Hu, 2002)

Il faut considérer au moins 3 réalisations conditionnelles à la fois

On va conserver les valeurs conditionnantes et la structure spatiale si on forme :

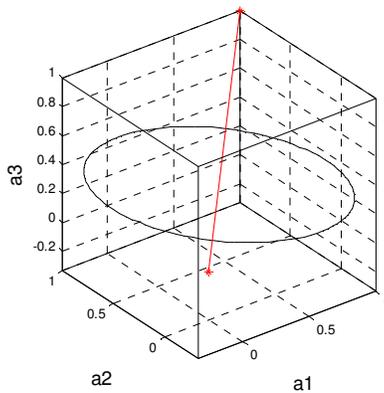
$$Z = a_1 Z_1 + a_2 Z_2 + a_3 Z_3$$

avec :  $a_1 + a_2 + a_3 = 1$   
 et  $a_1^2 + a_2^2 + a_3^2 = 1$

Ces équations imposent que les poids se retrouvent sur l'intersection d'une sphère unité avec le plan perpendiculaire au vecteur (1,1,1) et passant par les points (1,0,0), (0,1,0) et (0,0,1).

Hu (2002) propose la paramétrisation suivante :

$$\begin{aligned} a_1 &= 1/3 + 2/3 \cos(t) \\ a_2 &= 1/3 + 2/3 \sin(-\pi/6 + t) \\ a_3 &= 1/3 + 2/3 \sin(-\pi/6 - t) \end{aligned}$$



### 3- Une amélioration pour accélérer la convergence

Hu et al. (2004) proposent de combiner plusieurs réalisations pour en former une seule qui soit le plus susceptible d'améliorer la situation actuelle. Pour cela, ils s'inspirent des techniques d'optimisation basées sur le gradient de la fonction objectif.

Soit un état actuel. On peut calculer le gradient de la fonction objectif dans l'espace N-dimensionnel des N points simulés ( $dO/dZ_j$ ,  $j=1\dots N$ ). On peut aussi calculer le gradient en fonction des valeurs  $N(0,1)$  sous-jacentes à Z, soit  $v = dO/dY_j$  ( $j=1\dots N$ ) obtenu avec la réalisation actuelle de Y. On génère K nouvelles réalisations indépendantes de Y. On cherche la combinaison linéaire  $Y_u$  de ces K réalisations qui soit la plus proche possible de « v » (note si  $K \geq N$ , alors cette combinaison linéaire sera « v »). Il faut que la somme des coefficients au carré donne 1, mais on peut calculer directement ces coefficients sans passer par un processus d'optimisation. On définit :

$$Y_u = \sum_{i=1}^K a_i Y_i \text{ avec } \sum_{i=1}^K a_i^2 = 1. \text{ On veut minimiser la somme du carré des erreurs entre}$$

« v » et sa valeur approchée  $Y_u$ . On reconnaît là un problème de régression. Comme les réalisations sont indépendantes, on a  $b_i \propto v' Y_i$ . Pour que  $a'a = 1$ , on pose

$$a_i = \frac{b_i}{\sqrt{\sum_{i=1}^K v' Y_i}} .$$

pour respecter la contrainte.

Le reste de l'algorithme s'applique comme précédemment.  $u$  donne  $Y_u$ , qui donne  $Z_u$ . On fusionne ensuite le  $Z$  actuel avec ce  $Z_u$ . L'optimisation du paramètre  $t$  pour la fusion se fait comme dans l'algorithme de base, (soit sur  $Z$  et  $Z_u$  s'il n'y a pas de données conditionnantes, soit sur  $Y$  et  $Y_u$  s'il y en a avec une étape de post-conditionnement entre les deux).

Remarques :

- 1- La génération de  $K$  réalisations de  $Y$  est très rapide car il s'agit de simulations non-conditionnelles.
- 2- Le calcul des  $a_i$  est aussi très rapide.
- 3- La fusion est similaire à ce qui se fait habituellement. Elle requiert peu d'étapes pour chaque itération.
- 4- L'étape lente est le calcul du gradient (vecteur «  $v$  »). Pour certains problèmes, il existe des trucs très rapides (ex. méthode adjointe en hydrogéologie où l'on peut reformuler le problème pour obtenir toutes les dérivées en une seule itération du simulateur d'écoulement).
- 5- Le modèle direct (i.e. calcul de  $O$  étant donné  $Z$ ) est aussi une étape qui peut être lente, dépendant du problème. D'où l'importance d'appeler le modèle direct le moins souvent possible.

Références :

Hu L.Y, 2000. Gradual deformation and iterative calibration of gaussian related stochastic models. *Math. Geology*, 32, 1, 87-108.

Hu L.Y, Blanc et Noettinger, 2001. Gradual deformation and iterative calibration of sequential stochastic simulations. *Math. Geology*, 33, 4, 475-490.

Hu L.Y, 2002. Combination of dependent realizations within the gradual deformation method. *Math. Geology*, 34, 8, 953-964.

Hu L.Y, Le-Ravalec, M., 2001. An improved gradual deformation method for reconciling random and gradient searches in stochastic optimizations. *Math. Geology*, 36, 6, 703-720.

### **Méthode de simulation séquentielle d'indicatrice (SIS)**

Cette méthode est presque en tout point identique à la méthode SGS. La seule différence est dans la façon dont on obtient les distributions conditionnelles. Dans le SGS, la distribution conditionnelle est obtenue par KS sous une hypothèse de multinormalité. Dans le SIS, la distribution conditionnelle est obtenue par krigeage d'indicatrice (KI) ou cokrigeage d'indicatrice (CoKI) ou krigeage de probabilité (i.e. cokrigeage de  $I(x)$  et  $U(x)$ ,  $U(x)$  la variable uniforme associée au rang des valeurs observées (PK)). Pour chaque point visité, on effectue le krigeage d'indicatrices aux différents seuils, on corrige les relations d'ordre (le cas échéant) et l'on tire une valeur de la distribution.

Problème connu : La distribution est discrétisée à des seuils fixes et l'interpolation entre les seuils estimés génère du bruit. On se retrouve donc parfois avec une structure de petite échelle différente (effet de pépite trop grand par exemple) de ce que l'on voudrait avoir. De plus ce que l'on reproduit directement ce sont les variogrammes d'indicatrices, pas celui de  $Z$ . Rien ne garantit que le variogramme de  $Z$  est bien représenté. Finalement, la simulation n'est pas exacte aux points échantillons. En effet la valeur à un point échantillon tombe entre deux seuils. La probabilité sera 0 pour le seuil immédiatement inférieur et 1 pour le seuil immédiatement supérieur. Comme on interpole linéairement entre les deux seuils, la valeur estimée sera une valeur représentative de la classe entre les deux seuils, et non la valeur observée. La valeur simulée sera tirée aléatoirement sur cet intervalle.

### **Simulation de variables discrètes**

On désire maintenant simuler des variables discrètes représentant par exemple des faciès ou des types de roche.

#### **Méthode de simulation gaussienne tronquée**

##### **Cas non-conditionnel :**

L'idée est de générer un champ gaussien (on sait le faire) puis de tronquer la distribution gaussienne en la découpant en classes, chacune représentant une catégorie. Les seuils  $c_i$   $i=1...k$  définissant les  $k$  classes sont choisis de façon à avoir les bonnes proportions de chaque catégorie.

Exemple : Supposons que l'on simule une  $N(0,1)$  avec une certaine fonction de covariance. On a 3 types de roche (A, B et C) dans les proportions  $p_A=1/3$ ,  $p_B=1/2$ , et  $p_C=1/6$ . Les deux seuils associés aux trois types de roche peuvent être dépendants du choix de l'ordre des catégories :

S1	$c_i$	S2	$c_i$	S3	$c_i$	S4	$c_i$	S5	$c_i$	S6	$c_i$
A	-0.43	A	-0.43	B	0	B	0	C	-0.967	C	-0.967
B	0.967	C	0	A	0.967	C	0.43	A	0	B	0.43
C		B		C		A		B		A	

Une fois la simulation effectuée, on n'a qu'à coder en A, B ou C selon la valeur gaussienne simulée, l'ordre adopté et les seuils calculés. On notera que seulement trois codages différents sont possibles ici car S1 est équivalent à S6, S2 à S4 et S3 à S5 (il suffit de prendre  $-Z$  pour avoir exactement le même codage).

### Remarques :

- Comme la simulation est gaussienne, en absence d'effet de pépite et en présence de corrélation spatiale, il y a transition naturelle entre les catégories qui suit l'ordre adopté. Ainsi dans l'ordre S1 (A-B-C), il ne peut y avoir de contact direct spatialement entre la catégorie A et la catégorie C. Ces contraintes doivent guider le choix de l'ordre le plus approprié. Ainsi dans une séquence sédimentaire, si l'on a des contacts entre sable grossier (A) et sable fin (B) et entre sable fin (B) et argile (C) mais jamais de contact entre argile et sable grossier, on pourra choisir l'ordre A-B-C ou C-B-A indifféremment mais les autres seront interdits. Si tous les contacts sont possibles, la méthode gaussienne tronquée doit être rejetée.
- Marcotte et al. (2008) ont montré que les indicatrices de catégorie correspondant aux seuils dans les queues de la distribution d'une gaussienne montrent plus de corrélation spatiale que les indicatrices de catégorie autour de la médiane pour une même proportion.
- La difficulté principale est la détermination de la covariance du champ gaussien qu'il faut simuler. Pour la déterminer, on calcule les valeurs expérimentales et théoriques des variogrammes et variogrammes croisés (ou covariances et covariances croisées) des indicatrices de catégories. Les valeurs théoriques de ces fonctions peuvent toutes être dérivées des fonctions de répartition bigaussiennes. Ainsi, avec k faciès (1,...k) la covariance non-centrée des indicatrices est donnée par :

$$E[I_{c_i}(x)I_{c_j}(x+h)] = P(I_{c_i}(x)=1 \cap I_{c_j}(x+h)=1) = P(c_{i-1} < Z(x) \leq c_i \cap c_{j-1} < Z(x+h) \leq c_j)$$

avec  $c_0 = -\infty$  et  $c_k = \infty$ .  $I_{c_i}$  est l'indicatrice de la catégorie i.

- Si l'ajustement ne convient pas, on modifie les paramètres de la covariance gaussienne jusqu'à ce que l'on obtienne un ajustement satisfaisant. Si l'on ne réussit pas à avoir un ajustement satisfaisant c'est que le modèle « gaussien tronqué » est inadéquat.

### Cas conditionnel :

Normalement il y a des données conditionnantes, i.e. on a observé les types de roche ou faciès en un certain nombre de points (e.g. forages). Le problème est de simuler le champ gaussien que l'on va tronquer en s'assurant que les codes aux points échantillons vont correspondre aux observations. Bref, déterminer quelle valeur gaussienne associer à chaque valeur discrète observée.

On recourt à l'échantillonneur de Gibbs :

- On décide des seuils pour avoir le bon agencement des faciès et les bonnes proportions.
- On assigne arbitrairement aux points échantillons des valeurs gaussiennes tirées de l'intervalle correspondant à chaque faciès.
- On choisit un point au hasard. On retire ce point et on le krige par krigeage simple avec les autres points. On tire une valeur de la distribution  $N(Z_{KS}^*, \sigma_{KS}^2)$ . Si la valeur tombe dans l'intervalle, on la garde, sinon on remet l'ancienne valeur et on retourne choisir un autre point. Si l'algorithme est appliqué assez longtemps, on converge vers la bonne distribution.

Quand cette étape est terminée, on effectue une simulation conditionnelle classique par la méthode gaussienne de notre choix.

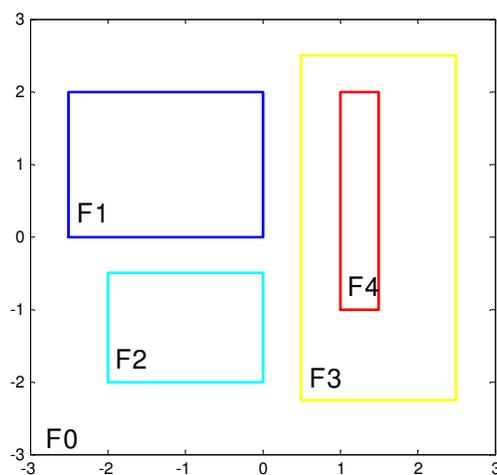
## Proportions variables

La méthode gaussienne tronquée a beaucoup été utilisée pour simuler des réservoirs pétroliers. Dans ce type d'applications, il n'est pas rare que les proportions des différents faciès varient par exemple en fonction de la profondeur.

On peut facilement adapter la méthode pour tenir compte de ce facteur. On simule toujours un seul champ gaussien, mais les seuils de codage deviennent fonction de la profondeur puisque les proportions de chaque faciès varient en fonction de celle-ci. Les covariances/variogrammes simples/croisés d'indicatrices deviennent alors non-stationnaire : ils dépendent de la profondeur des 2 points considérés. On peut intégrer ce fait dans le calcul des valeurs théoriques des variogrammes pour effectuer l'ajustement du modèle gaussien sur les indicatrices.

## Méthode de simulation plurigaussienne

Cette méthode est une généralisation de la précédente. L'idée est alors de simuler plus d'une variable gaussienne. Le plus souvent on se limite à 2 variables. Le codage des faciès est alors bidimensionnel, on parle alors du drapeau de la méthode. En voici un exemple :

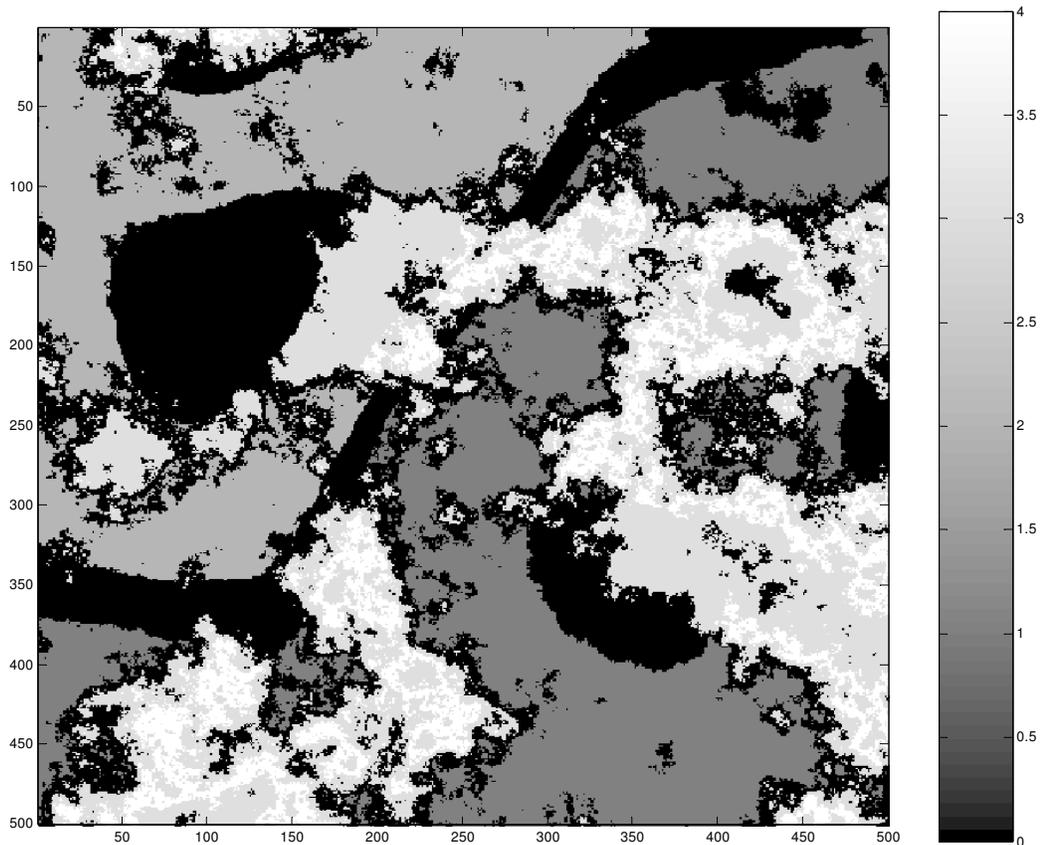


Les axes  $x$  et  $y$  représentent deux v.a. gaussiennes  $N(0,1)$ . Ces gaussiennes ont une corrélation «  $r$  » (qui peut être 0). Les rectangles découpent les zones de codage de chaque faciès (de F0 à F4, il y a donc 5 faciès). Le faciès F4 est en contact uniquement avec F3. Les faciès F1 et F2 ne sont en contact qu'avec F0. F3 est en contact soit avec F4, soit avec F0. Les proportions de chaque faciès s'obtiennent par intégration de la fonction de densité bigaussienne ayant corrélation «  $r$  » entre le point inférieur gauche et le point supérieur droit de chaque rectangle. Ainsi, pour F4, si les limites sont  $x_{41}$  et  $x_{42}$  selon  $x$  et  $y_{41}$  et  $y_{42}$  selon l'axe  $y$ , la proportion  $p_4$  est donnée par :

$$p_4 = F(x_{42}, y_{42}) + F(x_{41}, y_{41}) - F(x_{42}, y_{41}) - F(x_{41}, y_{42})$$
 où  $F$  est la fonction de répartition de la loi  $N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}\right)$ . Celle-ci peut facilement être évaluée dans Matlab par exemple.

Une fois le drapeau décrit, on simule le champ bigaussien (simulation spatiale). Il faut spécifier les covariances de chaque variable gaussienne et la covariance croisée (à  $h=0$ ,  $C_{XY}(0) = r$ ). Une fois la simulation bigaussienne effectuée, on effectue le codage selon les couples de valeurs en chaque point et le drapeau spécifié.

Exemple de simulation produite avec un drapeau similaire à la figure précédente :



Remarques :

- Il y a beaucoup de « degrés de liberté », et donc de flexibilité, dans cette méthode. Des proportions données et des relations entre faciès peuvent être obtenues par un nombre infini de drapeaux. Chaque drapeau peut fournir des images passablement différentes. Il faut spécifier aussi les covariances et covariances croisées des gaussiennes (incluant d'éventuelles anisotropies). Le réel problème est donc de déterminer le modèle correspondant à une image que l'on souhaite reproduire. On peut, comme précédemment, utiliser les covariances et covariances croisées d'indicatrices de catégories pour la modélisation, mais un nombre infini de choix et de combinaisons risque de fournir des ajustements pratiquement équivalents sur ces fonctions.
- Pour le conditionnement, on peut procéder de façon similaire au cas univariable. On utilise un échantillonneur de Gibbs. On génère initialement deux valeurs gaussiennes pour chaque

point tirées de la zone correspondante du drapeau. Ensuite, on tire un point au hasard, une variable au hasard. On retire cette variable. On cokrige (cokrigeage simple) la valeur gaussienne. On tire une valeur au hasard de la distribution conditionnelle ainsi obtenue. Si la valeur garde le point dans la bonne zone, on la garde, sinon, on remet l'ancienne valeur et on retourne sélectionner un nouveau point et une nouvelle variable. On répète le tout un très grand nombre de fois. La distribution finale converge vers la bonne distribution.

### ***Méthode de simulation séquentielle d'indicateur (SIS)***

Soit les indicateurs des faciès ou types de roche. Supposons que l'on a 3 faciès (A, B ou C). On effectue le krigage simple (ou cokrigeage simple d'indicateurs), pour chaque variable. Les valeurs krigées (ou cokrigées) estiment les probabilités que le faciès soit, en un point donné, A, B ou C. On s'assure que la somme de ces probabilités donne 1 en chaque point (c'est assuré si l'on a fait un cokrigeage en incluant l'effet de la contrainte dans la détermination des modèles de covariance/covariance croisée), sinon on normalise. On tire une valeur d'une  $U(0,1)$ . Sa valeur détermine le faciès à attribuer au point simulé. On inclut ce point simulé et on en visite un autre. On procède ainsi jusqu'à ce que tous les points aient été simulés. Chilès et Delfiner (1999, p. 522) démontrent que cet algorithme reproduit la bonne covariance d'indicateur et la bonne proportion lorsqu'il y a une seule variable indicatrice (i.e. 2 faciès).

Remarque : -tous les estimés dont la valeur est négative sont mis à 0 et ceux dont la valeur excède 1 sont ramenés à 1 avant de les normaliser.