Propriétés : (Box-Muller et simulation de vecteurs gaussiens)

a) Si U et V sont deux v.a. indépendantes de loi uniforme sur $[0;1]$ alors $X = \sqrt{-2\ln(U)}\cos(2\pi V)$ et $Y = \sqrt{-2\ln(U)}\sin(2\pi V)$ sont deux variables aléatoires indépendantes et telles que (X,Y) est un vecteur gaussien centré de matrice de covariance I_2.

b) Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $X = (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ un vecteur gaussien de matrice de covariance Γ et d'espérance le vecteur m.
Il existe une matrice A telle que $\Gamma = A\,{}^tA$ et si $Y = (Y_1, \ldots, Y_n)$ est un vecteur gaussien centré réduit alors $(AY + m)$ suit la même loi que X.

preuve :

a) C'est la méthode usuel pour déterminer la densité jointe d'un couple (X,Y). Soit $g : \mathbb{R}^2 \to \mathbb{R}$ une fonction continue et bornée et $\varphi : [0;1]^2 \to \mathbb{R}^2$ telle que $\varphi(u,v) = \left(\sqrt{-2\ln(u)}\cos(2\pi v), \sqrt{-2\ln(u)}\sin(2\pi v)\right)$.
On a :

$$\mathbb{E}(g(X,Y)) = \mathbb{E}(g \circ \varphi(U,V)) = \int_{[0;1]}\int_{[0;1]} g \circ \varphi(u,v)dudv$$

On va appliquer le théorème de changement de variable avec la fonction φ^{-1}. On a facilement : $u = e^{-\frac{x^2+y^2}{2}} = \varphi_1^{-1}(x,y)$ et $v = \frac{\arctan(y/x)}{2\pi} = \varphi_2^{-1}(x,y)$ si $x, y > 0$ (*dans les autres cas, selon le signe de x et y il faut ajouter une constante, changer un signe ou x et y mais cela ne change rien au jacobien*).
On a alors en notant abusivement $\varphi^{-1}(x,y) = \left(e^{-\frac{x^2+y^2}{2}}, \frac{\arctan(y/x)}{2\pi}\right)$:

$$Jac_{\varphi^{-1}}(x,y) = \begin{vmatrix} \frac{\partial\varphi_1^{-1}}{\partial x} & \frac{\partial\varphi_1^{-1}}{\partial y} \\ \frac{\partial\varphi_2^{-1}}{\partial x} & \frac{\partial\varphi_2^{-1}}{\partial y} \end{vmatrix} = \frac{1}{2\pi}\begin{vmatrix} -xe^{-\frac{x^2+y^2}{2}} & -ye^{-\frac{x^2+y^2}{2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{vmatrix}$$

Soit en prenant la valeur absolue du jacobien :

$$\int_{[0;1]}\int_{[0;1]} g \circ \varphi(u,v)dudv = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\frac{e^{-\frac{x^2+y^2}{2}}}{2\pi}dxdy$$

La densité jointe du couple (X,Y) s'écrit donc comme un produit tensoriel de deux densités marginales de loi normale centrée réduite $f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, X et Y sont donc indépendantes et de même loi. De plus, avec la densité de (X,Y) on voit que ce couple est gaussien centré de matrice de covariance I_2.

b) La matrice Γ est évidemment symétrique et de plus elle est positive car pour tout réels $\alpha_1, \ldots, \alpha_n$,

$$\sum_{i,j} \alpha_i\alpha_j\mathbb{E}\left((X_i - \mathbb{E}(X_i))(X_j - \mathbb{E}(X_j))\right) = \mathbb{E}\left[\left(\sum_{1\leqslant i\leqslant n} \alpha_i(X_i - \mathbb{E}(X_i))\right)^2\right] \geqslant 0$$

Donc, d'après le théorème spectral, il existe une matrice orthogonale P telle que ${}^tP\Gamma P = D$ avec $D = diag(\lambda_1, \ldots, \lambda_n)$ et où les $(\lambda_i)_{1\leqslant i\leqslant n}$ sont positifs ou nuls. On a alors en posant $D' = diag(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$ et $A = PD'$: $A\,{}^tA = PD'D'\,{}^tP = PD\,{}^tP = \Gamma$.
On peut remarquer aussi que le vecteur AY est bien un vecteur gaussien car toute combinaison linéaire de ses composantes est une combinaison linéaire des composantes de Y qui est gaussien. Reste à montrer que la matrice de covariance de AY est Γ on aura alors le résultat voulu pour $(AY + m)$ (par linéarité de l'espérance car $\mathbb{E}(AY) = 0$).
En notant $(AY)_i$ la i^{ieme} composante du vecteur AY, on a :
$\mathbb{E}\left((AY)_i(AY)_j\right) = \mathbb{E}\left(\left(\sum_{1\leqslant k\leqslant n} a_{i,k}Y_k\right)\left(\sum_{1\leqslant l\leqslant n} a_{j,l}Y_l\right)\right) = \sum_{1\leqslant k\leqslant n} a_{i,k}a_{j,k} = \Gamma_{i,j} = \mathbb{E}(X_iX_j)$, ce qui achève la démonstration.

José Gregorio : http://agregorio.net

Scilab :

Petit programme Scilab pour simuler un couple gaussien de matrice de covariance G :

```
A=chol(G);
U=rand(1,1000); V=rand(1,1000);
X=sqrt(-2*log(U)).*cos(2*%pi*V);
Y=sqrt(-2*log(U)).*sin(2*%pi*V);
C=A*[X;Y];
Z=C(1, :); T=C(2, :);
plot2d(Z,T,0);
```

Remarques :

1. La méthode décrite en a) pour simuler un vecteur gaussien à partir de variable uniforme sur $[0;1]$ est connue sous le nom de Box-muller, elle est très classique, pas très efficace s'il faut simuler beaucoup de gaussiennes mais elle a l'avantage d'être simple.

2. Toujours dans a), la remarque faite entre parenthèses et en italique est très importante, le jury peu (va ?) poser la question. Il faut l'avoir fait au propre au moins une fois de façon rigoureuse pour bien comprendre ce qu'il se passe et bien avoir en tête le théorème de changement de variable avec $\mathcal{C}^1$-difféomorphisme.

3. Dans b) on montre juste que la matrice de covariance de AY est Γ car on a montré avant que AY est gaussien et un vecteur gaussien est caractérisé par sa moyenne et sa matrice de covariance. Le fait que la matrice de covariance est symétrique et semi-définie positive est général et toujours vrai.

Thèmes abordés :

* Probabilités
* Covariance et indépendance
* Densité
* Simulation
* Théorème de changement de variable intégrale
* Matrice symétrique semi-définie positive et théorème spectral.

Bibliographie :

Pour le a), je n'ai pas trouvé de référence papier (on trouve des documents sur internet...). J'ai écrit cela suite à une petite parenthèse lors d'un TD de préparation à l'agrégation dirigé par M. Djalil Chafaï que je remercie au passage. Donc si jamais vous trouvez des coquilles, c'est de ma faute mais soyez gentils de me le dire que je puisse corriger.

Pour le b), "Probabilité. De la licence à l'agrégation" de Ph. Barbe et M. Ledoux, chez BELIN.

José Gregorio : http://agregorio.net