Processus aléatoires : généralités & exemples

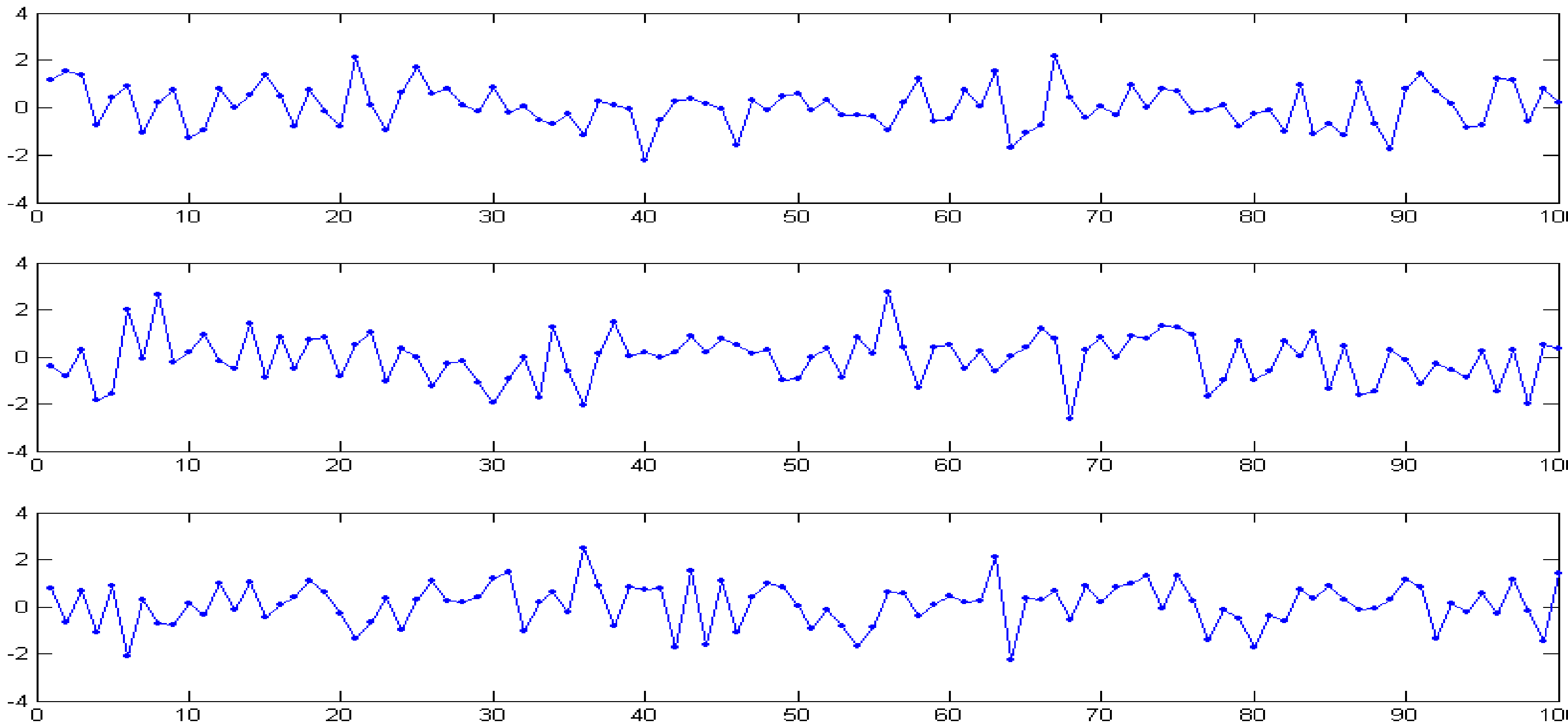
O. Roustant, 2020

Définition

- *Processus aléatoire, ou proc. stochastique :* famille de variables aléatoires $\{Z(x)\}_{x \in X}$
- « L'observable » : *réalisations*, *trajectoires*
- Un même processus, une infinité de réalisations

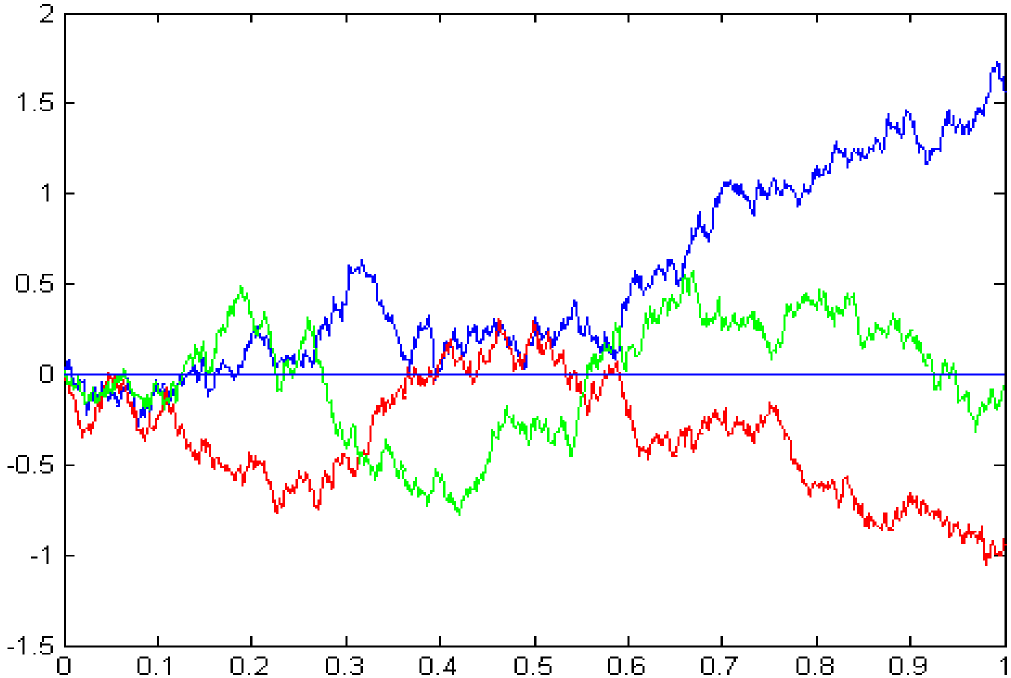
Exemples (1)

- Quelques réalisations d’un *bruit blanc*



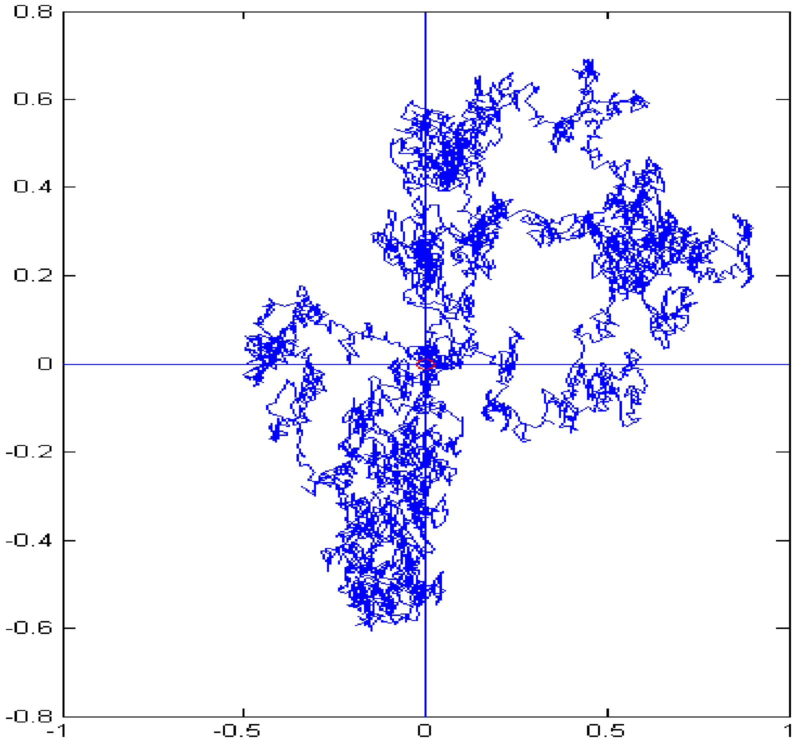
Exemples (2)

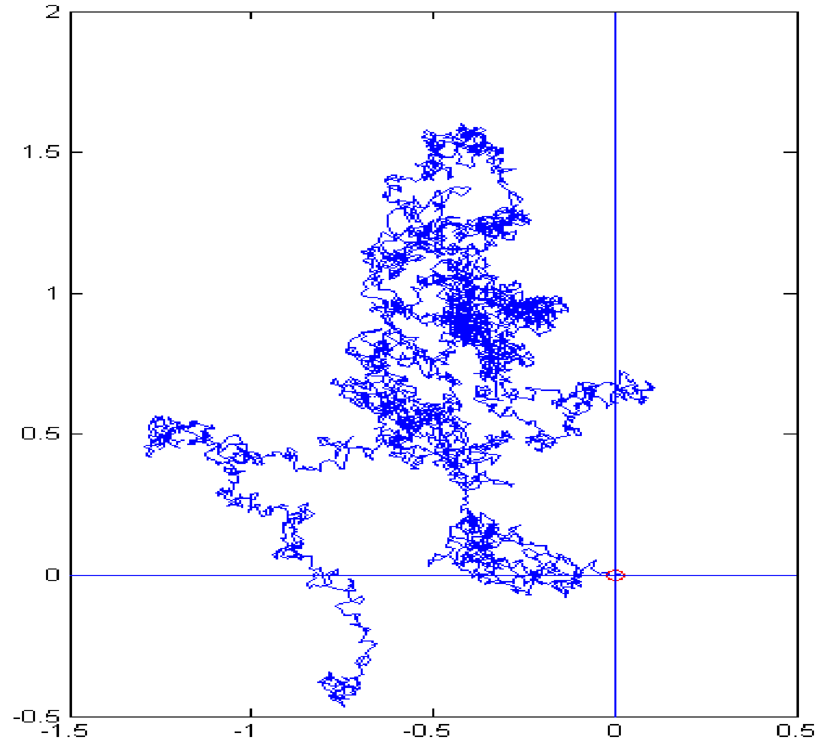
- Quelques réalisations d’un *mouvement brownien* en dimension 1…



Exemples (3)

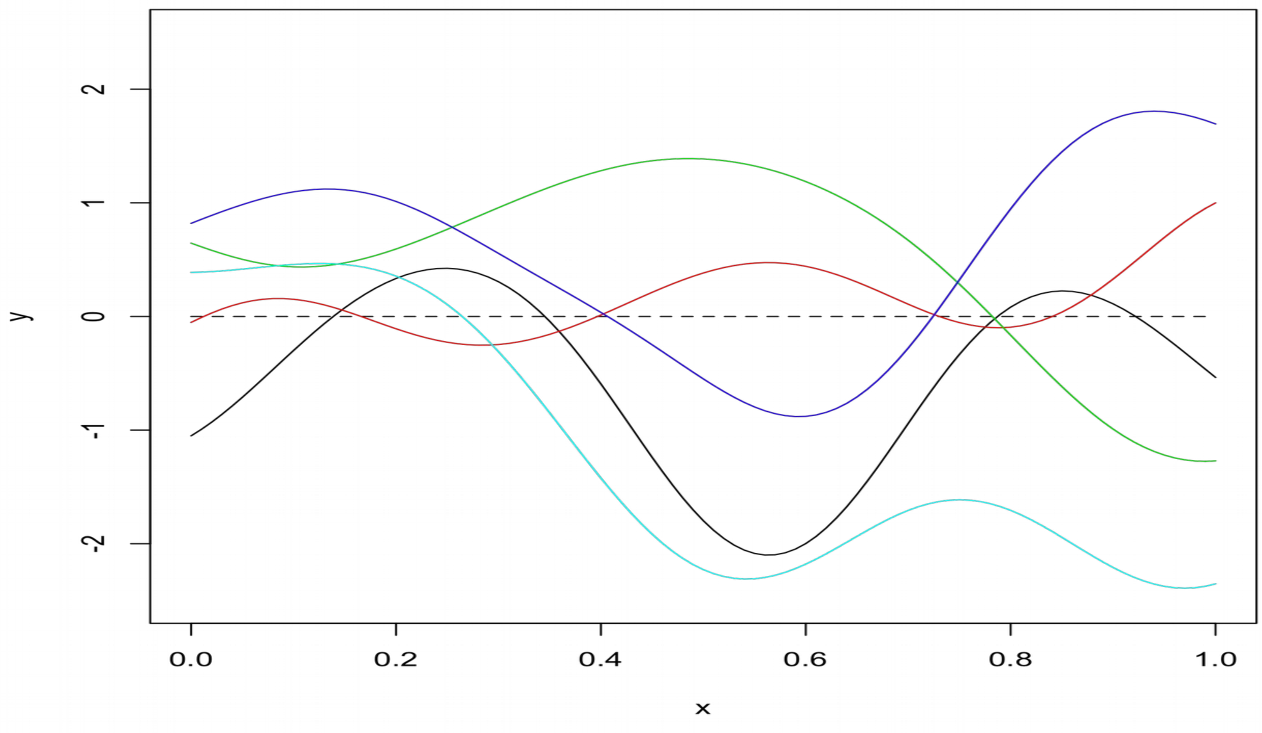
- … et en dimension 2





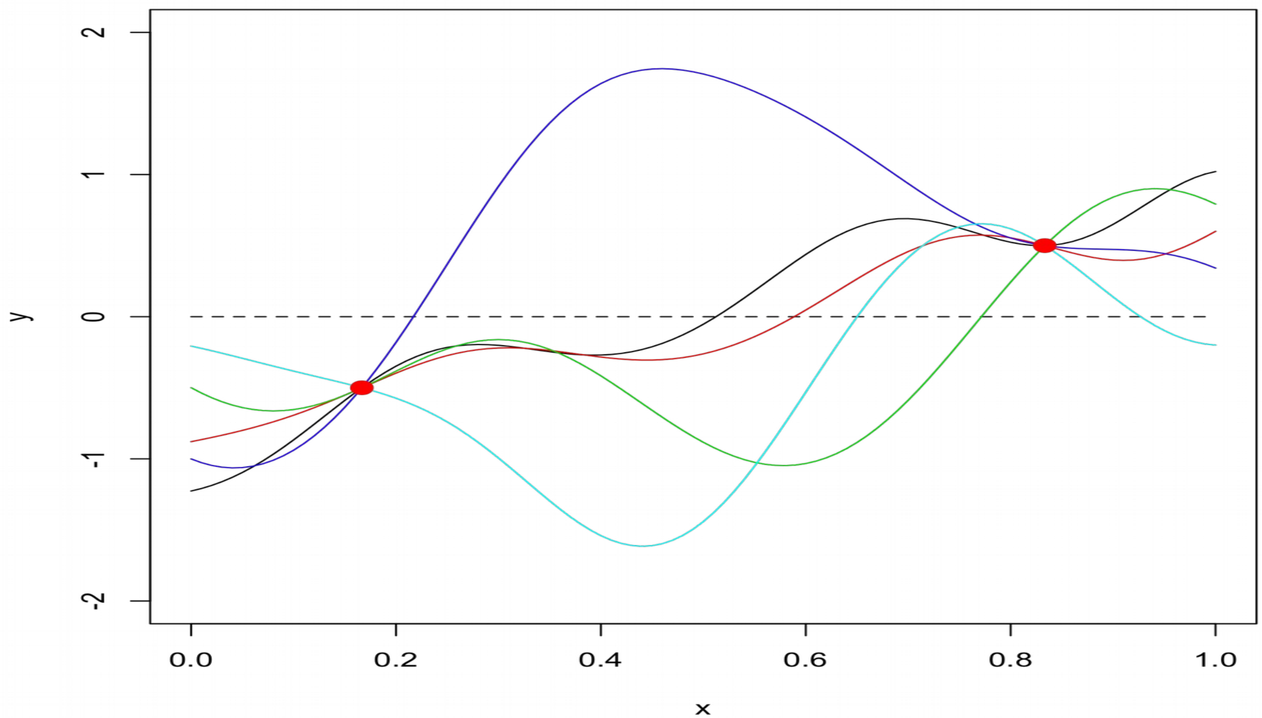
Exemples (4)

- Quelques réalisations d'un *processus gaussien*, avec trajectoires C^∞



Exemples (5)

- Le même, mais *conditionnel* à :
  $Z(x^{(1)}) = -0.5$ et $Z(x^{(2)}) = 0.5$



Processus du second ordre

- Un processus $(Z(x))$ est :
  - *centré* ssi pour tout x, $E(|Z(x)|)<+\infty$ et $E(Z(x))=0$
  - du *second ordre* ssi pour tout x, $E(Z(x)^2)<+\infty$
  - $\Rightarrow$ p.t. x,y, $E(|Z(x)|)<+\infty$ et $\mathrm{cov}(Z(x),Z(y))<+\infty$

<u>*Rappel*</u>

$\mathrm{cov}(X,Y) = E(\,(X-E(X))(Y-E(Y))\,)$	*covariance*
$\mathrm{var}(X) = \mathrm{cov}(X,X) = E(\,(X-E(X))^2\,)$	*variance*
$s(X) = (\,\mathrm{var}(X)\,)^{1/2}$	*écart-type*
$\mathrm{corr}(X,Y) = \mathrm{cov}(\,X/s(X),Y/s(Y)\,)$	*corrélation*

Cadre géométrique

- *(Espace L^2)*
  - L'espace vectoriel des v.a. qui ont un moment d'ordre 2 est un espace de Hilbert pour le produit scalaire $\langle X,Y \rangle = E(XY)$
  - Interprétation de la covariance et de la corrélation :

$$\mathrm{cov}(X,Y) = \langle X - \mu_X, Y - \mu_Y \rangle$$

$$\mathrm{cor}(X,Y) = \cos(X-\mu_X, Y-\mu_Y)$$

avec $\mu_X = E(X)$ et $\mu_Y = E(Y)$

Non corrélation et indépendance

- *Définition : X et Y sont non corrélées ou orthogonales si*

$$cov(X,Y)=0$$

- *Proposition :*
  - *si X et Y sont indépendantes, alors X et Y sont non corrélées*
  - *La réciproque est fausse !!*
    - *Contre-exemple :* $X = U$, $Y = U^2$ *avec U uniforme sur [-1,1] (ou plus généralement de densité paire sur [-a, a])*

- *Moralité : la covariance est un indicateur de dépendance* ***linéaire***

Stationnarité (1)

Le processus $(Z(x))$ est :

- *fortement stationnaire* ssi :
  - $(Z(x_1),\dots,Z(x_L))$ et $(Z(x_1+h),\dots,Z(x_L+h))$ ont la même loi, pour tous $x_1, \dots, x_L$ et pour tout h
- *faiblement stationnaire* ssi :
  - $E(Z(x))$, $E(Z(x)^2)$ et $\mathrm{cov}(Z(x),Z(x-h))$ sont finis et ne dépendent pas de x, pour tout x et pour tout h

Stationnarité (2)

- Relations entre les deux notions :
  - Fortement stationnaire $\Rightarrow$ faiblement stat.
  - La réciproque est fausse en général… mais vraie pour les processus *gaussiens*

Autocovariance, autocorrélation (1)

Pour un processus stationnaire d'ordre 2, on définit :

- La *fonction d'autocovariance* :
  - $\gamma(h) = \mathrm{cov}(Z(x), Z(x-h))$
- La *fonction d'autocorrélation* :
  - $\rho(h) = \mathrm{corr}(Z(x), Z(x-h))$

Autocovariance, autocorrélation (2)

- Propriétés :
  - $\rho(h) = \gamma(h)/\ \gamma(0)$
  - $\gamma(-h) = \gamma(h)$
  - $|\rho(h)| \leq 1$
- Interprétation géométrique :
  - $\rho(h)$ est le cosinus de l'angle formé par

    Z(x)-m et Z(x-h)-m

    où m = E(Z(x)) = E(Z(x-h))

Espérance conditionnelle (1)

- Définition dans L^2 (vision géométrique)
- *L'espérance conditionnelle (linéaire)* de X sachant Y, est une v.a. notée

$$E(X|Y_1,\ldots,Y_k), \text{ (resp. } E_L(X|Y_1,\ldots,Y_k))$$

qui est la meilleure approximation de X comme fonction (resp. comb. linéaire) de $Y_1,\ldots,Y_k$ au sens L^2 :

$$E(X|\, Y_1,\ldots,Y_k) = \text{argmin } (\|X-f(Y_1,\ldots,Y_k)\|_{L2})$$

Espérance conditionnelle (2)

- Il s'agit donc de projections orthogonales. Pour cette raison, on appelle souvent :
  - $E_L(X|Y)$ *régression linéaire* de X sur Y
  - $E(X|Y)$ *régression (non linéaire)* de X sur Y
- Cas particulier remarquable : si (X,Y) est un *vecteur gaussien*, les régressions linéaires et non linéaires coïncident

Espérance conditionnelle (3)

- Quelques propriétés
  - Linéarité : $E(aX+b|Y) = aE(X|Y) + b$
  - $EX = E(E(X|Y))$
  - $E(f(Y) \mid Y) = f(Y)$
  - $E(Xf(Y) \mid Y) = f(Y) E(X|Y)$
  - Si Z est relative à une information contenue dans Y, $E(X \mid Z) = E(E(X \mid Y) \mid Z)$
  - Si X et Y sont indépendants, $E(X|Y) = E(X)$
  - …

Variance conditionnelle

- Définition :

$$\mathrm{var}[f(X)|\ Y_1,\ldots,Y_k] = E([X-E(X|\ Y_1,\ldots,Y_k)]^2\ |\ Y_1,\ldots,Y_k)$$

- Propriété (Pythagore !):

$$\mathrm{var}(f(X)) = E(\mathrm{var}[f(X)|\ Y_1,\ldots,Y_k])$$
$$+ \mathrm{var}(E(f(X)|\ Y_1,\ldots,Y_k))$$

Loi conditionnelle

- On admet qu'il existe une probabilité P telle que pour toute fonction f,

$$E[f(X)|\, Y_1,\ldots,Y_k] = \int f(x)\,dP(x)$$

  P est appelé *loi conditionnelle* de X sachant $Y_1, \ldots, Y_k$, et notée $d\mu_{X|Y1,\ldots,Yk}$

- En particulier,
  - $E[X|\, Y_1,\ldots,Y_k] = \int x\,d\mu_{X|Y_1,\ldots,Y_k}(x)$

Processus gaussien (1)

- Rappel. La densité de la *loi normale d-dimensionnelle* est donnée par :

$$f_{\mu,\Sigma}(x_1,\dots,x_d) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(\Sigma)}}\exp\left(-\frac{1}{2}(X-\mu)'\Sigma^{-1}(X-\mu)\right)$$

avec

  - $X = (x_1, \dots, x_d)'$,
  - μ un vecteur d*1, s'interprétant comme la moyenne
  - Σ une matrice symétrique (définie) positive, s'interprétant comme la matrice de covariance

Processus gaussien (2)

- En fait, la densité est définie uniquement en fonction des 2 premiers moments, comme en dim. 1:
  - *Dans le cas d'indépendance*,

$$f_{(X1,\ldots,Xd)}(x_1,\ldots,x_d) = f_{(X1)}(x_1)*\ldots* f_{(Xd)}(x_d)$$

  - *Dans le cas général*, on exige que la loi d'un vecteur gaussien centré et réduit

$$(\Sigma^{1/2})^{-1}(X-\mu)$$

    soit la même que dans le cas de l'indépendance

Conséquence immédiate : orthogonalité ⇒ indépendance

Processus gaussien (3)

- Un processus $(Z(x))$ est *gaussien* ssi
  - pour tous $x_1, \ldots, x_n$, la loi du vecteur $(Z(x_1), \ldots, Z(x_n))$ est gaussienne

  ou, de façon équivalente, ssi :
  - pour tous $x_1, \ldots, x_n$, la loi de toute combinaison linéaire de $(Z(x_1), \ldots, Z(x_n))$ est gaussienne

Processus gaussien (4)

- ATTENTION !
  - si X et Y sont de loi normale et indépendantes, alors X+Y est de loi normale
  - si on oublie l'indépendance, le résultat est faux !
    - Contre-exemple : X et εX, où X est de loi normale et ε est le résultat obtenu au jeu de pile ou face avec une pièce équilibrée, en supposant ε et X indépendantes

Processus gaussien (5)

- Pour un processus gaussien:
  - Faiblement stationnaire ⇒ stationnaire
  - Orthogonalité ⇒ indépendance
  - E cond. Linéaire = E conditionnelle
  - Les lois conditionnelles sont gaussiennes :

  Si $X=[X_1; X_2]$ est gaussien, de moyenne $[m_1; m_2]$, et de matrice de covariance $S = [S_{11}\ S_{12}; S_{21}\ S_{22}]$, alors $X_1|X_2$ est gaussien, avec :

  $$E(X_1|X_2) = m_1 + S_{12}(S_{22})^{-1}(X_2 - m_2)$$

  $$Cov(X_1|X_2) = S_{11} - S_{12}(S_{22})^{-1} S_{21}$$

Processus gaussien (6)

- Comment simuler un processus gaussien ?
- On utilise la *décomposition de Choleski* : toute matrice symétrique définie positive admet la décomposition (unique)

$$S = T'*T$$

avec T matrice triangulaire supérieure

- Puis on remarque que si E est de loi normale $N(0, I_n)$, alors T'.E est de loi N(0,S)

Mouvement brownien (1)

- Définition 1 : on appelle *mouvement brownien* (B_t) tout processus vérifiant :
  - (B_t) est un *processus à accroissements indépendants et stationnaires (PAIS)* : pour tout instant t et tout horizon h, l'accroissement $(B_{t+h}-B_t)$ est indépendant des variables B_s, $s \leq t$, et de loi indépendante de t
  - Les trajectoires sont continues

Mouvement brownien (2)

- Définition 2 : on appelle *mouvement brownien standard* (W_t), ou *processus de Wiener*, un processus défini par :
  - $W_0 = 0$
  - (W_t) est un processus gaussien
  - (W_t) est centré, et $\mathrm{cov}(W_s,W_t)=\min(s,t)$
  - Les trajectoires sont continues

Mouvement brownien (3)

- Il est facile de voir que tout processus de Wiener est un mouvement brownien
- Réciproquement (plus difficile), si (B_t) est un mouvement brownien, alors il existe a, b et $\sigma>0$, tels que

$$B_t = a + bt + \sigma W_t$$

où W_t est un processus de Wiener

Mouvement brownien (4)

- Simulation : en discret, il s'agit d'une *marche au hasard*

$$W_{nh} = W_0 + \varepsilon_1 + \ldots + \varepsilon_n$$

- Donc simulation par intégration d'un bruit blanc gaussien

Mouvement brownien (5)

- *Mouvement brownien multi-dimensionnel* : il s'agit d'un processus dont les coordonnées sont des mouvements browniens (standard) indépendants

$$\mathbf{W}_t = (W_{1,t},\ldots,W_{d,t})$$

Pont brownien (1)

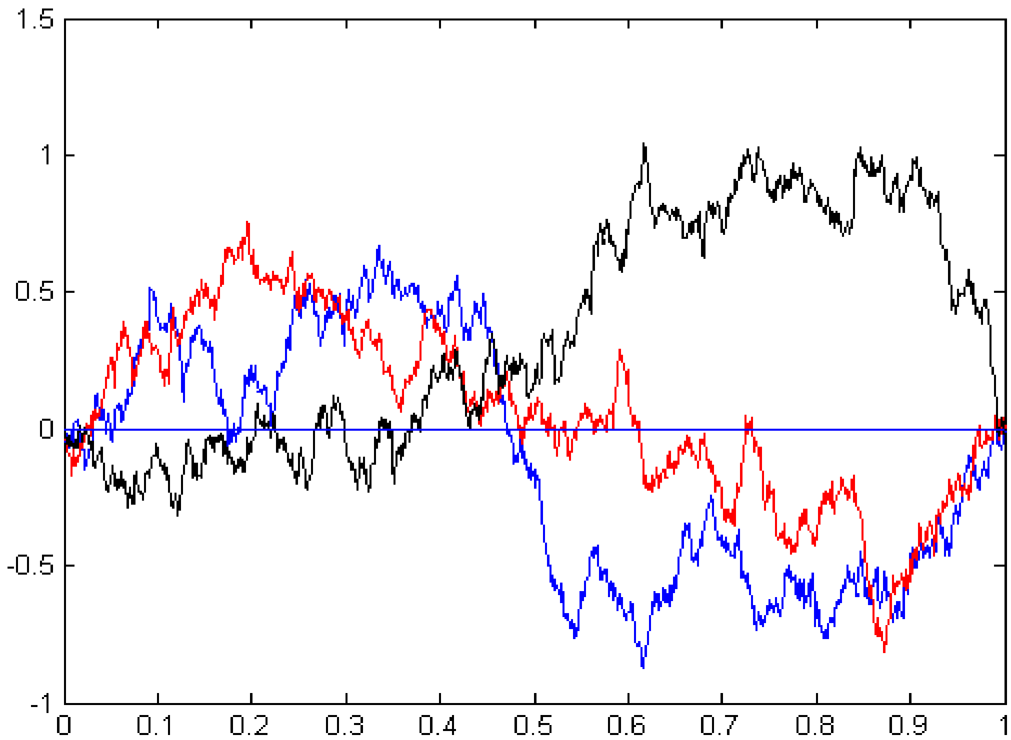
- On appelle *pont brownien* le processus (P_t) défini sur [0,1] par :

$$P_t = W_t - t.W_1$$

  où (W_t) est un mouvement brownien standard.
- Le processus est contraint de passer par 0 aux bords ⇒ interpolation

Pont brownien (2)

- Quelques réalisations :



Pont brownien (3)

- Quelques propriétés
  - (P_t) est un processus gaussien
  - P_t est de loi $N(0, t(1-t))$
  - $\mathrm{cov}(P_s,P_t) = \min(s,t) - st$

Modèle de krigeage (1)

- $y(x) = \mu + Z(x)$

  μ : moyenne de $y(x)$, constante

  $Z(x)$: processus gaussien stationnaire centré, de matrice de covariance $\sigma^2 R$, avec $R(h)$ décroissant / h

<u>Exemples :</u>

- $R(x,y) = \exp(-[\theta_1|x_1-y_1|^{p(1)} + \ldots + \theta_d|x_d-y_d|^{p(d)}])$
  où $p(1), \ldots, p(d) \in]0,2]$
- Si $p(i) = 2$, les réalisations sont de classe C^∞

Modèle de krigeage (2)

- Simulation : il s’agit d’un processus gaussien particulier…
- Simulation du processus conditionnel : c’est encore un processus gaussien

Modèle de krigeage (3)

- Prédiction : dans le cas où μ, σ et R sont connus,

  $$\text{Ypred}(x) = E[y(x)|y(x^{(1)}), \ldots, y(x^{(n)})]$$
  $$= \mu + r'R^{-1}(y(x^{(1)}) - \mu, \ldots, y(x^{(n)}) - \mu)'$$

  avec $r = (R(x,x^{(1)}), \ldots, R(x,x^{(n)}))'$

- En pratique, on remplace μ, σ et R par leurs estimations (à suivre…)

Modèle de krigeage (4)

- Exemple en dimension 1, avec $\theta = 30$