
—Texte—

Vitesse et position d'une particule

Mots-clefs : vecteurs gaussiens, maximum de vraisemblance, martingale.

1 Le phénomène observé et un peu d'histoire

En 1828, un botaniste du nom de Robert Brown rend compte de ses observations : une particule de pollen (observable au microscope) plongée dans l'eau est animée de mouvements aussi rapides qu'irréguliers. De nombreuses explications plus ou moins farfelues furent proposées pour rendre compte de ce phénomène mais peu à peu une thèse s'impose : le mouvement serait lié aux chocs de la particule de pollen avec les molécules d'eau. À la fin du 19^e siècle, les points suivants avaient été établis empiriquement :

1. le mouvement est très irrégulier et la trajectoire semble ne pas admettre de tangente,
2. deux particules (de pollen) semblent se déplacer indépendamment (si elles ne se choquent pas), même si la distance les séparant ne dépasse pas leur diamètre,
3. le mouvement est d'autant plus rapide que la particule est petite,
4. la composition et la densité de la particule n'ont aucun effet,
5. le mouvement est d'autant plus lent que le fluide est visqueux,
6. le mouvement est d'autant plus rapide que la température est élevée,
7. le mouvement est incessant.

Remarquons que le dernier point fut admis après une expérience de trente ans et l'observation de liquides prisonniers de quartz vieux de plusieurs milliers d'années.

L'intuition qui précède la modélisation est la suivante. La particule de pollen est « énorme » par rapport aux molécules du fluide. Nous dirons que les déplacements observés à l'échelle de la particule de pollen sont *macroscopiques* : ils peuvent être observés, au moins à l'aide d'une bonne loupe, et mesurés. Les molécules qui composent le fluide sont quant à elles beaucoup trop petites et surtout beaucoup trop nombreuses pour être observées une à une. Nous parlerons d'échelle *microscopique* pour les mouvements de ces molécules. À chaque instant, d'innombrables chocs ont lieu entre la particule de pollen et les molécules qui l'entourent. Ces chocs se répartissent en moyenne de manière uniforme sur la surface de la particule. Pourtant à un instant donné il peut y avoir plus de chocs sur une partie que sur une autre.

Nombre d'historiens et surtout des physiciens, présentent les travaux d'A. Einstein, en 1905, comme les premiers pas mathématiques de ce qui va devenir la théorie du mouvement brownien. Citons au passage N. Wiener qui fournira l'essentiel du travail de la

mathématisation de la théorie du mouvement brownien (on parle aussi de processus de Wiener). Cependant, un peu avant les travaux d'Einstein, et dans un tout autre contexte, L. Bachelier avait déjà obtenu, en 1900, la loi du mouvement brownien dans sa thèse intitulée « La théorie de la spéculation ». Le modèle étudié ici a été introduit par P. Langevin et perfectionné par L.S. Ornstein et G.E. Uhlenbeck dans les années 1930.

2 Distribution du couple position et vitesse à temps discret

Proposons à présent un modèle mathématique à temps discret pour la dynamique d'une particule. Discrétisons le temps avec un pas $\alpha > 0$. Nous noterons X_n et V_n la position et la vitesse de la particule à l'instant $n\alpha$.

L'évolution de la vitesse est donnée par une version à temps discret du principe fondamental de la dynamique qui assure que l'accélération (ou la variation de la vitesse) d'une particule est proportionnelle à la somme des forces qui s'exercent sur elle. Elles sont au nombre de quatre (les deux premières se compensant) : le poids, la poussée d'Archimède, la force de frottement due au déplacement dans un liquide et l'action des chocs des molécules d'eau sur la particule.

Nous supposons que l'intensité de la force de frottement est proportionnelle à la vitesse de la particule : si la vitesse est v , la force de frottement est $-\lambda v$. La modélisation de l'effet des chocs des molécules d'eau demande un peu plus d'attention. Pendant un intervalle de temps α , la particule de pollen subit un nombre de chocs proportionnel à α , disons $N\alpha$ où N , le nombre de chocs dans une unité de temps, est supposé très grand. Les chocs $(\varepsilon_i)_i$ sont supposés indépendants et identiquement distribués d'intensité aléatoire centrée de carré intégrable (de variance σ^2)

$$V_{n+1} - V_n = -\lambda V_n \alpha + \sum_{i=1}^{N\alpha} \varepsilon_i = -\lambda V_n \alpha + \underbrace{\sqrt{N\sigma^2\alpha} \frac{1}{\sqrt{N\sigma^2\alpha}} \sum_{i=1}^{N\alpha} \varepsilon_i}_{Y_{n+1}}$$

Pour peu que $N\alpha$ soit grand, le théorème limite central suggère que la loi de Y_{n+1} soit (proche d')une loi gaussienne centrée réduite. De plus, ces forces aléatoires seront supposées indépendantes : les molécules d'eau responsables des déplacements de la particule aux temps $n\alpha$ et $(n+1)\alpha$ ne sont pas les mêmes car la particule se déplace beaucoup plus vite que les molécules d'eau.

Nous modéliserons la variation de la vitesse de la manière suivante :

$$V_{n+1} - V_n = -\lambda\alpha V_n + \sqrt{2D\lambda^2\alpha} Y_{n+1},$$

où les variables $(Y_n)_n$ sont des variables aléatoires indépendantes et identiquement distribuées de loi gaussienne centrée réduite. Le paramètre $\lambda > 0$ représente le coefficient de frottement tandis que $2D\lambda^2\alpha$ représente la variance infinitésimale de la variation aléatoire

de la vitesse de la particule¹. Ils seront interprétés dans la suite. Adoptons les notations suivantes :

$$a = 1 - \alpha\lambda \quad \text{et} \quad b = \sqrt{2D\lambda^2\alpha}.$$

Nous supposons que le pas de discrétisation α est assez petit pour que a soit positif (il est aussi strictement inférieur à 1). On peut alors réécrire l'évolution des vitesses de la façon suivante :

$$V_{n+1} = aV_n + bY_{n+1}. \quad (1)$$

L'évolution de la position entre les instants αn et $\alpha(n+1)$ est quant à elle donnée par une version discrète de la définition de la vitesse :

$$X_{n+1} - X_n = \alpha V_n.$$

Grâce aux propriétés de stabilité des lois gaussiennes, il est possible de décrire la loi du couple (X_n, V_n) .

Proposition 2.1. *Lorsque $X_0 = x$ et $V_0 = v$, la loi de (X_n, V_n) est un vecteur gaussien de moyenne*

$$\mathbb{E}(X_n) = x + v \frac{1 - a^n}{\lambda} \quad \text{et} \quad \mathbb{E}(V_n) = va^n,$$

et de matrice de covariance définie par :

$$\begin{aligned} \text{Var}(X_n) &= \frac{\alpha^2 b^2}{(1-a)^2} \left[n - 1 - \frac{2+a}{1-a^2} + \frac{2a^n}{1-a} - \frac{a^{2n}}{1-a^2} \right], \\ \text{Var}(V_n) &= \frac{b^2}{1-a^2} (1 - a^{2n}), \\ \text{Cov}(X_n, V_n) &= \frac{\alpha a b^2}{(1-a)^2} \left[\frac{a}{1+a} - a^n + \frac{a^{2n}}{1+a} \right]. \end{aligned}$$

Remarque 2.2. Pour $n_1 < \dots < n_k$ entiers, on peut également déterminer la distribution du k -uplet $(V_{n_1}, \dots, V_{n_k})$ en remarquant qu'il s'agit d'un vecteur gaussien dont on peut calculer la moyenne et la matrice de covariance.

3 Passage au temps continu

On souhaite à présent comprendre le phénomène à temps continu. Soit $t > 0$. Pour tout $N \geq 1$, considérons la suite $(X_n^{(N)}, V_n^{(N)})_{n \geq 0}$ définie par

$$V_{n+1}^{(N)} = a_N V_n^{(N)} + b_N Y_{n+1}^{(N)} \quad \text{et} \quad X_{n+1}^{(N)} = X_n^{(N)} + \alpha_N V_n^{(N)},$$

¹Le fait de noter $2D\lambda^2$ la constante est arbitraire et s'expliquera par une jolie formule plus loin...

où $a_N = 1 - \lambda/N$ et $b_N = \sqrt{\lambda^2 D/N}$ (c'est-à-dire que $\alpha = 1/N$). On lui associe le processus $(\bar{X}_t^{(N)}, \bar{V}_t^{(N)})_{t \geq 0}$ défini par

$$\bar{V}_t^{(N)} = V_{[tN]}^{(N)} \quad \text{et} \quad \bar{X}_t^{(N)} = X_{[tN]}^{(N)},$$

où $[x]$ désigne la partie entière de $x \in \mathbb{R}$.

Remarque 3.1. Pour N fixé, on a découpé le temps en intervalles de longueur $1/N$. Le temps t pour le processus $(\bar{X}_t^{(N)}, \bar{V}_t^{(N)})_{t \geq 0}$ correspond au temps $[tN]$ pour la suite $(X_n^{(N)}, V_n^{(N)})_{n \geq 0}$.

Proposition 3.2. *Pour tout $t \geq 0$, la suite de couples aléatoires $(\bar{X}_t^{(N)}, \bar{V}_t^{(N)})$ converge en loi vers un vecteur gaussien (\bar{X}_t, \bar{V}_t) de moyenne*

$$\mathbb{E}(\bar{X}_t) = x + v \frac{1 - e^{-\lambda t}}{\lambda} \quad \text{et} \quad \mathbb{E}(\bar{V}_t) = v e^{-\lambda t}$$

et de matrice de covariance

$$K_t = \begin{pmatrix} 2Dt + \frac{D}{\lambda} [-3 + 4e^{-\lambda t} - e^{-2\lambda t}] & D(1 - e^{-\lambda t}) \\ D(1 - e^{-\lambda t}) & \lambda D(1 - e^{-2\lambda t}) \end{pmatrix}.$$

Remarque 3.3. Lorsque t tend vers l'infini, la distribution de \bar{V}_t converge vers la loi $\mathcal{N}(0, \lambda D)$. La vitesse de convergence à l'équilibre est exponentielle et dépend de λ . On appelle parfois la constante λ^{-1} temps de relaxation. Pour fixer les idées, λ^{-1} est en pratique de l'ordre de 10^{-8} seconde. La quantité λ est donc homogène à une fréquence (inverse du temps) et est donc mesurée en s^{-1} par exemple. La constante D quant à elle est appelée diffusivité et est mesurée en $m^2 s^{-1}$. Citons quelques exemples numériques : la diffusivité de l'hémoglobine dans le sang est de l'ordre de $10^{-7} cm^2 \cdot s^{-1}$, tandis que celle de l'oxygène, toujours dans le sang, est de l'ordre de $10^{-5} cm^2 \cdot s^{-1}$.

4 Estimation des paramètres du modèle

On suppose dans cette section que l'on a observé les positions de la particule de pollen à des instants régulièrement espacés : $0, \alpha, 2\alpha, \dots, (n+1)\alpha$. On souhaite proposer une méthode d'estimation des paramètres λ et D . On peut déduire des positions successives observées x_0, \dots, x_n les vitesses v_0, \dots, v_n . En utilisant les notations précédentes, on peut donc ramener le problème initial à celui de l'estimation des paramètres a et b de la relation (1) à partir des observations des vitesses. On supposera que $V_0 = 0$.

La vraisemblance du modèle est donnée par :

$$L(a, \sigma, V_1, \dots, V_n) = \frac{1}{\sqrt{2\pi b^2}^n} \prod_{k=1}^n \exp\left(-\frac{(V_k - aV_{k-1})^2}{2b^2}\right). \quad (2)$$

Proposition 4.1. *Les estimateurs obtenus par la méthode du maximum de vraisemblance sont*

$$\hat{a}_n = \frac{\sum_{k=1}^n V_{k-1} V_k}{\sum_{k=1}^n V_{k-1}^2},$$

$$\hat{b}_n^2 = \frac{1}{n} \sum_{k=1}^n (V_k - \hat{a}_n V_{k-1})^2.$$

Démonstration. Les estimateurs \hat{a}_n et \hat{b}_n sont les valeurs des paramètres a et b rendant maximale la quantité (2) (ou son logarithme)... \square

On peut montrer que ces estimateurs ont de bonnes propriétés asymptotiques.

Théorème 4.2. *Supposons que $a \in]0, 1[$. Les estimateurs \hat{a}_n et \hat{b}_n^2 sont fortement consistants, c'est-à-dire qu'ils convergent presque sûrement vers a et b^2 respectivement. De plus,*

$$\sqrt{n}(\hat{a}_n - a) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1 - a^2) \quad \text{et} \quad \sqrt{n}(\hat{b}_n^2 - b^2) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 2b^4)$$

La suite du paragraphe montre comment obtenir le comportement asymptotique de l'estimateur \hat{a}_n de a . On supposera dans la suite b connu et fixé. On réécrit \hat{a}_n de la manière suivante :

$$\hat{a}_n = a + b \frac{\sum_{k=1}^n V_{k-1} Y_k}{\sum_{k=1}^n V_{k-1}^2}.$$

On définit alors la suite aléatoire $(M_n)_{n \in \mathbb{N}}$ en posant

$$M_0 = 0 \quad \text{et, pour } n \geq 1, \quad M_n = \sum_{k=1}^n V_{k-1} Y_k.$$

La suite $(M_n)_{n \in \mathbb{N}}$ est une martingale par rapport à la filtration $(\mathcal{F}_n)_n$, où \mathcal{F}_n est la tribu engendrée par les variables aléatoires $(Y_k)_{1 \leq k \leq n}$. On lui associe sa variation quadratique $(\langle M \rangle_n)_n$ définie de la manière suivante : c'est la seule (p.s.) suite $(A_n)_n$ (*a priori* aléatoire) telle que $M_n^2 - A_n$ soit une martingale par rapport à (\mathcal{F}_n) .

Théorème 4.3. *La variation quadratique de M est donnée par*

$$\langle M \rangle_0 = 0 \quad \text{et, pour } n \geq 1, \quad \langle M \rangle_n = \sum_{k=1}^n V_{k-1}^2.$$

De plus, si $|a| < 1$ alors

$$\frac{\langle M \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \frac{1}{(1 - a^2)}, \quad \frac{M_n}{\langle M \rangle_n} \xrightarrow[n \rightarrow \infty]{p.s.} 0 \quad \text{et} \quad \frac{M_n}{\sqrt{\langle M \rangle_n}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

Remarque 4.4. Dans le cas $|a| < 1$, on a en particulier $\langle M \rangle_n \xrightarrow[n \rightarrow \infty]{p.s.} +\infty$.

Démonstration. Par une récurrence immédiate, on a

$$V_n = a^n V_0 + b \sum_{k=1}^n a^{n-k} Y_k.$$

Supposons que V_0 soit nul. L'inégalité de Cauchy-Schwarz assure que

$$V_n^2 \leq \frac{b^2}{1 - |a|} \sum_{k=1}^n |a|^{n-k} Y_k^2.$$

Posons $S_n = \sum_{k=1}^n V_{k-1}^2$ et $L_n = \sum_{k=1}^n Y_{k-1}^2$. La majoration ci-dessus montre que $S_n = O(L_n)$. D'autre part, en vertu de la loi des grands nombres, $L_n = O(n)$ et par suite $S_n = O(n)$ et $\langle M \rangle_n = O(n)$. Un résultat de martingale assure qu'alors $M_n = o(n)$.

On peut également écrire

$$S_n = a^2 S_{n-1} + 2abM_n + b^2 L_n \quad \text{et donc} \quad (1 - a^2) \frac{S_n}{n} = -V_{n-1}^2 + \frac{2abM_n}{n} + b^2 \frac{L_n}{n}.$$

Ainsi donc, la suite S_n/n converge-t-elle presque sûrement vers $b^2/(1 - a^2)$. On a donc montré que

$$\frac{\langle M \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \frac{b^2}{(1 - a^2)}.$$

La fin du théorème repose sur la loi des grands nombres et le théorème limite central pour les martingales de carré intégrable. \square

5 L'équation sur les vitesses

Dans ce paragraphe, on ne s'intéresse qu'à l'évolution temporelle de la distribution des vitesses. Dans les deux modèles, à temps discret comme à temps continu, on connaît la loi de la vitesse à un instant donné conditionnellement à l'événement $V_0 = v$. Nous nous plaçons à présent dans le modèle à temps continu. Soit f une fonction de classe \mathcal{C}^∞ à support compact. Introduisons la notation suivante :

$$P_t f(v) = \mathbb{E}(f(V_t) | V_0 = v).$$

Proposition 5.1. *Pour tout $t \geq 0$ et tout $v \in \mathbb{R}$, $P_t f(v)$ se représente de la façon suivante :*

$$\begin{aligned} P_t f(v) &= \int f(y) \exp\left(-\frac{(y - ve^{-\lambda t})^2}{2\lambda D(1 - e^{-2\lambda t})}\right) \frac{dy}{\sqrt{2\pi\lambda D(1 - e^{-2\lambda t})}} \\ &= \mathbb{E}[f(\sigma_t Y + ve^{-\lambda t})], \end{aligned}$$

où $\sigma_t^2 = \lambda D(1 - e^{-2\lambda t})$ et Y suit une loi gaussienne centrée réduite. La famille $(P_t)(\cdot)(v)_{t \geq 0, v \in \mathbb{R}}$ vérifie les propriétés suivantes :

1. Pour tout $v \in \mathbb{R}$, $\lim_{t \rightarrow 0} P_t f(v) = P_0 f(v) = f(v)$.
2. Pour tous $s, t \geq 0$ et $v \in \mathbb{R}$, $P_t \circ P_s f(v) = P_{t+s} f(v)$.
3. La fonction u définie par $u(t, v) = P_t f(v)$ est solution de l'équation aux dérivées partielles suivante :

$$\partial_t u(t, v) = \lambda^2 D \partial_v^2 u(t, v) - \lambda v \partial_v u(t, v),$$

de condition initiale $u(0, \cdot) = f(\cdot)$.

Démonstration. La représentation intégrale de $P_t f(v)$ découle de la proposition 3.2. Le point 2 s'obtient en écrivant :

$$P_t \circ P_s f(v) = \mathbb{E}[P_s f(\sigma_t Y + v e^{-\lambda t})] = \mathbb{E}[f(\sigma_s Z + (\sigma_t Y + v e^{-\lambda t}) e^{-\lambda s})],$$

où Y et Z sont des variables aléatoires indépendantes de même loi $\mathcal{N}(0, 1)$. On conclut en utilisant les propriétés de stabilité des lois gaussiennes. Le point 3 s'obtient en dérivant la représentation intégrale. \square

Remarque 5.2. Les points 1 et 2 de la proposition 5.1 traduisent que $(P_t)_t$ est un semi-groupe. De plus, l'opérateur $f \mapsto P_t f$ préserve la positivité et les constantes. On dit alors que c'est un semi-groupe de Markov. Enfin, la fonction $(t, v) \mapsto P_t f(v)$ est solution de l'équation

$$\partial_t P_t f(v) = L P_t f(v) \quad \text{où} \quad L g(v) = \lambda^2 D g''(v) - \lambda v g'(v).$$

L'opérateur L apparaît comme la dérivée logarithmique de P_t et l'on pourrait être tenté de représenter P_t comme

$$P_t f = e^{tL} f.$$

Cette écriture qui ne semble que formelle peut être rendue rigoureuse et traduit parfaitement l'intuition. On pourra en particulier penser à l'analogie en temps discret.

6 Suggestions

1. On pourra commenter la remarque 3.3.
2. On pourra démontrer la proposition 4.1.
3. On pourra démontrer le début de la proposition 4.3.
4. On pourra illustrer par la simulation les propriétés asymptotiques des estimateurs \hat{a}_n et \hat{b}_n .
5. On pourra démontrer tout ou partie de la proposition 5.1.