Approche probabiliste des liens entre distances et maillages

Application à l'exploitation d'enquêtes origine-destination

Olivier BONIN

UPE – IFSTTAR – LVMT

Mohamed Hamza LEMSSOUGUER

Ecole des Ponts ParisTech

Enquêtes origine-destination

Les enquêtes origine-destination décrivent les déplacements à l'aide d'un zonage (régulier ou irrégulier). Pour chaque déplacement, on connaît la maille de départ et la maille d'arrivée.

Deux problèmes :

- on ne connaît pas le point de départ du déplacement dans la zone de départ, ni le point d'arrivée dans la zone d'arrivée ;
- on ne connaît pas le trajet réel entre les deux zones.

Méthodologie du CERTU

En appelant D la vraie distance parcourue et D_{vo} la distance à vol d'oiseau, les formules calées sont pour des trajets en voiture et deux-roues motorisés :

- si $D_{vo} \leq 1$km alors $D = (2,2 - 0,72 D_{vo}) D_{vo}$
- si $D_{vo} \geq 1$km alors $D = 1,4 D_{vo}$

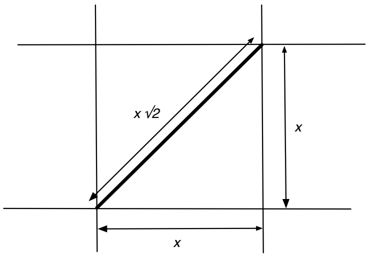
pour les trajets en transport collectifs :

- $D = 1,5 D_{vo}$

et pour les trajets en vélo :

- $D = 1,35 D_{vo}$

Un facteur de l'ordre de 1,4



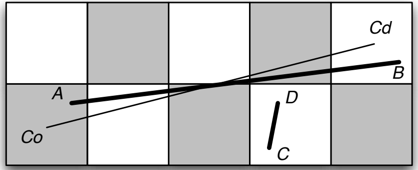
Cas du déplacement sur un réseau carré : l'écart le plus important entre le trajet à vol d'oiseau et le trajet sur le réseau est obtenu pour une diagonale du réseau. Cet écart vaut $2/\sqrt{2} = 1,414$.
La distance de Manhattan correspond intuitivement à des parcours de réseaux routiers en zone urbaine dense.

«Géomathématique des flux» de Ch. Terrier

Ch. Terrier a étudié le problème de la mesure des longueurs des déplacements dans les enquêtes origine-destination (JMS09) : quelle est l'influence du maillage sur les mesures de flux entre zones ? Deux constats :

- le nombre de déplacements diminue lorsqu'on augmente la taille des mailles ;
- pour un grand nombre de voyages simulés, la somme des distances entre mailles est à peu près égale à la somme des distances entre points, avec des écarts relatifs compris entre $0,8$ et $1,4$.

Maillage régulier



d(A, B) est approché par d(C_o, C_d) et le trajet DC est interne à une maille.

Formalisation du problème

À la suite de Ch. Terrier, nous considérons deux problèmes :

Déplacements internes à une maille

Quelle est la probabilité qu'un déplacement soit interne à une maille, lorsque le maillage est régulier ?

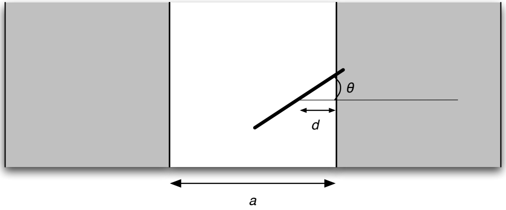
Déplacements quelconques

Quelle est la loi de probabilité de la distance entre deux points de position aléatoire, selon la distribution spatiale du point de départ et du point d'arrivée ?

Variations autour de l'aiguille de Buffon

Problème de l'aiguille de Buffon

Expérience menée en 1733 par Georges-Louis Leclerc de Buffon et qui a permis une approximation du nombre π. Elle consiste en un lancer répété d'une aiguille de longueur l sur un parquet constitué de lattes parallèles de largeur a.



La probabilité que l'aiguille soit à cheval sur deux lattes lors d'un lancer vaut $\frac{2l}{\pi a}$.

Résolution par calcul

La probabilité recherchée est l'aire de la zone de chevauchement (zone favorable) divisée par l'aire de la zone totale.
L'aiguille est à cheval sur deux lattes si et seulement si $\frac{l \cos \theta}{2} \geq d$, d'où :

$$\begin{aligned}
\frac{\int_{-\pi/2}^{\pi/2} \int_0^{l\cos\theta/2} \mathrm{d}\theta \mathrm{d}d}{\int_{-\pi/2}^{\pi/2} \int_0^{a/2} \mathrm{d}\theta \mathrm{d}d} &= \frac{\int_{-\pi/2}^{\pi/2} (l \cos\theta/2) \mathrm{d}\theta}{\pi a/2} \\
&= \frac{l(\sin(\pi/2) - sin(-\pi/2))/2}{\pi a/2} \\
&= \frac{2l}{\pi a}
\end{aligned}$$

Problème 95 de Loïc Terrier

Revenons à notre problème.

Problème 95 de Loïc Terrier, posé par Ch. Terrier aux JMS en 2009

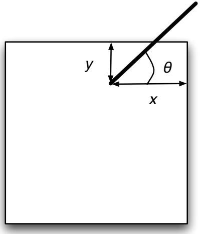
On lance un spaghetti de longueur d sur un sol carrelé, les carreaux étant des carrés de côté unité. Quelle est la probabilité que le spaghetti soit à l'intérieur d'un des carreaux ?

Trois cas :

- $d \leq 1$ (aiguille plus petite que le côté d'un carreau)
- $1 \leq d \leq \sqrt{2}$ (aiguille plus petite que la diagonale de longueur $\sqrt{2}$ du carreau)
- $d \geq \sqrt{2}$.

Paramétrage

Nous repérons la position de l'aiguille (du spaghetti) par (x, y, θ) avec x et y qui varient entre 0 et 1, et $\theta \in [0, \pi/2]$.

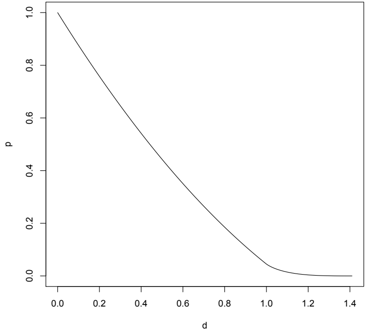


Résultat

Une série de calculs donne :

$$
\begin{aligned}
p &= 1 - \frac{4d - d^2}{\pi} \text{ pour } d \leq 1 \\
&= 1 - \frac{d^2 - 4\sqrt{d^2 - 1} - 4\arcsin(1/d) + 2\pi + 2}{\pi} \text{ pour } 1 \leq d \leq \sqrt{2} \\
&= 0 \text{ pour } d \geq \sqrt{2}
\end{aligned}
$$

Evolution de p en fonction de d



Cas d'une grande aiguille sur un sol carrelé

Nous revenons au sujet des longueurs des trajets calculés à partir des enquêtes origine-destination en prenant le cas d'une grande aiguille, figurant le trajet, sur un carroyage décrivant les zones d'enquêtes tel que celui de l'EGT.

Pour simplifier les calculs, nous revenons au cas simple du parquet, mais les résultats s'étendent de la même façon que plus haut au cas d'un dallage à partir des probabilités que nous avons calculées.

Espérance du nombre de lames traversées

Soit X la variable aléatoire qui décrit le nombre de lames intersectées par l'aiguille lors d'un lancer. On remarque que $E(X)$ est une fonction de la longueur de l'aiguille d, et nous notons cette fonction $f(d) = E(X)$.

Si $d \leq a$, X est une variable de Bernoulli de paramètre $p = \frac{2d}{\pi a}$ donc

$$E(X) = \frac{2d}{\pi a} = f(d).$$

Espérance du nombre de lames traversées

Si $d \geq a$, on remarque que $f(d)$ est linéaire en d et on écrit $d = na + d_0$ avec $d_0 \leq a$ et $n \in \mathbb{N}$i :

$$f(d) = f(na + d_0) = nf(a) + f(d_0) = n\frac{2a}{\pi a} + \frac{2d_0}{\pi a} = \frac{2d}{\pi a}$$

d'où le même résultat :

$$E(X) = \frac{2d}{\pi a}.$$

Loi des distances entre deux points

Notations

À partir d'un maillage quelconque, nous supposons qu'un trajet a été effectué entre deux points A et B, où A appartient à la zone de départ o de centroïde $C_o = A_0$ et B à la zone d'arrivée de centroïde $C_d = B_0$.
La vraie distance du parcours est $\mathrm{d}(A, B)$, et la distance imputée à vol d'oiseau $\mathrm{d}(C_o, C_d) = \mathrm{d}(A_0, B_0)$.

En plaçant un repère orthonormé dont l'origine est en A_0 et un des axes de même direction que $\overrightarrow{A_0B_0}$ on a dans ce repère $A_0 = (0, 0)$, $B_0 = (d_0, 0)$, $A = (x_A, y_A)$ et $B = (x_B, y_B)$.
On peut supposer que x_A, x_B, y_A et y_B sont des réalisations de variables aléatoire indépendantes X_A, X_B, Y_A et Y_B.

Expression générale de $d(A, B)$

Un calcul de densité donne, après un passage en polaires, la densité suivante pour la loi de $Z = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2}$ qui est la distance (aléatoire) entre A et B :

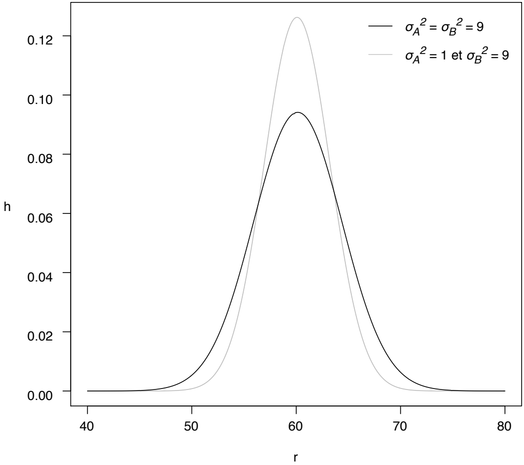
$$h(r) = r \int_0^{2\pi} \left(\int f_A(v) f_B(v - r\cos\theta) \mathrm{d}v \int g_A(v) g_B(v - r\sin\theta) \mathrm{d}v \right) \mathrm{d}\theta$$

Cas de lois normales

Trajet entre deux mailles dont les centroïdes sont distants de 60km.

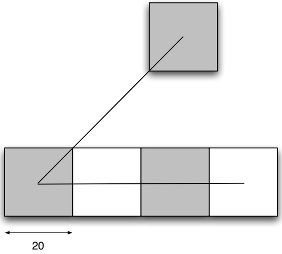
$$\begin{aligned} X_A &\sim N(0, \sigma_A^2) \\ Y_A &\sim N(0, \sigma_A^2) \\ X_B &\sim N(60, \sigma_B^2) \\ Y_B &\sim N(0, \sigma_B^2) \end{aligned}$$

avec $\sigma_A^2 = \sigma_B^2 = 9$, puis $\sigma_A^2 = 1$ et $\sigma_B^2 = 9$.


$\sigma_A{}^2 = \sigma_B{}^2 = 9$
$\sigma_A{}^2 = 1$ et $\sigma_B{}^2 = 9$
h
0.12
0.10
0.08
0.06
0.04
0.02
0.00
40
50
60
70
80
r

Cas de lois uniformes

Maillage carré avec deux mailles sur une même ligne horizontale séparées d'une distance centre à centre de 60km, et deux mailles séparées par une même distance en diagonale.

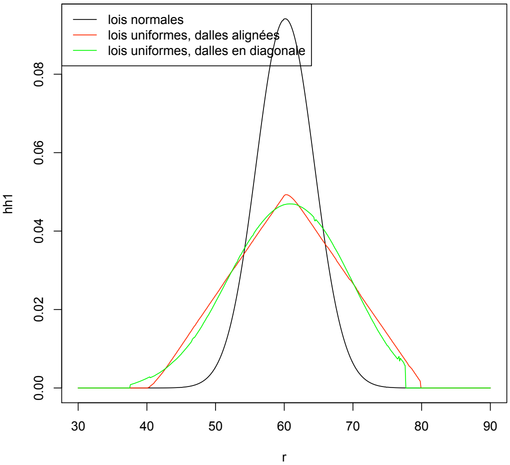


$$
\begin{aligned}
X_A &\sim U[-10, 10] \\
Y_A &\sim U[-10, 10] \\
X_B &\sim U[50, 70] \\
Y_B &\sim U[-10, 10]
\end{aligned}
$$

pour les mailles alignées, et

$$
\begin{aligned}
X_A &\sim U[-10, 10] \\
Y_A &\sim U[-10, 10] \\
X_B &\sim U[\sqrt{60}/2 - 10, \sqrt{60}/2 + 10] \\
Y_B &\sim U[\sqrt{60}/2 - 10, \sqrt{60}/2 + 10]
\end{aligned}
$$

pour les mailles en diagonale.


lois normales
lois uniformes, dalles alignées
lois uniformes, dalles en diagonale
hn1
0.00
0.02
0.04
0.06
0.08
30
40
50
60
70
80
90
r

Lois normales et uniformes