Celia CONSTANTINI

Le problème du voyageur de commerce

Mémoire d'initiation à la recherche

Sous la direction de Monsieur **Julien Poisat**

PSL★ RESEARCH UNIVERSITY

DAUPHINE UNIVERSITÉ PARIS

Cycle Pluridisciplinaire d'Études Supérieures
Juin 2017

Table des matières

Introduction

Le problème du voyageur de commerce consiste en la traversée de n villes par un voyageur qui doit minimiser la distance parcourue : n villes sont disposées dans l'espace. Un voyageur doit parcourir ces n villes puis revenir à son point de départ, tout en minimisant la longueur du trajet. Chaque chemin possible correspond à une permutation des villes. Comme le trajet effectué est un cycle, on peut se fixer une ville de départ. Il y a donc $(n-1)!$ chemins possibles. On pourrait tester toutes les mais cela prendrait un temps trop important étant donné le nombre de chemins possibles. On utilise donc un algorithme appelé recuit simulé donnant une approximation de ce plus court chemin. Un exemple de code pour cet algorithme ainsi que ses résultats sont donnés en annexe. On peut aussi donner une estimation de la longueur de ce chemin lorsque les villes sont aléatoirement réparties. Dans une première partie, nous étudierons les chaînes de Markov et leur propriétés permettant de construire les algorithmes de Metropolis-Hastings et du recuit simulé qui font l'objet de la seconde partie. Enfin la troisième partie est consacrée à l'estimation de la longueur du trajet optimal.

1 Chaînes de Markov

On considère dans toute cette partie un espace d'états E fini.

1.1 Définition

Définition 1.1. *Une chaine de Markov est un processus aléatoire $(X_n)_{n\in\mathbb{N}}$ à valeur dans E vérifiant la proprieté de Markov : pour tout $n \in \mathbb{N}$, pour tous $x_0, \dots, x_n, x \in E$,*

$$\mathbb{P}(X_{n+1} = x | X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_{n+1} = x | X_n = x_n).$$

La chaine de Markov est dite homogène si $\mathbb{P}(X_{n+1} = x | X_n = y)$ ne dépend pas de n mais seulement de x et $y \in E$.

On peut alors considérer la matrice P telle que pour $x, y \in E$, $P(x,y) = \mathbb{P}(X_1 = y | X_0 = x)$, appelée matrice de transition associée à (X_n). Une telle matrice vérifie :
- $\forall (x,y) \in E^2, P(x,y) \geq 0$
- $\forall x,\ \sum_y P(x,y) = 1$

On a alors :

Proposition 1.1. *Soit $n \in \mathbb{N}$. Étant donnée la loi μ_0 de X_0, la loi de X_n est donnée par :*

$$\mu_n = \mu_0 P^n$$

Démonstration. Par récurrence. On a bien $\mu_0 = \mu_0 P^0$.
Soit $n \geq 1$ tel que $\mu_{n-1} = \mu_0 P^{n-1}$.

$$\begin{aligned}
\mathbb{P}(X_n = x) &= \sum_{x_0,\dots,x_{n-1}} \mathbb{P}(X_0 = x_0, \dots, X_{n-1} = x_{n-1}, X_n = x) \\
&= \sum_{x_0,\dots,x_{n-1}} \mathbb{P}(X_0 = x_0, \dots, X_{n-1} = x_{n-1}) \mathbb{P}(X_n = x | X_{n-1} = x_{n-1}) \\
&= \sum_{x_{n-1}} \sum_{x_0,\dots,x_{n-2}} \mathbb{P}(X_0 = x_0, \dots, X_{n-1} = x_{n-1}) P(x_{n-1}, x) \\
&= \sum_{x_{n-1}} (\mu P^{n-1})(x_{n-1}) P(x_{n-1}, x) \\
&= (\mu P^n)(x)
\end{aligned}$$

où la première ligne découle de la formule des probabilités totales, la seconde de la propriété de Markov.

□

1.2 Irréductibilité et apériodicité

Définition 1.2. *Soient x, y dans E. On dit que x et y communiquent s'il existe m et n entiers tels que $P^m(x, y) > 0$ et $P^n(y, x) > 0$.*
Si pour tous x, y dans E, x et y communiquent, la chaîne de Markov associée à P est dite irréductible.

Exemple 1.1. La matrice de transition, $\begin{pmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix}$ est irréductible.

En revanche, $\begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ ne l'est pas car pour tout $n \in \mathbb{N}, P^n(3, 2) = 0$.

Définition 1.3. *La période de $x \in E$ est pgcd$\{n, P^n(x, x) > 0\}$. Un état $x \in E$ est dit apériodique si sa période est 1. Si tous les états sont apériodiques la chaîne est dite apériodique.*

Exemple 1.2. La matrice

Proposition 1.2. *Si (X_n) est irréductible et qu'il existe $x \in E$ apériodique alors pour tout y dans E, y est apériodique.*

Démonstration. Soit $y \in E$. On appelle p sa période. Il existe $k, l \in \mathbb{N}$ tels que $P^k(x, y) > 0$ et $P^l(y, x) > 0$. Soit $n \geq 1$ tel que $P^n(x, x) > 0$ alors,

$$P^{k+l+n}(y, y) \geq P^l(y, x) P^n(x, x) P^k(x, y) > 0$$

et

$$P^{k+l}(y, y) \geq P^l(y, x) P^k(x, y) > 0.$$

Ainsi p divise $k + l + n$ et $k + l$ donc il divise n. C'est vrai pour tout $n \in \{k, P^k(x, x) > 0\}$ donc p divise le *pgcd*$\{k, P^k(x, x) > 0\} = 1$, soit $p = 1$. □

1.3 Mesure invariante, réversibilité et convergence

Définition 1.4. *Soit π une mesure sur E. On dit que π est invariante ou stationnaire, si $\pi = \pi P$*

Définition 1.5. *Une mesure π sur E est dite réversible si pour tous x, y dans E,*

$$\pi(x)P(x,y) = \pi(y)P(y,x)$$

Proposition 1.3. *Soit π une mesure réversible. Alors π est invariante.*

Démonstration. On a pour tout $x \in E$

$$\pi(x) = \pi(x)\sum_y P(x,y) = \sum_y \pi(y)P(y,x)$$

□

Théorème 1.1. *Soit (X_n) une chaîne de Markov sur E irréductible. Alors il existe une unique mesure de probabilité invariante pour (X_n). .*

Démonstration. (Tirée du polycopié de processus discrets de J. Lehec])
Existence
Soit P la matrice de transition associée à (X_n). On pose $k = \text{card } E$. On peut identifier l'ensemble des mesures de probabilité sur E à l'ensemble $A = \{x \in \mathbb{R}^k, x_i \geq 0 \ \forall i, \sum x_i = 1\}$ qui est un sous ensemble convexe compact de $\mathbb{R}^k$. Soit $\mu \in A$. On pose $\mu_n = \frac{1}{n}\sum_{i=0}^{n-1} \mu P^i \in A$. A est compact donc on peut en extraire une sous-suite qui converge vers $\pi \in A$. D'autre part on a,

$$\mu_n P = \frac{1}{n}\sum_{i=0}^{n-1} \mu P^{i+1} = \mu_n + \frac{1}{n}(\mu P^{n+1} - \mu)$$

$\mu P^{n+1} - \mu$ est bornée donc $\mu_n - \mu_n P \to 0$. Donc $\pi = \pi P$ et π est donc invariante.
Unicité
Soient π et μ deux mesures invariantes. π est une mesure de probabilité alors il existe $x \in E$ tel que $\pi(x) > 0$. Donc il existe $y \in E$ tel que $\frac{\mu(y)}{\pi(y)}$ soit minimal. On a donc pour tout $x \in E,\ \mu(x) \geq \frac{\mu(y)}{\pi(y)}\pi(x)$ On a donc pour tout k,

$$\begin{array}{rcl} \mu(y) & = & \mu P^k(y) \\ & = & \sum_{x\in E} \mu(x)P^k(x,y) \\ & \geq & \sum_{x\in E} \frac{\mu(y)}{\pi(y)}\pi(x)P^k(x,y) \\ & = & \mu P^k(y) \end{array}$$

donc l'inégalité est une égalité et pour tout x et pour tout $k \in \mathbb{N}$,

$$\mu(x)P^k(x,y) = \frac{\mu(y)}{\pi(y)}\pi(x)P^k(x,y)$$

Or la chaîne est irréductible alors pour tout $x \in E$, il existe $k \geq 1$ tel que $P^k(x, y) > 0$ et donc $\mu(x) = \frac{\mu(y)}{\pi(y)}\pi(x)$.

Alors μ et π sont deux mesures de probabilité proportionnelles donc $\mu = \pi$

$\square$

Théorème 1.2. *Soit (X_n) une chaîne de Markov sur E (fini) irréductible et apériodique. Alors (X_n) converge en loi vers sa loi invariante.*

2 Algorithmes de Metropolis-Hastings et du recuit-simulé

Soit E l'ensemble des permutations des n villes (chemins possibles). Soit $x \in E$, on note $L(x)$ la longueur du chemin x.

Pour tout $\beta \in \mathbb{R}$, on appelle mesure de Gibbs μ_β la mesure de probabilité sur E donnée par :

$$\mu_\beta(x) = \frac{1}{Z_\beta} e^{-\beta L(x)}, \quad \text{avec } Z_\beta = \sum_{x \in E} e^{-\beta L(x)}$$

On remarque que μ_β est d'autant plus grande que $L(x)$ est petit. Ainsi, simuler cette distribution permettrait d'obtenir avec une probabilité élevée un chemin de petite longueur.

Cependant on ne peut pas avoir exactement la mesure de Gibbs car l'obtention de la constante de normalisation nécessite de calculer $e^{-\beta L(x)}$ pour tout x dans E, ce qu'on veut éviter.

On peut néanmoins approcher μ_β en construisant une chaîne de Markov récurrente, apériodique de loi invariante μ_β.

Théorème 2.1 (Metropolis-Hastings)**.** *Étant donnée une matrice de transition Q irréductible, et une fonction $\alpha : \mathbb{R} \mapsto [0,1]$ telle que $\forall u \in \mathbb{R}_+, \alpha(u) = u\alpha(\frac{1}{u})$, on définit P par :*

$$P(x,y) = \begin{cases} Q(x,y)\rho(x,y) & si\ x \neq y \\ 1 - \sum_{z \neq x} P(x,z) & si\ x = y \end{cases}$$

avec

$$\rho(x,y) = \alpha\left(\frac{\mu_\beta(y)Q(y,x)}{\mu_\beta(x)Q(x,y)}\right) \mathbb{1}_{Q(x,y)>0}$$

Alors, P est une matrice de transition irréductible, apériodique si Q l'est ou si $\alpha < 1$, de loi invariante μ_β.

Démonstration. On a, pour tous x, y dans E, $P(x,y) \geq 0$ et $\sum_{y \in E} P(x,y) = 1$ par construction donc P est une matrice de transition. Q est irréductible donc pour tous $x, y \in E$ il existe $n \in \mathbb{N}*$ tel que $Q^n(x,y) > 0$ et comme $\alpha > 0$, $P^n(x,y) > 0$ et est irreductible et de même si Q est apériodique, P l'est car $\{n \geq 1, (Q^n)(x,x) > 0\} \subset \{n \geq 1, (P^n)(x,x) > 0\}$.

Si $\alpha < 1$, $P(x,y) < Q(x,y)$ pour $x \neq y$ et donc pour tout x dans E, $P(x,x) > Q(x,x) \geq 0$ et P apériodique. Il suffit en fait de l'existence de $(x,y) \in E^2$ tel que $\rho(x,y) < 1$ pour que P soit apériodique. On a alors $P(x,y) < Q(x,y)$ d'où $P(x,x) > 0$ et par la propriété 1.2 Enfin μ_β est la mesure invariante de P car elle est réversible. En effet, pour tous $x, y, Q(x,y) \neq 0$

$$\begin{aligned}
\mu_\beta(x)P(x,y) &= \mu_\beta(x)Q(x,y)\alpha\left(\frac{\mu_\beta(y)Q(y,x)}{\mu_\beta(x)Q(x,y)}\right) \\
&= -\mu_\beta(x)Q(x,y)\frac{\mu_\beta(y)Q(y,x)}{\mu_\beta(x)Q(x,y)}\alpha\left(\frac{\mu_\beta(x)Q(x,y)}{\mu_\beta(y)Q(y,x)}\right) \\
&= \mu_\beta(y)Q(y,x)\alpha\left(\frac{\mu_\beta(x)Q(x,y)}{\mu_\beta(y)Q(y,x)}\right) \\
&= \mu_\beta(y)P(y,x)
\end{aligned}$$

Si $Q(x,y) = 0$, on a $\mu_\beta(x)P(x,y) = 0 = \mu_\beta(y)P(y,x)$. □

Cela correspond à accepter la transition x à y effectuée selon Q avec une probabilité $\rho(x,y)$.

Prenons par exemple $\alpha : u \mapsto \min(u,1)$ et pour Q la marche aléatoire sur les transposition :

$$Q(x,y) = \begin{cases} \frac{2}{n(n-1)} & \text{si il existe une transposition } \tau \text{ telle que } y = x\tau \\ 0 & \text{sinon} \end{cases}$$

On a alors pour $x, y \in E$, $Q(x,y) = Q(y,x)$ et $\rho(x,y) = \min(e^{-\beta(L(y)-L(x))}, 1)$. Ainsi, on accepte systématiquement y si $L(y) < L(x)$ et on le rejette avec une probabilité non nulle $E^{-\beta(L(y)-L(x))}$ sinon. Cette probabilité est d'autant plus élevée que $\beta(L(y) - L(x))$ est grand.

Ainsi si on a $x \in E$ qui n'est pas optimal mais tel que pour tout y tel que $Q(x,y) > 0$, $L(x) < L(y)$, on peut accepter l'état y et pas rester au niveau d'un minimum local.

On constate que plus β est grand moins les cycles de grande longueur seront acceptés.

En effet on a pour tout $x \in E$, $\lim_{\beta\to\infty} \mu_\beta(x) = \frac{\mathbb{1}_M}{|M|}$ où $M = x \in E, L(x) = \inf L$.

$$\begin{aligned}
\mu_\beta(x) &= \frac{e^{-\beta L(x)}}{\sum_{x\in E} e^{-\beta L(x)}} \\
&= \frac{e^{-\beta(L(x)-L_{min})}}{\sum_{x\notin M} e^{-\beta(L(x)-L_{min})} + \sum_{x\in M} 1}, \text{ avec } L_{min} = \min L
\end{aligned}$$

C'est ce qu'exploite l'algorithme du recuit simulé. Il s'agit du même algorithme que Metropolis-Hastings, en faisant varier le paramètre β au cours du temps.

3 Estimation de la longueur du trajet minimal

Dans la section précédente, nous avons vu une méthode permettant d'approcher la solution optimale étant donnés les n points à parcourir. On s'intéresse dorénavant à un aspect plus théorique, estimer la longueur du chemin optimal lorsque les points sont répartis selon une loi donnée.

On considère n points de l'espace, $X_1, \ldots X_n$, variables aléatoires indépendantes et de même distribution. Dans toute la suite, on prendra une distribution uniforme sur $[0,1]^2$.
On note $L_n(X_1, \ldots, X_n)$ la longueur du plus court chemin passant par les points $X_1, \ldots, X_n$.

Théorème 3.1. *Il existe des constantes positives A et B telles que p.s. à partir d'un certain rang $A\sqrt{n} < L_n(X_1, \ldots, X_n) < B\sqrt{n}$*

La preuve de ce théorème repose sur deux résultats. On montre d'abord que l'espérance L_n est de l'ordre $\sqrt{n}$, puis que L_n est proche de son espérance.

Proposition 3.1. *Il existe des réels c^+ et c^-, $0 < c^- < c^+$ tels que pour tout $n \geq 1$,*

$$c^- \sqrt{n} \leq \mathbb{E}(L_n(X_1, \ldots, X_n)) \leq c^+ \sqrt{n}$$

Proposition 3.2. *Il existe $c > 0$ tel que pour tout $n \geq 1$ et tout réel $t \geq 0$, on a :*

$$\mathbb{P}(|L_n - \mathbb{E}(L_n)| \geq t) \leq 2 \exp\left(-\frac{ct^2}{\ln(n)}\right)$$

On peut alors démontrer le théoreme 3.1.

Démonstration. Par la proposition 3.2, on a pour tout n≥ 1 :

$$\mathbb{P}\left(|L_n - \mathbb{E}(L_n)| \geq \frac{2\ln(n)}{\sqrt{c}}\right) \leq 2\exp\left(-4\ln(n)\right) = n^{-4}$$

d'où,

$$\sum_{n \in \mathbb{N}^*} \mathbb{P}\left(|L_n - \mathbb{E}(L_n)| \geq \frac{2\ln(n)}{\sqrt{c}}\right) < \infty$$

Alors par le lemme de Borel-Cantelli, presque sûrement à partir d'un certain rang, $|L_n - \mathbb{E}(L_n)| < C\ln(n)$ pour une certaine constante C.
Or pour n assez grand, il existe $a > 0$ tel que $0 < \ln n < a\sqrt{n}$.
Donc par la proposition 3.1, on a bien l'existence de A, B tels que $A\sqrt{n} < L_n(X_1, \ldots, X_n) < B\sqrt{n}$ □

Montrons maintenant la proposition 3.1.

Démonstration. $[0,1]^2$ est l'union de n petits cubes de côté $\frac{1}{\sqrt{n}}$.

Soit $(x_1,\dots,x_n) \in ([0,1]^2)^n$. On distingue deux cas : soit les x_i sont chacun dans un carré différent et on a $\min\limits_{x_i \neq x_j} |x_i - x_j| < \frac{c}{\sqrt{n}}$, où c=3 convient, soit il existe deux points x_i et x_j dans le même carré et $|x_i - x_j| < \frac{\sqrt{2}}{\sqrt{n}}$. Donc, il existe $c > 0$ pour tous $x_1,\dots,x_n$ dans $[0,1]^2$, $\min\limits_{x_i \neq x_j} |x_i - x_j| < \frac{c}{\sqrt{n}}$. Ainsi, si on pose $l_n = \max\limits_{x_1,\dots,x_n} L_n(x_1,\dots,x_n)$ on a, $l_n \leq l_{n-1} + 2\frac{c}{\sqrt{n}}$ et donc

$$\|L_n(X_1,\dots,X_n)\|_\infty \leq 2c\sum_{k=2}^{n} \frac{1}{\sqrt{k}} < c^+\sqrt{n}$$

par comparaison somme integrale, pour une certaine constante $c^+ > 0$. On a donc $\mathbb{E}(L_n(X_1,\dots,X_n)) \leq c^+\sqrt{n}$.

Montrons la seconde inégalité.
Soient $x \in [0,1]^2$ et $r \leq 1/2$. On a $\mathbb{P}(X_1 \in B(x,r)) \leq \pi r^2$. D'où,

$$\begin{aligned}
\mathbb{P}(\min_{1\leq i\leq n-1} |X_i - x| \geq r) &= \mathbb{P}(|X_1 - x| \geq r)^{n-1} \\
&= (1 - \mathbb{P}(X_1 \in B(x,r)))^{n-1} \\
&\geq (1 - \pi r^2)^{n-1}
\end{aligned}$$

Alors,

$$\begin{aligned}
\mathbb{E}(\min_{1\leq i\leq n-1} |X_i - x|) &= \int_0^\infty \mathbb{P}(\min_{1\leq i\leq n-1} |X_i - x| \geq r)dr \\
&\geq \int_0^{1/2} (1-\pi r^2)^{n-1} dr \\
&\geq \int_0^{1/n-1} (1-\pi r^2)^{n-1} dr \\
&\geq \int_0^{\sqrt{(1/\pi(n-1))}} 1 - (n-1)\pi r^2 dr \ (*) \\
&= \frac{1}{\sqrt{\pi(n-1)}} - \frac{\pi(n-1)}{3(\pi(n-1))^{3/2}} \\
&\geq \frac{c}{\sqrt{n}} \text{ pour une constante } c > 0
\end{aligned}$$

avec (*) qui découle de l'inégalité $(1-u)^\alpha \geq 1 - \alpha u$ pour $u \geq 1/\alpha$.

Enfin le chemin optimal passe par chacun des $X_1,\dots,X_n$ et on peut donc écrire $L_n(X_1,\dots,X_n) = \sum_{i=1}^n |X_i - X_{j_k}|$ avec $j_k \in \{1,\dots,n\}$

$$\begin{aligned}
\mathbb{E}(L_n(X_1,\dots,X_n)) &\geq \sum_{i=1}^n \mathbb{E}\left(\min_{j\neq i} |X_i - X_j|\right) \\
&= \sum_{i=1}^n \mathbb{E}\left(\mathbb{E}\left(\min_{j\neq i} |X_i - X_j||X_j\right)\right) \\
&\geq n\frac{c}{\sqrt{n}}
\end{aligned}$$

□

On peut en fait montrer qu'il existe une constante γ telle que $L_n \sim_{n\to\infty} \gamma\sqrt{n}$ p.s.

Ce résultat peut s'étendre à d'autres distributions des X_i dans $\mathbb{R}^d$.

Théorème 3.2 (Bearwood-Halton-Hammersley)**.** *Il existe une constante* $0 < \gamma_d < \infty$ *telle que si la distribution* μ *des* $X_i,\ i = 1, \ldots, n$ *à support compact et de densité* $f : \mathbb{R}^d \to \mathbb{R}_+$ *par rapport à la mesure de Lebesgue alors,*

$$\frac{L_n(X_1, \ldots, X_n)}{n^{(d-1)/d}} \xrightarrow[n \to +\infty]{p.s.} \gamma_d \int_{\mathbb{R}^d} f(x)^{(d-1)/d} dx$$

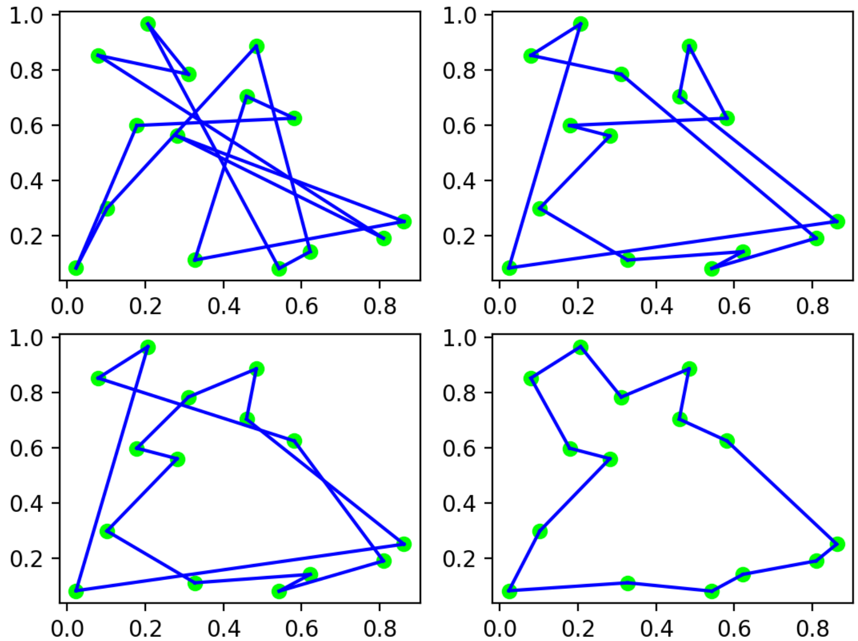
Annexe

```

from numpy import *
from matplotlib import pyplot
from random import *


def positions(n):
    # renvoie un tableau x avec les abscisses de n points tirés
    #uniformément et y contenant les ordonnées
    x=[]
    y=[]
    for i in range (n):
        a=random()
        b=random()
        x.append(a)
        y.append(b)
    return x,y

def distancesVilles(n,x,y):
    A=[[sqrt((x[i]-x[j])**2 + (y[i]-y[j])**2) for i in range(n)] for j in range(n)]
    return A

def longueurCycle(cycle,A,n):
    l=A[0][n-1]
    for i in range (n-1):
        l+= A[cycle[i+1]][cycle[i]]
    return l

def affichageCycle(cycle, x, y,n):
    pyplot.plot(x,y, linestyle = ’none’, marker = ’o’, c = ’lime’)
    for i in range (n-1):
        pyplot.plot([x[cycle[i]],x[cycle[i+1]]],[y[cycle[i]],y[cycle[i+1]]], ’b’ )
    pyplot.plot([x[cycle[n-1]],x[cycle[0]]],[y[cycle[n-1]],y[cycle[0]]], ’b’ )

def xy(n):
   #renvoie les positions de n villes placées aux sommets d’un polygone regulier
    return [cos(i*2*pi/n) for i in range (n) ], [sin(i*2*pi/n) for i in range (n) ]

def listealea(n):
    #genere aleatoirement un cycle de n elements
    i=0
    P=[]
    while i<n:
```

```
        a=randint(0,n-1)
        if (a not in P):
            P.append(a)
            i+=1
    return P

def affichageDansFonction(i,P,x,y,n):
    if i==0:
        affichageCycle(P,x,y,n)
        pyplot.subplot(2,2,2)
    if i==n//3:
        affichageCycle(P,x,y,n)
        pyplot.subplot(2,2,3)
    if i==2*n//3:
        affichageCycle(P,x,y,n)
        pyplot.subplot(2,2,4)

def genereQ(cycle, n):
    #simule une chaine de Markov de noyau Q, ici la marche aléatoire
    #sur les transposition
    Q=[]
    for i in cycle:
        Q.append(i)
    i=randint(0,n-1)
    j=randint(0,n-1)
    if n!=0:
        Q[i]=Q[j]
        Q[j]=cycle[i]
    return Q

def genereP (cycle, Q, A, beta,n):
    Hy= longueurCycle(Q, A,n)
    Hx= longueurCycle(cycle, A,n)
    if Hy<Hx:
        ro=1
    else :
        ro=exp(-beta*(Hy-Hx))
    p=random()
    if p<ro :
        return Q,Hy
    return cycle, Hx

def HastingsMetropolis(P,n, x, y,A, it, beta):
    lmin =longueurCycle(P,A,n)
    Pmin=P
    pyplot.subplot(2,2,1)
```

```
    for i in range(it):
        Q=genereQ(P,n)
        P,l=genereP (P, Q, A, beta,n)
        if lmin>l:
            lmin=l
            Pmin=P

        affichageDansFonction(i,P,x,y,n)
    affichageCycle(Pmin,x,y,n)
    return Pmin


def RecuitSimule(P,n,x,y,A,it):
#beta qui varie en log(i)/c
    lmin =longueurCycle(P, A,n)
    Pmin=P
    pyplot.subplot(2,2,1)
    for i in range(it):
        beta=log((i+1))
        Q=genereQ(P,n)
        P,l=genereP (P, Q, A, beta,n)
        if lmin>l:
            lmin=l
            Pmin=P

        affichageDansFonction(i,P,x,y,n)
    affichageCycle(Pmin,x,y,n)
    return Pmin


def bouclePrinc(n, it, beta):
#Si beta vaut 0 recuit simule
#chacun des deux algos affiche les resultats au début, apres1/3, 2/3
#et a la fin des itérations
    P=listealea(n)
    x,y=positions(n)
    A=distancesVilles(n,x,y)
    if beta!=0 :
        return HastingsMetropolis(P,n, x, y,A, it, beta),x,y,A
    else :
        return RecuitSimule(P,n,x,y,A,it),x,y,A

R,x,y,A=bouclePrinc(15, 500000, 0)
```

Une simulation pour 15 villes et 500000 itérations donne le résultat suivant :



Bibliographie

- Chafaï, D., Malrieu, F. *Recueil de modèles aléatoires*, Springer (2016)
- Häggström O. *Finite Markov Chains and Algorithmic Applications*, Cambridge University Press, 2002
- Hastings, W.K. Monte Carlo sampling methods using Markov chains and their applications, Biometrika,57 (1970).
- Lehec J. Processus discrets