

Méthode probabiliste (dite de Monte Carlo) pour résoudre un système d'équations

Considérons un système linéaire de N équations à N inconnues x_1, x_2, \dots, x_N , soit

$$M X = B \text{ avec } X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_N \end{pmatrix} \text{ matrice colonne des inconnues, } B = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_N \end{pmatrix} \text{ matrice colonne de}$$

constantes b_i , M étant la matrice $N \times N$ du système avec des coefficients donnés réels. En prenant $M = Id - A$, le système peut toujours se réécrire :

$$X = AX + B, \text{ } A \text{ étant une matrice de dimensions } N \times N \text{ de coefficients } a_{ij}.$$

Par exemple, pour $N = 2$, le système

$$\begin{cases} 0,9x_1 - 0,2x_2 = 0,7 \\ -0,2x_1 + 1,3x_2 = 1,1 \end{cases}$$

équivalent au système :

$$\begin{cases} x_1 = 0,1x_1 + 0,2x_2 + 0,7 \\ x_2 = 0,2x_1 - 0,3x_2 + 1,1 \end{cases} \text{ où } A = \begin{pmatrix} 0,1 & 0,2 \\ 0,2 & -0,3 \end{pmatrix} \text{ et } B = \begin{pmatrix} 0,7 \\ 1,1 \end{pmatrix}$$

C'est sous la forme $X = AX + B$ que nous écrirons dorénavant le système, avec une contrainte : pour que la méthode de Monte Carlo fonctionne, on doit prendre la norme de A inférieure à 1 ($\|A\| < 1$), ce qui signifie que la somme des coefficients en valeur absolue sur chaque ligne de A doit être inférieure à 1. A partir de là, on construit une matrice P de dimensions $(N+1) \times (N+1)$ avec comme coefficients :

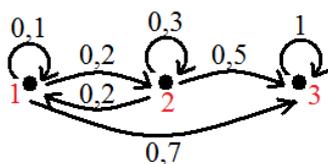
- $p_{ij} = v_{ij} a_{ij}$ (pour i et j de 1 à N) où $v_{ij} = \pm 1$ selon que a_{ij} est positif ou négatif, de façon que tous les p_{ij} soient positifs (ils sont aussi inférieurs à 1)
- $p_{i \ N+1} = 1 - \sum_{j=1}^n p_{ij}$ pour i de 1 à N , et sur la dernière ligne $N+1$, des 0 sauf un 1 en dernier.

Dans l'exemple de la matrice A précédente, cela donne :

$$P = \begin{pmatrix} 0,1 & 0,2 & 0,7 \\ 0,2 & 0,3 & 0,5 \\ 0 & 0 & 1 \end{pmatrix} \text{ avec } v_{11} = v_{12} = v_{21} = 1 \text{ et } v_{22} = -1$$

Ainsi construite, la matrice P va pouvoir être vue comme une matrice de probabilités. Pour cela considérons une particule circulant sur un graphe orienté. Celui-ci possède $N+1$ nœuds numérotés de 1 à $N+1$, et tous reliés entre eux deux à deux par des flèches, sauf le nœud $N+1$ qui n'est relié qu'à lui-même (il constitue une frontière absorbante, la particule qui y arrive y reste). La probabilité de passer du nœud i au nœud j est justement le coefficient p_{ij} de la matrice P . Comme la somme des coefficients de chaque ligne vaut 1, cela signifie que le passage du nœud i à n'importe quel nœud est un évènement sûr, de probabilité 1. La dernière

ligne de la matrice P indique que le nœud $N+1$ constitue bien une fin de trajectoire pour la particule.



Prenons un exemple de trajectoire avec la matrice P ($N = 2$) précédente : 1223. Cela signifie que la particule part du nœud 1 pour aller vers le nœud 2, puis elle revient en 2, et enfin elle va vers 3, et c'est fini. La probabilité d'avoir cette trajectoire est $p_{12} p_{22} p_{23} = 0,2 \times 0,3 \times 0,5 = 0,03$.

Prenons une trajectoire T démarrant au nœud i (i de 1 à N) et se terminant au nœud $N+1$, et associons-lui une variable aléatoire $X(T)$. Si la trajectoire s'écrit sous forme de succession des nœuds $i j k l m$, où i est le nœud de départ, et m le dernier nœud absorbant $N+1$, on prend :

$$X(T) = b_i + v_{ij} b_j + v_{ij} v_{jk} b_k + v_{ij} v_{jk} v_{kl} b_l$$

où les coefficients du type b sont ceux du système linéaire initial. La probabilité associée à $X(T)$ est $p_{ij} p_{jk} p_{kl} p_{lm}$. Cette définition de $X(T)$ se généralise à toute trajectoire T , quelle que soit sa longueur.

Plus précisément, appelons $X(T_i)$, ou X_i , la variable aléatoire associée à une trajectoire T_i ¹ quand celle-ci part du nœud i (i de 1 à N). On a alors la propriété suivante :

Les valeurs moyennes (espérances) des X_i sont égales aux solutions x_i du système initial : $E(X_i) = x_i$.

On admettra ici que l'existence de ces espérances est assurée dès que la norme de la matrice A est inférieure à 1, comme on l'a supposé au début.

Démonstration de la propriété

Pour simplifier l'écriture, nous allons supposer que $N = 2$, et prendre $i = 1$, ce qui signifie que la particule part du nœud 1.

Par définition, $E(X_1) = \sum_{T_1} X(T_1) p(T_1)$, la sommation étant étendue à toutes les trajectoires T_1 démarrant au nœud 1. Après son départ en 1, la particule va vers le nœud suivant qui peut être 1, 2 ou 3. Maintenant regroupons les trajectoires en trois catégories T_{11} , T_{12} , T_{13} , selon le numéro du nœud qui suit le nœud de départ. S'il y a une infinité de chemins T_{11} et T_{12} , il n'y a qu'un chemin T_{13} , On a alors :

$$E(X_1) = \sum_{T_{11}} X(T_{11}) p(T_{11}) + \sum_{T_{12}} X(T_{12}) p(T_{12}) + X(T_{13}) p(T_{13})$$

Pour un chemin T_{11} , par exemple 11213, on a :

¹ Ces notations ne servent qu'à simplifier l'écriture. Nous donnons les mêmes noms T_i et X_i aux multiples trajectoires partant du nœud i et aux variables aléatoires associées à chacune.

$X(T_{11}) = b_1 + v_{11} b_1 + v_{11} v_{12} b_2 + v_{11} v_{12} v_{21} b_1 = b_1 + v_{11} (b_1 + v_{12} b_2 + v_{12} v_{21} b_1)$
 $= b_1 + v_{11} X(T_1)$ où T_1 est le reste de la trajectoire T_{11} à partir du deuxième nœud, avec $p(T_{11}) = p_{11} p(T_1)$. Ainsi :

$$\begin{aligned} E(X_1) &= \sum_{T_1} (b_1 + v_{11} X(T_1)) p_{11} p(T_1) + \sum_{T_2} (b_1 + v_{12} X(T_2)) p_{12} p(T_2) + b_1 p_{13} \\ &= \sum_{T_1} v_{11} p_{11} X(T_1) p(T_1) + \sum_{T_2} v_{12} p_{12} X(T_2) p(T_2) + b_1 (\sum_{T_1} p_{11} p(T_1) + \sum_{T_2} p_{12} p(T_2) + p_{13}) \\ &= v_{11} p_{11} \sum_{T_1} X(T_1) p(T_1) + v_{12} p_{12} \sum_{T_2} X(T_2) p(T_2) + b_1 (p_{11} \sum_{T_1} p(T_1) + p_{12} \sum_{T_2} p(T_2) + p_{13}) \end{aligned}$$

Mais $\sum_{T_1} X(T_1) p(T_1) = E(X_1)$, $\sum_{T_2} X(T_2) p(T_2) = E(X_2)$, et $\sum_{T_1} p(T_1) = 1$, $\sum_{T_2} p(T_2) = 1$.

Finalement :

$$E(X_1) = v_{11} p_{11} E(X_1) + v_{12} p_{12} E(X_2) + b_1 (p_{11} + p_{12} + p_{13}). \text{ Et comme } p_{11} + p_{12} + p_{13} = 1,$$

$$E(X_1) = v_{11} p_{11} E(X_1) + v_{12} p_{12} E(X_2) + b_1$$

$E(X_1) = a_{11} E(X_1) + a_{12} E(X_2) + b_1$, ce qui prouve que les espérances $E(X_1)$ et $E(X_2)$ vérifient la première équation du système initial, et on ferait de même pour les autres équations. On vient de trouver la solution du système, à savoir $E(X_1)$, $E(X_2)$.

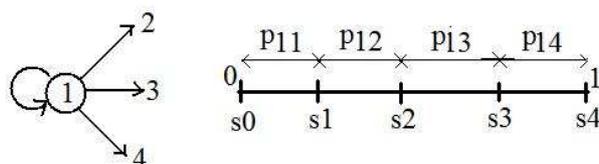
Traitement expérimental

- **Conditions initiales, avec la fonction *init()***

On commence par se donner la matrice P de probabilités, en prenant les p_{ij} au hasard, pour i et j de 1 à N , mais dans les limites d'une somme inférieure à 1 sur chaque ligne, puis en complétant chaque ligne par la probabilité complémentaire $p_{i, N+1}$. Remarquons que la dernière ligne $N+1$ n'est pas utile. Ensuite, on prend au hasard les v_{ij} (i et j de 1 à N) égaux à ± 1 , ce qui permet d'obtenir la matrice A du système que l'on veut résoudre. Enfin, on se donne les b_i (i de 1 à N) au hasard. On a alors toutes les données correspondant au système à résoudre.

Il reste à préparer le mouvement aléatoire de la particule. On sait qu'à partir de chaque nœud (de 1 à N) il y a des jonctions vers les nœuds de 1 à $N+1$, chacune avec sa propre probabilité p_{ij} (de i vers j). Pour chaque nœud i , on prend l'intervalle $[0, 1]$ et on le découpe en intervalles successifs de longueur p_{ij} avec j de 1 à $N+1$. Les graduations obtenues sont appelées *seuil*[i][j] avec j allant de 0 à $N+1$.

Sur le
ses



dessin ci-joint, où $N = 3$, on a pris l'exemple du nœud 1, avec le découpage de l'intervalle $[0, 1]$ et graduations de s_0 à s_4 (il s'agit des seuils).

Pourquoi fait-on cela ? Pour pouvoir obtenir les trajectoires aléatoires de la particule. Il suffira de tirer un nombre au hasard entre 0 et 1, et de déterminer dans quel intervalle il se

trouve. Par exemple, s'il se trouve dans l'intervalle $[s_2, s_3]$ pour le nœud 1, la particule ira du nœud 1 vers le nœud 3.

```
void init(void) /* toutes les variables ont été déclarées en global */
{ srand(time(NULL));
  for(i=1;i<=N;i++) /* la matrice P */
  { do
    { somme=0.;
      for(j=1;j<=N;j++)
        { p[i][j]=(float)rand()/((float)RAND_MAX+1.); somme+=p[i][j]; }
      }
    while (somme>0.9);
    p[i][N+1]=1.-somme;
  }
  /* la matrice A */
  for(i=1;i<=N;i++) for(j=1;j<=N;j++) v[i][j]=2*(rand()%2)-1;
  for(i=1;i<=N;i++) for(j=1;j<=N;j++) a[i][j]=v[i][j]*p[i][j];
  /* les b_i */
  for(i=1;i<=N;i++) b[i]=(rand()%10)*(2*(rand()%2)-1);
  afficher

  for(i=1;i<=N;i++) /* calcul des seuils de probabilité*/
  { cumul=0.; seuil[i][0]=0.;
    for(j=1;j<=N+1;j++) { cumul+=p[i][j]; seuil[i][j]=cumul; }
  }
}
```

- **Programme principal**

Maintenant, à partir de chaque nœud de 1 à N , pris à tour de rôle, on lance des trajectoires, au nombre de $NBEXP$. Chaque trajectoire est placée dans un tableau $t[]$, indexé de 0 à $lastindex$. Pour déterminer le passage d'un nœud au suivant sur la trajectoire, on prend un nombre au hasard $h01$ sur $[0, 1[$, et l'on cherche le numéro du seuil situé juste après lui, ce qui donnera le numéro du nœud suivant. La trajectoire se poursuit jusqu'à ce que l'on arrive au nœud final $N+1$. Pour chaque trajectoire obtenue, on calcule la variable aléatoire correspondante X , puis on cumule les résultats obtenus pour les X de chaque trajectoire. Il suffira de diviser le cumul des X , noté $cumulX$, par le nombre des trajectoires $NBEXP$, pour avoir la valeur moyenne de X , $E(X)$, ce qui donne, comme on l'a vu, les solutions du système d'équations. Enfin, pour tester la validité des résultats, on calcule les seconds membres de chaque équation ($AX + B$), avec les solutions expérimentales trouvées, ce qui redonne dans le membre de gauche les valeurs des solutions. La comparaison des deux valeurs obtenues pour chaque solution indique le degré de précision des résultats.

```
#define N 8
#define NBEXP 5000 /* nombre de trajectoires à partir de chaque nœud */
void init(void);
float p[ N+2][N+2],v[ N+1][N+1], seuil[N+1][N+2],cumul,b[N+1];
int t[10000],lastindex ,i,j,k ,traj;;
float sum[N+1],a[N+1][N+1], h01,X,cumulv,cumulX,x[N+1],somme;

int main()
{
  init();
}
```

```

for(i=1;i<=N;i++)
{ srand(time(NULL));
  cumulX=0.;
  for(traj=1;traj<=NBEXP; traj++)
  { t[0]=i; k=0;
    do
      { h01=(float)rand()/((float)RAND_MAX+1.);
        j=0; while(seuil[t[k]][j]<h01) j++;
          k++; t[k]=j;
        }
      while (t[k]!=N+1);
      lastindex=k;
      X=b[i]; cumulv=1.;
      for(j=1;j<lastindex;j++) { cumulv*=v[t[j-1]][t[j]]; X+=cumulv*b[t[j]]; }
      cumulX+=X;
    }
  x[i]=cumulX/(float)NBEXP;
}
/* affichage */
for(i=1;i<=N;i++)
{ sum[i]=0; for(j=1;j<=N;j++) sum[i]+=a[i][j]*x[j]; sum[i]+=b[i]; }
for(i=1;i<=N;i++) printf("\n%6.2f %6.2f",x[i],sum[i]);
getch(); return 0;
}

```

- Un exemple de résultats, ici pour $N = 8$

La matrice de probabilités P , sans sa dernière ligne, avec ses $N+1$ colonnes

```

0.03  0.01  0.14  0.09  0.01  0.10  0.37  0.09  0.18
0.05  0.05  0.04  0.03  0.02  0.25  0.22  0.02  0.31
0.02  0.38  0.04  0.29  0.05  0.06  0.01  0.00  0.15
0.05  0.13  0.04  0.03  0.04  0.08  0.09  0.13  0.40
0.15  0.15  0.03  0.12  0.01  0.12  0.08  0.08  0.26
0.05  0.26  0.09  0.03  0.12  0.09  0.08  0.00  0.28
0.02  0.09  0.03  0.16  0.21  0.30  0.05  0.03  0.12
0.00  0.03  0.14  0.14  0.08  0.21  0.05  0.05  0.30

```

La matrice A , avec ses N lignes et N colonnes, et dans la colonne $N+1$ les b_i (en ne gardant qu'un chiffre derrière la virgule) :

```

-0.0  -0.0  0.1  0.1  -0.0  0.1  0.4  0.1  -6.00
-0.1  -0.1  0.0  0.0  -0.0  0.3  -0.2  -0.0  2.00
0.0  0.4  -0.0  0.3  0.1  0.1  0.0  -0.0  -2.00
-0.1  0.1  -0.0  0.0  0.0  -0.1  0.1  0.1  -2.00
-0.1  0.1  -0.0  -0.1  -0.0  -0.1  0.1  0.1  -8.00
-0.1  0.3  -0.1  0.0  0.1  0.1  0.1  0.0  1.00
-0.0  0.1  0.0  -0.2  -0.2  0.3  -0.0  -0.0  0.00
0.0  0.0  -0.1  -0.1  0.1  -0.2  -0.1  0.0  3.00

```

La solution pour les $N = 8$ inconnues, avec les deux valeurs approchées obtenues pour chacune

-5.08	-4.99
2.02	2.01
-1.82	-1.84
-1.19	-1.28
-6.36	-6.45
1.56	1.49
2.19	2.06
2.69	2.63