

IFT 6561

Simulation et modèles

Fabian Bastin
DIRO
Université de Montréal

Automne 2013

Considérons le sous-ensemble de $[0, 1]^t$ construit à partir des différents états initiaux possibles du générateur.

$$\Psi_t = \{\mathbf{u} = (u_0, \dots, u_{t-1}) = (g(s_0), \dots, g(s_{t-1})), s_0 \in \mathcal{S}\}.$$

Un critère majeur est que Ψ_t doit recouvrir $[0, 1]^t$ très uniformément, et ce pour “tout” t .

Par généralisation, nous chercherons également à mesurer l'uniformité de $\Psi_I = \{(u_{i_1}, \dots, u_{i_t}) \mid s_0 \in \mathcal{S}\}$ pour une classe choisie d'ensembles d'indices de forme $I = \{i_1, i_2, \dots, i_t\}$. La récurrence linéaire à la base d'un MRG a comme conséquence majeure de produire une structure pour l'ensemble Ψ_t .

Structure de réseau

Soit $(x_0, \dots, x_{k-1})$ dans $\{0, 1, \dots, m-1\}^k$, et la base canonique de $\mathcal{R}^k$:

$$\{\mathbf{e}_i, i = 1, \dots, k\},$$

Si $(x_0, \dots, x_{k-1}) = \mathbf{e}_1 = (1, 0, \dots, 0)$, la récurrence du MRG donne

$$\begin{aligned}(x_1, \dots, x_k) &= (0, \dots, a_k), \\(x_2, \dots, x_k, x_{k+1}) &= (0, \dots, a_k, a_1 a_k \bmod m), \\(x_3, \dots, x_{k+2}) &= (0, \dots, (a_1^2 + a_2) a_k \bmod m), \dots\end{aligned}$$

Si $(x_0, \dots, x_{k-1}) = \mathbf{e}_2 = (0, 1, \dots, 0)$, alors

$$\begin{aligned}(x_1, \dots, x_k) &= (1, 0, \dots, a_{k-1}), \\(x_2, \dots, x_k, x_{k+1}) &= (0, \dots, a_{k-1}, (a_1 a_{k-1} + a_k) \bmod m), \\(x_3, \dots, x_{k+2}) &= (0, \dots, (a_1^2 a_{k-1} + a_1 a_k + a_2 a_{k-1}) \bmod m), \dots\end{aligned}$$

Nous pouvons continuer de la sorte jusqu'à considérer $(x_0, \dots, x_{k-1}) = \mathbf{e}_k = (0, \dots, 0, 1)$, ce qui produit

$$\begin{aligned}(x_1, \dots, x_k) &= (0, \dots, 1, a_1), \\(x_2, \dots, x_k, x_{k+1}) &= (0, \dots, 1, a_1, (a_1^2 + a_2) \bmod m), \dots\end{aligned}$$

Or tout vecteur $(x_n, \dots, x_{n+t-1})$ qui obéit à la récurrence, pour $t \geq k$, est une combinaison linéaire à coefficients entiers de ces k vecteurs de base.

Pour le voir, notons $x_{i,0}, x_{i,1}, x_{i,2}, \dots$ la suite obtenue à partir du vecteur de base $\mathbf{e}_i$. Un état initial $(x_0, \dots, x_{k-1}) = (z_1, \dots, z_k)$ peut s'écrire comme $z_1 \mathbf{e}_1 + \dots + z_k \mathbf{e}_k$ et produit la suite

$$z_1(x_{1,0}, x_{1,1}, \dots) + \dots + z_k(x_{k,0}, x_{k,1}, \dots) \bmod m,$$

et réciproquement.

Structure de réseau

La réduction modulo m se fait en soustrayant des vecteurs $m\mathbf{e}_j$. Ainsi, pour $t \geq k$, $(x_0, x_1, \dots, x_{t-1})$ suit la récurrence si et seulement s'il s'agit d'une combinaison linéaire à coefficients entiers de

$$(1, 0, \dots, 0, x_{1,k}, \dots, x_{1,t-1})$$

$$\vdots$$

$$(0, 0, \dots, 1, x_{k,k}, \dots, x_{k,t-1})$$

$$(0, 0, \dots, 0, m, \dots, 0)$$

$$\vdots$$

$$(0, 0, \dots, 0, 0, \dots, m).$$

Structure de réseau

En divisant par m , on obtient que $(u_0, \dots, u_{t-1}) \in [0, 1]^t$ est dans Ψ_t si et seulement si c'est une combinaison linéaire (sur les entiers) de

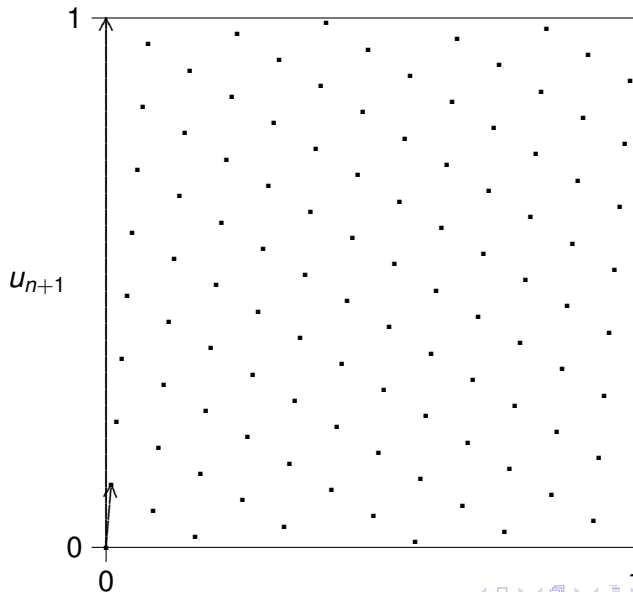
$$\begin{aligned}\mathbf{v}_1 &= (1, 0, \dots, 0, x_{1,k}, \dots, x_{1,t-1})^T / m \\ &\vdots \\ \mathbf{v}_k &= (0, 0, \dots, 1, x_{k,k}, \dots, x_{k,t-1})^T / m \\ \mathbf{v}_{k+1} &= (0, 0, \dots, 0, 1, \dots, 0)^T \\ &\vdots \\ \mathbf{v}_t &= (0, 0, \dots, 0, 0, \dots, 1)^T.\end{aligned}$$

Si

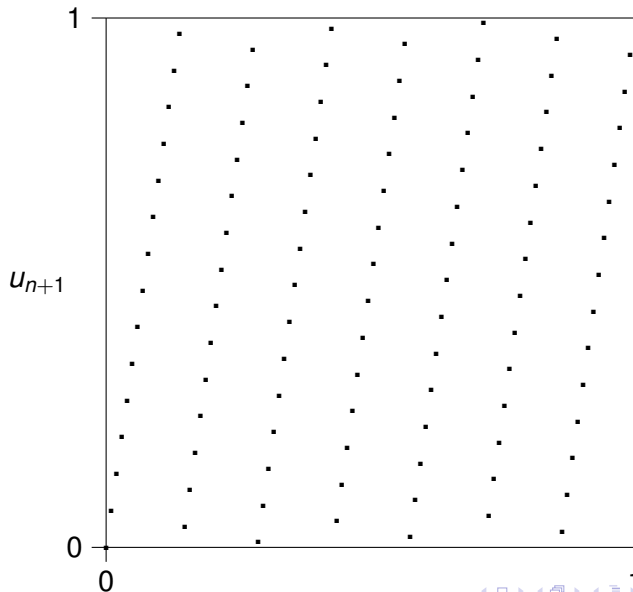
$$L_t = \left\{ \mathbf{v} = \sum_{i=1}^t z_i \mathbf{v}_i \mid z_i \in \mathcal{Z} \right\}$$

est le réseau ayant ces vecteurs pour base, alors
 $\Psi_t = L_t \cap [0, 1)^t$.

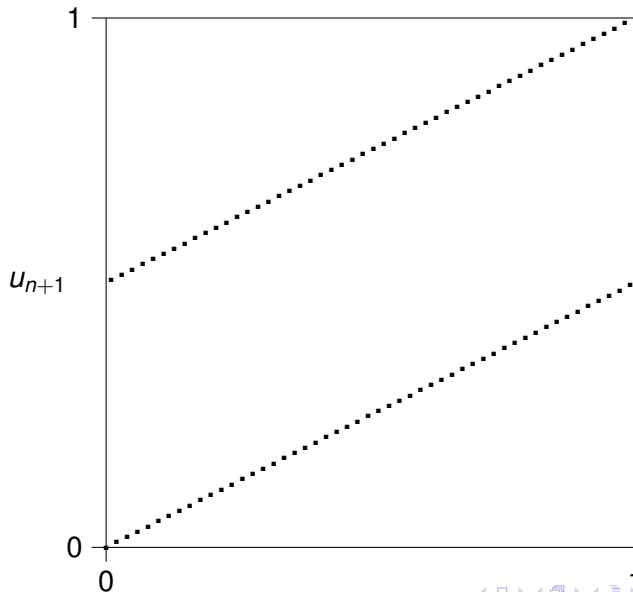
$$m = 101, a = 12; \mathbf{v}_1 = (1, 12)/101, \mathbf{v}_2 = (0, 1)$$



LCG with $m = 101$ and $a = 7$



LCG with $m = 101$ and $a = 51$



LCG with $m = 101$ and $a = 51$

Pour $t > k$, il y a m^t vecteurs dont les coordonnées sont des multiples de $1/m$, mais seulement m^k sont dans Ψ_t (il n'y a que m^k états initiaux possible), soit une proportion de $1/m^{t-k}$.

Cette structure de réseau implique que les points de Ψ_t sont distribués suivant un schéma très régulier.

Par exemple, chaque point de L_t a un plus proche voisin à la même distance et dans la même direction que n'importe quel autre point, et il y a des familles d'hyperplans parallèles équidistants qui couvrent tous les points.

Au regard de cette structure particulière, des manières naturelles de mesurer l'uniformité de ce genre d'ensemble de points Ψ_t incluent:

- 1 la distance d'un point à son plus proche voisin, qui est aussi la distance de l'origine au point le plus proche, ou de manière équivalente le vecteur non-nul le plus court dans L_t ;
- 2 La distance entre les deux hyperplans les plus éloignés qui couvrent une région ne contenant aucun point de L_t ;
- 3 le nombre minimum d'hyperplan équidistants parallèles qui peuvent couvrir tous les points de Ψ_t .

Distance au plus proche voisin

Pour obtenir la première mesure, nous devons calculer le plus court vecteur non nul dans un réseau de base $\mathbf{v}_1, \dots, \mathbf{v}_t$:

$$\text{Minimiser } \|\mathbf{v}\|_2^2 = \sum_{i=1}^t \sum_{j=1}^t z_i \mathbf{v}_i^T \mathbf{v}_j z_j$$

sous les contraintes que $z_1, \dots, z_t$ soient entiers et non tous nuls.

Il s'agit par conséquent d'un problème d'optimisation quadratique en nombres entiers, lequel est difficile à résoudre.

Distance entre les hyperplans

Pour traiter les deux autres mesures proposées, nous avons tout d'abord besoin de définir le réseau dual, dénoté L_t^* , comme suit:

$$L_t^* = \{\mathbf{h} \in \mathcal{R}^t : \mathbf{h} \cdot \mathbf{v} \in \mathcal{Z} \text{ pour tout } \mathbf{v} \in L_t\}.$$

Pour chaque vecteur $\mathbf{h} \in L_t^*$ et chaque entier z , l'ensemble $\{\mathbf{v} \in \mathcal{R}^t : \mathbf{h}^T \mathbf{v} = z\}$ est un hyperplan orthogonal à $\mathbf{h}$. Quand z parcourt tous les entiers, nous obtenons dès lors une famille d'hyperplans parallèles qui couvrent tous les points $\mathbf{v}$ de L_t , car $\mathbf{h}^T \mathbf{v} \in \mathcal{Z}$ pour $\mathbf{v} \in L_t$.

La distance entre deux hyperplans est la distance entre l'hyperplan défini avec $z = 1$ et celui défini avec $z = 0$, i.e., à l'origine, puisque ce dernier contient l'origine.

Distance entre les hyperplans

Il est possible de montrer que cette distance vaut $1/\|\mathbf{h}\|_2$. Si ℓ_t est la longueur du plus court vecteur $\mathbf{h}$ non nul dans L_t^* , alors la distance entre les hyperplans pour la famille où ils sont le plus éloignés est $1/\ell_t$. Par conséquent, nous cherchons à maximiser ℓ_t .

Un petit nombre d'hyperplans parallèles couvrant tous les points de ψ_t peut être trouvé en calculant le vecteur non-nul le plus court dans L_t^* en utilisant la norme $\mathcal{L}_1$ au lieu de la norme euclidienne.

Plutôt de que considérer des indices consécutifs, nous pouvons considérer n'importe quel ensemble de t index (distincts)

$I = \{i_1, i_2, \dots, i_t\}$. Dans ce cas, nous avons

$$\begin{aligned}\Psi_I &= \{(u_{i_1}, \dots, u_{i_t}) \mid s_0 = (x_0, \dots, x_{k-1}) \in \mathcal{Z}_m^k\}, \\ &= L_I \cap [0, 1)^t,\end{aligned}$$

$$1/\ell_I = \text{distance entre les hyperplans dans } L_I.$$

Note: $\mathcal{Z}_m = \{0, 1, 2, \dots, m-1\}$.

Mesures d'uniformité

Des bornes supérieures de la forme $\ell_t \leq \ell_t^*(n)$ pour un réseau général de densité n dans $\mathcal{R}^t$ existent.

Nous pouvons dès lors standardiser ℓ_t par $\ell_t/\ell_t^*(m^k)$ pour avoir une mesure dans $[0, 1]$, ce qui permet d'obtenir la figure de mérite générale:

$$M_{\mathcal{J}} = \min_{I \in \mathcal{J}} \ell_I/\ell_{|I|}^*(m^k)$$

où $\mathcal{J}$ est une famille d'ensembles $I = \{i_1, i_2, \dots, i_t\}$.

La recherche de générateurs potentiellement intéressants passe alors par le calcul numérique des paramètres qui maximisent cette mesure.

Si $i \in I$ lorsque $a_{k-i} \neq 0$ (avec $a_0 = -1$), alors

$$\ell_i^2 \leq 1 + a_1^2 + \cdots + a_k^2.$$

Il faut donc que cette somme soit grande!

Preuve partielle.

Considérons $I = \{0, \dots, t-1\}$ et soit

$$\mathbf{h} = (-a_k, \dots, -a_1, 1, 0, \dots, 0)^T.$$

Si $\mathbf{v} = (v_0, v_1, \dots, v_{t-1})^T \in L_t$, alors $v_k = (a_1 v_{k-1} + \cdots + a_k v_0) \bmod 1$.

Ainsi, $\mathbf{h}^T \mathbf{v}$ doit être un entier, i.e., $\mathbf{h} \in L_t^*$.

Donc $\ell_t^2 \leq \|\mathbf{h}\|_2^2 = 1 + a_1^2 + \cdots + a_k^2$.

Se généralise au cas de ℓ_I .



Exemple: Lagged-Fibonacci

$$x_n = (\pm x_{n-r} \pm x_{n-k}) \mod m.$$

Pour $I = \{0, k-r, k\}$, on a $1/\ell_I \geq 1/\sqrt{3} \approx .577$.

Les vecteurs $(u_n, u_{n+k-r}, u_{n+k})$ sont tous contenus dans deux plans!

Fonction `random` de la glibc

L'exemple à éviter...

Le générateur commence par initialiser un tableau de 34 nombres sur le principe du générateur Standard Minimal:

- 1 $x_0 = s; x_i = 16807 * x_{i-1} \bmod 2^{31} - 1$ (pour $i = 1, \dots, 30$);
- 2 $x_i = x_{i-31}$ (pour $i = 31, \dots, 33$).

Pour $i \geq 34$, l'algorithme suit l'algorithme de Lagged-Fibonacci, avec $m = 2^{31}$:

$$x_i = (x_{i-3} + x_{i-31}) \bmod 2^{31}.$$

Enfin, la fonction ignore les 344 premiers nombres et supprime le bit de poids faible:

$$o_i = x_{i+344} \gg 1.$$

Fonction `random` de la glibc: observations

Le générateur résume ce qu'il convient de ne pas faire:

- 1 le générateur Standard Minimal est dépassé;
- 2 les vecteurs x_i, x_{i+28}, x_{i+31} sont contenus dans deux plans;
- 3 la linéarité n'est pas complètement supprimée en enlevant le bit de poids faible, vu que pour $i \geq 34$,

$$o_i = o_{i-31} + o_{i-3} \mod 2^{31} \text{ ou } o_i = o_{i-31} + o_{i-3} + 1 \mod 2^{31},$$

et on conserve la majeure partie des inconvénients liés à $m = 2^{31}$.

RANDU (IBM)

$$x_{n+1} = 65539x_n \mod 2^{31}$$

