

XIV Conferencia Latinoamericana de Informática
17avas. Jornadas Argentinas de Informática
e Investigación Operativa.
Buenos Aires, Setiembre 1988.

**UN ALGORITMO PARA EL CALCULO DE LA PRUEBA ESPECTRAL
EN EL CASO DE GENERADORES CONGRUENCIALES LINEALES
GCL(A,C,M) CON PERIODO MAXIMO**

Rafael A. Pastoriza

**ESCUELA DE COMPUTACION, FACULTAD DE CIENCIAS,
UNIVERSIDAD CENTRAL DE VENEZUELA**

Apartado 47002, Caracas, Venezuela.

RESUMEN

Se propone un nuevo algoritmo computacional - SP - para resolver la prueba espectral. Está basado en la sucesiva resolución de problemas simples de programación entera cuadrática sujeta a restricciones lineales. El algoritmo propuesto evita problemas numéricos y de terminación.

1. INTRODUCCION

La prueba espectral para generadores congruenciales lineales - GCL - es una prueba de carácter "teórico" que tiene por objeto evaluar a priori la performance relativa de un GCL(A,C,M) con respecto a las propiedades aleatorias teóricas esperadas, conociendo sólo el valor numérico de los parámetros A, C, M que especifican un determinado generador.

$$X_{n+1} = A * X_n + C \pmod{M}$$

Al decir de Knuth [8, p.82] la importancia de la prueba espectral radica no sólo en el hecho de que todos los buenos generadores conocidos la pasan, sino que para todas aquellas sucesiones, reconocidamente malas, los generadores de números aleatorios - GNA - que las producen no pasan dicha prueba.

Es bien conocido (ver Marsaglia [10]) que si se dibujan las sucesivas n-uplas $Y_i = (X_i, X_{i+1}, \dots, X_{i+n-1})$, $Y_{i+1} = (X_{i+1}, X_{i+2}, \dots, X_{i+n})$, $\dots, Y_k, \dots$ donde los valores que toman las coordenadas X_i provienen de un generador GCL(A,C,M) los puntos Y_i caen todos en un reticulado del espacio de n-dimensiones. En general ese reticulado es regular, y este es el caso para generadores GCL con período máximo.

La prueba espectral, originalmente propuesta por Coveyou y Macpherson [1], proporciona un método para encontrar la distancia M/F_n entre la familia de hiperplanos que está más separada, en n-dimensiones, y calcula un valor G_n que mide esa distancia. La idea presentada por esos autores era desarrollar un método de análisis de los GNA, aplicable a todos los métodos prácticos de generación, y ellos propusieron con ese fin realizar el análisis de Fourier de las sucesiones resultantes de operar los GNA.

Se tiene (ver Knuth [8, p.83]) que para sucesiones "verdaderamente" al azar provenientes de la distribución uniforme, la transformada de Fourier de la función constante $1/M^n$ será siempre uno para los valores congruentes con cero (mod M), y cero en todo otro caso.

Cuando un GCL(A,C,M,X0) tiene período máximo su transformada de Fourier $F(S_1, \dots, S_n)$ de las n-uplas es :

$$e^{(-2\pi i C i/M) * (S(A)-S(1))/(A-1)}$$

$$\text{si } S(A) = 0 \pmod{M}$$

$$F(S_1, \dots, S_n) =$$

$$0 \quad \text{si no}$$

donde

$$S(A) = S_1 + S_2 * A + \dots + S_n * A^{n-1}$$

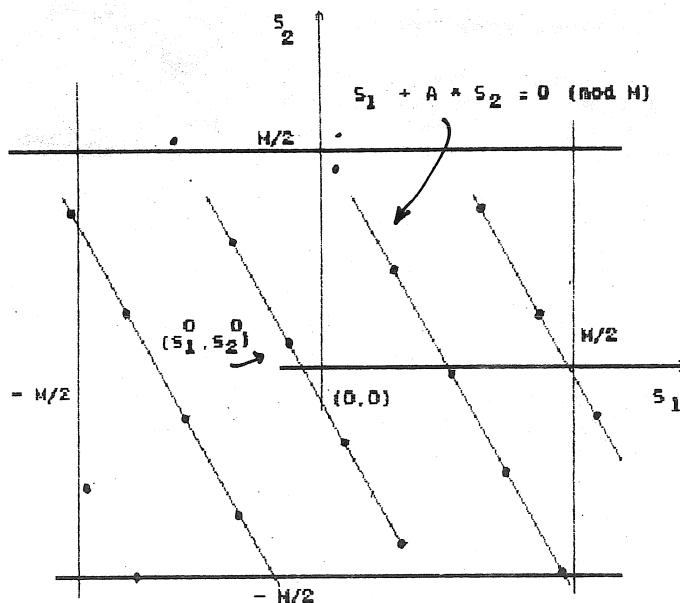
que es la expresión de lo que se denomina el espectro de un $GCL(A, C, M, X_0)$.

En el texto de Fishman [2, p.363] se pueden encontrar también las expresiones de las transformadas de Fourier para los generadores GNA que no tienen periodo máximo, como son los generadores congruenciales lineales multiplicativos modulo primo - GCLMMP -.

Entonces una manera de medir la uniformidad de un $GCL(A, C, M)$, en n dimensiones, es medir para cuales puntos $(S_1, S_2, \dots, S_n)$ se aproxima $F(S_1, S_2, \dots, S_n)$, la transformada de Fourier de la sucesión a $f(S_1, S_2, \dots, S_n)$, ya que se puede ver que la transformada para las n-uplas es :

$$f(S_1, S_2, \dots, S_n) = \begin{cases} 1 & \text{si } S_1=0 \pmod{M}, \dots, S_n=0 \pmod{M} \text{ y } S(A)=0 \pmod{M} \\ 0 & \text{si no} \end{cases}$$

Se puede entonces tratar de identificar el punto $(S_1, S_2, \dots, S_n)$ en $S(A) = \pmod{M}$ más cercano al punto $(0, 0, \dots, 0)$. La figura siguiente proporciona una imagen del problema asociado para el caso de dos dimensiones.



En la prueba espectral se usa la norma euclídea para caracterizar el vector $(S_1, S_2, \dots, S_n)$:

$$N_n(A) = (S * S^t)^{1/2} = (S_1^2 + \dots + S_n^2)^{1/2}$$

Luego para seleccionar el peor caso posible, se debe encontrar el punto

$S = (S_1^0, \dots, S_n^0)$ con S_i enteros
que
minimiza $N_n(A)$

sujeto a

$$S(A) = 0 \pmod{M}$$

El análisis se facilita (ver Knuth [1, p.104]) si se cambia el dominio de los valores $0 \leq S_i < M$ al $-M/2 \leq S_i < M/2$, y entonces los S_i no deben ser todos cero.

2. LA PRUEBA ESPECTRAL

En resumen, dado un generador congruencial lineal $GCL(A,C,M)$ con los coeficientes especificados de manera tal que tenga periodo máximo (ver Hull y Dobell [7, p.233]), la prueba espectral se reduce al siguiente procedimiento :

1. Dado n , valor de la n -uniformidad que se desea evaluar apriori, resolver el siguiente problema :

problema.PRESP :

$$F_n = (S_1^2 + \dots + S_n^2)^{1/2} \longrightarrow \min$$

sujeto a:

$$-M/2 \leq S_i < M/2 \quad \text{con } i = 1, 2, \dots, n$$

$$(S_1, S_2, \dots, S_n) \neq (0, 0, \dots, 0)$$

$$S_i \text{ en } \mathbb{Z}[-M/2, M/2], \text{ enteros, con } i = 1, 2, \dots, n$$

$$S(A) = S_1 + A * S_2 + \dots + A^{n-1} * S_n = 0 \pmod{M}$$

2. Evaluar el valor C_n , propuesto por Knuth [8, p.87] donde :

$$C_n = (\pi F_n)^n / ((n/2)! M)$$

donde el símbolo $(n/2)!$ está definido por la expresión

$$(n/2)! = (n/2) (n/2-1) \dots (1/2) (\pi)^{1/2} \text{ para } n \text{ impar}$$

El valor de C_n es una cantidad que mide la efectividad de multiplicador A en una sucesión proveniente de un generador GCL con periodo máximo. Valores pequeños de C_n corresponden a generadores GCL no aleatorios.

El algoritmo S propuesto por Knuth [8, p.93] proporciona un procedimiento computacional para realizar la prueba espectral, y fué implantado en FORTRAN por Golder [3, p.173] con algunas incorrecciones. Tiempo después las condiciones correctas fueron dadas por Hoaglin y King [5, p.376].

Posteriormente Hopkins [6, p.328] en 1983 publica otra version FORTRAN del algoritmo S, que se corresponde con la nueva versión corregida para evitar problemas de terminación, e incluida por Knuth en la segunda edición del volumen II de su texto The Art of Computer Programming [9, p.98].

Cualquiera sea el caso dicho algoritmo requiere de un esfuerzo computacional muy grande, y en la práctica no se lo ha usado en equipos mainframe para evaluar la performance de los GCL más conocidos, más allá de $n = 6$. La búsqueda exhaustiva implícita en el algoritmo hace que el tiempo de ejecución crezca como 3^n , (ver Knuth [9, p.100]).

3. EL ALGORITMO SP

Aquí se propone un procedimiento computacional diferente y muy simple para resolver el problema PRESP.

Se trata de resolver de manera secuencial el problema asociado, definido por el funcional cuadrático, y como el término $S(A)$ se puede escribir

$$S(A) = S_1 + A(S_2 + A(S_3 + \dots + A(S_{n-1} + AS_n)) \dots) = 0 \pmod{M}$$

emplear entonces el algoritmo SP siguiente :

ALGORITMO SP

Resolver sucesivamente los problemas auxiliares :

$$\text{PAUX1 : } \min F_1(S_{n-1}, S_n) = S_{n-1}^2 + S_n^2$$

sujeto a:

$$-M/2 \leq S_{n-1} < M/2$$

$$-M/2 \leq S_n < M/2$$

$$S_i \text{ en } \mathbb{Z}[-M/2, M/2], \text{ enteros, } i = n-1, n$$

$$H_1(A) = S_{n-1} + A * A_n = 0 \pmod{M}$$

$$(S_{n-1}, S_n) \neq (0, 0)$$

Sea (S_{n-1}^*, S_n^*) una solución óptima de PAUX1 (si existe), luego resolver el siguiente problema auxiliar :

$$\text{PAUX2 : } \min F_2(S_{n-2}, H_1) = S_{n-2}^2 + F_1(S_{n-1}, S_n)$$

sujeto a:

$$-M/2 \leq S_{n-2} < M/2$$

$$-M/2 \leq H_1 < M/2$$

$$S_{n-2}, H_1 \text{ en } \mathbb{Z}[-M/2, M/2], \text{ enteros}$$

$$H_2(A) = S_{n-2} + A * H_1 = 0 \pmod{M}$$

$$(S_{n-2}, H_1) \neq (0, 0)$$

Sea entonces (S_{n-2}^*, H_1^*) una solución óptima de PAUX2 (si existe), lo cual implica resolver también el problema PAUX1 con el valor $H_1 = S_n^*$, cuya solución, si existe, será (S_{n-1}^{**}, S_n^{**}) . Luego una solución óptima de PAUX2 será $(S_{n-2}^*, S_{n-1}^{**}, S_n^{**})$. Y así sucesivamente.

Es claro que este procedimiento se deberá repetir tantas veces como sea necesario para resolver el problema PRESP.

Luego para resolver el problema PRESP mediante el algoritmo SP, basta saber resolver el problema PAUX con un término genérico B cualquiera :

$$\text{PAUX : } \min F(x, y) = x^2 + y^2$$

sujeto a :

$$-M/2 \leq x < M/2$$

$$-M/2 \leq y < M/2$$

$$x, y \text{ en } \mathbb{Z}[-M/2, M/2], \text{ enteros}$$

$$H(A, B) = x + A * y = B \pmod{M}$$

$$(x, y) \neq (0, 0)$$

y construir las soluciones $(S_1^0, S_2^0, \dots, S_n^0)$ al problema original PRESP en función de las soluciones (x^0, y^0) a los sucesivos problemas auxiliares. Luego, con el valor de las soluciones se deberá, como antes, evaluar el valor de C_n .

4. REFERENCIAS

1. Coveyou, R.R. y Macpherson, R.D. : "Fourier Analysis of Uniform Random Number Generators", J.A.C.M., vol. 14, num. 1, p.100/109, (1967).
2. Fishman, G.S. : "Principles of Discrete Event Simulation", ed. J.Wiley, N.Y., (1970).
3. Golder, E.R. : "The Spectral Test for the Evaluation of Congruential Pseudo-random Generators", Appl. Stat., vol. 25, num. 2, p.173/180, (1976).
4. Golder, E.R. : "Remark ASR18 : The Spectral Test for the Evaluation of Congruential Pseudo-random Generators", Appl. Stat., vol. 25, num. 3, p.324, (1976).
5. Hoaglin, D.C. y King, M.L. : "Remark ASR24: A Remark on Algorithm AS 98 : The Spectral Test for the Evaluation of Congruential Pseudo-random Generators", Appl. Stat., vol. 27, num. 3, p.375/377, (1978).
6. Hopkins, T.R. : "A Revised Algorithm for the Spectral Test", Appl. Stat., vol. 32, num.3, p.328/334, (1983).
7. Hull, T.E. y Dobell, A.R. : "Random Number Generators", SIAM Rev., vol. 4, p.229/254, (1962).
8. Knuth, D.E. : "The Art of Computer Programming", vol. 2, ed. Addison-Wesley, Reading, Massachusetts, p.82, (1969).
9. Knuth, D.E. : "The Art of Computer Programming", vol. 2, second edition, ed. Addison-Wesley, Reading, Massachusetts, p.89, (1969).
10. Marsaglia, G. : "Random numbers fall mainly in the planes", Proc. Mat. Acad. Sci. U.S.A., vol. 61, p.25/28, (1968).