

AUTÓMATAS CELULARES COMO MODELOS ESTOCÁSTICOS ESPACIO-TEMPORALES

Mario Rodríguez Riotorto¹

¹Departamento de Matemáticas
IES Punta Candieira (Cedeira)

RESUMEN

Los Autómatas Celulares (ACs) se utilizan como modelos discretos de procesos complejos en las áreas de investigación computacional, física, química, social y biológica; se caracterizan por estar formados por una población de agentes o células que interactúan localmente cambiando sus estados y produciendo configuraciones globales de difícil predicción. El punto de vista que se trata aquí es el de considerar los ACs como modelos estocásticos para el análisis estadístico de datos espacio-temporales, centrandó el estudio sobre los estimadores de máxima verosimilitud y sus propiedades.

Palabras e frases chave: autómatas celulares, máxima verosimilitud, estadística espacio-temporal.

Clasificación AMS: 37B15, 62M99, 62F12.

1. INTRODUCCIÓN

Introducido por John von Neumann (von Neumann, 1966), un AC se compone de un cierto número de objetos o *células* que interactúan localmente y que podemos imaginar situados en puntos de la recta o del plano. Cada una de estas células se encuentra en cada momento en uno de los estados internos del conjunto E , evolucionando éstos en tiempo discreto según una *regla de transición local* que hace depender el estado de un objeto en el momento t del estado de sus vecinos y del suyo propio en el momento anterior $t - 1$. Si bien la dinámica de un AC está determinada por reglas que actúan síncrona y localmente, la observación en conjunto de estos objetos interactuantes, como un todo, permite apreciar el surgimiento de estructuras y comportamientos globales emergentes, que difícilmente se pueden prever a partir de las especificaciones locales.

La interacción de múltiples agentes intercambiando información localmente es frecuente en muchos ámbitos de la naturaleza, cualquiera que sea la escala de observación. No es pues de extrañar que los ACs se utilicen como modelos formales en la simulación de procesos espacio-temporales. Es especialmente extensa la literatura sobre su aplicación a procesos físicos y químicos (Vichniac, 1984; Wolfram, 1986; Gerhardt et al., 1989), biológicos (Hogeweg, 1988; Luthi et al., 1998) o sociales (Couclelis, 1985; Deadman et al., 1993).

Ante un AC, la primera pregunta que se plantea es cómo serán sus configuraciones globales una vez transcurrido cierto tiempo y qué propiedades macroscópicas cualitativas y cuantitativas se podrán observar. Después de muchas simulaciones con ordenador de modelos deterministas, Wolfram (1984) consiguió fijar cuatro clases de comportamiento bien diferenciado. En la clase I, la configuración global tiende a un patrón espacialmente homogéneo, independientemente del estado inicial; la clase II la forman los ACs que alcanzan

un estado estable formado por estructuras periódicas espaciales; los que conforman la clase III siguen una dinámica caótica, por lo que pueden utilizarse como generadores de números pseudoaleatorios, tal como hace el *Mathematica* (Wolfram, 1996), el cual, por otra parte, ha sido el entorno de programación y simulación utilizado en este trabajo; por último, son los autómatas de la clase IV los que presentan estructuras más complejas, habiéndose conjeturado que estos ACs son capaces de ejecutar tareas computacionales equivalentes a las que realizan las *máquinas de Turing* (Wolfram, 1984). Como se verá, en los ACs estocásticos volverán a aparecer algunas de las características asociadas a las clases I, II y III.

2. AUTÓMATAS CELULARES ESTOCÁSTICOS

En general, un *proceso estocástico espacio-temporal* es una familia $\mathbf{X} = \{X_{i,j} : i \in S, j \in \mathbb{N}\}$ de variables aleatorias definidas sobre un mismo espacio probabilístico $(\Omega, \mathcal{B}, \Pr)$. La variable $X_{i,j}$ se asocia a la célula o localización i del espacio S , finito o infinito, y al instante j de una escala temporal que aquí se supondrá discreta. Todas las variables $X_{i,j}$ son susceptibles de tomar valores en lo que se llamará *conjunto de estados* $E = X_{i,j}(\Omega)$. En ocasiones será útil hacer referencia a subprocesos de \mathbf{X} ; por ejemplo, $\mathbf{X}_{A,j} = \{X_{i,j} : i \in A\}$ da cuenta de la configuración del subespacio $A \subset S$ justo en el instante $j \in \mathbb{N}$; también, si se hace $N_k = \{0, 1, \dots, k\}$, $\mathbf{X}_{A,N_k} = \{X_{i,j} : i \in A, j \in N_k\}$ se referirá al proceso temporal multivariante de los estados de las células en A desde el momento inicial $j = 0$ hasta $j = k$; casos particulares son \mathbf{X}_{i,N_k} y \mathbf{X}_{S,N_k} . Como suele ser habitual, la misma notación, pero en minúsculas, servirá para indicar realizaciones de estas mismas familias de variables. En general, se identificará por f a la función de probabilidad de cualquiera de los subprocesos anteriores, quedando claro del contexto a cuál de ellos se refiere. Finalmente, dos variables $X_{i,j}$ y $X_{i',j'}$ se dirá que son *contemporáneas* si $j = j'$.

Un *autómata celular estocástico* (ACE) se caracteriza por el hecho de que el estado de la célula en $i \in S$ en el instante $j > 0$ depende de los estados de las células vecinas a i en el momento $j - 1$. En general, la vecindad de i vendrá dada por un conjunto no vacío $V_i \subset S$, llamándose a $V = \{V_i : i \in S\}$ *esquema de vecindad* del ACE; por ejemplo, si S es un espacio métrico de distancia d , podemos definir, para un radio r fijo, $V_i = \{u \in S : d(i, u) \leq r\}$. Se puede ya definir un ACE:

Definición 1. *El proceso estocástico espacio-temporal \mathbf{X} es un ACE si para cada sitio i del espacio finito S y cada instante $j > 0$, la distribución de la variable $X_{i,j}$, condicionada a toda la historia anterior, $\mathbf{x}_{S,N_{j-1}}$, y a las otras variables contemporáneas, $\mathbf{x}_{S \setminus \{i\},j}$, es igual a la distribución de $X_{i,j}$ condicionada al estado de su vecindad en el instante anterior, $\mathbf{x}_{V_i,j-1}$. Más formalmente,*

$$f(x_{i,j} \mid \mathbf{x}_{S,N_{j-1}}, \mathbf{x}_{S \setminus \{i\},j}) = f(x_{i,j} \mid \mathbf{x}_{V_i,j-1}), \forall i \in S, j \in \mathbb{N}^*.$$

Estas probabilidades condicionadas, o reglas de transición local, junto con la distribución inicial $f(\mathbf{x}_{S,0})$ determinan la dinámica del ACE.

Si un ACE de $|S|$ células es observado hasta el momento $k \in \mathbb{N}^*$, el siguiente resultado permite factorizar la distribución de los datos observados:

Proposición 1. *Sea \mathbf{X} un ACE y considérese el subproceso \mathbf{X}_{S,N_k} , entonces su función de probabilidad marginal será*

$$f(\mathbf{x}_{S,N_k}) = f(\mathbf{x}_{S,0}) \cdot \prod_{i \in S} \prod_{j=1}^k f(x_{i,j} \mid \mathbf{x}_{V_i,j-1}). \quad (1)$$

Visto desde otra perspectiva, el ACE define un proceso temporal multivariante:

Proposición 2. *Sea \mathbf{X} un ACE. La sucesión de variables aleatorias $(\mathbf{X}_{S,0}, \mathbf{X}_{S,1}, \mathbf{X}_{S,2}, \dots)$ es una cadena de Markov de estados en $E^{|S|}$.*

En general, el número de estados de la cadena de Markov es potencialmente enorme, de forma que el cálculo de la matriz de transición se hará impracticable; por ejemplo, un ACE de $|S| = 50$ células y $|E| = 2$ estados dará lugar a una matriz de orden $2^{50} \approx 10^{15}$. Aunque sea difícil el cálculo de la distribución estacionaria, $f(\mathbf{x}_{S,\infty})$, interesa el siguiente resultado:

Proposición 3. *Sea \mathbf{X} un ACE con conjunto de estados E finito. La cadena de Markov $(\mathbf{X}_{S,0}, \mathbf{X}_{S,1}, \mathbf{X}_{S,2}, \dots)$ es ergódica si la regla de transición cumple la condición de positividad, $f(x_{i,j} | \mathbf{x}_{V_i,j-1}) > 0, \forall i, j$.*

Si la regla de transición local se especifica de forma paramétrica, lo que interesa a efectos de estimación, podemos utilizar la notación $f(x_{i,j} | \mathbf{x}_{V_i,j-1}; \theta)$, siendo θ el vector paramétrico correspondiente.

Para concretar, se realizará el estudio sobre un ACE de estados binarios que se llamará *logístico*, el cual se construye en los términos que se indica a continuación. El conjunto de estados es $E = \{0, 1\}$ y las células se sitúan sobre los puntos $S = \{1, 2, \dots, m\}$ de la recta entera; el esquema de vecindad es el llamado de radio unidad, $V_i = \{i-1, i, i+1\}$ para $i = 2, \dots, m-1$, uniendo los extremos de forma que $V_1 = \{m, 1, 2\}$ y $V_m = \{m-1, m, 1\}$. Finalmente, la distribución condicionada de probabilidad que define la regla de transición local, y que depende del vector paramétrico $\theta = (\alpha, \beta)$ es

$$f(x_{i,j} | \mathbf{x}_{V_i,j-1}; \alpha, \beta) = \frac{\exp(x_{i,j} \cdot (\alpha + \beta \cdot s_{i,j}))}{1 + \exp((\alpha + \beta \cdot s_{i,j}))},$$

siendo $x_{i,j} \in \{0, 1\}$ y $s_{i,j} = \sum_{u \in V_i} x_{u,j-1}$. Tomando ahora en consideración la factorización (1), la log-verosimilitud de la muestra \mathbf{x}_{S,N_k} viene dada por

$$\ell(\mathbf{x}_{S,N_k}; \alpha, \beta) = \ln f(\mathbf{x}_{S,0}) + \alpha \sum_{i \in S} \sum_{j=1}^k x_{i,j} + \beta \sum_{i \in S} \sum_{j=1}^k x_{i,j} s_{i,j} - \sum_{i \in S} \sum_{j=1}^k \ln(1 + \exp(\alpha + \beta s_{i,j})),$$

de donde los estimadores de máxima verosimilitud se obtienen haciendo

$$(\hat{\alpha}, \hat{\beta}) = \arg \max_{(\alpha, \beta)} \ell(\mathbf{x}_{S,N_k}; \alpha, \beta).$$

3. APLICACIÓN A DATOS SIMULADOS

En la figura 1 se representan nueve simulaciones para distintos valores de (α, β) ; en todos los casos, el tiempo transcurre de izquierda a derecha, partiendo de un mismo estado inicial generado de forma aleatoria, siendo el esquema de vecindad y la regla de transición local las descritas en la sección anterior. Como se puede apreciar, el simple cambio de los parámetros provoca en el ACE comportamientos dinámicos cualitativamente muy diferentes. Los casos *a)* y *b)* tienden rápidamente hacia configuraciones homogéneas con todas las células a 0 ó 1; el caso *c)* presenta un comportamiento casi cíclico, donde se alternan configuraciones con amplias zonas homogéneas separadas por vetas; el caso *d)* es singular, con $\alpha = \beta = 0$, ignora las dependencias locales, siendo el cuadro que presenta aleatorio, uniforme y falto de correlaciones entre las variables; los casos *e)* y *f)* presentan clústeres espaciales en los que pueden dominar células en estado 0 ó 1; finalmente, *g)*, *h)* e *i)* pueden ser quizás los casos

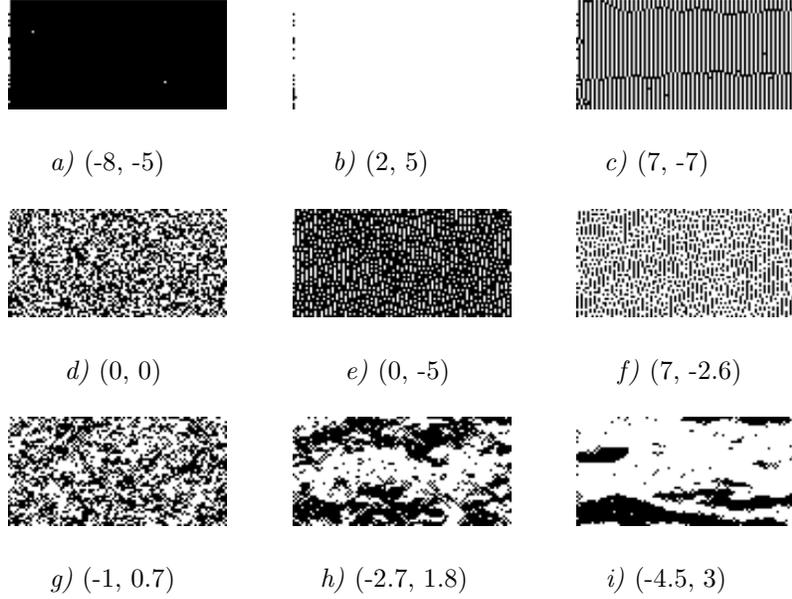


Figura 1: Diferentes simulaciones del ACE logístico para distintos valores del par paramétrico (α, β) . Puntos negros representan células en estado 0, puntos blancos en estado 1.

más interesantes, toda vez que la formación de clústeres espacio-temporales, con sus bordes más o menos nítidos, dan fe de cómo las interacciones que actúan a nivel local provocan estructuras globales.

En la figura 2 se representa una porción del plano paramétrico \mathbb{R}^2 en el que toma valores el vector paramétrico (α, β) ; las etiquetas 0, 1, 2 y A están colocadas sobre puntos en los que se han realizado simulaciones y tienen la siguiente interpretación: 0 si la dinámica converge rápidamente a una configuración homogénea cuasi-determinística en la que prácticamente todas las celdas están en estado cero, como por ejemplo el caso del apartado a) de la figura 1; etiqueta 1, cuando las células convergen al estado uno, como el b) del ejemplo; etiqueta 2, para comportamientos cíclicos de periodo dos, los únicos observados en este modelo, como en el caso c); finalmente, la etiqueta A, para aquellas dinámicas menos predecibles, más aleatorias, de mayor interés para la modelización estadística. Se observa que amplias zonas del plano se asocian a dinámicas cuasi-deterministas (0, 1 y 2), separadas entre sí por la región de dinámica aleatoria (A) con forma de horquilla.

Es bien conocido que bajo determinadas condiciones, los estimadores de máxima verosimilitud son asintóticamente eficientes y normales (Dudewicz, 1976; Rao, 1973), lo que en el caso del ACE se reduce a $(\hat{\alpha}, \hat{\beta})^T \stackrel{a}{\sim} N((\alpha, \beta)^T, I^{-1}(\alpha, \beta))$, siendo $I^{-1}(\alpha, \beta)$ la matriz de información de Fisher, cuyo resultado exacto no será conocido, pero que habitualmente se sustituye (*method of scoring*) por la matriz de información muestral, la cual en el caso que nos ocupa, y al no disponer de muestras independientes del mismo proceso espacio-temporal, toma la forma

$$M^{-1}(\hat{\alpha}, \hat{\beta}) = \begin{pmatrix} -\frac{\delta^2}{\delta\alpha^2} \ln f(\mathbf{x}_{S, N_k}; \hat{\alpha}, \hat{\beta}) & -\frac{\delta^2}{\delta\alpha\delta\beta} \ln f(\mathbf{x}_{S, N_k}; \hat{\alpha}, \hat{\beta}) \\ -\frac{\delta^2}{\delta\alpha\delta\beta} \ln f(\mathbf{x}_{S, N_k}; \hat{\alpha}, \hat{\beta}) & -\frac{\delta^2}{\delta\beta^2} \ln f(\mathbf{x}_{S, N_k}; \hat{\alpha}, \hat{\beta}) \end{pmatrix}. \quad (2)$$

Puesto que la matriz muestral \mathbf{x}_{S, N_k} guarda las observaciones correspondientes a una *única* realización del proceso espacio-temporal, desde la etapa inicial 0 hasta la k -ésima, que

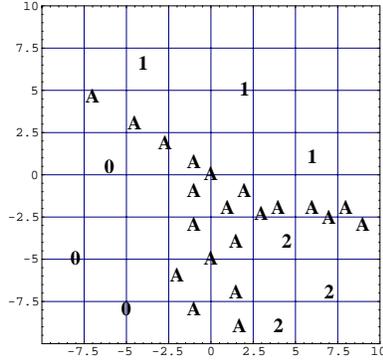


Figura 2: Etiquetas mostrando las dinámicas obtenidas simulando los modelos con diferentes pares paramétricos. Ver el texto para su interpretación.

(α, β)	$\hat{E}[\hat{\alpha}, \hat{\beta}]$	$\hat{V}[\hat{\alpha}]$	$\hat{V}[\hat{\beta}]$	$\hat{Cov}[\hat{\alpha}, \hat{\beta}]$
$(-8.0, -5.0)$	$(-22.10, -12.77)$ $(-7.67, -21.23)$	146.46 1.39	44.98 16143.9	-80.45 -1.38
$(-2.7, 1.8)$	$(-2.71, 1.80)$ $(-2.75, 1.79)$	0.020 0.013	0.006 0.005	-0.010 -0.007

Figura 3: Comparación de los momentos de los estimadores obtenidos por el método de Montecarlo y mediante la matriz de información muestral, correspondientes a los modelos *a)* y *h)* de la figura 1.

marca el final del periodo de observación, el tamaño muestral es la unidad, siendo el caso que sus componentes están sometidas a un cierto esquema de dependencia estocástica. Se plantea hasta qué punto es válida bajo estas circunstancias la aproximación a los momentos de segundo orden de los estimadores mediante la matriz de información muestral. El método de Montecarlo que se ha seguido consiste en realizar múltiples simulaciones del ACE para valores α y β fijados de antemano, obteniendo en cada caso sus estimadores $\hat{\alpha}$ y $\hat{\beta}$ por el método de Newton-Raphson; con todos los estimadores calculados se procede a evaluar su media y matriz de covarianzas muestrales, disponiendo así de sendos estimadores para $E[\hat{\alpha}]$, $E[\hat{\beta}]$, $V[\hat{\alpha}]$, $V[\hat{\beta}]$ y $Cov[\hat{\alpha}, \hat{\beta}]$, que luego se compararán con los obtenidos invirtiendo la matriz de información muestral (2). Adicionalmente, la normalidad de los estimadores Montecarlo se contrastó con el método gráfico *gamma-plot* (Johnson et al., 1998).

Las estimaciones se han realizado con diferentes números de etapas. A título de ejemplo, la figura 3 tabula los resultados obtenidos para los pares paramétricos correspondientes a los casos *a)* y *h)* de la figura 1.

Para cada par paramétrico se tabulan, en la fila superior, la media, varianzas y covarianza obtenida de los 30 estimadores obtenidos por el método de Montecarlo con simulaciones de $k = 45$ etapas, $|S| = 50$ celdas y radio de vecindad $r = 1$; en la fila inferior, se muestran los resultados correspondientes a una sola estimación, tomada al azar de las 30 disponibles, primero los estimadores $\hat{\alpha}$ y $\hat{\beta}$ y luego las varianzas y covarianza estimadas mediante la inversión de la matriz de información muestral. Se observa que en el caso $(\alpha, \beta) = (-8, 0; -5, 0)$, una típica situación de dinámica cuasi-determinista, los resultados difieren para ambos métodos de evaluación, lo que no sorprende habida cuenta de que amplias regiones del plano

paramétrico imponen comportamientos espacio-temporales prácticamente indistinguibles, lo cual permanece válido para las zonas que se vislumbran con las etiquetas 0, 1 y 2 de la figura 2. En el caso $(\alpha, \beta) = (-2, 7; 1, 8)$, donde el comportamiento azaroso es más evidente, se observa que los momentos muestrales de ambos métodos son más cercanos, y de hecho se aproximan tanto más cuanto mayor sea el número de etapas simuladas; esto mismo se observa para otros pares paramétricos pertenecientes a la zona de la figura 2 marcada con etiquetas A.

4. CONCLUSIONES

Las simulaciones sugieren que la distribución conjunta de los estimadores $\hat{\alpha}$ y $\hat{\beta}$ del ACE logístico se puede considerar, a efectos de contrastes, aproximadamente normal de media (α, β) y matriz de covarianzas $M^{-1}(\hat{\alpha}, \hat{\beta})$, siempre que el par paramétrico del modelo que ha generado los datos se encuentre en una determinada zona del plano, que aquí se ha llamado aleatoria y a la que pertenecen los puntos con la etiqueta A de la figura 2. Fuera de esta zona, y debido a que amplias regiones de \mathbb{R}^2 producen el mismo comportamiento espacio-temporal caracterizado por una dinámica prácticamente determinista y predecible, la conclusión anterior ya no se puede mantener; por otro lado, estos comportamientos cuasi-deterministas carecen de interés desde el punto de vista del análisis estadístico. La riqueza de comportamientos que el ACE aporta tan sólo variando sus valores paramétricos invita a considerarlos como procesos estocásticos para el ajuste de datos espacio-temporales, bien ampliando sus reglas de transición local a otras distribuciones, bien aumentando la dimensionalidad del espacio S , con vistas a su utilización con datos geográficos.

5. REFERENCIAS

- Couclelis, H. (1985). "Cellular worlds: a framework for modeling micro-macro dynamics". *Environment and Planning* 17, 585-596.
- Deadman, P., Brown, R., Gimblett, H. (1993). "Modeling rural residential settlement patterns with cellular automata". *Journal of Environmental Management* 37, 147.
- Dudewicz, E.I. (1976). *Introduction to Probability and Statistics*. Holt, Rinehart y Wilson.
- Gerhardt, M., Schuster, H. (1989). "A cellular automaton describing the formation of spatially ordered structures in chemical systems". *Physica D* 36, 209-221.
- Hogeweg, P. (1988). "Cellular automata as a paradigm for ecological modeling". *Applied Mathematics and Computation* 27, 81-100.
- Johnson, R.A., Wichern, D.W. (1998). *Applied Multivariate Statistical Analysis*. Prentice Hall.
- Luthi, P., Chopard, B., Preiss, A., Ramsden, J. (1998). "A cellular automaton model for neurogenesis in *Drosophila*". *Physica D* 118, 151-160.
- Rao, C.R. (1973). *Linear Statistical Inference and Its Applications*. J. Wiley.
- Vichniac, G. (1984). "Simulating physics with cellular automata". *Physica D* 10, 96-116.
- von Neumann, J. (1966). *Theory of Self-Reproducing Automata*. University of Illinois Press.
- Wolfram, S. (1984). "Computational Theory of Cellular Automata". *Communications in Mathematical Physics* 96, 15-57.
- Wolfram, S. (1986) "Cellular automaton fluids: basic theory". *Journal of Statistical Physics* 45, 471-526.
- Wolfram, S. (1996). *The Mathematica Book*. Cambridge University Press.