1
Cadenas de Markov

Kamal A. Romero S.

Sumario

Se repasan las principales propiedades de los proceso de Markov, la cual representa la principal herramienta que utilizaremos en nuestros modelos para representar la incertidumbre en relación a la ocurrencia de una variable.

Contenido

1. Introducción

La base de los modelos de agentes heterogéneos es la existencia de un choque idiosincrático, lo cual combinado con la ausencia de un mercado perfecto de seguros genera distribuciones de agentes.

Dado el rol clave de la realización de los choques, necesitamos una manera adecuada de representarlo. No obstante, no podemos escoger cualquier especificación para el proceso estocástico de la variable que genera incertidumbre en nuestra economía.

La primera restricción es que el proceso estocástico debe adaptarse a las herramientas de análisis recursivo que utilizamos para analizar las economías utilizadas en estas notas. Otra característica deseable es que el proceso estocástico se adapte de manera flexible a la aproximación discreta que realizamos para resolver numéricamente nuestras economías.

En este capítulo presentamos el proceso mayormente utilizado en el análisis recursivo, las cadenas de Markov.

Nuestro objetivo es describir de manera resumida las principales características de los proceso de Markov. Sólo mencionaremos lo mínimo necesario para que el lector pueda seguir estas notas[1] .

2. Procesos de Markov

Una característica de los modelos recursivos es que las decisiones de los agentes dependen sólo de una variable de estado que resume toda la información necesaria acerca de la posición en que se encuentra el sistema dinámico, y de una ley de movimiento que describe la dinámica de dicho sistema.

Para introducir choques estocásticos en este tipo de modelos desearíamos tener una especificación para los choques que tuvieran una estructura similar.

Existe una familia de procesos estocásticos en los cuales la probabilidad de ocurrencia de un evento específico depende de la historia pasada de realizaciones de la variable aleatoria, pero donde toda la historia se encuentra resumida en una variable de estado la cual cambia en el tiempo según unas probabilidades dadas.

De manera poco técnica definimos un *proceso de Markov* como un proceso estocástico donde la esperanza condicionada de la realización de una variable estocástica x viene dada por:

$$E(x_{t+1}|x_t, x_{t-1}, \ldots, x_0) = E(x_{t+1}|x_t) \tag{1}$$

Un proceso de Markov definido en un espacio de los estados $\mathcal{S}$ finito es llamado una *cadena de Markov*, donde $\mathcal{S}$ representa el conjunto de todos los posibles valores que puede adoptar x.

[1] Para un análisis introductorio del tema ver Bertsekas and Tsitsiklis (2002, Cap. 6), para un análisis más profundo en el marco de los métodos recursivos ver Stokey and Lucas (1989, Sección 3)

Si nos ceñimos a realizaciones discretas de la variable x nos referimos a cadenas de Markov en tiempo discreto.

Las cadenas de Markov se encuentran descritas en términos de probabilidades de transición $p_{i,j}$, las cuales nos indican la probabilidad de ocurrencia del estado j dado que el estado actual es i. Según la expresión 1, $p_{i,j}$ viene dado por:

$$p_{ij} = P(x_{t+1} = j | x_t = i) \tag{2}$$

Para especificar una cadena de Markov necesitamos tres elementos:

1. El espacio de los posibles estados $\mathcal{S} = \{1, \ldots, n\}$. $\mathcal{S}$ es un vector de tamaño n que nos indica los posibles valores que puede tomar la variable x

2. Las probabilidades de transición $p_{i,j}$ para cada par $(i, j) \in \mathcal{S}$. Estas probabilidades $p_{i,j}$ estarán representadas en una matriz de transición P de tamaño $(n \times n)$

3. La distribución inicial de estados π_0, que nos indica la probabilidad de ocurrencia del estado i en $t = 0$.

La matriz de transición P será de la siguiente forma:

$$\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \tag{3}$$

Dado que p_{ij} son probabilidades, se debe cumplir $\forall i, j \in \mathcal{S}$ que $\sum_{j=1}^{n} p_{ij} = 1$

Ejemplo 2.1. Supongamos una variable aleatoria x, donde x es el estado de empleo de un trabajador. Existen 2 posibles realizaciones de la variable x, empleado (i) o desempleado (j).

La probabilidad de mantenerse empleado p_{ii} es de 0,8, la de quedar desempleado estando empleado p_{ij} es de 0,2, la de quedarse desempleado p_{jj} es de 0,5 y la de encontrar un empleo estando desempleado p_{ji} de 0,5.

La matriz de transición P de esta cadena de markov viene dada por:

$$\begin{pmatrix} 0,8 & 0,2 \\ 0,5 & 0,5 \end{pmatrix}$$

La matriz de transición nos resume toda la información necesaria para saber las probabilidades de transitar aun determinado estado el próximo período dado el estado actual

Asimismo podemos conocer la probabilidad de transitar a un determinado estado dentro de 2 períodos a través de P^2, por ejemplo:

$$
\begin{aligned}
& P(x_{t+2} = j | x_t = i) \\
& = \sum_{h=1}^{n} P(x_{t+2} = j | x_{t+1} = h) P(x_{t+1} = h | x_t = i) \\
& = \sum_{h=1}^{n} P_{ih} P_{hj} = P_{ij}^2
\end{aligned}
$$

donde P_{ij}^2 es el elemento i, j de P^2

Repitiendo el procedimiento anterior par cualquier n es posible obtener que

$$P(x_{t+k} = j | x_t = i) = P_{ij}^k$$

A través de la propiedad anterior es posible generar una secuencia de P

Ejemplo 2.2. Dada la matriz del ejemplo anterior podemos calcular las probabilidades de transición para $t+2, t+3$ y $t+4$

$$
\begin{pmatrix} 0,74 & 0,26 \\ 0,65 & 0,35 \end{pmatrix}
\begin{pmatrix} 0,7166 & 0,2834 \\ 0,7085 & 0,2915 \end{pmatrix}
\begin{pmatrix} 0,7143 & 0,2857 \\ 0,7142 & 0,2858 \end{pmatrix}
$$

Dada la distribución inicial π_0, es posible generar una secuencia de distribuciones π_n de la siguiente forma:

$$
\begin{aligned}
\pi_1 &= \pi_0 P \\
\pi_2 &= \pi_0 P^2 \\
&\vdots \\
\pi_n &= \pi_0 P^n
\end{aligned}
$$

3. Clasificación de los Estados

Podemos clasificar los estados de una cadena de Markov según la frecuencia con que son “visitados”. Pueden existir estados a los cuales una vez se llegue no es posible transitar a otro (por ejemplo $p_{ii} = 1$ y $p_{ij} = 0$ $\forall j$) o estados en los cuales una vez se abandonado no es posible volver, etc.

Esto es de utilidad dado que estamos interesados en las características dinámicas de un modelo estocástico cuya dinámica está gobernada por un proceso de Markov.

Definicion 3.1. Un estado i es **accesible** desde un estado j, si para algún n p_{ji}^n es positivo

Denotemos por $A(j)$ el conjunto de estados que son accesibles desde j, decimos que un estado j es **recurrente** si para cada $i \in A(j)$ tenemos que $j \in A(i)$.

Un estado j es recurrente si para cualquier estado al cual se puede acceder desde j existe una probabilidad positiva de volver a dicho estado.

Un estado es j **trasiente** si no es recurrente, existen $i \in A(j)$ tales que j no es accesible desde i.

Existe una probabilidad positiva de una vez abandonado un estado trasiente no retornar al mismo.

Un conjunto $E \subseteq \mathcal{S}$ es **ergódico** si $P(E|s_i \in E) = 1$

Según lo anterior si en un subconjunto de $E \subseteq \mathcal{S}$ existe algún estado trasiente j donde hay algún $i \in A(j) \nsubseteq E$, dicho subconjunto no puede ser ergódico.

La existencia o no de conjuntos ergódicos es importante para el análisis de convergencia de las cadenas de Markov.

4. Distribuciones Estacionarias

En la mayoría de los modelos económicos dinámicos estamos interesados en el comportamiento de largo plazo de la economía, lo que solemos referirnos como el estado estacionario que se encuentra asociado al punto fijo de un sistema dinámico.

Siguiendo los ejemplos anteriores, nos gustaría saber en el largo plazo cual es la probabilidad de estar o no desempleado. Esta información nos la brinda la *distribución estacionaria*.

Definicion 4.1. Un proceso estocástico es **estacionario** (fuerte) si la distribución conjunta de $\{x_t, x_{t+1}, \ldots\}$ es la misma que $\{x_{t+k}, x_{t+k+1}, \ldots\}$ para cualquier k

Una cadena de Markov es estacionaria si $\pi_t = \pi$ para todo t. A la distribución π se le denomina distribución estacionaria.

Ejemplo 4.2. Supongamos que la distribución inicial entre empleados y no empleados viene dada por:

$$A = [0,5 \;\; 0,5]$$

La distribución de las unidades de eficiencia el próximo período vendrá dada por:

$$[\underbrace{(0,5 \times 0,8 + 0,5 \times 0,5)}_{s_1} \;\; \underbrace{(0,5 \times 0,5 + 0,5 \times 0,5)}_{s_2}]$$

Si deseamos saber la distribución de las unidades de eficiencia los próximos dos períodos, realizamos el producto $(A \times P) \times P = A \times (P \times P) = A \times P^2$.

Si continuamos por inducción, la distribución A los próximos n períodos es $A \times P^n$.

Si la secuencia P^n converge a algún punto, las filas de la matriz nos dará la distribución invariante P^*.

Ya habíamos observado en el ejemplo anterior que luego de tres pasos la matriz de transición viene dada por:

$$\begin{pmatrix} 0,7143 & 0,2857 \\ 0,7142 & 0,2858 \end{pmatrix}$$

Si continuamos el proceso anterior la matriz converge en aproximadamente 6 pasos a:

$$\begin{pmatrix} 0,7143 & 0,2857 \\ 0,7143 & 0,2857 \end{pmatrix}$$

Y la distribución invariante de empleo es

$$\begin{pmatrix} 0,7143 & 0,2857 \end{pmatrix}$$

Podemos de una manera muy sencilla obtener la distribución estacionaria en MATLAB.

4.1. Implementación Numérica

1. Definimos un nivel de tolerancia y el valor inicial de la matriz P^o y de la métrica.
2. Realizamos los productos $P' = P \times P$ y $A \times P$
3. Evaluamos la diferencia entre P^o y P' y verificamos si es inferior al criterio de tolerancia $|D^o - D'| \leq tol$
4. De cumplirse el paso 3 hemos acabado, en caso contrario hacemos $D^o = D'$ y volvemos al paso 1.

Dado lo anterior los pasos computacionales son:

- Paso 1: se escribe la matriz de transición, se postula una distribución inicial , el criterio de tolerancia y el valor inicial de la métrica.

  `p=Q;`

  `g=[0.5,0.5];`

  `distance=1;`

  `toll=`$1 * 10^(- 8)$`);`

- Paso 2: realizamos las multiplicaciones

  `while distance>toll`

  `p1=p*p;`

  `g1=g*p;`

- Paso 3: evaluamos la distancia entre la matriz de transición inicial y la obtenida en la multiplicación

  ```
  distance=norm(abs(p1-p));
  ```

- Paso 4: de no cumplirse el condicional `while`

  ```
  p=p1;
  g=g1;
  ```

- al cumplirse el condicional `while`

  ```
  end
  ```

Ejercicio 4.3. Encontrar las distribuciones estacionarias asociadas a las siguientes matrices de transición:

1. $P = \begin{pmatrix} 0,75 & 0,25 \\ 0,25 & 0,75 \end{pmatrix}$

2. $P = \begin{pmatrix} 0,8 & 0,2 \\ 0,6 & 0,4 \end{pmatrix}$

3. $P = \begin{pmatrix} 0,7 & 0,3 & 0 \\ 0 & 0,5 & 0,5 \\ 0 & 0,9 & 0,1 \end{pmatrix}$

Referencias

BERTSEKAS, D. P., AND J.Ñ. TSITSIKLIS (2002): *Introduction to Probability*. Athena Scientific.

STOKEY, N. L., AND R. E. LUCAS (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press.