Skattning av Markovprocesser

Jan Enger
Matematisk statistik
KTH

Ht 2001

Vi skall här undersöka hur man kan skatta okända övergångssannolikheter i en Markovkedja i diskret tid eller intensiteterna i en Markovprocess i kontinuerlig tid. Kedjan eller processen antas därför ha studerats under en viss tid varvid man observerat processens tillstånd och, om en process i kontinuerlig tid betraktats, tiderna den legat i de olika tillstånden. Vi skall ge intuitivt naturliga skattningar och visa att dessa är identiska med maximum-liklelihoodskattningarna.

1 Diskret tid

En Markovkedja med tillståndsrum, $\{1, 2, \ldots, N\}$, startas i tidpunkt 1 i tillstånd x_1 och hoppar vid tidpunkterna $2, 3, \ldots, n$ till tillstånden $x_2, x_3, \ldots, x_n$. Vi vill nu skatta den okända övergångsmatrisen $P = (p_{ij})_{N\times N}$.

Den intuitiva skattningen av övergångssannolikheten p_{ij} är den relativa frekvensen av hopp från tillstånd i till tillstånd j, d.v.s.

$$p^*_{ij} = \frac{n_{ij}}{n_i} \tag{1}$$

där n_{ij} är antalet observerade hopp från i till j och n_i är antalet gånger kedjan befunnit sig i tillstånd i fram till tid $n-1$ (observera att vi från tillståndet vid tid n inte vet vart hoppet gick). Om $n_i = 0$ kan vi inte skatta n_{ij}:na.

Man skulle kunna använda maximum-likelihoodmetoden för att skatta övergångssannolikheterna. Likelihoodfunktionen är sannolikheten få de observationer man faktiskt erhållit och därför gäller enligt markovegenskapen att

$$L(x_1, x_2, \ldots, x_n; P) = p_{x_1x_2}p_{x_2x_3}\cdots p_{x_{n-1}x_n}. \tag{2}$$

Genom eftertanke inses att

$$p_{x_1x_2}p_{x_2x_3}\cdots p_{x_{n-1}x_n} = \prod_{i=1}^{N}\prod_{j=1}^{N} p_{ij}^{n_{ij}} \tag{3}$$

där n_{ij} är antalet hopp från i till j. Denna funktion skall maximeras under bivillkoren att radsummorna i övergångsmatrisen är 1, d.v.s.

$$p_{i1} + p_{i2} + \cdots + p_{iN} = 1, \quad i = 1, 2, \ldots, N \tag{4}$$

Vi maximerar logaritmen för L dessa bivillkor med hjärlp av Lagrange multiplikatormetod. Betrakta alltså

$$\Lambda = \ln(L) - \lambda_1(p_{11} + p_{12} + \cdots + p_{1N}) - \lambda_2(p_{21} + p_{22} + \cdots + p_{2N}) - \cdots - \lambda_N(p_{11} + p_{N2} + \cdots + p_{NN}) =$$
$$\sum_{i=1}^{N}\sum_{j=1}^{N} n_{ij}\ln(p_{ij}) - \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i p_{ij} = \sum_{i=1}^{N}\sum_{j=1}^{N}(n_{ij}\ln(p_{ij}) - \lambda_i p_{ij}) \tag{5}$$

Deriverar vi med avseende på p_{ij} erhåller vi att

$$\frac{\partial \Lambda}{\partial p_{ij}} = \frac{n_{ij}}{p_{ij}} - \lambda_i = 0$$

vilket ger

$$\frac{p_{i1}}{n_{i1}} = \frac{p_{i2}}{n_{i2}} = \cdots = \frac{p_{iN}}{n_{iN}} = \frac{1}{\lambda_i}.$$

Det gäller alltså att

$$p_{ij} = \frac{n_{ij}}{\lambda_i}$$

Utnyttjas nu bivillkoret (4) fås

$$1 = \sum_{j=1}^{N} p_{ij} = \sum_{j=1}^{N} \frac{n_{ij}}{\lambda_i} = \frac{n_i}{\lambda_i} \tag{6}$$

där n_i är antalet gånger kedjan varit i tillstånd i fram till tid $n-1$. Vi erhåller alltså

$$\frac{1}{\lambda_i} = \frac{1}{n_i}, \text{ d.v.s. man får att } p^*_{ij} = \frac{n_{ij}}{n_i} \tag{7}$$

Som man ser erhålls samma skattningar som de intuitiva (1).

Vi har förutsatt att övergångssannolikheterna inte är relaterade till varandra, förutom att de radvis summerar sig till 1. Maximum-likelihood-metoden kan dock också användas då övergångssannolikheterna beror på en parameter, vilket vi visar genom ett exempel.

Exempel 1 Antag att en Markovkedja startar i tillstånd $x_1 = 1$ och har övergångsmatrisen

$$P = \begin{pmatrix} p & 1-p \\ 2p & 1-2p \end{pmatrix}$$

och att man observerat följden $1, 2, 2, 2, 1, 2, 1, 2, 2, 2$. Sannolikheten för denna följd är

$$L(p) = (1-p)(1-2p)(1-2p)2p(1-p)2p(1-p)(1-2p)1-2p) = \\ 4p^2(1-p)^3(1-2p)^4.$$

Maximum-likelihoodskattningen av p är det värde som maximerar denna sannolikhet. Derivera logaritmen och vi erhåller

$$\frac{d\ln(L)}{dp} = \frac{2}{p} - \frac{3}{1-p} - \frac{8}{1-2p} = \cdots = \frac{18p^2-20p+2}{p(1-p)(1-2p)}$$

som har nollställen $p = \frac{1}{9}$ och $p = 1$. Det sista värdet är otillåtet (alla övergångssannolikheter måste vara icke-negativa) men det första ger ett maximum. Maximum-likelihoodskattningen av övergångsmatrisen är alltså

$$\begin{pmatrix} 1/9 & 8/9 \\ 2/9 & 7/9 \end{pmatrix}$$

$\square$

2 Kontinuerlig tid

Antag nu att vi har en Markovprocess i kontinuerlig tid med intensitetsmatris

$$Q = \begin{pmatrix} -q_1 & q_{12} & \cdots & q_{1N} \\ q_{21} & -q_2 & \cdots & q_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ q_{N1} & q_{N2} & \cdots & -q_N \end{pmatrix}$$

där som vanligt $q_i = \sum_{j \neq i} q_{ij}$. Antag att processen startar i tillstånd x_1 och att vi observerar hoppen och tiderna (x_1, t_1), $(x_2, t_1), \ldots, (x_{n-1}, t_{n-1})$ och slutligen ett hopp till tillstånd x_n. Hur skall vi från dessa data skatta intensitetsmatrisen?

En intuitiv skattning är följande. Hoppen är oberoende av tiderna i de olika tillstånden. Hoppmatrisen är

$$P = \begin{pmatrix} 0 & q_{12}/q_1 & \cdots & q_{1N}/q_! \\ q_{21}/q_2 & 0 & \cdots & q_{2N}q_2 \\ \vdots & \vdots & \ddots & \vdots \\ q_{N1}/q_N & q_{N2}/q_N & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & p_{11} & \ldots & p_{1N} \\ p_{21} & 0 & \ldots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1} & p_{N2} & \ldots & 0 \end{pmatrix}$$

där $p_{ij} = \dfrac{q_{ij}}{q_i}$. Dessa hoppsannolikheter, övergångssannolikheter, kan man skatta som i avsnitt 1, d.v.s. med relativa frekvenser för hoppen. Diagonalelementen i Q kan man skatta genom att betrakta tiderna i de olika hoppen.

Väntevärdet för den tid som Markovprocessen ligger i tillstånd i innan den hoppar är $1/q_i$ och detta väntevärde skattas lämpligen med medelvärdet av de tider man varit i tillståndet, $1/q_i^* = \bar{t}_i$, där $\bar{t}_i$ medeltiden i tillstånd i. Det ger $q_i^* = 1/\bar{t}_i, i = 1, 2, \ldots, N$ och om de skattade hoppsannolikheterna är p_{ij}^* erhålls skattningen av övergångsintensiteterna $q_{ij} = q_i \dfrac{q_{ij}}{q_i} = q_i p_{ij}$ som

$$q_{ij}^* = p_{ij}^* q_i^* = p_{ij}^*/\bar{t}_i = \frac{n_{ij}}{n_i \bar{t}_i} = \frac{n_{ij}}{T_i} \tag{8}$$

där T_i är total tid som tillbringats i tillstånd i, n_{ij} antal hopp från tillstånd i till j och n_i antalet gånger processen varit i tillstånd i (sista tillståndet x_n oräknat). Skattningen av intensiteten att gå från tillstånd i till tillstånd j är alltså lika med antalet hopp från i till j per tidsenhet.

Dessa skattningar är även de maximum-likelihoodskattningarna. Likelihoodfunktionen är i detta fall en blandning av en diskret del och en kontinuerlig del. Vi bildar produkten av sannolikheten för de observerade hoppen samt täthetsfunktionen för de observerade tiderna i tillstånden;

$$\begin{aligned} L(\mathbf{x}, \mathbf{t}; Q) &= \prod_{k=1}^{n-1} p_{x_k x_{k+1}} \prod_{k=1}^{n-1} (e^{-t_k q_{x_k}} q_{x_k}) = \prod_{k=1}^{n-1} \Big(\frac{q_{x_k x_{k+1}}}{q_{x_k}}\Big) \prod_{k=1}^{n-1} (e^{-t_k q_{x_k}} q_{x_k}) \\ &= \prod_{k=1}^{n-1} q_{x_k x_{k+1}} e^{-t_k q_{x_k}} = (\text{tänk efter}) = \prod_{i=1}^{N} e^{-T_i q_i} \prod_{j=1}^{N} q_{ij}^{n_{ij}} \end{aligned} \tag{9}$$

där $\mathbf{x}$ och $\mathbf{t}$ är vektorn av de observerade tillstånden respektive tiderna i tillstånden och n_{ij} är antalet hopp från i till j och T_i total tid i tillstånd i. Logaritmen av L är

$$\ln(L) = \sum_{i=1}^{N} \Big(\sum_{j=1}^{N} n_{ij} \ln(q_{ij}) - T_i \sum_{j=1}^{N} q_{ij} \Big) \tag{10}$$

där vi satt in $q_i = \sum_{j=1}^{N} q_{ij}$. Denna funktion skall maximeras. Deriverar vi med avseende på q_{ij} erhålls

$$\frac{\partial \ln(L)}{\partial q_{ij}} = \frac{n_{ij}}{q_{ij}} - T_i \tag{11}$$

och den partiella derivatan lika med 0 ger omedelbart

$$q_{ij} = q_{ij}^* = \frac{n_{ij}}{T_i} \tag{12}$$

som är antalet hopp till j per tidsenhet i tillstånd i. Vi får alltså samma skattningar som de intuitiva.

Exempel 2 En Markovprocess med tillståndsrum $\{1, 2, 3\}$ startade i tillstånd 1 och hoppade vid tidpunkterna 2.6, 11.8, 15.8, 17.1, 21.1, 21.9, 22, 23.8, 25.4

och 26.4 till tillstånden 3, 1, 2, 1, 2, 1, 3, 2, 1, 3. Tiderna i de successiva tillstånden blir 2.6, $11.8 - 2.6 = 9.2$, $15.8 - 11.8 = 4.0$ o.s.v. I tabellform:

Tillstånd	1	3	1	2	1	2	1	3	2	1	3
Tid i tillstånd	2.6	9.2	4.0	1.3	4.0	0.8	0.1	1.8	1.6	1.0	–

Total tid i tillstånd 1 blir $T_1 = 2.6 + 4.0 + 4.0 + 0.1 + 1.0 = 11.7$. På liknande sätt får tiderna i tillstånd 2 och 3, $T_2 = 3.7$ och $T_3 = 11.0$. Antalet hopp från tillstånd 1 till 2 är två, $n_{12} = 2$. Man ser också att $n_{13} = 3$, $n_{21} = 3$, $n_{31} = n_{32} = 1$. Det ger oss skattningarna

$$q^*_{12} = \frac{n_{12}}{T_1} = \frac{2}{11.7} = 0.171, \quad q^*_{13} = \frac{3}{11.7} = 0.256$$

$$q^*_{21} = \frac{3}{3.7} = 0.811, \qquad q^*_{23} = \frac{0}{3.7} = 0, \qquad q^*_{31} = q^*_{32} = \frac{1}{11.0} = 0.091$$

Den skattade intensitetsmatrisen blir alltså

$$Q^* = \begin{pmatrix} -0.427 & 0.171 & 0.256 \\ 0.811 & -0.811 & 0 \\ 0.091 & 0.091 & -0.182 \end{pmatrix}$$

□