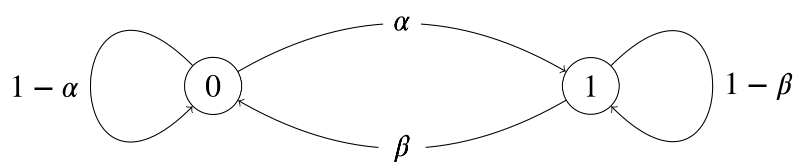
Cvičení - procesy

Štěpán Holub, Michal Kupsa

Uvažujme náhodný proces $X = (X_i)_{i \in \mathbb{N}}$ na dvouprvkové abecedě $A = \{0, 1\}$ daný tímto přechodovým grafem:



1. Napište přechodovou matici procesu.
2. Ověřte, že
$$\left(\frac{\beta}{\alpha + \beta}, \frac{\alpha}{\alpha + \beta} \right)$$
je stacionární rozdělení procesu.
3. Pro jaká α a β je proces i.i.d?
4. Určete entropii příslušného procesu při stacionárním rozdělení.
5. Dokažte, že $\left(X_{2i}\right)_{i \in \mathbb{N}}$ je také markovský proces, napište jeho přechodovou matici a určete jeho stacionární rozdělení.
6. Určete vzájemnou informaci $\mathcal{I}\left(X_i : X_{i+1}\right)$ a $\mathcal{I}\left(X_i : X_{i+2}\right)$ při stacionárním rozdělení.

Řešení

Poznamenejme nejprve, že zadání pokrývá zcela obecný případ náhodného markovského procesu na dvouprvkové abecedě. Tomuto případu se také věnuje CT, Example 4.1.1. na str. 73n.

1. Přechodová matice je matice $M_{i,j} = \mathbb{P}(X_{n+1} = j \mid X_n = i)$, tedy

$$M = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}.$$

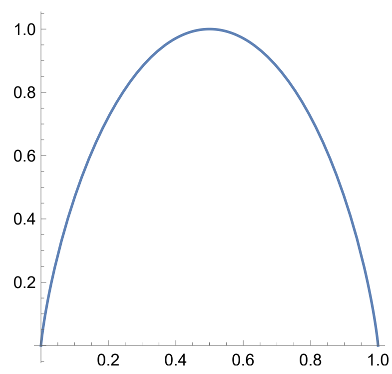
2. Stacionární distribuce je podle definice X_0 zaručující $X_0 = X_i$, $i \in \mathbb{N}$. Tomu odpovídá, že vektor $\mu = (\mathbb{P}(X_0 = 0), \mathbb{P}(X_0 = 1))$ splňuje $\mu M = \mu$. Ověřte rozepsáním, že tento vztah opravdu zaručuje $X_{n+1} = X_n$. Stačí tedy tento vztah ověřit pro navržený vektor.
3. Proces je i.i.d, pokud platí $\mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(X_{n+1} = j)$ pro všechna i, j. Pro přechodovou matici to znamená, že její řádky musí být stejné. Při pohledu na graf to znamená, že obě vstupní šipky musí mít stejnou hodnotu. Z každého z těchto pohledů lze získat podmínku $1 - \alpha = \beta$ neboli $\alpha + \beta = 1$. Všimněte si, že v tomto případě má X_i, $i \geq 1$, rozdělení $(\beta, 1-\beta) = (1-\alpha, \alpha)$ (které je zároveň stacionárním rozdělením tohoto procesu) bez ohledu na počáteční rozdělení X_0.
4. Entropie procesu je definována jako limita průměrné entropie počátečních vektorů $X_{[0..n)}$. Pro stacionární případ není třeba nic průměrovat, entropie procesu je rovna přírůstku entropie v jednom kroku, tedy podmíněné entropii $\mathcal{H}(X_{i+1} \mid X_i)$, neboli $\mathcal{H}(X_1 \mid X_0)$. Tuto hodnotu je zde vhodné vyjádřit jako střední hodnotu jednotlivých entropií:

$$\mathcal{H}(X_1 \mid X_0) = \frac{\beta}{\alpha+\beta} \cdot \mathcal{H}(X_1 \mid X_0 = 0) + \frac{\alpha}{\alpha+\beta} \cdot \mathcal{H}(X_1 \mid X_0 = 1),$$

přičemž $\mathcal{H}(X_1 \mid X_0 = 0) = h(\alpha)$ a $\mathcal{H}(X_1 \mid X_0 = 1) = h(\beta)$, kde

$$h(x) = \mathcal{H}(x, 1-x) = -x \cdot \log x - (1-x) \cdot \log(1-x)$$

je známá funkce entropie binární náhodné veličiny s rozdělením $(x, 1-x)$.



Máme tedy

$$\mathcal{H}(X) = \frac{\beta}{\alpha+\beta} \cdot h(\alpha) + \frac{\alpha}{\alpha+\beta} \cdot h(\beta).$$

5. Rozdělení $P_{X_{2i+2}}$ je $P_{X_{2i+1}} \cdot M$ a rozdělení $P_{X_{2i+1}} = P_{X_{2i}} \cdot M$. Tedy $P_{X_{2i+2}} = P_{X_{2i+2}} \cdot M^2$. Proces $\left(X_{2i}\right)_{i\in\mathbb{N}}$ je tedy opět markovský a jeho přechodová matice je druhá

mocnina M. Pro zápis M^2 je vhodné si uvědomit, co jsme řekli na začátku. Jedná se opět o markovský proces nad binární abecedou, jen s jinými parametry. Po úpravě dostáváme

$$M^2 = \begin{pmatrix} 1 - c \cdot \alpha & c \cdot \alpha \\ c \cdot \beta & 1 - c \cdot \beta \end{pmatrix},$$

kde $c = 2 - (\alpha + \beta)$. Není překvapením, že pro případ i.i.d., tedy pro $\alpha + \beta = 1$, dostáváme $M^2 = M$.

Stacionární rozdělení X' je stejné jako pro X. To lze zjistit několika způsoby. Jednak můžeme do již ověřeného vztahu pro stacionární vektor dosadit nové přechodové pravděpodobnosti a zkrátit c. Nebo můžeme (zbytečně pracně) ověřit $\mu M^2 = \mu$, což je ovšem zřejmé ze zápisu $(\mu M)M$. To také odpovídá jednoduché úvaze: pro současné X_0 mají všechna X_i stejné rozdělení, tedy i všechna sudá.
6. Pro výpočet vzájemné informace použijeme součtový vzorec

$$\mathcal{I}\left(X_2 : X_0\right) = \mathcal{H}\left(X_2\right) - \mathcal{H}\left(X_2 \mid X_0\right),$$

přičemž $\mathcal{H}(X_2) = h\left(\frac{\alpha}{\alpha+\beta}\right)$ a pro $\mathcal{H}(X_2 \mid X_0) = \mathcal{H}(X')$ můžeme použít již spočtenou hodnotu upravenou pro koeficienty matice M^2, tedy $\alpha' = c \cdot \alpha$ a $\beta' = c \cdot \beta$. Můžeme tedy rovnou napsat hodnoty pro oba procesy

$$\begin{aligned} \mathcal{I}\left(X_1 : X_0\right) &= h\left(\frac{\alpha}{\alpha+\beta}\right) - \frac{\beta}{\alpha+\beta}h(\alpha) - \frac{\alpha}{\alpha+\beta}h(\beta), \\ \mathcal{I}\left(X_2 : X_0\right) &= h\left(\frac{\alpha'}{\alpha'+\beta'}\right) - \frac{\beta'}{\alpha'+\beta'}h\left(\alpha'\right) - \frac{\alpha'}{\alpha'+\beta'}h\left(\beta'\right), \end{aligned}$$

kde ve druhém vzorci můžeme využít rovností

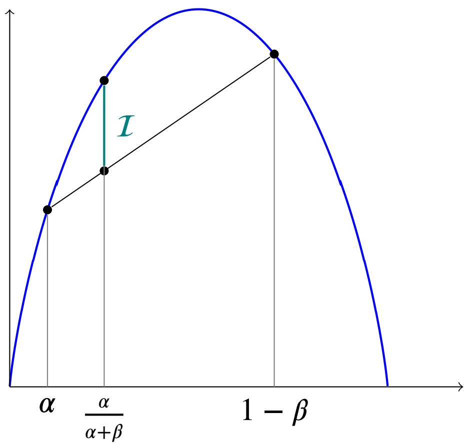
$$\frac{\alpha}{\alpha+\beta} = \frac{\alpha'}{\alpha'+\beta'}, \qquad \frac{\beta}{\alpha+\beta} = \frac{\beta'}{\alpha'+\beta'},$$

které už navíc známe z rovnosti stacionárních rozdělení pro procesy X a X'.

V tomto vzorci je pěkně vidět, že vzájemná informace je kladná díky konkávnosti entropie. S využitím rovnosti $h(\beta) = h(1-\beta)$ můžeme psát

$$\mathcal{I}\left(X_1 : X_0\right) = h\left(\frac{\beta}{\alpha+\beta}\alpha + \frac{\alpha}{\alpha+\beta}(1-\beta)\right) - \frac{\beta}{\alpha+\beta}h(\alpha) - \frac{\alpha}{\alpha+\beta}h(1-\beta)$$

a vidíme, že vzájemná informace je svislá vzdálenost mezi bodem na grafu h v bodě $\frac{\alpha}{\alpha+\beta}$ a tětivou mezi hodnotami v bodech α a $1-\beta$. Vzájemná entropie je tedy nezáporná a je nulová, jen pokud $\alpha = 1 - \beta$, což je již vyšetřený případ nezávislosti.



Předchozí postup můžeme opakovat a vyšetřovat vzájemnou entropii X_0 a X_{2^k}. Je intuitivně zřejmé, že závislost by měla postupně klesat, vzájemná entropie by se tedy měla blížit k nule. To se skutečně děje, protože $\alpha + \beta$ se blíží k jedné. To je vidět z hodnoty c. Je-li totiž $\alpha + \beta = 1 + \varepsilon$, pak $c = 2 - (\alpha + \beta) = 1 - \varepsilon$, a

$$\alpha' + \beta' = c \cdot (\alpha + \beta) = (1 - \varepsilon)(1 + \varepsilon) = 1 - \varepsilon^2.$$

Vzdálené veličiny procesu jsou tedy téměř nezávislé.

Pozn.: Vyšetřovali jsme binární markovské procesy se stacionárním rozdělením. Pro markovské procesy nad libovolnou (konečnou) abecedou, a to i pro ty nestacionární, platí, že entropie (definovaná limitou) existuje. Platí také, za jistých podmínek zkoumaných tzv. ergodickou teorií, že rozdělení procesu ke stacionárnímu konverguje.