Dinamikus rendszerek paramétereinek becslése

Bayes paraméterbecslés

Hangos Katalin

MTA SzTAKI, Folyamatirányítási Kutató Csoport

PE Számítástudomány Alkalmazása Tanszék

ISMÉTLÉS

Prediktív modellek paramétereinek ML becslése 1.

A teljes valószinuségi modell

$$\mathbb{M}(\theta): \qquad \hat{y}(k \mid \theta) = g(k, D^{k-1}; \theta)$$

$$\varepsilon(k,\theta) = y(k) - \hat{y}(k \mid \theta) \text{ függetlenek,}$$

$$\text{adott } \mathbf{f}_\varepsilon(\mathbf{x}; \mathbf{k}; \theta) \text{ suruségfüggvénnyel}$$

A likelihood függvény és a maximum likelihood becslés

A $\varepsilon(k,\theta)$ hibák függetlenek, így az együttes suruségfüggvény:

$$\hat{f}_y(\theta; y^N) = \prod_{i=1}^{N} f_\varepsilon(y(i) - g(i, D^{i-1}; \theta), i; \theta) = \prod_{i=1}^{N} f_\varepsilon(\varepsilon(i,\theta), i; \theta)$$

a log-likelihood függvény pedig

$$\ell(\varepsilon, \theta, k) = -\ln f_\varepsilon(\varepsilon, k; \theta)$$

Prediktív modellek paramétereinek ML becslése 2.

A ML becslés:

$$\hat{\theta}_{ML}(y^N) = \arg\min_{\theta} \frac{1}{N} \sum_{i=1}^{N} -\ell(\varepsilon(i,\theta), i; \theta)$$

A LKN becslés:

$$\hat{\theta}_{ML}(y^N) = \arg\min_{\theta \in D_M} V_N(\theta, y^N) \quad , \quad V_N(\theta, D^N) = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{2} \varepsilon(k,\theta)^2$$

Látható, hogy a ML becslés a LKN becslés abban a speciális esetben, ha a veszteségfüggvényt a

$$V_N(\theta, y^N) = \frac{1}{N} \sum_{i=1}^{N} - \ln f_{\varepsilon}(\varepsilon(i,\theta), i; \theta)$$

alakúra választjuk.

A BAYES BECSLÉS MATEMATIKAI ALAPJAI

A véletlen Bayes féle fogalma

Bayes értelemben *véletlen változónak tekintünk minden változót vagy akár paramétert is, amely nem ismert* előttünk, mint megfigyelők előtt.

Ilyen értelemben az ***ismeretlen*** θ ***rendszerparaméterek valószínűségi változóknak tekintendők***.

A Bayes formula: klasszikus eset

Legyenek adva az alábbi B_i események illetve $P(B_i)$ valószínűségeik:

- $B_1, B_2, \ldots, B_n \in \mathbb{B}$ eseményalgebra, ahol
- $B_1, B_2, \ldots, B_n$ teljes eseményrendszert alkot, azaz $\cup_{i=1}^{N} B_i = \Omega$ és $B_j \cap B_i = \emptyset$
- $P(B_i) > 0 \qquad\qquad i = 1, 2, \ldots, N$

$$P(B_k \mid A) = \frac{P(B_k \cap A)}{P(A)} = \frac{P(A \mid B_k)P(B_k)}{\sum_{i=1}^{N} P(A \mid B_i)P(B_i)}$$

Bayes formula feltételes valószínűségi sűrűségfüggvényekre

Adott

(1) az a valószínűségi változó sűrűségfüggvénye: $p(a) = f_a(x)$,
(2) az a és b együttes sűrűségfüggvénye: $p(a,b) = f_{ab}(x_1, x_2)$,
(3) az a-nak a b-re vonatkozó feltételes s.f.-e: $p(a|b) = f_{(a|b)}(x)$

$$p(b|a) = \frac{p(a,b)}{p(a)} \quad , \quad p(a) = \int_{\Omega_b} p(a,b)db$$

$$p(b|a) = \frac{p(a,b)}{p(a)} = \frac{p(a|b)p(b)}{\int p(a,b)db}$$

$$((P(B_k \mid A) = \frac{P(B_k \cap A)}{P(A)} = \frac{P(A \mid B_k)P(B_k)}{\sum_{i=1}^{N} P(A \mid B_i)P(B_i)}))$$

A Bayes formula több feltételes sűrűségfüggvényre

A feltételes és együttes sűrűségfüggvényekre vonatkozó összefüggés:

$$p(a, b|c) = p(a|b, c)p(b|c) \quad , \quad p(a, b|c) = p(b|a, c)p(a|c)$$

és

$$p(b|c) = \int_{\Omega_a} p(a, b|c)da$$

Ekkor a Bayes formula:

$$p(a|b, c) = \frac{p(a, b|c)}{p(b|c)} = \frac{p(a, b|c)}{\int p(a, b|c)da} = \frac{p(b|a, c)p(a|c)}{\int p(b|a, c)p(a|c)da}$$

A láncszabály

A láncszabály a feltételes és együttes sűrűségfüggvényekre vonatkozó

$$p(a,b) = p(a|b)p(b)$$

összefüggés általánosítása több (N) valószínűségi változóra.

A láncszabályt rekurzív módon vezethetjük le:

$$\begin{aligned}
p(x_N, x_{N-1}, \ldots, x_1) &= p(x_N|x_{N-1}, \ldots, x_1)p(x_{N-1}, x_{N-2}, \ldots, x_1) \\
&= p(x_N|x_{N-1}, \ldots, x_1)p(x_{N-1}|x_{N-2}, \ldots, x_1) \ldots p(x_1) \\
&= \cdots \\
&= \left(\prod_{k=2}^{N} p(x_k|x_{k-1}, \ldots, x_1)\right)p(x_1)
\end{aligned}$$

DINAMIKUS RENDSZEREK PARAMÉTEREINEK BAYES-BECSLÉSE

Dinamikus rendszerek prediktív Bayes modellje – 1

Egyszerűség kedvéért: SISO eset.

LTI rendszerek bemenet-kimenet modelljének Bayes féle alakja az alábbi feltételes sűrűségfüggvény:

$$p(y(k) \mid u(k), D^{k-1})$$

amely alkalmas előrebecslésre és szabályozótervezésre is, de nem szerepelnek benne a θ rendszerparaméterek.

Parametrizált prediktív bemenet-kimenet modell Bayes-féle alakja:

$$p(y(k) \mid u(k), D^{k-1}, \theta)$$

A fenti parametrizált rendszermodellt, azaz a fenti feltételes sűrűségfüggvény alakját a Bayes paraméterbecslésnél adottnak tételezzük fel.

Dinamikus rendszerek prediktív Bayes modellje – 2

A nem parametrizált bemenet-kimenet modell a parametrizált alakból:

$$
\begin{aligned}
p(y(k) \mid u(k), D^{k-1}) &= \int p(y(k), \theta \mid u(k), D^{k-1}) d\theta \\
&= \int p(y(k) \mid \theta, u(k), D^{k-1}) p(\theta \mid u(k), D^{k-1}) d\theta
\end{aligned}
$$

A θ paraméterek Bayes féle becslése a $p(\theta|u(k), D^{k-1})$ feltételes sűrűségfüggvény, amely azonban paraméterbecslés esetén nem függ $u(k)$-tól, mint feltételtől, hiszen az egy általunk beállított determinisztikus változó

$$
p(\theta \mid u(k), D^{k-1}) = p(\theta \mid D^{k-1})
$$

Rekurziv paraméterbecslés a Bayes formulával – 1

Egy rekurzív lépés: A θ paraméterek Bayes becslése a $p(\theta|D^k)$ feltételes sűrűségfüggvény, amelyben szétválasztjuk a jelen k időpillanat és a $k-1$ időpillanatig beérkezett adatokat:

$$p(\theta \mid D^k) = p(\theta \mid y(k), u(k), D^{k-1}) = p(\theta \mid y(k), D^{k-1})$$

A feltételes sűrűségfüggvényekre vonatkozó Bayes formula

$$p(a|b,c) = \frac{p(b|a,c)p(a|c)}{\int p(b|a,c)p(a|c)da}$$

$$b \sim y(k) \quad , \quad a \sim \theta \quad , \quad c \sim D^{k-1}$$

szereposztással alkalmazva következő rekurzív formulát adja:

$$p(\theta \mid D^k) = \frac{p(y(k) \mid D^{k-1}, \theta)p(\theta \mid D^{k-1})}{\int p(y(k) \mid D^{k-1}, \theta)p(\theta \mid D^{k-1})d\theta}$$

Rekurziv paraméterbecslés a Bayes formulával – 2

$$p(\theta \mid D^k) = \frac{p(y(k) \mid D^{k-1}, \theta)p(\theta \mid D^{k-1})}{\int p(y(k) \mid D^{k-1}, \theta)p(\theta \mid D^{k-1})d\theta}$$

A formula jobboldalán szereplő, **fizikai értelemmel is bíró részek**:

- $p(\theta \mid D^{k-1})$ az előző $k-1$ időpillanatbeli becslés,
- $p(y(k) \mid D^{k-1}, \theta)$ az adott parametrizált rendszermodell,
- a nevezőben pedig egy normalizálási tényező.

Rekurziv paraméterbecslés a Bayes formulával – 3

Nem rekurzív Bayes paraméterbecslési formula Az egylépéses

$$p(\theta \mid D^k) = \frac{p(y(k) \mid D^{k-1}, \theta) p(\theta \mid D^{k-1})}{\int p(y(k) \mid D^{k-1}, \theta) p(\theta \mid D^{k-1}) d\theta}$$

Bayes paraméterbecslő formulából a láncszabállyal:

$$p(\theta \mid D^N) = \frac{\left[\prod_{k=1}^{N} p(y(k) \mid D^{k-1}, \theta)\right] p^0(\theta)}{NORM}$$

ahol

$$p^0(\theta) = p(\theta \mid D^0)$$

az úgynevezett *prior vagy kezdeti becslés* és $NORM$ egy normalizáló tényező.

A Bayes paraméterbecslés tulajdonságai – 1

1. A becslési eljárás eredménye a $p(\theta|D^N)$ feltételes valószínűségi sűrűségfüggvény.

   Ez a módszer elméleti ereje és alkalmazhatósági gyengesége egyben.

   Elméletileg a becsült θ paraméterekre vonatkozó teljes statisztika rendelkezésre áll, nemcsak aszimptotikusan, hanem véges esetben is, ehhez azonban egy függvényt kell(ene) minden lépésben kiszámítani.

A Bayes paraméterbecslés tulajdonságai – 2

2. A **becslés maga a Bayes formulából adódóan természetében rekurzív**, végrehajtásához a $p(y(k)|D^{k-1}, \theta)$ parametrizált rendszermodellen kívül a $p^0(\theta)$ **prior vagy kezdeti becslés** is szükséges. A prior becsléssel a paraméterekről rendelkezésünkre álló technológiai, fizikai vagy üzemeltetői tudás építhető be a paraméterbecslésbe elméletileg tiszta, és jól követhető módon.
3. Belátható, hogy **autoregressziós bemenet-kimenet modell és normális eloszlású becslési hiba, valamint normális eloszlású prior becslés mellett a Bayes becslés a standard rekurzív legkisebb négyzetes (LKN) becslésre vezet**, tehát ilyen esetben jól számítható.

A MAXIMUM A POSTERIORI (MAP) BECSLÉS

Maximum a posteriori (MAP) becslés – 1

A nem rekurzív Bayes paraméterbecslő formula ML alakja A

$$p(\theta \mid D^N) = \frac{\left[\prod_{k=1}^{N} p(y(k) \mid D^{k-1}, \theta)\right] p^0(\theta)}{NORM}$$

formula számlálójában a mért $y(k), k = 1, ..., N$ kimenetek $p(y^N|\theta)$ együttes sűrűségfüggvénye áll

$$p(y^N|\theta) = \prod_{k=1}^{N} p(y(k)|D^{k-1}, \theta)$$

hiszen a Bayes becslésnél is a rendszer **prediktív teljes valószínűségi modelljét** használjuk fel.

Maximum a posteriori (MAP) becslés – 2

A maximum a posteriori becslés elve

A maximum a posteriori becslés **a Bayes becslésből származtatható** úgy, hogy a becslés eredményeképpen kapott **teljes** $p(\theta|D^N)$ **valószínűségi sűrűségfüggvény helyett** annak egy pontbecslését, méghozzá a **maximum likelihood elve (legnagyobb valószínűség elve) alapján képezett pontbecslését** használjuk.

Maximum a posteriori (MAP) becslés – 3

A Bayes paraméterbecslés nem rekurzív alakja:

$$p(\theta \mid y^N) = \frac{p(y^N \mid \theta)p^0(\theta)}{p(y^N)} \sim f_y(\theta; y^N)g_\theta(\theta)$$

Ebből a maximum likelihood elven képezett becslés a Maximum A Posteriori (MAP) becslés:

$$\hat{\theta}_{MAP}(y^N) = \arg\max_\theta \left[f_y(\theta; y^N)g_\theta(\theta)\right]$$