PENDUGAAN PARAMETER MODEL *HIDDEN* MARKOV *

BERLIAN SETIAWATY DAN LINDA KRISTINA

Departemen Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Pertanian Bogor
Jl Meranti, Kampus IPB Darmaga, Bogor 16680 Indonesia

ABSTRAK. Pendugaan parameter untuk model *Hidden* Markov Elliott et. al. (1995) dilakukan mengunakan Metode *Maximum Likelihood* dan pendugaan ulang menggunakan metode *Expectation Maximization* yang melibatkan perubahan ukuran. Dari metode tersebut diperoleh algoritma untuk menduga parameter model.

Kata kunci: Rantai Markov, model *Hidden* Markov, perubahan ukuran. metode *Expectation Maximization*.

1. PENDAHULUAN

Tulisan ini merupakan kajian pustaka tentang pendugaan parameter untuk model *Hidden* Markov Elliott, et. al. (1995). Pendugaan parameter menggunakan metode *Maximum Likelihood* dan pendugaan ulangnya menggunakan metode *Expectatition Maximization* (Metode EM) yang melibatkan perubahan ukuran. Dari kedua metode tersebut kemudian diturunkan suatu algoritma yang dapat dipakai secara umum untuk menduga parameter model *Hidden* Markov Elliott, et. al. (1995).

Tulisan ini dimulai dengan definisi model *Hidden* Markov beserta karakteristiknya. Pada bagian 3 dibahas Pendugaan Parameter model dan terakhir pada bagian 4 diturunkan algoritmanya.

*Tulisan ini merupakan bagian dari hasil penelitian yang didanai oleh Hibah Penelitian PHK A2 Departemen Matematika IPB tahun 2006

2. MODEL *HIDDEN* MARKOV

2.1 Definisi

Pasangan proses stokastik $\{(X_k, Y_k) : k \in \aleph\}$ yang terdefinisi pada ruang peluang $(\Omega, \mathcal{F}, P)$ dan mempunyai nilai pada $S \times \mathcal{Y}$ disebut model *hidden* Markov apabila $\{X_k\}$ adalah rantai Markov dengan *state* berhingga dan diasumsikan bahwa rantai Markov $\{X_k\}$ tidak diamati. Sehingga $\{X_k\}$ tersembunyi (*hidden*) di balik proses observasi $\{Y_k\}$. Banyaknya elemen dari *S* disebut ukuran (orde) dari model *hidden* Markov.

Pada tulisan ini dibahas model *hidden* Markov Elliot, et. al. (1995) yang berbentuk:

$$X_{k+1} = AX_k + V_{k+1}$$
$$Y_{k+1} = c(X_k) + \sigma(X_k)\omega_{k+1}$$

di mana $\{X_k\}$ adalah rantai Markov yang homogen dengan ruang state $S = \{e_1, e_2, ..., e_N\}$, dengan e_i vektor satuan di $\Re^N$ dan $A = (a_{ji})_{N\times N}$ merupakan matriks transisinya, dengan

$$a_{ji} = P(X_k = e_j \mid X_{k-1} = e_i) \qquad i, j = 1, 2, ..., N$$

$\{\omega_k\}$ adalah barisan peubah acak yang bebas stokastik identik dengan sebaran $N(0,1)$ dan

$$c(X_k) = \langle c, X_k \rangle \quad \text{dan} \quad \sigma(X_k) = \langle \sigma, X_k \rangle$$

dengan

$$c = (c_1, c_2, ..., c_N) \in \Re^N, \quad \sigma = (\sigma_1, \sigma_2, ..., \sigma_N) \in \Re^N$$
$$\langle .,. \rangle = \text{perkalian dalam di } \Re^N.$$

Asumsikan $\sigma_i > 0,$ untuk $i = 1, 2, ..., N.$

Misalkan $\{\mathcal{F}_k\}_k$ adalah filtrasi lengkap yang dibangun oleh $\{X_k\}_k$, $\{\mathcal{Y}_k\}$ adalah filtrasi lengkap yang dibangun oleh $\{Y_k\}$ dan $\{\mathcal{G}_k\}_k$ adalah filtrasi lengkap yang dibangun oleh $\{X_k\}_k$ dan $\{Y_k\}$.

Catatan 2.1.1 Karena $\omega_k, k \in \aleph$ adalah peubah acak yang bebas stokastik identik, maka ω_k bebas dari $\mathcal{G}_k$. Akibatnya ω_k juga bebas dari $\mathcal{F}_k$.

2.2 Nilai Harapan Bersyarat

Untuk sebarang $t \in \Re$, berlaku

$$\begin{aligned}
P(Y_{k+1} \le t \mid \mathrm{Y}_k) &= \sum_{i=1}^{N} P(Y_{k+1} \le t, X_k = e_i \mid \mathrm{Y}_k) \\
&= \sum_{i=1}^{N} P(Y_{k+1} \le t \mid X_k = e_i, \mathrm{Y}_k) \cdot P(X_k = e_i \mid \mathrm{Y}_k) \\
&= \sum_{i=1}^{N} P(Y_{k+1} \le t \mid X_k = e_i) \cdot P(X_k = e_i \mid \mathrm{Y}_k) \\
&= \sum_{i=1}^{N} P(c(X_k) + \sigma(X_k)\omega_{k+1} \le t \mid X_k = e_i) \cdot P(X_k = e_i \mid \mathrm{Y}_k) \\
&= \sum_{i=1}^{N} P(c_i + \sigma_i \omega_{k+1} \le t) \cdot P(X_k = e_i \mid \mathrm{Y}_k) \\
&= \sum_{i=1}^{N} P(\sigma_i \omega_{k+1} \le t - c_i) \cdot P(X_k = e_i \mid \mathrm{Y}_k).
\end{aligned}$$

Misalkan $\hat{X}_k := E\left[X_k \mid \mathrm{Y}_k\right]$ dan $\phi_i(x) = \dfrac{1}{\sqrt{2\pi\sigma_i}} e^{-\frac{1}{2\sigma_i}x^2}$ (fungsi kepadatan $N(0,\sigma_i)$)

maka

$$\begin{aligned}
\left\langle \hat{X}_k, e_i \right\rangle = \left\langle E\left[X_k \mid \mathrm{Y}_k\right], e_i \right\rangle &= \left\langle \sum_{j=1}^{N} e_j \cdot P\left(X_k = e_j \mid \mathrm{Y}_k\right), e_i \right\rangle \\
&= \sum_{j=1}^{N} P\left(X_k = e_j \mid \mathrm{Y}_k\right) \left\langle e_j, e_i \right\rangle = P\left(X_k = e_i \mid \mathrm{Y}_k\right).
\end{aligned}$$

Sehingga diperoleh

$$P(y_{k+1} \le t \mid \mathrm{Y}_k) = \sum_{i=1}^{N} P(\sigma_i \omega_{k+1} \le t - c_i) \cdot P(X_k = e_i \mid \mathrm{Y}_k) = \sum_{i=1}^{N} \left\langle \hat{X}_k, e_i \right\rangle \int_{-\infty}^{t-c_i} \phi_i(x)dx. \tag{2.1}$$

Jadi fungsi kepadatan bersyarat dari Y_{k+1} diketahui Y_k adalah

$$\sum_{j=1}^{N} \left\langle \hat{X}_k, e_j \right\rangle \phi_j(t - c_j).$$

Adapun sebaran gabungan dari X_k dan Y_{k+1} diketahui Y_k adalah

$$\begin{aligned}
P(X_k = e_i, Y_{k+1} \le t \mid \mathrm{Y}_k) &= P(\mathrm{Y}_{k+1} \le t \mid X_k = e_i, \mathrm{Y}_k) \cdot P(X_k = e_i \mid \mathrm{Y}_k) \\
&= P(c_i + \sigma_i \omega_{k+1} \le t) \cdot P(X_k = e_i \mid \mathrm{Y}_k) = \left\langle \hat{X}_k, e_i \right\rangle \int_{-\infty}^{t-c_i} \phi_i(x)dx.
\end{aligned}$$

Sehingga diperoleh fungsi kepadatan gabungan bersyarat dari X_k dan Y_{k+1} diketahui Y_k adalah $\left\langle \hat{X}_k, e_i \right\rangle \phi_i(t - c_i)$.

Berdasarkan aturan Bayes, diperoleh

$$E\left[\left\langle X_k, e_i \right\rangle \mid \mathrm{Y}_{k+1}\right] = P\left(X_k = e_i \mid Y_{k+1}, \mathrm{Y}_k\right) = \frac{P\left(X_k = e_i, Y_{k+1} \mid \mathrm{Y}_k\right)}{P\left(Y_{k+1} \mid \mathrm{Y}_k\right)} = \frac{\left\langle \hat{X}_k, e_i \right\rangle \phi_i(\mathrm{Y}_{k+1} - c_i)}{\sum_{j=1}^{N} \left\langle \hat{X}_k, e_j \right\rangle \phi_j(\mathrm{Y}_{k+1} - c_j)}.$$ Akibatnya didapat

$$E\left[X_k \mid \mathrm{Y}_{k+1}\right]=E\left[\sum_{i-1}^{N}\left\langle X_k,e_i\right\rangle e_i \middle| \mathrm{Y}_{k+1}\right]=\sum_{i-1}^{N}E\left[\left\langle X_k,e_i\right\rangle \middle| \mathrm{Y}_{k+1}\right] e_i=\frac{\sum_{i=1}^{N}\left\langle \hat{X}_k,e_i\right\rangle \phi_i(y_{k+1}-c_i)\ e_i}{\sum_{j=1}^{N}\left\langle \hat{X}_k,e_j\right\rangle \phi_j(y_{k+1}-c_j)}\ .$$

Teorema 2.2.1 (Elliot et. al. 1995)

$$\hat{X}_{k+1}=E\left[X_{k+1} \mid \mathrm{Y}_{k+1}\right]=\frac{\sum_{i=1}^{N}\left\langle \hat{X}_k,e_i\right\rangle \phi_i(y_{k+1}-c_i)\ A\ e_i}{\sum_{j=1}^{N}\left\langle \hat{X}_k,e_j\right\rangle \phi_j(y_{k+1}-c_j)}\ .$$

Catatan 2.2.2 Dari persamaan di atas diperoleh bahwa penduga $\hat{X}_{k+1}$ bergantung pada $\hat{X}_k$ secara tidak linear.

2.3 Perubahan Ukuran Pada Model *Hidden* Markov

Misalkan $\omega(\cdot)$ adalah peubah acak yang terdefinisi pada $(\Omega,\mathrm{F}\ ,P)$ dengan fungsi kepadatan $\phi(\omega)$ dan c, σ adalah konstanta yang diketahui. Diketahui $Y(\cdot)=c+\sigma\ \omega(\cdot)$.

Akan dikonstruksi ukuran peluang baru $\overline{P}$ pada $(\Omega,\mathrm{F}\ )$ sedemikian sehingga:

a. $\dfrac{d\overline{P}}{dP}=\lambda$

b. Di bawah $\overline{P}$ peubah acak *y* mempunyai fungsi kepadatan ϕ.

Karena

$$\begin{aligned}\overline{P}(Y\le t)&=\int_{-\infty}^{t}\phi(y)\ dy=\int_{\Omega}I_{\{Y\le t\}}\ d\overline{P}=\int_{\Omega}I_{\{Y\le t\}}\lambda\ dP=\int_{\Omega}I_{\left\{\omega\le\frac{t-c}{\sigma}\right\}}\lambda(\omega)\phi(\omega)\ d\omega\\&=\int_{-\infty}^{\frac{t-c}{\sigma}}\lambda(\omega)\phi(\omega)\ d\omega=\int_{-\infty}^{t}\lambda(\omega)\phi(\omega)\frac{1}{\sigma}\ dy\end{aligned}$$

maka diperoleh $\phi(y)=\dfrac{\lambda(\omega)\phi(\omega)}{\sigma}$ atau $\lambda(\omega)=\dfrac{\sigma\phi(y)}{\phi(\omega)}$.

Pada $(\Omega,\mathrm{F}\ ,P)$ proses observasi $\left\{Y_k\right\}$ mempunyai bentuk

$$Y_{k+1}=\left\langle c,X_k\right\rangle+\left\langle \sigma,X_k\right\rangle\omega_{k+1}$$

di mana $\left\{\omega_k\right\}$bebas stokastik identik menyebar *N*(0,1). Misalkan $\phi(\cdot)$ adalah fungsi kepadatan peluang *N*(0,1) dan

$$\lambda_l=\frac{\left\langle\sigma,X_{l-1}\right\rangle\phi\left(y_l\right)}{\phi\left(\omega_l\right)},\quad l\in\aleph;\quad \Lambda_0=1;\quad \Lambda_k=\prod_{l=1}^{k}\lambda_l\ ,\quad k\ge 1.$$

Definisikan ukuran peluang $\bar{P}$ pada $(\Omega, \mathcal{F})$ sebagai berikut $\left.\frac{d\bar{P}}{dP}\right|_{\mathcal{G}_k} = \Lambda_k$.

Eksistensi Λ_k dijamin oleh Teorema Radon-Nikodym dan eksistensi $\bar{P}$ dijamin oleh Teorema Perluasan Kolmogorov (Wong and Hajek, 1985).

Lemma 2.3.1 (Elliot et. al. 1995) Di bawah $\bar{P}$, $\{Y_k\}$ adalah barisan peubah acak yang bebas stokastik identik menyebar *N*(0,1).

Sebaliknya, dimulai dengan ukuran peluang $\bar{P}$ pada $(\Omega, \mathcal{F})$, di mana di bawah $\bar{P}$ berlaku:

a. $\{X_k\}$ adalah rantai Markov dengan matriks transisi *A* sehingga $X_{k+1} = AX_k + V_{k+1}$, dengan $\bar{E}\left[V_{k+1} \mid \mathcal{F}_k\right] = 0$

b. $\{Y_k\}$ adalah barisan peubah acak yang bebas stokastik identik menyebar *N*(0,1) dan bebas dari X_k,

akan dikonstruksi *P* dari $\bar{P}$ sehingga di bawah *P* berlaku:

$$\omega_{k+1} = \frac{Y_{k+1} - \langle c, X_k\rangle}{\langle \sigma, X_k\rangle}, \quad k \in \aleph \quad \left(\langle \sigma, X_k\rangle \neq 0\right)$$

adalah barisan peubah acak yang bebas stokastik identik dan menyebar *N*(0,1).

Untuk mengkonstruksi *P* dari $\bar{P}$, definisikan

$$\bar{\lambda}_l = \frac{1}{\lambda_l} = \frac{\phi(\omega_l)}{\langle \sigma, X_{l-1}\rangle \phi(y_l)}, \quad l \in \aleph; \quad \bar{\Lambda}_0 = 1; \quad \bar{\Lambda}_k = \prod_{l=1}^{k} \bar{\lambda}_l, \quad k \geq 1; \quad \left.\frac{d\bar{P}}{dP}\right|_{\mathcal{G}_k} = \bar{\Lambda}_k.$$

Lemma 2.3.2 Di bawah P, $\{\omega_k\}$ adalah barisan peubah acak yang bebas stokastik identik menyebar *N*(0,1).

Catatan 2.3.3 Untuk selanjutnya kita akan bekerja pada ruang peluang $(\Omega, \mathcal{F}, \bar{P})$.

3. PENDUGAAN PARAMETER

Sifat statistik model *Hidden* Markov ditentukan secara lengkap oleh himpunan parameter

$$\theta = \left\{(a_{ji}), 1 \leq i, j \leq N;\ c_i, 1 \leq i \leq N;\ \sigma_i, 1 \leq i \leq N\right\}.$$

Pada bagian ini dibahas proses pendugaan parameter tersebut menggunakan Algoritma EM (*Expectation maximization algorithm*).

3.1 Pendugaan Rekursif

Definisi 3.1.1 (Elliot et. al. 1995) Barisan peubah acak $\{\phi_k\}$ disebut *adapted*-$\mathcal{G}$, atau *adapted* terhadap filtrasi $\{\mathcal{G}_k\}$, apabila untuk setiap *k*, ϕ_k terukur-$\mathcal{G}_k$.

Definisi 3.1.2 Jika $\{H_k\}$ adalah barisan peubah acak yang *adapted* terha-dap $\{\mathcal{G}_k\}$, definisikan $\gamma_k(H_k) := \bar{E}\left[\Lambda_k H_k \middle| \mathcal{Y}_k\right]$ dan $\hat{H}_k := E\left[H_k \middle| \mathcal{Y}_k\right]$.

Menurut Teorema Bayes bersyarat

$$\hat{H}_k := E\left[H_k \middle| \mathcal{Y}_k\right] = \frac{\bar{E}\left[\Lambda_k H_k \middle| \mathcal{Y}_k\right]}{\bar{E}\left[\Lambda_k \middle| \mathcal{Y}_k\right]} = \frac{\gamma_k(H_k)}{\gamma_k(1)},$$

sehingga

$$\hat{X}_0 = E\left[X_0 \middle| \mathcal{Y}_0\right] = \frac{\gamma_0(X_0)}{\gamma_0(1)} = \gamma_0(X_0)$$

karena $\gamma_0(1) = 1$.

Catatan 3.1.3 Pada proses pendugaan rekursif, $\gamma_0(X_0) = E[X_0]$ diambil sebagai nilai awal.

Lemma 3.1.4. (Elliot et. al. 1995)
Misalkan $\{H_k\}$ adalah barisan peubah acak bernilai skalar, maka

a. $\gamma_k(H_k) = \left\langle \gamma_k(H_k X_k), \bar{1} \right\rangle$
b. $\gamma_k(1) = \left\langle \gamma_k(X_k), \bar{1} \right\rangle$.

Catatan 3.1.5 Dari persamaan di atas diperoleh bahwa pendugaan $\gamma_{k+1}(H_{k+1})$ bergantung pada $\gamma_k(H_k X_k)$.

Definisi 3.1.6 Barisan peubah acak $\{\phi_k\}$ disebut *predictable* terhadap filtrasi $\{\mathcal{G}_k\}$, apabila untuk setiap *k*, ϕ_k terukur-$\mathcal{G}_{k-1}$.

Teorema 3.1.7 Misalkan $\{H_k\}$ adalah proses bernilai skalar yang *adapted* terhadap filtrasi $\{\mathcal{G}_k\}$ dan mempunyai bentuk

a. H_0 terukur-$\mathcal{F}_0$
b. $H_{k+1} = H_k + \alpha_{k+1} + \left\langle \beta_{k+1}, V_{k+1} \right\rangle + \delta_{k+1} f(y_{k+1}), \quad k \geq 1$

di mana

- $V_{k+1} = X_{k+1} - AX_k$
- f adalah fungsi bernilai skalar
- α, β, δ adalah proses yang *predictable*-$\mathcal{G}$

- β adalah proses bernilai vektor berdimensi-N

maka

$$\gamma_{k+1}\left(H_{k+1}X_{k+1}\right) := \gamma_{k+1,k+1}\left(H_{k+1}\right)$$
$$= \sum_{i=1}^{N}\Big\{\left\langle \gamma_k\left(H_k X_k\right), \Gamma^i\left(y_{k+1}\right)\right\rangle a_i + \gamma_k\left(\alpha_{k+1}\left\langle X_k, \Gamma^i\left(y_{k+1}\right)\right\rangle\right) a_i$$
$$+ \gamma_k\left(\delta_{k+1}\left\langle X_k, \Gamma^i\left(y_{k+1}\right)\right\rangle\right) f(y_{k+1})\, a_i + \left(\mathrm{diag}(a_i) - a_i a_i^T\right)\gamma_k\left(\beta_{k+1}\left\langle X_k, \Gamma^i\left(y_{k+1}\right)\right\rangle\right)\Big\}$$

di mana $a_i = Ae_i$ dan $\Gamma^{(\cdot)}\left(y_k\right) := \dfrac{\phi\left(\dfrac{y_k - c_{(\cdot)}}{\sigma_{(\cdot)}}\right)}{\sigma_{(\cdot)}\phi\left(y_k\right)}\, e_{(\cdot)}$.

3.2 Penduga Parameter

3.2.1 Penduga untuk *State*

Teorema 3.2.1 (Elliott et. al. 1995)

$$\gamma_{k+1}(X_{k+1}) = \sum_{i=1}^{N}\left\langle \gamma_k(X_k), \Gamma^i(y_{k+1})\right\rangle a_i\,. \tag{3.1}$$

Bukti:

Dengan mengambil $H_k = H_0 = 1,\ \alpha_k = 0,\ \beta_k = 0$ dan $\delta_k = 0$ pada Teorema 3.1.7 diperoleh $\gamma_{k+1}(X_{k+1}) = \sum_{i=1}^{N}\left\langle \gamma_k(X_k), \Gamma^i(y_{k+1})\right\rangle a_i$.

Catatan 3.2.2 Persamaan (3.1) menunjukkan bahwa $\gamma_{k+1}(X_{k+1})$ bergantung pada $\gamma_k(X_k)$ secara linear.

3.2.2 Penduga untuk *Number of Jumps*

Jika rantai Markov melompat dari state e_r pada waktu ke-k, ke state e_s pada waktu ke-$k+1$, $1 \le r, s \le N$, maka $\left\langle X_k, e_r\right\rangle\left\langle X_{k+1}, e_s\right\rangle = 1$, sehingga banyaknya lompatan (*number of jumps*) dari state e_r ke state e_s pada waktu ke-$k+1$ adalah

$$\mathrm{J}_{k+1}^{rs} := \sum_{n=1}^{k+1}\left\langle X_{n-1}, e_r\right\rangle\left\langle X_n, e_s\right\rangle$$
$$= \mathrm{J}_k^{rs} + \left\langle X_k, e_r\right\rangle\left\langle X_{k+1}, e_s\right\rangle$$
$$= \mathrm{J}_k^{rs} + \left\langle X_k, e_r\right\rangle\left\langle AX_k + V_{k+1}, e_s\right\rangle$$
$$= \mathrm{J}_k^{rs} + \left\{\left\langle X_k, e_r\right\rangle\left\langle AX_k, e_s\right\rangle + \left\langle X_k, e_r\right\rangle\left\langle V_{k+1}, e_s\right\rangle\right\}$$
$$= \mathrm{J}_k^{rs} + \left\{\left\langle X_k, e_r\right\rangle a_{sr} + \left\langle X_k, e_r\right\rangle\left\langle V_{k+1}, e_S\right\rangle\right\}.$$

Teorema 3.2.2 (Elliott et. al. 1995)

$$\gamma_{k+1,k+1}(\mathrm{J}_{k+1}^{rs}) = \sum_{i=1}^{N} \left\langle \gamma_{k,k}(\mathrm{J}_{k}^{rs}), \Gamma^i(y_{k+1}) \right\rangle a_i + \left\langle \gamma_k(X_k), \Gamma^r(y_{k+1}) \right\rangle a_{sr} e_s \,. \tag{3.2}$$

Bukti: Dengan mengambil
$H_{k+1} = \mathrm{J}_{k+1}^{rs}$, $H_0 = 0$, $\alpha_{k+1} = \langle X_k, e_r \rangle a_{sr}$, $\beta_{k+1} = \langle X_k, e_r \rangle e_s$ dan $\delta_{k+1} = 0$ pada Teorema 3.1.7 diperoleh

$$\gamma_{k+1,k+1}\left(\mathrm{J}_{k+1}^{rs}\right) = \sum_{i=1}^{N} \Big\{ \left\langle \gamma_k\left(\mathrm{J}_{k}^{rs} X_k\right), \Gamma^i\left(y_{k+1}\right) \right\rangle a_i + \gamma_k\left(\left\langle X_k, e_r \right\rangle a_{sr} \left\langle X_k, \Gamma^i\left(y_{k+1}\right) \right\rangle \right) a_i$$
$$+ 0 + \left(\mathrm{diag}(a_i) - a_i a_i^T\right) \gamma_k\left(\left\langle X_k, e_r \right\rangle e_s \left\langle X_k, \Gamma^i\left(y_{k+1}\right) \right\rangle \right) \Big\}.$$

Sedangkan

$$\sum_{i=1}^{N} \gamma_k\left(\left\langle X_k, e_r \right\rangle a_{sr} \left\langle X_k, \Gamma^i\left(y_{k+1}\right) \right\rangle \right) a_i = \sum_{i=1}^{N} \gamma_k\left(\left\langle X_k, e_r \right\rangle \left\langle X_k, \Gamma^i\left(y_{k+1}\right) \right\rangle \right) a_{sr} a_i$$
$$= \gamma_k\left(\left\langle X_k, \Gamma^r\left(y_{k+1}\right) \right\rangle \right) a_{sr} a_r$$

dan

$$\sum_{i=1}^{N} \left(\mathrm{diag}(a_i) - a_i a_i^T\right) \gamma_k\left(\left\langle X_k, e_r \right\rangle e_s \left\langle X_k, \Gamma^i\left(y_{k+1}\right) \right\rangle \right)$$
$$= \sum_{i=1}^{N} \gamma_k\left(\left\langle X_k, e_r \right\rangle \left\langle X_k, \Gamma^i\left(y_{k+1}\right) \right\rangle \right) \left(\mathrm{diag}(a_i) - a_i a_i^T\right) e_s$$
$$= \gamma_k\left(\left\langle X_k, \Gamma^r\left(y_{k+1}\right) \right\rangle \right) \left(\mathrm{diag}(a_r) - a_r a_r^T\right) e_s$$
$$= \gamma_k\left(\left\langle X_k, \Gamma^r\left(y_{k+1}\right) \right\rangle \right) \left(a_{sr} e_s - a_{sr} a_r\right).$$

Sehingga

$$\gamma_{k+1,k+1}(\mathrm{J}_{k+1}^{rs}) = \sum_{i=1}^{N} \left\langle \gamma_k(\mathrm{J}_{k}^{rs} X_k), \Gamma^i(y_{k+1}) \right\rangle a_i + \gamma_k\left(\left\langle X_k, \Gamma^r(y_{k+1}) \right\rangle \right) a_{sr} e_s$$
$$= \sum_{i=1}^{N} \left\langle \gamma_{k,k}(\mathrm{J}_{k+1}^{rs}), \Gamma^i(y_{k+1}) \right\rangle a_i + \left\langle \gamma_k(X_k), \Gamma^r(y_{k+1}) \right\rangle a_{sr} e_s .$$

3.2.3 Penduga untuk *Occupation Time*

Banyaknya kejadian rantai Markov berada pada state e_r, $1 \le r \le N$ sampai waktu ke-k didefinisikan sebagai berikut.

$$O_{k+1}^{r} := \sum_{n=1}^{k+1} \left\langle X_{n-1}, e_r \right\rangle = O_k^{rs} + \left\langle X_k, e_r \right\rangle .$$

Teorema 3.2.3 (Elliott et. al. 1995)

$$\gamma_{k+1,k+1}(O_{k+1}^{r}) = \sum_{i=1}^{N} \left\langle \gamma_{k,k}(O_k^{r}), \Gamma^i(y_{k+1}) \right\rangle a_i + \left\langle \gamma_k(X_k), \Gamma^r(y_{k+1}) \right\rangle a_r \,. \tag{3.3}$$

Bukti: Dengan mengambil
$H_{k+1} = O_{k+1}^{rs},\ H_0 = 0,\ \alpha_{k+1} = \langle X_k, e_r \rangle,\ \beta_{k+1} = 0$ dan $\delta_{k+1} = 0$ pada Teorema 3.1.7 diperoleh

$$\begin{aligned}\gamma_{k+1,k+1}\left(O_{k+1}^{r}\right) &= \sum_{i=1}^{N} \left\langle \gamma_k\left(O_k^r X_k\right), \Gamma^i\left(y_{k+1}\right)\right\rangle a_i + \gamma_k\left(\left\langle X_k, e_r\right\rangle \left\langle X_k, \Gamma^i\left(y_{k+1}\right)\right\rangle\right) a_i \\ &= \sum_{i=1}^{N} \left\langle \gamma_k\left(O_k^r X_k\right), \Gamma^i\left(y_{k+1}\right)\right\rangle a_i + \gamma_k\left(\left\langle X_k, \Gamma^r\left(y_{k+1}\right)\right\rangle\right) a_r \\ &= \sum_{i=1}^{N} \left\langle \gamma_{k,k}\left(O_k^r\right), \Gamma^i\left(y_{k+1}\right)\right\rangle a_i + \left\langle \gamma_k(X_k), \Gamma^r(y_{k+1})\right\rangle a_r.\end{aligned}$$

3.2.4 Penduga untuk Proses Observasi

Untuk menduga parameter $\sigma = (\sigma_1, \sigma_2, ..., \sigma_N)^T$ dan $c = (c_1, c_2, ..., c_N)^T$ pada proses observasi $y_{k+1} = \langle c, X_k \rangle + \langle \sigma, X_k \rangle \omega_{k+1}$, definisikan

$$\begin{aligned}\tau_{k+1}^r(f) &:= \sum_{l=1}^{k+1} \langle X_{l-1}, e_r \rangle f(y_l) \qquad 1 \le r \le N \\ &= \tau_k^r(f) + \langle X_k, e_r \rangle f(y_{k+1})\end{aligned}$$

di mana $f(y) = y$ atau $f(y) = y^2$.

Teorema 3.2.4 (Elliott et. al. 1995)

$$\gamma_{k+1,k+1}(\tau_{k+1}^r(f)) = \sum_{i=1}^{N} \left\langle \gamma_{k,k}(\tau_k^r(f)), \Gamma^i(y_{k+1})\right\rangle a_i + \left\langle \gamma_k(X_k), \Gamma^r(y_{k+1})\right\rangle f(y_{k+1}) a_r . \quad (3.4)$$

Bukti:
Dengan mengambil
$H_{k+1} = \tau_{k+1}^r(f),\ H_0 = 0,\ \alpha_{k+1} = 0,\ \beta_{k+1} = 0$ dan $\delta_{k+1} = \langle X_k, e_r \rangle$ pada Teorema 3.1.7 diperoleh

$$\begin{aligned}\gamma_{k+1,k+1}\left(\tau_{k+1}^r(f)\right) &= \sum_{i=1}^{N} \left\langle \gamma_k\left(\tau_k^r(f) X_k\right), \Gamma^i\left(y_{k+1}\right)\right\rangle a_i + \gamma_k\left(\left\langle X_k, e_r\right\rangle \left\langle X_k, \Gamma^i\left(y_{k+1}\right)\right\rangle\right) f(y_{k+1}) a_i \\ &= \sum_{i=1}^{N} \left\langle \gamma_k\left(\tau_k^r(f) X_k\right), \Gamma^i\left(y_{k+1}\right)\right\rangle a_i + \gamma_k\left(\left\langle X_k, \Gamma^r\left(y_{k+1}\right)\right\rangle\right) f(y_{k+1}) a_r \\ &= \sum_{i=1}^{N} \left\langle \gamma_{k,k}\left(\tau_k^r(f)\right), \Gamma^i\left(y_{k+1}\right)\right\rangle a_i + \left\langle \gamma_k\left(X_k\right), \Gamma^r\left(y_{k+1}\right)\right\rangle f(y_{k+1}) a_r.\end{aligned}$$

3.3 *Expectation Maximization Algorithm* (Algoritma EM)

Algoritma EM dikembangkan oleh Baum and Petrie (1966) dengan ide dasar sebagai berikut.

Misalkan $\{P_\theta : \theta \in \Theta\}$ adalah koleksi ukuran peluang yang terdefinisi pada ruang (Ω, G) dan kontinu absolut terhadap P_0. Misalkan $\mathrm{Y} \subset \mathrm{G}$. Definisikan fungsi *likelihood* untuk menentukan penduga parameter θ berdasarkan informasi Y sebagai

$$L(\theta) = E_0\left[\left.\frac{dP_\theta}{dP_0}\right| \mathrm{Y}\right]$$

dan penduga maksimum *likelihood* didefinisikan sebagai

$$\hat{\theta} \in \arg\max_{\theta \in \Theta} L(\theta).$$

Secara umum penduga maksimum likelihood $\hat{\theta}$ sulit dihitung secara langsung. Algoritma EM memberikan suatu metode iteratif untuk mengaproksimasi $\hat{\theta}$, dengan prosedur sebagai berikut.

Langkah 1: Set $p = 0$ dan pilih $\hat{\theta}_0$.

Langkah 2: [Langkah-E]

Set $\theta^* = \hat{\theta}_p$ dan hitung $Q\left(\theta, \theta^*\right) = E_{\theta^*}\left[\left.\log \frac{dP_\theta}{dP_{\theta^*}}\right| \mathrm{Y}\right]$.

Langkah 3: [Langkah-M]

Tentukan $\hat{\theta}_{p+1} \in \arg\max_{\theta \in \Theta} Q\left(\theta, \theta^*\right)$.

Langkah 4: $p \leftarrow p + 1$

Ulangi langkah 2 sampai kriteria berhenti dipenuhi.

Catatan 3.3.1

1. Barisan $\left\{\hat{\theta}_p : p \geq 0\right\}$ memberikan barisan $\left\{L\left(\hat{\theta}_p\right) : p \geq 0\right\}$ yang tak turun.
2. Menurut ketaksamaan Jensen,
$$Q\left(\hat{\theta}_{p+1}, \hat{\theta}_p\right) \leq \log L(\hat{\theta}_{p+1}) - \log L(\hat{\theta}_p).$$
3. $Q\left(\theta, \theta^*\right)$ disebut *pseudo-loglikelihood* bersyarat.

Model *Hidden* Markov yang akan diduga parameternya berbentuk

$$X_{k+1} = AX_k + V_{k+1}$$
$$y_{k+1} = \langle c, X_k \rangle + \langle \sigma, X_k \rangle \omega_{k+1} \quad k \in \aleph$$

di mana V_{k+1} adalah martingale *increments* dan ω_k adalah peubah acak yang bebas stokastik identik menyebar $N(0,1)$. Parameter model diberikan oleh himpunan

$$\theta = \left\{(a_{ji}), 1 \leq i, j \leq N;\ c_i, 1 \leq i \leq N;\ \sigma_i, 1 \leq i \leq N\right\}$$

di mana $\sum_{i=1}^{N} a_{ji} = 1,\ 1 \le i \le N$.

Akan ditentukan himpunan parameter baru

$$\hat{\theta}(k) = \left\{ (\hat{a}_{ji}(k)), 1 \le i,\ j \le N;\ \hat{c}_i(k), 1 \le i \le N;\ \hat{\sigma}_i(k), 1 \le i \le N \right\}$$

di mana $\sum_{i=1}^{N} \hat{a}_{ji}(k) = 1,\ 1 \le i \le N$ yang memaksimumkan *pseudolog-likelihood* bersyarat.

Untuk mengubah parameter a_{ji} menjadi $\hat{a}_{ji}(k)$ pada rantai Markov X, misalkan

$$\lambda_k = \prod_{r,s=1}^{N} \left[\frac{\hat{a}_{sr}(k)}{a_{sr}} \right]^{\langle X_k, e_s \rangle \langle X_{k-1}, e_r \rangle}$$

$$\Lambda_k = \prod_{l=1}^{k} \prod_{r,s=1}^{N} \left[\frac{\hat{a}_{sr}(k)}{a_{sr}} \right]^{\langle X_l, e_s \rangle \langle X_{l-1}, e_r \rangle}$$

dan definisikan peluang $P_{\hat{\theta}}$ sehingga

$$\left. \frac{dP_{\hat{\theta}}}{dP_{\theta}} \right|_{\mathrm{F}_k} = \Lambda_k .$$

Lemma 3.3.2

Di bawah $P_{\hat{\theta}}$, jika $X_k = e_r$, maka $E_{\hat{\theta}} \left[\langle X_{k+1}, e_s \rangle \mid \mathrm{F}_k \right] = \hat{a}_{sr}(k+1)$.

Teorema 3.3.3 (Elliott et. al. 1995)

$$\hat{a}_{sr}(k) = \frac{\hat{\mathrm{J}}_k^{rs}}{\hat{O}_k^r} = \frac{\gamma_k \left(\mathrm{J}_k^{rs} \right)}{\gamma_k \left(O_k^r \right)} . \tag{3.5}$$

Bukti: Berdasarkan definisi

$$\begin{aligned}
\log \Lambda_k &= \log \left(\prod_{l=1}^{k} \prod_{r,s=1}^{N} \left[\frac{\hat{a}_{sr}(k)}{a_{sr}} \right]^{\langle X_l, e_s \rangle \langle X_{l-1}, e_r \rangle} \right) \\
&= \sum_{l=1}^{k} \sum_{r,s=1}^{N} \langle X_l, e_s \rangle \langle X_{l-1}, e_r \rangle \left(\log \hat{a}_{sr}(k) - \log a_{sr} \right) \\
&= \sum_{r,s=1}^{N} \mathrm{J}_k^{rs} \log \hat{a}_{sr}(k) + R(a)
\end{aligned}$$

di mana $R(a)$ tidak bergantung pada $\hat{a}$. Sehingga *pseudo-loglikelihood* bersyaratnya menjadi

$$E\left[\log \Lambda_k \mid Y_k\right] = \sum_{r,s=1}^{N} \hat{J}_k^{rs} \log \hat{a}_{sr}(k) + \hat{R}(a). \tag{3.6}$$

Parameter $\hat{a}_{sr}(k)$ harus memenuhi

$$\sum_{s=1}^{N} \hat{a}_{sr}(k) = 1$$

atau dalam bentuk dinamik

$$\sum_{l=1}^{k}\sum_{s=1}^{N} \left\langle X_{l-1}, e_r \right\rangle \hat{a}_{sr}(k) = k \cdot$$

Bentuk dinamik di atas dapat dituliskan dalam bentuk bersyarat

$$\sum_{r,s=1}^{N} \hat{O}_k^r \hat{a}_{sr}(k) = k \cdot \tag{3.7}$$

Masalah optimasinya sekarang menjadi memilih $\hat{a}_{sr}(k)$ yang memaksimumkan (3.7) dengan kendala (3.8). Definisikan fungsi Lagrange

$$L(\hat{a}, \lambda) = \sum_{r,s=1}^{N} \hat{J}_k^{rs} \log \hat{a}_{sr}(k) + \lambda \left(\sum_{r,s=1}^{N} \hat{O}_k^r \hat{a}_{sr}(k) - k \right)$$

Dari $\dfrac{\partial L}{\partial \lambda} = 0$ dan $\dfrac{\partial L}{\partial \hat{a}_{sr}(k)} = 0$ diperoleh

$$\sum_{r,s=1}^{N} \hat{O}_k^r \hat{a}_{sr}(k) = k \tag{3.8}$$

$$\frac{\hat{J}_k^{rs}}{\hat{a}_{sr}(k)} + \lambda \hat{O}_k^r = 0. \tag{3.9}$$

Dari persamaan (3.8) dan (3.9) diperoleh $\lambda = -1$. Dari persamaan (3.9) diperoleh

$$\hat{a}_{sr}(k) = \frac{\hat{J}_k^{rs}}{\hat{O}_k^r} = \frac{\gamma_k(J_k^{rs})}{\gamma_k(1)} \bigg/ \frac{\gamma_k(O_k^r)}{\gamma_k(1)} = \frac{\gamma_k(J_k^{rs})}{\gamma_k(O_k^r)}.$$

Untuk mengubah parameter c_i menjadi $\hat{c}_i(k)$, misalkan

$$\lambda_{k+1}^*(X_k, y_{k+1}) = \exp\left(\frac{1}{2\left\langle \sigma, X_k \right\rangle} \left\{ \left\langle c, X_k \right\rangle^2 - \left\langle \hat{c}, X_k \right\rangle^2 - 2y_{k+1}\left\langle c, X_k \right\rangle + 2y_{k+1}\left\langle \hat{c}, X_k \right\rangle \right\} \right)$$

$$\Lambda_k^* = \prod_{l=1}^{k} \lambda_l^*(X_{l-1}, y_l)$$

dan definisikan peluang P^* sehingga $\left.\dfrac{dP^*}{dP}\right|_{G_k} = \Lambda_k^* \cdot$

Lemma 3.3.4

Di bawah P^*, $\left\{ y_k - \left\langle \hat{c}, X_{k-1} \right\rangle \right\}$ $k \in \aleph$ adalah barisan peubah acak yang bebas stokastik identik menyebar $N(0,\ \sigma)$.

Bukti: Menurut Teorema Bayes bersyarat

$$P^*\left(y_{k+1}-\langle \hat{c}, X_k\rangle \le t \middle| \mathcal{G}_k\right) = E^*\left[I_{\{y_{k+1}-\langle \hat{c},X_k\rangle\le t\}} \middle| \mathcal{G}_k\right] = \frac{E\left[\Lambda^*_{k+1} I_{\{y_{k+1}-\langle \hat{c},X_k\rangle\le t\}} \middle| \mathcal{G}_k\right]}{E\left[\Lambda^*_{k+1}\middle|\mathcal{G}_k\right]}$$

$$= \frac{E\left[\Lambda^*_k \lambda^*_{k+1} I_{\{y_{k+1}-\langle \hat{c},X_k\rangle\le t\}} \middle| \mathcal{G}_k\right]}{E\left[\Lambda^*_k\lambda^*_{k+1}\middle|\mathcal{G}_k\right]} = \frac{\Lambda^*_k E\left[\lambda^*_{k+1} I_{\{y_{k+1}-\langle \hat{c},X_k\rangle\le t\}} \middle| \mathcal{G}_k\right]}{\Lambda^*_k E\left[\lambda^*_{k+1}\middle|\mathcal{G}_k\right]} = \frac{E\left[\lambda^*_{k+1} I_{\{y_{k+1}-\langle \hat{c},X_k\rangle\le t\}} \middle| \mathcal{G}_k\right]}{E\left[\lambda^*_{k+1}\middle|\mathcal{G}_k\right]}.$$

Di bawah P, $y_{k+1}-\langle c, X_k\rangle = \langle c, X_k\rangle \omega_{k+1}$ dengan ω_{k+1} menyebar $N(0,1)$, sehingga $y_{k+1}-\langle c, X_k\rangle$ menyebar $N\left(\langle c, X_k\rangle, \langle \sigma, X_k\rangle\right)$. Jadi

$$E\left[\lambda^*_{k+1}\middle|\mathcal{G}_k\right] = \int_{-\infty}^{\infty} \exp\left(\frac{1}{2\langle\sigma, X_k\rangle}\left\{\langle c, X_k\rangle^2 - \langle \hat{c}, X_k\rangle^2 - 2y_{k+1}\langle c, X_k\rangle + 2y_{k+1}\langle \hat{c}, X_k\rangle\right\}\right)$$

$$\times \frac{1}{\sqrt{2\pi\langle\sigma, X_k\rangle}} \exp\left(-\frac{1}{2\langle\sigma, X_k\rangle}\left\{y_{k+1}-\langle c, X_k\rangle\right\}^2\right) dy_{k+1}$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\langle\sigma, X_k\rangle}} \exp\left(-\frac{1}{2\langle\sigma, X_k\rangle}\left\{y_{k+1}-\langle \hat{c}, X_k\rangle\right\}^2\right) dy_{k+1} = 1.$$

Sehingga

$$P^*\left(y_{k+1}-\langle \hat{c}, X_k\rangle \le t \mid \mathcal{G}_k\right) = E\left[\lambda^*_{k+1} I_{\{y_{k+1}-\langle \hat{c},X_k\rangle\le t\}} \middle| \mathcal{G}_k\right]$$

$$= \int_{-\infty}^{\infty} \exp\left(\frac{1}{2\langle\sigma, X_k\rangle}\left\{\langle c, X_k\rangle^2 - \langle \hat{c}, X_k\rangle^2 - 2y_{k+1}\langle c, X_k\rangle + 2y_{k+1}\langle \hat{c}, X_k\rangle\right\}\right)$$

$$\times \frac{1}{\sqrt{2\pi\langle\sigma, X_k\rangle}} \exp\left(-\frac{1}{2\langle\sigma, X_k\rangle}\left\{y_{k+1}-\langle c, X_k\rangle\right\}^2\right) I_{\{y_{k+1}\le t+\langle \hat{c},X_k\rangle\}}\; dy_{k+1}$$

$$= \int_{-\infty}^{t-\langle \hat{c},X_k\rangle} \frac{1}{\sqrt{2\pi\langle\sigma, X_k\rangle}} \exp\left(-\frac{1}{2\langle\sigma, X_k\rangle}\left\{y_{k+1}-\langle \hat{c}, X_k\rangle\right\}^2\right) dy_{k+1}$$

$$= \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi\langle\sigma, X_k\rangle}} \exp\left(-\frac{1}{2\langle\sigma, X_k\rangle}u^2\right) du$$

$$= P^*\left(U \le t\right)$$

$$= P^*\left(y_{k+1}-\langle \hat{c}, X_k\rangle \le t\right).$$

Teorema 3.3.5 (Elliott et. al. 1995)

$$\hat{c}_r(k) = \frac{\hat{\tau}^r_k(y)}{\hat{O}^r_k} = \frac{\gamma_k\left(\tau^r_k(y)\right)}{\gamma_k\left(O^r_k\right)} \; . \tag{3.10}$$

Bukti: Dari definisi

$$
\begin{aligned}
\log \Lambda_k^* &= \sum_{l=1}^{k} \frac{\langle c, X_{l-1}\rangle^2 - \langle \hat{c}, X_{l-1}\rangle^2 - 2y_l\langle c, X_{l-1}\rangle + 2y_l\langle \hat{c}, X_{l-1}\rangle}{2\langle \sigma, X_{l-1}\rangle} \\
&= \sum_{l=1}^{k}\sum_{r=1}^{N} \frac{\langle X_{l-1}, e_r\rangle\left(c_r^2 - \hat{c}_r^2(k) - 2y_l\ c + 2y_l\ \hat{c}_r(k)\right)}{2\sigma_r} \\
&= \sum_{r=1}^{N} \frac{2\tau_k^r(y)\ \hat{c}_r(k) - O_k^r\ \hat{c}_r^2(k)}{2\sigma_r} + R(c)
\end{aligned}
$$

di mana $R(c)$ tidak bergantung pada $\hat{c}$.
Jadi

$$
E\left[\log \Lambda_k^* \mid \mathrm{Y}_k\right] = \sum_{r=1}^{N} \frac{2\hat{\tau}_k^r(y)\ \hat{c}_r(k) - \hat{O}_k^r\ \hat{c}_r^2(k)}{2\sigma_r} + \hat{R}(c)
$$

sehingga

$$
\frac{d}{d\hat{c}_r(k)} E\left[\log \Lambda_k^* \mid \mathrm{Y}_k\right] = \frac{2\hat{\tau}_k^r(y) - 2\hat{O}_k^r\hat{c}_r(k)}{2\sigma_r} = 0
$$

memberikan

$$
2\hat{\tau}_k^r(y) - 2\hat{O}_k^r\hat{c}_r(k) = 0
$$

$$
\hat{c}_r(k) = \frac{\hat{\tau}_k^r(y)}{\hat{O}_k^r} = \frac{\gamma_k(\tau_k^r(y))}{\gamma_k(1)} \Bigg/ \frac{\gamma_k(O_k^r)}{\gamma_k(1)} = \frac{\gamma_k(\tau_k^r(y))}{\gamma_k(O_k^r)}.
$$

Untuk mengubah parameter σ_i menjadi $\hat{\sigma}_i(k)$ (ambil c_i tetap), definisikan

$$
\tilde{\lambda}_{k+1}(X_k, y_{k+1}) = \sqrt{\frac{\langle \sigma, X_k\rangle}{\langle \hat{\sigma}, X_k\rangle}}\ \frac{\exp\left(\frac{1}{2\langle \hat{\sigma}, X_k\rangle}\left\{y_{k+1} - \langle c, X_k\rangle\right\}^2\right)}{\exp\left(\frac{1}{2\langle \sigma, X_k\rangle}\left\{y_{k+1} - \langle c, X_k\rangle\right\}^2\right)}
$$

$$
\tilde{\Lambda}_k = \prod_{l=1}^{k} \tilde{\lambda}_l(X_{l-1}, y_l)
$$

dan definisikan peluang $\tilde{P}$ sehingga $\left.\frac{d\tilde{P}}{dP}\right|_{\mathrm{G}_k} = \tilde{\Lambda}_k$.

Teorema 3.3.6 (Elliott et. al. 1995)

$$
\hat{\sigma}_i(k) = \frac{\hat{\tau}_k^i(y^2) - 2c_i\hat{\tau}_k^i(y) + c_i^2\hat{O}_k^i}{\hat{O}_k^i} = \frac{\gamma_k\left(\tau_k^i(y^2)\right) - 2c_i\gamma_k\left(\tau_k^i(y)\right) + c_i^2\gamma_k\left(O_k^i\right)}{\gamma_k\left(O_k^i\right)} \quad . \qquad (3.11)
$$

Bukti: Dari definisi

$$
\log \tilde{\Lambda}_k = \sum_{l=1}^{k} -\frac{1}{2}\log\langle \hat{\sigma}, X_{l-1}\rangle - \frac{\left(y_l - \langle c, X_{l-1}\rangle\right)^2}{2\langle \hat{\sigma}, X_{l-1}\rangle} + R(c, \sigma)
$$

di mana $R(c,\sigma)$ tidak bergantung pada $\hat{\sigma}$. Karena

$$E\left[\log\tilde{\Lambda}_k \mid \Upsilon_k\right] = E\left[\sum_{l=1}^{k}\sum_{i=1}^{N} -\frac{1}{2}\langle X_{l-1}, e_i\rangle \log\hat{\sigma}_i(k) - \frac{\langle X_{l-1}, e_i\rangle}{2\hat{\sigma}_i(k)} + \left(y_l^2 - 2y_l c_i + c_i^2\right)\middle| \Upsilon_k\right] + \hat{R}(c,\sigma)$$
$$= -\frac{1}{2}\sum_{i=1}^{N}\left\{\hat{O}_k^i \log\hat{\sigma}_i(k) + \frac{1}{\hat{\sigma}_i(k)}\left(\hat{\tau}_k^i(y^2) - 2c_i\hat{\tau}_k^i(y) + c_i^2\hat{O}_k^i\right)\right\} + \hat{R}(c,\sigma),$$

maka

$$\frac{d}{d\hat{\sigma}_i(k)} E\left[\log\tilde{\Lambda}_k \mid \Upsilon_k\right] = \frac{1}{2}\left\{-\frac{\hat{O}_k^i}{\hat{\sigma}_i(k)} + \frac{1}{\hat{\sigma}_i(k)}\left(\hat{\tau}_k^i(y^2) - 2c_i\hat{\tau}_k^i(y) + c_i^2\hat{O}_k^i\right)\right\} = 0$$

memberikan

$$\frac{\hat{O}_k^i}{\hat{\sigma}_i(k)} = \frac{1}{\hat{\sigma}_i^2(k)}\left(\hat{\tau}_k^i(y^2) - 2c_i\hat{\tau}_k^i(y) - c_i^2\hat{O}_k^i\right)$$

atau

$$\hat{\sigma}_i(k) = \frac{\hat{\tau}_k^i(y^2) - 2c_i\hat{\tau}_k^i(y) - c_i^2\hat{O}_k^i}{\hat{O}_k^i} = \frac{\dfrac{\gamma_k(\tau_k^i(y^2))}{\gamma_k(1)} - 2c_i\dfrac{\gamma_k(\tau_k^i(y))}{\gamma_k(1)} - c_i^2\dfrac{\gamma_k(O_k^i)}{\gamma_k(1)}}{\dfrac{\gamma_k(O_k^i)}{\gamma_k(1)}}$$
$$= \frac{\gamma_k(\tau_k^i(y^2)) - 2c_i\gamma_k(\tau_k^i(y)) - c_i^2\gamma_k(O_k^i)}{\gamma_k(O_k^i)}.$$

Catatan 3.3.7 Berdasarkan observasi sampai waktu ke-k, parameter model yang baru yaitu $(\hat{a}_{ji}(k)), 1 \le i, j \le N;\ \hat{c}_i(k), 1 \le i \le N;\ \hat{\sigma}_i(k), 1 \le i \le N$ diberikan oleh persamaan (3.5), (3.10) dan (3.11). Nilai $\gamma_k\left(\mathcal{J}_k^{rs}\right),\ \gamma_k\left(O_k^{rs}\right),\ \gamma_k\left(\tau_k^r(y)\right),\ \gamma_k\left(\tau_k^r(y^2)\right)$ dan $\gamma_k\left(X_k\right)$ kemudian dapat dihitung kembali menggunakan parameter yang baru dan data pengamatan yang baru.

4. ALGORITMA MENDUGA PARAMETER MODEL

Diketahui parameter berbentuk

$$\theta = \left\{(a_{ji}),\ 1 \le i, j \le N;\ c_i,\ 1 \le i \le N;\ \sigma_i, 1 \le i \le N\right\}.$$

Akan ditentukan parameter

$$\hat{\theta}(k) = \left\{(\hat{a}_{ji}(k)),\ 1 \le i, j \le N;\ \hat{c}_i(k),\ 1 \le i \le N;\ \hat{\sigma}_i(k), 1 \le i \le N\right\}$$

yang memaksimumkan *pseudo-loglikelihood* bersyarat seperti pada bagian 3. Algoritma untuk memperoleh parameter tersebut adalah sebagai berikut.

Algoritma untuk menentukan parameter $\hat{\theta}(k)$

Langkah 1: Tetapkan N (banyaknya *state*)
M (banyak data)
Input data $\{y_k\}$.

Langkah 2: Tetapkan Nilai awal

$$\pi = (\pi_i)_{N\times 1}$$
$$A = (a_{ji})_{N\times N}$$
$$c = (c_i)_{N\times 1}$$
$$\sigma = (\sigma_i)_{N\times 1}.$$

Catatan: $\pi = E\left[X_0\right]$ dan memenuhi $A\pi = \pi$

Langkah 3: Lakukan untuk $l = 0$ sampai dengan M

1. Tetapkan

$$a_i = Ae_i$$
$$\gamma_0\left(X_0\right) = \pi$$
$$\gamma_0\left(J_0^{rs}\right) = 0$$
$$\gamma_0\left(O_0^r\right) = 0$$
$$\gamma_0\left(\tau_0(y)\right) = 0$$
$$\gamma_0\left(\tau_0(y^2)\right) = 0.$$

2. Lakukan untuk $k = 0$ sampai dengan $l-1$

a. Hitung penduga rekursif

$$\gamma_{k+1}(X_{k+1}) = \sum_{i=1}^{N}\left\langle \gamma_k(X_k), \Gamma^i(y_{k+1})\right\rangle a_i$$

$$\gamma_{k+1,k+1}(J_{k+1}^{rs}) = \sum_{i=1}^{N}\left\langle \gamma_{k,k}(J_k^{rs}), \Gamma^i(y_{k+1})\right\rangle a_i + \left\langle \gamma_k(X_k), \Gamma^r(y_{k+1})\right\rangle a_{sr}e_s$$

$$\gamma_{k+1}(J_{k+1}^{rs}) = \left\langle \gamma_{k+1,k+1}(J_{k+1}^{rs}), \overline{1}\right\rangle$$

$$\gamma_{k+1,k+1}(O_{k+1}^r) = \sum_{i=1}^{N}\left\langle \gamma_{k,k}(O_k^r), \Gamma^i(y_{k+1})\right\rangle a_i + \left\langle \gamma_k(X_k), \Gamma^r(y_{k+1})\right\rangle a_r$$

$$\gamma_{k+1}(O_{k+1}^r) = \left\langle \gamma_{k+1,k+1}(O_{k+1}^r), \overline{1}\right\rangle$$

$$\gamma_{k+1,k+1}(\tau_{k+1}^r(y)) = \sum_{i=1}^{N}\left\langle \gamma_{k,k}(\tau_k^r(y)), \Gamma^i(y_{k+1})\right\rangle a_i + \left\langle \gamma_k(X_k), \Gamma^r(y_{k+1})\right\rangle y_{k+1}a_r$$

$$\gamma_{k+1}(\tau_{k+1}^r(y)) = \left\langle \gamma_{k+1,k+1}(\tau_{k+1}^r(y)), \overline{1}\right\rangle$$

$$\gamma_{k+1,k+1}(\tau_{k+1}^r(y^2)) = \sum_{i=1}^{N}\left\langle \gamma_{k,k}(\tau_k^r(y^2)), \Gamma^i(y_{k+1})\right\rangle a_i + \left\langle \gamma_k(X_k), \Gamma^r(y_{k+1})\right\rangle y_{k+1}^2 a_r$$

$$\gamma_{k+1}(\tau_{k+1}^r(y^2)) = \left\langle \gamma_{k+1,k+1}(\tau_{k+1}^r(y^2)), \overline{1}\right\rangle.$$

b. Hitung penduga parameter

$$\hat{a}_{sr}(k+1) = \frac{\gamma_{k+1}\left(J_{k+1}^{rs}\right)}{\gamma_{k+1}\left(O_{k+1}^r\right)}$$

$$\hat{c}_r(k+1) = \frac{\gamma_{k+1}\left(\tau_{k+1}^r(y)\right)}{\gamma_{k+1}\left(O_{k+1}^r\right)}$$

$$\hat{\sigma}_i(k+1) = \frac{\gamma_{k+1}\left(\tau_{k+1}^i(y^2)\right) - 2c_i\gamma_{k+1}\left(\tau_{k+1}^i(y)\right) + c_i^2\gamma_{k+1}\left(O_{k+1}^i\right)}{\gamma_{k+1}\left(O_{k+1}^i\right)}.$$

c. Tuliskan

$$\hat{A}(k+1) = \left(\hat{a}_{sr}(k+1)\right)$$
$$\hat{c}(k+1) = \left(\hat{c}_r(k+1)\right)$$
$$\hat{\sigma}(k+1) = \left(\hat{\sigma}_i(k+1)\right).$$

d. Tentukan

$\hat{\pi}(k+1)$ dari persamaan $\hat{A}(k+1)\hat{\pi}(k+1) = \hat{\pi}(k+1)$.

e. Ulangi langkah a sampai dengan d untuk k berikutnya.

3. Beri nilai

$$A \leftarrow \hat{A}(k)$$
$$c \leftarrow \hat{c}(k)$$
$$\sigma \leftarrow \hat{\sigma}(k).$$

3. Ulangi langkah 1 s/d 3 untuk l berikutnya.

Langkah 4: Untuk $k = 1$ sampai dengan M, cetak

$\hat{A}(k)$, $\hat{\pi}$(k), $\hat{c}(k)$, $\hat{\sigma}(k)$, $\gamma_k(X_k)$

DAFTAR PUSTAKA

[1].Baum, L. E. and Petrie, T. 1966. Statistical inference for probabilistic functions of finite state Markov chains. *Annals of the Institute of Statistical Mathematics*, 37:1554-1563.

[2]. Elliot, R. J., Aggoun, L. dan Moore, J. B. 1995. *Hidden Markov models*, Springer Verlag, New York.

[3]. Wong, E and Hajek, B. 1985. *Stochastic Process in Engineering System*. Springer Verlag, Berlin.