PENDUGAAN PARAMETER DAN KEKONVERGENAN PENDUGA PARAMETER MODEL POISSON *HIDDEN* MARKOV

M. FIKRI[1], B. SETIAWATY[2], I G. P. PURNABA[2]

Abstrak

Model *hidden* Markov terdiri dari sepasang proses stokastik yaitu proses observasi dan proses yang memengaruhi observasi. Proses stokastik yang memengaruhi observasi ini diasumsikan membentuk rantai Markov dan tidak diamati. Model Poisson *hidden* Markov (MPHM) adalah salah satu model *hidden* Markov diskret dan proses observasinya jika diketahui proses yang memengaruhinya diasumsikan menyebar Poisson. Salah satu ciri MPHM adalah bersifat *overdispersi*, yaitu ragam data lebih besar dari rataannya. Permasalahan utama MPHM ialah menduga parameter yang memaksimumkan fungsi *likelihood*. Fungsi *likelihood* dihitung menggunakan algoritme *Forward-Backward*. Algoritme *Expectation Maximization* (algoritme EM) digunakan untuk memaksimumkan fungsi *likelihood*. Penduga parameter MPHM yang diperoleh menggunakan algoritme EM konvergen ke titik stasioner dari fungsi *likelihood*.

Kata Kunci: model Poisson *hidden* Markov, *overdispersi*, algoritme *Forward-Backward*, algoritme *Expectation Maximization.*

1. PENDAHULUAN

Terdapat banyak kejadian atau fenomena yang terjadi dalam kehidupan sehari-hari yang bersifat tidak pasti. Ketidakpastian ini dapat dijelaskan dengan proses stokastik. Hal ini dikarenakan proses stokastik merupakan suatu model yang dibangun dengan aturan-aturan peluang. Model ini dapat diterapkan dalam berbagai bidang pada kehidupan sehari-hari seperti nilai tukar rupiah, kedatangan pelanggan ke suatu pusat layanan, dan banyaknya klaim pada suatu perusahaan asuransi.

Ketidakpastian pada suatu fenomena dapat disebabkan oleh beberapa faktor. Faktor-faktor penyebab ini seringkali sulit diamati. Model *hidden* Markov dapat diandalkan untuk memodelkan permasalahan ini.

Model *hidden* Markov terdiri dari sepasang proses stokastik, yaitu proses observasi dan proses yang memengaruhi observasi. Proses stokastik penyebab observasi ini diasumsikan tidak diamati dan membentuk rantai Markov, yaitu

---

[1]Mahasiswa S2, Program Studi Matematika Terapan, Sekolah Pascasarjana IPB. E-mail: miftahulfikri0@gmail.com

[2]Departemen Matematika, Fakultas Ilmu Pengetahuan Alam, Jalan Meranti Kampus IPB Dramaga Bogor,16680.

peluang terjadinya penyebab kejadian pada suatu waktu tertentu hanya bergantung pada penyebab kejadian pada satu satuan waktu sebelumnya. Penyebab observasi ini biasa disebut *state*.

Model Poisson *hidden* Markov (MPHM) adalah salah satu model *hidden* Markov diskret di mana observasi (kejadian yang diamati) jika diketahui penyebab kejadiannya diasumsikan menyebar Poisson. Salah satu ciri MPHM adalah bersifat *overdispersi*, yaitu ragam dari data observasi pada setiap waktu lebih dari rataannya.

Permasalahan utama pada MPHM ialah menduga parameter yang memaksimumkan fungsi *likelihood*. Karya ilmiah ini akan membahas pendugaan parameter MPHM menggunakan algoritme *Expectation Maximization* (EM) berdasarkan karya ilmiah Dempster *et al.* [2], kemudian membahas kekonvergenan parameter MPHM berdasarkan karya ilmiah Wu [6].

2. MODEL POISSON *HIDDEN* MARKOV

2.1 Definisi

Model Poisson *hidden* Markov adalah model dengan waktu diskret yang terdiri dari sepasang proses stokastik $\{X_t, Y_t\}_{t\in\mathbb{N}}$. $\{X_t\}_{t\in\mathbb{N}}$ merupakan penyebab kejadian yang diasumsikan tidak diamati dan membentuk suatu rantai Markov. Sedangkan $\{Y_t\}_{t\in\mathbb{N}}$ adalah proses observasinya yang hanya bergantung pada $\{X_t\}_{t\in\mathbb{N}}$. Kemudian peubah acak Y_t diketahui X_t adalah diasumsikan peubah acak Poisson, untuk setiap $t \in \mathbb{N}$.

2.2 Karakteristik Model Poisson *Hidden* Markov

1. Diasumsikan $\{X_t\}_{t\in\mathbb{N}}$ adalah rantai Markov diskret, homogen dan *ergodic* dengan ruang *state* $S_X = \{1,2,\cdots,m\}$.
2. Matriks peluang *state* transisi $\Gamma = [\gamma_{ij}]$, di mana Γ matriks berukuran $m \times m$ dan $i, j \in S_X$, memenuhi:
   - $\gamma_{ij} = P(X_t = j | X_{t-1} = i) = P(X_2 = j | X_1 = i)$,
   - $\gamma_{ij} \geq 0$,
   - $\sum_{j=1}^{m} \gamma_{ij} = 1$, untuk setiap $i = 1,2,\ldots,m$.

   Dalam model Poisson *hidden* Markov, Saat X_t berada pada *state* i $(i \in S_X)$, maka sebaran bersyarat Y_t jika diketahui $X_t = i$ $(t \in \mathbb{N})$ adalah peubah acak Poisson dengan parameter λ_i. Untuk setiap $y \in \{0,1,2,\ldots\}$, matriks peluang dari proses observasi $\Pi = [\pi_{yi}]$, dengan

$$\pi_{yi} = P(Y_t = y | X_t = i) = e^{-\lambda_i} \frac{{\lambda_i}^y}{y!},$$

$$\sum_{y=0}^{\infty} \pi_{yi} = 1.$$

3. Vektor peluang *state* awal $\delta = [\delta_i]$, di mana δ merupakan vektor berukuran $m \times 1$ dan $i \in S_X$, dengan
$$\delta_i = P(X_1 = i),$$
$$\sum_{i=1}^{m} \delta_i = 1.$$
Karena rantai Markov $\{X_t\}_{t\in\mathbb{N}}$ diasumsikan rantai Markov yang *ergodic*, δ merupakan sebaran yang stasioner sehingga memenuhi
$$\Gamma\delta = \delta. \tag{1}$$
4. Untuk setiap $t \in \mathbb{N}$ dan $y \in \{0,1,2,\ldots\}$, fungsi sebaran marginal dari Y_t, yaitu
$$P(Y_t = y) = \sum_{=i=1}^{m} P(Y_t = y|X_t = i)P(X_t = i) = \sum_{i=1}^{m} \delta_i \pi_{yi}\,.$$
5. Nilai harapan dari Y_t diberikan oleh
$$E(Y_t) = \sum_{y=0}^{\infty} y\, P(Y_t = y) = \sum_{i=1}^{m}\sum_{y=0}^{\infty} y\, P(Y_t = y|X_t = i)P(X_t = i)$$
$$= \sum_{i=1}^{m} E(Y_t|X_t = i)P(X_t = i) = \sum_{i=1}^{m} \lambda_i\delta_i = \delta'\lambda,$$
dengan λ merupakan vektor yang didefinisikan sebagai $(\lambda_1, \ldots, \lambda_m)$.
6. Ragam dari Y_t diberikan oleh
$$\begin{aligned}\mathrm{Var}(Y_t) &= E(Y_t^2) - \big(E(Y_t)\big)^2\\ &= \sum_{i=1}^{m}(\lambda_i^2 + \lambda_i)\delta_i - \left(\sum_{i=1}^{m}\lambda_i\delta_i\right)^2\\ &= \lambda' D\lambda + \delta'\lambda - (\delta'\lambda)^2,\end{aligned}$$
dengan D merupakan $\mathrm{diag}(\delta)$. Berdasarkan [3] dapat ditunjukkan bahwa terjadi *overdispersi*.

Berdasarkan pembahasan karakteristik MPHM di atas, model Poisson *hidden* Markov $\{X_t, Y_t\}_{t\in\mathbb{N}}$ dicirikan oleh parameter $\phi = (\Gamma, \lambda, \delta)$, dengan
$$\delta = [\delta_i] \quad i \in S_X,$$
$$\Gamma = [\gamma_{ij}] \quad i,j \in S_X,$$
$$\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)'.$$

Hal yang sangat penting pada MPHM ialah mengestimasi parameter model. Akan tetapi berdasarkan persamaan (1), δ dengan mudah diperoleh ketika Γ diperoleh, sehingga cukup mengestimasi parameter $\phi = (\Gamma, \lambda)$. Pada tesis ini untuk mengestimasi parameter model dilakukan menggunakan metode maksimum *likelihood*.

3. PENDUGAAN PARAMETER

Misalkan T merupakan banyaknya observasi, m banyaknya *state*, dan $y = (y_1, y_2, \dots, y_T)$ merupakan barisan observasi. Diberikan sebarang $\varepsilon > 0$ cukup kecil mendekati 0. $\boldsymbol{\Phi} = \left\{\phi = (\Gamma, \lambda) : \Gamma \in [0,1]^{\mathrm{m}^2},\ \lambda \in \left[\varepsilon, \frac{1}{\varepsilon}\right]^m\right\}$ merupakan ruang parameter MPHM.

Untuk setiap $\in \boldsymbol{\Phi}$, $\Gamma(\phi) = \left(\gamma_{ij}(\phi)\right)$, $\lambda(\phi) = \left(\lambda_i(\phi)\right)$, $\delta(\phi) = \left(\delta_i(\phi)\right)$, memenuhi asumsi sebagai berikut.

Asumsi Kekontinuan Parameter MPHM

1. $\gamma_{ij}: \boldsymbol{\Phi} \to \mathbb{R}$, dengan $\gamma_{ij}(\phi) = \gamma_{ij}$ merupakan fungsi yang kontinu di $\boldsymbol{\Phi}$, $\forall\ i, j \in S_X$,
2. $\lambda_i: \boldsymbol{\Phi} \to \mathbb{R}$, dengan $\lambda_i(\phi) = \lambda_i$ merupakan fungsi yang kontinu di $\boldsymbol{\Phi}$, $\forall\ i \in S_X$,
3. $\delta_i: \boldsymbol{\Phi} \to \mathbb{R}$, dengan $\delta_i(\phi) = \delta_i$ merupakan fungsi yang kontinu di $\boldsymbol{\Phi}$, $\forall\ i \in S_X$.

Selanjutnya, misalkan:

1. $y = (y_1, y_2, \dots, y_T)$, data dari proses $\{Y_t\}_{t=1}^T$ (biasa disebut data tak lengkap),
2. $x = (i_1, i_2, \dots, i_T)$, penyebab dari y yang tak diamati dan merupakan data dari proses $\{X_t\}_{t=1}^T$,
3. $z = (i_1, y_1, \dots, i_T, y_T) = (x, y)$, data dari proses $\{X_t, Y_t\}_{t=1}^T$ (biasa disebut data lengkap),
4. $Y = \{Y_t\}_{t=1}^T$,
5. $X = \{X_t\}_{t=1}^T$,
6. $Z = \{X_t, Y_t\}_{t=1}^T$,
7. $P(Z = z|\phi) = p(z;\phi) = p(x, y|\phi)$, merupakan fungsi massa peluang dari Z,
8. $P(Y = y|\phi) = p(y|\phi)$ merupakan fungsi massa peluang dari Y,
9. $L_T^c(\phi) = p(z|\phi) = p(x, y|\phi)$, merupakan fungsi *likelihood* dari data lengkap,
10. $P(X = x|Y = y, \phi) = p(x|y, \phi)$merupakan fungsi massa peluang dari X dengan syarat $Y = y$ diberikan, yaitu $p(x|y, \phi) = \frac{p(z|\phi)}{p(y|\phi)} = \frac{p(x,y|\phi)}{p(y|\phi)} = \frac{L_T^c(\phi)}{L_T(\phi)}$.

Fungsi *likelihood* dari proses observasi Y didefinisikan sebagai berikut:

$$\begin{aligned} L_T(\phi) &= P(Y_1 = y_1, Y_2 = y_2, \dots, Y_T = y_T|\phi) \\ &= p(y_1, y_2, \dots, y_T|\phi) \\ &= p(y|\phi) \\ &= \sum_{i_1=1}^{m} \dots \sum_{i_T=1}^{m} \left(\pi_{y_1 i_1}\pi_{y_2 i_2} \dots \pi_{y_T i_T}\right) \times \left(\delta_{i_1}\gamma_{i_1 i_2}\gamma_{i_2 i_3} \dots \gamma_{i_{T-1} i_T}\right) \\ &= \sum_{i_1=1}^{m} \dots \sum_{i_T=1}^{m} \delta_{i_1}\pi_{y_1 i_1} \prod_{t=2}^{T} \gamma_{i_{t-1} i_t}\pi_{y_t i_t}. \end{aligned} \tag{2}$$

Permasalahan utama pada MPHM ialah mencari parameter $\phi^* \in \mathbf{\Phi}$ yang memaksimumkan fungsi *likelihood* $L_T(\phi)$. Untuk banyaknya data observasi T yang cukup besar, menghitung fungsi *likelihood* ini dibutuhkan waktu yang cukup lama. Algoritme *forward-backward* dapat digunakan untuk menangani masalah ini.

3.1 Algoritme *Forward-Backward*

Algoritme *forward-backward* digunakan untuk menghitung peluang terjadinya barisan observasi $(y_1, y_2, \cdots, y_T)$ secara rekursif, hal ini sangat berguna untuk mempercepat waktu komputasi. Algoritme ini terbagi menjadi dua, yaitu algoritme *forward* dan algoritme *backward*.

Baum *et al.* [1] mendefinisikan peluang *forward* sebagai berikut:

$$\alpha_t(i|\phi) = P(Y_1 = y_1, Y_2 = y_2, \dots, Y_t = y_t, X_t = i|\phi),$$

dan peluang *backward*

$$\beta_t(i|\phi) = P(Y_{t+1} = y_{t+1}, \dots, Y_T = y_T|X_t = i, \phi),$$

untuk $t = 1,2, \dots T$, dan $i \in S_X$.

MacDonald dan Zucchini [3] merumuskan peluang *forward* dan peluang *backward* secara rekursif, yang biasa disebut algortime *forward* sebagai berikut:

$$\alpha_1(i|\phi) = \pi_{y_1 i}\delta_i,$$

$$\alpha_{t+1}(j|\phi) = \left(\sum_{i \in S_X} \alpha_t(i|\phi)\gamma_{ij}\right)\pi_{y_{t+1}j},$$

dan algoritme *backward*

$$\beta_T(j|\phi) = 1,$$

$$\beta_t(j|\phi) = \sum_{i \in S_X} \beta_{t+1}(i|\phi)\pi_{y_{t+1}i}\,\gamma_{ji},$$

untuk $t = 1, \dots, T-1$, dan $i, j \in S_X$.

Kemudian MacDonald dan Zucchini [3] menggunakan algoritme *forward* dan algoritme *backward* untuk menghitung fungsi *likelihood* $L_T(\phi)$, yang biasa disebut algoritme *forward-backward* sebagai berikut:

$$L_T(\phi) = \sum_{i \in S_X} \alpha_t(i|\phi)\beta_t(i|\phi),$$

untuk sebarang $t = 1,2, \dots, T$, dan $i \in S_X$.

Fungsi *likelihood* dari data lengkap MPHM adalah sebagai berikut:

$$L_T^c(\phi) = \delta_{i_1}\pi_{y_1 i_1}\prod_{t=2}^{T}\gamma_{i_{t-1}i_t}\pi_{y_t i_t}. \qquad (3)$$

Berdasarkan persamaan (2) dan (3), keterkaitan antara fungsi *likelihood* data tak-lengkap dan data lengkap adalah sebagai berikut:

$$
\begin{aligned}
L_T(\phi) = p(y|\phi) &= \sum_{i_1=1}^{m} \dots \sum_{i_T=1}^{m} \delta_{i_1}\pi_{y_1 i_1} \prod_{t=2}^{T} \gamma_{i_{t-1}i_t}\pi_{y_t i_t} \\
&= \sum_x p(y, x|\phi) \\
&= \sum_x L_T^c(\phi).
\end{aligned}
$$

Untuk memperoleh $\phi^* \in \mathbf{\Phi}$ yang memaksimumkan $L_T(\phi)$ merupakan masalah yang sulit. $\phi^* \in \mathbf{\Phi}$ yang memaksimumkan $\ln L_T(\phi)$ juga akan memaksimumkan $L_T(\phi)$.

Untuk $\phi \in \mathbf{\Phi}$, berlaku

$$
\ln p(x|y, \phi) = \ln \frac{L_T^c(\phi)}{L_T(\phi)} \quad \Longrightarrow \ln L_T(\phi) = \ln L_T^c(\phi) - \ln p(x|y, \phi).
$$

Perhatikan bahwa untuk sebarang $\hat{\phi} \in \mathbf{\Phi}$ juga berlaku

$$
E_{\hat{\phi}}(\ln L_T(\phi)\,|y) = E_{\hat{\phi}}(\ln L_T^c(\phi)\,|y) - E_{\hat{\phi}}(\ln p(x|y, \phi)|y), \tag{4}
$$

dan

$$
\begin{aligned}
E_{\hat{\phi}}(\ln L_T(\phi)\,|y) &= \sum_x \ln L_T(\phi)\, p\big(x|y, \hat{\phi}\big) \\
&= \sum_x \ln p(y|\phi)\, p\big(x|y, \hat{\phi}\big) \\
&= \sum_x \ln p(y|\phi) \frac{p\big(x, y|\hat{\phi}\big)}{p\big(y|\hat{\phi}\big)} \\
&= \frac{\ln p(y|\phi)}{p\big(y|\hat{\phi}\big)} \sum_x p\big(x, y|\hat{\phi}\big) \\
&= \frac{\ln p(y|\phi)}{p\big(y|\hat{\phi}\big)} p\big(y|\hat{\phi}\big) \\
&= \ln p(y|\phi) \\
&= \ln L_T(\phi),
\end{aligned} \tag{5}
$$

sehingga berdasarkan persamaan (4), dan (5) diperoleh

$$
\ln L_T(\phi) = Q\big(\phi|\hat{\phi}\big) - H\big(\phi|\hat{\phi}\big), \tag{6}
$$

dengan $Q\big(\phi|\hat{\phi}\big) = E_{\hat{\phi}}(\ln L_T^c(\phi)|y)$ dan $H\big(\phi|\hat{\phi}\big) = E_{\hat{\phi}}(\ln p(x|y, \phi)\,|y)$.

Langkah pertama untuk memperoleh ϕ^* yang memaksimumkan $\ln L_T(\phi)$ adalah menyelesaikan persamaan $\partial_\phi(\ln L_T(\phi)) = 0$ untuk mendapatkan titik stasioner. Dengan mengikuti pola persamaan (4), dengan mudah diperoleh

$$\partial_\phi(\ln L_T(\phi)) = E_{\hat{\phi}}\big(\partial_\phi(\ln L_T(\phi))|y\big). \tag{7}$$

Akibat persamaan (6) dan (7), maka

$$\begin{aligned}\partial_\phi(\ln L_T(\phi)) &= E_{\hat{\phi}}\big(\partial_\phi(\ln L_T(\phi))|y\big)\\ &= E_{\hat{\phi}}\big(\partial_\phi \ln L_T^c(\phi)\,|y\big) - E_{\hat{\phi}}\big(\partial_\phi \ln p(x|y,\phi)|y\big).\end{aligned} \tag{8}$$

Definisikan

$$D^{10}Q\big(\phi|\hat{\phi}\big) = E_{\hat{\phi}}\left(\frac{\partial}{\partial\phi}\ln L_T^c(\phi)|y\right), \tag{9}$$

dan

$$D^{10}H\big(\phi|\hat{\phi}\big) = E_{\hat{\phi}}\left(\frac{\partial}{\partial\phi}\ln p(x|y,\phi)\,|y\right), \tag{10}$$

sehingga dengan mensubstitusikan persamaan (9) dan (10) ke persamaan (8) akan diperoleh

$$\partial_\phi(\ln L_T(\phi)) = D^{10}Q\big(\phi|\hat{\phi}\big) - D^{10}H\big(\phi|\hat{\phi}\big).$$

Karena $D^{10}H\big(\hat{\phi}|\hat{\phi}\big) = 0$ untuk setiap $\hat{\phi} \in \mathbf{\Phi}$ dan $H\big(\phi|\hat{\phi}\big) \leq H\big(\hat{\phi}|\hat{\phi}\big)$ untuk setiap $\phi, \hat{\phi} \in \mathbf{\Phi}$, sehingga berdasarkan [6] untuk mencari titik stasioner dari $\ln L_T(\phi)$ cukup mencari titik stasioner dari $Q\big(\phi|\hat{\phi}\big)$ terhadap $\phi \in \mathbf{\Phi}$. Akan tetapi $D^{10}Q\big(\phi|\hat{\phi}\big)$ merupakan fungsi tak-linear dan sulit diselesaikan secara eksplisit terhadap parameter $\phi \in \mathbf{\Phi}$, akibatnya untuk memperoleh titik stasioner dari $Q\big(\phi|\hat{\phi}\big)$ terhadap $\phi \in \mathbf{\Phi}$ secara analitik merupakan persoalan yang sulit, sehingga permasalahan ini diselesaikan dengan pendekatan numerik. Pada tesis ini digunakan algoritme *Expectation Maximization*.

3.2 Algoritme *Expectation Maximization*

Algoritme *Expectation Maximization* (EM) adalah algoritme iteratif yang terdiri atas dua langkah pada setiap iterasinya, yaitu langkah E dan langkah M. Langkah-langkah dalam algoritme EM yaitu, ambil $\phi^{(k)}$ sebagai penduga parameter MPHM yang didapat pada iterasi ke-(k). Pada iterasi ke-$(k+1)$ langkah E dan langkah M didefinisikan sebagai berikut:

1. Berikan nilai awal parameter $\phi^{(k)}$ untuk $k = 0$,
2. **Langkah E** – diberikan $\phi^{(k)}$, hitung
   $Q\big(\phi;\phi^{(k)}\big) = E_{\phi^{(k)}}(\ln L_T^c(\phi)|Y = y)$

$$= \sum_{i\in S_X} \frac{\alpha_1(i|\phi^{(k)})\beta_1(i|\phi^{(k)})}{\sum_{l\in S_X}\alpha_t(l\,|\phi^{(k)})\beta_t(l|\phi^{(k)})}\ln\delta_i(\phi)$$
$$+\sum_{i\in S_X}\frac{\sum_{t=1}^{T}\alpha_t(i|\phi^{(k)})\beta_t(i|\phi^{(k)})}{\sum_{l\in S_X}\alpha_t(l|\phi^{(k)})\beta_t(l|\phi^{(k)})}\ln P(Y_t=y_t|X_t=i,\phi)$$
$$+\sum_{i\in S_X}\sum_{j\in S_X}\frac{\sum_{t=1}^{T-1}\gamma_{ij}(\phi^{(k)})\alpha_t(i|\phi^{(k)})P(Y_{t+1}=y_{t+1}|X_{t+1}=j,\phi^{(k)})\beta_{t+1}(j|\phi^{(k)})}{\sum_{l\in S_X}\alpha_t(l|\phi^{(k)})\beta_t(l|\phi^{(k)})}\ln\gamma_{ij}(\phi),$$

3. **Langkah M** – cari $\phi^{(k+1)}$ yang memaksimumkan $Q(\phi;\phi^{(k)})$, sehingga
$$Q(\phi^{(k+1)}|\phi^{(k)})\geq Q(\phi|\phi^{(k)}),$$
untuk setiap $\phi\in\mathbf{\Phi}$,

4. Ganti k dengan $k+1$ dan ulangi langkah 2 sampai langkah 4 hingga $\left|\ln L_T(\phi^{(k+1)})-\ln L_T(\phi^{(k)})\right|$ kurang dari galat yang diinginkan, dengan kata lain $\{\ln L_T(\phi^{(k+1)})\}$ konvergen.

Pada **Langkah M**, untuk memperoleh parameter $\gamma_{ij}(\phi^{(k+1)})$ yang memaksimumkan $Q(\phi|\phi^{(k)})$ terhadap $\phi\in\mathbf{\Phi}$ ialah menggunakan metode pengali Lagrange dengan kendala $\sum_{j=1}^{m}\gamma_{ij}(\phi)=1$, untuk $i=1,2,\dots,m$. Misalkan $G(\phi|\phi^{(k)})=Q(\phi|\phi^{(k)})-\sum_{i=1}^{m}\theta_i\left(\sum_{\forall j}\gamma_{ij}(\phi)-1\right)$, untuk sebarang $\theta_i\in\mathbb{R}$. Maka menurut Paroli dan Spezia [5], $\frac{\partial G(\phi|\phi^{(k)})}{\partial\gamma_{ij}(\phi)}=0$ mengimplikasikan

$$\gamma_{ij}(\phi^{(k+1)})=\frac{\sum_{t=1}^{T-1}\gamma_{ij}(\phi^{(k)})\alpha_t(i|\phi^{(k)})P(Y_{t+1}=y_{t+1}|X_{t+1}=j,\phi^{(k)})\beta_{t+1}(j|\phi^{(k)})}{\sum_{t=1}^{T-1}\alpha_t(i|\phi^{(k)})\beta_t(i|\phi^{(k)})}.$$

Kemudian, untuk memperoleh parameter $\lambda_i(\phi^{(k+1)})$ yang memaksimumkan $Q(\phi|\phi^{(k)})$ terhadap $\phi\in\mathbf{\Phi}$ dari persamaan

$$\frac{\partial Q(\phi|\phi^{(k)})}{\partial\lambda_i(\phi)}=0,$$

berdasarkan [5] diperoleh penduga λ_i ke $k+1$ adalah

$$\lambda_i(\phi^{(k+1)})=\frac{\sum_{t=1}^{T}\alpha_t(i|\phi^{(k)})\,\beta_t(i|\phi^{(k)})\,y_t}{\sum_{t=1}^{T}\alpha_t(i|\phi^{(k)})\,\beta_t(i|\phi^{(k)})}.$$

4. KEKONVERGENAN PENDUGA PARAMETER

Akan dibuktikan barisan $\{\ln L_T(\phi^{(k)})\}$ konvergen ke $\ln L_T(\phi^*)$ menggunakan algoritme EM, dengan $\phi^{(k)}$ merupakan penduga parameter MPHM pada iterasi ke$-k$ dan ϕ^* merupakan titik stasioner dari fungsi $\ln L_T(\phi)$. Hal ini akan dibahas pada Teorema Wu. Sebelum membahas Teorema Wu, perlu dibahas kondisi Wu untuk kasus MPHM yang dibuktikan oleh Paroli *et al*. [4] sebagai berikut.

Kondisi Wu
Misalkan T merupakan banyaknya observasi, dan m banyaknya *state.* Diberikan sebarang $\varepsilon > 0$ cukup kecil mendekati 0. $\mathbf{\Phi} = \left\{\phi = (\Gamma, \lambda) : \Gamma \in [0,1]^{\mathrm{m}^2},\ \lambda \in \left[\varepsilon, \frac{1}{\varepsilon}\right]^m\right\}$ merupakan ruang parameter MPHM. Maka 4 kondisi berikut terpenuhi.

1. $\mathbf{\Phi}$ adalah himpunan bagian terbatas dari $\mathbb{R}^{m^2+m}$,
2. $\ln L_T(\phi)$ adalah kontinu pada $\mathbf{\Phi}$ dan terdiferensialkan di dalam interior $\mathbf{\Phi}$,
3. $\mathbf{\Phi}_{\phi^{(0)}} = \{\phi \in \mathbf{\Phi} : \ln L_T(\phi) \geq \ln L_T(\phi^{(0)})\}$ adalah himpunan kompak, untuk sebarang $\phi^{(0)} \in \mathbf{\Phi}$, dengan $\ln L_T(\phi^{(0)}) > -\infty$,
4. $Q(\varphi|\phi)$ adalah fungsi kontinu terhadap φ dan ϕ pada $\mathbf{\Phi} \times \mathbf{\Phi}$.

Teorema Wu
Misalkan $Q(\varphi|\phi)$ fungsi yang kontinu terhadap φ, ϕ di $\mathbf{\Phi} \times \mathbf{\Phi}$. Misalkan $\{\phi^{(k)}\}$ adalah barisan penduga MPHM yang diperoleh menggunakan algoritme EM. Maka,
1. Jika $\lim_{k\to\infty} \phi^{(k)} = \phi^*$, maka ϕ^* adalah titik stasioner dari fungsi $\ln L_T(\phi)$,
2. $\lim_{k\to\infty} \ln L_T(\phi^{(k)}) = \ln L_T(\phi^*)$, di mana kekonvergenannya monoton naik.

Berdasarkan Teorema Wu di atas, kekonvergenan fungsi *likelihood* yang diperoleh hanya akan menuju titik stasioner dari fungsi *likelihood* tersebut. Sehingga menentukan nilai awal penduga parameter MPHM pada algoritme EM merupakan hal yang sangat penting.

5. SIMPULAN

MPHM yang diasumsikan rantai Markovnya homogen dan *ergodic* serta memenuhi Asumsi Kekontinuan Parameter, maka
1. Pendugaan parameter MPHM menggunakan algoritme EM menghasilkan rumusan yang memaksimumkan fungsi *likelihood,*
2. Penduga parameter yang diperoleh konvergen ke titik stasioner dari fungsi *likelihood* tersebut.

DAFTAR PUSTAKA

[1] Baum LE, Petrie T, Soules G, Weiss N. 1970. A maximization technique occuring in the statistical estimation for probabilistic functions of Markov chains. *The Annals of Statistics*. 41: 164-171.

[2] Dempster AP, Laird NM, Rubin DB. 1977. Maximum likelihood from incomplete data via the EM algorithm*. Journal of the Royal Statistical Society.* 39: 1-38.

[3] MacDonald IL, Zucchini W. 1997. *Hidden Markov and Other Models for Discrete- Valued Time Series*. Chapman and Hall, London.

[4] Paroli R, Redaelli G, Spezia L. 2000. Poisson hidden Markov models for time series of overdispersed insurance counts. *Astin Colloquium*. 461-474.

[5] Paroli R, Spezia L. 1999. Gaussian hidden Markov models: parameters estimation and applications to air pollution data. [E.P.n. 94]. Milan (IT): Universita Cattolica Del Sacro Cuore.

[6] Wu CFJ. 1983. On the convergence properties of the EM algorithm. *The Annals of Statistics*. 11: 95-103.