

BAB III

MARKOV SWITCHING AUTOREGRESSIVE (MSAR)

3.1 Model Markov Switching Autoregressive

Model runtun waktu *Markov Switching Autoregressive* adalah salah satu model runtun waktu yang merupakan perluasan dari model *Autoregressive* (AR). Ide dasar dari model ini adalah membuat model yang dinamis seiring dengan terjadinya perubahan pola data. Perubahan pola data dianggap dipengaruhi oleh peubah acak diskrit tak teramati S_t yang biasa disebut *state* atau *regime*, dimana peubah acak diskrit S_t diasumsikan mengikuti rantai Markov orde pertama. Suatu rantai Markov dikatakan berorde satu jika nilai suatu *state* pada periode tertentu hanya bergantung pada *state* satu periode sebelumnya. Sebagai contoh, pada kasus hanya dua *regime*, S_t dapat diasumsikan memiliki nilai 1 dan 2, sehingga model dengan AR(1) adalah sebagai berikut

$$y_t = \begin{cases} \phi_{0,1} + \phi_{1,1}y_{t-1} + \varepsilon_t & , \text{jika } S_t=1 \\ \phi_{0,2} + \phi_{1,2}y_{t-1} + \varepsilon_t & , \text{jika } S_t=2 \end{cases}$$

yang bisa disederhanakan sebagai berikut

$$y_t = \phi_{0,S_t} + \phi_{1,S_t}y_{t-1} + \varepsilon_t.$$

Karena *regime* saat ini yaitu S_t hanya tergantung pada *regime* satu periode sebelumnya S_{t-1} . Oleh karena itu, model dilengkapi dengan mendefinisikan peluang transisi dari satu *state* ke *state* lainnya,

$$P(S_t = 1 | S_{t-1} = 1) = p_{11}$$

$$P(S_t = 2 | S_{t-1} = 1) = p_{12}$$

$$P(S_t = 1 | S_{t-1} = 2) = p_{21}$$

$$P(S_t = 2 | S_{t-1} = 2) = p_{22}$$

koleksi peluang transisi dapat dituliskan dalam matriks transisi $\mathbf{P}$

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix}$$

dimana $p_{11} + p_{12} = 1$ dan $p_{21} + p_{22} = 1$.

Peluang tak bersyarat proses berada pada tiap *regime* yaitu $P(S_t = i)$ untuk $i = 1, 2$, berdasarkan persamaan (2.35) dan (2.36) untuk model *Markov Switching* dua *state* memiliki peluang tak bersyarat sebagai berikut:

$$P(S_t = 1) = \frac{1 - p_{22}}{2 - p_{11} - p_{22}}, \quad (3.1)$$

$$P(S_t = 2) = \frac{1 - p_{11}}{2 - p_{11} - p_{22}}. \quad (3.2)$$

Multiple Regimes

Markov switching Autoregressive dengan m *regimes* di definisikan sebagai berikut

$$y_t = \phi_{0,S_t} + \phi_{1,S_t}y_{t-1} + \dots + \phi_{p,S_t}y_{t-p} + \varepsilon_t, \quad \text{jika } S_t = j \quad (3.3)$$

dimana $\varepsilon_t \sim i.i.d. N(0, \sigma^2)$

$j = 1, \dots, m$, dengan peluang transisi

$$p_{ij} = P(S_t = j | S_{t-1} = i), \quad i, j = 1, \dots, m,$$

$p_{ij} \geq 0$ untuk $i, j = 1, \dots, m$ dan $\sum_{j=1}^m p_{ij} = 1$ untuk semua $i = 1, \dots, m$

Langkah-langkah untuk membentuk model *Markov Switching Autoregressive* yaitu:

1. Memodelkan data dengan proses AR yang sesuai
 - a. Membuat plot runtun waktu untuk melihat kestasioneran data. Apabila data belum stasioner lakukan penyeselisihan (*differencing*) dan atau transformasi data.
 - b. Menentukan model AR berdasarkan plot fak dan fakp

- c. Mengestimasi parameter dari model AR yang dipilih
2. Uji Perubahan Struktur
3. Estimasi parameter dari model *MarkovSwitching Autoregressive*
4. Uji nonlinearitas
5. Gunakan model untuk tujuan peramalan.

3.2 Estimasi Parameter

Parameter dalam model *Markov Switching* dapat diestimasi menggunakan teknik maksimum likelihood. Namun, dikarenakan kenyataan bahwa proses Markov S_t tidak terobservasi, masalah estimasinya tidaklah standar. Tujuan dari prosedur estimasi tidak hanya untuk mendapatkan estimasi parameter dalam model autoregressive dalam *state* yang berbeda dan peluang transisi dari satu *state* ke *state* yang lain, tetapi juga untuk mendapatkan perkiraan *state* yang terjadi pada tiap titik *sample*.

Markov Switching 2 regime

$$y_t = \begin{cases} \phi_{0,1} + \phi_{1,1}y_{t-1} + \dots + \phi_{p,1}y_{t-p} + \varepsilon_t & \text{jika } S_t = 1 \\ \phi_{0,2} + \phi_{1,2}y_{t-1} + \dots + \phi_{p,2}y_{t-p} + \varepsilon_t & \text{jika } S_t = 2 \end{cases} \quad (3.4)$$

dapat disederhanakan

$$y_t = \phi_{0,s_t} + \phi_{1,s_t}y_{t-1} + \dots + \phi_{p,s_t}y_{t-p} + \varepsilon_t \quad (3.5)$$

di bawah asumsi ε_t berdistribusi normal, fungsi kepadatan peluang dari y_t dengan syarat S_t dan Ω_{t-1} adalah berdistribusi normal dengan rata-rata $\phi_{0,s_t} + \phi_{1,s_t}y_{t-1} + \dots + \phi_{p,s_t}y_{t-p}$ dan variansi σ^2 ,

$$f(y_t | S_t = j, \Omega_{t-1}; \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(y_t - \phi'_j x_t)^2}{2\sigma^2} \right\} \quad (3.6)$$

dimana $x_t = (1, y_{t-1}, \dots, y_{t-p})'$, $\phi_j = (\phi_{0,j}, \phi_{1,j}, \dots, \phi_{p,j})'$ untuk $j = 1, 2$ dan θ adalah vektor yang berisi semua parameter dalam model, $\theta = (\phi'_1, \phi'_2, p_{11}, p_{22},$

$\sigma^2)$, Ω_{t-1} menotasikan himpunan observasi sebelumnya $\Omega_{t-1} = \{y_{t-1}, y_{t-2}, \dots, y_{t-p}\}$. p_{11} dan p_{22} sepenuhnya menjelaskan semua peluang transisi, sebagai contoh, $p_{12} = 1 - p_{11}$.

Log likelihood bersyarat untuk observasi ke t , $l_t(\theta)$ adalah log dari fungsi kepadatan peluang dari y_t dengan syarat Ω_{t-1} yaitu $l_t(\theta) = \ln f(y_t | \Omega_{t-1}; \theta)$. Fungsi kepadatan dari $f(y_t | \Omega_{t-1}; \theta)$ diperoleh dari fungsi kepadatan bersama dari y_t dan S_t sebagai berikut (Hamilton, 1994)

$$\begin{aligned} f(y_t | \Omega_{t-1}; \theta) &= f(y_t, S_t = 1 | \Omega_{t-1}; \theta) + f(y_t, S_t = 2 | \Omega_{t-1}; \theta) \\ &= \sum_{j=1}^2 f(y_t, S_t = j | \Omega_{t-1}; \theta) \\ &= \sum_{j=1}^2 f(y_t | S_t = j, \Omega_{t-1}; \theta) \cdot P(S_t = j | \Omega_{t-1}; \theta) \end{aligned} \quad (3.7)$$

persamaan kedua pada (3.7) mengikuti hukum dasar dari peluang bersyarat, yang menyatakan bahwa peluang gabungan dari dua kejadian A dan B yaitu $P(A \text{ dan } B) = P(A|B) \cdot P(B)$ (Van Djick, 2003).

Untuk dapat menghitung fungsi kepadatan peluang $f(y_t | \Omega_{t-1}; \theta)$ pada persamaan (3.7) perlu dihitung peluang *regime* dengan syarat sejarah dari proses Ω_{t-1} yaitu $P(S_t = j | \Omega_{t-1}; \theta)$. Pada kenyataannya untuk mengembangkan estimasi maksimum *likelihood* dari parameter pada model, tiga estimasi berbeda dari peluang pada tiap *regime* pada saat t yang dibutuhkan yaitu estimasi dari peluang proses berada pada *regime* j pada saat t dengan syarat semua observasi sampai $t - 1$ atau $\hat{\xi}_{t|t-1}$, dengan syarat semua observasi sampai t atau $\hat{\xi}_{t|t}$, dengan syarat semua observasi di seluruh sampel atau $\hat{\xi}_{t|n}$, secara berurutan estimasi ini disebut *forecast*, *inference* dan *smooth inference* dari peluang *regime*

Jika *regime* yang terjadi pada saat $t - 1$ diketahui termasuk himpunan informasi Ω_{t-1} , peramalan optimal dari peluang *regime* secara sederhana sama dengan peluang transisi dari proses Markov S_t . Secara formal dituliskan sebagai berikut

$$\hat{\xi}_{t|t-1} = P \cdot \xi_{t-1} \quad (3.8)$$

$\hat{\xi}_{t|t-1}$ menotasikan vektor peluang bersyarat berukuran 2×1 yaitu $\hat{\xi}_{t|t-1} = (P(S_t = 1|\Omega_{t-1}; \theta), P(S_t = 2|\Omega_{t-1}; \theta))'$, $\xi_{t-1} = (1, 0)'$ jika $S_t = 1$ dan $\xi_{t-1} = (0, 1)'$ jika $S_t = 2$ dan $\mathbf{P}$ adalah matriks yang berisi peluang transisi,

$$\mathbf{P} = \begin{pmatrix} p_{11} & 1 - p_{22} \\ 1 - p_{11} & p_{22} \end{pmatrix} \quad (3.9)$$

Pada praktiknya *regime* pada saat $t - 1$ tidak diketahui, oleh karena itu disebut tak teramati. Salah satu cara terbaik adalah dengan mengganti ξ_{t-1} pada (3.8) dengan sebuah estimasi peluang dari masing-masing *regime* yang terjadi pada saat $t - 1$ dengan syarat observasi sampai $t - 1$. Notasikan vektor berukuran 2×1 berisi inferensi optimal peluang *regime* yang terjadi sebagai $\hat{\xi}_{t-1|t-1}$. Berikan nilai awal $\hat{\xi}_{1|0}$ dan nilai parameter di dalam θ , dapat dihitung peramalan optimal dan inferensi dari peluang *regime* bersyarat adalah dengan mengiterasi persamaan berikut

$$\hat{\xi}_{t|t} = \frac{\hat{\xi}_{t|t-1} \odot f_t}{\mathbf{1}'(\hat{\xi}_{t|t-1} \odot f_t)} \quad (3.10)$$

$$\hat{\xi}_{t+1|t} = P \cdot \hat{\xi}_{t|t} \quad (3.11)$$

untuk $t = 1, \dots, n$, dimana f_t menotasikan vektor yang berisi fungsi kepadatan peluang pada persamaan (3.6) untuk dua *regime*, $\mathbf{1}$ adalah vektor berukuran 2×1 dari satu dan simbol $\odot$ ialah perkalian elemen per elemen. Nilai awal yang diperlukan $\hat{\xi}_{1|0}$ dapat di ambil dari vektor yang berisi konstanta dengan jumlah satu (Hamilton, 1994).

Misalkan $\hat{\xi}_{t|n}$ menotasikan vektor yang berisi inferensi pemulusan pada peluang *regime* yaitu estimasi dari peluang bahwa *regime* j terjadi pada saat t dengan syarat semua observasi yang mungkin, $P(S_t = j | \Omega_n; \theta)$. Kim (1993) mengembangkan sebuah algoritma untuk mendapatkan peluang *regime* ini dari peluang bersyarat $\hat{\xi}_{t|t}$ dan $\hat{\xi}_{t+1|t}$

Inferensi pemulusan peluang *regime* pada saat t dihitung

$$\hat{\xi}_{t|n} = \hat{\xi}_{t|t} \odot (P'[\hat{\xi}_{t+1|n} \div \hat{\xi}_{t+1|t}]) \quad (3.12)$$

dimana $\div$ adalah pembagian elemen per elemen. Algoritma berjalan mundur melalui sampel yang dimulai dengan $\hat{\xi}_{n|n}$ dari (3.10) yang diterapkan pada persamaan (3.12) untuk $t = n - 1, n - 2, \dots, 1$. Secara detail Kim(1993) dan Hamilton(1994), memenuhi rumus (3.12) (Lampiran 2).

Estimasi maksimum likelihood dari peluang transisi adalah sebagai berikut (Hamilton, 1990)

$$\hat{p}_{ij} = \frac{\sum_{t=2}^n P(S_t = j, S_{t-1} = i | \Omega_n; \hat{\theta})}{\sum_{t=2}^n P(S_{t-1} = i | \Omega_n; \hat{\theta})}, \quad (3.13)$$

dimana $\hat{\theta}$ menotasikan estimasi maksimum likelihood dari θ . Untuk mendapatkan nilai estimasi dari $\hat{\phi}_j$ dan $\hat{\sigma}^2$ harus memenuhi persamaan berikut, bukti lihat Hamilton (1990).

$$\sum_{t=1}^n (y_t - \hat{\phi}'_j x_t) x_t P(S_t = j | \Omega_n; \hat{\theta}) = 0, \quad j = 1, 2, \quad (3.14)$$

dan

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n \sum_{j=1}^2 (y_t - \hat{\phi}'_j x_t)^2 x_t P(S_t = j | \Omega_n; \hat{\theta}) \quad (3.15)$$

$\hat{\phi}_j$ adalah estimasi kuadrat terkecil tertimbang dari regresi y_t pada x_t , dengan bobot diberikan oleh akar dari peluang pemulusan dari terjadinya *regime* j . Oleh karena itu, estimasi $\hat{\phi}_j$ dapat diperoleh dengan

$$\hat{\phi}_j = \left(\sum_{t=1}^n x_t(j) x_t(j)' \right)^{-1} \left(\sum_{t=1}^n x_t(j) y_t(j) \right) \quad (3.16)$$

dimana

$$y_t(j) = y_t \sqrt{P(S_t = j | \Omega_n; \hat{\theta})}$$

$$x_t(j) = x_t \sqrt{P(S_t = j | \Omega_n; \hat{\theta})}$$

Estimasi Maksimum Likelihood dari variansi residual diperoleh dengan menggunakan (3.15) sebagai rata-rata residual kuadrat dari dua regresi kuadrat terkecil tertimbang.

Tempatkan semua elemen bersama-sama, prosedur iteratif untuk mengestimasi parameter dari model *Markov Switching*. Berikan nilai awal untuk vektor parameter $\hat{\theta}^{(0)}$, pertama hitung peluang pemulusan *regime* (*smoothed regime probabilities*). Selanjutnya peluang pemulusan $\hat{\xi}_{t|n}$ dikombinasikan dengan estimasi awal dari peluang transisi $\hat{p}_{ij}^{(0)}$ untuk mendapatkan estimasi baru dari peluang transisi $\hat{p}_{ij}^{(1)}$ dari (3.13). Akhirnya, (3.16) dan (3.15) dapat digunakan untuk mendapatkan estimasi baru dari parameter *autoregressive* dan variansi residual. Kombinasikan dengan estimasi baru dari peluang transisi, ini memberikan estimasi baru untuk semua parameter dalam model, $\hat{\theta}^{(1)}$. Iterasi prosedur ini menghasilkan estimasi $\hat{\theta}^{(2)}, \hat{\theta}^{(3)}, \dots$ dan prosedur ini berlanjut sampai konvergensi terjadi, yaitu sampai estimasi di iterasi berikutnya cukup dekat. Prosedur ini merupakan aplikasi dari algoritma *Expectation Maximization* (EM) yang dikembangkan oleh Dempster, Laird dan Rubin (1977). Setiap iterasi dari prosedur ini meningkatkan nilai dari

fungsi likelihood, yang menjamin bahwa estimasi akhirnya adalah estimasi Maksimum Likelihood.

Untuk selanjutnya perhitungan akan dilakukan dengan menggunakan *software* Oxmetrics dan Eviews.

3.3 Uji Nonlinearitas

Persamaan (3.3) sebelumnya dapat dituliskan sebagai

$$y_t = \phi_{0,S_t} + \phi_{1,S_t}y_{t-1} + \dots + \phi_{p,S_t}y_{t-p} + \varepsilon_t, \text{ jika } S_t = j$$

Pendekatan alami dari uji nonlinearitas adalah mengambil model linear sebagai hipotesis nol dan model *regime switching* sebagai hipotesis alternatif. Pada kasus model 2 regime, hipotesis nol dapat dinyatakan sebagai kesamaan parameter *autoregressive* pada kedua *regime* yaitu

$$H_0 = \phi_1 = \phi_2 \text{ (model linear), dan}$$

$$H_1 = \text{terdapat minimal satu } \phi_{i,1} \neq \phi_{i,2} \text{ } i = 1, \dots, p \text{ (model nonlinear)}$$

ϕ_1, ϕ_2 merepresentasikan koefisien AR dari model *Markov switching* pada kedua *regime*

Tidak semua parameter dalam model tercakup dibawah hipotesis nol, dalam model nonlinear *markov switching* terdapat parameter gangguan atau *nuisance parameter* yang tidak teridentifikasi dibawah asumsi hipotesis nol dan tidak hadir dalam model linear, dalam hal ini adalah p_{ij} atau peluang transisi. Sebagai konsekuensinya teori statistik konvensional tidak dapat diterapkan untuk memperoleh distribusi dari tes statistik hal ini dijelaskan (Davies 1977, 1987) (Garcia 1995) dan (Hansen 1996). Tes statistiknya cenderung memiliki distribusi yang tidak standar, yang seringkali ekspresi analisisnya tidak tersedia. Hanya setelah melakukan estimasi *markov switching* dapat dilihat hubungan atau relevansi antara model Markov Switching dan model linear AR.

Ketika mengkaji relevansi dari model MSW, pendekatan alami nya adalah menggunakan Likelihood Ratio (LR) statistik, yang menguji hipotesis nol model

linear melawan hipotesis alternative yaitu model MSW, $H_0 = \varphi_1 = \varphi_2$ diuji dengan menggunakan tes statistik berikut

$$LR_{MSW} = \mathcal{L}_{MSW} - \mathcal{L}_{AR}$$

dimana $\mathcal{L}_{MSW}$ dan $\mathcal{L}_{AR}$ adalah nilai dari fungsi log likelihood yang berhibungan dengan MSW dan model AR. Dalam kasus dua *regime*, parameter p_{11} dan p_{22} yang mendefinisikan peluang transisi dalam model MSW adalah parameter gangguan tak teridentifikasi di hipotesis nol. LR statistik memiliki distribusi yang tidak standar yang tidak dapat dikarakterisasi secara analisis. Nilai kritis (*critical value*) untuk menentukan signifikansi dari uji statistik harus ditentukan dengan cara simulasi. Struktur dasar dari percobaan simulasi tersebut salah satunya menghasilkan sejumlah besar *time series artificialy_t^{*}* menurut model dibawah hipotesis nol. Kemudian, estimasi antara AR dan model MSW untuk masing-masing artificial time series dan hitung LR statistik LR_{MSW}^* . Tes statistik ini dapat digunakan untuk mendapatkan sebuah perkiraan dari distribusi lengkap dari tes statistik dibawah hipotesis nol, atau sekedar untuk menghitung p-value dari LR statistik untuk time series yang asli. Estimasi model MSW bisa memakan waktu dan prosedurnya menuntut jumlah dalam komputasi (Van Dijk, 2003). Perhitungan selanjutnya menggunakan Pcgive 14 Oxmetrics.

3.4 Peramalan

Peramalan adalah salah satu tujuan utama dari pemodelan runtun waktu, hasil peramalan yang baik jika nilai estimasi tidak jauh dari nilai sebenarnya.

Peramalan untuk nilai dari runtun waktu y_{t+h} adalah peramalan dari y_{t+h} dengan syarat *regime* pada saat $t + h$, yaitu S_{t+h} dikali dengan peramalan dari peluang pada masing-masing *regime* pada saat $t + h$. Sebagai contoh, peramalan satu langkah ke depan untuk model *Markov Switching Autoregressive* dengan banyaknya *state* sebanyak dua adalah sebagai berikut:

$$\hat{y}_{t+1|t} = E[y_{t+1} | S_{t+1} = 1, \Omega_t] \cdot P(S_{t+1} = 1 | \Omega_t; \theta)$$

$$\times E[y_{t+1}|S_{t+1} = 2, \Omega_t] \cdot P(S_{t+1} = 2|\Omega_t; \theta).$$

Peramalan dari y_{t+1} dengan syarat *regime* pada saat $t + 1$ mengikuti langsung dari persamaan (3.5), contoh untuk AR(1)

$$E[y_{t+1}|S_{t+1} = j, \Omega_t] = \phi_{0,j} + \phi_{1,j}y_t,$$

dimana

$P(S_{t+1} = j|\Omega_t; \theta), j = 1, 2$ adalah peramalan optimal peluang *regime* yaitu $\hat{\xi}_{t+1|t}$

Matriks peluang transisi m periode kedepan untuk rantai Markov 2 *state*

$$\begin{aligned} \mathbf{P}^m &= \begin{bmatrix} \frac{1-p_{22}}{2-p_{11}-p_{22}} & -1 \\ \frac{1-p_{11}}{2-p_{11}-p_{22}} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2^m \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \frac{-(1-p_{11})}{2-p_{11}-p_{22}} & \frac{1-p_{22}}{2-p_{11}-p_{22}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{(1-p_{22}) + \lambda_2^m(1-p_{11})}{2-p_{11}-p_{22}} & \frac{(1-p_{22}) - \lambda_2^m(1-p_{22})}{2-p_{11}-p_{22}} \\ \frac{(1-p_{11}) - \lambda_2^m(1-p_{11})}{2-p_{11}-p_{22}} & \frac{(1-p_{11}) + \lambda_2^m(1-p_{22})}{2-p_{11}-p_{22}} \end{bmatrix} \end{aligned}$$

sebagai contoh jika proses berada pada *state* 1, peluang m periode berikutnya akan ada di *state* 2 adalah

$$P\{S_{t+m} = 2|S_t = 1\} = \frac{(1-p_{11}) - \lambda_2^m(1-p_{11})}{2-p_{11}-p_{22}},$$

dimana $\lambda_2 = -1 + p_{11} + p_{22}$ (lampiran 3)