Kode / Nama Rumpun : 121 / Matematika

LAPORAN AKHIR
PENELITIAN FUNDAMENTAL

METODE REVERSIBLE JUMP MARKOV CHAIN MONTE CARLO UNTUK ESTIMASI BAYESIAN DALAM MODEL REGRESI LINEAR PER POTONGAN

Tahun ke 1 dari rencana 2 tahun

TIM PENGUSUL

Ketua	: Dr. Suparman, M.Si., DEA	NIDN : 0517046902
Anggota	: Drs. Abdul Taram, M.Si	NIDN : 0505035801

UNIVERSITAS AHMAD DAHLAN
NOVEMBER 2014

Kode / Nama Rumpun : 121 / Matematika

LAPORAN AKHIR
PENELITIAN FUNDAMENTAL

METODE REVERSIBLE JUMP MARKOV CHAIN MONTE CARLO UNTUK ESTIMASI BAYESIAN DALAM MODEL REGRESI LINEAR PER POTONGAN

Tahun ke 1 dari rencana 2 tahun

TIM PENGUSUL

Ketua	**: Dr. Suparman, M.Si., DEA**	**NIDN : 0517046902**
Anggota	**: Drs. Abdul Taram, M.Si**	**NIDN : 0505035801**

UNIVERSITAS AHMAD DAHLAN
NOVEMBER 2014

HALAMAN PENGESAHAN
PENELITIAN FUNDAMENTAL

Judul Kegiatan	:	Metode Reversible Jump Markov Chain Monte Carlo untuk Estimasi Bayesian dalam Model Regresi Linear per Potongan.
Kode/Nama Rumpun Ilmu	:	121 / Matematika
Ketua Peneliti		
A. Nama Lengkap	:	SUPARMAN
B. NIDN	:	0517046902
C. Jabatan Fungsional	:	Lektor
D. Program Studi	:	Pendidikan Matematika
E. Nomor HP	:	081328201198
F. Surel (e-mail)	:	suparmancict@yahoo.co.id
Anggota Peneliti (1)		
A. Nama Lengkap	:	Drs. ABDUL TARAM M.Si.
B. NIDN	:	0505035801
C. Perguruan Tinggi	:	Universitas Ahmad Dahlan
Lama Penelitian Keseluruhan	:	2 Tahun
Penelitian Tahun ke	:	1
Biaya Penelitian Keseluruhan	:	Rp 124.584.000,00
Biaya Tahun Berjalan	:	- diusulkan ke DIKTI Rp 62.160.000,00
		- dana internal PT Rp 0,00
		- dana institusi lain Rp 0,00
		- inkind sebutkan Rp 4.800.000,00

Mengetahui,
Dekan FKIP UAD

(Dra. Trikinasih Handayani, M.Si)
NIP/NIK 195909071985032002

Yogyakarta, 30 - 10 - 2014,
Ketua Peneliti,

(SUPARMAN)
NIP/NIK 60110621

Menyetujui,
Kepala LPP UAD

(Dr. Widodo, M.Si)
NIP/NIK 196002211987091001

RINGKASAN

Model regresi linear per potongan merupakan model yang sangat berguna dalam ekonometrik. Jika data ekonometrik dicocokkan terhadap model regresi linear per potongan maka banyaknya titik ambang, koefisien regresi tiap potongan dan variansi galat tiap potongan umumnya tidak diketahui. Penelitian ini bertujuan untuk mengestimasi banyaknya titik ambang, koefisien regresi tiap potongan dan variansi galat tiap potongan.

Estimasi banyaknya titik ambang, koefisien regresi tiap potongan dan variansi galat tiap potongan dilakukan dalam kerangka Bayesian. Mula-mula distribusi prior untuk banyaknya ambang, koefisien regresi tiap potongan, dan variansi galat tiap potongan ditetapkan. Kemudian, distribusi prior ini dikombinasikan dengan fungsi kemungkinan maksimum dari data untuk mendapatkan distribusi posterior. Berdasarkan pada distribusi posterior ini, penaksir Bayes untuk banyaknya titik ambang, koefisien regresi tiap potongan dan variansi galat tiap potongan dihitung. Akan tetapi, penaksir Bayes tidak dapat ditentukan secara eksak atau analitis. Untuk menyelesaikan masalah ini, digunakan metode *reversible jump* MCMC (Monte Carlo Markov Chain).

Metode reversible jump MCMC merupakan salah satu metode baru yang dapat digunakan untuk mengestimasi parameter model regresi linear per potongan meskipun banyaknya titik ambang tidak diketahui. Keunggulan dari metode ini adalah estimasi banyaknya titik ambang dan estimasi parameter regresi tiap potongan dapat diestimasi secara bersama-sama.

PRAKATA

Puji syukur kepada Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya sehingga pelaksanaan kegiatan Penelitian skema Penelitian Fundamental yang kami laksanakan dapat berjalan dengan baik hingga tersusun laporan akhir penelitian ini.

Kami menyadari bahwa kegiatan penelitian dengan judul " Metode Reversible Jump Markov Chain Monte Carlo untuk Estimasi Bayesian dalam Model Regresi Linear per Potongan." yang kami laksanakan ini, tidak akan dapat berjalan baik tanpa dukungan LPP UAD sebagai kepanjangan dari Direktorat Penelitian dan Pengabdian kepada Masyarakat, Direktorat Jenderal pendidikan Tinggi Kementrian Pendidikan dan Kebudayaan, yang telah memberikan dana hibah penelitian melalui skema Penelitian Fundamental tahun 2014. Untuk itu kami mengucapkan terima kasih yang sebesar-besarnya.

Meskipun kami telah berusaha untuk menyusun laporan ini sebaik-baiknya, namun tidak bisa dipungkiri masih ada kekurangsempurnaan. Untuk itu masukan dan saran senantiasa kami harapkan untuk perbaikan di masa mendatang.

Yogyakarta, 30 Oktober 2014
Ketua,

Dr. Suparman, M.Si., DEA

DAFTAR ISI

BAB 1. PENDAHULUAN

1. Latar Belakang Masalah

Dalam analisis regresi, jika variabel terikat meningkat secara linier bersama dengan variabel bebas sampai nilai ambang tertentu dan setelah nilai ambang tadi variabel terikat mengingkat secara linier bersama dengan variabel bebas tapi pada tingkat yang berbeda sampai nilai ambang kedua dan demikian seterusnya, maka model tadi disebut model regresi linear per potongan. Sebagai contoh, suatu perusahaan memberikan komisi pada petugas penjualannya. Perusahaan tadi membayar komisi yang didasarkan pada penjualan dengan cara sedemikian rupa sehingga sampai suatu tingkat tertentu, yang disebut ambang ada satu struktur komisi dan selewat di atas tadi struktur komisi lainnya. Asumsikan bahwa komisi penjualan meningkat secara linear bersama dengan penjualan.

Model regresi linear per potongan merupakan model yang sangat fleksibel dan sering untuk memodelkan data dalam banyak bidang, antara lain dalam bidang ekonometri (Gujarati, 2006), bidang Geofisika (Stewart and Whaler, 1995), bidang kesehatan (Shi *et al*, 2011) dan bidang ekologi (Toms and Lesperance, 2003). Meskipun model seperti itu sangat berguna dalam banyak bidang, model tadi memiliki beberapa masalah estimasi karena dimensi dari parameter tidak diketahui. Dengan kata lain, dimensi dari parameter (banyaknya titik ambang) juga merupakan parameter yang harus diestimasi.

2. Permasalahan

Berdasarkan latar belakang di atas, maka yang menjadi permasalahan dalam penelitian ini adalah :

a. Bagaimana cara mengestimasi banyaknya titik ambang ?
b. Bagaimana cara mengestimasi koefisien regresi untuk masing-masing potongan ?
c. Bagaimana cara mengestimasi variansi galat untuk masing-masing potongan ?
d. Bagaimana cara memprakirakan nilai variabel terikat jika diketahui nilai variabel bebasnya ?
e. Bagaimana cara membuat algoritma untuk memprakirakan nilai variabel terikat jika diketahui nilai dari variabel bebasnya ?

BAB 2. TINJAUAN PUSTAKA

Metode *reversible jump* MCMC telah berhasil dengan baik diterapkan diberbagai bidang termasuk pada pemrosesan sinyal dan analisis data deret berkala. Beberapa hasil penelitian yang telah dilakukan oleh ketua peneliti terkait penerapan metode *reversible jump* MCMC adalah : Suparman *et al* (2002) menerapkan metode *reversible jump* MCMC untuk mendeteksi sinyal yang dikacau oleh galat multiplikatif, Suparman (2008) mengunakan metode *reversible jump* MCMC untuk mengestimasi model ARMA, Suparman (2010a) menggunakan metode *reversible jump* MCMC untuk mensegmentasi sinyal MA konstan per segmen, dan Suparman (2010b) menggunakan metode *reversible jump* MCMC untuk mensegmentasi sinyal AR konstan per segmen.

Dalam usulan penelitian ini metode *reversible jump* MCMC (Green, 1995) akan digunakan mengestimasi regresi linear per potongan. Misalkan y menyatakan variabel terikat dan x menyatakan variabel bebas. Regresi linear per potongan dari variabel terikat y_t terhadap variabel bebas *t* untuk $t = 1, 2, \cdots, n$ dapat dituliskan dalam persamaan stokastik berikut :

$$y_t = \begin{cases} \alpha_1 + \beta_1 t + z_t & \tau_1 < t \le \tau_2 \\ \alpha_2 + \beta_2 t + z_t & \tau_2 < t \le \tau_3 \\ \vdots & \\ \alpha_k + \beta_k t + z_t & \tau_k < t \le \tau_{k+1} \end{cases}$$

dengan $\tau_1 = 0$, $\tau_{k+1} = n$ dan

$$\begin{aligned} z_t &\sim N(0, \sigma_1^2) \quad \tau_1 \le t < \tau_2 \\ z_t &\sim N(0, \sigma_2^2) \quad \tau_2 \le t < \tau_3 \\ &\vdots \\ z_t &\sim N(0, \sigma_k^2) \quad \tau_k \le t < \tau_{k+1} \end{aligned}$$

Dalam persamaan di atas :

a. k menyatakan banyaknya titik ambang

b. $\tau_1, \tau_2, \cdots, \tau_{k+1}$ menyatakan titik-titik ambang yang bersesuaian

c. $\alpha_1, \cdots, \alpha_k$ dan $\beta_1, \cdots, \beta_k$ menyatakan koefisien regresi

d. $\sigma_1^2, \sigma_2^2, \cdots, \sigma_k^2$ menyatakan variansi galat.

Jika θ menyatakan parameter model regresi linear di atas, maka

$$\theta = \left(k, \tau_1, \cdots, \tau_{k+1}, \alpha_1, \cdots, \alpha_k,\ \beta_1, \cdots, \beta_k, \sigma_1^2,, \cdots, \sigma_k^2\right).$$

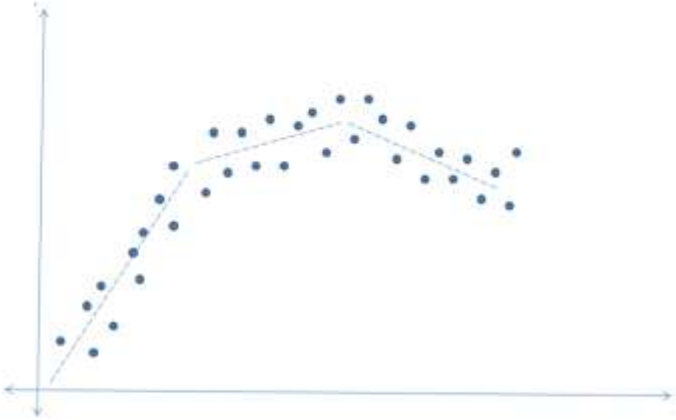
Sebagai ilustrasi untuk k = 3, maka $\theta = (3, \tau_1, \tau_2, \tau_3, \tau_4, \alpha_1, \alpha_2, \alpha_3,$ $\beta_1, \beta_2, \beta_3, \sigma_1^2, \sigma_2^2, \sigma_3^2 \ )$. Persamaan regresinya menjadi

$$y_t = \begin{cases} \alpha_1 + \beta_1 t + z_t & \tau_1 < t \le \tau_2 \\ \alpha_2 + \beta_2 t + z_t & \tau_2 < t \le \tau_3 \\ \alpha_3 + \beta_3 t + z_t & \tau_3 < t \le \tau_4 \end{cases}$$

dan

$$\begin{aligned} z_t &\sim N(0, \sigma_1^2) \quad \tau_1 < t \le \tau_2 \\ z_t &\sim N(0, \sigma_2^2) \quad \tau_2 < t \le \tau_3 \\ z_t &\sim N(0, \sigma_3^2) \quad \tau_3 < t \le \tau_4 \end{aligned}$$

Kurva dari persamaan regresinya dapat dilihat pada gambar berikut.



Gambar 1 : Kurva regresi linear per potongan untuk k = 3

Sebelum membicarakan estimasi regresi linear per potongan dengan banyak titik ambang tidak diketahui, terlebih dahulu akan dilihat beberapa hasil penelitian untuk kasus banyaknya titik ambang diketahui. Jika banyak titik ambang diketahui, permasalahan estimasi model regresi linear per potongan dapat disederhanakan menjadi permasalahan estimasi koefisien model regresi linear dan estimasi variansi galat yang bersesuaian untuk masing-masing segmen. Estimasi koefisien regresi linear dapat dilakukan dengan menggunakan berbagai metode. Metode-metode tersebut diantaranya diusulkan oleh Bain dan Engelhardt (1992), Efron dan Tibshirani (1993), dan Robert (2001). Dalam mengestimasi koefisien regresi linear, para penulis dan peneliti menggunakan pendekatan yang berbeda. Bain dan Engelhardt (1992) menggunakan Metode Kuadrat Terkecil, Efron dan Tibshirani (1993) menggunakan Metode Bootstrap. Sedangkan, Robert (2001) menggunakan Metode Bayesian.

Namun kenyataannya banyaknya titik ambang umumnya tidak diketahui. Jika banyaknya titik ambang tidak diketahui, permasalahan estimasi model regresi linear per potongan menjadi lebih kompleks. Untuk mengestimasi parameter θ digunakan metode Bayes. Misalkan $\pi(\theta)$ merupakan distribusi prior untuk parameter θ dan $f(y_1, y_2, \cdots, y_n | \theta)$ merupakan fungsi kemungkinan untuk data $y_1, y_2, \cdots, y_n$, maka distribusi posterior untuk parameter θ dapat ditentukan dengan Teorema Bayes

$$\pi(\theta | y_1, y_2, \cdots, y_n) \propto f(y_1, y_2, \cdots, y_n | \theta)\pi(\theta)$$

di mana $\propto$ berarti "sebanding dengan". Distribusi posterior ini mempunyai bentuk yang sangat kompleks mengakibatkan estimasi parameter tidak dapat ditentukan secara eksak. Untuk menyelesaikan masalah tersebut digunakan metode *reversible jump* MCMC.

Ide dasar dari metode *reversible jump* MCMC adalah pembuatan rantai markov yang rekuren dan ireduktibel sedemikian sehingga distribusi limit dari rantai Markov tersebut akan sama dengan distribusi posterior. Selanjutnya rantai Markov yang dihasilkan digunakan untuk menghitung estimator untuk parameter θ. Misalkan

$$\hat{\theta} = \left(\hat{k}, \hat{\tau}_1, \cdots, \hat{\tau}_{k+1}, \hat{\alpha}_1, \cdots, \hat{\alpha}_k, \hat{\beta}_1, \cdots, \hat{\beta}_k, \hat{\sigma}_1^2, \cdots, \hat{\sigma}_k^2\right)$$

merupakan estimator untuk parameter θ yang didapat dari metode *reversible jump* MCMC. Maka model regresi linear per potongan untuk data dapat ditulis sebagai

$$y_t = \begin{cases} \hat{\alpha}_1 + \hat{\beta}_1 t + z_t & \hat{\tau}_1 < t \le \hat{\tau}_2 \\ \hat{\alpha}_2 + \hat{\beta}_2 t + z_t & \hat{\tau}_2 < t \le \hat{\tau}_3 \\ \vdots & \\ \hat{\alpha}_k + \hat{\beta}_{\hat{k}} t + z_t & \hat{\tau}_{\hat{k}} < t \le \hat{\tau}_{\hat{k}+1} \end{cases}$$

Agar model yang diperoleh layak, maka dilakukan proses pengujian. Pengujian dilakukan dengan dengan menggunakan metode simulasi. Setelah model dinilai layak, maka model tersebut digunakan untuk peramalan.

Penelitian ini juga merupakan bagian dari peta jalan penelitian yang telah dan akan dilakukan dalam kurun waktu tertentu (Tabel 1). Berdasarkan hasil pada penelitian tahun pertama penerapan metode *reversible jump* MCMC untuk mengestimasi model regresi polinomial per potongan dengan prakiraan waktu penelitian selama 2 tahun. Regresi linear per potongan merupakan bentuk khusus dari regresi polinomial per potongan jika polinomialnya berorde satu. Peta jalan penelitian ini sejalan dengan Rencana Induk Penelitian (RIP) Universitas Ahmad Dahlan Tahun 2013-2017.

Tabel 1 : Peta Jalan Penelitian

Topik	Sub Topik	Tahun			
		2014	2015	2016	2017
Estimasi Bayesian dengan Metode Reversible Jumps Markov Chain Monte Carlo	Estimasi Parameter Model Regresi Linear Konstan per Potongan	Metode estimasi parameter	Algoritma komputer untuk estimasi parameter		
	Estimasi Parameter Model Regresi Polinomial Konstan per Potongan			Metode estimasi parameter	Algoritma komputer untuk estimasi parameter
	Estimasi Parameter Model Moving Average Inversible per Potongan				
	Estimasi Parameter Model Autoregresif Stasioner per Potongan				

BAB 3. TUJUAN DAN MANFAAT PENELITIAN

Tujuan

Kegiatan penelitian ini bertujuan untuk mencari pemecahan dari masalah-masalah yang diuraikan di atas. Tujuan kegiatan penelitian untuk masing-masing tahun adalah sebagai berikut :

Tujuan Penelitian Tahun Pertama :

a. Menemukan estimasi banyaknya titik ambang.
b. Menemukan estimasi koefisien regresi untuk masing-masing potongan.
c. Menemukan estimasi variansi galat untuk masing-masing potongan.

Tahun Penelitian Tahun Kedua :

a. Menemukan cara meramalkan nilai variabel terikat jika diketahui nilai variable bebasnya.
b. Membuat algoritma untuk memprakirakan nilai variabel terikat jika diketahui nilai dari variabel bebasnya.

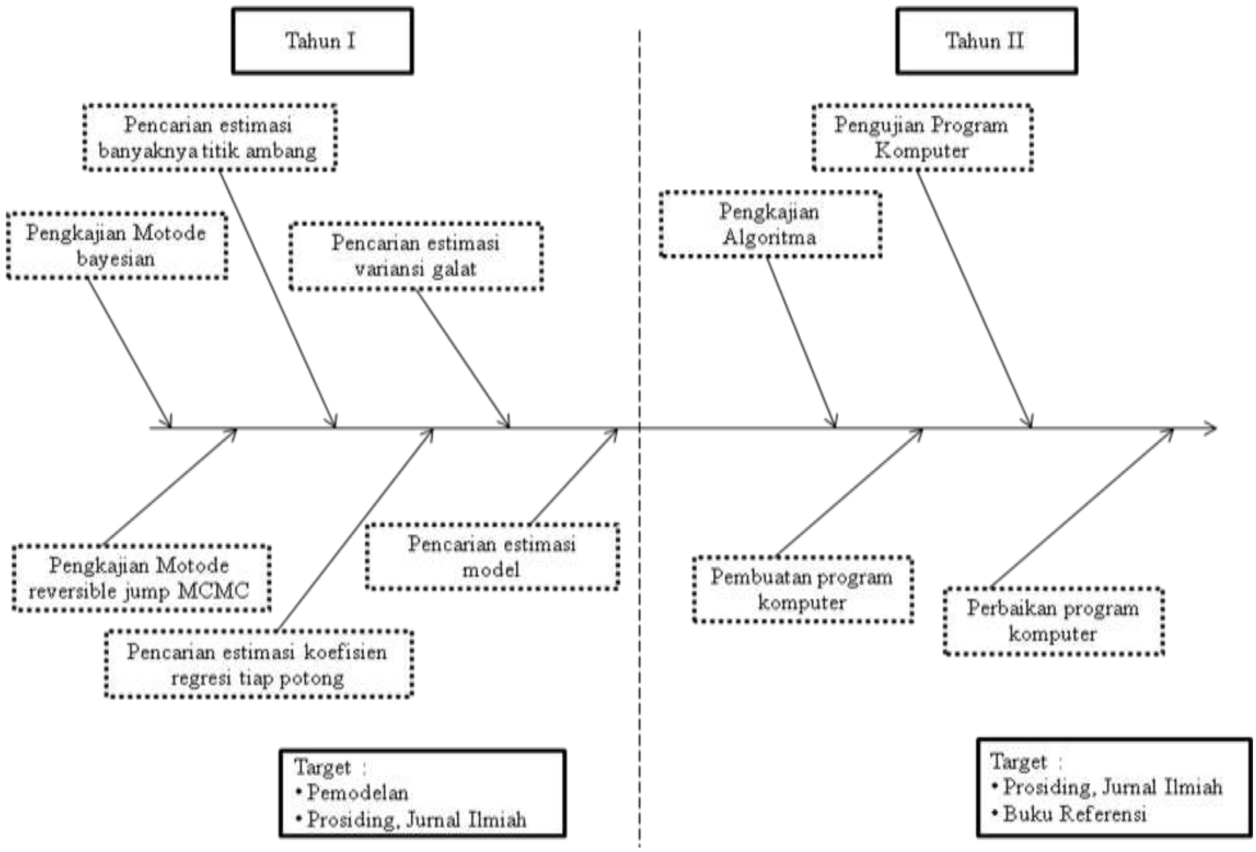
Manfaat

Penelitian ini menghasilkan metode baru untuk mengestimasi model regresi linear per potongan dan memprakirakan nilai variabel terikat jika diketahui nilai variabel bebasnya pada regresi linear per potongan. Dengan diketahuinya metode tersebut, para pembuat kebijakan dapat menerapkan hasil penelitian ini untuk mengembangkan berbagai alternatif kebijakan dalam variabel bebas dan memprakirakan dampaknya terhadap variabel terikatnya. Pembuat kebijakan dapat membandingkan hasil simulasi dari berbagai alternatif kebijakan yang dibuat, dan menentukan alternatif mana yang paling baik untuk diimplementasikan. Alternatif kebijakan yang baik adalah alternatif kebijakan yang memiliki dampak positif yang paling besar.

BAB 4. METODE PENELITIAN

Kegiatan penelitian fundamental ini termasuk dalam jenis penelitian pengembangan ipteks. Tahap pertama adalah penelusuran pustaka (artikel jurnal ilmiah, buku, prosiding seminar ilmiah) dan bahan yang relevan dan mutakhir dengan masalah penelitian. Tahap berikutnya adalah mengkaji dan mengulas pustaka tersebut. Kemudian menulis gagasan fundamental dan orisinal mengenai Algoritma Reversible Jump Markov Chain Monte Carlo yang berkaitan dengan estimasi parameter model regresi linear konstan per potongan. Tahap selanjutnya adalah membuat algoritmanya dalam bahasa pemrograman Matlab (Chapman, 2009). Kinerja algoritma diuji dengan data simulasi. Selanjutnya algoritma diimplementasikan pada data real yang diambil di lapangan.

Bagan alir sistematika kegiatan penelitian fundamental selama 2 tahun disajikan dalam gambar berikut.



Gambar 2 : Bagan alir kegiatan penelitian

BAB 5. HASIL DAN PEMBAHASAN

Sampai dengan disusunnya laporan akhir ini, kegiatan penelitian telah menyelesaikan semua kajian teori yang merupakan target penelitian tahun pertam. Hasil kajian teori yang dihasilkan dalam penelitian tahun pertama ini adalah sebagai berikut :

1. Fungsi kemungkinan
2. Distribusi prior
3. Distribusi posterior
4. Estimasi Parameter Model dengan Reversible Jump MCMC

Hasil penelitian tahun pertama ini didesiminasi dan dipublikasikan baik pada seminar ilmiah nasional (nasional, internasional) maupun jurnal ilmiah (nasional, internasional), sebagai berikut:

1. Publikasi Ilmiah (Jurnal AdMathEdu Vol.3 No.1, Juni 2014, ISSN : 2088-687X), naskah dalam status sudah terbit.
2. Pembicara pada Pertemuan Ilmiah (40th IEEE International Conference on Acoustic, Speech and Signal Processing, 19-24 April 2015, Brisbane Australia), makalah dalam status sedang ditelaah.
3. Pembicara pada Pertemuan Ilmiah (Seminar Nasional Riset Inovatif II, 21-22 November 2014, Bali), makalah dalam status sudah diterima.
4. Publikasi ilmiah (International Journal of Computer Technology and Application (Suparman and Michel Doisy, dalam status sudah dikirim)

Pada bagian berikut diuraikan hasil kajian teori mengenai metode estimasi parameter model regresi linear per potongan secara lebih detail.

Fungsi Kemungkinan

Misalkan y menyatakan variabel terikat dan x menyatakan variabel bebas. Regresi linear per potongan dari variabel terikat y_t terhadap variabel bebas t untuk $t = 1,2,\cdots,n$ dapat dituliskan dalam persamaan stokastik berikut :

$$y_t = \begin{cases} \alpha_1 + \beta_1 t + z_t & \tau_1 < t \le \tau_2 \\ \alpha_2 + \beta_2 t + z_t & \tau_2 < t \le \tau_3 \\ \vdots & \\ \alpha_k + \beta_k t + z_t & \tau_k < t \le \tau_{k+1} \end{cases}$$

dengan $\tau_1 = 0$, $\tau_{k+1} = n$ dan

$$\begin{aligned} z_t &\sim N(0,\sigma_1^2) \quad \tau_1 \le t < \tau_2 \\ z_t &\sim N(0,\sigma_2^2) \quad \tau_2 \le t < \tau_3 \\ &\vdots \\ z_t &\sim N(0,\sigma_k^2) \quad \tau_k \le t < \tau_{k+1} \end{aligned}$$

Untuk $i = 1, 2, \cdots, k$ dan $\tau_i < t \le \tau_{i+1}$, karena z_t (untuk) berdistribusi normal dengan 0 dan variansi σ_i^2, maka fungsi kepadatan dari z_t adalah

$$f(z_t | \sigma_i^2) = \left(\frac{1}{\sqrt{2\pi\sigma_i^2}} \right) \exp - \frac{1}{2\sigma_i^2} z_t^2$$

Sehingga untuk $z_i = (z_{\tau_i + 1}, \cdots, z_{\tau_{i+1}})$ fungsi kepadatan gabungan dari z_i

$$f(z_i | \sigma_i^2) = \prod_{t=\tau_i+1}^{\tau_{i+1}} f(z_t | \sigma_i^2)$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma_i^2}} \right)^{\tau_{i+1}-\tau_i} \exp - \frac{1}{2\sigma_i^2} \sum_{t=\tau_i}^{\tau_{i+1}} z_t^2$$

Dengan menggunakan transformasi variabel $y_t = \alpha_t + \beta_t t + z_t$, maka $z_t = y_t - \alpha_t - \beta_t t$ dan $\frac{dz_t}{dy_t} = 1$. Sehingga fungsi kepadatan dari y_i adalah

$$f(y_i | \alpha_i, \beta_i, \sigma_i^2) = \left(2\pi\sigma_i^2\right)^{-\frac{1}{2}(\tau_{i+1}-\tau_i)} \exp - \frac{1}{2\sigma_i^2} \sum_{t=\tau_i}^{\tau_{i+1}} (y_t - \alpha_i - \beta_i t)^2$$

Akhirnya fungsi kemungkinan maksimum untuk $y = (y_1, y_2, \cdots, y_n)$

$$L(y | \theta) = \prod_{i=1}^{k} \left\{ (2\pi\sigma_i^2)^{-\frac{1}{2}(\tau_{i+1}-\tau_i)} \exp - \frac{1}{2\sigma_i^2} \sum_{t=\tau_i}^{\tau_{i+1}} (y_t - \alpha_i - \beta_i t)^2 \right\}$$

Catatan :

$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k)$

$\beta = (\beta_1, \beta_2, \cdots, \beta_k)$

$\sigma^2 = (\sigma_1^2, \sigma_2^2, \cdots, \alpha_k^2)$

$\tau = (\tau_1, \tau_2, \cdots, \tau_{k+1})$

$\theta = (k, \alpha, \beta, \sigma^2, \tau)$

Distribusi Prior

Metode bayes digunakan untuk mengestimasi parameter $\theta = (k, \alpha\ \ , \beta\ \ , \sigma^2, \tau\ \ )$. Untuk itu, distribusi prior dipilih. Misalkan k_{maks} menyatakan maksimum banyaknya potongan. Distribusi banyaknya potongan mengikuti distribusi binomial dengan hiper parameter λ, yaitu:

$$\pi(k) = C_k^{k_{maks}} \lambda^k (1-\lambda)^{k_{maks}-k} \text{ untuk } k = 1, 2, \cdots, k_{maks}$$

Distribusi posisi titik-titik ambang mengikuti distribusi posisi indek genap terurut pada 2k+1 titik yang diambil secara acak tanpa pengembalian dalam {1, 2, ..., n-1).

$$\pi(\tau \mid k) = \frac{(2k+1)!}{n^{2k}} \frac{1}{2^k} \prod_{i=1}^{k} (\tau_{i+1} - \tau_i)$$

Selanjutnya, untuk banyaknya potongan yang diberikan distribusi parameter regresi, sebagai berikut :

$$\pi(\alpha \mid k) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi a^2}} \exp - \frac{1}{2a^2} \alpha_i^2 \text{ (distribusi N(0,1)}$$

$$\pi(\beta \mid k) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi b^2}} \exp - \frac{1}{2b^2} \beta_i^2 \text{ (distribusi N(0,1)}$$

$$\pi(\sigma^2 \mid k) = \prod_{i=1}^{k} c(\sigma_i^2)^{-2} \exp - \frac{c}{\sigma_i^2} \text{ (distribusi IG(1,c))}$$

Persoalan yang timbul, diantaranya munculnya hiperparameter $\varphi = (\lambda, c)$ dalam distribusi prior di atas. Selanjutnya hiperparameter φ dipandang sebagai suatu variabel acak dengan distribusi tertentu, yaitu :

$$\pi(\lambda) \sim U(0,1)$$

$$\pi(c) \sim Gamma(1,1)$$

Misalkan $\pi(\theta, \varphi)$ menyatakan distribusi prior untuk (θ, φ). Karena $\pi(\theta \mid \varphi) = \dfrac{\pi(\theta, \varphi)}{\pi(\varphi)}$ maka distribusi prior untuk (θ, φ) dapat ditentukan sebagai berikut

$$\pi(\theta, \varphi) = \pi(\theta \mid \varphi) \pi(\varphi)$$

atau

$$\pi(\theta, \varphi) \propto C_k^{k_{maks}} \lambda^k (1-\lambda)^{k_{maks}-k} \frac{(2k+1)!}{n^{2k}} \frac{1}{2^k} \prod_{i=1}^{k} (\tau_i - \tau_{i-1}) \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi}} \exp - \frac{1}{2} \alpha_i^2$$

$$\prod_{i=1}^{k} \frac{1}{\sqrt{2\pi}} \exp - \frac{1}{2} \beta_i^2 \prod_{i=1}^{k} \frac{c}{2} (\sigma_i^2)^{-2} \exp - \frac{c}{2\sigma_i^2}$$

Distribusi Posterior

Misalkan $\pi(\theta,\varphi \mid y)$ merupakan distribusi posterior untuk (θ,φ). Dengan menggunakan Teorema Bayes, maka distribusi posterior untuk parameter (θ,φ) dapat dinyatakan sebagai hasil kali fungsi kemungkinan dan distribusi prior

$$\pi(\theta,\varphi \mid y) \propto f(y \mid \theta)\pi(\theta,\varphi)$$

$$\propto f(y \mid \theta)\pi(\theta,\varphi)\ \pi(\theta \mid \varphi)\pi(\varphi)$$

$$\propto \prod_{i=1}^{k} \left\{ (\sigma_i^2)^{-\frac{1}{2}(\tau_{i+1}-\tau_i)} \exp - \frac{1}{2\sigma_i^2} \sum_{t=\tau_i}^{\tau_{i+1}} (y_t - \alpha_i - \beta_i\, t)^2 \right\}$$

$$C_k^{k_{maks}} \lambda^k (1-\lambda)^{k_{maks}-k} \frac{(2k+1)!}{n^{2k}} \frac{1}{2^k} \prod_{i=1}^{k} (\tau_i - \tau_{i-1}) \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi}} \exp - \frac{1}{2}\alpha_i^2$$

$$\prod_{i=1}^{k} \frac{1}{\sqrt{2\pi}} \exp - \frac{1}{2}\beta_i^2 \prod_{i=1}^{k} c(\sigma_i^2)^{-2} \exp - \frac{c}{2\sigma_i^2}$$

Metode Reversible Jump MCMC

Misalkan $M = (\theta,\varphi)$. Secara umum, metode MCMC merupakan suatu metode sampling dengan cara membuat rantai Markov homogen $M_1, M_2, \cdots, M_m$ yang memenuhi sifat aperiodik dan irreduktibel (Robert, 1996) sedemikian hingga $M_1, M_2, \cdots, M_m$ dapat dipertimbangkan sebagai variabel acak yang berdistribusi $\pi(\theta,\varphi \mid y)$. Dengan demikian $M_1, M_2, \cdots, M_m$ dapat digunakan untuk menaksir parameter M. Untuk merealisasikan hal tersebut, diadopsi Algoritma Gibbs (Robert, 1996) yang terdiri dari dua tahap :

1. Simulasi distribusi $\pi(\varphi \mid \theta, y)$
2. Simulasi distribusi $\pi(\theta \mid \varphi, y)$

Distribusi $\pi(\varphi \mid \theta, y)$ mempunyai bentuk eksplisit. Sehingga Agoritma Gibbs dapat digunakan untuk membuat simulasi distribusi $\pi(\varphi \mid \theta, y)$.

$$\pi(\varphi \mid \theta, y) = B(k+1, k_{maks} - k + 1) \otimes \prod_{i=1}^{k} G(2k, \sum_{i=1}^{k} \frac{1}{2\sigma_i^2})$$

Sebaliknya, distribusi $\pi(\theta \mid \varphi, y)$ tidak mempunyai bentuk eksplisit. Sehingga simulasi eksak tidak mungkin dilakukan. Penyelesaiannya menggunakan Algoirtma Hibrida yang terdiri dari tiga tahap sebagai berikut :

2.1. Simulasi $\pi(\sigma^2 | k,\tau ,\alpha,\beta,\varphi,y)$

2.2. Simulasi $\pi(\alpha | k,\tau ,\beta,\sigma^2,\varphi,y)$

2.3. Simulasi $\pi(\beta | k,\tau ,\alpha,\sigma^2,\varphi,y)$

2.4. Simulasi $\pi(k,\tau | \alpha,\beta,\sigma^2,\varphi,y)$

Distribusi $\pi(\sigma^2 | k,\tau ,\varphi,y)$ mempunyai bentuk eksplisit, sehingga Algoritma Gibbs dapat digunakan untuk membuat simulasi distribusi $\pi(\sigma^2 | k,\tau ,\varphi,y)$.

$$\pi(\sigma^2 | k,\tau ,\varphi,y) = \otimes_{i=1}^{k} IG(a_i, b_i)$$

di mana $a_i = 1 + \frac{1}{2}(\tau_{i+1} - \tau_i)$ dan $b_i = c + \sum_{t=\tau_i+1}^{\tau_{i+1}} (y_t - \alpha_i - \beta_i t)^2$. Demikian pula, distribusi $\pi(\alpha | k,\tau ,\beta,\sigma^2,\varphi,y)$ dan $\pi(\beta | k,\tau ,\alpha,\sigma^2,\varphi,y)$ masing-masing mempunyai bentuk ekplisit, sehingga Algoritma giss dapat digunakan untuk membuat simulasi distribusi-distribusi tersebut.

Sebaliknya, karena k tidak diketahui maka Algoritma MCMC biasa tidak bisa digunakan untuk membuat simulasi distribusi $\pi(k,\tau | \varphi,y)$. Sebagai gantinya digunakan Algoritma Reversible Jump MCMC untuk membuat simulasi distribusi $\pi(k,\tau | \varphi,y)$.

Misalkan $\omega = (k,\tau)$ adalah titik aktual dari rantai markov. Terdapat 3 jenis transformasi yang digunakan, yaitu : kelahiran segmen, kematian segmen dan perubahan terjadinya perubahan model. Selanjutnya misalkan N_k adalah probabilitas transformasi dari k menuju k+1 (kelahiran), D_k adalah probabilitas transformasi dari k+1 menuju k (kematian), dan P_k adalah probabilitas transformasi dari k menuju k (perubahan).

Kelahiran/Kematian Sebuah Segmen

Transformasi kelahiran membuat banyaknya segmen berubah dari k menuju k+1. Jika transformasi kelahiran yang dipilih, transformasi kelahiran dari titik $\omega = (k,\tau)$ didefinisikan dengan cara berikut : Pilih titik z secara acak dalam $\{1,\cdots,n-1\} \setminus \tau$. Misalkan titik z dalam $\{\tau_i + 1,\cdots,\tau_{i+1} - 1\}$ maka

$$\tau_1, ..., \tau_i, \tau_{i+1} = z, \tau_{i+2}, ..., \tau_{k+2}$$

Sehingga diperoleh titik $\omega^* = (k+1,\tau)$ yaitu titik yang dihasilkan dari transformasi kelahiran. Sebaliknya transformasi kematian membuat banyaknya segmen berubah dari k+1 menuju k. Jika transformasi kematian yang dipilih, maka transformasi kematian dari titik $\omega^* = (k+1, \tau)$ didefinisikan dengan cara berikut : Pilih secara acak sebuah titik dalam τ. Misalkan titik tersebut adalah τ_{i+1} maka

$$\tau_1, \ldots, \tau_i, \tau_{i+2}, \ldots, \tau_{k+1}$$

Misalkan a_n dan a_d masing-masing merupakan probabilitas penerimaan untuk kelahiran dan probabilitas penerimaan untuk kematian. Maka probabilitas penerimaan untuk kelahiran adalah sebagai berikut

$$a_n(\omega, \omega^*) = \min\left\{1, \frac{\pi(\omega^* | \varphi, y)}{\pi(\omega | \varphi, y)} \frac{q(\omega^*, \omega)}{q(\omega, \omega^*)}\right\}$$

Sedangkan probabilitas untuk kematian adalah sebagai berikut\

$$a_d(\omega, \omega^*) = \min\left\{1, \frac{1}{a_n(\omega^*, \omega)}\right\}$$

di mana

$$\frac{q(\omega^*, \omega)}{q(\omega, \omega^*)} = \frac{D_{k+1}}{N_k} \quad \frac{n-1-k}{k+1}$$

Perubahan Sebuah Segmen

Transformasi perubahan tidak membuat banyaknya segmen berubah dari k menuju k, akan tetapi membuat batas segmennya ada yang berubah. Jika transformasi perubahan yang dipilih, transformasi perubahan dari titik $\omega = (k, \tau)$ didefinisikan dengan cara berikut : Pilih titik secara acak dalam τ. Misalkan titik tersebut adalah τ_i maka titik ini digantikan dengan z yang dibangkitkan dari distribusi seragam pada himpunan $\{1, \cdots, n-1\} \setminus \tau$.

Misalkan a_p merupakan probabilitas penerimaan untuk perubahan. Maka probabilitas penerimaan untuk perubahan adalah sebagai berikut

$$a_p(\omega, \omega^*) = \min\left\{1, \frac{\pi(\omega^* | \varphi, y)}{\pi(\omega | \varphi, y)} \frac{q(\omega^*, \omega)}{q(\omega, \omega^*)}\right\}$$

Sedangkan probabilitas untuk kematian adalah sebagai berikut
di mana

$$\frac{q(\omega^*, \omega)}{q(\omega, \omega^*)} = 1$$

BAB 6. KESIMPULAN DAN SARAN

Dalam penelitian ini dikaji estimasi parameter model regresi linear konstan per potongan. Jika banyaknya regresi diketahui, maka estimasi parameter dapat dilakukan dengan metode markov chain monte carlo. Namun dalam penelitian ini, banyaknya regresi tidak diketahui. Dengan kata lain, banyaknya regresi juga merupakan variabel. Sehingga metode markov chain monte carlo tidak dapat digunakan. Penggantinya digunakan metode reversible jump markov chain monte carlo.

Metode reversible jump MCMC merupakan salah satu metode baru yang dapat digunakan untuk mengestimasi parameter model regresi linear per potongan meskipun banyaknya titik ambang tidak diketahui. Keunggulan dari metode ini adalah estimasi banyaknya titik ambang dan estimasi parameter regresi tiap potongan dapat diestimasi secara bersama-sama.

DAFTAR PUSTAKA

Bain, L.J. and Engelhardt, M. 1992. *Introduction to Probability and Mathematical Statistics*, Duxbury Press, California.

Chapman, S.J. (2009) Essentials of Matlab Programming, Cengage Learning.

Efron, B. and Tibshirani, R.J. 1993. *An Introduction to the Bootstrap*, Chapman & Hill, New York.

Green, P.J. 1995. Reversible Jump MCMC Computation and Bayesian Model Determination. *Biometrika*, Vol. 82, pp. 711-732.

Gujarati, D.N. 2006. *Essensials of Econometrics*, McGraw-Hill, New York.

Robert, C.P. 2001. *The Bayesian Choice : A Decision-Theory Motivation*, Springer, New York.

Shi, H.Y., Lee, H.H, Tsai, M.H, Chiu, C.C, Uen, Y.H. and Lee, K.T. 2011. Long-term Outcomes of Laparoscopic Colecystectomy: a Perpective piecewise linear regression analysis. *Surg Endosc*, pp. 2132-2140.

Stewart, D.N. and Whaler, K.A. 1995. Optimal Piecewise Regression Analysis and Its Application to Geomagnetic Time Series. *Geophysical Journal International*, pp. 710-724.

Suparman, Doisy, M., and Tourneret, J.Y. 2002. Changepoint Detection using Reversible Jump MCMC Methods, *Proceedings of the IEEE ICASSP*, pp. 1569-1573.

Suparman. 2008. Pemilihan Model Bayesian Hirarki Dalam Model ARMA Menggunakan Algoritma *Reversible Jump* MCMC, *Jurnal Eksakta*, Vol. 10, No. 2, Hal 66-76.

Suparman. 2009. Estimator Bayesian Hirarki untuk Parameter Model Sinyal Multiplikatif Menggunakan Algoritma MCMC Hibrida, *Prosiding Seminar Nasional Teknologi V*, Buku 8, Hal. 41-47.

Suparman. 2010a. Segmentasi Bayesian Hirarki untuk Model MA Inversiblel Konstan Sepotong demi Sepotong Berbasis Algoritma Reversibel Jump MCMC, *Jurnal Eksakta*, Vol. 11, No. 1, Hal 9-15.

Suparman. 2010b. Inferensi Bayesian Hirarki untuk Model AR Stasioner Konstan per Segmen Menggunakan Algoritma *Reversible Jump* MCMC, *Jurnal Kadikma*, Vol. 2, No 1, Hal 81-95.

Suparman, 2010c. *Pengantar Reversible Jump Markov Chain Monte Carlo dan Aplikasinya*, Penerbit UTY, Yogyakarta.

Suparman dan Taram, 2014. Metode Reversible Jump MCMC untuk Estimasi Model Regresi Linier Konstan per Potongan, *Jurnal AdmathEdu*, Vol. 4 No. 1, Hal. 79-86.

Toms, J.D. and Lesperance, M.L. 2003. Piecewise Regression: a Tool Identifiying Ecological Thresholds. *Ecology*, pp. 2034-2041.

LAMPIRAN-LAMPIRAN

Lampiran 1 : Personalia Tenaga Peneliti beserta Kualifikasinya

BIODATA KETUA PENELITI

A. Identitas Diri		
1.	Nama Lengkap (dengan gelar)	Dr. Suparman, M.Si., DEA
2.	Jenis Kelamin	L
3.	Jabatan Fungsional	Lektor
4.	NIY	60110621
5.	NIDN	0517046902
6.	Tempat dan Tanggal Lahir	Bantul, 17 April 1969
7.	E-mail	suparmancict@yahoo.co.id
8.	Nomor Telepon /HP	081328201198
9.	Alamat Kantor	Kampus 3 UAD Jl. Prof. Dr. Soepomo, SH Yogyakarta
10.	Nomor Telepon/Faks	0274 563515 / 0274 564604
11.	Lulusan yang Telah Dihasilkan	S-1 = 48 orang
12	Mata Kuliah yang Diampu	1. Statistika Matematika
		2. Teori Peluang
		3. Metodologi Penelitian
		4. Analiis Desain dan Eksperimen
		5. Statistika Nonparametrik

B. Riwayat Pendidikan				
	S1	S2		S3
Nama Perguruan Tinggi	UNILA	UGM	Universitas Toulouse 3	Universitas Toulouse 3
Bidang Ilmu	Pendidikan Matematika	Matematika	Matematika Terapan	Matematika Terapan
Tahun Masuk-Lulus	1988 – 1992	1994-1997	1999-2000	2000-2003
Judul Skripsi/Tesis/Disertasi	Perbandingan Nilai Mhs Ditinjau dari Segi Asal Sekolah, Jurusan dan Ekonomi Orang Tua	Estimasi Bayesian Model runtun waktu ARMA	Methode de Green Applications a la Detection de Ruptures dans un signal	Problems de choix de models par simulation de type monte carlo par chaines de markov a sauts reversibles

Nama Pembimbing/Promotor	Drs. Buchori Kifli /Drs maksum	Prof. Zanzawi S, Ph.D	Prof. Dr. Michel Doisy	Prof Dr. Jean-Marc Azais / Prof. Dr. Michel Doisy

C. Pengalaman Penelitian Dalam 5 Tahun Terakhir				
No.	Tahun	Judul Penelitian	Pendanaan	
			Sumber	Jml (Juta Rp)
1.	2009	Segmentasi Bayesian Hirarkis untuk Data Deret Waktu Berkala Model MA Inversibel Konstan Sepotong-potong Menggunakan Algoritma Reversible Jump MCMC.	Kopertis V	1,5
2.	2011	Segmentasi Bayesian Hirarkis untuk Data Deret Waktu Berkala Model SAR Stasioner Konstan Sepotong-potong Menggunakan Algoritma Reversible Jump MCMC.	Kopertis V	1,675
3.	2014	Metode Reversible Jump Markov Chain Monte Carlo untuk Estimasi Bayesian dalam Model Regresi Linear per Potongan	Dikti	50

D. Pengalaman Pengabdian Kepada Masyarakat				
No.	Tahun	Judul Pengabdian Kepada Masyarakat	Pendanaan	
			Sumber	Jml (Juta Rp)
1	2009	Tim Pemantau Independen D Ujian Nasional di SMPN 1 Godean Sleman	LPMP DIY	0,8
2	2010	Tim Pemantau Independen D Ujian Nasional di SMPN 1 Godean Sleman	LPMP DIY	0,9
3	2013	IbM Guru SD di Gunung Kidul	Dikti	50
4	2014	IbM Guru SMK di Klaten	Dikti	36

E. Publikasi Artikel Ilmiah Dalam Jurnal 5 Tahun Terakhir			
No.	Judul Artikel Ilmiah	Nama Jurnal	Volume/ Nomor
1.	Inferensi Bayesian Hirarki untuk Model AR Stasioner Konstan per segmen Menggunakan Algoritma Reversible Jump MCMC	Jurnal Kadikma	Vol. 2 / No.1 Hal. 81-95 Tahun 2010
2.	Segmentasi Bayesian Hirarki untuk Model MA Konstan Sepotong demi Sepotong Berbasis Algoritma	Jurnal Eksakta	Vol. 11/No 1 Hal 9-15 Tahun 2010

	Reversible Jump MCMC		
3.	Implementasi Metode Bootstrap untuk Menentukan Prediksi Selang pada Model Regresi Subset Polinomial	Jurnal Konvergensi	Vol. 1 / No. 1 Hal 13-21 Tahun 2011
4-	Implementasi Metode Bootstrap untuk Pengujian Hipotesis Mengenai dua Mean Populasi	Jurnal Admathedu	Vol. 2 / No. 1 Hal 41-50 Tahun 2012
5.	Metode Reversible Jump Markov Chain Monte Carlo untuk Estimasi Bayesian dalam Model Regresi Linear per Potongan	Jurnal Admathedu	Vol. 4 No. 1 Hal 79-86 Tahun 2014

F. Pemakalah Seminar Ilmiah dalam 5 tahun Terakhir			
No.	Nama Pertemuan Ilmiah / Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1.	Seminar Nasional Teknologi V	Estimator Bayesian Hirarki untuk Parameter Model Sinyal Multiplikatif Menggunakan Algoritma MCMC Hibrida	2009 Yogyakarta
2.	Seminar Nasional Teknologi VI	Implementasi Metode Bootstrap untuk Menentukan Selang Prediksi pada Regresi Ganda	2010 Yogyakarta
3.	Seminar Nasional Kopertis V	Seleksi Orde dan Estimasi Parameter dalam Model Regresi Polinomial dengan Menggunakan Metode Bootstrap	2010 Yogyakarta
4.	Seminar Nasional Sistem Informasi Indonesia (SESINDO), ISBN : 978-979-18985-6-0	Segmentasi Bayesian Hirarki untuk Model AR Stasioner Konstan per Segmen Menggunakan Algoritma Reversible Jump MCMC	2-4 Desember 2013 Bali

G. Karya Buku dalam 5 Tahun Terakhir				
No.	Judul Buku Ajar	Tahun	Jumlah Halaman	Penerbit
1.	Kalkulus Lanjutan ISBN : 978-979-98964-7-6	2009	92	FST Press Yogyakarta
2.	Pengantar Reversible Jump MCMC dan Aplikasinya ISBN : 978-979-1334-30-0	2010	75	FST Press Yogyakarta
3.	Pengantar Statistika Teknik dan Bisnis ISBN : 978-979-505-213-8	2011	230	Muara Indah Bandung
4.	Statistika Matematika ISBN : 978-602-989-196-6	2012	130	MIPA Press Yogyakarta
5.	Analisis Runtun Waktu ISBN : 978-602-17339-4-3	2013	62	JPMIPA Press Yogyakarta

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggung jawabkan secara hukum. Dan apabila dikemudian hari ternyata dijumpai ketidak sesuaian dengan kenyataan, saya sanggup menerima resikonya.

Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi persyaratan sebagai salah satu syarat pengajuan hibah penelitian fundamental.

Yogyakarta, 30 Oktober 2014

Pengusul,

Dr. Suparman, M.Si., DEA

BIODATA ANGGOTA PENELITI

A. Identitas Diri

1.	Nama Lengkap (dengan gelar)	Drs. Abdul Taram, M.Si.
2.	Jenis Kelamin	L
3.	Jabatan Fungsional	Lektor Kepala
4.	NIY	60900070
5.	NIDN	0505035801
6.	Tempat dan Tanggal Lahir	Indramayu, 5 Maret 1958
7.	E-mail	taromahmad@yahoo.com
8.	Nomor Telepon /HP	0818268305
9.	Alamat Kantor	Kampus 3 UAD Jl. Prof. Dr. Soepomo, SH Yogyakarta
10.	Nomor Telepon/Faks	0274 563515 / 0274 564604
11.	Lulusan yang Telah Dihasilkan	S-1 = 72 orang
12	Mata Kuliah yang Diampu	1. Statistika Elementer
		2. Teori Peluang
		3. Analisis Runtun Waktu
		4. Analiis Desain dan Eksperimen

	S1	S2
Nama Perguruan Tinggi	IKIP Muhammadiyah	UGM
Bidang Ilmu	Pendidikan Matematika	Matematika
Tahun Masuk-Lulus	1982 – 1988	1994-1997
Judul Skripsi/Tesis/Disertasi	Studi Komparasi Proses Belajar Mengajar dengan Cara Belajar Siswa Aktif (CBSA) dengan Non-CBSA terhadap Prestasi Belajar Matematika Siswa Sekolah Menengah Ekonomi Atas (SMEA) N 3 Gowongan Yogyakarta Tahun Ajaran 1988/1989.	Estimasi Maksimum Likelihood untuk Model Analisis Runtun Waktu
Nama Pembimbing/Promotor	Prof.Drs.Hirdjan.	Prof.Dr.H.Zanzawi Soejoeti, M.Sc.

B. Pengalaman Penelitian Dalam 5 Tahun Terakhir

No.	Tahun	Judul Penelitian	Pendanaan	
			Sumber	Jml (Juta Rp)
1	2007	Identifikasi Orde dan Penaksiran Parameter Model AR untuk Data Deret Berkala Menggunakan Algoritma Simulated Annealing	Fundamental	27,5
2	2011	Pemilihan Model dan Estimasi Parameter Bayesian Hirarki untuk Model Subset ARMA Menggunakan Algoritma Reversible Jump Simulated Annealing MCMC	Kopertis V	1,5

C. Publikasi Artikel Ilmiah Dalam Jurnal 5 Tahun Terakhir

No.	Judul Artikel Ilmiah	Nama Jurnal	Volume/ Nomor
1.	Pemilihan model dan estimasi parameter bayesian hirarki untuk model subset arma menggunakan algoritma reversible jump simulated annealing MCMC.	Jurnal Admathedu	Vol. 1 / No. 2 Hal 115-142 Tahun 2011

D. Pemakalah Seminar Ilmiah dalam 5 tahun Terakhir

No.	Nama Pertemuan Ilmiah / Seminar	Judul Artikel Ilmiah	Waktu dan Tempat
1.	The 2nd International Conference on Green Wold in Business and Technology	The Effectifeness of Mathematics learning With Quiz Seen from The Point of View of Learning Time in The Achievement of Student Study Performace Grade VIII Semester of II SMP 2 Kasihan Regency Of Bantul Year 2009/2010	23 Maret 2013 Yogyakarta

E. Karya Buku dalam 5 Tahun Terakhir

No.	Judul Buku Ajar	Tahun	Jumlah Halaman	Penerbit
1.	Teori Peluang ISBN : 978-602-17339-5-0	2013	96	JPMIPA Press Yogyakarta

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggung jawabkan secara hukum. Dan apabila dikemudian hari ternyata dijumpai ketidak sesuaian dengan kenyataan, saya sanggup menerima resikonya.

Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi persyaratan sebagai salah satu syarat pengajuan hibah penelitian fundamental.

Yogyakarta, 30 Oktober 2014

Pengusul,

Drs. Abdul Taram, M.Si.

Lampiran 2 : Publikasi Ilmiah (Jurnal AdMathEdu Vol.3 No.1, Juni 2014, ISSN : 2088-687X), naskah dalam status sudah terbit.

METODE REVERSIBLE JUMP MARKOV CHAIN MONTE CARLO UNTUK ESTIMASI BAYESIAN DALAM MODEL REGRESI LINEAR PER POTONGAN

Suparman, Abdul Taram
Program Studi Pendidikan Matematika FKIP UAD
Jl. Prof. Dr. Soepomo, SH. Janturan Yogyakarta
suparmancict@yahoo.co.id

ABSTRAK

Model regresi linear per potongan merupakan model yang sangat fleksibel untuk memodelkan data. Jika model regresi linear per potongan dicocokkan terhadap data, maka umumnya parameternya tidak dketahui. Tulisan ini mengkaji masalah penaksiran parameter model regresi linear per potongan.

Metode yang digunakan untuk mengestimasi parameter regresi linear per potongan adalah Metode Bayesian. Namun Penaksir Bayes tidak dapat ditemukan secara analitik. Untuk mengatasi masalah tersebut diusulkan Algoritma Reversible Jump MCMC. Algoritma Reversible Jump MCMC menghasilkan Rantai Markov yang distribusi limitnya konvergen menuju distribusi posterior dari parameter model regresi linear per potongan.

Penaksir Bayes untuk parameter model regresi linear per potongan diperoleh dengan Rantai Markov tersebut.

Kata Kunci : Regresi linear per potongan, Penaksir Bayesian, Reversible jump MCMC.

ABSTRACT

Piecewise linear regression models is very flexible models for modeling the data. If the piecewise linear regression models matched against the data, then the parameters are generally not known. This paper examines the problem of picewise linear regression model parameter estimation.

The method used to estimate the parameters of picewise linear regression is Bayesian method. But the Bayes estimator can not be found analytically. To overcome these problems the proposed Reversible Jump MCMC Algorithm. Reversible Jump MCMC Algorithm generates the Markov chain converges to the limit distribution of the posterior distribution of the parameters of picewise linear regression models.

Bayes estimator for the parameters of picewise linear regression models obtained by the Markov chain.

Keywords : Piecewise linear regression models, Bayesian estimator, Reversible jump MCMC

Pendahuluan

Model regresi linear per potongan merupakan model yang sering digunakan dalam bidang ekonometri dan demografi. Dalam bidang ekonometri, model regresi linear linear per potongan digunakan untuk modelisasi pemberian komisi. Sebagai contoh, suatu perusahaan memberikan komisi pada petugas penjualannya. Perusahaan tadi membayar komisi yang didasarkan pada penjualan dengan cara sedemikian rupa sehingga sampai suatu tingkat tertentu, yang disebut ambang ada satu struktur komisi dan selewat di atas tadi struktur komisi lainnya (Gujarati, 1978). Dalam bidang demografi, model regresi lineaar per potongan dapat digunakan untuk modelisasi pertumbuhan populasi manusia atas berbagai periode waktu yang berbeda. Sebagai contoh, selama suatu interval waktu populasi menunjukkan pertumbuhan linear namun pada periode waktu yang lain menunjukkan pertumbuhan linear yang lain.

Jika model regresi linear konstan per segmen dicocokkan terhadap data, maka umumnya parameter model tidak diketahui. Mengingat begitu banyak model regresi linear konstan per potongan, di sini akan dibatasi pada model regresi linear konstan per potongan dengan galat berdistribusi normal dengan mean 0 dan variansi σ^2 di mana σ^2 adalah parameter.

Regresi linear per potongan dari variabel terikat y_t terhadap variabel bebas x_t untuk $t = 1,2,\cdots,n$ dapat dituliskan dalam persamaan stokastik berikut :

$$y_t = \begin{cases} \alpha_1^{(k)} + \beta_1^{(k)} x_t + z_t & \tau_0^{(k)} < t \le \tau_1^{(k)} \\ \alpha_2^{(k)} + \beta_2^{(k)} x_t + z_t & \tau_1^{(k)} < t \le \tau_2^{(k)} \\ & \vdots \\ \alpha_k^{(k)} + \beta_k^{(k)} x_t + z_t & \tau_{k-1}^{(k)} < t \le \tau_k^{(k)} \end{cases}$$

dengan $\tau_0^{(k)} = 0$, $\tau_k^{(k)} = n$ dan

$$\begin{array}{ll} z_t \sim N(0, \sigma_1^{2(k)}) & \tau_0^{(k)} \le t < \tau_1^{(k)} \\ z_t \sim N(0, \sigma_2^{2(k)}) & \tau_1^{(k)} \le t < \tau_2^{(k)} \\ \vdots & \\ z_t \sim N(0, \sigma_k^{2(k)}) & \tau_{k-1}^{(k)} \le t < \tau_k^{(k)} \end{array}$$

Dalam persamaan di atas : (a) k menyatakan banyaknya titik ambang, (b) $\tau_0^{(k)}, \tau_1^{(k)}, \cdots, \tau_k^{(k)}$ menyatakan titik-titik ambang yang bersesuaian, (c) $\alpha_1^{(k)}, \cdots, \alpha_k^{(k)}$ dan $\beta_1^{(k)}, \cdots, \beta_k^{(k)}$ menyatakan koefisien regresi, (d) $\sigma_1^{2(k)}, \sigma_2^{2(k)}, \cdots, \sigma_k^{2(k)}$ menyatakan variansi galat. Jika θ menyatakan parameter model regresi linear di atas, maka

$$\theta = \left(k, \tau_1^{(k)}, \cdots, \tau_k^{(k)}, \alpha_1^{(k)}, \cdots, \alpha_k^{(k)}, \beta_1^{(k)}, \cdots, \beta_k^{(k)}, \sigma_1^{2(k)}, , \cdots, \sigma_k^{2(k)}\right)$$

Misalkan $x_1, x_2, \cdots, x_n$ merupakan sampel random yang diambil dari suatu populasi yang bermodel regresi linear konstan per potongan. Berdasarkan sampel random tersebut, permasalahan utama adalah cara mengestimasi parameter θ. Parameter θ diestimasi dengan menggunakan Metode Bayesian. Kajian mengenai Metode Bayesian dapat ditemukan diberbagai literatur, misalnya Robert (2001). Estimasi parameter dengan menggunakan Metode Bayesian tidak dapat ditentukan secara analitik karena fungsi kemungkinan untuk parameter θ mempunyai bentuk yang rumit. Untuk mengatasi masalah tersebut, dalam penelitian ini digunakan Algoritma Reversible Jump MCMC (Green, 1995).

Metode Penelitian

Penelitian dimulai dengan mengkaji berbagai pustaka terkait dengan regresi linear konstan per potongan. Di samping itu, dikaji juga fungsi kemungkinan maksimum, distribusi prior, distribusi posterior, dan metode reversible jump MCMC.

Hasil dan Pembahasan

Fungsi Kemungkinan Maksimum

Untuk $i=1,2,\cdots,k$, oleh karena z_t berdistribusi normal dengan mean 0 dan variansi $\sigma_i^{2(i)}$ untuk $\tau_{i-1}^{(i)} < t \le \tau_i^{(i)}$, maka fungsi kepadatan dari z_t adalah

$$f(z_t|\ \sigma_i^{2(i)}) = \left(\frac{1}{\sqrt{2\pi\sigma_i^{2(i)}}}\right) \exp - \frac{1}{2\sigma_i^{2(i)}} z_t^2$$

Kajian mengenai fungsi kemungkinan dapat ditemukan dalam berbagai literatur, misalnya Bain and Engelhardt (1992). Sehingga untuk $z_i = (z_{\tau_{i-1}^{(i)}+1}, \cdots, z_{\tau_i^{(i)}})$ fungsi kepadatan gabungan dari z_i

$$f(z_i|\ \sigma_i^{2(i)}) = \left(\frac{1}{\sqrt{2\pi\sigma_i^{2(k)}}}\right)^{\tau_i^{(k)}-\tau_{i-1}^{(k)}}$$

$$\exp - \frac{1}{2\sigma_i^{2(k)}} \sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}} z_t^2$$

Dengan menggunakan transformasi variabel $y_t = \alpha_t^{(k)} + \beta_t^{(k)} x_t + z_t$, maka $z_t = y_t - \alpha_t^{(k)} - \beta_t^{(k)} x_t$ dan $\frac{dz_t}{dy_t} = 1$.

Sehingga fungsi kepadatan dari y_i adalah

$$f(y_i|\ \alpha_i^{(i)}, \beta_i^{(i)}, \sigma_i^{2(i)}) = \left(2\pi\sigma_i^{2(k)}\right)^{-\frac{1}{2}(\tau_i^{(k)}-\tau_{i-1}^{(k)})}$$
$$\exp - \frac{1}{2\sigma_i^{2(k)}} \sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}} (y_t - \alpha_i^{(k)} - \beta_i^{(k)} x_t)^2$$

Akhirnya fungsi kemungkinan maksimum untuk $y = (y_1, y_2, \cdots, y_n)$

$$L(y|\ \theta) = \prod_{i=1}^{k} \prod_{t=\tau_{i-1}^{(k)}+1}^{\tau_i^{(k)}} \frac{1}{\sqrt{2\pi\sigma_i^{2(k)}}}$$

$$\exp - \frac{1}{2\sigma_i^{2(k)}} (y_t - \alpha_i^{(k)} - \beta_i^{(k)} x_t)^2$$

atau

$$L(y|\ \theta) = \prod_{i=1}^{k} \left\{ (2\pi\sigma_i^{2(k)})^{-\frac{1}{2}(\tau_i^{(k)}-\tau_{i-1}^{(k)})} \right.$$
$$\left. \exp - \frac{1}{2\sigma_i^{2(k)}} \sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}} (y_t - \alpha_i^{(k)} - \beta_i^{(k)} x_t)^2 \right\}$$

Distribusi Prior

Untuk mendapatkan distribusi posterior, maka terlebih dahulu ditentukan distribusi prior untuk parameter $\theta=(k,\tau^{(k+1)},\alpha^{(k)},\beta^{(k)},\sigma^{2(k)})$, sebagai berikut :

$\pi(k)=C_k^{k_{maks}}\lambda^k(1-\lambda)^{k_{maks}-k}$ untuk $k=1,2,\cdots,k_{maks}$

$$\pi(\tau^{(k)}|\ k)=\frac{(2k+1)!}{n^{2k}}\frac{1}{2^k}\prod_{i=1}^{k}\ (\tau_i-\tau_{i-1})$$

$$\pi(\alpha^{(k)}|\ k)=\prod_{i=1}^{k}\ \frac{1}{\sqrt{2\pi a^2}}\exp-\frac{1}{2a^2}\alpha_i^2$$

$$\pi(\beta^{(k)}|\ k)=\prod_{i=1}^{k}\ \frac{1}{\sqrt{2\pi b^2}}\exp-\frac{1}{2b^2}\beta_i^2$$

$$\pi(\sigma^{2k)}|\ k)=\prod_{i=1}^{k}\ \frac{a}{2}(\sigma_i^{2(i)})^{-2}\exp-\frac{a}{2\sigma_i^{2(i)}}$$

Persoalan yang timbul, diantaranya munculnya hiperparameter $\varphi=(\lambda,c,d,a,b)$ dalam distribusi prior di atas. Selanjutnya hiperparameter φ dipandang sebagai suatu variabel acak dengan distribusi tertentu, yaitu:

$$\pi(\lambda)\propto\frac{1}{\lambda}\qquad \pi(c)\propto\frac{1}{c}$$

$$\pi(a)\propto\frac{1}{a}\qquad \pi(b)\propto\frac{1}{b}$$

Misalkan $\pi(\theta,\varphi)$ menyatakan distribusi prior untuk (θ,φ). Karena $\pi(\theta|\ \varphi)=\dfrac{\pi(\theta,\varphi)}{\varphi}$ maka distribusi prior untuk (θ,φ) dapat ditentukan sebagai berikut

$$\pi(\theta,\varphi)=\pi(\theta|\ \varphi)\pi(\varphi)$$

atau

$$\pi(\theta,\varphi)\propto C_k^{k_{maks}}\lambda^k(1-\lambda)^{k_{maks}-k}\ \frac{(2k+1)!}{n^{2k}}$$

$$\frac{1}{2^k}\prod_{i=1}^{k}\ (\tau_i-\tau_{i-1})$$

$$\prod_{i=1}^{k}\ \frac{1}{\sqrt{2\pi a^2}}\exp-\frac{1}{2a^2}\alpha_i^2$$

$$\prod_{i=1}^{k}\ \frac{1}{\sqrt{2\pi b^2}}\exp-\frac{1}{2b^2}\beta_i^2$$

$$\prod_{i=1}^{k}\ \frac{a}{2}(\sigma_i^{2(i)})^{-2}\exp-\frac{a}{2\sigma_i^{2(i)}}$$

$$\frac{1}{\lambda}\ \frac{1}{c}\ \frac{1}{a}\ \frac{1}{b}$$

Distribusi Posterior

Misalkan $\pi(\theta,\varphi|\ y)$ merupakan distribusi posterior untuk (θ,φ). Dengan menggunakan Teorema Bayes, maka distribusi posterior untuk parameter (θ,φ) dapat dinyatakan sebagai hasil kali fungsi kemungkinan dan distribusi prior

$$\pi(\theta,\varphi|\ y)\propto f(y|\ \theta)\pi(\theta,\varphi)$$

$$\propto f(y|\ \theta)\pi(\theta,\varphi)\ \pi(\theta|\ \varphi)\pi(\varphi)$$

$$\propto\prod_{i=1}^{k}\ \Big\{\ (\sigma_i^{2(k)})^{-\frac{1}{2}(\tau_i^{(k)}-\tau_{i-1}^{(k)})}$$

$$\exp-\frac{1}{2\sigma_i^{2(k)}}\sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}}\ (y_t-\alpha_i^{(k)}-\beta_i^{(k)}x_t)^2\ \Big\}$$

$$C_k^{k_{maks}}\lambda^k(1-\lambda)^{k_{maks}-k}\ \frac{(2k+1)!}{n^{2k}}$$

$$\frac{1}{2^k}\prod_{i=1}^{k}\ (\tau_i-\tau_{i-1})$$

$$\prod_{i=1}^{k} \frac{1}{\sqrt{a^2}} \exp - \frac{1}{2a^2} \alpha_i^2$$

$$\prod_{i=1}^{k} \frac{1}{\sqrt{b^2}} \exp - \frac{1}{2b^2} \beta_i^2$$

$$\prod_{i=1}^{k} \frac{a}{2} (\sigma_i^{2(i)})^{-2} \exp - \frac{a}{2\sigma_i^{2(i)}}$$

$$\frac{1}{\lambda} \quad \frac{1}{c} \frac{1}{a} \frac{1}{b}$$

Metode Reversible Jump MCMC

Misalkan $M = (\theta, \varphi)$. Secara umum, metode MCMC merupakan suatu metode sampling dengan cara membuat rantai Markov homogen $M_1, M_2, \cdots, M_m$ yang memenuhi sifat aperiodik dan irreduktibel (Robert and Casella, 1999) sedemikian hingga $M_1, M_2, \cdots, M_m$ dapat dipertimbangkan sebagai variabel acak yang berdistribusi $\pi(\theta, \varphi | \ y)$. Dengan demikian $M_1, M_2, \cdots, M_m$ dapat digunakan untuk menaksir parameter M. Untuk merealisasikan hal tersebut, diadopsi Algoritma Gibbs (Robert and Casella, 1999) yang terdiri dari dua tahap : (1) Simulasi distribusi $\pi(\varphi | \ \theta, y)$ dan (2) Simulasi distribusi $\pi(\theta | \ \varphi, y)$.

Distribusi $\pi(\varphi | \ \theta, y)$ mempunyai bentuk eksplisit. Sehingga Agoritma Gibbs dapat digunakan untuk membuat simulasi distribusi $\pi(\varphi | \ \theta, y)$. Sebaliknya, distribusi $\pi(\theta | \ \varphi, y)$ tidak mempunyai bentuk eksplisit. Sehingga simulasi eksak tidak mungkin dilakukan. Penyelesaiannya menggunakan Algoirtma Hibrida yang terdiri dari tiga tahap sebagai berikut : (2.1) Simulasi $\pi(k, \tau^{(k)} | \ \varphi, y)$ dan (2.2) Simulasi $\pi(\sigma^{2(k)} | \ k, \tau^{(k)}, \varphi, y)$. Karena k tidak diketahui maka Algoritma MCMC biasa tidak bisa digunakan untuk membuat simulasi distribusi $\pi(k, \tau^{(k)} | \ \varphi, y)$. Sebagai gantinya digunakan Algoritma Reversible Jump MCMC untuk membuat simulasi distribusi $\pi(k, \tau^{(k)} | \ \varphi, y)$. Kajian mengenai aplikasi Algoritma Reversible Jump MCMC dapat ditemukan diberbagai literatur, misalnya *Suparman et al.* (2002), Suparman (2008), Suparman (2009), Suparman (2010a), Suparman (2010b) dan Suparman (2010c). Sebaliknya distribusi $\pi(\sigma^{2(k)} | \ k, \tau^{(k)}, \varphi, y)$ mempunyai bentuk eksplisit, sehingga $\pi(\sigma^{2(k)} | \ k, \tau^{(k)}, \varphi, y)$ Algoritma Gibbs dapat digunakan untuk membuat simulasi distribusi $\pi(\sigma^{2(k)} | \ k, \tau^{(k)}, \varphi, y)$.

Kesimpulan

Dalam penelitian ini dikaji estimasi parameter model regresi linear konstan per potongan. Jika banyaknya regresi diketahui, maka estimasi parameter dapat dilakukan dengan Metode Markov Chain Monte Carlo. Namun dalam

ISSN: 2088-687X

penelitian ini, banyaknya regresi tidak diketahui. Dengan kata lain, banyaknya regresi juga merupakan variabel. Sehingga Metode Markov Chain Monte Carlo tidak dapat digunakan. Penggantinya digunakan Metode Reversible Jump Markov Chain Monte Carlo.

Ucapan Terima Kasih

Kegiatan penelitian ini tidak akan dapat berjalan baik tanpa dukungan dana dari Direktorat Penelitian dan Pengabdian kepada Masyarakat, Direktorat Jenderal pendidikan Tinggi Kementrian Pendidikan dan Kebudayaan, yang telah memberikan dana hibah penelitian melalui skema Penelitian Fundamental Tahun 2014 dengan Surat Kontrak Penelitian No. PF.02/LPP-UAD/V/2014. Untuk itu penulis mengucapkan terima kasih yang sebesar-besarnya atas bantuan dana hibah penelitian tersebut.

Pustaka

Bain, L.J. and Engelhardt, M. 1992. *Introduction to Probability and Mathematical Statistics*, Duxbury Press, California.

Gujarati, D. 1978. *Ekonometrika Dasar*, Erlangga, Jakarta.

Green, P.J. 1995. Reversible Jump MCMC Computation and Bayesian Model Determination. *Biometrika*, Vol. 82, pp. 711-732.

Robert, C.P. 2001. *The Bayesian Choice : A Decision-Theory Motivation*, Springer, New York.

Robert, C.P. and Casella, G. 1999. *Monte Carlo Statistical Methods*, Springer, New York.

Suparman, Doisy, M., and Tourneret, J.Y. 2002. Changepoint Detection using Reversible Jump MCMC Methods, *Proceedings of the IEEE ICASSP*, pp. 1569-1573.

Suparman. 2008. Pemilihan Model Bayesian Hirarki Dalam Model ARMA Menggunakan Algoritma *Reversible Jump* MCMC, *Jurnal Eksakta*, Vol. 10, No. 2, Hal 66-76.

Suparman. 2009. Estimator Bayesian Hirarki untuk Parameter Model Sinyal Multiplikatif Menggunakan Algoritma MCMC Hibrida, *Prosiding Seminar Nasional Teknologi V*, Buku 8, Hal. 41-47.

Suparman. 2010a. Segmentasi Bayesian Hirarki untuk Model MA Inversiblel Konstan Sepotong demi Sepotong Berbasis Algoritma Reversibel Jump MCMC, *Jurnal Eksakta*, Vol. 11, No. 1, Hal 9-15.

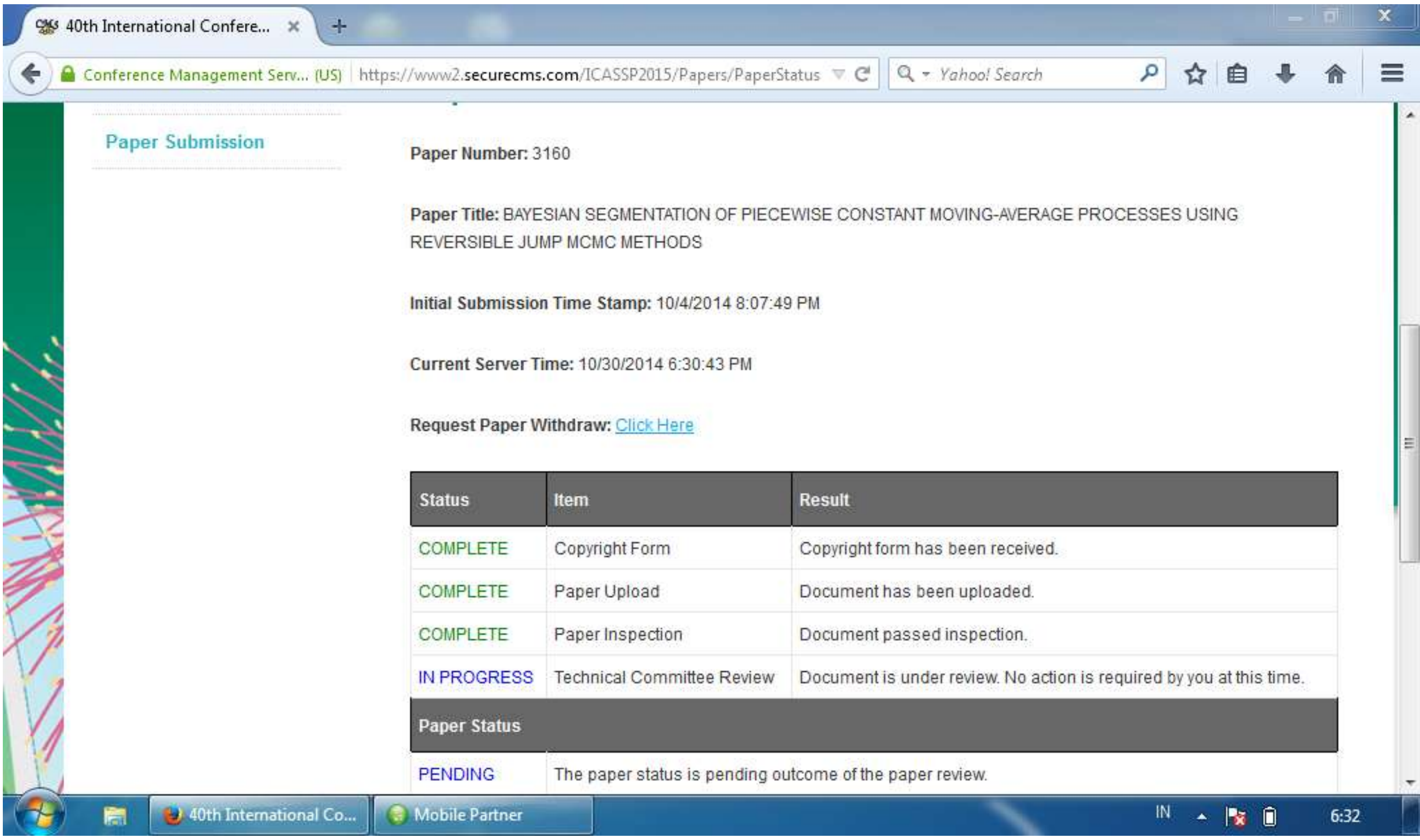
Suparman. 2010b. Inferensi Bayesian Hirarki untuk Model AR Stasioner Konstan per Segmen Menggunakan Algoritma

Reversible Jump MCMC, *Jurnal Kadikma*, Vol. 2, No 1, Hal 81-95.

Suparman, 2010c. *Pengantar Reversible Jump Markov Chain Monte Carlo dan Aplikasinya*, Penerbit UTY, Yogyakarta.

Lampiran 3 : Pembicara pada Pertemuan Ilmiah (40th IEEE International Conference on Acoustic, Speech and Signal Processing, 19-24 April 2015, Brisbane Australia), makalah dalam status sedang ditelaah.

Bukti Submit a Paper dengan judul "Bayesian Segmentation of Piecewise Constant Moving Aerage Processses Using Reversible Jump MCMC Methods" pada XV th International Conference on Acoustics, Speech, and Signal Processing



40th International Confere...

Conference Management Serv... (US) https://www2.securecms.com/ICASSP2015/Papers/PaperStatus

Yahoo! Search

Paper Submission

Paper Number: 3160

Paper Title: BAYESIAN SEGMENTATION OF PIECEWISE CONSTANT MOVING-AVERAGE PROCESSES USING REVERSIBLE JUMP MCMC METHODS

Initial Submission Time Stamp: 10/4/2014 8:07:49 PM

Current Server Time: 10/30/2014 6:30:43 PM

Request Paper Withdraw: Click Here

Status	Item	Result
COMPLETE	Copyright Form	Copyright form has been received.
COMPLETE	Paper Upload	Document has been uploaded.
COMPLETE	Paper Inspection	Document passed inspection.
IN PROGRESS	Technical Committee Review	Document is under review. No action is required by you at this time.
Paper Status		
PENDING	The paper status is pending outcome of the paper review.	

40th International Co... Mobile Partner IN 6:32

BAYESIAN SEGMENTATION OF PIECEWISE CONSTANT MOVING-AVERAGE PROCESSES USING REVERSIBLE JUMP MCMC METHODS

Suparman[1], *Michel Doisy*[2]

[1]Department of Mathematical Education, University of Ahmad Dahlan, Yogyakarta, Indonesia
[2]ENSEEIHT/TESA, Toulouse, France

ABSTRACT

This paper addresses the problem of the signal segmentation within a Bayesian framework by using reversible jump MCMC sampling. The signal is modelled by piecewise constant Moving-Average (MA) processes where the numbers of segments, the position of abrupt, the order and the coefficients of the MA processes for each segment are unknown. The reversible jump MCMC algorithm is then used to generate samples distributed according to the joint posterior distribution of the unknown parameters. These samples allow to compute some interesting features of the a posterior distribution. The performance of the this methodology is illustrated via several simulation results.

Index Terms— Bayesian model, Reversible Jump MCMC methods, Signal Segmentation, MA Processes

1. INTRODUCTION

Consider a signal y = (y_1,...,y_n) where n is the number of observations, modelled as a MA process with piecewise constant parameter and k (k = 0,...,k_{max}) changepoints. Mathematically, the model is the following:

$$y_t = z_t + \sum_{j=1}^{q_{i,k}} \phi_{i,k,j}^{(q_{i,k})} z_{t-j}, \qquad \text{..... (1)}$$

for $\tau_{i,k} \le t < \tau_{i+1,k}$ and $i = 0,1,\cdots,k$ where under the assumption of k changepoints: $\tau_{i,k}$ is the ith changepoint (defined as the index of the observation just before the ith changepoint, with the usual convention $\tau_{0,k} = 1$ and $\tau_{k+1,k} = n$, and for each ith segmen:

- $q_{i,k}$ and $\phi_{i,k}^{(q_{i,k})} = (\phi_{i,k,1}^{(q_{i,k})},\cdots,\phi_{i,k,q_{i,k}}^{(q_{i,k})})'$ are the model order and the parameter of the MA process associated with this segment.
- z_t is the Gaussian noise of variance $\sigma_{i,k}^2$ associated with the MA process in this segment, i.e. $z_t \sim N(0,\sigma_{i,k}^2)$ for $\tau_{i,k} \le t < \tau_{i+1,k}$.

As shown in [2], for each the ith segment (i = 0,...,k), the MA process of the order $q_{i,k}$ is invertible if $\phi_{i,k}^{(q_{i,k})}$ belongs to

$$I_{q,k} = \{\phi_{i,k}^{(q_{i,k})} \in \Re^{q_{i,k}} \mid 1 + \phi_{i,k,1}^{(q_{i,k})} x + \cdots$$
$$\cdots + \phi_{i,k,q_{i,k}}^{(q_{i,k})} x^{q_{i,k}} \ne 0, x \in C\}$$

This invertibility region $I_{q_{i,k}}$ is complicated to identify for $q_{i,k} > 2$. To solve this problem, a reparameterization of a $MA(q_{i,k})$ process is adopted [6]. For a $MA(q_{i,k})$ process, Monahan [6] shown that there is a one-to-one transformation

$$G : \rho_{i,k}^{(q_{i,k})} \in (-1,1)^{q_{i,k}} \to \phi_{i,k}^{(q_{i,k})} \in I_{q_{i,k}}$$

where $\rho_{i,k}^{(q_{i,k})} = (\rho_{i,k,1}^{(q_{i,k})},\cdots,\rho_{i,k.q_{i,k}}^{(q_{i,k})})'$ is the vector of the first $q_{i,k}$ inverse partial autocorrelations of the MA($q_{i,k}$) process (see [1] for a definition of the vector $\rho_{i,k}^{(q_{i,k})}$). The results of this reparametrization, the condition that $\phi_{i,k}^{(q_{i,k})} \in I_{q_{i,k}}$ become $\left|\rho_{i,k}^{(q_{i,k})}\right| < 1$ ($j = 1,\cdots,q_{i,k}$).

From the observed data y, our aim is to estimate the unknown parameter vector

$$\theta = (k,\tau^{(k)},q^{(k)},\phi^{(k)},\sigma^{(k)})'$$

where

$$\tau^{(k)} = (\tau_1,\cdots,\tau_k)',$$
$$q^{(k)} = (q_{0,k},\cdots,q_{k,k})',$$
$$\phi^{(k)} = (\phi_{0,k}^{(q_{0,k})},\cdots,\phi_{k,k}^{(q_{k,k})})',$$

and

$\sigma^{(k)} = (\sigma_{0,k}, \cdots, \sigma_{k,k})'$.

This paper formulate this problem as a Bayesian estimation problem.

2. APPROXIMATION LIKELIHOOD

In lieu of the exact likelihood, we propose to develop an approximation to the likelihood. Let $q_{\max}$ be the maximum number of orders. Since the residual sequence is Gaussian, the appoximate likelihood takes the form [3]:

$$f(s|\ \theta) = \prod_{i=0}^{k} \ (2\pi\sigma_{i,k}^2)^{-\frac{1}{2}(\tau_{i+1,k}-\tau_{i,k})} \exp - \frac{1}{2\sigma_{i,k}^2}$$

$$\sum_{t=\tau_{i,k}+1}^{\tau_{i+1,k}} \ (y_t - \sum_{j=1}^{q_{i,k}} \ G(\rho_{i,k}^{(q_{i,k})})\hat{z}_{t-j})^2 \qquad \text{..... (2)}$$

where $s = (y_{q_{\max}+1}, \cdots, y_n)'$ and with letting $\hat{z}_1 = \cdots \hat{z}_{q_{\max}} = 0$ the t[th] residual ($t = q_{\max} + 1, \cdots, n$) is calculated by

$$\hat{z}_t = y_t + \sum_{j=1}^{q_{i,k}} G(\rho_{i,k,j}^{(q_{i,k})})\hat{z}_{t-j},$$

for $\tau_{i,k} \le t < \tau_{i+1,k}$, and $i = 0,1,\cdots,k$

3. BAYESIAN INFERENCE

A Bayesian approach is adopted in this work. It implies the choice of priors. Denote $k_{\max}$ as the maximum number of changepoint. The number k of positions is drawn following a binomial distribution $B(k_{\max}, \lambda)$ with parameter $\lambda(0 \le \lambda \le 1)$ and $k_{\max}$.

For k fixed, the changepoint positions are distributed as the even numbered positions of the order statistics of $2k+1$ points uniformly drawn without repetitions in $\{2, \cdots, n\}$. This choice avoid too small interval between changes. We obtain ($k << n$ large). The model order $q_{i,k}$ ($i = 0,1,\cdots,k$) are independent with the same binomial distribution $B(q_{\max}, \mu)$ with parameter $\mu(0 \le \mu \le 1)$ and $q_{\max}$. For $q_{i,k}$ fixed, the $\rho_{i,k,j}^{(q_{i,k})}$ are independent with the same uniform distribution $\rho_{i,k,j}^{(q_{i,k})} \sim U(0-1,1)$ and the $\sigma_{i,k}^2$ are independent and distributed according to $IG(\alpha/2, \beta/2)$ ($\alpha > 0, \beta > 0$). In order to robustify the prior, we consider the hyperparameter vector $(\lambda, \mu, \alpha, \beta)$ to be random. The hyperpameters λ and μ are drawn following the same uniform prior on (0,1), i.e. $\lambda \sim U(0,1)$ and $\mu \sim U(0,1)$. Set $\alpha = 2$ and we choose a non-informative improper Jeffreys' prior for β. Then the prior distribution of the parameters θ is given by:

$$\pi(\theta;\xi) = \pi(k;\lambda)\pi(\tau^{(k)};k)\pi(q^{(k)};k)$$

$$\pi(\rho^{(k)};k;q^{(k)})\pi(\sigma^{2(k)};\alpha,\beta,k) \qquad \text{..... (3)}$$

where $\xi = (\lambda, \mu, \beta)$. By the classical Bayes's formula, the posterior distribution is

$$\pi(\theta, \xi|s) \propto f(s|\theta)\pi(\theta;\xi) \qquad \text{..... (4)}$$

that is, the product of likelihood function (2) and prior distribution (3).

Bayesian inference on (θ, ξ) is classicaly based on its posterior distribution defined by (4). In our case, however, it is not possible to obtain these quantities analytically. Therefore, we apply MCMC methods.

4. MCMC METHODS

The key idea is to build an ergodic Markov chain $\pi(\theta, \xi)_{(j)}$ ($j = 1, \cdots, M$) whose equilibrium distribution is the posterior distribution $\pi(\theta, \xi|\ s)$. This samples generated by the Markov chain can be used to estimate all posterior features of interest.

If the number of changepoints k and the model orders $q^{(k)}$ are known, then the Metropolis-Hasting algorithm [5] can be use to simulate a process according to this posterior distribution. In our case, since both k and $q^{(k)}$ unknown, the chain must jump from the model $(k, q^{(k)})$ with parameters $(\tau^{(k)}, \rho^{(k)}, \sigma^{2(k)})$ to the model $(k', q^{(k')})$ with parameters $(\tau^{(k')}, \rho^{(k')}, \sigma^{2(k')})$. Green [4] has proposed a solution to such problems of model selection. This model selection is done in two stages:

- First stage}, the reversible jump MCMC algorithm is used to define the jump between models of differing dimensionality in term of k. In this work, the moves are chosen to be: birth of a changepoint, death of a changepoint and update of the changepoint positions.
- Second stage}, for k fixed we use the reversible jump MCMC algorithm to define the jump between models of differing dimensionality, but in each term of $q_{i,k}$ ($i = 0,1,\cdots,k$). For the moves, the following transitions are used: birth of the model parameter, death of the model parameter and update of the parameter.

The technical detail of the calculus of the acceptation probability for first stage and second stage can be found in [7] and [8] respectively.

5. ESTIMATION OF THE PARAMETER

Using the Markov chain previously defined, to simulate - after a burn-in period - a random vector distributed as the posterior distribution $\pi(\theta,\xi|\ s)$.

In the proposed implementation, the samples $k_{(j)}$ from the joint posterior distribution $\pi(\theta,\xi|\ s)$ are collected after ignoring the other parameters. This strategy provides the marginal distribution of the changepoint number k. Consequently, the Marginal Maximum A Posteriori (MMAP) estimator of parameter k can be easily determined:

$$\hat{k} = \arg\max\ \hat{p}[k_{(j)} = l|s] \quad k_{(j)} \in \{0,\dots,k_{\max}\}$$

Once parameter k has been estimated, the changepoint locations and the orders can be estimated as follows:

$$\tau^{(\hat{k})} = \arg\max\ \hat{p}[\tau_{(j)}^{(k)}|k=\hat{k},s]\ \tau_{(j)}^{(k)} \in \{2,\dots,n-1\}_k$$

and

$$\hat{q}_{i,k} = \arg\max\ \hat{p}[q_{i,k_{(j)}}|k=\hat{k},s]\ q_{i,k_{(j)}} \in \{0,\dots,q_{\max}\}$$

For $k=\hat{k}$ and $q_{i,k}=\hat{q}_{i,k}\ (i=1,\dots,\hat{k})$, the MA parameter and the noise variance associated are estimated by the same way (using MMAP).

6. SIMULATION RESULTS

The simulation results are presented for a synthetic signal. 250 samples of synthetic signal are generated from a signal model (1) whose parameters are $k=1$, $\tau^{(1)}=125$, and the model order, the MA parameter and the noise variance for each segment are summarized in the Table 1.

Table 1 The synthetic MA parameter

i^{th} segment	$\sigma_{i,1}$	$q_{i,1}$	$\phi_{i,1}^{(q_{i,1})}$
0	0.5	1	(0.78)’
1	1.5	3	(0.52,-0.09,0.96)’

The MCMC simulation is run for 60.000 iterations, after a burn-in period of 10.000 iterations $(k_{\max}=10, q_{\max}=15)$.

The histogram of the marginal a posteriori distribution k is plotted Figure 1, and we obtain the MMAP of k is $\hat{k}=1$. For $k=\hat{k}$ and $q_{i,k}=\hat{q}_{i,k}\ (i=1,\dots,\hat{k})$, the MA parameter Then, for fixed $\hat{k}=k$, the histogram of the conditional marginal a posterior distribution of $\tau^{(1)}$ is given Figure 2.

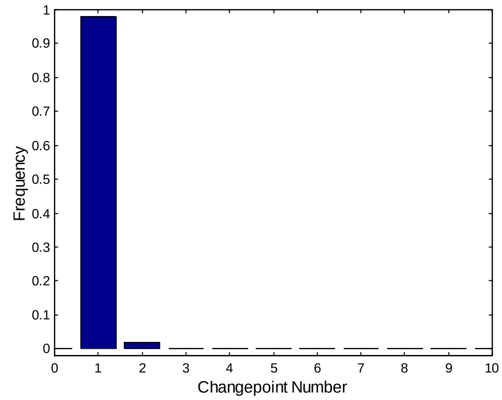


Figure 1 The histogram of the marginal a posteriori distribution k

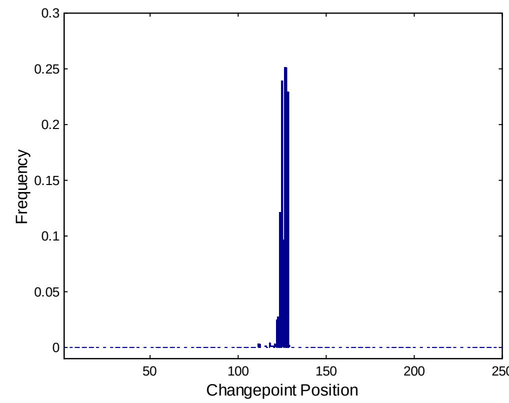


Figure 2 The histogram of the conditional marginal a posterior distribution of $\tau^{(1)}$

We obtain the MMAP of $\tau^{(1)}\ is \cdots \hat{\tau}^{(1)}=127$. This estimated changepoint and the signal are plotted Figure 3.

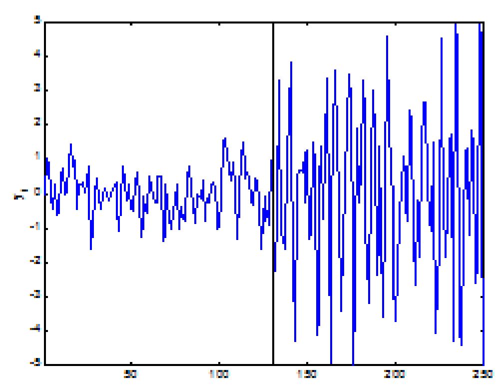


Figure 3 The estimted changepoint and the synthetic signal

We also obtain the MMAP of $q_{0,1}$ and $q_{1,1}$ are $\hat{q}_{0,1}=1$ and $\hat{q}_{1,1}=3$.

7. CONCLUSION

This paper studied new MA signal segmentation strategies based on reversible jump MCMC algorithms. The first algorithm generated Markov chain samples distributed according to the joint posterior distribution of the unknown parameters. These samples were then used to derive marginal MAP estimators. This algorithm showed good performance for image segmentation. A comparison with other existing approaches on real data is currently under investigation.

9. REFERENCES

[1]. Bhansali, R.J. 1983. The inverse partial correlation function of a time series and its applications. *J.Multivar. Anal.*, 13: 310-327.

[2]. Brockwell, P.J. and Davis, R.A. 1993. *Time Series: Theory and Methods*. Springer-Verlag, New York.

[3]. Box, G.E.P., Jenkins, G.M., Reinsel, G.C. 1994. *Time Series Analysis: forecasting and control*. Holden Day.

[4]. Green, P.J. 1994. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82**: 711-732.

[5]. Hastings, W. K. 1970. Monte Carlo sampling methods using Markov Chains and their application. *Biometrika,* **57**: 97-109.

[6]. Monahan, J.F. 1984. A note on enforcing stationarity in autoregressive-moving average models, *Biometrika*, **71**: 403-404.

[7]. Punskaya, E., Andrieu, C, Doucet, A, Fitzgerald, W.J. 2000. Bayesian Curve Fitting Using MCMC With Applications to Signal Segmentation. *IEEE Trans. on Sign. Processing*, **50**: 747-758.

[8]. Suparman and Doisy, M. 2014. Hierarchical Bayesian of ARMA Models Using Simulated Annealing Algorithm. *Telkomnika*, **12** : 87-98

Lampiran 4 : Pembicara pada Pertemuan Ilmiah (Seminar Nasional Riset Inovatif II, 21-22 November 2014, Bali), makalah dalam status sudah diterima.

Metode Reversible Jump Markov Chain Monte Carlo
Estimasi Bayesian dalam Model Regresi Linear per Potongan

Suparman [1*], Abdul Taram [2]

Program Studi Pendidikan Matematika, Universitas Ahmad Dahlan, Yogyakarta, Indonesia[1]*

Email : suparmancict@yahoo.co.id

Abstrak

Model regresi linear per potongan merupakan model yang sangat fleksibel untuk memodelkan data. Jika model regresi linear per potongan dicocokkan terhadap data, maka umumnya parameternya tidak dketahui. Tulisan ini mengkaji masalah penaksiran parameter model regresi linear per potongan. Metode yang digunakan untuk mengestimasi parameter regresi linear per potongan adalah Metode Bayesian. Namun Penaksir Bayes tidak dapat ditemukan secara analitik. Untuk mengatasi masalah tersebut diusulkan Algoritma Reversible Jump MCMC. Algoritma Reversible Jump MCMC menghasilkan Rantai Markov yang distribusi limitnya konvergen menuju distribusi posterior dari parameter model regresi linear per potongan.

Kata kunci: Regresi, Bayesian, MCMC

Abstract

The method used to estimate the parameters of picewise linear regression is Bayesian method. But the Bayes estimator can not be found analytically. To overcome these problems are proposed the Reversible Jump MCMC Algorithm. Reversible Jump MCMC Algorithm generates the Markov chain converges to the limit distribution of the posterior distribution of the parameters of picewise linear regression models. Bayes estimator for the parameters of piecewise linear regression models obtained by the Markov chain.

Keywords : Regression, Bayesian, MCMC

1. Pendahuluan

Model regresi linear per potongan merupakan model yang sering digunakan dalam berbagai bidang. Dalam bidang ekologi, model regresi linear per potongan digunakan untuk memodelkan hubungan antara suhu organisme dan suhu eksternal (Meunier, 2014). Dalam bidang ekonometri, model regresi linear linear per potongan digunakan untuk modelisasi pemberian komisi (Gujarati, 1978). Dalam bidang demografi, model regresi lineaar per potongan dapat digunakan untuk modelisasi pertumbuhan populasi manusia atas berbagai periode waktu yang berbeda. Sebagai contoh, selama suatu interval waktu populasi menunjukkan pertumbuhan linear namun pada periode waktu yang lain menunjukkan pertumbuhan linear yang lain.

Jika model regresi linear konstan per potongan dicocokkan terhadap data, maka umumnya parameter model tidak diketahui. Mengingat begitu banyak model regresi linear konstan per potongan, di sini akan dibatasi pada model regresi linear konstan per potongan dengan galat berdistribusi normal dengan mean 0 dan variansi σ^2 di mana σ^2 adalah parameter.

Regresi linear per potongan dari variabel terikat y_t terhadap variabel bebas x_t untuk $t = 1, 2, \cdots, n$ dapat dituliskan dalam persamaan berikut :

$$y_t = \begin{cases} \alpha_1^{(k)} + \beta_1^{(k)} x_t + z_t & \tau_0^{(k)} < t \le \tau_1^{(k)} \\ \alpha_2^{(k)} + \beta_2^{(k)} x_t + z_t & \tau_1^{(k)} < t \le \tau_2^{(k)} \\ \vdots & \\ \alpha_k^{(k)} + \beta_k^{(k)} x_t + z_t & \tau_{k-1}^{(k)} < t \le \tau_k^{(k)} \end{cases}$$

dengan $\tau_0^{(k)} = 0$, $\tau_k^{(k)} = n$ dan

$$z_t \sim N(0, \sigma_1^{2(k)}) \quad \tau_0^{(k)} \le t < \tau_1^{(k)}$$
$$z_t \sim N(0, \sigma_2^{2(k)}) \quad \tau_1^{(k)} \le t < \tau_2^{(k)}$$
$$\vdots$$
$$z_t \sim N(0, \sigma_k^{2(k)}) \quad \tau_{k-1}^{(k)} \le t < \tau_k^{(k)}$$

Dalam persamaan di atas :

a) k menyatakan banyaknya titik ambang
b) $\tau_0^{(k)}, \tau_1^{(k)}, \cdots, \tau_k^{(k)}$ menyatakan titik-titik ambang yang bersesuaian,
c) $\alpha_0^{(k)}, \cdots, \alpha_k^{(k)}$ dan $\beta_0^{(k)}, \beta_1^{(k)}, \cdots, \beta_k^{(k)}$ menyatakan koefisien regresi,
d) $\sigma_1^{2(k)}, \sigma_2^{2(k)}, \cdots, \sigma_k^{2(k)}$ menyatakan variansi galat.

Jika θ menyatakan parameter model regresi linear di atas, maka

$$\theta = \left(k, \tau_1^{(k)}, \cdots, \tau_k^{(k)}, \alpha_1^{(k)}, \cdots, \alpha_k^{(k)}, \beta_1^{(k)}, \cdots, \beta_k^{(k)}, \sigma_1^{2(k)}, , \cdots, \sigma_k^{2(k)}\right)$$

Misalkan $x_1, x_2, \cdots, x_n$ merupakan sampel random yang diambil dari suatu populasi yang bermodel regresi linear konstan per potongan. Berdasarkan sampel random tersebut, permasalahan utama adalah bagaimana cara mengestimasi parameter θ. Dalam paper ini, metode Bayesian digunakan untuk mengestimasi parameter θ. Kajian mengenai Metode Bayesian dapat ditemukan diberbagai literatur, misalnya Robert (2001). Namun fungsi kemungkinan untuk parameter θ mempunyai bentuk yang sangat rumit sehingga penaksir parameter θ tidak dapat ditentukan secara analitik. Untuk mengatasi masalah tersebut, dalam penelitian ini digunakan Algoritma Reversible Jump MCMC (Green, 1995)..

2. Metode

Penelitian dimulai dengan mengkaji berbagai pustaka terkait dengan regresi linear konstan per potongan. Di samping itu, dikaji juga fungsi kemungkinan maksimum, distribusi prior, distribusi posterior, dan metode reversible jump MCMC.

3. Hasil dan Pembahasan

3.1 Fungsi Kemungkinan

Untuk $i = 1, 2, \cdots, k$, oleh karena z_t berdistribusi normal dengan mean 0 dan variansi $\sigma_i^{2(i)}$ untuk $\tau_{i-1}^{(i)} < t < \tau_i^{(i)}$, maka fungsi kepadatan dari z_t berbentuk

$$f(z_t | \sigma_i^{2(i)}) = \left(\frac{1}{\sqrt{2\pi\sigma_i^{2(i)}}} \right) \exp - \frac{1}{2\sigma_i^{2(i)}} z_t^2$$

Kajian mengenai fungsi kemungkinan dapat ditemukan dalam berbagai literatur, misalnya Bain and Engelhardt (1992). Sehingga untuk $z_i = (z_{\tau_{i-1}^{(i)}+1}, \cdots, z_{\tau_i^{(i)}})$ fungsi kepadatan gabungan dari z_i dapat dituliskan sebagai berikut:

$$f(z_i | \sigma_i^{2(i)}) = \left(\frac{1}{\sqrt{2\pi\sigma_i^{2(k)}}} \right)^{\tau_i^{(k)} - \tau_{i-1}^{(k)}} \exp - \frac{1}{2\sigma_i^{2(k)}} \sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}} z_t^2$$

Dengan menggunakan transformasi variabel $y_t = \alpha_t^{(k)} + \beta_t^{(k)} x_t + z_t$, maka diperoleh $z_t = y_t - \alpha_t^{(k)} + \beta_t^{(k)} x_t$ dan $\frac{dz_t}{dy_t} = 1$. Sehingga fungsi kepadatan dari y_i dapat dituliskan sebagai berikut:

$$f(y_i | \alpha_i^{(i)}, \beta_i^{(i)}, \sigma_i^{2(i)}) = \left(2\pi\sigma_i^{2(k)}\right)^{-\frac{1}{2}(\tau_i^{(k)} - \tau_{i-1}^{(k)})} \exp - \frac{1}{2\sigma_i^{2(k)}} \sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}} (y_t - \alpha_i^{(k)} - \beta_i^{(k)} x_t)^2$$

Akhirnya fungsi kemungkinan maksimum untuk $y = (y_1, y_2, \cdots, y_n)$ dapat dituliskan sebagai berikut

$$L(y | \theta) = \prod_{i=1}^{k} \prod_{t=\tau_{i-1}^{(k)}+1}^{\tau_i^{(k)}} \frac{1}{\sqrt{2\pi\sigma_i^{2(k)}}} \exp - \frac{1}{2\sigma_i^{2(k)}} (y_t - \alpha_i^{(k)} - \beta_i^{(k)} x_t)^2$$

atau

$$L(y | \theta) = \prod_{i=1}^{k} \left\{ (2\pi\sigma_i^{2(k)})^{-\frac{1}{2}(\tau_i^{(k)} - \tau_{i-1}^{(k)})} \exp - \frac{1}{2\sigma_i^{2(k)}} \sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}} (y_t - \alpha_i^{(k)} - \beta_i^{(k)} x_t)^2 \right\}$$

3.2 Distribusi Prior

Untuk mendapatkan distribusi posterior, terlebih dahulu ditentukan distribusi prior untuk parameter

$$\theta = (k, \tau^{(k+1)}, \alpha^{(k)}, \beta^{(k)}, \sigma^{2(k)}),$$

dengan cara sebagai berikut :

$$\pi(k) = C_k^{k_{maks}} \lambda^k (1-\lambda)^{k_{maks}-k}$$

untuk $k = 1,2,\cdots,k_{maks}$

$$\pi(\tau^{(k)}| \ k) = \frac{(2k+1)!}{n^{2k}} \frac{1}{2^k} \prod_{i=1}^{k} (\tau_i - \tau_{i-1})$$

$$\pi(\alpha^{(k)}| \ k) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi a^2}} \exp - \frac{1}{2a^2} \alpha_i^2$$

$$\pi(\beta^{(k)}| \ k) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi b^2}} \exp - \frac{1}{2b^2} \beta_i^2$$

$$\pi(\sigma^{2k)}| \ k) = \prod_{i=1}^{k} \frac{a}{2} (\sigma_i^{2(i)})^{-2} \exp - \frac{a}{2\sigma_i^{2(i)}}$$

Namun timbul permasalahan baru yaitu adalah dengan hadirnya hiperparameter $\varphi = (\lambda, c, d, a, b)$ dalam distribusi prior di atas. Selanjutnya hiperparameter φ dipandang sebagai suatu variabel acak dengan distribusi tertentu, yaitu:

$$\pi(\lambda) \propto \frac{1}{\lambda} \qquad \pi(c) \propto \frac{1}{c}$$

$$\pi(a) \propto \frac{1}{a} \qquad \pi(b) \propto \frac{1}{b}$$

Misalkan $\pi(\theta,\varphi)$ menyatakan distribusi prior untuk (θ,φ). Karena $\pi(\theta| \ \varphi) = \frac{\pi(\theta,\varphi)}{\pi(\varphi)}$ maka distribusi prior untuk (θ,φ) dapat ditentukan sebagai berikut

$$\pi(\theta,\varphi) = \pi(\theta| \ \varphi)\pi(\varphi)$$

atau

$$\pi(\theta,\varphi) \propto C_k^{k_{maks}} \lambda^k (1-\lambda)^{k_{maks}-k} \frac{(2k+1)!}{n^{2k}}$$

$$\frac{1}{2^k} \prod_{i=1}^{k} (\tau_i - \tau_{i-1})$$

$$\prod_{i=1}^{k} \frac{1}{\sqrt{2\pi a^2}} \exp - \frac{1}{2a^2} \alpha_i^2$$

$$\prod_{i=1}^{k} \frac{1}{\sqrt{2\pi b^2}} \exp - \frac{1}{2b^2} \beta_i^2$$

$$\prod_{i=1}^{k} \frac{a}{2} (\sigma_i^{2(i)})^{-2} \exp - \frac{a}{2\sigma_i^{2(i)}}$$

$$\frac{1}{\lambda} \ \frac{1}{c} \ \frac{1}{a} \ \frac{1}{b}$$

3.3 Distribusi Posterior

Misalkan $\pi(\theta,\varphi| \ y)$ merupakan distribusi posterior untuk (θ,φ). Dengan menggunakan Teorema Bayes, maka distribusi posterior untuk parameter (θ,φ) dapat dinyatakan sebagai hasil kali fungsi kemungkinan dan distribusi prior

$$\pi(\theta,\varphi| \ y) \propto f(y| \ \theta)\pi(\theta,\varphi)$$

$$\propto f(y| \ \theta)\pi(\theta,\varphi) \ \pi(\theta| \ \varphi)\pi(\varphi)$$

$$\propto \prod_{i=1}^{k} \left\{ (\sigma_i^{2(k)})^{-\frac{1}{2}(\tau_i^{(k)} - \tau_{i-1}^{(k)})} \right.$$

$$\left. \exp - \frac{1}{2\sigma_i^{2(k)}} \sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}} (y_t - \alpha_i^{(k)} - \beta_i^{(k)} x_t)^2 \right\}$$

$$C_k^{k_{maks}} \lambda^k (1-\lambda)^{k_{maks}-k} \frac{(2k+1)!}{n^{2k}}$$

$$\frac{1}{2^k} \prod_{i=1}^{k} (\tau_i - \tau_{i-1})$$

$$\prod_{i=1}^{k} \frac{1}{\sqrt{a^2}} \exp - \frac{1}{2a^2} \alpha_i^2$$

$$\prod_{i=1}^{k} \frac{1}{\sqrt{b^2}} \exp - \frac{1}{2b^2} \beta_i^2$$

$$\prod_{i=1}^{k} \frac{a}{2} (\sigma_i^{2(i)})^{-2} \exp - \frac{a}{2\sigma_i^{2(i)}}$$

$$\frac{1}{\lambda} \ \frac{1}{c} \ \frac{1}{a} \ \frac{1}{b}$$

3.4 Reversible Jump MCMC

Misalkan $M(\theta,\varphi)$. Secara umum, metode MCMC merupakan suatu metode sampling dengan cara membuat rantai Markov homogen $M_1, M_2, \cdots, M_m$ yang memenuhi sifat aperiodik dan irreduktibel (Robert and Casella, 1999) sedemikian hingga $M_1, M_2, \cdots, M_m$ dapat dipertimbangkan sebagai variabel acak yang berdistribusi $\pi(\theta,\varphi| \ y)$. Dengan demikian $M_1, M_2, \cdots, M_m$ dapat digunakan untuk menaksir parameter M. Untuk merealisasikan hal tersebut, diadopsi Algoritma Gibbs (Robert and Casella, 1999) yang terdiri dari dua tahap :

(1) Simulasi distribusi $\pi(\theta,\varphi| \ y)$

(2) Simulasi distribusi $\pi(\theta,\varphi|\ y)$.

Distribusi $\pi(\theta,\varphi|\ y)$ mempunyai bentuk eksplisit. Sehingga Agoritma Gibbs dapat digunakan untuk membuat simulasi distribusi $\pi(\theta,\varphi|\ y)$. Sebaliknya, distribusi $\pi(\theta,\varphi|\ y)$ tidak mempunyai bentuk eksplisit. Sehingga simulasi eksak tidak mungkin dilakukan. Penyelesaiannya menggunakan Algoirtma Hibrida yang terdiri dari tiga tahap sebagai berikut:

(2.1) Simulasi $\pi(k,\tau^{(k)}|\ \varphi,y)$

(2.2) Simulasi $\pi(\sigma^{2(k)}|\ k,\tau^{(k)},\varphi,y)$.

Karena k tidak diketahui maka algoritma MCMC biasa tidak bisa digunakan untuk membuat suatu simulasi distribusi $\pi(k,\tau^{(k)}|\ \varphi,y)$. Sebagai gantinya, digunakan Algoritma Reversible Jump MCMC untuk membuat simulasi distribusi $\pi(k,\tau^{(k)}|\ \varphi,y)$. Kajian mengenai aplikasi Algoritma Reversible Jump MCMC dapat ditemukan diberbagai literatur, misalnya *Suparman et al.* (2002), Suparman (2008), Suparman (2009), Suparman (2010a), Suparman (2010b) dan Suparman (2010c). Sebaliknya distribusi

$$\pi(\sigma^{2(k)}|\ k,\tau^{(k)},\varphi,y)$$

mempunyai bentuk eksplisit, sehingga Algoritma Gibbs dapat digunakan untuk membuat suatu simulasi distribusi

$$\pi(\sigma^{2(k)}|\ k,\tau^{(k)},\varphi,y)$$

4. Kesimpulan

Dalam penelitian ini dikaji estimasi parameter model regresi linear konstan per potongan. Jika banyaknya regresi diketahui, maka estimasi parameter dapat dilakukan dengan Metode Markov Chain Monte Carlo. Namun dalam penelitian ini, banyaknya regresi tidak diketahui. Dengan kata lain, banyaknya regresi juga merupakan variabel. Sehingga Metode Markov Chain Monte Carlo tidak dapat digunakan. Penggantinya digunakan Metode Reversible Jump Markov Chain Monte Carlo.

5. Ucapan Terima Kasih

Kegiatan penelitian ini tidak akan dapat berjalan baik tanpa dukungan dana dari Direktorat Penelitian dan Pengabdian kepada Masyarakat, Direktorat Jenderal pendidikan Tinggi Kementrian Pendidikan dan Kebudayaan, yang telah memberikan dana hibah penelitian melalui skema Penelitian Fundamental Tahun 2014 dengan Surat Kontrak Penelitian No. PF.02/LPP-UAD/V/2014. Untuk itu penulis mengucapkan terima kasih yang sebesar-besarnya atas bantuan dana hibah penelitian tersebut.

6. Daftar Pustaka

Gujarati, D. 1978. *Ekonometrika Dasar*, Erlangga, Jakarta.

Green, P.J. 1995. Reversible Jump MCMC Computation and Bayesian Model Determination. *Biometrika*, Vol. 82, pp. 711-732.

Meunier, C.L., Malzahn, A.M., Boersma, M. 2014 A New Approach to Homeostatisc Regulation: Towards a Unified View of Physiological and Ecological Concepts. *Plos ONE*, Vol. 9 Issue 9, pp. 1-7.

Robert, C.P. 2001. *The Bayesian Choice : A Decision-Theory Motivation*, Springer, New York.

Robert, C.P. and Casella, G. 1999. *Monte Carlo Statistical Methods*, Springer, New York.

Suparman, Doisy, M., and Tourneret, J.Y. 2002. Changepoint Detection using Reversible Jump MCMC Methods, *Proceedings of the IEEE ICASSP*, pp. 1569-1573.

Suparman. 2008. Pemilihan Model Bayesian Hirarki Dalam Model ARMA Menggunakan Algoritma *Reversible Jump* MCMC, *Jurnal Eksakta*, Vol. 10, No. 2, Hal 66-76.

Suparman. 2009. Estimator Bayesian Hirarki untuk Parameter Model Sinyal Multiplikatif Menggunakan Algoritma MCMC Hibrida, *Prosiding Seminar Nasional Teknologi V*, Buku 8, Hal. 41-47.

Suparman. 2010a. Segmentasi Bayesian Hirarki untuk Model MA Inversiblel Konstan Sepotong demi Sepotong Berbasis Algoritma Reversibel Jump MCMC, *Jurnal Eksakta*, Vol. 11, No. 1, Hal 9-15.

Suparman. 2010b. Inferensi Bayesian Hirarki untuk Model AR Stasioner Konstan per Segmen Menggunakan Algoritma *Reversible Jump* MCMC, *Jurnal Kadikma*, Vol. 2, No 1, Hal 81-95.

Suparman, 2010c. *Pengantar Reversible Jump Markov Chain Monte Carlo dan Aplikasinya*, Penerbit UTY, Yogyakarta.

ACCEPTANCE LETTER

Nomor : 170.10/SENARI II/X/2014 22 Oktober 2014
Lampiran : 1 (satu) gabung
Perihal : Pemberitahuan hasil review artikel SENARI II 2014

Yth, Bapak/Ibu :
Suparman

Dengan Hormat,
Terima kasih telah mengirimkan usulan abstrak/artikel untuk dipresentasikan pada Seminar Nasional Riset Inovatif (SENARI) II Tahun 2014 yang diselenggarakan oleh Lembaga Penelitian Universitas Pendidikan Ganesha. Setelah melalui proses review, maka bersama ini kami dengan senang hati menginformasikan bahwa artikel Bapak/Ibu yang berjudul :
Estimasi Bayesian dalam Model Regresi Linear per Potongan Menggunakan Metode Reversible Jump MCMC
Dinyatakan **Dapat dimuat dengan revisi**.

Selanjutnya kami persilakan Bapak/Ibu untuk segera mengirimkan artikel yang sudah direvisi (jika ada) dalam format DOC/DOCX melalui web SENARI (http://lemlit.undiksha.ac.id/senari2014/) dengan login terlebih dahulu dan selanjutnya mengunggah pada menu "UPLOAD CAMERA READY". Artikel tersebut paling lambat kami terima tanggal 5 November 2014 dan batas pembayaran tanggal 11 November 2014 (Bagi pemakalah/peserta yang akan menginap diharuskan sudah melakukan pembayaran paling lambat tanggal 7 Nopember 2014).

Jika Bapak/Ibu membutuhkan penjelasan lebih lanjut bisa menghubungi kami melalui email senari.undiksha@gmail.com atau melalui :

- Terkait Pelaksanaan (Made Karunia 081 353 161 225 atau Wage 087 760 280 597)
- Terkait Pembayaran dan Penginapan (Ari Astrini 081 797 166 47)
- Terkait Teknis Web (Resika Arthana 085 737 515 515)

Kami tunggu kehadiran Bapak/Ibu pada acara SENARI II Tahun 2014 di Hotel Grand Inna Kuta Bali, pada tanggal 21 – 22 Nopember 2014. Terima Kasih.

Hormat Kami,
Ketua Panitia,

Dr. Gede Rasben Dantes, S.T., MTI
NIP. 197502212003121001

Sekretaris,

I Putu Ngurah Wage Myartawan, S.Pd, M.Pd
NIP. 198210052006041005

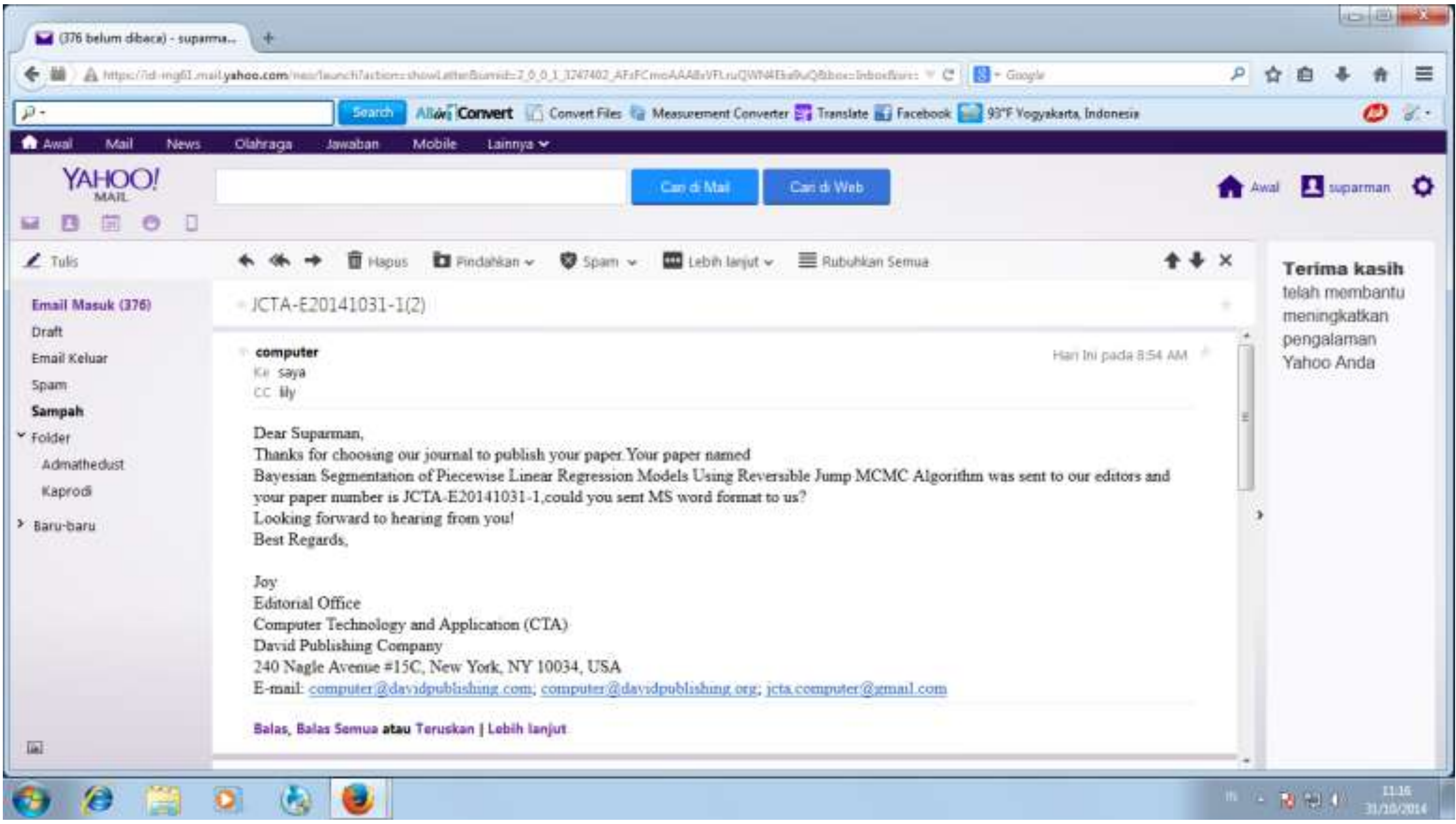
LAMPIRAN DAFTAR REVISI :

Tata tulis ada salah ketik

Lembaga Penelitian
Universitas Pendidikan Ganesha
Jalan Udayana, Kampus Tengah
Singaraja, Bali, 81116
telp. 0362-22358, Fax. 0362-22358

Lampiran 5 : Publikasi ilmiah (International Journal of Computer Technology and Application (Suparman and M. Doisy, dalam status sudah dikirim)

Bukti Pengiriman Email tentang Submit a Paper dengan judul "**Bayesian Segmentation of Piecewise Linear Regression Models Using Reversible Jump MCMC Algorithm**" ke Jurnal Internasional Computer Technology and Application (ISSN : 1934-7332)

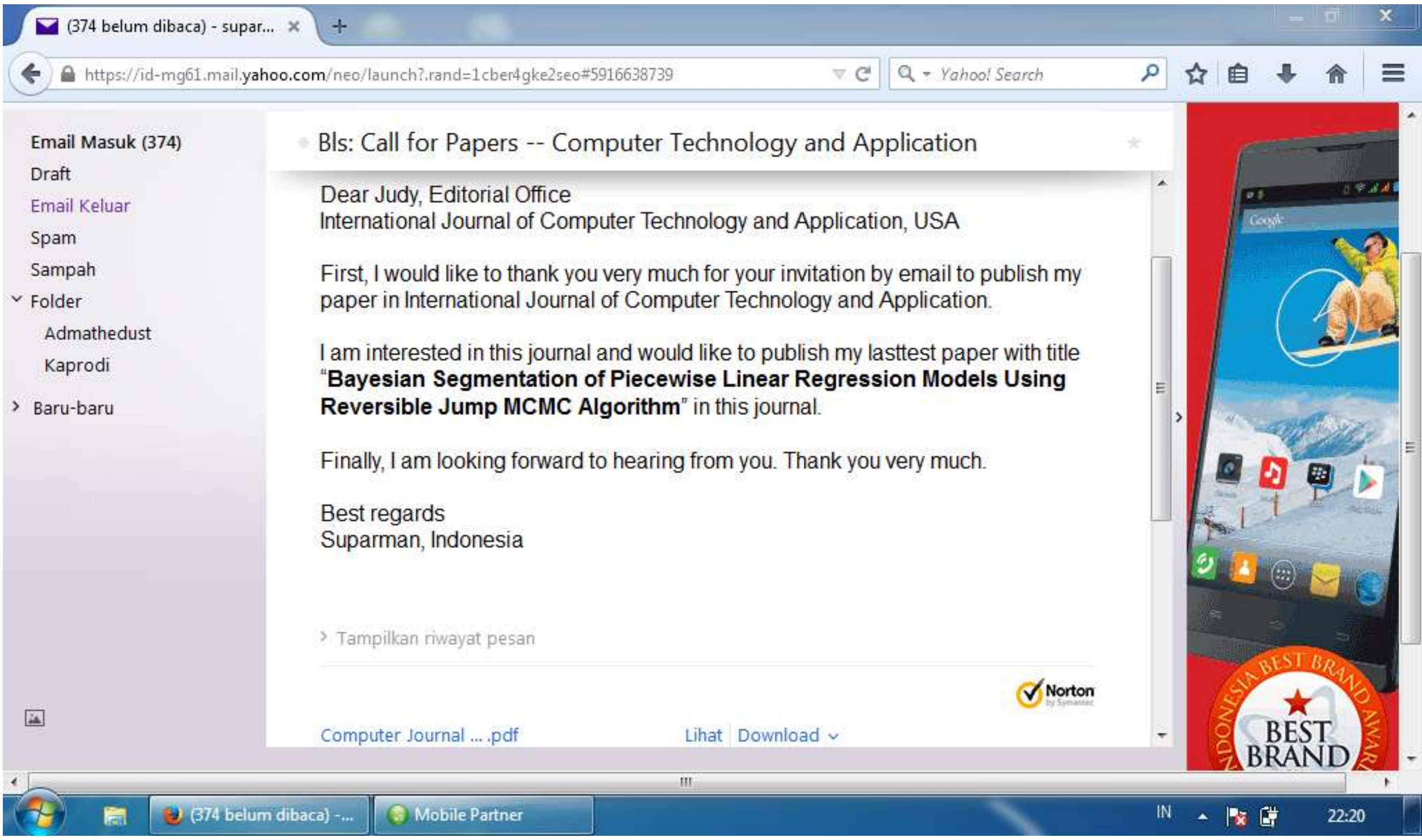

JCTA-E20141031-1(2)

computer
Ke saya
CC lily

Hari Ini pada 8:54 AM

Dear Suparman,
Thanks for choosing our journal to publish your paper.Your paper named
Bayesian Segmentation of Piecewise Linear Regression Models Using Reversible Jump MCMC Algorithm was sent to our editors and your paper number is JCTA-E20141031-1,could you sent MS word format to us?
Looking forward to hearing from you!
Best Regards,

Joy
Editorial Office
Computer Technology and Application (CTA)
David Publishing Company
240 Nagle Avenue #15C, New York, NY 10034, USA
E-mail: computer@davidpublishing.com; computer@davidpublishing.org; jcta.computer@gmail.com



Bls: Call for Papers -- Computer Technology and Application

Dear Judy, Editorial Office
International Journal of Computer Technology and Application, USA

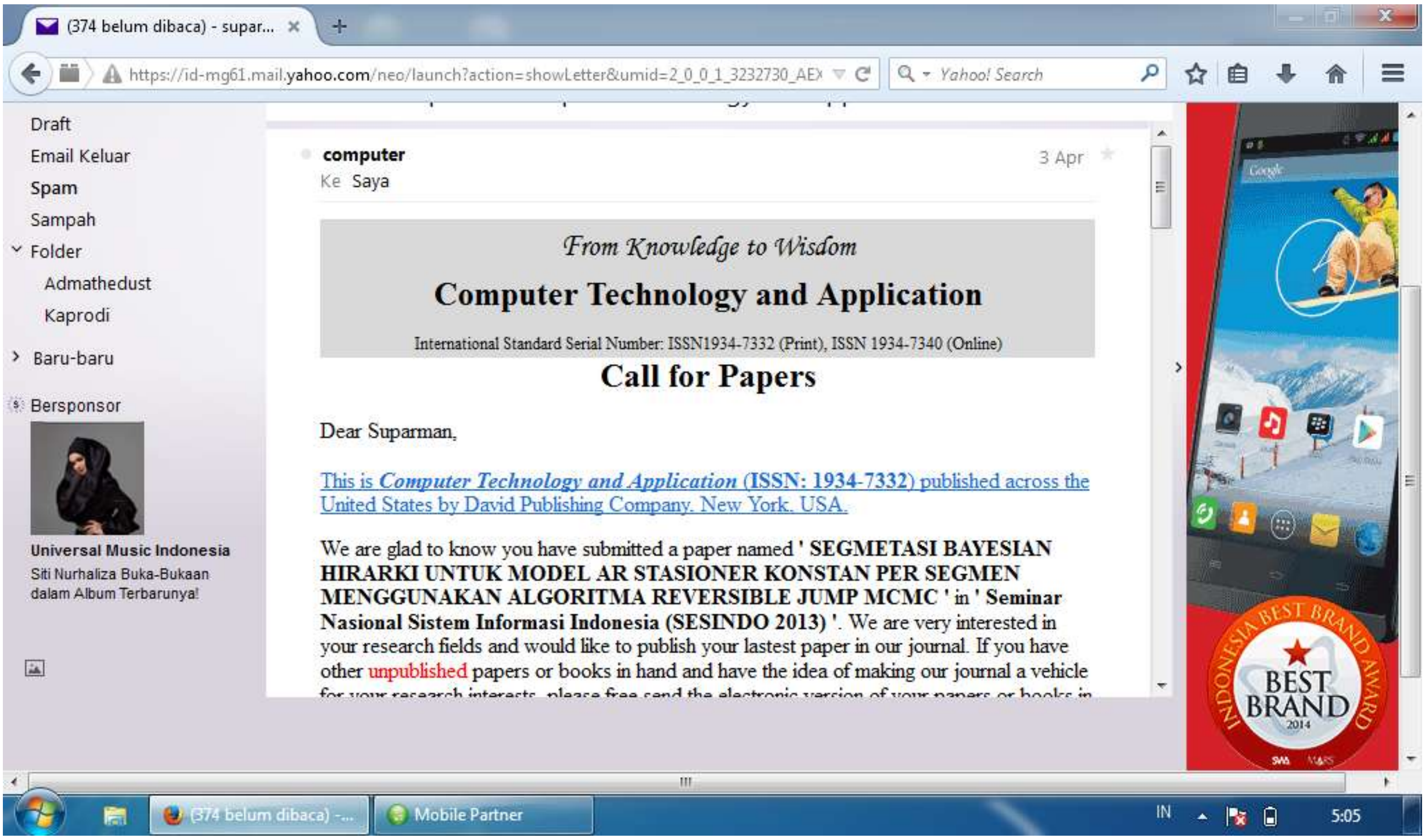
First, I would like to thank you very much for your invitation by email to publish my paper in International Journal of Computer Technology and Application.

I am interested in this journal and would like to publish my lasttest paper with title "**Bayesian Segmentation of Piecewise Linear Regression Models Using Reversible Jump MCMC Algorithm**" in this journal.

Finally, I am looking forward to hearing from you. Thank you very much.

Best regards
Suparman, Indonesia

Bukti Pengiriman Email tentang Call for Paper dari Jurnal Internasional Computer Technology and Application (ISSN : 1934-7332)



Computer Technology and Application xx (20xx) x-x

Bayesian Segmentation of Piecewise Linear Regression Models Using Reversible Jump MCMC Algorithm

Suparman[1], Michel Doisy[2]

1. Department of Mathematical Education, Faculty of Teacher Training and Education, Yogyakarta, 55161, Indonesia

2. ENSEEIHT/TESA, Toulouse, 31071, France

Received: ** **, 2014 / Accepted: ** **, 20xx / Published: xxxxx xx, 20xx.

Abstract: Piecewise linear regression models are very flexible models for modeling the data. If the piecewise linear regression models are matched against the data, then the parameters are generally not known. This paper studies the problem of parameter estimation of piecewise linear regression models. The method used to estimate the parameters of picewise linear regression models is Bayesian method. But the Bayes estimator can not be found analytically. To overcome these problems, the reversible jump MCMC algorithm are proposed. Reversible jump MCMC algorithm generates the Markov chain converges to the limit distribution of the posterior distribution of the parameters of picewise linear regression models. The resulting Markov chain is used to calculate the Bayes estimator for the parameters of picewise linear regression models.

Key words: Piecewise linear regression models, Hierarchical Bayesian, Reversible jump MCMC.

1. Introduction

Piecewise linear regression models are a model that is often used in many fiels. For example, it is used in the field of econometrics [1], geophysics [2], health [3], and ecology [4]. In the field of econometrics, piecewise linear regression models used to model the commission. For example, a company pays a commission to the sales clerk. The company was paid a commission based on the sales in a way such that up to a certain level, called the threshold one and after a commission structure on top of earlier structures commissions.

If the piecewise linear regression models are matched against the data, then the model parameters are generally unknown. There are many piecewise linear regression models. In this paper, the error distribution for each piece will be assumed has the gaussian distribution with mean 0 and variance σ^2.

For $t = 1, 2, \cdots, n$, let y_t be a dependet variable and let x_t be independent variable. Then the piecewise linear regression models can be written in the following equation:

$$y_t = \begin{cases} \alpha_1^{(k)} + \beta_1^{(k)} x_t + z_t & \tau_0^{(k)} < t \le \tau_1^{(k)} \\ \alpha_2^{(k)} + \beta_2^{(k)} x_t + z_t & \tau_1^{(k)} < t \le \tau_2^{(k)} \\ \vdots & \\ \alpha_k^{(k)} + \beta_k^{(k)} x_t + z_t & \tau_{k-1}^{(k)} < t \le \tau_k^{(k)} \end{cases}$$

with $\tau_0^{(k)} = 0$, $\tau_k^{(k)} = n$ and

$$\begin{array}{ll} z_t \sim N(0, \sigma_1^{2(k)}) & \tau_0^{(k)} \le t < \tau_1^{(k)} \\ z_t \sim N(0, \sigma_2^{2(k)}) & \tau_1^{(k)} \le t < \tau_2^{(k)} \\ \vdots & \\ z_t \sim N(0, \sigma_k^{2(k)}) & \tau_{k-1}^{(k)} \le t < \tau_k^{(k)} \end{array}$$

Fig. 1 shows the graph of the four piecewise linear regression.

Full Name, Academic title, degree, research field: signal processing.

Corresponding author: Suparman, Dr, Ph.D degree, research fields: time series, bayesian, reversible jump MCMC, signal processing. E-mail: suparmancict@yahoo.co.id

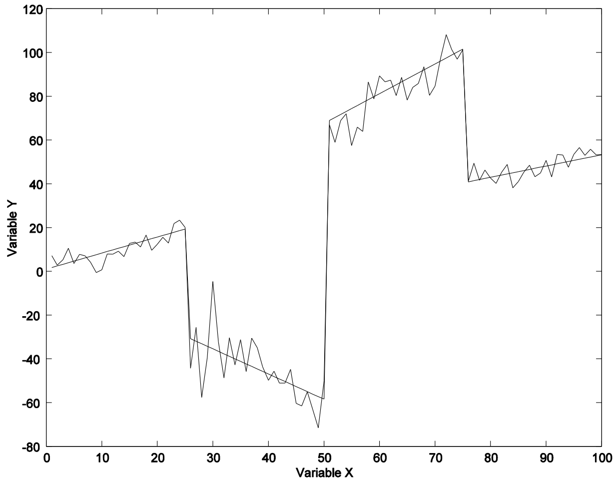


Fig. 1 Four piecewise linear regression.

In the above equation : (a) k is the number of threshold point, (b) $\tau_0^{(k)}, \tau_1^{(k)}, \cdots, \tau_k^{(k)}$ are the point corresponding threshold, (c) $\alpha_1^{(k)}, \cdots, \alpha_k^{(k)}$ and $\beta_1^{(k)}, \cdots, \beta_k^{(k)}$ are the regression coefficients, (d) $\sigma_1^{2(k)}, \sigma_2^{2(k)}, \cdots, \sigma_k^{2(k)}$ are the error variance. If θ is the pamameter of the piecewise linear regression models above, then

$$\theta = \Big(k, \tau_1^{(k)}, \cdots, \tau_k^{(k)}, \alpha_1^{(k)}, \cdots, \alpha_k^{(k)},$$

$$\beta_1^{(k)}, \cdots, \beta_k^{(k)}, \sigma_1^{2(k)}, , \cdots, \sigma_k^{2(k)}\Big)$$

Suppose $x_1, x_2, \cdots, x_n$ are a random sample drawn from a population having a piecewise linear regression models. Based on the random sample, the main problem is how to estimate the parameters θ. Parameter θ are estimated using Bayesian method. The study of the Bayesian method can be found in the literature, for example [5]. Parameter estimation using the Bayesian method can not be determined analytically because of the likelihood function for the parameter θ has a complicated shape. To overcome these problems, in this study Reversible Jump MCMC Algorithm is used.

2. Maximum Likelihood Function

For $i = 1, 2, \cdots, k$ and $\tau_{i-1}^{(i)} < t \le \tau_i^{(i)}$, suppose z_t has a gaussian distribution with mean 0 and variance $\sigma_i^{2(i)}$. Then the density function of z_t is

$$f(z_t | \ \sigma_i^{2(i)}) = \left(\frac{1}{\sqrt{2\pi\sigma_i^{2(i)}}}\right) \exp - \frac{1}{2\sigma_i^{2(i)}} z_t^2$$

Let $z_i = (z_{\tau_{i-1}^{(i)}+1}, \cdots, z_{\tau_i^{(i)}})$, then it has a joint density function as follow:

$$f(z_i | \ \sigma_i^{2(i)}) = \left(\frac{1}{\sqrt{2\pi\sigma_i^{2(k)}}}\right)^{\tau_i^{(k)} - \tau_{i-1}^{(k)}}$$

$$\exp - \frac{1}{2\sigma_i^{2(k)}} \sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}} z_t^2$$

By using variable transformation

$$y_t = \alpha_t^{(k)} + \beta_t^{(k)} x_t + z_t ,$$

then $z_t = y_t - \alpha_t^{(k)} - \beta_t^{(k)} x_t$ and $\frac{dz_t}{dy_t} = 1$. So that the density function of y_i is

$$f(y_i | \ \alpha_i^{(i)}, \beta_i^{(i)}, \sigma_i^{2(i)}) = \left(2\pi\sigma_i^{2(k)}\right)^{-\frac{1}{2}(\tau_i^{(k)} - \tau_{i-1}^{(k)})}$$

$$\exp - \frac{1}{2\sigma_i^{2(k)}} \sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}} (y_t - \alpha_i^{(k)} - \beta_i^{(k)} x_t)^2$$

Finally, the maximum likelihood function for $y = (y_1, y_2, \cdots, y_n)$ is as follow

$$L(y | \ \theta) = \prod_{i=1}^{k} \ \prod_{t=\tau_{i-1}^{(k)}+1}^{\tau_i^{(k)}} \frac{1}{\sqrt{2\pi\sigma_i^{2(k)}}}$$

$$\exp - \frac{1}{2\sigma_i^{2(k)}} (y_t - \alpha_i^{(k)} - \beta_i^{(k)} x_t)^2$$

or

$$L(y|\ \theta) = \prod_{i=1}^{k} \left\{ (2\pi\sigma_i^{2(k)})^{-\frac{1}{2}(\tau_i^{(k)}-\tau_{i-1}^{(k)})} \right.$$

$$\left. \exp - \frac{1}{2\sigma_i^{2(k)}} \sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}} (y_t - \alpha_i^{(k)} - \beta_i^{(k)} x_t)^2 \right\}$$

3. Prior Distribution

To obtain the posterior distribution, first it must to be determined the prior distribution of parameter

$$\theta = (k, \tau^{(k+1)}, \alpha^{(k)}, \beta^{(k)}, \sigma^{2(k)}),$$

as follow (see [6]):

$$\pi(k) = C_k^{k_{maks}} \lambda^k (1-\lambda)^{k_{maks}-k}$$

for $k = 1, 2, \cdots, k_{maks}$.

$$\pi(\tau^{(k)}|\ k) = \frac{(2k+1)!}{n^{2k}} \frac{1}{2^k} \prod_{i=1}^{k} (\tau_i - \tau_{i-1})$$

$$\pi(\alpha^{(k)}|\ k) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi a^2}} \exp - \frac{1}{2a^2} \alpha_i^2$$

$$\pi(\beta^{(k)}|\ k) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi b^2}} \exp - \frac{1}{2b^2} \beta_i^2$$

$$\pi(\sigma^{2k)}|\ k) = \prod_{i=1}^{k} \frac{a}{2} (\sigma_i^{2(i)})^{-2} \exp - \frac{a}{2\sigma_i^{2(i)}}$$

Let $\varphi = (\lambda, c, d, a, b)$ is hyperparameter of the prior distributions. Generally this hyperparameter is unknown. Furthermore hyperparameter φ viewed as a random variable with a certain distribution, i.e. :

$$\pi(\lambda) \propto \frac{1}{\lambda} \qquad \pi(c) \propto \frac{1}{c}$$

$$\pi(a) \propto \frac{1}{a} \qquad \pi(b) \propto \frac{1}{b}$$

Suppose $\pi(\theta,\varphi)$ expressed a prior distribution for (θ,φ). Because $\pi(\theta|\ \varphi) = \frac{\pi(\theta,\varphi)}{\varphi}$ then the prior distribution of (θ,φ) can be determined as follows

$$\pi(\theta,\varphi) = \pi(\theta|\ \varphi)\pi(\varphi)$$

or

$$\pi(\theta,\varphi) \propto C_k^{k_{maks}} \lambda^k (1-\lambda)^{k_{maks}-k} \frac{(2k+1)!}{n^{2k}}$$

$$\frac{1}{2^k} \prod_{i=1}^{k} (\tau_i - \tau_{i-1}) \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi a^2}} \exp - \frac{1}{2a^2} \alpha_i^2$$

$$\prod_{i=1}^{k} \frac{1}{\sqrt{2\pi b^2}} \exp - \frac{1}{2b^2} \beta_i^2$$

$$\prod_{i=1}^{k} \frac{a}{2} (\sigma_i^{2(i)})^{-2} \exp - \frac{a}{2\sigma_i^{2(i)}}$$

$$\frac{1}{\lambda} \ \frac{1}{c} \frac{1}{a} \frac{1}{b}$$

4. Posterior Distribution

Suppose $\pi(\theta,\varphi|\ y)$ is a posterior distribution for θ. By using Bayes theorem (θ,φ), then the posterior distribution for the parameter (θ,φ) can be expressed as the product of the likelihood function and the prior distribution

$$\pi(\theta,\varphi|\ y) \propto f(y|\ \theta)\pi(\theta,\varphi)$$

$$\propto f(y|\ \theta)\pi(\theta,\varphi)\ \pi(\theta|\ \varphi)\pi(\varphi)$$

$$\propto \prod_{i=1}^{k} \left\{ (\sigma_i^{2(k)})^{-\frac{1}{2}(\tau_i^{(k)}-\tau_{i-1}^{(k)})} \right.$$

$$\exp - \frac{1}{2\sigma_i^{2(k)}} \sum_{t=\tau_{i-1}^{(k)}}^{\tau_i^{(k)}} (y_t - \alpha_i^{(k)} - \beta_i^{(k)} x_t)^2 \ \}$$

$$C_k^{k_{maks}} \lambda^k (1-\lambda)^{k_{maks}-k} \frac{(2k+1)!}{n^{2k}}$$

$$\prod_{i=1}^{k} \frac{1}{\sqrt{a^2}} \exp - \frac{1}{2a^2} \alpha_i^2$$

$$\prod_{i=1}^{k} \frac{1}{\sqrt{b^2}} \exp - \frac{1}{2b^2} \beta_i^2$$

$$\prod_{i=1}^{k} \frac{a}{2} (\sigma_i^{2(i)})^{-2} \exp - \frac{a}{2\sigma_i^{2(i)}}$$

$$\frac{1}{\lambda} \quad \frac{1}{c} \frac{1}{a} \frac{1}{b}$$

5. Reversible Jump MCMC Algorithm

Suppose $M = (\theta, \varphi)$. In general, The MCMC algorithm is a method of sampling to make a homogeneous Markov Chain $M_1, M_2, \cdots, M_m$ that satisfies aperiodic and irreductibel [7] such that $M_1, M_2, \cdots, M_m$ can be considered as a random variable whose distribution $\pi(\theta, \varphi| \ y)$. Thus $M_1, M_2, \cdots, M_m$ it can be used to estimate parameters M. To realize this, the Gibbs algorithm is adopted [7] which consists of two phases :

(1) Simulate $\pi(\varphi| \ \theta, y)$

(2) Simulate $\pi(\theta| \ \varphi, y)$

Distribution $\pi(\varphi| \ \theta, y)$ has an axplicit form. So that the Gibbs algorithm can be used to simulate the distribution of $\pi(\varphi| \ \theta, y)$. On the contrary, the distribution $\pi(\theta| \ \varphi, y)$ has not an explicit form. So the exact simulation is possible. The solution is to use a hybrid algorithm consisting of three stages as follows :

(2.1) Simulate $\pi(\sigma^{2(k)}| \ k, \tau^{(k)}, \varphi, y)$

(2.2) Simulate $\pi(k, \tau^{(k)}| \ \varphi, y)$

The distribution $\pi(\sigma^{2(k)}| \ k, \tau^{(k)}, \varphi, y)$ has the form explicit, so that the Gibbs algorithm can be used to simulate $\pi(\sigma^{2(k)}| \ k, \tau^{(k)}, \varphi, y)$. On the contrary, because the value k is not known then the MCMC algorithm can not be used to simulate the distribution $\pi(k, \tau^{(k)}| \ \varphi, y)$. Here, reversible jump MCMC algorithm [8] is used to simulate $\pi(k, \tau^{(k)}| \ \varphi, y)$.

Let $\omega = (k, \tau)$ is the actual point of the Markov chain. There are 3 types of transformations are used, namely: the birth of the threshold point, the death of the threshold point and the change of the threshold point. Further suppose that N_k is the probability of transformation from k to k + 1, D_k is the probability of transformation from k + 1 to k, and P_k is the probability of transformation from k to k.

Birth/Death of the Threshold Point

The transdormstion of the birth of the threshold will change the number of threshold point, from k to the k

+ 1. If the birth of the threshold is selected, then the birth of the threshold from a point $\omega = (k, \tau)$ is defined in the following way: Choose a random point z in the $\{1, \cdots, n-1\} \setminus \tau$. Suppose the point z in the $\{\tau_i + 1, \cdots, \tau_{i+1} - 1\}$. Next, create a new point $\omega^* = (k+1, \tau)$ with

$$\tau_1, \ldots, \tau_i, \tau_{i+1} = z, \tau_{i+2}, \ldots, \tau_{k+2}$$

Otherwise, the transdormstion of the birth of the threshold will change the number of threshold point, from k+1 to the k . If the death of the threshold is selected, then the death of the threshold from a point $\omega^* = (k+1, \tau)$ is defined in the following way: Choose randomly a point in τ. Suppose then that point is τ_{i+1}. Next, create a new point $\omega = (k, \tau)$ with

$$\tau_1, \ldots, \tau_i, \tau_{i+2}, \ldots, \tau_{k+1}$$

Suppose that a_n and a_d are respectively a probability of acceptance for birth and death. Then the probability of acceptance for birth is as follows:

$$a_n(\omega, \omega^*) = \min\left\{1, \frac{\pi(\omega^* | \varphi, y)}{\pi(\omega | \varphi, y)} \frac{q(\omega^*, \omega)}{q(\omega, \omega^*)}\right\}$$

While the probability of death is as follows:

$$a_d(\omega, \omega^*) = \min\left\{1, \frac{1}{a_n(\omega^*, \omega)}\right\}$$

where

$$\frac{q(\omega^*, \omega)}{q(\omega, \omega^*)} = \frac{D_{k+1}}{N_k} \frac{n-1-k}{k+1}$$

Change of Threshold Point

The transformation of the change of threshol will not change the number of threshold point. But this transformation makes to change the position of the threshol point. If the change of the threshold is selected, then the change of the threshold point from a $\omega = (k, \tau)$ is defined in the following way: Choose a random point in τ. Suppose that point is τ_i. Next, create a new point $\omega^* = (k, \tau)$ where this point τ_i is replaced with z generated from the uniform distribution on the set $\{1, \cdots, n-1\} \setminus \tau$.

Let a_p is the probability of acceptance to the change. Then the probability of acceptance for change is as follows:

$$a_p(\omega, \omega^*) = \min\left\{1, \frac{\pi(\omega^* | \varphi, y)}{\pi(\omega | \varphi, y)} \frac{q(\omega^*, \omega)}{q(\omega, \omega^*)}\right\}$$

where

$$\frac{q(\omega^*, \omega)}{q(\omega, \omega^*)} = 1$$

6. Conclusion

The purpose of this study is to examine how to estimate the parameters of piecewise linear regression models when the number of regression is unknown. If the number of regression is unknown, the estimated parameters can not be done by Markov chain Monte Carlo algorithm.

The reversible jump Markov chain Monte Carlo algorithm is one of the new methods that can be used to estimate the parameters of piecewise linear regression models although the number of regression is unknown. The advantages of this method is both the number of regression and the estimation of parameter of linear regression models per piece can be estimated simultaneously.

6. Acknowledgements

This research activity will not be able to walk without the financial support of the Directorate of Research and Community Service, the Directorate General of Higher Education Ministry of Education and Culture, Republic of Indonesia, which has provided research grants through the scheme of Fundamental Research in 2014 with the Number of Research Contract : PF.02/LPP-UAD/V/2014. The authors would like to thank profusely for these research grants.

Reference

[1] D.N. Gujarati, Essensials of Econometrics, McGraw-Hill, New York, 2006.

[2] D.N. Stewart and K.A. Whaler, Optimal Piecewise Regression Analysis and Its Application to Geomagnetic Time Series, Geophysical Journal International, 1995, 710-724.

[3] H.-Y. Shi, H.-H. Lee, M.-H. Tsai, C.-C. Chiu, Y.-H. Uen and K.-T. Lee, Long-term outcomes of laparoscopic cholecystectomy: a prospective piecewise linear regression analysis, Surg Endosc, 2011, 2132-2140.

[4] J.D. Toms and M.L. Lesperance, Piecewise Regression : A Tool Identifying Ecological Thresholds, Ecology, 2003, 2034-2041.

[5] C.P. Robert, The Bayesian Choice : A Decision-Theory Motivation, Springer, New York, 2001.

[6] Suparman, M. Doisy, J.Y. Tourneret, Changepoint Detection Using Reversible Jump MCMC Methods, Proceedings of IEEE ICASSP, 2002, pp. 159-1573.

[7] C.P. Robert, G. Casella, Monte Carlo Statistical Methods, Springer, New York, 1999.

[8] P.J. Green, Reversible Jump MCMC Computation and Bayesian Model Determination, Biometrika, 1995, 711-732.