Institut Teknologi Sepuluh Nopember

Seminar Hasil Tesis

Regresi Campuran Nonparametrik Spline Linier Truncated dan Fungsi Kernel untuk Pemodelan Data Kemiskinan di Provinsi Papua

- Oleh : R o r y
- Pembimbing : Prof. Dr. I Nyoman Budiantara, M.Si.
- Co-Pembimbing : Dr. Wahyu Wibowo, M.Si.

Surabaya, 25 Januari 2016

ITS Institut Teknologi Sepuluh Nopember

Outline

PENDAHULUAN

KAJIAN PUSTAKA

METODE PENELITIAN

HASIL & PEMBAHASAN

KESIMPULAN

DAFTAR PUSTAKA

Institut
Teknologi
Sepuluh Nopember

PENDAHULUAN

Latar Belakang

Perumusan Masalah

Tujuan Penelitian

Manfaat Penelitian

Batasan Penelitian

Analisis Regresi

Regresi Nonparametrik

Memiliki fleksibilitas yang tinggi, dimana data mencari sendiri bentuk kurva regresinya tanpa dipengaruhi oleh subyektifitas peneliti (Eubank, 1999)

Regresi nonparametrik, diantaranya:

1. Spline
   Data yang perilakunya berubah-ubah pada sub-sub interval tertentu.
2. Kernel
   Data yang tidak mempunyai pola tertentu.

ITS
Institut
Teknologi
Sepuluh Nopember

Regresi Nonparametrik

Asumsi

Masing-masing prediktor dalam kurva nonparametrik multiprediktor dianggap memiliki pola yang sama.

Estimator

Menggunakan hanya satu bentuk estimator model untuk setiap variabel prediktornya.

Akibat

- Estimator kurang sesuai dengan pola data.
- Error yang besar.

Estimator Campuran Regresi Nonparametrik

Estimator campuran → Kurva dihampiri dengan estimator yang sesuai

Sudiarsa, Budiantara, Suhartono, & Purnami (2015)

- Spline Truncated
- Deret Fourier
- → Estimator Campuran Spline Truncated dan Deret Fourier

Budiantara, Ratnasari, Ratna, & Zain (2015)

- Spline Truncated
- Kernel
- → Estimator Campuran Spline Truncated dan Kernel

Institut Teknologi Sepuluh Nopember

Masalah Kemiskinan

Regresi Nonparametrik → Masalah Kemiskinan → Agenda MDGs

Indonesia
11,47 persen

Provinsi Papua
31,52 persen

ITS Institut Teknologi Sepuluh Nopember

Perumusan Masalah

Satu pola data berubah-ubah pada setiap sub-sub interval tertentu → Spline

Beberapa pola data tidak diketahui atau tidak memiliki pola tertentu → Kernel

Spline + Kernel → Estimator Campuran Nonparametrik → Kemiskinan

Tujuan Penelitian

1. Melakukan kajian mengenai estimator kurva regresi campuran nonparametrik spline dan kernel dalam model regresi campuran nonparametrik multiprediktor aditif.

2. Melakukan kajian mengenai sifat estimator kurva regresi campuran nonparametrik spline dan kernel.

3. Memodelkan kemiskinan di Provinsi Papua menggunakan model regresi campuran nonparametrik spline dan kernel.

ITS Institut Teknologi Sepuluh Nopember

Manfaat Penelitian

- Mengembangkan wawasan keilmuan bidang statistika khususnya pada keilmuan analisis regresi pendekatan spline dan kernel dalam model regresi nonparametrik.
- Memberikan model alternatif untuk menganalisis masalah kemiskinan yaitu dengan menggunakan model regresi campuran nonparametrik.

ITS Institut Teknologi Sepuluh Nopember

Batasan Masalah

- Fungsi kernel yang digunakan adalah fungsi kernel Gaussian.
- Pemilihan titik-titik knot dan bandwidth optimum menggunakan metode Generalized Cross Validation (GCV).
- Penggunaan titik knot dibatasi sampai dengan tiga titik knot.

KAJIAN PUSTAKA

Analisis Regresi

Regresi Spline Truncated

Regresi Kernel

Tinjauan Kemiskinan

ITS
Institut
Teknologi
Sepuluh Nopember

Analisis Regresi

y_i

x_i

$$y_i = f(x_i) + \varepsilon_i$$

$i = 1, 2, \ldots, n$
$\varepsilon_i \sim \mathrm{N}(\mathbf{0}, \sigma^2)$.
$f(x_i)$ merupakan kurva regresi
Bentuk model regresi tergantung pada bentuk kurva f

ITS Institut Teknologi Sepuluh Nopember

Regresi Spline Truncated

Model Regresi

$$y_i = g(u_i) + \varepsilon_i$$

Kurva Regresi Spline Truncated

$$g(u_i) = \sum_{k=0}^{p} \beta_k u_i^k + \sum_{l=0}^{q} \lambda_l (u_i - \xi_l)_+^p$$

Parameter β_k dan λ_l tidak diketahui

Bentuk Matriks

$$\tilde{g}(u) = \mathbf{G}(\tilde{\xi})\tilde{\theta}$$

Parameter $\tilde{\theta}$ tidak diketahui

Estimasi Parameter $\tilde{\theta}$

$$\hat{\tilde{\theta}} = \left[\left(\mathbf{G}(\tilde{\xi})\right)^T \mathbf{G}(\tilde{\xi})\right]^{-1} \left(\mathbf{G}(\tilde{\xi})\right)^T \tilde{y}$$

Pemilihan Titik-titik Knot Optimum

$$\mathrm{GCV}(\tilde{\xi}) = \frac{n^{-1} \sum_{i=1}^{n} \left(y_i - \hat{g}(u_i)\right)^2}{\left(n^{-1} tr\left(\mathbf{I} - \mathbf{S}(\tilde{\xi})\right)\right)^2}$$

$$\mathrm{GCV}(\tilde{\xi}_{(\mathrm{opt})}) = \underset{\tilde{\xi},\tilde{\phi}}{\mathrm{Min}}\{\mathrm{GCV}(\tilde{\xi})\}$$

Regresi Kernel

Model Regresi

$$y_i = h(v_i) + \varepsilon_i$$

Kernel Nadaraya-Watson

$$\hat{h}_\phi(v_i) = n^{-1}\sum_{i=1}^{n} W_{\phi i}(v) y_i$$

dimana

$$W_{\phi i}(v) = \frac{K_\phi(v - v_i)}{n^{-1}\sum_{i=1}^{n} K_\phi(v - v_i)}$$

$$K_\phi(v - v_i) = \frac{1}{\phi} K\left(\frac{v - v_i}{\phi}\right)$$

Pemilihan Bandwidth Optimum

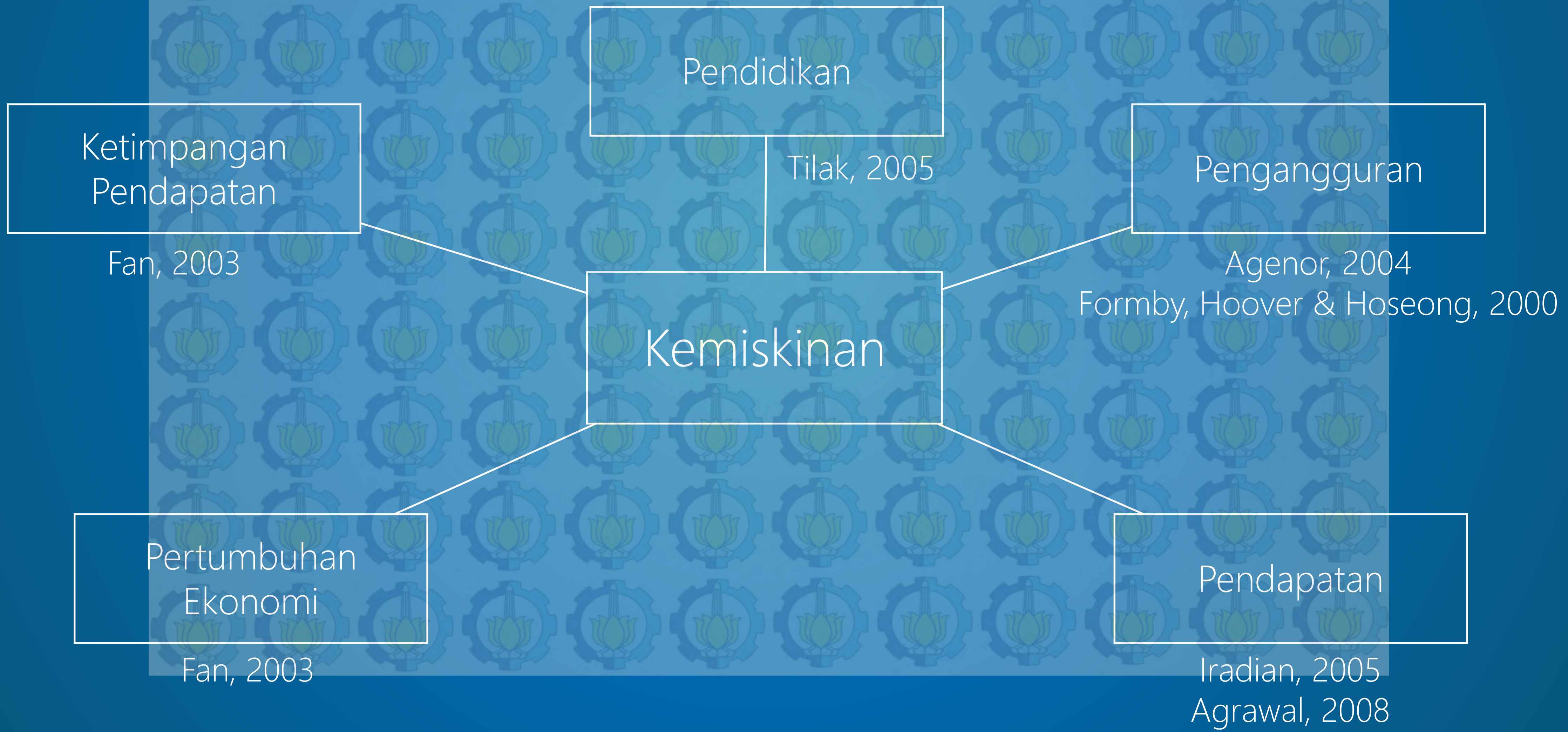
$$\text{GCV}(\phi) = \frac{n^{-1}\sum_{i=1}^{n}\left(y_i - \hat{h}_\phi(v_i)\right)^2}{\left(n^{-1} tr\left(\mathbf{I} - \mathbf{V}(\boldsymbol{\phi})\right)\right)^2}$$

$$\text{GCV}(\phi) = \underset{\tilde{\xi},\tilde{\phi}}{\text{Min}}\{\text{GCV}(\phi)\}$$

Fungsi Kernel Gaussian

$$K(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z^2\right), \quad -\infty < z < \infty$$

Tinjauan Kemiskinan



Institut Teknologi Sepuluh Nopember

METODE PENELITIAN

Sumber Data

Variabel Penelitian

Tahapan Penelitian

Jadwal Penelitian

ITS Institut Teknologi Sepuluh Nopember

Sumber Data

Data makro seluruh kabupaten/kota di Provinsi Papua kondisi tahun 2013. Jumlah kabupaten/kota di Provinsi Papua kondisi tahun 2013 adalah 29 kabupaten/kota.

iTS Institut Teknologi Sepuluh Nopember

Variabel Penelitian

Variabel Respon

- Persentase penduduk miskin

Variabel Prediktor

- Rata-rata lama sekolah
- Tingkat Pengangguran Terbuka
- Gini ratio
- Laju pertumbuhan ekonomi
- PDRB perkapita

iTS Institut Teknologi Sepuluh Nopember

Tahapan Penelitian (Tujuan 1 dan 2)

1. Menyajikan model regresi nonparametrik additif
2. Menghampiri kurva dengan spline linier truncated dan kernel Nadaraya-Watson
3. Mendapatkan estimasi kernel Nadaraya-Watson
4. Menyajikan model regresi campuran nonparametrik spline dan kernel
5. Menyajikan bentuk optimasi dari $\tilde{\theta}$ dengan menggunakan MLE
6. Menyelesaikan hasil optimasi menggunakan derivatif partial
7. Mendapatkan estimator kurva regresi campuran nonparametrik spline dan kernel
8. Mendapatkan sifat estimator kurva regresi campuran nonparametrik

ITS
Institut Teknologi Sepuluh Nopember

Tahapan Penelitian (Tujuan 3)

1. Pasangan data $(u_i, v_{1i}, v_{2i} \dots, v_{4i}, y_i)$
2. Scatter plot data
3. Menentukan variabel prediktor yang menggunakan spline dan kernel
4. Memodelkan data dengan estimator campuran
5. Memilih banyak dan lokasi titik knot optimum dan bandwidth optimum
6. Menetapkan model terbaik
7. Menghitung R^2
8. Menarik kesimpulan

HASIL DAN PEMBAHASAN

Model Regresi Campuran Nonparametrik

Estimasi Kurva Regresi

Sifat Estimator Kurva Regresi

Pemilihan Titik-titik Knot dan Bandwidth

Aplikasi Model pada Data Kemiskinan

Model Regresi Campuran Nonparametrik

Model Regresi Nonparametrik

$$y_i = f(u_i, v_{1i}, v_{2i}, \dots, v_{mi}) + \varepsilon_i \;, \quad i = 1,2,\dots,n$$

$$= f(u_i, \tilde{v}_i) + \varepsilon_i$$

Kurva Regresi Nonparametrik $f(u_i, \tilde{v}_i)$ Bersifat Aditif

$$f(u_i, \tilde{v}_i) = g(u_i) + \sum_{j=1}^{m} h_j(v_{ji})$$

- Komponen $g(u_i)$ dihampiri dengan spline linier truncated
- Komponen $h_j(v_{ji})$ dihampiri dengan kernel Nadaraya-Watson

Institut Teknologi Sepuluh Nopember

Kurva Regresi Campuran Nonparametrik

Spline Linier Truncated

$$g(u_i) = \beta_0 + \beta_1 u_i + \sum_{l=1}^{q} \lambda_l (u_i - \xi_l)_+ ,$$

dimana

$$(u_i - \xi_l)_+ = \begin{cases} (u_i - \xi_l) & , u_i \geq \xi_l \\ 0 & , u_i < \xi_l . \end{cases}$$

Kernel Nadaraya-Watson

$$\hat{h}_{j\phi_j}(v_{ji}) = n^{-1} \sum_{i=1}^{n} W_{\phi_j i}(v_j) y_i ,$$

dimana

$$W_{\phi_j i}(v_j) = \frac{K_{\phi_j}(v_j - v_{ji})}{n^{-1} \sum_{i=1}^{n} K_{\phi_j}(v_j - v_{ji})}$$

$$K_{\phi_j}(v_j - v_{ji}) = \frac{1}{\phi_j} K\left(\frac{v_j - v_{ji}}{\phi_j}\right) .$$

Estimasi Kurva Regresi (1)

Spline Linier Truncated

$$g(u_i) = \beta_0 + \beta_1 u_i + \sum_{l=1}^{q} \lambda_l (u_i - \xi_l)_+$$

$$\tilde{g}(u) = \mathbf{G}(\tilde{\xi})\tilde{\theta}$$

$$\tilde{g}(u) = \begin{bmatrix} g(u_1) \\ g(u_2) \\ \vdots \\ g(u_n) \end{bmatrix} \quad \mathbf{G}(\tilde{\xi}) = \begin{bmatrix} 1 & u_1 & (u_1 - \xi_1)_+ & \cdots & (u_1 - \xi_q)_+ \\ 1 & u_2 & (u_2 - \xi_1)_+ & \cdots & (u_2 - \xi_q)_+ \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u_n & (u_n - \xi_1)_+ & \cdots & (u_n - \xi_q)_+ \end{bmatrix} \quad \tilde{\theta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_q \end{bmatrix}$$

Estimasi Kurva Regresi (2)

Kernel Nadaraya-Watson

$$\sum_{j=1}^{m} \hat{h}_{j\phi_j}(v_{ji}) = \sum_{j=1}^{m} n^{-1} \sum_{i=1}^{n} W_{\phi_j i}(v_j) y_i$$

⇓

$$\sum_{j=1}^{m} \hat{\hat{h}}_{j\phi_j}(v_j) = \mathbf{V}(\tilde{\boldsymbol{\phi}})\tilde{y}$$

$$\hat{\hat{h}}_{j\phi_j}(v_j) = \begin{bmatrix} \hat{h}_{j\phi_j}(v_{j1}) \\ \hat{h}_{j\phi_j}(v_{j2}) \\ \vdots \\ \hat{h}_{j\phi_j}(v_{jn}) \end{bmatrix} \quad \mathbf{V}(\tilde{\boldsymbol{\phi}}) = \begin{bmatrix} n^{-1}\sum_{i=1}^{n} W_{\phi_j 1}(v_{j1}) & n^{-1}\sum_{i=1}^{n} W_{\phi_j 2}(v_{j1}) & \cdots & n^{-1}\sum_{i=1}^{n} W_{\phi_j n}(v_{j1}) \\ n^{-1}\sum_{i=1}^{n} W_{\phi_j 1}(v_{j2}) & n^{-1}\sum_{i=1}^{n} W_{\phi_j 2}(v_{j2}) & \cdots & n^{-1}\sum_{i=1}^{n} W_{\phi_j n}(v_{j2}) \\ \vdots & \vdots & \ddots & \vdots \\ n^{-1}\sum_{i=1}^{n} W_{\phi_j 1}(v_{jn}) & n^{-1}\sum_{i=1}^{n} W_{\phi_j 2}(v_{jn}) & \cdots & n^{-1}\sum_{i=1}^{n} W_{\phi_j n}(v_{jn}) \end{bmatrix} \quad \tilde{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Estimasi Kurva Regresi (3)

Model Regresi Campuran Nonparametrik

$$\tilde{y} = \mathbf{G}(\tilde{\boldsymbol{\xi}})\tilde{\theta} + \mathbf{V}(\tilde{\boldsymbol{\phi}})\tilde{y} + \tilde{\varepsilon}$$

$$\tilde{\varepsilon} \sim \mathrm{N}(0, \sigma^2\mathbf{I})$$

Fungsi Likelihood

$$L(\tilde{\theta}, \sigma^2 | \tilde{\phi}, \tilde{\xi}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\varepsilon_i^2\right)$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\|\tilde{\varepsilon}\|^2\right)$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\left\|\left(\mathbf{I} - \mathbf{V}(\tilde{\boldsymbol{\phi}})\right)\tilde{y} - \mathbf{G}(\tilde{\boldsymbol{\xi}})\tilde{\theta}\right\|^2\right)$$

Estimasi Kurva Regresi (4)

Estimator dari $\tilde{\theta}$ diperoleh dari optimasi

$$\underset{\tilde{\theta}\in R^{q+2}}{\text{Max}}\left\{L\left(\tilde{\theta},\sigma^2\middle|\tilde{\phi},\tilde{\xi}\right)\right\} = \underset{\tilde{\theta}\in R^{q+2}}{\text{Max}}\left\{(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left(-\frac{1}{2\sigma^2}\left\|\left(\mathbf{I}-\mathbf{V}(\tilde{\boldsymbol{\phi}})\right)\tilde{y}-\mathbf{G}(\tilde{\xi})\tilde{\theta}\right\|^2\right)\right\}$$

$$= \underset{\tilde{\theta}\in R^{q+2}}{\text{Max}}\left\{-\frac{n}{2}\ln(2\pi\sigma^2)-\frac{1}{2\sigma^2}\left\|\left(\mathbf{I}-\mathbf{V}(\tilde{\boldsymbol{\phi}})\right)\tilde{y}-\mathbf{G}(\tilde{\xi})\tilde{\theta}\right\|^2\right\}$$

$$= \underset{\tilde{\theta}\in R^{q+2}}{\text{Min}}\left\{\left\|\left(\mathbf{I}-\mathbf{V}(\tilde{\boldsymbol{\phi}})\right)\tilde{y}-\mathbf{G}(\tilde{\xi})\tilde{\theta}\right\|^2\right\}$$

Derivatif parsial

$$\frac{\partial}{\partial\tilde{\theta}}\left[\left\|\left(\mathbf{I}-\mathbf{V}(\tilde{\boldsymbol{\phi}})\right)\tilde{y}-\mathbf{G}(\tilde{\xi})\tilde{\theta}\right\|^2\right] = 0$$

$$\left(\mathbf{G}(\tilde{\xi})\right)^T\mathbf{G}(\tilde{\xi})\tilde{\theta} = \left(\mathbf{G}(\tilde{\xi})\right)^T\left(\mathbf{I}-\mathbf{V}(\tilde{\boldsymbol{\phi}})\right)\tilde{y}$$

Estimasi Kurva Regresi (5)

Estimator dari parameter $\tilde{\theta}$

$$\hat{\tilde{\theta}}(\tilde{\xi},\tilde{\phi}) = \left[\left(\mathbf{G}(\tilde{\xi})\right)^T \mathbf{G}(\tilde{\xi})\right]^{-1}\left(\mathbf{G}(\tilde{\xi})\right)^T\left(\mathbf{I}-\mathbf{V}(\tilde{\boldsymbol{\phi}})\right)\tilde{y}$$

$$= \mathbf{B}(\tilde{\xi},\tilde{\boldsymbol{\phi}})\tilde{y}$$

dimana

$$\mathbf{B}(\tilde{\xi},\tilde{\boldsymbol{\phi}}) = \left[\left(\mathbf{G}(\tilde{\xi})\right)^T \mathbf{G}(\tilde{\xi})\right]^{-1}\left(\mathbf{G}(\tilde{\xi})\right)^T\left(\mathbf{I}-\mathbf{V}(\tilde{\boldsymbol{\phi}})\right)$$

Estimator Kurva Regresi Spline Linier Truncated

$$\hat{\tilde{g}}_{\tilde{\phi},\tilde{\xi}}(u,\tilde{v}) = \mathbf{G}(\tilde{\xi})\hat{\tilde{\theta}}(\tilde{\xi},\tilde{\phi})$$

$$= \mathbf{S}(\tilde{\xi},\tilde{\boldsymbol{\phi}})\tilde{y}$$

dimana

$$\mathbf{S}(\tilde{\xi},\tilde{\boldsymbol{\phi}}) = \mathbf{G}(\tilde{\xi})\left[\left(\mathbf{G}(\tilde{\xi})\right)^T \mathbf{G}(\tilde{\xi})\right]^{-1}\left(\mathbf{G}(\tilde{\xi})\right)^T\left(\mathbf{I}-\mathbf{V}(\boldsymbol{\phi})\right)$$

Estimator Kurva Regresi Campuran Nonparametrik Spline dam Kernel

$$\hat{\tilde{f}}_{\tilde{\phi},\tilde{\xi}}(u,\tilde{v}) = \hat{\tilde{g}}_{\tilde{\phi},\tilde{\xi}}(u,\tilde{v}) + \sum_{j=1}^{m}\hat{\tilde{h}}_{j\phi_j}(v_j)$$

$$= \mathbf{S}(\tilde{\xi},\tilde{\boldsymbol{\phi}})\tilde{y} + \mathbf{V}(\tilde{\boldsymbol{\phi}})\tilde{y}$$

$$= \mathbf{Z}(\tilde{\xi},\tilde{\boldsymbol{\phi}})\tilde{y}$$

dimana

$$\mathbf{Z}(\tilde{\xi},\tilde{\boldsymbol{\phi}}) = \mathbf{S}(\tilde{\xi},\tilde{\boldsymbol{\phi}}) + \mathbf{V}(\tilde{\boldsymbol{\phi}})$$

Sifat Estimator Kurva Regresi

$$\mathrm{E}\left[\hat{\tilde{\theta}}(\tilde{\xi},\tilde{\phi})\right] \neq \tilde{\theta}$$

$$\mathrm{E}\left[\hat{\tilde{g}}_{\tilde{\phi},\tilde{\xi}}(u,\tilde{v})\right] \neq \tilde{g}(u)$$

$$\mathrm{E}\left[\sum_{j=1}^{m} \hat{\tilde{h}}_{j\phi_j}(v_j)\right] \neq \sum_{j=1}^{m} \tilde{h}_{j\phi_j}(v_j)$$

$$\mathrm{E}\left[\hat{\tilde{f}}_{\tilde{\phi},\tilde{\xi}}(u,\tilde{v})\right] \neq \tilde{f}_{\tilde{\phi},\tilde{\xi}}(u,\tilde{v})$$

$$\hat{\tilde{\theta}}(\tilde{\xi},\tilde{\phi}) = \mathbf{B}(\tilde{\xi},\tilde{\boldsymbol{\phi}})\tilde{y}$$

$$\hat{\tilde{g}}_{\tilde{\phi},\tilde{\xi}}(u,\tilde{v}) = \mathbf{S}(\tilde{\xi},\tilde{\boldsymbol{\phi}})\tilde{y}$$

$$\sum_{j=1}^{m} \hat{\tilde{h}}_{j\phi_j}(v_j) = \mathbf{V}(\tilde{\boldsymbol{\phi}})\tilde{y}$$

$$\hat{\tilde{f}}_{\tilde{\phi},\tilde{\xi}}(u,\tilde{v}) = \mathbf{Z}(\tilde{\xi},\tilde{\boldsymbol{\phi}})\tilde{y}$$

Pemilihan Titik Knot dan Bandwidth Optimum

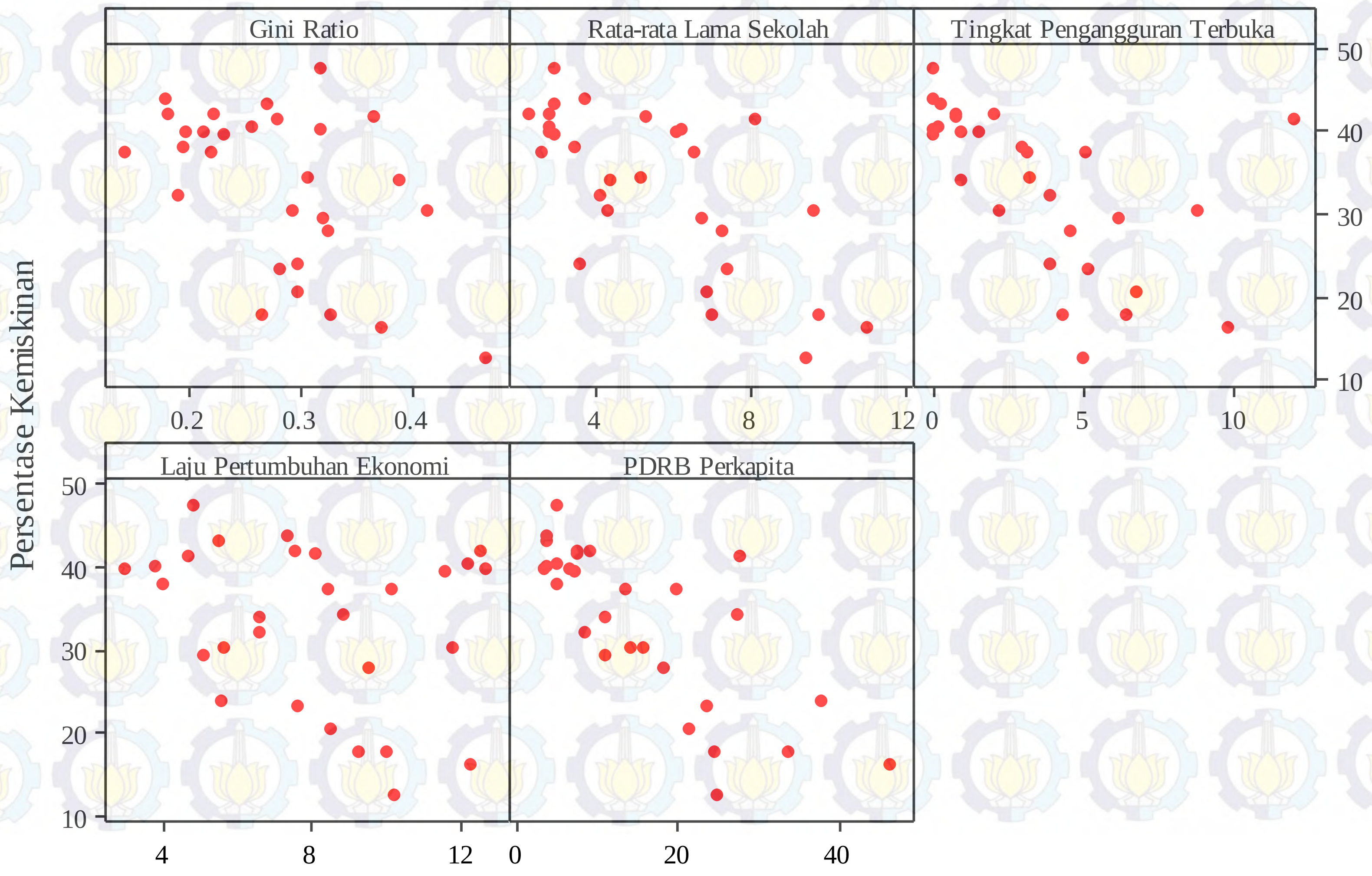
$$\mathrm{GCV}(\tilde{\xi},\tilde{\phi}) = \frac{\mathrm{MSE}(\tilde{\xi},\tilde{\phi})}{\left(n^{-1}\mathrm{trace}\left(\mathbf{I}-\mathbf{Z}(\tilde{\xi},\tilde{\phi})\right)\right)^2}$$

dimana

$$\mathrm{MSE}(\tilde{\xi},\tilde{\phi}) = n^{-1}\sum_{i=1}^{n}\left(y_i - \hat{f}_{\tilde{\phi},\tilde{\xi}}(u_i,\tilde{v}_i)\right)^2$$

$$\mathrm{GCV}\left(\tilde{\xi}_{(\mathrm{opt})},\tilde{\phi}_{(\mathrm{opt})}\right) = \underset{\tilde{\xi},\tilde{\phi}}{\mathrm{Min}}\left\{\mathrm{GCV}(\tilde{\xi},\tilde{\phi})\right\}$$

Eksplorasi Data



Statistik Deskriptif

Variabel-variabel

y = persentase penduduk miskin

u = PDRB perkapita

v_1 = gini ratio

v_2 = rata−rata lama sekolah

v_3 = tingkat pengangguran terbuka

v_4 = laju pertumbuhan ekonomi

Tabel Statistik Deskriptif Variabel Respon dan Prediktor

Statistik Deskriptif	y	u	v_1	v_2	v_3	v_4
Banyak Data	29	29	29	29	29	29
Min	12,33	3,5871	0,15	2,30	0,00	3,05
Max	47,52	46,5405	0,47	11,07	11,98	12,69
Range	35,19	42,9534	0,32	8,77	11,98	9,64
Rata-rata	33,2124	15,5921	0,2822	5,4914	3,4617	8,1034
Standar Deviasi	9,6856	11,4664	0,0765	2,5277	3,1856	2,8926

Model Regresi Campuran Nonparametrik

Model Regresi Campuran Nonparametrik Spline dan Kernel dengan q Titik Knot

$$y_i = \beta_0 + \beta_1 u_i + \lambda_1 (u_i - \xi_1)_+ + \cdots + \lambda_q (u_i - \xi_q)_+ + \sum_{i=1}^{29} \frac{\frac{1}{\phi_1} K\left(\frac{v_1 - v_{1i}}{\phi_1}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_1} K\left(\frac{v_1 - v_{1i}}{\phi_1}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_2} K\left(\frac{v_2 - v_{2i}}{\phi_2}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_2} K\left(\frac{v_2 - v_{2i}}{\phi_2}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_3} K\left(\frac{v_3 - v_{3i}}{\phi_3}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_3} K\left(\frac{v_3 - v_{3i}}{\phi_3}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_4} K\left(\frac{v_4 - v_{4i}}{\phi_4}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_4} K\left(\frac{v_4 - v_{4i}}{\phi_4}\right)} y_i + \varepsilon_i$$

Model dengan 1 Titik Knot

$$y_i = \beta_0 + \beta_1 u_i + \lambda_1 (u_i - \xi_1)_+ + \sum_{i=1}^{29} \frac{\frac{1}{\phi_1} K\left(\frac{v_1 - v_{1i}}{\phi_1}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_1} K\left(\frac{v_1 - v_{1i}}{\phi_1}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_2} K\left(\frac{v_2 - v_{2i}}{\phi_2}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_2} K\left(\frac{v_2 - v_{2i}}{\phi_2}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_3} K\left(\frac{v_3 - v_{3i}}{\phi_3}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_3} K\left(\frac{v_3 - v_{3i}}{\phi_3}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_4} K\left(\frac{v_4 - v_{4i}}{\phi_4}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_4} K\left(\frac{v_4 - v_{4i}}{\phi_4}\right)} y_i + \varepsilon_i$$

Model dengan 2 Titik Knot

$$y_i = \beta_0 + \beta_1 u_i + \lambda_1 (u_i - \xi_1)_+ + \lambda_2 (u_i - \xi_2)_+ + \sum_{i=1}^{29} \frac{\frac{1}{\phi_1} K\left(\frac{v_1 - v_{1i}}{\phi_1}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_1} K\left(\frac{v_1 - v_{1i}}{\phi_1}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_2} K\left(\frac{v_2 - v_{2i}}{\phi_2}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_2} K\left(\frac{v_2 - v_{2i}}{\phi_2}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_3} K\left(\frac{v_3 - v_{3i}}{\phi_3}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_3} K\left(\frac{v_3 - v_{3i}}{\phi_3}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_4} K\left(\frac{v_4 - v_{4i}}{\phi_4}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_4} K\left(\frac{v_4 - v_{4i}}{\phi_4}\right)} y_i + \varepsilon_i$$

Model dengan 3 Titik Knot

$$y_i = \beta_0 + \beta_1 u_i + \lambda_1 (u_i - \xi_1)_+ + \lambda_2 (u_i - \xi_2)_+ + \lambda_3 (u_i - \xi_3)_+ + \sum_{i=1}^{29} \frac{\frac{1}{\phi_1} K\left(\frac{v_1 - v_{1i}}{\phi_1}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_1} K\left(\frac{v_1 - v_{1i}}{\phi_1}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_2} K\left(\frac{v_2 - v_{2i}}{\phi_2}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_2} K\left(\frac{v_2 - v_{2i}}{\phi_2}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_3} K\left(\frac{v_3 - v_{3i}}{\phi_3}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_3} K\left(\frac{v_3 - v_{3i}}{\phi_3}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{\phi_4} K\left(\frac{v_4 - v_{4i}}{\phi_4}\right)}{\sum_{i=1}^{29} \frac{1}{\phi_4} K\left(\frac{v_4 - v_{4i}}{\phi_4}\right)} y_i + \varepsilon_i$$

ITS Institut Teknologi Sepuluh Nopember

GCV pada Model dengan Satu Titik Knot

Tabel GCV pada Model dengan Satu Titik Knot

No	Titik Knot	Bandwidth				GCV
	ξ_1	ϕ_1	ϕ_2	ϕ_3	ϕ_4	
1	14,0186	0,1139	0,0696	0,0389	0,0077	22,0305
2	16,0129	0,1160	0,0686	0,0388	0,0078	21,9732
3	18,0839	0,1155	0,0693	0,0386	0,0077	21,8351
4	20,9986	0,1153	0,0712	0,0387	0,0083	21,5841
5	**24,1434**	**0,1182**	**0,0751**	**0,0395**	**0,0077**	**21,2932**
6	25,0638	0,1192	0,0765	0,0395	0,0078	21,3468
7	26,1376	0,1165	0,0769	0,0397	0,0076	21,7099
8	27,0581	0,1143	0,0770	0,0398	0,0076	21,9926
9	28,0552	0,1123	0,0768	0,0398	0,0075	22,1998
10	30,9699	0,1125	0,0781	0,0398	0,0078	21,9427

ITS Institut Teknologi Sepuluh Nopember

GCV pada Model dengan Dua Titik Knot

Tabel GCV pada Model dengan Dua Titik Knot

No.	Titik Knot		Bandwidth				GCV
	ξ_1	ξ_2	ϕ_1	ϕ_2	ϕ_3	ϕ_4	
1	12,6984	21,8098	0,1150	0,0706	0,0388	0,0077	22,8073
2	16,6033	23,1114	0,1139	0,0726	0,0395	0,0029	22,4439
3	17,9049	27,0162	0,1135	0,0659	0,0383	0,0077	22,6556
4	20,5081	28,3178	0,1110	0,0617	0,0369	0,0078	21,9038
5	21,8098	29,6195	0,1129	0,0601	0,0373	0,0078	21,4083
6	23,1114	28,3178	0,1145	0,0557	0,0369	0,0077	19,5870
7	**25,7146**	**28,3178**	**0,1387**	**0,0540**	**0,0361**	**0,0076**	**17,1844**
8	25,7146	30,9211	0,1249	0,0671	0,0393	0,0021	21,5130
9	27,0162	34,8259	0,1109	0,0793	0,0401	0,0077	22,6506
10	28,3178	34,8259	0,1071	0,0812	0,0399	0,0078	22,4255

GCV pada Model dengan Tiga Titik Knot

Tabel GCV pada Model dengan Tiga Titik Knot

No.	Titik Knot			Bandwidth				GCV
	ξ_1	ξ_2	ξ_3	ϕ_1	ϕ_2	ϕ_3	ϕ_4	
1	9,3142	23,6320	29,3591	0,1181	0,0577	0,0383	0,0076	21,7166
2	12,1778	23,6320	26,4956	0,1065	0,0562	0,0379	0,0076	20,9848
3	15,0413	23,6320	29,3591	0,1142	0,0570	0,0382	0,0076	20,7097
4	17,9049	23,6320	26,4956	0,1016	0,0567	0,0399	0,0076	19,3065
5	20,7685	23,6320	26,4956	0,1097	0,0585	0,0448	0,0076	17,7708
6	23,6320	26,4956	40,8134	0,1019	0,0637	0,0377	0,0036	19,2524
7	**26,4956**	**29,3591**	**35,0863**	**0,1411**	**0,0676**	**0,0344**	**0,0207**	**13,4836**
8	26,4956	29,3591	40,8134	0,1357	0,0853	0,0361	0,0228	16,1498
9	29,3591	35,0863	37,9498	0,1064	0,0877	0,0401	0,0072	22,7724
10	32,2227	37,9498	43,6769	0,1065	0,0891	0,0398	0,0092	22,1233

Banyaknya Titik Knot dan Bandwidth

Perbandingan GCV Minimum

No	Model	GCV
1	1 Titik Knot 4 Bandwidth	21,2932
2	2 Titik Knot 4 Bandwidth	17,1844
3	**3 Titik Knot 4 Bandwidth**	**13,4836**

Lokasi titik-titik knot dan bandwidth optimum model dengan 3 titik knot

$\xi_1 = 26{,}4956$
$\xi_2 = 29{,}3591$
$\xi_3 = 35{,}0863$
$\phi_1 = 0{,}1411$
$\phi_2 = 0{,}0676$
$\phi_3 = 0{,}0344$
$\phi_4 = 0{,}0207$

ITS Institut Teknologi Sepuluh Nopember

Estimasi Parameter

Estimasi parameter

$\hat{\beta}_0 = 3,6965$
$\hat{\beta}_1 = -0,2832$
$\hat{\lambda}_1 = 3,6986$
$\hat{\lambda}_2 = -5,5289$
$\hat{\lambda}_3 = 2,3139$

Model Terbaik

$$\hat{f}_{\tilde{\Phi},\tilde{\xi}}(u_i, \tilde{v}_i) = 3,6965 - 0,2832u_i + 3,6986(u_i - 26,4956\,)_+ - 5,5289(u_i - 29,3591)_+ + 2,3139(u_i - 35,0863)_+ + \sum_{i=1}^{29} \frac{\frac{1}{0,1411} K\left(\frac{v_1 - v_{1i}}{0,1411}\right)}{\sum_{i=1}^{29} \frac{1}{0,1411} K\left(\frac{v_1 - v_{1i}}{0,1411}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{0,0676} K\left(\frac{v_2 - v_{2i}}{0,0676}\right)}{\sum_{i=1}^{29} \frac{1}{0,0676} K\left(\frac{v_2 - v_{2i}}{0,0676}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{0,0344} K\left(\frac{v_3 - v_{3i}}{0,0344}\right)}{\sum_{i=1}^{29} \frac{1}{0,0344} K\left(\frac{v_3 - v_{3i}}{0,0344}\right)} y_i + \sum_{i=1}^{29} \frac{\frac{1}{0,0207} K\left(\frac{v_4 - v_{4i}}{0,0207}\right)}{\sum_{i=1}^{29} \frac{1}{0,0207} K\left(\frac{v_4 - v_{4i}}{0,0207}\right)} y_i$$

Estimasi model memberikan R^2 sebesar 92,02%.

Kesimpulan

Model regresi nonparametrik memiliki kurva regresi bersifat aditif

$$f(u_i, \tilde{v}_i) = g(u_i) + \sum_{j=1}^{m} h_j(v_{ji}) \ ,$$

dimana
$g(u_i)$ dihampiri spline linier truncated,
$h_j(v_{ji})$ dihampiri kernel Nadaraya-Watson.

Estimator

$$\hat{\tilde{\theta}}(\tilde{\xi}, \tilde{\phi}) = \mathbf{B}(\tilde{\xi}, \tilde{\boldsymbol{\phi}})\tilde{y} \ , \qquad \hat{\tilde{g}}_{\tilde{\phi},\tilde{\xi}}(u, \tilde{v}) = \mathbf{S}(\tilde{\xi}, \tilde{\boldsymbol{\phi}})\tilde{y} \ ,$$

$$\sum_{j=1}^{m} \hat{\tilde{h}}_j(v_j) = \mathbf{V}(\tilde{\boldsymbol{\phi}})\tilde{y} \ , \qquad \hat{\tilde{f}}_{\tilde{\phi},\tilde{\xi}}(u, \tilde{v}) = \mathbf{Z}(\tilde{\xi}, \tilde{\boldsymbol{\phi}})\tilde{y} \ .$$

Estimator-estimator tersebut bersifat bias, namun masih merupakan kelas estimator linier.

Pemilihan banyaknya titik knot, lokasi titik-titik knot dan bandwidth optimum diperoleh dari metode **GCV**.

Model regresi campuran nonparametrik spline dan kernel tersebut diterapkan pada data kemiskinan di Provinsi Papua, dimana variabel yang digunakan adalah

y = persentase penduduk miskin
u = PDRB perkapita
v_1 = gini ratio
v_2 = rata-rata lama sekolah
v_3 = tingkat pengangguran terbuka
v_4 = laju pertumbuhan ekonomi

Model terbaik adalah model dengan 3 titik knot. Estimasi model memberikan $\mathbf{R}^2$ sebesar 92,02%.

DAFTAR PUSTAKA (1)

Agenor, R., & Pierre. (2004). *Unemployment-Poverty Trade-Offs*. Washington DC: The World Bank.

Agrawal, P. (2008). "Economic Growth and Poverty Reduction: Evidence from Kazakhstan". *Asian Development Review*, Vol. 24, No. 2, hal. 90-115.

Antoniadis, A., Gregoire, G., & McKeague, I. W. (1994). "Wavelet Methods for Curve Estimation". *Journal of the American Statistical Association*, Vol. 89, hal. 1340-1353.

Bilodeau, M. (1992). "Fourier Smoother and Additive Models". *Canadian Journal of Statistics*, Vol. 20, No.3, hal. 257-269.

Boediono. (1981). *Teori Pertumbuhan Ekonomi.* Yogyakarta: BPFE Yogyakarta.

Budiantara, I. N. (2009). *Spline Dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Surabaya: ITS Press.

Budiantara, I. N., & Mulianah. (2007). "Pemilihan Bandwidth optimum dalam Regresi Semiparametrik Kernel dan Aplikasinya". *SIGMA: Jurnal Sains dan Teknologi*, Vol. 10, No. 2, hal 159-166.

Budiantara, I. N., Diana, R., Purhadi, & Darmesto, S. (2011). "Relationship Pattern of Poverty and Unemployement in Indonesia with Bayesian Spline Approach". *International Journal of Basic & Applied Sciences IJBAS-IJENS*, Vol. 11, No.6, hal. 119-127.

Budiantara, I. N., Ratna, M., Zain, I., & Wibowo, W. (2012). "Modeling the Percentage of Poor People in Indonesia Using Spline Nonparametric Regression Approach". *International Journal of Basic & Applied Sciences IJBAS-IJENS*, Vol.12 No. 06, hal. 119-124.

Budiantara, I. N., Ratnasari, V., Ratna, M., & Zain, I. (2015). "The Combination of Spline and Kernel Estimator for Nonparametric Regression and Its Properties". *Applied Mathematical Sciences*, Vol. 9, No. 122, hal. 6083-6094.

DAFTAR PUSTAKA (2)

Cox, D. D., & O'Sullivan, F. (1996). "Penalized Type Estimator for Generalized Nonparametric Regression". *Journal of Multivariate Analysis*, Vol. 56, No. 10, hal. 185-206.

Ellies, S. (1994). *The Dimension of Poverty.* Kumarian Press.

Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing.* New York: Marcel Dekker, Inc.

Fan, S. (2003). "Public Investment and Poverty Reduction Case Studies From Asia and Implications For Latin America". *Tendencias Y Desafion Del Gato Publiko Para El Desarollo Agricola Y Rural En America Latina Y El Caribe.*

Formby, J. P., Hoover, G. A., & Hoseong, K. (2000). *Economic Growth and Poverty in the United States: Comparisons of Estimates Based Upon Official Poverty Statistics and Sen's Index of Poverty.* Working Paper No. 00-11-01, Univ. of Alabama, Department of Economics, Finance, and Legal Studies.

Gasser, T., & Muller, H.-G. (1979). *Kernel Estimation of Regression Functions.* Springer Berlin Heidelberg.

Hardle, W. (1994). *Applied Nonparametric Regression.* Berlin: Humboldt-Universität zu Berlin.

Iradian, G. (2005). *Inequality, Poverty, and Growth: Cross-Country Evidence.* IMF Working Paper, 1-39.

Islamiyati, A., & Budiantara, I. N. (2007). "Model Spline dengan Titik-titik Knots dalam Regresi Nonparametrik". *Jurnal INFERENSI*, Vol. 3, hal. 11-21.

Jalan, J., & Ravallion, M. (2003). "Estimating The Benefit Incidence of An Antipoverty Program". *Journal of Business & Economic Statistics*, Vol. 21, No. 1, hal. 19-30.

Jung, S. (2009). *Analyzing Poverty in the Southern United States.* University of T ennessee.

Institut Teknologi Sepuluh Nopember

DAFTAR PUSTAKA (3)

Klemela, J. (2014). *Multivariate Nonparametric Regression and Visualization: with R and Applications to Finance.* New Jersey: John Wiley & Sons, Inc.

Lestari, B., & Budiantara, I. N. (2010). "Spline Estimator of Triple Response Nonparametric Regression Model". *Jurnal Ilmu Dasar*, Vol. 11, hal. 17-22.

Merdekawati, I. P., & Budiantara, I. N. (2013). "Pemodelan Regresi Spline Truncated Multivariabel pada Faktor-faktor yang Mempengaruhi Kemiskinan di Kabupaten/Kota Provinsi Jawa Tengah". *Jurnal Sains dan POMITS*, Vol. 2, No. 1, hal. 2337-3520.

Muller, C. (1997). *Transient Seasonal and Chronic Poverty of Peasants Evidence from Rwanda.* Oxford: Centre for the Study of African Economies.

Nadaraya, E. A. (1989). *Nonparametric Estimation of Probability Densities and Regression Curves.* Kluwcer Academic Publishers.

Naja, A. H. (2006). "Pendidikan Berkualitas dan Pembangunan Sumber Daya Manusia: Solusi Utama Masalah Pengangguran dan Kemiskinan di Indonesia". *Jurnal Bisnis dan Ekonomi Politik*, Vol. 7 No. 1, hal. 67-79.

Ratnasari, V., Budiantara, I. N., Zain, I., Ratna, M., & Mirah, M. N. (2015, Agustus). "Comparison Truncated Spline and Fourier Series in Multivariable Nonparametric Regression Models (Applications: Data of Poverty in Papua, Indonesia)". *International Journal of Basic & Applied Sciences IJBAS-IJENS*, Vol. 15 No. 04, hal. 9-12.

Reinsch, C. H. (1967). Smoothing by Spline Functions. *Numerische Mathematik*, hal. 77-183.

Rencher, A. C., & Schaalje, G. B. (2008). *Linears Models in Statistics* (2nd ed.). New Jersey: A John Wiley & Sons.

Sahdan, G. (2005). "Menanggulangi Kemiskinan di Desa". *Jurnal Ekonomi Rakyat dan Kemiskinan*, Vol. 3, hal. 1-10.

DAFTAR PUSTAKA (4)

Sala-i-Martin, X. (2006). "The World Distribution of Income: Falling Poverty and... Convergence, Period". *The Quarterly Journal of Economics*, Vol. 121, No. 2, hal. 351-397.

Silverman, B. W. (1985). "Some Aspects of The Spline Smoothing Approach to Non-parametric Regression Curve Fitting". *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 47, No. 1, hal. 1-52.

Stich, A. (1997). *Poverty and Life Cycle Effects: A Nonparametric Analysis for Germany.* (No. 5/96 [rev.]). Discussion Papers in Statistics and Econometrics.

Sudiarsa, I. W., Budiantara, I. N., Suhartono, & Purnami, S. W. (2015). "Combined Estimator Fourier Series and Spline Truncated Multivariable Nonparametric Regression". *Applied Mathematical Sciences*, Vol. 9, No. 100, hal. 4997-5010.

Suharto, E. (2009). *Kemiskinan dan Perlindungan Sosial di Indonesia: Menggagas Model Jaminan Sosial Universal Bidang Kesehatan.* Bandung: Alfabeta.

Tilak, J. B. (2005). *Post-Elementary Education, Poverty and Development in India.* New Delhi: Working Paper Series - No. 6, Centre of African Studies, University of Edinburgh.

Wahba, G. (1990). *Spline Models for Observational Data.* Philadelphia: Society for Industrial and Applied Mathematics.

Zhang, Y., & Wan, G. (2006). *Globalization and The Urban Poor in China.* Helsinki: UNU-WIDER Working Paper No. 42, United Nations University, World Institute for Development Economics Research.

Institut Teknologi Sepuluh Nopember

Seminar Hasil Tesis

Regresi Campuran Nonparametrik Spline Linier Truncated dan Fungsi Kernel untuk Pemodelan Data Kemiskinan di Provinsi Papua

- Oleh : R o r y
- Pembimbing : Prof. Dr. I Nyoman Budiantara, M.Si.
- Co-Pembimbing : Dr. Wahyu Wibowo, M.Si.

Surabaya, 25 Januari 2016