ESTIMASI KURVA REGRESI NONPARAMETRIK DENGAN MENGGUNAKAN METODE FOURIER

(Skripsi)

Oleh

Neli Rohmatilah

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS LAMPUNG

BANDAR LAMPUNG

2019

ABSTRACT

ESTIMATION OF NONPARAMETRIC REGRESSION CURVE USING FOURIER METHOD

By

Neli Rohmatilah

This study aims to determine the optimal bandwidth in the nonparametric regression of the Fourier method using the Generalized Cross Validation (GCV) and to estimate the nonparametric regression curve in the generation data with different errors. The results of this study indicate that the estimation of the nonparametric regression curve uses the Fourier method on the periodic wave function and linearly affected by an error. The smaller the error used, the better the estimation curve obtained. Estimation of the Fourier method with optimal bandwidth selection, the smaller the bandwidth produced, the smoother the regression curve obtained or vice versa.

Keywords: Nonparametic Regression, Fourier Series Estimator, *Generalized Cross Validation* (GCV), Error

ABSTRAK

ESTIMASI KURVA REGRESI NONPARAMETRIK DENGAN MENGGUNAKAN METODE FOURIER

Oleh

Neli Rohmatilah

Penelitian ini bertujuan untuk menentukan *bandwitdth* optimal pada regresi nonparametrik metode Fourier dengan menggunakan *Generalized Cross Validation* (GCV) dan untuk mengestimasi kurva regresi nonparametrik pada data bangkitan dengan galat berbeda.Hasil penelitian ini menunjukkan bahwa estimasi kurva regresi nonparametrik menggunakan metode Fourier pada fungsi gelombang periodik dan linear dipengaruhi oleh galat. Semakin kecil galat yang digunakan, maka semakin baik kurva estimasi yang didapat. Estimasi metode Fourier dengan pemilihan *bandwidth* yang optimal, semakin kecil *bandwidth* yang dihasilkan maka semakin mulus kurva regresi yang didapat atau sebaliknya.

Kata kunci: Regresi Nonparametik, Penduga Deret Fourier, *Generalized Cross Validation (GCV)*, Galat

ESTIMASI KURVA REGRESI NONPARAMETRIK DENGAN MENGGUNAKAN METODE FOURIER

Oleh

Neli Rohmatilah

Skripsi
Sebagai salah satu syarat untuk mencapai gelar
Sarjana Sains

Pada
Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2019

Judul Skripsi	: **ESTIMASI KURVA REGRESI NONPARAMETRIK DENGAN MENGGUNAKAN METODE FOURIER**
Nama Mahasiswa	: *Neli Rohmatilah*
No. Pokok Mahasiswa	: 1517031005
Jurusan	: Matematika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam

MENYETUJUI

1. Komisi Pembimbing

Ir. Netti Herawati, M.Sc., Ph.D.	**Dra. Dorrah Aziz, M.Si.**
NIP 19650125 199003 2 001	NIP 19610128 198811 2 001

2. Ketua Jurusan Matematika

Prof. Dra. Wamiliana, M.A., Ph.D.
NIP 19631108 198902 2 001

MENGESAHKAN

1. Tim Penguji

Ketua	: **Ir. Netti Herawati, M.Sc., Ph.D.**	....
Sekretaris	: **Dra. Dorrah Aziz, M.Si.**	....
Penguji Bukan Pembimbing	: **Widiarti, S.Si., M.Si.**	....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Drs. Suratman, M.Sc.
NIP 19640604 199003 1 002

Tanggal Lulus Ujian Skripsi : **20 November 2019**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini :

Nama	:	Neli Rohmatilah
Nomor Pokok Mahasiswa	:	1517031005
Jurusan	:	Matematika
Judul Skripsi	:	Estimasi Kurva Regresi Nonparametrik dengan menggunakan Metode Fourier

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri, bukan hasil orang lain, dan semua hasil tulisan yang ada dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya berseedia menerima sanksi dengan ketentuan akademik yang berlaku.

Bandar Lampung, 20 November 2019

Penulis

Neli Rohmatilah
NPM. 1517031005

RIWAYAT HIDUP

Penulis dilahirkan pada tanggal 22 Februari 1997 di Purwodadi. Terlahir dari keluarga yang sederhana dari pasangan Bapak Muhammad Bajuri dan Ibu Suryati, merupakan anak ketiga dari tiga bersaudara. Adik dari Syaiful Mukmin dan Ina Wati.

Penulis menyelesaikan pendidikan taman kanak-kanak di TK Islam pada tahun 2003. Sekolah dasar di SD Negeri 1 Pasar Madang pada tahun 2009. Pendidikan sekolah menengah pertama di Mts Matlaul Anwar Landbaw pada tahun 2012. Pendidikan sekolah menengah atas di SMA Negeri 1 Kotaagung jurusan IPA pada tahun 2015. Kemudian penulis melanjutkan pendidikan di perguruan tinggi dan terdaftar sebagai mahasiswa Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SNMPTN pada tahun 2015 dan mendapatkan bantuan beasiswa dari ristekdikti yaitu Bidikmisi.

Pada periode 2016/2017 penulis terdaftar sebagai anggota bidang Kesekretariatan UKMF Natural, lalu pada periode 2017 menjadi ketua bidang Kesekretariatan UKMF Natural Fakultas Matermatika dan Ilmu Pengetahuan Alam Universitas Lampung.

Sebagai bentuk aplikasi bidang ilmu di dunia kerja, penulis telah mlaksanakan Kerja Praktik (KP) selama empat puluh hari di Badan Pusat Statistik Kabupaten Tanggamus. Dan sebagai bentuk aplikasi bidang ilmu kepada masyarakat, penulis telah melaksanakan Kuliah Kerja Nyata (KKN) selama 32 hari di Desa Sinar Mancak, Kecamatan Pulau Panggung, Kabupaten Tanggamus, Provinsi Lampung.

KATA INSPIRASI

"Sukses itu ketika kita gagal lalu bangkit sampai Berhasil"

"Allah tidak akan membebani seseorang melainkan sesuai dengan kadar kesanggupannya"

(Q.S. Al Baqarah : 286)

"Sesungguhnya Allah tidak akan mengubah keadaan suatu kaum hingga mereka mengubah diri mereka sendiri"

(Q.S. Ar – Ra'd : 11)

"Sesungguhnya bersama kesulitan ada kemudahan"

(Q.S. Al-insyirah:6)

"Hanya karena kamu gagal satu kali, bukan berarti kamu akan gagal terus menerus. Kamu hanya gagal ketika kamu berhenti mencoba"

(Anonim)

PERSEMBAHAN

Alhamdulillah, puji syukur kehadirat Allah SWT yang Maha Pengasih lagi Maha Penyayang. Dengan segala ketulusan hati penulis persembahkan skripsi ini kepada:

Bapak Bajuri dan ibu Sur

Terimakasih kepada bapak dan ibu yang selalu tulus mendoakan setiap waktu, membimbing, dan selalu memberikan semangat untuk keberhasilan penulis.

Mas Ipul dan Mbak Ina

Terimakasih karena selalu memberikan keceriaan,semangat dan dukungan serta do’a yang tak pernah henti untukku.

Dosen pembimbing dan pembahas

Terimakasih telah memberikan bimbingan serta saran terbaiknya dalam penyelesaian skripsi ini.

Orang terdekat dan sahabat-sahabatku

Almamaterku Universitas Lampung.

SANWACANA

Dengan menyebut nama Allah SWT, yang Maha Pengasih lagi Maha Penyayang puji dan syukur atas kehadirat Allah SWT berkat segala rahmat dan hidayah-Nya sehingga skripsi dengan Judul "Estimasi Kurva Regresi Nonparametrik dengan menggunakan Metode Fourier" dapat penulis selesaikan dengan baik dan tepat pada waktunya.

Dalam penyusunan skripsi ini banyak kendala yang dihadapi, diantaranya keterbatasan waktu dan materi yang diperlukan dalam penyelesaian tulisan dan pengalaman untuk menulis yang masih terbatas. Namun, atas bantuan dari berbagai pihak skripsi ini dapat diselesaikan tepat pada waktunya. Oleh karena itu, penulis mengucapkan terimakasih kepada:

1. Ibu Ir. Netti Herawati M.Sc., Ph.D., selaku dosen pembimbing akademik dan pembimbing utama yang telat bersedia untuk membimbing, memberi saran dan masukan demi terselesaikannya skripsi ini.
2. Ibu Dra. Dorrah Azis, M.Si., selaku dosen membimbing kedua yang telah membimbing dan memberikan pengarahan pada penulis selama proses penyelesaian skripsi ini.
3. Ibu Widiarti, S.Si., M.Si., selaku dosen pembahas yang telah memberikan kritik dan saran dalam menyelesaikan skripsi ini.

4. Ibu Prof. Dra. Wamiliana, M.A., Ph.D., selaku ketua Jurusan Matematika
5. Bapak Drs. Suratman, M.Sc., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Seluruh Dosen dan staf Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Kedua orangtuaku Bapak Muhammad Bajuri dan Ibu Suryati yang tiada henti-hentinya selalu mendoakan.
8. Kakak-kakakku Syaiful Mukmin, Ina Wati, Dewi Yuliana dan Nelson Alwi Mandala yang selalu memberikan dukungan kepada penulis dalam menyelesaikan skripsi ini.
9. Afrisca, Annisa, Eka, Fransiska, Elita, Hanny, Riswanti, Putri, Azam, nuy, Freta, Meilinda, Irma, Dita, El dan Windi yang selalu memberikan semangat serta motivasi selama masa perkuliahan dan saling membantu dalam menyelesaikan skripsi.
10. Teman-teman KKN desa Sinar Mancak, Pulau Panggung (Tri, Rika, Arum, Harlika, Boby, Zakky) yang telah memberi dukungan.
11. Teman-teman matematika 2015.
12. Seluruh pihak yang telah membantu penulis yang tidak dapat disebutkan satu persatu, atas peran dan dukungannya dalam menyusun skripsi ini.

Akhir kata, penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan akan tetapi semoga skripsi ini dapat berguna dan bermanfaat bagi kita semua.

Bandar Lampung, 20 November 2019
Penulis

Neli Rohmatilah

DAFTAR ISI

Halaman

DAFTAR TABEL

DAFTAR GAMBAR

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis regresi merupakan metode statistika untuk mengetahui hubungan antara satu variabel respon (*dependent*) dengan satu atau lebih variabel prediktor (*independent*). Pada regresi harus ada variabel yang ditentukan dan variabel yang menentukan atau dengan kata lain ada ketergantungan antara variabel yang satu dengan yang lainnya (Hosmer and Lemeshow, 2000). Cara sederhana untuk mengetahui informasi mengenai hubungan antara variabel respon dengan variabel prediktor dapat dilihat melalui bentuk pola hubungan pada diagram pencar, dengan mengetahui pola hubungan yang terbentuk, dapat ditentukan pendekatan yang sesuai untuk mengestimasi fungsi regresi.

Mengestimasi fungsi regresi ada dua pendekatan yang dapat dilakukan yaitu pendekatan secara parametrik dan nonparametrik. Pendekatan parametrik dilakukan apabila bentuk kurva regresi diketahui, dalam regresi parametrik data harus memenuhi asumsi normalitas. Sedangkan pendekatan nonparametrik dilakukan apabila bentuk kurva tidak diketahui.

Regresi nonparametrik dapat digunakan untuk memodelkan data yang berbentuk apasaja, baik berbentuk linear maupun nonlinear dikarenakan tidak adanya asumsi yang harus dipenuhi. Fungsi regresi nonparametrik dilakukan berdasarkan data pengamatan menggunakan teknik pemulusan. Antara lain histogram, *estimator* kernel, *estimator* spline, Fourier, dan wavelet (Eubank, 1998).

Salah satu teknik pemulusan dalam regresi nonparametrik adalah metode Fourier. Metode Fourier mampu mengatasi data yang mempunyai sebaran trigonometri, dalam hal ini adalah sinus dan cosinus. Pola data yang sesuai dengan pendekatan Forier merupakan pola data yang berulang, yaitu pengulangan terhadap nilai variabel *dependent* untuk variabel *independent* yang berbeda-beda. Penelitian menggunakan Fourier untuk nonparametrik maupun semiparametrik telah banyak dikembangkan antara lain: 1. Prahutama (2013) telah meneliti metode pendekatan Fourier pada data pengangguran terbuka di Jawa Timur. 2. Tripena (2007) mengkaji estimator deret Fourier pada regresi nonparametrik. 3. Rudianto (2015) penduga kurva regresi nonparametrik linear dan nonlinear dengan metode Fourier dan metode Nadaraya-Watson.

Penelitian ini akan membandingkan estimasi regresi nonparametrik menggunakan metode Fourier pada beberapa kondisi galat yang berbeda dengan pemilihan *bandwidth* menggunakan metode *Generalized Cross Validation* (GCV).

1.2 Tujuan

Tujuan dari penelitian ini adalah sebagai berikut:

1. Untuk menentukan *bandwitdth* optimal pada regresi nonparametrik metode Fourier dengan menggunakan *Generalized Cross Validation* (GCV) pada data bangkitan dengan galat berbeda.
2. Untuk melakukan estimasi k urva regresi nonparametrik menggunakan metode Fourier pada data bangkitan dengan beberapa kondisi galat yang berbeda.

1.3 Manfaat

Manfaat dari penelitian ini adalah:

1. Mampu mengetahui dan memahami regresi nonparametrik menggunakan metode Fourier.
2. Menambah pengetahuan tentang pemilihan *bandwidth* optimal menggunakan metode *Generalized Cross Validation (GCV)* pada metode Fourier.

II. TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi merupakan salah satu teknik statistika yang sering digunakan untuk mengetahui hubungan antara variabel respon dan variabel prediktor. Untuk analisis yang berhubungan antara satu variabel respon dengan satu variabel prediktor disebut analisis regresi linier sederhana. Sedangkan untuk data yang memiliki lebih dari satu variabel prediktor disebut dengan analisis regresi linier berganda. Jadi, analisis regresi adalah hubungan antara variabel respon dengan satu atau lebih variabel prediktor (Hosmer and Lemeshow, 2000).

Misalkan X adalah variabel prediktor dan Y adalah variabel respon untuk n data pengamatan berpasangan $\{(x_i, y_i)\}_{i=1}^{n}$, maka hubungan antara variabel prediktor dan variabel respon tersebut dapat dinyatakan sebagai berikut:

$$y_i = f(x_i) + \varepsilon_i \qquad i = 1,2,\dots,n \qquad (2.1)$$

Dengan ε_i adalah galat yang berdistribusi normal dengan mean 0 dan variansi σ^2 $(0, \sigma^2)$. y_i adalah variabel respon pada pengamatan ke-i, sedangkan $f(x_i)$ merupakan kurva regresi dengan x_i adalah variabel prediktor (Hardle,1994).

2.2 Regresi Nonparametrik

Regresi nonparametrik merupakan pendekatan metode regresi dimana bentuk kurva dari fungsi regresinya tidak diketahui, kurva regresi diasumsikan termuat dalam ruang fungsi tertentu (Eubank, 1988). Model regresi nonparametrik secara matematis dapat ditulis:

$$y_i = m(x_i) + \varepsilon_i \qquad i = 1,2,3,\dots,n \tag{2.2}$$

dengan y_i adalah variabel tak bebas pada pengamatan ke-i, $m(x_i)$ merupakan kurva fungsi regresi yang tidak diketahui bentuknya dengan x_i merupakan variabel indepenen, ε_i merupakan galat yang berdistribusi normal,mean 0 dan varian σ^2 $(0, \sigma^2)$. Dalam hal ini, $f(x_i)$ hanya diasumsikan termuat dalam suatu ruang fungsi tertentu, dimana pemilihan ruang fungsi tersebut biasanya termotivasi oleh sifat kemulusan (*smooth*) yang dimiliki oleh fungsi $f(x_i)$ (Silverman and Green, 1994).

Estimasi fungsi regresi nonparametrik dilakukan berdasarkan data pengamatan dengan menggunakan teknik pemulusan. Terdapat beberapa teknik pemulusan dengan model regresi nonparametrik antara lain penduga kernel, penduga spline, fourier dan wavelet (Silverman, 1986).

2.3 Fungsi Periodik

Menurut Tolstov (1962), suatu fungsi f(x) dikatakan periodik jika terdapat konstanta T>0, sehingga memenuhi $f(x + T) = f(x)$ untuk setiap x anggota domain f(x), selanjutnya T disebut dengan periode dari fungsi $f(x)$. Jika T adalah periodik dari suatu fungsi f(x), maka …,-2T, -T, 2T, 3T, … juga merupakan periode dari fungsi $f(x)$. Salah satu contoh fungsi periodik adalah $f(x) = sin(x)$, dengan periode 2π, karena $sin(x + 2\pi) = sin(x)$.

2.4 Deret Fourier

Menurut Tolstov (1962), jika fungsi f(x) terdefinisi pada interval $[-L, L\ ]$ dan diluar selang ini oleh $f(x \pm 2L) = f(x)$, sehingga $f(x)$ merupakan fungsi periodik dengan periode 2L. $f(x)$ dapat direpresentasikan dengan deret perluasan Fourier sebagai berikut :

$$f(x) = \frac{1}{2}a_0 + \sum_{j=1}^{\infty}\left[a_j \cos\left(\frac{2\pi jx}{2L}\right) + b_j \sin\left(\frac{2\pi jx}{2L}\right)\right] \tag{2.3}$$

$$f(x) = \frac{1}{2}a_0 + \sum_{j=1}^{\infty}\left[\ a_j \cos\left(\frac{\pi jx}{L}\right) + b_j \sin\left(\frac{\pi jx}{L}\right)\right]$$

dengan

$$a_j = \frac{1}{L}\int_{-L}^{L} f(x) \cos\left(\frac{j\pi x}{L}\right) dx$$

$$b_j = \frac{1}{L}\int_{-L}^{L} f(x) \sin\left(\frac{j\pi x}{L}\right) dx$$

$$j = 1, 2, 3, ...$$

Nilai $\frac{2\pi}{T}$ (dengan T adalah periode $f(x)$) merupakan faktor pengali agar x dalam satuan radian.

2.5 Estimator Fourier

Diberikan n data pengamatan $\{(x_i, y_i)\}_{i=1}^{n}$yang memenuhi persamaan (2.3). Jika $X_i \in [L, -L]$ dan $Y_i \in R$, dan diasumsikan periode m(x) adalah $T = 2L$, maka penduga m(x) dapat didekati oleh Fourier yang didefinisikan sebagai berikut:

$$\widehat{m_f}(j, x_i) = \frac{1}{2}a_0 + \sum_{j=1}^{J}\left[a_j \cos\left(\frac{2\pi jx}{2L}\right) + b_j \sin\left(\frac{2\pi jx}{2L}\right)\right] \quad (2.4)$$

Sehingga:

$$\widehat{m_f}(j, x_i) = \frac{1}{2}a_0 + \sum_{j=1}^{J}\left[a_j \cos(wjx_i) + b_j \sin(wjx_i) \right] \quad w = \frac{2\pi}{2L} \quad (2.5)$$

dengan a_0, a_j dan b_j adalah koefisien Fourier (Bilodeau, 1992).

Untuk menentukan koefisien deret Fourier pada persamaan (2.5), dapat dilakukan perhitungan seperti pada metode parametrik sebagai berikut:

$$\widehat{m_f}(j, x_i) = \frac{1}{2}a_0 + [a_1 \cos(wx_i) + b_1 \sin(wx_i) + a_2 \cos(2wx_i)$$

$$+ b_2 \sin(2wx_i) + \cdots + a_J \cos(wJx_i) + b_J \sin(wJx_i)]$$

$$\widehat{m_f}(j, x_i) = \frac{1}{2}a_0 + [a_1 \cos(wx_i) + a_2 \cos(2wx_i) + \cdots + a_J \cos(wJx_i)$$

$$+ b_1 \sin(wx_i) + \sin(2wx_i) + \cdots + b_J \sin(wJx_i)]$$

Sehingga persamaan regresi nonparametrik menjadi:

$$Y = \widehat{m_f}(j, x_i) + \varepsilon$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \widehat{m_f}(j, x_1) \\ \widehat{m_f}(j, x_2) \\ \vdots \\ \widehat{m_f}(j, x_n) \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dengan:

$\widehat{m_f}(J,X)$

$$\begin{bmatrix} \frac{1}{2}a_0 + [a_1\cos(wx_1) + a_1\cos(2wx_1) + \cdots + a_J\cos(wJx_1) + b_1\sin(wx_1) + b_2\sin(2wx_1) + \cdots + b_J\sin(wJx_1)] \\ \frac{1}{2}a_0 + [a_1\cos(wx_2) + a_1\cos(2wx_2) + \cdots + a_J\cos(wJx_2) + b_1\sin(wx_2) + b_2\sin(2wx_2) + \cdots + b_J\sin(wJx_2)] \\ \vdots \\ \frac{1}{2}a_0 + [a_1\cos(wx_n) + a_1\cos(2wx_n) + \cdots + a_J\cos(wJx_n) + b_1\sin(wx_n) + b_2\sin(2wx_n) + \cdots + b_J\sin(wJx_n)] \end{bmatrix}$$

Misalkan $\widehat{m_f}(J,X) = A\theta$, maka

$$A = \begin{bmatrix} 1 & \cos(wx_1) & \cos(2wx_1) & \cdots & \cos(w\_x_1) & \sin(wx_1) & \sin(2wx_1) & \cdots & \sin(wJx_1) \\ 1 & \cos(wx_2) & \cos(2wx_2) & \cdots & \cos(wJx_2) & \sin(wx_2) & \sin(2wx_2) & \cdots & \sin(wJx_2) \\ & & & & \vdots & & & & \\ 1 & \cos(wx_n) & \cos(2wx_n) & \cdots & \cos(wJx_n) & \sin(wx_n) & \sin(2wx_n) & \cdots & \sin(wJx_n) \end{bmatrix}$$

$$\theta = [\tfrac{a_0}{2} \quad a_1 \quad a_2 \quad \cdots \quad a_J \quad b_1 \quad b_2 \quad \cdots \quad b_J]^T$$

Menurut Chaterjee (2006), nilai θ terbaik dapat diperoleh dengan meminimalkan nilai kuadrat galat (metode *Least Square*). Karena $Y = A\theta + \varepsilon$, maka:

$$\varepsilon = Y - A\theta$$

$$\varepsilon^T\varepsilon = (Y - A\theta)^T(Y - A\theta)$$

Nilai $\varepsilon^T\varepsilon$ minimum diperoleh jika diferensial pertamanya sama dengan nol, sehingga:

$$\frac{d\varepsilon^T\varepsilon}{d\theta} = \frac{d}{d\theta}((Y - A\theta)^T(Y - A\theta))$$

$$\frac{d\varepsilon^T\varepsilon}{d\theta} = \frac{d}{d\theta}(Y^T - \theta^T A^T)(Y - A\theta)$$

$$\frac{d\varepsilon^T\varepsilon}{d\theta} = \frac{d}{d\theta}(Y^TY - \theta^T A^T Y - \theta A Y^T + \theta^T A^T A\theta)$$

$* \; \theta^T A^T Y = \theta A Y^T$, karena konstanta

$$\frac{d\varepsilon^T\varepsilon}{d\theta} = \frac{d}{d\theta}(Y^TY - \theta^T A^T Y - \theta^T A^T Y + \theta^T A^T A\theta)$$

$$\frac{d\varepsilon^T\varepsilon}{d\theta} = \frac{d}{d\theta}(Y^TY - 2\theta^T A^T Y + \theta^T A^T A\theta)$$

$$\frac{d\varepsilon^T\varepsilon}{d\theta} = (-2A^TY + 2A^TA\theta)$$

$$0 = 2(-A^TY + A^TA\theta)$$

$$0 = -A^TY + A^TA\theta$$

$$A^TA\theta = A^TY$$

Jika A^tA memiliki invers, persamaan di atas

$$\theta = (A^tA)^{-1}A^tY \quad (2.6)$$

$\widehat{m_f}(J,X) = A\theta,$ karena $\hat{\theta} = (A^tA)^{-1}A^tY$ maka:

$$\widehat{m_f}(J,X) = A(A^tA)^{-1}A^tY$$

$$\widehat{m_f}(J,X) = [A(A^tA)^{-1}A^t]Y$$

Sehingga $A_J = A(A^tA)^{-1}A^t$ (2.7)

Persamaan (2.6) merupakan penduga $m(x)$ dengan metode Fourier untuk bentuk data gelombang periodik. Sedangkan A_J pada persamaan (2.7) disebut sebagai *Hat Matrix* penduga $m(x)$ pada metode Fourier.

Tingkat kemulusan estimator Fourier ditentukan oleh pemilihan parameter pemulus . Semakin kecil parameter pemulusan, semakin mulus estimasinya dan semakin besar parameter pemulus , semakin kurang mulus estimasi dari *f*. Oleh karena itu, perlu dipilih *bandwidth* yang optimal.

2.6 Pemilihan *Bandwidth* Optimal

Pada pemodelan regresi nonparametrik dengan menggunakan Fourier, hal yang perlu diperhatikan adalah menentukan nilai J. Salah satu metode yang dapat digunakan adalah metode *Generalized Cross Validation (GCV).* Penentuan J optimal akan menghasilkan nilai koefisien determinasi (R^2) yang tinggi. *Generalized Cross Validation (GCV)* didefiniskan sebagai berikut:

$$GCV(J) = \frac{MSE(J)}{[1-\left(\frac{trace\left(A_j\right)}{n}\right)]^2} \quad (2.8)$$

Sehingga, GCV untuk penduga deret Fourier pada persamaan (2.5) adalah

$$GCV\ (J) = \frac{n^{-1}\sum_{i=1}^{n}(y_i-\widehat{m_f}(j,x_i))^2}{[1-(\frac{trace(A(A^tA)^{-1}A^t)}{n})]^2} \quad (2.9)$$

dengan $MSE(J) = n^{-1}\sum_{i=1}^{n}(y_i - \hat{m}(x_i))^2$dan A_J adalah matriks berukuran nxn yang memenuhi $\hat{m}(X) = A_J Y$ dan disebut juga *Hat Matrix.* Nilai GCV terkecil akan menghasilkan nilai J yang optimal (Craven and Wahba, 1979).

2.7 Ukuran Kebaikan *Bandwidth* Optimal

Kebaikan suatu penduga dapat dilihat dari tingkat kesalahannya. Semakin kecil tingkat kesalahan suatu pendugaan maka semakin baik estimasinya. Menurut Chaterjee (2006), kriteria untuk mentukan estimator terbaik dalam model regresi antara lain nilai *Mean Square Error* (MSE) dan nilai koefisien determinasi *R-Square* (R^2).

MSE didefinisikan sebagai berikut:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})^2. \qquad (2.10)$$

sedangkan koefisien determinasi didefinisikan sebagai berikut:

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^{n}(\hat{y_i} - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (2.11)$$

y_i adalah data variabel respon ke-i, $\bar{y}$ adalah mean data variabel respon, sedangkan $\hat{y_i}$adalah nilai hasil estimasi variabel respon ke-i. *Sum of Square Regression* (SSR) adalah jumlah kuadrat simpangan hasil dugaan terhadap rata-rata variabel respon. Sedangkan *Sum of Square Total* (SST) adalah jumlah kuadrat simpangan variabel respon. SSR berfungsi untuk mengukur kualitas variabel prediktor sebagai prediktor variabel respon. Sehingga, koefisien determinasi dapat diartikan sebagai proporsi keragaman total variabel respon yang diukur oleh variabel prediktor.

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat

Penelitian ini dilakukan pada semester genap tahun akademik 2018/2019 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data simulasi menggunakan *software* Matlab R2013b dengan variabel bebas $x \sim U(0,3)$ dengan n= 100 dan dipetakan pada fungsi gelombang periodik dan linear:

$f(x) = \sin(4x) + \varepsilon_i$.

$f(x) = \cos(4x) + \varepsilon_i$.

$f(x) = 2x + 1 + \varepsilon_i$

Pada penelitian ini, fungsi yang digunakan adalalah fungsi gelombang periodik dan linear, pada fungsi linear yang digunkan untuk membuktikan apakah fungsi tersebut baik atau tidak digunakan pada metode Fourier. Dengan $\varepsilon_i$merupakan noise variabel *independen* berdistribusi $\varepsilon_i \sim N(0,1), \varepsilon_i \sim N(0,0.5),\ \varepsilon_i \sim U(-2,2)$ dan $\varepsilon_i \sim 0.9N(0,1) + 0.1N(5,15)$.

Pada penelitian ini, galat yang dibangkitkan adalah galat yang berdistribusi normal, uniform dan tidak normal (mengandung pencilan) agar diperoleh informasi apakah metode Fourier baik dalam mengestimasi data pada kondisi galat tersebut.

Sehingga diperoleh data variabel respon (y_i) melalui persamaan:

$y_i = \sin\ (4x_i) + \varepsilon_I$

$y_i = \cos\ (4x_i) + \varepsilon_i$

$y_i = 2x_i + 1 + \varepsilon_i$

3.3 Metode Penelitian

Dalam penelitian ini akan dilakukan perbandingan pendugaan kurva regresi nonparametrik dengan menggunakan parameter pemulusan (*bandwidth)* optimal dilakukan dengan menggunakan *Generalized Cross Validation* (GCV) minimum. Data diolah dengan menggunakan *software* Matlab R2013b. Adapun langkah-langkah yang dilakukan adalah sebagai berikut:

1. Membuat diagram pencar antara variabel *dependen* dan variabel *independen* yang telah dibangkitkan.
2. Menentukan nilai *bandwidth* optimal berdasarkan nilai GCV dan MSE minimum.
3. Menentukan garis duga regresi berdasarkan nilai *bandwidth* optimal dengan metode Fourier.
4. Membandingkan hasil estimasi regresi nonparametrik dengan metode Fourier dengan galat yang berbeda.

5. Menguji kebaikan model regresi dengan menghitung R^2.
6. Membuat kesimpulan.

V. KESIMPULAN

5.1. Kesimpulan

Berdasarkan hasil penelitian dapat disimpulkan bahwa estimasi kurva regresi nonparametrik menggunakan metode Fourier pada fungsi gelombang periodik dipengaruhi oleh galat, semakin kecil galat yang digunakan, maka semakin baik kurva estimasi yang didapat. Estimasi metode Fourier dengan pemilihan *bandwidth* yang optimal, semakin kecil *bandwidth* yang dihasilkan maka kurva regresi yang didapat semakin mulus atau sebaliknya. Berikut adalah tabel untuk hasil yang didapat dari ketiga fungsi:

Tabel 4.16 Daftar nilai *bandwidth* optimal yang didapat dari ketiga fungsi

Fungsi	ε_i	*Bandwidth Optimum*	R^2	MSE
$y = sin\ 4x + \varepsilon_i$	$N(0{,}0.5)$	1	0.7113	0.2444
$y = cos\ 4x + \varepsilon_i$	$N(0{,}0.5)$	1	0.6270	0.2481
$y = 2x + 1 + \varepsilon_i$	$N(0{,}0.5)$	2	0.9196	0.2418

DAFTAR PUSTAKA

Bilodeu, M. 1992. Fourier Smoother and Additive Models. *The Canadian Journal of Statistics.* **3**: 236-241.

Chaterjee, S. 2006. *Regression Analysis by Example.* 4^{th} Edition. Jhon Wiley and Sons, New Jersey.

Craven, P. and Wahba, G. 1979. Smoothing Noisy Data with Spline Functions: Estimating the Correct Degree of Smooting by The Method of Generalized Cross Validation. *Numer Math University of Wisconsin*. **31**: 377-403.

Eubank, R.L. 1988. *Spline Smoothing and Nonparametrik Regression.* Marcel Dekker, New York.

Hardle, W. 1991. *Smoothing Techniques with Regression.* Cambridge University Press, New York.

Hardle, W. 1994. *Applied Nonparametrik Regression.* Cambridge University Press, New York.

Hosmer, D.W. dan Lemeshow, S. 2000. *Applied Logistic Regression,* 2^{nd}ed. John Wiley and Sons, New York.

Prahutama, A. 2013. Model Regresi Nonparametrik dengan Pendekatan Deret Fourier pada Kasus Tingkat Penggangguran Terbuka di JawaTimur. Prosiding Seminar Nasional Statistika Universitas Diponegoro.

Rudianto, J. 2015. Penduga Kurva Regresi Nonparametrik Linear dan Nonlinear dengan Metode Fourier dan Metode Nadaraya-Watson. Skripsi. Jurusan Matematika FMIPA Universitas Lampung, Bandar Lampung.

Silverman, B.W. 1986. *Dendity Estimation for Statistics and Data Analysis.* Chapman and Hall, London.

Silverman, B.W. and Green, P.J. 1994. *Nonparametric Regression and Generalized Linear Models.* Chapman and Hall/CRC, London.

Tripena, A. 2007. *Estimator Deret Fourier dalam Regresi Nonparametrik*. Tesise. ITS, Surabaya.

Tolstov, G.P. 1962. *Fourier Series Translated from the Russian by Richard A. Silverman.* Dover Publications, New York.