

28293/H/06



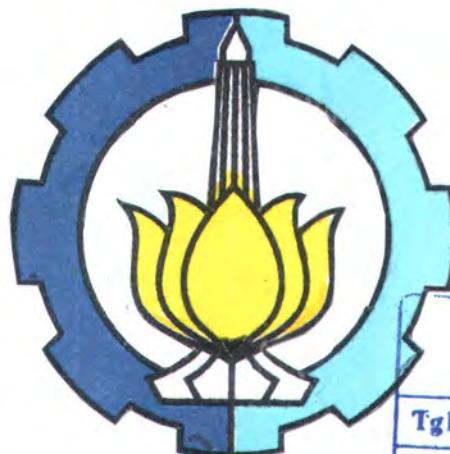
TESIS

PEMILIHAN BANDWIDTH UNTUK REGRESI NONPARAMETRIK
KERNEL DENGAN METODE
RULE OF THUMB, SHIBATA DAN CROSS VALIDATION

Oleh :

MOH. HAFIYUSHOLEH
NRP. 1304 201 019

RTS
519.536
Haf
P-1
2006



PERPUSTAKAAN ITS	
Tgl. Terima	6 - 9 - 06
Terima Dari	H
No. Agenda Prp.	226606

PROGRAM STUDI MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2006

**PEMILIHAN BANDWIDTH UNTUK REGRESI
NONPARAMETRIK KERNEL DENGAN METODE
RULE OF THUMB, SHIBATA DAN CROSS VALIDATION**

Telah disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si)
di
Institut Teknologi Sepuluh Nopember Surabaya

Oleh :
MOH. HAFIYUSHOLEH
NRP. 1304 201 019

Tanggal Ujian : 31 Juli 2006
Periode Wisuda : September 2006

Disetujui oleh Tim Pengaji Tesis :

1. Dr. Drs. I Nyoman Budiantara, MS (Pembimbing I)
NIP. 131 843 382

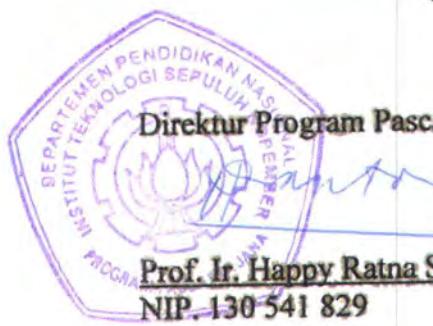
2. Ir. Mutiah Salamah Chamid, M.Kes (Pembimbing II)
NIP. 131 283 368

3. Prof. Dra. Susanti Linuwih, M.Stats., Ph.D (Pengaji)
NIP. 130 368 808

4. Prof. Drs. Nur Iriawan, M.IKom., Ph.D (Pengaji)
NIP. 131 782 011

5. Dr. Purhadi, M.Sc. (Pengaji)
NIP. 131 652 051

6. Dr. Sony Sunaryo, M.Si. (Pengaji)
NIP. 131 843 380



Direktur Program Pascasarjana

Prof. Ir. Happy Ratna S., M.Sc., Ph.D
NIP. 130 541 829

BANDWIDTH SELECTOR FOR NONPARAMETRIC KERNEL REGRESSION WITH RULE OF THUMB METHOD, SHIBATA AND CROSS VALIDATION

By : Moh. Hafiyusholeh
Supervisor : Dr. Drs. I Nyoman Budiantara, MS.
Co-Supervisor : Ir. Mutiah Salamah Chamid, M.Kes.

ABSTRACT

The choice of bandwidth is crucial in the nonparametric estimation procedure. When bandwidth is too small, the resulting curve is too under smoothing. When bandwidth is too large, the resulting curve is over smoothing.

In this research we study three existing bandwidth selectors for nonparametric kernel regression. The first method is based on Mean Squared Error estimate. With penalize Shibata and leave-one-out estimate, bias term of Mean Square Error cancel asymptotically. The other method is based on $Q = E[\hat{m}_h(x) - m(x)]^2$ by combining bias and variance and then derive it to h , we obtained Rule of Thumb bandwidth selector \hat{h}_{ROT} .

With illustrate three methods to the data we examine represent the voltage drops in the battery of a guided missile motor during its flight (Eubank, 1988). We obtained h_{opt} *Shibata* 0,9500, \hat{h}_{ROT} 0,3402 and h_{opt} *Cross Validation* 0,3050. The value of MSE and R² is 0,0137 and 0,9978 for *Cross Validation*, 0,0617 and 0,9898 for Rule of Thumb and 0,0727, 0,9878 for *Shibata*.

Key words: Nonparametric kernel regression, Rule-of-Thumb bandwidth,
Penalize Shibata, Cross Validation

**PEMILIHAN BANDWIDTH UNTUK REGRESI NONPARAMETRIK
KERNEL DENGAN METODE
RULE OF THUMB, SHIBATA DAN CROSS VALIDATION**

Nama mahasiswa : Moh. Hafiyusholeh
NRP : 1304 201 019
Pembimbing : Dr. Drs. I Nyoman Budiantara, MS.
Co-Pembimbing : Ir. Mutiah Salamah Chamid, M.Kes.

ABSTRAK

Pemilihan bandwidth dalam estimasi nonparametrik merupakan sesuatu yang sangat krusial, jika ditetapkan *bandwidth* yang terlalu kecil, maka akan berakibat estimasi kurva regresi *undersmoothing*. Begitu juga sebaliknya jika ditetapkan *bandwidth* yang terlalu besar, maka akan menyebabkan estimasi kurva *oversmoothing*.

Dalam penelitian ini kita mengkaji tiga metode pemilihan bandwidth untuk regresi nonparametrik kernel . Metode pertama berdasar pada estimasi *Mean Square Error*. Dengan menetapkan *penalize Shibata* dan *leave-one-out estimate*, maka bias dari *Mean Square Error* (*MSE*) akan terhapuskan. Metode lain berdasar pada $Q = E[\hat{m}_h(x) - m(x)]^2$ dengan mengkombinasikan antara bias dan varians dan menurunkannya terhadap h didapatkan *bandwidth Rule of Thumb* \hat{h}_{ROT} .

Dengan menerapkan ketiga metode tersebut ke dalam data penurunan *voltage* baterai dari motor peluru kendali selama waktu penerbangan (Eubank, 1988) diperoleh h_{opt} *Shibata* sebesar 0,9500, *Rule of Thumb* sebesar 0,8170 dan *Cross Validation* sebesar 0,3050. Nilai *MSE* dan R^2 masing-masing metode sebesar 0,0137 dan 0,9978 untuk *Cross Validation*, sedangkan untuk *ROT* sebesar 0,0617 dan 0,9898 serta *Shibata* sebesar 0,0727, dan 0,9878.

Kata kunci: Regresi nonparametrik kernel, Rule-of-Thumb bandwidth, Penalize Shibata, Cross Validation

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Alhamdulillaahi Rabbil Alamiin, segala puji bagi Allah SWT yang telah memberikan rahmat, hidayah, petunjuk, kekuatan dan kesabaran pada penulis sehingga dapat menyelesaikan tesis ini dengan judul : **Pemilihan Bandwidth untuk Regresi Nonparametrik Kernel dengan Metode Rule of Thumb, Shibata dan Cross Validation.** Shalawat dan salam penulis haturkan kepada Nabi besar Muhammad SAW yang telah membawa *dienul Islam*. Tesis ini disusun untuk memenuhi salah satu persyaratan mencapai jenjang Magister Sains (M.Si.) pada Program Magister Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, ITS Surabaya.

Penulis menyadari sepenuhnya bahwa tesis ini dapat diselesaikan, tidak terlepas dari berbagai pihak baik secara langsung maupun tidak, dengan memberikan dorongan, bantuan dan bimbingan kepada penulis. Untuk itu pada kesempatan ini penulis mengucapkan banyak terima kasih yang sedalam-dalamnya kepada :

1. Bapak Dr. Drs. I Nyoman Budiantara, MS. selaku Koordinator Program Studi S2 Statistika ITS, sekaligus sebagai dosen pembimbing utama, yang dengan penuh kesabaran dan tidak kenal lelah, memberikan bimbingan, arahan serta motivasi.
2. Ibu Ir. Mutiah Salamah Chamid, M.Kes., selaku Ketua Jurusan Statistika ITS sekaligus dosen pembimbing kedua yang juga telah banyak memberikan pengarahan dan motivasi selama proses penulisan.
3. Para dosen pengajar yang dengan tulus telah memberikan wacana keilmuan kepada penulis, beserta staf administrasi akademik, laboratorium komputasi dan ruang baca Jurusan Statistika yang telah banyak membantu penulis selama masa studi.

4. Ibu Evawati Alisah, M.Pd dan Pimpinan STIMIK Malang yang telah memberikan bantuan dan kesempatan kepada penulis untuk melanjutkan pendidikan di Program Magister Jurusan Statistika ITS.
5. Orang tua dan istriku yang tercinta, Siti Shun'ah Hasanah yang telah banyak berkorban dengan segala kasih sayang, do'a dan motivasi kepada penulis. Semoga Allah SWT memberikan yang terbaik untuk mereka.
6. Saudara - saudariku yang telah banyak memberikan pelajaran akan kehidupan.
7. Saudara-saudaraku seangkatan 2004 di Program Magister Statistika ITS yang namanya tidak sempat penulis sebutkan satu persatu, atas kebersamaan, kekompakan, bantuan dan masukannya selama masa perkuliahan. Semoga kita mendapat ilmu yang Nafi', diberi ke-Selamet-an, Maunah, dan diberi kehidupan yang Sriyani.

Akhirnya penulis berharap dan memohon semoga bantuan yang telah diberikan mendapat Ridho serta balasan yang setimpal dari-Nya. Sebagai suatu karya ilmiah penulis berharap masukan dan kritikan. Semoga tesis ini dapat memberikan manfaat dan mendapat ridho dari Allah SWT. Amin.

Surabaya, Agustus 2006

Penulis

DAFTAR ISI

	Halaman
LEMBAR PENGESAHAN	<i>i</i>
ABSTRACT	<i>ii</i>
ABSTRAK	<i>iii</i>
KATA PENGANTAR	<i>iv</i>
DAFTAR ISI	<i>vi</i>
DAFTAR TABEL	<i>viii</i>
DAFTAR GAMBAR	<i>ix</i>
DAFTAR LAMPIRAN	<i>x</i>
BAB I. PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian	4
1.4. Manfaat Penelitian	4
1.5. Batasan Masalah	4
BAB II TINJAUAN PUSTAKA	5
2.1. Regresi Nonparametrik	5
2.1.1. Estimator Kernel	6
2.1.2. Regresi Nonparametrik Kernel	7
2.1.3. Estimator Least Square	8
2.2. Rule of Thumb (ROT) Bandwidth	9
2.2.1. Rumus Taylor	10
2.2.2. Turunan Matriks	10
2.2.3. Kriteria Asimtotik	11
2.3. Fungsi Penalize	12
2.4. Cross Validation (CV)	12

BAB III METODOLOGI PENELITIAN	13
3.1. Bahan dan Alat	13
3.1. Metode Penelitian	13
 BAB IV HASIL DAN PEMBAHASAN	15
4.1. Rule of Thumb (ROT)	15
4.2. Bandwidth Shibata	29
4.3. Cross Validation (CV)	39
4.4. Aplikasi	42
4.3.1. Bandwidth Shibata	45
4.3.2. Rule of Thumb Bandwidth	48
4.3.3. Metode Cross Validation	50
 BAB V KESIMPULAN DAN SARAN	55
5.1. Kesimpulan	55
5.2. Saran	57
 DAFTAR PUSTAKA	58

DAFTAR TABEL

Tabel	Keterangan	hal
Tabel 4. 1	Berbagai nilai bandwidth yang diperoleh dengan metode Shibata	45
Tabel 4. 2	Berbagai nilai CV dan Bandwidth (h)	50
Tabel 4. 3	Nilai SSE, MSE dan R^2 dari masing-masing metode	54

DAFTAR GAMBAR

Gambar	Keterangan	hal
Gambar 4.1	Plot antara voltage drop dengan waktu	42
Gambar 4.2	Estimasi kurva regresi nonparametrik dengan $h = 0,100$	43
Gambar 4.3	Estimasi kurva regresi nonparametrik dengan $h = 3,000$	44
Gambar 4.4	Plot antara bandwidth dengan Shibata	46
Gambar 4.5	Plot estimasi nonparametrik kernel dengan $h = 0,9500$	47
Gambar 4.6	Plot estimasi kurva regresi dengan bandwidth ROT	49
Gambar 4.7	Nilai CV dan bandwidth	51
Gambar 4.8	Estimasi kurva regresi dengan bandwidth CV	52
Gambar 4.9	Plot gabungan taksiran kurva regresi	53

DAFTAR LAMPIRAN

Lampiran	Keterangan	hal
A	Data penurunan voltage baterai dari motor peluru kendali selama waktu penerbangan	60
B	Program Regresi Nonparametrik Kernel, SSE, MSE, dan R ²	61
C	Program mencari bandwidth optimal dengan:	
	○ Rule of Thumb	62
	○ Shibata	63
	○ Cross Validation	64
D	Pengujian Asumsi Residual	65
E	Pengujian Data dengan Kuadratik dan Kubik	71

BAB I

PENDAHULUAN

BAB I

PENDAHULUAN

1.1. Latar Belakang

Analisis regresi merupakan salah satu alat statistik yang banyak digunakan untuk mengetahui hubungan antara sepasang variabel atau lebih. Hubungan yang didapat pada umumnya dinyatakan dalam bentuk persamaan yang menyatakan hubungan fungsional antara variabel-variabel. Misalkan Y adalah variabel respon dan $m(X)$ adalah variabel prediktor, maka untuk n pengamatan hubungan variabel itu dapat dinyatakan dalam bentuk :

$$Y_i = m(X_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (1.1)$$

dimana $m(X_i)$ adalah fungsi regresi atau kurva regresi, dan ε_i adalah suatu variabel random yang diasumsikan independen identik dengan mean 0 dan varians σ^2 .

Ada dua pendekatan yang dapat digunakan untuk mengestimasi m yaitu pendekatan parametrik dan nonparametrik. Pendekatan parametrik digunakan bila bentuk fungsi m diketahui berdasarkan pada teori atau pengalaman masa lalu. Sedangkan pendekatan nonparametrik digunakan bila tidak ada asumsi tentang bentuk kurva atau fungsi regresi.

Dalam regresi nonparametrik, ada beberapa pendekatan yang dapat digunakan antara lain histogram, kernel, deret orthogonal, spline, dan lain-lain (Härdle, 1990). Pendekatan kernel diperkenalkan oleh Rosenblatt (1956) untuk mengestimasi fungsi densitas $f(x)$ yaitu :

$$f_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right) \quad (1.2)$$

dengan $K(\cdot)$ adalah fungsi kernel dan h adalah *bandwidth*.

Dalam pendekatan kernel, bentuk estimasinya dipengaruhi oleh fungsi kernel $K(\cdot)$ dan *bandwidth* h . Bandwidth h adalah parameter penghalus yang bersfungsi untuk mengontrol kemulusan dari kurva estimasi. Bandwidth yang terlalu kecil akan menghasilkan kurva yang *undersmoothing*, yaitu sangat kasar dan sangat fluktuatif dan sebaliknya jika bandwidth terlalu besar, maka akan menghasilkan kurva yang *oversmoothing*.

Cara untuk menentukan *bandwidth* dapat dilakukan dengan cara coba-coba, sehingga yang harus dilakukan adalah mencoba beberapa nilai *bandwidth* kemudian nilai tersebut dimasukkan dalam taksiran fungsi kernelnya. Namun cara ini kurang efektif karena memerlukan *sense* yang cukup tinggi untuk mengambil keputusan bahwa taksiran cukup halus. Dari sinilah kemudian para ahli berusaha mengembangkan metode pemilihan *bandwidth* yang bisa digunakan untuk mendapatkan taksiran yang tepat.

Metode yang dimaksud adalah dengan menurunkan suatu formula untuk mendapatkan suatu *bandwidth* yang optimal melalui $Q = E[\hat{m}_h(x) - m(x)]^2$. Tapi sayangnya hasil *bandwidth* optimal yang diperoleh masih mengandung suatu fungsi m dan σ yang tidak diketahui. Silverman (1986) menawarkan metode pemilihan *bandwidth* yang kemudian dikenal dengan *Rule of Thumb (ROT)*. Prinsip yang mendasarinya adalah dengan mendekati fungsi yang tidak diketahui tersebut dengan pendekatan parametrik biasa.

Selain dengan metode diatas, *bandwidth* yang optimal dapat juga diperoleh melalui pendekatan fungsi *Loss* ($L(h)$), *Rice* ($R(h)$) atau *Prediction Rice* ($p(h)$), tapi sayangnya diantara ketiganya tidak satupun yang dapat dihitung karena mengandung

fungsi $m(x)$ yang tidak diketahui. Oleh karena itu, hal yang memungkinkan agar didapatkan *bandwidth* yang optimal adalah dengan mengganti fungsi $m(x)$ tersebut dengan Y yang oleh Hardle (1990) disebut dengan istilah *resubstitution estimate* atau *Mean Square Error (MSE)*. Tapi ternyata dengan pendekatan *MSE* tersebut masih mengandung bias untuk $p(h)$.

Shibata (1980) dan Shibata (1981) mengkonstruksi suatu *penalize* pada *Mean Square Error (MSE)* agar taksiran yang dihasilkan tidak bias terhadap $p(h)$, atau dengan kata lain bias pada *MSE* setelah diberi *penalize* *Shibata* akan terhapuskan. Cara lain untuk menghilangkan bias dari *MSE* adalah dengan menggunakan *leave-one-out estimate* yang kemudian lebih dikenal dengan istilah *Cross Validation* (Hardle, 1991).

Untuk mendapatkan gambaran seberapa baik masing-masing metode pemilihan *bandwidth*, maka akan diterapkan pada data penurunan voltase baterai motor peluru kendalai selama waktu penerbangan. Data diambil dari Eubank (1988).

1.2. Rumusan Masalah

Berdasarkan uraian diatas, maka rumusan masalah yang dapat penulis kemukakan adalah:

1. Bagaimana mendapatkan *bandwidth* dengan metode *Rule-of-Thumb* ?
2. Bagaimana mendapatkan *bandwidth* dengan metode *Shibata* ?
3. Bagaimana mendapatkan *bandwidth* dengan metode *Cross Validation* ?
4. Bagaimana penerapan pemilihan *bandwidth* diatas untuk regresi nonparametrik kernel dalam data penurunan voltase baterai motor peluru kendali ?

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka tujuan dari penelitian ini adalah:

1. Mengkaji pemilihan *bandwidth* dengan metode *Rule-of-Thumb* untuk model regresi nonparametrik kernel.
2. Mengkaji pemilihan *bandwidth* dengan metode *Shibata*.
3. Mengkaji pemilihan *bandwidth* dengan metode *Cross Validation*.
4. Menerapkan metode pemilihan *bandwidth* untuk regresi nonparametrik kernel pada data penurunan voltase baterai motor peluru kendali.

1.4. Manfaat Penelitian

Manfaat yang diharapkan dari hasil penelitian ini adalah mampu meningkatkan wawasan keilmuan khususnya yang berkaitan dengan model regresi nonparametrik dengan menggunakan estimator kernel. Disamping itu juga dengan adanya penelitian ini diharapkan dapat memberikan informasi alternatif metode pemilihan *bandwidth* yang optimal dari estimator kernel.

1.5. Batasan Masalah

Mengacu pada rumusan masalah diatas, maka ruang lingkup dari penelitian ini akan dibatasi pada regresi nonparametrik kernel dengan menggunakan pendekatan lokal konstan dengan metode *Rule of Thumb*, *Shibata* dan *Cross Validation*.



BAB II

TINJAUAN PUSTAKA

BAB II

TINJAUAN PUSTAKA

2.1 Regresi Nonparametrik

Istilah regresi diperkenalkan pertama kali oleh Francis Galton pada tahun 1885 berkaitan dengan eksperimen antara hubungan tinggi orangtua dengan anaknya (Draper and Smith, 1966). Dalam regresi parametrik diasumsikan bahwa bentuk kurva regresi diketahui. Pembuatan asumsi tersebut berdasarkan pada teori, pengalaman masa lalu atau tersedianya sumber-sumber lain yang dapat memberikan pengetahuan atau informasi yang terperinci.

Estimasi dapat juga dilakukan berdasarkan suatu pendekatan yang tidak terikat dengan asumsi bentuk kurva regresi tertentu, yang memberikan fleksibilitas yang lebih besar. Metode dengan pendekatan ini dinamakan dengan pendekatan nonparametrik yang mulai dikenal sejak abad ke-19.

Misalkan y adalah variabel respon dan x adalah variabel prediktor. Secara umum hubungan variabel dapat dinyatakan sebagai $Y_i = m(X_i) + \varepsilon_i, i = 1, 2, 3, \dots, n$ dimana ε_i adalah variabel random yang diasumsikan identik independent dengan mean nol dan variansi σ^2 . Fungsi $m(x)$ adalah fungsi yang tidak diketahui yang disebut sebagai fungsi regresi atau kurva regresi. Fungsi regresi $m(x)$ hanya diasumsikan termuat dalam suatu ruang fungsi yang berdimensi tidak hingga. Kemudian estimasi $m(x)$ dilakukan berdasarkan data pengamatan dengan teknik penghalus tertentu (Härdle, 1990).

Ada beberapa teknik penghalus dalam regresi non parametrik, antara lain histogram, estimasi kernel, estimasi deret orthogonal, estimasi spline, k-NN, deret Fourier, dan Wavelet.

2.1.1 Estimator Kernel

Estimator kernel diperkenalkan pertamakali oleh Rosenblatt pada tahun 1956 dan Parzen pada tahun 1962 dengan menentukan pembobot pada masing-masing observasi dalam himpunan data (Kweon dan Koekelman, 2004). Salah satu kelebihan dari estimator kernel adalah bentuknya lebih fleksibel dan mudah disesuaikan.

Beberapa jenis fungsi kernel dapat disajikan sebagai berikut: (Härdle, 1990):

- a. Kernel Gaussian : $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}u^2) I_{(-\infty, \infty)}(u)$
- b. Kernel Epanechnikov : $K(u) = \frac{3}{4} (1 - u^2) I_{[-1,1]}(u)$
- c. Kernel Kuadrat : $K(u) : \frac{15}{8} (1 - 4u^2)^2 I_{[-0,5; 0,5]}(u)$
- d. Kernel Uniform : $K(u) = \frac{1}{2} I_{[-1,1]}(u)$
- e. Kernel Segitiga : $K(u) = (1 - |u|) I_{[-1,1]}(u)$

Secara umum fungsi kernel untuk dimensi satu didefinisikan sebagai berikut :

$$K_h(u) = \frac{1}{h} K\left(\frac{u}{h}\right); \quad -\infty < u < \infty \quad (2.1)$$

Kernel K merupakan fungsi yang kontinu, terbatas, simetrik, dan terintegral ke satu, $\int_{-\infty}^{\infty} K(u)du=1$. (Härdle, 1990).

Jika suatu kernel memenuhi syarat berikut,: (Silverman, 1986)

- i. $\int_{-\infty}^{\infty} K(u) du = 1$
- ii. $\int_{-\infty}^{\infty} uK(u) du = 0$
- iii. $\int_{-\infty}^{\infty} u^2 K(u) du \neq 0$
- iv. $\int_{-\infty}^{\infty} K(u)^2 du < \infty,$

maka kernel K disebut berordo 2.

2.1.2 Regresi Nonparametrik Kernel

Estimasi regresi nonparametrik kernel dengan menggunakan pendekatan linear konstan menggunakan sampel random $(x_i, y_i), i = 1, \dots, n$. Estimasi kurva $\hat{y} = \hat{m}(x)$ diperoleh dengan meminimumkan:

$$\sum_{i=1}^n \varepsilon_i^2 K_h(x - X_i) = \sum_{i=1}^n \{y_i - [\delta_0 + \delta_1(x - X_i) + \dots + \delta_p(x - X_i)^p]\}^2 K_h(x - X_i)$$

K disebut fungsi kernel dan h disebut sebagai *bandwidth*. Estimator $\hat{\delta}$ merupakan hasil dari optimasi *least square* yang terboboti.

Pemilihan bandwidth yang optimal tergantung pada kriteria yang digunakan. Kriteria yang biasa digunakan adalah: (Eubank, 1988) dan (Härdle, 1990) serta (Rice, 1984).

- a. Fungsi Loss $L(h) = n^{-1} \sum_{j=1}^n (m(X_j) - \hat{m}_h(X_j))^2$
- b. Fungsi Risk $R(h) = E[n^{-1} \sum_{j=1}^n (m(X_j) - \hat{m}_h(X_j))^2]$

2.1.3 Estimator Least Square

Ide yang melatar belakangi metode *Least Square* adalah untuk memilih estimator dengan jalan meminimumkan *sum of square residual* guna mendapatkan estimator $\beta_0, \beta_1, \beta_2, \dots, \beta_k$. Misal diberikan model :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Dimana \mathbf{y} adalah vektor $n \times 1$, $\boldsymbol{\varepsilon}$ adalah vektor random $n \times 1$ dari residual dan $\boldsymbol{\beta}$ adalah vektor $(k+1) \times 1$ dari estimator untuk koefisien $\beta_0, \beta_1, \beta_2, \dots, \beta_k$, serta \mathbf{X} adalah matrik full rank $n \times (k+1)$ dari bilangan riil.

Untuk mendapatkan $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ dengan metode *least square*, dapat diperoleh dengan cara meminimumkan *sum of square* dari residual.

$$\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = \sum_{i=1}^n \varepsilon_i^2$$

Teorema 2.1

Jika $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ dimana \mathbf{X} adalah matrik full rank $n \times (k+1)$, $\boldsymbol{\beta}$ adalah vektor $(k+1) \times 1$ dari parameter yang tidak diketahui, dan $\boldsymbol{\varepsilon}$ adalah vektor random $n \times 1$ dengan mean 0 dan variansi $\sigma^2 I$. Estimator least square untuk $\boldsymbol{\beta}$ dinotasikan $\hat{\boldsymbol{\beta}}$ diberikan oleh $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$ (Myers dan Milton, 1991)

Bukti:

Vektor dari residual $\boldsymbol{\varepsilon}$ dapat ditulis sebagai: $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$. Sehingga :

$$\begin{aligned}\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= (\mathbf{y}^T - \boldsymbol{\beta}^T \mathbf{X}^T)(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{y}^T \mathbf{y} - \mathbf{y}^T \mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X}\boldsymbol{\beta}\end{aligned}$$

Karena $\beta^T X^T y$ adalah 1×1 , $\beta^T X^T y = (\beta^T X^T y)^T = y^T X \beta$, maka dengan substitusi

$$\varepsilon^T \varepsilon = y^T y - 2y^T X \beta + \beta^T (X^T X) \beta$$

$$= y^T y - 2(X^T y)^T \beta + \beta^T (X^T X) \beta$$

Untuk meminimumkan $\varepsilon^T \varepsilon$ sebagai fungsi dari β , maka kita turunkan terhadap β kemudian menyamakannya dengan nol

$$\frac{\partial \varepsilon^T \varepsilon}{\partial \beta} = -2X^T y + 2(X^T X)\beta = 0$$

$$\Leftrightarrow (X^T X)\beta = X^T y$$

untuk mendapatkan $\hat{\beta}$ kita kalikan kedua ruas dengan $(X^T X)^{-1}$,

sehingga didapat :

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

2.2. Rule of Thumb (ROT) Bandwidth

Taksiran fungsi regresi $m_h(x)$ bergantung pada h . Lebih khusus, jika nilai h semakin kecil, maka bias juga semakin kecil akan tetapi varians akan bertambah besar. Begitujuga sebaliknya jika nilai h semakin besar, maka bias akan cenderung besar dan varian akan semakin kecil. Berkaitan dengan bias dan varian, maka h yang optimal dapat diperoleh dengan meminimumkan $Q = E[\hat{m}_h(x) - m(x)]^2$ terhadap h .

Rule of Thumb pertama kali ditemukan oleh Silverman (1986) kaitannya dengan *pdf* kernel. Prinsip yang mendasarinya adalah dengan mendekati fungsi yang tidak diketahui dengan parametrik, sehingga

$$\hat{h}_{ROT} = n^{-1/5} \left[\frac{\hat{\sigma}^2 \|K\|_2^2}{\{\mu_2(K)\}^2 \{\hat{m}''(x)\}^2} \right]^{1/5}$$

2.2.1. Rumus Taylor (Purcell dan Varberg, 1995)

Andaikan f adalah suatu fungsi dengan turunan ke $(n+1)$, $f^{(n+1)}(x)$ ada untuk setiap x pada suatu selang buka I yang mengandung a , maka untuk setiap x di I

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x)$$

Dimana sisa $R_n(x)$ diberikan oleh rumus $R_n(x) = \frac{f^{(n+1)}(a)}{(n+1)!}(x-a)^{n+1} + \dots$

2.2.2. Turunan Matriks

Misal $\mathbf{x}^T = (x_1, x_2, \dots, x_p)$, $\mathbf{y}^T = (y_1, y_2, \dots, y_p)$, $\mathbf{a}^T = (a_1, a_2, \dots, a_p)$ dan $\mathbf{A} = [a_{ij}]$ simetris orde p serta $\mathbf{B} = [b_{ij}]$ berukuran $p \times q$.

1. Jika $Z = \mathbf{a}^T \mathbf{x}$, maka turunan Z ke \mathbf{x} diberikan oleh:

$$\frac{\partial Z}{\partial \mathbf{x}} = \left(\frac{\partial Z}{\partial x_1}, \frac{\partial Z}{\partial x_2}, \dots, \frac{\partial Z}{\partial x_p} \right) = \mathbf{a}^T$$

2. Jika $Z = \mathbf{a}^T \mathbf{B} \mathbf{y}$, maka turunan Z ke \mathbf{B} diberikan oleh:

$$\frac{\partial Z}{\partial \mathbf{B}} = \left(\frac{\partial Z}{\partial b_{ij}} \right) = \mathbf{x} \mathbf{y}^T$$

3. Jika $Z = \mathbf{x}^T \mathbf{A} \mathbf{x}$, maka turunan Z ke \mathbf{x} diberikan oleh:

$$\frac{\partial Z}{\partial \mathbf{x}} = 2 \mathbf{A} \mathbf{x}$$

4. Jika $Z = \mathbf{x}^T \mathbf{A} \mathbf{x}$, maka turunan Z ke \mathbf{A} diberikan oleh:

$$\frac{\partial Z}{\partial \mathbf{A}} = 2 \mathbf{x} \mathbf{x}^T - \text{diag}(\mathbf{x} \mathbf{x}^T)$$

2.2.3. Kriteria asimtotik

Dalam teori estimasi, terdapat beberapa pendekatan untuk mengetahui sifat-sifat baik suatu estimator. Dalam banyak kasus, tidak semua persoalan teori estimasi dapat didekati dengan sampel berukuran kecil (eksak). Tetapi sering suatu persoalan estimasi, menyarankan kita menggunakan sampel berukuran besar.

Berkaitan dengan kriteria asimtotik estimator bandwidth, maka akan dibahas mengenai kriteria asimtotik *big Oh* dan *little Oh*, serta sifat kekonvergenannya.

Big Oh dan Little Oh (Rohatgi, 1976)

Notasi yang sering digunakan dari *Big Oh* dan *Little Oh* adalah $O(\cdot)$ dan $o(\cdot)$. Misalkan f dan g dua fungsi yang terdefinisikan pada bilangan real x .

a. Fungsi $f(x)$ dikatakan *most of order* dari $g(x)$, $x \rightarrow \infty$, ditulis dengan simbol:

$$f(x) = O[g(x)], \quad x \rightarrow \infty, \text{ jika terdapat suatu konstan } M < \infty \text{ sehingga } \left| \frac{f(x)}{g(x)} \right| < M$$

b. Fungsi $f(x)$ dikatakan *smaller order* dari $g(x)$, $x \rightarrow \infty$, ditulis dengan simbol:

$$f(x) = o[g(x)], \quad x \rightarrow \infty, \text{ jika } \lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 0.$$

Beberapa sifat yang dimiliki oleh $O(\cdot)$ dan $o(\cdot)$ adalah sebagai berikut :

- (i) Jika $f_1(x) = O[g_1(x)]$, $f_2(x) = O[g_2(x)]$ maka $f_1(x) + f_2(x) = O[g_1(x) + g_2(x)]$
- (ii) Jika $\alpha > 0$ konstan, $f(x) = O[\alpha g(x)]$ maka $f(x) = O[g(x)]$
- (iii) Jika $f_1(x) = O[g_1(x)]$ dan $f_2(x) = O[g_2(x)]$ maka $f_1(x)f_2(x) = O[g_1(x)g_2(x)]$
- (iv) Jika $f(x) = o[g(x)]$ maka $f(x) = O[g(x)]$
- (v) Jika $f_1(x) = O[g_1(x)]$, $f_2(x) = o[g_2(x)]$ maka $f_1(x)f_2(x) = o[g_1(x)g_2(x)]$

2.3. Fungsi Penalize

Beberapa fungsi penalize antara lain : (Härdle, 1990)

1. Shibata's model selector $\Xi_S(u) = 1 + 2u$
2. Generalized cross validation $\Xi_{GCV}(u) = (1 - u)^{-2}$
3. Akaike's information criterion $\Xi_{AIC}(u) = \exp(2u)$
4. Finite prediction error $\Xi_{FPE}(u) = \frac{1+u}{1-u}$
5. Rice's T $\Xi_T(u) = (1 - 2u)^{-1}$

2.4. Cross Validation (CV)

Metode *cross validation* pertama kali diperkenalkan oleh Rudemo pada tahun 1982 (Silverman, 1986). Cross validation merupakan *leave-one-out method* yang berdasar pada *regression smoother* dimana observasi ke-j dihilangkan. (Härdle, 1990).

$$\hat{m}_{h,j}(X_j) = (n-1)^{-1} \sum_{i \neq j} W_{hi}(X_j) Y_i$$

Sehingga :

$$CV(h) = n^{-1} \sum_{i \neq j} (Y_i - \hat{m}_{h,j}(X_j))^2$$

BAB III

METODOLOGI PENELITIAN

BAB III

METODOLOGI PENELITIAN

3.1 Bahan dan Alat

Bahan dan alat yang digunakan dalam penelitian ini adalah:

1. Jurnal dan referensi yang terkait dengan permasalahan di atas.
2. Software minitab, matlab, dan microsoft excel.
3. Data penurunan voltase baterai motor peluru kendali selama waktu penerbangan dengan data sebanyak 41 (Eubank, 1988).

3.2 Metode Penelitian

Metode yang akan dilakukan berkaitan dengan penelitian ini adalah:

1. Mendapatkan *Rule of Thumb* (ROT) dengan langkah-langkah sebagai berikut:

- a. Mendefinisikan Q dimana :

$$Q = E[(\hat{m}_h(x) - m(x))^2] = ([Bias\{\hat{m}_h(x)\}]^2 + Var\{\hat{m}_h(x)\})$$

- b. Memperlihatkan bahwa;

$$Bias\{\hat{m}_h(x)\} = \left\{ \int x^2 K(x) dx \right\} \frac{1}{(2)!} m'' h^2 + o(h)$$

$$Var\{\hat{m}_h(x)\} = \left\{ \int K^2(x) dx \right\} \frac{\sigma^2}{nh}$$

- c. Meminimumkan Q terhadap h sehingga didapat h_{opt} :

$$h_{opt} = n^{-1/5} \left[\frac{\sigma^2 \|K\|_2^2}{(\mu_2(K))^2 \{m''(x)\}^2} \right],$$

2. Mengkaji ketidakbiasaan metode *Shibata* dengan langkah-langkah sebagai berikut:
- Mendefinisikan fungsi *Loss* $L(h) = n^{-1} \sum_{j=1}^n (m(X_j) - \hat{m}_h(X_j))^2$
 - Mendefinisikan fungsi *Risk* $R(h) = E[L(h)]$
 - Mendefinisikan *prediction risk* $p(h) = \sigma^2 + R(h)$
 - Menunjukkan bias dari *MSE* terhadap $p(h)$ adalah $-2n^{-2}\sigma^2 \sum_{j=1}^n W_{h,j}(X_j)$
 - Memberikan *penalize* *Shibata* kedalam model, sehingga diperoleh *unbiased estimate* untuk $p(h)$.
3. Mengkaji ketidakbiasaan CV, dengan langkah-langkah sebagai berikut:
- Mendefinisikan CV dimana:
$$CV(h) = n^{-1} \sum_{j=1}^n [Y_j - \hat{m}_{h,j}(X_j)]^2$$
 - Mendapatkan *unbiased estimate* untuk $p(h)$.
4. Membandingkan metode *Rule of Thumb*, *Shibata* dan *Cross Validation* pada data penurunan voltase baterai motor peluru kendali dengan langkah-langkah sebagai berikut:
- Plot data antara variabel prediktor (x) dengan variabel respon (y).
 - Mendapatkan *bandwidth* optimal dengan metode *Shibata* dan *Rule of Thumb (ROT)* serta *Cross Validation (CV)*.
 - Mendapatkan estimasi kurva regresi nonparametrik kernel.
 - Menghitung dan membandingkan nilai *MSE* dan R^2 dari metode *Shibata*, *Rule of Thumb (ROT)*, dan *Cross Validation (CV)*.

BAB IV

HASIL DAN PEMBAHASAN

BAB IV

HASIL DAN PEMBAHASAN

Pemilihan *bandwidth* dalam regresi nonparametrik kernel merupakan suatu hal yang sangat penting. Jika *bandwidth* h terlalu kecil, maka estimasi kurva yang dihasilkan akan cenderung kasar, sebaliknya jika *bandwidth* h terlalu besar, maka akan menghasilkan estimasi yang *oversmoothing*. Berikut ini diberikan pemilihan *bandwidth* optimal dengan metode *Rule of Thumb*, *Shibata* dan *Cross Validation*.

4.1 Rule of Thumb (ROT)

Rule of Thumb memilih *bandwidth* optimal dengan meminimumkan $Q = E[\hat{m}_h(x) - m(x)]^2$ terhadap h . Pertama ditunjukkan Q dapat dikomposisi menjadi bias kuadrat dan varians, sebagai berikut:

$$\begin{aligned} Q &= E[\hat{m}_h(x) - m(x)]^2 \\ &= E[\hat{m}_h(x) - E(\hat{m}_h(x)) + E(\hat{m}_h(x)) - m(x)]^2 \\ &= E[(\hat{m}_h(x) - E(\hat{m}_h(x)))^2 + (E(\hat{m}_h(x)) - m(x))^2 + \\ &\quad + 2(\hat{m}_h(x) - E(\hat{m}_h(x)))(E(\hat{m}_h(x)) - m(x))] \\ &= E[(\hat{m}_h(x) - E(\hat{m}_h(x)))^2] + (E(\hat{m}_h(x)) - m(x))^2 + \\ &\quad + 2(E(\hat{m}_h(x)) - E(\hat{m}_h(x)))(E(\hat{m}_h(x)) - m(x)) \\ &= E[(\hat{m}_h(x) - E(\hat{m}_h(x)))^2] + (E(\hat{m}_h(x)) - m(x))^2 + 0 \\ &= (E(\hat{m}_h(x)) - m(x))^2 + E[(\hat{m}_h(x) - E(\hat{m}_h(x)))^2] \\ &= (\text{Bias}(\hat{m}_h(x)))^2 + \text{Var}(\hat{m}_h(x)). \end{aligned}$$

dengan

$$(Bias(\hat{m}_h(x)))^2 = (E(\hat{m}_h(x)) - m(x))^2 \quad (4.1)$$

$$Var(\hat{m}_h(x)) = E[(\hat{m}_h(x) - E(\hat{m}_h(x))^2] \quad (4.2)$$

Untuk mendapatkan \hat{h}_{ROT} terlebih dahulu akan dikaji bias dan varians dari $\hat{m}_h(x)$. Merujuk pada hasil penguraian Q diatas, maka perlu untuk mengetahui $\hat{m}_h(x)$. Misal diberikan data (X_i, Y_i) dan diasumsikan mengikuti model regresi :

$$Y_i = m(X_i) + \varepsilon_i, \quad \varepsilon_i \sim iid(0, \sigma^2), \quad i = 1, 2, 3, \dots, n. \quad (4.3)$$

Fungsi $m(X_i)$ didekati dengan polinomial derajad p (Jia dan Schucany, 2004).

$$\begin{aligned} m(X_i) &= \sum_{j=0}^p \delta_j (x - X_i)^j \\ &= \delta_0 + \delta_1 (x - X_i) + \delta_2 (x - X_i)^2 + \dots + \delta_p (x - X_i)^p \end{aligned}$$

dengan δ_i konstan yang tidak diketahui.

Untuk mendapatkan estimasi model tersebut, digunakan metode *weighted least square*, dengan memminimumkan :

$$\begin{aligned} \sum_{i=1}^n \varepsilon_i^2 K_h(x - X_i) &= \sum_{i=1}^n \{y_i - m(X_i)\}^2 K_h(x - X_i) \\ &= \sum_{i=1}^n \left\{ y_i - \sum_{j=0}^p \delta_j (x - X_i)^j \right\}^2 K_h(x - X_i) \end{aligned} \quad (4.4)$$

dimana K adalah fungsi kernel dan h adalah bandwidth.

Misalkan

$$\mathbf{Y} = (Y_1, Y_2, Y_3, \dots, Y_n)^T$$

$$\boldsymbol{\delta} = (\delta_0, \delta_1, \delta_2, \dots, \delta_p)^T$$

$$K_x = diag [K_h(x - X_i), i = 1, 2, 3, \dots, n],$$

dan

$$\mathbf{X}_{p,x} = \begin{bmatrix} 1 & (x - X_1) & \cdots & (x - X_n)^p \\ 1 & (x - X_2) & \cdots & (x - X_n)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (x - X_n) & \cdots & (x - X_n)^p \end{bmatrix},$$

maka persamaan (4.4) dapat ditulis dalam bentuk matriks:

$$\begin{aligned} \mathbf{M} &= (\mathbf{Y} - \mathbf{X}_{p,x} \boldsymbol{\delta})^T \mathbf{K}_x (\mathbf{Y} - \mathbf{X}_{p,x} \boldsymbol{\delta}) \\ &= (\mathbf{Y}^T - (\mathbf{X}_{p,x} \boldsymbol{\delta})^T) \mathbf{K}_x (\mathbf{Y} - \mathbf{X}_{p,x} \boldsymbol{\delta}) \\ &= \mathbf{Y}^T \mathbf{K}_x \mathbf{Y} - \mathbf{Y}^T \mathbf{K}_x \mathbf{X}_{p,x} \boldsymbol{\delta} - (\mathbf{X}_{p,x} \boldsymbol{\delta})^T \mathbf{K}_x \mathbf{Y} + (\mathbf{X}_{p,x} \boldsymbol{\delta})^T \mathbf{K}_x \mathbf{X}_{p,x} \boldsymbol{\delta} \\ &= \mathbf{Y}^T \mathbf{K}_x \mathbf{Y} - (\mathbf{X}_{p,x} \boldsymbol{\delta})^T \mathbf{K}_x \mathbf{Y} - (\mathbf{X}_{p,x} \boldsymbol{\delta})^T \mathbf{K}_x \mathbf{Y} + (\mathbf{X}_{p,x} \boldsymbol{\delta})^T \mathbf{K}_x \mathbf{X}_{p,x} \boldsymbol{\delta} \\ &= \mathbf{Y}^T \mathbf{K}_x \mathbf{Y} - 2(\mathbf{X}_{p,x} \boldsymbol{\delta})^T \mathbf{K}_x \mathbf{Y} + (\mathbf{X}_{p,x} \boldsymbol{\delta})^T \mathbf{K}_x \mathbf{X}_{p,x} \boldsymbol{\delta} \\ &= \mathbf{Y}^T \mathbf{K}_x \mathbf{Y} - 2\boldsymbol{\delta}^T \mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{Y} + \boldsymbol{\delta}^T \mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{X}_{p,x} \boldsymbol{\delta} \end{aligned}$$

Untuk mendapatkan estimasi $\boldsymbol{\delta}$ maka kita turunkan \mathbf{M} terhadap $\boldsymbol{\delta}$ kemudian menyamakannya dengan 0.

$$\mathbf{M} = \mathbf{Y}^T \mathbf{K}_x \mathbf{Y} - 2\boldsymbol{\delta}^T \mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{Y} + \boldsymbol{\delta}^T \mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{X}_{p,x} \boldsymbol{\delta}$$

$$\frac{\partial M}{\partial \delta} = -2\mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{Y} + 2\mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{X}_{p,x} \boldsymbol{\delta} = 0$$

$$\Leftrightarrow \mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{X}_{p,x} \boldsymbol{\delta} = \mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{Y}$$

Dengan mengalikan masing-masing ruas dengan $(\mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{X}_{p,x})^{-1}$, didapat:

$$(\mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{X}_{p,x})^{-1} \mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{X}_{p,x} \boldsymbol{\delta} = (\mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{X}_{p,x})^{-1} \mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{Y}$$

Sehingga $\hat{\boldsymbol{\delta}} = (\mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{X}_{p,x})^{-1} \mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{Y}$

Estimator untuk $m(X_i)$ diberikan oleh:

$$\hat{m}_h(x; p) = \xi_1^T (\mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{X}_{p,x})^{-1} \mathbf{X}_{p,x}^T \mathbf{K}_x \mathbf{Y}$$

Dimana ξ_1 merupakan vektor $(1, 0, 0, \dots, 0)^T$ yang berukuran $(p+1) \times 1$.
(Wand dan Jones, 1995).

Jika ditetapkan $p = 0$ maka

$$\hat{m}_h(x, 0) = \xi_1^T (\mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{0,x})^{-1} \mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{Y} \quad (4.5)$$

Estimator yang diperoleh disebut dengan estimator nonparametrik *local constant* (Jia dan Schucany, 2004) dengan $\xi_1^T = (1)$ dan $\mathbf{X}_{0,x} = (1 \ 1 \ \dots \ 1)^T$.

Atau dengan perhitungan lebih lanjut akan diperoleh estimator nonparametrik kernel sebagai berikut:

$$\hat{m}_h(x) = \frac{n^{-1} \sum_{i=1}^n K_h(x - X_i) Y_i}{n^{-1} \sum_{i=1}^n K_h(x - X_i)}$$

Persamaan tersebut dapat ditulis: $\hat{m}_h(x) = n^{-1} \sum_{i=1}^n W_{hi}(x) Y_i$ (4.6)

Dengan

$$W_{hi}(x) = \frac{h^{-1} K\left(\frac{x - X_i}{h}\right)}{\hat{f}_h(x)}$$

dan

$$\hat{f}_h(x) = n^{-1} \sum_{i=1}^n K_h(x - X_i)$$

Berdasarkan persamaan (4.5), maka bias dari regresi nonparametrik kernel dapat dicari dengan :

$$\begin{aligned} E[\hat{m}_h(x; 0)] &= \xi_1^T (\mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{0,x})^{-1} \mathbf{X}_{0,x}^T \mathbf{K}_x E[Y] \\ &= \xi_1^T (\mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{0,x})^{-1} \mathbf{X}_{0,x}^T \mathbf{K}_x \Omega \end{aligned} \quad (4.7)$$

dimana $\Omega = (m(X_1), m(X_2), \dots, m(X_n))^T$

Selanjutnya kita deretkan $m(X_i)$ berdasarkan ekspansi Taylor disekitar titik x , didapat :

$$m(X_i) = m(x) + \frac{m'(x)}{1!}(X_i - x) + \frac{m''(x)}{2!}(X_i - x)^2 + \dots \quad (4.8)$$

dimana $i = 1, 2, 3, \dots, n$.

Jika ditulis dalam bentuk matriks, maka persamaan (4.8) menjadi :

$$\Omega = m(x)X_{0,x} + m'(x)X_{1,x} + \frac{1}{2!}m''(x)X_{2,x} + \dots \quad (4.9)$$

dimana

$$\Omega = \begin{bmatrix} m(X_1) \\ m(X_2) \\ \vdots \\ m(X_n) \end{bmatrix}, \quad X_{0,x} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad X_{1,x} = \begin{bmatrix} (X_1 - x) \\ (X_2 - x) \\ \vdots \\ (X_n - x) \end{bmatrix} \quad \text{dan} \quad X_{2,x} = \begin{bmatrix} (X_1 - x)^2 \\ (X_2 - x)^2 \\ \vdots \\ (X_n - x)^2 \end{bmatrix},$$

sehingga persamaan (4.7) menjadi :

$$\begin{aligned} E[\hat{m}_h(x; 0)] &= \xi_1^T (X_{0,x}^T K_x X_{0,x})^{-1} X_{0,x}^T K_x \left\{ m(x)X_{0,x} + m'(x)X_{1,x} + \frac{1}{2!}m''(x)X_{2,x} + \dots \right\} \\ &= m(x) + m'(x)\xi_1^T (X_{0,x}^T K_x X_{0,x})^{-1} X_{0,x}^T K_x X_{1,x} + \\ &\quad + \frac{1}{2}m''(x)\xi_1^T (X_{0,x}^T K_x X_{0,x})^{-1} X_{0,x}^T K_x X_{2,x} + \dots \end{aligned}$$

Karena $\xi_1^T = (1)$, maka persamaan diatas dapat ditulis menjadi :

$$\begin{aligned} E[\hat{m}_h(x; 0)] &= m(x) + m'(x)(X_{0,x}^T K_x X_{0,x})^{-1} X_{0,x}^T K_x X_{1,x} + \\ &\quad + \frac{1}{2}m''(x)(X_{0,x}^T K_x X_{0,x})^{-1} X_{0,x}^T K_x X_{2,x} + \dots \end{aligned}$$

Dengan merujuk pada persamaan (4.1) bahwa Bias $\hat{m}_h(x) = E(\hat{m}_h(x)) - m(x)$,

maka diperoleh :

$$\begin{aligned} \text{Bias}[\hat{m}_h(x)] &= m'(x)(\mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{0,x})^{-1} \mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{1,x} + \\ &+ \frac{1}{2} m''(x)(\mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{0,x})^{-1} \mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{2,x} + \dots \end{aligned} \quad (4.10)$$

Untuk menyelesaikan persamaan (4.10) lebih lanjut, maka dimisalkan :

$$\begin{aligned} A &= n^{-1} \mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{0,x} \\ &= n^{-1} \left(\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} K_h(x-X_1) & 0 & \cdots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & K_h(x-X_n) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right) \\ &= n^{-1} \left(\begin{bmatrix} K_h(x-X_1) & K_h(x-X_2) & \cdots & K_h(x-X_n) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right) \\ &= n^{-1} \sum_{i=1}^n K_h(x-X_i) = \hat{\nu}_0 \end{aligned} \quad (4.11)$$

$$\begin{aligned} B &= n^{-1} \mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{1,x} \\ &= n^{-1} \left(\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} K_h(x-X_1) & 0 & \cdots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & K_h(x-X_n) \end{bmatrix} \begin{bmatrix} (X_1-x) \\ (X_2-x) \\ \vdots \\ (X_n-x) \end{bmatrix} \right) \end{aligned}$$

$$\begin{aligned}
&= n^{-1} \left(\begin{bmatrix} K_h(x-X_1) & K_h(x-X_2) & \cdots & K_h(x-X_n) \end{bmatrix} \begin{bmatrix} (X_1-x) \\ (X_2-x) \\ \vdots \\ (X_n-x) \end{bmatrix} \right) \\
&= n^{-1} \left[\sum_{i=1}^n (X_i - x) K_h(x - X_i) \right] = \hat{\nu}_1
\end{aligned} \tag{4.12}$$

$$C = n^{-1} \mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{2,x}$$

$$\begin{aligned}
&= n^{-1} \left(\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} K_h(x-X_1) & 0 & \cdots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & K_h(x-X_n) \end{bmatrix} \begin{bmatrix} (X_1-x)^2 \\ (X_2-x)^2 \\ \vdots \\ (X_n-x)^2 \end{bmatrix} \right) \\
&= n^{-1} \left(\begin{bmatrix} K_h(x-X_1) & K_h(x-X_2) & \cdots & K_h(x-X_n) \end{bmatrix} \begin{bmatrix} (X_1-x)^2 \\ (X_2-x)^2 \\ \vdots \\ (X_n-x)^2 \end{bmatrix} \right) \\
&= n^{-1} \left[\sum_{i=1}^n (X_i - x)^2 K_h(x - X_i) \right] = \hat{\nu}_2
\end{aligned} \tag{4.13}$$

Secara umum persamaan (4.11), (4.12) dan (4.13) dapat ditulis :

$$\hat{\nu}_k = n^{-1} \sum_{i=1}^n (X_i - x)^k K_h(x - X_i) , k = 0, 1, 2, \dots$$

Karena $K_h(x - X_i) = \frac{1}{h} K\left(\frac{X_i - x}{h}\right)$, maka:

$$\hat{\nu}_k = n^{-1} \sum_{i=1}^n (X_i - x)^k \frac{1}{h} K\left(\frac{X_i - x}{h}\right).$$



Misal $u_i = \frac{X_i - x}{h}$, maka $X_i - x = hu_i$, sehingga

$$\hat{v}_k = \sum_{i=1}^n n^{-1} (hu_i)^k \frac{1}{h} K(u_i)$$

$$\text{Selanjutnya } \Delta u_i = \frac{X_i - x}{h} - \frac{X_{i-1} - x}{h} = \frac{X_i - X_{i-1}}{h}, \text{ sehingga } h\Delta u_i = X_i - X_{i-1}.$$

Karena fungsi K merupakan fungsi yang kontinu, terbatas, simetrik dan terintegral, maka :

$$\hat{v}_k = \sum_{i=1}^n h \Delta u_i (hu_i)^k \frac{1}{h} K(u_i)$$

$$= \sum_{i=1}^n (hu_i)^k K(u_i) \Delta u_i$$

$$= \int (hu)^k K(u) du \text{ dengan syarat } \lim_{\Delta u \rightarrow 0} \sum_{i=1}^n (hu_i)^k K(u_i) \Delta u_i \quad (4.14)$$

Dengan memperhatikan syarat kernel berordo 2 yaitu: $\int K(u) du = 1$ dan

$\int u K(u) du = 0$, maka (4.11), (4.12) dan (4.13) menjadi :

$$A = n^{-1} \mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{0,x}^T = \hat{v}_0 = \int K(u) du = 1 \quad (4.15)$$

$$B = n^{-1} \mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{1,x}^T = h \int u K(u) du = 0 \quad (4.16)$$

$$C = n^{-1} \mathbf{X}_{0,x}^T \mathbf{K}_x \mathbf{X}_{2,x}^T = \hat{v}_2 = h^2 \int u^2 K(u) du \quad (4.17)$$

Selanjutnya dengan merujuk pada persamaan (4.15), (4.16) dan (4.17) maka persamaan (4.10) menjadi :

$$\begin{aligned}
 Bias[\hat{m}_h(x)] &= m'(x)(X_{0,x}^T K_x X_{0,x})^{-1} X_{0,x}^T K_x X_{1,x} + \frac{1}{2} m''(x)(X_{0,x}^T K_x X_{0,x})^{-1} X_{0,x}^T K_x X_{2,x} + \dots \\
 &= m'(x)A^{-1}B + \frac{1}{2} m''(x)A^{-1}C + \dots \\
 &= m'(x).(1)^{-1}.0 + \frac{1}{2} m''(x).(1)^{-1} h^2 \int u^2 K(u) du + \dots \\
 &= \frac{1}{2} m''(x)h^2 \int u^2 K(u) du + \dots \\
 &= \frac{1}{2} h^2 m''(x) \int u^2 K(u) du + o(h), \quad h \rightarrow 0
 \end{aligned}$$

Jadi $Bias[\hat{m}_h(x)] = \frac{1}{2} h^2 m''(x) \int u^2 K(u) du + o(h)$ (4.18)

Langkah selanjutnya akan ditentukan variansi dari estimator nonparametrik kernel *local constan*. Dari persamaan (4.5), maka :

$$\begin{aligned}
 Var[\hat{m}_h(x)] &= Var[\xi_1^T (X_{0,x}^T K_x X_{0,x})^{-1} X_{0,x}^T K_x Y] \\
 &= \xi_1^T (X_{0,x}^T K_x X_{0,x})^{-1} X_{0,x}^T K_x Var[Y] [\xi_1^T (X_{0,x}^T K_x X_{0,x})^{-1} X_{0,x}^T K_x]^T \\
 &= \xi_1^T (X_{0,x}^T K_x X_{0,x})^{-1} (X_{0,x}^T K_x) Var[Y] K_x^T X_{0,x} (X_{0,x}^T K_x X_{0,x})^{-1} \xi_1 \\
 &= \xi_1^T (X_{0,x}^T K_x X_{0,x})^{-1} (X_{0,x}^T K_x) \sigma^2 I K_x^T X_{0,x} (X_{0,x}^T K_x X_{0,x})^{-1} \xi_1 \\
 &= \xi_1^T \sigma^2 (X_{0,x}^T K_x X_{0,x})^{-1} X_{0,x}^T K_x K_x^T X_{0,x} (X_{0,x}^T K_x X_{0,x})^{-1} \xi_1
 \end{aligned}$$

Dengan menggunakan pendekatan yang sama, maka varians dapat dihitung sebagai berikut:

Misal :

$$R = n^{-1} X_{0,x}^T K_x K_x^T X_{0,x}$$

maka

$$\begin{aligned}
 R &= n^{-1} [1 \ 1 \ \dots \ 1] \begin{bmatrix} K_h(x - X_1)^2 & 0 & \dots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & K_h(x - X_n)^2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \\
 &= n^{-1} \sum_{i=1}^n K_h(x - X_i)^2 \\
 &= \sum_{i=1}^n n^{-1} (X_i - x)^0 h^{-2} K\left(\frac{X_i - x}{h}\right)^2 \\
 &= \sum_{i=1}^n n^{-1} (hu_i)^0 h^{-2} K(u_i)^2 \\
 &= \sum_{i=1}^n h^{-1} K(u_i)^2 \Delta u_i \\
 &= h^{-1} \int K(u)^2 du
 \end{aligned} \tag{4.19}$$

Dengan memperhatikan persamaan (4.15) dan (4.19) maka diperoleh :

$$\begin{aligned}
 Var[\hat{m}(x)] &= \sigma^2 n^{-1} \xi_1^T (X_{0,x}^T K_x K_x^T X_{0,x})^{-1} X_{0,x}^T K_x K_x^T X_{0,x} (X_{0,x}^T K_x K_x^T X_{0,x})^{-1} \xi_1 \\
 &= \sigma^2 n^{-1} \xi_1^T A^{-1} R A^{-1} \xi_1 \\
 &= \sigma^2 n^{-1} (1)^{-1} h^{-1} \int K(u)^2 du (1)^{-1} \\
 &= \sigma^2 n^{-1} h^{-1} \int K(u)^2 du \\
 &= \frac{1}{nh} \sigma^2 \int K(u)^2 du
 \end{aligned} \tag{4.20}$$

Dari bias dan varians yang diperoleh dari persamaan (4.18) dan (4.20), maka

$$\begin{aligned} Q &= (\text{Bias}(\hat{m}(x)))^2 + \text{Var}(\hat{m}(x)) \\ &= \left[\frac{1}{2} h^2 m''(x) \int u^2 K(u) du \right]^2 + \left[\frac{1}{nh} \sigma^2 \int K(u)^2 du \right] \end{aligned}$$

Dengan memisalkan

$$\mu_2(K) = \int u^2 K(u) du$$

dan

$$\|K\|_2^2 = \int K(u)^2 du,$$

maka

$$Q = \frac{h^4}{4} (\mu_2(K))^2 (m''(x))^2 + \frac{1}{nh} \sigma^2 \|K\|_2^2 \quad (4.21)$$

Bandwidth (h) yang optimal diperoleh dari meminimumkan Q terhadap h , kemudian menyamakannya dengan 0.

$$\frac{\partial Q}{\partial h} = h^3 (\mu_2(K))^2 (m''(x))^2 - \frac{\sigma^2}{nh^2} \|K\|_2^2 = 0$$

$$\Leftrightarrow h^3 (\mu_2(K))^2 (m''(x))^2 = \frac{\sigma^2}{nh^2} \|K\|_2^2$$

$$h^5 = \frac{\sigma^2 \|K\|_2^2}{n (\mu_2(K))^2 (m''(x))^2}$$

$$h_{opt} = n^{-\frac{1}{5}} \left(\frac{\sigma^2 \|K\|_2^2}{(\mu_2(K))^2 (m''(x))^2} \right)^{1/5}$$

$$\text{Jadi diperoleh } h_{opt} = n^{-\frac{1}{5}} \left(\frac{\sigma^2 \|K\|_2^2}{(\mu_2(K))^2 (m''(x))^2} \right)^{1/5} \quad (4.22)$$

$$\text{Jadi diperoleh } h_{opt} = n^{-\frac{1}{5}} \left(\frac{\sigma^2 \|K\|_2^2}{(\mu_2(K))^2 (m''(x))^2} \right)^{1/5} \quad (4.22)$$

Berdasarkan uraian di atas, penentuan *bandwidth* optimal tergantung pada $m''(x)$ dan σ^2 yang pada kenyataannya tidak diketahui, sehingga pendekatan ini sulit dilakukan. Untuk mengatasi hal ini maka Lee dan Hill (2001) serta Ye, et al (2006) mendekati m dengan polinomial derajat p .

$$m(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + \dots + \beta_p x^p \quad (4.23)$$

Untuk mendapatkan estimator dari β , maka digunakan metode least square sehingga didapat :

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (4.24)$$

Akibatnya diperoleh estimasi untuk m :

$$\hat{m}(x) = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 + \hat{\beta}_3 x^3 + \hat{\beta}_4 x^4 + \dots + \hat{\beta}_p x^p$$

Untuk mendapatkan $\hat{m}''(x)$, kita turunkan $\hat{m}(x)$ sebanyak dua kali terhadap x .

$$\begin{aligned} \hat{m}(x) &= \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 + \hat{\beta}_3 x^3 + \hat{\beta}_4 x^4 + \dots + \hat{\beta}_p x^p \\ \hat{m}'(x) &= \hat{\beta}_1 + 2\hat{\beta}_2 x + 3\hat{\beta}_3 x^2 + 4\hat{\beta}_4 x^3 + \dots + p\hat{\beta}_p x^{p-1} \\ \hat{m}''(x) &= 2\hat{\beta}_2 + 6\hat{\beta}_3 x + 12\hat{\beta}_4 x^2 + \dots + p(p-1)\hat{\beta}_p x^{(p-1)-1} \end{aligned} \quad (4.25)$$

Sedangkan untuk taksiran variansnya

$$\hat{\sigma}^2 = \frac{1}{n-p} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4.26)$$

Dengan mensubstitusikan persamaan (4.25) dan (4.26) ke dalam (4.22) didapat :

Langkah selanjutnya adalah mendapatkan nilai $\|K\|_2^2 = \int K(u)^2 du$ dan $\mu_2(K) = \int u^2 K(u) du$ pada persamaan (4.27). Jika digunakan kernel Gaussian, diperoleh :

$$\begin{aligned}\|K\|_2^2 &= \int K(u)^2 du \\ &= \int \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \right)^2 du \\ &= \int \frac{1}{2\pi} e^{-u^2} du \\ &= \frac{1}{2\pi} \int e^{-u^2} du\end{aligned}$$

Misalkan

$$I = \int_{-\infty}^{\infty} e^{-u^2} du \text{ maka } I = 2 \int_0^{\infty} e^{-u^2} du$$

$$\begin{aligned}I^2 &= 2 \int_0^{\infty} e^{-u^2} du \cdot 2 \int_0^{\infty} e^{-v^2} dv \\ &= 4 \int_0^{\infty} \int_0^{\infty} e^{-(u^2+v^2)} du dv \\ &= \pi\end{aligned}$$

sehingga didapat

$$\begin{aligned}\|K\|_2^2 &= \frac{1}{2\pi} \int e^{-u^2} du \\ &= \frac{1}{2\pi} \sqrt{\pi} \\ &= \frac{1}{2\sqrt{\pi}}\end{aligned}\tag{4.28}$$

Sedangkan

$$\begin{aligned}
 \mu_2(K) &= \int u^2 K(u) du \\
 &= \int u^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
 &= \frac{1}{\sqrt{2\pi}} \int u^2 e^{-\frac{1}{2}u^2} du \\
 &= \frac{1}{\sqrt{2\pi}} \int u \cdot ue^{-\frac{1}{2}u^2} du
 \end{aligned}$$

Persamaan diatas dapat diselesaikan dengan integral parsial :

$$\int U dV = U \cdot V - \int V dU$$

Misal untuk $U = u$ maka $dU = du$ dan $dV = ue^{-\frac{1}{2}u^2} du$ maka $V = -e^{-\frac{1}{2}u^2}$ oleh karena itu,

$$\begin{aligned}
 \mu_2(K) &= \frac{1}{\sqrt{2\pi}} \int u \cdot ue^{-\frac{1}{2}u^2} du \\
 &= \frac{1}{\sqrt{2\pi}} \left(-ue^{-\frac{1}{2}u^2} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2} du \right) \\
 &= \frac{1}{\sqrt{2\pi}} \left(-ue^{-\frac{1}{2}u^2} \Big|_{-\infty}^{\infty} + \sqrt{2\pi} \right) \\
 &= \frac{1}{\sqrt{2\pi}} (0 + \sqrt{2\pi}) \\
 &= 1
 \end{aligned} \tag{4.29}$$

Selanjutnya untuk mendapatkan *Rule of Thumb (ROT)* maka kita substitusikan persamaan (4.28) dan (4.29) ke dalam persamaan (4.27) didapat :

$$\hat{h}_{ROT} = n^{-\frac{1}{5}} \left(\frac{\frac{1}{2\sqrt{\pi}} \hat{\sigma}^2}{(\hat{m}''(x))^2} \right)^{1/5} \quad (4.30)$$

4.2 Bandwidth Shibata

Pada bagian kedua ini, kita akan menunjukkan prosedur pemilihan *bandwidth* yang lain. Pertama-tama kita definisikan dulu fungsi *loss*.

Fungsi *Loss* didefinisikan sebagai

$$L(h) = n^{-1} \sum_{j=1}^n [m(X_j) - \hat{m}_h(X_j)]^2 \quad (4.31)$$

Rata-rata atau nilai harapan dari *Loss* adalah $R(h) = E[L(h)]$ yang disebut sebagai *Risk*.

$$R(h) = E(n^{-1} \sum_{j=1}^n [m(X_j) - \hat{m}_h(X_j)]^2) \quad (4.32)$$

Baik $R(h)$ maupun $L(h)$ merupakan kriteria yang digunakan untuk mengukur kebaikan suatu estimator. Nilai terkecil dari kriteria adalah indikasi dari estimator yang terbaik. Menurut Eubank (1988), pengukuran yang lain yang berhubungan erat dengan $R(h)$ adalah prediksi *risk* yang kadang-kadang disebut prediksi *mean square error* yang berhubungan erat dengan $R(h)$, yaitu:

$$p(h) = \sigma^2 + R(h). \quad (4.33)$$

Sehingga sebuah estimator yang meminimumkan $R(h)$ juga akan meminimumkan $p(h)$ dan juga sebaliknya estimator yang meminimumkan $p(h)$ juga akan meminimumkan $R(h)$.

Jadi hal yang mungkin untuk mengestimasi $p(h)$ adalah dengan *mean square error*.

$$MSE(h) = n^{-1} \sum_{j=1}^n (Y_j - \hat{m}_h(X_j))^2 \quad (4.34)$$

Akan tetapi $MSE(h)$ ini merupakan estimator yang bias terhadap $p(h)$. Untuk melihat bahwa $MSE(h)$ bias terhadap $p(h)$, kita uraikan sebagai berikut :

$$\begin{aligned} MSE(h) &= n^{-1} \sum_{j=1}^n [Y_j - \hat{m}_h(X_j)]^2 \\ &= n^{-1} \sum_{j=1}^n [Y_j^2 + \hat{m}_h^2(X_j) - 2(Y_j \hat{m}_h(X_j))] \end{aligned}$$

Dengan memperhatikan $Y_j = m(X_j) + \varepsilon_j$, maka :

$$\begin{aligned} MSE(h) &= n^{-1} \sum_{j=1}^n [(m(X_j) + \varepsilon_j)^2 + \hat{m}_h^2(X_j) - 2((m(X_j) + \varepsilon_j)\hat{m}_h(X_j))] \\ &= n^{-1} \sum_{j=1}^n [m^2(X_j) + \varepsilon_j^2 + 2m(X_j)\varepsilon_j + \hat{m}_h^2(X_j) + \\ &\quad - 2(m(X_j)\hat{m}_h(X_j) + \varepsilon_j\hat{m}_h(X_j))] \\ &= n^{-1} \sum_{j=1}^n [m^2(X_j) + \varepsilon_j^2 + 2m(X_j)\varepsilon_j + \hat{m}_h^2(X_j) + \\ &\quad - 2m(X_j)\hat{m}_h(X_j) - 2\varepsilon_j\hat{m}_h(X_j)] \\ &= n^{-1} \sum_{j=1}^n [\varepsilon_j^2 + m^2(X_j) + \hat{m}_h^2(X_j) - 2m(X_j)\hat{m}_h(X_j) \\ &\quad - 2\varepsilon_j\hat{m}_h(X_j) + 2m(X_j)\varepsilon_j] \\ &= n^{-1} \sum_{j=1}^n [\varepsilon_j^2 + m^2(X_j) + \hat{m}_h^2(X_j) - 2m(X_j)\hat{m}_h(X_j) + \\ &\quad - 2\varepsilon_j(\hat{m}_h(X_j) - m(X_j))] \quad (4.35) \end{aligned}$$

Karena

$$\begin{aligned} L(h) &= n^{-1} \sum_{j=1}^n [m(X_j) - \hat{m}_h(X_j)]^2 \\ &= n^{-1} \sum_{j=1}^n [m^2(X_j) + \hat{m}_h^2(X_j) - 2m(X_j)\hat{m}_h(X_j)] \end{aligned}$$

maka persamaan (4.35) menjadi :

$$MSE(h) = n^{-1} \sum_{j=1}^n \varepsilon_j^2 + L(h) - 2n^{-1} \sum_{j=1}^n \varepsilon_j (\hat{m}_h(X_j)) - m(X_j)).$$

Dengan memperhatikan bahwa error ε_i adalah iid dengan ekspektasi nol dan variansi σ^2 maka :

$$\begin{aligned} E[MSE(h)] &= n^{-1} \sum_{j=1}^n E(\varepsilon_j^2) + E(L(h)) - 2n^{-1} E\left(\sum_{j=1}^n \varepsilon_j (\hat{m}_h(X_j)) - m(X_j)\right) \\ &= \sigma^2 + R(h) + EC_{ln}, \end{aligned} \quad (4.36)$$

dengan

$$EC_{ln} = -2n^{-1} E\left(\sum_{j=1}^n \varepsilon_j (\hat{m}_h(X_j) - m(X_j))\right). \quad (4.37)$$

Selanjutnya untuk menguraikan persamaan (4.37) lebih lanjut, maka perlu diperhatikan kembali persamaan (4.6) yaitu :

$$\hat{m}_h(x) = n^{-1} \sum_{i=1}^n W_{hi}(x) Y_i$$

yang berakibat :

$$\hat{m}_h(X_j) = n^{-1} \sum_{i=1}^n W_{hi}(X_j) Y_i,$$

sehingga persamaan (4.37) dapat ditulis :

$$EC_{ln} = -2n^{-1}E\left(\sum_{j=1}^n \varepsilon_j (\hat{m}_h(X_j) - m(X_j))\right)$$

$$= -2n^{-1}E\left(\sum_{j=1}^n \varepsilon_j \left[n^{-1}\sum_{i=1}^n W_{hi}(X_i)Y_i - m(X_j)\right]\right)$$

Karena $Y_i = m(X_i) + \varepsilon_i$, maka

$$EC_{ln} = -2n^{-1}E\left(\sum_{j=1}^n \varepsilon_j \left[n^{-1}\sum_{i=1}^n W_{hi}(X_i)(m(X_i) + \varepsilon_i) - m(X_j)\right]\right)$$

$$= -2n^{-1}E\left(\sum_{j=1}^n \varepsilon_j \left[n^{-1}\sum_{i=1}^n W_{hi}(X_i)m(X_i) + n^{-1}\sum_{i=1}^n W_{hi}(X_i)\varepsilon_i - m(X_j)\right]\right)$$

$$= -2n^{-1}\sum_{j=1}^n E(\varepsilon_j)n^{-1}\sum_{i=1}^n W_{hi}(X_i)m(X_i) +$$

$$-2n^{-1}E\left(\sum_{j=1}^n \varepsilon_j n^{-1}\sum_{i=1}^n W_{hi}(X_i)\varepsilon_i\right) + 2n^{-1}\sum_{j=1}^n E(\varepsilon_j)m(X_j) \quad (4.38)$$

Selanjutnya dengan memperhatikan bahwa $E(\varepsilon_j) = 0$ dan $E(\varepsilon_j^2) = \sigma^2$, maka persamaan (4.38) menjadi :

$$EC_{ln} = -2n^{-2}E\left(\sum_{j=1}^n \varepsilon_j \sum_{i=1}^n \varepsilon_i W_{hi}(X_j)\right)$$

$$= -2n^{-2}E\left(\sum_{j=1}^n \varepsilon_j [\varepsilon_1 W_{h1}(X_j) + \varepsilon_2 W_{h2}(X_j) + \dots + \varepsilon_n W_{hn}(X_j)]\right)$$

$$= -2n^{-2}E\left(\sum_{j=1}^n [\varepsilon_j \varepsilon_1 W_{h1}(X_j) + \varepsilon_j \varepsilon_2 W_{h2}(X_j) + \dots + \varepsilon_j \varepsilon_n W_{hn}(X_j)]\right)$$

$$= -2n^{-2}E[\{\varepsilon_1 \varepsilon_1 W_{h1}(X_1) + \varepsilon_2 \varepsilon_1 W_{h1}(X_2) + \dots + \varepsilon_n \varepsilon_1 W_{h1}(X_n)\} +$$

$$+ \{\varepsilon_1 \varepsilon_2 W_{h2}(X_1) + \varepsilon_2 \varepsilon_2 W_{h2}(X_2) + \dots + \varepsilon_n \varepsilon_2 W_{h2}(X_n)\} + \dots +$$

$$+ \{\varepsilon_1 \varepsilon_n W_{hn}(X_1) + \varepsilon_2 \varepsilon_n W_{hn}(X_2) + \dots + \varepsilon_n \varepsilon_n W_{hn}(X_n)\}]$$

$$= -2n^{-2} \left[\{\sigma^2 W_{h1}(X_1)\} + \{\sigma^2 W_{h2}(X_2)\} + \dots + \{\sigma^2 W_{hn}(X_n)\} \right]$$

$$= -2n^{-2} \sigma^2 \sum_{j=1}^n W_{hj}(X_j)$$

sehingga

$$E[MSE(h)] = \sigma^2 + R(h) + EC_{1n}$$

$$= \sigma^2 + R(h) - 2n^{-2} \sigma^2 \sum_{j=1}^n W_{hj}(X_j)$$

$$= p(h) - 2n^{-2} \sigma^2 \sum_{j=1}^n W_{hj}(X_j)$$

$$\neq p(h)$$

Dari persamaan diatas terlihat bahwa $MSE(h)$ adalah bias untuk $p(h)$, dengan nilai biasnya :

$$-2n^{-2} \sigma^2 \sum_{j=1}^n W_{hj}(X_j). \quad (4.39)$$

Untuk mengatasi hal ini maka kita harus memodifikasi $MSE(h)$ sedemikian rupa sehingga suku bias seperti pada persamaan (4.39) terhapuskan (Härdle, 1990 dan Rice (1984) serta Härdle, 1991). Untuk cara ini, $MSE(h)$ diberi *penalizing function sibata's* sehingga $MSE(h)$ menjadi :

$$\begin{aligned} G(h) &= n^{-1} \sum_{j=1}^n (Y_j - \hat{m}_h(X_j))^2 \cdot (1 + 2n^{-1}W_{hj}(X_j)) \\ &= n^{-1} \sum_{j=1}^n (Y_j^2 + \hat{m}_h^2(X_j) - 2Y_j \hat{m}_h(X_j) \cdot (1 + 2n^{-1}W_{hj}(X_j))) \\ &= n^{-1} \sum_{j=1}^n \left[(m(X_j) + \varepsilon_j)^2 + \hat{m}_h^2(X_j) - 2(m(X_j) + \varepsilon_j) \hat{m}_h(X_j) \cdot (1 + 2n^{-1}W_{hj}(X_j)) \right] \end{aligned}$$

$$\begin{aligned}
&= n^{-1} \sum_{j=1}^n [(m^2(X_j) + \varepsilon_j^2 + 2m(X_j)\varepsilon_j + \hat{m}_h^2(X_j) + \\
&\quad - 2m(X_j)\hat{m}_h(X_j) - 2\varepsilon_j\hat{m}_h(X_j) \cdot (1 + 2n^{-1}W_{hj}(X_j))] \\
&= n^{-1} \sum_{j=1}^n [\varepsilon_j^2 + m^2(X_j) + \hat{m}_h^2(X_j) - 2m(X_j)\hat{m}_h(X_j) + \\
&\quad - 2\varepsilon_j\hat{m}_h(X_j) + 2m(X_j)\varepsilon_j \cdot (1 + 2n^{-1}W_{hj}(X_j))] \\
&= n^{-1} \sum_{j=1}^n [\varepsilon_j^2 + m^2(X_j) + \hat{m}_h^2(X_j) - 2m(X_j)\hat{m}_h(X_j) + \\
&\quad + 2\varepsilon_j(m(X_j) - \hat{m}_h(X_j)) \cdot (1 + 2n^{-1}W_{hj}(X_j))] \\
&= n^{-1} \sum_{j=1}^n [\varepsilon_j^2 + (m(X_j) - \hat{m}_h(X_j))^2 + 2\varepsilon_j(m(X_j) - \hat{m}_h(X_j)) \cdot (1 + 2n^{-1}W_{hj}(X_j))] \\
&= n^{-1} \sum_{j=1}^n \left[\varepsilon_j^2 + (m(X_j) - \hat{m}_h(X_j))^2 + 2\varepsilon_j(m(X_j) - \hat{m}_h(X_j)) + \right. \\
&\quad \left. + \varepsilon_j^2 \cdot 2n^{-1}W_{hj}(X_j) + (m(X_j) - \hat{m}_h(X_j))^2 \cdot 2n^{-1}W_{hj}(X_j) + \right. \\
&\quad \left. + 2\varepsilon_j(m(X_j) - \hat{m}_h(X_j)) \cdot 2n^{-1}W_{hj}(X_j) \right] \\
&= \left[n^{-1} \sum_{j=1}^n \varepsilon_j^2 + n^{-1} \sum_{j=1}^n (m(X_j) - \hat{m}_h(X_j))^2 + 2n^{-1} \sum_{j=1}^n \varepsilon_j(m(X_j) - \hat{m}_h(X_j)) + \right. \\
&\quad \left. + 2n^{-2} \sum_{j=1}^n \varepsilon_j^2 W_{hj}(X_j) + 2n^{-1} \sum_{j=1}^n (m(X_j) - \hat{m}_h(X_j))^2 \cdot n^{-1}W_{hj}(X_j) + \right. \\
&\quad \left. + 4n^{-2} \sum_{j=1}^n \varepsilon_j(m(X_j) - \hat{m}_h(X_j)) W_{hj}(X_j) \right]
\end{aligned}$$

Dengan cara yang sama, karena error ε adalah iid dengan ekspektasi nol dan variansi σ^2 maka persamaan diatas menjadi:

$$\begin{aligned}
E[G(h)] &= \left[n^{-1} \sum_{j=1}^n E(\varepsilon_j)^2 + n^{-1} \sum_{j=1}^n E\{(m(X_j) - \hat{m}_h(X_j))^2\} + \right. \\
&\quad + 2n^{-1} \sum_{j=1}^n E\{\varepsilon_j(m(X_j) - \hat{m}_h(X_j))\} + 2n^{-2} \sum_{j=1}^n E(\varepsilon_j)^2 W_{hj}(X_j) + \\
&\quad + 2n^{-1} \sum_{j=1}^n E\{(m(X_j) - \hat{m}_h(X_j))^2\} \cdot n^{-1} W_{hj}(X_j) + \\
&\quad \left. + 4n^{-2} \sum_{j=1}^n E\{\varepsilon_j(m(X_j) - \hat{m}_h(X_j)) W_{hj}(X_j)\} \right] \\
&= \sigma^2 + R(h) + P + Q + R + S, \tag{4.40}
\end{aligned}$$

dimana

$$\begin{aligned}
P &= 2n^{-1} \sum_{j=1}^n E\{\varepsilon_j(m(X_j) - \hat{m}_h(X_j))\} \\
Q &= 2n^{-2} \sum_{j=1}^n E(\varepsilon_j)^2 W_{hj}(X_j) \\
R &= 2n^{-1} \sum_{j=1}^n E\{(m(X_j) - \hat{m}_h(X_j))^2\} \cdot n^{-1} W_{hj}(X_j) \\
S &= 4n^{-2} \sum_{j=1}^n E\{\varepsilon_j(m(X_j) - \hat{m}_h(X_j)) W_{hj}(X_j)\}
\end{aligned}$$

Dengan menguraikan masing-masing persamaan didapat :

$$P = -2n^{-2}\sigma^2 \sum_{j=1}^n W_{hj}(X_j) \tag{4.41}$$

$$\begin{aligned}
Q &= 2n^{-2} \sum_{j=1}^n E(\varepsilon_j^2) W_{hj}(X_j) \\
&= 2n^{-2}\sigma^2 \sum_{j=1}^n W_{hj}(X_j) \tag{4.42}
\end{aligned}$$

sedangkan

$$R = 2n^{-1} \sum_{j=1}^n E\{(m(X_j) - \hat{m}_h(X_j))^2\} \cdot n^{-1} W_{hj}(X_j) \quad (4.43a)$$

Karena

$$n^{-1} W_{hj}(X_j) = n^{-1} h^{-1} K(0) / \hat{f}_h(X_j), \text{ (Härdle, 1990)}$$

maka dengan memperhatikan juga bahwa :

$$R(h) = n^{-1} \sum_{j=1}^n E\{(m(X_j) - \hat{m}_h(X_j))^2\},$$

maka persamaan (4.2.18a) dapat ditulis

$$R = 2n^{-1} h^{-1} \frac{K(0)}{\hat{f}_h(X_j)} R(h) \quad (4.43b)$$

dan

$$\begin{aligned} S &= 4n^{-2} \sum_{j=1}^n E\{\varepsilon_j(m(X_j) - \hat{m}_h(X_j))W_{hj}(X_j)\} \\ &= 4n^{-2} \sum_{j=1}^n E\{\varepsilon_j m(X_j)W_{hj}(X_j)\} - 4n^{-2} \sum_{j=1}^n E\{\varepsilon_j \hat{m}_h(X_j)W_{hj}(X_j)\} \\ &= -4n^{-2} \sum_{j=1}^n E\{\varepsilon_j \hat{m}_h(X_j)W_{hj}(X_j)\}, \end{aligned}$$

Karena

$$\hat{m}_h(x) = n^{-1} \sum_{i=1}^n W_{hi}(x) Y_i,$$

maka

$$S = -4n^{-2} \sum_{j=1}^n E\{\varepsilon_j \left(n^{-1} \sum_{i=1}^n W_{hi}(X_j) Y_i \right) W_{hj}(X_j)\}$$

Dengan memperhatikan bahwa $Y_i = m(X_i) + \varepsilon_i$ maka :

$$\begin{aligned}
S &= -4n^{-2} \sum_{j=1}^n W_{hj}(X_j) E\{\varepsilon_j \left(n^{-1} \sum_{i=1}^n W_{hi}(X_i) [m(X_i) + \varepsilon_i] \right)\} \\
&= -4n^{-2} \sum_{j=1}^n W_{hj}(X_j) E\{\varepsilon_j \left[n^{-1} \sum_{i=1}^n W_{hi}(X_i) m(X_i) + n^{-1} \sum_{i=1}^n W_{hi}(X_i) \varepsilon_i \right]\} \\
&= -4n^{-2} \sum_{j=1}^n W_{hj}(X_j) E\{\varepsilon_j n^{-1} \sum_{i=1}^n W_{hi}(X_i) m(X_i)\} + \\
&\quad -4n^{-2} \sum_{j=1}^n W_{hj}(X_j) E\{\varepsilon_j n^{-1} \sum_{i=1}^n W_{hi}(X_i) \varepsilon_i\}
\end{aligned}$$

Karena $E(\varepsilon_j) = 0$ maka suku pertama sama dengan nol, sehingga

$$\begin{aligned}
S &= -4n^{-2} \sum_{j=1}^n W_{hj}(X_j) E\{\varepsilon_j n^{-1} \sum_{i=1}^n \varepsilon_i W_{hi}(X_i)\} \\
&= -4n^{-3} \sum_{j=1}^n W_{hj}(X_j) E\{\varepsilon_j (\varepsilon_1 W_{h1}(X_j) + \dots + \varepsilon_n W_{hn}(X_j))\} \\
&= -4n^{-3} E \sum_{j=1}^n (\varepsilon_j \varepsilon_1 W_{h1}(X_j) W_{hj}(X_j) + \dots + \varepsilon_j \varepsilon_n W_{hn}(X_j) W_{hj}(X_j)) \\
&= -4n^{-3} E (\varepsilon_1 \varepsilon_1 W_{h1}(X_1) W_{h1}(X_1) + \dots + \varepsilon_n \varepsilon_n W_{hn}(X_n) W_{hn}(X_n)) \\
&= -4n^{-3} E (\varepsilon_1^2 W_{h1}^2(X_1) + \dots + \varepsilon_n^2 W_{hn}^2(X_n)) \\
&= -4n^{-3} (\sigma^2 W_{h1}^2(X_1) + \dots + \sigma^2 W_{hn}^2(X_n)) \\
&= -4n^{-2} \sigma^2 (n^{-1} W_{h1}^2(X_1) + \dots + n^{-1} W_{hn}^2(X_n)) \\
&= -4n^{-2} \sigma^2 \sum_{j=1}^n n^{-1} W_{hj}^2(X_j) \\
&= -4n^{-2} \sigma^2 \sum_{j=1}^n n^{-1} (h^{-1} K(0) / \hat{f}_h(X_j))^2 \\
&= -4n^{-2} \sigma^2 (h^{-1} K(0) / \hat{f}_h(X_j))^2
\end{aligned}$$

$$= -4\sigma^2 \left(\frac{h^{-1}K(0)/\hat{f}_h(X_j)}{n} \right)^2 \quad (4.44)$$

Dengan mensubstitusikan persamaan (4.41), (4.42), (4.43b) dan (4.44) kedalam persamaan (4.40) akan didapat

$$\begin{aligned} E[G(h)] &= \sigma^2 + R(h) - 2n^{-2} \sum_{j=1}^n W_{hj}(X_j) \sigma^2 + 2n^{-2} \sum_{j=1}^n W_{hj}(X_j) \sigma^2 + \\ &\quad + 2n^{-1} h^{-1} \frac{K(0)}{\hat{f}_h(X_j)} R(h) - 4\sigma^2 \left(\frac{h^{-1}K(0)/\hat{f}_h(X_j)}{n} \right)^2 \\ &= \sigma^2 + R(h) + 2n^{-1} h^{-1} \frac{K(0)}{\hat{f}_h(X_j)} R(h) - 4\sigma^2 \left(\frac{h^{-1}K(0)/\hat{f}_h(X_j)}{n} \right)^2 \end{aligned}$$

Selanjutnya dengan mengabaikan *lower order term* (Rice, 1984), (Härdle, 1990) dan (Härdle, 1991) maka persamaan diatas menjadi :

$$\begin{aligned} E[G(h)] &= \sigma^2 + R(h) \\ &= p(h) \quad (4.45) \end{aligned}$$

Dari sini dapat disimpulkan bahwa dengan menambahkan *penalize Shibata* maka $G(h)$ merupakan estimator *unbiased* untuk $p(h)$.

4.3. Cross Validation (CV)

Hampir sama dengan metode *shibata*, ide dari *cross validation* adalah untuk menghilangkan suku bias pada *mean square error* (MSE).

$$CV(h) = n^{-1} \sum_{j=1}^n (Y_j - \hat{m}_{h,j}(X_j))^2, \quad (4.46)$$

dengan $\hat{m}_{h,j}(X_j)$ adalah *leave-one-out method* dimana observasi ke- j dihilangkan.

$$\hat{m}_{h,j}(X_j) = (n-1)^{-1} \sum_{i \neq j} W_h(X_j) Y_i$$

Selanjutnya untuk menunjukkan bahwa dengan *leave-one-out estimate* pada $\hat{m}_h(X_j)$ maka bias pada MSE akan sama dengan nol, dibawah ini diberikan penguraianya.

$$\begin{aligned} CV(h) &= n^{-1} \sum_{j=1}^n (Y_j - \hat{m}_{h,j}(X_j))^2 \\ &= n^{-1} \sum_{j=1}^n \left[Y_j^2 + \hat{m}_{h,j}^2(X_j) - 2Y_j \hat{m}_{h,j}(X_j) \right], \text{ karena } Y_j = m(X_j) + \varepsilon_j \text{ maka:} \\ &= n^{-1} \sum_{j=1}^n \left[(m(X_j) + \varepsilon_j)^2 + \hat{m}_{h,j}^2(X_j) - 2(m(X_j) + \varepsilon_j) \hat{m}_{h,j}(X_j) \right] \\ &= n^{-1} \sum_{j=1}^n \left[m(X_j)^2 + \varepsilon_j^2 + 2m(X_j)\varepsilon_j + \hat{m}_{h,j}^2(X_j) + \right. \\ &\quad \left. - 2(m(X_j)\hat{m}_{h,j}(X_j) + \varepsilon_j \hat{m}_{h,j}(X_j)) \right] \\ &= n^{-1} \sum_{j=1}^n \left[\varepsilon_j^2 + m(X_j)^2 + \hat{m}_{h,j}^2(X_j) - 2m(X_j)\hat{m}_{h,j}(X_j) + \right. \\ &\quad \left. + 2m(X_j)\varepsilon_j - 2\varepsilon_j \hat{m}_{h,j}(X_j) \right] \\ &= n^{-1} \sum_{j=1}^n \left[\varepsilon_j^2 + (m(X_j) - \hat{m}_{h,j}(X_j))^2 - 2\varepsilon_j(m(X_j) + \hat{m}_{h,j}(X_j)) \right] \end{aligned}$$

$$= n^{-1} \sum_{j=1}^n \left[\varepsilon_j^2 + n^{-1} \sum_{j=1}^n (m(X_j) - \hat{m}_{h,j}(X_j))^2 + \right. \\ \left. - 2n^{-1} \sum_{j=1}^n \varepsilon_j (m(X_j) + \hat{m}_{h,j}(X_j)) \right]$$

Dengan mengambil ekspektasi dari masing-masing persamaan didapat :

$$E[CV(h)] = n^{-1} \sum_{j=1}^n E(\varepsilon_j^2) + n^{-1} E \left(\sum_{j=1}^n (m(X_j) - \hat{m}_{h,j}(X_j))^2 \right) + \\ - 2n^{-1} E \left(\sum_{j=1}^n \varepsilon_j (m(X_j) + \hat{m}_{h,j}(X_j)) \right)$$

Denga memperhatikan bahwa error ε adalah iid dengan mean nol dan varians σ^2 maka diperoleh

$$E[CV(h)] = \sigma^2 + R_c(h) + C_{h2}, \quad (4.47)$$

dengan

$$C_{h2} = -2n^{-1} E \left(\sum_{j=1}^n \varepsilon_j (m(X_j) + \hat{m}_{h,j}(X_j)) \right) \quad (4.48)$$

selanjutnya karena

$$\hat{m}_{h,j}(X_j) = (n-1)^{-1} \sum_{i \neq j} W_{hi}(X_j) Y_i,$$

maka persamaan (4.48) menjadi:

$$C_{h2} = -2n^{-1} E \left(\sum_{j=1}^n \varepsilon_j \left[m(X_j) + (n-1)^{-1} \sum_{i \neq j} W_{hi}(X_j) Y_i \right] \right).$$

Karena $Y_i = m(X_i) + \varepsilon_i$ maka :

$$C_{h2} = -2n^{-1} E \left(\sum_{j=1}^n \varepsilon_j \left[m(X_j) + (n-1)^{-1} \sum_{i \neq j} W_{hi}(X_j) \{m(X_i) + \varepsilon_i\} \right] \right)$$

$$\begin{aligned}
C_{h2} &= -2n^{-1}E\left(\sum_{j=1}^n \varepsilon_j \left[m(X_j) + (n-1)^{-1} \sum_{i \neq j} W_{hi}(X_j)m(X_i) + (n-1)^{-1} \sum_{i \neq j} W_{hi}(X_i)\varepsilon_i \right]\right) \\
&= -2n^{-1} \sum_{j=1}^n E(\varepsilon_j)m(X_j) - 2n^{-1} \sum_{j=1}^n E(\varepsilon_j)(n-1)^{-1} \sum_{i \neq j} W_{hi}(X_j)m(X_i) + \\
&\quad - 2n^{-1}E\left(\sum_{j=1}^n \varepsilon_j(n-1)^{-1} \sum_{i \neq j} W_{hi}(X_j)\varepsilon_i\right) \\
&= -2n^{-1} \sum_{j=1}^n E(\varepsilon_j)m(X_j) - 2n^{-1}(n-1)^{-1} \sum_{j=1}^n E(\varepsilon_j) \sum_{i \neq j} W_{hi}(X_j)m(X_i) \\
&\quad - 2n^{-1}(n-1)^{-1} \sum_{j=1}^n \sum_{i \neq j} W_{hi}(X_j)E(\varepsilon_j\varepsilon_i)
\end{aligned}$$

Karena $E(\varepsilon_j) = 0$ dan $E(\varepsilon_j\varepsilon_i) = 0$, $j \neq i$ maka persamaan diatas akan sama dengan 0

Sehingga persamaan (4.47) menjadi :

$$\begin{aligned}
E[CV(h)] &= \sigma^2 + R_c(h) + C_{h2} \\
&= \sigma^2 + R_c(h) \\
&= p_c(h)
\end{aligned}$$

Dari sini dapat dibuat kesimpulan bahwa dengan *leave-one-out estimate* pada $\hat{m}_h(X_j)$ akan menyebabkan bias pada *MSE* sama dengan nol.

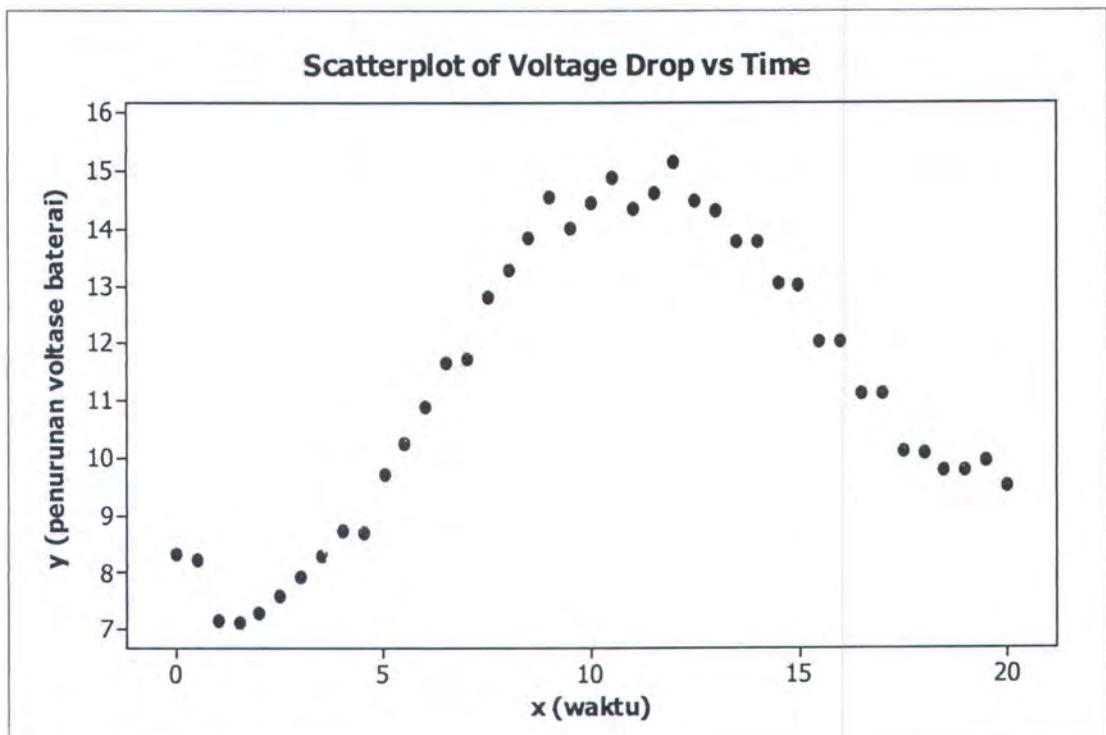
4.4. Aplikasi

Untuk menerapkan pemilihan bandwidth yang optimal pada regresi nonparametrik kernel, maka berikut ini akan diaplikasikan dalam data penurunan voltase baterai dari motor peluru kendali selama waktu penerbangan (Eubank, 1988).

Dalam pembahasan ini akan dilihat seberapa tepat bandwidth yang dikaji akan memberikan pengaruh terhadap pengestimasian model regresi nonparametrik. Sebagai variabel prediktor (x) adalah waktu (dalam satuan seconds) sedangkan variabel respon (y) adalah penurunan voltase baterai. Hubungan kedua variabel diberikan oleh model sebagai berikut:

$$Y_t = m(X_t) + \varepsilon_t, \quad t = 1, 2, 3, \dots, 41$$

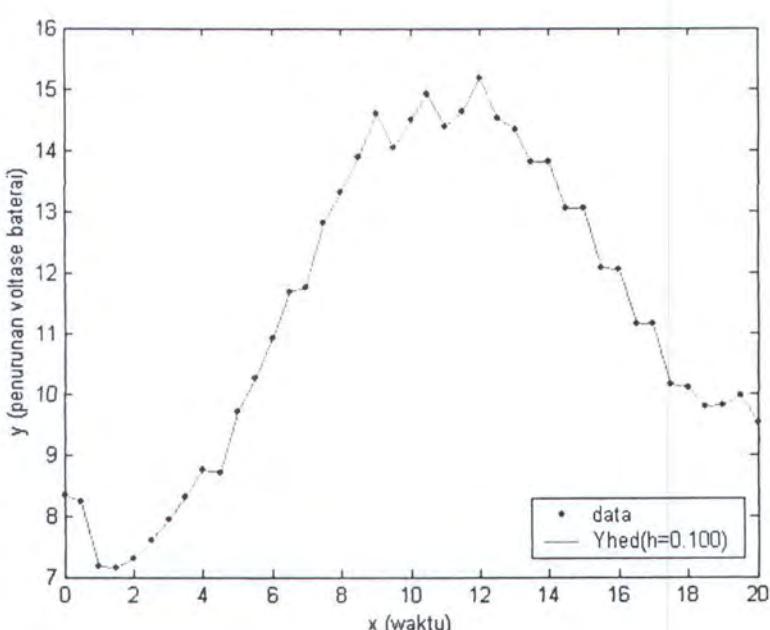
dengan error diasumsikan identik, independen dengan mean nol dan varians σ^2 . Plot antara x (waktu) dan y (penurunan voltase baterai) diberikan dalam Gambar 4.1.



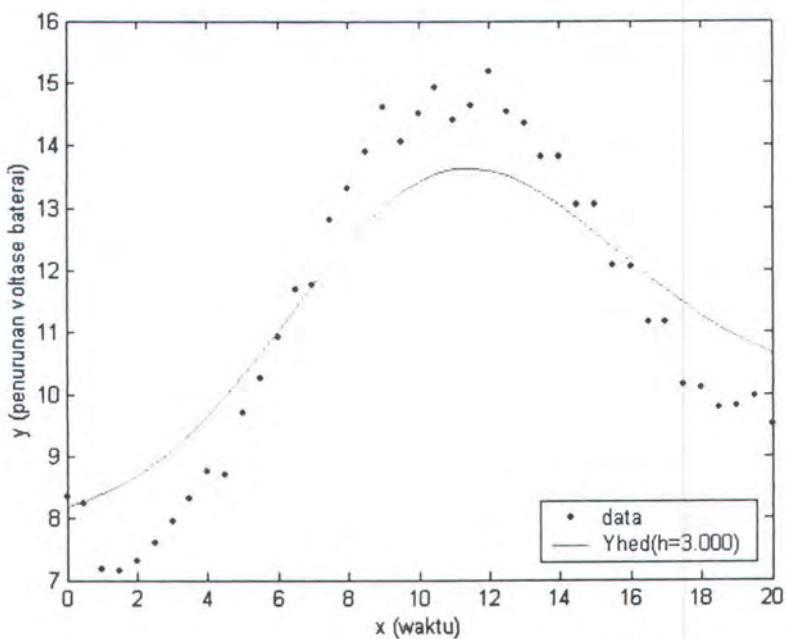
Gambar 4.1. Plot antara penurunan voltase baterai (y) dengan waktu (x)

Dari Gambar 4.1 terlihat bahwa plot data antara penurunan voltase baterai dengan waktu cenderung tidak berpola, dalam kata lain plot yang ada tidak memberikan informasi yang pasti mengenai bentuk dari kurva regresi. Ini terlihat pada interval waktu antara 0 detik pertama sampai 5 detik. Begitujuga pada interval waktu antara 8 detik hingga 14 detik dan interval waktu antara 16 detik sampai 20 detik tidak ada informasi atau asumsi yang kuat untuk mengestimasi kurva regresi.

Karena tidak ada informasi mengenai bentuk kurva regresi, yang diperkuat dengan hasil estimasi parametrik pada lampiran, maka akan dicoba dengan menggunakan pendekatan regresi nonparametrik kernel yang notabene sangat dipengaruhi oleh nilai bandwidth. Akan tetapi sebelum diterapkan pemilihan *bandwidth* yang optimal, terlebih dahulu akan diilustrasikan estimasi kurva regresi yang dihasilkan dari pemilihan *bandwidth* yang tidak optimal.



Gambar 4.2. Estimasi kurva regresi dengan $h = 0,100$



Gambar 4.3. Estimasi kurva regresi dengan $h = 3,000$

Untuk h yang terlalu kecil, misal diambil $h = 0,200$ (Gambar 4.2), maka estimasi kurva regresi akan cenderung menuju ke data, seluruh data dihubungkan oleh garis regresi sehingga diperoleh model yang kasar (*undersmooth*). Sebaliknya jika h yang kita pilih terlalu besar, misalkan dalam kasus ini kita pilih $h = 3.000$ (Gambar 4.3), maka akan mengakibatkan estimasi kurva sangat halus (*oversmooth*).

Dari ilustrasi diatas, maka tampak bahwa *bandwidth* optimal sangat penting di dalam estimasi regresi nonparametrik kernel. Selanjutnya dari plot yang ada (Gambar 4.1) akan diestimasi bentuk kurva regresinya dengan menggunakan pemilihan bandwidth yang sudah dikaji diawal pembahasan kedalam estimator nonparametrik kernel.

$$\hat{m}_h(x) = \frac{n^{-1} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x-X_i}{h}\right) Y_i}{n^{-1} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x-X_i}{h}\right)}$$

Pada pembahasan ini akan dikaji aplikasi dari *bandwidth* dengan menggunakan metode *shibata* dan *Rule of Thumb* serta *Cross Validation*. *Bandwidth* optimal yang sudah dipilih akan digunakan untuk mengestimasi kurva regresi. Hasil dari estimasi akan dilihat nilai MSE dan R^2 nya yang kemudian akan dibandingkan mana diantara ketiga metode yang memiliki MSE terkecil serta R^2 terbesar.

4.3.1. Bandwidth Shibata

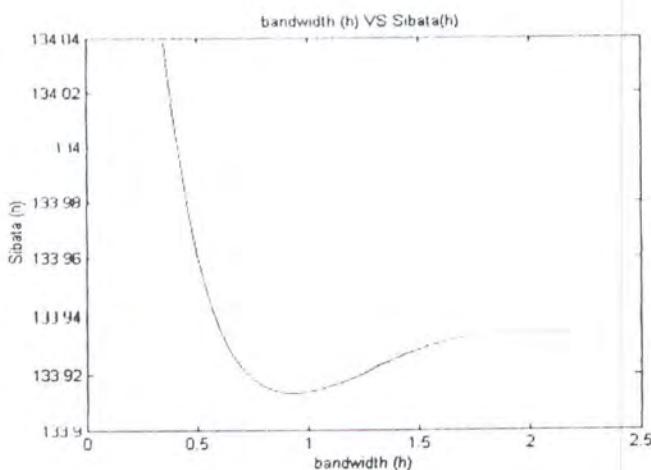
Sebagaimana sebagian metode pemilihan *bandwidth* yang ada, *bandwidth* yang optimal dengan menggunakan metode *Shibata* juga diperoleh dari nilai *Shibata* yang paling kecil. Selanjutnya setelah dilakukan perhitungan (program untuk memperoleh *bandwidth* optimal dengan metode *Shibata* diberikan dalam lampiran) diperoleh berbagai nilai *Shibata* dan *bandwidth* (h) sebagai berikut:

Tabel 4.1 Berbagai nilai *bandwidth* yang diperoleh dengan menggunakan metode *Shibata*

No	Bandwidth (h)	Nilai Sibata's	No	Bandwidth (h)	Nilai Sibata's
1	0,3500	134,0384	22	1,4000	133,9256
2	0,4000	134,0056	23	1,4500	133,9272
3	0,4500	133,9798	24	1,5000	133,9286
4	0,5000	133,9605	25	1,5500	133,9299
5	0,5500	133,9462	26	1,6000	133,9310
6	0,6000	133,9356	27	1,6500	133,9320
7	0,6500	133,9278	28	1,7000	133,9327
8	0,7000	133,9222	29	1,7500	133,9333
9	0,7500	133,9183	30	1,8000	133,9337
10	0,8000	133,9157	31	1,8500	133,9340
11	0,8500	133,9142	32	1,9000	133,9342
12	0,9000	133,9134	33	1,9500	133,9343
13	0,9500	133,9133	34	2,0000	133,9343
14	1,0000	133,9137	35	2,0500	133,9342
15	1,0500	133,9145	36	2,1000	133,9341
16	1,1000	133,9157	37	2,1500	133,9340

17	1,1500	133,9171	38	2,2000	133,9338
18	1,2000	133,9187	39	2,2500	133,9337
19	1,2500	133,9204	40	2,3000	133,9335
20	1,3000	133,9222	41	2,3500	133,9334
21	1,3500	133,9240			

Dari hasil perhitungan terlihat bahwa nilai terkecil dari shibata adalah 133,9133 dengan bandwidth sebesar 0,9500. Untuk lebih jelasnya dibawah ini diberikan plot antara nilai Shibata dengan bandwidth



Gambar 4.4. Plot antara *bandwidth* dengan Shibata

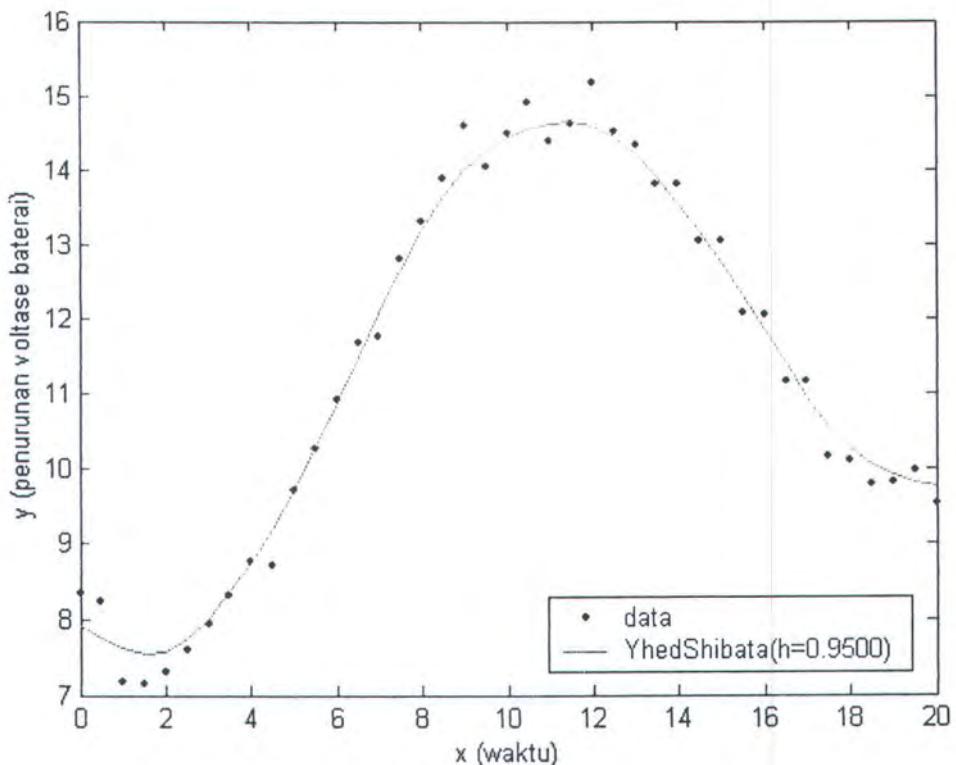
Dari plot diatas diperoleh nilai bandwidth optimal dengan menggunakan metode Shibata's adalah 0,9500, sehingga hubungan antara waktu dengan penurunan voltase baterai pada motor peluru kendali akan memenuhi persamaan :

$$\hat{m}_h(x) = \frac{41^{-1} \sum_{i=1}^{41} \frac{1}{0,9500} K\left(\frac{x - X_i}{0,9500}\right) Y_i}{41^{-1} \sum_{i=1}^{41} \frac{1}{0,9500} K\left(\frac{x - X_i}{0,9500}\right)}$$

Untuk mengestimasi kurva regresi, penulis menggunakan fungsi kernel Gaussian;

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} u^2\right)$$

Hasil plot estimasi nonparametrik kernel dengan menggunakan bandwidth shibata sebesar 0,9500 diberikan dalam Gambar 4.5 berikut :



Gambar 4.5. Plot estimasi nonparametrik kernel dengan bandwidth 0,9500

Dari Gambar 4.5 telihat bahwa dengan menggunakan *bandwidth* 0,9500 kurva regresi yang dihasilkan cenderung *smooth* (halus). Artinya kurva regresi tidak kasar dan tidak *oversmoothing*, sebagian besar data terwakili oleh estimasi kurva regresi. Ini diperkuat oleh nilai MSE sebesar 0,0727 dan R^2 0,9878.

4.3.2. Rule of Thumb Bandwidth

Dengan menggunakan metode *rule of thumb*, kita bisa langsung mendapatkan nilai *bandwidth* yang optimal tanpa mencari nilai yang paling minimal dari metode yang bersangkutan sebagaimana metode *shibata*.

Dengan memperhatikan plot data antara penurunan voltase baterai (y) dan waktu (x), maka kita coba dengan menggunakan pendekatan polinomial derajat tiga.

$$\hat{m}(x) = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 + \hat{\beta}_3 x^3$$

Persamaan tersebut kemudian diturunkan dua kali terhadap x untuk mendapatkan :

$$\hat{m}''(x) = 2\hat{\beta}_2 + 6\hat{\beta}_3 x$$

Selanjutnya dengan perhitungan sederhana didapat $(\hat{m}''(x))^2 = 0,1006$ dan $\hat{\sigma}^2 = 0,8977$ yang jika disubstitusikan kedalam \hat{h}_{ROT} ,

$$\hat{h}_{ROT} = n^{-\frac{1}{5}} \left(\frac{\frac{1}{2\sqrt{\pi}} \hat{\sigma}^2}{(\hat{m}''(x))^2} \right)^{1/5}$$

maka akan didapat bandwidth optimal sebesar 0,8170, sehingga estimasi kurva regresi atau pola hubungan antara waktu dengan penurunan voltase baterai akan memenuhi persamaan:

$$\hat{m}_h(x) = \frac{41^{-1} \sum_{i=1}^{41} \frac{1}{0,8170} K\left(\frac{x - X_i}{0,8170}\right) Y_i}{41^{-1} \sum_{i=1}^{41} \frac{1}{0,8170} K\left(\frac{x - X_i}{0,8170}\right)}$$

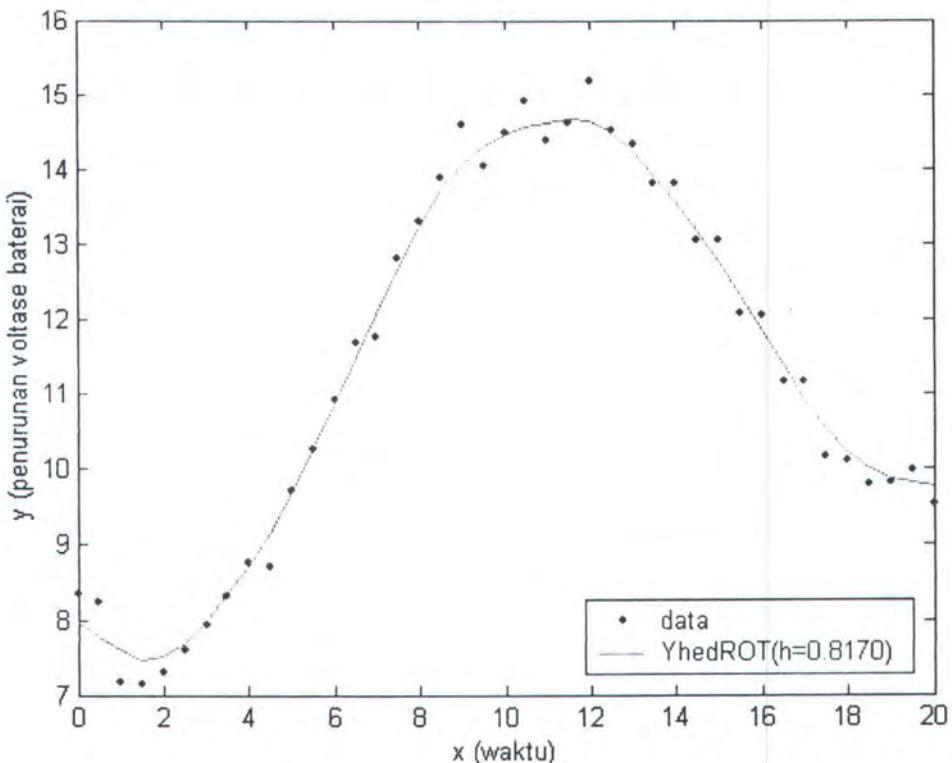
Dimana K adalah fungsi kernel.

Sebagaimana aplikasi metode Shibata pada (4.3.1), untuk mengestimasi kurva regresi digunakan fungsi kernel Gaussian yaitu :

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right)$$

Estimasi kurva regresi dengan menggunakan metode pemilihan *bandwidth*

Rule of Thumb diberikan sebagai berikut :



Gambar 4.6. Plot estimasi kurva regresi dengan bandwidth ROT sebesar 0,8170

Dari Gambar 4.6 terlihat bahwa estimasi kurva yang dihasilkan cukup halus. Selanjutnya dengan perhitungan yang sama diperoleh nilai MSE sebesar 0,0617 dan R^2 nya sebesar 0,9898

4.3.3 Metode Cross Validation (CV)

Sebagai bahan perbandingan, dibawah ini akan diberikan pengestimasian kurva regresi dengan menggunakan metode pemilihan bandwidth yang lain. Metode yang dimaksud adalah metode CV.

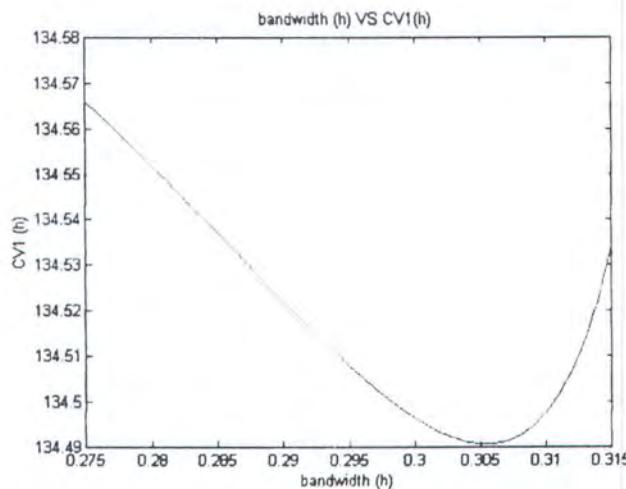
Sebagaimana yang kita ketahui bahwa pengembangan dari metode ini banyak digunakan dalam mendukung hasil penelitian yang ada yang terkait dengan masalah pengestimasian kurva regresi nonparametrik baik spline maupun kernel.

Prinsip yang mendasarinya sama dengan *Shibata* yaitu nilai terkecil dari CV menunjukkan *bandwidth* yang optimal. Selanjutnya setelah dilakukan perhitungan dibawah ini diberikan berbagai nilai bandwidth dan skor CV.

Tabel 4.2. Berbagai nilai CV dan *bandwidth* (*h*)

No	Bandwidth (h)	Nilai CV	No	Bandwidth (h)	Nilai CV
1	0,2750	134,5660	22	0,2960	134,5053
2	0,2760	134,5632	23	0,2970	134,5028
3	0,2770	134,5604	24	0,2980	134,5005
4	0,2780	134,5576	25	0,2990	134,4983
5	0,2790	134,5547	26	0,3000	134,4963
6	0,2800	134,5518	27	0,3010	134,4946
7	0,2810	134,5488	28	0,3020	134,4931
8	0,2820	134,5458	29	0,3030	134,4919
9	0,2830	134,5429	30	0,3040	134,4910
10	0,2840	134,5399	31	0,3050	134,4906
11	0,2850	134,5369	32	0,3060	134,4906
12	0,2860	134,5339	33	0,3070	134,4912
13	0,2870	134,5309	34	0,3080	134,4924
14	0,2880	134,5279	35	0,3090	134,4945
15	0,2890	134,5249	36	0,3100	134,4974
16	0,2900	134,5219	37	0,3110	134,5015
17	0,2910	134,5190	38	0,3120	134,5068
18	0,2920	134,5161	39	0,3130	134,5138
19	0,2930	134,5133	40	0,3140	134,5227
20	0,2940	134,5105	41	0,3150	134,5341
21	0,2950	134,5079	42		

Jika dilakukan plot antara nilai-nilai CV dengan *bandwidth* (*h*), maka akan tampak seperti gambar berikut:



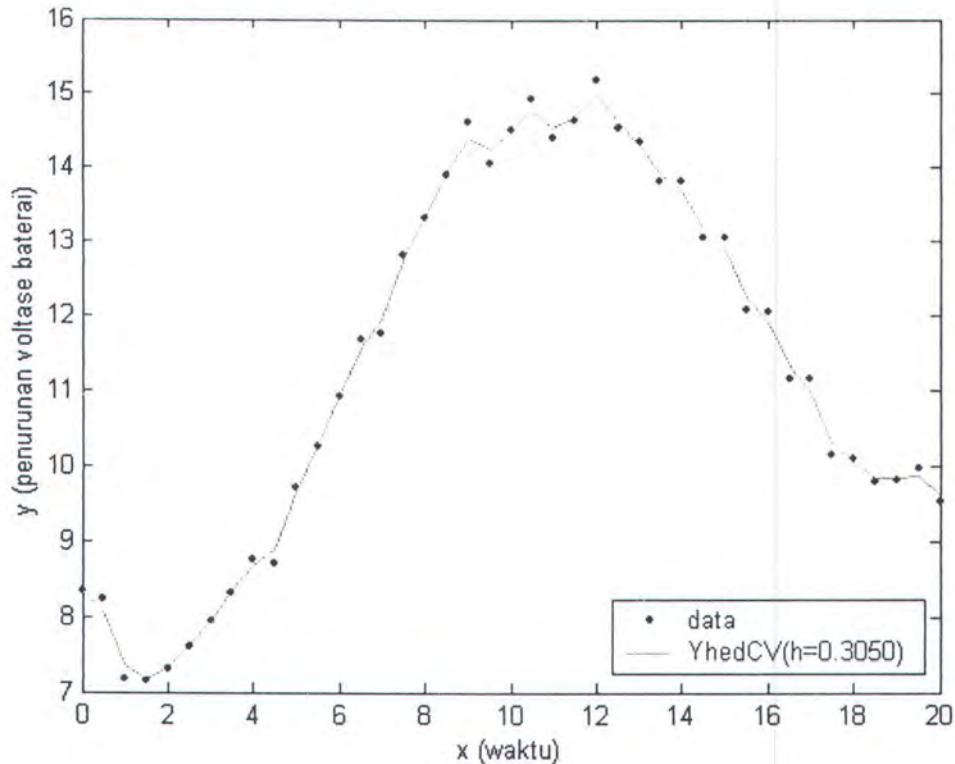
Gambar 4.7. Nilai CV dan bandwidth

Dari Tabel 4.2 dan hasil plot seperti pada Gambar 4.7 diatas, ternyata nilai CV yang paling minimum adalah 134,4906. Ini menunjukkan bahwa *bandwidth* yang optimal ada pada nilai 0,3050, sehingga hubungan antara penurunan voltase baterai motor peluru kendali selama waktu penerbangan akan memenuhi persamaan:

$$\hat{m}_h(x) = \frac{41^{-1} \sum_{i=1}^{41} \frac{1}{0,3050} K\left(\frac{x - X_i}{0,3050}\right) Y_i}{41^{-1} \sum_{i=1}^{41} \frac{1}{0,3050} K\left(\frac{x - X_i}{0,3050}\right)}$$

Selanjutnya untuk membandingkan dengan metode yang sudah dikaji (4.3.1 dan 4.3.2), maka berikut diberikan estimasi kurva regresi dan nilai MSE serta R^2 nya.

Dengan menggunakan fungsi kernel Gaussian dan bandwidth CV sebesar 0,3050, maka estimasi kurva regresinya seperti tampak pada gambar berikut:

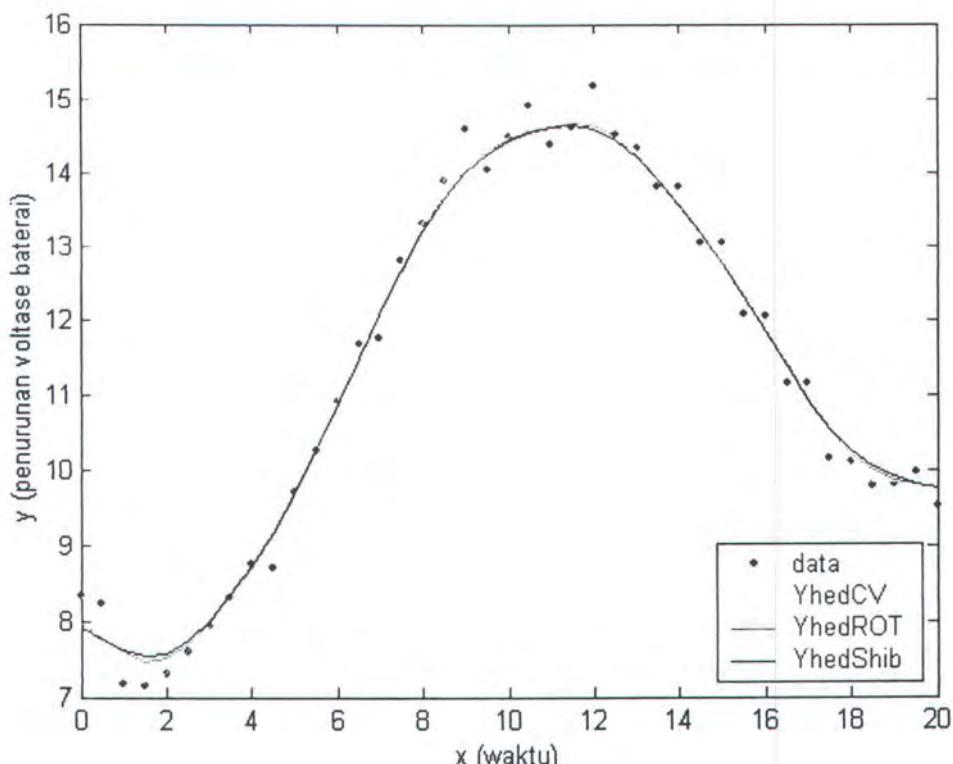


Gambar 8. Estimasi kurva regresi dengan bandwidth CV sebesar 0,3050

Nilai MSE dan R^2 dari hasil estimasi dengan menggunakan metode pemilihan *bandwidth Cross Validation* berturut-turut 0,0137 dan 0,9978

Lebih jelasnya, untuk melihat perbandingan antara metode *Rule of Thumb*, *Shibata*, dan metode *Cross Validation*, maka dibawah ini diberikan plot gabungan antara data yang disimbolkan dengan titik warna biru, taksiran kurva regresi dengan menggunakan bandwidth *Rule of Thumb* yang disimbolkan dengan garis berwarna

merah, dan taksiran kurva regresi dengan menggunakan metode *Shibata* yang disimbolkan dengan garis berwarna hitam serta taksiran kurva regresi dengan *bandwidth CV* dengan garis warna hijau.



Gambar 4.9. Plot gabungan taksiran kurva regresi dengan bandwidth ROT , Shibata dan CV

Dari Gambar 4.9 tampak bahwa estimasi kurva regresi dengan menggunakan metode *Rule of Thumb* dan *Shibata* hampir sama didalam mendekati data. Ini berbeda dengan metode *Cross Validation* yang cenderung sensitif didalam mendekati data.

Untuk melihat lebih lanjut perbedaan antara ketiga metode pemilihan *bandwidth*, dibawah ini diberikan nilai MSE, dan R^2 .

Tabel 4.3. Nilai SSE, MSE dan R^2 dari masing-masing metode

Bandwidth	<i>h</i>	MSE	R^2
CV Bandwidth	0,3050	0,0137	0,9978
Rule of Thumb (ROT)	0,8170	0,0617	0,9898
Shibata,s Bandwidth	0,9500	0,0727	0,9878

Dari Tabel 4.3 diatas terlihat bahwa nilai MSE untuk CV lebih kecil dibandingkan dengan ROT dan Shibata yaitu sebesar 0,0137 dengan R^2 sebesar 0,9978. Adapun perbandingan antara ROT dan Shibata sendiri menunjukkan bahwa MSE dari metode *Rule of Thumb* cenderung lebih kecil dibanding dengan metode *Shibata* yaitu 0,0617. Begitujuga dengan R^2 nya, terlihat bahwa koefisien determinasi model dengan menggunakan ROT cenderung agak besar dengan nilai 0,9898 dibandingkan dengan metode *shibata* yang memiliki nilai R^2 sebesar 0,9878 dan MSE sebesar 0,0727.

Secara umum, walaupun dalam Tabel 4.3 terdapat perbedaan nilai MSE dan R^2 dari masing-masing metode pemilihan *bandwidth*, akan tetapi perbedaan yang ada tidak terlalu signifikan karena hanya terpaut beberapa angka dibelakang koma.

BAB V

KESIMPULAN DAN SARAN

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Diberikan model regresi : $Y_i = m(X_i) + \varepsilon_i$, $\varepsilon_i \sim iid(0, \sigma^2)$ dengan $i = 1, 2, 3, \dots, n$. dimana $m(X_i)$ adalah kurva regresi yang diasumsikan tidak diketahui. Dari analisis dan pembahasan yang telah dilakukan, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Bandwidth optimal yang diperoleh dari meminimumkan Q masih mengandung fungsi m dan σ yang tidak diketahui. Metode *Rule of Thumb (ROT)* mendekati fungsi yang tidak diketahui tersebut dengan pendekatan parametrik sehingga diperoleh :

$$\hat{h}_{ROT} = n^{-\frac{1}{5}} \left(\frac{\frac{1}{2\sqrt{\pi}} \hat{\sigma}^2}{(\hat{m}''(x))^2} \right)^{1/5}$$

2. Dengan menggunakan penalize shibata, diperoleh sifat :

$$E[G(h)] = \sigma^2 + R(h) = p(h),$$

dengan

$$G(h) = n^{-1} \sum_{j=1}^n (Y_j - \hat{m}_h(X_j))^2 \cdot (1 + 2n^{-1}W_{hj}(X_j))$$

dan

$$R(h) = n^{-1} E \sum_{j=1}^n (m(X_j) - \hat{m}_h(X_j))^2$$

3. Dengan menggunakan *leave-one-out estimate* diperoleh :

$$\begin{aligned}
 ECV(h) &= n^{-1} \sum_{j=1}^n E(\varepsilon_j^2) + n^{-1} E \sum_{j=1}^n (m(X_j) - \hat{m}_{h,j}(X_j))^2 + \\
 &\quad - 2n^{-1} E \sum_{j=1}^n \varepsilon_j (m(X_j) + \hat{m}_{h,j}(X_j)) \\
 &= \sigma^2 + R_c(h) \\
 &= p_c(h)
 \end{aligned}$$

4. Untuk data penurunan *voltage* baterai dari motor peluru kendali selama waktu penerbangan, diperoleh *bandwidth* untuk masing-masing metode sebagai berikut :

- Metode Shibata diperoleh *bandwidth* sebesar 0,9500 sehingga persamaan kurva regresinya :

$$\hat{m}_h(x) = \frac{41^{-1} \sum_{i=1}^{41} \frac{1}{0,9500} K\left(\frac{x-X_i}{0,9500}\right) Y_i}{41^{-1} \sum_{i=1}^{41} \frac{1}{0,9500} K\left(\frac{x-X_i}{0,9500}\right)}$$

dan nilai MSE yang diperoleh adalah 0,0727 dengan R^2 sebesar 0,9878.

- Metode ROT diperoleh *bandwidth* sebesar 0,8170. Persamaan kurva regresinya :

$$\hat{m}_h(x) = \frac{41^{-1} \sum_{i=1}^{41} \frac{1}{0,8170} K\left(\frac{x-X_i}{0,8170}\right) Y_i}{41^{-1} \sum_{i=1}^{41} \frac{1}{0,8170} K\left(\frac{x-X_i}{0,8170}\right)}$$

dengan nilai MSE sebesar 0,0617 dan R^2 sebesar 0,9898.

- Metode CV diperoleh *bandwidth* sebesar 0,3050. Persamaan kurva regresinya :

$$\hat{m}_h(x) = \frac{n^{-1} \sum_{i=1}^n \frac{1}{0,3050} K\left(\frac{x-X_i}{0,3050}\right) Y_i}{n^{-1} \sum_{i=1}^n \frac{1}{0,3050} K\left(\frac{x-X_i}{0,3050}\right)}$$

dengan MSE dan R^2 masing-masing sebesar 0,0137, dan 0,9978.

Dari nilai MSE dan R^2 yang diperoleh, maka dapat disimpulkan untuk aplikasi data riil ini metode *Cross Validation* relatif lebih baik dari pada *Rule of Thumb* dan *Shibata*. Namun begitu, perbedaan nilai MSE dan R^2 dari masing-masing metode tidak terlalu signifikan. Masing-masing nilai hanya terpaut beberapa angka dibelakang koma.

5.2. Saran

Saran yang dapat diberikan untuk penelitian selanjutnya adalah :

1. Perlu kiranya dikaji lagi metode plug-in yang mendekati fungsi yang tidak diketahui pada *bandwidth optimal* (yang diperoleh dengan meminimumkan MSE atau MISE terhadap h) dengan pendekatan *pilot kernel*.
2. Perlu dilakukan simulasi untuk membandingkan metode pemilihan *bandwidth* yang optimal agar hasil yang diperoleh lebih mencerminkan kondisi sesungguhnya dari prilaku *bandwidth*.

DAFTAR PUSTAKA

DAFTAR PUSTAKA

- Draper, N.R., dan Smith, (1966), *Applied Regression Analysis*, Second Edition (alih Bahasa Ir. Bambang Sumantri, *Analisis Regresi Terapan*, PT Gramedia Pustaka Utama, Jakarta, 1992).
- Eubank, R.L. (1988). *Spline Smoothing and Nonparametric Regression*, Marcel Dekker, Inc. New York
- Härdle, W. (1990). *Applied Nonparametric Regression*, Cambridge University Press. New York
- _____. (1991). *Smoothing Techniques With Implementation in S*, Springer-Verlag. New York
- Jia, A. dan Schucany, W. R. (2004) Recursive Partitioning for Kernel Smoother: A tree-based Approach for Estimating Variable Bandwidths in Local Linear Regression. <http://www.smu.edu/Statistics/TechReports/TR316.pdf>
- Kweon, Y. J., dan Kockelman, K. M., (2004) Nonparametric regression estimation of household VMT, *Submitted for Presentation & Publication at the 2004 Annual Meeting of the Transportation Research Board*.
- Lee, K.S. dan Hill, CR.. Performance of Bandwidth selection rules for the local linear regression ,
- Myers, R. H., dan Milton, J. S., (1991), *A first course in the theory of linear statistical Models*, PWS – KENT Publishing Company, Boston
- Purcell, E. J., dan Varberg, D., (1995), Kalkulus dan Geometri Analitis, jilid 1, (alih Bahasa Drs. I. Nyoman Susila, M.Sc., Banu Kartasasmita, Ph.D dan Drs. Rawuh, Erlangga, Jakarta).
- Rice, J. (1984), Bandwidth choice for nonparametric regression, *The Annal of Statistics, Vol.12, No. 4, 1215 – 1230.*
- Rohatgi, V. K., (1976), *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons, USA.

Rosenblat, M. (1956), Remark on some Nonparametric Estimates of a Density Function, *Annals of Mathematical Statistics*, 27, 642-669.

Shibata, R (1980), Asymptotically efficient selection of the order of the model for estimating parameters of a linear process, *The Annals of Statistics*, Vol. 8, No. 1, 147 – 164.

_____ (1981), An optimal selection of regression variables, *Biometrika*, 68, 45 – 54.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, New York

Wand, M.P. dan Jones, M.C. (1995), *Kernel Smoothing*, Chapman and Hall, New York

Ye, A., Hyndman, R. J., dan Li, Z., (2006) Local linear multiple regression with variable bandwidth in the presence of heteroscedasticity,

. Februari 2006

LAMPIRAN

Lampiran A

Data penurunan Voltage Baterai dari Motor Peluru Kendali selama Waktu penerbangan

Nomor	Time	Voltage Drop
1	0.000	8.330
2	0.500	8.230
3	1.000	7.170
4	1.500	7.140
5	2.000	7.310
6	2.500	7.600
7	3.000	7.940
8	3.500	8.300
9	4.000	8.760
10	4.500	8.710
11	5.000	9.710
12	5.500	10.260
13	6.000	10.910
14	6.500	11.670
15	7.000	11.760
16	7.500	12.810
17	8.000	13.300
18	8.500	13.880
19	9.000	14.590
20	9.500	14.050
21	10.000	14.480
22	10.500	14.920
23	11.000	14.370
24	11.500	14.630
25	12.000	15.180
26	12.500	14.510
27	13.000	14.340
28	13.500	13.810
29	14.000	13.790
30	14.500	13.050
31	15.000	13.040
32	15.500	12.060
33	16.000	12.050
34	16.500	11.150
35	17.000	11.150
36	17.500	10.140
37	18.000	10.080
38	18.500	9.780
39	19.000	9.800
40	19.500	9.950
41	20.000	9.510

Lampiran B
Program Regresi Nonparametrik Kernel
SSE, MSE dan R²

```
h=input('h= ')
n=41;
fh=zeros(1,n);
for i=1:n
    for j=1:n
        d1(i,j)=(x(i)-x(j));
        da(i,j)=d1(i,j)/h;
        ker1(i,j)=exp((-da(i,j))^2/2)/sqrt(2*pi);
        fh(j)=fh(j)+(ker1(i,j))/n;
        rh(j)=rh(j)+((ker1(i,j))/n)*y(i);
    end
end
for j=1:n
    LC(j)=rh(j)/fh(j);
end
Yhed_LC=LC';
plot(x,y,'.',x,LC,'-');
%%R_Kuadrat dan MSE%%
residual= y-Yhed_LC;
ybar= sum(y)/n;
jke=(y-Yhed_LC)*(y-Yhed_LC)
mse=jke/n
jkr=(Yhed_LC-ybar)*(Yhed_LC-ybar);
jktot= jke+jkr;
Rsquare= jkr/jktot
```

Lampiran C
Program mencari Bandwidth Optimal dengan
ROT, Shibata, dan CV

Program mencari Bandwidth Optimal dengan ROT

```
n=41;
for i=1:n
    I(i)=1;
end
X = [I' x x.^2 x.^3];
S = X' * X;
bhet = inv (S) * X' * y;
yhet = X * bhet;
p =length(bhet);
mhet2 = 2*bhet(3) + 6 * bhet(4) *mean(x)
skuadrat=1/(n-p)*(y-yhet)'*(y-yhet)
hrotgaus= n^(-0.2) * (( 1/(2 * sqrt(pi)) * skuadrat )/mhet2^2)^0.2
```

Mencari Bandwidth Optimal dengan Shibata

```
z=length(x1)
for I=1:z
    x1kecil=ones(size(h(I)))*h(I);
    k1=(x1-x1kecil);
    K=k1./h(I);
    u=K.^2/-2;
    Kh=exp(u)./(sqrt(2*pi));
    kernel=diag(Kh);
    p=k1'*kernel*k1;
    Pin=inv(p);
    xw=k1'*kernel;
    a1=eye(size(k1));
    Hh=a1*(Pin.*xw);
    aa=eye(size(kernel));
    het=aa-Hh;
    msein=y'*het'*het*y;
    mse=msein/z;
    trc1=trace(Hh);
    tr1=trc1/z;
    Sibata(I)=mse*I+2*tr1;
end
Sib=Sibata';
s=min(Sib);
disp('=====');
disp('bandwidth (h) Sibata(h)');
disp('=====');
disp([h Sib]);
disp('=====');
figure(2);
plot(h,Sib);
title('bandwidth (h) VS Sibata(h)');
xlabel('bandwidth (h)');
ylabel('Sibata (h)');
for i=1:z
    if Sib(i)==s
        h2=h(i)
    end;
end;
```



Mencari Bandwidth Optimal dengan CV

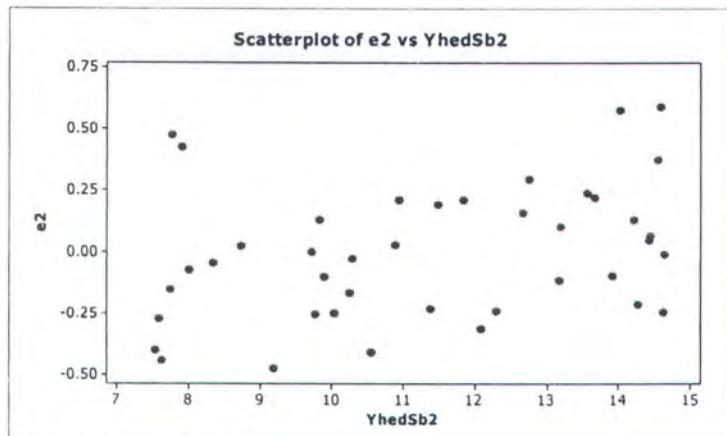
```
z=length(x)
for l=1:z
    xkecil=ones(size(h(l)))*h(l);
    k1=(x-xkecil);
    K=k1./h(l);
    u=K.^2/-2;
    Kh=exp(u)./(sqrt(2*pi));
    kernel=diag(Kh);
    p=k1'*kernel*k1;
    Pin=inv(p);
    xw=k1'*kernel;
    a1=eye(size(k1));
    Hh=a1*(Pin.*xw);
    aa=eye(size(kernel));
    het=aa-Hh;
    atas=y'*het;
    bawah=eye(size(x))'*(1-Hh);
    C=atas./bawah;
    CC=1/z*(C*C');
    CV1(l)=CC';
end
g=CV1';
disp('=====');
disp('bandwidth (h) CV1(h)');
disp('=====');
disp([h g]);
disp('=====');
k=min(g);
figure(1);
plot(h,g);
title('bandwidth (h) VS CV1(h)');
xlabel('bandwidth (h)');
ylabel('CV1 (h)');
for i=1:z
    if g(i)==k
        h1=h(i)
    end;
end;
```

Lampiran D

Pengujian Asumsi Residual

Bandwidth Sibata

1. Uji Identik



Regression Analysis: Abs_e2 versus Time

The regression equation is
Abs_e2 = 0.250 - 0.00302 Time

Predictor	Coef	SE Coef	T	P
Constant	0.25020	0.04872	5.14	0.000
Time	-0.003023	0.004193	-0.72	0.475

S = 0.158830 R-Sq = 1.3% R-Sq(adj) = 0.0%

Analysis of Variance

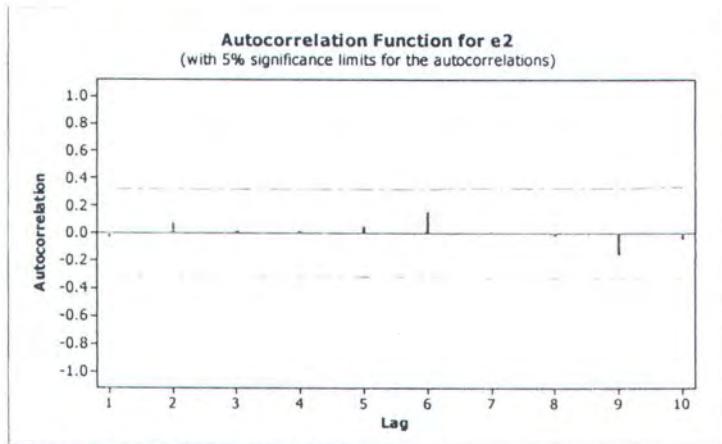
Source	DF	SS	MS	F	P
Regression	1	0.01311	0.01311	0.52	0.475
Residual Error	39	0.98385	0.02523		
Total	40	0.99696			

Unusual Observations

Obs	Time	Abs_e2	Fit	SE Fit	Residual	St Resid
19	9.0	0.5794	0.2230	0.0252	0.3564	2.27R
25	12.0	0.5949	0.2139	0.0262	0.3810	2.43R

R denotes an observation with a large standardized residual.

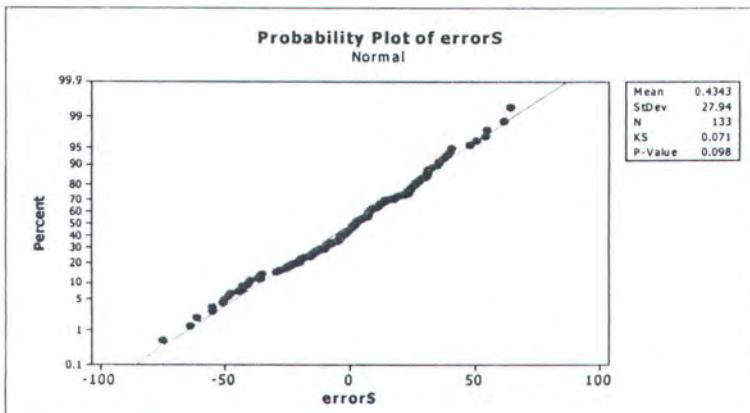
2. Uji Independen



Autocorrelation Function: e_2

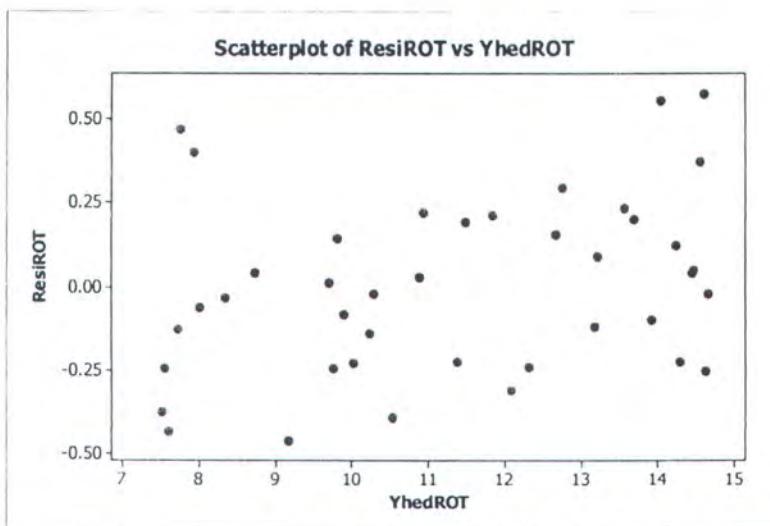
Lag	ACF	T	LBQ
1	-0.031681	-0.20	0.04
2	0.074824	0.48	0.30
3	0.013605	0.09	0.31
4	0.019277	0.12	0.32
5	0.050808	0.32	0.45
6	0.159520	1.01	1.73
7	0.008301	0.05	1.74
8	-0.017661	-0.11	1.75
9	-0.150876	-0.93	3.01
10	-0.039174	-0.24	3.09

3. Uji Normal



Rule of Thumb Bandwidth

1. Uji Identik



Regression Analysis: AbsResROT versus Time

The regression equation is
AbsResROT = 0.239 - 0.00259 Time

Predictor	Coef	SE Coef	T	P
Constant	0.23895	0.04726	5.06	0.000
Time	-0.002593	0.004067	-0.64	0.527

S = 0.154072 R-Sq = 1.0% R-Sq(adj) = 0.0%

Analysis of Variance

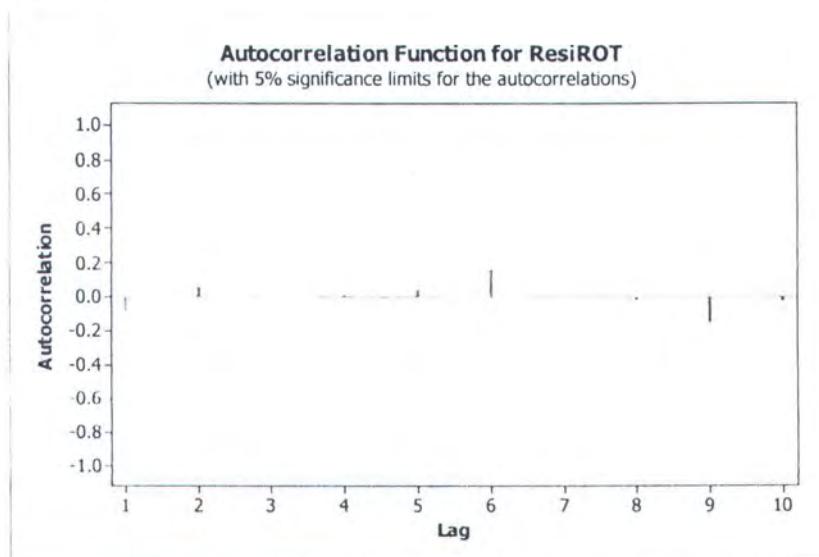
Source	DF	SS	MS	F	P
Regression	1	0.00965	0.00965	0.41	0.527
Residual Error	39	0.92579	0.02374		
Total	40	0.93544			

Unusual Observations

Obs	Time	AbsResROT	Fit	SE Fit	Residual	St. Resid
19	9.0	0.5579	0.2156	0.0244	0.3423	2.25R
25	12.0	0.5756	0.2078	0.0254	0.3678	2.42R

R denotes an observation with a large standardized residual.

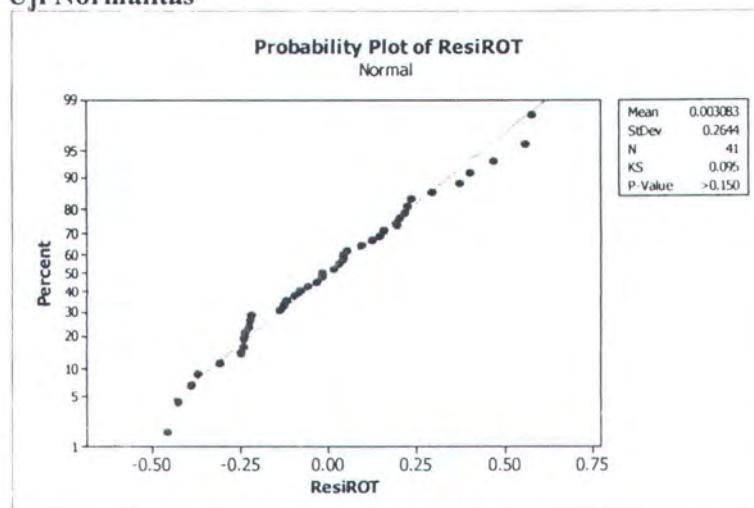
2. Uji Independent



Autocorrelation Function: ResiROT

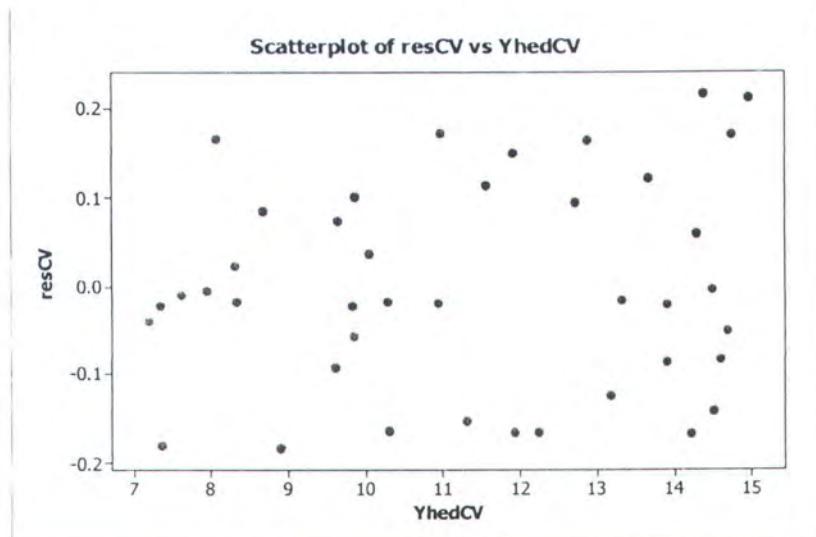
Lag	ACF	T	LBQ
1	-0.083946	-0.54	0.31
2	0.056329	0.36	0.45
3	0.002231	0.01	0.45
4	0.009302	0.06	0.46
5	0.039821	0.25	0.54
6	0.157538	1.00	1.79
7	0.003019	0.02	1.79
8	-0.013381	-0.08	1.80
9	-0.147206	-0.91	2.99
10	-0.028998	-0.18	3.04

3. Uji Normalitas



Metode Cross Validation

1. Uji Identik



Regression Analysis: AbsResCV versus Time

The regression equation is
AbsResCV = 0.0711 + 0.00259 Time

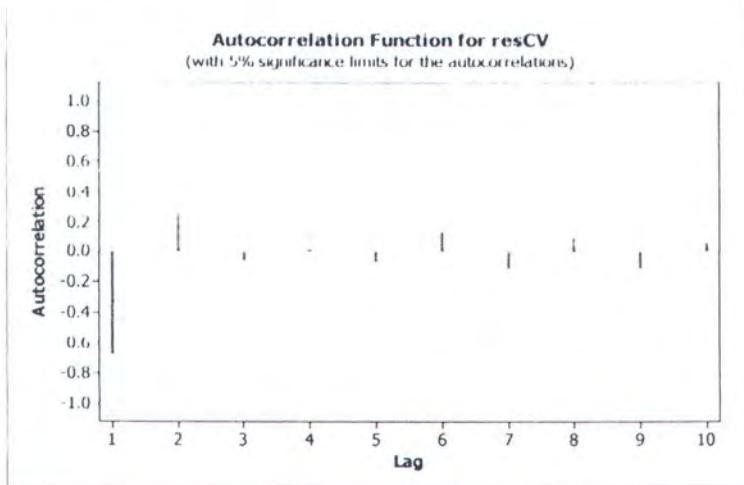
Predictor	Coef	SE Coef	T	P
Constant	0.07105	0.02019	3.52	0.001
Time	0.002586	0.001737	1.49	0.145

S = 0.0618385 R-sq(Adj) = 2.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.009593	0.009593	2.21	0.145
Residual Error	39	0.168951	0.004332		
Total	40	0.178544			

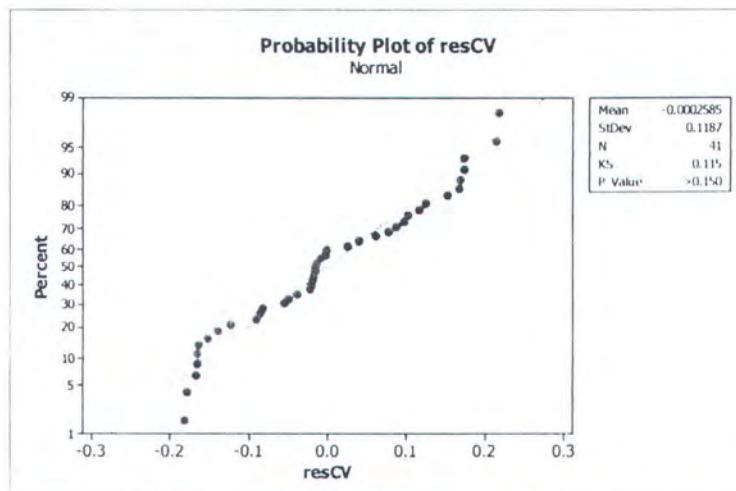
2. Uji Independen



Autocorrelation Function: resCV

Lag	ACF	T	LBQ
1	-0.674089	-4.32	20.03
2	0.258881	1.20	23.06
3	-0.057503	-0.26	23.21
4	0.015473	0.07	23.22
5	-0.064663	-0.29	23.43
6	0.137848	0.62	24.38
7	-0.108550	-0.48	24.99
8	0.095122	0.42	25.48
9	-0.108562	-0.48	26.13
10	0.059966	0.26	26.33

3. Uji Normalitas



Lampiran E

Pengujian Data dengan Kuadratik, dan Kubik

a. Kuadratik

Polynomial Regression Analysis: Voltage Drop versus Time

The regression equation is

$$\text{Voltage Drop} = 5.277 + 1.484 \text{ Time} - 0.06513 \text{ Time}^{**2}$$

$$S = 1.08054 \quad R-\text{Sq} = 83.0\% \quad R-\text{Sq}(\text{adj}) = 82.1\%$$

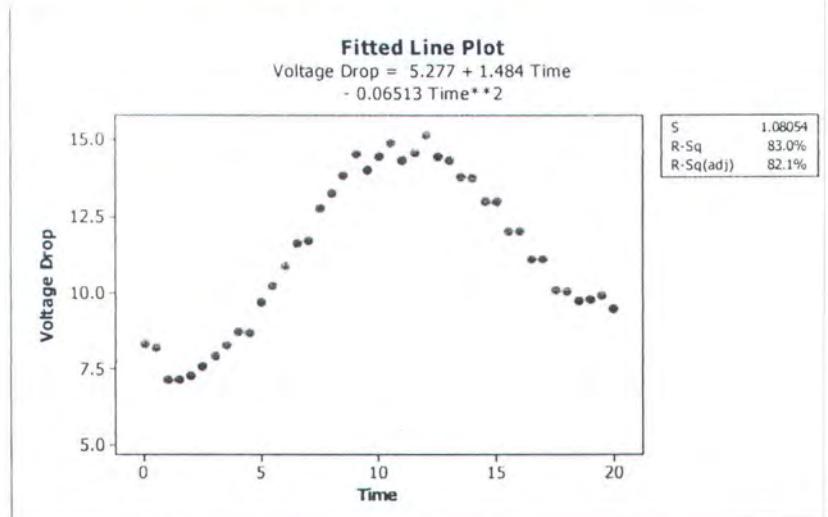
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	211.740	105.870	93.03	0.000
Error	38	44.368	1.168		
Total	40	256.108			

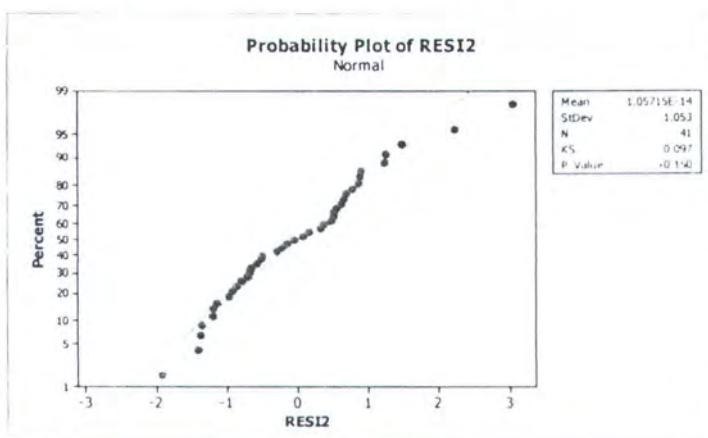
Sequential Analysis of Variance

Source	DF	SS	F	P
Linear	1	47.081	8.56	0.006
Quadratic	1	170.159	145.74	0.000

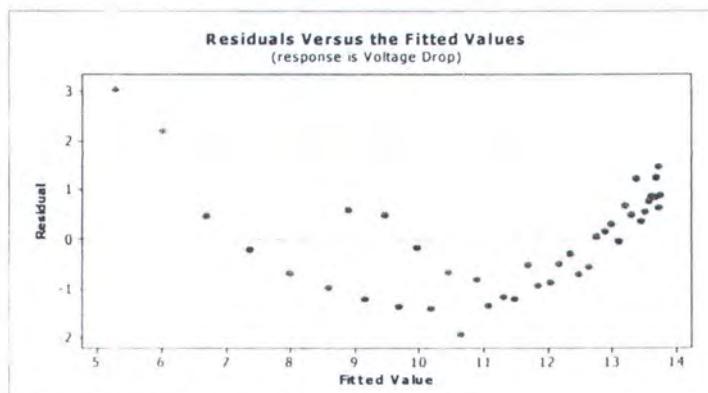
Fitted Line: Voltage Drop versus Time



Normplot of Residuals for Voltage Drop



Residuals vs Fits for Voltage Drop



b. Kubik

Polynomial Regression Analysis: Voltage Drop versus Time

The regression equation is
 $\text{Voltage Drop} = 6.469 + 0.7212 \text{ Time} + 0.03138 \text{ Time}^{**2} - 0.003217 \text{ Time}^{**3}$

S = 0.947446 R-Sq = 87.3% R-Sq(adj) = 86.3%

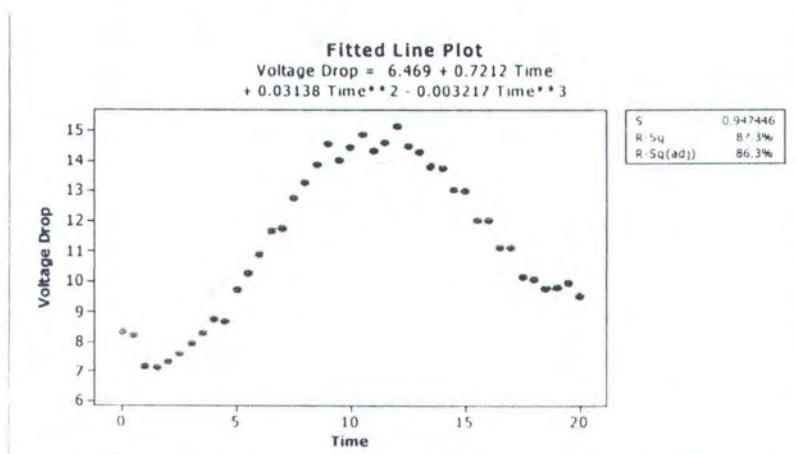
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	228.394	76.1315	84.81	0.000
Error	37	33.213	0.8977		
Total	40	261.608			

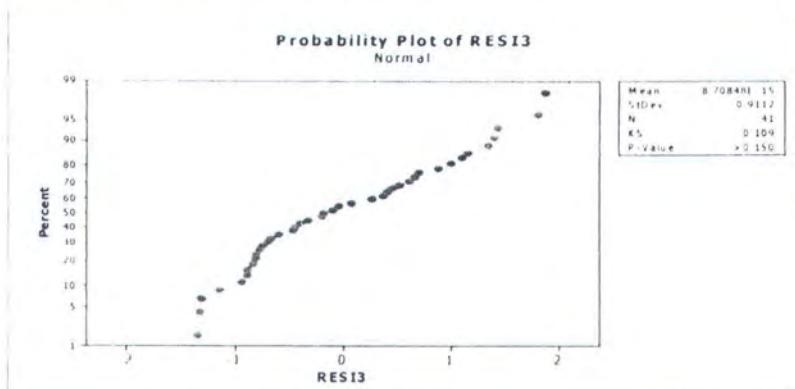
Sequential Analysis of Variance

Source	DF	SS	F	P
Linear	1	47.081	8.56	0.006
Quadratic	1	170.159	145.74	0.000
Cubic	1	11.155	12.43	0.001

Fitted Line: Voltage Drop versus Time



Normplot of Residuals for Voltage Drop



Residuals vs Fits for Voltage Drop

