

REGRESI SPLINE

SKRIPSI

**untuk memenuhi sebagian persyaratan
mencapai derajat Sarjana S-1**



Diajukan oleh

Wempy Eka Saputra

05610025

Kepada

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UIN SUNAN KALIJAGA
YOGYAKARTA**

2011

PERNYATAAN

Dengan ini saya menyatakan bahwa skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 01 Febuari 2011



Wempy Eka Saputra
Wempy Eka Saputra



PENGESAHAN SKRIPSI/TUGAS AKHIR

Nomor : UIN.02/D.ST/PP.01.1/370/2011

Skripsi/Tugas Akhir dengan judul : Regresi Spline/Spline Regression

Yang dipersiapkan dan disusun oleh :

Nama : Wempy Eka Saputra

NIM : 05610025

Telah dimunaqasyahkan pada : 16 Februari 2011

Nilai Munaqasyah : A / B

Dan dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga

TIM MUNAQASYAH :

Ketua Sidang

Sri Utami Zuliana, M.Sc
NIP. 19741003 200003 2 002

Penguji I

Sugiyarto, Ph.D

Penguji II

Ki Hariyadi, M.Ph

Yogyakarta, 24 Februari 2011

UIN Sunan Kalijaga

Fakultas Sains dan Teknologi

Dekan



Prof. Dr. H. Akh. Minhaji, M.A, Ph.D

NIP. 19580919 198603 1 002

KATA PENGANTAR

Puji syukur penulis panjatkan atas kehadiran Allah SWT. Sesungguhnya karena Dialah yang telah mengkaruniakan rahmatNya sehingga skripsi/ tugas akhir ini dapat terselesaikan. Tak lupa shalawat dan salam selalu tecurah atas Nabi kita Muhammad SAW yang selalu menuntun kita dengan sunah-sunah Beliau hingga akhir zaman.

Skripsi/ tugas akhir merupakan SKS yang wajib ditempu oleh mahasiswa dimana meupakan sebagian prasyarat dalam mencapai derajat sarjana S-1. Diharapkan skripsi/ tugas akhir dapat memberikan manfaat bagi mahasiswa pada umumnya dan bagi penulis pada khususnya. Dalam kesempatan ini penulis mengucapkan rasa terimakasih kepada seluruh pihak yang telah membantu dalam penyusunan skripsi maupun penyusunan laporannya, antara lain:

1. Bapak Prof. Dr. H. Akh. Minhaji, M.A., Ph.D selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
2. Ibu Meizer Said Nahdi, M.Si selaku mantan Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
3. Ibu Sri Utami Zuliana, M.Sc selaku ketua Prodi Matematika maupun selaku Pembimbing I.
4. Bapak M. Farhan Quadratullah, M.Si selaku Pembimbing II.
5. Bapak Sugiyarto, Ph.D selaku Penguji I.
6. Bapak Ki Hariyadi, M.Ph selaku penguji II.

7. Bapak Sugiyanto, S.T., M.Si selaku Penasehat Akademik.
8. Bapak Zaim, S.Si bagian Tata Usaha.
9. Ayah dan Ibu tercinta yang telah memberikan kasih sayang dan doa selalu.
10. Adikku Ricky dan Vanda yang selalu mendukungku dengan doa.
11. Adik-adik tingkatku Nyawang, Freza, Indra, Tomy, Heri dan lainnya.
12. Semua pihak yang tidak dapat disebutkan satu persatu.

Penulis menyadari bahwa dalam penyusunan skripsi/ tugas akhir ini masih terdapat kekurangan-kekurangan. Oleh karena itu segala kritik dan saran dari pembaca yang membangun sangat penulis harapkan.

Akhirnya penulis berharap semoga skripsi ini dapat memberikan manfaat bagi almamater dan semua pihak.

Yogyakarta, 1 Februari 2011

Penulis

Wempy Eka Saputra

HALAMAN PERSEMBAHAN

SKRIPSI INI AKU DEDIKASIKAN KEPADA

Allah SWT, terimakasih Tuhan karena semua berkat Engkau
Ayahku tercinta, terimakasih atas pengorbanan, kesabaran dan doamu untuk anakmu
Ibuku tercinta, meski engkau jauh tapi aku yakin ini berkat doamu juga
Almamaterku, UIN Sunan Kalijaga

DAFTAR ISI

Halaman Judul.....	i
Halaman Pengesahan.....	ii
Halaman Pernyataan.....	iii
Kata Pengantar.....	iv
Halaman Persembahan.....	vi
Daftar Isi.....	vii
Daftar Gambar.....	x
Daftar Tabel.....	xi
Daftar Lampiran.....	xii
Arti Lambang dan Singkatan.....	xiii
Intisari.....	xiv
BAB I PENDAHULUAN.....	1
A. Latar Belakang Masalah.....	1
B. Batasan Masalah.....	3
C. Tujuan Penelitian.....	4
D. Manfaat Penelitian.....	4
G. Keaslian Penelitian.....	5

BAB II	LANDASAN TEORI.....	6
2.1	Regresi Linear.....	6
2.1.1	Model Regresi Linear Sederhana.....	6
2.2.2	Asumsi Regresi Linear.....	7
2.2	Regresi Linear Parametrik.....	8
2.2.1	Estimasi Kuadrat Terkecil.....	9
2.2.2	Sifat Estimasi Kuadrat Terkecil.....	18
2.2.3	Regresi Nonparametrik.....	26
2.2.3.1	Kriteria Pemilihan Estimator.....	26
2.2.3.2	Estimasi $P(\lambda)$ dan $R(\lambda)$	30
2.2.3.3	Cross-Validation.....	33
BAB III	REGRESI SPLINE.....	37
3.1	Fungsi Spline.....	37
3.2	Ide Dasar Penghalusan regresi.....	41
3.3	Penghalusan Fungsi Spline.....	43
3.3	Regresi Spline.....	48
3.4	Pemilihan Parameter.....	51
3.4.1	Fungsi Resiko dan prediksi.....	53

3.4.2	Estimasi $P(\lambda)$ dan $R(\lambda)$	55
3.4.3	Generalized Cross-Validation.....	56
BAB IV	HASIL PENELITIAN	61
A.	Model Regresi Spline Linear.....	61
B.	Model Regresi Spline Kuadrat.....	63
C.	Model Regresi Spline Kubik.....	65
D.	Penghalusan Spline Kubik.....	68
E.	Hasil Output.....	69
BAB V	KESIMPULAN DAN SARAN	70
A.	Kesimpulan.....	70
B.	Saran.....	71
	DAFTAR PUSTAKA	72

DAFTAR GAMBAR

Gambar 1 Plot regresi spline linear dari pengambilan sebanyak 10 knot.....	60
Gambar 2 Hubungan kurva regresi spline linear dengan plot dari data.....	61
Gambar 3 Plot regresi spline kuadrat dari pengambilan sebanyak 10 knot.....	62
Gambar 4 Hubungan kurva regresi spline kuadrat dengan plot dari data.....	63
Gambar 5 Plot regresi spline kubik dari pengambilan sebanyak 10 knot.....	65
Gambar 6 Hubungan kurva regresi spline kubik denganplot dari data.....	65
Gambar 7 Penghalusan spline kubik dengan parameter penghalus $\lambda = 1,5$	67

DAFTAR TABEL

Tabel 1 Perolehan MSE dan GCV pada masing-masing regresi spline.....	68
Tabel 2 Data rata-rata berat badan bayi usia 0,5 – 58,5 bulan.....	87
Tabel 3 Prediksi rata-rata berat badan bayi dengan regresi spline linear.....	89
Tabel 4 Prediksi rata-rata berat badan bayi dengan regresi spline kuadrat.....	93
Tabel 5 Prediksi rata-rata berat badan bayi dengan regresi spline kubik.....	94
Tabel 6 Prediksi rata-rata berat badan bayi dengan penghalusan spline kubik $\lambda=1,5$	96

DAFTAR LAMPIRAN

Lampiran 1 Data rata-rata berat badan bayi usia 0,5 – 58,5 bulan.....	87
Lampiran 2 Prediksi rata-rata berat badan bayi dengan program MATLAB Ver 7.0.....	89
Lampiran 3 Listing program model regresi spline (linear, kuadrat dan kubik).....	99

ARTI LAMBANG DAN SINGKATAN

E	=	Ekspektasi
\sim	=	Berdistribusi
MSE	=	Mean Square Error
SSE	=	Sum Square Error
Cov	=	Covariance
Var	=	Variance
\perp	=	Orthogonal
Tr	=	Trace
$ \cdot $	=	Harga mutlak
$\ \cdot \ $	=	Norm Vektor
R	=	Risk/ Resiko
P	=	Prediction/ Prediksi
L	=	Loss/ Kerugian
CV	=	Cross-Validation
GCV	=	Generalized Cross-Validation
H	=	Matriks Hat

REGRESI SPLINE/SPLINE REGRESSION

Oleh :

Wempy Eka Saputra

05610025

Intisari

Jika sebanyak n buah data observasi $\{(X_i, Y_i)\}_{i=1}^n$ diambil dari sampel, maka dapat dimodelkan hubungan dalam regresi seperti berikut $Y_i = \mu(X_i) + \varepsilon_i, i=1,2,\dots,n$. Dalam mengestimasi kurva regresi dapat didekati dengan dua pendekatan, yaitu pendekatan parametrik dan pendekatan nonparametrik. Perbedaan antara keduanya adalah pendekatan parametrik estimasinya ditentukan dari percobaan sedangkan pendekatan nonparametrik hanya ditentukan dari data.

Teknik estimasi dalam regresi nonparametrik ada banyak, antara lain : estimator kernel, estimator spline, histogram, estimator deret orthogonal maupun estimator wavelet. Model nonparametrik dibangun dengan memilih ruang fungsi yang sesuai dimana fungsi regresi diyakini termasuk didalamnya, sedangkan model spline dibangun dari knot. Oleh karena itu penentuan jumlah dan posisi knot dalam regresi spline memegang peran yang sangat penting.

Fungsi spline kubik yang dilengkapi dengan parameter penghalus (*smoothing parameter*) sering disebut dengan penghalusan spline kubik. Penghalusan spline kubik diperoleh dengan meminimumkan penalized least square/ PLS. Pemilihan parameter penghalus dalam regresi ini menjadi penting tanpa mengabaikan bias dan variansi data.

Dalam memilih model spline terbaik dapat digunakan fungsi prediksi, uji CV, uji GCV, uji GML maupun uji UBR. Dari beberapa pilihan tersebut yang paling sering digunakan adalah fungsi prediksi dan uji GCV.

Kata Kunci : Nonparametric Regression, Spline, Smoothing Spline, GCV.

BAB I

PENDAHULUAN

A. Latar Belakang

Kajian data dalam statistik secara umum dibagi menjadi tiga yaitu : data *time series*, data *cross-section* dan data panel. Data *time series* lebih banyak merefleksikan perubahan subyek dalam kurun waktu tertentu. Data *cross-section* lebih banyak merefleksikan perbedaan antar subyek. Sedangkan data panel merupakan penggabungan data *time series* dan data *cross-section*. Penyelesaian yang sering digunakan pada data *cross-section* adalah analisis regresi. Meskipun analisis regresi sangat bervariasi ragamnya, namun keseluruhan analisis regresi dalam statistik bertujuan untuk prediksi atau peramalan data. Dalam analisis regresi akan diestimasi hubungan antara variabel x_i dan variabel y_i .

Jika diberikan sejumlah n buah pengamatan yaitu $\{(x_i, y_i), i=1, \dots, n\}$ dimana $x_i, y_i \in R$, maka hubungan antara x_i dan y_i adalah mengikuti model berikut:

$$y_i = f(t) + \varepsilon_i \quad (1.1.1)$$

dimana $f(t)$ merupakan fungsi kurva regresi dalam t .

y_i merupakan variabel tidak bebas dalam regresi.

ε_i merupakan residu acak yang berdistribusi normal.

permasalahan dalam analisis regresi adalah bagaimana menentukan estimasi kurva regresi yang diperoleh dari sampel. Berkaitan dengan estimasi tersebut ada dua metode pendekatan yang sering digunakan, yaitu pendekatan parametrik dan pendekatan nonparametrik¹. Dalam pendekatan parametrik, bentuk kurva regresi f harus diketahui dengan kata lain terdapat vektor $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{R}^n$ dengan $f(\beta, \cdot) = f(\cdot)$ sehingga berakibat estimasi f diperoleh dengan mengestimasi β . Jika fungsi regresi dinyatakan sebagai:

$$f(t) = \sum_{i=1}^n \beta_i x_i(t) \quad (1.1.2)$$

dimana β_i adalah vektor parameter dari populasi.

$x_i(t)$ adalah variabel bebas linear.

$f(t)$ adalah fungsi regresi.

Kurva estimasi f terhadap β merupakan kurva seleksi keluarga kurva-kurva yang memenuhi model regresi tersebut dan sesuai dengan datanya. Asumsi terhadap bentuk kurva regresi parametrik memerlukan pengalaman masa lalu atau terdapat sumber-sumber lain yang tersedia dalam penelitian sehingga dapat memberikan informasi yang detail tentang proses penyelidikan. Apabila asumsi untuk kurva regresi tidak diketahui atau informasi yang tersedia tentang kurva regresi tidak ada atau sangat kurang, maka untuk

¹ R.L Eubank, 1988, Spline Smoothing and Nonparametric Regression, Marcel Dekker Inc, New York, halaman 2.

mengestimasi kurva regresi hanya akan bergantung pada data sehingga dapat digunakan pendekatan nonparametrik. Resiko yang terjadi apabila bentuk kurva regresi tidak diketahui, namun tetap menggunakan pendekatan parametrik, maka inferensi yang dihasilkan tidak dapat dipertanggungjawabkan validitasnya. Dalam model regresi nonparametrik tidak memberikan asumsi terhadap bentuk kurva regresinya. Hal ini memberikan fleksibilitas yang lebih besar di dalam mengestimasi bentuk yang mungkin dari kurva regresi. Pemilihan kurva regresi tersebut biasanya dimotivasi oleh sifat kemulusan yang diasumsikan dimiliki oleh kurva regresi.

Ada beberapa teknik dalam mengestimasi kurva regresi dalam regresi nonparametrik, diantaranya estimator kernel, estimator spline, estimator histogram, estimator deret orthogonal maupun estimator wavelet. Penelitian skripsi ini hanya akan dibahas teknik estimasi menggunakan spline.

B. Batasan Masalah

Skripsi ini hanya akan membahas estimator spline dalam penyelesaian menggunakan pendekatan nonparametrik. Secara umum fungsi spline yang sering digunakan antara lain spline linear, spline kuadrat dan spline kubik. Untuk itu akan diasumsikan variabel bebas dan variabel tidak bebas yang kontinu dalam fungsi regresinya.

Sebagai pembanding akan diperkenalkan estimator spline menggunakan metode penghalusan spline kubik (*cubic smoothing spline*).

C. Tujuan Penulisan

Tujuan yang ingin dicapai dalam penulisan skripsi:

1. Mengkaji estimator spline dengan fungsi spline linear, spline kuadrat dan spline kubik serta mempelajari perkembangannya.
2. Mengkaji simulasi estimator spline dalam regresi spline linear, spline kuadrat dan spline kubik pada data yang tidak memenuhi asumsi linearitas.
3. Membandingkan estimator dalam fungsi spline linear, spline kuadrat dan spline kubik serta metode penghalusan spline kubik.
4. Menyelidiki uji generalized cross validation/ GCV dalam pemilihan model terbaik yang identik dengan pemilihan kriteria GCV minimum.

D. Manfaat Penelitian

Hasil penelitian ini diharapkan dapat bermanfaat dalam memprediksi observasi mendatang yang tidak memenuhi asumsi linearitas dengan regresi spline kubik maupun dengan penghalusan, tergantung metode mana yang menghasilkan nilai residu terkecil.

F. Keaslian Penelitian

Penulisan skripsi ini adalah merupakan studi literatur yang membahas estimator spline dalam regresi spline beserta perkembangannya. Regresi spline yang dijabarkan antara lain menggunakan fungsi spline linear, spline kuadrat dan spline kubik maupun dengan metode penghalusan spline kubik. Dalam penelitian ini juga akan menjabarkan pemilihan model regresi terbaik dengan uji GCV pada setiap model regresi spline linear, spline kuadrat, spline kubik maupun penghalusan dengan spline kubik. Setiap model regresi spline akan disimulasikan menggunakan data penelitian Purnomo (2004). Dalam penelitian skripsi sebelumnya oleh Purnomo (2004) yang berjudul *cubic smoothing spline* menjabarkan penghalusan dalam regresi spline kubik sebagai salah satu metode penyelesaian dalam regresi spline serta mensimulasikannya pada data rata-rata berat bayi usia 0.5 sampai 58.5 bulan yang diperoleh dari data primer.

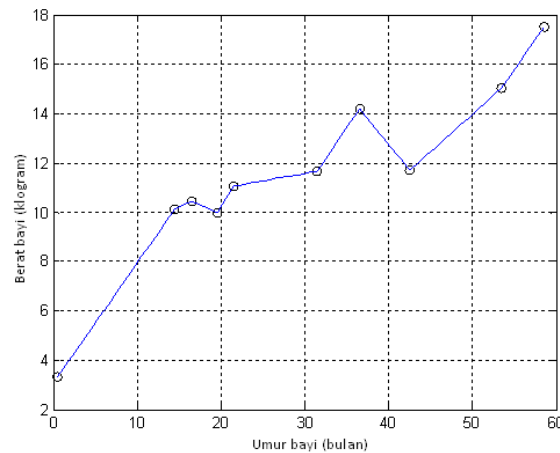
BAB IV

HASIL PENELITIAN

Regresi spline pada penelitian ini akan diterapkan pada data rata-rata berat badan bayi usia 0,5 bulan sampai 58,8 bulan. Data didapatkan dari penelitian Purnomo (2004) yang dicantumkan pada lampiran 1.

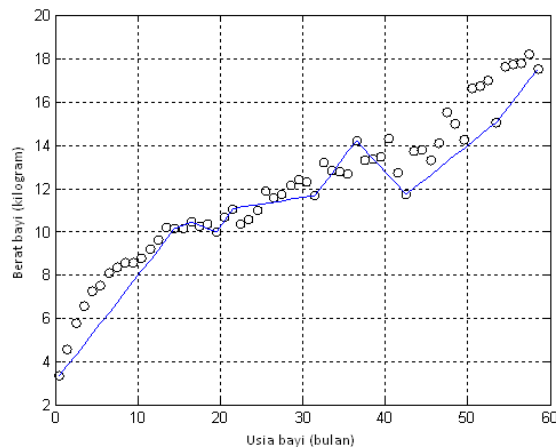
A. Model Spline Linear

Menggunakan bantuan program MATLAB VER 7.0, data tersebut diolah menggunakan metode spline linear. Jumlah knot yang dipilih adalah sebanyak 10 buah titik. Titik – titik tersebut diambil dari titik-titik yang mengalami perubahan kemiringan/ slope pada kurva datanya. Untuk memperoleh gambaran pemilihan posisi knot yang diambil terhadap penyelesaian yang diharapkan, dapat ditunjukkan pada gambar berikut:



Gambar 1. Plot regresi spline linear dari pengambilan sebanyak 10 knot.

penentuan jumlah knot secara teoritis tidak dibatasi. Apabila knot yang dipilih terlalu banyak tentu akan menyulitkan dalam menyusun persamaan regresinya, sehingga pemilihan knot dapat diambil seperlunya sesuai dengan petunjuk awal yang telah diuraikan sebelumnya. Untuk melihat hubungan penyelesaian regresi spline linear dengan plot dari data ditunjukkan pada gambar berikut:



Gambar 2. Hubungan kurva regresi spline linear dengan plot dari data.

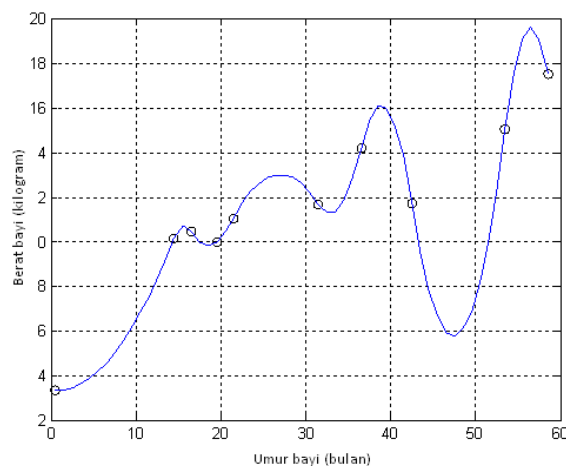
Dengan pemilihan jumlah knot sebanyak 10 buah, maka diperoleh 9 buah potongan polinomial (*piecewise polinomial*) dengan orde tertinggi tingkat satu dari masing-masing bagiannya. Berdasarkan output program yang terlampir pada lampiran diperoleh model regresi spline linear sebagai berikut:

$$S(x) = (0,4856x + 3,0841) + (0,164x - 7,751) + (-0,1523x + 12,9705) + (0,525x - 0,2375) + (0,0617x + 9,7235) + (0,5066x - 4,2909) + (-0,4167x + 29,4083) + (0,3045x - 1,2432) + (0,49x - 11,165).$$

penyelesaian akhir regresi spline linear, diperoleh bahwa nilai MSE sebesar 1.2294 dan nilai GCV sebesar 1,5170 (lampiran 2).

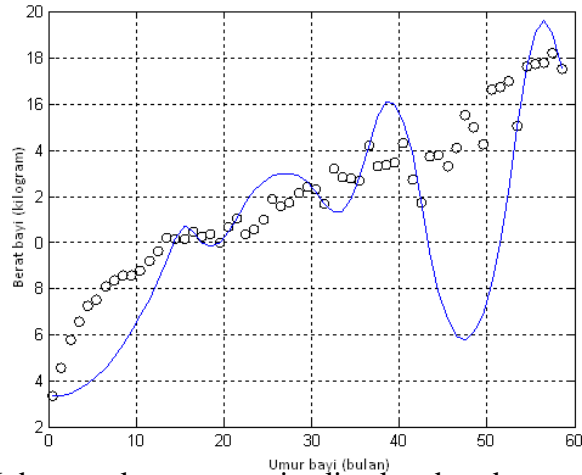
B. Model Spline Kuadrat

Dengan pemilihan jumlah knot sebanyak 10 buah dalam regresi spline kuadrat, maka diperoleh 9 fungsi potongan polinomial (*piecewise polinomial*) dengan orde tertinggi tingkat dua dari masing-masing bagiannya dan dapat ditemukan turunan tingkat satu yang kontinu pada masing-masing bagian polinomialnya. Pemilihan sebanyak 10 knot tersebut didasarkan pada petunjuk awal yang telah diuraikan sebelumnya, yaitu memilih titik knot yang dekat nilai maksimum lokal, minimum lokal atau titik dimana terjadi perubahan data. Untuk memperoleh gambaran pemilihan posisi knot yang diambil terhadap penyelesaian yang diharapkan, maka dapat ditunjukkan pada gambar berikut:



Gambar 3. Plot regresi spline kuadrat dari pengambilan sebanyak 10 knot.

Untuk melihat hubungan penyelesaian dalam regresi spline kuadrat dengan plot dari data ditunjukkan pada gambar berikut :



Gambar 4. Hubungan kurva regresi spline kuadrat dengan plot dari data.

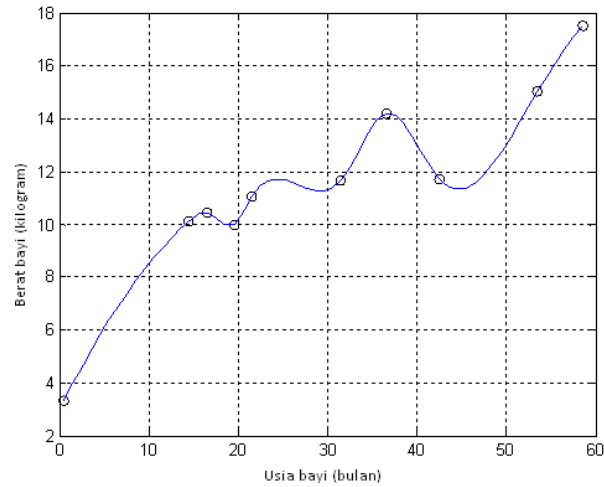
Berdasarkan output program yang terlampir pada lampiran dapat diperoleh model regresi spline kuadrat berikut:

$$\begin{aligned}
 S(x) &= 0,0347(x - 0,5)^2 + 3,327 - 0,4039(x - 14,5)^2 + 0,9717(x - 14,5) + 10,129 + \\
 &0,1638(x - 16,5)^2 - 0,6437(x - 16,5) + 10,457 + 0,093(x - 19,5)^2 + 0,339 \\
 &(x - 19,5) + 10 - 0,0649(x - 21,5)^2 + 0,711(x - 21,5) + 11,05 + 0,2188 \\
 &(x - 31,5)^2 - 0,5876(x - 31,5) + 11,667 - 0,3362(x - 36,5)^2 + 1,6008 \\
 &(x - 36,5) + 14,2 + 0,249(x - 42,5)^2 - 2,4341(x - 42,5) + 11,7 - 0,5106 \\
 &(x - 53,5)^2 + 3,0432(x - 53,5) + 15,05.
 \end{aligned}$$

penyelesaian akhir regresi spline kuadrat, diperoleh bahwa nilai MSE sebesar 11,2386 dan nilai GCV sebesar 13,8748 (lampiran 2).

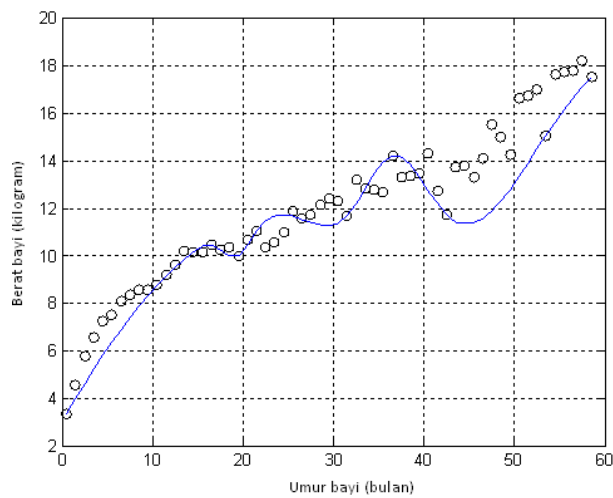
C. Model Spline Kubik

Dengan pemilihan jumlah knot sebanyak 10 buah dalam regresi spline kubik, maka diperoleh 9 buah potongan polinomial (*piecewise polinomial*) dengan orde tertinggi tingkat tiga dari masing-masing bagiannya dan dapat ditemukan turunan tingkat dua dan tingkat satu yang kontinu pada masing-masing bagian polinomialnya. Pemilihan sebanyak 10 knot tersebut diatas didasarkan pada petunjuk awal yang telah diuraikan sebelumnya, yaitu memilih titik knot dekat titik dimana terjadi perubahan data dengan tidak lebih dari satu titik ekstrem (maksimum atau minimum) dan satu titik pembelokan dimana terjadi diantara dua knot dengan sebisa mungkin titik ekstremnya harus menjadi pusat pada ruas. Untuk memperoleh gambaran pemilihan posisi knot yang diambil terhadap penyelesaian yang diharapkan, maka dapat ditunjukkan pada gambar berikut:



Gambar 5. Plot regresi spline kubik dari pengambilan sebanyak 10 knot.

Untuk melihat hubungan penyelesaian dalam regresi spline kubik dengan plot dari data ditunjukkan pada gambar berikut:



Gambar 6. Hubungan kurva regresi spline kubik dengan plot dari data.

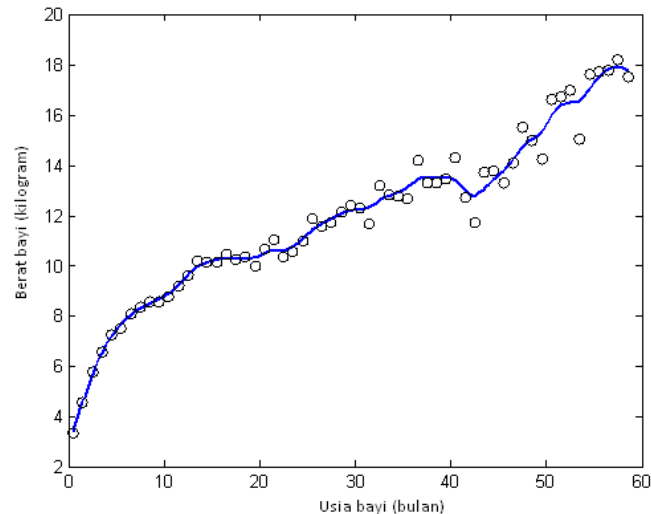
Berdasarkan output program yang terlampir pada lampiran dapat diperoleh model regresi spline kubik berikut:

$$\begin{aligned}
 S(x) &= -0,0026(x - 0,5)^3 + 0,0049(14,5 - x)^3 + 1,2388(x - 0,5) - 0,722(14,5 - x) \\
 &- 0,026(x - 14,5)^3 - 0,0185(16,5 - x)^3 + 5,3322(x - 14,5) + 5,1381(16,5 - x) \\
 &+ 0,030(x - 16,5)^3 - 0,0173(19,5 - x)^3 + 3,0575(x - 16,5) + 3,6412(19,5 - x) \\
 &- 0,0217(x - 19,5)^3 + 0,046(21,5 - x) + 5,6116(x - 19,5) + 4,8161(21,5 - x) \\
 &+ 0,0039(x - 21,5)^3 - 0,0043(31,5 - x)^3 + 0,7744(x - 21,5) + 1,5381(31,5 - x) \\
 &- 0,012(x - 31,5)^3 + 0,00785(36,5 - x)^3 + 3,1388(x - 31,5) + 2,1373(36,5 - x) \\
 &+ 0,00543(x - 36,5)^3 - 0,00996(42,5 - x)^3 + 1,7546(x - 36,5) + 2,7253 \\
 &(42,5 - x) - 0,0029(x - 42,5)^3 + 0,0029(53,5 - x)^3 + 1,3959(x - 42,5) \\
 &+ 0,7053(53,5 - x) - 0,00369(x - 53,5)^3 - 0,005(58,5 - x)^3 + 3,5923(x - 53,5) \\
 &+ 3,0226(58,5 - x).
 \end{aligned}$$

penyelesaian akhir regresi spline kubik diperoleh nilai MSE sebesar 1,7092 dan nilai GCV sebesar 1,7686 (lampiran 2).

D. Model Cubic Smoothing Spline

Metode penghalusan spline kubik (*cubic smoothing spline*) merupakan metode penyelesaian regresi spline kubik dengan mengalikan suatu parameter penghalus kurva regresi dalam model. Berawal dari penyelesaian menggunakan regresi spline kubik, kemudian memilih suatu parameter penghalus spline dengan pertimbangan tidak bias dan variansi yang minimum. Untuk melihat hubungan dalam penyelesaian penghalusan spline kubik dengan plot dari data ditunjukkan pada gambar berikut:



Gambar 7 Penghalusan spline kubik dengan parameter penghalus $\lambda = 1,5$.

Penyelesaian akhir penghalusan spline kubik diperoleh nilai MSE sebesar 0,699 dan nilai GCV sebesar 0,7233.

E. Hasil Output

Hasil dari nilai MSE dan nilai GCV dari masing-masing metode regresi dapat disusun dalam tabel berikut:

No	Metode Regresi	MSE	GCV
1	Regresi spline linear	1,2294	1,5178
2	Regresi spline kuadrat	11,2386	13,8748
3	Regresi spline kubik	1,7092	1,7686
4	Penghalusan spline kubik	0,6990	0,7233

Sumber : Pengolahan Data Sekunder (2004)

Tabel 1 Perolehan MSE dan GCV pada masing-masing regresi spline

Berdasarkan output pengolahan data dapat disimpulkan bahwa metode penghalusan spline kubik merupakan metode terbaik dalam pendekatan nonparametrik dibandingkan dengan metode lain yang ada karena penghalusan spline kubik dilengkapi dengan parameter penghalus yang dapat dikontrol. Dapat pula dikatakan bahwa nilai MSE dan GCV sama-sama dapat digunakan untuk memilih model terbaik yang meminimumkan *penalized least square*, dimana pada pembahasan sebelumnya telah diuraikan bahwa $MSE(\lambda)$ merupakan estimator tidak bias lain yang juga merupakan kriteria GCV untuk fungsi regresi dalam lemma 4.5.2.

BAB V

KESIMPULAN

A. Kesimpulan

Berdasarkan uraian pada bab-bab sebelumnya, maka dapat disimpulkan beberapa hal sebagai berikut:

1. Regresi spline merupakan teknik pencocokan potongan polinomial (*piecewise polynomial*) yang memberikan keleluasaan bagi kurva regresi untuk menentukan sendiri bentuk kurvanya. spline orde $r-1$ dengan n buah knot dengan $\xi_1 < \dots < \xi_n$ dan memiliki $r-2$ derivatif yang kontinu dinyatakan sebagai berikut:

$$S(x) = \sum_{i=0}^{r-1} \beta_{0i} x^i + \sum_{i=1}^n \delta_i (x - \xi)_+^{r-1}$$

dengan $\sum_{i=1}^n (x - \xi)_+^{r-1}$ merupakan fungsi subplus jumlahan dari $1, x, \dots, x^{r-1}, (x - \xi_1)_+^{r-1}, \dots, (x - \xi_n)_+^{r-1}$.

β_{0i} dan δ_i merupakan suatu parameter tidak diketahui.

2. Penghalusan spline kubik merupakan metode perluasan dari polinomial orde 3 yang dilengkapi dengan parameter penghalus yang dapat dikontrol didalamnya agar

diperoleh model terbaik yang sesuai dengan datanya. Penghalusan spline kubik dinyatakan sebagai berikut:

$$S(\xi) = n^{-1} \sum_{i=1}^n (y_i - f(\xi))^2 + \lambda \int_b^a f^m(\xi)^2 d\xi$$

3. Metode GCV merupakan metode yang sering digunakan dalam memilih model terbaik yang sesuai digunakan dalam pendekatan regresi nonparametrik yang dinyatakan sebagai berikut:

$$GCV(\lambda) = \frac{MSE(\lambda)}{[n^{-1}Tr(I - S)]^2}$$

4. Penyelesaian estimasi kurva regresi menggunakan metode penghalusan spline kubik merupakan metode terbaik karena dihasilkan nilai GCV minimum sebesar 0,7233 dibandingkan metode lainnya.
5. Penyelesaian estimasi kurva regresi menggunakan metode regresi spline kuadrat merupakan metode terburuk Karena dihasilkan nilai GCV maksimum sebesar 13,8746. Adapun faktor penyebab yang dapat diamati adalah terbentuknya redaman/overshoot pada kurva spline kuadratnya.

B. Saran

Dari kesimpulan diatas penulis dapat memberikan saran sebagai berikut:

1. Penentuan jumlah knot dalam regresi spline diusahakan tidak terlalu banyak sebab akan menyulitkan analisis maupun membentuk model regresinya.
2. Pemilihan posisi knot dalam regresi spline sebaiknya mengikuti petunjuk awal sebab dapat mempengaruhi dalam hasil model regresi spline nya.
3. Penggunaan metode regresi spline sebaiknya hanya digunakan pada data-data yang tidak memenuhi asumsi linearitas.
4. Untuk penelitian selanjutnya penulis menyarankan supaya regresi spline dapat dikembangkan lagi pada regresi spline adaptif maupun pada fungsi B-spline yang terbobot.

DAFTAR PUSTAKA

- Drapper, N.R & Smith H., 1998, "Applied Regression Analysis", John Wiley & Sons Inc, Canada.
- Eubank, R.L., 1988, "Spline Smoothing and Nonparametric Regression", Marcel Darker Inc, New York.
- Hardle, W., 1989, "Applied Nonparametric Regression", Cambridge University Press, Cambridge.
- Montgomery, Dauglas C & Elizabeth A Peack., 1982, "Introduction to Linear Regression Analysis", John Wiley & Sons Inc, New York.
- Rencher, Alvin C., 2001, "Linear Models in Statistics", John Wiley & Sons Inc, Canada.
- Sahid., 2005, "Pengantar Komputasi Numerik dengan Matlab", Andi Offset, Yogyakarta.
- Seber, G.A.F., 1977, "Linear Regression Analysis" John Wiley & Sons Inc, Canada.
- Sembiring, R.K., 1995, "Analisis Regresi", Penerbit ITB, Bandung.

LAMPIRAN

LAMPIRAN (1)

Data rata-rata berat badan bayi usia 0,5 bulan sampai 58,5 bulan yang diambil dari penelitian Purnomo, 2004:

No	Umur Bayi (bulan)	Berat Badan (Kg)
1	0,5	3,327
2	1,5	4,550
3	2,5	5,780
4	3,5	6,544
5	4,5	7,256
6	5,5	7,509
7	6,5	8,100
8	7,5	8,322
9	8,5	8,571
10	9,5	8,578
11	10,5	8,788
12	11,5	9,167
13	12,5	9,583
14	13,5	10,200
15	14,5	10,129
16	15,5	10,113
17	16,5	10,457
18	17,5	10,243
19	18,5	10,343
20	19,5	10,000
21	20,5	10,643
22	21,5	11,050
23	22,5	10,350
24	23,5	10,543
25	24,5	10,971
26	25,5	11,860
27	26,5	11,563
28	27,5	11,700
39	28,5	12,120
30	29,5	12,383
31	30,5	12,300

No	Umur Bayi (bulan)	Berat Bayi (Kg)
32	31,5	11,667
33	32,5	13,160
34	33,5	12,825
35	34,5	12,780
36	35,5	12,650
37	36,5	14,200
38	37,5	13,267
39	38,5	13,267
40	39,5	13,433
41	40,5	14,300
42	41,5	12,700
43	42,5	11,700
44	43,5	13,733
45	44,5	13,767
46	45,5	13,267
47	46,5	14,067
48	47,5	15,500
49	48,5	15,000
50	49,5	14,250
51	50,5	16,600
52	51,5	16,700
53	52,5	17,000
54	53,5	15,050
55	54,5	17,600
56	55,5	17,700
57	56,5	17,750
58	57,5	18,200
59	58,5	17,500

Tabel 2 Data rata-rata berat badan bayi usia 0,5 bulan sampai 58,5 bulan.

LAMPIRAN (2)

Prediksi rata-rata berat badan bayi dengan menggunakan program MATLAB Ver 7.0 adalah:

```
>>x = [0.5 14.5 16.5 19.5 21.5 31.5 36.5 42.5 53.5 58.5];
```

```
>>y = [3.327 10.129 10.457 10 11.05 11.667 14.2 11.7 15.05 17.5];
```

```
>>z = 0.5:1:58.5;
```

Sedangkan matriks Hat yang digunakan dalam uji GCV:

```
>>H = x*inv(x'*x)*x';
```

```
>>GCV = MSE/(1-Trace(H)/n)^2;
```

Hasil olah data dengan regresi spline linear:

```
>>[a,b] = spline(x,y)
```

```
>>S = interspline(x,y,z)'
```

Tabel prediksi rata-rata berat badan bayi dengan mnggunakan regresi spline linear:

No	Umur Bayi (bulan)	Berat Bayi (Kg)
1	0,5	3,327
2	1,5	3,813
3	2,5	4,299
4	3,5	4,785

No	Umur Bayi (bulan)	Berat Bayi (Kg)
5	4,5	5,270
6	5,5	5,756
7	6,5	6,242
8	7,5	6,728
9	8,5	7,214
10	9,5	7,699
11	10,5	8,186
12	11,5	8,671
13	12,5	9,157
14	13,5	9,643
15	14,5	10,129
16	15,5	10,293
17	16,5	10,457
18	17,5	10,305
19	18,5	10,152
20	19,5	10,000
21	20,5	10,525
22	21,5	11,050
23	22,5	11,112
24	23,5	11,173
25	24,5	11,235
26	25,5	11,297
27	26,5	11,359
28	27,5	11,420
39	28,5	11,482
30	29,5	11,454
31	30,5	11,605
32	31,5	11,667
33	32,5	12,174
34	33,5	12,680
35	34,5	13,187
36	35,5	13,693
37	36,5	14,200
38	37,5	13,783
39	38,5	13,367
40	39,5	12,950
41	40,5	12,533
42	41,5	12,117

No	Umur Bayi (bulan)	Berat Bayi (Kg)
43	42,5	11,700
44	43,5	12,005
45	44,5	12,310
46	45,5	12,614
47	46,5	12,918
48	47,5	13,223
49	48,5	13,527
50	49,5	13,832
51	50,5	14,136
52	51,5	14,440
53	52,5	14,746
54	53,5	15,050
55	54,5	15,540
56	55,5	16,030
57	56,5	16,520
58	57,5	17,010
59	58,5	17,500

Tabel 3 Prediksi rata-rata berat badan bayi dengan regresi spline linear.

Hasil olah data dengan regresi spline kuadrat:

```
>>m = spline2(x,f);
```

```
>>S = interspliner2(x,f,z)';
```

Tabel prrediksi rata-rata berat badan bayi dengan menggunakan regresi spline kuadrat:

No	Umur Bayi (bulan)	Berat Badan (Kg)
1	0,5	3,327
2	1,5	3,362
3	2,5	3,466
4	3,5	3,639

No	Umur Bayi (bulan)	Berat Bayi (Kg)
5	4,5	3,882
6	5,5	3,195
7	6,5	3,576
8	7,5	5,028
9	8,5	5,548
10	9,5	6,138
11	10,5	6,797
12	11,5	7,526
13	12,5	8,324
14	13,5	9,192
15	14,5	10,129
16	15,5	10,697
17	16,5	10,457
18	17,5	9,977
19	18,5	9,825
20	19,5	10,000
21	20,5	10,432
22	21,5	10,050
23	22,5	10,696
24	23,5	12,212
25	24,5	12,599
26	25,5	12,855
27	26,5	12,982
28	27,5	12,978
39	28,5	12,845
30	29,5	12,582
31	30,5	12,189
32	31,5	11,667
33	32,5	11,298
34	33,5	11,367
35	34,5	11,874
36	35,5	12,818
37	36,5	14,200
38	37,5	15,465
39	38,5	16,057
40	39,5	15,976
41	40,5	15,223
42	41,5	13,798

No	Umur Bayi (bulan)	Berat Bayi (Kg)
43	42,5	11,700
44	43,5	9,515
45	44,5	7,828
46	45,5	6,628
47	46,5	5,957
48	47,5	5,754
49	48,5	6,058
50	49,5	6,861
51	50,5	8,161
52	51,5	9,959
53	52,5	12,256
54	53,5	15,050
55	54,5	17,583
56	55,5	19,094
57	56,5	19,584
58	57,5	19,053
59	58,5	17,500

Tabel 4 Prediksi raa-rata berat badan bayi dengan regresi spline kuadrat.

Hasil olah data dengan regresi spline kubik:

```
>>S3 = spline(x,y,z,3,0,0);
```

Tabel prediksi rata-rata berat badan bayi dengan menggunakan regresi spline kubik:

No	Umur Bayi (bulan)	Berat Badan (Kg)
1	0,5	3,327
2	1,5	3,986
3	2,5	4,618
4	3,5	5,223
5	4,5	5,802
6	5,5	6,355

No	Umur Bayi (bulan)	Berat Bayi (Kg)
7	6,5	6,880
8	7,5	7,379
9	8,5	7,852
10	9,5	8,298
11	10,5	8,717
12	11,5	9,110
13	12,5	9,476
14	13,5	9,816
15	14,5	10,129
16	15,5	10,388
17	16,5	10,457
18	17,5	10,250
19	18,5	9,996
20	19,5	10,000
21	20,5	10,449
22	21,5	11,050
23	22,5	11,467
24	23,5	11,672
25	24,5	11,715
26	25,5	11,647
27	26,5	11,516
28	27,5	11,372
39	28,5	11,266
30	29,5	11,246
31	30,5	11,364
32	31,5	11,667
33	32,5	12,178
34	33,5	12,807
35	34,5	13,432
36	35,5	13,937
37	36,5	14,200
38	37,5	14,139
39	38,5	13,810
40	39,5	13,309
41	40,5	12,728
42	41,5	12,161
43	42,5	11,700
44	43,5	11,421

No	Umur Bayi (bulan)	Berat Bayi (Kg)
45	44,5	11,321
46	45,5	11,379
47	46,5	11,576
48	47,5	11,889
49	48,5	12,298
50	49,5	12,782
51	50,5	13,320
52	51,5	13,891
53	52,5	14,475
54	53,5	15,050
55	54,5	15,599
56	55,5	16,119
57	56,5	16,609
58	57,5	17,069
59	58,5	17,500

Tabel 5 Prediksi rata-rata berat badan bayi dengan regresi spline kubik.

Hasil olah data menggunakan metode penghalusan spline kubik tanpa knot:

```
>>x = csaps(x,y,0.5);
```

```
>>xx =
```

```
Form: 'pp'
```

```
Breaks: [1x59 double]
```

```
Coefs: [58x4 double]
```

```
Pieces: 58
```

```
Order: 4
```

Dim:4

>>ta = fnplt(xx)

Diperoleh tabel prediksi rata-rata berat bayi berikut:

No	Umur Bayi (bulan)	Berat Badan (Kg)
1	0,5	2,751
2	1,5	5,225
3	2,5	6,308
4	3,5	6,226
5	4,5	6,436
6	5,5	7,615
7	6,5	8,692
8	7,5	8,696
9	8,5	8,079
10	9,5	8,089
11	10,5	8,939
12	11,5	9,687
13	12,5	9,712
14	13,5	9,572
15	14,5	10,029
16	15,5	10,607
17	16,5	10,558
18	17,5	10,127
19	18,5	9,946
20	19,5	10,268
21	20,5	10,625
22	21,5	10,878
23	22,5	10,900
24	23,5	10,485
25	24,5	10,412
26	25,5	11,342
27	26,5	12,528
28	27,5	12,389
39	28,5	11,365
30	29,5	11,375

No	Umur Bayi (bulan)	Berat Bayi (Kg)
31	30,5	12,669
32	31,5	13,293
33	32,5	12,336
34	33,5	11,915
35	34,5	12,856
36	35,5	13,793
37	36,5	13,626
38	37,5	13,295
39	38,5	13,176
40	39,5	13,369
41	40,5	13,775
42	41,5	13,859
43	42,5	12,586
44	43,5	11417
45	44,5	13,009
46	45,5	15,454
47	46,5	15,269
48	47,5	13,525
49	48,5	13,845
50	49,5	15,877
51	50,5	16,796
52	51,5	16,553
53	52,5	16,229
54	53,5	16,327
55	54,5	16,369
No	Umur bayi (bulan)	Berat Bayi (Kg)
56	55,5	17,437
57	56,5	18,653
58	57,5	18,450
59	58,5	17,063

Tabel 6 Prediksi rata-rata betat badan bayi dengan penghalusan spline kubik.

LAMPIRAN (3)

Listing program model regresi spline (spline linear, spline kuadrat maupun spline kubik) sebagai berikut:

```
Function [a,b] = spliner(x,f);
%Menghitung koefisien pada model spline linear:
%Sk(x) = akx + b, k = 1,2,...,(n-1); xk <= x <= xk+1
n = length(x);
for k=1:(n-1),
    a(k) = (f(k+1)-f(k))/(x(k+1)-x(k));
    b(k) = f(k)-a(k)*x(k);
end
```

```
Function S = interspliner(x,y,z)
%Menghitung nilai spline linear
%S(z) = fk + mk(z - xk) dengan
%mk = (fk+1 - fk)/(xk+1 - xk);
n = length(x);
for j=1:length(z),
    for k=1:(n-1),
        if z(j)>=x(k)&z(j)<=x(k+1),
```

```

        m = (f(k+1) - f(k))/(x(k+1) - x(k));
        S(j) = f(k) + m*(z(j) - x(k));
    end
end
end

Function S = interspliner2(x,f,z)
%Menghitung nilai spline kuadrat alami (m_1 = 0)
%S(z) = (m_(k+1) - m_k)/(2(x_(k+1) - x_k))[(z - x_k)^2 + m_k(z - x_k) + f_k
%m_k = 2(f_k - f_(k-1))/(x_k - x(k-1)) - m_(k+1)
n = length(x);
m = spline2(x,f);
for j=1:length(z),
    for k=1:(n-1),
        if z(j)>=x(k)&z(j)<=x(k+1),
            S(j) = (m(k+1) - m_k)/(2*(x(k+1) - x(k)))*(z(j) - x(k))^2 + m(k)*(z(j) -
            x(k)) = f(k);
        end
    end
end
end
end

```

```

Function S3 = spline3(x,y,z,st,b1,bn)

n = length;

h = x(2:n) - x(1:n-1);

d = (y(2:n) - y(1:n-1))./h;

u = 2*(h(1:n-2)+h(2:n-1));

v = 6*d(2:n-1) - d(1:n-1));

V=v';

dia = h(2:n-2);

dib = dia;

if    st==1,

    u(1) = 3/2*h(1)+2*h(2);

    V(1) = V(1) - 3*(bn - d(n-1));

    u(n-2) = 2*h(n-2)+3/2*h(n-1);

    V(n-2) = V(n-2) - 3*(bn - d(n-1));

    A = diag(u) + diag(dib,-1) + diag(dia,1);

    m = A\V; m = [0;m;0];

    V

    m

end

if    st==2,

    A = diag(u) + diag(dib,-1) + diag(dia,1);

```

```

V
m = A\V; m = [0;m;0];
end
if st==3,
u(1) = 3*h(1) + 2*h(2) + h(1)^2/h(2);
dia(1) = h(2) - h(1)^2/h(2);
dib(n-3) = h(n-2) - h(n-1)^2/h(n-2);
u(n-2) = 2*h(n-2) + 3*h(n-1) + h(n-1)^2/h(n-2);
A = diag(u) + diag(dib,-1) + diag(dia,1);
m = A\V; m = [0;m;0];
m(1) = m(2) - h(1)*(m(3) - m(2))/h(2);
V
m
end
if st==4,
u(1) = 3*h(1) + 2*h(2);
u(n-2) = 2*h(n-2) + 3*h(n-1);
A = diag(u) + diag(dib,-1) + diag(dia,1);
m = A\V; m = [0;m;0];
m(1) = m(2); m(n) = m(n-1);
V

```



```

        m
end
if    st==5,
    V(1) = V(1) - h(1)*b(1);
    V(n-2) = V(n-2) - h(n-1)*bn;
    A = diag(u) + diag(dib,-1) + diag(dia,1);
    V
    m = A\V; m = [b1;m;bn];
end
C = y(2:n)./h - h.*m(2:n)'/6;
D = y(1:n-1)./h - h.*m(1:n-1)'/6;
For    j=1:length(z),
    For    k=1:n-1,
        If    z(j) >=x(k)&z(j)<=x(k+1),
            S3 = (m(k+1)*(z(j) - x(k)).^3 + m(k)*(x(k+1) - z(j)).^3)/(6*h(k)) +
                C(k)*(z(j) - x(k)) D(k)*(x(k+1) - z(j));
        end
    end
end
end
end

```

Wempy Eka Saputra

Blimbingsari CT IV no 55 RT 03/ RW 16 Sleman, Yogyakarta 55284

Jl. Dr Sutomo Gg. Seroja No. 17 RT 03/RW 08 Kelurahan Pandaan Pasuruan

East Java 47156

Email : Wempymat@yahoo.com

Phone : +62 878 390 22494



CURRICULUM VITAE

PERSONAL DATA	
Name	: Wempy Eka Saputra
Place/ Date of Birth	: Situbondo, October 20 th 1982
Sex	: Male
Marrial Status	: Single
Religion	: Moeslem
Citizenship	: Indonesia
Adress	: Jl. Dr Sutomo Gg. Seroja No. 17 RT 03/RW 08 Kelurahan Pandaan Pasuruan Jawa timur 47165
Present Adress	: Blimbingsari CT IV no 55 RT 03/ RW 16 Sleman, Yogyakarta 55284
Telephone No.	: +62 878 390 22494
E-mail	: Wempymat@yahoo.com
EDUCATION	
2005 – 2011	: Mathematics Department, Faculty of Sains & Technology Islamic State University of Sunan Kalijaga Yogyakarta. Final Assignment : Spline Regression.

2003 – 2008	:	D-3 Electrical Engineering, Faculty of Engineering Gadjah Mada University Yogyakarta
1998 – 2001		SMU Negeri 1 Pandaan Major in Science
HOBBIES / INTERESTS		
<ol style="list-style-type: none">1. Playing and listening music2. Playing computer game3. Internet browsing4. Travelling5. Swimming		