LAPORAN

PRAKTIKUM KOMPUTASI BIOMEDIS

Chapter 6 : Regression: Linear & Polynomials

Pelaksanaan Praktikum:

Hari: Selasa   Tanggal: 17 September 2019   Jam ke: 9-10

Oleh:

Nama : M. Thoriqul Aziz E

NIM : 081711733002

Dosen Pembimbing : Endah Purwanti, S. Si, M. T.

LABORATORIUM KOMPUTER

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS AIRLANGGA

SURABAYA

2019

A. TUJUAN

Mahasiswa dapat mendefinisikan korelasi variable dalam sebuah fungsi menggunakan metode regresi.

B. DASAR TEORI

Regresi adalah sebuah persmaan matematika sederhana yang digunkan untuk meramalkan nilai sutau peubah tak bebas pada data dari nilai-nilai satu atau lebih peubah bebas(Walpole,1995) . sedangkan secara umum yang digunakan adalah metode regresi linier yang mana bentukan persamaan yang digunkan adalah persmaan garis lurus f(x)= ax+b . Regresi yang lain yaitu regresi non linier, dimana persamaan matematika yan gdigunkan unutk meramal perubahan data memiliki persamaan polynomial, eksponensial, logaritmik atau yang lain. Penggunaan jenis regresi ditentukan dari tren data yang muncul hasil pembuatan grafik. Apabila tren data menunjukan kenaikaan konstan antara parameter nilai fungsi f(x) terhadap variable bebas nya yaitu x ,maka regresi yang digunakan adalah regresi linier. Begitu pula berlaku sebaliknya. Dengan persaman matematika regresi linier:

$$f(x) = ax + b$$

Maka akan diperoleh persamaan matriks untuk menemukan nilai koefisien a dan b :

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum x^2 & \sum x \\ \sum x & n \end{bmatrix}^{-1} \begin{bmatrix} \sum xy \\ \sum y \end{bmatrix}$$

Sehingga, dapat diperoleh persamaan regresi linier f(x) untuk kemudian meramal data data yang akan muncul dari nilai x pada nilai fungsi berikutnya.

C. TUGAS

1. Make a linear regression program based on data as shown below.

X	8	17	20	25	31	42	50	59	65	72	80
Y	100	130	209	276	330	359	420	487	550	645	700

2. Using hemoglobin as a blood substitute: hemoglobin-oxygen binding.

The explain of problem in reference King M.R and Mody N.A, 2010, “Numerical and Statistical Methods for Bioengineering”, Cambridge University Press, New York, page 102 . Have several equation:

$$S = \frac{[Hb(O_2)_n]}{[Hb(O_2)_n] + [Hb]} = \frac{(pO_2)^n}{P_{50}^n + (pO_2)^n} \quad (1)$$

From equation (1) is non linier but can be converted to linear form by first rearranging to :

$$\frac{S}{1-S} = \frac{(pO_2)^n}{P_{50}^n}$$

And then taking the logarithm of both sides to obtain:

$$\ln \frac{S}{1-S} = n \ln(pO_2) - n \ln P_{50}$$

Plotting $\ln \frac{S}{1-S}$ as a function of $n \ln(pO_2)$ produces a line with slope n and intercept with $-n \ln P_{50}$.

Equation (3) is the functional form for the dependency of oxygen saturation of hemoglobin as a function of the oxygen partial pressure, and is linear in the regression parameters

Table 1. Fractional saturation of hemoglobin as a function of partial pressure of oxygen in blood

pO_2 (mmHg)	10	20	30	40	50	60	70	80	90	100	110	120
S	0.18	0.4	0.65	0.8	0.87	0.92	0.94	0.95	0.95	0.96	0.96	0.97

Question:

a. Please derive equation (1) to equation (3)
b. Make a program of linear regression based on the data above by following the given information.(Remember: the equation is non-linear at first, so it is needed to be linearized first. Then, input the data in the linearized equation to get the unknown variables).
c. Find what is n and P_{50}
d. Plot the data before and after linearization.

3. Make a polynomial regression program to solve a nonlinear data problem.

D. PEMBAHASAN

1. Permasalahan yaitu terdapat beberapa data nilai x dan y yang kemudian dibaut sebuah regresi linier dari data tersebut. Penentuan jenis regresi dilihat dari tren data yang terbentuk. Pada kasus ini menggunakan regresi linier. Algoritma yang digunakan yaitu dengan penyusunan matriks matriks perhitungan, dengan nilai variable dari tiap sigma yang diperlukan. Nilai koefisien dari fungsi regresi diperoleh dari perkalian invers matriks dengan matriks hasil. Berikut adalah flowchart dari program:

   Berikut kode program dalam IDE Octave 5.10:

```

x=[ 8 17 20 25 31 42 50 59 65 72 80];
y=[100 130 209 276 330 359 420 487 550 645 700];
n=length(x);
%Inisialisasi sigma
sigmay=0;
sigmax=0;
sigmaxy=0;
sigmaxx=0;
for i=1:n
  sigmax=sigmax+x(i);
  sigmay=sigmay+y(i);
  sigmaxx=sigmaxx+x(i)*x(i);
  sigmaxy=sigmaxy+x(i)*y(i);
endfor
parameter = inv([sigmaxx sigmax;sigmax n])*[sigmaxy;sigmay];
a = parameter(1);
b = parameter(2);
fungsi = a*x + b;
plot(x,y,'*',x,fungsi)
```

   Pada syntax

```
x=[ 8 17 20 25 31 42 50 59 65 72 80];
y=[100 130 209 276 330 359 420 487 550 645 700];
n=length(x);
```

   menunjukan nilai matriks satu baris yang mana terdapat matriks x yang beranggotakan nilai nilai x dan begitu pula dengan matriks y. syntax `n=length(x);` menunjukan panjang matriks x dalam n dimana fungsinya nanti digunakan sebagai batas perulangan sigma.

   Pada syantax:

```
sigmay=0;
sigmax=0;
sigmaxy=0;
sigmaxx=0;
```

keempat variable tersebut menetukan nilai awalan dari sigma yang akan digunakan sebagai perhitungan koefisien regresi linier. Jumlah dari variable ditentukan dari jenis regresi yang digunakan.

Pada syntax:

```
for i=1:n
  sigmax=sigmax+x(i);
  sigmay=sigmay+y(i);
  sigmaxx=sigmaxx+x(i)*x(i);
  sigmaxy=sigmaxy+x(i)*y(i);
endfor
```

menunjukan perulangan yang dilakukan dari 1 hingga ke n pada tiap variable yang digunakan. Variable tersebut akan diupdate nilainya sedemikian sehingga nanti dapat diketahui nilai akhirnya.

Pada syntax:

```
parameter = inv([sigmaxx sigmax;sigmax n])*[sigmaxy;sigmay];
```

menunjukkan bahwa nilai variable sigma yang telah terupdate, kemudian disusun dalam sebuah matriks yang kemdudian langsung dihitung nilai parameter a dan b nya sesuai dengan rumus umum yang digunakan. Nilai perhitungan parameter menggunakan perkalian dengan matriks invers.

Pada syntax:

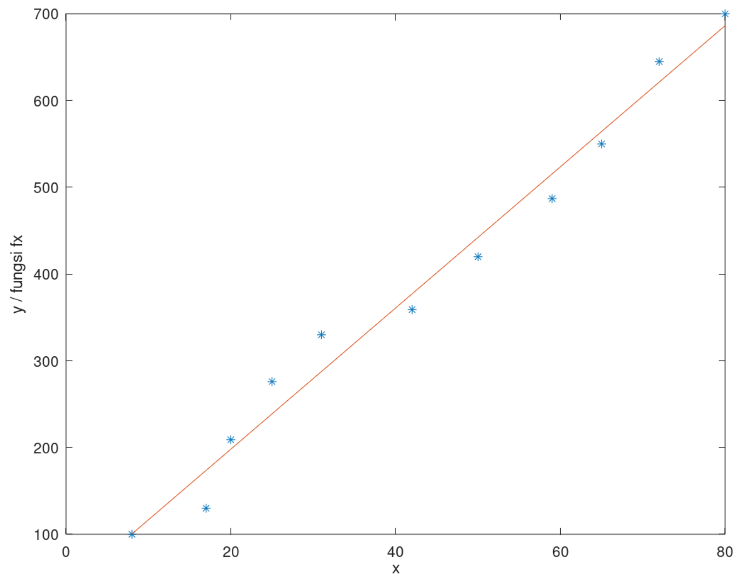
```
a = parameter(1);
b = parameter(2);
```

menunjukan pendefinisian variable a dan b dimana variable a bernilai dengan parameter anggota 1 dan variable b bernilai sesuai dengan parameter anggota 2. Hal ini sesuai dengan aturan perkalian matriksnya.

Pada syntax:

```
fungsi = a*x + b;
plot(x,y,'*',x,fungsi)
```

variable fungsi menunjukan hasil dari nilai x ketika dimasukan kedalam fungsi regresi linier, sedemikian sehingga nantinya dapat ditentukan titik titik untuk menentukan garis regersi. Syntax `plot(x,y,'*',x,fungsi)` berfungsi untuk mengeluarkan grafik dari nilai yang ada dengan perincian nilai data sebenarnya menggunakan "*" dan regresi menggunakan garis. Berikut adalah hasil *plotting* program :



2. Permasalahan dari kasus tersebut ialah persamaan yang diperoleh tidak membentuk tren persamaan linier sehingga penggunaan regresi linier tidak lagi relevan. Oleh karena itu diperlukan beberapa fungsi lain yang dengan fungsi tersebut dapat diolah data yang ada menajdi data yang memiliki tren linear.
   a. Dalam menemukan persamaan (1) menuju ke (3) maka diperoleh hasil matematik sebagai berikut :

$$S = \frac{(pO_2)^n}{P_{50}^n + (pO_2)^n} \; (1)$$

$$S.P_{50}^n + S.(pO_2)^n = (pO_2)^n$$

$$S.P_{50}^n = (pO_2)^n(1 - S)$$

$$\frac{S}{1 - S} = \frac{(pO_2)^n}{P_{50}^n} \; (2)$$

   Dari persmaan 2 menuju persamaan 3 cukup di logaritma natural kan sehingga diperoleh persmaaan:

$$\ln\frac{S}{1-S} = n\ln(pO_2) - n\ln P_{50} \quad (3)$$

b. Dari persmaan tersebut, kemudian buat program dari data hasil yang dimasukkan kedalam persamaan yang sudah dilinierkan.

Maka dari persamaan yang telah dimiliki, persaman 3 adalah persaman linier dalam fungsi fx dengan penjabaran:

$$f(x) = ax + b$$

Sedangkan persamaan fungsi (3) :

$$\ln\frac{S}{1-S} = n\ln(pO_2) - n\ln P_{50}$$

maka dapat dikorelasikan :

$$f(x) = \ln\frac{S}{1-S}$$

$$a = n; x = \ln(pO_2)$$

$$b = -n\ln P_{50}$$

Sehingga, dari rumusan tersebut dapat kemudian disusun algoritma dari data data yang telah ada. Beriut adalah flowchart dari kasus ini:

Berikut kode program dalam IDE Octave 5.10:

```
po=[10 20 30 40 50 60 70 80 90 100 110 120];
s=[0.18 0.4 0.65 0.8 0.87 0.92 0.94 0.95 0.95 0.96 0.96 0.97];
n=length(po)
%%%%% di lienierkan %%%%%%%%%%
for i=1:n
  y(i)=log(s(i)/(1-s(i)));
  x(i)=log(po(i));
endfor
y
x
%%%%%%%% inisialisasi sigma%%%%%%%%
sigmay=0;
sigmax=0;
sigmaxy=0;
sigmaxx=0;
for i=1:n
  sigmax=sigmax+x(i);
  sigmay=sigmay+y(i);
  sigmaxx=sigmaxx+x(i)*x(i);
  sigmaxy=sigmaxy+x(i)*y(i);
endfor
parameter = inv([sigmaxx sigmax;sigmax n])*[sigmaxy;sigmay];
a = parameter(1);
b = parameter(2);
fungsi = a*x + b;
```

Pada syntax:

```
po=[10 20 30 40 50 60 70 80 90 100 110 120];
s=[0.18 0.4 0.65 0.8 0.87 0.92 0.94 0.95 0.95 0.96 0.96 0.97];
n=length(po)
```

menujukan nilai matriks po dan s yang belum dilikukan linieraisasi. Nilai 2 matriks tersebut didapatkan dari tabel yang terdpat pada soal. Kemudian syntax `n=length(po);` menunjukan panjang dari matriks po, yang nantinya akan digunaka dalam perhituangan nilai nilainya.

Pada syntax:

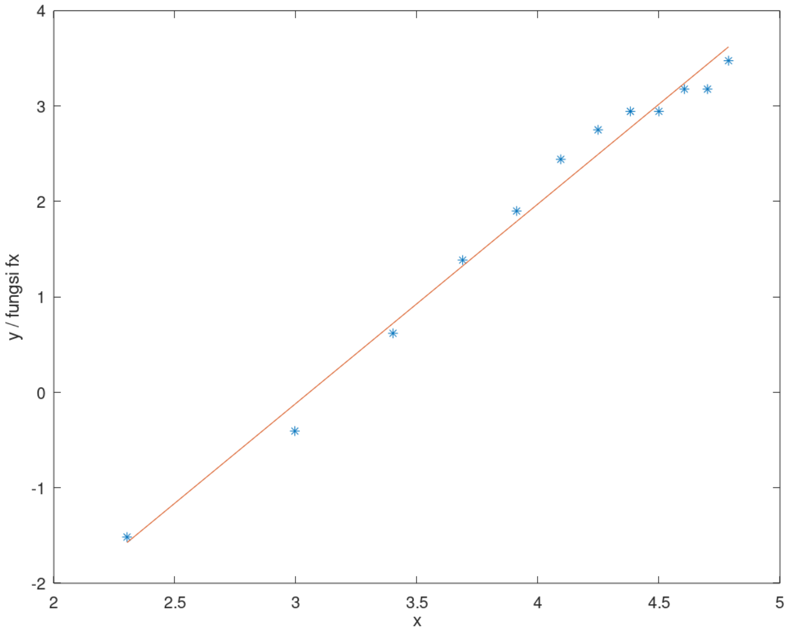
```
for i=1:n
  y(i)=log(s(i)/(1-s(i)));
  x(i)=log(po(i));
endfor
```

`y`

`x`

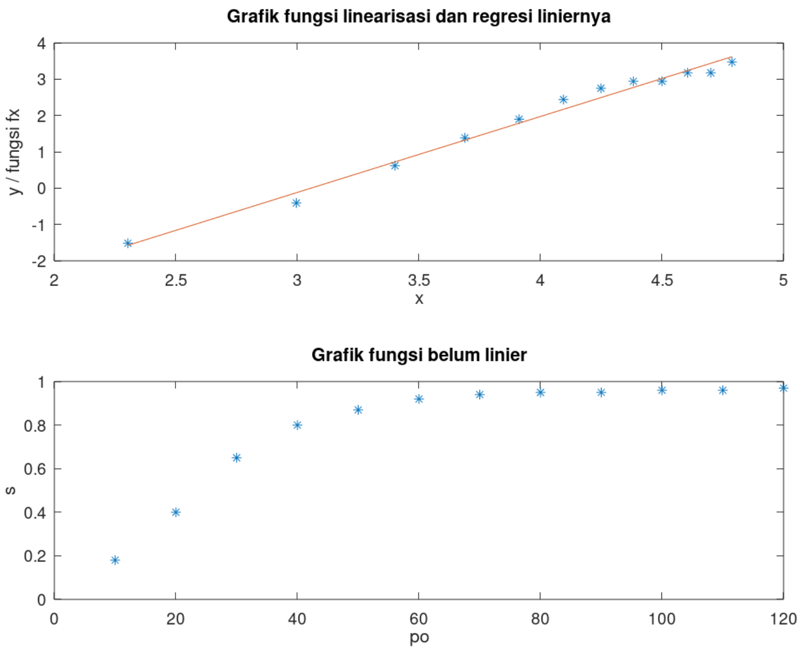
sesuai dengan rumus pada persmaan (3) maka, kemudian nilai matriks s dan po dirubah sedemikian sehingga telah membentuk grafik regresi dengan tren linear. Dari hasil update nilai tiap anggota matriks, kemudian nilai nilai nya disusun dalam matriks bar x dan y sesuai dengan panjang matriks sebelumnya. Hasil update terakhir kemudian akan dipasang dalam matriks baru dan ditampilkan.

Pada syntax berikutnya sama dengan program sebelumnya yaitu mencari nilai sigma yang diperlukan dan kemudian dapat diplot hasil dari pengolahan datanya seperti berikut:



c. Dari hasil pemrograman sebelumnya dan berdasarkan persamaan matematika sebelumnya maka diperoleh nilai n adalah sama dengan nilai a dan nilai P_{50} sesuai dengan kode pemrograman yang telah dibuat yaitu `p50=e.^(-b/a).` sehingga dapat diketahuin nila a atau n adalah 2,0906 dan nilai P_{50} adalah 21,230.

d. Plot data hasil linerisasi dan yang non lineraisasi. Maka akan diperoleh hasil sebagai berikut:



3. Kasus berikutnya adalah membuat persamaan non linier untuk menentukan regresinya. Maka salah satu yang bisa diterapkan adalah menggunakan regresi persamaan kuadrat dengan persamaan fungsi:

$$f(x) = ax^2 + bx + c$$

Sehingga dari persmaan tersebut diperoleh perhitungan parameter yaitu:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum x^4 & \sum x^3 & \sum x^2 \\ \sum x^3 & \sum x^2 & \sum x \\ \sum x^2 & \sum x & n \end{bmatrix}^{-1} \begin{bmatrix} \sum x^2 y \\ \sum xy \\ \sum y \end{bmatrix}$$

Maka dari persamaan matematika tersebut akan diperoleh flowchart program:

Sehingga, dapat dibuat program perhitungan regresi non lineier(kuadrat) pada Octave 5.10 sebagai berikut:

```
y=[100 130 209 276 330 359 420 487 550 645 700];
n=length(x);
%Inisialisasi sigma
sigma4x=0;
sigma3x=0;
sigma2x=0;
sigmax=0;
sigmax2y=0;
sigmaxy=0;
sigmay=0;
for i=1:n
    sigma4x=sigma4x+x(i).^4;
    sigma3x=sigma3x+x(i).^3;
    sigma2x=sigma2x+x(i).^2;
    sigmax=sigmax+x(i);
    sigmax2y=sigmax2y+x(i)*x(i)*y(i);
    sigmaxy=sigmaxy+x(i)*y(i);
    sigmay=sigmay+y(i);
endfor
par = inv([sigma4x sigma3x sigma2x;sigma3x sigma2x sigmax;sigma2x sigmax n])*[sigmax2y;sigmaxy;sigmay];
a = par(1);
b = par(2);
c = par(3);
fungsi=a*(x.*x) + b*x + c;
plot(x,y,'*',x,fungsi)
```

E. KESIMPULAN

Dari hasil percobaan, dapat disimpulkan bahwa regresi mempunyai korelasi dengan data yang dihasilkan. Korelasi tersebut menyesuaikan dengan bentukan grafik data yang ada dimana bisa berbentuk linier maupun non linier. Jika korelasi tersebut berbentuk linier, maka dapat digambarkan dengan persamaan matematik f(x)= ax+b s.

F. DAFTAR PUSTAKA

Capra, Steven C and Canale.1991. “**Numerical Methods for Engineers with Personal Computers Applications**”. MacGraw-Hill Book Company.

King M.R and Mody N.A .2010. “**Numerical and Statical Methods for Bioengineering**”.Cambridge University Press. New York.

Patel VA. 1994. “**Numerical Analysis**” .Saunders College Publishing

Walpole. 1995. Diakses di library.binus.ac.id›eColls›eThesisdoc›Bab2 , pada tanggal 23 September 2019.