

Nonlinear Regression Model Comparison - with Data linearization vs. without Data linearization

©2007 Autar Kaw, Jamie Trahan

University of South Florida

United States of America

kaw@eng.usf.edu

Introduction

This worksheet illustrates comparing a nonlinear regression model determined with data linearization versus a model without data linearization. The comparison is illustrated for three common nonlinear models.

Exponential: $y = ae^{bx}$

Power: $y = ax^b$

Saturation Growth: $y = (ax)/(b+x)$

where a and b are the constants of the above regression models.

Given n data points $(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)$ you can best fit one of the above nonlinear models to the data. To learn more about nonlinear regression models see the [Nonlinear Regression](#) worksheet.

Section 1: Input Data

Below are the input parameters to begin the simulation. This is the only section that requires user input. The user can specify the nonlinear model that is used with the *model_type* variable. Once the values are entered, Maple will return the constants of the nonlinear regression model that are derived with data linearization and without using data linearization.

NOTE: Before evaluating the worksheet, the user must enter initial guesses of the constants of the model a and b for the nonlinear regression without data linearization procedure. For

reasonable initial guesses, use the solution from the [nonlinear model with data linearization worksheet](#). For convergence use initial guesses for a and b close to the values of a and b obtained by using data linearization.

Input Parameters:

n = number of data points

\mathbf{X} = Array of x values

\mathbf{Y} = Array of y values

model_type = type of model the user wishes to evaluate

Ainit = initial guess value for a

Binit = initial guess value for b

Note: Reasonable initial guesses for a and b can be obtained from data linearization models.

```
> restart;
n:=5;
X:=[10,16,25,40,60];
Y:=[94,118,147,180,230];
#For exponential model call model type to be "exponential"
#For power model assign the model type variable as "power"
#For saturation growth model, assign the model type variable to be "growth"
model_type:="exponential";
Ainit:=86.9;
Binit:=0.0159;
```

```
          n := 5
          X := [10, 16, 25, 40, 60]
          Y := [94, 118, 147, 180, 230]
          model_type := "exponential"
          Ainit := 86.9
          Binit := 0.0159
```

(2.1)

Section 2: Nonlinear model with data linearization procedure

In the procedure below, the data is first linearized so that least squares linear regression method for a linear model can be used. Once the coefficients of the linear model are determined, the constants of the nonlinear regression model a and b can be calculated. Linearizing the data is a useful technique to estimate the parameters of a nonlinear model because it does not require iterative methods to solve for the model constants. Note that data linearization is only done for mathematical convenience. For details, see the [Nonlinear Regression worksheet](#).

```
> datalinearized:=proc(x,y,n)
  local z,q,sumq,sumz,sumqz,sumqq,i,a0,a1,a,b;
```

```

z:=array(1..n):
q:=array(1..n):
sumq:=0:
sumz:=0:
sumqq:=0:
sumqz:=0:
#Data linearization step
for i from 1 by 1 to n do
    if (model_type="exponential") then
        z[i]:=ln(y[i]):
        q[i]:=x[i]:
    end if:
    if (model_type="power") then
        z[i]:=ln(y[i]):
        q[i]:=ln(x[i]):
    end if:
    if (model_type="growth") then
        z[i]:=1/y[i]:
        q[i]:=1/x[i]:
    end if:
#Calculating the constants of the linear model,  $z = a_0 + a_1 * q$ 
    sumq:=sumq+q[i]:
    sumz:=sumz+z[i]:
    sumqz:=sumqz+q[i]*z[i]:
    sumqq:=sumqq+(q[i])^2:
end do:
a1:=evalf((n*sumqz-sumq*sumz)/(n*sumqq-sumq^2)):
a0:=evalf((sumz/n)-a1*(sumq/n)):
#Calculating the constants of the original Nonlinear Regression model,  $a$  and  $b$ 
if (model_type="exponential") then
    a:=exp(a0):
    b:=a1:
end if:
if (model_type="power") then
    a:=exp(a0):
    b:=a1:
end if:
if (model_type="growth") then
    a:=1/a0:
    b:=a1/a0:
end if:
return(a,b);

```

```
end proc:
```

Calling the data linearization procedure for the input data set and returning the constants of the nonlinear regression model a and b .

```
> a:=datalinearized(X,Y,n)[1]:
b:=datalinearized(X,Y,n)[2]:
if (model_type="exponential") then
    g:=x->a*exp(b*x);
end if:
if (model_typ="power") then
    g:=x->a*x^b;
end if:
if (model_type="growth") then
    g:=x->(a*x)/(b+x):
end if:
print(`The constants of the `,model_type, ` model found with
data linearization are a= `, a , ` and b= `, b );
```

The constants of the "exponential" model found with data linearization are a= , 87.80451272, and b= , 0.01695287194

(3.1)

Section 3: Nonlinear model without data linearization procedure

In the following procedure, the constants of the nonlinear regression model a and b are found without linearizing the data. This requires the *fsolve* command which utilizes numerical techniques to converge to a real solution. In this section, the initial guess inputs in Section 1, *Ainit* and *Binit*, are used as the starting values.

Assigning the proper regression model.

```
> if (model_type="exponential") then
    f:=x->A*exp(B*x);
end if:
if (model_type="power") then
    f:=x->A*x^B;
end if:
if (model_type="saturation") then
    f:=x->A*x/(B+x);
end if:
```

Calculating the sum of the square of the residuals,

```
> Sr:=0:
```

```

for i from 1 by 1 to n do
Sr:=Sr+(Y[i]-f(X[i]))^2:
end do:

```

Differentiating the sum of the square of the residuals with respect to the constants of the model a and b , to setup two simultaneous nonlinear equations and two unknowns.

```

> eqn1:=expand(diff(Sr,A)=0):
eqn2:=expand(diff(Sr,B)=0):

```

Solving the two simultaneous nonlinear equations. We are using *fsolve* since we are looking for real solutions.

```

> soln:=fsolve({eqn1,eqn2},{A=Ainit,B=Binit}):
assign(soln):
print(`The constants of the `,model_type,` model without data
linearization are a= `,A,` and b= `,B);

```

The constants of the "exponential" model without data linearization are a= , 92.33998795, and b= , 0.01558509828 (4.1)

Section 4: Table of constants

The following table allows you to compare the coefficients a and b that were found with data linearization versus the coefficients found without linearizing the data.

	a		b	
with data linearization	a	87.80451272 (5.1)	b	0.01695287194 (5.2)
without data linearization	A	92.33998795 (5.3)	B	0.01558509828 (5.4)

Section 5: Plot of nonlinear model with data linearization and without data linearization

Plotting the observed values and both predicted curves.

```

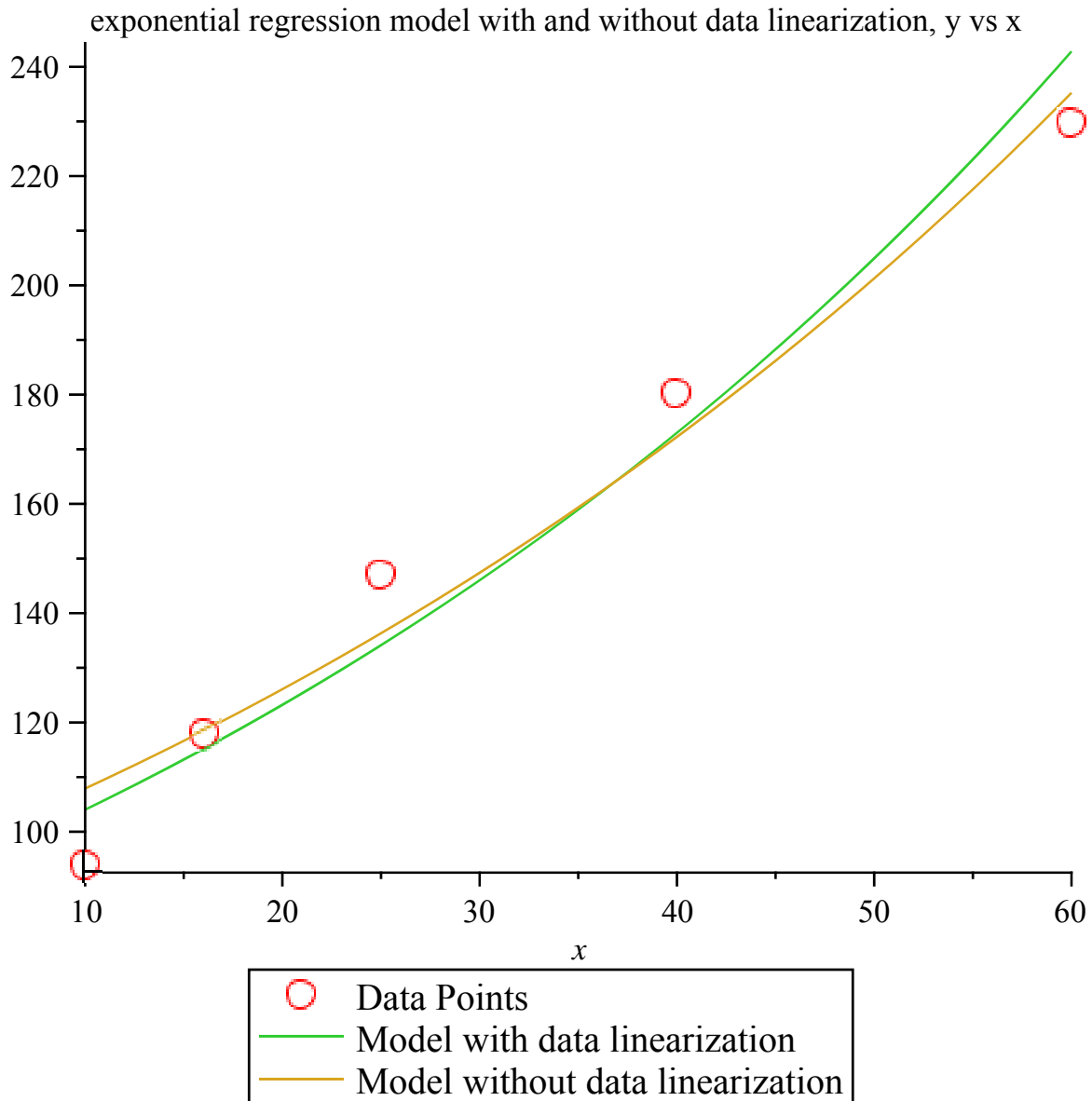
> observed:=[seq([X[i],Y[i]],i=1..n)];
predictedLinear:=g(x):
predictedNonLinear:=f(x):

```

```

ttl:=cat(model_type, ` regression model with and without data
linearization, y vs x`):
plot([observed,predictedLinear,predictedNonLinear],x=X[1]..X
[n],style=[POINT,LINE,LINE],symbol=CIRCLE,symbolsize=20,
title=ttl,legend=["Data Points","Model with data
linearization","Model without data linearization"]);
observed := [[10, 94], [16, 118], [25, 147], [40, 180], [60, 230]]

```



References

[1] Autar Kaw, *Holistic Numerical Methods Institute*, <http://numericalmethods.eng.usf.edu/mws>, See [How does Nonlinear Regression work?](#)

Conclusion

Maple helped us to compare a nonlinear regression model that was determined by linearizing the data to one that was found without linearizing the data.

Question 1: A functional relationship, $\rho = k_1 \exp(-k_2 h)$ exists between the mass density, ρ of air and the altitude above sea level, h .

Altitude above sea level, h (km)	0.32	0.56	0.64	1.28	1.35	1.60
Mass Density, ρ $\left(\frac{\text{kg}}{\text{m}^3}\right)$	1.15	1.12	1.10	1.05	1.03	0.95

- Find the constants, k_1 and k_2 of the model using data linearization
- Find the constants, k_1 and k_2 of the model without data linearization
- Compare the two models from (a) and (b) on a plot
- Find the sum of the square of the residuals for both the models. Which one is smaller? Is that what you expected?

Question 2: Theoretical considerations assume that the rate of flow from a fire hose is proportional to some power of the nozzle pressure. Determine whether this conjecture is true.

Flow rate, F (gallons/min)	91	120	127	190	240	310	409
Pressure, p (psi)	9	15	23	40	61	72	90

- Find the constants, a and b of the model using data linearization
- Find the constants, a and b of the model without data linearization
- Compare the two models on a plot
- Find the sum of the square of the residuals for both the models. Which one is smaller? Is that what you expected?

Legal Notice: The copyright for this application is owned by the author(s). Neither Maplesoft nor the author are responsible for any errors contained within and are not liable for any damages resulting from the use of this material. This application is intended for non-commercial, non-profit use only. Contact the author for permission if you wish to use this application in for-profit activities.