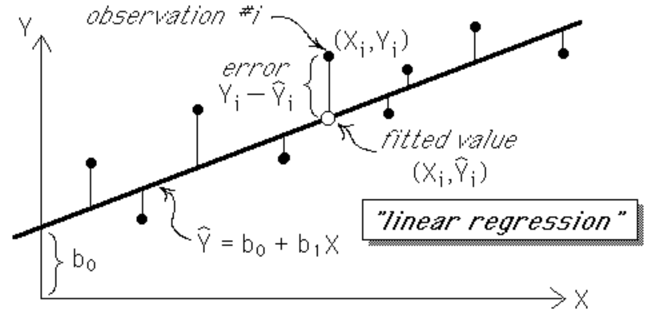
Curve Fitting

This Hypercard stack was prepared by:
Dennis L. Bricker,
Dept. of Industrial Engineering,
University of Iowa,
Iowa City, Iowa 52242
e-mail: dbricker@icaen.uiowa.edu

author



Given a set of points, we wish to draw a line which approximately passes through them!

Error: $\varepsilon_i = Y_i - \hat{Y}_i = Y_i - aX_i - b$

Choose a&b to Minimize $\sum_i \varepsilon_i^2 = \sum_i (Y_i - aX_i - b)^2$

optimality cond'ns

$$\Longrightarrow \begin{cases} \dfrac{\partial}{\partial a}\sum_i \varepsilon_i^2 = -2\sum_i X_i (Y_i - aX_i - b) = 0 \\ \dfrac{\partial}{\partial b}\sum_i \varepsilon_i^2 = -2\sum_i (Y_i - aX_i - b) = 0 \end{cases}$$

linear eqns in a&b

$$\Longrightarrow \begin{cases} \sum_i Y_i = nb + a\sum_i X_i \quad \Longrightarrow b = \dfrac{\sum_i Y_i - a\sum_i X_i}{n} \\ \sum_i X_i Y_i = b\sum_i X_i + a\sum_i X_i^2 \end{cases}$$

$$\Longrightarrow \boxed{a = \frac{\sum_i X_i Y_i - n\overline{X}\,\overline{Y}}{\sum_i X_i^2 - n\overline{X}^2}, \quad b = \overline{Y} - a\overline{X}}$$

note: $\overline{X} = \frac{1}{n}\sum_i X_i$, $\overline{Y} = \frac{1}{n}\sum_i Y_i$

Suppose that an experiment is performed N times, each time varying certain "independent" variables and observing the value of a "dependent" variable:

observation #	Dependent variable Y	INDEPENDENT VARIABLES U	V	W	X	...
1	Y_1	U_1	V_1	W_1	X_1	...
2	Y_2	U_2	V_2	W_2	X_2	...
3	Y_3	U_3	V_3	W_3	X_3	...
4	Y_4	U_4	V_4	W_4	X_4	...
⋮	⋮	⋮	⋮	⋮	⋮	
N	Y_N	U_N	V_N	W_N	X_N	...

Suppose that we suspect the relationship

$$Y = b_0 + b_1U + b_2V + b_3W + b_4X + \ldots$$

where $b_0, b_1, b_2, \ldots$ are unknown to us.

The "least square error" fit to the data is the set of coefficients $b_0, b_1, b_2, \ldots$ which

$$\text{Minimizes } \sum_{i=1}^{N} (Y_i - b_0 - b_1U_i - b_2V_i - b_3W_i - b_4X_i - \ldots)^2$$

Note that $b_0, b_1, b_2, \ldots$ are the variables of this problem, while U_i, V_i, W_i, X_i, etc., are constants!

$$\boxed{\text{Minimize } \sum_{i=1}^{N}(Y_i - b_0 - b_1U_i - b_2V_i - b_3W_i - b_4X_i - \ldots)^2}$$

This is an unconstrained nonlinear optimization problem with a ***quadratic*** objective function $\Phi(b_0, b_1, b_2, \ldots)$

Therefore, its gradient, $\nabla\Phi$, is a ***linear*** function and the necessary condition for an optimum

$$\nabla\Phi(b_0, b_1, b_2, \ldots) = \mathbf{0}$$

is a ***linear*** system of equations.

system of linear equations

$$\begin{cases} \dfrac{\partial\Phi}{\partial b_0} = 0 \\ \dfrac{\partial\Phi}{\partial b_1} = 0 \\ \vdots \end{cases}$$

Other possible objectives, for example

$$\text{Min } \sum_{i=1}^{N} | Y_i - b_0 - b_1U_i - b_2V_i - \ldots |$$

requires LP optimum

are more difficult to achieve, and are therefore seldom used.

To solve:

$$\text{Min} \sum_{i=1}^{N} | Y_i - b_0 - b_1 U_i - b_2 V_i - \ldots |$$

Define new variables z_i, i=1,2, ... n

Linear Program #1

Minimize $\sum_{i=1}^{n} \mathbf{z_i}$

subject to

$$\left.\begin{array}{l} \mathbf{z_i \geq Y_i - b_0 - b_1 U_i - b_2 V_i - \ldots} \\ \mathbf{z_i \geq -Y_i + b_0 + b_1 U_i + b_2 V_i + \ldots} \end{array}\right\} i=1,2, \ldots n$$

To solve:

$$\text{Min} \sum_{i=1}^{N} | Y_i - b_0 - b_1 U_i - b_2 V_i - \ldots |$$

Define new variables z_i^+ & z_i^-, i=1,2, ... n

Linear Program #2

Minimize $\sum_{i=1}^{n} (z_i^+ + z_i^-)$

subject to

$$z_i^+ - z_i^- = \mathbf{Y_i - b_0 - b_1 U_i - b_2 V_i - \ldots} \quad \forall i$$

$$z_i^+ \geq 0, \qquad z_i^- \geq 0 \qquad \forall i$$

Fitting Nonlinear Curves

To fit a polynomial curve

$$Y = b_0 + b_1 X + b_2 X^2 + b_3 X^3 + \ldots$$

make a transformation of variables

$U = X^2$

$V = X^3$

$W = X^4$, etc.

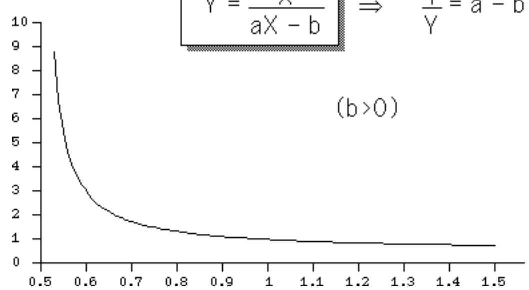
Then the problem becomes that of fitting a linear function: $Y = b_0 + b_1 X + b_2 U + b_3 V + \ldots$

A wide variety of other curves are linearizable:

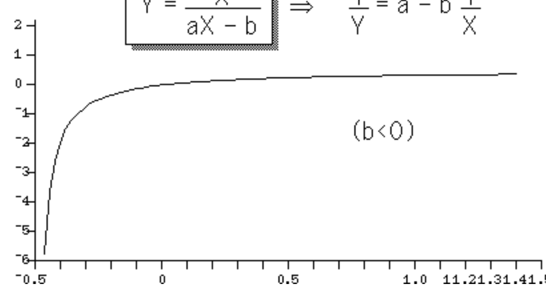
- $Y = aX^b \Rightarrow \ln Y = \ln a + b \ln X$
- $Y = a e^{bX} \Rightarrow \ln Y = \ln a + bX$
- $Y = a e^{b/X} \Rightarrow \ln Y = \ln a + b \frac{1}{X}$
- $Y = aX^b e^{cX} \Rightarrow \ln Y = \ln a + b \ln X + cX$
- $Y = \frac{X}{aX - b} \Rightarrow \frac{1}{Y} = a - b \frac{1}{X}$
- $Y = \frac{1}{a + be^{-X}} \Rightarrow \frac{1}{Y} = a + be^{-X}$
- $Y = a + b \ln X$

Warning

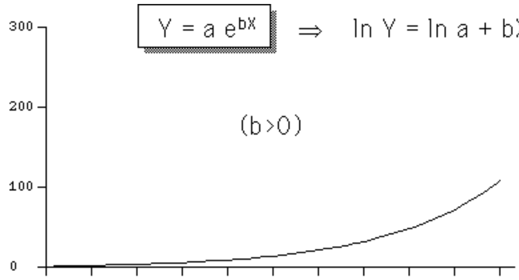
$Y = \frac{X}{aX - b} \Rightarrow \frac{1}{Y} = a - b \frac{1}{X}$



$Y = \frac{X}{aX - b} \Rightarrow \frac{1}{Y} = a - b \frac{1}{X}$

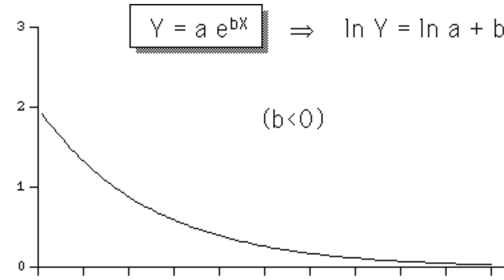


$Y = a e^{bX} \Rightarrow \ln Y = \ln a + bX$

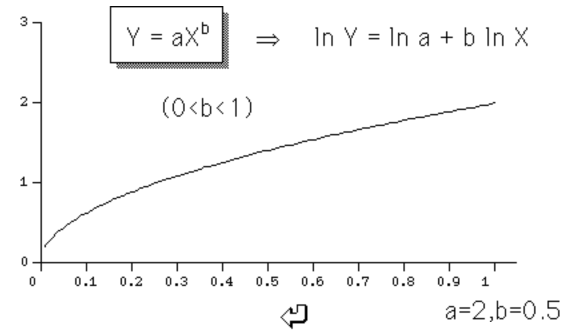


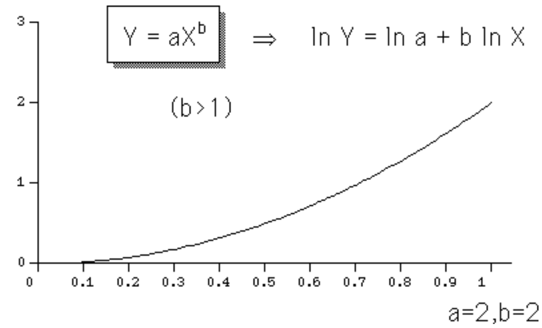
(a=2,b=4)

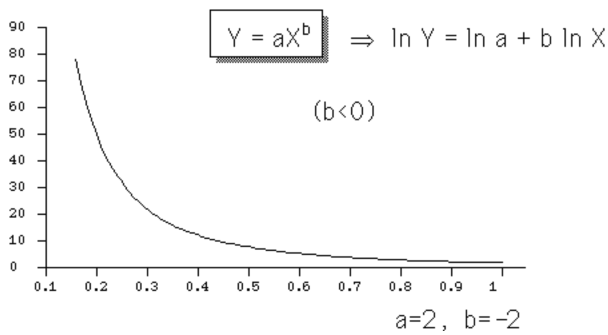
$Y = a e^{bX} \Rightarrow \ln Y = \ln a + bX$

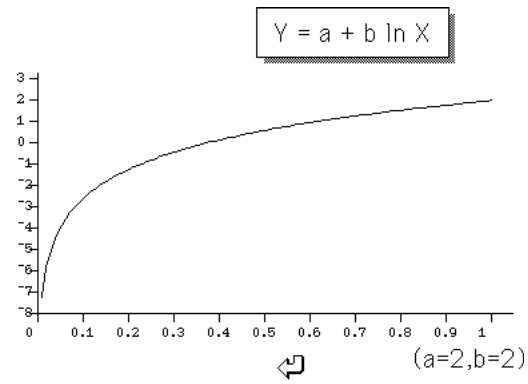


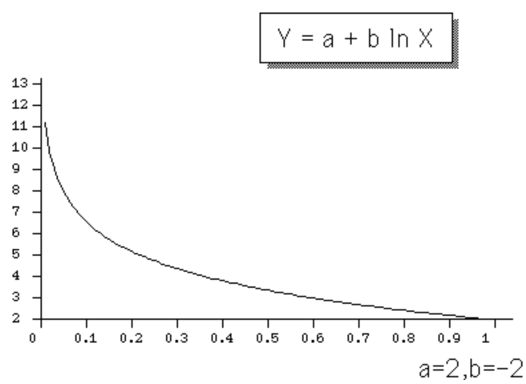
a=2,b=−4

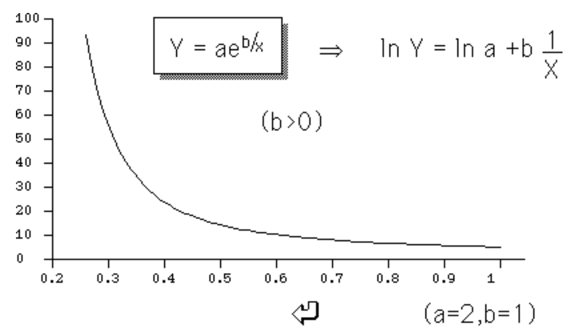

$Y = aX^b$ ⇒ ln Y = ln a + b ln X
(0<b<1)
a=2,b=0.5

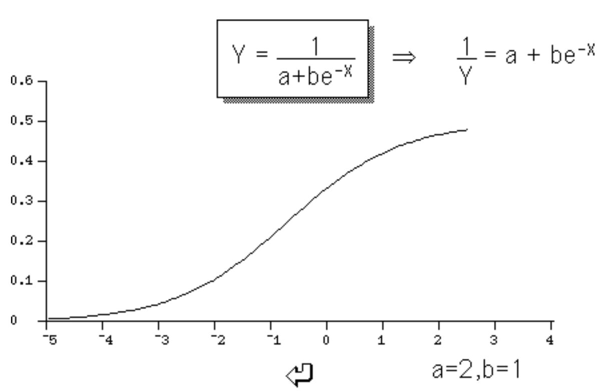

$Y = aX^b$ ⇒ ln Y = ln a + b ln X
(b>1)
a=2,b=2

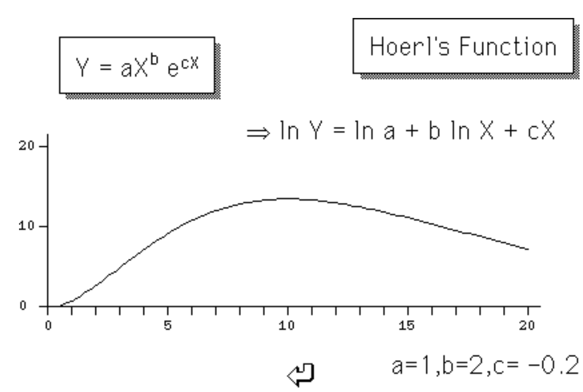

$Y = aX^b$ ⇒ ln Y = ln a + b ln X
(b<0)
a=2, b=−2


Y = a + b ln X
(a=2,b=2)


Y = a + b ln X
a=2,b=−2


$Y = ae^{b/X}$ ⇒ $\ln Y = \ln a + b \frac{1}{X}$
(b>0)
(a=2,b=1)

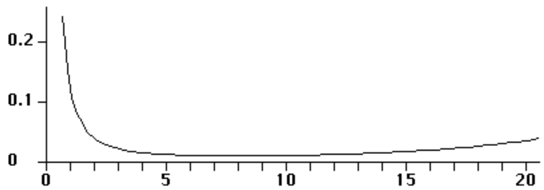

$Y = \frac{1}{a+be^{-X}}$ ⇒ $\frac{1}{Y} = a + be^{-X}$
a=2,b=1


$Y = aX^b e^{cX}$
Hoerl's Function
⇒ ln Y = ln a + b ln X + cX
a=1,b=2,c= −0.2

$Y = aX^b e^{cX}$

Hoerl's Function

$\Rightarrow \ln Y = \ln a + b \ln X + cX$

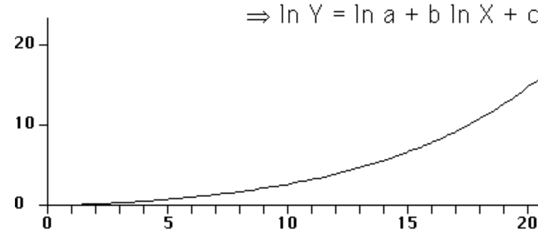


a=0.1, b=-2, c=0.25

$Y = aX^b e^{cX}$

Hoerl's Function

$\Rightarrow \ln Y = \ln a + b \ln X + cX$

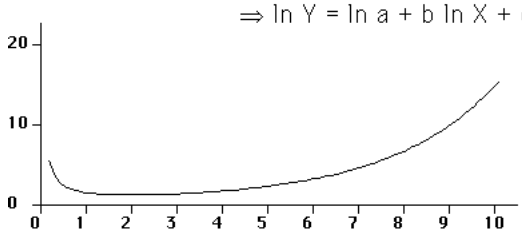


a=0.1, b=1, c=0.1

$Y = aX^b e^{cX}$

Hoerl's Function

$\Rightarrow \ln Y = \ln a + b \ln X + cX$



a=1, b=-1, c=0.5

Warning!

Minimizing the sum of squares of the errors in the linearized form of the fitted curve does not , in general, minimize the sum of squares of the errors in the nonlinear fitted curve.

For example, suppose that we wish to fit a curve $Y = ab^X$ to our data.

The linearized form is $\ln Y = (\ln b)X + \ln a$

original error $\} \quad \varepsilon_i = Y_i - \widehat{Y}_i = Y_i - ab^{X_i}$

error in linearized eqn $\} \quad \tilde{\varepsilon}_i = \ln Y_i - \ln \widehat{Y}_i = \ln Y_i - (\ln b)X_i - \ln a$

The 2 optimization problems:

$$\begin{cases} \text{Minimize } \sum_i \varepsilon_i^2 \\ \text{Minimize } \sum_i \tilde{\varepsilon}_i^2 \end{cases}$$

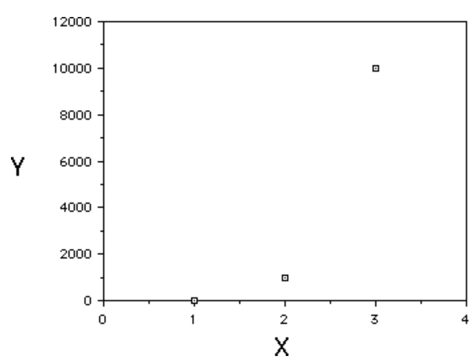
will give different results!

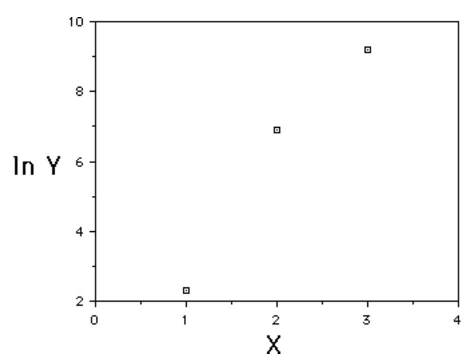
Example:

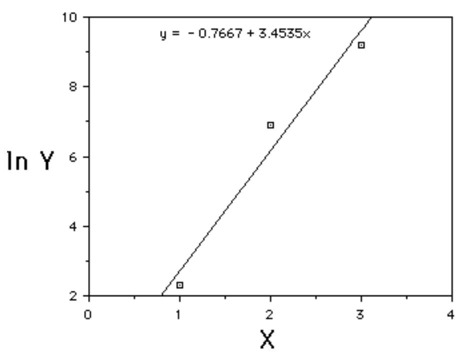
i	1	2	3
X_i	1	2	3
Y_i	10	1000	10000

By inspection, we might guess that a fairly good fit is given by

$$\widehat{Y} = 10 \times 10^X$$







least-square error fit of a line to the (transformed) data

Fitted line: $\ln Y = (\ln b)X + (\ln a)$

$$\ln Y = -0.7667 + 3.4535\,X$$

$\ln a = -0.7667 \Rightarrow a = e^{-0.7667} = 0.46454$

$\ln b = 3.4535 \Rightarrow b = e^{3.4535} = 31.611$

$$\boxed{Y = 0.46454 \times 31.611^X}$$

We have two "fits" of curves of the form $Y = ab^X$ ***to the data:***

based on fit of $\ln Y = \ln a + (\ln b)X$ *(the "least-square error" fit)*

$$\boxed{\hat{Y}' = 0.4645 \times 31.611^X}$$

based upon inspection:

$$\boxed{\hat{Y}'' = 10 \times 10^X}$$

based on fit of $\ln Y = \ln a + (\ln b)X$

X_i	Y_i	$\hat{Y}_i'$	ε_i
1	10	14.685	4.685
2	1000	464.19	535.81
3	10000	14673.64	4673.64

$$\sum_i \tilde{\varepsilon}_i^2 = 22{,}130{,}012.5$$

$$\boxed{\hat{Y}' = 0.4645 \times 31.611^X}$$

X_i	Y_i	$\hat{Y}_i''$	ε_i
1	10	100	90
2	1000	1000	0
3	10000	10000	0

$$\sum_i \varepsilon_i^2 = 8100$$

$$\boxed{\hat{Y}'' = 10 \times 10^X}$$

Only about 0.04% of the sum of the squared errors for the other curves!

The curve which minimizes the squared error in the linear fit to the transformed data does not minimize the squared error in the nonlinear fit to the raw data!

X	Y
667	54
727	42
823	34
1086	75
1529	103
1941	87
2266	53
2515	113
3187	137
3218	114

X	Y
3619	106
3865	98
4266	261
4299	197
4382	106
5560	216
5955	251
6358	347
7165	339
7910	282

It is expected that $Y = aX^b$