Data Linearization

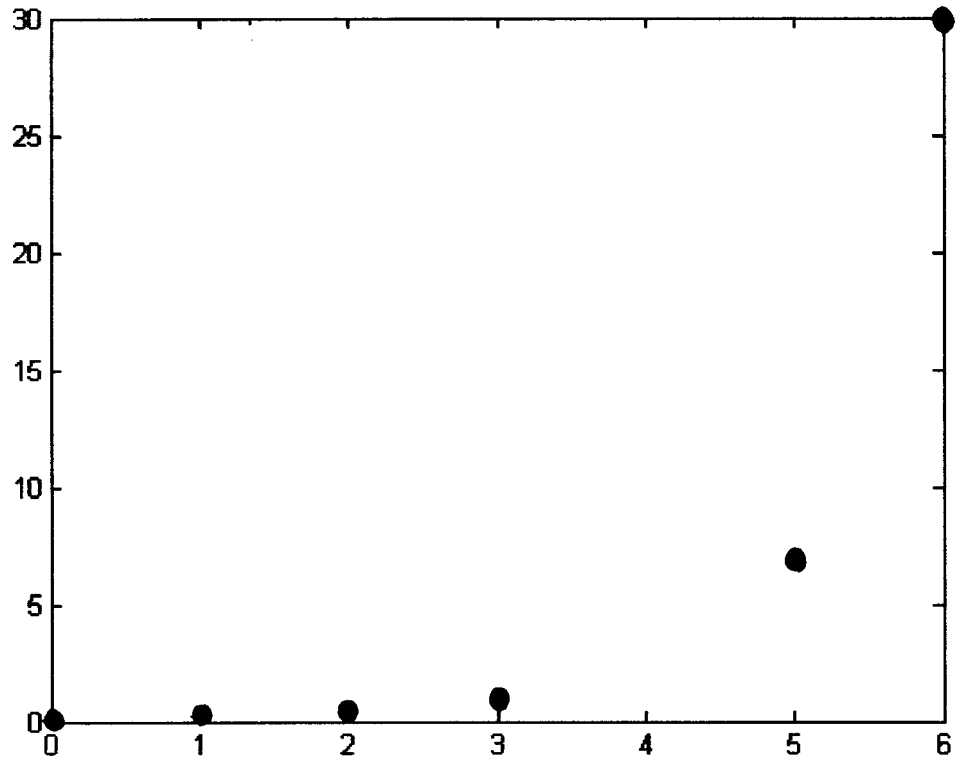
The table below gives the number of computers connected to the internet for selected years.

Year	1989	1990	1991	1992	1994	1995
Numbers (in millions)	.08	.2	.4	.8	6.9	30

1. Based on the data, find the best fitting exponential model. Let $x=0$ correspond to 1989. (General form: $y = C \cdot 10^{bx}$)
2. Use the model to predict the number of computers connected to the internet in the year 2005.

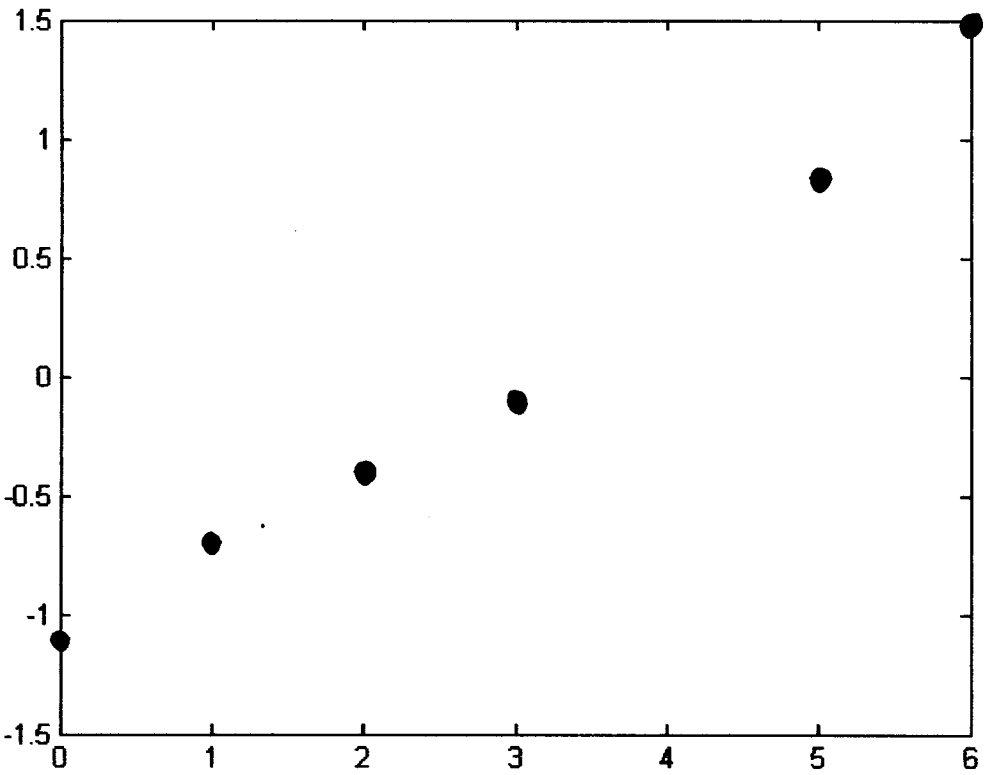
Scatterplot for dataset (x,y)

x	0	1	2	3	5	6
y	.08	.2	.4	.8	6.9	30



Scatterplot for dataset (x, $\log_{10} y$)

x	0	1	2	3	5	6
y	.08	.2	.4	.8	6.9	30
$\log_{10} y$	-1.0969	-.6990	-.3979	-.0969	.8388	1.4771



$y = C \cdot 10^{bx}$ *Start with the exponential model*

$\log_{10} y = \log_{10}(C \cdot 10^{bx})$ *take the log of both sides*

$\log_{10} y = \log_{10} C + \log_{10} 10^{bx}$ *use the log properties to simplify*

$\log_{10} y = \log_{10} C + bx$ *use the log inverse property*

Notice that the last line is a linear equation in the variables (x, $\log_{10} y$).
The slope is b and the y-intercept is the constant, $\log_{10} C$.
We will use the Least Squares Method to find the coefficients $\log_{10} C$ and b. From these we can find C and b for the exponential model.