Chapter 06.04
Nonlinear Models for Regression

After reading this chapter, you should be able to

1. *derive constants of nonlinear regression models,*
2. *use in examples, the derived formula for the constants of the nonlinear regression model, and*
3. *linearize (transform) data to find constants of some nonlinear regression models.*

From fundamental theories, we may know the relationship between two variables. An example in chemical engineering is the Clausius-Clapeyron equation that relates vapor pressure P of a vapor to its absolute temperature, T.

$$\log(P) = A + \frac{B}{T} \quad (1)$$

where A and B are the unknown parameters to be determined. The above equation is not linear in the unknown parameters. Any model that is not linear in the unknown parameters is described as a nonlinear regression model.

Nonlinear models using least squares

The development of the least squares estimation for nonlinear models does not generally yield equations that are linear and hence easy to solve. An example of a nonlinear regression model is the exponential model.

Exponential model

Given (x_1, y_1), (x_2, y_2), . . . (x_n, y_n), best fit $y = ae^{bx}$ to the data. The variables a and b are the constants of the exponential model. The residual at each data point x_i is

$$E_i = y_i - ae^{bx_i} \quad (2)$$

The sum of the square of the residuals is

$$S_r = \sum_{i=1}^{n} E_i^2$$

$$= \sum_{i=1}^{n} \left(y_i - ae^{bx_i} \right)^2 \tag{3}$$

To find the constants a and b of the exponential model, we minimize S_r by differentiating with respect to a and b and equating the resulting equations to zero.

$$\frac{\partial S_r}{\partial a} = \sum_{i=1}^{n} 2\left(y_i - ae^{bx_i} \right)\left(-e^{bx_i} \right) = 0$$

$$\frac{\partial S_r}{\partial b} = \sum_{i=1}^{n} 2\left(y_i - ae^{bx_i} \right)\left(-ax_i e^{bx_i} \right) = 0 \tag{4a,b}$$

or

$$-\sum_{i=1}^{n} y_i e^{bx_i} + a\sum_{i=1}^{n} e^{2bx_i} = 0$$

$$\sum_{i=1}^{n} y_i x_i e^{bx_i} - a\sum_{i=1}^{n} x_i e^{2bx_i} = 0 \tag{5a,b}$$

Equations (5a) and (5b) are nonlinear in a and b and thus not in a closed form to be solved as was the case for linear regression. In general, iterative methods (such as Gauss-Newton iteration method, method of steepest descent, Marquardt's method, direct search, etc) must be used to find values of a and b.

However, in this case, from Equation (5a), a can be written explicitly in terms of b as

$$a = \frac{\sum_{i=1}^{n} y_i e^{bx_i}}{\sum_{i=1}^{n} e^{2bx_i}} \tag{6}$$

Substituting Equation (6) in (5b) gives

$$\sum_{i=1}^{n} y_i x_i e^{bx_i} - \frac{\sum_{i=1}^{n} y_i e^{bx_i}}{\sum_{i=1}^{n} e^{2bx_i}} \sum_{i=1}^{n} x_i e^{2bx_i} = 0 \tag{7}$$

This equation is still a nonlinear equation in b and can be solved best by numerical methods such as the bisection method or the secant method.

Example 1

Many patients get concerned when a test involves injection of a radioactive material. For example for scanning a gallbladder, a few drops of Technetium-99m isotope is used. Half of the techritium-99m would be gone in about 6 hours. It, however, takes about 24 hours for the radiation levels to reach what we are exposed to in day-to-day activities. Below is given the relative intensity of radiation as a function of time.

Table 1 Relative intensity of radiation as a function of time

t (hrs)	0	1	3	5	7	9
γ	1.000	0.891	0.708	0.562	0.447	0.355

If the level of the relative intensity of radiation is related to time via an exponential formula $\gamma = Ae^{\lambda t}$, find

a) the value of the regression constants A and λ,
b) the half-life of Technium-99m, and
c) the radiation intensity after 24 hours.

Solution

a) The value of λ is given by solving the nonlinear Equation (7),

$$f(\lambda) = \sum_{i=1}^{n} \gamma_i t_i e^{\lambda t_i} - \frac{\sum_{i=1}^{n} \gamma_i e^{\lambda t_i}}{\sum_{i=1}^{n} e^{2\lambda t_i}} \sum_{i=1}^{n} t_i e^{2\lambda t_i} = 0 \tag{8}$$

and then the value of A from Equation (6),

$$A = \frac{\sum_{i=1}^{n} \gamma_i e^{\lambda t_i}}{\sum_{i=1}^{n} e^{2\lambda t_i}} \tag{9}$$

Equation (8) can be solved for λ using bisection method. To estimate the initial guesses, we assume $\lambda = -0.23105$ and $\lambda = -0.077016$. We need to check whether these values first bracket the root of $f(\lambda) = 0$. At $\lambda = -0.23105$, the table below shows the evaluation of $f(-0.23105)$.

Table 2 Summation value for calculation of constants of model

i	t_i	γ_i	$\gamma_i t_i e^{\lambda t_i}$	$\gamma_i e^{\lambda t_i}$	$e^{2\lambda t_i}$	$t_i e^{2\lambda t_i}$
1	0	1	0.00000	1.00000	1.00000	0.00000
2	1	0.891	0.70719	0.70719	0.62996	0.62996
3	3	0.708	1.0620	0.35400	0.25000	0.75000
4	5	0.562	0.88509	0.17702	0.09921	0.49606
5	7	0.447	0.62087	0.08870	0.03937	0.27560
6	9	0.355	0.39937	0.04437	0.01562	0.14062
$\sum_{i=1}^{6}$			3.6745	2.3713	2.0342	2.2922

From Table 2

$$n = 6$$

$$\sum_{i=1}^{6} \gamma_i t_i e^{-0.23105 t_i} = 3.6745$$

$$\sum_{i=1}^{6} \gamma_i e^{-0.23105 t_i} = 2.3713$$

$$\sum_{i=1}^{6} e^{2(-0.23105) t_i} = 2.0342$$

$$\sum_{i=1}^{6} t_i e^{2(-0.23105)t_i} = 2.2922$$

$$f(-0.23105) = (3.6745) - \frac{2.3713}{2.0342}(2.2922)$$
$$= 1.0024$$

Similarly

$$f(-0.0077016) = -3.9201$$

Since

$$f(-0.23105) \times f(-0.0077016) < 0,$$

the value of λ falls in the bracket of $[-0.23105, -0.0077016]$. The next guess of the root then is

$$\lambda = \frac{-0.23105 + (-0.0077016)}{2}$$
$$= -0.11938$$

Continuing with the bisection method, the root of $f(\lambda) = 0$ is found as $\lambda = -0.11508$. This value of the root was obtained after 20 iterations with an absolute relative approximate error of less than 0.0002%.

From Equation (9), A can be calculated as

$$A = \frac{\sum_{i=1}^{6} \gamma_i e^{\lambda t_i}}{\sum_{i=1}^{6} e^{2\lambda t_i}}$$

$$= \frac{1 \times e^{-0.11508(0)} + 0.891 \times e^{-0.11508(1)} + 0.708 \times e^{-0.11508(3)} + 0.562 \times e^{-0.11508(5)} + 0.447 \times e^{-0.11508(7)} + 0.355 \times e^{-0.11508(9)}}{e^{2(-0.11508)(0)} + e^{2(-0.11508)(1)} + e^{2(-0.11508)(3)} + e^{2(-0.11508)(5)} + e^{2(-0.11508)(7)} + e^{2(-0.11508)(9)}}$$

$$= \frac{2.9373}{2.9378}$$
$$= 0.99983$$

The regression formula is hence given by

$$\gamma = 0.99983 e^{-0.11508t}$$

b) Half life of Technetium-99m is when $\gamma = \frac{1}{2}\gamma\Big|_{t=0}$

$$0.99983 \times e^{-0.11508t} = \frac{1}{2}(0.99983)e^{-0.11508(0)}$$

$$e^{-0.11508t} = 0.5$$

$$-0.11508t = \ln(0.5)$$

$$t = 6.0232 \text{ hours}$$

c) The relative intensity of the radiation after 24 hrs is

$$\gamma = 0.99983 \times e^{-0.11508(24)}$$
$$= 6.3160 \times 10^{-2}$$

This implies that only $\dfrac{6.3160 \times 10^{-2}}{0.99983} \times 100 = 6.3171\%$ of the initial radioactive intensity is left after 24 hrs.

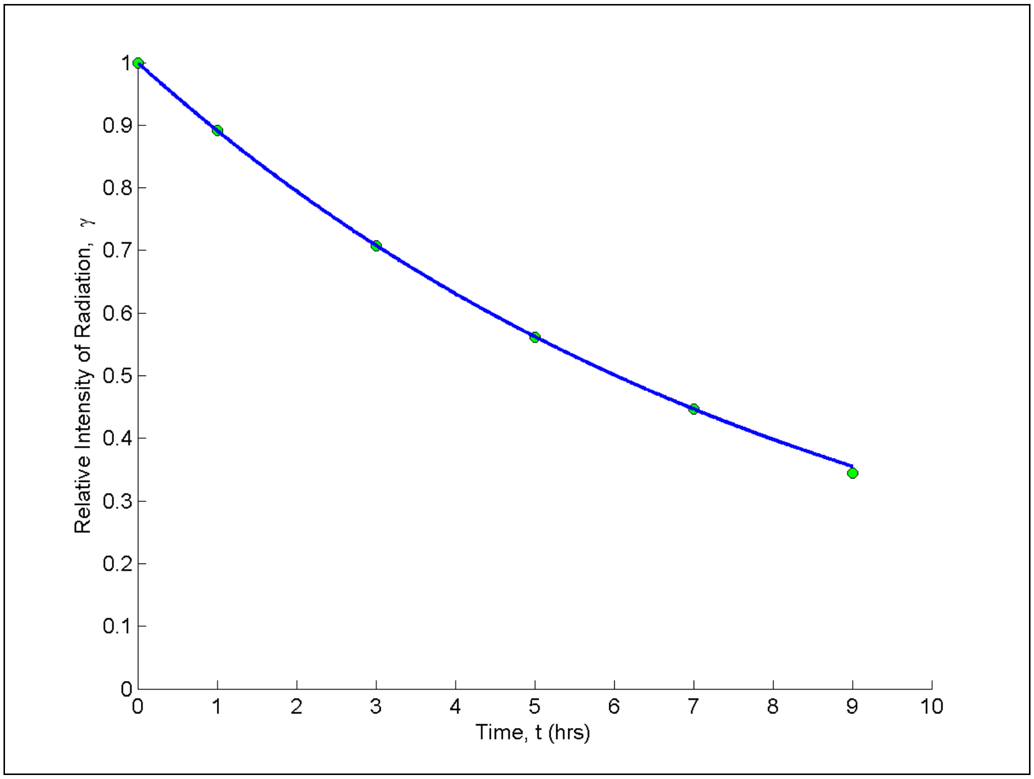

Figure 1 Relative intensity of radiation as a function of temperature using an exponential regression model.

<u>Growth model</u>

Growth models common in scientific fields have been developed and used successfully for specific situations. The growth models are used to describe how something grows with changes in the regressor variable (often the time). Examples in this category include growth of thin films or population with time. Growth models include

$$y = \frac{a}{1 + be^{-cx}} \tag{10}$$

where a, b and c are the constants of the model. At $x = 0$, $y = \dfrac{a}{1+b}$ and as $x \to \infty$, $y \to a$.

The residuals at each data point x_i, are

$$E_i = y_i - \frac{a}{1 + be^{-cx_i}} \tag{11}$$

The sum of the square of the residuals is

$$S_r = \sum_{i=1}^{n} E_i^2$$
$$= \sum_{i=1}^{n}\left(y_i - \frac{a}{1+be^{-cx_i}}\right)^2 \qquad (12)$$

To find the constants a, b and c we minimize S_r by differentiating with respect to a, b and c, and equating the resulting equations to zero.

$$\frac{\partial S_r}{\partial a} = \sum_{i=1}^{n}\left(\frac{2e^{cx_i}\left[ae^{c\times x_i} - y_i\left(e^{cx_i}+b\right)\right]}{\left(e^{cx_i}+b\right)^2}\right) = 0,$$
$$\frac{\partial S_r}{\partial b} = \sum_{i=1}^{n}\left(\frac{2ae^{cx_i}\left[by_i + e^{cx_i}\left(y_i - a\right)\right]}{\left(e^{cx_i}+b\right)^3}\right) = 0,$$
$$\frac{\partial S_r}{\partial c} = \sum_{i=1}^{n}\left(\frac{-2abx_ie^{cx_i}\left[by_i + e^{cx_i}\left(y_i - a\right)\right]}{\left(e^{cx_i}+b\right)^3}\right) = 0. \qquad (13a,b,c)$$

One can use the Newton-Raphson method to solve the above set of simultaneous nonlinear equations for a, b and c.

Example 2

The mechanism of polymer degradation reaction kinetics is suspected to follow Avrami or random nucleation reaction,

$$f(\alpha) = A\frac{(T-T_0)}{b}e^{-\frac{E}{RT}}$$

where $f(\alpha) = -\ln(1-\alpha)$, T is the absolute temperature (K), b is the heating rate in K/min, A is the frequency factor with units of rate constant, R is the gas constant (8.314 kJ/kmol-K) and T_0 is the activation temperature. Given that $T_0 = 338.75\,\text{K}$, $b = 10$ K/min and conversion, α, at different temperatures are as given in table 3. Use the method of least squares to determine the values of A and E.

Table 3 Conversion at given different temperatures

Temp (K)	360	370	380	390	400	410	420	430	440
Conversion, α	0.1055	0.2010	0.3425	0.5146	0.6757	0.8026	0.8924	0.9544	1.00

Solution

To set-up the table, we must re-write equation

$$-\ln(1-\alpha) = A\frac{(T-T_0)}{b}e^{-\frac{E}{RT}}$$

as

$$-\frac{b\ln(1-\alpha)}{(T-T_0)} = Ae^{-\frac{E}{RT}}$$

Taking natural log of both sides of the above equation, we obtain

$$\ln\left[-\frac{b\ln(1-\alpha)}{(T-T_0)}\right] = \ln(A) - \frac{E}{RT}$$

so that the equation is in the form $y = \beta_0 + \beta_1 x$ where

$$y = \ln\left[-\frac{b\ln(1-\alpha)}{(T-T_0)}\right]$$

$$\beta_0 = \ln(A)$$

$$\beta_1 = -\frac{E}{R}$$

$$x = \frac{1}{T}$$

$$\beta_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_1^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$\beta_0 = \left(\frac{\sum_{i=1}^{n} y_i}{n}\right) - \beta_1\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)$$

Table 4 Example on nonlinear exponential problem.

i	T	α	x	y	x^2	$x \times y$
1	360	0.1055	2.7778×10^{-3}	-2.9476	7.7160×10^{-6}	-8.1877×10^{-3}
2	370	0.2010	2.7027×10^{-3}	-2.6338	7.3046×10^{-6}	-7.1183×10^{-3}
3	380	0.3425	2.6316×10^{-3}	-2.2862	6.9252×10^{-6}	-6.0163×10^{-3}
4	390	0.5146	2.5641×10^{-3}	-1.9588	6.5746×10^{-6}	-5.0225×10^{-3}
5	400	0.6757	2.5000×10^{-3}	-1.6936	6.2500×10^{-6}	-4.2341×10^{-3}
6	410	0.8026	2.4390×10^{-3}	-1.4796	5.9488×10^{-6}	-3.6088×10^{-3}
7	420	0.8924	2.3810×10^{-3}	-1.2932	5.6689×10^{-6}	-3.0791×10^{-3}
8	430	0.9544	2.3256×10^{-3}	-1.0835	5.4083×10^{-6}	-2.5199×10^{-3}
$\sum$			2.0322×10^{-2}	-1.5376×10^{1}	5.1797×10^{-5}	-3.9787×10^{-2}

$$n = 8$$

$$\sum_{i=1}^{8} x_i = 2.0322\times10^{-2}$$

$$\sum_{i=1}^{8} y_i = -1.5376\times10^{1}$$

$$\sum_{i=1}^{8} x_i y_i = -3.9787\times10^{-2}$$

$$\sum_{i=1}^{8} x_i^2 = 5.1797.10^{-5}$$

$$\beta_1 = \frac{8\left(-3.9787\times10^{-2}\right)-\left(2.0322\times10^{-2}\right)\left(-1.5376\times10^{-1}\right)}{8\left(5.1797\times10^{-5}\right)-\left(2.0322\times10^{-2}\right)^2}$$

$$= -4.1561\times10^{3}$$

$$\beta_0 = \frac{-1.5376\times10^{1}}{8} - \left(-4.1561\times10^{3}\right)\frac{2.0322\times10^{-2}}{8}$$

$$= 8.6352$$

$$A = e^{\beta_0}$$

$$= e^{8.6352}$$

$$= 5.6264\times10^{3}$$

$$E = -\beta_1 R$$

$$= -\left(-4.1561.10^{3}\right)\times 8.3140$$

$$= 3.4553\times10^{4}$$

This gives the model as

$$-\ln(1-\alpha) = 5.6264\times10^{3}\ \frac{(T-338.75)}{10}\times e^{-\frac{4.1561\times10^{3}}{T}}$$

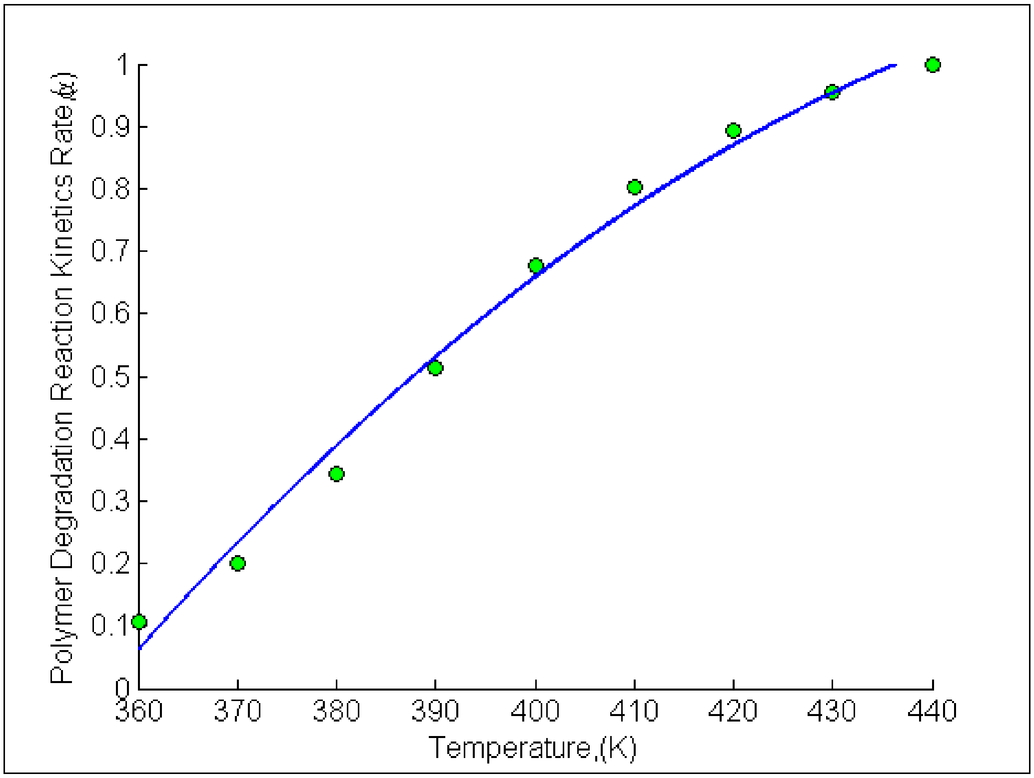

Figure 2 Polymer degradation reaction kinetics rate as a function of temperature.

Example 3

Below is given the FT-IR (Fourier Transform Infra Red) data of a 1:1 (by weight) mixture of ethylene carbonate (EC) and dimethyl carbonate (DMC). Absorbance P is given as a function of wavenumber, m.

Table 5 Absorbance as a function of wavenumber.

Wavenumber, m (cm^{-1})	Absorbance, P (arbitrary unit)
804.184	0.1591
827.326	0.0439
846.611	0.0050
869.753	0.0073
889.038	0.0448
892.895	0.0649
900.609	0.1204

Regress the above data to a second order polynomial

$$P = a_0 + a_1 m + a_2 m^2$$

Find the absorbance at $m = 1000\,\text{cm}^{-1}$

Solution

Table 6 shows the summations needed for the calculations of the constants of the regression model.

Table 6 Summations for calculating constants of model.

i	Wavenumber, m (cm^{-1})	Absorbance, P (arbitrary unit)	m^2	m^3	m^4	$m\times P$	$m^2\times P$
1	804.18	0.1591	6.4671×10^5	5.2008×10^8	4.1824×10^{11}	127.95	1.0289×10^5
2	827.33	0.0439	6.8447×10^5	5.6628×10^8	4.6849×10^{11}	36.319	3.0048×10^4
3	846.61	0.0050	7.1675×10^5	6.0681×10^8	5.1373×10^{11}	4.233	3.583×10^3
4	869.75	0.0073	7.5647×10^5	6.5794×10^8	5.7225×10^{11}	6.349	5.522×10^3
5	889.04	0.0448	7.9039×10^5	7.0269×10^8	6.2471×10^{11}	39.828	3.5409×10^4
6	892.90	0.0649	7.9726×10^5	7.1187×10^8	6.3563×10^{11}	57.948	5.1742×10^4
7	900.61	0.1204	8.1110×10^5	7.3048×10^8	6.5787×10^{11}	108.43	9.7655×10^4
$\sum_{i=1}^{7}$	6030.4	0.4454	5.2031×10^6	4.4961×10^9	3.8909×10^{12}	381.06	3.2685×10^5

$P = a_0 + a_1 m + a_2 m^2$ is the quadratic relationship between the absorbance and the wavenumber where the coefficients a_0, a_1, a_2 are found as follows

$$\begin{bmatrix} n & \left(\sum_{i=1}^{n} m_i\right) & \left(\sum_{i=1}^{n} m_i^2\right) \\ \left(\sum_{i=1}^{n} m_i\right) & \left(\sum_{i=1}^{n} m_i^2\right) & \left(\sum_{i=1}^{n} m_i^3\right) \\ \left(\sum_{i=1}^{n} m_i^2\right) & \left(\sum_{i=1}^{n} m_i^3\right) & \left(\sum_{i=1}^{n} m_i^4\right) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} P_i \\ \sum_{i=1}^{n} m_i P_i \\ \sum_{i=1}^{n} m_i^2 P_i \end{bmatrix}$$

$$n = 7$$

$$\sum_{i=1}^{7} m_i = 6.0304\times10^3$$

$$\sum_{i=1}^{7} m_i^2 = 5.2031\times10^6$$

$$\sum_{i=1}^{7} m_i^3 = 4.4961\times10^9$$

$$\sum_{i=1}^{7} m_i^4 = 3.8909\times10^{12}$$

$$\sum_{i=1}^{7} P_i = 0.4454$$

$$\sum_{i=1}^{7} m_i P_i = 381.06$$

$$\sum_{i=1}^{7} m_i^2 P_i = 3.2685\times10^5$$

We have

$$\begin{bmatrix} 7.0000 & 6.0304\times10^3 & 5.2031\times10^6 \\ 6.0304\times10^3 & 5.2031\times10^6 & 4.4961\times10^9 \\ 5.2031\times10^6 & 4.4961\times10^9 & 3.8909\times10^{12} \end{bmatrix}\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0.4454 \\ 381.06 \\ 3.2685\times10^5 \end{bmatrix}$$

Solve the above system of simultaneous linear equations, we get

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -8.6623 \\ 2.0798\times10^{-2} \\ -1.2365\times10^{-5} \end{bmatrix}$$

The polynomial regression model is

$$P = a_0 + a_1 m + a_2 m^2$$
$$= -8.6623 + 2.0798\times10^{-2} m - 1.2365\times10^{-5} m^2$$

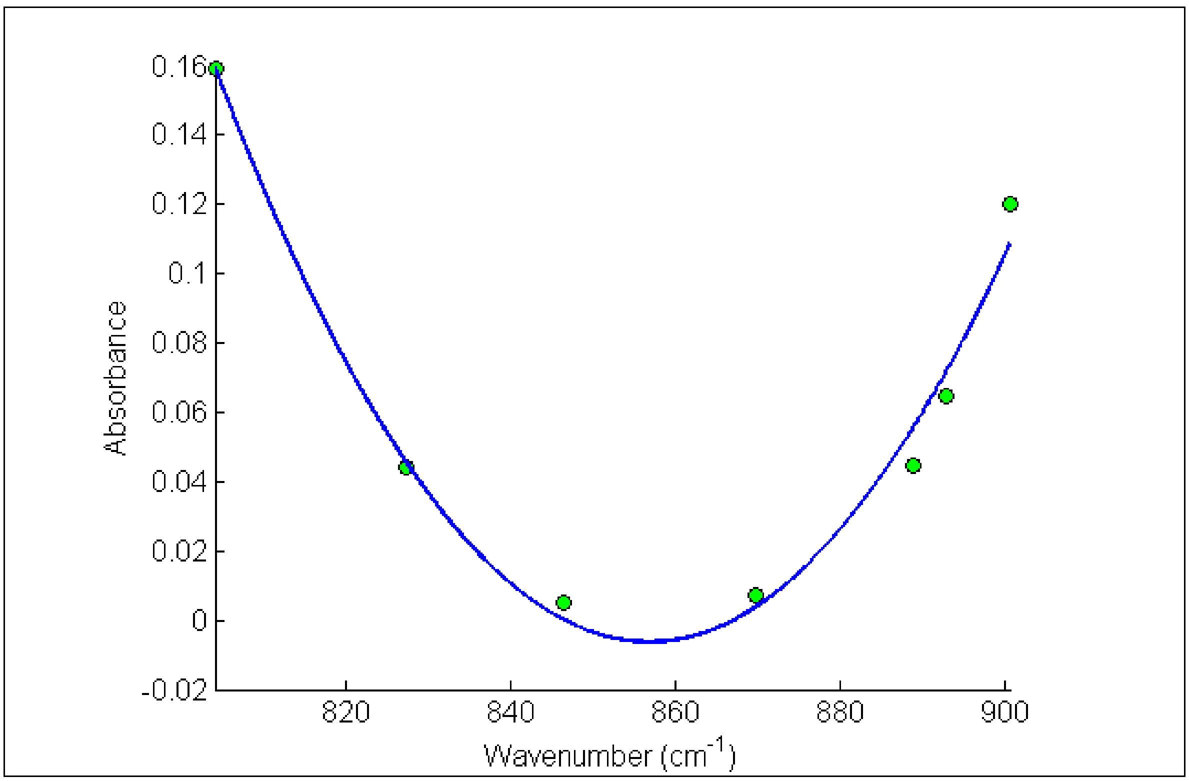

Figure 3 Second order polynomial regression model for absorbance as a function of wavenumber.

To find P where $m = 1000\,\text{cm}^{-1}$:

$$P = a_0 + a_1 m + a_2 m^2$$
$$= -8.6623 + 2.0798\times10^{-2} m - 1.2365\times10^{-5} m^2$$

$$= -836623 + 2.0798 \times 10^{-2} \times (1000) - 1.2365 \times 10^{-5} \times (1000)^2$$
$$= -0.22970$$

Transforming the data to use linear regression formulas

Examination of the nonlinear models above shows that in general iterative methods are required to estimate the values of the model parameters. It is sometimes useful to use simple linear regression formulas to estimate the parameters of a nonlinear model. This involves first transforming the given data such as to regress it to a linear model. Following the transformation of the data, the evaluation of model parameters lends itself to a direct solution approach using the least squares method. Data for nonlinear models such as exponential, power, and growth can be transformed.

<u>Exponential Model</u>

As given in Example 1, many physical and chemical processes are governed by the exponential function.

$$\gamma = ae^{bx} \tag{20}$$

Taking natural log of both sides of Equation (20) gives

$$\ln \gamma = \ln a + bx \tag{21}$$

Let

$$z = \ln \gamma$$
$$a_0 = \ln a \text{ implying } a = e^{a_o}$$
$$a_1 = b$$

then

$$z = a_0 + a_1 x \tag{22}$$

The data z versus x is now a linear model. The constants a_0 and a_1 can be found using the equation for the linear model as

$$a_1 = \frac{n\sum_{i=1}^{n} x_i z_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} z_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{23a,b}$$

$$a_0 = \bar{z} - a_1 \bar{x}$$

Now since a_0 and a_1 are found, the original constants with the model are found as

$$b = a_1$$
$$a = e^{a_0} \tag{24a,b}$$

Example 4

Repeat Example 1 using linearization of data.

Solution

$$\gamma = Ae^{\lambda t}$$

$$\ln(\gamma) = \ln(A) + \lambda t$$

Assuming

$$y = \ln\gamma$$
$$a_0 = \ln(A)$$
$$a_1 = \lambda$$

We get

$$y = a_0 + a_1 t$$

This is a linear relationship between y and t.

$$a_1 = \frac{n\sum_{i=1}^{n} t_i y_i - \sum_{i=1}^{n} t_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} t_i^2 - \left(\sum_{i=1}^{n} t_i\right)^2}$$

$$a_0 = \bar{y} - a_1\bar{t} \qquad \text{(25a,b)}$$

Table 6 Summations of data to calculate constants of model.

i	t_i	γ_i	$y_i = \ln\gamma_i$	$t_i y_i$	t_i^2
1	0	1	0.00000	0.0000	0.0000
2	1	0.891	-0.11541	-0.11541	1.0000
3	3	0.708	-0.34531	-1.0359	9.0000
4	5	0.562	-0.57625	-2.8813	25.000
5	7	0.447	-0.80520	-5.6364	49.000
6	9	0.355	-1.0356	-9.3207	81.000
$\sum_{i=1}^{6}$	25.000		-2.8778	-18.990	165.00

$$n = 6$$
$$\sum_{i=1}^{6} t_i = 25.000$$
$$\sum_{i=1}^{6} y_i = -2.8778$$
$$\sum_{i=1}^{6} t_i y_i = -18.990$$
$$\sum_{i=1}^{6} t_i^2 = 165.00$$

From Equation (25a,b) we have

$$a_1 = \frac{6(-18.990) - (25)(-2.8778)}{6(165.00) - (25)^2}$$
$$= -0.11505$$
$$a_0 = \frac{-2.8778}{6} - (-0.11505)\frac{25}{6}$$

$$= -2.6150 \times 10^{-4}$$

Since

$$a_0 = \ln(A)$$

$$A = e^{a_0}$$

$$= e^{-2.6150 \times 10^{-4}}$$

$$= 0.99974$$

$$\lambda = a_1 = -0.11505$$

The regression formula then is

$$\gamma = 0.99974 \times e^{-0.11505t}$$

Compare the formula to the one obtained without data linearization,

$$\gamma = 0.99983 \times e^{-0.11508t}$$

b) Half-life is when

$$\gamma = \frac{1}{2}\gamma\Big|_{t=0}$$

$$0.99974 \times e^{-0.11505t} = \frac{1}{2}(0.99974)e^{-0.11505(0)}$$

$$e^{-0.11508t} = 0.5$$

$$-0.11505t = \ln(0.5)$$

$$t = 6.0248 \text{ hours}$$

c) The relative intensity of radiation, after 24 hours is

$$\gamma = 0.99974e^{-0.11505(24)}$$

$$= 0.063200$$

This implies that only $\frac{6.3200 \times 10^{-2}}{0.99974} \times 100 = 6.3216\%$ of the initial radioactivity is left after 24 hours.

<u>Logarithmic Functions</u>

The form for the log regression models is

$$y = \beta_0 + \beta_1 \ln(x) \quad (26)$$

This is a linear function between y and $\ln(x)$ and the usual least squares method applies in which y is the response variable and $\ln(x)$ is the regressor.

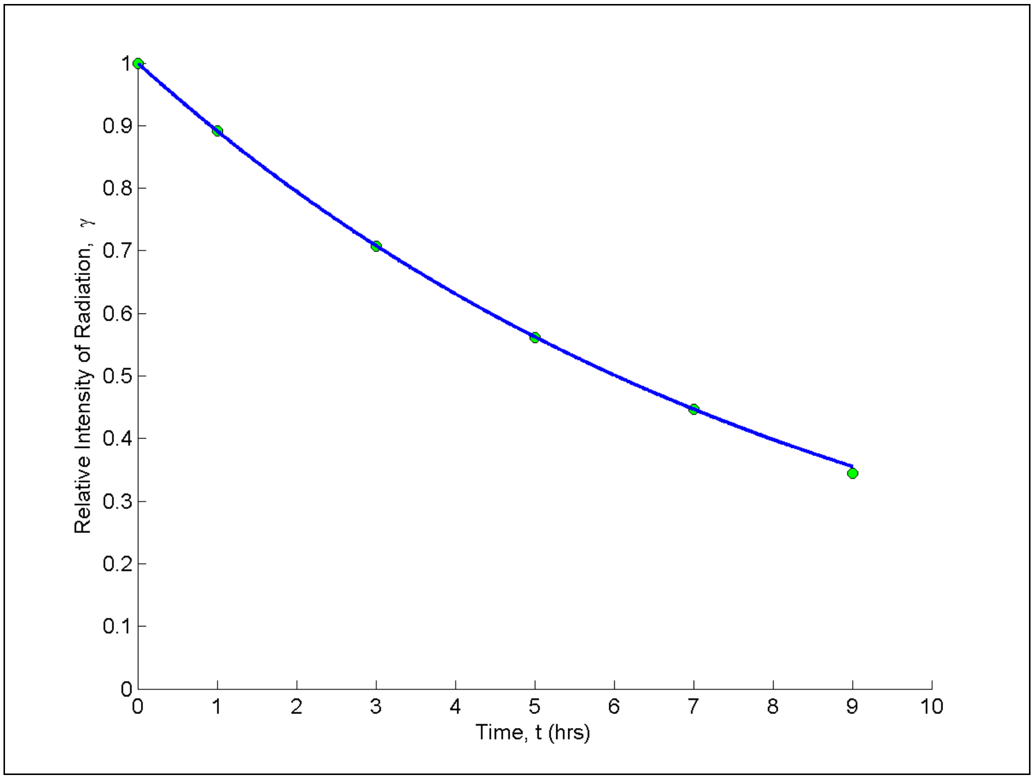

Figure 4 Exponential regression model with transformed data for relative intensity of radiation as a function of temperature.

Example 5

Sodium borohydride is a potential fuel for fuel cell. The following overpotential (η) vs. current (i) data was obtained in a study conducted to evaluate its electrochemical kinetics.

Table 7 Electrochemical Kinetics of borohydride data.

$\eta\,(V)$	-0.29563	-0.24346	-0.19012	-0.18772	-0.13407	-0.0861
$i\,(A)$	0.00226	0.00212	0.00206	0.00202	0.00199	0.00195

At the conditions of the study, it is known that the relationship that exists between the overpotential (η) and current (i) can be expressed as

$$\eta = a + b \ln i \tag{27}$$

where a is an electrochemical kinetics parameter of borohydride on the electrode. Use the data in Table 7 to evaluate the values of a and b.

Solution

Following the least squares method, Table 8 is tabulated where

$$x = \ln i$$
$$y = \eta$$

We obtain

$$y = a + bx \tag{28}$$

This is a linear relationship between y and x, and the coefficients b and a are found as follow

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_1^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$a = \bar{y} - b\bar{x} \tag{29a,b}$$

Table 8 Summation values for calculating constants of model

#	i	$y=\eta$	$x=\ln(i)$	x^2	$x \times y$
1	0.00226	-0.29563	-6.0924	37.117	1.8011
2	0.00212	-0.24346	-6.1563	37.901	1.4988
3	0.00206	-0.19012	-6.1850	38.255	1.1759
4	0.00202	-0.18772	-6.2047	38.498	1.1647
5	0.00199	-0.13407	-6.2196	38.684	0.83386
6	0.00195	-0.08610	-6.2399	38.937	0.53726
$\sum_{i=1}^{6}$	0.012400	-1.1371	-37.098	229.39	7.0117

$$n = 6$$

$$\sum_{i=1}^{6} x_i = -37.098$$

$$\sum_{i=1}^{6} y_i = -1.1371$$

$$\sum_{i=1}^{6} x_i y_i = 7.0117$$

$$\sum_{i=1}^{6} x_i^2 = 229.39$$

$$b = \frac{6(7.0117) - (-37.098)(-1.1371)}{6(229.39) - (-37.098)^2}$$

$$= -1.3601$$

$$a = \frac{-1.1371}{6} - (-1.3601)\frac{-37.098}{6}$$

$$= -8.5990$$

Hence

$$\eta = -8.5990 - 1.3601 \times \ln i$$

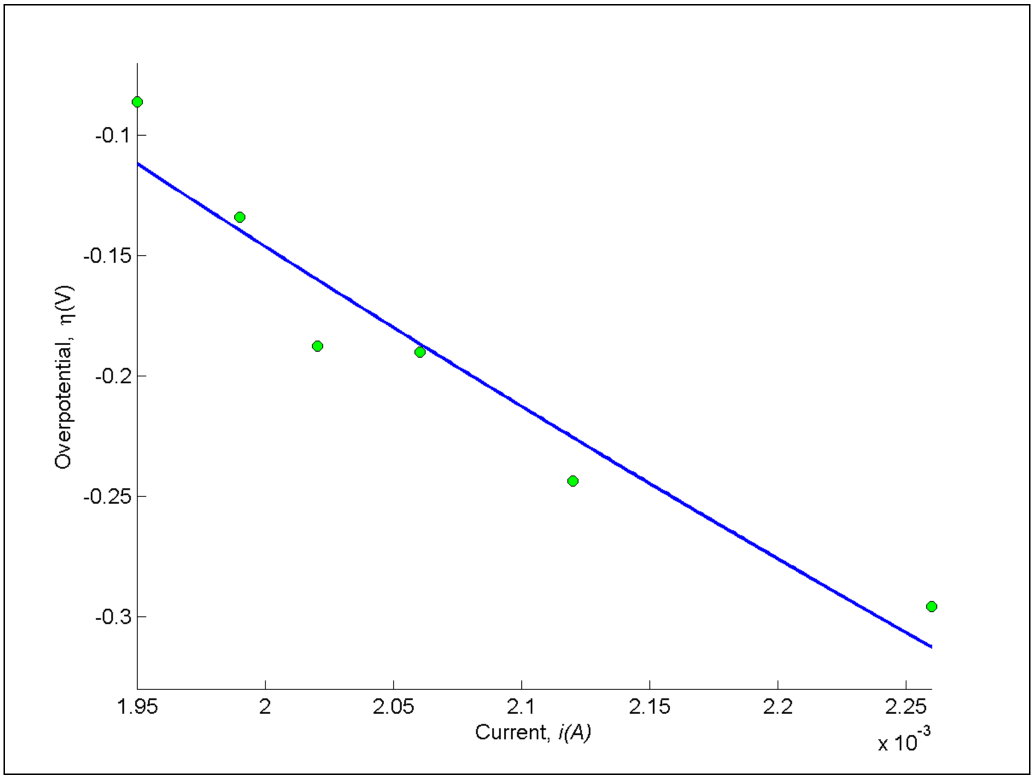

Figure 5 Overpotential as a function of current. $\eta(V)$

Power Functions

The power function equation describes many scientific and engineering phenomena. In chemical engineering, the rate of chemical reaction is often written in power function form as

$$y = ax^b \quad (30)$$

The method of least squares is applied to the power function by first linearizing the data (the assumption is that b is not known). If the only unknown is a, then a linear relation exists between x^b and y. The linearization of the data is as follows.

$$\ln(y) = \ln(a) + b\ln(x) \quad (31)$$

The resulting equation shows a linear relation between $\ln(y)$ and $\ln(x)$.

Let

$$z = \ln y$$

$$w = \ln(x)$$

$$a_0 = \ln a \text{ implying } a = e^{a_0}$$

$$a_1 = b$$

we get

$$z = a_0 + a_1 w \quad (32)$$

$$a_1 = \frac{n\sum_{i=1}^{n} w_i z_i - \sum_{i=1}^{n} w_i \sum_{i=1}^{n} z_i}{n\sum_{i=1}^{n} w_i^2 - \left(\sum_{i=1}^{n} w_i\right)^2} \qquad (33a,b)$$

$$a_0 = \frac{\sum_{i=1}^{n} z_i}{n} - a_1 \frac{\sum_{i=1}^{n} w_i}{n}$$

Since a_0 and a_1 can be found, the original constants of the model are

$$b = a_1$$
$$a = e^{a_0} \qquad (34a,b)$$

Example 6

The progress of a homogeneous chemical reaction is followed and it is desired to evaluate the rate constant and the order of the reaction. The rate law expression for the reaction is known to follow the power function form

$$-r = kC^n \qquad (35)$$

Use the data provided in the table to obtain n and k.

Table 9 Chemical kinetics.

C_A (gmol/l)	4	2.25	1.45	1.0	0.65	0.25	0.006
$-r_A$ (gmol/l·s)	0.398	0.298	0.238	0.198	0.158	0.098	0.048

Solution

Taking the natural log of both sides of Equation (35), we obtain

$$\ln(-r) = \ln(k) + n\ln(C)$$

Let

$$z = \ln(-r)$$
$$w = \ln(C)$$
$$a_0 = \ln(k) \text{ implying that } k = e^{a_0} \qquad (36)$$
$$a_1 = n \qquad (37)$$

We get

$$z = a_0 + a_1 w$$

This is a linear relation between z and w, where

$$a_1 = \frac{n\sum_{i=1}^{n} w_i z_i - \sum_{i=1}^{n} w_i \sum_{i=1}^{n} z_i}{n\sum_{i=1}^{n} w_i^2 - \left(\sum_{i=1}^{n} w_i\right)^2}$$

$$a_0 = \left(\frac{\sum_{i=1}^{n} z_i}{n}\right) - a_1\left(\frac{\sum_{i=1}^{n} w_i}{n}\right) \qquad (38a,b)$$

Table 10 Kinetics rate law using power function

i	C	$-r$	w	z	$w \times z$	w^2
1	4	0.398	1.3863	-0.92130	-1.2772	1.9218
2	2.25	0.298	0.8109	-1.2107	-0.9818	0.65761
3	1.45	0.238	0.3716	-1.4355	-0.5334	0.13806
4	1	0.198	0.0000	-1.6195	0.0000	0.00000
5	0.65	0.158	-0.4308	-1.8452	0.7949	0.18557
6	0.25	0.098	-1.3863	-2.3228	3.2201	1.9218
7	0.006	0.048	-5.1160	-3.0366	15.535	26.173
$\sum_{i=1}^{7}$			-4.3643	-12.391	16.758	30.998

$$n = 7$$
$$\sum_{i=1}^{7} w_i = -4.3643$$
$$\sum_{i=1}^{7} z_i = -12.391$$
$$\sum_{i=1}^{7} w_i z_i = 16.758$$
$$\sum_{i=1}^{7} w_i^2 = 30.998$$

From Equation (38a,b)

$$a_1 = \frac{7 \times (16.758) - (-4.3643) \times (-12.391)}{7 \times (30.998) - (-4.3643)^2}$$
$$= 0.31943$$

$$a_0 = \frac{-12.391}{7} - (.31943)\frac{-4.3643}{7}$$
$$= -1.5711$$

From Equation (36) and (37), we obtain

$$k = e^{-1.5711}$$
$$= 0.20782$$
$$n = a_1$$
$$= 0.31941$$

Finally, the model of progress of that chemical reaction is

$$-r \ = 0.20782 \times C^{0.31941}$$

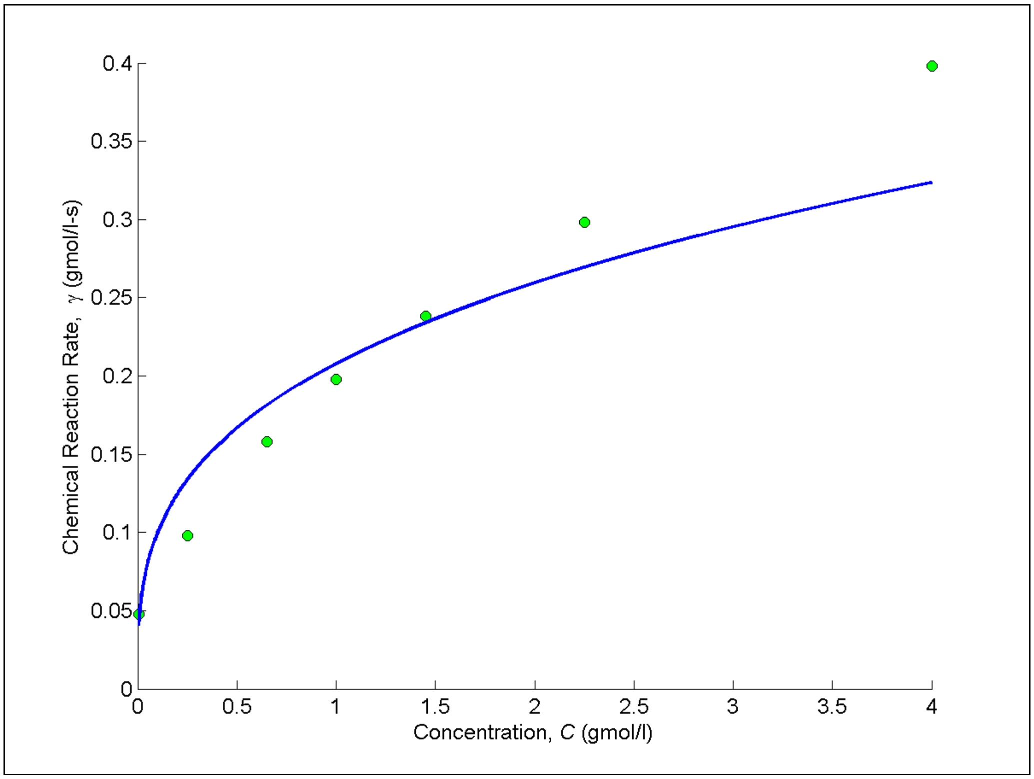

Figure 6 Kinetic chemical reaction rate as a function of concentration.

Growth Model

Growth models common in scientific fields have been developed and used successfully for specific situations. The growth models are used to describe how something grows with changes in a regressor variable (often the time). Examples in this category include growth of thin films or population with time. In the logistic growth model, an example of a growth model in which a measurable quantity y varies with some quantity x is

$$y = \frac{ax}{b+x} \tag{39}$$

For $x = 0$, $y = 0$ while as $x \to \infty$, $y \to a$. To linearize the data for this method,

$$\begin{aligned} \frac{1}{y} &= \frac{b+x}{ax} \\ &= \frac{b}{a}\frac{1}{x} + \frac{1}{a} \end{aligned} \tag{40}$$

Let

$$z = \frac{1}{y}$$

$$w \ = \frac{1}{x},$$

$$a_0 = \frac{1}{a} \text{ implying that } a = \frac{1}{a_0}$$

$$a_1 = \frac{b}{a} \text{ implying } b = a_1 \times a = \frac{a_1}{a_0}$$

Then

$$z = a_0 + a_1 w \tag{41}$$

The relationship between z and w is linear with the coefficients a_0 and found as follows.

$$a_1 = \frac{n\sum_{i=1}^{n} w_i z_i - \sum_{i=1}^{n} w_i \sum_{i=1}^{n} z_i}{n\sum_{i=1}^{n} w_i^2 - \left(\sum_{i=1}^{n} w_i\right)^2}$$

$$a_0 = \left(\frac{\sum_{i=1}^{n} z_i}{n}\right) - a_1\left(\frac{\sum_{i=1}^{n} w_i}{n}\right) \tag{42a,b}$$

Finding a_0 and a_1, then gives the constants of the original growth model as

$$a = \frac{1}{a_0}$$

$$b = \frac{a_1}{a_0} \tag{43a,b}$$

NONLINEAR REGRESSION	
Topic	Nonlinear Regression
Summary	Textbook notes of Nonlinear Regression
Major	Chemical Engineering
Authors	Egwu Kalu, Autar Kaw, Cuong Nguyen
Date	September 10, 2009
Web Site	http://numericalmethods.eng.usf.edu