United Arab Emirates University
College of Sciences
Department of Mathematical Sciences

HOMEWORK 1 – SOLUTION

Sec 1.1 – Sec 1.3

Mathematical Modeling
MATH 470 SECTION **51** CRN 31749
8:00 – 9:15 on Sunday, Monday, Tuesday, & Wednesday
Due Date: Wednesday, February 17, 2010

ID No: *Solution*

Name: *Solution*

Score: *Solution*

Section 1.1 Modeling Change with Difference Equations

1. Find a formula for the n^{th} term of the sequence.

(1.1) $\{1, 4, 16, 64, 256, \ldots\}$

Answer. Look for the pattern:

$$a_1 = 1 = 2^0 = (2^0)^2 \qquad a_2 = 4 = 2^2 = (2^1)^2 \qquad a_3 = 16 = 4^2 = (2^2)^2$$

$$a_4 = 64 = 8^2 = (2^3)^2 \qquad a_5 = 256 = 16^2 = (2^4)^2 \qquad a_n = (2^n)^2 = 2^{2n}. \qquad \square$$

(1.2) $\{1, 3, 7, 15, 31, \ldots\}$

Answer. Look for the pattern:

$$\triangle a_1 = a_2 - a_1 = 3 - 1 = 2 = 2^1 \qquad \triangle a_2 = a_3 - a_2 = 7 - 3 = 4 = 2^2$$

$$\triangle a_3 = a_4 - a_3 = 15 - 7 = 8 = 2^3 \qquad \triangle a_4 = a_5 - a_4 = 31 - 15 = 16 = 2^4$$

$$\triangle a_n = a_{n+1} - a_n = 2^n \implies a_{n+1} = 2^n a_n. \qquad \square$$

- *Remark:* If we add all the equations up to $\triangle a_{n-1}$ above, by the telescoping cancelation,

$$\sum_{i=1}^{n-1} \triangle a_i = a_n - a_1 = 2^1 + 2^2 + 2^3 + \cdots + 2^{n-1} = \sum_{i=1}^{n-1} 2^i = 2^n - 2$$

$$\implies \quad a_n = 2^n - 2 + a_1 = 2^n - 2 + 1 = 2^n - 1 \quad \implies \quad a_n = 2^n - 1.$$

2. By examining the following sequences, write a difference equation to represent the change during the n^{th} interval as a function of the previous term in the sequence.

(2.1) $\{1, 2, 5, 11, 23, \ldots\}$

Answer. Ignore! $\square$

(2.2) $\{1, 8, 29, 92, \ldots\}$

Answer. Look for the pattern:

$$\triangle a_1 = a_2 - a_1 = 8 - 1 = 7, \qquad \triangle a_2 = a_3 - a_2 = 29 - 8 = 21 = 3 \cdot 7 = 3\triangle a_1$$

$$\triangle a_3 = a_4 - a_3 = 92 - 29 = 63 = 3 \cdot 21 = 3\triangle a_2$$

$$\triangle a_n = 3\triangle a_{n-1} \quad \implies \quad a_{n+1} - a_2 = \sum_{i=2}^{n} \triangle a_i = 3\sum_{i=2}^{n} \triangle a_{i-1} = 3\left(a_n - a_1\right)$$

$$a_{n+1} = 3a_n - 3a_1 + a_2 = 3a_n + 5 \quad \overset{\textit{formula on linear system}}{\implies} \quad a_n = \frac{3^n 7 - 15}{6}. \qquad \square$$

3. Formulate a dynamical system that models change exactly for the described situation.

(3.1) You currently have $\$5000$ in a savings account that pays 0.5% interest each month. You add another $\$200$ each month.

Answer. Let a_n be the amount of money in the account after n months. Then we have

$$\triangle a_n = 0.005a_n + 200 \qquad \Longrightarrow \qquad a_{n+1} = 1.005a_n + 200 \quad with \quad a_0 = 5000.$$

The formula on the linear dynamical system implies

$$a_n = 1.005^n 45000 - 40000 = 45000\left(1.005^n - \frac{8}{9}\right).$$

□

(3.2) Your parents are considering a 30–year, $\$100,000$ mortgage that charges 0.5% interest each month. Formulate a model in terms of a monthly payment p that allows the mortgage (loan) to be paid off after 360 payments. (*Hint:* If a_n represents the amount owed after n months, what are a_0 and a_{360}?)

Answer. Let a_n be the amount owed after n months. Then we have

$$\triangle a_n = 0.005a_n - p \qquad \Longrightarrow \qquad a_{n+1} = 1.005a_n - p \quad with \quad a_0 = 1000000.$$

The formula on the linear dynamical system implies

$$\begin{aligned} a_n &= 1.005^n\,(1000000 - 200p) + 200p = 200\,[1.005^n(5000 - p) + p] \\ a_0 &= 200\left[1.005^0(5000 - p) + p\right] = 1000000 \\ a_{360} &= 200\left[1.005^{360}(5000 - p) + p\right] = 200(6.02258(5000 - p) + p) \\ &\approx 6022575 - 1004.52p \end{aligned}$$

Since $a_{360} \approx 6022575 - 1004.52p = 0$ at $p \approx 5995.51$, so they need to pay about $\$6000$ each month for the complete pay off after 360 payments, i.e., 30 years. □

- *Remark:* If we increase the interest rate from 0.5% to 1%, then how much money should they pay for the complete pay off after 360 payments, i.e., 30 years?

Section 1.2 Approximating Change with Difference Equations

4. The following data were obtained for the growth of a sheep population introduced into a new environment on the island of Tasmania.

Year	1814	1824	1834	1844	1854	1864
Population	125	275	830	1200	1750	1650

Plot the data. Is there a trend? Plot the change in population versus years elapsed after 1814. Formulate a discrete dynamical system that reasonably approximates the change you have observed.

Answer. Ignore! $\square$

5. Sociologists recognize a phenomenon called *social diffusion*, which is the spreading of a piece of information, a technological innovation, or a cultural fad among a population. The members of the population can be divided into two classes: those who have the information and those who do not. In a fixed population whose size is known, it is reasonable to assume that the rate of diffusion is proportional to the number who have the information times the number yet to receive it. If a_n denotes the number of people who have the information in a population N people after n days, formulate a dynamical system to approximate the change in the number of people in the population who have the information.

Answer. It is clear to see

$$\text{rate of diffusion} = ka_n\left(N - a_n\right).$$

Here k is the proportionality constant. In the discrete case like this problem, the rate means $\dfrac{a_{n+1} - a_n}{n+1-n} = \triangle a$
(In the continuous case, it means the derivative.) So, the equation is in fact

$$\triangle a_n = ka_n\left(N - a_n\right). \qquad \square$$

Section 1.3 Solutions to Dynamical Systems

6. Find the solution to the difference equations in the following problems.

(6.1) $a_{n+1} = 3a_n/4$, $a_0 = 64$

Answer. The formula on the solution of the linear dynamical system implies

$$a_n = \left(\frac{3}{4}\right)^n a_0 = \left(\frac{3}{4}\right)^n 64. \qquad \square$$

(6.2) $a_{n+1} = 0.1a_n + 3.2$, $a_0 = 1.3$

Answer. The formula on the solution of the linear dynamical system implies

$$a_n = -\frac{203}{90}\,0.1^n + \frac{32}{9}. \qquad \square$$

7. For the following problems, find an equilibrium value if one exists. Classify the equilibrium value as stable or unstable.

(7.1) $a_{n+1} = 0.9a_n$

Answer. Suppose the system has the equilibrium value E. Then

$$E = 0.9E \quad \Longrightarrow \quad E = 0.$$

In the given system, $|r = 0.9| < 1$ and so the equilibrium value is stable. $\square$

(7.2) $a_{n+1} = -0.8a_n + 100$

Answer. Suppose the system has the equilibrium value E. Then

$$E = -0.8E + 100 \quad \Longrightarrow \quad E = \frac{500}{9}.$$

In the given system, $|r = -0.8| < 1$ and so the equilibrium value is stable. $\square$

(7.3) $a_{n+1} = a_n - 100$

Answer. In the given system, $r = 1$ and so the equilibrium value does not exist. $\square$

8. Build a numerical solution for the *initial value problem*, $a_{n+1} = 0.8a_n - 100$, $a_0 = 500$. Plot your data to observe pattern in the solution. Is there an equilibrium value? Is it stable or unstable?

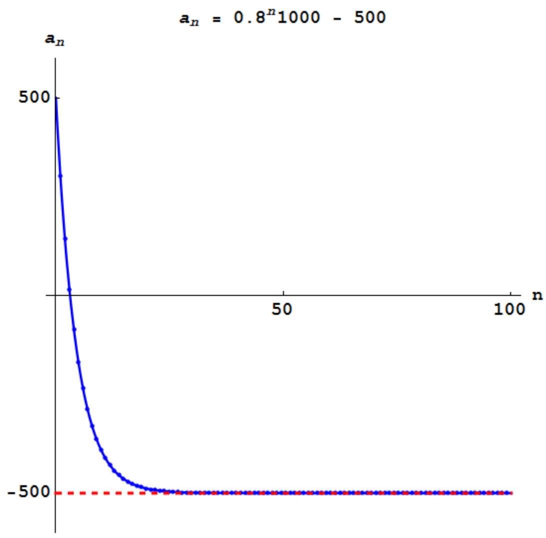

Figure 1: Continuous graph of solution a_n

Answer. Let E be the equilibrium value of the system. Then, we have

$$E = 0.8E - 100 \quad \Longrightarrow \quad E = -\frac{100}{0.2} = -500.$$

By the formula, the dynamical system has the solution

$$a_n = 0.8^n (a_0 + 500) - 500 = 0.8^n 1000 - 500.$$

As $n \to \infty$, $0.8^n \to 0$ and so $a_n \to -500 = E$. Thus, the equilibrium value is stable. See the figure 1. $\square$

9. Consider $a_{n+1} = a_n + 2$, $a_0 = -1$. Find the solution to the difference equation and the equilibrium value if one exists. Discuss the long–term behavior of the solution for the various initial values. Classify the equilibrium values as stable or unstable.

Answer. Suppose E is the equilibrium value of the system. Then,

$$E = E + 2,$$

which cannot be true for any E. Thus, there is no equilibrium value.

Now we find the solution of the dynamical system as follows.

$$\triangle a_n = a_{n+1} - a_n = 2 \quad \Longrightarrow \quad \triangle a_n = 2 \quad \Longrightarrow \quad a_n - a_0 = \sum_{i=0}^{n-1} \triangle a_i = \sum_{i=0}^{n-1} 2 = 2n$$

$$\Longrightarrow \quad a_n = 2n + a_0 = 2n - 1.$$

$\square$

10. You currently have \$5000 in a savings account that pays 0.5% interest each month. You add \$200 each month. Build a numerical solution to determine when the account reaches \$20,000.

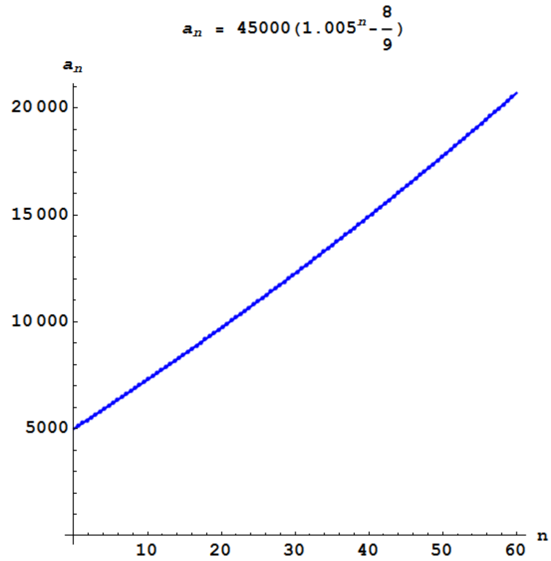

Figure 2: Continuous graph of solution a_n

Answer. Let a_n be the amount owed after n months. Then we have

$$\triangle a_n = 0.005a_n + 200 \qquad \Longrightarrow \qquad a_{n+1} = 1.005a_n + 200 \quad \textit{with} \quad a_0 = 5000.$$

The formula on the linear dynamical system implies

$$a_n = 1.005^n 45000 - 40000 = 45000\left(1.005^n - \frac{8}{9}\right).$$

Now we solve $a_n = 20000$ for n:

$$20000 = a_n = 45000\left(1.005^n - \frac{8}{9}\right) \quad \Longrightarrow \quad \frac{4}{9} = 1.005^n - \frac{8}{9} \quad \Longrightarrow \quad n = \frac{\ln(4/3)}{\ln 1.005} \approx 57.6801.$$

Thus, after 58 months, the account will reach \$20000. $\square$

United Arab Emirates University
College of Sciences
Department of Mathematical Sciences

HOMEWORK 2 – SOLUTION

Sec 2.2 – Sec 3.4

Mathematical Modeling
MATH 470 SECTION **51** CRN 31749
8:00 – 9:15 on Sunday, Monday, Tuesday, & Wednesday
Due Date: Monday, March 15, 2010

Section 2.2 Modeling Using Proportionality

1. If a spring is stretched 0.37 in. by a 14–lb force, what stretch will be produced by a 9–lb force? By a 22–lb force? Assume Hooke's law, which asserts the distance stretched is proportional to the force applied.

Answer. Hooke's Law says $F = kS$, where F is the restoring force in a spring stretched or compressed a distance S. The given condition implies

$$4 \text{ lb} = k(0.37 \text{ in}) \quad \Longrightarrow \quad k = \frac{4 \text{ lb}}{0.37 \text{ in}} = 10.8108 \text{ lb/in}.$$

Using this proportional constant, we get

$$\begin{aligned} 9 \text{ lb} &= (10.8108 \text{ lb/in})S, \quad i.e., \quad S = \frac{9}{10.8108} \text{ in} = 0.8325 \text{ in}. \\ 22 \text{ lb} &= (10.8108 \text{ lb/in})S, \quad i.e., \quad S = \frac{22}{10.8108} \text{ in} = 2.035 \text{ in}. \end{aligned}$$

□

2. If an architectural drawing is scaled so that 0.75 in. represents 4 ft, what length represents 27 ft?

Answer. Let x be the length representing 27 ft. Then we observe a simple ratio:

$$\frac{4 \text{ ft}}{0.75 \text{ in}} = \frac{27 \text{ ft}}{x \text{ in}}, \qquad i.e., \qquad x \text{ in} = 27 \text{ ft} \frac{0.75 \text{ in}}{4 \text{ ft}} = 5.0625 \text{ in}.$$

□

3. Determine whether the following data support a proportionality argument for $y \propto z^{1/2}$. If so, estimate the slope.

y	3.5	5	6	7	8
z	3	6	9	12	15

Answer 1. Transformation via Square. First we have to determine whether or not y and $z^{1/2}$ are proportional, i.e., whether or not there is a positive constant k satisfying $y = kz^{1/2}$. If they are not, we don't have to proceed. For this purpose, we compute the ratio y^2/z, because y^2 and z are proportional if and only if y and $z^{1/2}$ are proportional. (One can compute the ratio $y/z^{1/2}$ or $z^{1/2}/y$ or z/y^2 or $(y/z^{1/2})^p$, where p is any real number except 0.)

From the following table, we are allowed to say that the given data can be approximated by

(y^2, z)	$(12.25, 3)$	$(25, 6)$	$(36, 9)$	$(49, 12)$	$(64, 15)$
y^2/z	4.0833	4.1667	4	4.0833	4.2667

$$\frac{y^2}{z} = 4.12, \qquad i.e., \qquad y^2 = 4.12z, \qquad i.e., \qquad y = \sqrt{4.12}z^{1/2} = 2.0298z^{1/2},$$

where 4.12 is the average of the ratio y_i^2/z_i. The figure 1 below supports our result.

□

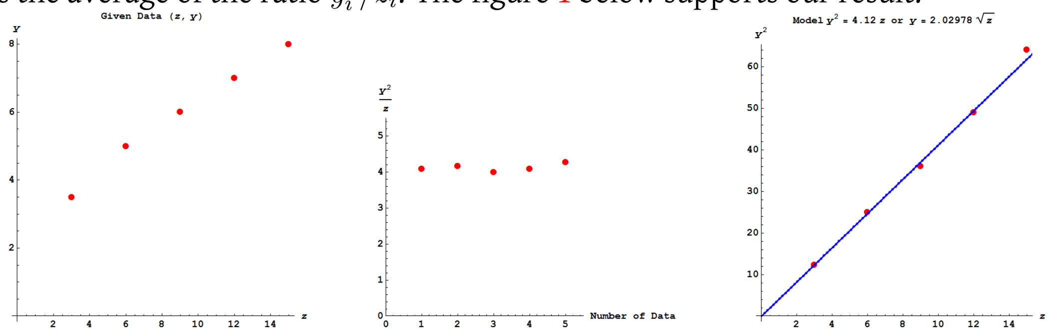

Figure 1: Data and Fitting Model $y = 2.0298z^{1/2}$

(y, z)	$(3.5, 3)$	$(5, 6)$	$(6, 9)$	$(7, 12)$	$(8, 15)$
$\ln y - (\ln z)/2$	0.7035	0.7136	0.6931	0.7035	0.7254

Answer 2. Transformation via Natural Logarithmic Function. We introduce another way to determine whether or not y and $z^{1/2}$ are proportional and to find the constant of proportionality. A simple observation shows that y and $z^{1/2}$ are proportional if and only if the difference between $\ln y$ and $\ln z^{1/2} = (\ln z)/2$ is constant.

From the following table,

the third row shows that $\ln y - (\ln z)/2$ is close to a constant 0.7078, which is the average of them. Hence, we can deduce

$$\ln y = \frac{1}{2}\ln z + 0.7078 \quad \Longrightarrow \quad y = e^{0.7078} z^{1/2} = 2.0295 z^{1/2}.$$

The figure 2 below supports our result. $\square$

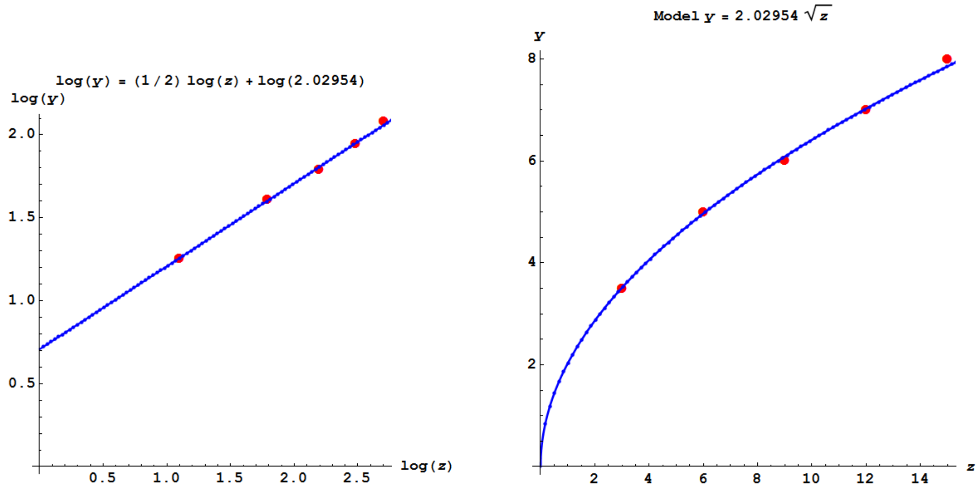

Figure 2: Fitting Model $y = 2.02954 z^{1/2}$

4. A new planet is discovered beyond Pluto at a mean distance to the sun of 4004 million miles. Using Kepler's third law, determine an estimate for the time T to travel around the sun in an orbit.

Answer. Kepler's Third Law says $T = cR^{3/2}$, where T is the period (days) and R is the mean distance to the sun. From the result of the Example 1 on page 67 of the textbook, we take $c = 0.410$. Then we have $T = 0.410(4004^{3/2}) = 103878$. $\square$

Section 2.3 Modeling Using Geometric Similarity

5. The stilt, a little long–legged bird, was described in *Gulliver's Travels* as weighing 4.5 ounces and having legs that are 8 in long. A flamingo has a similar shape and weighs 4 lb. Apply scaling arguments to show that flamingo's legs should be about 20 in long (as they actually are!).

Answer. Since a stilt and a flamingo have the similar shape, we can assume the geometric similarity between two species and moreover, we assume the weight is proportional to the cube of the leg length,

$$W \propto l^3 \qquad W = kl^3,$$

where l is the length of the leg and W is the weight and k is the constant of the proportionality. Then, the given information implies

$$k = \frac{W}{l^3} = \frac{4.5 \text{ oz}}{8^3 \text{ in}^3} = 0.00879 \text{ oz/in}^3.$$

Using the equation $W = (0.00879 \text{ oz/in}^3)\, l^3$, we deduce

$$\begin{aligned} & 4 \text{ lb} = (0.00879 \text{ oz/in}^3)\, l^3 \\ \Longrightarrow \qquad & l = \left(\frac{4 \text{ lb}}{0.00879 \text{ oz/in}^3}\right)^{1/3} = \left(\frac{64 \text{ oz}}{0.00879 \text{ oz/in}^3}\right)^{1/3} = 19.3826 \text{ in}, \end{aligned}$$

which is almost 20 in and here, 1 lb $= 16$ oz is used. $\square$

Section 3.1 Fitting Models to Data Graphically

6. In the following data, V represents a mean walking velocity and P represents the population size. *We wish to know if we can predict the population size P by observing how fast people walk.*

(1) Plot the data.

(2) What kind of a relationship is suggested?

(3) Test the given models by plotting the appropriate transformed data.

V	2.27	2.76	3.27	3.31	3.70	3.85	4.31	4.39	4.42	4.81	4.90
P	2500	365	23700	5491	14000	78200	70700	138000	304500	341948	49375
V	5.05	5.21	5.62	5.88							
P	260200	867023	1340000	1092759							

(6.1) $P = aV^b$.

Answer. (1) See the figure 3.
(2) Based on the plot of data, we may expect the exponential relationship such as $P = aV^b$.
(3) It is not easy to determine the parameters a and b such that the graph of $P = aV^b$ is close to the data. So we take the natural logarithmic function on the suggested model $P = aV^b$, $\ln P = \ln a + b \ln V$, which is the equation of a line on the $(\ln V)(\ln P)$–plane with the slope b and $(\ln P)$–intercept $(0, \ln a)$. To use this model, we need to transform the data $(\ln V_i, \ln P_i)$. We plot the transformed data and find a straight line which is close to data as many as possible. It seems that the slope $b = 9.55$ and the intercept $\ln a = -2.88$. That is, we deduce

$$\ln P = -2.88 + 9.55 \ln V \qquad \Longrightarrow \qquad P = e^{-2.88} V^{9.55} = 0.05613 V^{9.55}. \tag{6.1}$$

One may have other a and b as long as those values of the parameters are reasonable. □

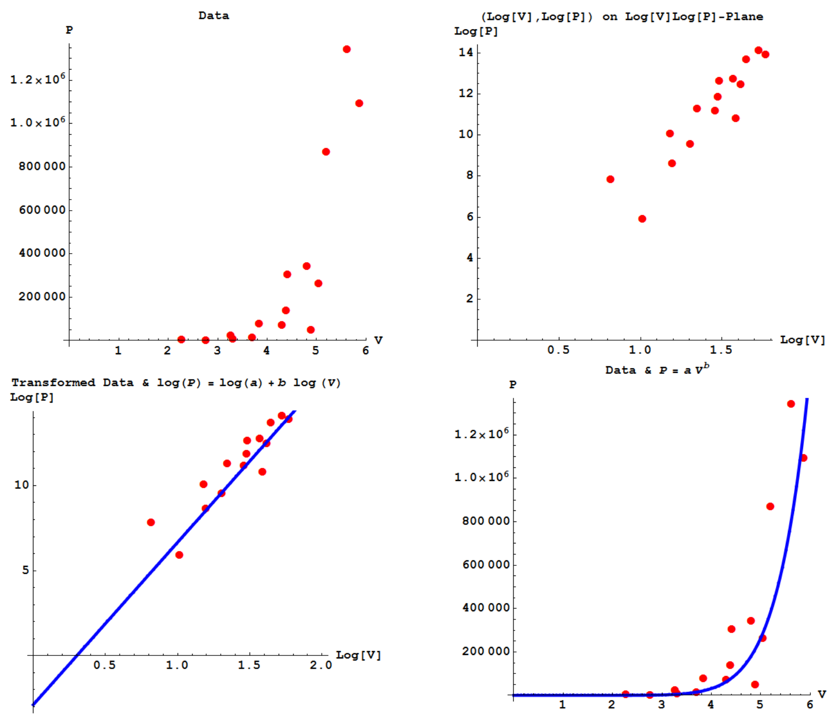

Figure 3: Fitting Model $P = 0.05613V^{9.55}$

(6.2) $P = a \ln V$.

Answer. The plot of transformed data $(\ln V, P)$ is almost similar to the plot of (V, P), which looks like the exponential graph. For this reason, the suggested model $P = a \ln V$ cannot be used to predict the population size P by the walking speed V. See the figure. See the figure 4. $\square$

(6.3) $P = a(b^V)$.

Answer. (1) See the figure 5.

(2) Based on the plot of data, we may expect the exponential relationship such as $P = ab^V$.

(3) Since the previous suggested model is not good, let us consider another exponential type of model. Transformation:

$$\ln P = V \ln b + \ln a,$$

which is a line on the V versus $\ln P$ plane. It has the slope $\ln b$ and the vertical intercept $(0, \ln a)$. Using two points of $(V, \ln P)$, we find the slope and the intercept. If we choose $(V, \ln P) = (4.31, \ln 70700)$ and $(5.88, \ln 1092759)$, we have

$$\begin{aligned} \ln P &= \frac{\ln 1092759 - \ln 70700}{5.88 - 4.31}(V - 4.31) + \ln 70700 = 1.7440V + 3.6497 \\ \Longrightarrow \quad P &= 38.4646(5.7199^V). \end{aligned} \tag{6.2}$$

$\square$

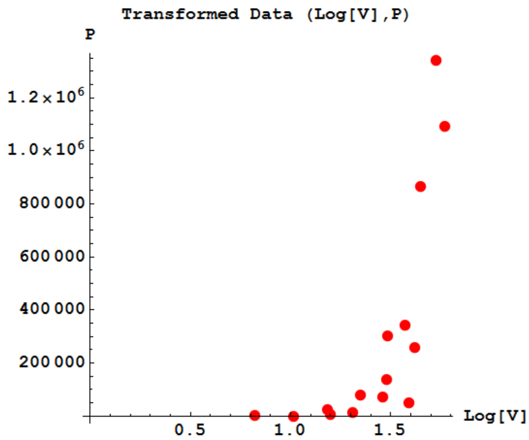

Figure 4: Fitting Model $P = a \ln V$

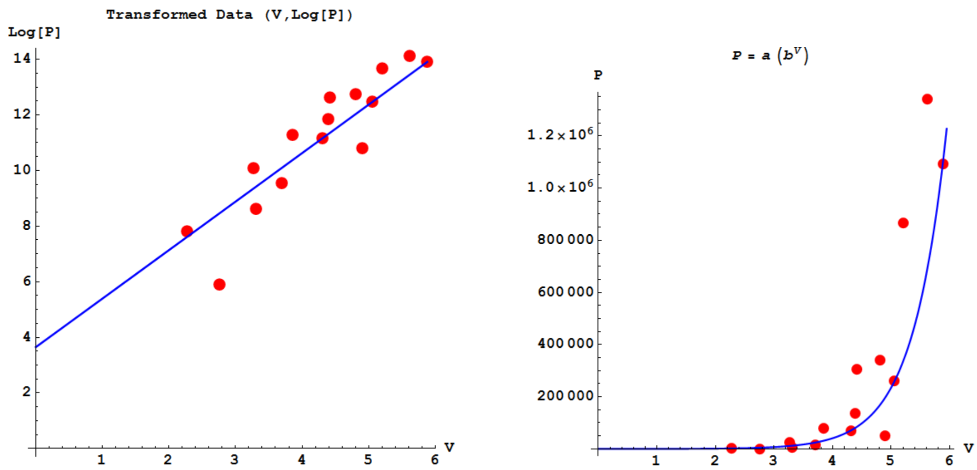

Figure 5: Fitting Model $P = 38.4646(5.7199^V)$

7. The following data represent the growth of a population (P) of fruit flies over a 6–week period.

(1) Test the following models by plotting an appropriate set of data.

(2) Estimate the parameters of the given model.

t (days)	7	14	21	28	35	42
P (population)	8	41	133	250	280	297

(7.1) $P = ct$

Answer. (1) It does not seem that the plot of the data have the linearity. So the model $P = ct$ is not appropriate for the given data.

(2) If we insist the linear model $P = ct$, then by choosing one of the data points, we can estimate c. When we choose the third $(t, P) = (21, 133)$, it seems that the line $P = 6.33t$ passing through the origin and looks close to all data. See the figure 6. □

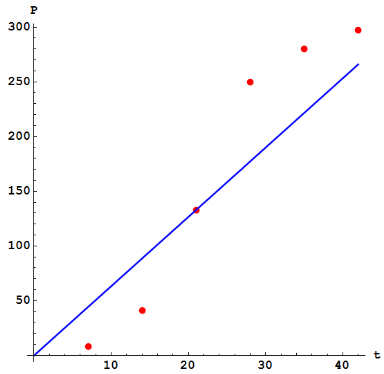

Figure 6: Model $P = 6.33t$

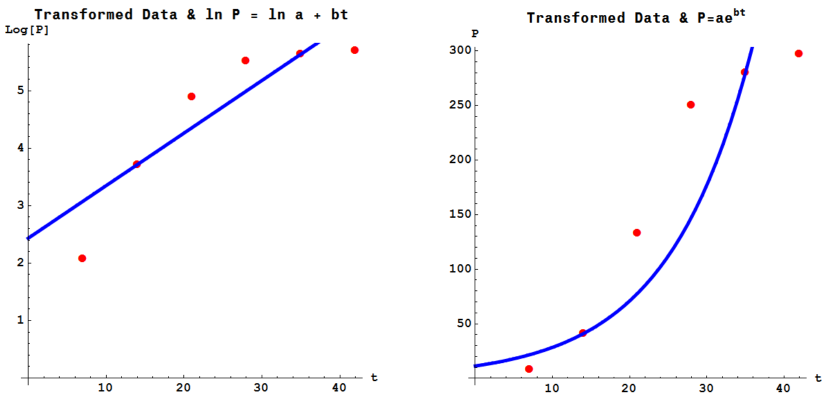

Figure 7: Model $P = 11.3903e^{0.0915t}$

(7.2) $P = ae^{bt}$

Answer. (1) The plot of the data looks like the exponential graph and so the suggested model is more appropriate. See the figure 7.

(2) We transform the data and also the model by taking the natural logarithmic function:

$$\ln P = \ln a + bt,$$

which is a line on the t versus $\ln P$ plane. The line has the slope b and the $(\ln P)$–intercept $(0, \ln a)$. We plot the data $(t, \ln P)$ and look for the line close to all the data. Using the second and the four data, $(t, \ln P) = (14, \ln 41)$ and $(t, \ln P) = (45, \ln 280)$,

$$\ln P = 0.0915t + 2.4328 \qquad \Longrightarrow \qquad P = 11.3903e^{0.0915t}. \tag{7.1}$$

□

• Aside: As we studied in Section 1.2 Approximating Change with Difference Equations, specifically, *Example 2 Growth of a Yeast Culture Revisited* on page 10 of the textbook, the population $p(t)$ follows the differential equation,

$$\frac{dp(t)}{dt} = (\mu - bp)\,p$$

where μ and b are constants. Solving the differential equation, we have

$$p(t) = \frac{\mu}{b + ce^{-\mu t}},$$

where c is another constant. Thus, this model should be better than both of them tested above. However, it is not easy to determine the constants/parameters μ, b and c. See the figure 8. □

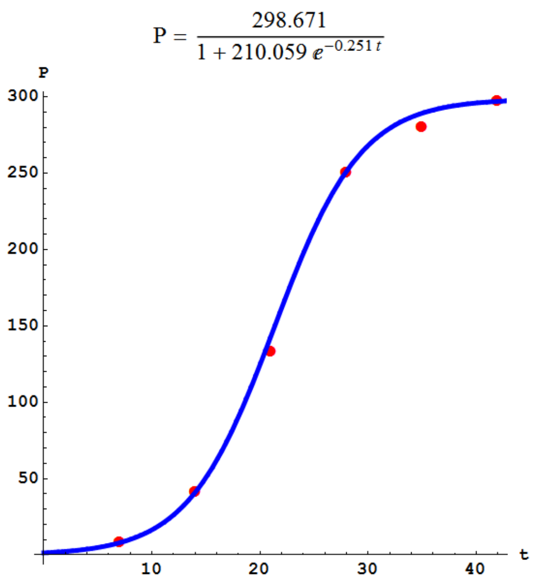

Figure 8: Model $P = \dfrac{298.671}{1 + 210.059e^{-0.251t}}$

8. The following data represent (hypothetical) energy consumption normalized to the year 1900. Plot the data. Test the model $Q = ae^{bx}$ by plotting the transformed data. Estimate the parameters of the model graphically.

x	Year	Consumption Q
0	1900	1.00
10	1910	2.01
20	1920	4.06
30	1930	8.17
40	1940	16.44
50	1950	33.12
60	1960	66.89
70	1970	134.29
80	1980	270.43
90	1990	544.57
100	2000	1096.63

Answer. When we transform the data and the model, $\ln Q = bx + \ln a$, we observe a straight line passing through the origin, because $(x, Q) = (0, 1)$ is transformed to $(x, \ln Q) = (0, 0)$. Using the origin and the last data $(x, \ln Q) = (100, 7)$, we have the equation of the line,

$$\ln Q = 0.07x \qquad \Longrightarrow \qquad Q = e^{0.07x}.$$

See the figure 9. $\square$

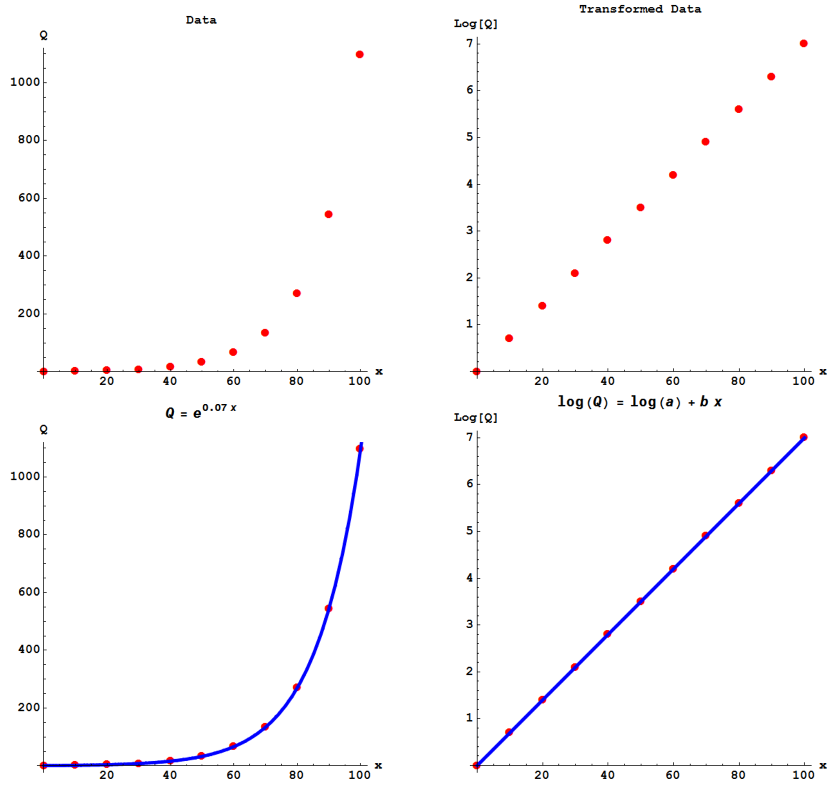

Figure 9: Model $Q = e^{0.07x}$

Section 3.2 Analytic Methods of Model Fitting

9. [1] Using elementary calculus, show that the minimum and maximum points for $y = f(x)$ occur among the minimum and maximum points for $y = f^2(x)$. Assuming $f(x) \geq 0$, why can we minimize $f(x)$ by minimizing $f^2(x)$?

Answer. Logic Behind the Strategy: Let A and B be the sets of all extremum points of $y = f(x)$ and $y = f^2(x)$, respectively. Then the problem says A is a subset of B. So we choose any element of A and show that it belongs to B.

Let (x_0, y_0) be an extremum of $y = f(x)$.

Step 1: From Calculus, $x = x_0$ must be a critical number and so the derivative of $y = f(x)$ at $x = x_0$ should vanish, i.e.,

$$f'(x_0) = 0 \qquad \Longrightarrow \qquad 2f(x_0)f'(x_0) = 0, \quad i.e., \quad (f^2)'(x_0) = 0,$$

implying that $x = x_0$ is a critical number of $y = f^2(x)$.

Step 2: We introduce the sign function $\mathrm{sgn}(x)$ defined by

$$\mathrm{sgn}(x) = \begin{cases} -1 & \text{for } x < 0 \\ 0 & \text{for } x = 0 \\ 1 & \text{for } x > 0. \end{cases}$$

It is straightforward to understand $\mathrm{sgn}(ab) = \mathrm{sgn}(a)\,\mathrm{sgn}(b)$ for any real numbers a and b.

We recall the First Derivative Test from Calculus saying that for a critical number $x = x_0$,

$$\mathrm{sgn}(f'(a))\,\mathrm{sgn}(f'(b)) < 0 \qquad \text{at any } a \text{ and } b \text{ near } x_0 \text{ with } a \leq x_0 \leq b \tag{9.1}$$

if and only if the critical number yields the extremum.

Since the critical number $x = x_0$ gives the extremum of $y = f(x)$, there *exists* a small interval $x_1 < x_0 < x_2$ such that for any $a \in (x_1, x_0)$ and $b \in (x_0, x_2)$ in the interval (x_1, x_2),

$$\mathrm{sgn}(f(a))\,\mathrm{sgn}(f(b)) > 0. \tag{9.2}$$

If there does not exist such an interval (x_1, x_2), then $x = x_0$ cannot give the extremum of $y = f(x)$. Why?

Proof. Suppose there exists an interval $x_1 < x_0 < x_2$ such that $\mathrm{sgn}(f(a))\,\mathrm{sgn}(f(b)) < 0$, specifically, $\mathrm{sgn}(f(a)) > 0$ and $\mathrm{sgn}(f(b)) < 0$, for any $a \in (x_1, x_0)$ and any $b \in (x_0, x_2)$.

Then it implies $f(x_0) = 0$. (Why?) So we have $f(b) < f(x_0) = 0 < f(a)$ for any $a \in (x_1, x_0)$ and any $b \in (x_0, x_2)$. This is a contradiction on the big assumption that $f(x_0)$ is a extremum. $\square$

In a nutshell, in general, if a function g has a critical number $x = c$, then

$$\begin{aligned} & g \text{ has an extremum at } x = c \\ \Longleftrightarrow \quad & \mathrm{sgn}(g'(a))\,\mathrm{sgn}(g'(b)) < 0 \text{ and } \mathrm{sgn}(g(a))\,\mathrm{sgn}(g(b)) > 0 \end{aligned}$$

for any $a \in (x_1, c)$ and $b \in (c, x_2)$ near $x = c$.

Step 3: Now we prove that $x = x_0$ yields the extremum of $y = f^2(x)$ by showing (9.1) and (9.2) on $y = f^2(x)$:

$$(f^2)'(x) = 2f(x)f'(x), \qquad \mathrm{sgn}((f^2)'(x)) = \mathrm{sgn}(2)\,\mathrm{sgn}(f(x))\,\mathrm{sgn}(f'(x)) = \mathrm{sgn}(f(x))\,\mathrm{sgn}(f'(x)).$$

(1) For any $a \in (x_1, x_0)$ and $b \in (x_0, x_2)$ in the small interval (x_1, x_2) of x_0,

$$\mathrm{sgn}((f^2)'(a))\,\mathrm{sgn}((f^2)'(b)) = \mathrm{sgn}(f(a))\,\mathrm{sgn}(f'(a))\,\mathrm{sgn}(f(b))\,\mathrm{sgn}(f'(b))$$

[1]Through the books or professors or by yourself, one may find a nicer and simpler proof of the given problem.

$$= \operatorname{sgn}(f(a))\operatorname{sgn}(f(b))\operatorname{sgn}(f'(a))\operatorname{sgn}(f'(b)) < 0$$

by the inequalities (9.1) and (9.2) on the extremum of $x = x_0$ for $y = f(x)$.

(2) For any $a \in (x_1, x_0)$ and $b \in (x_0, x_2)$ in the small interval (x_1, x_2) of x_0,

$$\operatorname{sgn}(f^2(a))\operatorname{sgn}(f^2(b)) = \operatorname{sgn}(f(a))\operatorname{sgn}(f(a))\operatorname{sgn}(f(b))\operatorname{sgn}(f(b)) = [\operatorname{sgn}(f(a))\operatorname{sgn}(f(b))]^2 > 0,$$

by the inequality (9.2) on the extremum of $x = x_0$ for $y = f(x)$.

Step 4 Conclusion: Thus, we can conclude $x = x_0$ belongs to B, i.e., A is a subset of B, i.e., any extremum of $y = f(x)$ is an extremum of $y = f^2(x)$. Specifically, if $y = f(x)$ has a (local) maximum at $x = c$, then $y = f^2(x)$ has also a (local) maximum at $x = c$. If $y = f(x)$ has a (local) minimum at $x = c$, then $y = f^2(x)$ has also a (local) minimum at $x = c$.

False Converse: Is the converse true? That is, if $y = f^2(x)$ has an extremum at $x = c$, then can $y = f(x)$ also have an extremum at $x = c$? No. We can explain it by giving a counterexample.

Counterexample) Let $f(x) = x$, i.e., $y = f^2(x) = x^2$. Then it has the minimum at $x = 0$. But $y = f(x) = x$ does not have any extremum.

True Converse: Suppose $f \geq 0$ for all x in the domain. Then those two sets A and B are same, i.e., the converse becomes true, i.e., if $y = f^2(x)$ has an extremum at $x = x_0$, then $y = f(x)$ also has an extremum at $x = x_0$.

Proof. Case 1. Let $x = x_0$ give an extremum of $y = f^2(x)$ and $f^2(x_0) = 0$. Then clearly it gives the absolute minimum value of $y = f^2(x)$, because $f^2(x_0) = 0 \leq f^2(x)$ for all x. Moreover, due to the assumption $f(x) \geq 0$ for all x, $x = x_0$ gives the absolute minimum value of $y = f(x)$, too, because of $f^2(x_0) = 0 = f(x_0) \leq f(x)$ for all x.

Case 2. Let $x = x_0$ give an extremum of $y = f^2(x)$ but $f^2(x_0) \neq 0$. Then since $x = x_0$ is the critical number of $y = f^2(x)$, so it is easy to see $f'(x_0) = 0$, i.e., $x = x_0$ is a critical number of $y = f(x)$:

$$0 = (f^2)'(x_0) = 2f(x_0)f'(x_0) \quad \text{and} \quad f(x_0) \neq 0 \qquad \Longrightarrow \qquad f'(x_0) = 0.$$

Since $y = f^2(x)$ has an extremum at $x = x_0$, it satisfies the inequalities (9.1) and (9.2), i.e., for any $a \in (x_1, x_0)$ and $b \in (x_0, x_2)$ near $x = x_0$ (where we choose the small interval (x_1, x_2) containing x_0 *not* to have any root of $f^2(x) = 0$):

$$\operatorname{sgn}((f^2)'(a))\operatorname{sgn}((f^2)'(b)) < 0 \qquad \text{and} \qquad \operatorname{sgn}(f^2(a))\operatorname{sgn}(f^2(b)) > 0.$$

When we expand them, we have

$$\begin{aligned} 0 &> \operatorname{sgn}((f^2)'(a))\operatorname{sgn}((f^2)'(b)) \\ &= \operatorname{sgn}(2f(a)f'(a))\operatorname{sgn}(2f(b)f'(b)) \\ &= \operatorname{sgn}(2)\operatorname{sgn}(f(a))\operatorname{sgn}(f'(a))\operatorname{sgn}(2)\operatorname{sgn}(f(b))\operatorname{sgn}(f'(b)) \\ &= \operatorname{sgn}(f(a))\operatorname{sgn}(f(b))\operatorname{sgn}(f'(a))\operatorname{sgn}(f'(b)). \end{aligned}$$

Since $f \geq 0$ for all x and the interval (x_1, x_2) does not have any root of $f^2(x) = 0$, i.e., $f \gneqq 0$ in the interval (x_1, x_2), so the result above becomes

$$\begin{aligned} 0 &> \operatorname{sgn}((f^2)'(a))\operatorname{sgn}((f^2)'(b)) \\ &= \operatorname{sgn}(f(a))\operatorname{sgn}(f(b))\operatorname{sgn}(f'(a))\operatorname{sgn}(f'(b)) \\ &= \operatorname{sgn}(f'(a))\operatorname{sgn}(f'(b)), \qquad \text{and} \qquad \operatorname{sgn}(f(a))\operatorname{sgn}(f(b)) > 0. \end{aligned}$$

Therefore, again by the inequalities (9.1) and (9.2) on $y = f(x)$, $y = f(x)$ has an extremum at $x = x_0$. □

□

10. For each of the following data sets,

1. formulate the mathematical model that minimizes the largest deviation between the data and the line $y = ax + b$.
2. If a computer is available, solve for the estimates of a and b.

(a)

x	1.0	2.3	3.7	4.7	6.1	7.0
y	3.6	3.0	3.2	5.1	5.3	6.8

Answer. *Part 1. Observing Plot:* From the plot of data points, we find a straight line $y = ax + b$ which is close to all the points. One can choose $(x, y) = (2.3, 3.0)$ and $(x, y) = (6.1, 5.3)$ to find the slope a and the y–intercept b:

$$y = 0.6053x + 1.6079. \tag{10.1}$$

It is fine that you use any other choice of pairs of points to find the slope a and the intercept b.

Part 2. Chebyshev (Optional): When we use the Chebyshev criterion via the computer,

$$y = 0.5333x + 2.1467 \tag{10.2}$$

minimizes the largest absolute deviations of the given data with the minimum value 0.92. See the two plots in the left–hand side of the figure 10. □

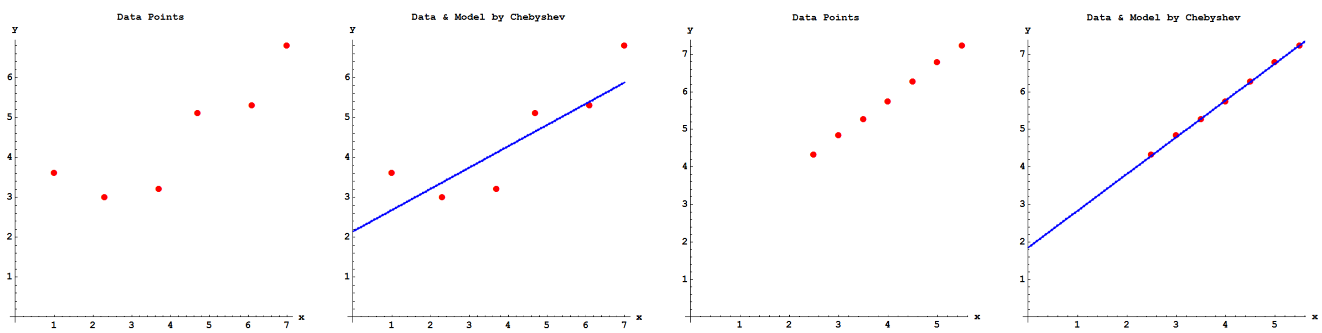

Figure 10: Model $y = ax + b$

(b)

x	2.5	3.0	3.5	4.0	4.5	5.0	5.5
y	4.32	4.83	5.27	5.74	6.26	6.79	7.23

Answer. *Part 1. Observing Plot:* From the plot of data points, we find a straight line $y = ax + b$ which is close to all the points. One can choose $(x, y) = (2.5, 4.32)$ and $(x, y) = (5.5, 7.23)$ to find the slope a and the y–intercept b:

$$y = 0.97x + 1.895. \tag{10.3}$$

It is fine that you use any other choice of pairs of points to find the slope a and the intercept b.

Part 2. Chebyshev (Optional): When we use the Chebyshev criterion via the computer,

$$y = 0.98x + 1.855. \tag{10.4}$$

minimizes the largest absolute deviations of the given data with the minimum value 0.035. See the two plots in the right–hand side of the figure 10. □

11. For the following data,

1. formulate the mathematical model that minimizes the largest deviation between the data and the model $y = c_1x^2 + c_2x + c_3$.
2. If a computer code is available, solve for the estimates of c_1, c_2 and c_3.

x	0.1	0.2	0.3	0.4	0.5
y	0.06	0.12	0.36	0.65	0.95

Answer. *Part 1. Observing Plot:* From the plot of data points, we find a quadratic curves $y = c_1x^2 + c_2x + c_3$ which is close to all the points. One can choose $(x, y) = (0.1, 0.06)$ and $(0.3, 0.36)$ and $(0.5, 0.95)$ to find the coefficients c_1, c_2 and c_3:

$$y = 3.625x^2 + 0.05x + 0.0188. \tag{11.1}$$

It is fine that you use any other choice of points to find the coefficients.

Part 2. Chebyshev (Optional): When we use the Chebyshev criterion via the computer,

$$y = 4x^2 - 0.0333x - 0.0050 \tag{11.2}$$

minimizes the largest absolute deviations of the given data with the minimum value 0.0283. See the figure 11. □

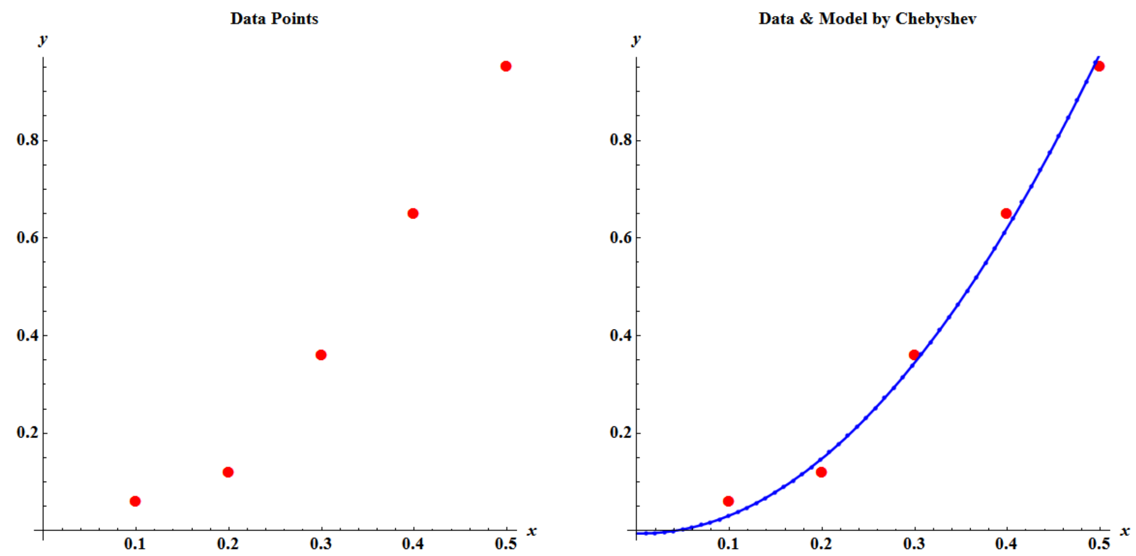

Figure 11: Model $y = c_1x^2 + c_2x + c_3$

12. For the following data,

1. formulate the mathematical model that minimizes the largest deviation between the data and the model $P = ae^{bt}$.
2. If a computer code is available, solve for the estimates of a and b.

t	7	14	21	28	35	42
P	8	41	133	250	280	297

Answer. Taking the natural logarithmic function on the suggested model $P = ae^{bt}$, it becomes

$$\ln P = \ln(ae^{bt}) = \ln a + \ln e^{bt} = \ln a + bt \ln e = \ln a + bt, \quad i.e., \quad \ln P = bt + \ln a.$$

Now we transform the data, only P values:

t	7	14	21	28	35	42
P	8	41	133	250	280	297
$\ln P$	2.0794	3.7136	4.8903	5.5215	5.6348	5.6937

Part 1. Observing Plot: From the plot of transformed data points, we find a straight line $\ln p = bt + \ln a$ which is close to all the points. One can choose $(t, \ln P) = (14, 3.71357)$ and $(t, \ln P) = (35, 5.63479)$ to find the slope b and the $(\ln P)$-intercept $\ln a$:

$$\ln P = 0.0915t + 2.4328, \qquad i.e., \qquad P = e^{2.4328}e^{0.0915t} = 11.3903e^{0.0915t}. \tag{12.1}$$

It is fine that you use any other choice of pairs of points to find the slope b and the intercept $\ln a$.

Part 2. Chebyshev (Optional): When we use the Chebyshev criterion via the computer,

$$\ln P = 0.1033t + 2.0387, \qquad i.e., \qquad P = e^{2.0387}e^{0.1033t} = 7.6810e^{0.1033t} \tag{12.2}$$

minimizes the largest absolute deviations of the given data with the minimum value 0.6822. See the figure 12. $\square$

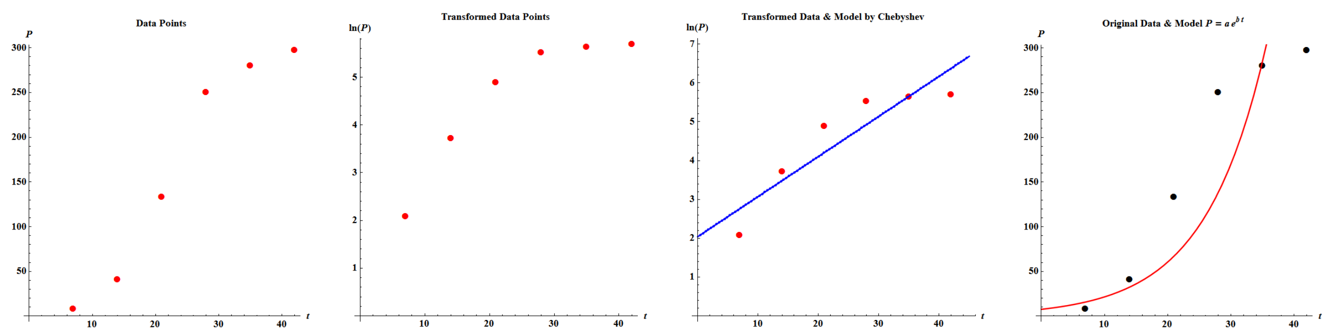

Figure 12: Model $P = ae^{bt}$

- Aside: Again, the data on the population can be approximated best by the model

$$P = \frac{\mu}{b + ce^{-\mu t}}.$$

However, it is not easy to estimate the parameters μ, b and c. See the Aside on page 7.

Section 3.3 Applying the Least–Squares Criterion

13. 1. Use the solutions of the normal equations coming from the least–squares criterion to estimate the coefficients of the line $y = ax + b$ such that the sum of the squared deviations between the line and the following data points is minimized.

For each problem,

2. compute $D = \sqrt{(\sum_{i=1}^{m} d_i^2)/m}$ and d_{max} to bound c_{max}, where $d_i = |y_i - ax_i - b|$ and $d_{max} = \max_i d_i$.

3. Compare the results to your solutions to Problem 10 in Section 3.2 on page 11.

(a)

x	1.0	2.3	3.7	4.7	6.1	7.0
y	3.6	3.0	3.2	5.1	5.3	6.8

Answer. *Part 1. Least–Squares:* For the suggested model $y = ax + b$, the Least–Squares criterion gives the normal equations,

$$a\|\boldsymbol{x}\|^2 + b\boldsymbol{x}\cdot\boldsymbol{i} = \boldsymbol{x}\cdot\boldsymbol{y} \quad \text{and} \quad a\boldsymbol{x}\cdot\boldsymbol{i} + b\|\boldsymbol{i}\|^2 = \boldsymbol{y}\cdot\boldsymbol{i},$$

where $\boldsymbol{x} = \langle x_1, x_2, \ldots, x_m\rangle$ and $\boldsymbol{y} = \langle y_1, y_2, \ldots, y_m\rangle$ and $\boldsymbol{i} = \langle 1, 1, \ldots, 1\rangle$ and m is the number of the given data. Solving the equations for a and b, we have

$$a = \frac{1}{z}\left[(\boldsymbol{x}\cdot\boldsymbol{y})\|\boldsymbol{i}\|^2 - (\boldsymbol{x}\cdot\boldsymbol{i})(\boldsymbol{y}\cdot\boldsymbol{i})\right], \qquad b = \frac{1}{z}\left[(\boldsymbol{y}\cdot\boldsymbol{i})\|\boldsymbol{x}\|^2 - (\boldsymbol{x}\cdot\boldsymbol{y})(\boldsymbol{x}\cdot\boldsymbol{i})\right], \tag{13.1}$$

where $z = \|\boldsymbol{x}\|^2\|\boldsymbol{i}\|^2 - (\boldsymbol{x}\cdot\boldsymbol{i})^2$.

Putting $\boldsymbol{x} = \langle 1.0, 2.3, 3.7, 4.7, 6.1, 7.0\rangle$ and $\boldsymbol{y} = \langle 3.6, 3.0, 3.2, 5.1, 5.3, 6.8\rangle$ and $\boldsymbol{i} = \langle 1, 1, 1, 1, 1, 1\rangle$ into the formula (13.1) on a and b above, we have

$$a = 0.5680, \qquad b = 2.1522$$

and so the suggested model is in fact $y = 0.5680x + 2.1522$.

Part 2. Bounds: A computation gives the table:

x	1.0	2.3	3.7	4.7	6.1	7.0
y	3.6	3.0	3.2	5.1	5.3	6.8
d_i	0.8798	0.4586	1.0539	0.2781	0.3171	0.6717

where $d_i = |y_i - 0.5680x_i - 2.1522|$. From the table, the maximum value is $d_{max} = 1.0539$ and $D = 0.6738$.

$$D = 0.6738 \le c_{max} \le 1.0539 = d_{max}.$$

Part 3. Comparison with Problem 10: Using the result (10.2), $y = 0.5333x - 2.1467$ (One can use the line (10.1) by observing the plot), we have the table:

x	1.0	2.3	3.7	4.7	6.1	7.0
y	3.6	3.0	3.2	5.1	5.3	6.8
c_i	0.92	0.3733	0.92	0.4467	0.1	0.92

where $c_i = |y_i - 0.5333x_i - 2.1467|$. From the table, the maximum value is $c_{max} = 0.92$. Hence,

$$D = 0.6738 \le c_{max} = 0.92 \le 1.0539 = d_{max}.$$

See the left one of the figure 13. □

(b)

x	2.5	3.0	3.5	4.0	4.5	5.0	5.5
y	4.32	4.83	5.27	5.74	6.26	6.79	7.23

Answer. *Part 1. Least-Squares:* Since we have the same type of model as above, we use the formula (13.1) in the solution above. So putting the vectors coming from the given data,

$$\boldsymbol{x} = \langle 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5 \rangle, \qquad \boldsymbol{y} = \langle 4.32, 4.83, 5.27, 5.74, 6.26, 6.79, 7.23 \rangle,$$

and $\boldsymbol{i} = \langle 1, 1, 1, 1, 1, 1, 1 \rangle$, into the formula (13.1), we have

$$a = 0.9743, \qquad b = 1.88$$

and so the suggested model is in fact $y = 0.9743x + 1.88$.

Part 2. Bounds: A computation gives the table:

x	2.5	3.0	3.5	4.0	4.5	5.0	5.5
y	4.32	4.83	5.27	5.74	6.26	6.79	7.23
d_i	0.0043	0.0271	0.02	0.0371	0.0043	0.0386	0.0086

where $d_i = |y_i - 0.9743x_i - 1.88|$. From the table, the maximum value is $d_{max} = 0.0386$ and $D = 0.0242$.

$$D = 0.0242 \leq c_{max} \leq 0.0386 = d_{max}.$$

Part 3. Comparison with Problem 10: Using the result (10.4), $y = 0.98x + 1.855$ (One can use the line (10.3) by observing the plot), we have the table:

x	2.5	3.0	3.5	4.0	4.5	5.0	5.5
y	4.32	4.83	5.27	5.74	6.26	6.79	7.23
c_i	0.015	0.035	0.015	0.035	0.005	0.035	0.015

where $c_i = |y_i - 0.98x_i - 1.855|$. From the table, the maximum value is $c_{max} = 0.035$. Hence,

$$D = 0.0242 \leq c_{max} = 0.035 \leq 0.0386 = d_{max}.$$

See the right one of the figure 13. □

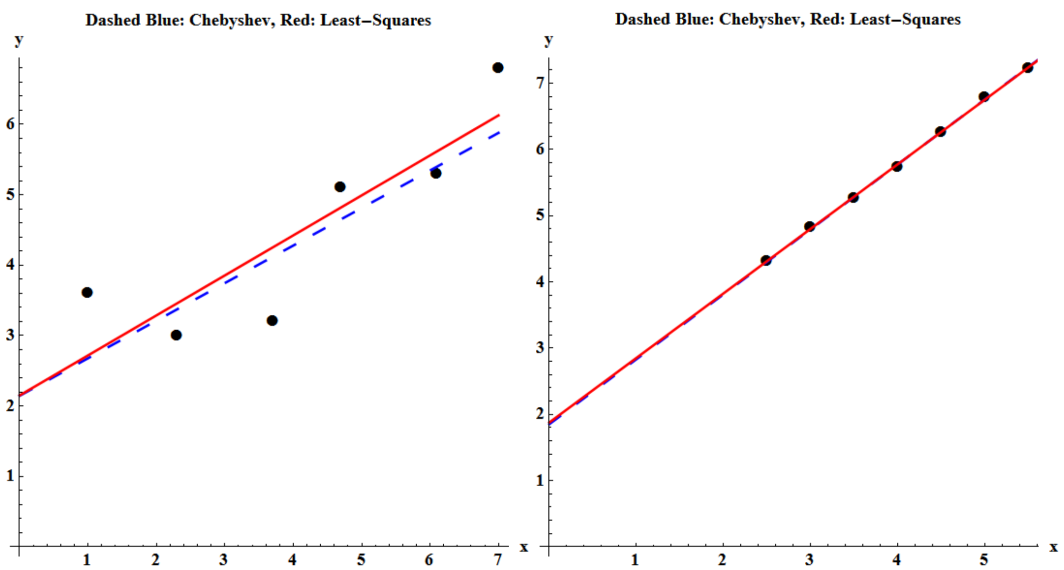

Figure 13: Model $y = ax + b$

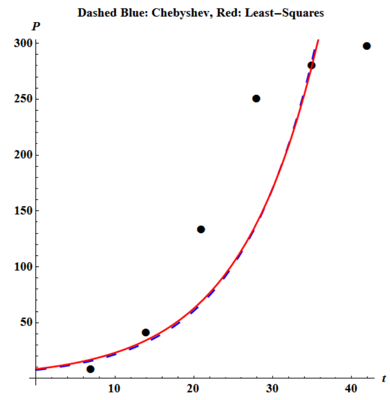

Figure 14: Model $P = ae^{bt}$

14. 1. Make an appropriate transformation to fit the model $P = ae^{bt}$ using the normal equations coming from the least–squares criterion.

2. Estimate a and b.

3. Compute $D = \sqrt{(\sum_{i=1}^{m} d_i^2)/m}$ and d_{max} to bound c_{max}, where $d_i = |y_i - ax_i - b|$ and $d_{max} = \max_i d_i$.

4. Compare the results to your solutions to Problem 12 in Section 3.2 on page 13.

t	7	14	21	28	35	42
P	8	41	133	250	280	297

Answer. *Part 1. Least–Squares:* For the suggested model $P = ae^{bt}$, we transform the data and the model (resulting in $\ln P = bt + \ln a$) as we already did in the Problem 12, Section 3.2 on page 13. To the transformed data, we apply the Least–Squares criterion, which gives the normal equations,

$$b\|\boldsymbol{X}\|^2 + (\ln a)\boldsymbol{X}\cdot\boldsymbol{i} = \boldsymbol{X}\cdot\boldsymbol{Y} \quad \text{and} \quad b\boldsymbol{X}\cdot\boldsymbol{i} + (\ln a)\|\boldsymbol{i}\|^2 = \boldsymbol{Y}\cdot\boldsymbol{i},$$

where $\boldsymbol{X} = \langle t_1, t_2, \ldots, t_m\rangle$ and $\boldsymbol{y} = \langle \ln P_1, \ln P_2, \ldots, \ln P_m\rangle$ and $\boldsymbol{i} = \langle 1, 1, \ldots, 1\rangle$ and m is the number of the given data. Solving the equations for b and $\ln a$, we have

$$b = \frac{1}{Z}\left[(\boldsymbol{X}\cdot\boldsymbol{Y})\|\boldsymbol{i}\|^2 - (\boldsymbol{X}\cdot\boldsymbol{i})(\boldsymbol{Y}\cdot\boldsymbol{i})\right], \qquad \ln a = \frac{1}{Z}\left[(\boldsymbol{Y}\cdot\boldsymbol{i})\|\boldsymbol{X}\|^2 - (\boldsymbol{X}\cdot\boldsymbol{Y})(\boldsymbol{X}\cdot\boldsymbol{i})\right], \tag{14.1}$$

where $Z = \|\boldsymbol{X}\|^2\|\boldsymbol{i}\|^2 - (\boldsymbol{X}\cdot\boldsymbol{i})^2$.

Putting $\boldsymbol{x} = \langle 7, 14, 21, 28, 35, 42\rangle$ and $\boldsymbol{y} = \langle \ln 8, \ln 41, \ln 133, \ln 250, \ln 280, \ln 297\rangle$ and $\boldsymbol{i} =< 1, 1, 1, 1, 1, 1 >$ into the formula (14.1) on b and $\ln a$ above, we have

$$b = 0.0999, \qquad \ln a = 2.1423.$$

and so the suggested model is in fact $\ln P = 0.0999t + 2.1423$, i.e., $P = e^{2.1423}e^{0.0999t} = 8.5188e^{0.0999t}$.

Part 2. Bounds: A computation gives the table:

(t, P)	$(7, 8)$	$(14, 41)$	$(21, 133)$	$(28, 250)$	$(35, 280)$	$(42, 297)$
d_i	9.1381	6.5214	63.6356	110.452	0.7440	267.803

where $d_i = |P_i - 8.5188e^{0.0999t_i}|$. From the table, the maximum value is $d_{max} = 267.803$ and $D = 121.171$.

$$D = 121.171 \le c_{max} \le 267.803 = d_{max}.$$

Part 3. Comparison with Problem 12: Using the result (12.2), $P = 7.6810e^{0.1033t}$ (One can use the line (12.1) by observing the plot), we have the table:

(t, P)	$(7, 8)$	$(14, 41)$	$(21, 133)$	$(28, 250)$	$(35, 280)$	$(42, 297)$
c_i	7.8252	8.3951	65.8236	111.5956	5.1564	290.5115

where $c_i = |P_i - 7.6810e^{0.1033t_i}|$. From the table, the maximum value is $c_{max} = 290.5115$.

$$D = 121.171 \le c_{max} = 290.5115 \le 267.803 = d_{max}.$$

See the left one of the figure 14. $\square$

Section 3.4 Choosing a Best Model

15. For each of the following problems, answer to the following questions.

1. Find a model using the least–squares criterion either on the data or on the transformed data (as appropriate).
2. Compare your results with the graphical fits obtained in the problem set 3.1 by computing the deviations, maximum absolute deviations and the sum of the squared deviations for each model.
3. Find a bound on c_{max} if the model was fit using the least–squares criterion.

(15.1) Problem (6.1) and (6.3) in Section 3.1 on page 4.

Answer. • Problem (6.1) $P = aV^b$: 1. For the transformed data $(\ln V, \ln P)$, the least–squares criterion gives

$$\ln P = 8.0063 \ln V - 0.2266 \qquad \Longrightarrow \qquad P = 0.7973V^{8.0063}. \tag{15.1}$$

2. Let f_G and f_L be the models obtained by the graphical fits in Section 3.1 and the one obtained by the least–squares criterion, respectively. Then, by the results (6.1) and (15.1),

$$P = f_G(V) = 0.0561V^{9.55}, \qquad P = f_L(V) = 0.7973V^{8.0063}.$$

	$\max\|P_i - f(V_i)\|$	$\sum_i (P_i - f(V_i))^2$
Model from (6.1)	528639	6.4005×10^{11}
Least–Squares (15.1)	537950	5.8242×10^{11}

That is, the least–squares result is *better* than the graphical fit in the sense of minimizing the sum of squares of deviations.

3. Bounds.

$$D = 197049 \le c_{max} \le 537950 = d_{max}.$$

• Problem (6.3) $P = a(b^V)$: 1. For the transformed data $(V, \ln P)$, the least–squares criterion gives

$$\ln P = 2.0691V + 2.3097 \qquad \Longrightarrow \qquad P = 10.0717\left(7.9180^V\right). \tag{15.2}$$

2. Let f_G and f_L be the models obtained by the graphical fits in Section 3.1 and the one obtained by the least–squares criterion, respectively. Then, by the results (6.2) and (15.2),

$$P = f_G(V) = 38.4646\left(5.7199^V\right), \qquad P = f_L(V) = 10.0717\left(7.9180^V\right).$$

	$\max\|P_i - f(V_i)\|$	$\sum_i (P_i - f(V_i))^2$
Model from (6.2)	645613	8.0043×10^{11}
Least–Squares (15.2)	843499	1.0182×10^{12}

That is, the least–squares result is *worse* than the graphical fit in the sense of minimizing the sum of squares of deviations.

3. Bounds.

$$D = 260534 \le c_{max} \le 843499 = d_{max}.$$

□

(15.2) Problem 7.2 in Section 3.1 on page 6.

Answer. 1. For the transformed data $(t, \ln P)$, the least–squares criterion gives

$$\ln P = 0.0999t + 2.1423 \qquad \Longrightarrow \qquad P = 8.5187e^{0.0999t} \tag{15.3}$$

2. Let f_G and f_L be the models obtained by the graphical fits in Section 3.1 and the one obtained by the least–squares criterion, respectively. Then, by the results (7.1) and (15.3),

$$P = f_G(t) = 11.3903e^{0.0915t}, \qquad P = f_L(t) = 8.5187e^{0.0999t}.$$

	$\max \lvert P_i - f(t_i) \rvert$	$\sum_i (P_i - f(t_i))^2$
Model from (7.1)	234.531	68714.3
Least–Squares (15.3)	268.7	88536.4

That is, the least–squares result is *worse* than the graphical fit in any sense. How can this happen? See the Aside following the answer.

3. Bounds.

$$D = 121.475 \leq c_{max} \leq 268.7 = d_{max}.$$

See the figure 15. □

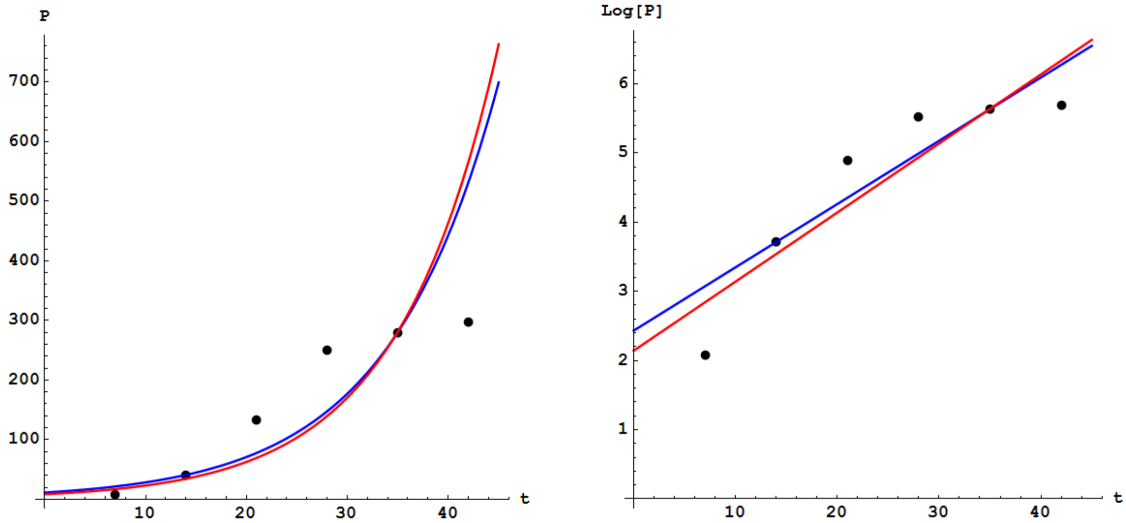

Figure 15: Blue: f_G (Graphical Fit), Red: f_L (Least-Squares Fit)

• Aside In the problem above, we have the case that the least–square fit is worse than the graphical fit. How can this happen? Let us redescribe what we have done.

1. Graphical Fit: We chose two points $(t, P) = (14, 41)$ and $(35, 280)$. Using these two points, we found a and b in the model $P = ae^{bt}$. Even if we find $\ln a$ and b in the transformed data and transformed model, the values of a and b are same. That is, before and after the transformation, the values of a and b do not change and the graphical fit model passes through two points, the second and the fifth.

2. Least–Squares Fit: We used the least–squares fit on the original data (t, P) and got $a = 8.51901$ and $b = 0.0999$ via a computer. Even if we use the least–squares fit on the transformed data $(t, \ln P)$, we have the same parameters. But the resulting model is worse than the graphical fit. Why? When we look for a line, $\ln P = bt + \ln a$, the least–squares fit gives the best a and b. However those best constants *become worse via the backward transformation*, i.e., $P = ae^{bt}$, because those constants were chosen to be THE BEST FOR THE STRAIGHT LINE rather than a curve. As a proof, look at the table:

t $\ln P$	7 $\ln 8$	14 $\ln 41$	21 $\ln 133$	28 $\ln 250$	35 $\ln 280$	42 $\ln 297$
$\lvert \ln P_i - 0.0915t_i - 2.43276 \rvert$ (Graphical)	0.9937	0	0.5364	0.5271	0	0.5815
$\lvert \ln P_i - 0.0999t_i - 2.14227 \rvert$ (Least–Square)	0.7619	0.1732	0.6510	0.5831	0.0027	0.6427

	$\max \lvert \ln P_i - bt_i - \ln a \rvert$	$\sum_i (\ln P_i - bt_i - \ln a)^2$
Graphical Fit on Transformed Data	0.9937	1.8911
Least–Squares Fit on Transformed Data	0.7619	1.7873

That is, on the transformed data, the Least–Squares Fit is better in any sense. See the one in the right–hand side of the figure 15.