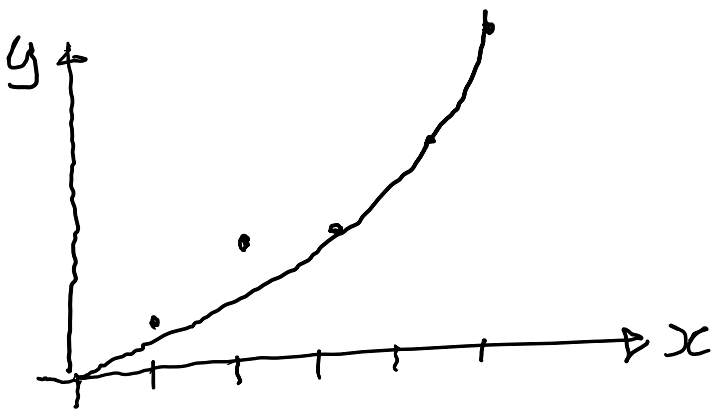
Example

x_i	0	1	2	3	4	5
y_i	2.1	7.7	13.6	27.2	40.9	61.1

Fit a second order polynomial to the data.

Solution:



$$\boxed{y = a_0 + a_1 x + a_2 x^2}$$

$$\left\{\begin{array}{ccccc} n a_0 & \left(\sum x_i\right) a_1 & \left(\sum x_i^2\right) a_2 & = & \sum y_i \\ \left(\sum x_i\right) a_0 & \left(\sum x_i^2\right) a_1 & \left(\sum x_i^3\right) a_2 & = & \sum y_i x_i \\ \left(\sum x_i^2\right) a_0 & \left(\sum x_i^3\right) a_1 & \left(\sum x_i^4\right) a_2 & = & \sum y_i x_i^2 \end{array}\right.$$

$$n = 6$$

$$\sum x_i = 15$$
$$\sum x_i^2 = 55$$
$$\sum x_i^3 = 225$$
$$\sum x_i^4 = 979$$

$$\sum y_i = 152.6$$
$$\sum y_i x_i = 585.6$$
$$\sum y_i x_i^2 = 2488.8$$

The linear equations:

$$\begin{pmatrix} 6 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{pmatrix} \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \end{Bmatrix} = \begin{Bmatrix} 152.6 \\ 585.6 \\ 2488.8 \end{Bmatrix}$$

Then:

$a_0 = 2.47857$
$a_1 = 2.35929$
$a_2 = 1.86071$

$$\boxed{\therefore y = 2.47857 + 2.35929x + 1.86071x^2}$$

Since

$$\bar{y} = \frac{\sum y_i}{n} = \frac{152.6}{6} = 25.433$$
$$S_y = \sum (y_i - y)^2 = 2513.39$$
$$S_r = \sum \left(y_i - a_0 - a_1 x_i - a_2 x_i^2\right)^2 = 3.74657$$

Standard error:

$$s_{y|x} = \sqrt{\frac{S_r}{n-(m+1)}} = \sqrt{\frac{3.74657}{6-(2+1)}} = 1.12$$

The coefficient of determination:

$$r^2 = \frac{S_y - S_r}{S_y}$$
$$r^2 = \frac{2513.39 - 3.74657}{2513.39}$$
$$r^2 = 0.99851$$

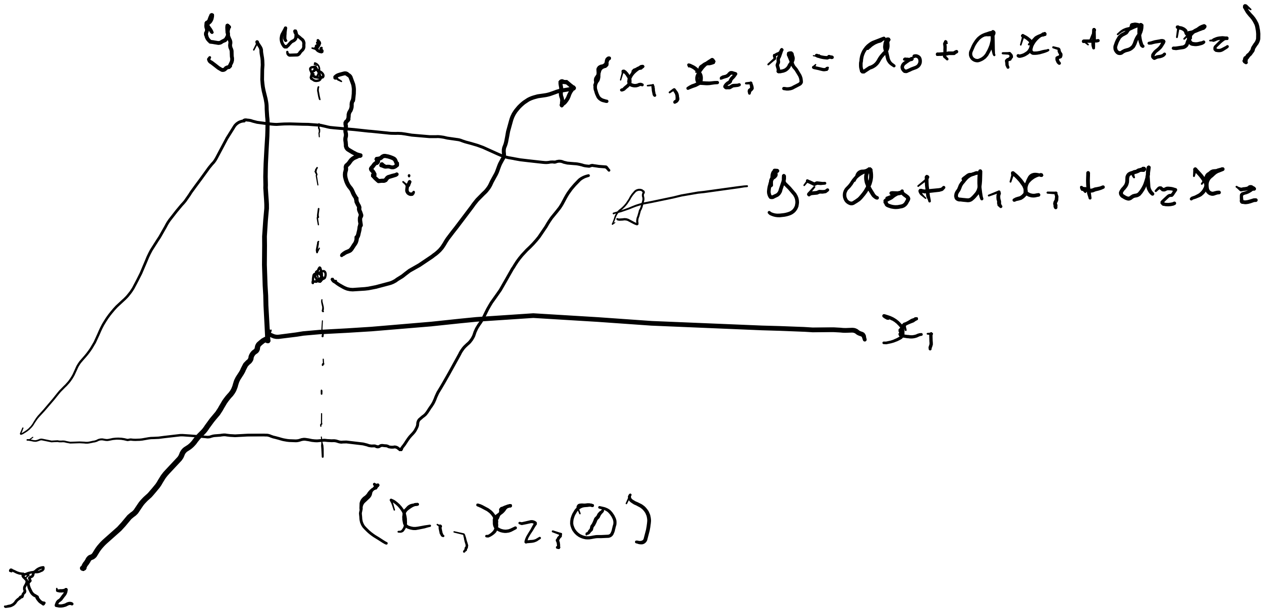
What do you do if you have a polynomial? It's the same procedure:

$$y = a_0 + a_1 x^2 + a_2 x^2 + \cdots + a_m x^m + e$$

$m + 1$ unknown: a_0 a_1 ... a_m

Multiple linear regression

$$y = a_0 + a_1 x_1 + a_2 x_2 + e$$



Given data:

$$\begin{matrix} (x_{11} & x_{21} & y_1) \\ (x_{12} & x_{22} & y_2) \\ & \cdots & \\ (x_{1n} & x_{2n} & y_n) \end{matrix}$$

Consider:

$$S_r = \sum e_i^2 = \sum (y_i - a_0 - a_1 x_{1i} - a_2 x_{2i})^2$$
$$S_r = S_r(a_0, a_1, a_2)$$

$$\frac{\delta S_r}{\delta a_0} = \sum 2(y_i - a_0 - a_1 x_{1i} - a_2 x_{2i}) \cdot (-1) = 0$$
$$\frac{\delta S_r}{\delta a_1} = \sum 2(y_i - a_0 - a_1 x_{1i} - a_2 x_{2i}) \cdot (-x_{1i}) = 0$$
$$\frac{\delta S_r}{\delta a_2} = \sum 2(y_i - a_0 - a_1 x_{1i} - a_2 x_{2i}) \cdot (-x_{2i}) = 0$$

$$\sum a_0 + \sum a_1 x_{1i} + \sum a_2 x_{2i} = \sum y_i$$

$$\begin{matrix} n a_0 & (\sum x_{1i}) a_1 & (\sum x_{2i}) a_2 & = & \sum y_i \\ (\sum x_{1i}) a_0 & (\sum x_{1i}^2) a_1 & (\sum x_{1i} x_{2i}) a_2 & = & \sum x_{1i} y_i \\ (\sum x_{2i}) a_0 & (\sum x_{1i} x_{2i}) a_1 & (\sum x_{2i}^2) a_2 & = & \sum x_{2i} y_i \end{matrix}$$

$$y = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_m x_m + e$$

Each data point:

$$x_{1i}, x_{2i} \dots x_{mi}, y_i \; (i = 1,2, \dots, n)$$
$$S_r = \sum e_i^2 = \sum (y_i - a_0 - a_1 x_{1i} - \cdots - a_m x_{mi})^2$$

Standard error:

$$s_{y|x} = \sqrt{\frac{S_r}{n - (m+1)}}$$

General linear least-squares

$$\boxed{y = a_0 z_0 + a_1 z_1 + \cdots + a_m z_m + e}$$

z_0, z_1, z_m: the basis functions

In the multiple linear regression:

$$z_0 = 1,\ z_1 = x_1,\ z_2 = x_2,\ z_m = x_m$$

Polynomial regression:

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m + e$$
$$z_0 = 1, z_1 = x, z_2 = x^2, \dots, z_m = x^m$$

For example:

$$z_0 = 1, z_1 = \cos \omega t\,, z_2 = \sin \omega t$$
$$\boxed{y = a_0 + a_1 \cos \omega t + a_2 \sin \omega t}$$

Note: this is the first three terms of the Fourier expansion.

Fort the sample point

$$\underbrace{z_{0i}, z_{1i}, \dots, z_{mi}, y_i}_{\text{data}} \; (i = 1,2, \dots, n)$$

$$e_i = y_i - a_0 z_{0i} - a_1 z_{1i} - \cdots - a_m z_{mi}$$
$$i = 1, 2, \dots, n$$

$$S_r = \sum e_i^2$$

$$\frac{\delta S_r}{\delta a_0} = 0, \quad \frac{\delta S_r}{\delta a_1} = 0, \qquad \dots \qquad \frac{\delta S_r}{\delta a_m} = 0$$

In the matrix form:

$$\boxed{[Z]^T [Z] \{A\} = [Z]^T \{Y\}}$$

Here

$$\{A\} = \begin{Bmatrix} a_0 \\ a_1 \\ \dots \\ a_m \end{Bmatrix} \qquad \{Y\} = \begin{Bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{Bmatrix}$$

$$[Z] = \begin{bmatrix} z_{01} & z_{11} & z_{21} & \cdots & z_{m1} \\ z_{02} & z_{12} & z_{22} & \cdots & z_{m2} \\ z_{03} & z_{13} & z_{23} & \cdots & z_{m3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ z_{0n} & z_{1n} & z_{2n} & \cdots & z_{mn} \end{bmatrix}_{n\times(m+1)} \quad \text{Where } n > m+1$$

$[Z]$ is a tall matrix

To solve the final linear equations,
LU decomposition
Cholesky's method

$$\{A\} = ([Z]^T[Z])^{-1}[Z]^T\{Y\}$$

Let

$$([Z]^T[Z])^{-1} = \begin{bmatrix} z_{11}^{-1} & z_{12}^{-1} & \cdots & z_{1,m+1}^{-1} \\ z_{12}^{-1} & z_{22}^{-1} & \cdots & z_{2,m+1}^{-1} \\ \cdots & \cdots & \cdots & \cdots \\ z_{m+1,1}^{-1} & z_{m+1,2}^{-1} & \cdots & z_{m+1,m+1}^{-1} \end{bmatrix}$$

The diagonal of the matrix:
${z_{ii}^{-1}}$: The variance of a_{i-1} $(i = 1, 2, \dots, m+1)$

The off-diagonal of the matrix (basically, not the diagonal):
z_{ij}^{-1}: The covariance of a_{i-1} and a_{j-1}

$$var(a_{i-1}) = z_{ii}^{-1} s_{y|x}^2$$
$$cov(a_{i-1}, a_{j-1}) = z_{ij}^{-1} s_{y|x}^2$$
$$s_{y|x} = \sqrt{\frac{S_r^2}{n-(m+1)}}$$

For one independent variable, the linear regression:

$$\boxed{y = a_0 + a_1 x + e}$$

The lower and upper bounds of a_0:

$$L = a_0 - t_{\alpha/2,n-2} \cdot s(a_0)$$
$$U = a_0 + t_{\alpha/2,n-2} \cdot s(a_0)$$

The lower and upper bounds of a_1:

$$L = a_1 - t_{\alpha/2,n-2} \cdot s(a_1)$$
$$U = a_1 - t_{\alpha/2,n-2} \cdot s(a_1)$$

$t_{\alpha/2,n}$: $\frac{the\ student\ distribution}{two\ sided\ interval}$

$s(a_i)$ = the standard error of the coefficient a_i

$s(a_i) = \sqrt{var\ (a_i)} \quad (i = 0, 1)$

(x) (y)

Time, s	Measured v, m/s (a)	Model-calculated v, m/s (b)
1	10.00	8.953
2	16.30	16.405
3	23.00	22.607
4	27.50	27.769
5	31.00	32.065
6	35.60	35.641
7	39.00	38.617
8	41.50	41.095
9	42.90	43.156
10	45.00	44.872
11	46.00	46.301
12	45.50	47.490
13	46.00	48.479
14	49.00	49.303
15	50.00	49.988

measure

$$\boxed{y = a_0 + a_1 x + e}$$

model

Since

$$y = a_0 + a_1 x + e$$

$$\left\{\begin{aligned} y_1 &= a_0 + a_1 x_1 + e_1 \\ y_2 &= a_0 + a_1 x_2 + e_2 \\ &\cdots \\ y_n &= a_0 + a_1 x_n + e_n \end{aligned}\right.$$

"eliminated" through standard least square procedure.

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} = \begin{Bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{Bmatrix}$$

$$[Z] = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{bmatrix} \quad \{A\} = \begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} \quad \{Y\} = \begin{Bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{Bmatrix}$$

$$[Z]\{A\} = \{Y\}$$

$$[Z]^T[Z]\{A\} = [Z]^T\{Y\}$$

$$\begin{bmatrix} 15 & 548.3 \\ 548.3 & 22191.21 \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} = \begin{Bmatrix} 552.74 \\ 22421.43 \end{Bmatrix}$$

$$y = a_0 + a_1 x + e$$

Model testing

$$\Downarrow$$

$$[Z]^T[Z]\{A\} = [Z]^T\{y\}$$

$$\begin{bmatrix} 15 & 548.3 \\ 548.3 & 22191.21 \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} = \begin{Bmatrix} 552.74 \\ 22421.43 \end{Bmatrix}$$

$$\begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} = \left[\underbrace{\begin{bmatrix} 0.688414 & -0.01701 \\ -0.01701 & 0.000405 \end{bmatrix}}_{([Z]^T[Z])^{-1}} \right] \cdot \begin{Bmatrix} 552.74 \\ 22421.43 \end{Bmatrix}$$

$$= \begin{Bmatrix} -0.85872 \\ 1.031592 \end{Bmatrix}$$

$$a_0 = -0.85872$$
$$a_1 = 1.031592$$

Standard error of the estimation:

$$s_{y|x} = \sqrt{\frac{S_r}{n-(m+1)}}$$

Here

$$S_r = \sum (y_i - a_0 - a_1 x_i)^2$$
$$S_r = 9.69104$$

$$\therefore s_{y|x} = \sqrt{\frac{9.69104}{15-(1+1)}} = 0.863403$$

Since

$$z_{11}^{-1} = 0.688414$$
$$z_{22}^{-1} = 0.000465$$

$$s(a_0) = \sqrt{z_{11}^{-1}\left(s_{y|x}\right)^2}$$
$$s(a_0) = \sqrt{(0.688414)(0.863403)^2}$$
$$s(a_0) = 0.716372$$

$$s(a_1) = \sqrt{z_{22}^{-1}\left(s_{y|x}\right)^2}$$
$$s(a_1) = \sqrt{(0.000465)(0.863403)^2}$$
$$s(a_1) = 0.018625$$

For a 95% confidence interval,

$$n = 15$$
$$\alpha = 0.05$$

$$t_{\alpha/2,\ n-2} = t_{0.05/2,\ 13} = 2.160368$$

NOTE: You can find this value in excel by using TINV(0.05, 13)

For a_0:
The lower bound
$L(a_0) = a_0 - t_{\alpha/2,\ n-2} \cdot S(a_0)$
$L(a_0) = (-0.85872) + (2.160368) \cdot (0.716372)$
$L(a_0) = -2.40634$

The upper bound
$U(a_0) = a_0 + t_{\alpha/2,\ n-2} \cdot S(a_0)$
$U(a_0) = (-0.85872) + (2.160368) \cdot (0.716372)$
$U(a_0) = 0.688912$

$\therefore\ -2.40634 < a_0 < 0.688912$

For a_1:
The lower bound
$L(a_1) = a_1 - t_{\alpha/2,\ n-2} \cdot S(a_1)$
$L(a_1) = (1.031592) - (2.160368) \cdot (0.018625)$
$L(a_1) = 0.991355$

The upper bound
$U(a_1) = a_1 + t_{\alpha/2,\ n-2} \cdot S(a_1)$
$U(a_1) = (1.031592) + (2.160368) \cdot (0.018625)$
$U(a_1) = 1.071828$

$\therefore 0.991355 < a_1 < 1.071828$

NOTE: Lets look at the slope – when we use our hypothesis testing, and we provide our model, we try to test our model. Ideally the measured data fits the model exactly. So, we expect the slope of the fit line to be close to 1, or equal to 1. By our estimation, we find that our slope is between 0.99 and 1.07.

Therefore, the test result support our hypothesis from the slope point of view because the target slope equals 1 and by our estimation the 1 is between our interval for a_1.

Non-linear regression

$$f(x) = a_0(1 - e^{-a_1 x})$$

Using Gauss-Newton method to solve the problem.

Data:

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

Curve to fit:

$$y = f(x_i, a_0, a_1, \dots, a_m) + e$$

$$\begin{cases} y_1 = f(x_1, a_0, a_1, \dots, a_m) + e_1 \\ y_2 = f(x_2, a_0, a_1, \dots, a_m) + e_2 \\ \qquad \dots \\ y_n = f(x_n, a_0, a_1, \dots, a_m) + e_n \end{cases}$$

$$y_i = f(x_i) + e_i \qquad (i = 1, 2, \dots, n)$$

Iteration:

$$f(x_i)_{j+1} = \boxed{f(x_i)_j} + \boxed{\frac{\delta f(x_i)_j}{\delta a_0}} \Delta a_0 + \boxed{\frac{\delta f(x_j)}{\delta a_1}} \Delta a_1$$

$$j = 1, 2, 3, \dots$$

Note: we’re using the first few terms of the Taylor expansion to determine approximate results.

The error equation:

$$y_i - f(x_i) = e_i$$

$$\Rightarrow \quad \boxed{y_i - f(x_i)_j = \frac{\delta f(x_i)_j}{\delta a_0} \Delta a_0 + \frac{\delta f(x_i)_j}{\delta a_1} \Delta a_1 + e_i}$$

Here $m = 1$

$$\{D\} = [Z]\{\Delta A\} + \{E\}$$

Here

$$\{D\} = \begin{Bmatrix} y_i - f(x_1)_j \\ y_2 - f(x_2)_j \\ \dots \\ y_n - f(x_n)_j \end{Bmatrix}$$

$$\{E\} = \begin{Bmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{Bmatrix}$$

$$\{\Delta A\} = \begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix}$$

$$[Z] = \begin{bmatrix} \frac{\delta f(x_1)_j}{\delta a_0} & \frac{\delta f(x_1)_j}{\delta a_1} \\ \frac{\delta f(x_2)_j}{\delta a_0} & \frac{\delta f(x_2)_j}{\delta a_1} \\ \dots & \dots \\ \frac{\delta f(x_n)_j}{\delta a_0} & \frac{\delta f(x_n)_j}{\delta a_1} \end{bmatrix}$$

tall matrix

(Generally, no Solution)

$$\boxed{[Z]^T[Z]\{\Delta A\} = \{Z\}^T\{D\}}$$

$$\{\Delta A\} = ([Z]^T[Z])^{-1}[Z]^T\{D\}$$

$$(a_0)_{j+1} = (a_0)_j + \Delta a_0$$
$$(a_1)_{j+1} = (a_1)_j + \Delta a_1$$

Find the error:

$$\epsilon_k = \left|\frac{(a_k)_{j+1} - (a_k)_j}{(a_k)_{j+1}}\right| \cdot 100\% \qquad (k = 0, 1)$$

Example

x	0.25	0.75	1.25	1.75	2.25
y	0.28	0.57	0.68	0.74	0.79

Use the data to fit:

$$y = a_0(1 - e^{-a_1 x})$$

Using the initial guess of $a_0 = 1$ and $a_1 = 1$

Solution

$$f(x) = a_0(1 - e^{-a_1 x})$$

The partial derivatives are:

$$\begin{cases} \frac{\delta f}{\delta a_0} = 1 - e^{-a_1 x} \\ \frac{\delta f}{a_1} = a_0 x e^{-a_1 x} \end{cases}$$

The first iteration

$$a_0 = 1$$
$$a_1 = 1$$

$$[Z] = \begin{bmatrix} \frac{\delta f(x_1)}{\delta a_0} & \frac{\delta f(x_1)}{\delta a_1} \\ \frac{\delta f(x_2)}{\delta a_0} & \frac{\delta f(x_2)}{\delta a_1} \\ \dots & \dots \\ \frac{\delta f(x_5)}{\delta a_0} & \frac{\delta f(x_5)}{\delta a_1} \end{bmatrix} = \begin{bmatrix} 1 - e^{-a_1 x_1} & a_0 x_1 e^{-a_1 x_1} \\ 1 - e^{-a_1 x_2} & a_0 x_1 e^{-a_1 x_2} \\ \dots & \dots \\ 1 - e^{-a_1 x_5} & a_0 x_5 e^{-a_1 x_5} \end{bmatrix}$$

$$[Z] = \begin{bmatrix} 0.2212 & 0.1947 \\ 0.5276 & 0.3543 \\ 0.7135 & 0.3581 \\ 0.8262 & 0.3041 \\ 0.8946 & 0.2371 \end{bmatrix}$$

$$\{D\}_0 = \begin{Bmatrix} y_1 - f(x_1) \\ y_2 - f(x_2) \\ \dots \\ y_5 - f(x_5) \end{Bmatrix} = \begin{Bmatrix} y_1 - a_0(1 - e^{-a_1 x_1}) \\ y_2 - a_0(1 - e^{-a_1 x_2}) \\ \dots \\ y_5 - a_0(1 - e^{-a_1 x_5}) \end{Bmatrix}$$

$$\{D\}_0 = \begin{Bmatrix} 0.0588 \\ 0.0424 \\ -0.0335 \\ -0.0862 \\ -0.1046 \end{Bmatrix}$$

$$[Z]_0^T [Z]_0 \{\Delta A\} = [Z\}_0^T [D\}$$

$$\begin{bmatrix} 2.3193 & 0.9489 \\ 0.9489 & 0.4404 \end{bmatrix} \begin{Bmatrix} \Delta a_0 \\ \Delta a_1 \end{Bmatrix} = \begin{Bmatrix} -0.1533 \\ -0.0365 \end{Bmatrix}$$

$$\begin{Bmatrix} \Delta a_0 \\ \Delta a_1 \end{Bmatrix} = \begin{Bmatrix} -0.2714 \\ 0.5019 \end{Bmatrix}$$

$$\begin{Bmatrix} \Delta a_0 \\ \Delta a_1 \end{Bmatrix} = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} + \begin{Bmatrix} -0.2714 \\ 0.5019 \end{Bmatrix} = \begin{Bmatrix} 0.7286 \\ 1.5109 \end{Bmatrix}$$

The relative error

For a_0:

$$\left| \frac{0.7286 - 1}{0.7286} \right| \cdot 100\% = 37\%$$

For a_1:

$$\left| \frac{1.5109 - 1}{1.5019} \right| \cdot 100\% = 33\%$$

The second iteration:

$$a_0 = 0.7286$$
$$a_1 = 1.5019$$

$$[Z]_1 = \begin{bmatrix} 1 - e^{-a_1 x_1} & a_0 x_1 e^{-a_1 x_1} \\ \dots & \dots \\ 1 - e^{-a_1 x_5} & a_0 x_5 e^{-a_1 x_5} \end{bmatrix}$$

$$[Z]_1 = \begin{bmatrix} 0.3130 & 0.1251 \\ 0.6758 & 0.1771 \\ 0.8470 & 0.1393 \\ 0.9278 & 0.09204 \\ 0.9659 & 0.05585 \end{bmatrix}$$

$$\{D\}_1 = \begin{Bmatrix} y_1 = a_0(1 - e^{-a_1 x_1}) \\ \dots \\ y_5 = a_0(1 - e^{-a_1 x_5}) \end{Bmatrix} = \begin{Bmatrix} 0.05194 \\ 0.07765 \\ 0.06293 \\ 0.06407 \\ 0.08630 \end{Bmatrix}$$

$$\{\Delta A\} = \begin{Bmatrix} 0.06252 \\ 0.1758 \end{Bmatrix}$$

$$\begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} = \begin{Bmatrix} 0.7286 \\ 1.5019 \end{Bmatrix} + \begin{Bmatrix} 0.06252 \\ 0.1758 \end{Bmatrix} = \begin{Bmatrix} 0.7910 \\ 1.6777 \end{Bmatrix}$$

The relative error

For a_0:

$$\left| \frac{0.7910 - 0.7286}{0.7910} \right| \cdot 100\% = 7.9\%$$

Fora_1:

$$\left| \frac{1.6777 - 1.5019}{1.6777} \right| \cdot 100\% = 10.5\%$$

The 3rd iteration:

$$\begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} = \begin{Bmatrix} 0.7919 \\ 1.6753 \end{Bmatrix}$$

Relative errors are 0.1% and 0.15%

The 4th iteration:

$$\begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} = \begin{Bmatrix} 0.7919 \\ 1.6751 \end{Bmatrix}$$

Thus,

$$\boxed{\therefore y = f(x) = 0.7919(1 - e^{-1.6751x})}$$