Total Least Squares

Definition:

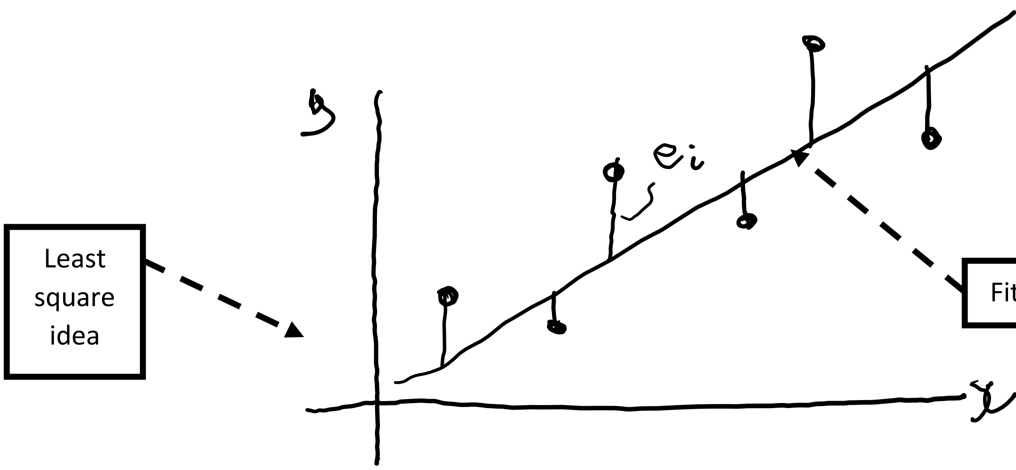
Given a matrix $A_{m \cdot n}$, $m > n$ (tall matrix), and a vector $b \in R^m$, find residuals $E \in R^{m \cdot n}$ and $r \in R^m$ that minimize the Frobenius norm $||E \vdots r||_F$ subject to the conditions $b + r \in I_m(A + E)$

The least-squares

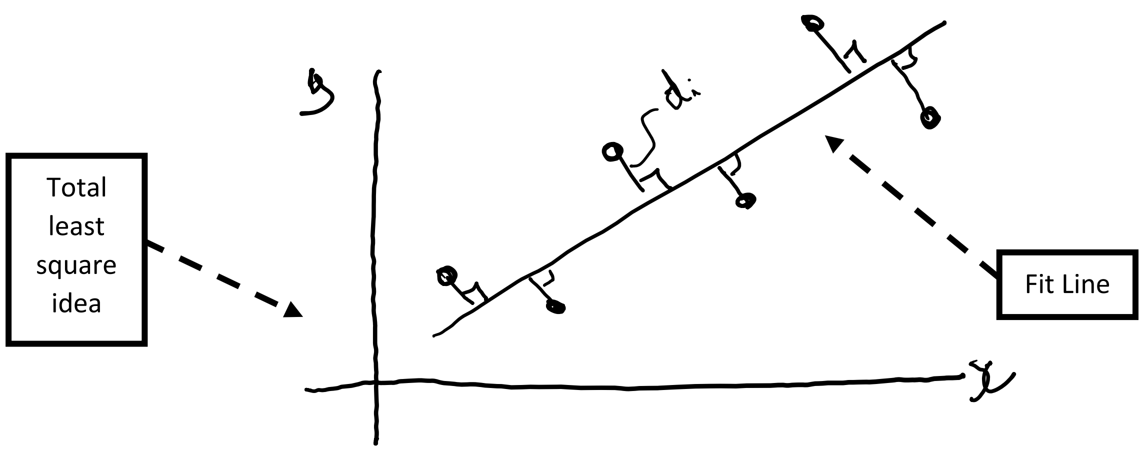
$$y = a_0 + a_1 x + \boxed{e}$$

Given data set:

$$(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$$

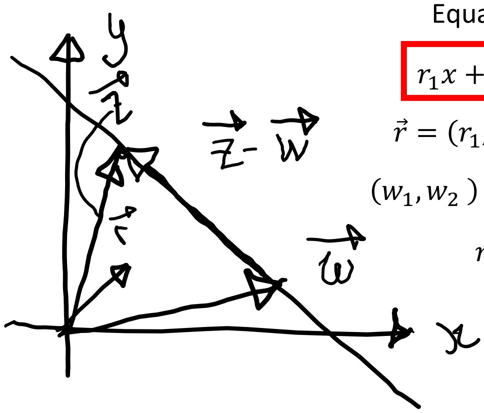


$$S_r = \sum_i (y_i - a_0 - a_1 x_i)^2$$



$$\sum_i d_i^2$$

Distance of a point to a line:



Equation of the line

$$\boxed{r_1 x + r_2 y - \vec{r} \cdot \vec{w} = 0}$$

$$\vec{r} = (r_1, r_2), \vec{w} = (w_1, w_2)$$

(w_1, w_2) is a point on the line

$$r_1^2 + r_2^2 = 1$$

$$\left\|r^2\right\|_2 = 1$$

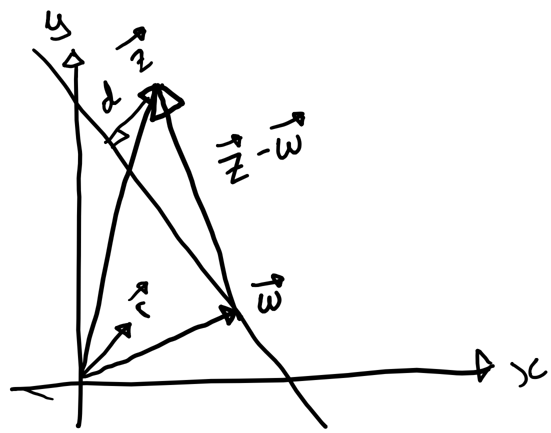
Where $\vec{r} \perp$ line

Take $\vec{z} = (x, y)$

Equation:

$$\boxed{\vec{r} \cdot (\vec{z} - \vec{w}) = 0}$$

$$d = |\vec{r} \cdot (\vec{z} - \vec{w})|$$



Total least squares: find $\vec{r}$ and $\vec{w}$ that minimizing the (error) functional

$$S(\vec{r}, \vec{r}) = \sum_i \left(\vec{r} \cdot (\vec{z} - \vec{w})\right)^2$$

Here

$$\vec{z_i} = (x_i, y_i)$$

Define:

$$\vec{r} = (r_1, r_2)$$
$$\vec{w} = (w_1, w_2)$$

$$S(\vec{r},\vec{w}) = \sum\nolimits_i (r_1 x_i + r_2 y_i - r_1 w_1 - r_2 w_2)^2$$

The centroid of the data set:

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\bar{y} = \frac{\sum y_i}{n}$$

$$S(\vec{r},\vec{w}) = \sum_i \boxed{(r_1(x_i - \bar{x}) + r_2(y_i - \bar{y})} + r_1\bar{x} + r_2\bar{y} - r_1 w_1 - r_2 w_2)^2$$

$$= \sum\nolimits_i \{[r_1(x_i - \bar{x}) + r_2(y_i - \bar{y})]^2 + [r_1(\bar{x} - w_1) + r_2(\bar{y} - w_2)]^2$$
$$+ 2[r_1(x_i - \bar{x}) + r_2(y_i - \bar{y})][r_1(\bar{x} - w_1) + r_2(\bar{y} - w_2)]\}$$

$$= \sum\nolimits_i \left\{[r_1(x_i - \bar{x}) + r_2(y_i - \bar{y})]^2 + n[r_1(\bar{x} - w_1) + r_2(\bar{y} - w_2)]^2 + 2[r_1(x_i - \bar{x}) + r_2(y_i - \bar{y})]\right.$$
$$\left. \cdot \sum_i [r_1(x_i - \bar{x}) + r_2(y_i - \bar{y})]\right\}$$

Since

$$\sum_i [r_1(x_i - \bar{x}) + r_2(y_i - \bar{y})]$$
$$= \sum_i r_1(x_i - \bar{x}) + \sum_i r_1(y_i - \bar{y})$$
$$= r_1 \sum_i (x_i - \bar{x}) + r_2 \sum_i (y_i - \bar{y})$$
$$= r_1(\textstyle\sum_i x_i - \sum_i \bar{x}) + r_2(\sum_i y_i - \sum_i \bar{y})$$
$$= r_1(\textstyle\sum_i x_i - n\bar{x}) + r_2(\sum_i y_i - n\bar{y})$$
$$= 0$$

$$S(\vec{r},\vec{w}) = \sum_i [r_1(x_i - \bar{x}) + r_2(y_i - \bar{y})]^2 + n[r_1(\bar{x} - w_1) + r_2(\bar{y} - w_2)]^2$$

The centroid $\bar{z} = (\bar{x},\bar{y})$ minimizes$\left(r_1(\bar{x} - w_1) + r_2(\bar{y} - w_2)\right)^2$

So,

$$w_1 = \bar{x}, w_2 = \bar{y}$$

Therefore the fitting line passes though the centroid of the data.

$$S(\vec{r},\vec{w}) = \sum_i [r_1(x_i - \bar{x}) + r_2(y_i - \bar{y})]^2$$

Define

$$B = \begin{bmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ x_2 - \bar{x} & y_2 - \bar{y} \\ \vdots & \\ x_n - \bar{x} & y_n - \bar{y} \end{bmatrix}$$

$$r = \begin{Bmatrix} r_1 \\ r_2 \end{Bmatrix}$$

$$\begin{aligned} &S(\vec{r}, \vec{w}) = (B\,r)^T (B\,r) \\ &= r^T B^T B\,r \end{aligned}$$

$$\boxed{\text{Find the vector } r = \begin{Bmatrix} r_1 \\ r_2 \end{Bmatrix} \text{ with } r_1^2 + r_2^2 = 1 \text{ minimizing} \quad S(\vec{r}, \vec{w}) = r^T B^T B\,r}$$

$$\boxed{\text{The right singular vector of} B \text{ corresponding to the smaller singular value of } B, \sigma_2, \text{ is the vector } \vec{r}.}$$

For matrix $B_{n\,x\,2}$, the singular value decomposition is given by:

$$B = U_{n\,x\,2} \textstyle\sum_{2\,x\,2} V_{2\,x\,2}^T$$

Where

$$\textstyle\sum_{2\,x\,2} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \text{ where } \sigma_1 \geq \sigma_2 \geq 0 \text{ (singular values)}$$

The columns of $U_{n\,x\,2}$ are the left singular vectors, the columns of V are the right singular vectors.

$$\begin{aligned} &B^T B = (U \textstyle\sum V^T)^T (U \textstyle\sum V^T) \\ &= V \textstyle\sum^T U^T U \textstyle\sum V^T \\ &= V \textstyle\sum^T \textstyle\sum V^T \\ &= V \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} V^T \end{aligned}$$

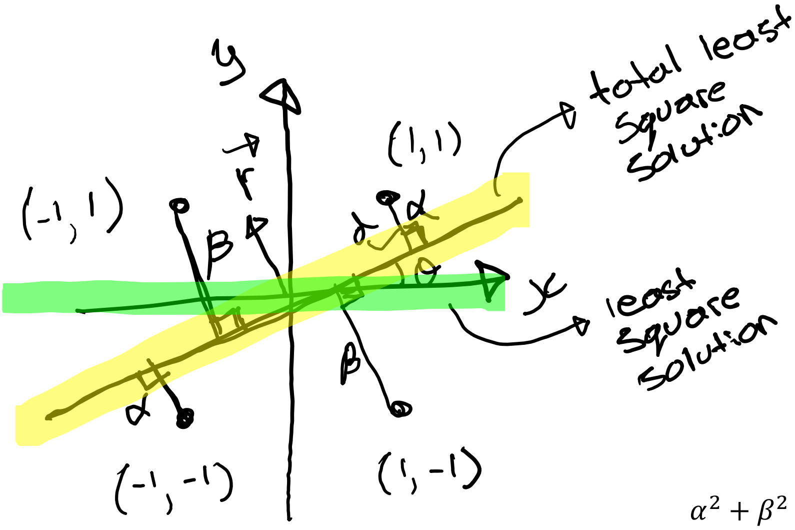
$$\boxed{B^T B V = V \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}}$$

* The TLS (total least square) solution always exists and is given by the line through the centroid orthogonal to the smaller singular vector of B.

* The solution is unique if $\sigma_1 \neq \sigma_2$

Example: (1, 1), (-1, 1), (1, -1), and (-1, -1)

The least squares is the line:



$$y = 0$$

$$S(\vec{r}, \vec{w}) = 2(\alpha^2 + \beta^2)$$

$$r_1 = -\sin\theta$$
$$r_2 = \cos\theta$$

$$\alpha = |\vec{r} \cdot (\vec{z} - \vec{w})|$$
$$\alpha = |(\vec{z} - \vec{w})|$$
$$\alpha = |-\sin\theta\,(1) + \cos\theta\ \,(1)|$$
$$\alpha = |\cos\theta - \sin\theta|$$

$$\beta = |\vec{r} \cdot \vec{z}|$$
$$= |-\sin\theta(1) + cos\theta\,(-1)|$$
$$= |\cos\theta + \sin\theta|$$

$$\alpha^2 + \beta^2 = (\cos\theta - \sin\theta)^2 + (\cos\theta + \sin\theta)^2$$
$$= 2$$

$$\rightarrow S(\vec{r}, \vec{w}) = 4$$

Example

x	1	2	3	4	5	6	7
y	0.5	2.5	2	4	3.5	6	3.5

Using TLS fit a line

Solution:

$$x = \frac{\sum x_i}{n} = \frac{1 + 2 + \cdots + 7}{7} = 4$$

$$y = \frac{\sum y_i}{n} = \frac{0.5 + 2.5 + \cdots + 3.5}{7} = 3.42857$$

$$B = \begin{bmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ x_2 - \bar{x} & y_2 - \bar{y} \\ \vdots & \vdots \\ x_3 - \bar{x} & y_3 - \bar{y} \end{bmatrix} = \begin{bmatrix} 1 - 4 & 0.5 - 3.42857 \\ 2 - 4 & 2.5 - 3.42857 \\ \vdots & \vdots \\ 7 - 4 & 3.5 - 3.42857 \end{bmatrix}$$

$$B == \begin{bmatrix} -3 & -2.9285 \\ -2 & -0.92857 \\ \vdots & \vdots \\ 3 & 2.0714286 \end{bmatrix}_{7\,x\,2}$$

$$B^T B = \begin{bmatrix} 28 & 23.5 \\ 23.5 & 22.714286 \end{bmatrix}$$

Eigenvalues 1.70900 and 49.005286 the corresponding eigenvectors are:

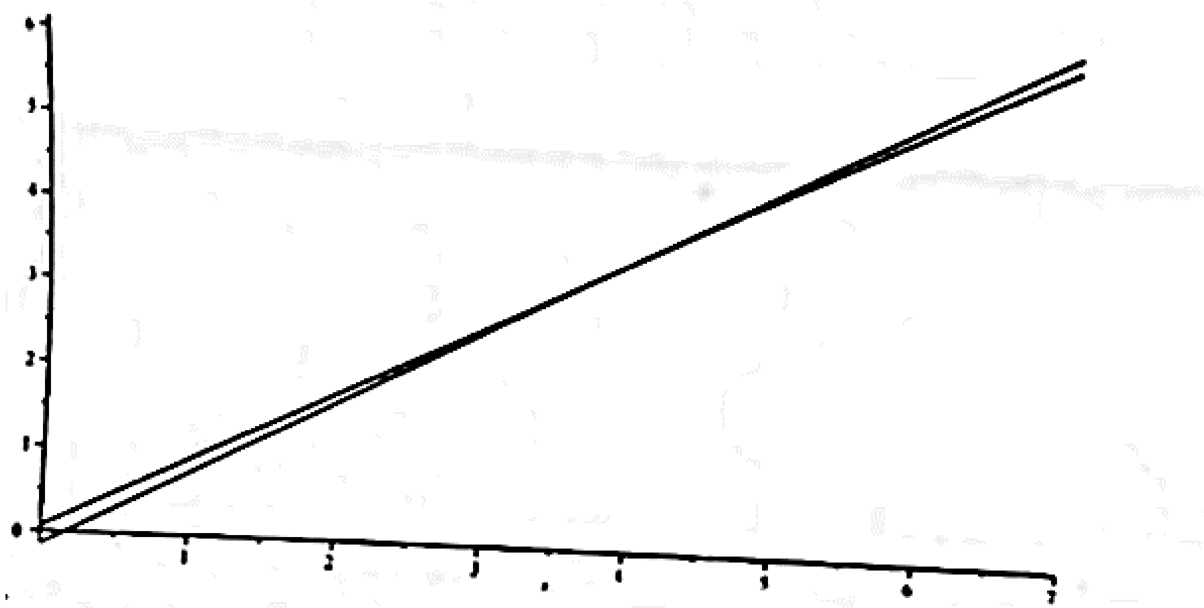
$$\begin{Bmatrix} 0.666424 \\ -0.745573 \end{Bmatrix} \text{ and } \begin{Bmatrix} -0.745573 \\ -0.666424 \end{Bmatrix}$$

$$\begin{aligned} r_1 &= 0.666424 \\ r_2 &= -0.745573 \\ w_1 &= \bar{x} = 4 \\ w_2 &= \bar{y} = 3.42857 \end{aligned}$$

The line

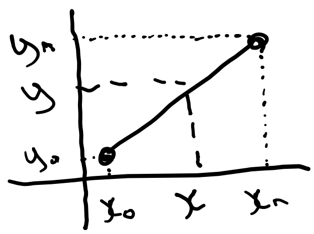
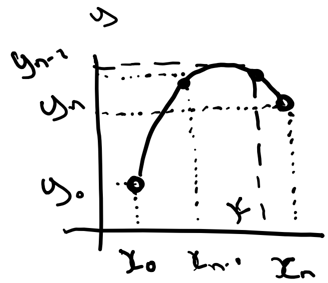
$$r_1 x + r_2 y - r_1 w_1 - r_2 w_2 = 0$$

$$\boxed{0.666424x - 0.745573y - 0.109447 = 0}$$



least square $y = 0.0714286 + 0.839286x$

Part 2: Interpolation

Given a data set:

$$(x_0, y_o), (x_1, y_1), \dots, (x_n, y_n)$$

Fit a polynomial of degree n:

$$
\begin{aligned}
&f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = y_0 \\
&f(x_0) = a_0 + a_1 x_0 + a_2 x_1^2 + \cdots + a_n x_1^n = y_1 \\
&\dots \\
&f(x_n) = a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_n x_n^n = y_n
\end{aligned}
$$

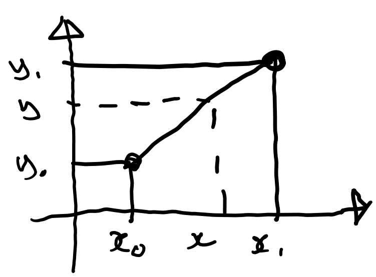
$$
\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^2 \\ 1 & x_1 & x_1^2 & \cdots & x_1^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^2 \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{Bmatrix} = \begin{Bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{Bmatrix}
$$

↑ Vandermonde matrix

2.1 Newton's divided difference interpolating polynomials

Liner interpolation:

Slope:



$$
\begin{aligned}
\frac{y_1 - y_0}{x_1 - x_0} &= \frac{y - y_0}{x - x_0} \\
y &= y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0)
\end{aligned}
$$

$$\boxed{f(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0)}$$

$y = f(x)$

$\frac{f(x_1) - f(x_0)}{x_1 - x_0}: \textit{the first finite divided difference}$

Quadratic interpolation:

$$f_2(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)(x - x_1)$$
$$f_2(x_0) = b_0 = f(x_o)$$
$$f_2(x_1) = b_0 + b_1(x_1 - x_0) = f(x_1)$$
$$f_2(x_2) = b_0 + b_1(x_2 - x_0) + b_2(x_2 - x_0)(x_2 - x_1) = f(x_2)$$

$$b_0 = f(x_0)$$
$$b_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$
$$b_2 = \frac{1}{(x_2 - x_0)}\left[f(x_2) - f(x_0) - \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x_2 - x_0)\right]$$
$$= \frac{1}{(x_2 - x_0)}\left[f(x_2) - f(x_1) + f(x_1) - f(x_0) - \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x_2 - x_0)\right]$$
$$= \frac{1}{(x_2 - x_0)}\left[f(x_2) - f(x_1) + f(x_1) - f(x_0) \cdot \left(1 - \frac{x_2 - x_0}{x_1 - x_0}\right)\right]$$
$$= \frac{1}{(x_2 - x_0)(x_2 - x_1)}[f(x_2) - f(x_1) + f(x_1) - f(x_0)\left(1 - \frac{x_2 - x_0}{x_1 - x_0}\right)$$
$$= \frac{1}{(x_2 - x_0)(x_2 - x_1)}\left[f(x_2) - f(x_1) + f(x_1) - f(x_0)\left(\frac{x_1 - x_2}{x_1 - x_0}\right)\right]$$
$$= \frac{1}{x_2 - x_0} \cdot \left(\frac{f(x_2) - f(x_1)}{x_2 - x_1} - \frac{f(x_1) - f(x_0)}{x_1 - x_0}\right)$$
$$= \frac{1}{x_2 - x_0}\left(\frac{f(x_2) - f(x_1)}{x_2 - x_1} - \frac{f(x_1) - f(x_0)}{x_1 - x_0}\right)$$

This is the second finite divided difference

Example: Fit data points
$x_0 = 1, f(x_0) = 0$
$x_1 = 4, f(x_1) = 1.386294$
$x_2 = 6, f(x_2) = 1.791759$

using quadratic polynomial.

Solution: $f_2(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)(x - x_1)$
$b_0 = f(x_0) = 0$

$$b_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{1.386294 - 0}{4 - 1} = 0.4620981$$
$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{1.791759 - 1.386294}{6 - 4} = 0.2027325$$
$$b_2 = \frac{\frac{f(x_2) - f(x_1)}{x_2 - x_1} - \frac{f(x_1) - f(x_0)}{x_1 - x_0}}{x_2 - x_0}$$
$$b_2 = \frac{0.2027325 - 0.4620981}{6 - 1}$$

$\therefore f_2(x) = 0.4620981(x-1) - 0.0518731(x-1)(x-4)$

Use this polynomial to evaluate $f(2)$:
$f(2) = f_2(2) = 0.5658444$
$f(x) = \ln x$
$f(2) = \ln 2 = 0.6931472$

$$Relative\ error\ = \left|\frac{0.5658444 - 0.6931472}{0.6931472}\right| \cdot 100\%$$
$$Relative\ error = 18.4\ \%$$

Using the first two data points to find $f(2)$:
$f_1(x) = b_0 + b_1(x - x_0) = 0.4620981(x-1)$
$f_1(2) = 0.4620981$

$$Relative\ error\ = \left|\frac{0.4620581 - 0.6931472}{0.6931472}\right| \cdot 100\%$$
$$Relative\ error = 33.3\%$$

General form of Newton's interpolation:

$$f_n(x) = b_0 + b_1(x - x_0) + b_2(x - x_0)(x - x_1) + \cdots + b_n(x - x_0)(x - x_1)\dots(x - x_n)$$

Here:

$$b_0 = f(x_0)$$
$$b_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} \triangleq f[x_1, x_0]$$
$$b_2 = \frac{f[x_2, x_1] - f[x_1, x_2]}{x_2 - x_0} \triangleq f[x_2, x_1, x_0]$$
$$b_3 = \frac{f[x_3, x_2, x_1] - f[x_2, x_1, x_0]}{x_3 - x_0} \triangleq f[x_3, x_2, x_1, x_0]$$
$$Recursive$$

$$f[x_n, x_{n-1}, x_1, x_0] = \frac{f[x_n, x_{n-1}, \dots, x_1] - f[x_{n-1}, \dots, x_1, x_0]}{x_n - x_0}$$
$\uparrow$ $nth\ finite\ divided\ difference$

Example: Estimate $f(2)$ using a third-order Newton's interpolating polynomial:

$$x_0 = 1 \quad ; \quad f(x_0) = 0$$
$$x_1 = 4 \quad ; \quad f(x_1) = 1.386294$$
$$x_2 = 6 \quad ; \quad f(x_2) = 1.791759$$
$$x_3 = 5 \quad ; \quad f(x_3) = 1.609438$$

i	x_i	$f(x_i)$	$f[x_i, x_j] = \frac{f(x_i) - f(x_j)}{(x_i - x_j)}$	$f[x_i, x_j, x_k] = \frac{f[x_i, x_j] - f[x_j, x_k]}{(x_i - x_k)}$	$f[x_i, x_j, x_k, x_l]$
0	1	$\boxed{0}$			
1	4	1.386	$\boxed{0.4620981}$		
2	6	1.792	0.2027326	$\boxed{-0.05187311}$	
3	5	1.609	0.1823266	-0.0204110	$\boxed{0.007865539}$

$$\therefore f_3(x) = 0.4620981(x-1) - 0.0518711(x-1)(x-4) + 0.007865539(x-1)(x-4)(x-6)$$

$f_3(x) = 0.6287686$
$RE = 9.3\%$

Errors (Newton's interpolating polynomial)

$$f(x) = f_n(x) + R_n(x)$$

The error:

$$R_n(x) = \frac{f^{n+1}(c)}{(n+1)!} \cdot (x - x_0)(x - x_1) \ldots (x - x_n)$$

Here c is the interval containing the data using the finite divided difference.

$$R_n = \underbrace{f[x_1, x_n, x_{n+1}, \ldots, x_1, x_0]}_{(n+1)^{th}}(x - x_0)(x - x_1) \ldots (x - x_n)$$

If there is extra data. $[x_{n+1}, f(x_{n+1})]$:
Then:

$$R_n \approx f[x_{n+1}, x_n, \ldots, x_0](x - x_0) \ldots (x - x_n)$$

Example: Using quadratic polynomial: $f_2(x) = 0.4620981(x-1) - 0.0518731(x-1)(x-4)$ using $x_3 = 5$, $f(x_3) = 1.609438$ to estimate the error.

Double - check (incorrect ?)

Solution:

$$R_2 = f[x_3, x_2, x_1, x_0](x_3 - x_0)(x_3 - x_1)(x_3 - x_2)$$
$$R_2 = 0.00786553(5-1)(5-4)(5-6)$$

Error at $x = 2$:

$$R_2 = f[x_3. x_2, x_1, x_0](2-1)(2-4)(2-6)$$
$$R_2 = 0.0629$$