보간법 : 불연속적으로 주어진 데이터 점들을 이용하여 그 점들 사이의 값을 추정하는 방법의 한 종류, 주어진 데이터 값이 매우 정밀한 경우에 적용하는 방법

보간법의 절차:

1. 먼저 주어진 데이터 점들을 모두 지나는 추정 함수(estimated function)
2. 추정 함수를 이용하여 알고자 하는 함수 값을 계산

즉, 추정 함수에 독립 변수의 값을 대입하여 데이터 값이 주어져있지 않은 점의 함수값을 계산

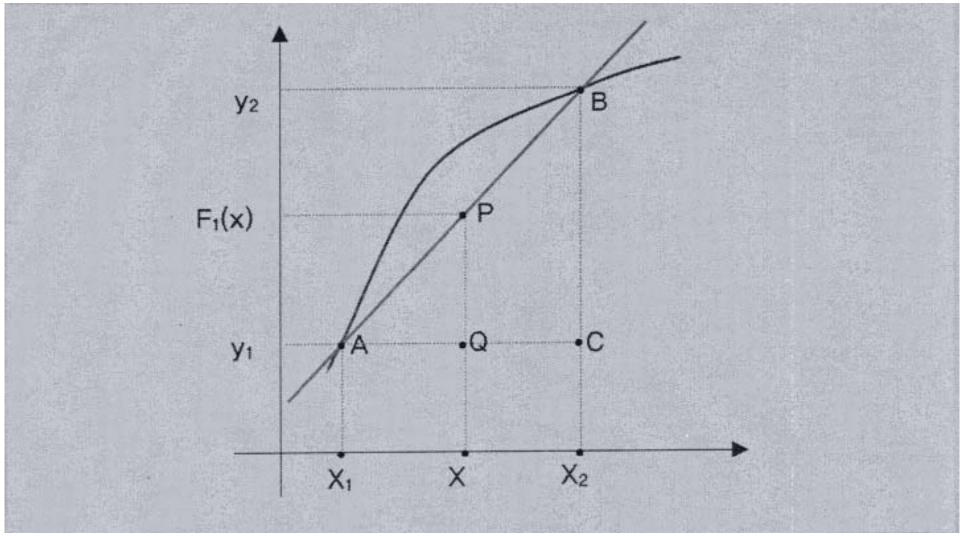
일반적으로 n+1 개의 데이터 점이 주어져 있을 경우에 추정 함수는 n차 다항식

1. Newton 보간법

1) **선형 보간**(Linear interpolation)

- 주어진 2개의 점을 직선으로 연결하여 추정 함수
- 선형 보간의 원리

1) 선형 보간(Linear interpolation)



$$\frac{\overline{BC}}{\overline{AC}} = \frac{\overline{PQ}}{\overline{AQ}}$$

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{F_1(x) - y_1}{x - x_1}$$

$$F_1(x) = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$$

2개의 데이터 점 $(x_1,\ y_1), (x_2,\ y_2)$가 주어져 있을 때

(단, $x_1 < x_2$)

1차 보간 함수: $F_1(x) = b_0 + b_1(x - x_1)$

$b_0 = y_1$, $b_1 = \dfrac{y_2 - y_1}{x_2 - x_1}$

2) 2차 보간(Quadratic interpolation)

- 주어진 3개의 데이터 점들을 2차식의 곡선으로 연결하여 추정 함수를 구하는 방법
  - 3개의 점 A(x_1 , y_1), B(x_2 , y_2), C(x_3 , y_3)를 지나는 함수(2차식):

$$F_2(x) = b_0 + b_1(x - x_1) + b_2(x - x_1)(x - x_2)$$

단, $x_1 < x_2 < x_3$

전개하여 오름차순으로 정리하면

$$F_2(x) = b_0 + b_1 x - b_1 x_1 + b_2 x^2 - b_2 x_1 x - b_2 x_2 x + b_2 x_1 x_2$$
$$= (b_0 - b_1 x_1 + b_2 x_1 x_2) + (b_1 - b_2 x_1 - b_2 x_2)x - (b_2)x^2$$

2차식의 일반꼴:

$$F_2(x) = a_0 + a_1 x + a_2 x^2$$

2) 2차 보간(Quadratic interpolation)

- 주어진 3개의 데이터 점들을 2차식의 곡선으로 연결하여 추정 함수를 구하는 방법

식 (4.4)=식 (4.6)이므로 두 식 사이에 동류항의 계수가 각각 같아야

$$a_0 = b_0 - b_1 x_1 + b_2 x_1 x_2$$
$$a_1 = b_1 - b_2 x_1 - b_2 x_2$$
$$a_2 = b_2$$

주어진 3점의 데이터를 이용하여 b_0, b_1, b_2 를 구하여 2차 다항식의 추정 함수를 결정

추정 함수 $F_2(x)$ 의 그래프는 주어진 데이터 점들을 모두 지나야 하므로 다음 관계가 성립

$$F_2(x_1) = y_1 \,, F_2(x_2) = y_2 \,, F_2(x_3) = y_3$$

- b_0, b_1, b_2 결정하기

$$F_2(x_1) = b_0 + b_1(x_1 - x_1) - b_2(x_1 - x_1)(x_1 - x_2) = b_0$$

$F_2(x_1) = y_1$ 이므로

$$b_0 = y_1$$

2) 2차 보간(Quadratic interpolation)

- 주어진 3개의 데이터 점들을 2차식의 곡선으로 연결하여 추정 함수를 구하는 방법

식 (4.4)=식 (4.6)이므로 두 식 사이에 동류항의 계수가 각각 같아야

$$a_0 = b_0 - b_1 x_1 + b_2 x_1 x_2$$
$$a_1 = b_1 - b_2 x_1 - b_2 x_2$$
$$a_2 = b_2$$

주어진 3점의 데이터를 이용하여 b_0, b_1, b_2 를 구하여 2차 다항식의 추정 함수를 결정

추정 함수 $F_2(x)$ 의 그래프는 주어진 데이터 점들을 모두 지나야 하므로 다음 관계가 성립

$$F_2(x_1) = y_1 , F_2(x_2) = y_2 , F_2(x_3) = y_3$$

- b_0 , b_1 , b_2 결정하기

$$F_2(x_1) = b_0 + b_1(x_1 - x_1) - b_2(x_1 - x_1)(x_1 - x_2) = b_0$$

$F_2(x_1) = y_1$ 이므로

$$b_0 = y_1$$

2) 2차 보간(Quadratic interpolation)

- 주어진 3개의 데이터 점들을 2차식의 곡선으로 연결하여 추정 함수를 구하는 방법

$$F_2(x_2) = y_1 + b_1(x_2 - x_1) + b_2(x_2 - x_1)(x_2 - x_2) = y_1 + b_1(x_2 - x_1)$$

$F_2(x_2) = y_2$ 이므로

$$y_1 + b_1(x_2 - x_1) = y_2$$

$$b_1 = \frac{y_2 - y_1}{x_2 - x_1}$$

$$F_2(x_3) = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x_3 - x_1) + b_2(x_3 - x_1)(x_3 - x_2)$$

$F_2(x_3) = y_3$ 이므로

$$y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x_3 - x_1) + b_2(x_3 - x_1)(x_3 - x_2) =$$

$$b_2 = \frac{y_3 - y_1}{(x_3 - x_1)(x_3 - x_2)} - \frac{y_2 - y_1}{x_2 - x_1} \square \frac{1}{x_3 - x_2}$$

2) 2차 보간(Quadratic interpolation)

- 주어진 3개의 데이터 점들을 2차식의 곡선으로 연결하여 추정 함수를 구하는 방법

$$= \frac{\dfrac{y_3 - y_1}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1} \cdot \dfrac{x_3 - x_1}{x_3 - x_2}}{x_3 - x_1}$$

$$= \frac{\dfrac{y_3 - y_2}{x_3 - x_2} + \dfrac{y_2 - y_1}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1} \cdot \dfrac{x_3 - x_1}{x_3 - x_2}}{x_3 - x_1}$$

$$= \frac{\dfrac{y_3 - y_2}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1}\left[\dfrac{x_3 - x_1}{x_3 - x_2} - \dfrac{x_2 - x_1}{x_3 - x_2}\right]}{x_3 - x_1}$$

$$\therefore b_2 = \frac{\dfrac{y_3 - y_2}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1}}{x_3 - x_1}$$

3개의 데이터 점 $(x_1,\ y_1), (x_2,\ y_2), (x_3,\ y_3)$이 주어져 있을 때

(단, $x_1 < x_2 < x_3$)

2차 보간 함수: $F_2(x) = b_0 + b_1(x - x_1) + b_2(x - x_1)(x - x_2)$

$$b_0 = y_1,\ b_1 = \frac{y_2 - y_1}{x_2 - x_1},\ \therefore b_2 = \frac{\dfrac{y_3 - y_2}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1}}{x_3 - x_1}$$

2) 2차 보간(Quadratic interpolation)

- 주어진 3개의 데이터 점들을 2차식의 곡선으로 연결하여 추정 함수를 구하는 방법

$$= \frac{\dfrac{y_3 - y_1}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1} \square \dfrac{x_3 - x_1}{x_3 - x_2}}{x_3 - x_1}$$

$$= \frac{\dfrac{y_3 - y_2}{x_3 - x_2} + \dfrac{y_2 - y_1}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1} \square \dfrac{x_3 - x_1}{x_3 - x_2}}{x_3 - x_1}$$

$$= \frac{\dfrac{y_3 - y_2}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1}\left[\dfrac{x_3 - x_1}{x_3 - x_2} - \dfrac{x_2 - x_1}{x_3 - x_2}\right]}{x_3 - x_1}$$

$$\therefore b_2 = \frac{\dfrac{y_3 - y_2}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1}}{x_3 - x_1}$$

3개의 데이터 점 $(x_1,\ y_1), (x_2,\ y_2), (x_3,\ y_3)$이 주어져 있을 때

(단, $x_1 < x_2 < x_3$)

2차 보간 함수: $F_2(x) = b_0 + b_1(x - x_1) + b_2(x - x_1)(x - x_2)$

$$b_0 = y_1,\ b_1 = \frac{y_2 - y_1}{x_2 - x_1},\ \therefore b_2 = \frac{\dfrac{y_3 - y_2}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1}}{x_3 - x_1}$$

3) Newton interpolation

- 주어진 n+1개의 데이터 점들을 n차식의 곡선으로 연결하여 추정 함수

$$F_n(x) = b_0 + b_1(x - x_1) + \ldots + b_n(x - x_1)(x - x_2)\ldots(x - x_n)$$

$$b_0 = y_1$$

$$b_1 = f[x_2, x_1]$$

$$b_2 = f[x_3, x_2, x_1]$$

$$\vdots$$

$$b_n = f[x_{n+1}, x_n, \ldots, x_2, x_1]$$

- **1계 차분:** $f[x_2, x_1] = \dfrac{y_2 - y_1}{x_2 - x_1}$, $f[x_3, x_2] = \dfrac{y_3 - y_2}{x_3 - x_2}$, …
- **2계 차분:** $f[x_3, x_2, x_1] = \dfrac{f[x_3, x_2] - f[x_2, x_1]}{x_3 - x_1}$
- **3계 차분:** $f[x_4, x_3, x_2, x_1] = \dfrac{f[x_4, x_3, x_2] - f[x_3, x_2, x_1]}{x_4 - x_1}$

…

- **n계 차분:**

$$f[x_{n+1}, x_n, \ldots, x_2, x_1] = \frac{f[x_{n+1}, x_n, \ldots, x_2] - f[x_n, x_{n-1}, \ldots, x_1]}{x_{n+1} - x_1}$$

2. Lagrange interpolation

- Lagrange 보간다항식은 Newton의 다항식을 재구성

Newton의 선형 보간식:

$$F_1(x) = y_1 + f[x_2, x_1](x - x_1)$$

차분식:

$$f[x_2, x_1] = \frac{y_2 - y_1}{x_2 - x_1} = \frac{y_2}{x_2 - x_1} - \frac{y_1}{x_1 - x_2}$$

$$F_1(x) = y_1 + \left[\frac{y_2}{x_2 - x_1} + \frac{y_1}{x_1 - x_2}\right](x - x_1) = \frac{x - x_1}{x_2 - x_1} y_2 + \frac{x - x_1}{x_1 - x_2} y_1$$

여기에서 구하는 Lagrange 선형 보간식은:

$$F_1(x) = \frac{x - x_2}{x_1 - x_2} y_1 + \frac{x - x_1}{x_2 - x_1} y_2$$

2. Lagrange interpolation

- Lagrange 보간다항식은 Newton의 다항식을 재구성

일반적인 경우에 (n+1)개의 데이터점 (x_1, y_1) , (x_2, y_2) , ... ,(x_{n+1}, y_{n+1}) 통과하는 다항식에 대한 n차의 Lagrange 선형 보간식은

$$F_n(x) = \sum_{i=1}^{n+1} L_i(x) y_i$$

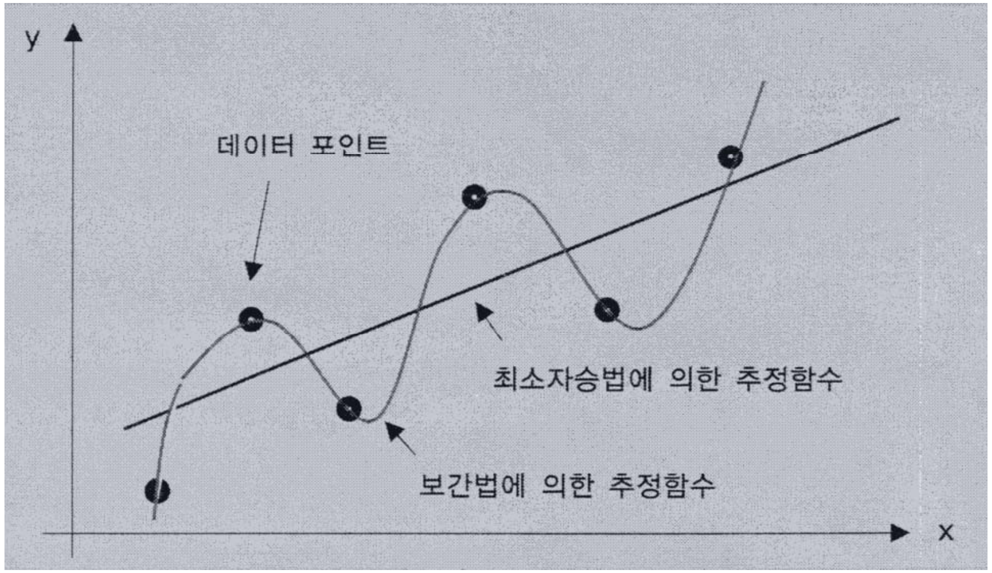
단,

$$L_i(x) = \frac{(x-x_1)(x-x_2)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_{n+1})}{(x_i-x_1)(x_i-x_2)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_{n+1})}$$

$$= \prod_{\substack{j=1 \\ j \neq i}}^{n+1} \left[\frac{x-x_j}{x_i-x_j} \right]$$

Regression Analysis

	보간법	최소자승법
추정함수 요건	데이터 포인트를 모두 지나는 함수	전체 데이터의 일반적인 경향을 나타내는 함수
추정함수 형태	다항식	다항식, 지수식, 다변수 함수
데이터 정밀도	높음	상대적으로 낮음
데이터 수	적음	많음



1. 선형회귀
2. 지수곡선 회귀
3. 다중 회귀
4. 타원 회귀

Regression Analysis

1) 선형 회귀(Liner Regression Analysis)

- 전체 데이터 포인트에 가장 적합한 직선식,
- 즉 전체 데이터의 경향을 나타내는, 오차가 최소인 1차 방정식을 구하는 것

n개의 데이터 포인트 $(x_1, f_1), (x_2, f_2), \cdots, (x_n, f_n)$ 에 적합한 직선식 가정:

$$y(x) = c_1 + c_2 x$$

근사식에 $x_1, x_2, \ldots, x_n$ 대입:

$$\begin{aligned} y(x_1) &= c_1 + c_2 x_1 \\ y(x_2) &= c_1 + c_2 x_2 \\ &\vdots \\ y(x_n) &= c_1 + c_2 x_n \end{aligned}$$

추정함수값 $y(x_i)$ 와 주어진 데이터 f_i 의 차이를 오차 r_i 로 간주하면:

$$\begin{aligned} r_1 &= y(x_1) - f_1 = c_1 + c_2 x_1 - f_1 \\ r_2 &= y(x_2) - f_2 = c_1 + c_2 x_2 - f_2 \\ &\vdots \\ r_n &= y(x_n) - f_n = c_1 + c_2 x_n - f_n \end{aligned}$$

오차를 제곱하면:

$$
\begin{aligned}
(r_1)^2 &= (c_1 + c_2 x_1 - f_1)^2 \\
(r_2)^2 &= (c_1 + c_2 x_2 - f_2)^2 \\
&\vdots \\
(r_n)^2 &= (c_1 + c_2 x_n - f_n)^2
\end{aligned}
$$

오차 제곱의 합을 s로 두면:

$$S = \sum_{i=1}^{n} (r_i)^2 = \sum_{i=1}^{n} (c_1 + c_2 x_i - f_i)^2$$

여기서 s가 최소가 되는 조건은:

$$\frac{\partial S}{\partial c_1} = 0, \qquad \frac{\partial S}{\partial c_2} = 0$$

$$S = (c_1 + c_2 x_1 - f_1)^2 + (c_1 + c_2 x_2 - f_2)^2 + \cdots + (c_1 + c_2 x_n - f_n)^2$$

$$S = (c_1 + c_2 x_1 - f_1)^2 + (c_1 + c_2 x_2 - f_2)^2 + \cdots + (c_1 + c_2 x_n - f_n)^2$$

편미분항을 구하면:

$$\frac{\partial S}{\partial c_1} = 2(c_1 + c_2 x_1 - f_1) + 2(c_1 + c_2 x_2 - f_2) + \cdots + 2(c_1 + c_2 x_n - f_n)$$

$$= 2(nc_1 + c_2 \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} f_i) = 0$$

$$\frac{\partial S}{\partial c_2} = 2(c_1 + c_2 x_1 - f_1)x_1 + 2(c_1 + c_2 x_2 - f_2)x_2 + \cdots + 2(c_1 + c_2 x_n - f_n$$

$$= 2(c_1 \sum_{i=1}^{n} x_i + c_2 \sum_{i=1}^{n} x_i^{\,2} - \sum_{i=1}^{n} x_i f_i) = 0$$

c_1, c_2 에 대한 2원 1차 연립방정식을 얻는다.

$$nc_1 + (\sum_{i=1}^{n} x_i)c_2 = \sum_{i=1}^{n} f_i$$
$$(\sum_{i=1}^{n} x_i)c_1 + (\sum_{i=1}^{n} x_i^{\,2})c_2 = \sum_{i=1}^{n} x_i f_i$$

Crammer 공식을 이용하여 다음과 같이 계수 c_1, c_2 를 구한다.

$$c_1 = \frac{\begin{vmatrix} \sum f_i & \sum x_i \\ \sum x_i f_i & \sum x_i^{\,2} \end{vmatrix}}{\begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^{\,2} \end{vmatrix}}, \qquad c_2 = \frac{\begin{vmatrix} n & \sum f_i \\ \sum x_i & \sum x_i f_i \end{vmatrix}}{\begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^{\,2} \end{vmatrix}}$$

$$y(x) = c_1 + c_2 x$$

2) 지수 곡선(Exponential Function)

전체 데이터의 경향을 나타내는, 오차가 최소인 지수 곡선식을 구하는 것.

n 개의 데이터 포인트 $(x_1, f_1), (x_2, f_2), \cdots, (x_n, f_n)$ 에 대하여 지수곡선식을 가정:

$$y(x) = ab^x$$

지수곡선 회귀의 목표는 주어진 데이터를 이용해서 계수 a,b,를 구하는 것, 양변에 자연대수를 취하면:

$$\ln y = \ln a + x \ln b$$

$Y = \ln y, \quad A = \ln a, \quad B = \ln b$ 로 치환하면:

$$Y = A + Bx \quad \text{(근사값)}$$

데이터의 자연대수를 취하면:

$$\ln f = F \quad (\text{참값})$$

선형회귀'의 절차에 따라 $|Y-F|$를 최소로 하는 A,B에 대하여 다음의 연립방정식

$$nA + \left(\sum_{i=1}^{n} x_i\right)B = \sum_{i=1}^{n} F_i$$

$$\left(\sum_{i=1}^{n} x_i\right)A + \left(\sum_{i=1}^{n} x_i^{2}\right)B = \sum_{i=1}^{n} x_i F_i$$

Crammer 공식에 의하여 A,B를 구한다.

$$A = \frac{\begin{vmatrix} \sum F_i & \sum x_i \\ \sum x_i F_i & \sum x_i^{2} \end{vmatrix}}{\begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^{2} \end{vmatrix}}, \qquad B = \frac{\begin{vmatrix} n & \sum F_i \\ \sum x_i & \sum x_i F_i \end{vmatrix}}{\begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^{2} \end{vmatrix}}$$

$A = \ln a, \quad \ln b = B$ 의 관계에서:

$$a = e^{A}$$
$$b = e^{B}$$

최종적으로 지수 곡선 회귀식을 결정

$$y(x) = ab^{x}$$

3) 다항식 회기(Polynomial Regression Analysis)

전체 데이이터의 경향을 나타내는, 오차가 최소인 고차 다항식을 구하는 것.

선형 회귀'와 동일

n 개의 측정점 $(x_1, f_1), (x_2, f_2), \cdots, (x_n, f_n)$ 에 대한 m차 다항식 (단, m<n) :

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

각 점에서의 오차 r_i는:

$$r_i = y(x_i) - f_i$$

오차 제곱의 합을 S라 하면:

$$S = \sum_{i=1}^{n} \left(y(x_i) - f_i \right)^2$$
$$= \sum_{i=1}^{n} \left(a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_m x_i^m - f_i \right)^2$$

S를 최소로 하는 $a_0, a_1, a_2, \cdots, a_m$ 등을 다음과 같이

$$\frac{\partial S}{\partial a_j} = 0, \quad (j = 0, 1, 2, \ldots, m)$$

즉,

$$\frac{\partial S}{\partial a_0} = 2\sum_{i=1}^{n}\left(a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_m x_i^m - f_i\right) = 0$$

$$\frac{\partial S}{\partial a_1} = 2\sum_{i=1}^{n}\left(a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_m x_i^m - f_i\right) x_i = 0$$

$$\cdots\cdots\cdots\cdots$$

$$\frac{\partial S}{\partial a_m} = 2\sum_{i=1}^{n}\left(a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_m x_i^m - f_i\right) x_i^m = 0$$

위 식에서 a_i 에 관한 (m+1)원 연립방정식을

$$
\begin{aligned}
&a_0 n + a_1 \sum x_i + \cdots + a_m \sum x_i^m = \sum f_i \\
&a_0 \sum x_i + a_1 \sum x_i^2 + \cdots + a_m \sum x_i^{m+1} = \sum f_i x_i \\
&\qquad \cdots\cdots\cdots\cdots\cdots \\
&a_0 \sum x_i^m + a_1 \sum x_i^{m+1} + \cdots + a_m \sum x_i^{2m} = \sum f_i x_i^m
\end{aligned}
$$

Gauss 정규방정식

행렬과 벡터로 표시하면:

Aa=b

단,

$$
A = \begin{bmatrix} n & \sum x_i & \cdots & \sum x_i^m \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{m+1} \\ \cdots & \cdots & \cdots & \cdots \\ \sum x_i^m & \sum x_i^{m+1} & \cdots & \sum x_i^{2m} \end{bmatrix}, \quad a = \begin{bmatrix} a_0 \\ a_1 \\ \cdots \\ a_m \end{bmatrix}, \quad b = \begin{bmatrix} \sum f_i \\ \sum f_i x_i \\ \cdots \\ \sum f_i x_i^m \end{bmatrix}
$$

계수 $a_0, a_1, a_2, \cdots, a_m$ 을 구하여 m차 다항식을 결정.

$$C = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix}$$

계수행렬은 대칭행렬이므로 다음과 같이 표현될 수 있다.

$$A = C^T C = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_1^m & x_2^m & x_3^m & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix}$$

$$= \begin{bmatrix} n & \sum x_i & \cdots & \sum x_i^m \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{m+1} \\ \cdots & \cdots & \cdots & \cdots \\ \sum x_i^m & \sum x_i^{m+1} & \cdots & \sum x_i^{2m} \end{bmatrix}$$

$$b = C^T f = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_1^m & x_2^m & x_3^m & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ \cdots \\ f_n \end{bmatrix} = \begin{bmatrix} \sum f_i \\ \sum f_i x_i \\ \cdots \\ \sum f_i x_i^m \end{bmatrix}$$

3) 다중 회기(Multiple Regression Analysis)

독립변수가 2개 이상일 때 전체 데이터의 경향을 오차가 최소인 다변수 함수로

독립 변수가 x,y, 종속변수가 z인 경우에 n개의 측정점 $(x_1, y_1, f_1), (x_2, y_2, f_2), \cdots, (x_n, y_n, f_n)$ 에

대한 다중 회귀 방정식은 다음과 같이

z(x,y)=a+bx+cy

각 점에서의 오차 r_i 는 측정치 f_i와 추정치 $z_i(x_i, y_i)$의 차이로 표현

$$r_i = z_i(x_i, y_i) - f_i$$

오차 제곱의 합을 S라 하면:

$$S = \sum_{i=1}^{n} (r_i)^2 = \sum_{i=1}^{n} (a + bx_i + cy_i - f_i)^2$$

S가 최소일 조건:

$$\frac{\partial S}{\partial a} = 0, \frac{\partial S}{\partial b} = 0, \frac{\partial S}{\partial c} = 0$$

3) 다중 회기(Multiple Regression Analysis)

$$na+\left(\sum x_i\right)b+\left(\sum y_i\right)c=\sum f_i$$

$$\left(\sum x_i\right)a+\left(\sum x_i^2\right)b+\left(\sum x_i y_i\right)c=\sum x_i f_i$$

$$\left(\sum y_i\right)a+\left(\sum x_i y_i\right)b+\left(\sum y_i^2\right)c=\sum y_i f_i$$

행렬과 벡터로 표현하면:

$$\begin{bmatrix} n & \sum x_i & \sum y_i \\ \sum x_i & \sum x_i^2 & \sum x_i y_i \\ \left(\sum y_i\right) & \sum x_i y_i & \sum y_i^2 \end{bmatrix} \begin{Bmatrix} a \\ b \\ c \end{Bmatrix} = \begin{Bmatrix} \sum f_i \\ \sum x_i f_i \\ \sum y_i f_i \end{Bmatrix}$$

연립방정식을 풀어서 계수 a,b,c를 결정하고, 다중 회귀식 결정

z(x,y)=a+bx+cy