Universitas Gadjah Mada
Departemen Teknik Sipil dan Lingkungan
Prodi Teknik Sipil (Program Sarjana)

https://istiarto.staff.ugm.ac.id

REGRESI DAN INTERPOLASI

Curve Fitting

Curve Fitting

- Acuan
  - Chapra, S.C., Canale R.P., 1990, *Numerical Methods for Engineers*, 2nd Ed., McGraw-Hill Book Co., New York.
    - Chapter 11 dan 12, pp. 319-398.

Curve Fitting

- Mencari garis/kurva yang mewakili serangkaian titik data
- Ada dua cara untuk melakukannya, yaitu
  - Regresi
  - Interpolasi
- Aplikasi di bidang enjiniring
  - Pola perilaku data (*trend analysis*)
  - Uji hipotesis (*hypothesis testing*)

Curve Fitting

- Pemakaian regresi
  - Apabila data menunjukkan tingkat kesalahan yang cukup signifikan atau menunjukkan adanya noise
  - Untuk mencari satu kurva tunggal yang mewakili pola umum perilaku data
  - Kurva regresi tidak perlu melewati setiap titik data

Curve Fitting

- Interpolasi
  - Diketahui bahwa data sangat akurat
  - Untuk mencari satu atau serangkaian kurva yang melewati setiap titik data
  - Untuk memperkirakan nilai-nilai di antara titik-titik data
- Extrapolasi
  - Mirip dengan interpolasi, tetapi untuk memperkirakan nilai-nilai di luar kisaran titik-titik data

Curve Fitting terhadap Data Pengukuran

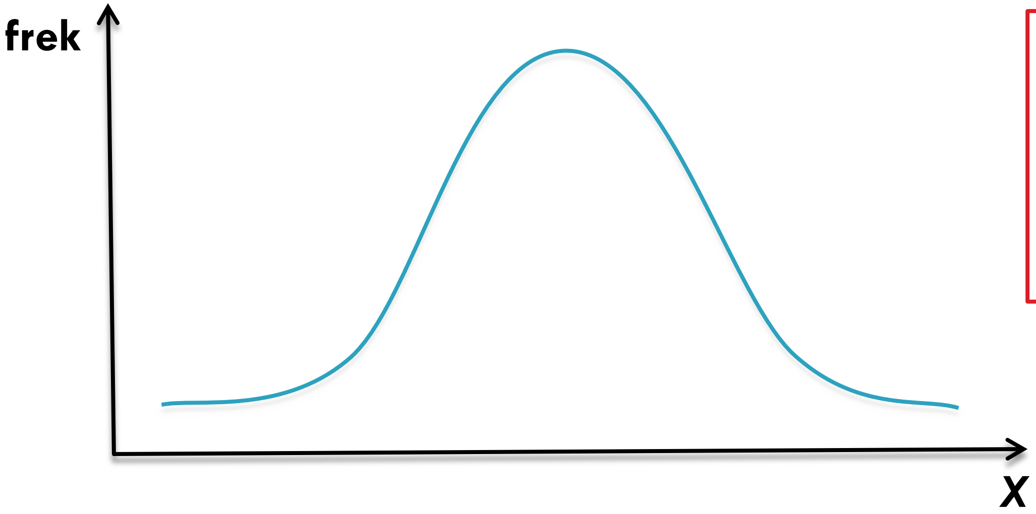
- Analisis pola perilaku data
  - Pemanfaatan pola data (pengukuran, experimen) untuk melakukan perkiraan
  - Apabila data persis (akurat): interpolasi
  - Apabila data tak persis (tak akurat): regresi
- Uji hipotesis
  - Pembandingan antara hasil teori atau hasil hitungan dengan hasil pengukuran

Beberapa Parameter Statistik

merepresentasikan sebaran data

- Rerata aritmatik, *mean* ➡ $\bar{y} = \frac{1}{n}\sum y_i$
- Deviasi standar, simpangan baku, *standard deviation* ➡ $s_y = \sqrt{\frac{S_t}{n-1}}$ $\quad S_t = \sum (y_i - \bar{y})^2$
- Varian ('ragam'), *variance* ➡ $s_y{}^2 = \frac{S_t}{n-1}$
- *Coefficient of variation* ➡ $c.v. = \frac{s_y}{\bar{y}} 100\%$

Distribusi Probabilitas



Distribusi Normal

salah satu distribusi/sebaran data yang sering dijumpai adalah distribusi normal

Regresi

Regresi linear

Regresi non-linear

Regresi: Metode Kuadrat Terkecil

- Mencari satu kurva atau satu fungsi (pendekatan) yang sesuai dengan pola umum yang ditunjukkan oleh data
  - Datanya menunjukkan kesalahan yang cukup signifikan
  - Kurva tidak perlu memotong setiap titik data
- Regresi linear
- Regresi persamaan-persamaan tak-linear yang dilinearkan
- Regresi tak-linear

Regresi: Metode Kuadrat Terkecil

- Bagaimana caranya?
  - Program komputer
  - *Spreadsheet* (Microsoft Excel)

Regresi Linear

- Mencari suatu kurva lurus yang cocok menggambarkan pola serangkaian titik data: (x_1,y_1), (x_2,y_2) … (x_n,y_n)

$$y_{reg} = a_0 + a_1 x$$

a_0 : intercept

a_1 : slope

- Microsoft Excel
  - INTERCEPT(y_1:y_n;x_1:x_n)
  - SLOPE(y_1:y_n;x_1:x_n)

Regresi Linear

- Kesalahan atau residu (e) adalah perbedaan antara nilai *y* sesungguhnya (data *y*) dan *y* nilai pendekatan (y_{reg}) menurut persamaan linear $a_0 + a_1 x$.

$$e = y - (a_0 + a_1 x) = y - a_0 - a_1 x$$

- Minimumkan jumlah kuadrat residu tersebut

$$\min[S_r] = \min\left[\sum {e_i}^2\right] = \min\left[\sum (y_i - a_0 - a_1 x_i)^2\right]$$

Regresi Linear

- Bagaimana cara mencari koefisien a_0 dan a_1?
  - Diferensialkan persamaan tersebut dua kali, masing-masing terhadap a_0 dan a_1.
  - Samakan kedua persamaan hasil diferensiasi tersebut dengan nol.

$$\frac{\partial S_r}{\partial a_0} = -2\sum(y_i - a_0 - a_1 x_i) = 0$$

$$\frac{\partial S_r}{\partial a_1} = -2\sum(y_i - a_0 - a_1 x_i)x_i = 0$$

Regresi Linear

- Selesaikan persamaan yang didapat untuk mencari a_0 dan a_1

$$a_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum {x_i}^2 - (\sum x_i)^2}$$

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

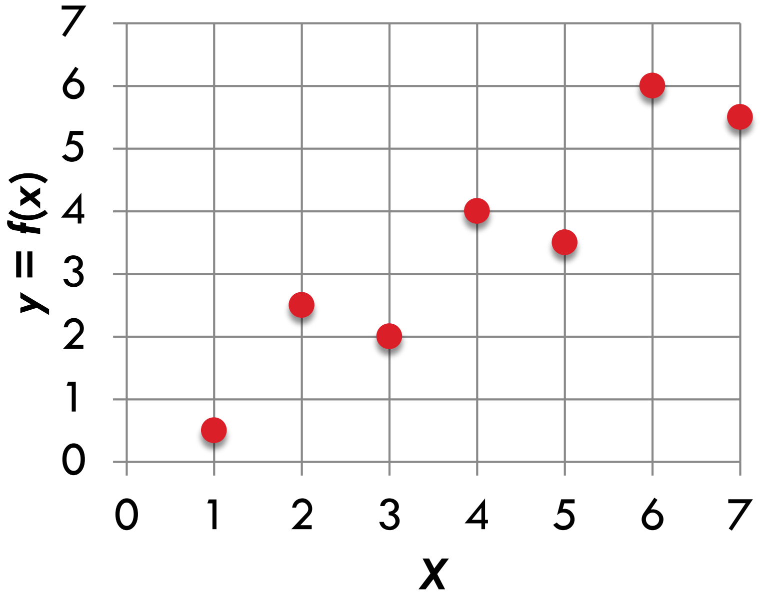
- dalam hal ini, $\bar{y}$ dan $\bar{x}$ masing-masing adalah nilai *y* rerata dan *x* rerata

Contoh Regresi Linear

Tabel data

i	x_i	$y_i = f(x_i)$
0	1	0.5
1	2	2.5
2	3	2
3	4	4
4	5	3.5
5	6	6
6	7	5.5

Grafik/kurva data



Hitungan regresi linear

i	x_i	y_i	$x_i\ y_i$	x_i^2	y_{reg}	$(y_i-y_{reg})^2$	$(y_i-y_{mean})^2$
0	1	0.5	0.5	1	0.910714	0.168686	8.576531
1	2	2.5	5	4	1.75	0.5625	0.862245
2	3	2.0	6	9	2.589286	0.347258	2.040816
3	4	4.0	16	16	3.428571	0.326531	0.326531
4	5	3.5	17.5	25	4.267857	0.589605	0.005102
5	6	6.0	36	36	5.107143	0.797194	6.612245
6	7	5.5	38.5	49	5.946429	0.199298	4.290816
$\sum =$	**28**	**24.0**	**119.5**	**140**	**$\sum =$**	**2.991071**	**22.71429**

Hitungan regresi linear

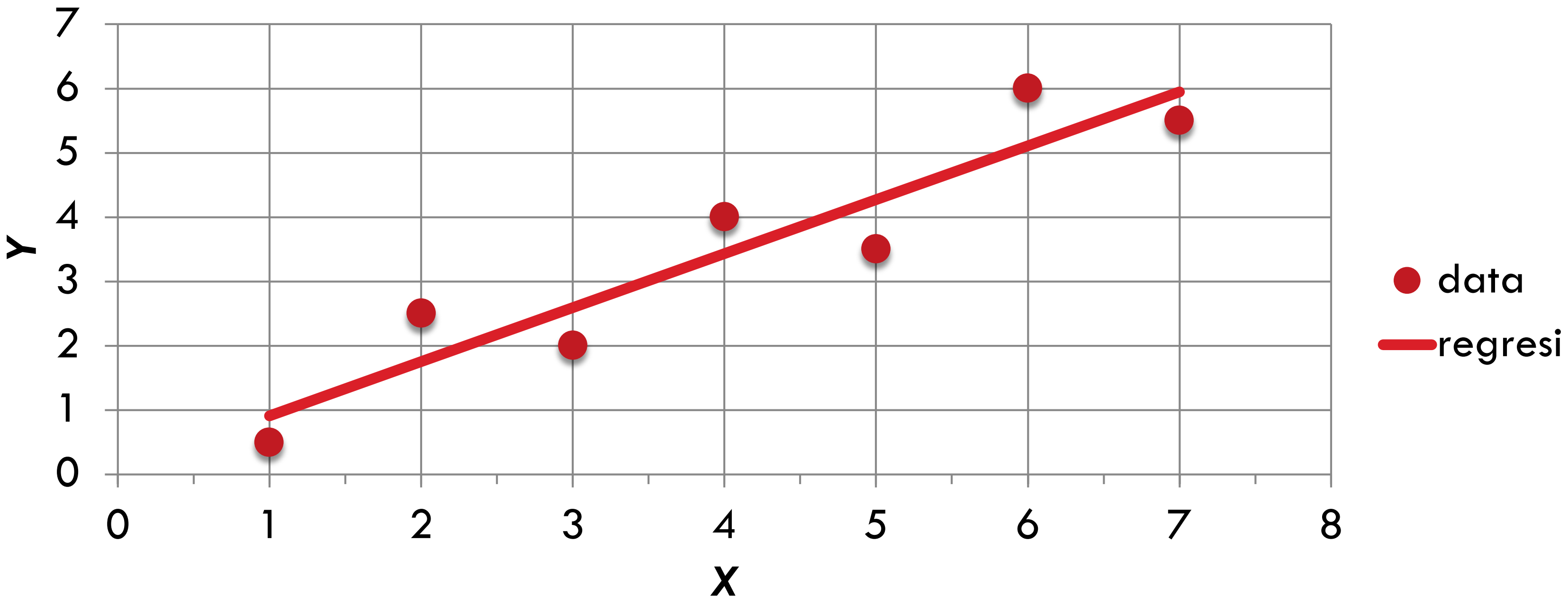
$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum {x_i}^2 - \left(\sum x_i\right)^2} = \frac{7\times 119.5 - 28\times 24}{7\times 140 - 28^2} = 0.839286$$

$$\bar{y} = \frac{24}{7} = 3.4$$

$$\bar{x} = \frac{28}{7} = 4$$

$$a_0 = 3.4 - 0.839286\times 4 = 0.071429$$

Hitungan regresi linear



Regresi Linear

- Kuantifikasi kesalahan
  - Kesalahan standar

$$s_{y/x} = \sqrt{\frac{S_r}{n-2}} \qquad S_r = \sum (y_i - a_0 - a_1 x_i)^2$$

  - Perhatikan kemiripannya dengan simpangan baku

$$s_y = \sqrt{\frac{S_t}{n-1}} \qquad S_t = \sum (y_i - \bar{y})^2$$

Regresi Linear

- Beda antara kedua kesalahan tersebut menunjukkan perbaikan atau pengurangan kesalahan

$$r^2 = \frac{S_t - S_r}{S_r}$$ → koefisien determinasi (*coefficient of determination*)

$$r = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i{}^2 - \left(\sum x_i\right)^2}\ \sqrt{n\sum y_i{}^2 - \left(\sum y_i\right)^2}}$$ → koefisien korelasi (*correlation coefficient*)

Hitungan regresi linear

$$S_r = \sum (y_i - a_0 - a_1 x_i)^2 = 2.991071$$

$$S_t = \sum (y_i - \bar{y})^2 = 22.71429$$

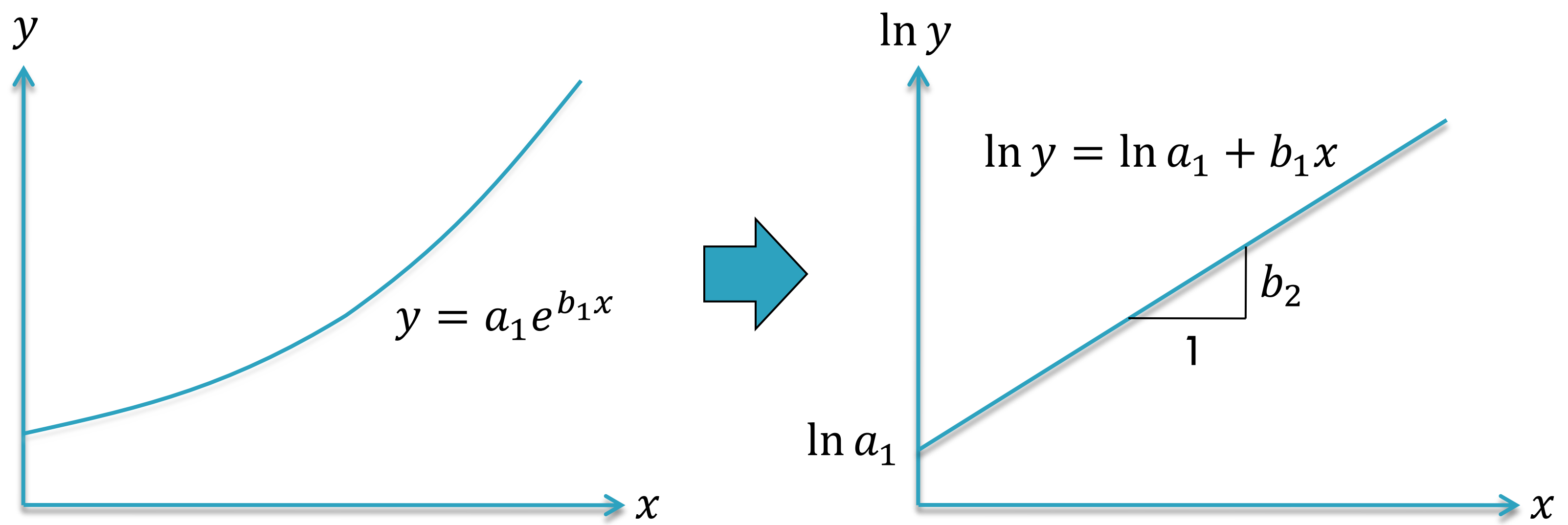
$$r^2 = \frac{S_t - S_r}{S_t} = \frac{22.71429 - 2.991071}{22.71429}$$

$$r = 0.931836$$

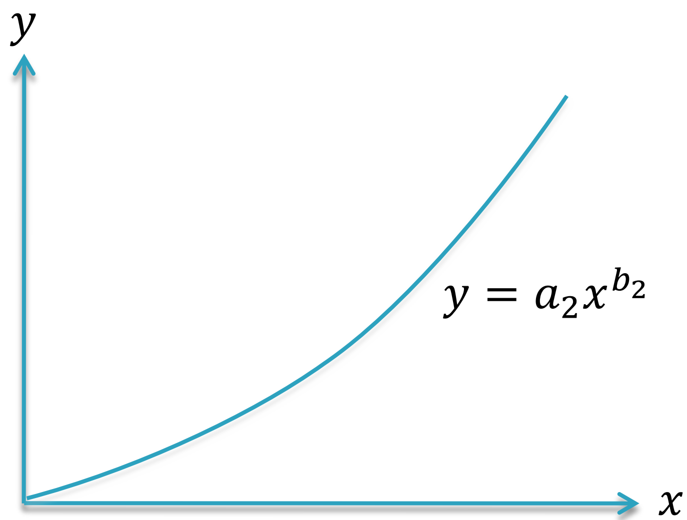
Regresi Linear

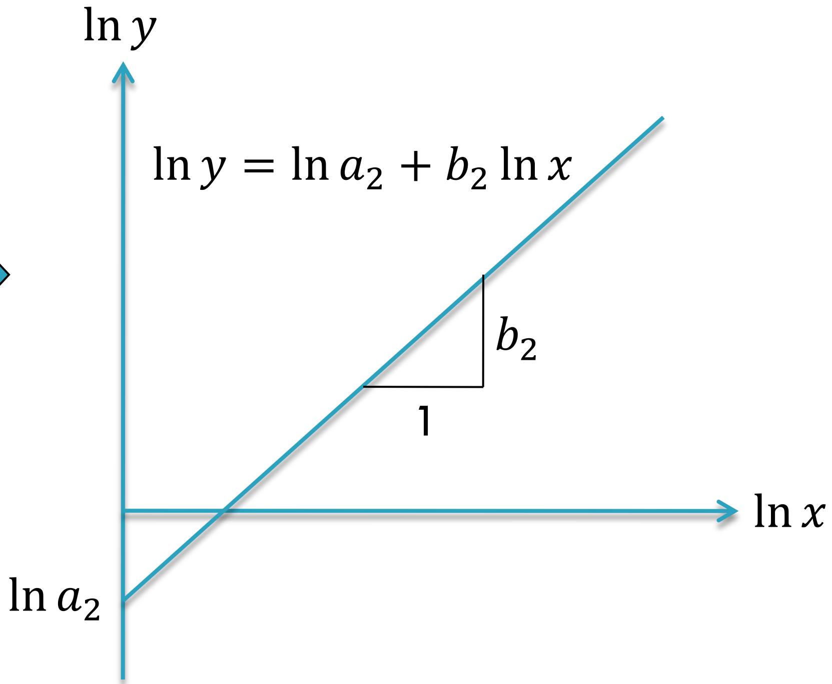
- Linearisasi persamaan-persamaan tak-linear
  - Logaritmik menjadi linear
  - Eksponensial menjadi linear
  - Pangkat (polinomial tingkat $n > 1$) menjadi linear (polinomial tingkat 1)
  - Dll.

Linearisasi persamaan non-linear

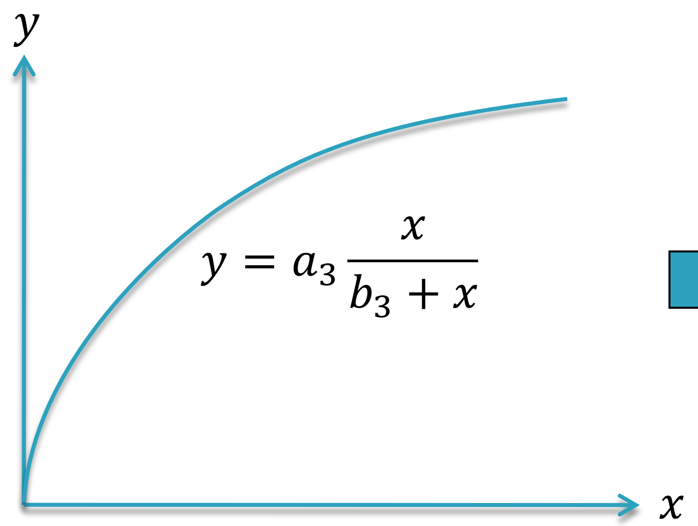


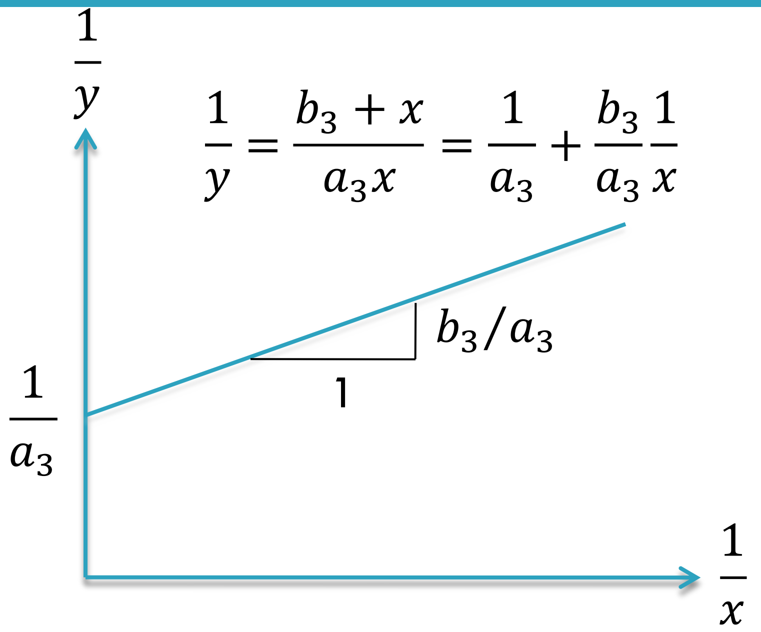
Linearisasi persamaan non-linear





Linearisasi persamaan non-linear





Regresi Polinomial Orde 2

- Persamaan polinomial orde 2

$$y = a_0 + a_1 x + a_2 x^2$$

- Persamaan regresi polinomial orde 2

$$y_{reg} = a_0 + a_1 x + a_2 x^2 + e$$

- Error (residue)

$$e = y - (a_0 + a_1 x + a_2 x^2)$$

Regresi Polinomial Orde 2

- Jumlah data (x_i, y_i) adalah n

$$e_i = y_i - \left(a_0 + a_1 x_i + a_2 {x_i}^2\right)$$

$$S_r = \sum_{i=1}^{n} {e_i}^2 = \sum_{i=1}^{n} \left(y_i - a_0 - a_1 x_i - a_2 {x_i}^2\right)^2$$

- Metode kuadrat terkecil untuk memperoleh a_0, a_1, a_2

$$\frac{\partial S_r}{\partial a_0} = 0 \qquad \frac{\partial S_r}{\partial a_1} = 0 \qquad \frac{\partial S_r}{\partial a_2} = 0$$

Regresi Polinomial Orde 2

$$\frac{\partial S_r}{\partial a_0} = -2\sum(y_i - a_0 - a_1 x_i - a_2 {x_i}^2) = 0$$

$$\frac{\partial S_r}{\partial a_1} = -2\sum(y_i - a_0 - a_1 x_i - a_2 {x_i}^2) x_i = 0$$

$$\frac{\partial S_r}{\partial a_2} = -2\sum(y_i - a_0 - a_1 x_i - a_2 {x_i}^2) {x_i}^2 = 0$$

$$-2(\Sigma y_i - \Sigma a_0 - a_1 \Sigma x_i - a_2 \Sigma {x_i}^2) = 0$$

$$-2(\Sigma x_i y_i - a_0 \Sigma x_i - a_1 \Sigma {x_i}^2 - a_2 \Sigma {x_i}^3) = 0$$

$$-2(\Sigma {x_i}^2 y_i - a_0 \Sigma {x_i}^2 - a_1 \Sigma {x_i}^3 - a_2 \Sigma {x_i}^4) = 0$$

Regresi Polinomial Orde 2

$$-2(\Sigma y_i - \Sigma a_0 - a_1\Sigma x_i - a_2\Sigma {x_i}^2) = 0$$

$$-2(\Sigma x_i y_i - a_0\Sigma x_i - a_1\Sigma {x_i}^2 - a_2\Sigma {x_i}^3) = 0$$

$$-2(\Sigma {x_i}^2 y_i - a_0\Sigma {x_i}^2 - a_1\Sigma {x_i}^3 - a_2\Sigma {x_i}^4) = 0$$

$$(\Sigma y_i - \Sigma a_0 - a_1\Sigma x_i - a_2\Sigma {x_i}^2) = 0$$

$$(\Sigma x_i y_i - a_0\Sigma x_i - a_1\Sigma {x_i}^2 - a_2\Sigma {x_i}^3) = 0$$

$$(\Sigma {x_i}^2 y_i - a_0\Sigma {x_i}^2 - a_1\Sigma {x_i}^3 - a_2\Sigma {x_i}^4) = 0$$

Regresi Polinomial Orde 2

$$\left(\Sigma y_i - \Sigma a_0 - a_1\Sigma x_i - a_2\Sigma {x_i}^2\right) = 0$$

$$\left(\Sigma x_i y_i - a_0\Sigma x_i - a_1\Sigma {x_i}^2 - a_2\Sigma {x_i}^3\right) = 0$$

$$\left(\Sigma {x_i}^2 y_i - a_0\Sigma {x_i}^2 - a_1\Sigma {x_i}^3 - a_2\Sigma {x_i}^4\right) = 0$$

$$\Sigma a_0 + a_1\Sigma x_i + a_2\Sigma {x_i}^2 = \Sigma y_i$$

$$a_0\Sigma x_i + a_1\Sigma {x_i}^2 + a_2\Sigma {x_i}^3 = \Sigma x_i y_i$$

$$a_0\Sigma {x_i}^2 + a_1\Sigma {x_i}^3 + a_2\Sigma {x_i}^4 = \Sigma {x_i}^2 y_i$$

Regresi Polinomial Orde 2

$$\Sigma a_0 + a_1 \Sigma x_i + a_2 \Sigma {x_i}^2 = \Sigma y_i$$

$$a_0 \Sigma x_i + a_1 \Sigma {x_i}^2 + a_2 \Sigma {x_i}^3 = \Sigma x_i y_i$$

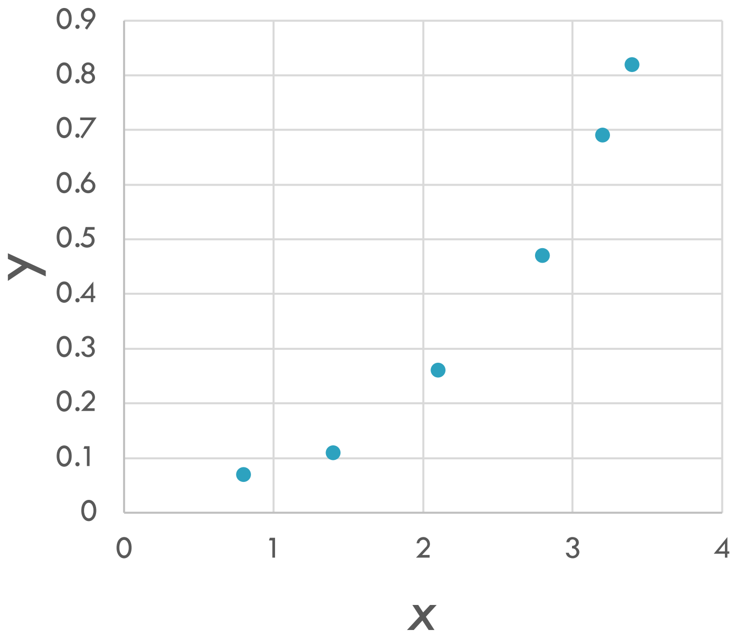
$$a_0 \Sigma {x_i}^2 + a_1 \Sigma {x_i}^3 + a_2 \Sigma {x_i}^4 = \Sigma {x_i}^2 y_i$$

$$\begin{bmatrix} n & \Sigma x_i & \Sigma {x_i}^2 \\ \Sigma x_i & \Sigma {x_i}^2 & \Sigma {x_i}^3 \\ \Sigma {x_i}^2 & \Sigma {x_i}^3 & \Sigma {x_i}^4 \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \end{Bmatrix} = \begin{Bmatrix} \Sigma y_i \\ \Sigma x_i y_i \\ \Sigma {x_i}^2 y_i \end{Bmatrix}$$

$$\begin{Bmatrix} a_0 \\ a_1 \\ a_2 \end{Bmatrix} = \begin{Bmatrix} \cdots \\ \cdots \\ \cdots \end{Bmatrix}$$

Contoh Regresi Polinomial Orde 2

x_i	y_i
0.8	0.07
1.4	0.11
2.1	0.26
2.8	0.47
3.2	0.69
3.4	0.82



Contoh Regresi Polinomial Orde 2

i	x_i	y_i	x_i^2	x_i^3	x_i^4	$x_i\,y_i$	$x_i^2\,y_i$
1	0.8	0.07	0.64	0.512	0.4096	0.056	0.0448
2	1.4	0.11	1.96	2.744	3.8416	0.154	0.2156
3	2.1	0.26	4.41	9.261	19.4481	0.546	1.1466
4	2.8	0.47	7.84	21.952	61.4656	1.316	3.6848
5	3.2	0.69	10.24	32.768	104.8576	2.208	7.0656
6	3.4	0.82	11.56	39.304	133.6336	2.788	9.4792
Σ	**13.7**	**2.42**	**36.65**	**106.541**	**323.6561**	**7.068**	**21.6366**

Contoh Regresi Polinomial Orde 2

$$\begin{bmatrix} n & \Sigma x_i & \Sigma {x_i}^2 \\ \Sigma x_i & \Sigma {x_i}^2 & \Sigma {x_i}^3 \\ \Sigma {x_i}^2 & \Sigma {x_i}^3 & \Sigma {x_i}^4 \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \end{Bmatrix} = \begin{Bmatrix} \Sigma y_i \\ \Sigma x_i y_i \\ \Sigma {x_i}^2 y_i \end{Bmatrix}$$

$$\begin{bmatrix} 6 & 13.7 & 36.65 \\ 13.7 & 36.65 & 106.541 \\ 36.65 & 106.541 & 323.6561 \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \end{Bmatrix} = \begin{Bmatrix} 2.42 \\ 7.068 \\ 21.6366 \end{Bmatrix}$$

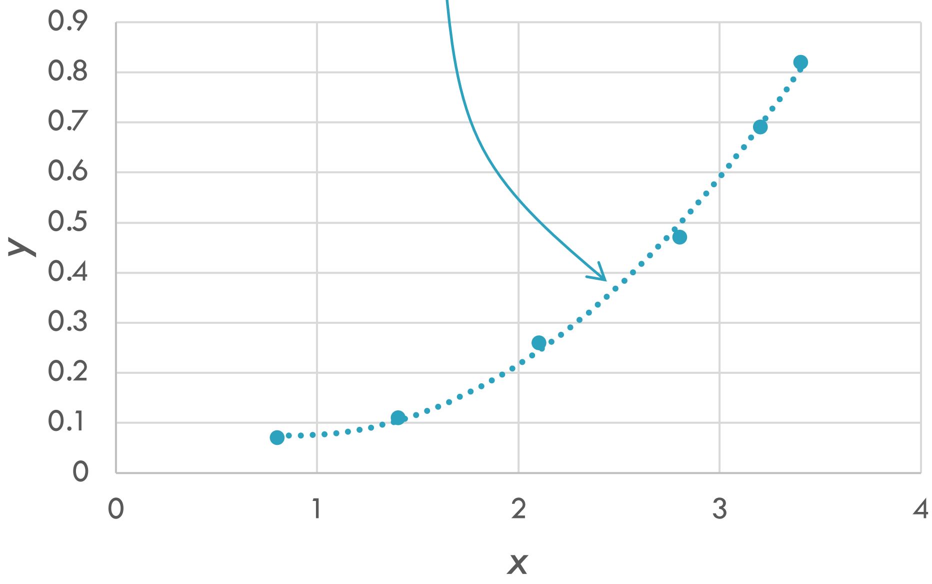
a_0, a_1, a_2 dapat diperoleh dengan salah satu metode penyelesaian sistem persamaan linear (eliminasi Gauss, Gauss-Jordan, Gauss-Seidel, SOR, …)

Contoh Regresi Polinomial Orde 2

$$\begin{Bmatrix} a_0 \\ a_1 \\ a_2 \end{Bmatrix} = \begin{Bmatrix} 0.1685 \\ -0.2089 \\ 0.1165 \end{Bmatrix}$$

$$y_{reg} = 0.1685 - 0.2089x + 0.1165x^2$$

x_i	y_i
0.8	0.0759
1.4	0.1044
2.1	0.2437
2.8	0.4972
3.2	0.6934
3.4	0.8054

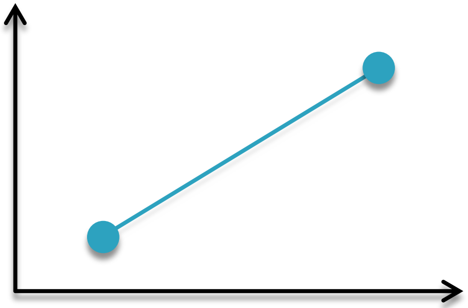


Interpolasi

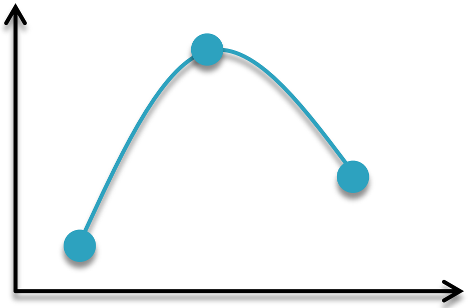
Metode Newton

Metode Lagrange

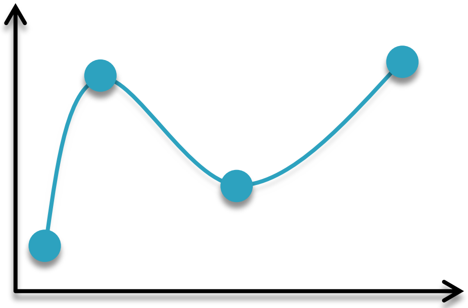
Interpolasi



linear



kuadratik

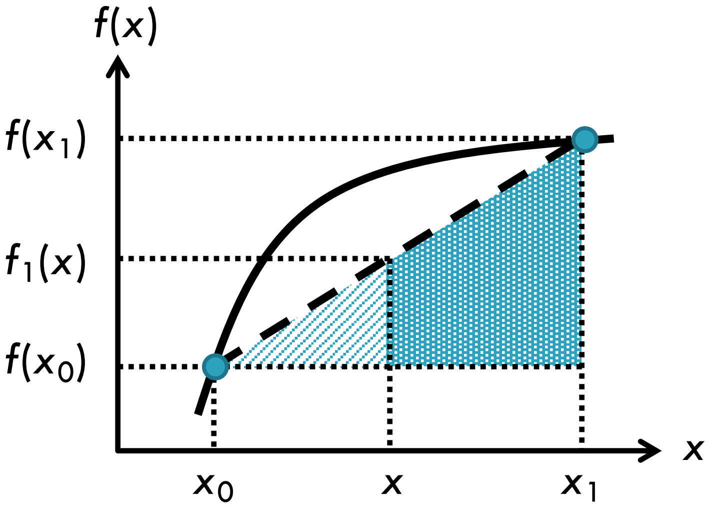


kubik

Interpolasi

- Penyelesaian persamaan polinomial tingkat n membutuhkan sejumlah $n + 1$ titik data
- Metode untuk mencari polinomial tingkat n yang merupakan interpolasi sejumlah $n + 1$ titik data:
  - Metode Newton
  - Metode Lagrange

Interpolasi Linear: Metode Newton



$$\frac{f_1(x)-f(x_0)}{x-x_0}=\frac{f(x_1)-f(x_0)}{x_1-x_0}$$

$$f_1(x)=f(x_0)+\frac{f(x_1)-f(x_0)}{x_1-x_0}(x-x_0)$$

Interpolasi Kuadratik: Metode Newton

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

$$= b_0 + b_1 x - b_1 x_0 + b_2 x^2 + b_2 x_0 x_1 - b_2 x x_0 - b_2 x x_1$$

$$= (b_0 - b_1 x_0 + b_2 x_0 x_1) + (b_1 - b_2 x_0 - b_2 x_1)x + (b_2)x^2$$

$$f_2(x) = a_0 + a_1 x + a_2 x^2 \quad \begin{cases} a_1 = b_0 - b_1 x_0 + b_2 x_0 x_1 \\ a_2 = b_1 - b_2 x_0 - b_2 x_1 \\ a_2 = b_2 \end{cases}$$

Interpolasi Kuadratik: Metode Newton

$$b_0 = f(x_0)$$

$$b_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = f[x_1, x_0]$$

$$b_1 = \frac{\dfrac{f(x_2) - f(x_1)}{x_2 - x_1} - \dfrac{f(x_1) - f(x_0)}{x_1 - x_0}}{x_2 - x_0} = f[x_2, x_1, x_0] = \frac{f[x_2, x_1] - f[x_1, x_0]}{x_2 - x_0}$$

Interpolasi Polinomial: Metode Newton

$$f_n(x) = b_0 + b_1(x - x_0) + \cdots + b_n(x - x_0)(x - x_1)\cdots(x - x_{n-1})$$

$$b_0 = f(x_0)$$

$$b_1 = f[x_1, x_0]$$

$$b_2 = f[x_2, x_1, x_0]$$

.

.

.

$$b_n = f[x_n, x_{n-1}, \cdots, x_1, x_0]$$

Interpolasi Polinomial: Metode Newton

$$f[x_i, x_j] = \frac{f(x_i) - f(x_j)}{x_i - x_j}$$

$$f[x_i, x_j, x_k] = \frac{f[x_i, x_j] - f[x_j, x_k]}{x_i - x_k}$$

$$f[x_n, x_{n-1}, \ldots, x_1, x_0] = \frac{f[x_n, x_{n-1}, \ldots, x_2, x_1] - f[x_{n-1}, x_{n-2}, \ldots, x_1, x_0]}{x_n - x_0}$$

$$\boxed{\begin{aligned} f_n(x) = f(x_0) + (x - x_0)f[x_1, x_0] + (x - x_0)(x - x_1)f[x_2, x_1, x_0] + \cdots \\ (x - x_0)(x - x_1)\cdots(x - x_{n-1})\, f[x_n, x_{n-1}, \ldots, x_0] \end{aligned}}$$

Interpolasi Polinomial: Metode Newton

i	x_i	$f(x_i)$	langkah hitungan		
			ke-1	ke-2	ke-3
0	x_0	$f(x_0)$	$f[x_1,x_0]$	$f[x_2,x_1,x_0]$	$f[x_3,x_2,x_1,x_0]$
1	x_1	$f(x_1)$	$f[x_2,x_1]$	$f[x_3,x_2,x_1]$	
2	x_2	$f(x_2)$	$f[x_3,x_2]$		
3	x_3	$f(x_3)$			

Interpolasi Polinomial: Metode Lagrange

$$f_n(x) = \sum_{i=0}^{n} L_i(x) f(x_i) \qquad L_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^{n} \frac{x - x_k}{x_i - x_k}$$

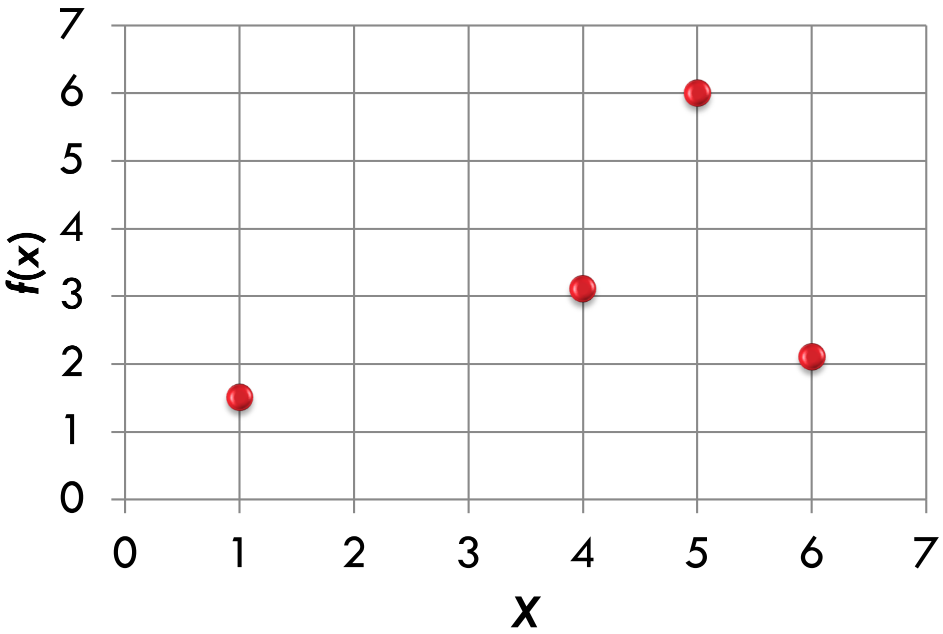
Contoh interpolasi polinomial order 3:

$$f_3(x) = \left[\frac{x - x_1}{x_0 - x_1} \frac{x - x_2}{x_0 - x_2} \frac{x - x_3}{x_0 - x_3}\right] f(x_0) + \left[\frac{x - x_0}{x_1 - x_0} \frac{x - x_2}{x_1 - x_2} \frac{x - x_3}{x_1 - x_3}\right] f(x_1) + \cdots$$

$$+ \left[\frac{x - x_0}{x_2 - x_0} \frac{x - x_1}{x_2 - x_1} \frac{x - x_3}{x_2 - x_3}\right] f(x_2) + \left[\frac{x - x_0}{x_3 - x_0} \frac{x - x_1}{x_3 - x_1} \frac{x - x_2}{x_3 - x_2}\right] f(x_3)$$

Contoh interpolasi

i	x_i	$f(x_i)$
0	1	1.5
1	4	3.1
2	5	6
3	6	2.1



Sekian