Universitas Gadjah Mada
Fakultas Teknik
Departemen Teknik Sipil dan Lingkungan
Prodi Magister Teknik Pengelolaan Bencana Alam

Teknik Pengolahan Data

Curve Fitting: Regresi dan Interpolasi

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
  - Uji hipotesis

Curve Fitting

- 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 yang dicari tidak perlu melewati setiap titik data
- 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

Curve Fitting

- Ekstrapolasi
  - Mirip dengan interpolasi, tetapi untuk memperkirakan nilai-nilai di luar kisaran titik-titik data
  - Ekstrapolasi tidak disarankan

Curve Fitting terhadap Data Pengukuran

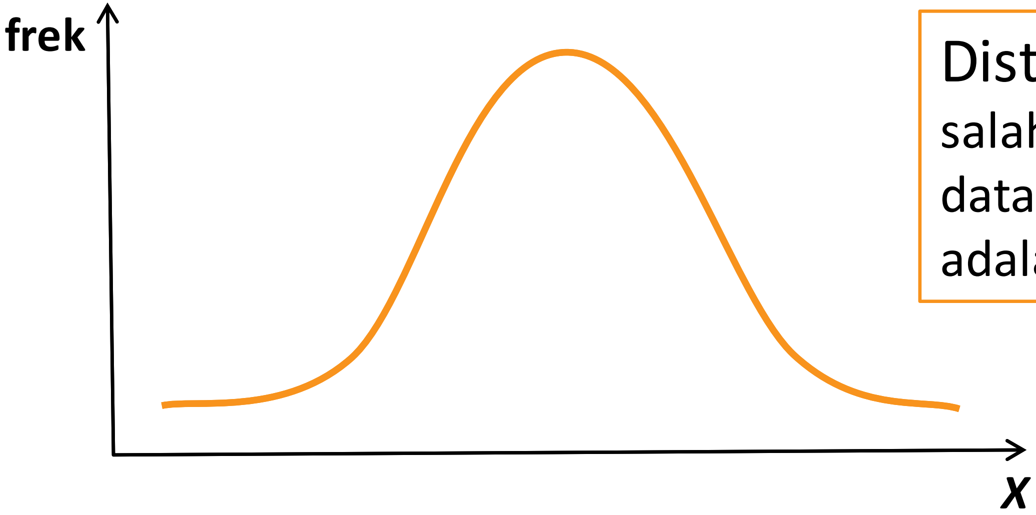
- Analisis pola perilaku data
  - Pemanfaatan pola data (pengukuran, eksperimen) untuk melakukan perkiraan
  - Apabila data persis (akurat): interpolasi
  - Apabila data tak persis (tak akurat): regresi
- Uji hipotesis
  - Pembandingan antara hasil teori atau hasil hitungan dengan hasil pengukuran

Beberapa Besaran Statistik

merepresentasikan sebaran data

- Rata-rata aritmatik, *mean* ⇨ $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$
- Deviasi standar, simpangan baku, *standard deviation* ⇨ $s_Y = \sqrt{\frac{S_t}{n-1}}$ $\quad S_t = \sum_{i=1}^{n}\left(y_i - \bar{Y}\right)^2$
- Varian (‘ragam’), *variance* ⇨ ${s_Y}^2 = \frac{S_t}{n-1}$
- *Coefficient of variation* ⇨ $c.v. = \frac{s_Y}{\bar{Y}} \times 100\%$

Distribusi Probabilitas



Distribusi Normal

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

Regresi dan Interpolasi

Regresi 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
- Metode
  - Regresi linear
  - Regresi persamaan-persamaan tak-linear yang dilinearkan
  - Regresi polinomial
  - Regresi linear ganda
  - Regresi tak-linear

Regresi: Metode Kuadrat Terkecil

- Bagaimana caranya?
  - Program komputer
  - *Spreadsheet* (Microsoft Excel)
- Program aplikasi gratis, mirip MatLab
  - Octave
  - Scilab
  - Freemat

Regresi Linear

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

  $y_{reg} = a_0 + a_1 x$

  $a_0 \;\; intercept$

  $a_1 \;\; slope$

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

$$e = y - y_{reg} = y - (a_0 + a_1 x) = y - a_0 - a_1 x$$

- 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

- Cara mendapatkan koefisien a_0 dan a_1
  - Diferensialkan persamaan tersebut dua kali, masing-masing terhadap a_0 dan a_1

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

  - Samakan kedua persamaan hasil diferensiasi tersebut dengan nol

$$\sum(y_i - a_0 - a_1 x_i) = 0 \qquad \sum(y_i - a_0 - a_1 x_i)x_i = 0$$

Regresi Linear

- Cara mendapatkan koefisien a_0 dan a_1
  - Selesaikan persamaan tersebut untuk mendapatkan a_0 dan a_1

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

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

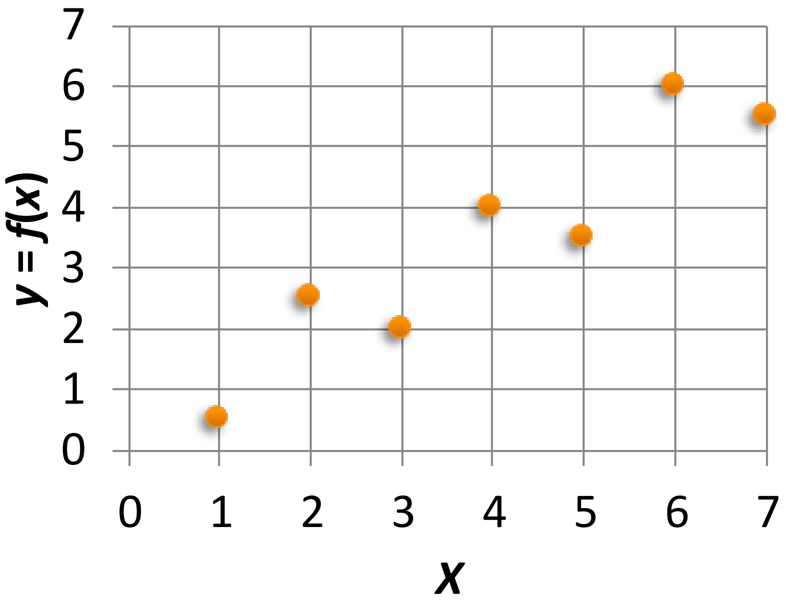
  - Dalam hal ini, $\bar{y}$ dan $\bar{x}$ masing-masing adalah nilai *y* rerata dan *x* rerata

Contoh #1

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 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
∑ =	**28**	**24.0**	**119.5**	**140**	**∑ =**	**2.991071**	**22.71429**

Hitungan Regresi Linear

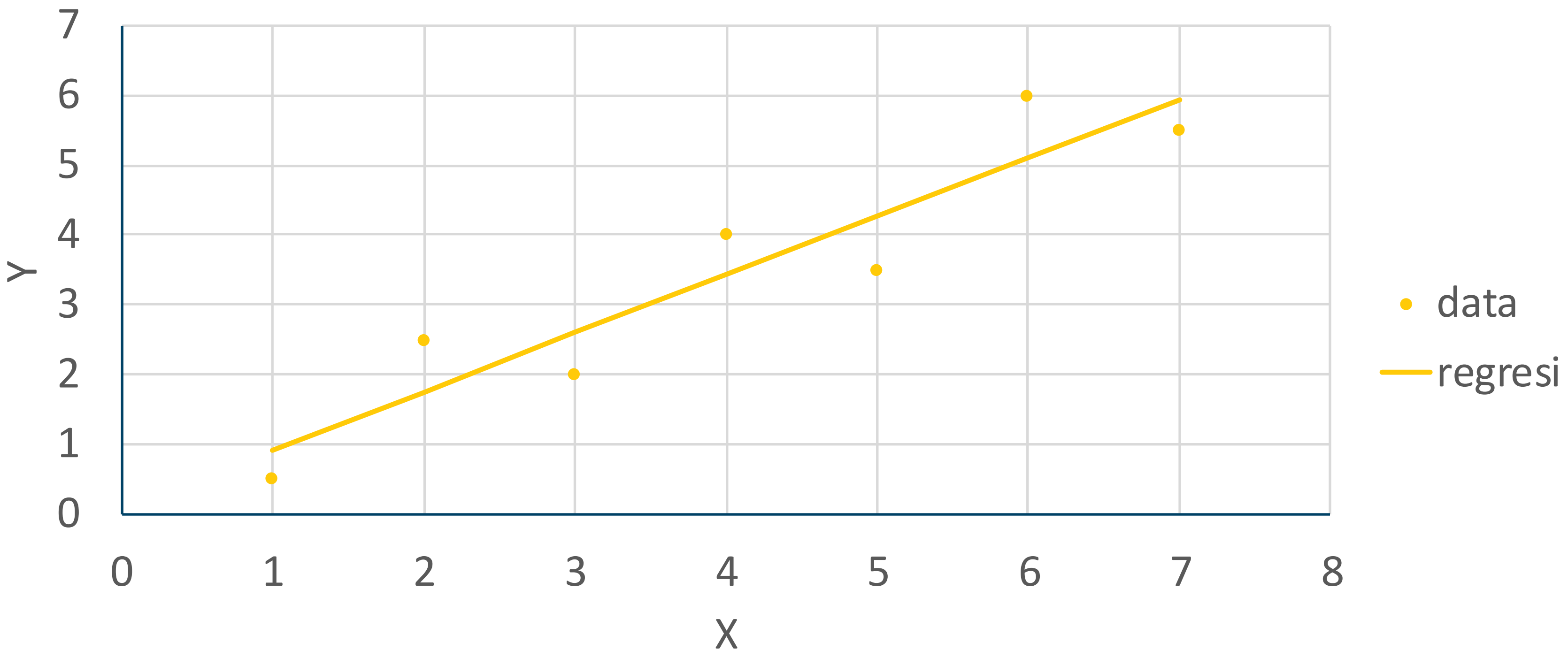
$$a_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum {x_i}^2 - \left(\sum x_i\right)^2} = \frac{7 \times 119.5 - 28 \times 24}{7 \times 140 - 28^2} = 0.839286$$

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

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

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

Grafik 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

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

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

Hitungan Regresi Linear

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

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

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

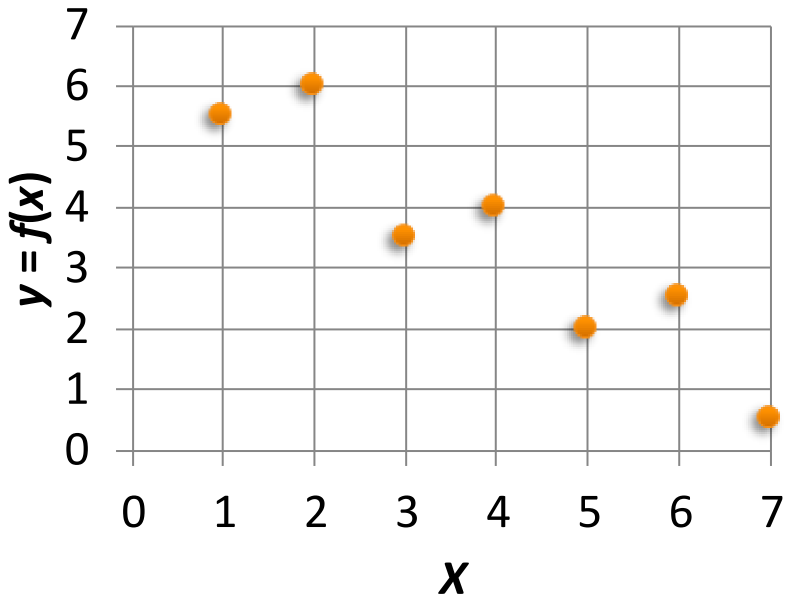
$$r = 0.931836$$

Contoh #2

Tabel data

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

Grafik/kurva data



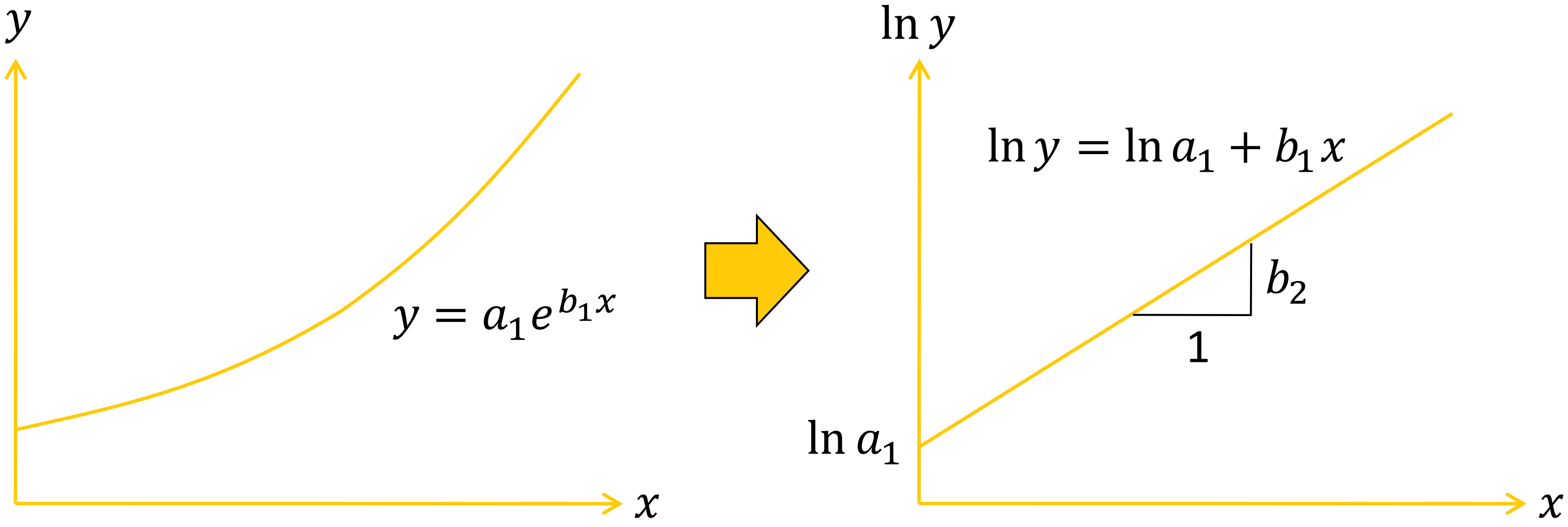
Regresi dan Interpolasi

Regresi persamaan non-linear

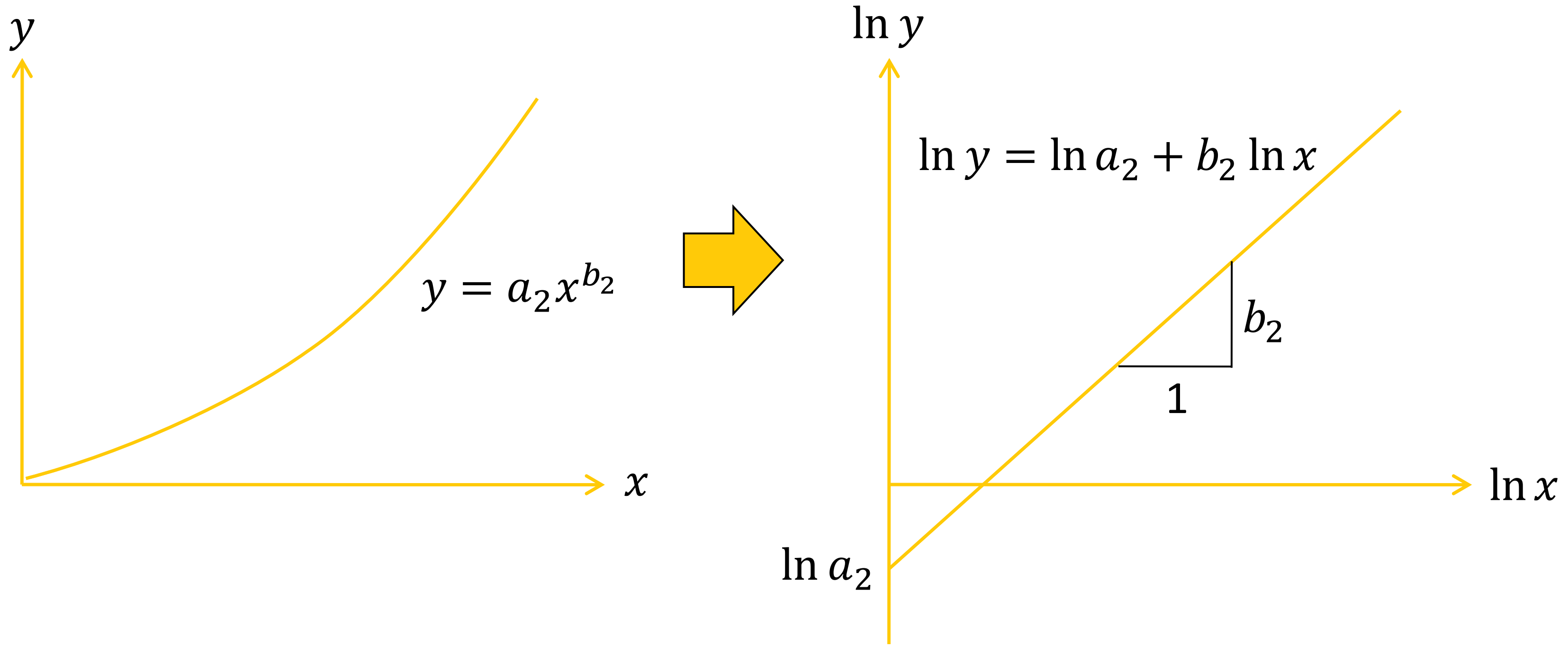
Regresi Linear

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

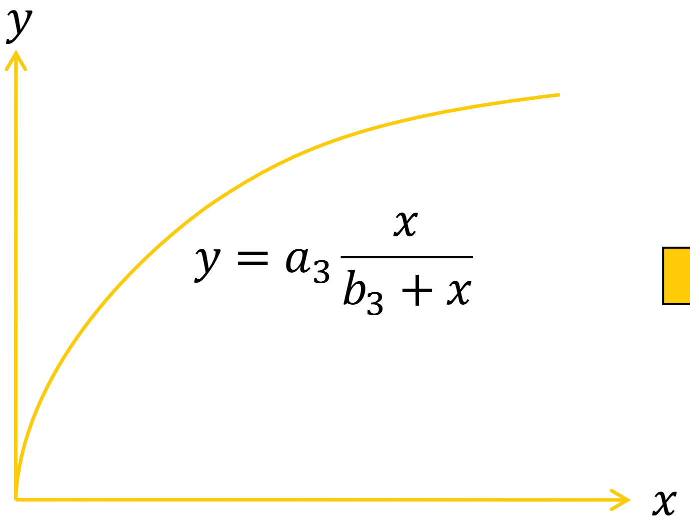
Linearisasi Persamaan Non-linear

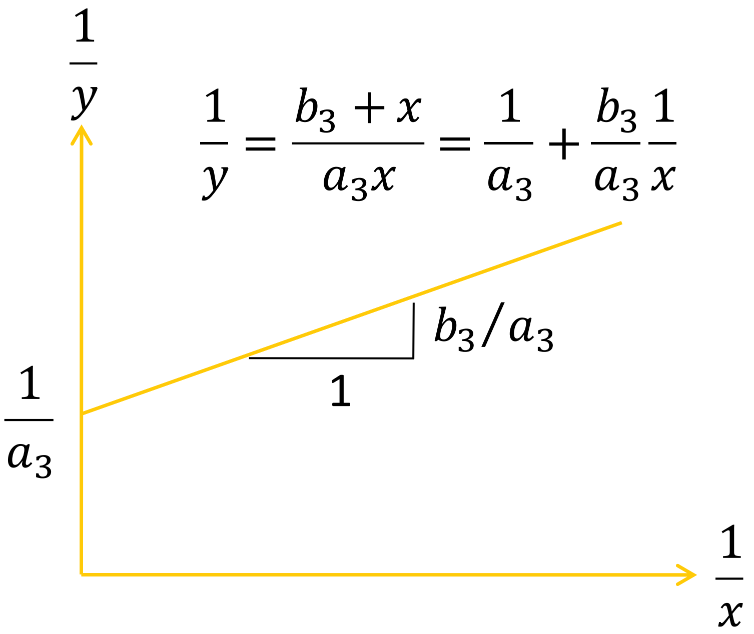


Linearisasi Persamaan Non-linear



Linearisasi Persamaan Non-linear





Regresi dan Interpolasi

Regresi Polinomial

Regresi Polinomial

- Sebagian data bidang teknik, walaupun menunjukkan pola yang jelas, namun pola tsb tidak dapat diwakili oleh sebuah garis lurus
  - Metode 1: transformasi koordinat (linearisasi persamaan non-linear)
  - Metode 2: regresi polinomial
    - Polinomial tingkat m

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

    - Jumlah residu kuadrat

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

- Metode kuadrat terkecil yang diperluas untuk regresi polinomial tingkat *m*

$$\frac{\partial S_r}{\partial a_0} = -2\sum_{i=1}^{n}\left(y_i - a_0 - a_1x_i - a_2{x_i}^2 - \cdots - a_m{x_i}^m\right)$$

$$\frac{\partial S_r}{\partial a_1} = -2\sum_{i=1}^{n}x_i\left(y_i - a_0 - a_1x_i - a_2{x_i}^2 - \cdots - a_m{x_i}^m\right)$$

$$\frac{\partial S_r}{\partial a_2} = -2\sum_{i=1}^{n}{x_i}^2\left(y_i - a_0 - a_1x_i - a_2{x_i}^2 - \cdots - a_m{x_i}^m\right)$$

.
.
.

$$\frac{\partial S_r}{\partial a_m} = -2\sum_{i=1}^{n}{x_i}^m\left(y_i - a_0 - a_1x_i - a_2{x_i}^2 - \cdots - a_m{x_i}^m\right)$$

- Persamaan-persamaan tersebut disamakan dengan nol dan disusun sedemikian rupa menjadi sistem persamaan linear

$$\sum_{i=1}^{n}\left(y_i - a_0 - a_1 x_i - a_2 {x_i}^2 - \cdots - a_m {x_i}^m\right) = 0$$

$$\sum_{i=1}^{n} x_i\left(y_i - a_0 - a_1 x_i - a_2 {x_i}^2 - \cdots - a_m {x_i}^m\right) = 0$$

$$\sum_{i=1}^{n} {x_i}^2\left(y_i - a_0 - a_1 x_i - a_2 {x_i}^2 - \cdots - a_m {x_i}^m\right) = 0$$

.

.

.

$$\sum_{i=1}^{n} {x_i}^m\left(y_i - a_0 - a_1 x_i - a_2 {x_i}^2 - \cdots - a_m {x_i}^m\right) = 0$$

- Ada m+1 persamaan linear dengan m+1 variabel tak diketahui, yaitu a_0, a_1, a_2, …, a_m

$$a_0 n + a_1 \sum_{i=1}^{n} x_i + a_2 \sum_{i=1}^{n} {x_i}^2 + \cdots + a_m \sum_{i=1}^{n} {x_i}^m = \sum_{i=1}^{n} y_i$$

$$a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} {x_i}^2 + a_2 \sum_{i=1}^{n} {x_i}^3 + \cdots + a_m \sum_{i=1}^{n} {x_i}^{m+1} = \sum_{i=1}^{n} x_i y_i$$

$$a_0 \sum_{i=1}^{n} {x_i}^2 + a_1 \sum_{i=1}^{n} {x_i}^3 + a_2 \sum_{i=1}^{n} {x_i}^4 + \cdots + a_m \sum_{i=1}^{n} {x_i}^{m+2} = \sum_{i=1}^{n} {x_i}^2 y_i$$

$$\vdots$$

$$a_0 \sum_{i=1}^{n} {x_i}^m + a_1 \sum_{i=1}^{n} {x_i}^{m+1} + a_2 \sum_{i=1}^{n} {x_i}^{m+2} + \cdots + a_m \sum_{i=1}^{n} {x_i}^{2m} = \sum_{i=1}^{n} {x_i}^m y_i$$

- Persamaan-persamaan linear ini dapat diselesaikan dengan metode
  - Eliminasi Gauss
  - Gauss-Jordan
  - Iterasi Jacobi
  - Inversi matriks

- Inversi matriks

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} {x_i}^2 & \dots & \sum_{i=1}^{n} {x_i}^m \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} {x_i}^2 & \sum_{i=1}^{n} {x_i}^3 & \dots & \sum_{i=1}^{n} {x_i}^{m+1} \\ \sum_{i=1}^{n} {x_i}^2 & \sum_{i=1}^{n} {x_i}^3 & \sum_{i=1}^{n} {x_i}^4 & \dots & \sum_{i=1}^{n} {x_i}^{m+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^{n} {x_i}^m & \sum_{i=1}^{n} {x_i}^{m+1} & \sum_{i=1}^{n} {x_i}^{m+2} & \dots & \sum_{i=1}^{n} {x_i}^{2m} \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} {x_i}^2 y_i \\ \vdots \\ \sum_{i=1}^{n} {x_i}^m y_i \end{Bmatrix}$$

- Inversi matriks

$$
\begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{Bmatrix} =
\begin{bmatrix}
n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} {x_i}^2 & \dots & \sum_{i=1}^{n} {x_i}^m \\
\sum_{i=1}^{n} x_i & \sum_{i=1}^{n} {x_i}^2 & \sum_{i=1}^{n} {x_i}^3 & \dots & \sum_{i=1}^{n} {x_i}^{m+1} \\
\sum_{i=1}^{n} {x_i}^2 & \sum_{i=1}^{n} {x_i}^3 & \sum_{i=1}^{n} {x_i}^4 & \dots & \sum_{i=1}^{n} {x_i}^{m+2} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\sum_{i=1}^{n} {x_i}^m & \sum_{i=1}^{n} {x_i}^{m+1} & \sum_{i=1}^{n} {x_i}^{m+2} & \dots & \sum_{i=1}^{n} {x_i}^{2m}
\end{bmatrix}^{-1}
\begin{Bmatrix}
\sum_{i=1}^{n} y_i \\
\sum_{i=1}^{n} x_i y_i \\
\sum_{i=1}^{n} {x_i}^2 y_i \\
\vdots \\
\sum_{i=1}^{n} {x_i}^m y_i
\end{Bmatrix}
$$

Contoh

- Temukanlah kurva polinomial tingkat 2 yang mewakili pola sebaran data pada tabel di sisi kanan ini

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

- Jawab

  $y = 2.47857 + 2.35929x + 1.86071x^2$

  $r^2 = 1 - \frac{S_r}{S_t} = 1 - \frac{3.74657}{2513.39} = 0.99851$

  $r = 0.99925$

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

Regresi dan Interpolasi

Regresi Linear Ganda (*Multiple Linear Regression*)

Regresi Linear Ganda

- Misal variabel y adalah fungsi linear dua variabel bebas x_1 dan x_2

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

- Koefisien a_0, a_1, a_2 dalam persamaan di atas dapat ditemukan dengan metode kuadrat terkecil kesalahan

$$S_r = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_1 - a_2 x_2)^2$$

Regresi Linear Ganda

- Diferensial parsial persamaan tersebut terhadap masing-masing koefisien

$$\frac{\partial S_r}{\partial a_0} = -2\sum_{i=1}^{n}(y_i - a_0 - a_1 x_{1i} - a_2 x_{2i})$$

$$\frac{\partial S_r}{\partial a_1} = -2\sum_{i=1}^{n}x_{1i}(y_i - a_0 - a_1 x_{1i} - a_2 x_{2i})$$

$$\frac{\partial S_r}{\partial a_2} = -2\sum_{i=1}^{n}x_{2i}(y_i - a_0 - a_1 x_{1i} - a_2 x_{2i})$$

Regresi Linear Ganda

- Samakan persamaan diferensial tsb dengan nol dan atur suku-suku dalam persamaan

$$na_0 + \sum_{i=1}^{n} x_{1i}a_1 + \sum_{i=1}^{n} x_{2i}a_2 = \sum_{i=1}^{n} y_i$$

$$\sum_{i=1}^{n} x_{1i}a_0 + \sum_{i=1}^{n} x_{1i}x_{1i}a_1 + \sum_{i=1}^{n} x_{1i}x_{2i}a_2 = \sum_{i=1}^{n} x_{1i}y_i$$

$$\sum_{i=1}^{n} x_{2i}a_0 + \sum_{i=1}^{n} x_{2i}x_{1i}a_1 + \sum_{i=1}^{n} x_{2i}x_{2i}a_2 = \sum_{i=1}^{n} x_{2i}y_i$$

Regresi Linear Ganda

- Persamaan-persamaan linear tersebut dapat dituliskan dalam bentuk persamaan matriks

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{2i} \\ \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{1i}x_{1i} & \sum_{i=1}^{n} x_{2i}x_{1i} \\ \sum_{i=1}^{n} x_{2i} & \sum_{i=1}^{n} x_{2i}x_{1i} & \sum_{i=1}^{n} x_{2i}x_{2i} \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{1i}y_i \\ \sum_{i=1}^{n} x_{2i}y_i \end{Bmatrix}$$

Regresi Linear Ganda

- Inversi matriks

$$\begin{Bmatrix} a_0 \\ a_1 \\ a_2 \end{Bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{2i} \\ \sum_{i=1}^{n} x_{1i} & \sum_{i=1}^{n} x_{1i}x_{1i} & \sum_{i=1}^{n} x_{2i}x_{1i} \\ \sum_{i=1}^{n} x_{2i} & \sum_{i=1}^{n} x_{2i}x_{1i} & \sum_{i=1}^{n} x_{2i}x_{2i} \end{bmatrix}^{-1} \begin{Bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{1i}y_i \\ \sum_{i=1}^{n} x_{2i}y_i \end{Bmatrix}$$

Contoh

- Temukanlah persamaan linear yang mewakili pola sebaran data dalam tabel di samping ini
- Jawab

$$y = 5 + 4x_1 - 3x_2$$

$$r^2 = 1$$

x_1	x_2	y
0	0	5
2	1	10
2.5	2	9
1	3	0
4	6	3
7	2	27

Regresi Linear Ganda

- Regresi linear ganda dapat dipakai dalam kasus hubungan antar variabel yang berupa persamaan pangkat (*power equations*)

$$y = a_0 {x_1}^{a_1} {x_2}^{a_2} \dots {x_m}^{a_m}$$

  - Persamaan di atas sangat bermanfaat pada kasus *fitting* data eksperimen
  - Persamaan di atas ditransformasikan menjadi persamaan linear

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

Regresi dan Interpolasi

Bentuk Umum Persamaan Regresi Linear (Metode Kuadrat Terkecil)

Regresi Linear (Kuadrat Terkecil)

- Tiga jenis regresi yang telah dipaparkan, yaitu regresi linear, regresi polinomial, dan regresi linear ganda dapat dituliskan dalam bentuk umum model kuadrat terkecil

$$y = a_0 z_0 + a_1 z_1 + a_2 z_2 + \ldots + a_m z_m$$

  - $z_0, z_1, \ldots, z_m$ adalah fungsi-fungsi yang berjumlah $m + 1$
  - $m + 1$ adalah jumlah variabel bebas
  - n adalah jumlah data

- Persamaan di atas dapat dituliskan dalam bentuk persamaan matriks

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

Regresi Linear (Kuadrat Terkecil)

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

$$[Z] = \begin{bmatrix} z_{01} & z_{11} & . & . & . & z_{m1} \\ z_{11} & z_{12} & . & . & . & a_{m2} \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ a_{0n} & a_{1n} & . & . & . & a_{mn} \end{bmatrix}$$

- $\{Y\}$ adalah vektor kolom variabel tak bebas
- $[Z]$ adalah matriks data nilai variabel bebas
- $\{A\}$ adalah vektor kolom koefisien yang tidak diketahui

$$S_r = \sum_{i=1}^{n} \left(y_i - \sum_{j=0}^{m} a_j z_{ji} \right)^2$$

Regresi Linear (Kuadrat Terkecil)

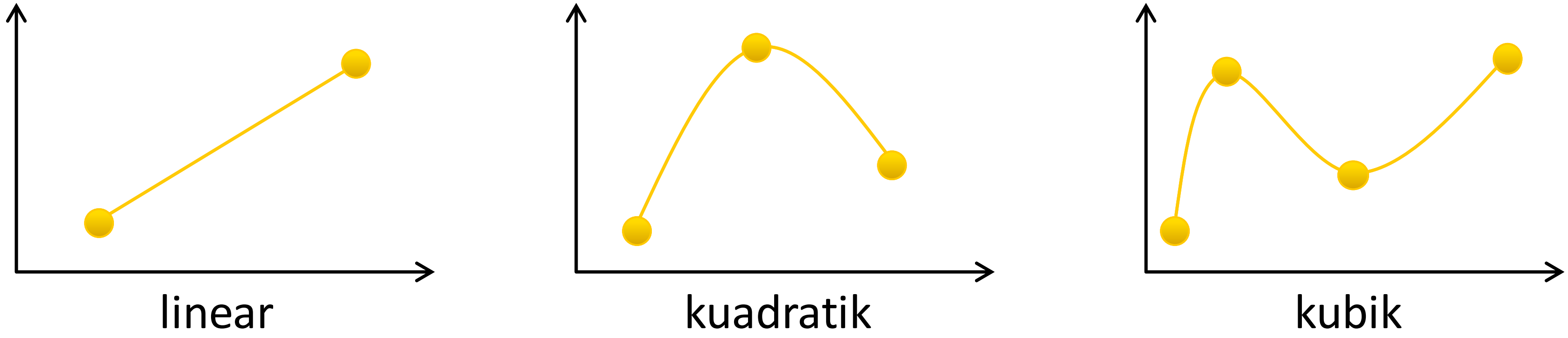
$$[Z]^T[Z]\{A\} = [Z]^T\{Y\}$$

- Strategi penyelesaian
  - Dekomposisi LU
  - Metode Cholesky
  - Inversi matriks $\longrightarrow$ $\{A\} = \left[[Z]^T[Z]\right]^{-1}[Z]^T\{Y\}$

Regresi dan Interpolasi

Interpolasi

Interpolasi



Interpolasi

- Situasi
  - Keperluan untuk memperkirakan nilai variabel di antara data akurat yang diketahui
  - Metode yang paling sering dipakai untuk keperluan tersebut adalah interpolasi polinomial
- Bentuk umum persamaan polinomial tingkat n

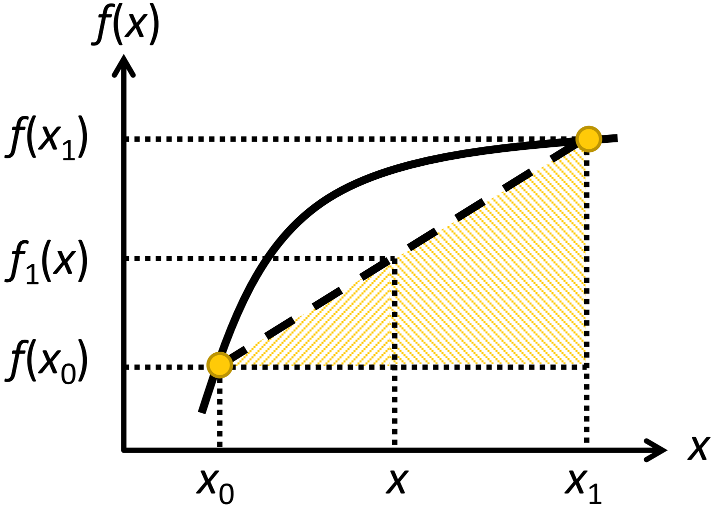
$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

- Hanya ada satu polinomial tingkat n atau tingkat yang lebih kecil yang melalui semua $n + 1$ titik data

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

$$b_0 = f(x_0)$$

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

$$b_2 = \frac{\frac{f(x_2) - f(x_1)}{x_2 - x_1} - \frac{f(x_1) - f(x_0)}{x_1 - x_0}}{x_2 - x_0} = f[x_2, x_1, x_0] = \frac{f[x_2, x_1] - f[x_1, x_0]}{x_2 - x_0}$$

Interpolasi Polinomial: Metode Newton

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

$$\Downarrow$$

$$b_0 = f(x_0)$$
$$b_1 = f[x_1, x_0]$$
$$b_2 = f[x_2, x_1, x_0]$$
$$\vdots$$
$$b_n = f[x_n, x_{n-1}, \cdots, x_1, x_0]$$

Interpolasi Polinomial: Metode Newton

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

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

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

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

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

$$f_n(x) = \sum_{i=0}^{n} L_i(x) f(x_i)$$

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

Interpolasi Polinomial: Metode Lagrange

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

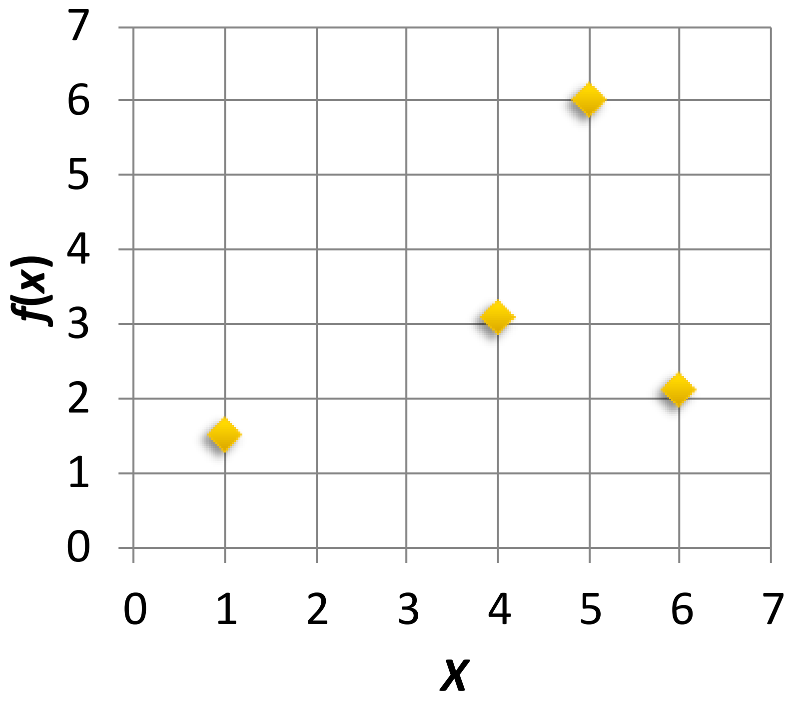
- Contoh interpolasi polinomial order 3:

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

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

Contoh interpolasi

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



Regresi dan Interpolasi

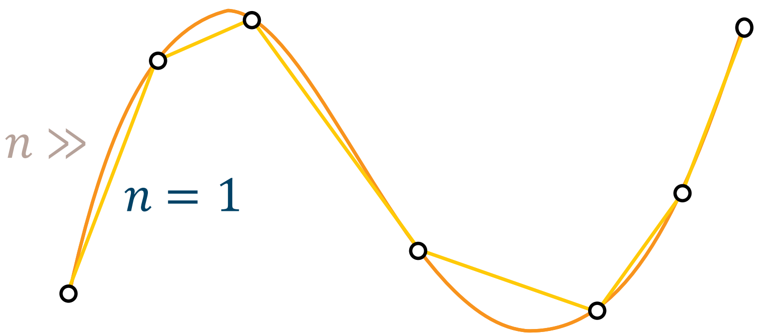
Spline

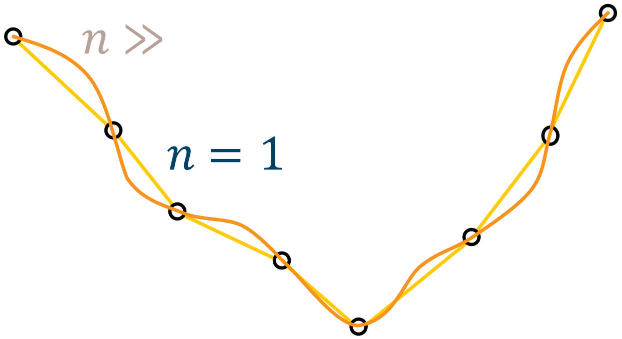
Interpolasi: Spline

- Jumlah titik data $n+1 \rightarrow$ interpolasi polinomial tingkat n
  - Tingkat besar, $n \gg$, mengalami kesulitan apabila titik-titik data menunjukkan adanya perubahan tiba-tiba di suatu titik tertentu (perubahan gradien secara tiba-tiba)
  - Dalam situasi tsb, polinomial tingkat kecil, $n \ll$, dapat lebih representatif untuk mewakili pola data
  - Spline
    - *Cubic splines* $(n = 3)$
    - *Quadratic splines*
    - *Linear splines*

Interpolasi Polinomial vs *Spline*

- Polinomial tingkat n





Linear Splines

- *Spline* tingkat 1 : garis lurus
- Data urut : $x_0, x_1, x_3, \dots, x_n$

$$f(x) = f(x_0) + m_0(x - x_0) \qquad x_0 \le x \le x_1$$
$$f(x) = f(x_1) + m_1(x - x_1) \qquad x_1 \le x \le x_2$$

.

.

.

$$f(x) = f(x_{n-1}) + m_{n-1}(x - x_{n-1}) \qquad x_{n-1} \le x \le x_n$$

Gradien:

$$m_i = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}$$

Linear Splines

- *Linear spline*
  - Dengan demikian, *linear spline* adalah sama dengan interpolasi linear
  - Kekurangan *linear spline* adalah ketidak-mulusan kurva interpolasi
  - Terdapat perubahan *slope* yang sangat tajam di titik-titik data atau di titik-titik pertemuan kurva *spline* (*knot*)
  - Derivatif pertama fungsi *linear spline* diskontinu di titik-titik knot
  - Kelemahan *linear spline* tersebut diatasi dengan pemakaian polinomial yang memiliki tingkat lebih tinggi yang menjamin kemulusan kurva *spline* di *knots* dengan cara menyamakan nilai derivatif di titik-titik *knot*

Quadratic Splines

- *Quadratic splines*
  - Untuk mendapatkan kurva yang memiliki diferensial/laju-perubahan ke-m kontinu di titik *knot*, maka diperlukan kurva *spline* yang bertingkat paling kecil $m + 1$
  - Yang paling banyak dipakai adalah *spline* tingkat 3 (*cubic spline*): diferensial pertama dan kedua kontinu di titik-titik *knot*
    - Ketidak-mulusan diferensial ketiga, keempat, dst. umumnya tidak begitu tampak secara visual

Quadratic Splines

- Tujuan: mendapatkan polinomial tingkat 2 untuk setiap interval titik-titik data
- Polinomial tingkat 2 tsb harus memiliki diferensial pertama (laju perubahan) yang kontinu di titik-titik data
- Polinomial tingkat 2:

$$f(x) = a_i x^2 + b_i x^2 + c_i$$

- Untuk $n + 1$ titik data ($i = 1,2, \dots, n$), terdapat n interval, sehingga terdapat $3n$ koefisien yang harus dicari $(a_i, b_i, c_i), i = 1,2, \dots, n$
- Perlu persamaan sejumlah $3n$

Quadratic Splines

- Ke- $3n$ persamaan tsb adalah (1)
  - Kurva *spline* memotong titik-titik data (*knot*): interval $i-1$ dan i bertemu di titik data $\{x_{i-1}, f(x_{i-1})\}$

$$a_{i-1}{x_{i-1}}^2 + b_{i-1}x_{i-1} + c_{i-1} = f(x_{i-1})$$

$$a_i{x_{i-1}}^2 + b_i x_{i-1} + c_i = f(x_{i-1})$$

i = 2, 3, …, *n*

2(*n* – 1) pers.

Quadratic Splines

- Ke- $3n$ persamaan tsb adalah (2)
  - Kurva *spline* di interval pertama memotong titik data pertama ($i = 1$) dan kurva *spline* di interval terakhir memotong titik data terakhir ($i = n$)

$$a_1 {x_0}^2 + b_1 x_0 + c_1 = f(x_0)$$

$$a_n {x_n}^2 + b_n x_n + c_n = f(x_n)$$

2 pers.

Quadratic Splines

- Ke- $3n$ persamaan tsb adalah (3)
  - Diferensial (gradien) kurva *spline* di dua interval berurutan adalah sama di titik data yang bersangkutan

    $f'^{(x)} = 2ax + b \Longrightarrow 2a_{i-1}x_{i-1} + b_{i-1} = 2a_i x_{i-1} + b_i$ *i* = 2, 3, …, *n*

    (*n* – 1) pers.
  - Diferensial kedua (laju perubahan gradien) kurva *spline* di titik data pertama sama dengan nol

    $a_i = 0$ 1 pers.

➡ Konsekuensi: 2 titik data pertama (*i* = 0 dan *i* = 1) dihubungkan dengan garis lurus

Quadratic Splines

- Dengan demikian, jumlah persamaan seluruhnya adalah $2(n-1)+2+(n-1)+1=3n$

Cubic Splines

- Tujuan: mencari polinomial tingkat 3 untuk setiap interval titik-titik data.
  - Polinomial tingkat 3 tsb harus memiliki diferensial pertama (gradien) dan diferensial kedua (laju perubahan gradien) yang kontinu di titik-titik data.
  - Polinomial orde 3:

$$f_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i$$

  - Untuk $(n + 1)$ titik data ($i = 0,1,2, \dots, n$), terdapat n interval, shg. terdapat $4n$ koefisien yang harus dicari (a_i, b_i, c_i, d_i), $i = 1,2, \dots, n$
    - Perlu persamaan sejumlah $4n$

Cubic Splines

- Ke- $4n$ persamaan tsb adalah sbb.
  - Kurva spline memotong titik-titik data (knot): interval $i-1$ dan i bertemu di titik data $\{x_{i-1}, f(x_{i-1})\}$ → (2*n* – 2) pers.
  - Kurva *spline* di interval pertama memotong titik data pertama dan kurva *spline* terakhir memotong titik data terakhir → 2 pers.
  - Diferensial pertama kurva *spline* di dua interval berurutan adalah sama di titik data ybs. → (n – 1) pers.
  - Diferensial kedua kurva *spline* di dua interval berurutan adalah sama di titik data ybs. → (n – 1) pers.
  - Diferensial kedua kurva *spline* di titik data pertama dan terakhir sama dengan nol → 2 pers.

Cubic Splines

- Ke- $4n$ persamaan tsb.
  - Syarat kelima membawa konsekuensi sbb.
    - Kurva *spline* di interval pertama dan interval terakhir berupa garis lurus
      - dua titik data pertama dihubungkan dengan sebuah garis lurus
      - dua titik data terakhir dihubungkan dengan sebuah garis lurus
  - Ada sebuah syarat alternatif sebagai pengganti syarat kelima tsb
    - Derivatif kedua di titik *knot* terakhir diketahui

Cubic Splines

- Diperoleh $4n$ persamaan yang harus diselesaikan untuk mencari $4n$ koefisien, a_i, b_i, c_i, d_i

$$2(n-1)+2+(n-1)+(n-1)+2=4n$$

- Dimungkinkan untuk melakukan manipulasi matematis shg diperoleh suatu teknik cubic splines yang hanya memerlukan $n-1$ penyelesaian (lihat uraian di buku acuan)
    - Chapra, S.P., Canale, R.P., 1985, *Numerical Methods for Engineers*, McGraw-Hill Book Co., New York, hlm. 395-396).

Cubic Splines

2 *unknows* di setiap interval:
$f''(x_{i-1})$ dan $f''(x_i)$

$$f_i(x) = \frac{f''(x_{i-1})}{6(x_i - x_{i-1})}(x_i - x)^3 + \frac{f''(x_{i-1})}{6(x_i - x_{i-1})}(x - x_{i-1})^3 +$$

$$+ \left[\frac{f(x_{i-1})}{(x_i - x_{i-1})} - \frac{f''(x_{i-1})(x_i - x_{i-1})}{6}\right](x_i - x) +$$

$$+ \left[\frac{f(x_i)}{(x_i - x_{i-1})} - \frac{f''(x_{i-1})(x_i - x_{i-1})}{6}\right](x - x_{i-1})$$

$$(x_i - x_{i-1})f''(x_{i-1}) + 2(x_{i+1} - x_{i-1})f''(x_i) + (x_{i+1} - x_i)f''(x_{i+1}) =$$

$$\frac{6}{(x_{i+1} - x_i)}[f(x_{i+1}) - f(x_i)] + \frac{6}{(x_i - x_{i-1})}[f(x_{i-1}) - f(x_i)]$$

$$\left.\begin{array}{r} n \text{ interval} \\ f''(x_0) = 0 \\ f''(x_n) = 0 \end{array}\right\} \Rightarrow (n-1) \text{ pers.}$$

Terima kasih