

# **PENCOCOKAN KURVA:**

## **ANALISA REGRESI**

## **LEAST SQUARE FIT**

- ▶ Pencocokan Kurva
- ▶ Regresi Kuadrat Terkecil (RKT)
  - RKT untuk Kurva Linier
  - RKT untuk Kurva Non-Linier
- ▶ Regresi Polynomial
- ▶ Least Square Fit
- ▶ *Tugas VI*

# Pencocokan Kurva

Seringkali data tersajikan dalam bentuk serangkaian nilai diskrit (deretan angka<sup>2</sup> dalam urutan yang kontinu).

Jika kita tidak mengetahui bentuk fungsi yang menghasilkan data-data tersebut, maka pendekatan terbaik yang dapat kita lakukan adalah dengan men-“generate” fungsi<sup>2</sup> sederhana untuk mengaproksimasi ‘fungsi yang tak diketahui tersebut’.

Mengapa bentuk fungsi begitu penting bagi kita?...

Selain untuk memenuhi kebutuhan proses numeris, seperti integrasi atau mendapatkan solusi pendekatan dari persamaan differensial, seringkali kita harus menganalisa tren atau melakukan pengujian hipotesis terhadap nilai-nilai diskrit yang dihasilkan oleh fungsi tsb.

Atas keberadaan serangkaian data diskrit yang tidak diketahui fungsi penghasilnya, ada 2 hal yang dapat anda lakukan :

Pertama, berusaha mencari bentuk kurva yang mewakili rangkaian data diskrit tersebut.

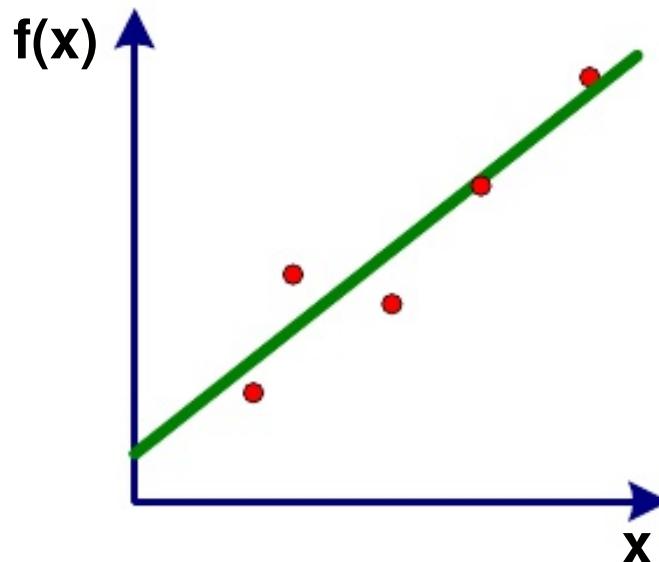
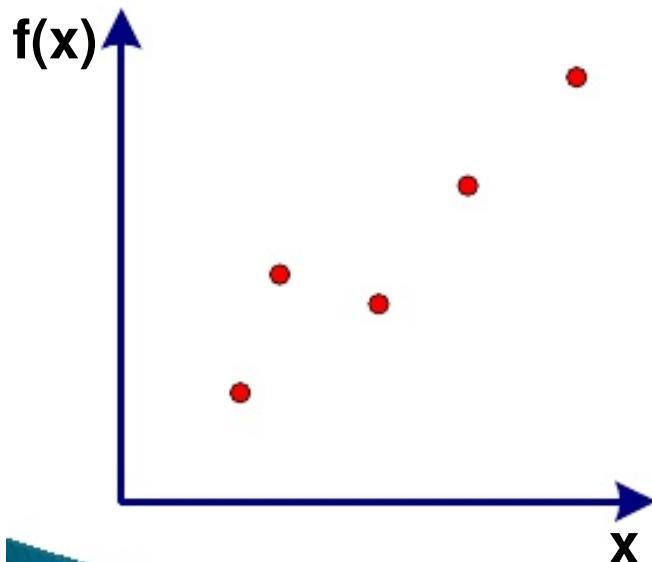
Kedua, berusaha meng-estimasi/memperkirakan nilai data pada titik<sup>2</sup> di antara nilai<sup>2</sup> yang diketahui.

Keduanya dikenal sebagai teknik Pencocokan Kurva (curve fitting).

Pendekatan<sup>2</sup> yang lazim digunakan untuk melakukan Pencocokan Kurva antara lain adalah :

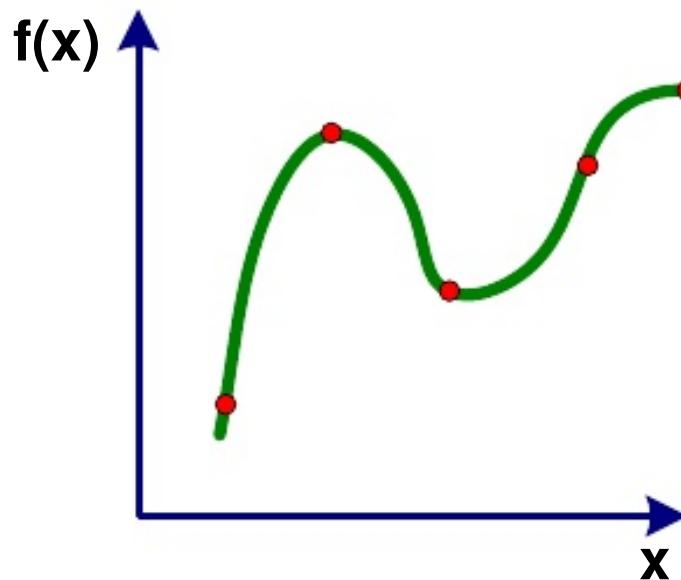
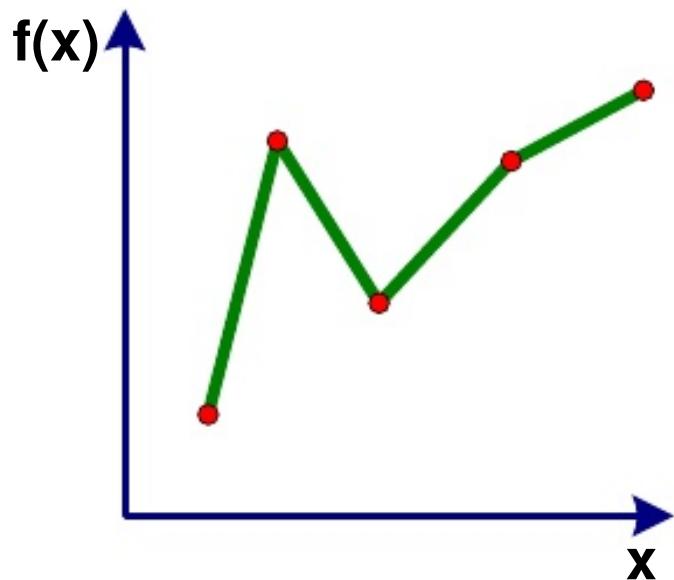
## Regresi Kuadrat Terkecil (*least-square regression*)

Metode ini digunakan apabila data (yang tersaji) memiliki tingkat kesalahan berarti. Anda hanya perlu membuat sebuah kurva yang dapat merepresentasikan tren umum dari data-data tersebut. Kurva ini harus dapat menyatakan kecenderungan umum dari data yang tersaji.



## Interpolasi

Jika tingkat ketelitian data yang disajikan lebih baik, maka metode Interpolasi ini dapat dipakai. Anda dapat menggunakan segmen<sup>2</sup> garis lurus untuk menghubungkan titik<sup>2</sup> data (interpolasi linier). Atau, cara kedua dengan menggunakan kurva untuk menghubungkan titik<sup>2</sup> data.



Analisis regresi menggunakan sedikit notasi dan perhitungan statistik. Ini artinya, ada sedikit yang perlu anda ingat kembali...

# Sedikit Notasi Statistik

Tahun	Debit Air = $y_i$ (m <sup>3</sup> /det)
1991	8,52
1992	3,33
1993	7,85
1994	7,65
1995	10,91
1996	4,17
1997	3,40
1998	8,00
1999	13,4
2000	5,40
2001	8,87
2002	4,73
2003	7,40
2004	6,88
2005	5,00

Rerata/rata-rata data ( $\bar{y}$ ) adalah jumlah nilai data ( $\Sigma y$ ) dibagi jumlah data (n) :

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

Deviasi Standar ( $\sigma$ ) atau penyebaran nilai data adalah :

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}}$$

Atau dapat pula dinyatakan dalam bentuk Varians ( $\sigma^2$ ) :

$$\sigma^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}$$

# Regresi Kuadrat Terkecil

Jika anda tahu bahwa data yang tersaji memiliki tingkat kesalahan yang cukup signifikan, maka penggunaan interpolasi polynomial rasanya bukanlah pilihan yang bijaksana. Karena (kemungkinan besar) hasil pendekatannya akan kurang memuaskan.

Ada cara yang lebih mudah, yaitu dengan membuat kurva yang 'kurang-lebih' dapat mewakili titik<sup>2</sup> data tersebut.

Katakan kurva ini adalah kurva dari fungsi  $g(x)$  yang 'mirip' dengan fungsi aslinya.

Tetapi jika penyebarannya titik<sup>2</sup> datanya sangat besar? ...

Rasanya koq sulit cara di atas bisa berhasil, ya?! ...

Kecuali jika, ... kita membuat kurva  $g(x)$  yang sekaligus mampu meminimalkan perbedaan (selisih) antara titik<sup>2</sup> data dengan kurva  $g(x)$  sendiri.

Teknik ini yang disebut dengan metode Regresi Kuadrat Terkecil.

Secara umum prosedur untuk mengaplikasikan metode Regresi Kuadrat Terkecil (RKT) ini adalah sbb :

1. Titik<sup>2</sup> data diplot ke dalam koordinat cartesian. Dari pola datanya bisa dilihat, apakah kurva pendekatan yang akan kita buat berbentuk garis lurus atau garis lengkung;
2. Tentukanlah sebuah fungsi  $g(x)$  yang kira<sup>2</sup> bisa mewakili fungsi titik<sup>2</sup> data  $f(x)$  :

$$g(x) = a_0 + a_1x + a_2x^2 + \dots + a_rx^r$$

3. Jika  $a_0, a_1, \dots, a_r$  adalah parameter fungsi  $g(x)$ , maka tentukan nilai parameter<sup>2</sup> tsb sdmk hingga garis  $g(x)$  akan mendekati titik<sup>2</sup> data;
4. Jika koordinat titik<sup>2</sup> data adalah  $M(x_i, y_i)$ , maka selisihnya dengan fungsi  $g(x)$  adalah :

$$E_i = M_i G_i$$

$$= y_i - g(x_i; a_0, a_1, \dots, a_r)$$

$$= y_i - (a_0 + a_1x_i + a_2x_i^2 + a_3x_i^3 + \dots + a_rx_i^r)$$

5. Tingkat kesalahan fungsi  $g(x)$  diukur melalui rumus berikut :

$$D^2 = \sum_{i=1}^n E_i^2 = \sum_{i=1}^n \{y_i - g(x_i)\}^2$$

6. Cari nilai parameter  $a_0, a_1, \dots, a_r$  sdmk hingga  $D^2$  dapat seminimum mungkin.  $D^2$  bernilai minimum jika turunan pertamanya terhadap  $a_0, a_1, \dots, a_r = 0$ .

$$\frac{\partial D^2}{\partial a_0} = 0, \quad \frac{\partial D^2}{\partial a_1} = 0, \quad \dots \quad \frac{\partial D^2}{\partial a_r} = 0$$

7. Persamaan no 6 di atas akan memberikan nilai parameter  $a_0, a_1, \dots, a_r$ . Sehingga persamaan kurva terbaik yang mewakili titik<sup>2</sup> data dapat diperoleh.

# RKT untuk Kurva Linier

Permasalahan kita akan menjadi sederhana apabila kurva (yang diwakili oleh titik<sup>2</sup> data) berbentuk garis lurus.

Seperti anda ketahui, bahwa persamaan umum sebuah garis lurus adalah :

$$g(x) = a + bx$$

Dan jumlah-kuadrat-kesalahan dapat dihitung melalui persamaan :

$$D^2 = \sum_{i=1}^n E_i^2 = \sum_{i=1}^n \{y_i - a - bx_i\}^2$$

Agar nilai  $D^2$  dapat seminimal mungkin, maka persamaan jumlah-kuadrat-kesalahan di atas harus diturunkan terhadap parameter  $a$  dan  $b$  (untuk kemudian disamadengankan 0).

Karena turunan pertama thd parameter a dan b masing-masing adalah nol:

$$\frac{\partial D^2}{\partial a} = 0 \text{ dan } \frac{\partial D^2}{\partial b} = 0$$

maka diperoleh:

$$a = \bar{y} - b\bar{x}$$

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

Kedua persamaan ini dapat dimanfaatkan untuk mendapatkan koefisien a dan b, sehingga fungsi  $g(x)$  dapat diperoleh.

Sedangkan untuk mengetahui derajat kesesuaian dari persamaan yang dicari, dapat dihitung melalui koefisien korelasi yang berbentuk :

$$r^2 = \frac{D_t^2 - D^2}{D_t^2}$$

dengan,

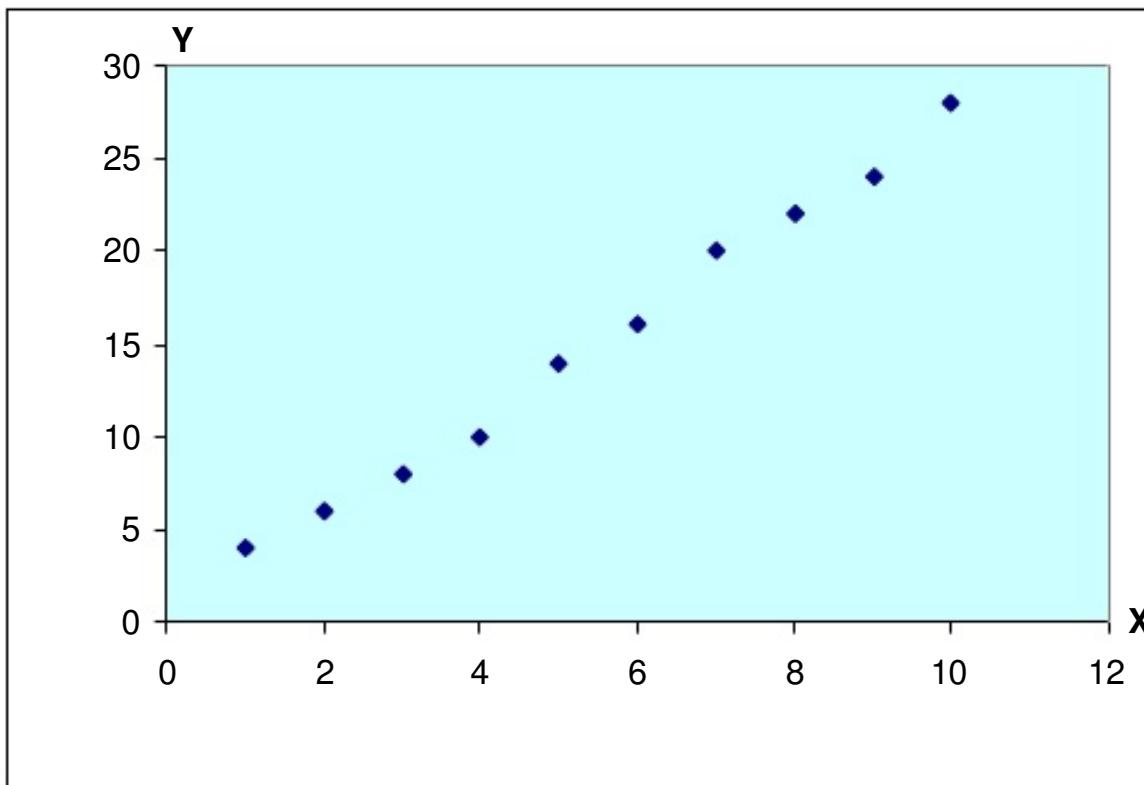
$r$  = koefisien korelasi

$$D_t^2 = \sum_{i=1}^n (y_i - \bar{y})^2 \quad D^2 = \sum_{i=1}^n (y_i - a - bx_i)^2$$

Jika  $r = 1$ , maka fungsi  $\bar{g}(x)$  memiliki tingkat kesesuaian yang tinggi dengan persamaan aslinya ( $f(x)$ ). Atau  $r = 0$  jika sebaliknya.

**contoh :** tentukan persamaan garis yang mewakili data berikut :

x	1	2	3	4	5	6	7	8	9	10
y	4	6	8	10	14	16	20	22	24	28



No	$x_i$	$y_i$	$x_i \cdot y_i$	$x_i^2$
1	1	4	4	1
2	2	6	12	4
3	3	8	24	9
4	4	10	40	16
5	5	14	70	25
6	6	16	96	36
7	7	20	140	49
8	8	22	176	64
9	9	24	216	81
10	10	28	280	100
$\Sigma$	55	152	1058	385

Nilai rerata untuk  $x$  dan  $y$  adalah :

$$\bar{x} = \sum x / n = 55/10 = 5,5$$

$$\bar{y} = \sum y / n = 152/10 = 15,2$$

Jika persamaan umum garis dinyatakan sebagai :  $y = a + bx$ ,  
dan,

$$b = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} = \frac{10 \cdot 1058 - 55 \cdot 152}{10 \cdot 385 - (55)^2} = \frac{2220}{825} = 2,690909$$

$$a = \bar{y} - b\bar{x} = 15,2 - 2,690909 \cdot 5,5 = 0,4$$

Jadi persamaan garis yang mendekati rangkaian data tersebut adalah :

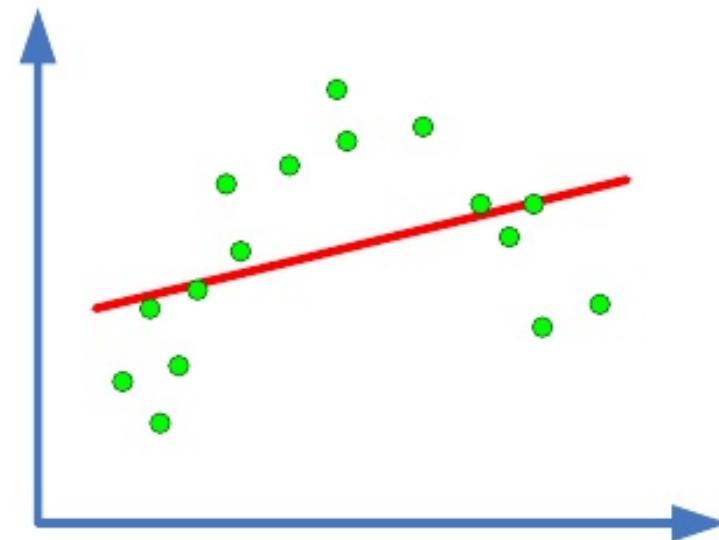
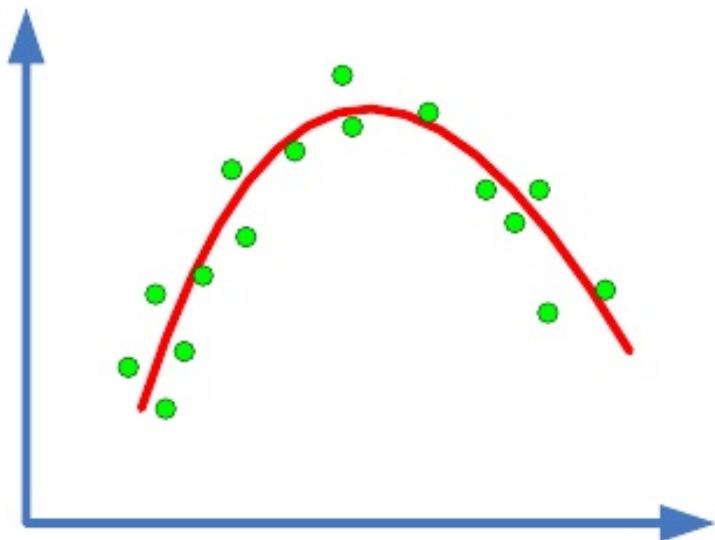
$$y = 0,4 + 2,69x$$

# RKT untuk Kurva Non-Linier

Di dalam praktek akan sering kita jumpai kasus dimana plotting titik<sup>2</sup> data memiliki tren berupa kurva lengkung.

Sehingga persamaan yang sudah dikenalkan pada subbab sebelumnya tidak dapat (langsung) digunakan.

Oleh karenanya perlu dilakukan 'sedikit' transformasi koordinat agar plotting data<sup>2</sup> tersebut dapat dipresentasikan dalam kurva linier.



Tetapi bisa dilihat bahwa kurva lengkung yang didekati dengan garis lurus akan menimbulkan kesalahan yang berarti.

Kecuali untuk beberapa bentuk fungsi yang memang dapat didekati dengan metode Linierisasi Kurva Non-Linier. Fungsi<sup>2</sup> tersebut antara lain :

## 1. Fungsi Eksponensial

$$y = a e^{bx} \quad \text{dengan } a_1 \text{ dan } b_1 \text{ adalah konstanta}$$

persamaan di atas dapat dilinierkan dengan logaritma-natural spt berikut :

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

$$\text{jika } \ln e = 1, \text{ maka } \ln y = \ln a + b x$$

persamaan di atas berbentuk garis lurus dengan kemiringan  $b$ , dan memotong sumbu  $\ln y$  di  $\ln a$ .

## 2. Fungsi Berpangkat

fungsi berpangkat adalah contoh lain fungsi (dengan kurvanya yang) non-linier.

$$y = a x^b \quad \text{dengan } a \text{ dan } b \text{ adalah konstanta}$$

pe-linier-an fungsi di atas juga dapat dilakukan menggunakan persamaan logaritmik spt berikut :

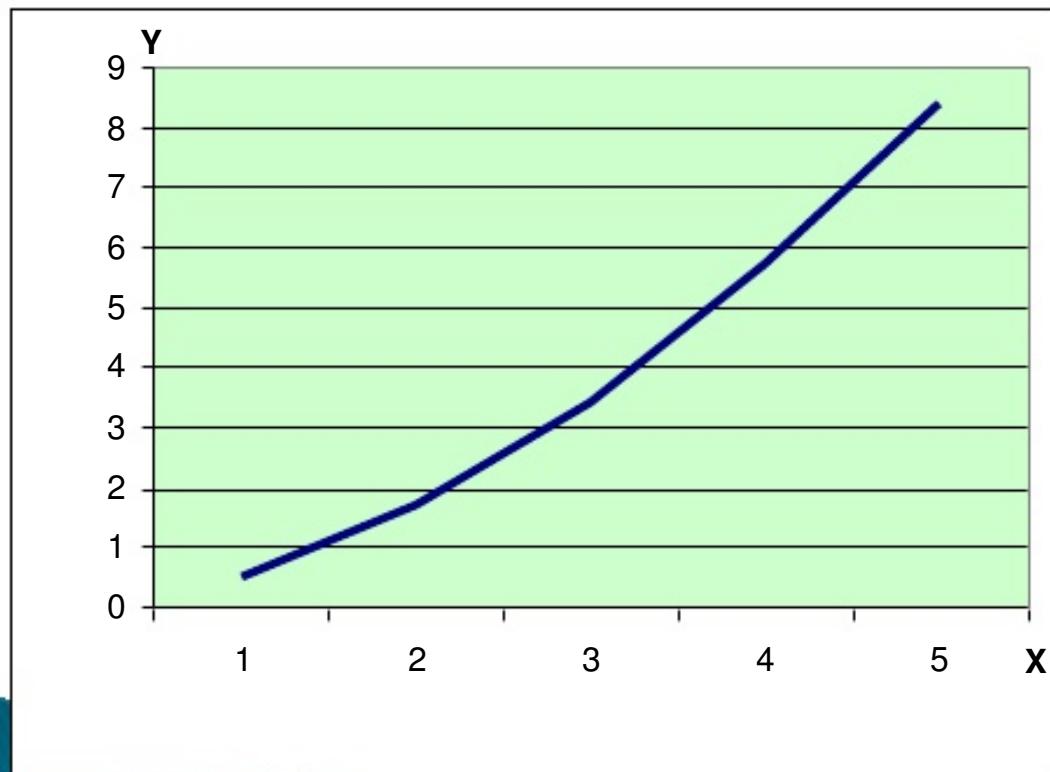
$$\log y = b \log x + \log a$$

persamaan di atas berbentuk garis lurus dengan kemiringan  $b$  dan memotong sumbu  $\log y$  di  $\log a$ .

**Contoh :** Tentukan persamaan kurva lengkung yang diwakili serangkaian data berikut :

x	1	2	3	4	5
y	0,5	1,7	3,4	5,7	8,4

Penyelesaian masalah di atas dilakukan melalui 2 *fashion*, transformasi log dan ln.



## Transformasi log

Misal persamaan yang dicari adalah :  $y = ax^b$

Persamaan tersebut dapat dilinierisasi melalui fungsi logaritmik sbb :  $\log y = \log ax^b$

atau dapat pula dinyatakan sebagai :  $\log y = \log a + b \log x$

Atau jika  $p = \log y$ ;  $A = \log a$ ;  $B = b$ ;  $q = \log x$ ;

Maka persamaan log di atas dapat ditulis menjadi lebih sederhana :  $p = A + Bq$

No	$x_i$	$y_i$	$q_i (= \log x_i)$	$p_i (= \log y_i)$	$q_i p_i$	$q_i^2$
1	1	0,5	0	- 0,301	0	0
2	2	1,7	0,3010	0,2304	0,0693	0,0906
3	3	3,4	0,4771	0,5315	0,2536	0,2276
4	4	5,7	0,6020	0,7559	0,4550	0,3624
5	5	8,4	0,6990	0,9243	0,6461	0,4886
$\Sigma$		19,7	2,0791	2,1411	1,4240	1,1692

Dari tabel tersebut dapat diperoleh beberapa parameter penting, seperti :

$$\bar{y} = \sum y_i / n = 19,7 / 5 = 3,94$$

$$\bar{q} = \sum \log x_i / n = 2,0791 / 5 = 0,4158$$

$$\bar{p} = \sum \log y_i / n = 2,1411 / 5 = 0,42822$$

Sedangkan koefisien A dan B dihitung melalui persamaan (4) dan (6) :

$$B = \frac{n \sum q_i p_i - \sum q_i \sum p_i}{n \sum q_i^2 - (\sum q_i)^2} = \frac{5 (1,4240) - 2,0791 (2,1411)}{5 (1,1692) - 2,0791 (2,0791)} = \frac{2,6684}{1,5233} = 1,75172$$

$$A = \bar{p} - B \bar{q} = 0,42822 - 1,75172 \cdot 0,4158 = - 0,30015$$

Karena  $A = \log a \rightarrow$  maka  $a = 10^A = 0,5010$

Karena  $B = b \rightarrow$  maka  $b = 1,75172$

Dengan demikian fungsi yang dicari adalah :  $y = 0,5010 \times 10^{1,75172}$

## Transformasi ln

Misal persamaan yang dicari adalah :  $y = a e^{bx}$

Persamaan tersebut dapat dilinierisasi melalui fungsi logaritmik sbb :  $\ln y = \ln a e^{bx}$

atau dapat pula dinyatakan sebagai :  $\ln y = \ln a + \ln e^{bx}$  atau  $\ln y = \ln a + bx$

Atau jika  $p = \ln y$ ;  $A = \ln a$ ;  $B = b$ ;  $q = x$ ;

Maka persamaan log di atas dapat ditulis menjadi lebih sederhana :  $p = A + Bq$

No	$x_i = q_i$	$y_i$	$q_i^2 (= x_i^2)$	$p_i (= \ln y_i)$	$q_i p_i$
1	1	0,5	1	- 0,6931	- 0,6931
2	2	1,7	4	0,5306	1,0612
3	3	3,4	9	1,2238	3,6714
4	4	5,7	16	1,7405	6,962
5	5	8,4	25	2,1282	10,641
$\Sigma$	15	19,7	55	4,93	21,6425

Dari tabel tersebut dapat diperoleh beberapa parameter penting, seperti :

$$y = \sum y_i / n = 19,7 / 5 = 3,94$$

$$q = \sum q_i / n = 15 / 5 = 3,0$$

$$p = \sum p_i / n = 4,93 / 5 = 0,986$$

Sedangkan koefisien A dan B dihitung melalui persamaan (4) dan (6) :

$$B = \frac{n \sum q_i p_i - \sum q_i \sum p_i}{n \sum q_i^2 - (\sum q_i)^2} = \frac{5 (21,6425) - 15 (4,93)}{5 (55) - (15)^2} = \frac{34,2625}{50} = 0,6852$$

$$A = p - B q = 0,986 - 0,6852 \cdot 3,0 = -1,0696$$

Karena  $A = \ln a \rightarrow$  maka  $a = e^A = 0,3431$

Karena  $B = b \rightarrow$  maka  $b = 0,6852$

Dengan demikian fungsi yang dicari adalah :  $y = 0,3431 e^{0,6852x}$

Sekarang waktunya memilih. Mana di antara 2 pendekatan yang memberikan akurasi lebih bagus. Caranya adalah dengan menghitungnya melalui nilai koefisien korelasi (7) :

$$r^2 = \frac{D_t^2 - D^2}{D_t^2}$$

dengan

$$D_t^2 = \sum_{i=1}^n (y_i - \bar{y})^2$$

$$D^2 = \sum_{i=1}^n (y_i - ax^b)^2$$

$$D^2 = \sum_{i=1}^n (y_i - ae^{bx})^2$$

No	x <sub>i</sub>	y <sub>i</sub>	Transformasi log			Transformasi ln		
			g(x <sub>i</sub> )	D	D <sub>t</sub>	g(x <sub>i</sub> )	D	D <sub>t</sub>
1	1	0,5	0,4984	0,000003	11,8336	0,6835	0,03367	11,8336
2	2	1,7	1,6848	0,000231	5,0176	1,3563	0,11813	5,0176
3	3	3,4	3,4354	0,00125	0,2916	2,6912	0,50240	0,2916
4	4	5,7	5,6953	0,000022	3,0976	5,3401	0,12953	3,0976
5	5	8,4	8,4296	0,000876	19,8916	10,5963	4,82373	19,8916
$\Sigma$	15	19,7		0,00238	40,132		5,60746	40,132

Dari tabel tersebut dapat dicari nilai r untuk transformasi log :

$$r = \left[ \frac{D_f^2 - D^2}{D_f^2} \right]^{\frac{1}{2}} = \left[ \frac{40,132 - 0,00238}{40,132} \right]^{\frac{1}{2}} = 0,99997$$

Sedangkan untuk transformasi ln :

$$r = \left[ \frac{D_f^2 - D^2}{D_f^2} \right]^{\frac{1}{2}} = \left[ \frac{40,132 - 5,60746}{40,132} \right]^{\frac{1}{2}} = 0,92751$$

Dapat dilihat bahwa **koefisien korelasi r untuk transformasi log lebih besar dibanding transformasi ln**. Sehingga bisa disimpulkan bahwa persamaan yang didapat dari transformasi log memberikan pendekatan yang lebih baik.

# Regresi Polynomial

Persamaan sebuah garis lurus dapat didekati melalui **Metode Kuadrat Terkecil**.

Sementara untuk kurva lengkung, pendekatan yang cukup logis adalah melalui transformasi data aslinya ke bentuk lain yang sesuai.

Atau, untuk kurva lengkung dapat pula didekati dengan **Regresi Polynomial**.

Persamaan polynomial orde- $r$  mempunyai bentuk :

$$y = a_0 + a_1x + a_2x^2 + \dots + a_rx^r$$

Jumlah kuadrat dari kesalahan adalah :

$$D^2 = \sum_{i=1}^n (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_rx_i^r)^2 \quad \dots (9)$$

Jika didiferensialkan terhadap setiap koefisiennya, maka persamaan (9) akan menjadi :

$$\frac{\partial D^2}{\partial a_0} = - 2 \sum_{i=1}^n (a_0 - a_1x_i - a_2x_i^2 - \dots - a_rx_i^r)$$

$$\frac{\partial D^2}{\partial a_1} = - 2 \sum_{i=1}^n x_i (a_0 - a_1x_i - a_2x_i^2 - \dots - a_rx_i^r)$$

$$\frac{\partial D^2}{\partial a_2} = - 2 \sum_{i=1}^n x_i^2 (a_0 - a_1x_i - a_2x_i^2 - \dots - a_rx_i^r)$$

•  
•  
•

$$\frac{\partial D^2}{\partial a_r} = - 2 \sum_{i=1}^n x_i^r (a_0 - a_1x_i - a_2x_i^2 - \dots - a_rx_i^r)$$

Bisa anda lihat,  
bentuk persamaan<sup>2</sup>  
di samping  
menyerupai sistem  
persamaan,  
bukan?!...

$$\begin{bmatrix}
 n & \sum x_i & \sum x_i^2 & \cdots & \sum x_i^r \\
 \sum x_i & \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^{r+1} \\
 \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \cdots & \sum x_i^{r+2} \\
 \cdot & \cdot & \cdot & & \cdot \\
 \cdot & \cdot & \cdot & & \cdot \\
 \cdot & \cdot & \cdot & & \cdot \\
 \sum x_i^r & \sum x_i^{r+1} & \sum x_i^{r+2} & \cdots & \sum x_i^{r+r}
 \end{bmatrix}
 \begin{bmatrix}
 a_0 \\
 a_1 \\
 a_2 \\
 \cdot \\
 \cdot \\
 \cdot \\
 a_r
 \end{bmatrix}
 =
 \begin{bmatrix}
 \sum y_i \\
 \sum x_i y_i \\
 \sum x_i^2 y_i \\
 \cdot \\
 \cdot \\
 \cdot \\
 \sum x_i^r y_i
 \end{bmatrix}$$

Berarti anda bisa menyatakan himpunan persamaan turunan tersebut menjadi  $AX = B$ . Dan selanjutnya kita dapat menggunakan metode untuk mencari himpunan solusi pada sistem persamaan (eliminasi Gauss, Gauss-Jordan, dll) untuk menentukan nilai  $a_0, a_1, \dots, a_r$ .

**contoh :** carilah persamaan kurva polynomial orde-2 yang mewakili data berikut :

x	0	1	2	3	4	5
y	2,1	7,7	13,6	27,2	40,9	61,1

Persamaan polynomial orde-2 memiliki bentuk :

$$g(x) = a_0 + a_1x + a_2x^2$$

$$E_i = y_i - g(x)$$

$$E_i^2 = \sum (y_i - a_0 - a_1x - a_2x^2)^2$$

$$D^2 = \sum E_i^2$$

Diferensialisasi  $D^2$  terhadap setiap koefisiennya akan menghasilkan bentuk :

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

No	$x_i$	$y_i$	$x_i^2$	$x_i^3$	$x_i^4$	$x_i y_i$	$x_i^2 y_i$
1	0	2,1	0	0	0	0	0
2	1	7,7	1	1	1	7,7	7,7
3	2	13,6	4	8	16	27,2	54,4
4	3	27,2	9	27	81	81,6	244,8
5	4	40,9	16	64	256	163,6	654,4
6	5	61,1	25	125	625	305,5	1527,5
$\Sigma$	15	152,6	55	225	979	585,6	2488,8

$$\begin{pmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \end{pmatrix}$$

$$\begin{pmatrix} 6 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 152,6 \\ 585,6 \\ 2488,8 \end{pmatrix}$$

Dengan menggunakan data tabel, maka diperoleh :

$$6a_0 + 15a_1 + 55a_2 = 152,6$$

$$15a_0 + 55a_1 + 225a_2 = 585,6$$

$$55a_0 + 225a_1 + 979a_2 = 2488,8$$

$$a_0 = 2,4786$$

$$a_1 = 2,3593$$

$$a_2 = 1,8607$$



Jadi persamaan kurva yang dicari adalah :

$$y = 2,4786 + 2,3593 x + 1,8607 x^2$$

## 5.1. Curve Fitting Untuk Formula Linier

Jika fungsi yang hendak difit dapat ditulis sebagai

$$y = f(x, c_1, c_2, c_3, \dots, c_n) = c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) + \dots + c_n f_n(x) \text{ atau } y = \sum_{i=1}^n c_i f_i(x),$$

karena hubungan antara  $c_1, c_2, c_3, \dots, c_n$  adalah linier maka dikatakan formula linier. Di sini nilai  $c_i$  ditentukan melalui curve fitting. Dengan diketahuinya nilai  $c_i$  maka kesalahan

pada setiap titik data dapat ditentukan dengan  $\sum_{i=1}^n c_i f_i(x) - y = \delta$ . Jika jumlah titik yang

diketahui adalah  $m$  maka dapat ditulis  $\sum_{i=1}^n c_i f_i(x_j) - y_j = \delta_j$  untuk  $j=1, 2, 3, \dots, m$ .

Persamaan ini menunjukkan bahwa ada  $n$  bilangan yang tidak diketahui dari  $m$  persamaan yang ada. Persamaan ini merupakan persamaan linier yang dapat ditulis:

$$\sum_{i=1}^n c_i f_i(x_1) - y_1 = \delta_1 \quad \sum_{i=1}^n c_i f_i(x_2) - y_2 = \delta_2 \quad \dots \quad \sum_{i=1}^n c_i f_i(x_m) - y_m = \delta_m$$

Agar kriteria kuadrat terkecil dipenuhi maka  $\sum_{j=1}^m \delta_j^2 = \text{minimum}$ . Untuk mengimplementasikan kriteria ini dapat dilakukan dengan melakukan diferensial bagian terhadap persamaan  $\sum_{j=1}^m \delta_j^2 = \text{minimum}$  dan hasilnya dibuat sama dengan nol seperti berikut:

$\frac{\partial}{\partial c_k} \sum_{j=1}^m \delta_j^2 = 0 \Rightarrow \sum_{j=1}^m \delta_j \frac{\partial \delta_j}{\partial c_k} = 0$ . Karena  $\delta_j = \sum_{i=1}^n c_i f_i(x_j) - y_j$  dan  $\frac{\partial \delta_j}{\partial c_k} = f_k(x_j)$  untuk  $k=1,2,3,\dots,n$ , dengan mensubstitusikannya ke persamaan sebelumnya maka diperoleh  $\sum_{j=1}^m \left\{ \sum_{i=1}^n c_i f_i(x_j) - y_j \right\} f_k(x_j) = 0 \Rightarrow \sum_{i=1}^n c_i \sum_{j=1}^m f_i(x_j) f_k(x_j) = \sum_{j=1}^m y_j f_k(x_j)$  untuk  $k=1,2,3,\dots,n$ .

**Contoh 5.1.** Tentukan nilai  $c_1$  dan  $c_2$  dalam persamaan  $y = c_1 e^x + c_2 e^{2x}$  yang memenuhi kriteria kuadrat terkecil terhadap titik-titik  $(0,1)$ ,  $(1,-2)$  dan  $(2,-40)$ .

**Penyelesaian:** Dari soal diperoleh bahwa  $n=2$  dan  $m=3$ ,  $f_1(x) = e^x$  dan  $f_2(x) = e^{2x}$ .

Persamaan linier LSF adalah  $\sum_{i=1}^2 c_i \sum_{j=1}^3 f_i(x_j) f_k(x_j) = \sum_{j=1}^3 y_j f_k(x_j)$  untuk  $k=1,2$ . Dari sini maka dapat dibuat tabel seperti berikut:

$J$	$x_j$	$y_j$	$f_1(x_j) = e^{x_j}$	$f_2(x_j) = e^{2x_j}$	
1	0	1	1	1	
2	1	-2	$e$	$e^2$	
3	2	-40	$e^2$	$e^4$	

Untuk  $k = 1$ .

$$\begin{aligned}c_1 \{f_1(x_1)f_1(x_1) + f_1(x_2)f_1(x_2) + f_1(x_3)f_1(x_3)\} + \\c_2 \{f_2(x_1)f_1(x_1) + f_2(x_2)f_1(x_2) + f_2(x_3)f_1(x_3)\} = y_1 f_1(x_1) + y_2 f_1(x_2) + y_3 f_1(x_3) \\c_1 \{1 * 1 + e * e + e^2 * e^2\} + c_2 \{1 * 1 + e^2 * e + e^4 * e^2\} = 1 * 1 + (-2) * e + (-40) * e^2 \quad (*)\end{aligned}$$

Untuk  $k = 2$ .

$$\begin{aligned}c_1 \{f_1(x_1)f_2(x_1) + f_1(x_2)f_2(x_2) + f_1(x_3)f_2(x_3)\} + \\c_2 \{f_2(x_1)f_2(x_1) + f_2(x_2)f_2(x_2) + f_2(x_3)f_2(x_3)\} = y_1 f_2(x_1) + y_2 f_2(x_2) + y_3 f_2(x_3) \\c_1 \{1 * 1 + e * e^2 + e^2 * e^4\} + c_2 \{1 * 1 + e^2 * e^2 + e^4 * e^4\} = 1 * 1 + (-2) * e^2 + (-40) * e^4 \quad (**)\end{aligned}$$

Dari (\*) dan (\*\*) diperoleh  $63c_1 + 424c_2 = -299$  dan  $424c_1 + 3036c_2 = -2198$ . Dengan menyelesaikan persamaan linier di atas maka diperoleh  $c_1 \approx 2$  dan  $c_2 \approx -1$ . Jadi persamaan sebelumnya adalah  $y = 2e^x - e^{2x}$ .

Dengan menggunakan metode seperti di atas kelihatan bahwa resiko terjadi kesalahan sangat besar. Untuk memperkecil resiko ini biasanya digunakan metode matriks yaitu dengan mendefinisikan:

$$a_{ki} = \sum_{j=1}^m f_i(x_j) f_k(x_j) \text{ dan } b_k = \sum_{j=1}^3 y_j f_k(x_j) \text{ maka diperoleh } \sum_{i=1}^n c_i a_{ki} = b_k$$

dengan  $b_k$  dan  $a_{ki}$  adalah elemen-elemen matriks. Dengan mengandaikan matrik

$$F_{mn} = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \dots & f_n(x_1) \\ f_1(x_2) & f_2(x_2) & \dots & f_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_m) & f_2(x_m) & \dots & f_n(x_m) \end{bmatrix} \text{ maka } \bar{F}_{nm} = \begin{bmatrix} f_1(x_1) & f_1(x_2) & \dots & f_1(x_m) \\ f_2(x_1) & f_2(x_2) & \dots & f_2(x_m) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ f_n(x_1) & f_n(x_2) & \dots & f_n(x_m) \end{bmatrix}$$

dengan matrik  $\bar{F}_{nm}$  adalah transpose matrik  $F_{mn}$ . Seterusnya karena suatu matrik, yang jumlah baris dan kolomnya tidak sama, jika dikalikan dengan transposenya maka jumlah kolom dan barisnya menjadi sama, maka dengan mengandaikan:

$$A_{nn} = \bar{F}_{nm} F_{mn} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad B_{n1} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad Y_{m1} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \quad C_{n1} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

hubungan antara matrik dapat ditulis sebagai  $A_{nn}C_{n1} = B_{n1}$  atau  $\bar{F}_{nm}F_{mn}C_{n1} = \bar{F}_{nm}Y_{m1}$ .

Yang artinya jika suatu persamaan linier yang elemen-elemennya dinyatakan dengan matrik, dengan jumlah yang tidak diketahui lebih sedikit dibandingkan jumlah persamaan maka dengan mengalikan ruas kiri dan ruas kanan dengan matrik transpose koefisiennya akan diperoleh jumlah persamaan menjadi sama dengan jumlah yang tidak diketahui.

**Contoh 5.2.** Selesaikan Contoh 5.1 sebelumnya dengan metoda matrik.

**Penyelesaian:** Dengan menggunakan metode matrik dan tabel pada Contoh 5.1 maka diperoleh:

$$F_{32} = \begin{bmatrix} f_1(x_1) & f_2(x_1) \\ f_1(x_2) & f_2(x_2) \\ f_1(x_3) & f_2(x_3) \end{bmatrix} \quad F_{32} = \begin{bmatrix} 1 & 1 \\ e & e^2 \\ e^2 & e^4 \end{bmatrix} \quad \bar{F}_{23} = \begin{bmatrix} 1 & e & e^2 \\ 1 & e^2 & e^4 \end{bmatrix}$$

$$C_{21} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad Y_{31} = \begin{bmatrix} 1 \\ -2 \\ -40 \end{bmatrix}$$

Dengan menggunakan persamaan  $\bar{F}_{nm} F_{mn} C_{nl} = \bar{F}_{nm} Y_{ml}$  maka diperoleh:

$$\bar{F}_{23} F_{32} C_{21} = \bar{F}_{23} Y_{31} \quad \begin{bmatrix} 1 & e & e^2 \\ 1 & e^2 & e^4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ e & e^2 \\ e^2 & e^4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 & e & e^2 \\ 1 & e^2 & e^4 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ -40 \end{bmatrix}$$

$$A_{22} C_{21} = B_{21}$$

Elemen-elemen matrik dapat diperoleh dari perkalian matrik.

$$a_{11} = \{1 * 1 + e * e + e^2 * e^2\} = 63 \quad a_{12} = \{1 * 1 + e * e^2 + e^2 * e^4\} = 424$$

$$a_{21} = \{1 * 1 + e^2 * e + e^4 * e^2\} = 424 \quad a_{22} = \{1 * 1 + e^2 * e^2 + e^4 * e^4\} = 3036$$

$$b_{11} = \{1 * 1 + (-2) * e + (-40) * e^2\} = -299 \quad b_{21} = \{1 * 1 + (-2) * e^2 + (-40) * e^4\} = -2198$$

Dengan memasukkan nilainya masing-masing maka diperoleh nilai masing-masing elemen seperti berikut:

$$\begin{bmatrix} 63 & 424 \\ 424 & 3036 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -299 \\ -2198 \end{bmatrix}$$

Dari sini diperoleh  $63c_1 + 424c_2 = -299$  dan  $424c_1 + 3036c_2 = -2198$ . Dengan menyelesaikan persamaan linier di atas maka diperoleh  $c_1 \approx 2$  dan  $c_2 \approx -1$ , yang hasilnya sama dengan hasil yang diperoleh sebelumnya yaitu  $y = 2e^x - e^{2x}$ .

## 5.2. Kuadrat Terkecil Berbobot (Weighted Least Square Fit)

Sering pula ditemui bahwa grafik yang diperoleh diketahui seharusnya lebih dekat ke titik-titik data tertentu. Jika hal ini ditemui maka harus diberikan bobot titik-titik data tertentu tersebut lebih besar dibandingkan bobot titik-titik data yang lainnya. Untuk

mengimplementasikan hal ini maka perhatikan persamaan sebelumnya  $\sum_{j=1}^m \delta_j^2 = \text{minimum}$

$\Rightarrow \delta_1^2 + \delta_2^2 + \dots + \delta_m^2 = \text{minimum}$ . Di sini bobot masing-masing titik dibuat sama yaitu

1. Jika bobotnya tidak sama maka dapat ditulis  $w_1\delta_1^2 + w_2\delta_2^2 + \dots + w_m\delta_m^2 = \text{minimum}$ , dengan,  $w_1, w_2, \dots, w_m$  adalah bobot masing-masing titik. Perhatikan bahwa sebelumnya kita menganggap  $w_1=w_2=\dots=w_m=1$ , atau bobot masing-masing titik adalah sama yaitu 1. Selanjutnya metode penyelesaiannya sama dengan metode sebelumnya.

**Contoh 5.3.** Untuk Contoh 5.1 sebelumnya, jika titik ke-2 dibuat bobotnya 2 kali bobot titik-titik data yang lain maka tentukan nilai  $c_1$  dan  $c_2$ .

**Penyelesaian:** Jika  $w_1=w_3=1$  maka  $w_2=2$ . Dengan demikian maka dapat ditulis  $1\delta_1^2 + 2\delta_2^2 + 1\delta_3^2 = \text{minimum}$  atau  $1\delta_1^2 + 1\delta_2^2 + 1\delta_2^2 + 1\delta_3^2 = \text{minimum}$ . Hal ini menunjukkan bahwa seolah-olah dimiliki ada 4 titik data yaitu  $(0,1)$ ,  $(1,-2)$ ,  $(1,-2)$  dan  $(2,-40)$ . Dengan menyelesaiakannya dengan metode kwadrat terkecil maka diperoleh:

$$F_{42} = \begin{bmatrix} f_1(x_1) & f_2(x_1) \\ f_1(x_2) & f_2(x_2) \\ f_1(x_3) & f_2(x_3) \\ f_1(x_4) & f_2(x_4) \end{bmatrix} \quad F_{42} = \begin{bmatrix} 1 & 1 \\ e & e^2 \\ e & e^2 \\ e^2 & e^4 \end{bmatrix} \quad \bar{F}_{24} = \begin{bmatrix} 1 & e & e & e^2 \\ 1 & e^2 & e^2 & e^4 \end{bmatrix}$$

$$C_{21} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad Y_{41} = \begin{bmatrix} 1 \\ -2 \\ -2 \\ -40 \end{bmatrix}$$

$$\bar{F}_{24} F_{42} C_{21} = \bar{F}_{24} Y_{41}$$

$$\begin{bmatrix} 1 & e & e & e^2 \\ 1 & e^2 & e^2 & e^4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ e & e^2 \\ e & e^2 \\ e^2 & e^4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 & e & e & e^2 \\ 1 & e^2 & e^2 & e^4 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ -2 \\ -40 \end{bmatrix}$$

$$A_{22}C_{21} = B_{21}$$

$$a_{11} = \{1 * 1 + e * e + e * e + e^2 * e^2\} = 70 \quad a_{12} = \{1 * 1 + e * e^2 + e * e^2 + e^2 * e^4\} = 445$$

$$a_{21} = \{1 * 1 + e^2 * e + e^2 * e + e^4 * e^2\} = 445 \quad a_{22} = \{1 * 1 + e^2 * e^2 + e^2 * e^2 + e^4 * e^4\} = 3091$$

$$b_{11} = \{1 * 1 + e * (-2) + e * (-2) + e^2 * (-40)\} = -305$$

$$b_{21} = \{1 * 1 + e^2 * (-2) + e^2 * (-2) + e^4 * (-40)\} = -2212$$

$$\begin{bmatrix} 70 & 445 \\ 445 & 3091 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -305 \\ -2212 \end{bmatrix}$$

Dari sini diperoleh  $70 c_1 + 445 c_2 = -305$  dan  $445 c_1 + 3091 c_2 = -2212$ . Dengan menyelesaikan persamaan linier di atas maka diperoleh  $c_1 \approx 2$  dan  $c_2 \approx -1$ , yang hasilnya sama dengan hasil yang diperoleh sebelumnya yaitu  $y = 1,746e^x - 0,96e^{2x}$ .

# Latihan

1. Tentukan: (a) rerata; (b) deviasi standar; dan (c) varian; dari data-data berikut

0,95	1,42	1,54	1,55	1,63
1,32	1,15	1,47	1,95	1,25
1,46	1,47	1,92	1,35	1,05
1,85	1,74	1,65	1,78	1,71
2,39	1,82	2,06	2,14	2,27

2. Gunakan regresi kuadrat terkecil untuk menaksি fungsi garis lurus dari data berikut :

x	1	3	5	7	10	12	13	16	18	20
y	3	2	6	5	8	7	10	9	12	10

3. Gunakan regresi kuadrat terkecil untuk menaksি fungsi garis lurus dari data berikut :

x	4	6	8	10	14	16	20	22	24	28	28	34	36	38
y	30	18	22	28	14	22	16	8	20	8	14	14	0	8

4. Gunakan regresi kuadrat terkecil untuk menaksি fungsi kurva dari data berikut :

x	1	2	2,5	4	6	8	8,5
y	0,4	0,7	0,8	1,0	1,2	1,3	1,4

5. Gunakan regresi kuadrat terkecil untuk menaksi fungsi kurva dari data berikut :

x	2,5	3,5	5	6	7,5	10	12,5	15	17,5	20
y	5	3,4	2	1,6	1,2	0,8	0,6	0,4	0,3	0,3

6. Gunakan regresi polynomial untuk menaksi fungsi kurva dari data berikut :

x	0,05	0,4	0,8	1,2	1,6	2,0	2,4
y	550	750	1.000	1.400	2.000	2.700	3.750

7. Gunakan regresi polynomial untuk menaksi fungsi kurva dari data berikut :

x	0	2	4	6	9	11	13	15	17	19	23	25	28
y	1,2	0,6	0,4	-0,2	0	-0,6	-0,4	-0,2	-0,4	0,2	0,4	1,2	1,8