



# REGRESI

TEKNIK INFORMATIKA

JURUSAN TEKNOLOGI INFORMASI

POLITEKNIK NEGERI MALANG

# PENDAHULUAN

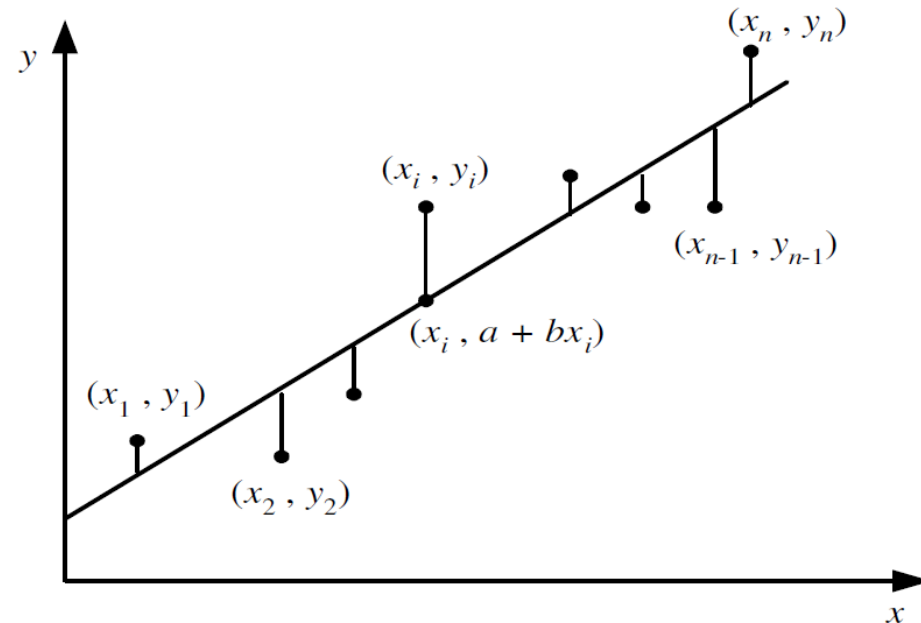
- Regresi adalah teknik pencocokan kurva untuk data yang berketelitian rendah.
- Contoh data yang berketelitian rendah data hasil pengamatan, percobaan di laboratorium, atau data statistik. Data seperti itu kita sebut data hasil pengukuran.
- Untuk data hasil pengukuran, pencocokan kurva berarti membuat fungsi menghampiri (approximate) titik-titik data.
- Kurva fungsi hampiran tidak perlu melalui semua titik data tetapi dekat dengannya tanpa perlu menggunakan polinom berderajat tinggi.

# PENDAHULUAN

- Prinsip penting yang harus diketahui dalam mencocokkan kurva untuk data hasil pengukuran adalah:
  1. Fungsi mengandung sesedikit mungkin parameter bebas
  2. Deviasi fungsi dengan titik data dibuat minimum.
- Kedua prinsip di atas mendasari metode **regresi kuadrat terkecil**.
- Teknik regresi yang dibahas di sini hanya regresi linier, yaitu pencocokan kurva untuk data yang memiliki hubungan linier antara peubah bebas dan peubah terikatnya.

# REGRESI LANJAR/ LINIER

- Misalkan  $(x_i, y_i)$  adalah data hasil pengukuran. Kita akan menghampiri titik-titik tersebut dengan sebuah garis lurus. Garis lurus tersebut dibuat sedemikian sehingga galatnya sekecil mungkin dengan titik-titik data.



# REGRESI LANJAR

Karena data mengandung galat, maka nilai data sebenarnya,  $g(x_i)$  dapat ditulis sebagai

$$g(x_i) = y_i + e_i \quad i = 1, 2, \dots, n$$

yang dalam hal ini,  $e_i$  adalah galat setiap data.

Diinginkan fungsi lanjar

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

Yang mencocokkan data sedemikian sehingga deviasinya,

$$r_i = y_i - f(x_i) = y_i - (a + bx_i)$$

minimum.

## REGRESI LANJAR

Total kuadrat deviasi persamaan  $r_i = y_i - f(x_i) = y_i - (a + bx_i)$  adalah

$$R = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - a - bx_i)^2$$

Agar  $R$  minimum, maka harus

$$\frac{\partial R}{\partial a} = -2 \sum (y_i - a - bx_i) = 0$$

$$\frac{\partial R}{\partial b} = -2 \sum x_i (y_i - a - bx_i) = 0$$

Untuk selanjutnya, notasi ditulis “ $\sum$ ” saja.

# REGRESI LANJAR

## ***Penyelesaian:***

Masing-masing ruas kedua persamaaan dibagi dengan -2:

$$\sum (y_i - a - bx_i) = 0$$

$$\sum y_i - \sum a - \sum bx_i = 0$$

$$\sum x_i(y_i - a - bx_i) = 0$$

$$\sum x_i y_i - \sum ax_i - \sum bx_i^2 = 0$$

Selanjutnya,

$$\sum a + \sum bx_i = \sum y_i$$

$$\sum ax_i + \sum bx_i^2 = \sum x_i y_i$$

atau

$$na + b \sum x_i = \sum y_i$$

$$a \sum x_i + b \sum x_i^2 = \sum x_i y_i$$

## REGRESI LANJAR

Kedua persamaan terakhir ini dinamakan **persamaan normal**, dan dapat dapat ditulis dalam bentuk persamaan matriks:

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

Solusi (nilai a dan b) bisa dicari dengan metode eliminasi Gauss atau langsung dengan rumus :

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

$$a = \bar{y} - b\bar{x} \rightarrow \bar{y} = \frac{\sum y_i}{n}; \bar{x} = \frac{\sum x_i}{n}$$



## CONTOH

Tentukan persamaan garis lurus yang mencocokkan data pada tabel di bawah ini.  
Kemudian, perkirakan nilai  $y$  untuk  $x = 1.0$ .

**Penyelesaian :**

$i$	$x_i$	$y_i$	$x_i^2$	$x_i y_i$
1	0.1	0.61	0.01	0.061
2	0.4	0.92	0.16	0.368
3	0.5	0.99	0.25	0.495
4	0.7	1.52	0.49	1.064
5	0.7	1.47	0.49	1.029
6	0.9	2.03	0.81	1.827
	$\Sigma x_i = 3.3$	$\Sigma y_i = 7.54$	$\Sigma x_i^2 = 2.21$	$\Sigma x_i y_i = 4.844$

Diperoleh sistem persamaan linier:

$$\begin{bmatrix} 6 & 3.3 \\ 3.3 & 2.21 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 7.54 \\ 4.844 \end{bmatrix}$$

Solusi SPL di atas adalah:  $a = 0.2862$  dan  $b = 1.7645$

## CONTOH

Persamaan garis regresinya adalah:  $f(x) = 0.2862 + 1.7645x$

Perbandingan antara nilai  $y_i$  dan  $f(x_i)$ :

$i$	$x_i$	$y_i$	$f(x_i) = a + bx_i$	deviasi	(deviasi) <sup>2</sup>
1	0.1	0.61	0.46261	0.147389	0.02172
2	0.4	0.92	0.99198	-0.07198	0.00518
3	0.5	0.99	1.16843	-0.17844	0.03184
4	0.7	1.52	1.52135	-0.00135	0.00000
5	0.7	1.47	1.52135	-0.05135	0.00264
6	0.9	2.03	1.87426	0.15574	0.02425
					$\Sigma = 0.08563$

Taksiran nilai  $y$  untuk  $x = 1.0$  adalah  $y = f(1.0) = 0.2862 + 1.7645(1.0) = 2.0507$

# LATIHAN SOAL

I. Berikut ini data mengenai pengalaman kerja dan penjualan

$X$  = pengalaman kerja (tahun)

$Y$  = omzet penjualan (ribuan)

X	2	3	2	5	6	1	4	1
Y	5	8	8	7	11	3	10	4

Tentukan nilai  $a$  dan  $b$  serta buatkan persamaan regresinya!

## LATIHAN SOAL

2. Hubungan antara variable X dan variable Y

X	1	2	3	4	5	6
Y	6	4	3	5	4	2

- Buatkan persamaan regresinya
- Hitung deviasi rata-ratanya
- Tentukan nilai dugaan Y, jika  $X = 8$