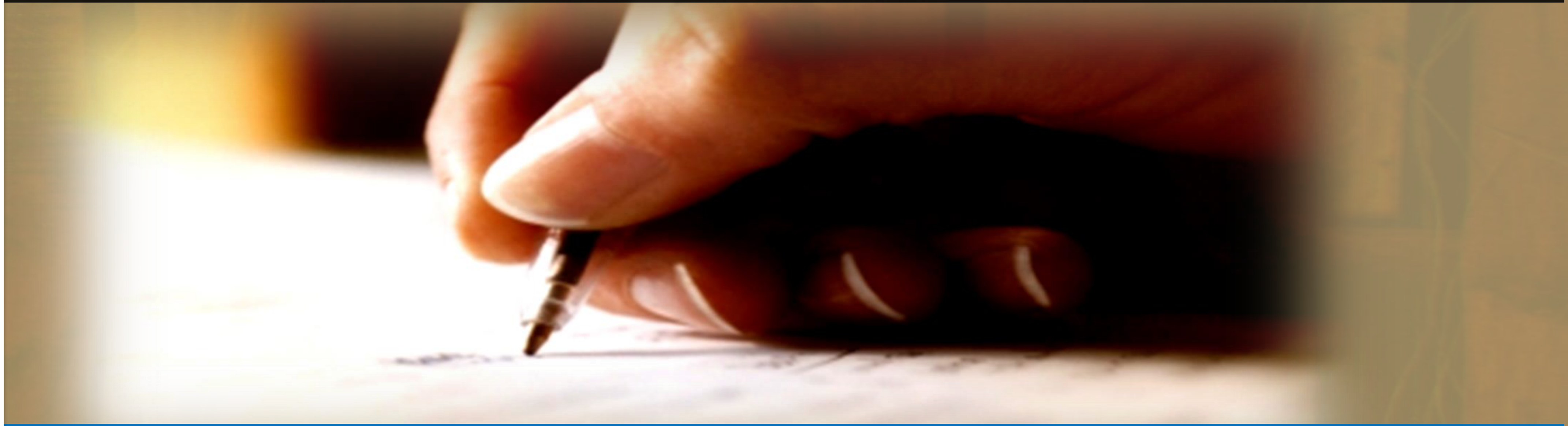
Institut Teknologi Sepuluh Nopember
Surabaya

JDEPARTEMEN TEKNIK FISIKA - FTIRS



REGRESI LINIER

Oleh: Syamsul Arifin

REGRESI LINIER

Capaian Pembelajaran:

Mampu menentukan persamaan regresi linear atas variable dependen sebagai fungsi dari variable independent nya

Kajian:

1. Regresi Linier
2. Penentuan parameter regresi

1. Konsep Regresi Linier

Regresi

- ✓ Teknik yang digunakan untuk pemodelan dan analisa data numerik
- ✓ Hubungan antara dua atau lebih variable, sehingga informasi dari variable tersebut akan menyebabkan diketahui nilai variable yang lain.
- ✓ Regresi dapat digunakan untuk:
  1. Prediksi
  2. Estimasi
  3. Test hipotesa
  4. Pemodelan hubungan sebab akibat

$$Y = X_1 + X_2 + X_3$$

- ✓ Variabel terikat
- ✓ Variabel Outcome
- ✓ Variabel Response

- ✓ Variable bebas
- ✓ Variabel predictor
- ✓ Variabel Explanatory

Mengapa Regresi Linier?

- Contoh, misalkan di inginkan model dari variable bebas Y dalam bentuk tiga variable predictor X1, X2, X3

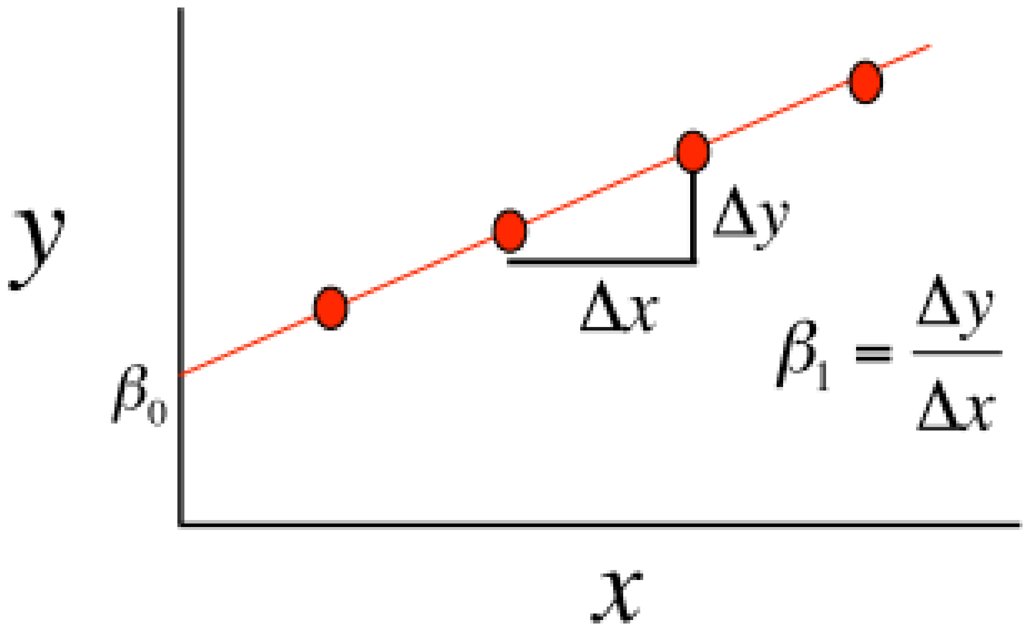
$$Y = f(X1, X2, X3)$$

- Bila tidak cukup data untuk melakukan estimasi secara langsung terhadap Y
- Diasumsikan bahwa bentuk persamaan nya:

$$Y = X1 + X2 + X3$$

Regresi Linier merupakan Model Probabilitas

$$y = \beta_0 + \beta_1 x$$



Regresi Sederhana

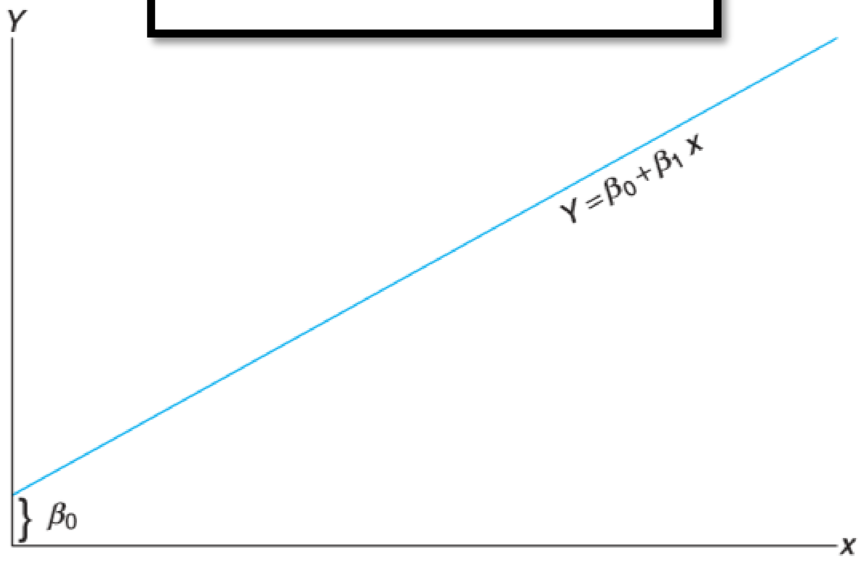
Hubungan antara variabel terikat dan variabel bebas

$$X \rightarrow Y$$

Suhu	PH
Tekanan	Rapat massa
Laju aliran	…
Massa	Panas laten
…	…

$$Y = \beta_0 + \beta_1 x + \epsilon.$$

Model regresi sederhana

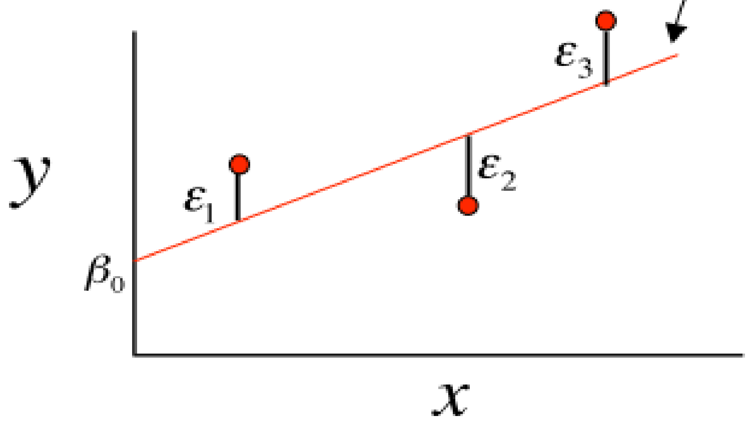


Model Probabilitas Linier

Model dari Variabel terikatnya – sebagai fungsi variable bebas:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

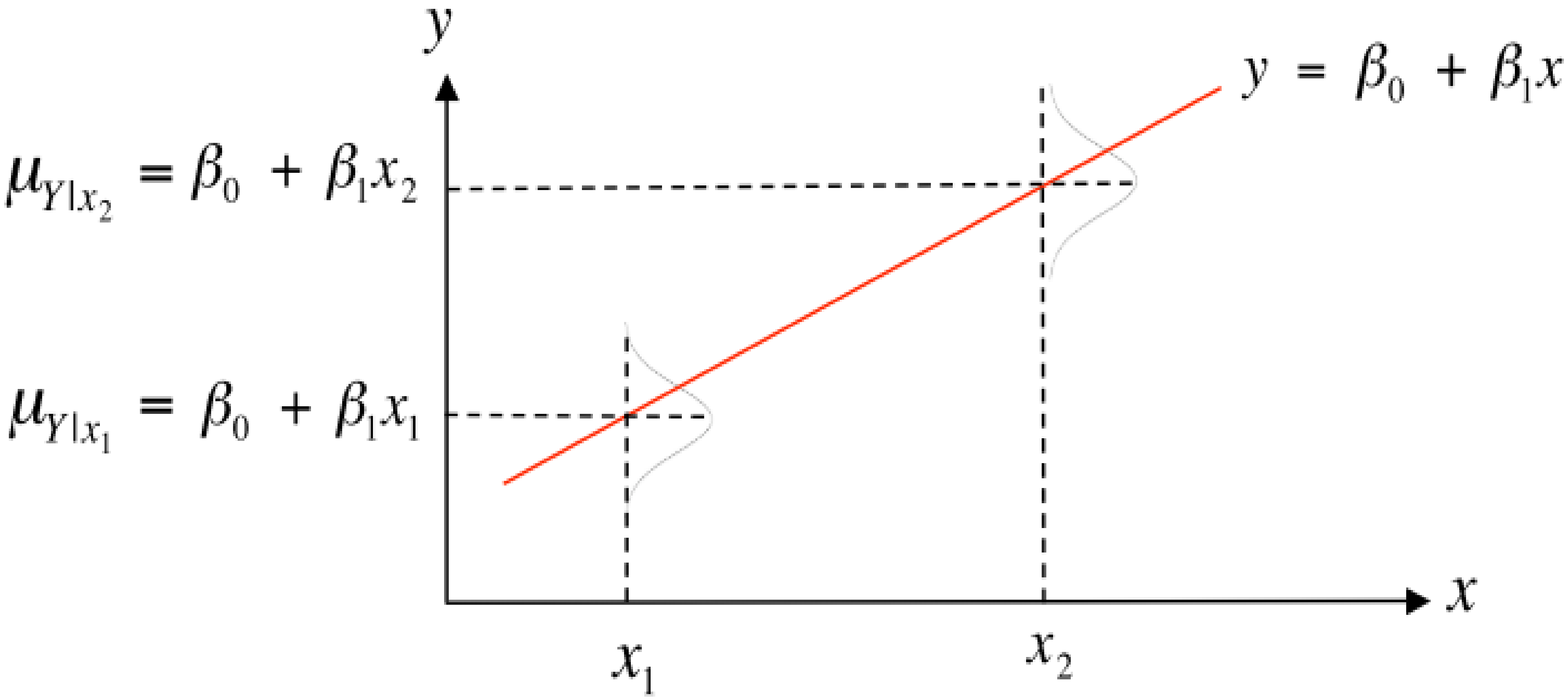
ε diasumsikan $N(0, \sigma^2)$



Pers. Regresi yang benar

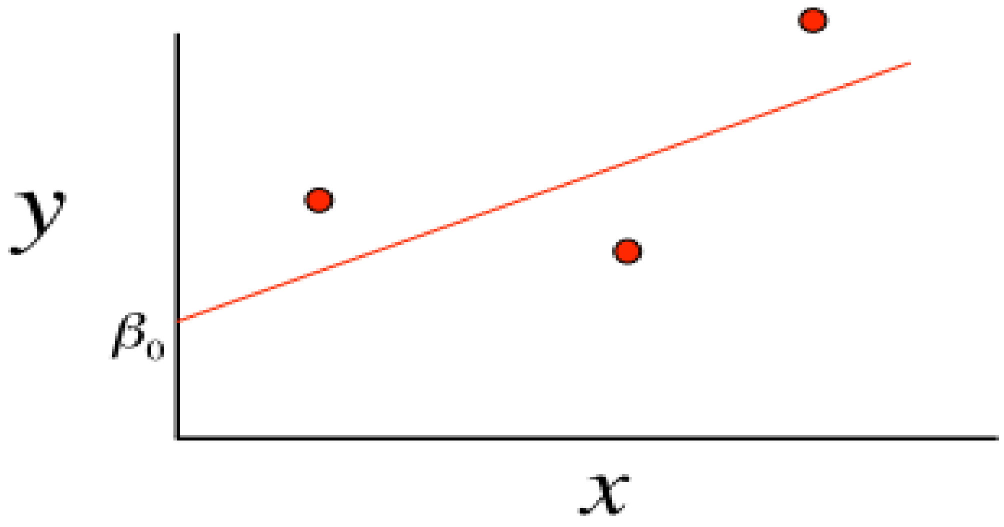
$$y = \beta_0 + \beta_1 x$$

Interprestasi Grafik



Estimasi parameter

$$f(\beta_0, \beta_1) = \sum_{i=1}^{n} [y_i - (\beta_0 + \beta_1 x_i)]^2$$



- $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

Prediksi atau fitting – dalam bentuk persamaan linier
Untuk nilai x1, x2

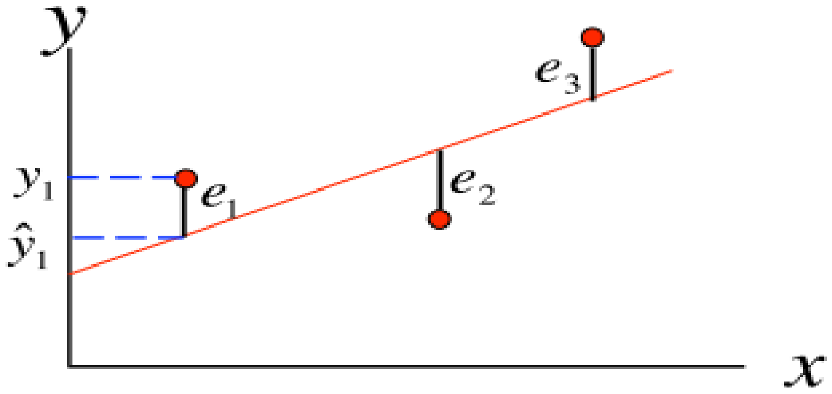
$$\hat{y}_1 = \hat{\beta}_0 - \hat{\beta}_1 x_1$$

$$\hat{y}_2 = \hat{\beta}_0 - \hat{\beta}_1 x_2$$

Residual atau Deviasi – selisih antara hasil pengamatan (nilai actual) dengan nilai prediksi

$$e_1 = y_1 - \hat{y}_1$$

$$e_2 = y_2 - \hat{y}_2$$



Sum – square error (SSE)

$$SSE = \sum_{i=1}^{n} (e_i)^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Estimasi terhadap σ^2:

$$\hat{\sigma}^2 = \frac{SSE}{n-2}$$

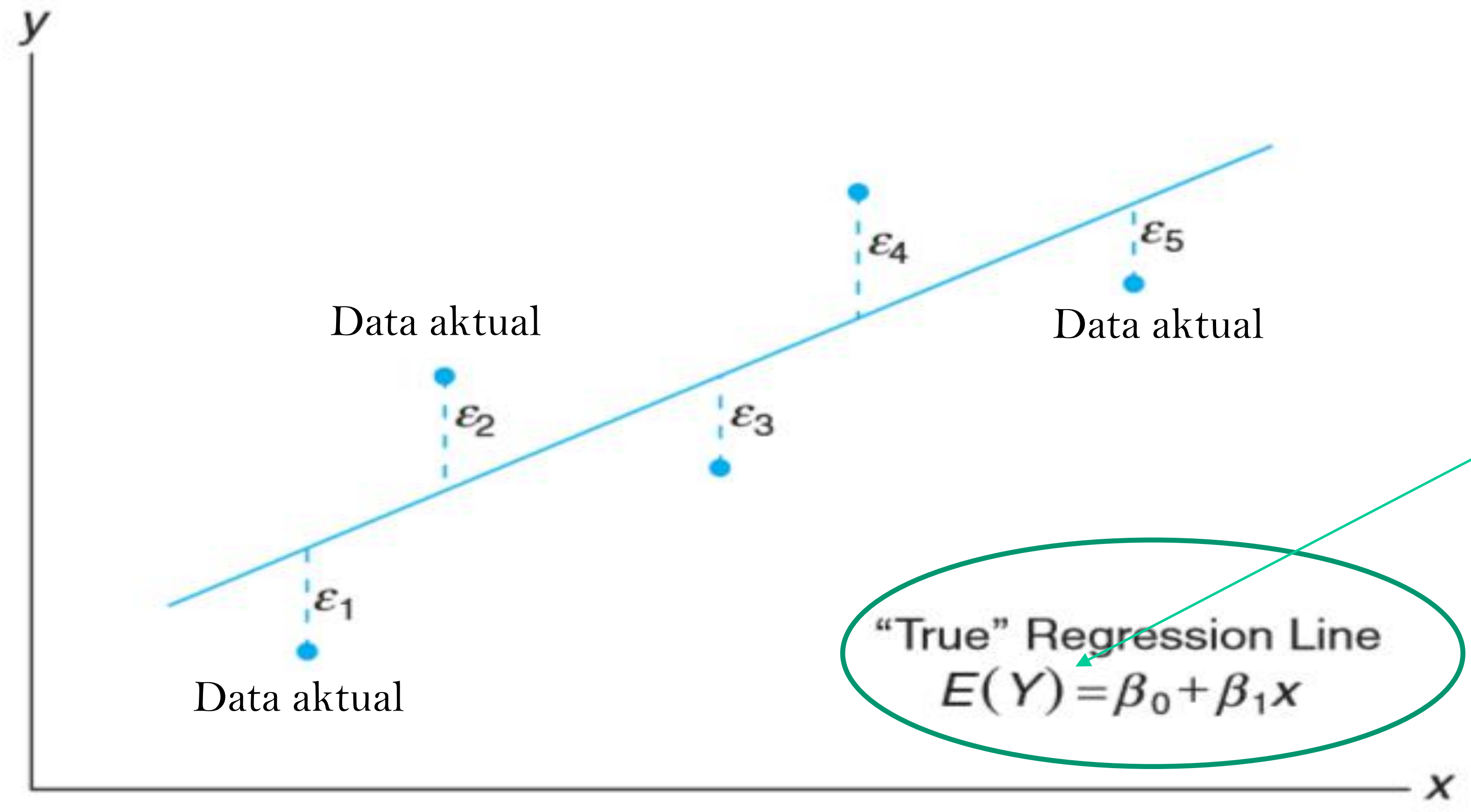
Penentuan koefisien

$$r^2 = 1 - \frac{SSE}{SST}$$

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

Regresi yang benar

Dalam eksperimen,
tidak selalu delta x
adalah tetap



Ekspektasi dari y

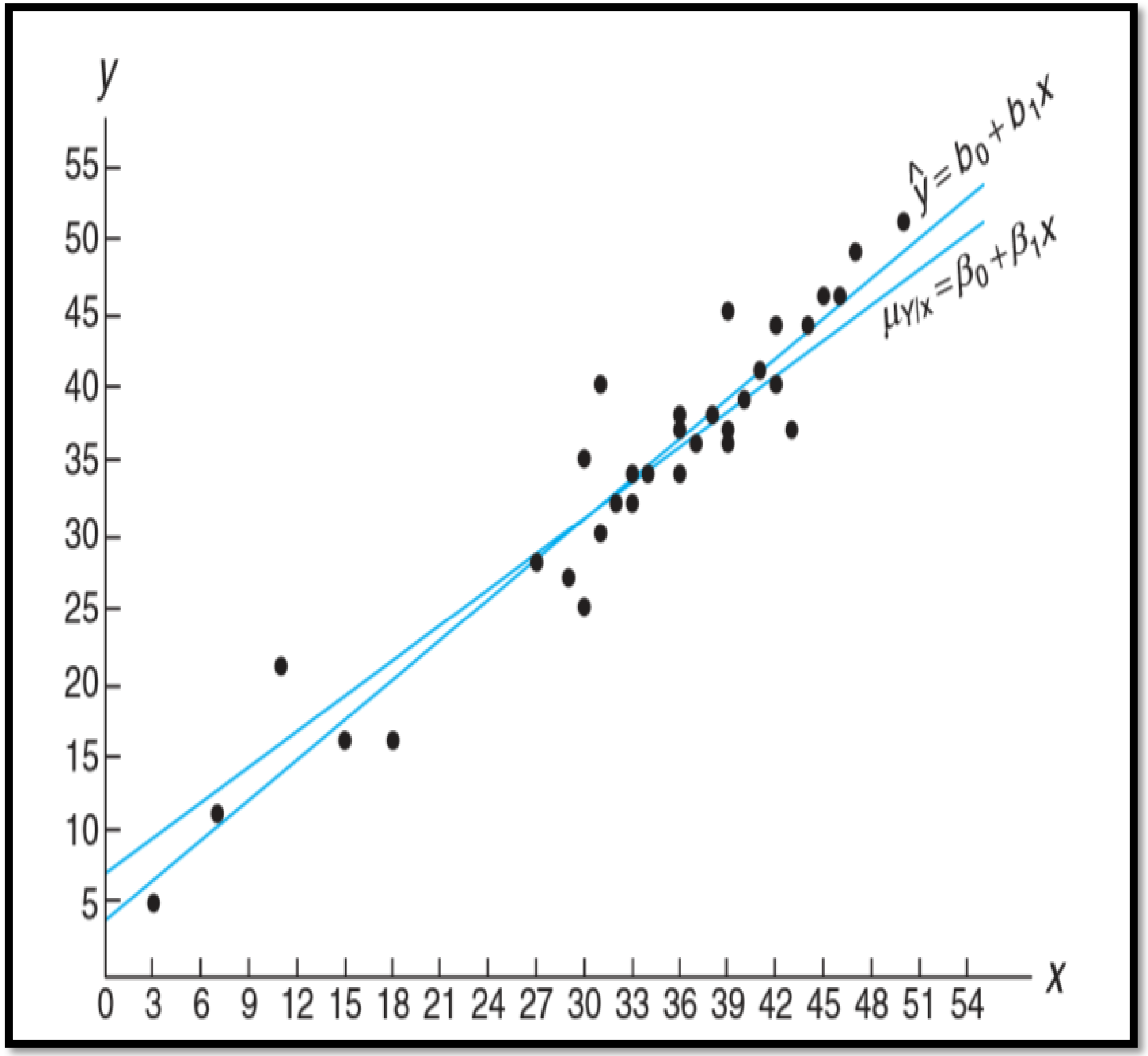
estimasi

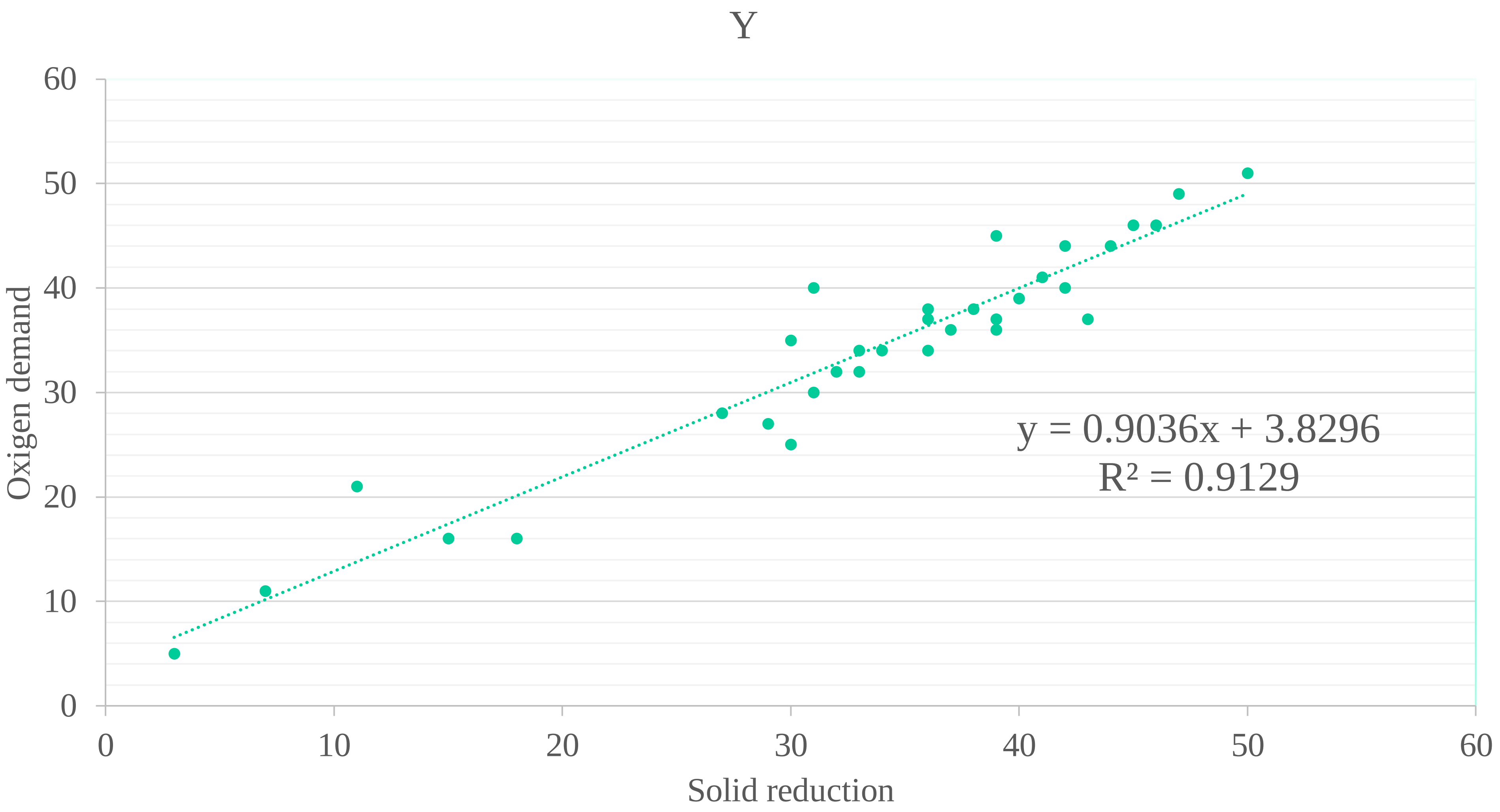
$$\hat{y} = b_0 + b_1 x,$$

Contoh

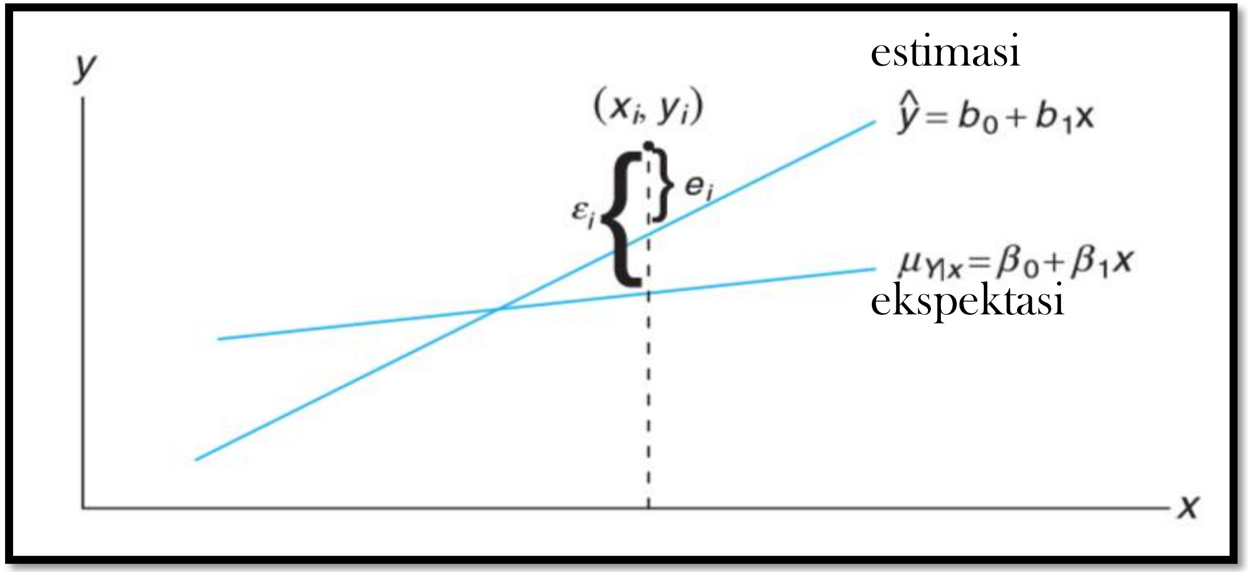
Table 11.1: Measures of Reduction in Solids and Oxygen Demand

Solids Reduction, x (%)	Oxygen Demand Reduction, y (%)	Solids Reduction, x (%)	Oxygen Demand Reduction, y (%)
3	5	36	34
7	11	37	36
11	21	38	38
15	16	39	37
18	16	39	36
27	28	39	45
29	27	40	39
30	25	41	41
30	35	42	40
31	30	42	44
31	40	43	37
32	32	44	44
33	34	45	46
33	32	46	46
34	34	47	49
36	37	50	51
36	38		




Y
60
50
40
30
20
10
0
Oxigen demand
0
10
20
30
40
50
60
Solid reduction
y = 0.9036x + 3.8296
R² = 0.9129

Metode Least Square



$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2.$$

Differentiating SSE with respect to b_0 and b_1, we have

$$\frac{\partial(SSE)}{\partial b_0} = -2\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i), \quad \frac{\partial(SSE)}{\partial b_1} = -2\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)x_i.$$

Setting the partial derivatives equal to zero and rearranging the terms, we obtain the equations (called the **normal equations**)

$$nb_0 + b_1\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \quad b_0\sum_{i=1}^{n} x_i + b_1\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i,$$

which may be solved simultaneously to yield computing formulas for b_0 and b_1.

Estimating the Regression Coefficients Given the sample $\{(x_i, y_i);\ i = 1, 2, \ldots, n\}$, the least squares estimates b_0 and b_1 of the regression coefficients β_0 and β_1 are computed from the formulas

$$b_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \text{ and}$$

$$b_0 = \frac{\sum_{i=1}^{n} y_i - b_1\sum_{i=1}^{n} x_i}{n} = \bar{y} - b_1\bar{x}.$$

The calculations of b_0 and b_1, using the data of Table 11.1, are illustrated by the following example.

Materi

Catat Penjelasan saat Kuliah

Materi

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$= \sum_{i=1}^{n} (y_i - b_1 - b_2 x_i)^2$$

$$SSR = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2$$

$$= \sum_{i=1}^{n} \left(y_i^2 - 2y_i(\beta_0 + \beta_1 x_i) + \beta_0^2 + 2\beta_0\beta_1 x_i + \beta_1^2 x_i^2 \right)$$

Materi

$$\frac{\partial \sum_{i=1}^{n} e_i^2}{\partial b_1} = -2\sum_{i=1}^{n}(y_i - b_1 - b_2 x_i) = 0 \qquad (1)$$

$$\frac{\partial \sum_{i=1}^{n} e_i^2}{\partial b_2} = -2\sum_{i=1}^{n}x_i(y_i - b_1 - b_2 x_i) = 0 \qquad (2)$$

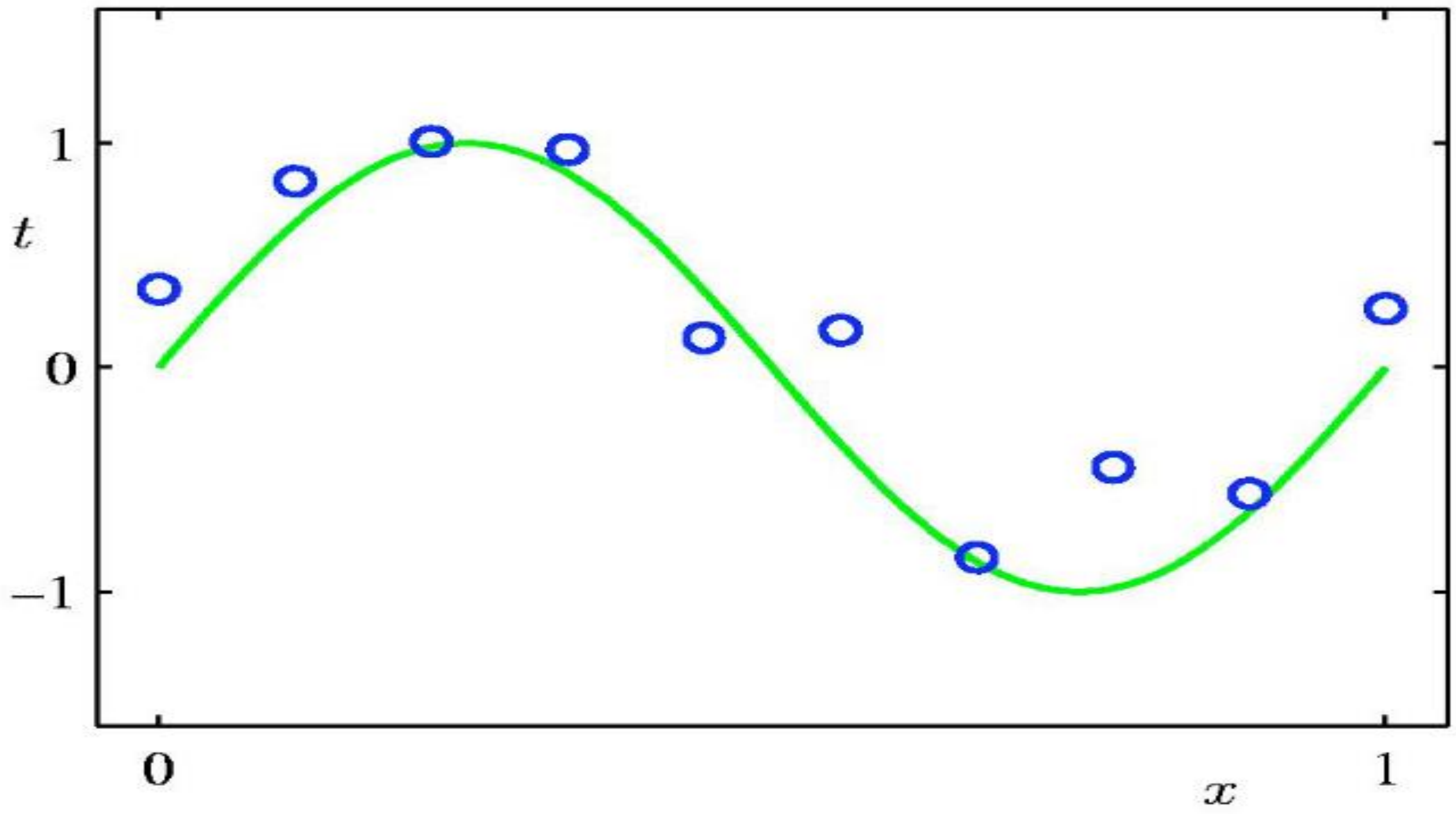
$$
\begin{aligned}
\frac{\partial SSR}{\partial \beta_0} &= \sum_{i=1}^{n} (-2y_i + 2\beta_0 + 2\beta_1 x_i) \\
0 &= \sum_{i=1}^{n} \left(-y_i + \hat{\beta}_0 + \hat{\beta}_1 x_i\right) \\
0 &= -n\bar{y} + n\hat{\beta}_0 + \hat{\beta}_1 n\bar{x} \\
\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x} \\
\frac{\partial SSR}{\partial \beta_1} &= \sum_{i=1}^{n} \left(-2x_i y_i + 2\beta_0 x_i + 2\beta_1 x_i^2\right) \\
0 &= -\sum_{i=1}^{n} x_i y_i + \hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 \\
0 &= -\sum_{i=1}^{n} x_i y_i + (\bar{y} - \hat{\beta}_1 \bar{x}) \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 \\
\hat{\beta}_1 &= \frac{\sum_{i=1}^{n} x_i (y_i - \bar{y})}{\sum_{i=1}^{n} x_i (x_i - \bar{x})}
\end{aligned}
$$

Materi

$$
\begin{aligned}
b_2 &= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
b_1 &= \bar{y} - b_2\bar{x}
\end{aligned}
$$

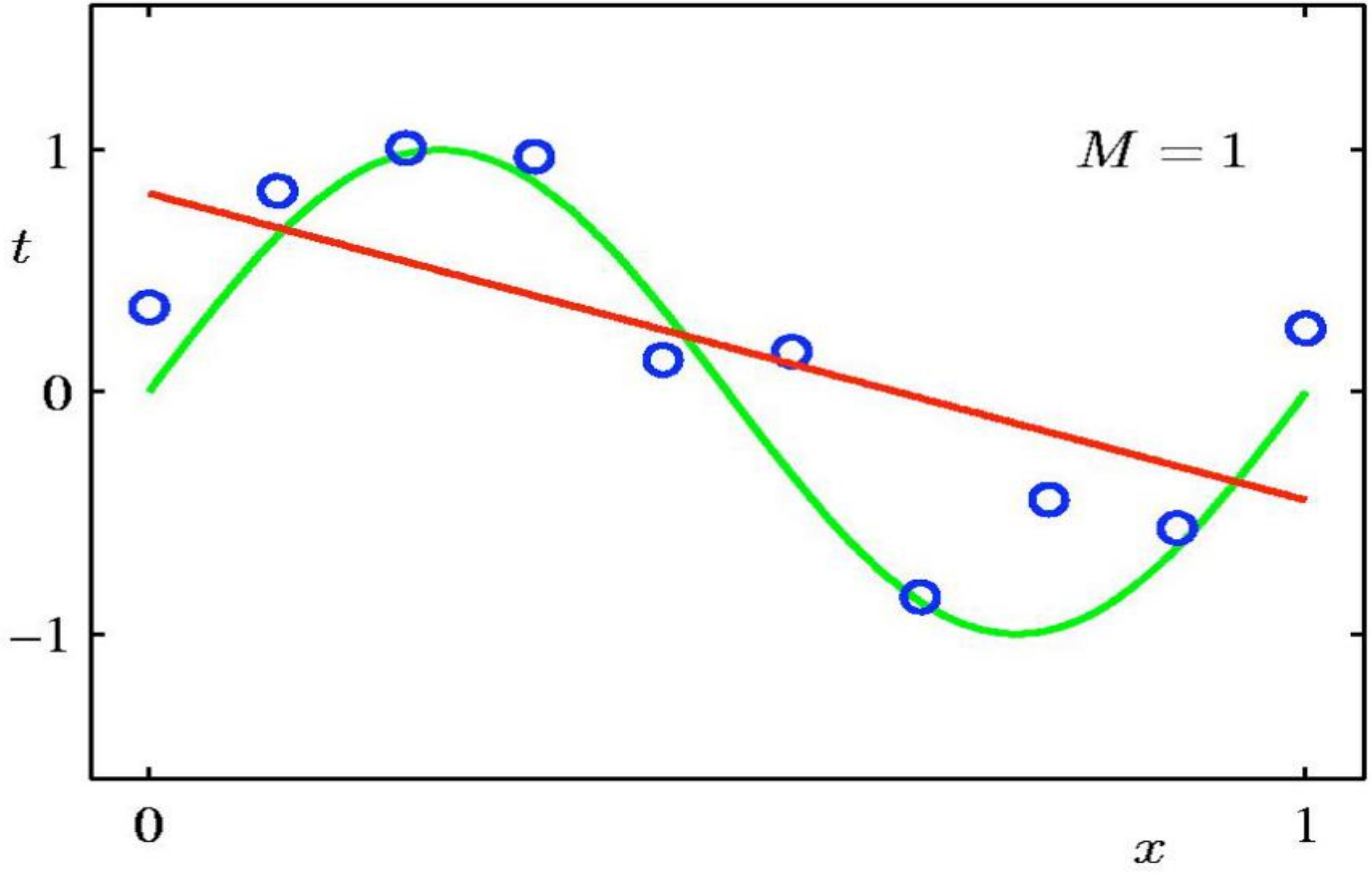
$$
\begin{aligned}
b_2 &= \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} \\
b_1 &= \bar{y} - b_2\bar{x}
\end{aligned}
$$

Materi

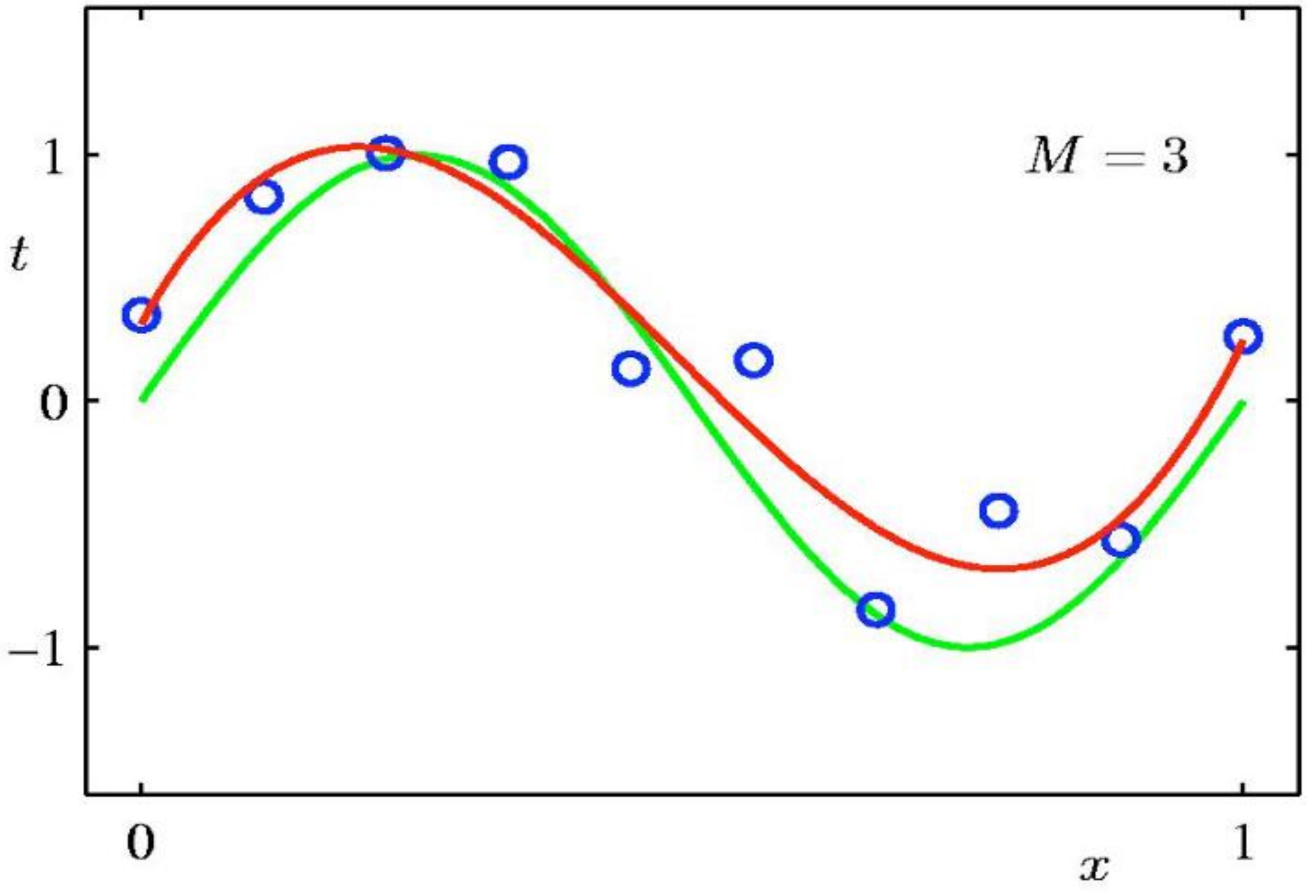


$$y(x, \mathbf{w}) = w_0 + w_1 x + w_2 x^2 + \ldots + w_M x^M = \sum_{j=0}^{M} w_j x^j$$

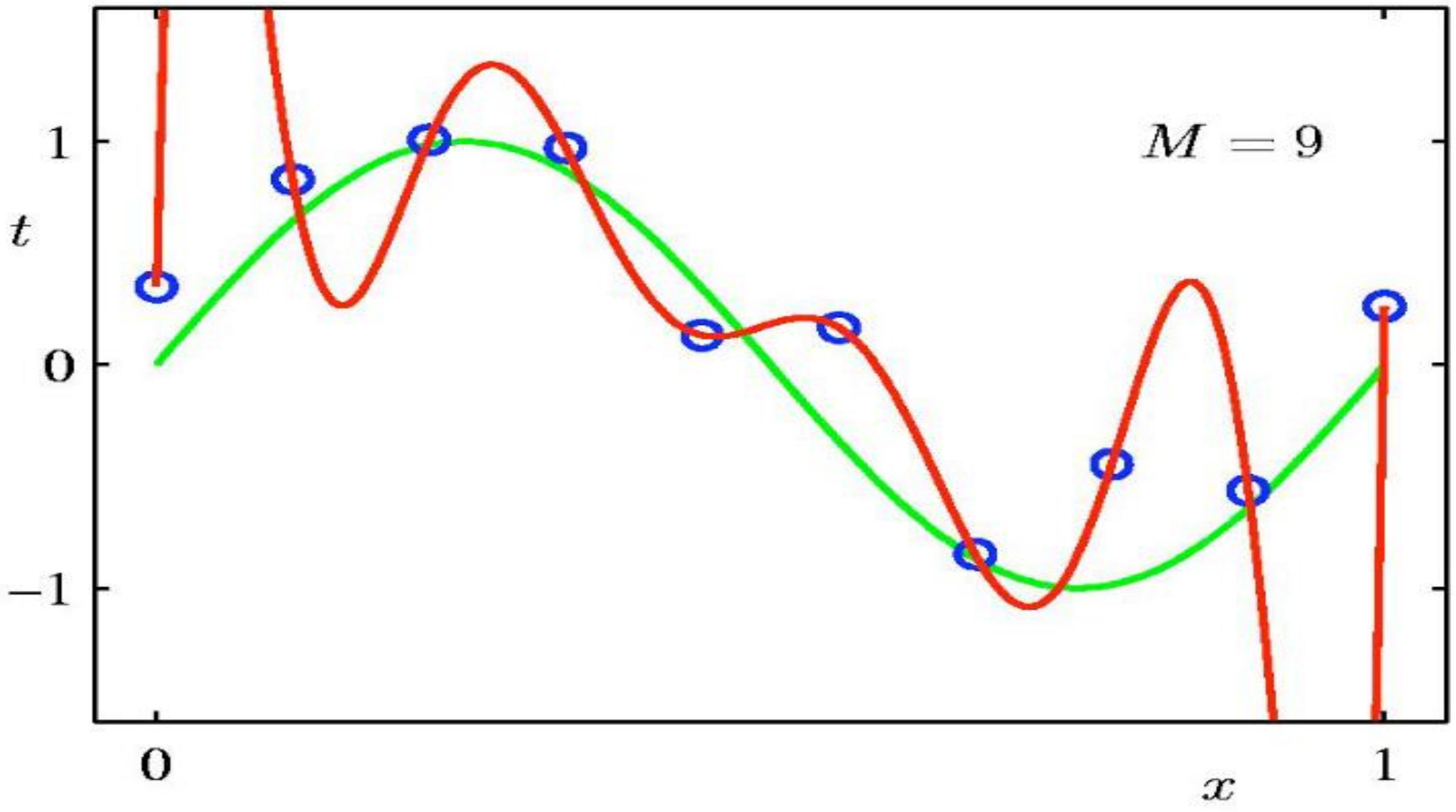
Contoh – pendekatan regresi linier



Materi



Materi



Materi

Bantuan software Excel

11.6 In a certain type of metal test specimen, the normal stress on a specimen is known to be functionally related to the shear resistance. The following is a set of coded experimental data on the two variables:

Normal Stress, x	Shear Resistance, y
26.8	26.5
25.4	27.3
28.9	24.2
23.6	27.1
27.7	23.6
23.9	25.9
24.7	26.3
28.1	22.5
26.9	21.7
27.4	21.4
22.6	25.8
25.6	24.9

Tentukan bentuk regresi linier $y = f(x)$

Terimakasih