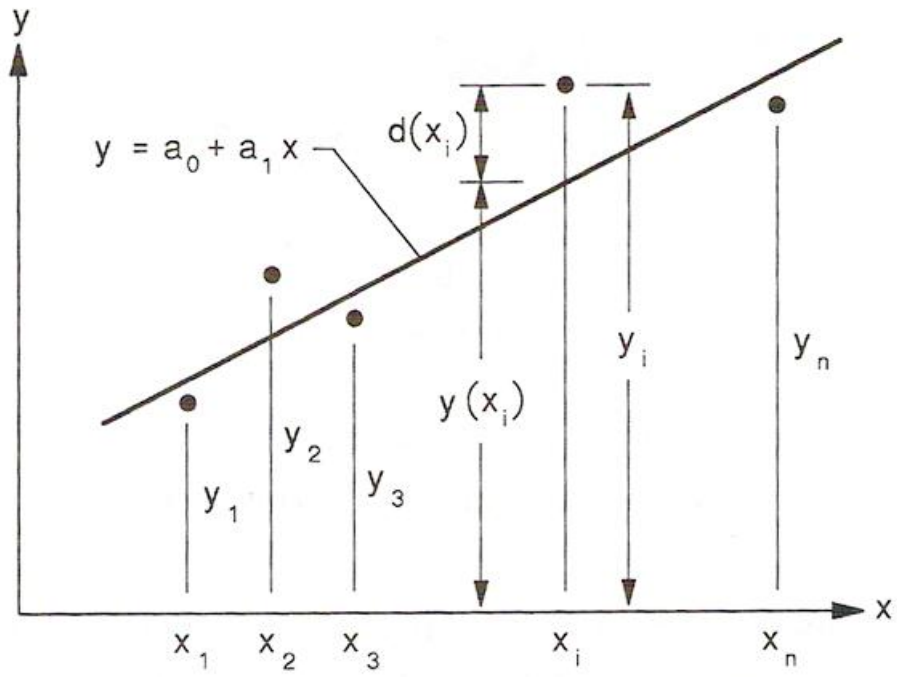
<u>การถดถอยแบบเชิงเส้น</u> <u>(linear regression)</u>

เป็นระเบียบวิธีที่ใช้ในการสร้างฟังก์ชันเส้นตรงสำหรับชุดของข้อมูลที่กำหนดให้ พิจารณากราฟจากภาพ



ชุดข้อมูลจากภาพประกอบด้วย x_i, y_i โดยที่ i = 1, 2 , ...n ซึ่งเมื่อเราต้องการสร้างสมการเส้นตรง จะได้ว่า

$$y = a_0 + a_1 x \qquad (1)$$

โดยที่ a_0 และ a_1 เป็นค่าคงที่ ที่จะต้องคำนวณหาจากเงื่อนไข สมการเส้นตรงที่สร้างขึ้นมานี้ จะทำให้เกิดค่าความผิดพลาดโดยเฉลี่ยที่น้อยที่สุดจากทุกข้อมูลที่กำหนด ด้วยการใช้วิธีกำลังสองน้อยสุด (least square)

พิจารณาจากภาพ ณ ตำแหน่ง x_i ของข้อมูลใดๆ ค่าของ $y(x_i)$ ที่ได้จากสมการ จะมีค่าที่แตกต่างไปจากข้อมูล y_i เท่ากับ $d(x_i)$

ดังนั้น ค่าความผิดพลาด E ทั้งหมดที่เกิดจากข้อมูล n ข้อมูล เมื่อเขียนอยู่ในรูปกำลังสองของค่าแตกต่าง $d(x_i)$ จะได้ว่า

$$E = \sum d(x_i)^2 \qquad (2)$$

เพราะว่า $d(x_i) = y_i - y(x_i)$ แทนลงในสมการ (2) จะได้ว่า

$$E = \sum [y_i - y(x_i)]^2 \tag{3}$$

แทนค่า $y(x_i)$ จากสมการ (1) ในสมการ (3) จะได้ว่า

$$E = \sum [y_i - (a_0 + a_1 x_i)]^2 \tag{4}$$

จากสมการ (4) เราสามารถหาค่า a_0 และ a_1 ด้วยวิธีกำลังสองน้อยที่สุด โดยการหาค่าต่ำสุดของค่าความผิดพลาด นั่นคือหาค่า $\frac{\partial E}{\partial a_0} = 0$ และ $\frac{\partial E}{\partial a_1} = 0$ ในสมการ (4) จะได้ว่า

$$na_0 + \left(\sum x_i\right)a_1 = \sum y_i \tag{5}$$

$$\left(\sum x_i\right)a_0 + \left(\sum x_i^{\,2}\right)a_1 = \sum x_i\, y_i \tag{6}$$

ดังนั้น

$$a_0 = \frac{\left(\sum y_i\right)\left(\sum x_i^{\,2}\right) - \left(\sum x_i\, y_i\right)\left(\sum x_i\right)}{n\left(\sum x_i^{\,2}\right) - \left(\sum x_i\right)^2} \tag{7}$$

$$a_1 = \frac{n\left(\sum x_i\, y_i\right) - \left(\sum x_i\right)\left(\sum y_i\right)}{n\left(\sum x_i^{\,2}\right) - \left(\sum x_i\right)^2} \tag{8}$$

เมื่อแทนค่า a_0 จากสมการ (7) และ a_1 จากสมการ (8) ลงในสมการ (1) ก็จะได้สมการถดถอยแบบเชิงเส้นที่ต้องการ

ตัวอย่างที่ 1 จงคำนวณหาสมการถดถอยเชิงเส้น สำหรับข้อมูลความเร็วลมตามระดับความสูงของอาคารดังตาราง

ความสูงของอาคาร, x (m)	**ความเร็วลม, y (m/sec)**
10	2.2
15	4.6
20	4.2
25	7.0
30	6.6
35	9.2

วิธีทำ

	x_i	y_i	$x_i\ y_i$	x_i^2
	10	2.2		
	15	4.6		
	20	4.2		
	25	7.0		
	30	6.6		
	35	9.2		
Σ				

a_0 =

=

a_1 =

=

เพราะฉะนั้น

y = #

ตัวอย่างที่ 2 จงคำนวณหาสมการถดถอยเชิงเส้น สำหรับข้อมูลต่อไปนี้

x	0	1	2	4	6	7	9	10	12
y	1	6	10	23	32	39	49	56	65

วิธีทำ

	x_i	y_i	$x_i\ y_i$	x_i^2
Σ				

a_0 =

=

a_1 =

=

เพราะฉะนั้น

y = #

แบบฝึกหัด

1) จงคำนวณหาสมการถดถอยเชิงเส้น สำหรับข้อมูลต่อไปนี้

x	1000	1500	2000	2500	3000	3500
y	1457	1282	932	788	465	264

2) จงคำนวณหาสมการถดถอยเชิงเส้น จากข้อมูลการทดสอบระบบเบรกของรถยนต์ ตั้งแต่เริ่มเบรกจนกระทั่งรถหยุดสนิท ดังตาราง และใช้สมการนี้คำนวณหาระยะหยุดรถเมื่อวิ่งด้วยความเร็ว 65 กม./ชม.

ความเร็วรถ (กม./ชม.)	10	15	20	30	40	50	60	70	80
ระยะหยุด (ม.)	5	9	15	18	22	30	35	38	43

3) จงคำนวณหาสมการถดถอยเชิงเส้น จากข้อมูลการทดสอบดินดังตาราง และใช้สมการนี้คำนวณค่าหน่วยแรงเฉือนที่ความลึก 4 m

ความลึก (m)	1.9	3.1	5.1	6.9	9.3
หน่วยแรงเฉือน (ksc)	0.3	0.6	0.9	1.1	1.2