

رگرسیون خطی

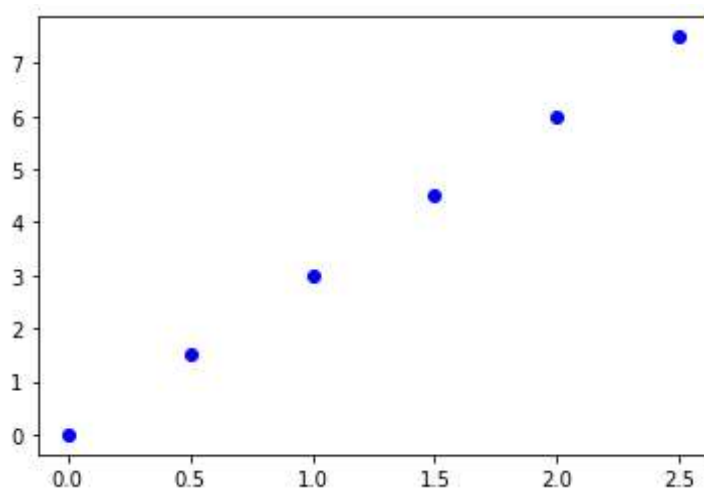
• طراح سوال اوستا صباي مقدم

فرض کنید تعدادی داده‌ی زیر را داریم:

$$x = \{0, 0.5, 1, 1.5, 2, 2.5\}$$

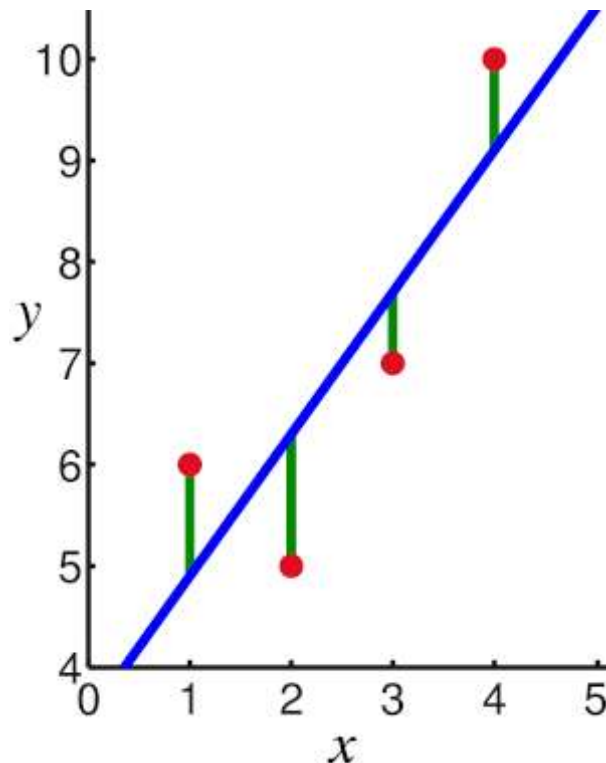
$$y = \{0, 1.5, 3, 4.5, 6, 7.5\}$$

نمودار پراکندگی داده‌های زیر به شکل زیر است:



واضح هست که با تابعی خطی به فرم $y = ax + b$ می‌توان رفتار این تابع را توصیف کرد اما سوال اینجاست که مقادیر a و b چیست؟

به زبان ساده رگرسیون خطی پیدا کردن تابعی خطی است که رفتار داده‌ها را توصیف کند یا همان پیدا کردن a و b . از نظر ما تابعی خوب است که مجموع فاصله‌اش با هر کدام از نقاط یا به عبارتی $\sum (y_i - f(x_i))^2$ مینیمم باشد. برای درک بهتر موضوع به شکل زیر نگاه کنید:



خطی که مجموع خطوط سبز رنگ برایش کمترین مقدار باشد بهترین گزینه است. برای یافتن مینیمم از مشتق گیری استفاده می‌کنیم داریم:

$$err = \sum (y_i - f(x_i))^2 = \sum (y_i - (ax_i + b))^2$$

$$\frac{\partial err}{\partial a} = -2 \sum_{i=1}^n x_i (y_i - ax_i - b) = 0$$

$$\frac{\partial err}{\partial b} = -2 \sum_{i=1}^n (y_i - ax_i - b) = 0$$

با کمی ساده سازی به دو معادله‌ی زیر می‌رسیم:

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

$$a \sum x_i + b * n = \sum y_i$$

که در اینجا تنها a و b مجهول هستند پس دو معادله و دو مجهول داریم. در این سوال از شما می‌خواهیم برنامه‌ای نوشته که مجموعه‌ای از داده‌ها را از کاربر دریافت کند و سپس ضرایب تابع خطی که رفتار آن داده‌ها را توصیف کند به ما نشان دهد. ابتدا توابع زیر را به صورت بازگشتی تعریف کنید:

این تابع یک لیست را دریافت کرده و مجموع اعضای آن را برمی‌گرداند `sum`.

این تابع یک لیست را دریافت کرده و مجموع مربع اعضای آن را برمی‌گرداند `sumsq`.

دو لیست را گرفته و عناصر آنها را نظیر به نظیر ضرب کرده و لیست حاصل را برگرداند `multiply`.

سپس مقادیر لیست‌های x و y را در ورودی از کاربر گرفته و با استفاده از توابع فوق و دو معادله و دو مجهول بالا مقادیر a و b را حساب کرده و به کاربر نمایش دهید.

توجه: a و b را با استفاده از تابع `round` تا سه رقم اعشار گرد نمایید.

ورودی

در خط اول برنامه عدد n که بیانگر تعداد اعضای x و y است وارد می‌شود. در n خط بعدی اعضای x و در n خط بعدتر اعضای y وارد می‌شوند.

خروجی

خروجی برنامه شامل دو خط است که در خط اول مقدار b و در خط دوم مقدار a نمایش داده می‌شود. (مجدداً تاکید می‌گردد که فراموش نکنید که اعداد را تا سه رقم اعشار گرد کنید)

مثال

ورودی نمونه ۱

6

0

0.5

1.0

Copy

1.5
2
2.5
0
1.5
3
4.5
6
7.5

خروجى نمونه ۱

0.0
3.0

Copy