

最小自乗法とは重回帰分析の偏回帰係数を求める方法の 1 つです。実際にはソフトウェアが計算してくれるので、手計算をする必要はありません。ですがどのようにして計算されているのかくらいは知っていて損はないでしょう。

なお、最小自乗法によって偏回帰係数を求めるには逆行列という行列についての知識が必要となりますが、これも実際にはコンピュータで計算できるのであまり細かいことは気にしなくてもよいです。

売り上げ (y)	広告費 (x1)	人材教育費 (x2)
126	56	21
67	44	18
41	43	14
29	37	11
26	30	10

1. 売り上げ (目的変数) を次のようにおきます。

$$y = \begin{bmatrix} 126 \\ 67 \\ 41 \\ 29 \\ 26 \end{bmatrix}$$

2. 広告費と人材教育費を次のようにおきます (1 番左の列の要素には 1 を配置する)。

$$X = \begin{bmatrix} 1 & 56 & 21 \\ 1 & 44 & 18 \\ 1 & 43 & 14 \\ 1 & 37 & 11 \\ 1 & 30 & 10 \end{bmatrix}$$

3. $X^T X$ を計算します。

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 56 & 44 & 43 & 37 & 30 \\ 21 & 18 & 14 & 11 & 10 \end{bmatrix} \begin{bmatrix} 1 & 56 & 21 \\ 1 & 44 & 18 \\ 1 & 43 & 14 \\ 1 & 37 & 11 \\ 1 & 30 & 10 \end{bmatrix}$$

行列の積は R を使えば簡単に計算できます。

```
> X
      [,1] [,2] [,3]
[1,]    1   56   21
[2,]    1   44   18
[3,]    1   43   14
[4,]    1   37   11
[5,]    1   30   10
> t(X)%*%X
      [,1] [,2] [,3]
[1,]    5  210   74
[2,]  210 9190 3277
[3,]   74 3277 1182
```

4. $X^T X$ の逆行列 $(X^T X)^{-1}$ を求めます。これも R を使えば簡単に計算できます。

```
> X2
      [,1] [,2] [,3]
[1,]    5  210   74
[2,]  210 9190 3277
[3,]   74 3277 1182
> X3
      [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    1    0
[3,]    0    0    1
> solve(X2,X3)
      [,1]      [,2]      [,3]
[1,] 6.9677075 -0.32191280  0.45625879
[2,] -0.3219128  0.02441632 -0.04753868
[3,]  0.4562588 -0.04753868  0.10407876
```

5. 続いて $(X^T X)^{-1} X^T$ を計算します。

```
> X3
      [,1]      [,2]      [,3]
[1,] 6.9677075 -0.32191280  0.45625879
[2,] -0.3219128  0.02441632 -0.04753868
[3,]  0.4562588 -0.04753868  0.10407876
```

```

> X3%%t(X)
           [,1]      [,2]      [,3]      [,4]      [,5]
[1,] -1.47797468  1.0162025 -0.48691983  0.07578059  1.87291139
[2,]  0.04708861 -0.1032911  0.06244726  0.05856540 -0.06481013
[3,] -0.02025316  0.2379747 -0.13080169 -0.15780591  0.07088608

```

5. 先ほど得られた行列 $(X^T X)^{-1} X^T$ と y の積を求めます。

```

> y
           [,1]
[1,] 126
[2,]  67
[3,]  41
[4,]  29
[5,]  26
> X4
           [,1]      [,2]      [,3]      [,4]      [,5]
[1,] -1.47797468  1.0162025 -0.48691983  0.07578059  1.87291139
[2,]  0.04708861 -0.1032911  0.06244726  0.05856540 -0.06481013
[3,] -0.02025316  0.2379747 -0.13080169 -0.15780591  0.07088608
> X4%%y
           [,1]
[1,] -87.209620
[2,]  1.586329
[3,]  5.296203

```

以上の計算結果から,

$$y = -87.21 + 1.59x_1 + 5.30x_2$$

となり,無事,係数と定数項を求めることができました。