

行列の計算ができれば多変量間の相関係数も簡単に求めることができます。

X1	X2	X3
1	6	4
2	10	4
4	15	7
2	9	3
3	7	4
6	21	8

1. まずは行列  $X$  として上の表のデータを表現してみましょう。

$$X = \begin{bmatrix} 1 & 6 & 4 \\ 2 & 10 & 4 \\ 4 & 15 & 7 \\ 2 & 9 & 3 \\ 3 & 7 & 4 \\ 6 & 21 & 8 \end{bmatrix}$$

2. 次に  $X$  を標準化します (これを行列  $Z$  とします)。

$$Z = \begin{bmatrix} -1.11 & -0.93 & -0.50 \\ -0.55 & -0.23 & -0.50 \\ 0.55 & 0.64 & 1.00 \\ -0.55 & -0.41 & -1.00 \\ 0.00 & -0.76 & -0.50 \\ 1.67 & 1.70 & 1.50 \end{bmatrix}$$

3.  $Z^T Z$  を計算します ( $Z^T$  は  $Z$  の転置行列です)。途中の計算手順は省略して、ここでは  $R$  を使います。

```
> t(z) %*% z
      [,1] [,2] [,3]
[1,] 5.000000 4.625363 4.472136
[2,] 4.625363 5.000000 4.577162
[3,] 4.472136 4.577162 5.000000
```

4.  $ZTZ = S$  として,相関行列  $R$  は

$$R = S / (n - 1)$$

で求めることができます。n はデータ数です。今回の場合はデータ数が 6 なので,  $6 - 1 = 5$  ということになります。

```
> s <- t(z) %*% z #ZTZ を s に代入。
> s / (6 - 1) #s を n-1 でわる。
      [1] [2] [3]
[1,] 1.0000000 0.9250726 0.8944272
[2,] 0.9250726 1.0000000 0.9154324
[3,] 0.8944272 0.9154324 1.0000000
```

これで相関行列の出来上がりです。試しに相関行列を求める関数 `cor()` を使ってみると, 下のようになんと結果が一致していることが分かります。

```
> dat
      [1] [2] [3]
[1,]  1   6   4
[2,]  2  10   4
[3,]  4  15   7
[4,]  2   9   3
[5,]  3   7   4
[6,]  6  21   8
> cor(dat)
      [1] [2] [3]
[1,] 1.0000000 0.9250726 0.8944272
[2,] 0.9250726 1.0000000 0.9154324
[3,] 0.8944272 0.9154324 1.0000000
```