Best answer with Matrix: Moore-Penrose Pseudo-inverse 行列で最善の解: ムーアペンローズ疑似逆行列

ryamada

2016年12月27日

1 連立方程式の解と最小二乗法による線形回帰 Least square method for linear regression

When the number of variables and the number of equations are the same, and the inverse of matrix is calculable, the system is solvable.

変数の数と等式の数が一致しているとき、うまく逆行列が取れれば

$$y = X\mathbf{a}$$

はきっちりと解けて$\mathbf{a}$が一意に求まる。

その際、

$$\mathbf{a} = X^{-1}y = (X^TX)^{-1}X^Ty$$

であった。

In the case you do linear regression, the number of samples is more than the number of variables, therefore it does not happen. Instead the formula below should be minimized.

線形回帰の場合は、サンプルの数(ベクトルyの長さ:nとする)に対して、説明変数の数Xの列数(mとする)が小さいので

$$y \sim X\mathbf{a}$$

として、

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

が最小になるような$\mathbf{a}$を推定するわけだが、そのとき

$$\mathbf{a} = (X^T X)^{-1} X^{-1}$$

で求まるのだった。

結局、行列Xの列数によらず、$\mathbf{a}$について、「これが一番」という答えが

$$\mathbf{a} = (X^T X)^{-1} X^T y$$

で求まる。

The “minimizig” $\mathbf{a}$ is called the least square solution.

ここで、「これが一番」というのは、最小二乗が0であったり、「かっちり連立方程式が解ける」ということだったりした。

2 連立方程式でも線形回帰でもない場合 The case where it is not the solvable system or linear regression

The number of rows of X is n and the number of columns is m. When n=m, it is the solvable system. when n>m, it is linear regression.

行列Xの行数n,列数mについて

- $n = m$の場合：連立方程式を解く
- $n > m$の場合：線形回帰をする

だった。

The condition n < m has not been discussed yet.

網羅されていないのは

- $n < m$の場合である

In this case, psuedo inverse matrix is used to solove.

この場合にも、うまく行く方法があり、これをムーアペンローズ疑似逆行列と読んだり、一般化逆行列と読んだり、疑似逆行列と呼んだりする。

The inverse matrix is not specifically determined when n is not m.

$n < m$の場合には、$(X^T X)^{-1}$がうまく計算されないので、別の式を用いる。

$(X'^T X')^{-1} X'^T$ is the formula to get one type of psuedo-inverse matrices, called Moore-Penrose type.

Xの行数が列数より小さい場合には、転置して考える。 転置($X' = X^T$)すると、そのムーアペンローズ疑似逆行列は、$(X'^T X')^{-1} X'^T$と計算できるので、それを再度転置すれとうまく行く((($X'^T X')^{-1} X'^T)^T$)。

結局、覚えるべきは In summary,

- $n \geq m$ のときは $gen.inv(X) = (X^T X)^{-1} X^T$
- $n < m$のときは、$X_{gen.inv} = (gen.inv(X^T))^T$

$n < m$の場合を$X^T(X^TX)^{-1}$と書くこともある。

```
n <- 3
m <- 2 # m < n
X <- matrix(rnorm(n*m),nrow=n)
solve(t(X)%*%X)%*%t(X)
```

```
##           [,1]      [,2]       [,3]
## [1,] -0.625956 -0.9879096 -0.1472827
## [2,]  1.614524  2.6134061 -1.4664492
```

```
n <- 3
m <- 3 # m = n
X <- matrix(rnorm(n*m),nrow=n)
solve(t(X)%*%X)%*%t(X)
```

```
##           [,1]      [,2]        [,3]
## [1,] -0.209077 0.8225860 -0.32208917
## [2,]  1.897980 1.6719556 -1.92967732
## [3,]  1.264410 0.2809747  0.03628408
```

```
n <- 2
m <- 3
X <- matrix(rnorm(n*m),nrow=n)
# solve(t(X)%*%X) %*% t(X) # Error
tX <- t(X)
t(solve(t(tX)%*%tX)%*%t(tX))
```

```
##            [,1]        [,2]
## [1,] -1.2493320 -0.72758821
## [2,]  0.1234655  0.06926064
## [3,]  0.7660636 -0.09626852
```

```
t(X)%*%solve(X%*%t(X))
```

```
##            [,1]        [,2]
## [1,] -1.2493320 -0.72758821
## [2,]  0.1234655  0.06926064
## [3,]  0.7660636 -0.09626852
```

You will find ginv() function in MASS package that returns Moore-Penrose psuedoinverse matrix.

Rでは、MASS パッケージにginv() 関数があり、Xの行数・列数に関係なく、ムーアペンローズ疑似逆行列を算出できる。

任意の行列を次のように分解することを特異値分解と言うが The following is singular value decomposition.

$$X = U\Sigma V^T$$

ムーアペンローズ疑似逆行列は

$$gen.inv(X) = V\Sigma^+ U^T$$

Where Σ is diagonal and whose diagonal elements are singular values and Σ^+ is diagonal and its diagonal values are the inverse of sigular values.

で与えられる。ただしΣは対角成分に特異値を持つ行列であり、$/Simga^+$ は特異値の逆数を対角成分に持つ。

The ginv() function in MASS package uses this to return the psuedo-inverse matrix.

実際、Rではこれを利用して、MASSパッケージのginv()関数(g:generalized, inv:inverse)がXの行数・列数を場合分けすることなく、ムーアペンローズ疑似逆行列を返す。

```
library(MASS)
```

```
## Warning: package 'MASS' was built under R version 3.6.2
```

```
n <- 3
m <- 2 # m < n
X <- matrix(rnorm(n*m),nrow=n)
solve(t(X)%*%X)%*%t(X)
```

```
##            [,1]      [,2]        [,3]
## [1,] -0.1116292 0.1222668 -0.33937634
## [2,] -0.2737158 0.5894858  0.07843593
```

```
ginv(X)
```

```
##            [,1]      [,2]        [,3]
## [1,] -0.1116292 0.1222668 -0.33937634
## [2,] -0.2737158 0.5894858  0.07843593
```

```
n <- 3
m <- 3
X <- matrix(rnorm(n*m),nrow=n)
solve(t(X)%*%X)%*%t(X)
```

```
##            [,1]       [,2]       [,3]
## [1,]  0.9431798 -0.2103496 -0.6185412
## [2,] -0.9681423 -0.5082898  0.0414813
## [3,] -0.3874916  0.6476825 -0.4760076
```

```
ginv(X)
```

```
##            [,1]       [,2]       [,3]
## [1,]  0.9431798 -0.2103496 -0.6185412
## [2,] -0.9681423 -0.5082898  0.0414813
## [3,] -0.3874916  0.6476825 -0.4760076
```

```
n <- 2
m <- 3
X <- matrix(rnorm(n*m),nrow=n)
tX <- t(X)
t(solve(t(tX)%*%tX)%*%t(tX))
```

```
##              [,1]        [,2]
## [1,]  0.002383139 -0.04619755
## [2,] -0.638975002  0.73034089
## [3,] -0.445238855  2.84775425
```

```
ginv(X)
```

```
##              [,1]        [,2]
## [1,]  0.002383139 -0.04619755
## [2,] -0.638975002  0.73034089
## [3,] -0.445238855  2.84775425
```

3 Exercise 1

What does the Moore-Penrose pseudoinverse matrix indicate geometrically? To answer this question, go through the steps of the exsecises below.

ムーアペンローズ疑似逆行列の解の幾何的な意味を以下の手順で確認せよ。

3.1 Exercise 1-1

$n = m$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_{11}, x_{12} \\ x_{21}, x_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

To get the values of a_1, a_2 means to know the crossing point of two straight lines in 2D space.

今、a_1, a_2の値を求めるというのは、(a_1, a_2)を2次元座標として、2次元平面にある2本の直線の交点座標を求めることである。

$$\begin{aligned} x_{11} a_1 + x_{12} a_2 &= y_1 \\ x_{21} a_1 + x_{22} a_2 &= y_2 \end{aligned}$$

今、$y_1 = 3, y_2 = 1, x_{11} = 4, x_{12} = 2, x_{21} = -1, x_{22} = 3$のとき の２直線を描き、その交点をMASS::ginv()関数を求め、その点$(\hat{a}_1, \hat{a}_2)$が交点にあることを、打点することによって示せ。

Draw two straight lines and demonstrate $(\hat{a}_1, \hat{a}_2)$ is actually the crossing point visually.

3.2 Exercise 1-2

The follwoing is the case of three lines. $n > m$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_{11}, x_{12} \\ x_{21}, x_{22} \\ x_{31}, x_{32} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

これは３直線の場合である。

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} x_{11}, x_{12} \\ x_{21}, x_{22} \\ x_{31}, x_{32} \end{pmatrix} = \begin{pmatrix} 2, 3 \\ 1, -2 \\ -3, -4 \end{pmatrix}$$ とする。

Draw three straight lines and draw a point $(\hat{a_1}, \hat{a_2})$ in it.

３直線を描図し、ムーアペンローズ疑似逆行列による解$(\hat{a_1}, \hat{a_2})$を打点せよ。

3.3 Exercise 1-3

Excerise 1-2 の例で、第１の直線は(a_1, a_2)座標について、$y_1 = x_{11}a_1 + x_{12}a_2$で表されているのに対し、解$(\hat{a_1}, \hat{a_2})$は、別の直線、$\hat{y_1} = x_{11}\hat{a_1} + x_{12}\hat{a_2}$を表している。

In the case of Exercise 1-2 has the 1st line is $y_1 = x_{11}a_1 + x_{12}a_2$ but solution indicates another line $\hat{y_1} = x_{11}\hat{a_1} + x_{12}\hat{a_2}$. Draw two lines.

2本の直線を描け。

同様に、第１、第２、第３の直線と、それに対応する、$(\hat{a_1}, \hat{a_2})$を通る３直線を描け。

Similarly, draw a pair of lines for 1st, 2nd and 3rd lines.

3.4 Exercise 1-4

第１の直線は、$(a_1 = \frac{y_1}{x_{11}}, 0)$と$(0, \frac{y_1}{x_{12}})$を通る直線である。

The 1st line connecting $(a_1 = \frac{y_1}{x_{11}}, 0)$ and $(0, \frac{y_1}{x_{12}})$.

それに対応する$(\hat{a_1}, \hat{a_2})$を通る直線は$(a_1 = \frac{\hat{y_1}}{x_{11}} = \frac{\hat{a_1}}{x_{11}}, 0)$と$(0, \frac{\hat{y_1}}{x_{12}} = \frac{\hat{a_2}}{x_{12}})$を通る直線である。

The corresponding estimated straight line goes through $(a_1 = \frac{\hat{y_1}}{x_{11}} = \frac{\hat{a_1}}{x_{11}}, 0)$ and $(0, \frac{\hat{y_1}}{x_{12}} = \frac{\hat{a_2}}{x_{12}})$.

These two lines are in parallel.

点$(\hat{a_1}, \hat{a_2})$を通り、この2直線と垂直な直線を引け。その直線の傾きは、$(x_{12}, -x_{11})$である。 この垂直な直線が第１の直線と交わる交点P_1を求めよ。

Draw a line perpendicular to the parallel two lines that has $(\hat{a_1}, \hat{a_2})$ on it. Its slope is $(x_{12}, -x_{11})$. Calculate the coordinate of point P_1 that is a crossing point of the 1st line and the perpendicular line.

$(\hat{a_1}, \hat{a_2})$と交点P_1との距離を求めよ。

Calculate the distance between $(\hat{a_1}, \hat{a_2})$ and P_1.

同様に、行列が定める３つの直線のそれぞれについて、行え。

Do similarly for three lines.

$||y_i - \hat{y_i}||$とこの距離の関係は何か。

Explain the relation between the calculated distances and $||y_i - \hat{y_i}||$.

3.5 Exercise 1-5

(a_1, a_2)平面の任意の点に関して、その点から、直線への垂線の足を求める関数を作成せよ。

Make a function that takes (a_1, a_2) and parameters of a straight line and returns the distance to the perpendicular foot from (a_1, a_2).

その関数を用いて、行列が定める３直線への足、３点が求まる。 $||y_i - \hat{y_i}||$がわかるので、$\sum_{i=1}^{3}(y_i - \hat{y_i})^2$も求まるはずである。

Using the function, you can calculate the distance of three line cases. You know $||y_i - \hat{y_i}||$ and you can calculate $\sum_{i=1}^{3}(y_i - \hat{y_i})^2$ as well.

この値が(a_1, a_2)平面に、どのような高低を作るかを図示し、$(\hat{a_1}, \hat{a_2})$の意味を確認せよ。

Draw the value of $||y_i - \hat{y_i}||$ as function of (a_1, a_2) and explain what $(\hat{a_1}, \hat{a_2})$ means.

3.6 Exercise 1-6

$n < m$

$$\left(y_1 \right) = \left(x_{11}, x_{12} \right) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

この場合、直線は１本引ける。 ムーアペンローズ疑似逆行列の解は直線上の1点である。

In this case only one line is drawn and Moore-Pensore psuedoinverse solution is one point on the line.

$||(a_1, a_2)||$が原点から最短距離になっていることを確かめよ。

$||(a_1, a_2)||$ is the closest point from the origin.

3.7 Exercise 1-7

$n = m$,$n > m$,$n < m$の３つの場合について、ムーアペンローズ疑似逆行列の解の幾何学的な意味を説明せよ。

For the three cases of $n = m$,$n > m$,$n < m$, describe geometric meaning of Moore-Penrose solution.

3.8 Exercise 1-8

$n < m$の場合に、$y = X\mathbf{a}$の解は、直線上のいずれの点でもよかったが、ムーアペンローズ疑似逆行列は、原点からの距離が最短になるような点$||(\hat{a_1}, \hat{a_2})||$を選んだ。

When n < m, the solution of $y = X\mathbf{a}$ can be any point on the straight line. Moore-Penrose psuedoinverse method specifically takes a point $||(\hat{a_1}, \hat{a_2})||$.

In machine learning, a procedure called regularization is taken to select THE solution among multiple candidates, with the following regularization formula.

正規化手法と呼ばれる手法では、

$$\sum_{i=1}^{n}(y_i - \hat{y_i})^2 + \lambda||\mathbf{a}||^p$$

を最小にすることで、残差$\sum_{i=1}^{n}(y_i - \hat{y_i})^2$を小さくことのみを目指すのではなく、$||\mathbf{a}||^p$という形で、解の取り方に制約を持たせること、その制約の強さをパラメタλでコントロールする。

This regularization attempts not to minimize the residuals but put a restriction by introducing the term with λ to select the “best” solution .

ムーアペンローズによる解の選び方と、正則化手法での解の選び方との類似性についてコメントせよ。

Do you have any comments on the Moore-Penrose solution and regularization?