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

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、

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

であった。

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。

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

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。

網羅されていないのは

- $n < m$の場合である

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。

$n < m$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

- $n \geq m$ 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$X^T (X^T X)^{-1}$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,] 1.347294 -2.0673768 0.2431391
## [2,] 1.180096 -0.7110244 0.1672885
```

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

```
##           [,1]       [,2]       [,3]
## [1,]  1.860546  0.5413008 -1.1951049
## [2,] -2.003575 -0.2372205  0.4930090
## [3,]  2.890075  1.1240450 -0.8753272
```

```
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,]  0.5906884  0.01287776
## [2,]  0.4153723  0.01793991
## [3,] -0.1167761 -0.56387476
```

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

```
##            [,1]        [,2]
## [1,]  0.5906884  0.01287776
## [2,]  0.4153723  0.01793991
## [3,] -0.1167761 -0.56387476
```

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

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

ムーアペンローズ疑似逆行列は

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

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。

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

```
##             [,1]      [,2]       [,3]
## [1,] -0.65929868  1.432681 -0.3101911
## [2,] -0.08738356 -0.806935  0.4711483
```

```
ginv(X)
```

```
##             [,1]      [,2]       [,3]
## [1,] -0.65929868  1.432681 -0.3101911
## [2,] -0.08738356 -0.806935  0.4711483
```

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

```
##           [,1]      [,2]       [,3]
## [1,] -25.51992 -7.142307 -12.741628
## [2,]  15.32677  3.927049   8.075916
## [3,] -23.13079 -6.570553 -12.417258
```

```
ginv(X)
```

```
##           [,1]      [,2]       [,3]
## [1,] -25.51992 -7.142307 -12.741628
## [2,]  15.32677  3.927049   8.075916
## [3,] -23.13079 -6.570553 -12.417258
```

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

```
##            [,1]      [,2]
## [1,] -1.3768647 0.3987604
## [2,] -1.4370526 1.6462194
## [3,]  0.5656033 1.2149662
```

```
ginv(X)
```

```
##            [,1]      [,2]
## [1,] -1.3768647 0.3987604
## [2,] -1.4370526 1.6462194
## [3,]  0.5656033 1.2149662
```

3 Exercise 1

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

$n = m$

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

今、a_1, a_2の値を求めるというのは、(a_1, a_2)を次元座標として、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。

$$x_{11}a_1 + x_{12}a_2 = y_1$$
$$x_{21}a_1 + x_{22}a_2 = y_2$$

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。

3.2 Exercise 1-2

$n > m$

\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。

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 1-3 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。

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。

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。

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。 これらは平行である。

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。

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

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。

$||y_i, \hat{y_i}||$とこの距離の関係は何か。

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。

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。

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。

3.6 Exercise 1-6

$n < m$

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

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。

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

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。

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。

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。

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。