

《统计计算与SAS软件》

实验7 IML过程、作图

1907402030 熊雄

2021年11月17日

1 实验目的

掌握SQL和IML过程的基础部分。

2 实验内容

2.1 IML过程

1. IML中矩矩阵的输入及显示；
2. IML中矩矩阵的拼接（水平及垂直拼接）；
3. IML中矩矩阵的基本运算（矩矩阵的加减乘，求逆，矩矩阵的数乘）；
4. IML中矩矩阵关于行，列求和，求平均，取最大值，最小值等；
5. IML中统计函数举例（自己找）；
6. IML中创建SAS数据集。

2.2 作图（用sashelp.class数据集作图）

1. 画height和weight两变量的散点图，数据点用“红色三角形”表示；
2. 画height和weight两变量的连线图，用不同的线型，颜色，粗细表示连线特征；
3. 画height的直方图、年龄的饼图（可以设置值，填充颜色、类型等）。

3 结果与分析

3.1 IML过程

3.1.1 IML中矩阵的输入及显示

输入以下SAS代码：

```
1 proc iml;
2     A = {[2]7 6,4 [2]2,[2]2 9};/*有重复元素时的列示法*/
3     B = 2 # A; /*赋值法*/
4     print A, B; /*按列排列*/
5 quit;
```

提交后便生成了矩阵 $A = \begin{bmatrix} 7 & 7 & 6 \\ 4 & 2 & 2 \\ 2 & 2 & 9 \end{bmatrix}$, $B = \begin{bmatrix} 14 & 14 & 12 \\ 8 & 4 & 4 \\ 4 & 4 & 18 \end{bmatrix}$ 。

3.1.2 IML中矩阵的拼接（水平及垂直拼接）

水平拼接：将具有相同行数的两个矩阵水平拼接为一个矩阵，使用||算符；

垂直拼接：将具有相同列数的两个矩阵垂直拼接为一个矩阵，使用//算符。

输入以下SAS代码：

```
1 proc iml;
2     A = {[2]7 6,4 [2]2,[2]2 9};/*有重复元素时的列示法*/
3     B = 2 # A; /*赋值法*/
4     C = A || B; /*水平拼接*/
5     D = A // B; /*垂直拼接*/
6     PRINT C,D;
7 quit;
```

提交后可以得到矩阵 $C = \begin{bmatrix} 7 & 7 & 6 & 14 & 14 & 12 \\ 4 & 2 & 2 & 8 & 4 & 4 \\ 2 & 2 & 9 & 4 & 4 & 18 \end{bmatrix}$, 矩阵 $D = \begin{bmatrix} 7 & 7 & 6 \\ 4 & 2 & 2 \\ 2 & 2 & 9 \\ 14 & 14 & 12 \\ 8 & 4 & 4 \\ 4 & 4 & 18 \end{bmatrix}$ 。

3.1.3 IML中矩阵的基本运算（矩阵的加减乘，求逆，矩阵的数乘）

输入以下代码，依次验证矩阵的加减乘，求逆与矩阵的数乘：

```

1  proc iml;
2      k = 5;
3      A = {[2]7 6,4 [2]2,[2]2 9};/*有重复元素时的列示法*/
4      B = 2 # A;    /*赋值法*/
5      E = A + B;    /*矩阵相加*/
6      F = B - A;    /*矩阵相减*/
7      G = A * B;    /*矩阵相乘*/
8      H = inv(A);    /*矩阵A的逆矩阵*/
9      I = k * A;    /*矩阵的数乘*/
10     PRINT E,F,G,H,I;
11  quit;

```

提交后得到结果查看器输出为：矩阵 $E = \begin{bmatrix} 21 & 21 & 18 \\ 12 & 6 & 6 \\ 6 & 6 & 27 \end{bmatrix}$ ，矩阵 $F = \begin{bmatrix} 7 & 7 & 6 \\ 4 & 2 & 2 \\ 2 & 2 & 9 \end{bmatrix}$ ，
 矩阵 $G = \begin{bmatrix} 178 & 150 & 220 \\ 80 & 72 & 92 \\ 80 & 72 & 194 \end{bmatrix}$ ，矩阵 $H = \begin{bmatrix} -0.137255 & 0.5 & -0.019608 \\ 0.3137255 & -0.5 & -0.098039 \\ -0.039216 & 0 & 0.1372549 \end{bmatrix}$ ，矩阵
 $I = \begin{bmatrix} 35 & 35 & 30 \\ 20 & 10 & 10 \\ 10 & 10 & 45 \end{bmatrix}$ 。

3.1.4 IML中矩阵关于行，列求和，求平均，取最大值，最小值等

输入以下代码：

```

1  proc iml;
2      A = {[2]7 6,4 [2]2,[2]2 9};/*有重复元素时的列示法*/
3      SUM_OF_ROW = A[,+]; /*矩阵A关于行求和*/
4      SUM_OF_COLUMN = A[+,]; /*矩阵A关于列求和*/
5      AVERAGE = A[:]; /*A的所有元素的平均*/
6      MAX = A[<>]; /*A中元素的最大值*/
7      MIN = A[><]; /*A中元素的最小值*/
8      PRINT A, SUM_OF_ROW, SUM_OF_COLUMN, AVERAGE, MAX, MIN;
9  quit;

```

得到输出如下：

A		
7	7	6
4	2	2
2	2	9

SUM_OF_ROW	
	20
	8
	13

SUM_OF_COLUMN		
13	11	17

AVERAGE
4.5555556

MAX
9

MIN
2

3.1.5 IML中统计函数举例

VEC_VAR: The variance of the elements of the first row of the matrix X;

MAT_VAR: The variance of the elements of the matrix X;

SM_VAR: The variance of the elements of a submatrix including the first two rows and the first two columns of the matrix X.

```

1  PROC IML;
2      x = {1 2 3,4 5 6,7 8 9};
3      ssq = SSQ(x[1,]);
4      sum = x[1,+];
5      vec_var = (ssq - sum * sum / NCOL(x)) / (NCOL(x) - 1);
6      jn = J(NROW(x));
7      mat_mean = x[:];
8      mat_var = SSQ(x - mat_mean * jn) / (NROW(x) * NCOL(x) - 1);
9      submat = x[1:2,1:2];
10     ssq_sub = SSQ(submat);
11     sum_sub = SUM(submat);
12     sm_var = (ssq_sub - sum_sub * sum_sub / (NCOL(submat) *
NROW(submat))) / (NCOL(submat) * NROW(submat)-1);
13     PRINT vec_var mat_var sm_var;
14  QUIT;

```

提交后得到输出如下:

vec_var	mat_var	sm_var
1	7.5	3.3333333

3.1.6 IML中创建SAS数据集

输入以下代码：

```

1  proc iml;
2      use sashelp.class;
3      read all var{name age height weight};
4      hwt = height / weight;
5      new = age || hwt;
6      create ratio from new[colname={'age' 'h_w'}];
7      append from new;
8      close ratio;
9  quit;
10 proc print;
11 run;

```

输出如下：

Obs	age	h_w
1	14	0.61333
2	13	0.67262
3	13	0.66633
4	14	0.61268
5	14	0.61951
6	12	0.69036
7	12	0.70769
8	15	0.55556
9	13	0.74405
10	12	0.59296
11	11	1.01584
12	14	0.71444
13	12	0.73117
14	15	0.59375
15	16	0.48000
16	12	0.50625
17	15	0.50376
18	11	0.67647
19	15	0.59375

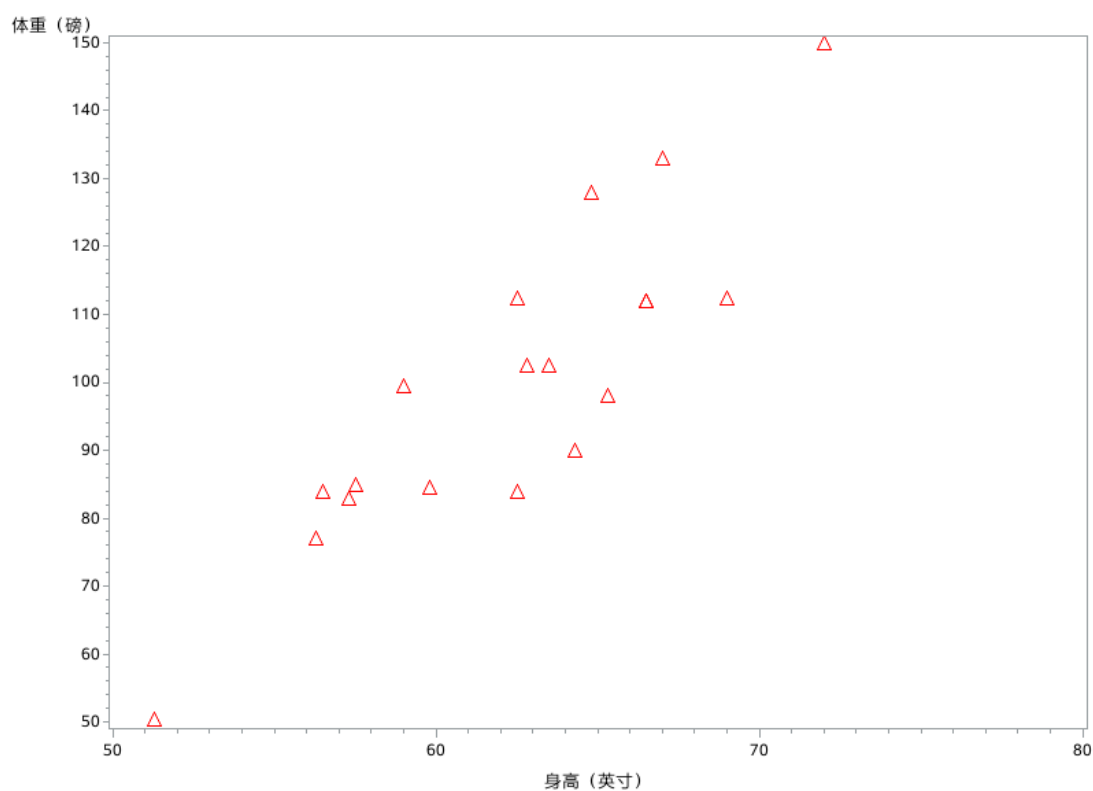
3.2 作图（用sashelp.class数据集作图）

3.2.1 height和weight两变量的散点图，数据点用“红色三角形”表示

输入以下SAS代码：

```
1 proc gplot data = sashelp.class;
2     symbol v = TRIANGLE, cv = RED, h = 2;
3     plot weight * height;
4 run;
```

提交后得到如下散点图：

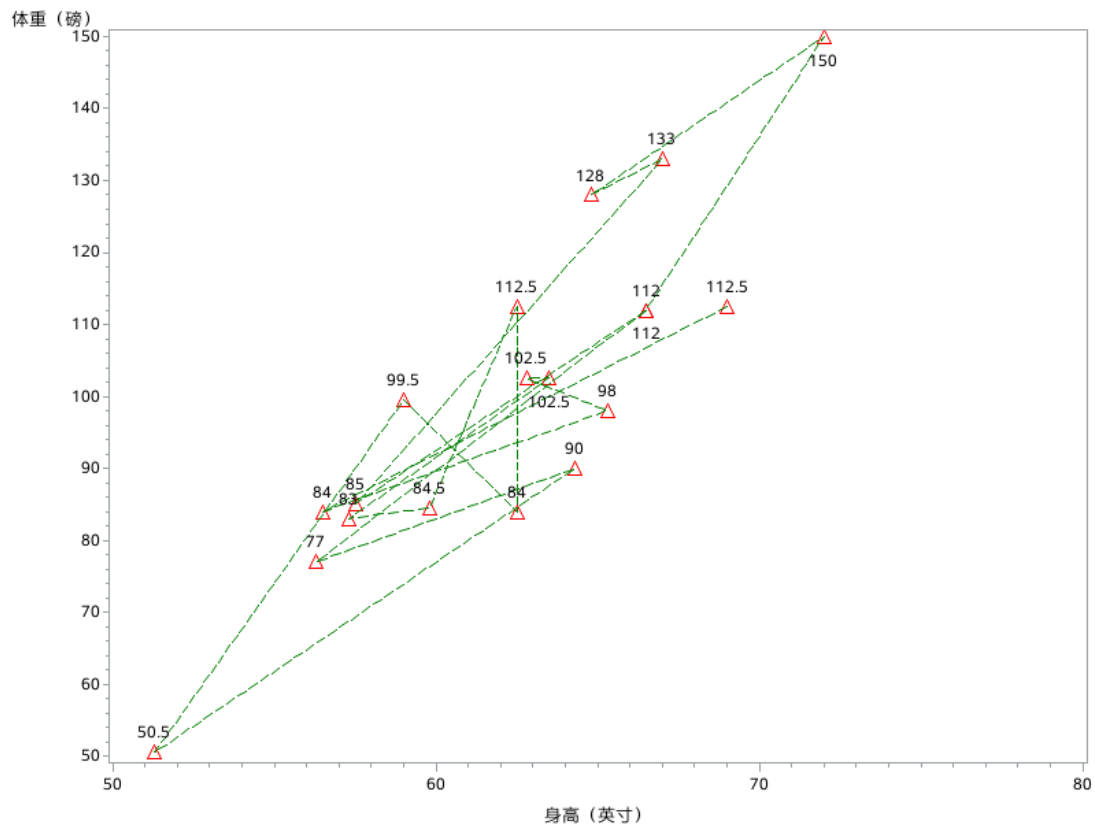


3.2.2 height和weight两变量的连线图

输入以下SAS代码：

```
1 proc gplot data = sashelp.class;
2     symbol i = JOIN, CI = GREEN, v = TRIANGLE, LINE = 3, cv = red, h
3     = 2, pointlabel;
4     plot weight * height;
```

提交后得到如下散点图：



3.2.3 画height的直方图、年龄的饼图

3.2.3.1 height的直方图

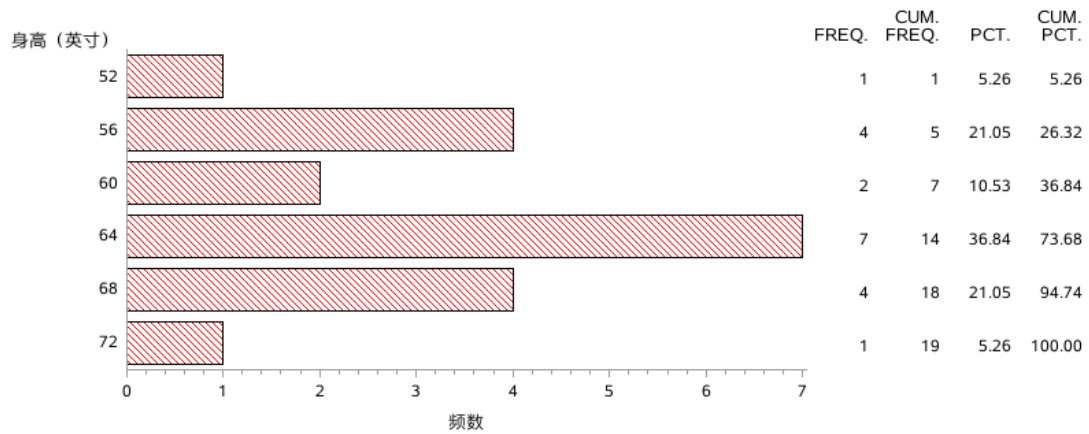
输入以下SAS代码：

```

1  GOPTIONS RESET = ALL;
2  proc gchart data = sashelp.class;
3      HBAR height / levels = 6 CFRAME = WHITE NOFRAME;
4      PATTERN V = L2, C = red;
5  run;

```

提交后得到如下柱状图：



3.2.3.2 age的饼图

输入以下SAS代码:

```

1  GOPTIONS RESET = ALL;
2  proc gchart data = sashelp.class;
3      PIE age / slice = arrow percent = inside value = arrow;
4      PATTERN1 v = p3x70, c = green;
5      PATTERN2 v = p2n45, c = red;
6      PATTERN3 v = p5, c = blue;
7      PATTERN4 v = ps, c = orange;
8      PATTERN5 v = p2n45, c = brown;
9  run;

```

提交后得到如下饼图:

