

数据采集方法作业

姓名：蒋贵豪 学号：B+X9bo

2021 年 11 月 29 日

题目 1. 对于一个物质提取试验，共有 6 个试验因子：甲醇 (x_1)、乙醇 (x_2)、丙醇 (x_3)、丁醇 (x_4)、PH 值 (x_5)、时间 (x_6)。我们的试验次数少于 15 次，有先验信息表明， $\text{INT}(x_2, x_3)$, $\text{INT}(x_5, x_6)$ 显著。利用坐标交换算法，寻找 D-最优设计，计算相应设计的 D 值。并说明选择初始设计的个数。

解答. 因为非正规设计中，试验次数为 4 的倍数，而我们需要次数需要少于 15 次，因此我们选择试验次数 $n = 12$ ，因为我们有 6 个试验因子，因此我们的设计矩阵为 $\mathbf{D}_{12 \times 6}$ 。

运用坐标变换算法，我们的试验因子都是连续的变量，因此我们对设计矩阵 \mathbf{D} 中的每个位置 x_{ij} ，均产生 $[-1, 1]$ 上的随机数，然后改变 x_{ij} 为 -1 和 1 ，计算对应的 $\mathbf{X}^T \mathbf{X}$ 的行列式，然后选取 x_{ij} 为 -1 或 1 。重复上述过程，直到信息矩阵行列式不再增大。

我们编写的 **Matlab** 代码见附录，下面是我们寻找 D-最优设计的结果。为方便展示，我们对得到的设计矩阵 \mathbf{D} 按因子水平取值进行了排序。当初始设计的个数 $N = 100$ 时，得到的设计矩阵为如表1所示，信息矩阵如式1所示。对应的信息矩阵的行列式为： $|\mathbf{X}^T \mathbf{X}| = 2684354560$ 。

$$\mathbf{X}^T \mathbf{X} = \begin{bmatrix} 12 & 0 & 0 & 0 & -4 & 0 & 0 & 0 & 0 \\ 0 & 12 & 0 & 4 & 0 & -4 & 0 & 0 & 0 \\ 0 & 0 & 12 & 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 4 & 0 & 12 & 0 & -4 & 0 & 0 & 0 \\ -4 & 0 & 0 & 0 & 12 & 0 & 0 & 0 & 0 \\ 0 & -4 & 0 & -4 & 0 & 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 12 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 12 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 12 \end{bmatrix} \quad (1)$$

表 1: 初始设计个数 $N = 100$ 得到的设计矩阵 \mathbf{D}

试验号	x_1	x_2	x_3	x_4	x_5	x_6
1	-1	-1	-1	-1	-1	1
2	-1	-1	-1	-1	1	-1
3	-1	-1	1	1	1	-1
4	-1	1	-1	-1	1	1
5	-1	1	-1	1	-1	-1
6	-1	1	1	-1	1	1
7	1	-1	-1	1	1	1
8	1	-1	1	-1	-1	-1
9	1	-1	1	-1	-1	1
10	1	1	-1	-1	1	-1
11	1	1	1	-1	-1	-1
12	1	1	1	1	-1	1

改变初始设计的个数为 $N = 1000$, 得到的设计矩阵为如表2所示, 信息矩阵如式2所示。对应的信息矩阵的行列式为: $|\mathbf{X}^T \mathbf{X}| = 2818572288$ 。

$$\mathbf{X}^T \mathbf{X} = \begin{bmatrix} 12 & 4 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\ 4 & 12 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 12 & 0 & 0 & -4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 12 & 0 & 0 & 4 & 0 & 0 \\ 4 & 0 & 0 & 0 & 12 & 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 & 0 & 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 12 \end{bmatrix} \quad (2)$$

表 2: 初始设计个数 $N = 1000$ 得到的设计矩阵 \mathbf{D}

试验号	x_1	x_2	x_3	x_4	x_5	x_6
1	-1	-1	-1	-1	1	-1
2	-1	-1	1	1	1	1
3	-1	1	-1	1	-1	1
4	-1	1	1	1	-1	-1
5	1	-1	-1	1	-1	-1
6	1	-1	-1	1	1	1
7	1	-1	1	-1	-1	1
8	1	-1	1	1	1	-1
9	1	1	-1	-1	-1	-1
10	1	1	-1	1	1	-1
11	1	1	1	-1	1	1
12	1	1	1	1	-1	1

同样，我们改变初始设计的个数为 $N = 10000$ ，得到的设计矩阵和信息矩阵的结果同 $N = 1000$ 。说明我们在 $N = 1000$ 得到的解已经接近了全局最优，因此我们的 D-最优设计如表2所示。也就是：

$$\mathbf{D} = \begin{bmatrix} -1 & -1 & -1 & -1 & 1 & -1 \\ -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & 1 \end{bmatrix} \quad (3)$$

附录

本次作业用到的所有代码如下：

```
1 N = 100;%重复次数，获取全局最优
2 rng(1129);
3 Information_Max = 0;
4 for k = 1:N
5     D = 2*rand(12,6)-1;
6     X(:,1) = linspace(1,1,12)';
7     for i = 1:12
8         for j = 1:6
9             D(i,j) = -1;
10            X(:,2:7) = D;
11            X(:,8) = D(:,2).*D(:,3);
12            X(:,9) = D(:,5).*D(:,6);
13            a = det(X'*X);
14            D(i,j) = 1;
15            X(:,2:7) = D;
16            X(:,8) = D(:,2).*D(:,3);
17            X(:,9) = D(:,5).*D(:,6);
18            b = det(X'*X);
19            if a ≥ b
20                D(i,j) = -1;
21            end
22        end
23    end
24    if det(X'*X) > Information_Max
25        Information_Max = det(X'*X);
26        D_k = D;
27    end
28 end
```