MAE 5350: HW #2
Multidisciplinary Design Optimization

Hanfeng Zhai

hz253@cornell.edu

Sibley School of Mechanical and Aerospace Engineering, Cornell University

October 9, 2021

Q1. Design of experiments

Recall the airplane design experiment we did in Lecture 4. You can download the results from Canvas. (*The code for Q1 can be downloaded through* https://hanfengzhai.net/data/MDO_A2_Q1.mlx)

(a) Calculate the mean and variance for each experiment (9 experiments).

SOLUTION: The mean value is calculated through

$$\overline{\mathcal{D}}_i = \frac{\sum_i^N \mathcal{D}_i}{N}$$

where $\mathcal{D}$ stands for the distance and N stands for the total numbers of attempts, i is the experiments numbers. Based on such we can calculate the following mean for the nine experiments in Table 1.

The variance is calculated through the absolute minus between the mean values per experiments to the overall mean value, written as $\mathcal{V}_i$:

$$\mathcal{V}_i = \frac{\sum_{i=1}^N (x_i - \overline{x_i})^2}{N}$$

The variance for the nine experiments in Table 2.

(b) Calculate the effect of each design variable setting (12 effects).

$\overline{\mathcal{D}}_1$	$\overline{\mathcal{D}}_2$	$\overline{\mathcal{D}}_3$	$\overline{\mathcal{D}}_4$	$\overline{\mathcal{D}}_5$	$\overline{\mathcal{D}}_6$	$\overline{\mathcal{D}}_7$	$\overline{\mathcal{D}}_8$	$\overline{\mathcal{D}}_9$
19.415	16.43	10.94	15.138	17.207	18.438	21.955	16	14.855

Table 1: The mean value for the nine experiments.

$\overline{\mathcal{V}}_1$	$\overline{\mathcal{V}}_2$	$\overline{\mathcal{V}}_3$	$\overline{\mathcal{V}}_4$	$\overline{\mathcal{V}}_5$	$\overline{\mathcal{V}}_6$	$\overline{\mathcal{V}}_7$	$\overline{\mathcal{V}}_8$	$\overline{\mathcal{V}}_9$
3.2547	22.8058	2.7175	19.0957	45.8479	157.8374	0.0156	7.1289	0.7731

Table 2: The variance for the nine experiments.

$\overline{\mathcal{M}}_{A_1}$	$\overline{\mathcal{M}}_{A_2}$	$\overline{\mathcal{M}}_{A_3}$	$\overline{\mathcal{M}}_{B_1}$	$\overline{\mathcal{M}}_{B_2}$	$\overline{\mathcal{M}}_{B_3}$
-0.9944	-0.0366	0.2075	0.9913	-0.0912	-1.9645
$\overline{\mathcal{M}}_{C_1}$	$\overline{\mathcal{M}}_{C_2}$	$\overline{\mathcal{M}}_{C_3}$	$\overline{\mathcal{M}}_{D_1}$	$\overline{\mathcal{M}}_{D_2}$	$\overline{\mathcal{M}}_{D_3}$
1.6323	-1.1568	-1.0587	0.4505	1.3113	-2.8262

Table 3: The main effect for the twelve variables.

SOLUTION: Calculating the main effect of A_i, named as $\mathcal{M}_{A_i}$:

$$\mathcal{M}_{A_i} = \overline{\mathcal{D}_{(A_i)}} - \overline{\sum_i^N \overline{\mathcal{D}}_i}$$

With such an equation we compute the effects for 12 variables as in Table 3.

(c) What are the design variable settings of the predicted "optimal" airplane?

SOLUTION: From Table 3 comparing $\mathcal{M}_{A_i}$ we can deduce that the optimal design is A_3. Similarly, observing $\mathcal{M}_{B_i}$ we can deduce that the optimal design is B_1; observing $\mathcal{M}_{C_i}$ we can deduce that the optimal design is C_1. And similarly we have D_2 for optimal design.

(d) Assuming that the effects can be added linearly, estimate the range of the predicted optimal airplane.

SOLUTION:

Since the main effects can be added linearly, then the optimal design is $\{A_3,\ B_1,\ C_1,\ D_2\}$. We first calculate the sum main effect by adding the main effects:

$$\sum \overline{\mathcal{M}} = \overline{\mathcal{M}}_{A_3} + \overline{\mathcal{M}}_{B_1} + \overline{\mathcal{M}}_{C_1} + \overline{\mathcal{M}}_{D_2} = 4.1424$$

Then we can calculate the range by adding such to the overall mean

$$\sum \mathcal{M} + \overline{\mathcal{D}} = 20.4963$$

(e) Download the airplane template from Canvas, build the predicted optimal airplane, and fly it 5 times*. Report your results and discuss similarities/differences between the predicted and actual performance.

SOLUTION:

Expts. #	1	2	3	4	5
Dist. (ft)	29.8	25.3	25.9	18.2	20.1

Expts. # / Factor	A	B	C	D
1	A_3	B_1	C_1	D_2
2	A_2	B_1	C_1	D_2
3	A_1	B_1	C_1	D_2
4	A_3	B_2	C_1	D_2
5	A_3	B_3	C_1	D_2
6	A_3	B_1	C_2	D_2
7	A_3	B_1	C_3	D_2
8	A_3	B_1	C_1	D_1
9	A_3	B_1	C_1	D_3

Table 4: The experiment matrix for parameter study.

Most of the experimental distance was larger than the predicted value, three reasons might be accounted for this "inaccuracy": (1) I got better throwing skills (doubt about it); (2) There are errors for my measurement; (3) Errors of measurement by other students.

(f) If we were now to perform a parameter study using the predicted optimal airplane as the baseline, what would the experimental matrix be? Comment on what, if any, new information this new experiment might bring.

SOLUTION:

Compared with the original design table, this table could present us information on when three parameters are fixed at the optimal design, whether the changing of the other parameters may variate the results. It also give us numerous experiments that the original table does not cover. So we may be able to see how each factors variate the final distance more specifically through different experiments.

(g) In one paragraph, discuss the variance results computed in part (a). Explain the possible significance of these results and how they might be used to inform an actual design process.

SOLUTION: In probability theory and statistics, variance is the expectation of the squared deviation of a random variable from its population mean or sample mean. Variance is a measure of dispersion, meaning it is a measure of how far a set of numbers is spread out from their average value [Wikipedia]. Based on such definition, we can deduce that the bigger the variant the higher error of the experiments. Or in other words, when we are approaching the more "accurate" experimental results we should expect to see a smaller variance. Here, from Table 2 we can see that experiments # 2, 4, 5, 6 have a general larger variance, especially for Exp. # 6. Here we can say that the reliability of these experiments are lower than the rest.

Q2. Gradient-Based Optimization: optimal can sizing problem

Consider the simple constrained optimization problem of minimizing the surface area of a cylinder subject to an equality constraint on its volume:

$$\begin{aligned} \min_{x} f(x_1, x_2) &= 2\pi x_1(x_1 + x_2) \\ \text{subject to} \quad h(x_1, x_2) &= \pi x_1^2 x_2 - V = 0 \end{aligned}$$

where x_1 is the radius of the cylinder, x_2 is the height of the cylinder, and V is the required volume.

(a) Formulate the Lagrangian function, derive the optimality conditions, and solve the resulting system of equations to determine the dimensions of the minimum-surface-area cylinder that has a volume of 1 liter (1000 cm^3). (We know $V = 0.001 m^3$ (SI Unit))

Solution: According to the definition of a Lagrangian function:

$$L(\mathbf{x}, \lambda) = J(\mathbf{x}) + \sum_{j=1}^{m_1} \lambda_i g_j(\mathbf{x}) + \sum_{k=1}^{m_2} \lambda_{m_1+k} h_k(\mathbf{x})$$

Substituting the given condition to the function L we have:

$$L(\mathbf{x}, \lambda) = 2\pi x_1(x_1 + x_2) + \lambda(\pi x_1^2 x_2 - V) \tag{1}$$

By solving $\nabla_{x_1,x_2,\lambda} L(\mathbf{x}, \lambda) = 0$, we have:

$$\left(\frac{\partial L}{\partial x_1}, \frac{\partial L}{\partial x_2}, \frac{\partial L}{\partial \lambda}\right) = 0 \Longleftrightarrow \begin{cases} 2\pi x_1 + 2\pi(x_1 + x_2) + 2\pi\lambda x_1 x_2 = 0 \\ \lambda \pi x_1^2 + 2\pi x_1 = 0 \\ \pi x_2 x_1^2 - V = 0 \end{cases} \tag{2}$$

Solving the above equations, applying the `vpasolve` function we have[1]:

$$\begin{aligned} x_1 &= 0.054192607013928900874456136482964 \\ x_2 &= 0.10838521402785780174891227296593 \\ \lambda &= -36.905402972880568381936077591 78 \end{aligned}$$

NOTE THAT THE UNIT IN THIS COMPUTATION IS IN **SI UNIT** (*The answer hasn't change since my last version but I switched all the answers to SI Unit to make it more clear and avoid confusion*) I switched the answer from the original cm to m is because Jiayi asked how did I solve the problem and

[1] Codes can be downloaded through https://hanfengzhai.net/data/MDO_A2_Q2.mlx

I explained it to him, and he get confused about my unit, so here I switch all the unit to SI Unit to avoid similar issues.

(b) Repeat the optimization for a constrained volume of 12 US fl oz (355 ml) and compare your dimensions against the standard U.S. beverage can (https://en.wikipedia.org/wiki/Beverage_can) and explain any differences.

SOLUTION: For this problem, an inequality constraint was added as $V \leq 355 \times 10^{-6}[m^3]$, which is written as $V - 0.355 = 0$ in the standard form. Hence the Lagrangian (Eq. 1) should be rewritten in the form:

$$\min_x f(x_1, x_2) = 2\pi x_1(x_1 + x_2)$$

$$\text{subject to} \quad h(x_1, x_2) = \pi x_1^2 x_2 = V, \text{ where V} = 355 \times 10^{-6}$$

$$\rightarrow L(\mathbf{x}, \lambda) = 2\pi x_1(x_1 + x_2) + \lambda(\pi x_1^2 x_2 - 355 \times 10^{-6})$$

With the given setup, we write out the Karush-Kuhn-Tucker (KKT) Conditions:

$$\left(\frac{\partial L}{\partial x_1}, \frac{\partial L}{\partial x_2}, \frac{\partial L}{\partial \lambda}\right) = 0 \Longleftrightarrow \begin{cases} 2\pi x_1 + 2\pi(x_1 + x_2) + 2\pi\lambda x_1 x_2 = 0 \\ \lambda\pi x_1^2 + 2\pi x_1 = 0 \\ \pi x_2 x_1^2 - V = 0 \end{cases} \tag{3}$$

By solving Eq. 3 with `vpasolve` we have:

$$x_1 = 0.038372152480156730624817304791733$$

$$x_2 = 0.076744304960313461249634609583465$$

$$\lambda = -52.121131360411789506046960254401$$

According to the provided reference, *The US standard can is 4.83 in or 12.3 cm high, 2.13 in or 5.41 cm in diameter at the lid, and 2.6 in or 6.60 cm in diameter at the widest point of the body.* So here our optimized results can be identified as a "shorter" and "fatter" version compared with the standard soda can.

(c) Formulate the can optimization problem as a geometric program in standard form.

SOLUTION: Write such a problem into geometric form:

$$\min_x 2\pi x_1^2 + x_1 x_2$$

$$\text{subject to} \quad 2.8169 \times 10^3 \pi x_1^2 x_2 = 1$$

(d) Use Python/MATLAB/R/Julia to independently solve the problem numerically using either the cvxpy library for Python, the CVX library for MATLAB, CVXR for R or Convex.jl in Julia*. Does the solution agree with your analytical solution found in part (a)?

SOLUTION: We construct the following code based on the course materials:

```
cvx_begin gp
    variables x1 x2 lambda
    minimize( 2*pi*x1*(x1 + x2) )
    subject to
        pi*x1*x2 == 355e-6
cvx_end
```

And we obtain the following results:

$$x_1 = 8.4782 \times 10^{-21}$$

$$x_2 = 1.3328 \times 10^{16}$$

$$\lambda = 1$$

Which is very different from what we obtain in (a) with gradient based methods.

Q3. Gradient-Based Optimization: Newton's method

Consider the function

$$f(x_1, x_2) = \frac{1}{2}(x_1^2 - x_2)^2 + \frac{1}{2}(1 - x_1)^2$$

(a) At what point does f attain a minimum?

SOLUTION: Solving $\nabla_{x_1} f = 0$ & $\nabla_{x_2} f = 0$, we have $x_1 = x_2 = 1$.

(b) Perform (by hand) one iteration of Newton's method using the starting point:$x_1 = 2, x_2 = 2$.

SOLUTION: Based on Newton's method, the search direction can be computed as

$$\mathbf{S} = -[\mathbf{H}(x^0)]^{-1}\nabla J(x^0)$$

Here, $J = f(\mathbf{x})$, and $x^0 = (2, 2)$. So we first compute the gradient and Hessian:

$$\nabla J = \begin{bmatrix} x_1 - 2x_1(-x_1^2 + x_2) - 1 \\ -x_1^2 + x_2 \end{bmatrix} \ \& \ \mathbf{H} = \begin{bmatrix} 6x_1^2 - 2x_2 + 1 & -2x_1 \\ -2x_1 & 1 \end{bmatrix}$$

For the first step, substituting the first coordinate into the direction $\mathbf{S}$ we have

$$\mathbf{S}(2, 2) = -[\mathbf{H}(2, 2)]^{-1}\nabla J(2, 2) = [-0.2, 1.2]^T$$

Q4. Gradient-Based Optimization: constraint qualifications

Consider the problem

$$\min_x f(x_1, x_2) = x_1^2 + x_2^2$$

$$\text{subject to } h(x_1, x_2) = x_2^2 - (x_1 - 1)^3 = 0$$

(a) Formulate the Lagrangian function and derive the KKT optimality conditions. Can you solve the resulting system of equations to determine the optimal solution? Explain why this method might fail for this problem.

SOLUTION: We first formulate the Lagrangian function:

$$L(\mathbf{x}, \lambda) = J(\mathbf{x}) + \sum_{j=1}^{m_1} \lambda_i g_j(\mathbf{x}) + \sum_{k=1}^{m_2} \lambda_{m_1+k} h_k(\mathbf{x})$$

Substitute $f = J$ and the equality constraint h we have:

$$L(\mathbf{x}, \lambda) = x_1^2 + x_2^2 + \lambda \left(x_2^2 - (x_1 - 1)^3\right)$$

By solving the Lagrangian we have

$$\left(\frac{\partial L}{\partial x_1}, \frac{\partial L}{\partial x_2}, \frac{\partial L}{\partial \lambda}\right) = 0 \Longleftrightarrow \begin{cases} 2\,x_1 - 3\,\lambda\,(x_1 - 1)^2 = 0 \\ 2\,x_2 + 2\,\lambda\,x_2 = 0 \\ {x_2}^2 - (x_1 - 1)^3 = 0 \end{cases} \tag{4}$$

We then derive the KKT condition: if there exists a feasible optimal point $\mathbf{x}^*$,

$$\lambda_i g_j(\mathbf{x}^*) = 0,\ j = 1, ..., m_i, \lambda_i \geq 0$$

$$\begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + \lambda_i \begin{bmatrix} -3\,(x_1 - 1)^2 \\ 2\,x_2 \end{bmatrix} = 0$$

Try solving this equation with MATLAB Built-in function `vpasolve` we generate the following codes:

```
>> syms x1 x2 lambda
>> L = x1^2 + x2^2 + lambda*(x2^2 - (x1 - 1)^3);lagran = [diff(L,x1); diff(L,x2); diff(L,lambda)]
>> eq1 = lagran(1); eq2 = lagran(2); eq3 = lagran(3)
>> [x1sol , x2sol , lambdasol] = vpasolve([eq1,eq2,eq3],[x1,x2,lambda])
x1sol =
0.66666666666666666666666666666667 - 0.74535599249992989880305788957709i
0.66666666666666666666666666666667 + 0.74535599249992989880305788957709i
0.66666666666666666666666666666667 + 0.74535599249992989880305788957709i
```

```
0.66666666666666666666666666666667 - 0.74535599249992989880305788957709i

x2sol =
- 0.72898887936316829444036282580464 - 0.11360575564930422273334997589706i
- 0.72898887936316829444036282580464 + 0.11360575564930422273334997589706i
  0.72898887936316829444036282580464 - 0.11360575564930422273334997589706i
  0.72898887936316829444036282580464 + 0.11360575564930422273334997589706i

lambdasol =
-1.0
-1.0
-1.0
-1.0
```

From the results we can then deduce that the system cannot be solved. Based on the *Nunemacher, 2003* paper we know that the reason why the Lagrangian multiplier fails when the constraints' geometry is not smooth or the critical point is neglected ("*But to be assured that the method succeeds, we must know that the geometry is right—that is, the set defined by $g(x, y) = k$ is a smooth curve in the plane*")

(b) Solve this problem using the exterior penalty method discussed in L8, using a quadratic penalty function. (Hint: derive an expression for the solution as a function of the penalty parameter ρ, and then take the limit as $\rho \to +\infty$)

SOLUTION: we first formulate a pseudo-objective Φ_Q with the Quadratic Penalty Function for exterior penalty method:

$$\Phi_Q(\mathbf{x}, \rho_p) = x_1^2 + x_2^2 + \rho_p \left(x_2^2 - (x_1 - 1)^3\right)^2$$

We then first calculate the gradient of the pseudo-objective function:

$$\nabla\Phi_Q = \begin{pmatrix} 2\,x_1 + 6\,\rho_p \left((x_1 - 1)^3 - {x_2}^2\right)(x_1 - 1)^2 \\ 2\,x_2 - 4\,\rho_p\,x_2 \left((x_1 - 1)^3 - {x_2}^2\right) \end{pmatrix}$$

and also the Hessian:

$$\mathbf{H}\Phi_Q =$$
$$\begin{pmatrix} 18\,\rho_p\,(x_1 - 1)^4 + 6\,\rho\,(2\,x_1 - 2)\left((x_1 - 1)^3 - {x_2}^2\right) + 2 & -12\,\rho_p\,x_2\,(x_1 - 1)^2 \\ -12\,\rho_p\,x_2\,(x_1 - 1)^2 & 8\,\rho_p\,{x_2}^2 - 4\,\rho_p\left((x_1 - 1)^3 - {x_2}^2\right) + 2 \end{pmatrix}$$

Here, by analyzing the second term of the gradient $\nabla\Phi_Q$ we can rewrite it into the form:

$$2x_2\left(1 - 4\rho_p((x_1 - 1)^3 - x_2^2)\right)$$

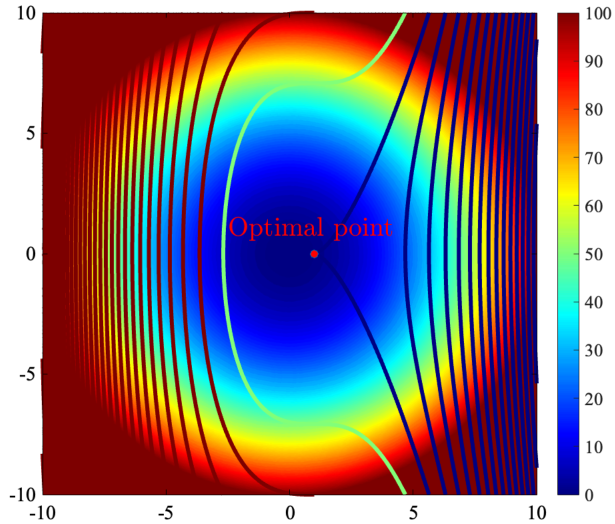

Figure 1: The sketch for the objective and constraint.

And we can deduce that if we let $x_2 = 0$ agrees for $\nabla \Phi_Q = 0$; therefore the first term of the gradient can be written as $2x_1 + 6\rho_p(x_1 - 1)^5 = 0$, taking $\rho_p \to \infty$ we get $x_1 = 1$.

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

(c) Plot the contours of the objective function and the equality constraint. Explain how the plot relates to your answers in (a) and (b).

The codes for running the program are shown as belows:

```
x1 = -10:0.1:10;x2 = x1;
[x1,x2] = meshgrid(x1,x2);
f =x1.^2 + x2.^2;
h  = x2.^2 - (x1 - 1).^3;
x1 = -10:0.1:10;x2 = x1;
contour(x1,x2,f);hold on
contour(x1,x2,h)
```

The figure containing the objective and the constraint are shown as in the following figure, where our estimated optimal point are plotted in the red dot.

Appendix. Data & Code

Code for Q1

The following code used for Q1 is written in MATLAB R2021a.

```
>> %% --------------Code for Q1 - (a)--------------
```

```
>> %% Calculating the mean value
>> D_1 = mean(data(1,2:5)); D_2 = mean(data(2,2:7)); D_3 = mean(data(3,2:5)); D_4 = mean(data(4,2:7));
    D_5 = mean(data(5,2:5)); D_6 = mean(data(6,2:5)); D_7 = mean(data(7,2:3)); D_8 = mean(data(8,2:3));
     D_9 = mean(data(9,2:5));
>> D_matrix = [D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9]
D_matrix =
       19.415   16.43   10.94   15.138  17.207  18.438   21.955   16  14.855
>> % Calculate the overall mean value
>> D_overall = mean(D_matrix)
D_overall =
       16.709
>> %% Calculate the variance
>> V_1 = D_1 - D_overall; V_2 = D_2 - D_overall; V_3 = D_3 - D_overall; V_4 = D_4 - D_overall; V_5 =
    D_5 - D_overall; V_6 = D_6 - D_overall; V_7 = D_7 - D_overall; V_8 = D_8 - D_overall; V_9 = D_9 -
    D_overall;
>> V_matrix = [V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9]
V_matrix =
       2.7063   -0.2787 -5.7687 -1.5704 0.4988  1.7288  5.2463   -0.7087 -1.8537
```

```
>> %% --------------Code for Q1 - (b)--------------
>> M_A_1 = mean([D_1,D_2,D_3]) - D_overall; M_A_2 = mean([D_4,D_5,D_6]) - D_overall; M_A_3 = mean([D_7,
    D_8,D_9]) - D_overall;
>> M_A = [M_A_1, M_A_2, M_A_3];
>> M_B_1 = mean([D_1,D_4,D_7]) - D_overall; M_B_2 = mean([D_2,D_5,D_8]) - D_overall; M_B_3 = mean([D_3,
    D_6,D_9]) - D_overall;
>> M_B = [M_B_1, M_B_2, M_B_3];
>> M_C_1 = mean([D_1,D_6,D_8]) - D_overall; M_C_2 = mean([D_2,D_4,D_9]) - D_overall; M_C_3 = mean([D_3,
    D_5,D_7]) - D_overall;
>> M_C = [M_C_1, M_C_2, M_C_3];
>> M_all = [M_A; M_B; M_C]
M_all =
      -1.1137       0.21907       0.89463
       2.1274      -0.16287       -1.9645
       1.2421       -1.2343    -0.0078704
```