National Cheng Kung University

Optimization Design Homework 1

Andrés Ponce 彭思安 P76107116

March 18, 2022

1 Find the stationary points for the following functions. Also identify(for each stationary point), the local maximum, minimum, or neither (by using second order derivative or Hessian matrix.)

1.1 $f(x) = x^3 \exp(-x^2)$, for $-2 < x < 2$.

The stationary points are those where $f'(x) = 0$, so we first find the first derivative. By using the product rule,

$$f'(x) = 3x^2 \exp(-x^2) - 2x^4 \exp(-x^2) = -x^2 \exp(-x^2)(2x^4 - 3x^2) \tag{1.1}$$

which is zero at $x = 0$, and by solving $2x^4 - 3x^2$ we get the other roots $x = \pm\frac{\sqrt{3}}{\sqrt{2}}$. Theorem 2.2 gives the sufficient condition for a minimum or maximum point for single variables. We must find a point where $f^m(x^*) \neq 0$, so we find $f''(x)$. We again use the product rule on Equation 1.1 and factor the result to obtain

$$f''(x) = -x \exp(-x^2)(4x^4 + 2x^2 - 6) \tag{1.2}$$

where $f''(\frac{\sqrt{3}}{\sqrt{2}}) < 0$ and $f''(-\frac{\sqrt{3}}{\sqrt{2}}) > 0$ However, $f''(0) = 0$, so we find the third derivative by using the same procedure

$$f'''(x) = -x \exp(-x^2)(20x^3 + 6x + 8x^4 + 4x^2 - 12) - 6\exp(-x^2) \tag{1.3}$$

where $f'''(0) = 6$, however since $n = 3$, 0 does not correspond to either a maximum nor minimum point. In the end, $-\frac{\sqrt{3}}{\sqrt{2}}$ is a minimum, 0 is neither, and $\frac{\sqrt{3}}{\sqrt{2}}$ is a maximum.

1.2 $f(x,y) = -x^2 - 3y^2 + 12xy$

For a multivariable equation, solving for x and y in the partial derivatives will give us the stationary points.

$$\frac{\partial f}{\partial x} = -2x + 12y \tag{1.4}$$

$$\frac{\partial f}{\partial y} = -6y + 12x \tag{1.5}$$

We first solve for y in Equation 1.5 and we get $y = 2x$. Substituing for y in Equation 1.4, we get

$$-2x + 24x = 22x$$

$$22x = 0$$

which means $x = 0$. Substituting again $x = 0$ in Equation 1.5 we get $-6y = 0$ which means $y = 0$. The stationary point for f is $(0,0)$.
To determine the nature of the stationary points, we have to find the Hessian matrix, which involves finding the second order partial derivatives.

$$\frac{\partial^2 f}{\partial^2 x} = -2 \quad \frac{\partial^2 f}{\partial^2 y} = -6 \tag{1.6}$$

$$\frac{\partial^2 f}{\partial y \partial x} = 12 \quad \frac{\partial^2 f}{\partial x \partial y} = 12 \tag{1.7}$$

and build the Hessian matrix

$$H = \begin{bmatrix} -2 & 12 \\ 12 & -6 \end{bmatrix}$$

We can check for positive or negative definiteness by looking at the sign of $|H_1|$ and $|H_2|$. Here $|H_1| = -2$ and $|H_2| = (-6)(-2) - (12)(12) = -136$. Since both determinants are negative, we conclude $(0,0)$ is neither a maximum nor minimum of f.

2 A quadratic function of n variables has the following standard form $$f(\boldsymbol{x}) = \boldsymbol{x}^T A \boldsymbol{x}/2 + \boldsymbol{b}^T \boldsymbol{x} + c$$, where $\boldsymbol{x}$ is the vector containing the n variables. Vector $\boldsymbol{b}(n \times 1)$ and symmetric matrix $A(n \times n)$ contain constant coefficients. For the following two quadratic functions $f_a(\boldsymbol{x}) = x_1^2 + 2x_1x_2 + 3x_2^2$ and $f_b(\boldsymbol{x}) = -x_1^2 - x_2^2 - x_3^2 + 2x_1x_2 + 6x_1x_3 + 4x_1 - 5x_3 + 7$, please

2.1 Rewrite these functions in the standard quadratic function form.

In the standard quadratic form, the first product $\boldsymbol{x}^T A \boldsymbol{x}$ produces the quadratic terms and the combinations $\boldsymbol{x}_i \boldsymbol{x}_j, i \neq j$. For f_a, on the diagonals of A we place the coefficients of the quadratic terms, in this case 2 and 6. Outside of the diagonals we place the half of the combinations $\boldsymbol{x}_1 \boldsymbol{x}_2$, or 2 in this case. The product $\boldsymbol{b}^T \boldsymbol{x}$ would take care of any single first order variables $x_i, i = 1,2$; however f_a does not have such factors. The final quadratic form for $f_a(\boldsymbol{x})$ is

$$f_a(\boldsymbol{x}) = \frac{1}{2} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + 0 \tag{2.1}$$

The process is similar for f_b. On the diagonals we place the coefficients of the second order components, and off the diagonals the coefficients of $x_i x_j, i \neq j$. Next, $\boldsymbol{b}$ will contain the first order coefficients, here $4, 0, -5$ for x_1, x_2, x_3, respectively. The standard quadratic formula for $f_b(\boldsymbol{x})$ is

$$f_b(\boldsymbol{x}) = \frac{1}{2}\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} -2 & 2 & 6 \\ 2 & -2 & 0 \\ 6 & 0 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 4 & 0 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + 7 \tag{2.2}$$

2.2 For these functions, show that $\nabla f = A\boldsymbol{x} + b$.

The gradient vector $\nabla f(\boldsymbol{x})$ is the vector containing the partial derivatives $\frac{\partial f}{\partial x_i}$.

$$\nabla f_a(\boldsymbol{x}) = \begin{bmatrix} 2x_1 + 2x_2 \\ 2x_1 + 6x_2 \end{bmatrix}$$

and the product $A\boldsymbol{x} + \boldsymbol{b}$ is

$$\begin{bmatrix} 2 & 2 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2x_1 + 2x_2 \\ 2x_1 + 6x_2 \end{bmatrix}$$

which shows that $\nabla f_a(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$.
For the second equation

$$\nabla f_b(\boldsymbol{x}) = \begin{bmatrix} -2x_1 + 2x_2 + 6x_3 + 4 \\ -2x_2 + 2x_1 \\ -2x_3 + 6x_1 - 5 \end{bmatrix}$$

and the product $A\boldsymbol{x} + \boldsymbol{b}$ is

$$\begin{bmatrix} -2 & 2 & 6 \\ 2 & -2 & 0 \\ 6 & 0 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 4 \\ 0 \\ -5 \end{bmatrix} = \begin{bmatrix} -2x_1 + 2x_2 + 6x_3 + 4 \\ -2x_2 + 2x_1 \\ -2x_3 + 6x_1 - 5 \end{bmatrix}$$

which again shows $\nabla f_b(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$.

2.3 For these functions, determine whether the matrix A is positive definite, negative definite, or indefinite.

The textbook mentions two methods to check whether the matrix is positive definite, negative definite, or semi-definite. The first method involves checking the signs of the eigenvalues of A. The second involves checking the sign of the determinants of the submatrices $A_1, A_2, \ldots, A_n$. For $f_a(\boldsymbol{x})$ this would be

$$|A_1| = |2| \quad \text{and} \quad |A_2| = \begin{vmatrix} 2 & 2 \\ 2 & 6 \end{vmatrix} = (2)(6) - (2)(2) = 8$$

Since $|A_1|$ and $|A_2|$ are positive, A is positive definite.
For $f_b(\boldsymbol{x})$ we repeat the same procedure, we check the signs of $|A_1|, |A_2|$ and $|A_3|$.

$$|A_1| = -2$$

$$|A_2| = (-2)(-2) - (2)(2) = 0$$

$$|A_3| = -2(4) - 2(-4) + 6(-12) = 72$$

Since there are both negative and positive determinants for the submatrices of A, the matrix of $f_b(\boldsymbol{x})$ is indefinite.

3 Minimize the following functions subject to equality and/or inequality constraints. For problems with inequality constraints, identify the active ones. To know whether a point is local minimum, perform function evalutation at 2 nearby points (which satisfy constraints) to show that what you found is local minimum.

3.1 $f(x,y) = 1/(xy)^2$, subject to $x^2 + y^2 = 1$

Using the Lagrange multiplier method, we first define our Lagrange function L

$$L(x,y,\lambda) = \frac{1}{(xy)^2} + \lambda(1 - x^2 - y^2)$$

and find the partial derivatives of x, y, and λ.

$$\frac{\partial L}{\partial x} = -2x^{-3}y^{-2} - 2\lambda x \tag{3.1}$$

$$\frac{\partial L}{\partial y} = -2x^{-2}y^{-3} - 2\lambda y \tag{3.2}$$

$$\frac{\partial L}{\partial \lambda} = 1 - x^2 - y^2 \tag{3.3}$$

When we solve for λ in Equation 3.1, we get $\lambda = -x^{-4}y^{-2}$. Substituting for λ in Equation 3.2, we get

$$-2x^{-2}y^{-3} - 2y(-x^{-4}y^{-2})$$

which in the end gives us $y = x$. Plugging this result in Equation 3.3, we get the result

$$x = y = \frac{1}{\sqrt{2}}$$

which gives us a value of 4 in f. To show $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ is a local minimum, we plot some nearby points. We chose the points $(0.69, 0.7238)$ and $(0.72, 0.694)$, which lie on either side of our proposed minimum, and evaluate

$$f(0.69, 0.7238) = 4.009$$

$$f(0.72, 0.694) = 4.005$$

which supports our claim that $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ is a minimum.

3.2 $f(x,y) = x^2 - 2x + y^2 - 10y$, subject to $-x^2 + y \leq 4$ and $-(x-2)^2 + y \leq 3$

We can rewrite the constraints and leave 0 on the right hand side

$$-x^2 + y - 4 \leq 0 \tag{3.4}$$

$$-(x-2)^2 + y - 3 \leq 0 \tag{3.5}$$

Since we are dealing with inequality constraints, we use the Kuhn-Tucker conditions. The Kuhn-Tucker conditions for satisfying a constrained problem when the active constraints are unknown are

$$\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial g_j}{\partial x_i} = 0 \tag{3.6}$$

$$\lambda_j > 0 \tag{3.7}$$

We first want to write the conditions for x and y as in Equation 3.6

$$2x - 2 - 2\lambda_1 x - 2\lambda_2 x + 4\lambda_2 \tag{3.8}$$

$$2y - 10 + \lambda_1 + \lambda_2 \tag{3.9}$$

We can solve for y in Equation 3.5 and get

$$y = 4 + x^2 \tag{3.10}$$

We then plug Equation 3.10 into Equation 3.5,

$$-(x-2)^2 + 4 + x^2 - 3 = 0$$

Expanding the $-(x-2)^2$ and solving for x, we get $x = \frac{4}{3}$. We can solve for λ_2 by plugging $4 + x^2$ in Equation 3.9.

$$10 - 2(4 + x^2) - \lambda_1 = \lambda_2$$

$$\lambda_2 = 2 - 2x^2 - \lambda_1 \tag{3.11}$$

We then use Equation 3.11 in Equation 3.8 to solve for λ_1

$$2x - 2 - 2x\lambda_1 - 2x(2 - x^2 - \lambda_1) + 4(2 - 2x^2 - \lambda_1)$$

Solving for λ_1, we have

$$\lambda_1 = \frac{1}{2}x^3 - 2x^2 - \frac{1}{2}x + 2 \tag{3.12}$$

And putting $\frac{4}{3}$ in for x, we get $\lambda_1 = \frac{91}{128}$. And also in Equation 3.11 we get $\lambda_2 = \frac{21}{128}$. Finally, for Equation 3.10, we get $y = \frac{91}{128}$. The solution $x = \frac{4}{3}, y = \frac{73}{16}, \lambda_1 = \frac{91}{128}, \lambda_2 = \frac{21}{128}$ meets the constraints. Since $\lambda_i \neq 0, i = 1,2$, all the constraints in this problem are active. To show that this point is a minimum, we again pick two close points and evaluate the function.

$f(\frac{5}{3}, \frac{73}{16})$	-25.36
$f(\frac{4}{3}, \frac{73}{16})$	-25.69
$f(\frac{4}{3}, \frac{72}{16})$	-25.63

3.3 $f(x,y) = (x-1)^2 + y - 2$, subject to $x - y - 1 = 0$ and $x + y - 2 \leq 0$

Here we have both equality and inequality constraints, so the Kuhn-Tucker conditions are

$$\nabla f + \sum_{j=1}^{m} \lambda_j \nabla g_j - \sum_{k=1}^{p} \beta_k \nabla h_k$$

with the conditions

$$\lambda_j g_j = 0, \quad g_j \leq 0, \quad h_k = 0 \quad \lambda_j \geq 0$$

Expanding the form above for x and y, we get

$$2(x-1)+\lambda+\beta \tag{3.13a}$$

$$1-\lambda-\beta \tag{3.13b}$$

subject to the constraints

$$\lambda(x+y-2)=0 \tag{3.14}$$

$$x-y-1=0 \tag{3.15}$$

For Equation 3.14, we have $x+y-2=0$ or $\lambda=0$.

3.3.1 Case 1: $x+y-2=0$

We choose to solve for x, where $x=2-y$. Substituting this result in Equation 3.15, we get

$$(2-y)-y-1=0$$

$$1-2y-0$$

$$y=\frac{1}{2}$$

Then Equation 3.14 gives us $x=\frac{3}{2}$. These points meet the constraints; however $f(\frac{3}{2},\frac{1}{2})=-\frac{5}{4}$ and $f(1,0)=-2$, which shows that the proposed point is not a minimum.

3.3.2 Case 2: $\lambda=0$

With Equation 3.15, we get $x=1+y$ which we substitute in Equation 3.13. This gives us

$$2y+\beta=0$$

$$\beta=-2y$$

which we substitute in Equation 3.13b

$$1-0-(-2y)=0$$

$$y=-\frac{1}{2}$$

Now, we substitute this into the equality constraint Equation 3.15

$$x-(-\frac{1}{2})-1=0$$

$$x=\frac{1}{2}$$

so $x=\frac{1}{2}$. To see if this point is a minimum, we compare two other points that match the constraints.

$f(1,0)$	-2
$f(0.75,-0.25)$	-2.1875
$f(\frac{1}{2},-\frac{1}{2})$	-2.25

$(\frac{1}{2},-\frac{1}{2})$ meets our requirements, and the inequality constraint is inactive since $\lambda=0$.