National Cheng Kung University

Optimization Design Homework 1

Andrés Ponce 彭思安 P76107116

March 7, 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

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

We can check for positive or negative definiteness by looking at the sign of $|\boldsymbol{H}_1|$ and $|\boldsymbol{H}_2|$. Here $|\boldsymbol{H}_1| = -2$ and $|\boldsymbol{H}_2| = (-6)(-2) - (12)(12) = -136$. 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 $\boldsymbol{A}(n \times n)$ 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 $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$ produces the quadratic terms and the combinations $\boldsymbol{x}_i\boldsymbol{x}_j, i \neq j$. For f_a, on the diagonals of $\boldsymbol{A}$ 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 sub-matrices $A_1, A_2, \ldots, A_n$. 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 $x^2+y^2=1$

We can use constrained variation to find the maximum values of x and y subject to the constraint. According to Equation 2.25 in the textbook, constrained variation takes the form

$$\left(\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}\right)\Big|_{(x^*,y^*)} = 0 \tag{3.1}$$

We first find the four partial derivatives required

$$\frac{\partial f}{\partial x} = -2x^{-3}y^{-2} \quad , \quad \frac{\partial f}{\partial y} = -2x^{-2}y^{-3}$$

$$\frac{\partial g}{\partial x} = 2x \quad , \quad \frac{\partial g}{\partial y} = 2y$$

and arrange the terms according to Equation 3.1.

$$(-2x^{-3}y^{-2})(2y) - (-2x^{-2}y^{-3})(2x) = 0$$

Solving for y first, we get

$$4x^{-1}y^{-3} = 4x^{-3}y^{-1}$$
$$x^{-1}y^{-3} = x^{-3}y^{-1}$$
$$x^2 = y^2$$
$$x^* = y^*$$

And now, we can plug this result into our constraint.

$$x^2 + (x)^2 = 1$$
$$2x^2 = 1$$
$$x = \pm\frac{1}{\sqrt{2}}$$

When we substitute for x, we also get $y = \pm\frac{1}{\sqrt{2}}$. Now, we have to show that these points indeed are minima by seeing to points closeby. The four combinations of points all evaluate to 4. Since $x = y = \pm\frac{1}{\sqrt{2}}$, we have four points to check. Below we compare points close to our suspected minima, one point on either side.

$(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$	$f(x,y)$
(0.607, 0.795)	4.294
(0.77, 0.638)	4.14

$(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$	$f(x,y)$
(-0.607, 0.795)	4.294
(-0.77, 0.638)	4.14

$(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$	$f(x,y)$
(0.674, -0.739)	4.03
(0.74, -0.673)	4.03

$(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$	$f(x,y)$
(-0.674, -0.739)	4.03
(-0.74, -0.673)	4.03

As we can see, due to the quadratic nature of $f(x,y)$, all our points are local minima when compared to points in their neighborhood.

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