



**BURSA TECHNICAL UNIVERSITY**  
**INSTITUTE OF NATURAL AND APPLIED SCIENCES**  
**DEPT. OF MECHANICAL ENGINEERING**

<b>STUDENT NAME:</b>	<b>BURAK ER</b>
<b>COURSE TITLE:</b>	<b>MECH 601 OPTIMIZATION IN ENGINEERING</b>
<b>ASSIGNMENT TITLE:</b>	
<b>ASSIGNMENT NUMBER:</b>	<b>3-4</b>

**ISSUE DATE** **5 DEC 2013**

**DEADLINE DATE** **19 DEC 2013**

**SCORE**

**LECTURER NAME:** Assoc. Prof. Dr. Ali Rıza YILDIZ

**SIGNED:**.....

**DATE:**.....

**Student declaration: (must be signed before submission)**

**I declare that this assignment is my own work**

Student signature: .....

Date:.....

**NOTE:** Please solve the following problems, giving sufficient explanations for the stages of solution. Solutions have to be presented as a ms Word file. If your assignments need Matlab solution, add your codes to assignment. Please submit your assignment into a CD.

Homework 3-4 Solution  
of  
Optimization in Engineering Class  
(MECH601)

MSc. Burak ER

December 29, 2013

### **Abstract**

Solved problems in this homework are some of the problems at the end of chapter 4, 10 and 11 of the book Introduction to Optimum Design by Jasbir Arora[1]. The solution includes the following problems; Problem 4.7, 4.9, 4.22, 4.32, 4.47, 4.60, 4.85, 4.101, 4.138, 4.150, 10.11, 10.16, 10.27, 10.53, 10.59, 10.68, 11.10, 11.12, 11.33.

**Problem 4.7**

Write down the Taylor Series representation of  $e^x$  including quadratic terms around  $x^* = 2$ .

**Solution**

Taylor series of any function around  $x_0$  is given by

$$f(x_0 + \Delta x) = x_0 + f'(x_0) \Delta x + \frac{f''(x)}{2!} (\Delta x)^2 + \dots \quad (1)$$

by putting  $\Delta x = x - x_0$  and neglecting the higher order terms, it is found that

$$\begin{aligned} f(x) &= 2 + e^2(x - 2) + \frac{e^2}{2!}(x - 2)^2 \\ &= 2 + e^2\left(\frac{x^2}{2} - x\right) \end{aligned}$$

Therefore, Taylor series of  $e^x$  around the point  $x^* = 2$  is found as

$$f(x) = e^2\left(\frac{x^2}{2} - x\right) + 2$$

**Problem 4.9**

Determine the nature of the following quadratic form.

$$F(\mathbf{x}) = x_1^2 + 4x_1x_2 + 2x_1x_3 - 7x_2^2 - 6x_2x_3 + 5x_3^2 \quad (2)$$

**Solution**

The quadratic form of the equation can be written as

$$F(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x}$$

$$\mathbf{x} = [x_1 \ x_2 \ x_3]^T$$

$$\mathbf{P} = \begin{bmatrix} 1 & 4 & 2 \\ 0 & -7 & -6 \\ 0 & 0 & 5 \end{bmatrix}$$

we can now determine the definition of the matrix  $\mathbf{P}$  by finding the eigen values of it.

$$\det(\mathbf{P} - \lambda \mathbf{I}) = \det \left( \begin{bmatrix} 1 - \lambda & 4 & 2 \\ 0 & -7 - \lambda & -6 \\ 0 & 0 & 5 - \lambda \end{bmatrix} \right) = 0$$

From this determinant eigen value equation is found as

$$(-1)^2 (1 - \lambda) (-7 - \lambda) (5 - \lambda) + (-1)^3 (4) (0) (0) + (-1)^4 (2) (0) (0) = 0$$

or

$$(1 - \lambda) (-7 - \lambda) (5 - \lambda) = 0$$

The eigen values of the  $\mathbf{P}$  are  $\lambda_1 = 1$ ,  $\lambda_2 = -7$ ,  $\lambda_3 = 5$ , thus, it is indefinite because of having at least one negative eigen value. Therefore, the function in equation 2 is not convex.

### Problem 4.22

Find the stationary points for the following function. Also, determine the local minimum, local maximum and inflection points.

$$f(\mathbf{x}) = 3x_1^2 + 2x_1x_2 + 2x_2^2 + 7 \quad (3)$$

### Solution

The function is written in quadratic form as

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x} + \mathbf{b}$$

$$\mathbf{x} = [x_1 \ x_2]^T$$

$$\mathbf{P} = \begin{bmatrix} 3 & 2 \\ 0 & 2 \end{bmatrix}$$

$$\mathbf{b} = [0 \ 7]^T$$

stationary points of the function is found by setting its variation to zero

$$\begin{aligned} \delta f(\mathbf{x}) &= 0 \\ &= 2\delta \mathbf{x}^T \mathbf{P} \mathbf{x} \end{aligned}$$

by setting  $\delta \mathbf{x}$  arbitrary it is found

$$\mathbf{P} \mathbf{x} = 0$$

Resulting two linear equations

$$\begin{aligned} 6x_1 + 2x_2 &= 0 \\ 4x_2 &= 0 \end{aligned}$$

$x_1$  and  $x_2$  is found

$$\begin{aligned}x_1 &= 0 \\x_2 &= 0\end{aligned}$$

From the definition  $\mathbf{H} = 2\mathbf{P}$ . The eigen values of  $\mathbf{H}$ ;

$$\det(\mathbf{H} - \lambda\mathbf{I}) = \det\left(\begin{bmatrix} 6 - \lambda & 4 \\ 0 & 4 - \lambda \end{bmatrix}\right) = 0$$

$$(6 - \lambda)(4 - \lambda) = 0 \implies \lambda_1 = 6, \lambda_2 = 4$$

Eigen values of Hessian matrix are both positive thus it is positive definite. Therefore, the point  $\mathbf{x} = (0, 0)$  is a global minimum.

### Problem 4.32

Annual operating cost  $U$  for an electrical line system is given by the following expression

$$U = \frac{21.9 \times 10^7}{V^2 C} + 3.9 \times 10^6 C + 1000V \quad (4)$$

$V$  : line voltage ,  $C$  : line capacitance

Find the stationary points for the function and determine  $V$  and  $C$  to minimize the operating cost.

### Solution

At stationary points, variation of function is equal to zero, thus,

$$\delta U = \frac{\partial U}{\partial \mathbf{x}} \delta \mathbf{x} = 0$$

$$\delta U = \left[ \frac{43.8 \times 10^7}{V^3 C} + 1000 - \frac{21.9 \times 10^7}{V^2 C^2} + 3.9 \times 10^6 \right] \begin{bmatrix} \delta V \\ \delta C \end{bmatrix}$$

the stationary points should satisfy the equations

$$\begin{aligned}\frac{43.8 \times 10^7}{V^3 C} + 1000 &= 0 \\ -\frac{21.9 \times 10^7}{V^2 C^2} + 3.9 \times 10^6 &= 0\end{aligned}$$

Using Newton-Raphson method for the solution of this two equations, stationary point of the function is found at

$$\begin{aligned} V &= 241.764\text{kV} \\ C &= 0.031\text{mhos} \end{aligned}$$

This value is just a stationary point. Determining whether it is a min or max requires Hessian matrix to be evaluated. Hessian of the function is

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 U}{\partial V^2} & \frac{\partial^2 U}{\partial V \partial C} \\ \frac{\partial^2 U}{\partial V \partial C} & \frac{\partial^2 U}{\partial C^2} \end{bmatrix}$$

Evaluating at stationary point  $\mathbf{H}$  is found as

$$\mathbf{H} = \begin{bmatrix} 12.408 & 32262.9502 \\ 32262.9502 & 2.516 \times 10^8 \end{bmatrix}$$

Eigen values of the Hessian matrix are found as  $\lambda_1 > 0$  &  $\lambda_2 > 0$ , thus, it is positive definite which makes the stationary point minimum.

#### Problem 4.47

Minimize:

$$f(\mathbf{x}) = 4x_1^2 + 9x_2^2 + 6x_2 - 4x_1 + 13$$

Subject to:

$$x_1 - 3x_2 + 3 = 0$$

#### Solution

The function to be minimized includes only equality constraint. It can be solved analytically by using Lagrange Multipliers Method. Minimization of the problem function with constraints is equivalent to minimization of Lagrangian. Lagrangian of a problem is given as

$$L(\mathbf{x}) = f(\mathbf{x}) + \mathbf{v}^T \mathbf{h}(\mathbf{x})$$

Lagrangian of the problem is found as

$$L(\mathbf{x}) = 4x_1^2 + 9x_2^2 + 6x_2 - 4x_1 + 13 + v_1(x_1 - 3x_2 + 3)$$

Stationary point of the Lagrangian is found as

$$\delta L(\mathbf{x}) = (8x_1 - 4)\delta x_1 + (18x_2 + 6 - 3v_1)\delta x_2 + v_1(x_1 - 3x_2 + 3)\delta v_1 = 0$$

For arbitrary variations, this results to three equations with three unknowns as

$$\begin{aligned} 8x_1 - 4 + v_1 &= 0 \\ 18x_2 + 6 - 3v_1 &= 0 \\ x_1 - 3x_2 + 3 &= 0 \end{aligned}$$

Solution of this equations is  $x_1 = 0.4$ ,  $x_2 = 1.133$  and  $v_1 = 8.798$ .

### Problem 4.60

Minimize:

$$f(\mathbf{x}) = (x-4)^2 + (y-6)^2$$

Subject to:

$$\begin{aligned} 12 &\geq x+y \\ x &\geq 6 \\ y &\geq 0 \end{aligned}$$

### Solution

The problem can be solved by Lagrange Multipliers method. It includes three inequality constraints. Lagrangian of a problem with both equality and inequality constraints is given by

$$L = f(\mathbf{x}) + \mathbf{v}^T \mathbf{h} + \mathbf{u}^T (\mathbf{g} + \mathbf{s}^2)$$

$$L = (x-4)^2 + (y-6)^2 + 0 + u_1(x+y-12+s_1^2) + u_2(6-x+s_2^2) + u_3(-y+s_3^2)$$

Lagrangian should be stationary for an extremum value

$$\delta L = 0$$

$$\begin{aligned} \delta L = & (2(x-4) + u_1 - u_2) \delta x + (2(y-6) + u_1 - u_3) \delta y + \\ & (x+y-12+s_1^2) \delta u_1 + (6-x+s_2^2) \delta u_2 + \\ & (-y+s_3^2) \delta u_3 + (x+y-12+s_1^2) \delta s_1 + \\ & (x+y-12+s_1^2) \delta s_2 + (x+y-12+s_1^2) \delta s_3 \end{aligned}$$

This condition results eight independent nonlinear equations with eight unknowns as

$$\begin{aligned} 2(x-4) + u_1 - u_2 &= 0 \\ 2(y-6) + u_1 - u_3 &= 0 \\ x+y-12+s_1^2 &= 0 \\ 6-x+s_2^2 &= 0 \\ -y+s_3^2 &= 0 \\ 2s_1u_1 &= 0 \\ 2s_2u_2 &= 0 \\ 2s_3u_3 &= 0 \end{aligned}$$

solving these nonlinear equations results the following values,

$$\begin{aligned} x = 6, y = 6, u_1 = 4.3e^{-19}, u_2 = 4, u_3 = 4.028e^{-19}, s_1 = 5.63e^{-19}, \\ s_2 = 4.49e^{-19}, s_3 = 2.4495e^{-19} \end{aligned}$$



### Problem 4.85

Formulate and solve the following problem graphically. Verify the KKT conditions at the solution point and show the gradients of the cost function and active constraints on the graph.

Minimize:

$$f(x, y) = -\pi x^2 y$$

Subject to:

$$\begin{aligned} x - 20 &\leq 0 \\ y - 20 &\leq 0 \\ 5 - x &\leq 0 \\ -x &\leq 0 \\ 2\pi xy - 900 &\leq 0 \end{aligned}$$

### Solution

Graphical solution is given in fig. 1.

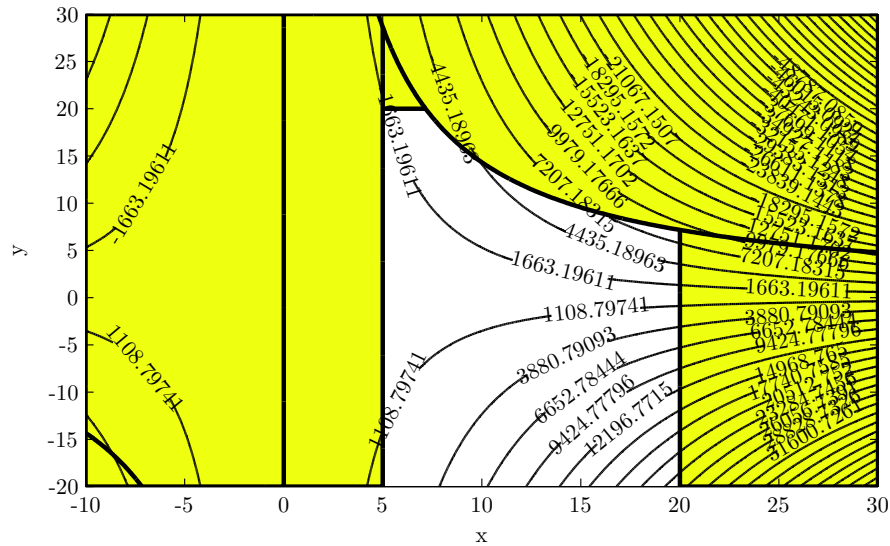


Figure 1: Plot of the cost function and constraints of problem 4.85.

From fig. 1, it is seen that increasing value of the both  $x$  and  $y$  decreases the value of the cost function. Within the feasible region, the minimum value of the function is found to be 9000 at  $(x^*, y^*) = (20, 7.162)$ .

For the verification of KKT conditions of the problem we should write Lagrangian which is given by the following equation

$$L = f(\mathbf{x}) + \mathbf{v}^T \mathbf{h} + \mathbf{u}^T (\mathbf{g} + \mathbf{s}^2)$$

Thus, Lagrangian of the problem is

$$\begin{aligned} L(x, y) = & -\pi x^2 y + u_1 (x - 20 + s_1^2) + u_2 (y - 20 + s_2^2) + \\ & u_3 (5 - x + s_3^2) + u_4 (-x + s_4^2) + u_5 (2\pi xy - 900 + s_5^2) \end{aligned}$$

The stationary point conditions are found as

$$\begin{aligned} \frac{\partial L}{\partial x} &= 0 ; -2\pi xy + u_1 - u_3 - u_4 + 2\pi u_5 y = 0 \\ \frac{\partial L}{\partial y} &= 0 ; -\pi x^2 y + 2\pi u_5 x = 0 \\ \frac{\partial L}{\partial u_1} &= 0 ; x - 20 + s_1^2 = 0 \\ \frac{\partial L}{\partial u_2} &= 0 ; y - 20 + s_2^2 = 0 \\ \frac{\partial L}{\partial u_3} &= 0 ; 5 - x + s_3^2 = 0 \\ \frac{\partial L}{\partial u_4} &= 0 ; -x + s_4^2 = 0 \\ \frac{\partial L}{\partial u_5} &= 0 ; -2\pi xy - 900 + s_5^2 = 0 \\ \frac{\partial L}{\partial s_1} &= 0 ; 2u_1 s_1 = 0 \\ \frac{\partial L}{\partial s_2} &= 0 ; 2u_2 s_2 = 0 \\ \frac{\partial L}{\partial s_3} &= 0 ; 2u_3 s_3 = 0 \\ \frac{\partial L}{\partial s_4} &= 0 ; 2u_4 s_4 = 0 \\ \frac{\partial L}{\partial s_5} &= 0 ; 2u_5 s_5 = 0 \end{aligned}$$

Evaluating the conditions at the graphically found solution point  $(x^*, y^*) = (20, 7.162)$ , the unknowns of the problem are found as

$$\begin{aligned} x = 20, y = 7.16, s_1 = 0, u_2 = 0, u_3 = 0, u_4 = 0, s_5 = 0, \\ u_5 = 10, u_1 = 449.876, s_2 = 3.583, s_3 = 3.873, s_4 = 4.472 \end{aligned}$$

Slack variables of constraints of  $g_1$  and  $g_5$  is found to be zero at the solution point, thus, they are both active.

### Problem 4.101

Solve the following problem graphically, verify the KKT necessary conditions for the solution points and study the effect on the cost function of changing the boundary of the equality constraint by one unit.

Minimize:

$$f(\mathbf{x}) = 9x_1^2 + 18x_1x_2 + 13x_2^2 - 4$$

Subject to:

$$x_1^2 + x_2^2 + 2x_1 = 16$$

### Solution

Plots of the cost function and equality constraints of the problem is given in fig. 2. From the graphical solution of this problem, optimum point is found as

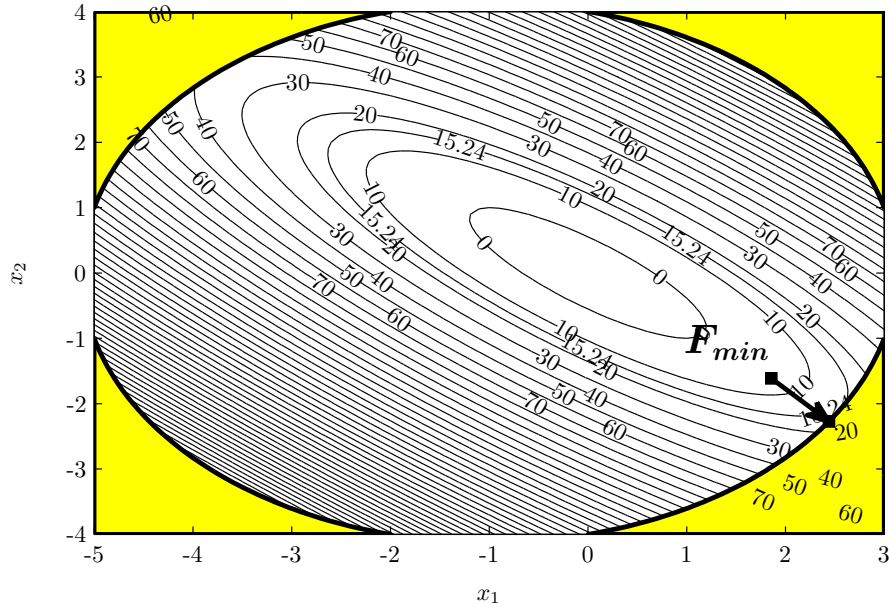


Figure 2: Plot of the cost function and constraint of the problem 4.101.

to be  $x_1^* = -2$ ,  $x_2^* = 2.6$  and  $f^* = 15.24$ .

The verification of the KKT conditions can be found by defining the Lagrangian as

$$\begin{aligned} L(\mathbf{x}) &= f(\mathbf{x}) + \mathbf{u}^T \mathbf{h} \\ &= 9x_1^2 + 18x_1x_2 + 13x_2^2 - 4 + u_1 (x_1^2 + x_2^2 + 2x_1 - 16) \end{aligned}$$

The stationary condition for the Lagrangian

$$\delta L(\mathbf{x}) = 0$$

That result is three equations with three unknowns as

$$\begin{aligned} 18x_1 + 18x_2 + 2u_1x_1 &= 0 \\ 18x_1 + 26x_2 + 2u_2x_2 + 2 &= 0 \\ x_1^2 + x_2^2 + 2x_1 - 16 &= 0 \end{aligned}$$

Solution to these equations is  $x_1^* = 2.5491$ ,  $x_2^* = -2.0985$ ,  $\lambda^* = 1.5909$ . As it can be seen the graphical solution and the Lagrange Multipliers Method solution is nearly same.

By changing the value of the constraint by one unit we find

$$\begin{aligned} f(\mathbf{h} + \mathbf{b}) &= -\mathbf{b}^T \mathbf{u} \\ &= 1.15909 \end{aligned}$$

which means changing the constraint by one unit will increase the cost function.

### Problem 4.138

Check for the convexity of the problem. If the function is not convex everywhere, then, determine the domain(feasible set  $S$ ) over which the function is convex.

$$f(\mathbf{x}) = 5x_1 - \frac{1}{16}x_1^2x_2^2 + 2x_2^2 + 5x_1^2 + 3x_2$$

### Solution

Convexity of a function can be found from the definition of its Hessian Matrix. If its Hessian matrix is positive or semidefinite then it is convex else it is not. Hessian is defined as

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 U}{\partial V^2} & \frac{\partial^2 U}{\partial V \partial C} \\ \frac{\partial^2 U}{\partial V \partial C} & \frac{\partial^2 U}{\partial C^2} \end{bmatrix}$$

where for this problem it is found as

$$\mathbf{H} = \begin{bmatrix} -\frac{1}{8}x_2^2 + 10 & \frac{1}{4}x_1x_2 \\ -\frac{1}{4}x_1x_2 & -\frac{1}{8}x_1^2 + 4 \end{bmatrix}$$

Eigen values of the Hessian can be found as

$$\det(\mathbf{H} - \lambda \mathbf{I}) = \det \left( \begin{bmatrix} -\frac{1}{8}x_2^2 + 10 - \lambda & \frac{1}{4}x_1x_2 \\ -\frac{1}{4}x_1x_2 & -\frac{1}{8}x_1^2 + 4 - \lambda \end{bmatrix} \right) = 0$$

$$\lambda^2 + \left( \frac{x_1^2 + x_2^2}{8} - 14 \right) \lambda - \frac{3x_1^2 x_2^2}{64} - \frac{5x_1^2}{4} - \frac{x_2^2}{2} + 40 = 0 \quad (5)$$

For a convex behavior,  $\lambda$  roots of the ?? should be positive. If we find the roots we get

$$\begin{aligned} \lambda_1 &= \frac{\sqrt{x_1^4 + 14x_1^2 x_2^2 + 96x_1^2 + x_2^4 - 96x_2^2 + 2304}}{16} - \frac{x_2^2}{16} - \frac{x_1^2}{16} + 7 \\ \lambda_2 &= 7 - \frac{x_2^2}{16} - \frac{\sqrt{x_1^4 + 14x_1^2 x_2^2 + 96x_1^2 + x_2^4 - 96x_2^2 + 2304}}{16} - \frac{x_1^2}{16} \end{aligned}$$

If these roots are plot for the determination of their positive value range fig. 3 is found. It is seen from the figure that first of the eigen value is positive for all design variables while second eigen value is negative for all design variables. Therefore, there is no feasible set  $S$  in that the problem is convex.

### Problem 4.150

The problem of minimum weight deisgn of the symmetric three bar truss system is formulated as follows:

Minimize:

$$f(\mathbf{x}) = 2\sqrt{2}x_1 + x_2$$

Subject to:

$$\begin{aligned} g_1 &= \frac{1}{2} \left[ \frac{P_u}{x_1} + \frac{P_v}{(x_1 + \sqrt{2}x_2)} \right] - 20000 \leq 0 \\ g_2 &= \frac{\sqrt{2}P_v}{(x_1 + \sqrt{2}x_2)} - 20000 \leq 0 \\ g_3 &= -x_1 \leq 0 \\ g_4 &= -x_2 \leq 0 \end{aligned}$$

where  $x_1$  is the cross sectional area of members1 and 3(symmetric structure) and  $x_2$  is the cross-sectional area of member 2,  $P_u = P \cos \theta, P_v = P \sin \theta$ , with  $P > 0$  and  $0 \leq \theta \leq 90$ . Check for the convexity of the problem for  $\theta = 60$ .

### Solution

The cost function of the problem is convex because it is linear in terms of design variables. The constraints  $g_3$  and  $g_4$  are also convex. The behavior of constraints  $g_1$  and  $g_2$  an be determined from their hessian matrix. Hessian matrix of the this two constraints are found as

$$\nabla^2 g_1 = \begin{bmatrix} \sqrt{2} \left( \frac{P \sin \theta}{(x_1 + \sqrt{2}x_2)^3} + \frac{P \cos \theta}{x_1^3} \right) & \frac{2P \sin \theta}{(x_1 + \sqrt{2}x_2)^3} \\ \frac{2P \sin \theta}{(x_1 + \sqrt{2}x_2)^3} & \frac{2\sqrt{2}P \sin \theta}{(x_1 + \sqrt{2}x_2)^3} \end{bmatrix}$$

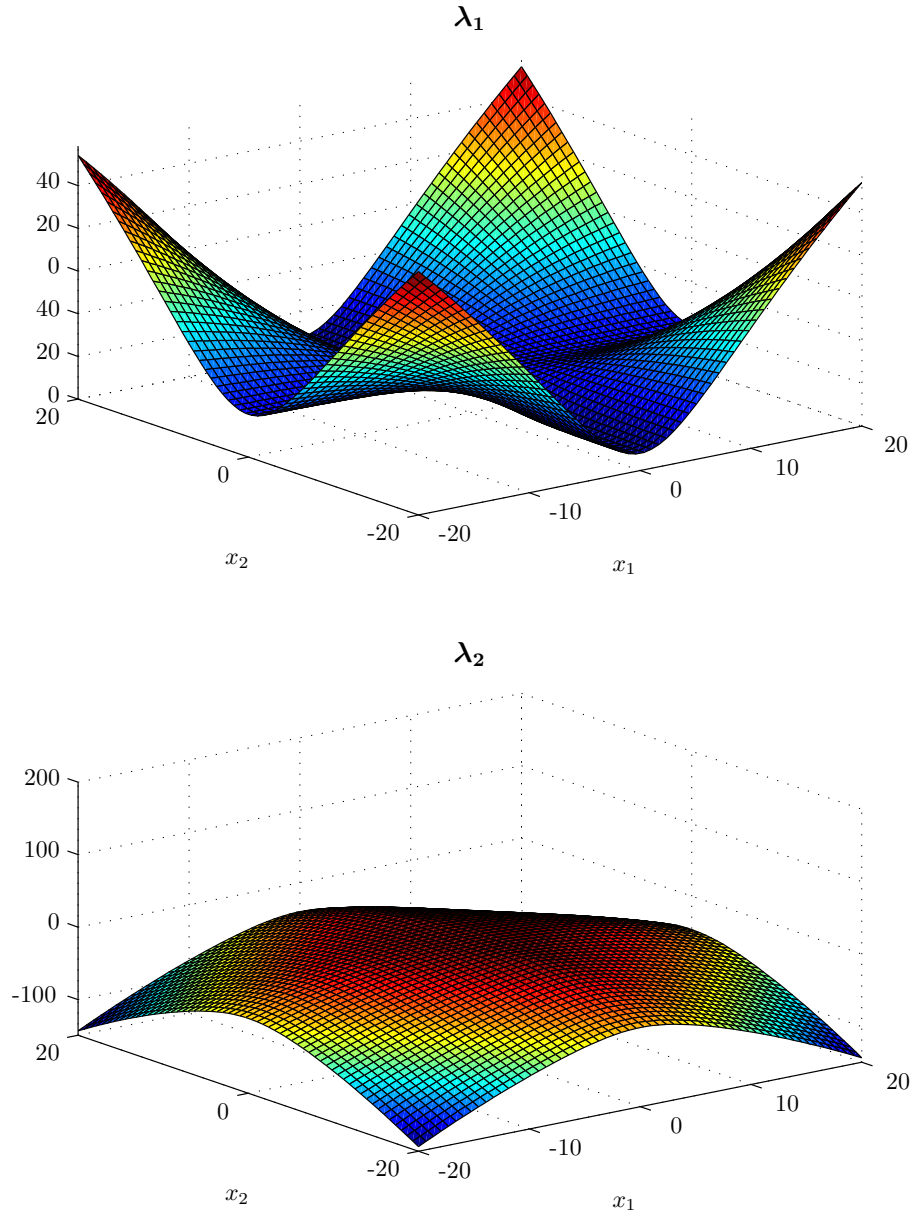


Figure 3: Plot of the eigen values of the Hessian of the problem 4.138 with respect to design variables.

$$\nabla^2 g_2 = \begin{bmatrix} \frac{2\sqrt{2}P \sin \theta}{(x_1 + \sqrt{2}x_2)^3} & \frac{4P \sin \theta}{(x_1 + \sqrt{2}x_2)^3} \\ \frac{4P \sin \theta}{(x_1 + \sqrt{2}x_2)^3} & \frac{4\sqrt{2}P \sin \theta}{(x_1 + \sqrt{2}x_2)^3} \end{bmatrix}$$

The Hessian matrices of both the constraints are positive definite for  $\theta = \frac{\pi}{3}$  and  $x_1 \geq 0, x_2 \geq 0$ . Therefore, the problem is strictly convex.

### Problem 10.11

Determine whether the given direction is descent for the given function.

$$f(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2; \quad \mathbf{d} = (2, 4, -2) \quad \text{at} \quad \mathbf{x} = (1, 2, -1)$$

#### Solution

Direction's projection on the gradient's direction should be find in order to determine whether it is a descent direction. Gradient vector is

$$\begin{aligned} \nabla f &= \begin{bmatrix} 2x_1 & 2x_2 & 2x_3 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 4 & -2 \end{bmatrix} \end{aligned}$$

projection of direction on the gradient vector is

$$\begin{aligned} \mathbf{d} &= \begin{bmatrix} 2 & 4 & -2 \end{bmatrix} \\ \mathbf{d} \cdot \nabla f &= \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 \end{bmatrix} = 20 \\ \mathbf{d} \cdot \nabla f &> 0 \end{aligned}$$

Projection of direction on the gradient of the function is found to be positive, therefore, the direction is not descent.

### Problem 10.16

Find the minimum of the function  $f(\alpha) = 7\alpha^2 - 20\alpha + 22$  using equal-interval-search method within accuracy of 0.001. Use  $\delta = 0.05$ .

#### Solution

In equal interval equal-interval-search design variable range is determined by  $x_l$  and  $x_u$  then this range is divided to equal intervals. Function is evaluated at the end of the intervals. A change in function behavior at the interval  $k$  is a sign for a root.

For the start of iteration initial values are selected to be as  $\alpha_l = 1.3$  and  $\alpha_u = 1.5$ . The iterations for the solutions are

$k=0$

$$\alpha_l^{(0)} = 1.3, \alpha_u^{(0)} = 1.5, \delta^{(0)} = 0.5$$

$$\alpha_l^{(0)} = 1.3 \implies f_l^{(0)} = 7.83$$

$$\alpha_1^{(0)} = 1.3 + 0.05 \implies f_1^{(0)} = 7.7575$$

$$\alpha_2^{(0)} = 1.3 + 0.1 \implies f_2^{(0)} = 7.72$$

$$\alpha_3^{(0)} = 1.3 + 0.15 \implies f_3^{(0)} = 7.7175$$

$$\alpha_u^{(0)} = 1.3 + 0.2 \implies f_u^{(0)} = 7.75$$

$k=1$

$$\alpha_l^{(1)} = 1.4, \alpha_u^{(1)} = 1.5, \delta^{(1)} = 0.025$$

$$\alpha_l^{(1)} = 1.4 \implies f_l^{(1)} = 7.72$$

$$\alpha_1^{(1)} = 1.4 + 0.025 \implies f_1^{(1)} = 7.714375$$

$$\alpha_2^{(1)} = 1.4 + 0.05 \implies f_2^{(1)} = 7.7175$$

$$\alpha_3^{(1)} = 1.4 + 0.075 \implies f_3^{(1)} = 7.729375$$

$$\alpha_u^{(1)} = 1.5 \implies f_u^{(1)} = 7.75$$

$k=2$

$$\alpha_l^{(2)} = 1.4, \alpha_u^{(2)} = 1.45, \delta^{(2)} = 0.01$$

$$\alpha_l^{(2)} = 1.4 \implies f_l^{(2)} = 7.72$$

$$\alpha_1^{(2)} = 1.4 + 0.01 \implies f_1^{(2)} = 7.7167$$

$$\alpha_2^{(2)} = 1.4 + 0.02 \implies f_2^{(2)} = 7.7148$$

$$\alpha_3^{(2)} = 1.4 + 0.03 \implies f_3^{(2)} = 7.7143$$

$$\alpha_4^{(2)} = 1.4 + 0.04 \implies f_4^{(2)} = 7.7152$$

$$\alpha_u^{(2)} = 1.45 \implies f_u^{(2)} = 7.7175$$

$k=3$

$$\alpha_l^{(3)} = 1.42, \alpha_u^{(3)} = 1.44, \delta^{(3)} = 0.005$$

$$\alpha_l^{(3)} = 1.42 \implies f_l^{(3)} = 7.7148$$

$$\alpha_1^{(3)} = 1.42 + 0.005 \implies f_1^{(3)} = 7.714375$$

$$\alpha_2^{(3)} = 1.42 + 0.01 \implies f_2^{(3)} = 7.7143$$

$$\alpha_3^{(3)} = 1.42 + 0.015 \implies f_3^{(3)} = 7.714575$$

$$\alpha_u^{(3)} = 1.44 \implies f_u^{(3)} = 7.7152$$

$k=4$

$$\alpha_l^{(4)} = 1.425, \alpha_u^{(4)} = 1.435, \delta^{(4)} = 0.0025$$

$$\alpha_l^{(4)} = 1.425 \implies f_l^{(4)} = 7.714375$$



$$\begin{aligned}
\alpha_1^{(4)} &= 1.425 + 0.0025 \implies f_1^{(4)} = 7.71429375 \\
\alpha_2^{(4)} &= 1.425 + 0.005 \implies f_2^{(4)} = 7.7143 \\
\alpha_3^{(4)} &= 1.425 + 0.0075 \implies f_3^{(4)} = 7.71439375 \\
\alpha_u^{(4)} &= 1.435 \implies f_u^{(4)} = 7.7152
\end{aligned}$$

$k = 5$

$$\begin{aligned}
\alpha_l^{(5)} &= 1.425, \alpha_u^{(5)} = 1.43, \delta^{(5)} = 0.001 \\
\alpha_l^{(5)} &= 1.425 \implies f_l^{(5)} = 7.714375 \\
\alpha_1^{(5)} &= 1.425 + 0.001 \implies f_1^{(5)} = 7.714332 \\
\alpha_2^{(5)} &= 1.425 + 0.002 \implies f_2^{(5)} = 7.714303 \\
\alpha_3^{(5)} &= 1.425 + 0.003 \implies f_3^{(5)} = 7.714288 \\
\alpha_4^{(5)} &= 1.425 + 0.004 \implies f_4^{(5)} = 7.734287 \\
\alpha_u^{(5)} &= 1.43 \implies f_u^{(5)} = 7.7143
\end{aligned}$$

solution is found as  $\alpha^* = 1.429$ ,  $f^* = 7.734287$ .

### Problem 10.27

For the following function, direction of change is given. Derive the function of one variable that can be used to determine optimum step size.

$$f(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2; \quad \mathbf{d} = (-2, -4, 2) \quad \text{at} \quad \mathbf{x} = (1, 2, -1)$$

### Solution

Changing the direction of function in the direction  $\mathbf{d}$  by  $\alpha$  times at  $\mathbf{x}_0$  is given by

$$\mathbf{x} = \mathbf{x}_0 + \alpha \mathbf{d}$$

Using problem values, the function for the step size determination is found as

$$\begin{aligned}
f(\mathbf{x}_0 + \alpha \mathbf{d}) &= (1 - 2\alpha)^2 + (2 - 4\alpha)^2 + (-1 + 2\alpha)^2 \\
&= 4\alpha^2 - 2\alpha + 1 + 16\alpha^2 - 16\alpha + 4 + 4\alpha^2 - 2\alpha + 1 \\
&= 24\alpha^2 - 20\alpha + 6
\end{aligned}$$

which should be minimized.

### Problem 10.53

For the following problem complete two iterations of the steepest descent method.

$$f(\mathbf{x}) = 12.096x_1^2 + 21.504x_2^2 - 1.7321x_1 - x_2; \quad \mathbf{x}_0 = (1, 1)$$

#### Solution

Steepest descent algorithm is given by the following steps

- Step 1 : estimate initial value for the design variables and set  $\mathbf{x}_0$  .
- Step 2 : calculate gradient of function  $f(\mathbf{x})$  as  $\mathbf{c}^{(k)} = \nabla f(\mathbf{x}^{(k)})$ .
- Step 3 : calculate length of gradient. If  $|\mathbf{c}| < \epsilon$  then break the iteration and set  $\mathbf{x}^* = \mathbf{x}^{(k)}$ .
- Step 4 : let the search direction to be  $\mathbf{d}^{(k)} = -\mathbf{c}^{(k)}$ .
- Step 5 : calculate step size  $\alpha$  that minimizes  $f(\mathbf{x}^{(k)} + \alpha\mathbf{d}^{(k)})$ .
- Step 6 : update  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{d}^{(k)}$  .
- Step 7 : go to Step 2.

The iterations for the solution of problem with Steepest Descent Method are given below

For k=0

- (1)  $\mathbf{x}^{(0)} = (1, 1)$  ,  $k = 0$  ,  $\epsilon = 10^{-3}$
- (2)  $\mathbf{c}^{(0)} = [22.4599, 42.008]$
- (3)  $|\mathbf{c}^{(0)}| = 47.6352 > \epsilon$
- (4)  $\mathbf{d} = [1.7321 - 24.192x_1, 1 - 43.008x_2]$
- (5) 
$$f(\mathbf{x}^{(0)} + \alpha^{(0)}\mathbf{d}^{(0)}) = \underbrace{(12.096d_1^2 + 21.504d_2^2)}_A \alpha^2 + \underbrace{(24.192d_1x_1 + 43.008d_2x_2 - 1.7321d_1 - d_2)}_B \alpha$$

$$\underbrace{12.096x_1^2 + 21.504x_2^2 - 1.7321x_1 - x_2}_C$$

$$\delta f = 0; 2A\alpha + B = 0;$$

$$\implies \alpha = \frac{-B}{2A}$$

$$\alpha = 0.0257$$

$$(6) \quad \mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{d}^{(0)} = [0.4215, -0.0819]$$

For k=1

$$(2) \quad \mathbf{c}^{(1)} = [8.4650, -4.5259]$$

$$(3) \quad |\mathbf{c}^{(1)}| = 9.599 > \epsilon$$

$$(4) \quad \mathbf{d}^{(1)} = [-8.4650, 4.5259]$$

$$(5) \quad \delta f(\mathbf{x}^{(1)} + \alpha^{(1)} \mathbf{d}^{(1)}) = 0$$

$$\alpha = 0.0230$$

$$(6) \quad \mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = [0.2265, 0.0222]$$

### Problem 10.59

For the following problem complete two iterations of the Steepest Descent Method.

$$f(\mathbf{x}) = 9x_1^2 + 9x_2^2 - 100\sqrt{x_1^2 + x_2^2 - 20x_2 + 100} + 80\sqrt{x_1^2 + x_2^2 + 20x_2 + 100} - 5x_1 - 41x_2, \quad \mathbf{x}_0 = (5, 2)$$

### Solution

Steepest Descent Algorithm is given in problem 10.53. Therefore, it is not given here. For the solution, gradient of the cost function is needed which is found as

$$\mathbf{c} = \nabla f(\mathbf{x}) = \begin{bmatrix} 18x_1 - \frac{100x_1}{\sqrt{x_1^2 + x_2^2 - 20x_2 + 100}} + \frac{80x_1}{\sqrt{x_1^2 + x_2^2 + 20x_2 + 100}} \\ 18 - x_2 - \frac{100x_2 - 10}{\sqrt{x_1^2 + x_2^2 - 20x_2 + 100}} + \frac{80x_2 + 10}{\sqrt{x_1^2 + x_2^2 + 20x_2 + 100}} - 41 \end{bmatrix}^T$$

Iterations done for the solution is given below,

For k=0

$$(1) \quad \mathbf{x}^{(0)} = (5, 2), \quad k = 0, \quad \epsilon = 10^{-3}$$

$$(2) \quad \mathbf{c}^{(0)} = [62.7693, 153.6460]$$

$$(3) \quad |\mathbf{c}^{(0)}| = 165.9731 > \epsilon$$

$$(4) \quad \mathbf{d} = [-62.7693, -153.6460]$$

$$(5) \quad \delta f(\mathbf{x}^{(0)} + \alpha^{(0)} \mathbf{d}^{(0)}) = 0$$

$$\alpha = 0.0669$$

$$(6) \quad \mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{d}^{(0)} = [0.802, -8.279]$$

For k=1

$$(2) \quad \mathbf{c}^{(1)} = [38.8, -17.5]$$

$$(3) \quad |\mathbf{c}^{(1)}| = 42.6 > \epsilon$$

$$(4) \quad \mathbf{d}^{(1)} = [-38.8, 17.5]$$

$$(5) \quad \delta f(\mathbf{x}^{(1)} + \alpha^{(1)} \mathbf{d}^{(1)}) = 0$$

$$\alpha = 0.0191$$

$$(6) \quad \mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = [0.063, -7.9]$$

### Problem 10.68

For the following problem, complete two iterations of the Conjugate Gradient Method.

$$f(\mathbf{x}) = 9x_1^2 + 9x_2^2 - 100\sqrt{x_1^2 + x_2^2 - 20x_2 + 100} + 80\sqrt{x_1^2 + x_2^2 + 20x_2 + 100} - 5x_1 - 41x_2, \quad \mathbf{x}_0 = (5, 2)$$

### Solution

Conjugate gradient method differs from steepest descent on the way of choosing the search direction. The search direction is chosen as a combination of current gradient and the previous search direction.

It's algorithm is

Step 1 : estimate initial value for the design variables and set  $\mathbf{x}_0$  .

Step 2 : compute gradient of function  $f(\mathbf{x})$  at current point as  $\mathbf{c}^{(k)} = \nabla f(\mathbf{x}^{(k)})$ .

Step 3 : calculate  $|\mathbf{c}^{(k)}|$  , if  $|\mathbf{c}^{(k)}| < \epsilon$  then break the iteration and set  $\mathbf{x}^* = \mathbf{x}^{(k)}$ .

Step 4 : calculate conjugate search direction as

$$\mathbf{d}^{(k)} = -\mathbf{c}^{(k)} + \beta^{(k)}\mathbf{d}^{(k-1)}$$

$$\beta^{(k)} = \left( \frac{|\mathbf{c}^{(k)}|}{|\mathbf{c}^{(k-1)}|} \right)$$

Step 5 : compute step size  $\alpha^{(k)}$  that minimizes  $f(\mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{d}^{(k)})$ .

Step 6 : update  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{d}^{(k)}$  .

Step 7 : go to Step 2.

Solution steps of the problem is given below. Before going into the solution its useful to compute gradient of  $f$  with respect to design variables.

$$\nabla f = [24.192x_1 - 1.7321, 43.008x_2 - 1]$$

For k=0

$$(1) \quad \mathbf{x}^{(0)} = (1, 1) , k = 0 , \epsilon = 10^{-3}$$

$$(2) \quad \mathbf{c}^{(0)} = [22.4599, 42.008]$$

$$(3) \quad |\mathbf{c}^{(0)}| = 47.6352 > \epsilon$$

$$(4) \quad \mathbf{d} = [-22.4599, -42.008]$$

$$(5) \quad \delta f(\mathbf{x}^{(0)} + \alpha^{(0)}\mathbf{d}^{(0)}) = 0$$

$$\alpha = 0.0257$$

$$(6) \quad \mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{d}^{(0)} = [0.4215, -0.0819]$$

For k=1

$$(2) \quad \mathbf{c}^{(1)} = [8.465, -4.5259]$$

$$(3) \quad |\mathbf{c}^{(1)}| = 9.559 > \epsilon$$

$$(4) \quad \mathbf{d}^{(1)} = -\mathbf{c}^{(1)} + \beta^{(1)}\mathbf{d}^{(0)}$$

$$\begin{aligned} \mathbf{d}^{(1)} &= -[8.465, -4.5259] + 0.2015[-22.4599, 42.008] \\ &= [-12.9909, 12.9909] \end{aligned}$$

$$(5) \quad \delta f(\mathbf{x}^{(1)} + \alpha^{(1)}\mathbf{d}^{(1)}) = 0$$

$$\alpha = 0.01488$$

$$(6) \quad \mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \mathbf{d}^{(1)} = [0.2294, 0.1133]$$

### Problem 11.10

For the following problem, complete one iteration of modified Newton's method, also check descent condition for the search direction.

$$f(\mathbf{x}) = x_1^2 + 2x_2^2 - 4x_1 - 2x_1x_2 \quad ; \quad \mathbf{x}_0 = (1, 1)$$

### Solution

In modified Newton Algorithm, function to be minimized is expanded to Taylor series that is quadratic in terms of the design variables. Then, the search direction is chosen as the minimizing value of the design variables in the series representation. Because of Hessian is also considered within this method, the converge ratio is quicker than Steepest Descent Method.

Taylor series of a function in quadratic terms is given by

$$f(\mathbf{x} + \Delta \mathbf{x}) = f(\mathbf{x}) + \mathbf{c}^T \Delta \mathbf{x} + \frac{1}{2} \Delta \mathbf{x}^T \mathbf{H} \Delta \mathbf{x}$$

minimum value of it is found where

$$\frac{\partial f}{\partial \Delta \mathbf{x}} = 0$$

which results the following relations

$$\begin{aligned} \mathbf{c} + \mathbf{H} \Delta \mathbf{x} &= \mathbf{0} \\ \Delta \mathbf{x} &= -\mathbf{H}^{-1} \mathbf{c} \end{aligned}$$

the descent direction is selected to be

$$\mathbf{d} = \alpha \Delta \mathbf{x}$$

Modified Newton's Method's algorithm is given below,

Step 1 : estimate initial value for the design variables and set  $\mathbf{x}_0$  .

Step 2 : compute gradient of function  $f(\mathbf{x})$  at current point as  $\mathbf{c}^{(k)} = \nabla f(\mathbf{x}^{(k)})$ .

Step 3 : calculate  $|\mathbf{c}^{(k)}|$  , if  $|\mathbf{c}^{(k)}| < \epsilon$  then break the iteration and set  $\mathbf{x}^* = \mathbf{x}^{(k)}$ .

Step 4 : calculate hessian of cost function as

$$\mathbf{H}^{(k)} = -\nabla^2 f(\mathbf{x}^{(k)})$$

$$\mathbf{d}^{(k)} = \mathbf{H}^{(k)^{-1}} \mathbf{c}^{(k)}$$

Step 5 : compute step size  $\alpha^{(k)}$  that minimizes  $f(\mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{d}^{(k)})$ .

Step 6 : update  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{d}^{(k)}$  .

Step 7 : go to Step 2.

Gradient and the Hessian of the problem is found as,

$$\mathbf{c} = [2x_1 - 2x_2 - 4, 4x_2 - 2x_1] , \mathbf{H} = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}$$

The iterations of the problem is given below.

For k=0

$$(1) \quad \mathbf{x}^{(0)} = (1, 1) , k = 0 , \epsilon = 10^{-3}$$

$$(2) \quad \mathbf{c}^{(0)} = [-4, 2]$$

$$(3) \quad |\mathbf{c}^{(0)}| = 4.2426 > \epsilon$$

$$(4) \quad \mathbf{d} = \mathbf{H}^{(0)^{-1}} \mathbf{c}^{(0)}$$

$$= \begin{bmatrix} 1 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$$

$$-\mathbf{c}^{(0)T} \mathbf{H}^{(0)} \mathbf{c}^{(0)} < 0 \implies \text{descent direction}$$

$$(5) \quad \delta f(\mathbf{x}^{(0)} + \alpha^{(0)} \mathbf{d}^{(0)}) = 0$$

$$\alpha = -1$$

$$(6) \quad \mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{d}^{(0)} = [4, 2]$$

### Problem 11.12

For the following problem, complete two iterations of Davidon-Fletcher-Powell method, also check descent condition for the search direction.

$$f(\mathbf{x}) = x_1^2 + 2x_2^2 - 4x_1 - 2x_1x_2 \quad ; \quad \mathbf{x}_0 = (1, 1)$$

### Solution

Davidon-Fletcher-Powell method differs from Generalized Newton Method only on the calculation of the Hessian of the cost function. In this algorithm inverse of the Hessian is approximated and calculated by the approximation in the previous step.

Davidon-Fletcher-Powell algorithm is given below,

Step 1 : estimate initial value for the design variables  $\mathbf{x}_0$  and inverse of Hessian  $\mathbf{A}^{(0)}$ .

Step 2 : compute gradient of function  $f(\mathbf{x})$  at current point as  $\mathbf{c}^{(k)} = \nabla f(\mathbf{x}^{(k)})$ .

Step 3 : calculate  $|\mathbf{c}^{(k)}|$ , if  $|\mathbf{c}^{(k)}| < \epsilon$  then break the iteration and set  $\mathbf{x}^* = \mathbf{x}^{(k)}$ .

Step 4 : calculate search direction as

$$\mathbf{d}^{(k)} = \mathbf{A}^{(k)} \mathbf{c}^{(k)}$$

Step 5 : compute step size  $\alpha^{(k)}$  that minimizes  $f(\mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{d}^{(k)})$ .

Step 6 : update  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{d}^{(k)}$  and calculate  $\mathbf{A}^{(k+1)}, \mathbf{B}^{(k)}, \mathbf{C}^{(k)}$

Step 7 : go to Step 2.

The problem solution steps are given below.

For k=0

$$(1) \quad \mathbf{x}^{(0)} = (1, 1), k = 0, \epsilon = 10^{-3}, \mathbf{A}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$(2) \quad \mathbf{c}^{(0)} = [-4, 2]$$

$$(3) \quad |\mathbf{c}^{(0)}| = 4.2426 > \epsilon$$



$$(4) \quad \mathbf{d} = \mathbf{A}^{(0)} \mathbf{c}^{(0)}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

$$(5) \quad \delta f(\mathbf{x}^{(0)} + \alpha^{(0)} \mathbf{d}^{(0)}) = 0$$

$$\alpha = -0.25$$

$$(6) \quad \mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha^{(0)} \mathbf{d}^{(0)} = [2, 0.5]$$

$$\mathbf{A}^{(1)} = \mathbf{A}^{(0)} + \mathbf{B}^{(0)} + \mathbf{C}^{(0)}$$

$$\mathbf{y}^{(0)} = \mathbf{c}^{(1)} - \mathbf{c}^{(0)} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$$

$$\mathbf{z}^{(0)} = \mathbf{A}^{(0)} \mathbf{y}^{(0)} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$$

$$\mathbf{s}^{(0)} = \alpha \mathbf{d}^{(0)} = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$$

$$\mathbf{B}^{(0)} = \frac{\mathbf{s}^{(0)} \mathbf{s}^{\text{T}(0)}}{\mathbf{s}^{(0)} \mathbf{y}^{(0)}} = \begin{bmatrix} 0.2 & -0.1 \\ -0.1 & 0.05 \end{bmatrix}$$

$$\mathbf{C}^{(0)} = -\frac{\mathbf{z}^{(0)} \mathbf{z}^{\text{T}(0)}}{\mathbf{y}^{(0)} \mathbf{z}^{(0)}} = \begin{bmatrix} -0.36 & -0.48 \\ 0.48 & -0.64 \end{bmatrix}$$

$$\mathbf{A}^{(1)} = \begin{bmatrix} 0.84 & -0.38 \\ 0.38 & 0.41 \end{bmatrix}$$

For k=1

$$(2) \quad \mathbf{c}^{(1)} = [-1, 2]$$

$$(3) \quad |\mathbf{c}^{(1)}| = 2.2361 > \epsilon$$

$$(4) \quad \mathbf{d} = \mathbf{A}^{(1)} \mathbf{c}^{(1)}$$

$$= \begin{bmatrix} 0.84 & 0.38 \\ -0.38 & 0.41 \end{bmatrix} \begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} -0.13 \\ 1.2 \end{bmatrix}$$

$$(5) \quad \delta f(\mathbf{x}^{(1)} + \alpha^{(1)} \mathbf{d}^{(1)}) = 0$$

$$\alpha = 0.3537$$

$$(6) \quad \mathbf{x}^{(1)} = \mathbf{x}^{(1)} + \alpha^{(1)} \mathbf{d}^{(1)} = [1.954, 0.9244]$$

$$\mathbf{A}^{(2)} = \mathbf{A}^{(1)} + \mathbf{B}^{(1)} + \mathbf{C}^{(1)}$$

$$\mathbf{y}^{(1)} = \mathbf{c}^{(2)} - \mathbf{c}^{(1)} = \begin{bmatrix} 2.0591 \\ -2.2103 \end{bmatrix}$$

$$\mathbf{z}^{(1)} = \mathbf{A}^{(1)} \mathbf{y}^{(1)} = \begin{bmatrix} 0.9927 \\ -1.6887 \end{bmatrix}$$

$$\mathbf{s}^{(1)} = \alpha^{(1)} \mathbf{d}^{(1)} = \begin{bmatrix} -0.0460 \\ 0.4244 \end{bmatrix}$$

$$\mathbf{B}^{(1)} = \frac{\mathbf{s}^{(1)} \mathbf{s}^{T(1)}}{\mathbf{s}^{(1)} \mathbf{y}^{(1)}} = \begin{bmatrix} -0.0092 & 0.0849 \\ 0.0849 & -0.7836 \end{bmatrix}$$

$$\mathbf{C}^{(1)} = -\frac{\mathbf{z}^{(1)} \mathbf{z}^{T(1)}}{\mathbf{y}^{(1)} \mathbf{z}^{(1)}} = \begin{bmatrix} 0.5888 & -0.4210 \\ -0.4210 & 0.3010 \end{bmatrix}$$

$$\mathbf{A}^{(2)} = \begin{bmatrix} 1.4696 & 0.0439 \\ -0.7161 & -0.0726 \end{bmatrix}$$

### Problem 11.33

For the following problem, complete two iterations of BFGS method.

$$f(\mathbf{x}) = x_1^2 + 2x_2^2 - 4x_1 - 2x_1x_2 \quad ; \quad \mathbf{x}_0 = (1, 1)$$

### Solution

BFGS method differs from Davidon-Fletcher-Powell from only usage approximation of Hessian rather than approximation of inverse of Hessian.

BFGS algorithm is as follows

Step 1 : estimate initial value for the design variables  $\mathbf{x}_0$  and inverse of Hessian  $\mathbf{A}^{(0)}$ .

Step 2 : compute gradient of function  $f(\mathbf{x})$  at current point as  $\mathbf{c}^{(k)} = \nabla f(\mathbf{x}^{(k)})$ .

Step 3 : calculate  $|\mathbf{c}^{(k)}|$ , if  $|\mathbf{c}^{(k)}| < \epsilon$  then break the iteration and set  $\mathbf{x}^* = \mathbf{x}^{(k)}$ .

Step 4 : calculate search direction from the equation

$$\mathbf{H}^{(k)} \mathbf{d}^{(k)} = \mathbf{c}^{(k)}$$

Step 5 : compute step size  $\alpha^{(k)}$  that minimizes  $f(\mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{d}^{(k)})$ .

Step 6 : update  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{d}^{(k)}$  and calculate  $\mathbf{H}^{(k+1)}, \mathbf{D}^{(k)}, \mathbf{E}^{(k)}$

Step 7 : go to Step 2.

Iterations of BFGS solution of the problem is given below,

For k=0

$$(1) \quad \mathbf{x}^{(0)} = (1, 1), \quad k = 0, \quad \epsilon = 10^{-3}, \quad \mathbf{H}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$(2) \quad \mathbf{c}^{(0)} = [-4, 2]$$

$$(3) \quad |\mathbf{c}^{(0)}| = 4.2426 > \epsilon$$

$$(4) \quad \mathbf{d} = \mathbf{H}^{-1(0)} \mathbf{c}^{(0)}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

$$(5) \quad \delta f(\mathbf{x}^{(0)} + \alpha^{(0)} \mathbf{d}^{(0)}) = 0$$

$$\alpha = -0.25$$

$$(6) \quad \mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha^{(0)} \mathbf{d}^{(0)} = [2, 0.5]$$

$$\mathbf{H}^{(1)} = \mathbf{H}^{(0)} + \mathbf{D}^{(0)} + \mathbf{E}^{(0)}$$

$$\mathbf{y}^{(0)} = \mathbf{c}^{(1)} - \mathbf{c}^{(0)} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$$

$$\mathbf{s}^{(0)} = \alpha \mathbf{d}^{(0)} = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$$

$$\mathbf{D}^{(0)} = \frac{\mathbf{y}^{(0)} \mathbf{y}^{\text{T}(0)}}{\mathbf{y}^{(0)} \mathbf{s}^{(0)}} = \begin{bmatrix} 0.2000 & -0.1000 \\ -0.1000 & 0.0500 \end{bmatrix}$$

$$\mathbf{E}^{(0)} = \frac{\mathbf{c}^{(0)} \mathbf{c}^{\text{T}(0)}}{\mathbf{c}^{(0)} \mathbf{d}^{(0)}} = \begin{bmatrix} 5.4795 & -2.7397 \\ -2.7397 & 1.3699 \end{bmatrix}$$

$$\mathbf{H}^{(1)} = \begin{bmatrix} 8.2795 & -5.1397 \\ -5.1397 & 5.5699 \end{bmatrix}$$

For k=1

$$(2) \quad \mathbf{c}^{(1)} = [-1, -2]$$

$$(3) \quad |\mathbf{c}^{(1)}| = 2.2361 > \epsilon$$

$$(4) \quad \mathbf{H}^{(1)} \mathbf{d} = \mathbf{c}^{(1)}$$

$$\mathbf{d}^{(1)} = \begin{bmatrix} -0.8046 \\ -1.1015 \end{bmatrix}$$

$$(5) \quad \delta f(\mathbf{x}^{(1)} + \alpha^{(1)} \mathbf{d}^{(1)}) = 0$$

$$\alpha = -1.1555$$

$$(6) \quad \mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \alpha^{(1)} \mathbf{d}^{(1)} = [-0.3110, -0.0777]$$

$$\mathbf{H}^{(2)} = \mathbf{H}^{(1)} + \mathbf{D}^{(1)} + \mathbf{E}^{(1)}$$

$$\mathbf{y}^{(1)} = \mathbf{c}^{(2)} - \mathbf{c}^{(1)} = \begin{bmatrix} 2.9297 \\ 1.7728 \end{bmatrix}$$

$$\mathbf{s}^{(1)} = \alpha^{(1)} \mathbf{d}^{(1)} = \begin{bmatrix} 0.9297 \\ 1.2728 \end{bmatrix}$$

$$\mathbf{D}^{(1)} = \frac{\mathbf{y}^{(1)} \mathbf{y}^{\text{T}(1)}}{\mathbf{y}^{(1)} \mathbf{s}^{(1)}} = \begin{bmatrix} -0.2045 & 0.9632 \\ 0.9632 & -4.5369 \end{bmatrix}$$

$$\mathbf{E}^{(1)} = \frac{\mathbf{c}^{(1)} \mathbf{c}^{\text{T}(1)}}{\mathbf{c}^{(1)} \mathbf{d}^{(0)}} = \begin{bmatrix} 0.3325 & 0.6650 \\ 0.6650 & 1.3300 \end{bmatrix}$$

$$\mathbf{H}^{(1)} = \begin{bmatrix} 8.4075 & -3.5115 \\ -3.5115 & 2.3630 \end{bmatrix}$$

Approximation of Hessian is getting closer to the real Hessian of the cost function by subsequent iterations.

# Bibliography

- [1] J. Arora. *Introduction to Optimum Design*. Elsevier Science, 2004.