3.5 or 4.6 Global Optimality & Convex Programming Problems

Where can we have global optimum?

1. If the cost function *f(x)* is continuous on a closed and bounded feasible set, then Weierstrauss Theorem guarantees the existence of a global minimum. Therefore, if we calculate all the local minimum points, then the point that gives the least value to the cost function can be selected as a global minimum for the function. This is called exhaustive search.
2. If the optimization problem can be shown to be convex, then any local minimum is also a global minimum; also the KKT necessary conditions are sufficient for the minimum point.

Global Optimality

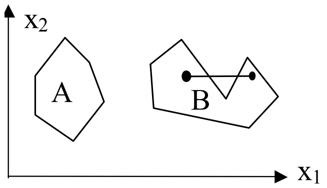
It is difficult to investigate global optimal conditions, except for:

- ✓ The *convex problems* to be discussed next.
- ✓ By using stochastic optimization methods such as simulated annealing or genetic algorithms if you are lucky (These methods are likely to give global optimum).

In general, additional conditions must be imposed upon the model, called *convexity conditions.* These conditions must be satisfied to guarantee that a local constrained optimum (minimum) is also global.

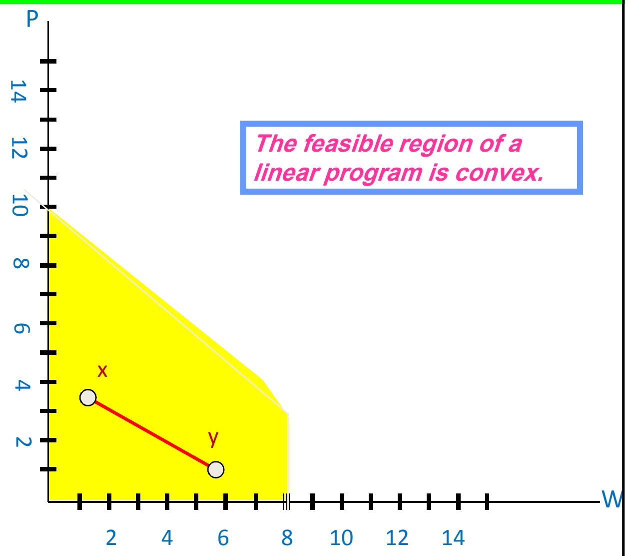
Convex Sets

A set is convex if a line connecting any two points in the feasible space always lies in the feasible region of the design space.



A: Convex

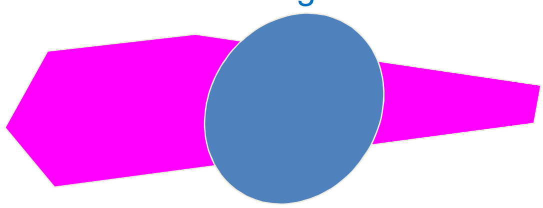
B: Not convex



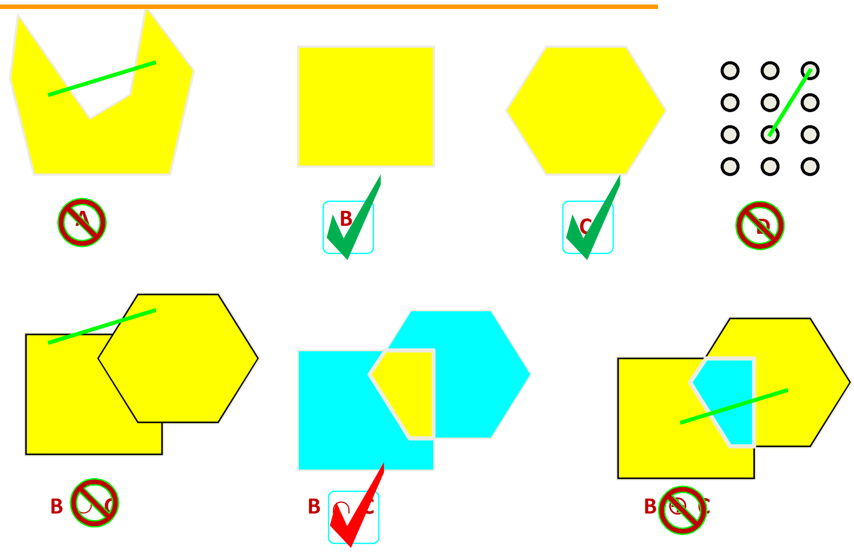
On Convex Feasible Regions

- If all constraints are linear, the feasible region is convex.

- The intersection of convex regions is convex.

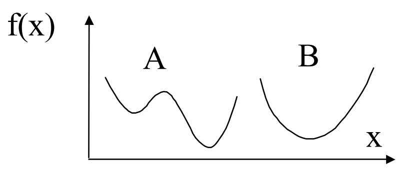


Which one is convex?



Convex Function

- Convex Functions:



A: Not convex

B: Convex

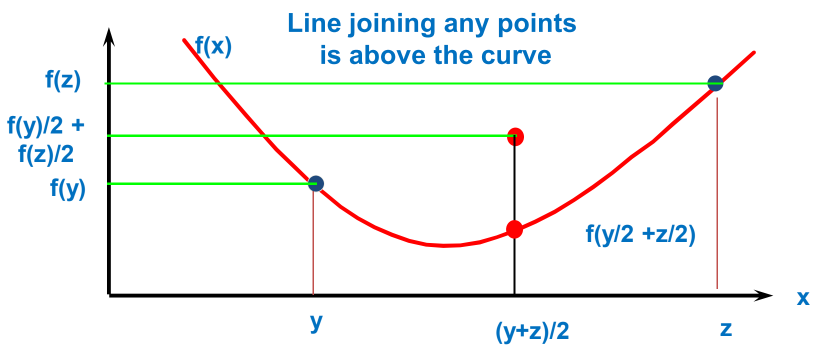
Mathematical Def.: If X is a convex set in R^n. f(x): X $\rightarrow$ R is a **convex function,** if:

$$f(\alpha x_2 + (1-\alpha)x_1) \le \alpha f(x_2) + (1-\alpha) f(x_1) \quad \text{for } 0 \le \alpha \le 1$$

Convex Functions

Example: $\alpha = 1/2$; $f(y/2 + z/2) \leq f(y)/2 + f(z)/2$

We say "*strict" convexity* if sign is "<" for 0< α <1.



Convex Function

Theorem 4.8 Check for Convexity of a Function

A function of n variables $f(x_1, x_2, \cdots, x_n)$ defined on a convex set *S* is convex if and only if the Hessian matrix of the function is positive semidefinite or positive definite at all points in the set *S*.

If the Hessian matrix is positive definite for all points in the feasible set, then *f* is called a strictly convex function.

(Note that the converse of this is not true, i.e., a strictly convex function may have only positive semidefinite Hessian at some points; e.g., $f(x) = x^4$ is a strictly convex (e.g. from its graph) function but its second derivative is zero at *x=0*.)

EXAMPLE 4.37 Check for Convexity of a Function

$$f(X) = x_1^2 + x_2^2 - 1$$

Solution. The domain for the function (which is all values of x_1 and x_2) is convex. The gradient and Hessian of the function are given as

$$\nabla f = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \qquad H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

By either of the tests given in Theorems 4.2 and 4.3 ($M_1=2$, $M_2=4$, $\lambda_1=2$, $\lambda_2=2$), we see that **H** is positive definite everywhere. Therefore, f is a strictly convex function.

EXAMPLE 4.38 Check for Convexity of a Function

$$f(X) = 10 - 4x + 2x^2 - x^3$$

Solution. The second derivative of the function is $d^2f/dx^2=4-6x$. For the function to be convex, $d^2f/dx^2 \geq 0$. Thus, the function is convex only if $4-6x \geq 0$ or $x \leq 2/3$.

The convexity check actually defines a domain for the function over which it is convex. It can be seen in Fig. 4-30 that the function is convex for $x \leq 2/3$ and concave for $x \geq 2/3$ [a function f(x) is called concave if -f(x) is convex].

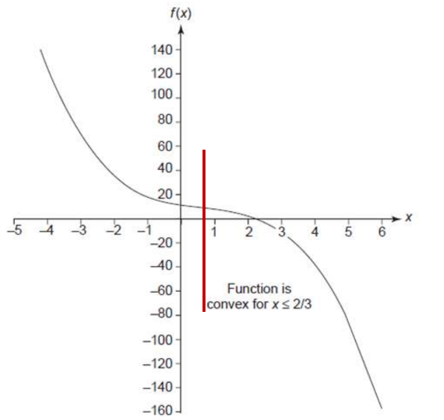


FIGURE 4-30 Graph of the function $f(x) = 10 - 4x + 2x^2 - x^3$ of Example 4.38.

Theorem 4.9 Convex Functions and Convex Sets

Let a set *S* be defined with constraints of the general optimization problem in Eqs (4.37) to (4.39) as

$$S = \{x \mid h_j(x) = 0;\ j = 1 \text{ to } p,\ g_i(x) \le 0;\ i = 1 \text{ to } m\}$$

Then *S* is a convex set if functions g_i are convex and h_j are linear.

It is important to note that Theorem 4.9 does not say that the feasible set S cannot be convex if a constraint function $g_i(x)$ fails the convexity check, i.e., it is not an "if and only if" theorem. There are some problems having inequality constraint functions that fail the convexity check, but the feasible set is still convex. Thus, the condition that $g_i(x)$ be convex for the region $g_i(x) \le 0$ to be convex are only sufficient but not necessary.

Theorem 4.10 Global Minimum

If *f(x*)* is a local minimum for a convex function *f(x)* defined on a convex feasible set *S*, then it is also a global minimum.

Important notes:

• The theorem does not say that x* cannot be a global minimum point if functions of the problem fail the convexity test.

The point may indeed be a global minimum; however, we cannot claim global optimality using Theorem 4.10. We will have to use some other procedure, such as exhaustive search.

• The theorem does not say that the global minimum is unique; i.e., there can be multiple minimum points in the feasible set, all having the same cost function value.

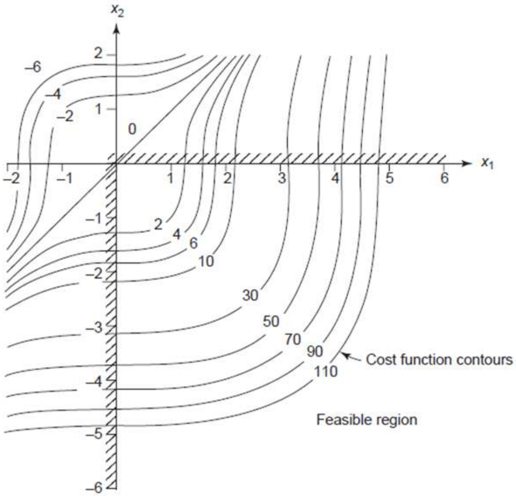
EXAMPLE 4.39 Check for Convexity of a Problem

Minimize $f(x_1,x_2)=x_1^3-x_2^3$ subject to the constraints $x_1 \geq 0$, $x_2 \leq 0$.

This domain is convex.

$$H = \begin{bmatrix} 6x_1 & 0 \\ 0 & -6x_2 \end{bmatrix}$$

The Hessian is positive semidefinite or positive definite over the domain defined by the constraints ($x_1 \geq 0$, $x_2 \leq 0$). Therefore, the cost function is convex and the problem is convex.



EXAMPLE 4.40 Check for Convexity of a Problem

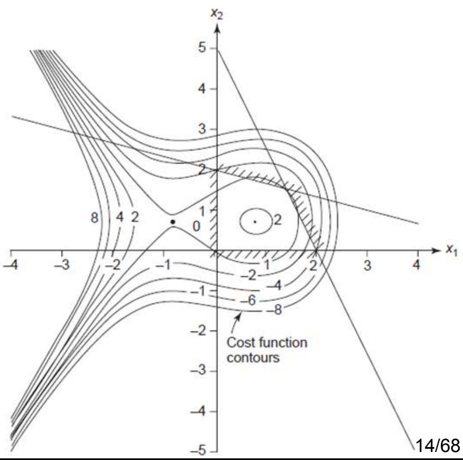
Minimize $f(x_1,x_2)=2x_1+3x_2-x_1^3-2x_2^2$ subject to the constraints

$$x_1+3x_2 \leq 6,\ 5x_1+2x_2 \leq 10,\ x_1,x_2 \geq 0$$

Since all the constraint functions are linear in the variables x_l and x_2, *the* feasible set for the problem is convex. If the cost function *f* is also convex, then the problem is convex. The Hessian of the cost function is

$$H = \begin{bmatrix} -6x_1 & 0 \\ 0 & -4 \end{bmatrix}$$

The eigenvalues of *H* are $-6x_1$ and -4. Since the first eigenvalue is nonpositive for $x_1 \geq 0$, and the second eigenvalue is negative, the function is not convex (Theorem 4.8), so the problem cannot be classified as a convex programming problem.



4.6.4 Transformation of a Constraint

Transformation of a constraint function (the constraint boundary and the feasible set for the problem do not change), however, may affect its convexity check, i.e., transformed constraint function may fail the convexity check.

Convexity of the feasible set is, however, not affected by the transformation.

Example:

$$g_1 = \frac{a}{x_1 x_2} - b \le 0 \qquad \text{(a)}$$

with $x_1>0$, $x_2>0$, and a and b as the given positive constants.

To check convexity:

$$\nabla^2 g_1 = \frac{2a}{x_1^2 x_2^2}\begin{bmatrix} \frac{x_2}{x_1} & 0.5 \\ 0.5 & \frac{x_2}{x_1} \end{bmatrix} \qquad \text{(b)}$$

Both eigenvalues as well as the two leading principal minors of the preceding matrix are strictly positive, so the matrix is positive definite and the constraint function g_1 is convex. The feasible set for g_1 is convex.

Now: Transform the constraint by multiplying by x_1x_2
(since $x_1>0$, $x_2>0$, the sense of the inequality is not changed)

$$g_2 = a - bx_1x_2 \le 0 \qquad \text{(c)}$$

The constraints g_1 and g_2 are equivalent and will give the same optimum solution for the problem.

$$\nabla^2 g_2 = \begin{bmatrix} 0 & -b \\ -b & 0 \end{bmatrix} \qquad \text{(d)}$$

The eigenvalues of the preceding matrix are: $\lambda_1=-b$ and $\lambda_2=b$.

The matrix is indefinite by Theorem 4.2, and by Theorem 4.8, the constraint function g_2 is not convex.

Thus, we lose convexity of the constraint function and we cannot claim convexity of the feasible set by Theorem 4.9.
Since the problem cannot be shown to be convex, we cannot use results related to convex programming problems.

4.6.5 Sufficient Conditions for Convex Programming Problems

Theorem 4.11 Sufficient Condition for Convex Programming Problem

If *f (x)* is a convex cost function defined on a convex feasible set, then the first-order KKT conditions are necessary as well as sufficient for a global minimum.

TABLE 4-3 Convex Programming Problem-Summary of Results

The problem must be written in the standard form: Minimize *f (x)* subject to $h_i(x) = 0, g_j \le 0$

1. *Convex set.* The geometrical condition that a line joining two points in the set is to be in the set, is an "if and only if" condition for convexity of the set.
2. *Convexity of feasible set S.* All the constraint functions should be convex. This condition is only sufficient but not necessary; i.e., functions failing the convexity check may also define convex sets.
   - nonlinear equality constraints always give nonconvex sets
   - linear equalities or inequalities always give convex sets

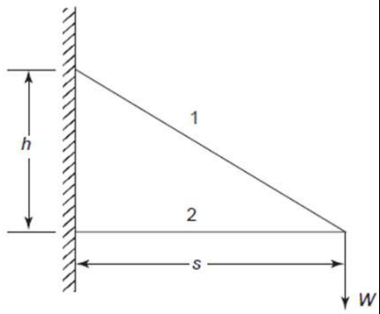
3. *Convex functions.* A function is convex if and only if its Hessian is at least positive semidefinite everywhere.
   A function is strictly convex if its Hessian is positive definite everywhere.
   The converse is not true:
   A strictly convex function may not have a positive definite Hessian everywhere.
   Thus this condition is only sufficient but not necessary.
4. *Form of constraint function.* Changing the form of a constraint function can result in failure of the convexity check for the new constraint or vice versa.
5. *Convex programming problem.* *f(x)* is convex over the convex feasible set S.
   - KKT first order conditions are necessary as well as sufficient for global minimum
   - Any local minimum point is also a global minimum point

Nonconvex programming problem:

If a problem fails convexity checks, it does not imply that there is no global minimum for the problem. It could also have only one local minimum in the feasible set *S* which would then be a global minimum as well.

Design of a Wall Bracket

The wall bracket is to be designed to support a load of W=1.2 MN. The material for the bracket should not fail under the action of forces in the bars. These are expressed as the following stress constraints:



h=30 cm, s=40 cm

Bar 1: $\sigma_1 \le \sigma_a$

Bar 2: $\sigma_2 \le \sigma_a$

σ_a= allowable stress for the material (16,000 N/cm^2)
σ_1 = stress in Bar 1 which is given as F_1/A_1, N/cm^2
σ_2 = stress in Bar 2 which is given as F_2/A_2, N/cm^2
A_1 = cross-sectional area of Bar 1 (cm^2)
A_2 = cross-sectional area of Bar 2 (cm^2)
F_1 = force due to load W in Bar 1 (N)
F_2 = force due to load W in Bar 2 (N)

Total volume of the bracket is to be minimized.

Design variables: A_1 and A_2

Objective function: $f(A_1, A_2) = l_1 A_1 + l_2 A_2, \quad \text{cm}^3$ (a)

l_1= =50 cm and l_2 = 40 cm

Forces on bar 1 and bar 2 are:

$$F_1 = (2.0E+06)\ \text{N},\ F_2 = (1.6E+06)\ \text{N}$$

Stress constraints:

$$g_1 = \frac{(2.0E+06)}{A_1} - 16000 \le 0 \quad \text{(b)}$$

$$g_2 = \frac{(1.6E+06)}{A_2} - 16000 \le 0 \quad \text{(c)}$$

$$g_3 \equiv -A_1 < 0, \quad g_4 \equiv -A_2 < 0, \quad \text{(d)}$$

Convexity

Since the cost function is linear in terms of design variables, it is convex.

For constraints:

$$\nabla^2 g_1 = \begin{bmatrix} \frac{(4.0\times10^6)}{A_1^3} & 0 \\ 0 & 0 \end{bmatrix}$$

which is a positive semidefinite matrix for $A_1>0$, so g_1 is convex. Similarly, g_2 is convex, and since g_3 and g_4 are linear, they are convex. Thus the problem is convex, and KKT necessary conditions are also sufficient and any design satisfying the KKT conditions is a global minimum.

KKT Necessary Conditions

$$L = (l_1 A_1 + l_2 A_2) + u_1\left[\frac{(2.0E+06)}{A_1} - 16000 + s_1^2\right]$$

$$+u_2\left[\frac{(1.6E+06)}{A_2} - 16000 + s_2^2\right] + u_3(-A_1 + s_3^2) + u_4(-A_2 + s_4^2) \quad \text{(e)}$$

$$\frac{\partial L}{\partial A_1} = l_1 - u_1 \frac{(2.0E+06)}{A_1^2} - u_3 = 0 \qquad (f)$$

$$\frac{\partial L}{\partial A_2} = l_2 - u_2 \frac{(1.6E+06)}{A_1^2} - u_4 = 0 \qquad (g)$$

$$u_i s_i = 0, u_i \geq 0, g_i + s_i^2 = 0, s_i^2 \geq 0; \; i = 1 \text{ to } 4 \qquad (h)$$

TABLE 4-4 Definition of Karush-Kuhn-Tucker Cases with Four Inequalities

No.	Case	Active Constraints
1	$u_1 = 0, u_2 = 0, u_3 = 0, u_4 = 0$	No inequality active
2	$s_1 = 0, u_2 = 0, u_3 = 0, u_4 = 0$	One inequality active at a time
3	$u_1 = 0, s_2 = 0, u_3 = 0, u_4 = 0$	
4	$u_1 = 0, u_2 = 0, s_3 = 0, u_4 = 0$	
5	$u_1 = 0, u_2 = 0, u_3 = 0, s_4 = 0$	
6	$s_1 = 0, s_2 = 0, u_3 = 0, u_4 = 0$	Two inequalities active at a time
7	$u_1 = 0, s_2 = 0, s_3 = 0, u_4 = 0$	
8	$u_1 = 0, u_2 = 0, s_3 = 0, s_4 = 0$	
9	$s_1 = 0, u_2 = 0, u_3 = 0, s_4 = 0$	
10	$s_1 = 0, u_2 = 0, s_3 = 0, u_4 = 0$	
11	$u_1 = 0, s_2 = 0, u_3 = 0, s_4 = 0$	
12	$s_1 = 0, s_2 = 0, s_3 = 0, u_4 = 0$	Three inequalities active at a time
13	$u_1 = 0, s_2 = 0, s_3 = 0, s_4 = 0$	
14	$s_1 = 0, u_2 = 0, s_3 = 0, s_4 = 0$	
15	$s_1 = 0, s_2 = 0, u_3 = 0, s_4 = 0$	
16	$s_1 = 0, s_2 = 0, s_3 = 0, s_4 = 0$	Four inequalities active at a time

Note that any case that requires s_3=0 (i.e., g_3=0) makes the area A_1=0. For such a case the constraint g_1 of Eq. (b) is violated, so it does not give a candidate solution.

Similarly, s_4=0 makes A_2=0, which violates the constraint of Eq. (c).

In addition, A_1 and A_2 cannot be negative because the corresponding solution has no physical meaning.

Therefore, all the cases requiring either s_3=0 and/or s_4=0 do not give any candidate solution.

This leaves only cases 1 to 3 and 6 for further consideration.

Case 1: u_1=0, u_2=0, u_3=0, u_4=0. This case gives l_1=0 and l_2=0 in Eqs. (f) and (g) which is not acceptable.

Case 2: s_1=0, u_2=0, u_3=0, u_4=0. This gives l_2=0 in Eq. (g) which is not acceptable.

Case 3: u_1=0, s_2=0, u_3=0, u_4=0. This gives l_1=0 in Eq. (f) which is not acceptable.

Case 6: s_1=0, s_2=0, u_3=0, u_4=0.

From Equations (b) and (c): A_1*=125 cm^2, A_2*=100 cm^2.

From Equations (f) and (g): u_1=0.391>0 and u_2=0.25>0,

Therefore, all the KKT conditions are satisfied; and A_1*=125 cm^2, A_2*=100 cm^2 is a global minimum.

The cost function at optimum: f*=50(125)+40(100) or f*=10,250 cm^3.

The gradients of active constraints: (-(2.0*10^6)/A_1^2, 0) and (0, -(1.0*10^6)/A_2^2). These vectors are linearly independent, and so the minimum point is a regular point of the feasible set.

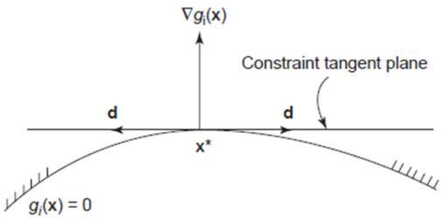
مسائل زیر را حل کرده و تا دو هفته دیگر تحویل فرمایید:

3) 61, 69

تمرین‌های قبلی:

3) 1,8,16,17, 29, 30, 46

Second Order Conditions for General Constrained Problems



In the constrained case, we must also consider active constraints at x^* to determine feasible changes d. We will consider only the points $x=x^*+d$ in the neighborhood of x^* that satisfy the active constraint equations. Any $d \neq 0$ satisfying active constraints to the first order must be in the constraint tangent hyperplane. Such d's are then orthogonal to the gradients of the active constraints since constraint gradients are normal to the constraint tangent hyperplane.

$$g(x) = g(x^*) + \nabla g^T(x^*)d + \ldots$$

Therefore,

$$\nabla h_i^T d = 0 \text{ and } \nabla g_i^T d = 0.$$

Second Order Necessary Conditions for General Constrained Problems

Theorem 3.11 or **Theorem 5.1.**

Second-order necessary condition for general constrained problem.

Let x^* satisfy the first order K-T necessary conditions. Define Hessian of the Lagrange function L at x^* as:

$$\nabla^2 L = \nabla^2 f + \sum_{i=1}^{p} v_i^* \nabla^2 h_i + \sum_{i=1}^{m} u_i^* \nabla^2 g_i$$

Let there be nonzero feasible directions $(d \neq 0)$ satisfying the following linear systems at the point x*:

$$\begin{cases} \nabla h_i^T \mathbf{d} = 0; i = 1 \text{ to p} \\ \nabla g_i^T \mathbf{d} = 0 \text{ (for all active inequalities (i.e. for those i with } g_i(x^*)=0) \end{cases}$$

Then if x* is a local minimum point, then it must be true that:

$Q \geq 0$ where $Q = \mathbf{d}^T \nabla^2 L(x^*)\mathbf{d}$

Note that any point that does not satisfy the second-order necessary conditions cannot be a local minimum point.

Sufficient Conditions for General Constrained Problems

Theorem 3.12 or Theorem 5.2.

Let x^* satisfy the first-order K-T necessary conditions for the general optimum design problem. Define Hessian of the Lagrange function L at x^* as:

$$\nabla^2 L = \nabla^2 f + \sum_{i=1}^{p} v_i^* \nabla^2 h_i + \sum_{i=1}^{m} u_i^* \nabla^2 g_i$$

Define nonzero feasible conditions $(d \neq 0)$ as solutions of the linear systems:

$$\begin{cases} \nabla h_i^T \mathbf{d} = 0; i = 1 \text{ to p} & (5.10) \\ \nabla g_i^T \mathbf{d} = 0 \text{ (for active inequalities with } u_i>0) & (5.11) \end{cases}$$

Also let $\nabla g_i^T \mathbf{d} \le 0$ for those constraints with $u_i=0$. (5.12)

If $Q > 0$ where $Q = \mathbf{d}^T \nabla^2 L(x^*)\mathbf{d}$ then

x^* is an isolated local minimum point (it means that there are no other local minimum points in the neighborhood of x^*)

Strong Sufficient Condition

Theorem 3.13 or Theorem 5.3.

Let x* satisfy the first-order K-T necessary conditions for the general optimum design problem. Define Hessian $\nabla^2 L(x^*)$ for the Lagrange function at x* as:

$$\nabla^2 L = \nabla^2 f + \sum_{i=1}^{p} v_i^* \nabla^2 h_i + \sum_{i=1}^{m} u_i^* \nabla^2 g_i$$

Then if $\nabla^2 L(x^*)$ is positive definite, x* is an isolated minimum point.

A Special Case

At the candidate minimum point *x :**

The total number of active constraints (with at least one inequality) =

The number of independent design variables

That is, there are no design degrees of freedom.

حل دستگاه معادلات قیود فعال در آن نقطه یک جواب خواهد داشت.

Since x* satisfies KKT conditions, gradients of all the active constraints are linearly independent. Thus, the only solution for the system of Eqs.(5.10)($\nabla h_i^T \mathbf{d} = 0; i = 1 \text{ to } p$) and (5.11) ($\nabla g_i^T \mathbf{d} = 0$) is d=0 and Theorem 5.2 cannot be used. However, since d=0 is the only solution, there are no feasible directions in the neighborhood that can reduce the cost function any further. Thus, the point **x*** is indeed a local minimum for the cost function (see the next example)

Check for Sufficient Conditions

EXAMPLE 5.4

minimize $f(x) = \frac{1}{3}x^3 - \frac{1}{2}(b+c)x^2 + bcx + f_0$

subject to $a \le x \le d$ where 0<a<b<c<d and f_0 are specified constants

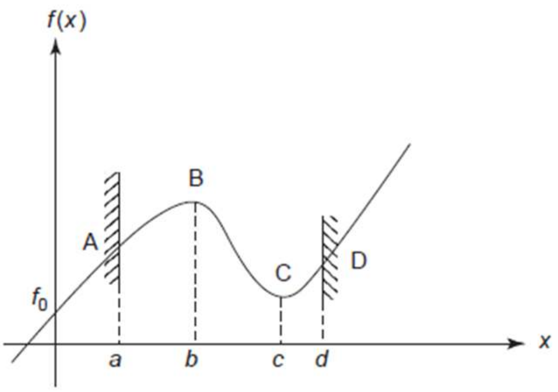


FIGURE 4-20 Graphical representation.
Point A, constrained local minimum;
B, unconstrained local maximum;
C, unconstrained local minimum;
D, constrained local maximum.

$g_1 = a - x \le 0; \quad g_2 = x - d \le 0$

$L = \frac{1}{3}x^3 - \frac{1}{2}(b+c)x^2 + bcx + f_0 + u_1(a - x + s_1^2) + u_2(x - d + s_2^2)$

$$\frac{\partial L}{\partial x} = x^2 - (b+c)x + bc - u_1 + u_2 = 0 \qquad (c)$$

$$(a-x) + s_1^2 = 0,\ s_1^2 \ge 0;\ (x-d) + s_2^2 = 0,\ s_2^2 \ge 0 \qquad (d)$$

$$u_1 s_1 = 0,\ u_2 s_2 = 0 \qquad (e)$$

$$u_1 \ge 0,\ u_2 \ge 0 \qquad (f)$$

4 normal cases:

Case 1: u_1=0, u_2=0. (no constraint is active)

Eq. (c) gives two solutions as *x=b* and *x=c*.

for *x=b*: s_1^2=b-a>0 ; s_2^2=d-b>0 (g)

for *x=c*: s_1^2=c-a>0 ; s_2^2=d-c>0 (h)

Both points are candidate of minimum points (unconstrained)

Check for sufficient conditions:

$$x = b; \quad \frac{d^2 f}{dx^2} = 2x - (b+c) = b - c < 0$$ a local maximum point

$$x = c; \quad \frac{d^2 f}{dx^2} = c - b > 0$$ a local minimum point

35/68

Case 2: u_1=0, s_2=0.

$$\text{from (d)} \rightarrow x = d$$

$$u_2 = -[d^2 - (b+c)d + bc] = -(d-c)(d-b)$$

$$\text{Since } d > c > b,\ u_2 \text{ is } < 0$$

The KKT necessary condition is violated

Case 3: s_1=0, u_2=0.

↓

x=a is a candidate minimum point. g_1 is active

$$u_1 = a^2 - (b+c)a + bc = (a-b)(a-c) > 0$$

Case 4: s_1=0, s_2=0.

does not give any valid solution since x cannot be simultaneously equal to *a* and *d*.

گردآوری و تنظیم: محمدحسین ابوالبشری 36/68

Solution. There is only one constrained candidate local minimum point, *x=a.* Since there is only one design variable and one active constraint, the condition $\nabla g_1 \bar{d} = 0$ of Eq. (5.11) gives $\bar{d} = 0$ as the only solution (note that $\bar{d}$ is used as a direction for sufficiency check since *d* is used as a constant in the example). Therefore, Theorem 5.2 cannot be used for a sufficiency check.

Also note that at *x=a,* $d^2L/dx^2 = 2a - b - c$ which can be positive, negative, or zero depending on the values of *a, b,* and *c.* So, we cannot use curvature of Hessian to check the sufficiency condition (Strong Sufficient).

However, from Fig. 4-20 we observe that *x=a* is indeed an isolated local minimum point.

From this example we can conclude that if the number of active inequality constraints is equal to the number of independent design variables and all other KKT conditions are satisfied, then the candidate point is indeed a local minimum design.

Check for Sufficient Conditions

minimize $f(x)=x_1^2+x_2^2-3x_1x_2$

subject to $g=x_1^2+x_2^2-6\leq 0.$

$L=x_1^2+x_2^2-3x_1x_2+u(x_1^2+x_2^2-6+s^2)$

$$\frac{\partial L}{\partial x_1} = 2x_1 - 3x_2 + 2ux_1 = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 3x_1 + 2ux_2 = 0$$

$x_1^2+x_2^2-6+s^2=0,\ s^2\geq 0,\ u\geq 0$

$us=0$

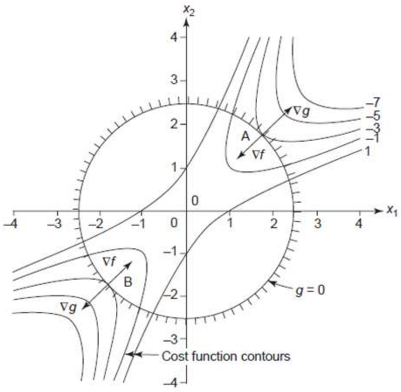


FIGURE 4-21 Graphical solution for Example 4.31. Local minimum points, A and B.

3 possible cases:

Case 1: u=0.

$$2x_1 - 3x_2 = 0;\ 2x_2 - 3x_1 = 0$$

$x_1^* = 0, x_2^* = 0; f(0,0) = 0$ $s^2=6$ candidate minimum point

Case 2: s=0.

$$x_1 = x_2 = \sqrt{3}, u = \frac{1}{2} \quad x_1 = x_2 = -\sqrt{3}, u = \frac{1}{2}$$

$$x_1 = -x_2 = \sqrt{3}, u = -\frac{5}{2} \quad x_1 = -x_2 = \sqrt{3}, u = -\frac{5}{2}$$

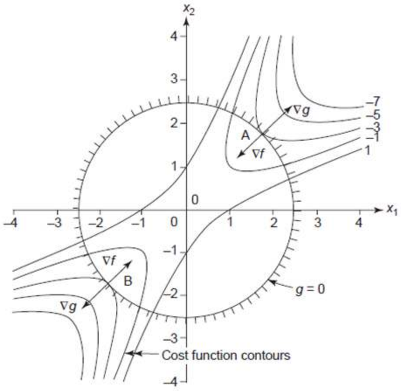


FIGURE 4-21 Graphical solution for Example 4.31. Local minimum points, A and B.

~~Case 3: u=0, s=0.~~

$s^2 = 6 \neq 0$ KKT conditions cannot be satisfied

1. $x_1{}^* = 0,\ x_2{}^* = 0,\ u^* = 0,\ f(0,0) = 0,$ Point O in Fig. 4-21

2. $x_1{}^* = x_2{}^* = \sqrt{3},\ u^* = \frac{1}{2},\ f(\sqrt{3},\sqrt{3}) = -3,$ Point A in Fig. 4-21

3. $x_1{}^* = x_2{}^* = -\sqrt{3},\ u^* = \frac{1}{2},\ f(-\sqrt{3},-\sqrt{3}) = -3,$ Point B in Fig. 4-21

Check for Sufficient Conditions

$(i)\ x^* = (0,0),\ u^* = 0,\ \ (ii)\ x^* = (\sqrt{3},\sqrt{3}),\ u^* = \frac{1}{2},\ \ (iii)\ x^* = (-\sqrt{3},-\sqrt{3}),\ u^* = \frac{1}{2} \quad (a)$

$$\nabla^2 f = \begin{bmatrix} 2 & -3 \\ -3 & 2 \end{bmatrix} \qquad \nabla^2 g = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

since eigenvalues of $\nabla^2 f$ are -1 and 5 — f is not convex

eigenvalues of $\nabla^2 g$ are $\lambda_1 = 2$ and $\lambda_2 = 2$ — function g is convex

Therefore, it cannot be classified as a convex programming problem

$$\nabla^2 L = \nabla^2 f + u\nabla^2 g = \begin{bmatrix} 2+2u & -3 \\ -3 & 2+2u \end{bmatrix}$$

For the first point x* = (0,0), u*=0, $\nabla^2 L$ becomes $\nabla^2 f$ (the constraint g(x)≤0 is inactive). In this case the problem is unconstrained and the local sufficiency requires $d^T \nabla^2 f(x^*) d > 0$ for all **d**. Or, $\nabla^2 f$ should be positive definite at x*. Since both eigenvalues of $\nabla^2 f$ are not positive, we conclude that the above condition is not satisfied.

Therefore, x*=(0,0) does not satisfy the second-order sufficiency condition.

Note that since $\lambda_1 = -1$ and $\lambda_2 = 5$, the matrix $\nabla^2 f$ is indefinite at x^*

Therefore, the point x*=(0,0) violates the second-order necessary condition of Theorem 4.4

(شرط لازم و کافی برای مینیمم محلی مسایل نامقید)

requiring $\nabla^2 f$ to be positive semidefinite or definite at the candidate local minimum point. Thus, x*=(0,0) cannot be a local minimum point.

At points $x^*(\sqrt{3}, \sqrt{3}), u^* = \frac{1}{2}$ and $x^*(-\sqrt{3}, -\sqrt{3}), u^* = \frac{1}{2}$

$$\nabla^2 L = \nabla^2 f + u\nabla^2 g = \begin{bmatrix} 2+2u & -3 \\ -3 & 2+2u \end{bmatrix} = \begin{bmatrix} 3 & -3 \\ -3 & 3 \end{bmatrix}$$

Eigenvalues:

$$(3-\lambda)^2 - 9 = 0$$

$$\lambda_1 = 0, \lambda_2 = 6$$

$$\therefore \text{Semidefinite}$$

$$\nabla g = \pm(2\sqrt{3}, 2\sqrt{3}) = \pm 2\sqrt{3}(1,1)$$

It may be checked that $\nabla^2 L$ is not positive definite at either of the two points. Therefore, we cannot use Theorem 5.3 to conclude that x* is a minimum point. We must find d satisfying Eqs. (5.10) and (5.11). If we let $d=(d_1,d_2)$, then $\nabla g^T d = 0$ gives

$$\pm 2\sqrt{3}(1,1)\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0; \quad \text{or} \quad d_1 + d_2 = 0$$

Thus, $d_1=-d_2=c$, where $c\neq 0$ is an arbitrary constant, and a $d\neq 0$ satisfying $\nabla g^T d = 0$ is given as d=c(1,-1). The sufficiency condition of Eq. (5.12) gives

$$Q = d^T(\nabla^2 L)d = c\,[1,-1]\begin{bmatrix} 3 & -3 \\ -3 & 3 \end{bmatrix} c \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 12c^2 > 0 \quad \text{for} \quad c \neq 0$$

The points $x^* = (\sqrt{3}, \sqrt{3})$ and $x^* = (-\sqrt{3}, -\sqrt{3})$ satisfy the sufficiency conditions. They are therefore isolated local minimum points as was observed graphically in Example 4.31 and Fig. 4-21. We see for this example that $\nabla^2 L$ is not positive definite, but x* is still an isolated minimum point.

Note that since f is continuous and the feasible set is closed and bounded, we are guaranteed the existence of a global minimum by the Weierstrass Theorem 4.1.

Also we have examined every possible point satisfying necessary conditions. Therefore, we must conclude by elimination that $x^* = (\sqrt{3}, \sqrt{3})$ and $x^* = (-\sqrt{3}, -\sqrt{3})$ are global minimum points. The value of the cost function for both points is f(x*)=-3.

Check for Sufficient Conditions

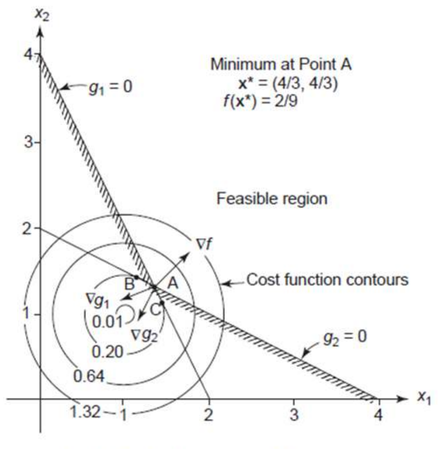


FIGURE 4-22 Graphical solution for Example 4.32.

minimize $f(x_1,x_2)=x_1^2+x_2^2-2x_1-2x_2+2$

subject to $g_1=-2x_1-x_2+4\le 0.$

$g_2=-x_1-2x_2+4\le 0.$

The KKT necessary conditions are satisfied for the point $x_1^* = \frac{4}{3},\ x_2^* = \frac{4}{3},\ u_1^* = \frac{2}{9},\ u_2^* = \frac{2}{9},$

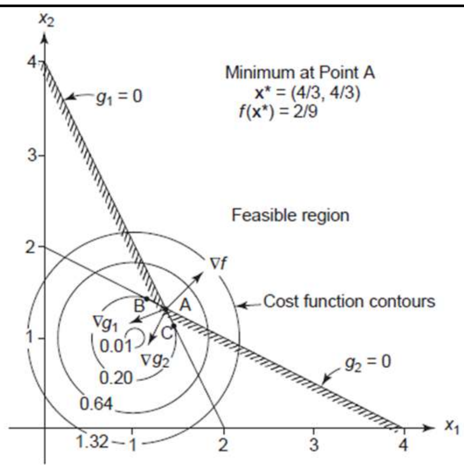


FIGURE 4-22 Graphical solution for Example 4.32.

We have already recognized the Geometrical sufficient condition

It can be observed in Fig. 4-22 that the vector $-\nabla f$ can be expressed as a linear combination of the vectors ∇g_1 and ∇g_2 at point A. This satisfies the necessary condition of Eq. (4.52). It can also be seen from the figure that point A is indeed a local minimum because any further reduction in the cost function is possible only if we go into the infeasible region. Any feasible move from point A results in an increase in the cost function.

Since all the constraint functions are linear, the feasible set *S* is convex. The Hessian of the cost function is positive definite. Therefore, it is also convex and the problem is convex. By Theorem 4.11, $x_1^* = \frac{4}{3},\ x_2^* = \frac{4}{3}$ *satisfies sufficiency conditions for a global minimum* with the cost function as $f(x^*) = \frac{2}{9}$

Note that local sufficiency cannot be shown by the method of Theorem 5.2. The reason is that the conditions of ($\nabla g_i^T \mathbf{d} = 0$) give two equations in two unknowns:

$$-2d_1 - d_2 = 0,\ -d_1 - 2d_2 = 0 \qquad \text{(b)}$$

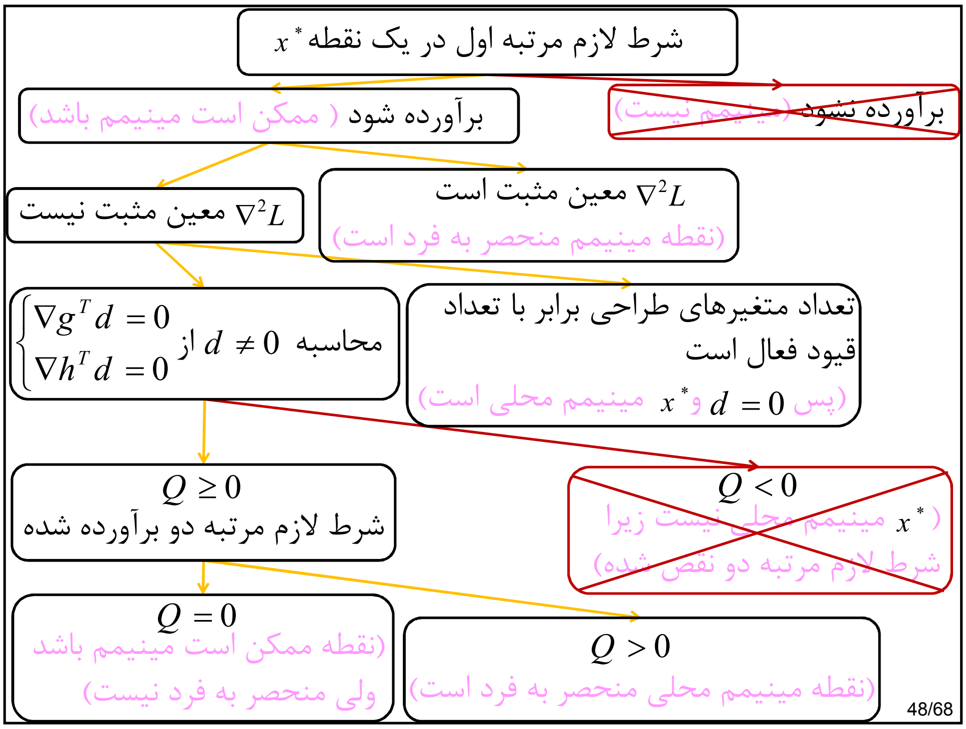
This is a homogeneous system of equations with a nonsingular coefficient matrix. Therefore, its only solution is $d_1 = d_2 = 0$. Thus, we cannot find a $d \neq 0$ for use in the condition of ($Q > 0$), and Theorem 5.2 cannot be used. However, we have seen in the foregoing and in Fig. 4-22 that the point is actually an isolated global minimum point. Since it is a two-variable problem and two inequality constraints are active at the KKT point, the condition for local minimum is satisfied.

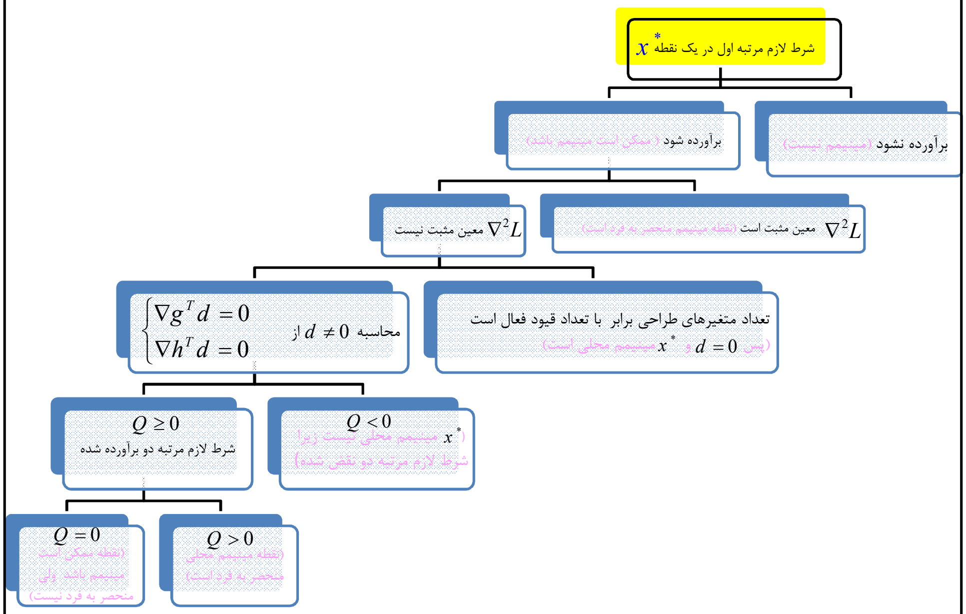
Two Points

1. Global optimum solutions can be obtained for problems that cannot be classified as convex programming problems. We cannot show global optimality unless we find all the local optimum solutions in the closed and bounded set (Weierstrass Theorem 4.1).

2. If sufficiency conditions are not satisfied, the only conclusion we can draw is that the candidate point need not be an isolated minimum. It may have many local optima in the neighborhood, and they may all be actually global solutions.

47/68





مسائل زیر را حل کرده و تا دو هفته دیگر تحویل فرمایید:

3) 84, 93

تمرین‌های قبلی:

3) 1,8,16,17, 29, 30, 46, 61, 69

4.5 Postoptimality Analysis: Physical Meaning of Lagrange Multipliers

۳.۷ تحلیل پس بهینگی (تحلیل حساسیت)

- مطالعه تغییرات جواب بهین نسبت به تغییر بعضی پارامترهای مسئله اولیه به عنوان تحلیل پس بهینگی یا تحلیل حساسیت شناخته می‌شود که موضوعی مهم در زمینه طراحی بهین است.
- تغییرات جواب بهین (تابع هزینه و متغیرهای طراحی) در اثر تغییر پارامترهای مختلفی می‌تواند مورد بررسی قرار گیرد ولی ما روی حساسیت تابع هزینه نسبت به تغییرات حدود قیود تأکید خواهیم کرد.

- در این تحلیل فرض خواهیم کرد که مسئله مینیمم سازی با قیود:

$$h_i(X) = 0 \text{ and } g_j(X) \le 0$$

(یعنی با حدود فعلی قیود که صفرند) حل شده است. می‌خواهیم بدانیم وقتی حدود قید از صفر تغییر کنند چه اتفاقی برای تابع هزینه می‌افتد.

- بحث حساسیت به تعبیر فیزیکی ضریب لاگرانژ ختم می‌شود که در کاربردهای عملی می‌تواند بسیار مفید باشد. تعبیر فیزیکی نیز نشان خواهد داد که چرا ضریب لاگرانژ قیود "نوع $\le$" باید نامنفی باشند. افزون بر این ضرایب می‌توانند برای مطالعه معایب محدود کردن دامنه قیود یا مزایای گسترش آن‌ها استفاده شوند.
- گسترش قیود، ناحیه قابل قبول (مجموعه قید) را بزرگ می‌کند در صورتی که محدود کردن قیود، دامنه آن را کوچک می‌کند.

اثرات تغییر در حدود قیود

- برای بحث تغییرات جواب بهین (تابع هزینه و متغیرهای طراحی) نسبت به حدود قیود، یک مسئله بهینه‌سازی با رابطه‌سازی استاندارد زیر را در نظر می‌گیریم:

$$\begin{aligned} &\min f(X) \\ &h_i(X) = 0; \quad i = 1 \text{ to } p \\ &g_j(X) \le 0; \quad j = 1 \text{ to } m \end{aligned}$$

- حال با تغییرات کوچک (در همسایگی صفر) در حدود قیود داریم:

$$\begin{aligned} &\min f(\mathrm{X}) \\ &h_i(\mathrm{X}) = b_i; \quad i = 1 \text{ to } p \\ &g_j(\mathrm{X}) \le e_j; \quad j = 1 \text{ to } m \end{aligned}$$

53/68

- روشن است که نقطه بهین برای مسئله جدید به بردارهای b و e بستگی دارد :

$$X^* = X^*(b,e)$$

- همچنین، مقدار بهین تابع هزینه نیز به b و e وابسته است:

$$f = f(b,e)$$

ولی وابستگی صریح تابع هزینه به b و e معلوم نیست یعنی عبارتی برای f که بر حسب b_i و e_j باشد در دست نیست.

- قضیه 3.14 راهی را برای محاسبه مشتقات جزیی زیر به ما نشان می‌دهد.

$$\frac{\partial f}{\partial e_j} \text{ and } \frac{\partial f}{\partial b_i}$$

54/68

Theorem 4.7 Constraint Variation Sensitivity Theorem

قضیه 3.14 حساسیت تغییر قید

فرض کنید توابع $g_j(X)$, j=1 to m, $h_i(X)$, i=1 to p, $f(X)$ تا دو مرتبه مشتق پیوسته داشته باشند. همچنین فرض کنید X^* یک نقطه منظم باشد که همراه ضرایب v_j^* و u_j^* ، هم شرایط لازم کان تاکر و هم شرایط کافی قضیهٔ 3.12 را برای یک نقطهٔ مینیمم محلی منحصر به فرد برای مسئله‌ای که در معادلات زیر تعریف شده‌اند، برآورده کنند.

$$\min f(\mathrm{X})$$

$$h_i(\mathrm{X}) = 0; \quad \mathrm{i=1\ to\ p} \qquad (3.69)$$

$$g_j(\mathrm{X}) \le 0; \quad \mathrm{j=1\ to\ m} \qquad (3.70)$$

- اگر برای هر $g_j(X)=0$ درست باشد که بگوییم $u^*_j > 0$ ، آنگاه جواب $X^*(b,e)$ مسئلهٔ بهینه‌سازی فوق، که در معادله‌های (3.69) تا (3.70) تعریف شده، یک تابع پیوسته مشتق پذیر از *b* و *e* در همسایگی *b*=0 و *e*=0 خواهد بود. افزون بر این:

$$L = f + vh + ug, \mathrm{Now}: h = b,\ g \le e$$

$$\frac{\partial L}{\partial b} = \frac{\partial f}{\partial b} + v^* = 0 \rightarrow \frac{\partial f(x^*(0,0))}{\partial b_i} = -v_i^*,\ \mathrm{i=1\ to\ p} \qquad (3.71)$$

$$\frac{\partial L}{\partial e} = \frac{\partial f}{\partial e} + u^* = 0 \rightarrow \frac{\partial f(x^*(0,0))}{\partial e_j} = -u_j^*,\ j=1\ \mathrm{to\ m} \qquad (3.72)$$

- قضیه فقط برای قیود نامساوی که به شکل ($\le$) نوشته می‌شوند به کار می‌رود.

حدس تغییرات تابع هزینه با تغییر طرف راست قیود در همسایگی صفر با استفاده از قضیه 3.14:

بسط تیلور مرتبه اول تابع هزینه بر حسب b و e به صورت زیر است:

$$f(b_i, e_j) = f(0,0) + \frac{\partial f(0,0)}{\partial b_i} b_i + \frac{\partial f(0,0)}{\partial e_j} e_j$$

با جایگزینی معادله‌های (3.71) و (3,72) داریم :

$$f(b_i, e_j) = f(0,0) - v_i^* b_i - u_j^* e_j$$

$$\Delta f = f(b_i, e_j) - f(0,0) = -v_i^* b_i - u_j^* e_j$$

$$\Delta f = -\sum_i v_i^* b_i - \sum_j u_j^* e_j$$

- توجه:

اگر شرایط قضیه 3.14 برآورده نشود، وجود مشتق‌های ضمنی معادلات (3.71) و (3.72) با این قضیه نفی نمی‌شود. یعنی این مشتقات ممکن است وجود داشته باشند ولی با این قضیه وجودشان تضمین نمی‌شود.

- دلیل $u \geq 0$ برای قیود فعال نامساوی در نقطهٔ بهین:

با آزاد کردن قید فعال داریم:

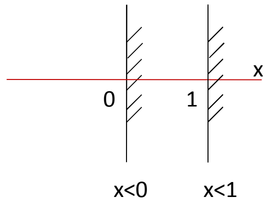
$$g_j = 0 \quad \to \quad g_j = e_j > 0$$

حداقل انتظار ما اینست که با آزاد سازی قید، جواب بهین بدون تغییر بماند و یا بهبود یابد (بدتر نشود).

if:

$u_j < 0 \, , e_j > 0$

$\Delta f = f(e_j) - f(0) = -u_j^* e_j > 0$



غیر قابل قبول است زیرا انتظار ما این است $f(e_j)$ کمتر از $f(0)$ یا حداقل مساوی با آن باشد.

در نتیجه ضریب لاگرانژ این قیود باید نامنفی باشد. $u_j \ge 0$

مثال: اثر تغییر حدود قید روی تابع هزینه

- مسئلهٔ ذکر شده در مثال 3.28 :

$$\min f(x_1,x_2) = x_1^2 + x_2^2 - 3x_1x_2$$
$$g(x_1,x_2) = x_1^2 + x_2^2 - 6 \le 0$$

یک نقطه که هم شرط لازم و هم شرط کافی را برآورده کند:

$$x_1^* = x_2^* = \sqrt{3} \, , u^* = \frac{1}{2} \, , f(X^*) = -3$$

- با تبدیل قید داریم:

$$if \quad g \le 0 \Rightarrow g \le e$$

- از طرفی با استفاده از قضیه 3.14 داریم:

$$\frac{\partial f}{\partial e} = -u^* = -\frac{1}{2}$$

- اگر e=1 باشد، آن‌گاه:

$$\Delta f = -\sum_i v_i^* b_i - \sum_j u_j^* e_j = -(0.5)\times(1) = -0.5$$

$$f = f(0,0) + \Delta f$$
$$f = -3 - 0.5 = -3.5$$

- اگر e=-1 باشد، آن‌گاه:

$$\Delta f = -(0.5)\times(-1) = 0.5$$
$$f = -3 + (0.5) = -2.5$$

اثر مقدار ضریب لاگرانژ

- هر چه ضریب لاگرانژ بزرگ‌تر باشد حساسیت بیشتر است یعنی هم آزاد کردن قید مزایای بیشتری دارد و هم زیان‌های ناشی از محدود کردن قید بیشتر است.
- این امر در مورد قیود فعال است.

اثر مقیاس بندی تابع هزینه روی ضرایب لاگرانژ

می‌دانیم که مقیاس بندی تابع هزینه، نقطهٔ بهین را تغییر نمی‌دهد ولی مقدار تابع هزینه مسلماً عوض می‌شود.

$$\begin{cases} f(X^*) \\ v_j^*,\ u_i^* \end{cases} \Rightarrow \begin{cases} \bar{f}(X) = Kf(X) \\ \bar{u}_i^*\ ,\ \bar{v}_j^* \end{cases} \quad K>0$$

آن‌گاه :

$$\bar{u}_i^* = Ku_i^* \ ; \quad i = 1 \text{ to } m$$
$$\bar{v}_j^* = Kv_j^* \ ; \quad j = 1 \text{ to } p$$

مثال: اثر مقیاس بندی تابع هزینه روی ضرایب لاگرانژ

- اثرات مقیاس بندی تابع هزینه را روی جواب بهین تابع زیر بررسی کنید.

$$\begin{cases} f(x_1,x_2) = x_1^2 + x_2^2 - 3x_1x_2 \\ g(x_1,x_2) = x_1^2 + x_2^2 - 6 \le 0 \end{cases}$$

یک نقطه که هم شرط لازم و هم شرط کافی را برآورده کند:

$$x_1^* = x_2^* = \sqrt{3}\ ,\ u^* = \frac{1}{2}\ ,\ f(X^*) = -3$$

- لاگرانژین مسئله عبارت است از:

$$L = K(x_1^2 + x_2^2 - 3x_1x_2) + \bar{u}(x_1^2 + x_2^2 - 6 + \bar{s}^2)$$

- شرایط لازم کان تاکر معادله‌های زیر را می‌دهد:

$$\begin{cases} \frac{\partial L}{\partial x_1} = 2Kx_1 - 3Kx_2 + 2\bar{u}x_1 = 0 \\ \frac{\partial L}{\partial x_2} = 2Kx_2 - 3Kx_1 + 2\bar{u}x_2 = 0 \\ x_1^2 + x_2^2 - 6 + \bar{s}^2 = 0 \\ \bar{u}\,\bar{s} = 0 \quad , \quad \bar{u} \geq 0 \end{cases}$$

- در اینجا، حالت $\bar{s} = 0$ نقطهٔ نامزد مینیمم را می‌دهد. با حل معادلات اسلاید قبلی داریم:

$$a)\ x_1^* = x_2^* = \sqrt{3} \quad , \quad \bar{u}^* = K/2 \ , \ \bar{f}(X^*) = -3K$$
$$b)\ x_1^* = x_2^* = -\sqrt{3} \ , \ \bar{u}^* = K/2 \ , \ \bar{f}(X^*) = -3K$$

- بنابراین مشاهده می‌کنیم که:

$$\bar{u}^* = Ku^*$$

اثر مقیاس بندی قید روی ضرایب لاگرانژ

- مقیاس بندی یک قید مرز قید را تغییر نمی‌دهد، بنابراین روی جواب بهین اثری نمی‌گذارد. فقط ضرایب لاگرانژ قید مقیاس بندی شده متأثر می‌شود.
- فرض کنید $P_i\ , M_j > 0$ دو پارامتر مقیاس برای ***j*** امین قید نامساوی و ***i*** امین قید مساوی باشند، آن‌گاه:

$$\bar{u}_j^* = u_j^* / M_j$$

$$\bar{v}_i^* = v_i^* / P_i$$

مسائل زیر را حل کرده و تا دو هفته دیگر تحویل فرمایید:

3) 112, 113, 114

تمرین‌های قبلی:

3) 1,8,16, 17, 29, 30, 46, 61, 69, 84, 93