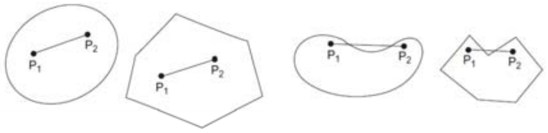
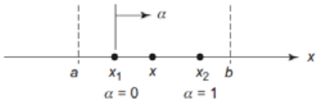
Lecture 08: Convexity





Global Optimality and Convexity

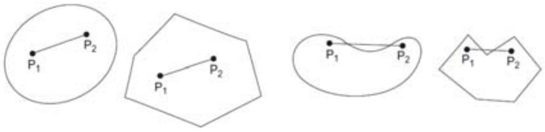
- If the optimization problem can be shown to be convex, then any local minimum is also a global minimum.
- Also the KKT necessary conditions are sufficient for the minimum point.

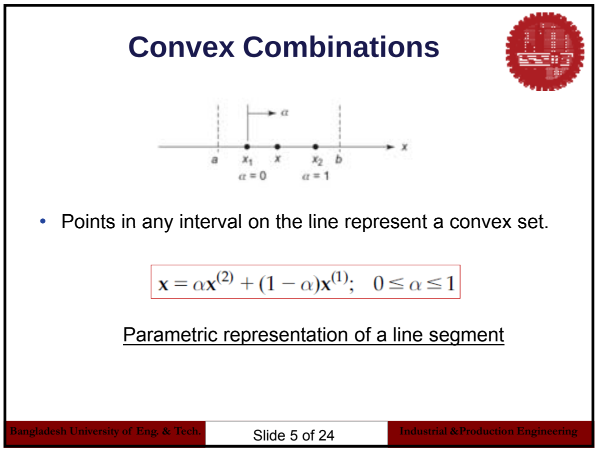
Related Terms

- Convexity
- Convex programming problems
- Convex sets
- Convex functions

Convex Sets

- A convex set *S* is a collection of points (vectors ***x***) having the following property:
  - If P_1 and P_2 are any points in *S*, then the entire line segment P_1 - P_2 is also in *S*.
  - This is a necessary and sufficient condition for convexity of the set *S*.



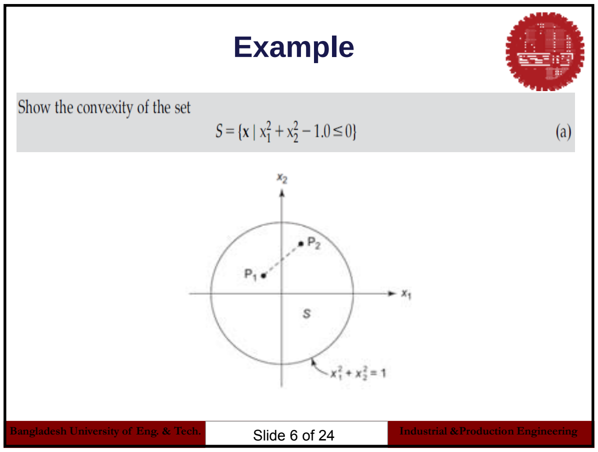


Convex Combinations

- Points in any interval on the line represent a convex set.

$$\boxed{\mathbf{x} = \alpha\mathbf{x}^{(2)} + (1-\alpha)\mathbf{x}^{(1)}; \quad 0 \le \alpha \le 1}$$

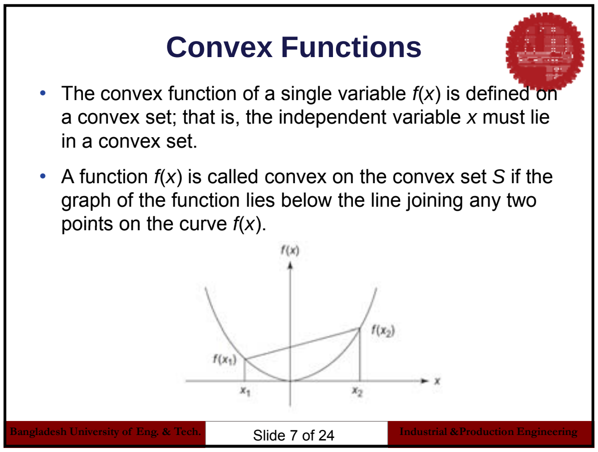
Parametric representation of a line segment



Example

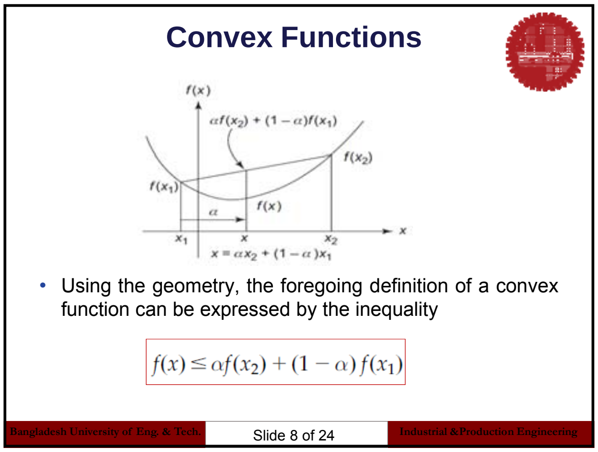
Show the convexity of the set

$$S = \{\mathbf{x} \mid x_1^2 + x_2^2 - 1.0 \le 0\} \qquad \text{(a)}$$



Convex Functions

- The convex function of a single variable *f*(*x*) is defined on a convex set; that is, the independent variable *x* must lie in a convex set.
- A function *f*(*x*) is called convex on the convex set *S* if the graph of the function lies below the line joining any two points on the curve *f*(*x*).



Convex Functions

- Using the geometry, the foregoing definition of a convex function can be expressed by the inequality

$$\boxed{f(x) \le \alpha f(x_2) + (1-\alpha) f(x_1)}$$

Convex Functions

$$f(x) \le \alpha f(x_2) + (1-\alpha) f(x_1)$$

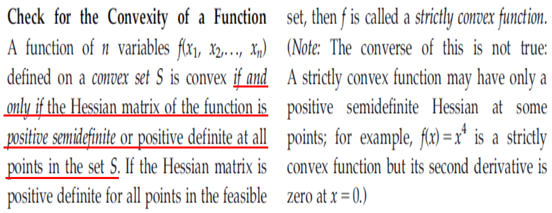
$$x = \alpha x_2 + (1-\alpha)x_1$$

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

- The foregoing definition of a convex function of one variable can be generalized to functions of *n* variables.

$$f(\alpha \mathbf{x}^{(2)} + (1-\alpha)\mathbf{x}^{(1)}) \le \alpha f(\mathbf{x}^{(2)}) + (1-\alpha)f(\mathbf{x}^{(1)}) \quad \text{for} \quad 0 \le \alpha \le 1$$

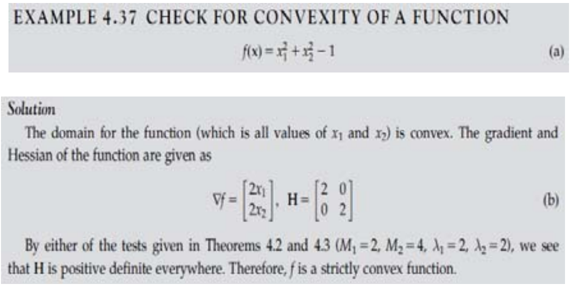
Theorem: Check for Convexity of a Function



Check for the Convexity of a Function A function of *n* variables $f(x_1, x_2, \ldots, x_n)$ 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:* The converse of this is not true: A strictly convex function may have only a positive semidefinite Hessian at some points; for example, $f(x) = x^4$ is a strictly convex function but its second derivative is zero at $x = 0$.)

In one dimension, the convexity check of the theorem reduces to the condition that the second derivative (curvature) of the function be non-negative.

Example



EXAMPLE 4.37 CHECK FOR CONVEXITY OF A FUNCTION

$$f(\mathbf{x}) = x_1^2 + x_2^2 - 1 \qquad \text{(a)}$$

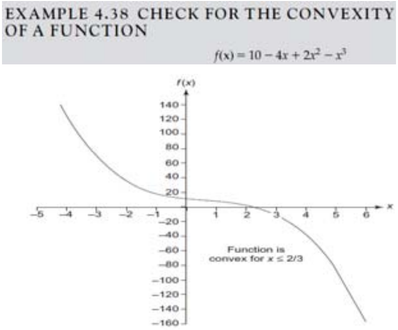
Solution

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

$$\nabla f = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad \text{(b)}$$

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.

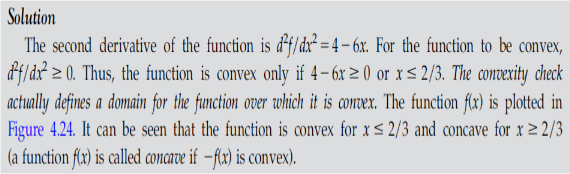
Example



EXAMPLE 4.38 CHECK FOR THE CONVEXITY OF A FUNCTION

$$f(\mathbf{x}) = 10 - 4x + 2x^2 - x^3$$

Example



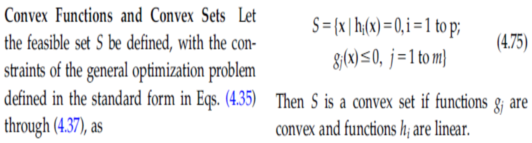
Solution

The second derivative of the function is $d^2f/dx^2 = 4 - 6x$. For the function to be convex, $d^2f/dx^2 \ge 0$. Thus, the function is convex only if $4 - 6x \ge 0$ or $x \le 2/3$. *The convexity check actually defines a domain for the function over which it is convex.* The function $f(x)$ is plotted in Figure 4.24. It can be seen that the function is convex for $x \le 2/3$ and concave for $x \ge 2/3$ (a function $f(x)$ is called *concave* if $-f(x)$ is convex).

Convex Programming Problem

- If a function $g_i(\mathbf{x})$ is convex, then the set $g_i(\mathbf{x}) \le e_i$ is convex, where e_i is any constant.
- If functions $g_i(\mathbf{x})$ for $i = 1$ to m are convex, then the set defined by $g_i(\mathbf{x}) \le e_i$ $i = 1$ to m is also convex.
- The set $g_i(\mathbf{x})$ for $i = 1$ to m is the intersection of sets defined by the individual constraints $g_i(\mathbf{x}) \le e_i$.
- Therefore, the intersection of convex sets is itself a convex set.

Theorem: Convex Functions and Convex Sets



Convex Functions and Convex Sets Let the feasible set S be defined, with the constraints of the general optimization problem defined in the standard form in Eqs. (4.35) through (4.37), as

$$S = \{x \mid h_i(x) = 0, i = 1 \text{ to } p; \quad g_j(x) \le 0, \; j = 1 \text{ to } m\} \tag{4.75}$$

Then S is a convex set if functions g_j are convex and functions h_i are linear.

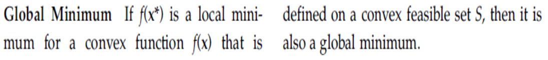
Convex Functions and Convex Sets

- If all inequality constraint functions for an optimum design problem are convex, and all equality constraint are linear, then the feasible set S is convex.
- If the cost function is also convex over the set S, then we have what is known as a convex programming problem.
- Such problems have a very useful property, which is that KKT necessary conditions are also sufficient and any local minimum is also a global minimum.

Observations Regarding the Theorem

- The theorem does not say that the feasible set *S* cannot be convex if a constraint function fails the convexity check (i.e., it is not an “if and only if” theorem).
- There are some problems where the constraint functions fail the convexity check, but the feasible set is still convex.
- Thus, <u>the conditions of the theorem are only sufficient but not necessary for the convexity of the problem</u>.

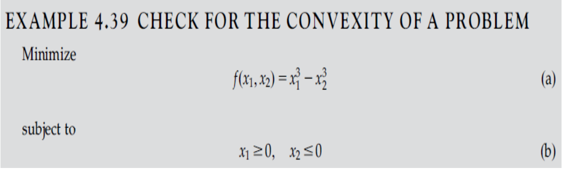
Theorem: Global Minimum



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

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

Example



EXAMPLE 4.39 CHECK FOR THE CONVEXITY OF A PROBLEM

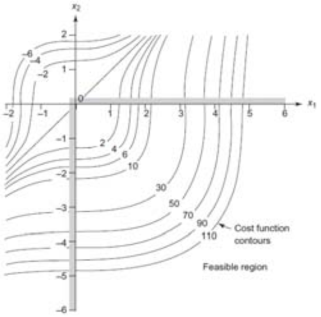
Minimize

$$f(x_1, x_2) = x_1^3 - x_2^3 \qquad \text{(a)}$$

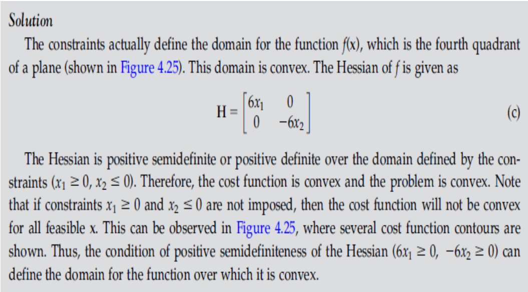
subject to

$$x_1 \geq 0, \quad x_2 \leq 0 \qquad \text{(b)}$$

Example



Example



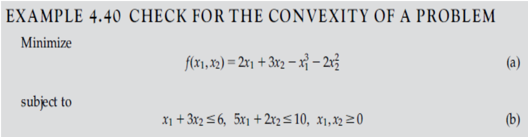
Solution

The constraints actually define the domain for the function $f(\mathbf{x})$, which is the fourth quadrant of a plane (shown in Figure 4.25). This domain is convex. The Hessian of f is given as

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

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. Note that if constraints $x_1 \geq 0$ and $x_2 \leq 0$ are not imposed, then the cost function will not be convex for all feasible $\mathbf{x}$. This can be observed in Figure 4.25, where several cost function contours are shown. Thus, the condition of positive semidefiniteness of the Hessian ($6x_1 \geq 0$, $-6x_2 \geq 0$) can define the domain for the function over which it is convex.

Example



EXAMPLE 4.40 CHECK FOR THE CONVEXITY OF A PROBLEM

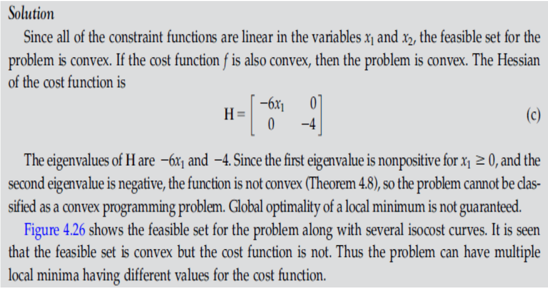
Minimize

$$f(x_1, x_2) = 2x_1 + 3x_2 - x_1^3 - 2x_2^2 \tag{a}$$

subject to

$$x_1 + 3x_2 \leq 6, \quad 5x_1 + 2x_2 \leq 10, \quad x_1, x_2 \geq 0 \tag{b}$$

Example



Solution

Since all of the constraint functions are linear in the variables x_1 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

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

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. Global optimality of a local minimum is not guaranteed.

Figure 4.26 shows the feasible set for the problem along with several isocost curves. It is seen that the feasible set is convex but the cost function is not. Thus the problem can have multiple local minima having different values for the cost function.

Example