Non-Linear Programming

Fatih Cavdur

fatihcavdur@uludag.edu.tr

Differential Calculus

Limit

$$\boxed{\lim_{x \to a} f(x) = c}$$

Continuity

A function $f(x)$ is continuous at point a if

$$\boxed{\lim_{x \to a} f(x) = f(a)}$$

Differential Calculus

Differentiation

The derivative of a function $f(x)$ at $x = a$ is defined as

$$\boxed{f'(a) = \lim_{\Delta x \to 0} \frac{f(x) - f(a)}{\Delta x}}$$

If $f'(a) > 0$, then, $f(x)$ is increasing at $x = a$.

If $f'(a) < 0$, then, $f(x)$ is decreasing at $x = a$.

Differential Calculus

Taylor Series Expansion

The nth order Taylor series expansion of a function of $f(x)$ about a (given that $f^{(n+1)}(x)$ exists for all points on the interval $[a, b]$), for any h such that $0 \leq h \leq b - a$, is defined as follows where it holds for some number p between a and $a + h$:

$$\boxed{f(a+h) = f(a) + \sum_{i=1}^{n} \frac{f^{(i)}(a)}{i!} h^i + \frac{f^{(n+1)}(p)}{(n+1)!} h^{n+1}}$$

Differential Calculus

Example:

Find the Taylor series expansion of e^{-x} about $x = 0$.

Since $f^{'}(x) = -e^{-x}$ and $f^{''}(x) = e^{-x}$, the Taylor series expansion will hold for $[0, b]$. Also, since $f(0) = 0$, $f^{'}(0) = -1$ and $f^{''}(x) = e^{-x}$, we have

$$f(h) = e^{-h} = 1 - h + \frac{h^2 e^{-p}}{2}, \quad p \in [0, h]$$

Differential Calculus

Partial Derivatives

The partial derivative of function $f(x_1, x_2, \ldots, x_n)$ with respect to x_i is defined as

$$\boxed{\frac{\partial f}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{f(x_1, \ldots, x_i + \Delta x_i, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{\Delta x_i}}$$

Introduction to NLP

A feasible region for an NLP is the set of points $(x_1, x_2, \ldots, x_n)$ that satisfy the constraint equations of the NLP. Any point $\bar{x}$ in the feasible region for which $f(\bar{x}) \geq f(x)$ holds for $\forall x$ in the feasible region is an optimal solution to the NLP (for a max problem).

Introduction to NLP

Example:

It costs c dollars per unit to manufacture a product. If the manufacturer charges p dollars per unit for the product, customer demand dp units. What price should be charged to maximize profit?

We have the following unconstrained NLP:

$$\max(p - c)dp$$

Introduction to NLP

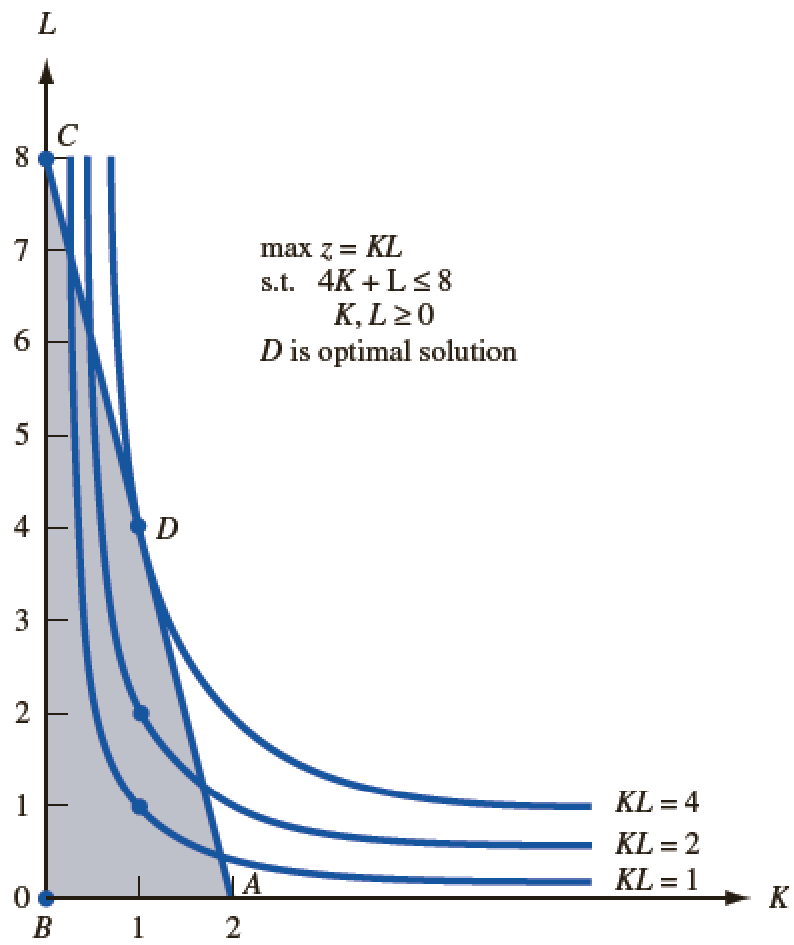
Example:

If C units of capital and L units of labor are used, a company can produce KL units of a manufactured good. Capital can be purchased at $4/unit and labor can be purchased at $1/unit. A total of $8 is available to purchase capital and labor. How can the firm maximize the quantity of the good that can be manufactured?

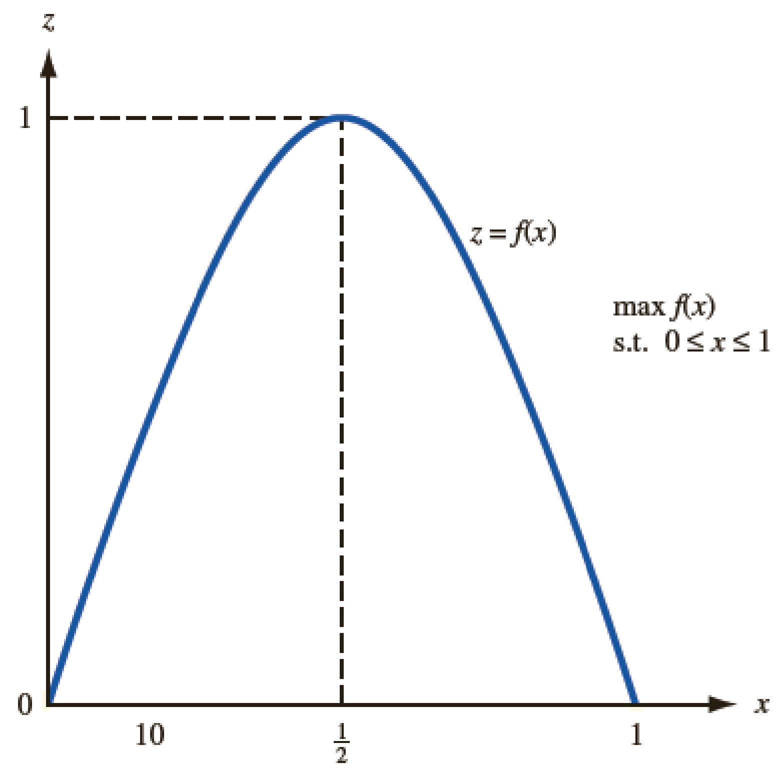
We have the following NLP:

$$\max z = KL \qquad \text{S.T.} \qquad 4K + L \leq 8; K, L \geq 0$$

Introduction to NLP



Introduction to NLP



Introduction to NLP

Local Extremum

For an NLP (max), a feasible point $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is a local maximum if for sufficiently small ε, $\mathbf{x}' = (x_1{}', x_2{}', \dots, x_n{}')$ having $|x_i - x_i{}'| < \varepsilon$ for $\forall i$ satisfies $f(\mathbf{x}) \geq f(\mathbf{x}')$.

Introduction to NLP

Example:

Truckco is trying to determine where it should locate a single warehouse. The positions in the $x - y$ plane (in miles) of four customers and the number of shipments made annually to each customer are given in the below table. Truckco wants to locate the warehouse to minimize the total distance trucks must travel annually from the warehouse to the four customers.

Customer	x	y	# of Shipments
1	5	10	200
2	10	5	150
3	0	12	200
4	0	0	300

Introduction to NLP

We let

- $x = x$-coordinate of the warehouse
- $y = y$-coordinate of the warehouse
- d_i = distance from customer i to the warehouse

We then have

$$\min z = 200d_1 + 150d_2 + 200d_3 + 300d_4$$

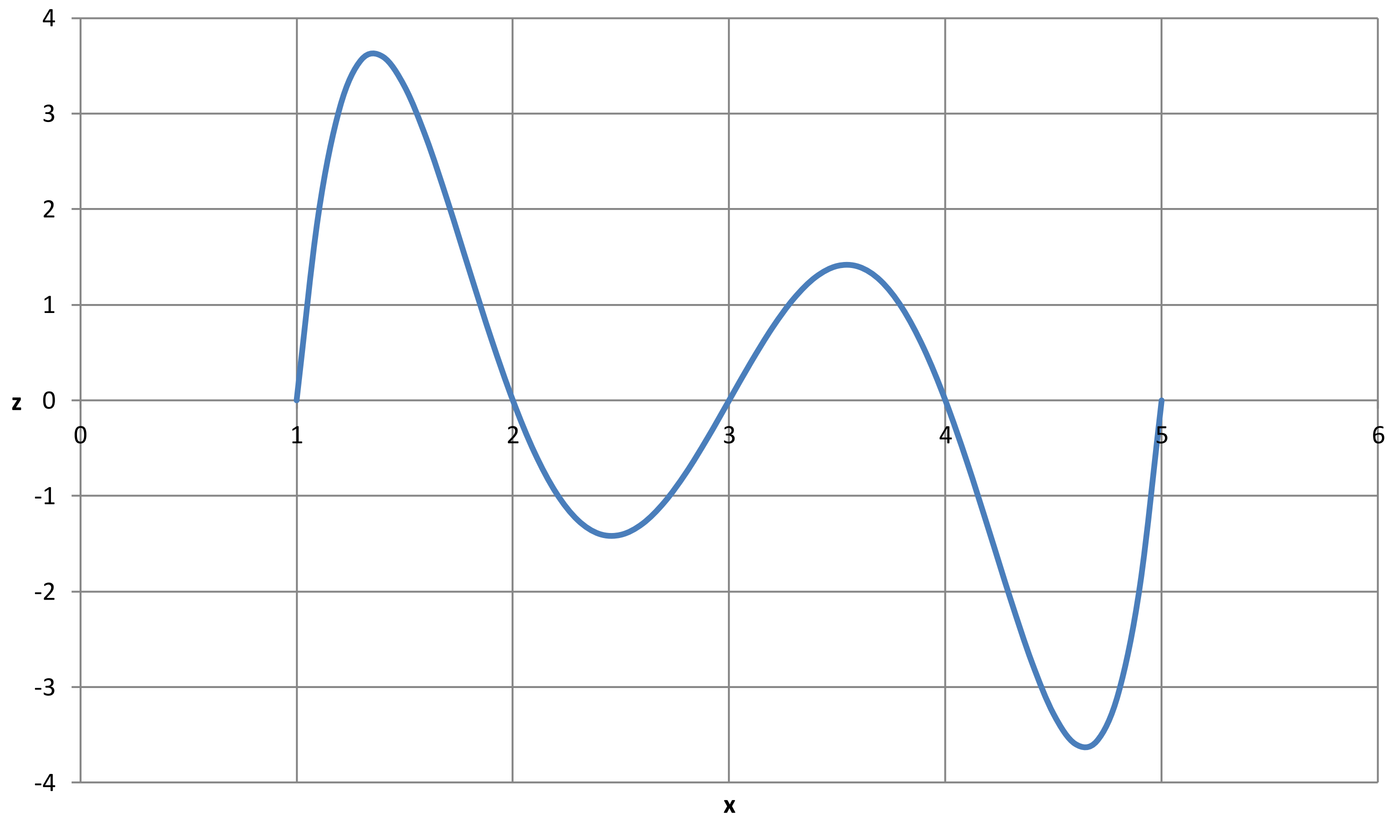
$$d_1 = \sqrt{(x-5)^2 + (y-10)^2}$$
$$d_2 = \sqrt{(x-10)^2 + (y-5)^2}$$
$$d_3 = \sqrt{x^2 + (y-12)^2}$$
$$d_4 = \sqrt{(x-12)^2 + y^2}$$

$z = 5{,}456.540$; $x = 9.314$, $y = 5.029$, $d_1 = 6.582$, $d_2 = 0.686$, $d_3 = 11.634$, $d_4 = 5.701$

Introduction to NLP

Example:

$$\max z = (x-1)(x-2)(x-3)(x-4)(x-5),\ x \geq 1;\ x \leq 5$$



Convex and Concave Functions

Convex and Concave Functions

A function $f(x_1, x_2, \dots, x_n)$ is a convex function on a convex set S if, for any $\mathbf{x}' \in S$ and $\mathbf{x}'' \in S$, and for $0 \le \lambda \le 1$, the following expression holds:

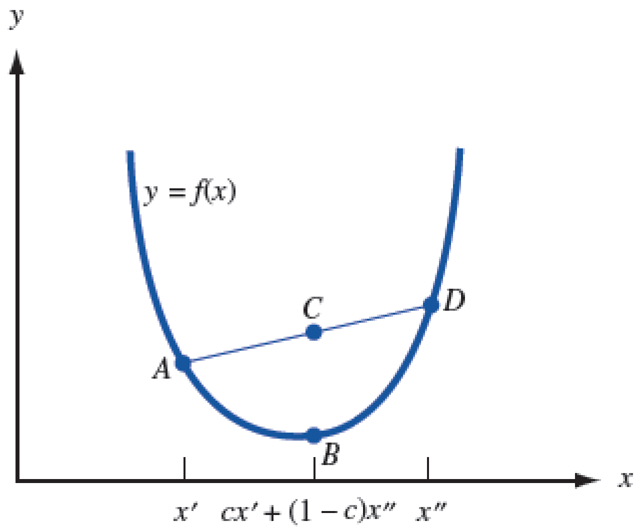
$$\boxed{f[\lambda \mathbf{x}' + (1-\lambda)\mathbf{x}''] \le \lambda f(\mathbf{x}') + (1-\lambda) f(\mathbf{x}'')}$$

A function $f(x_1, x_2, \dots, x_n)$ is a concave function on a convex set S if, for any $\mathbf{x}' \in S$ and $\mathbf{x}'' \in S$, and for $0 \le \lambda \le 1$, the following expression holds:

$$\boxed{f[\lambda \mathbf{x}' + (1-\lambda)\mathbf{x}''] \ge \lambda f(\mathbf{x}') + (1-\lambda) f(\mathbf{x}'')}$$

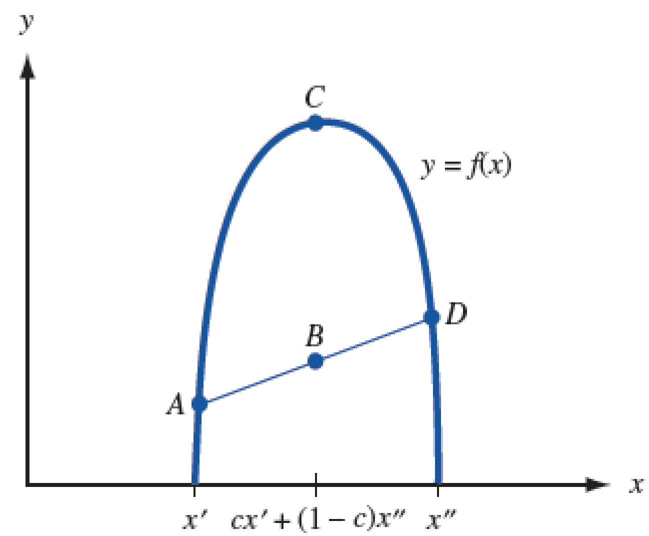
Convex and Concave Functions

A Convex Function



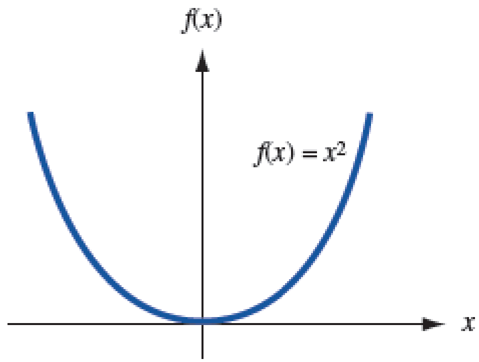
Point $A = (x', f(x'))$
Point $D = (x'', f(x''))$
Point $C = (cx' + (1 - c)x'', cf(x') + (1 - c)f(x''))$
Point $B = (cx' + (1 - c)x'', f(cx' + (1 - c)x''))$
From figure: $f(cx' + (1 - c)x'') \leq cf(x') + (1 - c)f(x'')$

A Concave Function

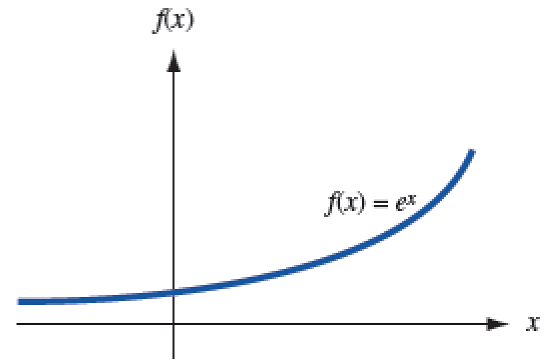


Point $A = (x', f(x'))$
Point $D = (x'', f(x''))$
Point $C = (cx' + (1 - c)x'', f(cx' + (1 - c)x''))$
Point $B = (cx' + (1 - c)x'', cf(x') + (1 - c)f(x''))$
From figure: $f(cx' + (1 - c)x'') \geq cf(x') + (1 - c)f(x'')$

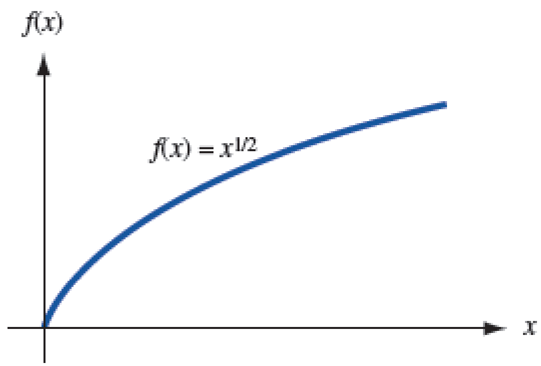
Convex and Concave Functions



a Convex

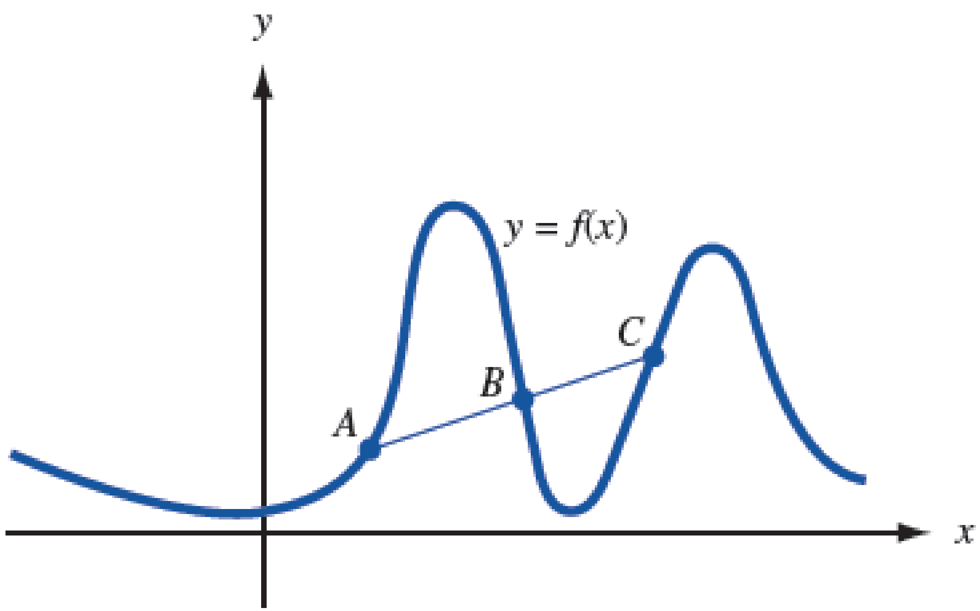


b Convex



c Concave

Convex and Concave Functions



Convex and Concave Functions

Theorem:

The sum of two convex (concave) functions is also convex (concave).

Theorem:

A linear function is both convex and concave.

Theorem:

If a feasible region S for a max NLP is a convex set, and if $f(\mathbf{x})$ is concave on S, then, any local maximum for the NLP is also the global maximum (the optimal solution).

Convex and Concave Functions

Theorem:

If a feasible region S for a min NLP is a convex set, and if $f(\mathbf{x})$ is convex on S, then, any local minimum for the NLP is also the global minimum (the optimal solution).

Theorem:

If $f''(\mathbf{x})$ exists for all $\mathbf{x}$ in a convex set S, then, $f(\mathbf{x})$ is a convex function on S iif $f''(\mathbf{x}) \geq 0$ for $\forall \mathbf{x} \in S$.

Theorem:

If $f''(\mathbf{x})$ exists for all $\mathbf{x}$ in a convex set S, then, $f(\mathbf{x})$ is a concave function on S iif $f''(\mathbf{x}) \leq 0$ for $\forall \mathbf{x} \in S$.

Convex and Concave Functions

Examples:

$$f(x) = x^2 \text{ is convex on } S = R^1 \text{ since } f''(x) = 2 \geq 0$$

$$f(x) = e^x \text{ is convex on } S = R^1 \text{ since } f''(x) = e^x \geq 0$$

$$f(x) = \sqrt{x} \text{ is concave on } S = (0, \infty) \text{ since } f''(x) = -\frac{x^{-3/2}}{4} \leq 0$$

$$f(x) = ax + b \text{ is both convex and concave on } S = R^1 \text{ since } f''(x) = 0$$

Convex and Concave Functions

The Hessian of $f(x_1, x_2, \dots, x_n)$ is the $n \times n$ matrix whose (i, j)th entry is given by

$$h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

An ith principal minor of an $n \times n$ matrix is the determinant of any $i \times i$ matrix obtained by deleting $n - i$ rows and the corresponding $n - i$ columns of the matrix.

The kth leading principal minor of an $n \times n$ matrix is the determinant of the $k \times k$ matrix obtained by deleting the last $n - k$ rows and columns of the matrix.

Convex and Concave Functions

Theorem:

Suppose $f(x_1, x_2, \dots, x_n)$ has continuous second-order partial derivatives for $\forall \mathbf{x} \in S$. Then, $f(x_1, x_2, \dots, x_n)$ is a convex function on S iif for $\forall \mathbf{x} \in S$, all principal minors of $\mathbf{H}$ are non-negative.

Theorem:

Suppose $f(x_1, x_2, \dots, x_n)$ has continuous second-order partial derivatives for $\forall \mathbf{x} \in S$. Then, $f(x_1, x_2, \dots, x_n)$ is a concave function on S iif for $\forall \mathbf{x} \in S$ and $k = 1,2, \dots, n$, all non-zero principal minors of $\mathbf{H}$ have the same sign as $(-1)^k$.

Convex and Concave Functions

Example:

$$f(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2$$

$$H(x_1, x_2) = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

Note that the first principal minors are both 2 and the second principal minor is 2 (2) – 2 (2) = 0

Since all principal minors are nonnegative, the function is convex on R^2.

Convex and Concave Functions

Example:

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + 2x_3^2 - x_1x_2 - x_1x_3 - x_2x_3$$

$$H(x_1, x_2, x_3) = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 4 \end{bmatrix}$$

Note that the first principal minors are 4, 2 and 2 and the second principal minor is

$$\det\begin{vmatrix} 2 & -1 \\ -1 & 4 \end{vmatrix} = 7, \det\begin{vmatrix} 2 & -1 \\ -1 & 4 \end{vmatrix} = 7, \det\begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3$$

Finally, the third principal minor is

$$\det\begin{vmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 4 \end{vmatrix} = 6$$

Since all principal minors are nonnegative, the function is convex on R^3.

NLPs with 1 Variable

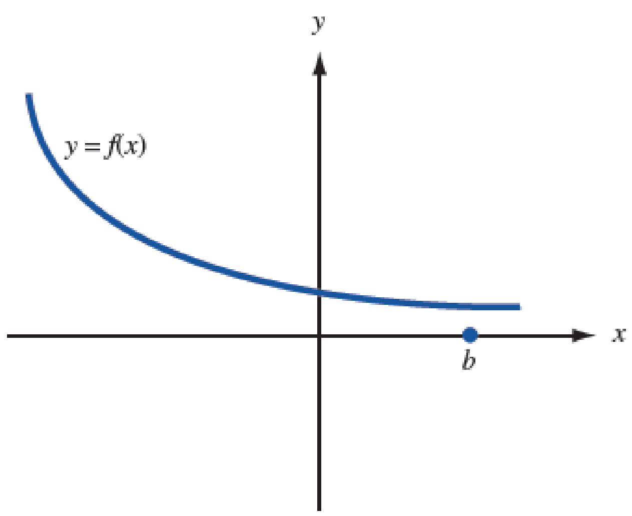
We assume that we have the following LP:

$$\max f(x),\ x \in [a, b]$$

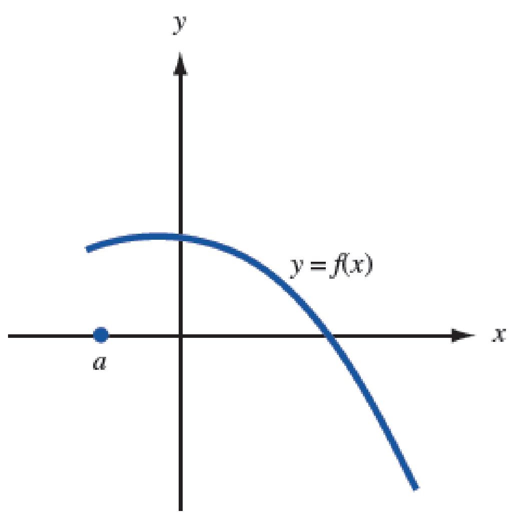
There are 3 types of points for which we can have local max or min points (extremum candidates):

Case 1. Points where $a \leq x \leq b$ and $f'(x) = 0$ (stationary point)
Case 2. Points where $f'(x)$ does not exist.
Case 3. Endpoints of the interval $[a, b]$.

NLPs with 1 Variable: Case 1

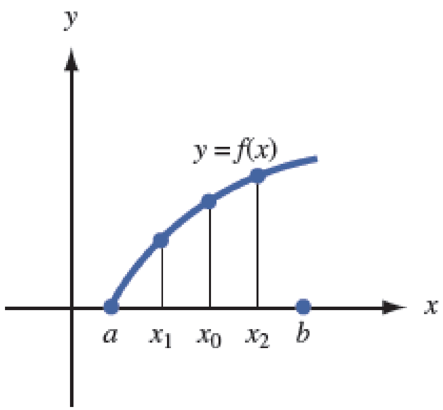


a max $f(x)$
s.t. $x \in (-\infty, b]$

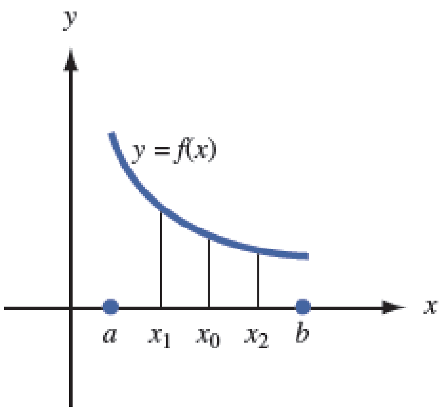


b max $f(x)$
s.t. $x \in [a, \infty)$

NLPs with 1 Variable: Case 1



a $f'(x_0) > 0$
$f(x_1) < f(x_0)$
$f(x_2) > f(x_0)$
x_0 not a local extremum

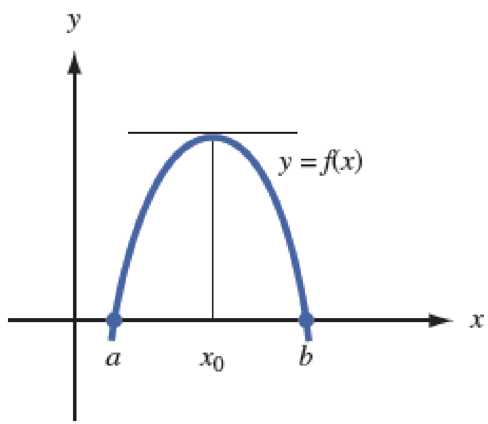


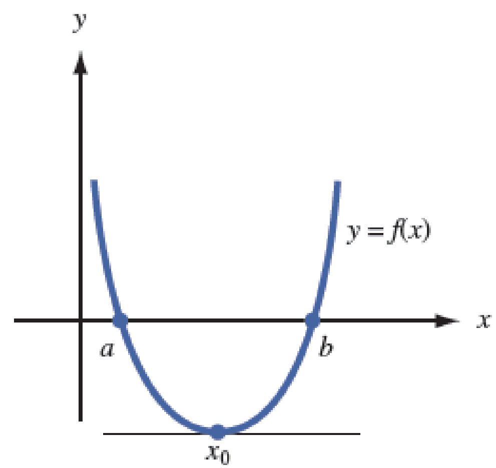
b $f'(x_0) < 0$
$f(x_1) > f(x_0)$
$f(x_2) < f(x_0)$
x_0 not a local extremum

NLPs with 1 Variable: Case 1

a $f'(x_0) > 0$
$f(x_1) < f(x_0)$
$f(x_2) > f(x_0)$
x_0 not a local extremum

b $f'(x_0) < 0$
$f(x_1) > f(x_0)$
$f(x_2) < f(x_0)$
x_0 not a local extremum

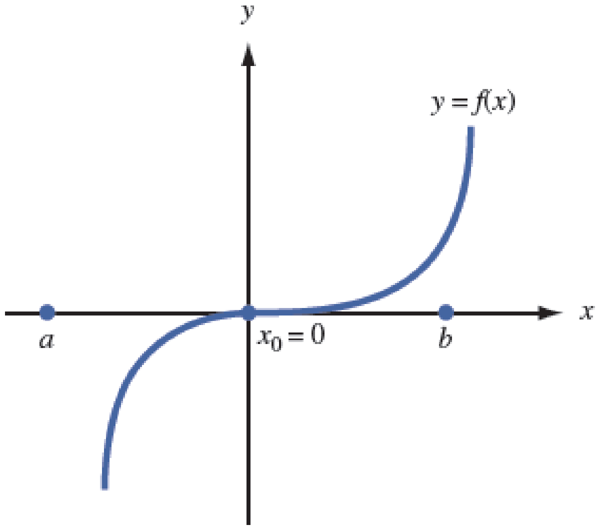




c $f'(x_0) = 0$
For $x < x_0$, $f'(x) > 0$
For $x > x_0$, $f'(x) < 0$
x_0 is a local maximum

d $f'(x_0) = 0$
For $x < x_0$, $f'(x) < 0$
For $x > x_0$, $f'(x) > 0$
x_0 is a local maximum

NLPs with 1 Variable: Case 1



e $x_0 = 0$ not a local maximum
or a local minimum
but $f'(x_0) = 0$

NLPs with 1 Variable: Case 1

Theorem:

If $f'(x_0) = 0$ and $f''(x_0) < 0$, then, x_0 is a local maximum.
If $f'(x_0) = 0$ and $f''(x_0) > 0$, then, x_0 is a local minimum.

If both derivatives are zero, we can use the following theorem:

Theorem:

If $f'(x_0) = 0$, and

- if the first non-zero derivative at x_0 is an odd-order derivative, then, x_0 is a not a local max or min.
- if the first non-zero derivative at x_0 is positive and an even-order derivative, then, x_0 is a local min.
- if the first non-zero derivative at x_0 is negative and an even-order derivative, then, x_0 is a local max.

NLPs with 1 Variable: Case 2

If $f'(x)$ does not exist, we can use the followings to determine if x_0 is a local min or max based on the relationships between $f(x_0)$, $f(x_1)$ and $f(x_2)$:

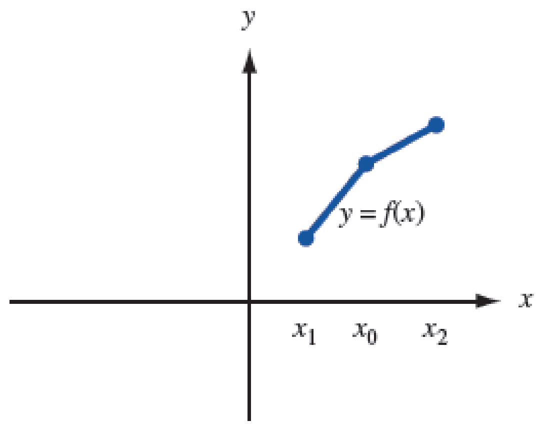
$$f(x_0) > f(x_1) \text{ and } f(x_0) < f(x_2) \Rightarrow x_0 \text{ not a local extremum (fig a)}$$

$$f(x_0) < f(x_1) \text{ and } f(x_0) > f(x_2) \Rightarrow x_0 \text{ not a local extremum (fig b)}$$

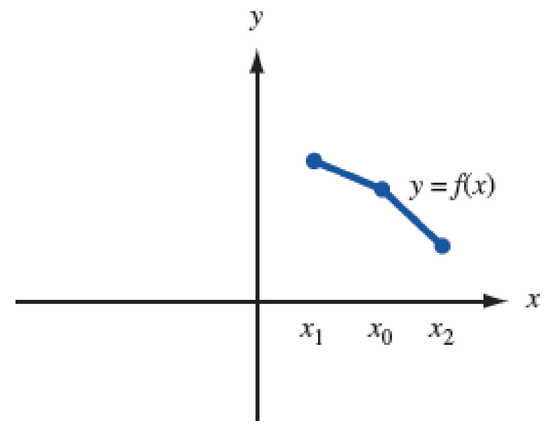
$$f(x_0) \geq f(x_1) \text{ and } f(x_0) \geq f(x_2) \Rightarrow x_0 \text{ a local max (fig c)}$$

$$f(x_0) \leq f(x_1) \text{ and } f(x_0) \leq f(x_2) \Rightarrow x_0 \text{ a local min (fig d)}$$

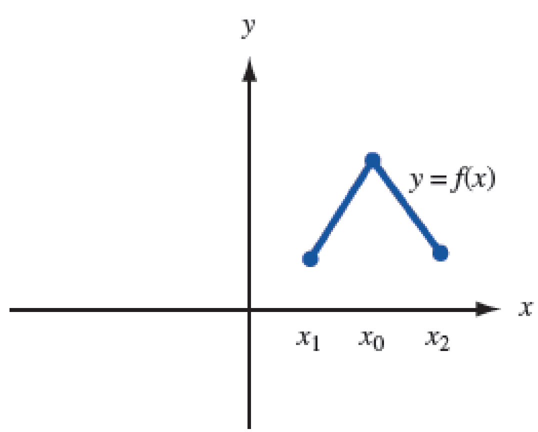
NLPs with 1 Variable: Case 2



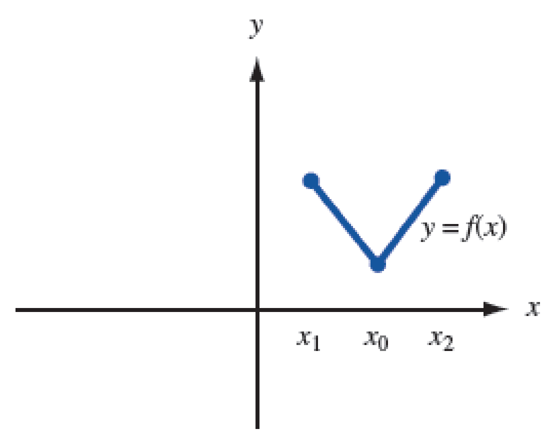
a x_0 not a local extremum



b x_0 not a local extremum



c x_0 is a local maximum



d x_0 is a local minimum

fatihcavdur@uludag.edu.tr

NLPs with 1 Variable: Case 3

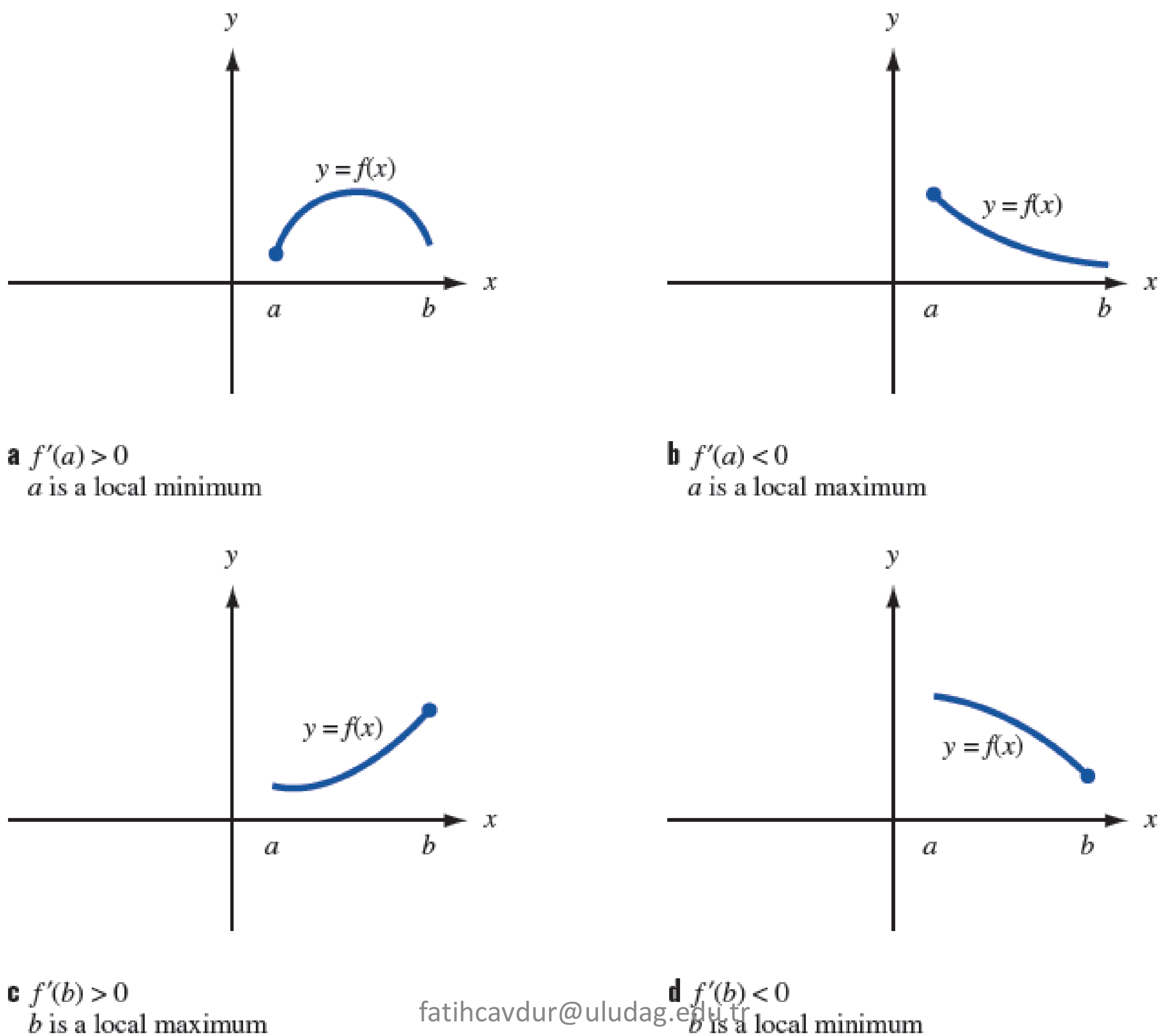
$f'(a) > 0 \Rightarrow a$ is a local min

$f'(a) < 0 \Rightarrow a$ is a local max

$f'(b) > 0 \Rightarrow b$ is a local max

$f'(b) < 0 \Rightarrow b$ is a local min

NLPs with 1 Variable: Case 3



a $f'(a) > 0$
a is a local minimum

b $f'(a) < 0$
a is a local maximum

c $f'(b) > 0$
b is a local maximum

d $f'(b) < 0$
b is a local minimum

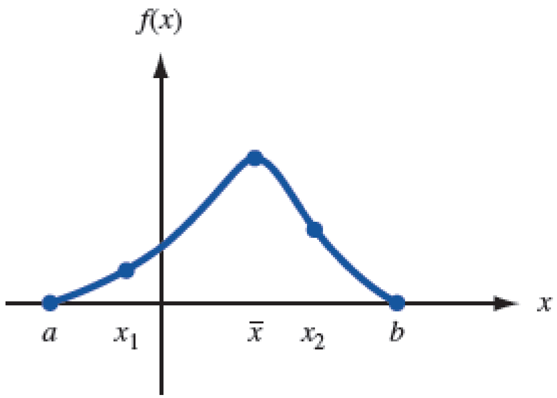
Golden Section Search

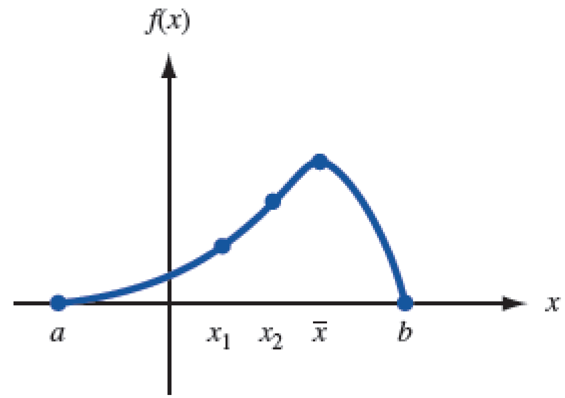
Suppose that we have the following NLP:

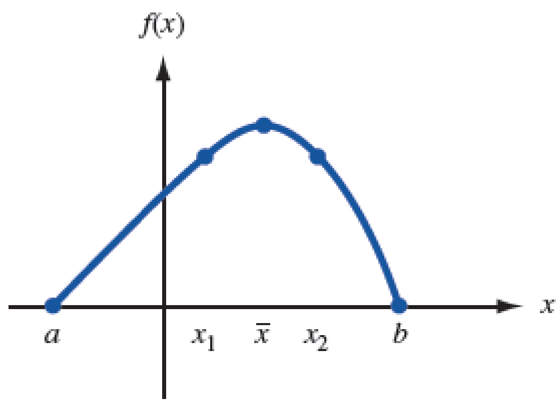
$$\max f(x),\ a \leq x \leq b$$

If $f'(x)$ does not exist for some x, or if it may be difficult to solve $f'(x) = 0$. In this case we can use a new approach if $f(x)$ is a unimodal function.

Golden Section Search

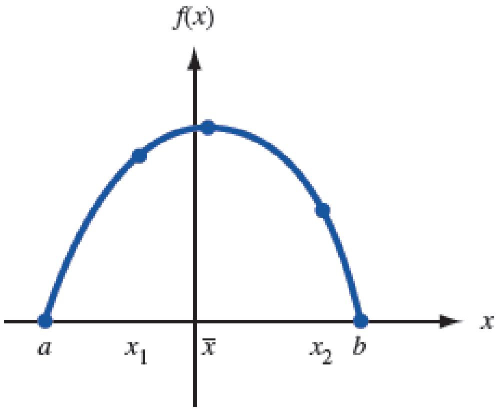


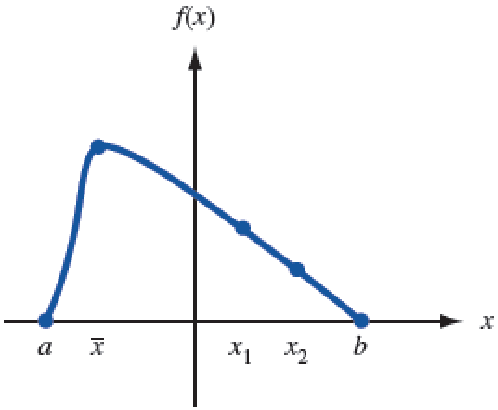




fatihcavdur@uludag.edu.tr

Golden Section Search





Golden Section Search

A function $f(x)$ is unimodal on $[a, b]$ if for some $\bar{x} \in [a, b]$, $f(x)$ is strictly increasing on $[a, \bar{x}]$ and strictly decreasing on $[\bar{x}, b]$.

If $f(x)$ is unimodal on $[a, b]$, then, $f(x)$ will have only one local maximum $\bar{x}$ on $[a, b]$ and that local maximum will solve the NLP. By evaluating $f(x)$ at two points x_1 and x_2 on $[a, b]$ where $x_1 < x_2$, we may reduce the size of the interval in which the solution to the NLP must lie. After evaluation of the function, 3 cases might be possible:

Case 1: $f(x_1) < f(x_2)$

Case 2: $f(x_1) = f(x_2)$

Case 3: $f(x_1) > f(x_2)$

Golden Section Search

Case 1: If $f(x_1) < f(x_2)$, since the function is unimodal the optimal solution cannot be on $[a, x_1]$. So, we have

$$f(x_1) < f(x_2) \Rightarrow \bar{x} \in (x_1, b]$$

Case 2: If $f(x_1) = f(x_2)$, since the function is unimodal the optimal solution must have $\bar{x} < x_2$. So, we have

$$f(x_1) = f(x_2) \Rightarrow \bar{x} \in [a, x_2]$$

Case 3: If $f(x_1) > f(x_2)$, since the function is unimodal the optimal solution must have $\bar{x} < x_2$. So, we have

$$f(x_1) > f(x_2) \Rightarrow \bar{x} \in [a, x_2]$$

The interval in which $\bar{x}$ lies is called the interval of uncertainty. We can use a search algorithm to find $\bar{x}$.

Search Algorithm

Step 1: Begin with $[a, b]$. Evaluate $f(x)$ at two judiciously points x_1 and x_2.

Step 2: Determine the case (1, 2 or 3) and reduce the interval.

Step 3: Evaluate $f(x)$ at two new points. Return to Step 2 unless the interval is small enough.

Golden Section Search Algorithm

We find r as the root of the equation:

$$r^2 + r - 1 = 0 \Rightarrow r = \frac{\sqrt{5} - 1}{2} = 0.618$$

Golden Section Search starts with points

$$x_1 = b - r(b - a)$$

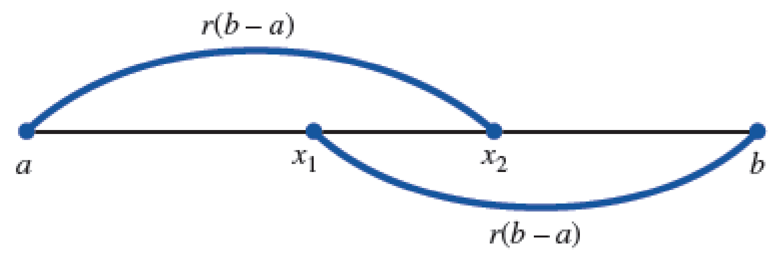
$$x_2 = a + r(b - a)$$

Each time $f(x)$ is evaluated the interval of uncertainty is reduced, an iteration of the Golden Section Search is completed.

$L_k =$ the length of the interval of uncertainty after k iterations

$I_k =$ the interval of uncertainty after k iterations

Golden Section Search Algorithm

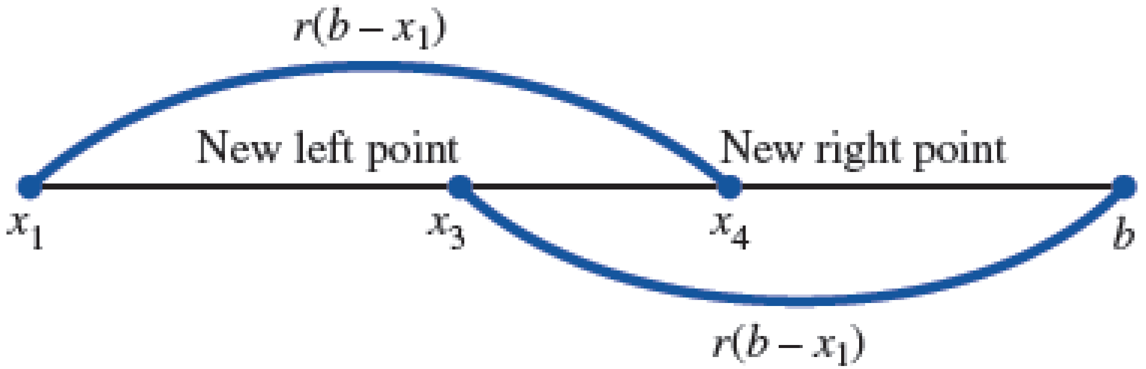


Golden Section Search Algorithm

If $f(x_1) < f(x_2)$, then

$$x_3 = b - r(b - x_1) = b - r^2(b - a)$$

$$x_4 = x_1 + r(b - x_1)$$



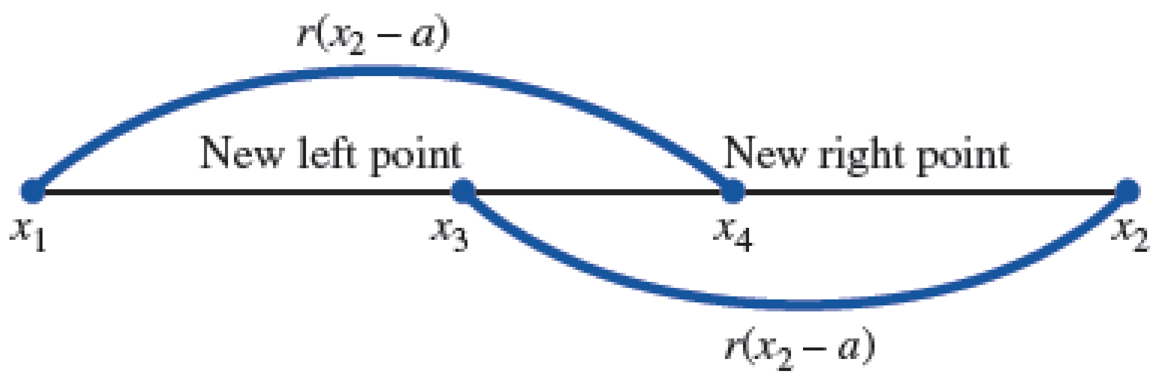
a If $f(x_1) < f(x_2)$, new interval of uncertainty is $(x_1, b]$

Golden Section Search Algorithm

If $f(x_1) \geq f(x_2)$, then

$$x_3 = x_2 - r(x_2 - a)$$

$$x_4 = a + r(x_2 - a) = a + r^2(b - a)$$



Golden Section Search Algorithm

Example: Solve the following problem with the final interval of uncertainty's length is less than 1/4.

$$\max -x^2 - 1\,, \quad -1 \leq x \leq 0.75$$

We have $a = -1$ and $b = 0.75$ and $b - a = 1.75$. We should perform k iterations where

$$1.75(0.618)^k \leq 0.25 \Rightarrow k > \frac{\ln\left(\frac{1}{7}\right)}{\ln(0.618)} = 4.06$$

Golden Section Search Algorithm

$$x_1 = 0.75 - (0.618)(1.75) = -0.3315 \Rightarrow f(x_1) = -1.1099$$

$$x_2 = -1 + (0.618)(1.75) = 0.0815 \Rightarrow f(x_2) = -1.0066$$

$$f(x_1) < f(x_2) \Rightarrow I_1 = (x_1, b] = (-0.3315,0.75]$$

We also have $L_1 = 0.75 + 0.3315 = 1.0815$. We can now write

$$x_3 = x_2 = 0.0815 \Rightarrow f(x_3) = -1.0066$$

$$x_4 = -3315 + 0.618(1.0815) = 0.3369 \Rightarrow f(x_4) = -1.1135$$

$$f(x_3) > f(x_4) \Rightarrow I_1 = [-0.3315, x_4) = [-0.3315,0.3369)$$

Continuing in a similar manner, we obtain

$I_5 = (x_9, 0.0815] = (-0.0762,0.0815]$ and $L_5 = 0.1577 < 0.25$

Unconstrained Problems

Consider the following NLP:

$$\max f(\mathbf{x}),\ x \in R^n$$

or

$$\min f(\mathbf{x}),\ x \in R^n$$

We assume that the first and second partial derivatives exist and continuous for $\forall x_i$.

Unconstrained Problems

Theorem:

If $\bar{x}$ is a local extremum, then,

$$\frac{\partial f(\bar{x})}{\partial x_i} = 0$$

$\bar{x}$ is called a stationary point of f.

Unconstrained Problems

Theorem:

If $H_k(\bar{x}) > 0$, $k = 1,2, \dots, n$, then, a stationary point $\bar{x}$ is local min for the NLP.

Theorem:

If $H_k(\bar{x}) \neq 0$, $k = 1,2, \dots, n$, and has the same sign with $(-1)^k$, then, a stationary point $\bar{x}$ is local max for the NLP.

Unconstrained Problems

Theorem:

If $H_k(\bar{x}) \neq 0$, $k = 1,2,\ldots,n$, and the above theorems do not hold , then, a stationary point $\bar{x}$ is not a local extremum.

Theorem:

If a stationary point is not a local extremum, then, it is a saddle point.

Unconstrained Problems

Example:

$$f(x_1, x_2) = x_1^2 x_2 + x_1 x_2^3 - x_1 x_2$$

$$\frac{\partial f}{\partial x_1} = 2x_1 x_2 + x_2^3 - x_2$$

$$\frac{\partial f}{\partial x_2} = x_1^2 + 3x_1 x_2^2 - x_1$$

$$\frac{\partial f}{\partial x_1} = 0 \Rightarrow x_2(2x_1 + x_2^2 - 1) = 0 \Rightarrow x_2 = 0 \lor 2x_1 + x_2^2 - 1 = 0$$

$$\frac{\partial f}{\partial x_1} = 0 \Rightarrow x_1(x_1 + 3x_2^2 - 1) = 0 \Rightarrow x_1 = 0 \lor x_1 + 3x_2^2 - 1 = 0$$

Unconstrained Problems

If we write,

(i). $x_2 = 0$
(ii). $2x_1 + x_2^2 - 1 = 0$
(iii). $x_1 = 0$
(iv). $x_1 + 3x_2^2 - 1 = 0$

We obtain the following stationary points:

- (i) and (iii) hold. $(x_1, x_2) = (0,0)$
- (i) and (iv) hold. $(x_1, x_2) = (1,0)$
- (ii) and (iii) hold. $(x_1, x_2) = (0,1)$ and $(x_1, x_2) = (0,-1)$
- (ii) and (iv) hold. $(x_1, x_2) = \left(\frac{2}{5}, \frac{\sqrt{5}}{5}\right)$ and $(x_1, x_2) = \left(\frac{2}{5}, -\frac{\sqrt{5}}{5}\right)$

Unconstrained Problems

$$H(x_1, x_2) = \begin{bmatrix} 2x_2 & 2x_1 + 3x_2^2 - 1 \\ 2x_1 + 3x_2^2 - 1 & 6x_1 x_2 \end{bmatrix}$$

Note that $H_1(0,0) = 0$, and $H_2(0,0) = -1 \neq 0$, then, $(0,0)$ is a saddle point.

Also note that $H_1(1,0) = 0$ and $H_2(1,0) = -1 \neq 0$, then, $(1,0)$ is also a saddle point.

Also note that $H_1(0,1) = 2$ and $H_2(0,1) = -4 \neq 0$, then, $(1,0)$ is also a saddle point.

Unconstrained Problems

Finally, since

$$H_1\left(\frac{2}{5},-\frac{\sqrt{5}}{5}\right)=-\frac{2}{\sqrt{5}}<0,\quad H_2\left(\frac{2}{5},-\frac{\sqrt{5}}{5}\right)=\frac{4}{5}>0$$

and

$$H_1\left(\frac{2}{5},\frac{\sqrt{5}}{5}\right)=\frac{2}{\sqrt{5}}>0,\quad H_2\left(\frac{2}{5},\frac{\sqrt{5}}{5}\right)=\frac{4}{5}>0$$

Points $\left(\frac{2}{5},-\frac{\sqrt{5}}{5}\right)$ and $\left(\frac{2}{5},\frac{\sqrt{5}}{5}\right)$ are local max and local min, respectively.

Steepest Ascent

We can use the method of steepest ascent to approximate a stationary point. Given a vector $\mathbf{x} = (x_1, x_2, \dots, x_n) \in R^n$, the length of $\mathbf{x}$ is

$$\left|\left|\mathbf{x}\right|\right| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

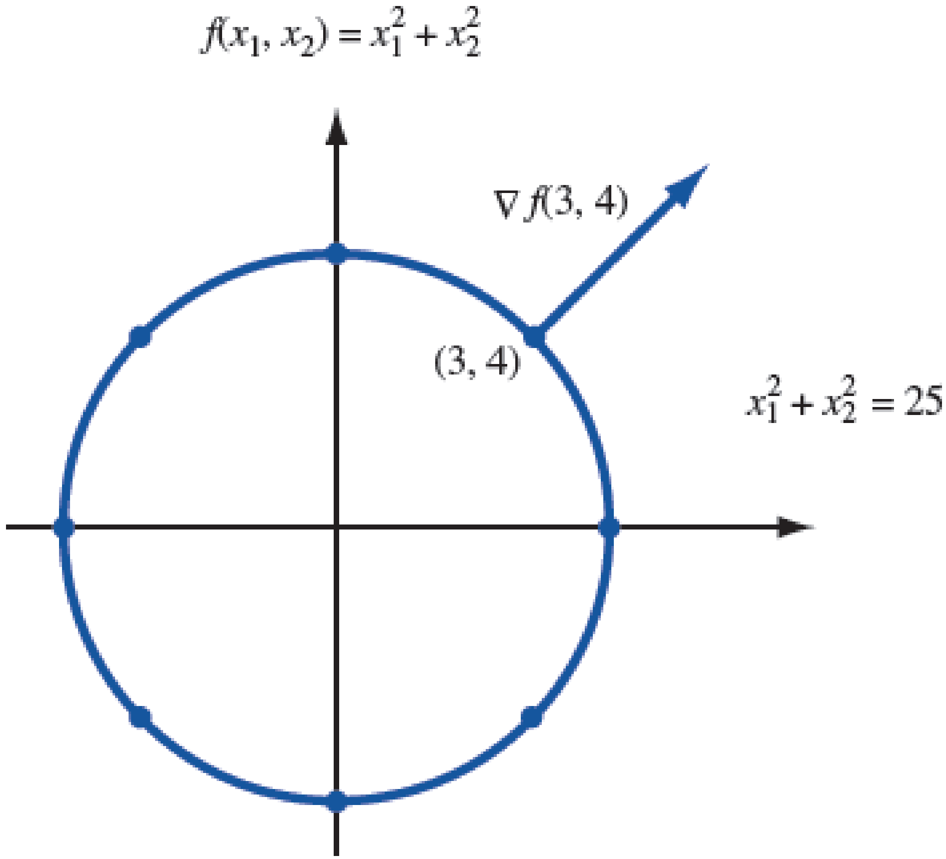
The gradient vector for $f(x_1, x_2, \dots, x_n)$ is

$$\nabla f(\mathbf{x}) = \left[\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \dots, \frac{\partial f(\mathbf{x})}{\partial x_n}\right]$$

$\nabla f(\mathbf{x})$ defines the direction

$$\frac{\nabla f(\mathbf{x})}{\left|\left|\nabla f(\mathbf{x})\right|\right|}$$

Steepest Ascent



Steepest Ascent

Example: Use the method of steepest ascent to find the approximate solution of the following NLP:

$$\max f(x_1, x_2) = -(x_1 - 3)^2 - (x_2 - 2)^2, \quad x_1, x_2 \in R^2$$

We arbitrarily choose $v_0 = (1,1)$.

$$\nabla f(x_1, x_2) = [-2(x_1 - 3) \quad -2(x_2 - 2)] \Rightarrow \nabla f(1,1) = [4 \quad 2]$$

Steepest Ascent

We thus choose t_0 as to maximize

$$
\begin{aligned}
f(t_0) &= f[(1,1) + t_0(4,2)] \\
&= f(1 + 4t_0, 1 + 2t_0) \\
&= -(-2 + 4t_0)^2 - (-1 + 2t_0)^2
\end{aligned}
$$

$$f'(t_0) = 0 \Rightarrow t_0 = 0.5$$

$$v_1 = (1,1) + 0.5(4,2) = (3,2) \Rightarrow\Rightarrow \nabla f(1,1) = [0 \quad 0]$$

Since f is concave, $(3,2)$ is the optimal solution.

Lagrange Multipliers

If all constraints are equalities, we can use the Lagrange Multipliers to solve such NLPs. Consider an NLP as follows:

$$\max(\min) f(\mathbf{x})$$

$$\begin{aligned}
g_1(\mathbf{x}) &= b_1 \\
g_2(\mathbf{x}) &= b_2 \\
&\dots \dots \dots \dots \dots \\
g_m(\mathbf{x}) &= b_m
\end{aligned}$$

Lagrange Multipliers

If we associate a multiplier λ_i with the ith constraint, we perform the Lagrangian as follows:

$$L(x_1, x_2, \dots, x_n; \lambda_1, \lambda_2, \dots, \lambda_m) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \left[b_i - g_i(\mathbf{x})\right]$$

We then find,

$$\frac{\partial L}{\partial \lambda_i} = b_i - g_i(\mathbf{x}) = 0$$

If we have $\max L(x_1, x_2, \dots, x_n; \lambda_1, \lambda_2, \dots, \lambda_m)$, then, it is necessary that $\left(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n; \bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_m\right)$ satisfies the following condition:

$$\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial x_2} = \cdots = \frac{\partial L}{\partial x_n} = \frac{\partial L}{\partial \lambda_1} = \frac{\partial L}{\partial \lambda_2} = \cdots = \frac{\partial L}{\partial \lambda_m} = 0$$

fatihcavdur@uludag.edu.tr

Lagrange Multipliers

Theorem:

If we have a max NLP, and if $f(\mathbf{x})$ is a concave function and each $g_i(\mathbf{x})$ is a linear function, then, any point $\left(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n; \bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_m\right)$ satisfying the following equation $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ is optimal:

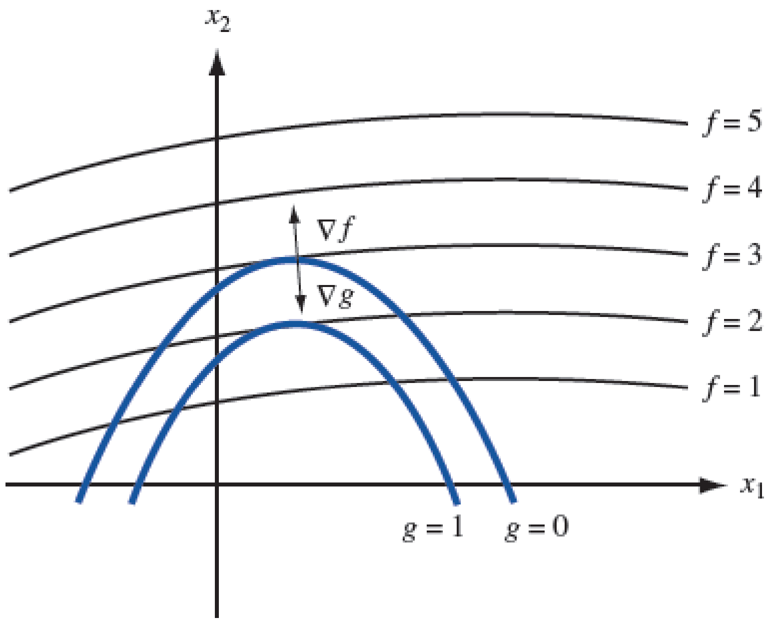
$$\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial x_2} = \cdots = \frac{\partial L}{\partial x_n} = \frac{\partial L}{\partial \lambda_1} = \frac{\partial L}{\partial \lambda_2} = \cdots = \frac{\partial L}{\partial \lambda_m} = 0$$

Theorem:

If we have a min NLP, and if $f(\mathbf{x})$ is a convex function and each $g_i(\mathbf{x})$ is a linear function, then, any point $\left(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n; \bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_m\right)$ satisfying the following equation $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ is optimal:

$$\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial x_2} = \cdots = \frac{\partial L}{\partial x_n} = \frac{\partial L}{\partial \lambda_1} = \frac{\partial L}{\partial \lambda_2} = \cdots = \frac{\partial L}{\partial \lambda_m} = 0$$

Lagrange Multipliers



The Kuhn-Tucker Conditions

In this section, we discuss the necessary and sufficient conditions for $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ to be an optimal solution to the following NLP:

$$\max(\min) f(\mathbf{x})$$

$$\begin{aligned}
&g_1(\mathbf{x}) \leq b_1 \\
&g_2(\mathbf{x}) \leq b_2 \\
&\ldots\ldots\ldots\ldots\ldots \\
&g_m(\mathbf{x}) \leq b_m
\end{aligned}$$

The Kuhn-Tucker Conditions

Theorem (*):

If $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ is an optimal solution to the NLP which is a max problem, then, $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ must satisfy the m constraints in the NLP and there must exist multipliers $\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_m$ satisfying

$$\frac{\partial f(\bar{\mathbf{x}})}{\partial x_j} - \sum_{i=1}^{m} \bar{\lambda}_i \frac{\partial g_i(\bar{\mathbf{x}})}{\partial x_j} = 0; \quad \forall j$$

$$\bar{\lambda}_i [b_i - g_i(\bar{\mathbf{x}})] = 0; \quad \forall i$$

$$\bar{\lambda}_i \geq 0; \quad \forall i$$

The Kuhn-Tucker Conditions

Theorem (**):

If $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ is an optimal solution to the NLP which is a min problem, then, $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ must satisfy the m constraints in the NLP and there must exist multipliers $\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_m$ satisfying

$$\frac{\partial f(\bar{\mathbf{x}})}{\partial x_j} + \sum_{i=1}^{m} \bar{\lambda}_i \frac{\partial g_i(\bar{\mathbf{x}})}{\partial x_j} = 0; \quad \forall j$$

$$\bar{\lambda}_i [b_i - g_i(\bar{\mathbf{x}})] = 0; \quad \forall i$$

$$\bar{\lambda}_i \geq 0; \quad \forall i$$

The Kuhn-Tucker Conditions

Theorem (***):

If $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ is an optimal solution to the NLP which is a max problem, then, $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ must satisfy the m constraints in the NLP and there must exist multipliers $\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_m; \bar{\mu}_1, \bar{\mu}_2, \dots, \bar{\mu}_n$ satisfying

$$\begin{aligned}
\frac{\partial f(\bar{\mathbf{x}})}{\partial x_j} - \sum_{i=1}^{m} \bar{\lambda}_i \frac{\partial g_i(\bar{\mathbf{x}})}{\partial x_j} + \mu_j &= 0; \quad \forall j \\
\bar{\lambda}_i [b_i - g_i(\bar{\mathbf{x}})] &= 0; \quad \forall i \\
\left[\frac{\partial f(\bar{\mathbf{x}})}{\partial x_j} - \sum_{i=1}^{m} \bar{\lambda}_i \frac{\partial g_i(\bar{\mathbf{x}})}{\partial x_j}\right] \bar{x}_j &= 0; \quad \forall j \\
\bar{\lambda}_i &\geq 0; \quad \forall i \\
\bar{\mu}_j &\geq 0; \quad \forall j
\end{aligned}$$

fatihcavdur@uludag.edu.tr

The Kuhn-Tucker Conditions

Theorem (****):

If $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ is an optimal solution to the NLP which is a min problem, then, $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ must satisfy the m constraints in the NLP and there must exist multipliers $\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_m; \bar{\mu}_1, \bar{\mu}_2, \dots, \bar{\mu}_n$ satisfying

$$\frac{\partial f(\bar{\mathbf{x}})}{\partial x_j} - \sum_{i=1}^{m} \bar{\lambda}_i \frac{\partial g_i(\bar{\mathbf{x}})}{\partial x_j} - \mu_j = 0; \quad \forall j$$

$$\bar{\lambda}_i [b_i - g_i(\bar{\mathbf{x}})] = 0; \quad \forall i$$

$$\left[\frac{\partial f(\bar{\mathbf{x}})}{\partial x_j} + \sum_{i=1}^{m} \bar{\lambda}_i \frac{\partial g_i(\bar{\mathbf{x}})}{\partial x_j}\right] \bar{x}_j = 0; \quad \forall j$$

$$\bar{\lambda}_i \geq 0; \quad \forall i$$

$$\bar{\mu}_j \geq 0; \quad \forall j$$

The Kuhn-Tucker Conditions

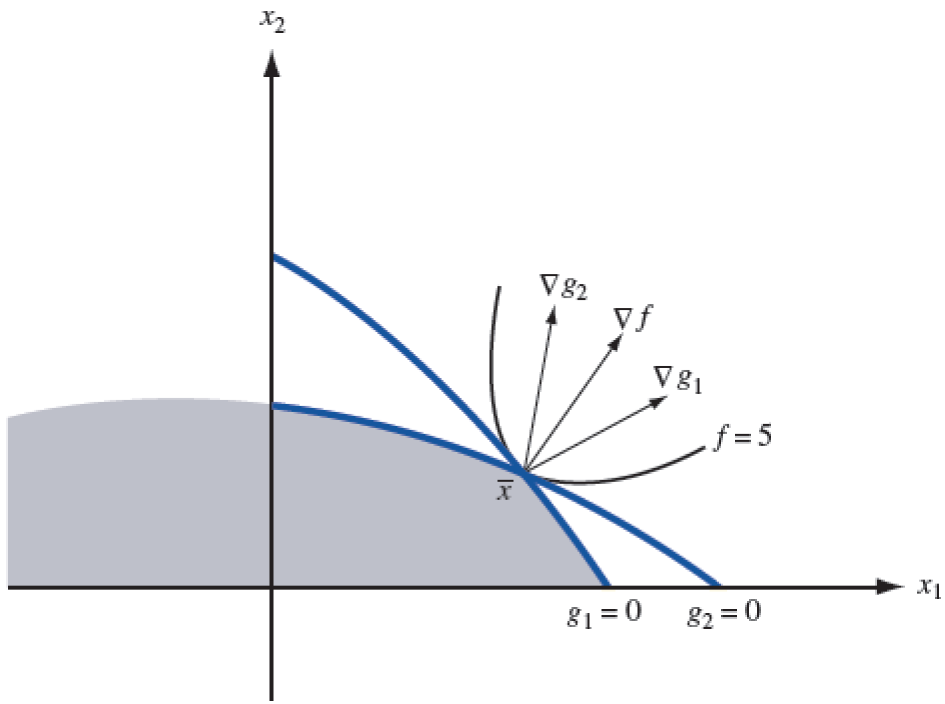
Theorem:

If $f(\mathbf{x})$ is a concave function and if $g_i(\mathbf{x})$ are convex functions for $\forall i$, then, any point $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ satisfying Theorem (*) is an optimal solution to the NLP which is a max problem.
If $f(\mathbf{x})$ is a concave function and if $g_i(\mathbf{x})$ are convex functions for $\forall i$, then, any point $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ satisfying Theorem (***) is an optimal solution to the NLP which is a max problem.

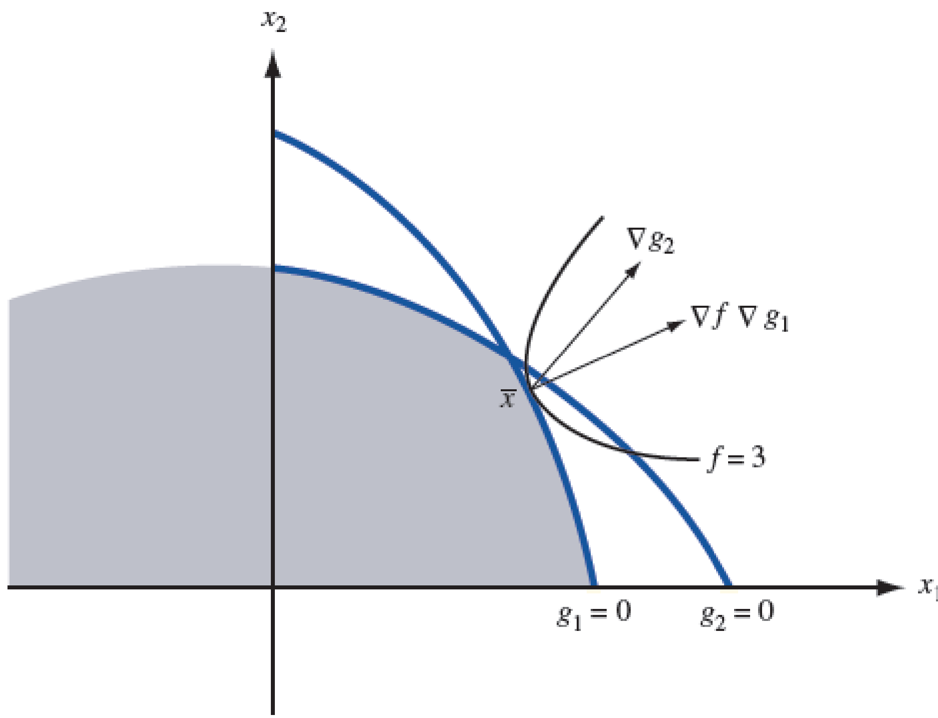
Theorem:

If $f(\mathbf{x})$ is a convex function and if $g_i(\mathbf{x})$ are convex functions for $\forall i$, then, any point $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ satisfying Theorem (**) is an optimal solution to the NLP which is min problem.
If $f(\mathbf{x})$ is a concave function and if $g_i(\mathbf{x})$ are convex functions for $\forall i$, then, any point $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ satisfying Theorem (****) is an optimal solution to the NLP which is a min problem.

The Kuhn-Tucker Conditions



The Kuhn-Tucker Conditions



The Kuhn-Tucker Conditions

Example:

$$\max z = x_1(30 - x_1) + x_2(50 - 2x_2) - 3x_1 - 5x_2 - 10x_3$$

$$\begin{aligned} x_1 + x_2 - x_3 &\leq 0 \\ x_3 &\leq 17.25 \\ x_i &\geq 0, \forall i \end{aligned}$$

The K-T conditions are

$$\begin{aligned} 30 - 2x_1 - 3 - \lambda_1 &= 0 \\ 50 - 4x_2 - 5 - \lambda_1 &= 0 \\ -10 + \lambda_1 - \lambda_2 &= 0 \\ \lambda_1(-x_1 - x_2 + x_3) &= 0 \\ \lambda_2(17.25 - x_3) &= 0 \\ \lambda_1 &\geq 0 \\ \lambda_2 &\geq 0 \end{aligned}$$

The Kuhn-Tucker Conditions

Case 1: $\lambda_1 = \lambda_2 = 0$. It violates the third constraint.

Case 2: $\lambda_1 = 0, \lambda_2 > 0$. It violates the third constraint.

Case 3: $\lambda_1 > 0, \lambda_2 = 0$. By solving the above system, we have $x_1 = 8.5$, $x_2 = 8.75$, $x_3 = 17.25$, $\lambda_1 = 10$ and $\lambda_2 = 0$, which satisfies the K-T conditions.

Case 4: $\lambda_1 > 0, \lambda_2 > 0$. Since Case (3) gives the optimal solution, it is not necessary to consider this.

Quadratic Programming

An NLP whose constraints are linear and whose objective is the sum of the terms of the form $x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$ with each term having a degree of 0, 1 or 2 is a Quadratic Programming Problem (QPP).

Portfolio Selection

Example: I have \$1,000 to invest in three stocks. Let S_i be the random variable representing the annual return on \$1 invested in stock i. Thus, if $S_i = 0.12$, \$1 invested in stock i at the beginning of a year was worth \$1.12 at the end of the year. We are given the following information:

$$E(S_1) = 0.14, E(S_2) = 0.11, E(S_3) = 0.10$$

$$V(S_1) = 0.20, V(S_2) = 0.08, V(S_3) = 0.18$$

$$C(S_1, S_2) = 0.05, C(S_1, S_3) = 0.02, C(S_2, S_3) = 0.03$$

Formulate a QPP that can be used to find the portfolio that attains an expected annual return of at least 12% and minimizes the variance of the annual dollar return on the portfolio.

Portfolio Selection

Let x_i be the number of dollars invested in stock i.

$$\min 0.20x_1^2 + 0.08x_2^2 + 0.18x_3^2 + 0.10x_1x_2 + 0.04x_1x_3 + 0.06x_2x_3$$

$$0.14x_1 + 0.11x_2 + 0.10x_3 \geq 120$$

$$\sum_{i=1}^{3} x_i = 1$$

$$x_i \geq 0, \forall i$$

Wolfe's Method for QPs

We can use the Wolfe's method to solve QPPs with non-negative variables. Consider the following example:

$$\min z = -x_1 - x_2 + \frac{x_1^2}{2} + x_2^2 - x_1 x_2$$

$$\begin{aligned} x_1 + x_2 &\leq 3 \\ -2x_1 - 3x_2 &\leq -6 \\ x_i &\geq 0; \quad \forall i \end{aligned}$$

Wolfe's Method for QPs

We can write the followings where all variables are non-negative:

$$
\begin{aligned}
x_1 - 1 - x_2 + \lambda_1 - 2\lambda_2 - e_1 &= 0 \\
2x_1 - 1 - x_1 + \lambda_1 - 3\lambda_2 - e_2 &= 0 \\
x_1 + x_2 + s_1{}' &= 3 \\
2x_1 + 3x_2 - e_2{}' &= 6 \\
\lambda_2 e_2' &= 0 \\
\lambda_1 s_1' &= 0 \\
e_1 x_1 &= 0 \\
e_2 x_2 &= 0
\end{aligned}
$$

Wolfe's Method for QPs

We note that except the last 4 equations, all equations are linear. To apply the Wolfe's method, we need to solve the following LP with all non-negative variables:

$$\min w = a_1 + a_2 + a_2{}'$$

$$\begin{aligned} x_1 - x_2 + \lambda_1 - 2\lambda_2 - e_1 + a_1 &= 1 \\ -x_1 + 2x_2 + \lambda_1 - 3\lambda_2 - e_2 + a_2 &= 1 \\ x_1 + x_2 + s_1{}' &= 3 \\ 2x_1 + 3x_2 + e_2' + a_2{}' &= 6 \end{aligned}$$

Wolfe’s Method for QPs

The optimal solution of the LP is shown in the below table where

$$w = 0; x_1 = \frac{9}{5}, x_2 = \frac{6}{5}, \lambda_1 = \frac{2}{5}, \lambda_2 = 0 \text{ (since } e_2^{'} = \frac{6}{5}, \lambda_2 = 0)$$

Wolfe’s method is guaranteed to obtain the optimal solution to a QPP if all leading principal minors of the objective function’s Hessian are positive. Otherwise, Wolfe’s method may not converge in a finite number of pivots. In practice, the method of complementary pivoting is most often used to solve QPPs which will not be discussed in this class.

Wolfe's Method for QPs

w	x_1	x_2	λ_1	λ_2	e_1	e_2	$s_1{}'$	$e_2{}'$	a_1	a_2	$a_2{}'$	RHS
1	0	0	0	0	0	0	0	0	-1	-1	-1	0
0	0	0	1	-12/5	-3/5	-2/5	-1/5	0	3/5	2/5	0	2/5
0	0	1	0	-1/5	1/5	-1/5	2/5	0	-1/5	1/5	0	6/5
0	0	0	0	-1/5	1/5	-1/5	12/5	1	-1/5	1/5	-1	6/5
0	1	0	0	1/5	-1/5	1/5	3/5	0	1/5	-1/5	0	9/5

Separable Programming

$$\max z = \sum_{i=1}^{n} f_i(x_i)$$

$$\sum_{i=1}^{n} g_{ij}(x_i) \leq b_j, \quad j = 1, \ldots, m$$

Separable Programing Problems are often solved by approximating $f_i(x_i)$ and $g_{ij}(x_i)$ by a piecewise linear function.

Method of Feasible Directions

A modification of the method of steepest descent, the method of feasible directions, can be used to solve NLPs with linear constraints.

$$\max z = f(\mathbf{x})$$

$$\begin{aligned} \mathbf{Ax} &\leq \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

We assume that $f(\mathbf{x})$ is a concave function.

Method of Feasible Directions

We start with $\mathbf{x}^0$ that satisfies the constraints and try to find a direction in which we can move away from $\mathbf{x}^0$ which has the following properties:

- When we move away from $\mathbf{x}^0$, we remain feasible.
- When we move away from $\mathbf{x}^0$, we increase z.

Method of Feasible Directions

We choose to move away from $\mathbf{x}^0$ in a direction $\mathbf{d}^0 - \mathbf{x}^0$, where $\mathbf{d}^0$ is an optimal solution to the following LP:

$$\max z = \nabla f(\mathbf{x}^0) \cdot \mathbf{d}$$

$$\mathbf{Ad} \leq \mathbf{b}$$
$$\mathbf{d} \geq \mathbf{0}$$

We now choose our new point $\mathbf{x}^1 = \mathbf{x}^0 + t_0(\mathbf{d}^0 - \mathbf{x}^0)$ where t_0 solves

$$\max f[\mathbf{x}^0 + t_0(\mathbf{d}^0 - \mathbf{x}^0)]$$

$$0 \leq t_0 \leq 1$$

Pareto Optimality

Step 1: Choose an objective (say objective 1) and determine the best value of this objective that can be attained (call it v_1). For this best solution, find the value of objective 2 (call it v_2). (v_1, v_2) is then a point on the trade-off curve.

Step 2: For values v of objective 2 that are better than v_2, solve the optimization problem in Step (1) with the additional constraint: The value of objective 2 is at least as good as v. Varying v will give you other points on the trade-off curve.

Step 3: In Step 1, we obtained one end point of the trade-off curve. If we determine the best value of objective 2 that can be attained, we obtain the other end point of the trade-off curve.

The End