ENGINEERING OPTIMIZATION

B.TECH. V SEMESTER

MECHANICAL ENGINEERING

PREPARED BY:

T.VANAJA, ASSISTANT PROFESSOR

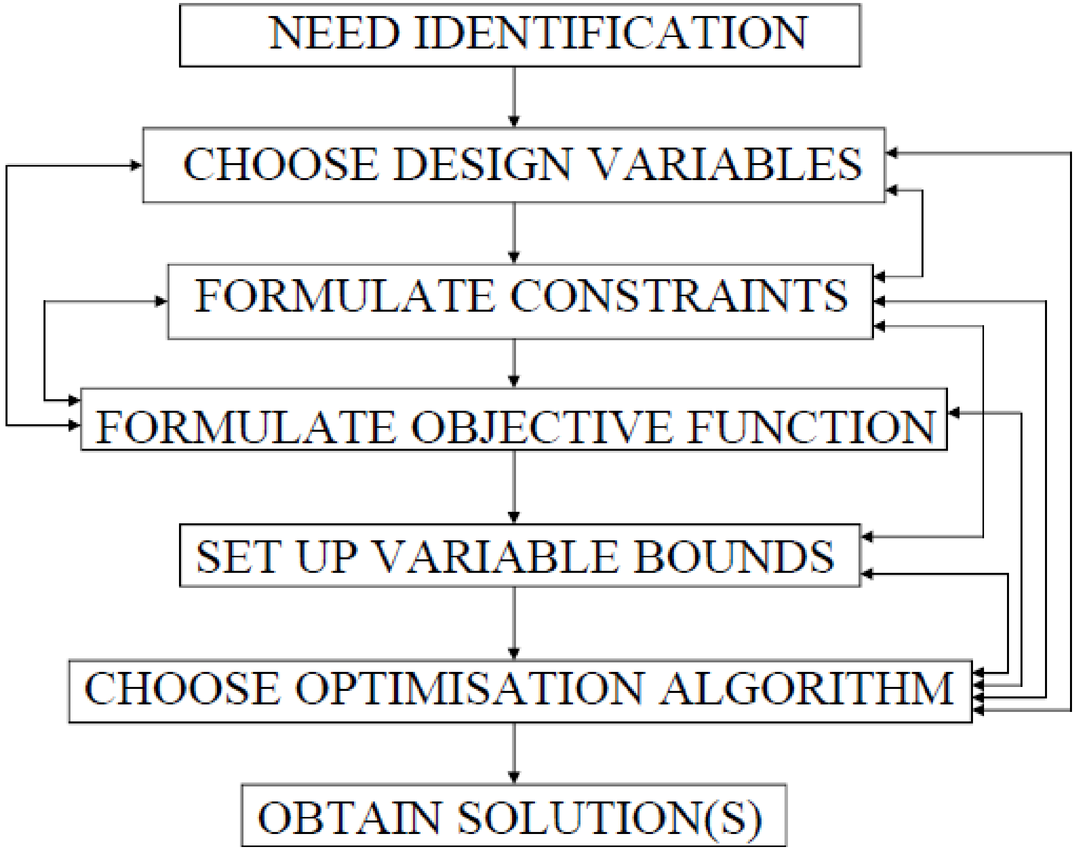
Institute of Aeronautical Engineering (Autonomous)

UNIT – I

INTRODUCTION TO OPTIMIZATION

Introduction to Optimization

STEPS IN FORMULISATION OF AN OPTIMISATION PROBLEM



Optimization is the act of obtaining the best result under given circumstances.

Optimization can be defined as the process of finding the conditions that give the maximum or minimum of a function.

The optimum seeking methods are also known as *mathematical programming techniques* and are generally studied as a part of operations research.

Operations research is a branch of mathematics concerned with the application of scientific methods and techniques to decision making problems and with establishing the best or optimal solutions.

TABLE 1.1 Methods of Operations Research

Mathematical Programming Techniques	Stochastic Process Techniques	Statistical Methods
Calculus methods	Statistical decision theory	Regression analysis
Calculus of variations	Markov processes	Cluster analysis, pattern recognition
Nonlinear programming	Queueing theory	Design of experiments
Geometric programming	Renewal theory	Discriminate analysis (factor analysis)
Quadratic programming	Simulation methods	
Linear programming	Reliability theory	
Dynamic programming		
Integer programming		
Stochastic programming		
Separable programming		
Multiobjective programming		
Network methods: CPM and PERT		
Game theory		
Simulated annealing		
Genetic algorithms		
Neural networks		

Mathematical optimization problem:

$$\text{minimize } f_0(x)$$

$$\text{subject to } g_i(x) \leq b_i, \quad i = 1, \ldots, m$$

f_0 : $\boldsymbol{R^n}$ $\boldsymbol{R}$**:** objective function

$x=(x_1, \ldots\ldots, x_n)$: design variables (unknowns of the problem, they must be linearly independent)

g_i : $\boldsymbol{R^n}$ $\boldsymbol{R}$**:** $(i=1, \ldots, m)$: inequality constraints

The problem is a constrained optimization problem

- If a point x^* corresponds to the minimum value of the function $f(x)$, the same point also corresponds to the maximum value of the negative of the function, $-f(x)$. Thus optimization can be taken to mean minimization since the maximum of a function can be found by seeking the minimum of the negative of the same function.

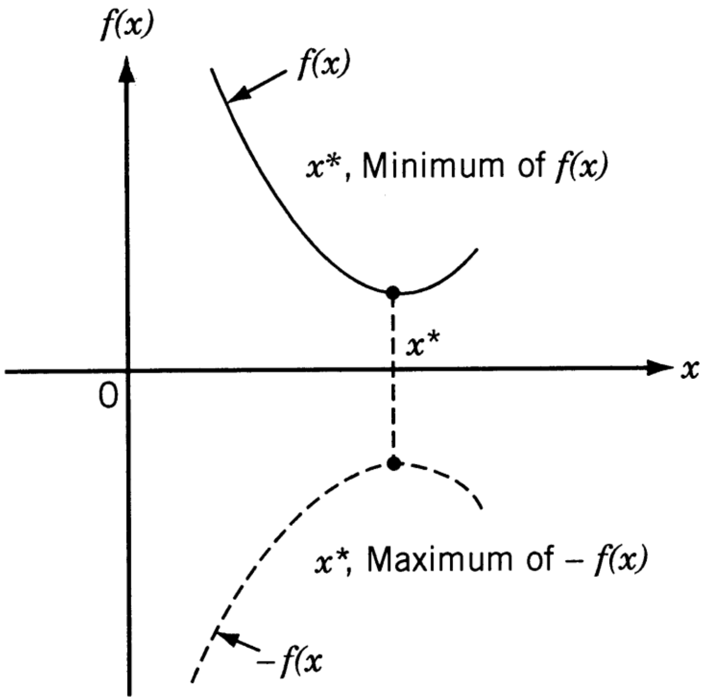


Figure 1.1 Minimum of $f(x)$ is same as maximum of $-f(x)$.

Constraints

Behaviour constraints: Constraints that represent limitations on the behaviour or performance of the system are termed *behaviour* or *functional constraints.*

Side constraints: Constraints that represent physical limitations on design variables such as manufacturing limitations.

In civil engineering, the objective is usually taken as the minimization of the cost.

In mechanical engineering, the maximization of the mechanical efficiency is the obvious choice of an objective function.

In aerospace structural design problems, the objective function for minimization is generally taken as weight.

In some situations, there may be more than one criterion to be satisfied simultaneously. An optimization problem involving multiple objective functions is known as a **multiobjective programming problem.**

With multiple objectives there arises a possibility of conflict, and one simple way to handle the problem is to construct an overall objective function as a linear combination of the conflicting multiple objective functions.

Thus, if f_1 **(X)** and f_2 **(X)** denote two objective functions, construct a new (overall) objective function for optimization as:

where α_1 and α_2 are constants whose values indicate the relative importance of one objective function to the other.

Classification of optimization problems

Classification based on:

Constraints

- Constrained optimization problem
- Unconstrained optimization problem

Nature of the design variables

- Static optimization problems
- Dynamic optimization problems

Classification based on:

Physical structure of the problem

Optimal control problems

Non-optimal control problems

Nature of the equations involved

Nonlinear programming problem

Geometric programming problem

Quadratic programming problem

Linear programming problem

Classification based on:

Permissable values of the design variables

Integer programming problems

Real valued programming problems

Deterministic nature of the variables

Stochastic programming problem

Deterministic programming problem

UNIT – II

SINGLE VARIABLE OPTIMIZATION

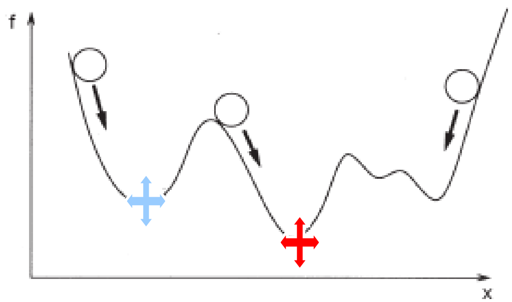
Useful in finding the optimum solutions of continuous and differentiable functions

These methods are analytical and make use of the techniques of differential calculus in locating the optimum points.

Since some of the practical problems involve objective functions that are not continuous and/or differentiable, the classical optimization techniques have limited scope in practical applications.

A function of one variable $f(x)$ has a relative or local minimum at $x = x^*$ if $f(x^*) \leq f(x^*+h)$ for all sufficiently small positive and negative values of h

A point x^* is called a relative or local maximum if $f(x^*) \geq f(x^*+h)$ for all values of h sufficiently close to zero.



Global minima

Local minima

A function $f(x)$ is said to have a global or absolute minimum at x^* if $f(x^*) \leq f(x)$ for all x, and not just for all x close to x^*, in the domain over which $f(x)$ is defined.

Similarly, a point x^* will be a global maximum of $f(x)$ *if* $f(x^*) \geq f(x)$ for all x in the domain.

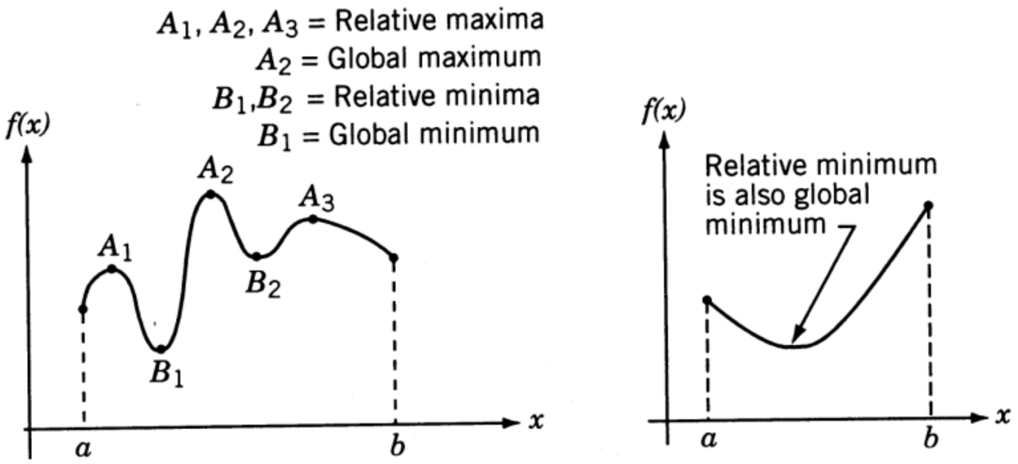
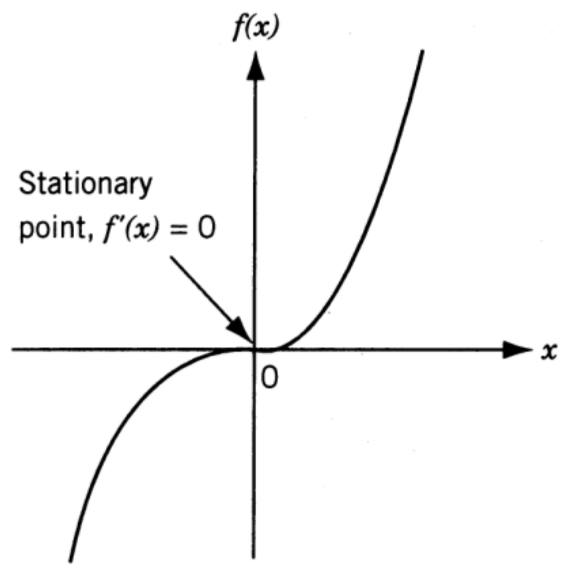


Figure 2.1 Relative and global minima.

derivative $df(x) / dx = f'(x)$ exists as a finite number at $x = x^*$, then $f'(x^*)=0$

The theorem does not say that the function necessarily will have a minimum or maximum at every point where the derivative is zero. *e.g.* $f'(x)=0$ at $x= 0$ for the function shown in figure. However, this point is neither a minimum nor a maximum. In general, a point x^* at which $f'(x^*)=0$ is called a *stationary point.*



The theorem does not say what happens if a minimum or a maximum occurs at a point x^* where the derivative fails to exist. For example, in the figure

$$\lim_{h \to 0} \frac{f(x^*+h) - f(x^*)}{h} = m^+ \text{ (positive) or } m^- \text{ (negative)}$$

depending on whether h approaches zero through positive or negative values, respectively. Unless the numbers or are equal, the derivative $f'(x^*)$ does not exist. If $f'(x^*)$ does not exist, the theorem is not applicable.

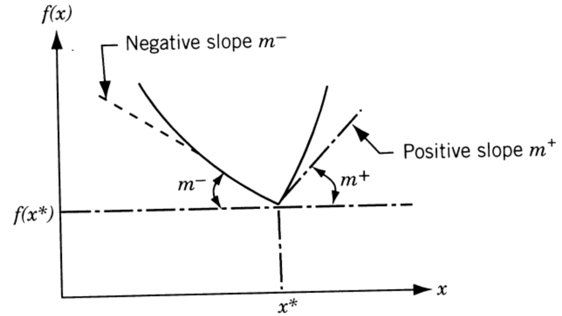


Figure 2.2 Derivative undefined at x^*.

Let $f'(x^*)=f''(x^*)=\ldots=f^{(n-1)}(x^*)=0$, but $f^{(n)}(x^*) \neq 0$. Then $f(x^*)$ is

A **minimum** value of $f(x)$ if $f^{(n)}(x^*) > 0$ and n is **even**

A **maximum** value of $f(x)$ if $f^{(n)}(x^*) < 0$ and n is **even**

Neither a minimum nor a maximum if n is **odd**

Determine the maximum and minimum values of the function:

Solution: Since $f'(x)=60(x^4-3x^3+2x^2)=60x^2(x-1)(x-2)$,

$f'(x)=0$ at $x=0, x=1$, and $x=2$.

The second derivative is:

At $x=1$, $f''(x)=-60$ and hence $x=1$ is a relative maximum. Therefore,

$$f_{max}= f(x=1) = 12$$

At $x=2$, $f''(x)=240$ and hence $x=2$ is a relative minimum. Therefore,

$$f_{min}= f(x=2) = -11$$

Solution cont'd:

At *x=0, f''(x)=0* and hence we must investigate the next derivative.

Since at *x=0, x=0* is neither a maximum nor a minimum, and it is an inflection point.

- <u>Exhaustive search algorithm</u> (given f(x), a & b)

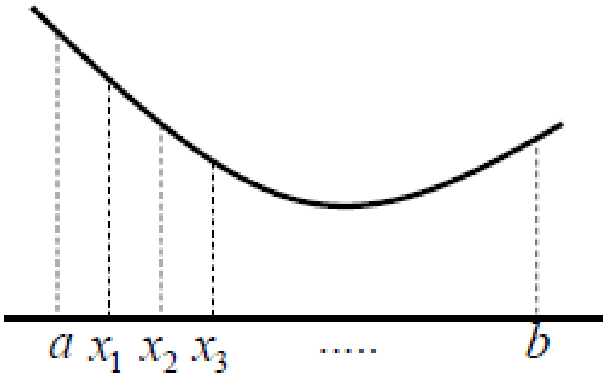
Step 1 set $x_1 = a$, $\Delta x = (b-a)/n$ (n is the number of intermediate points), $x_2 = x_1 + \Delta x$, and $x_3 = x_2 + \Delta x$.

Step 2 If $f(x_1) \geq f(x_2) \leq f(x_3)$, the minimum point lies in (x_1, x_3), **Terminate;**

Else $x_1 = x_2, x_2 = x_3, x_3 = x_2 + \Delta x$, and go to Step 3.

Step 3 Is $x_3 \leq b$? If yes, go to Step 2;

Else no minimum exists in (a,b) or a boundary point (a or b) is the minimum point.



Interval Halving Method

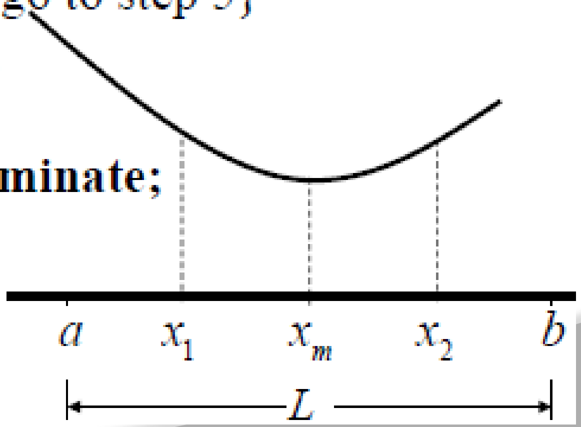
Step 1 Choose a lower bound a and an upper bound b. Choose also a small number ε. Let $x_m = (a+b)/2$, $L_o = L = b - a$. Compute $f(x_m)$.

Step 2 Set $x_1 = a + L/4$, $x_2 = b - L/4$. Compute $f(x_1)$ and $f(x_2)$.

Step 3 If $f(x_1) < f(x_m)$ set $b = x_m; x_m = x_1$ go to Step 5;
Else go to Step 4.

Step 4 If $f(x_2) < f(x_m)$ {$a = x_m$; $x_m = x_2$; go to step 5}
Else {$a = x_1$; $b = x_2$; go to step 5}.

Step 5 Calculate L = b-a If |L| < ε **Terminate;**
Else go to Step 2.



EXAMPLE

- Minimize $f(x) = x^2 + 54/x$ in the interval (0,5).

- Step 1: $\varepsilon = 10^{-3}$; a=0; b=5; $L_0 = 5$; $x_m = 2.5$; $f(x_m) = 27.85$.
- Step 2: $x_1 = 1.25; x_2 = 3.75$

  $f(x_1) = 44.7; f(x_2) = 28.4$

- Step 3: IS $f(x_1) < f(x_m)$? NO.
- Step 4: IS $f(x_2) < f(x_m)$? NO.

  hence [1.25 - 3.75] i.e a=1.25; b=3.75.

- Step 5: L=2.5; a=1.25;b=3.75; $x_m = 2.5$;

  $x_1 = 1.25 + 2.5/4 = 1.875$; $x_2 = 3.75 - 2.5/4 = 3.125$;

  $f(x_1) = 32.3$; $f(x_2) = 27.05$

- Step 3 IS $f(x_1) < f(x_m)$? NO.

- Step 4 IS $f(x_2) < f(x_m)$? YES.

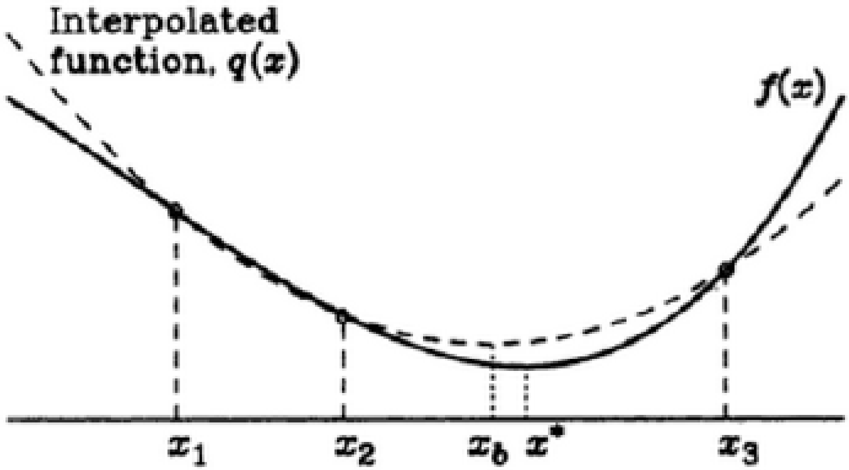
  a=2.5; b=3.75; $x_m = 3.125$

- Step 5 L= 1.25 (3.75 - 2.5)

Iteration continues

Successive Quadratic Estimation

- A quadratic curve is fitted through three points
- Start with three points x_1, x_2 and x_3



- A general quadratic function is

$$\bar{f}(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2)$$

- If (f_1, x_1), (f_2, x_2) and (f_3, x_3) are known,

$$a_0 = f_1 \qquad a_1 = \frac{(f_2 - f_1)}{(x_2 - x_1)}; a_2 = \frac{1}{(x_3 - x_2)}\left(\frac{(f_3 - f_1)}{(x_3 - x_1)} - a_1\right)$$

- Minima lies at

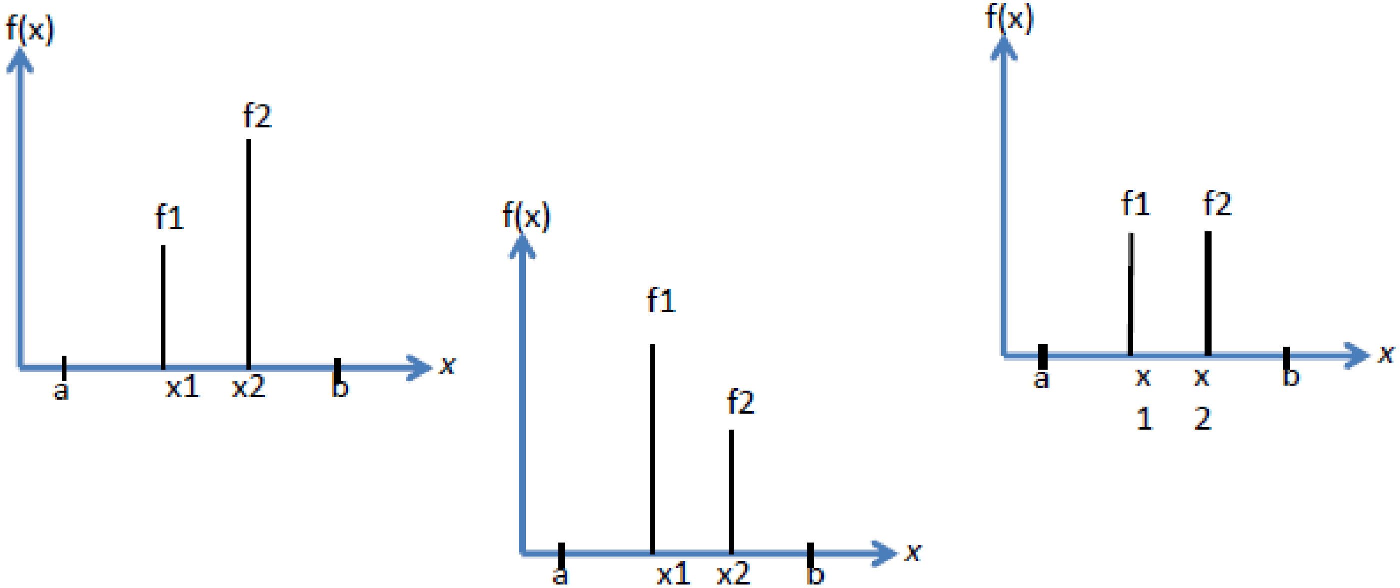
$$\bar{x} = \frac{x_1 + x_2}{2} - \frac{a_1}{2a_2}$$

Algorithm

- S1: Let x_1 be the initial point, Δ be the step size, $x_2 = x_1 + \Delta$
- S2: Compute $f(x_1)$ and $f(x_2)$
- S3: If $f(x_1) > f(x_2)$ Let $x_3=x_1+2\Delta$ else $x_3=x_1-\Delta$, Compute $f(x_3)$
- S4: Determine $f_{min} = \min(f_1,f_2,f_3)$ and X_{min}
- S5: Compute $\bar{x}$
- S 6: If $\left|f_{min} - f(\bar{x})\right|$ and $\left|x_{min} - \bar{x}\right|$ are small, optima is best of all known points
- S7: Save the best point and neighbouring points. Goto 4.

Bracketing Method based on unimodal property of objective function

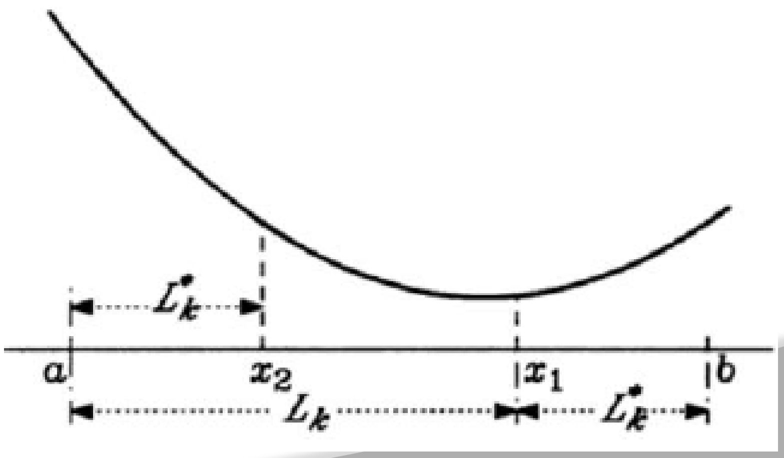
1) Assume an interval [a,b] with a minima in the range
2) Consider x1 and x2 within the interval.
3) Find out values of f at x=x1,x2 and Compare f(x1) and f(x2)
    a) If f(x1) < f(x2) then eliminate [x2,b] and set new interval [a,x2]
    b) If f(x1) > f(x2) then eliminate [a,x1] and set new interval [x1,b]
    c) If f(x1) = f(x2) then eliminate [a,x1] & [x2,b] and set new interval [x1,x2]



Advantage of Fibonacci Series

- Fibonacci Series: $F_0=F_1=1$; $F_n=F_{n-1}+F_{n-1}$
- Hence: 1,1,2,3,5,8,13,21,34,…

- $L_k-L_k^*=L_{k+1}^*$
- Hence one point is always pre-calculated



Fibonacci Search Algorithm

- Step 1: L=b-a; k = 2; Decide n;
- Step 2: $L_k^* = \left(F_{n-k+1} / F_{n+1} \right) L$ $\quad x_1 = a + L_k^*$ $\quad x_2 = b - L_k^*$
- Step 3:
  - Compute either f(x1) or f(x2) (whichever is not computed earlier)
  - Use region elimination rule
  - Set new a and b
- If k=n, TERMINATE else k=k+1 and GOTO Step 2

Golden Section Search Algorithm

- Step 1: $L=b-a$; $k = 1$; Decide ε; map a, b to $a'=0$; $b'=1$
- Step 2: $L_w=b'-a'$
  - $w_1=a'+0.618L_w$
  - $w_2=b'-0.618L_w$
- Step 3:
  - Compute either $f(w_1)$ or $f(w_2)$ (whichever is not computed earlier)
  - Use region elimination rule
  - Set new a' and b'
- If $|L_w|\leq\varepsilon$ TERMINATE else $k=k+1$ and GOTO Step 2

Function f(x) is

$0.65 - [0.75/(1+x^2)] - 0.65*x*\tan^{-1}(1/x)$ is minimized using golden section method with n=6

A=0 B=3

❖ *The location of first two expt. Points are defined by* $L_2^* = 0.382L_0 = (0.382)(3.0) = 1.146$

X1 = 1.1460   x2 = 3.0-1.1460 = 1.854   with   f1 =-0.208   f2 = -0.115

Since f1<f2   delete $[x_2, 3.0]$ *based on assumption of unimodality and new interval of uncertainty obtained is [0,x2] = [0, 1.854]*

❖ *The third experiment point is placed at x3 = 0 + (x2-x1) = 1.854 – 1.146 = 0.708.*

f3 = -0.288943   f1 = -0.208

Since f3<f1 delete interval [x1,x2].

The new interval of uncertainty is [0,x1] = [0, 1.146]

- *The fourth experiment point is placed at x4 = 0 + (x1-x3) = 0.438.*

  f4 = -0.308951 f3 = -0.288943

Since f4<f3 delete interval [x3,x1].

The new interval of uncertainty is [0,x3] = [0,0.7080]

- *The fifth experiment point is placed at x5 = 0 + (x3-x4) = 0.27.*

  f4 = -0.308951 f5 = -0.278

Since f4<f5 delete interval [0,x5].

The new interval of uncertainty is [x5,x3] = [0.27,0.7080]

- *The last experiment point is placed at x6 = x5 + (x3-x4) = 0.54.*

  f4 = -0.308951 f6 = -0.308234

Since f4<f6 delete interval [x6,x3].

The new interval of uncertainty is [x5,x6] = [0.27,0.54]

Ratio of final interval to initial interval is 0.09.

Bounding Phase Method

- Step 1

  Choose an initial guess X0 and an increment D. Set K = 0

- Step 2

  If $f(X0 - |D|) > f(X0) > f(X0 + |D|)$ then D is +ve

  $f(X0 - |D|) < f(X0) < f(X0 + |D|)$ then D is –ve

  else goto step 1

- Step3

  Set $X_{K+1} = X_K + 2^{K*}D$

- Step4

  If $f(X_{K+1}) < f(X_K)$. Set K =K+1 and goto step 3

  else, the minimum lies in the interval(X_{K-1}, X_{K+1}) and terminate

If D is large, accuracy is poor.

Minimize $f(x) = x^2 + 54/x$;

❖Step 1 Choose an initial guess X_0 = 0.6 and an increment D = 0.5. Set K = 0

❖Step 2 Calculate $f(X_0 - |D|) = f(0.1) = 540.010$ $f(X_0) = f(0.6) = 90.36$

$f(X_0 + |D|) = f(1.1) = 50.301$ We observe f(0.1)>f(0.6)>f(1.1) therefore D is +ve = 0.5

❖Step3 Set $X_1 = X_0 + 2^0 * D$ ➔ $X_1 = 1.1$

❖Step4 If $f(X_1) = 50.301 < f(X_0)$. Set K = 1 and goto step 3

Next guess $X_2 = X_1 + 2 * D = 2.1$

- $f(X_2) = 30.124 < f(X_1)$ therefore set K=2 and goto step 3
- Next guess $X_3 = X_2 + 2^2 * D = 4.1$
- $f(X_3) = 29.981 < f(X_2)$ therefore set K=3 and goto step 3
- Next guess $X_4 = X_3 + 2^3 * D$ $f(X_4) = 72.277 > f(X_3)$.
- Thus terminate with interval (2.1,8.1)

With D = 0.5 the bracketing is poor.

If D = 0.001, the obtained interval is (1.623,4.695).

UNIT – III

MULTI VARIABLE UNCONSTRAINED OPTIMIZATION

- If a function $f(x)$ is defined in the interval $a \leq x \leq b$ and has a relative minimum at $x = x^*$, where $a < x^* < b$, and if the derivative $df(x)/dx = f'(x)$ exists as a finite number at $x = x^*$, then $f'(x^*)=0$

Proof :

- It is given that

$$f'(x) = \lim_{h \to 0} \frac{f(x^*+h) - f(x^*)}{h} \quad \dots\dots\dots\dots\dots\dots 1$$

Exists as a definite number, which we want to prove to be zero. Since x^* is a relative minimum, we have

$f \leq f(x^*+h)$

for all values of h sufficiently close to *zero*. Hence

$$\frac{f(x^*+h) - f(x^*)}{h} \geq 0 \qquad if \quad h > 0$$

$$\frac{f(x^*+h) - f(x^*)}{h} \leq 0 \qquad if \quad h < 0$$

Thus equation *1* gives the limit as h tends to zero through positive values as

$f'(x^*) \geq 0$

while *1* gives the limit as h tends to *zero* through negative values as

$f'(x^*) \leq 0$

The only way to satisfy both the equations above is to have

$f'(x^*) = 0$

This proves the theorem

- Let $f'(x^*) = f''(x^*) = \ldots\ldots\ldots\ldots = f^{(n+1)}(x^*)$, $(x^*)=0$, but $f^{(n)}(x^*) \neq 0$
- Then $f(x)$ is

  (i) a minimum value of $f(x)$ if $f^{(n)}(x^*)>0$ and n is even ;
- (ii) a maximum value of $f(x)$ if $f^{(n)}(x^*)<0$ and n
  – is even ;
- (iii) neither a maximum nor a minimum if n is odd.

- *Proof* : Applying Taylor's theorem with reminder after n terms, we have

$$f(x^*+h) = (f(x^*) + hf'(x^*) + \frac{h^2}{2!}f''(x^*) + \ldots\ldots + \frac{h^{n-1}}{(n-1)!}f^{(n-1)}(x^*)$$

$$+ \frac{h^n}{n!}f^{(n)}(x^*+\theta h) \quad for \quad 0 < \theta < 1$$

Since $\quad f'(x^*) = f''(x^*) = \ldots\ldots = f^{(n-1)}(x^*) = 0$

Hence the above equation becomes

$$f(x^*+h) - f(x^*) = \frac{h^n}{(n)!}f^{(n)}(x^*+\theta h)$$

When n is even, $h^n/n!$ is positive irrespective of wether h is positive or negative, and hence $f^{(n)}(x^*+h)-f(x^*)$ will have the same sign as that of $f^{(n)}(x^*)$

Thus x^* will be

Relative minimum if $f^{(n)}(x^*)$ is positive

Relative maximum if $f^{(n)}(x^*)$ is negative

When n is odd $h^n/n!$ changes sign with the change in the sign of h and hence the point x^* is neither maximum nor a minimum.

In this case point x^* is called a point of ***inflection***

Gradient based Methods

- Algorithms require derivative information
- Many real world problems, difficult to obtain information about derivatives
  - Computations involved
  - Nature of problem
- Still gradient methods are effective and popular
- Recommended to use in problems where derivative information available
- Global optimum occurs where gradient is zero
- the search process terminates where gradient is zero

$$\text{Min} \quad f(x_1, x_2, x_3, \text{-----}\ x_n)$$

UNIDIRECTIONAL SEARCH

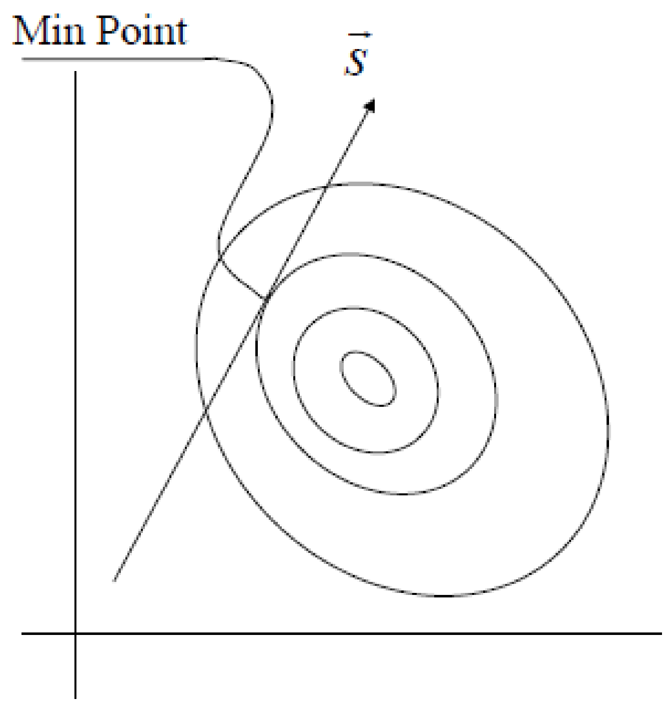
- CONSIDER A DIRECTION S

$$x(\alpha) = \vec{x} + \alpha\, \vec{s}$$

- REDUCE TO

$$Min \quad f(\alpha)$$

- SOLVE AS
  A SINGLE VARIABLE PROBLEM

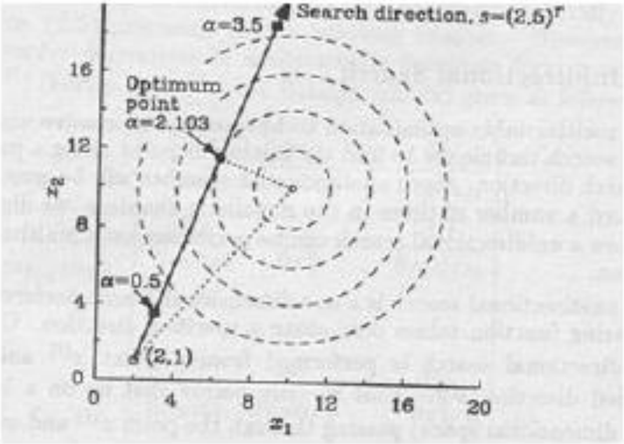


Uni directional search (example)

Min $f(x_1, x_2) = (x_1-10)^2 + (x_2-10)^2$

S = (2, 5) (search direction)

X = (2, 1) (Initial guess)



Hooke Jeeves pattern search

- Pattern Search ---
  - Create a set of search directions iteratively
  - Should be linearly independent
- A combination of exploratory and pattern moves
  - Exploratory – find the best point in the vicinity of the current point
  - Pattern – Jump in the direction of change, if better then continue, else reduce size of exploratory move and continue

Exploratory move

- Current solution is x^c; set $i = 1$; $x = x^c$
- S1: $f = f(x)$, $f^+ = f(x_i + \Delta_i)$, $f^- = f(x_i - \Delta_i)$
- S2: $f_{min} = \min(f, f^+, f^-)$; set x corresponding to f_{min}
- S3: If $i = N$, go to 4; else $i = i + 1$, go to 1
- S4: If $x \neq x^c$, success, else failure

Pattern Move

- S1: Choose $x^{(0)}$, Δ_I, for $I = 1, 2, \ldots N$, ε, and set $k = 0$
- S2: Perform exploratory move with x^k as base point;
  - If success, $x^{k+1} = x$, go to 4 else goto 3
- S3: If $|\Delta| < \varepsilon$, terminate
  - Else set $\Delta_i = \Delta_i / \alpha$ --- $\forall i$, go to 2

Example :

- Consider the Himmelblau function:

$$f(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$

- Solution

Step 1 *Selection of initial conditions*

1. Initial Point : $x^{(0)} = (0,0)^T$
2. Increment vector : $\Delta = (0.5, 0.5)^T$
3. Reduction factor : $\alpha = 2$
4. Termination parameter : $\varepsilon = 10^{-3}$
5. Iteration counter : $k = 0$

- Step 2

  Perform an Iteration of the exploratory move with **base point** as $x = x^{(0)}$

  Thus we set $x = x^{(0)} = (0,0)^T \quad and \;\; i = 1$

  The exploratory move will be performed with the following steps

Steps for the *Exploratory move*

Step 1 : Explore the vicinity of the variable x_1

Calculate the function values at three points

$$(x^{(0)} + \Delta_1 x^{(0)})^T = (0.5, 0.5)^T \qquad f^+ = f((0.5, 0.5)^T) = 157.81$$

$$x^{(0)} = (0, 0)^T \qquad f = f((0, 0)^T) = 170$$

$$(x^{(0)} - \Delta_1 x^{(0)})^T = (-0.5, 0.5)^T \qquad f^- = f((-0.5, 0.5)^T) = 171.81$$

Step 2 : Take the Minimum of above function and corresponding point

Step 3 : As $i \neq 1$: all variables are not explored

Increment counter *i=2* and explore second variable.

First iteration completed

Step1: At this point the base point is $x = (0.5, 0)^T$ explore the variable x2 and calculate the function values.

$$f^+ = f((0.5, 0.5)^T) = 144.12$$

$$f = f((0.5, 0)^T) = 157.81$$

$$f^- = f((0.5, -0.5)^T) = 165.62$$

Step 2 : $f_{min} = 144.12$ and point , $x=(0.5,0.5)$

Step 3 : As $i=2$ move to the step 4 of the exploratory move

Step 4 : (of the Exploratory move)

Since $x \neq x^c$ the move is success and we set

$$x = (0.5, 0.5)^T$$

- As the move is success, set $x^{(1)} = x = (0.5, 0.5)^T$ move to step 4

STEP 4 : We set $k=1$ and perform *Pattern move*

$$x^{(2)}{}_p = (x^{(1)} + (x^{(1)} - x^{(0)})^T) = 2(0.5, 0.5)^T - (0,0)^T = (1,1)^T$$

Step 5 : Perform another exploratory move as before and with $x_p^{(2)}$ as the base point.

The new point is $x = (1.5, 1.5)^T$

Set the new point $x^{(2)} = x = (1.5, 1.5)^T$

Step 6 : $f(x^{(2)}) = 63.12$ is smaller than $f(x^{(1)}) = 144.12$

Proceed to next step to perform another pattern move

STEP 4 : Set *k=2* and create a new point

$$x^{(3)}{}_p = (2x^{(2)} - x^{(1)}) = (2.5, 2.5)^T$$

Note: as $x^{(2)}$ *is better than* $x^{(1)}$*, a jump along the direction* $(x^{(2)} - x^{(1)})$ *is made, this will take search closer to true minimum*

STEP 5 : Perform another exploratory move to find any better point around the new point.

Performing the move on both variables we have

New point

$$x^{(3)} = (3.0, 2.0)^T$$

This point is the *true minimum point*

In the example the minimum of the Hookes-Jeeves algorithm happen in two iterations : ***this may not be the case always***

Even though the **minimum point** is ***reached*** there is ***no way*** of finding whether the **optimum** is ***reached*** or ***not***

The algorithm proceeds until the norm if the ***increment vector is small.***

STEP 6 : function value at new point

$$f(x^3) = 0 < f(x^2) = 63.12$$

Thus move on to step 4

Step 2 : Perform an exploratory move with the following as current point $x^{(3)} = (3.0, 2.0)^T$

The exploratory move on both the variables is failure and we obtain $x^{(3)} = (3.0, 2.0)^T$

thus we proceed to Step 3

Step 3 : Since $\|\Delta\|$ is not small reduce the increment vector and move to Step 2.

The new increment vector is $\Delta = (0.125, 0.125)^T$

The algorithm now continues with step 2 and step 3 until $\|\Delta\|$ *is smaller than the termination factor.*

The final solution is $x^* = (3.0, 2.0)^T$ *with the function value 0*

POWELL'S CONJUGATE DIRECTION METHOD

For a quadratic function IN 2 VARIABLES

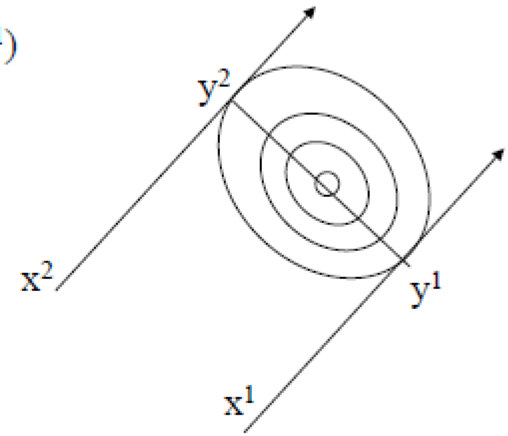
- TAKE 2 POINTS x^1 & x^2 AND
- A DIRECTION 'd'

IF y^1 IS A SOLUTION OF MIN $f(x^1 + \lambda d)$ &

y^2 IS A SOLUTION OF MIN $f(x^2 + \lambda d)$

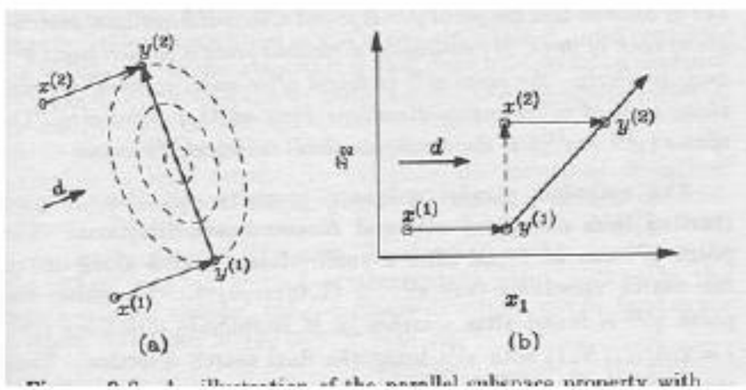
THEN $(y^2 - y^1)$ IS CONJUGATE TO d

OPTIMUM LIES ALONG $(y^2 - y^1)$



Alternate to the above method

- One point (x^1) and both coordinate directions $((1,0)^T \; and \; (0,1)^T)$
- Can be used to create a pair of conjugate directions $(d \; and \; (y^2 - y^1))$



- Point (y^1) obtained by unidirectional search along $(1,0)^T$ from the point (x^1) .
- Point (x^2) obtained by unidirectional search along $(0,1)^T$ from the point (y^1) .
- Point (y^2) obtained by unidirectional search along $(1,0)^T$ from the point(x^2) .

The figure shown also follows the ***Parallel Subspace Property.***

This requires ***Three Unidirectional*** searches.

Thus the point x^p can be written as $x^p=(\alpha,4)^T$

Now the two variable function can be expressed in terms of one variable

$$F(\alpha) = (\alpha^2 - 7)^2 + (\alpha + 9)^2$$

We are looking for the point which the function value is minimum.

== Following the procedure of numerical differentiation. Using the bounding phase method, the minimum is bracketed in the interval $(1,4)$, and using the golden search method the minimum $\alpha^*=2.083$ with three decimals places of accuracy. Thus $x^1 = (2.083,4.00)^T$

Step 3 : According to the parallel subspace property, we find the new conjugate direction

$$s^{(2)} = (1,0)^T$$

Step 4 : the magnitude of search vector d is not small. Thus the new conjugate search direction are

$$s^{(1)} = (0.798,-1.592)^T / \left\|(0.798,-1.592)^T\right\|$$

$$= (0.448,-0.894)^T$$

This completes one iteration of Powell's conjugate direction method.

Step 2 : A single variable minimization along the search direction $s^{(1)}$ from the point $x^{(3)}=(2.881,2.408)^T$ results in the new point x = (3.063,2.045)

One more unidirectional search along the s^2 from the point x^4 results in the point x^5. Another minimization along s^1 results in x^6

Step 3 : the new conjugate direction is

$$d = (x^{(6)} - x^{(4)}) = (0.055,-0.039)^T$$

The unit vector along this direction is (0.816,-0.578)

Step 4 : The new pair of conjugate search direction are

$$s^{(1)} = (0.448, -0.894)^T \qquad s^{(2)} = (0.055, -0.039)^T$$

The search direction d (before normalizing) may be considered to be small and therefore the algorithm may be terminated

UNIT – IV

MULTIVARIABLE CONSTRAINED OPTIMIZATION

Langrangian Method

- The Lagrangian approach transfers a constrained optimization problem into
  - an unconstrained optimization problem and
  - a pricing problem.
- The new function to be optimized is called the **Lagrangian**.
- For each constraint, a *shadow price* is introduced, called a **Lagrange multiplier**.
- In the new unconstrained optimization problem a constraint can be violated, but only at a cost.
- The pricing problem is to find shadow prices for the constraints such that the solutions to the new and the original optimization problem are identical.
- $L(x, y) = f(x, y) + \lambda_1 g_1(x, y) + \lambda_2 g_2(x, y) + \ldots + \lambda_K g_K(x, y)$.

1. Make an informed guess about which constraints are binding at the optimum.
2. Suppose there are k^* such constraints. Set the Lagrange multipliers for all other $k - k^*$ constraints to zero, i.e. ignore these constraints.
3. Solve the first-order conditions

$$\partial\mathcal{L}/\partial x = 0, \partial\mathcal{L}/\partial y = 0,$$

   together with the conditions that the k^* constraints hold with equality. Note that we obtain a system of $k^* + 2$ equations for the same number of unknowns.
4. Check whether the solution is indeed an unconstrained optimum of the Lagrangian. This may be difficult.
5. Check that the Lagrange multipliers are all non-negative and that the solution (x^*, y^*) satisfies all the constraints.
6. If 4) or 5) are violated, start again at 1) with a new guess.

- Suppose we are given numbers $\lambda_1, \lambda_2, ...\lambda_K$ and a pair of numbers (x^*, y^*) such that
  - $\lambda_1, \lambda_2, ...\lambda_K$, i.e. Lagrange multipliers are nonnegative;
  - (x^*, y^*) satisfies all the constraints, i.e. $g_k(x^*, y^*) \geq 0, \forall k = 1, 2, .., K$;
  - (x^*, y^*) is an unconstrained maximum of the Lagrangian $\mathcal{L}$;
  - The complementary slackness conditions $\lambda_k g_k(x^*, y^*) = 0$ are satisfied, i.e. either the kth Lagrange multiplier is zero or the kth constraint binds.
- Then (x^*, y^*) is a maximum for the constrained maximization problem.

- The Lagrangian approach does not immediately tell us, which constraints are binding in the optimum. We have to start with an informed guess using all problem-specific information.
- We write down the Lagrangian assuming that only certain constraints bind.
- We solve the system of simultaneous equations consisting of the FOCs and the binding constraints.
- We check if the solution satisfies the other constraints and that Lagrange multipliers are nonnegative.
- We check if the solution found is an unconstrained optimum of the Lagrangian.

- Suppose only constraint g_1 is binding.
- The FOCs are then

$$\frac{\partial f}{\partial x} = -\lambda_1 \frac{\partial g_1}{\partial x},$$

$$\frac{\partial f}{\partial y} = -\lambda_1 \frac{\partial g_1}{\partial y}.$$

- We can get rid of λ_1 to obtain the following system to solve:

$$\frac{\partial f/\partial x}{\partial f/\partial y} = \frac{\partial g_1/\partial x}{\partial g_1/\partial y},$$

$$g_1(x, y) = 0.$$

Kuhn Tucker Conditions

Given general problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{subject to} \quad & h_i(x) \le 0, \quad i = 1, \ldots m \\ & \ell_j(x) = 0, \quad j = 1, \ldots r \end{aligned}$$

The **Karush-Kuhn-Tucker conditions** or **KKT conditions** are:

- $0 \in \partial f(x) + \sum_{i=1}^{m} u_i \partial h_i(x) + \sum_{j=1}^{r} v_j \partial \ell_j(x)$ (stationarity)
- $u_i \cdot h_i(x) = 0$ for all i (complementary slackness)
- $h_i(x) \le 0, \; \ell_j(x) = 0$ for all i, j (primal feasibility)
- $u_i \ge 0$ for all i (dual feasibility)

Let $x^\star$ and $u^\star, v^\star$ be primal and dual solutions with zero duality gap (strong duality holds, e.g., under Slater's condition). Then

$$\begin{aligned} f(x^\star) &= g(u^\star, v^\star) \\ &= \min_{x \in \mathbb{R}^n} f(x) + \sum_{i=1}^{m} u_i^\star h_i(x) + \sum_{j=1}^{r} v_j^\star \ell_j(x) \\ &\le f(x^\star) + \sum_{i=1}^{m} u_i^\star h_i(x^\star) + \sum_{j=1}^{r} v_j^\star \ell_j(x^\star) \\ &\le f(x^\star) \end{aligned}$$

In other words, all these inequalities are actually equalities

Two things to learn from this:

- The point $x^\star$ minimizes $L(x, u^\star, v^\star)$ over $x \in \mathbb{R}^n$. Hence the subdifferential of $L(x, u^\star, v^\star)$ must contain 0 at $x = x^\star$—this is exactly the **stationarity** condition
- We must have $\sum_{i=1}^m u_i^\star h_i(x^\star) = 0$, and since each term here is ≤ 0, this implies $u_i^\star h_i(x^\star) = 0$ for every i—this is exactly **complementary slackness**

Primal and dual feasibility obviously hold. Hence, we've shown:

> If $x^\star$ and $u^\star, v^\star$ are primal and dual solutions, with zero duality gap, then $x^\star, u^\star, v^\star$ satisfy the KKT conditions

(Note that this statement assumes nothing a priori about convexity of our problem, i.e. of f, h_i, ℓ_j)

If there exists $x^\star, u^\star, v^\star$ that satisfy the KKT conditions, then

$$\begin{aligned} g(u^\star, v^\star) &= f(x^\star) + \sum_{i=1}^{m} u_i^\star h_i(x^\star) + \sum_{j=1}^{r} v_j^\star \ell_j(x^\star) \\ &= f(x^\star) \end{aligned}$$

where the first equality holds from stationarity, and the second holds from complementary slackness

Therefore duality gap is zero (and $x^\star$ and $u^\star, v^\star$ are primal and dual feasible) so $x^\star$ and $u^\star, v^\star$ are primal and dual optimal. I.e., we've shown:

If $x^\star$ and $u^\star, v^\star$ satisfy the KKT conditions, then $x^\star$ and $u^\star, v^\star$ are primal and dual solutions

Quadratic programming

Quadratic with equality constraints

Consider for $Q \succeq 0$,

$$\min_{x \in \mathbb{R}^n} \ \frac{1}{2} x^T Q x + c^T x$$
$$\text{subject to} \ \ Ax = 0$$

E.g., as in Newton step for $\min_{x \in \mathbb{R}^n} f(x)$ subject to $Ax = b$

Convex problem, no inequality constraints, so by KKT conditions: x is a solution if and only if

$$\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} -c \\ 0 \end{bmatrix}$$

for some u. Linear system combines stationarity, primal feasibility (complementary slackness and dual feasibility are vacuous)

Example from B & V page 245: consider problem

$$\min_{x\in\mathbb{R}^n} \; -\sum_{i=1}^{n} \log(\alpha_i + x_i)$$
$$\text{subject to } \; x \geq 0, \; 1^T x = 1$$

Information theory: think of $\log(\alpha_i + x_i)$ as communication rate of ith channel. KKT conditions:

$$-1/(\alpha_i + x_i) - u_i + v = 0, \quad i = 1, \ldots n$$
$$u_i \cdot x_i = 0, \quad i = 1, \ldots n, \quad x \geq 0, \quad 1^T x = 1, \quad u \geq 0$$

Eliminate u:

$$1/(\alpha_i + x_i) \leq v, \quad i = 1, \ldots n$$
$$x_i(v - 1/(\alpha_i + x_i)) = 0, \quad i = 1, \ldots n, \quad x \geq 0, \quad 1^T x = 1$$

Let's return the lasso problem: given response $y \in \mathbb{R}^n$, predictors $A \in \mathbb{R}^{n \times p}$ (columns $A_1, \ldots A_p$), solve

$$\min_{x \in \mathbb{R}^p} \frac{1}{2}\|y - Ax\|^2 + \lambda\|x\|_1$$

KKT conditions:

$$A^T(y - Ax) = \lambda s$$

where $s \in \partial\|x\|_1$, i.e.,

$$s_i \in \begin{cases} \{1\} & \text{if } x_i > 0 \\ \{-1\} & \text{if } x_i < 0 \\ [-1, 1] & \text{if } x_i = 0 \end{cases}$$

Now we read off important fact: if $|A_i^T(y - Ax)| < \lambda$, then $x_i = 0$

... we'll return to this problem shortly

UNIT – V

GEOMETRIC AND INTEGER PROGRAMMING

Integer Programming

Integer programming is a branch of mathematical programming or optimization.

A general mathematical programming problem can be stated as

$$\max f(x) \qquad x \in S \subset I\!R^n, \tag{1}$$

where f is called the objective function and it is a function defined on S, and S is the so-called constraint set or admissible set.

Every $x \in S$ is called a feasible solution. Moreover, if there is $x^\star$ such that

$$\infty > f(x^\star) \geq f(x)$$

for all $x \in S$, then $x^\star$ is called an optimal solution to (1).

The goal of mathematical programming is to establish if an optimal solution exists and to find one, or all, optimal solutions.

An integer programming problem is a mathematical programming problem in which

$$S \subset Z^n \subset \mathbb{R}^n,$$

where Z^n is the set of all n-dimensional vectors with integer components.

A mixed integer programming problem is a mathematical programming problem in which at least one, but not all, of the components of $x \in S$ are required to be integer.

From an applied point of view, it is convenient to regard problem (1) as a model of decision making in which S represents the set of admissible decisions and f assigns a utility or profit to each $x \in S$.

The problem (1) is called a linear programming (LP) problem if

$$f = cx \qquad S = \{x \mid Ax = b, x \geq 0\},$$

where $c \in \mathbb{R}^{1\times n}$, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^{m\times 1}$. Moreover, the inequality $x \geq 0$ has to be understood componentwise, *i.e.* $x_i \geq 0$ for all i.

Note that the set S is convex, *i.e.* if $x \in S$ and $y \in S$ then $\alpha x + (1-\alpha)y \in S$ for all $\alpha \in [0, 1]$.

A set defined by linear constraints is called a polyhedron or a polytope.

Minimization problems can be rewritten as maximization problems noting that

$$-\min(-f(x)) = \max f(x).$$

Inequality constraints can be converted into equality constraints by adding auxiliary variables. For example

$$ax \leq b \quad \Leftrightarrow \quad ax + s = b \quad s \geq 0,$$

and

$$ax \geq b \quad \Leftrightarrow \quad ax - t = b \quad t \geq 0.$$

The variables s and t are known as slack or surplus variables.

The LP problem obtained by dropping the integrality constraint from the ILP problem (2) will be referred to as the corresponding LP problem.

In general, the problem

$$P_1 : \max f(x) \qquad x \in S_1$$

is said to be a relaxation of the problem

$$P_2 : \max f(x) \qquad x \in S_2$$

if

$$S_1 \supseteq S_2.$$

Similarly, P_2 is said to be a restriction of P_1.

The concepts of relaxation and restriction are often used in mathematical programming. Note that if x° is an optimal solution to P_1 and $x^\star$ is an optimal solution to P_2 then

$$f(x^\circ) \geq f(x^\star).$$

Moreover, if $x^\circ \in S_2$ then x° is an optimal solution to P_2.

An important special case of the ILP problem is the so-called binary ILP problem described by

$$\begin{aligned} &\max cx \\ Ax &= b \\ x &\geq 0 \text{ binary.} \end{aligned} \tag{3}$$

(x binary means $x_i = 0$ or $x_i = 1$ for all i.)

Capital budgeting. A firm has n projects to undertake but, because of budget restrictions, not all can be selected.

Project j has a present value of c_j, and requires an investment of a_{ij} in the time period i, where $i = 1, \cdots, m$. The capital available in time period i is b_i.

The problem of maximizing the total present value subject to the budget constraints can be written as

$$\begin{array}{ll} \max & \sum_{j=1}^{m} c_j x_j \\ & \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \qquad i = 1, \cdots, m \\ & x_j = 0, 1, \qquad j = 1, \cdots, n \end{array}$$

where $x_j = 1$ if the project j is selected and $x_j = 0$ if the project j is not selected.

Dichotomies. Consider the problem max $f(x)$ with $x \in S$ subject to

$$g(x) \geq 0 \ \text{ or } \ h(x) \geq 0. \tag{4}$$

This is in general a difficult problem. However, the dichotomy (4) is equivalent to

$$\begin{aligned} g(x) &\geq \delta \underline{g} \\ h(x) &\geq (1-\delta)\underline{h} \\ \delta & \quad \text{binary}, \end{aligned}$$

where $\underline{g}$ and $\underline{h}$ are known finite lower bounds on g and h. In fact,

$$\delta = 0 \Rightarrow g(x) \geq 0 \ \text{ and } \ h(x) \geq \underline{h}$$

$$\delta = 1 \Rightarrow g(x) \geq \underline{g} \ \text{ and } \ h(x) \geq 0.$$

The fixed charge problem. In general the cost of an activity is a nonlinear function of the activity level x, given by

$$f(x) = \begin{cases} d + cx & \text{if } x > 0 \\ 0 & \text{if } x = 0. \end{cases}$$

If $d > 0$ and f is to be minimized, we have the problem

$$\begin{aligned} &\min cx + dy \\ x &\geq 0 \\ x - uy &\leq 0 \\ y &= 0, 1, \end{aligned}$$

where y is an indicator of whether or not the activity is undertaken, and u is a known, finite, upper bound for x. The second constraint guarantees that $x > 0$ implies $y = 1$.

The plant location problem. Consider n customers, the j-th one requiring b_j units of a commodity. There are m locations in which plants may operate to satisfy the demands.

There is a fixed charge of d_i for opening plant i, and the unity cost for supplying customer j from plant i is c_{ij}. The capacity of plant i is h_i.

The problem is

$$\min \sum_{i=1}^{m} \left(\sum_{j=1}^{n} c_{ij} x_{ij} + d_i y_i \right)$$

$$\sum_{j=1}^{m} x_{ij} = b_i$$

$$\sum_{j=1}^{n} x_{ij} - h_i y_i \leq 0$$

$$x_{ij} \geq 0, \qquad y_i = 0,\ 1.$$

The knapsack problem. Suppose n different types of scientific equipment are considered for inclusion on a space vehicle.

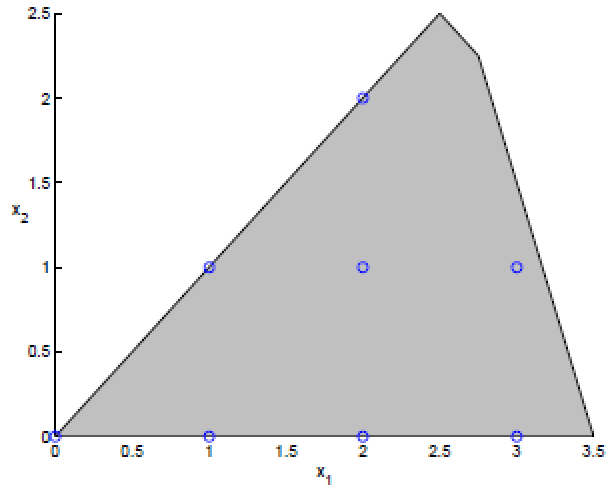
Let c_j be the scientific value per unit and a_j the weight per unit of the j-th type.

If the total weight limitation is b, the problem of maximizing the total value of the equipment taken is

$$\max \sum_{j=1}^{n} c_j x_j$$
$$\sum_{j=1}^{n} a_j x_j \leq b$$
$$x_j \geq 0, \quad \text{integer},$$

where x_j is the number of units of the j-th type included.

The feasible region is the shaded area in the figure.



The optimum of the relaxed (non-integer) problem is located at $(\frac{11}{4}, \frac{9}{4})$ with a value of the objective function equal to $7\frac{3}{4}$.

Enumeration. Without plotting the admissible set it is possible to obtain an upper bound on the number of feasible points.

The first constraint, together with nonnegativity of the x_i, implies $0 \leq x_i \leq 5$.

The third constraint implies $0 \leq x_1 \leq 3$.

This limits the feasible points to 24 (16 are infeasible, and 8 are feasible).

By total enumeration one could find the optimal point $(x_1, x_2) = (3, 1)$.

With some work one can reduce the number of candidate optimal solutions.

Adding the first and second constraints yields $2x_2 \leq 5$, which implies $x_2 \leq 2$, and reduces the upper bound on the number of feasible points to 12.

Note that the feasible point $(3, 0)$ yields a value of the objective function equal to 6. Thus every optimal solution should be such that $2x_1 + x_2 \geq 6$.

The above, together with $x_2 \leq 2$ yields $2x_1 \geq 4$.

In summary, we have reduced the number of candidate optimal points to 6:

$$(2, 0) \quad (2, 1) \quad (2, 2) \quad (3, 0) \quad (3, 1) \quad (3, 2).$$

Of these points, $(2, 0)$ and $(2, 1)$ yields a value of the objective smaller than 6.

Moreover, since the non-integer optimum of the objective is $7\frac{3}{4}$, it follows that $2x_1 + x_2 \leq 7$, which rules out $(3, 2)$.

The candidates for optimality have been reduced to

$$(2, 2) \quad (3, 0) \quad (3, 1),$$

from which, by direct computation, one obtains the optimum $(3, 1)$.

The main idea of enumeration methods is thus to explore, explicitly or implicitly, a set of integer points containing the set of admissible points.

Suppose the set $S = \{x \mid Ax = b, x \geq 0 \text{ integer}\}$ of feasible solutions of an ILP problem is bounded, hence contains a finite number of points.

Define the convex hull of S, namely

$$S^+ = \{y \mid y = \sum \alpha_i x_i,\ \alpha \geq 0,\ \sum \alpha = 1, x_i \in S\}.$$

Then

$$S \subseteq S^+ \subseteq T = \{x \mid Ax = b, x \geq 0\}$$

and the optimal solution of

$$\max cx \qquad x \in S$$

can be computed solving

$$\max cx \qquad x \in S^+.$$

The computation of S^+ is in general very difficult, and involves several cuts.

In practice, a small number of *good* cuts is enough to generate a LP problem with an integer solution, which coincides with the solution of the given ILP problem.

For the considered example, from the optimal solution of the corresponding LP problem one has

$$2x_1 + x_2 \leq 7\frac{3}{4} \Rightarrow 2x_1 + x_2 \leq 7.$$

Moreover

$$2x_1 + x_2 \leq 7 \text{ and } x_2 \geq 0 \Rightarrow 2x_1 \leq 7 \Rightarrow x_1 \leq 3.$$

$$\max 2x_1 + x_2$$

$$x_1 + x_2 \leq 5 \qquad -x_1 + x_2 \leq 0 \qquad 6x_1 + 2x_2 \leq 21$$

$$x_i \geq 0 \quad \text{integer}$$

and

$$\max 2x_1 + x_2$$

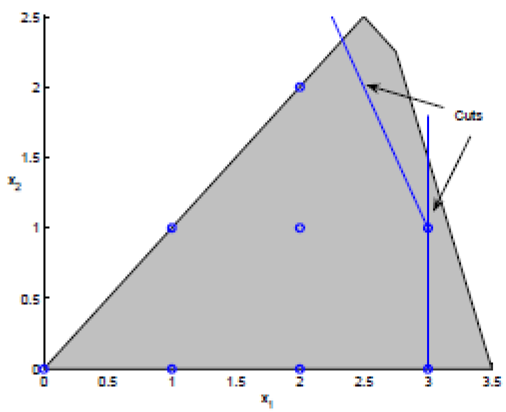
$$x_1 + x_2 \leq 5 \qquad -x_1 + x_2 \leq 0 \qquad 6x_1 + 2x_2 \leq 21$$

$$x_1 \leq 3 \qquad 2x_1 + x_2 \leq 7$$

$$x_i \geq 0$$

have the same optimal solution (the point $(3, 1)$) which is integer.


S+


Cuts
x2
x1

Thank you