Research Methodology Tools:
2. Optimization

What is Management Science / Operations Research?

Operations Research and Management Science means:

What is Management Science / Operations Research?

Operations Research and Management Science means:
the use of mathematical models in providing guidance to management in order to make an effective and wise decisions based on available information, or seek additional information if the information is not sufficient to make accurate decisions.

Definition

Operations Research (Management Science) is a scientific approach to decision making that seeks to best design and operate a system, usually under conditions requiring the allocation of scarce resources.

Definition

Operations Research (Management Science) is a scientific approach to decision making that seeks to best design and operate a system, usually under conditions requiring the allocation of scarce resources.

A system is an organization of interdependent components that work together to accomplish the goal of the system.

Definition

Operations Research (Management Science) is a scientific approach to decision making that seeks to best design and operate a system, usually under conditions requiring the allocation of scarce resources.

A system is an organization of interdependent components that work together to accomplish the goal of the system.

Modeling: The scientific approach to decision making requires the use of one or more mathematical models.

Definition

Operations Research (Management Science) is a scientific approach to decision making that seeks to best design and operate a system, usually under conditions requiring the allocation of scarce resources.

A system is an organization of interdependent components that work together to accomplish the goal of the system.

Modeling: The scientific approach to decision making requires the use of one or more mathematical models.

A mathematical model is a mathematical representation of the actual situation that may be used to make better decisions or clarify the situation.

Operations Research Over the Years

- 1947: Project Scoop (Scientific Computation of Optimum Programs) with George Dantzig and others. Developed the simplex method for linear programs.
- 1950's: Lots of excitement, mathematical developments, queuing theory, mathematical programming. (A.I. in the 1960's)
- 1970's: Disappointment, and a settling down. NP-completeness. More realistic expectations.
- 1980's: Widespread availability of personal computers. Increasingly easy access to data. Widespread willingness of managers to use models.
- 1990's: Improved use of O.R. systems.Further inroads of O.R. technology, e.g., optimization and simulation add-ins to spreadsheets, modeling languages, large scale optimization. More intermixing of A.I. and O.R.

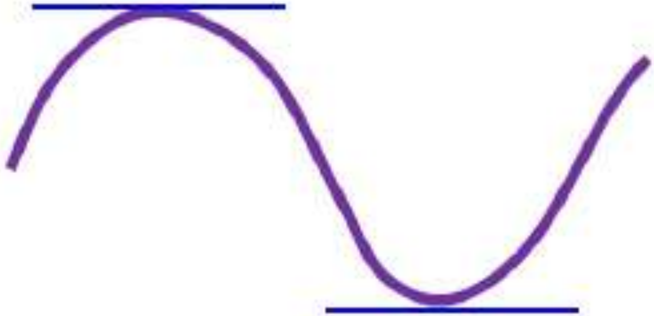
Operations Research in the 00s

- Lots of opportunities for OR as a field
- Data, data, data
  - E-business data (click stream, purchases, other transactional data, E-mail and more)
  - The human genome project and its outgrowth
- Need for more automated decision making
- Need for increased coordination for efficient use of resources (Supply chain management)

OPTIMIZATION

As Ageless as time

CALCULUS



Fermat, Newton, Euler, Lagrange, Gauss, and more

Points to Ponder

- Construction of Panama Canal (1880s)
- Golden Gate Bridge (San Francisco) (1930's)
    - Using Slide Rules
    - No Computer, No Calculators!!
- Eiffel Tower
- Mostar Bridge Bosnia (400 Years ago)

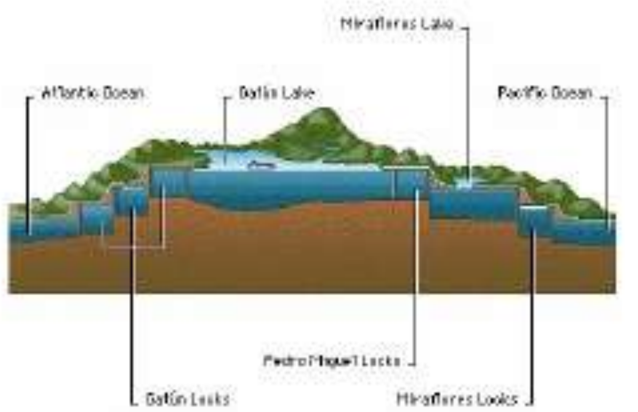


Figure: Construction Panama Canal. Source: Wikipedia

Figure: Panama Canal. Source: Google image

Figure: Golden Gate Bridge. Source: Google image

Figure: Eiffel Tower. Source: Google image

Figure: Eiffel Tower. Source: Google image

Figure: Mostar Bridge. Source: Google image

Some Themes

- Optimization is everywhere
- Needs scientific or mathematical model
- The purpose of the model gives an insight to the actual situation or physical problem
- The appropriate and efficient algorithms are very important

Optimization is everywhere

It is embedded in language, and part of the way we think.

- Firms want to maximize value to shareholders
- People want to make the best choices
- We want the highest quality at the lowest price
- When playing games, we want the best strategy
- When we have too much to do, we want to optimize the use of our time
- etc.

Mathematical Optimization is nearly everywhere

- Finance
- Marketing
- E-business
- Telecommunications
- Games
- Operations Management
- Production Planning
- Transportation Planning
- System Design

Look for it! You will see opportunities for its use.

Optimization Tools

Some goals :

- To present a variety of tools for optimization
- Illustrate applications in manufacturing, finance, e-business, marketing and more.
- Prepare ourselves to recognize opportunities for mathematical optimization as they arise

Some Success Stories

- Optimal crew scheduling saves American Airlines $20 million/yr.
- Improved shipment routing saves Yellow Freight over $17.3 million/yr.
- Improved truck dispatching at Reynolds Metals improves on-time delivery and reduces freight cost by $7 million/yr.
- Optimizing global supply chains saves Digital Equipment over $300 million.
- Restructuring North America Operations, Proctor and Gamble reduces plants by 20%, saving $200 million/yr.
- GTE local capacity expansion saves $30 million/yr.

Some Success Stories

- Optimal traffic control of Hanshin Expressway in Osaka saves 17 million driver hours/yr.
- Better scheduling of hydro and thermal generating units saves southern company $140 million.
- Improved production planning at Sadia (Brazil) saves $50 million over three years.
- Production Optimization at Harris Corporation improves on-time deliveries from 75% to 90%.
- Tata Steel (India) optimizes response to power shortage contributing $73 million.
- Optimizing police patrol officer scheduling saves police department $11 million/yr.
- Gasoline blending at Texaco results in saving of over $30 million/yr.

What is optimization about?

1. Is to find the BEST solution from the domain or feasible set of solution.
   - Feasible set or region is a set that contain all solution point.
2. Provide mathematical tools that allow to search the best solution in efficient way.
3. However, before these tolls can be applied the design problem needs to be formulate in an appropriate form.

There are 2 type of the optimization.

- Unconstrained Optimization
  - No restrictions are imposed in decision variables
  - Either Maximize or Minimize
- Constrained Optimization
  - There are a lot of restriction are imposed in decision variables.
  - Either Maximize or Minimize.

Basic of optimization Problem

In mathematical point of view, its can be written as

In mathematical point of view, its can be written as

$$Min_{x\in\Omega} \quad f(x) \quad \rightarrow \text{UO Problem} \tag{1}$$

In mathematical point of view, its can be written as

$$Min_{x\in\Omega} \quad f(x) \quad \rightarrow \text{UO Problem} \qquad (1)$$

$$\left.\begin{array}{ll} Min & f(x) \\ g(x) \in S_1, & x \in S_2 \end{array}\right\} \rightarrow \text{CO Problem} \qquad (2)$$

where

In mathematical point of view, its can be written as

$$Min_{x\in\Omega} \quad f(x) \quad \rightarrow \text{UO Problem} \tag{1}$$

$$\left.\begin{array}{ll} Min & f(x) \\ g(x)\in S_1, & x\in S_2 \end{array}\right\} \rightarrow \text{CO Problem} \tag{2}$$

where

- $\mathbf{x} = (x_1, \cdots, x_n)^T \in \Re$ is decision variables. $\mathbf{x}$ is chosen so that the objective $f(x)$ is optimized (maximum or minimum).

In mathematical point of view, its can be written as

$$Min_{x\in\Omega} \quad f(x) \quad \rightarrow \text{UO Problem} \tag{1}$$

$$\left.\begin{array}{ll} Min & f(x) \\ g(x) \in S_1, & x \in S_2 \end{array}\right\} \rightarrow \text{CO Problem} \tag{2}$$

where

- $\mathbf{x} = (x_1, \cdots, x_n)^T \in \Re$ is decision variables. $\mathbf{x}$ is chosen so that the objective $f(x)$ is optimized (maximum or minimum).
- $f(x)$ is called the objective function. It is what will be maximized or minimized.

Basic of optimization Problem

- $g(x)$ is called constraint function. It is the restriction.

Basic of optimization Problem

- $g(x)$ is called constraint function. It is the restriction.
- In choosing **x** the choice is made subject to a set of constraints $g(x) \in S_1$ and $x \in S_2$ must be obeyed.

- $g(x)$ is called constraint function. It is the restriction.
- In choosing **x** the choice is made subject to a set of constraints $g(x) \in S_1$ and $x \in S_2$ must be obeyed.
- However, the constrained optimization can be transformed to unconstrained optimization problem.

$$L(x, \lambda_1, \lambda_2) = f(x) + \lambda_1 g(x) + \lambda_2 x$$

Basic of optimization Problem

- $g(x)$ is called constraint function. It is the restriction.
- In choosing **x** the choice is made subject to a set of constraints $g(x) \in S_1$ and $x \in S_2$ must be obeyed.
- However, the constrained optimization can be transformed to unconstrained optimization problem.

$$L(x, \lambda_1, \lambda_2) = f(x) + \lambda_1 g(x) + \lambda_2 x$$

- Maximize $f(x) = -$ Minimize $f(x)$

Components of Optimization

Now, we describe the components of optimization model

- Objective Function

Now, we describe the components of optimization model

- Objective Function
  - Represents the quantity, f which is to be maximize or minimize as a function of one or more variable $(x_1, \cdots, x_n)$

Now, we describe the components of optimization model

- Objective Function
  - Represents the quantity, f which is to be maximize or minimize as a function of one or more variable $(x_1, \cdots, x_n)$
- Decision Variables

Now, we describe the components of optimization model

- Objective Function
  - Represents the quantity, f which is to be maximize or minimize as a function of one or more variable $(x_1, \cdots, x_n)$
- Decision Variables
  - The independent variable $(x_1, \cdots, x_n)$ that the objective function depends on. They tell us how much of something to do:

Now, we describe the components of optimization model

- Objective Function
  - Represents the quantity, f which is to be maximize or minimize as a function of one or more variable $(x_1, \cdots, x_n)$
- Decision Variables
  - The independent variable $(x_1, \cdots, x_n)$ that the objective function depends on. They tell us how much of something to do:
    - Shares to be purchased
    - Length and width an electronics board
    - Truckloads of oil to move

Now, we describe the components of optimization model

- Objective Function
  - Represents the quantity, f which is to be maximize or minimize as a function of one or more variable $(x_1, \cdots, x_n)$
- Decision Variables
  - The independent variable $(x_1, \cdots, x_n)$ that the objective function depends on. They tell us how much of something to do:
    - Shares to be purchased
    - Length and width an electronics board
    - Truckloads of oil to move
  - Generally assumed to be nonnegative & continuous.

Now, we describe the components of optimization model

- Objective Function
  - Represents the quantity, f which is to be maximize or minimize as a function of one or more variable $(x_1, \cdots, x_n)$
- Decision Variables
  - The independent variable $(x_1, \cdots, x_n)$ that the objective function depends on. They tell us how much of something to do:
    - Shares to be purchased
    - Length and width an electronics board
    - Truckloads of oil to move
  - Generally assumed to be nonnegative & continuous.
  - Sometimes they are problematic. For instance, when the items modeled can not have a fractional part.

Components of Optimization

- Constraints

Components of Optimization

- Constraints
  - Restriction
    1. how much of a resource can be used: Amount of money to be invested
    2. what must be done: hour of labor or production requirements

- Constraints
  - Restriction
    1. how much of a resource can be used: Amount of money to be invested
    2. what must be done: hour of labor or production requirements
  - Comes in two mathematical sign
    1. Equality constraints $(=) \rightarrow$ often come from fundamental physics considerations. For instance: Kirchhoff's law, First law of thermodynamics, etc.
    2. Inequality constraints $(\leq, \geq) \rightarrow$ cost, space, material, labor, etc.

Mathematical Formulation

Let us consider

Mathematical Formulation

Let us consider

$$Minimize \quad z = f(x_1, x_2, \cdots, x_n) \tag{3}$$

subject to

Mathematical Formulation

Let us consider

$$Minimize \quad z = f(x_1, x_2, \cdots, x_n) \tag{3}$$

subject to

$$g_i(x_1, ..., x_n) - b_i = 0, \quad i = 1, ..., m \tag{4}$$

and

$$(x_1, x_2, ..., x_n) \geq (0, 0, ..., 0) \tag{5}$$

Let us consider

$$Minimize \quad z = f(x_1, x_2, \cdots, x_n) \tag{3}$$

subject to

$$g_i(x_1, ..., x_n) - b_i = 0, \quad i = 1, ..., m \tag{4}$$

and

$$(x_1, x_2, ..., x_n) \geq (0, 0, ..., 0) \tag{5}$$

- If $n > m \rightarrow$ a minimization problem results
- If $n = m \rightarrow$ a unique solution exists.
- If $n < m \rightarrow$ no solution which satisfies all of the constraints.

Classification of Optimization Techniques

- Calculus Based Techniques
  - Lagrange multiplier, Kuhn-Tucker, Penalty etc.
- Search Techniques
  - Elimination Techniques: Golden Section Search & Fibonacci
  - Hill-Climbing Techniques: Steepest descent, Newton's Method, Quasi-Newton, etc.
- Programming Techniques
  - Linear Programming, Convex programming, Nonlinear Programming, Dynamic Programming, Integer Programming, Geometric Programming, Non-Linear Programming, etc.

Classification

1. Calculus Based Techniques

1. Calculus Based Techniques
   - Related to principle of the simple Min and Max techniques that you used in calculus.
     - Require all equation to be differentiable and continuous.
     - There are limitations.

1. Calculus Based Techniques
   - Related to principle of the simple Min and Max techniques that you used in calculus.
     - Require all equation to be differentiable and continuous.
     - There are limitations.
   - Fundamental concepts to optimization techniques. For $f'(c) = 0$ or $f'(c)$ does not exist then c is called as critical or extreme point.
     - If $f''(c) > 0 \Rightarrow c$ is minimum point
     - If $f''(c) < 0 \Rightarrow c$ is maximum point.

Fundamental definition in calculus.

Definition

Absolute or Global Minimum: $f(x)$ defined over a closed set X in E^n is said to take on its absolute or global minimum over X at the point x^* if $f(x) \geq f(x^*)\ \forall x \in X$

Definition

Strong Relative or Local Minimum: $f(x)$ is said to have a strong relative minimum or local minimum at x^* if there exists an ϵ, $0 < \epsilon < \delta$, such that $\forall x$, $0 < \|x - x^*\| < \delta$, $f(x) > f(x^*)$

Classification

- Global Minimum at x^* implies that the function takes on its lowest value at that point no matter where else we may search in the set X

Classification

- Global Minimum at x^* implies that the function takes on its lowest value at that point no matter where else we may search in the set X
- Strong relative or local minimum. Only guarantees that the value of $f(x^*)$ is a minimum with respect to other point nearby, specifically in an ϵ-region about x^*.

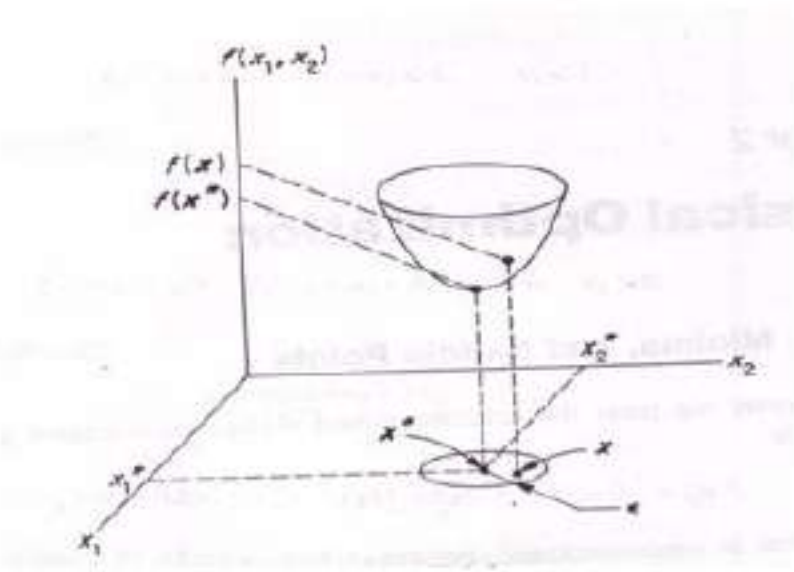


Figure: The Global Minimum

1.1 Lagrange Multipliers

1.1 Lagrange Multipliers

Let us consider problem (3) - (5), with $n > m$. To solve the problem, we introduced the *lagrange multiplier*, λ. The problem becomes:

pause

$$L(x_1, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_m) \equiv f(x_1, ..., x_n) - \sum_{i=1}^{m} \lambda_i \left(g_i(x) - b_i \right) \quad (6)$$

where $\lambda_i (i = 1, ..., m)$. Then solve the system of $(n + m)$ equations

1.1 Lagrange Multipliers

Let us consider problem (3) - (5), with $n > m$. To solve the problem, we introduced the *lagrange multiplier*, λ. The problem becomes:

pause

$$L(x_1, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_m) \equiv f(x_1, ..., x_n) - \sum_{i=1}^{m} \lambda_i \Big(g_i(x) - b_i \Big) \quad (6)$$

where $\lambda_i (i = 1, ..., m)$. Then solve the system of $(n + m)$ equations

$$\frac{\partial L}{\partial x_j} = 0 \quad (j = 1, ..., n) \qquad (7)$$

1.1 Lagrange Multipliers

Let us consider problem (3) - (5), with $n > m$. To solve the problem, we introduced the *lagrange multiplier*, λ. The problem becomes:

pause

$$L(x_1, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_m) \equiv f(x_1, ..., x_n) - \sum_{i=1}^{m} \lambda_i \Big(g_i(x) - b_i \Big) \quad (6)$$

where $\lambda_i (i = 1, ..., m)$. Then solve the system of $(n + m)$ equations

$$\frac{\partial L}{\partial x_j} = 0 \quad (j = 1, ..., n) \quad (7)$$

$$\frac{\partial L}{\partial \lambda_i} = 0 \quad (i = 1, ..., m) \quad (8)$$

Example

Given a fixed area of cardboard, we wish to construct a closed cardboard box with maximum volume. We can formulate and solve this problem using Lagrange Method. Denote the dimensions of the box with maximum volume by x_1, x_2 and x_3, and let the given fixed area of cardboard be A. The problem then can be formulated as:

maximize $\quad x_1x_2x_3$

subject to $\quad x_1x_2 + x_2x_3 + x_3x_1 = \frac{A}{2}$

Solution

Lagrangian function:

$L(x_1, x_2, x_3, \lambda) = x_1 x_2 x_3 - \lambda(x_1 x_2 + x_2 x_3 + x_3 x_1 - \frac{A}{2})$

$$\frac{\partial L}{\partial x_1} = x_2 x_3 - \lambda(x_2 + x_3) = 0 \tag{9}$$

$$\frac{\partial L}{\partial x_2} = x_1 x_3 - \lambda(x_1 + x_3) = 0 \tag{10}$$

$$\frac{\partial L}{\partial x_3} = x_1 x_2 - \lambda(x_1 + x_2) = 0 \tag{11}$$

$$\frac{\partial L}{\partial \lambda} = x_1 x_2 + x_2 x_3 + x_3 x_1 - \frac{A}{2} = 0 \tag{12}$$

Example

Consider the optimal allocation of a scarce resource between two process where the total amount of resource available is b (see figure below). If $f_1(x_1) = 50(x_1 - 2)^2$, and $f_2(x_2) = 50(x_2 - 2)^2$. Maximize the return from both process.

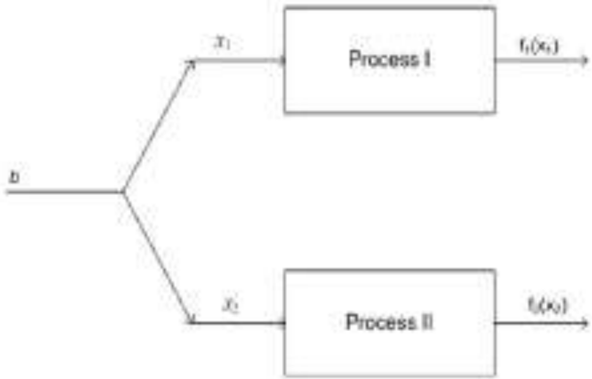


Figure: Two process

1.2 Kuhn-Tucker Theorem

- Similar with Lagrange Method.
- Sign of constraints ($\leq$, or $\geq$)
- Necessary condition of (x^*, λ^*) becomes optimal solution are:
  1. $\frac{\partial L}{\partial x_i^*} = 0$
  2. $\frac{\partial L}{\partial \lambda_i^*} \leq 0$ or $\frac{\partial L}{\partial \lambda_i^*} > 0$
  3. $(\lambda^*)^T g_i(x^*) = 0$
- Unfortunately Kuhn-Tucker does not give any information to obtain (x^*, λ^*).

2. **Search Techniques**

You keep trying different solutions and searching for the best:

- More flexible
- Provided algorithm which help to find the best solution
- Useful when a design variable only take on certain discrete values
- Find the best solution in a finite amount of iterations and time.

Several well-known methods:

1. Steepest Descent Method
2. Newton's method
3. Quasi-Newton Method

2.1 **Steepest Descent Method**

1. Locally but super linear convergence
2. Based on gradient methods
3. Needs initial point

New point could be calculated by

$$x^{(k+1)} = x^k - \alpha_k \nabla f(x^k) \tag{13}$$

where α is step size and can be determined by

$$\alpha_k = arg\ min_{\alpha \geq 0}\ f(x^k - \alpha_k \nabla f(x^k)). \tag{14}$$

Search Techniques

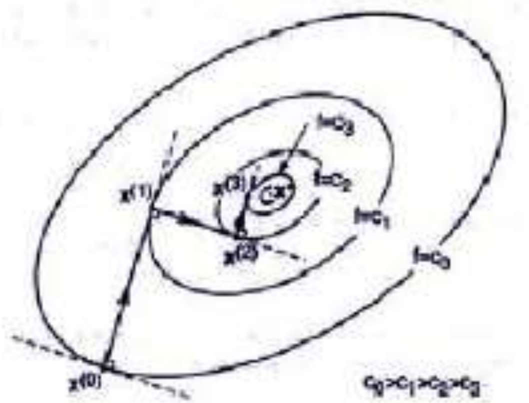


Figure: Procedure of Steepest Descent

Algorithm of Steepest descent method:

1. Start at an initial point, $x^{(0)}$
2. Determine a step size, α_k
3. Find the first order derivative of f at $x^{(0)}$, $\nabla f(x^k)$.
4. Generate new point, by (13)
5. Repeat the process until satisfied the minimum tolerance, $\|\nabla f(x^k)\| < \epsilon$ where ϵ might be 5×10^{-5}. Alternatively, we may compute the absolute difference $|f(x^{(k+1)} - f(x^{(k)})|$ between objective function values for every two successive iterations. Stop if $|f(x^{(k+1)} - f(x^{(k)})| < \epsilon$.

Example

We use the method of steepest descent to find the minimizer of $f(x_1, x_2, x_3) = (x_1 - 4)^4 + (x_2 - 3)^2 + 4(x_3 + 5)^4$. The initial point is $x^{(0)} = [4, 2, 1]^T$. We perform three iterations as follows:

We find,

$$\nabla f(x) = [4(x_1 - 4)^3, 2(x_2 - 3), 16(x_3 + 5)^3]^T.$$

Hence,

$$\nabla f(x^{(0)}) = [0, -2, 1024]^T$$

To compute $x^{(1)}$, we need

$$\begin{aligned} \alpha_0 &= arg\ min_{\alpha \geq 0}\ f(x^{(0)} - \alpha \nabla f(x^{(0)}) \\ &= arg\ min_{\alpha \geq 0} \left(0 + (2 + 2\alpha - 3)^2 + 4(-1 - 1024\alpha + 5)^4\right) \\ &\leq arg\ min_{\alpha \geq 0}\ \varphi_0(\alpha). \end{aligned}$$

By differentiating α_0 w.r.t. α and make it equal to zero, we obtain $\alpha_0 = 3.967 \times 10^{-3}$, $\alpha = 0.38758 \times 10^{-2} \pm j0.52720 \times 10^{-4}$. Since $\alpha > 0$, so, we choose $\alpha_0 = 3.967 \times 10^{-3} =$

Search Techniques

For illustrative purpose, we show a plot of $\varphi_0(\alpha)$ versus α in Figure 1.1 below, obtained using Maple.13

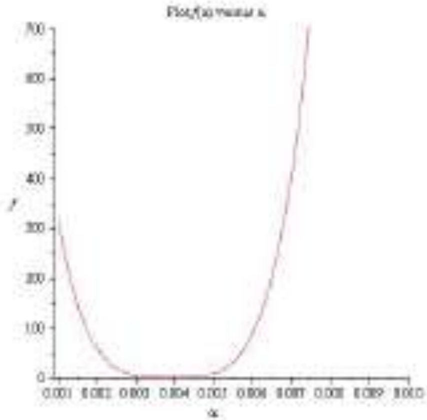


Figure: 1.1

Thus

$$x^{(1)} = x^{(0)} - \alpha_0 \nabla f(x^{(0)}) = [4.000, 2.008, -5.062]^T$$

To find $x^{(2)}$, we first determine

$$\nabla f(x^{(1)}) = [0.000, -1.984, -0.003875]^T.$$

Next we find α_1, where

$$\begin{aligned} \alpha_1 &= arg\ min_{\alpha \geq 0}\ (0 + (2.008 + 1.984\alpha - 3)^2 + \\ &\quad 4(-5.062 + 0.003875\alpha + 5)^4) \\ &= arg\ min_{\alpha \geq 0}\ \varphi_1(\alpha). \end{aligned}$$

Similarly, we obtain $\alpha_1 = 0.5000$. Figure 1.2 depicts a plot of $\varphi_1(\alpha)$ versus α

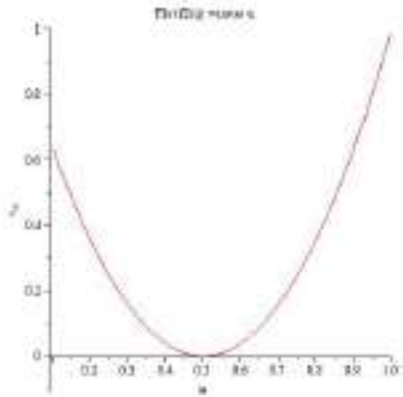


Figure: 1.2

Thus,

$$x^{(2)} = x^{(1)} - \alpha_1 \nabla f(x^{(1)}) = [4.000, 3.000, -5.060]^T$$

To find $x^{(3)}$, we first determine

$$\nabla f(x^{(2)}) = [0.000, 0.000, -0.003525]^T.$$

and

$$\begin{aligned}\alpha_2 &= arg\ min_{\alpha \geq 0} \left(0 + 0 + 4(-5.060 + 0.003525\alpha + 5)^4\right) \\ &= arg\ min_{\alpha \geq 0}\ \varphi_2(\alpha).\end{aligned}$$

We proceed as in the previous iterations to obtain $\alpha_2 = 16.29$. The value of $x^{(3)}$ is

$$x^{(3)} = x^{(2)} - \alpha_2 \nabla f(x^{(2)}) = [4.000, 3.000, -5.002]^T.$$

Note that the minimizer of f is $[4.000, 3.000, -5.002]^T$, and hence it appears that we have arrived at the minimizer in only three iterations.

2.2 Newton's Method

Recall that the method of SD uses only first derivatives (gradients) in selecting a suitable search direction. This strategy is not always the most effective.

2.2 Newton's Method

Recall that the method of SD uses only first derivatives (gradients) in selecting a suitable search direction. This strategy is not always the most effective.

If higher derivatives are used, the resulting iterative algorithm may perform better than the SD method.

2.2 Newton's Method

Recall that the method of SD uses only first derivatives (gradients) in selecting a suitable search direction. This strategy is not always the most effective.

If higher derivatives are used, the resulting iterative algorithm may perform better than the SD method.

Newton's method uses first and second derivatives and indeed does perform better than the SD if the initial point is close to the minimizer.

The idea behind this method is as follows:

Search Techniques

The idea behind this method is as follows:

- Given starting point

The idea behind this method is as follows:

- Given starting point
- construct a quadratic approximation to the objective function that matches the first and second derivative values at the point

The idea behind this method is as follows:

- Given starting point
- construct a quadratic approximation to the objective function that matches the first and second derivative values at the point
- Then minimize the quadratic function instead of the original objective function

The idea behind this method is as follows:

- Given starting point
- construct a quadratic approximation to the objective function that matches the first and second derivative values at the point
- Then minimize the quadratic function instead of the original objective function
- use the minimizer of the approximate function as a starting point in the next iteration

The idea behind this method is as follows:

- Given starting point
- construct a quadratic approximation to the objective function that matches the first and second derivative values at the point
- Then minimize the quadratic function instead of the original objective function
- use the minimizer of the approximate function as a starting point in the next iteration
- repeat the procedure iteratively

Search Techniques

If the objective function is quadratic, then the approximation is exact, and the method yields the true minimizer in one step.

But, if the objective function is not a quadratic then the approximation will provide only an estimate of the position of the true minimizer.

If the objective function is quadratic, then the approximation is exact, and the method yields the true minimizer in one step.

But, if the objective function is not a quadratic then the approximation will provide only an estimate of the position of the true minimizer.

We can obtain a quadratic approximation to the given twice continuous differentiable objective function $f : \Re^n \to \Re$ using the Taylor's series expansion of f about the current point $x^{(k)}$, neglecting terms of order three and higher.

We obtain

$$f(x) \approx f(x^{(k)}) + (x - x^{(k)})^T g^{(k)} + \frac{1}{2}(x - x^{(k)})^T F(x^{(k)})(x - x^{(k)}) \triangleq q(x),$$

where $g^{(k)} = \nabla f(x^{(k)})$. Applying the FONC to q yields

We obtain

$$f(x) \approx f(x^{(k)}) + (x - x^{(k)})^T g^{(k)} + \frac{1}{2}(x - x^{(k)})^T F(x^{(k)})(x - x^{(k)}) \triangleq q(x),$$

where $g^{(k)} = \nabla f(x^{(k)})$. Applying the FONC to q yields

$$0 = \nabla q(x) = g^{(k)} + F(x^{(k)})(x - x^{(k)}).$$

If $F(x^{(k)}) > 0$ then q achieves a minimum at

$$x^{(k+1)} = x^{(k)} - F(x^{(k)})^{-1} g(x^{(k)}). \tag{15}$$

We obtain

$$f(x) \approx f(x^{(k)}) + (x - x^{(k)})^T g^{(k)} + \frac{1}{2}(x - x^{(k)})^T F(x^{(k)})(x - x^{(k)}) \triangleq q(x),$$

where $g^{(k)} = \nabla f(x^{(k)})$. Applying the FONC to q yields

$$0 = \nabla q(x) = g^{(k)} + F(x^{(k)})(x - x^{(k)}).$$

If $F(x^{(k)}) > 0$ then q achieves a minimum at

$$x^{(k+1)} = x^{(k)} - F(x^{(k)})^{-1} g(x^{(k)}). \quad (15)$$

This recursive formula represents Newton's method.

Example

Consider Powell function using Newton's method

$$f(x_1, x_2, x_3, x_4) = (x_1+10x_2)^2+5(x_3-x_4)^2+(x_2-2x_3)^4+10(x_1-x_4)^4.$$

Use as the starting point $x^{(0)} = [3, -1, 0, 1]^T$. Perform three iterations

Solution:

Example

Consider Powell function using Newton's method

$f(x_1, x_2, x_3, x_4) = (x_1+10x_2)^2+5(x_3-x_4)^2+(x_2-2x_3)^4+10(x_1-x_4)^4.$

Use as the starting point $x^{(0)} = [3, -1, 0, 1]^T$. Perform three iterations

Solution:
Note that $f(x^{(0)}) = 215$. We have

$$\nabla f(x) = \begin{pmatrix} 2(x_1+10x_2)+40(x_1-x_4)^3 \\ 20(x_1+10x_2)+4(x_2-2x_3)^3 \\ 10(x_3-x_4)-8(x_2-2x_3)^3 \\ -10(x_3-x_4)-40(x_1-x_4)^3 \end{pmatrix},$$

and $F(x)$ is given by

$$\begin{pmatrix} 2+120(x_1-x_4)^2 & 20 & 0 & -120(x_1-x_4)^2 \\ 20 & 200+12(x_2-2x_3)^2 & -24(x_2-2x_3)^2 & 0 \\ 0 & -24(x_2-2x_3)^2 & 10+48(x_2-x_3)^2 & -10 \\ -120(x_1-x_4)^2 & 0 & -10 & 10+20(x_1-x_4)^2 \end{pmatrix}$$

Iteration 1:

$$g^{(0)} = [306, -144, -2, -310]^T$$

$$\nabla F(x^{(0)}) = \begin{pmatrix} 482 & 20 & 0 & -480 \\ 20 & 212 & -24 & 0 \\ 0 & -24 & 58 & -10 \\ -480 & 0 & -10 & 490 \end{pmatrix},$$

$$\nabla F(x^{(0)})^{-1} = \begin{pmatrix} .1126 & -.0089 & .0154 & .1106 \\ -.0089 & .0057 & .0008 & -.0087 \\ .0154 & .0008 & .0203 & 0.0155 \\ .1106 & -.0087 & .0155 & .1107 \end{pmatrix},$$

$$F(x^{(0)})^{-1}g^{(0)} = [1.4127, -0.8413, -0.2540, 0.7460]^T.$$

Hence,

$$\begin{aligned} x^{(1)} &= x^{(0)} - F(x^{(0)})^{-1}g^{(0)} = [1.5873, -0.1587, 0.2540, 0.2540]^T, \\ f(x^{(1)}) &= 31.8. \end{aligned}$$

Iteration 2:

$$g^{(1)} = [94.81, -1.179, 2.371, -94.81]^T$$

$$\nabla F(x^{(1)}) = \begin{pmatrix} 215.3 & 20 & 0 & -213.3 \\ 20 & 205.3 & -10.67 & 0 \\ 0 & -10.67 & 31.34 & -10 \\ -213.3 & 0 & -10 & 233.3 \end{pmatrix},$$

$$F(x^{(1)})^{-1}g^{(1)} = (0.5291, -0.0529, 0.0846, 0.0846)^T$$

Hence,

$$\begin{aligned} x^{(2)} &= x^{(1)} - F(x^{(1)})^{(-1)}g^{(1)} \\ &= (1.0582, -0.1058, 0.1694, 0.1694)^T, \\ f(x^{(2)}) &= 6.28 \end{aligned}$$

Iteration 3:

$$g^{(2)} = [28.09, -0.3475, 0.7031, -28.08]^T$$

$$\nabla F(x^{(2)}) = \begin{pmatrix} 96.80 & 20 & 0 & -94.80 \\ 20 & 202.4 & -4.744 & 0 \\ 0 & -4.744 & 19.49 & -10 \\ -94.80 & 0 & -10 & 104.80 \end{pmatrix},$$

Hence,

$$\begin{aligned} x^{(3)} &= x^{(2)} - F(x^{(2)})^{(-1)} g^{(2)} \\ &= (0.7037, -0.0704, 0.1121, 0.1111)^T, \\ f(x^{(3)}) &= 1.24. \end{aligned}$$

Observe that the k-th iteration of Newton's method can be written in two steps as

Step 1 Solve $F(x^{(k)})d^{(k)} = -g^{(k)}, \quad \text{for } d^{(k)};$

Step 2 Set $x^{(k+1)} = x^{(k)} + d^{(k)}$

Step 1 requires the solution of an $n \times n$ system of linear equations. Thus, an efficient method for solving system of linear equations is essential when using Newton's method.

3. Programming Techniques

3.1 Linear Programming

The theory of LP is concerned with the problems of constraints minimization (maximization) where the constraint functions and the objective function are linear.

3. Programming Techniques

3.1 Linear Programming

The theory of LP is concerned with the problems of constraints minimization (maximization) where the constraint functions and the objective function are linear.

If the solution optimum of the LP exists, then it is either an extreme points or convex combination of extreme points.

3. Programming Techniques

3.1 Linear Programming

The theory of LP is concerned with the problems of constraints minimization (maximization) where the constraint functions and the objective function are linear.
If the solution optimum of the LP exists, then it is either an extreme points or convex combination of extreme points.
There are several methods to solve the LP. The most powerful is simplex method.

Let us consider the LP as follows:

$$minimize \quad \sum_{i=1}^{n} c_i^T x_i,$$

subject to

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_j \le \sum_{i=1}^{m} b_i,$$

and

$$(x_1, x_2, \cdots, x_n) \ge 0.$$

Actually, the sign of constraints could be $\{\le, =, \ge\}$.

Feasible region of LP is defined as

$$F = \{x| \in \Re^n, Ax \leq b\}$$

(simplex method)