ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES
DEPARTMENT OF MATHEMATICS

Quadratic Programming Problem

By

Tariku Damtew

Advisor: **Berhanu Guta (PhD)**

June, 2020

Acknowledgment

First of all I would like to thanks my God which help me in all my ways from the beginning to the end.

My special tanks go to my adviser Dr.Berhanu Guta for his help, suggestion and comments. I am also glad to thank to my friends Mis, Wuletaw Yemataw and Atalay Mekonnen for their suggestion and material support.

It is a great pleasure intended to acknowledge all my families specially wife Roza Demile who helped me though the study of graduate program.

Abstract

Quadratic programming is one of the type of nonlinear optimization problems in which the objective function is quadratic and all the constraints are linear. There is no single method available for solving a quadratic programming problem. So a number of methods have been developed for solving quadratic programming problems. In this project concerns for convex and non convex quadratic programming problems with Lagrangian, Wolfe's and Beale methods.

Contents

CHAPTER ONE

1 Introduction

A quadratic programming problem (QPP) involve the maximization or minimization of a quadratic objective function subject to linear equality and inequality constraints on the variables.

The general form of quadratic programming problem is given by:

$$\begin{aligned} min \quad f(y) = &\quad c^T y + \frac{1}{2} y^T D y \\ s.t: &\quad By \leq d \\ &\quad y \geq 0 \end{aligned}$$

where D is an $k \times k$ symmetric matrix, y and c are $k \times 1$ vectors, d is an $m \times 1$ vector and B is an $m \times k$ matrix .

Now this project focuses on Lagrangian method and wolfe algorithm which are used to solve quadratic programming problems. Solving quadratic programming problems using only this method is difficult. So they work in existence with Beale algorithm to reach an optimal solution with in a finite iteration.

This chapter introduces the basic concepts of quadratic programming problem, its optimality condition and solution methods as well as theorems which are needed later for a better understanding of the next chapters.

1.1 Preliminary concept

Consider an optimization problem

$$\begin{gathered} (P) \quad min \quad f(x) \\ s.t: x \in S \end{gathered}$$

- The function f(x) is called objective function of (p).
- The components of variable x are called decision variables.
- S is called feasible (constraint) set.
- The condition(inequality /equality) describing S are called constraints.

1.1.1 Optimal Solution and Optimal Value

Optimal solution :- Let $f(x)$ be objective function and $S \subseteq R^n$ be its feasible set.

- $x^* \in S$ is called optimal solution for minimization problem.

$$min \quad f(x)$$
$$s.t : x \in S$$

If $f(x^*) \leq f(x)$ for all $x \in S$
Such x^* is also called global minimize of f on S .

[x^* is called a local minimizer f on S if there is $\epsilon > 0$ such that $f(x^*) \leq f(x)$ for all $x \in N_\epsilon(x^*) \cap S$ where $N_\epsilon(x^*) = \{x \in R^n : \|x - x^*\| < \epsilon,\}$ $\epsilon - nbh \quad of \quad x^*$].

Optimal value :- If $x^* \in S$ is an optimal solution then $f(x^*)$ is called the optimal (objective) value.

1.1.2 Derivative of Function Over R^k

Definition ;- Let a function $q : R^k \longrightarrow R$ be differentiable, we define:-

A. The gradient of q at $x \in R^k$ denoted by $\bigtriangledown q(x)$,is a vector given by:

$$\bigtriangledown q(x)^T = \left(\frac{\partial q}{\partial x_1}, \frac{\partial q}{\partial x_2}, ...\frac{\partial q}{\partial x_k}\right). \quad \text{(if each partial derivative exist)}$$

B. The hessian matrix, $\bigtriangledown^2 q(x)$ of q for $x \in R^k$ is an $k \times k$ matrix given by:

$$\bigtriangledown^2 q(x) = \left(q_{ij}\right)_{i=1,2,3,...,k \quad j=1,2,3,...,k}, \text{ where } q_{ij} = \frac{\partial^2 q}{\partial x_i \partial x_j}, \text{ (symmetric)}$$

Theorem : First Order Taylor's Theorem:

Iff is differentiable and $x_0, x \in R^n$, then

$f(x) = f(x_0) + \bigtriangledown f(\hat{x})^T (x - x_0)$, for some $\hat{x} \in (x_0, x)$.

Theorem : Second Order Taylor's Theorem:

Iff is twice differentiable and $x_0, x \in R^n$, then

$f(x) = f(x_0) + \bigtriangledown f(x_0)^T (x - x_0) + \frac{1}{2}(x - x_0)^T \bigtriangledown^2 f(\hat{x})(x - x_0)$, for some $\hat{x} \in (x_0, x)$.

1.1.3 Convex Set and Convex Function

Definition :- A non empty set S is called convex if for any $x_1, x_2 \in S$

$$\lambda x_1 + (1-\lambda)x_2 \in S, \quad \forall \lambda \in [0,1]$$

Definition:- Let $S \subseteq R^n$ be a convex set. A function $f : S \longrightarrow R$ is said to be convex (over S)

if for any $x_1, x_2 \in S$ and any $\lambda \in [0,1]$,

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2).$$

- f is said to be concave if and only if $-f$ convex
  [so, f is concave if and only if $f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda) f(x_2)$]

- f is strictly convex if for every distinct $x_1, x_2 \in S$ and $\lambda \in (0,1)$,

$$f(\lambda x_1 + (1-\lambda)x_2) < \lambda f(x_1) + (1-\lambda) f(x_2).$$

Theorem 1:-Let $S \subseteq R^n$ is convex and $f : S \longrightarrow \mathbb{R}$ is differentiable. f is convex on S if and only if $f(y) \geq f(x) + \bigtriangledown f(x)^T (y-x)$ for every $x, y \in S$.

Proof

($\Leftarrow$) Suppose $f(y) \geq f(x) + \bigtriangledown f(x)^T (y-x) \quad , \forall x, y \in S$

Let $x_1, x_2 \in S$ and let $\bar{x} = \lambda x_1 + (1-\lambda)x_2$ for any $\lambda \in [0,1]$.

So from the given taking x_1 and x_2 as y, and $\bar{x}$ as x we get:

$f(x_1) \geq f(\bar{x}) + \bigtriangledown f(\bar{x})^T (x_1 - \bar{x})$ and

$f(x_2) \geq f(\bar{x}) + \bigtriangledown f(\bar{x})^T (x_2 - \bar{x})$

$\Rightarrow f(x_1) \geq f(\bar{x}) + (1-\lambda) \bigtriangledown f(\bar{x})^T (x_1 - x_2)$
and $f(x_2) \geq f(\bar{x}) - \lambda \bigtriangledown f(\bar{x})^T (x_1 - x_2)$

$\Rightarrow \lambda f(x_1) \geq \lambda f(\bar{x}) + \lambda(1-\lambda) \bigtriangledown f(\bar{x})^T (x_1 - x_2)$ and
$(1-\lambda) f(x_2) \geq (1-\lambda) f(\bar{x}) - \lambda(1-\lambda) \bigtriangledown f(\bar{x})^T (x_1 - x_2)$

Then adding the two inequalities we get:

$\lambda f(x_1) + (1-\lambda) f(x_2) \geq \lambda f(\bar{x}) + (1-\lambda) f(\bar{x}) = f(\bar{x})$

Therefor a function f is convex .

($\Rightarrow$) Let a function f is convex

$\Rightarrow f(x+t(y-x)) \leq f(x)+t(f(y)-f(x)), t \in (0,1)$

$\Rightarrow t(f(y)-f(x)) \geq f(x+t(y-x))-f(x)$

$\Rightarrow f(y)-f(x) \geq \frac{f(x+t(y-x))-f(x)}{t}, \quad \forall t \in (0,1)$

So taking $t \longrightarrow 0^{+}$, we get:

$f(y)-f(x) \geq \lim_{t\to 0^{+}} \frac{f(x+t(y-x))-f(x)}{t} = \bigtriangledown f(x)^{T}(y-x)$

Therefore $f(y) \geq f(x)+\bigtriangledown f(x)^{T}(y-x)$.

Corollary:- Suppose $S \subseteq R^{n}$ is convex, $x_0 \in S$ and $f: S \longrightarrow \mathbb{R}$ is a differentiable convex function. If $\bigtriangledown f(x_0)=0$, then x_0 is a global minimizer of f on S.

Proof

Let $\bigtriangledown f(x_0)=0$ and let f be a differentiable convex function

$\Rightarrow f(x) \geq f(x_0)+\bigtriangledown f(x_0)^{T}(x-x_0), \quad \forall x, x_0 \in S$
$\Rightarrow f(x) \geq f(x_0), \quad \forall x \in S$

Therefore x_0 is global min of f.

1.1.4 Minimizing Convex Function

- A problem of minimizing a convex function over a convex set is called convex programming.

- That is,if S is a convex set and f is a convex function(on S), then

$$\begin{gathered} min \quad f(x) \\ s.t: x \in S \end{gathered}$$

  is convex programming.

- In particular, if

$f, q_j : R^n \longrightarrow R, j = 1,2,3,...,n$ are all convex function and

$r_i : R^n \longrightarrow R, i = 1,2,3,...,m$ are all affine functions, then

$$\begin{aligned} min \quad & f(x) \\ s.t: \quad & q_j(x) \leq 0, j = 1,2,3,...,n \\ & r_i(x) = 0, i = 1,2,3,...,m \\ & x \geq 0 \end{aligned}$$

is a convex programming problem.

Theorem 2:- Let S be anon empty open convex set in R^n and $f : S \longrightarrow \mathbb{R}$ be a convex function. If x_0 is a local minimizer of f over S, then x_0 is global minimum of f over S.

Proof
Given $x_0 \in S$ is local minimum of f, implies that $f(x_0) \leq f(x), \quad \forall x \in N_\epsilon(x_0)$ and $N_\epsilon(x_0) \subseteq S$ for some $\epsilon > 0$.

WTS:- $f(x_0) \leq f(x), \quad \forall x \in S$

Assume that $f(x^*) < f(x_0)$ for some $x^* \in S, \quad x^* \neq x_0$

Now take the open line segment $(x_0, x^*) \subseteq S$ and take $\bar{x} \in (x_0, x^*)$ such that $\bar{x} \in N_\epsilon(x_0)$

By convexity of f:

$$\begin{aligned} f(\bar{x}) &= f(x_0 + \lambda(x^* - x_0)) \\ &\leq f(x_0) + \lambda(f(x^*) - f(x_0)) \\ &< f(x_0) \end{aligned}$$

So we got $f(\bar{x}) < f(x_0)$ where $\bar{x} \in N_\epsilon(x_0)$ which contradiction to the fact that x_0 is local minimum.

Theorem 3:- Let S be anon empty open convex set in R^n and $f : S \longrightarrow \mathbb{R}$ is twice continuously differentiable on S. f is convex if and only if its hessian matrix $\bigtriangledown^2 f(\hat{x})$ is positive semi definite at each $\hat{x} \in S$.

Proof
$(\Leftarrow)$Suppose $\bigtriangledown^2 f(\hat{x})$ is positive semi-definite at any $\hat{x} \in S$

$\Rightarrow (x - \bar{x})^T \bigtriangledown^2 f(\hat{x})(x - \bar{x}) \geq 0$ for any $x, \bar{x} \in S$

By $2^n d$ order Taylor's theorem we also have:

$f(x) - \left(f(\bar{x}) + \bigtriangledown f(\bar{x})^T (x - \bar{x})\right) = \frac{1}{2}(x - \bar{x})^T \bigtriangledown^2 f(\hat{x})(x - \bar{x}) \geq 0$

$\Rightarrow f(x) \geq f(\bar{x}) + \triangledown f(\bar{x})^T (x - \bar{x})$ for any $x, \bar{x} \in S$

Hence f is convex by theorem 1.

($\Rightarrow$)If f is convex on S, then $f(x + th) \geq f(x) + t \triangledown f(x) h$ for any $h \in R^n$, $t \in (0,1)$

and from $2^n d$ order Taylor's theorem:

$f(x + th) = f(x) + t \triangledown f(x)^T h + \frac{t^2}{2} h^T \triangledown^2 f(x + \theta th) h, \quad \theta \in (0,1)$

i.e $\quad \frac{t^2}{2} h^2 \triangledown^2 f(x + \theta th) h = f(x + th) - (f(x) + t \triangledown f(x)^T h), \quad \forall (t) > 0, \quad \theta \in (0,1)$

$\Rightarrow \frac{t^2}{2} h^T \triangledown^2 f(x + \theta th) h \geq 0, \quad \forall t > 0, \quad h \in R^n$

$\Rightarrow h^T \triangledown^2 f(x + \theta th) h \geq 0$

Take $t \to 0, \quad h^T \triangledown^2 f(x) h \geq 0 \quad$ for any $h \in R^n$

Therefore $\triangledown^2 f(x)$ is positive semi-definite.

1.1.5 Optimal Condition

- Optimality condition is a condition that has to be satisfied for a feasible point x^* to be an optimal solution.

Theorem. (First Order Necessary Condition)
Let $f : R^n \to R$, differentiable.
If f has local min value at $x^* \in R^n$, then $\triangledown f(x^*) = 0$.

Theorem. (Second Order Sufficient Condition)
Let $f : R^n \to R$, twice differentiable at $x^* \in R^n$.
If

i. $\triangledown f(x^*) = 0$

ii. $\triangledown^2 f(x^*)$ is positive definite

then x^* is a local minimum point of f.

1.1.6 Definiteness of Matrices

Definition :- Let A be a square matrix of order n. Then A is said to be :

- Positive definite (PD) if $X^T BX > 0$ for all non zero$X \in R^n$.

- Positive semi- definite (PSD) if $X^T BX \geq 0$ for all non zero$X \in R^n$.
  (Indefinite $X^T BX > 0$ is Positive for some non zero $X \in R^n$ and negative for some non zero $X \in R^n$).
  B is said to be negative(semi)definite if $-B$ is PD(PSD).

- The following characterizes a PD or PSD matrix.

Let B_k be the $k \times k$ minor sub-matrix of B: $k = 1, 2, 3, ..., n$.

1. B is PD iff $\det(B_k) > 0$ for all $k = 1, 2, 3, ..., n$.

2. B is PSD iff $\det(B_k) \geq 0$ for all $k = 1, 2, 3, ..., n$.

1.1.7 Linear Programming Problem

A linear program (LP) is an optimization problem in which the objective function is linear in the unknowns and the constraints consist of linear equalities and linear inequalities. The exact form of these constraints may differ from one problem to another, but as shown below, any linear program can be transformed in to the following standard form:

$$\begin{aligned}
Z(max/min) = \quad & c_1x_1 + c_2x_2 + ... + c_nx_n \\
s.t: \quad & a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\
& a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\
& . \\
& . \\
& . \\
& a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m \\
& and \quad x_1, x_2, ..., x_n \quad \geq 0
\end{aligned}$$

where the b_i's, c_i's and a_{ij}'s are fixed real constants, and the x_i's are real numbers to be determined. We always assume that each equation ha been multiplied by minus unity, if necessary, so that each $b_i \geq 0$.

In more compact vector notation, this standard problem becomes:

$$\begin{aligned}
Z(max/min) = \quad & c^T x \\
s.t: \quad & Bx = b \\
and \quad & x \geq 0.
\end{aligned}$$

Here x is an n-dimensional column vector, c^T is an n-dimensional row vector, B is an $m \times n$ matrix, and b is an m-dimensional column vector. The vector inequality $x \geq 0$ means that each component of x is non negative.

1.1.8 Nonlinear Programming Problem

A general optimization problem is to select n decision variables $x_1, x_2, ..., x_n$ from a given feasible region in such a way as to optimize (minimize or maximize) a given objective function

$$f(x_1, x_2, ..., x_n)$$

of the decision variables. The problem is called a nonlinear programming problem (NLP) if the objective function is nonlinear and/or the feasible region is determined by nonlinear constraints. Thus, in maximization or minimization form, the general nonlinear program is stated as:-

$$\begin{aligned}
Z(max/min) = \quad & f(x_1, x_2, ..., x_n) \\
s.t: \quad & g_1(x_1, x_2, ..., x_n) \quad (\leq, =, \geq) b_1 \\
& g_2(x_1, x_2, ..., x_n) \quad (\leq, =, \geq) b_2 \\
& . \\
& . \\
& . \\
& g_m(x_1, x_2, ..., x_n) \quad (\leq, =, \geq) b_m \\
& and \quad x_1, x_2, ..., x_n \quad \geq 0
\end{aligned}$$

where each of the constraint function g_1 through g_m is given.

1.2 Lagrangian Method for Constrained Problems and Karush-Kuhn-Tucker Optimal Conditions

Let $f : R^n \longrightarrow \mathbb{R}$ be differentiable, and $S \subseteq R^n$, we consider a constrained non linear minimization problem:

$$(p) \quad min \quad f(x)$$
$$s.t: \quad x \in S$$

- To convert p in to unconstrained problem /or relatively easier problem / we try to fined a function f_0 on R^n with the following property.

  $f_0 : R^n \longrightarrow \mathbb{R}$ be differentiable, and $f_0(x) = 0, \quad \forall x \in S$

- Then we consider an auxiliary function L such that

  $L : R^n \longrightarrow \mathbb{R}$ given by

  $L(x) = f(x) + f_0(x)$

- Such L is called Lagrange function /Lagrangian/ for (p).

- Then we consider the auxiliary /Lagrangian/ problem:

$$(p_L) \quad min \quad L(x)$$
$$s.t: \quad x \in R^n$$

- Note that L is differentiable .

- The relation between $(p)$and (p_L) given the following theorem.

Theorem (Lagrangian Lemma)

A. If x^* is a global minimizer of L and $x^* \in S$ then x^* is optimal solution of (p).

B. If x^* is a local minimizer of L and $x^* \in S$ then x^* is local optimal solution of (p).

Proof

Since L is given on R^n, If x^* is a local/global minimizer of L, then $\quad \triangledown L(x^*) = 0$

$i.e \quad \triangledown L(x^*) = 0$ is a necessary condition for x^* to be a minimizer of L on R^n.

- If so then $\triangledown L(x^*) = 0$ and $x^* \in S$ are necessary condition for x^* to be aminimizer of (p).

1.3 Lagrangian Method for Equality Constraint

- Let $\quad f : R^n \longrightarrow \mathbb{R}, \quad$ be differentiable, and

  $h_j : R^n \longrightarrow \mathbb{R}^k, \quad$ be differentiable for each $\quad j \in J = \{1, 2, 3, \ldots k\}$

- We consider, the problem

$$(p =) \quad min \quad f(x)$$
$$s.t: \quad x \in S = \{x \in R^n / h_j(x) = 0, j \in J\}$$

- Now if we take

  $$f_0(x) = \sum_{j=1}^{k} \mu_j h_j(x)$$
  where $\mu_j \in R$ then $f_0(x) = 0$ for all$x \in S$

- So we define the Lagrangian function for $(p =)$as follows:

- For any $\mu = (\mu_1, \mu_2, ..., \mu_k)^T \in R^k$ called Lagrangian multipliers, let

  $L : R^n \times R^k \longrightarrow \mathbb{R}$, given by

$$L(x,\mu) = f(x) + \sum_{j=1}^{k} \mu_j h_j(x)$$

- Thus, the corresponding auxiliary (Lagrange) problem is:

$$\begin{array}{lll} (L_\mu) & min & L(x,\mu) \\ & s.t: & x \in R^n, \quad \mu \in R^k \end{array}$$

- Then from the Lagrange Lemma, we have following result:

Theorem:-If (x^*, μ^*) is an optimal solution of (L_μ), for some $x^* \in R^n$ and $\mu^* \in R^k$, then

1. $x^* \in S$ and
2. x^* is an optimal solution of $(p =)$

proof

1. $\bigtriangledown L_\mu(x^*, \mu^*) = h(x^*) = 0, \quad \Rightarrow x^* \in S$

2. Follow from the lagrange lemma

- Here $(x^*, \mu^*) \in R^n \times R^k$ is an optimal solution of (L_μ), implise

  $\bigtriangledown L_x(x^*, \mu^*) = 0$

  and $\bigtriangledown L_\mu(x^*, \mu^*) = h(x^*) = 0 \quad i.e \quad h_j(x^*) = 0$ for all $j = 1, 2, 3, ..., k$
  Thus we have the following necessary condition for x^* to be an optimal solution of $(p =)$.

- Karush-Kuhn-Tucker (KKT) condition (for equality constrained problems):

  If x^* is an optimal solution of $(p =)$ then for some $\mu^* = (\mu_1^*, \mu_2^*, ..., \mu_k^*)^T \in R^k$ the following hold:

  1. $\bigtriangledown f(x^*) + \sum_{j=1}^{k} \mu_j^* \bigtriangledown h_j(x^*) = 0, \quad \leftarrow$ **Dual feasibility**
  2. $h_j(x^*) = 0, \quad j = 1, 2, 3, ..., k, \quad \leftarrow$ **primal feasibility**

- If x^* together with some $\mu^* \in R^k$ satisfy the KKT conditions, then x^* is called a KKT point for $(p =)$.

- The KKT conditions are only necessary conditions. However, the KKT conditions are sufficient condition under some convexity assumption.

Theorem:- Let x^* be a KKT point for $(p =)$ with a corresponding Lagrange multiplier μ^* and $L(.,\mu^*)$ be a convex function then x^* is an optimal solution of $(p =)$.

proof

$L(.,\mu^*)$ is convex and $\quad \bigtriangledown L_x(x^*,\mu^*) = 0$

$\Rightarrow x^*$ is optimal solution of (L_{μ^*}). furthermore, $x^* \in S$ Hence, by the Lagrange Lemma, x^* is optimal solution of $(P =)$.

Corollary : Consider the problem

$$\begin{aligned} (p =) \quad min \quad & f(x) \\ s.t: \quad & h_j(x) = 0, \quad j \in J = \{1,2,3,...,k\} \end{aligned}$$

where f is a convex function.

If x^* is a KKT point of $(p =)$ with some Lagrangian multipliers $\mu^* = (\mu_1^*, \mu_2^*, ..., \mu_k^*)^T$ and

$h_{\mu^*}(x) = \mu_1^* h_1(x) + \mu_2^* h_2(x) + ... + \mu_k^* h_k(x)$ is a convex function then x^* is an optimal solution.

Note: $\quad h_\mu(x) \quad$ is convex if

1. Each h_j affine (' linear ') function or

2. h_j is convex and $\mu_j^* \geq 0$ for each $j \in J$.

1.4 Lagrangian Method for Inequality Constraint

- Let $\quad f : R^n \longrightarrow \mathbb{R}, \quad$ be differentiable, and

  $g_i : R^n \longrightarrow \mathbb{R}^m$, be differentiable for each $\quad i \in I = \{1,2,3,...,m\}$.

- We consider, the problem

$$\begin{aligned} (p <) \quad min \quad & f(x) \\ s.t: \quad & x \in S = \{x \in R^n / g_i(x) \leq 0, i \in I\} \end{aligned}$$

- Now we define Lagrangian function for $(p <)$ as follows :
  for any $\lambda = (\lambda_1, \lambda_2, \lambda_3, ..., \lambda_m)^T \in R_+^m, \quad i.e \quad \lambda_i \geq 0 \quad \forall i$, define

$L : R^n \times R_+^m \longrightarrow R \quad$ given by

$$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

(L is called Lagrange function and each λ_i are called Lagrange multiplier)

- Thus corresponding auxiliary problem is:

$$\begin{array}{lll} (L_\lambda) & min & L(x,\lambda) \\ & s.t: & x \in R^n \quad \forall \lambda \in R^m_+ \end{array}$$

- The following is the corresponding Lagrange Lemma relating $(p <)$ and (L_λ) :

Theorem:- Let $<\lambda^*, g(x^*)>= 0$ for some $\lambda^* \in R^m_+$. If (x^*, λ^*) is an optimal solution of (L_λ), for some $\lambda^* \in R^m_+$ and $x^* \in S$, then x^* is an optimal solution of $(P <)$.

- Thus, under the condition $\lambda^* \in R^m_+$ and $<\lambda^*, g(x^*)>= 0$, x^* an optimal solution of (L_λ) over R^n

  implies $\triangledown L_x(x^*, \lambda^*) = 0$, and
  $x^* \in S$ implies $g_i(x^*) \leq 0$, for all $i = 1,2,3,...,m$

  Hence we have the following necessarily condition for x^* to be optimal solution of $(p <)$.

- Karush-Kuhn-Tucker(KKT)condition (for inequality constraint problem):

  If x^* is an optimal solution of $(p <)$, then for some $\lambda^* = (\lambda_1^*, \lambda_2^*, \lambda_3^*, ..., \lambda_m^*)^T \in R^m_+$
  the following hold:

  1. $\triangledown f(x^*) + \sum_{i=1}^{m} \lambda_i^* \triangledown g_i(x^*) = 0$ $\leftarrow$ **Dual feasibility**
     $\lambda_i^* \geq 0, \quad i = 1,2,3,...,m$
  2. $\lambda_i^* g_i(x^*) = 0, \quad i = 1,2,3,...,m$ $\leftarrow$ **complementary slackness**
  3. $g_i(x^*) \leq 0, \quad i = 1,2,3,...,m$ $\leftarrow$ **primal feasibility**

- If x^* together with some $\lambda^* \in R^m_+$ satisfy the KKT condition then $x^*$is called a KKT point for $(p <)$.

Theorem:- Let x^* be a KKT point of $(p <)$ with a corresponding Lagrange multiplier $\lambda^* \in R^m_+$. If f is a convex function over S and each constraint function g_i is convex , then x^* is an optimal solution of $(p <)$.

1.5 Lagrange Method for Mixed constraint

- Let

$$\begin{aligned} g(x) = &\quad (g_1(x), g_2(x), g_3(x), ..., g_m(x))^T \\ h(x) = &\quad (h_1(x), h_2(x), h_3(x), ..., h_k(x))^T \end{aligned}$$

- We consider, the problem

$$\begin{aligned} &(p) \quad min \quad f(x) \\ &s.t: \quad x \in S := \quad \{x \in R^n / g(x) \leq 0, \quad h(x) = o\} \end{aligned}$$

- Now, we define Lagrange function for (p) as follows:

  for any $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, ...\lambda_m)^T \in R^m_+, \quad i.e \quad \lambda_i \geq 0, \quad \forall i$ and $\mu = (\mu_1, \mu_2, \mu_3, ..., \mu_k) \in R^k$ define,

  $L : R^n \times R^m_+ \times R^k \longrightarrow R$ given by

  $$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{k} \mu_j h_j(x)$$

- Thus, corresponding auxiliary problem is:

$$\begin{aligned} &(L_{\lambda,\mu}) \quad min \quad L(x, \lambda, \mu) \\ &Subject \quad to: \quad x \in R^n \\ &\qquad \forall \lambda \in R^m_+ \quad and \quad \mu \in R^k \end{aligned}$$

  Hence, we have the following necessary condition for x^* to be optimal solution of (p).

- Karush-Kuhn-Tucker (KKT)condition (for mixed constraint problem):

  If x^* is an optimal solution of (p), then for some $\lambda^*_i, \mu^*_j \in R$ the following hold:

  1. $\bigtriangledown f(x^*) + \sum_{i=1}^{m} \lambda^*_i \bigtriangledown g_i(x^*) + \sum_{j=1}^{k} \mu^*_j \bigtriangledown h_j(x^*) = 0$ $\leftarrow$ **Dual feasibility**
     $\lambda^*_i \geq 0, \quad i = 1, 2, 3, ..., m$
  2. $\lambda^*_i g_i(x^*) = 0, \quad i = 1, 2, 3, ..., m$ $\leftarrow$ **complementary slackness**
  3. $g_i(x^*) \leq 0, \quad i = 1, 2, 3, ..., m$ $\leftarrow$ **primal feasibility**
     $h_j(x^*) = 0, \quad j = 1, 2, 3, 4, ..., k$

- If x^* together with some $\lambda^* \in R^m_+, \mu^* \in R^k$ satisfy the KKT condition, then x^* is called a KKT point for (p).

Example:- solve the following problem using Lagrange method.

$$\begin{aligned} min \quad f(x) = &\quad x_1^2 - x_2 - 2x_1 \\ s.t: &\quad 2x_1 + 3x_2 = 6 \\ &\quad 2x_1 + x_2 \leq 4 \\ &\quad x_1, x_2 \geq 0 \end{aligned}$$

solution

$f(x) = x_1^2 - x_2 - 2x_1, \quad g(x) = 2x_1 + x_2 - 4 \quad and \quad h(x) = 2x_1 + 3x_2 - 6$

The corresponding Lagrange function is:

$$\begin{aligned} L(x, \lambda, \mu) =& f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{k} \mu_j h_j(x) \\ =& x_1^2 - x_2 - 2x_1 + \lambda(2x_1 + x_2 - 4) + \mu(2x_1 + 3x_2 - 6) \end{aligned}$$

This implies that:

$$\begin{aligned} min &\quad L(x, \lambda, \mu) \\ Subject \quad to: &\quad x \in R^2 \\ &\quad \forall \lambda \in R^+ \quad and \quad \mu \in R \end{aligned}$$

is the corresponding unconstrained optimization problem.

The Karush-Kuhn-Tucker (KKT) condition is:

$\bigtriangledown f(x) + \lambda \bigtriangledown g(x) + \mu \bigtriangledown h(x) = 0$

$$\begin{pmatrix} 2x_1 - 2 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow 2x_1 + 2\lambda + 2\mu - 2 = 0 \tag{1}$$

$$\lambda + 3\mu - 1 = 0 \tag{2}$$

$$\lambda \geq 0 \tag{3}$$

$\lambda g(x) = 0$

$$\Rightarrow \lambda(2x_1 + x_2 - 4) = 0 \tag{4}$$

$h(x) = 0$

$$\Rightarrow 2x_1 + 3x_2 - 6 = 0 \tag{5}$$

$g(x) \leq 0$

$$\Rightarrow 2x_1 + x_2 - 4 \leq 0 \tag{6}$$

To solve this system of equation and inequalities we consider the following cases.

case one:- If $\lambda > 0$, then from equation (4) and equation (5), we get $(x_1, x_2) = (\frac{3}{2}, 1)$ this substitute in to equation (1) and equation (2) to get $\lambda = \frac{-5}{4}$, but this is a contradiction to equation (3).

Hence no KKT point corresponding to this case.

case two:- If $\lambda = 0$, from equation (2) we have to get $\mu = \frac{1}{3}$ substitute this in to equation (1) to get $x_1 = \frac{2}{3}$ and from equation (5) we get $x_2 = \frac{14}{9}$.

Therefor $(x_1, x_2) = (\frac{2}{3}, \frac{14}{9})$ is KKT point.

Hence by convexity of the given problem $(x_1, x_2) = (\frac{2}{3}, \frac{14}{9})$ is an optimal solution.

CHAPTER TWO

2 Quadratic programming

In mathematics, Optimization means to maximize or minimize a function $f(x)$, where the values of x have to satisfy a set of prescribed mathematical relation. There are different methods available for solving all optimization problems efficiently. Hence, number of methods developed for solving quadratic programming problem which is a type of optimization problem.

A quadratic programming problem (QPP) is the special class of non linear optimization problems in which the objective function is quadratic and all the constraints are linear .

The general mathematical formulation of a (QPP) is as follow:

$$(QPP) \quad min \quad f(y) = c^T y + \frac{1}{2} y^T D y$$
$$s.t: By \leq d$$
$$y \geq 0$$

where D is an $k \times k$ symmetric matrix, y and c are $k \times 1$ vectors, d is an $m \times 1$ vector and B is an $m \times k$ matrix .

Consider the following optimization problem:

$$\begin{aligned} min \quad f(y) = \quad & 3y_1^2 + 4y_2^2 + 2y_1 y_2 - 2y_1 - 3y_2 \\ s.t: \quad & 3y_1 + 2y_2 \leq 6 \\ & y_1 + y_2 \leq 2 \\ & y_1, y_2 \geq 0 \end{aligned}$$

This can be equivalent to:-

$$\begin{aligned} min \quad f(y) = \quad & \begin{pmatrix} -2 & -3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} y_1 & y_2 \end{pmatrix} \begin{pmatrix} 6 & 2 \\ 2 & 8 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ s.t: \quad & \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \leq \begin{pmatrix} 6 \\ 2 \end{pmatrix} \\ & \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

$\Rightarrow$ This optimization problem is called quadratic programming problem.

The convexity of the quadratic objective function plays a defining role in the computational convexity of quadratic programming problem.

Quadratic programming problems are typically classified as being convex or non convex depending on whether D is positive definite or indefinite .

- **Convex Quadratic Programming Problem**

  The QPP,

$$\begin{aligned} min \quad f(y) = c^T y + \frac{1}{2} y^T D y \\ s.t: By \le d \\ y \ge 0 \end{aligned}$$

  is convex quadratic programming problem when D is PSD.

  The QPP has a unique local minimum (also called global minimum) if the objective function $f(y)$ is strictly convex for all feasible points.

Theorem:-If D be a symmetric positive semi-definite matrix of order k, then for any $x, y \in R^k$

i. $2y^T Dx \le (y^T Dy + x^T Dx)$ and

ii. $f(y) = y^T Dy + c^T y, \quad y \in R^k$ (in QPP) is a convex function.

Proof

i. Let D be a symmetric positive semi-definite matrix of order k.

$\Rightarrow (y-x)^T D(y-x) \ge 0, \forall x, y \in R^k$

WTS:- $2y^T Dx \le (y^T Dy + x^T Dx)$

Suppose $(y-x)^T D(y-x) \ge 0, \quad \forall x, y \in R^k$

$\Rightarrow y^T Dy - y^T Dx - x^T Dy + x^T Dx \ge 0$

$\Rightarrow y^T Dx + x^T Dy \le y^T Dy + x^T Dx$

$\Rightarrow 2y^T Dx \le (y^T Dy + x^T Dx), \quad \forall x, y \in R^k$

ii. Let D be a symmetric positive semi-definite matrix of order k.

WTS:- f is convex

For any $y_1, y_2 \in R^k, \quad \lambda \in [0,1]$

$$\begin{aligned}
f(\lambda y_1 + (1-\lambda)y_2) - \lambda f(y_1) - (1-\lambda) f(y_2) &= (\lambda y_1 + (1-\lambda)y_2)^T D(\lambda y_1 + (1-\lambda)y_2) - \lambda[y_1^T D y_1 + c^T y_1] \\
&- (1-\lambda)[y_2^T D y_2 + c^T y_2] + c^T[\lambda y_1 + (1-\lambda)y_2] \\
&= \lambda^2 y_1^T D y_1 + \lambda(1-\lambda) y_1^T D y_2 + \lambda(1-\lambda) y_2^T D y_1 + (1-\lambda)^2 y_2^T D y_2 \\
&- \lambda y_1^T D y_1 - (1-\lambda) y_2^T D y_2 \\
&= \lambda(\lambda-1) y_1^T D y_1 + (1-\lambda) y_2^T D y_2 (1-\lambda-1) + 2\lambda(1-\lambda) y_1^T D y_2 \\
&\le \lambda(\lambda-1) y_1^T D y_1 - \lambda(1-\lambda) y_2^T D y_2 + \lambda(1-\lambda) y_1^T D y_1 \\
&+ \lambda(1-\lambda) y_2^T D y_2 \\
&= 0
\end{aligned}$$

$\Rightarrow f(\lambda y_1 + (1-\lambda)y_2) \le \lambda f(y_1) - (1-\lambda) f(y_2), \quad \forall y_1, y_2 \in R^k, \quad \lambda \in [0,1]$

Hence a function f is convex.

- **Non Convex Quadratic Programming Problem:-** In the quadratic programming when D is indefinite or negative definite the quadratic programming problem is non convex and a local optimal points may not be global optimal solution .

When the matrix D is negative semi-definite that is all its eigenvalues are non positive, then it is called a concave quadratic programming problem.

Example:- Let $f(y) = y^T D y + c^T y$, where D is $k \times k$ symmetric matrix and $c, y \in R^k$.

1. f is concave if D is negative semi-definite.
2. f is strictly concave if D is negative definite.

Note:- D is negative semi-definite $\Rightarrow 2y^T D x \ge (y^T D y + x^T D x), \quad \forall x, y \in R^k$ as

$(y-x)^T D(y-x) \le 0$ and inequalities are strict when D is negative definite.

2.1 Lagrangian Method for Equality Constraint QPP

- consider the following QPP:-

$$\begin{aligned}
(p=) \quad min \quad f(y) = \quad & c^T y + \frac{1}{2} y^T D y \\
s.t: \quad & y \in S = \{y \in R^n / h_j(y) = Py - q = 0, \quad j = 1,2,3,...,s\}
\end{aligned}$$

- Now if we take

  $$f_0(y) = \sum_{j=1}^{s} \mu_j h_j(y)$$

  where $\mu_j \in R$ then $f_0(y) = 0$ for all$y \in S$

- So, we define the Lagrangian function for $(p =)$as follows:

- For any $\mu = (\mu_1, \mu_2, ..., \mu_s)^T \in R^s$, called Lagrangian multipliers, let

  $L : R^k \times R^s \longrightarrow \mathbb{R}$ given by

$$L(y, \mu) = f(y) + \sum_{j=1}^{s} \mu_j h_j(y)$$

- Thus, the corresponding auxiliary (Lagrange) problem is:

$$\begin{aligned} (L_\mu) \quad min \quad & L(y, \mu) \\ s.t: \quad & y \in R^k, \quad \mu \in R^s \end{aligned}$$

Then from the Lagrange Lemma, we have following result:

Theorem:-If (y^*, μ^*) is an optimal solution of (L_μ), for some$y* \in R^k$ and $\mu^* \in R^s$, then

1. $y^* \in S$ and
2. y^* is an optimal solution of $(p =)$.

Proof

1. $\triangledown L_\mu(y^*, \mu^*) = h(y^*) = 0, \quad \Rightarrow y^* \in S.$

2. follow from the lagrange lemma.

- Here $(y^*, \mu^*) \in R^k \times R^s$ is an optimal solution of (L_μ), implies

  $$\triangledown L_y(y^*, \mu^*) = 0,$$

  and $\triangledown L_\mu(y^*, \mu^*) = h(y^*) = 0 \quad i.e \quad h_j(y^*) = 0$ for all $j = \{1, 2, 3, ..., s\}$

  Thus, we have the following necessary condition for y^* to be an optimal solution of $(p =)$.

- Karush-Kuhn-Tucker (KKT) condition (for equality constrained QPP):

  If y^* is an optimal solution of $(p=)$ then for some $\mu^* = (\mu_1^*, \mu_2^*, ..., \mu_s^*)^T \in R^s$ the following hold:

  $1. C + Dy^* + P^T\mu^* = 0$ $\leftarrow$ **Dual feasibility**

  $2. Py^* = q$ $\leftarrow$ **primal feasibility**

- If y^* together with some $\mu^* \in R^s$ satisfy the KKT conditions, then y^* is called a KKT point for $(p=)$.

- Let y^* be a KKT point of equality constrained QPP with a corresponding Lagrange multiplier $\mu^* \in R^s$. If the given QPP is a convex, then y^* is an optimal solution of QPP.

2.2 Karush-Kuhn-Tucker (KKT) Condition for Inequality Constraint QPP

Consider the following QPP:-

$$
\begin{aligned}
min \quad f(y) = \quad & c^T y + \frac{1}{2} y^T D y \\
s.t: \quad & By \leq d && (7) \\
& y \geq 0 && (8)
\end{aligned}
$$

where y and c are k component column vectors, B is an $m \times k$ matrix, d is an $m \times 1$ vector and D is a $k \times k$ symmetric matrix .

Let the KKT multiplier associated with the constraint (7) and (8) $u \in R_+^m$ and $v \in R_+^k$ respectively.

Then the KKT condition for the problem (QPP)are as follows:-

$$
\begin{aligned}
c^T + y^T D + u^T B - v^T (I) &= 0 \\
u^T (By - d) - v^T (y) &= 0 \\
By - d &\leq 0 \\
y \geq 0, \quad u \geq 0, \quad v &\geq 0
\end{aligned}
$$

OR

$$
\begin{aligned}
c + Dy + B^T u - v(I) &= 0 \\
By + s &= d \\
u_i s_i = 0, \quad i &= 1,2,3,4,...,m \\
v_j y_j = 0, \quad j &= 1,2,3,4,...,k
\end{aligned}
$$

- If y together with some $u \in R_{+}^{m}, v \in R_{+}^{k}$ satisfy the KKT condition then y is called a KKT point for QPP.

- Let y be a KKT point of inequality constraint QPP with a corresponding Lagrange multiplier $u \in R_{+}^{m}, v \in R_{+}^{k}$. If the given QPP is a convex, then y is an optimal solution of QPP.

2.3 Karush-Kuhn-Tucker (KKT) Condition for Mixed Constraint QPP

Consider the following QPP:-

$$\begin{aligned} min \quad f(y) = \quad & c^T y + \frac{1}{2} y^T D y \\ s.t: \quad & By \leq d \qquad (9) \\ & Py = q \qquad (10) \\ & y \geq 0 \qquad (11) \end{aligned}$$

where y and c are k component column vectors, B is an $m \times k$ matrix, P is an $l \times k$ matrix, d is an $m \times 1$ vector, q is an $l \times 1$ vector and D is a $k \times k$ symmetric matrix .

Let the KKT multiplier associated with the constraint (9), (10) and (11)be $u \in R_{+}^{m}$, $\mu \in R^{l}$ and $v \in R_{+}^{k}$ respectively.

Then the KKT condition for the problem (QPP)are as follows:-

$$\begin{aligned} c^T + y^T D + u^T B - v^T (I) + \mu^T B &= 0 \\ u^T (Bx - d) - v^T (y) &= 0 \\ By - d &\leq 0 \\ Py - q &= 0 \\ y \geq 0, \quad u \geq 0, \quad v \geq 0 \end{aligned}$$

OR

$$\begin{aligned} c + Dy + B^T u - v(I) + B^T \mu &= 0 \\ Py &= q \\ By + s &= d \\ u_i s_i = 0, \quad i = 1,2,3,4,...,m \\ v_j y_j = 0, \quad j = 1,2,3,4,...,k \end{aligned}$$

- If y together with some $u \in R_{+}^{m}, v \in R_{+}^{k}$ and $\mu \in R^{l}$ satisfy the KKT condition then y is called a KKT point for QPP.

- Let y be a KKT point of mixed constraint QPP with a corresponding Lagrange multiplier $u \in R_{+}^{m}, v \in R_{+}^{k}$, and $\mu \in R^{l}$. If the given QPP is a convex, then y is an optimal solution of QPP.

CHAPTER THREE

3 Methods for Solving Quadratic Programming Problem

We can solve quadratic programming problem by one of the following methods.

1. Algorithm of wolfe.
2. Beale method.

3.1 Algorithm of Wolfe

Consider the following convex QPP

$$\begin{aligned} min \quad f(y) = \quad & c^T y + \frac{1}{2} y^T D y \\ s.t: \quad & By \leq d \\ & y \geq 0 \end{aligned}$$

We can formulate an algorithm to solve a convex quadratic programming problem using algorithm of wolf .

We consider the following steps:-

step 1. Formulate KKT condition for the given quadratic programming problem.

That is the Karush-Kuhn-Tucker(KKT) condition:-

$$\begin{aligned} c + Dy + B^T u - v(I) &= 0 \\ u^T(By - d) - v^T(y) &= 0 \\ By - d &\leq 0 \\ y \geq 0, \quad u \geq 0, \quad v &\geq 0 \end{aligned}$$

step 2. Introduce slack and surplus variable in order to have equation. Check, wether there is a feasible basis. If there is a feasible basis, then the corresponding basic solution is optimal. If there is no feasible basis, then go step 3.

step 3. Introduce artificial variables in order to get a feasible basis, Go to step 4.

step 4. Apply simplex algorithm and solve it.

Example:- Solve the following convex QPP by using wolfe algorithm.

$$\begin{aligned} min \quad f(y) = &\; y_1^2 + 2y_2^2 - 2y_1y_2 - 6y_1 - 8y_2 \\ s.t: &\; 2y_1 - y_2 \le 13 \\ &\; y_1 \ge 0 \\ &\; y_2 \ge 0 \end{aligned}$$

solution

Obviously we have here

$\mathrm{D} = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}$, $\mathrm{c} = \begin{pmatrix} -6 \\ -8 \end{pmatrix}$, $\mathrm{Y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, $\mathrm{B} = \begin{pmatrix} 2 & -1 \end{pmatrix}$, and $\mathrm{d} = \begin{pmatrix} 13 \end{pmatrix}$

step 1. As KKT condition we get:-

$c + Dx + B^T u - v = 0$

$$\begin{pmatrix} -6 \\ -8 \end{pmatrix} + \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix} u - \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow 2y_1 - 2y_2 + 2u - v_1 = 6$$
$$\Rightarrow -2y_1 + 4y_2 - u - v_2 = 8$$

$u^T(By - d) - v^T(y) = 0$

$$u(2y_1 - y_2 - 13) - v_1y_1 - v_2y_2 = 0$$

$By - d \le 0$

$$\Rightarrow 2y_1 + 3y_2 - 6 = 0$$

$g(y) \le 0$

$$2y_1 - y_2 \le 13$$

$y \ge o, u \ge 0, v_1 \ge 0, v_2 \ge 0$

step 2. By introducing of slack variable we get:-

$$\begin{aligned} 2y_1 - 2y_2 + 2u - v_1 &= 6 \\ -2y_1 + 4y_2 - u - v_2 &= 8 \\ 2y_1 - y_2 + s &= 13 \\ us = v_1y_1 = v_2y_2 &= 0 \\ all \quad variable &\ge o \end{aligned}$$

step 3. In order to get a feasible basis we have introduce artificial variable a_1 and a_2 for the first and the second row:

$$\begin{aligned} max \quad & -a_1 - a_2 \\ s.t: \quad & 2y_1 - 2y_2 + 2u - v_1 + a_1 = 6 \\ & -2y_1 + 4y_2 - u - v_2 + a_2 = 8 \\ & 2y_1 - y_2 + s = 13 \\ & all \quad variable \geq o \\ & us = v_1 y_1 = v_2 y_2 = 0 \end{aligned}$$

step 4 . Now applying simplex algorithm technique, then the result have as follow:-

	C_j	0	0	0	0	0	-1	-1	0		
C_B	B.V	y_1	y_2	u	v_1	v_2	a_1	a_2	s_1	sol	ratio
	z	0	-2 $\downarrow$	-1	1	1	0	0	0		
-1	a_1	2	-2	2	-1	0	1	0	0	6	-
-1	$\leftarrow a_2$	-2	4	-1	0	-1	0	1	0	8	2
0	s_1	2	-1	0	0	0	0	0	1	13	-
	z	-1 $\downarrow$	0	$\frac{-3}{2}$	1	$\frac{1}{2}$	0	$\frac{1}{2}$	0		
-1	$\leftarrow a_1$	1	0	$\frac{3}{2}$	-1	$\frac{-1}{2}$	1	$\frac{1}{2}$	0	10	10
0	y_2	$\frac{-1}{2}$	1	$\frac{-1}{4}$	0	$\frac{-1}{4}$	0	$\frac{1}{4}$	0	2	-
0	s_1	$\frac{3}{2}$	0	$\frac{-1}{4}$	0	$\frac{-1}{4}$	0	$\frac{1}{4}$	1	15	10
	z	0	0	0	0	0	1	1	0		
0	y_1	1	0	$\frac{3}{2}$	-1	$\frac{-1}{2}$	1	$\frac{1}{2}$	0	10	
0	y_2	0	1	$\frac{1}{2}$	$\frac{-1}{2}$	$\frac{-1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	7	
0	s_1	0	0	$\frac{-5}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{-3}{2}$	$\frac{-1}{2}$	1	0	

Hence from the last simplex iteration table we have the KKT point $(y_1, y_2) = (10, 7)$. Since D is positive definite, the program is convex. Therefore the point (10,7) is optimal solution.

3.1.1 Convergence of Wolfe Algorithm

Consider the following QPP

$$\begin{aligned} min \quad f(y) =& c^T y + \frac{1}{2} y^T D y \\ s.t: \quad & By \leq d \\ & y \geq 0 \end{aligned} \tag{12}$$

where D is an $k \times k$ symmetric matrix , y and c are $k \times 1$ vectors, d is an $m \times 1$ vector and B is an $m \times k$ matrix .

Now before attempting to prove that whenever D is positive semi definite Wolfe's algorithm will eventually find a feasible point satisfying the KKT condition, We must first demonstrate that such a point actually exists. To obtain this result we need only show that when D is positive semi definite, then problem (12) can not have an unbounded minimum. It will then follow that some feasible point must be a global minimum and therefore a constrained local minimum at which the KKT condition necessary hold.

When Wolfe's algorithm is applied to a quadratic problem of the form(12), with D is positive definite or positive semi definite and $c = 0$, it will obtain with in a finite number of iteration a feasible solution at which KKT conditions are satisfied. Because if D is positive semi definite, then (12) is a convex program, and sufficient for the KKT condition thus guarantees that y_0 is an optimal solution.

3.2 Beale method

Beale has developed a method for solving quadratic programming problem which is based on the basic principles of classical calculus rather than the KKT condition. Beale algorithm is directly applicable to any QPP of the following form:-

$$\begin{aligned} min \quad f(y) = & \quad c^T y + \frac{1}{2} y^T D y \\ s.t: & \quad By \leq d \\ & \quad y \geq 0 \end{aligned}$$

where B is an $m \times k$ matrix , d is an $m \times 1$ vector , c and y are $k \times 1$ vectors, and D is a $k \times k$ symmetric matrix. we do not require D is positive semi definite.

$i.e$ We don't require the objective function to be convex.

The algorithm generally yield a local minimum of a non convex quadratic function, of course if the objective function is convex, the local solution obtained will be global. It is assumed that the problem is degenerate.

3.2.1 Algorism of Beale

Beale method is an iterative procedure and begin with any basic feasible solution of the problem.

Consider the following QPP

$$
\begin{aligned}
min \quad f(y) &= c^T y + \frac{1}{2} y^T D y \\
s.t: \quad & By \leq d \\
& y \geq 0
\end{aligned}
$$

OR

$$
\begin{aligned}
min \quad f(y) = \quad & \sum_{i=1}^{k} c_i y_i + \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} d_{ji} y_i y_j \\
s.t: \quad & \sum_{i=1}^{k} b_{1i} y_i \leq d_1 \\
& \sum_{i=1}^{k} b_{2i} y_i \leq d_2 \\
& . \\
& . \\
& . \\
& \sum_{i=1}^{k} b_{mi} y_i \leq d_m \\
& y \geq 0
\end{aligned}
$$

By introducing slack variable we have the problem:-

$$
\begin{aligned}
min \quad f(y) = \quad & \sum_{i=1}^{k} c_i y_i + \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} d_{ji} y_i y_j \\
s.t: \quad & \sum_{i=1}^{k} b_{1i} y_i + \bar{y}_1 = d_1 \\
& \sum_{i=1}^{k} b_{2i} y_i + \bar{y}_2 = d_2 \\
& . \\
& . \\
& . \\
& \sum_{i=1}^{k} b_{mi} y_i + \bar{y}_m = d_m \\
& y \geq 0, \quad and \quad \bar{y}_m \geq 0
\end{aligned}
$$

This can be equivalent to:

$$\begin{aligned}
\bar{y}_1 = &\quad d_1 - \sum_{i=1}^{k} b_{1i} y_i \\
\bar{y}_2 = &\quad d_2 - \sum_{i=1}^{k} b_{2i} y_i \\
&\quad . \\
&\quad . \\
&\quad . \\
\bar{y}_m = &\quad d_m - \sum_{i=1}^{k} b_{mi} y_i \\
&\quad y \geq 0, \quad and \quad \bar{y}_m \geq 0
\end{aligned}$$

Let $\bar{y}_1, \bar{y}_2, \bar{y}_3, ..., \bar{y}_m$ are basic variable and y_i be non basic variable for $i = 1, 2, 3, 4, ..., m$

The basic variable can thus be written interns of non basic variables as:-

$$\bar{y}_h = d_h - \sum_{i=1}^{k} b_{hi} y_i$$

where $\bar{y}_h$ is the h^{th} basic variable.

We can formulate an algorithm to solve a quadratic programming problem using algorithm of Beale .

We consider the following steps:-

step 1. Select arbitrary basic and non basic variable(s).

step 2. Express the basic variables $\bar{y}_h, \quad h = 1, 2, 3, ..., m$ in terms of non basic variables

we will have:

$$\bar{y}_h = d_h - \sum_{i=1}^{k} b_{hi} y_i$$

step 3. Express the objective function $f(y)$ in terms of the non basic variables in the symmetric form:

$$f(y) = \sum_{i=1}^{k} c_i y_i + \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} d_{ji} y_i y_j$$

step 4. Consider the partial derivative of $f(y)$ with respect to any one of non basic variables say y_t.

$$\frac{1}{2}\frac{\partial f}{\partial y_t}=\frac{1}{2}(c_t+\sum_{i=1}^{k}d_{ti}y_i), \quad for \quad t=1,2,....,k$$

With other non basic variables held at zero.

A If $\frac{\partial f}{\partial y_t}\geq 0 \quad \forall t$, then the current solution is optimal.

B If $\frac{\partial f}{\partial y_t}<0$ for at least one t, then y_t increase until either:

- Some basic variables become zero , or
- $\frac{\partial f}{\partial y_t}$ vanish.

step 5. To determine the change in the basis , calculate

$$\min \quad \Big[\frac{d_h}{|b_{ht}|},\frac{|c_t|}{d_{tt}}, \quad t=1,2,3,...,k\Big] \quad for \quad b_{ht}>0, \quad q_{tt}>0$$

I. If the minimum occurs for some $h=v$, then y_v become non basic.

II. If the minimum occurs for the second term , introduce anew non basic variable u_j defined by:

$$u_j=\frac{1}{2}\frac{\partial f}{\partial y_t}=\frac{1}{2}\Big(c_t+\sum_{i=1}^{k}d_{ti}y_i\Big)$$

u_j is unrestricted and is called a free variable.
Thus will lead to one equation and one more basic variable then before.

step 6. Go to step 2 and repeat the process until $\frac{\partial f}{\partial y_t}\geq 0$
for all t and $\frac{\partial f}{\partial u_j}=0$ for all free variable.

step 7. Obtain the local (global if f is convex) optimal solution and the value of $min \quad f$ by setting non basic variable equal to zero in their expression .

Example: Consider the following QP problem.

$$\begin{aligned} min \quad f(x) = & \; -y_1 - 2y_2 + y_2^2 \\ s.t: & \; y_1 + 2y_2 \le 4 \\ & \; 3y_1 + 2y_2 \le 6 \\ & \; y_1, y_2 \ge 0 \end{aligned}$$

By introducing slack variable , we have the problem,

$$\begin{aligned} min \quad f(y) = & \; -y_1 - 2y_2 + y_2^2 \\ s.t: & \; y_1 + 2y_2 + y_3 = 4 \\ & \; 3y_1 + 2y_2 + y_4 = 6 \\ & \; y_1, y_2, y_3, y_4 \ge 0 \end{aligned}$$

Now to solve this problem using Beale method:

Step 1. Take y_1 and y_2 as basic variable.

Step 2. Express them in terms of non basic variable as:

$$y_1 = 1 + \frac{1}{2}y_3 - \frac{1}{2}y_4 \quad and \quad y_2 = \frac{3}{2} + \frac{1}{4}y_4 - \frac{3}{4}y_3$$

Step 3. $f(y) = -y_1 - 2y_2 + y_2^2$

$\Rightarrow \quad f(y_3, y_4) = -4 + y_3 + (\frac{3}{2} + \frac{1}{4}y_4 - \frac{3}{4}y_3)^2$

Step 4. $\frac{\partial f}{\partial y_3} = \frac{-5}{4} - \frac{3}{8}y_4 + \frac{9}{8}y_3 \quad , \left(\frac{\partial f}{\partial y_3}\right)_{at \quad y_3, y_4 = 0} = \frac{-5}{4} < 0 \quad and \quad \frac{1}{2}\frac{\partial f}{\partial y_4} = \frac{-5}{8} - \frac{3}{16}y_4 + \frac{9}{16}y_3$

$\Rightarrow \quad \frac{\partial f}{\partial y_4} = \frac{3}{4} + \frac{1}{8}y_4 - \frac{3}{8}y_3, \quad \left(\frac{\partial f}{\partial y_4}\right)_{at \quad y_3, y_4 = 0} = \frac{3}{4} > 0 \quad and \quad \frac{1}{2}\frac{\partial f}{\partial y_4} = \frac{3}{8} + \frac{1}{16}y_4 - \frac{3}{16}y_3$

Step 5. Hence it is profitable to increase y_3 and we calculate:

$$\min \quad \left[\frac{2}{1}, \frac{10}{9}\right] = \frac{10}{9}$$

We introduce a free non basic variable

$$u_1 = \frac{1}{2}\frac{\partial f}{\partial y_3} = \frac{13}{8} + \frac{3}{16}y_4 - \frac{9}{8}y_3$$

We again express basic variables in terms of non basic variables:

$$\begin{aligned} y_3 &= \frac{10}{9} + \frac{1}{3}y_4 + \frac{16}{9}u_1 \\ y_2 &= \frac{2}{3} - \frac{4}{3}u_1 \\ y_1 &= \frac{14}{9} - \frac{1}{3}y_4 + \frac{8}{9}u_1 \\ f(y) = & \; -y_1 - 2y_2 + y_2^2 \\ f(y_4, u_1) &= \frac{-26}{9} + \frac{1}{3}y_4 + \frac{16}{9}u_1 + \left(\frac{2}{3} - \frac{4}{3}u_1\right)^2 \end{aligned}$$

Step 6. Now from $f(y_4, u_1) = \frac{-26}{9} + \frac{1}{3}y_4 + \frac{16}{9}u_1 + \left(\frac{2}{3} - \frac{4}{3}u_1\right)^2$, we get

$\frac{\partial f}{\partial u_1} = \frac{32}{9}u_1$, $\left(\frac{\partial f}{\partial u_1}\right)_{at \quad u_1=0} = 0$, and $\frac{\partial f}{\partial y_4} = \frac{1}{3} > 0$

Step 7. Here, the minimum solution is achieved

$$y_1 = \frac{14}{9}, y_2 = \frac{2}{3}, \quad and \quad \min f = \frac{-22}{9}$$

4 Conclusions

For any given quadratic programming problem the major factor limiting the choice among the solution method is character of D in the objective function. If D is indefinite or negative semi definite, convergence cannot be guaranteed for either Wolfe's or lagrangian method, and it becomes necessary to use some type of constrained hill-climbing approach to find the various local optima that may exist. This strategy is embodied in the quadratic programming algorithm of Beale, as well as in many other more general methods that can be applied to all linearly constrained problem, regardless of the nature of their objective function.

In choosing among the algorithms, there are a number of different consideration to bear in mind. For example, suppose analyst wishes to solve quadratic programming problem with a convex objective function as quickly or as cheaply readily available, so that he is forced to modify an existing linear programming code. In such a case his primary consideration will be that of expedience, and it is clear that he should plan to use Wolfe's algorithm. Conversion from the simplex to the "Wolfe-simplex" method requires little more than modifying the logic by which the entering variable is chosen, so that the complementary slackness condition $x_i v_i = 0, \quad i = 1, 2, 3, 4, \ldots.., n$ are preserved. This may be a decisive advantage for users with limited budgets.

5 References

1. Bazara M.S, Sherall H.D, Shetty C.M:*Nonlinear programming Theory and Algorithms,* ($2^{nd} edition$) 1993

2. Jorge Nocedal, Stephen J.Wright:*Numerical Optimization,* Springer operation research.

3. Paul A. Jensen and Jonathan F: *Operations Research Model and Methods,* *http://www.me.utexas.edu/ jensen/ORMM/supplements/methods/nlpmethod/ S2 quadratic.pdf*

4. R.Fletcher. *A general quadratic programming algorithem.*J.Inst.math.Applics.,1971

5. Donald M. Simmons: *Nonlinear programming for Operation Research.*

6. Kirtiwant P.$Ghadle^1$,Tanaji S.$Pawar^2$ New Approach For *Wolfe's Modified Simplex Method to solve quadratic programming problem* (International journal of research in Engineering and Technology)

7. Ding Zhudu *Optimization theory and methods non linear programming*

8. Sinha S.M: *Mathematical Programming Theory and Algorithms,* (First edition)2006

9. John Burkardt :*Lemke solver for quadratic programming,* *http://people.sc.fsu.edu/ burkardt/msrc/lemke/lemke.html,05 December 2007*

10. Deumlich R : *Optimization Theory II*, Undergraduate text, Addis Ababa 1996.

11. Bapi Chatterjee : *Quadratic Programming by Wolf method,* *htt://www.Mathworks.com*, 28 April 2010