A Graduate Seminar Report

On

The Theory of Duality in Mathematical Programming

By

Berhanu Tekle

School of Graduate Studies

Addis Ababa University

June, 1995

Addis Ababa

PREFACE

This is a report of two semesters work , which is divided into two parts.

PartI, which is the prelimenary part, deals with convex sets, convex functions and so on. Here, we give basic definitions and facts that are important for the development of the subsequent section.

Part II, in this part we formulate relations of duality and prove equivalence theorems with respect to saddle-point , duality and minimax theorems. Here what intersts us most is what we call a Lagrangian-form and its corresponding dual problems.

ACKNOWLEDGEMENT

I would like to express my special appreciation to my adviser and instructor Prof.Dr.R.Deumlich,for his constructive criticism and valuable advice.

TABLE OF CONTENTS.

PART I

PRELIMENARIES

1.1. Convex Sets .

Def. 1.1.1: A set $C \subseteq \mathbb{R}^n$ is *convex*, if for any two elements $u, v \in C$ and any real number $\lambda, 0 \le \lambda \le 1$ we have ,

$$x = \lambda u + (1-\lambda)v \in C .$$

An element $x = \sum_{i=1}^{r} \lambda_i x_i$ is a convex combination of vectors $x_1, \ldots, x_r \in \mathbb{R}^n$ if the real numbers $\lambda_1, \ldots, \lambda_r$ are non-negative and $\sum_{i=1}^{r} \lambda_i = 1$.

Theorem 1.1.1: Let $C \subseteq \mathbb{R}^n$ be a convex set. Every convex combinition of vectors of C is an element of C.

Proof: For $r = 1$ the assertion of the theorem is trivial.

Assume that the theorem is true for $r-1$ where $r > 1$. Now consider the vectors $x_1, x_2, \ldots, x_r \in C$, and the non-negative numbers $\lambda_1, \lambda_2, \ldots, \lambda_r$, where $\sum_{i=1}^{r} \lambda_i = 1$.

If $\sum_{i=1}^{r-1} \lambda_i = 0$, then $x = \sum_{i=1}^{r} \lambda_i x_i = x_r \in C$.

If $\lambda = \sum_{i=1}^{r-1} \lambda_i \neq 0$, then we have

$$x = \sum_{i=1}^{r} \lambda_i x_i = \lambda \sum_{i=1}^{r} \frac{\lambda_i}{\lambda} x_i + \lambda_r x_r$$

clearly $\sum_{i=1}^{r-1} \frac{\lambda_i}{\lambda} = 1$, thus by induction assumption

$$\bar{x} = \sum_{i=1}^{r-1} \frac{\lambda_i}{\lambda} x_i \in C .$$

Moreover $\sum_{i=1}^{r} \lambda_i = 1$ implies $0 < \lambda \leq 1$ and $\lambda_r = 1-\lambda$.

Thus, by definition $x = \lambda\bar{x} + \lambda_r x_r \in C$.

1.2 Convex and concave functions.

Def. 1.2.1: Let $C \subseteq \mathbb{R}^n$ be a convex set, we call a function f defined on C a *convex function* ,if for any element x,y in C and $0 \leq \lambda \leq 1$,we have

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y).$$

If f is a convex function we call -f a *concave* function. Thus f is concave if -f is convex.

Fact: Let f be a convex function defined on a convex set $C \subseteq \mathbb{R}^n$. If x^0 is an interior point of C,then f is continous at x^0.

Note: A convex function need not necessarily be continuous on the boundary. For example consider the function,

$$f(x)=\begin{cases} 2 & \text{for } x = 0 \\ 1 & \text{for } x \in (0,1] \end{cases}$$

f is convex on [0,1] but not continuous at 0.

1.3 Frechet Differentiablity .

Let f be a real valued function defined on $C \subseteq \mathbb{R}^n$. If x_0 is an interior point of C then for a suitable radius $\rho > 0$ the sphere

$$S_\rho(x_0) = \{ x \in \mathbb{R}^n : \|x - x_0\| < \rho \},$$

is completely contained in C. If $t \in \mathbb{R}^n$ is a vector such that $\|t\| < \rho$, then obviously the point $x = x_o + t$ belongs to $S_\rho(x_o)$.

Def. 1.3.1 : A linear mapping l from $\mathbb{R}^n$ into $\mathbb{R}$ is called a *Frechet differential* of the function f at x_0 in the direction of t, if there exist a function $a(x_0, t)$ such that,

$$f(x_0 + t) = f(x_0) + l(t) + a(x_0, t)$$

with, $$\lim_{\|t\| \to 0} \frac{|a(x_0, t)|}{\| t \|} = 0$$

Note: The F-differential dependes on the point x_0. If f possesses at a point x_0 a Frechet differential, we say that f is Frechet differentiable.

Notation: We use the notation $(df)_{x_0}$ for the Frechet differential

Some facts:

1. The Frechet differential, at a point is uniquely defined.
2. If a function f is F-differentiable at a point x_0, then f is continuous at x_0.

3. If f possesses, at an interior point $x_0 \in C$ $(C \subseteq \mathbb{R}^n)$, first continuous partial derivative, then the F-differential of f at x_0 exists and we have,

$$(df)_{x_0} = \mathrm{grad} f(x_0).$$

Next we state and proof two theorems which will be useful later.

Theorem 1.3.1: If x_0 is an interior point of the set

$$\mathbb{R}^{n+} = \{ x \in \mathbb{R}^n : x \geq 0 \}$$

and if the differentiable function f possesses a local maximum at x_0, then we have

$$(df)_{x_0} = \mathrm{grad} f(x_0) \leq \mathbf{0} ,$$

$$\langle (df)_{x_0}, x_0 \rangle = \langle \mathrm{grad}\ f(x_0), x_0) = 0.$$

Proof: Let f be F-differentiable, and x_0 an interior point of $\mathbb{R}^{n+}$. Because f has at $x_0 \in \mathbb{R}^{n+}$ a local maximum there exists an $\varepsilon > 0$ such that for any $x \in \mathbb{R}^{n+} \cap S_\varepsilon(x_0)$ we have

$$0 \geq f(x) - f(x_0) = f(x_0 + x - x_0) - f(x_0) = (df)_{x_0} + a(x_0, x - x_0)$$

If we choose a real number η with $o < \eta < \varepsilon$ then all points $x = x_0 + \eta e_i$ belongs to $\mathbb{R}^{n+} \cap S_\varepsilon(x_0)$ and we get by substitution,

$$0 \geq \frac{f(x_0 + \eta e_i) - f(x_0)}{\eta} = (df)_{x_0}(e_i) + \frac{a(x_0, \eta e_i)}{\eta}$$

on taking the limit as $\eta \longrightarrow 0$,

$$0 \geq D_i f(x_0) = (df)_{x_0}(e_i)$$

this inequality for i= 1,2,...,n leads to

$\operatorname{grad} f(x_0) \leq 0.$

This proves the first part.

To prove the second part consider the scalar product

$$\langle \operatorname{grad} f(x_0), x_0 \rangle = \sum_{i=1}^{n} \alpha_i \frac{\partial f}{\partial x_i}(x_0) \ , x_o = (\alpha_1^0, \ldots, \alpha_n^0)$$

For $x_0 \in \mathbb{R}^{n+}$ we have $\alpha_i^0 \geq 0$, i= 1,...,n , thus only those strictly positive components of x_0 are of interest here. If α_k^0 were strictly positive, then there exists $\eta > o$ such that, $x = x_0 - \eta e_k \geq 0$

Furthermore if $\eta \leq \varepsilon$,we have ,

$$0 \geq f(x_0 - \eta e_k) - f(x_0) = (df)_{x_0}(-\eta e_k) + a(x_0, -\eta e_k)$$

$$\Rightarrow 0 \geq \frac{f(x_o - \eta e_k) - f(x_0)}{\eta} = -(df)_{x_0}(e_k) + \frac{a(x_0, -\eta e_k)}{\eta}$$

from this inequality we get putting $\beta = -\eta$,

$$0 \geq \frac{f(x_0 + \beta e_k) - f(x_0)}{-\beta} = -(df)_{x_0}(e_k) + \frac{a(x_0, \beta e_k)}{-\beta}$$

Thus as $\beta \longrightarrow 0$ we get,

$$0 \geq -D_k f(x_0) = -(df)_{x_0}(e_k)$$

$$\Rightarrow D_k f(x_0) = (df)_{x_0}(e_k) \geq 0.$$

Because in general we have $\operatorname{grad} f(x_0) \leq 0$ which implies

$D_k f(x_0) \leq 0$ for all k ,it follows that

$D_k f(x_0) = 0$,in case $\alpha_k^0 > 0$

This proves the theorem.

Theorem 1.3.2: Let f be a convex function defined on a convex set C of which x_0 is an interior point. If f is Frechet differentiable at x_0, then the inequality

$$f(x)-f(x_0) \geq (df)_{x_0}(x-x_0)= \langle \text{grad } f(x_0), x-x_0 \rangle$$

holds for all $x \in C$.

Proof: Let $x \in C$ be a point and s be a vector. Let $t > 0$ be such

that $x+ts \in C$

Then $g(s) = \dfrac{f(x+ts)-f(x)}{t\| s \|}$, is increasing.

Now let s be any vector with $\|s\| > 0$ and $x_0+s \in C$.

If f is F-differentiable at an interior point x_0, then

$$\frac{f(x_0+ s)-f(x_0)}{\| s \|} = (df)_{x_0}\left(\frac{s}{\|s\|}\right) + \frac{a(x_0,s)}{\|s\|}.$$

In this equation, we calculate the limits in such a way that we do not change the direction of the vector s, but the length of s tends to 0. The vector $\frac{s}{\|s\|}$ is constant, namely the unit vector in the direction of s. Thus we get

$$\lim_{\|s\| \to 0} \frac{f(x_0+s) - f(x_0)}{\| s \|} = (df)_{x_0}\left(\frac{s}{\|s\|}\right).$$

From this follows, because of the monotonocity of the difference quotient,

$$\frac{f(x_0+s)-f(x_0)}{\|s\|} \geq (df)_{x_0}\left(\frac{s}{\|s\|}\right),$$

$$\Rightarrow \; f(x_0+s)-f(x_0) \geq (df)_{x_0}\left(\frac{s}{\|s\|}\right).$$

If we put $x = x_0 + s$ we get the assertion of the theorem.

For the case $x = x_0$ the assertion is trivial.

PART II

THEORY OF DUALITY.

2.1. Basic Notions.

Generally speaking the basic idea of duality theory consists in considering a given minimum problem

$$\inf_{x \in A} f(x) = \alpha \qquad (P)$$

together with a maximum problem

$$\sup_{y \in B} g(y) = \beta \qquad (D)$$

where, $\beta \le \alpha$. Then we will try to establish relationships between (P) and (D).

For example, the study of duality in linear programming sharpens our understanding of the simplex procedure and motivates certain alternative metods. Indeed the simultaneous consideration of a problem from both the primal and the dual point of views often provides significant computational advantages as well as economic insightes.

2.2. Fundemental theorems.

In view of constructing a general theory of duality of mathematical programming, we conisider what we call a *Lagrangian-form* which produces two corrsponding dual problems. Then we will attempt to find relations between the Lagrangian form and the dual pair of problems. Such relations are expressed in the form of duality theorems, saddle-point theorems or minimax theorems.

Here we will confine our attention to finite dimensional Euclidean spaces. And the situtation will be such that we always assume the existence of solutions of one of the three theorems; and we try to find a possible equivalence to one another.

Now let μ and ν be real valued functions defined on $M \subseteq \mathbb{R}^n$ and $N \subseteq \mathbb{R}^n$ respectively. If for any $(x,y) \in M \times N$ we have,

$$\mu(x) \leq \nu(y)$$

then the supremum of μ and the infimum of ν exists and

$$\sup_{x \in M} \mu(x) \leq \inf_{y \in N} \nu(y)$$

In the theory of mathematical programming, the following results are of high theoritical value.

(A_1) Under certain necessary and suifficeny conditions we have,

$$\sup_{x \in M} \mu(x) = \inf_{y \in N} \nu(y)$$

A statement of type (A_1) is called *duality theorem.*

(A_2) Let Ψ be a real valued function defined on $K \times L \subseteq \mathbb{R}^n \times \mathbb{R}^m$.

Then under certain necessary and sufficent conditions we have,

$$\sup_{x \in K} \inf_{y \in L} \Psi(x,y) = \inf_{y \in L} \sup_{x \in K} \Psi(x,y).$$

A statement of type (A_2)is called a *minimax theorem.*

(A_3) Under certain conditions a pair of vectors $(x^0,y^0) \in K \times L$ exists such that for any $(x,y) \in K \times L$ we have,

$$\Psi(x,y^0) \le \Psi(x^0,y^0) \le \Psi(x^0,y).$$

A statement of type (A_3) is called *saddle-point theorem.*

The main purpose of the theory of duality in mathematical programing is to make inquiries about a corresponding pair of optimization problems, namely the primal problem to find,

$$W = \sup_{x \in M} \mu(x) \qquad (P)$$

and the dual problem to find,

$$W^* = \inf_{y \in N} \nu(y), \qquad (D)$$

and to discover relations between the corresponding duality, minimax and saddle point theorems.

Suppose for the problems (P) and (D), there exist a pair $(x^0,y^0) \in M \times N$ such that,

$$\sup_{x \in M} \mu(x) = \max_{x \in M} \mu(x) = \mu(x^0),$$

$$\inf_{y \in N} \nu(y) = \min_{y \in N} \nu(y) = \nu(y^0).$$

Then we say that the problems (P) and (D) are *realizable* with respect to M and N respectively. In this case we call x^0 and y^0 *optimal solutions.* To recognize this idea we add to the problems (A_1) to (A_3) the following problem.

(A_4) Find necessary and sufficient conditions such that problems (P) and (D) are realizable. Assertations in this direction are called *realization theorems*.

2.3. Relation between Saddle Point and Minimax Theorem.

Def 2.3. 1: Let $K \subseteq \mathbb{R}^n$ and $L \subseteq \mathbb{R}^m$ and Ψ be a real valued function defined on $K \times L$. A point (x^0, y^0) is said to be a saddle point of Ψ with respect to $K \times L$ iff for all $(x,y) \in K \times L$ the following holds:

$$\Psi(x, y^0) \le \Psi(x^0, y^0) \le \Psi(x^0, y).$$

In this section we confine to the problem of finding relations between duality theorem (A_1), minimax theorem (A_2) and saddle point theorem (A_3). To do so we proceed as follows. Let Ψ be a real valued function defined on $K \times L \subseteq \mathbb{R}^n \times \mathbb{R}^m$. Then we define:

$$\mu(x) = \inf_{y \in L} \Psi(x, y),$$

$$\nu(y) = \sup_{x \in K} \Psi(x, y).$$

Let $M \times N \subseteq K \times L$ be the domain of definition of the functions μ and ν. This means M contains all points x in K for which the infimum exists (i.e. the infimum is finite); and N contains those points $y \in L$ for which the supreimum is finite. The sets M and N can be empty or we may have the case $M = K$ and $N = L$. Then for any $(x,y) \in M \times L$ we have,

$$\mu(x) = \inf_{y \in L} \Psi(x,y) \le \Psi(x,y),$$

and for any $(x,y) \in K \times N$

$$\nu(y) = \sup_{x \in K} \Psi(x,y) \ge \Psi(x,y).$$

Thus it follows that for any $(x,y) \in M \times L \cap K \times N = M \times N$

$$\mu(x) \le \Psi(x,y) \le \nu(y).$$

Consequently, there exists a supremum for μ over M and an infimum for ν over N . Thus we are able to formulate a primal problem (P) for the function μ and a dual problem (D) for the function ν, namely :

$$\text{Find } \alpha = \inf_{N} \nu(y) \text{ , and} \qquad \text{(D)}$$

$$\text{Find } \beta = \sup_{M} \mu(x) \qquad \text{(P)}$$

If this two problems are realizable i.e. if there exists a pair $(x^0,y^0) \in M \times N$ such that

$$\sup_{x \in M} \mu(x) = \max_{x \in M} \mu(x) = \mu(x^0),$$

$$\inf_{y \in N} \nu(y) = \min_{y \in N} \nu(y) = \nu(y^0),$$

and if ,

$$\inf_{y \in L} \Psi(x,y) = \min_{y \in L} \Psi(x,y),$$

$$\sup_{x \in K} \Psi(x,y) = \max_{x \in K} \Psi(x,y),$$

then the(theorem of duality),

$$\mu(x^0) = \max_{x \in M} \mu(x) = \min_{y \in N} \nu(y) = \nu(y^0)$$

is obviously equivalent to (the minimax theorem):

$$\max_{x \in M} \min_{y \in L} \Psi(x,y) = \min_{y \in N} \max_{x \in K} \Psi(x,y).$$

Also we find the following connection between minimax theorem and saddle point theorem

Theorem 2.3.1: If there exist a minimax point $(x^0, y^0) \in K \times L$ such that

$$\psi(x^0, y^0) = \max_{x \in K} \min_{y \in L} \Psi(x,y) = \min_{y \in L} \max_{x \in K} \Psi(x,y),$$

then (x^0, y^0) is a saddle point of Ψ with respect to $K \times L$.

Proof: Let $(x^0, y^0) \in K \times L$ be a minimax point such that

$$\Psi(x^0, y^0) = \max_{x \in K} \min_{y \in L} \Psi(x,y) = \min_{y \in L} \max_{x \in K} \Psi(x,y)$$

$$\Rightarrow \max_{x \in K} \Psi(x, y^0) = \Psi(x^0, y^0) = \min_{y \in L} \Psi(x^0, y)$$

consequently, for any $(x,y) \in K \times L$, we have

$$\Psi(x, y^0) \le \Psi(x^0, y^0) \le \Psi(x^0, y).$$

Hence the theorem is proved.

Remark: We cannot conclude the existence of a minimax point, as we have done in Theorem 2.3.1, from the existence of saddle point. This is because

$$\min_{y \in L} \Psi(x,y) \text{ may not exist for all } x \in K \text{ , or}$$

$$\max_{x \in K} \Psi(x,y) \text{ may not exist for all } y \in L.$$

We will demonstrate this idea with example.

Example 1. Let $\Psi: \mathbb{R} \times \mathbb{R}^+ - \mathbb{R}$, be given by

$$\Psi(x,y) = x - xy.$$

since, $\Psi(x,1) \le \Psi(0,1) \le \Psi(0,y)$, $\forall\ (x,y) \in \mathbb{R} \times \mathbb{R}^+$

the point (0,1) is a saddle point of Ψ.

But,

$$\min_{y \in \mathbb{R}^+} \Psi(x,y) = \min_{y \in \mathbb{R}^+} x(1-y) \quad \text{does not exist } \forall\ x \in \mathbb{R}\backslash\{0\}.$$

Let now $M \subseteq K$ be the domain of definition for

$$\mu(x) = \min_{y \in L} \Psi(x,y),$$

and $N \subseteq L$ be the domain of definition for

$$\nu(y) = \max_{x \in K} \Psi(x,y).$$

Then under the above assumptions we have the following theorem

Theorem 2.3.2: If $\Psi(x,y)$ possesses a saddle-point (x^0, y^0) with respect to $K \times L$, then for $\Psi(x,y)$ it holds:

$$\Psi(x^0,y^0) = \max_{x \in M} \min_{y \in L} \Psi(x,y) = \min_{y \in N} \max_{x \in K} \Psi(x,y).$$

Proof: (x^0, y^0) is a saddle point of Ψ with respect to $K \times L$

$$\text{implies } \Psi(x,y^0) \le \Psi(x^0,y^0) \le \Psi(x^0,y),\ \forall\ (x,y) \in K \times L$$

$$\Rightarrow \nu(y^0) = \max_{x\in K} \Psi(x,y^0) = \Psi(x^0,y^0) = \min_{y\in L} \Psi(x^0,y) = \mu(x^0)$$

$$\Rightarrow x^0 \in M \text{ and } y^0 \in N.$$

Hence M and N are non empty.

Furthermore, since $\mu(x) = \min_{y \in L} \Psi(x,y)$ and

$$\nu(y) = \max_{x \in K} \Psi(x,y)$$

we have that $\mu(x) \le \nu(y)$, $\forall\ (x,y) \in M \times N$,and thus

$$\mu(x) \le \nu(y) \quad \forall\ x \in M$$

$$\Rightarrow \mu(x) \le \mu(x^0) \quad \forall\ x \in M\ ;$$

and $$\nu(y^0) = \mu(x^0) \le \nu(y) \quad \forall\ y \in N$$

$$\Rightarrow \nu(y^0) \le \nu(y)\ , \quad \forall\ y \in N.$$

From this it follows that ,

$$\max_{x \in M} \mu(x) = \mu(x^0) = \nu(y^0) = \min_{y \in N} \nu(y)$$

$$\Rightarrow \max_{x\in M}\min_{y\in L} \Psi(x,y) = \min_{y\in N}\max_{x\in K} \Psi(x,y) = \Psi(x^0,y^0).$$

This proves the theorem.

Remark: From Theorem 2.3.2, it also follows immediately that the existence of a saddle point of Ψ implies that there exist exist functions μ and ν with,

$$\mu(x) = \min_{y \in L} \Psi(x,y),$$

$$\nu(y) = \max_{x \in K} \Psi(x,y)$$

and the domains of definition M and N of μ and ν are nonempty sets. Furthermore, a relation of duality of the form

$$\max_{x \in M} \mu(x) = \min_{y \in N} \nu(y)$$

is true.

In this case the problems (P) and (D) namely,

$$W = \sup_{x \in M} \mu(x) = \max_{x \in M} \mu(x) \qquad \text{(P)}$$

$$W^{*} = \inf_{y \in N} \nu(y) = \min_{y \in N} \nu(y) \qquad \text{(D)}$$

are realizable ,and x^0 and y^0 are optimal solutions of the problems (P) and (D) respectively.

2.4. Special Lagrangian Form.

In this section we treat special types of Lagrangian-forms which are particularly suitable to construct well known problems of optimizations. For this purpose the idea of congugate functions play an important role.

Let φ be a real valued function defined on $K \times L \subseteq \mathbb{R}^n \times \mathbb{R}^m$, α be a function defined on K and β be a function defined on L. α and β are real valued Consider the Lagrangian form

$$\Psi(x,y) = \varphi(x,y)-\alpha(x)-\beta(y) \ ,$$

defined on $K \times L$.

Now we define the functions

$$\alpha^{\bullet}(y) = \sup_{x \in K} \{\varphi(x,y)-\alpha(x)\},$$

$$\beta^{\bullet}(x) = \inf_{y \in L} \{\varphi(x,y)-\beta(y)\}.$$

Let $N \subseteq L$ and $M \subseteq K$ be the domains of definition of the functions $\alpha^{\bullet}$ and $\beta^{\bullet}$ respectively. $\alpha^{\bullet}$ and $\beta^{\bullet}$ are called the *conjugate functions* of α and β respectively .

For example: If $M = N$, $\varphi(x,y) = \langle x,y\rangle$ and α is convex and β is concave functions, the functions $\alpha^{\bullet}$ and $\beta^{\bullet}$ are respectively conjugate convex and conjugate concave functions.

Now as we have done in section 2.3, we consider the functions:

$$\mu(x) = \inf_{y \in L} \Psi(x,y) = \inf_{y \in L} \{\varphi(x,y)-\alpha(x)-\beta(y)\}$$

$$= \beta^{\bullet}(x)-\alpha(x)$$

defined on $M \subseteq K$, and

$$\nu(y) = \sup_{x \in K} \Psi(x,y) = \sup_{x \in K} \{\varphi(x,y)-\alpha(x)-\beta(y)\}$$

$$= \alpha^{\bullet}(y)-\beta(y)$$

defined on $N \subseteq L$.

Then for all $(x,y) \in M \times L$ we have

$$\mu(x) \le \Psi(x,y)$$

and for all $(x,y) \in K \times N$ we have

$$\nu(y) \geq \Psi(x,y).$$

Hence $\forall\ (x,y) \in M \times L \cap K \times N = M \times N$ we have

$$\left.\begin{aligned} &\mu(x) \leq \Psi(x,y) \leq \nu(y), \\ &\Rightarrow \beta^{\bullet}(x)-\alpha(x) \leq \alpha^{\bullet}(y)-\beta(y). \end{aligned}\right\} \tag{2.4.1}$$

From this inequality we get a pair of dual problems :

$$\text{Find } W = \sup_{x \in M} \{\beta^{\bullet}(x)-\alpha(x)\}, \tag{P}$$

$$\text{Find } W^{\bullet} = \inf_{y \in N} \{\alpha^{\bullet}(y)-\beta(y)\}. \tag{D}$$

Regarding this two problems, we have the following theorem.

Theorem 2.4.1: The following three statements are equivalent.

(i). There exist a pair $(x^0,y^0) \in K \times L$ such that for any $(x,y) \in K \times L$, we have

$$\Psi(x\ ,y^0) \leq \Psi(x^0,y^0) \leq \Psi(x^0,y)$$

(ii). It is true that

$$\inf_{y \in L} \Psi(x^0,y) = \min_{y \in L} \Psi(x^0,y) = \beta^{\bullet}(x^0)-\alpha(x^0) = \Psi(x^0,y^0)$$

$$\sup_{x \in K} \Psi(x,y^0) = \max_{x \in K} \Psi(x,y^0) = \alpha^{\bullet}(y^0)-\beta(y^0) = \Psi(x^0,y^0).$$

Furthermore the domains of definition M and N of μ and ν respectively are non-empty sets and the problems (P) and (D) are realziable.and,

$$\max_{x \in M} \{\beta^{\bullet}(x)-\alpha(x)\} = \min_{y \in N} \{\alpha^{\bullet}(y)-\beta(y)\}.$$

(iii). There exist a pair $(x^0, y^0) \in M \times N$ such that

$$\varphi(x^0, y^0) = \alpha(x^0) + \alpha^*(y^0)$$

$$\varphi(x^0, y^0) = \beta(y^0) + \beta^*(y^0).$$

Proof: We prove the theorem by showing, (i)⇒ (iii)⇒ (ii)⇒ (i).

(i)⇒ (iii). Suppose (i) holds true, then

$$\varphi(x, y^0)-\alpha(x)-\beta(y^0)\le \varphi(x^0, y^0)-\alpha(x^0)-\beta(y^0)\le \varphi(x^0, y)-\alpha(x^0)-\beta(y) \qquad (1)$$

this implies for all $x \in K$

$$\varphi(x, y^0) - \alpha(x) \le \varphi(x^0, y^0) - \alpha(x^0)$$

$$\Rightarrow \sup_{x \in K} \{\varphi(x, y^0)-\alpha(x)\} \le \varphi(x^0, y^0) - \alpha(x^0).$$

But since $x^0 \in K$ we also have

$$\varphi(x^0, y^0)-\alpha(x^0) \le \sup_{x\in K} \{\varphi(x, y^0)-\alpha(x)\}$$

consequently we have,

$$\alpha^*(y^0) = \sup_{x \in K} \{\varphi(x, y^0)-\alpha(x)\}= \max_{x \in K} \{\varphi(x, y^0)-\alpha(x)\}= \varphi(x^0, y^0)-\alpha(x^0). \qquad (2)$$

Thus $y^0 \in N \subseteq L$ i.e. N is non-empty.

Again from inequality (1), we get

$$\varphi(x^0, y^0)-\beta(y^0) \le \varphi(x^0, y)-\beta(y) \;, \forall\, y \in L$$

$$\Rightarrow \varphi(x^0, y^0)-\beta(y^0) \le \inf_{y \in L} \{\varphi(x^0, y)-\beta(y)\}$$

and since $y^0 \in L$ we have ,

$$\inf_{y \in L} \{\varphi(x^0, y)-\beta(y)\} \le \varphi(x^0, y^0)-\beta(y^0),$$

consequently,

$$\beta^{\bullet}(x^0) = \varphi(x^0,y^0)-\beta(y^0)= \inf_{y \in L} [\varphi(x^0,y)-\beta(y)]= \min_{y \in L} [\varphi(x^0,y)-\beta(y)]. \quad (3)$$

Thus $x^0 \in M \subseteq K$; i.e. M is non-empty.

Now from (2) and (3) it follows that, there is a pair (x^0,y^0) in $M \times N$ such that,

$$\varphi(x^0,y^0) = \alpha^{\bullet}(y^0) + \alpha(x^0),$$

$$\varphi(x^0,y^0) = \beta^{\bullet}(x^0) + \beta(y^0).$$

This proves (i) $\Rightarrow$ (iii).

(iii) $\Rightarrow$ (ii); Suppose (iii) holds.

Then there exist a pair $(x^0,y^0) \in M \times N$ such that,

$$\left.\begin{aligned} \varphi(x^0,y^0) &= \alpha(x^0)+ \alpha^{\bullet}(y^0) \\ \varphi(x^0,y^0) &= \beta(y^0)+ \beta^{\bullet}(x^0) \end{aligned}\right\} \quad (4)$$

Clearly M and N are non empty. Furthermore, from (4) and definition of $\alpha^{\bullet}$ and $\beta^{\bullet}$ it follows that:

$$\alpha^{\bullet}(y^0)= \sup_{x\in K}[\varphi(x,y^0)-\alpha(x)]= \varphi(x^0,y^0)-\alpha(x^0)= \max_{x\in K}[\varphi(x,y^0)-\alpha(x)], \text{and}$$

$$\beta^{\bullet}(x^0)= \inf_{y\in L}[\varphi(x^0,y)-\beta(y)]= \varphi(x^0,y^0)-\beta(y^0)= \min_{y\in L}[\varphi(x^0,y)-\beta(y)].$$

Equivalently we have:

$$\sup_{x \in K} \Psi(x,y^0)= \max_{x \in K} \Psi(x,y^0)= \alpha^{\bullet}(y^0)-\beta(y^0) = \Psi(x^0,y^0), \text{and}$$

$$\inf_{y \in L} \Psi(x^0,y)= \min_{y \in L} \Psi(x^0,y) = \beta^{\bullet}(x^0)-\alpha(x^0) = \Psi(x^0,y^0).$$

(The above follows trivally from and the definition of Ψ)

From (2.4.1) we have,

$$\mu(x) \le \beta^{\bullet}(x)-\alpha(x) \le \alpha^{\bullet}(y)-\beta(y) \le \nu(y), \ \forall \ (x,y)\in M \times N$$

which implies for any $(x,y)\in M \times N$,

$$\beta^{\bullet}(x)-\alpha(x) \le \alpha^{\bullet}(y^{0})-\beta(y^{0}), \text{and}$$

$$\beta^{\bullet}(x^{0})-\alpha(x^{0}) \le \alpha^{\bullet}(y)-\beta(y).$$

Now by (4)

$$\beta^{\bullet}(x^{0})-\alpha(x^{0}) = \alpha^{\bullet}(y^{0})-\beta(y^{0}), \quad \text{this}$$

and the above inequalities leads to:

$$\max_{x \in M} [\beta^{\bullet}(x)-\alpha(x)]= \beta^{\bullet}(x^{0})-\alpha(x^{0})= \alpha^{\bullet}(y^{0})-\beta(y^{0}) = \min_{y\in N} [\alpha^{\bullet}(y)-\beta(y)],$$

$$\Rightarrow \max_{x \in M} [\beta^{\bullet}(x)-\alpha(x)] = \min_{y \in N} [\alpha^{\bullet}(y)-\beta(y)].$$

And this proves (iii) $\Rightarrow$ (ii).

(ii) $\Rightarrow$ (i). Suppose (ii) holds.

Let $(x^{0},y^{0}) \in M \times N$ be an optimal solution of the pair of problems (P) and (D). Then statement (ii) implies,

$$\beta^{\bullet}(x^{0})-\alpha(x^{0}) = \alpha^{\bullet}(y^{0})-\beta(y^{0}). \tag{5}$$

But, $$\beta^{\bullet}(x^{0}) = \inf_{y \in L} [\varphi(x^{0},y)-\beta(y)], \text{and}$$

$$\alpha^{\bullet}(y^{0}) = \sup_{x \in K} [\varphi(x,y^{0})-\alpha(x)]$$

implies that for all $(x,y) \in K \times L$,

$$\left.\begin{aligned} \beta^{*}(x^{0}) &\le \varphi(x^{0},y)-\beta(y), \text{ and} \\ \alpha^{*}(y^{0}) &\ge \varphi(x,y^{0})-\alpha(x). \end{aligned}\right\} \tag{6}$$

From (5) and (6) $\forall\ (x,y) \in K \times L$ it follows that,

$$\beta^{*}(x^{0})-\alpha(x^{0}) \le \varphi(x^{0},y)-\beta(y)-\alpha(x^{0}) = \Psi(x^{0},y), \text{ and}$$

$$\alpha^{*}(y^{0})-\beta(y^{0}) \ge \varphi(x,y^{0})-\alpha(x)-\beta(y^{0}) = \Psi(x,y^{0}).$$

Thus we have for all $(x,y) \in K \times L$,

$$\Psi(x,y^{0}) \le \alpha^{*}(y^{0})-\beta(y^{0}) = \beta^{*}(x^{0})-\alpha(x^{0}) \le \Psi(x^{0},y) \tag{7}$$

Furthermore,

$$\alpha^{*}(y^{0})-\beta(y^{0}) = \beta^{*}(x^{0})-\alpha(x^{0}), \text{ and}$$

$$\beta^{*}(x^{0})-\alpha(x^{0}) \le \varphi(x^{0},y^{0})-\beta(y^{0})-\alpha(x^{0}) \le \alpha^{*}(y^{0}) - \beta(y^{0})$$

implies that,

$$\Psi(x^{0},y^{0})= \varphi(x^{0},y^{0})-\beta(y^{0})-\alpha(x^{0})= \beta^{*}(x^{0})-\alpha(x^{0})= \alpha^{*}(y^{0})-\beta(y^{0}).$$

This together with (7) gives us

$$\Psi(x,y^{0}) \le \Psi(x^{0},y^{0}) \le \Psi(x,y^{0}),\ \forall\ (x,y) \in K \times L.$$

This proves (ii) $\Rightarrow$ (i), and the theorem is proved.

Note that : In Theorem 2.4.1 the saddle-point coordinates (x^{0},y^{0}) of statement (i) are optimal solutions of (P) and (D) and satisfy statement (ii) as well as (iii) .

DEF.2.4.1: Points (x^{0},y^{0}) which satisfy statement (iii) of Theorem 2.4.1 are called *conjugate points.*

Remark: From Theorem 2.4.1 follows the following conclusion.

Consider the sets,

$$L_\alpha = \{(x,y) \in K \times N \ / \ \varphi(x,y) = \alpha(x) + \alpha^\bullet(y)\}, \text{and}$$

$$L_\beta = \{(x,y) \in M \times L \ / \ \varphi(x,y) = \beta(y) + \beta^\bullet(x)\}.$$

This sets contains community of conjugate points corresponding to the functions α and β.

If the statement (iii) of Theorem 2.4.1 is true, the pair $(x^0, y^0) \in M \times N$ belongs to L_α as well as L_β. Consequently we have $L_\alpha \cap L_\beta \neq \emptyset$.

On the other hand if $L_\alpha \cap L_\beta \neq \emptyset$ and say $(x^0, y^0) \in L_\alpha \cap L_\beta$, then the pair (x^0, y^0) belongs to $M \times N$ and satisfies statement (iii) of Theorem 2.4.1. Thus $L_\alpha \cap L_\beta \neq \emptyset$ iff statement (iii), and hence, either statement (i) or (ii) of Teorem 2.4.1 is true.

Let us now consider the special case m = n. Let the sets K and L be subsets of one and the same space $\mathbb{R}^n$, and α, β be given functions defined on K and L respectively. Further we demand that the function φ defined on $\mathbb{R}^n \times \mathbb{R}^n$ is symmetric; i.e.

$$\varphi(x,y) = \varphi(y,x) \quad \forall \ (x,y) \in \mathbb{R}^n \times \mathbb{R}^n .$$

We assume that the functions α and β posses the following properties.

If

$$\beta^{\bullet}(y) = \inf_{x \in L} [\varphi(x,y)-\beta(x)] ,$$

is aconjugate function of β defined on $M \subseteq K$,and if

$$\alpha^{\bullet}(y) = \sup_{x \in K} [\varphi(x,y)-\alpha(x)]$$

is a conjugate function of α defined on $N \subseteq L$,then

we may be able to reproduce the functions α and β in the form

$$\beta(x) = \inf_{y \in M} [\varphi(x,y)-\beta^{\bullet}(y)] \quad \forall\, x \in L \text{ , and}$$

$$\alpha(x) = \sup_{y \in N} [\varphi(x,y)-\alpha^{\bullet}(y)] \quad \forall\, x \in K .$$

Let us denote the transition to a conjugate function by a star and we put:

$$[\alpha(x)]^{\bullet} = \alpha^{\bullet}(y) \text{ , } [\beta(x)]^{\bullet} = \beta^{\bullet}(y) \text{ , } [\varphi(x,y)]^{\bullet}=\varphi(y,x).$$

Let

$$\Psi(x,y) = \varphi(x,y)-\alpha(x)-\beta^{\bullet}(y)$$

be a Lagrangian form defined on $K \times M$.

We define the conjugate Lagrangian-form by,

$$\begin{aligned}\Psi^{\bullet}(x,y) &= [\varphi(x,y)]^{\bullet} - [\alpha(x)]^{\bullet} -[\beta^{\bullet}(y)]^{\bullet} \\ &= \varphi(y,x) - \alpha^{\bullet}(y) - \beta^{\bullet\bullet}(x) \\ &= \varphi(x,y) - \alpha^{\bullet}(y) - \beta(x) .\end{aligned}$$

Clearly the conjugate function $\Psi^{\bullet}$ is defined on $L \times N$.

Moreover, $\Psi^{\bullet\bullet}(x,y) = \Psi(x,y)$. We now consider:

$$\mu(x) = \inf_{y \in M} \Psi(x,y) = \beta(x) - \alpha(x), \quad \text{defined on } K \cap L \text{ ,and}$$

$$\nu(y) = \sup_{x \in K} \Psi(x,y) = \alpha^{\bullet}(y) - \beta^{\bullet}(y), \quad \text{defined on } M \cap N .$$

From ,

$$\mu(x) = \inf_{y \in M} \Psi(x,y) \le \Psi(x,y) \quad \forall\ (x,y) \in K \cap L \times M \text{ ,and}$$

$$\nu(y) = \sup_{x \in K} \Psi(x,y) \ge \Psi(x,y)\ ,\ \forall\ (x,y) \in K \times M \cap N$$

it follows that:

$$\mu(x) = \beta(x)-\alpha(x) \le \nu(y) = \alpha^{\bullet}(y)-\beta^{\bullet}(y)$$

for all $(x,y) \in K \cap L \times M \cap N$.

If the sets $K \cap L$ and $M \cap N$ are non empty, the formulation of the of the following dual problems make sense.

$$\text{Find } W_{\Psi} = \sup_{x\in K\cap L} [\beta(x)-\alpha(x)] \qquad (P)$$

$$\text{Find } W_{\Psi}^{\bullet} = \inf_{y\in M\cap N} [\alpha^{\bullet}(y)-\beta^{\bullet}(y)] \qquad (D)$$

Similarly for the conjugate Lagrangian form

$$\Psi^{\bullet}(x,y) = \varphi(x,y)-\alpha^{\bullet}(y)-\beta(x)$$

we get:

$$\mu^{\bullet}(x) = \sup_{y \in N} \Psi^{\bullet}(x,y) = \alpha(x)-\beta(y)$$

defined on $K \cap L$, and

$$\nu^{\bullet}(y) = \inf_{y \in L} \Psi^{\bullet}(x,y) = \beta^{\bullet}(y)-\alpha^{\bullet}(y)$$

defined on $M \cap N$. Furthermore, in this case

$$\nu^{\bullet}(y) = \beta^{\bullet}(y)-\alpha^{\bullet}(y) \le \alpha(x)-\beta(x) = \mu^{\bullet}(x) \ ,$$

$\forall\ (x,y) \in K \cap L \times M \cap N$.

And in case $K \cap L$ and $M \cap N$ are non empty, we get the following dual pair of problems.

$$\text{Find } \overline{W}_{\Psi^\bullet} = \sup_{y \in M \cap N} [\beta^\bullet(y) - \alpha^\bullet(y)] \qquad (P^\bullet)$$

$$\text{Find } W_{\Psi^\bullet} = \inf_{x \in K \cap L} [\alpha(x) - \beta(x)] \,. \qquad (D^\bullet)$$

Clearly, we have $W_\Psi^\bullet = -\overline{W}_{\Psi^\bullet}$, $W_\Psi = -W_{\Psi^\bullet}$. Furthermore,

$$[\mu(x)]^\bullet = [\beta(x)]^\bullet - [\alpha(x)]^\bullet = \beta^\bullet(y) - \alpha^\bullet(y) = \nu^\bullet(y),$$

$$[\nu^\bullet(y)]^\bullet = [\alpha^\bullet(y)]^\bullet - [\beta^\bullet(y)]^\bullet = \alpha(x) - \beta(x) = \mu^\bullet(x) \,.$$

Now similar to Theorem 2.4.1 , we have the following theorem.

Theorem 2.4.2: The following three statements are equivalent.

(i). There exist a pair $(x^0,y^0) \in K \cap L \times M \cap N$ such that for any $(x,y) \in K \times M$ we have

$$\Psi(x,y^0) \le \Psi(x^0,y^0) \le \Psi(x^0,y),$$

and for any $(x,y) \in L \times N$, we have,

$$\Psi^\bullet(x^0,y) \le \Psi^\bullet(x^0,y^0) \le \Psi(x,y^0).$$

(ii). It is true that $K \cap L$ and $M \cap N$ are non-empty and

$$\max_{x \in K \cap L} [\beta(x) - \alpha(x)] = \min_{y \in M \cap N} [\alpha^\bullet(y) - \beta^\bullet(y)] \,,$$

$$\min_{x \in K \cap L} [\alpha(x) - \beta(x)] = \max_{y \in M \cap N} [\beta^\bullet(y) - \alpha^\bullet(y)].$$

(iii). There exist a pair $(x^0,y^0) \in K \cap L \times M \cap N$ such that,

$$\varphi(x^0,y^0) = \alpha(x^0) + \alpha^\bullet(y^0),$$

$$\varphi(x^0,y^0) = \beta(x^0) + \beta^\bullet(y^0).$$

Proof: The proof of this theorem is just like the proof of Theorem 2.4.1 .

The development of Theorem 2.4.1 is based on the assumptions that the functions α and β have the properties $\alpha^{\bullet\bullet} = \alpha$ and $\beta^{\bullet\bullet} = \beta$. Now we find the conditions under which this properties are true.

Let,

$$L_{\beta} = \{(x,y) \in L \times M \,/\, \varphi(x,y) = \beta^{\bullet}(y) + \beta(x)\}.$$

We consider the set for $y \in M$ of the form ,

$$L_{\beta^{\bullet}(y)} = \{x \in L \,/\, \varphi(x,y) = \beta^{\bullet}(y) + \beta(x)\},$$

and the set for $x \in L$ of the form

$$L_{\beta(x)} = \{y \in M \,/\, \varphi(x,y) = \beta^{\bullet}(y) + \beta(x)\}.$$

we assume that $L_{\beta^{\bullet}(y)}$ is not empty for any $y \in M$.

Then for any $y \in M$ there exist at least one $x_y \in L$ such that,

$$\varphi(x_y,y) = \beta^{\bullet}(y) + \beta(x_y).$$

Consequently, we get for any $y \in M$

$$\beta^{\bullet}(y) = \inf_{x \in L} [\varphi(x,y)-\beta(x)] = \varphi(x_y,y)-\beta(x_y)$$

$$= \min_{x \in L} [\varphi(x,y)-\beta(x)].$$

Thus, we can replace infimum by minimum when we define the conjugate function $\beta^{\bullet}$ if $L_{\beta^{\bullet}(y)} \neq \emptyset$ for any $y \in M$.

Theorem 2.4.3 : Let $Ł_{\beta^{\bullet}(y)}$ be a non-mpty set for any $y \in M$. Then $Ł_{\beta(x)} \neq \emptyset$ for any $x \in L$ iff for all $x \in L$ we have,

$$\beta(x) = \min_{y \in M} [\varphi(x,y)-\beta^{\bullet}(y)].$$

Proof: Suppose for all $x \in L$, we have

$$\beta(x) = \min_{y \in M} [\varphi(x,y)-\beta^{\bullet}(y)]$$

then for all $x \in L$, there exist at least one $y_x \in M$ such that

$$\beta(x) = \min_{y \in M} [\varphi(x,y)-\beta^{\bullet}(x)] = \varphi(x,y_x)-\beta^{\bullet}(y_x),$$

thus $y_x \in Ł_{\beta(x)}$ for all $x \in L \Rightarrow Ł_{\beta(x)} \neq \emptyset$.

Conversly, suppose $Ł_{\beta(x)} \neq \emptyset$ for all $x \in L$; this implies for any $x \in L$, there exist $y_x \in M$ such that

$$\varphi(x,y_x) = \beta^{\bullet}(y_x) + \beta(x).$$

Furthermore, using the definition of $\beta^{\bullet}$ for any $(x,y) \in L \times M$ we get:

$$\beta^{\bullet}(y) \leq \varphi(x,y)-\beta(x).$$

Consequently,

$$\varphi(x,y)-\beta^{\bullet}(y) \geq \beta(x) = \varphi(x,y_x)-\beta^{\bullet}(y_x).$$

Thus,

$$\min_{y \in M} [\varphi(x,y)-\beta^{\bullet}(y)] = \varphi(x,y_x)-\beta^{\bullet}(y_x) = \beta(x).$$

This proves the theorem .

Similarly for the function α , let

$$Ł_{\alpha} = \{(x,y) \in K \times N \ / \ \varphi(x,y) = \alpha^{*}(y) + \alpha(x)\},$$

and consider the sets

$$Ł_{\alpha^{*}(y)} = \{x \in K \ / \ \varphi(x,y) = \alpha(x) + \alpha^{*}(y)\}$$

$$Ł_{\alpha(x)} = \{y \in N \ / \ \varphi(x,y) = \alpha(x) + \alpha^{*}(y)\}$$

If we assume that for any $y \in N$ the set $Ł_{\alpha^{*}(y)} \neq \emptyset$, then there exist for each $y \in N$ at least one $x_y \in K$ such that,

$$\varphi(x_y,y) = \alpha^{*}(y) + \alpha(x_y).$$

Consequently, for any $y \in N$ we get,

$$\alpha^{*}(y) = \sup_{x \in K} [\varphi(x,y)-\alpha(x)] = \max_{x \in K} [\varphi(x,y)-\alpha(x)] = \varphi(x_y,y)-\alpha(x_y).$$

Thus we can replace supremum by maximum when we define the conjugate function α^{*} if $Ł_{\alpha^{*}(y)} \neq \emptyset$, $\forall\, y \in N$.

And we get a theorem which is similar to Theorem 2.4.3 .

Theorem 2.4.4 : Let $Ł_{\alpha^{*}(y)}$ be a non-empty set for any $y \in N$. Then $Ł_{\alpha(x)} \neq \emptyset$ for any $x \in K$ iff for all $x \in K$ we have,

$$\alpha(x) = \max_{y \in N} [\varphi(x,y)-\alpha^{*}(y)] .$$

Proof: The proof of this theorem is just like the proof of Theorem 2.4.3 .

Illustration.

This is an application of Theorem 2.4.2.

Consider:

$$\varphi(x,y) = x^2y^2,$$

$$\alpha(x) = \;^{2}, x \in K = \{x \in \mathbb{R} \,/\, x > 0\}$$

$$\beta(x) = x \;,\; x \in L = \{x \in \mathbb{R} \,/\, x \geq 1\}.$$

The conjugate function β^{*} is:

$$\beta^{*}(y) = \inf_{x \in L} [\varphi(x,y) - \beta(x)]$$

$$\beta^{*}(y) = \inf_{L} [x^2y^2 - x] = \begin{cases} {}^{-1}/_{4y^2} & \text{if } 0 < y \leq {}^{1}/_{\sqrt{2}} \\ y^2 - 1 & \text{if } {}^{1}/_{\sqrt{2}} \leq y \end{cases}$$

and β^{*} is defined on $M = K$. Furthermore we have

$$L_{\beta^{*}(y)} = \{x \in L \,/\, \varphi(x,y) = \beta^{*}(y) + \beta(x)\},$$

$$= \{x \geq 1 \,/\, x^2y^2 = \beta^{*}(y) + x\}.$$

Case(i): For any $y \leq {}^{1}/\sqrt{2}$, we have

$$L_{\beta^{*}(y)} = \{x \geq 1 \,/\, x^2y^2 = {}^{-1}/_{4y^2} + x\},$$

since at least $x = {}^{1}/_{2y^2} \geq 1$ is contained in $L_{\beta^{*}(y)}$ we have

$L_{\beta^{*}(y)} \neq \emptyset$.

Case(ii): If $y \geq {}^{1}/\sqrt{2}$, then

$$L_{\beta^{*}(y)} = \{x \geq 1 \,/\, x^2y^2 = y^2 - 1 + x\},$$

and the point $x = 1$ is in $L_{\beta^{*}(y)}$ for all $y \geq {}^{1}/\sqrt{2}$.

Hence we have $L_{\beta^\bullet(y)} \neq \emptyset$.

Consequently from (i) and (ii) we get

$$L_{\beta^\bullet(y)} \neq \emptyset \text{ for any } y \in M .$$

Next we consider the set

$$L_{\beta(x)} = \{y > 0 \;/\; x^2y^2 = \beta^\bullet(y) + x \}.$$

Case(i): If $x = 1$, then

$$L_{\beta(1)} = \{y > 0 \;/\; y^2 = \beta^\bullet(y) + 1 \} \neq \emptyset \text{ ,this is}$$

because all elements y with $y \geq 1/\sqrt{2}$ belongs to $L_{\beta(1)}$.

Case(ii): If $x > 1$,then

$$L_{\beta(x)} = \{y > 0 \;/\; x^2y^2 = \beta^\bullet(y) + x \}.$$

Any y with $y^2 = 1/2x$ has the property $y \leq 1/\sqrt{2}$,

consequently ,y is contained in $L_{\beta(x)}$.

Thus from (i) and (ii) $L_{\beta(x)} \neq \emptyset$ for any $x \in L$.

Hence by Theorem 2.4.3 $\beta^{\bullet\bullet} = \beta$. (1)

Next we consider,

$$\alpha^\bullet(y) = \sup_{x \in K} [\varphi(x,y) - \alpha(x)] = \sup_{x>0} [x^2y^2 - x].$$

Unless $y = 1$ the supremum does not exist (i.e. the supremum is not finite). Hence,

$$\alpha^\bullet(y) = \sup_{x>0} [x^2y^2 - x] = 0 \text{ ,and } \alpha^\bullet \text{ is defined on } N = \{1\}.$$

Then, $\alpha^{**}(x) = [\alpha^{*}(y)]^{*} = \sup [x^2y^2] = x^2 = \alpha(x), \quad \forall\, x \in K$

$$\Rightarrow \alpha^{**} = \alpha \,. \tag{2}$$

(1) and (2) implies that some of the assumption for Theorem 2.4.2 are fuilfilled .

Now the lagrangian-form Ψ and its conjugate Ψ^{*} are given by:

$$\Psi(x,y) = \varphi(x,y)-\alpha(x)-\beta^{*}(y)$$

$$= x^2y^2 - x^2 - \beta^{*}(y)$$

defined on $K \times M = K \times K$;and

$$\Psi^{*}(x,y) = \varphi(x,y)-\alpha^{*}(y)-\beta(x)$$

$$=x^2y^2 - x$$

defined on $L \times N = L \times \{1\}$.

We now have,

$$\mu(x) = \inf_{y>0} \Psi(x,y) = \beta(x)-\alpha(x)$$

$$\Rightarrow \mu(x) = x - x^2 \,,\ x \in K \cap L = L \,.$$

And ,

$$\nu(y) = \sup_{x>0} \Psi(x,y) = \alpha^{*}(y)-\beta^{*}(y) \,, y \in M \cap N = K \cap \{1\}= \{1\}$$

$$\Rightarrow \nu(y) = 1 - y^2 = 0.$$

Also:

$$\mu^{*}(x)= \sup_{y=1} \Psi^{*}(x,y) = x^2 - x \,,\ x \in K \cap L = L \text{ , and}$$

$$\nu^{*}(y)= \inf_{x\geq 1} \Psi^{*}(x,y) = \beta^{*}(y)-\alpha^{*}(y)= \beta^{*}(y) \,, y \in K\cap \{1\} = \{1\}$$

$$\Rightarrow \nu^{*}(y) = y^2 - 1 = 0 \,.$$

Then the dual problems take the form:

$$\sup_{x\geq 1} [x - x^2] = 0 = W_\Psi \ , \ x^0 = 1 \tag{P}$$

$$\inf_{y=1} [1 - y^2] = 0 = W_\Psi^{\bullet} \ , \ y^0 = 1 \tag{D}$$

$$\sup_{y=1} [y^2 - 1] = 0 = \overline{W}_{\Psi^\bullet} \ , \ y^0 = 1 \tag{P$^\bullet$}$$

$$\inf_{x\geq 1} [x^2 - x] = 0 = W_{\Psi^\bullet} \ , \ x^0 = 1 \tag{D$^\bullet$}$$

The pair $(x^0, y^0) = (1,1)$ satisfies statements (i), (ii) and (iii) of Theorem 2.4.2 .

Remark: The statement (ii) of Theorem 2.4.1 and Theorem 2.4.2 is a duality theorem together with the assertion about the existence of optimal solutions by which we mean realizability.

2.5 Necessary and Sufficient Conditions for the Existence of Saddle Points

Theorem 2.4.1 and Theorem 2.4.2 show us that saddle point theorem, duality theorem ,realizability of problems (P) and (D) and conjugate points are very closely related to one another. Besides,this relationships,we are intersted in theorems of existence of saddle points. So in this section we deal with *Kuhn-Tucker* conditions which ensure the existence of saddle-points and minimax theorem corresponding to Theorems 2.3.1 , Theorem 2.4.1 ,and Theorem 2.4.2 .

Theorem 2.5.1: Let $\Psi(x,y)$ be a function defined on $K \times L = \mathbb{R}^{n+} \times \mathbb{R}^{m+}$ with the property that Ψ is concave in x and convex in y, and Ψ is Frechet differentiable both in x and y. Then $(x^0,y^0) \in \mathbb{R}^{n+} \times \mathbb{R}^{m+}$ is a saddle point of Ψ with respect to $K \times L$ if and only if

$$\left.\begin{array}{ll} (1) & d_y\Psi(x^0,y^0) = \mathrm{grad}_y\Psi(x^0,y^0) \geq \mathbf{0}, \\ & y^0 \in \mathbb{R}^{m+}, \quad \langle d_y\Psi(x^0,y^0), y^0\rangle = 0 \\ (2) & d_x\Psi(x^0,y^0) = \mathrm{grad}_x\Psi(x^0,y^0) \leq \mathbf{0}, \\ & x^0 \in \mathbb{R}^{n+}, \quad \langle d_x\Psi(x^0,y^0), x^0\rangle = 0 \end{array}\right\} \quad (2.5.1)$$

Proof: First we will show that (2.5.1) implies the saddle point inequality. According to the assumption, $-\Psi(x,y)$ is convex with respect to x.

Now by Theorem 1.3.2 for any $x \in \mathbb{R}^{n+}$

$$\Psi(x^0,y^0)-\Psi(x,y^0) \geq -\langle d_x\Psi(x^0,y^0), x-x^0\rangle.$$

$$\Rightarrow \Psi(x^0,y^0)-\Psi(x,y^0) \geq \langle d_x\Psi(x^0,y^0), x\rangle + \langle d_x\Psi(x^0,y^0), x^0\rangle.$$

But, then from (2.5.1) we have,

$$d_x\Psi(x^0,y^0) \leq 0, \text{and } \langle d_x\Psi(x^0,y^0), x^0\rangle = 0.$$

Thus we get,

$$\Psi(x^0,y^0) - \Psi(x,y^0) \geq -\langle d_x\Psi(x^0,y^0), x\rangle + 0$$

$$\Rightarrow \Psi(x,y^0) - \Psi(x^0,y^0) \leq -\langle d_x\Psi(x^0,y^0), x\rangle \leq 0$$

$$\Rightarrow \Psi(x,y^0) \leq \Psi(x^0,y^0), \quad \forall\, x \in \mathbb{R}^{n+}. \qquad \text{(i)}$$

Similarly because Ψ is convex with respect to y

$$\Psi(x^0,y) - \Psi(x^0,y^0) \geq \langle d_y\Psi(x^0,y^0), y-y^0\rangle = \langle d_y\Psi(x^0,y^0), y\rangle \geq 0$$

$$\Rightarrow \Psi(x^0,y) \geq \Psi(x^0,y^0) \ , \ \forall \ y \in \mathbb{R}^{m+} . \qquad \text{(ii)}$$

From (i) and (ii) it follows that,

$$\Psi(x,y^0) \leq \Psi(x^0,y^0) \leq \Psi(x^0,y), \quad \forall \ (x,y) \in \mathbb{R}^{n+} \times \mathbb{R}^{m+}.$$

Thus (x^0,y^0) is a saddle point of Ψ with respect to $\mathbb{R}^{n+} \times \mathbb{R}^{m+}$.

Conversly, suppose $(x^0,y^0) \in \mathbb{R}^{n+} \times \mathbb{R}^{m+}$ be a saddle point of Ψ.

If we put

$$\phi(x) = \Psi(x,y^0) \ , \text{ for } x \in \mathbb{R}^{n+} \ ,$$

then clearly

$$\phi(x) \leq \phi(x^0) \ , \ \forall \ x \in \mathbb{R}^{n+} .$$

Thus the function ϕ possesses at least a local maximum at x^0.

Then by Theorem 1.3.1 we have,

$$(d\phi)_{x^0} = d_x\Psi(x^0,y^0) \leq \mathbf{0} \ , \text{ and}$$

$$\langle (d\phi)_{x^0}, x^0\rangle = \langle d_x\Psi(x^0,y^0), x^0\rangle = 0.$$

On the other hand if we put

$$\phi(y) = -\Psi(x^0,y)$$

the function ϕ posseses at y^0 a local maximum with respect to $\mathbb{R}^{m+}$. Again by Theorem 1.3.1 we have,

$$(d\phi)_{y^0} = -d_y\Psi(x^0,y^0) \le \mathbf{0}, \text{ and}$$

$$\langle (d\phi)_{y^0}, y^0\rangle = -\langle d_y\Psi(x^0,y^0), y^0\rangle = 0 .$$

This proves the theorem.

Note: The conditions (2.5.1) are called Kuhn-Tucker conditions and Theorem 2.5.1 is the Kuhn-Tucker Theorem .

Corollary 2.5.1 If we espically consider the Lagrangian form

$$\Psi(x,y) = \varphi(x,y)-\alpha(x)-\beta(y)$$

defined on $K \times L = \mathbb{R}^{n+} \times \mathbb{R}^{m+}$, where α is convex, β concave and φ convex with respect to y and concave with respect to x ;then Ψ satisifies the assumptions of Theorem 2.5.1 ,if the functions $\varphi, \alpha,$ and β are Frechet differentiable in x and y. In this case conditions (2.5.1) has the form:

$$\left.\begin{aligned} &(1).\ d_y\varphi(x^0,y^0) \ge d\beta(y^0) \ , y^0 \in \mathbb{R}^{m+} \ , \\ &\qquad \langle d_y\varphi(x^0,y^0), y^0\rangle = \langle d\beta(y^0), y^0\rangle; \\ &(2).\ d_x\varphi(x^0,y^0) \le d\alpha(x^0),\ x^0 \in \mathbb{R}^{n+} , \\ &\qquad \langle d_x\varphi(x^0,y^0), x^0\rangle = \langle d\alpha(x^0), x^0\rangle. \end{aligned}\right\} \quad (2.5.2)$$

Proof: Similar to the proof of Theorem 2.5.1 .

Remarks :

(1). Condition (2.5.2) in the above corollary satisfied means that the point (x^0, y^0) is a saddle point of the Lagrangian-form Ψ with respect to $K \times L$. Hence if the functions φ, α, and β has the desired properties mentioned above, then condition (2.5.2) is equivalent to the statements (i), (ii), and (iii) of Theorem 2.4.1, and regarding our inquiries in section 2.4 the conditions (2.5.2) are also equivalent to $L_\alpha \cap L_\beta \neq \emptyset$.

(2) If we drop the demand that Ψ is concave in x and convex in y, the conditions (2.5.1) and (2.5.2) remain as necessary conditions for the existence of saddle points, but they are not sufficient any more. This result follows immediately from the proof of Theorem 2.5.1 where we only used convexity to prove sufficency.

(3) We can also generalize the theorem with respect to the domain of definitions $\mathbb{R}^{n+} \times \mathbb{R}^{m+}$ of the function Ψ. To do so consider,

$$x = (u, v) \in \mathbb{R}^{r+} \times \mathbb{R}^{s}.$$

Let $\Psi(u, v; y)$ be a function defined on $(\mathbb{R}^{r+} \times \mathbb{R}^{s}) \times \mathbb{R}^{m+}$ with the following properties.

(a) Either Ψ is concave in (u, v), convex in y and Frechet-differentiable in all variables.

(b) Or Ψ possesses all partial derivatives of first order which are continuous. Then it follows that Ψ is F-differentiable with respect to all variables and that the derivatives can be represented by the gradient of Ψ.

Now, put $v = w-z$, with $w \geq 0$, $z \geq 0$ and conisder:

$$\Psi(u, w-z; y) = \phi(u, w, z, y)$$

defined on $K \times L = (\mathbb{R}^{r+} \times \mathbb{R}^{s+} \times \mathbb{R}^{s+}) \times \mathbb{R}^{m+}$, then condition (2.5.1) for ϕ has the form :

(1) $\text{grad}_y \phi = \text{grad}_y \Psi \geq \mathbf{0}$, $y^0 \in \mathbb{R}^{m+}$,

$$\langle \text{grad}_y \phi, y^0 \rangle = \langle \text{grad}_y \Psi, y^0 \rangle = 0$$

(2) $\text{grad}_x \phi = (\text{grad}_u \phi,\ \text{grad}_w \phi,\ \text{grad}_z \phi) \leq 0$, $(u^0, w^0, z^0) \geq \mathbf{0}$,

$$\langle \text{grad}_x \phi, x^0 \rangle = \langle \text{grad}_u \phi, y^0 \rangle + \langle \text{grad}_w \phi, w^0 \rangle + \langle \text{grad}_z \phi, z^0 \rangle = 0,$$

where $x^0 = (u^0, w^0, z^0)$

Furthermore,

$$\text{grad}_u \phi = \text{grad}_u \Psi,\quad \text{grad}_w \phi = \text{grad}_v \Psi,\quad \text{grad}_z \phi = -\text{grad}_v \Psi.$$

Then we have :

$$\left.\begin{array}{l} \text{grad}_y \Psi \geq \mathbf{0},\ y^0 \in \mathbb{R}^{m+},\ \langle \text{grad}_y \Psi, y^0 \rangle = 0 \\ (\text{grad}_u \Psi, \text{grad}_v \Psi, -\text{grad}_v \Psi) \leq \mathbf{0}, \\ \langle \text{grad}_u \Psi, u^0 \rangle + \langle \text{grad}_v \Psi, w^0 - z^0 \rangle = 0, \end{array}\right\} \qquad (2.5.3)$$

Or

$$\text{grad}_u \Psi \leq \mathbf{0},\quad \text{grad}_v \Psi = \mathbf{0},\quad \langle \text{grad}_u \Psi, v^0 \rangle = 0.$$

The condition (2.5.3) is necessary for the existence of saddle points for ϕ and therefore necessary for the existance of saddle points for Ψ . If Ψ satisfies the condition of convexity, then the conditions in (2.5.3) are also sufficient for the existence of saddle points for Ψ and ϕ .

We can also have analogous situtation and similar consideration with respect to the variable y.

BIBLOGRAPHY

1. M.Walk : Theory of Duality in Mathematical Programming; Akedemia-Verlag Berlin 1989.

2. J.Stoer ,and C.Witzgal :Convexity and Optimization in Finite Dimention I ; Springer-Verlag 1970.

3. V.G.Karmanov : Mathematical Programming ; Mir Publishers 1989.

4. E.Zeidler : Nonlinear Functional Analysis and its Applications; III.Varitation Methods and Optimization . Springer-Verlag , New York Inc. 1985.