DUALITY IN CONVEX PROGRAMMING

by

DAVID CHARLES WILLIAM MUIR
B.Sc., University of British Columbia, 1962

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE— ARTS

in the Department
of
MATHEMATICS

We accept this thesis as conforming to the
required standard

THE UNIVERSITY OF BRITISH COLUMBIA
April, 1966
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Mathematics

The University of British Columbia
Vancouver 8, Canada

Date May 12, 1966.
Problems of minimizing a convex function or maximizing a concave function over a convex set are called convex programming problems. Duality principles relate two problems, one a minimization problem, the other a maximization problem, in such a way that a solution to one implies a solution to the other and that the minimum value of one is equal to the maximum value of the other.

When the functions are linear and the constraint sets are polyhedral, the problems are called linear programming problems. Their duality is well-known. Certain duality results of linear programming can be extended to convex programming by means of the theory of conjugate convex functions introduced by Fenchel ([1], [2]).

In this thesis the theory of conjugate functions is generalized and applied to convex programming problems. In particular a duality theorem is given for a class of convex programming problems. This theorem is compared with a duality theorem for convex programming problems given by Dorn [3].
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Introduction</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Section 1:</strong> Review of Duality Theory for Conjugate Convex Functions . . .</td>
<td>3</td>
</tr>
<tr>
<td><strong>Section 2:</strong> Generalized Conjugate Sets and Functions . . . . . . . . . . .</td>
<td>11</td>
</tr>
<tr>
<td><strong>Section 3:</strong> Concave Functions and their Duals . . . . . . . . . . . . . . .</td>
<td>20</td>
</tr>
<tr>
<td><strong>Section 4:</strong> Duality for Convex and Concave Functions . . . . . . . . . . .</td>
<td>22</td>
</tr>
<tr>
<td><strong>Section 5:</strong> Lagrangians and Dual Convex Programming Problems . . . . . . .</td>
<td>28</td>
</tr>
<tr>
<td><strong>Section 6:</strong> Comparison of Duality Theorems . . . . . . . . . . . . . . . .</td>
<td>35</td>
</tr>
<tr>
<td><strong>Bibliography</strong> . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .</td>
<td>42</td>
</tr>
<tr>
<td><strong>Appendix</strong> . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .</td>
<td>43</td>
</tr>
</tbody>
</table>
ERRATA

<table>
<thead>
<tr>
<th>Page</th>
<th>Line</th>
<th>Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>12, 19</td>
<td>The assumption of relative interior is not required.</td>
</tr>
<tr>
<td>11</td>
<td>9, 12</td>
<td>Only $\geq$ has been proved. The equality requires the definition of $f^*$ and $f^{**}$.</td>
</tr>
<tr>
<td>13</td>
<td>9</td>
<td>For $\hat{C}$ read $C_A$.</td>
</tr>
<tr>
<td>17</td>
<td>6</td>
<td>&quot;then $y_0$ is on the boundary of $C_A$. Should be deleted.&quot;</td>
</tr>
</tbody>
</table>
INTRODUCTION

Problems which seek to maximize or minimize a function of several variables which are subject to certain constraints, are called programming problems. Those for which the functions are convex or concave, and the constraint sets are convex, are called convex programming problems.

Duality principles relate two programming problems, one of which, the primal, is a maximization problem and the other the dual, is a minimization problem. They are related in such a way that the existence of a solution to one insures a solution to the other. Furthermore the respective extreme values are equal. Problems in which the constraint sets are polyhedral and the functions linear are called linear programming problems. Their duality is well-known.

Primal Linear Programming Problem

Given the mXn matrix $A$ and constant vectors $c$ in $\mathbb{R}^n$ and $b$ in $\mathbb{R}^m$, maximize $(c,x)$ subject to the constraints $Ax \leq b$, $x \geq 0$.

Dual Linear Programming Problem

To minimize $(\xi,b)$ subject to the constraints $\xi A \geq c$, $\xi \geq 0$.

Associated with these problems is the Lagrangian function

$$L(x,\xi) = (c,x) + (\xi,b) - \xi Ax$$

defined for $x \geq 0$, $\xi \geq 0$. The components of the vector $\xi$ in the dual problem are the "Lagrange Multipliers" for the primal problem, hence the name Lagrangian function for $L(x,\xi)$. 
A saddle-point of the Lagrangian function $L(x, \xi)$ is a pair of vectors $\{x^*, \xi^*\}$ for which $L(x^*, \xi) \geq L(x^*, \xi^*) \geq L(x, \xi^*)$, for all $\{x, \xi\}$ in the domain of $L(x, \xi)$.

The connection between the two linear programming problems and their associated Lagrangian function is given by the following "duality" theorem.

Theorem:
(1) if there exists a vector $x^*$ which maximizes $(c, x)$ in the primal problem then there also exists a vector $\xi^*$ which minimizes $(\xi, b)$ in the dual problem, and vice versa.
(2) if there exists a saddle-point $\{x^*, \xi^*\}$ of the Lagrangian $L(x, \xi) = (c, x) + (\xi, b) - \xi^* Ax$ then $x^*$ maximizes $(c, x)$ in the primal problem and $\xi^*$ minimizes $(\xi, b)$ in the dual problem, and vice versa. Furthermore $(c, x^*) = L(x^*, \xi^*) = (\xi^*, b)$.

Some results for linear programming problems can be extended to convex programming problems by means of the theory of conjugate functions introduced by Fenchel ([1], [2]).

The purpose of this thesis is to obtain duality results for a larger class of convex programming problems. This is achieved by generalizing the concept of conjugate convex functions and applying the resulting theory. In particular a duality theorem for a class of convex programming problems is given, which includes the well-known duality theorem for linear programming stated above. This theorem can also be specialized to a duality theorem for convex programming due to Dorn [3].
Notation: In this thesis the following notation will be used.

- $A = (a_{ij})$ is an $m \times n$ matrix over the real numbers,
- $x$ is a column vector in $E^n$,
- $y$ is a column vector in $E^m$,
- $h$ is a row vector in $E^n$,
- $\xi$ is a row vector in $E^m$.

For any real numbers $t$ and $c$, $[t, x]$ will denote a column vector in $E^{n+1}$ and $[c, \xi]$ will denote a row vector in $E^{m+1}$.

The usual inner product in $E^n$ or $E^m$ will be denoted by $(\ , \ )$. It will be defined only between a row vector and a column vector. Thus

$$\langle h, x \rangle = \sum_{j=1}^{n} h_j x_j$$

$$\langle \xi, y \rangle = \sum_{i=1}^{m} \xi_i y_i.$$ 

Also $\xi \geq 0$ means $\xi_i \geq 0$ for all $i; \ i = 1, \ldots, m$.

Sets will be denoted by capital letters.

Section 1: Review of Duality Theory for Conjugate Convex Functions.

The theory of conjugate functions introduced by Fenchel [1], can be used to extend certain results of the duality theory of linear programming to convex programming problems.

This section reviews the main results of the theory of conjugate convex functions, most of which can be found in Fenchel ([1],[2]) or in the book by Karlin [4] (p.218-229).
In the following sections these results will be generalized and used to prove a duality theorem for a class of convex programming problems.

**Definition 1.1: Convex Set**

A set \( C \) of vectors in \( \mathbb{E}^n \) is called a convex set if for any two vectors \( x_1 \) and \( x_2 \) in \( C \), and for any real number \( \lambda \), \( 0 \leq \lambda \leq 1 \), the vector \( \lambda x_1 + (1-\lambda)x_2 \) is in the set \( C \).

**Definition 1.2: Convex Function**

A function \( f \) defined on a convex set \( C \) is called convex if for any two points \( x_1 \) and \( x_2 \) in \( C \) and for all \( \lambda \), \( 0 \leq \lambda \leq 1 \),

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

**Definition 1.3: Concave Function**

A function \( g \) defined on a convex set \( D \) is called concave if the function \( -g \) is a convex function on the convex set \( D \).

**Lemma 1.1:** Let \( f \) be a convex function defined on a convex set \( C \). Then

1. \( f \) is continuous in the relative interior of \( C \),
2. for a boundary point \( x_0 \) of \( C \),

\[
\lim_{x \to x_0} \inf f(x) \leq f(x_0).
\]
**Proof:** The result is well-known, it can be found in Courant [5], p.326.

**Definition 1.4:** A convex function $f$ with domain the convex set $C$, is said to be closed if $x_0$ in the boundary of $C$ and $\liminf_{x \to x_0} f(x) < \infty$ implies $x_0$ is in $C$ and $f(x_0) = \liminf_{x \to x_0} f(x)$.

**Definition 1.5:** If $f$ is a convex function with domain $C$, then the set $[f,C]$ is defined by

$$[f,C] = \{ [t,x] \in \mathbb{R}^{n+1} : x \in C, t > f(x) \}.$$

**Lemma 1.2:** If $f$ is a closed convex function then $[f,C]$ is a closed convex set.

**Proof:** The result follows immediately from the definition.

**Lemma 1.3:** If $f$ is a closed convex function defined on the convex set $C$, then for every point $[s_0, x_0] \notin [f,C]$ there exists a non-vertical hyperplane that strictly separates $[s_0, x_0]$ from $[f,C]$.

**Proof:** $f$ closed implies $[f,C]$ is a closed convex set. Hence $[s_0, x_0] \notin [f,C]$ implies there exists a sphere $S$ of radius $\epsilon$ about $[s_0, x_0]$ such that $S \cap [f,C] = \emptyset$. 
Then $S$ and $[f,C]$ can be separated by a hyperplane $P$, say

$$P: ([\tau, h], [t, x]) = k$$

so that

$$\tau t + (h, x) \leq k, \quad \text{for all } [t, x] \in [f, C] \tag{1.1}$$

$$\tau s + (h, x) \geq k, \quad \text{for all } [s, x] \in S. \tag{1.2}$$

Since $[s_o, x_o]$ is in the interior of $S$,

$$\tau s_o + (h, x_o) > k. \tag{1.3}$$

Since $[t, x]$ is still in $[f, C]$ for $t$ large, $\tau < 0$.

**Case 1:** $x_o$ is an interior point of $C$. Then $\tau \neq 0$, for if $\tau = 0$, from (1.1), $(h, x_o) \leq k$ and from (1.3), $(h, x_o) > k$. This is a contradiction; therefore $\tau < 0$ and $P$ is a non-vertical hyperplane.

**Case 2:** $x_o$ is on the boundary of $C$. Suppose $\tau = 0$, then $(h, x_o) > k$ from (1.3). Since $x_o$ is on the boundary of $C$ there exists a sequence $\{x_n\}$, $x_n \rightarrow x_o$, with $x_n \in C$ and $f(x_n)$ defined. Therefore $[f(x_n), x_n] \in [f, C]$ for all $n$, and hence by (1.1) $\tau f(x_n) + (h, x_n) \leq k$ for all $n$ or since $\tau = 0$, $(h, x_n) \leq k$. Therefore $(h, x_o) = \lim_{n \to \infty} (h, x_n) \leq k$, which is a contradiction. Therefore $\tau \neq 0$ and $P$ is a non-vertical hyperplane.
Case 3: \( x_0 \) is not in the closure of \( C \). Then there exists a hyperplane \( \overline{P}_1 \) in \( E^n \), \( \overline{P}_1 : (n_1, x) = k_1 \), that strictly separates \( x_0 \) from \( C \). Thus

\[
(1.4) \quad (n_1, x) \leq k_1 \text{ for all } x \text{ in } C
\]

\[
(1.5) \quad (n_1, x_0) > k_1.
\]

This hyperplane can be extended vertically to \( E^{n+1} \) obtaining the vertical hyperplane \( P_1 \) in \( E^{n+1} \), \( P_1 : ([0, n_1], [t, x]) = k_1 \).

Next take any non-vertical hyperplane \( P_2, P_2 : ([\tau, n_2], [t, x]) = k_2 \), that keeps \([f, C]\) on one side, say

\[
(1.6) \quad \tau t + (n_2, x) \leq k_2
\]

for all \([t, x]\) in \([f, C]\). The existence of the plane is guaranteed by case 1 or 2.

Form the hyperplane \( P = (1-\alpha)P_1 + \alpha P_2 \), where \( 1 > \alpha > 0 \),

\[
P : \left( [\omega t, (1-\alpha)n_1 + \alpha n_2], [t, x] \right) = (1-\alpha)k_1 + \alpha k_2.
\]

It will be shown that for \( \alpha \) small enough \( P \) separates \([s_0, x_0]\) from \([f, C]\). For \([t, x]\) in \([f, C]\),

\[
\omega t + (1-\alpha)(n_1, x) + \alpha(n_2, x) \leq (1-\alpha)k_1 + \alpha k_2.
\]

If \([s_0, x_0]\) is on the opposite side of \( P_2 \) from \([f, C]\), there is nothing to prove, any \( \alpha, 0 < \alpha < 1 \), would be acceptable. Hence suppose \([s_0, x_0]\) is on the same side of \( P_2 \) as \([f, C]\); then from (1.6), \( \tau s_0 + (n_2, x_0) \leq k_2 \), say

\[
\tau s_0 + (n_2, x_0) = k_2 - \varepsilon_2, \quad \varepsilon_2 > 0.
\]

From (1.5), \( (n_1, x_0) > k_1 \), say
\((h_1, x_0) = k_1 + \epsilon_1, \epsilon_1 > 0\). Then

\[a\sigma s_o + a(h_2, x_0) + (1-a)(h_1, x_0) = a(k_2 - \epsilon_2) + (1-a)(k_1 + \epsilon_1) .\]

Choose \(\alpha > 0\) small enough so that \((1-a)\epsilon_1 > \alpha \epsilon_2\), then

\[a\sigma s_o + a(h_2, x_0) + (1-a)(h_1, x_0) > \alpha k_2 + (1-a)k_1 .\]

Therefore \(P\) strictly separates \([s_o, x_0]\) from \([f, u]\), and \(P\) is non-vertical because \(\alpha \tau \neq 0\).

**Remark:** It can easily be verified that the statement

"There exists a non-vertical hyperplane below the set \([f, C]\) with normal \([-1, h]\)"

is equivalent to the statement

"\(\sup_{x \in C} [(h, x) - f(x)] < \infty\)",

as long as the relative interior of \(C\) is non-empty.

**Definition 1.6:** Given the set \([f, C]\), the set \(C^*\) and the function \(f^*\) are defined by

\[C^* = \{h \in \mathbb{R}^n : \sup_{x \in C} [(h, x) - f(x)] < \infty\}\]

\[f^*(h) = \sup_{x \in C} [(h, x) - f(x)], \text{ for } h \in C^* .\]

The function \(f^*\) is referred to as the conjugate function of \(f\). From the previous remarks if \(C\) contains an interior point then the set \(C^*\) is non-empty.
Lemma 1.4: \( C^* \) is a convex set and \( f^* \) is a convex function on \( C^* \).

Proof: If \( n_1 \) and \( n_2 \) are in \( C^* \) and \( 0 < \lambda < 1 \), then

\[
\sup_{x \in C} [(\lambda n_1 + (1-\lambda)n_2, x) - f(x)] = \sup_{x \in C} [\lambda (n_1, x) + (1-\lambda)(n_2, x) - \lambda f(x) - (1-\lambda)f(x)] \\
\leq \sup_{x \in C} \lambda [(n_1, x) - f(x)] + \sup_{x \in C} (1-\lambda)[(n_2, x) - f(x)] \\
= \lambda f^*(n_1) + (1-\lambda)f^*(n_2) < \infty.
\]

Hence \( \lambda n_1 + (1-\lambda)n_2 \in C^* \) and

\[
f^*[\lambda n_1 + (1-\lambda)n_2] \leq \lambda f^*(n_1) + (1-\lambda)f^*(n_2)
\]

completing the proof of the lemma.

Lemma 1.5: \( f^* \) is a closed convex function of \( C^* \).

Proof: Suppose \( n_0 \) is on the boundary of \( C^* \) and there exists a sequence \( \{n_i\} \), with \( n_i \) in \( C^* \), \( n_i \rightarrow n_0 \) and

\[ \lim \inf f^*(n_i) < \infty; \]

then

\[
f^*(n_1) = \sup_{x \in C} [(n_1, x) - f(x)] \geq (n_1, x) - f(x) \text{ for all } x \in C.
\]

Now

\[ \infty > \lim \inf f^*(n_1) \geq \lim \inf [(n_1, x) - f(x)] = (n_0, x) - f(x) \text{ for all } x \in C. \]

Therefore

\[
\sup_{x \in C} [(n_0, x) - f(x)] \leq \lim \inf f^*(n_1) < \infty
\]
which implies $h_o$ is in $C^*$ and $f^*(h_o) \leq \lim \inf f^*(h_1)$.

Since $f^*$ is convex, the opposite inequality holds, therefore $f^*(h_o) = \lim \inf f^*(h_1)$ and $f^*$ is a closed convex function.

**Definition 1.7:** The conjugate set of the set $[f, \mathcal{C}]$, denoted by $[f, \mathcal{C}]^*$, is defined by

$$[f, \mathcal{C}]^* = \{[c,h] : c \geq (h,x) - t, \text{ for all } [t,x] \in [f, \mathcal{C}] \}.$$  

**Remark:** From definition 1.6 and the remark preceding it, $[f, \mathcal{C}]^* = [f^*, \mathcal{C}^*]$.

**Theorem 1.1:** If $[f, \mathcal{C}]$ is a closed convex set, then $[[f, \mathcal{C}]^*]^* = [f, \mathcal{C}]$.

**Remarks:**

(1) This theorem expresses the duality principle for conjugate convex functions and sets. It is a well-known theorem, Fenchel ([1], [2]). A generalized version of this theorem, developed in the next section, will be used to obtain certain results on duality for a class of convex programming problems.

(2) Since $[f, \mathcal{C}]^* = [f^*, \mathcal{C}^*]$, this theorem implies that $\mathcal{C} = \mathcal{C}^{**}$, $f = f^{**}$ and in particular that

$$f(x) = f^{**}(x) = \sup_{h \in \mathcal{C}^*} [(h,x) - f^*(h)].$$

(3) The basic relation between $[f, \mathcal{C}]$ and $[f^*, \mathcal{C}^*]$ is that

$$f^*(h) + f(x) \geq (h,x), \text{ for all } x \in \mathcal{C}, h \in \mathcal{C}^*.$$
For consider

\[(1.7) \quad r^*(h) = \sup_{x \in C}(h,x)-f(x) \geq (h,x)-f(x) \text{ for all } x \in C.\]

\[(1.8) \quad r(x) = \sup_{h \in \mathcal{C}^*}(h,x)-r^*(h) \geq (h,x)-r^*(h) \text{ for all } h \in \mathcal{C}^*.\]

Then (1.7) and (1.8) imply

\[(1.9) \quad f^*(h) + f(x) \geq (h,x) \text{ for all } h \in \mathcal{C}^*, \ x \in C.\]

Conversely (1.9) implies (1.7) and (1.8) since,

(a) \( f^*(h) \geq (h,x)-f(x) \text{ for all } x \in C, \) implies

\[r^*(h) \geq \sup_{x \in C}(h,x)-f(x).\] Therefore \( h \) is in \( C^* \) and

\[f^*(h) = \sup_{x \in C}(h,x)-f(x).\]

(b) \( r(x) \geq (h,x)-f^*(h) \text{ for all } h \in C^*, \) implies

\[f(x) \geq \sup_{h \in \mathcal{C}^*}(h,x)-r^*(h); \text{ so that } x \in C^{**} = C \text{ and}

\[r(x) = \sup_{h \in \mathcal{C}^*}[(h,x)-r^*(h)].\]

**Section 2: Generalized Conjugate Sets and Functions.**

In this section the concept of conjugate convex sets and functions, as developed in the previous section, is generalized. Results corresponding to those in the previous section are obtained, including a generalized version of the main theorem of the last section.
Definition 2.1: Let $A$ be an $m \times n$ matrix over the reals, let $f$ be a convex function defined on a convex set $C$ in $\mathbb{R}^n$. Then the set $C_A$ and the function $f_A$ are defined by

$$C_A = \{y \in \mathbb{R}^m : y = Ax \text{ for some } x \in C\},$$

$$f_A(y) = \inf \{f(x) : y = Ax, x \in C\}, \; y \in C_A.$$ 

Lemma 2.1: $C_A$ is a convex set and $f_A$ is a convex function on $C_A$.

Proof: Let $0 < \lambda < 1$, $y_1, y_2 \in C_A$, then there exists $x_1, x_2$ in $C$ such that $y_1 = Ax_1$, $y_2 = Ax_2$. Then

$$y = \lambda y_1 + (1 - \lambda)y_2 = \lambda Ax_1 + (1 - \lambda)Ax_2 = A(\lambda x_1 + (1 - \lambda)x_2) = Ax$$

where $x = \lambda x_1 + (1 - \lambda)x_2 \in C$, since $C$ is convex. Hence $y \in C_A$ and $C_A$ is convex.

For all $\varepsilon > 0$ and for all $y_1, y_2$ in $C_A$ there exists $x_1, x_2$ in $C$ such that $y_1 = Ax_1$, $y_2 = Ax_2$ and

$$f_A(y_1) + \varepsilon \geq f(x_1), \; f_A(y_2) + \varepsilon \geq f(x_2).$$

Therefore

$$f_A[\lambda y_1 + (1 - \lambda)y_2] \leq f[\lambda x_1 + (1 - \lambda)x_2] \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

$$\leq \lambda [f_A(y_1) + \varepsilon] + (1 - \lambda)[f_A(y_2) + \varepsilon].$$

Since $\varepsilon$ is arbitrary this argument shows that $f_A$ is a convex function.
**Definition 2.2:** Let \( A \) be an \( m \times n \) matrix and let \( f \) be a convex function defined on the convex set \( C \) in \( E^n \). The set \( \hat{\mathcal{C}} \) and the function \( \hat{f} \) are defined by

\[
\hat{\mathcal{C}} = \{ \hat{x} \in E^n : A\hat{x} = Ax \text{ for some } x \in C \} \quad \text{and} \quad \hat{f}(\hat{x}) = \inf \{ f(x) : Ax = A\hat{x}, \ x \in C \}, \ \hat{x} \in \hat{\mathcal{C}}.
\]

**Remarks:**

(1) As for \( C_A \) and \( r_A \) it is easily verified that \( \hat{\mathcal{C}} \) is a convex set and \( \hat{f} \) is a convex function on \( \hat{\mathcal{C}} \). Furthermore \( \hat{\mathcal{C}} = \mathcal{C} \) and \( \hat{f} = f_A \).

(2) If \( y \) is in \( C_A \), \( \hat{x} \) is in \( \hat{\mathcal{C}} \) and \( y = Ax \), then

\[
\hat{r}_A(y) = \inf \{ f(x) : y = Ax = A\hat{x}, \ x \in C \} = \hat{f}(\hat{x}).
\]

**Definition 2.3:** Given a \( m \times n \) matrix \( A \) and the set \([f,C]\), the set \( \Gamma' \) and the function \( \varphi \) are defined by

\[
\Gamma' = \{ \xi \in E^m : \sup_{x \in C} [\xi A x - f(x)] < \infty \}.
\]

\[
\varphi(\xi) = \sup_{x \in C} [\xi A x - f(x)], \ \xi \in \Gamma'.
\]

**Remark:** \( \varphi \) is called the conjugate function of \( f \) with respect to the matrix \( A \).

**Lemma 2.3:** \( \Gamma' \) is a convex set and \( \varphi \) is a closed convex function on \( \Gamma' \).

**Proof:** Similar to the proofs of lemmas (1.4) and (1.5).
Remark: It is clear that the set $\Gamma$, 
\[ \hat{\Gamma} = \{ \xi \in \mathbb{E}^m : \xi A = \xi A, \text{ for some } \xi \in \Gamma \} \]
is just the set $\Gamma$, and that $\hat{\varphi}(\xi) = \varphi(\xi)$. Furthermore 
\[ \varphi(\xi) = \sup_{x \in C}[\xi Ax - f(x)] = \sup_{x \in \hat{C}}[\xi Ax - \hat{f}(x)] \text{ for } \xi \in \Gamma. \]

Definition 2.4: The conjugate set with respect to the matrix $A$ of the set $[f, C]$ is denoted by $[f, C]_A^*$ and is defined by 
\[ [f, C]_A^* = \{ [c, \xi] \in \mathbb{E}^{m+1} : c \geq \xi Ax - t, \text{ for all } [t, x] \in [f, C] \}. \]

Remark: When $m = n$ and $A = I_n$, the identity matrix, then $[f, C]_A^* = [f, C]^*$. 

Lemma 2.4: $[f, C]_A^* = [\varphi, \Gamma]$. 

Proof: Let $[c, \xi] \in [f, C]_A^*$; then $c \geq \xi Ax - f(x)$ for all $x$ in $C$. 

Hence $c \geq \sup_{x \in C}[\xi Ax - f(x)]$, so that $\xi \in \Gamma$ and $c \geq \varphi(\xi)$. 

Hence $[c, \xi] \in [\varphi, \Gamma]$. 

Conversely, let $[c, \xi] \in [\varphi, \Gamma]$, then 
\[ c \geq \varphi(\xi) = \sup_{x \in C}[\xi Ax - f(x)] \geq \xi Ax - f(x), \text{ for all } x \in C \]
\[ \geq \xi Ax - t, \text{ for all } [t, x] \in [f, C], \]

Hence $[c, \xi] \in [f, C]^*$ . Therefore $[f, C]_A^* = [\varphi, \Gamma]$. 

Lemma 2.5: $[f_A, C_A]^* = [f, C]_A^*$. 
Proof  Let \([c, \xi] \in [f_A, c_A]^*\) then \(c \geq (\xi, y) - f_A(y)\) for all \(y\) in \(C_A\); therefore

\[c \geq \xi A x - f(x), \text{ for all } x \text{ in } C\]

\[\geq \xi A x - t, \text{ for all } [t, x] \in [f, C],\]

Hence \([c, \xi] \in [f, C]^*_A\).

Conversely let \([c, \xi] \in [f, C]^*_A\), then \(c \geq \xi A x - f(x)\) for all \(x \in C\). Therefore \(f(x) \geq \xi A x - c\) for all \(x \in C\).

Let \(y \in C_A\), then

\[f_A(y) = \inf \{f(x); y = Ax, x \in C\} \geq \inf \{\xi A x - c; y = Ax, x \in C\}\]

\[= (\xi, y) - c .\]

Since this holds for any \(y\) in \(C_A\),

\[c \geq (\xi, y) - f_A(y), \text{ for all } y \in C_A\]

\[\geq (\xi, y) - t, \text{ for all } [t, y] \in [f_A, C_A]^* ,\]

Hence \([c, \xi] \in [f_A, C_A]^*\). Therefore \([f_A, C_A]^* = [f, C]^*_A\).

Remarks:

(1) Since \(C_A = C\) and \(f_A = f\), the previous two lemmas show that

\([f_A, C_A]^* = [f_A, C_A]^* = [f, C]^*_A = [\varphi, \Gamma]\).

(2) Let \(A'\) denote the transpose of the matrix \(A\). Then the set \([\Gamma, C_A]_{A'}\) and the conjugate function to \(\varphi\)
with respect to \( A' \), \( \varphi_{A'} \), are given by

\[
\Gamma_{A'} = \{ n \in \mathbb{E}^n : n = \xi A : \text{for some } \xi \in \Gamma \}
\]

\[
\varphi_{A'}(n) = \inf\{ \varphi(\xi) : n = \xi A, \xi \in \Gamma \}, \ n \in \Gamma_{A'}.
\]

(5) the conjugate set to \( [\varphi_{A'}, \Gamma_{A'}] \) is given by

\[
[\varphi_{A'}, \Gamma_{A'}]^* = \{ [t, x] \in \mathbb{E}^{n+1} : t \geq \langle n, x \rangle - c, \forall [c, n] \in [\varphi_{A'}, \Gamma_{A'}] \}.
\]

(4) The generalized conjugate set to \( [\varphi, \Gamma] \) with respect to \( A' \) is given by

\[
[\varphi, \Gamma]^*_{A'} = \{ [t, x] \in \mathbb{E}^{n+1} : t \geq \xi A x - c, \forall [c, \xi] \in [\varphi, \Gamma] \}.
\]

(5) The previous lemmas can be applied to the set \( [\varphi, \Gamma] \) to obtain \( [\varphi, \Gamma]^*_{A'} = [\varphi_{A'}, \Gamma_{A'}]^* \).

**Lemma 2.0:** \([f, c] = [\hat{A}, \hat{c}] \subseteq [\varphi, \Gamma]^*_{A'}\).  

**Proof:** The definition of \([\hat{A}, \hat{c}]\) implies immediately that \([f, c] \subseteq [\hat{A}, \hat{c}]\). Suppose \([t, x_o]\) is in \([\hat{A}, \hat{c}]\); then \(t \geq \hat{A}(x_o), x_o \) is in \( \hat{c} \) and

\[
\varphi(\xi) = \sup_{x \in \hat{c}} [\xi A x - f(x)] = \sup_{x \in \hat{c}} [\xi A x - \hat{A}(x_o)] \geq \xi A x_o - \hat{A}(x_o), \xi \in \Gamma.
\]

Hence

\[
t \geq \hat{A}(x_o) \geq \xi A x_o - \varphi(\xi) \geq \xi A x_o - c, \forall [c, \xi] \in [\varphi, \Gamma].
\]

Therefore \([t, x_o] \in [\varphi, \Gamma]^*_{A'}\), and hence \([\hat{A}, \hat{c}] \subseteq [\varphi, \Gamma]^*_{A'}\).
Lemma 2.7: \( \hat{f} \) is a closed convex function on \( \mathcal{C} \) if and only if \( f_A \) is a closed convex function on \( \mathcal{C}_A \).

Proof: Suppose \( f_A \) is closed. Let \( x_0 \) be on the boundary of \( \mathcal{C} \) and let \( \{x_i\} \) be a sequence of vectors in \( \mathcal{C} \) such that \( x_i \to x_0 \) and \( \lim \inf \hat{f}(x_i) < \infty \). Let \( y_i = Ax_i \in \mathcal{C}_A \), \( y_0 = Ax_0 \); then \( y_0 \) is on the boundary of \( \mathcal{C}_A \). Furthermore \( y_i \to y_0 \) and

\[
\lim \inf \hat{f}(x_i) = \lim \inf f_A(y_i).
\]

Since \( f_A \) is closed, \( y_0 \in \mathcal{C}_A \), hence \( x_0 \in \mathcal{C} \) and

\[
\hat{f}(x_0) = f_A(y_0) = \lim \inf f_A(y_i) < \infty.
\]

Therefore \( \hat{f} \) is a closed convex function.

Conversely suppose \( \hat{f} \) is closed. Let \( y_0 \) be on the boundary of \( \mathcal{C}_A \) and let \( \{y_i\} \) be a bounded sequence of vectors in \( \mathcal{C}_A \) such that \( y_i \to y_0 \) and \( \lim \inf f_A(y_i) < \infty \). Then by Theorem A, appendix, there exists a set of vectors \( \{x_i\} \) which is bounded and such that \( y_i = Ax_i \) for all \( i \). Furthermore since \( \{x_i\} \) is bounded there exists a convergent subsequence, which we assume to be the whole sequence, say \( x_i \to x_0 \). Now \( \Rightarrow \lim \inf f_A(y_i) = \lim \inf \hat{f}(x_i) \) and since \( \hat{f} \) is closed, \( x_0 \in \mathcal{C} \) and hence \( y_0 = Ax_0 \in \mathcal{C}_A \).

Furthermore \( \Rightarrow \lim \inf f_A(y_i) = \lim \inf \hat{f}(x_i) = \hat{f}(x_0) = f_A(y_0) \) so that \( f_A \) is closed.
Lemma 2.8: \([t_o, x_o] \in [\hat{f}, \hat{c}]\) if and only if \([t_o, Ax_o] \in [f_A, C_A]\).

Proof: \([t_o, x_o] \in [f, C]\) if and only if \(t_o \geq \hat{f}(x_o) = f_A(Ax_o)\), if and only if \([t_o, Ax_o] \in [f_A, C_A]\).

Theorem 2.1: If \([f_A, C_A]\) is a closed convex set then \([\varphi, \Gamma]\), \(= [\hat{f}, \hat{c}]\).

Proof: Since \([f_A, C_A]\) is a closed convex set, by lemmas (1.2) and (2.7), \([\hat{f}, \hat{c}]\) is also a closed convex set. Assume \([t_o, x_o]\) is not in \([\hat{f}, \hat{c}]\); then by lemma (2.8), if \(y_o = Ax_o\), then \([t_o, y_o]\) is not in \([f_A, C_A]\). Since \([f_A, C_A]\) is a closed convex set, by lemma (1.3), there exists a non-vertical hyperplane that strictly separates \([t_o, y_o]\) from the set \([f_A, C_A]\).

Let the hyperplane \(P\) be
\[
P : \langle [\tau, \xi_o], [t, y] \rangle = k.
\]
Since \(\tau \neq 0\) we can take \(\tau = -1\), then
\[
(2.1) \quad -t + (\xi_o, y) \leq k \text{ for all } [t, y] \in [f_A, C_A],
\]
\[
(2.2) \quad -t_o + (\xi_o, y_o) > k.
\]
Therefore \(k \geq (\xi_o, y) - f_A(y), \quad \forall y \in C_A\), and
\[
k \geq \sup_{y \in C_A} [(\xi_o, y) - f_A(y)] = \sup_{x \in C_o} [\xi_o \cdot Ax - \hat{f}(x)] = \sup_{x \in C_o} [\xi_o \cdot Ax - f(x)].
\]
hence \( \xi_0 \) is in \( \Gamma \) and \( k, \xi_0 \). Therefore \([k, \xi_0]\) is in \([\varphi, \Gamma]\). From (2.2) \( t_0 < (\xi_0, y_0) - k \), and since 
\[ k, \varphi(\xi_0), \]
\[ t_0 < (\xi_0, y_0) - \varphi(\xi_0) = \xi_0 \mathbf{A} \mathbf{x}_0 - \varphi(\xi_0), \]
hence \([t_0, x_0]\) is not in \([\varphi, \Gamma]_A^* \). Therefore \([\varphi, \Gamma]_A^* \subset [\hat{\varphi}, \hat{\varphi}] \). By lemma (2.6) the opposite inclusion holds, hence there is equality, \([\varphi, \Gamma]_A^* = [\hat{\varphi}, \hat{\varphi}] \).

**Remarks:** If \( f_A \) is a closed convex function then

1. The above theorem and the remarks following lemma (2.5), show that
\[ [\hat{\varphi}, \hat{\varphi}] = [\varphi, \Gamma]_A^* = [\varphi_A', \Gamma_A', \Gamma_A'^*]. \]

2. Since \([\varphi, \Gamma] = [\hat{\varphi}, \hat{\varphi}]_A^* \), the above remark shows that
\[ [[\hat{\varphi}, \hat{\varphi}]_A^*]^* = [\hat{\varphi}, \hat{\varphi}] \].

3. Since \([\hat{\varphi}, \hat{\varphi}] = [\varphi, \Gamma]^* \), \( \hat{\varphi} \) is conjugate to \( \varphi \) with respect to \( A' \), hence
\[ \hat{\varphi}(x) = \sup_{\xi \in \Gamma} [\xi A - \varphi(\xi)] \).

4. The basic relation between a closed convex function and its generalized dual function is \( \hat{\varphi}(x) + \varphi(\xi) \geq \xi A \mathbf{x} \) for all \( x \in \hat{\varphi} \) and for all \( \xi \in \Gamma \).
Section 3: Concave Functions and their Duals

A theory of conjugate concave functions can be constructed analogous to the theory given in section 1 for convex functions. The facts and definitions which will be needed are obtained from those in section 1 by interchanging ≤ with ≥, +∞ with −∞, and infimum with supremum, wherever these occur. The results for generalized convex functions obtained in section 2 also carry over to concave functions with the above mentioned interchanges. A list of definitions and results is given, for completeness. This will serve to introduce the notation used in the following sections.

Definition 3.1: Let g be a concave function defined on the convex set D in $E^n$. The set $[g,D]$ in $E^{n+1}$ is defined by

$$[g,D] = \{ [s,x] \in E^{n+1} : x \in D, s < g(x) \}.$$ 

Remark: $[g,D]$ is a closed convex set if g is a closed concave function, that is, −g is a closed convex function.

Definition 3.2: Let $A$ be an $m \times n$ matrix over the real numbers. Let $[g,D]$ be defined as above, then

$$D_A = \{ y \in E^m : y = Ax \text{ for some } x \in D \}$$

$$g_A(y) = \sup \{ g(x) : x \in D, y = Ax \}, \quad y \in D_A$$
\[ \hat{D} = \{ \hat{x} \in \mathbb{E}^n : A\hat{x} = Ax \text{ for some } x \in D \} \]
\[ \hat{g}(\hat{x}) = \sup \{ g(x) : A\hat{x} = Ax, x \in D \}, \ \hat{x} \in \hat{D} \]
\[ \Delta = \{ \xi \in \mathbb{E}^m : \inf_{x \in D} [\xi A - g(x)] > -\infty \} \]
\[ \psi(\xi) = \inf_{x \in D} [\xi A - g(x)], \ \xi \in \Delta. \]

Remarks: Using the arguments of the previous sections it is easy to verify that

(1) \( D_A \) is a convex set.
(2) \( g_A \) is a concave function on \( D_A \).
(3) \( \hat{D}_A = D_A \)
(4) \( \hat{g}_A = g_A \)
(5) \( \Delta \) is a convex set.
(6) \( \psi \) is a closed concave function on \( \Delta \).

Definition 3.3: Given the matrix \( A \) and the set \([g,D]\), the conjugate set to \([g,D]\) with respect to the matrix \( A \) is given by
\[ [g,D]^*_A = \{ [c,\xi] \in \mathbb{E}^{m+1} : c \leq \xi A - t, \forall [t,x] \in [g,D] \}. \]

Remarks:

(1) As for convex functions, the following relation holds \( [g,D]^*_A = [\psi,\Delta] = [g_A,D_A]^* \).
(2) The main theorem of the previous section becomes under similar restrictions \( \hat{[g,D]} = [\psi_A,\Delta_A]^* = [\psi,\Delta]_A^* \) and hence \( \hat{g}(\hat{x}) = \inf_{\xi \in \Delta} [\xi A - \psi(\xi)] \).
(3) For closed concave functions the basic relation between the function and its generalized dual function is
\[ \hat{g}(x) + \psi(\xi) \leq \xi Ax \quad \text{for all } x \in \hat{\Omega}, \]
and for all \( \xi \in \Delta \), provided \( g_A \) is closed.

Section 4: Duality for Convex and Concave Functions

The basic relations between convex and concave functions, given in the previous sections, imply that if \([f_A, \xi_A]\) and \([g_A, \xi_D]\) are closed then
\[ \hat{f}(x) + \phi(\xi) \geq \xi Ax \geq \hat{g}(x) + \psi(\xi) \]
for all \( x \in \hat{\Omega} \Delta \) and for all \( \xi \in \Gamma \cap \Delta \). Therefore

\[ \inf_{\xi \in \Gamma \cap \Delta} \{ \phi(\xi) + \psi(\xi) \} \geq \sup_{x \in \hat{\Omega} \Delta} \{ \hat{g}(x) - \hat{f}(x) \}. \tag{4.1} \]

The main purpose of this section is to answer the question of equality in (4.1). The following theorem due to Fenchel [1], answers this question for the special case of ordinary dual functions \( f^* \) and \( g^* \).

**Theorem 4.1:** Let \( f \) be a closed convex function on the convex set \( C \), let \( g \) be a closed concave function on the convex set \( D \). Let \( C \) and \( D \) have relative interior points in common, similarly for \( C^* \) and \( D^* \). Then
\[ \sup_{x \in \hat{\Omega} \Delta} \{ g(x) - f(x) \} = \inf_{\xi \in \Gamma \cap \Delta} \{ f^*(\xi) - g^*(\xi) \}, \]
and the supremum and infimum are attained.
The general version of this theorem will follow from the existence of a non-vertical hyperplane separating two special convex sets, as will be shown by the next lemma. The existence of the non-vertical hyperplane is then assured by lemma (4.2). The general theorem is then stated and several results obtained.

Lemma 4.1: Suppose that \( \mu = \sup_{x \in \mathbb{C} \cap \mathbb{D}} [\hat{g}(x) - \hat{f}(x)] < \infty \). Let \( \bar{f}_A(y) = f_A(y) + \mu \), for all \( y \in C_A \). If the sets \([\bar{f}_A, C_A]\) and \([g_A, D_A]\) can be separated by a non-vertical hyperplane, then there exists a \( \xi_0 \in \Gamma \cap \Delta \) such that

\[
\inf_{\xi \in \Gamma \cap \Delta} [\varphi(\xi) - \psi(\xi)] = \varphi(\xi_0) - \psi(\xi_0) = \sup_{x \in \mathbb{C} \cap \mathbb{D}} [\hat{g}(x) - \hat{f}(x)].
\]

Proof: Let \( x \in \mathbb{C} \cap \mathbb{D} \) and \( y = Ax \), then \( y \in C_A \cap D_A \) and \( \hat{f}_A(y) = \hat{f}(x) \), \( g_A(y) = \hat{g}(x) \). Therefore

\[
\mu = \sup_{x \in \mathbb{C} \cap \mathbb{D}} [\hat{g}(x) - \hat{f}(x)] = \sup_{y \in C_A \cap D_A} [g_A(y) - f_A(y)].
\]

Let the non-vertical separating hyperplane \( P \) be given by \( P : ([t, \xi_0], [t, y]) = c_0 \), normalized so that \( \tau = -1 \). Then

\[
(t + (\xi_0, y) \leq c_0 \quad \text{for all } [t, y] \in [\bar{f}_A, C_A] \quad (4.2)
\]

\[
(t + (\xi_0, y) \geq c_0 \quad \text{for all } [t, y] \in [g_A, D_A] \quad (4.3)
\]

From (4.3), \( (\xi_0, y) - g_A(y) \geq c_0 \), for all \( y \in D_A \), hence \( \xi_0 \in \Delta \) and \( \psi(\xi_0) \geq c_0 \). From (4.2), \( (\xi_0, y) - \bar{f}_A(y) \leq c_0 \) \( \forall y \in C_A \).
hence $(\xi_0, y) - f_A(y) \leq c_0 + \mu$ so that $\xi_0 \in \Gamma$ and $c_0 + \mu \geq \varphi(\xi_0)$.

Therefore $\xi_0 \in \Gamma \cap \Delta$ and

$$\sup_{x \in \Gamma \cap \Delta} [\hat{\varphi}(x) - \hat{\psi}(x)] = \mu \geq \varphi(\xi_0) - c_0 \geq \varphi(\xi_0) - \psi(\xi_0) \geq \inf_{\xi \in \Gamma \cap \Delta} [\varphi(\xi) - \psi(\xi)].$$

Since the opposite inequality (4.1) always holds, there must be equality. This completes the proof of the lemma.

Lemma 4.2: Suppose $[f_A, C_A]$ and $[g_A, D_A]$ are closed convex sets such that $C_A, D_A$ have relative interior points in common.

Suppose $\mu = \sup_{y \in C_A \cap D_A} [g_A(y) - f_A(y)] < \infty$. Let $\bar{f}_A(y) = f_A(y) + \mu$ for all $y \in C_A$. Then there exists a non-vertical hyperplane which separates $[g_A, D_A]$ and $[\bar{f}_A, C_A]$.

Proof: By the choice of $\mu$ the sets $[\bar{f}_A, C_A]$ and $[g_A, D_A]$ come arbitrary close to each other but do not overlap at any interior points. Hence there is a separating hyperplane $P$ which is a supporting hyperplane to both sets. We need to show it is non-vertical.

If $P$ is a vertical hyperplane, project $[\bar{f}_A, C_A]$, $[g_A, D_A]$ and $P$ vertically onto $E^m$. The two convex sets project onto $C_A$ and $D_A$, and the projection of $P$ is a hyperplane $P_1$ of $E^m$ which must separate $C_A$ and $D_A$. 
This contradicts the hypothesis that \( C_A \) and \( D_A \) have relative interior points in common. Hence \( P \) cannot be a vertical hyperplane.

Jewellery theorem for convex and concave functions.

**Theorem 4.2**: Let \( f \) be a closed convex function on the convex set \( C \), let \( g \) be a closed concave function on the convex set \( D \). Suppose \([f_A,C_A]\) and \([g_A,D_A]\) are closed convex sets. Suppose \( C_A,D_A \) have relative interior points in common and \( \Gamma_A,\Gamma_A' \) have relative interior points in common. Then the following statements are equivalent

1. \( \sup_{x \in \partial C} [\hat{g}(x)-\hat{f}(x)] \) is finite.
2. \( \inf_{\xi \in \Gamma \cap \Delta} [\varphi(\xi)-\psi(\xi)] \) is finite.

In either case

\[
\sup_{x \in \partial C \cap \partial D} [\hat{g}(x)-\hat{f}(x)] = \inf_{\xi \in \Gamma \cap \Delta} [\varphi(\xi)-\psi(\xi)]
\]

and the infimum and supremum are attained.

**Proof**: If (1) holds then by lemmas (4.1) and (4.2), there exists \( \xi_0 \in \Gamma \cap \Delta \) such that

\[
\sup_{x \in \partial C \cap \partial D} [\hat{g}(x)-\hat{f}(x)] = \varphi(\xi_0)-\psi(\xi) = \inf_{\xi \in \Gamma \cap \Delta} [\varphi(\xi)-\psi(\xi)]
\]

and hence (2) holds.

Conversely suppose (2) holds. Since \([\phi,\hat{\Gamma}] = [\varphi,\Gamma] \)
is closed, by lemmas (1.2) and (2.7), \([\varphi_{A_1}, \Gamma_{A_1}]\) is closed. Similarly \([\psi_{A_1}, \Delta_{A_1}]\) is closed. Since \(\Gamma_{A_1}\) and \(\Delta_{A_1}\) have relative interior points in common, lemmas (4.1) and (4.2) can be applied to the functions \(\varphi\) and \(\psi\). Since \([\hat{\varphi}, \hat{\Delta}] = [\varphi, \Gamma]^{*}_{A_1}\) and \([\hat{\psi}, \hat{\Delta}] = [\psi, \Delta]^{*}_{A_1}\), by theorem (2.1), lemma (4.1) shows there exists \(x_0 \in \hat{\Omega} \cap \hat{\Delta}\) such that

\[
\sup_{\xi \in \hat{\Omega} \cap \hat{\Delta}} [\varphi(\xi) - \varphi(\xi)] = \hat{\varphi}(x_0) - \hat{\psi}(x_0) = \inf_{x \in \hat{\Omega} \cap \hat{\Delta}} [\hat{\varphi}(x) - \hat{\psi}(x)].
\]

Hence (1) holds and furthermore

\[
\hat{\varphi}(x_0) - \hat{\varphi}(x_0) = \sup_{x \in \hat{\Omega} \cap \hat{\Delta}} [\hat{\varphi}(x) - \hat{\varphi}(x)] = \inf_{\xi \in \hat{\Omega} \cap \hat{\Delta}} [\varphi(\xi) - \psi(\xi)]
\]

\[
= \varphi(x_0) - \psi(x_0).
\]

**Remarks:** From theorem (2.1) and the remarks following definition (3.3)

1. \(\hat{\varphi}(x) = \inf_{\xi \in \Delta} \{\xi A - \varphi(\xi)\}\)
2. \(\varphi(x) = \sup_{x \in \mathcal{C}} \{\xi A - \hat{\varphi}(x)\}\)
3. \(\hat{\varphi}(x) = \sup_{\xi \in \Gamma} \{\xi A - \varphi(\xi)\}\)
4. \(\psi(x) = \inf_{x \in \mathcal{D}} \{\xi A - \hat{\psi}(x)\}\)

These relations and the previous theorem give the following results.
\[ (4.4) \quad \sup_x \inf_{\xi \in \Delta} [\xi \问询(x) - \hat{f}(x)] = \inf_{\xi \in \Gamma} \sup_{x \in D} [\xi \问询(x) - \hat{f}(x)] \]

\[ (4.5) \quad \sup_x \inf_{\xi \in \Gamma} [\hat{g}(x) + \varphi(\xi) - \xi \问询(x)] = \inf_{\xi \in \Gamma} \sup_{x \in D} [\hat{g}(x) + \varphi(\xi) - \xi \问询(x)] \]

**Theorem 4.3:** Under the hypothesis of theorem (4.2), the infimum and supremum in equations (4.4) and (4.5) can be taken over the smaller sets \( \Gamma \cap \Delta \) and \( \Delta \cap \Gamma \) respectively.

**Proof:** To show that the infimum and supremum can be taken over the smaller sets \( \Gamma \cap \Delta \) and \( \Delta \cap \Gamma \), first note that

\[ (4.6) \quad \varphi(\xi) - \psi(\xi) = \sup_{x \in C} [\xi \问询(x) - \hat{f}(x)] \geq \sup_{x \in \Gamma} [\xi \问询(x) - \hat{f}(x)] \]

\[ (4.7) \quad \hat{g}(x) - \hat{f}(x) = \inf_{\xi \in \Delta} [\xi \问询(x) - \hat{f}(x)] \leq \inf_{\xi \in \Gamma} [\xi \问询(x) - \hat{f}(x)] \]

Hence

\[ (4.8) \quad \sup_x \inf_{\xi \in \Gamma} [\xi \问询(x) - \hat{f}(x)] \geq \sup_{x \in \Gamma} [\hat{g}(x) - \hat{f}(x)] = \inf_{\xi \in \Delta} [\varphi(\xi) - \psi(\xi)] \geq \inf_{\xi \in \Gamma} \sup_{x \in D} [\xi \问询(x) - \hat{f}(x)] . \]

Since the opposite inequality always holds, there is equality.
Therefore

\[(4.9) \quad \sup_{x \in \mathcal{D}} \inf_{\xi \in \Gamma \setminus \Delta} [\xi A x - \psi(\xi) - \hat{g}(x)] = \inf_{\xi \in \Gamma \setminus \Delta} \sup_{x \in \mathcal{D}} [\xi A x - \psi(\xi) - \hat{g}(x)].\]

In the same way it can be shown that

\[(4.10) \quad \sup_{x \in \mathcal{D}} \inf_{\xi \in \Gamma \setminus \Delta} [\hat{g}(x) + \varphi(\xi) - \xi A x] = \inf_{\xi \in \Gamma \setminus \Delta} \sup_{x \in \mathcal{D}} [\hat{g}(x) + \varphi(\xi) - \xi A x].\]

**Remarks:** Under the hypothesis of theorem (4.2), it is easily verified that

1. The supremum and infimum in equation (4.4) can be taken over the larger sets $\mathcal{D}$ and $\Delta$ respectively.

2. The supremum and infimum in equation (4.5) can be taken over the larger sets $\mathcal{D}$ and $\Gamma$ respectively.

**Section 5: Lagrangians and Dual Convex Programming Problems**

A convex programming problem will mean either minimize a convex function over a convex set, or maximize a concave function over a convex set.

In this section the close connection between a saddle-point of an appropriate Lagrangian function and the solutions to a pair of dual convex programming problems will be studied.
These problems and their associated Lagrangian functions will extend the duality theory of the previous sections. In particular a duality theorem for a pair of convex programming problems is given.

**Definition 5.1:** Let $f$ be a closed convex function on the convex set $C$ and let $g$ be a closed concave function on the convex set $D$. Let $[f,C]_A^* = \{\varphi, \Gamma\}$, $[g,D]_A^* = \{\varphi, \Delta\}$. Then the Lagrangian functions $F, G, \hat{F}, \hat{G}$ are defined by

\begin{align*}
\text{(5.1)} & \quad F(x, \xi) = \xi \text{Ax} - f(x) - \psi(\xi), \quad \{x, \xi\} \in C \times \Delta \\
\text{(5.2)} & \quad G(x, \xi) = g(x) + \varphi(\xi) - \xi \text{Ax}, \quad \{x, \xi\} \in D \times \Gamma \\
\text{(5.3)} & \quad \hat{F}(x, \xi) = \xi \text{Ax} - \hat{f}(x) - \psi(\xi), \quad \{x, \xi\} \in C \times \Delta \\
\text{(5.4)} & \quad \hat{G}(x, \xi) = \hat{g}(x) + \varphi(\xi) - \xi \text{Ax}, \quad \{x, \xi\} \in D \times \Gamma.
\end{align*}

**Definition 5.2:** The pair of vectors $\{x^*, \xi^*\}$ is called a saddle-point of a Lagrangian function $L(x, \xi)$ if

\begin{align*}
\text{(5.5)} & \quad L(x^*, \xi^*) \geq L(x^*, \xi) \geq L(x, \xi^*), \quad \forall \{x, \xi\} \text{ in the domain of } L.
\end{align*}

**Lemma 5.1:** Let $g_A$ be a closed concave function. Suppose $\{x_o, \xi_o\}$ is a saddle-point of $F(x, \xi)$ such that $F(x_o, \xi_o)$ is finite; then

\begin{align*}
\text{(1)} & \quad x_o \in C \cap \hat{D} \text{ and } \xi_o \in \Gamma \cap \Delta \\
\text{(2)} & \quad \varphi(\xi_o) - \psi(\xi_o) = F(x_o, \xi_o) = \hat{g}(x_o) - \hat{f}(x_o) \\
\text{(3)} & \quad \varphi(\xi_o) - \psi(\xi_o) = \inf_{\xi \in \Gamma} [\varphi(\xi) - \psi(\xi)] \\
\text{(4)} & \quad \hat{g}(x_o) - \hat{f}(x_o) = \sup_{x \in C \cap \hat{D}} [\hat{g}(x) - \hat{f}(x)].
\end{align*}
Title: "Cancer Programming"

Set: MUIR

Vol.

Years

Donor

No of bd vols

Tie-up:

Note for Cat:

2 copies

Date

No of vols

Rebind:

Bind everything "As is"

Bind all but back covers & stiff advts

Vols. split into book

Put TPC in book

Put Index in books

No main TP

Use front cover no

No main Contents

No main Index

Individual contents

Tog. at front

Where they appear

Include all text in Front advts

Back advts.

Vols. sep by Kraft front cover

Pocket

Folded material

No Trim

Color NAVY

Arrangement
Proof: Since \( \{x_0, \xi_0\} \) is a saddle-point of \( F(x, \xi) \),

\[
\xi A x_0 - f(x_0) - \psi(\xi) \geq F(x_0, \xi_0) \geq \xi_0 A x - f(x) - \psi(\xi_0)
\]

for all \( \xi \) in \( \Delta \) and for all \( x \) in \( C \).

The first inequality implies

\[
\inf_{\xi \in \Delta} [\xi A x_0 - f(x_0)] \geq F(x_0, \xi_0) > -\infty
\]

therefore \( x_0 \in \overset{\text{\( \hat{\Delta} \)}}{\Delta} \) and

\[
F(x_0, \xi_0) \leq \overset{\text{\( \hat{\Delta} \)}}{g}(x_0) - f(x_0), \quad x_0 \in \cap \Delta.
\]

The second inequality implies

\[
\infty > F(x_0, \xi_0) \geq \sup_{x \in C} [\xi_0 A x - f(x) - \psi(\xi_0)]
\]

therefore \( \xi_0 \in \Gamma \) and \( F(x_0, \xi_0) \geq \varphi(\xi_0) - \psi(\xi_0), \quad \xi_0 \in \Gamma \cap \Delta \).

Since \( \inf_{\xi \in \Gamma \cap \Delta} [\varphi(\xi) - \psi(\xi)] \geq \sup_{x \in \cap \Delta} [\overset{\text{\( \hat{\Delta} \)}}{g}(x) - f(x)] \) always holds,

the infimum and supremum are attained at the points \( \xi_0 \) and \( x_0 \) respectively.

Lemma 5.2: Let \( f_A \) be a closed convex function. Suppose

\( \{x_1, \xi_1\} \) is a saddle-point of \( G(x, \xi) \) such that \( G(x_1, \xi_1) \)

is finite, then

(1) \( x_1 \in \overset{\text{\( \hat{\Delta} \)}}{\Delta} \cap \Delta \) and \( \xi_1 \in \Gamma \cap \Delta \)

(2) \( \varphi(\xi_1) - \psi(\xi_1) = G(x_1, \xi_1) = g(x_1) - \overset{\text{\( \hat{\Delta} \)}}{f}(x_1) \)
(3) \[ \varphi(\xi_1) - \psi(\xi_1) = \inf_{\xi \in \Gamma \Delta} [\varphi(\xi) - \psi(\xi)] \]

(4) \[ g(x_1) - \hat{f}(x_1) = \sup_{x \in \mathcal{D}} [g(x) - \hat{f}(x)] \]

**Proof:** Similar to lemma 5.1.

**Remarks:** In terms of Lagrangian functions, the remarks following theorem (4.3) show that if there exist vectors \( x^* \in \hat{U} \hat{D} \) and \( \xi^* \in \Gamma \Delta \) such that

\[ \hat{g}(x^*) - \hat{f}(x^*) = \varphi(\xi^*) - \psi(\xi^*) \]

then \( \{x^*, \xi^*\} \) is a saddle-point for both \( \hat{F}(x, \xi) \) and \( \hat{G}(x, \xi) \).

The converse statement is also true, as is shown by the following theorem.

**Theorem 5.1:** Let \( f_A \) and \( g_A \) be closed. Then the statement

(i) \( \{x^*, \xi^*\} \) is a saddle-point of \( \hat{F}(x, \xi) \) and \( \hat{F}(x^*, \xi^*) \) is finite.

is equivalent to the statement

(ii) \( \{x^*, \xi^*\} \) is a saddle-point of \( \hat{G}(x, \xi) \) and \( \hat{G}(x^*, \xi^*) \) is finite.

Furthermore

(1) \[ x^* \in \hat{U} \hat{D}, \xi^* \in \Gamma \Delta \]

(2) \[ \hat{g}(x^*) - \hat{f}(x^*) = \varphi(\xi^*) - \psi(\xi^*) \]

(3) \[ \hat{g}(x^*) - \hat{f}(x^*) = \sup_{x \in \mathcal{D}} [\hat{g}(x) - \hat{f}(x)] \]

(4) \[ \varphi(\xi^*) - \psi(\xi^*) = \inf_{\xi \in \Gamma \Delta} [\varphi(\xi) - \psi(\xi)] \]
\( F(x^*,\xi^*) = \hat{G}(x^*,\xi^*) = \varphi(\xi^*) - \psi(\xi^*) \).

**Proof:** If \( \{x^*,\xi^*\} \) is a saddle-point of \( \hat{F} \) for which \( \hat{F}(x^*,\xi^*) \) is finite, then using the proof of lemma (5.1) with \( f \) replaced by \( \hat{f} \), \( C \) with \( \hat{C} \), it is easy to verify that (1), (2), (3) and (4) hold. But the remarks preceding this theorem show that if (1) and (2) hold, then \( \{x^*,\xi^*\} \) is a saddle-point of \( \hat{G}(x,\xi) \), and then (5) holds.

Conversely a saddle-point \( \{x^*,\xi^*\} \) of \( \hat{G} \) is also a saddle-point of \( \hat{F} \) by a similar argument.

**Lemma 5.3:** Let \( f_A, g_A \) be closed. Let \( \{x_o^*,\xi_o^*\} \) be a saddle-point of \( \hat{F} \), \( \{x^*,\xi^*\} \) be a saddle-point of \( \hat{F} \), and \( \{x_1^*,\xi_1^*\} \) be a saddle-point of \( G \), such that \( F(x_o^*,\xi_o^*) \), \( \hat{F}(x^*,\xi^*) \) and \( G(x_1^*,\xi_1^*) \) are all finite. Then

1. \( F(x_o^*,\xi_o^*) = \hat{F}(x^*,\xi^*) = G(x_1^*,\xi_1^*) \)
2. \( f(x_o^*) = \hat{f}(x_o^*) \), \( g(x_1^*) = \hat{g}(x_1^*) \)
3. \( \{x_o^*,\xi_o^*\} \) is a saddle-point of \( \hat{F}(x,\xi) \)
4. \( \{x_1^*,\xi_1^*\} \) is a saddle-point of \( \hat{F}(x,\xi) \).

**Proof:** By the preceding theorem and lemmas (5.1) and (5.2),

\[
F(x_o^*,\xi_o^*) = \varphi(\xi_o^*) - \psi(\xi_o^*) = \inf_{\xi \in \mathbb{R}^\Delta} [\varphi(\xi) - \psi(\xi)]
\]

\[
\hat{F}(x^*,\xi^*) = \varphi(\xi^*) - \psi(\xi^*) = \inf_{\xi \in \mathbb{R}^\Delta} [\varphi(\xi) - \psi(\xi)]
\]

\[
G(x_1^*,\xi_1^*) = \varphi(\xi_1^*) - \psi(\xi_1^*) = \inf_{\xi \in \mathbb{R}^\Delta} [\varphi(\xi) - \psi(\xi)]
\]
therefore $F(x_0,\xi_0) = \hat{F}(x^*,\xi^*) = G(x^*,\xi^*)$.

Since $\{x_0,\xi_0\}$ is a saddle-point of $F(x,\xi)$ and $\{x^*,\xi^*\}$ is a saddle-point of $\hat{F}(x,\xi)$, by Lemma (5.1) and Theorem (5.1), $\xi, \xi_0 \in \Gamma\Delta$. Hence

$$F(x_0,\xi_0) = \xi_0 A x_0 - f(x_0) - \psi(\xi_0)$$
$$\leq \xi^* A x_0 - f(x_0) - \psi(\xi^*)$$
$$\leq \xi^* A x^* - \hat{f}(x^*) - \psi(\xi^*)$$
$$\leq \xi^* A x^* - \hat{f}(x^*) - \psi(\xi^*)$$
$$= \hat{F}(x^*,\xi^*) = F(x_0,\xi_0).$$

Therefore there is equality throughout; thus $f(x_0) = \hat{f}(x_0)$ and $\hat{F}(x_0,\xi_0) = F(x_0,\xi_0) = \hat{F}(x^*,\xi^*)$.

Now $\{x_0,\xi_0\}$ is a saddle-point of $\hat{F}$ since

$$\hat{F}(x_0,\xi) = F(x_0,\xi) \geq F(x_0,\xi_0) = \hat{F}(x_0,\xi_0), \forall \xi \in \Delta$$

and

$$\hat{F}(x_0,\xi_0) = F(x_0,\xi_0) = \xi(\xi_0) - \psi(\xi_0)$$
$$\geq \hat{F}(x,\xi_0), \forall x \in \hat{C}.$$

Similarly it is easy to verify that

(i) $g(x_1) = \hat{g}(x_1)$.

(ii) $\hat{G}(x_1,\xi_1) = G(x_1,\xi_1) = \hat{G}(x^*,\xi^*) = \hat{F}(x^*,\xi^*)$.

(iii) $\{x_1,\xi_1\}$ is saddle-point of $\hat{G}$ and hence of $\hat{F}$.

(iv) $\hat{G}(x_1,\xi_1) = \hat{F}(x_1,\xi_1) = \hat{F}(x^*,\xi^*)$. 
Remarks: The Lagrangian functions of this section, the duality theorem (4.2) and the results of the previous section are all related to the following convex programming problems.

Problem I:
Find \( \sup [g(x) - f(x)] \) for \( x \in \text{CND} \).

Problem II:
Find \( \inf [\varphi(\xi) - \psi(\xi)] \) for \( \xi \in \text{FND} \).

Problem I:
Find \( \sup [\hat{g}(x) - \hat{f}(x)] \) for \( x \in \hat{\text{CND}} \).

The Lagrangian functions \( F \) and \( G \) are associated with problems I and II, and the Lagrangian functions \( \hat{F} \) and \( \hat{G} \) are associated with problems \( \hat{I} \) and \( \hat{II} \).

The duality theorem (4.2) and the results of this section can now be summarized in the following duality theorem for convex programming problems.

**Theorem 5.2:** Let \( f \) be a closed convex function on the convex set \( C \) and let \( g \) be a closed concave function on the convex set \( D \). Suppose that \( f_A \) and \( g_A \) are closed, \( C_A, D_A \) have relative interior points in common, and \( \cap A, \Delta A \) have relative interior points in common. Then the following statements are equivalent:

1. \( \{x^*, \xi^*\} \) is a saddle-point of \( \hat{F} \) and \( \hat{f}(x^*, \xi^*) \) is finite.
Section 6: Comparison of Duality Theorems

In this section the well-known duality theorem for linear programming will be obtained from the duality theorem (5.2).

Theorems (4.2) and (5.2) are then compared with a duality theorem due to Dorn [3]. Dorn's theorem will be derived from theorem (4.2) under the assumption that the functions involved have continuous first partial derivatives.

Lemma 6.1: Let \( b \) be a constant vector in \( E^m \), \( c \) a constant vector in \( E^n \) and \( A \) a \( m \times n \) matrix of rank \( m \leq n \).

Let \( C = \{x \in E^n : Ax \leq b\} \), \( D = \{x \in E^n : x \geq 0\} \). If \( f(x) = 0 \) for all \( x \in C \), \( g(x) = (c,x) \) for \( x \in D \), then the convex programming problems I and II become the following linear programming problems.

Problem I': Maximize \((c,x)\) subject to \( Ax \leq b, \ x \geq 0\).

Problem II': Minimize \((\xi,b)\) subject to \( \xi A \geq c, \ \xi \geq 0\).

Proof: Clearly under the conditions stated the convex programming problem I becomes problem I'. Now
\[ C = \{ x \in \mathbb{E}^n : A x \leq b \} \]
\[ = \{ x \in \mathbb{E}^n : \xi A x \leq (\xi, b), \; \xi \geq 0 \} \]

hence by definition

\[ \Gamma' = \{ \xi \in \mathbb{E}^m : \sup_{x \in C} \xi A x < \infty \} \]
\[ = \{ \xi \in \mathbb{E}^m : \xi \geq 0 \} \]

and for \( \xi \in \Gamma' \),

\[ \phi(\xi) = \sup_{x \in C} \xi A x = (\xi, b) \]

since \( \text{rank}(A) = m \). Now

\[ D = \{ x \in \mathbb{E}^n : x \geq 0 \} \]

hence by definition

\[ \Delta = \{ \xi \in \mathbb{E}^m : \inf_{x \in D} [\xi A x - (c, x)] > -\infty \} \]
\[ = \{ \xi \in \mathbb{E}^m : \inf_{x \geq 0} (\xi A - c, x) > -\infty \} \]
\[ = \{ \xi \in \mathbb{E}^m : \xi A - c > 0 \} \]

and for \( \xi \in \Delta \),

\[ \psi(\xi) = \inf_{x > 0} (\xi A - c, x) = 0 . \]

Hence problem II becomes problem II'.

Remarks: The associated Lagrangian functions \( F \) and \( G \) for problems I and II reduce to the functions:

\[ G(x, \xi) = (c, x) + (\xi, b) - \xi A x \]

defined on \( D \times \Gamma' \), and \( F(x, \xi) = \xi A x \) defined on \( C \times \Delta \).
\[ C = \{ x \in \mathbb{E}^n : Ax \preceq b \} \]
\[ = \{ x \in \mathbb{E}^n : \xi A x \preceq (\xi, b), \quad \xi \geq 0 \} \]

hence by definition
\[ \Gamma' = \{ \xi \in \mathbb{E}^m : \sup_{x \in C} \xi A x < \infty \} \]
\[ = \{ \xi \in \mathbb{E}^m : \xi \geq 0 \} \]

and for \( \xi \in \Gamma' \),
\[ \varphi(\xi) = \sup_{x \in C} \xi A x = (\xi, b) \]

since \( \text{rank}(A)= m \). Now
\[ D = \{ x \in \mathbb{E}^n : x \succeq 0 \} \]

hence by definition
\[ \Delta = \{ \xi \in \mathbb{E}^m : \inf_{x \in D} [\xi A x - (c, x)] > -\infty \} \]
\[ = \{ \xi \in \mathbb{E}^m : \inf_{x \succeq 0} (\xi A - c, x) > -\infty \} \]
\[ = \{ \xi \in \mathbb{E}^m : \xi A - c \succeq 0 \} \]

and for \( \xi \in \Delta \),
\[ \psi(\xi) = \inf_{x \succeq 0} (\xi A - c, x) = 0 . \]

Hence problem II becomes problem II'.

**Remarks:** The associated Lagrangian functions \( F \) and \( G \) for problems I and II reduce to the functions:
\[ G(x, \xi) = (c, x) + (\xi, b) - \xi A x \]
defined on \( D \times \Gamma' \), and \( F(x, \xi) = \xi A x \) defined on \( C \times \Delta \).
Definition 6.1: Let $A$ be an $m \times n$ matrix with rank $m \leq n$ and let $b$ be a constant vector in $\mathbb{E}^m$. Let $C = \{x \in \mathbb{E}^n : Ax \leq b\}$, $D = \{x \in \mathbb{E}^n : x \geq 0\}$ and suppose $f(x) = 0$ on $C$. Let $g$ be a closed concave function defined on the set $D$ and suppose that $g$ has continuous first partial derivatives on $D$. Then the following convex programming problems are defined:

**Problem 1:** Maximize $g(x)$ subject to $Ax \leq b$, $x \geq 0$

**Problem 2:** Minimize $H(x, \xi) = (\xi, b) + g(x) - (\nabla g(x), x)$ subject to $\xi A \geq \nabla g(x)$, $\xi \geq 0$.

Note that these problems reduce to the linear programming problems I' and I'I when $g(x) = (c, x)$, $c$ a constant vector in $\mathbb{E}^n$.

The following duality theorem for the programming problems 1 and 2 is due to Dorn [3].

**Theorem 6.1:** (Dorn)

Let the programming problems 1 and 2 be given as in definition (6.1). Let $R = \{x \in \mathbb{E}^n : \xi A \geq \nabla g(x), \xi \geq 0\}$, $S = \{y \in \mathbb{E}^n : \xi A \geq y, \xi \geq 0\}$. The transformation $y = \nabla g(x)$ maps $R$ onto $S$, and there exists a function which maps $S$ onto $R$. If this function is once differentiable and one-to-one in a neighborhood of the maximum, then the following statements hold.
(1) If there exists a vector \( x^* \) which maximizes \( g(x) \) in problem 1, then there also exist vectors \( x = x^*, \xi = \xi^* \) which minimize \( H(x, \xi) \) in problem 2.

(2) Conversely if \( x^*, \xi^* \) are vectors which minimize \( H(x, \xi) \) in problem 2, then \( x = x^* \) maximizes \( g(x) \) in problem 1.

Furthermore maximum \( g(x) = \) minimum \( H(x, \xi) \).

In order to compare Dorn's results with our duality theorems, the following lemmas will be needed.

Lemma 6.2: Let \( g(x) \) be a concave function defined over the convex set \( D \), with continuous first partial derivatives on \( D \). Then for all points \( x, x^* \) in \( D \),

\[
(\nabla g(x), x^* - x) \geq g(x^*) - g(x).
\]

Proof: The result is well-known. A proof can be found in Dorn's paper [3].

Lemma 6.3: Let \( g(x) \) be a closed concave function defined on the convex set \( D = \{ x \in \mathbb{R}^n : x \geq 0 \} \), with continuous first partial derivatives on \( D \). Let

\[
Y = \{ \xi \in \mathbb{R}^m : \min_{x \geq 0} [\xi A x - g(x)] = \xi A x^* - g(x^*) > -\infty, x^* \geq 0 \}
\]

Then for any \( \xi \in Y \) there exists \( x^* \geq 0 \) such that

(1) \( \xi A \geq \nabla g(x^*) \)

(2) \( \xi A x^* = (\nabla g(x^*), x^*) \).
Proof: Let $\xi \in Y$ be fixed.

(1) By definition of $Y$ there exists $x^* \geq 0$ such that $\xi A x^* - g(x^*) \leq \xi A x - g(x)$, $\forall x \geq 0$. Hence by lemma (6.1),

$$(\nabla g(x), x^* - x) \geq g(x^*) - g(x) \geq (\xi A, x^* - x), \forall x \geq 0$$

therefore $(\xi A - \nabla g(x), x - x^*) \geq 0, \forall x \geq 0$.

Let $x = x^* + \lambda e_i, \lambda > 0$, $e_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{E}^n$; then $(\xi A - \nabla g(x^* + \lambda e_i), \lambda e_i) \geq 0$ for each $i$. Therefore

$$(\xi A)^i \geq \frac{\partial g(x^* + \lambda e_i)}{\partial x^i}$$

for each $i$. Hence by the continuity of the partial derivatives of $g$, taking the limit as $\lambda \rightarrow 0$ gives

(6.1) $\xi A \geq \nabla g(x^*)$.

(2) If $x^* \neq 0$ and $x^*_i = 0$ for $i \in I$, $I$ a subset of $\{1, 2, \ldots, n\}$, then for $i \not\in I$ and for $\lambda > 0$ small enough, $x^* - \lambda e_i \geq 0$. Then

$$(\xi A - \nabla g(x^* - \lambda e_i), -\lambda e_i) \geq 0$$

implies $(\xi A)^i - \frac{\partial g(x^* - \lambda e_i)}{\partial x^i} \leq 0$ for $i \not\in I$. Hence by continuity of the derivatives,

$$(\xi A)^i \leq \frac{\partial g(x^*)}{\partial x^i}$$

for $i \not\in I$.

Now by the equation (6.1) the opposite inequality holds, hence $(\xi A)^i = \frac{\partial g(x^*)}{\partial x^i}$ for $i \not\in I$. Therefore
\[(6.2)\] \[\xi Ax^* = (\nabla g(x^*), x^*).\]

**Lemma 6.4:** Let \(g\) be a closed concave function defined on the set \(D = \{x \in \mathbb{R}^n : x \geq 0\}\), with continuous first partial derivatives on \(D\). Let \(Y\) be the set defined in lemma (6.3) and let \(\xi \in Y\) be fixed. Let \(W = \{x : \xi A \geq \nabla g(x), x \geq 0\}\). Let \(x^* > 0\) be such that

\[h(\xi) = \min_{x \geq 0}[\xi Ax - g(x)] = \xi Ax^* - g(x^*).\]

Then

\[-h(\xi) = \min_{x \in W}[g(x) - (\nabla g(x), x)].\]

**Proof:** By lemma (6.3), \(x^* \in W\) and \(\xi Ax^* = (\nabla g(x^*), x^*)\); hence

\[
\min_{x \in W}[g(x) - (\nabla g(x), x)] \leq g(x^*) - (\nabla g(x^*), x^*) = g(x^*) - \xi Ax^* = -h(\xi).
\]

Also for \(x \in W\), \(g(x) - (\nabla g(x), x) \geq g(x) - \xi Ax\); hence

\[
\min_{x \in W}[g(x) - (\nabla g(x), x)] \geq \min_{x \in W}[g(x) - \xi Ax] \geq \min_{x > 0}[g(x) - \xi Ax] = g(x^*) - \xi Ax^* = -h(\xi).
\]

Since both inequalities hold, there is equality and

\[-h(\xi) = \min_{x \in W}[g(x) - (\nabla g(x), x)].\]

The comparison with Dorn's results is contained in the following lemma.
Lemma 6.5: The convex programming problems I and II reduce to the programming problems 1 and 2 if the sets C and D and the functions f and g are as in definition (6.1), and if
\[ \inf_{x \geq 0} [\xi Ax - g(x)] \]
is attained for all \( \xi \in \Delta \).

Proof:
(1) The convex programming problem I clearly reduces to problem 1 under the conditions stated.

(2) As in the proof of lemma (6.1) it is clear that
\[
\Gamma = \{ \xi \in \mathbb{R}^n : \xi \geq 0 \}
\]
\[
\varphi(\xi) = \sup_{x} [\xi Ax : Ax \leq b]
\]
\[
= (\xi, b) \quad \text{for} \quad \xi \in \Gamma,
\]
since the rank of \( A \) is \( m \).

(3) \( \Delta = \{ \xi \in \mathbb{R}^m : \inf_{x \geq 0} [\xi Ax - g(x)] > -\infty \} = \{ \xi \in \mathbb{R}^m : \min_{x \geq 0} [\xi Ax - g(x)] > -\infty \} \).

Since by assumption the infimum is attained at some \( x \geq 0 \), \( \Delta = \bar{Y} \) given in lemma (6.3) and (6.4), for \( \xi \in \Delta \),
\[
-\psi(\xi) = -\min_{x \geq 0} [\xi Ax - g(x)] = \min_{x \geq 0} \{ g(x) - (\nabla g(x), x) : \xi A \geq \nabla g(x) \}.
\]

Therefore the convex programming problem II;
\[
\begin{align*}
\text{minimize} & \quad (\xi, b) - \psi(\xi) \quad \text{for} \quad \xi \in \Gamma \cap \Delta \\
\text{subject to} & \quad \xi \geq 0, \quad \xi A \geq \nabla g(x), \quad x \geq 0.
\end{align*}
\]
becomes problem 2;
\[
\begin{align*}
\text{minimize} & \quad H(x, \xi) = (\xi, b) + g(x) - (\nabla g(x), x) \\
\text{subject to} & \quad \xi \geq 0, \quad \xi A \geq \nabla g(x), \quad x \geq 0.
\end{align*}
\]
BIBLIOGRAPHY


Theorem A: Let $A$ be any matrix over the real numbers. Let $\{y_i\}_{i \in I}$ be a bounded set of vectors in the range of $A$. Then there exists a set of vectors $\{x_i\}, i \in I$, which is bounded and for which

$$Ax_i = y_i \text{ for all } i \in I.$$ 

Proof: Let $A$ be $m \times n$, let $r = \text{rank } A \leq \min (m,n)$. Then $A$ can be reduced by elementary row and column operations, to a matrix $\bar{A}$,

$$\bar{A} = \begin{pmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{pmatrix}$$

where $I_r$ is a $r \times r$ identity matrix, $0_1$ is a $r \times (n-r)$ zero matrix, $0_2$ is a $(m-r) \times r$ zero matrix, $0_3$ is a $(m-r) \times (n-r)$ zero matrix.

That is, there exist non-singular matrices $P$ and $Q$, $P$ is $(m \times m)$, $Q$ is $(n \times n)$, such that $PAQ = \bar{A}$.

The system of equations

$$A \begin{pmatrix} x_i \end{pmatrix} = \begin{pmatrix} y_i \end{pmatrix}, \quad i \in I$$

is equivalent to the system

$$PAQQ^{-1} \begin{pmatrix} x_i \end{pmatrix} = P \begin{pmatrix} y_i \end{pmatrix}, \quad i \in I.$$
or

(3) \( \bar{A}z_i = w_i \)

where \( w_i = Py_i, \ z_i = Q^{-1}x_i, \ i \in I. \)

Now the set \( \{w_i\} = \{Py_i\} \) is bounded since \( P \) is non-singular and \( \{y_i\} \) is bounded. Furthermore (3) obviously has a bounded solution \( z_i \) for all \( i \in I. \) Hence taking \( x_i = Qz_i, \ i \in I, \) one has a set \( \{x_i\}, \ i \in I, \) which is bounded since \( Q \) is non-singular, and for which \( Ax_i = y_i \) for all \( i \in I. \)