Convex Transformable Programming
Problems and Invexity

by

M. A. Hanson
Department of Statistics
The Florida State University
Tallahassee, Florida 32306-3033

B. Mond
Department of Pure Mathematics
La Trobe University
Bundoora, Victoria 3083
Australia

FSU Statistics Report M715

November, 1985

The Florida State University
Department of Statistics
Tallahassee, Florida 32306-3033
1. Introduction

G. Heal [1] has considered some properties of nonconvex programming problems which can be transformed into convex programming problems by means of diffeomorphisms. He does not explicitly describe how to construct such a diffeormorphism or whether it exists in a given case, although he points out a connection with the generalized convex functions given by Ben-Tal [2].

In this paper we show that programming problems which can be transformed in this manner are a strict subset of invex programming problems. A differentiable function \( f(x) \) is invex over a set \( S \subseteq \mathbb{R}^n \) if there exists an n-dimensional function \( \eta(x, u) \) such that for all \( x \in S \)

\[
f(x) - f(u) \geq [\nabla f(u)]^t \eta(x, u). \tag{1.1}
\]

More generally, following upon a result of Ben-Israel and Mond [3], one could define a more general class of functions, without regard to differentiability, by defining \( f(x) \) to be a member of this class if there exists an n-dimensional function \( \eta(x, u) \) such that, for all \( x \in S, u \in S \) and \( 0 \leq \lambda \leq 1 \),

\[
f(u + \lambda \eta(x, u, \lambda)) \leq \lambda f(x) + (1 - \lambda)f(u). \tag{1.2}
\]

Rewriting (1.2) as

\[
f(u + \lambda \eta(x, u, \lambda)) - f(u) \leq \lambda[f(x) - f(u)]
\]

we obtain

\[
limit_{\lambda \to 0^+} \frac{1}{\lambda} [f(u) + \lambda \eta(x, u, \lambda) - f(u)] \leq f(x) - f(u).
\]
So if \( f(x) \) is differentiable the definition (1.2) reduces to (1.1).

There are the usual extensions of invex functions: A differentiable function \( f(x) \) is pseudo-invex over \( S \) if for all \( x \in S, u \in S \)

\[
[\forall f(u)]^\tau \eta(x, u) \geq 0 \Rightarrow f(x) - f(u) \geq 0,
\]

(1.3)

and \( f(x) \) is quasi-invex over \( S \) if for all \( x \in S, u \in S \)

\[
f(x) - f(u) \leq 0 \Rightarrow [\forall f(u)]^\tau \eta(x, u).
\]

(1.4)

2. Convex Transformable Functions

In considering the programming problem: Minimize \( f(x) \) subject to \( g_i(x) \leq 0, \ i = 1, 2, \ldots, m \), where \( f: \mathbb{R}^n \rightarrow \mathbb{R}^1 \) is differentiable and \( g_i: \mathbb{R}^n \rightarrow \mathbb{R}^1, \ i = 1, 2, \ldots, m \), is differentiable, Heal [1] has made the definition:

The programming problem is **convex transformable** if there exists a regular \( C^1 \) diffeomorphism \( T: \mathbb{R}^n \rightarrow \mathbb{R}^n \) such that the problem is transformed into a convex programming problem; that is \( f \) and \( g_i, \ i = 1, 2, \ldots, m \) are transformed into convex functions. In this case we shall also say that \( f \) and \( g_i \) are convex transformable in the context of this programming problem.

We will show that convex transformable functions are a sub-class of differentiable invex functions.

Considering a differentiable function \( f: \mathbb{R}^n \rightarrow \mathbb{R}^1 \) and using the regular \( C^1 \) diffeomorphism \( T \) Heal makes the transformation \( f(x) \rightarrow \hat{f}(x') \) where \( \hat{f}(x') = f(T^{-1}(x')) \) and \( x' = Tx, T^{-1}x' = x \).

Now suppose \( \hat{f}(x') \) is convex. That is
\[ \hat{f}(x') - \hat{f}(u') \geq [\nabla u \cdot \hat{f}(u')]^T [x' - u']. \]

So \( \hat{f}(Tx) - \hat{f}(Tu) \geq [\nabla u \cdot \hat{f}(Tu)]^T [Tx - Tu] \); that is

\[ f(x) - f(u) \geq [\nabla f(u)]^T J(u, u') [Tx - Tu], \]

where \( J \) is the Jacobian \( \begin{vmatrix} 2u \\ 2u' \end{vmatrix} \) of the transformation, which exists since \( T \) is \( C^1 \) diffeomorphic.

Defining \( \eta(x, u) = J(u, Tu) [Tx - Tu] \) we see that \( f \) is invex. So all convex transformable functions are invex.

Note that \( f(x) \) can be a vector in the above since all components of \( f(x) \) would be transformed by the same \( T \).

On the other hand all differentiable invex functions are not convex transformable. Consider the two-dimensional function \( f(x) = -x_1x_2 + 1 \) and the set \( S \) defined by \( -x_1x_2 \leq 0 \).

By putting \( \eta(x, u) = \begin{bmatrix} f(x) - f(u) \\ -2u_2 \\ f(x) - f(u) \\ -2u_1 \end{bmatrix} \) we see that \( f(x) \) is invex on \( S \).

(This result is also evident from Corollary 1 of [3], namely: If \( f \) has no stationary points then \( f \) is invex.) In \( \mathbb{R}^n \) the function \( -x_1x_2 + 1 \) has a critical point \( x_1 = x_2 = 0 \), but this is not in \( S \). The set consists of two disjoint regions, and it follows from the discontinuity of the values of \( (x_1, x_2) \) that there is no \( C^1 \) diffeomorphism (or even homeomorphism) that would transform \( S \) into a convex set.

So not all invex functions are convex transformable, and thus convex transformable functions are a sub-class of invex functions.
In [1] Heal shows that if \( f \) and \( g_i \), \( i = 1, 2, \ldots, m \), are all convex transformable then any critical point is a global minimum. Ben-Israel and Mond [3] note that "a function is invex if every stationary point is a global minimum." So we have the following relationships:

convex transformable \( \implies \) a critical point is a global minimum \( \implies \) invex.

To illustrate the abovementioned procedures we use Heal's example to show directly that the functions involved, which he shows to be convex transformable, are invex with respect to a common \( \eta(x, u) \).

He considers the problem

\[
\text{minimize } -[ax_1^3(\beta x_2^3 - \gamma)]^{1/2}
\]

subject to \(-A + \delta x_1^3 + \varepsilon x_2^3 \leq 0,\)

and uses the transformation \( T = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \) defined by

\[
T_1(x_1) = ax_1^3 = x_1^{\ast}
\]

\[
T_2(x_2) = \beta x_2^3 - \gamma = x_2^{\ast}.
\]

Hence \( \frac{\partial x_1^{\ast}}{\partial x_1} = 3ax_1^2; \frac{\partial x_1^{\ast}}{\partial x_2} = 0; \frac{\partial x_2^{\ast}}{\partial x_1} = 0; \frac{\partial x_2^{\ast}}{\partial x_2} = 3\beta x_2^2.\)

Therefore

\[
J(u, u^{\ast}) = \begin{bmatrix}
\frac{1}{3au_1^2} & 0 \\
0 & \frac{1}{3\beta u_2^2}
\end{bmatrix}
\]
Also
\[
Tx - Tu = \begin{bmatrix}
\frac{x_1^3 - u_1^3}{3u_1^2} \\
\frac{x_2^3 - u_2^3}{3u_2^2}
\end{bmatrix}.
\]

So
\[
\eta(x, u) = \begin{bmatrix}
\frac{x_1^3 - u_1^3}{3u_1^2} \\
\frac{x_2^3 - u_2^3}{3u_2^2}
\end{bmatrix}.
\]

Now let
\[
f_1(x) = -\{ax_1^3(\beta x_2^3 - \gamma)\}^{1/2}.
\]

We have
\[
\nabla f_1(u) = -\frac{1}{2}\{au_1^3(\beta u_2^3 - \gamma)\}^{-1/2} \begin{bmatrix}
3au_1^2(\beta x_2^3 - \gamma) \\
a u_1^3(3\beta u_2^3)
\end{bmatrix}.
\]

So \([\nabla f_1(u)]^T\eta(x, u)\)
\[
= -\frac{1}{2}\{au_1^3(\beta u_2^3 - \gamma)\}^{-1/2} \left\{ \frac{x_1^3 - u_1^3}{u_1^3} + \frac{\beta(x_2^3 - u_2^3)}{\beta u_2^3 - \gamma} \right\}
\]
\[- \frac{1}{2} \{ \alpha x_1^3 (\beta u_2^3 - \gamma) \}^{1/2} \left\{ \frac{x_1^3}{u_1^3} + \frac{\beta x_2^3 - \gamma}{\beta u_2^3 - \gamma} - 2 \right\} \]

\[
= \{ \alpha u_1^3 (\beta u_2^3 - \gamma) \}^{1/2} \left\{ \frac{a^{1/2} x_1^3}{(u_1^3)^{1/2}} (\beta u_2^3 - \gamma)^{1/2} + \frac{(au_1^3)^{1/2} (\beta x_2^3 - \gamma)^{1/2}}{(\beta u_2^3 - \gamma)^{1/2}} \right\} \]

\[
= \{ \alpha u_1^3 (\beta u_2^3 - \gamma) \}^{1/2} \left\{ \frac{ax_1^3 (\beta u_2^3 - \gamma)}{au_1^3 (\beta x_2^3 - \gamma)} \right\}^{1/2} + \left\{ \frac{au_1^3 (\beta x_2^3 - \gamma)}{ax_1^3 (\beta u_2^3 - \gamma)} \right\}^{1/2} \]

\[
\leq -f_1(u) + f_1(x),
\]

since the sum of a nonnegative number and its reciprocal is at least 2. (It is implicit in the definition of \( f_1 \) that \( ax_1^3 (\beta x_2^3 - \gamma) \) is nonnegative for all values of the argument.)

So \( f_1(x) \) is invex with respect to \( \eta(x, u) \).

Now let

\[
f_2(x) = -A + \delta x_1^3 + \epsilon x_2^3.
\]

We have

\[
\nabla f_2(u) = \begin{bmatrix}
3\delta u_1^2 \\
3\epsilon u_2^2
\end{bmatrix}.
\]
So $[\varphi \varepsilon_2 (u)]^T \eta(x, u) = [3\delta u_1^2 \ 3\varepsilon u_2^2]$

$$\begin{bmatrix}
\frac{x_1^3 - u_1^3}{3u_1^2} \\
\frac{3u_1^2}{x_2^3 - u_2^3} \\
\frac{3u_2^2}{3u_2^2}
\end{bmatrix}$$

$$= \delta(x_1^3 - u_1^3) + \varepsilon(x_2^3 - u_2^3)$$

$$= f_2(x) - f_2(u).$$

So $f_2(x)$ is invex with respect to $\eta(x, u)$.

3. $H$-Convex Functions

Heal has shown that convex transformable functions are closely related to another class of functions studied by Ben-Tal [2] which we will call $H$-convex functions.

Let $H: \mathbb{R}^n \to \mathbb{R}^n$ be a bijective function with inverse $H^{-1}$. For $x \in S$, $u \in S$, and $0 \leq \lambda \leq 1$ define the $H$-weighted mean

$$M_H((x, u), \lambda) = H^{-1}(\lambda H(x) + (1 - \lambda)H(u)).$$

Then define $f(x)$ to be $H$-convex if for all $x \in S$, $u \in S$ and $0 \leq \lambda \leq 1$,

$$M_H((x, u), \lambda) \text{ exists in } S \text{ such that}$$

$$f(M_H((x, u), \lambda)) \leq \lambda f(x) + (1 - \lambda)f(u). \quad (3.1)$$

Putting $\eta(x, u, \lambda) = \frac{M_H((x, u), \lambda) - u}{\lambda}$ we see that this is a specialization to bijective functions of the class of functions defined by (1.2).
4. References

