QUASI-CONJUGACY, QUASI-SUBGRADIENTS, AND SURROGATE DUALITY IN NONCONVEX PROGRAMMING

Dennis L. Bricker

Department of Industrial Engineering
The University of Iowa
Iowa City, Iowa 52242

Conjugate functions have played an important role in the theory of convex programming. (For example, see [4].) An analogous role in quasi-convex programming is played by quasi-conjugate functions. Conjugates relate to epigraph supports, whereas quasi-conjugates relate to level set supports and barriers; conjugate functions provide a basis for Lagrangian duality, whereas quasi-conjugate functions provide a basis for surrogate duality. In this paper, we shall briefly survey the existing theory of quasi-conjugacy and surrogate duality as developed by Greenberg and Pierskalla ([2] and [3]) as it relates to nonconvex programming, interpreting it geometrically, and shall then add several extensions to this theory.

**QUASI-CONJUGATES**

A hyperplane in \( E^n \) is a set, with parameters \( u \in E^n \), \( u \not= 0 \), and \( c \in E^1 \), of the form

\[
H^c_u = \{ x \in E^n : (u, x) = c \}
\]

(1.1)

where \((u, x)\) and \(ux\) will interchangeably denote the inner product of \( u \) and \( x \). The parameter \( u \) determines the orientation of \( H^c_u \) and may be referred to as its direction vector. In particular, the hyperplane with direction vector \( u \) passing through the fixed point \( x^0 \) is \( H^c_{u^0} \). (See Figure 1.)

A hyperplane \( H^c_u \) determines two closed halfspaces, one of which we will denote by

\[
H^c_u = \{ x \in E^n : (u, x) \geq c \}
\]

(1.2)

If \( f \) is a function from \( E^n \) into the (extended) real line, \( E^1 = [-\infty, +\infty] \), i.e., a functional, then we denote its \( c \)-level-sets by

\[
L_c f = \{ x \in E^n : f(x) \leq c \}
\]

(1.3)

and

\[
L'_c f = \{ x \in E^n : f(x) < c \}
\]

(1.4)

Figure 2 denotes \( L_c f \) for a case in which \( n=1 \).
Our interest in level sets results chiefly from the fact that quasi-convex functions may be defined to be functions all of whose level sets are convex.

For many functions, \( L_c f = \text{cl} L_c^0 f \) (the closure of \( L_c^0 f \)), but such is not always the case, as demonstrated by Figure 3. Here \( L_c^0 f \subset L_c f \), but it is neither the case that \( \text{cl} L_c^0 f = L_c f \), nor even that \( \text{cl} L_c^0 f \subset L_c f \). Note that \( f \) is not lower semi-continuous, nor explicitly quasi-convex (because of the "flat" spot in the graph).

Figure 4 depicts a \( c \)-level-set of a function defined on \( \mathbb{E}^2 \). The boundaries of level sets are simply the contour curves of the function. Given a point \( x \in \mathbb{E}^n \), a level set of particular interest is \( L_{f(x)} f \), depicted in Figures 5 a&b. In Figure 5b we note that \( x \) need not be a boundary point of \( L_{f(x)} f \).

We next define the \( z \)-quasi-conjugate function \( f^+_z : \mathbb{E}^n \rightarrow \mathbb{E}^1 \) where \( z \in \mathbb{E}^1 \) and

\[
 f^+_z (u) = z - \inf \left\{ f(x) : (u, x) \geq z \right\} \quad (1.5)
\]

Note that it is helpful to consider \( f^+_z \) as a function of direction vectors, i.e.,

\[
 f^+_z (u) = z - \inf \left\{ f(x) : x \in H^z_u \right\}. \quad (1.6)
\]

If \( f \) is a quasi-convex function, as in Figure 6, and \( H^z_u \) is a supporting hyperplane for some level set \( L_c f \), then \( f^+_z (u) = z - c \), provided that the global minimum point \( x^* = \text{argmin} f(x) \) does not lie in \( H^z_u \) (in which case \( f^+_z (u) = z - f(x^*) \)).

One important property which should be noted is that \( f^+_z \) is quasi-convex (without assuming any properties of \( f \)).

We now consider the second \( z \)-quasi-conjugate \( (f^+_z)^+_z (x) \), defined in the obvious way as the \( z \)-quasi-conjugate of \( f^+_z \), and define the normalized second quasi-conjugate of \( f \) as

\[
 f^{++}(x) = \sup_{z \in \mathbb{E}^1} (f^+_z)^+_z (x) \quad (1.7)
\]

Example:

Let \( f(x) = -e^{-x^2} = -\exp(-x^2) \), for \( x \in \mathbb{E}^1 \) (see Figure 7a). Note that \( f \) is quasi-convex. Then
\[ f^*_z(u) = z - \inf \{ f(x) : ux \geq z \} = z - \inf \{ -\exp(-x^2) : ux \geq z \} \]
\[
\begin{cases} 
  z + \exp\left( -\frac{z^2}{u^2} \right) & \text{if } z > 0 \text{ and } u \neq 0 \\
  -\infty & \text{if } z > 0 \text{ and } u = 0 \\
  z + 1 & \text{if } z \leq 0 
\end{cases}
\]

The second z-quasi-conjugate is
\[
(f^*_z)^+ = z - \inf \{ f^*_z(u) : ux \geq z \} = z - (z+1) \text{ if } z \leq 0 \\
\quad z - \infty \text{ if } z > 0 \text{ and } x = 0 \\
\quad z - \inf \left\{ z + \exp\left( -\frac{z^2}{u^2} \right) : \frac{z}{u} \leq 0 \right\} \text{ if } z > 0 \text{ and } x < 0 \\
\quad z - \inf \left\{ z + \exp\left( -\frac{z^2}{u^2} \right) : \frac{z}{u} \geq x \right\} \text{ if } z > 0 \text{ and } x < 0 \\
\quad -1 \text{ if } z \leq 0 \\
\quad -\infty \text{ if } z > 0 \text{ and } x = 0 \\
\quad -\exp(-x^2) \text{ if } x > 0 \text{ and } x \neq 0
\]

And hence
\[
f^{++}(x) = \sup (f^*_z)^+(x) = \begin{cases} 
  \max \{ -1, -\infty \} & \text{if } x = 0 \\
  \max \{ -1, -\exp(-x^2) \} & \text{if } x \neq 0 
\end{cases} = -\exp(-x^2) = f(x)
\]

(See Figure 7b.)

The function \( f^{++} \) has several important properties:

**Property (i):** (see [2]): \( f^{++} \) is quasi-convex, and
\[
f(x) \geq f^{++}(x) \geq f^{\vee\vee}(x),
\]
where \( f^{\vee\vee} \) is the second (convex) conjugate. That is, \( f^{++} \) provides a quasi-convex approximation to \( f \), from below, which is better than the convex approximation provided by \( f^{\vee\vee} \).
PROPERTY (ii) (see [2]):

\[ f^{**}(x) = \sup_u \inf_w \{ f(w) : (u, w) \succeq (u, x) \} \]

\[ = \sup_u \inf_w \left\{ f(w) : w \in H_u^{ax} \right\} \] (1.8)

This relaxation is easier to interpret geometrically than the definition of \( f^{**} \). In Figure 8, we see depicted the three hyperplanes through \( x \) with the direction vectors \( u^2, u^1, \) and \( u^0 \), together with the points

\[ \hat{w} = \arg \min_w \left\{ f(w) : w \in H_u^x \right\} \]

It is clear that for the function depicted, rotating a hyperplane clockwise from \( H_u^{ax} \) through \( H_u^x \) to \( H_u^{0x} \) (which supports the contour curve through \( x \) of the function \( f \)) produces a maximizing sequence \( \{ \hat{w}_i \} \) converging to \( x \), and \( f^{**}(x) = f(x) \).

Figure 9 depicts a function which (unlike that in Figure 8) is not quasi-convex. Again, rotating a hyperplane clockwise from \( H_u^{ax} \) to \( H_u^{0x} \) produces a maximizing sequence \( \{ \hat{w}_i \} \), which does not, however, converge to \( x \). Moreover,

\[ f \left( w^{*0} \right) = f \left( w^0 \right) = f^{**}(x) < f(x). \]

Furthermore, it is shown in [2] that if \( f \) is an isotonic function, i.e.,

\[ w \succeq v \Rightarrow f(w) \succeq f(v), \]

the optimal \( u \) in equation (1.8) has the property \( u \in E^x_c \), i.e., \( u \geq 0 \). Hence, if \( f \) is isotonic,

\[ f^{**}(x) = \sup_{u \in E^x_c} \inf_w \left\{ f(w) : uw \succeq ux \right\} \] (1.9)

PROPERTY (iii): If \( L_c f \) is compact for all \( c \), then

\[ L_c f^{**} = \text{conv} L_c f \]

for all \( c \) (cf. [3]). More generally, for all \( c \),

\[ L_c f^{**} \supset \text{cl conv} L_c f \]

and

\[ L_c f^{**} \supset \text{conv} L_c f. \] (1.10)
Proof: The proof of (1.10) is a trivial result of property (i). Let \( x \notin \text{cl conv } L_c f \). Then \( x \) may be separated from \( \text{cl conv } L_c f \), i.e., there is a \( y \) such that \( xy > wy \) for all \( w \in \text{cl conv } L_c f \). By Property (ii),

\[
f^{++} (x) = \sup_{u} \inf_{w} \{ f(w) : wu \geq xu \}
\]

and so, in particular,

\[
f^{++} (x) \geq K
\]

where

\[
K = \inf_{w} \{ f(w) : wy \geq xy \}.
\]

Now, given \( \delta > 0 \), there must exist \( w_\delta \) such that \( w_\delta y \geq xy \) and \( f(w_\delta) < K + \delta \). But \( w_\delta y \geq xy \) implies that \( w_\delta \notin \text{conv } L_c f \) and hence \( w_\delta \notin L_c f \), i.e., \( f(w_\delta) > c \). Therefore, we have, for all \( \delta > 0 \),

\[
c < f(w_\delta) < K + \delta \leq f^{++} (x) + \delta
\]

or simply \( c - \delta < f^{++} (x) \) for all \( \delta > 0 \). Therefore, \( c \leq f^{++} (x) \) and so \( x \notin L_c f^{++}, \) proving that

\[
L_c f^{++} \subset \text{cl conv } L_c f.
\]

We are now in a position to introduce the concept of surrogate mathematical programming.
SURROGATE MATHEMATICAL PROGRAMMING

Consider the family of mathematical programs obtained by parameterizing the constraint right-hand-side vector and whose optimal value $F:E^m \to \mathbb{R}$ is defined by

\[
F(b) = \inf \left\{ f(x) : g(x) \geq b, x \in S \right\}
\]

where $f:S \to E^1$, $S \subseteq E^n$, $g:S \to E^m$, and $b \in E^m$. (If the problem is infeasible, then we define $F(b) = +\infty$.)

Note that if $b^1 \geq b^2$, then

\[
\left\{ x : g(x) \geq b^1, x \in S \right\} \subseteq \left\{ x : g(x) \geq b^2, x \in S \right\},
\]

and so

\[
\inf \left\{ x : g(x) \geq b^1, x \in S \right\} \geq \inf \left\{ x : g(x) \geq b^2, x \in S \right\},
\]

i.e., $F(b^1) \geq F(b^2)$. Thus $F$ is isotonic.

A surrogate problem, parameterized by $b$ and the surrogate multiplier vector $u \in E^m_+$, is defined to be that of computing

\[
S(u, b) = \inf \left\{ f(x) : u g(x) \geq ub, x \in S \right\}.
\]

This is equivalent to

\[
S(u, b) = \inf \left\{ F(\beta) : u \beta \geq ub, x \in S \right\}.
\]

We further define the surrogate dual problem to be that of computing

\[
\hat{S}(b) = \sup_{u \geq 0} S(b, u)
\]

\[
= \sup_{u \geq 0} \inf_{x} \left\{ f(x) : u g(x) \geq ub, x \in S \right\}
\]

\[
= \sup_{u \geq 0} \inf \left\{ F(\beta) : u \beta \geq ub \right\}.
\]

Without affecting the supremum we may perform the outer optimization over the subset of surrogate multipliers

\[
\mathcal{U} = \left\{ u \in E_+^m : \sum_{i=1}^m u_i = 1 \right\}
\]

which is both convex and compact. Any direction in $E_+^m$ has a representative vector in $\mathcal{U}$. We may then write
\[ \hat{S}(b) = \sup_{u \in U} S(b, u). \]

Comparison of (2.4) with equation (1.9) shows that, since F is isotonic,
\[ \hat{S}(b) = F^{++}(b). \tag{2.5} \]

We know that \( F^{++}(b) \leq F(b) \), and we are naturally interested in knowing under what conditions equality holds. That is, when does there exist a \( u \geq 0 \) such that solving the surrogate problem \( S(b, u) \) solves our original problem, and \( S(u, b) = F(b) \)? If such a \( u \) does not exist, \( b \) is said to lie in a surrogate gap. The point \( b^0 \) is in such a gap in Figure 11, where
\[ F^{++}(b) = F(b^i) < F(b^0). \]

This figure also illustrates one of the results stated in [2]. Suppose, for some \( u^* \geq 0, b^0 \) is a convex combination of points in the set \( \arg\min\{F(\beta) : u\beta \geq ub^0\} \). Then either some solution \( x \) of the surrogate problem \( S(b^0, u^*) \) is a solution of \( F(b^0) \), or else \( b^0 \) is in a surrogate gap.

The quasi-subgradient, to be introduced next, will help to characterize the surrogate gaps of a mathematical program.
The conjugate inequality [4], namely,
\[(x, y) \leq f(x) + f^\vee(y),\]
with equality if and only if \(y \in \partial f(x)\), where \(\partial f(x)\) is the subgradient set of \(f\) at the point \(x\), and \(f^\vee\) is the convex conjugate of the function \(f\) has an analogue in quasi-conjugate theory. It is easy to derive the result
\[(u, x) \leq f(x) + f^\vee_{ax}(u)\]  \tag{3.1}
and we shall define \(\partial^+ f(x)\), the set of quasi-subgradients of \(f\) at \(x\), to be those vectors \(u\) such that equality holds in (3.1), i.e.,
\[(u, x) = f(x) + f^\vee_{ax}(u),\]
with equality if and only if \(u \in \partial^+ f(x)\).

Equivalently,
\[
\partial^+ f(x) = \left\{ u : (u, w) \geq (u, x) \implies f(w) \geq f(x) \right\}
= \left\{ u : w \in H^ax \implies f(w) \geq f(x) \right\}
= \left\{ u : f(w) < f(x) \implies w \notin H^ax \right\}
= \left\{ u : L^u_{f(x)} \cap H^ax = \emptyset \right\}.
\]
That is, \(u\) is a quasi-subgradient of \(f\) at \(x\) if \(L^u_{f(x)}\) lies entirely on one side of the hyperplane through \(x\) with direction vector \(u\), or equivalently, \(H^ax\) is a non-intersecting barrier of \(L^u_{f(x)}f\). (\(H^ax\) is a barrier for a set \(S\) if
\[
\sup_{x \in S} (u, x) \leq z.
\]
In many cases (e.g., as we shall see, when \(f\) is continuous and convex or explicitly quasi-convex), there is a one-to-one correspondence between quasi-subgradients and level set supports (see Figure 12). (This assumes, of course, that the vectors in \(\partial^+ f(x)\) are normalized in some manner, since any multiple of a quasi-subgradient is also a quasi-subgradient.) However, Figure 13 depicts quasi-subgradients which do not produce
corresponding level set supports. Any \( u \) which is a convex combination of \( u^0 \) and \( u^1 \) is a quasi-subgradient in Figure 13.

To see that level set supports, conversely, do not necessarily correspond to quasi-subgradients, consider the function \( f : E^2 \) defined by
\[
f(x_1, x_2) = \begin{cases} 
0 & \text{if } x_1 + x_2 < 1, \text{ or } x_1 + x_2 = 1 & x_i \geq 0.5 \\
x_1 + x_2 & \text{otherwise } (x_i \geq 0 & x_i \geq 0) 
\end{cases}
\]
whose graph and level sets are illustrated in Figure 14 J&B. The set \( L_f \) is supported at the point \( x=(0.5, 0.5) \) by the hyperplane \( x_1 + x_2 = 1 \) (i.e., \( H_{(1,1)}^1 \)) but unfortunately \( L_f \) has a nonempty intersection with this hyperplane, and so \( u=(1,1) \) is not a quasi-subgradient of \( f \) at \( x=(0.5, 0.5) \).

The correspondence between level set supports and quasi-subgradients failed for the function in Figure 13 because \( \text{cl } L_f \neq L_f \), while the failure for the function in Figure 14 results from the fact that \( L_f \) contained boundary points. In general, if
\[
\text{cl } L_f = L_f \Rightarrow \text{ supports } L_f \text{ at } x. \text{ Conversely, if } L_f \text{ is open, then } H_u \text{ is a barrier (or support) for } L_f \text{ at } x \Rightarrow u \in \partial f(x).
\]

The importance of the quasi-subgradient derives mainly from the following properties:

(i) \( 0 \in \partial^+ f(x) \iff x \in \text{argmin } f(x) \)

(ii) \( \partial^+ f(x) \neq \Phi \Rightarrow f(x) = f^{++}(x) \)

Thus our question "does \( b^0 \) lie in a surrogate gap?" is equivalent to the question "does \( F \) have a quasi-subgradient at \( b^0 \)?". Toward answering this question, we may use the following sufficient conditions, the proofs of which are very straightforward. (Note that any support is a barrier, but not conversely.)
(i) If \( L'_{f(x)}f \) is a non-empty open set, and if \( H^{xx}_{u} \) is a barrier for \( L'_{f(x)}f \), then \( u \in \partial^+ f(x) \).

(ii) If \( L'_{f(x)}f \) is non-empty and \( f \) is upper semi-continuous on some set containing \( L'_{f(x)}f \), and if \( H^{xx}_{u} \) is a barrier for \( L'_{f(x)}f \), then \( u \in \partial^+ f(x) \).

(iii) If \( L'_{f(x)}f \) is non-empty and \( f \) is upper semi-continuous on some set containing \( L'_{f(x)}f \), and if \( H^{xx}_{u} \) supports \( L_{f(x)}f \), then \( u \in \partial^+ f(x) \).

(iv) If \( f \) is quasi-convex and \( x \notin \text{cl} \ L'_{f(x)}f \), then \( \partial^+ f(x) \) is non-empty.

(v) If \( f \) is a quasi-convex function which is upper semi-continuous on \( L_{f(x)}f \) for some \( x \), then \( \partial^+ f(x) \) is non-empty.

**Examples**

The following examples will help to illustrate the concepts which have been presented.

**Example 1.** Consider the problem

\[
\text{Minimize } f(x) = x_1^2 + x_2^2
\]

subject to

\[
x_1 + x_2 \geq 1 = b_1^0
\]
\[
x_1 - x_2 \geq 1 = b_2^0
\]

Our optimal response function, \( F(b) \), is

\[
F(b_1, b_2) = \min \left\{ x_1^2 + x_2^2 : x_1 + x_2 \geq b_1, x_1 - x_2 \geq b_2 \right\}
\]

\[
= \begin{cases} 
0.5(b_1^2 + b_2^2) & \text{if } b_1 \geq 0, b_2 \geq 0 \\
0.5b_1^2 & \text{if } b_1 > 0, b_2 < 0 \\
0.5b_2^2 & \text{if } b_1 < 0, b_2 > 0 \\
0 & \text{if } b_1 < 0, b_2 < 0
\end{cases}
\]

as can be seen graphically (see Figure 15a). Its contours are depicted in Figure 15b.
\[ U = \{(u_1, u_2) : u_1 + u_2 = 1, u_i \geq 0, u_2 \geq 0\} \]

is
\[
S(b^0, u) = \inf \left\{ \sum_{i=1}^{m} x^2_i + \sum_{j=1}^{n} x_j : \sum_{i=1}^{m} x^2_i (x_i + x_j) + \sum_{j=1}^{n} x_j (u_i + u_j) \geq 1 \right\}
\]
\[
= \inf \left\{ \sum_{i=1}^{m} x^2_i + \sum_{j=1}^{n} x_j : \sum_{i=1}^{m} x^2_i (x_i + x_j) + \sum_{j=1}^{n} x_j (u_i - u_j) \geq 1 \right\}
\]
\[
= \inf \left\{ \sum_{i=1}^{m} x^2_i + \sum_{j=1}^{n} x_j : \sum_{i=1}^{m} x^2_i + \sum_{j=1}^{n} x_j (2u_i - 1) \geq 1 \right\}
\]
which has the solution (see Figure 15c):
\[
S(b^0, u) = \frac{1}{1 + (2u_i - 1)^2} \quad \text{for } 0 \leq u_i \leq 1, u_i = 1 - u_i
\]

The surrogate dual is therefore
\[
\hat{S}(b) = \sup_{u \in U} S(b, u)
\]
\[
= S(b^0, u^0), \quad \text{where } u^0 = (0.5, 0.5)
\]
\[= 1.\]

Thus \( b^0 = (1, 1) \) is not in a surrogate gap, since
\[
\hat{S}(b^0) = F(b^0) = 1
\]

and it is evident from Figure 15 that \( F \) has no surrogate gaps whatsoever.


Our next example illustrates the existence of surrogate gaps.

**Example 2.**

Consider the problem

\[
\text{Minimize} \quad x_1 + x_2
\]

subject to
\[
x_1 + 2x_2 \geq 4 = b_1^0
\]
\[
2x_1 + x_2 \geq 3 = b_2^0
\]
\[x_1 \text{ and } x_2 \text{ both nonnegative and integer}\]

The graph of our optimal response function, \( F \), is sketched in Figure 16a and its contours are shown in Figure 16b. Note that \( F(b) \) is both isotonic and lower semi-continuous everywhere, but clearly is not quasi-convex.
The surrogate problem with parameter \( u \in U \), where (as before),

\[
U = \{(u_1, u_2) : u_1 + u_2 = 1, u_1 \geq 0, u_2 \geq 0\}
\]

is

\[
S(\beta^0, u) = \min_{x \in \{0,1,2\ldots\}} \{ x_1 + x_2 : u_1 (x_1 + 2x_2) + u_2 (2x_1 + x_2) \geq 4u_1 + 3u_2 \}
\]

\[
= \min_{x \in \{0,1,2\ldots\}} \{ x_1 + x_2 : x_1 (1+u_1) + x_2 (1+u_2) \geq 3 + u_1 \}
\]

which has the solution

\[
S(\beta^0, u) = \begin{cases} 
\left\lceil \frac{(3+u_1)}{(1+u_1)} \right\rceil = 1 + \left\lceil \frac{2}{(1+u_1)} \right\rceil & \text{if } u_1 \geq u_2, \text{ i.e., } 0.5 \leq u_1 \leq 1 \\
\left\lceil \frac{(3+u_1)}{(1+u_2)} \right\rceil = \left\lceil \frac{5}{2-2u_1} \right\rceil & \text{if } u_1 < u_2, \text{ i.e., } 0 \leq u_1 < 0.5
\end{cases}
\]

where \( \lceil z \rceil \) denotes the smallest integer greater than or equal to \( z \). (That this is the solution may be seen in Figure 16c: the minimum will always be attained at a point on a coordinate axis.) This solution is graphed as a function of \( u \) in Figure 16d.

The surrogate dual is \( \hat{S}(\beta^0) = \sup_{u \in U} S(\beta^0, u) \) and its solution, obtained from Figure 16d, is \( \hat{S}(\beta^0) = 3 \), and

\[
\arg\min_u S(\beta^0, u) = \{u \in U : \frac{1}{3} \leq u_1 \leq 1\}
\]

We see, therefore, that \( b^0 = (4,3) \) is not in a surrogate gap, since from Figure 16b,

\[
F(4,3) = 3,
\]

It follows then that any optimal multiplier \( u \) is a quasi-subgradient, so

\[
\partial^+ F (4,3) = \{u : \frac{1}{3} \leq u_1 \leq 1, u_2 = 1-u_1\}.
\]

An examination of Figure 16e confirms this; any direction between \( u^1 = \left(\frac{\sqrt{3}}{3}, \frac{2}{\sqrt{3}}\right) \) and \( u^2 = (1,0) \) is a barrier of \( L_x^2 F = L_z^2 F \). (It was demonstrated in [2] that \( b^0 = (4,3) \) is in a GLM (generalized Lagrangian multiplier) duality gap. This is evident from Figure 16a: the epigraph of \( F(b) \) has supports only at the points indicated in Figure 16f, and all other points must be in a GLM duality gap.)
We might now ask, "does our $F$ have any surrogate gaps?". Further inspection indicates that the areas indicated in Figure 16g, for example, are surrogate gap regions. That is, the triangular area
\[
\{(b_1, b_2) : b_1 > 3, b_2 < 2, b_1 + b_2 \leq 6\}
\]
is a surrogate gap region. For any point $b$ in these regions, we cannot construct a hyperplane which acts as a non-intersecting barrier of $L^*_{\{i\}} F$.

An important relationship is illustrated here, namely, that surrogate gaps form a subset of the GLM duality gaps, i.e., if $b^0$ is in a surrogate gap, so that no surrogate multiplier vector $u \geq 0$ can be found such that
\[
S(b^0, u) = F(b^0),
\]
then it is also true that no GLM multiplier vector $u \geq 0$ may be found such that
\[
\min_x \left\{ f(x) + u \left[ b^0 - g(x) \right] \right\} = F(b^0).
\]

**Summary**

We have seen that quasi-conjugacy and the quasi-subgradient provide a basis for interpreting surrogate duality, much as conjugacy and the subgradient provide a basis for understanding Lagrangian duality.

While the Lagrangian dual has gaps when $F$ is not convex, i.e.,
\[
F^{\vee\vee}(b) < F(b),
\]
the surrogate dual has a reduced gap region, as a consequence of the property
\[
F^{\vee\vee}(b) \leq F^{++}(b) \leq F(b).
\]
That is, $F^{++}$ provides a better approximation to $F$ than does $F^{\vee\vee}$.

A much more complete discussion of the relationship between the surrogate and Lagrangian dual may be found in [2]. Other important properties of the quasi-conjugates and quasi-subgradients are reported in [3].

**References**


Figure 1. A hyperplane with direction vector $u$ through $x^0$.

Figure 2. The $c$-level set of the function $f$. 
Figure 3. An example illustrating $L_c f \neq \text{cl } L_c^* f$.

Figure 4. A c-level set of a function defined on $E^2$. 

Quasi-conjugacy & Surrogate Duality
Figure 5. The level set $L_{f(x)}f$ corresponding to a point $x$. 
Figure 6. The hyperplane $H_u^z$ corresponding to $f^+_z (u)$.

Figure 7a. The function $f(x) = -\exp(-x^2)$.
Figure 7b. Graphs of selected z-quasi-conjugates of $f(x) = -\exp(-x^2)$

Figure 8. Geometric interpretation of $f^{++}$ (where $f$ is quasi-convex)
Figure 9. Geometric interpretation of $f^{++}$ (where $f$ is not quasi-convex).

Figure 10. Level curves of an isotonic function
Figure 11. Illustration of a surrogate gap (at $b^0$).

Figure 12. The hyperplane $H_u^{ux}$ corresponding to quasi-subgradient $u$ of the function $f$ is a support of the level set $L_{f(x)}f$ (where $f$ is explicitly quasi-convex).
Figure 13. The quasi-subgradient set of \( f \) is the convex hull of \( u^0 \) and \( u^1 \), which do not correspond to supports of the level set \( L_{f(x)} f \).
Figure 14. The graph (a) and the 1-level set (b) of an example function $f$
Figure 15. Example 1: (a) graphical solution; (b) contours of optimal response function $F$; (c) graphical solution of surrogate problem.
Figure 16. Example 2: (a) graph of optimal response function $F$; (b) contours of optimal response function $F$
Figure 16 (continued). Example 2: (c) graphical solution of surrogate problem; (d) graphical solution of surrogate dual problem; (e) the quasi-subgradient set of $F$ at $b^0$
Figure 16 (continued). Example 2: (f) Lagrangian duality gap region; (g) surrogate duality gap regions