A duality principle for non-convex optimization in $\mathbb{R}^n$

Fabio Silva Botelho
Departamento de Matemática
Universidade Federal de Santa Catarina, UFSC
Florianópolis, SC - Brazil

April 2, 2019

Abstract

This article develops a duality principle for a class of optimization problems in $\mathbb{R}^n$. The results are obtained based on standard tools of convex analysis and on a well known result of Toland for D.C. optimization. Global sufficient optimality conditions are also presented as well as relations between the critical points of the primal and dual formulations. Finally we formally prove there is no duality gap between the primal and dual formulations in a local extremal context.

1 Introduction

Consider a function $J : \mathbb{R}^n \to \mathbb{R}$ defined by

$$J(x) = -G_1(x) + G_2(x, 0),$$

where

$$G_1(x) = -\frac{x^T Ax}{2} + \frac{K}{2} x^T x - f^T x$$

and

$$G_2(x, v) = \sum_{j=1}^N \frac{\gamma_j}{2} \left( \frac{x^T B_j x}{2} + c_j + v_j \right)^2 + \frac{K}{2} x^T x,$$

and where $x \in \mathbb{R}^n$, $v \in \mathbb{R}^N$, $A$ is a $n \times n$ real symmetric matrix, $B_j$ is a $n \times n$ real symmetric matrix and $c_j, \gamma_j \in \mathbb{R}$, where $\gamma_j > 0$, $\forall j \in \{1, \ldots, N\}$.

Finally, $f \in \mathbb{R}^n$ as well.

Observe that

$$J(x) = \frac{x^T Ax}{2} + \sum_{j=1}^N \frac{\gamma_j}{2} \left( \frac{x^T B_j x}{2} + c_j \right)^2 + f^T x.$$
We shall develop a duality principle which has no restriction concerning \( n \) and \( N \), so that it includes the case \( n \neq N \).

Also, we establish a relation between the corresponding critical points of the primal and dual formulations.

The main result is established through an extension of a Toland result found in [7].

Indeed, we must emphasize our work is a kind of extension and continuation of the original works of Bielski and Telega [1, 2] combined with the work of Toland [7]. The technical details follow in some extent the results in [3]. Anyway, we highlight once more our work in some sense complements the results in [1, 2] but now applied to a \( \mathbb{R}^n \) simpler context.

Similar problems have been addressed in [5, 6], among others.

2 The main result

We start this section with a remark.

Remark 2.1. About the notation we denote the canonical basis of \( \mathbb{R}^n \) by

\[
\{e_1, \ldots, e_n\}
\]

and we recall that in general \( A^T \) denotes the transpose of the matrix \( A \). For a \( n \times n \) matrix \( A \) we denote \( A \succ 0 \) if \( A \) is positive definite. Finally, \( I_d \) denotes the identity matrix \( n \times n \) and by \( \{\delta_{ij}\} \) we denote the standard \( N \times N \) Kronecker delta, that is,

\[
\delta_{ij} = \begin{cases} 
1, & \text{if } i = j \\
0, & \text{otherwise}
\end{cases}
\]

\( \forall i, j \in \{1, \ldots, N\} \).

Our main result is summarized by the following theorem.

**Theorem 2.2.** Let \( J : \mathbb{R}^n \to \mathbb{R} \) be defined by

\[
J(x) = \frac{x^T Ax}{2} + \sum_{j=1}^{N} \frac{\gamma_j}{2} \left( \frac{x^T B_j x}{2} + c_j \right)^2 + f^T x
\]

\[= -G_1(x) + G_2(x, 0) \tag{2}\]

where

\[
G_1(x) = -\frac{x^T Ax}{2} + \frac{K}{2} x^T x - f^T x
\]

and

\[
G_2(x, v) = \sum_{j=1}^{N} \frac{\gamma_j}{2} \left( \frac{x^T B_j x}{2} + c_j + v_j \right)^2 + \frac{K}{2} x^T x.
\]

Assume \( A \) is a \( n \times n \) symmetric matrix and \( B_j \) are \( n \times n \) symmetric matrices \( \forall j \in \{1, \ldots, N\} \) such that

\[J(x) \to +\infty\]

as \( |x| \to \infty \), and \( K > 0 \) is such that \( K I_d > A \).
Define also \( G^*_1 : \mathbb{R}^n \to \mathbb{R} \) by
\[
G^*_1(v^*) = \sup_{x \in \mathbb{R}^n} \{(v^*)^T x - G_1(x)\}
= \frac{1}{2} (v^* + f)^T (K I_d - A)^{-1} (v^* + f) \tag{3}
\]
and \( G^*_2 : \mathbb{R}^n \times C^* \to \mathbb{R} \) by
\[
G^*_2(v^*, v^*_0) = \sup_{(x, v) \in \mathbb{R}^n \times \mathbb{R}^N} \{(v^*)^T x + (v^*_0)^T v - G_2(x, v)\}
= \frac{1}{2} (v^*)^T \left( \sum_{j=1}^N (v^*_0)_j B_j + K I_d \right)^{-1} v^* + \sum_{j=1}^N \frac{1}{2 \gamma_j} (v^*_0)_j^2
- \sum_{j=1}^N c_j (v^*_0)_j \tag{4}
\]
where
\[
C^* = \left\{ v^*_0 \in \mathbb{R}^N : \sum_{j=1}^N (v^*_0)_j B_j + K I_d > 0 \right\}.
\]
Moreover, define
\[
B^* = \left\{ v^*_0 \in \mathbb{R}^N : A + \sum_{j=1}^N (v^*_0)_j B_j > 0 \right\}
\]
and
\[
A^* = B^* \cap C^*.
\]
At this point we denote
\[
J^*(v^*, v^*_0) = G_1(v^*) - G^*_2(v^*, v^*_0),
\]
and define
\[
\tilde{J}^*(v^*) = \sup_{v^*_0 \in C^*} J^*(v^*, v^*_0).
\]
Assume \( x_0 \in \mathbb{R}^n \) is such that \( \delta J(x_0) = 0 \) and define
\[
(\tilde{v}^*_0)_j = \gamma_j \left( \frac{x^T_0 B_j x_0}{2} + c_j \right),
\]
\[
\hat{v}^* = \sum_{j=1}^N (\tilde{v}^*_0)_j B_j x_0 + K x_0,
\]
\[
H_3 = P_1 \bar{E} P_2,
\]
\[
\alpha \equiv (\alpha)_{n \times n} = (I_d - H_3) D - I_d,
\]
\[ \alpha_1 = - \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + K I_d \right)^{-1} \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + K I_d \right) \]

where

\[ P_1 = \begin{bmatrix} B_1 x_0 & B_2 x_0 & \cdots & B_N x_0 \end{bmatrix}_{n \times N} \quad (5) \]

and

\[ P_2 = \begin{bmatrix} x_0^T B_1 \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + K I_d \right)^{-1} \\ x_0^T B_2 \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + K I_d \right)^{-1} \\ \vdots \\ x_0^T B_N \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + K I_d \right)^{-1} \end{bmatrix}_{N \times n} \quad (6) \]

where

\[ E = \{ E_{lq} \} = \gamma_l \left( x_0^T B_l \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + K I_d \right)^{-1} B_n x_0 \right) + \delta_{lq} \]

and

\[ \overline{E} = \{ \overline{E}_{lq} \} = \{ E_{lq} \}^{-1}. \]

Furthermore,

\[ D = \hat{B} \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + K I_d \right)^{-1} + I_d \]

where

\[ \hat{B}_{n \times n} = \{ \hat{B}_{jk} \} = \left\{ \sum_{l=1}^{N} \sum_{s,q=1}^{n} \gamma_l (x_0)_s (B_l)_{js} (B_l)_{qk} (x_0)_q \right\}. \]

Under such assumptions and notation, we have,

1. If \( \delta^2 J(x_0) > 0 \), \( \delta^2 J(x_0) + (K I_d - A)(\alpha_1) > 0 \) and \( \hat{v}_0^* \in C^* \), then

\[ \delta \hat{J}(\hat{v}^*) = 0, \]

and

\[ \delta^2 \hat{J}(\hat{v}^*) > 0, \]

so that there exist \( r > 0 \) and \( r_1 > 0 \) such that

\[ J(x_0) = \inf_{x \in B_r(x_0)} J(x) \]

\[ = \inf_{v^* \in B_{r_1}(\hat{v}^*)} \hat{J}^*(v^*) \]

\[ = \hat{J}^*(\hat{v}^*) \]

\[ = \inf_{v^* \in B_{r_1}(\hat{v}^*)} \sup_{v_0^* \in C^*} J^*(v^*, v_0^*) \]

\[ = J^*(\hat{v}^*, \hat{v}_0^*). \]
2. If \( \hat{v}_0^* \in A^* \) so that
\[
\delta^2 J(x_0) > 0,
\]
define
\[
J_2^*(v^*) = \sup_{v_0^* \in A^*} J^*(v^*, v_0^*).
\]
Thus in such a case, we have
\[
\delta J_2^*(\hat{v}^*) = 0,
\]
\[
\delta^2 J_2^*(\hat{v}^*) > 0,
\]
and
\[
J(x_0) = \inf_{x \in \mathbb{R}^n} J(x)
= \inf_{v^* \in \mathbb{R}^n} J_2^*(v^*)
= J_2^*(\hat{v}^*)
= \inf_{v^* \in \mathbb{R}^n} \sup_{v_0^* \in A^*} J^*(v^*, v_0^*)
= J^*(\hat{v}^*, \hat{v}_0^*). \tag{8}
\]

3. If \( \delta^2 J(x_0) < 0, \delta^2 J(x_0) + (KI_d - A)(\alpha_1) < 0 \) and \( \hat{v}_0^* \in C^* \) then
\[
\delta \bar{J}(\hat{v}^*) = 0
\]
and
\[
\delta^2 \bar{J}^*(\hat{v}^*) < 0,
\]
so that there exist \( r > 0 \) and \( r_1 > 0 \) such that
\[
J(x_0) = \sup_{x \in B_r(x_0)} J(x)
= \sup_{v^* \in B_{r_1}(\hat{v}^*)} J^*(v^*)
= J^*(\hat{v}^*)
= \sup_{v^* \in B_{r_1}(\hat{v}^*)} \sup_{v_0^* \in C^*} J^*(v^*, v_0^*)
= J^*(\hat{v}^*, \hat{v}_0^*). \tag{9}
\]

Proof. From \( \delta J(x_0) = 0 \) we obtain
\[
Ax_0 + \sum_{j=1}^N \gamma_j \left( \frac{x_0^T B_j x_0}{2} + c_j \right) B_j x_0 + f = 0.
\]
Hence
\[
-Ax_0 + Kx_0 - f = \sum_{j=1}^N \gamma_j \left( \frac{x_0^T B_j x_0}{2} + c_j \right) B_j x_0 + Kx_0
= \sum_{j=1}^N (\hat{v}_0^*)_j B_j x_0 + Kx_0
= \hat{v}^* \tag{10}
\]
Thus,
\[ x_0 = (KI_d - A)^{-1}(\hat{v}^* + f), \]
so that
\[
(KI_d - A)^{-1}(\hat{v}^* + f) - \left( \sum_{j=1}^{N} (\hat{v}_0^*)_j B_j + KI_d \right)^{-1} \hat{v}^* = x_0 - x_0
\]
\[ = 0, \tag{11} \]
and therefore
\[
\frac{\partial J^*(\hat{v}^*, \hat{v}_0^*)}{\partial v^*} = 0.
\]

From this and the implicit function theorem, we get
\[
\frac{\partial J^*(\hat{v}^*, \hat{v}_0^*)}{\partial v^*} = \frac{\partial J^*(\hat{v}^*, \hat{v}_0^*)}{\partial v^*} + \sum_{j=1}^{N} \frac{\partial J^*(\hat{v}^*, \hat{v}_0^*)}{\partial (v_0^*)_j} \frac{\partial (v_0^*)_j}{\partial v^*}.
\tag{12} \]

However, from
\[
(\hat{v}_0^*)_j = \gamma_j \left( \frac{x_0^T B_j x_0}{2} + c_j \right),
\]
we have
\[
0 = -\frac{\hat{v}_0^*)_j}{\gamma_j} + \frac{x_0^T B_j x_0}{2} + c_j
\]
\[
= \frac{\partial J^*(\hat{v}^*, \hat{v}_0^*)}{\partial (v_0^*)_j}, \forall j \in \{1, \ldots, N\}, \tag{13} \]
so that from (12), we obtain
\[
\frac{\partial \tilde{J}^*(\hat{v}^*)}{\partial v^*} = \frac{\partial J^*(\hat{v}^*, \hat{v}_0^*)}{\partial v^*} = 0
\tag{14} \]
Hence, we may denote
\[
\delta \tilde{J}^*(\hat{v}^*) = 0.
\]
On the other hand from (10), we have
\[
G_1^*(\hat{v}^*) = (\hat{v}^*)^T x_0 - \frac{K}{2} \hat{v}_0^T x_0 + \frac{1}{2} x_0^T A x_0 + f^T x_0
\]
\[= (\hat{v}^*)^T x_0 - G_1(x_0), \tag{15} \]
and
\[
G_2^*(\hat{v}^*, \hat{v}_0^*) = (\hat{v}^*)^T x_0 + (\hat{v}_0^*)^T x_0 - G_2(x_0, 0).
\]
Therefore
\[
\tilde{J}^*(\hat{v}^*) = J^*(\hat{v}^*, \hat{v}_0^*)
\]
\[= G_1^*(\hat{v}^*) - G_2^*(\hat{v}^*, \hat{v}_0^*)
\]
\[= -G_1(x_0) + G_2(x_0, 0)
\]
\[= J(x_0). \tag{16} \]
Observe also that

\[
\delta^2 \tilde{J}^*(\hat{v}^*) = \left\{ \frac{\partial^2 \tilde{J}^*(\hat{v}^*)}{\partial v_j^* \partial v_k^*} \right\} = \left\{ \frac{\partial^2 J^*(\hat{v}^*, \hat{v}_0^*)}{\partial v_j^* \partial v_k^*} + \sum_{l=1}^N \frac{\partial^2 J^*(\hat{v}^*, \hat{v}_0^*)}{\partial v_j^* \partial (v_0^*)_l} \frac{\partial (v_0^*)_l}{\partial v_k^*} \right\},
\]

where \( \hat{v}_0^* \) is such that

\[
\frac{\partial J^*(\hat{v}^*, \hat{v}_0^*)}{\partial (v_0^*)_l} = 2\left( v_0^* \right)^T \left( \sum_{j=1}^N (\hat{v}_0^*)_j B_j + K I_d \right)^{-1} B_l \left( \sum_{j=1}^N (\hat{v}_0^*)_j B_j + K I_d \right)^{-1} (v_0^*) \\
- \frac{(v_0^*)_l}{\gamma_l} + c_l \]

\[
= 0.
\]

Taking the variation of this last equation in \( v_k^* \), we get

\[
\epsilon_k^T \left( \sum_{j=1}^N (\hat{v}_0^*)_j B_j + K I_d \right)^{-1} B_l x_0 \\
- \sum_{\eta=1}^N \left( x_0^T B_l \left( \sum_{j=1}^N (\hat{v}_0^*)_j B_j + K I_d \right)^{-1} B_\eta x_0 \frac{\partial (\hat{v}_0^*)_\eta}{\partial v_k^*} \right) \\
- \frac{1}{\gamma_l} \frac{\partial (\hat{v}_0^*)_l}{\partial v_k^*} \]

\[
= 0
\]

From this, denoting

\[
- \frac{1}{\gamma_l} \frac{\partial (\hat{v}_0^*)_l}{\partial v_k^*} = - \frac{1}{\gamma_l} \sum_{\eta=1}^N \delta_{l\eta} \frac{\partial (\hat{v}_0^*)_\eta}{\partial v_k^*}
\]

we obtain

\[
\left\{ \frac{\partial (\hat{v}_0^*)_l}{\partial v_k^*} \right\} = \left[ x_0^T B_l \left( \sum_{j=1}^N (\hat{v}_0^*)_j B_j + K I_d \right)^{-1} B_\eta x_0 + \frac{1}{\gamma_l} \delta_{l\eta} \right]^{-1} \\
\times \left[ x_0^T B_\eta \left( \sum_{j=1}^N (\hat{v}_0^*)_j B_j + K I_d \right)^{-1} e_k \right] \\
= \mathcal{E} P_2.
\]

\[7\]
Also
\[
\left\{ \frac{\partial^2 J^*(\hat{v}^*, \hat{v}^*_0)}{\partial v^*_j \partial (v^*_0)_l} \right\}
\]
\[
= \begin{bmatrix} e_j^T \left( \sum_{p=1}^{N} (\hat{v}^*_0)_p B_p + K I_d \right) \end{bmatrix}_{n \times N}^{-1} B_1 x_0
\]
\[
= \left( \sum_{p=1}^{N} (\hat{v}^*_0)_p B_p + K I_d \right)^{-1} \begin{bmatrix} B_1 x_0 & B_2 x_0 & \cdots & B_N x_0 \end{bmatrix}_{n \times N}
\]
\[
= \sum_{p=1}^{N} (\hat{v}^*_0)_p B_p + K I_d )^{-1} P_1,
\]  
so that
\[
\begin{bmatrix} \sum_{l=1}^{N} \frac{\partial^2 J^*(\hat{v}^*, \hat{v}^*_0)}{\partial v^*_j \partial (v^*_0)_l} \end{bmatrix}
\]
\[
= \left( \sum_{p=1}^{N} (\hat{v}^*_0)_p B_p + K I_d \right)^{-1} \begin{bmatrix} P_1 E P_2 \end{bmatrix}
\]
\[
= \left( \sum_{p=1}^{N} (\hat{v}^*_0)_p B_p + K I_d \right)^{-1} \begin{bmatrix} H_3 \end{bmatrix}.
\]  
Therefore, recalling that
\[
D = \hat{B} \left( \sum_{p=1}^{N} (\hat{v}^*_0)_p B_p + K I_d \right)^{-1} + I_d
\]
where
\[
\hat{B}_{n \times n} = \left\{ \hat{B}_{jk} \right\} = \left\{ \sum_{l=1}^{N} \sum_{s,q=1}^{n} \gamma_l (x_0)_s (B_l)_{js} (B_l)_{qk} (x_0)_q \right\},
\]

8
we may write

\[
\delta^2 \tilde{J}^*(\hat{v}^*) D
= - \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + KI_d \right)^{-1} (I_d - H_3) D \\
+ (KI_d - A)^{-1} D
= - \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + KI_d \right)^{-1} \left( (I_d - H_3) D - I_d + I_d \right)
+ (KI_d - A)^{-1} D
= - \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + KI_d \right)^{-1} (I_d + \alpha)
+ (KI_d - A)^{-1} D
= (KI_d - A)^{-1} \left( -(KI_d - A)(I_d - \alpha_1) \right)
+ \hat{B} + \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + KI_d \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + KI_d \right)^{-1}
\]

Therefore, denoting also

\[
H_1 = (KI_d - A)^{-1}, \quad H_2 = \left( \sum_{p=1}^{N} (\hat{v}_0^*)_p B_p + KI_d \right)^{-1},
\]

we have

\[
\delta^2 \tilde{J}^*(\hat{v}^*) D
= H_1 \left( A(I_d - \alpha_1) + \hat{B} + \sum_{\eta=1}^{N} (\hat{v}_0^*)_{\eta} B_{\eta} + KI_d(\alpha_1) \right) H_2
= H_1(\delta^2 J(x_0) + (KI_d - A)(\alpha_1)) H_2.
\]

Since \( D, H_1 \) and \( H_2 \) are symmetric positive definite matrices, assuming \( \delta^2 J(u_0) > 0 \) and \( \delta^2 J(x_0) + (KI_d - A)(\alpha_1) > 0 \), we have

\[
\delta^2 \tilde{J}(\hat{v}^*) > 0.
\]

9
so that there exist $r > 0$ and $r_1 > 0$ such that

$$J(x_0) = \inf_{x \in B_r(x_0)} J(x) = \inf_{v^* \in B_{r_1}(\hat{v}^*)} \tilde{J}(v^*) = \tilde{J}(\hat{v}^*) = \inf_{v^* \in B_{r_1}(\hat{v}^*)} \sup_{v^* \in C^*} J^*(v^*, v^*_0) = J^*(\hat{v}^*, \hat{v}^*_0).$$  \hfill (26)

Assume now $\hat{v}^*_0 \in A^*$ so that $\delta^2 J(x_0) > 0$.

Observe that if $v_0^* \in A^*$, then

$$J^*(v^*, v^*_0) = G_1^*(v^*) - G_2^*(v^*, v^*_0)$$

is such that

$$\frac{\partial J^*(v^*, v^*_0)}{\partial (v^*)} = (KI_d - A)^{-1} - \left( \sum_{j=1}^{N} (v_0^*)_j B_j + KI_d \right)^{-1} > 0,$$

so that defining

$$J_2^*(v^*) = \sup_{v_0^* \in A^*} J^*(v^*, v_0^*)$$

we have that $J_2^*$ is convex as the supremum of a family of convex functions.

Similarly as above, we may obtain

$$\delta J_2^*(\hat{v}^*) = 0$$

and

$$J_2^*(\hat{v}^*) = J(x_0) = J^*(\hat{v}^*, \hat{v}^*_0).$$

From this, since $J_2^*$ is convex, from the min-max theorem and from the general result in Toland [7], we may infer that

$$J_2^*(\hat{v}^*) = \inf_{v^* \in \mathbb{R}^n} J_2^*(v^*) = \inf_{v^* \in \mathbb{R}^n} \sup_{v_0^* \in A^*} J^*(v^*, v_0^*)$$

$$\leq \sup_{v_0^* \in A^*} \left\{ -G_1(x) + \frac{K}{2} x^T x + \sum_{j=1}^{N} \left( (v_0^*)_j \left( \frac{x^T B_j x}{2} + c_j \right) - \frac{(v_0^*)_j^2}{2 \gamma_j} \right) \right\}$$

$$\leq \sup_{v_0^* \in \mathbb{R}^N} \left\{ -G_1(x) + \frac{K}{2} x^T x + \sum_{j=1}^{N} \left( (v_0^*)_j \left( \frac{x^T B_j x}{2} + c_j \right) - \frac{(v_0^*)_j^2}{2 \gamma_j} \right) \right\}$$

$$= -G_1(x) + G_2(x, 0)$$

$$= J(x), \ \forall x \in \mathbb{R}^n.$$  \hfill (27)
Hence
\[ \inf_{x \in \mathbb{R}^n} J(x) \geq J^*_2(\tilde{v}^*) = J(x_0), \]
so that
\[
\begin{align*}
J(x_0) &= \inf_{x \in \mathbb{R}^n} J(x) \\
&= \inf_{v^* \in \mathbb{R}^n} J_2(v^*) \\
&= J_2(\tilde{v}^*) \\
&= \inf_{v^* \in \mathbb{R}^n} \sup_{v^*_0 \in A^*} J^*(v^*, v^*_0) \\
&= J^*(\tilde{v}^*, \tilde{v}^*_0).
\end{align*}
\] (28)

Finally, the proof of third item is similar to that of the first one.
This would complete the proof. □

**Remark 2.3.** For the special case in which \( n = N = 1 \) we obtain \( \alpha_1 = 0 \).

**Remark 2.4.** We may obtain an even more interesting result if we consider a more general case in which \( K \) is a symmetric matrix \( n \times n \). Specifically for the case
\[ K = K I_d = A + \varepsilon I_d \]
we get
\[ K I_d - A = \varepsilon I_d, \]
and in such a case
\[
\begin{align*}
\delta^2 \tilde{J}^*(\tilde{v}^*) D \\
&= H_1(\delta^2 J(x_0) + (K I_d - A)(\alpha_1)) H_2 \\
&= H_1(\delta^2 J(x_0) + \varepsilon I_d(\alpha_1)) H_2 \\
&= H_1(\delta^2 J(x_0) + O(\varepsilon) I_d) H_2
\end{align*}
\] (29)
so that we recover at least approximately a correspondence between \( \delta^2 J(x_0) \) and \( \delta^2 \tilde{J}^*(\tilde{v}^*) \), up to considering the sign of \( H_2 \) as well.
Observe that in this last context,
\[ H_1 = \frac{1}{\varepsilon} I_d \]
and
\[ H_2 = \left( A + \sum_{p=1}^{N} (\tilde{v}^*_0)_p B_p + \varepsilon I_d \right)^{-1}. \]

**Remark 2.5.** Let us now consider a dual functional proposed in the current literature (see [6], for example). For the model addressed in this article, such a functional is expressed as
\[
-J_1^*(v_0^*) = \frac{1}{2} f^T \left( \sum_{p=1}^{N} (v_0^*)_p B_p + A \right)^{-1} f + \sum_{p=1}^{N} \left( (v_0^*)_p \right)^2 - \sum_{p=1}^{N} c_p(v_0^*)_p.
\]
Taking the variation (in fact derivative) of such a function al in \((v_0^*)_j\), since the matrices in question are symmetric, we obtain

\[
\frac{\partial J_1^* (v_0^*)}{\partial (v_0^*)_j} = -\frac{1}{2} f^T \left( \sum_{p=1}^N (v_0^*)_p B_p + A \right)^{-1} B_j \left( \sum_{p=1}^N (v_0^*)_p B_p + A \right)^{-1} f + \frac{(v_0^*)_j}{\gamma_j} - c_j
\]

(30)

Now taking the derivative of this expression relating \((v_0^*)_k\) we get

\[
\left\{ -\frac{\partial^2 J_1^* (v_0^*)}{\partial (v_0^*)_j \partial (v_0^*)_k} \right\} = \left\{ f^T \left( \sum_{j=1}^N (v_0^*)_j B_j + A \right)^{-1} \right. \\
\times B_k \left( \sum_{j=1}^N (v_0^*)_j B_j + A \right)^{-1} f + \frac{\delta_{jk}}{\gamma_j} \right\}
\]

(31)

Since the matrices in question are symmetric, at a critical point as specified in the last theorem, we obtain,

\[
\left\{ -\frac{\partial^2 J_1^* (v_0^*)}{\partial (v_0^*)_j \partial (v_0^*)_k} \right\} = \left\{ x_0^T B_j \left( \sum_{p=1}^N (\hat{v}_0^*)_p B_p + A \right)^{-1} \right. \\
B_k x_0 + \frac{\delta_{jk}}{\gamma_j} \right\}
\]

(32)

On the other hand, for the functional \(J(x)\) we obtain

\[
\delta^2 J(x_0) = A + \hat{B} + \sum_{p=1}^N (\hat{v}_0^*)_p B_p
\]

(33)

where

\[
\hat{B} = \hat{B}_{n \times n} = \{ \hat{B}_{jk} \} = \left\{ \sum_{l=1}^N \sum_{s,q=1}^n \gamma_l (x_0)_s (B_l)_{js} (B_l)_{qk} (x_0)_q \right\}
\]

From this we may see that there exists a qualitative correspondence (in terms of positivity or negativity in a matrix sense) between the two second derivative matrices only for the special case \(n = N = 1\). Even so we have to consider the sign of \(\sum_{p=1}^N (\hat{v}_0^*)_p B_p + A\) to get a right conclusion.

For a general case such a correspondence may not hold even if \(n = N\).
3 Conclusion

In this article we have developed a duality principle for a class of non-convex optimization problems in $\mathbb{R}^n$. For such a class of problems we address the case in which for the variables in question, $n \neq N$.

We believe to have obtained a very interesting way of developing the dual formulation, establishing a correct relation between the critical points of the primal and dual problems, with no duality gap between such primal and dual formulations.

This problem has been addressed in similar form in [5, 6], for example. It is not our objective here to comment extensively such previous results, but just offer a new possibility of obtaining the dual formulations for such a class of problems.

References

[1] W.R. Bielski, A. Galka, J.J. Telega, The Complementary Energy Principle and Duality for Geometrically Nonlinear Elastic Shells. I. Simple case of moderate rotations around a tangent to the middle surface. Bulletin of the Polish Academy of Sciences, Technical Sciences, Vol. 38, No. 7-9, 1988.

[2] W.R. Bielski and J.J. Telega, A Contribution to Contact Problems for a Class of Solids and Structures, Arch. Mech., 37, 4-5, pp. 303-320, Warszawa 1985.

[3] F. Botelho, Functional Analysis and Applied Optimization in Banach Spaces, (Springer Switzerland, 2014).

[4] F. Botelho, Real Analysis and Applications, (Springer Switzerland, 2018).

[5] D.Y. Gao and H.F. Yu, Multi-scale modelling and canonical dual finite element method in phase transition in solids. Int. J. Solids Struct., 45, 3660-3673 (2008).

[6] D.Y.Gao and C. Wu, On the Triality Theory in Global Optimization, Arxiv: 1104.2970 - v2, February, 2012.

[7] J.F. Toland, A duality principle for non-convex optimisation and the calculus of variations, Arch. Rath. Mech. Anal., 71, No. 1 (1979), 41-61.