A Note on Equation (4) in "A Computationally Useful Algebraic Representation of Nonlinear Disjunctive Convex Sets Using the Perspective Function"

The purpose of this note is to provide more insight around a condition that appears in [1] as equation (4). In [1], Furman, Sawaya & Grossmann study the disjunctive set

$$F = \underset{j \in J}{\cup} C_j ,$$

where J is finite, $C_j \equiv \left\{ x \in \mathbb{R}^n \middle| G_j(x) \le 0 \right\}$ and $G_j : \mathbb{R}^n \to \left(\mathbb{R} \cup \{\infty\} \right)^{m_j}$ are vector mappings whose components $g_{ij}, i = 1 \ldots m_j$ are proper closed convex functions. Furthermore, $C_j, \forall j \in J$ are assumed to be compact, though not necessarily non-empty.

The set $eps-rel\ F(\varepsilon)$ is defined to be all those $(x, v, \lambda) \in \mathbb{R}^{n+n|J|+|J|}$ that satisfy the following set of constraints for some $0 < \varepsilon < 1$:

$$x = \sum_{j \in J} v_j$$

$$((1-\varepsilon)\lambda_j + \varepsilon) g_{ij} \left(\frac{v_j}{(1-\varepsilon)\lambda_j + \varepsilon} \right) - \varepsilon g_{ij}(0)(1-\lambda_j) \le 0, \quad i = 1 \ldots m_j,\ j \in J$$

$$\sum_{j \in J} \lambda_j = 1$$

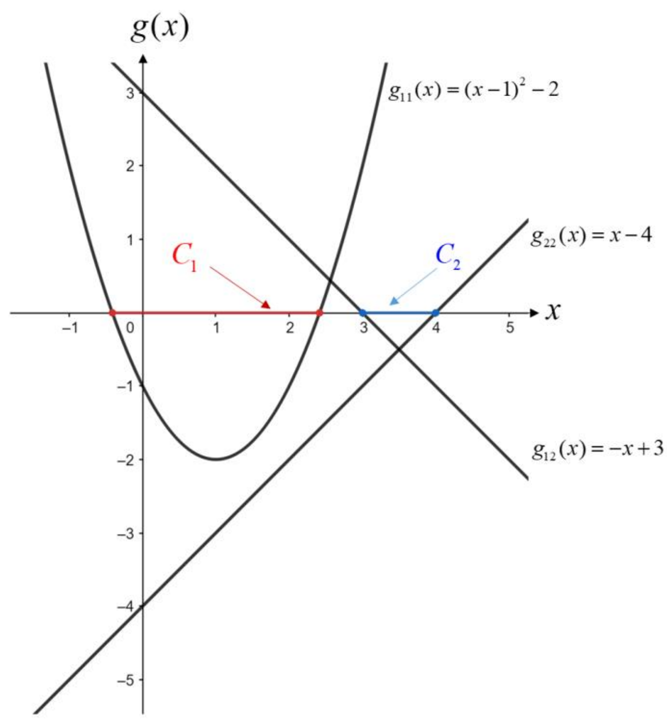
$$\lambda_j \ge 0, \quad j \in J ,$$

and $eps-MIP\ F(\varepsilon) \equiv \left\{ (x, v, \lambda) \in eps-rel\ F(\varepsilon) \middle| \lambda_j \in \{0,1\},\ j \in J \right\}$. The projection of $eps-MIP\ F(\varepsilon)$ onto the x space is defined as $proj_{(x)}(eps-MIP\ F(\varepsilon)) \equiv \left\{ x \in eps-MIP\ F(\varepsilon) \middle| (v, \lambda) \in eps-MIP\ F(\varepsilon) \right\}$, and in Proposition 1 of [1], it was proven that for any $0 < \varepsilon < 1$, $proj_{(x)}(eps-MIP\ F(\varepsilon)) = F$ under the assumption that the sets $C_j \equiv \left\{ x \in \mathbb{R}^n \middle| G_j(x) \le 0 \right\}$ are such that $G_j(0)$ is defined and the following condition holds:

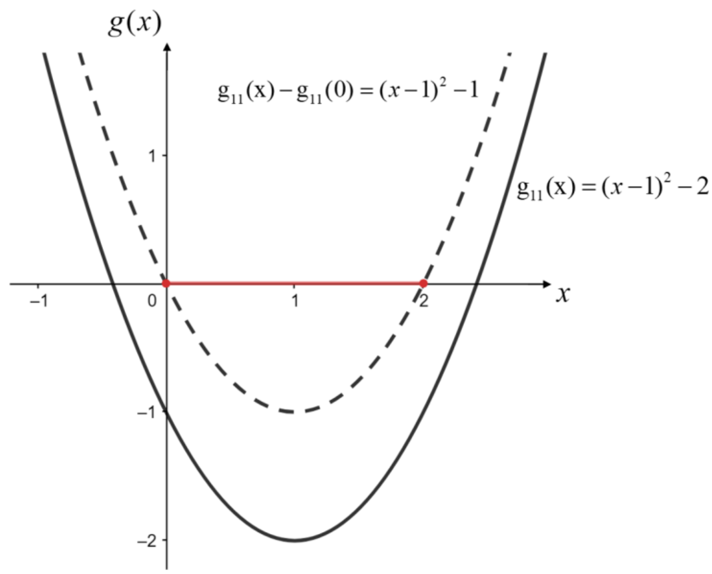
$$\left\{ x \in \mathbb{R}^n \middle| G_j(x) - G_j(0) \le 0 \right\} = \{0\}, \quad \forall j \in J \,. \tag{1}$$

This is the condition that appears in [1] as equation (4). If this condition holds, then the vector function $G_j(x)$ is such that the only point $x \in \mathbb{R}^n$ satisfying $G_j(x) - G_j(0) \le 0$ $\forall j \in J$ is the "0" point. This condition is needed in order to ensure that $\nu_{j'} = 0$ when $\lambda_j = 1$ for $\forall j' \setminus j \in J$ in $eps-MIP\ F(\varepsilon)$ (see equations (14)-(15) in the proof of Proposition 1 in [1]). On the other hand, if the condition in (1) does not hold, then $proj_{(x)}(eps-MIP\ F(\varepsilon)) \ne F$. We show this through the following simple example.

Example 1. Let $F = C_1 \cup C_2$, where $C_1 \equiv \left\{ x \in \mathbb{R} \middle| (x-1)^2 - 2 \le 0 \right\}$ and $C_2 \equiv \left\{ x \in \mathbb{R} \middle| 3 \le x \le 4 \right\}$. The feasible regions of sets C_1 and C_2 are plotted below.



The condition described in (1) doesn't hold for set C_1 since $\left\{x \in \mathbb{R} \middle| \left((x-1)^2 - 2\right) - (-1) \le 0\right\} = \left\{x \in \mathbb{R} \middle| x^2 - 2x \le 0\right\} = \left\{x \in \mathbb{R} \middle| x(x-2) \le 0\right\}$, which implies that $0 \le x \le 2$ (which is clearly not $\{0\}$). We can also see this graphically:



As such, if we generate $eps-MIP\ F(\varepsilon)$ for this set, $proj_{(x)}(eps-MIP\ F(\varepsilon)) \ne F$. Indeed, the $eps-MIP\ F(\varepsilon)$ formulation of this disjunctive set for some $0 < \varepsilon < 1$ is the following:

$$x = v_1 + v_2$$

$$((1-\varepsilon)\lambda_1 + \varepsilon)\left[\left(\frac{v_1}{(1-\varepsilon)\lambda_1 + \varepsilon} - 1\right)^2 - 2\right] - \varepsilon(-1)(1-\lambda_1) \le 0$$

$$((1-\varepsilon)\lambda_2 + \varepsilon)\left[3 - \frac{v_2}{(1-\varepsilon)\lambda_2 + \varepsilon}\right] - \varepsilon(3)(1-\lambda_2) \le 0$$

$$((1-\varepsilon)\lambda_2 + \varepsilon)\left[\frac{v_2}{(1-\varepsilon)\lambda_2 + \varepsilon} - 4\right] - \varepsilon(-4)(1-\lambda_2) \le 0$$

$$\lambda_1 + \lambda_2 = 1$$

$$\lambda_1 \in \{0,1\}.$$

We can simplify this to:

$$x = \nu_1 + \nu_2$$

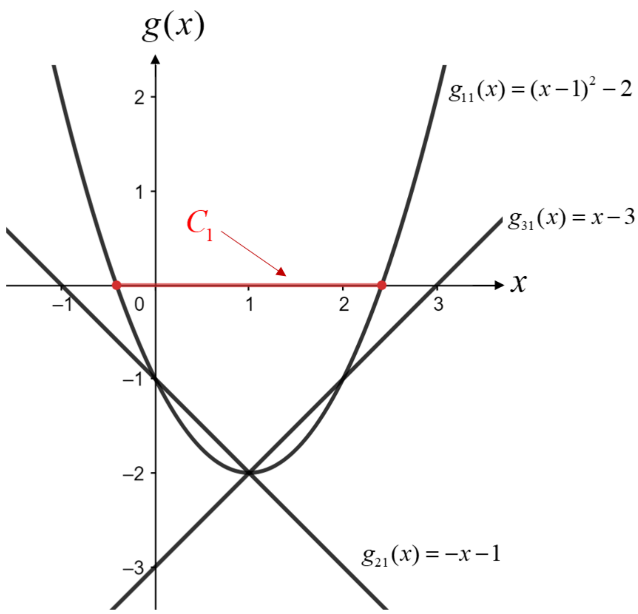
$$((1-\varepsilon)\lambda_1 + \varepsilon)\left[\left(\frac{\nu_1}{(1-\varepsilon)\lambda_1 + \varepsilon} - 1\right)^2 - 2\right] - \varepsilon(-1)(1-\lambda_1) \le 0$$

$$3(1-\lambda_1) \le \nu_2 \le 4(1-\lambda_1)$$

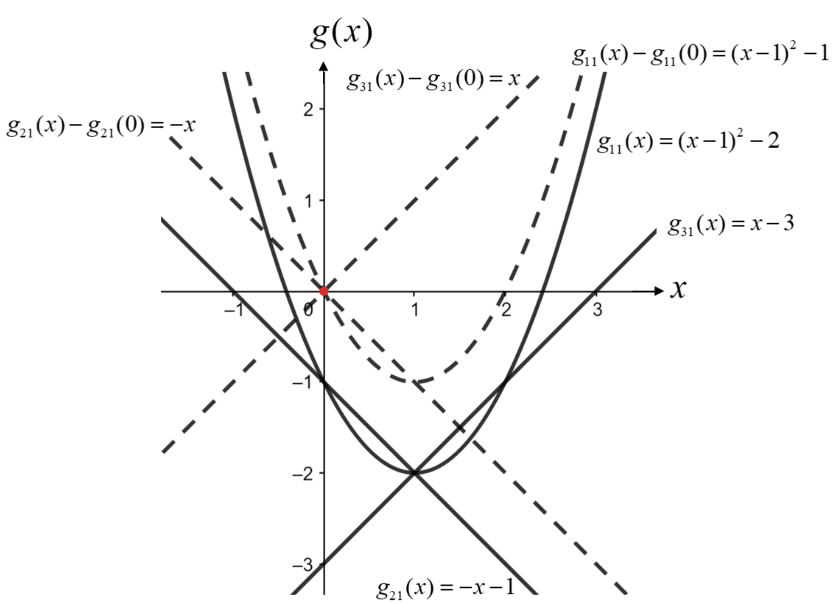
$$\lambda_1 \in \{0,1\}.$$

When $\lambda_1 = 1$, then $\nu_2 = 0$; thus $x = \nu_1$ and $(x-1)^2 - 2 \le 0$, which checks out. However, when $\lambda_1 = 0$, $3 \le \nu_2 \le 4$ and $\varepsilon\left[\left(\frac{\nu_1}{\varepsilon} - 1\right)^2 - 2\right] + \varepsilon \le 0 \Rightarrow \varepsilon\left(\frac{\nu_1^{\,2}}{\varepsilon^2} - 2\frac{\nu_1}{\varepsilon} + 1 - 2\right) + \varepsilon \le 0$ $\Rightarrow \frac{\nu_1^{\,2}}{\varepsilon} - 2\nu_1 \le 0 \Rightarrow \nu_1\left(\frac{\nu_1}{\varepsilon} - 2\right) \le 0$, which implies that $0 \le \nu_1 \le 2\varepsilon$; therefore, ν_1 is not necessarily equal to 0 when $\lambda_1 = 0$. As such, our reformulation is not an exact representation of the original disjunctive problem.

Example 2. Let us now slightly modify the example above by imposing bounds inside the term C_1 such that the new set $C_1 \equiv \left\{ x \in \mathbb{R} \,\middle|\, \begin{matrix} (x-1)^2 - 2 \le 0 \\ -1 \le x \le 3 \end{matrix} \right\}$; we keep $C_2 \equiv \left\{ x \in \mathbb{R} \,\middle|\, 3 \le x \le 4 \right\}$. Clearly, the feasible region for C_1 remains the same as in Example 1 since the bounds are technically redundant. We can see this graphically:



However, the condition described in (1) now holds for the set C_1 since $\left\{ x \in \mathbb{R} \middle| \begin{array}{l} \left((x-1)^2-2\right)-(-1) \le 0 \\ -x-1-(-1) \le 0 \\ x-3-(-3) \le 0 \end{array} \right\} = \left\{ x \in \mathbb{R} \middle| \begin{array}{l} x^2-2x \le 0 \\ x \ge 0 \\ x \le 0 \end{array} \right\} = \left\{ x \in \mathbb{R} \middle| \begin{array}{l} x(x-2) \le 0 \\ x = 0 \end{array} \right\}$, which implies that $x = 0$. We can also see this graphically:



If we generate $eps-MIP\ F(\varepsilon)$ for this set, $proj_{(x)}(eps-MIP\ F(\varepsilon))=F$. Indeed, the $eps-MIP\ F(\varepsilon)$ formulation of this disjunctive set for some $0<\varepsilon<1$ is the following:

$$x=\nu_1+\nu_2$$

$$((1-\varepsilon)\lambda_1+\varepsilon)\left[\left(\frac{\nu_1}{(1-\varepsilon)\lambda_1+\varepsilon}-1\right)^2-2\right]-\varepsilon(-1)(1-\lambda_1)\le 0$$

$$((1-\varepsilon)\lambda_1+\varepsilon)\left[-1-\frac{\nu_1}{(1-\varepsilon)\lambda_1+\varepsilon}\right]-\varepsilon(-1)(1-\lambda_1)\le 0$$

$$((1-\varepsilon)\lambda_1+\varepsilon)\left[\frac{\nu_1}{(1-\varepsilon)\lambda_1+\varepsilon}-3\right]-\varepsilon(-3)(1-\lambda_1)\le 0$$

$$((1-\varepsilon)\lambda_2+\varepsilon)\left[3-\frac{\nu_2}{(1-\varepsilon)\lambda_2+\varepsilon}\right]-\varepsilon(3)(1-\lambda_2)\le 0$$

$$((1-\varepsilon)\lambda_2+\varepsilon)\left[\frac{\nu_2}{(1-\varepsilon)\lambda_2+\varepsilon}-4\right]-\varepsilon(-4)(1-\lambda_2)\le 0$$

$$\lambda_1+\lambda_2=1$$

$$\lambda_1\in\{0,1\}.$$

We can simplify this to:

$$x=\nu_1+\nu_2$$

$$((1-\varepsilon)\lambda_1+\varepsilon)\left[\left(\frac{\nu_1}{(1-\varepsilon)\lambda_1+\varepsilon}-1\right)^2-2\right]-\varepsilon(-1)(1-\lambda_1)\le 0$$

$$-\lambda_1\le\nu_1\le 3\lambda_1$$

$$3(1-\lambda_1)\le\nu_2\le 4(1-\lambda_1)$$

$$\lambda_1\in\{0,1\}.$$

When $\lambda_1=1$, then $\nu_2=0$; thus $x=\nu_1$ and $(x-1)^2-2\le 0$, which checks out. When $\lambda_1=0$, then $\nu_1=0$; thus $x=\nu_2$ and $3\le x\le 4$ and $\varepsilon\left[(0-1)^2-2\right]+\varepsilon\le 0\Rightarrow\varepsilon(-1)+\varepsilon\le 0$ $\Rightarrow 0\le 0$, which is redundant. As such, our reformulation is an exact representation of the original disjunctive problem.

It turns out that having bounds on all variables inside every term of the disjunction automatically insures that the condition in (1) holds. In fact, if a subset of the constraints g_{ij} make up a polytope $a_{ij}x \le b_{ij},\ i = 1\ldots l_j < m_j$, then the condition in (1) always holds. We prove this in the following proposition:

Proposition 1. If the sets $C_j \equiv \left\{ x \in \mathbb{R}^n \middle| \begin{matrix} a_{ij}x \le b_{ij},\ i = 1\ldots l_j \\ g_{ij}(x) \le 0, i = 1 + l_j \ldots m_j \end{matrix} \right\},\ \forall j \in J$, with $a_{ij}x \le b_{ij},\ i = 1\ldots l_j$ defining a polytope, then the condition in (1) holds.

Proof: For simplicity of notation, let us replace $a_{ij}x \le b_{ij},\ i = 1\ldots l_j$ by $A_j x \le b_j$, with A_j an $l_j \times n$ matrix. It is well known that the recession cone of $A_j x \le b_j$ is $A_j x \le 0$ (see for e.g. p. 39 in [2]). Furthermore, by definition, a polytope is a closed and bounded polyhedron; therefore it is compact. As such, $\left\{ x \in \mathbb{R}^n \middle| A_j x \le 0 \right\} = \{0\}$ ([2] Section A Proposition 2.2.3). This implies that $\left\{ x \in \mathbb{R}^n \middle| \begin{matrix} A_j x \le 0 \\ g_{ij}(x) - g_{ij}(0) \le 0, i = 1 + l_j \ldots m_j \end{matrix} \right\} = \{0\}$. But this corresponds precisely to (1), with $G_j(x) - G_j(0) = \begin{cases} (A_j x - b_j) - (-b_j) \\ g_{ij}(x) - g_{ij}(0), i = 1 + l_j \ldots m_j \end{cases}$.

■

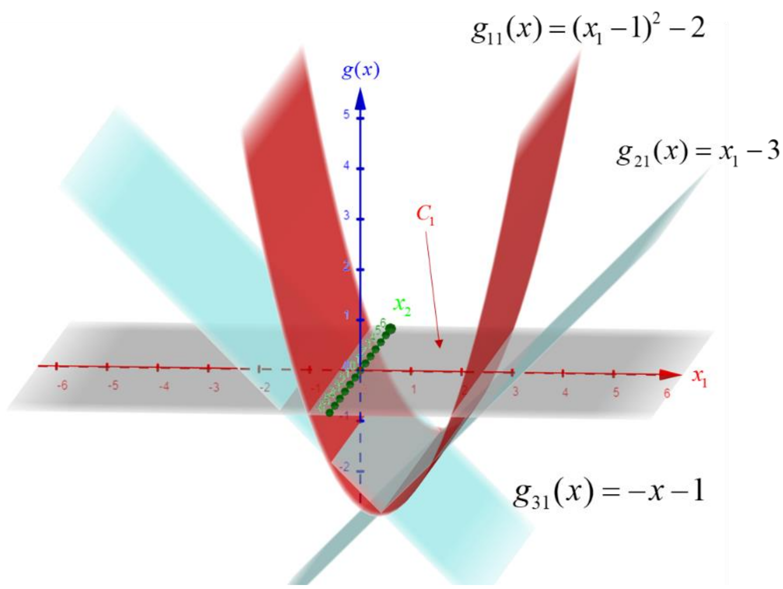
Corollary 1. If the sets $C_j,\ \forall j \in J$ contain explicit bounds $x_k^{LB_j} \le x_k \le x_k^{UB_j}$ on every variable $x_k,\ k = 1...n$, then the condition in (1) holds.

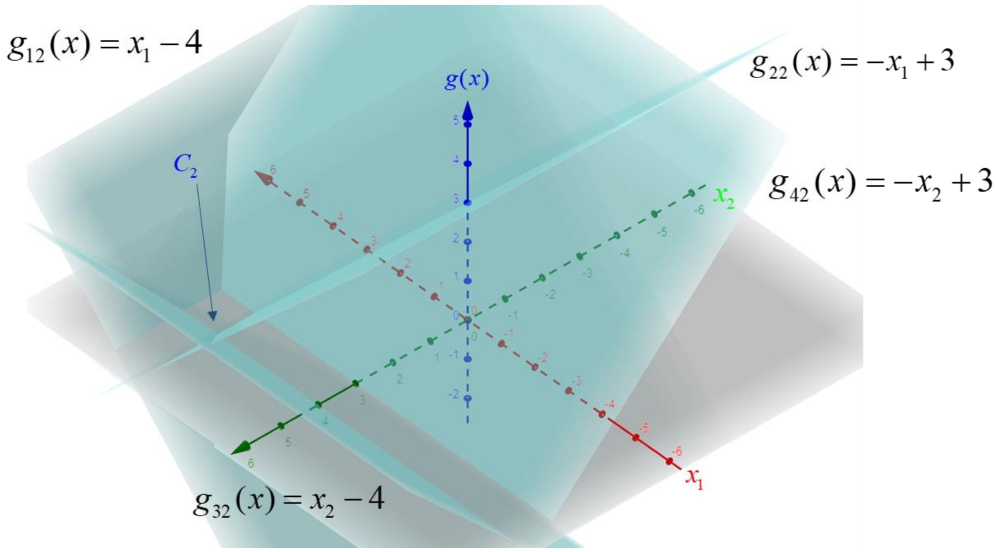
Proof: This automatically follows from Proposition 1 by virtue of the fact that the explicit bounds on every variable $x_k,\ k = 1...n$ constitute a polytope.

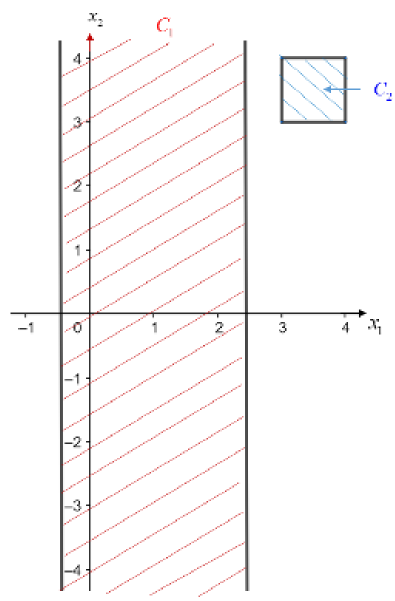
■

Remark 1. In order for Corollary 1 to apply, every term C_j must contain bounds on every variable x_k, $k = 1...n$ that appears in any term C_j; otherwise, Corollary 1 fails to hold. We show the latter in the following example.

Example 3. Let $F = C_1 \cup C_2$, where $C_1 \equiv \left\{ (x_1, x_2) \in \mathbb{R}^2 \left| \begin{matrix} (x_1 - 1)^2 - 2 \le 0 \\ -1 \le x_1 \le 3 \end{matrix} \right. \right\}$ and $C_2 \equiv \left\{ (x_1, x_2) \in \mathbb{R}^2 \left| \begin{matrix} 3 \le x_1 \le 4 \\ 3 \le x_2 \le 4 \end{matrix} \right. \right\}$. The feasible regions of sets C_1 and C_2 are plotted below (in separate graphs for clarity of exposition, followed by the projection onto the (x_1, x_2) space):







Since the variables x_1 and x_2 appear in C_2, bounds on both x_1 and x_2 are needed in both C_1 and C_2 in order for Corollary 1 to hold. However, in Example 3, only bounds on x_1 appear; therefore Corollary 1 does not hold, and the condition in (1) does not hold. As such, if we generate $eps-MIP\ F(\varepsilon)$ for this set, $proj_{(x)}(eps-MIP\ F(\varepsilon)) \neq F$. Indeed, the $eps-MIP\ F(\varepsilon)$ formulation of this disjunctive set for some $0 < \varepsilon < 1$ is the following:

$$x_1 = \nu_{11} + \nu_{12}$$

$$x_2 = \nu_{21} + \nu_{22}$$

$$((1-\varepsilon)\lambda_1 + \varepsilon)\left[\left(\frac{\nu_{11}}{(1-\varepsilon)\lambda_1 + \varepsilon} - 1\right)^2 - 2\right] - \varepsilon(-1)(1-\lambda_1) \le 0$$

$$-\lambda_1 \le \nu_{11} \le 3\lambda_1$$

$$3\lambda_2 \le \nu_{12} \le 4\lambda_2$$

$$3\lambda_2 \le \nu_{22} \le 4\lambda_2$$

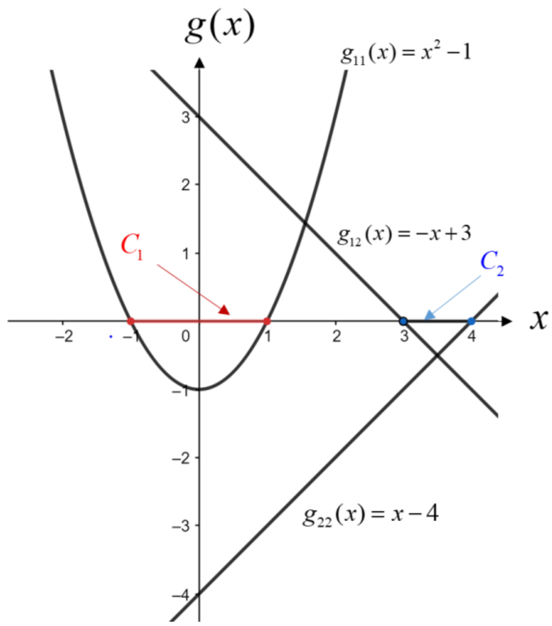
$$\lambda_1 + \lambda_2 = 1$$

$$\lambda_1 \in \{0,1\}.$$

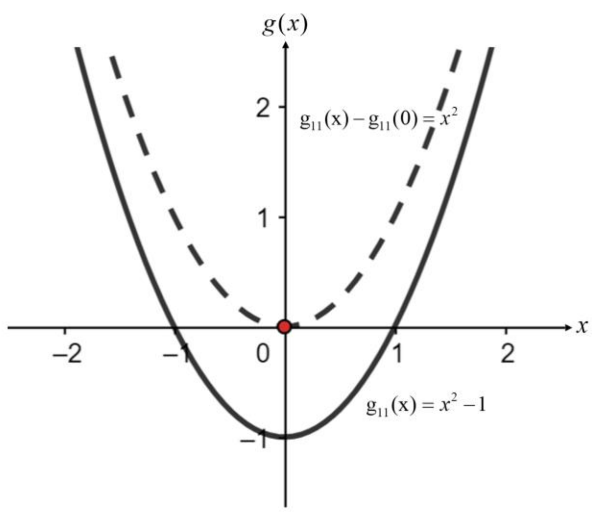
When $\lambda_1 = 1$, then $\lambda_2 = 0$ and $\nu_{12} = \nu_{22} = 0$; thus $x_1 = \nu_{11}$, $(x_1 - 1)^2 - 2 \le 0$ and $-1 \le x_1 \le 3$, which checks out. However, when $\lambda_1 = 0$, although $\nu_{11} = 0$ and therefore $x_1 = \nu_{12}$ (which implies that $\lambda_2 = 1$ and $3 \le x_1 \le 4$), there are no bounds on ν_{21} to drive it to 0 (since there are no bounds on x_2 in C_1). As such, $x_2 = \nu_{21} + \nu_{22}$, and although $3 \le \nu_{22} \le 4$, since ν_{21} is unbounded, there is no guarantee that $3 \le x_2 \le 4$. As such, our reformulation is not an exact representation of the original disjunctive problem.

Although Proposition 1 provides a *sufficient* condition for (1) to hold, having a subset of the constraints $g_{ij}, i = 1\ldots l_j < m_j$ making up a polytope (or having bounds on every variable) is not *necessary* for (1) to hold. We show this via the following example.

Example 4. Let $F = C_1 \cup C_2$, where $C_1 \equiv \left\{x \in \mathbb{R} \middle| x^2 - 1 \le 0\right\}$ and $C_2 \equiv \left\{x \in \mathbb{R} \middle| 3 \le x \le 4\right\}$. The feasible regions of sets C_1 and C_2 are plotted below:



The condition described in (1) holds for set C_1 since $\left\{x\in\mathbb{R}\,\middle|\,\left(x^2-1\right)-(-1)\le 0\right\}$ $=\left\{x\in\mathbb{R}\,\middle|\,x^2\le 0\right\}$, which implies that $x=0$. We can also see this graphically:



Of course, the condition also holds for set C_2 because we have a bounded range on x (see Corollary 1). Indeed, $\left\{x\in\mathbb{R}\,\middle|\,\left(x-4\right)-(-4)\le 0\cap\left(-x+3\right)-(3)\le 0\right\}$ $=\left\{x\in\mathbb{R}\,\middle|\,x\le 0\cap x\ge 0\right\}$, which implies that $x=0$. Now if we generate the $eps-MIP\ F(\varepsilon)$ formulation of this disjunctive set for some $0<\varepsilon<1$, we get (after simplifying):

$$x = \nu_1 + \nu_2$$

$$((1-\varepsilon)\lambda_1 + \varepsilon)\left[\left(\frac{\nu_1}{(1-\varepsilon)\lambda_1 + \varepsilon}\right)^2 - 1\right] - \varepsilon(-1)(1-\lambda_1) \le 0$$

$$3(1-\lambda_1) \le \nu_2 \le 4(1-\lambda_1)$$

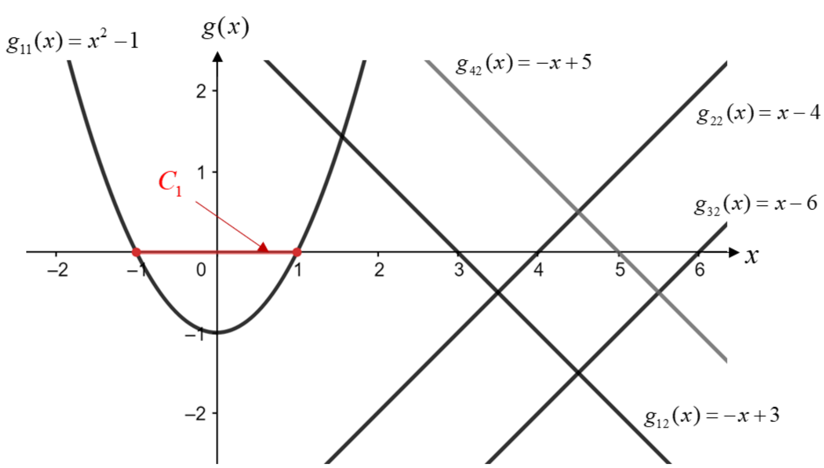
$$\lambda_1 \in \{0,1\}.$$

When $\lambda_1 = 1$, then $\nu_2 = 0$; thus $x = \nu_1$ and $x^2 - 1 \le 0$, which checks out. When $\lambda_1 = 0$, $\varepsilon\left[\left(\frac{\nu_1}{\varepsilon}\right)^2 - 1\right] + \varepsilon \le 0 \Rightarrow \frac{\nu_1^2}{\varepsilon} - \varepsilon + \varepsilon \le 0 \Rightarrow \frac{\nu_1^2}{\varepsilon} \le 0$, which implies that $\nu_1 = 0$. Therefore, $x = \nu_2$ and $3 \le x \le 4$, which checks out. As such, our reformulation is an exact representation of the original disjunctive problem.

Our final example highlights another nice feature of the condition in (1). If (1) holds, then we can write an equivalent algebraic reformulation to the disjunctive convex set $F = \underset{j\in J}{\cup} C_j$ using the formulation in [1] even when some (or all) of the sets C_j are empty.

Example 5. Let $F = C_1 \cup C_2$, where $C_1 \equiv \left\{x \in \mathbb{R} \,\middle|\, x^2 - 1 \le 0\right\}$ and $C_2 \equiv \left\{x \in \mathbb{R} \,\middle|\, \begin{matrix} 3 \le x \le 4 \\ 5 \le x \le 6 \end{matrix}\right\}$.

The feasible regions of sets C_1 and C_2 are plotted below:



Clearly, the set C_2 is empty. However, the condition in (1) holds even for C_2 since $\left\{ x \in \mathbb{R} \left| \begin{matrix} (x-4)-(-4) \le 0 \cap (-x+3)-(3) \le 0 \\ (x-6)-(-6) \le 0 \cap (-x+5)-(5) \le 0 \end{matrix} \right. \right\} = \left\{ x \in \mathbb{R} \left| \begin{matrix} x \le 0 \cap x \ge 0 \\ \| x \le 0 \cap x \ge 0 \end{matrix} \right. \right\}$, which implies that $x = 0$ (we've already seen that (1) holds for C_1 per Example 4) . Now if we generate the $eps-MIP\ F(\varepsilon)$ formulation of this disjunctive set for some $0 < \varepsilon < 1$, we get (after simplifying):

$$x = \nu_1 + \nu_2$$

$$((1-\varepsilon)\lambda_1 + \varepsilon)\left[\left(\frac{\nu_1}{(1-\varepsilon)\lambda_1 + \varepsilon}\right)^2 - 1\right] - \varepsilon(-1)(1-\lambda_1) \le 0$$

$$3(1-\lambda_1) \le \nu_2 \le 4(1-\lambda_1)$$

$$5(1-\lambda_1) \le \nu_2 \le 6(1-\lambda_1)$$

$$\lambda_1 \in \{0,1\}.$$

When $\lambda_1 = 1$, then $\nu_2 = 0$; thus $x = \nu_1$ and $x^2 - 1 \le 0$, which checks out. When $\lambda_1 = 0$, $\varepsilon\left[\left(\frac{\nu_1}{\varepsilon}\right)^2 - 1\right] + \varepsilon \le 0 \Rightarrow \frac{\nu_1^2}{\varepsilon} - \varepsilon + \varepsilon \le 0 \Rightarrow \frac{\nu_1^2}{\varepsilon} \le 0$, which implies that $\nu_1 = 0$. Therefore, $x = \nu_2$ and $3 \le x \le 4$ and $5 \le x \le 6$, which is empty (and therefore checks out). As such, our reformulation is an exact representation of the original disjunctive problem.

Bibliography

1. Furman K., Sawaya N., Grossmann I, "A Computationally Useful Algebraic Representation of Nonlinear Disjunctive Convex Sets Using the Perspective Function", *submitted*; preprint available at:

2. Hiriart-Urruty J. and C. Lemaréchal, "Fundamentals of Convex Analysis", 2nd edition, Springer-Verlag, (2004).