On convex envelopes and underestimators for bivariate functions

Marco Locatelli

Dipartimento di Ingegneria Informatica, Università di Parma
Via G.P. Usberti, 181/A, 43124 Parma, Italy
e-mail: locatell@ce.unipr.it

Fabio Schoen

Dipartimento di Sistemi e Informatica, Università di Firenze
Via di Santa Marta, 3, 50139 Firenze, Italy
e-mail: fabio.schoen@unifi.it

Abstract

In this paper we discuss convex underestimators for bivariate functions. We first present a method for deriving convex envelopes over the simplest two-dimensional polytopes, i.e., triangles. Next, we propose a technique to compute the value at some point of the convex envelope over a general two-dimensional polytope, together with a supporting hyperplane of the convex envelope at that point. Noting that the envelopes might be of quite complicated form and their computation not easy, we move later on to the discussion of a method, based on the solution of a semidefinite program, to derive convex underestimators (not necessarily convex envelopes) of simple enough form for bivariate quadratic functions over general two-dimensional polytopes.

KEYWORDS: convex envelopes, convex underestimators, semidefinite programming, quadratic problems.

1 Introduction

Convex underestimators of a nonconvex function f over some region X are very important in the development of branch-and-bound techniques for the solution of global optimization problems. The best (largest) possible of such underestimators is called convex envelope of f over X. Unfortunately, the derivation of the convex envelope is not a simple task. For instance, in [6] it is proved that finding the convex envelope of a multilinear function (linear combination of products of variables) over the unit hypercube is NP-hard, although envelopes can be computed for special cases (see, e.g., [7, 16, 17, 19]). In general, the difficulty is related to the nonlinear function f to be underestimated and/or the region X over which we want to compute the envelope (see also a similar comment, e.g., in [9]). Usually, we are not able to find the envelope of f over the region X we are interested at, but it is sometimes possible to find the envelope over a region $M \supseteq X$. Of course, the closer is M to X, the better is the approximation delivered by the envelope. Usual choices for M are (hyper)rectangles. For the bilinear function xy convex and concave envelopes over rectangles have been derived in [14] and later validated in [2]. For the fractional term y/x in [24] the concave envelope over a rectangle in the positive orthant has been derived, while in [20] the result has been extended also to rectangles which do not lie in the positive orthant. In [18] convex envelopes of the bilinear term over special polytopes, the D-polytopes, have been derived. D-polytopes are characterized by the fact that they do not contain edges with finite positive slope. In [4] the envelopes of the fractional term over special bounded quadrilaterals over $\mathbb{R}^2$ (parallelograms and trapezoids) have been derived. In [13] the convex envelope of the bilinear term over triangles with a single edge having positive slope is presented, while in [3] a representation of the convex envelope over general triangles through doubly nonnegative matrices (i.e., matrices which are both semidefinite and nonnegative) is given. Some literature about convex envelopes is dedicated to the detection of the cases with a polyhedral convex envelope, i.e., cases where the convex envelope is the maximum of a finite number of affine functions. In [15, 23] edge-concave functions over a polytope P (i.e., functions which are concave on all segments in P parallel to an edge of P) are discussed and it is proved that they have a polyhedral convex envelope. Another relevant result about polyhedral convex envelopes is presented in [16] and will be discussed later. Other results about convex envelopes can be found in the literature (e.g., in [12] the convex envelope for monomials with odd degree is derived) but those discussed up to now already show that convex envelopes can be computed only for simple enough functions f and regions M. Even when they are computable, they might be of quite complicated form. Then we consider the following question. If the computation of the convex envelope of a function f over some region M is too hard and/or the functional form of the envelope is too complicated, what could we do? An obvious answer is to give up the requirement of finding the convex envelope and being satisfied with a convex underestimator of f over M. Reasonable requirements for such underestimator are the following:

- should be simple enough to compute;

- should be close enough to the function f;
- M should approximate as tightly as possible the region X at which we are interested;
- should be of simple enough form.

For instance, in [1] it is proposed to underestimate a general nonconvex term over a box by summing up to it a convex quadratic function (nonpositive over the box), so that the positive eigenvalues of the quadratic term compensate the negative ones of the nonconvex term, making the result of the sum a convex underestimator.
The paper is organized as follows. In Section 2 we will propose a way to compute the convex envelope of some bivariate functions over quite simple domains like triangles and show that for some triangles (with two or three edges along which the function is convex) the functional form of the envelope is not simple at all. In Section 3 we further show that the computation of the value at some point of the convex envelope over a general two-dimensional polytope, together with a supporting hyperplane of the convex envelope at that point, requires the solution of a number of three-dimensional convex subproblems which might increase exponentially with the number of edges along which the function is convex. In Section 4 we propose convex undestimators for indefinite bivariate quadratic functions over general two-dimensional polytopes. These will be called Quadratic Convex Underestimators (QCUs). Their computation will involve the solution of a semidefinite program whose size increases with the number of vertices of the polytopes.

2 Convex envelopes of bivariate functions over triangles

In this section we propose a method to derive the convex envelope over triangles for bivariate functions satisfying some conditions. Let T be a triangle with vertices V_1, V_2, V_3. Let $f(x, y)$ be a bivariate function. We assume the following conditions are satisfied:

Condition 1 the Hessian of f is indefinite in the interior of the triangle;

Condition 2 the restriction of f along each edge of the triangle is either concave or strictly convex;

Condition 3 if f is strictly convex over all the three edges, then there exist two edges such that f is also strictly convex along each segment joining two points belonging to such edges.

For instance, function $f(x,y) = xy$ satisfies the above conditions for all possible triangles T. In particular, in this case edges along which the function is strictly convex are those lying along lines with positive slope and the two edges satisfying Condition 3 are those forming an obtuse angle. The bivariate function $f(x,y) = x\log(1+y)$ satisfies the above conditions, e.g., over the triangle T with

$$V_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad V_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad V_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

We recall that, according to Caratheodory's theorem, given a polytope $P \subset \mathbb{R}^n$ and a function f, the convex envelope of f at a point $K \in P$ is defined as follows

$$\begin{aligned} CE_{f,P}(K) \quad = \min \quad & \left\{ \sum_{i=1}^{n+1} \lambda_i f(Q_i) \ : \ Q_i \in P, \ i = 1, \ldots, n+1, \right. \\ & \left. \sum_{i=1}^{n+1} \lambda_i = 1, \ \sum_{i=1}^{n+1} \lambda_i Q_i = K, \ \lambda_i \geq 0 \right\}. \end{aligned}$$

A relevant subset of P is the so called *generating set* denoted by $X(f)$ and defined as the smallest subset of P such that

$$\begin{aligned} CE_{f,P}(K) \quad = \min \quad & \left\{ \sum_{i=1}^{n+1} \lambda_i f(Q_i) \ : \ Q_i \in X(f), \ i = 1, \ldots, n+1, \right. \\ & \left. \sum_{i=1}^{n+1} \lambda_i = 1, \ \sum_{i=1}^{n+1} \lambda_i Q_i = K, \ \lambda_i \geq 0 \right\}, \end{aligned} \tag{1}$$

i.e., the convex envelope of f over P is equal to its convex envelope over $X(f)$.

The simplest case to deal with is when the convex envelope over a triangle is polyhedral, i.e., $X(f)$ is equal to the set of vertices of the triangle. A fundamental result, which we report here for sake of completeness, is the following:

Theorem 2.1 (from Theorem 1.2 in [16]) *: Let $f(x)$ be a lower semi-continuous function on a compact polytope P and for every point x_0 which is not a vertex there exists a line ℓ_x such that $f(x)$ is a concave function in a neighborhood of x_0 on a segment $(\ell_x \bigcap P)$ and $x_0 \in ri[\ell_x \bigcap P]$. Then $conv_P f(x)$ is a polyhedral function [...].*

Therefore, under the given conditions, this theorem clearly defines the cases where the convex envelope over a triangle is polyhedral: this holds when f is concave over all the three edges of the triangle. But the theorem actually tells us something more. Indeed, its proof can be employed to show that the generating set $X(f)$ is made up by all the vertices of the triangle plus all the edges along which the function f is convex and not concave (see also [21]). Therefore, besides the simplest polyhedral case we need to deal with three more cases, those where the number of edges along which f is convex and not concave is equal respectively to 1, 2 or 3. The analysis of such cases is strictly related to the one in [11] and [20]. Both papers deal with multivariate functions, but here we describe their contents only for bivariate functions. In [11] techniques to derive convex envelopes of 1-convex bivariate functions over boxes are discussed. 1-convex functions are bivariate functions $f(x, y)$ whose Hessian is always indefinite but which are convex with respect to one of the two variables when the other one is fixed. In [20] the convex envelope over boxes is derived for bivariate functions $f(x, y)$ which are convex with respect to x when y is fixed, and concave with respect to y when x is fixed. In particular, the results are applied to derive the convex envelope of the fractional function y/x over a box. The results of both papers are based on the fact that for functions satisfying the given assumptions over boxes, it turns out that the minimum in (1) for a point $K \notin X(f)$, and belonging to the box, is attained with just two points Q_1, Q_2 belonging to $X(f)$. Here we prove that, for the three cases mentioned above, the same holds for bivariate functions satisfying Conditions 1-3 over triangles. This is the result of the following lemma.

Lemma 2.1 *Given a triangle with vertices V_1, V_2, V_3, let ℓ be the number of edges along which the function is convex and not concave. Then the minimum in (1) for a point $K \notin X(f)$ and belonging to the triangle is attained:*

1. *with $Q_i = V_i$, $i = 1, 2, 3$, i.e. points Q_i's are the three vertices belonging to $X(f)$ for $\ell = 0$;*

2. *with just two points $Q_1, Q_2 \in X(f)$, one being the vertex opposite to the convex edge and the other lying along the convex edge itself for $\ell = 1$;*

3. *with just two points $Q_1, Q_2 \in X(f)$, one belonging to a convex edge and the other lying along the other convex edge for $\ell = 2$;*

4. *with just two points $Q_1, Q_2 = X(f)$, one belonging to one of the two edges satisfying Condition 3 and the other belonging to the third edge.*

Proof.

1. immediately follows from Theorem 2.1.

2. the minimum in (1) is attained for (at most) three points lying in $X(f)$. First of all we notice that the three points can not all lie along the convex edge, because no convex combination of three points along the edge can return a point outside the edge. Therefore, one of the points must be in the unique point in $X(f)$ outside the convex edge, i.e., the vertex opposite to the convex edge, say V_1. Next, let us assume that the other two points, denoted by Q' and Q'', lie along the convex edge. Then, we must have that

$$K = \lambda_1 Q' + \lambda_2 Q'' + \lambda_3 V_1 \quad \text{and} \quad CE_{f,T}(K) = \lambda_1 f(Q') + \lambda_2 f(Q'') + \lambda_3 f(V_1),$$

with $\lambda_1, \lambda_2, \lambda_3 > 0$. But if we consider the point

$$Q = \frac{\lambda_1}{\lambda_1 + \lambda_2} Q' + \frac{\lambda_2}{\lambda_1 + \lambda_2} Q'',$$

it holds that

$$K = (\lambda_1 + \lambda_2) Q + \lambda_3 V_1$$

and, in view of the convexity of f along the edge,

$$(\lambda_1 + \lambda_2) f(Q) \leq \lambda_1 f(Q') + \lambda_2 f(Q''),$$

so that we can replace the two points Q', Q'' with the single one Q belonging to the convex edge.

3. in case we have exactly three points, at least two of them must belong to the same edge. As in the proof of Point 2, convexity of f over such edge allows to substitute the two points along the edge with their convex combination, so that the conditions of the lemma are satisfied.

4. by the standard argument already employed for the proofs of Points 2 and 3, we know that the three points, say Q', Q'', Q''', must all lie on distinct edges. Now, let Q' and Q'' lie along the two edges satisfying Condition 3. Then, function f is strictly convex along the line through Q' and Q'', so that we could substitute these two points with their convex combination, contradicting the optimality of the three points Q', Q'', Q'''. Therefore, we can restrict our attention to just two points, one of which should lie along one of the two edges satisfying Condition 3 and the other belonging to the third edge.

□

The above remark can be employed to derive the convex envelope of function f over triangles. It will turn out that when $\ell = 0, 1$, for a given point within the triangle, we will be able to immediately identify the points giving the minimum value in (1), while for the cases $\ell = 2, 3$ the identification of the points will require the solution of a one-dimensional problem.

Theorem 2.2 *For a given point K in the triangle, it holds that the convex envelope of f at K is:*

$\ell = 0$ *the value at K of the unique affine function whose value at the three vertices are equal to the corresponding values of f at the vertices;*

$\ell = 1$ *the value*

$$\lambda^* f(V_1) + (1 - \lambda^*) f(H),$$

where H is the point at the intersection of the convex edge and the line through V_1 and K and $\lambda^* \in [0, 1]$ is the unique value such that

$$K = \lambda^* V_1 + (1 - \lambda^*) H.$$

$\ell = 2, 3$ *the solution of a properly defined minimization problem of a one-dimensional function.*

Proof.

$\ell = 0$ this case is well known and does not need to be proved.

$\ell = 1$ The result is an immediate consequence of Point 2 in Lemma 2.1, stating that the minimum of (1) is attained at points V_1 and H, where H is the unique point along the convex edge lying along the line through V_1 and K.

$\ell = 2$ let $\bar{V_1 V_3}$ and $\bar{V_2 V_3}$ be the two edges along which function f is convex. There exist unique values $\lambda_1, \lambda_2, \lambda_3 \geq 0$, $\sum_{i=1}^{3} \lambda_i = 1$, such that

$$K = \lambda_1 V_1 + \lambda_2 V_2 + \lambda_3 V_3.$$

In view of Point 3 in the Lemma 2.1, it holds that the optimal value of (1) is attained at two points

$$Q = \mu V_1 + (1 - \mu) V_3 \quad \mu \in [0, 1]$$

$$Q' = \alpha V_2 + (1-\alpha)V_3 \quad \alpha \in [0,1]$$

such that $K \in \bar{QQ'}$, i.e.

$$K = \xi Q + (1-\xi)Q' \quad \xi \in [0,1].$$

By substitution we get

$$\begin{aligned} K &= \xi(\mu V_1 + (1-\mu)V_3) + (1-\xi)(\alpha V_2 + (1-\alpha)V_3) = \\ &= \mu\xi V_1 + (1-\xi)\alpha V_2 + [(1-\mu)\xi + (1-\xi)(1-\alpha)]V_3 \end{aligned}$$

By the uniqueness of the λ_i values, it must hold that

$$\begin{aligned} \xi\mu &= \lambda_1 \\ \alpha(1-\xi) &= \lambda_2 \end{aligned}$$

and, consequently

$$\begin{aligned} \xi &= \lambda_1/\mu \\ \alpha &= \frac{\mu\lambda_2}{\mu-\lambda_1} \end{aligned}$$

Note that

$$\xi \in [0,1] \;\Rightarrow\; \mu \geq \lambda_1, \qquad \alpha \in [0,1] \;\Rightarrow\; \mu \geq \lambda_1/(1-\lambda_2).$$

Then, if we want to detect the value of the convex envelope at K, we need to solve the following one-dimensional minimization problem

$$\min_{\mu \in [\lambda_1/(1-\lambda_2),1]} \quad \frac{\lambda_1}{\mu} f(Q) + \left(1 - \frac{\lambda_1}{\mu}\right) f(Q'),$$

or, equivalently

$$\begin{aligned} \min_{\mu \in [\lambda_1/(1-\lambda_2),1]} \quad & \frac{\lambda_1}{\mu} f\left(\mu V_1 + (1-\mu)V_3\right) + \\ & + \left(1 - \frac{\lambda_1}{\mu}\right) f\left(\frac{\mu\lambda_2}{\mu-\lambda_1} V_2 + \left(1 - \frac{\mu\lambda_2}{\mu-\lambda_1}\right) V_3\right). \end{aligned}$$

$\ell = 3$ Let $\bar{V_1V_2}$ and $\bar{V_2V_3}$ be the two edges satisfying Condition 3. As before, there exist unique values $\lambda_1, \lambda_2, \lambda_3 \geq 0$, $\sum_{i=1}^{3} \lambda_i = 1$, such that

$$K = \lambda_1 V_1 + \lambda_2 V_2 + \lambda_3 V_3.$$

In view of Point 4 in Lemma 2.1, it holds that the optimal value of (1) is attained at two points, one belonging to the edge $V_1\bar{V}_3$ different from the two satisfying Condition 3

$$Q = \mu V_1 + (1-\mu)V_3 \quad \mu \in [0,1],$$

the other one either belonging to the edge $V_2\bar{V}_3$:

$$Q' = \alpha V_2 + (1-\alpha)V_3 \quad \alpha \in [0,1]$$

or to the edge $V_1\bar{V}_2$:

$$Q'' = \beta V_1 + (1-\beta)V_2 \quad \beta \in [0,1]$$

Therefore, either $K \in Q\bar{Q}'$, i.e.

$$K = \xi Q + (1-\xi)Q' \quad \xi \in [0,1]$$

or $K \in Q\bar{Q}''$, i.e.

$$K = \eta Q + (1-\eta)Q'' \quad \eta \in [0,1].$$

In particular, by taking the intersection of $V_1\bar{V}_3$ with the line through V_2 and K, it holds that

$$K \in Q\bar{Q}' \quad \Rightarrow \quad \mu \in [\lambda_1/(1-\lambda_2), 1]$$

while

$$K \in Q\bar{Q}'' \quad \Rightarrow \quad \mu \in [0, \lambda_1/(1-\lambda_2)]$$

The case $K \in Q\bar{Q}'$ leads to a result completely analogous to what already seen in Point 3 and leads to the following one-dimensional function

$$g_1(\mu) = \frac{\lambda_1}{\mu} f\left(\mu V_1 + (1-\mu)V_3\right) + \left(1 - \frac{\lambda_1}{\mu}\right) f\left(\frac{\mu\lambda_2}{\mu-\lambda_1}V_2 + \left(1 - \frac{\mu\lambda_2}{\mu-\lambda_1}\right)V_3\right)$$

The case $K \in Q\bar{Q}''$ can be dealt with in a completely similar way. Indeed, by substitution we get

$$\begin{aligned} K &= \eta(\mu V_1 + (1-\mu)V_3) + (1-\eta)(\beta V_1 + (1-\beta)V_2) = \\ &= [\eta\mu + (1-\eta)\beta]V_1 + (1-\eta)(1-\beta)V_2 + (1-\mu)\eta V_3 \end{aligned}$$

$$\begin{aligned} \eta &= \lambda_3/(1-\mu) \\ \beta &= \frac{\lambda_1 - \mu\lambda_1 - \mu\lambda_3}{1-\mu-\lambda_3}, \end{aligned}$$

which leads to the one-dimensional function

$$\begin{aligned} g_2(\mu) &= \tfrac{\lambda_3}{1-\mu} f\left(\mu V_1 + (1-\mu)V_3\right) + \\ & \left(1 - \tfrac{\lambda_3}{1-\mu}\right) f\left(\tfrac{\lambda_1 - \mu\lambda_1 - \mu\lambda_3}{1-\mu-\lambda_3} V_1 + \tfrac{(1-\mu)\lambda_2}{1-\mu-\lambda_3} V_2\right) \end{aligned}$$

After defining the one-dimensional function

$$g(\mu) = \begin{cases} g_2(\mu) & \mu \in [0, \lambda_1/(1-\lambda_2)] \\ g_1(\mu) & \mu \in [\lambda_1/(1-\lambda_2), 1], \end{cases}$$

we have that the convex envelope of f at K is equal to the solution of the following one-dimensional problem:

$$\min_{\mu \in [0,1]} g(\mu).$$

□

Examples

In what follows we will derive some convex envelopes exploiting Theorem 2.2.

$f(x,y) = xy$**, vertices of** T**:** $V_1 = (1\ 0)$, $V_2 = (0\ 0)$, $V_3 = (1\ 1)$

First of all we remark that the formula of the convex envelope of xy for triangles with $\ell = 1$ has been already given in [13]. We derive it again here using Theorem 2.2. Given a point $K = (x, y) \in T$, it holds that

$$K = (1-\lambda)V_1 + \lambda Q, \quad \lambda \in [0,1],$$

for some $Q \in \bar{V_2 V_3}$, i.e.

$$Q = (1-\mu)V_2 + \mu V_3, \quad \mu \in [0,1].$$

In view of Theorem 2.2 it holds that

$$CE_{f,T}(K) = (1-\lambda)f(V_1) + \lambda f(Q) = \lambda\mu^2$$

Imposing

$$K = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (1-\lambda) + \lambda\mu \\ \lambda\mu \end{pmatrix},$$

we have that

$$\lambda = 1 + y - x, \quad \mu = \frac{y}{1 + y - x},$$

so that

$$CE_{f,T}(K) = \frac{y^2}{1 + y - x}.$$

It is worthwhile to remark that in the approach based on triangular subdivisions proposed in [13], Linderoth did not use the above convex envelope formula but a polyhedral approximation of it, in order to avoid the lower reliability of nonlinear solvers (see also the comment in the following Section 3). This suggests that one should employ convex underestimators which are not only good approximations of the function at hand, but which are also of simple enough form.

$f(x, y) = xlog(1 + y)$**, vertices of** T**:** $V_1 = (1\ 0), \quad V_2 = (0\ 0), \quad V_3 = (1\ 1)$

As in the previous example, we are in a case with $\ell = 1$. For a point $K = (x, y) \in T$, it holds that

$$K = (1 - \lambda)V_1 + \lambda Q, \quad \lambda \in [0, 1],$$

for some $Q \in \bar{V_2 V_3}$, i.e.

$$Q = (1 - \mu)V_2 + \mu V_3, \quad \mu \in [0, 1].$$

In view of Theorem 2.2 it holds that

$$CE_{f,T}(K) = (1 - \lambda)f(V_1) + \lambda f(Q) = \lambda\mu \log(1 + \mu)$$

Imposing

$$K = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (1 - \lambda) + \lambda\mu \\ \lambda\mu \end{pmatrix},$$

we have that

$$\lambda = 1 + y - x, \quad \mu = \frac{y}{1 + y - x},$$

so that

$$CE_{f,T}(K) = y \log \left(1 + \frac{y}{1 + y - x}\right).$$

$f(x, y) = xy$**, vertices of** T**:** $V_1 = (0\ 1), \quad V_2 = (0\ 0), \quad V_3 = (2\ 2)$

In this case $\ell = 2$ holds. As seen in the proof of Theorem 2.2, we need to solve the following one-dimensional problem

$$\begin{aligned}\min_{\mu\in[\lambda_1/(1-\lambda_2),1]} \quad & \frac{\lambda_1}{\mu} f\left(\mu V_1 + (1-\mu)V_3\right) + \\ & + \left(1 - \frac{\lambda_1}{\mu}\right) f\left(\frac{\mu\lambda_2}{\mu-\lambda_1} V_2 + \left(1 - \frac{\mu\lambda_2}{\mu-\lambda_1}\right) V_3\right),\end{aligned}$$

where

$$\lambda_1 = y - x, \quad \lambda_2 = 1 - y + x/2.$$

Elementary but quite tedious computations show that the result is the following function

$$CE_{f,T}(K) = \begin{cases} \frac{x^2}{1-y+x} & \frac{\sqrt{2}}{2}x + y < 1 \\ (3-2\sqrt{2})x^2 + (6-4\sqrt{2})y^2 + (6\sqrt{2}-8)xy- & \\ -(4\sqrt{2}-6)x + (4\sqrt{2}-6)y & \text{otherwise} \end{cases}$$

3 Convex envelope of bivariate functions over polytopes

Let us consider a polytope $P \subset \mathbb{R}^2$ and a bivariate function $f(x, y)$ satisfying Conditions 1 and 2 of the previous section, with the triangle replaced by polytope P. Moreover, we require that $f \in \mathcal{C}^2$. Note that these conditions are satisfied by well known bivariate functions such as the bilinear one xy and the ratio one y/x (in the latter case we require that P lies in the interior of the half-space $x \geq 0$, or in the interior of the half-space $x \leq 0$). We would like to compute the value of the convex envelope of f over P at some point $(x_0, y_0) \in P$ and to return a supporting hyperplane for the convex envelope at that point. Knowledge of a supporting hyperplane is quite relevant for the definition of a polyhedral convex underestimator of the convex envelope (maximum of the supporting hyperplanes at a finite set of points in P). As remarked in [22], the use of polyhedral underestimators of the convex envelope rather than the convex envelope itself is advisable in view of the higher speed and stability of linear programming solvers with respect to nonlinear ones.

Among the different definitions of convex envelope of f over P, we also have the one of pointwise supremum of the underestimating affine functions of f over P. Therefore, given some point $(x_0, y_0) \in P$, the value of the convex envelope of f over P is given by the solution of the following optimization problem in the three variables a, b and c with an infinite number of constraints

$$\begin{aligned} CE_{f,P}(x_0, y_0) = \quad & \max \ c \\ & f(x, y) - [a(x - x_0) + b(y - y_0) + c] \geq 0 \quad \forall \ (x, y) \in P. \end{aligned}$$

The infinite number of constraints can be substituted by a single one involving, however, a further optimization problem

$$\begin{aligned} CE_{f,P}(x_0, y_0) = \quad & \max \ c \\ & \min_{(x,y) \in P} f(x, y) - [a(x - x_0) + b(y - y_0) + c] \geq 0 \end{aligned}$$

The optimal solution (a^*, b^*, c^*) of this problem defines a supporting hyperplane for the convex envelope of f over P at point (x_0, y_0).
In view of Condition 1, the minimum of $f(x, y) - [a(x - x_0) + b(y - y_0) + c]$ can not be attained (only) in the interior of P, and is always attained at a vertex of P or along an edge of P such that the restriction of f along the edge is a convex (and not concave, i.e., not affine) function. Therefore, the constraint

$$\min_{(x,y) \in P} f(x, y) - [a(x - x_0) + b(y - y_0) + c] \geq 0$$

can be rewritten as

$$\begin{aligned} & f(x_{v_i}, y_{v_i}) - [a(x_{v_i} - x_0) + b(y_{v_i} - y_0) + c] \geq 0 \qquad \forall \ (x_{v_i}, y_{v_i}) \in V(P) \\ & \min_{(x,y) \in e_j} f(x, y) - [a(x - x_0) + b(y - y_0) + c] \geq 0 \quad \forall \ e_j \in E'(P) \end{aligned}$$

where $V(P)$ denotes the vertex set of P, while $E'(P)$ denotes the set of edges of P along which f is strictly convex. More precisely, we can substitute the requirement for all the vertices in $V(P)$ with that for all the vertices in $V'(P) \subseteq V(P)$, where $V'(P)$ is the set of vertices which do *not* belong to edges in $E'(P)$. The constraints related to the vertices in $V'(P)$ are simple linear ones with respect to the unknowns a, b and c. Let us now consider the constraints related to the edges in $E'(P)$ which still involve optimization problems. For some $e_j \in E'(P)$, let us assume that the edge belongs to the line

$$y = mx + q.$$

Then, in view of the strict convexity assumption for f over the edge, the following condition holds

$$f''_{e_j}(x) > 0 \quad \forall \ x \in [x_1^{e_j}, x_2^{e_j}], \tag{2}$$

where $f_{e_j}(x) = f(x, mx + q)$ denotes the restriction of f along the edge $e_j \in E'(P)$, while $x_1^{e_j}, x_2^{e_j}$, $x_1^{e_j} < x_2^{e_j}$, denote the x-coordinates of the two vertices in $V(P)$ defining e_j. To be more precise, we should also consider the case where the edge lies along a line $x = \beta$ for some constant β, but this can be dealt with in a completely analogous way. The constraint related to e_j is then

$$\min_{x_1^{e_j} \le x \le x_2^{e_j}} f_{e_j}(x) - [a(x - x_0) + b(mx + q - y_0) + c] \ge 0 \tag{3}$$

Let us denote by $s(a, b)$ the minimum point of

$$f_{e_j}(x) - (a + bm)x. \tag{4}$$

We also allow for $s(a, b) = +\infty$ $(-\infty)$ if the function is decreasing (increasing). In particular, note that if $s(a, b) \neq \pm\infty$, then

$$f'_{e_j}(s(a, b)) = a + bm. \tag{5}$$

Therefore, the minimum point in (3) is

$$\begin{cases} x_1^{e_j} & \text{if } s(a,b) < x_1^{e_j} \\ x_2^{e_j} & \text{if } s(a,b) > x_2^{e_j} \\ s(a,b) & \text{otherwise} \end{cases}$$

with the minimum value

$$\begin{cases} f_{e_j}(x_1^{e_j}) - (a + bm)x_1^{e_j} + ax_0 + by_0 - bq - c & \text{if } s(a,b) < x_1^{e_j} \\ f_{e_j}(x_2^{e_j}) - (a + bm)x_2^{e_j} + ax_0 + by_0 - bq - c & \text{if } s(a,b) > x_2^{e_j} \\ f_{e_j}(s(a,b)) - (a + bm)s(a,b) + ax_0 + by_0 - bq - c & \text{otherwise} \end{cases}$$

Taking into account that, in view of (2) the first derivative of f_{e_j} is not decreasing along the interval $[x_1^{e_j}, x_2^{e_j}]$, then we can rewrite the above minimum value also as

$$\begin{cases} f_{e_j}(x_1^{e_j}) - (a + bm)x_1^{e_j} + ax_0 + by_0 - bq - c & \text{if } f'_{e_j}(x_1^{e_j}) - (a + bm) \ge 0 \\ f_{e_j}(x_2^{e_j}) - (a + bm)x_2^{e_j} + ax_0 + by_0 - bq - c & \text{if } f'_{e_j}(x_2^{e_j}) - (a + bm) \le 0 \\ f_{e_j}(s(a,b)) - (a + bm)s(a,b) + ax_0 + by_0 - bq - c & \text{otherwise} \end{cases}$$

Therefore, each constraint related to an edge $e_j \in E'(P)$ needs to be split into three different sets of constraints: one pair of linear constraints

$$f_{e_j}(x_1^{e_j}) - (a + bm)x_1^{e_j} + ax_0 + by_0 - bq - c \ge 0$$
$$f'_{e_j}(x_1^{e_j}) - (a + bm) \ge 0$$

another pair of linear constraints

$$f_{e_j}(x_2^{e_j}) - (a+bm)x_2^{e_j} + ax_0 + by_0 - bq - c \geq 0$$
$$f'_{e_j}(x_2^{e_j}) - (a+bm) \leq 0$$

and a third set with two linear constraints and a further constraint whose nature has to be clarified

$$f_{e_j}(s(a,b)) - (a+bm)s(a,b) + ax_0 + by_0 - bq - c \geq 0$$
$$f'_{e_j}(x_1^{e_j}) - (a+bm) \leq 0$$
$$f'_{e_j}(x_2^{e_j}) - (a+bm) \geq 0$$

Taking into account that such splitting into three groups of constraints needs to be done for all edges in $E'(P)$, if we denote by t the cardinality of $E'(P)$, we see that we need to solve (at most) 3^t three-dimensional subproblems in order to compute the value of the convex envelope of f over P at (x_0, y_0). What we still need to clarify is the nature of the constraints

$$f_{e_j}(s(a,b)) - (a+bm)s(a,b) + ax_0 + by_0 - bq - c \geq 0. \tag{6}$$

What we will prove in the following theorem is that such constraints are convex ones, so that all the subproblems to be solved in order to compute the convex envelope are convex ones.

Theorem 3.1 *Under the given conditions, function*

$$g(a,b) = f_{e_j}(s(a,b)) - (a+bm)s(a,b)$$

is concave, so that constraint (6) is a convex one.

Proof. Throughout the proof we will omit the dependency of s form a, b. Let us compute the Hessian of g. We have that

$$g_{aa} = f''_{e_j}(s)s_a^2 - 2s_a + s_{aa}[f'_{e_j}(s) - (a+bm)] = f''_{e_j}(s)s_a^2 - 2s_a$$

where the last equality follows from (5). Analogously, we have that

$$g_{bb} = f''_{e_j}(s)s_b^2 - 2ms_b + s_{bb}[f'_{e_j}(s) - (a+bm)] = f''_{e_j}(s)s_b^2 - 2ms_b,$$

$$g_{ab} = f''_{e_j}(s)s_a s_b - ms_a - s_b + s_{ab}[f'_{e_j}(s) - (a+bm)] = f''_{e_j}(s)s_a s_b - ms_a - s_b.$$

Then, the determinant of the Hessian is equal to

$$-(ms_a - s_b).$$

It turns out that the determinant is null. Indeed, deriving both members of (5) with respect to a we get

$$f''_{e_j}(s)s_a = 1, \tag{7}$$

while deriving with respect to b we get

$$f''_{e_j}(s)s_b = m,$$

so that $s_b = ms_a$ and the determinant is equal to 0. Then , in order to prove that the Hessian is semidefinite negative (i.e., that g is concave), we only need to show that $g_{aa}, g_{bb} \leq 0$. We notice that g_{aa} (similar for g_{bb}) is equal to

$$s_a[f''_{e_j}(s)s_a - 2],$$

but from (7) this is equal to $-s_a$, so that

$$g_{aa} \leq 0 \quad \Leftrightarrow \quad s_a \geq 0.$$

Since, again from (7)

$$s_a = \frac{1}{f''_{e_j}(s)}$$

and since $f''_{e_j}(s) > 0$ in view of the strict convexity of f_{e_j}, we can conclude that s_a is positive. □

In order to illustrate the technique proposed in this section we consider the following example.

$f(x, y) = y/x$**,** P **triangle with vertices:** $V_1 = (1\ 1), \quad V_2 = (1\ 2), \quad V_3 = (2\ 1)$

There is a single edge, the one joining vertices V_1 and V_3, along which f is strictly convex. Therefore, given $(x_0, y_0) \in P$, we need to solve the three subproblems

$$\begin{array}{ll} \max & c \\ & c \leq 2 - a + ax_0 - 2b + 2y_0 \\ & c \leq 1 - a + ax_0 - b + by_0 \\ & a \leq -1 \\ \max & c \\ & c \leq 2 - a + ax_0 - 2b + by_0 \\ & c \leq \frac{1}{2} - 2a + ax_0 - b + by_0 \\ & a \geq -1/4 \end{array}$$

$$\begin{aligned} \max \quad & c \\ & c \leq 2 - a + ax_0 - 2b + by_0 \\ & c \leq 2\sqrt{-a} + ax_0 - b + by_0 \\ & -1 \leq a \leq -1/4 \end{aligned}$$

Simple but tedious computations show that the optimal values of the three subproblems are, respectively,

$$y_0 + 1 - x_0, \quad \frac{5}{4}y_0 - \frac{1}{4}x_0 - \frac{1}{4}, \quad \frac{(2-y_0)^2}{x_0 - y_0 + 1} + 2(y_0 - 1),$$

and it can be proved that the maximum of these quantities over P is always the last optimal value, so that

$$CE_{f,P}(x_0, y_0) = \frac{(2-y_0)^2}{x_0 - y_0 + 1} + 2(y_0 - 1).$$

4 Quadratic Convex Underestimators

As observed in Section 2, even the convex envelope of a simple bilinear term over the simplest possible polytope, a triangle, can be of quite complicated form. Therefore, we might wonder whether we are able to define "good enough" convex undestimators which are simple enough to compute, which are of simple enough form and, finally, which are defined over more general polytopes. We will propose how to derive such underestimators for bivariate nonconvex quadratic functions

$$f(x, y) = ax^2 + by^2 + 2cxy$$

and general polytopes $P \subset \mathbb{R}^2$ with vertices $v_1, \ldots, v_t$.

In order to detect a good convex underestimator, as a first step we will move to the t-dimensional unit simplex

$$\Delta_t = \{\lambda \in \mathbb{R}^t \ : \ \sum_{i=1}^{t} \lambda_i = 1, \ \lambda \geq 0\}.$$

Then, we can define P as follows

$$P = \{V\lambda \ : \ \lambda \in \Delta_t\},$$

where $V \in \mathbb{R}^{2\times t}$ is the matrix whose columns are the coordinates of the vertices of P. Next, we rewrite function f as follows

$$h(\lambda) = \lambda^T Q \lambda$$

where

$$Q = V^T \begin{pmatrix} a & c \\ c & b \end{pmatrix} V.$$

A generic quadratic function over the t-dimensional unit simplex is the following:

$$q_{S,c,v}(\lambda) = \lambda^T[S + 1/2(ce^T + ec^T) + vE]\lambda$$

where $e \in \mathbb{R}^t$ is the vector whose components are all equal to 1, while $E \in \mathbb{R}^{t\times t}$ is the square matrix whose entries are all equal to 1. Note that over the unit simplex Δ_t this function is equal to

$$\lambda^T S\lambda + c^T\lambda + v.$$

In other words, we can always incorporate the linear terms in the matrix $1/2(ce^T + ec^T)$ and the constant ones in the matrix vE, in order to work with a pure quadratic form with no linear and constant terms. Now, in order to have a good convex underestimator we would like to choose S, c, v so that $q_{S,c,v}$ is convex, underestimates h over the unit simplex Δ_t (or, equivalently, f over P) and the volume between h and $q_{S,c,v}$ is as small as possible. Basically, we aim at solving a problem with:

variables : $S \in R^{t\times t}$, $c \in R^t$ e $v \in R$;

constraints :

- convexity, i.e., $S \succeq O$ (S is a semidefinite matrix);
- $q_{S,c,v}$ underestimates h over Δ_t. This means that

  $$\lambda^T[Q - S - 1/2(ce^T + ec^T) - vE]\lambda \geq 0 \quad \forall\, \lambda \in \Delta_t$$

  or, equivalently

  $$Q - S - 1/2(ce^T + ec^T) - vE \in \mathcal{C}_t,$$

  where $\mathcal{C}_t$ is the set of copositive matrices of order t.

objective : we would like to minimize the integral

$$\int_{\Delta_t} \lambda^T[Q - S - 1/2(ce^T + ec^T) - vE]\lambda.$$

The integral of a quadratic function over the unit simplex can be expressed in analytic form. It holds that the value of the integral defining our objective function is

$$\frac{2}{(t+1)!}E\bullet[Q-S-1/2(ce^T+ec^T)-vE] = \frac{2}{(t+1)!}\sum_{i,j=1}^{t}[Q-S-1/2(ce^T+ec^T)-vE]_{ij}.$$

Therefore, we are finally left with the following problem to be solved

$$\begin{aligned} \min \quad & E \bullet [Q - S - 1/2(ce^T + ec^T) - vE] \\ & S \succeq O \\ & Q - S - 1/2(ce^T + ec^T) - vE \in \mathcal{C}_t \end{aligned} \tag{8}$$

The copositive constraint

$$Q - S - 1/2(ce^T + ec^T) - vE \in \mathcal{C}_t,$$

makes the problem a difficult one, but we might simplify it by substituting the copositive cone with the cone obtained as a sum of the semidefinite cone and the cone of nonnegative matrices, that is

$$Q - S - 1/2(ce^T + ec^T) - vE \in \mathcal{P}_t + \mathcal{N}_t,$$

where $\mathcal{P}_t$ is the set of semidefinite matrices of order t, and $\mathcal{N}_t$ the set of nonnegative matrices of order t. In particular, for $t \leq 4$ the two cones are equivalent, while more generally it holds that

$$\mathcal{P}_t + \mathcal{N}_t \subseteq \mathcal{C}_t$$

(see, e.g., [8] for an example of copositive matrix of order 5 which can not be obtained as the sum of a semidefinite and a nonnegative matrix). Let us consider the simplified problem

$$\begin{aligned} \min \quad & E \bullet [Q - S - 1/2(ce^T + ec^T) - vE] \\ & S \succeq O \\ & Q - S - 1/2(ce^T + ec^T) - vE \in \mathcal{P}_t + \mathcal{N}_t \end{aligned} \tag{9}$$

We can prove the following result.

Observation 4.1 *The constraint*

$$Q - S - 1/2(ce^T + ec^T) - vE \in \mathcal{P}_t + \mathcal{N}_t$$

can be replaced by the constraint

$$Q - S - 1/2(ce^T + ec^T) - vE \geq O.$$

Moreover, for the optimal solution S^*, c^*, v^* it holds that

$$q_{S^*,c^*,v^*}(e_i) = h(e_i) \quad \forall\ i = 1, \ldots, t.$$

where $e_i \in \mathbb{R}^t$ is the i-th vertex of the unit simplex Δ_t.

Proof. Notice that, given a feasible solution with

$$Q - S - 1/2(ce^T + ec^T) - vE = P + N, \quad P \in \mathcal{P}_t,\ N \in \mathcal{N}_t,$$

we can obtain a further feasible solution by substituting S with $S' = S + P$, so that

$$Q - S' - 1/2(ce^T + ec^T) - vE = N, \quad N \in \mathcal{N}_t.$$

If we compute the objective (the volume) in the new feasible solution, this is reduced, with respect to the starting feasible solution, by the quantity $E \bullet P$, which is certainly nonnegative in view of the semidefiniteness of E and P. Therefore, the new feasible solution is at least as good as the starting one and we can substitute the constraint

$$Q - S - 1/2(ce^T + ec^T) - vE \in \mathcal{P}_t + \mathcal{N}_t$$

with the following constraint

$$Q - S - 1/2(ce^T + ec^T) - vE \geq O.$$

Moreover, we also notice that

$$N' = N - Diag(N) \in \mathcal{N}_t,$$

where $Diag(N)$ is the diagonal matrix whose diagonal elements are equal to those of matrix N. Then, we have that

$$Q - S' - Diag(N) - 1/2(ce^T + ec^T) - vE = N', \quad N \in \mathcal{N}_t.$$

and setting $S'' = S' + Diag(N)$ (being N nonnegative, we have that $Diag(N)$ is positive semidefinite) we still get a feasible solution with objective function not worse than the original one. In other words, we can impose the further constraint

$$diag(Q) = diag(S + 1/2(ce^T + ec^T) + vE),$$

(given a matrix A, $diag(A)$ is the vector whose components are the diagonal elements of matrix A) or, equivalently, we can impose that our convex quadratic underestimator has the same values of h at the vertices of the unit simplex (i.e., the same values of f at the vertices of P). $\square$

Now, let us assume that we solved problem (9) and that we have an optimal solution (S^*, c^*, v^*). This is a convex quadratic function which underestimates h over Δ_t. But what happens if we want to get back in the original two-dimensional space? In case P is a triangle, we do not encounter any difficulty: matrix V is invertible and we can derive a two-dimensional convex quadratic function depending on the original variables x and y, starting from function q_{S^*,c^*,v^*} over Δ_t. But what happens for polytopes with more than three vertices? In order to have a convex function which underestimates f over P, we need to use the following

$$\begin{aligned} g(x,y) = \min \quad & \lambda^T[S^* + 1/2(c^*e^T + ec^{*T}) + v^*E]\lambda \\ & V\lambda = \begin{pmatrix} x \\ y \end{pmatrix} \\ & \lambda \in \Delta_t \end{aligned} \tag{10}$$

Basically, among all possible representations of a point $\begin{pmatrix} x \\ y \end{pmatrix} \in P$, we consider the one with the minimum value of the quadratic function.

In fact, if we use (10) to define a convex underestimator, we might try to improve problem (9). Indeed, in such problem we are currently asking that q_{S^*,c^*,v^*} underestimates h *for any possible representation* $\lambda \in \Delta_t$ *of a point* $\begin{pmatrix} x \\ y \end{pmatrix} \in P$, while it would be enough to impose that only for the representation with the minimum value. In other words, a better solution can be obtained by solving the following problem

$$\begin{aligned} \min \quad & E \bullet [Q - S - 1/2(ce^T + ec^T) - vE] \\ & S \succeq O \\ & f(x,y) - \min_{\lambda \in \Delta_t \,:\, V\lambda = \begin{pmatrix} x \\ y \end{pmatrix}} \lambda^T[S + 1/2(ce^T + ec^T) + vE]\lambda \geq 0 \qquad \forall \begin{pmatrix} x \\ y \end{pmatrix} \in P \end{aligned} \tag{11}$$

Such problem appears to be more complicated to solve with respect to (9). Indeed, it has an infinite number of constraints, one for each $(x, y) \in P$. However, we can simplify it following a strategy similar to the one discussed in Section 3. First we observe that

$$f(x,y) - \min_{\lambda \in \Delta_t \,:\, V\lambda = \begin{pmatrix} x \\ y \end{pmatrix}} \lambda^T [S + 1/2(ce^T + ec^T) + vE]\lambda \geq 0 \quad \forall \begin{pmatrix} x \\ y \end{pmatrix} \in P$$

is equivalent to

$$\min_{(x,y)\in P} \left\{ f(x,y) - \min_{\lambda \in \Delta_t \,:\, V\lambda = \begin{pmatrix} x \\ y \end{pmatrix}} \lambda^T [S + 1/2(ce^T + ec^T) + vE]\lambda \right\} \geq 0.$$

The optimal solution of the minimum problem on the left-hand side of this inequality can not be a point in the interior $int(P)$ of P. Indeed

$$- \min_{\lambda \in \Delta_t \,:\, V\lambda = \begin{pmatrix} x \\ y \end{pmatrix}} \lambda^T [S + 1/2(ce^T + ec^T) + vE]\lambda$$

is a concave function and if we consider $(x, y) \in int(P)$ we are always able to find a direction along which also $f(x, y)$ is concave (unless f is already convex, but in such case we do not need any convex underestimator) . Therefore, along the segment in P lying on the direction, we have a minimum value at one of the two extremes of the segment, that is along an edge of P. In a completely analogous way, if along an edge we have that f is concave, then the minimum is attained at one of the two extremes of the edge. Therefore, we can finally rewrite the infinite set of constraints as follows

$$f(x,y) - \min_{\lambda \in \Delta_t \,:\, V\lambda = \begin{pmatrix} x \\ y \end{pmatrix}} \lambda^T [S + 1/2(ce^T + ec^T) + vE]\lambda \geq 0 \quad \forall \begin{pmatrix} x \\ y \end{pmatrix} \in X_f(P)$$

where $X_f(P)$ is the generating set of the convex envelope of f over P, i.e., the set of edges of P along which f is not concave plus the vertices of P which do not belong to edges where f is not concave. The fact that we are able to restrict the attention to points along edges of P is an advantage. Indeed, for a point along an edge its representation through points $\lambda \in \Delta_t$ is unique.

Now, let us recall the notation $V'(P)$ and $E'(P)$ introduced in Section 3. For the vertices in $V'(P)$ we need to impose the linear constraints

$$q_{kk} - s_{kk} - c_k - v \geq 0 \quad \forall\, v_k \in V'(P)$$

where q_{kk} coincides with the value of f at vertex v_k. Along the edges in $E'(P)$ we can impose the copositivity condition on the restriction of matrix

$$Q - S - 1/2(ce^T + ec^T) - vE$$

to the two rows and columns corresponding to the two vertices identifying an edge. Then, if $(v_i, v_j) \in E'(P)$, we need to impose

$$[Q - S - 1/2(ce^T + ec^T) - vE]_{ij} \in \mathcal{C}_2$$

where $[Q - S - 1/2(ce^T + ec^T) - vE]_{ij}$ is the 2×2 submatrix of $[Q - S - 1/2(ce^T + ec^T) - vE]$ formed by rows and columns i and j, while $\mathcal{C}_2$ is the copositive cone of order 2 which can be substituted by the cone $\mathcal{P}_2 + \mathcal{N}_2$ (sum of the semidefinite and nonnegative cones of order 2). Then, we can rewrite (11) as

$$\begin{array}{lll} \min & E \bullet [Q - S - 1/2(ce^T + ec^T) - vE] & \\ & S \succeq O & \\ & q_{kk} - s_{kk} - c_k - v \geq 0 & k \in V'(P) \\ & [Q - S - 1/2(ce^T + ec^T) - vE]_{ij} \in \mathcal{P}_2 + \mathcal{N}_2 & \forall\ (v_i, v_j) \in E'(P) \end{array} \tag{12}$$

First of all, it is worthwhile to make a remark. A bit counterintuitively, the optimal value of (12) can also be negative. This might be surprising because if g underestimates f, we would expect a nonnegative value of the integral. Here the trick is that in the objective we do not have the integral of $f - g$ over P but the integral of $h - q_{S^*,c^*,v^*}$ over Δ_t. In particular, if we consider the definition (10) of g, we notice that $g(x, y)$ is the *minimum* of q_{S^*,c^*,v^*} over all possible representations of a point (x, y). Then, for all other representations λ of the same point (x, y) it might happen that $h(\lambda) < q_{S^*,c^*,v^*}(\lambda)$, so that the value of the integral might be negative. Of course, a direct minimization of the integral of $f - g$ would be better, but, unfortunately, this makes the computation of the underestimator much harder. Then, it seems that minimizing the integral of $h - q_{S,c,v}$ is still a reasonable choice, allowing for a compromise between ease of computation and quality of the underestimator.

We are able to prove the following result.

Observation 4.2 *If the edges in $E'(P)$ are not adjacent (they do not have* *common vertices), then we can rewrite (12) as follows*

$$\begin{array}{lll} \min & E \bullet [Q - S - 1/2(ce^T + ec^T) - vE] & \\ & S \succeq O & \\ & q_{ii} - s_{ii} - c_i - v = 0 & \forall\ v_i \in V(P) \\ & q_{ij} - s_{ij} - 1/2(c_i + c_j) - v \geq 0 & \forall\ (v_i, v_j) \in E'(P) \end{array} \tag{13}$$

Proof. The proof is similar to that of Observation 4.1. Let us denote by (S^*, c^*, v^*) an optimal solution of the problem. It holds that

$$[Q - S^* - 1/2(c^* e^T + ec^{*T}) - v^* E]_{ij} = M_{ij} + N_{ij} \quad \forall\ (v_i, v_j) \in E'(P),$$

where $M_{ij} \in \mathcal{P}_2$ and $N_{ij} \in \mathcal{N}_2$. If we consider the matrix M'_{ij} of order t which is equal to M_{ij} when restricted to rows and columns i and j and with all the other entries equal to 0, we have that $M'_{ij} \in \mathcal{P}_t$ and we have a new optimal solution if we set

$$S' = S^* + M'_{ij}.$$

Then, we can substitute constraints

$$[Q - S - 1/2(ce^T + ec^T) - vE]_{ij} \in \mathcal{P}_2 + \mathcal{N}_2$$

with

$$[Q - S - 1/2(ce^T + ec^T) - vE]_{ij} \in \mathcal{N}_2 \quad \forall\ (v_i, v_j) \in E'(P) \tag{14}$$

With a similar proof, we can show that we can fix all the diagonal elements of matrices N_{ij} to 0. Therefore constraints (14) can be further rewritten as follows

$$\begin{array}{ll} q_{ii} - s_{ii} - c_i - v = 0 & \forall\ v_i \in V(P) \setminus V'(P) \\ q_{ij} - s_{ij} - 1/2(c_i + c_j) - v \geq 0 & \forall\ (v_i, v_j) \in E'(P) \end{array}$$

Taking into account that the constraints

$$q_{kk} - s_{kk} - c_k - v \geq 0 \quad v_k \in V'(P)$$

can be substituted by equalities (a positive value of the left-hand side can be summed up to s_{kk}), we can finally rewrite (12) as (13). □

The following easy observations also hold.

Observation 4.3 *Under the assumptions of Observation 4.2, the returned convex underestimator is equal to f at the vertices of P.*

Proof. This is an immediate consequence of the equality constraints

$$q_{ii} - s_{ii} - c_i - v = 0 \quad \forall\, v_i \in V(P).$$

$\square$

Observation 4.4 *For the polyhedral cases, i.e., those for which $E'(P) = \emptyset$, the returned convex underestimators is the convex envelope of f over P.*

Proof. We first observe that in this case

$$\bar{S} = O, \quad \bar{c}_i = q_{ii} \ \ \forall\, v_i \in V(P), \quad \bar{v} = 0$$

is a feasible solution for (13). Moreover, for any other feasible solution (S, c, v) of (13) it holds, in view of convexity, that

$$\begin{array}{rcl} q_{S,c,v}(\lambda) & = & (\sum_{i=1}^t \lambda_i e_i)^T S (\sum_{i=1}^t \lambda_i e_i) + c(\sum_{i=1}^t \lambda_i e_i) + v \\ & \leq & \sum_{i=1}^t \lambda_i (e_i^T S e_i + c e_i + v) = \sum_{i=1}^t \lambda_i (s_{ii} + c_i + v) = \sum_{i=1}^t \lambda_i q_{ii} = q_{\bar{S},\bar{c},\bar{v}}(\lambda), \end{array}$$

i.e., $(\bar{S}, \bar{c}, \bar{v})$ is also optimal for (13). The corresponding function

$$\begin{array}{rl} g(x,y) = \min & \sum_{i=1}^t q_{ii}\lambda_i \\ & V\lambda = \begin{pmatrix} x \\ y \end{pmatrix} \\ & \lambda \in \Delta_t \end{array}$$

is exactly the convex envelope. $\square$

Now, let us consider non polyhedral cases. We might wonder whether the returned solution is always at least as good as the convex envelope over the smallest box $B \supseteq P$. The following example shows that this is not always the case.

Example 4.1 *Let $f(x,y) = xy$ and T be the triangle with vertices $(0,0)$, $(0,1)$ and $(1,1)$. The smallest box B containing T is the unit square, over which the convex envelope is*

$$h(x,y) = \max\{0, x + y - 1\}.$$

With some computations we have that the solution of (13), when expressed in terms of the original variables x, y is the function

$$g(x,y) = \frac{1}{2}x^2 + \frac{1}{4}y^2 + \frac{1}{4}xy - \frac{1}{4}y + \frac{1}{4}x.$$

Even if the volume between xy and g over T is lower than that between xy and h (3/96 the former, 1/24 the latter), it does not hold that $g(x,y) \leq h(x,y)$ $\forall\ (x,y) \in T$ (e.g., this does not hold for $x = 0, y = 1$).

However, in such cases we can define a new convex underestimator ℓ taking the maximum between h and g. In the example we have that

$$\ell(x,y) = \max\left\{0, x+y-1, \frac{1}{2}x^2 + \frac{1}{4}y^2 + \frac{1}{4}xy - \frac{1}{4}y + \frac{1}{4}x\right\}.$$

We conclude this section remarking that a similar development could be extended to bivariate polynomial functions satisfying Conditions 1 and 2 of Section 2. Indeed, in such cases we might exploit along each edge in $E'(P)$ the semidefinite representation of nonnegative one-dimensional polynomials (see, e.g., [5])

5 Conclusions

In this paper we discussed convex underestimators for functions over two-dimensional polytopes. In Section 2 we proposed a method to derive the convex envelope of bivariate functions, satisfying three conditions, over the simplest two-dimensional polytopes: triangles. In Section 3 we have shown that the computation of the value at some point of the convex envelope over a general two-dimensional polytope, together with a supporting hyperplane of the convex envelope at that point, requires the solution of some three-dimensional convex subproblems. Having remarked that the convex envelope might be of quite complicated form (even over triangles), in Section 4 we proposed a method, based on the solution of a semidefinite program, to derive convex underestimators for bivariate quadratic functions over general two-dimensional polytopes.

References

[1] C. S. Adjiman, I. P. Androulakis, C. A. Floudas, A global optimization method, α-BB, for general twice-differentiable constrained NLPs - II. Implementation and Computational Results, *Comp. Chem. Eng.*, 22: 1159-1179 (1998)

[2] F. A. Al-Khayyal, J. E. Falk, Jointly constrained biconvex programming, *Mathematics of Operations Research*, 8: 273-286 (1983)

[3] K. M. Anstreicher, S. Burer, Computable representations for convex hulls of low-dimensional quadratic forms, to appear in *Mathematical Programming* (2010)

[4] H. P. Benson, On the construction of convex and concave envelope formulas for bilinear and fractional functions on quadrilaterals, *Computational Optimization and Applications*, 27: 5-22 (2004)

[5] L. Brickman, L. Steinberg, On nonnegative polynomials, *The American Mathematical Monthly*, 69: 218-221 (1962)

[6] Y. Crama, Recognition problems for polynomials in 0-1 variables, *Mathematical Programming*, 44: 139-155 (1989)

[7] Y. Crama, Concave extensions for nonlinear 0-1 maximization problems, *Mathematical Programming*, 61: 53-60 (1993)

[8] P. H. Diananda, On non-negative forms in real variables some or all of which are non-negative, *Proc. Cambridge Philos. Soc.* , 58: 17-25 (1967).

[9] R. Horst, H. Tuy, *Global optimization: deterministic approaches*, Springer Verlag, Berlin, 3rd edition (1996).

[10] T. Kuno, A branch-and-bound algorithm for maximizing the sum of several linear ratios, *Journal of Global Optimization*, 22: 155-174 (2002)

[11] M. Jach, D. Michaels, R. Weismantel, The Convex Envelope of (n–1)-Convex Functions, *SIAM Journal on Optimization*, 19(3): 1451-1466 (2008)

[12] L. Liberti, C. C. Pantelides, Convex Envelopes of Monomials of Odd Degree, *Journal of Global Optimization*, 25(2): 157-168 (2003)

[13] J. Linderoth, A simplicial branch-and-bound algorithm for solving quadratically constrained quadratic programs, *Mathematical Programming*, 103: 251-282 (2005)

[14] G. P. McCormick, Computability of global solutions to factorable nonconvex programs – Part I – Convex underestimating problems, *Mathematical Programming*, 10:147–175 (1976).

[15] C.A. Meyer, C.A. Floudas, Convex envelopes for edge-concave functions. *Mathematical Programming*, 103:207–224 (2005).

[16] A. Rikun, A convex envelope formula for multilinear functions, *Journal of Global Optimization*, 10: 425-437 (1997)

[17] H. S. Ryoo, N. V. Sahinidis, Analysis of bounds for multilinear functions, *Journal of Global Optimization*, 19: 403-424 (2001)

[18] H. D. Sherali, A. Alameddine, An explicit characterization of the convex envelope of a bivariate bilinear function over special polytopes, *Annals of Operations Research*, 27: 197-210 (1992).

[19] H. D. Sherali, Convex envelopes of multilinear functions over a unit hypercube and over special discrete sets, *Acta Mathematica Vietnamica*, 22: 245-270 (1997)

[20] M. Tawarmalani, N. V. Sahinidis, Semidefinite relaxations of fractional programs via novel convexification techniques, *Journal of Global Optimization*, 20: 137-158 (2001).

[21] M. Tawarmalani, N. V. Sahinidis, Convex extensions and envelopes of lower semicontinuous functions, *Mathematical Programming*, 93: 247-263 (2002)

[22] M. Tawarmalani, N. V. Sahinidis, A polyhedral branch-and-cut approach to global optimization, *Mathematical Programming*, 103: 225-249 (2005)

[23] F. Tardella. On the existence of polyhedral convex envelopes, in C.A. Floudas, P.M. Pardalos (eds.), *Frontiers in Global Optimization*, Dordrecht, Kluwer Academic Publishers, 563-574 (2003).

[24] J.M. Zamora, I.E. Grossmann, A Branch and Contract algorithm for problems with concave univariate, bilinear and linear fractional terms, *Journal of Global Optimization*, 14, 217-249 (1999).