On the computation of convex envelopes for bivariate functions through KKT conditions

Marco Locatelli

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

Abstract

In this paper we exploit a slight variant of a result previously proved in [11] to define a procedure which delivers the convex envelope of some bivariate functions over polytopes. The procedure is based on the solution of a KKT system and simplifies the derivation of the convex envelope with respect to previously proposed techniques. The procedure is applied to derive the convex envelope of the bilinear function xy over any polytope, and the convex envelope of functions $x^n y^m$ over boxes.

KEYWORDS: Global Optimization, Convex Envelope, KKT Conditions.

1 Introduction

The best convex underestimator of a non convex function f over some region $X \subset \mathbb{R}^n$ is called convex envelope of f over X, is denoted by $conv_{f,X}$, and is equal to the supremum of all affine underestimators of f over X, i.e., for each $\mathbf{x} \in X$

$$conv_{f,X}(\mathbf{x}) = \sup\{\mathbf{a}^T\mathbf{x} + a_0 \ : \ \mathbf{a}^T\mathbf{y} + a_0 \leq f(\mathbf{y}) \ \forall \mathbf{y} \in X\}.$$

An alternative definition of the convex envelope is the following

$$\begin{array}{rll} conv_{f,X}(\mathbf{x}) = \min & \sum_{i=1}^k \lambda_i f(\mathbf{x}_i) & \\ & \sum_{i=1}^k \lambda_i \mathbf{x}_i = \mathbf{x} & \\ & \sum_{i=1}^k \lambda_i = 1 & \\ & \boldsymbol{\lambda} \geq \mathbf{0} & \\ & \mathbf{x}_i \in X & i = 1, \ldots, k \\ & k \in \{1, \ldots, n+1\}. & \end{array} \quad (1)$$

Note that we can impose $k \leq n+1$ in view of Carathéodory's theorem.

Convex envelopes are widely used to define convex relaxations and, thus, lower bounds, of non convex problems. The literature about convex envelopes encompasses many results, ranging from general theoretical results to more specific results about certain functions.

Results about specific functions include those for the bilinear function [1, 2, 3, 10, 15, 20], the fractional function [4, 6, 22, 26], and the multilinear functions over the unit hypercube [5, 18, 19, 21]. More general results include those about polyhedral convex envelopes, i.e., convex envelopes which are the maximum of a finite number of affine functions [16, 18, 23, 24], and results about non polyhedral convex envelopes [7, 8, 9, 11, 12, 22, 25]. In this paper we discuss the derivation of non-polyhedral convex envelopes for some bivariate functions. The derivation is based on the solution of a KKT system. More specifically, the paper is structured as follows. In Section 2 we discuss a slight variant of a result proved in [11] and we introduce the notation which will be used throughout the paper. The result shows that, under suitable conditions, the convex envelope of bivariate functions over polytopes can be computed by solving a convex problem. In Section 3 we show how to employ the KKT system associated to the convex problem in order to derive the formula of the convex envelope. In this section a procedure is proposed for the computation of the convex envelope. Two special cases are discussed in the following sections. In Section 4 we exploit the results in Section 3 to derive the convex envelope of the bilinear function over general polytopes, while in Section 5 we do the same for the product of power functions over boxes. These two examples have been already discussed in previous papers. In particular, in [11] and [12] it is shown how to derive these convex envelopes through the solution of some one-dimensional problems. However, the aim of this paper is to show that the derivation of the convex envelope is simplified if the convex envelope is derived by solving the above mentioned KKT system.

2 A slight variant of a previous result

In this section we discuss a slight variant of a result proved in [11]. First, we need to introduce an assumption.

Assumption 2.1 *Let $f \in \mathcal{C}^2$ be a bivariate function and $P \subset \mathbb{R}^2$ be a polytope with vertex set $V(P)$ such that*

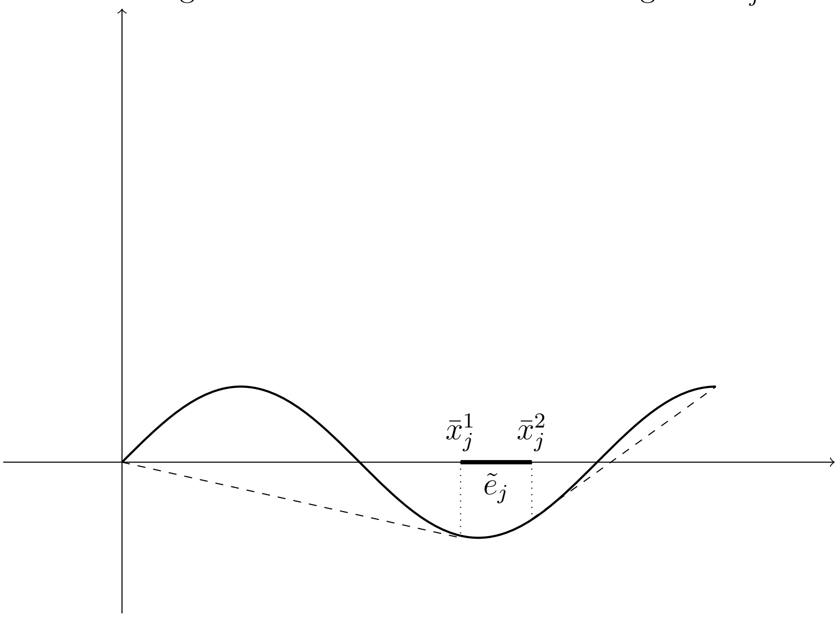
Condition 1 *the Hessian of f is indefinite in the interior of P;*

Condition 2 *the restriction of f along each edge of P is strictly convex over a nonempty (and, possibly, coinciding with the whole edge) subsegment of the edge, and is otherwise concave.*

The result proved in [11] covered the cases where each "convex" subsegment coincides with an edge of the polytope. Under the above assumption, the semi-infinite linear programming definition of the convex envelope of f over P as the pointwise supremum of the underestimating affine functions of f over P, i.e., for each $(x, y) \in P$

$$\begin{aligned} Conv_{f,P}(x,y) = \quad & \max_{a,b,c} \ c \\ & f(\chi,\gamma) - [a(\chi - x) + b(\gamma - y) + c] \geq 0 \quad \forall (\chi,\gamma) \in P, \end{aligned}$$

Figure 1: Identification of the subsegment $\tilde{e}_j$ of the edge $e_j \in \bar{E}(P)$.



can be converted into the following problem with a finite number of constraints

$$\begin{array}{lll} Conv_{f,P}(x,y) = & \max_{a,b,c}\ c & \\ & f(x_{v_i}, y_{v_i}) - [a(x_{v_i} - x) + b(y_{v_i} - y)] \geq c & \forall v_i \in V(P) \\ & \min_{(\chi,\gamma)\in e_j} f(\chi,\gamma) - [a(\chi - x) + b(\gamma - y)] \geq c & \forall e_j \in \bar{E}(P), \end{array} \quad (2)$$

where $\bar{E}(P)$ denotes the set of edges of P containing a non-empty subsegment along which f is strictly convex. For each $e_j \in \bar{E}(P)$

$$\min_{(\chi,\gamma)\in e_j} f(\chi,\gamma) - [a(\chi - x) + b(\gamma - y)] \geq c, \quad (3)$$

let $y = m_j x + q_j$ be the line along which the edge lies, and x_j^1, x_j^2 be the x-coordinates at the extreme points of the edge (edges parallel to the y-axis should be dealt with in a different, but analogous way, by considering the interval of the y-coordinates). Let

$$g_j(x) = f(x, m_j x + q_j),$$

be the restriction of f along the line. Let $\bar{x}_j^1$ be the tangent point of the tangent to g_j interpolating g_j at x_j^1, and $\bar{x}_j^2$ be the tangent point of the tangent to g_j interpolating g_j at x_j^2 (see also Figure 1). If $\bar{x}_j^1 > \bar{x}_j^2$, then the minimum on the left hand-side of (3) is attained at the vertex $(x_j^1, m_j x_j^1 + q_j)$ or at the vertex $(x_j^2, m_j x_j^2 + q_j)$. In this case (3) reduces to the pair of linear constraints

$$f(x_j^1, m_j x_j^1 + q_j) - [a(x_j^1 - x) + b(m_j x_j^1 + q_j - y)] \geq c$$
$$f(x_j^2, m_j x_j^2 + q_j) - [a(x_j^2 - x) + b(m_j x_j^2 + q_j - y)] \geq c.$$

In other words, if $\bar{x}_j^1 > \bar{x}_j^2$ we can remove e_j from $\bar{E}(P)$.

If $\bar{x}_j^1 \leq \bar{x}_j^2$, $\tilde{e}_j \in \tilde{E}(P)$ is the subsegment of the edge $e_j \in \bar{E}(P)$ with x-coordinates in the interval $[\bar{x}_j^1, \bar{x}_j^2]$. The minimization problem on the left-hand side of the constraints associated to edges in $\bar{E}(P)$ can be solved as follows. Let

$$\begin{aligned} D_j^- &= \{(a,b) \ : \ a + bm_j \leq g_j'(\bar{x}_j^1)\} \\ D_j^+ &= \{(a,b) \ : \ a + bm_j \geq g_j'(\bar{x}_j^2)\} \\ D_j &= \{(a,b) \ : \ g_j'(\bar{x}_j^1) \leq a + bm_j \leq g_j'(\bar{x}_j^2)\}. \end{aligned} \tag{4}$$

Let

$$x_j(a,b) = \begin{cases} x_j^1 & \text{if } (a,b) \in int[D_j^-] \\ s_j(a,b) & \text{if } (a,b) \in D_j \\ x_j^2 & \text{if } (a,b) \in int[D_j^+] \end{cases} \tag{5}$$

where, *int* denotes the interior of a set, and $s_j(a,b)$ is a stationary point of

$$g_j(\chi) - (a + bm_j)\chi, \tag{6}$$

i.e.,

$$g_j'(s_j(a,b)) = a + bm_j. \tag{7}$$

Note that, after the definition of $x_j(a,b)$ we can redefine D_j^-, D_j, D_j^+ as follows

$$\begin{aligned} D_j^- &= \{(a,b) \ : \ s_j(a,b) \leq \bar{x}_j^1\} \\ D_j^+ &= \{(a,b) \ : \ s_j(a,b) \geq \bar{x}_j^2\} \\ D_j &= \{(a,b) \ : \ \bar{x}_j^1 \leq s_j(a,b) \leq \bar{x}_j^2\}. \end{aligned} \tag{8}$$

Finally, let

$$\mathbf{x}_j(a,b) = (x_j(a,b), y_j(a,b)) \quad \text{where} \quad y_j(a,b) = m_j x_j(a,b) + q_j.$$

For each $\mathbf{v}_i \in V(P)$ we set

$$x_i(a,b) = x_{v_i}, \quad y_i(a,b) = y_{v_i},$$

and

$$\mathbf{x}_i(a,b) = (x_i(a,b), y_i(a,b)).$$

Now, let

$$G(P) = \{\Omega_i \ : \ \Omega_i \in V(P) \cup \tilde{E}(P)\}. \tag{9}$$

For each k such that $\Omega_k \in G(P)$, let

$$\eta_k(a,b) = f(\mathbf{x}_k(a,b)) - ax_k(a,b) - by_k(a,b).$$

Then, we can rewrite (2) as follows

$$\begin{aligned} Conv_{f,P}(x,y) = \quad & \max \ c + ax + by \\ & \eta_k(a,b) \geq c \qquad\qquad k \ : \ \Omega_k \in G(P). \end{aligned} \tag{10}$$

In [11] it has been proved that problem (10) is a continuously differentiable convex problem.

3 KKT system and convex envelope

In [13] it has been proved that, given an optimal solution (a^*, b^*, c^*) of problem (10),

$$(x,y) \in chull\{\mathbf{x}_k(a^*,b^*) \ : \ k \in \mathcal{A}(a^*,b^*,c^*)\}, \tag{11}$$

where

$$\mathcal{A}(a^*,b^*,c^*) = \{k \ : \ \eta_k(a^*,b^*) = c^*\}, \tag{12}$$

is the set of active constraints at such optimal solution. This result is a consequence of the KKT conditions for problem (10). In fact, we can solve the problem if we are able to solve the KKT conditions. Thus, we report them in what follows:

$$\begin{array}{ll} \sum_{k \ : \ \Omega_k \in G(P)} \lambda_k^* x_k(a^*,b^*) = x & \\ \sum_{k \ : \ \Omega_k \in G(P)} \lambda_k^* y_k(a^*,b^*) = y & \\ \sum_{k \ : \ \Omega_k \in G(P)} \lambda_k^* = 1 & \\ \lambda_k^* \geq 0 & \forall k \ : \ \Omega_k \in G(P) \\ \eta_k(a^*,b^*) \geq c^* & \forall k \ : \ \Omega_k \in G(P) \\ \lambda_k^*(\eta_k(a^*,b^*) - c^*) = 0 & \forall k \ : \ \Omega_k \in G(P). \end{array} \tag{13}$$

Note that

$$\mathbf{x}_k(a^*,b^*) \ : \ k \in \mathcal{A}(a^*,b^*,c^*), \quad \lambda_k^* \geq 0 \ \ \forall k \ : \ \Omega_k \in G(P),$$

is an optimal solution of (1). Since $k \leq 3$ can be imposed in (1) in view of Caratheodory's theorem, we can restrict our attention to solutions of (13) with at most three positive values λ_k^*, i.e., with at most three active constraints. Thus, if we set

$$\mathcal{J} = \{J = \{k \ : \ \Omega_k \in G(P)\}, \ : \ 0 < |J| \leq 3\},$$

in order to compute the formula of the convex envelope, we should solve the following system:

$$\begin{array}{l} \sum_{k \in J} \lambda_k x_k(a,b) = x \\ \sum_{k \in J} \lambda_k y_k(a,b) = y \\ \sum_{k \in J} \lambda_k = 1 \\ \eta_k(a,b) - \eta_h(a,b) = 0 \quad \forall k,h \in J, \ k \neq h, \end{array} \tag{14}$$

Algorithm 1: Procedure for the computation of the convex envelope.

Data: Function $f : \mathbb{R}^2 \to \mathbb{R}$; polytope P.
Result: The convex envelope $conv_{f,P}$.
for $J \in \mathcal{J}$ **do**
  Solve the system (14) parametrically with respect to $(x, y) \in P$;
  Let $\lambda_k^J(x, y), k \in J,\ a_J(x, y),\ b_J(x, y)$ be the solution of the system;
  Set $\Gamma_J = \{(x, y)\ :\ \lambda_k^J(x, y) \geq 0, \forall k \in J,\ \ \eta_r(a_J(x, y), b_J(x, y)) \geq \eta_k(a_J(x, y), b_J(x, y)),\ \forall r \notin J,\ k \in J\}$;
  Set $conv_{f,P}(x, y) = \eta_k(a_J(x, y), b_J(x, y)) + a_J(x, y)x + b_J(x, y)y,\ k \in J \quad \forall (x, y) \in \Gamma_J$;
end

for each $J \in \mathcal{J}$. We end up with the procedure described in Algorithm 1 for the computation of the convex envelope. In fact, the collection $\mathcal{J}$ can be reduced in different ways. We first prove the following observation.

Observation 3.1 *Let $\Omega_k \in \tilde{E}(P)$ and let $e_k = [\mathbf{v}_1, \mathbf{v}_2]$, i.e., $\mathbf{v}_1, \mathbf{v}_2$ are the vertices of the edge e_k. Let $\Omega_i = \mathbf{v}_1$ and $\Omega_j = \mathbf{v}_2$. Then:*

- *if $k \in J$, for some $J \in \mathcal{J}$, we can restrict the attention to solutions of the system (14) for which $(a, b) \in D_k$;*
- *if $\mathbf{v}_1 \in \tilde{e}_k$, then discard all sets $J \supseteq \{i, k\}$, and similarly if $\mathbf{v}_2 \in \tilde{e}_k$, then discard all sets $J \supseteq \{j, k\}$;*
- *if $\mathbf{v}_1 (or\ \ \mathbf{v}_2) \notin \tilde{e}_k$, then the equality $\eta_k(a, b) = \eta_i(a, b)$ (or $\eta_k(a, b) = \eta_j(a, b)$) reduces to the linear equation*

$$a + m_k b = g_k'(\bar{x}_k^1)\ \ (or\ g_k'(\bar{x}_k^2)). \tag{15}$$

Proof. Let $k \in J$. For $(a, b) \notin D_k$, then $\mathbf{x}_k(a, b) = (x_k(a, b), y_k(a, b)) \in V(P)$. If $(a, b) \in int(D_k^-)$, then $\mathbf{x}_k(a, b) = \mathbf{v}_1$ and $\eta_k \equiv \eta_i$, i.e., the two functions η_i, η_k are equivalent over $int(D_k^-)$. This means that we can replace k with i in J. In a completely similar way, if $(a, b) \in int(D_k^+)$ we see that k can be replaced by j in J.

Now, if $k \in J$, in view of the previous part we must have $(a, b) \in D_k$. If $(a, b) \in D_k \setminus D_k^-$, then $\eta_i(a, b) > \eta_k(a, b)$ and $i, k \in J$ can not hold. If $(a, b) \in D_k \cap D_k^-$ and $\mathbf{v}_1 \in \tilde{e}_k$ we must have $\eta_i \equiv \eta_k$ and we can thus discard i from J. If $\mathbf{v}_1 \notin \tilde{e}_k$, then i, k can both lie in J and $(a, b) \in D_k \cap D_k^-$ is equivalent to the linear equality (15). In a completely similar way it can be seen that if $(a, b) \in D_k \setminus D_k^+$, then $j, k \in J$ can not hold, while for $(a, b) \in D_k \cap D_k^+$ we can discard j from J if $\mathbf{v}_2 \in \tilde{e}_k$, and reduce to the linear equality (15) otherwise. □

According to the result above, in Algorithm 1 we need to:

- introduce for each set $J \in \mathcal{J}$, the set
$$J' = \{k \in J \ : \ \Omega_k \in \tilde{E}(P)\};$$
- restrict the attention to solutions of the system (14) for which $(a, b) \in \cap_{k \in J'} D_k$ (in case $J' = \emptyset$, we consider $(a, b) \in \mathbb{R}^2$);
- redefine Γ_J as follows:
$$\Gamma_J = \{(x, y) \ : \ \eta_r(a_J(x, y), b_J(x, y)) \geq \eta_k(a_J(x, y), b_J(x, y)), \ \forall r \notin J, \ \forall k \in J, \\ \lambda_k^J(x, y) \geq 0, \ \forall k \in J, \quad (a_J(x, y), b_J(x, y)) \in \cap_{k \in J'} D_k\}.$$

Note that the restriction to solutions for which $(a, b) \in \cap_{k \in J'} D_k$ allows to overcome the difficulty related to the fact that for each $k \in J'$, $\eta_k(a, b), x_k(a, b), y_k(a, b)$ are piecewise defined over $\mathbb{R}^2$.

The procedure can be made more efficient with the exclusion of further subsets J. Some subsets J can be removed from $\mathcal{J}$ according to the following observation.

Observation 3.2 *If*
$$dim(chull(\cup_{k \in J} \Omega_k)) < 2,$$
then we can exclude J from further consideration.

Proof. For a subset J for which the condition is fulfilled, the corresponding optimal solution of (1) must lie over a region with dimension lower than two (a point or a segment). Over such regions the convex envelope can be recovered by continuity. □

The following is an immediate consequence of Observation 3.2.

Corollary 3.1 *We can restrict our attention to $J \in \mathcal{J}$ such that $|J| \geq 2$. Moreover, we can also discard all subsets $J = \{i, j\}$ such that Ω_i, Ω_j both reduce to a single point (a vertex of P).*

Further subsets can be removed. We first need to make an observation. For some subset J, if $(x, y) \in \Gamma_J$, then the solution of system (14)
$$\mathbf{x}_k^J(x, y) = (x_k(a_J(x, y), b_J(x, y)), y_k(a_J(x, y), b_J(x, y))), \quad \lambda_k^J(x, y) \geq 0, \ k \in J, \tag{16}$$
is an optimal solution for (1). In what follows we will omit the dependency from (x, y).

Observation 3.3 *Let $\{i, j\} \subset J$ and consider the optimal solution (16) with $\lambda_i^J, \lambda_j^J > 0$. Then, the following can not hold:*
$$\begin{aligned} &\forall \mathbf{z} = \lambda \mathbf{x}_i^J + (1 - \lambda)\mathbf{x}_j^J, \quad \lambda \in (0, 1) \\ &\exists \mathbf{z}_1, \mathbf{z}_2 \in [\mathbf{x}_i^J, \mathbf{x}_j^J] \ \textit{(not necessarily distinct), such that:} \\ &\qquad \mathbf{z} = \mu \mathbf{z}_1 + (1 - \mu)\mathbf{z}_2, \ \mu \in [0, 1] \\ &\qquad \mu f(\mathbf{z}_1) + (1 - \mu) f(\mathbf{z}_2) < \lambda f(\mathbf{x}_i^J) + (1 - \lambda) f(\mathbf{x}_j^J). \end{aligned} \tag{17}$$

Proof. The optimal value of (1), attained at the optimal solution (16), is

$$\sum_{k\in J} \lambda_k^J f(\mathbf{x}_k^J).$$

Now, let

$$\begin{aligned} \mathbf{x} &= \mathbf{x}_i^J \in \Omega_i \\ \mathbf{y} &= \mathbf{x}_j^J \in \Omega_j \\ \lambda &= \tfrac{\lambda_i^J}{\lambda_i^J+\lambda_j^J} \in (0,1). \end{aligned}$$

Then, if (17) holds, for some $\mu \in [0,1]$

$$\begin{aligned} &\textstyle\sum_{k\in J\setminus\{i,j\}} \lambda_k^J f(\mathbf{x}_k^J) + (\lambda_i^J + \lambda_j^J)\,[\lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})] > \\ &\textstyle\sum_{k\in J\setminus\{i,j\}} \lambda_k^J f(\mathbf{x}_k^J) + (\lambda_i^J + \lambda_j^J)\,[\mu f(\mathbf{z}_1) + (1-\mu) f(\mathbf{z}_2)]\,. \end{aligned}$$

Thus, it turns out that

$$\begin{array}{lll} \mathbf{x}_k^J, & \lambda_k^J & k \in J\setminus\{i,j\} \\ \mathbf{z}_1, & (\lambda_i^J+\lambda_j^J)\mu & \\ \mathbf{z}_2, & (\lambda_i^J+\lambda_j^J)(1-\mu) & \end{array}$$

is a feasible solution for (1) with a lower function value with respect to the optimal solution (16), which is a contradiction. □

The following are two corollaries of the previous observation.

Corollary 3.2 *Let $\{i,j\} \subset J$ and consider the optimal solution (16). Then, function f can not be strictly convex along the segment $[\mathbf{x}_i^J, \mathbf{x}_j^J]$.*

Proof. It is enough to observe that if f is strictly convex along the segment $[\mathbf{x}_i^J, \mathbf{x}_j^J]$, then $\forall \mathbf{z} = \lambda\mathbf{x}_i^J + (1-\lambda)\mathbf{x}_j^J$, $\lambda \in (0,1)$, the points $\mathbf{z}_1 = \mathbf{z}_2 = \mathbf{z}$ satisfy (17). □

Corollary 3.3 *Let $i,j \in \{k\ :\ \Omega_k \in G(P)\}$, $i \neq j$. If $\forall \mathbf{x} \in \Omega_i$, $\forall \mathbf{y} \in \Omega_j$, (17) holds after replacing $\mathbf{x}_i^J, \mathbf{x}_j^J$ with $\mathbf{x}, \mathbf{y}$, then, we can exclude from further consideration all subsets J such that $\{i,j\} \subseteq J$.*

A final remark is that we can deal with the remaining subsets $J \in \mathcal{J}$ in a different way according to their cardinality. Indeed, if $|J| = 3$ we notice that we can first solve over the region $\cap_{k\in J'} D_k$ at which we are interested, the following system with two equations and two unknowns

$$\eta_k(a,b) - \eta_h(a,b) = 0 \quad \forall k,h \in J,\ k \neq h. \tag{18}$$

In case $J' = \emptyset$, the system is a linear one and needs to be solved over $\mathbb{R}^2$. In [12] it has been proved that such system always has a finite number of solutions over the region $\cap_{k\in J'} D_k$ at which we are interested. Thus, for each of these solutions (a^*, b^*) we only need to check whether $\eta_r(a^*, b^*) \geq \eta_k(a^*, b^*)$ for each $r \notin J$, $k \in J$ and, in such case to define Γ_J as follows

$$\Gamma_J = chull\left(\left\{\mathbf{x}^k(a^*, b^*)\ :\ k \in J\right\}\right).$$

Note that over this region the convex envelope is defined by an affine function. If $J = \{k, h\}$, i.e., $|J| = 2$, we first remark that, in view of Corollary 3.1, at least one of the two sets Ω_h and Ω_k must belong to $\tilde{E}(P)$. Moreover (14) reduces to

$$\begin{aligned}
&\lambda x_k(a, b) + (1 - \lambda)x_h(a, b) = x \\
&\lambda(m_k x_k(a, b) + q_k) + (1 - \lambda)(m_h x_h(a, b) + q_h) = y \\
&\eta_k(a, b) - \eta_h(a, b) = 0.
\end{aligned} \tag{19}$$

Thus, we end up with Algorithm 2, which is a revised and more detailed version of Algorithm 1. In the following sections we will apply this procedure to two special cases.

4 The bilinear function

In [12] it has been observed that the convex envelope of the bilinear function over any polytope P is characterized by a polyhedral subdivision of P, and that over each member of the polyhedral subdivision the functional form of the convex envelope is one of three possible forms: affine, quadratic, and ratio of a quadratic to an affine function. In this section we exploit the results of the previous section in order to identify the polyhedral subdivision, given by the different sets Γ_J, $J \in \mathcal{J}$, and the functional form over each member of the subdivision.

When $f(x, y) = xy$ we observe that $\tilde{E}(P) = \bar{E}(P)$. For each edge $e_j \in E(P)$ we denote by $y = m_j x + q_j$ the line along which the edge lies, and by (x_j^1, y_j^1) and (x_j^2, y_j^2) the vertices of the edge. It turns out that

$$\bar{E}(P) = \{e_j\ :\ m_j > 0\},$$

i.e., the edges along which f is strictly convex are those lying along a line with positive slope. Moreover, we can further split $\bar{E}(P)$ into (at most) two subsets $\bar{E}^\ell(P)$ and $\bar{E}^u(P)$ defined as follows

$$\begin{aligned}
\bar{E}^u(P) &= \{e_k \in E(P)\ :\ P \subset \{(x, y)\ :\ y \leq m_k x + q_k\}\} \\
\bar{E}^\ell(P) &= \{e_k \in E(P)\ :\ P \subset \{(x, y)\ :\ y \geq m_k x + q_k\}\}.
\end{aligned}$$

We can introduce some simplifications within the general Algorithm 2. Before proceeding, we report the definition of $x_j(a, b)$ and of the function $\eta_j(a, b)$ for $\Omega_j \in \bar{E}(P)$. We have

$$x_j(a, b) = s_j(a, b) = \frac{a + m_j b - q_j}{2m_j} \text{ over } D_j = \left\{(a, b)\ :\ x_1^j \leq s_j(a, b) \leq x_2^j\right\}, \tag{20}$$

Algorithm 2: Revised procedure for the computation of the convex envelope.

Data: Function $f : \mathbb{R}^2 \to \mathbb{R}$; polytope P.
Result: The convex envelope $conv_{f,P}$.
for $J \in \mathcal{J}$ **do**
  Set $J' = \{k \in J \ : \ \Omega_k \in \tilde{E}(P)\}$;
  if $|J| = 3$ **then**
    Solve the system (18) over $\cap_{k \in J'} D_k$ (over $\mathbb{R}^2$ in case $J' = \emptyset$), and let T be the finite set of solutions;
    foreach $(a^*, b^*) \in T$ **do**
      if $\eta_r(a^*, b^*) \geq \eta_k(a^*, b^*) \ \forall r \notin J, \ k \in J$ **then**
        Set $\Gamma_J = chull\left(\left\{\mathbf{x}^k(a^*, b^*) \ : \ k \in J\right\}\right)$;
        Set $conv_{f,P}(x, y) = \eta_k(a^*, b^*) + a^*x + b^*y, \ k \in J \quad \forall (x, y) \in \Gamma_J$;
      end
    end
  end
  else
    Solve the system (19) over $\cap_{k \in J'} D_k$ parametrically with respect to $(x, y) \in P$;
    Let $\lambda^J(x, y), \ a_J(x, y), \ b_J(x, y)$ be the solution of the system;
    Set $\Gamma_J = \{(x, y) \ : \ \lambda^J(x, y) \in [0, 1], \quad \eta_r(a_J(x, y), b_J(x, y)) \geq \eta_k(a_J(x, y), b_J(x, y)), \ \forall r \notin J, \ k \in J, \ (a_J(x, y), b_J(x, y)) \in \cap_{k \in J'} D_k\}$;
    Set $conv_{f,P}(x, y) = \eta_k(a_J(x, y), b_J(x, y)) + a_J(x, y)x + b_J(x, y)y, \ k \in J \quad \forall (x, y) \in \Gamma_J$;
  end
end

and $\forall (a,b) \in D_j$

$$\eta_j(a,b) = -m_j x_j(a,b)^2 - bq_j. \tag{21}$$

Recall that we are interested at the definition of the function η_j only over D_j.

The following observation is a consequence of Corollary 3.3.

Observation 4.1 *We can exclude from $\mathcal{J}$ all subsets J such that $\{i,j\} \subseteq J$, $i \neq j$, and $\Omega_i, \Omega_j \in \bar{E}^\ell(P)$ or $\Omega_i, \Omega_j \in \bar{E}^u(P)$.*

Proof. It is enough to observe that f is strictly convex along each line through any $\mathbf{x} \in \Omega_i$ and $\mathbf{y} \in \Omega_j$, so that for each $\mathbf{z} \in [\mathbf{x}, \mathbf{y}]$ we can always choose $\mathbf{z}_1 = \mathbf{z}_2 = \mathbf{z}$. □

Next, we prove the following observation, which shows that the only subsets $J = \{i,j,k\}$ which need to be considered are those for which $\Omega_i, \Omega_j, \Omega_k \in V(P)$. A generalization of this result to other functions is given in Appendix A.

Observation 4.2 *Let $J = \{i,j,k\} \in \mathcal{J}$. Then, the three points $\mathbf{x}_i^J, \mathbf{x}_j^J, \mathbf{x}_k^J$ in (16) must all belong to $V(P)$.*

Proof. Let $\mathbf{x}_h^J = (x_h, y_h)$, $h \in \{i,j,k\}$. We assume that $\Omega_k \in \bar{E}^u(P)$ (the case when $\Omega_k \in \bar{E}^\ell(P)$ is completely analogous) and, by contradiction, we also assume that $(x_k, y_k) \notin V(P)$. In view of Observation 4.1 at least one among Ω_i and Ω_j belongs to $V(P)$. Without loss of generality we assume that:

- $(x_i, y_i) = (0,0)$ (this can always be made true by a translation);
- $x_j \geq 0$.

Note that $\Omega_k \in \bar{E}^u(P)$ and $(x_i, y_i) = (0,0)$ imply $m_k, q_k > 0$. Moreover, since $y_k = m_k x_k + q_k$, we have that $y_k < 0$ implies $x_k < 0$. But in this case f is strictly convex along the segment between (x_k, y_k) and (x_i, y_i), which can not hold in view of Corollary 3.2. Similarly, $x_k > 0$ implies $y_k > 0$, which can not hold for the same reason. Thus, only $x_k \leq 0, y_k \geq 0$ is possible. In view of (20), we have

$$x_k = \frac{a + m_k b - q_k}{2m_k}, \quad y_k = m_k x_k + q_k,$$

and

$$\begin{aligned}
\eta_k(a,b) &= x_k y_k - a x_k - b y_k = -\tfrac{(a+bm_k-q_k)^2}{4m_k} - bq_k \\
\eta_j(a,b) &= x_j y_j - a x_j - b y_j \\
\eta_i(a,b) &= x_i y_i - a x_i - b y_i = 0.
\end{aligned}$$

After setting

$$z = a + bm_k - q_k = 2m_k x_k \leq 0, \tag{22}$$

and imposing $\eta_k(a,b) = \eta_j(a,b) = \eta_i(a,b)$, we have the equivalent system

$$\begin{aligned} & b = -\tfrac{z^2}{4m_k q_k} \\ & a = z - m_k b + q_k \\ & (m_k x_j - y_j)z^2 + 4m_k q_k x_j z - 4m_k q_k x_j (y_j - q_k) = 0. \end{aligned} \tag{23}$$

The determinant of the last quadratic equation is

$$\Delta = y_j(m_k x_j - y_j + q_k).$$

We first discuss the case $x_j = 0$. Then, $z = 0$ and, thus, also $x_k = 0$. This is not possible since the three points $(x_i, y_i), (x_j, y_j), (x_k, y_k)$ can not be aligned along the line $x = 0$. Thus, we impose $x_j > 0$. We notice that $y_j > 0$ can not hold since in this case the restriction of f along $(0,0)$ and (x_j, y_j) is strictly convex, which is not possible in view of Corollary (3.2). If $y_j < 0$, we first remark that $m_k x_j - y_j + q_k > 0$, since $(x_j, y_j) \in P$ and can not lie along the line $y = m_k x + q_k$, i.e., $y_j < m_k x_j + q_k$. Then, $\Delta < 0$, i.e., the quadratic equation has no root and, thus, also the system (23) has no solution, from which we are lead to a contradiction. Thus, we are only left with $x_j > 0$ and $y_j = 0$, for which $\Delta = 0$ and the unique root of the quadratic equation is

$$z = -\frac{2m_k q_k x_k}{m_k x_j - y_j} = -\frac{2q_k x_k}{x_j} \geq 0.$$

If $x_k < 0$, this is not possible in view of the non positivity of z. Thus, we are only left with the case $x_k = 0$. But $x_k, y_j = 0$ is only possible if $z = a = b = 0$, which, however, can not be a solution of the system (23). □

Thus, for the bilinear function the systems (18) are always simple linear systems.

In conclusion, for the bilinear case in Algorithm 2 we can:

- restrict the attention to subsets $J = \{i, j, k\}$ with $\Omega_i, \Omega_j, \Omega_k \in V(P)$ and, moreover, such that no pair of these vertices lies along a line with positive slope;
- restrict the attention to subsets $J = \{h, k\}$ such that $\Omega_h \in \bar{E}^u(P)$, $\Omega_k \in \bar{E}^\ell(P)$, or $\Omega_h \in V(P)$ and either $\Omega_k \in \bar{E}^\ell(P)$ or $\Omega_k \in \bar{E}^u(P)$.

For the triples J made up by three vertices the computation of the convex envelope is rather simple: the convex envelope is the affine function interpolating xy at the three vertices, while Γ_J is the convex hull of the three vertices. Thus, we will only discuss the subsets $J = \{i, j\}$ with cardinality two. We need to discuss the two subcases mentioned before. It will turn out that for these subsets the definition of Γ_J will only depend on the vertices of the edge(s) if Ω_i and/or Ω_j belong to $\bar{E}(P)$, and on the adjacent edges if Ω_i or Ω_j (not both) belongs to $V(P)$.

4.1 $\Omega_i \in V(P)$, $\Omega_j \in \bar{E}(P)$

We denote by (v_x, v_y) the coordinates of the vertex $\Omega_i \in V(P)$. We will omit in what follows the dependency of x_j from (a, b). Recalling (21), system (19) is equivalent to

$$\begin{aligned} &\lambda x_j + (1-\lambda)v_x = x \\ &\lambda(m_j x_j + q_j) + (1-\lambda)v_y = y \\ &-m_j x_j^2 - bq_j = v_x v_y - av_x - bv_y. \end{aligned}$$

The parametric solution of this system is easy to derive. It follows from (20) that $a = 2m_j x_j + q_j - m_j b$. Then, the system can be rewritten as follows

$$\begin{aligned} &\lambda x_j + (1-\lambda)v_x = x \\ &\lambda(m_j x_j + q_j) + (1-\lambda)v_y = y \\ &m_j x_j^2 - 2m_j x_j v_x - q_j v_x + b(q_j + m_j v_x - v_y) + v_x v_y = 0. \end{aligned}$$

The first two equations lead to

$$\begin{aligned} \lambda &= \tfrac{x - v_x}{x_j - v_x} \\ \lambda &= \tfrac{y - v_y}{m_j x_j + q_j - v_y}, \end{aligned}$$

so that

$$x_j = \frac{x(v_y - q_j) - v_x(y - q_j)}{m_j(x - v_x) + v_y - y}. \tag{24}$$

Then

$$\begin{aligned} \lambda(x,y) &= \tfrac{m_j(x - v_x) + v_y - y}{v_y - q_j - m_j v_x} \\ b(x,y) &= \tfrac{m_j x_j^2 - 2m_j v_x x_j - q_j v_x + v_x v_y}{v_y - m_j v_x - q_j} \\ a(x,y) &= 2m_j x_j + q_j - m_j b(x,y). \end{aligned} \tag{25}$$

The convex envelope is equal to

$$conv_{f,P}(x,y) = v_x v_y + a(x,y)(x - v_x) + b(x,y)(y - v_y), \tag{26}$$

over the set

$$\begin{aligned} \Gamma_J = \{(x,y) \ : \ &\eta_r(a_J(x,y), b_J(x,y)) \geq \eta_k(a_J(x,y), b_J(x,y)), \ \forall r \notin J, \ k \in J, \\ &\lambda^J(x,y) \in [0,1], \quad (a_J(x,y), b_J(x,y)) \in D_j\}. \end{aligned}$$

Note that, according to definition (8) $(a_J(x,y), b_J(x,y)) \in D_j$ can also be written as $x_j^1 \leq x_j \leq x_j^2$. In the following observation we derive a simplified definition of the set Γ_J. It will turn out that the restrictions $\eta_r(a_J(x,y), b_J(x,y)) \geq \eta_k(a_J(x,y), b_J(x,y))$, $\forall r \notin J, \ k \in J$, can be replaced by simple restrictions on the values of x_j. We will assume in what follows that $v_x = v_y = 0$. This is without loss of generality since it can always be made true by a translation. We also assume that $\Omega_j \in \bar{E}^u(P)$ (the analysis for the case $\Omega_j \in \bar{E}^\ell(P)$ is analogous).

Observation 4.3 *Let* $(v_x, v_y) = (0, 0)$ *and* $\Omega_j \in \bar{E}^u(P)$. *Then, the following holds.*

- *If* $P \cap \{(\alpha_x, \alpha_y) \ : \ \alpha_x > 0, \alpha_y < 0\} \neq \emptyset$, *then* $\Gamma_J = \emptyset$;
- *If* $P \subseteq \{(\alpha_x, \alpha_y) \ : \ \alpha_x \leq 0, \alpha_y \geq 0\}$, *then*

$$\Gamma_J = \{(x, y) \ : \ \lambda^J(x, y) \in [0, 1], \quad x_j^1 \leq x_j \leq x_j^2\}. \tag{27}$$

- *If* $P \cap \{(\alpha_x, \alpha_y) \ : \ \alpha_x > 0, \alpha_y \geq 0\} \neq \emptyset$, *then we need to add the restriction*

$$x_j \leq \min\left\{x_j^2, -\frac{q_j}{m_j + \sqrt{m_j m_k}}\right\}$$

  in the definition (27) of Γ_J.
- *If* $P \cap \{(\alpha_x, \alpha_y) \ : \ \alpha_x < 0, \ \alpha_y \leq 0\} \neq \emptyset$, *then we need to add the restriction*

$$x_j \geq \max\left\{x_j^1, -\frac{q_j}{m_j + \sqrt{m_j m_h}}\right\}$$

  in the definition (27) of Γ_J.

Proof. Since $\Omega_j \in \bar{E}^u(P)$ and $(v_x, v_y) = (0, 0)$, it holds that $q_j > 0$ and

$$a(x, y) = \frac{(m_j x_j + q_j)^2}{q_j} \geq 0, \quad b(x, y) = -\frac{m_j x_j^2}{q_j} \leq 0.$$

Moreover, $x_j \leq 0, y_j = m_j x_j + q_j \geq 0$, otherwise the bilinear function is strictly convex along the segment between $(0, 0)$ and (x_j, y_j), which can not hold in view of Corollary 3.2. Thus, we can impose $-\frac{q_j}{m_j} \leq x_j \leq 0$. Next, we need to impose that for each $(\alpha_x, \alpha_y) \in P$,

$$\alpha_x \alpha_y - a(x, y)\alpha_x - b(x, y)\alpha_y \geq \eta_i(a(x, y), b(x, y)) = 0. \tag{28}$$

We remark that we could restrict the attention to $(\alpha_x, \alpha_y) \in \Omega_k$, for all $\Omega_k \in G(P)$, but this would not simplify the following analysis. The inequality (28) can be rewritten as follows

$$m_j x_j^2(\alpha_y - m_j \alpha_x) - 2m_j q_j \alpha_x x_j + q_j \alpha_x(\alpha_y - q_j) \geq 0. \tag{29}$$

If we consider the above inequality as a quadratic inequality with respect to x_j, then its determinant is

$$m_j q_j \alpha_x \alpha_y (m_j \alpha_x + q_j - \alpha_y).$$

This is < 0 for $\alpha_x \alpha_y < 0$ and $(\alpha_x, \alpha_y) \notin \Omega_j$. Thus, if $\alpha_x \alpha_y < 0$ and $\alpha_y - m_j \alpha_x < 0$, i.e., $\alpha_x \alpha_y < 0, \ \alpha_y < 0, \ \alpha_x > 0$, for some $(\alpha_x, \alpha_y) \in P$, then $\Gamma_J = \emptyset$. Otherwise, if $\alpha_x \alpha_y \leq 0$ and $\alpha_y - m_j \alpha_x \geq 0$, i.e., $\alpha_x \alpha_y \leq 0, \ \alpha_y \geq 0, \ \alpha_x \leq 0$ for all $(\alpha_x, \alpha_y) \in P$, then

$$\Gamma_J = \{(x, y) \ : \ \lambda^J(x, y) \in [0, 1], \quad x_j^1 \leq x_j \leq x_j^2\}.$$

Next, let us assume that $\alpha_x > 0,\ \alpha_y \geq 0$ for some $(\alpha_x, \alpha_y) \in P$. In this case the origin is the vertex of an edge of P lying along a line $y = m_k x$ for some $m_k \geq 0$. Moreover, $P \subseteq \{(x, y) :\ y \geq m_k x\}$. Thus, since $b(x, y) \leq 0$, we have

$$\alpha_x\alpha_y - a(x, y)\alpha_x - b(x, y)\alpha_y \geq \alpha_x\alpha_y - a(x, y)\alpha_x - b(x, y)m_k\alpha_x.$$

Taking into account the definitions of $a(x, y)$ and $b(x, y)$ we have that, in view of $\alpha_x > 0$, (28) is satisfied if

$$\alpha_y \geq \frac{(m_j x_j + q_j)^2}{q_j} - \frac{m_j m_k x_j^2}{q_j},$$

Since $\alpha_y \geq 0$, $q_j > 0$, $x_j \leq 0$, and $y_j = m_j x_j + q_j \geq 0$, the above inequality is satisfied if

$$m_j x_j + q_j + \sqrt{m_j m_k} x_j \leq 0,$$

or, equivalently

$$x_j \leq -\frac{q_j}{m_j + \sqrt{m_j m_k}},$$

which, combined with $x_j \leq x_j^2$ proves the result.

Finally, let us assume that $\alpha_x < 0,\ \alpha_y \leq 0$ for some $(\alpha_x, \alpha_y) \in P$. In this case the origin is the vertex of an edge of P lying along a line $y = m_h x$ for some $m_h \geq 0$. Moreover, $P \subseteq \{(x, y)\ :\ y \geq m_h x\}$. Thus, since $b(x, y) \leq 0$, we have

$$\alpha_x\alpha_y - a(x, y)\alpha_x - b(x, y)\alpha_y \geq \alpha_x\alpha_y - a(x, y)\alpha_x - b(x, y)m_h\alpha_x.$$

Taking into account the definitions of $a(x, y)$ and $b(x, y)$ we have that, in view of $\alpha_x < 0$, (28) is satisfied if

$$\alpha_y \leq \frac{(m_j x_j + q_j)^2}{q_j} - \frac{m_j m_h x_j^2}{q_j},$$

Since $\alpha_y \leq 0$, $q_j > 0$, $x_j \leq 0$, and $y_j = m_j x_j + q_j \geq 0$, the above inequality is satisfied if

$$m_j x_j + q_j + \sqrt{m_j m_h} x_j \geq 0,$$

or, equivalently

$$x_j \geq -\frac{q_j}{m_j + \sqrt{m_j m_h}},$$

which, combined with $x_j \geq x_j^1$ proves the result. □

4.2 $\Omega_i, \Omega_j \in \bar{E}(P)$

We will assume that $\Omega_j \in \bar{E}^\ell(P)$, $\Omega_i \in \bar{E}^u(P)$, so that

$$y \geq m_j x + q_j, \quad y \leq m_i x + q_i \quad \forall (x, y) \in P. \tag{30}$$

System (19) becomes (once again we omit the dependency of x_i and x_j from a, b)

$$\begin{aligned}
&\lambda x_j + (1-\lambda)x_i = x\\
&\lambda(m_j x_j + q_j) + (1-\lambda)(m_i x_i + q_i) = y\\
&m_j x_j^2 + b q_j = m_i x_i^2 + b q_i.
\end{aligned}$$

The case $m_i = m_j$ is a simpler one for which a solution of the system is easily derived. Indeed, in this case it follows from (20) that

$$x_j = x_i + \frac{q_i - q_j}{2m_i}.$$

The first two equations lead to

$$\begin{aligned}
\lambda &= \tfrac{2m_i(x-x_i)}{q_i-q_j}\\
\lambda &= \tfrac{2(y-m_i x_i-q_i)}{q_j-q_i},
\end{aligned}$$

so that

$$x_i = \frac{y + m_i x - q_i}{2m_i}.$$

The third equation reduces to

$$b(x,y) = x_i + \frac{q_i - q_j}{4},$$

while (20) implies

$$a(x,y) = 2m_i x_i - m_i b(x,y) + q_i.$$

If $m_i \neq m_j$ the solution of the system is a bit more cumbersome although always based on standard computations. It follows from (20) that

$$\begin{aligned}
b &= \tfrac{2m_j x_j + q_j - 2m_i x_i - q_i}{m_j - m_i}\\
a &= 2m_j x_j + q_j - m_j b.
\end{aligned} \tag{31}$$

The first two equations lead to

$$\begin{aligned}
\lambda &= \tfrac{x-x_i}{x_j-x_i}\\
\lambda &= \tfrac{y-m_i x_i-q_i}{m_j x_j+q_j-m_i x_i-q_i}.
\end{aligned}$$

The system can be rewritten as follows

$$\begin{aligned}
&(m_i - m_j)x_i x_j + x_i(y - m_i x - q_j) + x_j(m_j x + q_i - y) + (q_j - q_i)x = 0\\
&(\sqrt{m_i}x_i + \sqrt{m_j}x_j)(\sqrt{m_j}x_j - \sqrt{m_i}x_i) = (q_i - q_j)\tfrac{2m_j x_j + q_j - 2m_i x_i - q_i}{m_j - m_i}.
\end{aligned}$$

In order to solve it parametrically with respect to (x, y), it is worthwhile to make the following change of variables

$$Z = \sqrt{m_j}x_j - \sqrt{m_i}x_i$$
$$W = \sqrt{m_j}x_j + \sqrt{m_i}x_i.$$

After that, the system is rewritten as follows

$$(m_i - m_j)\tfrac{W^2-Z^2}{4\sqrt{m_i m_j}} + \tfrac{W-Z}{2\sqrt{m_i}}(y - q_j - m_i x) + \tfrac{W+Z}{2\sqrt{m_j}}(-y + q_i + m_j x) + (q_j - q_i)x = 0$$
$$[(\sqrt{m_j} + \sqrt{m_i})Z + q_j - q_i][(\sqrt{m_j} - \sqrt{m_i})W + q_j - q_i] = 0.$$

It is immediately seen that the possible solutions of the second equation are

$$Z = \frac{q_i - q_j}{\sqrt{m_i} + \sqrt{m_j}}, \tag{32}$$

or

$$W = \frac{q_i - q_j}{\sqrt{m_j} - \sqrt{m_i}}, \tag{33}$$

If (32) holds, then, after a few computations, it can be seen that the solutions of the first equation are

$$W_1 = 2\ell_1(x, y) + \frac{q_i - q_j}{\sqrt{m_i} - \sqrt{m_j}} \quad \text{or} \quad W_2 = \frac{q_i - q_j}{\sqrt{m_j} - \sqrt{m_i}},$$

where

$$\ell_1(x, y) = \frac{\sqrt{m_i}(m_j x - y + q_i) - \sqrt{m_j}(m_i x - y + q_j)}{m_j - m_i}.$$

Solution W_2 can be discarded. Indeed, in this case both Z and W do not depend from x, y and

$$x_i = x_j = \frac{q_i - q_j}{m_j - m_i}.$$

This is only possible if $(x_j, y_j) \equiv (x_i, y_i)$ and the point is the intersection of the two lines $y = m_j x + q_j$ and $y = m_i x + q_i$, so that we can discard this case. If we consider the solution W_1, we have

$$x_i = \tfrac{y+\sqrt{m_i m_j}x - q_i}{\sqrt{m_i}(\sqrt{m_i}+\sqrt{m_j})}$$
$$x_j = \tfrac{y+\sqrt{m_i m_j}x - q_j}{\sqrt{m_j}(\sqrt{m_i}+\sqrt{m_j})}. \tag{34}$$

Next, let us assume that (33) holds. In this case we have the two solutions

$$Z_1 = -2\ell_2(x, y) + \frac{q_j - q_i}{\sqrt{m_i} + \sqrt{m_j}} \quad \text{or} \quad Z_2 = \frac{q_i - q_j}{\sqrt{m_i} + \sqrt{m_j}},$$

where

$$\ell_2(x, y) = \frac{\sqrt{m_i}(y - m_j x - q_i) - \sqrt{m_j}(m_i x - y + q_j)}{m_i - m_j}.$$

Once again, Z_2 can be discarded. If we consider Z_1 we end up with

$$\begin{array}{l} x_i = \frac{y-\sqrt{m_i m_j}x-q_i}{\sqrt{m_i}(\sqrt{m_i}-\sqrt{m_j})} \\ x_j = \frac{-y+\sqrt{m_i m_j}x+q_j}{\sqrt{m_j}(\sqrt{m_i}-\sqrt{m_j})}. \end{array}$$

However, in this case we observe that (30) implies that for each $(x, y) \in P \setminus [\Omega_i \cup \Omega_j]$ $x_i, x_j < x$ if $m_i > m_j$, or $x_i, x_j > x$ if $m_i < m_j$ (note that over Ω_i and over Ω_j the convex envelope is equal to the restriction of the bilinear function to that edge and we do not need to consider these points). Then, $\lambda(x, y) \notin [0, 1]$, so that we can discard this solution. In conclusion, the only acceptable solution of the system is (34). Then, we can derive $a(x, y)$ and $b(x, y)$ from (31). We have

$$conv_{f,P}(x, y) = g(x, y) = -m_j x_j^2 - b(x, y)q_j + a(x, y)x + b(x, y)y, \tag{35}$$

over the set

$$\begin{aligned} \Gamma_J = \{(x, y) \ : \ & \eta_r(a_J(x, y), b_J(x, y)) \geq \eta_k(a_J(x, y), b_J(x, y)), \ \forall r \notin J, \ k \in J \\ & \lambda^J(x, y) \in [0, 1], \quad (a_J(x, y), b_J(x, y)) \in D_j\} . \end{aligned}$$

In fact, the following observation gives a simplified definition of the set Γ_J, showing that we can omit $\eta_r(a_J(x, y), b_J(x, y)) \geq \eta_k(a_J(x, y), b_J(x, y)), \ \forall r \notin J, \ k \in J$.

Observation 4.4 *We have that*

$$\Gamma_J = \{(x, y) \ : \ \lambda^J(x, y) \in [0, 1], \quad (a_J(x, y), b_J(x, y)) \in D_i \cap D_j\}. \tag{36}$$

Proof. Let $P' = chull(\Omega_i \cup \Omega_j) \subseteq P$. It turns out that $conv_{f,P'}(x, y)$ is equal to the function g defined in (35) over the set $\Gamma_J \subseteq P'$ defined in (36). This is an immediate consequence of the fact that in this case $r \notin J$ implies that Ω_r is a vertex of one of the two edges Ω_i and Ω_j of P', and $\eta_r(a_J(x, y), b_J(x, y)) \geq \eta_i(a_J(x, y), b_J(x, y))$ for all $r \notin J$ follows from $(a_J(x, y), b_J(x, y)) \in D_i \cap D_j$. What we will prove now is that g is a convex underestimator of f over the whole polytope P. Then, by definition of convex envelope as the largest convex underestimator of f over P and observing that $P' \subseteq P$ implies $conv_{f,P'}(x, y) \geq conv_{f,P}(x, y)$ $\forall (x, y) \in P'$, we must have $conv_{f,P}(x, y) = conv_{f,P'}(x, y) = g(x, y)$ over Γ_J.

In what follows we will omit the dependency of a and b from J. In order to see that g is a convex underestimator of f over the whole polytope P, we need to check whether

$$g(x, y) = -m_j x_j^2 - b(x, y)q_j + a(x, y)x + b(x, y)y \leq xy \quad \forall (x, y) \in P.$$

We rewrite this as

$$-m_j x_j^2 - b(x, y)q_j + a(x, y)x + b(x, y)(y + m_j x - q_j) + b(x, y)q_j - b(x, y)m_j x - xy \leq 0,$$

or, equivalently

$$-m_j x_j^2 + (a(x,y) + m_j b(x,y))x + b(x,y)(y - m_j x - q_j) - xy \le 0.$$

Since $a(x,y) + m_j b(x,y) = 2m_j x_j + b(x,y)q_j$, and after adding and subtracting $m_j x^2$, we end up, after a few elementary computations, with

$$m_j(x_j - x)^2 \ge (x - b(x,y))(m_j x + q_j - y).$$

By definition of x_j and $b(x,y)$ the inequality reduces to

$$y - m_i x - q_i \le 0,$$

which always holds over P. □

4.3 An example

In order to illustrate the results of this section we consider the following example taken from [17]. Let

$$P = \{(x,y)\ :\ x,y \ge 0,\ x \le 5,\ y \le x+1\},$$

i.e., P is the polytope with vertices $\mathbf{v}_1 = (0,0)$, $\mathbf{v}_2 = (5,0)$, $\mathbf{v}_3 = (0,1)$, and $\mathbf{v}_4 = (5,6)$. We set $\Omega_i = \{\mathbf{v}_i\}$, $i = 1,\dots,4$. We have that $\bar{E}^u(P)$ is made up by the edge $\Omega_5 = [\mathbf{v}_3, \mathbf{v}_4]$, while $\bar{E}^\ell(P) = \emptyset$. The only set J with cardinality three which needs to be considered is $J = \{1,2,3\}$, from which we have

$$conv_{xy,P}(x,y) = 0 \quad \forall\ (x,y) \in chull\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}.$$

The only set with cardinality two which needs to be considered is $J = \{2,5\}$. Following the development of Section 4.1, we have from (24) that

$$x_j = \frac{x+5y-5}{y+5-x}, \tag{37}$$

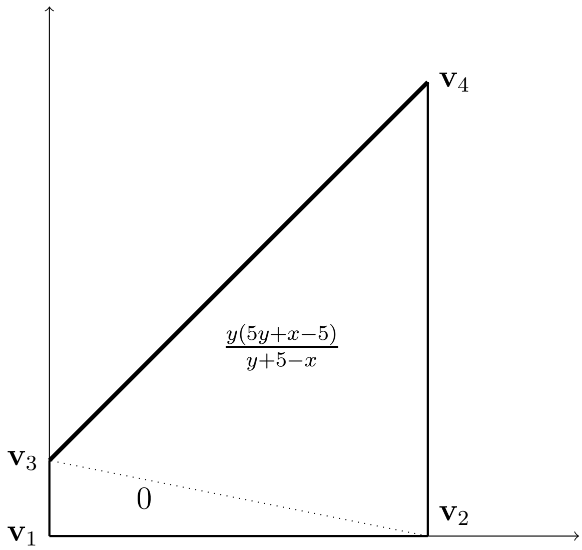
from (25) that

$$\begin{aligned} \lambda(x,y) &= \tfrac{y+5-x}{6} \\ b(x,y) &= -\tfrac{x_j^2-10x_j-5}{6} \\ a(x,y) &= x_j + 1 - b(x,y), \end{aligned}$$

so that it follows from (26) that

$$conv_{xy,P} = -\frac{x_j^2 - 10x_j - 5}{6}(y - x + 5) + (2x_j + 1)(x - 5),$$

Figure 2: Polyhedral subdivision for the convex envelope of $f(x, y) = xy$ over P.



over Γ_J. In order to define Γ_J, we notice that: (i) $\lambda(x, y) \in [0, 1]$ $\forall (x, y) \in P$; (ii) after translating $\mathbf{v}_2$ into the origin, we remark that we are in the second subcase of Observation 4.3. Thus,

$$\Gamma_J = \{(x, y) \ : \ 0 \leq x_j \leq 5\}.$$

Recalling (37), we conclude that

$$conv_{xy,P}(x, y) = \begin{cases} 0 & x + 5y - 5 \leq 0 \\ \frac{y(5y+x-5)}{y+5-x} & \text{otherwise.} \end{cases}$$

Figure 2 reports polytope P and the corresponding polyhedral subdivision.

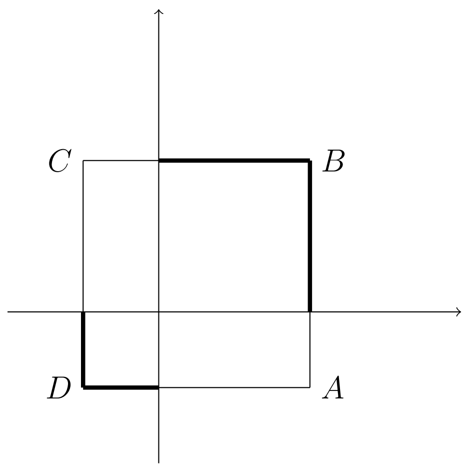
5 Functions $x^n y^m$ over boxes

In this section we consider functions $f(x, y) = x^n y^m$, with m, n positive integers, over boxes $\mathcal{B} = [\ell_x, u_x] \times [\ell_y, u_y]$. We will discuss the more complicated cases, namely those where $m, n > 1$ are odd integer values, and $\ell_x u_x, \ell_y u_y < 0$, i.e., the origin lies in the interior of the box $\mathcal{B}$. The vertices of the box will be denoted as

$$A \quad (u_x, l_y) \quad B \quad (u_x, u_y) \quad C \quad (l_x, u_y) \quad D \quad (l_x, l_y).$$

The subsegments of the edges along which f is strictly convex are displayed as thick lines in Figure 3. We remark that the case $m = n$ has been already discussed in [13]. Here, however, we propose a simplified way to derive the convex envelope also for that case.

Figure 3: The box $\mathcal{B} = [l_x, u_x] \times [l_y, u_y]$. The thick lines are the subsegments of the edges along which the restriction of $x^n y^m$, $n, m > 1$ and odd, is strictly convex.



Following the notation of Section 2 (see also Figure 1), for what concerns the edge $e_1 = BC$, we have $\bar{x}_1^2 = u_x$, while $\bar{x}_1^1$ is the solution of the following equation

$$(1-n)x^n + nl_x x^{n-1} = l_x^n. \tag{38}$$

If $\bar{x}_1^1 < u_x$, then $\tilde{e}_1 = KB$, where $K = (\bar{x}_1^1, u_y)$. Similarly, for what concerns the edge $e_2 = AB$, we have $\bar{y}_2^2 = u_y$, while $\bar{y}_2^1$ is the solution of the following equation

$$(1-m)y^m + ml_y y^{m-1} = l_y^m.$$

If $\bar{y}_2^1 < u_y$, then $\tilde{e}_2 = LB$, where $L = (u_x, \bar{y}_1^2)$. In a similar way we can define the values $\bar{x}_3^2$ associated to the edge e_3, $\bar{y}_4^2$ associated to the edge e_4, the segment $\tilde{e}_3$, associated to the edge $e_3 = AD$, and the segment $\tilde{e}_4$, associated to the edge $e_4 = CD$. The collection $G(\mathcal{B})$ is made up by the following sets

$$\Omega_1 = \tilde{e}_1, \quad \Omega_2 = \tilde{e}_2, \quad \Omega_3 = \tilde{e}_3, \quad \Omega_4 = \tilde{e}_4$$
$$\Omega_5 = A, \quad \Omega_6 = B, \quad \Omega_7 = C, \quad \Omega_8 = D.$$

In what follows we will assume that $\bar{x}_1^1 < u_x$, $\bar{y}_2^1 < u_y$, $\bar{x}_3^2 > l_x$, and $\bar{y}_4^2 > l_y$, since the analysis of the other cases is simpler. After simple computations, it can be seen that:

$$
\begin{aligned}
D_1 &= \{(a,b)\ :\ a \in [nu_y^m[\bar{x}_1^1]^{n-1}, nu_y^m u_x^{n-1}]\}\\
D_2 &= \{(a,b)\ :\ b \in [mu_x^n[\bar{y}_2^1]^{m-1}, mu_x^n u_y^{m-1}]\}\\
D_3 &= \{(a,b)\ :\ a \in [nl_y^m l_x^n, nl_y^m[\bar{x}_3^2]^{n-1}]\}\\
D_4 &= \{(a,b)\ :\ b \in [ml_x^n l_y^m, ml_x^n[\bar{y}_4^2]^{m-1}]\}\\
x_1(a) &= \left(\tfrac{a}{nu_y^m}\right)^{\frac{1}{n-1}} \quad \forall (a,b) \in D_1\\
y_2(b) &= \left(\tfrac{b}{mu_x^n}\right)^{\frac{1}{m-1}} \quad \forall (a,b) \in D_2\\
x_3(a) &= -\left(\tfrac{-a}{nl_y^m}\right)^{\frac{1}{n-1}} \quad \forall (a,b) \in D_3\\
y_4(b) &= -\left(\tfrac{-b}{ml_x^n}\right)^{\frac{1}{m-1}} \quad \forall (a,b) \in D_4.
\end{aligned}
$$

Let $y = mx + q$ be the line along which the diagonal of the box through the two vertices A and C lies, and let

$$
\begin{aligned}
\mathcal{B}^\ell &= \{(x,y) \in \mathcal{B}\ :\ y \leq mx + q\}\\
\mathcal{B}^u &= \{(x,y) \in \mathcal{B} :\ y \geq mx + q\},
\end{aligned}
$$

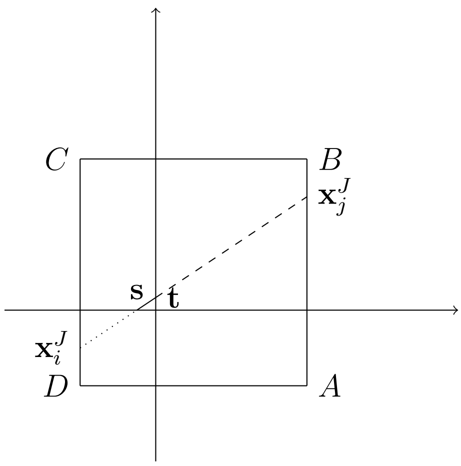
be the points in the box below and above the diagonal, respectively. The following observation shows that we can deal with $\mathcal{B}^\ell$ and $\mathcal{B}^u$ separately.

Observation 5.1 *For some $J \in \mathcal{J}$, let $i, j \in J$. Then, we must have either $\Omega_i, \Omega_j \subset \mathcal{B}^\ell$ or $\Omega_i, \Omega_j \subset \mathcal{B}^u$.*

Proof. Let us assume by contradiction that $\Omega_i \subset \mathcal{B}^\ell \setminus \mathcal{B}^u$ and $\Omega_j \subset \mathcal{B}^u \setminus \mathcal{B}^\ell$. If we are able to show that (17) holds, then we are lead to a contradiction. To see that (17) holds, consider $\mathbf{z} \in [\mathbf{x}_i^J, \mathbf{x}_j^J]$. We subdivide the segment $[\mathbf{x}_i^J, \mathbf{x}_j^J]$ into the three portions $[\mathbf{x}_i^J, \mathbf{s}]$, $[\mathbf{s}, \mathbf{t}]$ and $[\mathbf{t}, \mathbf{x}_j^J]$, where $\mathbf{s}$ and $\mathbf{t}$ are the intersection of the segment with the x-axis and with the y-axis, respectively (see also Figure 4). We should also consider the case when $\mathbf{s}$ and $\mathbf{t}$ are the intersection of the segment with the y-axis and with the x-axis, respectively, but we restrict to the previous case, since the discussion is analogous.

We first remark that the function $f(x,y) = x^n y^m$ is certainly strictly convex along $[\mathbf{x}_i^J, \mathbf{s}]$ and along $[\mathbf{t}, \mathbf{x}_j^J]$. Moreover, $f(\mathbf{x}_i^J), f(\mathbf{x}_j^J) > 0$. Then, if, e.g., $\mathbf{z} \in [\mathbf{x}_i^J, \mathbf{s}]$, (17) holds after choosing $\mathbf{z}_1 = \mathbf{x}_i^J$ and $\mathbf{z}_2 = \mathbf{s}$ (similar if $\mathbf{z} \in [\mathbf{t}, \mathbf{x}_j^J]$). If $\mathbf{z} \in [\mathbf{s}, \mathbf{t}]$, then (17) holds after choosing $\mathbf{z}_1 = \mathbf{s}$ and $\mathbf{z}_2 = \mathbf{t}$, since $f(\mathbf{s}) = f(\mathbf{t}) = 0 < f(\mathbf{x}_i^J), f(\mathbf{x}_j^J)$.

Figure 4: Subdivision of the segment $[\mathbf{x}_i^J, \mathbf{x}_j^J]$.



□

It is possible to restrict the attention to the subsets J such that $\Omega_k \subset \mathcal{B}^u$ for all $k \in J$. Indeed, the formulae for the other subsets can be immediately derived due to the symmetry of the function with respect to the origin. Thus, the sets to be considered are

$$\Omega_1 = \tilde{e}_1, \quad \Omega_2 = \tilde{e}_2, \quad \Omega_5 = A, \quad \Omega_6 = B, \quad \Omega_7 = C.$$

We remark that in view of Observation 3.1, $6 \in J$ implies $1, 7 \notin J$ and also $2, 5 \notin J$. Thus, Observation 3.2 allows to discard any set J such that $6 \in J$. After simple computations, it can be seen that:

$$\begin{aligned}
\eta_1(a,b) &= (1-n)u_y^m[x_1(a)]^n - bu_y \quad \forall (a,b) \in D_1 \\
\eta_2(a,b) &= (1-m)u_x^n[y_2(b)]^m - au_x \quad \forall (a,b) \in D_2 \\
\eta_5(a,b) &= u_x^n l_y^m - au_x - bl_y \\
\eta_7(a,b) &= l_x^n u_y^m - al_x - bu_y.
\end{aligned}$$

The pairs which need to be considered are:

$$J = \{2,7\}, \quad J = \{1,5\}, \quad J = \{1,2\}.$$

All the others can be discarded in view of Observation 3.2. The triples which need to be considered are

$$J = \{1,5,7\}, \quad J = \{2,5,7\}, \quad J = \{1,2,7\}, \quad J = \{1,2,5\}.$$

We remark that, according to Observation 3.1

$$1,7 \in J \quad \Rightarrow \quad x_1(a) = \bar{x}_1^1$$
$$2,5 \in J \quad \Rightarrow \quad y_2(b) = \bar{y}_2^1,$$

and, equivalently

$$1,7 \in J \quad \Rightarrow \quad a = \bar{a} = n u_y^m [\bar{x}_1^1]^{n-1}$$
$$2,5 \in J \quad \Rightarrow \quad b = \bar{b} = m u_x^n [\bar{y}_2^1]^{m-1}.$$

Now, let us consider the two triples $\{1,5,7\}$ and $\{2,5,7\}$. In both cases $\eta_5(a,b) = \eta_7(a,b)$, so that

$$b = \frac{a(u_x - l_x) + l_x^n u_y^m - u_x^n l_y^m}{u_y - l_y}. \tag{39}$$

If $J = \{1,5,7\}$, then

$$\eta_2(a,b) \geq \eta_5(a,b) \quad \Rightarrow \quad b \leq \bar{b}, \quad \eta_1(a,b) = \eta_7(a,b) \quad \Rightarrow \quad a = \bar{a}, \tag{40}$$

while if $J = \{2,5,7\}$, then

$$\eta_1(a,b) \geq \eta_7(a,b) \quad \Rightarrow \quad a \leq \bar{a}, \quad \eta_2(a,b) = \eta_5(a,b) \quad \Rightarrow \quad b = \bar{b}. \tag{41}$$

Due to (39), (40) and (41) can hold at the same time only if $\eta_1(a,b) = \eta_2(a,b) = \eta_5(a,b) = \eta_7(a,b)$. In all the other cases, only one of the two triples is acceptable. In particular, if

$$\frac{\bar{a}(u_x - l_x) + l_x^n u_y^m - u_x^n l_y^m}{u_y - l_y} < \bar{b}, \tag{42}$$

then, we can restrict our attention to $J = \{1,5,7\}$, while if

$$\frac{\bar{a}(u_x - l_x) + l_x^n u_y^m - u_x^n l_y^m}{u_y - l_y} > \bar{b} \tag{43}$$

then, we can restrict our attention to $J = \{2,5,7\}$. In fact, the two conditions (42) and (43) allow to restrict the attention to a collection $\mathcal{J}$ made up by only four sets. For instance, if (42) holds, then $J = \{2,7\}$ can be removed. Indeed, in such case

$$\eta_1(a,b) > \eta_7(a,b) \;\Rightarrow\; a < \bar{a}, \quad \eta_5(a,b) > \eta_2(a,b) \;\Rightarrow\; b > \bar{b},$$

while

$$\eta_2(a,b) = \eta_7(a,b) \;\Rightarrow (1-m) u_x^n [y_2(b)]^m + b u_y = a(u_x - l_x) + l_x^n u_y^m.$$

The left-hand side is an increasing function with respect to a, so that the above equation implies that b is an increasing function with respect to a. Moreover, for $b = \bar{b}$ the equation reduces to

$$\bar{b} = \frac{a(u_x - l_x) + l_x^n u_y^m - u_x^n l_y^m}{u_y - l_y},$$

which, in view of (42) implies $a > \bar{a}$, which is not possible. In a similar way it can be seen that if (42) holds, then we can remove the set $J = \{1, 2, 7\}$. In conclusion, if (42) holds, then we can restrict our attention to the four sets

$$\{1,5,7\}, \quad \{1,5\}, \quad \{1,2,5\}, \quad \{1,2\},$$

while in an analogous way it can be seen that if (43) holds, we can restrict our attention to the four sets

$$\{2,5,7\}, \quad \{2,7\}, \quad \{1,2,7\}, \quad \{1,2\}.$$

The case

$$\frac{\bar{a}(u_x - l_x) + l_x^n u_y^m - u_x^n l_y^m}{u_y - l_y} = \bar{b}$$

is a rather peculiar one where

$$conv_{f,\mathcal{B}}(x,y) = l_x u_y - \bar{a}x - \bar{b}y, \quad \forall (x,y) \in chull\{A,C,K,L\},$$

while over $chull\{B,K,L\}$ the convex envelope can be computed by solving the system associated to the set $J = \{1, 2\}$.

In what follows we only discuss the case where (42) holds, since the other cases are analogous. Figure 5 displays the polyhedral subdivision induced by the convex envelope in this case.

5.1 Set $\{1,5,7\}$

In this case

$$\eta_1(a,b) = \eta_7(a,b) \;\Rightarrow\; a = \bar{a}, \quad \eta_2(a,b) > \eta_5(a,b) \;\Rightarrow\; b < \bar{b}.$$

The solution of the system (18) is

$$a_1 = \bar{a}, \qquad b_1 = \frac{a_1(u_x - l_x) + l_x^n u_y^m - u_x^n l_y^m}{u_y - l_y} < \bar{b}.$$

We should also check whether $\eta_3(a,b), \eta_4(a,b), \eta_8(a,b) \geq \eta_5(a,b)$. It is enough to observe that $a_1 > 0$ and that $b_1 > 0$. The latter follows by observing that

$$a_1(u_x - l_x) + l_x^n u_y^m = u_y^m \left[n[\bar{x}_1^1]^{n-1}(u_x - l_x) + l_x^n \right] > 0.$$

In view of the definition (38) of $\bar{x}_1^1$, this is equivalent to prove that

$$n[\bar{x}_1^1]^{n-1} u_x + (1-n)[\bar{x}_1^1]^n > 0,$$

or, equivalently

$$n u_x > (n-1)\bar{x}_1^1,$$

which certainly holds. Then, $a_1 \in D_3^+$ and $b_1 \in D_4^+$. Indeed, D_3 and D_4 only contain negative a and b values, respectively. Thus, for these a and b values it holds that $\eta_8(a,b) > \eta_3(a,b) = \eta_4(a,b) = \eta_5(a,b) = \eta_7(a,b)$.

In conclusion, we have $\Gamma_J = chull\{A, C, K\}$, and

$$conv_{f,\mathcal{B}}(x,y) = u_x^n l_y^m - a_1(u_x - x) - b_1(l_y - y) \quad \forall (x,y) \in \Gamma_J.$$

5.2 Set $\{1, 2, 5\}$

In this case

$$\eta_7(a,b) > \eta_1(a,b) \;\Rightarrow\; a > \bar{a}, \quad \eta_2(a,b) = \eta_5(a,b) \;\Rightarrow\; b = \bar{b},$$

and the solution of the system (18) is

$$b_2 = \bar{b}, \quad a_2 = \frac{b_2(u_y - l_y) - l_x^n u_y^m + u_x^n l_y^m}{u_x - l_x} > \bar{a}.$$

Then, after defining $M = (x_1(a_2), u_y)$, we have $\Gamma_J = chull\{A, M, L\}$, and

$$conv_{f,\mathcal{B}}(x,y) = u_x^n l_y^m - a_2(u_x - x) - b_2(l_y - y) \quad \forall (x,y) \in \Gamma_J.$$

Note that also in this case we should check whether $\eta_3(a,b), \eta_4(a,b), \eta_8(a,b) \geq \eta_5(a,b)$, but the proof is analogous to the one in the previous case.

5.3 Set $\{1, 5\}$

In this case (19) is

$$\begin{aligned} &\lambda x_1(a) + (1-\lambda)u_x = x \\ &\lambda u_y + (1-\lambda) l_y = y \\ &(1-n)u_y^m [x_1(a)]^n + a u_x = u_x^n l_y^m + b(u_y - l_y), \end{aligned}$$

whose solution (parametric with respect to x, y) is

$$\begin{aligned} &\lambda_3(x,y) = \tfrac{y - l_y}{u_y - l_y} \\ &a_3(x,y) = n u_y^m \left[\tfrac{u_x y - u_x u_y + u_y x - l_y x}{y - l_y}\right]^{n-1} \qquad \text{(follows from } \; x_1(a) = u_x - \tfrac{u_x - x}{\lambda}) \\ &b_3(x,y) = \frac{(1-n)u_y^m \left[\frac{u_x y - u_x u_y + u_y x - l_y x}{y - l_y}\right]^n + a_3(x,y)u_x - u_x^n l_y^m}{u_y - l_y}. \end{aligned}$$

Then, $\Gamma_J = chull\{A, M, K\}$, and

$$conv_{f,\mathcal{B}}(x,y) = u_x^n l_y^m - a_3(x,y)(u_x - x) - b_3(x,y)(l_y - y) \quad \forall (x,y) \in \Gamma_J$$

(we omit to prove that $\eta_3(a,b), \eta_4(a,b), \eta_8(a,b) \geq \eta_5(a,b)$).

5.4 Set $\{1,2\}$

In this case (19) is

$$\begin{aligned} &\lambda x_1(a) + (1-\lambda)u_x = x \\ &\lambda u_y + (1-\lambda)y_2(b) = y \\ &(1-n)u_y^m[x_1(a)]^n + au_x = (1-m)u_x^n[y_2(b)]^m + bu_y. \end{aligned} \tag{44}$$

If we denote by $\lambda_4(x,y), a_4(x,y), b_4(x,y)$ the parametric solution of the system, then $\Gamma_J = chull\{B, M, L\}$, and $\forall (x,y) \in \Gamma_J$

$$conv_{f,\mathcal{B}}(x,y) = [x_1(a_4(x,y))]^n[y_2(b_4(x,y))]^m - a_4(x,y)(x_1(a_4(x,y))-x) - b_4(x,y)(y_2(b_4(x,y))-y)$$

(again, we omit to prove that $\eta_3(a,b), \eta_4(a,b), \eta_8(a,b) \geq \eta_1(a,b)$). If $m \neq n$, we are not able to derive a closed form formula for the parametric solution of the system (44). Thus, a_4, b_4 are implicitly defined as solutions of the system (44). If $m = n$ (the case already discussed in [13]), then it turns out that the last equation in (44) is equivalent to

$$y_2(b) = \frac{u_y}{u_x}x_1(a), \quad b = \frac{u_x}{u_y}a.$$

Thus, the first two equations become

$$\begin{aligned} &\lambda x_1(a) + (1-\lambda)u_x = x \\ &\lambda u_y + (1-\lambda)\tfrac{u_y}{u_x}x_1(a) = y, \end{aligned}$$

or, equivalently

$$\begin{aligned} \lambda &= \tfrac{u_x - x}{u_x - x_1(a)} \\ \lambda &= \tfrac{u_x y - u_y x_1(a)}{u_x u_y - u_y x_1(a)}, \end{aligned}$$

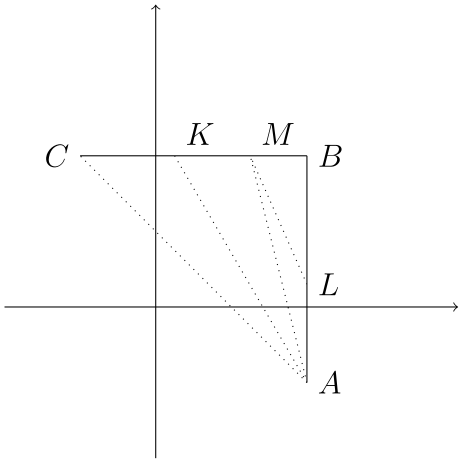
and, consequently

$$x_1(a) = \frac{u_x y + u_y x - u_x u_y}{u_y}, \quad y_2(b) = \frac{u_x y + u_y x - u_x u_y}{u_x}.$$

In view of the definition of $x_1(a)$ and $y_2(b)$ we also have

$$\begin{aligned} a_4(x,y) &= m u_y \left[u_x y + u_y x - u_x u_y\right]^{m-1} \\ b_4(x,y) &= m u_x \left[u_x y + u_y x - u_x u_y\right]^{m-1}. \end{aligned}$$

Figure 5: Polyhedral subdivision for the convex envelope of $f(x,y) = x^n y^m$ over $\mathcal{B}^u$ if (43) holds.



6 Conclusions

In this paper we proposed a technique for the computation of the convex envelope of some bivariate functions over polytopes based on the parametric solution of a KKT system. The technique has been applied in order to derive the convex envelope of the bilinear function over any polytope, and the convex envelope of the product of power functions over boxes. The technique could be applied to derive the convex envelope of other functions, such as the fractional one over polytopes. Moreover, the proposed approach could be extended to n-dimensional functions. In this case the computation of the n-dimensional functions η_k and the derivation of closed-form formulae of the convex envelope can be much more complicated. However, some special cases are still manageable (we refer to [14] for the derivation of the convex envelope of some quadratic functions over the n-dimensional unit simplex).

References

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

[2] K. M. Anstreicher, S. Burer, Computable representations for convex hulls of low-dimensional quadratic forms, *Mathematical Programming B*, 124:33–43, (2010)

[3] K.M. Anstreicher, On convex relaxations for quadratically constrained quadratic programming, *Mathematical Programming*, 136:233-251 (2012)

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

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

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

[8] A. Khajavirad, N.V. Sahinidis, Convex envelopes of products of convex and component-wise concave functions, *Journal of Global Optimization*, 51:391-409 (2012)

[9] A. Khajavirad, N.V. Sahinidis, Convex envelopes generated from finitely many compact convex sets, *Mathematical Programming*, 137:371-408 (2013)

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

[11] M. Locatelli, F.Schoen, On convex envelopes for bivariate functions over polytopes, *Mathematical Programming*, 144:65–91 (2014)

[12] M. Locatelli, Polyhedral subdivisions and functional forms for the convex envelopes of bilinear, fractional and other bivariate functions over general polytopes, *Journal of Global Optimization*, 66: 629–668 (2016)

[13] M. Locatelli, A technique to derive the analytical form of convex envelopes for some bivariate functions, *Journal of Global Optimization*, 59, 477-501 (2014)

[14] M. Locatelli, Convex envelopes of some quadratic functions over the n-dimensional unit simplex, *SIAM Journal on Optimization*, 25, 589–621 (2015)

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

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

[17] J. E. Mitchell, J.-S. Pang, B. Yu, Convex Quadratic Relaxations of Nonconvex Quadratically Constrained Quadratic Programs, *Optimization Methods and Software*, 29(1), 120–136 (2014)

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

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

[20] 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).

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

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

[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] F. Tardella, Existence and sum decomposition of vertex polyhedral convex envelopes, *Optimization Letters*, 2: 363-375 (2008)

[25] M. Tawarmalani, J. -P. P. Richard, C. Xiong, Explicit convex and concave envelopes through polyhedral subdivisions, *Mathematical Programming*, 138:531-577 (2013)

[26] 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).

A Conditions under which $J = \{i, j, k\}$ implies $\Omega_i, \Omega_j, \Omega_k \in V(P)$

In Section 4 we have seen that when $f(x, y) = xy$, $J = \{i, j, k\}$ is possible only if $\Omega_i, \Omega_j, \Omega_k \in V(P)$. In other words, this means that affine pieces of the convex envelope are possible only over polytopes whose vertices are a subset of $V(P)$. This is not always true. For instance, consider the function $f(x, y) = (x-1)^2(y-1)^2$ and $P = [0, 2]^2$. In this case the convex envelope is equal to 0 over the square with vertices $(0, 1)$, $(1, 2)$, $(2, 1)$, and $(1, 0)$, and none of these points is a vertex of P.

We might wonder whether there are general conditions under which we can guarantee the same result as for the bilinear case. We introduce the following assumption.

Assumption A.1 *Function $f(x, y)$ is either strictly convex or concave over each line $y = mx + q$. Moreover, for each triple of not aligned points (x_r, y_r), $r = i, j, k$,*

$$(x_k, y_k) \in \bar{E}(P)\setminus V(P), \quad m_k(x_j - x_i) + y_i - y_j = 0 \;\Rightarrow\; g_k'(x_k)(x_i - x_j) \neq f(x_i, y_i) - f(x_j, y_j),$$

where $y = m_k x + q_k$ is the line along which the edge containing (x_k, y_k) lies, while

$$g_k(x_k) = f(x_k, m_k x_k + q_k)$$

is the restriction of f along this line.

Note that the previous example does not fulfill this assumption. Indeed, consider the points $(x_k, y_k) = (1, 2)$, $(x_i, y_i) = (0, 1)$, and $(x_j, y_j) = (2, 1)$. We have that $m_k = 0$ and $q_k = 2$. Then, the left-hand side of the implication holds, but $g_k'(x_k) = 0$, $f(x_i, y_i) = f(x_j, y_j) = 0$, so that Assumption A.1 is not fulfilled. We prove the following observation.

Observation A.1 *Let Assumption A.1 hold and $J = \{i, j, k\}$. Then $\Omega_r \in V(P)$, $r = i, j, k$.*

Proof. We prove this by contradiction. Let $\Omega_k \in \bar{E}(P)$ and let $\mathbf{x}_r^J = (x_r, y_r)$, $r = i, j, k$, be derived from the solution of system (18). Assume by contradiction that $(x_k, y_k) \in \bar{E}(P)\setminus V(P)$, i.e., $x_k \in (x_k^1, x_k^2)$. Let $y = m_k x + q_k$ be the line along which Ω_k lies. Without loss of generality we assume that $P \subset \{(x, y) \;:\; y \leq m_k x + q_k\}$. We also assume that $dim(chull\{\mathbf{x}_i^J, \mathbf{x}_j^J, \mathbf{x}_k^J\}) = 2$ (otherwise, similarly to Observation 3.2, we can discard J). Let $(x, y) \in int(chull\{\mathbf{x}_i^J, \mathbf{x}_j^J, \mathbf{x}_k^J\})$, i.e.,

$$(x, y) = \lambda_k^*(x_k, y_k) + \lambda_j^*(x_j, y_j) + \lambda_i^*(x_i, y_i),$$

with $\lambda_k^*, \lambda_j^*, \lambda_i^* > 0$ and $\lambda_k^* + \lambda_j^* + \lambda_i^* = 1$. In (1) fix $\lambda_s = 0$, $\forall s \notin J$, fix (x_i, y_i) and (x_j, y_j) and let $x_k', \lambda_i, \lambda_j, \lambda_k$ be variables, while we impose $y_k' = m_k x_k' + q_k$, i.e., we allow to move the point (x'_k, y'_k) only along the edge Ω_k. Then, we end up with the following problem

$$
\begin{aligned}
\min\quad & \lambda_k f(x'_k, m_k x'_k + q_k) + \lambda_j f(x_j, y_j) + \lambda_i f(x_i, y_i) \\
& \lambda_k x'_k + \lambda_j x_j + \lambda_i x_i = x \\
& \lambda_k (m_k x'_k + q_k) + \lambda_j y_j + \lambda_i y_i = y \\
& \lambda_k + \lambda_j + \lambda_i = 1 = 1 \\
& \lambda_k, \lambda_j, \lambda_i \geq 0 \\
& x_k^1 \leq x'_k \leq x_k^2.
\end{aligned}
$$

In view of the assumption by contradiction we must have that $x_k, \lambda_k^*, \lambda_j^*, \lambda_i^*$ is an optimal solution of this problem. Thus, it must satisfy the KKT conditions (it is easily seen that a constraint qualification is satisfied). We remark that $\lambda_k^*, \lambda_j^*, \lambda_i^* > 0$ and $x_k \in (x_k^1, x_k^2)$ imply that the Lagrange multipliers of the corresponding constraints are null. Thus, the KKT conditions read as follows:

$$
\begin{aligned}
& g_k(x_k) + \mu_1 x_k + \mu_2 m_k x_k + \mu_2 q_k + \mu_3 = 0 \\
& f(x_i, y_i) + \mu_1 x_i + \mu_2 y_i + \mu_3 = 0 \\
& f(x_j, y_j) + \mu_1 x_i + \mu_2 y_i + \mu_3 = 0 \\
& \lambda_1^* g'_k(x_k) + \lambda_1^* \mu_1 + \lambda_1^* m_k \mu_2 = 0,
\end{aligned}
$$

where μ_1, μ_2, μ_3 are the Lagrange multipliers of the three equality constraints. Since $\lambda_1^* > 0$, the last equation becomes $g'_k(x_k) + \mu_1 + m_k \mu_2 = 0$. After a few computations, the KKT conditions lead to the following equation

$$
g'_k(x_k)[(x_i - x_k)A_j + (x_j - x_k)B_i] + g_k(x_k)(A_j + B_i) = B_i f(x_j, y_j) + A_j f(x_i, y_i), \tag{45}
$$

where

$$
\begin{aligned}
A_j &= m_k x_j + q_k - y_j > 0 \\
B_i &= y_i - m_k x_i - q_k < 0.
\end{aligned}
$$

Note that $A_j + B_i = m_k(x_j - x_i) + y_i - y_j$. If $A_j + B_i = 0$, then the above equation becomes

$$
g'_k(x_k)(x_i - x_j) = f(x_i, y_i) - f(x_j, y_j),
$$

which is not possible in view of Assumption (A.1). Thus, let us assume that $A_j + B_i \neq 0$. Then, after setting

$$
\gamma = \frac{A_j}{A_j + B_i},
$$

we can rewrite (45) as follows

$$
g'_k(x_k)[\gamma x_i + (1 - \gamma)x_j - x_k] + g_k(x_k) = (1 - \gamma)f(x_j, y_j) + \gamma f(x_i, y_i). \tag{46}
$$

We remark that $A_j > 0, B_i < 0$ imply that $\gamma \notin [0,1]$. We also remark that in view of Assumption A.1, g_k is strictly convex over the whole line $y = m_k x + q_k$, so that

$$g_k(\gamma x_i + (1-\gamma)x_j) > g_k'(x_k)[\gamma x_i + (1-\gamma)x_j - x_k] + g_k(x_k),$$

provided that $\gamma x_i + (1-\gamma)x_j \neq x_k$. We first prove that equality can not hold. Assume that $\gamma x_i + (1-\gamma)x_j = x_k$. By the definition of γ, it turns out that

$$(1-\gamma)y_j + \gamma y_i = m_k((1-\gamma)x_j + \gamma x_i) + q_k, \tag{47}$$

which, together with $\gamma x_i + (1-\gamma)x_j = x_k$, implies that (x_i, y_i), (x_j, y_j) and (x_k, y_k) are aligned, which is not possible in view of Observation 3.2. Moreover, in view of Corollary 3.2, f must be concave along the line through (x_i, y_i) and (x_j, y_j). Thus, $\gamma \notin [0,1]$ implies

$$(1-\gamma)f(x_j, y_j) + \gamma f(x_i, y_i) \geq f((1-\gamma)x_j + \gamma x_i, (1-\gamma)y_j + \gamma y_i).$$

Thus, from (46) we should have

$$g_k(\gamma x_i + (1-\gamma)x_j) > f((1-\gamma)x_j + \gamma x_i, (1-\gamma)y_j + \gamma y_i).$$

In view of (47) we have

$$f((1-\gamma)x_j + \gamma x_i, (1-\gamma)y_j + \gamma y_i) = g_k((1-\gamma)x_j + \gamma x_i),$$

i.e.,

$$g_k(\gamma x_i + (1-\gamma)x_j) > g_k(\gamma x_i + (1-\gamma)x_j),$$

which is, obviously, a contradiction. □

We remark that the bilinear function satisfies Assumption A.1. Indeed, along each line $y = mx + q$ the restriction of the bilinear function is $mx^2 + qx$ which is either a concave (if $m \leq 0$) or a strictly convex (if $m > 0$) function. Moreover, $m_k(x_j - x_i) + y_i - y_j = 0$ can not hold, since in this case we would have

$$\frac{y_i - y_j}{x_i - x_j} = m_k > 0,$$

which is not possible in view of Corollary 3.2.

In some cases, Assumption A.1 is not fulfilled since the restriction of the function along a line is not defined over the whole line. This is the case, e.g., of the fractional function $\frac{y}{x}$, which is not defined at $x = 0$. All the same the above proof can be sometimes adapted. We show this for the fractional function.

Observation A.2 *Let $f(x,y) = \frac{y}{x}$ and $P \subset \{(x,y) \ : \ x > 0\}$. Then, for $J = \{i,j,k\}$ it must hold that $\Omega_r \in V(P)$, $r = i,j,k$.*

Proof. Note that the restriction of $f(x,y) = \frac{y}{x}$ over a line $y = mx + q$ is either concave or strictly convex for $x > 0$. Indeed,

$$f(x, mx+q) = m + \frac{q}{x},$$

so that the restriction is concave if $q \le 0$, and strictly convex if $q > 0$. We remark that in view of Corollary 3.2, we must have that the restriction of f along the segments through the points (x_i, y_i) and (x_j, y_j), through (x_i, y_i) and (x_k, y_k), and through (x_k, y_k) and (x_j, y_j) must be concave, which implies that

$$\begin{aligned} (x_i - x_j)\left(\tfrac{y_i}{x_i} - \tfrac{y_j}{x_j}\right) &\ge 0 \\ (x_i - x_k)\left(\tfrac{y_i}{x_i} - \tfrac{y_k}{x_k}\right) &\ge 0 \\ (x_k - x_j)\left(\tfrac{y_k}{x_k} - \tfrac{y_j}{x_j}\right) &\ge 0, \end{aligned} \tag{48}$$

and, moreover, one of the two factors in each of the above inequalities must be non null. The proof of Observation A.1 can be repeated up to

$$g_k'(x_k)[(x_i - x_k)A_j + (x_j - x_k)B_i] + g_k(x_k)(A_j + B_i) = B_i f(x_j, y_j) + A_j f(x_i, y_i).$$

We observe that in case $A_i + B_j = 0$,

$$g_k'(x_k)(x_i - x_j) = f(x_i, y_i) - f(x_j, y_j) = \frac{y_i}{x_i} - \frac{y_j}{x_j},$$

can not hold. Indeed, in view of (48) the left-hand side has opposite sign with respect to the right-hand side since $g_k'(x_k) < 0$ (g_k is a decreasing function). If $A_j + B_i \ne 0$, we end up again with (46). If $\gamma x_i + (1-\gamma)x_j > 0$, then we can repeat the previous proof. But if $\gamma x_i + (1-\gamma)x_j \le 0$, we can not repeat it and we need some different argument. Let us rewrite (46) as follows

$$g_k'(x_k)[x_j + \gamma(x_i - x_j)] = (1-\gamma)\left[\frac{y_j}{x_j} - \frac{y_k}{x_k}\right] + \gamma\left[\frac{y_i}{x_i} - \frac{y_k}{x_k}\right].$$

We assume that $x_i > x_j$ (the proof for the other case is analogous). Then, (48) implies $\frac{y_i}{x_i} - \frac{y_j}{x_j} > 0$. Recall that $\gamma x_i + (1-\gamma)x_j \le 0$, otherwise we can repeat the previous proof. In particular, this implies $\gamma < 0$, i.e., $A_j + B_i < 0$. Since $y_i > x_i \frac{y_j}{x_j}$,

$$A_j + B_i = y_i - y_j + m_k(x_j - x_i) > \frac{y_j}{x_j}(x_i - x_j) + m_k(x_j - x_i).$$

Moreover,

$$\frac{y_k}{x_k} = m_k + \frac{q_k}{x_k},$$

so that

$$A_j + B_i > (x_i - x_j)\left[\frac{y_j}{x_j} - \frac{y_k}{x_k} + \frac{q_k}{x_k}\right].$$

If $x_j \geq x_k$, then $\frac{y_j}{x_j} - \frac{y_k}{x_k} > 0$, so that $A_j + B_i > 0$ and $\gamma > 0$. Thus, we restrict the attention to $x_k > x_j$. If $x_j < x_k \leq x_i$, then

$$g_k'(x_k)[x_j + \gamma(x_i - x_j)] > 0,$$

while

$$(1-\gamma)\left[\frac{y_j}{x_j} - \frac{y_k}{x_k}\right] + \gamma\left[\frac{y_i}{x_i} - \frac{y_k}{x_k}\right] < 0,$$

which is not possible. Thus, we are only left with $x_k > x_i > x_j$. In this case the left-hand side is still positive. We prove that the right-hand side is negative. We need to show that

$$(1-\gamma)\left[\frac{y_k}{x_k} - \frac{y_j}{x_j}\right] > -\gamma\left[\frac{y_k}{x_k} - \frac{y_i}{x_i}\right].$$

This is true since $1 - \gamma > -\gamma > 0$, while, in view of (48)

$$\frac{y_k}{x_k} - \frac{y_j}{x_j} > \frac{y_k}{x_k} - \frac{y_i}{x_i}.$$

□