Operations Research Letters 45 (2017) 421–426

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Convex and concave envelopes: Revisited and new perspectives

Walid Ben-Ameur [a], Adam Ouorou [b], Guanglei Wang [a,*]

[a] Samovar, Télécom Sudparis, CNRS, University Paris-Saclay, 9 rue Charles Fourier, 91011 Evry, France
[b] Orange Labs, 38-40 rue du General Leclerc, 92794 Issy-les-Moulineaux Cedex 9, France

ARTICLE INFO

Article history:
Received 24 October 2016
Received in revised form 28 June 2017
Accepted 28 June 2017
Available online 5 July 2017

Keywords:
Convex envelope
Finite cover
Polyhedral approximation

ABSTRACT

We discuss two approaches to approximate the convex and concave envelopes of bilinear functions over hypercubes. The first approach is based on a semidefinite program. The second approach considers some predefined cover sets of a hypercube and leads to a linear program. Then we establish a connection between the convex envelope of a bilinear function and the concave envelope of a piecewise linear function. Numerical experiments are conducted to compare the two approaches. As an extension, a novel approach is discussed.

1. Introduction

There has been a great amount of studies devoted to developing convex underestimators and concave overestimators of nonlinear functions $f(x)$ over various polyhedral sets. One of important motivations in optimization community is that, one can reformulate a complicated usually non-convex problem as an easier problem with convex representations of the objective function and constraints. Such convex relaxations can then be solved repeatedly in branch-and-bound algorithms, where searching spaces are refined in a convergent way. In general, the computational efficiency of branch-and-bound is greatly influenced by the strength of convex relaxations.

Among such investigations, the construction and approximation of convex and concave envelopes of bilinear functions over boxes draw much attention. This is mainly due to the following reasons. First, many important problems involve bilinear terms. Second, boxes are naturally outer approximations of polytopes. Therefore the estimators of convex and concave envelopes of f over a box are also valid estimators for f over polytopes contained in the box. In addition, most branch-and-bound algorithms partition the searching space by divisions of the feasible region into boxes. This also motivates investigations in deriving strong envelope estimators over boxes. It is well known that an n-dimensional box is a linear transformation of a hypercube. Thus we consider in this paper the convex (resp. concave) envelope of a bilinear function (1) over a hypercube $H_n = [0, 1]^n$:

$$f(x) = \frac{1}{2}x^T Qx = \sum_{i<j} Q_{ij}x_ix_j \tag{1}$$

where $Q_{ij} \in \mathbb{R}$ $(1 \le i < j \le n)$, $Q_{ii} = 0$ $(i = 1, \ldots, n)$.

There are two main research directions on convex and concave envelopes of f over H_n. The first direction is motivated by the definition that the convex envelope of any bounded f over any non-empty compact set P is a function $\check{f} : P \ni x \mapsto \check{f}(x) = \inf\{u : (x, u) \in \text{conv epi} f\}$, where $\text{epi} f = \{(x, u) : f(x) \le u\}$ and $\text{conv epi} f$ represents its convex hull. Geometrically, one can interpret $\check{f}$ as the bottom of the convex hull of the epigraph of f over P. Since P is compact and f is continuous, this definition is equivalent to

$$\min_{\lambda_i} \left\{ \sum_{v_i \in P} \lambda_i f(v_i) : \lambda_i v_i = x, \sum_{v_i \in P} \lambda_i = 1, \lambda_i \ge 0 \right\}. \tag{2}$$

The concave envelope $\widehat{f}(x)$ of f over P is defined symmetrically by replacing minimization to maximization. When $P = H_n$, a nice property of bilinear function f due to Sherali [20] and Rikun [19] is that its convex and concave envelopes over a hypercube H_n are vertex polyhedral, i.e., the envelopes of f on H_n coincide with the envelopes of its restriction to the vertices of H_n. This property allows us to simplify definition (2) to

$$\min_{\lambda_i} \left\{ \sum_{i=1}^{2^n} \lambda_i f(v_i) : \sum_{i=1}^{2^n} \lambda_i v_i = x, \sum_{i=1}^{2^n} \lambda_i = 1, \lambda_i \ge 0, \forall i \right\}, \tag{3}$$

where v_i, $i \in \{1, \ldots, 2^n\}$ are vertices of H_n. Owning to the exponential number of vertices, formula (3) is rarely exploited.

The second research line can be treated as the dual of the first since a convex (resp. concave) envelope can be interpreted as the

* Corresponding author.
E-mail addresses: walid.benameur@telecom-sudparis.eu (W. Ben-Ameur), adam.ouorou@orange.com (A. Ouorou), guanglei.wang13@gmail.com (G. Wang).

http://dx.doi.org/10.1016/j.orl.2017.06.008
0167-6377/

pointwise supremum (resp. infimum) of an affine under-(resp. over-) estimator of f over P. Hence, we have

$$\check{f}(x) = \max_{\alpha,\gamma} \left\{\gamma : \alpha^T(x - v) + \gamma \le f(v),\ \forall\, v \in P\right\}, \tag{4}$$

where $(\alpha, \gamma) \in \mathbb{R}^{n+1}$ defines the supporting hyperplane of $\check{f}$ at point x. When $P = H_n$, one can replace set P with its extreme points due to the vertex polyhedral property of $\check{f}$. However, the number of extreme points is exponential. A cutting-plane algorithm is used in [5] to find a facet of the convex envelope by separating the supporting function $(\alpha, \gamma) \in \mathbb{R}^{n+1}$ at a point. Perhaps, the most known convex and concave envelopes approximation of f over a hypercube is McCormick inequalities [16]:

$$\max\{x_i + x_j - 1, 0\} \le x_i x_j \le \min\{x_i, x_j\},\ i < j. \tag{5}$$

It is known that term-wise McCormick relaxation characterizes the convex and concave envelope over rectangles [3]. In general, this is not true for higher dimensional cases. The triangle inequalities [18] in concert with McCormick inequalities characterize the convex hull of $f(x)$ over a cube ($n = 3$). As noticed in [17], the number of facets defining the convex and concave envelopes is indicated by the number of simplices induced by the *triangulations* of a hypercube. But it grows super-exponentially in the number of dimension.

A number of papers are also dedicated to giving explicit characterizations of convex and concave envelopes over special polytopes. Locatelli and Schoen [13] characterize convex envelopes for some general bivariate functions over various polytopes in $\mathbb{R}^2$. Linderoth [12] presents analytical formulas for f over disjoint triangular regions. Tawarmalani et al. [22] derive envelopes for several nonlinear functions via polyhedral divisions of a hyper-rectangle. Recently, Hijazi [10] characterized the convex hull of a bilinear term xy over the intersection between a rectangle and a dominance constraint $x \le y$.

To deal with difficulties in (3) and (4), one may consider extended formulations with moderate sizes. However, as observed in [15], extended formulations do not always bring improvements compared with McCormick relaxation (5). In particular, the authors show that if the coefficients of a bilinear function as (1) are nonnegative, the ratio of the gap between McCormick's over and under estimators over the gap of the convex and concave envelopes is always less than 2. Lasserre and Thanh [11] construct a convex polynomial $p_d \in \mathbb{R}[\mathbf{x}]$ (degree $\le d$) underestimating f, which can be approximated by solving a hierarchy of semidefinite programs (SDP).

We investigate in this paper the strength of different approaches on approximating the convex and concave envelopes of the bilinear f over H_n. In Section 2, we review an SDP based estimator. Then we propose a novel approximation method using linear programs (LP). In addition, we show that the convex (resp. concave) envelope of a bilinear function f is affinely equivalent to the concave (resp. convex) envelope of a piecewise linear function. In Section 3, we report a set of numerical results to compare these two approaches. Finally, we propose an interesting research direction.

Notation. For a matrix C, C^T denotes its transpose and C_{ij} represents its ith row jth column component. $C_{I,J}$ represents the submatrix composed by set I of columns and set J of rows of C. When $I = J$, C_I stands for $C_{I,I}$. For a set S, conv S denotes its convex hull and $|S|$ represents its cardinality. H_n stands for an n-dimensional hypercube. For a general function $f : H_n \mapsto \mathbb{R}$, $\check{f}$ (resp. $\widehat{f}$) denotes the convex (resp. concave) envelope of f over H_n. The inner product between two matrices $A, B \in \mathbb{R}^{m\times n}$ is denoted by $\langle A, B\rangle$. For a square matrix $A \in \mathbb{R}^{n\times n}$, $\mathrm{diag}(A) \in \mathbb{R}^n$ represents the vector of diagonal elements. For a vector $a \in \mathbb{R}^n$, $\mathrm{Diag}(a) \in \mathbb{S}^n$ stands for the $n \times n$ diagonal matrix filled by vector a. For a symmetric matrix $X \in \mathbb{S}^n$ and vector $x \in \mathbb{R}^n$, $\mathcal{S}(x, X)$ refers to the matrix $\begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}$ in $\mathbb{S}^{n+1}$. $\tau(X) = (X_{ij})_{1\le i\le j\le n}$ is the vectorization of the upper triangular part of the matrix X and $\tau_+(X) = (X_{ij})_{1\le i<j\le n}$ denotes the strictly upper triangular part.

2. Main results

In this section, we review an SDP based convex (resp. concave) underestimator (resp. overestimator) and propose a novel way to construct estimators based on LPs. Several theoretical results are established.

2.1. SDP based estimators

Given a general bilinear function f, we are interested in constructing convex and concave estimators that are valid over a hypercube. To this end, we consider the following assumption.

Assumption 1. A set of valid inequalities involving quadratic terms is available,

$$x^T A_k x + b_k^T x + d_k \le 0,\ \forall x \in H_n,\ \forall k \in \mathcal{K}, \tag{6}$$

where $\mathcal{K}$ is an index set.

Note that Assumption 1 in general holds for bilinear optimization problems over any compact set. For instance, such valid inequalities can be McCormick's inequalities (5), Padberg's triangle inequalities [18], RLT based inequalities [1]. The idea is to use these valid inequalities to construct strong estimators. We start by constructing a convex underestimator. We associate nonnegative multipliers $\alpha \in \mathbb{R}_+^{|\mathcal{K}|}$ with the set of inequalities. Then we add the negative terms $\sum_{k\in\mathcal{K}} \alpha_k(x^T A_k x + b_k^T x + d_k)$ to $f(x)$. For ease of presentation, we define $f_\alpha(x) = x^T Q(\alpha)x + c(\alpha)^T x + p(\alpha)$, where $Q(\alpha) = \frac{1}{2}Q + \sum_{k\in\mathcal{K}} \alpha_k A_k$, $c(\alpha) = \sum_{k\in\mathcal{K}} \alpha_k b_k$ and $p(\alpha) = \sum_{k\in\mathcal{K}} \alpha_k d_k$.

Recall that f_α is a convex underestimator of f over H_n if and only if $f_\alpha \le f$ and f_α is convex over the hypercube H_n. Obviously, it holds that $f_\alpha(x) \le f(x),\ \forall x \in H_n$. The convexity of f_α can be ensured by restricting $Q(\alpha)$ to be positive semidefinite. Let us denote by $\mathcal{A}$ the set of feasible choices of α with $\mathcal{A} := \left\{\alpha \in \mathbb{R}_+^n : Q(\alpha) \in \mathbb{S}_+^n\right\}$.

We ensure the non-emptiness of set $\mathcal{A}$ by explicitly adding valid inequalities $x_i^2 \le x_i$ $(i = 1, \ldots, n)$ to inequality system (6). As a consequence, it also follows that $\mathcal{A}$ has a nonempty interior. The strongest convex underestimator of the form f_a can be obtained by solving the following problem

$$\sup_{\alpha\in\mathcal{A}} f_\alpha(x). \tag{7}$$

Problem (7) is then strictly feasible. The dual of (7) reads

$$\begin{aligned} \inf_X \quad & \frac{1}{2}\langle Q, X + xx^T\rangle \\ \text{s.t.} \quad & \langle A_k, X + xx^T\rangle + b_k^T x + d_k \le 0,\ \forall k \in \mathcal{K}, \\ & X \in \mathbb{S}_+^n. \end{aligned} \tag{8}$$

The feasible region of (8) is nonempty (e.g., take X as the null matrix of size n). In addition, as (7) is strictly feasible, the dual optimum is attained and is equal to the primal optimum of (7). By change of variables $X' = X + xx^T$, one can rewrite (8) as

$$\min_{X'} \left\{\frac{1}{2}\langle Q, X'\rangle : (x, X') \in \mathbf{K}\right\}, \tag{9}$$

where $\mathbf{K} := \{(x, X') \in H_n \times \mathbb{S}^n : \langle A_k, X'\rangle + b_k^T x + d_k \le 0,\ \forall k \in \mathcal{K}, \mathcal{S}(x, X') \succeq 0\}$. Notice that when x is not fixed, problem (9) is exactly the Shor's relaxation in conjunction with valid inequalities. Formally, we summarize the foregoing as follows.