Using mixed volume theory to compute the convex hull volume for trilinear monomials

23rd Combinatorial Optimization Workshop, Aussois
January 9, 2019

Institute of Mathematical Optimization, Otto-von-Guericke-University, Magdeburg, Germany.
• Using volume to compare relaxations for mixed-integer nonlinear optimization

• The convex hull of the graph of a trilinear monomial over a box

• Use techniques from mixed volume theory to obtain an alternative proof
Global optimization of non-convex functions is hard!

\[
\min_{x \in \mathbb{R}^n, z \in \mathbb{Z}^m} \{ f(x, z) : (x, z) \in \mathcal{F} \}
\]

- Global optimization of a mixed integer non-linear optimization (MINLO) problem
- Not necessarily convex sets/functions
Global optimization of non-convex functions is hard!

\[
\min_{x \in \mathbb{R}^n, z \in \mathbb{Z}^m} \{ f(x, z) : (x, z) \in \mathcal{F} \}
\]

- Global optimization of a mixed integer non-linear optimization (MINLO) problem
- Not necessarily convex sets/functions

**Software:** Baron, Couenne, Scip, Antigone...
Global optimization of non-convex functions is hard!

\[
\min_{x \in \mathbb{R}^n, z \in \mathbb{Z}^m} \{ f(x, z) : (x, z) \in \mathcal{F} \}
\]

- Global optimization of a mixed integer non-linear optimization (MINLO) problem
- Not necessarily convex sets/functions

Software: Baron, Couenne, Scip, Antigone...

- Each of these perform some variation on the algorithm known as **Spatial Branch-and-Bound** (sBB)
Global optimization of non-convex functions is hard!

\[
\min_{x \in \mathbb{R}^n, z \in \mathbb{Z}^m} \{ f(x, z) : (x, z) \in \mathcal{F} \}
\]

- Global optimization of a mixed integer non-linear optimization (MINLO) problem
- Not necessarily convex sets/functions

**Software:** Baron, Couenne, Scip, Antigone...

- Each of these perform some variation on the algorithm known as **Spatial Branch-and-Bound** (sBB)
- sBB has a daunting complexity
Global optimization of non-convex functions is hard!

\[
\min_{x \in \mathbb{R}^n, z \in \mathbb{Z}^m} \{ f(x, z) : (x, z) \in \mathcal{F} \}
\]

- Global optimization of a mixed integer non-linear optimization (MINLO) problem
- Not necessarily convex sets/functions

**Software:** Baron, Couenne, Scip, Antigone...

- Each of these perform some variation on the algorithm known as **Spatial Branch-and-Bound** (sBB)
- sBB has a daunting complexity
- How can we effectively tune/engineer this software?
  - experimentally
  - mathematically
Key algorithm: sBB

To find optimal solutions we use **Spatial Branch-and-Bound**: 

- Create sub-problems by branching on individual variables
- Generate convex relaxations of the graph of the function at each node
The choice of convexification method is important

Computational tractability of sBB depends on the quality of the convexifications we use.

We want both:

- Tight relaxations (good bounds)
- Simple algebraic representations of feasible regions (solve quickly)

We have a trade off:
We compare the tightness of convexifications via \( n \)-dimensional volume

- Allows us to quantify the difference between formulations analytically

- The optimal solution could occur anywhere in the feasible region, and therefore the volume measure corresponds to a uniform distribution on the location of the optimal solution

- Volume was introduced as a means of comparing formulations by Lee and Morris (1994)

- Recent survey on volumetric comparison of polyhedral relaxations for optimization Lee, Skipper, Speakman (2018)
Trilinear monomials: some notation

Assume we have a monomial of the form: $f = x_1x_2x_3$. 
Trilinear monomials: some notation

Assume we have a monomial of the form: $f = x_1x_2x_3$.

- $x_i \in [a_i, b_i]$, where $0 \leq a_i < b_i$, for $i = 1, 2, 3$
Trilinear monomials: some notation

Assume we have a monomial of the form: $f = x_1x_2x_3$.

- $x_i \in [a_i, b_i]$, where $0 \leq a_i < b_i$, for $i = 1, 2, 3$
- Label the variables $x_1$, $x_2$, and $x_3$ such that:

$$\frac{a_1}{b_1} \leq \frac{a_2}{b_2} \leq \frac{a_3}{b_3} \; \; \; \; (\Omega)$$
Trilinear monomials: some notation

Assume we have a monomial of the form: \( f = x_1x_2x_3. \)

- \( x_i \in [a_i, b_i], \) where \( 0 \leq a_i < b_i, \) for \( i = 1, 2, 3 \)
- Label the variables \( x_1, x_2 \) and \( x_3 \) such that:
  \[
  \frac{a_1}{b_1} \leq \frac{a_2}{b_2} \leq \frac{a_3}{b_3} \tag{\Omega}
  \]

The convex hull of the graph \( f = x_1x_2x_3 \) (on the domain \( x_i \in [a_i, b_i] \)) is polyhedral, we refer to this polytope as \( \mathcal{P}_H. \)
**Trilinear monomials: some notation**

Assume we have a monomial of the form: \( f = x_1 x_2 x_3 \).

- \( x_i \in [a_i, b_i] \), where \( 0 \leq a_i < b_i \), for \( i = 1, 2, 3 \)
- Label the variables \( x_1, x_2 \) and \( x_3 \) such that:
  \[
  \frac{a_1}{b_1} \leq \frac{a_2}{b_2} \leq \frac{a_3}{b_3}
  \]  
  \((\Omega)\)

The convex hull of the graph \( f = x_1 x_2 x_3 \) (on the domain \( x_i \in [a_i, b_i] \)) is polyhedral, we refer to this polytope as \( P_H \).

- The facet description was given by Meyer and Floudas (2004)
- There are alternative convexifications for trilinear monomials (based on the well-know McCormick inequalities)
- S. and Lee (2017) consider these alternatives and compute their volumes
- Here we focus on the convex hull i.e. \( P_H \)
Convex hull of the graph $f = x_1x_2x_3$ over a box

- The extreme points of $\mathcal{P}_H$ are the 8 points that correspond to the $2^3 = 8$ choices of each $x$-variable at its upper or lower bound:

\[
\begin{align*}
v^1 &:= \begin{bmatrix} b_1a_2a_3 \\ b_1 \\ a_2 \\ a_3 \end{bmatrix} \\
v^2 &:= \begin{bmatrix} a_1a_2a_3 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} \\
v^3 &:= \begin{bmatrix} a_1a_2b_3 \\ a_1 \\ a_2 \\ b_3 \end{bmatrix} \\
v^4 &:= \begin{bmatrix} a_1b_2a_3 \\ a_1 \\ b_2 \\ a_3 \end{bmatrix} \\
v^5 &:= \begin{bmatrix} a_1b_2b_3 \\ a_1 \\ b_2 \\ b_3 \end{bmatrix} \\
v^6 &:= \begin{bmatrix} b_1b_2b_3 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} \\
v^7 &:= \begin{bmatrix} b_1b_2a_3 \\ b_1 \\ b_2 \\ a_3 \end{bmatrix} \\
v^8 &:= \begin{bmatrix} b_1a_2b_3 \\ b_1 \\ a_2 \\ b_3 \end{bmatrix}
\end{align*}
\]
Convex hull of the graph $f = x_1 x_2 x_3$ over a box

• The extreme points of $P_H$ are the 8 points that correspond to the $2^3 = 8$ choices of each $x$-variable at its upper or lower bound:

\[
\begin{align*}
v^1 &:= \begin{bmatrix} b_1 a_2 a_3 \\ b_1 \\ a_2 \\ a_3 \end{bmatrix} & v^2 &:= \begin{bmatrix} a_1 a_2 a_3 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} & v^3 &:= \begin{bmatrix} a_1 a_2 b_3 \\ a_1 \\ a_2 \\ b_3 \end{bmatrix} & v^4 &:= \begin{bmatrix} a_1 b_2 a_3 \\ a_1 \\ b_2 \\ a_3 \end{bmatrix} \\
v^5 &:= \begin{bmatrix} a_1 b_2 b_3 \\ a_1 \\ b_2 \\ b_3 \end{bmatrix} & v^6 &:= \begin{bmatrix} b_1 b_2 b_3 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} & v^7 &:= \begin{bmatrix} b_1 b_2 a_3 \\ b_1 \\ b_2 \\ a_3 \end{bmatrix} & v^8 &:= \begin{bmatrix} b_1 a_2 b_3 \\ b_1 \\ a_2 \\ b_3 \end{bmatrix}
\end{align*}
\]

• Meyer and Floudas (2004) completely characterized the facets of $P_H$

• They did this for our special case (non-negative) and also in general
**Facet Description**

\( \mathcal{P}_H \) (Meyer and Floudas, 2004)

\[
\begin{align*}
 f - a_2 a_3 x_1 - a_1 a_3 x_2 - a_1 a_2 x_3 + 2a_1 a_2 a_3 & \geq 0 \\
 f - b_2 b_3 x_1 - b_1 b_3 x_2 - b_1 b_2 x_3 + 2b_1 b_2 b_3 & \geq 0 \\
 f - a_2 b_3 x_1 - a_1 b_3 x_2 - b_1 a_2 x_3 + a_1 a_2 b_3 + b_1 a_2 b_3 & \geq 0 \\
 f - b_2 a_3 x_1 - b_1 a_3 x_2 - a_1 b_2 x_3 + b_1 b_2 a_3 + a_1 b_2 a_3 & \geq 0 \\
 f - \frac{\eta_1}{b_1 - a_1} x_1 - b_1 a_3 x_2 - b_1 a_2 x_3 + \left( \frac{\eta_1 a_1}{b_1 - a_1} + b_1 b_2 a_3 + b_1 a_2 b_3 - a_1 b_2 b_3 \right) & \geq 0 \\
 f - \frac{\eta_2}{a_1 - b_1} x_1 - a_1 b_3 x_2 - a_1 b_2 x_3 + \left( \frac{\eta_2 b_1}{a_1 - b_1} + a_1 a_2 b_3 + a_1 b_2 a_3 - b_1 a_2 a_3 \right) & \geq 0 \\
 -f + a_2 a_3 x_1 + b_1 a_3 x_2 + b_1 b_2 x_3 - b_1 b_2 a_3 - b_1 a_2 a_3 & \geq 0 \\
 -f + b_2 a_3 x_1 + a_1 a_3 x_2 + b_1 b_2 x_3 - b_1 b_2 a_3 - a_1 b_2 a_3 & \geq 0 \\
 -f + a_2 a_3 x_1 + b_1 b_3 x_2 + b_1 a_2 x_3 - b_1 a_2 b_3 - b_1 a_2 a_3 & \geq 0 \\
 -f + b_2 b_3 x_1 + a_1 a_3 x_2 + a_1 b_2 x_3 - a_1 b_2 b_3 - a_1 b_2 a_3 & \geq 0 \\
 -f + a_2 b_3 x_1 + b_1 b_3 x_2 + a_1 a_2 x_3 - a_1 a_2 b_3 - a_1 a_2 b_3 & \geq 0 \\
 -f + b_2 b_3 x_1 + a_1 b_3 x_2 + a_1 a_2 x_3 - a_1 b_2 b_3 - a_1 a_2 b_3 & \geq 0 \\
 a_i \leq x_i \leq b_i, \quad i = 1..3
\end{align*}
\]

where \( \eta_1 = b_1 b_2 a_3 - a_1 b_2 b_3 - b_1 a_2 a_3 + b_1 a_2 b_3 \) and \( \eta_2 = a_1 a_2 b_3 - b_1 a_2 a_3 - a_1 b_2 b_3 + a_1 b_2 a_3 \).
Volume of $P_H$

Theorem (S. and Lee, 2017)

Under $\Omega$, the volume of $P_H$ is given by:

$$\frac{1}{24}(b_1 - a_1)(b_2 - a_2)(b_3 - a_3) \times$$

$$\left( b_1(5b_2b_3 - a_2b_3 - b_2a_3 - 3a_2a_3) + a_1(5a_2a_3 - b_2a_3 - a_2b_3 - 3b_2b_3) \right).$$
Our contribution

• We provide an alternative way to obtain the formula for the convex hull volume

• Observe a special structure in the convex hull polytope

• Allows us to use theory from so-called Mixed Volumes
Mixed Volumes

\( \mathcal{K}^n \) is the set of all nonempty compact convex sets in \( \mathbb{R}^n \)

### Theorem

There is a unique, nonnegative function, \( V : (\mathcal{K}^n)^n \to \mathbb{R} \), the **mixed volume**, which is invariant under permutation of its arguments, such that, for every positive integer \( m > 0 \), one has

\[
\text{Vol}(t_1 K_1 + t_2 K_2 + \cdots + t_m K_m) = \sum_{i_1, \ldots, i_n = 1}^{m} t_{i_1} \cdots t_{i_n} V(K_{i_1}, \ldots, K_{i_n}),
\]

for arbitrary \( K_1, \ldots, K_m \in \mathcal{K}^n \) and \( t_1, t_2, \ldots, t_n \in \mathbb{R}_+ \).
The mixed volume function satisfies the following properties:

(i) \( \text{Vol}(K_1) = V(K_1, \ldots, K_1) \).

(ii) \( V(t'K_1' + t''K_1'', K_2, \ldots, K_n) = t'V(K_1', K_2, \ldots, K_n) + t''V(K_1'', K_2, \ldots, K_n) \),

for \( K_1, \ldots K_n \in \mathcal{K}^n \) and \( t', t'' \in \mathbb{R}_+ \).
The mixed volume function satisfies the following properties:

(i) \( \text{Vol}(K_1) = V(K_1, \ldots, K_1) \).

(ii) \( V(t'K_1' + t''K_1'', K_2, \ldots, K_n) = t'V(K_1', K_2, \ldots, K_n) \)
     \[ + t''V(K_1'', K_2, \ldots, K_n), \]
for \( K_1, \ldots K_n \in \mathcal{K}^n \) and \( t', t'' \in \mathbb{R}_+ \).

There is a great deal of rich theory, see Schneider (2014).
Visualizing four dimensions

A 2d projection of a standard 4d cube that preserves the usual 2d projection of a 3d cube
Visualizing the convex hull

- 2d representation of the extreme points of $\mathcal{P}_H \in \mathbb{R}^4$
- $2^3 = 8$ extreme points
- $\mathbb{R}^4$ since points have form: $(f = x_1x_2x_3, x_1, x_2, x_3)$
- $f$ variable represented by ‘4th dimension’, inner cube to outer cube
- $v^2 = [a_1a_2a_3, a_1, a_2, a_3]$
  $v^6 = [b_1b_2b_3, b_1, b_2, b_3]$
- The original proof constructed a triangulation of the polytope
Another way of viewing the polytope

- Consider the $x_3$ component (could be $x_1$ or $x_2$)
- Observe that four of the points lie in the $x_3 = a_3$ hyperplane and form a 3d simplex, $S$
- Furthermore the remaining four points lie in the $x_3 = b_3$ hyperplane and form a simplex, $T$
- Consider calculating the volume of $\mathcal{P}_H = \text{conv}(S \cup T)$ via an integral as $x_3$ varies from $a_3$ to $b_3$
Alternative volume computation

\[ \text{Vol}(\mathcal{P}_H) = \text{Vol}(\text{conv}(S \cup T)) \]

\[ = \int_{a_3}^{b_3} \text{Vol} \left( \frac{b_3 - t}{b_3 - a_3} S + \frac{t - a_3}{b_3 - a_3} T \right) dt \]

\[ = (b_3 - a_3)^{-3} \int_{a_3}^{b_3} \text{Vol}((b_3 - t)S + (t - a_3)T) dt \]

\[ = (b_3 - a_3)^{-3} \int_{a_3}^{b_3} (b_3 - t)^3 \text{Vol}(S) + 3(b_3 - t)^2(t - a_3)V(S, S, T) \]
\[ + 3(b_3 - t)(t - a_3)^2V(S, T, T) + (t - a_3)^3 \text{Vol}(T) \] dt,

Where \( V(S, S, T) \) and \( V(S, T, T) \) are the mixed volumes.
Calculating the mixed volumes

• $S$ and $T$ are simplicies, therefore we can compute their volume via a simple determinant calculation
• All that remains is to calculate $V(S, S, T)$ and $V(S, T, T)$
• To do this we need a couple of definitions

Definition (Support function)
For $K \subseteq \mathbb{R}^n$, the support function, $h_K : \mathbb{R}^n \to \mathbb{R}$, is defined by

Definition
For a full dimensional polytope, $P \subseteq \mathbb{R}^n$, $U(P) =$ the set of all outer facet normals, $u$, such that the length of $u$ is the $(n-1)$-dimensional volume of the respective facet.
Calculating the mixed volumes

- \( S \) and \( T \) are simplicies, therefore we can compute their volume via a simple determinant calculation
- All that remains is to calculate \( V(S, S, T) \) and \( V(S, T, T) \)
- To do this we need a couple of definitions

**Definition (Support function)**

For \( K \in \mathcal{K}^n \), the support function, \( h_K : \mathbb{R}^n \rightarrow \mathbb{R} \), is defined by

\[
h_K(u) = \sup_{x \in K} x^T u.
\]
Calculating the mixed volumes

- $S$ and $T$ are simplicies, therefore we can compute their volume via a simple determinant calculation
- All that remains is to calculate $V(S, S, T)$ and $V(S, T, T)$
- To do this we need a couple of definitions

**Definition (Support function)**

For $K \in \mathcal{K}^n$, the support function, $h_K : \mathbb{R}^n \to \mathbb{R}$, is defined by

$$h_K(u) = \sup_{x \in K} x^T u.$$

**Definition**

For a full dimensional polytope, $P \subseteq \mathbb{R}^n$,

$$\mathcal{U}(P) = \text{the set of all outer facet normals, } u, \text{ such that the length of } u \text{ is the } (n - 1)\text{-dimensional volume of the respective facet.}$$
Calculating the mixed volumes

Now, to calculate $V(S, S, T)$ and $V(S, T, T)$, we make use of the following fact:

For $P \subseteq \mathbb{R}^n$, a full-dimensional polytope, and $K \in \mathcal{K}^n$,

$$V(P, P, \ldots, P, K) = \frac{1}{n} \sum_{u \in \mathcal{U}(P)} h_K(u).$$
Calculating the mixed volumes

Now, to calculate \( V(S, S, T) \) and \( V(S, T, T) \), we make use of the following fact:

For \( P \subseteq \mathbb{R}^n \), a full-dimensional polytope, and \( K \in \mathcal{K}^n \),

\[
V(P, P, \ldots, P, K) = \frac{1}{n} \sum_{u \in \mathcal{U}(P)} h_K(u).
\]

• Thus, we need: \( \mathcal{U}(S), \mathcal{U}(T), h_S(\cdot) \) and \( h_T(\cdot) \)
Calculating the mixed volumes

Now, to calculate $V(S, S, T)$ and $V(S, T, T)$, we make use of the following fact:

For $P \subseteq \mathbb{R}^n$, a full-dimensional polytope, and $K \in \mathcal{K}^n$,

$$V(P, P, \ldots, P, K) = \frac{1}{n} \sum_{u \in \mathcal{U}(P)} h_K(u).$$

• Thus, we need: $\mathcal{U}(S), \mathcal{U}(T), h_S(\cdot)$ and $h_T(\cdot)$

• Given that $S, T$ are tetrahedra we are able to compute these
  • Four facet normals to compute
  • Four extreme points to check
Calculating the mixed volumes

Now, to calculate $V(S, S, T)$ and $V(S, T, T)$, we make use of the following fact:

For $P \subseteq \mathbb{R}^n$, a full-dimensional polytope, and $K \in \mathcal{K}^n$,

$$V(P, P, \ldots, P, K) = \frac{1}{n} \sum_{u \in \mathcal{U}(P)} h_K(u).$$

• Thus, we need: $\mathcal{U}(S), \mathcal{U}(T), h_S(\cdot)$ and $h_T(\cdot)$

• Given that $S, T$ are tetrahedra we are able to compute these
  • Four facet normals to compute
  • Four extreme points to check

• We can therefore evaluate the integral and in doing so we obtain the same formula as the triangulation method
Conclusions and future work

• Because of some nice properties of the convex hull polytope, we are able to use mixed volume theory as an alternative method for computing the volume, this gives a more compact proof.

• However, the method does not naturally extend to the other volume proofs, (unlike the original triangulation method) and here is where the majority of the technical difficulties lay.

• Interestingly, in both proof methods for the convex hull, the technical difficulties were similar – establishing the sign of polynomial expressions in the parameters.
Conclusions and future work

• It seems likely that using this method we can extend to the case of negative bounds (current work)

• This method gives a more natural way to extend to $n = 4$ than we had before, however, it still seems like there will be difficulties doing this in practice

• Another tool in the volume toolbox!
Thank you for listening!

References:


