Mix-integer semidefinite programming with one application

CHEN Qifan

CUHK

August 16, 2019

Overview

- Outer Approximation method
  - Outer Approximation Approach for mix integer nonlinear programming
  - second order based Outer Approximation(non-polyhedral)
- Application: Joint Multicast Beamforming and Antenna Selection
- The problem on duality we met (left for Fang Cunling next week)

Overview

- Outer Approximation method
  - Outer Approximation Approach for mix integer nonlinear programming
  - second order based Outer Approximation(non-polyhedral)
- Application: Joint Multicast Beamforming and Antenna Selection
- The problem on duality we met (left for Fang Cunling next week)

Overview

- Outer Approximation method
  - Outer Approximation Approach for mix integer nonlinear programming
  - second order based Outer Approximation(non-polyhedral)
- Application: Joint Multicast Beamforming and Antenna Selection
- The problem on duality we met (left for Fang Cunling next week)

Overview

- Outer Approximation method
  - Outer Approximation Approach for mix integer nonlinear programming
  - second order based Outer Approximation(non-polyhedral)
- Application: Joint Multicast Beamforming and Antenna Selection
- The problem on duality we met (left for Fang Cunling next week)

Overview

- Outer Approximation method
  - Outer Approximation Approach for mix integer nonlinear programming
  - second order based Outer Approximation(non-polyhedral)
- Application: Joint Multicast Beamforming and Antenna Selection
- The problem on duality we met (left for Fang Cunling next week)

Mix-integer

Mixed-integer programming (MIP) refers to an optimization problem involving both continuous and discrete decision variables

NP-hard and nonconvex
even more challenging when positive semidefinite (PSD) conic constraints are present

Selection procedure,LASSO,source allocation,....

Goal—tractable

theoretically tractable:polynomially solvable.

practically tractable:not necessarily P.
Worst case NP
"In practice" acceptable.
Varies with each individual.

Anthony So told me:"Sometimes tractable means a good approximation quality; while sometimes one day for calculation is fine."

"Tractable" here:solvable by Gurobi(commercial solver):
Reduction to following forms:
Mixed-integer quadratically-constrained programming ,Mixed-integer quadratic programming ...

Example: mixed-integer nonlinear programming(MINLP)

The key idea is to develop outer approximations and inner approximations of the feasible region, thus reducing the MINLP to a finite sequence of "tractable" forms(Duran,1986).

$$\begin{array}{ll} \min_{x,z} & c^T z + f(x) \\ \texttt{s.t.} & g(x) + Bz \leq 0 \\ & x \in X \subset \mathbb{R}^n \\ & z \in U \subset \mathbb{R}^m_+ \end{array}$$

`(Slater's constraint qualification assumed to hold)`

where the nonlinear functions $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^p$ are assumed to be continuously differentiable and convex on the n -dimensional compact polyhedral convex set $\boldsymbol{X} = \{\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{A}_1\boldsymbol{x} \leq \boldsymbol{a}_1\}$. Besides, $\boldsymbol{U} = \{\boldsymbol{z} : \boldsymbol{z} \in \{0,1\}^m, \boldsymbol{A}_2\boldsymbol{z} \leq \boldsymbol{a}_2\}$ is a finite discrete set.

Outer Approximation Algorithm

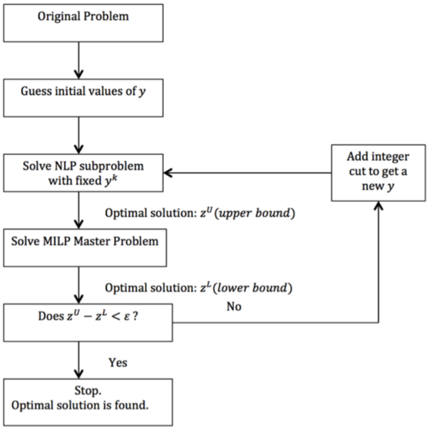

Figure 1: outer approximation

MINLP algorithm

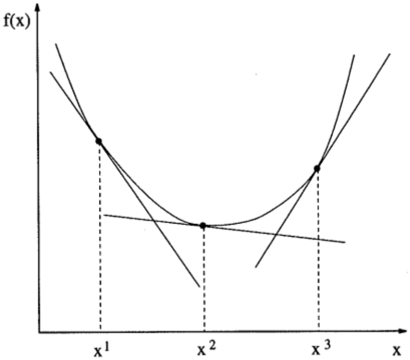

Figure 2: outer approximation

MINLP algorithm

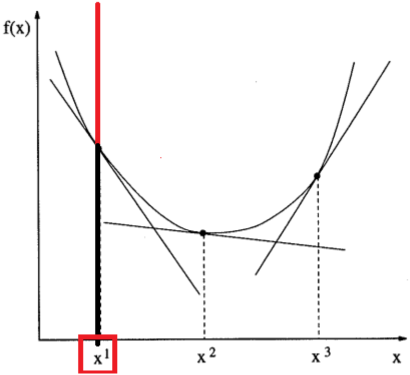

Figure 3: inner approximation

outer and inner approximation

Outer approximation

$$
\begin{aligned}
\Phi_L = \min_{x,z,\mu} \quad & c^T z + \mu \\
\text{s.t.} \quad & f\left(x^i\right) + \nabla f\left(x^i\right)^T \left(x - x^i\right) - \mu \leq 0, \forall x^i \in T \\
& g\left(x^i\right) + \nabla g\left(x^i\right)^T \left(x - x^i\right) + Bz \leq 0, \forall x^i \in T \\
& x \in X, z \in U, \mu \in [f_L, f_U]
\end{aligned}
$$

inner approximation(integer variables fixed at z)

$$
\begin{aligned}
\Phi_U = \min_{x} \quad & c^T \hat{z} + f(x) \\
\text{s.t.} \quad & g(x) + B\hat{z} \leq 0 \\
& x \in X
\end{aligned}
$$

Algorithm 2.1: Outer Approximation (OA) for MINLP

Initialize: $\phi_U \leftarrow \infty$, $\phi_L \leftarrow -\infty$, $\boldsymbol{T} \leftarrow \emptyset$; Tol $= \epsilon$.
while $\phi_U - \phi_L >$ Tol **do**
  if OA($\boldsymbol{T}$) is infeasible **then**
    (2.4) is also infeasible, terminate.
  end if
  Solve OA($\boldsymbol{T}$) for an optimal integer solution $\hat{\boldsymbol{z}}$ with optimal value Φ_L. Update lower bound: $\phi_L \leftarrow \Phi_L$.
  Solve IA($\hat{\boldsymbol{z}}$) with integer variables fixed at $\hat{\boldsymbol{z}}$.
  Obtain an optimal continuous solution $\hat{\boldsymbol{x}}$ with its optimal value Φ_U.
  if $\Phi_U < \phi_U$ **then**
    Update upper bound: $\phi_U \leftarrow \Phi_U$; Opt $\leftarrow (\hat{\boldsymbol{z}}, \hat{\boldsymbol{x}})$.
  end if
  Update $\boldsymbol{T} \leftarrow \boldsymbol{T} \cup \{\hat{\boldsymbol{x}}\}$.
end while

Mechanism

Outer approximation provides lower bounds and feeds integer assignments to the inner approximation.

The updating rule of T eliminates the current integer point, if not optimal.

Inner approximation provides lower bound,
update integer variables in T

Integer variabls are finite
Guaranteed to converge

Proof on convergence and optimality: section 7 in Duran,1986

Algorithm 2.1: Outer Approximation (OA) for MINLP

Initialize: $\phi_U \leftarrow \infty$, $\phi_L \leftarrow -\infty$, $\boldsymbol{T} \leftarrow \emptyset$; Tol $= \epsilon$.
while $\phi_U - \phi_L >$ Tol **do**
  if OA($\boldsymbol{T}$) is infeasible **then**
    (2.4) is also infeasible, terminate.
  end if squeeze
  Solve OA($\boldsymbol{T}$) for an optimal integer solution $\hat{z}$ with optimal value Φ_L.
  Update lower bound: $\phi_L \leftarrow \Phi_L$.
  Solve IA($\hat{z}$) with integer variables fixed at $\hat{z}$.
  Obtain an optimal continuous solution $\hat{x}$ with its optimal value Φ_U.
  if $\Phi_U < \phi_U$ **then**
    Update upper bound: $\phi_U \leftarrow \Phi_U$; Opt $\leftarrow (\hat{z}, \hat{x})$.
  end if
  Update $\boxed{\boldsymbol{T} \leftarrow \boldsymbol{T} \cup \{\hat{x}\}}$. effective update
end while

Objective Function

$$\begin{array}{ll} \min_{X,z} & C \bullet X \\ \text{s.t.} & \mathcal{A}(X,z) = b \\ & L \le z \le U, z \in \mathbb{Z}^M \\ & X \succeq 0 \end{array}$$

z vector denotes bounded discrete variables
X denotes the continuous variables restricted in a PSD cone
A represents the constraints as below
$D_i \bullet X = (b - A_z z)_i$, for $i \in [m]$

Build bridges from MISDP to MISOCP

Second-order conic constraints serve as a tighter approximation
Solvers such as Gurobi are mature for solving MISOCP with considerably low complexity

Recall the semidefinite problem

$$\begin{array}{ll} \min_{X,z} & C \bullet X \\ \text{s.t.} & D_i \bullet X = (b - A_z z)_i\,, \text{ for } i \in [m] \\ & L \leq z \leq U, z \in \mathbb{Z}^M \\ & X \succeq 0 \end{array}$$

Second-order Conic Relaxation

$\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^H \in \mathbb{H}_+^n$
which is equivalent to $Y \in \mathbb{H}_+^n, Y \bullet \left(X - xx^H\right) \geq 0$
or $\boldsymbol{Y} \bullet \boldsymbol{X} - \boldsymbol{x}^H \boldsymbol{Y} \boldsymbol{x} \geq 0$

Second-order Conic Relaxation

For any subset $\mathcal{T} \subset \mathbb{H}_+^n$, if we define

$S_{\mathcal{T}} = s\left\{ \begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^H & 1 \end{bmatrix} \in \mathbb{H}^{n+1} : \boldsymbol{Y} \bullet \boldsymbol{X} - \boldsymbol{x}^H \boldsymbol{Y} \boldsymbol{x} \geq 0, \forall \boldsymbol{Y} \in \mathcal{T} \right\}$ then

$\mathbb{H}_+^{n+1} \subset S_T$

Therefore, S_T forms a non-polyhedral outer description for $\mathbb{H}_+^{n+1}$

SOC-based Outer Approximation Algorithm (SOC-OA)

Initialize: $\phi_U \leftarrow \infty$, $\phi_L \leftarrow -\infty$, Opt$\leftarrow \emptyset$, $\mathcal{T}_2 \leftarrow \emptyset$, Tol $= \epsilon$.
while $\phi_U - \phi_L >$ Tol **do**
  Solve MISOC($\mathcal{T}$)
  if MISOC($\mathcal{T}$) is infeasible **then**
    MISDP is also infeasible, terminate.
  end if
  Solve MISOC($\mathcal{T}$) for an optimal solution $(\hat{\boldsymbol{z}}, \bar{\boldsymbol{X}}_{\hat{z}})$ with optimal value $\phi_{\mathcal{T}}$
  Update lower bound: $\phi_L \leftarrow \phi_{\mathcal{T}}$.
  Solve CP($\hat{\boldsymbol{z}}$).
  if CP($\hat{\boldsymbol{z}}$) is feasible **then**
    Obtain an optimal primal-dual solution pair $((\hat{\boldsymbol{z}}, \bar{\boldsymbol{X}}_{\hat{z}}), (\boldsymbol{\lambda}_{\hat{z}}, \bar{\boldsymbol{Y}}_{\hat{z}}))$ with it optimal value $v_{\hat{z}}$.
    Construct $\boldsymbol{Y}$ according to Lemma 2.1.
    if $v_{\hat{z}} < \phi_U$ **then**
      Update upper bound: $\phi_U \leftarrow v_{\hat{z}}$; Opt $\leftarrow (\hat{\boldsymbol{z}}, \bar{\boldsymbol{X}}_{\hat{z}})$.
    end if
  else
    Construct $(\boldsymbol{\lambda}_{\hat{z}}, \bar{\boldsymbol{Y}}_{\hat{z}})$ and $\boldsymbol{Y}$ according to Lemma 2.2.
  end if
  Update $\mathcal{T} \leftarrow \mathcal{T} \cup \{\boldsymbol{Y}\}$.
end while

context

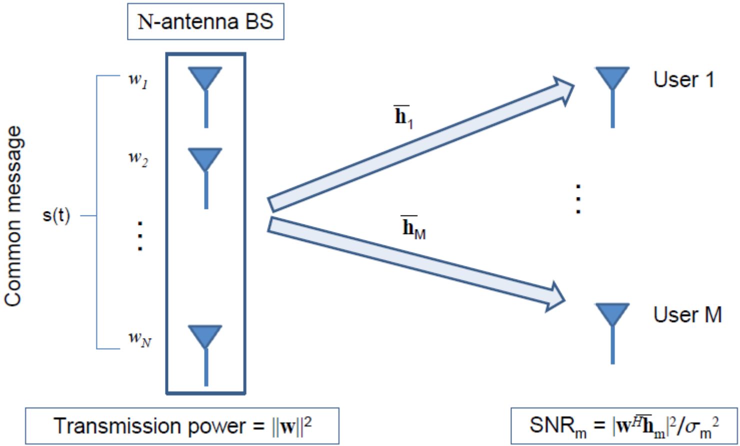

Figure 4: Minimal the SNR

Definition

A multicasting scenario: one N-antennae transmitter and M single-antenna users.

h_m : downlink channel vector between the transmitter and the m-th user.

$w \in C^N$: beaforming vector used to convey a zero mean and unit variance multicast signal s.

Only a subset of Q antennae can be active during transimission.

Received signal at m-th user: $y_m = h_m^H ws + z_m$ where z_m denotes the white Gaussian noise $(0, \sigma_m^2)$

^{H} means hermitian transpose

context

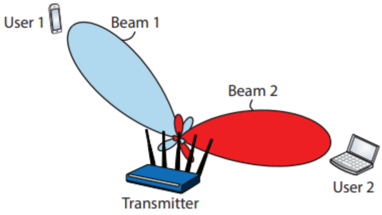

Figure 5: Visualization of transmit beamforming

Context

write $H_m = \frac{h_m h_m^H}{\sigma_m^2}$,signal-to-noise:$w^M H_m w$

Aim: select the best Q out of M antennas, and find the corresponding beamforming vector that maximizes the minimal SNR:

$$\max_{w \in \mathbb{C}^*} \min_{m \in [M]} w^M H_m w$$
$$\text{s.t. } \|w\|^2 \leq P, \|w\|_0 \leq Q$$

Equivalent Rewritten

joint admission control and multicast downlink beamforming (JABF) (Omar Mehanna,2013),(Xue-Ying Ni,2018)

Minimize transmit-power, subject to receive-SNR constraints per user
$\min_{\mathbf{w}\in\mathbb{C}^N} \|\mathbf{w}\|_2^2$ s.t.: $\left|\mathbf{w}^{\mathcal{H}}\mathbf{h}_m\right|^2 \geq 1, \quad m=1,\ldots,M$

Theorem 1 (Semidefinite Program relaxation(Luo,2010))

let $W = ww^H$ *and* $W \succeq 0$ *is equivalent to*
$W \cdot Y^\gamma - w^{II} Y w \geq 0, \forall Y \in S_+^n$ *by the self-duality of PSD cones*

Corollary 2

$\|\mathbf{w}\|_2^2 = \operatorname{tr}\left(\mathbf{w}\mathbf{w}^{\mathcal{H}}\right) = \operatorname{tr}(\mathbf{X})$

Corollary 3

write norm 0: $W_{ii} \leq \beta_i P, \forall i \in [N] \sum \beta_i \leq Q, \quad \beta \in \{0,1\}^N$

Problem modeling

incorporate binary variable for selection

$$
\begin{array}{rl}
v^* = & \min_{w\in\mathbb{C}^n,\gamma,\beta} \ -\gamma \\
\text{s.t.} & w^H H_m w \geq \gamma, \forall m \in [M] \\
& \|w\|^2 \leq P \\
& w_i^2 \leq \beta_i P, \forall i \in [N] \\
& \sum \beta_i \leq Q, \beta \in \{0,1\}^N
\end{array}
$$

Standard semidefinite relaxation to beamformer variable w(outer approximation)

$$
\begin{array}{rl}
v^*_{\text{OA}}(\mathcal{T}) = & \min_{W,w,\gamma,\beta} -\gamma \\
\text{s.t.} & H_m \cdot W \geq \gamma, \forall m \in [M] \\
& \text{tr}(W) \leq P \\
& W \cdot U - w^H U w \geq 0, \forall U \in T \\
& W_{ii} \leq \beta_i P, \forall i \in [N] \\
& \sigma \beta_i \leq Q, \beta \in \{0,1\}^N
\end{array}
$$

inner approximation and its duality form

fixing the binary variables to β_T:inner approxiamtion

$$
\begin{aligned}
& V_{IA}(\beta_T) = min_{W,\gamma} - \gamma \\
\text{s.t.} \quad & II_m \cdot W \geq \gamma, \forall m \in [M] \\
& \operatorname{tr}(W) \leq P \\
& W \succeq 0 \\
& \mathrm{W}'_{ii} \leq (\beta_T^*)_i \, P, \forall i \in [N] \ \gamma \geq 0
\end{aligned}
$$

Its dual form (standard lagrange dual)

$$
\begin{aligned}
& v^*_{\mathrm{OA}}(\mathcal{T}) = \min_{W,w,\gamma,\beta} -\gamma \\
\text{s.t.} \quad & H_m \cdot W \geq \gamma, \forall m \in [M] \\
& \operatorname{tr}(W) \leq P \\
& W \cdot U - w^H U w \geq 0, \forall U \in T \\
& W_{ii} \leq \beta_i P, \forall i \in [N] \\
& \sigma \beta_i \leq Q, \beta \in \{0,1\}^N
\end{aligned}
$$

if CP($\hat{z}$) is feasible **then**
  Obtain an optimal primal-dual solution pair $((\hat{z}, \tilde{\boldsymbol{X}}_{\hat{z}}), (\boldsymbol{\lambda}_{\hat{z}}, \tilde{\boldsymbol{Y}}_{\hat{z}}))$ with its optimal value $v_{\hat{z}}$.
  Construct $\boldsymbol{Y}$ according to Lemma 2.1.
  if $v_{\hat{z}} < \phi_U$ **then**
    Update upper bound: $\phi_U \leftarrow v_{\hat{z}}$; Opt $\leftarrow (\hat{z}, \tilde{\boldsymbol{X}}_{\hat{z}})$.

Necessary assumption to argue the finite-time convergence:
For any given integer solution $\hat{z}$ for MISOC(T) , if $CP(\hat{z})$ is feasible, then $DCP(\hat{z})$ is also feasible and there exists a primal-dual optimal solu- tion pair $\left(\left(\hat{z}, \tilde{\boldsymbol{X}}_{\hat{z}}\right), \left(\boldsymbol{\lambda}_{\hat{z}}, \tilde{\boldsymbol{Y}}_{\hat{z}}\right)\right)$; besides, strong duality holds at this optimal solution

Reference

M. A. Duran and I. E. Grossmann. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. Mathematical programming, 36(3):307–339, 1986.

S. X. Ni and A. M. So, "Mixed-Integer Semidefinite Relaxation of Joint Admission Control and Beamforming: An SOC-Based Outer Approximation Approach with Provable Guarantees," 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Kalamata, 2018, pp. 1-5.

Z. Luo, W. Ma, A. M. So, Y. Ye and S. Zhang, "Semidefinite Relaxation of Quadratic Optimization Problems," in IEEE Signal Processing Magazine, vol. 27, no. 3, pp. 20-34, May 2010.

Thanks!

Reference I