Optimization problems with bilinear matrix inequalities: the problem and algorithms

Michal Kočvara

Institute of Information Theory and Automation

Academy of Sciences of the Czech Republic

and

Faculty of Electrical Engineering

Czech Technical University

kocvara@utia.cas.cz

http://www.utia.cas.cz/kocvara

SDP notations

$\mathbb{S}^n$ . . . symmetric matrices of order $n \times n$
$A \succeq 0$ . . . A positive semidefinite
$A \succeq B$ . . . $A - B \geq 0$
$\langle A, B \rangle :=$ Trace(AB) . . . inner product on $\mathbb{S}^n$
$\mathcal{A}[\mathbb{R}^m \to \mathbb{S}^n]$. . . linear matrix operator defined by

$$\mathcal{A}(y) := \sum_{i=1}^{m} y_i A_i \quad \text{with } A_i \in \mathbb{S}_n$$

$\mathcal{A}^*[S^n \to \mathbb{R}^m]$ . . . adjoint operator to $\mathcal{A}$ defined by

$$\mathcal{A}^*(X) := [\langle A_1, X \rangle, \ldots, \langle A_1, X \rangle]^T$$

and satisfying

$$\langle \mathcal{A}^*(X), y \rangle = \langle \mathcal{A}(y), X \rangle \qquad \text{for all } y \in \mathbb{R}^m$$

BMI: what it is?

Optimization problem with Bilinear Matrix Inequality:

$$\min_{x\in\mathbb{R}^n} b^T x$$

s.t.

$$A_0 + \sum_{k=1}^{n} x_k A_k + \sum_{k=1}^{n}\sum_{\ell=1}^{n} x_k x_\ell K_{k\ell} \preccurlyeq 0 \qquad \text{(BMI)}$$

where $A_k, K_{k\ell} \in \mathbb{S}^m,\ b \in \mathbb{R}^n$.

Problem is nonlinear and nonconvex.

Search for

- global minimum
- local minimum
- feasible point

BMI feasibility problem

Find a feasible point of BMI

$$A_0 + \sum_{k=1}^{n} x_k A_k + \sum_{k=1}^{n} \sum_{\ell=1}^{n} x_k x_\ell K_{k\ell} \preccurlyeq 0$$

Formulated as an optimization problem (minimize the maximum eigenvalue)

$$\min_{x \in \mathbb{R}^n, \lambda \in \mathbb{R}} \quad \lambda$$

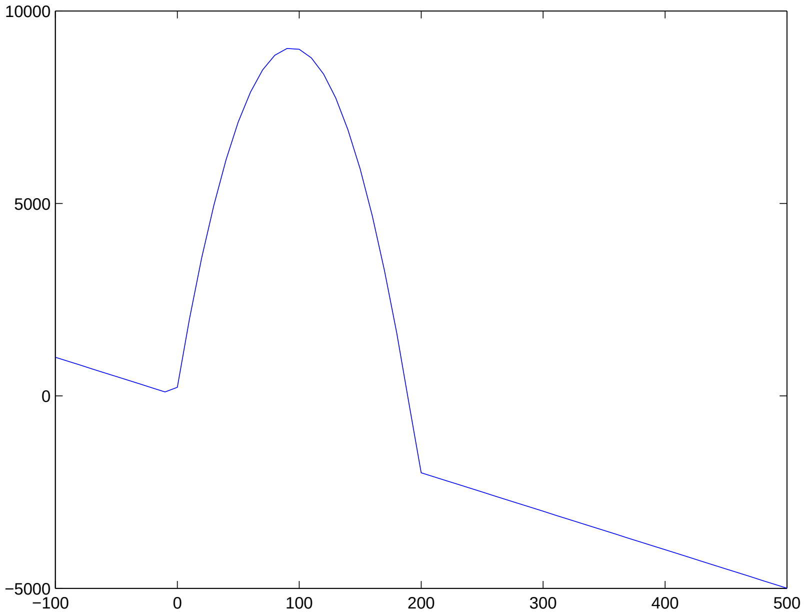
$$\text{s.t.} \quad A_0 + \sum_{k=1}^{n} x_k A_k + \sum_{k=1}^{n} \sum_{\ell=1}^{n} x_k x_\ell K_{k\ell} \preccurlyeq \lambda I_{m \times m}$$

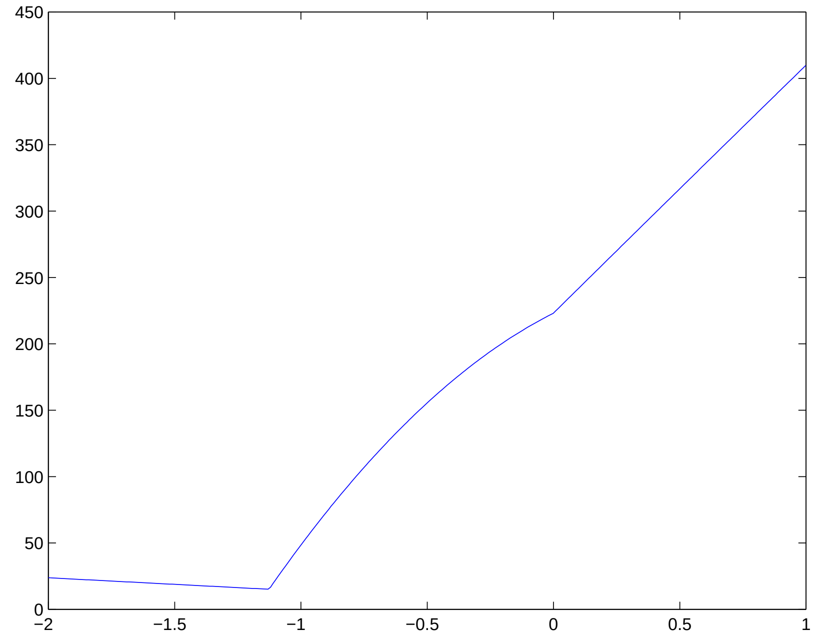
if $\lambda_{\text{opt}} < 0$, x_{opt} strictly feasible
if $\lambda_{\text{opt}} = 0$, x_{opt} marginally feasible, BMI has no interior point
if $\lambda_{\text{opt}} > 0$, BMI (may be) infeasible (when λ_{opt} local minimum)

Example: BMI is nonconvex

One-dimensional BMI feasibility problem.

Dependence of $\lambda_{\max}$ on x:





Consider BMI

$$\min_{x \in \mathbb{R}^n} b^T x$$
$$\text{s.t.} \quad A_0 + \sum_{k=1}^{n} x_k A_k + \sum_{k=1}^{n} \sum_{\ell=1}^{n} x_k x_\ell K_{k\ell} \preccurlyeq 0$$
$$\underline{x} \leq x \leq \overline{x}$$

Equivalent problem: define $w_{k\ell} = x_k x_\ell$:

$$\min_{x \in \mathbb{R}^n, W \in \mathbb{R}^{n \times n}} b^T x$$
$$\text{s.t.} \quad A_0 + \sum_{k=1}^{n} x_k A_k + \sum_{k=1}^{n} \sum_{\ell=1}^{n} w_{k\ell} K_{k\ell} \preccurlyeq 0$$
$$\underline{x} \leq x \leq \overline{x}$$
$$W = x x^T$$

Global methods, convex relaxations

Convex relaxations (to get lower bounds)

Goh et al.: replace $W = xx^T$ by $\underline{w}_{ij} \le w_{ij} \le \overline{w}_{ij}$, where

$$\underline{w}_{ij} = \min\{\underline{x}_i\underline{x}_j, \underline{x}_i\overline{x}_j, \overline{x}_i\underline{x}_j, \overline{x}_i\overline{x}_j\}$$
$$\overline{w}_{ij} = \max\{\underline{x}_i\underline{x}_j, \underline{x}_i\overline{x}_j, \overline{x}_i\underline{x}_j, \overline{x}_i\overline{x}_j\}$$

Fujioka & Hoshijima: tighter relaxation by

$$x_i\underline{x}_j + \underline{x}_i x_j - \underline{x}_i\underline{x}_j \le w_{ij} \le x_i\overline{x}_i + \underline{x}_i x_j - \underline{x}_i\overline{x}_j$$
$$x_i\overline{x}_j + \overline{x}_i x_j - \overline{x}_i\overline{x}_j \le w_{ij} \le x_i\underline{x}_j + \overline{x}_i x_j - \overline{x}_i\underline{x}_j$$

Theorem: For BMI of type $\sum_{\ell=1}^{n} y x_\ell K_\ell$ in variables $(x, y) \in \mathbb{R}^{n+1}$, the Fu-Ho relaxation cannot be improved.

Global methods, convex relaxations

Fukuda: still a tighter relaxation

Theorem: For BMI of type $\sum_{k=1}^{2} \sum_{\ell=1}^{n} y_k x_\ell K_{k\ell}$ in variables $(x, y) \in \mathbb{R}^{n+2}$, the Fukuda relaxation cannot be improved.

Number of inequalities to represent each convex polytope:

Goh	Fu-Ho	Fukuda
$2n^2 + 4n$	$4n^2$	$2n^2(n-1)^2 + 4n^2$

E.g., for $n = 10$, Goh $\longrightarrow 240$, Fu-Ho $\longrightarrow 400$, Fuk. $\longrightarrow 16600$

Global methods, approximations of local optima

Consider BMI in the form

$$\min_{x \in \mathbb{R}^n_x,\, y \in \mathbb{R}^n_y} b_x^T x + b_y^T y$$

$$\text{s.t.} \quad A_0 + \sum_{k=1}^{n_x} x_k (A_x)_k + \sum_{k=1}^{n_y} y_k (A_y)_k + \sum_{k=1}^{n_x} \sum_{\ell=1}^{n_y} x_k y_\ell K_{k\ell} \preccurlyeq 0$$

For a fixed x or y, we get a linear SDP. This suggest an Alternating SDP (or Gauss-Seidel) method:

Fix x, minimize w.r.t. y, get new y
Fix y, minimize w.r.t. x, get new x
etc.

Does not necessarily converge to a local optimum.

Local methods

- Interior point methods
- Sequential SDP methods
- (Generalized) Augmented Lagrangian methods

BMI is a special case of a general nonlinear SDP (NSDP) problem

$$\min_{x \in \mathbb{R}^n} b^T x$$

subject to

$$\mathcal{A}(x) \preccurlyeq 0$$

$b \in \mathbb{R}^n$ and $\mathcal{A} : \mathbb{R}^n \longrightarrow \mathbb{S}^m$ nonlinear, nonconvex

An interior-point—Conceptual Algorithm

Jarre '98:
Set $\mu_0 = 1$. For $k = 0, 1, 2, 3, \ldots$ do

1. Find an approximate local minimizer $x = x^{(k)}$ of

$$\varphi(x, \mu_k) := b^T x - \mu_k \log\det \mathcal{A}(x)$$

2. Find a predictor direction $\Delta x^{(k)}$ minimizing $b^T \Delta x_i$ s.t. $\Delta x^T H_k \Delta x \leq 1$.

   Here, H_k is the Hessian matrix of the linearized barrier function $\log\det \mathcal{A}_k^{lin}(x)$ with
   $\mathcal{A}_k^{lin}(x) = \mathcal{A}(x^{(k)}) + \nabla \mathcal{A}(x^{(k)})(x - x^{(k)})$.

3. Set $\hat{\mu}_k = \mu_k / 10$ and minimize $\varphi(x^{(k)} + \alpha \Delta x^{(k)},\ \hat{\mu}_k)$ with respect to α. Let α_k be the minimizer.

   Then, let μ_{k+1} be the minimizer of $\varphi(x^{(k)} + \alpha_k \Delta x^{(k)},\ \mu)$ with respect to μ.

Sequential SDP methods

Given (x_0, S_0), for $k = 0, 1, 2 \ldots$ solve

$$\min_{d \in \mathbb{R}^n} b^T d + \frac{1}{2} d^T M_k d$$

$$\text{s.t.} \quad \mathcal{A}_k(d) \preccurlyeq 0$$

where

$\mathcal{A}_k(d) := \mathcal{A}(x_k) + \sum_{i=1}^n d_i \frac{\partial \mathcal{A}(x_k)}{\partial x_i}$

$M_k \succeq 0$… "proper" approximation of the Hessian of the Lagrangian

$L(x_k, S_k) = b^T x_k + \langle \mathcal{A}(x_k), S_k \rangle$

Set $x_{k+1} = x_k + d_k$

Sequential SDP methods

The “quadratic SDP” solved by solving the KKT system

$$\begin{aligned} b + D\mathcal{A}(x_k)^* S_{k+1} + M_k d_k &= 0 \\ \mathcal{A}_k(d_k) &\preccurlyeq 0 \\ S_{k+1} &\succcurlyeq 0 \\ \langle \mathcal{A}_k(d_k), S_{k+1} \rangle &= 0 \end{aligned}$$

where $D\mathcal{A}(x) = \left(\frac{\partial \mathcal{A}(x)}{\partial x_1}, \ldots, \frac{\partial \mathcal{A}(x)}{\partial x_n} \right)^T$

Theory: global/local quadr. convergence (Correa-Ramirez)

Practice: ??? How to solve the above system? How to choose M_k???