Large-scale Semidefinite Programming in Computation of Many-Body Electronic Structures

Marlene L. Lund
PhD student, Chemical Engineering, NTNU
IWR School 2017

Overview

- Motivation
- Semidefinite programming
- Density matrix theory and semidefinite programming
- State-of-the-art solution methods in semidefinite programming
- New approaches to solve large-scale systems

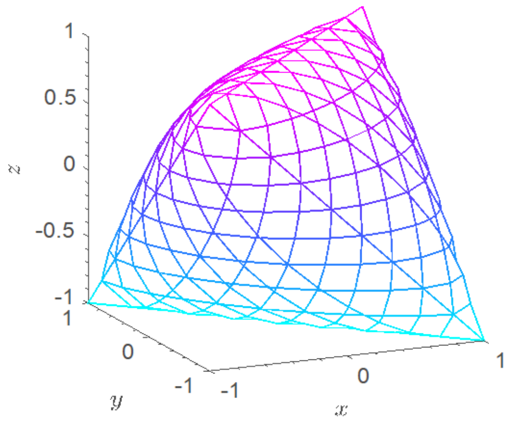
Motivation

- Through Hartree Fock, DFT and so on, an upper bound on the ground state energy is provided
- Want to improve the credibility of quantum chemical calculations
  - Perform computations instead of experiments
  - Reduce risk of investment based on such calculations to an acceptable level

Need to introduce a lower bound to estimate risk

Semidefinite Programming (SDP)

Spectrahedron



- Minimizing a linear objective subject to a set of linear constraints as well as constraining the solution to be positive semidefinite

Primal problem

$$\begin{aligned} \mathcal{P}: \quad & \min_{\mathbf{X}} \mathbf{C} \bullet \mathbf{X} \\ & \text{s.t} \quad \mathbf{A}_i \bullet \mathbf{X} = b_i, \quad \forall i \in \{1, \dots, p\}, \\ & \qquad \mathbf{X} \succeq \mathbf{0}. \end{aligned}$$

Dual problem

$$\begin{aligned} \mathcal{D}: \quad & \max_{\mathbf{S},\mathbf{y}} \mathbf{b}^{\mathsf{T}} \mathbf{y} \\ & \text{s.t} \quad \mathbf{S} + \sum_{i=1}^{p} y_i \mathbf{A}_i = \mathbf{C}, \\ & \qquad \mathbf{S} \succeq \mathbf{0}. \end{aligned}$$

SDP Optimality Conditions (KKT)

$$
\begin{aligned}
&\mathbf{A}_i \bullet \mathbf{X} = b_i, \quad \forall i \in \{1, \ldots, p\}, \\
&\mathbf{S} + \sum_{i=1}^{p} y_i \mathbf{A}_i = \mathbf{C}, \\
&\mathbf{X} \bullet \mathbf{S} = 0, \\
&\mathbf{X} \succeq \mathbf{0}, \\
&\mathbf{S} \succeq \mathbf{0}.
\end{aligned}
$$

- Sufficient and necessary if the problem is convex and a Slater condition holds
- Slater condition: there exists a strictly feasible point

SDP in Density Matrix Theory

- The problem of estimating the ground state energy can be written as an SDP
- Introducing reduced density matrices (RDM) gives a large-scale but practical problem

$$\mathcal{P}: \quad \min_{\mathbf{\Gamma}} \mathbf{H} \bullet \mathbf{\Gamma} \quad \text{s.t} \quad \mathbf{\Gamma} \bullet \mathbf{I} = 1, \quad \mathbf{\Gamma} \succeq \mathbf{0}.$$

$$\mathcal{P}_{RDM}: \quad \min_{\mathbf{\Gamma}^1, \mathbf{\Gamma}^2} \mathbf{H}_1 \bullet \mathbf{\Gamma}^1 + \mathbf{H}_2 \bullet \mathbf{\Gamma}^2 \quad \text{s.t} \quad \mathbf{\Gamma}^1 \bullet \mathbf{I} = N, \quad \mathbf{\Gamma}^2 \bullet \mathbf{I} = N(N-1), \quad \text{and } N\text{-representability.}$$

$\mathbf{H}$ - Hamiltonian
$\mathbf{\Gamma}$ - Density matrix
N - No. of electrons

$$\mathbf{\Gamma}^m = \frac{N!}{(N-m)!} \sum_{i_{m+1}, \ldots, i_N = 1}^{r} \mathbf{\Gamma}(i_1, \ldots, i_m, i_{m+1}, \ldots, i_N; i'_1, \ldots, i'_m, i'_{m+1}, \ldots, i'_N)$$

M. Fukuda, B. Braams, M. Nakata, M. Overton, J. Percus, M. Yamashita, and Z. Zhao, Mathematical Programming, vol. 109, pp. 553–580, 3 2007.

SDP in Density Matrix Theory

$$\mathcal{D}: \quad \max_{\mathbf{S},\mathbf{y}} \mathbf{b}^{\mathsf{T}}\mathbf{y}$$

Dual problem

$$\text{s.t} \quad \mathbf{S} + \sum_{i=1}^{p} y_i \mathbf{A}_i = \mathbf{C},$$

$$\mathbf{S} \succeq \mathbf{0}.$$

$$\mathbf{b} = \left[svec(\mathbf{H}_1)^{\mathsf{T}} \quad svec(\mathbf{H}_2)^{\mathsf{T}} \right]^{\mathsf{T}},$$

$$\mathbf{y} = \left[svec(\mathbf{\Gamma}_1)^{\mathsf{T}} \quad svec(\mathbf{\Gamma}_2)^{\mathsf{T}} \right]^{\mathsf{T}},$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{\Gamma}_1 & & & & & & \\ & \mathbf{I} - \mathbf{\Gamma}_1 & & & & & \\ & & \mathbf{\Gamma}_2 & & & & \\ & & & \mathbf{Q} & & & \\ & & & & \mathbf{G} & & \\ & & & & & \mathbf{T}_1 & \\ & & & & & & \mathbf{T}_2 \end{bmatrix}$$

$$svec(\mathbf{A}) = \left[a_{11} \quad \sqrt{2}a_{12} \quad a_{22} \quad \sqrt{2}a_{13} \quad \sqrt{2}a_{23} \quad \ldots \right], \quad \mathbf{A} \in \mathcal{S}^n$$

SDP in Density Matrix Theory

- Known issues:
  - Determine the N-representability conditions
    - Only necessary conditions are known → relaxation
  - Solving the problem: need a solver that can solve large-scale problems to sufficient accuracy

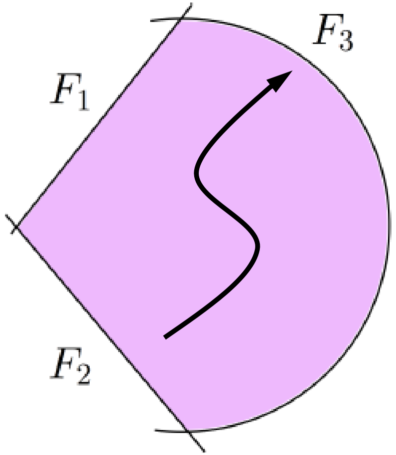
State-of-the-art SDP Solvers: Interior-Point Methods

- Primal-dual interior-point methods - procedure:
    1. Generate an equation system from the optimality conditions to determine a step direction

$$\mathbf{A}_i \bullet (\mathbf{X} + d\mathbf{X}) = b_i, \quad \forall i \in \{1, \dots, p\},$$
$$\mathbf{S} + d\mathbf{S} + \sum_{i=1}^{p} (y_i + dy_i)\mathbf{A}_i = \mathbf{C},$$
$$(\mathbf{X} + d\mathbf{X})(\mathbf{S} + d\mathbf{S}) = \mu \mathbf{I}.$$

    2. Compute step lengths such that X and S are still positive semidefinite
    3. Apply the step, and return to 1 if not optimal

$$\mathbf{X}_{k+1} = \mathbf{X}_k + \alpha_p d\mathbf{X} \succ \mathbf{0},$$
$$\mathbf{S}_{k+1} = \mathbf{S}_k + \alpha_d d\mathbf{S} \succ \mathbf{0},$$
$$\mathbf{y}_{k+1} = \mathbf{y}_k + \alpha_d d\mathbf{y}.$$



$$\mathcal{P}: \quad \min_{\mathbf{X}} \mathbf{C} \bullet \mathbf{X} \quad \text{s.t} \quad \mathbf{A}_i \bullet \mathbf{X} = b_i, \quad \forall i \in \{1, \dots, p\}, \quad \mathbf{X} \succeq \mathbf{0}.$$

Primal problem

M. Yamashita, K. Fujisawa, and M. Kojima, Optimization Methods and Software, vol. 18, no. 4, pp. 491–505, 2003.

State-of-the-art SDP Solvers: Interior-Point Methods

+ Avoids singular points by following a feasible trajectory

+ Postitive semidefiniteness is maintained with a step length

- A completely dense matrix (Schur complement) must be generated and factorized in each iteration
  – More than 10k constraints = trouble

$$\mathbf{A}_i \bullet (\mathbf{X} + d\mathbf{X}) = b_i, \;\; \forall i \in \{1, \dots, p\},$$

$$\mathbf{S} + d\mathbf{S} + \sum_{i=1}^{p} (y_i + dy_i)\mathbf{A}_i = \mathbf{C},$$

$$(\mathbf{X} + d\mathbf{X})(\mathbf{S} + d\mathbf{S}) = \mu \mathbf{I}.$$

$$\Rightarrow \quad \mathbf{B} = \begin{bmatrix} b_{11} & \dots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{p1} & \dots & b_{pp} \end{bmatrix}$$

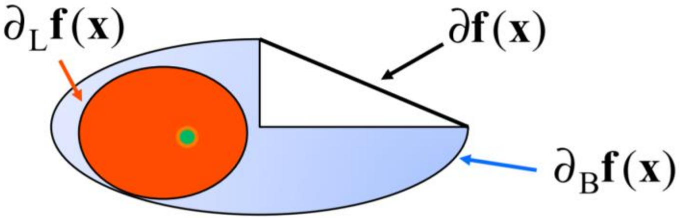
Completely dense

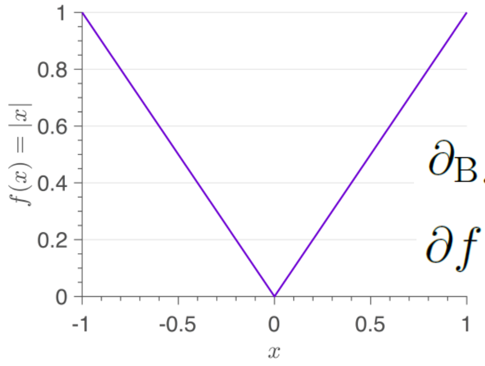
Nonsmooth Analysis in SDP Solving

- An alternative method to provide the semidefiniteness property of the variables

$\mathbf{X} \succeq \mathbf{0}$ ➡ $\lambda(\mathbf{X}) \geq \mathbf{0}$ ➡ $\min\{\lambda(\mathbf{X})\} \geq 0$ **Nonsmooth NLP**

- Lexicographic derivatives
  - Opens for automatic differentiation
  - For PC1 functions, an element of the B-subdifferential is guaranteed
  - Nonsmooth equation solvers have shown to be efficient





$$\partial_{\mathrm{B}} f(0) = \{-1, 1\}$$

$$\partial f(0) = [-1, 1]$$

Khan and Barton, Opt. Meth. & Soft. 30 (2015): 1185-1212, Khan and Barton, Journal Opt. Theory. Appl. 163 (2014): 355-386

Thank you!