Module 09
Relaxations of Nonconvex Problems

Ahmad F. Taha

CE 5999-02 Special Topics — Intro to Optimization

Email: ahmad.taha@vanderbilt.edu

Webpage: http://lab.vanderbilt.edu/taha

VANDERBILT UNIVERSITY

April 7, 2023

In this lecture

- Convex hull relaxations of mixed integer linear programs (MILPs)
- Semidefinite relaxation of nonconvex QCQPs and MILPs

See Z.-Q. Luo, W.-K. Ma, A. Man-Cho So, Y. Ye, and S. Zhang, *Semidefinite Relaxation of Quadratic Optimization Problems*, in IEEE Signal Processing Magazine, May 2010 (and references therein).

Hard problems

- Traditional view: Linear is easy; nonlinear is hard
- Modern view: Convex is easy; nonconvex is hard
- We can use convex optimization to approximately solve nonconvex problems

Nonconvex QCQP

Nonconvex QCQP with variable $x \in \mathbb{R}^n$

$$\begin{array}{ll} \text{minimize} & x^T P_0 x + q_0^T x \\ \text{subject to} & x^T P_i x + q_i^T x \trianglerighteq_i b_i, \quad i = 1, \ldots, m \end{array}$$

- Symbol $\trianglerighteq_i$ means $\geq$, $\leq$, or $=$, potentially different for each i
- $P_i \in \mathbb{S}^n$ $(i = 0, \ldots, n)$ but not necessarily positive semidefinite
- Quadratic constraint becomes linear if $P_i = 0$

Example: Boolean Least Squares

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|_2^2 \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n \end{array}$$

- Allow only $x_i = 1$ or $x_i = -1$
- Basic problem in digital communications

Variables $x \in \mathbb{R}^n$, $u_i \in \{0,1\}$ $(i = 1,\ldots,q)$, optimal value f^*

$$\begin{aligned} f^* = \text{minimize} \quad & c^T x + e^T u \\ \text{subject to} \quad & Fx + Gu \preceq h \\ & Ax + Bu = d \\ & x_i \in \mathbb{R} \ \ (i = 1,\ldots,n) \\ & u_i \in \{0,1\} \ \ (i = 1,\ldots,q) \end{aligned}$$

- If we ignore the integrality constraint $u_i \in \{0,1\}$, the resulting problem is an LP

Convex hull relaxation of MILP

- Only nonconvex constraint in MILP is $u_i \in \{0, 1\}$
- **Trick:** Relax to $0 \leq u_i \leq 1$ to obtain an LP

$$\begin{aligned} \check{f} = \text{minimize} \quad & c^T x + e^T u \\ \text{subject to} \quad & Fx + Gu \preceq h \\ & Ax + Bu = d \\ & x_i \in \mathbb{R} \; (i = 1, \ldots, n) \\ & 0 \leq u_i \leq 1 \; (i = 1, \ldots, q) \end{aligned}$$

- New problem is called a *relaxation* of the original MILP
- It has a *larger* feasible set than the original MILP
- We have for the optimal values of the original and relaxed problems: $\check{f} \leq f^*$

Advantages and disadvantages

- Solving the LP is very easy (and fast)
- Suppose that solution of the LP is $(\check{x}, \check{u})$
- **Lower bound** on MILP optimal value: $\check{f} = c^T \check{x} + e^T \check{u}$
- Main challenge: we don't know if $(\check{x}, \check{u})$ is feasible, that is, if $(\check{x}, \check{u})$ satisfies

$$\begin{aligned} & F\check{x} + G\check{u} \preceq h \\ & A\check{x} + B\check{u} = d \\ & \check{u}_i \in \{0, 1\} \; (i = 1, \ldots, q) \end{aligned}$$

Recovering an integer solution

- Straightforward idea: Slice $\check{u}_i$ to get binary u_i
- Specifically, set
$$\hat{u}_i = \begin{cases} 1, & \text{if } \check{u}_i \geq 1/2 \\ 0, & \text{otherwise} \end{cases}$$
- Still challenging to ensure $Fx + Gu \preceq h$, $Ax + Bu = d$ (may need to adjust $\check{x}$)
- For example, after having determined $\hat{u}$, solve the following LP with variable $x \in \mathbb{R}^n$
$$\begin{aligned} \hat{f} = \text{minimize} \quad & c^T x + e^T \hat{u} \\ \text{subject to} \quad & Fx + G\hat{u} \preceq h \\ & Ax + B\hat{u} = d \\ & x_i \in \mathbb{R} \ \ (i = 1, \ldots, n) \end{aligned}$$
- Optimal value $\hat{f}$ is an **upper bound** to the optimal value f^* of the MILP: $f^* \leq \hat{f}$

Assessment of the suboptimality

- Lower bound $\check{f} \leq f^*$
- Upper bound $\hat{f} \geq f^*$ with $\hat{x}, \hat{u}$ feasible for the MILP
- What we return to the user is $\hat{x}, \hat{u}$
- How good is that solution?
- **Approximation ratio:** $\hat{f}/f^*$
- Upper bound on approximation ratio: $\hat{f}/\check{f} \geq \hat{f}/f^*$
  - Approximation ratio is always ≥ 1, we want it to be as close to one as possible
- **Optimality gap:** $\hat{f} - \check{f}$
  - Optimality gap is always ≥ 0, we want it to be as close to zero as possible

Some general remarks

In general, all convex relaxation approaches involve two steps:

1. A method to obtain a convex relaxation of the nonconvex problem
2. A method to recover from the solution of the convex problem a solution that is *feasible* for the nonconvex problem

- Both steps are largely **problem-dependent**
- We will see the basic tricks in this lecture
- For example, one method for MILPs is to convert the constraint $u_i \in \{0, 1\}$ into $u_i \in [0, 1]$
- Another method is given next

MILP as (nonconvex) QCQP

Main trick:

$$u_i \in \{0,1\} \Longleftrightarrow u_i(u_i - 1) = 0$$

So MILP is equivalent to the following nonconvex QCQP with variables $x \in \mathbb{R}^n$, $u \in \mathbb{R}^q$

$$\begin{array}{ll} \text{minimize} & c^T x + e^T u \\ \text{subject to} & Fx + Gu \preceq h \\ & Ax + Bu = d \\ & u_i^2 - u_i = 0 \ \ (i = 1, \ldots, q) \end{array}$$

Next: Let's see relaxation methods for nonconvex QCQPs.

Homogeneous nonconvex QCQP with variable $x \in \mathbb{R}^n$:

$$\begin{aligned} \text{minimize} \quad & x^T P_0 x \\ \text{subject to} \quad & x^T P_i x \trianglerighteq_i 0 \end{aligned}$$

- Symbol $\trianglerighteq_i$ means $\geq$, $\leq$, or $=$, potentially different for each i
- $P_i \in \mathbb{S}^n$ but not necessarily positive semidefinite
- Homogeneous: we don't have the term $q_i^T x$
- Generalization to nonhomogeneous case is easy and will be shown later

Equivalent formulation of homogeneous nonconvex QCQP

- Semidefinite relaxation relies on the following **very important trick**:

$$x^T P_0 x = \mathrm{trace}(x^T P_0 x) = \mathrm{trace}(P_0 x x^T)$$

- QCQP constraint and objectives are **linear** in xx^T
- So let's define a matrix optimization variable $X = xx^T \in \mathbb{S}^n$
- Require X to be positive semidefinite and have rank 1
- Equivalent reformulation of the QCQP

$$\begin{array}{ll} \text{minimize} & \mathrm{trace}(P_0 X) \\ \text{subject to} & \mathrm{trace}(P_i X) \trianglerighteq_i 0 \\ & X \succeq 0 \\ & \mathrm{rank}(X) = 1 \end{array}$$

- Because of the rank constraint, the solution can always be written as $X = xx^T$, and x will be the solution of the QCQP

Implication of rank constraint

- If we have a matrix $X \succeq 0$ with $\text{rank} X = 1$, we can always write it is as $X = xx^T$ for some vector x
- Indeed, by eigendecomposition of X, we have that

$$X = \lambda_1 q_1 q_1^T$$

where $\lambda_1 > 0$ is the only nonzero eigenvalue, as we assumed $\text{rank} X = 1$
- Therefore, we can take

$$x = \sqrt{\lambda_1} q_1$$

Semidefinite relaxation (SDR) of homogeneous nonconvex QCQP

$$\begin{aligned} \text{minimize} \quad & \operatorname{trace}(P_0 X) \\ \text{subject to} \quad & \operatorname{trace}(P_i X) \trianglerighteq_i 0 \\ & X \succeq 0 \\ & \operatorname{rank}(X) = 1 \end{aligned}$$

- Where is the nonconvexity in the above formulation? $\operatorname{rank}(X) = 1$
- Semidefinite relaxation (SDR): Drop the rank constraint to get an SDP

$$\begin{aligned} \text{minimize} \quad & \operatorname{trace}(P_0 X) \\ \text{subject to} \quad & \operatorname{trace}(P_i X) \trianglerighteq_i 0 \\ & X \succeq 0 \end{aligned}$$

From the SDR solution to the QCQP solution

- Let $\check{X}$ be the solution of the SDR problem
- If it happens that $\text{rank}(\check{X}) = 1$, then we readily obtain the solution of the nonconvex QCQP. It is always possible (e.g., by eigendecomposition) to find $\check{x}$ so that
$$\check{X} = \check{x}\check{x}^T$$
- What if $\text{rank}(\check{X}) \neq 1$?
- We need to to use $\check{X}$ to obtain a vector $\hat{x}$, and we need to ensure that $\hat{x}$ is feasible for the QCQP constraints
$$\hat{x}^T P_i \hat{x} \trianglerighteq_i 0$$

Recovery of a feasible solution

There are two main ways to (hopefully) to recover a feasible solution:

1. *(Principal Eigenvector)* Using the eigenvalue decomposition $\check{X} = \sum_{i=1}^{n} \lambda_i q_i q_i^T$, select the maximum eigenvalue (say, it is j) and set

$$\hat{x} = \sqrt{\lambda_j} q_j$$

2. *(Randomization)* Draw samples

$$\hat{x} \sim \mathcal{N}(0, \check{X})$$

   and from the ones that satisfy the constraints

$$\hat{x}^T P_i \hat{x} \trianglerighteq_i 0, \quad i = 1, \ldots, m$$

   select the one yielding the smallest objective value

Interpretation of randomization

- x is a random vector from the Gaussian distribution with zero mean, covariance X
$$E[xx^T] = X$$
- The expected value of the objective function is
$$E[x^T P_0 x] = E[\text{trace}(x^T P_0 x)] = E[\text{trace}(P_0 x x^T)] = \text{trace}(P_0 E[xx^T]) = \text{trace}(P_0$$
- SDR looks for the covariance matrix that minimizes the original noncovex objective in expectation
- Likewise for the constraints
- It seems reasonable to use samples $x \sim \mathcal{N}(0, X)$ to generate an approximate solution to the QCQP

Implementation of randomization

Generate random vectors $x \sim \mathcal{N}(m, C)$ for given mean m and covariance matrix C

- Generate a random vector $y \in \mathbb{R}^n$ from the standard normal $\mathcal{N}(0, I_n)$
- In Matlab: `y = rand(n,1)`
- Find a matrix R such that $X = RR^T$
  - $C \succeq 0$ because it is a covariance matrix, so eigenvalues $(\lambda_1, \ldots, \lambda_n)$ are nonnegative
  - Eigendecomp. $C = Q\Lambda Q^T$; set $R = Q\sqrt{\Lambda}Q^T$ with $\sqrt{\Lambda} = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$

$$RR^T = Q\sqrt{\Lambda}Q^T Q\sqrt{\Lambda}^T Q^T = Q\sqrt{\Lambda}\sqrt{\Lambda}Q^T = Q\Lambda Q^T = C$$

    where we used that Q is orthonormal: $QQ^T = Q^T Q = I$
  - Matlab `[Q,D] = eig(C)` returns orthonormal eigenvectors in Q if C is real symmetric
  - A simpler option is $R = Q\sqrt{\Lambda}$, because $RR^T = Q\sqrt{\Lambda}\sqrt{\Lambda}^T Q^T = Q\Lambda^T Q^T = C$
  - Another option is Cholesky decomposition $C = LL^T$ (see Lecture 0)
- Then, $x = m + Ry \sim \mathcal{N}(m, C)$

$$E[x] = E[m + Ry] = m + RE[y] = m$$

$$\mathrm{Cov}[x] = E[(x - E[x])(x - E[x])^T] = E[Ryy^T R^T] = RE[yy^T]R^T = RI_nR^T = C$$

Relaxation

$$\begin{aligned} \check{f} = \text{min} \quad & \mathrm{trace}(P_0 X) \\ \text{subj. to} \quad & \mathrm{trace}(P_i X) \trianglerighteq_i 0 \\ & X \succeq 0 \end{aligned}$$

Nonconvex problem

$$\begin{aligned} f^* = \text{min} \quad & x^T P_0 x \\ \text{subj. to} \quad & x^T P_i x \trianglerighteq_i 0 \end{aligned}$$

Feasible point

$$\begin{aligned} \hat{f} &= \hat{x}^T P_0 \hat{x} \\ \text{where } & \hat{x}^T P_i \hat{x} \trianglerighteq_i 0 \end{aligned}$$

$$\boxed{\check{f} \leq f^* \leq \hat{f}}$$

- If $\hat{f} - \check{f}$ is small, we have approximately solved the problem
- If we generate the feasible $\hat{x}$ via randomization, draw many samples attempting to make $\hat{f} - \check{f}$ small

Nonconvex QCQP with variable $x \in \mathbb{R}^n$ ($P_i \in \mathbb{S}^n$ not necessarily positive semidefinite):

$$\begin{array}{rl} \text{minimize} & x^T P_0 x + q_0^T x \\ \text{subject to} & x^T P_i x + q_i^T x \trianglerighteq r_i, i = 1, \ldots, m \\ & Ax = b;\ Cx \preceq d \end{array}$$

- Same trick $[x^T P_i x = \text{trace}(P_i x x^T)]$ but now also keep vector x as variable
- We have the following equivalences (the latter is using Schur complement)

$$(X = xx^T,\ \text{rank}(X) = 1) \Longleftrightarrow (X \succeq xx^T,\ \text{rank}(X) = 1)$$
$$\Longleftrightarrow \left(\begin{bmatrix} X & x^T \\ x & 1 \end{bmatrix} \succeq 0,\ \text{rank}(X) = 1 \right)$$

$$\begin{array}{ll} \text{minimize} & x^T P_0 x + q_0^T x \\ \text{subject to} & x^T P_i x + q_i^T x \trianglerighteq r_i, i = 1, \ldots, m \\ & Ax = b;\ Cx \leq d \end{array}$$

SDP relaxation with variables $X \in \mathbb{S}^n$ and $x \in \mathbb{R}^n$

$$\begin{array}{ll} \text{minimize} & \operatorname{trace}(P_0 X) + q_0^T x \\ \text{subject to} & \operatorname{trace}(P_i X) + q_i^T x \trianglerighteq r_i, i = 1, \ldots, m \\ & Ax = b;\ Cx \preceq d \\ & \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0 \end{array}$$

Nonconvex QCQP with variable $x \in \mathbb{R}^n$ ($P_i \in \mathbb{S}^n$ not necessarily positive semidefinite):

$$\begin{aligned} \text{minimize} \quad & x^T P_0 x + q_0^T x \\ \text{subject to} \quad & x^T P_i x + q_i^T x \trianglerighteq r_i, i = 1, \dots, m \end{aligned} \tag{P1}$$

Introduce auxiliary variable $t \in \mathbb{R}$ and form the following homogeneous nonconvex QCQP:

$$\begin{aligned} \text{minimize} \quad & \begin{bmatrix} x \\ t \end{bmatrix}^T \begin{bmatrix} P_0 & q_0 \\ q_0^T & 0 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \\ \text{subject to} \quad & t^2 = 1 \\ & \begin{bmatrix} x \\ t \end{bmatrix}^T \begin{bmatrix} P_i & q_i \\ q_i^T & 0 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \trianglerighteq b_i, i = 1, \dots, m \end{aligned} \tag{P2}$$

- Let (x^*, t^*) be the optimal solution of (P2)
- Then x^* is the solution of (P1) if $t^* = 1$, and $-x^*$ is the solution of (P1) if $t^* = -1$
- We can now form the SDR of (P2) as we did in the homogeneous case

Randomization for Method 1

SDP relaxation with variables $X \in \mathbb{S}^n$ and $x \in \mathbb{R}^n$

$$\begin{array}{ll} \text{minimize} & \operatorname{trace}(P_0 X) + q_0^T x \\ \text{subject to} & \operatorname{trace}(P_i X) + q_i^T x \trianglerighteq r_i, i = 1, \ldots, m \\ & Ax = b;\ Cx \leq d \\ & \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0 \end{array}$$

- Let the solution be $(\check{X}, \check{x})$
- Generate $x \sim \mathcal{N}(\check{x}, \check{X} - \check{x}\check{x}^T)$
- Still not obvious how to make x feasible; see an example next

$$\begin{array}{ll} \text{minimize} & \|Ax-b\|_2^2 \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n \end{array}$$

- As before, introduce variable $X = xx^T$
- Notice that $X_{ii} = x_i^2$
- Equivalent problem with variables $X \in \mathbb{S}^n$, $x \in \mathbb{R}^n$

$$\begin{array}{ll} \text{minimize} & \operatorname{trace}(A^T AX) - 2b^T Ax + b^T b \\ \text{subject to} & X_{ii} = 1, i = 1, \ldots, n; \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0, \operatorname{rank}(X) = 1 \end{array}$$

SDR of Boolean Least Squares: Randomization and recovery of feasible solution

$$
\begin{array}{ll}
\text{minimize} & \operatorname{trace}(A^TAX) - 2b^TAx + b^Tb \\
\text{subject to} & X_{ii} = 1, i = 1, \ldots, n; \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0
\end{array}
$$

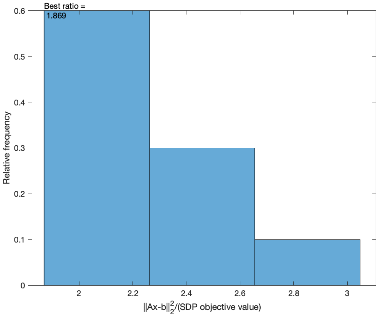
- Let the solution be $(\check{X}, \check{x})$
- Generate samples $x \sim \mathcal{N}(\check{x}, \check{X} - \check{x}\check{x}^T)$
- Obtain feasible solution as

$$
\hat{x}_i = \operatorname{sign}(x_i) = \begin{cases} 1, & \text{if } x_i \geq 0 \\ -1, & \text{if } x_i < 0 \end{cases}
$$

- Select the $\hat{x}$ that yields the smallest objective

- $A \in \mathbb{R}^{50\times 25}$, $b \in \mathbb{R}^{50}$ with Gaussian $\mathcal{N}(0,1)$ entries
- Solve SDP relaxation and use randomization to obtain a feasible solution
- We get the following histogram for $\hat{f}/\check{f}$
    - $\hat{f} = \|A\hat{x} - b\|_2^2$ from randomization; $\check{f}$ = SDP objective

10 samples



100 samples