Mini Project Report

Semidefinite programming method for integer convex quadratic minimization

Braghadeesh L

MTech, ECE, IISc

April 25, 2019

Outline

Introduction

Problem

Consider NP hard problem:

$$\begin{aligned} &\text{minimize} \quad f(x) = x^T P x + 2q^T x \\ &\text{subject to} \quad x \in \mathbf{Z}^n \end{aligned} \tag{1}$$

- Example: MIMO detection, solving least square problems etc
- Exhaustive search to find the optimal solution over the entire n-dimensional integer lattice is difficult.
- So, suboptimal is obtained by finding strong lower and upper bounds to optimal solution of (1)

Goal

Find strong lower and upper bounds for optimal value of (1)

Simple bounds for optimal value

- Simple lower bound: Remove integer constraints and solve unconstrained version of (1).
  Let x^{cts} and f^{cts} be the solution point and optimal value respectively to unconstrained version of (1).

- Simple upper bound is obtained by choosing any random integer point in n-dimensional lattice.

- Another upper bound is obtained by rounding off each entry of x^{cts} to nearest integer.

Formulating Lagrange duality

- Without loss of generality, we can assume that $x^{cts} \in [0,1]^n$ i.e, If $x^{cts} \notin [0,1]^n$, it can be translated to $[0,1]^n$ by solving following optimization problem:

$$\text{minimize}\,(x-v)^T P(x-v) + 2(Pv+q)^T(x-v) + f(v) \qquad (2)$$

$$\text{where } v = \text{floor}(x^{cts})$$

- Notice every integer point x satisfies $x_i \leq 0$ or $x_i \geq 1$ where x_i is the i^{th} coordinate of x.
- Using this fact, we relax the integer constraint $x \in \mathbf{Z}^n$ into set of nonconvex quadratic constraints:

$$x_i(x_i - 1) \geq 0 \quad \forall i = 1, 2, \ldots n$$

- The following non-convex problem is relaxation of (1):

$$\begin{aligned} \text{minimize} \quad & f(x) = x^T P x + 2q^T x \\ \text{subject to } & x_i(x_i - 1) \geq 0, \quad \forall i \in \{1,2,...n\} \end{aligned} \tag{3}$$

- The dual of the above problem is given by:

Dual

$$\begin{aligned} \text{maximize} \quad & -\tilde{q}(\lambda)^T (P - \mathbf{diag}(\lambda)) \tilde{q}(\lambda) \\ \text{subject to } & P - \mathbf{diag}(\lambda) \succeq 0 \\ & \tilde{q} \in R(P - \mathbf{diag}(\lambda)) \\ & \lambda \geq 0 \end{aligned} \tag{4}$$

where $\tilde{q}(\lambda) := q + (1/2)\lambda$

- Let $g(\lambda) = -\tilde{q}(\lambda)^T (P - \mathbf{diag}(\lambda)) \tilde{q}(\lambda)$

Comparison to simple lower bound

- Due to weak duality $g(\lambda) \leq \mathbf{f^*}$ where $\mathbf{f^*}$ is the optimal value to (1).

Theorem

Let $f^{sdp} = sup_{\lambda \geq 0}\, g(\lambda)$ be thee lower bound obtained by solving the lagrangian dual. Let ω_{max} and ω_{min} be the maximum and minimum eigenvalue of P. Let $\mathbf{1}$ be a vector whose entries are 1. Then we have

$$f^{sdp} - f^{cts} \geq \frac{n\omega_{min}^2}{4\omega_{max}} \left(1 - \frac{\|x^{cts} - (1/2)\mathbf{1}\|_2^2}{n/4}\right)^2. \tag{5}$$

where f^{sdp} is the optimal value to the dual (4).

- From the above result, $f^{sdp} \geq f^{cts}$, which indicates that f^{sdp} is a better lower bound to $\mathbf{f^*}$ than f^{cts}.

Semidefinite relaxation

- Problem (3) can be reformulated as

$$\begin{aligned} \text{minimize} \quad & \mathbf{Tr}(PX) + 2q^T x \\ \text{subject to} \quad & \mathbf{diag}(X) \geq x \\ & X = xx^T \end{aligned} \tag{6}$$

in the variables $X \in \mathbf{R}^{n \times n}$ and $x \in \mathbf{R}^n$

- Then we can relax the non-convex constraint $X = xx^T$ into $X \succeq xx^T$ and rewrite it using the Schur's complements to obtain a convex relaxation:

$$\begin{aligned} \text{minimize} \quad & \mathbf{Tr}(PX) + 2q^T x \\ \text{subject to} \quad & \mathbf{diag}(X) \geq x \\ & \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0 \end{aligned} \tag{7}$$

- The optimal value of problem(7) gives a lower bound on **f***, just as Lagrangian dual (4) gives a lower bound f^{sdp} on **f***.

- Problem(7) and problem (4) are duals of each other and hence give the same lower bound f^{sdp}.

The semidefinite relaxation (7) has a natural probabilistic interpretation, which can be used to construct a simple randomized algorithm for obtaining good suboptimal solutions. That is, optimal values of (7), denoted by X^* and x^* are the parameters of Gaussian distribution $\mathcal{N}(\mu, \sum)$ where $\mu = x^*$ and $\sum = X^* - x^* {x^*}^T$.The algorithm is listed below:

Given number of iterations K

1. Solve (7) to get X^* and x^*.
2. Form covariance matrix $\sum := X^* - x^* {x^*}^T$.
3. Initialize best point $x^{best} := 0$ and $f^{best} := 0$. for k=1,2....K
4. Random sampling: $z^k := x^*$, where $x^* \sim \mathcal{N}(x^*, \sum)$.
5. Round to nearest integer: $x^k := round(z^k)$.
6. Update the best point: If $f^{best} > f(x^k)$, then set $x^{best} := x^k$ and $f^{best} := f(x^k)$.

Greedy 1-opt

Definition

we say that $x \in \mathbf{Z}^n$ is 1-opt if the objective value at x does not improve by changing single coordinate, i.e., $f(x + ce_i) \geq f(x)$ for all indices i and integer c, where $f(.)$ is the objective function in (1).

we shall look at the algorithm:
given an initial point $x \in \mathbf{Z}^n$.

1. Compute initial gradient $g = 2(Px + q)$, where g is the gradient
   repeat
2. Stopping criterion: **quit** if $\mathbf{diag(P)} \geq |g|$
3. Find descent direction: Find index i and integer c minimizing $c^2 P_{ii} + cg_i$.
4. Update x. $x_i := x_i + c$.
5. Update gradient: $g := g + 2cP_i$.

Method for obtaining bounds for optimal value

1. Find x^{cts} by removing integer constraints in (1) and find $f(x^{cts}) = f^{cts}$. This is the simple lower bound on **f***.
2. Choose any random $x^{rnd} \in \mathbf{Z}^n$. Find $f(x^{rnd}) = f^{rnd}$. f^{rnd} is an upper bound on **f***. Now use 1-opt algorithm by taking initial point as x^{rnd} to get $\hat{x}^{rnd}$. $f(\hat{x}^{rnd})$ is an upper bound (better than f^{rnd}).
3. Now we do SDP relaxation on (1) to find lower bound f^{sdp} $(> f^{cts})$ on **f***. We then run randomized algorithm (3.1) to find x^{best} and find $f(x^{best}) = f^{best}$, which is an upper bound on **f***.
4. We also run 1-opt algorithm on every feasible solution point obtained in randomized algorithm (3.1) (after SDP relaxation) and find $\hat{x}^{best}$.
5. Finally we get $f(x^{rnd})$, $f(\hat{x}^{rnd})$, f^{best}, and $\hat{f}^{best}$ as upper bounds and f^{cts} and f^{sdp} as lower bounds.

Results

Sr.No.	Dimension(n)	Lower bound (f^{sdp})	Upper bound(f^{best})	L2 norm of suboptimal solution(x^{best})
1	130	141.90350	151.12220	11.35782
2	120	118.38522	125.58434	10.58301
3	110	105.65600	112.46488	11.26943
4	100	123.34624	128.95646	12.36932
5	90	75.97197	80.18095	10.04988

Table: Output

Conclusion

1. Obtained simple lower bound on optimal value of (1) by removing the integer constraint.
2. The lower bound is further improved by making SDP relaxation.
3. In order to get better upper bounds for the optimal value of (1), randomized algorithm and greedy 1-opt algorithm are carried out.
4. Instead of searching over the entire n- dimensional integer lattice, search is done over restricted space by making some relaxation in order to obtain suboptimal solution to (1).

References

J. Park and S. Boyd.
A semidefinite programming method for integer convex quadratic minimization.
Springer-Verlag, Berlin, Heidelberg, 2017.

S. Boyd and L. Vandenberghe
Convex Optimization.
Cambridge University Press, 2004