The Journal of
Mathematics and Computer Science

Available online at
http://www.TJMCS.com

The Journal of Mathematics and Computer Science Vol .5 No.4 (2012) 304 - 312

Nonconvex Optimization with Dual Bounds and Application in Communication Systems

Akram Ahadizadeh
Msc student of mathematics of sheikhbahaei university,Iran,Isfahan
Ahadi168@yahoo.com
Sadaf Anbarzadeh
Assistant professor of sheikhbahaei university, Iran,Isfahan
Anbarzadeh00@yahoo.com

Received: February 2012, Revised: November 2012
Online Publication: December 2012

Abstract

Convex optimization has provided both a powerful tool and an intriguing mentality to the analysis and design of communication systems over the last few years.
This paper presents the results of investigation on dual bounds for nonconvex quadratic programming with a nonlinear constraint and an overview of the nonconvex optimization problem in the networked communication systems.

Keywords: Dual bound, Quadratic programming, Duality bound method, Nonconvex optimization, Network utility maximization.

2010 Mathematics Subject Classification: Primary 54A40; Secondary 46S40.

1. Introduction.

There has been two major "waves" in the history of optimization theory: the first started with linear programming and simplex method in late 1940s, and the second with convex optimization and interior point method in late 1980s. Each has been followed by a transforming period of appreciation-application cycle: as more people appreciate the use of LP/convex optimization, more look for their formulations in various applications; then more work on its theory, efficient algorithms and softwares; the more powerful the tools become; then more people appreciate its usage. Communication systems benefit significantly from both waves, including multicommodity flow solutions (e.g., Bellman Ford algorithm) from LP, and basic network utility maximization. and robust transceiver design from convex optimization.
Much of the current research frontier is about the potential of the third wave,on nonconvex optimization. A variety of approaches have been proposed, from nonlinear transformation to turn an apparently nonconvex problem into a convex problem, to characterization of attraction regions

and systematically jumping out of a local optimum, from successive convex approximation to dualization, from leveraging the specific structures of the problems (e.g., Difference of Convex functions, concave minimization, low rank nonconvexity) to developing more efficient branch-and-bound procedures. Researchers in communications and networking have been examining nonconvex optimization using domain-specific structures in important problems in the areas of wireless networking, Internet engineering, and communication theory. Perhaps four typical topics best illustrate the variety of challenging issues arising from nonconvex optimization in communication systems:

- Nonconvex objective to be minimized. An example is congestion control for inelastic applications.
- Nonconvex constraint set. An example is power control in low SIR regimes.
- Integer constraints. Two important examples are single path routing and multiuser detection.
- Constraint sets that are convex but require an exponential number of inequalities to explicitly describe. An example is optimal scheduling.

The last results of latest researches about two first options are studied in [4].

2. On duality bound for general quadratic programming

The general quadratic programming problem with one additional quadratic constraint is formulated as follows:

$$\begin{aligned} \min \quad & f(x) = \langle x, Qx \rangle + \langle q, x \rangle \\ s.t \quad & Ax + d \le \circ \\ & \langle x, Cx \rangle + \langle c, x \rangle + h \le \circ \\ & x \in R^{\circ} \subset \Re^{n}_{+} \end{aligned} \qquad (Q)$$

where Q, C are real $n \times n$ matrices,

$$\begin{aligned} & q, c \in \Re^{n}, A \in \Re^{m \times n} \\ & a^{\circ}, b^{\circ} \in \Re^{n}, \ R^{\circ} = \left\{ x \in \Re^{n} : \circ \le a^{\circ} \le x \le b^{\circ} \right\} \\ & h \in \Re \\ & D = \left\{ x \in \Re^{n} : Ax + d \le \circ, \langle x, Cx \rangle \right. \\ & \qquad \left. + \langle c, x \rangle + h \le \circ, x \in R^{\circ} \right\} \end{aligned}$$

2.1. Lower bound

For each rectangle

$$R = \left\{ x \in \Re^{n} : a \le x \le b \right\} \subseteq R^{\circ} (\subset \Re^{n}_{+})$$

We intend to compute a lower bound $\mu(R)$ of the function f over R∩D.

In other words, we compute a lower bound for the optimal value of the problem

$$\begin{aligned} \min \quad & f(x) = \langle x, Qx \rangle + \langle q, x \rangle \\ s.t \quad & Ax + d \le \circ \\ & \langle x, Cx \rangle + \langle c, x \rangle + h \le \circ \\ & x \in R \end{aligned} \qquad (Q(R))$$

The main idea of our method for computing a lower bound *μ(R)* is to construct from Problem (Q(R)) an equivalent problem by replacing the quadratic terms with corresponding bilinear forms, and then solving the Lagrange dual of the resulting problem ([1], Proposition 2.6.1)

For this purpose let $\alpha, \overline{\alpha}, \beta, \overline{\beta}$ be vectors of $\Re^n$ defined by

$$\begin{aligned} \alpha_i &= \min\{< Q_i, x >: x \in R\} \\ \overline{\alpha}_i &= \max\{< Q_i, x >: x \in R\} - \alpha_i \end{aligned} \qquad (1)$$

$$\begin{aligned} \beta_i &= \min\{< C_i, x >: x \in R\} \\ \overline{\beta}_i &= \max\{< C_i, x >: x \in R\} - \beta_i \end{aligned} \qquad (2)$$

where for each $i = 1, \ldots, n$, Q_i and C_i are the i th rows of the matrices Q and C, respectively.
with the help of two additional variables $y, z \in \Re^n$ the problem can be written as

$$\begin{aligned} \min \quad & g(x,y) = < x, y > + < q + \alpha, x > \\ s.t \quad & Qx - y - \alpha \le \circ, \\ & Ax + d \le \circ, \\ & Cx - z - \beta \le \circ, \\ & < x, z > + < \beta + c, x > + h \le \circ, \\ & y \le \overline{\alpha},\ z \le \overline{\beta}, \\ & x \in R, y \ge \circ, z \ge \circ. \end{aligned} \qquad (P(R))$$

It is well known that the optimal value of the lagrangian dual of problem P(R) gives a lower bound for the
optimal value of Q(R) ([7], Proposition 2.1).Denote this lower bound by *μ(R)*.The following was shown in[7]:

(i) ([7], Proposition 2.2) *μ(R)* is the optimal value of the linear program LP (R):

$$\begin{aligned} \max \quad & < q + \alpha, a > + \{< a - b, s > + < Qa - \alpha, u > + < Aa + d, v > \\ & \quad + < Ca - \beta, w > + (< a, \beta > + < a, c > + h)t - \\ & \quad < \overline{\alpha}, \overline{u} > - < \overline{\beta}, \overline{w} > \ \} \\ s.t. \quad & s + Q^T u + A^T v + C^T w + \\ & t(\beta + c) \ge -(q + \alpha), \\ & -u + \overline{u} \ge -a, -w + ta + \overline{w} \ge \circ \\ & (s, u, v, w, t, \overline{u}, \overline{w}) \ge \circ. \end{aligned} \qquad (LP(R))$$

(ii) ([7], Propositions 2.3) The bound *μ(R)* is at least as good as any bound obtained by convex relaxation of P(R).

Assertion (i) is proved in [7] by showing, through tiring computations, that the lagrangian dual of P(R) reduces to the linear program LP(R). However, it will be shown shortly that these computations are unnecessary because μ(R) is merely the optimal value of a trivial linear relaxation of Q(R), which can be derived immediately without any use of Lagrange relaxation. On the other hand, (ii) is refuted by the following counter-example.

Counter-example. Consider the quadratic program with one variable $x \in R$:

$$\min \{x^2 - 3x : -x^2 + 4x \le \circ, \quad \circ \le x \le 4\} \quad (3)$$

whose optimal value is obviously 0. As can easily be verified, here $\overline{\alpha} = 4, \alpha = \circ, < q + \alpha, a >= \circ, \overline{\beta} = 4, \beta = -4$ and the associated problem P(R) is

$$\begin{aligned} \min \quad & xy - 3x \\ s.t \quad & x - y \le \circ \\ & -x - z + 4 \le \circ \\ & xz \le \circ \\ & \circ \le x \le 4 \quad , \ \circ \le y \le 4, \ \circ \le z \le 4 \end{aligned}$$

The linear program LP(R) is then

$$\begin{aligned} \max \quad & -4s + 4w - 4\overline{u} - 4\overline{w} \\ s.t \quad & s + u - w \ge 3 \\ & -u + \overline{u} \ge \circ, -w + \overline{w} \ge \circ \\ & (s, u, w, \overline{u}, \overline{w}) \ge \circ \quad . \end{aligned}$$

Solving this linear program yields $\mu(R) = -12$, a bound much inferior to the optimal value 0 of Q(R) Furthermore, μ(R) is inferior to the bound provided by the convex relaxation of P(R) that is obtained by replacing xy and xz with their convex envelopes on R, namely

$$\max \{\circ, 4x + 4y - 16\}, \ \max \{\circ, 4x + 4z - 16\}$$

This convex relaxation of P(R) reduces to the linear program

$$\begin{aligned} \min \quad & t - 3x \\ s.t \quad & 4x + 4y - 16 \le t, \ x - y \le \circ \\ & -x - z + 4 \le \circ, \ 4x + 4z - 16 \le \circ \\ & \circ \le x \le 4, \ \circ \le y \le 4, \ \circ \le z \le 4, \ t \ge \circ \end{aligned} \quad (4)$$

whose optimal value equals − 6, a much better bound than $\mu(R) = -12$. Therefore, assertion (ii) is not true. □

It is also worth noticing that the lagrangian dual of problem (3) itself is

$$\sup_{u \ge \circ} \inf \{x^2 - 3x + (-x^2 + 4x)u : \quad \circ \le x \le 4\}$$

Since for u = 1:

$$\inf \{x^2 - 3x + (-x^2 + 4x)u : \circ \le x \le 4\} =$$
$$\inf \{x : \quad \circ \le x \le 4\} = \circ$$

while the optimal value of (3) is 0 we have by the weak duality theorem[2]

$$\sup_{u \geq \circ} \inf \left\{ x^2 - 3x + (-x^2 + 4x)u : \quad \circ \leq x \leq 4 \right\} = \circ$$

Thus, for certain problems the lagrangian dual of the original quadratic program Q(R) may give a better bound than the lagrangian dual of P(R). It may even happen, as in the example under consideration, that the duality gap is zero for Q(R), but is positive for P(R), although the two problems are equivalent.This should be kept in mind when using lagrangian relaxation to estimate bounds.

It is now easy to see why *μ(R)* can be sometimes very poor. In fact, as can easily be checked, the linear program DLP(R) dual to LP(R) is simply (see [7], where $\xi + a = x$):

$$\begin{array}{ll} \min & < a, y > + < q + \alpha, x > \\ s.t & Qx - y - \alpha \leq \circ \\ & Ax + d \leq \circ \\ & Cx - z - \beta \leq \circ \\ & < a, z > + < \beta + c, x > + h \leq \circ \\ & y \leq \overline{\alpha} \\ & z \leq \overline{\beta} \\ & a \leq x \leq b, y \geq \circ \quad , \ z \geq \circ \end{array} \qquad \text{(DLP(R))}$$

Which is nothing but the linear program obtained from P(R) by replacing $< x, y >$ with $< a, y >$ and $< x, z >$ with $< a, z >$. Since $< a, y >$, $< a, z >$ are linear functions underestimating $< x, y >$, $< x, z >$ for $y \geq \circ, \circ \leq a < x \leq b$, z $\geq \circ$, DLP(R) is a most common linear relaxation of P(R).
We see that no involved computation of the type used in [7] is needed for deriving *μ(R)*.

Besides, by writing Q(R) as

$$Q(R) \qquad \begin{array}{ll} \min & < x, Qx - \alpha > + < q + \alpha, x > \\ s.t & Ax + d \leq \circ \\ & < x, Cx - \beta > + < \beta + c, x > + h \leq \circ \\ & a \leq x \leq b, \end{array}$$

it is easily seen that a linear relaxation of Q(R) is

$$\text{LRQ(R)} \qquad \begin{aligned} \min \quad & < a, Qx - \alpha > + < q + \alpha, x > \\ s.t \quad & Ax + d \le \circ \\ & < a, Cx - \beta > + < \beta + c, x > + h \le \circ \\ & a \le x \le b \ , \end{aligned}$$

which is exactly the same as DLP(R), though looking much simpler.
To sum up, the bound $\mu(R)$ can be computed by simple linear relaxation of either P(R) or Q(R).
No wonder that it is generally poorer than the bounds computed by tighter convex relaxations[3].
The thing is,one must be aware that the lagrangian duals of two equivalent problems may be quite different , so that by taking inappropriate equivalent formulations of the problem one can obtain even very poor bounds, as was demonstrated by example (3). In a general manner, the quality of a dual bound depends on the specific equivalent formulation of the original problem used for generating the lagrangian dual problem. The formulation P(R) used in [7] is a bad choice, whereas by adding the superfluous constraint $x(x-4) \le \circ$ (implied by $\circ \le x \le 4$) we get another equivalent formulation of (3) with no duality gap:

$$\begin{aligned} & \sup_{u \in \Re} \inf \left\{ x^2 - 3x + (-x^2 + 4x)u : \quad \circ \le x \le 4 \right\} \\ & = \min \left\{ x^2 - 3x : \ -x^2 + 4x = \circ . \quad \circ \le x \le 4 \right\} \\ & = \min \left\{ x : \circ \le x \le 4 \right\} = \circ \end{aligned}$$

Therefore, contrary to Proposition 2.3 in [7], the bounds computed according to [7] are even inferior to those earlier proposed in the literature.

3. Internet Congestion Control

3.1. Basic network utility maximization

Since the publication of the seminal paper [5] in 1998, the framework of Network Utility Maximization (NUM) has found many applications in network rate allocation algorithms and Internet congestion control protocols.
It has also lead to a systematic understanding the entire network protocol stack in the unifying framework.
By allowing nonlinear concave utility objective functions, NUM substantially expands the scope of the classical LP-based Network Flow Problems.

Consider a communication network with L links, each with a fixed capacity of c_l bps, and S sources (i.e., end users), each transmitting at a source rate of x_s bps. Each source s emits one flow, using a fixed set $L(s)$ of links in its path, and has a utility function $U_s(x_s)$. Each link l is shared by a set S(l) of sources.
Network Utility Maximization (NUM), in its basic version, is the following problem of maximizing the total utility of the network $\sum_s U_s(x_s)$, over the source rates X , subject to linear flow constraints $\sum_{s:l \in L(s)} (x_s) \le c_l$ for all links :

$$\begin{aligned} \max \quad & \sum_s U_s(x_s) \\ s.t \quad & \sum_{s\in S(l)} x_s \le c_l \quad ,\forall l, \quad (5) \\ & X \ge \circ \end{aligned}$$

where the variables are $X \in \Re^s$.

There are many nice properties of the basic NUM model due to several simplifying assumptions of the utility functions and flow constraints, which provide the mathematical tractability of problem (5) but also limit its applicability. In particular, the utility functions { U_s} are often assumed to be increasing and strictly concave functions.

Assuming that $U_s(x_s)$ becomes concave for large enough x_s is reasonable, because the law of diminishing marginal utility eventually will be effective.

However, U_s may not be concave throughout its domain. In his seminal paper published a decade ago, Shenker [6] differentiated inelastic network traffic from elastic traffic. Utility functions for elastic traffic were modeled as strictly concave functions. While inelastic flows with nonconcave utility functions represent important applications in practice, they have received little attention and rate allocation among them has scarcely any mathematical foundation, except three recent publications.

3.2. Canonical distributed algorithm

For network rate allocation problems, a dual-based distributed algorithm has been widely studied and is summarized below.

Zero duality gap for (5) states that the solving the Lagrange dual problem is equivalent to solving the primal problem (5). The Lagrange dual problem is readily derived. We first form the Lagrangian of (5):

$$L(x,\lambda) = \sum_s U_S(x_S) + \sum_l \lambda_l \left(c_l - \sum_{s\in S(l)} x_s \right)$$

Where $\lambda_l \ge \circ$ is the Lagrange multiplier (link congestion price) associated with the linear flow constraint on link *l*. Additivity of total utility and linearity of flow constraints lead to a Lagrangian dual decomposition into individual source terms:

$$\begin{aligned} L(X,\lambda) &= \sum_S \left[U_S(x_S) - (\sum_{l\in L(S)} \lambda_l) x_S \right] + \sum_l c_l \lambda_l \\ &= \sum_S L_S(x_s, \lambda^s) + \sum_l c_l \lambda_l \end{aligned}$$

Where $\lambda^s = \sum_{l\in L(s)} \lambda_l$. For each source s, $L_s(x_s, \lambda^s) = U_S(x_s) - \lambda^s x_s$ only depends on local x_s and the link prices λ_l on those links used by source s.

The Lagrange dual function g(λ) is defined as the maximized L(X, λ) over X. This "net utility" maximization obviously can be conducted distributively by the each source, as long as the

aggregate link price $\lambda^s = \sum_{l \in L(s)} \lambda_l$ is available to source s, where source s maximizes a strictly concave function $L_s(x_s, \lambda^s)$ over x_s for a given λ^s

$$x_s^*(\lambda^s) = \arg\max\left[U_s\ (x_s) - \lambda^s\ x_s\right], \quad \forall s. \tag{6}$$

The Lagrange dual problem is

$$\begin{aligned} \min \quad & g(\lambda) = L(X^*(\lambda), \lambda) \\ s.t \quad & \lambda \geq \circ \end{aligned} \tag{7}$$

where the optimization variable is λ.

The concavity assumption on U_s is also related to the elasticity assumption on rate demands by users.

When demands for x_s are not perfectly elastic, $U_s(x_s)$ may not be concave.

Suppose we remove the critical assumption that $\{U_s\}$ are concave functions, and allow them to be any nonlinear functions. The resulting NUM becomes nonconvex optimization and significantly harder to be analyzed and solved, even by centralized computational methods. In particular, a local optimum may not be a global optimum and the duality gap can be strictly positive. The standard distributive algorithms that solve the dual problem may produce infeasible or suboptimal rate allocation, as illustrated in (3).

4. Conclusion

Removing the concavity assumption on utility functions, NUM is turned into a nonconvex optimization problem with a strictly positive duality gap. Such problems in general are NP hard, thus extremely unlikely to be polynomial-time solvable even by centralized computations. Using a family of convex semidefinite programming (SDP) relaxations based on the sum-of-squares (SOS) relaxation and the Positivstellensatz Theorem in real algebraic geometry, we apply a centralized computational method to bound the total network utility in polynomial-time[4].

Acknowledgement. We would like to express our gratitude to all those who gave us the possibility to complete this paper.

References

[1] J.P. Aubin: *Applied Functional Analysis* , University of Paris – Dauphine, 1999.

[2] D.G. Luenberger: Linear *and Nonlinear Programming, Second* Edition, Standford University, 1984.

[3] H. Tuy: Convex *Analysis and Global Optimization*, kluwer, 1998.

[4] M. Chiang *''Nonconvex Optimization for Communication Systems*", Electrical Engineering Department, Princeton University , Princeton , NJ 08544,USA.

[5] F.P. Kelly, A. Maulloo and D. Tan "*Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability*", Journal of Operations Research Society, Vol.49, no .3, PP.237-252, March 1998.

[6] S. Shenker "*Fundamental Design Issues for the Future Internet* ". IEEE J.Scl .Area Comm., Vol.13, no.7, PP.1176-1188, Sept.1995.

[7] N.V. Thoai "*Duality Bound Method for the General Quadratic Progrmming Problem with Quadratic Constraints* ", Journal of Optimization Theory and Applications, Vol.107, pp.331-354, 2000.
[8] H. Tuy "*On Dual Bounds for Nonconvex Optimization*"Institute of Mathematics, 18 Hoang Quoc Viet, 10307, Hanoi, Vietnam.