Nonconvex Quadratic Problems: Strong Duality Results

Javier Zazo

Universidad Politécnica de Madrid
Department of Telecommunications Engineering

javier.zazo.ruiz@upm.es

March 17, 2017

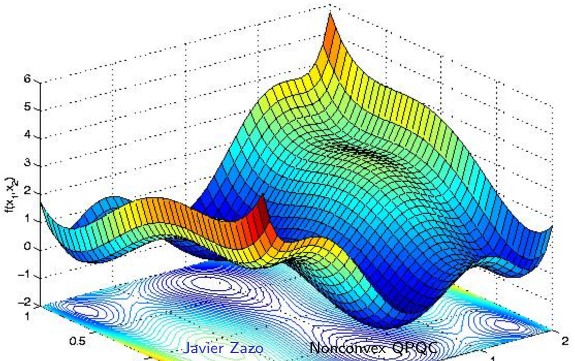
Outline

Quadratic Problem with Quadratic Constraints (QPQC)

- Let's consider a general QPQC:

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^p}{\text{minimize}} \quad & \mathbf{x}^T A_0 \mathbf{x} + 2b_0^T \mathbf{x} + c_0 \\ \text{subject to} \quad & \mathbf{x}^T A_i \mathbf{x} + 2b_i^T \mathbf{x} + c_i \le 0 \quad \forall i = 1, \dots, m. \end{aligned}$$

  where A_0, A_i are symmetric matrices and b_0, bi, $\mathbf{x} \in \mathbb{R}^p$, c_0, $c_i \in \mathbb{R}$.
- If $A_0 \succeq 0$ and every $A_i \succeq 0$ the problem is convex ($\approx$ easy to solve).
- Otherwise, the problem is non-convex (local minima may exist).
- These problems are generally NP-Hard.

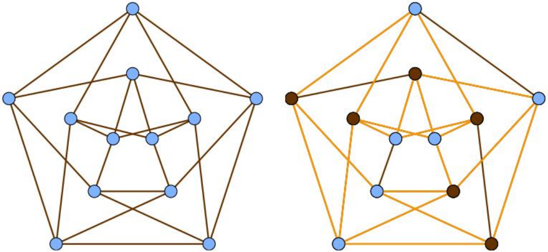


Max-Cut Problem

Nonconvex Examples I

- Given an undirected graph, with no self-loops
- Vertex set $V = \{1, \ldots, n\}$ and edge set E.
- For a subset $S \subset V$, the capacity of S is the number of edges connecting a node in S to a node not in S.

find $S \subset V$ with maximum capacity



Partinioning Problems

Nonconvex Examples II

- Also called "Boolean Optimization":

$$\begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & \mathbf{x}^T Q \mathbf{x} \\ \text{subject to} & x_i \in \{-1, 1\} \end{array}$$

- The problem is NP-complete (even if $Q \succeq 0$).
- Constraints of the form $x_i \in \{-1, 1\} \Leftrightarrow x_i^2 = 1$.
- The MAXCUT

$$\begin{array}{ll} \text{maximize} & \displaystyle\sum_{(v_i, v_j) \in E} w_{ij} \frac{1 - x_i x_j}{2} \\ \text{subject to} & x_i^2 = 1 \quad \forall v_i \in V. \end{array}$$

Polynomial Minimization

Nonconvex Examples III

- Minimize a polynomial over a set of of polynomial inequalities:

$$\begin{array}{ll} \text{minimize} & p_0(\mathbf{x}) \\ \text{subject to} & p_i(\mathbf{x}) \le 0, \quad i = 1, \dots, m. \end{array}$$

- Example:

$$\begin{array}{ll} \underset{x,y,z}{\text{minimize}} & x^3 - 2xyz + y + 2 \\ \text{subject to} & x^2 + y^2 + z^2 - 1 = 0. \end{array}$$

Introducing change of variables $u = x^2$, $v = yz$, we get

$$\begin{array}{ll} \text{minimize} & ux - 2vx + y + 2 \\ \text{subject to} & x^2 + y^2 + z^2 - 1 = 0 \\ & u - x^2 = 0 \\ & v - yz = 0. \end{array}$$

Semidefinite Relaxation

- Given a QPQC:

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^p}{\text{minimize}} \quad & \mathbf{x}^T A_0 \mathbf{x} + 2b_0^T \mathbf{x} + c_0 \\ \text{subject to} \quad & \mathbf{x}^T A_i \mathbf{x} + 2b_i^T \mathbf{x} + c_i \leq 0 \quad \forall i = 1, \dots, m, \end{aligned} \tag{1}$$

we can transform to

$$\begin{aligned} \underset{X\in\mathbb{R}^{p\times p}, \mathbf{x}\in\mathbb{R}^p}{\text{minimize}} \quad & \text{tr}(A_0 X) + 2b_0^T \mathbf{x} + c_0 \\ \text{subject to} \quad & \text{tr}(A_i X) + 2b_i^T \mathbf{x} + c_i \leq 0 \quad \forall i = 1, \dots, m \\ & X = \mathbf{x}\mathbf{x}^T. \end{aligned}$$

- We relax the rank constraint to $X \succeq \mathbf{x}^T \mathbf{x}$ and reformulate to

$$\begin{aligned} \underset{X\in\mathbb{R}^{p\times p}, \mathbf{x}\in\mathbb{R}^p}{\text{minimize}} \quad & \text{tr}(A_0 X) + 2b_0^T \mathbf{x} + c_0 \\ \text{subject to} \quad & \text{tr}(A_i X) + 2b_i^T \mathbf{x} + c_i \leq 0 \quad \forall i = 1, \dots, m \\ & \begin{bmatrix} X & \mathbf{x} \\ \mathbf{x}^T & 1 \end{bmatrix} \succeq 0. \end{aligned} \tag{2}$$

- **Strong duality**: problems (1) and (2) attain the same solution.

QP with a single QC

Strong duality I

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^T A_0 \mathbf{x} + 2b_0^T \mathbf{x} + c_0 \\
\text{subject to} \quad & \mathbf{x}^T A_1 \mathbf{x} + 2b_1^T \mathbf{x} + c_1 \leq 0
\end{aligned}
$$

Strong duality holds provided Slater's condition holds:

$$\exists \hat{\mathbf{x}} \mid \hat{\mathbf{x}}^T A_1 \hat{\mathbf{x}} + 2b_1^T \hat{\mathbf{x}} + c_1 < 0$$

Applications:

- Principal Component Analysis (PCA):

$$\underset{\|\mathbf{x}\| \leq 1}{\arg\max} \quad \|Q\mathbf{x}\|^2 = \underset{\|\mathbf{x}\| \leq 1}{\arg\max} \quad \mathbf{x}^T Q^T Q \mathbf{x} \tag{3}$$

- Trust Region methods

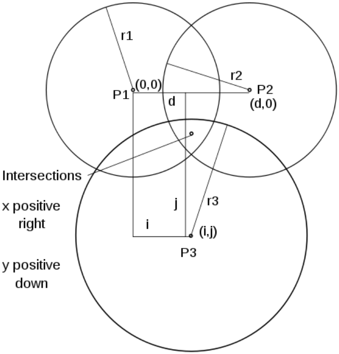
QP with Equality QC

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^T A_0 \mathbf{x} + 2b_0^T \mathbf{x} + c_0 \\ \text{subject to} \quad & g_1(\mathbf{x}) = \mathbf{x}^T A_1 \mathbf{x} + 2b_1^T \mathbf{x} + c_1 = 0 \end{aligned}$$

Strong duality holds provided $A_1 \neq 0$ and

$$\exists \hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2 \,|\, g_1(\hat{\mathbf{x}}_1) < 0 \wedge g_1(\hat{\mathbf{x}}_2) > 0.$$

Application: Localization problem



$$\underset{\mathbf{x}}{\text{minimize}} \quad \sum_i (\|\mathbf{x} - s_i\|^2 - d_i^2)^2$$

$$\begin{aligned} \underset{\mathbf{x},\alpha}{\text{minimize}} \quad & \sum_i (\alpha - 2s_i^T \mathbf{x} + s_i^T s_i)^2 \\ \text{subject to} \quad & \alpha = x^T x \end{aligned}$$

QP with two QC

Strong Duality II

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^T A_0 \mathbf{x} + 2b_0^T \mathbf{x} + c_0 \\ \text{subject to} \quad & \mathbf{x}^T A_1 \mathbf{x} + 2b_1^T \mathbf{x} + c_1 \leq 0 \\ & \mathbf{x}^T A_2 \mathbf{x} + 2b_2^T \mathbf{x} + c_2 \leq 0 \end{aligned}$$

Strong duality holds provided Slater's condition holds and

$$\exists \lambda_1, \lambda_2 \geq 0 \,|\, \lambda_1 A_1 + \lambda_2 A_2 \succ 0. \tag{4}$$

Application: Trust Region methods. Problem: $\min\ f(x)$ s.t. $c(x) = 0$

$$\underset{\mathbf{x}}{\text{minimize}} \quad \mathbf{x}^T \nabla f(\mathbf{x}^k) + \frac{1}{2} d^T B_k d \tag{5}$$

$$\text{subject to} \quad \|c(x^k) + \nabla c(x^k)^T d\| \leq \xi_k \tag{6}$$

$$\|d\| \leq \Delta_k \tag{7}$$

General QP with Z-matrices

We form the following matrices out of the QP:

$$H_i = \begin{bmatrix} A_i & b_i \\ b_i^T & c_i \end{bmatrix}$$

Definition

A matrix A is called a Z-matrix if every off-diagonal element is non-positive.

Suppose that Slater's condition is satisfied.
If every H_i, $i = 0, \ldots, m$ is a Z-matrix, strong duality holds.

S-property $\iff$ Strong Duality

Definition

A QP satisfies the S-property if and only if the following statements are strong alternatives:

- $\exists \mathbf{x} \in \mathbb{R}^p \,|\, f(\mathbf{x}) < 0 \,\wedge\, g_i(\mathbf{x}) \leq 0 \quad i \in 1, \dots, m$
- $\nexists \lambda_i \geq 0,\, i \in 1, \dots, m \,|\, f(\mathbf{x}) + \sum_i \lambda_i g_i(\mathbf{x}) \geq 0$

Theorem (Jeyakumar et. al 2006)

Suppose the QP satisfies the S-property. Let $\mathbf{x}^*$ be a feasible point of the QP. Then, $\mathbf{x}$ is a global minimizer of the QP if and only if:

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \sum_i \lambda_i \nabla_{\mathbf{x}} g_i(\mathbf{x}^*) = 0$$

$$\sum_i \lambda_i g_i(\mathbf{x}^*) = 0, \quad A_0 + \sum_i \lambda_i A_i \succeq 0.$$

Theory of Strong Alternatives

We say that two systems are strong alternatives if the systems cannot be feasible at the same time, but one of them must be. For instance, the following systems are strong alternatives:

- $\lambda_i \in \mathbb{R}$, $H_0 + \lambda_1 H_1 + \ldots + \lambda_m H_m \succeq 0$
- $Z \succeq 0$, $\mathrm{tr}(H_0 Z) < 0 \wedge \mathrm{tr}(H_i Z) = 0$

Requirement: H_i have a positive and negative eigenvalue $i \geq 1$.

$$H_i = \begin{bmatrix} A_i & b_i \\ b_i^T & c_i \end{bmatrix}$$

QP with Equality Quadratic Constraints

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^T A_0 \mathbf{x} + 2 b_0^T \mathbf{x} + c_0 \\ \text{subject to} \quad & x_i^T \widetilde{A}_i x_i + 2 \widetilde{b}_i^T x_i + c_i = 0 \quad \forall i = 1, \ldots, m \end{aligned}$$

where all constraints are separable and $\mathbf{x} = [x_1, \ldots, x_m]$.

Definition

A set of matrices $\{H_1, H_2, \ldots, H_m\}$ is said to be simultaneously diagonalizable via congruence, if there exists a nonsingular matrix P such that $P^T H_i P$ is diagonal for every matrix H_i.

Simultaneous Diagonalization via Congruence

$$P = \begin{pmatrix} P_1 & 0_{n_1 \times n_2} & \cdots & p_1 \\ 0_{n_2 \times n_1} & P_2 & \cdots & p_2 \\ \vdots & \ddots & \ddots & \vdots \\ 0_{n_N \times n_1} & \cdots & P_N & p_m \\ 0_{1 \times n_1} & 0_{1 \times n_2} & \cdots & 1 \end{pmatrix}$$

$$F_i = P^T H_i P = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & P_i^T \widetilde{A}_i P_i & 0 & P_i^T (\widetilde{A}_i p_i + \widetilde{b}_i) \\ \vdots & \cdots & \ddots & 0 \\ 0 & (p_i^T \widetilde{A}_i^T + \widetilde{b}_i^T) P_i & 0 & p_i^T \widetilde{A}_i p_i + 2\widetilde{b}_i^T p_i + c_i \end{pmatrix}$$

Sufficient condition: If $\widetilde{b}_i \in \text{range}(\widetilde{A}_i)$ and $\widetilde{A}_i = Q_i \Delta_i Q_i^T$, choose

$$P_i = Q_i \Delta_i^{\dagger/2} \wedge p_i = -\widetilde{A}_i^{\dagger} \widetilde{b}_i$$

(F_0 is not diagonal, only the constraints)

Existence of rank 1 solutions

Because P is a nonsingular matrix, we can transform the system

- $\lambda_i \in \mathbb{R}$, $H_0 + \lambda_1 H_1 + \ldots + \lambda_m H_m \succeq 0$
- $Z \succeq 0$, $\mathrm{tr}(H_0 Z) < 0 \wedge \mathrm{tr}(H_i Z) = 0$, $i \geq 1$

into a new system of alternatives that presents diagonal matrices

- $\lambda_i \in \mathbb{R}$, $F_0 + \lambda_1 F_1 + \ldots + \lambda_m F_m \succeq 0$
- $Y \succeq 0$, $\mathrm{tr}(F_0 Y) < 0 \wedge \mathrm{tr}(F_i Y) = 0$, $i \geq 1$

where we made the change of variables $Y = P^{-1} Z P^{-T}$.

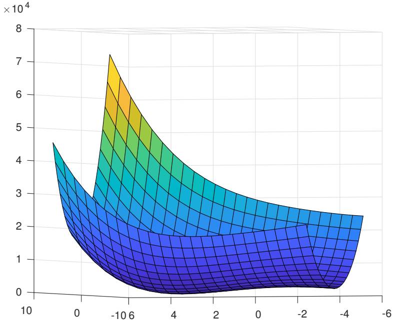
Back to the QP:

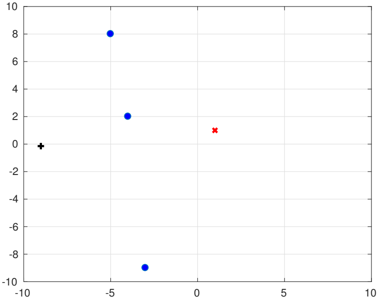
- Assume Y has rank R:
$$Y = VV^T$$
where $V = [v_1, v_2, \ldots, v_R]$.
- We can find a vector y such that
$$y^T F_0 y \leq \mathrm{tr}(F_0 Y)$$
$$y^T F_i y = \mathrm{tr}(F_i Y) \quad \forall i \geq 1.$$

  if F_0 is a Z-matrix.

Distributed Localization Problem

- We already saw an **optimal** algorithm for the localization problem (solving the dual)
- Every descent method that solves the primal problem, may get stuck into a local optimal point.





Distributed Localization Problem

$$\underset{\mathbf{x}}{\text{minimize}} \quad \sum_i (\|\mathbf{x} - s_i\|^2 - d_i^2)^2$$

Introduce variables t_i, $\mathbf{x} = x_i = z$ and rewrite constraints as

$$\|x_i - z\| - t_i = 0 \quad \forall i \geq 1 \tag{8}$$

Solve individual instances

$$\begin{aligned} \underset{x_i, z}{\text{minimize}} \quad & (t_i - d_i^2)^2 + \mu_i^T (x_i - z) + \rho \|x_i - z\|^2 \\ \text{subject to} \quad & \|x_i - z\| - t_i = 0. \end{aligned}$$

and iterate over the dual ascent method:

$$\mu_i^{k+1} = \mu_i^{k+1} + \alpha(x_i^k - z^k) \tag{9}$$

The dual method will force consensus and the algorithm will converge to the optimal point.