On Optimization Problems with Cardinality Constraints

Christian Kanzow

Joint work with
Oleg P. Burdakov, Michal Červinka, and Alexandra Schwartz

International Conference on Complementarity Problems, Berlin 2014

Outline

- Cardinality constrained optimization problems
- Reformulations
- Relation between local and global minima
- Problem-tailored constraint qualifications
- Stationarity conditions
- A regularization method
- Numerical results
- Comparison with MPCCs
- Conclusions and outline

Cardinality Constrained Optimization Problems

Cardinality constrained optimization problem:

$$\min_x \; f(x) \quad \text{s.t.} \quad x \in X, \|x\|_0 \leq \kappa$$

with $X \subseteq \mathbb{R}^n$ described by some standard constraints

$$X := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0 \ (i = 1, \ldots, m), \ h_i(x) = 0 \ (i = 1, \ldots, p)\}$$

and

$$\|x\|_0 := \text{number of nonzero components of the vector x.}$$

Functions $f, g_i, h_i : \mathbb{R}^n \to \mathbb{R}$ are assumed to be continuously differentiable, and parameter $\kappa < n$.

Cardinality Constrained Optimization Problems

Cardinality constrained optimization problem:

$$\min_x \; f(x) \quad \text{s.t.} \quad x \in X, \|x\|_0 \leq \kappa$$

with $X \subseteq \mathbb{R}^n$ described by some standard constraints

$$X := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0 \ (i = 1, \ldots, m), \ h_i(x) = 0 \ (i = 1, \ldots, p)\}$$

and

$$\|x\|_0 := \text{number of nonzero components of the vector x.}$$

Functions $f, g_i, h_i : \mathbb{R}^n \to \mathbb{R}$ are assumed to be continuously differentiable, and parameter $\kappa < n$.

Example: Portfolio Optimization

Portfolio selection problem:

$$\min_x x^T Q x \quad \text{s.t.} \quad \begin{aligned} &\mu^T x \geq \rho, \\ &e^T x \leq 1, \\ &0 \leq x_i \leq u_i \quad \forall i = 1, \ldots, n, \\ &\|x\|_0 \leq \kappa. \end{aligned}$$

Q and μ are the covariance matrix and mean of n possible assets and $e^T x \leq 1$ is a resource constraint.

Existing Literature

First (?) paper:

- D. Bienstock: *Computational study of a family of mixed-integer quadratic programming problems.* Mathematical Programming 74, 1996, pp. 121–140.

Many subsequent papers including

- Bertsimas and Shioda (2009), Di Lorenzo et al. (2012), Frangioni and Gentile (2007), Murray and Shek (2012), Ruiz-Torrubiano et al. (2010), Sun et al. (2013), Zheng et al. (2013).

Existing Literature

First (?) paper:

- D. Bienstock: *Computational study of a family of mixed-integer quadratic programming problems.* Mathematical Programming 74, 1996, pp. 121–140.

Many subsequent papers including

- Bertsimas and Shioda (2009), Di Lorenzo et al. (2012), Frangioni and Gentile (2007), Murray and Shek (2012), Ruiz-Torrubiano et al. (2010), Sun et al. (2013), Zheng et al. (2013).

Cardinality constrained problems related to sparse optimization. Many more papers there, particularly relevant for our approach is

- M. Feng, J.E. Mitchell, J.-S. Pang, X. Shen, and A. Wächter: *Complementarity formulation of ℓ_0-norm optimization problems.* Technical Report, September 2013.

Existing Literature

First (?) paper:

- D. Bienstock: *Computational study of a family of mixed-integer quadratic programming problems.* Mathematical Programming 74, 1996, pp. 121–140.

Many subsequent papers including

- Bertsimas and Shioda (2009), Di Lorenzo et al. (2012), Frangioni and Gentile (2007), Murray and Shek (2012), Ruiz-Torrubiano et al. (2010), Sun et al. (2013), Zheng et al. (2013).

Cardinality constrained problems related to sparse optimization. Many more papers there, particularly relevant for our approach is

- M. Feng, J.E. Mitchell, J.-S. Pang, X. Shen, and A. Wächter: *Complementarity formulation of ℓ_0-norm optimization problems.* Technical Report, September 2013.

Reformulations

1) Cardinality constrained optimization problem:

$$\min_{x} \; f(x) \quad \text{s.t.} \quad x \in X, \; \|x\|_0 \le \kappa \; .$$

2) Mixed integer program:

$$\begin{aligned} \min_{x,y} \; f(x) \quad \text{s.t.} \quad & x \in X \\ & e^T y \ge n - \kappa, \\ & x_i y_i = 0 \quad \forall i = 1, \ldots, n, \\ & y_i \in \{0, 1\} \quad \forall i = 1, \ldots, n \; . \end{aligned}$$

Reformulations

1) Cardinality constrained optimization problem:

$$\min_x \; f(x) \quad \text{s.t.} \quad x \in X, \; \|x\|_0 \le \kappa \; .$$

2) Mixed integer program:

$$\begin{aligned} \min_{x,y} \; f(x) \quad \text{s.t.} \quad & x \in X \\ & e^T y \ge n - \kappa, \\ & x_i y_i = 0 \quad \forall i = 1, \ldots, n, \\ & y_i \in \{0, 1\} \quad \forall i = 1, \ldots, n \; . \end{aligned}$$

3) Relaxation of mixed integer program:

$$\begin{aligned} \min_{x,y} \; f(x) \quad \text{s.t.} \quad & x \in X \\ & e^T y \ge n - \kappa, \\ & x_i y_i = 0 \quad \forall i = 1, \ldots, n, \\ & 0 \le y_i \le 1 \quad \forall i = 1, \ldots, n. \end{aligned}$$

Reformulations

1) Cardinality constrained optimization problem:

$$\min_{x} \; f(x) \quad \text{s.t.} \quad x \in X, \; \|x\|_0 \leq \kappa \; .$$

2) Mixed integer program:

$$\begin{aligned} \min_{x,y} \; f(x) \quad \text{s.t.} \quad & x \in X \\ & e^T y \geq n - \kappa, \\ & x_i y_i = 0 \quad \forall i = 1, \ldots, n, \\ & y_i \in \{0, 1\} \quad \forall i = 1, \ldots, n \; . \end{aligned}$$

3) Relaxation of mixed integer program:

$$\begin{aligned} \min_{x,y} \; f(x) \quad \text{s.t.} \quad & x \in X \\ & e^T y \geq n - \kappa, \\ & x_i y_i = 0 \quad \forall i = 1, \ldots, n, \\ & 0 \leq y_i \leq 1 \quad \forall i = 1, \ldots, n. \end{aligned}$$

Relation between Global Minima

Theorem

The following statements are equivalent:

(a) x^* *is a solution (=global minimum) of the cardinality constrained optimization problem.*

(b) *There exists a vector* y^* *such that* (x^*, y^*) *is a solution of the mixed-integer problem.*

(c) *There exists a vector* $y^* \in \mathbb{R}^n$ *such that* (x^*, y^*) *is a solution of the relaxed problem.*

Relation between Local Minima

Theorem

(a) *If x^* is a local minimum of the cardinality constrained optimization problem, then there exists a vector y^* such that (x^*, y^*) is a local minimum of the relaxed problem.*

(b) *If (x^*, y^*) is a local minimizer of the relaxed problem satisfying $\|x^*\|_0 = \kappa$, then x^* is a local minimum of the cardinality constrained problem.*

(c) *If (x^*, y^*) is a local minimum of the relaxed problem, then $\|x^*\|_0 = \kappa$ holds if and only if y^* is unique, i.e. if there is exactly one y^* such that (x^*, y^*) is a local minimum of the relaxed program. In this case, the components of y^* are binary.*

Some Tangent Cones

Let Z denote the feasible set of the relaxed program, and let $(x^*, y^*) \in Z$ be any feasible point. Define the following three cones:

$$
\begin{aligned}
\mathcal{T}_Z(x^*, y^*) &:= \text{standard (Bouligand) tangent cone of } Z \text{ at } (x^*, y^*), \\
\mathcal{L}_Z(x^*, y^*) &:= \text{standard linearization cone of } Z \text{ at } (x^*, y^*), \\
\mathcal{L}_Z^{CC}(x^*, y^*) &:= \text{CC-linearization cone of } Z \text{ at } (x^*, y^*) \\
&:= \big\{(d_x, d_y) \mid (d_x, d_y) \in \mathcal{L}_Z(x^*, y^*) \text{ and} \\
&\qquad (e_i^T d_x)(e_i^T d_y) = 0 \;\; \forall i \in I_{00}(x^*, y^*)\big\},
\end{aligned}
$$

where

$$
I_{00}(x^*, y^*) := \{i \in \{1, \ldots, n\} \mid x_i^* = 0, y_i^* = 0\}.
$$

Problem-tailored Constraint Qualifications

Theorem

It holds that $\mathcal{T}_Z(x^*, y^*) \subseteq \mathcal{L}_Z^{CC}(x^*, y^*) \subseteq \mathcal{L}_Z(x^*, y^*)$.

Definition

We say that

(a) CC-ACQ holds at (x^*, y^*) if $\mathcal{T}_Z(x^*, y^*) = \mathcal{L}_Z^{CC}(x^*, y^*)$.

(b) CC-GCQ holds at (x^*, y^*) if $\mathcal{T}_Z(x^*, y^*)^\circ = \mathcal{L}_Z^{CC}(x^*, y^*)^\circ$.

Problem-tailored Constraint Qualifications

Theorem

It holds that $\mathcal{T}_Z(x^*, y^*) \subseteq \mathcal{L}_Z^{CC}(x^*, y^*) \subseteq \mathcal{L}_Z(x^*, y^*)$.

Definition

We say that

(a) CC-ACQ holds at (x^*, y^*) if $\mathcal{T}_Z(x^*, y^*) = \mathcal{L}_Z^{CC}(x^*, y^*)$.

(b) CC-GCQ holds at (x^*, y^*) if $\mathcal{T}_Z(x^*, y^*)^\circ = \mathcal{L}_Z^{CC}(x^*, y^*)^\circ$.

Remark

CC-ACQ holds, in particular, if all constraints g_i, h_i are affine.

Problem-tailored Constraint Qualifications

Theorem

It holds that $\mathcal{T}_Z(x^*, y^*) \subseteq \mathcal{L}_Z^{CC}(x^*, y^*) \subseteq \mathcal{L}_Z(x^*, y^*)$.

Definition

We say that

(a) CC-ACQ holds at (x^*, y^*) if $\mathcal{T}_Z(x^*, y^*) = \mathcal{L}_Z^{CC}(x^*, y^*)$.

(b) CC-GCQ holds at (x^*, y^*) if $\mathcal{T}_Z(x^*, y^*)^\circ = \mathcal{L}_Z^{CC}(x^*, y^*)^\circ$.

Remark

CC-ACQ holds, in particular, if all constraints g_i, h_i are affine.

Stationarity Conditions

Definition

A feasible point $(x^*, y^*) \in Z$ of the relaxed program is called

(a) S-stationary if there exist multipliers $\lambda_i, \mu_i, \gamma_i$ such that

$$\nabla f(x^*) + \sum_{i \in I_g(x^*)} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i: y_i^* \neq 0} \gamma_i e_i = 0,$$

$$\lambda_i \geq 0 \quad \forall i \in I_g(x^*);$$

(b) M-stationary if there exist multipliers $\lambda_i, \mu_i, \gamma_i$ such that

$$\nabla f(x^*) + \sum_{i \in I_g(x^*)} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) + \sum_{i: x_i^* = 0} \gamma_i e_i = 0,$$

$$\lambda_i \geq 0 \quad \forall i \in I_g(x^*).$$

Optimality Conditions

Theorem

Let $(x^*, y^*) \in Z$ be feasible for the relaxed problem. Then CC-GCQ holds in (x^*, y^*) if and only if GCQ holds there.

Theorem

Let (x^*, y^*) be a local minimum of the relaxed program such that CC-GCQ holds at (x^*, y^*). Then (x^*, y^*) is an S-stationary point.

Optimality Conditions

Theorem

Let $(x^*, y^*) \in Z$ be feasible for the relaxed problem. Then CC-GCQ holds in (x^*, y^*) if and only if GCQ holds there.

Theorem

Let (x^*, y^*) be a local minimum of the relaxed program such that CC-GCQ holds at (x^*, y^*). Then (x^*, y^*) is an S-stationary point.

Corollary

Let (x^*, y^*) be a local minimum of the relaxed program, and suppose that g_i, h_i are affine mappings. Then (x^*, y^*) is an S-stationary point.

Optimality Conditions

Theorem

Let $(x^*, y^*) \in Z$ be feasible for the relaxed problem. Then CC-GCQ holds in (x^*, y^*) if and only if GCQ holds there.

Theorem

Let (x^*, y^*) be a local minimum of the relaxed program such that CC-GCQ holds at (x^*, y^*). Then (x^*, y^*) is an S-stationary point.

Corollary

Let (x^*, y^*) be a local minimum of the relaxed program, and suppose that g_i, h_i are affine mappings. Then (x^*, y^*) is an S-stationary point.

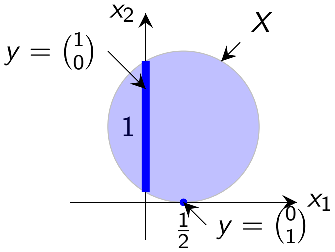
Counterexample where Solution is not S-stationary

Example

Consider the convex, but not polyhedral convex, set

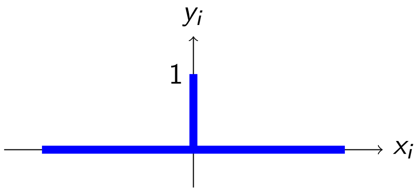
$$X := \{x \in \mathbb{R}^2 \mid (x_1 - \tfrac{1}{2})^2 + (x_2 - 1)^2 \leq 1\}$$

and $f(x) = x_1 + cx_2$ with $c > 0$. Choosing $\kappa = 1$ and c sufficiently large, $x^* = (\frac{1}{2}, 0)$, $y^* = (0, 1)$ is unique solution of the relaxed problem. But (x^*, y^*) is not a KKT point, hence GCQ is violated in (x^*, y^*).

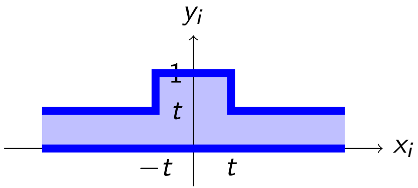


Geometric Idea of Regularization

Replace the cardinality constraints $x_i y_i = 0,\ 0 \leq y_i \leq 1$



geometrically by something like



Anlaytic Realization of Regularization

Define the functions

$$\phi(a, b; t) := \begin{cases} (a-t)(b-t) & \text{if } a+b \geq 2t, \\ -\frac{1}{2}\left[(a-t)^2 + (b-t)^2\right] & \text{if } a+b < 2t \end{cases}$$

as well as

$$\tilde{\phi}(a, b; t) := \phi(-a, b; t).$$

Remark

(a) The functions ϕ and $\tilde{\phi}$ are continuously differentiable everywhere.

(b) For $t = 0$, it holds that

$$\phi(a, b; 0) = 0 \quad \Longleftrightarrow \quad a \geq 0, b \geq 0, ab = 0,$$

i.e. ϕ is an NCP-function.

Regularization Method

Replace the cardinality constraints

$$x_i y_i = 0, \quad 0 \le y_i \le 1$$

by the inequalities

$$0 \le y_i \le 1, \ \phi(x_i, y_i; t) \le 0, \ \tilde{\phi}(x_i, y_i; t) \le 0$$

for some parameter $t > 0$ denotes yields the following regularization of the relaxed program:

$$\begin{aligned} \min_{x,y} f(x) \quad \text{s.t.} \quad & g_i(x) \le 0 \quad \forall i = 1, \dots, m, \\ & h_i(x) = 0 \quad \forall i = 1, \dots, p, \\ & e^T y \ge n - \kappa, \\ & \phi(x_i, y_i; t) \le 0 \quad \forall i = 1, \dots, n, \\ & \tilde{\phi}(x_i, y_i; t) \le 0 \quad \forall i = 1, \dots, n, \\ & 0 \le y_i \le 1 \quad \forall i = 1, \dots, n. \end{aligned}$$

Convergence Result

Theorem

Let $\{t_k\} \downarrow 0$ and $\{(x^k, y^k, \lambda^k, \mu^k, \delta^k, \tau^k, \tilde{\tau}^k, \nu^k)\}$ be a corresponding sequence of KKT points of NLP(t_k) such that $(x^k, y^k) \to (x^*, y^*)$. Assume that the limit point satisfies CC-CPLD. Then (x^*, y^*) is an M-stationary point of the relaxed program.

Theorem

Let (x^*, y^*) be feasible for the relaxed problem such that CC-CPLD is satisfied in (x^*, y^*). Then there is a $\bar{t} > 0$ and an $r > 0$ such that the following holds for all $t \in (0, \bar{t}]$: Is $(\hat{x}, \hat{y}) \in B_r(x^*) \times B_r(y^*)$ feasible for NLP(t), then standard GCQ for NLP(t) holds there.

Convergence Result

Theorem

Let $\{t_k\} \downarrow 0$ and $\{(x^k, y^k, \lambda^k, \mu^k, \delta^k, \tau^k, \tilde{\tau}^k, \nu^k)\}$ be a corresponding sequence of KKT points of NLP(t_k) such that $(x^k, y^k) \to (x^*, y^*)$. Assume that the limit point satisfies CC-CPLD. Then (x^*, y^*) is an M-stationary point of the relaxed program.

Theorem

Let (x^*, y^*) be feasible for the relaxed problem such that CC-CPLD is satisfied in (x^*, y^*). Then there is a $\bar{t} > 0$ and an $r > 0$ such that the following holds for all $t \in (0, \bar{t}]$: Is $(\hat{x}, \hat{y}) \in B_r(x^*) \times B_r(y^*)$ feasible for NLP(t), then standard GCQ for NLP(t) holds there.

Test Problems

Portfolio selection problem:

$$\min_x x^T Q x \quad \text{s.t.} \quad \begin{aligned} & \mu^T x \geq \rho, \\ & e^T x \leq 1, \\ & 0 \leq x_i \leq u_i \quad \forall i = 1, \ldots, n, \\ & \|x\|_0 \leq \kappa. \end{aligned}$$

Test examples created using the same randomly generated data Q, μ, ρ, and u which were used by Frangioni and Gentile (2007), available at their webpage

`http://www.di.unipi.it/optimize/Data/MV.html`.

We use 30 test instances for each of the three dimensions $n = 200, 300, 400$. In addition, for every example three cardinality constraints $\kappa = 5, 10, 20$ are used ($= 270$ test problems altogether).

Solvers

We use three different approaches for solving the test problems:

(a) `GUROBI`: Solves a mixed-integer formulation of the problem (used as a benchmark for our approach) (allowing approximately two hours of computation time for each test problem)

(b) Use `SNOPT` applied directly to the relaxed problem.

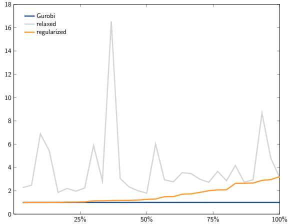
(c) Use our regularization approach with `SNOPT` applied to the regularized programs

Starting point for all three approaches:

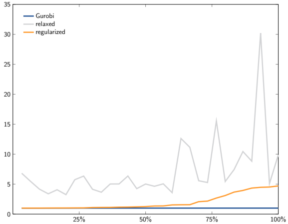
$$x^0 := (0, \ldots, 0)^T, y^0 := (1, \ldots, 1)^T.$$

The following figures present the optimal function values, normalized by the one found by GUROBI, and in increasing order for the regularization approach.

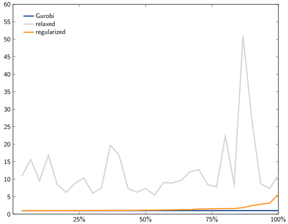
Numerical Results (Part 1)



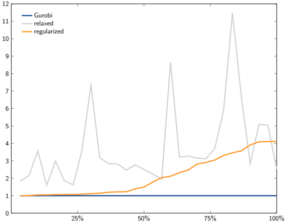
(a) $n = 200, \kappa = 5$



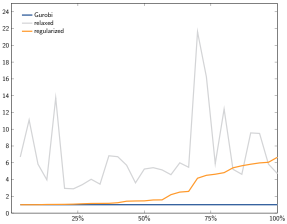
(b) $n = 200, \kappa = 10$



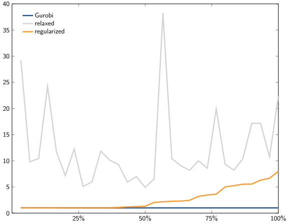
(c) $n = 200, \kappa = 20$



(d) $n = 300, \kappa = 5$

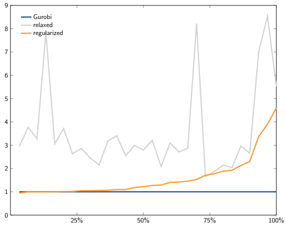


(e) $n = 300, \kappa = 10$

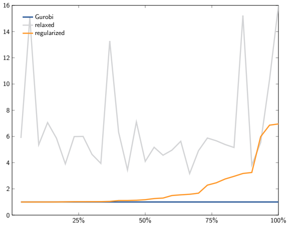


(f) $n = 300, \kappa = 20$

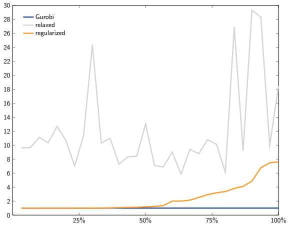
Numerical Results (Part 3)



(g) $n = 400, \kappa = 5$



(h) $n = 400, \kappa = 10$



(i) $n = 400, \kappa = 20$

MPCC-Formulation of a Class of Cardinality Constraints

Consider cardinality constrained problem with nonnegativity constraints:

$$\begin{aligned} \min_{x,y}\ f(x) \quad \text{s.t.} \quad & g_i(x) \le 0,\ x_i \ge 0 \quad \forall i = 1, \dots, m, \\ & h_i(x) = 0 \quad \forall i = 1, \dots, p, \\ & e^T y \ge n - \kappa, \\ & x_i y_i = 0 \quad \forall i = 1, \dots, n, \\ & 0 \le y_i \le 1 \quad \forall i = 1, \dots, n, \end{aligned}$$

Moving the nonnegativity constraints to the cardinality constraints yields the following MPCC:

$$\begin{aligned} \min_{x,y}\ f(x) \quad \text{s.t.} \quad & g_i(x) \le 0 \quad \forall i = 1, \dots, m, \\ & h_i(x) = 0 \quad \forall i = 1, \dots, p, \\ & e^T y \ge n - \kappa, \\ & x_i y_i = 0 \quad \forall i = 1, \dots, n, \\ & y_i \le 1 \quad \forall i = 1, \dots, n, \\ & x_i \ge 0,\ y_i \ge 0,\ x_i y_i = 0 \quad \forall i = 1, \dots, n. \end{aligned}$$

Relation between MPCC and CC-Problems

Remark

(a) (x^*, y^*) is S-stationary in the sense of MPCCs if and only if it is S-stationary in the sense of cardinality constrained problems.

(b) (x^*, y^*) is M-stationary in the sense of MPCCs if and only if it is M-stationary in the sense of cardinality constrained problems.

(c) For general MPCCs, M-, C-, and W-stationarity are different stationarity concepts, but for the MPCC arising from cardinality constrained problems, these concepts coincide.

(d) For general MPCCs, S-stationarity may not hold for affine functions g_i, h_i, whereas S-stationarity holds in this case for CC-problems.

(e) MPCC-LICQ implies a *piecewise LICQ* for general MPCCs, whereas a corresponding observation is not true for CC-problems

(f) MPCC-LICQ and MPCC-MFCQ are likely to be violated at a solution $(x^*, y^)$ of the cardinality constrained problem.

Conclusions and Outline

- We reformulated the cardinality constrained problem as an optimization problem in continuous variables.
- This allows application of results and techniques from continuous optimization to obtain, e.g., optimality conditions.
- A specialized analysis is necessary in order to take into account the particular structure of cardinality constrained problems.
- Cardinality constrained optimization problems have different properties than MPCCs and should therefore be treated separately.
- Other solution methods are possible, but Scholtes regularization, for example, seems to cause some troubles.
- Similar results seem to hold for sparse optimization problems.
- Application of ideas from mixed integer problems principally possible.

Many thanks for your attention!