Convex Optimization II

Lecture 17: Convex-Cardinality Problems

Hamed Shah-Mansouri

Department of Electrical Engineering

Sharif University of Technology

1400-2

Motivations

In several applications, we encounter the cardinality (i.e., the number of non-zero) of decision variables. Applications include

- sparse design
- robust estimation in statistics
- classification
- support vector machine (SVM) in machine learning
- compressed sensing

Thanks to Professor John Duchi, Stanford University, for the slides.

OUTLINE

- Convex-Cardinality Problems
- Solution
- Application
- Convex Relaxation

Cardinality

Cardinality: denoted by $\mathbf{card}(x)$ for $x \in \mathbb{R}^n$, is the number of non-zero components of x.

Arises in many applications including

- sparse design
- robust estimation in statistics
- Classification
- support vector machine (SVM) in machine learning
- compressed sensing

CONVEX-CARDINALITY PROBLEMS

A **convex-cardinality problem** is one that would be convex, except for appearance of **card** in the objective or constraints.

- convex minimum cardinality problem (sparse design)

$$\begin{array}{ll} \text{minimize} & \mathbf{card}(x) \\ \text{subject to} & x \in \mathcal{C} \end{array}$$

- convex problem with cardinality constraint

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & \mathbf{card}(x) \leq k \\ & x \in \mathcal{C}. \end{array}$$

Solving Convex-Cardinality Problems

- if we fix the sparsity pattern of x (i.e., which entries are zero/nonzero) we get a convex problem.
- by solving 2^n convex problems associated with all possible sparsity patterns, we can solve convex-cardinality problem.
- convex-cardinality problem is NP–hard in general

SOLUTION: ℓ_1-NORM HEURISTIC

Replace $\mathbf{card}(z)$ with $\delta||z||_1$, or add regularization term $\delta||z||_1$ to the objective.

Convex minimum cardinality problem

$$\begin{array}{ll} \text{minimize} & \mathbf{card}(x) \\ \text{subject to} & x \in \mathcal{C} \end{array}$$

replace by

$$\begin{array}{ll} \text{minimize} & ||x||_1 \\ \text{subject to} & x \in \mathcal{C} \end{array}$$

SOLUTION: ℓ_1-NORM HEURISTIC

convex problem with cardinality constraint

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & \mathbf{card}(x) \leq k \\ & x \in \mathcal{C}. \end{array}$$

replace by ℓ_1-constrained problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & ||x||_1 \leq \beta \\ & x \in \mathcal{C}. \end{array}$$

or ℓ_1-regularized problem

$$\begin{array}{ll} \text{minimize} & f(x) + \delta ||x||_1 \\ \text{subject to} & x \in \mathcal{C}. \end{array}$$

ILP AS CONVEX-CARDINALITY PROBLEMS

ILP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \\ & x \in \{0,1\}. \end{array}$$

can be transformed into

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \\ & \mathbf{card}(x) + \mathbf{card}(1-x) \leq n. \end{array}$$

where $x \in \mathbb{R}^n$.

Application: Sparse Design

Convex minimum cardinality problem

$$\begin{array}{ll} \text{minimize} & \mathbf{card}(x) \\ \text{subject to} & x \in \mathcal{C} \end{array}$$

Application: Sparse modeling / Regressor selection

Fit vector $b \in \mathbb{R}^m$ as a linear combination of k regressors (chosen from n possible regressors)

$$\begin{array}{ll} \text{minimize} & ||Ax - b||_2 \\ \text{subject to} & \mathbf{card}(x) \leq k \end{array}$$

choose subset of k regressors that (together) best fit or explain b

Variations:

$$\begin{array}{ll} \text{minimize} & \mathbf{card}(x) \\ \text{subject to} & ||Ax - b||_2 \leq \epsilon \end{array}$$

$$\text{minimize} \quad ||Ax - b||_2 + \delta\mathbf{card}(x)$$

Application: Sparse Signal Reconstruction

Estimate signal x, given

- noisy measurement $y = Ax + v$, v is unknown noise.
- prior information $\mathbf{card}(x) \leq k$

Maximum likelihood estimate is the solution of

$$\begin{array}{ll} \text{minimize} & ||Ax - y||_2 \\ \text{subject to} & \mathbf{card}(x) \leq k \end{array}$$

Application: Minimum Number of Violations

- Consider set of convex inequalities

$$f_1(x) \leq 0, \ldots, f_m(x) \leq 0$$

- Choose x to minimize the number of violated inequalities

$$\begin{array}{ll} \text{minimize} & \mathbf{card}(t) \\ \text{subject to} & f_i(x) \leq t_i, \text{for } i = 1, \ldots, m \end{array}$$

- Determining whether zero inequalities can be violated is (easy) convex feasibility problem.

Application: Linear classifier with fewest errors

- given data $(x_1, y_1), \ldots, (x_m, y_m) \in \mathbb{R}^n \times \{-1, 1\}$
- we seek linear (affine) classifier $y \approx \textbf{sign}(w^T x + v)$
- classification error corresponds to $y_i(w^T x_i + v) \leq 0$
- to find w and v that give fewest classification errors:

$$\begin{array}{ll} \text{minimize} & \textbf{card}(t) \\ \text{subject to} & y_i(w^T x_i + v) + t_i \geq 0, \text{ for } i = 1, \ldots, m \end{array}$$

Interpretation as Convex Relaxation

$$\begin{array}{ll} \text{minimize} & \textbf{card}(x) \\ \text{subject to} & x \in \mathcal{C} \\ & ||x||_\infty \leq R \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T z \\ \text{subject to} & |x_i| \leq R z_i, \quad \text{for } i = 1, \ldots, n \\ & x \in \mathcal{C} \\ & z_i \in \{0, 1\}, \quad \text{for } i = 1, \ldots, n. \end{array}$$

INTERPRETATION AS CONVEX RELAXATION

Now, relax binary variables

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T z \\ \text{subject to} & |x_i| \leq R z_i, \quad \text{for } i = 1, \ldots, n \\ & x \in \mathcal{C} \\ & 0 \leq z_i \leq 1, \quad \text{for } i = 1, \ldots, n. \end{array}$$

which is equivalent to

$$\begin{array}{ll} \text{minimize} & ||x||_1 \\ \text{subject to} & x \in \mathcal{C} \\ & ||x||_\infty \leq R. \end{array}$$