Convex Optimization II

Lecture 16: Regularization and Convexification

Hamed Shah-Mansouri

Department of Electrical Engineering

Sharif University of Technology

1400-2

Motivations

- Regularization is a common approach to solve an ill-posed problem–one whose solution is not unique or is acutely sensitive to data perturbations. Regularization techniques construct a related problem whose solution is well behaved and deviates only slightly from a solution of the original problem.
- It also provides a platform to choose a solution with a desired property among many solutions of the problem.
- Regularization helps find sparse solutions with applications in signal processing, machine learning, and statistics.

[1] M.P. Friedlander and P. Tseng, "Exact regularization of convex programs", *SIAM Journal on Optimization*. vol. 18, no. 4, pp.1326–1350, Nov. 2007.

[2] D.P. Bertsekas, "Convexification procedures and decomposition methods for non–convex optimization problems", *Journal of Optimization Theory and Applications*. Vol. 9, no. 2, pp. 169–197, Oct.1979.

Outline

- Regularization
- Penalization
- Sparse Design
- Convexification

Regularization

- **Regularization**: Construct a related problem whose solution is well behaved and has desired properties.
- **Deviations**: Deviations from solutions of the original problem are generally accepted as a trade-off for obtaining solutions with other desirable properties.
- **Exact Regularization**: However, it would be more desirable if solutions of the regularized problem are also solutions of the original problem.

Regularization of Convex Programs

Consider the general convex program

$$\mathbf{P}: \quad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in \mathcal{C}. \end{array}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a convex function, $\mathcal{C} \subset \mathbb{R}^n$ is a nonempty closed convex set.

A popular technique is to regularize by adding a convex function to the objective with a nonnegative regularization parameter δ.

$$\mathbf{P}_\delta: \quad \begin{array}{ll} \text{minimize} & f(x) + \delta\phi(x) \\ \text{subject to} & x \in \mathcal{C}. \end{array}$$

The regularization is exact if the solutions of Problem $\mathbf{P}_\delta$ are also solutions of Problem $\mathbf{P}$ for all values of δ below some positive threshold value.

Regularization of Convex Programs

A related convex program that selects solutions of **P** of least ϕ-value

$$\mathbf{P}^{\phi}: \qquad \begin{array}{ll} \text{minimize} & \phi(x) \\ \text{subject to} & f(x) \leq p^* \\ & x \in \mathcal{C}. \end{array}$$

where p^* denotes the optimal value of Problem **P**.

Any solution of Problem $\mathbf{P}^{\phi}$ is also a solution of P. The converse, however, does not generally hold.

Regularization of Convex Programs

Let $\mathcal{S}$, $\mathcal{S}_\delta$, and $\mathcal{S}^\phi$ denote the solution sets of Problems $\mathbf{P}$, $\mathbf{P}_\delta$, and $\mathbf{P}^\phi$, respectively.

Theorem

1. For any $\delta > 0$, $\mathcal{S} \cap \mathcal{S}_\delta \subset \mathcal{S}^\phi$.
2. If there exists a Lagrange multiplier μ^* for Problem $\mathbf{P}^\phi$, then $\mathcal{S} \cap \mathcal{S}_\delta = \mathcal{S}^\phi$ for all $\delta \in (0, 1/\mu^*]$.

The theorem says the regularization $\mathbf{P}_\delta$ is exact if and only if the selection problem $\mathbf{P}^\phi$ has a Lagrange multiplier μ^*. Moreover, $\mathcal{S}_\delta = \mathcal{S}$ for all $\delta < 1/\mu^*$.

Penalization

Consider the convex program

$$\begin{array}{lll} \mathbf{Q} & \text{minimize} & \phi(x) \\ & \text{subject to} & g_i(x) \leq 0, \text{for } i = 1, \ldots, m \\ & & x \in \mathcal{C}. \end{array}$$

where ϕ, $\mathbf{g} = (g_1, \ldots, g_m)$ are real-valued convex functions defined on $\mathbb{R}^n$, and $\mathcal{C} \in \mathbb{R}^n$ is a nonempty closed convex set.

The penalized form of above problem with positive penalty parameter is

$$\begin{array}{lll} \mathbf{Q}^P & \text{minimize} & \phi(x) + \sigma P(\mathbf{g}(x)) \\ & \text{subject to} & x \in \mathcal{C}. \end{array}$$

where $P : \mathbb{R}^m \to [0, \infty)$ is a non-negative convex function having the property that $P(u) = 0$ if and only if $u \leq 0$.

Exact Penalization

Theorem

Suppose that Problem $\mathbf{Q}$ has a nonempty compact solution set. If there exist Lagrange multipliers μ^* for $\mathbf{Q}$, then the penalized problem $\mathbf{Q}^P$ has the same solution set as $\mathbf{Q}$ for all $\sigma > w(\mu^*)$[1].

[1] Refer to [1] for more details on function w.

Error Bounds

Even when the exact regularization cannot be achieved, we can still estimate the distance from $\mathcal{S}_\sigma$ to $\mathcal{S}$ in terms of σ and the growth rate of f. Refer to [1] for more details.

Example: Sparse Solutions of LP

An LP has many optimal solutions. In some applications, we may be interested in sparse or even the sparsest solutions.

- Consider a system of linear equations $Ax = b$
- The sparsest solution can be obtained via

$$\begin{array}{ll} \text{minimize} & ||x||_0 \\ \text{subject to} & Ax = b. \end{array}$$

  which is computationally intractable. This can construct non-convex regularization.

- A sparse solution can be obtained via

$$\begin{array}{ll} \text{minimize} & ||x||_1 \\ \text{subject to} & Ax = b. \end{array}$$

  which can be transformed into LP.

Example: Sparse Solutions of LP

Consider a general LP problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \end{array}$$

To find a sparse solution, we use the ℓ_1-regularization

$$\begin{array}{ll} \text{minimize} & c^T x + \delta ||x||_1 \\ \text{subject to} & Ax = b \end{array}$$

First solve the equivalent $\mathbf{P}^{\phi}$ to find μ^*, then choose $\delta \in (0, 1/\mu^*)$.

Non-convex Problems

Consider the following possibly non-convex problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \leq 0 \\ & h(x) = 0 \\ & x \in \mathbb{R}^n. \end{array}$$

The problem cannot be solved using primal-dual approach if it is not convex.

It is either impossible to define a dual function or the maximal value of the dual function is not equal to the optimal value of the primal problem.

Question: Is there any way to convert the problem into a convex one at least in a local area?

Convexification

We consider the problem

$$
\begin{array}{lll}
\mathbf{C} & \text{minimize} & f(x) + \delta||y - x||_2^2 \\
& \text{subject to} & g(x) \leq 0 \\
& & h(x) = 0 \\
& & x, y \in \mathbb{R}^n.
\end{array}
$$

where $\delta > 0$ is some fixed scalar and y represents a vector of additional variables.

Clearly, a vector x^* is a local minimum of the original problem **P** if and only if (x^*, x^*) is a local minimum of Problem **C**.

Problem **C** has a locally convex structure for δ large enough provided suitable second-order sufficiency conditions are satisfied at x^*.

Thus, Problem **C** may be solved by primal-dual methods. Moreover, if the original problem is separable, so is the covexified problem.