ORIE 6326: Convex Optimization

Algorithms for convex optimization

Professor Udell

Operations Research and Information Engineering
Cornell

March 21, 2017

Outline

- gradient descent
- convex optimization as unconstrained minimization
- room for improvement

Challenge

convex problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 0, \ldots, m$, convex functions

how to solve it?

Challenge

convex problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 0, \ldots, m$, convex functions

how to solve it?

- idea: gradient descent
- difficulty: not differentiable unconstrained minimization problem

Idea 1: get rid of equality constraints

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- find an x^0 such that $Ax^0 - b$
- compute basis V for null space $\mathcal{N}(A)$

$$x \in \mathcal{N}(A) \quad \Longleftrightarrow \quad x = Vz$$

  (takes $O(n^3)$ time using QR factorization)
- solve

$$\begin{array}{ll} \text{minimize} & f_0(x^0 + Vz) \\ \text{subject to} & f_i(x^0 + Vz) \le 0, \quad i = 1, \ldots, m \end{array}$$

Idea 1: get rid of equality constraints

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
$$

- find an x^0 such that $Ax^0 - b$
- compute basis V for null space $\mathcal{N}(A)$

$$x \in \mathcal{N}(A) \quad \Longleftrightarrow \quad x = Vz$$

  (takes $O(n^3)$ time using QR factorization)
- solve

$$
\begin{array}{ll}
\text{minimize} & f_0(x^0 + Vz) \\
\text{subject to} & f_i(x^0 + Vz) \leq 0, \quad i = 1, \ldots, m
\end{array}
$$

- achievement: no equality constraints
- remaining difficulty?

Idea 1: get rid of equality constraints

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- find an x^0 such that $Ax^0 - b$
- compute basis V for null space $\mathcal{N}(A)$

  $$x \in \mathcal{N}(A) \quad \iff \quad x = Vz$$

  (takes $O(n^3)$ time using QR factorization)
- solve

  $$\begin{array}{ll} \text{minimize} & f_0(x^0 + Vz) \\ \text{subject to} & f_i(x^0 + Vz) \leq 0, \quad i = 1, \ldots, m \end{array}$$

- achievement: no equality constraints
- remaining difficulty? we still have inequality constraints

Idea 2: Put the constraints in the objective

$$\begin{array}{ll} \text{minimize} & f_0(x^0 + Vz) \\ \text{subject to} & f_i(x^0 + Vz) \leq 0, \quad i = 1, \ldots, m \end{array}$$

- rewrite problem using the indicator of the positive reals

$$I_+(y) = \begin{cases} 0 & y \geq 0 \\ \infty & \text{otherwise} \end{cases}$$

as

$$\text{minimize} \quad f_0(x^0 + Vz) + \sum_{i=1}^{m} I_+(-f_i(x^0 + Vz))$$

Idea 2: Put the constraints in the objective

$$\begin{array}{ll} \text{minimize} & f_0(x^0 + Vz) \\ \text{subject to} & f_i(x^0 + Vz) \leq 0, \quad i = 1, \ldots, m \end{array}$$

- rewrite problem using the indicator of the positive reals

$$I_+(y) = \begin{cases} 0 & y \geq 0 \\ \infty & \text{otherwise} \end{cases}$$

as

$$\text{minimize} \quad f_0(x^0 + Vz) + \sum_{i=1}^{m} I_+(-f_i(x^0 + Vz))$$

- achievement: no constraints
- remaining difficulty?

Idea 2: Put the constraints in the objective

$$\begin{array}{ll} \text{minimize} & f_0(x^0 + Vz) \\ \text{subject to} & f_i(x^0 + Vz) \leq 0, \quad i = 1, \ldots, m \end{array}$$

- rewrite problem using the indicator of the positive reals

$$I_+(y) = \begin{cases} 0 & y \geq 0 \\ \infty & \text{otherwise} \end{cases}$$

as

$$\text{minimize} \quad f_0(x^0 + Vz) + \sum_{i=1}^{m} I_+(-f_i(x^0 + Vz))$$

- achievement: no constraints
- remaining difficulty?

$$\nabla I_+(-f_i(x^0 + Vz)) = \begin{cases} 0 & f_i(x^0 + Vz) \leq 0 \\ ? & \text{otherwise} \end{cases}$$

Idea 3: Add a smooth barrier

$$\text{minimize} \quad c^T Vz + I_+(GVz + h + Gx^0)$$

- Pick $\mu > 0$, and solve

$$\text{minimize} \quad f_0(x^0 + Vz) - \mu \sum_{i=1}^{m} \log(-f_i(x^0 + Vz))$$

- As $\mu \to 0$,

$$-\mu \sum_{i=1}^{m} \log(-f_i(x^0 + Vz)) \to \begin{cases} 0 & z \text{ feasible} \\ \infty & \text{otherwise} \end{cases}$$

Idea 3: Add a smooth barrier

$$\text{minimize} \quad c^T Vz + I_+(GVz + h + Gx^0)$$

- Pick $\mu > 0$, and solve

$$\text{minimize} \quad f_0(x^0 + Vz) - \mu \sum_{i=1}^{m} \log(-f_i(x^0 + Vz))$$

- As $\mu \to 0$,

$$-\mu \sum_{i=1}^{m} \log(-f_i(x^0 + Vz)) \to \begin{cases} 0 & z \text{ feasible} \\ \infty & \text{otherwise} \end{cases}$$

- achievement: smooth problem with no constraints
- remaining difficulty?

Idea 3: Add a smooth barrier

$$\text{minimize} \quad c^T Vz + I_+(GVz + h + Gx^0)$$

- Pick $\mu > 0$, and solve

$$\text{minimize} \quad f_0(x^0 + Vz) - \mu \sum_{i=1}^{m} \log(-f_i(x^0 + Vz))$$

- As $\mu \to 0$,

$$-\mu \sum_{i=1}^{m} \log(-f_i(x^0 + Vz)) \to \begin{cases} 0 & z \text{ feasible} \\ \infty & \text{otherwise} \end{cases}$$

- achievement: smooth problem with no constraints
- remaining difficulty?
  - fast method for unconstrained smooth minimization
  - (and how fast to send $\mu \to 0$)

What about nondifferentiable functions?

recall: we can write just about every convex problem in conic form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Fx + g \preceq_K 0 \\ & Ax = b \end{array}$$

so we just need a function for each cone K so

$$f_K(x) \geq 0 \quad \iff \quad x \in K$$

and $-\log f_K(x)$ is convex and differentiable.

examples:

- LP cone $K_+ = \{x : x \geq 0.\ f_+(x) = x$
- SDP cone $K_{\text{SDP}} = \{X : X \succeq 0.\ f_{\text{SDP}}(X) = det(X)$

A note of veracity

reduction presented above is **not** a practical method

Interior point methods resolve these problems:

- equality constraints
  - if they're sparse, eliminating them makes problem dense
- barrier
  - problem becomes terribly conditioned as $\mu \to 0$

Optimization algorithms: a general template

- . . . problem reductions . . .

What's wrong with gradient descent?

primary difficulty:

- doesn't handle constraints

secondary difficulties (depend on problem):

- function is not smooth
- function has bounded domain
- need high accuracy
- each iteration is too slow
- too many iterations
- function is not convex
- can't compute gradient

What's wrong with gradient descent?

primary difficulty:

- **Interior point.** Handles constraints

secondary difficulties (depend on problem):

- **Subgradient.** Generalizes gradient for non-smooth functions
- **Proximal and conjugate maps.** Generalizes gradient for functions with bounded domain
- **Newton methods.** Give high accuracy (even for ill-conditioned problems)
- **Stochastic and parallel methods.** Speed up each iteration
- **Acceleration.** Reduces number of iterations
- **Branch and bound.** Finds global minimum of non-convex function
- **Bayesian optimization.** Doesn't require gradient

References

- Lieven Vandenberghe, UCLA EE236C
- Wotao Yin, UCLA
- Stephen Boyd and John Duchi, Stanford EE364b