CS885 Reinforcement Learning
Lecture 14c: June 15, 2018

Trust Region Methods

[Nocedal and Wright, Chapter 4]

Optimization in ML

- It is common to formulate ML problems as optimization problems.
  – Min squared error
  – Min cross entropy
  – Max log likelihood
  – Max discounted sum of rewards

Two important classes

- Line search methods
  - Find a direction of improvement
  - Select a step length

- Trust region methods
  - Select a trust region (analog to max step length)
  - Find a point of improvement in the region

Trust Region Methods

- Idea:
  - Approximate objective f with a simpler objective $\tilde{f}$
  - Solve $\tilde{x}^* = argmin_x \tilde{f}(x)$

- **Problem:** The optimum $\tilde{x}^*$ might be in a region where $\tilde{f}$ poorly approximates f and therefore $\tilde{x}^*$ might be far from optimal
- **Solution:** restrict the search to a region where we trust $\tilde{f}$ to approximate f well.
  - Solve $\tilde{x}^* = argmin_{x \in trustRegion} \tilde{f}(x)$

Example

- $\tilde{f}$ often chosen to be a quadratic approximation of f

$$f(x) \approx \tilde{f}(x)$$
$$= \mathrm{f(c)} + \nabla f(c)^T (x-c) + \frac{1}{2!}(x-c)^T H(c)(x-c)$$

where ∇f is the gradient and H is the hessian

- Trust region often chosen to be a hypersphere

$$||x-c||_2 \leq \delta$$

Generic Algorithm

trustRegionMethod

Initialize δ, x_0^* and $n = 0$

Repeat

$n \leftarrow n + 1$

Solve $x_n^* = argmin_x \tilde{f}(x)$ subject to $\left|\left|x - x_{n-1}^*\right|\right|_2 \leq \delta$

If $\tilde{f}(x_n^*) \approx f(x_n^*)$ then increase δ

else decrease δ

Until convergence

Trust Region Subproblem

- $\tilde{f}$ often chosen to be a quadratic approximation of f

$$\min_x \mathrm{f(c)} + \nabla f(c)^T(x-c) + \frac{1}{2!}(x-c)^T H(c)(x-c)$$

subject to $||x-c||_2 \leq \delta$

- When H is positive semi-definite
  - Convex optimization
  - Simple and globally optimal solution
- When H is not positive semi-definite
  - Non-convex optimization
  - Simple heuristics that guarantee improvement