Lecture 20

Proof that backtracking converges to a point where the gradient vanishes.

Result 1: There is always a point that BT line search picks.

Assumptions of theorem and justification.
New term: Sufficient descent (vs. Sufficient decrease)

Outline:
- the function is bounded
- function values converge
- show lim a||g||^2 -> 0
- show a >= gamma*||g||
- conclude ||g|| -> 0

Quick Recall

If $\nabla_k^T g_k < 0$, then at some point, backtracking line search accepts.

Pf (are stable):

Your Qui!

Key Imagination

by Taylor thm, then must exist a first α^{th} s.t. the point also hold.

Then all $\alpha \in [0, \alpha^*]$ are accepted

=> will quickly test one!

Theorem

- $f: \mathbb{R}^n \to \mathbb{R}$
  Cont
  Cont gradient
- x_0: given
- $x_{k+1} = x_k + \alpha_k p_k$
  where α_k is the first value in BLS sequence s.t. satisfies Sufficient decrease

i.e. $f(x_{k+1}) \le f(x_k) + \beta \alpha_k \nabla_k^T g_k$ for $0 < \beta < 1$

i.e. seq "x_k, x_{k+1}, x_{k+2}" — the set $S = \{x \mid f(x) \le f(x_0)\}$ is bounded.

$g(x)$ is Lipschitz Continuous

i.e. $\|g(x) - g(y)\| \le L \cdot \|x - y\|$

$L < \infty$

On the set S.

Quick Aside on Lipschitz

$f(x) = x^2$

$|x^2 - y^2| \le L \cdot |x - y|$

for some fixed L.

this must hold for all x, y!

If $y = 0$,

$|x^2| \le L \cdot |x|$

for some L. @ Not true!!

x^2 is Lipschitz on any bounded domain!

Eg. $|x| \le 5$.

=>

$|x^2 - y^2| \le 10 \cdot |x - y|$

Back to Algorithm

the search direction satisfy Sufficient descent

$\cos(\nabla_k, -g_k) \ge \epsilon > 0$.

$\frac{-\nabla_k^T g_k}{\|\nabla_k\| \cdot \|g_k\|}$

why?

=> $\cos(\nabla_k, -g_k) = \frac{1}{2} \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = \frac{1}{4}$

- Also, the search direction are gradient related

$\|p_k\| \ge m \|g_k\|$

and bounded

$\|p_k\| \le M$.

for n, M, μ, ϵ fixed.

Then

$\|g_k\| \to 0$.

§ Step 1 Proof

1. f is bounded below
2. $\lim_{k \to \infty} f(x_k)$ Exist
3. $\lim_{k \to \infty} \alpha_k \|g_k\|^2 \to 0$
4. BLS + Lipschitz =>
   $\alpha_k \ge \min(1, \gamma \|g_k\|^2)$
   for some fixed γ
5. 3 & 4 => $\lim_{k \to \infty} \|g_k\| \to 0$.

Step 1

- f is Cont, $g(x)$ is Lipschitz
  So S is closed, bounded.
  => $f(x) \ge C$
  for all $x \in S$.
  Why not $f(x) \ge f(x_0)$?

Step 2

1. $f(x_k) \le f(x_0)$
   $f(x)$ is bounded below
   for $k \to \infty$ $f(x_k)$ Exist
2. $f(x_k) - f(x_{k+1}) \ge \beta \alpha_k \|g_k\|^2$
   $f(x_1) - f(x_2) \ge$
   $f(x_2) - f(x_3) \ge$
   ...
   $f(x_k) - \bar{f} \ge$
   $\sum_{k=0}^{\infty} (f(x_k) - f(x_{k+1}))$

Step 3

$f(x_0) - \bar{f} =$
$f(x_0) - f(x_1) +$
$f(x_1) - f(x_2) +$
$f(x_2) - f(x_3) +$
...
$f(x_k) - \bar{f} \ge$
$\sum_{k=0}^{\infty} (f(x_k) - f(x_{k+1}))$

$f(x_{k+1}) \le f(x_k) - \beta \mu \alpha_k \|g_k\|^2$

$f(x_0) - f(x_{k+1}) \ge -\mu \alpha_k \|g_k\|^2$

$\ge \mu \epsilon \alpha_k \|g_k\| \cdot \|g_k\|$

$\ge \mu \epsilon m \alpha_k \|g_k\|^2$

$\underbrace{f(x_0) - \bar{f}}_{\text{finite}} \ge \underbrace{\sum_{k=0}^{\infty} \mu \epsilon m \alpha_k \|g_k\|^2}$

must converge

=> $\lim_{k \to \infty} \alpha_k \|g_k\|^2 \to 0$.