Lecture 15

H867 posted.

Quick proof that a point always satisfies the Wolfe conditions.

Most of class.

- Backtracking line search (30 mins)
- Convergence of Backtracking line search
- One of our most technical results.
- Make sure you are looking at the handout on this proof.

We'll spend ~10 class just walking through the assumptions of the statement and explaining why they are there.

Wolfe Conditions

$f(x_k + \alpha p) \le f(x_k) + c_1 \alpha g^T p$

$g(x_k + \alpha p)^T p \ge c_2 g^T p$

$0 < c_1 < c_2 < 1$

$c_1 = 10^{-4}$

Theorem: If you always satisfy a proper step size satisfies the Wolfe Conditions

- f - C^1 cont diff
- f - bounded below.

Define $\phi(\alpha) = f(x_k + \alpha p_k)$

$\phi'(\alpha) = \nabla f(x_k + \alpha p_k)^T p_k$

Define $\ell(\alpha) = \phi(0) + c_1 \alpha \phi'(0)$

$\ell'(\alpha) = c_1 \phi'(0)$.

Intuit is Backtracking line search: (BLS)

→ Pick a decreasing step size.

→ Check Stopping Criteria

BLS is usually slower (2)

→ more expensive per iteration.

→ Search in a range. (hard, all the same point)

BLS

Given $f, p_k, g(x)$

f & ∇f ok

Algo: while not done

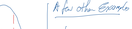
pick a search dir p_k

Set $x_{k+1} = x_k + \alpha_k p_k$

where α_k is the first element in the sequence $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$ satisfying

$f(x_k + \alpha_k p_k) \le f(x_k) + \alpha_k c_1 p_k^T g_k$

Algo Cannot miss if $\phi'(\alpha) \le c_1 \phi'(0)$ & $\phi(\alpha) \le c_1 \phi'(0)$ since $c_1 < c_2$



Low Curvature

BLS

Cannot miss if $\phi'(\alpha) \le c_2 \phi'(0)$

Proof: $\phi'(\alpha) \ge c_2 \phi'(0)$ & $\phi(\alpha) \le c_1 \phi'(0)$ since $c_1 < c_2$

A few other Examples

Quick Basic: If $\phi'(\alpha_k) < 0$, then set some point BLS accepts. (Sometimes)

So α will stop when $\ell'(\alpha^*) = \phi'(\alpha^*)$

ϕ' can not cross ℓ for $\alpha < \alpha^*$ (will be accepted!)

Theorem

f: $\mathbb{R}^n \to \mathbb{R}$

Cont diff

Cont grad.

x_0 given.

$x_{k+1} = x_k + \alpha_k p_k$ st

p_k descent dir

α_k satisfies Wolfe Conditions

Then $\lim_{k \to \infty} \|\nabla f(x_k)\| = 0$

Step Choice cond exists!

$\phi(\alpha) = f(x_k + \alpha p_k)$

α_k is the first element in the sequence $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$ satisfying

$f(x_k + \alpha_k p_k) \le f(x_k) + \alpha_k c_1 p_k^T g_k$

The set $S = \{x \mid f(x) \le f(x_0)\}$ is bounded.

$\|g(x)\|$ is Lipschitz Continuous.

$\|g(x) - g(y)\| \le L \|x - y\|$

$L < \infty$

(in S)

$x^* = \lim_{k \to \infty} x_k$ (exists by the bounded chosen.)

$\lim_{k \to \infty} \|x^* - x_k\| = 0$

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

Quick check on Lipschitz.

$f(x_k) - x^2$ is not a Lipschitz function.

Why not?

$|x_k^2 - y^2| \le L |x_k - y|$

$x_k \to \infty$ for fixed L.

If $y = 0$ then

$|x_k^2| \le L |x_k - 0|$

for some L. (Attention!)

Back to assumptions

The Second derivative is only Sufficient Condition.

$\cos(\angle g_k, p_k) \ge \varepsilon > 0$.

$-\frac{p_k^T g_k}{\|p_k\| \|g_k\|}$

$\|p_k\| \|g_k\|$

$\|p_k\| \ge \frac{\varepsilon}{\|g_k\|} \|g_k\|^2$

Quick check on $\|p_k\|$ and bounds

$\|p_k\| \le M$

$\|p_k\| \ge$ find (will prove)

conclusion

$\|g_k\| \to 0$

→ The second derivative are sufficient

$\|p_k\| \le M \|g_k\|$

and bounds

$\|p_k\| \le M$

$\Rightarrow \alpha_k \|p_k\| \ge$ find lower bound

$\alpha_k \ge \min(1, \frac{\varepsilon}{M} \|g_k\|^2)$

5. Step →

for example $\|g_k\| \to 0$.

Proof 5 Steps

1. f is bounded below.
2. $\lim_{k \to \infty} \phi'(\alpha_k) \text{ exists}$.
3. $\lim_{k \to \infty} \alpha_k \|g_k\|^2 \to 0$.
4. Use this step + Lipschitz to show $\alpha_k \ge \min(1, \frac{\varepsilon}{M} \|g_k\|^2)$
5. Step →

Lipschitz gives α_k not too small

Step 1

f is cont by S is closed bounded

$\Rightarrow \phi(\alpha)$ is bounded below.

$\phi(\alpha) \ge \phi(\alpha_k)$

$\phi(\alpha_k)$ is bounded below.

$\lim_{k \to \infty} \phi(\alpha_k) = \beta$

Step 2

$\phi(\alpha_k) - \beta \le \phi(\alpha_k) - \phi(\alpha_{k+1})$

$\phi(\alpha_k) - \phi(\alpha_{k+1}) \to 0$

Step 3

$\phi(\alpha_k) - \beta =$

$\phi(\alpha_k) - \phi(\alpha_{k+1})$

$+ \phi(\alpha_{k+1}) - \phi(\alpha_{k+2})$

$+ \phi(\alpha_{k+2}) - \phi(\alpha_{k+3})$

$+ \phi(\alpha_{k+3}) - \ldots$

$\sum_{i=0}^{\infty} \phi(\alpha_i) - \beta = \sum_{i=0}^{\infty} \phi(\alpha_i) - \phi(\alpha_{i+1})$

$= \sum_{i=0}^{\infty} \alpha_i c_1 \phi'(0) = c_1 \alpha_i \|g_i\|^2$

$= \sum_{i=0}^{\infty} \alpha_i \|g_i\|^2$

$\sum_{i=0}^{\infty} \alpha_i \|g_i\|^2 \le \frac{c_1}{\beta}$

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

Step 4

Take that not 1 step

$\phi(\alpha_k) - \phi(\alpha_{k+1})$

$\ge \sum_{i=0}^{k} \alpha_i c_1 \phi'(0) = c_1 \alpha_i \|g_i\|^2$

$= \sum_{i=0}^{k} \alpha_i c_1 \|g_i\|^2$

$\sum_{i=0}^{k} \alpha_i c_1 \|g_i\|^2 \le \frac{c_1}{\beta} \sum_{i=0}^{k} \alpha_i \|g_i\|^2$

$\lim_{k \to \infty} \sum_{i=0}^{k} \alpha_i \|g_i\|^2 \le \frac{c_1}{\beta} \lim_{k \to \infty} \sum_{i=0}^{k} \alpha_i \|g_i\|^2$

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

Step 5

Lipschitz gives α_k not too small

Step 1

Take $\phi(\alpha_k) - \phi(\alpha_{k+1})$

$\ge \sum_{i=0}^{k} \alpha_i c_1 \phi'(0) = c_1 \alpha_i \|g_i\|^2$

$\ge \sum_{i=0}^{k} \alpha_i c_1 \|g_i\|^2 \cdot \cos(\angle g_i, p_i)$

$\ge \sum_{i=0}^{k} \alpha_i c_1 \|g_i\|^2 \cdot \frac{\|p_i\|}{\|g_i\|}$

$\ge \sum_{i=0}^{k} \alpha_i c_1 \|g_i\|^2 \cdot \frac{\varepsilon}{M} \|g_i\|^2$

$\ge \sum_{i=0}^{k} \alpha_i c_1 \frac{\varepsilon}{M} \|g_i\|^4$

$\sum_{i=0}^{k} \alpha_i c_1 \frac{\varepsilon}{M} \|g_i\|^4 \le \phi(\alpha_k) - \phi(\alpha_{k+1})$

$\le \phi(\alpha_k) - \beta$

$\sum_{i=0}^{k} \alpha_i \|g_i\|^4 \le \frac{M}{c_1 \varepsilon} (\phi(\alpha_k) - \beta)$

$\lim_{k \to \infty} \sum_{i=0}^{k} \alpha_i \|g_i\|^4 \le \frac{M}{c_1 \varepsilon} \lim_{k \to \infty} (\phi(\alpha_k) - \beta)$

$\lim_{k \to \infty} \sum_{i=0}^{k} \alpha_i \|g_i\|^4 \le \frac{M}{c_1 \varepsilon} (\beta - \beta)$

$\lim_{k \to \infty} \sum_{i=0}^{k} \alpha_i \|g_i\|^4 \le 0$

$\lim_{k \to \infty} \alpha_k \|g_k\|^4 \to 0$

Step 6

Take step 4 + 5

$\lim_{k \to \infty} \alpha_k \|g_k\|^2 \le \frac{c_1}{\beta} \lim_{k \to \infty} \alpha_k \|g_k\|^4$

$\lim_{k \to \infty} \alpha_k \|g_k\|^2 \le \frac{c_1}{\beta} \lim_{k \to \infty} \alpha_k \|g_k\|^2 \cdot \|g_k\|^2$

$\lim_{k \to \infty} \alpha_k \|g_k\|^2 \le \frac{c_1}{\beta} \lim_{k \to \infty} \alpha_k \|g_k\|^2 \cdot \lim_{k \to \infty} \|g_k\|^2$

$\lim_{k \to \infty} \alpha_k \|g_k\|^2 \le \frac{c_1}{\beta} \lim_{k \to \infty} \alpha_k \|g_k\|^2 \cdot 0$

$\lim_{k \to \infty} \alpha_k \|g_k\|^2 \le 0$

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

Step 7

Take step 1 + 6

$\phi(\alpha_k) - \phi(\alpha_{k+1})$

$\ge \sum_{i=0}^{k} \alpha_i c_1 \phi'(0) = c_1 \alpha_i \|g_i\|^2$

$\ge \sum_{i=0}^{k} \alpha_i c_1 \frac{\varepsilon}{M} \|g_i\|^4$

$\ge \sum_{i=0}^{k} \alpha_i c_1 \frac{\varepsilon}{M} \|g_i\|^2 \cdot \|g_i\|^2$

$\ge \sum_{i=0}^{k} \alpha_i c_1 \frac{\varepsilon}{M} \|g_i\|^2 \cdot \frac{\varepsilon}{M} \|g_i\|^2$

$\ge \sum_{i=0}^{k} \alpha_i c_1 \frac{\varepsilon^2}{M^2} \|g_i\|^4$

$\sum_{i=0}^{k} \alpha_i c_1 \frac{\varepsilon^2}{M^2} \|g_i\|^4 \le \phi(\alpha_k) - \phi(\alpha_{k+1})$

$\le \phi(\alpha_k) - \beta$

$\sum_{i=0}^{k} \alpha_i \|g_i\|^4 \le \frac{M^2}{c_1 \varepsilon^2} (\phi(\alpha_k) - \beta)$

$\lim_{k \to \infty} \sum_{i=0}^{k} \alpha_i \|g_i\|^4 \le \frac{M^2}{c_1 \varepsilon^2} \lim_{k \to \infty} (\phi(\alpha_k) - \beta)$

$\lim_{k \to \infty} \sum_{i=0}^{k} \alpha_i \|g_i\|^4 \le \frac{M^2}{c_1 \varepsilon^2} (\beta - \beta)$

$\lim_{k \to \infty} \sum_{i=0}^{k} \alpha_i \|g_i\|^4 \le 0$

$\lim_{k \to \infty} \alpha_k \|g_k\|^4 \to 0$

Step 8

Take step 5 + 7

$\lim_{k \to \infty} \alpha_k \|g_k\|^2 \le \frac{c_1}{\beta} \lim_{k \to \infty} \alpha_k \|g_k\|^4$

$\lim_{k \to \infty} \alpha_k \|g_k\|^2 \le \frac{c_1}{\beta} \lim_{k \to \infty} \alpha_k \|g_k\|^2 \cdot \|g_k\|^2$

$\lim_{k \to \infty} \alpha_k \|g_k\|^2 \le \frac{c_1}{\beta} \lim_{k \to \infty} \alpha_k \|g_k\|^2 \cdot \lim_{k \to \infty} \|g_k\|^2$

$\lim_{k \to \infty} \alpha_k \|g_k\|^2 \le \frac{c_1}{\beta} \lim_{k \to \infty} \alpha_k \|g_k\|^2 \cdot 0$

$\lim_{k \to \infty} \alpha_k \|g_k\|^2 \le 0$

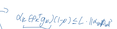
$\lim_{k \to \infty} \alpha_k \|g_k\|^2 = 0$

Step 9

$\alpha_k \ge \min(1, \frac{\varepsilon}{M} \|g_k\|^2)$

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

$\Rightarrow \lim_{k \to \infty} \|g_k\|^2 = 0$



Re-arrange in Lag

$\phi(\alpha_k) - \phi(\alpha_{k+1}) - \phi(\alpha_k) \ge$

$2 \mu_{\max} g_k$

$\phi(\alpha_k) - \phi(\alpha_{k+1}) - \phi(\alpha_k) \ge$

$- 2 \alpha_k \mu_{\max}^T g_k - 2 L \|\alpha_k p_k\|^2$

$0 \ge 2 \mu_{\max}^T g_k - 2 \alpha_k \mu_{\max}^T g_k$

$- 2 L \|\alpha_k p_k\|^2$

Cancel α $\mu_{\max}^T g_k$

$\alpha_k \mu_{\max}^T g_k (1 - \alpha_k) \le$

$- 2 L \|\alpha_k p_k\|^2$

$\alpha_k \mu_{\max}^T g_k (1 - \alpha_k) \le -L \|\alpha_k p_k\|^2$

$\|p_k\| \le M$

$\alpha_k L M^2 \ge \alpha_k \mu_{\max}^T g_k$

$\alpha_k \frac{\|\mu_{\max}\| \le L}{L M^2} \ge \|\alpha_k p_k\|^2$

$\alpha_k \ge \|\alpha_k p_k\|^2$

Compare $\|\alpha_k p_k\|$

Take step 10 + 11

$\alpha_k \ge \|\alpha_k p_k\|^2 \cdot \frac{1}{M}$

$\alpha_k \ge \|\alpha_k p_k\|^2 \cdot \frac{1}{M}$

$\Rightarrow \lim_{k \to \infty} \|\alpha_k p_k\|^2 = 0$