Lecture 19

Studying: Line search algorithms.

How to guarantee descent in line search via the strong wolfe conditions.

Example of Strong Wolfe (computer)

Why a point that satisfies strong wolfe must exist.
- A fun proof that involves "speeding"

Backtracking line search — what we'll use in a proof that's just a little bit easier

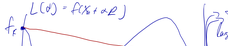
Iterative Alg $x^{(k)} \Rightarrow x^{(k+1)}$

- Pick Search direction p
- Pick Step Length α
- α needs to be chosen ch...
- α needs to satisfy some condition

Goal:

find condition on α so we make progress toward a minimum.

- Avoid tiny steps (unless they really help)
- Avoid short step unless very close to minimum



$1 > c_1 > 0$ $\quad c_1 = 0.0001$

The bulge avoid large, useless steps.

Small steps?

Idea: let the gradient on the line $L(\alpha)$ to flatten out. $\quad 1 > c_2 > c_1$, $c_2 = 0.999$

$L'(\alpha) > c_2 L'(0)$

Wolfe Conditions for choosing α.

$1 > c_2 \geq c_1 > 0$

1. $f(x_k + \alpha p_k) \leq f(x_k) + \alpha c_1 p^T g$

$L(\alpha) \leq L(0) + \alpha c_1 L'(0)$

Called Sufficient Decrease or the Armijo Condition)

2. $g(x_k + \alpha p_k)^T p \geq c_2 p^T g$

$L'(\alpha) \geq c_2 L'(0)$

Strong Wolfe

1. Same

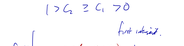
2. $|L'(\alpha)| \leq c_2 |L'(0)|$

1. it's okay
2. it's not for Wolfe
3. it's true for Strong Wolfe.

Theorem there exists a α s.t. α satisfies the Wolfe Conditions for a Cont. diff function f which is bounded below. (p is a descent dir.)

$1 > c_2 \geq c_1 > 0$

first interval.



← max is Cont. $\Rightarrow$ Some slope $c_2 L'(0)$ found!

$L'(\alpha) = c_2 L'(0)$

α $\quad \alpha_1 =$

By MVT there exists $\alpha_1 < \alpha < \alpha_2$ s.t. $L'(\alpha) = c_1 L'(0)$

The proof for Convergence of line Search methods we will do in Class Uses

- Backtracking Line Search.

↳ Usually slower (in cycles)

⇒ Requires more load/function calls.

↳ Much Easier to construct!

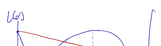
Backtracking line search alg.

Given f, x0, g(x), f : R^n -> R

Alg:

While not done.
pick a search direction pk,
set x_{k+1} = x_k + alpha_k p_k
where alpha_k is the first element of the sequence:
1, 1/2, 1/4, 1/8, ...
to give sufficient decrease, i.e.
f(x_k + alpha p_k) <=
f(x_k) + alpha*mu*g(x_k)^T p_k

where mu < 1.



pick one $1/4$ for the ex..

$\Rightarrow$ $\alpha \in (c_2 \alpha_0, \alpha_0)$