Lecture 18

* HW6 posted

Done with LPs/SIMPLEX
-> Do look at the phase 1/Crash codes.
(you may need them for HW!)
-> post questions on Piazza.

Algorithms for unconstrained optimization
- line search
- trust region
- simplifying notation
Line search alg overview
Descent direction
... into problems with
guaranteeing descent/convergence...

Back to

minimize $f(x)$

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

Opt Cond

1st order necessary: $g(x^*) = 0$

2nd order necessary: $H(x^*) \succeq 0$

2nd order sufficient: $g(x^*) = 0$, $H(x^*) \succ 0$.

All methods to solve this are iterative!

$x^{(k)} \longrightarrow x^{(k+1)}$

Line Search

Trust Region

I quickly tire of writing iteration index.

$\Rightarrow$ Simplified notation

$\underline{x} = x^{(k)}$

$\underline{x}^+ = x^{(k+1)}$

$\underline{g} = g(x^{(k)})$

$\underline{g}^+ = g(x^{(k+1)})$

$H = H(x^{(k)})$

$H^+ = H(x^{(k+1)})$

$f_k = f(x_k)$

$f^+ = f(x_{k+1})$

Example:

$f(\underline{x} - \alpha \underline{g}) = f(\underline{x}^{(k)} - \alpha g(x^{(k)}))$

Line Search Method

Step length

Search direction

$x^+ = \underline{x} + \alpha \underline{p}$

pick α

pick p

How to pick p?
- gradient descent
$p = -g$
- Newton ($H p = -g$)

Quasi-Newton method

How to pick α?
- Scrolling quick!

Make sure algo when

$f(x_{k+1}) < f(x_k)$

$f^+ < f_k$.

So p should help make f small.

Only p is a descent direction if $p^T g < 0$

Why Descent Direction?

Taylor Exam

$f(\underline{x} + \gamma \underline{p}) = f(\underline{x}) + \gamma \underline{p}^T g + O(\gamma^2)$

So if γ is tiny, and $\underline{p}^T g < 0$, then

$f(\underline{x} + \gamma \underline{p}) = f(x_k) + \underbrace{\gamma \underbrace{p^T g}_{<0}}_{<0} + O(\gamma^2)$

$< f(x_k)$

Lemma: Let B be a Symm. pos def Matrix. then any p st

$Bp = -g$ (solve the system)

gives a descent direction

Proof: $p = -B^{-1} g$

B is sym. pos def $\Rightarrow$

B^{-1} is sym. pos def $\Rightarrow$

$g^T p = \boxed{-g^T B^{-1} g} < 0$

> 0

$p^T g = p^T B p < 0$.

Why B?

If $B = I$

$p = -g \Rightarrow$ gradient descent

If $B = H(x_k) \Rightarrow$

$p = -H^{-1} g \Rightarrow$ Newton.

(and H is sym. pos def)

(B) Can $B(x_k)$ change p is always a descent direction

Is this enough to guarantee we find a minimum?

If $p^T g < 0$ and $f(x_{k+1}) < f(x_k)$

with $x_k \longrightarrow$ minimum?

Counter

$f(x) = x^2$

for 1

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

$\alpha_1 = \frac{1}{2}, \alpha_2 = \frac{1}{4}, \alpha_3 = \frac{1}{16}, \ldots$

$f(x_{k+1}) < f(x_k)$?

Yes!

$\Rightarrow$ Step sizes get too small.

$\alpha > 1$

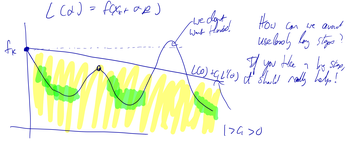
Sum of large steps, I can get to a minimum

Goal:

Opt Conditions on α

s.t. we make progress towards a minimum and avoid those scenarios.

$L(\alpha) = f(x_k + \alpha p)$



How can we avoid unbounded long steps?

If you take a large step it should really help!

CS 530 - Purdue University - Computer Science - Jeffrey M. Siskind - 2022