Numerical Optimization
Lecture Notes #3
Convergence; Line Search Methods

Peter Blomgren,
⟨`blomgren.peter@gmail.com`⟩

Department of Mathematics and Statistics
Dynamical Systems Group
Computational Sciences Research Center
San Diego State University
San Diego, CA 92182-7720

http://terminus.sdsu.edu/

Fall 2018

Outline

SAN DIEGO STATE
UNIVERSITY

Quick Recap: Last Time

Some fundamental building blocks of unconstrained optimization:

Theorem (Taylor)

For some $t \in (0, 1)$, we have

$$f(\bar{\mathbf{x}} + \bar{\mathbf{p}}) = f(\bar{\mathbf{x}}) + \bar{\mathbf{p}}^T \underbrace{\nabla f(\bar{\mathbf{x}})}_{gradient} + \frac{1}{2} \bar{\mathbf{p}}^T \underbrace{\left[\nabla^2 f(\bar{\mathbf{x}} + t\bar{\mathbf{p}}) \right]}_{Hessian} \bar{\mathbf{p}}.$$

4 theorems relating $f(\bar{\mathbf{x}})$ and its derivatives to optimal solutions.

[1] $\bar{\mathbf{x}}^*$ optimal $\Rightarrow \nabla f(\bar{\mathbf{x}}^*) = 0$.

[2] $\bar{\mathbf{x}}^*$ optimal $\Rightarrow \nabla f(\bar{\mathbf{x}}^*) = 0$, and $\nabla^2 f(\bar{\mathbf{x}}^*)$ positive semi-definite.

[3] $\nabla f(\bar{\mathbf{x}}^*) = 0$, and $\nabla^2 f(\bar{\mathbf{x}}^*)$ positive definite $\Rightarrow \bar{\mathbf{x}}^*$ optimal.

[4a] f convex, and $\bar{\mathbf{x}}^*$ local optimum $\Rightarrow \bar{\mathbf{x}}^*$ global optimum.

[4b] f convex, and $\nabla f(\bar{\mathbf{x}}^*) = 0 \Rightarrow \bar{\mathbf{x}}^*$ global optimum.

Note: The complete statement of the theorems require sufficient smoothness (existence) of derivatives of f.

Key Concepts   Rate of Convergence

Definition (Rate of Convergence, Sequences)

Suppose the sequence $\underline{\beta} = \{\beta_n\}_{n=1}^{\infty}$ converges to zero, and $\underline{\bar{\mathbf{x}}} = \{\bar{\mathbf{x}}_n\}_{n=1}^{\infty}$ converges to a point $\bar{\mathbf{x}}^*$.

If $\exists K > 0: \|\bar{\mathbf{x}}_n - \bar{\mathbf{x}}^*\| < K\beta_n$, for $n > N$ (*i.e.* for n large enough), then we say that $\{\bar{\mathbf{x}}_n\}_{n=1}^{\infty}$ converges to $\bar{\mathbf{x}}^*$ with a **Rate of Convergence** $\mathcal{O}(\beta_n)$ ("Big Oh of β_n").

We write

$$\bar{\mathbf{x}}_n = \bar{\mathbf{x}}^* + \mathcal{O}(\beta_n).$$

Note: The sequence $\underline{\beta} = \{\beta_n\}_{n=1}^{\infty}$ is usually chosen to be *e.g.*

$$\beta_n = \frac{1}{n^p}, \quad \text{for some value of } p.$$

Rates of Convergence

Let $\bar{\underline{\mathbf{x}}} = \{\bar{\mathbf{x}}_n\}_{n=1}^{\infty}$ be a sequence converging to $\bar{\mathbf{x}}^*$, the convergence rate is said to be

Q-linear (quotient-linear) if $\exists r \in (0, 1)$ and $K \in \mathbb{Z}$ such that

$$\frac{\|\bar{\mathbf{x}}_{k+1} - \bar{\mathbf{x}}^*\|}{\|\bar{\mathbf{x}}_k - \bar{\mathbf{x}}^*\|} \leq r, \quad \forall k \geq K.$$

Q-superlinear if

$$\lim_{k \to \infty} \frac{\|\bar{\mathbf{x}}_{k+1} - \bar{\mathbf{x}}^*\|}{\|\bar{\mathbf{x}}_k - \bar{\mathbf{x}}^*\|} = 0.$$

Q-quadratic if $\exists r \in \mathbb{R}^+$ and $K \in \mathbb{Z}$ such that

$$\frac{\|\bar{\mathbf{x}}_{k+1} - \bar{\mathbf{x}}^*\|}{\|\bar{\mathbf{x}}_k - \bar{\mathbf{x}}^*\|^2} \leq r, \quad \forall k \geq K.$$

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Line Search Methods

We now focus on line search methods where we *(i)* pick a **search direction** $\bar{\mathbf{p}}_k$ and, then *(ii)* solve the one-dimensional problem

$$\min_{\alpha > 0} f(\bar{\mathbf{x}}_k + \alpha \bar{\mathbf{p}}_k).$$

The solution gives us an optimal value for α_k, so the next point is given by

$$\bar{\mathbf{x}}_{k+1} = \bar{\mathbf{x}}_k + \alpha_k \bar{\mathbf{p}}_k,$$

where α_k is known as the **step length**.

In order for a line search method to be work well, we need good choices of the direction $\bar{\mathbf{p}}_k$ and the step length α_k.

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Steepest Descent Direction

Line Search

The intuitive choice for $\bar{\mathbf{p}}_k$ is to move in the direction of steepest descent, *i.e.* in the negative gradient direction.

Going back to the Taylor expansion

$$f(\bar{\mathbf{x}} + \alpha\bar{\mathbf{p}}) = f(\bar{\mathbf{x}}) + \alpha\bar{\mathbf{p}}^T \nabla f(\bar{\mathbf{x}}),$$

we immediately see that the direction of most rapid decrease gives

$$\min_{\|\bar{\mathbf{p}}\|=1} \bar{\mathbf{p}}^T \nabla f(\bar{\mathbf{x}}) = \min_{\theta \in [0,2\pi]} \cos\theta \, \|\nabla f(\bar{\mathbf{x}})\| = -\|\nabla f(\bar{\mathbf{x}})\|,$$

which is achieved when $\theta = \pi \Leftrightarrow \bar{\mathbf{p}} = -\nabla f(\bar{\mathbf{x}})/\|\nabla f(\bar{\mathbf{x}})\|$.

Recall: $\bar{\mathbf{v}}^T \bar{\mathbf{w}} = \cos\theta \, \|\bar{\mathbf{v}}\| \cdot \|\bar{\mathbf{w}}\|$, where θ is the angle between the vectors $\bar{\mathbf{v}}$ and $\bar{\mathbf{w}}$.
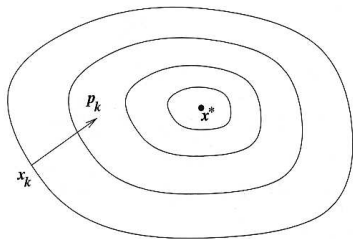
Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Steepest Descent Direction

Figure: The steepest descent direction $\bar{\mathbf{p}}_k$ is perpendicular to the contour lines of the objective.
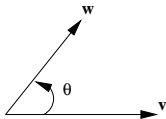

Figure: $\bar{\mathbf{v}}^T \bar{\mathbf{w}} = \cos\theta \, \|\bar{\mathbf{v}}\| \cdot \|\bar{\mathbf{w}}\|$.

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Newton Direction

Line Search

If f is smooth enough and the Hessian is positive definite, we can select $\bar{\mathbf{p}}_k$ to be the "Newton direction." We write down the second order Taylor expansion:

$$f(\bar{\mathbf{x}} + \bar{\mathbf{p}}) \approx f(\bar{\mathbf{x}}) + \bar{\mathbf{p}}^T \nabla f(\bar{\mathbf{x}}) + \frac{1}{2} \bar{\mathbf{p}}^T \left[\nabla^2 f(\bar{\mathbf{x}}) \right] \bar{\mathbf{p}}.$$

We seek the minimum of the right-hand-side by computing the derivative width respect to $\bar{\mathbf{p}}$ and set the result to zero

$$\nabla f(\bar{\mathbf{x}}) + \left[\nabla^2 f(\bar{\mathbf{x}}) \right] \bar{\mathbf{p}} = 0,$$

which gives the Newton direction

$$\bar{\mathbf{p}}^N = -\left[\nabla^2 f(\bar{\mathbf{x}}) \right]^{-1} \nabla f(\bar{\mathbf{x}}).$$

SAN DIEGO STATE
UNIVERSITY

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Newton Direction

Line Search

As long as the Hessian is positive definite, $\bar{\mathbf{p}}^N$ is a descent-direction:

$$\bar{\mathbf{p}}^N \nabla f(\bar{\mathbf{x}}) = -\nabla f(\bar{\mathbf{x}})^T \underbrace{\left[\nabla^2 f(\bar{\mathbf{x}})\right]^{-T}}_{\text{Pos. Def.}} \nabla f(\bar{\mathbf{x}}) < 0$$

Note: Clearly, the Newton direction is more "expensive" than the steepest descent direction — we must compute the Hessian matrix $\nabla^2 f(\bar{\mathbf{x}})$, and invert it (*i.e.* solve an $n \times n$ linear system).

Note: The convergence rate for steepest descent methods is **linear** and for Newton methods it is **quadratic**, hence there is a lot to gain by finding the Newton direction.

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Example: NW[1st]-2.2, p 30.

Problem: Show that the function $f(x) = 8x + 12y + x^2 - 2y^2$ has only one stationary point, and that it is neither a maximum nor a minimum, but a saddle point. Sketch the contours for f.

Solution: The gradient of f is

$$\nabla f = \left[\begin{array}{c} 8 + 2x \\ 12 - 4y \end{array} \right]$$

which has the stationary point $(x, y) = (-4, 3)$. Since the Hessian

$$\nabla^2 f = \left[\begin{array}{cc} 2 & 0 \\ 0 & -4 \end{array} \right]$$

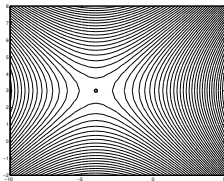has both positive and negative eigenvalues, the stationary point must be a saddle point.
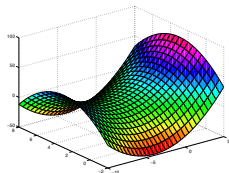

Figure: The contour lines for $f(x)$.

Figure: The function $f(x)$ around the stationary point.

SAN DIEGO STATE
UNIVERSITY

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Example: NW^{1st}-2.2, p 30.

2 of 3

If we start an iteration in $(x_0, y_0) = (0, 0)$:

The steepest descent direction is

$$\bar{\mathbf{p}}_0^{SD} = -\nabla f = -\left[\begin{array}{c} 8 + 2x \\ 12 - 4y \end{array}\right] = -\left[\begin{array}{c} 8 \\ 12 \end{array}\right]$$

and the Newton direction is

$$\bar{\mathbf{p}}_0^{N} = -[\nabla^2 f]^{-1}\, \nabla f = -\left[\begin{array}{cc} 2 & 0 \\ 0 & -4 \end{array}\right]^{-1} \left[\begin{array}{c} 8 \\ 12 \end{array}\right] = \left[\begin{array}{c} -4 \\ 3 \end{array}\right]$$

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1
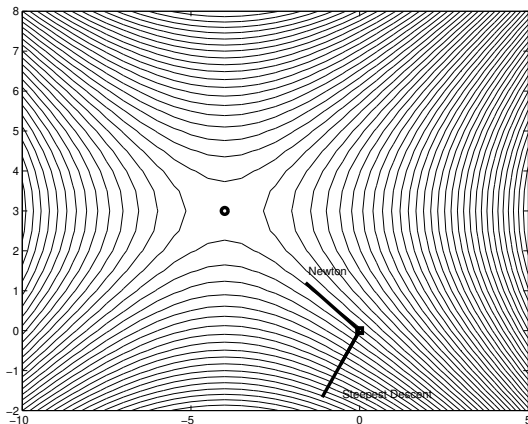
Example: NW^{1st}-2.2, p 30.

Figure: The **Newton** and **Steepest Descent** directions starting in $(0, 0)$. Note that the Newton method is heading to the saddle point, but the Steepest descent method will, in general, not converge to a non-minimum stationary point.

SAN DIEGO STATE
UNIVERSITY

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Modified (Convexified) Example

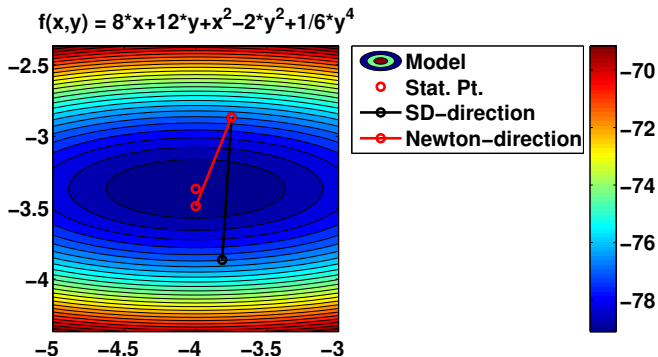**f(x,y) = 8*x+12*y+x²−2*y²+1/6*y⁴**

Figure: Convexification of the silly book problem. Same point of interest, $\nabla f = [8 + 2x, \ 12 - 4y + 2/3y^3]^T$, $\nabla^2 f = \begin{bmatrix} 2 & 0 \\ 0 & -4 + 2y^2 \end{bmatrix}$. Now, both the steepest descent and Newton directions are descent directions.

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Line Search Methods — Directions

Method	Search Direction	Convergence
Steepest Descent	$p_k = -\nabla f(\bar{\mathbf{x}}_k)/\|\nabla f(\bar{\mathbf{x}}_k)\|$	Linear
Quasi-Newton	$p_k = -H_k^{-1}\nabla f(\bar{\mathbf{x}}_k)$	Super-Linear
Newton	$p_k = -[\nabla^2 f(\bar{\mathbf{x}}_k)]^{-1}\nabla f(\bar{\mathbf{x}}_k)$	Quadratic

Table: Summary of search directions for different schemes. In **Quasi-Newton** schemes we do not explicitly compute the Hessian $\nabla^2 f(\bar{\mathbf{x}}_k)$ in each iteration, instead we use an approximation $H_k \approx \nabla^2 f(\bar{\mathbf{x}}_k)$ which is updated in some clever way [TO BE EXPLORED IN GREAT DETAIL LATER] (lecture 18→ . . .).

We will return to the selection of $\bar{\mathbf{p}}_k$, but let's consider the computation of the step length α_k...

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Line Search Methods: Step Length Selection

Given a descent direction $\bar{\mathbf{p}}_k$ we would like to find the global minimizer α_k^* of

$$\min_{\alpha > 0} f(\bar{\mathbf{x}}_k + \alpha \bar{\mathbf{p}}_k).$$

As this is just one of possible many steps in the iteration, it is not wise to expend too much time in finding α_k. We are faced with a trade-off:

— We want an α_k so that we get a **substantial reduction** in the objective f.

— We want to find α_k **fast**.

In practice we perform an **inexact line search** — settling for an α_k which gives **adequate reduction** in the objective.

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
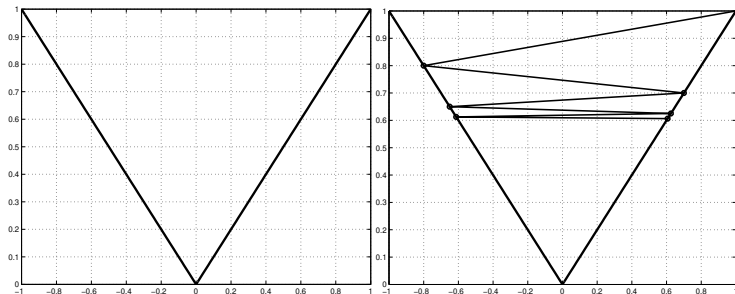Homework #1

What is "adequate reduction?"

Figure: Consider the objective $f(x) = \sqrt{x^2 + 10^{-8}}$, if we let $x_k = \{1, -0.8, 0.7, -0.65, 0.625, -0.6125, 0.60625, \dots\}$, then the descent directions are given by $p_k = \{-1, 1, -1, 1, -1, 1, -1, \dots\}$, so this generates a decreasing sequence $f(x_k + \alpha_k p_k) < f(x_k)$. However, with the current choice of $\alpha_k = \{-1.8, 1.5, -1.35, 1.275, -1.2375, 1.21875, \dots\}$ the convergence rate is less than spectacular.

Clearly, we need a stronger condition than $f(x_k + \alpha_k p_k) < f(x_k)$.

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

The Wolfe Conditions   1 of 2

There are many ways to enforce reduction in the objective, *e.g.*

Armijo Condition   (Wolfe Condition #1)

The **Armijo Condition**

$$f(\bar{\mathbf{x}}_k + \alpha \bar{\mathbf{p}}_k) \leq f(\bar{\mathbf{x}}_k) + c_1 \alpha \bar{\mathbf{p}}_k^T \nabla f(\bar{\mathbf{x}}), \quad c_1 \in (0, 1),$$

---

requires the reduction to be proportional to the step length α, as
well as the directional derivative $\bar{\mathbf{p}}_k^T \nabla f(\bar{\mathbf{x}})$. **In practice** c_1 is
usually set to be quite small, *e.g.* $\sim 10^{-4}$.

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

The Wolfe Conditions

To rule out unacceptably short steps, we additionally enforce

---

Curvature Coondition   (Wolfe Condition #2)

The **Curvature Condition**

$$\bar{\mathbf{p}}_k^T \nabla f(\bar{\mathbf{x}}_k + \alpha \bar{\mathbf{p}}_k) \geq c_2 \bar{\mathbf{p}}_k^T \nabla f(\bar{\mathbf{x}}_k), \quad c_2 \in (c_1, 1).$$

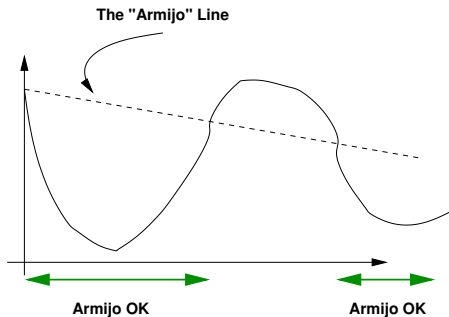It prevents us from stopping when more progress can be made by moving further (increasing α).

Together these two conditions are known as the **Wolfe conditions**.

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

The Wolfe Conditions: Part I — The Armijo Condition

The **Armijo Condition**

$$f(\bar{\mathbf{x}}_k + \alpha \bar{\mathbf{p}}_k) \leq f(\bar{\mathbf{x}}_k) + c_1 \alpha \bar{\mathbf{p}}_k^T \nabla f(\bar{\mathbf{x}}), \quad c_1 \in (0, 1)$$

requires the reduction to be proportional to the step length α, as well as the directional derivative. **In practice** c_1 is usually set to be quite small, e.g. $\sim 10^{-4}$.

The "Armijo" Line

Armijo OK   Armijo OK

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
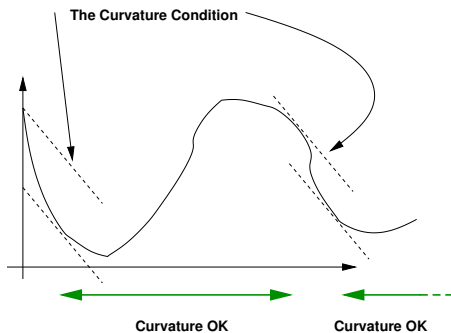Step Length Selection — The Wolfe Conditions
Homework #1

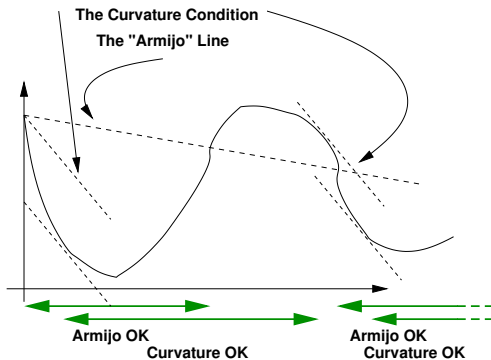The Wolfe Conditions: Part II — The Curvature Condition

To rule out unacceptable short steps, the **curvature condition**

$$\bar{\mathbf{p}}_k^T \nabla f(\bar{\mathbf{x}}_k + \alpha \bar{\mathbf{p}}_k) \geq c_2 \bar{\mathbf{p}}_k^T \nabla f(\bar{\mathbf{x}}_k), \quad c_2 \in (c_1, 1)$$

— it prevents us from stopping when more progress can be made by moving further (increasing α). Typical values: $c_2^{N,QN} = 0.9$, $c_2^{CG} = 0.1$.
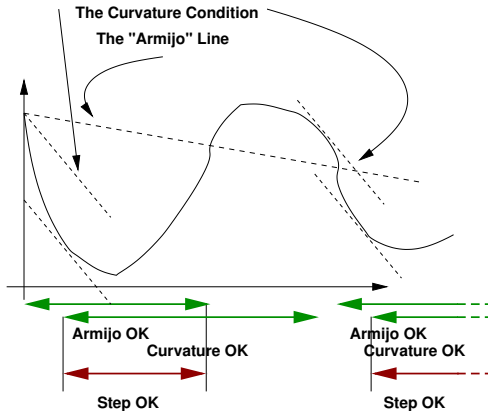

The Curvature Condition

Curvature OK   Curvature OK

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

The Wolfe Conditions: Part I+II — Acceptable Step   1/2

Together, the Armijo and Curvature conditions constitute the Wolfe Conditions.

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

The Wolfe Conditions: Part I+II — Acceptable Step   2/2

Together, the Armijo and Curvature conditions constitute the Wolfe Conditions.

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

The Strong Wolfe Conditions

A step length α may satisfy the **Wolfe Conditions**

$$
\begin{aligned}
f(\bar{\mathbf{x}}_k + \alpha\bar{\mathbf{p}}_k) &\leq f(\bar{\mathbf{x}}_k) + c_1\alpha\bar{\mathbf{p}}_k^T\nabla f(\bar{\mathbf{x}}), & c_1 \in (0,1) \\
\bar{\mathbf{p}}_k^T\nabla f(\bar{\mathbf{x}}_k + \alpha\bar{\mathbf{p}}_k) &\geq c_2\bar{\mathbf{p}}_k^T\nabla f(\bar{\mathbf{x}}_k), & c_2 \in (c_1,1)
\end{aligned}
$$

even though it is far from a minimizer of $f(\bar{\mathbf{x}}_k + \alpha\bar{\mathbf{p}}_k)$, the **Strong Wolfe Conditions**

$$
\begin{aligned}
f(\bar{\mathbf{x}}_k + \alpha\bar{\mathbf{p}}_k) &\leq f(\bar{\mathbf{x}}_k) + c_1\alpha\bar{\mathbf{p}}_k^T\nabla f(\bar{\mathbf{x}}), & c_1 \in (0,1) \\
|\bar{\mathbf{p}}_k^T\nabla f(\bar{\mathbf{x}}_k + \alpha\bar{\mathbf{p}}_k)| &\leq c_2|\bar{\mathbf{p}}_k^T\nabla f(\bar{\mathbf{x}}_k)|, & c_2 \in (c_1,1)
\end{aligned}
$$

further disallows values of

$$
\left[\bar{\mathbf{p}}_k^T\nabla f(\bar{\mathbf{x}}_k + \alpha\bar{\mathbf{p}}_k)\right]
$$

which are "too positive," thus excluding point that are far from the stationary points of $\bar{\mathbf{p}}_k^T\nabla f(\bar{\mathbf{x}}_k + \alpha\bar{\mathbf{p}}_k)$.

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Are the Wolfe Conditions too Restrictive?

It can be shown (see NW[2nd] pp.35–36) that there **exist** step lengths α which satisfy the Wolfe Conditions (and the Strong Wolfe Conditions) **for every** function f which is smooth and bounded below.

Formally —

Theorem (Existence of Acceptable α)

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable. Let $\bar{\mathbf{p}}_k$ be a descent direction at $\bar{\mathbf{x}}_k$, and assume that f is bounded below along the line $\{\bar{\mathbf{x}}_k + \alpha \bar{\mathbf{p}}_k : \alpha > 0\}$. Then if $0 < c_1 < c_2 < 1$, there exist intervals of step lengths satisfying the Wolfe conditions and the strong Wolfe conditions.

See also "Goldstein Conditions" (NW[2nd] p.36.)

SAN DIEGO STATE
UNIVERSITY

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Algorithm: Backtracking Linesearch

Algorithm: Backtracking Linesearch

```
[0] Find a descent direction p̄_k
[1] Set ᾱ > 0, ρ ∈ (0,1), c ∈ (0,1), set α = ᾱ
[2] While f(x̄_k + αp̄_k) > f(x̄_k) + cαp̄_k^T ∇f(x̄_k)
[3]     α = ρα
[4] End-While
[5] Set α_k = α
```

If an algorithm selects the step lengths appropriately (*e.g.* backtracking), we do not have to check the second inequality of the Wolfe conditions.

The algorithm above is especially well suited for use with Newton method ($\bar{\mathbf{p}}_k = \bar{\mathbf{p}}_k^N$), where $\overline{\alpha} = 1$. It is less successful for quasi-Newton and CG-based approaches.

The value of the **contraction factor** ρ can be allowed to vary at each iteration of the line search. (To be revisited)

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Homework #1 — Due at 12:00pm, Friday September 21, 2018

NW$^{\text{2nd}}$-3.1: Program the steepest descent and Newton algorithms using the backtracking line search. Use them to minimize the Rosenbrock function

$$f(\bar{\mathbf{x}}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

Set the initial step length $\alpha_0 = 1$ and report the step length used by each method at each iteration. First try the initial point $\bar{\mathbf{x}}_0^T = [1.2, 1.2]$ and then the more difficult point $\bar{\mathbf{x}}_0^T = [-1.2, 1]$.

Suggested values: $\overline{\alpha} = 1$, $\rho = \frac{1}{2}$, $c = 10^{-4}$.
Stop when: $|f(\vec{x}_k)| < 10^{-8}$, or $\|\nabla f(\vec{x}_k)\| < 10^{-8}$.

Note: The homework is due in Peter's mailbox in GMCS-411 or, office GMCS-587 (slide under the door if I'm not there).

Introduction
Line Search Methods

Search Direction: Steepest Descent, Newton, or Other?!?
Step Length Selection — 1D Minimization
Step Length Selection — The Wolfe Conditions
Homework #1

Index