Math 298
Fundamental Concepts in Computational and Applied Mathematics
Lecture 8: Optimization and Nonlinear Equations

Juan Meza
School of Natural Sciences
University of California, Merced

Optimization and Nonlinear Equations

- Optimization and nonlinear equations at the heart of many science and engineering problems
- Many of the ideas/methods used in optimization are the same as in nonlinear equations
- Can divide methods generically into derivative and non-derivative methods

For minimization we usually state the problem as: $\min f(x)$
For nonlinear equations we usually state the problem as : Solve $F(x) = 0$

Nonlinear Equations

Suppose we have:

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

then the problem of solving a set of nonlinear equations is given by:

$$\text{find } x_* \in \mathbb{R}^n \text{ such that } F(x_*) = 0 \in \mathbb{R}^n$$

Miminization

Suppose we have:

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

then the problem of minimization is given by:

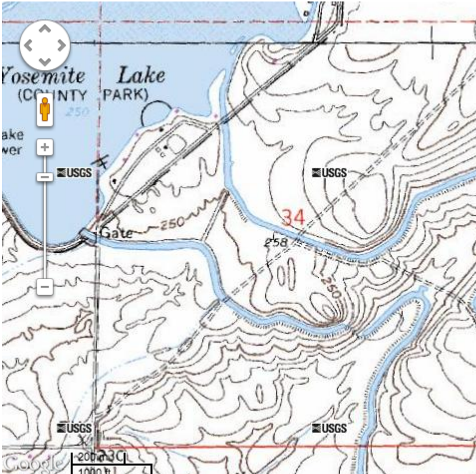
$$\text{find } x_* \in \mathbb{R}^n \text{ such that } f(x_*) \leq f(x), \ \forall x \in \mathbb{R}^n$$

Remark. Another variation of this is the constrained version, i.e.

$$\min_{x \in \Omega \subset \mathbb{R}^n} f : \mathbb{R}^n \rightarrow \mathbb{R}$$

Some important terminology

- Level sets (curves) are set of points where $f(x) = c$, for some constant; think topographic map
- gradient $g(x) = \nabla f(x)$ is the (column) vector of first derivatives of f
- Jacobian $J(x) = \nabla F^T(x)$ is the matrix of first derivatives of F
- Hessian $H(x) = \nabla^2 f(x)$ is the matrix of second derivatives of f



Motivation

Main approach to solving nonlinear equations/optimization can be viewed in several ways. We will consider the model-based approach

- Replace your nonlinear problem with a “model”
- Solve resulting model system
- Check for convergence, iterate
- So for example:
  - Nonlinear equations can be replaced by a linear problem
  - Minimization problem can be replaced by a quadratic model

Example 1

Consider the problem of solving one equation in one unknown, $f(x) = 0$. First replace $f(x)$ by a linear model $M_c(x)$:

$$M_c(x) = f(x_c) + f'(x_c)(x - x_c).$$

Solving for $M_c(x) = 0$ we get:

$$\begin{aligned} f(x_c) + f'(x_c)(x - x_c) &= 0 \\ x - x_c &= -\frac{f(x_c)}{f'(x_c)} \\ x &= x_c - \frac{f(x_c)}{f'(x_c)} \end{aligned}$$

Question: What does M_c look like in the general $n-$dimensional case?

Example 2

Consider the problem of minimizing a general nonlinear function, $f(x)$. First replace $f(x)$ by a quadratic model:

$$m_c(x) = f(x_c) + f'(x_c)(x - x_c) + \frac{1}{2}f''(x_c)(x - x_c)^2$$

Remark. We could have just as easily said that we would use our previous approach to solving $f'(x) = 0$!

Theory

- Under some fairly standard assumptions one can show convergence for Newton's method applied to NLE
  - $F(x)$ is continuously differentiable
  - A solution x_* exists s.t. $F(x_*) = 0$
  - $J(x)$ is Lipschitz continuous and $J(x_*)^{-1}$ exists
- For Newton's method one can show quadratic convergence, i.e.

$$||x_{k+1} - x_*|| \leq C||x_k - x_*||^2$$

- For variations on Newton's method still have fast convergence, i.e. superlinear convergence

Practicalities

- Many variations of based optimization based on problem characteristics
  - Convex
  - Non-smooth
  - Large-scale
  - Stochastic
  - ...
- Many science and engineering problems do not have the usual analytic assumptions
  - Smoothness
  - Availability of derivatives
  - Infinite precision

Class Participation

Question: Discuss the advantages and disadvantages to the model-based approach

Starter questions

Class Participation

Question: Discuss the advantages and disadvantages to the model-based approach

Starter questions

- What assumptions are we making, relying on?

Class Participation

Question: Discuss the advantages and disadvantages to the model-based approach

Starter questions

- What assumptions are we making, relying on?
- How do you choose a good model?

Class Participation

Question: Discuss the advantages and disadvantages to the model-based approach

Starter questions

- What assumptions are we making, relying on?
- How do you choose a good model?
- What does this look like in $\mathbb{R}^n$?

Class Participation

Question: Discuss the advantages and disadvantages to the model-based approach

Starter questions

- What assumptions are we making, relying on?
- How do you choose a good model?
- What does this look like in $\mathbb{R}^n$?
- What does convergence mean?

Advantages/Disadvantages of Newton's Method

Advantages
1. q-quadratic convergence if $J(x_*)$ is nonsingular
2. superlinear convergence for other variations of Newton's method
3. Exact solution in 1 iteration for affine F
Disadvantages
1. Not globally convergent for all problems
2. Requires $J(x_k)$ at each iteration
3. Need solution to system of linear equations at each iteration

Table: Newton's method

Numerical Methods for Unconstrained Optimization and Nonlinear Equations, J.E. Dennis, Jr. and Robert B. Schnabel, Prentice-Hall

Summary

- Newton-based methods provide some of the most popular and powerful methods for solving nonlinear equations and optimization
- Excellent software available
- Most practical problems do not lend themselves easily to standard set of assumptions

References

- **A Rapidly Convergent Descent Method for Minimization**, R. Fletcher and M.J.D. Powell