Optimization

Mario Gonzalez

Department of Economics
University of California, Santa Cruz

October 29, 2014

Outline

Introduction

- There is close relationship between the finite dimensional optimization problems and the root-finding and complementarity problem discussed in the previous session.
- The objective functions in optimization problems may be used in root-finding and complementary methods to determine whether iterations are converging on a solution.

Derivative Free Methods

- These methods simply place smaller brackets around a local maximum on a univariate function.
- Pros
  - They do not require evaluation of derivative functions
  - Guaranteed to find a local optimum
- Con
  - Slow

Derivative Free Methods: Golden Search Method

Basic Algorithm

- Suppose we wish to find a local maximum of $f(x)$ on the interval $[a, b]$.
- Pick two numbers x_1 x_2 in the interior of the interval, where $x_1 < x_2$.
- Evaluate the function and replace one of the end points of the interval by one these points, then a new interval is $[a, x_2]$.
- A local maximum must be contained in the new interval, otherwise the local maximum is at the end points.
- Repeat this procedure by choosing progressively smaller intervals

- Two criteria:
  - Length of the new interval should be independent of whether the upper bound or lower bound is replaced.
  - On successive iterations, one should be able to reuse an interior point from the previous iteration so that only one new function evaluation is performed per iteration.
- This two criteria are met by selecting

$$\mathrm{x_i} = \mathrm{a} + \alpha_{\mathrm{i}}\,(\mathrm{b} - \mathrm{a}) \qquad \text{where}$$

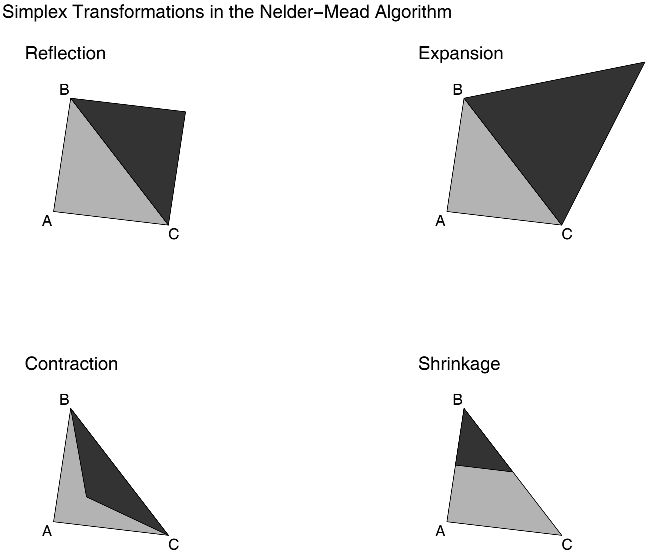
$$\alpha_1 = \frac{3 - \sqrt{5}}{2} \text{ and } \alpha_2 = \frac{\sqrt{5} - 1}{2}$$

- where α_2 is known as the Golden Ratio.

- The algorithm begins by evaluating the objective function at n+1 points
  - This way it forms a simplex in the n-dimensional decision space.
- At each iteration the algorithm determines the point on the simplex with the lowest function and alters that point by reflecting it through the opposite face of the simplex.
- If the reflection is successful and finds a new point that is higher than all the others in the simplex it continues expanding in this direction.
- If the reflection is not successful it contracts it contracts the simplex
- If contracting is not successful, the algorithm shrinks the entire simplex toward the best point.

Derivative Free Methods: Nelder-Mead Algorithm

Reflection, Expansion, Contraction and Shrinkage



Derivative Free Methods: Nelder-Mead Algorithm

- The method is simple; however, is slow and unreliable
- Is good for for problems that involve a single optimization or costly function and derivative evaluations.
- Better not to use it if the problem involves repeated iterations.

- Uses successive quadratic approximations to the objective hoping that the maxima of the approximations will converge to the maximum of the objective
- The method starts with a guess for the maximum of f.
- Given $x^{(k)}$, the following iteration $x^{(k+1)}$ is computed by maximizing the second order taylor approximation to f about $x^{(k)}$ i.e.

  $$f(x) \approx f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) + \frac{1}{2}\left(x - x^{(k)}\right)^{\top} f''(x^{(k)})(x - x^{(k)})$$

  and taking maximizing with respect to $\left(x - x^{(k)}\right)$ leads to

  $$f'\left(x^{(k)}\right) + f''\left(x^{(k)}\right)\left(x - x^{(k)}\right) = 0$$

- Rearranging leads to the iteration rule or "Newton Step"

  $$x^{(k+1)} \leftarrow x^{(k)} - \left[f''\left(x^{(k)}\right)\right]^{-1} f'\left(x^{(k)}\right)$$

Newton Raphson Method

- This method is rarely used and only if the objective function is globally concave
  - This method converges if f is twice continuously differentiable and if the initial guess is sufficiently close to a local maximum
  - No general formula for determining what is "sufficiently close"
  - This method can be very sensitive to starting point if function is not globally concave
- Drawbacks:
  - The algorithm requires computation of first and second derivatives
  - There is no guarantee that the objective function increases in the direction of the Newton Step

- Similar to Newton-Raphson method
- Replace the Hessian of the objective function with a negative definite approximation⇒this guarantees that value function increases in the direction of the Newton Step
- In this case the quasi-Newton step or search direction is of the form

$$
\begin{aligned}
d^{(k)} &= -\left[f''\left(x^{(k)}\right)\right]^{-1} f'\left(x^{(k)}\right) \\
d^{(k)} &= -A^{(k)} f'\left(x^{(k)}\right)
\end{aligned}
$$

- These methods avoid using second derivative, to ease implementation and calculation
- These methods differ in the way they approximate the Hessian approximation is constructed
- Method of the Steepest Ascent:
  - This method assumes that $A^k = -I$, then

$$d^{(k)} = f'\left(x^{(k)}\right)$$

    which leads to a Newton Step equal to the gradient of the objective function

- These methods face the same problems as the Newton-Raphson Method

-

- Davidson-Fletcher-Powell (DFP) method uses the updating scheme
$$A \leftarrow A + \frac{dd^\top}{d^\top u} - \frac{Auu^\top}{u^\top Bu},$$
where $d = x^{(k+1)} - x^{(k)}$ and $u = f'\left(x^{(k+1)}\right) - f'\left(x^{(k)}\right)$

- Broyden-Fletcher-Goldfrab-Shano (BFGS) method uses the updating scheme
$$A \leftarrow A + \frac{1}{d^\top u}\left(wd^\top + dw^\top - \frac{w^\top u}{d^\top u}dd^\top\right),$$
where $w = d - Au$

Line Search Methods

- Quasi-Newton methods results can be improved if the Newton Step is used as a search direction and stop short or move pass it.
- Two schemes are typically used:
  - Armijo Search: find the minimum power j such that

    $$\frac{f(x + sd) - f(x)}{s} \geq \mu f'(x)^\top d$$

    where $s = \rho^j$ and $0 < \mu < 0.5$.
  - Goldstein Search: similarly, the idea is to find a value of s such that

    $$\mu_0 f'(x)^\top d \leq \frac{f(x + sd) - f(x)}{s} \leq \mu_1 f'(x)^\top d$$

    where $0 < \mu_0 < 0.5 < \mu_1 < 1$.

Constrained Optimization

- Constrained optimization problems in economics do not come naturally in the form of root-finding or fixed point problems
  - The methods previously shown are not directly applicable
- These problems come naturally presented as a *Complementarity Problem*

- Reminder from previous session
  - In the complementarity problem two n-vectors a and b, with $a < b$ and a function f are given, then one must find an n-vector $x \in [a, b]$, that satisfies

$$\begin{aligned} x_i &> a_i \Rightarrow f_i(x) \geq 0 \;\; \forall i = 1, \ldots n \\ x_i &< b_i \Rightarrow f_i(x) \leq 0 \;\; \forall i = 1, \ldots n \end{aligned}$$

  - Therefore, it includes the root-finding problem when $a = -\infty$ and $b = \infty$.
  - However, the complementarity problem is not to find a root that lies within specified bounds.

- However, Complementarity Problems can be stated as a root-finding problems if they solve

$$\tilde{f}(x) = \min\left(\max\left(f'(x), a - x\right), b - x\right) = 0$$

- Once reformulated as a root-finding problem it can be solved using root-finding algorithms

Constrained Optimization
Example

$$\max_{c,h} \frac{c^{1-\sigma}}{1-\sigma} - \alpha \frac{h^{1+\gamma}}{1+\gamma} \text{ subject to}$$

$$\begin{aligned} c &= wh + z, \\ 0 &\leq h \leq 1, \\ c &\geq 0. \end{aligned}$$

Then the Kuhn Tucker conditions are

$$(\mathrm{wh} + \mathrm{z})^{-\sigma}\,\mathrm{w} - \alpha \mathrm{h}^{\gamma} + \lambda - \delta = 0$$

$$\mathrm{h} \geq 0, \quad \lambda \geq 0, \quad \lambda \mathrm{h} = 0$$

$$\mathrm{h} \leq 1, \quad \delta \geq 0, \quad \delta\,(\mathrm{h} - 1) = 0$$

Then in the complementarity problem form, and $\mathrm{h} \in [0,1]$ that satisfies the Kuhn-Tucker conditions will also satisfy

$$\begin{aligned} \mathrm{h} &> 0 \Rightarrow \left(\mathrm{wh} + \mathrm{z}\right)^{-\sigma} \mathrm{w} - \alpha \mathrm{h}^{\gamma} + \lambda - \delta \geq 0 \\ \mathrm{h} &< 1 \Rightarrow \left(\mathrm{wh} + \mathrm{z}\right)^{-\sigma} \mathrm{w} - \alpha \mathrm{h}^{\gamma} + \lambda - \delta \leq 0 \end{aligned}$$

So, this is equivalent to finding an h that solves the root-finding problem

$$\min\left(\max\left(\mathrm{w}\left(\mathrm{wh} + \mathrm{z}\right)^{-\sigma} - \alpha \mathrm{h}^{\gamma}, -\mathrm{h}\right), 1 - \mathrm{h}\right) = 0$$

To implement the Newton method, the Jacobian matrix must be used, the ith row of the $\tilde{\mathrm{J}}$ can be written as

$$\tilde{\mathrm{J}}_{\mathrm{i}}\left(\mathrm{x}\right) = \begin{cases} \mathrm{J}_{\mathrm{i}}\left(\mathrm{x}\right), & \mathrm{a}_{\mathrm{i}} - \mathrm{x}_{\mathrm{i}} < \mathrm{f}_{\mathrm{i}}\left(\mathrm{x}\right) < \mathrm{b}_{\mathrm{i}} - \mathrm{x}_{\mathrm{i}} \\ -\mathrm{I}_{\mathrm{i}}, & \text{otherwise} \end{cases}$$

DEMO

Conclusion

- Derivative free methods guarantee that a solution is going to be achieved, but they are slow.
- Newton Raphson method is based on the optimization of second order linear expansion. It is rarely used as there is no guarantee that it will find a solution.
- Quasi Newton methods use approximations of the Hessian matrix of the objective function to guarantee that a solution is going to be found.
- Constrained optimization problem can be rearranged as a root-finding problem and use any of the methods previously presented.