Numerical Optimization 09: Direct Methods

Qiang Zhu

University of Nevada Las Vegas

May 20, 2020

Overview

Direct method

Direct methods rely solely on the objective function f. They are usually called

- zero-orther
- black box
- pattern search
- derivative free

The most important feature is that they do not rely on derivative information. They use other criteria to choose the next search direction to judge if the search is converged.

Cyclic Coordinate Search

This method simply alternate coordinate directions for its line search. The search starts from an initial $\boldsymbol{x}^1$ and optimize the first input.

$$\boldsymbol{x}^2 = \arg\min_{x_1} f(x_1, x_2^1, x_3^1, \cdots, x_n^1)$$

Then, it moves to the next coordinate,

$$\boldsymbol{x}^3 = \arg\min_{x_1} f(x_1^2, x_2, x_3^2, \cdots, x_n^2)$$

This process is equivalent to doing a sequence of line searches along the set of n basis vectors. It is terminated after no significant improvement is made.

Cyclic Coordinate Search

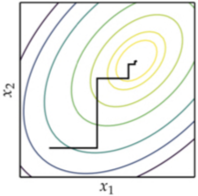
This method simply alternate coordinate directions for its line search. The search starts from an initial $\boldsymbol{x}^1$ and optimize the first input.

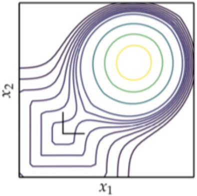
$$\boldsymbol{x}^2 = \underset{x_1}{\arg\min} f(x_1, x_2^1, x_3^1, \cdots, x_n^1)$$

Then, it moves to the next coordinate,

$$\boldsymbol{x}^3 = \underset{x_1}{\arg\min} f(x_1^2, x_2, x_3^2, \cdots, x_n^2)$$

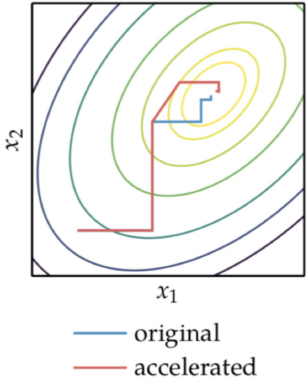
This process is equivalent to doing a sequence of line searches along the set of n basis vectors. It is terminated after no significant improvement is made.





Acceleration

Similar to the momentum method in the gradient descent, the cyclic method can be augmented with an acceleration step to help traverse diagonal valleys. For each full cycle starting with $\boldsymbol{x}^1$ from 1 to n, an additional line search is conducted along with the direction of $\boldsymbol{x}^{n+1} - \boldsymbol{x}^1$.



Powell's method

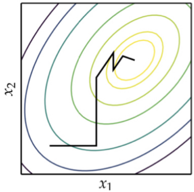
This algorithm maintains a list of search directions $\boldsymbol{u}^1, \cdots, \boldsymbol{u}^n$, which are initially the basis vectors. Starting at $\boldsymbol{x}^1$, Powell's method conduct a line search for each direction, updating the design point each time, Then shift each u by one index and drop u^1. The last direction is replaced with the direction of $\boldsymbol{x}^{n+1} - \boldsymbol{x}^1$.

$$\boldsymbol{x}^{i+1} \leftarrow \text{ line search}(f, \boldsymbol{x}^i, \boldsymbol{u}^i) \text{ for all} iin$$

$$\boldsymbol{u}^{i+1} \leftarrow \boldsymbol{u}^{i+1}$$

for all$iin1, \cdots, n-1$

$$\boldsymbol{u}^n \leftarrow \boldsymbol{x}^{n+1} - \boldsymbol{x}^n$$



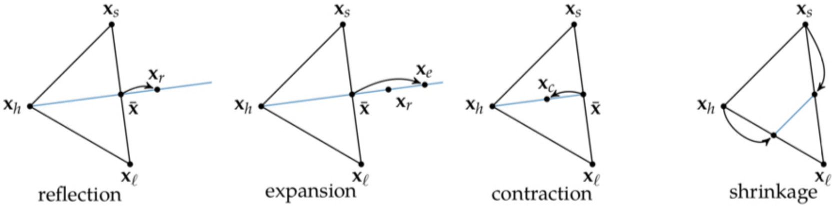
Powell showed that for quadratic functions, after k full iterations the last k direction will be mutually conjugate. It is recommended to reset every n or $n+1$ iterations.

Nelder-Mead Simplex Method

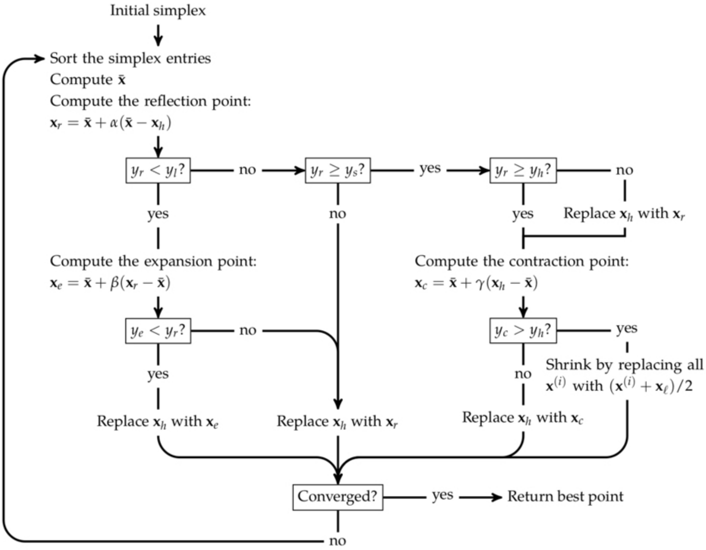
The Nelder-Mead simplex method uses a simplex to traverse the space in search of a minimum. A simplex is a $n+1$-vertices polyhedron in n-dimensional space.

- x_h, pt of highest f,
- x_s, pt of 2nd highest f,
- x_l, pt of lowest f,
- $\bar{x}$, mean pt excluding x_h.

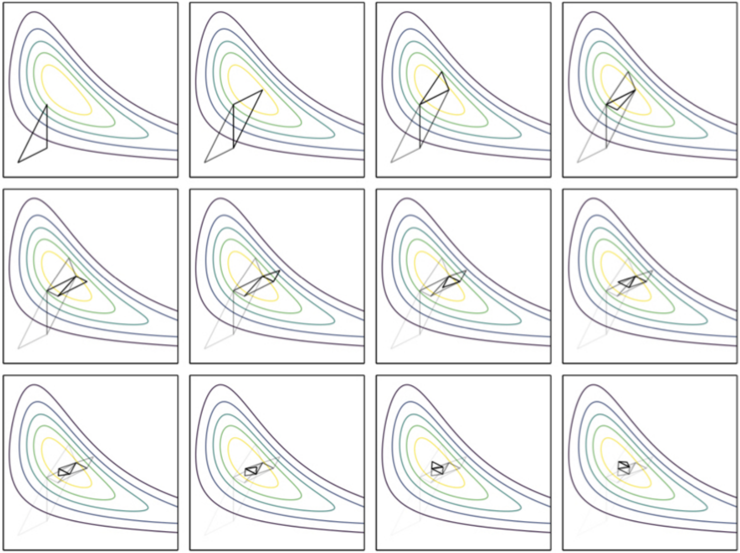
- Reflection. $x_r = \bar{x} + (\bar{x} - x_h)$,
- Expansion. $x_e = \bar{x} + 2(x_r - \bar{x})$,
- Contraction. $x_c = \bar{x} + 0.5(x_h - \bar{x})$,
- Shrinkage, halving the distance to x_l.



Nelder-Mead Simplex Algorithm



Nelder-Mead Simplex method in practice



Summary

- Direct methods rely solely on the objective function and do not use derivative information.
- Cyclic coordinate search optimizes one coordinate direction at a time.
- Powells method adapts the set of search directions based on the direction of progress.
- The Nelder-Mead simplex method uses a simplex to search the design space, adaptively expanding and contracting the size of the simplex in response to evaluations of the objective function.