AI505

Optimization

Derivative-Free Methods

Marco Chiarandini

Department of Mathematics & Computer Science
University of Southern Denmark

Outline

1. Nelder-Mead Simplex Method

2. Divided Rectangles

Derivative-Free Methods

- Also called **direct search methods**, zero-order, black box, pattern search
- Direct method search using function evaluations only

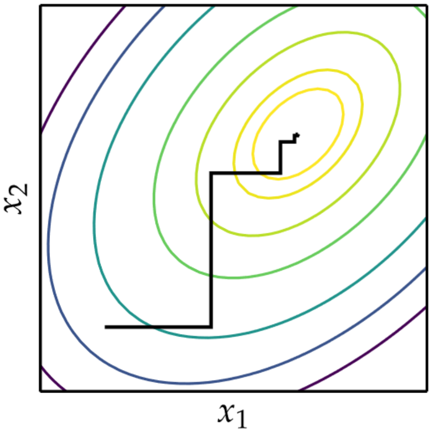
Cyclic Coordinate Search

- Also known as coordinate descent, or taxicab search

- Performs line search in alternating coordinate directions

$$x_{1,1} = \text{argmin}_{x_1} f(x_1, x_{2,0}, x_{3,0}, \ldots, x_{n,0})$$

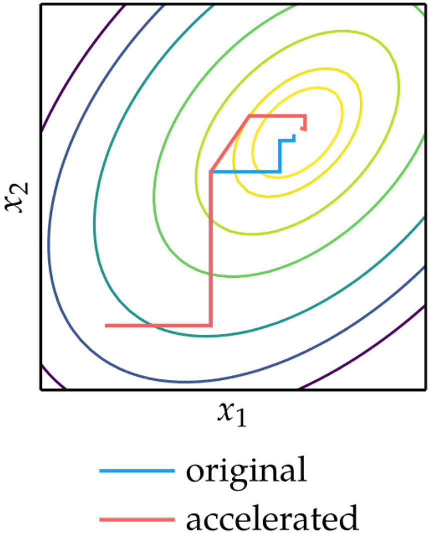
$$x_{2,1} = \text{argmin}_{x_2} f(x_{1,1}, x_2, x_{3,1}, \ldots, x_{n,1})$$



Cyclic Coordinate Search

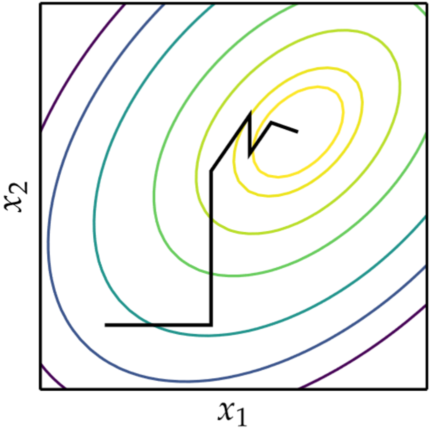
- Can be augmented to accelerate convergence

- For every full cycle starting with optimizing $\boldsymbol{x}_1$ along $[1, 0, \ldots, 0]$ and ending with $\boldsymbol{x}_{n+1}$ after optimizing along $[0, 0, \ldots, 1]$, an additional line search is conducted along the direction $\boldsymbol{x}_{n+1} - \boldsymbol{x}_1$.



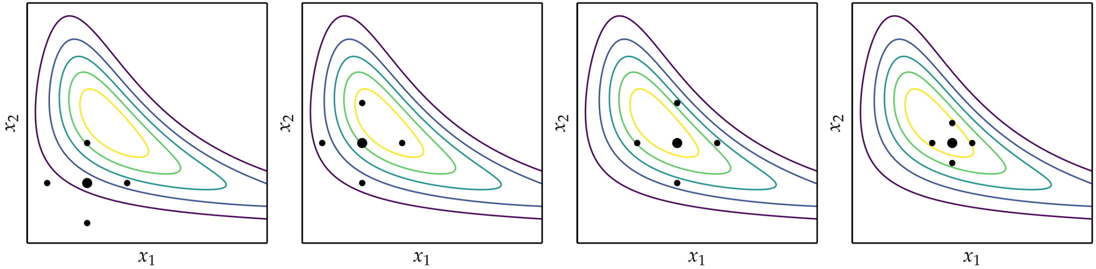
Powell's Method

- Similar to Cyclic Coordinate Search, but can search in non-orthogonal directions
- Drops the oldest search direction in favor of the overall direction of progress
- It can lead the search directions to become linearly dependent and the search directions can no longer cover the full design space, and the method may not be able to find the minimum



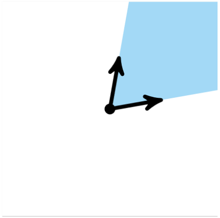
Hooke-Jeeves

- Evaluate $f(\boldsymbol{x})$ and $f(\boldsymbol{x} \pm \alpha \boldsymbol{e}_i)$ for a given step size α in every coordinate direction from an anchoring point $\boldsymbol{x}$.
- It accepts any improvement it may find.
- If no improvements are found, it decreases the step size.
- The process repeats until the step size is sufficiently small.
- $2n$ evaluations for an n-dimensional problem

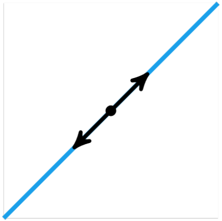


Generalized Pattern Search

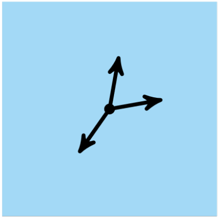
- Generalization of Hooke-Jeeves method
- A pattern P can be constructed from a set of directions D about an anchoring point $\boldsymbol{x}$ with a step size α according to: $P = \{\boldsymbol{x} + \alpha\boldsymbol{d} \text{ for each } \boldsymbol{d} \in D\}$
- Searches in set of directions that **positively spans** (nonnegative linear combination) search space. (if D has full row rank and if $D\boldsymbol{x} = -D1$ with $\boldsymbol{x} \geq 0$



only positively spans the cone

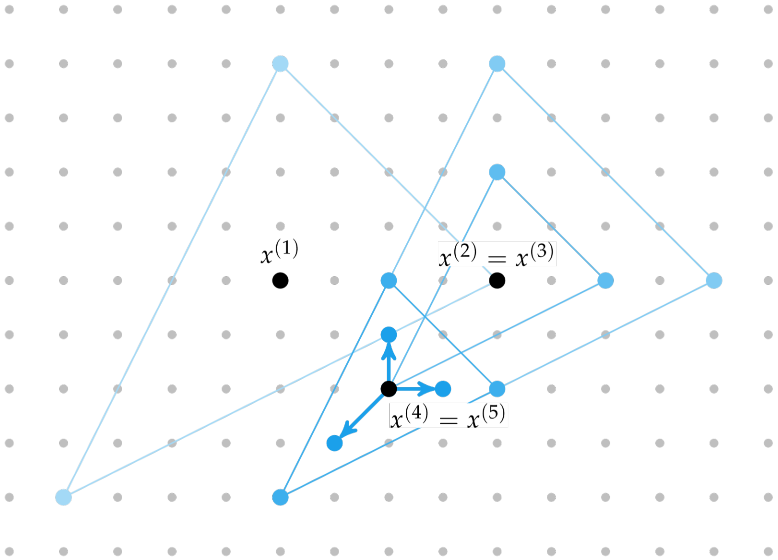


only positively spans 1d space



positively spans $\mathbb{R}^2$

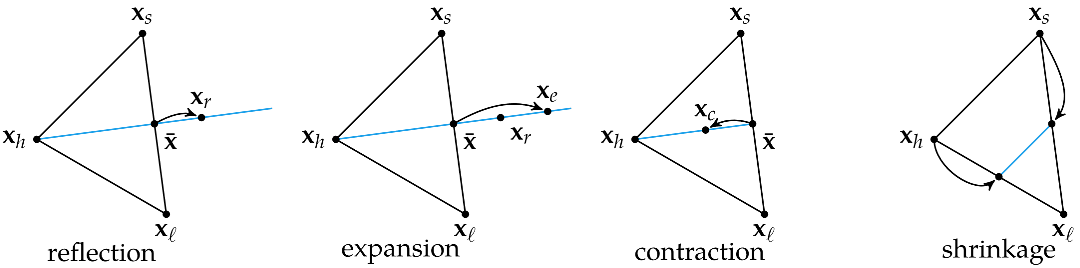
Generalized Pattern Search

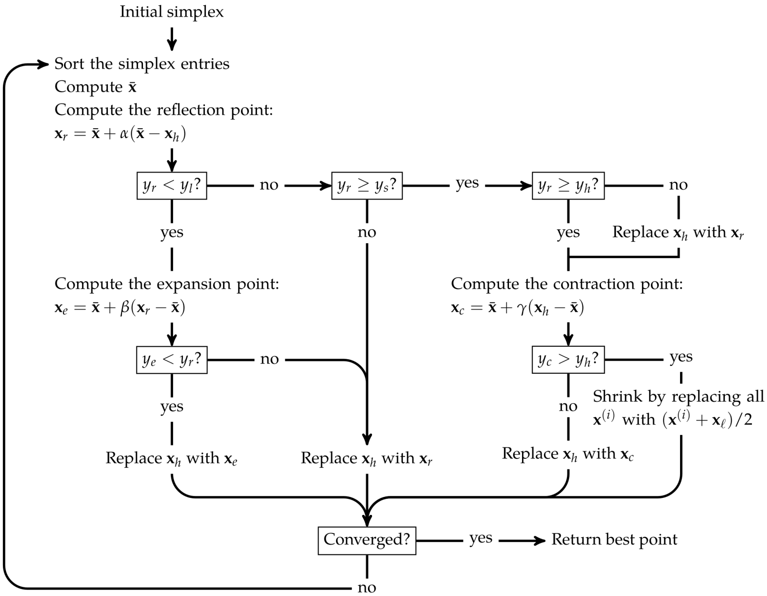


Outline

Nelder-Mead Simplex Method

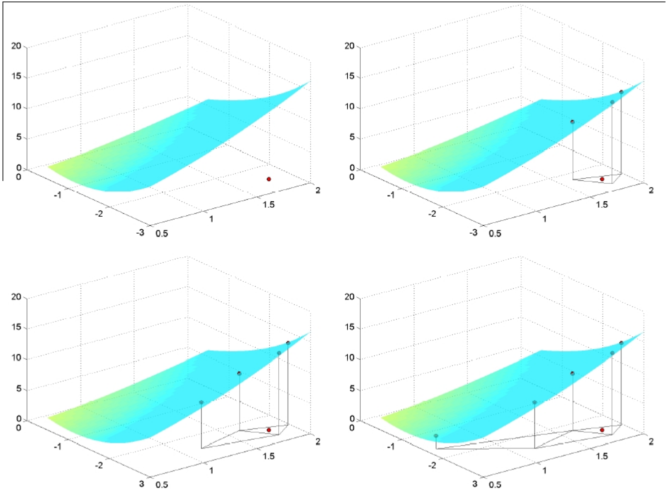
Uses simple algorithm to traverse search space using set of points defining a simplex



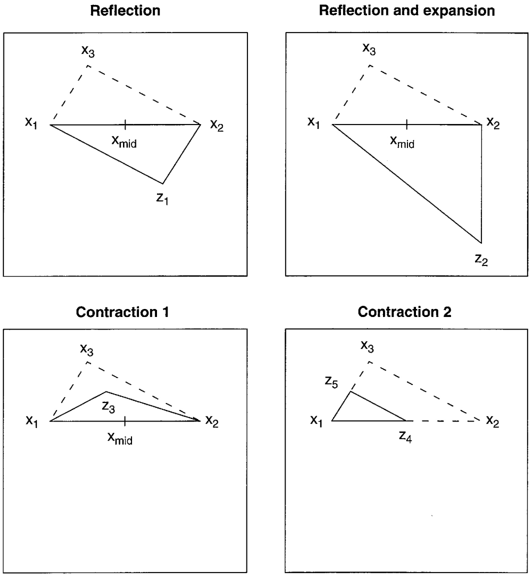

Initial simplex
Sort the simplex entries
Compute $\bar{\mathbf{x}}$
Compute the reflection point:
$\mathbf{x}_r = \bar{\mathbf{x}} + \alpha(\bar{\mathbf{x}} - \mathbf{x}_h)$
$y_r < y_l$?
no
$y_r \geq y_s$?
yes
$y_r \geq y_h$?
no
yes
no
yes
Replace $\mathbf{x}_h$ with $\mathbf{x}_r$
Compute the expansion point:
$\mathbf{x}_e = \bar{\mathbf{x}} + \beta(\mathbf{x}_r - \bar{\mathbf{x}})$
Compute the contraction point:
$\mathbf{x}_c = \bar{\mathbf{x}} + \gamma(\mathbf{x}_h - \bar{\mathbf{x}})$
$y_e < y_r$?
no
$y_c > y_h$?
yes
yes
no
Shrink by replacing all
$\mathbf{x}^{(i)}$ with $(\mathbf{x}^{(i)} + \mathbf{x}_\ell)/2$
Replace $\mathbf{x}_h$ with $\mathbf{x}_e$
Replace $\mathbf{x}_h$ with $\mathbf{x}_r$
Replace $\mathbf{x}_h$ with $\mathbf{x}_c$
Converged?
yes
Return best point
no

Nelder-Mead

Simplex based method [Spendley et al. (1962)]

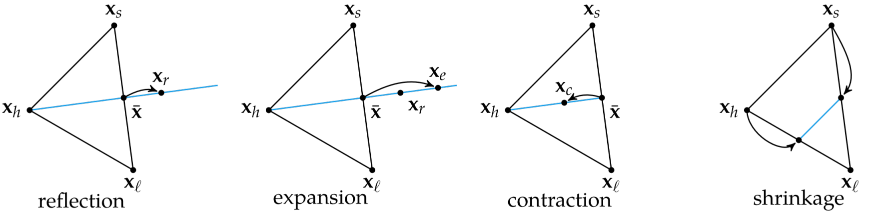


Nelder-Mead (cont.)



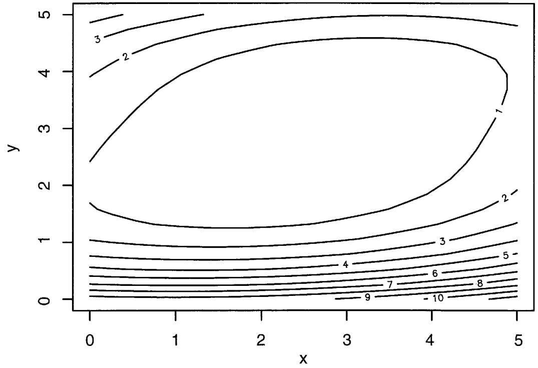
Nelder-Mead (cont.)

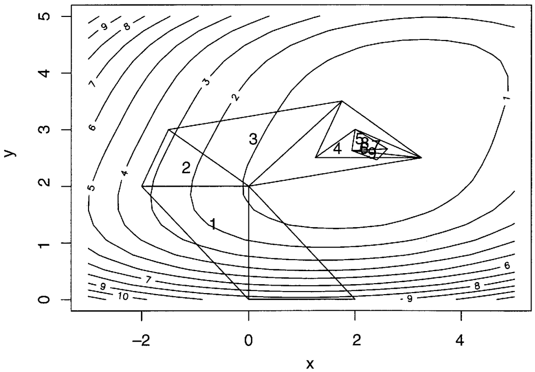
Nelder-Mead simplex method [Nelder and Mead, 1965]:

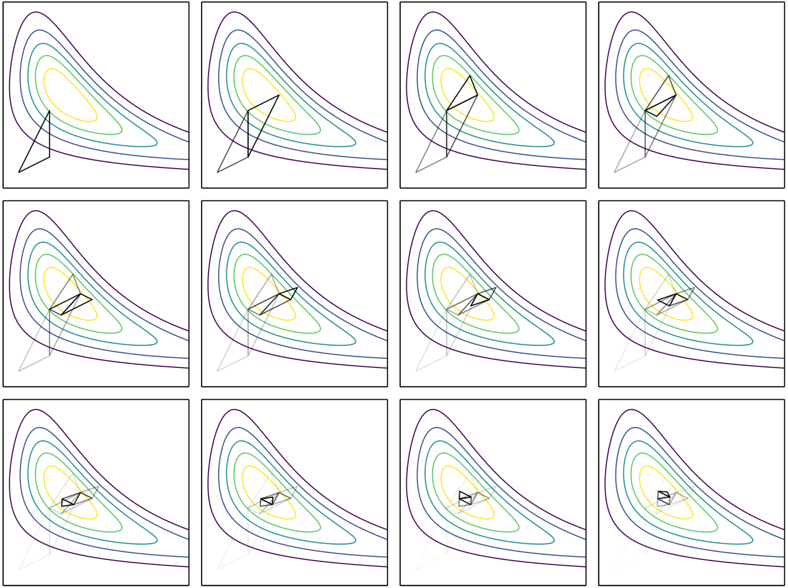


Nelder-Mead (cont.)

Example:







Nelder-Mead (cont.)

Algorithm: Simplex search

Let $x_1, \ldots, x_{n+1}$ be vertices of a simplex

Let h: $y_h = \max_i y_i = \max f(x_i)$ and

l: $y_l = \min_i y_i = \min f(x_i)$

Let $\bar{x}$ be the centroid of points with $i \neq h$ and

$d(x_i, x_j)$ the distance between two points x_i and x_j

$(k \leftarrow 0)$

Reflect iter. k: $(k \leftarrow k+1)$ Generate the reflection x_R of x_h

Case 1 **if** $y_l \leq y_R < y_h$ then $x_h \leftarrow x_R$ and go to **Reflect**

Case 2 **else if** $y_R < y_l$ **then**

generate the expansion x_E of x_h

if $y_E < y_l$ **then** $x_h \leftarrow x_E$ and go to **Reflect**

else $x_h \leftarrow x_R$ and go to **Reflect**

Case 3 **else if** $y_R > y_i, \forall i \neq h$ **then** $x_h \leftarrow \min\{y_h, y_R\}$ and generate the contraction 1

if $y_C \leq \min\{y_h, y_R\}$ then $x_h \leftarrow x_C$ and go to **Reflect**

else contraction 2 $x_i \leftarrow (x_i + x_l)/2$

Outline

1. Nelder-Mead Simplex Method

2. Divided Rectangles

DIRECT – Divided Rectangles

- Also called DIRECT for Divided RECTangles
- Same approach as Shubert-Piyavskii method, but does not need to specify Lipschitz constant and is more efficiently expanded to multiple dimensions

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.
- Powell's method adapts the set of search directions based on the direction of progress.
- Hooke-Jeeves searches in each coordinate direction from the current point using a step size that is adapted over time.
- Generalized pattern search is similar to Hooke-Jeeves, but it uses fewer search directions that positively span the design space.
- 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.
- The divided rectangles algorithm extends the Shubert-Piyavskii approach to multiple dimensions and does not require specifying a valid Lipschitz constant.