Optimization

Gursimran Singh Lnu, Yang Liu

Methods Implemented

- Gradient Descent
- Newton’s Method
- Quasi-Newton (BFGS)
- Conjugate Gradient Descent
- Nelder-Mead Method

Problem for function 2

$$f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} - \sum_{i=1}^{m} \log(b_i - \mathbf{a}_i^T\mathbf{x}) \text{ where } m = 500 \text{ and } n = 100$$

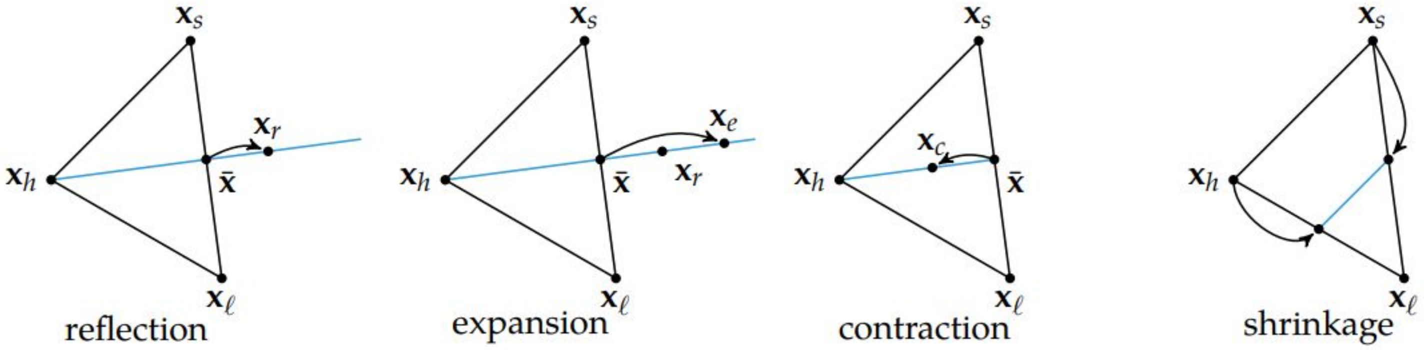
- Log can only accept positive values
- Optimization algorithm may step outside the domain.

Solution:

- Shrink the step size by half when such condition is met.

Nelder-Mead Method

Simplex operations



Algorithm for Optimization, M. J. Kochenderfer & T. A. Wheeler

Nelder-Mead Method

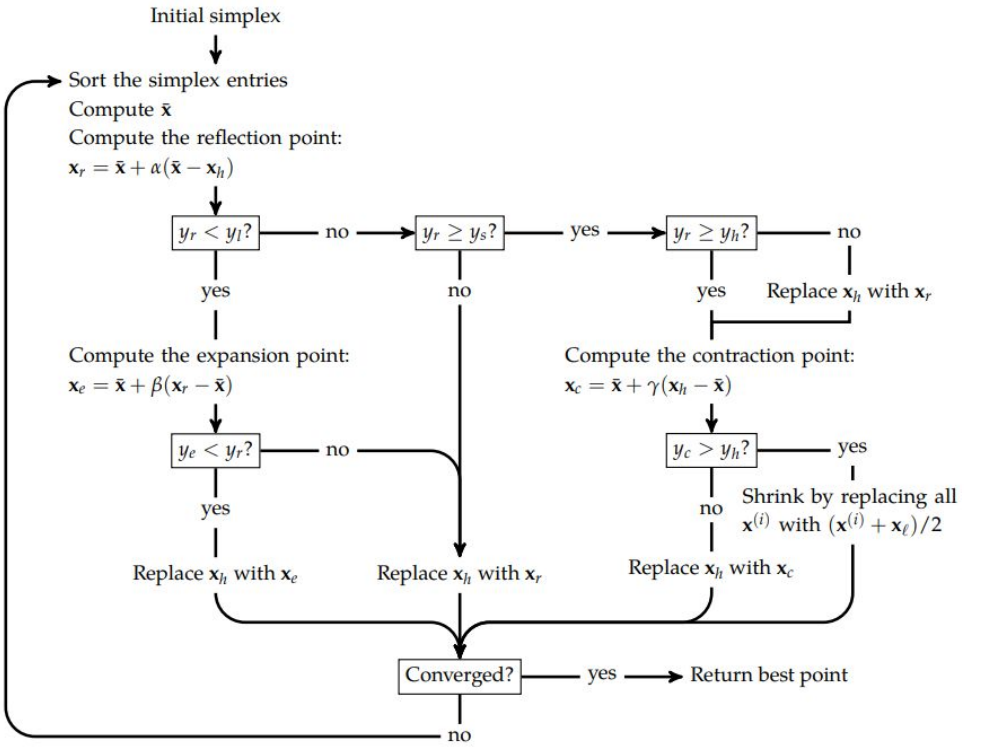
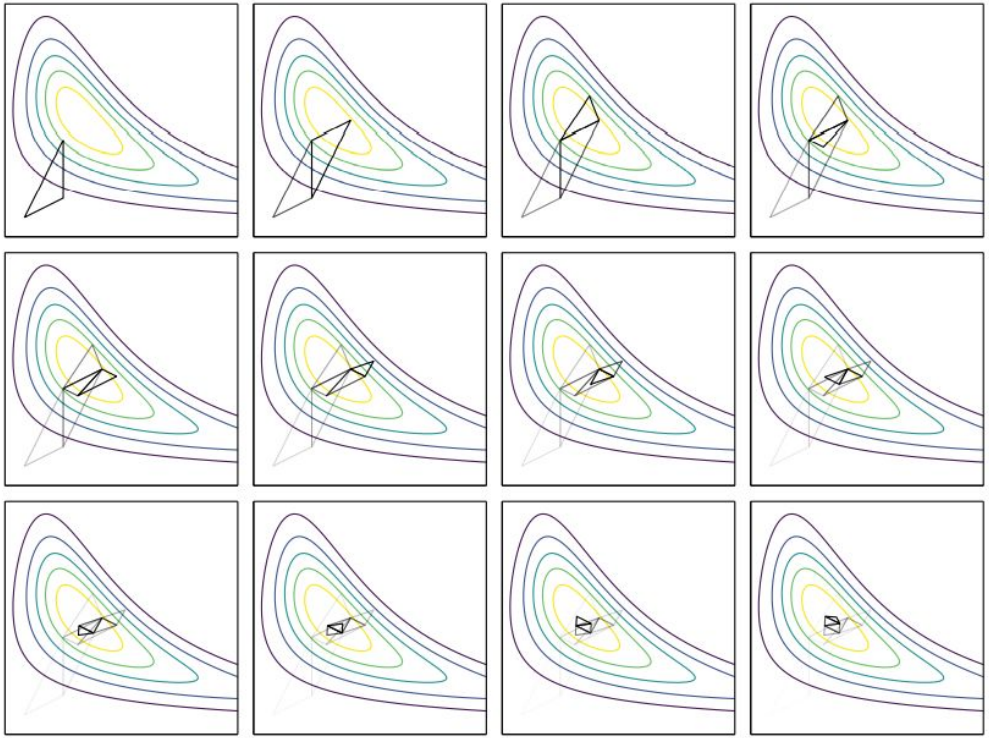


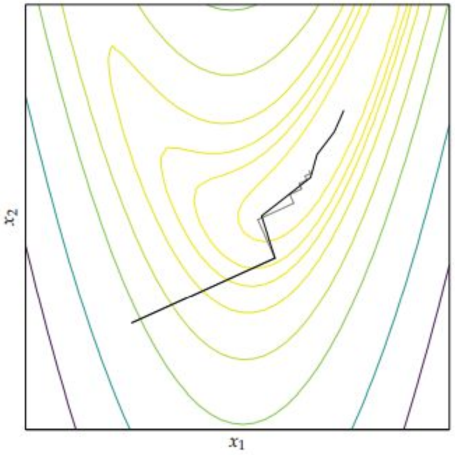
Figure 7.9. Flowchart for the Nelder-Mead algorithm.

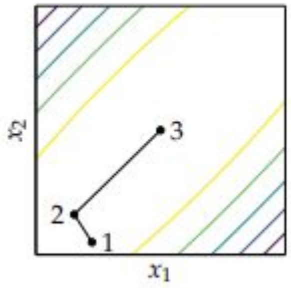
Nelder-Mead Method



Conjugate Gradient Descent

- Gradient Descent can perform poorly in narrow valleys.
- Similar to gradient descent, but differs in descent direction.
- It doesn't undo the previous steps, and optimizes in fewer steps.





Conjugate Gradient Descent

- The algorithm starts with the direction of steepest descent: **$d^{(1)} = -g^{(1)}$**
- We then use line search to find the next design point: **$x^{(2)} = x^{(1)} + \alpha^{(1)} d^{(1)}$**
- Subsequent iterations choose **$d^{(k+1)}$** based on the next gradient and a contribution from the current descent direction: **$d^{(k+1)} = -g^{(k+1)} + \beta^{(k)} d^{(k)}$**

Conjugate Gradient Descent

- Polak-Ribière:

  $\beta^{(k)} = g^{(k)\top} . (g^{(k)} - g^{(k-1)}) / g^{(k-1)} .^{\top} g^{(k-1)}$

- Fletcher-Reeves:

  $\beta^{(k)} = g^{(k)\top} g^{(k)} / g^{(k-1)\top} . g^{(k-1)}$

- Used to guarantee convergence - **$\beta \leftarrow \max(\beta, 0)$**

Summary

Functions	Function Type	Initial Values	Minimal value	Optimal solution	Achievable Algorithms
Fun 1	Convex	ones(100,1)	0	zeros(100,1)	Newton's, Nelder-mead
Fun 2	Non-Convex	zeros(100, 1)	-2.4432e3	[-1.1094 -1.0697 1.9717 -1.6617 0.4194 ...]	Newton's, BFGS
Fun 3	Non-Convex	[3, 3]	2.2775e-16	[1 1]	Newton's (All other methods close but not exact)