Nelder-Mead Method for Optimization

Stats 102A

Miles Chen

Department of Statistics

Week 8 Wednesday

UCLA

Section 1

Nelder-Mead Method

Nelder-Mead Method

Acknowledgments: Henri P. Gavin, Duke University

The Nelder-Mead is a numerical method to find the minimum (or maximum) of a multidimensional function.

It is a direct search method. This means it calculates the value of the function and compares it to other values. It does not use derivatives. The search method works well but is not guaranteed to converge.

Simplex

The Nelder-Mead method uses a simplex which is a shape consisting of $n + 1$ vertices in n dimensions.

- For a 2-dimensional function, the simplex is a triangle.
- For a 3-dimensional function, the simplex is a tetrahedron.
- For a 4-dimensional function, the simplex is a 5-cell Pentachoron https://en.wikipedia.org/wiki/5-cell
- and so on

For today's explanation, I will use the 2-dimensional triangle to illustrate the method.

Nelder-Mead method

We have some function in 2D space, $f(x, y)$. We wish to find a local minimum.

We begin with three arbitrarily selected points (in 2D) to form our simplex.

The steps of the method are as follows:

1. Sort
2. Reflect
3. Extend
4. Contract
5. Shrink
6. Check convergence

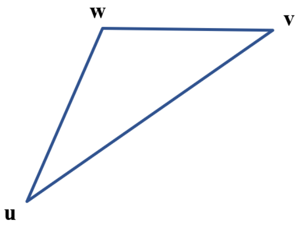
1. Sort

Evaluate the function f at the three points of the simplex.

Sort and label the three points $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ according to the functions value. That is:

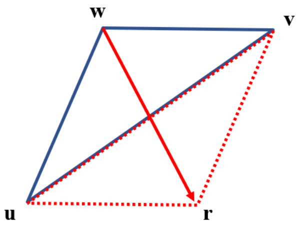
$$f(\mathbf{u}) < f(\mathbf{v}) < f(\mathbf{w})$$

According to the function, $\mathbf{u}$ is the best performing point (gives the smallest value of the three) and $\mathbf{w}$ is the worst performing point.



2. Reflect

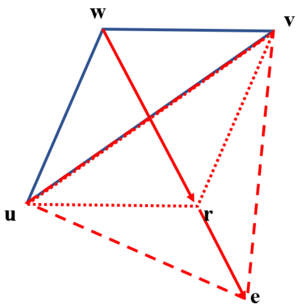
Reflect the worst point $\mathbf{w}$ through the centroid of the remaining points $\mathbf{u}$ and $\mathbf{v}$ to obtain a reflected point $\mathbf{r}$. Evaluate $f(\mathbf{r})$.



If the cost at the reflected point $f(\mathbf{r})$ is better than $f(\mathbf{v})$ but not better than $f(\mathbf{u})$, then replace $\mathbf{w}$ with $\mathbf{r}$. With this change, $\mathbf{v}$ will become the worst performing point in the simplex. Go to step 6 to check convergence.

3. Extend (if f(r) < f(u))

If the cost at the reflected point, $f(\mathbf{r})$ is better than $f(\mathbf{u})$, then extend the reflected point $\mathbf{r}$ even further to $\mathbf{e}$. Evaluate $f(\mathbf{e})$. (If $\mathbf{r}$ is better than $\mathbf{u}$, this step attempts to be a tiny bit greedy by trying something even further in that direction.)



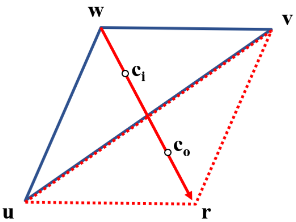
If the cost at the extended point $f(\mathbf{e})$ is even better than the cost at the reflected point, $f(\mathbf{r})$, replace the worst point $\mathbf{w}$ with the extended point $\mathbf{e}$. Go to step 6 to check convergence.

If the cost at the extended point $f(\mathbf{e})$ is not better than the cost at the reflected point, $f(\mathbf{r})$, replace the worst point $\mathbf{w}$ with the reflected point $\mathbf{r}$. Go to step 6 to check convergence.

4. Contract

If the checks in step 2 or step 3 are not satisfied, then the reflected point $\mathbf{r}$ is not better than $\mathbf{v}$ or $\mathbf{u}$ and it would be pointless trying to improve the simplex by using the reflected point $\mathbf{r}$ as it would only become the new worst performing point. Instead, contract the worst point $\mathbf{w}$ to points $\mathbf{c}_i$ and $\mathbf{c}_o$ along the line of reflection.

- $\mathbf{c}_i$ is the inside contracted point and is $1/4$ of the way between $\mathbf{w}$ and $\mathbf{r}$.
- $\mathbf{c}_o$ is the outside contracted point and is $3/4$ of the way between $\mathbf{w}$ and $\mathbf{r}$.

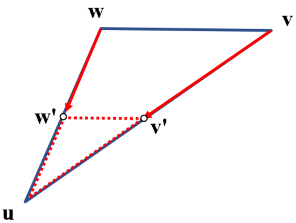


Evaluate the function at both contracted locations. If either of these points perform better than $f(\mathbf{v})$, then replace $\mathbf{w}$ with the better performing point. Go to step 6 to check for convergence.

5. Shrink

If neither of the contracted points perform better than $\mathbf{v}$, then we shrink the simplex towards the best performing point $\mathbf{u}$.

- Replace $\mathbf{v}$ with $\mathbf{v}'$ a point halfway between $\mathbf{u}$ and $\mathbf{v}$.
- Replace $\mathbf{w}$ with $\mathbf{w}'$ a point halfway between $\mathbf{u}$ and $\mathbf{w}$.



6. Check Convergence

Check convergence to see if the program should terminate.

One method is to use the sample standard deviation of the function values of the simplex. If the standard deviation is below some tolerance, the program quits and the best performing point is returned as a proposed optimum.

Animated Demo

animated demo

https://www.youtube.com/watch?v=j2gcuRVbwR0

optim()

R's `optim()` function can perform optimization of functions. It employs several optimization methods from which you can choose. The default method is the Nelder-Mead method.