3. UNCONSTRAINED MULTIVARIATE OPTIMIZATION

Problem: minimize $f(\mathbf{x})$
subject to $\mathbf{x} = (x_1,...,x_n) \in \mathbb{R}^n$

Algorithms can be divided into two categories:
1) Direct methods which are based on function comparisons and don't need derivatives or differentiability.
2) Derivative-based methods: gradient based methods and second derivative methods.

3.1 METHODS FOR NON-SMOOTH FUNCTIONS

Let f be non-smooth i.e. non-diffrentiable at some points. The preceding conditions cannot be used without precaution for verifying a minimum point.

Methods must be based on function comparisons.

Elimination type methods tend to be very inefficient ("curse of dimensionality"). A method using function comparisons should be used only when there is no suitable alternative method.

On the other hand, most of the methods are easy to implement and some may be used for global minimization.

3.1.1. THE POLYTOPE METHOD (THE SIMPLEX METHOD OF NELDER AND MEAD)

A *simplex* in n-dimensional space is a polytope or polygon with n+1 vertices. At each stage of the algorithm, n+1 points $\mathbf{x}_1, \mathbf{x}_2, ...,\mathbf{x}_{n+1}$ are retained together with their function values.

Denote $f_i = f(\mathbf{x}_i)$. The points are ordered so that $f_1 \leq f_2 \leq ... \leq f_{n+1}$. Then $\mathbf{x}_1$ is the best point and $\mathbf{x}_{n+1}$ is the worst point. The method repeats three basic operations: reflection, expansion and contraction. Every iteration starts with a reflection step.

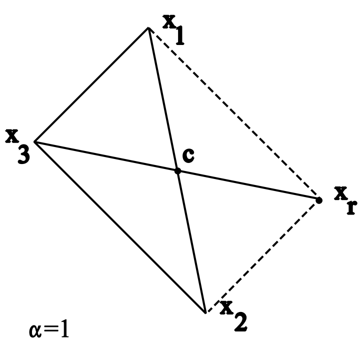
The polytope algorithm

Choose the vertices of the initial polytope around the anticipated minimum point. Calculate function values and order the points in ascending order with f values.

1. Let **c** be the centroid of the n best vertices: $\boldsymbol{c} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i$

REFLECTION: $\mathbf{x}_r = \mathbf{c} + \alpha(\mathbf{c} - \mathbf{x}_{n+1})$

where $\alpha>0$ is the reflection parameter.



Evaluate f_r.

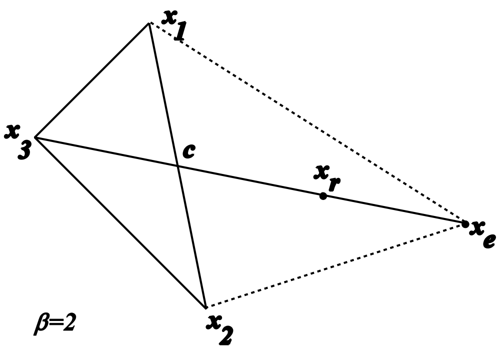
If $f_1 \leq f_r \leq f_n$, $\mathbf{x}_r$ replaces $\mathbf{x}_{n+1}$ and reflection step 1 is repeated.

If $f_r > f_n$, go to step 3.

2. If $f_r < f_1$, move on to the same direction, expanding the polytope:

EXPANSION: $\mathbf{x_e} = \mathbf{c} + \beta(\mathbf{x_r} - \mathbf{c})$

where $\beta>1$ is the expansion parameter.



Evaluate f_e.

If $f_e < f_r$. the expansion was successful and $\mathbf{x_e}$ replaces $\mathbf{x_{n+1}}$.

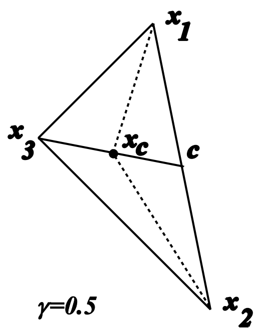
Otherwise $\mathbf{x_{n+1}}$ is relpaced with $\mathbf{x_r}$.

Repeat reflection step 1.

3. If $f_r > f_n$, the polytope is considered to be too large and will be contracted.

CONTRACTION: $\mathbf{x_c} = \mathbf{c} + \gamma(\mathbf{x_{n+1}} - \mathbf{c})$ if $f_r \geq f_{n+1}$
$\mathbf{x_c} = \mathbf{c} + \gamma(\mathbf{x_r} - \mathbf{c})$ if $f_r < f_{n+1}$

where γ is the contraction coefficient, $0\leq \gamma \leq 1$.



Evaluate f_c.

If $f_c < \min \{f_r, f_{n+1}\}$, $\mathbf{x_c}$ replaces $\mathbf{x_{n+1}}$.

If $f_c \geq \min \{f_r, f_{n+1}\}$, another contraction is carried out by halving the edges of the polytope (approaching the best vertex)::

$\mathbf{x_i} = (\mathbf{x_i} + \mathbf{x_1})/2$, $i=2,...,n+1$

Go to step 1 until the size of the simplex is below desired limit or if the standard deviation of the function values at the vertices gets small enough.

Remarks:
- If the polytope gets unbalanced, a restart with a regular polytope should be made.
- Easy to implement
- Small space requirement, essentially a $(n+1)\times(n+2)$-matrix.
- Convergence may be slow.
- Variables should be of the same magnitude → scaling.

When "default" parameter values $\alpha=1$, $\beta=2$, $\gamma=0.5$ are used,
$\mathbf{x_r} = 2\mathbf{c} - \mathbf{x_{n+1}}$
$\mathbf{x_e} = 2\mathbf{x_r} - \mathbf{c}$
$\mathbf{x_c} = (\mathbf{c} + \mathbf{x_{n+1}})/2$ if $f_r \geq f_{n+1}$
$\mathbf{x_c} = (\mathbf{c} + \mathbf{x_r})/2$ if $f_r < f_{n+1}$

3.1.2. OTHER DIRECT METHODS

Random search

The main idea is to make trial moves from the current point in random directions and replace the point whenever a lower function value is encountered. Step size is reduced when enough useless moves are made from one point.

Simulated annealing method is a refined version of random search and will be dealt with later.

Univariate search

Trial moves with a constant step size from the current point are made along each coordinate direction in turn. When a lower function value is encountered, a line search is carried out in that direction and the point is replaced.

Pattern search methods

The pattern search methods start with univariate search but modify the directions on the basis of the topology of the region. Information about the topology, e.g. curvature of the surface is accumulated without derivatives during the iterations.

M.J. Powell proved that using same search directions cyclically (like in the univariate search) is inefficient and convergence cannot be guaranteed.

Pattern search type algorithms:
Hooke and Jeeves method
Powell's method

Powell's method generates so called *conjugate directions* without calculating derivatives.