NUMERICAL METHODS FOR GENERAL NONLINEAR UNCONSTRAINED OPTIMIZATION

$$\min_{x \in \mathbb{R}^n} f(x) \qquad f: \mathbb{R}^n \longrightarrow \mathbb{R} \qquad n \text{ is possibly very large}$$

NELDER-MEAD method

this a "$\emptyset$-order" method meaning that it uses no information from derivatives of f. It belongs to the family of "direct search" methods which are based on the direct comparison of function values at selected points. NM is a simplex-type method: at every step the method starts with a given simplex and end with a different simplex that improves the approximation of the solution.

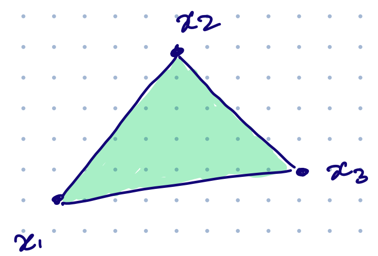
Definition A simplex S in $\mathbb{R}^n$ is the convex hull of $n+1$ points $x_i \in \mathbb{R}^n$, $i = 1, \dots, n+1$

$$S = \left\{ y \in \mathbb{R}^n : \; y = \sum_{i=1}^{n+1} \lambda_i x_i, \quad \lambda_i \geq 0, \; \sum_{i=1}^{n+1} \lambda_i = 1 \right\}$$

Example $\mathbb{R}^2$, $n=2$ $\quad x_1, x_2, x_3 \qquad y = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3$

$\lambda_1, \lambda_2, \lambda_3 \geq 0$

$\lambda_1 + \lambda_2 + \lambda_3 = 1$

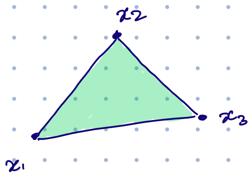


A simplex S in non-singular if the n vectors

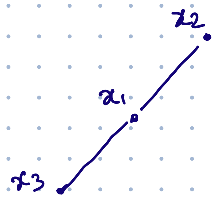
$$\left.\begin{array}{l} x_2 - x_1 \\ x_3 - x_1 \\ \vdots \\ x_{n+1} - x_1 \end{array}\right\} \text{ are linearly independent}$$

Otherwise, S is singular

In $\mathbb{R}^2$,



NON SINGULAR



SINGULAR

Nelder & Mead is based on four operations, each one depending on a parameter

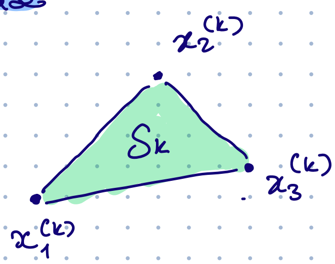
1. REFLECTION $\rho > 0$ (typically $\rho = 1$)
2. EXPANSION $\chi > 1$ (typically $\chi = 2$)
3. CONTRACTION $0 < \gamma < 1$ (typically $\gamma = 1/2$)
4. SHRINKING $0 < \sigma < 1$ (typically $\sigma = 1/2$)

<u>How does an iteration of NM method works.</u>

At step k, how do we reach step $k+1$?
We have at hand a non singular symplex S_k, which is characterized by $n+1$ points $x_i^{(k)}$, $i = 1, \dots, n+1$.
Assume that at every step we order the points $x_i^{(k)}$ in such a way that

$$f(x_1^{(k)}) \leq f(x_2^{(k)}) \leq \dots \leq f(x_{n+1}^{(k)})$$

Reflection phase

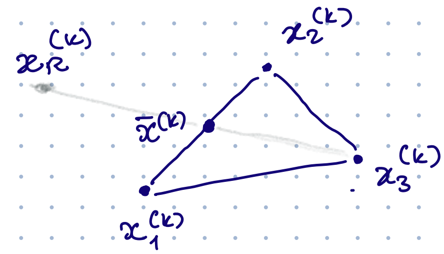


Let $\bar{x}^{(k)}$ denote the barycenter of the <u>n</u> best points:

$$\bar{x}^{(k)} = \frac{1}{n} \sum_{i=1}^{m} x_i^{(k)}.$$

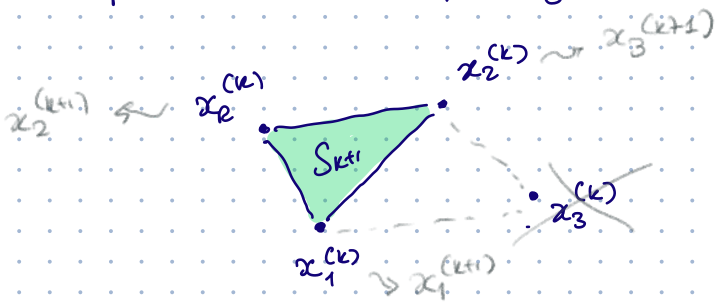
We compute a reflection with parameter ρ ($x_R^{(k)}$) of $x_{n+1}^{(k)}$ with respect to $\bar{x}^{(k)}$

$\rho = 1$



$$x_R^{(k)} = \bar{x}^{(k)} + \rho\left(\bar{x}^{(k)} - x_{n+1}^{(k)}\right)$$

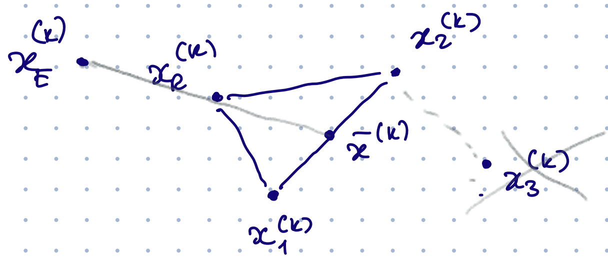
If $f(x_1^{(k)}) \leq f(x_R^{(k)}) < f(x_n^{(k)})$ THEN accept $x_R^{(k)}$ as new point for S_{k+1} (replacing $x_{n+1}^{(k)}$) and STOP



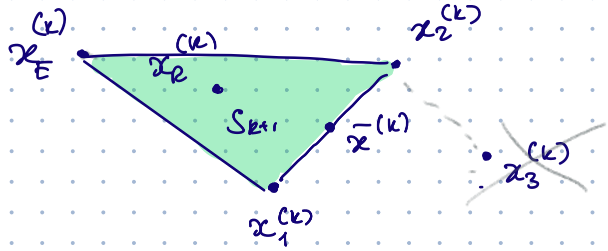
EXPANSION PHASE

If $f(x_R^{(k)}) < f(x_1^{(k)})$ (very good situation!) then we trust a lot the new symplex we want to expand it to explore a wider area. We compute an expansion point $x_E^{(k)}$ as

$$x_E^{(k)} = \bar{x}^{(k)} + \chi\left(x_R^{(k)} - \bar{x}^{(k)}\right)$$



If $f(x_E^{(k)}) < f(x_R^{(k)})$ then ACCEPT $x_E^{(k)}$ as a new point for S_{k+1} and STOP



CONTRACTION PHASE

If $f(x_R^{(k)}) \geq f(x_n^{(k)})$ THEN we perform a contraction between $\bar{x}^{(k)}$ and the best point among $x_R^{(k)}$ and $x_{n+1}^{(k)}$

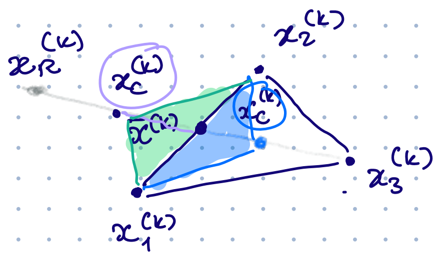
For example, if $x_{n+1}^{(k)}$ is the best among $x_R^{(k)}$ and $x_{n+1}^{(k)}$

- $x_C^{(k)} = \bar{x}^{(k)} - \gamma(\bar{x}^{(k)} - x_{n+1}^{(k)})$

Otherwise, if $x_R^{(k)}$ is the best among $x_R^{(k)}$ and $x_{n+1}^{(k)}$

- $x_C^{(k)} = \bar{x}^{(k)} - \gamma(\bar{x}^{(k)} - x_R^{(k)})$

$\gamma = 1/2$



If $f(x_C^{(k)}) < f(x_{n+1}^{(k)})$ then ACCEPT $x_C^{(k)}$ for S_{k+1} and STOP.

Otherwise:...

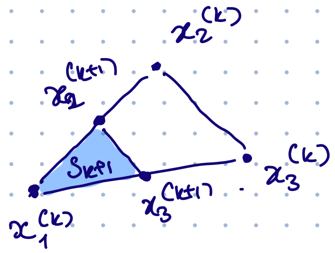
SHRINKAGE PHASE

We shrink the symplex around the best point:

$\forall\, i = 2, \dots, n+1 \qquad \tilde{x}_i^{(k+1)} = x_1^{(k)} + \delta(x_i^{(k)} - x_1^{(k)})$

$\tilde{x}_1^{(k+1)} = x_1^{(k)} \qquad 0 < \delta < 1$

(of course, $\tilde{x}_1^{(k+1)}, \tilde{x}_2^{(k+1)}, \dots, \tilde{x}_{n+1}^{(k+1)}$ will be reordered at the beginning of next step to get $x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{n+1}^{(k+1)}$)



$\sigma = 1/2$

Let us summarize iteration k, $S_k \longrightarrow S_{k+1}$

For simplicity, $f_i^{(k)} := f(x_i^{(k)})$ $f_R^{(k)} := f(x_c^{(k)})$, ... etc.