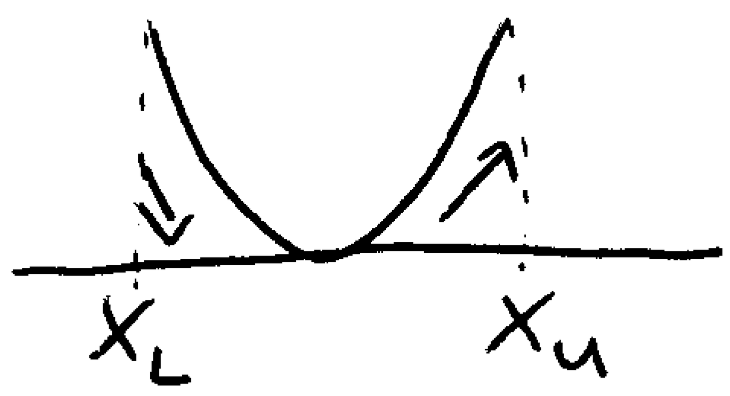
MINIMIZATION AND MAXIMIZATION

Univariate case:

$[X_L, X_u]$ brackets the extremum



unimodal behaviour

section method: (for min)

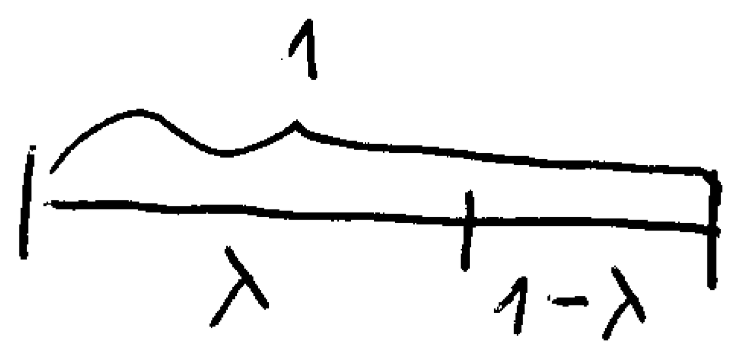
$X_L < X_1 < X_2 < X_L$, calculate $f(X_k)$-s then reduce the interval as:

$f(X_1) \le f(X_2) \Rightarrow$ new interval $[X_L, X_2]$

$>$ $\Rightarrow$ -"- $[X_1, X_u]$

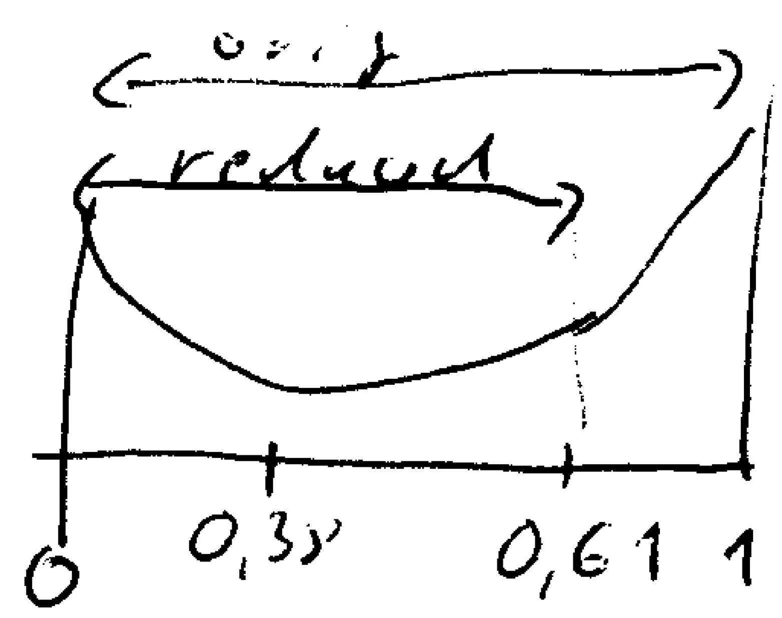
then add a new point within the interval to have 4 data.

golden section



$$1/\lambda = \lambda/(1-\lambda)$$

$$\lambda = 0{,}61803$$



$$X_1 = \lambda \cdot X_L + (1-\lambda) X_u$$

$$X_2 = (1-\lambda) \cdot X_L + \lambda \cdot X_u$$

repeat as in section methods, but point kept is also good in

parabolic interpolation

(3 points → fit parabola → min of parabola is X_4

calculate $x_1, f(x_1)$ $x_2, f(x_2)$ $x_3, f(x_3)$

$$X_4 = \frac{1}{2} \frac{(x_2^2 - x_3^2) f(x_1) + (x_3^2 - x_1^2) \cdot f(x_2) + (x_1^2 - x_2^2) f(x_3)}{(x_2 - x_3) f(x_1) + (x_3 - x_1) \cdot f(x_2) + (x_1 - x_2) \cdot f(x_3)}$$

sort x_1, x_2, x_3, x_4 and throw away x_i most far from x_4

Multivariate case

a) Only $f(\underline{x})$ is used $\quad \underline{x} \in \mathbb{R}^n$

Nelder-Mead simplex method

calculate $f(\underline{x})$ in $n+1$ points $f(\underline{x}_1), f(\underline{x}_2), \dots$

$n=2 \rightarrow$ triangle

$\rightarrow$ polyhedron (simplex)

definitions: $\min f(\underline{x}) \rightarrow \underline{x}_{min}$

$\max f(\underline{x}) \rightarrow \underline{x}_{max}$

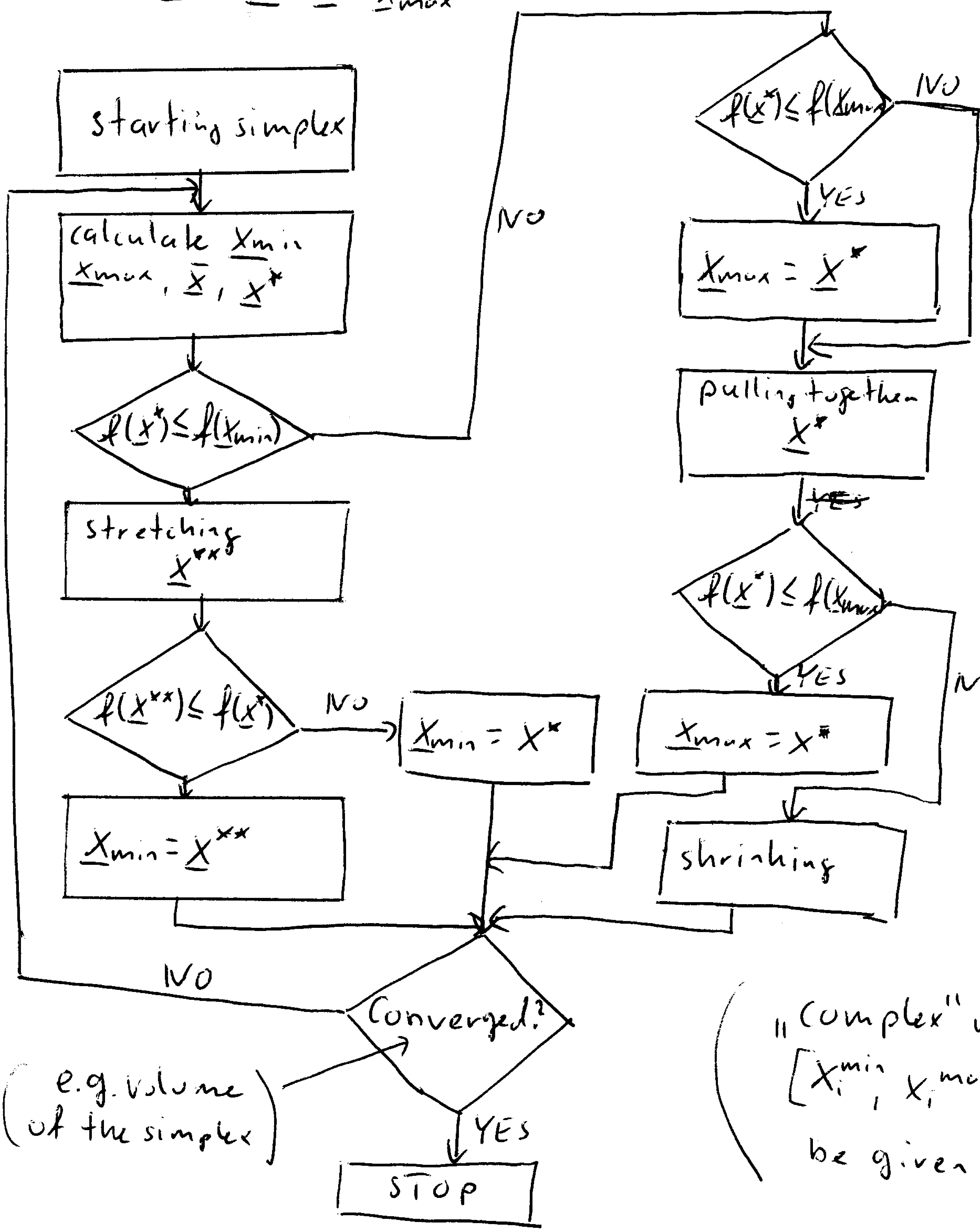
centre: $\bar{\underline{x}} = \left(\sum \underline{x}_i - \underline{x}_{max}\right)/n$

mirroring: $\underline{x}^* = 2\bar{\underline{x}} - \underline{x}_{max}$

stretching: $\underline{x}^{**} = \underline{x}^* + \bar{\underline{x}} - \underline{x}_{max}$

pulling together ~~shrinking~~: $\underline{x}^* = (\underline{x}_{max} + \bar{\underline{x}})/2$

shrinking: $\underline{x}_i = (\underline{x}_i + \underline{x}_{min})/2 \quad i = 1 \dots n+1$



(„complex" version: $[x_i^{min}, x_i^{max}]$ can be given)

b) gradient methods $f(\underline{x})$ and $\frac{\partial f(\underline{x})}{\partial x_i} = \underline{g}$ used

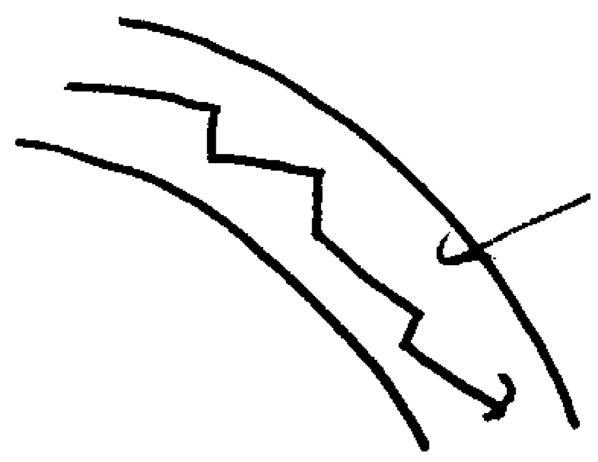
Steepest descent (for minima)

step in the direction of $-\underline{g}$

step size by line search as

- $f(\underline{x}^k - \lambda \cdot \underline{g}(\underline{x}^k))$ minimization with respect to λ ($\lambda \geq 0$)
- $\underline{x}^{k+1} = \underline{x}^k - \lambda_{min}\, \underline{g}(\underline{x}^k)$

but:



the next step is perpendicular to the previous one → bad convergence, small steps

valley of minima

Conjugate gradient methods

step direction is not equal to gradient direction

two vectors: $\underline{h}^i$ and $\underline{g}^i = \nabla f(\underline{x}^i)$

$\underline{h}^0 = \underline{r}^0$ $\quad \underline{x}^{i+1} = \underline{x}^i + \lambda_{min}\, \underline{h}^i$ $\quad$ (λ_{min} by line search)

in new position: $\underline{g}^{i+1}$ and $\underline{h}^{i+1} = \underline{g}^{i+1} + \gamma \cdot \underline{h}^i$

in a way, that

$$\underline{g}_i \cdot \underline{g}_j = 0 \quad \text{and} \quad \underline{g}_i \cdot \underline{h}_j = 0 \quad \text{for all } i \neq j$$

different forms for γ_i

Fletcher-Reeves
$$\gamma_i = \frac{\underline{g}^{i+1} \cdot \underline{g}^{i+1}}{\underline{g}^i \cdot \underline{g}^i}$$

Polak-Ribiere
$$\gamma_i = \frac{(\underline{g}^{i+1} - \underline{g}^i) \cdot \underline{g}^{i+1}}{\underline{g}^i\, \underline{g}^i}$$

Hestenes-Stiefel
$$\gamma_i = \frac{\underline{g}^{i+1}(\underline{g}^{i+1} - \underline{g}^i)}{\underline{h}^i(\underline{g}^{i+1} - \underline{g}^i)}$$

c) Newton-like methods

uses: $f(\underline{x})$, $\underline{g}(x)$, H ← second der. $H_{ij} = \frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j}$

$\underline{g}(x) = \nabla f(x)$ Hesse-matrix

series expansion:

$$f(\underline{x}) = f(\underline{x}^k) + (\underline{x} - \underline{x}^k) \cdot \nabla f(\underline{x}^k) + \frac{1}{2} (\underline{x} - \underline{x}^k)^T \cdot H (\underline{x} - \underline{x}^k)$$

$\nabla f(\underline{x}) = 0$ in extremum $\Rightarrow \Delta f(\underline{x}) = 0 = \nabla f(\underline{x}^k) + H(\underline{x} - \underline{x}^k)$

$$\underline{x}^{k+1} = \underline{x}^k - (H(\underline{x}^k))^{-1} \cdot g(\underline{x}^k)$$

↑ λ_{min} can put here for line search.

at the end positive definite, symmetric

Quasi-Newton methods (like Broyden earlier) ⇒ use B instead of H

BFGS (Broyden-Fletcher-Goldfarb-Shann)

1.) determination of step direction $\underline{p}^k$

$$B^k \cdot \underline{p}^k = -\nabla f(\underline{x}^k) \quad (\text{e.g. using } (B^k)^{-1})$$

2.) line search

$$\underline{x}^{k+1} = \underline{x}^k + \lambda \cdot \underline{p}^k \qquad \lambda \leftarrow \min f(\underline{x}^{k+1})$$

3.) $\Delta \underline{x}^k = \lambda^k \underline{p}^k$

$$\underline{y}^k = f(\underline{x}^{k+1}) - f(\underline{x}^k)$$

4.) $B^{k+1} = B^k + \ldots$ usually directly step 5

5.) $$(B^{k+1})^{-1} = (B^k)^{-1} + \frac{(\Delta \underline{x}^{kT} \cdot \underline{y}^k + \underline{y}^{kT} B^{k-1} \underline{y}^k)(\Delta \underline{x}^k \cdot \Delta \underline{x}^{kT})}{(\Delta \underline{x}^{kT} \cdot \underline{y}^k)^2} - \frac{B^{k-1} \underline{y}^k \cdot \Delta \underline{x}^{kT} + \Delta \underline{x}^k \cdot \underline{y}^{kT} \cdot B^{k-1}}{\Delta \underline{x}^{kT} \cdot \underline{y}^k}$$

B^0 e.g. identity matrix

for large molecules. L-BFGS

↑ (low memory)