Konvergenzbetrachtung zum Verfahren von Spendley et al (S (19))

Bei Abbruch sei oBdA $f(\hat{v}) \geq f(v^n)$, d.h. $\hat{v}$ ist der größte aller Funktionswerte der beiden letzten Simplices.

Setze $v^* := \frac{1}{n} \sum_{i=0}^{n-1} v^i$ und $g^* := Df(v^*)$.

<u>Annahme</u>: f sei stetig diff'bar.

Kantenlänge bei Abbruch des Verfahrens: δ

Es gilt: $f(v^i) = f(v^*) + g^{*T}(v^i - v^*) + o(\delta)$ für $i \leq n-1$

Aufsummiert und durch n dividiert:

$$\underbrace{\frac{1}{n} \sum_{i=0}^{n-1} f(v^i)}_{\leq f(v^n)} = f(v^*) + o(\delta)$$

ferner: $$f(v^n) = f(v^*) + g^{*T}(v^n - v^*) + o(\delta)$$
$$f(v^n) \geq f(v^*) + o(\delta)$$

$\Rightarrow g^{*T}(v^n - v^*) \geq o(\delta)$

Da $f(\hat{v}) \geq f(v^n)$ folgt analog $g^{*T}(\hat{v} - v^*) \geq o(\delta)$

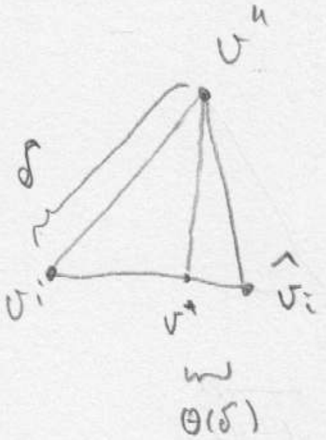
Aus $\hat{v} - v^* = -(v^n - v^*)$ folgt $g^{*T}(\hat{v} - v^*) = o(\delta)$.

Für die übrigen Ecken folgt

$$\begin{aligned} g^{*T}(v^i - v^*) &= f(v^i) - f(v^*) + o(\delta) \\ &\leq f(v^n) - f(v^*) + o(\delta) \\ &= g^*(v^n - v^*) + o(\delta) = o(\delta). \quad (*) \end{aligned}$$

Sei nun $\hat{v}^i$ der Punkt, an dem die Halbgerade $v^i + \lambda(v^* - v^i)$, $\lambda \geq 0$ den Simplex verlässt, $\hat{v}^i = \sum_{j=0}^{n-1} \vartheta_j v^j$ mit $\vartheta_j \geq 0$, $\sum_{j=0}^{n-1} \vartheta_j = 1$.

$\Rightarrow g^{*T}(\hat{v}^i - v^*) = \sum_{j=0}^{n-1} \vartheta_j g^{*T}(v^j - v^*) \underbrace{\leq}_{\text{s.o.}} o(\delta)$.



Da der Simplex regulär ist,
ist $\|\hat v^i - v^*\| = \Theta(\delta)$
(Übung $\approx \frac{\delta}{n-1}$)

Da $\hat v^i - v^*$ entgegengesetzt zu $v^i - v^*$ gerichtet ist
folgt $g^*(v^i - v^*) \overset{\geq}{=} o(\delta)$. Insgesamt $g^*(v^i - v^*) = o(\delta)$
Damit sind n linear unabhängige Richtungen r^i der Länge
$\Theta(\delta)$ gegeben, für die $g^{*T} r = o(\delta)$ gilt.
$\Rightarrow \|g^*\| = o(1)$.
(Beim Verfahren von Torczon & Dennis vereinfacht sich
obige Argumentation etwas.)

Das Verfahren von Nelder & Mead (1965)

(Lange Zeit das weltweit populärste Verfahren)

Wie bisher, $\hat v := \left(\frac{2}{n} \sum_{i=0}^{n-1} v^i\right) - v^n$ ⊛

1) Falls $f(v^0) \le f(\hat v) < f(v^{n-1})$ so ersetze v^n durch $\hat v$
Sortiere neu und gehe zu ⊛ "wie gehabt"

2) Falls $f(\hat v) < f(v^0)$ setze $\hat{\hat v} := \hat v + \frac{1}{2}(\hat v - v^n)$
falls $f(\hat{\hat v}) < f(\hat v)$ ersetze v^n durch $\hat{\hat v}$, sonst durch $\hat v$
Sortiere neu und gehe zu ⊛ "Extrapolation"

3) Falls $f(\hat v) \ge f(v^{n-1})$ (dann würde man im nächsten Schritt zurückspiegeln)

3a)) falls $f(\hat{v}) < f(v^n)$ setze $\hat{\hat{v}} := \frac{3}{4}\hat{v} + \frac{1}{4}v^n$

3b)) falls $f(\hat{v}) \geq f(v^n)$ setze $\hat{\hat{v}} := \frac{3}{4}v^n + \frac{1}{4}\hat{v}$ "kürzerer Schritt"

- falls nun $f(\hat{\hat{v}}) < f(v^n)$ ersetze v^n durch $\hat{\hat{v}}$, sortiere neu und gehe zu ⊗

  (im direkt nachfolgenden Schritt wird sicher nicht identisch zurückgespiegelt da die Schrittlänge verkürzt worden war.)

- sonst schrumpfe den Simplex zu v^0 hin.

<u>Bsp:</u> Mc Kinnon 1997 (SIOPT)

1)) Konstruktion einer Folge von Simplices, die gegen $\binom{0}{0}$ konvergiert

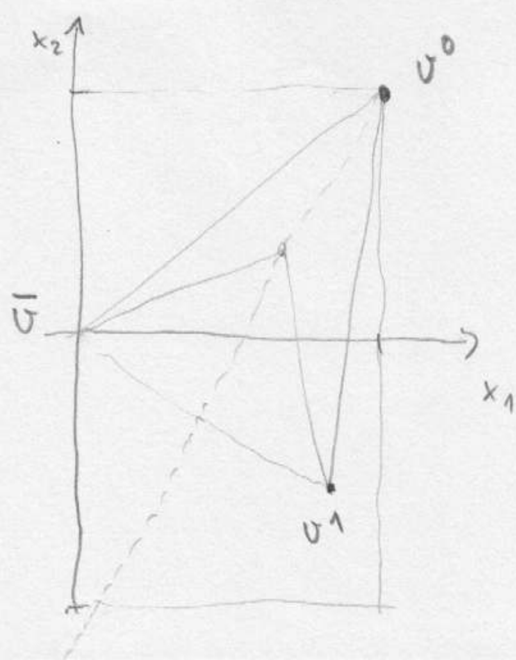
Nachfolgend ist die Nummerierung dahingehend geändert, dass immer eine neue Ecke v^{k+1} dazukommt und zum Start

$\bar{v} = \binom{0}{0}$, $v^0 = \binom{1}{1}$, $v^1 = \binom{\lambda_1}{\lambda_2}$ mit $\lambda_1 = \frac{1+\sqrt{33}}{8} \approx 0.843$

$\lambda_2 = \frac{1-\sqrt{33}}{8} \approx -0.593$

gegeben sind und $f(\bar{v}) < f(v^1) < f(v^0)$ gelte.

Es soll wiederholt Schritt 3b durchgeführt werden.



Die Dreiecke sollen dabei einer Rekursion folgen:

Erster Schritt:

$$\hat{v} = \frac{2}{2}(\bar{v} + v^1) - v^0 = v^1 - v^0$$

$$\hat{\hat{v}} = \frac{3}{4} v^0 + \frac{1}{4}\hat{v} = \frac{1}{2} v^0 + \frac{1}{4} v^1$$

<u>Beh:</u> $v^n = \begin{pmatrix} \lambda_1^n \\ \lambda_2^n \end{pmatrix}$ mit obigen λ_1, λ_2. $\quad e_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; e_2 := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$n = 0$ ✓

$n = 1$ ✓

$(n-1, n) \longrightarrow n+1$: $\quad \underbrace{\hat{\hat{v}} = \frac{1}{2} v^n + \frac{1}{4} v^{n-1}}_{\text{wie oben}}.$

$$= \frac{1}{4}(\lambda_1^n e_1 + \lambda_2^n e_2) + \frac{1}{2}(\lambda_1^{n-1} e_1 + \lambda_2^{n-1} e_2)$$

$$= \lambda_1^{n-1} \underbrace{\underbrace{\underbrace{\left(\frac{1}{4}\lambda_1 + \frac{1}{2}\right)}_{\frac{1}{32}(1+\sqrt{33}+16)}}_{\frac{1}{64}(1+\sqrt{33})^2}}_{\lambda_1^{n+1}} e_1 + \lambda_2^{n-1} \underbrace{\underbrace{\underbrace{\left(\frac{1}{4}\lambda_2 + \frac{1}{2}\right)}_{\frac{1}{32}(1-\sqrt{33}+16)}}_{\frac{1}{64}(1-\sqrt{33})^2}}_{\lambda_2^{n+1}} e_2$$

$$\leftarrow (1 \pm \sqrt{33})^2 = 1 \pm 2\sqrt{33} + 33 = 2(\pm\sqrt{33} + 17)$$

Also $\hat{\hat{v}} = \lambda_1^{n+1} e_1 + \lambda_2^{n+1} e_2$.

2) Eine streng konvexe Funktion, die ein Verhalten wie oben erzeugt: