1. Formulering af opgaven

At bestemme minimum for en funktion (af flere variable) er et kendt problem, og der findes talrige numeriske metoder. Jeg vil her præsentere en metode, som kræver kendskab til de partielle afledede af funktionen. Hvis der ikke findes et analytisk udtryk, kan de partielle afledede eventuelt beregnes numerisk, men i alle tilfælde kræver beregningen numeriske metoder.

Lad der være givet en reel funktion: $y = f(x_1, x_2, x_3 \dots x_n)$

Ekstrema for funktionen er givet ved betingelserne: $\frac{\partial f}{\partial x_1} = 0$, $\frac{\partial f}{\partial x_2} = 0$, $\frac{\partial f}{\partial x_3} = 0$, ... $\frac{\partial f}{\partial x_n} = 0$

I almindelighed er det ikke muligt at løse disse *n* ligninger med *n* ubekendte analytisk.

Idet vi skriver: $\overline{x} = (x_1, x_2, x_3 \dots x_n)$ og tilsvarende for andre variable.
Vi ud fra et punkt $\overline{x_0}$ og bevæger os i en retning $\overline{a}$, hvor $\overline{a} = (a_1, a_2, a_3, \dots, a_n)$ er en enhedsvektor.
Vi ser da på den reelle funktion af én variabel *t*:

$$y = f(t) = f(\overline{x_0} + t\overline{a}) = f(x_1 + ta_1, x_2 + ta_2, x_3 + ta_2 \dots, x_n + ta_n)$$

2. Ekstremum med bibetingelse. Method of steepest descend

Opgaven er da at bestemme $\overline{a}$, således at vi får den "stejleste nedgang" (steepest descent), og det gør vi ved at bestemme ekstremum for *f´(t)* ud fra $\overline{x_0}$, altså for $t = 0$

$$f'(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} a_i = F(\overline{a})$$

$F(\overline{a})$ kan betragtes som en funktion af de n variable $\overline{a} = (a_1, a_2, a_3, \dots, a_n)$ Opgaven er da, at bestemme ekstremum for denne funktion. Det er tilsyneladende den samme opgave, som vi begyndte med, men forskellen er den at $F(\overline{a})$ er en lineær funktion af $\overline{a} = (a_1, a_2, a_3, \dots, a_n)$.
Da $\overline{a}$ er en enhedsvektor, skal ekstremum bestemmes under *bibetingelsen* $|\overline{a}| = 1$ eller $\sum_{i=1}^{n} a_i^2 = 1$. Vi sætter: $G(\overline{a}) = \sum_{i=1}^{n} a_i^2$ Ifølge teorien for ekstremum med bibetingelse, kan dette gøres ved at søge minimum for funktionen: $F(\overline{a}) + \lambda G(\overline{a})$, hvor λ er en såkaldt Lagrange multiplikator, der bestemmes ved opgavens randbetingelser. Det gøres ved at udregne de partielle afledede og sætte dem lig med nul.

$$F(\overline{a}) + \lambda G(\overline{a}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} a_i + \lambda \sum_{i=1}^{n} a_i^2$$

$$\frac{\partial F}{\partial a_i} + \lambda \frac{\partial G}{\partial a_i} = 0 \quad \Leftrightarrow \quad \frac{\partial f}{\partial x_i} + 2\lambda a_i = 0$$

Hvoraf følger: $a_i = -\frac{1}{2\lambda}\frac{\partial f}{\partial x_i}$

λ bestemmes da ved normalitetsbetingelsen: $\sum_{i=1}^{n} a_i^{\,2} = 1$, som giver

$$4\lambda^2 = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)^2 \quad \Leftrightarrow \quad \lambda = \tfrac{1}{2}\sqrt{\sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)^2}$$

3. Numeriske metoder

Selvom vi har fundet den stejleste vej ned mod et minimum, så ved vi ikke hvor langt vis skal gå af denne vej. Man kunne gå et passende skridt og gentage beregningen, men det er ikke rigtig brugbart.
I stedet kunne man antage at minimum er parabolsk, og bestemme minimum som toppunktet af det 2.grads polynomium, der følger af udregningen af 1. og 2. afledede i punktet $t = 0$.
Et 2. gradspolynomium, der går gennem et punkt $(t_0, f(t_0))$, og som har de samme første og anden afledede som funktionen i punktet $t = 0$, kan skrives:

$$y = \tfrac{1}{2} f''(0)t^2 + f'(0)t + f(0)$$

Den afledede af en funktion i 0, kan udregnes numerisk til anden orden i h som:

$$f'(0) = \frac{f(\frac{h}{2}) - f(-\frac{h}{2})}{h}$$

Heraf følger det:

$$f'(-\tfrac{h}{2}) = \frac{f(0) - f(-h)}{h} \qquad f'(\tfrac{h}{2}) = \frac{f(h) - f(0)}{h} \qquad f''(0) = \frac{f'(\frac{h}{2}) - f'(-\frac{h}{2})}{h}$$

Bemærk, at til beregning af de tre afledede kun behøver at kende; $f(-h)$, $f(0)$ og $f(h)$.

$f(0) = f(\bar{x}_0) \qquad f'(0) = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}a_i$, hvor de partielle afledede beregnes efter formlen:

$$\left.\frac{\partial f}{\partial x_i}\right|_{t=0} = \frac{f(x_1, x_2, \ldots x_i + \frac{h}{2}, \ldots x_n) - f(x_1, x_2, \ldots x_i - \frac{h}{2}, \ldots x_n)}{h}$$

Når de første og 2. afledede er beregnet, kan man indsætte i formlen:

$$y = \tfrac{1}{2} f''(0)t^2 + f'(0)t + f(0)$$

Hvis vi differentierer med hensyn til t, får man $y' = f''(0)t + f'(0)$. Toppunktet findes, hvor $y' = 0$

Herefter følger det at toppunktet svarer til $f''(0)t + f'(0) = 0 \quad \Leftrightarrow \quad t = -\dfrac{f'(0)}{f''(0)}$

Hvis *f''(0)* ikke er negativ, er der enten ikke noget minimum i nærheden eller funktionen har muligvis et maximum.

Ovenstående er forarbejdet til et program, som jeg lavede, da jeg arbejdede i CERN i 1970. Det er uhyre effektivt for kvadratiske former i n variable findes minimum i et (højest to) skridt.

Programmet matematisk analyse, som findes på min hjemmeside, anvender denne metode.

Ole Witt-Hansen Oktober 2013 (CERN 1970)