Ekskurs i variationsregning

S. I. Olsen

14. januar 2003

Variationsregning er et værktøj til bestemmelse af den/de funktioner, der ekstremiserer en funktional. Dette er (i en vis forstand) en pendent til det værktøj vi benytter når vi ønsker at finde stationære punkter for en funktion. I sidstnævnte tilfælde differentierer vi funktionen mht. hver variabel og sætter hvert udtryk lig nul. Denne metode kan vi ikke bruge når vi ønsker at optimere over et rum af funktioner. I det følgende introduceres meget overfladisk til variationsregning. Alle beviser udelades.

Denne note er klippet fra mine *forelæsningsnoter til kurset: Introduktion til digital billedbehandling*. For en grundigere gennemgang henvises til et utal af lærebøger i matematik. En udmærket gennemgang findes på dansk i: *Erik Hansen, Variationsregning, Polyteknisk Forlag, Lyngby 1976.*

I variationsregning betragter vi en funktional $\Phi(f) : (\Omega \rightarrow \Re) \rightarrow \Re$, hvor f altså antages at afbilde Ω på de reelle tal, og hvor Φ er et *godhedsmål* på sådanne funktioner, dvs. et reelt tal. Ofte vælges Φ således at $\Phi(f_1) < \Phi(f_2)$ hvis f_1 er bedre end f_2. Lad os i det følgende antage at f er defineret på en delmængde Ω af $\Re$. I dette tilfælde antages at godhedsmålet har formen:

$$\Phi(f) \;=\; \int\int_{\Omega} F(x, f, f_x, f_{xx}, \ldots) dx \tag{1}$$

Det antages altså at godhedsmålet er bestemt ved funktionen F af stedkoordinaten x, af funktionen f selv, samt af de afledede af f. I de fleste tilfælde (fra billedanalyse) indgår hverken x eller afledede af orden større end 2. Til gengæld er det ofte at funktionen f er defineret på en delmængde af $\Re^2$. I dette tilfælde fås:

$$\Phi(f) \;=\; \int\int_{\Omega \subseteq \Re^2} F(f, f_x, f_y, f_{xx}, f_{xy}, f_{yy}) dxdy \tag{2}$$

I dette tilfælde har F således seks parametre. I det følgende betragtes mængden af *tilladelige* funktioner, der er $2N$ gange differentiable, hvor N er den højeste orden afledede af f som optræder i F. Desuden skal der gælde, at alle afledede af f af orden op til N enten har værdi lig nul på randen af definitionsområdet Ω eller opfylder et sæt på forhånd givne *randbetingelser*. Det kan vises at en nødvendig betingelse for at en tilladelig funktion f

ekstremiserer Φ er at funktionen tilfredsstiller *Euler-Lagrange ligningen* for funktionalen. For en funktion f af én variabel er Euler-Lagrange ligningen givet ved:

$$F_f + \sum_{n=1}^{N} (-1)^n \frac{\partial^n}{\partial x^n} F_{f^{(n)}} = 0 \tag{3}$$

Euler-Lagrange ligningen for funktioner af flere variable er analoge. F.eks. er Euler-Lagrange ligningen svarende til ovenstående eksempel lig:

$$F_f - \frac{\partial}{\partial x} F_{f_x} - \frac{\partial}{\partial y} F_{f_y} + \frac{\partial^2}{\partial x^2} F_{f_{xx}} + \frac{\partial^2}{\partial x \partial y} F_{f_{xy}} + \frac{\partial^2}{\partial y^2} F_{f_{yy}} = 0 \tag{4}$$

I ovenstående ligning skal faktoren $F_{f_{xx}}$ eksempelvis læses "den afledede af funktionen F med hensyn til den fjerde parameter f_{xx}". Størrelserne f_x, f_{xx}, etc. betragtes altså her som symboler, og ikke som funktionsudtryk.

Eksempel

Nedenstående anføres en række integrander F til godhedsmåled $\Phi(f)$, samt de tilhørende Euler-Lagrange ligninger.

$$\begin{array}{lll} F = f_x^2 + f_y^2 & : & f_{xx} + f_{yy} = 0 \\ F = f - f_x^2 + f_{xx}^2 & : & 1 + 2f_{xx} + 2f_{xxxx} = 0 \\ F = \sqrt{1 + f_x^2} & : & f_{xx} = 0 \end{array}$$

hvor der i sidste eksempel er udnyttet at kvadratrodsfunktionen er monotont voksende. Ekstremum for $\sqrt{g(x)}$ og for $g(x)$ antages derfor i de samme punkter.

Det er vigtigt ved anvendelse af variationsregning, at Euler-Lagrange ligningen kun angiver et nødvendigt krav til en funktion f om at ekstremiserer $\Phi(f)$. Selv om Euler-Lagrange ligningen er opfyldt er vi ikke garanteret at f er et globalt ekstremum. Det ikke let at opstille kriterier for en tilstrækkelig betingelse for ekstremum. Dog gælder, at hvis funktionen F er konveks, da angiver Euler-Lagrange ligningen en tilstrækkelig betingelse.

Eksempel

Betragt variationsregningsproblemet:

$$\Phi(f) = \int_a^b \sqrt{1 + f_x^2(x)} dx \ , \quad f(a) = f_a \ , \quad f(b) = f_b$$

Her søger vi altså den korteste glatte kurve gennem punkterne (a, f_a) og (b, f_b). Klassen af tilladelige funktioner er mængden af 2 gange differentiable relle funktioner på delmængden [a:b]. Integranden F er i dette tilfælde konveks. Euler-Lagrange ligningen er givet ved den anden ordens differentialligning $f_{xx} = 0$. Ved integration får vi derfor løsningen:

$f(x) = Ax + B$, hvor A og B er reelle integrationskonstanter. Vi har altså vist at den korteste vej mellem to punkter er en ret linie. Konstanterne A og B skal vælges således at *randbetingelserne* $f(a) = f_a$ og $f(b) = f_b$. Herved bliver løsningen specificeret fuldstændigt.

Hvis randbetingelserne i et variationsregningsproblem udelades kan det vises at et krav for ekstremitet (ud over Euler-Lagrange ligningen) er opfyldelsen af de såkaldte *naturlige randbetingelser*. I tilfældet $F = F(x, f, f')$ vil de naturlige randbetingelser diktere at $F_{f'}$ er lig nul på randen (i intervalgrænserne) for definitionsområdet. Hvis vi, som eksempel, i ovenstående eksempel undlod randbetingelserne $f(a) = f_a$ og $f(b) = f_b$ ville de naturlige randbetingelser angive at $f_x(a) = f_x(b) = 0$, svarende til kravet $A = 0$. Løsningen ville derfor ikke blive én funktion, men klassen af funktioner med konstant værdi.

I mange variationsregningsopgaver er der, ud over godhedsmålet F_0, givet en antal bibetingelser :

$$
\begin{aligned}
\Phi_1(f) &= \int_\Omega F_1(\ldots)dxdy = c_1 \\
\Phi_2(f) &= \int_\Omega F_2(\ldots)dxdy = c_2 \\
\vdots \quad & \quad \vdots \qquad\qquad\qquad\quad \vdots \\
\Phi_n(f) &= \int_\Omega F_n(\ldots)dxdy = c_n
\end{aligned}
$$

hvor alle integraler deler samme grænser. Det kan vises at løsningen til ovenstående variationsregningsproblem er den samme som løsningen til problemet:

$$
\Phi(f) = \int_\Omega \left[F_0 + \lambda_1 F_1 + \lambda_2 F_2 + \ldots + \lambda_n F_n\right] dxdy \tag{5}
$$

hvor F_0 er det oprindelige godhedsmål, og hvor sættet $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots \lambda_n) \neq (0, 0, \ldots, 0)$ vælges således at bibetingelserne er opfyldt. $\vec{\lambda}$ kaldes for et sæt af Lagranske multiplikatorer. Løsningen til Euler-Lagrange ligningen vil derfor være parametriseret ved $\vec{\lambda}$. Denne vektor bestemmes (efter at den generelle løsning er fundet) således at bibetingelserne bliver opfyldt.

Eksempel

Lad klassen af tilladelige funktioner være lig mængden af to gange differentiable funktioner på intervallet [-w:w]; lad godhedsmålet $\Phi_0(f)$, og bibetingelsen $\Phi_1(f)$ være defineret ved nedenstående ligninger:

$$
\begin{aligned}
\Phi_0(f) &= \int_{-w}^{w} f_x^2(x)dx \\
\Phi_1(f) &= \int_{-w}^{w} f^2(x)dx = 1
\end{aligned}
$$

Lad der endvidere gælde følgende randbetingelser: $f(0) = 0$, $f(w) = A$ og $f(w) = -A$. Først opstilles den sammensatte funktional:

$$\Phi(f) \;=\; \int_{\Re} [f_x^2(x) + \lambda f^2(x)]dx$$

Denne har Euler-Lagrange ligningen: $\lambda f - f_{xx} = 0$, der har løsningen: $f(x) = c_1 e^{mx} + c_2 e^{-mx}$, hvor $m = \sqrt{\lambda}$. Randbetingelsen $f(0) = 0$ giver nu at $c_1 + c_2 = 0$. Herved kan løsningen skrives: $f(x) = 2c_1 \sinh(mx)$. Konstanten c_1 kan nu bestemmes således at bibetingelsen bliver opfyldt. Endelig kan konstanten m bestemmes således at de to sidste randbetingelser bliver opfyldt.

Eksempel

Lad os antage at vi har målt en todimensional funktion i et antal punkter M, og at vi ved at funktionen er glat i den forstand at funktionen kan modelleres ved en membran. Hvis dette er tilfældet kan godheden af funktionen f måles ved den energi, som der skal til at bøje funktionen (membranen) således at den passerer tæt ved de givne målepunkter. Det kan vises, at integralet af kvadratet på gradientstørrelsen er et mål for den omtalte energi. Vi kan derfor opskrive kravet til f som et godhedsmål Φ_0 samt en bibetingelse Φ_1:

$$\begin{aligned}\Phi_0(f) &= \int_{\Re^2} [f_x^2(x,y) \;+\; f_y^2(x,y)]dxdy \\ \Phi_1(f) &= \int_{\Re^2} \delta_M(x,y)[f(x,y) \;-\; c(x,y)]^2 dxdy \;=\; 0\end{aligned}$$

hvor

$$\delta_M(x,y) \;=\; \begin{cases} 1 & \text{hvis } (x,y) \in M \\ 0 & \text{ellers} \end{cases}$$

og hvor

$$c(x,y) \;=\; \begin{cases} \text{den målte værdi} & \text{hvis } (x,y) \in M \\ \text{udefineret} & \text{ellers} \end{cases}$$

Den sammensatte funktional bliver derfor:

$$\Phi(f) \;=\; \int_{\Re^2} [f_x^2(x,y) \;+\; f_y^2(x,y) \;+\; \lambda\delta_M(x,y)[f(x,y) - c(x,y)]^2]dxdy$$

Euler-Lagrange ligningen bliver:

$$\lambda\delta_M(x,y)[f(x,y) \;-\; c(x,y)] \;+\; f_{xx}(x,y) \;+\; f_{yy}(x,y) = 0$$

Vi har således transformeret det globale optimeringsproblem til et problem om at løse en differentialligning, der skal holde for alle (x,y). Hvis vi kræver bibetingelsen overholdt eksakt (svarende til at vælge λ lig ∞) er løsningen (ud over overensstemmelserne mellem f og c på M) givet ved et sæt af koblede differentialligninger $f_{xx}(x,y) + f_{yy}(x,y) = 0$. Det er ikke muligt direkte at opskrive en analytisk løsning.

Antag i det følgende at vi ikke kræver at bibetingelsen overholdes eksakt, men kun at funktionen f passerer tæt forbi de målte værdier c. Denne situation svarer til at vælge $\lambda < \infty$. Antag endvidere at såvel f som c er diskrete funktioner. En ofte benyttet diskret approksimation til laplaceoperatoren er:

$$\nabla^2 f(x,y) \;=\; f_{xx}(x,y) \;+\; f_{yy}(x,y) \;\;=\;\; 4\bar{f}(x,y) - 4f(x,y)$$

hvor

$$\bar{f}(x,y) \;=\; \frac{1}{4}[f(x-1,y) \;+\; f(x+1,y) \;+\; f(x,y-1) \;+\; f(x,y+1)]$$

Indsættes den diskrete approksimation, og løses med hensyn til funktionsværdien $f(x,y)$ fås følgende relation:

$$f(x,y) \;=\; \frac{4\bar{f}(x,y) \;+\; \lambda\delta_M(x,y)c(x,y)}{4 \;+\; \lambda\delta_M(x,y)}$$

der skal være opfyldt i alle (x,y). I punkter hvori der ikke er nogen måling ses at værdien af f er bestemt ved gennemsnittet af de omkringliggende værdier. I punkter $(x,y) \in M$ er f bestemt ved et vægtet gennemsnit af denne værdi og af den målte værdi. Jo større λ vælges jo mere vægt lægges der på overensstemmelse med de målte værdier i forhold til glathedskriteriet. Ovenstående udledning kan sammenlignes med det tilsvarende udtryk for regularisering af et en-dimensionalt signal (se afsnit 1.7.3).