MAT 1110, 25. april 2022

∗ Ekstremalverdisetingen ∗ Lokale maks- og min-

Arne B. Sletsjøe
Universitetet i Oslo

Ekstremalverdisetningen

Definisjon

La $A \subset \mathbb{R}^m$ *og* $f : A \to \mathbb{R}$. *Vi sier at* f *er* **begrenset** *dersom det finnes reelle tall* K, M *slik at*

$$K \leq f(\mathbf{x}) \leq M$$

for alle $\mathbf{x} \in A$.

Vi sier at $\mathbf{c} \in A$ *er et* **globalt maksimumspunkt** *for* f *dersom*

$$f(\mathbf{c}) \geq f(\mathbf{x})$$

for alle $\mathbf{x} \in A$.

Vi sier at $\mathbf{d} \in A$ *er et* **globalt minimumspunkt** *for* f *dersom*

$$f(\mathbf{d}) \leq f(\mathbf{x})$$

for alle $\mathbf{x} \in A$.

Setning

La A være en lukket, begrenset delmengde av $\mathbb{R}^m$, og $f : A \to \mathbb{R}$ en kontinuerlig funksjon. Da har f maks- og min-punkter og er følgelig begrenset på A.

Bevis.

$$M = \sup\{f(\mathbf{x}) \,|\, \mathbf{x} \in A\} \in \mathbb{R} \cup \{\infty\}$$

La $\{x_n\}$ være en følge slik at $x_n \to M$ når $n \to \infty$. Siden A er lukket og begrenset har følgen en konvergent delfølge $\{\mathbf{x}_{n_k}\}$ (Bozano-Weierstrass) som konvergerer mot et punkt $\mathbf{c} \in A$ (siden A er lukket). Det gir

$$\lim_{n\to\infty} f(\mathbf{x}_{n_k}) = f(\mathbf{c})$$

Men vi har også

$$\lim_{n\to\infty} f(\mathbf{x}_{n_k}) = M$$

og det følger at $f(\mathbf{c}) = M$. Mao funksjonen er begrenset med maksimumspunkt $f(\mathbf{c})$. $\square$

Lokale maksimums- og minimumspunkter

Definisjon

La $A \subset \mathbb{R}^m$ *og* $f : A \to \mathbb{R}$*.*

(i) *Vi sier at* $\mathbf{a} \in A$ *er et* **lokalt maksimumspunkt** *for* f *dersom det finnes en ball* $B(\mathbf{a}, r)$ *med sentrum i* $\mathbf{a}$ *slik at*

$$f(\mathbf{a}) \geq f(\mathbf{x})$$

for alle $\mathbf{x} \in B(\mathbf{a}, r) \cap A$*.*

(ii) *Vi sier at* $\mathbf{a} \in A$ *er et* **lokalt minimumspunkt** *for* f *dersom det finnes en ball* $B(\mathbf{a}, r)$ *med sentrum i* $\mathbf{a}$ *slik at*

$$f(\mathbf{a}) \leq f(\mathbf{x})$$

for alle $\mathbf{x} \in B(\mathbf{a}, r) \cap A$*.*

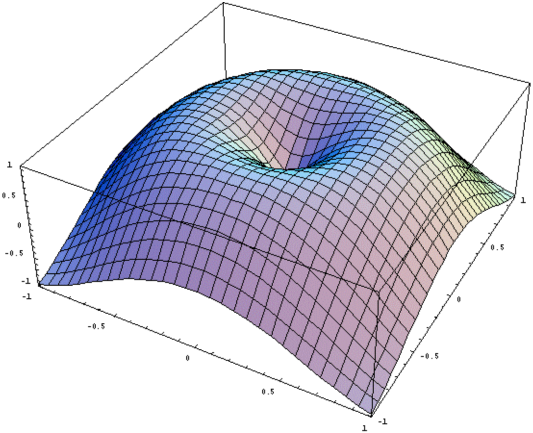
Fellesbetegnelsen for lokale maks og min er **lokale ekstremalpunkter***.*

Eksempel

Funksjonen $f(x,y) = x^2 + y^2$ *har et lokalt minimum i* $(0,0)$ *siden* $f(0,0) = 0$ *og* $x^2 + y^2 > 0$ *når* x *og* y *ikke begge er 0. Det betyr at uansett hvilken origo-sentrert sirkelskive vi ser på, så vil origo være minimum for funksjonen på denne sirkelskiven.*
Funksjonen $f(x,y) = x^2 + y^2$ *har ikke noen lokale maksimumpunkter siden funksjonsverdien vokser over alle grenser når vi beveger oss vekk fra origo.*

Eksempel

Funksjonen $z = \sin(\pi \cdot \sqrt{x^2 + y^2})$ *har et lokalt minimum i* $(0,0)$ *og lokale maksima i* $x^2 + y^2 = \frac{1}{2}$ *(pluss mange flere som vi ikke ser på figuren).* *Dette følger siden* $g(0,0) = 0$ *og* $g(x,y) = \sin(\pi \cdot \sqrt{x^2 + y^2}) > 0$ *i nærheten av origo. På sirkelen* $x^2 + y^2 = \frac{1}{4}$ *har vi at* $z = \sin(\pi \cdot \sqrt{\frac{1}{4}}) = \sin \frac{\pi}{2} = 1$*, som opplagt er maksimumsverdi for en sinus-funksjon.*



Definisjon

Et punkt $\mathbf{a} \in A$ *kalles et* **stasjonært punkt** *for* $f : A \to \mathbb{R}$ *gradienten* $\nabla f(\mathbf{a}) = 0$.

Setning

Et ekstremalpunkt i det indre av definisjonsområdet til en funksjon med kontinuerlige partiellderiverte vil være stasjonært.

Merk. Det motsatte er ikke tilfelle, stasjonære punkter trenger ikke å være ekstramalpunkter.

Definisjon

Et stasjonært punkt $\mathbf{a} \in A$ *som ikke er et ekstremalpunkt kalles et* **sadelpunkt**.

Bevis.

Vi gjør beviset for en funksjon i 2 variable.

Anta at $f(x,y)$ har et lokalt maksimum i (a,b). Det betyr at $f(x,y) - f(a,b) \leq 0$ for alle (x,y) i nærheten av (a,b). Anta videre

$$\frac{\partial f}{\partial x}(a,b) \approx \frac{f(a+h,b) - f(a,b)}{h} > 0$$

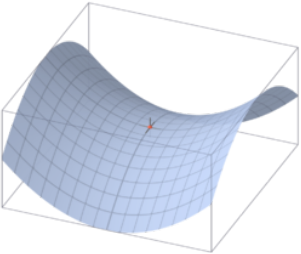
for alle små $h > 0$. Dette gir en motsigelse siden vi har antatt at $f(a+h,b) - f(a,b) \leq 0$. Anta så at

$$\frac{\partial f}{\partial x}(a,b) \approx \frac{f(a+h,b) - f(a,b)}{h} < 0$$

denne gangen for alle små $h < 0$. Det betyr $f(a+h,b) - f(a,b) > 0$ og igjen får vi en motsigelse. $\square$

Eksempel

Vi ser på grafen til funksjonen $f(x,y) = x^2 - y^2$.



De partielt deriverte er gitt ved

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = -2y$$

Punktet $(0,0)$ *er et stasjonært punkt. Det er et minimumspunkt dersom vi beveger oss gjennom det langs med* x*-aksen, og et maksimumspunkt langs med* y*-aksen. Dette er det lett å se ved å studere funksjonene*

$$f(x,0) = x^2, \quad f(0,y) = -y^2$$

Annenderiverttesten

Definisjon

La f være en 2 ganger deriverbar funksjon i m variable, og $\mathbf{a}$ et punkt i D_f. Vi kaller $m \times m$-matrisen

$$Hf(\mathbf{a}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_m} \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1} & \frac{\partial^2 f}{\partial x_m \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_m^2} \end{pmatrix}$$

for **Hesse-matrisen** *til f.*

Vi har

$$Hf(\mathbf{a}) = (\nabla f)'(\mathbf{a})$$

Setning

La $\mathbf{a}$ *være et stasjonært punkt for en funksjon* f *i* m *variable med kontinuerlige annenordens partiellderiverte i en omegn om* $\mathbf{a}$*. Da har vi*

(i) *Hvis alle egenverdiene til* $Hf(\mathbf{a})$ *er positive, så er punktet* $\mathbf{a}$ *et lokalt minimumspunkt.*

(ii) *Hvis alle egenverdiene til* $Hf(\mathbf{a})$ *er negative, så er punktet* $\mathbf{a}$ *et lokalt maksimumspunkt.*

(iii) *Hvis noen av egenverdiene til* $Hf(\mathbf{a})$ *er positive og noen er negative, så er punktet* $\mathbf{a}$ *et sadelpunkt.*

Hvis noen av egenverdiene til $Hf(\mathbf{a})$ *er 0, så gir ikke testen noe resultat.*

Setning

La $f(x, y)$ være en to ganger deriverbar funksjon i to variable og anta at (a, b) er et kritisk punkt for f. La videre $Df(a, b) = \det(Hf(a, b))$ Da har vi

- *Dersom $D(a, b) > 0$ og $\frac{\partial^2 f}{\partial x^2}(a, b) > 0$, så er (a, b) et lokalt minimumspunkt.*
- *Dersom $D(a, b) > 0$ og $\frac{\partial^2 f}{\partial x^2}(a, b) < 0$, så er (a, b) et lokalt maksimumspunkt.*
- *Dersom $D(a, b) < 0$, så er (a, b) et sadelpunkt.*
- *Dersom $D(a, b) = 0$ sier ikke denne testen oss noe.*

Bevis.

Taylorpolynomet av grad 2 til $g(t) = f(\mathbf{a} + t\mathbf{y})$ utviklet om 0, beregnet for $t = 1$ er gitt ved

$$g(1) = g(0) + g'(0) + \tfrac{1}{2} g''(c)$$

for en $c \in [0, 1]$.

Vi har

$$\begin{aligned} g'(t) &= \nabla f(\mathbf{a} + t\mathbf{y}) \cdot \mathbf{y} \\ g''(t) &= (\nabla f'(\mathbf{a} + t\mathbf{y})\mathbf{y}) \cdot \mathbf{y} = (Hf(\mathbf{a} + t\mathbf{y})\mathbf{y}) \cdot \mathbf{y} \end{aligned}$$

som gir

$$f(\mathbf{a} + \mathbf{y}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{y} + \tfrac{1}{2} (Hf(\mathbf{a} + c\mathbf{y})\mathbf{y}) \cdot \mathbf{y}$$

$\square$

Vi setter

$$A(\mathbf{y}) = \frac{1}{2}(Hf(\mathbf{a}+c\mathbf{y}) - Hf(\mathbf{a}))$$

og observerer at $A_{ij}(\mathbf{y}) \to 0$ når $\mathbf{y} \to \mathbf{0}$ for alle elementer i A.

Vi setter

$$\varepsilon(\mathbf{y}) = \frac{A(\mathbf{y})\mathbf{y}\cdot\mathbf{y}}{|\mathbf{y}|^2} = A(\mathbf{y})\frac{\mathbf{y}}{|\mathbf{y}|}\cdot\frac{\mathbf{y}}{|\mathbf{y}|}$$

og som en konsekvens

$$|\varepsilon(\mathbf{y})| \leq |\sum_{i,j=1}^{m} A_{ij}(\mathbf{y})|\frac{|\mathbf{y_i}|}{|\mathbf{y}|}\cdot\frac{|\mathbf{y_j}|}{|\mathbf{y}|} \leq \sum_{i,j=1}^{m}|A_{ij}(\mathbf{y})| \to 0 \text{ når } \mathbf{y} \to \mathbf{0}$$

og vi kan skrive

$$f(\mathbf{a}+\mathbf{y}) = f(\mathbf{a}) + \nabla f(\mathbf{a})\cdot\mathbf{y} + \frac{1}{2}(Hf(\mathbf{a})\mathbf{y})\cdot\mathbf{y} + \varepsilon(\mathbf{y})|\mathbf{y}|^2$$

Setning

La A være en symmetrisk $m \times m$-matrise.

a) *Bare positive egenverdier:* $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$. *Da er*

$$(A\mathbf{y}) \cdot \mathbf{y} \geq \lambda_1 |\mathbf{y}|^2$$

for alle $\mathbf{y} \in \mathbb{R}^m$.

b) *Bare negative egenverdier:* $0 > \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$. *Da er*

$$(A\mathbf{y}) \cdot \mathbf{y} \leq \lambda_1 |\mathbf{y}|^2$$

for alle $\mathbf{y} \in \mathbb{R}^m$.

Følger ved å skrive **y** som en sum av egenverdier.

I et stasjonært punkt $\mathbf{a}$ har vi $\nabla f(\mathbf{a}) = 0$ og derfor har vi i tilfelle (i)

$$\begin{aligned} f(\mathbf{a}+\mathbf{y}) - f(\mathbf{a}) &= \tfrac{1}{2}(Hf(\mathbf{a})\mathbf{y}) \cdot \mathbf{y} + \varepsilon(\mathbf{y})|\mathbf{y}|^2 \\ &\leq \lambda_1 |\mathbf{y}|^2 + \varepsilon(\mathbf{y})|\mathbf{y}|^2 \\ &= (\lambda_1 + \varepsilon(\mathbf{y}))|\mathbf{y}|^2 \end{aligned}$$

som har samme tegn som λ_1 når $|\mathbf{y}|$ er tilstrekkelig liten (husk at $|\varepsilon(\mathbf{y})| \to 0$ når $\mathbf{y} \to \mathbf{0}$)

Eksempel

Vi studerer funksjonen $f(x,y) = x^2 - y^2$. *De partielt deriverte er gitt ved*

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = -2y$$

Det betyr at vi har et kritisk punkt for $(x,y) = (0,0)$.
De dobbeltderiverte er gitt ved

$$\frac{\partial^2 f}{\partial x^2} = 2, \quad \frac{\partial^2 f}{\partial y^2} = -2$$

Det gir for Hesse-determinanten i punktet $(0,0)$;

$$D = 2 \cdot (-2) - 0^2 = -4 < 0$$

som sier at det kritiske punktet er et sadelpunkt.

Eksempel

Se på funksjonen $f(x,y) = x^3 - y^2 + xy$. *Vi setter de partielt deriverte* $\frac{\partial f}{\partial x} = 3x^2 + y$ *og* $\frac{\partial f}{\partial y} = -2y + x$ *til å være lik 0, noe som gir oss kritiske punkter* $(0,0)$ *og* $(-\frac{1}{6}, -\frac{1}{12})$. *Vi regner ut*

$$D(x,y) = 6x \cdot (-2) - 1^2 = -12x - 1$$

Dette gir $D(0,0) = -1$ *og* $H(-\frac{1}{6}, -\frac{1}{12}) = 1$. *Samtidig har vi* $\frac{\partial^2 f}{\partial x^2}(-\frac{1}{6}, -\frac{1}{12}) = -1 < 0$. *Det betyr at* $(0,0)$ *er et sadelpunkt og* $(-\frac{1}{6}, -\frac{1}{12})$ *er et lokalt maksimum.*

Eksempel

La $f(x,y) = x + y + 2$*. Vi skal finne maksimumsverdien til* $f(x,y)$ *på enhetskvadratet* $[0,1] \times [0,1]$*.*
De partielt deriverte til funksjonen er konstanter og det finnes derfor ikke noen lokale ekstremalpunkter inne i det indre av området. De fire kantene til enhetskvadratet er gitt ved $(x,0)$*,* $(1,y)$*,* $(x,1)$ *og* $(0,y)$*, hvor x og y ligger mellom 0 og 1. På de fire kantene vil funksjonen være* $f(x,0) = x + 2$*,* $f(1,y) = y + 3$*,* $f(x,1) = x + 3$ *og* $f(0,y) = y + 2$*. Det er enkelt å se at når x og y ligger mellom 0 og 1, så vil funksjonen ligge mellom 0 og 4. Det betyr at dette også er globale maksimums- og minimums-verdier.*

En interessant, men kanskje noe uventet anvendelse av partiell derivasjon er innen teorien for det som kalles **lineær regresjon** og da spesielt ved bruk av **minste kvadraters metode**. Vi skal se på dette via et konkret eksempel.
Forventet levealder for nyfødte jentebarn i Norge har utviklet seg slik de siste årene:

2001	81.53
2002	81.52
2003	81.93
2004	82.33
2005	82.52
2006	82.66
2007	82.66
2008	82.95

Vi skal bruke minste kvadraters metode til å finne en formel som beskriver disse tallene. Vi lar x være år og y være forventet levealder. Hypotesen vår er at det er en lineær sammenheng mellom disse, dvs. $y = ax + b$. Vi indekserer de 8 dataparene fra 1 til 8, slik at f.eks. $(x_3, y_3) = (2003, 81.93)$. For hvert par beregner vi avviket fra den antatte rette linja, gitt ved $ax_i + b - y_i$. Disse tallene kvadrerer vi og summerer over $1 = 1, 2, \ldots, 8$. Dernest skal vi finne de koeffisientene a og b slik at denne kvadratsummen blir minst mulig. Ideen er at vi betrakter kvadratsummen av avvikene fra den rette linja som en funksjon i de to konstantene a og b og bruker teorien til å finne minimumspunktet til kvadratsum-funksjonen uttrykt ved alle dataene vi har tilgjengelig.

For å gjøre regningene litt enklere bruker vi år 1,2,3 osv. Kvadratsummen blir

$$\sum_{i=1}^{8}(ax_i+b-y_i)^2 = a^2\sum_{i=1}^{8}x_i^2+2ab\sum_{i=1}^{8}x_i+\sum_{i=1}^{8}b^2 \\ -2a\sum_{i=1}^{8}x_iy_i-2b\sum_{i=1}^{8}y_i+\sum_{i=1}^{8}y_i^2$$

Vi setter inn verdiene og får dette kvadratiske avviket som en funksjon i a og b gitt ved

$$Q(a,b)=204a^2+72ab+8b^2-5941a-1316b+54139$$

Vi skal finne minimum for denne funksjonen.

Vi partiellderiverer med hensyn på a og b og setter uttrykkene lik 0. Det gir

$$\frac{\partial Q}{\partial a} = 408a + 72b - 5941$$
$$\frac{\partial Q}{\partial b} = 72a + 16b - 1316$$

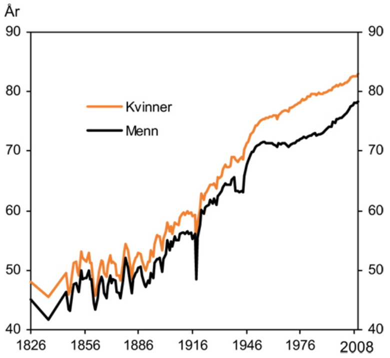
som gir $a = 0.226$ og $b = 81.23$ og forventet levealder som en funksjon av årstall blir

$$y = 0.226(x - 2000) + 81.23 = 0.226x - 370.77$$

Sammenlikner vi med de oppgitte tallene får vi

2001	81.53	81.46
2002	81.52	81.68
2003	81.93	81.91
2004	82.33	82.13
2005	82.52	82.36
2006	82.66	82.59
2007	82.66	82.81
2008	82.95	83.04

Forventet levealder ved fødselen



Vi ser at modellen gir oss en meget bra tilnærming tildataene.
Vi kan sette inn $x = 1866$ som er SSBs første registrering. Det gir forventet levealder $y = 50,95$, mens SSB oppgir 50,65 som gjennomsnitt for perioden 1866-1870. Vår lineære tilnærming er altså ut til å stemme godt.
I den motsatte tidsretning ser vi at $y = 100$ gir $x = 2083$, så med samme utvikling vil jentebarn født i 2083 ha en forventet levealder på 100 år.

Gjør vi den samme analysen i en generell setting med datapunkter

$$(x_i, y_i) \quad i = 1, 2, \ldots, n$$

får vi ut likninger for a og b;

$$a = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$
$$b = \frac{\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i}{n}$$

som kan brukes i alle eksempler.