MS-A0201
Differentiaali- ja integraalilaskenta 2 (TFM)
Luento 5: Gradientti ja suunnattu derivaatta. Vektoriarvoiset funktiot. Taylor-approksimaatio.

Harri Hakula

Matematiikan ja systeemianalyysin laitos[1]
Aalto-yliopisto

Kevät 2023

[1] Perustuu Antti Rasilan luentomonisteeseen vuodelta 2015 sekä Jarmo Malisen versioon vuodelta 2017.

Vektoriarvoiset funktiot

Olkoon $D \subset \mathbb{R}^n$ ja $\mathbf{f} = (f_1, f_2, \ldots, f_m)$ vektori, missä jokainen funktion f komponentti on funktio $f_j \colon D \to \mathbb{R}$ ja $m, n \geq 2$.

- Tällainen vektori määrittelee vektoriarvoisen funktion $\mathbf{f} \colon \mathbb{R}^n \to \mathbb{R}^m$, jota kutsutaan myös vektorikentäksi.
- Usein käytetään merkintää $\mathbf{y} = \mathbf{f}(\mathbf{x})$.
- Vektoriarvoisia funktiota esiintyy usein mm. fysiikassa sellaisten suureiden yhteydessä, joilla on voimakkuus ja suunta (esimerkiksi nopeus- ja voimakentät).

Huomaa, että f_j:t ovat tässä vektorin $\mathbf{f}$ komponentteja (eivät siis osittaisderivaattoja).

Vektoriarvoisen funktion derivointi

- Derivaatan luonnollinen vastine vektoriarvoisen funktion $\mathbf{f} = (f_1, f_2, \ldots, f_m)$ tapauksessa on Jacobin matriisi

$$D\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n}. \end{bmatrix}$$

- Jos $m = n$, Jacobin matriisi on neliömatriisi ja sen determinattia sanotaan funktion $\mathbf{f}$ Jacobin determinantiksi pisteessä $\mathbf{x}$. Tätä determinanttia tarvitaan kurssin loppuosassa.
- Jacobin matriiseilla ketjusääntö voidaan kirjoittaa yleisessä muodossa

$$D(\mathbf{f} \circ \mathbf{g})(\mathbf{x}) = D\mathbf{f}\big(\mathbf{g}(\mathbf{x})\big)\, D\mathbf{g}(\mathbf{x}).$$

Sovellus: implisiittifunktiolause

Oletetaan, että skalaarifunktiot $F_{(1)}, F_{(2)}, \ldots, F_{(n)}$ ovat derivoituvia.

- Tutkitaan yhtälöryhmää

$$\begin{cases} F_{(1)}(x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n) = 0, \\ F_{(2)}(x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n) = 0, \\ \vdots \\ F_{(n)}(x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n) = 0, \end{cases}$$

  pisteen $P_0 = (a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_n)$ lähellä.
- Muuttujat $\mathbf{y} = (y_1, \ldots, y_n)$ voidaan esittää muuttujien $\mathbf{x} = (x_1, \ldots, x_m)$ funktioina pisteen P_0 lähellä, jos funktion $\mathbf{F}(\mathbf{y}) = (F_{(1)}, \ldots, F_{(n)})(\mathbf{y})$ Jacobin determinatti

$$\det D\mathbf{F}(\mathbf{y})\Big|_{P_0} \neq 0.$$

Esimerkki 1

Osoitetaan, että (u, v) voidaan esittää muuttujien (x, y, z) funktiona systeemistä

$$\begin{cases} F(x, y, z, u, v) = xy^2 + xzu + yv^2 - 3 = 0, \\ G(x, y, z, u, v) = x^3yz + 2xv - u^2v^2 - 2 = 0, \end{cases}$$

pisteen $P_0 = (1, 1, 1, 1, 1)$ lähellä.

- Selvästi $F(P_0) = G(P_0) = 0$.
- Muodostetaan Jacobin determinatti

$$\left| \begin{array}{cc} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \end{array} \right| \Bigg|_{P_0} = \left| \begin{array}{cc} xz & 2yv \\ -2uv^2 & 2x - 2u^2v \end{array} \right| \Bigg|_{P_0} = \left| \begin{array}{cc} 1 & 2 \\ -2 & 0 \end{array} \right| = 4.$$

- Koska determinantti ei ole nolla, voidaan kirjoittaa $u = u(x, y, z)$ ja $v = v(x, y, z)$ kolmen muuttujan funktioina. Kaavoja näille funktioille ei kuitenkaan voida yleensä antaa.

Gradientti

- Olkoon $f \colon D \subset \mathbb{R}^n \to \mathbb{R}$, $n \geq 2$, derivoituva pisteessä $\mathbf{x} \in D$.

Määritelmä

Funktion f gradientti pisteessä $\mathbf{x}$ on vektori

$$\nabla f = \operatorname{grad} f = \left(\frac{\partial}{\partial x_1} f, \frac{\partial}{\partial x_2} f, \ldots, \frac{\partial}{\partial x_n} f \right) \in \mathbb{R}^n.$$

- Gradientti kertoo funktion f nopeimman kasvun suunnan. Se on vektoriarvoinen funktio $\nabla f \colon D \to \mathbb{R}^n$.
- Tapauksessa $n = 3$ voidaan kirjoittaa

$$\nabla = \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z}.$$

- Tapauksessa $n = 2$ kolmas termi jää pois.
- Gradientti on Jacobin matriisin erikoistapauksena kun $m = 1$.

Esimerkki 2

Olkoon $f(x, y) = x^2 + y^2$.

- Saadaan $\nabla f = 2x\mathbf{i} + 2y\mathbf{j}$.
- Erityisesti $\nabla f(a, b)$ on kohtisuorassa origokeskisen (yksikkö)ympyrän mielivaltaiseen pisteeseen (a, b) piirrettyä tangenttisuoraa vastaan.

Viimeinen väite on erikoistapaus yleisemmästä tasa-arvokäyriä koskevasta totuudesta.

Tasa-arvokäyrät (kertausta 2. luennolta)

Olkoon $c \in \mathbb{R}$ vakio, $D \subset \mathbb{R}^2$ ja $f\colon D \to \mathbb{R}$ funktio.

- Tällöin joukko $C = \{(x, y) : f(x, y) = c\}$ on usein tasokäyrä.
- Kyseinen pistejoukko voi olla myös tyhjä (jos f ei saa arvoa c) tai vaikkapa koko taso (jos f on vakio).
- Mikäli joukko C on tasokäyrä, sitä sanotaan funktion f arvoon c liittyväksi tasa-arvokäyräksi.
- **Esim.** Korkeuskäyrät kartalla ovat tasa-arvokäyriä funkiolle, joka liittää kartalla olevaan pisteeseen (x, y) sen korkeuden meren pinnasta.
- **Huom.** Jos tulivuoren kraaterissa on järvi, niin veden pinta on vakiokorkeudella meren pinnasta. Tulee ongelmia tulkita järven pinta tasa-arvokäyränä... mutta ei ajatella sitä nyt. Ajatellaan sen sijaan levadaa Madeiran rinteillä.

Gradientti ja tasa-arvokäyrät

Lause

Olkoon $D \subset \mathbb{R}^2$, $(a, b) \in D$ ja $f: D \to \mathbb{R}$ derivoituva pisteessä (a, b) siten, että $\nabla f(a, b) \neq \mathbf{0}$. Tällöin $\nabla f(a, b)$ on kohtisuorassa pisteen (a, b) kautta kulkevaa funktion f tasa-arvokäyrää (t.s., sen tangenttia) vasten.

Seuraus: Jos piste $\mathbf{x} \in D$ on funktion f paikallinen ääriarvo (minimi tai maksimi), niin $\nabla f(\mathbf{x}) = \mathbf{0}$.

Gradientin nollakohta ei kuitenkaan välttämättä ole funktion ääriarvo. Edes skalaarifunktion derivaatan nollakohta ei välttämättä ole minimi eikä maksimi, kuten nähdään jos $f(x) = x^3$.

Todistus

Olkoon $I = [-1, 1]$ ja $\mathbf{r}(t)\colon I \to \mathbb{R}^2$ tasa-arvokäyrän sellainen parametrisointi, että $\mathbf{r}(0) = (a, b)$.

Koska $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ on tasa-arvokäyrä, kaikilla $t \in I$ pätee $f(x(t), y(t)) = f(a, b)$ eli vakio.

Ketjusäännöstä saadaan (koska vakiofunktion derivaatta on nolla)

$$f_1\big(x(t), y(t)\big)x'(t) + f_2\big(x(t), y(t)\big)y'(t) = 0.$$

Erityisesti pisteessä $t = 0$ tämä tarkoittaa, että

$$\nabla f(a, b) \cdot \mathbf{r}'(0) = 0,$$

eli toisin sanoen vektori ∇f ja tangentin suuntainen $\mathbf{r}'(0)$ ovat kohtisuorassa. $\square$

Suunnattu derivaatta

- Edellinen tulos voidaan tulkita niin, että tasa-arvokäyrä(n tangentti) antaa suunnan, johon edettäessä funktio ei kasva eikä vähene. Niinpä funktio kasvaa jyrkimmin gradienttinsa suuntaan, joka on tasa-arvokäyrän normaalivektori.
- Muihin suuntiin liikuttessa kasvunopeuden antaa *suunnattu derivaatta*

$$D_{\mathbf{u}} f(a, b) = \frac{dg}{dt}(0) \text{ jossa } g(t) = f(a + tu_1, b + tu_2).$$

jossa $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$ on yksikkösuuntavektori.

Lause

Olkoon $f\colon D \subset \mathbb{R}^2 \to \mathbb{R}$ funktio, $(a, b) \in D$ ja $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$ sellainen vektori, että $\|\mathbf{u}\|^2 = u_1^2 + u_2^2 = 1$. Tällöin funktion f suunnattu derivaatta suuntaan $\mathbf{u}$ saadaan kaavasta

$$D_{\mathbf{u}} f(a, b) = \mathbf{u} \cdot \nabla f(a, b).$$

Esimerkki 3 1/2

- Olkoon $f(x,y) = y^4 + 2xy^3 + x^2y^2$. Etsitään $D_{\mathbf{u}}f(0,1)$, kun $\mathbf{u}$ on (a) $\mathbf{i} + 2\mathbf{j}$, (b) $\mathbf{j} - 2\mathbf{i}$, (c) $3\mathbf{i}$, (d) $\mathbf{i} + \mathbf{j}$.
- **Ratkaisu:** Lasketaan

$$\nabla f(x,y) = (2y^3 + 2xy^2)\mathbf{i} + (4y^3 + 6xy^2 + 2x^2y)\mathbf{j},$$

$$\nabla f(0,1) = 2\mathbf{i} + 4\mathbf{j}.$$

(a) $\|\mathbf{i} + 2\mathbf{j}\| = \sqrt{5}$ ja siten $\mathbf{u} = (\mathbf{i} + 2\mathbf{j})/\sqrt{5}$. Saadaan

$$D_{\mathbf{u}}f(0,1) = \frac{1}{\sqrt{5}}(\mathbf{i} + 2\mathbf{j}) \cdot (2\mathbf{i} + 4\mathbf{j}) = \frac{2+8}{\sqrt{5}} = 2\sqrt{5}.$$

Huomaa, että tässä $\mathbf{u}$ ja $\nabla f(0,1)$ ovat yhdensuuntaiset.

Esimerkki 3 2/2

(b) $\|\mathbf{j} - 2\mathbf{i}\| = \sqrt{5}$ ja siten $\mathbf{u} = (\mathbf{j} - 2\mathbf{i})/\sqrt{5}$. Saadaan

$$D_{\mathbf{u}} f(0,1) = \frac{1}{\sqrt{5}}(\mathbf{j} - 2\mathbf{i}) \cdot (2\mathbf{i} + 4\mathbf{j}) = \frac{-4+4}{\sqrt{5}} = 0.$$

Vektorit $\mathbf{u}$ ja $\nabla f(0,1)$ ovat siis kohtisuorassa.

(c) $\|3\mathbf{i}\| = 3$ ja siten $\mathbf{u} = \mathbf{i}$. Saadaan $D_{\mathbf{u}} f(0,1) = \mathbf{i} \cdot (2\mathbf{i} + 4\mathbf{j}) = 2$. Tämä on sama kuin $f_1(0,1)$.

(d) $\|\mathbf{i} + \mathbf{j}\| = \sqrt{2}$ ja siten $\mathbf{u} = (\mathbf{i} + \mathbf{j})/\sqrt{2}$. Saadaan

$$D_{\mathbf{u}} f(0,1) = \frac{1}{\sqrt{2}}(\mathbf{i} + \mathbf{j}) \cdot (2\mathbf{i} + 4\mathbf{j}) = \frac{2+4}{\sqrt{2}} = 3\sqrt{2}.$$

Huomaa, että $3\sqrt{2} \approx 4.243 < 2\sqrt{5} \approx 4.472$.

Taylorin kaava

- Yhden muuttujan tapauksessa $m+1$ kertaa jatkuvasti derivoituvaa funktiota $f\colon I \to \mathbb{R}$ voidaan approksimoida kaavalla

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(m)}(a)}{m!}(x-a)^m.$$

  kun $a, x \in I$.

- Tämä idea yleistyy usean muuttujan tapaukseen: Jos $\mathbf{a}, \mathbf{h} \in \mathbb{R}^n$, $n \geq 2$ ja funktiolla $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$ on jatkuvat kertaluvun $(m+1)$ osittaisderivaatat pisteitä $\mathbf{a}, \mathbf{a}+\mathbf{h}$ yhdistävällä janalla, niin

$$f(\mathbf{a}+\mathbf{h}) \approx \sum_{j=0}^{m} \frac{(\mathbf{h}\cdot\nabla)^j f(\mathbf{a})}{j!}.$$

Esimerkki 4

- Olkoon $(a, b) \in \mathbb{R}^2$ ja $f(x, y)$ neljä kertaa jatkuvasti derivoituva kiekossa (a, b)-keskisessä r-säteisessä kiekossa.
- Etsitään 3. asteen approksimaatio. Jos $\mathbf{h} = (h, k)$, niin

$$f(a+h, b+k) \approx f(a, b) + (hD_1 + kD_2)f(a, b) + \frac{1}{2!}(hD_1 + kD_2)^2 f(a, b)$$

$$+\frac{1}{3!}(hD_1 + kD_2)^3 f(a, b)$$

$$= f(a, b) + hf_1(a, b) + kf_2(a, b) + \frac{1}{2!}\Big(h^2 f_{11}(a, b) + 2hkf_{12}(a, b) + k^2 f_{22}(a, b)\Big)$$

$$+\frac{1}{3!}\Big(h^3 f_{111}(a, b) + 3h^2 kf_{112}(a, b) + 3hk^2 f_{122}(a, b) + k^3 f_{222}(a, b)\Big).$$

- **Huom.** 1. asteen Taylor-approksimaatio on sama kuin tangenttitaso.

Esimerkki 5 1/2

- Etsitään 2. asteen Taylor-approksimaatio funktiolle $f(x, y) = \sqrt{x^2 + y^3}$ pisteen $(1, 2)$ ympäristössä.
- Lasketaan $f(1, 2) = 3$,

$$f_1(x, y) = \frac{x}{\sqrt{x^2 + y^3}}, \quad f_2(x, y) = \frac{3y^2}{2\sqrt{x^2 + y^3}},$$

eli $f_1(1, 2) = 1/3$ ja $f_2(1, 2) = 2$.

Edelleen

$$f_{11}(x, y) = \frac{y^3}{(x^2 + y^3)^{3/2}}, \quad f_{11}(1, 2) = \frac{8}{27},$$

Esimerkki 5 2/2

$$f_{12}(x,y) = \frac{-3xy^2}{2(x^2+y^3)^{3/2}}, \quad f_{12}(1,2) = -\frac{2}{9},$$

$$f_{22}(x,y) = \frac{12x^2y+3y^4}{4(x^2+y^3)^{3/2}}, \quad f_{22}(1,2) = \frac{2}{3}.$$

- Siten

$$f(1+h,2+k) \approx 3 + \frac{1}{3}h + 2k + \frac{1}{2!}\Big(\frac{8}{27}h^2 + 2\Big(-\frac{2}{9}\Big)hk + \frac{2}{3}k^2\Big).$$