

MS-A0201  
Differentiaali- ja integraalilaskenta 2 (TFM)  
Luento 7: Lagrangen kertojat.  
Pienimmän neljösumman menetelmä.

Harri Hakula

Matematiikan ja systeemianalyysin laitos<sup>1</sup>  
Aalto-yliopisto

Kevät 2022

---

<sup>1</sup>Perustuu Antti Rasilan luentomonisteeseen vuodelta 2015 sekä Jarmo Malisen versioon vuodelta 2017.

## Lagrangen kertojat 1/3

Usein optimointitehtävissä halutaan asettaa rajoitusehtoja optimoitaville muuttujille.

- Tyypillinen esimerkki tällaisesta tehtävästä on peltipurkin muodon optimointi: Halutaan minimoida purkin pinta-ala (eli käytetty materiaali)  $A(h, r) = 2\pi rh + 2\pi r^2$  niin, että tilavuus  $V(r, h) = \pi r^2 h$  on vakio.
- Duaalitehtävä: Halutaan maksimoida purkin tilavuus  $V(r, h)$  siten, että pinta-ala  $A(h, r)$  on vakio.
- Primaali- ja duaalitehtävillä on sama ratkaisu. Tämän sanoo maalaisjärkikin, mutta itse asiassa ratkaisuun johtavat yhtälötkin ovat (olennaisesti) samoja.

## Lagrangen kertojat 2/3

“Minimoi  $f(x, y)$  ehdolla  $g(x, y) = 0$ .”

- Havaitaan, että mikäli ongelmalla on ratkaisu, niin ratkaisupisteessä  $(a, b)$  vektorien  $\nabla f$  ja  $\nabla g$  on oltava joko yhdensuuntaisia tai vastakkaissuuntaisia (mikäli  $\nabla g(a, b) \neq 0$ ).
- Miksi? Koska muussa tapauksessa funktiolla  $f$  olisi nollasta poikkeava suunnattu derivaatta käyrän  $g(x, y) = 0$  tangentin suuntaan pisteessä  $(a, b)$ , ja siis minimi ei voi olla pisteessä  $(a, b)$ .
- Entä jos tehtävänä olisi maksimoida  $f(x, y)$  ehdolla  $g(x, y) = 0$ ?
- Entä jos tehtävänä olisi maksimoida  $g(x, y)$  ehdolla  $f(x, y) = c$ ?

## Lagrangen kertojat 3/3

- Mikäli optimipiste on olemassa, se on Lagrangen funktion

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

kriittinen piste (eli gradientin nollakohta).

- Menetelmä yleistyy myös useammalle muuttujalle. Esimerkiksi kolmen muuttujan tapauksessa Lagrangen funktio on

$$L(x, y, z, \lambda, \mu) = f(x, y, z) + \lambda g(x, y, z) + \mu h(x, y, z),$$

missä  $f$  on minimoitava funktio ja rajoite-ehdot ovat  $g(x, y, z) = 0$  sekä  $h(x, y, z) = 0$ .

## Esimerkki 1 1/2

Minimoidaan funktio  $f(x, y) = x^2 + y^2$  ehdolla  $g(x, y) = x^2y - 16 = 0$ .

- Muodostetaan Lagrangen funktio

$$L(x, y, \lambda) = x^2 + y^2 + \lambda(x^2y - 16).$$

- Yhtälöt kriittisille pisteille ovat

$$0 = \frac{\partial L}{\partial x} = 2x(1 + \lambda y),$$

$$0 = \frac{\partial L}{\partial y} = 2y + \lambda x^2,$$

$$0 = \frac{\partial L}{\partial \lambda} = x^2y - 16,$$

joista viimeinen on aina itse rajoitusehto.

## Esimerkki 1 2/2

- Ensimmäisestä yhtälöstä saadaan  $x = 0$  tai  $\lambda y = -1$ , mutta  $x = 0$  on ristiriidassa kolmannen yhtälön kanssa.
- Siten toisesta yhtälöstä

$$0 = 2y^2 + \lambda yx^2 = 2y^2 - x^2.$$

- Tästä saadaan edelleen  $x = \pm\sqrt{2}y$ , ja  $2y^3 = 16$  eli  $y = 2$ .
- Ääriarvoja (mahdollisia minimejä) on siis kaksi  $(x, y) = (\pm 2\sqrt{2}, 2)$ . Pitää selvittää muilla keinoin, ovatko minimejä vai maksimeja.

## Esimerkki 2

Yritetään etsiä Lagrangen kertojien menetelmällä funktion  $f(x, y) = y$  minimi ehdolla  $g(x, y) = y^3 - x^2 = 0$ .

- **Selvästi nähdään**, että minimi  $f(x, y) = 0$  saavutetaan pisteessä  $(0, 0)$ .
- Muodostetaan Lagrangen funktio

$$L(x, y, \lambda) = y + \lambda(y^3 - x^2).$$

- Saadaan yhtälöt

$$-2\lambda x = 0, \quad 1 + 3\lambda y^2 = 0, \quad \text{ja} \quad y^3 - x^2 = 0.$$

- Nämä yhtälöt ovat keskenään ristiriidassa, joten ratkaisua niille ei ole.
- Huomaa, että  $\nabla g(0, 0) = \mathbf{0}$  minimipisteessä.
- Tästä nähdään, Lagrangen kertojat näkevät ääriarvoja vain pisteissä, joissa  $\nabla g(0, 0) \neq \mathbf{0}$ .

## Esimerkki 3 1/4

Etsitään ääriarvot funktiolle  $f(x, y, z) = xy + 2z$  ehdoilla  $x + y + z = 0$  ja  $x^2 + y^2 + z^2 = 24$ .

- Koska  $f$  on jatkuva ja annettujen leikkausjoukkojen leikkaus on ympyräviiva (eli rajoitettu ja suljettu joukko), niin ääriarvot ovat olemassa.
- Muodostetaan Lagrangen funktio

$$L(x, y, z, \lambda, \mu) = xy + 2z + \lambda(x + y + z) + \mu(x^2 + y^2 + z^2 - 24).$$



## Esimerkki 3 2/4

- Lagrangen funktion osittaisderivaatoista saadaan yhtälöt

$$y + \lambda + 2\mu x = 0,$$

$$x + \lambda + 2\mu y = 0,$$

$$2 + \lambda + 2\mu z = 0,$$

$$x + y + z = 0, \text{ ja}$$

$$x^2 + y^2 + z^2 - 24 = 0.$$

- Kahden ensimmäisen yhtälön erotus johtaa yhtälöön  $(x - y)(1 - 2\mu) = 0$ , joten joko  $\mu = 1/2$  tai  $x = y$ . Tutkitaan molemmat tapaukset.

## Esimerkki 3 3/4

- **Tapaus I** ( $\mu = 1/2$ ): Toisen ja kolmannen yhtälön perusteella

$$x + \lambda + y = 0 \text{ ja } 2 + \lambda + z = 0, \text{ siis } x + y = 2 + z.$$

- Neljännestä yhtälöstä saadaan  $z = -1$  ja  $x + y = 1$ . Viimeisen yhtälön perusteella  $x^2 + y^2 = 24 - z^2 = 23$ .
- Koska  $x^2 + y^2 + 2xy = (x + y)^2 = 1$ , saadaan  $2xy = 1 - 23 = -22$  ja  $xy = -11$ .
- Nyt  $(x - y)^2 = x^2 + y^2 - 2xy = 23 + 22 = 45$ , joten  $x - y = \pm 3\sqrt{5}$ .
- Yhdessä yhtälön  $x + y = 1$  tästä saadaan kaksi kriittistä pistettä  
 $P_1 = ((1+3\sqrt{5})/2, (1-3\sqrt{5})/2, -1)$  ja  $P_2 = ((1-3\sqrt{5})/2, (1+3\sqrt{5})/2, -1)$ .
- Kummassakin pisteessä  $f(x, y, z) = -11 - 2 = -13$ .

## Esimerkki 3 4/4

- **Tapaus II** ( $x = y$ ): Neljännestä yhtälöstä nähdään, että  $z = -2x$ , ja viimeisen yhtälön perusteella  $6x^2 = 24$  eli  $x = \pm 2$ .
- Näin ollen, kriittiset pisteet ovat

$$P_3 = (2, 2, -4) \text{ ja } P_4 = (-2, -2, 4).$$

- Saadaan

$$f(2, 2, -4) = 4 - 8 = -4 \text{ ja } f(-2, -2, 4) = 4 + 8 = 12.$$

- Siten funktion  $f$  maksimi on 12 ja minimi  $-13$ .

# Regressio-ongelma

- Regressioanalyysissä pyritään valitsemaan parametrin  $\beta$  arvo siten, että käyrä

$$y = f(x; \beta)$$

kulkisi mahdollisimman läheltä jokaista havaintopistettä

$$(x_j, y_j) \in \mathbb{R}^2, j = 1, 2, \dots, n.$$

- Tällaista optimaalisesti valittua käyrää kutsutaan *regressiomalliksi*  $y = f(x; \beta)$ , jossa funktion  $f$  muoto on valittu tilanteen ja harkinnan mukaan.
- Kunhan  $f$  valittu, niin eräs ratkaisu käyränsovitukseen on *pienimmän neliösumman menetelmä*.

# Pienimmän neliösumman menetelmä

- Pienimmän neliösumman menetelmässä pyritään minimoimaan regressiomallin virhetermien  $\varepsilon_j$

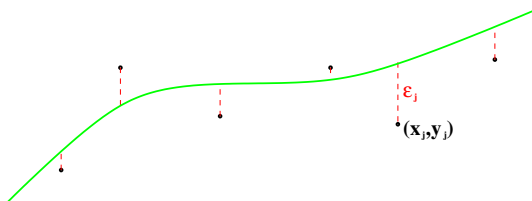
$$\varepsilon_j = y_j - f(x_j; \beta), \quad j = 1, 2, \dots, n$$

neliösummaa eli funktiota

$$F(\beta) = \sum_{j=1}^n \varepsilon_j^2 = \sum_{j=1}^n (y_j - f(x_j; \beta))^2.$$

muuttamalla parametrivektorin  $\beta = (\beta_0, \beta_1, \dots, \beta_m)$  arvoa.

- Optimaalinen  $\beta$ :n arvo on parametrin  $\beta$  *pienimmän neliösumman estimaatti* eli *PNS-estimaatti*.
- **Kysymys:** Miksi ei minimoitaisi lauseketta  $\sum_{j=1}^n |y_j - f(x_j; \beta)|$  neliösumman sijasta?



Kuvassa vihreällä parametreista  $\beta = (\beta_1, \beta_2, \dots, \beta_m)$  riippuva sovitettava funktio  $f(x; \beta)$  eräällä kiinteällä parametrin arvolla.

Datapisteet  $(x_j, y_j)$  ja vastaavat virhetermit  $\varepsilon_j$ , kun  $j = 1, \dots, n$ .

## Lineaarinen regressio 1/2

Linearisessa regressiossa  $f(x; \beta) = \beta_0 - \beta_1 x$  jossa  $\beta = (\beta_0, \beta_1)$  ja neliösumma on

$$F(\beta_0, \beta_1) = \sum_i (y_i - \beta_0 - \beta_1 x_i)^2.$$

Etsitään piste  $(\beta_0, \beta_1)$  siten, että  $\nabla F(\beta_0, \beta_1) = 0$ .

- Lasketaan osittaisderivaatta

$$\frac{\partial}{\partial \beta_0} F(\beta_0, \beta_1) = 2(\beta_1 \sum_i x_i + n\beta_0 - \sum_i y_i).$$

- Ratkaistaan nollakohta

$$\beta_0 = \frac{1}{n} \sum_i y_i - \frac{\beta_1}{n} \sum_i x_i = \bar{y} - \beta_1 \bar{x}$$

missä  $\bar{x}$  on datavektorin  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  komponenttien aritmeettinen keskiarvo.

## Lineaarinen regressio 2/2

- Lasketaan seuraavaksi osittaisderivaatta

$$\frac{\partial}{\partial \beta_1} F(\beta_0, \beta_1) = 2\left(\beta_0 \sum_i x_i + \beta_1 \sum_i x_i^2 - \sum_i x_i y_i\right).$$

- Sijoittamalla  $\beta_0$ :n lauseke, saadaan

$$n\bar{x}\bar{y} - n\beta_1\bar{x}^2 + \beta_1 \sum_i x_i^2 - \sum_i x_i y_i = 0.$$

- Ratkaistaan nollakohta

$$\beta_1 = \frac{n\bar{x}\bar{y} - \sum_i x_i y_i}{n\bar{x}^2 - \sum_i x_i^2} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}.$$

Tarkista jälkimmäinen yhtälö!



## Esimerkki 4

Sovita PNS-suora dataan

|       |      |      |      |      |      |
|-------|------|------|------|------|------|
| $x_i$ | 0.0  | 1.0  | 2.0  | 3.0  | 4.0  |
| $y_i$ | 2.10 | 1.92 | 1.84 | 1.71 | 1.64 |

Estimoi (ekstrapoloi)  $y$  kun  $x = 5$ .

- Saadaan  $\bar{x} = 2.0$ ,  $\bar{y} = 1.842$ , ja

$$\beta_1 = \frac{-1.13}{10.0} = -0.113.$$

- Siten  $\beta_0 = 1.842 + 0.113 \cdot 2.0 = 2.068$ .
- Näin ollen  $y = -0.113x + 2.068$ , ja haluttu estimaatti pisteessä  $x = 5$  on  $y = -0.113 \cdot 5 + 2.068 = 1.503$ .

## Esimerkki 5: Toisen asteen sovitus 1/2

Tutkitaan lisäaineen määrän  $x$  vaikutusta kuivumisaikaan  $y$ . Eri lisäaineen määrillä  $x_i$  (grammaa) saatiin kuivumisajat  $y_i$  (tuntia),  $i = 1, \dots, 9$ :

|       |      |     |     |     |     |     |     |     |     |
|-------|------|-----|-----|-----|-----|-----|-----|-----|-----|
| $x_i$ | 0.0  | 1.0 | 2.0 | 3.0 | 4.0 | 5.0 | 6.0 | 7.0 | 8.0 |
| $y_i$ | 11.0 | 9.4 | 9.1 | 7.0 | 6.2 | 7.1 | 6.6 | 7.5 | 8.2 |

- Huomataan, että kuivumisajan riippuvuus lisäaineen määrästä on epälineaarista. Minimikohdan estimoimiseksi sovitetaan havaintoihin paraabeli  $y = \beta_0 + \beta_1 x + \beta_2 x^2$ .
- Pienimmän neliösumman yhtälöryhmä mallille on

$$\frac{\partial}{\partial \beta_k} \sum (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2 = 0, \quad k = 0, 1, 2.$$

## Esimerkki 5: Toisen asteen sovitus 2/2

- Näistä saadaan yhtälöryhmä

$$\begin{cases} n\beta_0 + \beta_1 \sum x_i + \beta_2 \sum x_i^2 & = \sum y_i, \\ \beta_0 \sum x_i + \beta_1 \sum x_i^2 + \beta_2 \sum x_i^3 & = \sum x_i y_i \\ \beta_0 \sum x_i^2 + \beta_1 \sum x_i^3 + \beta_2 \sum x_i^4 & = \sum x_i^2 y_i. \end{cases}$$

- Laskemalla yhtälöryhmän kertoimet havainnoista saadaan

$$\begin{cases} 9\beta_0 + 36\beta_1 + 204\beta_2 & = 72.1 \\ 36\beta_0 + 204\beta_1 + 1296\beta_2 & = 266.6 \\ 204\beta_0 + 1296\beta_1 + 8772\beta_2 & = 1515.4 \end{cases}$$

- Ratkaisuna ovat  $\beta_0 = 11.15$ ,  $\beta_1 = -1.806$  ja  $\beta_2 = 0.1803$ . Pienimmän neliösumman mielessä parhaiten havaintoihin liittyvä paraabeli on siis

$$y = 11.15 - 1.806x + 0.1803x^2.$$