

MS-A0207 Differentiaali- ja integraalilaskenta 2 (Chem) Esimerkkejä ym., osa II

G. Gripenberg

Aalto-yliopisto

9. helmikuuta 2016

💡: Esimerkki: Implisiittinen lineaarinen approksimointi

Funktio $z = z(x, y)$ määräytyy yhtälöstä $xz + y^2z^2 + (x + y)z^3 = 4$ siten, että $z(-1, -1) = -1$. Jos nyt $|y + 1| \leq 0.003$ niin meidän pitää määrittää (implisiittistä derivointia käyttäen) approksimatiivinen yläraja lausekkeelle $|x + 1|$ siten, että lineaarisella approksimoinnilla saamme tulokseksi $|z(x, y) + 1| \leq 0.003$.

Merkitään $f(x, y, z) = xz + y^2z^2 + (x + y)z^3 - 4$. Nyt $f(-1, -1, -1) = 1 + 1 + (-2)(-1) - 4 = 0$ ja jos $z = z(x, y)$ niin määritelmän mukaan $f(x, y, z) = 0$ joten lineaarinen approksimaatio antaa

$$0 = f(x, y, z) - f(-1, -1, -1) \approx f_x(-1, -1, -1)(x - (-1)) \\ + f_y(-1, -1, -1)(y - (-1)) + f_z(-1, -1, -1)(z - (-1)),$$

eli

$$f_z(-1, -1, -1)(z + 1) \approx -f_x(-1, -1, -1)(x + 1) - f_y(-1, -1, -1)(y + 1).$$

💡: Esimerkki: Implisiittinen lineaarinen approksimointi, jatk.

Nyt

$$f_x(x, y, z) = z + z^3, \quad \Rightarrow \quad f_x(-1, -1, -1) = -2,$$

$$f_y(x, y, z) = 2yz^2 + z^3 \quad \Rightarrow \quad f_y(-1, -1, -1) = -3,$$

$$f_z(x, y, z) = x + 2y^2z + 3(x + y)z^2 \quad \Rightarrow \quad f_z(-1, -1, -1) = -9.$$

Tästä seuraa, että

$$9|z(x, y) + 1| \lesssim 2|x + 1| + 3|y + 1| \quad \text{eli} \quad |z(x, y) + 1| \lesssim \frac{2}{9}|x + 1| + \frac{1}{3}|y + 1|$$

Jos nyt haluamme, että $|z(x, y) + 1| \leq 0.003$ niin on meidän pitävä, koska $|y + 1| \leq 0.003$, vaatia, että

$$\frac{2}{9}|x + 1| + \frac{1}{3} \cdot 0.003 \leq 0.003,$$

josta seuraa, että

$$|x + 1| \leq \frac{9}{2}(0.003 - \frac{1}{3} \cdot 0.003) = 0.009.$$

😊 Eulerin ketjusääntö

Oleta, että funktio F on sellainen, että yhtälöstä $F(x, y, z) = 0$ voimme ratkaista x muuttujien y ja z funktiona, eli $x = x(y, z)$ mutta myös y muuttujien x ja z funktiona, eli $y = y(x, z)$ sekä z muuttujien x ja y funktiona, eli $z = z(x, y)$.

Jos $x = x(y, z)$ niin saamme kun derivoimme yhtälön $F(x(y, z), y, z) = 0$ molemmat puolet muuttujan y suhteen

$$F_x(x(y, z), y, z) \frac{\partial x(y, z)}{\partial y} + F_y(x(y, z), y, z) = 0.$$

Tästä seuraa, että kun kirjoitamme $x(y, z)$:n sijasta x niin

$$\frac{\partial x(y, z)}{\partial y} = - \frac{F_y(x, y, z)}{F_x(x, y, z)}.$$

Samanlaisilla laskuilla saamme

$$\frac{\partial y(x, z)}{\partial z} = - \frac{F_z(x, y, z)}{F_y(x, y, z)},$$

😊 Eulerin ketjusääntö, jatk.

ja

$$\frac{\partial z(x, y)}{\partial x} = -\frac{F_x(x, y, z)}{F_z(x, y, z)}.$$

Tästä seuraa nyt, että

$$\frac{\partial x(y, z)}{\partial y} \frac{\partial y(x, z)}{\partial z} \frac{\partial z(x, y)}{\partial x} = (-1)^3 \frac{F_y(x, y, z)}{F_x(x, y, z)} \frac{F_z(x, y, z)}{F_y(x, y, z)} \frac{F_x(x, y, z)}{F_z(x, y, z)} = -1.$$

Toisella tavalla esitettynä tämä kaava on

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1.$$

😊 Esimerkki: Implisiittinen lineaarinen approksimointi

Karttapaperi on venynyt niin, että origosta pisteestä (x, y) kulkevan janan pituus on kasvanut 2 % ja ja tämän janan ja x -akselin välinen kulma, joka alunperin oli 45° , on pienentynyt 1 %. Nyt meidän pitää laskea paljonko paperi on venynyt x -akselin suunnassa ja paljonko y -akselin suunnassa. Jos janan pituus on L ja kulma on φ niin pätee tietenkin

$$L = \sqrt{x^2 + y^2},$$
$$\varphi = \arctan\left(\frac{y}{x}\right).$$

Jos merkitsemme $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ ja $\mathbf{u} = \begin{bmatrix} L \\ \varphi \end{bmatrix}$, niin voimme kirjoittaa nämä yhtälöt muodossa

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} \sqrt{x^2 + y^2} - L \\ \arctan\left(\frac{y}{x}\right) - \varphi \end{bmatrix} = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

😊 Esimerkki: Implisiittinen lineaarinen approksimointi, jatk.

Kun karttapaperi muuttuu niin x ja y muuttuvat pisteiksi $x + \Delta x$ ja $y + \Delta y$ ja etäisyydeksi sekä kulmaksi tulee $L + \Delta L$ ja $\varphi + \Delta\varphi$ jolloin yhtälöksi tulee

$$\mathbf{f}(\mathbf{x} + \Delta\mathbf{x}, \mathbf{u} + \Delta\mathbf{u}) = \mathbf{0}.$$

Lineaarisella approksimoinnilla saamme nyt

$$\mathbf{0} = \mathbf{f}(\mathbf{x} + \Delta\mathbf{x}, \mathbf{u} + \Delta\mathbf{u}) - \mathbf{f}(\mathbf{x}, \mathbf{u}) \approx \mathbf{f}_x(\mathbf{x}, \mathbf{u})\Delta\mathbf{x} + \mathbf{f}_u(\mathbf{x}, \mathbf{u})\Delta\mathbf{u}.$$

Yksinkertaisella laskulla toteamme, että

$$\mathbf{f}_x(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{bmatrix} \quad \text{ja} \quad \mathbf{f}_u(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix},$$

jolloin yhtälösystemiksi tulee

$$\begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \approx \begin{bmatrix} \Delta L \\ \Delta\varphi \end{bmatrix}.$$

😊 Esimerkki: Implisiittinen lineaarinen approksimointi, jatk.

Koska

$$\mathbf{f}_x(\mathbf{x}, \mathbf{u})^{-1} = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{bmatrix}^{-1} = \sqrt{x^2+y^2} \begin{bmatrix} \frac{x}{x^2+y^2} & -\frac{y}{\sqrt{x^2+y^2}} \\ \frac{y}{x^2+y^2} & \frac{x}{\sqrt{x^2+y^2}} \end{bmatrix}.$$

niin

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \sqrt{x^2+y^2} \begin{bmatrix} \frac{x}{x^2+y^2} & -\frac{y}{\sqrt{x^2+y^2}} \\ \frac{y}{x^2+y^2} & \frac{x}{\sqrt{x^2+y^2}} \end{bmatrix} \begin{bmatrix} \Delta L \\ \Delta \varphi \end{bmatrix} = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} \Delta L - y \Delta \varphi \\ \frac{y}{\sqrt{x^2+y^2}} \Delta L + x \Delta \varphi \end{bmatrix}$$

Oletuksen mukaan $\Delta L = 0.02L = 0.02\sqrt{x^2+y^2}$ ja
 $\Delta \varphi = -0.01\varphi = -0.01 \arctan(\frac{y}{x})$ jolloin

$$\frac{\Delta x}{x} \approx 0.02 + 0.01 \cdot \frac{y}{x} \varphi,$$
$$\frac{\Delta y}{y} \approx 0.02 - 0.01 \cdot \frac{x}{y} \varphi.$$

😊 Esimerkki: Implisiittinen lineaarinen approksimointi, jatk.

Koska oletimme, että $\varphi = \frac{\pi}{4}$ niin $\frac{y}{x} = 1$ ja saamme

$$\frac{\Delta x}{x} \approx \frac{2}{100} + \frac{\pi}{400} \approx 0.028,$$
$$\frac{\Delta y}{y} \approx \frac{2}{100} - \frac{\pi}{400} \approx -0.012$$

Paperi on siis venynyt noin 2.8% x-akselin suunnassa ja kutistunut 1.2% y-akselin suunnassa.

💡 Esimerkki: Ääriarvo

Jos meidän pitää määrittää funktion $f(x, y) = -2x^4 - e^{4y} + 4x^2e^y$ paikalliset ääriarvot niin laskemme ensin funktion derivaatan joka on $f'(x, y) = [-8x^3 + 8xe^y, -4e^{4y} + 4x^2e^y]$. Tämä derivaatta eli gradientti on nollavektori kun

$$-8x^3 + 8xe^y = 0,$$

$$-4e^{4y} + 4x^2e^y = 0.$$

Jos $x = 0$ niin $-4e^{4y} + 4x^2e^y = -4e^{4y} < 0$ joten $x \neq 0$ ja ensimmäisestä yhtälöstä seuraa, että $e^y = x^2$ jolloin jälkimmäisestä seuraa, että $x^8 = x^4$ eli $x = \pm 1$ koska x on reaalinen ja $\neq 0$. Silloin $e^y = 1$ eli $y = 0$ ja kriittiset pisteet ovat $(\pm 1, 0)$.

Näiden kriittisten pisteiden luonteen selvittämiseksi meidän pitää laskea funktion toinen derivaatta ja se on

$$f''(x, y) = \begin{bmatrix} (-24x^2 + 8e^y) & 8xe^y \\ 8xe^y & (-16e^{4y} + 4x^2e^y) \end{bmatrix}.$$

💡 Esimerkki, jatk.

Näin ollen $f''(\pm 1, 0) = \begin{bmatrix} -16 & \pm 8 \\ \pm 8 & -12 \end{bmatrix}$. Näiden matriisien determinantti on $(-16) \cdot (-12) - 64 = 128 > 0$ joten ominaisarvot ovat samanmerkkiset ja koska lävistäjäalkiot ovat negatiiviset niin ominaisarvotkin ovat negatiiviset ja $(1, 0)$ ja $(-1, 0)$ ovat kaksi paikallista maksimipistettä. Yhden muuttujan derivoituvalla funktiolla ei voi olla täsmälleen kaksi kriittistä pistettä, jotka molemmat ovat maksimipisteitä.

- Funktio $f(x, y) = -2x^4 - e^{4y} + 4x^2e^y$ ei saavuta pienintä arvoa joukossa \mathbb{R}^2 koska $\lim_{x \rightarrow \infty} f(x, 0) = \lim_{x \rightarrow \infty} (-2x^4 - 1 + 4x^2) = -\infty$.
- Suurin arvo 1 saavutetaan pisteissä $(\pm 1, 0)$ ja tämän osoittamiseksi riittää osoittaa (mikä tässä nyt jää tekemättä), että $f(x, y) < \frac{1}{3}$ kun $(x, y) \notin \Omega = \{(x, y) : |x| < 3, |y| < \ln(3)\}$. Tämä seuraa siitä, että funktio saavuttaa suurimman arvonsa suljetussa joukossa $\Omega \cup \partial\Omega$ mutta jos pystymme osoittamaan että suurin arvo joukon Ω ulkopuolella (mukaanlukien reuna) on korkeintaan $\frac{1}{3}$ niin suurin arvo joukossa Ω ja samoin joukossa \mathbb{R}^2 saavutetaan jossain kriittisessä pisteessä, eli pisteissä $(\pm 1, 0)$ missä funktion arvo on 1.

😊 Esimerkki: Suurin ja pienin arvo

Jos meidän pitää määrittää funktion $f(x, y) = x^2 + y^2 - 4x - 2y + 7$ suurin ja pienin arvo joukossa $\Omega = \{ (x, y) : x \geq 0, y \geq 0, x + y \leq 5 \}$ niin laskemme ensin gradientin $\nabla f = (2x - 4)\mathbf{i} + (2y - 2)\mathbf{j}$ nollakohdat, ja näemme, että ainoa nollakohta on pisteessä $(2, 1)$, joka kuuluu kyseiseen joukkoon.

Seuraavaksi tutkimme ko. alueen reunaa joka koostuu kolmesta osasta:

$$A_1 = \{ (x, 0) : 0 \leq x \leq 5 \},$$

$$A_2 = \{ (0, y) : 0 \leq y \leq 5 \},$$

$$A_3 = \{ (x, y) : x + y = 5, 0 \leq x \leq 5 \}.$$

- Joukossa A_1 funktio on $g(x) = f(x, 0) = x^2 - 4x + 7$ ja koska $g'(2) = 0$ niin näemme että mahdolliset ääriarvopisteet ovat $(0, 0)$, $(2, 0)$ ja $(5, 0)$.

😊 Esimerkki: Suurin ja pienin arvo, jatk.

- Joukossa A_2 funktio on $h(y) = f(0, y) = y^2 - 2y + 7$ ja koska $h'(1) = 0$ niin näemme että mahdolliset ääriarvopisteet ovat $(0, 0)$, $(0, 1)$ ja $(0, 5)$.
- Joukossa A_3 funktio on $k(x) = f(x, 5 - x) = x^2 + (5 - x)^2 - 4x - 2(5 - x) + 7 = 2x^2 - 12x + 22$ ja koska $k'(3) = 0$ niin näemme että mahdolliset ääriarvopisteet ovat $(5, 0)$, $(3, 2)$ ja $(0, 5)$.

Koska

$$f(2, 1) = 2,$$

$$f(0, 0) = 7$$

$$f(2, 0) = 3$$

$$f(5, 0) = 12,$$

$$f(0, 1) = 6,$$

$$f(0, 5) = 22$$

$$f(3, 2) = 4,$$

niin päättelemme, että pienin arvo on $f(2, 1) = 2$ ja suurin $f(0, 5) = 22$.

😊 Esimerkki: Taylorin polynomi

Jos

$$f(x, y) = \frac{e^{x+y} - 1}{x + y} \ln(1 + x - y),$$

missä $\frac{e^0 - 1}{0} = 1$ ja oletamme, että $1 + x - y > 0$ niin voimme määrittää funktion f 2. asteen Taylorin polynomin pisteessä $(0, 0)$ käyttäen hyväksi Taylorin polynomin yksikäsitteisyyttä ja kehitelmiä

$$\frac{e^t - 1}{t} = 1 + \frac{1}{2}t + \frac{1}{6}t^2 + \frac{1}{24}t^3 + \dots \text{ ja } \ln(1 + t) = t - \frac{1}{2}t^2 + \frac{1}{3}t^3 + \dots$$

seuraavalla tavalla: Sijoitamme näihin kehitelmiin lausekkeet $x + y$ sekä $x - y$ ja otamme mukaan vain ne termit joista tulee korkeintaan astetta 2 olevia termejä. Näin saamme

$$\begin{aligned} f(x, y) &= \left(1 + \frac{1}{2}(x + y) + \frac{1}{6}(x + y)^2 + \dots\right) \left((x - y) - \frac{1}{2}(x - y)^2 + \dots\right) \\ &= x - y - \frac{1}{2}(x - y)^2 + \frac{1}{2}(x + y)(x - y) + \dots = x - y + xy - y^2 + \dots \end{aligned}$$

Näin ollen Taylorin polynomiksi tulee

$$x - y + xy - y^2.$$

😊 Taylorin 2. asteen polynomi ja ääriarvot

Oletamme, että $f : \mathbb{R}^d \rightarrow \mathbb{R}$ on kaksi kertaa jatkuvasti derivoituva ainakin pisteen \mathbf{x}_0 läheisyydessä. Merkitsemme $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$ ja määrittelemme $g(t) = f(\mathbf{x}_0 + t\mathbf{h})$. Silloin

$$\begin{aligned}g(1) &= g(0) + \int_0^1 g'(t) dt = g(0) - \int_0^1 (1-t)g'(t) dt + \int_0^1 (1-t)g''(t) dt \\ &= g(0) + g'(0) + \int_0^1 (1-t)g''(0) + \int_0^1 (1-t)(g''(t) - g''(0)) dt \\ &= g(0) + g'(0) + \frac{1}{2}g''(0) + \int_0^1 (1-t)(g''(t) - g''(0)) dt.\end{aligned}$$

Ketjusäännön nojalla $g'(t) = f'(\mathbf{x} + t\mathbf{h})\mathbf{h} = \mathbf{h}^\top f'(\mathbf{x} + t\mathbf{h})^\top$ josta seuraa, että $g''(t) = \mathbf{h}^\top f''(\mathbf{x} + t\mathbf{h})\mathbf{h}$ joten

$$f(\mathbf{x}) = f(\mathbf{x}_0) + f'(\mathbf{x})(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^\top f''(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \eta(\mathbf{x}),$$

missä

😊 Taylorin 2. asteen polynomi ja ääriarvot, jatk.

$$\eta(\mathbf{x}) = \int_0^1 (1-t)(\mathbf{x} - \mathbf{x}_0)^\top (f''(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) - f''(\mathbf{x}_0))(\mathbf{x} - \mathbf{x}_0) dt.$$

Nyt $\max_{t \in [0,1]} \|f''(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) - f''(\mathbf{x}_0)\| \rightarrow 0$ kun $\mathbf{x} \rightarrow \mathbf{x}_0$ koska oletamme, että f'' on jatkuva ja siitä seuraa, että $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\eta(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|^2} = 0$. Jos A on mikä tahansa $d \times d$ -matriisi, joka on symmetrinen ja reaalinen (kuten $f''(\mathbf{x}_0)$) niin

$$\lambda_{\min} \|\mathbf{h}\|^2 \leq \mathbf{h}^\top A \mathbf{h} \leq \lambda_{\max} \|\mathbf{h}\|^2.$$

Jos siis $f'(\mathbf{x}_0) = \mathbf{0}$ ja kaikki $f''(\mathbf{x}_0)$:n ominaisarvot ovat positiiviset, niin silloin pienin niistä λ_{\min} on myös positiivinen ja

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \left(\frac{1}{2} \lambda_{\min} + \frac{\eta(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|^2} \right) \|\mathbf{x} - \mathbf{x}_0\|^2 \geq f(\mathbf{x}_0),$$

kun $\|\mathbf{x} - \mathbf{x}_0\|$ on niin pieni, että $\frac{\eta(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|^2} > -\frac{1}{2} \lambda_{\min}$.

😊 Miksi $C = (A^T A)^{-1} A^T Y$ on neliösumman $C \mapsto \|AC - Y\|^2$ minimipiste?

Voimme kirjoittaa minimoitavan funktion muodossa

$$\begin{aligned} F(C) &= (AC - Y)^T (AC - Y) = C^T A^T AC - C^T A^T Y - Y^T AC + Y^T Y \\ &= C^T A^T AC - 2Y^T AC + Y^T Y, \end{aligned}$$

jolloin derivaatta on

$$F'(C) = C^T (A^T A + (A^T A)^T) - 2Y^T A.$$

Koska $(A^T A)^T = A^T A$ niin voimme kirjoittaa yhtälön $F'(C) = \mathbf{0}$ muodossa

$$C^T A^T A = Y^T A \Rightarrow A^T AC = A^T Y \Rightarrow C = (A^T A)^{-1} A^T Y.$$

Toisella tavalla: Jos määrittelemme $P = A(A^T A)^{-1} A^T$ niin $P^2 = P$ eli P on projektio ja $P^T A = PA = A$ josta seuraa, että jos kirjoitamme $C = \tilde{C} + (A^T A)^{-1} A^T Y$ niin $AC = A\tilde{C} + PY$ ja

$$\begin{aligned} F(C) &= \|A\tilde{C} - (Y - PY)\|^2 = \|A\tilde{C}\|^2 - 2(Y - PY)^T A\tilde{C} + \|PY - Y\|^2 \\ &= \|A\tilde{C}\|^2 - 2Y^T (A - P^T A)\tilde{C} + \|PY - Y\|^2 = \|A\tilde{C}\|^2 + \|PY - Y\|^2 \geq \|PY - Y\|^2. \end{aligned}$$

💡 Esimerkki: Linaarinen regressio

Meillä on seuraavat havainnot muuttujista x ja y :

x	96	110	103	127	60	54	43	36	20	11	22
y	76	74	76	87	66	59	63	60	55	52	46

Jos haluamme määrittää luvut a ja b siten, että $y \approx a + bx$ minimoimolla neliösummaa $\sum_{j=1}^{11} (a + bx_j - y_j)^2$ niin määrittelemme 11×2 matriisin A siten, että $A(j, 1) = 1$, $A(j, 2) = x_j$ ja 11×1 pystyvektorin Y siten, että $Y(j, 1) = y_j$. Silloin vektori $(A^T A)^{-1} A^T Y$ sisältää optimaaliset kertoimet a ja b .

Matlab/Octavessa lasku menee seuraavalla tavalla:

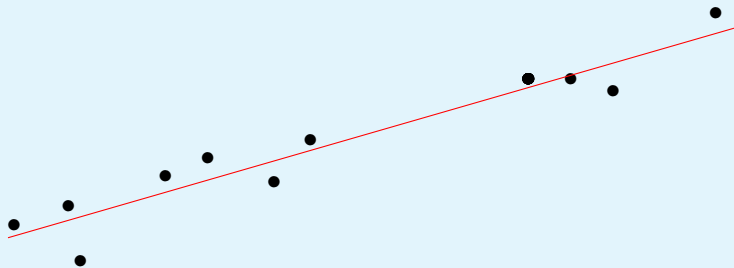
```
A=[1, 96; 1, 110; 1, 103; 1, 127; 1, 60; 1, 54; 1,43; 1,36;  
1,20; 1, 11; 1,22]
```

```
Y=[76; 74; 76; 87; 66; 59; 63; 60; 55; 52; 46]
```

```
C=inv(A'*A)*A'*Y jolloin saamme tulokseksi  $a = C(1) = 47.014$  ja
```

```
 $b = C(2) = 0.28863$ .
```

💡 Esimerkki: Linaarinen regressio, jatk.



Voimme myös laskea $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$ ja $\bar{y} = \frac{1}{n} \sum_{j=1}^n y_j$ jolloin

$b = \frac{\sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^n (x_j - \bar{x})^2}$ ja $a = \bar{y} - b\bar{x}$. Nämä laskut voimme suorittaa komennoilla

`x=[96, 110, 103, 127, 60, 54, 43, 36, 20, 11, 22];`

`y=[76, 74, 76, 87, 66, 59, 63, 60, 55, 52, 46];`

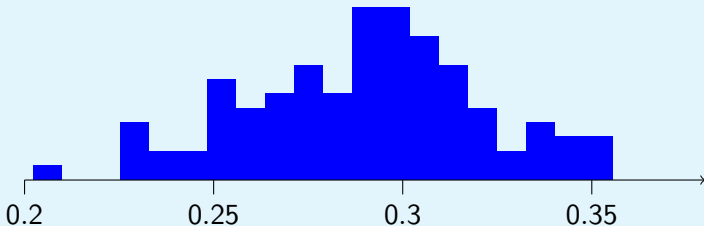
`b=cov(x,y)/var(x), a=mean(y)-b*mean(x)`

💡 Esimerkki: Linaarinen regressio, jatk.

Jos nyt haluaisimme arvioida miten luotettavia numeroarvot $a = 47.014$ ja $b = 0.28863$ ovat niin voimme joko käyttää perustilastotieteen standardioletuksia riippumattomuudesta ja normaalijakautuneisuudesta tai sitten menetellä seuraavalla tavalla: Valitsemme satunnaisesti (takaisinpanolla) 11 pistettä (x, y) joukosta $\{(x_j, y_j) : 1 \leq j \leq 11\}$ ja laskemme näillä arvoilla lineaariset regressiokertoimet a ja b ja tämän toistamme kunnes olemme saaneet riittävän monta arvoa. Laskut voimme tehdä näillä komennoilla (missä A on edellä määritelty matriisi):

```
for k=1:100, jj=floor(11*rand(11,1))+1;  
C(:,k)=(A(jj,:)'*A(jj,:))\A(jj,:)'*Y(jj);end
```

Silloin esimerkiksi b :n jakauma on seuraavanlainen:



😊 Esimerkki: Pienimmän neliösumman menetelmä

Meillä on pisteet (x_j, y_j) , $j = 1, \dots, n$,

$x = [0.24 \ 0.42 \ 0.63 \ 0.83 \ 1.28 \ 1.74 \ 2.12 \ 2.7 \ 2.89 \ 3.14]$

$y = [0.24 \ 0.41 \ 0.59 \ 0.74 \ 0.96 \ 0.99 \ 0.85 \ 0.43 \ 0.25 \ 0]$

oletamme, että $y \approx c_1 + c_2x + c_3x^2$ ja haluamme määrittää kertoimet c_1 , c_2 ja c_3 annettujen pisteiden perusteella.

Yhtä ainoata oikeata vastausta ei (tietenkään) ole olemassa joten on tärkeitä, että teemme selväksi miten määritämme kertoimet.

Yksinkertaisinta on valita c_j , $j = 1, 2, 3$ siten, että neliösumma

$$\sum_{j=1}^n (c_1 + c_2x_j + c_3x_j^2 - y_j)^2$$

on mahdollisimman pieni. (Tässä tapauksessa saisimme melkein saman tuloksen jos neliön paikalla olisi itseisarvo.) Voimme derivoida lausekkeen muuttujien c_1 , c_2 ja c_3 suhteen, ja hakea piste missä derivaatat ovat 0.

Tehokas tapa laskujen suorittamiseksi on kirjoittaa yhtälösystemi

$c_1 + c_2x_j + c_3x_j^2 - y_j \approx 0$ muodossa $AC - Y \approx 0$ missä $A(j, k) = x_j^{k-1}$, $C(k, 1) = c_k$ ja $Y(j, 1) = y_j$.

😊 Esimerkki: Pienimmän neliösumman menetelmä, jatk.

Funktion $C \mapsto \|AC - Y\|^2 = C^T A^T A C - 2Y^T A C + Y^T Y$ derivaatta on $2C^T A^T A - 2Y^T A$ ja tämä on $\mathbf{0}$ kun $C = (A^T A)^{-1} A^T Y$.

Matlab/Octavessa voimme kirjoittaa tämän seuraavalla tavalla

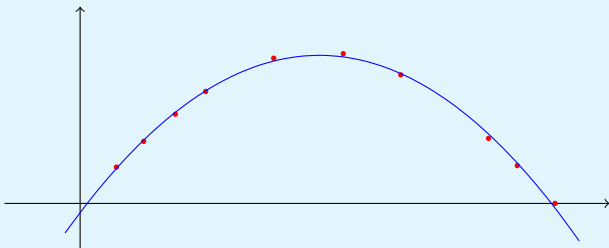
$A = [(1+0*x'), x', x'.^2]$; $Y = y'$;

ja sitten laskemme

$C = (A' * A) \backslash A' * Y$

Tästä saamme $C = \begin{bmatrix} -0.060453 \\ 1.314680 \\ -0.415687 \end{bmatrix}$. Pisteet ja käyrä näyttävät

seurvaavanlaisilta:



💡 Esimerkki: Lagrangen kerroin

Jos haluamme määrittää lyhimmän etäisyyden jostain paraabelin $y = x^2$ pisteestä pisteeseen $(3, 0)$ niin voimme käyttää Lagrangen kerrointa seuraavalla tavalla:

Minimoimme etäisyyden neliötä eli funktiota $f(x, y) = (x - 3)^2 + y^2$ kun $g(x, y) = x^2 - y = 0$ ja muodostamme Lagrangen funktion

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

Ehto, että F :n gradientti on 0 antaa yhtälösystemin

$$F_x(x, y, \lambda) = 2(x - 3) + 2\lambda x = 0,$$

$$F_y(x, y, \lambda) = 2y - \lambda = 0,$$

$$F_\lambda(x, y, \lambda) = x^2 - y = 0.$$

💡 Esimerkki: Lagrangen kerroin

Kolmannen yhtälön (eli rajoitusehdon) nojalla $y = x^2$ ja kun sijoitamme tämän toiseen yhtälöön saamme $\lambda = 2x^2$ jolloin ensimmäisen yhtälön nojalla näemme, että $x + 2x^3 = 3$. Nyt huomaamme, että $x = 1$ on tämän yhtälön ratkaisu ja muita ei ole koska funktio $x \mapsto x + 2x^3$ on aidosti kasvava. Silloin $y = 1$ ja $f(1, 1) = 5$.

Funktio f on jatkuva ja $\lim_{x^2+y^2 \rightarrow \infty} f(x, y) = \infty$ joten funktion f minimiarvo saavutetaan jossain pisteessä. Koska funktiot f ja g ovat jatkuvasti derivoituvia ja g :n derivaatta $2x\mathbf{i} - \mathbf{j} \neq \mathbf{0}$ niin minimipiste saavutetaan pistessä (joka yhdessä λ :n kanssa) on Lagrangen funktion kriittinen piste. Näin ollen voimme olla varmoja siitä, että $(1, 1)$ todella on minimipiste ja etäisyys on $\sqrt{5}$.

💡 Huom!

Edellisessä esimerkissä Lagrangen kertoimen käyttö ei oikeastaan tarjonnut mitään suurta etua verrattuna muihin mahdollisuuksiin mutta Lagrangen menetelmän vahvuudet tulevat esille kun ratkaistava ongelma on vaikeampi.

😊 Mitä Lagrangen kerroin "kertoo"?

Oleta, että funktion $f(\mathbf{x})$ suurin (tai pienin) arvo kun $g(\mathbf{x}) + c = 0$ saavutetaan pisteessä $\mathbf{x}(c)$ ja että tämä piste (yhdessä $\lambda(c)$:n kanssa) on Lagrangen funktion $L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda(g(\mathbf{x}) + c)$ derivaatan nollakohta eli

$$\begin{aligned}f'(\mathbf{x}(c)) + \lambda(c)g'(\mathbf{x}(c)) &= \mathbf{0}, \\g(\mathbf{x}(c)) + c &= 0.\end{aligned}$$

Jos nyt $h(c) = f(\mathbf{x}(c))$ on maksimiarvo (tai minimiarvo) niin ensimmäisestä yhtälöstä seuraa, että

$$h'(c) = f'(\mathbf{x}(c))\mathbf{x}'(c) = -\lambda(c)g'(\mathbf{x}(c))\mathbf{x}'(c).$$

Koska $g(\mathbf{x}(c)) = -c$ niin pätee $g'(\mathbf{x}(c))\mathbf{x}'(c) = -1$ ja silloin myös

$$h'(c) = \lambda(c).$$

Lagrangen kerroin siis kertoo miten maksimi- tai minimiarvo riippuu muutoksista rajoitusehdossa.

😊 Miksi Lagrangen menetelmä toimii?

Oleta, että f ja $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ ovat jatkuvasti derivoituvia, $g(x_0, y_0, z_0) = 0$ ja $f(x, y, z) \geq f(x_0, y_0, z_0)$ kaikilla $(x, y, z) \in \mathbb{R}^3$ joilla $g(x, y, z) = 0$ ja että $g'(x_0, y_0, z_0) \neq \mathbf{0}$. Voimmeko tästä päätellä, että on olemassa luku λ_0 siten, että $(x_0, y_0, z_0, \lambda_0)$ on funktion $L(x, y, z, \lambda) = f(x, y, z) + \lambda g(x, y, z)$ kriittinen piste eli derivaatan nollakohta?

Koska $g'(x_0, y_0, z_0) \neq \mathbf{0}$ niin ainakin yksi komponentti ei ole 0 ja oletamme, että esimerkiksi $g_z(x_0, y_0, z_0) \neq 0$. Implisiittifunktiolauseen nojalla tiedämme, että on olemassa funktio $h(x, y)$ siten, että $h(x_0, y_0) = z_0$ ja $g(x, y, h(x, y)) = 0$ kun $\|(x, y) - (x_0, y_0)\|$ on riittävän pieni. Mutta silloin $f(x, y, h(x, y)) \geq f(x_0, y_0, h(x_0, y_0))$ ja funktion $(x, y) \mapsto f(x, y, h(x, y))$ derivaatta pisteessä (x_0, y_0) on $\mathbf{0}$, eli

$$f_x(x_0, y_0, h(x_0, y_0)) + f_z(x_0, y_0, h(x_0, y_0))h_x(x_0, y_0) = 0,$$

$$f_y(x_0, y_0, h(x_0, y_0)) + f_z(x_0, y_0, h(x_0, y_0))h_y(x_0, y_0) = 0.$$

😊 Miksi Lagrangen menetelmä toimii? jatk.

Derivoimalla yhtälön $g(x, y, h(x, y)) = 0$ molemmat puolet saamme

$$g_x(x_0, y_0, h(x_0, y_0)) + g_z(x_0, y_0, h(x_0, y_0))h_x(x_0, y_0) = 0,$$

$$g_y(x_0, y_0, h(x_0, y_0)) + g_z(x_0, y_0, h(x_0, y_0))h_y(x_0, y_0) = 0.$$

josta seuraa, että kun ratkaisemme $h_x(x_0, y_0)$ ja $h_y(x_0, y_0)$ näistä yhtälöistä ja muistamme, että $z_0 = h(x_0, y_0)$ niin saamme

$$f_x(x_0, y_0, z_0) - \frac{f_z(x_0, y_0, z_0)}{g_z(x_0, y_0, z_0)}g_x(x_0, y_0, z_0) = 0,$$

$$f_y(x_0, y_0, z_0) - \frac{f_z(x_0, y_0, z_0)}{g_z(x_0, y_0, z_0)}g_y(x_0, y_0, z_0) = 0.$$

Jos nyt valitsemme $\lambda_0 = -\frac{f_z(x_0, y_0, z_0)}{g_z(x_0, y_0, z_0)}$ niin pätee $L'(x_0, y_0, z_0, \lambda_0) = 0$.

😊 Esimerkki: Regressiosuora ja Lagrangen kerroin

Meillä on pisteet (x_j, y_j) , $j = 1, \dots, n$ ja haluamme määrittää suoran $ax + by + c = 0$ siten, että etäisyyksien neliöiden summa pisteistä (x_j, y_j) suoraan $ax + by + c = 0$ on mahdollisimman pieni. Näin ollen meidän pitää minimoida lauseketta

$$F(a, b, c) = \sum_{j=1}^n \frac{(ax_j + by_j + c)^2}{a^2 + b^2}.$$

Derivoimalla tätä funktiota muuttujan c suhteen saamme

$F_c(a, b, c) = \frac{2}{a^2 + b^2} \sum_{j=1}^n (ax_j + by_j + c)$ ja ehdosta $F_c = 0$ seuraa $a\bar{x} + b\bar{y} + c = 0$ eli suora kulkee pisteen $(\bar{x}, \bar{y}) = (\frac{1}{n} \sum_{j=1}^n x_j, \frac{1}{n} \sum_{j=1}^n y_j)$ kautta. Näin ollen voimme, kun korvaamme pisteet (x_j, y_j) pisteillä $(x_j - \bar{x}, y_j - \bar{y})$ olettaa, että $\bar{x} = \bar{y} = 0$ jolloin $c = 0$ ja suoran yhtälö on $ax + by = 0$.

Koska meidän täytyy vaatia, että $(a, b) \neq (0, 0)$ voimme yhtä hyvin vaatia, että $a^2 + b^2 = 1$ jolloin meidän täytyy löytää funktion $(a, b) \mapsto \sum_{j=1}^n (ax_j + by_j)^2$ pienin arvo kun $a^2 + b^2 = 1$.

😊 Esimerkki: Regressiosuora ja Lagrangen kerroin, jatk.

Jos määrittelemme $X = \begin{bmatrix} x \\ y \end{bmatrix}$ ja $M = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{bmatrix}$ niin

$$\sum_{j=1}^n (ax_j + by_j)^2 = \|M^T X\|^2 = (M^T X)^T (M^T X) = X^T M M^T X,$$

ja rajoitusehto on $X^T X = 1$.

Tästä syystä muodostamme Lagrangen funktion

$$L(X, \lambda) = X^T M M^T X + \lambda(X^T X - 1).$$

Haemme tämän funktion derivaatan nollakohdat ja se on seuraavan yhtälösystemin ratkaisu:

$$L_X = X^T (M M^T + (M M^T)^T) + 2\lambda X^T = 0,$$

$$L_\lambda = X^T X - 1 = 0.$$

😊 Esimerkki: Regressiosuora ja Lagrangen kerroin, jatk.

Koska $(MM^T)^T = MM^T$ niin saamme transponoimalla

$$MM^T X = -\lambda X,$$

eli $-\lambda$ on matriisin MM^T ominaisarvo ja X on vastaava ominaisvektori.
Minimoitavan funktion arvo on

$$X^T MM^T X = -\lambda X^T X = -\lambda,$$

joten meidän pitää valita vektoriksi X matriisin MM^T pienempään ominaisarvoon liittyvä ominaisvektori.

Jos laskemme matriisin M singulaarihajotelman $M = USV^T$ niin matriisin U toinen sarake $U(:, 2)$ on hakemamme ratkaisu $\begin{bmatrix} a \\ b \end{bmatrix}$ ja suoran suuntavektori on ensimmäinen sarake $U(:, 1)$.

Jos meillä on vaakavektorit x ja y niin voimme suorittaa nämä laskut Matlab/Octavessa komennoilla:

$M = [x - \text{mean}(x); y - \text{mean}(y)]$; $[u, s, v] = \text{svd}(M)$;
jolloin $a = u(1, 2)$, $b = u(2, 2)$ ja $c = -a * \text{mean}(x) - b * \text{mean}(y)$

😊 Esimerkki: Boltzmannin jakauma

Kaasussa on N molekyyliä ja jokainen niistä voi olla tilassa j jolloin sen energia on E_j missä $j = 1, \dots, M$, ($1 \leq M \leq \infty$). Jos molekyyleistä N_j ovat tilassa j niin $\sum_{j=1}^M N_j = N$ ja jos niiden yhteenlaskettu energia on E , niin pätee $\sum_{j=1}^M N_j E_j = E$.

On olemassa

$$\frac{N!}{N_1! N_2! \dots N_M!},$$

erilaista tapaa millä N molekyyliä voivat sijoittua eri energialuokkiin siten, että luokassa j on N_j molekyyliä. Nyt meidän pitää määrittää osuudet $\frac{N_j}{N}$ siten, että tämä luku, tai yhtäpitävästi sen logaritmi, on suurimmillaan kun $\sum_{j=1}^M N_j = N$ ja $\sum_{j=1}^M N_j E_j = E$.

Niin sanotun Stirlingin kaavan mukaan

$\ln(k!) \approx (k + \frac{1}{2}) \ln(k) - k + \frac{1}{2} \ln(2\pi)$ joten me haemme lausekkeen $(N + \frac{1}{2}) \ln(N) - N - \sum_{j=1}^M (N_j + \frac{1}{2}) \ln(N_j) + \sum_{j=1}^M N_j$ suurinta arvoa emmekä välitä vaatimuksesta, että lukujen N_j pitäisi olla kokonaislukuja.

😊 Esimerkki: Boltzmannin jakauma, jatk.

Lagrangen funktioksi tulee

$$L(N_1, \dots, N_M, \lambda, \mu) = (N + \frac{1}{2}) \ln(N) - \sum_{j=1}^M (N_j + \frac{1}{2}) \ln(N_j) \\ + \lambda \left(\sum_{j=1}^M N_j - N \right) + \mu \left(\sum_{j=1}^M N_j E_j - E \right).$$

Derivaatan nollakohdat toteuttavat yhtälösystemin

$$\frac{\partial L}{\partial N_j} = -\ln(N_j) - 1 - \frac{1}{2N_j} + \lambda + \mu E_j = 0, \quad j = 1, \dots, M,$$

$$\frac{\partial L}{\partial \lambda} = \sum_{j=1}^M N_j - N = 0,$$

$$\frac{\partial L}{\partial \mu} = \sum_{j=1}^M N_j E_j - E = 0.$$

😊 Esimerkki: Boltzmannin jakauma, jatk.

Ensimmäisistä yhtälöistä saamme ratkaisuksi

$$N_j = e^{-1 - \frac{1}{2N_j} + \lambda} e^{\mu E_j} \approx e^{\lambda-1} e^{\mu E_j}$$

ja laskemalla yhteen saamme

$$N \approx e^{\lambda-1} \sum_{j=1}^M e^{\mu E_j}, \quad j = 1, \dots, M.$$

Näin ollen

$$\frac{N_j}{N} \approx \frac{e^{\mu E_j}}{\sum_{j=1}^M e^{\mu E_j}},$$

ja μ määräytyy ehdosta

$$\frac{E}{N} \approx \frac{\sum_{j=1}^M E_j e^{\mu E_j}}{\sum_{j=1}^M e^{\mu E_j}}.$$

Tavallisesti kirjoitetaan $\mu = -\frac{1}{kT}$ missä k on Boltzmannin vakio.

😊 Gradientti- ja konjugaatti-gradientti-menetelmä

Gradientti-menetelmä funktion f minimin löytämiseksi toimii seuraavalla tavalla: Aloitetaan jostain pisteestä \mathbf{x}_0 ja kun algoritmi on tullut pisteeseen \mathbf{x}_n niin valitaan suunta $\mathbf{s}_n = -f'(\mathbf{x}_n)^T$ (kun oletetaan, että $f'(\mathbf{x}_n)$ on rivivektori jolloin \mathbf{s}_n on pystyvektori kuten \mathbf{x}_n) ja seuraavaksi pisteeksi valitaan $\mathbf{x}_{n+1} = \mathbf{x}_n + t_n \mathbf{s}_n$ missä t_n määräytyy ehdosta $f(\mathbf{x}_n + t_n \mathbf{s}_n) = \min_{t \in \mathbb{R}} f(\mathbf{x}_n + t \mathbf{s}_n)$ (tai ainakin niin että löydetään paikallinen minimipiste).

Konjugaatti-gradientti-menetelmä toimii periaatteessa samalla tavalla mutta ainostaan kun $n = 0$ valitaan $\mathbf{s}_0 = -f'(\mathbf{x}_0)^T$ ja muuten kun on löydetty piste \mathbf{x}_{n+1} niin valitaan uudeksi suunnaksi

$$\mathbf{s}_{n+1} = -f'(\mathbf{x}_{n+1})^T + \frac{\|f'(\mathbf{x}_{n+1})\|^2}{\|f'(\mathbf{x}_n)\|^2} \mathbf{s}_n.$$

Seuraavaksi tutkimme miten nämä menetelmät toimivat kun

$$f(x, y) = x^2 + 1.96xy + y^2 = 0.02x^2 + 0.02y^2 + 0.98(x + y)^2 \text{ eli}$$

$$f(X) = X^T B X \text{ missä } B = \begin{bmatrix} 1 & 0.98 \\ 0.98 & 1 \end{bmatrix} \text{ ja } X = \begin{bmatrix} x \\ y \end{bmatrix}.$$

😊 Gradientti- ja konjugaatti-gradientti-menetelmä, jatk.

Funktion $t \mapsto f(X + tS)$ derivaatta on $2S^T BX + 2tS^T BS$ joten tämä funktio saavuttaa pienimmän arvonsa pisteessä

$$X - \frac{S^T BX}{S^T BS} S.$$

Käyttämällä tätä tulosta hyväksi saamme gradienttimenetelmällä esimerkiksi seuraavat pisteet:

$$\begin{aligned} X_0 &= \begin{bmatrix} -0.98000 \\ 1.00000 \end{bmatrix}, & X_1 &= \begin{bmatrix} -0.98000 \\ 0.96040 \end{bmatrix}, & X_2 &= \begin{bmatrix} -0.94119 \\ 0.96040 \end{bmatrix}, \\ X_3 &= \begin{bmatrix} -0.94119 \\ 0.92237 \end{bmatrix}, & X_4 &= \begin{bmatrix} -0.90392 \\ 0.92237 \end{bmatrix}, & X_5 &= \begin{bmatrix} -0.90392 \\ 0.88584 \end{bmatrix} \end{aligned}$$

Syy siihen, ettei gradienttimenetelmä toimi tätä paremmin, on että matriisin B ominaisarvot $\lambda_1 = 0.02$ ja $\lambda_2 = 1.98$ ovat niin erisuuret ja tässä esimerkissä alkuarvo X_0 oli valittu erityisen epäedullisella tavalla.

😊 Gradientti- ja konjugaatti-gradientti-menetelmä, jatk.

Konjugaatti-gradientti-menetelmällä saamme samalla alkuarvolla

$$X_0 = \begin{bmatrix} -0.98 \\ 1 \end{bmatrix} \text{ ensin saman pisteen } X_1 = \begin{bmatrix} -0.98000 \\ 0.96040 \end{bmatrix} \text{ kuin}$$

gradienttimenetelmällä mutta uusi suunta on (koska $f'(X)^T = -2BX$)

$$S_1 = -2BX_1 + \frac{X_1^T B B X_1}{X_0^T B B X_0} (-2BX_0) = \begin{bmatrix} 0.077616 \\ -0.076064 \end{bmatrix},$$

jolloin saamme

$$X_2 = X_1 - \frac{S_1^T B X_1}{S_1^T B S_1} S_1 = \begin{bmatrix} 0.0000e + 00 \\ -4.7740e - 15 \end{bmatrix},$$

ja tarkka lasku olisi antanut vastaukseksi origon.

Voidaan osoittaa, että jos $f(X) = X^T B X + C X$ missä B on $d \times d$ -matriisi joka on reaalinen, symmetrinen ja jonka ominaisarvot ovat positiiviset ja C on $1 \times d$ -rivivektori niin konjugaatti-gradientti-menetelmä tarvitsee (korkeintaan, jos lasketaan tarkasti) d askelta minimipisteen löytämiseksi. Käytännössä minimoitavat funktiot eivät tietenkään ole tätä muotoa mutta minimipisteen läheisyydessä ne voivat olla melkein tällaisia.

😊 Esimerkki: Lineaarinen optimointi

Jos meidän pitää ratkaista seuraava lineaarinen optimointiprobleema

$$\max 5x_1 + 3x_2 + 2x_3$$

$$x_1, x_2, x_3 \geq 0$$

$$4x_1 + 5x_2 + 2x_3 \leq 20$$

$$3x_1 + 4x_2 - x_3 \leq 30$$

niin voimme menetellä seuraavalla tavalla: Ensin lisäämme ylijäämämuuttujat s_1 ja s_2 rajoitusehtoihin eli kirjoitamme

$4x_1 + 5x_2 + 2x_3 + s_1 = 20$ ja $3x_1 + 4x_2 - x_3 + s_2 = 30$ jolloin epäyhtälöt muuttuvat ehdoiksi $s_1 \geq 0$ ja $s_2 \geq 0$. Maksimoitavan funktion kirjoitamme yhtälönä muodossa $q - 5x_1 - 3x_2 - 2x_3 = 0$. Nyt voimme kirjoittaa nämä yhtälöt seuraavana taulukkona

q	x_1	x_2	x_3	s_1	s_2	$=$
0	4	5	2	1	0	20
0	3	4	-1	0	1	30
1	-5	-3	-2	0	0	0

😊 Esimerkki: Lineaarinen optimointi

Yhtälösystemin ratkaisuksi otamme nyt $s_1 = 20$, $s_2 = 40$ ja $q = 0$ ja silloin $x_1 = x_2 = x_3 = 0$ koska niiden sarakkeissa ei ole tukialkoita **1**.

Kantamuuttujat ovat siis nyt $\{s_1, s_2\}$ (jos q :ta ei lasketa sellaiseksi).

Seuraavassa vaiheessa valitsemme uudeksi kantamuuttujaksi x_1 :n koska sen kerroin viimeisellä rivillä on negatiivisistä kertoimista itseisarvoltaan suurin. Kantamuuttujien joukosta poistamme muuttujan s_1 jonka tukialkio on rivillä 1 koska $\frac{20}{4} < \frac{30}{3}$. (Jos poistaisimme muuttujan s_2 niin s_1 saisi negatiivisen arvon.) Gaussin eliminaatiomenetelmän rivioperaatioiden avulla saamme nyt yhtälösystemiksi

q	x_1	x_2	x_3	s_1	s_2	=
0	1	1.25	0.5	0.25	0	5
0	0	0.25	-2.5	-0.75	1	15
1	0	3.25	0.5	1.25	0	25

Nyt kantamuuttujat ovat $\{x_1, s_2\}$ ja optimaalinen ratkaisu on $x_1 = 5$, $x_2 = x_3 = 0$ koska kolmannella rivillä ei ole yhtään negatiivista lukua jolloin emme pysty kasvattamaan q :n arvoa vaihtamalla kantamuuttujia.

😊 Esimerkki: Integroimisjärjestystä ei voi aina vaihtaa

Funktio $(x, y) \mapsto \frac{x^2 - y^2}{(x^2 + y^2)^2}$ on origoa lukuunottamatta jatkuva kaikissa pisteissä (ja integraalien kannalta on yhdentekevää miten se määritellään origossa). Erityisesti pätee silloin, että integraali $\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx$ on kaikin puolin hyvin määritelty kun $y > 0$.

Tämän integraalin laskemiseksi toteamme, että funktio $t \mapsto \int_0^t \frac{x^2 - y^2}{(x^2 + y^2)^2} dx$ saa arvon 0 origossa ja sen derivaatta on $\frac{t^2 - y^2}{(t^2 + y^2)^2}$. Mutta myös funktio $t \mapsto -\frac{t}{t^2 + y^2}$ saa arvon 0 kun $t = 0$ ja sen derivaatta on

$$-\frac{(t^2 + y^2) - t \cdot 2t}{(t^2 + y^2)^2} = \frac{t^2 - y^2}{(t^2 + y^2)^2}.$$

Näin ollen $\int_0^t \frac{x^2 - y^2}{(x^2 + y^2)^2} dx = -\frac{t}{t^2 + y^2}$ kaikilla t ja erityisesti

$$\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx = -\frac{1}{1 + y^2}.$$

😊 Esimerkki: Integroimisjärjestystä ei voi aina vaihtaa, jatk.

Funktio $y \mapsto -\frac{1}{1+y^2}$ on jatkuva ja koska $\frac{d}{dy} \arctan(y) = \frac{1}{1+y^2}$ niin

$$\int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx \right) dy = - \int_0^1 \frac{1}{1+y^2} dy = - \int_0^1 \arctan(y) = -\frac{\pi}{4}.$$

Koska $\frac{x^2 - y^2}{(x^2 + y^2)^2} = -\frac{y^2 - x^2}{(y^2 + x^2)^2}$ niin sama päättely kuin edellä osoittaa, että

$$\int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dy \right) dx = \frac{\pi}{4} \neq -\frac{\pi}{4} = \int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx \right) dy.$$

Fubinin lauseen nojalla voimme silloin myös vetää johtopäätöksen, että

$$\begin{aligned} \iint_{[0,1] \times [0,1]} \left| \frac{x^2 - y^2}{(x^2 + y^2)^2} \right| dA &= \int_0^1 \left(\int_0^1 \left| \frac{x^2 - y^2}{(x^2 + y^2)^2} \right| dx \right) dy \\ &= \int_0^1 \left(\int_0^1 \left| \frac{x^2 - y^2}{(x^2 + y^2)^2} \right| dy \right) dx = \infty. \end{aligned}$$

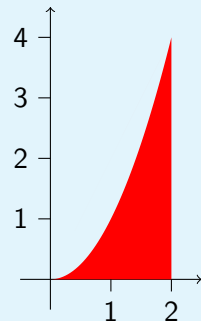
💡 Esimerkki: Integroimisjärjestyksen vaihto

Jos haluamme vaihtaa integroimisjärjestyksen integraaleissa

$$(a) \int_0^4 \int_{\sqrt{y}}^2 y \, dx \, dy, \quad (b) \int_0^4 \int_{\sqrt{y}}^4 y \, dx \, dy,$$

niin voimme menetellä seuraavalla tavalla:

(a) Integroimisalue näyttää seuraavanlaiselta:

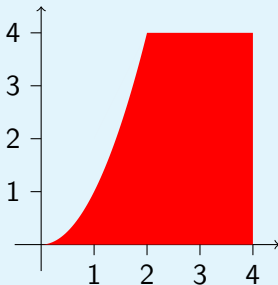


💡 Esimerkki: Integroimisjärjestyksen vaihto, jatk.

Koska $x = \sqrt{y}$ kun $y = x^2$ ja $x \geq 0$ niin voimme kuvion perusteella päätellä, että $0 \leq x \leq 2$ ja $0 \leq y \leq x^2$. Tästä seuraa, että

$$\begin{aligned} \int_0^4 \int_{\sqrt{y}}^2 y \, dx \, dy &= \int_0^4 \left(\int_{\sqrt{y}}^2 y \, dx \right) dy = \int_0^2 \left(\int_0^{x^2} y \, dy \right) dx \\ &= \int_0^2 \left(\int_0^{x^2} \frac{1}{2} y^2 \right) dx = \int_0^2 \frac{1}{2} x^4 \, dx = \int_0^2 \frac{1}{10} x^5 = \frac{16}{5}. \end{aligned}$$

(b) Integroimisalue näyttää seuraavanlaiselta:



💡 **Esimerkki: Integroimisjärjestyksen vaihto, jatk.**

Tässä tapauksessa $0 \leq x \leq 4$, $0 \leq y \leq x^2$ ja $y \leq 4$ joten kun $0 \leq x \leq 2$ niin pätee $0 \leq y \leq x^2$ ja kun $2 \leq x \leq 4$ niin pätee $0 \leq y \leq 4$. Näin ollen saamme (a)-kohdan nojalla

$$\begin{aligned} \int_0^4 \int_{\sqrt{y}}^4 y \, dx \, dy &= \int_0^2 \left(\int_0^{x^2} y \, dy \right) dx + \int_2^4 \left(\int_0^4 y \, dy \right) dx \\ &= \frac{16}{5} + \int_2^4 \left(\int_0^4 \frac{1}{2} y^2 \right) dx = \frac{16}{5} + \int_2^4 8 \, dx = \frac{16}{5} + 8(4 - 2) = \frac{96}{5}. \end{aligned}$$

Matlab/Octavessa voimme numeerisesti laskea tällaiset integraalit määrittelemällä integroitavaa funktiota komennolla

```
f=@(x,y) y.*(y< x.^2)
```

jolloin siis $f(x, y) = 0$ kun $y \geq x^2$ ja sitten (a)-tapauksessa laskea

```
dblquad(f,0,2,0,4)
```

ja (b)-tapauksessa

```
dblquad(f,0,4,0,4).
```

😊 Integroimisjärjestyksen vaihto

Jos haluamme vaihtaa integroimisjärjestystä integraalissa

$\int_0^1 \int_0^{\frac{1}{y^2}} \frac{yx}{1+x^2} dx dy$ niin meidän pitää ensin todeta, että

integroimisalueessa pätee $0 < y < 1$ ja $0 < x < \frac{1}{y^2}$ jolloin $0 < x < \infty$ ja

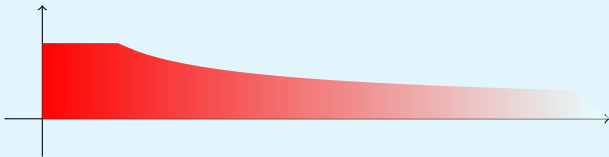
$y^2 < \frac{1}{x}$ josta seuraa, että $0 < y < \min(\frac{1}{\sqrt{x}}, 1)$. Tästä seuraa, että

$0 < y < 1$ kun $0 < x < 1$ ja $0 < y < \frac{1}{\sqrt{x}}$ kun $1 \leq x < \infty$. Näin ollen

saamme

$$\int_0^1 \int_0^{\frac{1}{y^2}} \frac{xy}{1+x^2} dx dy = \int_0^1 \int_0^1 \frac{xy}{1+x^2} dy dx + \int_1^{\infty} \int_0^{\frac{1}{\sqrt{x}}} \frac{xy}{1+x^2} dy dx.$$

Integroimisalue näyttää suunnilleen seuraavanlaiselta:



😊 Integroimisjärjestyksen vaihto

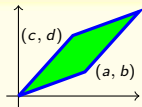
Jos lisäksi haluamme laskea integraalin, niin saamme

$$\begin{aligned} & \int_0^1 \int_0^1 \frac{xy}{1+x^2} dy dx + \int_1^\infty \int_0^{\frac{1}{\sqrt{x}}} \frac{xy}{1+x^2} dy dx \\ &= \int_0^1 \int_0^1 \frac{1}{2} \frac{xy^2}{1+x^2} dx + \int_1^\infty \int_0^{\frac{1}{\sqrt{x}}} \frac{1}{2} \frac{xy^2}{1+x^2} dx = \int_0^1 \frac{1}{2} \frac{x}{1+x^2} dx \\ &+ \int_1^\infty \frac{1}{2} \frac{1}{1+x^2} dx = \int_0^1 \frac{1}{4} \ln(1+x^2) + \int_1^\infty \frac{1}{2} \arctan(x) \\ &= \frac{1}{4} (\ln(2) - \ln(1)) + \frac{1}{2} \left(\lim_{x \rightarrow \infty} \arctan(x) - \arctan(1) \right) \\ &= \frac{1}{4} \ln(2) + \frac{1}{2} \left(\frac{\pi}{2} - \frac{\pi}{4} \right) = \frac{1}{4} \ln(2) + \frac{\pi}{8} \approx 0,5659858770. \end{aligned}$$

😊 Muuttujien vaihto

Oletamme, että D on suunnikas, jonka kulmapisteet ovat $(0, 0)$, (a, b) , (c, d) ja $(a + c, b + d)$.

Jos $x = as + ct$ ja $y = bs + dt$ niin



$$x\mathbf{i} + y\mathbf{j} = s(a\mathbf{i} + b\mathbf{j}) + t(c\mathbf{i} + d\mathbf{j}) \quad \text{eli} \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} = s \begin{bmatrix} a \\ b \end{bmatrix} + t \begin{bmatrix} c \\ d \end{bmatrix},$$

ja tämä on suunnikkaan sivujen $a\mathbf{i} + b\mathbf{j} = \begin{bmatrix} a \\ b \end{bmatrix}$ ja $c\mathbf{i} + d\mathbf{j} = \begin{bmatrix} c \\ d \end{bmatrix}$

linearikombinaatio. Tällä tavalla saamme kaikki suunnikkaan pisteet kun s ja $t \in [0, 1]$ eli $(x, y) \in D$ missä $x = as + ct$ ja $y = bs + dt$ jos ja vain jos $(s, t) \in D_* = [0, 1] \times [0, 1]$.

Suunnikkaan pinta-ala on

$$\|(a\mathbf{i} + b\mathbf{j}) \times (c\mathbf{i} + d\mathbf{j})\| = \left| \det \left(\begin{bmatrix} a & c \\ b & d \end{bmatrix} \right) \right| = |ad - bc| \text{ ja neliön } D_*$$

pinta-ala on tietenkin 1.

😊 Muuttujien vaihto, jatk.

Jos nyt määrittelemme

$$\frac{\partial(x, y)}{\partial(s, t)} = \det \left(\begin{bmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{bmatrix} \right) = \det \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = ad - bc,$$

niin näemme, että jos f on vakio niin

$$\iint_D f dx dy = \iint_{D_*} f \left| \frac{\partial(x, y)}{\partial(s, t)} \right| ds dt.$$

Jos teemme (mielivaltaisen) muuttujien vaihdon $x = x(s, t)$, $y = y(s, t)$ niin saamme lineaarisella approksimoinnilla

$$x \approx x(s_0, t_0) + x_s(t_0, s_0)(s - s_0) + x_t(s_0, t_0)(t - t_0),$$

$$y \approx y(s_0, t_0) + y_s(t_0, s_0)(s - s_0) + y_t(s_0, t_0)(t - t_0),$$

josta seuraa, että kun $(s, t) \in [s_0, s_0 + \Delta s] \times [t_0, t_0 + \Delta t]$ niin saamme alueen, joka on melkein suunnikas. Sitten meidän pitää yleisessä tapauksessa laskea yhteen integraalit "suunnikkaiden" yli ja tarkistaa, että virheiden summa pysyy mielivaltaisen pienenä.

💡 Esimerkki: Muuttujien vaihto ja $\int_0^\infty e^{-x^2} dx$

Olkoon $E = \int_0^\infty e^{-x^2} dx$. Silloin

$$\begin{aligned} E^2 &= \left(\int_0^\infty e^{-x^2} dx \right) \left(\int_0^\infty e^{-x^2} dx \right) = \left(\int_0^\infty e^{-x^2} dx \right) \left(\int_0^\infty e^{-y^2} dy \right) \\ &= \int_0^\infty \int_0^\infty e^{-x^2} e^{-y^2} dx dy = \int_0^\infty \int_0^\infty e^{-(x^2+y^2)} dx dy. \end{aligned}$$

Nyt teemme muuttujien vaihdon $x = r \cos(\theta)$ ja $y = r \sin(\theta)$ ja toteamme, että $(x, y) \in (0, \infty) \times (0, \infty)$ jos ja vain jos $r > 0$ ja $\theta \in (0, \frac{1}{2}\pi)$. Lisäksi pätee $dx dy = r dr d\theta$ ja $x^2 + y^2 = r^2$ joten

$$\begin{aligned} E^2 &= \int_0^{\frac{1}{2}\pi} \int_0^\infty e^{-r^2} r dr d\theta = \left(\int_0^{\frac{1}{2}\pi} 1 d\theta \right) \int_0^\infty \left(-\frac{1}{2} e^{-r^2} \right) \\ &= \frac{1}{2}\pi \frac{1}{2} (0 - (-1)) = \frac{1}{4}\pi, \end{aligned}$$

josta seuraa, että $E = \int_0^\infty e^{-x^2} dx = \frac{1}{2}\sqrt{\pi}$.

😊 Pyörähdyskappaleet I

Kun käyrä $y = f(x)$, $a \leq x \leq b$ tai $\{(t, f(t)) : t \in [a, b]\}$ pyörähtää x -akselin ympäri syntyy pyörähdyskappale joka sisältää kaikki pisteet (x, y, z) , joille pätee että niiden etäisyys x -akselista eli $\sqrt{y^2 + z^2}$ on korkeintaan $|f(x)|$ ja $a \leq x \leq b$. Voimme laskea tämän kappaleen tilavuuden sylinterikoordinaattien avulla seuraavalla tavalla:

$$x = x,$$

$$y = r \cos(\theta),$$

$$z = r \sin(\theta).$$

Silloin $dx dy dz = r dx dr d\theta$ ja integroimisrajat ovat $a \leq x \leq b$, $0 \leq r \leq |f(x)|$, $0 \leq \theta \leq 2\pi$. Näin ollen tilavuus on

$$\begin{aligned} & \int_a^b \left(\int_0^{2\pi} \left(\int_0^{|f(x)|} r dr \right) d\theta \right) dx \\ &= \left(\int_0^{2\pi} d\theta \right) \int_a^b \left(\int_0^{|f(x)|} \frac{1}{2} r^2 \right) dx = 2\pi \frac{1}{2} \int_a^b |f(x)|^2 dx = \pi \int_a^b f(x)^2 dx. \end{aligned}$$

😊 Pyörähdyskappaleet II

Kun x -akselin, suorien $x = a$ ja $x = b$ sekä funktion $y = f(x)$ kuvaajan rajoittama alue, missä $0 \leq a \leq x \leq b$ ja $f(x) \geq 0$, pyörähtää y -akselin ympyrä niin syntyy pyörähdyskappale

$$\Omega = \{ (x, y, z) : 0 \leq y \leq f(\sqrt{x^2 + z^2}), \quad a^2 \leq x^2 + z^2 \leq b^2 \}.$$

Tämän kappaleen tilavuuden voimme laskea seuraavanlaisten sylinterikoordinaattien avulla:

$$x = r \cos(\theta),$$

$$y = y,$$

$$z = r \sin(\theta),$$

jolloin $dx dy dz = r dr d\theta dy$ ja integroimisrajat ovat $a \leq r \leq b$, $0 \leq y \leq f(r)$, $0 \leq \theta \leq 2\pi$. Tilavuus tulee siten olemaan

$$\begin{aligned} V(\Omega) &= \int_0^{2\pi} \left(\int_a^b \left(\int_0^{f(r)} r dy \right) dr d\theta \right) = \left(\int_0^{2\pi} d\theta \right) \int_a^b r f(r) dr \\ &= 2\pi \int_a^b x f(x) dx. \end{aligned}$$