Tehtävä 1

Olkoon $\overline{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, $n \geq 2$. Osoita seuraava binomikaavan yleistys

$$(x_1 + x_2 + \cdots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \overline{x}^\alpha \text{ kaikilla } k \in \mathbb{N},$$

missä α on multi-indeksi.

Ratkaisu:

Osoitetaan väite induktiolla muuttujan n suhteen. Olkoon $k \in \mathbb{N}$. Kun $n = 2$, niin tavallisesta binomikaavasta saadaan

$$\begin{aligned}
(x_1 + x_2)^k &= \sum_{i=0}^{k} \binom{k}{i} x_1^k x_2^{k-i} = \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} x_1^k x_2^{k-i} \\
&= \sum_{\alpha_1+\alpha_2=k} \frac{k!}{\alpha_1!\alpha_2!} x_1^{\alpha_1} x_2^{\alpha_2} = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \overline{x}^\alpha,
\end{aligned}$$

sillä vaatimuksesta $\alpha_1 + \alpha_2 = k$ saadaan esimerkiksi, että $\alpha = (k - i, i)$ kaikilla $0 \leq i \leq k$. Siispä perusaskel pätee. Tehdään sitten induktio-oletus, että väite pätee jollakin $n = m > 2$. Tällöin, kun $n = m + 1$, saadaan induktio-oletuksesta ja jälleen tavallisesta binomikaavasta

$$\begin{aligned}
(x_1 + x_2 + \cdots + x_{m+1})^k &= ((x_1 + x_2 + \cdots + x_m) + x_{m+1})^k \\
&= \sum_{i=0}^{k} \binom{k}{i} (x_1 + x_2 + \cdots + x_m)^i x_{m+1}^{k-i} \\
&= \sum_{i=0}^{k} \binom{k}{i} x_{m+1}^{k-i} \sum_{\alpha_1+\cdots+\alpha_m=i} \frac{i!}{\alpha_1! \cdots \alpha_m!} x_1^{\alpha_1} \cdots x_m^{\alpha_m} \\
&= \sum_{i=0}^{k} \sum_{\alpha_1+\cdots+\alpha_m=i} \frac{k!}{\alpha_1! \cdots \alpha_m!(k-i)!} x_1^{\alpha_1} \cdots x_m^{\alpha_m} x_{m+1}^{k-i} \\
&= \sum_{\alpha_1+\cdots+\alpha_{m+1}=k} \frac{k!}{\alpha_1! \cdots \alpha_{m+1}!} x_1^{\alpha_1} \cdots x_{m+1}^{\alpha_{m+1}} \\
&= \sum_{|\alpha|=k} \frac{k!}{\alpha!} (x_1, x_2, \ldots, x_{m+1})^\alpha.
\end{aligned}$$

Väite seuraa nyt induktioperiaatteesta.

Tehtävä 2

Olkoon $P : \mathbb{R}^2 \to \mathbb{R}$ polynomi, jonka aste on korkeintaan 2 ja

$$\lim_{\overline{x}\to\overline{0}} \frac{P(\overline{x})}{\|\overline{x}\|^2} = 0.$$

Osoita, että $P(\overline{x}) = 0$ kaikilla $\overline{x} \in \mathbb{R}^2$.

Ratkaisu:

Merkitään mielivaltaista kahden muuttujan polynomia P siten, että

$$P(x, y) = Ax^2 + By^2 + Cxy + Dx + Ey + F,$$

missä $A, B, C, D, E, F \in \mathbb{R}$. Oletuksen $P(\overline{x})/\|\overline{x}\|^2 \to 0$, kun $\overline{x} \to \overline{0}$ nojalla raja-arvon täytyy olla sama kaikista mahdollisista lähestymissuunnista. Lähestytään siis origoa pitkin suoraa $y = x$, jolloin

$$\lim_{\substack{x\to 0\\ y=x}} \frac{P(\overline{x})}{\|\overline{x}\|^2} = \lim_{x\to 0} \frac{(A + B + C)x^2 + (D + E)x + F}{2x^2} = 0.$$

Jotta näin olisi, täytyy olla $A+B+C = 0$, $D+E = 0$ ja $F = 0$ (muuten raja-arvoa ei olisi olemassa tai se ei olisi nolla). Lähestytään sitten origoa pitkin suoraa $y = -x$. Tällöin

$$\lim_{\substack{x\to 0\\ y=-x}} \frac{P(\overline{x})}{\|\overline{x}\|^2} = \lim_{x\to 0} \frac{(A + B - C)x^2 + (D - E)x}{2x^2} = 0$$

jos ja vain jos $A+B-C = 0$ ja $D-E = 0$. Yhdistämällä saadut tulokset saadaan, että $A+B = 0$, $C = D = E = 0$. Lähestytään origoa sitten pitkin suoraa $y = 0$, jolloin

$$\lim_{\substack{x\to 0\\ y=0}} \frac{P(\overline{x})}{\|\overline{x}\|^2} = \lim_{x\to 0} \frac{Ax^2}{x^2} = 0$$

jos ja vain jos $A = 0$. Vastaavasti lähestymällä origoa pitkin suoraa $x = 0$ saadaan

$$\lim_{\substack{y\to 0\\ x=0}} \frac{P(\overline{x})}{\|\overline{x}\|^2} = \lim_{y\to 0} \frac{By^2}{y^2} = 0$$

jos ja vain jos $B = 0$. Siispä saatiin, että $A = B = C = D = E = F = 0$ eli $P(\overline{x}) = 0$ kaikilla $\overline{x} \in \mathbb{R}^2$ ja väite on siten osoitettu.

Tehtävä 3

Olkoot $f \in \mathcal{C}^2(\mathbb{R}^n)$ ja $\overline{x}_0, \overline{x} \in \mathbb{R}^n$. Määritellään $\varphi : \mathbb{R} \to \mathbb{R}$,

$$\varphi(t) = f(t\overline{x} + (1-t)\overline{x}_0).$$

Osoita ketjusäännön avulla, että $\varphi \in \mathcal{C}^2(\mathbb{R})$ ja

$$\varphi''(t) = \nabla^2 f(t\overline{x} + (1-t)\overline{x}_0)(\overline{x} - \overline{x}_0) \cdot (\overline{x} - \overline{x}_0) \text{ kaikilla } t \in \mathbb{R}.$$

Ratkaisu:

Edellisissä demoissa osoitettiin, että

$$\begin{aligned}
\varphi'(t) &= Df(t\overline{x} + (1-t)\overline{x}_0)D(t\overline{x} + (1-t)\overline{x}_0) \\
&= \nabla f(t\overline{x} + (1-t)\overline{x}_0)^T(\overline{x} - \overline{x}_0) \\
&= \nabla f(t\overline{x} + (1-t)\overline{x}_0) \cdot (\overline{x} - \overline{x}_0).
\end{aligned}$$

Jatketaan ratkaisua tästä. Olkoon $h : [0,1] \to \mathbb{R}^n$, $h(t) = t\overline{x} + (1-t)\overline{x}_0$, jolloin $\varphi(t) = f(h(t))$. Aiemman perusteella siis

$$\varphi'(t) = Df(h(t))(\overline{x} - \overline{x}_0).$$

Ketjusäännön nojalla kuvaus $f \circ h$ on kahdesti jatkuvasti derivoituva, jolloin Clairaut'n lauseen nojalla Hessen matriisi on symmetrinen. Tällöin ketjusäännöllä

$$\begin{aligned}
\varphi''(t) &= D[Df(h(t))(\overline{x} - \overline{x}_0)]Dh(t) \\
&= \left[\nabla^2 f(h(t))(\overline{x} - \overline{x}_0)\right](\overline{x} - \overline{x}_0) \\
&= \left[\nabla^2 f(h(t))(\overline{x} - \overline{x}_0)\right]^T(\overline{x} - \overline{x}_0) \\
&= \nabla^2 f(t\overline{x} + (1-t)\overline{x}_0)(\overline{x} - \overline{x}_0) \cdot (\overline{x} - \overline{x}_0)
\end{aligned}$$

ja saatiin väite.

Tehtävä 4

Määrää seuraavien funktioiden lokaalit ääriarvokohdat ja niiden laadut:

(a) $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x,y,z) = e^{x^2-4y+y^2} + z^2$.

(b) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \cos(x+y) + \sin(x+y)$.

Ratkaisu:

(a) Ketjusäännöllä gradientiksi tulee

$$\nabla f(x,y,z) = \left(2xe^{x^2-4y+y^2}, (2y-4)e^{x^2-4y+y^2}, 2z\right),$$

jolloin Lauseen 4.2 nojalla ääriarvopisteissä täytyy olla $\nabla f(x_0,y_0,z_0) = \overline{0}$. Koska

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

kaikilla $x, y \in \mathbb{R}$, niin

$$\nabla f(x,y,z) = \overline{0} \iff (2x, 2y-4, 2z) = \overline{0} \iff (x,y,z) = (0,2,0).$$

Hessen matriisi saadaan tulon derivaatan ja ketjusäännön avulla niin, että

$$\begin{aligned}\nabla^2 f(x,y,z) &= \begin{bmatrix} \partial_{11} f(x,y,z) & \partial_{21} f(x,y,z) & \partial_{31} f(x,y,z) \\ \partial_{12} f(x,y,z) & \partial_{22} f(x,y,z) & \partial_{32} f(x,y,z) \\ \partial_{13} f(x,y,z) & \partial_{23} f(x,y,z) & \partial_{33} f(x,y,z) \end{bmatrix} \\ &= \begin{bmatrix} (4x^2+2)e^{x^2-4y+y^2} & 4x(y-2)e^{x^2-4y+y^2} & 0 \\ 4x(y-2)e^{x^2-4y+y^2} & 2(2y^2-8y+9)e^{x^2-4y+y^2} & 0 \\ 0 & 0 & 2 \end{bmatrix}.\end{aligned}$$

Pisteessä $(0,2,0)$ Hessen matriisi saa arvon

$$\nabla^2 f(0,2,0) = \begin{bmatrix} 2e^{-4} & 0 & 0 \\ 0 & 2e^{-4} & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Kyseessä on diagonaalimatriisi, joten sen ominaisarvot ovat suoraan diagonaalialkiot, jotka selvästi ovat positiivisia. Siispä Lauseen 4.8 nojalla Hessen matriisin määräämä neliömuoto on positiivisesti definiitti ja siten Lauseesta 4.9 seuraa, että piste $(0,2,0)$ on minimipiste.

(b) Ketjusäännöllä gradientiksi saadaan

$$\nabla f(x,y) = (\cos(x+y) - \sin(x+y), \cos(x+y) - \sin(x+y)) .$$

Kriittisiä pisteitä ovat

$$\nabla f(x,y) = 0 \iff \sin(x+y) = \cos(x+y) \iff x+y = \frac{(4n+1)\pi}{4},$$

missä $n \in \mathbb{Z}$. Jos Hessen matriisin laskisi näissä pisteissä, vastaava neliömuoto olisi semidefiniitti. Ääriarvotestiä ei voi nyt käyttää. Huomataan kuitenkin, että kun $z \in \mathbb{R}$, niin

$$\sin z + \cos z = \sqrt{2}\sin(z + \pi/4).$$

Kun n on parillinen,

$$\frac{(4n+1)\pi}{4} = \frac{\pi}{4} + 2k\pi,\ \ k \in \mathbb{Z},$$

saadaan

$$\sin(x+y) + \cos(x+y) = \sqrt{2}\sin\left(\frac{\pi}{2} + 4k\pi\right) = \sqrt{2}.$$

Yksiulotteisesta analyysista tiedetään, että kyseessä on sinifunktion maksimipiste. $x+y$ kuvautuu nyt vakioksi, eli näitä suoria pitkin funktion arvot pysyvät vakiona. Tuloksena on siis ylösalainen kouru. Siispä kaikilla parillisilla n kuvaus saavuttaa (ei-aidon) lokaalin maksimin.

Vastaavasti, kun n on pariton,

$$\frac{(4n+1)\pi}{4} = \frac{5\pi}{4} + 2k\pi,\ \ k \in \mathbb{Z},$$

niin

$$\sin(x+y) + \cos(x+y) = \sqrt{2}\sin\left(\frac{3\pi}{2} + 4k\pi\right) = -\sqrt{2}$$

eli kyseessä on lokaali ei-aito minimi pitkin suoraa $x+y = 5\pi/4 + 2k\pi$, tällä kertaa normaalin kourun muodossa.

Tehtävä 5

Etsi funktion $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = e^{x+y}\cos(xy)$ Taylorin polynomi astetta 2 pisteessä $\overline{x}_0 = (0,0)$ sekä määritelmää että funktioiden $u \mapsto e^u$ ja $v \mapsto \cos v$ (tavallisia) Taylorin polynomeja käyttäen.

Ratkaisu:

Lasketaan aluksi osittaisderivaattoja tulon derivointikaavalla ja ketjusäännöllä:

$$\begin{aligned}
\partial_1 f(x,y) &= e^{x+y}(\cos xy - y\sin xy),\\
\partial_2 f(x,y) &= e^{x+y}(\cos xy - x\sin xy),\\
\partial_{12} f(x,y) &= -e^{x+y}(xy-1)\cos xy + (x+y+1)\sin xy),\\
\partial_{11} f(x,y) &= -e^{x+y}((y^2-1)\cos xy + 2y\sin xy),\\
\partial_{22} f(x,y) &= -e^{x+y}((x^2-1)\cos xy + 2x\sin xy).
\end{aligned}$$

Tällöin Taylorin polynomin määritelmästä

$$\begin{aligned}
T_{\overline{0},2} f(x,y) &= \sum_{|\alpha|\le 2} \frac{1}{\alpha!}\frac{\partial^\alpha}{\partial x^\alpha} f(0,0)(x,y)^\alpha\\
&= \overbrace{f(0,0)}^{\alpha=(0,0)} + \overbrace{\partial_1 f(0,0)x}^{\alpha=(1,0)} + \overbrace{\partial_2 f(0,0)y}^{\alpha=(0,1)} + \overbrace{\partial_{12} f(0,0)xy}^{\alpha=(1,1)} + \overbrace{\frac{1}{2}\partial_{11} f(0,0)x^2}^{\alpha=(2,0)} + \overbrace{\frac{1}{2}\partial_{22} f(0,0)y^2}^{\alpha=(0,2)}\\
&= 1 + x + y + xy + \frac{1}{2}x^2 + \frac{1}{2}y^2.
\end{aligned}$$

Toisaalta yksiulotteisesta analyysistä tiedetään, että funktiot $u \mapsto e^u$ ja $v \mapsto \cos v$ voidaan esittää Taylorin polynomien avulla siten, että

$$e^u = 1 + u + \frac{1}{2}u^2 + R_1(u),$$

missä $|R_1(u)| \le C_1 |u|^3$ ja

$$\cos v = 1 - \frac{1}{2}v^2 + R_2(v),$$

missä $|R_2(v)| \le C_2 |v|^3$. Tällöin

$$\begin{aligned}
f(x,y) &= e^{x+y}\cos xy\\
&= \left(1 + (x+y) + \frac{1}{2}(x+y)^2 + R_1(x,y)\right)\left(1 - \frac{1}{2}(xy)^2 + R_2(x,y)\right)\\
&= 1 + x + y + xy + \frac{1}{2}x^2 + \frac{1}{2}y^2 + R(x,y),
\end{aligned}$$

missä $|R(x,y)| \le C\|(x,y)\|^3$. Taylorin polynomin yksikäsitteisyydestä seuraa nyt, että $T_{\overline{0},2} f(x,y) = 1 + x + y + xy + \frac{1}{2}x^2 + \frac{1}{2}y^2$, kuten odotettua.

Tehtävä 6

Olkoot $a, b, c \in \mathbb{R}$ ja määritellään

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \text{ ja } Q : \mathbb{R}^2 \to \mathbb{R},\ Q(\overline{x}) = \overline{x}^T A \overline{x}.$$

Osoita, että

(a) neliömuoto Q on positiivisesti (negatiivisesti) definiitti jos ja vain jos $a > 0$ $(a < 0)$ ja $b^2 < ac$.

(b) neliömuoto Q on indefiniitti jos ja vain jos $b^2 > ac$.

Ratkaisu:

(a) Oletetaan aluksi, että $Q(\overline{x})$ on positiivisesti definiitti. Lauseen 4.8 nojalla matriisin A ominaisarvoille pätee $\lambda_1, \lambda_2 > 0$. Tällöin lineaarialgebran tiedoilla

$$\det A = ac - b^2 = \lambda_1\lambda_2 > 0,$$

joten $b^2 < ac$. Toisaalta

$$Q(1,0) = \left(\begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = a > 0,$$

sillä positiivisesti definiitille neliömuodolle pätee määritelmän nojalla $Q(\overline{x}) > 0$ kaikilla $\overline{x} \in \mathbb{R}^2 \setminus \{\overline{0}\}$.

Oletetaan sitten, että $b^2 < ac$ ja $a > 0$. Tällöin

$$\det A = \lambda_1\lambda_2 = ac - b^2 > 0,$$

joten ominaisarvoilla λ_1, λ_2 on sama merkki. Lauseesta 4.8 seuraa, että $Q(\overline{x})$ on joko positiivisesti tai negatiivisesti definiitti. Kuitenkin

$$Q(1,0) = a > 0$$

oletuksen nojalla, joten neliömuodon $Q(\overline{x})$ täytyy olla positiivisesti definiitti.

Väite negatiivisesta definiittisyydestä osoitetaan vastaavaan tapaan. Jos $Q(\overline{x})$ on negatiivisesti definiitti, Lauseen 4.8 nojalla $\lambda_1, \lambda_2 < 0$, jolloin

$$\det A = ac - b^2 = \lambda_1\lambda_2 > 0,$$

eli $b^2 < ac$. Toisaalta

$$Q(1,0) = a < 0,$$

sillä $Q(\overline{x})$ on negatiivisesti definiitti. Jos taas oletetaan, että $b^2 < ac$ ja $a < 0$, niin

$$\det A = \lambda_1 \lambda_2 = ac - b^2 > 0,$$

joten ominaisarvoilla on sama merkki. Koska

$$Q(1, 0) = a < 0,$$

täytyy olla, että $Q(\overline{x})$ on negatiivisesti definiitti Lauseen 4.9 nojalla.

(b)

$$\begin{aligned} Q(\overline{x}) \text{ on indefiniitti} &\overset{\text{L 4.8}}{\iff} \lambda_1 < 0 < \lambda_2 \text{ tai } \lambda_2 < 0 < \lambda_1 \\ &\iff \lambda_1 \lambda_2 < 0 \\ &\iff \det A = \lambda_1 \lambda_2 = ac - b^2 < 0 \\ &\iff b^2 > ac. \end{aligned}$$

Tehtävä 7

Olkoon $f \in \mathcal{C}^2(\mathbb{R}^3)$. Määritellään $\Phi_f : \mathbb{R}^3 \to \mathbb{R}$,

$$\Phi_f(\overline{x}) = \det(\nabla^2 f(\overline{x})).$$

(a) Osoita, että Φ_f on jatkuva joukossa $\mathbb{R}^3$.

(b) Oletetaan lisäksi, että funktiolla f on lokaali maksimi pisteessä $\overline{y}$ ja lokaali minimi pisteessä $\overline{z}$. Osoita, että on olemassa $\overline{w}$ siten, että $\Phi_f(\overline{w}) = 0$.

Ratkaisu:

(a) Koska $f \in \mathcal{C}^2(\mathbb{R}^3)$, Hessen matriisin kaikki alkiot

$$\nabla^2 f(\overline{x}) = \begin{bmatrix} \partial_{11} f(\overline{x}) & \partial_{21} f(\overline{x}) & \partial_{31} f(\overline{x}) \\ \partial_{12} f(\overline{x}) & \partial_{22} f(\overline{x}) & \partial_{32} f(\overline{x}) \\ \partial_{13} f(\overline{x}) & \partial_{23} f(\overline{x}) & \partial_{33} f(\overline{x}) \end{bmatrix}$$

ovat jatkuvia kuvauksia. Φ_f eli Hessen matriisin determinantti koostuu[1] näiden jatkuvien kuvausten tuloista ja summista, jolloin myös kuvaus Φ_f on jatkuva joukossa $\mathbb{R}^3$.

(b) Lokaaleissa ääriarvopisteissä $\overline{y} \in \mathbb{R}^3$ ja $\overline{z} \in \mathbb{R}^3$ pätee Lauseen 4.2 nojalla

$$\nabla f(\overline{y}) = \overline{0} = \nabla f(\overline{z}).$$

Toisaalta ääriarvotestistä seuraa, että pisteessä $\overline{y}$ Hessen matriisi on negatiivisesti definiitti ja vastaavasti pisteessä $\overline{z}$ se on positiivisesti definiitti. Tästä seuraa Lauseen 4.8 nojalla, että matriisin $\nabla^2 f(\overline{y})$ ominaisarvoille pätee $\lambda_1, \lambda_2, \lambda_3 < 0$ ja matriisin $\nabla^2 f(\overline{z})$ ominaisarvoille puolestaan pätee $\mu_1, \mu_2, \mu_3 > 0$. Lineaarialgebran tiedoilla Hessen matriisin determinantti on ominaisarvojen tulo, joten

$$\Phi_f(\overline{y}) = \det(\nabla^2 f(\overline{y})) = \lambda_1 \lambda_2 \lambda_3 < 0$$

ja

$$\Phi_f(\overline{z}) = \det(\nabla^2 f(\overline{z})) = \mu_1 \mu_2 \mu_3 > 0.$$

Valitaan[2] nyt $r > 0$ niin, että $\overline{y}, \overline{z} \in \overline{B}(\overline{0}, r)$. Pallo $\overline{B}(\overline{0}, r)$ on selvästi suljettu, rajoitettu ja murtoviivayhtenäinen. Vektorianalyysi 1-kurssin tiedoilla Φ_f saavuttaa joukossa $\overline{B}(\overline{0}, r)$ suurimman ja pienimmän arvonsa sekä kaikki arvot niiden välistä. Erityisesti, koska $\Phi_f(\overline{y}) < 0$ ja $\Phi_f(\overline{z}) > 0$, täytyy minimille m päteä $m < 0$ ja vastaavasti maksimille M päteä $M > 0$. Siispä on olemassa $\overline{w} \in \overline{B}(\overline{0}, r) \subset \mathbb{R}^3$ siten, että $\Phi_f(\overline{w}) = 0$.

[1] Toki determinantin voisi kirjoittaa auki eksplisiittisesti, mutta siitä tulee aikamoinen sotku eikä se oikeastaan lisää mitään tähän vastaukseen.

[2] Esimerkiksi $r = \max\{\|\overline{y}\|, \|\overline{z}\|\}$ kelpaa.