7. Többváltozós függvények

7.1. Definíció, ábrázolás

Az előző fejezetekben olyan függvényeket vizsgáltunk, amelyek értelmezési tartománya a valós számok részhalmaza. Ekkor a független változót gyakran az idővel azonosítjuk. A fizikai testek térben helyezkednek el. A tér pontjait az (x, y, z) koordinátahármasokkal azonosítjuk. Ezért egy test hőmérséklete, sűrűsége, egy oldat töménysége $f(x, y, z)$ alakú *háromváltozós függvényekkel* írhatók le, ahol x, y, z változók egymástól függetlenül változnak. De ha a koncentráció például az időben is változik, akkor a folyamatot, a koncentrációt az adott t pillanatban és (x, y, z) helyen $f(t, x, y, z)$ alakú *négyváltozós függvénnyel* jellemezhetjük.

Az olyan függvényeket, amelyek értelmezési tartománya valamely $\mathbf{R}^n$ *halmaz részhalmaza, más szavakkal, több egymástól független valós változója van,* **többváltozós függvényeknek** *nevezzük.* Egy kétváltozós függvény értelmezési tartománya a sík ($\mathbf{R}^2$) részhalmaza, egy háromváltozós függvény értelmezési tartománya $\mathbf{R}^3$-nak részhalmaza.

7.1. Példa. Az $f(x, y) = \sqrt{x} + \sqrt{y}$ függvénynek a $D = \{x \geq 0, \quad y \geq 0\}$ halmaz, a sík pozitív negyede az értelmezési tartománya. Az $f(x, y) = \sqrt{1 - x^2 - y^2}$ függvénynek az $\{x^2 + y^2 \leq 1\}$ halmaz, az egységsugarú zárt körlap az értelmezési tartománya.

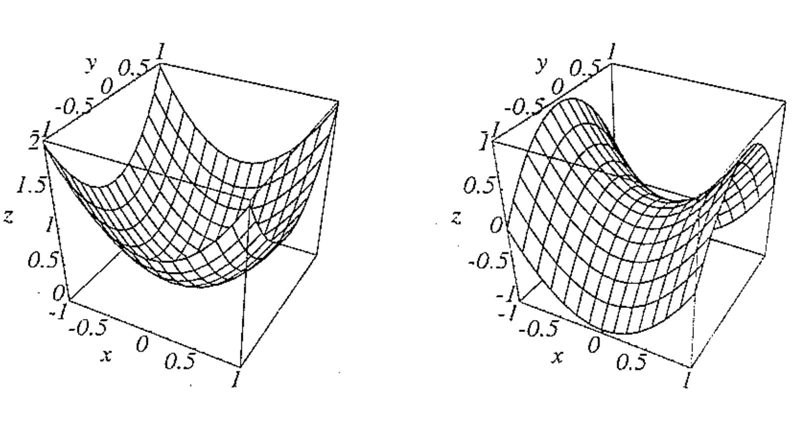
A többváltozós függvények lehetnek szám– és vektorértékűek is (lásd az előző fejezetet). Ebben a fejezetben a számértékű többváltozós függvényekkel foglalkozunk. Az egyszerűség kedvéért főként az $f : \mathbf{R}^2 \to \mathbf{R}$ függvényeket tekintjük. Meggondolásaink általában érvényesek maradnak $f : \mathbf{R}^3 \to \mathbf{R}$ és általánosabb $f : \mathbf{R}^n \to \mathbf{R}$ függvényekre is értelemszerű módosításokkal. Ezekre a megfelelő helyeken utalni fogunk.

7.1.1. Ábrázolás grafikonnal

A függvények grafikonja definció szerint az $(a, f(a))$ pontpárok halmaza. Ha $a = (x, y) \in \mathbf{R}^2$, akkor a grafikon egy $\{(x, y, f(x, y)) : (x, y) \in D_f \subseteq \mathbf{R}^2\}$ az $\mathbf{R}^3$ tér részhalmaza, amely az (x, y, z) koordinátarendszerben ábrázolható. A z koordinátatengely a függvényértékek koordinátája. Ha a függvény három– vagy többváltozós, akkor a grafikon $\mathbf{R}^4$ vagy valamely $\mathbf{R}^n$ $(n > 3)$, ezért nem tudjuk ábrázolni. Ezért más megjelenítési ábrázolási módszerekre is szükség lesz.

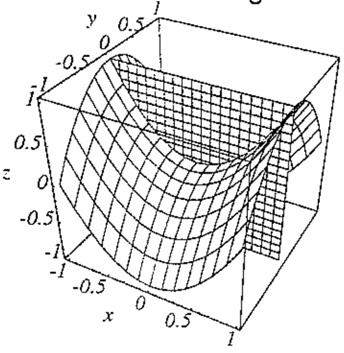
Kétváltozós függvények grafikonját síkban perspektivikusan ábrázoljuk. Azonban egy térbeli felületet nehéz olyan pontosan ábrázolni, mint például a síkbeli görbéket. Ezért a grafikont ún. vonalhálók segítségével jelenítjük meg.

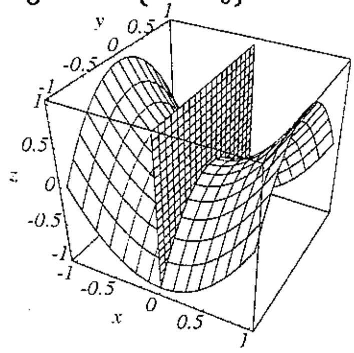
A változók egyikét rögzítjük valamely értéken. A függvényt $y = konstans$ és $x = konstans$ egyenesek mentén számoljuk ki. Ekkor egyváltozós függvényeket kapunk, amelyeket ábrázolni tudunk. Ezt több érték mellett mindkét változó esetén megtesszük. Általában egyenletes beosztásokat választunk. A kapott $(x, c, f(x, c))$ és $(c, y, f(c, y))$ görbéket ábrázoljuk az (x, y, z) koordinátarendszerben. Számítógéppel a vonalak közé felületet feszíthetünk ki, amely közelíti (!) a függvény grafikonját. A közelítés annál pontosabb, minél több $y = konstans$ és $x = konstans$ egyenes mentén számoltunk. Az eredményt az $f(x, y) = x^2 + y^2$ és $f(x, y) = x^2 - y^2$ függvények esetén az alábbi ábra mutatja.



Megjegyezzük, hogy valamely változó, vagy változók rögzítése általános eljárás, amely a vizsgált problémát egyszerűsíti különösen a három- és többváltozós esetekben.

A kétváltozós esetben az y változó y_0 értéknél való rögzítésének grafikus jelentése, hogy a grafikont $\mathbf{R}^3$-ban az $\{y = y_0\}$ síkkal (az (x, z) síkkal párhuzamos, az y tengelyt y_0-ban metsző sík) elmetszük. Az x változó rögzítése x_0-ban a grafikon $\{x = x_0\}$ síkkal való elmetszését jelenti.

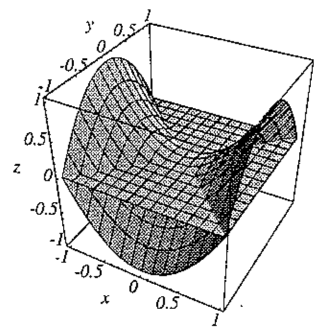




7.1.2. Ábrázolás szintvonalakkal, szintfelületekkel

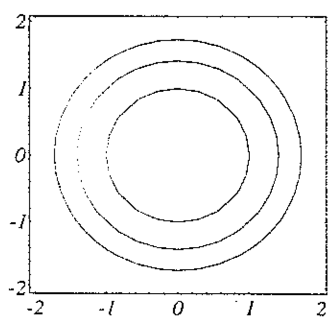
A többváltozós függvények megjelenítése lehetséges az ún. szintvonalak segítségével. Alkalmazzuk a térképek készítésénél a magasságok és mélységek jelzésére bevált technikát. Ott síkban rajzolják meg az azonos magasságokat jelző vonalakat.

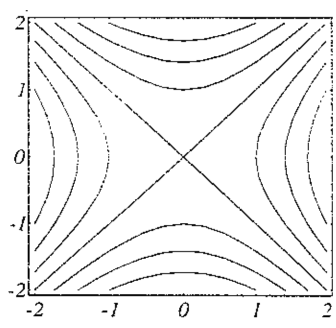
Legyen $f(x, y)$ kétváltozós függvény, és legyen a z_0 érték adott. A függvény $\mathbf{R}^3$-beli grafikonját metsszük el a $z = z_0$ síkkal, vagyis az (x, y) síkkal párhuzamos, a z-tengelyt a z_0 értéknél metsző síkkal. A metszésvonal bármely (x, y, z) pontjára $z = z_0$, és $f(x, y) = z_0$. Az $f(x, y) = x^2 - y^2$ függvény grafikonjának a $z = 0$ síkkal való metszetét láthatjuk az ábrán.



A metszésvonalat vetítsük az (x, y) síkra! ***A kapott görbe az értelmezési tartomány azon pontjainak a halmaza, amelyekre $f(x, y) = z_0$. Az $f(x, y) = z_0$ egyenletet kielégítő (x, y) pontok halmazát nevezzük az $f(x, y)$ függvény z_0-hoz tartozó*** **szintvonalának**.

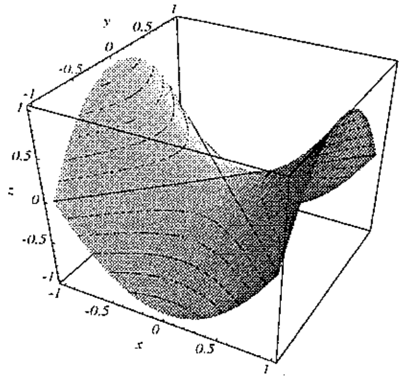
A függvény pontosabb ábrázolásához több szintvonalra van szükség. Ekkor a térképekhez hasonlóan a vonalakhoz írjuk a megfelelő függvényértéket. Az egyes szintek közötti átmeneteket gyakran a szín illetve színárnyalat változása jelzi. Az $f(x, y) = x^2 + y^2$ és a $g(x, y) = x^2 - y^2$ függvények $z_0 = -3, -2, -1, 0, 1, 2, 3$ szintvonalait látjuk a következő ábrákon.

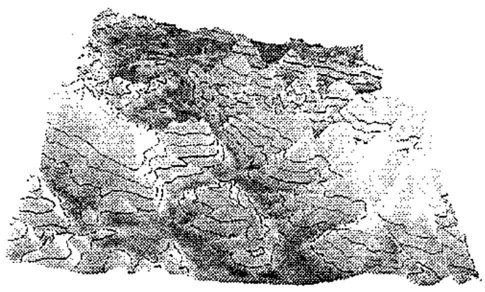




A grafikon és a szintvonalak együttesen adják a térképészetben elterjedt háromdimenziós térképet. Ekkor a grafikonon megjelöljük az (x, y) síkkal párhuzamos síkokkal való metszés során kapott metszésvonalakat.

Az ábrán az $f(x, y) = x^2 - y^2$ függvény "szintvonalas grafikonja" és egy valódi háromdimenziós szintvonalas térkép látható.





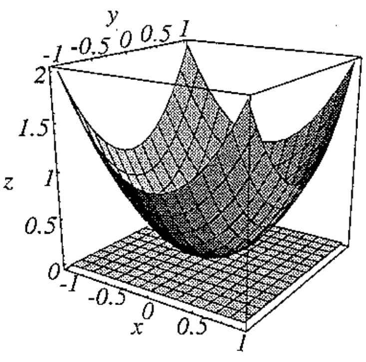
Ha a függvény háromváltozós, akkor az $f(x, y, z) = konstans$ egyenleteket kielégítő halmazok felületek $\mathbf{R}^3$-ban. Például az $f(x, y, z) = x^2 + y^2 + z^2$ függvény szintfelületei az origó középpontú gömbök.

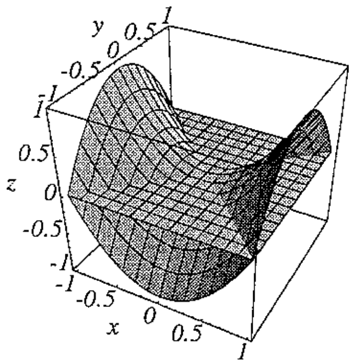
7.2. Szélsőértékek, parciális deriváltak

A szélsőértékek keresése a többváltozós függvények esetében is fontos probléma.

Az (x_0, y_0) pont lokális maximumhelye az $f(x, y)$ $\mathbf{R}^2 \to \mathbf{R}$ függvénynek, ha van egy (x_0, y_0) pont körüli körlap, melynek minden pontjában $f(x, y) \leq f(x_0, y_0)$. Az (x_0, y_0) pont lokális minimumhely, ha $f(x, y) \geq f(x_0, y_0)$ az (x_0, y_0) közelében.

A grafikonokat ismerve láthatjuk, hogy az $f(x, y) = x^2 + y^2$ függvénynek a $(0, 0)$ pont lokális minimumhelye (baloldali ábra), de az $f(x, y) = x^2 - y^2$ függvénynek (nyeregfelület) sem minimuma, sem maximuma (jobboldali ábra).





Az egyváltozós esetben a szélsőértékek keresése szoros kapcsolatban van a monotonitással. Mivel a sík és tér pontjai közt nincs értelmezve a klasszikus $a < b$ fogalom, ezért a monotonitás fogalmának a többváltozós függvények esetén nincs értelme. Értelmezhető azonban, ha az $f(x, y)$ függvényt egyváltozóssá egyszerűsítjük az egyik változó rögzítésével, mint ahogyan azt a fentiekben az ábrázolás kapcsán tettük.

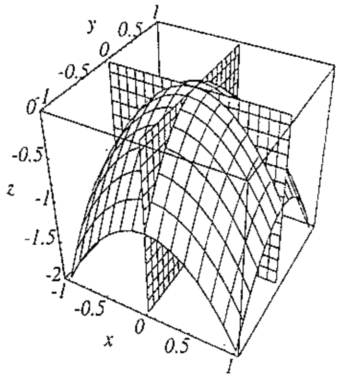
Legyen $(x_0, y_0) \in \mathbf{R}^2$ rögzített pontja az értelmezési tartománynak. A $g(x) = f(x, y_0)$ egyváltozós függvény monotonitása, szélsőértékei a derivált segítségével vizsgálhatók. Ha $f(x, y)$-nak (x_0, y_0)-ban szélsőértéke van, akkor $g(x)$-nek is szélsőértéke van x_0-ban és

$$g'(x_0) = 0.$$

A $h(y) = f(x_0, y)$ deriváltjára pedig

$$h'(y_0) = 0,$$

ha az $f(x, y)$-nek szélsőértéke van (x_0, y_0)-ban.



Egy fontos eredményre jutottunk. Az egyik változó rögzítésével kapott egyváltozós függvények deriváltjának speciális szerepe van a többváltozós függvények vizsgálatában.

7.1. Definíció. *Az $f(x, y)$ függvény y változóját rögzítve kapott x-től függő egyváltozós függvény deriváltját az $f(x, y)$ függvény x szerinti parciális deriváltjának nevezzük.*

Jelölése

$$\frac{\partial f(x, y)}{\partial x}.$$

Az x változót rögzítve az y-tól függő egyváltozós függvény deriváltja az $f(x, y)$ függvény y-szerinti parciális deriváltja, melynek jelölése

$$\frac{\partial f(x, y)}{\partial y}.$$

Formálisan

$$\frac{\partial f(x, y)}{\partial x} = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h},$$

$$\frac{\partial f(x, y)}{\partial y} = \lim_{h \to 0} \frac{f(x, y + h) - f(x, y)}{h}.$$

A parciális deriváltak kiszámítása a definíció alapján a deriválás ismert szabályai szerint történik. Eszerint

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

A $\partial f / \partial x$ parciális derivált az $f(x, y)$ függvény x-irányú, míg $\partial f / \partial y$ az $f(x, y)$ függvény y irányú változásának mértékét mutatja.

A parciális deriváltak fogalmának felhasználásával megfogalmazhatjuk a fentebb bebizonyított állítást a szélsőértékre vonatkozóan.

7.1. Tétel. *Legyen $f(x, y)$ parciálisan differenciálható valamely téglalapon $\mathbf{R}^2$-ben. Az $f(x, y)$ szélsőértékhelyei kielégítik a*

$$\frac{\partial f(x, y)}{\partial x} = 0, \quad \frac{\partial f(x, y)}{\partial y} = 0$$

egyenletrendszert.

Az egyváltozós függvényeknél tapasztaltak sejtetik, hogy ez nem elegendő. Grafikus tapasztalataink is ezt igazolják. Az $f(x, y) = x^2 - y^2$ függvényt parciális deriváltjainak zéróhelye az origó, de az nem szélsőértékhely.

A konvexség valamilyen általánosítására van szükségünk, hogy a tételbeli egyenletrendszer megoldásai közül kiválasszuk a szélsőértékhelyeket. A konvexitás többváltozós kiterjesztése nem tárgya a tananyagnak. Csak a legszükségesebb segédeszközöket tekintjük. Tegyük fel, hogy léteznek és folytonosak a

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2},$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$$

második parciális deriváltak is. Belátható, hogy ekkor

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Igaz a következő tétel.

7.2. Tétel. *Tegyük fel, hogy (x_0, y_0) megoldása az*

$$\frac{\partial f(x,y)}{\partial x} = 0, \quad \frac{\partial f(x,y)}{\partial y} = 0$$

egyenletnek.

Tegyük fel, hogy az $f(x,y)$ összes második parciális deriváltja folytonos (x_0, y_0)-ban és

$$\left(\frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2\right)\Bigg|_{(x,y)=(x_0,y_0)} > 0,$$

Ha $\partial^2 f/\partial x^2 > 0$, akkor (x_0, y_0)- ban $f(x,y)$-nak minimuma van.

Ha $\partial^2 f/\partial x^2 < 0$, akkor maximum van.

Ha

$$\left(\frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2\right)\Bigg|_{(x_0,y_0)} < 0,$$

akkor nincs szélsőérték.

Ellenőrizzük a fenti tétel segítségével az elemi megfontolások alapján már ismert tényeket. Legyen $f(x,y) = x^2 + y^2$. Ekkor

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

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

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 0.$$

Az $2x = 0$; $2y = 0$ rendszer megoldása $(0,0)$ és

$$\frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 = 4 > 0, \quad \frac{\partial^2 f}{\partial x^2} = 2 > 0,$$

tehát a $(0,0)$ minimum pont.

Ugyanígy belátható, hogy az $f(x,y) = x^2 - y^2$ függvénynek a $(0,0)$-ban nincs szélsőértéke. Ez a függvény grafikonjának ismeretében várható volt.

A kettőnél több változó esetén a parciális deriváltak hasonló módon definiálhatók. Az n változós $f(x_1, \ldots, x_n)$ függvény x_i szerinti parciális deriváltja a

$$\lim_{h \to 0} \frac{f(x_1, \ldots x_i + h, \ldots x_n) - f(x_1, \ldots x_i, \ldots x_n)}{h} = \frac{\partial f}{\partial x_i},$$

vagyis az x_i kivételével minden változót konstansnak tekintve, a kapott egyváltozós függvényt (x_i-től függ) deriváljuk. $\partial f / \partial x_i$ a függvény x_i irányú változását mutatja. A második parciális deriváltak hasonlóan definiálhatók:

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

Az $f(x_1, \ldots, x_n)$ függvény szélsőértékhelyei kielégítik az

$$\frac{\partial f}{\partial x_i} = 0, \; \ldots, \; \frac{\partial f}{\partial x_n} = 0$$

egyenletrendszert.

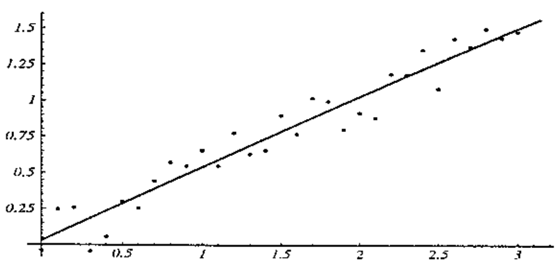
A kétváltozós esethez hasonló elegendő feltétel megfogalmazása túl bonyolult, ezért mellőzzük. Igaz azonban, ha a függvény

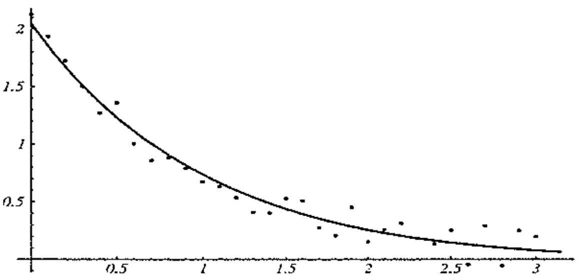
$$f(x_1, \cdots, x_n) = \sum_{i=1}^{n} (a_i x_i + b_i)^2$$

alakú, akkor a függvénynek van szélsőértéke. Ha azonban az összegben valamelyik $(a_i x_i + b_i)^2$ tag előjele negatív, akkor a függvény egy n-dimenziós nyeregfelület, és így nincs szélsőértéke.

7.3. A legkisebb négyzetes közelítés

Tegyük fel, hogy mérések során kaptuk az $(x_1, y_1), (x_2, y_2) \ldots (x_n, y_n)$ számpárokat, pontokat az (x, y) koordinátarendszerben. A pontok elhelyezkedése sejteti, hogy az x, y mennyiségek között milyen függvénykapcsolat állhat fenn. A baloldali esetben a pontok valamilyen egyenes mentén, míg a jobboldali ábrán egy csökkenő exponenciális függvény mentén helyezkednek el.





Feladatunk, hogy valamely **függvénycsaládból** *válasszuk ki azt a függvényt, amely az adott ponthalmazt legpontosabban közelíti meg. A legpontosabb közelítés azt jelenti, hogy a pontok és a keresett $h(x)$ függvény közti négyzetes távolság, azaz*

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

az adott függvénycsaládhoz tartozó függvények között minimális. A függvénycsalád megadása paraméterekkel történik. Ha a pontpárok egyenes mentén helyezkednek el, akkor a pontokat az

$$y = ax + b$$

alakú, míg az exponenciálisan kapcsolat esetén például $y = b10^{ax}$ alakú függvényekkel közelíthetjük. Bonyolultabb problémáknál több paraméter is szükséges lehet.

A feladat tehát a közelítő függvénycsalád paramétereinek meghatározása. *Ez egy többváltozós szélsőértékszámítási feladat.*

A probléma megoldását gyakran visszavezetjük a legegyszerűbb esetre, egyenessel való közelítés esetére. Ezért a lineáris közelítés problémáját vizsgáljuk a továbbiakban.

Tekintsük az

$$y = ax + b$$

alakú lineáris közelítést.

Az egyenes értéke az x_i helyen

$$ax_i + b.$$

Az y_i-től való eltérés

$$\delta_i = y_i - (ax_i + b).$$

Az (x_i, y_i) $(i = 1 \dots n)$ pontoktól való négyzetes eltérés akkor a legkisebb, ha az

$$S(a,b) = \sum_{i=1}^{n} \delta_i^2 = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

függvénynek minimuma van. Képezzük a $\frac{\partial S}{\partial a}, \frac{\partial S}{\partial b}$ parciális deriváltakat:

$$\frac{\partial S}{\partial a} = -\sum_{i=1}^{n} 2(y_i - ax_i - b)x_i, \quad \frac{\partial S}{\partial b} = -\sum_{i=1}^{n} 2(y_i - ax_i - b).$$

Oldjuk meg a

$$\sum_{i=1}^{n} 2(y_i - ax_i - b)x_i = 0, \quad \sum_{i=1}^{n} 2(y_i - ax_i - b) = 0$$

egyenletrendszert:

$$\sum_{i=1}^{n} y_i x_i - a\sum_{i=1}^{n} x_i^2 - b\sum_{i=1}^{n} x_i = 0, \quad \sum y_i - a\sum_{i=1}^{n} x_i - nb = 0.$$

Vezessük be az $\bar{x} = \frac{1}{n}\sum x_i$, $\bar{y} = \frac{1}{n}\sum y_i$ jelöléseket. Az $\bar{x}$ és $\bar{y}$ az x_i illetve az y_i számok $(i = 1 \dots n)$ átlagai. E jelölésekkel kapjuk, hogy

$$b = -a\bar{x} + \bar{y} \quad (\bar{y} = a\bar{x} + b),$$

$$\sum y_i x_i - a\sum x_i^2 - n\bar{x}(-a\bar{x} + \bar{y}) = 0,$$

$$a(\sum x_i^2 - n\bar{x}^2) = \sum y_i x_i - n\bar{x}\bar{y}.$$

Végül:

$$a = \frac{\sum y_i x_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}.$$

A kapott a együtthatót a statisztikában lineáris regressziós együtthatónak, a kapott

$$y = ax + b$$

egyenest pedig regressziós egyenesnek nevezzük. A b értékre kapott összefüggés azt fejezi ki, hogy az egyenes átmegy az $(\bar{x}, \bar{y})$ ponton. Itt nem foglalkozunk azzal, hogy a közelítés mennyire jó. Csak azt tudjuk, hogy az egyenesek között a legjobb. A közelítés jóságára vonatkozó részletesebb elemzések a statisztika tárgykörébe tartoznak.

Végül ejtsünk néhány szót a lineáris közelítésre való visszavezetésről. Ha exponenciális

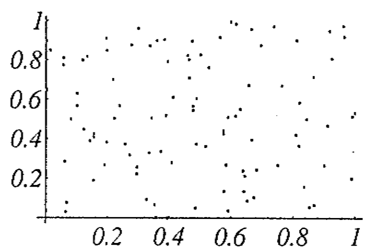
$$y = B10^{ax}$$

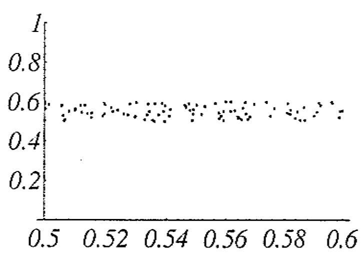
alakú összefüggést sejtünk, akkor mindkét oldal logaritmusát véve kapjuk az

$$\lg y = \lg B + ax$$

egyenlőséget. Ugyanazt tesszük, mint a függvények szemilogaritmikus ábrázolásánál (1.22. fejezet). Az y logaritmusa és x között már lineáris az összefüggés. Ezért az $\{(x_i, \lg y_i)\}$ értékpárokra a fenti lineáris közelítést alkalmazhatjuk. Ha a regressziós egyenes $y = ax + b$, akkor az eredeti ponthalmazt az $f(x) = 10^{ax+b}$ függvény közelíti legjobban. Ugyanezt a függvényt kapnánk, ha az eredeti pontsorozatot közvetlenül a $y = B10^{ax}$ alakú exponenciális függvénnyel közelítenénk.

Befejezésül tekintsünk néhány speciális tulajdonságú adathalmazt, melyeknél további vizsgálatok nélkül nincs értelme jól illeszkedő függvényt keresni.





A baloldali ábrán olyan ponthalmazt látunk, ahol az (x_i, y_i) pontok véletlenszerűen helyezkednek el. Láthatóan nincs kapcsolat a két koordináta között.

A jobboldali ábrán az adatok egy $y = konstans$ egyenest követnek. Az x változása esetén y konstans marad, ami arra utal, hogy az y kooordináta által leírt tulajdonság nem függ x-től.

Utolsó ábránkon az adathalmaz kettős tulajdonságú. A pontok két egymást metsző egyenest követnek. Ennek több oka lehet, de felmerül a gyanú, hogy két különböző adathalmaz összekeveredett, a mérések körülményei változtak, vagy a mintavétel nem volt megfelelő.