Bevezetés, tudnivalók, Valszám és statisztika ismétlés

1. előadás

2017. szeptember 5.

Tartalom

- Technikai kérdések: tudnivalók a félévről adminisztratív szempontból
- Bevezetés, alapgondolatok, célok és eszközök
- Valószínűség-számítási ismeretek - rövid ismétlés!!!
- Statisztika ismeretek - rövid ismétlés!!!
- Többváltozós statisztika - alapok

Adminisztratív ügyek

- **Tananyag:** elsősorban az előadás fóliák (fent lesznek a honlapon), másodsorban a honlapon lévő nyomtatott jegyzet! Vigyázat, a nyomtatott jegyzet régi, nincs is benne a teljes anyag!
- **honlap:** www.math.bme.hu/∼orlovits
- **Könyv:**
  - Ramu Ramanathan: Bevezetés az ökonometriába alkalmazásokkal, Panem Kiadó, 2003
  - G.S. Maddala: Bevezetés az ökonometriába, Nemzeti Tankönyvkiadó, 2004
  - Jeffrey M. Wooldridge: Introductory Econometrics, A Modern Approach

Adminisztratív ügyek

- **Számonkérések:** 1 db 90 perces zárthelyi a 13. héten, valamint a félév során egy ún. cikkfeldolgozási feladat, azaz
  - egy szabadon választható, ökonometria témájú cikk feldolgozása
  - max. 5 oldalas összefoglaló elkészítése
  - formai követelmények, útmutató és az ajánlott folyóiratok listája a honlapon
- Osztályozás a szokásos módon
- Minden egyéb kérdésre a válasz a tárgykövetelményekben a honlapon

Célok, feladatok

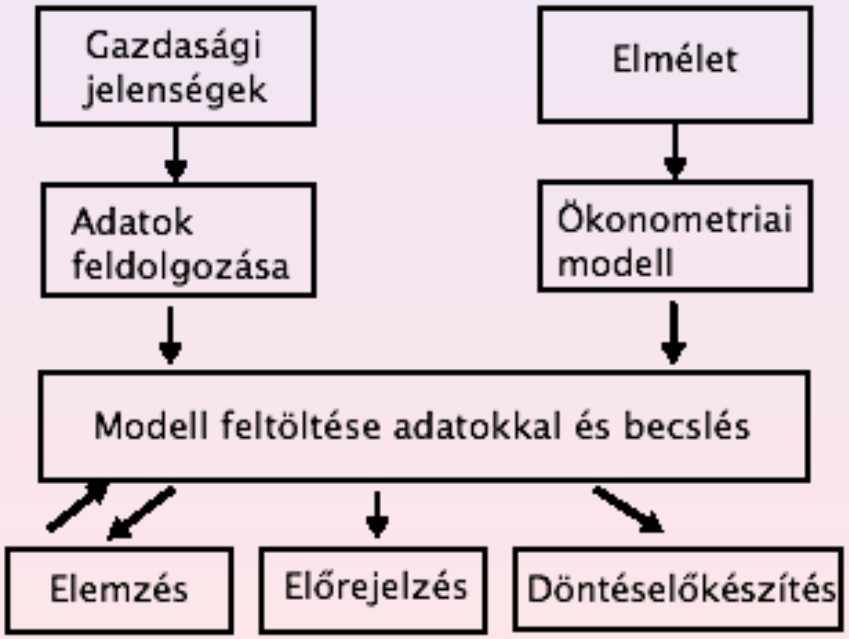
- Ismerkedés az ökonometriával: elmélet, módszerek, alkalmazások
- Mindezt modellorientált szemléletben!
- Cél: társadalmi-gazdasági jelenségek kvantitatív elemzése
- Szükséges matematikai alapok: lineáris algebra, valószínűségszámítás, statisztika

Féléves tematika

- Valószínűségszámítási alapismeretek rövid ismétlése
- Statisztika alapismeretek rövid ismétlése
- Többváltozós statisztika néhány alapvető elemének tárgyalása
- Lineáris regresszió-számítás
- Idősorelemzés

Mi az ökonometria?

- **Feladata:** társadalmi-gazdasági jelenségek statisztikai elemzése, modellezése
- A közgazdaságtan, matematika és statisztika kölcsönhatásából kialakuló, ún. határtudományról beszélünk majd.



Mi az ökonometria?

- A valóság egyszerű mását akarjuk létrehozni. Ám valóság túl bonyolult és összetett, másolni lehetetlen, így egyszerűsítünk, de ezzel persze torzítunk is! Erre majd vigyázni kell!
- A modellezés kulcsa az absztrakciós szint helyes megválasztása. Legyen a modell valósághű, de még kezelhető!
- A valóságban jelen lévő állandó bizonytalanság miatt ún. sztochasztikus modellekkel foglalkozunk.
- A struktúrát előre megadjuk, ismeretlen paraméterekkel. Feladatunk e paraméterek becslése a minta segítségével.
- Persze a struktúrán változtatni lehet, de csak bizonyos határokon belül. Ez lesz a modell-specifikáció és diagnosztika kérdése.
- Cél: elemzés, előrejelzés
- Lépések: hipotézis felállítása, adatgyűjtés, modellezés, diagnosztika

Valószínűségszámítási alapok

Kolmogorov-féle valószínűségi mező: $(\Omega, \mathcal{F}, \mathcal{P})$, ahol

- Ω: valószínűségi tér (mintatér), $\omega \in \Omega$ elemekkel (elemi események)
- $\mathcal{F}$: σ-algebra
- $\mathcal{P}$: valószínűségi mértékek családja $\mathcal{F}$-en

Valószínűségi változó (v.v.): olyan $X : \Omega \to \mathbb{R}$ függvény, melyre $X^{-1}(B) \in \Omega$ bármely B Borel-halmazra

Diszkrét valószínűségi változó: $R(X)$ értékhalmaza megszámlálható, és ekkor az eloszlás olyan (x_i, p_i) párok sorozata, ahol x_i az $R(X)$ értékhalmaz egy felsorolása, $p_i = P(X = x_i)$.

Valószínűségi vektorváltozó (v.v.v.): p darab v.v. vektora

Valószínűségszámítási alapok

Eloszlásfüggvény: X p-dimenziós v.v.v. esetén

$$F(x) = F(x_1, \ldots, x_p) = P(\omega | X_1(\omega) \leq x_1, \ldots, X_p(\omega) \leq x_p).$$

Sűrűségfüggvény: ha $F(x)$ folytonos a Lebesgue-mértékre, akkor létezik, és

$$f(x) = F'(x), \quad F(x) = \int_{-\infty}^{x} f(t)dt, \quad f(x) \geq 0, \quad \int_{-\infty}^{\infty} f(t)dt = 1.$$

Várható érték: diszkrét v.v. esetén

$$\mathrm{E}X = \sum_i x_i p_i,$$

ha a sor abszolút konvergens, míg folytonos esetben

$$\mathrm{E}X = \int_\Omega X(\omega)\mathrm{d}P(\omega) = \int_{\mathbb{R}} x\,\mathrm{d}F(x) = \int_{\mathbb{R}} xf(x)\mathrm{d}x$$

Valószínűségszámítási alapok

Momentumok: $\mathrm{E}(X^k)$ a k-dik momentum, $\mu_k = \mathrm{E}(X - \mathrm{E}X)^k$, $k \in \mathbb{N}$, a k-adik centrális momentum. Speciálisan

- $\mu_2 = \sigma^2 = Var(X)$ a szórásnégyzet
- μ_3/σ^3 az aszimmetria mértéke (skewness)
- μ_4/σ^4 a csúcsosság és farok-eloszlás mértéke (kurtosis)

Tétel 1.

A magasabb momentumok létezéséből következik az alacsonyabb momentumok létezése, a megfordítás viszont nem igaz.

Függetlenség: X és Y függetlenek, ha $\forall A, B$ Borel-halmazra

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

Speciálisan

$$F(x, y) = F_X(x)F_Y(y).$$

Valószínűségszámítási alapok

Ekkor $\mathrm{E}(XY) = \mathrm{E}X \cdot \mathrm{E}Y$, és $Var(X+Y) = Var(X) + Var(Y)$.

Feltételes valószínűség:

$$P(A|B) = \frac{P(AB)}{P(B)}, \quad \text{ha} \quad P(B) \neq 0.$$

Kovariancia, korreláció:

$$Cov(X,Y) = \mathrm{E}[(X - \mathrm{E}X)(Y - \mathrm{E}Y)].$$

Speciálisan $Cov(X,X) = Var(X)$. A két változó közötti kapcsolat erősségét méri, de számszerű értéke függ a változók értékeinek nagyságrendjétől.

- szimmetrikus, bilineáris függvény

Valószínűségszámítási alapok

Ezt a problémát küszöböli ki a **korrelációs együttható:**

$$R(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

Ismeretes, hogy $|R(X,Y)| \leq 1$, és ugyancsak a két változó közötti lineáris kapcsolat erősségét méri:

- ha $R \approx \pm 1$, akkor majdnem lineáris a kapcsolat a változók közt
- ha $R \approx 0$, akkor nemlineáris a kapcsolat a változók közt

Tétel 2.

Független v.v-k korrelálatlanok (azaz $R(X,Y) = 0$), de ez visszafelé nem igaz.

Mi a statisztika?

- "A statisztika a matematika azon ága, melynek alapfeladata az, hogy a politikus kezébe olyan eszközt adjon, mellyel tetszőleges állítás és annak ellentéte is tudományos alapon igazolható." (ismeretlen forrás)
- A statisztika a világ számszerűsíthető tényeinek szisztematikus összegyűjtésével és elemzésével foglalkozó tudományos módszer és gyakorlat.

Feladat, cél: a tapasztalati adatokból az információk kinyerése, statisztikai törvényszerűségek feltárása, következtetések levonása és felhasználása. Modellépítés, paraméterbecslés, következtetések, hipotézisek vizsgálata.

Alapfogalmak

Kiinduló pont: $(\Omega, \mathcal{F}, \mathcal{P})$ statisztikai mező, ahol a $\mathcal{P}$ mértékcsalád olyan P eloszlásokból áll, melyekkel $(\Omega, \mathcal{F}, P)$ valószínűségi mező. A probléma éppen a megfelelő eloszlás kiválasztása valamilyen $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, $\Theta \subset \mathbb{R}^k$ paraméteres családból.

Vizsgálódás tárgya: X v.v., pl. testmagasság, melynek pontos P eloszlását nem ismerjük, csak annyit tudunk, hogy $P \in \mathcal{P}$.

Cél: paraméterbecslés, hipotézis-vizsgálat

Eszköz: a statisztikai minta, mely valamilyen véletlen mennyiségre vonatkozó véges számú független megfigyelés (lehetséges) eredménye, azaz véges sok független, azonos eloszlású v.v.

Realizáció: egy konkrét kimenetel, $T(X) = T(X_1, \ldots, X_n)$ a statisztika, ahol T mérhető függvény.

Alapfogalmak

Alapstatisztikák: kiinduló tájékozódás az $X_1, \ldots, X_n$ mintáról

- Mintaátlag: $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- Tapasztalati szórásnégyzet: $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \overline{X^2} - \overline{X}^2$
- Korrigált tapasztalati szórásnégyzet:
  $S^{*2} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$
- Mintaátlag standardizált hibája: $\frac{\overline{X} \cdot \sqrt{n}}{S^*}$
- k-dik tapasztalati centrális momentum: $M_k^c = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k$

Alapfogalmak

- Ferdeség (szimmetria): $\dfrac{M_3^c}{(M_2^c)^{2/3}}$
- Lapultság: $\dfrac{M_4^c}{(M_2^c)^2 - 3}$
- Tapasztalati kovariancia: (X_i, Y_i), $i = 1, \ldots, n$, 2-dim i.i.d. minta,

$$C = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) = \frac{1}{n}\sum_{i=1}^{n} X_i Y_i - \overline{XY}$$

- Tapasztalati korrelációs együttható:

$$R = \frac{C}{S_X S_Y},$$

ahol S_X és S_Y a komponensek tapasztalati szórásai.

Feladat

Feladat: a θ paramétert vagy annak valamilyen $\Psi(\theta)$ függvényét szeretnénk becsülni az $X = (X_1, \ldots, X_n)$ i.i.d. minta alapján konstruált $\hat{\theta}_n = T_n(X)$ (avagy $\Psi(\hat{\theta}_n) = T_n(X)$) statisztika segítségével. Ő lesz majd a paraméter avagy megfelelő függvényének becslése.

Az olyan becsléseket, amelyek megadott paraméterek becslését adják, paraméter- (vagy pont)-becslésekнek nevezzük.

A becslés jóságának mérése: a valódi paraméter körüli ingadozás és a sztochasztikus konvergencia segítségével történik.

Alapfogalmak röviden

Torzítatlanság: $\Psi(\hat{\theta}_n) = T_n(X)$ torzítatlan becslés $\Psi(\theta)$-ra, ha

$$\mathrm{E}(\Psi(\hat{\theta}_n)) = \Psi(\theta), \quad \forall \theta \in \Theta.$$

Pl.: az átlag mindig torzítatlan becslése a várható értéknek, ha ez véges.

Aszimpotitkus torzítatlanság: a $T_n(X)$ statisztika-sorozat aszimptotikusan torzítatlan becslése $\Psi(\theta)$-nak, ha

$$\lim_{n\to\infty} \mathrm{E}(\Psi(\hat{\theta}_n)) = \Psi(\theta), \quad \forall \theta \in \Theta.$$

Pl.: szórásnégyzetre a tapasztalati szórásnégyzet. (A korrigált változat torzítatlan is!)

Alapfogalmak röviden

Konzisztencia: $\Psi(\hat{\theta}_n) = T_n(X)$ konzisztens becslés $\Psi(\theta)$-ra, ha $\Psi(\hat{\theta}_n) \to \theta$, $n \to \infty$ sztochasztikusan, azaz bármely $\varepsilon > 0$ esetén

$$P(|\Psi(\hat{\theta}_n) - \Psi(\theta)| > \varepsilon) \to 0, \quad n \to \infty, \quad \forall \theta \in \Theta.$$

Erős konzisztencia: $\Psi(\hat{\theta}_n) = T_n(X)$ erősen konzisztens becslés $\Psi(\theta)$-ra, ha konzisztens és szórása nullához tart $n \to \infty$ esetén.

Pl.: $\bar{X}$ konzisztens és erősen is konzisztens becslése a várható értéknek, és ez persze nem más, mint a nagy számok erős és gyenge törvénye.

Hatásosság: Ha T_1 és T_2 torzítatlan becslések $\Psi(\theta)$-ra, akkor T_1 hatásosabb, mint T_2, ha

$$D^2_\theta(T_1) \leq D^2_\theta(T_2), \quad \forall \theta \in \Theta.$$

Egy becslés hatásos, ha minden más becslésnél hatásosabb.

Becslési elvek

Maximum-likelihood becslés: az ismeretlen $\Psi(\theta)$ paraméter becsléseként azt a $\Psi(\hat{\theta}_n)$ értéket vesszük, amely mellett az $x_1, \ldots, x_n$ minta valószínűsége a legnagyobb.

- Diszkrét esetben: az

$$L_\theta(x_1, \ldots, x_n) = p_\theta(x_1) \cdot \ldots \cdot p_\theta(x_n)$$

  likelihood-függvény maximumát keressük.

- Folytonos esetben: az

$$L_\theta(x_1, \ldots, x_n) = f_\theta(x_1) \cdot \ldots \cdot f_\theta(x_n)$$

  likelihood-függvény maximumát keressük, ahol f a megfelelő sűrűségfüggvény.

Becslési elvek

Tétel 3.

Ha $\hat{\theta}$ a θ paraméter maximum-likelihood becslése, akkor a $\Psi(\theta)$ maximum-likelihood becslése éppen $\Psi(\hat{\theta})$.

Amivel most nem foglalkozunk:

- momentumok módszere
- legkisebb négyzetek módszere - ezt majd a regressziónál tárgyaljuk részletesen

Információs határ

Legyen paraméterünk egydimenziós, és $L(x, \theta)$ a likelihood-fv.

Definíció 1.

Az $X_1, \ldots, X_n$ statisztikai minta Fisher-féle információjának az

$$I_n(\theta) = \mathrm{E}_\theta \left[\left(\frac{\partial}{\partial \theta} \log L(X, \theta) \right)^2 \right] \geq 0$$

mennyiséget hívjuk, ha derivált létezik és a várható érték véges.

Tétel 4.

Ha a fenti becslés torzítatlan és $D_\theta^2(T) < \infty$ minden $\theta \in \Theta$ esetén, akkor

$$D_\theta^2(T) \geq \frac{(\Psi'(\theta))^2}{I_n(\theta)}.$$

Konfidencia-intervallumok

A pontbecslések helyett térjünk most át az intervallum-becslésekre.

Legyen $X_1, \ldots, X_n$ statisztikai minta a P_θ eloszláscsaládból. A θ paraméterhez olyan (T_1, T_2) véletlen hosszúságú intervallumot keresünk, melyre

$$P(T_1 \leq \theta < T_2) \geq 1 - \varepsilon,$$

ahol $\varepsilon > 0$ kicsi. Itt T_1 és T_2 maguk is v.v.-k, a minta valamilyen függvényei. Ezt az intervallumot a θ paraméterre vonatkozó legalább $1 - \varepsilon$ megbízhatósági szintű konfidencia-intervallumnak nevezzük.

Alapfogalmak: minta és adatmátrix

- Legyen $\underline{x} = (x_1, \ldots, x_p)^\intercal$ p-dimenziós valószínűségi vektorváltozó, azaz olyan véletlen vektor, melynek minden koordináta-változója egy-egy 1-dimenziós valószínűségi változó. Tegyük fel, hogy adott n darab $\underline{x}^1, \ldots, \underline{x}^n \in \mathbb{R}^p$ független realizáció. Ez a p-**dimenziós** n **elemű i.i.d. minta**.

- Azt a $p \times n$ dimenziós mátrixot, melynek oszlopai a mintaelemek, **adatmátrix**nak nevezzük. A pontos jelölés:

$$X = \begin{pmatrix} x_1^1 & x_1^2 & \ldots & x_1^n \\ x_2^1 & x_2^2 & \ldots & x_2^n \\ \vdots & \vdots & & \vdots \\ x_p^1 & x_p^2 & \ldots & x_p^n \end{pmatrix}$$

Alapfogalmak: várható érték vektor

- Tegyük fel, hogy az $\underline{X}$ p-dimenziós valószínűségi vektorváltozó minden komponensének létezik a várható értéke. Ekkor

$$\underline{\mu} = (\mu_1, \ldots, \mu_p)^\intercal = \mathrm{E}\underline{X} = (\mathrm{E}X_1, \ldots, \mathrm{E}X_p)^\intercal$$

  a v.v.v. **várható érték vektora**.

- n-elemű statisztikai minta esetén $\underline{\mu}$ mintából való becslése az **átlagvektor**, mely a komponensek átlagainak vektora. Azaz

$$\overline{\underline{X}} = \frac{1}{n}\sum_{i=1}^{n} \underline{X}_i$$

  A becslés torzítatlan és erősen konzisztens.

Alapfogalmak: kovariancia mátrix

- Legyen az $\underline{X}$ p-dimenziós valószínűségi vektorváltozó olyan, mely minden komponensének létezik a második momentuma. A Cauchy-Schwarz egyenlőtlenség alapján ekkor bármely két komponense szorzatának is létezik a várható értéke, így a kovarianciájuk is. Az ezen értékekből felépített mátrix lesz $\underline{X}$ **kovariancia-mátrix**a, azaz

$$\Sigma = (\sigma_{ij})_{i,j=1}^{p} = \mathrm{E}[(\underline{X} - \mathrm{E}\underline{X})(\underline{X} - \mathrm{E}\underline{X})^{\intercal}] = (\mathrm{cov}(X_i, X_j))_{i,j=1}^{p}$$

- Σ n-elemű mintából való becslése a **tapasztalati kovariancia-mátrix**:

$$S_n = \frac{1}{n}\sum_{i=1}^{n}(\underline{X}_i - \overline{\underline{X}})(\underline{X}_i - \overline{\underline{X}})^{\intercal}$$

A korrigált változat:

$$S_n^* = \frac{n}{n-1}S_n$$

Alapfogalmak: szimmetria és definitség

Mind S_n, mind S_n^* szimmetrikus és pozitív szemidefinit, azaz

- $S_n^\mathsf{T} = S_n$ és $S_n^{*\mathsf{T}} = S_n^*$
- bármely $\underline{a} \in \mathbb{R}^p$ esetén $\underline{a}^\mathsf{T} S_n \underline{a} = \sum_{i=1}^{p}\sum_{j=1}^{p} a_i a_j (S_n)_{i,j} \geq 0$ (S_n^*-ra hasonlóan)

Ebből következik, hogy létezik ún. **spektrálfelbontás**uk, azaz

$$S_n = V \Lambda V^\mathsf{T} = \sum_{i=1}^{p} \lambda_i v_i v_i^\mathsf{T}$$

ahol $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$ a sajátértékek diagonális mátrixa, $V = (\underline{v}_1, \ldots, \underline{v}_p)$ a normált sajátvektorok alkotta ortogonális mátrix (azaz $VV^\mathsf{T} = V^\mathsf{T} V = I$). A sajátértékek a szimmetria miatt valósak, a pozitív szemidefinitség miatt pedig nemnegatívak. Ugyanez igaz S_n^* esetén is.

Többdimenziós standard normális eloszlás

Definíció 2.

*Legyenek $\xi_1, \ldots, \xi_p$ független $N(0,1)$ eloszlású valószínűségi változók. Ekkor a $\xi = (\xi_1, \ldots, \xi_p)^\mathsf{T}$ valószínűségi vektorváltozót p-**dimenziós standard normális eloszlású** v.v.v-nak nevezzük. Jelölés: $\xi \sim \mathcal{N}_p(\underline{0}, I_{p\times p})$.*

- $\mathrm{E}\xi = \underline{0}$, hiszen $\mathrm{E}\xi_i = 0$ minden $i = 1, \ldots, p$ esetén.
- $\Sigma_\xi = I_{p\times p}$, hiszen $D^2(\xi_i) = 1$ minden $i = 1, \ldots, p$ esetén és $\mathrm{cov}(\xi_i, \xi_j) = 0$ minden $i \neq j$ esetén.
- Sűrűségfüggvény:

$$f_\xi(\underline{x}) = \frac{1}{\left(\sqrt{2\pi}\right)^p} \exp\left\{-\frac{1}{2}\underline{x}^\mathsf{T}\underline{x}\right\}$$

ahol $\underline{x} = (x_1, \ldots, x_p)^\mathsf{T} \in \mathbb{R}^p$.

Többdimenziós normális eloszlás

Definíció 3.

Legyenek $\underline{m} \in \mathbb{R}^p$, $A \in \mathbb{R}^{p \times p}$ *és* $\xi \sim \mathcal{N}_p(\underline{0}, I_{p \times p})$. *Ekkor az* $\eta = A\xi + m$ *valószínűségi vektorváltozót* p-***dimenziós normális eloszlású*** *v.v.v-nak nevezzük. Jelölés:* $\eta \sim \mathcal{N}_p(\underline{m}, \Sigma_\eta)$.

- $\mathrm{E}\eta = \underline{m}$.
- $\Sigma_\eta = AA^\intercal$ szimmetrikus pozitív szemidefinit mátrix.
- Ha Σ_η invertálható, akkor létezik sűrűségfüggvény:

$$f_\eta(\underline{y}) = \frac{1}{\left(\sqrt{2\pi}\right)^p \cdot (\det \Sigma_\eta)^{1/2}} \exp\left\{-\frac{1}{2}(y - \underline{m})^\intercal \Sigma_\eta^{-1}(y - \underline{m})\right\}$$

ahol $\underline{y} = (y_1, \ldots, y_p)^\intercal \in \mathbb{R}^p$.

Fordított irány

A dolog visszafelé is működik, azaz ha $\eta \sim \mathcal{N}_p(\underline{m}, \Sigma_\eta)$ adott és Σ_η invertálható, akkor vegyük a

$$\Sigma_\eta = AA^\intercal$$

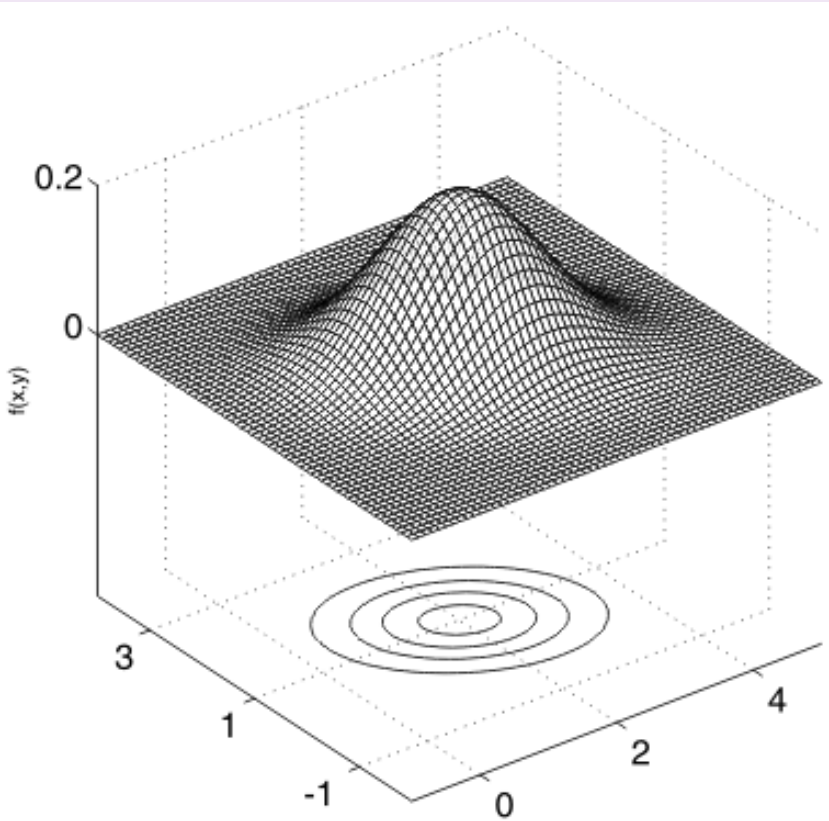
felbontást. Ez létezik, hiszen a spektrálfelbontásból $A = V\Lambda^{1/2}$ például jó választás lenne. A baj az, hogy ez még nem lesz egyértelmű felírás. Viszont, mivel Σ_η szimmetrikus és pozitív definit (hiszen feltettük, hogy invertálható), így az $A = V\Lambda^{1/2}$ felírás már egyértelmű. Ezt szokás $\Sigma_\eta^{1/2}$ jelöléssel is használni. Ekkor könnyen látszik, hogy a

$$\xi = A^{-1}(\eta - \underline{m})$$

lineáris transzformáció **standard normális eloszlású** valószínűségi vektorváltozót eredményez.

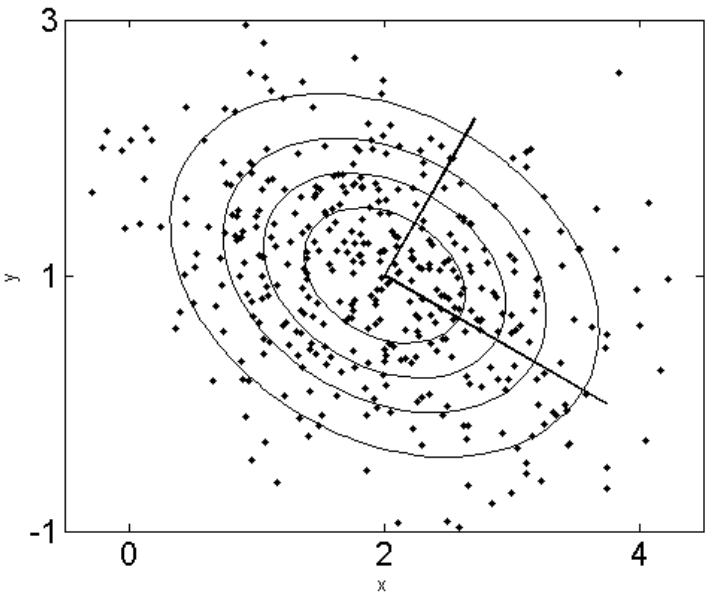
Szemléltetés: 2-dimenzió

A kétdimenziós standard normális eloszlás sűrűségfüggvényének képe egy harang alakú felület, mely tulajdonképpen a haranggörbe saját tengelye körüli megforgatottja. A szintvonalak körök lesznek. (Azaz függetlenek a komponensek és azonos szórásúak.)



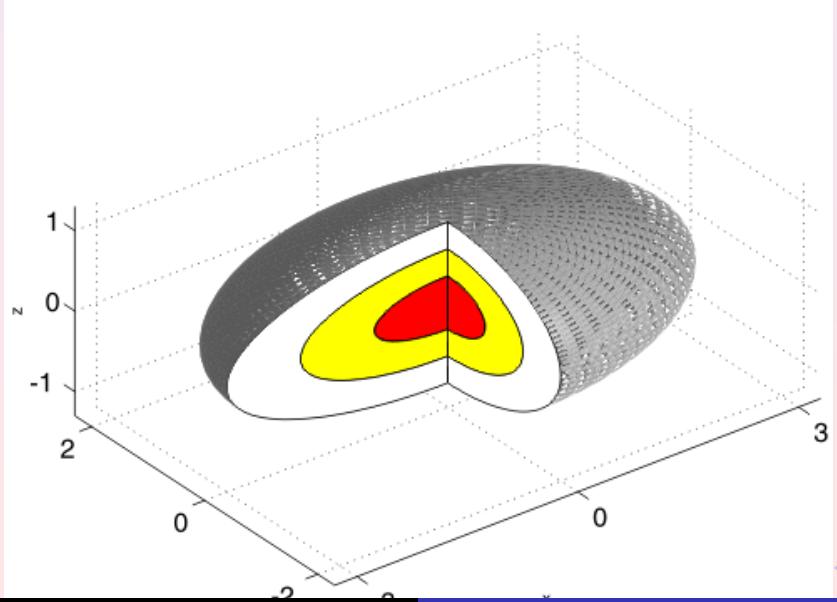
Szemléltetés: 2-dimenzió

Nem standard esetben ennek eltorzítottját kapjuk, melynek szintvonalai ekkor ellipszisek lesznek, $\underline{m}$ középponttal és $\sqrt{\lambda_1}, \sqrt{\lambda_2}$ hosszúsággal arányos, $\underline{v}_1, \underline{v}_2$ irányú tengelyekkel. Ezek az ún. koncentráció ellipszisek, melyek azt mutatják meg, hogy belülről kifelé haladva egyre kisebb valószínűséggel esnek a pontok az ellipszisekbe.



Szemléltetés: 3-dimenzió

Ebben az esetben a felületet már nem tudjuk felrajzolni, csak a szintvonalakat: azon pontok, melyeken a sűrűségfüggvény azonos értéket vesz fel, egy-egy ellipszoidon helyezkednek el. Ezek a koncentráció ellipszoidok, $\underline{m}$ középponttal és $\sqrt{\lambda_1}, \sqrt{\lambda_2}, \sqrt{\lambda_3}$ hosszúsággal arányos, $\underline{v}_1, \underline{v}_2, \underline{v}_3$ irányú tengelyekkel. Olyan, mint egy héjszerkezet, középen a legsűrűbb az anyag, kifelé pedig folyamatosan ritkul.



Paraméterbecslés: ML-módszer

Legyen $\underline{X}_1, \ldots, \underline{X}_n$ független minta az $\underline{X} \in \mathcal{N}_p(\underline{m}, C)$ vektorra, és tegyük fel, hogy $n \geq p$.

Feladat: a paraméterek, azaz $\underline{m}$ és C becslése, feltéve, hogy C pozitív definit.

Eszköz: maximum-likelihood (ML) becslés

Likelihood függvény:

$$L_{\underline{m},C}(\underline{X}_1, \ldots, \underline{X}_n) =$$

$$= \prod_{k=1}^{n} \frac{1}{(\sqrt{2\pi})^p \cdot (\det C)^{1/2}} \exp\left\{-\frac{1}{2}(X_k - \underline{m})^\intercal C^{-1}(X_k - \underline{m})\right\} =$$

$$= \frac{1}{(2\pi)^{\frac{np}{2}}(\det C)^{\frac{n}{2}}} \cdot \exp\left\{-\frac{1}{2}\sum_{k=1}^{n}(\underline{X}_k - \underline{m})^\intercal C^{-1}(\underline{X}_k - \underline{m})\right\}$$

Az ML-becslés tulajdonságai

- A várható érték becslése torzítatlan, míg a kovariancia-mátrix becslése aszimptotikusan torzítatlan, a korrigált változat persze torzítatlan lesz.
- A becsléseink konzisztensek és hatásosak.
- Alsó korlát adható az

$$I_n(\theta_0) = \frac{1}{n}\mathrm{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\ln L_{\theta_0}(\underline{X})\right)\left(\frac{\partial}{\partial\theta}\ln L_{\theta_0}(\underline{X})\right)^\mathsf{T}\right]$$

Fisher-információs mátrix segítségével, a

$$D_\theta^2(T) \geq I_n(\theta_0)^{-1}, \quad \theta \in \Theta$$

Cramér-Rao egyenlőtlenség alapján, ahol θ_0 és θ az igazi és a futó paramétervektorokat jelölik.