

A legkisebb négyzetes becslés tulajdonságainak vizsgálata egy qubit állapotára

tudományos diákköri dolgozat

Ruppert László matematikus hallgató
BME Analízis tanszék

Témavezetők: Magyar Attila ¹ és Petz Dénes ²

Kivonat

Egy kvantummechanikai rendszerről mérésekkel nyerünk információt. A mérések eredménye sztochasztikus, másrészt egy mérés lényegesen megváltoztatja a rendszer eredeti állapotát. Ezért sok azonos állapotban lévő rendszerre van szükség, ahhoz, hogy több mérést is elvégezhessünk. A mérési eredményekből statisztikus módszerekkel következtetünk a rendszer állapotára. Az egyik lehetőség a legkisebb négyzetes állapotbecslés, ami elég egyszerű, a hatékonyságát mutató szórása könnyen számolható. Ugyanakkor megvan az a hátránya, hogy nem mindig ad fizikailag lehetséges állapotot. Ha a kapott állapotbecslés értelmetlen, akkor módosításra szorul, helyette a legközelebbi fizikailag értelmes állapotot kell becslésként megadnunk.

A dolgozat célja az eredeti legkisebb négyzetes becslés szórásának összehasonlítása a módosított becslésével egy kvantum bit esetén. A kvantum bit állapotai egy gömb pontjaival adhatók meg. Ha az eredeti becslés a gömbön kívülre megy, akkor a módosítás a gömbfelszínre való vetítést jelenti. A dolgozat fő eredménye az, hogy az eredeti és a módosított becslés kovariancia mátrixai közötti különbség exponenciálisan tart nullához a mérések számának függvényében, amennyiben a rendszer valódi állapota gömb belsőjében van (kevert állapot). Az eredmény a nagyeltérés tétel alkalmazásával adódik. Emelett megvizsgáltunk néhány speciális állapotot, és számítógépes szimulációkat is végeztünk az eltérés nagyságára vonatkozóan. Amennyiben a valódi állapot nyomszerű, azaz a gömb középpontja, akkor a binomiális eloszlás normálissal való közelítése is eredményre vezet. Mivel az okoskodás nagyban kihasználja a szimmetriát, más pontokra való átvitele nem látszik járhatónak.

¹SZTAKI Rendszer- és Irányításelméleti Kutató Laboratórium

²BME Analízis tanszék

Tartalomjegyzék

1. Kvantum állapotbecslés	4
1.1. Bevezető	4
1.2. Definíciók, jelölések, fogalmak	5
1.2.1. A rendszer állapota	5
1.2.2. Mérések	5
1.2.3. A legkisebb négyzetes módszer	6
1.3. Eddigi irodalmi eredmények	7
1.3.1. Az eredeti becslés tulajdonságai	7
1.3.2. A megszorított becslés tulajdonságai	9
1.4. A probléma kitűzése	9
2. Egy speciális eset, a nyomszerű állapot varianciájának a vizsgálata	10
2.1. Áttérés binomiális eloszlásról normális eloszlásra	10
2.2. Binomális és normális eloszlást összekötő lemma	11
2.3. Becslés a normális eloszlás segítségével	12
2.4. További számítások	13
2.5. A konvergencia sebessége szimmetrikus esetben	14
3. A variancia megváltozása általános esetben	14
3.1. Nagy eltérés tétel alkalmazása 1 dimenzióban	15
3.2. Nagy eltérés tétel alkalmazása általánosan	16
3.3. A konvergencia sebessége általános esetben	17
3.4. Speciális eset: $\theta_1 = \theta_2 = \theta_3$	17
4. Szimulációk	20
4.1. Az varianciáról általában	20
4.2. A konvergencia sebessége	21
4.3. A becslések pontossága	23

1. Kvantum állapotbecslés

1.1. Bevezető

A kvantummechanikában a mérések sztochasztikus természete miatt nehezen rekonstruálható egy rendszer valódi állapota. Ezenkívül egy mérés megváltoztatja a rendszer eredeti állapotát is, így sok azonos állapotban lévő rendszerre van szükség, hogy egy adott állapotról több mérést is elvégezhessünk.

A fizikában az állapotbecslés kvantum tomográfia [1] néven közismert. Igencsak kiterjedt szakirodalommal rendelkezik, legismertebb módszerei közé tartozik a maximum likelihood becslés, Bayes becslés, és a dolgozat témájául szolgáló legkisebb négyzetes becslés. A Bayes becslés tulajdonságairól többet olvashatunk [6]-ban. Többek között a Bayes becslés és a legkisebb négyzetes becslés számítógépszimulációs összehasonlításával foglalkozik [4]. A vizsgált módszertől kissé eltérő van leírva [5]-ben, melyet dinamikusan alkalmazva egy érdekes számítógépes (numerikus) módszert kapunk az állapot becslésére. Azonban a dolgot meg lehet közelíteni konvex optimalizációs problémaként is [8].

A legkisebb négyzetes becslés népszerűségének titka egyszerűségében rejlik, egyszerűen számolható a becslés több tulajdonsága is, és emiatt könnyen általánosítható az eljárás magasabb dimenziókra. A kutatás tárgya a legegyszerűbb eset: egy qubit vizsgálata. Ebben az esetben az állapottér a Bloch-féle parametrizáció segítségével egy gömb lesz a háromdimenziós térben. Ám a standard becslés, különösen ha a mérésszám kicsi, illetve az eredeti állapot közel van a gömb felszínéhez, könnyen adhat gömbön kívüli eredményt, azaz egy olyan állapotot, aminek nincs valódi fizikai jelentése. Ez könnyen kiküszöbölhető, ha becslést megszorítjuk a valódi állapotok terére, ám ekkor a becslésünk elveszti kedvező tulajdonságait.

A dolgozat első részében e két módszer ismert tulajdonságai vannak felsorolva, valamint azok a fogalmak amelyek szükségesek a probléma tárgyalásához. Ezután következnek a saját eredmények: a két módszer varianciájának összehasonlítása, és az ezzel szorosan összefüggő, a fizikailag értelmetlen becslés valószínűségének becslése az eredeti módszerrel. A második részben egy speciális esetet (a teljesen kevert állapotot) vizsgálom meg a valószínűségszámítás eszközeivel, míg a harmadik részben a nagy eltérés tételek témakörét [2] hívom segítségül a probléma megoldásához. Az így kapott eredmények elég általánosak, felső becslést kapunk a konvergencia sebességére, ezért a negyedik részben számítógépes szimulációkkal megvizsgáljuk, hogy a kapott becslések mennyire térnek el a valóságtól.

1.2. Definíciók, jelölések, fogalmak

1.2.1. A rendszer állapota

Jelöljük θ -val az egyparaméteres rendszer (qubit) állapotát. Egy állapotot egyértelműen jellemez a sűrűségi mátrixa ($\rho(\theta)$), mely általában egy $n \times n$ -es mátrix, amely teljesíti az alábbi feltételeket:

- $\text{Tr}(\rho(\theta)) = 1$
- $\rho(\theta) \geq 0$

Esetünkben a sűrűségi mátrix 2×2 -es lesz. Valamint az előbbiből azt is tudjuk, hogy a főátlóban lévő számok összege rögzített, így az állapotterünk 3 paraméteres lesz. A pozitív szemidefinitivitás átírható geometriailag szemléletesebb alakba a Bloch-parametrizáció segítségével:

$$\rho(\theta) = \frac{1}{2} (I + \theta_1 \sigma_1 + \theta_2 \sigma_2 + \theta_3 \sigma_3),$$

ahol I az egységmátrix, továbbá σ_1, σ_2 és σ_3 a Pauli-mátrixok, lásd (2). Más formában

$$\rho(\theta) = \frac{1}{2} \begin{bmatrix} 1 + \theta_3 & \theta_1 - i\theta_2 \\ \theta_1 + i\theta_2 & 1 - \theta_3 \end{bmatrix} \quad (1)$$

Ezért a sűrűségi mátrix elemei helyett a megfelelő θ_i -kel tudjuk jellemezni az állapotot. Könnyű meggondolni, hogy ekkor a második feltétel az alábbira módosul:

$$\sqrt{\theta_1^2 + \theta_2^2 + \theta_3^2} \leq 1,$$

azaz az állapotterünk az egység sugarú (háromdimenziós) gömb. Azon állapotokat, amelyekre egyenlőség teljesül tiszta állapotoknak nevezzük, azon állapotokat amelyekre szigorú egyenlőtlenség teljesül kevert állapotoknak nevezzük. A továbbiakban az állapotot vagy a sűrűségi mátrix-szal, vagy a megfelelő θ vektorral jelöljük attól függően, hogy melyik jelölés a célszerűbb (ez nem okoz problémát, hiszen könnyedén átszámolható az egyik a másikba).

1.2.2. Mérések

A rendszer állapotát mérésel lehet rekonstruálni, azonban egy mérés eredménye csak tiszta állapot lehet, így ha az állapotunk kevert, akkor biztos, hogy 1 mérésel nem tudjuk megmérni az állapotát. Ezért felteszük, hogy sok azonos állapotú egymástól független qubitünk van.

A méréseink önadjungált 2×2 -es mátrixok. Ha a spektrális felbontása a mátrixnak $\sum_i \lambda_i P_i$, ahol P_i -k projekciók, akkor annak a valószínűsége, hogy a mérés során λ_i -t kapunk, feltéve, ha a kezdő állapotunk ρ :

$$\mathbf{P}_\rho(\lambda_i) = \text{Tr}(\rho P_i)$$

Háromfajta mérést hajtunk végre, ezek a

$$\sigma_1 := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (2)$$

Pauli-mátrixoknak felelnek meg. Mivel mindhárom mátrix unitér, sajátértékeik ± 1 . Ha σ_i mérést hajtunk végre a $\rho(\theta)$ állapotban, akkor a $+1$ kimenet valószínűsége

$$\frac{1 + \theta_i}{2}.$$

Fizikailag σ_1, σ_2 és σ_3 az x, y, z irányú spin mérést jelenti.

Tegyük fel, hogy mindhárom irányban m mérést hajtunk végre. Ekkor az adatokat tárolhatjuk, egy $3 \times m$ -es mátrixban, például:

$$D = \begin{bmatrix} +1_x & -1_x & -1_x & \dots & +1_x \\ -1_y & -1_y & +1_y & \dots & +1_y \\ +1_z & +1_z & -1_z & \dots & -1_z \\ 1. & 2. & 3. & \dots & m. \end{bmatrix} \begin{array}{l} \leftarrow x \text{ irányban} \\ \leftarrow y \text{ irányban} \\ \leftarrow z \text{ irányban} \end{array}$$

Látható lesz, hogy elég ha az egyes sorokban megjegyezzük a $+1$ -ek relatív gyakoriságát:

$$\nu_\xi := \frac{m_\xi}{m}, \quad \text{ahol } \xi = 1, 2, 3, \text{ és } m_\xi \text{ a } (+1)\text{-esek száma a } \xi\text{-edik sorban.}$$

1.2.3. A legkisebb négyzetes módszer

Az legkisebb négyzetes módszer általánosan ismert módszer, a célja a hiba négyzetének a minimalizálása. Ha a mérések során kapott relatív gyakoriságok ν_ξ -k, akkor a valódi értéktől való eltérés az alábbi négyzetösszeggel egyezik meg (a mérések ügyes megválasztása miatt nincsenek kereszttagok):

$$\varphi_{LS}^m(D, \theta) = \sum_{\xi=1,2,3} \left(\nu_\xi - \frac{1 + \theta_\xi}{2} \right)^2$$

Ezt a kifejezést akarjuk minimalizálni, ezt pedig nyilván akkor minimális, ha a zárójelben lévő kifejezések rendre 0-k. Így ν_ξ -k ismeretében az alábbi becslésünk adódik θ értékére:

$$\Phi_{uc}^m(D) = \begin{bmatrix} 2\nu_1 - 1 \\ 2\nu_2 - 1 \\ 2\nu_3 - 1 \end{bmatrix} \quad (3)$$

Az előző becslésnek van egy elég jelentős hibája, mégpedig hogy adhat eredményként egy gömbön kívüli állapotot. Ez különösen akkor valószínű, ha a mérések száma kicsi, vagy a rendszer valódi állapota közel van a Bloch-gömb felszínéhez. Ezért bevezetjük a megszorított legkisebb négyzetes becslésünket:

$$\Phi_c^m(D) = \begin{cases} \Phi_{uc}^m(D) & \text{if } \|\Phi_{uc}^m(D)\|_2 \leq 1 \\ \frac{\Phi_{uc}^m(D)}{\|\Phi_{uc}^m(D)\|_2} & \text{if } \|\Phi_{uc}^m(D)\|_2 > 1 \end{cases} \quad (4)$$

Ez mindig értelmes eredményt fog adni, hiszen megvizsgálja az eredeti becslésünk eredményét, ha az gömbön belüli, akkor változatlanul hagyja, ha gömbön kívüli, akkor pedig leosztja a hosszával, azaz lényegében ráhúzza a gömb felszínére.

1.3. Eddigi irodalmi eredmények

Az eddigi kutatások e két becslés tulajdonságainak meghatározására vonatkoznak. Az alább felsorolt eredmények főként Magyar Attila és Petz Dénes munkája, részletesebb leírásuk megtalálható [9] és [10]-ben.

1.3.1. Az eredeti becslés tulajdonságai

Az eredeti (3) becslés sokkal egyszerűbben kezelhető, így sokkal több mindent is tudunk róla:

1. Tétel. *Az eredeti becslés, (3) torzítatlan.*

Bizonyítás:

Mivel az egyes paraméterekre egymástól függetlenül adunk becslést, elég belátni az egyik irányban, például θ_1 -re. És tudjuk, hogy $E(t(X)) = \sum_x t(x)p(x)$, így:

$$\begin{aligned} E\{\Phi_1^m(D)\} &= \sum_{m_1=0}^m \left(2\frac{m_1}{m} - 1\right) \cdot p_{m_1} = \frac{2}{m} \sum_{m_1=0}^m m_1 \cdot p_{m_1} - \sum_{m_1=0}^m p_{m_1} = \\ &= \frac{2}{m} E(X_1) - 1 = \frac{2}{m} \cdot m \frac{1 + \theta_1}{2} - 1 = \theta_1, \end{aligned}$$

mivel X_1 egy $(m, \frac{1+\theta_1}{2})$ paraméterű binomiális valószínűségi változó.

2. Tétel. *Az eredeti becslés, (3) kovariancia mátrixa:*

$$V^m(\theta) = \frac{1}{m} \begin{bmatrix} 1 - \theta_1^2 & 0 & 0 \\ 0 & 1 - \theta_2^2 & 0 \\ 0 & 0 & 1 - \theta_3^2 \end{bmatrix}$$

Bizonyítás:

A mérések függetlenségéből adódik, hogy $V^m(\theta) = \frac{1}{m} \cdot V^1(\theta)$, így elég egy mérésnek a kovariancia mátrixát kiszámolni.

$$V^1(\theta) = E \{ (\Phi^1(D) - \theta)(\Phi^1(D) - \theta)^T \},$$

kiírva:

$$V^1(\theta)_{i,j} = E \{ (\Phi_i^1(D) - \theta_i)(\Phi_j^1(D) - \theta_j) \}, \quad i, j = 1, 2, 3.$$

Az egyes irányú mérések függetlenségéből, ha $i \neq j$, akkor $V^1(\theta)_{i,j} = 0$.

Másrészt:

$$V^1(\theta)_{i,i} = E \{ (\Phi_i^1(D) - \theta_i)^2 \} = E \left\{ \left(2 \frac{m_i}{m} - 1 - \theta_i \right)^2 \right\},$$

ahol m_i csak 0 vagy 1 lehet, így

$$V^1(\theta)_{i,i} = (1 - \theta_i)^2 \cdot \frac{1}{2}(1 + \theta_i) + (1 + \theta_i)^2 \cdot \frac{1}{2}(1 - \theta_i) = 1 - \theta_i^2$$

3. Tétel. *Az eredeti becslés, (3) Fisher információs mátrixa:*

$$I^m(\theta) = m \begin{bmatrix} \frac{1}{1-\theta_1^2} & 0 & 0 \\ 0 & \frac{1}{1-\theta_2^2} & 0 \\ 0 & 0 & \frac{1}{1-\theta_3^2} \end{bmatrix}.$$

Bizonyítás:

A Fisher információs mátrix definíció szerint:

$$I^m(\theta)_{i,j} = E \left\{ \left(\frac{\partial}{\partial \theta_i} \ln f_{\theta,m}(D) \right) \cdot \left(\frac{\partial}{\partial \theta_j} \ln f_{\theta,m}(D) \right) \right\}, \quad i, j = 1, 2, 3,$$

ahol $f_{\theta,m}(D)$ az előbbieken már említett binomiális eloszlás sűrűségfüggvénye. Végigszámolva valóban a fenti formula adódik.

4. Tétel. *Az eredeti becslés, (3) hatásos.*

Bizonyítás: Triviális az előző két tétel és a Cramer-Rao egyenlőtlenség [3] ismeretében. Ugyanis a Fisher-információ mátrix inverze alsó becslése a kovariancia mátrixnak, esetünkben pedig pont teljesül is az egyenlőség

Azaz arra jutottunk, hogy az eredeti legkisebb négyzetes becslés hatásos, azaz az összes torzítatlan becslés közül minimális kovariancia mátrixú. Ez és a könnyű számolhatósága roppant vonzóvá teszi ezt a becslést, ám ne feledjük, hogy könnyen adhat értelmetlen eredményt (Bloch-gömbön kívül).

1.3.2. A megszorított becslés tulajdonságai

Megjegyzés: A megszorított becslés, (4) már nem torzítatlan.

Könnyű ellenpéldát találni, hiszen egy mérés esetén:

$$E \{ \Phi_i^1(D) \} = \frac{1}{\sqrt{3}} \cdot \frac{1}{2} (1 + \theta_i) - \frac{1}{\sqrt{3}} \cdot \frac{1}{2} (1 - \theta_i) = \frac{\theta_i}{\sqrt{3}}, \quad i = 1, 2, 3.$$

Azaz ha bármely $\theta_i \neq 0$, akkor a becslés nem torzítatlan.

5. Tétel. *A megszorított becslés, (4) aszimptotikusan torzítatlan, a konvergencia sebessége exponenciális a mérések számának függvényében.*

Bizonyítás

A tétel bizonyítása megtalálható [9]-ben. Ide csak azt jegyezném meg, hogy a bizonyítás a nagy eltérés tételek segítségével történik.

1.4. A probléma kitűzése

Mint azt láttuk, az eredeti legkisebb négyzetes módszer hatásos. Felmerül a kérdés, hogy a megszorított becslésnek mi lesz a kovarianciamátrixa.

Mint látni fogjuk ez egy nehéz kérdés, így arra adunk becslést, hogy a megszorítás által mennyivel változik meg a kovarianciamátrix. Pontosabban azt fogjuk vizsgálni, hogy mennyivel változik meg a teljes variancia, azaz a kovariancia mátrix nyoma. Könnyű átgondolni, hogy a megszorított kovariancia mátrixa kisebb mint az eredeti kovariancia mátrixa, így ha a nyomuk konvergál, akkor a kovariancia mátrixok elemenként is konvergálni fognak.

Az eredeti becslés teljes varianciája:

$$VAR_{uc} = \int_{-1}^1 \int_{-1}^1 \int_{-1}^1 |(x, y, z) - (\theta_x, \theta_y, \theta_z)|^2 \cdot \mathcal{L}^m(r, \theta) dx dy dz$$

A megszorított becslés teljes varianciája:

$$VAR_c = \int_{-1}^1 \int_{-1}^1 \int_{-1}^1 |(x', y', z') - (\theta_x, \theta_y, \theta_z)|^2 \cdot \mathcal{L}^m(r, \theta) dx dy dz$$

ahol

$$(x', y', z') = \begin{cases} (x, y, z) & \text{if } x^2 + y^2 + z^2 \leq 1 \\ \frac{(x, y, z)}{x^2 + y^2 + z^2} & \text{if } x^2 + y^2 + z^2 > 1 \end{cases}$$

Különbség csak abban az esetben van, ha a vektor a Bloch-gömbön kívülre esik, s így $VAR_{uc} > VAR_c$:

$$VAR_{uc} - VAR_c = \iiint (|(x, y, z) - (\theta_x, \theta_y, \theta_z)|^2 - |(x', y', z') - (\theta_x, \theta_y, \theta_z)|^2) \cdot \mathcal{L}^m(r, \theta) dx dy dz$$

$$VAR_{uc} - VAR_c = \int_{|r|>1} (d(r, \theta)^2 - d(r', \theta)^2) \cdot \mathcal{L}^m(r, \theta) dr \quad (5)$$

ahol az integrálandó tartomány a Bloch-gömbön kívüli rész, $d(\cdot, \cdot)$ két pont távolságát jelöli, valamint $r = (x, y, z)$ és $r' = (x', y', z')$.

2. Egy speciális eset, a nyomszerű állapot varianciájának a vizsgálata

A probléma nehézsége miatt először egy egyszerűbb eset következik: vizsgáljuk meg a maximálisan kevert állapotot (nyomszerű állapotot), azaz tegyük fel, hogy a rendszer állapota a Bloch-gömb középpontjával reprezentálható ($\theta = (0, 0, 0)$).

Ekkor (5) jelentősen kiegyeszerűsödik:

$$VAR_{uc} - VAR_c = \int_{|r|>1} (|r|^2 - |r'|^2) \cdot \mathcal{L}^m(r, \theta) dr = \int_{|r|>1} (|r|^2 - 1) \cdot \mathcal{L}^m(r, \theta) dr \quad (6)$$

ahol $\mathcal{L}^m(r, \theta)$ a binomiális eloszlás likelihood függvénye. Sajnos a diszkrét eloszlásnál nehezen kezelhető, hogy bent vagyunk-e a Bloch-gömbben, így normális eloszlással közelítjük. Ezt elvégezve már elvégezhetőek a kívánt számítások.

2.1. Áttérés binomiális eloszlásról normális eloszlásra

A binomiális likelihood függvénye az alábbi:

$$\mathcal{L}^m(D, \theta) = \prod_{\xi=x,y,z} \binom{m}{m_\xi} \cdot \left(\frac{1+\theta_\xi}{2}\right)^{m_\xi} \cdot \left(\frac{1-\theta_\xi}{2}\right)^{m-m_\xi} = \prod_{\xi=x,y,z} \binom{m}{m_\xi} \cdot \frac{1}{2^m}$$

Ha ezt csak az egyik irányba nézzük, akkor ez egy binomiális eloszlás, melynek paraméterei: $m, m_\xi, \frac{1+\theta_\xi}{2}$.

Ennek a várható értéke $m \cdot \frac{1+\theta_\xi}{2}$, szórásnégyzete pedig: $m \cdot \frac{1+\theta_\xi}{2} \cdot \frac{1-\theta_\xi}{2} = m \cdot \frac{1-\theta_\xi^2}{4}$.

Jelen esetben $\theta_\xi = 0$, így a fenti kifejezések az alábbi egyszerűbb formára hozhatók: várható érték = $\frac{m}{2}$, szórásnégyzet = $\frac{m}{4}$. Legyen a közelítő normális eloszlásunk, ugyanekkor várható értékű és szórású. Ekkor az ő likelihood függvényének egy komponense:

$$\mathbf{L}_\xi^m(D, \theta) = \sqrt{\frac{2}{\pi m}} \cdot \exp \left[-2 \frac{(m_\xi - \frac{m}{2})^2}{m} \right]$$

Az $m_\xi = \frac{\xi+1}{2} m$ változócserevel a következőt kapjuk:

$$\begin{aligned} \mathbf{L}^m(r, 0) &= \prod_{\xi=x,y,z} \frac{m}{2} \sqrt{\frac{2}{\pi m}} \cdot \exp \left(-\frac{m \xi^2}{2} \right) = \prod_{\xi=x,y,z} \sqrt{\frac{m}{2\pi}} \cdot \exp \left(-\frac{m \xi^2}{2} \right) = \\ &= \left(\sqrt{\frac{m}{2\pi}} \right)^3 \cdot \exp \left[-\frac{m}{2} (x^2 + y^2 + z^2) \right] = \left(\sqrt{\frac{m}{2\pi}} \right)^3 \cdot \exp \left(-\frac{m}{2} r^2 \right) \end{aligned}$$

2.2. Binomális és normális eloszlást összekötő lemma

Ahhoz, hogy megbecsülhessük (6)-t szükségünk van az alábbi lemmára:

1. Lemma. $\exists N$, ha $m > N$, akkor $\mathcal{L}^m(r, 0) \leq \mathbf{L}^m(r, 0)$, ahol $|r| > 1$.

Bizonyítás: A binomiális eloszlás likelihood függvénye, $\mathcal{L}^m(r, 0)$ egy diszkrét függvény, azonban kiterjeszthetjük folytonos függvényé (hiszen a binomiálisok léteznek nem egész számokra is).

A normális eloszlás sűrűségfüggvénye, $\mathbf{L}^m(r, 0)$ konstans, ha $|r|$ konstans, azaz elég maximalizálni $\mathcal{L}^m(r, 0)$ -t a gömbfelületen.

Ha a folytonos binomiális likelihood függvénye kisebb, mint $\mathbf{L}^m(r, 0)$, akkor a diszkrétre is igaz, hogy $\mathcal{L}^m(r, 0) \leq \mathbf{L}^m(r, 0)$, hiszen az egész helyeken ugyanannyi az értékük, ha pedig nő a távolság a középponttól, mindkettőnek csökken az értéke, azaz mivel a folytonos eloszlást tudjuk értelmezni a gömbfelület minden pontjában (a diszkrétet pedig lehet, hogy csak a gömbön kívül lehet értelmezni), így a folytonos maximuma legalább akkora mint a diszkrété. $\mathcal{L}^m(r, 0)$ -nak maximuma van, ha $\log \mathcal{L}^m(r, 0)$ -nak maximuma van, és $\log \mathcal{L}^m(r, 0) = \log \left[\binom{m}{\frac{x+1}{2}} \binom{m}{\frac{y+1}{2}} \binom{m}{\frac{z+1}{2}} \frac{m^3}{8^{m+1}} \right] = \log \binom{m}{\frac{x+1}{2}} + \log \binom{m}{\frac{y+1}{2}} + \log \binom{m}{\frac{z+1}{2}} + c(m)$. Valamint teljesül a $x^2 + y^2 + z^2 = 1$ feltétel.

A maximumot kiszámolhatjuk Lagrange-módszerrel. A Lagrange függvény az alábbi lesz:

$$L(x, y, z, \lambda) = \log \mathcal{L}^m(r, 0) - \lambda \cdot (x^2 + y^2 + z^2 - 1)$$

És tudjuk, hogy csak akkor lehet maximuma az eredeti feladatnak, $\frac{dL}{dx} = 0$, $\frac{dL}{dy} = 0$, $\frac{dL}{dz} = 0$, amiből:

$$\frac{dL}{dx} = \frac{d}{dx} \log \left(\binom{m}{\frac{x+1}{2}} \right) - 2\lambda x = 0$$

A szimmetria miatt feltehetjük, hogy x, y, z pozitívak.

$\frac{d}{dx} \log \left(\frac{x+1}{2} \cdot m \right)$ egy páratlan függvény és konkáv, ha x pozitív, így minden λ -ra az előző egyenletnek, csak egyetlen megoldása van.

A másik két egyenlőség (y -ra, és z -re) teljesen ugyanilyen, így nekik is ugyanez az egy pozitív megoldásuk lesz, azaz tudjuk hogy $x = y = z$, ha $\mathcal{L}^m(r, 0)$ extremális, azaz tudjuk a megfelelő helyettesítő értéket, ha beírjuk a képletekbe, és a mérésszámmal tartva a végtelenbe, megkapjuk hogy a normális eloszlás aszimptotikusan lassabban tart nullához (valójában a hányadosuk nullához tart), amiből már következik a lemma állítása.

Megjegyzés: Számítógépes számításokból kiderül, hogy $N = 18$ már jó.

2.3. Becslés a normális eloszlás segítségével

Írjuk fel még egyszer (6)-t:

$$VAR_{uc} - VAR_c = \int_{|r|>1} (|r|^2 - 1) \cdot \mathcal{L}^m(r, 0) dr$$

Tudjuk, hogy $(|r|^2 - 1) \geq 0$, így az 1. Lemmából:

$$VAR_{uc} - VAR_c \leq \int_{|r|>1} (|r|^2 - 1) \cdot \mathbf{L}^m(r, 0) dr = \int_{|r|>1} |r|^2 \cdot \mathbf{L}^m(r, 0) dr - \int_{|r|>1} \mathbf{L}^m(r, 0) dr = A - B.$$

Az $\mathbf{L}^m(r, 0)$ csak a Bloch-vektor hosszától függ. Így logikusnak tűnik, hogy kiszámítjuk a sűrűségfüggvényét annak, hogy a becsült vektor r hosszú (rajta van az r sugarú gömbön).

$$\mathbf{P}_r = 4 r^2 \pi \cdot \mathbf{L}^m(r, 0) = 4 r^2 \pi \left(\sqrt{\frac{m}{2\pi}} \right)^3 \cdot \exp \left(-\frac{m}{2} r^2 \right) = \sqrt{\frac{2 m^3}{\pi}} r^2 \exp \left(-\frac{m}{2} r^2 \right)$$

Amiből:

$$A = \int_1^{\infty} r^2 \cdot \mathbf{P}_r dr, \text{ and } B = \int_1^{\infty} \mathbf{P}_r dr$$

A teljes variancia $\frac{3}{m}$, így

$$A = \frac{3}{m} - \int_0^1 \mathbf{P}_r \cdot r^2 dr = \frac{3}{m} - \int_0^1 \sqrt{\frac{2 m^3}{\pi}} r^4 \exp \left(-\frac{m}{2} r^2 \right) dr$$

másrészt:

$$B = 1 - \int_0^1 \mathbf{P}_r dr = 1 - \int_0^1 \sqrt{\frac{2 m^3}{\pi}} r^2 \exp \left(-\frac{m}{2} r^2 \right) dr$$

2.4. További számítások

Parciális integrálással:

$$\int r^2 \cdot \exp(c r^2) dr = \frac{1}{3} r^3 \exp(c r^2) - \int \frac{1}{3} r^3 \cdot 2cr \exp(c r^2) dr = \frac{1}{3} r^3 \exp(c r^2) - \frac{2c}{3} \int r^4 \exp(c r^2) dr$$

és

$$\int 1 \cdot \exp(c r^2) dr = r \exp(c r^2) - \int r \cdot 2cr \exp(c r^2) dr = r \exp(c r^2) - 2c \int r^2 \exp(c r^2) dr$$

Amiből:

$$\int r^2 \exp(c r^2) dr = \frac{1}{2c} \left(r \exp(c r^2) - \int \exp(c r^2) dr \right)$$

és

$$\begin{aligned} \int r^4 \exp(c r^2) dr &= \frac{3}{2c} \left(\frac{1}{3} r^3 \exp(c r^2) - \int r^2 \cdot \exp(c r^2) dr \right) = \\ &= \frac{3}{2c} \left(\frac{1}{3} r^3 \exp(c r^2) - \frac{1}{2c} \left(r \exp(c r^2) - \int \exp(c r^2) dr \right) \right) \\ &= \frac{1}{2c} r^3 \exp(c r^2) - \frac{3}{4c^2} r \exp(c r^2) + \frac{3}{4c^2} \int \exp(c r^2) dr \end{aligned}$$

Ezt felhasználva: esetünkben $c = -\frac{m}{2}$, az integrálás tartománya legyen $[0, 1]$, és az utolsó integrálban alkalmazva a $t := r \sqrt{m}$ helyettesítést, az alábbi képleteket kapjuk (ahol $\varphi(t)$ a standard normális eloszlás sűrűségfüggvénye):

$$A = \frac{3}{m} - \left[-3\sqrt{\frac{2}{m\pi}} \exp\left(-\frac{m}{2}\right) - \sqrt{\frac{2m}{\pi}} \exp\left(-\frac{m}{2}\right) + \frac{3}{m} \int_{-\sqrt{m}}^{\sqrt{m}} \varphi(t) dt \right]$$

és

$$B = 1 + \sqrt{\frac{2m}{\pi}} \exp\left(-\frac{m}{2}\right) - \int_{-\sqrt{m}}^{\sqrt{m}} \varphi(t) dt$$

Így a keresett mennyiség:

$$A - B = 3\sqrt{\frac{2}{m\pi}} \exp\left(-\frac{m}{2}\right) + \frac{3}{m} \left(1 - \int_{-\sqrt{m}}^{\sqrt{m}} \varphi(t) dt \right) - \left(1 - \int_{-\sqrt{m}}^{\sqrt{m}} \varphi(t) dt \right)$$

Használjuk fel az alábbi egyenlőtlenséget, ahol X standard normális eloszlású (könnyű belátni, elég megvizsgálni mindegyik kifejezés deriváltját):

$$\frac{1}{\sqrt{2\pi}} \left(\frac{1}{x} - \frac{1}{x^3} \right) e^{-\frac{x^2}{2}} < \mathbb{P}(X > x) < \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-\frac{x^2}{2}}$$

és szimmetriából:

$$\frac{2}{\sqrt{2\pi}} \left(\frac{1}{x} - \frac{1}{x^3} \right) e^{-\frac{x^2}{2}} < \mathbb{P}(|X| > x) = 1 - \int_{-x}^x \varphi(t) dt < \frac{2}{\sqrt{2\pi}} \frac{1}{x} e^{-\frac{x^2}{2}}$$

$A - B$ -ben a második tag elhanyagolható a többihez képest, így azt kapjuk, hogy

$$A - B \sim 3\sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{m}} \exp\left(-\frac{m}{2}\right) - \frac{2}{\sqrt{2\pi}} \frac{1}{\sqrt{m}} \exp\left(-\frac{m}{2}\right) = \left(3\sqrt{\frac{2}{\pi}} - \frac{2}{\sqrt{2\pi}}\right) \frac{1}{\sqrt{m}} \exp\left(-\frac{m}{2}\right)$$

$$A - B \sim \frac{2\sqrt{2}}{\sqrt{\pi}} \frac{1}{\sqrt{m}} \exp\left(-\frac{m}{2}\right)$$

2.5. A konvergencia sebessége szimmetrikus esetben

Mindezek tükrében kimondhatjuk a következő tételt:

6. Tétel.

$$VAR_{uc} - VAR_c \leq c \cdot \frac{1}{\sqrt{m}} \exp\left(-\frac{m}{2}\right)$$

Azt kaptuk amit vártunk: teljesül az aszimptotikus torzítatlanság, és a konvergencia sebessége nagyon gyors, több mint exponenciális.

3. A variancia megváltozása általános esetben

A nagy eltérés tételek arra mondanak becslést, hogy az állapotok egy halmazába mekkora valószínűséggel lesz a becslésünk.

$$VAR_{uc} - VAR_c = \int_{|r|>1} (|r - \theta|^2 - |r' - \theta|^2) \cdot \mathcal{L}^m(r, \theta) dr$$

Ebben a kifejezésben $|r - \theta|^2 - |r' - \theta|^2$ korlátos: $|r - \theta|^2 - |r' - \theta|^2 = (|r - \theta| + |r' - \theta|)(|r - \theta| - |r' - \theta|)$.

$$|r - \theta| + |r' - \theta| < 2 + (1 + \sqrt{3}) = 3 + \sqrt{3}, \text{ and } |r - \theta| - |r' - \theta| < \sqrt{3} - 1$$

Így $|r - \theta|^2 - |r' - \theta|^2 < (3 + \sqrt{3})(\sqrt{3} - 1) = 2\sqrt{3} =: K$, és eléri, ha $\theta = (-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}})$ és $r = (1, 1, 1)$.

Amiből

$$VAR_{uc} - VAR_c \leq K \cdot \int_{|r|>1} \mathcal{L}^m(r, \theta) dr = K \cdot \mathcal{P}^m(|r| > 1) \quad (7)$$

Tehát ha van egy jó becslésünk arra, hogy mekkora valószínűséggel esik a becslés a gömbön kívülre, akkor kapunk egy becslést a konvergencia sebességére. A nagy eltérés tételek pont ebben segítenek.

3.1. Nagy eltérés tétel alkalmazása 1 dimenzióban

7. Tétel (Cramer,[2]). *Legyen (X_i) független azonos eloszlású valós értékű valószínűségi változó, melyre*

$$\varphi(t) = \mathbb{E} e^{tX_1} < \infty$$

Legyen $S_n = \sum_{i=1}^n X_i$. Ekkor minden $a > \mathbb{E}X_1$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq an) = -I(a),$$

ahol

$$I(z) = \sup_{t \in \mathbb{R}} [zt - \log \varphi(t)].$$

Esetünkben az X_i változók a kísérletek kimenetele lesz a megfelelő ξ irányokban, ami binomiális eloszlást eredményez.

Speciálisan: $\mathbb{P}(X_i = 1) = p$, és $\mathbb{P}(X_i = -1) = q$, ahol $q = 1 - p$, és $p = \frac{1+\theta\xi}{2}$.
Akkor $\varphi(t)$ az alábbival lesz egyenlő:

$$\varphi(t) = \mathbb{E} e^{tX_1} = \mathbb{P}(X_1 = 1) \cdot e^{t \cdot 1} + \mathbb{P}(X_1 = -1) \cdot e^{t \cdot (-1)} = pe^t + qe^{-t}.$$

Ebből már kiszámolhatjuk a ráta függvényt($I(x)$ -t):

$$I(x) = \sup_{t \in \mathbb{R}} [xt - \log \varphi(t)] = \sup_{t \in \mathbb{R}} [xt - \log (pe^t + qe^{-t})].$$

Ha lederiváljuk a belső kifejezést megkapjuk a stacionárius pontokat:

$$(xt - \log(pe^t + qe^{-t}))' = x - \frac{pe^t - qe^{-t}}{pe^t + qe^{-t}} = 0$$

Ebből kifejezzük t -t (itt lesz maximális):

$$t = \frac{1}{2} \log \frac{q(1+x)}{p(1-x)}$$

Behelyettesítve a ráta függvénybe:

$$I(x) = \sup_{t \in \mathbb{R}} [xt - \log(pe^t + qe^{-t})] = x \cdot \frac{1}{2} \log \frac{q(1+x)}{p(1-x)} - \log \left(p \cdot \sqrt{\frac{q(1+x)}{p(1-x)}} + q \sqrt{\frac{p(1-x)}{q(1+x)}} \right).$$

Átalakítások után:

$$I(x) = \frac{x+1}{2} \log \left(\frac{x+1}{2p} \right) + \frac{1-x}{2} \log \left(\frac{1-x}{2q} \right) = S \left(\frac{x+1}{2} \parallel p \right) \quad (8)$$

Azaz megegyezik $\frac{x+1}{2}$ és p relatív entropiájával. A minimumot az $x = p - q = 2p - 1$ -ben veszi fel, így ha Taylor-sorba fejtjük e pont körül, akkor ezt kapjuk:

$$I(x) \sim \frac{(x - (p - q))^2}{8pq} \quad (9)$$

Legyen $a = \mathbb{E}X_1 + \lambda = m + \lambda$, ekkor a 7. Tételből:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq (m + \lambda)n) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left(\frac{S_n}{n} - m \geq \lambda \right) \sim -\frac{((m + \lambda) - (p - q))^2}{8pq},$$

és

$$m = \mathbb{E}X_1 = \mathbb{P}(X_1 = 1) \cdot 1 + \mathbb{P}(X_1 = -1) \cdot (-1) = p - q.$$

Így

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left(\frac{S_n}{n} - m \geq \lambda \right) \sim -\frac{\lambda^2}{8pq} = -\frac{\lambda^2}{2(1 - \theta_\xi^2)}$$

amiből

$$\mathbb{P} \left(\frac{S_n}{n} - m \geq \lambda \right) \leq c \cdot \exp \left[-n \frac{\lambda^2}{2(1 - \theta_\xi^2)} \right]$$

Ahol $\frac{S_n}{n}$ a legkisebb négyzetes becslés n mérés után, és m a valódi állapota a rendszernek. Azaz ez azt mondja, hogy annak a valószínűsége, hogy a valódi állapot és a becslés különbsége nagy (nagyobb λ -nál) exponenciálisan csökken, ha a mérések száma növekszik.

3.2. Nagy eltérés tétel alkalmazása általánosan

8. Tétel (Cramer, általánosan). X_i nem feltétlenül valós értékű független azonos eloszlású valószínűségi változó, és F egy zárt halmaz, akkor

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left(\frac{S_n}{n} \in F \right) = -\inf \{ I^*(u) : u \in F \}. \quad (10)$$

A mi esetünkben F a Bloch-gömbön kívüli tartomány lesz, amibe beletartozik a felszín is, hiszen F zárt.

A három dimenziós eloszlásunk független egy-dimenziós $\{-1, 1\}$ bináris eloszlásokból áll, melyek ráta függvényét az előbbi fejezetben kiszámoltuk. A legkisebb négyzetes becslésünk ezen vektorértékű valószínűségi változó empirikus várható értéke, azaz $\frac{S_n}{n}$. A függetlenség miatt a ráta függvény az egy-dimenziós ráta függvények összege lesz:

$$I^*(x, y, z) = I(x) + I(y) + I(z)$$

ami egy nemnegatív függvény, és $I^*(x, y, z) = 0$ pontosan akkor, ha $(x, y, z) = (\theta_x, \theta_y, \theta_z)$. Így ha $\|\theta\|_2 \neq 1$ akkor az összes (x, y, z) -re, ami a Bloch-gömbön kívül van $I^*(x, y, z) \geq \delta > 0$.

3.3. A konvergencia sebessége általános esetben

Az előbbi és (7) fényében kimondhatjuk a következő tételt:

9. Tétel. *Ha a valódi állapotunk nem tiszta, akkor:*

$$VAR_{uc} - VAR_c \leq C \exp(-K(\theta) \cdot m),$$

ahol K egy θ -tól függő konstans.

Azaz minden kevert állapotra beláttuk a konvergenciát, sőt azt is, hogy exponenciális a sebessége.

3.4. Speciális eset: $\theta_1 = \theta_2 = \theta_3$

Konkrétan a Bloch gömbön kívül keresett infimum, azaz (10) meghatározása nem végezhető el analitikusan az általánosság megszorítása nélkül, azonban (9) alapján adhatunk egy heurisztikus becslést, hiszen a Taylor-sorának első két tagja szerint, a ráta függvény egy irányban a valódi állapot és a vizsgált állapot különbségének négyzete osztva egy konstanssal. Ezeket összeadva azt kapjuk, hogy az ekvipotens felület egy ellipszoid, azaz abban a pontban veszi fel az infimumot, ahová a beírt ellipszoid érinti a gömb felszínét.

Ha $\theta_1 = \theta_2 = \theta_3 > 0$, akkor mindenütt ugyanaz a konstans, így az ellipszoid gömbbe megy át, és ez nyilvánvalóan az $\frac{1}{\sqrt{3}}(1, 1, 1)^T$ -ben érinti a Bloch-gömböt. Mint látni fogjuk ez utóbbi becslés meg is felel a valóságnak:

10. Tétel. *Ha $\theta_1 = \theta_2 = \theta_3 > 0$, akkor a ráta függvény az $\frac{1}{\sqrt{3}}(1, 1, 1)^T$ -ben veszi fel az infimumát a Bloch-gömbön kívül.*

Bizonyítás

Először is belátjuk, hogy ha $(x, y, z)^T$ -ben veszi fel I^* az infimumát, akkor x, y, z nemnegatívok. Ugyanis tegyük fel, hogy legalább az egyik negatív, például x , akkor $I(x) > I(-x)$, hiszen a ráta függvény szimmetrikus a valódi állapot értékére, ami jelen esetben pozitív, és így

$$I^*(x, y, z) = I(x) + I(y) + I(z) > I(-x) + I(y) + I(z) = I^*(-x, y, z),$$

azaz $(x, y, z)^T$ -ben nem lehet I^* -nak infimuma.

(8) alapján tudjuk, hogy

$$I(x) = \frac{x+1}{2} \log\left(\frac{x+1}{2p}\right) + \frac{1-x}{2} \log\left(\frac{1-x}{2q}\right), \quad \text{ahol } p = \frac{1+\theta_i}{2}, \quad q = 1-p$$

Mivel $\theta_1 = \theta_2 = \theta_3$, így y és z esetén is ugyanaz lesz p és q értéke.

Minél közelebb megyünk a gömb felszínéhez annál kisebb lesz $I^*(u)$, hiszen elérhető, hogy csak az egyik koordinátát változtatjuk úgy, hogy közelebb kerüljön a valódi állapothoz. Így az infimum mindenképp a gömb felületén fog felvétetni, és Lagrange-módszerrel kiszámíthatjuk az infimumot.

A Lagrange függvény az alábbi lesz:

$$L(x, y, z) = I(x) + I(y) + I(z) - \lambda \cdot (x^2 + y^2 + z^2 - 1)$$

És tudjuk, hogy csak akkor lehet maximuma az eredeti feladatnak, ha $\frac{dL}{dx} = 0$, $\frac{dL}{dy} = 0$, $\frac{dL}{dz} = 0$,
amiből:

$$I'(x) = 2\lambda x$$

Vizsgáljuk meg $f(x) = I'(x)$ -et!

$$f(x) = \frac{1}{2} \log\left(\frac{x+1}{2p}\right) - \frac{1}{2} \log\left(\frac{1-x}{2q}\right)$$

$$f(0) = \frac{1}{2} \log\left(\frac{1}{2p}\right) - \frac{1}{2} \log\left(\frac{1}{2q}\right) = \frac{1}{2} [\log(2q) - \log(2p)] < 0$$

$$f'(x) = \frac{1}{2} \left(\frac{1}{x+1} + \frac{1}{1-x} \right) > 0$$

$$f''(x) = \frac{1}{2} \left(\frac{1}{(1-x)^2} - \frac{1}{(x+1)^2} \right) > 0, \quad \text{ha } x > 0$$

Azaz ha $x > 0$, akkor $f(x)$ szigorúan monoton nő és konvex, valamint a 0-ban felvett értéke negatív, így bármely nullán átmenő egyenes legfeljebb egy pontban metszheti, azaz legfeljebb egy pozitív megoldása van.

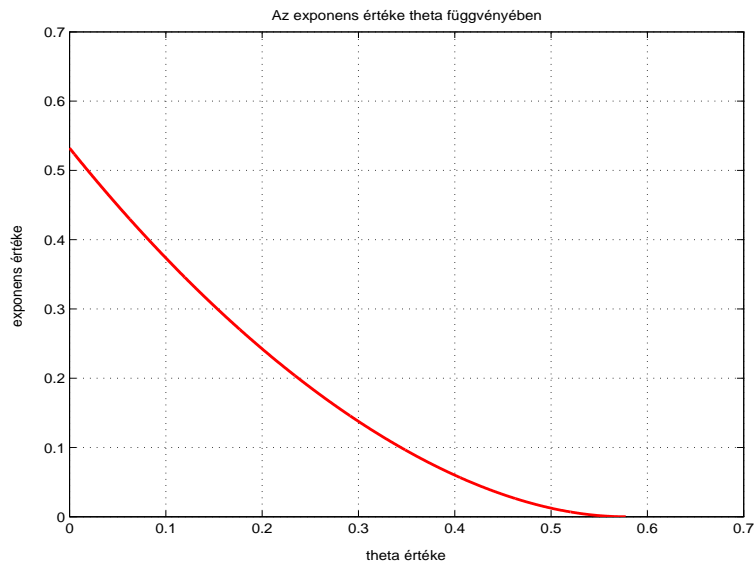
Mivel tudjuk, hogy folytonos függvény kompakt halmazon felveszi infimumát, így létezik olyan $(x, y, z)^T$ pontja a felszínnek, ahol I^* minimális, és mivel ugyanaz az egyenlet írható fel y és z -re, mint azt az előbb tettük x esetén, így mindegyiknek pontosan egy megoldása lesz, még hozzá ugyanaz, azaz $x = y = z$, és mivel a gömbfelszínen vagyunk, ez egyenlő lesz $\frac{1}{\sqrt{3}}$ -mal.

A 10. Tételt felhasználva már könnyen kiszámíthatjuk (10)-t:

$$K(\theta) = 3 \cdot I \left(\frac{1}{\sqrt{3}} \right) = 3 \cdot \left[\frac{\frac{1}{\sqrt{3}} + 1}{2} \log \left(\frac{\frac{1}{\sqrt{3}} + 1}{1 + \theta} \right) + \frac{1 - \frac{1}{\sqrt{3}}}{2} \log \left(\frac{1 - \frac{1}{\sqrt{3}}}{1 - \theta} \right) \right], \quad (11)$$

ahol $\theta = \theta_1 = \theta_2 = \theta_3 > 0$.

Ez így nem igazán mond sokat, de az 1. ábrán jól látszik, hogy 0-ban az értéke több mint a 0.5, valamint az, hogy szigorúan monoton csökken, ahogy tartunk a gömbfelszínhez, ahol is 0 lesz a függvény értéke, azaz a felszínen, már nem látjuk, az exponenciális konvergenciát.

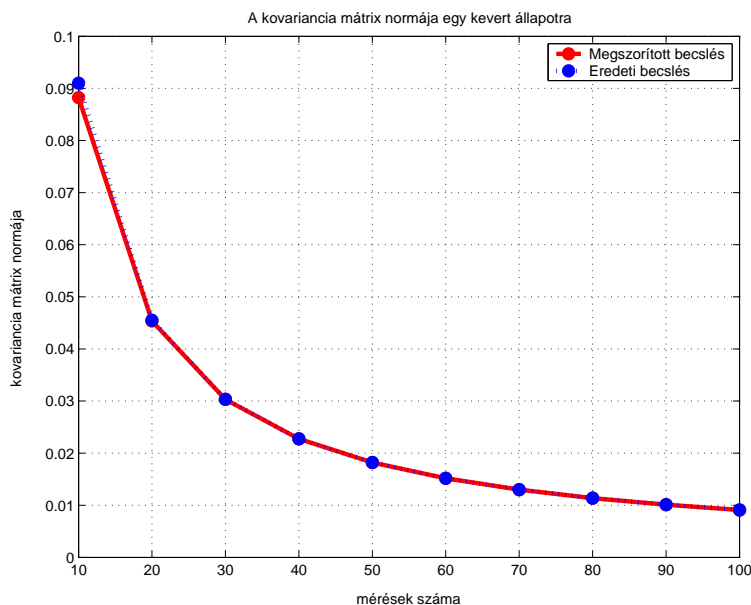


1. ábra. Az infimum értéke θ függvényében

4. Szimulációk

Ahogy az már a bevezetésben meg van fogalmazva, kaptunk az eltérésre felső becsléseket, és a most következő részben szeretnénk megvizsgálni, hogy mennyire tér el a becslésünk a valóságtól. Ennek vizsgálatához MATLAB [7] scriptek lettek írva, melyek kiszámolják a pontos szórásokat, valamint kirajzolják a kívánt függvényeket.

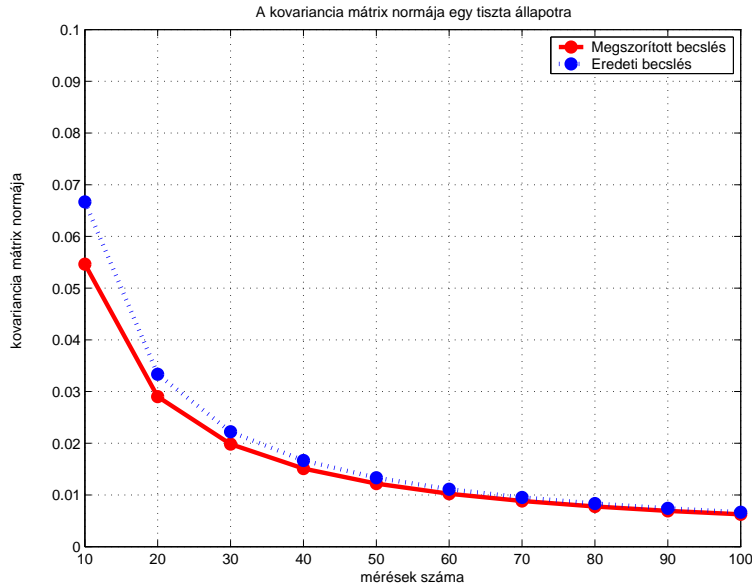
4.1. Az varianciáról általában



2. ábra. A kovariancia mátrix normája a $[0.3, -0.3, 0.3]^T$ kevert állapotra

Azt szeretnénk megtudni, hogy teljesül-e a konvergencia, valamint hogy mitől és hogyan függ a sebessége. Az elég nyilvánvalónak tűnik, hogy az eltérés függ a mérések számától. Az előző fejezetben a felső becslésünk függött a ráta függvényről, méghozzá oly módon, hogy minél inkább a gömb szélén vagyunk annál lassabb lesz a konvergencia. Így az a várakozásunk, hogy a valódi állapot kevertsége is összefüggésben van-e a konvergencia sebességével. Mivel a kovariancia mátrixot nehéz ábrázolni, így a legnagyobb sajátértékét ábrázoltuk a mérésszám függvényében (mivel olyan állapotokat vizsgálunk ahol a koordináták azonos nagyságúak, így a kovariancia mátrix főátlójában minden elem megegyezik a legnagyobb sajátértékkel). Az ábrán szaggatottal az eredeti, folytonossal a megszorított becslést jelöltük.

Az 2. ábrán egy kevert állapot látható. Jól megfigyelhető, hogy a variancia fordítottan arányos a mérésszámmal, valamint az is, hogy a két becslés közötti különbség már kis mérésszám esetén is (20-30 mérés) gyakorlatilag megegyezik. Ellenben a 3. ábrán egy tiszta állapot látható, és még akár 100 mérésnél is észrevehető a különbség. Az is látszik,



3. ábra. A kovariancia mátrix normája a $[0.5774, -0.5774, 0.5774]^T$ tiszta állapotra

hogy az eredeti szórása a nagyobb, hiszen a megszorításnál lecsökkennek a távolságok, így a szórás is.

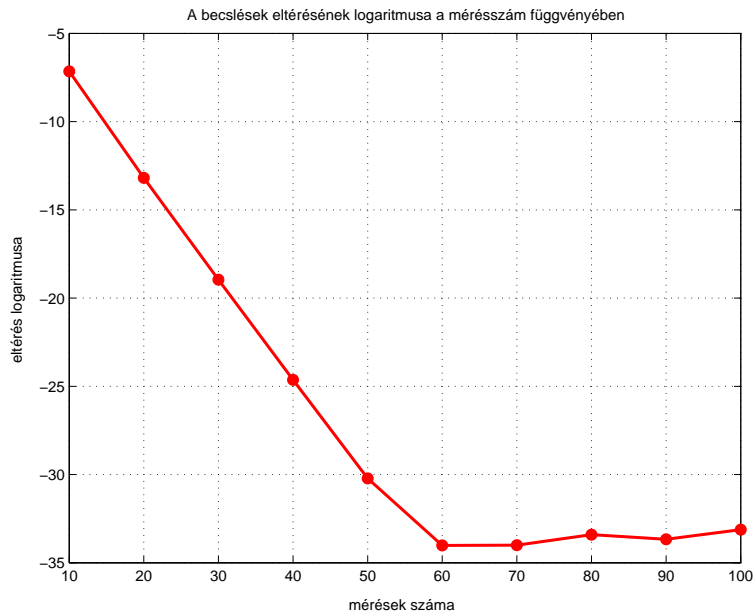
4.2. A konvergencia sebessége

Valójában nem a két varianciára vagyunk kíváncsiak, hanem a különbségükre, úgyhogy célszerűbb azt megvizsgálni. Ám azt tudjuk, hogy a konvergencia kevert állapotokra legalább exponenciális, így ha a különbséget ábrázolnánk pillanatok alatt 0 közelébe érnénk, és nem tudnánk összehasonlítani az egyes értékeket. Ezért ábrázoljuk a különbség logaritmusát. Ezzel elérjük, hogy nagy intervallumon összehasonlítható adatokat kapjunk, valamint mivel úgy gondoljuk, hogy a konvergencia exponenciális, így a függvényünk meredeksége meghatározza a valódi exponens értékét.

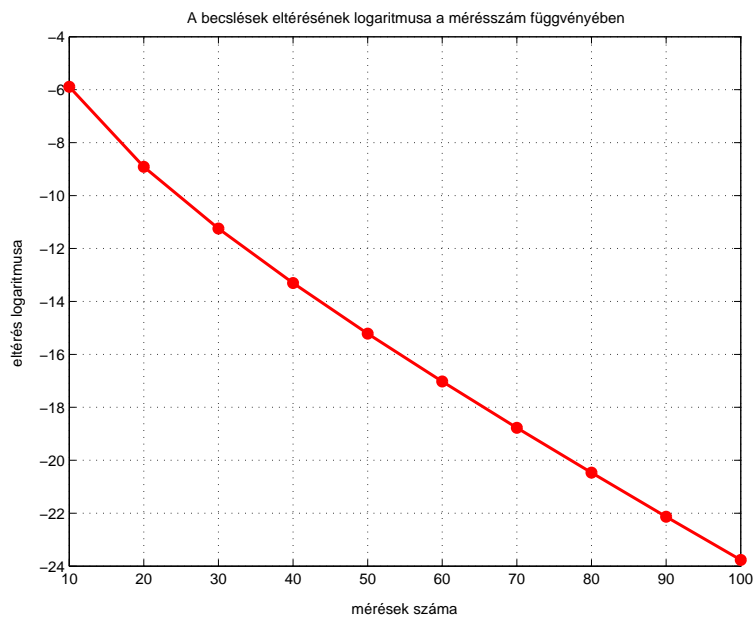
Azonban a 4. ábrán rögtön megfigyelhetünk egy furcsaságot, a függvény kezdetben szabályosan viselkedik, majd 50-60 mérés között megváltozik, majd teljesen véletlenszerű értékeket vesz fel. Ez abból adódik, hogy a konvergencia olyan gyors, hogy már 60 mérésnél kisebb az eltérés, mint a számítógép pontossága, így onnantól kezdve valóban véletlenszerűek az adatok. Azonban az első részén jól látható a linearitás, azaz tényleg exponenciális a konvergencia sebessége.

Ha távolodunk a Bloch-gömb középpontjától, már nem futunk bele ilyen számítási pontatlanságba. Egy tipikus helyzetet mutat az 5. ábra: a meredekség az elején változik, de egy idő után beáll egy határozott irányba.

A kérdés, hogy mi történik tiszta állapot esetén. Mind az a 6. ábrán látható, a

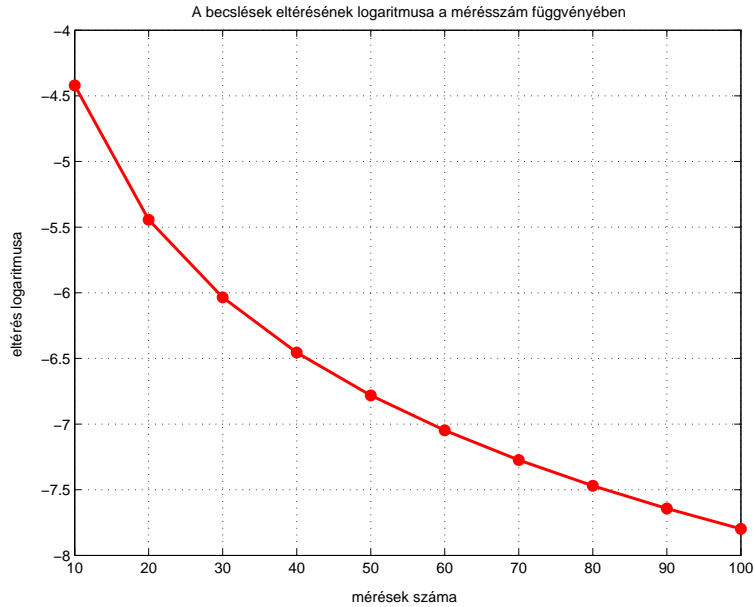


4. ábra. Az eltérés logaritmus a $[0, 0, 0]^T$ állapotra



5. ábra. Az eltérés logaritmus a $[0.3, -0.3, 0.3]^T$ kevert állapotra

meredekség egyre csökken, de nem tudjuk megmondani, hogy a nullához tart, vagy ő is megállapodik egy irányban, mint az az előző ábránál látható volt. Talán megállapítható lenne, ha nagyobb mérőszámig vizsgálnánk, azonban a szimulációk nagy futásideje ezt nem teszi lehetővé, azaz még egy kicsivel lehetne bővíteni az intervallumunkat, azonban nagyságrendileg nem, így nem vagyunk kíségetve.



6. ábra. Az eltérés logaritmusai a $[0.5774, -0.5774, 0.5774]^T$ tiszta állapotra

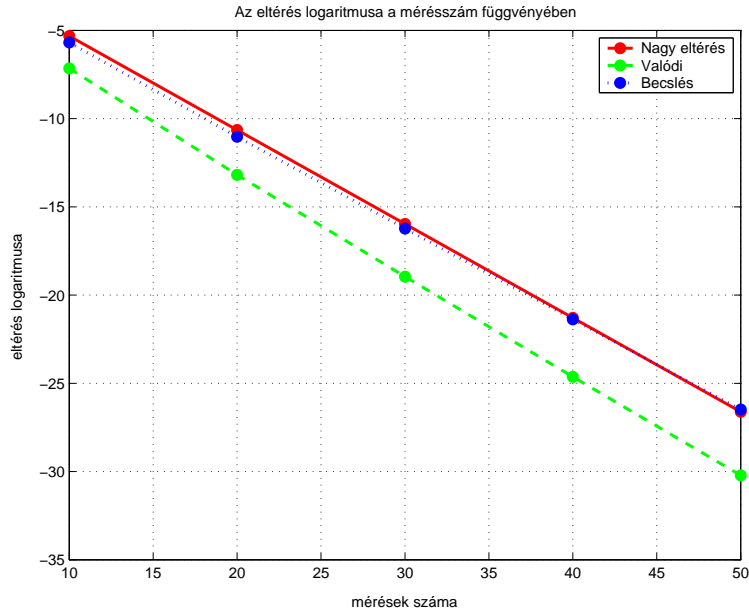
4.3. A becslések pontossága

Hasonlítsuk össze az előző két fejezet eredményei révén kapott becsléseket a valósággal nyomszerű állapot esetén. Ezt hasonlóképp végezzük mint az előbb, az eltérés logaritmusát rajzoltatjuk ki. Ha a 6. Tételben lévő becslést használjuk, akkor c helyére írjuk be az asszimptotikus értéket: $\frac{2\sqrt{2}}{\sqrt{\pi}}$ (nagy valószínűséggel ez is jó lesz, hiszen a binomiálist jóval felülbecsültük a normálissal). A 9. Tételben lévő becslésben szintén nem tudjuk C értékét, jelen esetben 1-nek vesszük, és csak a meredekséget figyeljük, ami (11) szerint körülbelül -0.53232133 . Tudjuk, hogy szimmetrikus állapotnál 60 mérésnél már hibás eredményt kapunk, így csak 50 mérésig rajzoljuk ki az értékeket.

Tudjuk, hogy a 2. fejezet becslése esetén az exponens -0.5 -höz tart. A nagy eltérések esetén, mint azt már említettük körülbelül -0.5323 , míg a valóságban a szimulációból az adódott, hogy nagyobb mint -0.56 , azaz kimondható, hogy mindkét becslés elég jól közelíti a sebesség nagyságrendjét, azaz az ábrán a meredekséget.

5. Összegzés

A dolgozat célja az eredeti legkisebb négyzetes becslés szórásának összehasonlítása a megszorított becslésével egy qubit esetén. A motivációt az adja, hogy a gyakorlati életben az eredeti becslést használják, mivel lényegesen egyszerűbben kezelhető, torzítatlan és hatásos. Azonban értelmetlen megoldást is eredményezhet, így ha matematikailag korrektek akarunk lenni a megszorított becslést kellene alkalmazni. Azt tudjuk, hogy a



7. ábra. Konvergencia sebességének összehasonlítása a $[0, 0, 0]^T$ állapotra

becslés aszimptotikusan torzítatlan, azaz a várható értéke nagy mérésszám esetén tart a valódi értékhez, sőt kevert állapotra a konvergencia sebessége exponenciális. Azt várjuk, hogy valami hasonló igaz a kovariancia mátrixra is.

A második fejezetben egy speciális állapotra, $\theta = 0$ -ra beláttuk, hogy a konvergencia exponenciális. Mivel a probléma nehéz, első megközelítésnél feltettünk bizonyos szimmetriákat, hogy a probléma könnyebben kezelhető legyen, és így már a valószínűségi számítás elemi eszközei is elegendőnek bizonyulnak.

A harmadik fejezetben általános kevert állapotra beláttuk az exponenciális konvergenciát. Ehhez a nagy eltérés tételek témakörét hívtuk segítségül. Ezenkívül a felső becslés exponensének explicit értékét is kiszámoltuk egy egyenes mentén. Ez általánosabb eredmény az előzőnél két értelemben is, hiszen nem csak egy konkrét állapotra láttuk be, azonban a felső becslésünk nem lesz annyira egzakt.

A negyedik fejezetben szimulációk segítségével vizsgáltuk a kovariancia mátrixok eltérést. Azt kaptuk, amit vártunk, hogy míg kevert állapotnál a konvergencia gyors, és ahogy haladunk ki a Bloch-gömb pereme felé a konvergencia egyre lassabb, hasonlóan (11) eredményéhez (1. ábra). Másrészt összehasonlítva a valódi eltérést a leírt felső becsléseinkkel azt kaptuk, hogy a becslésünk nem pontos, azonban nagyságrendileg közel állnak a valósághoz.

További kutatás tárgya lehet a tiszta állapot vizsgálata, ugyanis a fenti bizonyítás kis módosításával be lehet látni, hogy a konvergencia fent áll, azonban az hogy ez exponenciális-e avagy sem nem tisztázott (lásd 6. ábra). Ezenkívül a 3.2 rész bizonyítása teljesen hasonlóképp végigvihető többdimenziós esetben is, bár ott már egyáltalán nem

triviális a nem negatívitási feltételből fakadó határoló alakzat.

Hivatkozások

- [1] G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, Quantum tomography, *Advances in Imaging and Electron Physics*, Vol. 128, p. 205-308, 2003
- [2] F. Den Hollander, *Large Deviations*, Fields Institute Monograph, American Mathematical Society, Providence, 2000
- [3] Bolla M., Krámli A., *Statisztikai következtetések elmélete*, Typotex Kiadó, Budapest, 2005
- [4] Th. Baier, K. M. Hangos, A. Magyar, D. Petz, Comparasion of some methods of quantum state estimation, *arXiv*, quant-ph/0511263v1, 2005
- [5] J. Rehacek, B.-G. Englert, D Kaszlikowski, Minimal qubit tomography, *Physical Review A*, 70:052321, 2004
- [6] A. Magyar, D. Petz, K.M. Hangos, Bayesian qubit state estimation, *IFAC Conference on Identification and System Parameter Estimation*, Proceedings CD, p. 949-954, 2005
- [7] Mathworks Inc, MATLAB Software System, <http://www.mathworks.com/>, 2001
- [8] R. L. Kosut, I. Walmsley, H. Rabitz, Optimal experiment design for quantum state and process tomography and Hamiltonian parameter estimation, *arXiv*, quant-ph/0411093, 2004
- [9] D. Petz, K. M. Hangos, A. Magyar, Point Estimation of States of Finite Quantum Systems, *arXiv*, quant-ph/0610124v1, 2006
- [10] D. Petz, K. M. Hangos, A. Magyar, L. Ruppert, State estimation of N-level quantum systems, A Rendszer- és Irányításeleméleti Laboratórium SCL-007/2006 számú kutatási riportja, Budapest, MTA SZTAKI, 2006