



UNIVERSITATEA BABEȘ-BOLYAI
BABEȘ-BOLYAI TUDOMÁNYEGYETEM
BABEȘ-BOLYAI UNIVERSITÄT
BABEȘ-BOLYAI UNIVERSITY

FACULTATEA DE FIZICĂ
Str. Mihail Kogălniceanu nr.1
Cluj-Napoca, RO-400084
Tel: +4(0)264-405300 | FAX: +4(0)264-591906
secretariat.phys@ubbcluj.ro
www.phys.ubbcluj.ro



LUCRARE DE LICENȚĂ

Coordonator științific
Prof. Dr. Néda Zoltán

Absolvent
Roth Apor

2023



UNIVERSITATEA BABEȘ-BOLYAI
BABEȘ-BOLYAI TUDOMÁNYEGYETEM
BABEȘ-BOLYAI UNIVERSITÄT
BABEȘ-BOLYAI UNIVERSITY

FACULTATEA DE FIZICĂ
Str. Mihail Kogălniceanu nr.1
Cluj-Napoca, RO-400084
Tel: +4(0)264-405300 | FAX: +4(0)264-591906
secretariat.phys@ubbcluj.ro
www.phys.ubbcluj.ro



UNIVERSITATEA "BABEȘ-BOLYAI", CLUJ NAPOCA
FACULTATEA DE FIZICA
SPECIALIZAREA FIZICA

LUCRARE DE LICENȚĂ

Studiul seriilor financiare cu matricea stohastică a proceselor markoviene

Coordonator științific
Prof. dr. Néda Zoltán

Absolvent
Roth Apor

2023



UNIVERSITATEA BABEŞ-BOLYAI
BABEŞ-BOLYAI TUDOMÁNYEGYETEM
BABEŞ-BOLYAI UNIVERSITÄT
BABEŞ-BOLYAI UNIVERSITY

FACULTATEA DE FIZICĂ
Str. Mihail Kogălniceanu nr.1
Cluj-Napoca, RO-400084
Tel: +4(0)264-405300 | FAX: +4(0)264-591906
secretariat.phys@ubbcluj.ro
www.phys.ubbcluj.ro



BABEŞ-BOLYAI TUDOMÁNYEGYETEM, KOLOZSVÁR

FIZIKA KAR

FIZIKA SZAK

ÁLLAMVIZSGA DOLGOZAT

GAZDASÁGI IDŐSOROK TANULMÁNYOZÁSA A MARKOVI FOLYAMATOK
STOCHASZTIKUS MÁTRIXÁVAL

Témavezető tanár
Dr. Néda Zoltán, egyetemi tanár

Hallgató
Roth Apor

2023

Kivonat

A gazdasági fizika a statisztikus fizika elismert interdiszciplináris alterülete. Ezen modern kutatási irányzat egyik lényeges témája a gazdasági idősorok (részvények, indexek, árfolyamok, stb...) modellezése. A jelen dolgozat keretében a Markovi folyamatokra használt stochasztikus mátrixot használjuk arra, hogy a DJIA tőzsdeindex dinamikáját leírjuk. Napos és perces adatokat feldolgozva, megszerkesztjük az indexnek az átlaghoz viszonyított értékére, illetve ennek változására a stochasztikus mátrixot majd ezt egy kompakt matematikai formával megközelítjük. A stochasztikus mátrix sajátvektorait megszerkesztve igazoljuk, hogy azt, hogy az index értékének illetve ennek változása stacionárius eloszlást mutat. Összehasonlítsuk a stochasztikus mátrix alakját és ezen eloszlásokat különböző gazdaságtörténeti periódusokra és észrevesszük azon lényeges trendbeli változásokat amelyek a 2007 utáni évekre jellemzőek. A továbbiakban a stochasztikus mátrixot az idősorok előrejelzési eszközének tekintve megvizsgáljuk, hogy mennyire előrejelezhető a DJIA index dinamikája. Habár statisztikus előrejelzési lehetőségeket kapunk, ezeknek a felhasználása pénzügyi haszonszerzés céljára nem realiztikus.

Abstract

Econophysics is a consecrated interdisciplinary subfield of Statistical Physics. One extremely fashionable research direction in this subfield is the modeling of financial time series (stock prices, indexes, exchange rates, etc...). Here, we use the stochastic matrix characteristic for markovian processes to describe the dynamics of the DJIA stock index. Using the daily and minute resolution data, we first construct the stochastic matrix for the index value relative to the mean trend and the stochastic matrix for the changes in the index. A compact mathematical approximation is then given to describe the observed stochastic matrix. Calculating the eigenvectors of this stochastic matrix, we prove that the distribution of the index values relative to the mean trend and the distribution of the index changes is the one that is characteristic for the stationary limit of the implied stochastic process. We compare the form of the stochastic matrix and the stationary distributions for different economic periods in the history of the DJIA index and observe a striking difference in all trends after the years after 2007. Finally, we use the stochastic matrix to predict the evolution of the DJIA index. Although we observe possibilities for a statistical prediction, we conclude that this is not strong enough for gaining any serious financial gain.

Tartalomjegyzék

1. Bevezető	5
1.1. A gazdasági fizika szükségessége, elterjedése	5
1.2. A gazdasági fizika klasszikus problémái	6
1.2.1. Gazdasági idősorok tanulmányozása, bennük levő általános törvényszerűségek felfedezése	6
1.2.2. Vagyon-és jövedelemeloszlás tanulmányozása	7
1.2.3. Gazdasági folyamatok mint komplex hálózatok	7
1.3. A gazdasági fizikában használt modellek	7
1.3.1. Ideális gázmodell a vagyon eloszlására	7
1.3.2. Ising modell használata adóelkerülés modellezésére	8
1.3.3. Portfólió-optimalizálás korrelációs mátrixok segítségével	9
2. Pénzügyi idősorok elemzése	11
2.1. Részvényárak és tőzsdeindexek	11
2.1.1. S&P500 index	11
2.1.2. DJIA index	11
2.2. Tőzsdeindexek leírására szolgáló modellek	14
2.3. Markov-folyamatok	15
2.4. Stochasztikus mátrix	16
2.5. Folytonos állapottér esete	17
3. Stochasztikus mátrix használata	19
3.1. Stochasztikus mátrix szerkesztése	19
3.1.1. Stochasztikus mátrix értéke a napos adatokra	19
3.1.2. Napi adatok eloszlása a történelem során	20
3.1.3. Perces felbontású adatok vizsgálata	23
3.2. Stacionárius eloszlás vizsgálata	24
3.2.1. Stacionárius eloszlás napos adatokra	25
3.2.2. Eloszlások perces adatokra	27
3.3. Függvény illesztése a stochasztikus mátrixra	27
3.3.1. Az átlaghoz viszonyított értékekre megszerkesztett stochasztikus mátrixok analitikus megközelítése	28

3.3.2. A változásokra megszerkesztett stochasztikus mátrix analitikus megközelítése	29
4. Stochasztikus mátrix alkalmazásra az indexérték alakulásának előrejelzésére	33
4.1. Használt módszer	33
4.2. Eredmények	34
Következtetések	37

1. Bevezető

Dolgozatunk célja a gazdasági fizikában használt módszerek ismertetése és ezek alkalmazása gazdasági idősorok elemzésére. Konkrét alkalmazásként a Dow Jones Industrial Average (DJIA) index néhány aspektusát tanulmányozzuk.

A bevezetőben röviden áttekintjük a gazdasági fizika történetét, érdekeltségi területeit és modelljeit. A dolgozat második részében szót ejtünk a tőzsdeindexekről, bemutatjuk a DJIA index tulajdonságait, és bevezetjük az általunk használt statisztikai fogalmakat és módszereket. A további fejezetekben alkalmazzuk ezen módszereket: a harmadik részben a DJIA-index különböző jellemzőiből nyerhető idősorokat vizsgáljuk, az utolsó fejezetben pedig az index értékének növekvését és csökkenésének az előrejelezési lehetőségeit vizsgáljuk.

1.1. A gazdasági fizika szükségessége, elterjedése

A gazdasági fizika a gazdasági folyamatok és törvényszerűségek statisztikus fizikában használt modellekkel való leírásával foglalkozó tudományág. Megjelenése mind a fizikában, mind a közgazdaságtan számára előrelépést jelentett. A gazdaságban felmerülő problémák tanulmányozásával a fizikusok új területeken tesztelhetik és fejleszthetik az eddig használt modelljeiket. Ezen felül a gazdaságban megjelenő folyamatokat nem mindig magyarázzák megfelelően a meglévő gazdasági modellek, új, olyan nézőpontok is szükségesek amelyek a sokrészecske rendszerek tanulmányozási módszereit is figyelembe veszi [1].

A fizikának a többi komplex rendszereket tanulmányozó ágaival hasonlóan, a gazdasági fizika is a XX. század végétől kezdve terjedt el. Ez első sorban a számítógépek megjelenésének és fejlődésének köszönhető, mely által nagyszámú adathoz lett hozzáférésünk. Az 1970-as évektől kezdve a tőzsdéken átfolyó tranzakciókat digitalizálták, és tranzakciónyi pontossággal (ms)-ok tárolták [2]. Az 1990-es évekre robbanásszerűen megnőtt a feldolgozható adatok száma, ekkortól számítjuk a modern gazdasági fizika elterjedését. A tudományág angol megnevezését ("econophysics") H. E. Stanley az 1995-ben lezajló "Statphys-Kolkata" második konferenciáján vezette be [3]. Érdekes megjegyezni, hogy az első kimondottan gazdasági fizika konferenciát 1998-ban szervezték Budapesten.

A "big data" megjelenésével és elterjedésével a gazdasági fizika az elmúlt évtizedben is óriásit fejlődött. A még nagyobb számú adatmennyiség felhasználásával kiküszöbölhetőek az addigi adatsorban fellépő hibák, hiányosságok, valamint az adatfeldolgozásban használt módszerek segítségével leszűkíthetőek és finomíthatóak a jelenséget magyarázható modellek és feltételezések [4].

Jelenleg egy nagyon aktív kutatási terület, évente több nemzetközi konferenciát tartanak gazdasági fizika témakörben: Econophysics Colloquium (2005-től), Econophys-Kolkata (2005-től), International Conference on Econophysics (2011-től) és Workshop on Economics with Heterogeneous Interacting Agents (1995-től). Gazdasági fizika terén elért eredményeket főleg statisztikus fizikával foglalkozó folyóiratokba publikálnak, mint a Physica A , The European Physical Journal B, vagy a Physical Review Letters E.

Az első könyv a témában az "Introduction to Econophysics: Correlations and Complexity in Finance" [5], valamint további jelentősebb könyvek a témában a Classical Econophysics (Cockshott) [6], illetve A. Chakraborti könyvei: Econophysics: an Introduction [7], Econophysics of Income and Wealth Distributions [8], valamint New Perspectives and Challenges in Econophysics and Sociophysics [9].

1.2. A gazdasági fizika klasszikus problémái

A gazdasági fizika szerteágazó problémákkal foglalkozik. A következőkben csak a legfontosabbakról lesz szó.

1.2.1. Gazdasági idősorok tanulmányozása, bennük levő általános törvényszerűségek felfedezése

Az idősorok megadják egy adott tőzsdeindex értékének időbeli változását. A leggyakrabban egy index logaritmikus visszatérítési rátájának eloszlását és az index volatilitását vizsgálják, az utóbbi azt adja meg, hogy milyen gyorsan változik meg az index értéke egy bizonyos idő alatt. A logaritmikus visszatérítési ráta a következőképpen adható meg:

$$r_{\Delta t}(t) = \ln \left(\frac{S(t + \Delta t)}{S(t)} \right) \quad (1)$$

Az idősorok tanulmányozása közben általános törvényeket keresünk. Ilyen, jól ismert általános törvényszerűségek például a visszatérítési ráták hatványfüggvény-eloszlása ($\rho(x) = Cx^{-(1+\alpha)}$), az autokorrelációk hiánya adott visszatérítési ráta esetén, és a nagy változások összesűrűsödése ("volatility clustering") [10]. Fontos és nagyon tanulmányozott a különböző tőzsdeindexek árai közti korrelációk is.

1.2.2. Vagyon-és jövedelemeloszlás tanulmányozása

A gazdasági fizikának ezen ága a társadalomban fellelhető egyenlőtlenségeket modellezi. A rendelkezésünkre álló adat azt tanúsítja, hogy az alsó- és felsőosztály vagyoneloszlását különböző típusú eloszlásfüggvények írják le [11]. Az eloszlás időbeli változásának tanulmányozása során észrevették, hogy amíg az alsó osztály vagyoneloszlása az évek elteltével is nagyjából azonos maradt, addig a felső osztály hatványfüggvényszerű eloszlását leíró Pareto-eloszlás kitevője változott az évek során. A szegényebb rétegekben tapasztalt eloszlások, amint a későbbiekben látni fogjuk, magyarázható például a tranzakcióknak az ideális gázokban levő részecske ütközésekkel való analógia segítségével. [8].

1.2.3. Gazdasági folyamatok mint komplex hálózatok

A gazdaságban lejátszódó folyamatok tanulmányozhatóak komplex hálózatokként is, ahol a gazdasági szereplők a hálózat csomópontjai, melyek az interakciójukon keresztül kapcsolódnak egymáshoz. A gazdasági fizika ezeket a hálózatokat első sorban komplex rendszerként kezeli, és a hálózat szerkezetét, statisztikai leírását, szabályosságait tanulmányozza. Az adatbányászat útján kapott eredményeket sajátos modellekkel tanulmányozza [12]. A gazdasági hálózatokban nagy jelentősége van az egyes csúcsok centralitásának, mely a csúcsok fontosságának a mértéke. Erre több rangsorolási módszer is létezik, a legegyszerűbb fokszám szerint rangsorolni a csúcsokat.

Fontos témakör ezen hálózatok dinamikájának tanulmányozása, ugyanis ennek az ismeretében tanulmányozhatjuk a gazdasági folyamatok evolúcióját. Felfedték például, hogy a bankrendszer, mint hálózat fogékony a rendszerkockázatok megjelenésére csak a szerkezetükből adódóan [13], valamint a 2008-as válság idején a piacok hálózata erősen összefüggő, ezáltal sérülékeny volt [14].

1.3. A gazdasági fizikában használt modellek

A gazdasági fizika sok esetben a statisztikus fizika által kifejlesztett modelleket alkalmazza. A következőkben egy pár példa segítségével szemléltetjük ezen analógiákat:

1.3.1. Ideális gázmodell a vagyon eloszlására

A legegyszerűbb erre használt modell a tranzakciókban a pénz vándorlását az ideális gázban fellelhető ütközésekkel veszi analógnak, és a jól ismert Gibbs-Boltzmann-eloszlás megjelenését igazolja:

$$P(E) = c \cdot e^{-E/(k_B T)} \quad (2)$$

Mivel a tranzakció során a teljes pénzmennyiség megmarad, ezért a részecskék energiája azonosítható az egyének pénzmennyiségével, valamint a Boltzmann-állandót egységnyinek véve a hőmérséklet (T_m) az egyének átlag-pénzmennyiségével azonosítható. Ilyen alapon a vagyoneeloszlásra exponenciális adódik [11]:

$$P(m) = c \cdot e^{-m/T_m} \quad (3)$$

A modellben megengedhető negatív pénzmennyiség is, amely az adósságot jelképezi. Nincs figyelembe véve, hogy a rendszerbe vihető be pénz a központi bankok által, valamint bizonyos javak a tranzakció lebonyolítása után nem érnek pénzt.

Vagyoneeloszlás esetén figyelembe kell venni, hogy a fogyasztó vagyona részét képezik a pénze mellett a tulajdonjai is. Mivel a rendszerben a tulajdonok értéke és a pénz mennyisége csak rövid távon marad meg, ezért stochasztikus modelleket is használnak a vagyoneeloszlás leírására. A rendelkező adatokból az derül ki, hogy a szegényebb rétegek vagyoneeloszlását jól leírja a Boltzmann-Gibbs eloszlás, azonban a gazdagabb réteg vagyona és jövedelme Pareto-eloszlás (hatványfüggvény eloszlás) szerint oszlik meg [11].

1.3.2. Ising modell használata adóelkerülés modellezésére

Az Ising-modellt a ferromágnesség leírására dolgozták ki, melyben az atomi spinek közti kicserélődési kölcsönhatást veszi alapul:

$$H = - \sum_{i,j} J_{ij} S_i S_j - \sum_j B_j S_j \quad (4)$$

A modellben a H a Hamilton-függvénye az adott spinkonfigurációnak, S_i az i . atom spinje, J_{ij} két szomszédos spin közti kicserélődési kölcsönhatási együttható, valamint B_j a j . atomra ható külső mágneses térből származó indukcióvektor. A modell alkalmazhatósága kiterjeszhető a gazdasági fizikára is, és erre egy példa az adóelkerülés modelle, ahol az atomi momentumok a gazdaság szereplőinek felelnek meg, a spinek iránya megadja, hogy az adott szereplő befizeti-e vagy elkerüli az adót.

A modellbe a szereplők viselkedését is be lehet vinni [15]:

- Minden szereplőnek létezik egy T_i hőmérséklete, melynek a kicserélődési kölcsönhatással vett hányadosa megadja, hogy mennyire hat kölcsön a szomszédjaival (J/T_i kis értékei esetén individualisztikus)

- A B_i előjel meghatározza, hogy a szereplő alpból befizeti-e az adót, vagy nem
- Bevezethető az egyenletbe a felülvizsgálati valószínűség is, mely következtében a be nem fizetett adót egy bizonyos határidőn belül be kell fizetni.
- Az emberek így négy csoportra oszthatóak: akik nem fizetnek adót ($|B_i|/J > 1$ és $B_i < 0$), akik mindig fizetnek adót ($|B_i|/J > 1$ és $B_i > 0$), akik a környezetet másolják ($|B_i| \ll J$ és $J > T_i$), valamint akik véletlenszerűen viselkednek ($|B_i| \ll J$ és $J \ll T_i$)

Az így alkotott modell analitikus közelítésekkel vagy számítógép szimulációval tanulmányozható.

1.3.3. Portfólió-optimalizálás korrelációs mátrixok segítségével

A portfólió-optimalizálás arra keres választ, hogy hogyan fektessünk be nagyszámú részvénybe úgy, hogy egy adott várt hozam esetén minél kevesebb kockázatot vállaljunk [16]. Ekkor, ha R -el jelöljük az összhozam számát, és R_i -vel az egyes részvények hozamát, a kockázat minimalizálásához az R varianciáját kell minimálisnak tartani. Ha G_i egy meghatározott, Δt idő alatti részvényár-változás, w_i a befektetett részvényhez tartozó súly, akkor T idő alatt a variancia:

$$D^2 = \frac{T}{\Delta t} \sum_{i,j=1}^N (\langle G_i G_j \rangle - \langle G_i \rangle \langle G_j \rangle) w_i w_j \quad (5)$$

Bevezetve a korrelációs mátrixot:

$$C_{ij} = \frac{\langle G_i G_j \rangle - \langle G_i \rangle \langle G_j \rangle}{\sigma_i \sigma_j} \quad (6)$$

Felírható, hogy:

$$D^2 = \sum_{i,j} C_{ij} \sigma_i \sigma_j w_i w_j \quad (7)$$

Ahol:

$$\sigma_i = \sqrt{\frac{T}{\Delta t}} \cdot \sqrt{\langle G_i^2 \rangle - \langle G_i \rangle^2} \quad (8)$$

Ezt minimalizálva a Lagrange-féle multiplikátorok segítségével lehet, a két peremfeltétel adott: $\sum_i w_i = 1$, $\sum_i R_i w_i = R$. A minimalizálandó függvény a következő:

$$F = \sum_{i,j} C_{ij} \sigma_i \sigma_j w_i w_j - \lambda \sum_i R_i w_i - \mu \sum_i w_i \quad (9)$$

Ez látható, hogy analóg egy Ising-típusú spin-üveg modellel, ahol $C_{ij}\sigma_i\sigma_j$ felel meg a kicserélődési kölcsönhatásnak, w_i a spineknek, és R_i a külső mágneses térnek.

2. Pénzügyi idősorok elemzése

2.1. Részvényárak és tőzsdeindexek

A részvények egy adott vállalat értékpapírjai, melynek tulajdonosai részesednek a vállalat vagyonából. A részvényár egyetlen egy részvény értéke, és a részvényár meg a kibocsátott részvények száma megadja a vállalat piaci kapitalizációját. A részvényárak változása így az adott vállalat értékének változását fejezik ki.

A tőzsdeindexek több vállalat részvényeiből összeállított portfóliók értékét követő indexek, melyek egy adott ország/ régió gazdasági helyzetét és jelentős gazdasági eseményeit (növekedés/recesszió) jellemzik. Különböző indexeknek különbözőképpen határozzák meg az értékét, két használt módszer az részvényárak szerint súlyozott index, valamint a részvények piaci kapitalizációja szerint súlyozott index. Két ismert tőzsdeindex az S&P500 és a DJIA indexek, melyeket az Egyesült Államok tőzsdéin bejegyzett jelentősebb cégek részvényeiből állítanak össze.

2.1.1. S&P500 index

Az S&P tőzsdeindex az Egyesült Államok 500 legnagyobb vállalatának teljesítményét követi. Az értékét az öt alkotó részvények piaci kapitalizációja szerint súlyozzuk:

$$I = \frac{\sum_{i=1}^{500} P_i \cdot Q_i}{D} \quad (10)$$

Itt P_i az i . vállalathoz tartozó részvény ára, Q_i az i . vállalat részvénytulajdonosi részének száma. A D osztó az index értékének kalibrálására szolgál, mely akkor szükséges, mikor egy cég olyan változtatást hoz be, mely a piaci kapitalizáció értékének változtatásával jár, valamint ha az 500 cég közé egy új cég kerül.

2.1.2. DJIA index

A DJIA tőzsdeindex az Egyesült Államok 30 kiemelkedő cégének teljesítményét mutató index. Az értéke az öt alkotó részvények ára szerint van súlyozva:

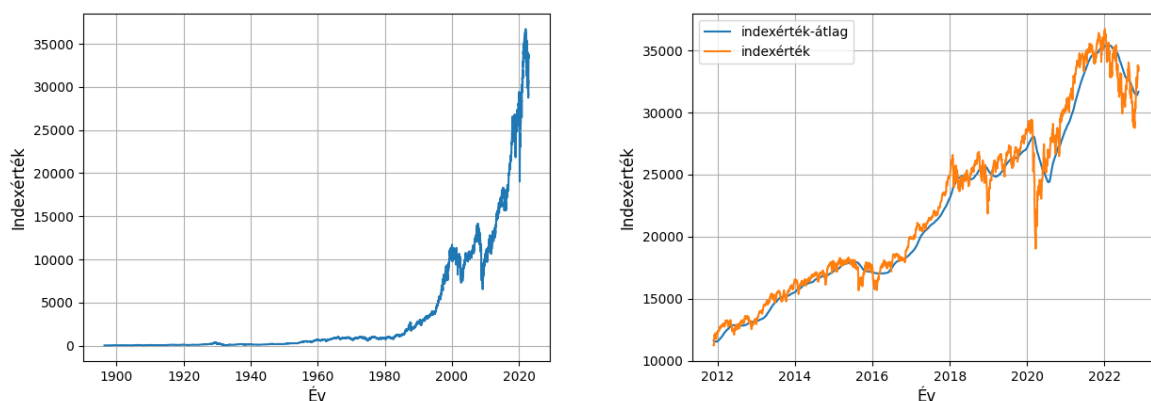
$$I = \frac{\sum_{i=1}^{30} P_i}{D} \quad (11)$$

Mivel az index a XIX. század végén 12 cég teljesítményét mérte, ezért újabb cég bekerülésekor az indexet úgy kellett kalibrálni, hogy ne változzon meg drasztikusan. Ezt, mint az S&P500

esetén, ugyancsak egy osztóval érik el. Az osztó értékét akkor is megváltoztatják, mikor egy cég részvényeit feldarabolják.

Az indexérték alakulásának napi adatai letölthetőek a [17] honlapról, nagyobb felbontású adatok azonban korlátozott mennyiségben, fizetősén érhetőek el. A perces adatokat a [18] oldalról töltöttük le, itt az elmúlt 15 évre találhatóak meg. Részletesebb, percnyi pontosságnál nagyobb adatok nem állnak rendelkezésünkre, valamint a számunkra hozzáférhető adatok sem aktuálisak az index pillanatnyi értékével.

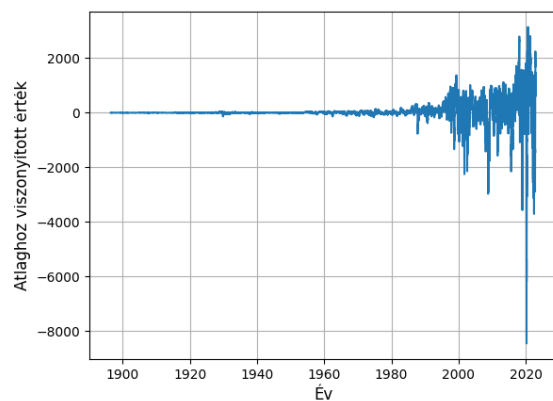
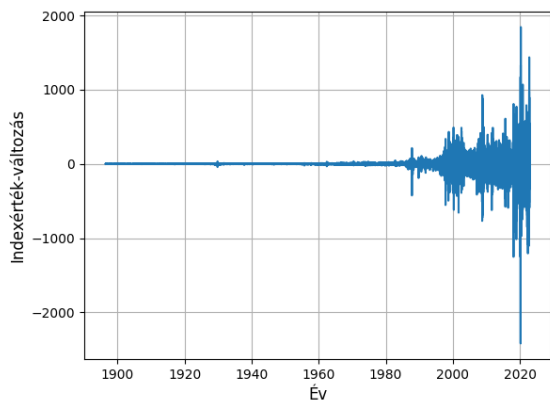
Mivel a DJIA index értéke a gazdaság növekedésével folyamatosan nő, ezért az index pillanatnyi alakulását az előző napok mozgó átlagához szokták mérni, így az átlagtól való távolság ad meg pontosabb információt arról, hogy az index értéke éppen magasan, vagy alacsonyan van. A következő ábrák az index napi átváltozását szemléltetik.



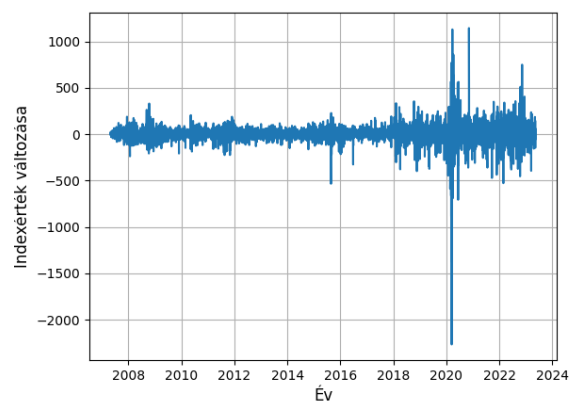
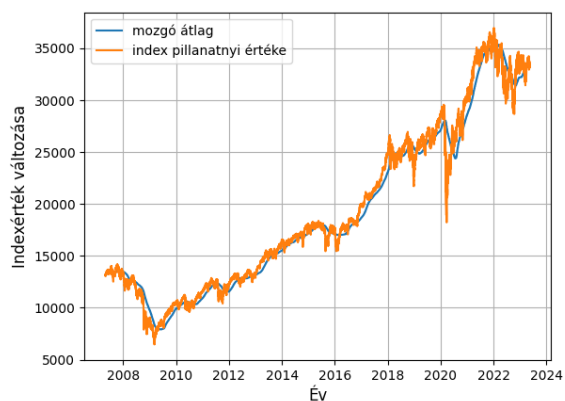
2.1. ábra. **Bal:** a DJIA index változása a XIX. századtól napjainkig **Jobb:** a DJIA index értéke és 100 napra vett mozgó átlaga az utóbbi évtizedben. Jól látható a COVID által előidézett gazdasági válság

A továbbiakban a DJIA index két mutatóját fogjuk részletesen vizsgálni: a trendvonalról hoz viszonyított értéket, valamint két egymást követő időpillanatban (napos felbontásban egy nap, perces felbontásban egy percnyi időintervallum) az index aktuális értékei közti változás nagyságát. Az index árát minden esetben a nyitóárnak rögzítettük. Ezt a két mutatót napi felbontás szerint ábráztuk:

Perces felbontású adataink 2007-től kezdve állnak rendelkezésünkre. Ezek reggel 9:30-tól 16:00-ig követik az index értékének alakulását, a napos adatokhoz hasonlóan itt is az index nyitóértékét vesszük figyelembe. A következő grafikonokon ábrázoljuk az index értékét és mozgó átlagát, valamint az indexérték változásának a perces felbontását:

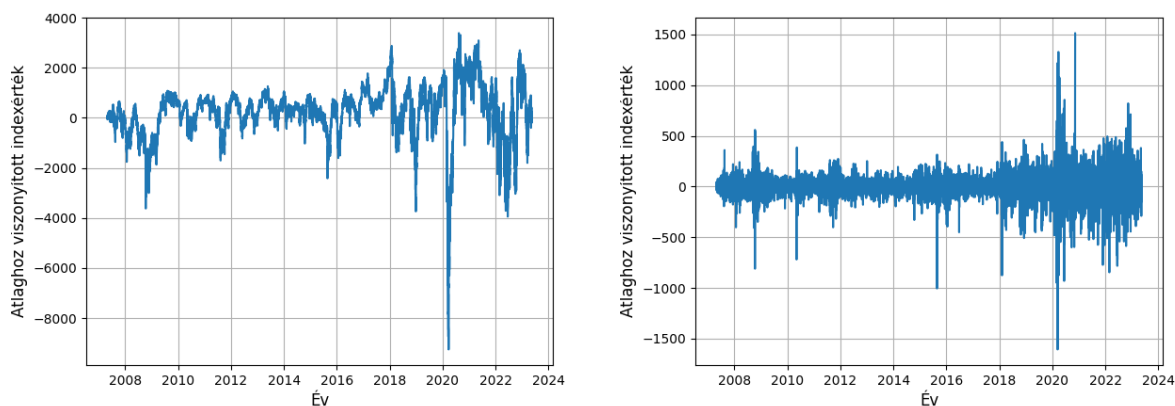


2.2. ábra. **Bal:** A DJIA-index értékének két egymás utáni időpillanat közötti változás napos felbontásban **Jobb:** A DJIA-index értéke a trendvonalhoz viszonyítva napos felbontásban



2.3. ábra. **Bal:** 2007-2022 között az index értéke és mozgó átlaga perces felbontásban. A mozgó átlagot 40000 percre szerkesztettük. **Jobb:** 2007-2022 között az index értékének percenkénti változásának nagysága

Percenkénti adatoknál az átlagolás ablakától függően is vizsgáljuk az index tulajdonságait. A következő ábrák alapján látható, hogy a trendvonalhoz képest vett érték erősen függ az átlagablak méretétől:



2.4. ábra. **Bal:** Az index értékének trendvonalhoz viszonyított értéke 40000 perces (kb. 100 napos) ablakkal **Jobb:** Az index értékének trendvonalhoz viszonyított értéke 60 perces ablakkal

2.2. Tőzsdeindexek leírására szolgáló modellek

Mivel a tőzsdeindexek változása jellemzi a gazdaság alakulását, ezért ennek előrejelzése segíthet a gazdasági fellendülések kihasználásában és a gazdasági válságok megjósolásában. Erre a közgazdaságtanban több módszert is kidolgoztak, melyek különböző regressziókon alapulnak, és amelyekre különböző machine learning algoritmusokat /neurális hálózatokat használnak [19].

A gazdasági fizikának a célja első sorban olyan modelleket alkotni, amelyek meg is magyarázzák a tőzsdeindexek alakulását. A legkorábbi modell, a véletlen bolyongás volt, mely szerint az index visszatérítési ráta eloszlásának $p(r, t)$ a változását a diffúziós egyenlet írja le:

$$\frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial r^2} \quad (12)$$

p az r visszatérítési ráta valószínűség-sűrűsége. Ennek a megoldása egy adott t időpillanatra Gauss-eloszlást ad vissza:

$$p(r, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{r^2}{4Dt}} \quad (13)$$

A visszatérítési rátát az inverz statisztikai módszereivel is tanulmányozták, arra keresve a választ, hogy egy index visszatérítési rátája mennyi idő alatt (τ_c) ér el egy kritikus ρ hozam értéket. Ennek az értéknek az eloszlása véletlen bolyongásra megadható, mint [20, 21]:

$$p(\tau_c) = \frac{|\rho|}{\sqrt{4\pi D \tau_c^3}} e^{-\frac{\rho^2}{4\pi D \tau_c}} \quad (14)$$

Ez alapján létezik egy τ_c^* , idő melyben az eloszlásnak maximuma van, ezt optimális befektetési horizontnak nevezzük [20]:

$$\tau_c^* = \frac{1}{6D}\rho^2 \quad (15)$$

Ennél jóval nagyobb értékeknél igaz a $\tau_c \propto \rho^{-\frac{3}{2}}$ skálázási törvény. Bár ez utóbbit a DJIA indexből származó adatok megerősítik, az optimális befektetési horizontra nem kapjuk vissza az előrejelzett értéket, mely jelzi a modell hibáját [22].

Egy másik, empirikusan meghatározott törvény a nyereség-veszteség asszimmetria, mely szerint a veszteség optimális befektetési horizontja hamarabbra esik, mint az ugyanolyan értékű nyereségé. Ezt többfajta modellel próbálták magyarázni, egyik magyarázat a részvények korrelációjának növekedése veszteségek idején [23].

Egy további univerzális törvény az emelő-hatás, mely szerint negatív visszatérítési ráta után a megnő az index volatilitása (melyet az index varianciájával szoktunk jellemezni). Az effektust leíró függvény az emelőfüggvény, melyet a következőképpen definiáltak [24]:

$$L(\tau) = \frac{\langle r_{\Delta t}^2(t + \tau)r_{\Delta t}(t) \rangle}{\langle r_{\Delta t}^2(t) \rangle^2} \quad (16)$$

A DJIA indexet vizsgálva látható, az emelőfüggvény negatív, ha $\tau > 0$, ami azt jelenti, hogy negatív visszatérítési ráták esetén jövőbeli volatilitás megnő. Egy magyarázata ennek, hogy a befektetők a vagyonukat más részvényekbe fektetik, miután lecsökken egy részvény értéke, ezáltal megnövelve ezek volatilitását [25].

Látható tehát, hogy tőzsdeindexek tulajdonságait nem mindig írja le a véletlen bolyongás modellje. A dolgozatunk további részében a DJIA index néhány jellemzőjét Markov-folyamatként kezeljük, a továbbiakban bemutatjuk a folyamatot leíró statisztikai eszközöket.

2.3. Markov-folyamatok

A Markov-folyamatok olyan stochasztikus folyamatok, mely során a rendszer elkövetkező állapotai csak a rendszer jelenlegi állapotától függ, és a rendszer múltja ezt nem befolyásolja. Ha a rendszer diszkrét időlépésű, ez a feltétel felírható matematikailag a következőképpen:

$$\mathcal{P}(X_n = x | X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = \mathcal{P}(X_n = x | X_{n-1} = x_{n-1}) \quad (17)$$

Ahol \mathcal{P} fejezi ki a valószínűségi függvényt, X_i az i . időpillanatnak megfelelő valószínűségi változó, melynek értéke fejezi ki a rendszer állapotát. Markov-folyamatok esetén megadható így az n . időpillanatban egy adott állapot valószínűsége az $n-1$. időpillanatban az összes lehetséges

állapot valószínűségének függvényében a stochasztikus mátrix segítségével.

2.4. Stochasztikus mátrix

A stochasztikus mátrix megadja egy több, diszkrét állapotú markovi rendszerben a különböző állapotok közti átmeneti valószínűséget. Tekintve egy n állapotú rendszert, a stochasztikus mátrix a következőképpen írható fel:

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix} \quad (18)$$

Ahol a P_{ij} az i . állapotból a j . állapotba való átmenet valószínűsége. A stochasztikus mátrix normált kell legyen, azaz $\sum_{j=1}^N P_{ij} = 1, \forall i = \overline{1, n}$.

Tekintsük a $\mathbf{v}^{(k)} = [v_1^{(k)} \ v_2^{(k)} \ \dots \ v_n^{(k)}]$ vektort, mely kifejezi, hogy a k . időpillanatokban a rendszer milyen valószínűséggel van egyes állapotokban. Ez a vektor is normált kell legyen, azaz: $\sum_{i=1}^N v_i^{(k)} = 1, \forall k \in \mathbb{N}$. Ekkor a $k+1$. időpillanatban felírható, hogy:

$$\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} \cdot \mathbf{P} \quad (19)$$

Ezáltal, ha ismerjük, hogy egy adott pillanatban a rendszer milyen valószínűséggel található egy adott állapotban, meghatározható az állapotok valószínűsége a következő pillanatban is. Az esetünkben azt tanulmányozzuk, hogy a vizsgált változó milyen valószínűséggel vesz fel egy adott intervallumból való értéket (ami lehet az index értéke vagy az indexérték változásának nagysága) egy adott időpillanatban. Ha a változó lehetséges értékeiből n intervallumot képezünk, akkor a felírt \mathbf{v} vektor k . eleme azt fejezi ki, hogy milyen valószínűséggel található a változó értéke a k . intervallumban, tehát a \mathbf{v} vektor fejezi ki egy adott pillanatban a változó értékének eloszlását. Ekkor a stochasztikus mátrix P_{ij} eleme a változó értékének az i . intervallumból a j . intervallumba történő átmenet valószínűségét fejezi ki. Ezáltal ismerve a stochasztikus mátrix együtthatóit, ismerve egy adott időpillanatban a változó értékének a valószínűség-eloszlását, meghatározható ennek az eloszlása a következő időpillanatokban is.

Nagy időskálán az indexértékek és az értékváltozások stacionárius eloszlása érdekel, ugyanis ezekből következtethetünk az index általános jellemzőire. A stacionárius eloszlás időfüggetlen, így a változó értéke ebben az esetben minden időpillanatban ugyanolyan valószínűséggel található az

adott intervallumban, ezért két egymás után következő pillanatban a \mathbf{v} vektor megegyezik. Tehát $\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)}$, bármilyen k pillanatban, így a 19-as egyenletből a következőt kapjuk:

$$\mathbf{v} = \mathbf{v} \cdot \mathbf{P} \quad (20)$$

A fenti egyenlet a mátrix baloldali sajátérték-egyenlete, és a stacionárius eloszlást kifejező \mathbf{v} vektor a mátrixnak az 1-es sajátértékhez tartozó sajátvektora.

2.5. Folytonos állapotter esete

Folytonos állapotterben bevezethetünk a stochasztikus mátrix helyett egy kétváltozós függvényt (kernelt), mely meghatározza két állapot közti átmenet valószínűségét a következőképpen:

$$\mathcal{P}(x \rightarrow (y, y + dy)) = w(x, y)dy \quad (21)$$

Ahol $\mathcal{P}(x \rightarrow (y, y + dy))$ a mutató x értékből az $(y, y + dy)$ intervallumba való átmenetének a valószínűsége. Az x érték a változó egy adott, k . időpillanatban felvett értéke, az y pedig a $k + 1$. időpillanatban felvett értéke. Ha a stochasztikus mátrixban a felosztott intervallumok Δ nagyságúak, és egy $(x_0, x_0 + \Delta)$ intervallumból vett tetszőleges két értékre ugyanaz az átmeneti valószínűség bármely intervallumra, a Δ értékével 0-hoz tartva a kernel megkapható, mint a stochasztikus mátrix értékeinek Δ -val való hányadosának a határértéke.

Mivel $\forall x$ -re az átmenet valamilyen y -ba kerül, átírható a stochasztikus mátrix esetén vett normálási feltétel a következő alakba:

$$\int_{-\infty}^{\infty} w(x, y)dy = 1, \forall x \in \mathbf{R} \quad (22)$$

Ha ismert egy k . időpillanatban a változó által felvett értékek lehetséges eloszlása, folytonos esetben is meghatározható a kernel segítségével a $k + 1$. időpillanatban az eloszlásfüggvény. Diszkrét stochasztikus mátrix esetben a \mathbf{v} vektor i . eleme a 19. egyenlet alapján meghatározható, mint:

$$v_i^{(k+1)} = \sum_j v_j^{(k)} P_{ji} \quad (23)$$

Ez általánosítható folytonos esetre:

$$\rho^{(k+1)}(y) = \int_{-\infty}^{\infty} \rho^{(k)}(x)w(x, y)dx \quad (24)$$

Ahol $\rho^{(k)}$ a változó eloszlásfüggvénye a k . időpillanatban. Stacionárius esetben az eloszlásfüggvény időben nem változik, tehát:

$$\rho(y) = \int_{-\infty}^{\infty} \rho(x)w(x, y)dx \quad (25)$$

3. Stochasztikus mátrix használata

Ebben a fejezetben a napi-és perces adatok esetén is megszerkesztjük a trendvonalhoz viszonyított értékek és az index egységnyi (napi vagy perces) időtartam alatt történő változások által meghatározott stochasztikus mátrixot, megvizsgáljuk a mátrix alakját különböző átlagablakok esetén. Ezután meghatározzuk a stochasztikus mátrix segítségével a stacionárius eloszlásfüggvényt, és összehasonlítjuk az adatsorból számolt eloszlással. Végül a kapott stochasztikus mátrixokra analitikus alakban felírható függvényeket illesztünk, és megvizsgáljuk ezek tulajdonságait.

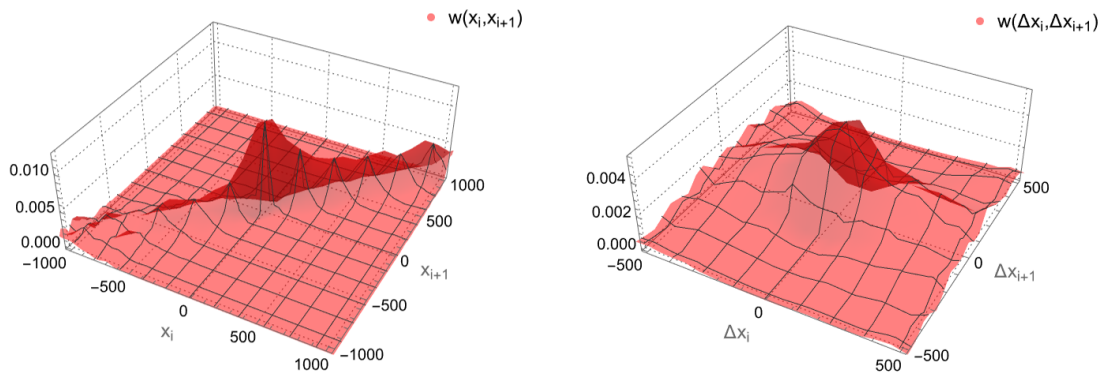
3.1. Stochasztikus mátrix szerkesztése

Mind az index trendvonalhoz viszonyított értéke, mind az indexérték változása esetén azt vizsgáltuk, hogy mekkora a valószínűsége annak, hogy egy adott értékű változás/ érték után következő időpillanatban a változás/ érték egy bizonyos értéket vegyen fel. Ennek meghatározására mindkét adatsorból háromdimenziós hisztogramot készítettünk, melyen azt ábrázoltuk, hogy tanulmányozni kívánt mutató (indexérték-változás/átlaghoz viszonyított érték) egy tetszőleges i . időpontban egy adott x_i (átlaghoz viszonyított esetén) vagy Δx_i (változás esetén) érték esetén mekkora következő időpillanatban egy tetszőleges x_{i+1} vagy Δx_{i+1} pillanatba való átmeneti valószínűsősűrűség (az ábrákon ezt w -vel jelöljük).

A hisztogram tartományaiból szerkesszük meg a stochasztikus mátrix P_{ij} elemeit. Ehhez normálnunk kell a hisztogramot úgy, hogy igaz legyen, hogy $\sum_j P_{ij} = 1$, azaz egy rögzített i -re leosztjuk a hisztogram $\{i., j.\}$ tartományának értékét ezeknek a tartományok értékének az összegével. Ahhoz, hogy sűrűségfüggvényt kapjunk, egy adott tartomány értékét leosztjuk a felvett intervallumok nagyságával.

3.1.1. Stochasztikus mátrix értéke a napos adatokra

Látható, hogy átlaghoz viszonyított értékekből nyert stochasztikus mátrix esetén egy adott indexérték esetén az átmeneti valószínűség értéke az értékhez közeli indexértékek esetén a legnagyobb, tehát egy adott értékről az átmenet valószínűsége csökken az átmenet nagyságának növekedésével. Az is látható azonban, hogy a stochasztikus mátrix főátló menti elemeinek nagysága csökken nagyobb értékek esetén, így ha egy időpillanatban távol vagyunk a trendvonaltól, nagyobb valószínűséggel változik nagyobbat a trendvonalhoz viszonyított érték, mint ha közel lennénk a trendvonalhoz.

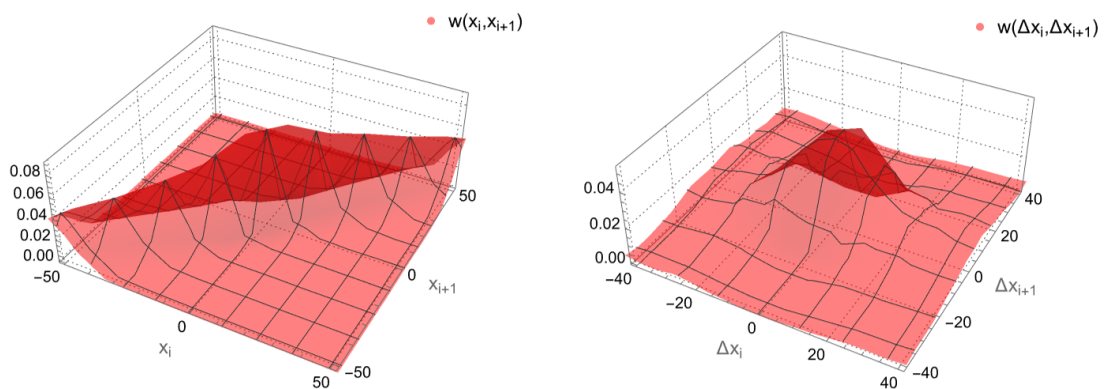


3.1. ábra. **Bal:** a mozgó átlaghoz viszonyított értékekből kapott stochasztikus mátrix napos adatokra **Jobb:** változásokból nyert stochasztikus mátrix napos adatokra

A változásokból nyert grafikonon az látható, hogy az index értékének kis változása esetén a következő változás értéke is nagy valószínűséggel kicsi. Az is látható, hogy minél nagyobbat változik az index értéke, a következő változás annál nagyobb valószínűséggel lesz szintén nagyobb értékű.

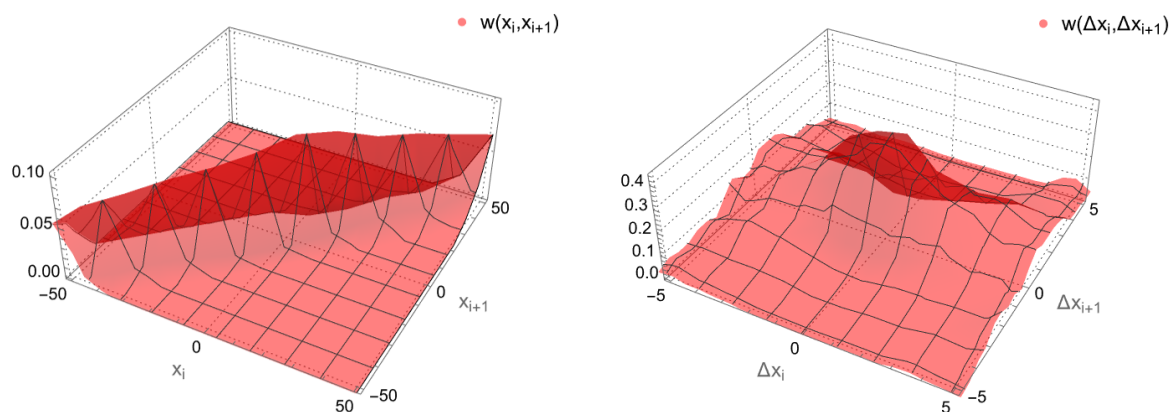
3.1.2. Napi adatok eloszlása a történelem során

A következőkben a napi adatsort négy részre bontottuk: 1940 előtti adatok, 1940 és 1980 közti értékek, 1980-2007, valamint 2007-2022 közti értékek. Mivel az 1980-as évek előtt az index értéke keveset változott, ezért mind az átlaghoz viszonyított értékek, mind az indexértékek változásai kis értékeket vesznek fel. A teljes napi adatsorral való összehasonlításához ezért ábrázoljuk a teljes adatsorból kapott, nagyobb felbontásban kapott átlaghoz viszonyított értékekből és indexérték-változásból szerkesztett stochasztikus mátrixot:

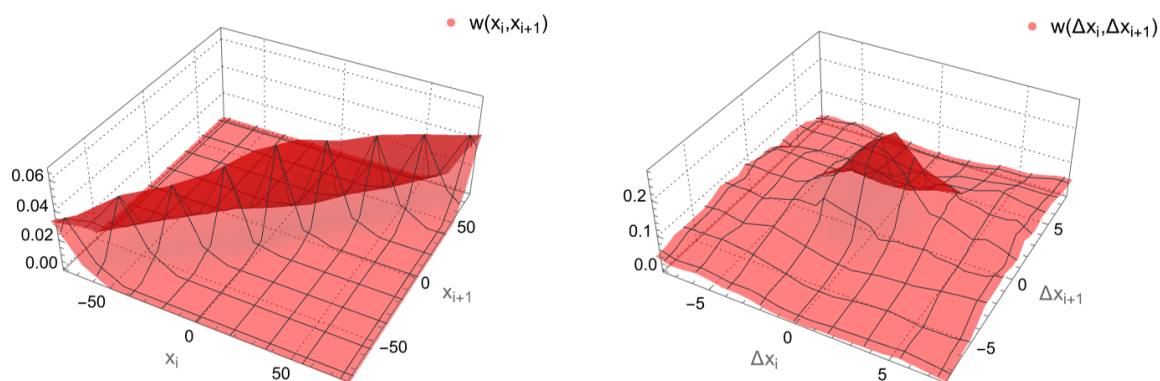


3.2. ábra. **Bal:** az összes napi adatból nyert stochasztikus mátrix nagyobb felbontásban, kisebb értékekre **Jobb:** az összes napi adatból nyert stochasztikus mátrix nagyobb felbontásban, kisebb változás-értékekre

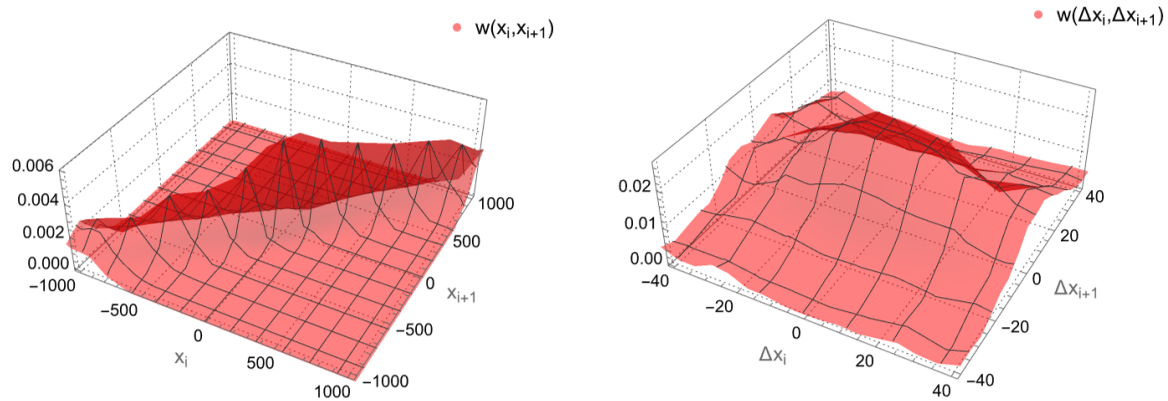
A következőkben ábrázoljuk a négy időszakból kapott stochasztikus mátrixokat:



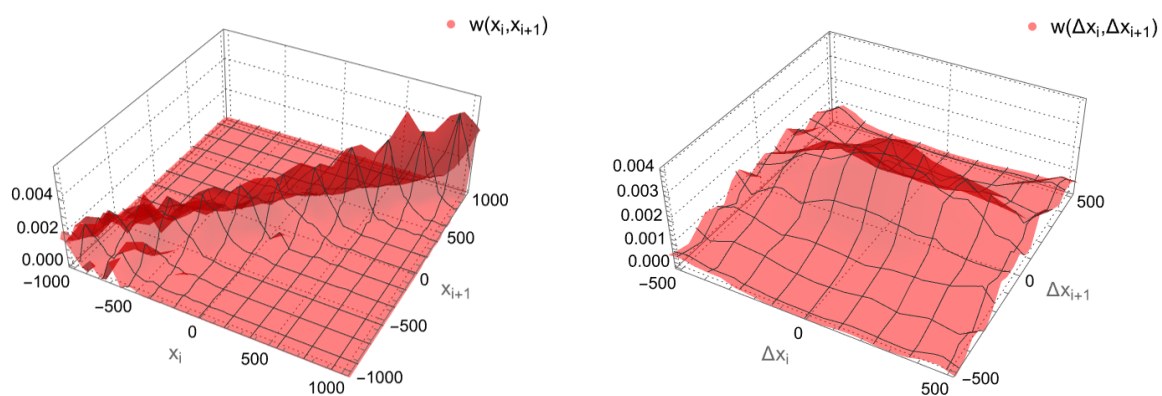
3.3. ábra. **Bal:** 1896-1940 közti átlaghoz viszonyított értékek napi felbontásából nyert stochasztikus mátrix **Jobb:** 1896-1940 közti változások napi felbontásából nyert stochasztikus mátrix



3.4. ábra. **Bal:** 1940-1980 közti átlaghoz viszonyított értékek napi felbontásából nyert stochasztikus mátrix **Jobb:** 1940-1980 közti változások napi felbontásából nyert stochasztikus mátrix



3.5. ábra. **Bal:** 1980-2007 közti átlaghoz viszonyított értékek napi felbontásából nyert stochasztikus mátrix **Jobb:** 1980-2007 közti változások napi felbontásából nyert stochasztikus mátrix



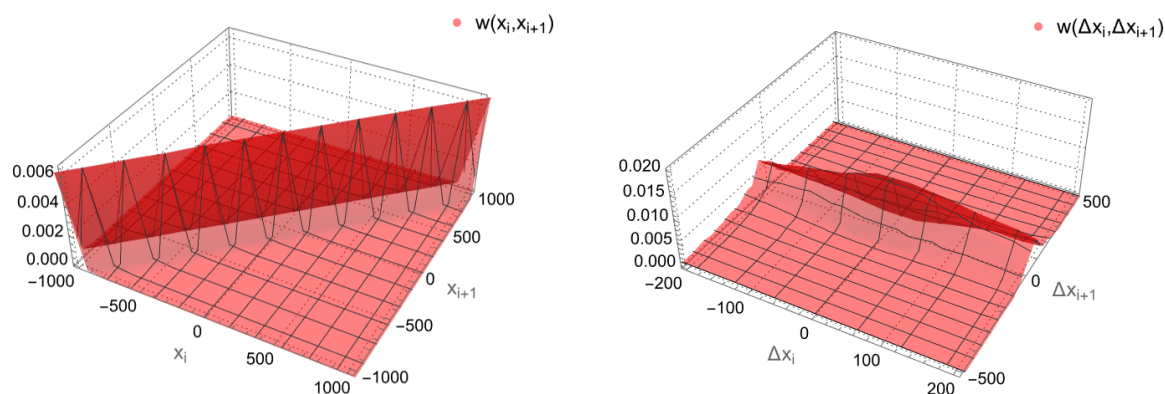
3.6. ábra. **Bal:** 2007-2022 közti átlaghoz viszonyított értékek napi felbontásából nyert stochasztikus mátrix **Jobb:** 2007-2022 közti változások napi felbontásából nyert stochasztikus mátrix

Látható, hogy az egész idősort leíró, nagy felbontású stochasztikus mátrix hasonlít az 1896-1940 és 1940-1980-as adatokból felépített stochasztikus mátrixra. Ennek az az oka, hogy az 1980-as évekig az index értéke viszonylag keveset változott, így mind a trendvonalhoz képest vett érték, mind az egyes napok közti értékek változása kicsi. 1980-tól 2007-ig vegyesen megtalálhatóak kisebb-és nagyobb átlagtól való eltérések és indexérték-változások, így a stochasztikus mátrix alakja még mindig hasonlít a teljes adatsorból vett mátrix alakjára, azonban már nagyobb skálán mozognak a trendvonalától való eltérések és a napi változások.

2007-től erre fele a korábbi adatokhoz képest az index átlaghoz viszonyított értékének és indexérték-változásának értéke drasztikusan megnőtt, így az értékekből készített stochasztikus mátrix már nem is hasonlít az eredeti adatsor stochasztikus mátrixára, a napi különbségek nagysága pedig jelentősen megnőtt. Ezen kívül a stochasztikus mátrix maximuma se az origóban található, így a legvalószínűbb átmenet az átlagtól való pozitív, nagy eltérésből hasonló nagyságú és előjelű eltérés esetén következik be. Ez betudható az indexérték nagyléptű növekvésének, így a mozgó átlag egy jelentősen későbbi időpillanatban éri el az index már valaha elért értékét.

3.1.3. Perces felbontású adatok vizsgálata

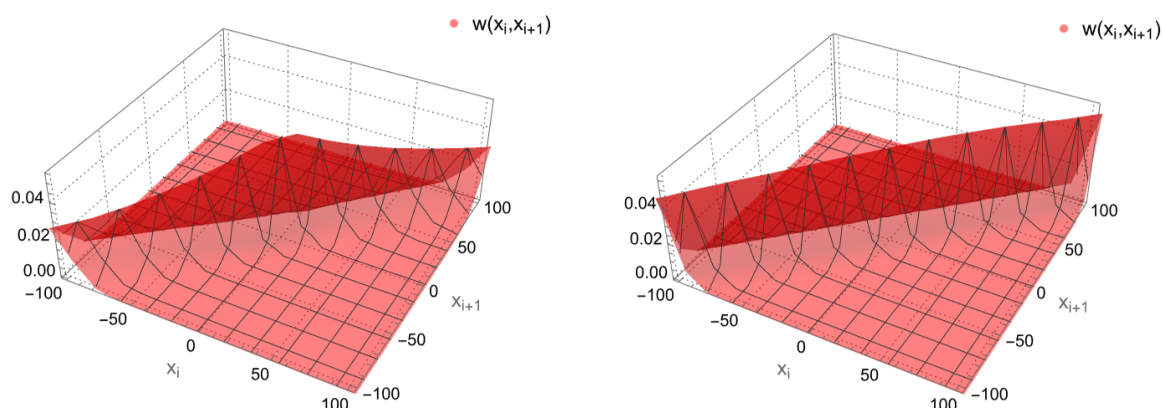
A következőben a napos adatoknál leírtak alapján elkészítjük a percenkénti trendhez viszonyított értékek- és értékváltozásokból származó stochasztikus mátrixot. Első lépésben a mozgó átlagot 40000 percre vesszük, mely nagyjából 100 napnak felel meg a tőzsdén.



3.7. ábra. **Bal:** Perces adatokból nyert átlaghoz viszonyított érték stochasztikus mátrixa 40000 perces átlagablakkal dolgozva **Jobb:** Perces adatokból nyert indexérték-változás stochasztikus mátrixa

A 3.7 ábrán látható, hogy perces adattrendhez viszonyított érték esetén a grafikon szimmetrikus a főátlóra nézve. Ez, amint látni fogjuk, annak köszönhető, hogy az átlagot sok pontra vettük, fel, így a trendvonal helye nem változik jelentősen egy perc alatt, valamint az index értékének

percenkénti változása is jelentősen kisebb, mint a napi változása.



3.8. ábra. **Bal:** 60 perces átlagból számított trendvonalhoz viszonyított értékekből szerkesztett stochasztikus mátrix **Jobb:** 1000 perces átlagból számított trendvonalhoz viszonyított értékekből szerkesztett stochasztikus mátrix

A fenti ábrán látható, hogy amíg 1000 percre (kb. 2,5 nap) számolt átlaghoz képest vett értékek percnyi változása is megközelítőleg csak a változás nagyságától függ, addig 60 percnyi átlagra véve már látható egy csökkenés az origótól való távolságot figyelve, azaz függ az előző pillanatban az érték átlagától viszonyított értéktől.

A 60 perces átlagablakkal szerkesztett trendvonalhoz viszonyított érték stochasztikus mátrixa hasonlít a legjobban a teljes idősről vett napi adatokból számított stochasztikus mátrixhoz. Ennek az oka, hogy ebben az esetben megközelítőleg ugyanannyi adatpontból vettünk átlagot, mint a napi felbontású adatok esetén (abban az esetben 100 adatpontból vettünk), így az átlagból egy érték megváltoztatásának mindkét esetben nagyjából ugyanolyan súlya van az átlagától való eltérés változásában. Mivel a percenkénti változások kismértékűek, hasonlóan az 1980-as évek előtti napi változásokhoz, ezért megfelelő átlagablakkal dolgozva hasonló alakú stochasztikus mátrixokat kapunk. Nagyobb átlagablakot véve azonban már jobban hasonlítanak a 2007-es után vett napi adatokhoz, ugyanis a perces adatok változása két időpillanat között jóval nagyobb, mint az átlag változása.

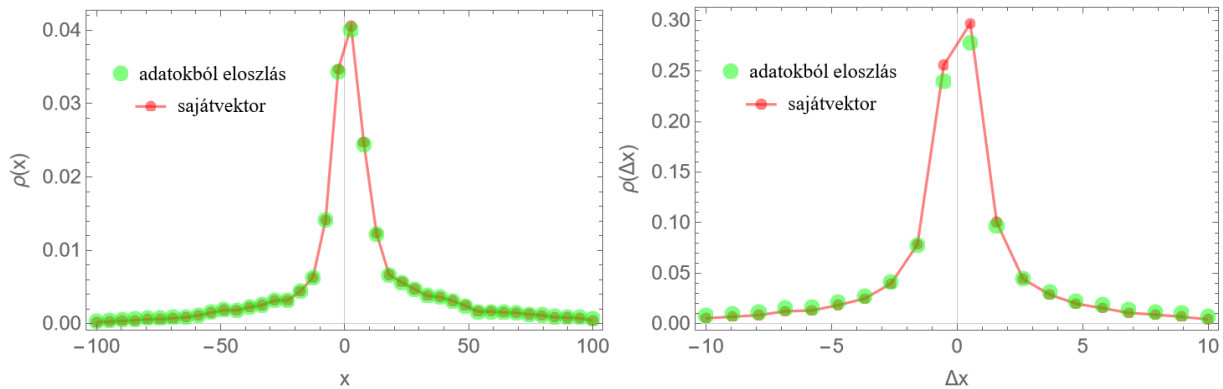
A változásokból felépített stochasztikus mátrix esetén az látható, hogy nagy változás esetén nem nő meg jelentősen az újabb nagy változás valószínűsége, mint a napos adatok esetén. Mivel kisebb időskálán vesszük a változásokat, ezért ezek értéke csak ritka esetben ér el egy nagy értéket.

3.2. Stacionárius eloszlás vizsgálata

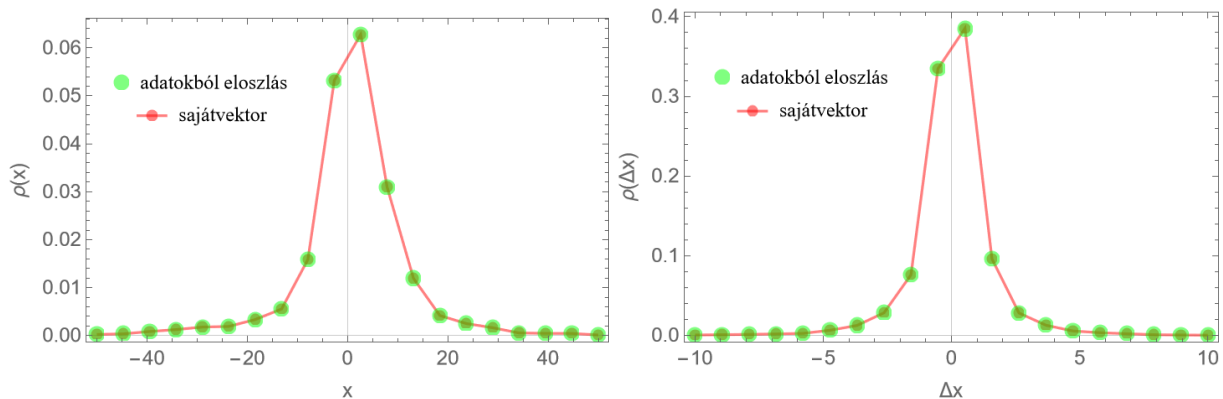
A stochasztikus mátrix ismeretében meghatározható, hogy a változás, illetve átlagától való eltérés értékeinek eloszlásai stacionárius eloszlást követnek-e. Ezek az 20-es egyenlet alapján a

mátrix baloldali sajátértékei, melyekből alkotott mátrix megadható, mint a jobboldali sajátvektorok által megadott mátrix inverze. Ebből kiszedhető az 1-es sajátértékhez tartozó baloldali sajátvektor. Ezután kiszámoljuk az adatsorból kapott eloszlást, majd a kettőt összehasonlítjuk.

3.2.1. Stacionárius eloszlás napos adatokra

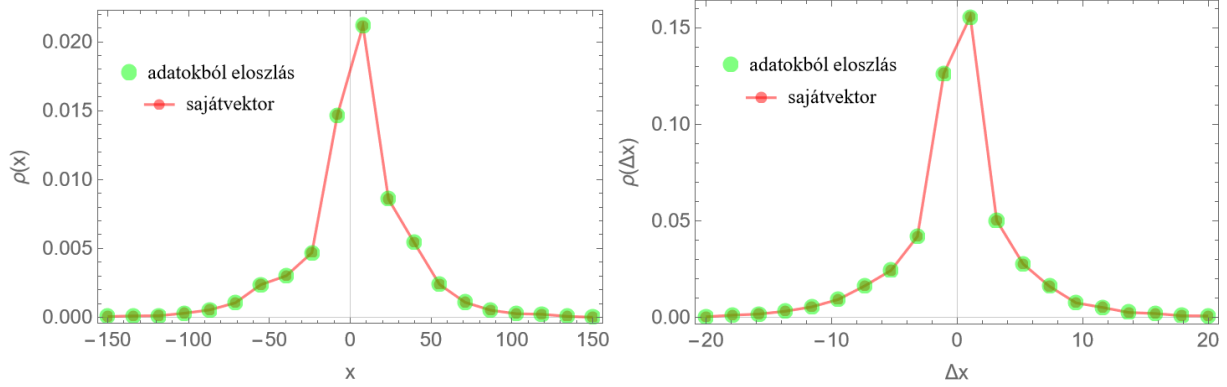


3.9. ábra. **Bal:** A napi adatokból nyert átlagtól való eltérésekből számolt sajátvektor (piros), és a napi adatokból számolt eloszlás (zöld) 1897-2022 között **Jobb:** a változásokból esetén számolt sajátvektor (piros) és az adatokból számolt eloszlás (zöld) a teljes időintervallumon

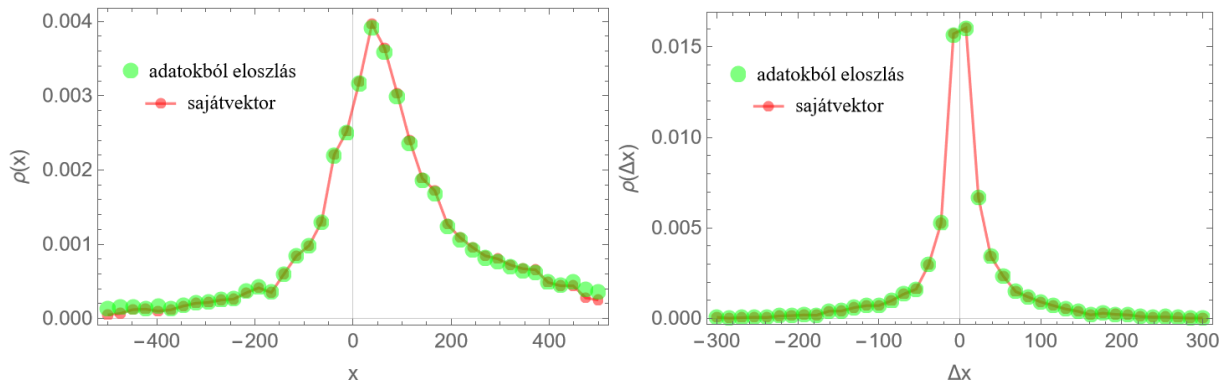


3.10. ábra. **Bal:** az átlagtól való eltérésekből számolt sajátvektor (piros), és az adatokból számolt eloszlás (zöld) 1940 előtti napi adatok esetén **Jobb:** a változásokból számolt sajátvektor (piros) és az adatokból számolt eloszlás (zöld) 1940 előtti napi adatok esetén

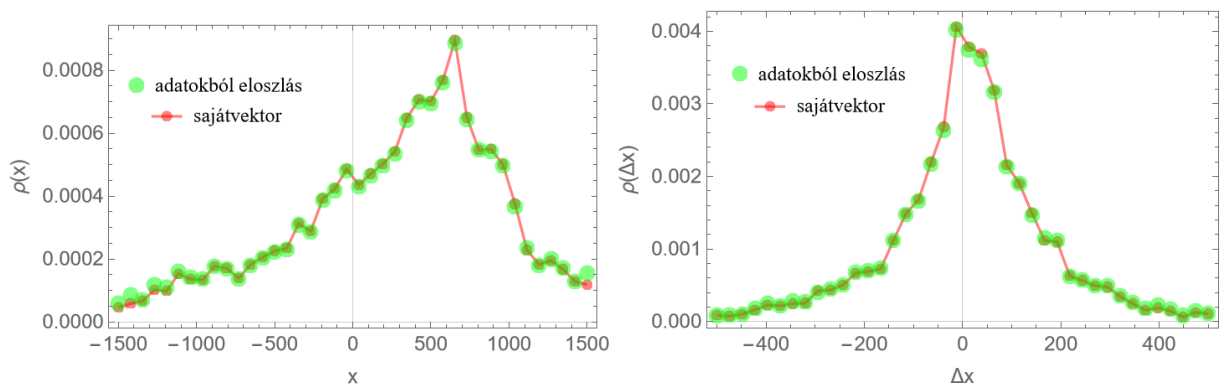
A következő ábrákon látható, hogy a stochasztikus mátrixból számolt eloszlás megegyezik az eredeti adatsorokból számolt eloszlással. Felfedezhető, hogy az 1980-as évek után változik az eloszlásfüggvény alakja, a 2007-es adatokban pedig a átlagtól való eltérésekből számolt eloszlás maximuma távol van a 0-tól. A kapott eloszlásokban minden esetben fellép egy asszimetria: a maximumaik értéke legtöbb esetben egy 0-hoz közeli pozitív érték. Ez a pozitív érték mindkét mutató esetén arra enged következtetni, hogy az indexérték az idő során nő: többször fordul elő pozitív irányú változás, mint negatív, valamint az indexérték kevésszer esik a mozgó átlag alá.



3.11. ábra. **Bal:** az átlagtól való eltérésekből számolt sajátvektor (piros), és az adatokból számolt eloszlás (zöld) 1940-1980 közti napi adatok esetén **Jobb:** a változásokból számolt sajátvektor (piros) és az adatokból számolt eloszlás (zöld) 1940-1980 közti napi adatok esetén

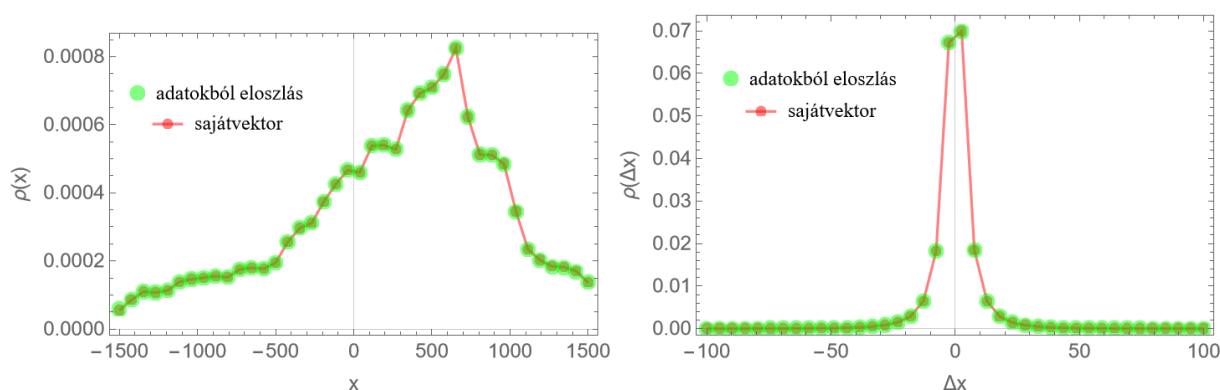


3.12. ábra. **Bal:** az átlagtól való eltérésekből számolt sajátvektor (piros), és az adatokból számolt eloszlás (zöld) 1980-2007 közti napi adatok esetén **Jobb:** a változásokból számolt sajátvektor (piros) és az adatokból számolt eloszlás (zöld) 1980-2007 közti napi adatok esetén



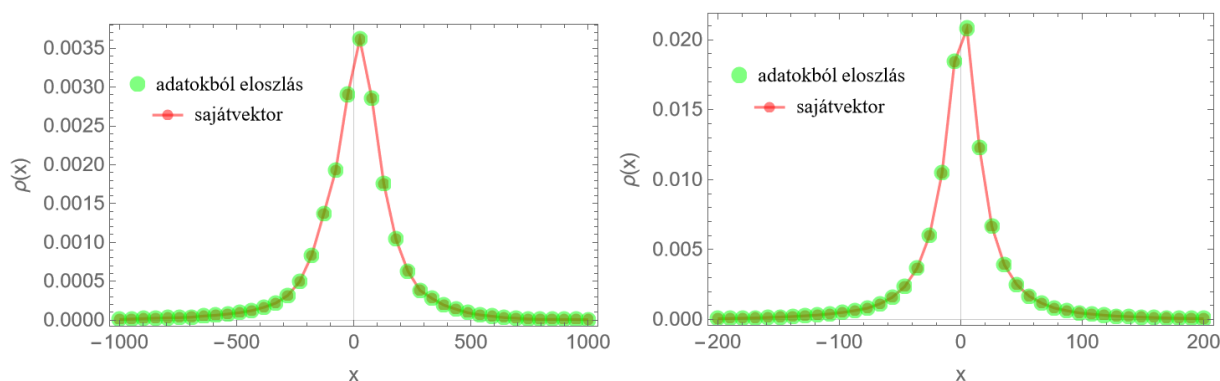
3.13. ábra. **Bal:** az átlagtól való eltérésekből számolt sajátvektor (piros), és az adatokból számolt eloszlás (zöld) 2007-2022 közti napi adatok esetén **Jobb:** a változásokból számolt sajátvektor (piros) és az adatokból számolt eloszlás (zöld) 2007-2022 közti napi adatok esetén

3.2.2. Eloszlások perces adatokra



3.14. ábra. **Bal:** az átlagtól való eltérések esetén számolt sajátvektor (piros), és az adatokból számolt eloszlás a perces adatokra, 40000 perces (100 napos) ablakkal **Jobb:** a változások esetén számolt sajátvektor (piros) és az adatokból számolt eloszlás (zöld) a perces adatokra

A fenti ábra alapján 40000 perces ablak alkalmazása esetén a trendvonaltól való eltérés értékének eloszlása hasonló a 2007-2022 közti napi adatokból vett eloszlásnak. A változások esetén látható, hogy a percenkénti változások eloszlása változott a napi értékekhez képest, ugyanis ezek jóval kisebbek, mint a napos különbségek.



3.15. ábra. **Bal:** az átlagtól való eltérések esetén számolt sajátvektor (piros), és az adatokból számolt eloszlás (zöld) 1000 perces átlagablakra **Jobb:** az átlagtól való eltérések esetén számolt sajátvektor (piros) és az adatokból számolt eloszlás (zöld) perces adatokból 60 perces átlagablakra

Kisebbségi átlagok esetén az átlagérték közelebb található az aktuális értékhez, így a nagy ablakkal ellentétben, ez látható a két eloszlásból is, melyeken a legvalószínűbb átlagtól való eltérés értéke 0-hoz közelebb. Az asszimetria itt is nyomon követhető, nagyobb valószínűsége van a nagyobb változásnak és a trendvonal feletti értékeknek.

3.3. Függvény illesztése a stochasztikus mátrixra

A következőkben a stochasztikus mátrix felhasználásával egy analitikus modellt keresünk, melyből következtethetünk arra, hogy az index értékének változása milyen ismert

véletlenszerű folyamatokkal közelíthető meg. Ehhez a kapott stochasztikus mátrixokra folytonos függvényt illesztünk, és ahol lehetséges, ennek segítségével határozzuk meg az adott mutató eloszlásfüggvényét.

3.3.1. Az átlaghoz viszonyított értékekre megszerkesztett stochasztikus mátrixok analitikus megközelítése

A napi adatokból származó átlagtól való eltérésekből szerkesztett stochasztikus mátrixok esetén az látható, hogy az értékek a főátló mentén csoportosulnak, azonban észlelhető egy függés az origótól való távolságtól is. A stochasztikus mátrixra illesztett függvényt megfelelően normálni is akarjuk.

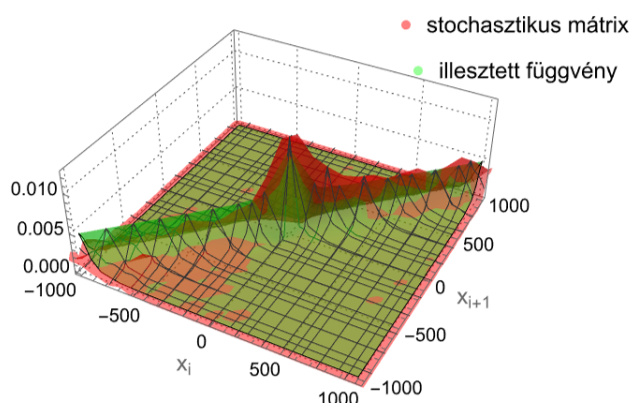
Ez alapján azt feltételeztük, hogy a függvénynek van egy komponense, mely a két átlagtól való eltérés közti különbségtől függ, ennek köszönhetően csökken az átmeneti valószínűség a főátlótól távolodva. Mivel az átmeneti valószínűség a kapott ábra alapján az origótól való távolsággal is csökken, ezért ezt is figyelembe vesszük. A (22) normálási feltételt figyelembe véve a következő függvényt illesztjük:

$$w(x, y) = C(x) \cdot \frac{1}{a + (x - y)^2} \cdot \frac{1}{b + x^2 + y^2} \quad (26)$$

A $C(x)$ a normálási tényező ebben az esetben:

$$C(x) = \pi \frac{-a + b + 2x^2 - b\sqrt{\frac{a}{b+x^2}} + \frac{a^{\frac{3}{2}}}{\sqrt{b+x^2}}}{\sqrt{a}(a^2 - 2ab + (b + 2x^2)^2)} \quad (27)$$

Ezt a függvényt ráillesztettük az összes napi adatból meghatározott, az átlaghoz viszonyított értékeket -1000-től 1000-ig ábrázoló függvényre.

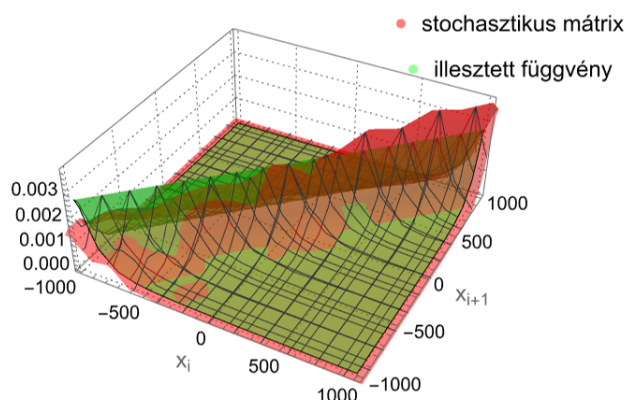


3.16. ábra. Analitikus megközelítés az átlagtól való eltérésekből szerkesztett stochasztikus mátrixra. A használt paraméterek: $a = 4210$, $b = 2800$.

Az illesztett függvény első tagja önmagában egy olyan folyamatot ír le, melyben egy, a trendvonalhoz viszonyított értékbe való átmeneti valószínűség nem függ az előző időpillanatbeli értéktől, csak a két érték közti különbségtől. A teljes idősorra kapott napos adatok leírására azonban szükség van a második tagot is bevezetni, mivel nagyobb értékről a következő időpillanatban hajlamosabb nagyobbat ugrani, mint kisebb értékről. Ez azzal magyarázható, hogy az index értéke és a trendtől való eltérése a 80-as évekig kicsi volt, és nem történtek nagy értékváltozások. A 2007-2022 közti napos adatok stochasztikus mátrixának a maximuma már nem az origóban található és nem tér el annyira a mátrix többi értékétől.

Erre a mátrixra mutató értékének változásától függő tagot illesztettünk:

$$w(x, y) = \frac{\sqrt{a}}{\pi} \cdot \frac{1}{a + (x - y)^2} \quad (28)$$



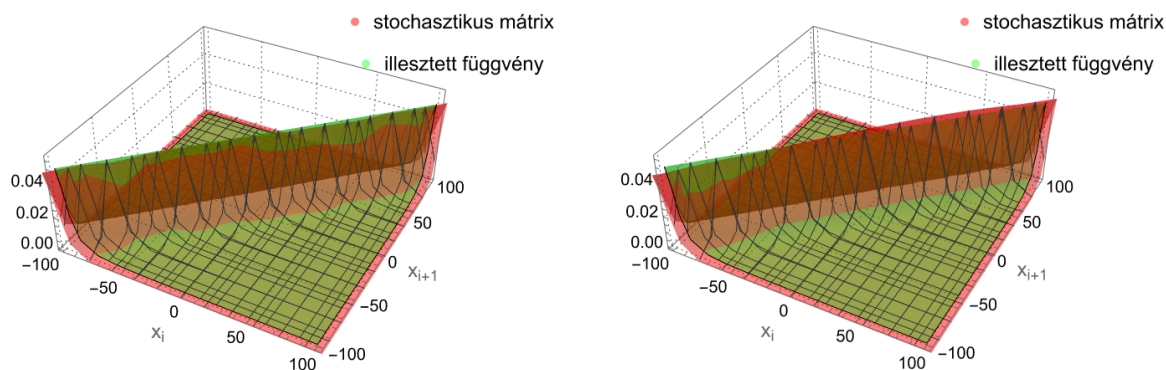
3.17. ábra. A megadott függvény illesztése a 2007-2022 közti napos adatokból nyert, átlagtól való eltéréseiből szerkesztett stochasztikus mátrixra. A használt paraméterek: $a = 13500$.

Perces adatok esetén láthattuk, hogy az 1000 és 40000 percre vett átlagok ablaka esetén a stochasztikus mátrix elemei nagyrészt a főátlóra koncentrálódnak, így ezekre is mutató változásától függő tagból származó függvényt illesztjük:

Perces adatok esetén mind az index értéke, mind az átlag értéke lassan változik egy időlépés alatt, így percenként az átlagtól való eltérés értéke nagy valószínűséggel keveset változik.

3.3.2. A változásokra megszerkesztett stochasztikus mátrix analitikus megközelítése

A változásokra szerkesztett stochasztikus mátrixokról megállapítható, hogy az y tengely (későbbi időpillanat) mentén az x tengelytől (korábbi időpont) távolodva gyorsan, az x tengely mentén pedig lassabban csökken. A függvényt így ismét két tag szorzataként keressük: az egyik az y irányban csökkenti a függvényt, a másik pedig szimmetrikusan az origótól távolodva:



3.18. ábra. **Bal:** A megadott függvény illesztése a perces adatokból készített, átlaghoz viszonyított értékekből szerkesztett stochasztikus mátrixra 40000-res ablakra való átlagolás esetén. Az a paraméter értéke 42. **Jobb:** A megadott függvény illesztése a perces adatokból készített, átlaghoz viszonyított értékekből szerkesztett stochasztikus mátrixra 1000-res ablakra való átlagolás esetén. Az a paraméter értéke 41.

$$w(x, y) = C(x) \cdot \frac{1}{a + y^2} \cdot \frac{1}{b + x^2 + y^2} \quad (29)$$

A normálási tényező:

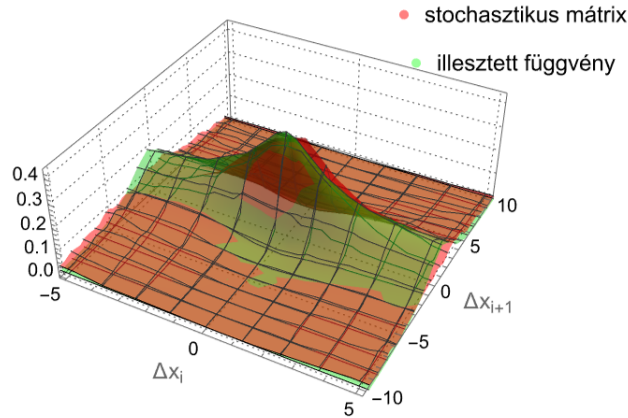
$$C(x) = \frac{\sqrt{a}(b + x^2 + \sqrt{a(b + x^2)})}{\pi} \quad (30)$$

A függvény első tagja a klasszikus véletlenszerű mozgásokhoz tartozó tag, mely csak a változás értékétől függ, és azt az esetet reprezentálja, mikor egy adott indexértékről való ugrás valószínűsége csak az ugrás nagyságától függ, és nem befolyásolják az addig történt ugrások. A második tag jelenléte kölcsönzi az index bizonyos értelemben tekintett nem Markovi jellegét, ugyanis a tag figyelembe vételekor az ugrás nagysága függ az előző ugrás nagyságától is, vagyis az index értéke egy adott pillanatban függ az előző két értékétől.

A napi adatokból számolt különbségfüggvény:

Ez a függvény akkor írja le jól a stochasztikus mátrixot, mikor a korai indexértékek vannak túlsúlyban, ezek alacsony értékeinél ugyanis jelentősen nagyobb a valószínűsége annak, hogy kis változás után megint kis változás következik. Az index frissebb értékeinél nagyobb ingadozás figyelhető meg, így ebben az esetben a nagyobb változásoknak is megnő a valószínűsége, így az előbb illesztett függvény x tengelye menti csökkenés ütemje nem írja le a mátrix alakját.

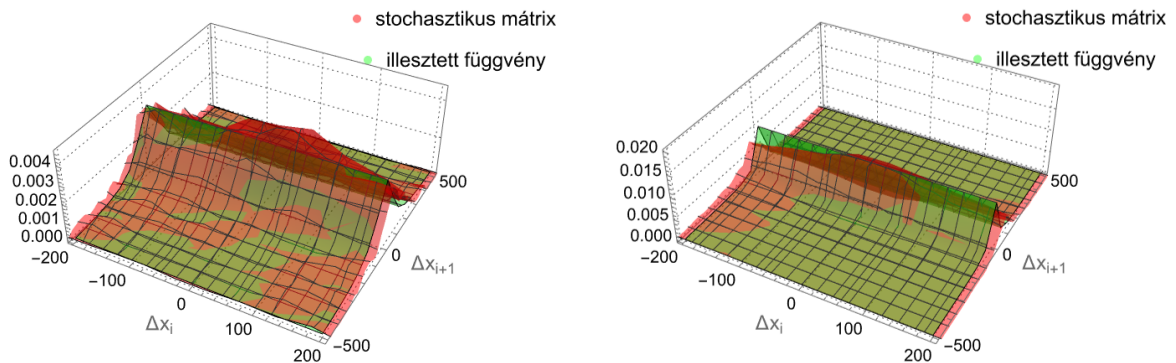
A 2007-esnél frissebb napos és perces adatok esetén a kisebb változásokra (100-ig) illesztünk függvényt, ugyanis itt még az x menti csökkenés értéke elhanyagolható. Ez alapján csak a változást tartalmazó függvényi alakot illesztjük a mátrixra:



3.19. ábra. A megadott függvény illesztése a változásokból szerkesztett stochasztikus mátrixra. A használt paraméterek: $a = 8$, $b = 1$.

$$w(x, y) = \frac{\sqrt{a}}{\pi} \cdot \frac{1}{a + y^2} \quad (31)$$

Ábrázolva a 2007 utáni adatokat a következő ábrákat kapjuk:



3.20. ábra. **Bal:** 2007 utáni napos adatokból nyert változások stochasztikus mátrixra illesztett függvény. Az a paraméter értéke 8000. **Jobb:** Perces adatokból nyert illesztett függvény, a értéke 1000.

Mivel ezekben az esetekben a stochasztikus mátrixra illesztett függvény csak az y értékétől függ, ezért a 25-es egyenlet így alakul:

$$\rho(y) = \int_{-\infty}^{\infty} w(y)\rho(x)dx \quad (32)$$

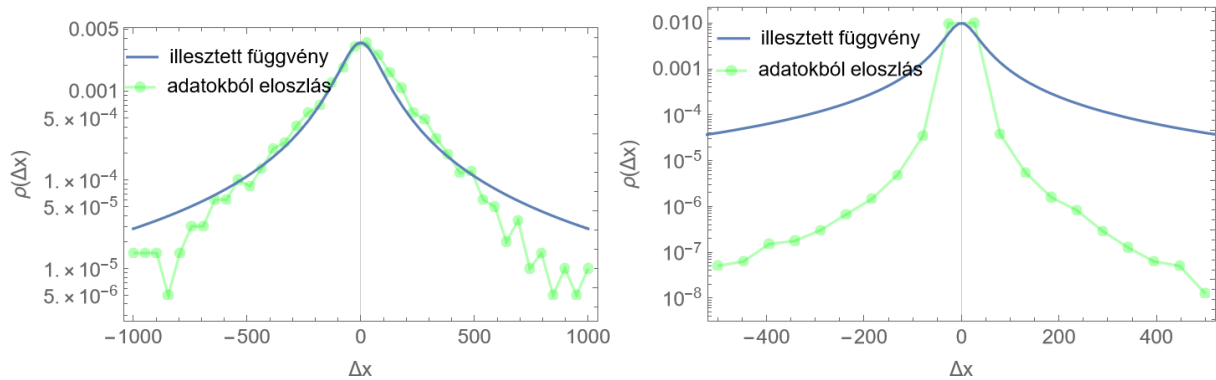
Ez tovább alakítható:

$$\rho(y) = w(y) \int_{-\infty}^{\infty} \rho(x)dx \quad (33)$$

Mivel a sűrűségfüggvény normált, azaz $\int_{-\infty}^{\infty} \rho(x)dx = 1$, ezért a következőt kapjuk:

$$\rho(y) = w(y) \quad (34)$$

Tehát a stochasztikus mátrixra illesztett függvény értéke megegyezik az eloszlásfüggvénnyel. Az illesztett függvény paraméterét ismerve összehasonlítjuk az így meghatározott eloszlásfüggvényt az adatokból megszerkesztett eloszlásfüggvényre:



3.21. ábra. **Bal:** 2007 utáni napos adatokból nyert eloszlás és stochasztikus mátrixra illesztett függvény összehasonlítása **Jobb:** Perces adatokból nyert eloszlás és stochasztikus mátrixra illesztett függvény összehasonlítása

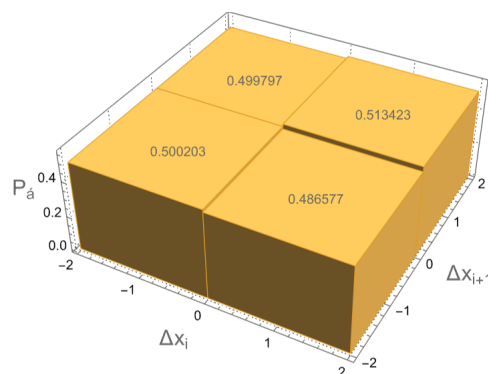
Látható, hogy az illesztett függvény napos, és perces adatokra is csak nagyon kis különbségekre adja vissza az adatokból számolt eloszlásfüggvényt. Ez tehát nem igazán járható út arra, hogy egyszerűen megközelítsük a stacionárius eloszlásokat. Ez alapján tehát a nagyobb változások esetén a következő változás nagysága nem lesz független az előző változás nagyságától, így az index értékének változásának leírásához szükséges egy bizonyos értelemben nem-markovi tagot is figyelembe kell veyünk.

4. Stochasztikus mátrix alkalmazásra az indexérték alakulásának előrejelzésére

Mivel az tőzsdeindexek értékének alakulása a gazdaság alakulását is tükrözi, ezért ennek az előrejelzése alapján következtethetünk egyéb idősorok, mint például valutaárfolyamok alakulására is. Az indexérték előrejelezhetőségét vizsgálva így előrevetíthetjük, hogy mekkora eséllyel térül meg bizonyos befektetés. Ebben a fejezetben azt vizsgáljuk meg, hogy az index értékének perces alakulásából milyen pontossággal lehet előrejelezni arra, hogy az index értéke egy adott időpillanatban nőni vagy csökkenni fog. Megvizsgáljuk, hogy mennyi ideig optimális követni az adatokat, hogy minél pontosabban következtethessünk az index értékének az alakulására.

4.1. Használt módszer

Egy adott t időpillanatban való tippeléshez a következőképp járunk el: ha előtte n perccel kezdjük el követni az index értékének alakulását (legyen ez a t_0 időpillanat), akkor megszerkesztjük a t_0 és t időpillanatok közti percenkénti különbségekből a 2×2 -es stochasztikus mátrixot. Ennek a négy eleme kifejezi, hogy mekkora valószínűséggel követ egy csökkenést növekedés vagy csökkenés, illetve egy növekedést növekedés vagy csökkenés. Egy példa a stochasztikus mátrixra a 4.1 ábrán található (P_a -val jelöltük az átmeneti valószínűséget).



4.1. ábra. Stochasztikus mátrix 10000-res időablakkal számítva, 2015 november 13-án 13:21-kor

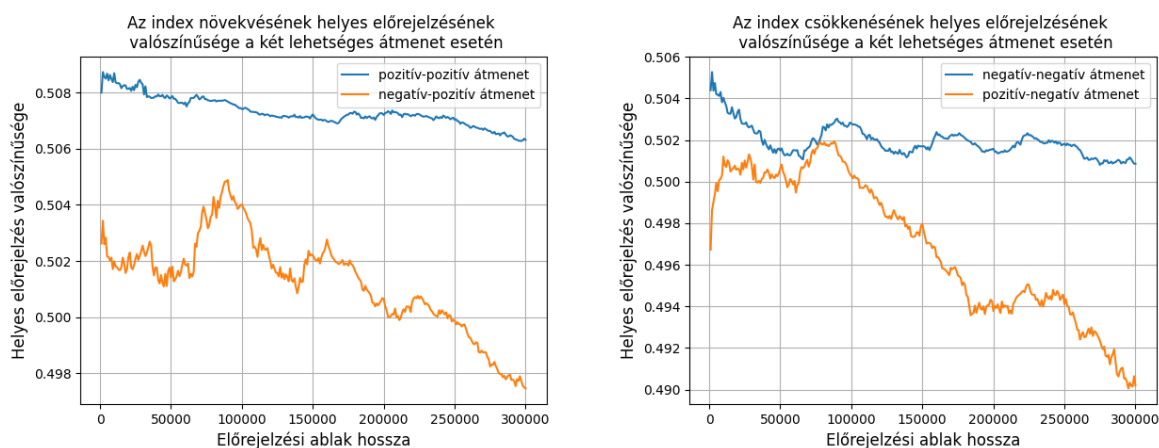
Ezután, attól függően, hogy az adott időpillanat előtt csökkent vagy nőtt az index értéke, megvizsgáljuk, hogy a stochasztikus mátrix szerint csökkenésnek vagy a növekedésnek nagyobb a valószínűsége, és a nagyobb valószínűségű folyamatra tippelünk.

Ezt elvégezzük minden időpillanatra, és megnézzük, hogy az esetek hányadában volt helyes a tippünk. Megvizsgáljuk, hogy a négyféle lehetőséget (csökkenést csökkenés, csökkenést növekvés, növekvést növekvés illetve növekvést csökkenés követ) milyen pontosan tippelhetjük meg ezzel

a módszerrel. Ezután változtatjuk a megfigyelt időpontok mennyiségét, majd megnéztük, hogy a négyféle lehetőség esetén mennyi ideig optimális az indexet megfigyelni.

4.2. Eredmények

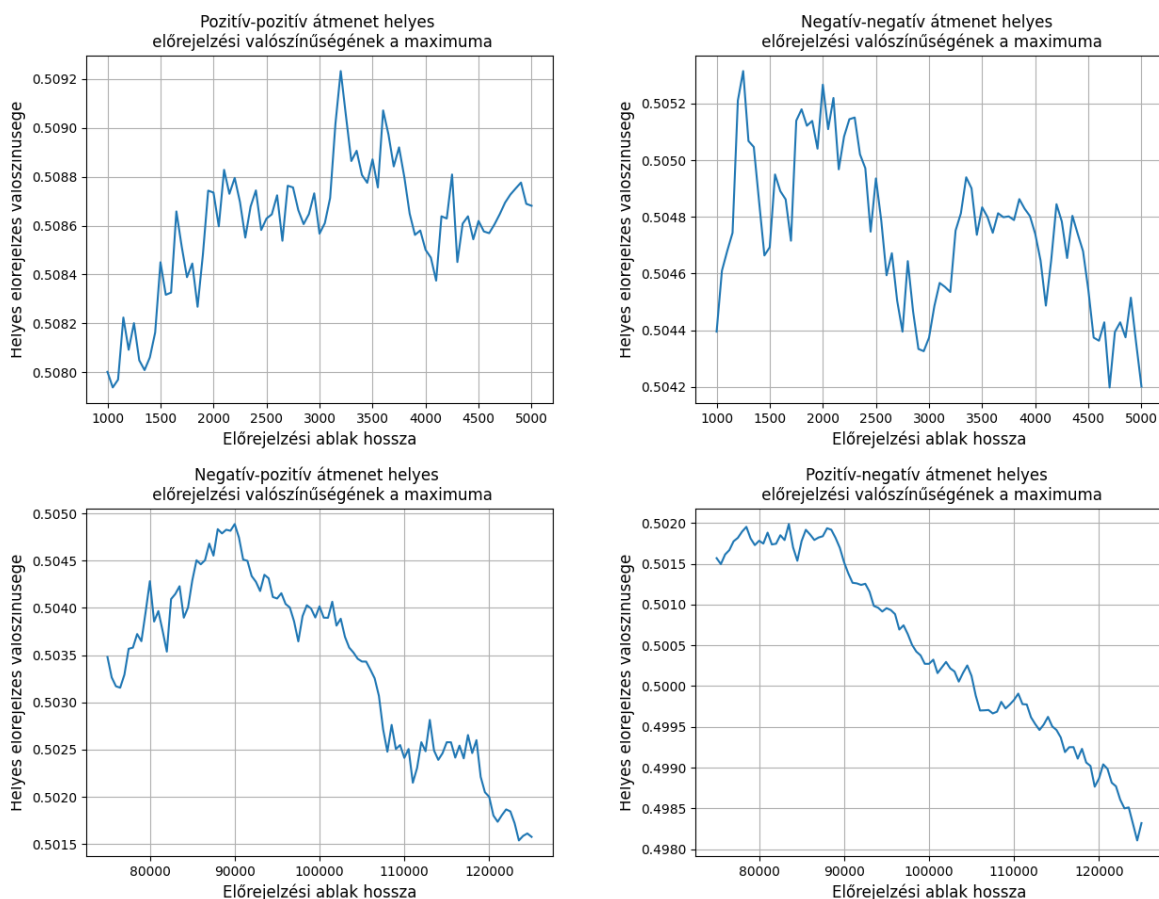
A következő ábrákon azt ábrázoljuk, hogy miként változik a megfigyelt ablak nagyságának a változtatásával az előrejelzés helyessége a négyféle folyamat esetén:



4.2. ábra. **Bal:** Növekedés helyes előrejelzésének valószínűsége kétféle átmenet esetén **Jobb:** Csökkenés helyes előrejelzésének valószínűsége kétféle átmenet esetén

Látható, hogy kisebb (pár napos) időablak optimális akkor, mikor a trend folytatását, és nagyobb (megközelítőleg egy éves) időablak optimális, ha a trend megfordulását tippeljük. Látható az is viszont, hogy mind a növekedést, mind a csökkenést bármilyen időablak esetén könnyebb megtippelni a trendet tartó esetben, valamint a növekedést a trenddel megegyező, illetve attól eltérő esetben is könnyebb előrejelezni, mint a csökkenést.

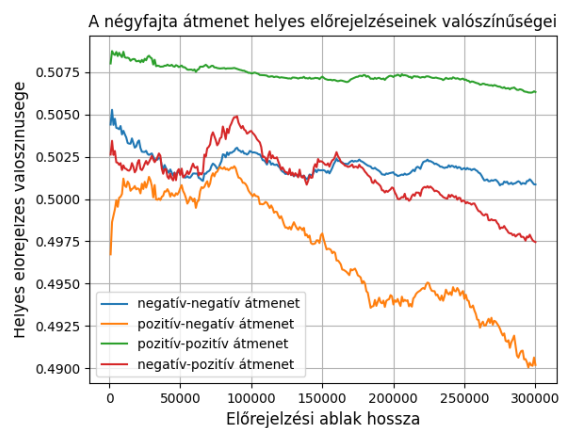
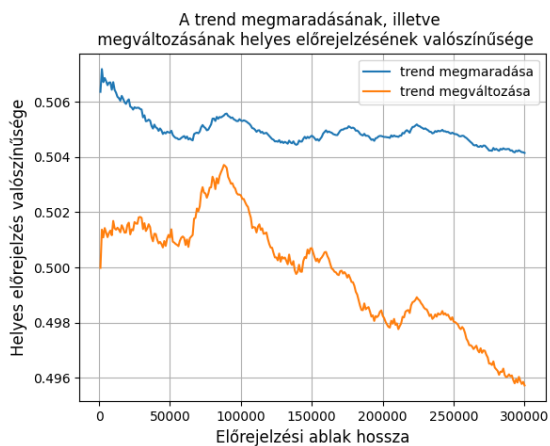
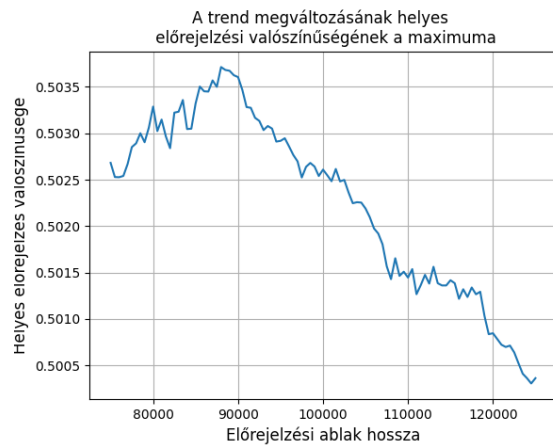
A pontosabb időablak meghatározása érdekében nagyobb felbontással ábrázoljuk a négy esetben az eredményeket a maximumok körül:



4.3. ábra. **Bal felső:** Pozitív-pozitív átmenet helyes előrejelzésének valószínűségének maximuma **Jobb felső:** Negatív-negatív átmenet helyes előrejelzésének valószínűségének maximuma **Bal alsó:** Pozitív-negatív átmenet helyes előrejelzésének valószínűségének maximuma **Jobb alsó:** Pozitív-negatív átmenet helyes előrejelzésének valószínűségének maximuma

A pozitív-pozitív előrejelzés helyességének a maximuma 3000 perces ablak körül 50,9 %-os, a negatív-negatív előrejelzés maximuma 50,5 % körüli, és 1200-2000-res ablak között található, a negatív-pozitív előrejelzés maximuma 50,45% körüli, és a 90000-res ablaknál található, a pozitív-negatív előrejelzés maximuma 50,2 %, 85000-90000-res ablak körül található.

A következő ábrán látható, hogy a trenddel való tartás esetén a helyes előrejelzés valószínűségének maximuma 50,7 %, és a 2000-res ablaknál található, a trenddel szembeni változtatás esetén a maximum 50,35 %, és a 90000-es ablak körül található. Az alábbi ábrán azonban látható, hogy minden időablaknál a pozitív-pozitív átmenet a legelőrejelezhetőbb, valamint a trenddel való tartás előrejelezhetőbb, mint a trend megváltozása. Megjegyzendő azonban, hogy mindegyik átmenetkor bármilyen időablakra az index értékének alakulása csak kismértékben jelezhető elő, így az indexbe befektetve csak hosszú távon, nagy mennyiségű alaptőkével rendelkezve termelhetünk nyereséget.



4.4. ábra. **Bal felső:** Trenddel való tartás helyes előrejelzésének valószínűségének maximuma **Jobb felső:** Trend megváltozásának helyes előrejelzésének valószínűségének maximuma **Bal alsó:** Trend megváltoztatása és trenddel való tartás előrejelezhetősége **Jobb alsó:** A négyféle átmenet előrejelezhetőségének összehasonlítása

Következtetések

Dolgozatunk első felében bemutattuk a gazdasági fizikában használt statisztikus fizikai modelleket, megközelítéseket, hangsúlyt helyezve az idősorok modellezésére. Ezek segítségével általános törvényszerűségeket állapíthatunk meg a gazdasági idősorokra vonatkozóan.

Ezt követően bemutattunk egy, a statisztikában sokat használt eszköz, stochasztikus mátrix alkalmazását egy tőzsdeindexből származó adatok, jelen esetben az időpillanatok közti értékváltozás és az egyes időpillanatokban az index átlaghoz viszonyított értékének dinamikájának a tanulmányozására. A stochasztikus mátrix felhasználásakor feltételeztük, hogy ezen mutatók változása diszkrét időlépésű Markov-folyamatként tekinthető. A dolgozat során láttuk, hogy a megszerkesztett stochasztikus mátrix baloldali sajátértékei visszaadják az adatokból szerkesztett eloszlásfüggvényt. Az eloszlásfüggvény alakjában felfedezhettünk egy asszimmetriát, az eloszlás maximuma mindig pozitív értéktartományban található, mely az indexértékek hosszú távon való növekedéséhez vezet.

A stochasztikus mátrixra analitikus megközelítéseket tettünk, melyeknek mind a trendtől való eltérések, mind az index értékére számított stochasztikus mátrix esetén két tagja volt. Az átlaghoz viszonyított értékek esetén az első tag csak a két időpillanatban az index átlaghoz viszonyított értékének változásától függött, valamint emellett egy másik tag, melyen keresztül függött az előző értéktől is. Perces felbontású adatok esetén, lassan változó átlagot és indexértéket tekintve a stochasztikus mátrixra illeszthető volt az átlaghoz viszonyított indexérték változását tartalmazó tagból álló függvény.

A változások vizsgálata során az első tag csak az adott időpillanat utáni változástól függött, mely így egy tipikusan Markovi tag. Csak ezt a tagot tekintve a stochasztikus mátrixot leíró függvény egyben a változások eloszlásfüggvénye is. Az így kapott függvény az eredeti eloszlásra csak kis értékekre illeszkedett, így az indexértékek alakulását nem tekinthetjük klasszikus értelemben vett Markov-folyamatnak.

Az utolsó fejezet keretében a stochasztikus mátrix segítségével előre próbáltuk jelezni az index értékeinek az alakulását. A kidolgozott módszerünk által az index trendjét (növekvő/ csökkenő) kevéssel több, mint 50 % valószínűséggel tudtuk előrejelezni, a legnagyobb pontossággal a pozitív trend megtartását lehet előrejelezni. Ez alapján az indexbe egy pozitív trend során érdemes befektetni, azonban jelentős nyereség nem származik ebből. Ezen felül ahhoz, hogy ténylegesen hozzájussunk a nyereségünkhöz, a megvásárolt befektetést el is kell adni, melyet negatív változás előtt érdemes. Azonban ennek az előrejelezhetősége még kisebb, mint a pozitív változásnak, így az így nyert nyereség értéke tovább csökken. Figyelembe véve az általunk nem tárgyalt befolyásoló

tényezőket, mint a tranzakciók kezelési költsége, vagy a nyereség után fizetett adók, az alapba való befektetés során keletkezett nyereség elhanyagolható értékű lesz a kezdőtőkénkhez képest.

A fentiek alapján arra jutottunk, hogy habár a Markov folyamatokra jellemző stochasztikus mátrixot használtuk, az index értékének a változása nem egy tökéletesen Markovi folyamatként tekinthető, az index értékének változása függ ugyanis az előtte történő változástól is, tehát egy az index értékének nagysága egy adott időpillanatban függ az előző két időpillanatban felvett értékétől. A trendhez viszonyított indexérték és a változás nagyságának eloszlásának maximuma világosan a pozitív irányba tolódik, mely következtetni enged az indexérték hosszútávon történő növekedésére is.

Hivatkozások

- [1] J.P. Huang. „Experimental econophysics: Complexity, self-organization, and emergent properties”. *Physics Reports* 564 (2015). Experimental econophysics: Complexity, self-organization, and emergent properties, 1–55. oldal. ISSN: 0370-1573. DOI: <https://doi.org/10.1016/j.physrep.2014.11.005>. URL: <https://www.sciencedirect.com/science/article/pii/S0370157314004074>.
- [2] FRANCK JOVANOVIĆ és CHRISTOPHE SCHINCKUS. „ECONOPHYSICS: A NEW CHALLENGE FOR FINANCIAL ECONOMICS?": *Journal of the History of Economic Thought* 35.3 (2013), 319–352. oldal. DOI: 10.1017/S1053837213000205.
- [3] Bikas Chakrabarti. „Econophys-Kolkata: A Short Story”. 2005. jan., 225–228. oldal. ISBN: 978-88-470-0329-3. DOI: 10.1007/88-470-0389-X_26.
- [4] Siew Ann Cheong. „Economics and econophysics in the era of Big Data”. *The European Physical Journal Special Topics* 225 (2016), 3159–3170. oldal.
- [5] Rosario N. Mantegna és H. Eugene Stanley. *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, 1999. DOI: 10.1017/CB09780511755767.
- [6] W Paul Cockshott és mások. *Classical econophysics*. 2. kötet. Routledge, 2009.
- [7] Sitabhra Sinha és mások. *Econophysics: an introduction*. John Wiley & Sons, 2010.
- [8] Bikas K Chakrabarti és mások. *Econophysics of income and wealth distributions*. Cambridge University Press, 2013.
- [9] Frédéric Abergel és mások. *New perspectives and challenges in econophysics and sociophysics*. Springer, 2019.
- [10] Anirban Chakraborti és mások. „Econophysics review: I. Empirical facts”. *Quantitative Finance* 11.7 (2011), 991–1012. oldal.
- [11] Victor M Yakovenko és J Barkley Rosser Jr. „Colloquium: Statistical mechanics of money, wealth, and income”. *Reviews of modern physics* 81.4 (2009), 1703. oldal.
- [12] Frank Schweitzer és mások. „Economic Networks: The New Challenges”. *Science* 325.5939 (2009), 422–425. oldal. DOI: 10.1126/science.1173644. eprint: <https://www.science.org/doi/pdf/10.1126/science.1173644>. URL: <https://www.science.org/doi/abs/10.1126/science.1173644>.
- [13] Andrew G Haldane és Robert M May. „Systemic risk in banking ecosystems”. *Nature* 469.7330 (2011), 351–355. oldal.
- [14] Paulo Ferreira, Éder JAL Pereira és Hernane BB Pereira. „From big data to econophysics and its use to explain complex phenomena”. *Journal of Risk and Financial Management* 13.7 (2020), 153. oldal.

- [15] Michael Pickhardt és Goetz Seibold. „Income tax evasion dynamics: Evidence from an agent-based econophysics model”. *Journal of Economic Psychology* 40 (2014), 147–160. oldal.
- [16] Bernd Rosenow és mások. „Portfolio optimization and the random magnet problem”. *Europhysics Letters* 59.4 (2002), 500. oldal.
- [17] URL: <https://stooq.com/q/d/?s=%5C%5Edji>.
- [18] URL: <https://firstratedata.com/i/index/DJI>.
- [19] Dattatray P Gandhmal és K Kumar. „Systematic analysis and review of stock market prediction techniques”. *Computer Science Review* 34 (2019), 100190. oldal.
- [20] Bulcsú Sándor és mások. „Time-scale effects on the gain-loss asymmetry in stock indices”. *Physical Review E* 94.2 (2016), 22311. oldal.
- [21] Geoffrey Grimmett és David Stirzaker. *Probability and random processes*. Oxford university press, 2020.
- [22] Ingve Simonsen, Mogens H Jensen és Anders Johansen. „Optimal investment horizons”. *The European Physical Journal B-Condensed Matter and Complex Systems* 27 (2002), 583–586. oldal.
- [23] Emeric Balogh és mások. „Persistent collective trend in stock markets”. *Physical Review E* 82.6 (2010), 66113. oldal.
- [24] Jean-Philippe Bouchaud, Andrew Matacz és Marc Potters. „Leverage effect in financial markets: The retarded volatility model”. *Physical review letters* 87.22 (2001), 228701. oldal.
- [25] Peter Toke Heden Ahlgren és mások. „Frustration driven stock market dynamics: Leverage effect and asymmetry”. *Physica A: Statistical Mechanics and its Applications* 383.1 (2007), 1–4. oldal.