FOGALMAK, MÓDSZEREK

A WAVELETEKRŐL[1]

HUNYADI LÁSZLÓ
Budapesti Corvinus Egyetem

A wavelet elemzést, melynek alapjait a Haar-féle függvényrendszer és az azzal történő függvényközelítés adta meg, elsődlegesen a változó hullámhosszú folytonos jelek és diszkrét idősorok hullámtermészetének vizsgálatára fejlesztették ki. A cikk ezen alapok felvázolása után bemutatja a kiinduló folytonos wavelet transzformációt, azt, hogy a számítástechnika miként járult hozzá a diszkrét wavelet transzformáció (DWT) és annak iteratív algoritmusai kialakulásához. Részletesen foglalkozik a népszerű piramis algoritmus egy egyszerű változatával, de a fő hangsúlyt a gazdasági – társadalmi modellezés alkalmazási lehetőségeire helyezi. Ezek keretein belül bemutatja a DWT alkalmazását idősorok hullámtulajdonságainak elemzésében, az információtömörítésben, az outlier szűrésben, az idősorok hasonlóságának vizsgálatában és az idősorok fordulópontjainak detektálásában. Végkövetkeztetése az, hogy a terület igen nagy lehetőségeket ígér az idősoros és területi soros alkalmazások terén, de ennek megvalósításához az érintett szakterületek (matematika, statisztika, modellezés, informatika) együttműködésére van szükség.

Bevezetés

A waveletek (kis hullámok) története sajátos, szinte regényes, de a tudományban nem egyedülálló. Sok tudományterületen találkozhattunk már olyannal, hogy egy kutatás a melléktermékeiről lett híres, hogy az eredeti célt nem érte ugyan el, ehelyett egy másodlagos eredmény legitimálta[2]. Így van ez a waveletekkel is: egy viszonylag egyszerű, általános, nem szakirányú matematikai probléma megoldása ötletet adott egy, a fizikai jelek hullámtermészetének vizsgálati módszerére. A számítógépek fejlődésével előtérbe került a probléma diszkrét (digitalizált) kezelése, az erre szolgáló hatékony algoritmusok kidolgozása és gépi megvalósítása. Ekkor már a diszkrét jelek (idősorok) alkalmazói is érdeklődéssel fordultak a waveletek felé, de az igazi fordulatot az alkalmazásokban az jelentette, hogy az informatikusok felismerték, milyen hatékony eszközt kapnak nagy hang- és képfájlok veszteség nélküli, ill. veszte-

---

[1] A szerző a Budapesti Corvinus Egyetem nyugalmazott egyetemi tanára. E-mail: `hunyadi44@gmail.com`. Beérkezett: 2019. január 4.

[2] A waveletek evolúciójáról jó leírás található Hubbard [1998] könyvében.

séges (de gazdaságos) tömörítésére. Ez pedig, a rohamosan fejlődő technikai eszközök lehetőségeinek kihasználása szempontjából kritikus fontosságú volt.

Mostanra a waveleteken alapuló elemzések elérték a társadalom- és közgazdaságtudományi alkalmazásokat is. Elsősorban a waveletek diszkrét kezelése, valamint a rohamosan növekvő adatmennyiség vezetett oda, hogy egyre több alkalmazás született és születik ezen a területen is. Ez indokolja számunkra azt, hogy a gazdaság- és társadalomtudományi szakmát megismertessük a waveletek elméletének alapjaival, és beszámoljunk e gyorsan fejlődő terület adta alkalmazási lehetőségekről.

Világszerte intenzív kutatások folynak a waveletek elméletét és főleg alkalmazását illetően. Természetesen Magyarországon sem más a helyzet, ám itthon főként matematikusok foglalkoznak ezzel a témával. Ezek és a kapcsolódó kutatások – melyek nagy része természetesen messze túlmutat e tanulmányon – jó áttekintése megtaláltató Schipp átfogó matematikai munkájában (Schipp [2015]).

Ez a cikk bevezető a waveletek elméletébe és alkalmazásába. Először bemutatja a kiinduló matematikai problémát és annak megoldását, majd a spektrálanalízis analógiájából származtatja a folytonos wavelet transzformációt. Ezt követően a számítógépes applikációhoz szükséges diszkrét átmenetet tárgyalja, majd bemutat egy egyszerű algoritmust a diszkrét wavelet transzformációra. A cikk talán leglényegesebb fejezete a gazdasági és társadalmi idősor-modellezési, statisztikai alkalmazásokat mutatja be, jellegzetes példákkal. A dolgozatot a következtetések és a felhasznált szakirodalom jegyzéke zárja.

1 A kezdetek

A waveletek történetét két ágról lehet indítani. Az egyik a matematikai indíttatású megközelítés. Haar Alfréd szegedi matematikus a XX. század elején definiált egy lépcsős alapfüggvényt,

$$\xi(x)=\begin{cases}1, & \text{ha } 0\le x<1/2,\\ -1, & \text{ha } 1/2\le x<1,\\ 0 & \text{különben}\end{cases}$$

formában, majd ebből az alapfüggvényéből és annak egyszerű transzformáltjaiból (eltolásaiból, és összenyomás-széthúzásaiból) ortogonális sort képezett (Haar [1910]). Ezt bázisnak tekintve, e sor segítségével közelített függvényeket. Maga az ötlet nem volt új, hiszen függvénysorok segítségével közelített más függvényeket egyebek közt Taylor, vagy Fourier (Hunyadi [2017]). A Haar-függvényrendszer általános tagja

$$\xi_n(x)=\begin{cases}2^{j/2}, & \text{ha } k2^{-j}\le x<(k+1/2)2^{-j},\\ -2^{j/2}, & \text{ha } (k+1/2)2^{-j}\le x<(k+1)2^{-j},\\ 0 & \text{különben,}\end{cases}$$

ahol $n=2^j+k$, $j\in N$ és $0\le k<2^j-1$.

Ha ezt a rendszert kiegészítjük egy meglehetősen egyszerű ún. skálázó (kalibráló) függvénnyel:

$$\xi_0(x) = \begin{cases} 1, & \text{ha } 0 \leq x < 1, \\ 0 & \text{különben}, \end{cases} \tag{1}$$

akkor az első 4 Haar-függvény a következő alakot ölti:

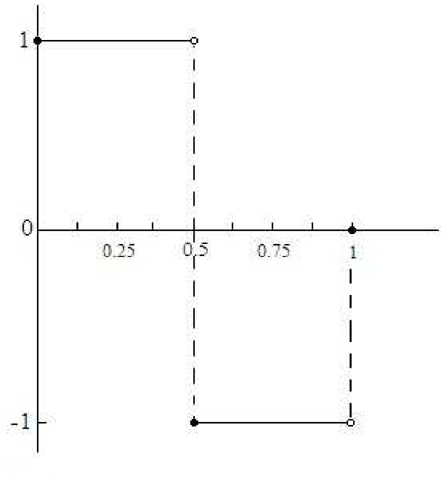
$$\xi_0(x) = \begin{cases} 1, & \text{ha } 0 \leq x < 1, \\ 0 & \text{különben}, \end{cases}$$

$$\xi_1(x) = \begin{cases} 1, & \text{ha } 0 \leq x < 1/2, \\ -1, & \text{ha } 1/2 \leq x < 1, \\ 0 & \text{különben}, \end{cases}$$

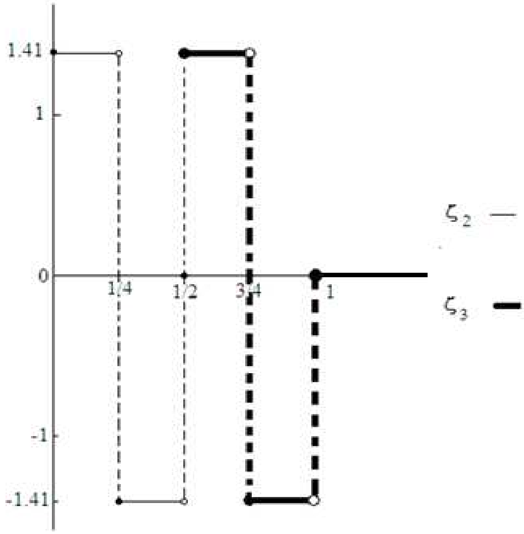
$$\xi_2(x) = \begin{cases} \sqrt{2}, & \text{ha } 0 \leq x < 1/4, \\ -\sqrt{2}, & \text{ha } 1/4 \leq x < 1/2, \\ 0 & \text{különben}, \end{cases}$$

$$\xi_3(x) = \begin{cases} \sqrt{2}, & \text{ha } 1/2 \leq x < 3/4, \\ -\sqrt{2}, & \text{ha } 3/4 \leq x < 1, \\ 0 & \text{különben}. \end{cases}$$

Mindez az 1. és 2. ábrán a következőképpen néz ki:

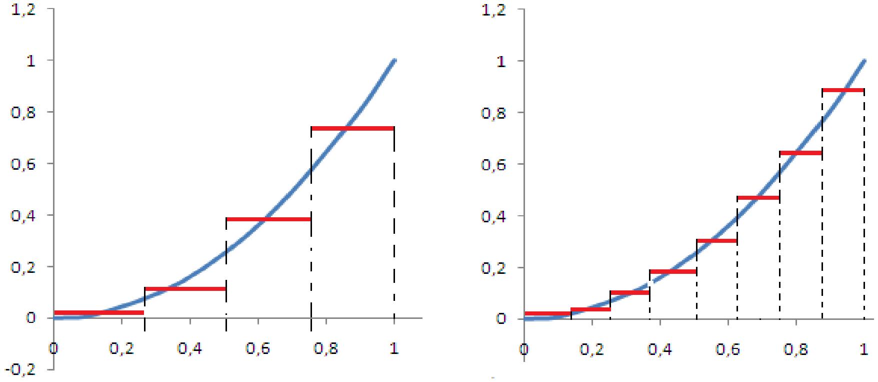


1. ábra. Haar alapfügggvénye (ξ_1). *Forrás:* saját szerkesztés.



2. ábra. A Haar-féle függvényrendszer 3. és 4. eleme.
Forrás: saját szerkesztés.

Már most előrebocsátjuk, hogy majd később, a waveletek bevezetése után, az azoknál alkalmazott terminológiával a ξ_0 skálázó függvényt apa-waveletnek, a ξ_1-et anya-waveletnek, a többi, származtatott függvényt gyermek-waveletnek nevezzük.



a) 4 taggal b) 8 taggal

3. ábra. Az $y = x^2$ függvény közelítése Haar-függvényekkel. *Forrás:* saját szerkesztés

A függvényközelítés pontos leírása megtalálható pl. Vidakovic [1999] művében, egy nagyon egyszerű példaként az $y = x^2$ függvényközelítését Haar-függvényekkel Hunyadi [2018a] mutatja be[3]. Mivel mindez csak előjáték a waveletekhez, itt csupán a 3. ábrán mutatjuk be azt, hogy 4 illetve 8 tagú

[3]A hivatkozott elektronikus füzetet a Szerző kérésre megküldi.

Haar-sorral milyen közelítést lehet elérni. Az ábrák magukért beszélnek, hiszen látható a jó közelítés. Az ábrák értékeléséhez csupán annyit kell hozzátennünk, hogy a közelítő függvények természetesen maguk is lépcsős függvények, és ha több tagú sort vonunk be a közelítésbe, a közelítés egyre pontosabb lesz.

Haar alapötlete akkoriban csak matematikai körökben talált visszhangra, alkalmazásokra a kor még nem volt kész. Így eredménye évtizedekig aludt, hogy aztán a manapság igen komoly alkalmazások alapja legyen.

A waveletek másik indító eleme az idősorok hullámelemzésének egyik eszköze, a Fourier-analízis, amelyik spektrálelemzés néven vonult be a statisztikába, és amelyik az idősorokban meglévő rejtett, de ismétlődő trigonometrikus hullámok segítségével írja le (közelíti) az idősorokat. Mind az egész alapjául szolgáló Fourier-analízis, mind pedig az azt felhasználó spektrálanalízis módszere részletesen dokumentált a szakirodalomban. Ezek közül a közgazdasági-társadalmi idősorok elemzésével foglalkozó szakirodalom néhány elemét: Granger – Hatanaka [1964] alapmunkáját, Harvey [1981] könyvének megfelelő fejezetét, Freschl [1978], valamint Pintér [2005] tanulmányait és Hunyadi [2017] elektronikus füzetét ajánljuk az olvasók figyelmébe. A spektrálanalízis lényege az, hogy trigonometrikus függvényekből és azok transzformáltjaikból alkotnak függvénysorokat, amelyeket bázisnak tekintve segítségükkel közelítenek tetszés szerinti idősorokat.

Ez a koncepció elsősorban a fizikában és a műszaki gyakorlatban – elsősorban persze a hullámokkal foglalkozó területeken – már jóval korábban lényeges alkalmazásokat nyert, ám amikor korlátai felszínre jöttek, a spektrálanalízis katalizátora lett a wavelet-elemzés kialakulásának.

2 Nemstacionárius hullámok kezelése

A Fourier-analízisen alapuló módszerek az időtérből a frekvenciatérbe transzformálják az időben lezajló jeleket, illetőleg az idősorokat. Segítségükkel megtudhatjuk, hogy egy jel kialakításában (a gazdasági modellezés fogalmai szerint a DGP-ben[4]) milyen frekvenciájú hullámok és milyen súllyal vesznek részt. Így fényt tudunk deríteni a folyamat bizonyos hullámtulajdonságaira. Ez az elemzés mellett lehetőséget ad – a hullámjellemzők állandóságát feltételezve – az előrejelzések pontosítására és finomítására, de emellett a különféle alkalmazott tulajdonságú szűrők tulajdonságainak vizsgálatára, a transzformációk hullámérzékenységére, és általában a jelek és idősorok elemzésére a hullámok terében. Ez a módszercsalád a fizikában, az akusztikában talált jelentős alkalmazásokra egyebek közt és jellemzően a jelek és zajok szétválasztásában, hangfelvételek minőségének javításában, hangtömörítésben stb. Közgazdasági és társadalomtudományi alkalmazásai azonban alig jutottak túl a kísérleti stádiumon.

A spektrálanalízis alaposan kidolgozott módszertana jó eszköz a hullámok elemzésére, de van egy komoly hiányossága, ami erősen korlátozza alkalma-

[4]DGP: data generating process – adatgeneráló folyamat (Hunyadi [1994]).

zását, ez pedig az, hogy *csak* a frekvenciatérbe transzformál, azaz ki tudja mutatni, hogy milyen frekvenciájú hullámok vannak az adott jelben, de azt nem, hogy ezek a hullámok hol, a megfigyelt tartomány mely részén jelennek meg, hiszen a transzformáció során elveszítjük az idődimenziót. Ez pedig azt is jelenti, hogy ez a módszer nem tudja kezelni a változó hosszúságú hullámokat, azaz nem alkalmas nem-stacionárius folyamatok elemzésére.[5]

Kezdetben a nem-stacionaritás problémáját úgy kezelték, hogy az eredeti intervallumot kis részekre osztották, és az egyes részekre külön-külön végeztek spektrálelemzést. Ez az ún. rövid távú Fourier-elemzés (Short Time Fourier Analysis) azonban elég nehézkes, kényelmetlen és meglehetősen munkaigényes eljárás volt (lásd pl. Polikar [1994]). A kutatók e helyett Haar eredményeihez nyúltak vissza, és kerestek olyan eszközt, amely segítségével egy lépésben, kényelmesen elemezhetők a nem-stacionárius idősorok (Kaiser [1994]). Így jutottak el a wavelet koncepcióhoz, amelynek lényege az, hogy definiál valamiféle hullámalakú alapfüggvényt (anya-wavelet), bevezet két transzformációt (eltolás és tágítás/összenyomás), majd az alapfüggvényből és a transzformáltakból függvénysort képez, amely segítségével elemezni lehet az időbeli jeleket. Mindez formálisan következőt jelenti.

A wavelet (hullámocska) eredendően olyan folytonos (vagy szakaszonként folytonos)

$$\Psi_{s,t}(t) = \frac{1}{\sqrt{s}}\Psi\Big(\frac{t-\tau}{s}\Big) \tag{2}$$

alakú függvény, amelyik

- a t időtérből a τ idő- és s frekvenciatérbe (2 dimenziós térbe) transzformál,
- hullám alakú (ezért integrálja 0), azaz $\int_{-\infty}^{\infty} \Psi(t)\,dt = 0$,
- és teljesít néhány, nem túlságosan korlátozó regularitási feltételt, pl. $\int_{-\infty}^{\infty} |\Psi(t)|^2\,dt < \infty$.

A wavelet függvények különböző paraméterek mellett ortogonális sort alkotnak, belőlük bázis készíthető. Maga a folytonos wavelet transzformáció (valós esetben):

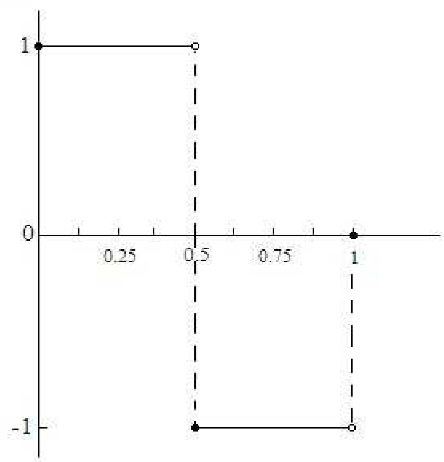
$$y = y(\tau, s) = CWT_x^{\Psi}(\tau, s) = CWT = \frac{1}{\sqrt{s}} \int x(t)\Psi\Big(\frac{t-\tau}{s}\Big)\,dt\,. \tag{3}$$

A (3) transzformáció, melynek elnevezése CWT, azaz continuous wavelet transformation (folytonos wavelet transzformáció), inverze is létezik és egyértelmű, azaz a folytonos wavelet transzformáció megfordítható (Vidakovic [1999]).

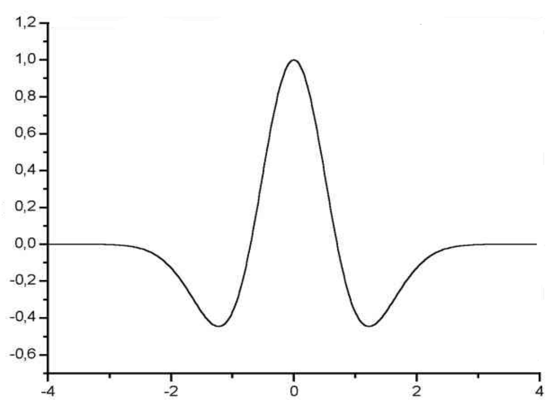
Az első és legegyszerűbb wavelet a Haar-függvény volt, de ahogy terjedt az alkalmazása, egyre bonyolultabb és kifinomultabb waveleteket készítettek.

[5] Nem-stacionárius jelen, vagy idősoron most pontos definíció híján azt értjük, hogy a hullámzás tulajdonságai nem állandók a vizsgált időtartományban.

Elsősorban fizikusok, geofizikusok alkalmazták őket. Eleinte különböző hullámfüggvényeket definiáltak, és azokból alakítottak ki a wavelet tulajdonságoknak megfelelő függvényeket, de később Daubechies [1992] már bizonyos tulajdonságokat definiált, és ezekből vezette le implicit wavelet-családját. Mára rengeteg különböző formájú wavelet létezik, és számuk alighanem még nőni fog. A következőkben, a 4-6. ábrákon csak néhány jellemző és gyakori wavelet (anya-wavelet) alakját mutatjuk meg[6]

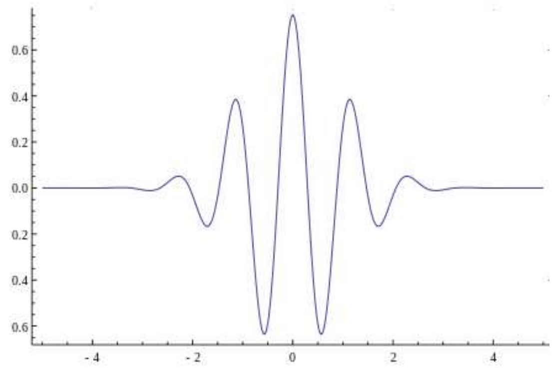


4. ábra. Haar wavelet. *Forrás:* saját szerkesztés



5. ábra. Ricker wavelet (Mexikói kalap). *Forrás:* http://en.wikipedia.org

[6]Csupán érdekességképp említjük arra, hogy maga a wavelet konkrét alakja mennyire nem lényeges, hogy a szakirodalomban találkozni lehet olyan tanulmányokkal, amelyek részletesen tárgyalják a wavelet elemzés lényegét, értelmét, módszereit, alkalmazását anélkül, hogy megmutatnának konkrétan akár egyetlen részletesen kifejtett waveletet is (!) Erre példa Polikar [1994] korábban már idézett tanulmánya.



6. ábra. Morlet (Gábor) wavelet.[7] *Forrás:* http://en.wikipedia.org

A folytonos wavelet transzformáció már alkalmas eszköz a nem-stacionárius hullámok elemzésére: megmutatja, hogy milyen frekvenciájú hullám, és hol (a tartomány melyik részén) helyezkedik el. Így alkalmazási lehetőségei sokkal szélesebb körűek, mint a korábban említett spektrálanalízisé, bár meg kell jegyeznünk, hogy a két módszer együttes, illetőleg párhuzamos használata jól kiemeli mindkét eszköz előnyös oldalát. A waveletek és a folytonos wavelet transzformáció korai alkalmazása elsősorban fizikai, műszaki területen volt jellemző: földrengések előrejelzése a geofizikában, napfoltok elemzése a csillagászatban, hangfelvételek javítása (zajszűrés) az akusztikában.

3 A waveletek diszkrét kezelése

A bemutatott CWT folytonos (vagy szakaszonként folytonos) függvényekkel dolgozik, ám a gyakorlatban a folytonos függvényeket digitális számítógépen csak közelíteni lehet, pontonként. A megoldást eleinte az jelentette, hogy a folytonos transzformáltat utólag pontonként értékelve diszkrétté alakították. Elég sűrű pontok („minták”) esetén ez természetesen megvalósítható volt, de itt is, mint más területeken, már korán felmerült az, hogy hatékonyabb megoldás lenne eleve diszkrét problémaként kezelni a feladatot. Ha diszkrét adatokból indulunk ki (és számos esetben ez a helyzet pl. digitalizált hang, kép, vagy gazdasági, társadalmi idősorok esetén), az eredményt diszkrét formában várjuk, akkor miért kell a folytonos esethez oda-vissza fordulni? Egyszerűbb eleve diszkrét problémaként kezelni: a bemenő adat diszkrét, a transzformáció csak diszkrét pontokat (idő és frekvencia) követ, így a kimenő adat is diszkrét (vektor), így az egész probléma viszonylag egyszerű lineáris algebrai feladattá válik. Maga a transzformáció

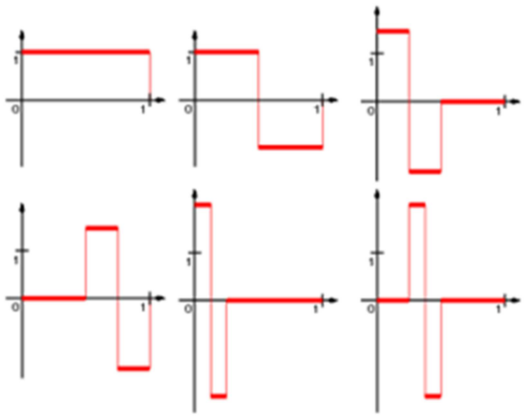
$$\mathbf{y} = \mathbf{W}\mathbf{x}$$

[7] A 6. ábrán látható waveletet a magyar származású fizikus, Gábor Dénes alkalmazta először, nevét mégis a tulajdonságait részletesen bemutató G. Morlettől kapta.

alakú lesz, ahol **x** a bemenő idősor, **W** a transzformáció négyzetes mátrixa, **y** pedig a diszkrét wavelet transzformált (DWT). A **W** négyzetes mátrix tartalma a wavelet fajtájától függ. Például 8 elemű idősorra a diszkrét Haar-transzformáció (DHT) (nem normált) mátrixa ($\mathbf{H} = \mathbf{W}$):

$$\mathbf{H}_8 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{pmatrix}$$

A mátrix értelmezéséhez elegendő emlékeztetni a Haar-függvényekre. A mátrix első sora az (1)-ben említett kalibráló (skálázó) függvénynek felel meg (apa-wavelet), a második sora az anya-wavelet, a harmadik és negyedik sora az első két transzformált gyermek-wavelet, és így tovább. A mátrix mérete természetesen igazodik a vizsgálandó idősor méretéhez, és elemeit úgy is lehet tekinteni, mint a kiinduló intervallumot egyenletesen felosztva n részre, minden intervallum közepén egy mintaelemet választunk, így alakítva diszkrétté a folytonos feladatot[8]. Ez az eljárás a 7. ábrán, ahol vázlatosan ismét bemutatjuk a Haar-függvényeket, jól követhető.



7. ábra. Az első 6 Haar-függvény. *Forrás:* http://en.wikipedia.org

Összefoglalva a **H** mátrixok tulajdonságait a következőket állíthatjuk:

- Haar-hullámokat képeznek le;

[8]Az itt bemutatott eljárás egyik komoly nehézsége, és egyben alkalmazásának korlátja az, hogy alapértelmezésben csak olyan méretű feladatokat tud kezelni, ahol a kiinduló adatok (és így az output) száma 2 egész kitevőjű hatványa. A problémára természetesen van megoldás (pl. Wickerhauser – Mladen [1994]), de mi itt csak az alapesettel foglalkozunk.

- ortogonálisak, sőt egyszerű, konstanssal történő normalizálás után ortonormáltak, azaz inverzük megegyezik a transzponáltjukkal;
- szorzásuk, invertálásuk csak egyszerű műveleteket tartalmaz, ami igen gyors oda-vissza transzformációt tesz lehetővé.

Ezek után a DHT példája alapján összefoglalhatjuk általában a DWT fontosabb tulajdonságait. Ezek pedig a következők:

- A DWT ugyanolyan méretű vektor, mint amilyen a kiinduló idősorvektor, azaz $\mathbf{x}$ és $\mathbf{y}$ elemszáma megegyezik, hiszen $\mathbf{W}$ (és $\mathbf{H}$) négyzetes mátrixok;
- a transzformált $\mathbf{y}$ vektor struktúrája olyan, hogy *elején* vannak a hosszú hullámok, közben a különböző hosszúságú középhullámok, a végén a rövid, átmeneti, zajnak tekinthető hullámok együtthatói;
- viszonylag könnyen belátható az információmegmaradás[9] egy sajátos esete: $\mathbf{x}$ és $\mathbf{y}$ elemeinek négyzetösszege megegyezik, de az eloszlásuk természetesen nem: $\mathbf{y}$ vektor sokkal koncentráltabb, mint $\mathbf{x}$;
- ezért a DWT jellemző alkalmazása az, hogy transzformáció után a *küszöböléssel* $\mathbf{y}$ kis elemeit elhagyjuk (0-vá tesszük), és az így küszöbölt DWT inverze jó közelítést ad a kiinduló vektorra, az eredetinél sokkal kevesebb elem felhasználásával. Ez az információtömörítés a modellezés másodlagos, de igen lényeges értelme.

Ezekre az eredményekre már sokan felkapták a fejüket, egyebek közt közgazdászok, társadalomtudósok, hiszen új lehetőségeket vetettek fel az idősorok hullámelemzésének terén. De még ennél is nagyobb volt az informatikusok érdeklődése, hiszen meglátták benne a lehetőséget nagy adattömegek gyors és egyszerű tömörítésére, kezelésére. Így ez vált napjainkra a waveletek alkalmazásának kiemelt területévé, ám az igazi áttöréshez még egy lépés kellett. A mátrixok nagy adattömeg esetén ugyanis igen nagy méretűek és elég üresek (sok bennük a 0 elem), így kezelésük felettébb nehézkessé tette volna az eljárást, aminek épp a könnyűsége lenne az előnye. Helyettük kellett valami más megoldás, ami nem is késett sokáig, hiszen a kutatók elég hamar eljutottak az eljárást tökéletesen leíró, ugyanakkor könnyen gépesíthető algoritmusokhoz.

4 DWT algoritmusok

Elsősorban a számítástechnika és az informatika gyors fejlődése vezetett oda, hogy a zárt formulák mellett, vagy inkább helyett, a könnyen programozható, egyszerű műveleteket nagy számban végző algoritmusok kerülnek előtérbe.

[9] Ezt az összefüggést a fizikai-műszaki alkalmazásokban energiamegmaradásként emlegetik.

Ez történt a DWT esetén is: bár lineáris algebrai megoldást akár zárt formában is lehet találni (ez bonyolultabb waveletek esetén azért nem nyilvánvaló), a gépi alkalmazásokhoz jobban illeszkedő algoritmusok váltak dominánssá. Ezek közül talán legismertebb az ún. Piramis algoritmus (Pyramide Algorithm), amit, illetőleg aminek változatait a szakirodalom sok más néven is ismeri (Cascade Algorithm, Subband Coding, Lifting Scheme). Ezek részletes ismertetése helyett itt csupán a piramis algoritmus működését mutatjuk be Haar-waveletek esetére egy igen egyszerű, könnyen nyomon követhető példán.

Legyen a bemenő idősor $\mathbf{x} = [1, 2, 3, 4]$ alakú, és tekintsük át lépésenként a diszkrét Haar-transzformáció piramis algoritmusát!

Az algoritmus alapötlete az, hogy az $\mathbf{x}$ elemeiből meghatározott rend szerint párosítva az elemeket, *összeg és különbség* jellegű mutatókat számol, majd a különbséget tárolva, az összegeket (melyek száma már az induló elemszám fele) újból összeg- és különbség típusú elemekké alakítja. Az egész eljárás addig folytatódik, ameddig csak egyetlen elem marad. Ha a kiinduló vektor elemszáma 2 egész kitevőjű hatványa volt ($n = 2^q$), akkor az eljárás q lépésben befejeződik. Ekkor az eredményvektor az egyes lépések (szintek) során képződött különbség jellegű elemekből, valamint az utolsó lépésben maradt összegből adódik. Az algoritmus pontos leírása, sőt lineáris algebrai reprezentációja is megtalálható pl. Vidakovic [1999] könyvében.

Példánkban ezt az algoritmust alkalmazva először particionáljuk az idősort $\mathbf{x} = [(1, 2), (3, 4)]$ alakban, majd a transzformáció első lépéseként készítsük el a normált *összegeket* és *különbségeket*

$$s_{11} = \frac{1+2}{\sqrt{2}} = \frac{3}{\sqrt{2}}, \; s_{12} = \frac{7}{\sqrt{2}}, \; d_{11} = \frac{1-2}{\sqrt{2}} = \frac{-1}{\sqrt{2}} \text{ és } d_{12} = \frac{-1}{\sqrt{2}},$$

majd foglaljuk össze ezeket egy munkavektorban:

$$\mathbf{x}_1 = \left[\frac{3}{\sqrt{2}}, \frac{7}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right].$$

Ezzel az algoritmus első lépése készen van (ezt nevezik *első szintnek*). A második lépésben ebből a vektorból kiindulva, a különbségeket (mint a végeredmény részét) félretesszük, a maradó kételemű összeg vektort újból felbontjuk, úgy, ahogy ezt az első lépésben tettük:

$$s_{21} = \frac{1}{\sqrt{2}}\left(\frac{3+7}{\sqrt{2}}\right) = \frac{10}{2} = 5 \quad \text{és} \quad d_{21} = \frac{1}{\sqrt{2}}\left(\frac{3-7}{\sqrt{2}}\right) = \frac{-4}{2} = -2\,.$$

Mivel 2 lépésben az eljárás végére értünk, most már összerakhatjuk a végeredményt:

$$\mathbf{y} = \left[5, -2, \frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right].$$

A feladat mérete 4 volt (4 elemű idősor-vektorból indultunk ki), azaz $2^2 = 4$, így az eljárás 2 lépésben befejeződik, az iterációnak *2 szintje* van.

Könnyű belátni, hogy a normalizált Haar-mátrixszal végzett transzformáció ugyanezt az eredményt adja (egy lépésben). Mivel

$$\mathbf{H}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & \sqrt{2} \end{bmatrix},$$

és normalizált változata $\mathbf{H}_4^* = \frac{1}{2}\mathbf{H}_4$, így

$$\mathbf{y} = \mathbf{H}_4^* \mathbf{x}(t) = \begin{bmatrix} 5 \\ -2 \\ -1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}.$$

A vissza-transzformáció (rekonstrukció) kézi számítással egy kicsit hosszadalmas, de működik, könnyen ellenőrizhető. A példából jól látható, hogy az eljárás nagy feladatok esetén is kényelmesen, gyorsan működik.

Más waveletek esetén természetesen más, de az itt bemutatotthoz hasonló, egyszerű, gyors és megfordítható algoritmusok készültek, amelyek számítógépes megvalósítása is elég egyszerű. Ezek az algoritmusok olyan gyorsak, hogy már alkalmasak nagy fájlok igen rövid idő alatt történő tömörítésére és kibontására.

Mivel a feladat (a wavelet transzformáció) problémája egyszerűen kiterjeszthető 2, sőt 3 dimenzióra, egy sor manapság igen elterjedt számítógépes alkalmazás alapjául szolgál. Jellemző alkalmazás egy dimenzióban a hangtömörítés, két dimenzióban a képtömörítés, 3 dimenzióban a mozgóképek tömörítése. Ezen alkalmazások során a mátrixokká alakított hang- és képfájlokat DWT-vel transzformálják, majd a kis, zavaró elemeket küszöbölik, így a feladatot méretben erősen redukálják (akár az eredeti méret tizedére), majd szükség esetén inverz DWT-vel visszaalakítják az eredeti dimenzióba. Tapasztalatok szerint az ilyen eljárás csak néhány %-kal rontja az eredeti hangilletve kép minőségét, ami általában az érzékelés számára észrevehetetlen marad. Az egyik első híres, ilyen irányú alkalmazást az FBI végezte el több millió darabot számláló ujjlenyomat bankjával, ahol a tömörítés után a felismerhetőség teljes értékű maradt (Vidakovic – Mueller [1991]). A számítógépes alkalmazások közül legismertebb a jpeg formátumú képtömörítés, amely szintén DWT alapokon nyugszik. A televíziózásban a HD képminőségű adás, illetve vétel gyors elterjedésének egyik fő akadálya az volt, hogy nagy információmennyiséget kellett gyorsan mozgatni. Amikor a HD szabványt készítették, egy Fourier-analízist felhasználó algoritmus teremtette meg erre a lehetőséget, de akkor a DWT még gyerekcipőben járt (Strang [1994]). Nincs információnk róla, de könnyen lehetséges, hogy az UHD kifejlesztése, a még nagyobb mennyiségű információ gyors átvitele DWT algoritmuson alapul.

A waveletek elmélete és alkalmazása igen gyorsan fejlődő terület a matematikában, a fizikában, az informatikában és más alkalmazott területen is. Elterjedéséhez nagyban hozzájárulhat az is, hogy hatékony algoritmusok mellett jól hozzáférhető számítógépes csomagok állnak rendelkezésre, elsősorban S és R nyelveken.

5 Nekünk mire jók a waveletek?

Eddig sokat írtunk a waveletek fizikai, műszaki alkalmazásairól, de természetesen felvetődik a kérdés, hogy mire alkalmazhatók a waveletek a társadalom- és gazdaságtudományok területén? Természetesen erre a kérdésre azonnal nem lehet egyszerű választ adni, hiszen maga az elmélet eléggé kevéssé ismert ezen tudományágak művelői, és elsősorban modellezői körében, így nem lehet tudni, hogy egy kicsit jobban megismertetett módszertani alapokkal milyen alkalmazásokra lehetne számítani. Ezért csupán azokra a triviálisnak tűnő alkalmazásokra mutatunk itt be példákat, amelyek részben a meglévő műszaki jellegű feladatokból analógiák folytán származtathatók, részben pedig a hagyományos idősorelemzési eszközök továbbfejlesztésével alakíthatók ki. Tehát a statisztika idősorelemzési alkalmazásaiból indulunk ki, de meg kell említenünk, hogy a területi sorok elemzése hasonlóképp érdekes és hasznos célterület lehet, bár a területi statisztika eszköztára ez ideig kevésbé kidolgozott, mint az idősoroké.

A következőkben tehát néhány jellemző alkalmazást mutatunk be példákkal illusztrálva[10].

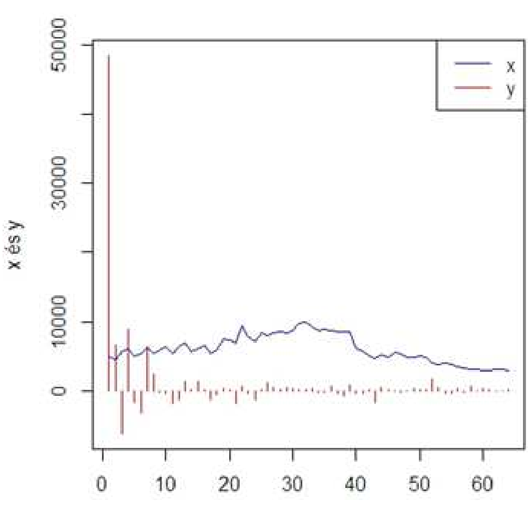
a) Hullámok kimutatása, elemzése

A waveletek elsődleges alkalmazása természetesen a hullámok elemzése, hiszen segítségükkel ki lehet mutatni, hogy a vizsgált idősorban hol és milyen hosszúságú hullám detektálható. Meg kell jegyeznünk, hogy a módszer korlátja lehet a rövid idősor, hiszen pl. a Haar-waveletekkel csak erősen korlátozott számú és nagyságú hullámot tudunk elemezni. Az itt következő példákban 64 vagy 32 elemű éves idősorokat vizsgálunk, ami a makromodellezésben hosszúnak minősül, ugyanakkor a wavelet elemzés legfeljebb 6 különböző, 1,2,4,8,16 és 32 éves ciklusokat tud azonosítani, és ilyen ciklusokat tud elhelyezni a teljes intervallum (64 év) megfelelő szakaszán. Ez természetesen elég durva eredmény, de ismételten hangsúlyozzuk, hogy az elemzés igazi területe a valóban hosszú idősorok (big data) vizsgálata.

Példaként a magyarországi sertésállomány 1960-2016 közötti éves idősorát és annak DHT-jét mutatjuk be (8. ábra). Ennél a feladatnál a diszkrét Haar-transzformációt használjuk fel annak érdekében, hogy megismerjük a jelenség hullámtermészetét, különös tekintettel arra, hogy korábban az ökonometria egyik ,,slágerfeladata" volt a sertésciklus kimutatása és elemzése (például Kornai [1981]). A waveletek segítségével most eggyel több, adekvát eszközünk adódik a sertésállomány ciklikus alakulásának vizsgálatára.

[10] A számításokat a KSH éves idősorai alapján végeztük saját fejlesztésű R-kódok segítségével. A számítások részletei megtekinthetők Hunyadi [2018a és 2018b] tanulmányaiban.

A 8. ábrán látható, hogy az első (pálcikával jelölt) együttható pozitív és nagy: ez a kalibráló változó együtthatója, és csak nagyságrendet jelöl: az átlag szorzódik egy konstanssal. A 2. együttható, amelyik ez esetben pozitív, de nem túl nagy, az alaptendenciát jelzi, hiszen ez a Haar-alapfüggvény (anyawavelet) együtthatója. Tárgyi jelentése az, hogy az idősor, akár az alapfüggvény az egész vizsgált intervallumon alapvetően csökkenő tendenciát mutat, de ez a tendencia (trend?) nem túl erős. Az ezt követő két együttható (a 3. negatív, a 4. pozitív) a hosszú hullám meglétére utal. A negatív együttható egy "fel" alakú hullámot jelöl az időszak első felében, míg a 4. pozitív együttható egy hasonló hosszúságú "le" alakot az időszak második felében. Ezek eredője adja ki az ábrán látható hosszú hullám konkáv szakaszát, ami lényegesen meghatározza az idősor lefutását. A DHT ábrán látszik még, hogy a 5-8 együtthatók viszonylag nagyok. Ezek a 16 éves középhullámokat mutatják. Ezekből az látszik, hogy ezek közül a 3. együttható (valójában a 7.) a legnagyobb, azaz az időszak 3. negyedében érvényesül leginkább egy középhullám csökkenő ága. Az ezt követő együtthatók, melyek a rövid hullámokat mutatják, kisebbek, nem jeleznek számottevő hullámzást. Ugyanakkor fontos megjegyezni, hogy ha célunk az lenne, hogy a sertésciklust, nem pedig a DHT tulajdonságait elemezzük a lehető legalaposabban, érdemes lenne megnézni, hogy pl. trendszűrés után (amikor a tartós tendenciát, esetleg a leghosszabb hullámot, vagy hullámokat kiszűrjük) mit mutat a transzformáció, hiszen egy ilyen eljárás felerősítheti a maradék idősorban lévő, nagyobb frekvenciás hullámokat, és jobb betekintést adhat az idősorban megjelenő rövid hullámok létébe és elhelyezkedésébe.

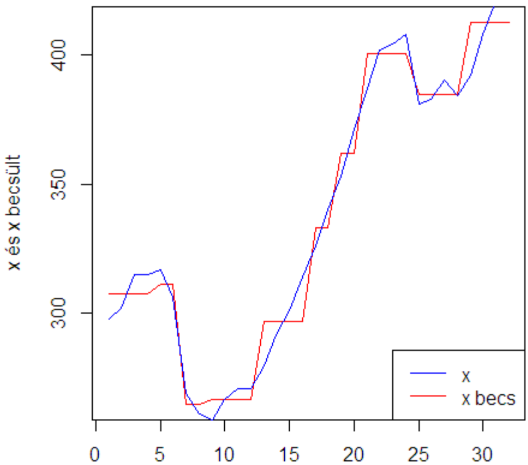


8. ábra. A sertésállomány idősora és annak DHT-je.
Forrás: saját szerkesztés.

b) Simítás, közelítés, outlier szűrés, tömörítés

A DWT a korábban elmondottak alpján egyre inkább elveszíti eredeti funkcióját, és a hullámmozgások elemzése helyett idősorok tömörített közelítésére, a kiugró értékektől való megszabadítására, általában egyszerűbb, modellszerű leírására használják. Ennek módja az, hogy a wavelet transzformáltból küszöböljük a kis elemeket, majd az így rövidített, tömörített transzformáltakból a vissza-transzformációt alkalmazva, megkapjuk az eredeti idősor közelítését. Ezt az eljárást úgy is fel lehet fogni, mint outlier szűrést, amikor a kis, eseti, tranziens hullámoktól megtisztítjuk az idősort.

A 9. ábrán a GDP indexének idősorát (1965-2018) és annak DHT-vel történő közelítését láthatjuk. Ebben az esetben úgy jártunk el, hogy a DHT **y** vektorának elemeiből lenulláztuk a legkisebb abszolút értékű 70%-ot, majd elvégeztük az inverz transzformációt. A 70% természetesesen csak egy lehetőség, a lényeg persze az, hogy az idősor egy nem elhanyagolható részét töröljük, és így a feladatot egyszerűsítsük. Több kísérlet után jutottunk el oda, hogy ez esetben a 70%-os szűrés látszik olyannak, ami lényegesen egyszerűsíti (modellezi) a feladatot, ugyanakkor az idősor fő jellegzetességeit is megtartja. Az ábráról ugyanis jól látható, hogy a közelítés a nehezen követhető ingadozások ellenére jónak mondható, hiszen a két sor között a determinációs együttható $r^2 = 0.97$, azaz bizonyos értelemben azt is állíthatjuk, hogy az eljárás csak 3% az információveszteséggel járt. Lényegileg ugyanezt az eljárást követik természetesen bonyolultabb módon a már említett hang- és képtömörítő eljárások. Ennél a kis példánál az eljárás előnyei nem tűnnek ki, de nagy adatmennyiség (pl. big data környezetben esetleg több százezer adat) esetén az adattárolás és manipuláció, valamint a feladat átláthatósága szempontjából ez már lényeges lehet.



9. ábra. A GDP idősora és annak DHT-vel történő közelítése.
Forrás: saját szerkesztés.

c) Hasonlóságok keresése idő- és területi sorokban

Ez az alkalmazás, bár érdekesnek tűnik, egyelőre nem valamiféle bizonyított elméleten, hanem egy sejtésen nyugszik. Lényege az, hogy ha idősorokat szándékozunk összehasonlítani, nem az eredeti idősorokat, hanem azok wavelet transzformáltjait hasonlítjuk össze. Emögött az a feltételezés áll, hogy a wavelet transzformáltak tömör és rendezett formában hordozzák azoknak az idősoroknak a tulajdonságait, amelyekből származnak. Ezért, ha a wavelet transzformáltakat hasonlítjuk össze, hatékonyabb lesz az összehasonlítás, az jobban kiemeli a közös tulajdonságokat, illetve jobban szelektál. Ezt a gondolatot még nem láttuk egzaktan bizonyítva, de alkalmazását már nyomon lehet követni. Uliha [2016] ezt az elvet használta a svéd és norvég makrogazdasági változók olajárral való együttmozgásának jellemzésére. Ezzel kapcsolatban az elmélet helyességén túlmenően természetesen egy sor kérdés vethető fel, így pl. az, hogy milyen waveletet célszerű használni, a teljes vagy a küszöbölt DWT vektort használjunk, vagy az, hogy a hasonlóságot milyen mutatóval mérjük?

Az elv bemutatására elkészítettük 4 makró változó (GDP, fogyasztás, ipari termelés, beruházás) 32 elemű (1985-2016) volumen-idősorainak összehasonlítását egyszerű korrelációs mutatók segítségével, ahol a korrelációs mátrixokat először az eredeti idősorokból, másodszor a DHT transzformáltakból számítottuk ki. Az eredményeket az *1. és 2. táblázat* korrelációs mátrixaiban foglaltuk össze.

	GDP	Fogyasztás	Ipari termelés	Beruházás
GDP	1	0.9435	0.9728	0.9286
Fogyasztás		1	0.8882	0.9051
Ipari termelés			1	0.9306
Beruházás				1

1. táblázat. Az eredeti sorokból számított korrelációs mátrix

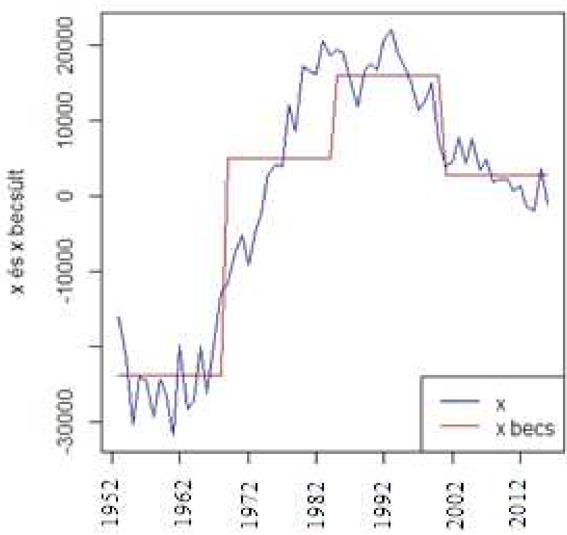
	GDP	Fogyasztás	Ipari termelés	Beruházás
GDP	1	0.9985	0.9804	0.9941
Fogyasztás		1	0.9710	0.9910
Ipari termelés			1	0.9885
Beruházás				1

2. táblázat. A DWT-ből számított korrelációs mátrix

A két mátrix elemenkénti összevetése az mutatja, hogy a transzformáltakból számított korrelációk *minden esetben* magasabbak, azaz a transzformáltak mindenütt erősebb kapcsolatot mutatnak, mint az eredeti sorok. Ha még azt is hozzávesszük, hogy az idősorok közti korrelációs együtthatók tartalmazzák a közös trendhatást, ami nyilván felfelé torzítja az eredeti sorokból számított mutatót, míg ez a transzformáltaknál nem, illetve másképp jelenik meg, azt mondhatjuk, hogy ezen a példán a wavelet transzformáltak felerősítik az együttmozgást leíró kapcsolatokat. Ez természetesen nem bizonyíték, csak indikáció, de maga a gondolat alighanem érdemes arra, hogy a későbbiek során alaposabban megvizsgáljuk.

d) Outlier szűrés és a legfontosabb fordulópontok elemzése

A 10. ábrán egy demográfiai idősort, a magyarországi halálozások számának idősorát és annak durva DHT közelítését láthatjuk (Hunyadi [2018b]). A függőleges skálán látható, hogy az alapadatokat a saját átlaguktól vett eltérésekkel helyettesítettük, ami az itt felvetendő kérdést – a kritikus pontok elhelyezkedését – nem érinti. Itt tehát a halálozások idősorának kritikus fordulópontjait kerestük $p = 0.96$ szűrés esetén, azaz úgy, hogy a DHT elemeinek legkisebb 96%-át 0-ra redukáltuk. A kapott fordulópontok: 1969, amikor megugrik a sor, 1985-ben kezdődik a lassulás, 1985 és 2001 közt megfordul a tendencia, végül 2001-től alacsony szinten stagnál, illetve lassan csökken.



10. ábra. A halálozások száma és lépcsős közelítése.
Forrás: saját szerkesztés.

Ez az alkalmazás hasonlít arra, amit korábban már említettünk. A különbség mindössze annyi, hogy a DHT-t ezúttal erősen küszöböljük (szűrjük), azaz pusztán 3-4 elemét hagyjuk meg, amelyek visszatranszformálás után kirajzolják azt a függvényt, amely durván ugyan, de megragadja az eredeti idősor lényegét, ugyanakkor megadja azt a néhány kulcsévet, amelyekben a folyamat lényeges fordulópontjai bekövetkeznek. Ehhez az elemzéshez kifejezetten előnyös a Haar-transzformáció lépcsős jellege.

A fenti néhány alkalmazás csak bemutató volt, azonban nem szabad elfelejteni, hogy a waveletek igazi elterjedése alighanem még előttünk áll. Korábban tettünk említést más, műszaki, fizikai, informatikai alkalmazásokról. A fent bemutatott néhány, társadalom- és gazdaságtudományi alkalmazáshoz még hozzá lehet tenni, hogy általában nagy lehetőség adódik a nem stacionárius idősorok elemzése esetén, hiszen ezen a téren a wavelet elemzés kiválthatja, sőt ki is kell váltania a hagyományos Fourier-elemzéseket. Már említettük, de talán nem árt újra hangsúlyozni, hogy a nagyon nagy adatmennyiségek (big data) esetén, ami jelenleg már a statisztika kézzel fogható valósága, a wavelet elemzés szintén igen sikeres lehet. Végül megemlítjük a területi statisztikai elemzéseket, amelyek az idősoros módszerek mintájára azokhoz hasonló, illetve közel álló módszereket használnak (Dusek T. – Kotosz B. [2016]), és a jövő előttük áll. Azt sem szabad elfelejteni, hogy a wavelet-elmélet még viszonylag új, társadalmi-gazdasági alkalmazásai még csak kialakulóban vannak, de biztosak lehetünk abban, hogy amint ez a nagyon hasznos terület megerősödik és elterjed, újabb és újabb alkalmazásokat indukál, olyanokat is, amilyenekre most talán még nem is gondolunk.

6 Záró gondolatok

Ez a cikk csupán bevezető a waveletek elméletének és alkalmazásának világába. Bevezető, mégpedig elég rövid, és talán felszínes bevezető, de tudatában kell lennünk annak, hogy a waveletek mélyebb megértése és alkalmazása transzdiszciplináris ismereteket követel. Maga a téma meglehetősen divatos és jól mutatja a modern matematika jellemzőit, a diszkrét probléma-kezelést, a számítástechnika felé fordulást, az alkalmazás-orientáltságot. Az átalakulóban, paradigmaváltás előtt álló statisztika jellemzői is megmutatkoznak ezen a területen, hiszen az induló adatok forrásainak minősége és mennyisége, a nagy számításigény, a sokoldalú felhasználhatóság, mind beleillik a statisztika átalakulásának egyre jobban kirajzolódó trendjébe.

A waveletek társadalmi-gazdasági modellezésben történő alkalmazásának is csak néhány egyszerű kezdetét villantja fel ez az összeállítás. Alighanem joggal gondolhatjuk, hogy az alkalmazásokban ennél jóval több lehetőség van. Ahhoz azonban, hogy ezeket a lehetőségeket jobban kihasználjuk, éppen a terület korábban említett transzdiszciplináris jellege miatt az érintett szakterületek, azaz a matematika, az informatika, a statisztika, valamint a társadalmi-gazdasági modellezés együttműködésére lenne szükség. Ennek a cikknek elsődleges célja és értelme az, hogy egy ilyen együttműködés szükségességére felhívja a figyelmet.

Irodalom

1. Daubechies, Ingrid [1992]: *Ten Lectures on Wavelets*, 2nd ed. Philadelphia: SIAM, CBMS-NSF Regional Conference Series in Applied Mathematics 61.
2. Dusek Tamás – Kotosz Balázs [2016]: *Területi statisztika.* Akadémiai Kiadó, Budapest
3. Freschl György [1978]: *Bevezetés a spektrálanalízisbe.* Ökonometriai Füzetek 15. KSH Ökonometriai Laboratórium, Budapest
4. Granger, Clive William John – Hatanaka, Michio [1964]: *Spectral Analysis of Economic Time Series.* Princeton University Press, Princeton
5. Haar Alfréd [1910]: Zur Theorie der orthogonalen Funktionensysteme, *Mathematische Annalen*, 69, pp. 331-371

6. Harvey, Andrew [1981]: *Time Series Models.* Philip Allan Publisher Limited, Oxford.
7. Hubbard, Barbara Burke [1998]: *The World according to Wavelets.* AK Peters, Natick, MA
8. Hunyadi László [1994]: Egységgyökök és tesztjeik. *Szigma* 25, 135–169.
9. Hunyadi László [2017]: Bevezetés a spektrálanalízisbe. HLM1, Budapest (kézirat)
10. Hunyadi László [2018a]: A wavelet elemzés alapjai, HLM3, Budapest (kézirat)
11. Hunyadi László [2018b]: Demográfiai változók hullámtulajdonságairól. In: *Emlékkötet Hoóz István 90 éves jubileuma alkalmából*, PTE, Pécs, 255–268.
12. Kaiser, Gerald [1994]: *A Friendly Guide to Wavelets.* Birkhäuser, Boston
13. Kornai Gábor [1981]: Megszűnt-e a sertésciklus? A magyarországi sertéstartás ökonometriai modellje. *Közgazdasági Szemle*, 28(3). 316–332.
14. Pintér József [2005]: A spektrálanalízisről. *Statisztikai Szemle,* 85(2), 130–156.
15. Polikar, Robi [1994]: *Wavelet Tutorial.* Iowa State University web.iitd.ac.in/~sumeet/WaveletTutorial.pdf
16. Schipp Ferenc [2015]: Hiperbolikus waveletek. Arató Mátyás professzor emlékére. *Alkalmazott Matematikai Lapok* 32, 1–39.
17. Strang, Gilbert [1994]: Wavelets. *American Scientist* 82, 250–255.
18. Uliha Gábor [2016]: Az olajár és a makrogazdaság kapcsolatának elemzése folytonos wavelet transzformáció segítségével. *Statisztikai Szemle.* 95(5), 506–534.
19. Vidakovic, Brani [1999]: *Statistical Modeling by Wavelets.* Wiley, New York
20. Vidakovic, Brani – Mueller, Peter [1991]: *Wavelets for Kids. A Tutorial Introduction*, Duke University Durham, NC 27708-0251 brani@isds.duke.edu pm@isds.duke.edu
21. Wickerhauser, Mladen Victor [1994]: *Adapted Wavelet Analysis from Theory to Software.* AK Peters Ltd, MA.
22. http://en.wikipedia.org

ON THE WAVELETS

The wavelet problem is not a new one, the first wavelet, that resembles to the ones used nowadays, were constructed by the Hungarian mathematician Alfred Haar in the beginning of the 20th century. Naturally, he tackled with a problem of pure mathematics, and he solved this one: using his famous stepwise function, among others, he could approximate any continuous functions. However, the results, obtained by Haar were rather unprecedented, so that time it did not obtain real applications outside the narrow field of theoretical mathematics. Many decades later, in the second half of the century, first in connection with physical, and later informatical problems, the real wavelet problem has been outlined and used. This proved to be a very effective tool to analyse non-stationary time series, a good means that complete the Fourier analysis, which can exceed it in some points. This article is an introductory study: it deals with some basic properties of wavelets, introduces the

discrete wavelets, shows an algorithm that is popular nowadays and useful, to accomplish wavelet transformation in both directions by using a computer. After this methodological section the article shows some eventual application in economic and social modelling and at the end it demonstrates some further hints of the hidden application facilities of the topic.

In the beginning, the article shows some results of Haar. Results, that later became the basic foundation of the whole topic. First it demonstrates the Haar-type basic stepwise function (the so called mother wavelet), the calibration function (or father wavelet), and the transformed (translated and dilated) versions of the mother wavelet (children wavelets). The series, formed from these elements – and this is a famous result of Haar – is orthogonal, and using this series, any continuous function can be approached with a voluntary precision. This result – in some aspects – resembles to the results of Taylor and Fourier, and the approximation is more or less similar, but evidently based on different functions. The article shows a very simple and well known quadratic function and its stepwise approximation by Haar series of 4 and 8 elements.

For a long time, a half century, this result was only that of mathematicians, it had not application outside this area, and only in the mid of the last century begin first of all physicists to reveal the connection of this type of methods and the Fourier transformation. Starting from Haar's results, constructing, introducing and using new wavelets they get efficient tools for time series analysis. So, new wavelets have been created (Gabor – Morlet wavelet, the Mexican hut wavelet, the implicitly defined Daubecies wavelet and a lot of other wavelets). This made necessary the unification of the theory based on wavelets. Following this, application of wavelets grown rapidly – this period is called by some authors the revolution of wavelets. The applications reached further territories like social and economic analysis, and not only time series analysis. In this field modern informatics was the first that used and applied wavelet analysis to manipulate (coding, storing and compressing of pictures) spatial data as well.

In connection with the applications, one of the main bottleneck of the original wavelet theory was, that it based on continuous (or at least partly continuous) functions, so there was a separate phase, when the results had to transform to discrete form or to digitalize them. A comparatively new result in the wavelet theory is the discrete transformation, which starts from the fact that the data are discrete (or at least discrete on the level of observation), so the wavelet mapping and compression of such data can be done by discrete methods only. The article deals with the discrete wavelet transformation in a detailed way. First the Haar matrix was introduced, but it was easy to understand that for large scale problems the quasi-empty transformation matrix (the large share of its elements is 0), is rather inconvenient for practical use. Therefore, starting from the original Haar transformation and using a rather simple but rather efficient stepwise method – the Cascade Algorithm – we show how to apply the wavelet transformation to discrete time or spatial series. The article shows this algorithm via a small numerical example. As the result of the transformation – which operates in this case as well in both directions – a series (a vector) of coefficients of the same size can be obtained. From this vector, by cutting or thresholding – i.e. deleting and replacing some coefficients by 0 – and an inverse transformation an approximation of the original series (vector) can be obtained.

After this transformation

- the DWT (Discrete Wavelet Transform) of a series remains a vector of the same size as the original one, that is the number of their elements is equal, since the transformation matrix is quadratic;
- the structure of the result vector ($\mathbf{y}$) is special: at the beginning (on the left hand side) the coefficients of the long waves, at the middle the coefficients of the middle waves, while on the right hand side of the short waves, which

refers to irregular sound, can be found;
- it is easy to see an interesting case of the information-conservation: the sum of squares of the elements of vector **x** and **y** are equal, but their distribution is very different: vector **y** is much more concentrated than the original vector (**x**);
- that is why a peculiar application of the DWT is as follows: after the transformation by mean of thresholding we delete the small absolute valued elements of **y** (i.e. they get 0 value), and after re-transformation the new vector **x** gives a good approximation of **y** but using much less nonzero-elements. This is a simple way of information compression, and so this is an important point of wavelet modelling.

As far as the applications are concern, the problem (the wavelet transformation) is easily extended to 2, or 3 dimensions, so it becomes the base of a lot of extended computer applications. A peculiar application in 1 dimension is the compression of voices in music files, in 2 dimensions the compression of images, in 3 dimensions the compression of TV pictures together with voices. In these applications first the files of sound or images are mapped to matrices, than transform by DWT. The small, nuisance elements are thresholding, so the original problem is reduced (sometimes by 90 percents), and then using the inverse DWT it will be re-transformed to the original scale. According to some reports, such double transformation will not reduce the quality of the original file in a considerable way, so in general it remains hidden for the users. The first famous application was due to the FBI with its famous fingerprint data bank: the data bank that contains millions of prints, which was compressed by using wavelet transformation. After a more then 90% compression rate, it remained fully operable. Among the computer application the jpeg image compression is well known: it is based a discrete wavelet transformation too. In the modern TV broadcasting a relevant obstacle of peak-quality HD (High Definition) transmission was the lack of means that could save and move a huge amount of information in a very short time. When the HD broadcasting was prepared, an algorithm, based on the Fourier-analysis was applied, but in that time there were no effective DWT algorithms. We have no information on the details, but it can easy be that the UHD television, where manipulation with much more information is necessary, is based on DWT algorithms.

In the last, but rather important part of the article some applications of wavelet transformation in the field of economic and social analysis are treated. The first one is the analysis of the Hungarian hog-cycle. According to former investigations the stock of the Hungarian hogs has a 3-5 year cyclical movement, which was analysed in details by many authors in the last decades. Now we tried to show the cycle from a 64 element annual time-series by means of wavelet analysis. The results are a bit ambiguous, perhaps, because in the last two decades there were huge changes of the system of the Hungarian economic management and foreign relations, but the analysis revealed remarkable cyclic tendencies.

The second example shows how the wavelet transformation smoothes and compresses time series. The base of this application was the annual series of Hungarian GDP. In this application after a DHT (Discrete Haar Transformation) 70 percents of the wavelet coefficients (obviously the less absolute valued coefficients) have been deleted and using the inverse transformation, a good fitness of the new series has been detected. This example shows in a rather convincing way the information compression of the wavelet transformation. It is to note that in case of a much extended data base (which can often be getting in the big data analysis), this method of information compression can yield much more spectacular results.

The next example was only a trial: according to some views the common features of time series are much more detectable from the transformed series than from the original ones. So our hypothesis was that the similarities of two or more series are higher in case of wavelet transformed series. The correlation matrix among the main macroeconomic indicators proved that the hypothesis is good – all of the elements of the new correlation matrix are higher than those, computed from the original series.

Finally in a demographic example we investigate the problem of Hungarian mortality. Our task was to separate the key-points of this 64 element time series, points that separate the particular parts of the mortality curves. Using the wavelet (Haar) transformation, 3 points of the curve has been detected and analysed. These points (1969, 1985, 2001) show those years, when the mortality curve has important, long-range turns.

The conclusion of the article is that this method is extremely promising in the field of time series and spatial series applications, but in order to realize it a more intensive cooperation among mathematics, statistics, modelling and informatics seems to be necessary.