

APLICAREA ANALIZEI COMPONENTELOR INDEPENDENTE ÎN EXTRAGEREA FACTORILOR FINANCIARI

Adela Buzuloiu

Institutul Național de Cercetare-Dezvoltare în Informatică,
ICI, București
e-mail: adela@ici.ro

Emil Sofron

Universitatea din Pitești
e-mail: sofron@upit.ro

Rezumat: Analiza componentelor independente este o tehnică apărută relativ recent, care încearcă să separe mărimi independente, care sunt inaccesibile direct din transformări ale acestora care sunt măsurabile. În lucrare, se studiază modul în care o metodă din această categorie preluată din tehnicile de prelucrare de semnale, analiza liniară a componentelor independente, poate fi aplicată la identificarea factorilor financiari, implicați în unul din modelele pieței de capital, și anume modelul APT. În prima parte, sunt prezentate succint modelul analizei ICA și modelul APT, iar apoi sistemul utilizat pentru studiul aplicabilității ICA în extragerea factorilor financiari, implicați în modelul APT. În a doua parte a lucrării, sunt prezentate rezultatele obținute de autori în demersul de a extrage factorii economici independenți, pe baza cotațiilor bursiere ale unor companii din lista Dow-Jones. Rezultatele obținute îndreptățesc utilizarea ICA în aplicații financiare datorită capacității metodei de a izola factorii care influențează piețele de capital.

Cuvinte cheie: analiza componentelor independente, piața de capital, modelare.

1. Introducere

Adeseori, apare în practică necesitatea obținerii unei informații „ascunse” în mărimi măsurate. Informația aceasta utilă există, dar reconstituirea ei este, adesea, foarte dificilă în absența unor date suplimentare despre informația căutată. Pentru acest caz, s-au dezvoltat tehnici care pornesc de la un minim de ipoteze, nu încearcă să modeleze matematic procesele și se bazează pe o fundamentare interdisciplinară, pornind de la relații provenind atât din statistică, cât și din teoria informației. Succesul acestor tehnici a depins în mare măsură de creșterea puterii de calcul a calculatoarelor, care a facilitat, în ultimii ani, dezvoltarea și implementarea unor metode performante și cu rezultate notabile, chiar și în astfel de cazuri, desigur în condițiile unor ipoteze și restricții.

Analiza componentelor independente (Independent Component Analysis, ICA) este o astfel de tehnică și este un instrument matematic, care își are originea în problema separării oarbe a surselor (problema care la rândul ei provine din prelucrarea semnalelor). Problema separării oarbe a surselor este cel mai bine ilustrată de problema „cocktail-party”: „dacă mai multe microfoane înregistrează simultan mai mulți vorbitori, cum se pot separa semnalele originale între ele?” Prin generalizare, se obține problema identificării unor influențe suprapuse, din măsurarea unor mărimi oarecare, ce reprezintă o transformare a datelor originale, necunoscute atât ca natură, cât și ca valoare. Aceste date originale pot fi inaccesibile, perturbate de zgomot sau necunoscute ca natură. Dacă se fac anumite ipoteze despre model, ICA demonstrează că se pot separa sursele și se poate determina transformarea. Două proprietăți esențiale caracterizează metoda ICA:

- se caută mărimi care pot fi surse, componente, variabile latente, care nu sunt vizibile de la început și care pot fi denumite generic „surse”;
- sursele sunt separate, pornind de la ipoteza independenței statistice mutuale.

2. Modelul ICA

Cel mai simplu model ICA presupune existența a n semnale independente $s_1(t), \dots, s_n(t)$ și m mixări de semnale $x_1(t), \dots, x_m(t)$, cu $m \geq n$, măsurate la momente discrete de timp k . Aceste mixări sunt liniare și instantanee. Modelul ICA este, în aceste condiții, descris de ecuația:

$$x_j(t) = \sum_{i=1}^n a_{ij} s_i(t) \quad (1)$$

sau sub formă vectorială:

$$\mathbf{x}(t) = \sum_{i=1}^n \mathbf{a}_i s_i(t)$$

Relația (1) poate fi scrisă în forma cea mai generală, cu păstrarea dependenței temporale și includerea componentei aditive de zgomot, sub forma:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) = \sum_{i=1}^n s_i(t) \mathbf{a}_i + \mathbf{n}(t) \quad (2)$$

unde $\mathbf{n}$ este componente aditivă de zgomot. Modelul simplificat ICA este, de regulă, scris sub forma compactă:

$$\mathbf{x} = \mathbf{A} \mathbf{s} + \mathbf{n} \quad (3)$$

în care vectorul $\mathbf{s}$ este de fapt vectorul sursă la momentul de eşantionare k : $\mathbf{s}(k) = [s_1(k), \dots, s_n(k)]^T$, care conţine semnalele originale, mutual independente. Matricea constantă $\mathbf{A}$ ($m \times n$) este matricea de mixare, cu coeficienţi necunoscuţi, care reprezintă chiar coeficienţii de mixare: $\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n]$. Vectorii $\mathbf{a}_i$ sunt vectori de bază.

În condiţiile enunţate mai sus, practic, problema ICA se reduce (formal) la găsirea unei transformări liniare:

$$\hat{\mathbf{s}} = \mathbf{B} \mathbf{x} \quad (4)$$

unde: $\hat{\mathbf{A}} = \mathbf{B}^{-1}$

iar zgomotul a fost inclus printre surse.

Dimensiunea vectorilor implicaţi în model reprezintă o limitare a modelului ICA. Raportul dintre numărul de surse necunoscute (n) şi semnale măsurate (m) trebuie să fie subunitar pentru a se putea estima sursele. Dacă $n > m$, se pot găsi, totuşi, soluţii în anumite cazuri speciale – de exemplu în cazul surselor binare [20], dar, în general, problema se consideră a fi *subdeterminată*. Mai mulţi senzori decât surse ($n < m$) duc la *supradeterminare* – măsurătorile suplimentare pot fi folosite pentru reducerea zgomotului, dar cazul $n < m$ ridică, adesea, probleme de convergenţă sau computaţionale.

O nedeterminare de bază este introdusă de model deoarece este imposibil a se determina o soluţie unică. Pornind de la soluţia $\hat{\mathbf{A}}\hat{\mathbf{s}}$ se pot construi alte soluţii $\hat{\mathbf{A}}'\hat{\mathbf{s}}'$ cu:

$$\mathbf{x} = \hat{\mathbf{A}}\hat{\mathbf{s}} = \hat{\mathbf{A}}\mathbf{P}\mathbf{A}\mathbf{A}^{-1}\mathbf{P}^T\hat{\mathbf{s}} = \hat{\mathbf{A}}'\hat{\mathbf{s}}' \quad (5)$$

unde $\mathbf{P}$ este o matrice de permutare şi $\mathbf{A}$ o matrice diagonală de scalare. Cu alte cuvinte, ordinea în care sunt identificate sursele rămâne arbitrară, ca şi amplitudinea semnalelor.

Pentru identificarea componentelor independente, se cunosc, până în prezent, mai multe clase de algoritmi „importanţi”, fiecare cu variante şi îmbunătăţiri. Marea lor majoritate se bazează pe reţele neuronale, şi presupun o anumită procesare prealabilă a datelor. Preprocesarea – aducerea la forma standardizată şi eliminarea surplusului de date prin reducerea dimensiunii datelor de intrare – elimină riscul supraînvăţării şi micşorează efortul de calcul.

Problema se reduce la determinarea unei estimate a inversei matricei de mixare, pornind de la o matrice oarecare, aceasta fiind apoi actualizată la fiecare pas cu regula de bază:

$$\mathbf{W} = \mathbf{W} + \Delta\mathbf{W} \quad (6)$$

în care $\mathbf{W}$ este matricea de ponderi, care se antrenează pe baza datelor de intrare, aplicate reţelei neuronale.

Majoritatea algoritmilor sunt, de fapt, familii de algoritmi parametrizaţi prin funcţiile neliniare utilizate [11]. Fiecare algoritm se bazează pe un criteriu de minimizare. Aceşti algoritmi sunt:

- algoritmul cu punct fix (FastICA), care se bazează pe informaţia mutuală şi care, deşi are dezavantajul că nu poate fi utilizat în aplicaţii on-line, are avantajul unei viteze de convergenţă mult mai mari decât a celorlalţi algoritmi [12];
- algoritmul bazat pe analiza neliniară a componentelor principale prin metoda recursivă a celor mai mici pătrate, algoritmul adaptiv bazat pe descreşterea de gradient [21];
- algoritmul Bell-Sejnowski, bazat pe maximizarea informaţiei şi pe o metodă de gradient stocastic [3];
- algoritmul Amari-Cichocki-Yang, care este derivat din algoritmul Bell-Sejnowski şi care utilizează o metodă de gradient natural [1];
- algoritmul adaptiv EASI (Equivariant Adaptive Source Separation via Independence), bazat pe echivarianţă [4].

Dintre algoritmii de mai sus, experimentele din lucrarea de faţă au fost făcute utilizându-se algoritmul cu punct fix FastICA*, datorită vitezei superioare de convergenţă. Criteriul de maximizare utilizat de FastICA este:

$$J_G(\mathbf{w}) = [E\{G(\mathbf{w}^T \mathbf{v})\} - E\{g'(\mathbf{v})\}]^2 \quad (7)$$

în care $\mathbf{v}$ este o variabilă standard cu repartiţie de probabilitate gaussiană şi G funcţia de contrast.

Regula de învăţare pentru extragerea unei componente este:

$$\mathbf{w}^+ = \mathbf{w} - \mu [E\{\mathbf{x} g(\mathbf{w}^T \mathbf{x})\} - \beta \mathbf{w}] / [E\{g'(\mathbf{w}^T \mathbf{x})\} - \beta] \quad (8)$$

*package Matlab de domeniu public, <http://www.cis.hut.fi/projects/ica/fastica/>

$$w^* = w^+ / \|w^+\|$$

Funcția g , care apare în relația de mai sus, este neliniaritatea introdusă, iar μ este rata de învățare.

3. Analiza evoluției cotațiilor bursiere

3.1. Contextul problemei

Evoluția prețului acțiunilor bursiere sau a indicilor bursieri interesează atât analiștii, cât și investitorii. Dinamica accentuată a piețelor de capital și multitudinea factorilor implicați fac din predicția evoluției cotațiilor bursiere o problemă dificilă, până la imposibil, de rezolvat. De regulă, evoluția prețurilor se analizează cu metode statistice și probabilistice. Studiul evoluției cotațiilor pe piața de capital se bazează pe *ipoteza pieței eficiente*, potrivit căreia prețul unui titlu financiar reflectă totalitatea informației disponibile în legătură cu acest titlu (și cu piața financiară din care face parte, evident). Orice informație nouă determină modificări de cotație, care se reflectă, practic, în modificări ale prețului din viitor. Cu alte cuvinte, informația duce la un proces rapid de ajustare a prețului titlului respectiv. Implicația este că seria de timp asociată prețului respectiv este o serie de timp pur aleatoare, iar efectul acesteia – imposibilitatea predicției acestor prețurilor.

Analiza distribuției a sugerat, încă din anul 1965 [12], faptul că acestea reprezintă serii de timp fără memorie („trecutul nu poate fi utilizat în nici o formă pentru a face predicții”), [16]. Această ipoteză, denumită ipoteza drumului aleator, se bazează pe ipoteza pieței eficiente și a fost multă vreme unanim acceptată în comunitatea academică. Practicienii au dezvoltat (totuși) în tot acest timp diferite metode de analiză tehnică, pentru a deduce din istoria unei cotații (exprimate în diferențe de preț, volum tranzacționat etc.) trendul probabil viitor al respectivei cotații [13]. O dată cu diversificarea metodelor de analiză statistică și cu creșterea capacității de calcul, în ultimul deceniu, ipoteza pieței eficiente a fost contestată și diverși cercetători au propus metode de a extrage trendul din seriile de timp ale cotațiilor bursiere [16], [2], [7], pentru a dovedi astfel că ipoteza drumului aleator nu este aplicabilă în cazul seriilor de timp asociate cotațiilor bursiere. În teoria modernă a portofoliilor, se acordă o importanță sporită analizei statistice a seriilor de timp, atașate prețurilor acțiunilor, și, în special, utilizării statisticilor de ordin superior. Modelul real al piețelor de capital este încă subiect de dezbatere și se mai caută dovezi care să susțină fie teoria predictibilității, fie cea a imposibilității predictibilității evoluției cotațiilor bursiere.

Pornind de la un model financiar clasic, modelul APT (Arbitrage Pricing Theory), se poate studia aplicabilitatea unor metode precum ICA pentru estimarea respectivelor factori și astfel, implicit, pentru îmbunătățirea performanțelor modelului. Accentul cade, astfel, pe *căutarea sensului* care poate fi atașat factorilor independenți identificați.

3.2. Modelul APT

O aserțiune importantă în teoria riscului financiar susține că riscul total al unui portofoliu din care fac parte mai multe titluri este mai mic decât riscul unui singur titlu *dacă* portofoliul este bine diversificat. Începând cu modelul medie – dispersie propus de H. Markovitz în 1956, diverși cercetători și-au pus problema modelării riscului unui titlu, și prin extensie a unui portofoliu de acțiuni. În esență, aceste modele exprimă o relație de dependență între profitul pe care îl poate înregistra un titlu și riscul asociat investiției respective. Trebuie precizat că profitul – adică diferența dintre prețul actual și cel viitor al titlului respectiv – reprezintă premiul de risc al acelei acțiuni, adică premiul care recompensează investitorul pentru faptul că a riscat investind în acel titlu.

Datorită dinamicii extreme a piețelor de capital, riscul – atât cel al unui singur titlu, cât și cel al unui portofoliu întreg – are o volatilitate foarte mare: zilnic trebuie reevaluate riscurile titlurilor din portofoliu și compunerea lor în riscul portofoliului pentru a se minimiza pierderile. Iar evaluarea riscului unui titlu pleacă, după cum am precizat, de la analiza evoluției prețului său, mai bine zis a profitului. În teoria modernă a riscului financiar, se pleacă cel mai des de la modelul medie – dispersie, care apreciază riscul pe baza datelor cunoscute la un moment dat (evoluția istorică a cotațiilor, dinamica și volatilitatea lor) pe baza cărora se determină caracteristicile statistice ale titlului respectiv.

Modelul APT presupune că profitul unui titlu depinde de o combinație liniară a mai multor influențe macroeconomice (numite factori) și de o componentă de zgomot, dar care nu este comun tuturor titlurilor, ci este un zgomot specific titlului respectiv:

$$\tilde{r}_i = \bar{r}_i + \sum_{k=1}^K b_{ik} \underbrace{(\tilde{F}_k - E(\tilde{F}_k))}_{\tilde{f}_k} + \varepsilon_i \quad (9)$$

unde:

- $\tilde{f}_k$ sunt factori care reprezintă evenimente privitoare la diversele influențe asupra evoluției profitului, precum produsul intern brut, rata dobânzii, nivelul inflației etc.

$$E\{\tilde{f}_k\} = 0 \quad \forall k$$

- b_{ik} (încărcarea factorială a titlului i cu factorul k) măsoară sensibilitatea profitului titlului i la evenimente care privesc factorul k .

- $\tilde{\varepsilon}_i$ (riscul idiosincratic) este componenta din profitul titlului i care nu este corelat cu profitul altor titluri:

$$\text{Cov}(\tilde{\varepsilon}_i, \tilde{\varepsilon}_j) = 0 \quad \forall i \neq j$$

$$\text{Cov}(\tilde{\varepsilon}_i, \tilde{F}_k) = 0 \quad \forall i, k$$

Evenimentele „generale” afectează rata profitului **tuturor** titlurilor, dar în măsuri diferite. După cum se poate observa, modelul APT este complex și are un neajuns (din punctul de vedere al aplicabilității lui în practică): nu specifică nici natura, nici numărul factorilor independenți, acești factori urmând a fi identificați pentru fiecare titlu în parte. Din acest motiv, în prezent, implementarea modelului APT este foarte dificilă și modelul este utilizat puțin în practică. El presupune întâi identificarea factorilor care se construiesc empiric, în funcție și de scenariul viitorului economic, specific unui anumit investitor. Observația interesantă este că modelul presupune că factorii independenți sunt mixați liniar, iar validitatea principală a modelului nu este contestată. Din acest motiv, prezintă interes separarea acestor factori, pornind de la observarea evoluției profiturilor, cu ajutorul unei tehnici precum „Analiza Componentelor Independente” (în engleză ICA).

3.3. Sistemul utilizat

Sistemul utilizat pentru extragerea factorilor independenți este prezentat în figura 1. Configurația datelor de intrare a fost aleasă în așa fel încât să fie determinați cât mai bine factori semnificativi dintr-un sector de activitate. Încercarea de a selecta în urma unei singure analize toți factorii care influențează piața de capital ar fi condus la prea multe rezultate care ar fi fost mai dificil de interpretat.

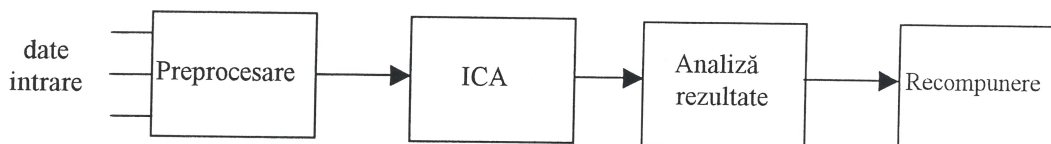


Figura 1. Sistemul utilizat pentru analiza factorilor independenți dintr-o serie de cotații financiare

a. Datele de intrare

În mod normal, aceiași factori care influențează o companie dintr-un domeniu oarecare influențează și celelalte companii din același domeniu, iar factorii semnificativi pentru domeniu se vor regăsi (cu o pondere diminuată) și printre factorii care influențează bursa de valori la nivelul unei țări. Pe de altă parte, trebuie să existe influențe resimțite de toate companiile, ca de exemplu, influența nivelului de referință al dobânzii din țara respectivă sau conjunctura internațională. Pentru a fi captați cât mai mulți factori semnificativi, s-au inclus, printre datele de intrare, 3 categorii de cotații, toate reprezentate de seriile de timp din perioada 01.01.2000 -31.12.2002 (și totalizând un număr de 650 de eşantioane), astfel:

- 7 indici bursieri majori: Dow-Jones, Amex, S&P 500, Nikkei, Toronto, FTSE și IXIC; acești indici reflectă evoluția în ansamblu a burselor de valori din mari centre bursiere din întreaga lume, pentru a putea fi surprinși atât factori macroeconomici, cât și factori legați de conjunctura internațională;
- pentru a completa informația legată de factorii macroeconomici și de conjunctură, au fost incluse printre datele de intrare și:
 - cursurile de schimb \$/Euro și \$/Yen, selectate deoarece factorii economici dominanți sunt SUA, Europa și Japonia;
 - evoluțiile obligațiunilor pe termen mediu (5 ani) și lung (30 ani) americane (în acest fel a fost inclusă în sistem dobânda de referință);
- din lista de 30 de companii pentru care se determină indicii Dow-Jones, au fost selectate 7 companii cu profil de activitate legat de tehnologia informației și telecomunicații: IBM, Intel, Microsoft, SBC Communications, Hewlett Packard și 3M. Astfel se încearcă modelarea factorilor care influențează companiile dintr-un anumit sector de activitate, în idea că analiza unui sector economic oferă informații care pot fi apoi extrapolate cu ușurință

pentru alte sectoare economice, iar reuniunea factorilor identificați pentru fiecare sector în parte reprezintă apoi factorii de la nivelul național.

Seriile de timp pentru un index – Dow - Jones, un curs de schimb (\$/Euro) și o companie (AT&T) sunt prezentate în figura 2.

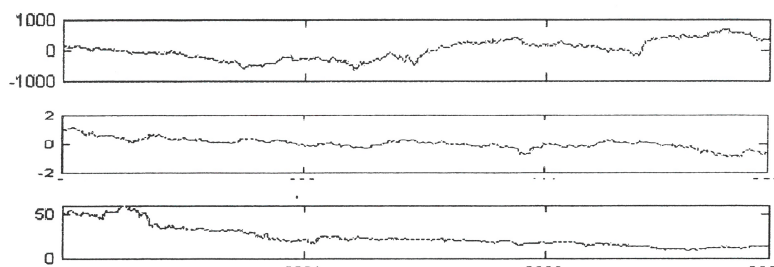


Figura 2. Seriile de timp atașate unor mărimi financiare pentru perioada 01.01. 2000 – 01. 01. 2003 (de sus în jos: indicele Dow-Jones, cursul (\$/Euro și cotația AT&T)

b. Preprocesarea

Prețurile cotațiilor bursiere au o dinamică și o volatilitate foarte mare. Din acest motiv, înaintea procesării uzuale pentru ICA, s-au extras din seriile de timp ale prețurilor seriile profiturilor procentuale, după relația:

$$x_i(t) = \frac{p_i(t) - p_i(t-1)}{p_i(t-1)} \quad (10)$$

Profiturile rezultate pentru cazul firmei AT&T sunt reprezentate în figura 3.

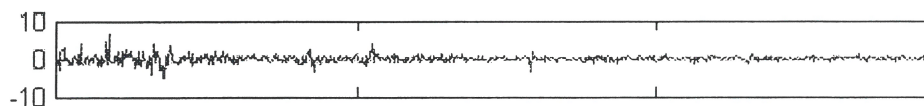
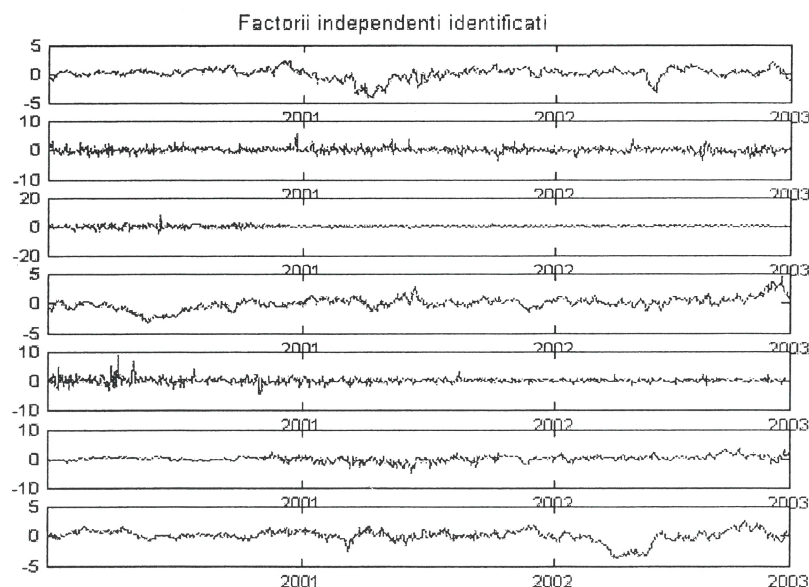


Figura 3. Profitul acțiunilor firmei AT&T, perioada 01. 01. 2000 – 01. 01. 2003

c. Identificarea factorilor independenți cu ajutorul ICA

Din mărimile reprezentând seriile de timp, ale celor 19 factori prezentați mai sus, s-au extras vectori de date cu câte 650 de valori, reprezentând perioada de timp 01 ianuarie 2000 – 31 decembrie 2002 și, din aceste date, s-au extras factorii independenți, după metoda „Analizei Componentelor Independente”.

În cadrul analizei, nu au rezultat componente corelate care să fie eliminate, astfel încât au putut fi izolate 19 componente independente, reprezentate în figura 4; după cum se poate observa, cele 19 mărimi sunt diferite.



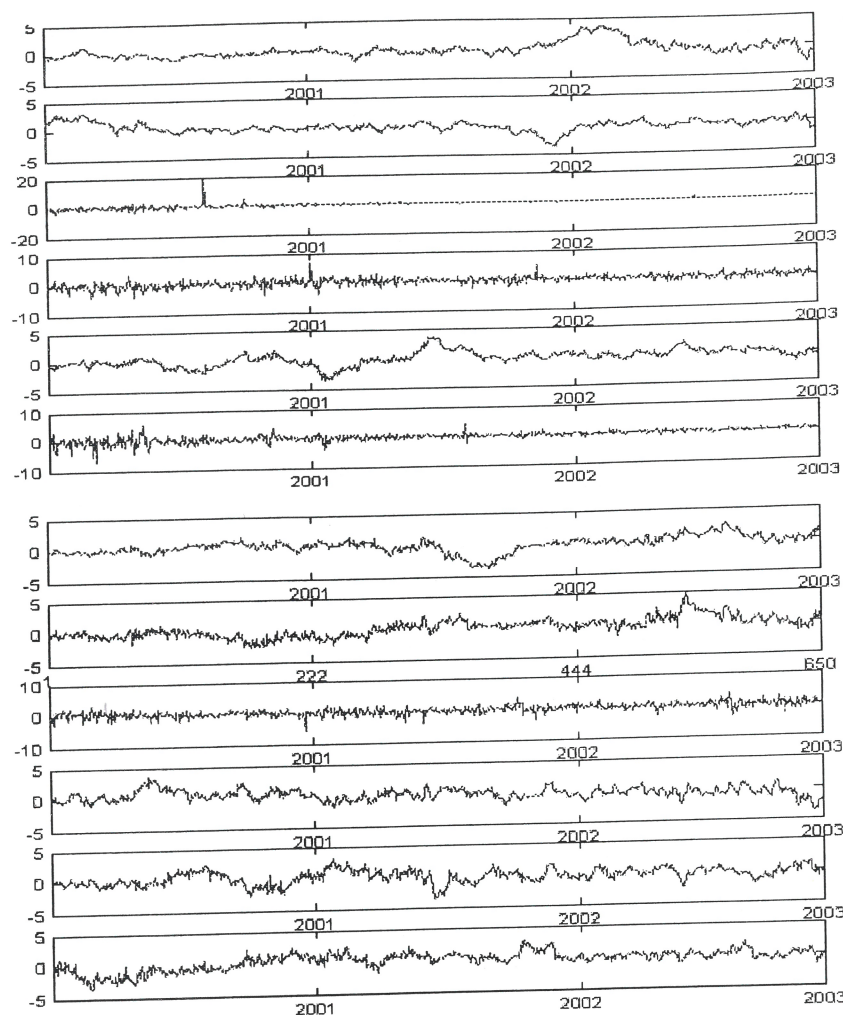


Figura 4. Factorii independenți identificați

d. Analiza rezultatelor

Toate mărimile izolate – care vor fi considerate, în continuare, factori independenți f_i – au un vector de bază aferent a_i coloană în matricea de mixare A . Conform ecuației 1, fiecare mărime x_i din setul de intrare poate fi exprimată ca:

$$x_i = \sum_{i=1}^n a_i f_i$$

Factorii independenți au rezultat de dispersie unitară (în urma procedurii de decorelare), astfel încât sortarea f_i se poate face doar după norma L_∞ . Pe de altă parte, vectorii de bază a_i - care pot fi asimilați cu factorii b_{ik} din ecuația 9 – au valori proprii diferite și pot fi ierarhizați după norma L_2 (vezi figura 5). După cum se poate observa, unii vectori de bază aduc un aport mai mare al factorului independent corespunzător în datele de intrare decât alții.

Deoarece spectrul de valori al vectorilor de bază este destul de mare, s-au reținut 10 vectori de bază, cu cele mai mari valori și anume $a_2, a_5, a_6, a_8, a_9, a_{10}, a_{14}, a_{15}, a_{17}, a_{19}$, într-o matrice de mixare numită A_{major} care captează influențele dominante ale factorilor independenți. Restul vectorilor de bază formează matricea A_{minor} .

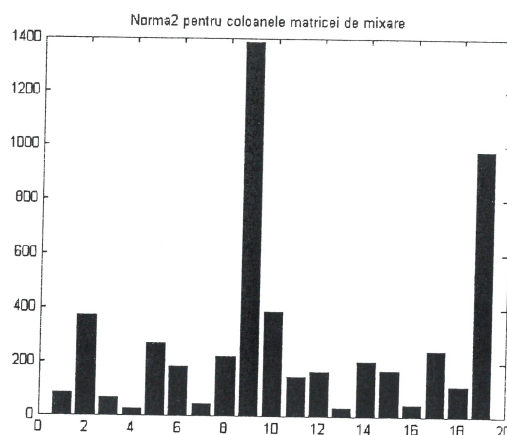


Figura 5. Ierarhizarea vectorilor de bază din matricea de mixare

Tabelul 1. Ierarhizarea vectorilor de bază din matricea de mixare

coloana	1	2	3	4	5	6	7	8	9
norma L_2	81	369	64	24	267	183	41	218	1387

Tabelul 1. Cont

coloana	10	11	12	13	14	15	16	17	18	19
L_2	389	146	164	31	204	168	43	245	110	981

e. Recompunerea rezultatelor

O ipoteză plauzibilă de care se poate face uz este că, reținând cei mai semnificativi 10 factori, se pot reconstitui semnalele originale, fără o pierdere semnificativă de precizie.

Pentru aceasta, s-au recompus mărimile:

$$\hat{x}_1(t) = A_{major} f_i$$

și s-au determinat abaterile înregistrate prin calculul abaterii medii absolute normalizate:

$$n m a e = \frac{\sum_{t=1}^N \|x(t) - \hat{x}_1(t)\|}{\sum_{t=1}^N \|x(t)\|}$$

Rezultatele sunt redată în Tabelul 1 și în graficele din figura 6, pentru fiecare din cele 19 mărimi. După cum se poate observa, se pot separa 2 categorii mari: pe de o parte, indicii bursieri și cotațiile valutare, care sunt modelate destul de bine cu ajutorul celor 10 factori, pe de altă parte, companiile care nu sunt, practic, deloc modelate de acești factori.

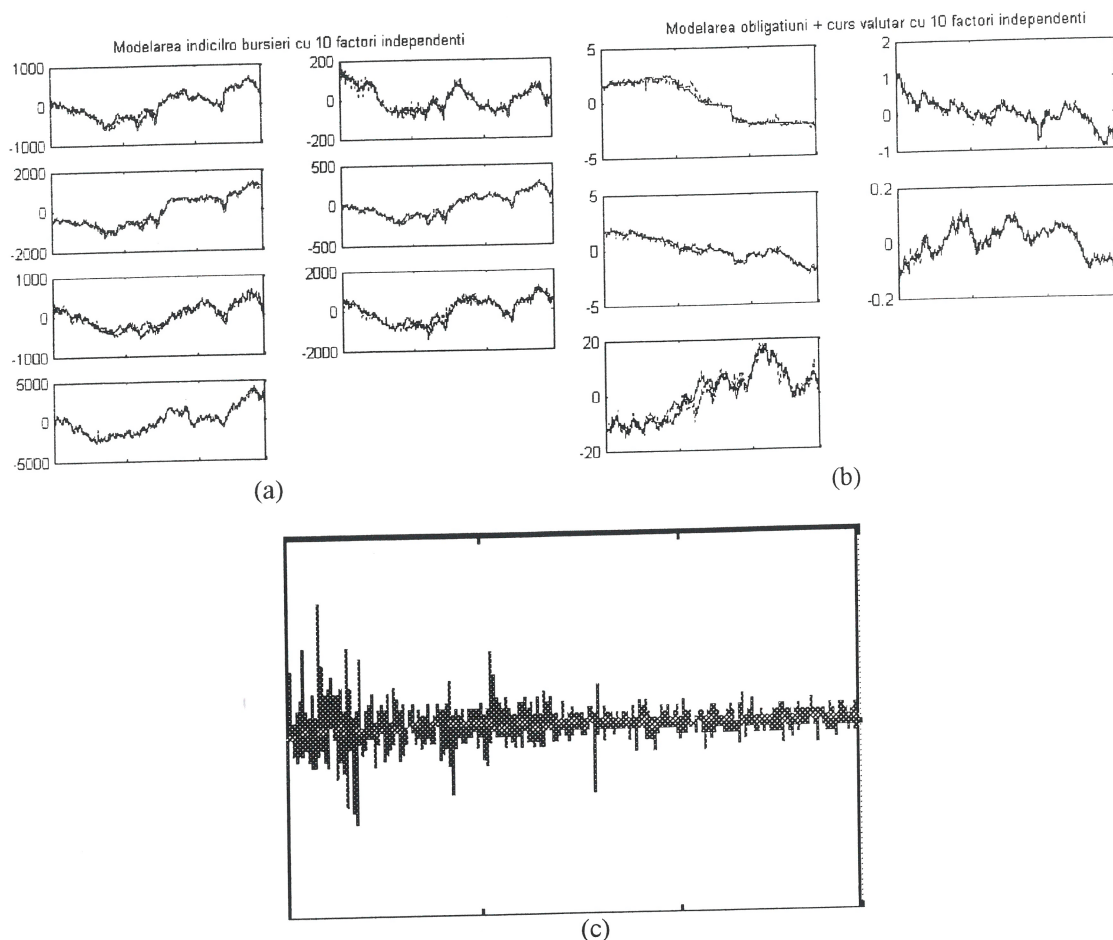


Figura 6. Modelarea cu ajutorul a 10 factori independenți a indicilor bursieri (a), a cursurilor de schimb (b) și a acțiunilor (c)

Tabelul 2. Eroarea medie absolută pentru reconstituirea cu factorii „majori”

mărime	abaterea medie absolută normalizată						
indici	0.2016	0.1963	0.0938	0.2047	0.2404	0.2611	0.0857
obligațiuni	0.1154	0.1851	0.1564	0.1428	0.2594		
acțiuni	0.9792	0.9801	0.9772	0.9877	0.9947	0.9717	0.9774

Pentru a se aprecia efectul celor 9 factori considerați reziduali, se repetă estimarea mărimilor cu ajutorul matricei de mixare A_{minor} .

$$\hat{\mathbf{x}}_2(t) = A_{\text{minor}} \mathbf{f}_i$$

cu rezultatele din Tabelul 3 și reprezentarea grafică din figura 7. De data aceasta, rezultatele sunt complementare: contribuția la evoluția indicilor bursieri există și, uneori, urmărește alura seriei originale de timp, în vreme ce cotațiile companiilor sunt descrise foarte bine de acești factori.

Tabelul 3. Eroarea medie absolută pentru reconstituirea cu factorii „minori”

mărime	abaterea medie absolută normalizată						
indici	0.9539	0.9578	0.9882	0.9618	0.9437	0.9403	0.9954
obligațiuni	1.0049	0.9697	0.9792	0.9802	0.9381		
acțiuni	0.1792	0.1917	0.1862	0.1362	0.1270	0.2364	0.1619

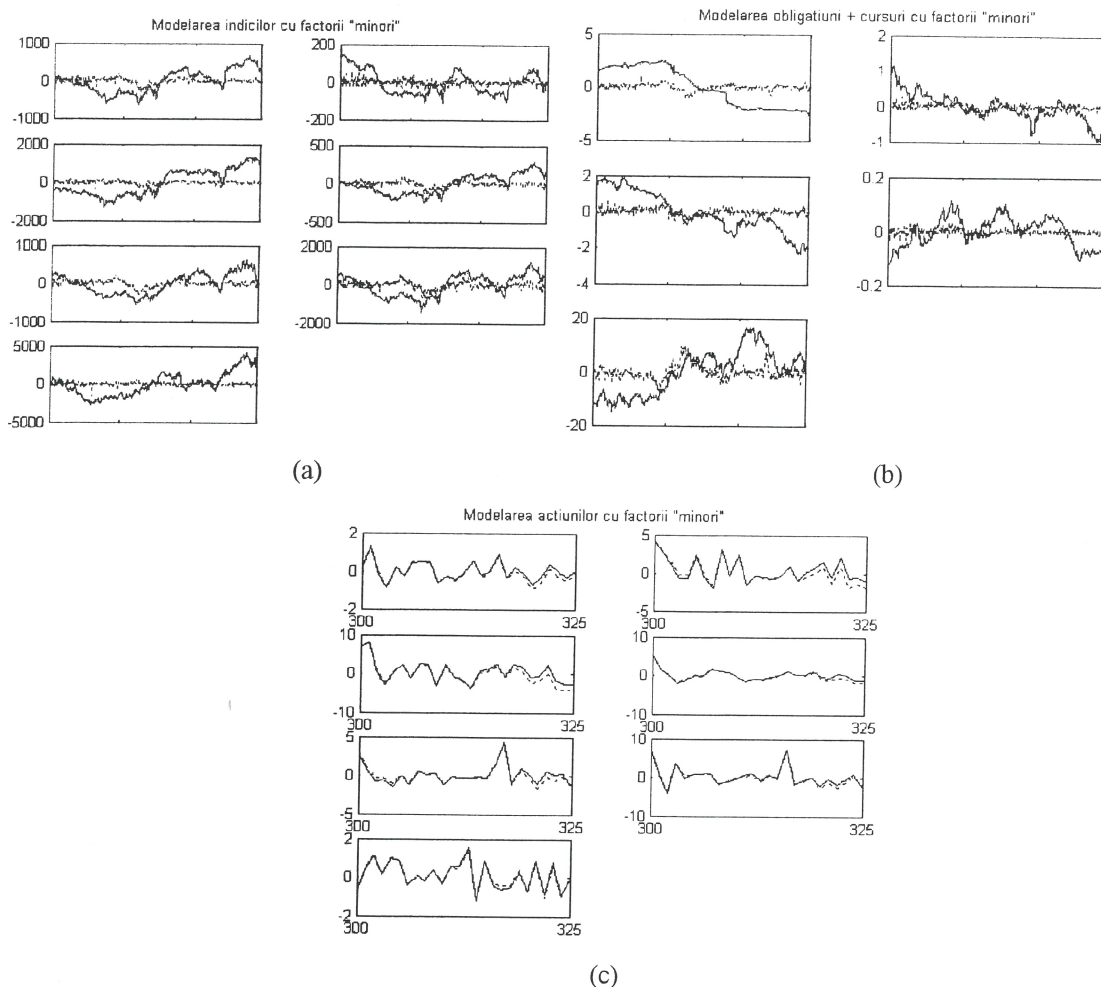


Figura 7. Modelarea cu ajutorul ultimilor 9 factori independenți a indicilor bursieri (a), a cursurilor de schimb (b) și a acțiunilor (c) (în cazul acțiunilor s-a reprezentat o lupă de timp)

În concluzie, printre cei 19 factori independenți izolați, se regăsesc 2 categorii:

- factorii macroeconomici care determină evoluția la nivelul bursei de valori în ansamblu,
- factori la nivel de sector economic (inclusiv de nivel internațional) care determină evoluția cotațiilor bursiere din domeniul respectiv.

4. Concluzii

Analiza datelor financiare este o problemă spinoasă, iar procesele care stau la baza funcționării piețelor de capital sunt complexe, insuficient cunoscute, eventual haotice, foarte dinamice și de complexitate din ce în ce mai mare datorită efectelor globalizării ([2], [16]). Dintre factorii macroeconomici semnificativi, în [9] sunt identificați un index al producției industriale, modificări în dobânda de referință, „răsuciri” (twist) în curba randamentului (diferența dintre curba randamentului promisă la obligațiuni guvernamentale pe termen lung și pe termen scurt) și inflația neanticipată. La aceștia este probabil că se mai adaugă și alții, legați de contextul internațional, de evenimente sociale etc. Încercarea de a identifica acești factori este dificilă. Prin analiza efectuată în lucrare s-a demonstrat că separarea acestor factori poate fi făcută și cu ajutorul „Analizei Componentelor Independente”. Procedul prezentat în această lucrare s-a utilizat pentru analiza profiturilor acțiunilor din domeniul IT și comunicații, dar acesta poate fi cu ușurință extins la alte sectoare economice urmând același algoritm. Selecția adecvată a mărimilor de intrare influențează performanțele sistemului, dar rezultatele obținute justifică interesul și permit extinderea ariei de aplicare a Analizei liniare a componentelor independente în zona aplicațiilor financiare.

Bibliografie

1. AMARI, S., A. CICHOCKI, H.H. YANG: A New Learning Algorithm for Blind Source Separation. În: Advances in Neural Information Processing, No. 8., MIT Press, Cambridge, MA, 1996, pp. 757-763.

2. **BACK, A. D., A. S. WEIGEND:** A First Application of Independent Component Analysis to Extracting Structure from Stock Returns. În: *Int. J. on Neural Systems*, 8(4), 1998, pp. 473-484.
3. **BELL, A.J., T.J. SEJNOWSKI:** An Information-Maximization Approach to Blind Separation and Blind Deconvolution. În: *Neural Computation* 7, 1995, pp. 1129-1159.
4. **CARDOSO, J.-F., B. HVAM LAHELD:** Equivariant Adaptive Source Separation. În: *IEEE Trans. on Signal Processing*, 44(12), 1996, pp. 3017-3030.
5. **CARDOSO, J.-F.:** Infomax and Maximum Likelihood for Source Separation. În: *IEEE Letters on Signal Processing* 4, 1997, pp. 112-114.
6. **CARDOSO, J.-F.:** Blind Signal Separation: Statistical Principles. În: *Proc. IEEE*, vol. 9, no. 10, 1998, pp. 2009 - 2025.
7. **CHA, S. M., L. W. CHAN:** Applying Independent Component Analysis to Factor Model in Finance, *Intelligent Data Engineering & Automated Learning – IDEAL 2000*, eds K. S. Leung, L. W. Chan, H. Meng, Springer, 2000, pp. 538-544.
8. **CHAN, L. W., S. M. CHA:** Selection of Independent Factor Model in Finance. În: *Proc. of 3rd International Conference on Independent Component Analysis and Blind Signal Separation*, 2001 December 9-12, 2001 - San Diego, California, USA.
9. **CHEN, N.F., R. ROLL, S. ROSS:** Economic Forces and the Stock Market. În: *Journal of Business*, Vol. 59, nr. 3, 1986, pp. 383-403.
10. **COMON, P.:** Independent Component Analysis – A New Concept? *Signal Processing*, 36, 1994, pp. 287-314.
11. **DABLEMONT, S., G. SIMON, A. LENDASSE, A. RUTTIENS, F. BLAYO, M. VERLEYSEN:** Time Series Forecasting with SOM and Local Non-linear Models – Application to the DAX30 Index Prediction WSOM'2003 Proc. – Workshop on Self - Organizing Maps Hibikino (Japan), 11-14 September 2003, pp. 340-345.
12. **FAMA, E.:** The Behavior of Stock Market Prices. În: *Journal of Business*, nr. 38, 1965, pp. 34-105.
13. **EDWARDS, FR.:** Do Futures Trading Increase Stock Market Volatility? În: *Financial Analysts Journal*, nr. 44, 1988, pp. 63-69.
14. **HYVÄRINEN, A., E. OJA:** A Fast Fixed-point Algorithm for Independent Component Analysis. În: *Neural Computation*, 9(7), 1997, pp. 1483-1492.
15. **HYVÄRINEN, A.:** Survey on Independent Component Analysis. În: *Neural Computing Surveys* 2, 1999, pp. 94-128.
16. **JUTTEN, C., J. HERAULT:** Blind Separation of Sources, Part I: An Adaptive Algorithm Based on Neuromimetic Architecture. În: *Signal Processing*, 24, 1991, pp. 1-10.
17. **KARHUNEN, J., E. OJA, L. WANG, R. VIGARIO, J. JOUTSENSALO:** A Class of Neural Networks for Independent Component Analysis. În: *IEEE Trans. on Neural Networks* 8, 1997, pp. 486 - 504.
18. **KIVILUOTO, K., E. OJA:** Independent Component Analysis for Parallel Financial Time Series. În: *Proc. ICONIP'98*, Vol. 2, 1998, Tokyo, Japan, pp. 895-898.
19. **LENDASSE, A., E. DE BODT, V. WERTZ, M. VERLEYSEN:** Non-Linear Financial Time Series Forecasting – Application to the Bel 20 Stock Market Index. În: *European Journal of Economic and Social Systems* 14 No. 1, 2000, pp. 81-91.
20. **MAGDON-ISMAIL, M., A. NICHOLSON, Y. ABU-MOSTAFA:** Financial Markets: Very Noisy Information Processing, În: *Proc. of the IEEE*, Vol. 86, no. 11, nov. 1998.
21. **MOISÃO, R. L. M., F. M. PIRES:** Prediction Model, Based on Neural Networks, for Time Series with Origin in Chaotic Systems. În: *Proc. of Workshop Artificial Intelligence Techniques for Financial Time Series Analysis*, EPIA'2001, Porto, Dec. 2001.
22. **OLDEMIRO FERNANDES, M., L. TORGO:** Predicting Daily Returns for the IBM Stock. În: *Proc. of Workshop Artificial Intelligence Techniques for Financial Time Series Analysis*, EPIA'2001, Porto, Dec. 2001.
23. **PAJUNEN, P.:** A Competitive Learning Algorithm for Separating Binary Sources. În: *Proc. European Symposium on Artificial Neural Networks (ESANN'97)*, Bruges, Belgium, 1997, pp. 255-260.
24. **PAJUNEN, P., J. KARHUNEN:** Least-Squares Methods for Blind Source Separation Based on Nonlinear PCA. În: *Int. Journal of Neural Systems*, vol. 8, oct./dec. 1998, pp. 601-612.
25. **SCHOBEN, D. W. E., K. TORKKOLA, P. SMARAGDIS:** Evaluation of Blind Signal Separation Methods. *First International Workshop on Independent Component Analysis and Blind Signal Separation*, Aussois, France, January 11-15, 1999.