

Capitolul 1. Studiul dependenței de timp

1.1 Serii de timp

Observând în diferite momente caracteristicile cantitative ale unei unități de observare sau ale unei populații întregi, datele pe care le obținem astfel formează o serie de timp sau o serie dinamică.

Temperaturile înregistrate într-un loc oarecare într-o anumită perioadă, populația unei țări de-a lungul unui număr de ani, sau importurile într-un șir de ani, greutatea unui animal înregistrată în diferite etape ale vieții sale, etc. sunt exemple curente ale fenomenelor care furnizează serii de valori la intervale de timp succesive.

Datele acestea sunt din punctul de vedere al statisticii importante atât prin valori cât și prin ordinea în care apar aceste valori.

Vom trata în special cazul foarte frecvent întâlnit în care seriile de mărimi sunt echidistante în timp. Egalând intervalul de timp cu unitatea putem considera seria ca fiind definită în momentele $t = 1, 2, 3, \dots$ și vom nota valorile respective ale seriei cu x_1, x_2, \dots valoarea seriei din momentul t fiind deci x_t . Dacă din vreo rațiune oarecare dorim să ne referim și la momente anterioare momentului $t = 0$ putem nota seria astfel: $\dots x_{-3}, x_{-2}, x_{-1}, x_0, x_1, x_2, x_3 \dots$

În practică, condiția intervalului de timp egal nu constituie un impediment serios. Cele mai multe serii în publicațiile oficiale, cum sunt seriile privind fenomenele economice, demografice sau meteorologice, au intervale egale, bunăoară zilele, sau aproape egale - anii, sau cu bunăvoință aproximativ egale - lunile.

Datele sunt de obicei culese la intervale egale; alteori (la barometre de pildă) dispunem de înregistrări continue, din care putem extrage date echidistante în timp.

Deși exemplificările noastre se referă la serii dinamice, trebuie să menționăm că teoria acestor serii este aplicabilă și altor tipuri de date statistice. De pildă, un fir de bumbac sub lentila microscopului se prezintă în lungime ca o succesiune de ondulații care seamănă cu o serie dinamică oscilantă; valorile corespunzătoare conținutului de azot în diferite puncte ale unei fâșii de teren pot fi considerate drept mărimi ale unei serii în care variabila timp a fost înlocuită cu variabila spațiu. Aceste metode sunt aplicabile, și uneori chiar potrivite, ori de câte ori o variabilă statistică u este o funcție a variabilei t , aceasta din urmă fiind timpul sau spațiul liniar.

Caracteristica u poate fi continuă sau discretă, în funcție de o singură variabilă sau de mai multe variabile.

Numărul de oameni, de exemplu, este în mod necesar un număr întreg și de aceea o serie referitoare la numărul populației nu poate fi decât discretă. În variațiile ei, ploaia și temperaturile sunt însă variabile continue.

Pe de altă parte este posibil să examinăm dinamica unei singure varietăți, cum ar fi prețul grâului, de pildă sau cum ar fi cea referitoare la salarii, gradul de ocupare a forței de muncă și volumul producției industriale. În acest al 2-lea caz este de preferat să considerăm că fiecare variabilitate creează o serie dinamică separată (unidimensională) și să privim relațiile dintre aceste serii drept o variație combinată a mai multor serii.

Exemple de serii dinamice sunt date la paragraful rezervat exemplilor și aplicațiilor [3.4.]

Urmărind aceste exemple observăm că seriile dinamice se compun din 3 elemente constitutive: o variație de lungă durată numită tendință, variații ritmice cu perioade scurte și variații neritmice, întâmplătoare.

Desigur sunt serii dinamice în care nu se manifestă toate aceste 3 elemente.

În exemplul 1 apare aproape exclusiv numai tendința, în exemplul 2 aproape numai variațiile ritmice, iar în exemplul 3 pare la prima vedere că este vorba doar de fluctuații întâmplătoare.

În cercetarea lor, izolarea fiecăreia din cele 3 componente este o problemă esențială.

Să subliniem și faptul că este necesară o distincție între variația în perioade de timp îndelungate și variația în perioade de timp scurte. Deși această delimitare e în mare măsură arbitrară este necesară mai ales din punctul de vedere al activității economice, unde fluctuația ciclică pare că se desfășoară în perioade îndelungate, maximele înviorării și minimele depresiunii repetându-se, în medie, o dată la 10 ani. Dar o perioadă de 10 ani poate apare neînsemnată dacă o raportăm la reaparițiile succesive ale glaciațiunilor sau la ascensiunea și decăderea civilizațiilor. Ceea ce se numește tendință este, pentru fiecare caz în parte o problemă de decizie.

Ar fi mai corect să vorbim de fluctuații de lungă durată și de fluctuații de scurtă durată. Dar și aici s-ar ridica problema ce este o perioadă de lungă durată și ce este o perioadă de scurtă durată.

În viziune clasică o serie dinamică cu conținut economic este interpretată ca fiind o sumă de 4 componente.

O primă componentă se referă la tendința generală de durată, ce se manifestă în evoluția procesului economic considerat. În literatura anglo-saxonă aceasta tendință poartă denumirea de trend. De aici am adoptat și noi această denumire.

În context, această componentă o vom nota cu $T(t)$.

A doua componentă se referă la variațiile sezoniere $S(t)$ specifice fenomenelor și proceselor economice. Se pot da numeroase exemple de ritmuri de dezvoltări în timp în care se manifestă anumite efecte de durate mai scurte - săptămânale, lunare, de anotimp - prezentând o oarecare regularitate aproape strict periodică.

Este vorba de așa-numitele efecte sezoniere pe care le putem observa și identifica prin înregistrări succesive în diferite epoci ale anului. Ele sunt generate de diferențele generale între anotimpuri, de obiceiurile specifice din diferitele luni ale anului, de cereri și oferte cu caracter sezonier (producția și consumul de răcoritoare, producția și consumul de gaze naturale, etc.)

A treia și a patra componentă se referă la mișcarea ciclică (ondulatorie) și, respectiv, la comportamentul abaterii aleatoare. Mișcarea ciclică o vom nota cu $C(t)$ iar cea aleatoare cu $e(t)$.

Prin urmare, în sens clasic, modelul unei serii dinamice este aditiv, dat de următoarea expresie:

$$X(t) = T(t) + S(t) + C(t) + e(t)$$

Să mai menționăm ca în funcție de obiectivele imediate ale cercetării poate fi determinat reziduul global $R(t) = C(t)+e(t)$ folosind relația evidentă:

$$R(t) = X(t) - T(t) - S(t)$$

urmând apoi ca acesta sa fie studiat separat.

Metodele de investigare a tendinței sunt descrise în paragraful [2.2.] și se bazează pe estimarea parametrilor unei curbe de regresie. Metodele de estimare a parametrilor sunt foarte variate și diferă în funcție de specificul curbei. Cea mai utilizată metodă, metoda celor mai mici pătrate este folosită în implementarea algoritmilor de trend tratați. Discuția celorlalte metode cât și legătura lor cu metoda celor mai mici pătrate e descrisă în paragrafele ce urmează.

O altă metodă, cea a mediilor mobile, care este un caz special de regresie parabolică este tratată pe larg în [20].

Problemele care se ridică la determinarea sezonaliității sunt legate de determinarea perioadei seriei și apoi construcția ei ca o serie de armonici sau, în cazul în care rezultatul eliminării $X(t)-T(t)$ o permite, în sensul că îndeplinește condițiile ca numărul de valori să fie putere a lui 2 și datele măsurate să fie repartizate pe un număr întreg de perioade aplicarea algoritmului Transformatei Fourier Rapide.

Cât privește reziduul $R(t)$ format din $C(t)$ și $e(t)$ acesta se poate analiza și interpreta ca un proces staționar în sensul definit în paragraful ce tratează tendința de stabilizare: [2.4.1.].

Exemple de serii dinamice sunt date la paragraful rezervat exemplilor și aplicațiilor [3.4.].

1.2. Teoria estimăției

Estimarea parametrilor curbei de regresie (trend) sau a parametrilor ce intervin în funcția de sezonabilitate se face prin metode specifice.

Acești parametri pot fi estimați teoretic prin mai multe metode plecând de la premise diferite.

O primă metodă de estimare este estimarea prin metoda mediei condiționate. Aceasta metodă denumită și metoda celor mai mici dispersii a fost fundamentată de A. N. Kolmogorov și estimăția se obține prin minimizarea riscului definit ca medie a funcției de pierdere pătratică $c(X,T)$ dată de:

$$c(X,T) = (X-T) \cdot (X-T).$$

O a doua metodă este estimarea parametrilor prin metoda probabilității a posteriori maxime. Ideea metodei de estimare a parametrilor în acest mod stă la baza formulei lui Bayes încă de la jumătatea secolului XVII. De aceea estimațiile obținute cu ajutorul acestei metode se numesc adesea estimații bayesiene. Dacă se introduce funcția de pierdere uniformă $c(X,T)=0$ dacă $X(t)-T(t) < D/2$ $c(X,T)=1$ în caz contrar și se minimizează riscul se obțin parametrii.

A treia metodă este metoda verosimilității maxime elaborată de R. Fischer în 1912. Principiul este că se alege cea estimăție pentru care funcția de verosimilitate $f(X(t), T(t)) = P(X(t)|T(t))P(T(t))$ ia valoarea maximă. Se demonstrează că acest lucru este echivalent cu alegerea funcției de pierdere :

$$c(X,T) = 1 - \exp(-(X-T)^2/2).$$

A patra metodă este metoda minimax fundamentată de J. Newman și dezvoltată mai târziu de A. Waald. Conform acestei metode, estimăția trebuie astfel determinată, încât riscul maxim să devină minim sau, altfel spus, erorile maxime să devină minime. Ca și celelalte metode, optimalitatea estimăției se caracterizează cu ajutorul funcției de pierdere. În acest caz funcția de pierdere este :

$$c(X,T) = |X-T|.$$

Se poate demonstra că estimăția minimax coincide cu estimăția prin metoda verosimilității maxime. Metoda minimax este întrucâtva mai dificilă din punctul de vedere al efectuării calculelor față de toate celelalte metode expuse anterior.

A cincea metodă, metoda celor mai mici pătrate o discutăm pe larg în paragraful următor.

1.3. Estimarea parametrilor prin metoda celor mai mici pătrate

Enunț

Să presupunem că avem o variabilă aleatoare $Y=Y(t)$ și variabilele aleatorii $Z_1=Z_1(t)$, $Z_2=Z_2(t)$, ..., $Z_m=Z_m(t)$ peste mulțimea $\{1, \dots, n\}$ cu valori reale.

Căutăm funcția liniară care să lege dependența lui Y de dependențele variabilelor Z_1 , Z_2 , ..., Z_m sub forma :

$$f(t) = a_1 Z_1(t) + a_2 Z_2(t) + a_3 Z_3(t) + \dots + a_m Z_m(t) + a_{m+1}$$

sau, introducând variabila aleatoare constantă:

$$Z_{m+1} = Z_{m+1}(t) = 1$$

funcția devine liniară și omogenă în Z_1, Z_2, \dots, Z_{m+1} :

$$f(t) = a_1 Z_1(t) + a_2 Z_2(t) + a_3 Z_3(t) + \dots + a_m Z_m(t) + a_{m+1} Z_{m+1}(t)$$

Așadar să se determine coeficienții $a_1, a_2, \dots, a_m, a_{m+1}$ astfel încât $f(t)$ să aproximeze cel mai bine pe $Y(t)$ în punctele t_1, t_2, \dots, t_n .

Expresia să aproximeze cel mai bine reprezintă chiar esența problemei celor mai mici pătrate, și după cum îi spune și numele. se minimizează pătratele erorilor de aproximare.

Observații

Problema este consistent enunțată când $n > m$ în caz contrar problema având o infinitate de soluții.

Deoarece $Y = Y(t)$ este o funcție de timp mărginită problema în cazul $n < m$ are întotdeauna soluție unică.

Variabilele Z_1, Z_2, \dots, Z_{m+1} nu sunt variabile independente, ele sunt legate prin intermediul variabilei timp.

Am ales să prezentăm metoda celor mai mici pătrate în cazul regresiei liniare multiple deoarece este prezent un grad mare de generalitate.

Rezolvare

Fie funcția de eroare $Er = Er(t)$ definită prin :

$$Er(t) = f(t) - Y(t) = a_1 Z_1(t) + a_2 Z_2(t) + a_3 Z_3(t) + \dots + a_m Z_m(t) + a_{m+1} Z_{m+1}(t) - Y(t)$$

Minimizând erorile de aproximare, condiția de minim devine :

$$S = \sum_{i=1}^n Er^2(t_i) = \min.$$

Mărimea S depinde valorile pe care le iau coeficienții $a_k, k = 1, \dots, m+1$ deci minimul se atinge când:

$$\begin{aligned} \frac{\partial S}{\partial a_k} = 0 &\Leftrightarrow \frac{\partial}{\partial a_k} \sum_{i=1}^n \text{Er}(t_i)^2 = 0 \Leftrightarrow \sum_{i=1}^n \frac{\partial}{\partial a_k} \text{Er}(t_i)^2 = 0 \Leftrightarrow \\ &\Leftrightarrow \sum_{i=1}^n 2\text{Er}(t_i) \frac{\partial \text{Er}(t_i)}{\partial a_k} = 0 \Leftrightarrow \sum_{i=1}^n \text{Er}(t_i) \text{Er}'_{a_k}(t_i) = 0, \quad k = \overline{1, m+1} \end{aligned}$$

Ținând seama că:

$$\begin{aligned} \text{Er}(t_i) &= \sum_{j=1}^{m+1} a_j Z_j(t_i) - Y(t_i) \Rightarrow \\ \sum_{i=1}^n \left(\sum_{j=1}^{m+1} a_j Z_j(t_i) - Y(t_i) \right) \cdot \left(\sum_{j=1}^{m+1} \frac{\partial a_j}{\partial a_k} Z_j(t_i) \right) &= 0, \quad k = \overline{1, m+1} \Leftrightarrow \\ \sum_{i=1}^n \left(\sum_{j=1}^{m+1} a_j Z_j(t_i) - Y(t_i) \right) Z_k(t_i) &= 0 \Leftrightarrow \\ \sum_{i=1}^n \sum_{j=1}^{m+1} a_j Z_j(t_i) Z_k(t_i) - \sum_{i=1}^n Z_k(t_i) Y(t_i) &= 0 \Leftrightarrow \\ \sum_{j=1}^{m+1} a_j \sum_{i=1}^n Z_j(t_i) Z_k(t_i) &= \sum_{i=1}^n Z_k(t_i) Y(t_i), \quad k = \overline{1, m+1} \end{aligned}$$

Făcând notația:

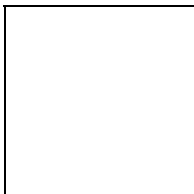
$$M(X) = \sum_{i=1}^n X(t_i) / n; \quad M(XY) = \sum_{i=1}^n X(t_i) Y(t_i) / n$$

sistemul de ecuații algebrice devine:

$$\sum_{j=1}^{m+1} a_j M(Z_j Z_k) = M(Z_k Y), \quad k = \overline{1, m+1}$$

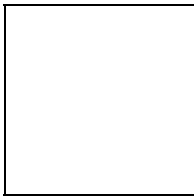
care este un sistem de $m+1$ ecuații și $m+1$ necunoscute cu determinantul sistemului ne nul ce admite o soluție unică. Să expunem cum se obține această soluție.

Scriind desfășurat sistemul de ecuații:



Pentru rezolvarea sa se folosește metoda lui Gauss cu pivot descrisă în continuare :

P0. Se construiește determinantul extins al sistemului



Matricea rezultată o numim $B=(b_{ij})_{i=1,m+2,i=1,m+1}$.

P1. $i=1$;

P2. Se caută $k \geq i$ pentru care coeficientul $b_{ki} = \text{maxim}$ în coloană.

P3. Se înlocuiește linia k cu linia i în această matrice.

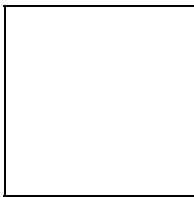
P4. Cu elementul b_{ii} facem 0 în coloană și împărțim linia i cu b_{ii} .

Deoarece transformările elementare nu modifică soluția sistemului aceasta nu e afectată de operațiile efectuate.

P5. Se face $i=i+1$;

P6. Dacă $i < m+1$ atunci salt la p2.

Aplicând acest algoritm asupra matricei sistemului obținem o matrice $(m+1) \times (m+2)$ ce are sub diagonala principală 0 și pe diagonala principală 1:



De aici se observă ușor soluția prin înlocuiri succesive în ecuații pornind de la ultima ecuație în sus :

$$a_{m+1} = b_{m+1,m+2} ;$$

$$a_m = b_{m,m+2} - a_{m+1} b_{m,m+1} ;$$

.....

$$a_1 = b_{1,m+2} - a_1 b_{1,2} - \dots - a_{m+1} b_{1,m+1} ;$$

1.4. Legătura metodei celor mai mici pătrate cu celelalte metode de estimare

Întrucât metoda celor mai mici pătrate este fundamentată teoretic pe baza repartiției normale, legătura ei cu celelalte metode de estimare a parametrilor poate fi ușor dovedită.

Într-adevăr, metoda celor mai mici pătrate reprezintă un caz particular al metodei de verosimilitate maximă. Anumiți parametri pot fi estimați prin metoda celor mai mici pătrate folosind relații matematice specifice altor metode de estimare.

Lucrările [15],[26],[27],[28],[29] aprofundează latura teoretică a metodei celor mai mici pătrate.

1.5. Analiza seriilor de timp prin metoda corelației

Considerând variabila $X=X(t)$ și variabila temporală t dacă valorile corespunzătoare celor 2 variabile sunt înregistrate putem grupa datele noastre într-un tabel cu dublă intrare (tabel de contingență) arătând frecvențele de valori care se află în interiorul unor intervale date.

Fiecare rând al acestui tabel dă repartițiile de frecvență ale primei variabile în care cea de-a doua e cuprinsă în intervalul respectiv.

O repartiție de frecvență grupată de acest tip, poate fi denumită o repartiție de frecvență bidimensională; dacă ne interesează însă în mod deosebit relația dintre cele 2 variabile, această repartiție poartă numele de tabel de corelație.

Un tabel de contingență are în plus față de un tabel de corelație faptul că poate fi folosit (și e de obicei folosit) în cazul acelor tabele în care gruparea se face în raport cu cantități nemăsurabile sau cu intervale definite mai puțin precis.

Suprafața de frecvență se obține prin ridicarea, în fiecare dreptunghi delimitat de 2 intervale liniare corespunzătoare celor 2 domenii pentru fiecare variabilă, a unui paralelipiped de înălțime direct proporțională cu valoarea numerică a frecvenței ce se află în acel domeniu bidimensional.

Cu cât micșorăm dreptunghiurile suprafața se apropie de ceea ce putem denumi o suprafață continuă.

Volumul delimitat de ea, planul orizontal și un domeniu dreptunghiular este chiar frecvența cazurilor din acel domeniu. Datorită complexității de reprezentare și formelor mult prea variate pe care le pot lua, aceste suprafețe nu ne oferă suficiente informații. În mod analog stau lucrurile și pentru stereograme care sunt similarul histogramei din cazul unei singure caracteristici.

Există o altă metodă de a reprezenta datele bidimensionale în mod grafic.

Fiecărui element al populației îi corespunde o pereche de valori $X(t), t$ care la rândul ei corespunde unui punct din plan de coordonate $(X(t), t)$.

O astfel de diagramă se numește o diagramă de împrăștiere.

Diagrama de împrăștiere dă o imagine mult mai bună a dependenței celor 2 variabile decât aceea obținută printr-o singură înregistrare a datelor ne grupate în tabel.

E clar că un tabel de corelație se poate trata prin metodele specifice tabelelor de contingență: calculul coeficientului de contingență medie pătratică al lui Pearson, calculul coeficientului lui

Cipurov, ș.a.m.d. și rezultatul obținut ne poate spune dacă variabilele sunt legate și cât de strânsă este această legătură.

Metodele pe care le prezentăm merg însă mult mai departe. Caracterul numeric al variabilelor și aranjamentul tabelului de corelație în intervale de grupă de mărime egală ne permit să abordăm problema investigației relațiilor dintre aceste variabile cu grad mult mai mare de precizie.

1.6. Corelație Simplă. Regresia liniară. Coeficientul de regresie

Să considerăm acum mediile aranjărilor.

Fie Ox, Oy două axe rectangulare reprezentând scalele celor două caracteristici variabile. Ca și în cazul diagramei de împrăștiere, putem trasa pozițiile mediilor; de exemplu dacă media unei linii a cărei valoare este centrată în y_1 este m_1 putem marca punctul a cărei abscisă este m_1 și a cărei ordonată este y_1 . Obținem astfel un punct corespunzător fiecărei linii și unul corespunzător fiecărei coloane.

Mediile liniilor și mediile coloanelor vor fi în general legate mai mult sau mai puțin de anumite curbe netede. De exemplu se pot grupa în jurul unei drepte.

Astfel de curbe sunt numite curbe de regresie, iar ecuațiile corespunzătoare referitoare la sistemul de axe Ox și Oy sunt numite ecuații de regresie. Dacă curbele de regresie sunt linii drepte regresia e numită liniară. În celelalte cazuri e numită curbilinie.

Dacă 2 caracteristici cantitative variabile sunt independente, liniile lor de regresie sunt drepte perpendiculare. Desigur că în cazurile luate cu date reale, mediile nu se vor înscrie exact pe linii drepte, din cauza fluctuațiilor de selecție.

Cazurile cu care are de-a face experimentatorul de exemplu chimistul sau fizicianul, cazuri în care observațiile sunt concentrate strâns în jurul unei singure linii, sunt cazuri opuse cazului independenței. Înscrierile în aceste tabele cad în numai câteva căsuțe pentru fiecare aranjare și mediile rândurilor și ale coloanelor cad aproximativ pe una și aceeași curbă.

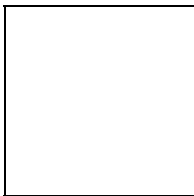
Cazurile întâlnite în mod obișnuit în statistică sunt între aceste 2 cazuri extreme, liniile mediilor nefiind nici perpendiculare și nici confundate.

O primă problemă a statisticianului este de a găsi expresiile cu care să se poată descrie liniile de regresie fie exact fie cu un grad de aproximare suficient.

În cazul în care mediile aranjărilor nu sunt situate pe o linie dreaptă căutăm acele linii drepte care sunt legate cel mai strâns de medii, evaluând coeficienții a_1 și b_1 din ecuația :

$$x = a_1 + b_1 y$$

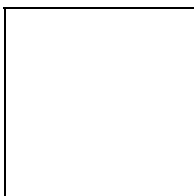
în care x și y reprezintă abaterile lui X și Y de la valorile medii $x = X - M(X)$ $y = Y - M(Y)$ astfel încât să minimizeze expresia :



așadar se poate aplica modelul regresiei liniare multiple luând

$$Z_1 = 1 \text{ și } Z_2 = Y$$

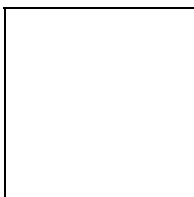
Rezolvarea sistemului ținând cont de semnificația lui x și y duce la :



Considerând în același mod pentru ecuația :

$$y = a_1 + b_1 x$$

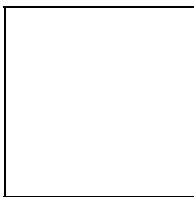
în care x și y reprezintă abaterile lui X și Y de la valorile medii astfel încât să minimizeze expresia :



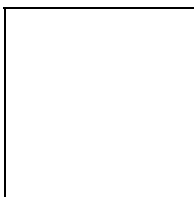
așadar se poate aplica modelul regresiei liniare multiple luând

$$Z_1 = 1 \text{ și } Z_2 = X$$

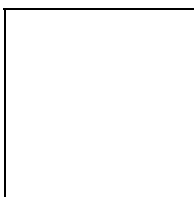
Rezolvarea sistemului ținând cont de semnificația lui x și y
duce la :



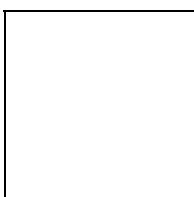
Facem notația :



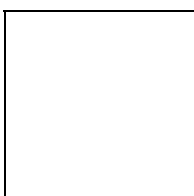
Știind că :



atunci relația dintre x și y se scrie :



Exprimând în raport cu variabilele măsurate avem :



Aceste ultime 4 ecuații au cea mai largă aplicare. Dacă regresia este exact liniară aceste ecuații dau liniile de regresie. Dacă regresia se îndepărtează de la liniaritate, fie datorită datelor

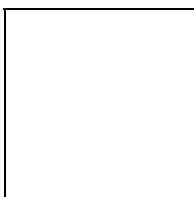
de selecție, fie datorită unor cauze reale, acestea dau " cea mai bună regresie liniară " admisă de datele considerate. Putem considera aceste ecuații ca :

(a) ecuații pentru a estima o valoare individuală a lui x cu ajutorul valorii asociate y (sau y cu ajutorul valorii asociate x), astfel încât suma pătratelor erorilor de estimare să fie minimă.

(b) ecuații care estimează media valorilor lui x asociate cu un anumit y (sau media valorilor lui y asociate cu un anumit x), astfel încât suma pătratelor de estimare să fie minimă, fiecare medie fiind calculată proporțional cu numărul de observații pe care se bazează.

1.7. Coeficientul de corelație. Covarianța

Coeficientul r definit anterior :



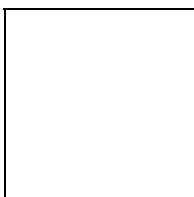
unde $x = X - M(X)$ și $y = Y - M(Y)$ se numește coeficient de corelație.

Este necesar să facem câteva observații în legătură cu acest coeficient :

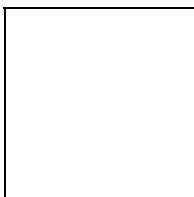
Teorema 1

Coeficientul r definit de relația anterioară nu poate să depășească $+1$ și nu poate fi mai mic decât -1 .

Demonstrație



care scrisă explicitând sumele:



este chiar inegalitatea Cauchy-Bunakovsky-Schwartz deci este evidentă.

Observații

Dacă $r=+1$ ecuațiile de regresie sunt identice și toate valorile se înscriu pe o singură linie dreaptă.

Valorile mari ale lui x corespund valorilor mari ale lui y . Variabilele se numesc corelate perfect pozitiv.

Dacă $r=-1$ ecuațiile de regresie sunt identice și toate valorile se înscriu pe o singură linie dreaptă. Valorile mari ale lui x corespund valorilor mici ale lui y . Variabilele se numesc corelate perfect negativ.

În sfârșit, dacă variabilele sunt independente, $r=0$, deoarece b_1 și b_2 sunt 0 și liniile de regresie sunt paralele cu Ox și Oy .

Faptul că $r=0$ nu implică că variabilele sunt independente ci doar că mediile aranjamentelor sunt cuprinse între 2 segmente care nu depășesc anumite limite pe care le-am definit înainte pe pe orizontale și pe verticale.

Două variabile pentru care $r=0$ sunt numite necorelate.

Dacă x și y sunt legate funcțional aceasta nu înseamnă că coeficientul lor de corelație este 1, în afara cazului în care această corelație este liniară.

Cantitatea :



o vom numi momentul mixt de ordinul I sau covarianța repartiției.

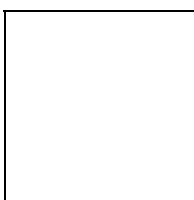
Teorema 2

Fie X_1 și X_2 două variabile și $Z=X_1+X_2$. În acest caz evident $z=x_1+x_2$

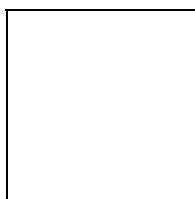


Atunci dacă x_1 și x_2 sunt necorelate :

Demonstrație.



Deoarece x_1 și x_2 sunt necorelate înseamnă că $r=0$ deci :



Rezultatul stabilit de teorema 2 poate fi aplicat în teoria erorilor de observație.

Să presupunem că orice valoare a lui X fiind observată de un număr mare de ori, media aritmetică a observațiilor este aproximativ valoarea reală, eroarea medie aritmetică fiind 0.

Atunci, deoarece eroarea medie aritmetică este egală cu 0 pentru toate valorile lui X , eroarea notată cu δ este necorelată cu X . În acest caz, dacă x_1 este o abatere observată de la media aritmetică și x este abaterea reală din cele de mai sus obținem :



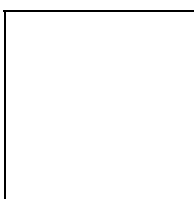
Efectul erorilor de observație se manifestă, prin urmare prin creșterea abaterii standard peste valoarea reală. Acest lucru nu implică independența completă a lui X și δ .

Putem presupune că fluctuația erorilor este mai mare, de exemplu, cu cât cresc valorile lui X , așa cum este de altfel foarte probabil să se întâmple. În acest caz coeficientul de contingență dintre X și δ nu este 0, cu toate că coeficientul de corelație se poate micșora până la anulare.

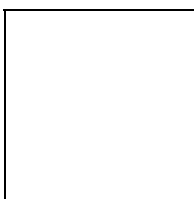
1.8. Raportul de corelație

Ținând seama de importanța liniarității regresiei este de dorit ca să existe anumite criterii care să permită aprecierea formei acestei regresii, între limitele permise de fluctuațiile de selecție și a condițiilor în care aceasta este liniară. Studiem un coeficient introdus în acest scop.

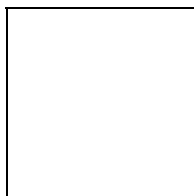
Să considerăm un tabel de frecvențe după 2 caracteristici și fie s_{px} abaterea standard a celei de-a p aranjări a X-ilor. Fie n_p numărul observațiilor în această aranjare. Fie



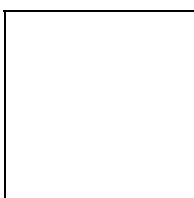
Atunci s_{px}^2 este media ponderată a dispersiilor aranjărilor obținute. Fie acum :



sau



Mărimea r_{xy} este numită raportul de corelație a lui X față de Y. Prin analogie raportul r_{yx} :



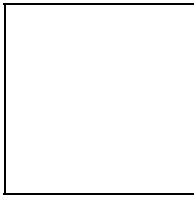
este denumit raportul de corelație al lui Y față de X.

Calculând acum:

$$\sum (x - b_{1y})^2 = \sum \left(x - \frac{1}{n} \frac{\sum xy}{\sigma_y^2} y \right)^2 = \sum x^2 - \frac{(\sum xy)^2}{\sum y^2} = (1 - r^2) \sum x^2$$

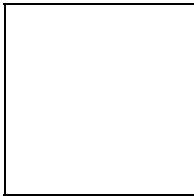
unde r coeficientul de corelație definit la [1.7.]

Împărțind cu n și ținând seama de semnificația mărimilor ce intervin avem :

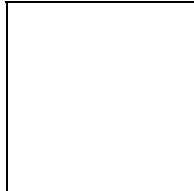


unde $x - b_1 y = 0$ este linia de regresie a lui x față de y iar x și y sunt abaterile lui X și Y de la media repartiției.

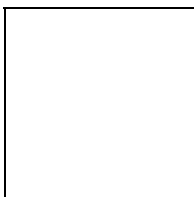
Pentru orice aranjare pentru care y este constant :



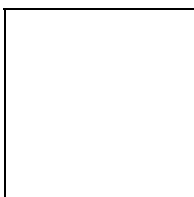
suma tuturor produselor anulându-se deoarece :



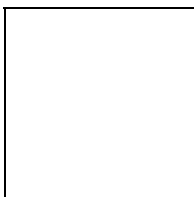
De aici însumând pentru toate aranjările lui y , obținem :



dar ținând seama că :



obținem :

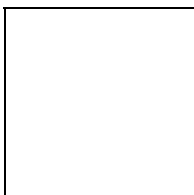


Din această relație se observă că η_{xy} nu poate fi mai mic decât r în valoare absolută.

Dacă $\eta_{xy}^2=r^2$ atunci :



adică :



pentru toate aranjările. Aceasta arată că media m_{px} trebuie să fie pe linia de regresie pentru toate aranjările, adică faptul că regresia este liniară.

Rezultă că diferența dintre η^2 și r^2 măsoară abaterea regresiei de la liniaritate. De aceea va trebui să observăm că fluctuațiile de selecție pot provoca variația cantității $\eta^2 - r^2$ chiar și atunci când regresia este riguros liniară.

1.9. Corelația rangurilor

În determinarea corelației dintre 2 variabile pot apărea obstacole mai serioase decât grupările imperfecte.

Să considerăm de exemplu relația dintre aptitudinile matematice și cele muzicale într-o clasă de elevi. Aptitudinea, într-o formă generală sau particulară este o variabilă, în sensul că variază de la o persoană la alta și poate fi o caracteristică numerică numai dacă putem decide într-un mod ce exclude ambiguitățile asupra măsurii sale. Un mod foarte comod este de a supune la diferite probe fiecare elev. Dar și acest gen de metode sunt susceptibile de multe obiecții, printre care și aceea că diverși examinatori vor aprecia în mod diferit aceeași persoană. De aceea corelația dintre notele obținute pentru matematici și cele obținute la

muzică va depinde probabil de anumite preferințe ale examinatorului și nu va reflecta deci în mod corect relația dintre cele 2 caracteristici calitative.

Dificultățile de acest fel dispar întrucâtva dacă aranjăm pe cei examinați în ordinea aptitudinilor lor și nu încercăm să evaluăm numeric acest parametru. Putem să atribuim atunci fiecărui ins examinat un număr care indică poziția sa în serie potrivit aptitudinii.

Cei examinați sunt astfel ordonați iar numărul unui anumit ins în șirul acestui aranjament se numește rang.

Un astfel de procedeu se folosește nu numai în studiul datelor care pot fi ordonate și nu și măsurate exact, ci și în studiul acelor care pot fi măsurate. De exemplu, putem ordona cu ușurință un număr de oameni fără să îi măsurăm.

De asemenea este relativ mai ușor să aranjăm un număr de nuanțe de culori, sau un număr de țări după importanța produselor de export acolo unde măsurătorile numerice ar fi foarte dificile.

În cazul extrem, putem avea situații în care diferite unități studiate pot fi aranjate într-o anumită ordine, dar în nici un caz teoretic ele nu pot fi măsurate.

Dacă avem o mulțime de elemente ordonate după 2 caracteristici calitative diferite este natural să ne întrebăm dacă rangurile lor ne pot oferi o măsură a gradului legăturii ce există între aceste 2 caracteristici.

Să presupunem că avem n elemente ale căror ranguri în raport cu caracteristica calitativă A sunt X_1, X_2, \dots, X_n și în raport cu însușirea B sunt Y_1, Y_2, \dots, Y_n unde cantitățile X și Y sunt simple permutări ale primelor n numere naturale.

Fie $d_k = X_k - Y_k$. Valorile lui d formează o măsură convenabilă a intensității de apropiere a corespondenței dintre A și B . Dacă toți d sunt 0 corespondența este perfectă, deoarece un element al cărui rang este X_k , după caracteristica calitativă A , va avea același rang și după caracteristica B . Nu putem totuși să luăm suma mărimilor d drept măsură a corespondenței, deoarece această sumă va fi întotdeauna egală cu 0 fiind aceleași numere naturale dar în altă ordine.

O măsură posibilă al cărei interes rezultă din însăși definiția ei este suma valorilor absolute a cantităților d , adică $\sum |d|$. Această măsură și media ei au fost de fapt folosite dar, ca și abaterea medie absolută ele ridică dificultăți de ordin analitic.

Un coeficient mai convenabil calculelor se obține în modul următor :

- valorile rangurilor X variază între 1 și n .

- suma lor este deci și media lor este prin urmare

- această valoare este de asemenea media Y -lor.

- notăm cu x_k valoarea X_k - adică abaterea lui X_k de la medie.

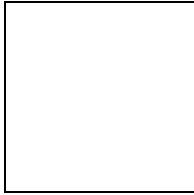
- în mod analog definim pe y_k prin Y_k -

- luăm

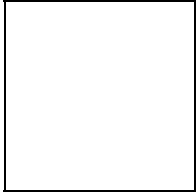
Acesta este coeficientul de corelație dintre X și Y după formula generală. Vom denumi cantitatea ρ , coeficientul de corelație a rangurilor al lui Spearman.

Dacă ținem seama de formula dată pentru d putem exprima ușor în funcție de n și d :

iar



De aici : și înlocuind în expresia lui ρ obținem :



Exemplu. Ordonarea a 10 elevi la matematică și muzică se prezintă după cum urmează :

- matematică : 1,2,3,4,5,6,7,8,9,10.

- muzică : 6,5,1,4,2,7,8,10,3,9.

Care este coeficientul de corelație al rangurilor?

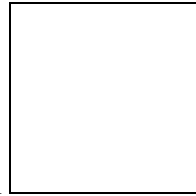
p1. Diferențele d sunt (rang matematică - rang muzică) :

-5 , -3 , +2 , 0 , +3 , -1 , -1 , -2 , +6 , +1.

Evident suma acestora este 0 cum ne așteptam.

p2. Pătratele mărimilor d sunt : 25 , 9 , 4 , 0 , 9 , 1 , 1 , 4 , 36 , 1 care adunate dau 90.

p3. Din relația ce dă pe ρ ca funcție de d și n obținem



Teoremă

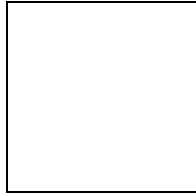
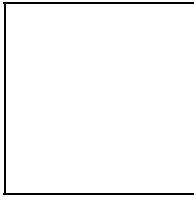
Coeficientul de corelație a rangurilor variază între +1 și -1. Dacă, pe de altă parte, rangurile sunt aranjate în ordine inversă $\rho = -1$.

Demonstrație

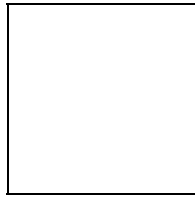
Faptul că ρ este între -1 și +1 este evident din faptul că el este un coeficient de corelație a seriilor de ranguri și pentru coeficientul de corelație în general am demonstrat acest lucru. Pentru cazul minimului -1 demonstrația este puțin diferită, după cum n este par sau impar.

- caz1 : $n=2m+1$:

valorile d sunt : $2m, 2m-2, \dots, 0, -2, \dots, -(2m-2), -2m$ și :



prin urmare :

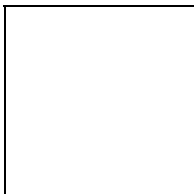


- caz2 : $n=2m$:

și $\rho = -1$ ca mai sus.

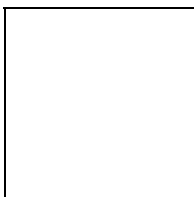
Observații

Proprietatea de a fi cuprins între +1 și -1 nu aparține și unui alt coeficient similar propus tot de Spearman, cunoscut sub denumirea de „foot – nule” adică regula grosolană și anume :

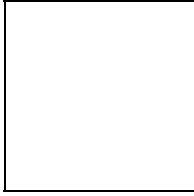


Poate fi arătat că \mathfrak{R} variază între -0,5 și +1 și, tocmai din acest motiv, \mathfrak{R} pare a fi un coeficient mai puțin recomandabil.

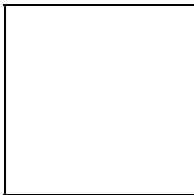
Există și un al doilea coeficient de corelație a rangurilor definit în modul următor :



Coeficientul τ poate fi considerat de asemenea, ca o corelație de tip general. Notând pentru două ranguri X_i și X_j cu a_{ij} mărimea :



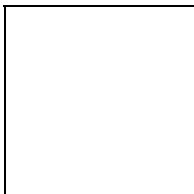
Analog notăm cu b_{ij} cantitatea corespunzătoare din a 2-a ordonare. Avem deci :



1.A. Autocorelația

Raportul dintre covarianța a 2 termeni alăturați și rădăcina pătrată a produsului dispersiilor celor 2 termeni se numește autocorelația de rangul 1.

În mod similar, vom defini autocorelația de rangul k , autocorelația între 2 termeni separați prin $k-1$ termeni :

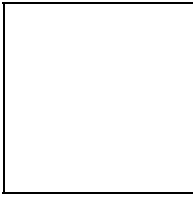


Aceste funcții sunt extrem de importante în teoria seriilor dinamice oscilante și domeniul lor de aplicare este mult mai larg decât acela la care ne referim aici.

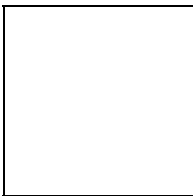
Dacă este necesar să distingem valorile populației originare și de selecție, vom denumi pe acestea din urmă corelații de serie și le vom nota cu r_k .

Deosebirea dintre autocorelație și corelația de serie ca și dintre notația ρ și r corespunde practicii statisticii de a folosi litere grecești pentru desemnarea valorilor populației originare și latine - pentru valorile de selecție.

Într-o serie de mare întindere, $D^2(u_t)$ și $D^2(u_{t+k})$ sunt practic identice și expresia autocorelației devine :



Pentru seriile empirice scurte este recomandabil să se ia în considerare dispersia întregii serii (calculată din n termeni) drept estimare a dispersiei u , deși covarianța se referă numai la $n-k$ termeni. De asemenea este bine să se calculeze abaterile valorilor u de la media întregii serii, atunci când se determină sumele produselor u_t și u_{t+k} . Dacă termenii seriei au drept origine media tuturor termenilor, ajungem la formula :

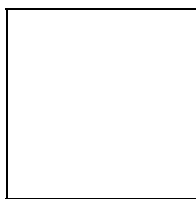


Testele privind caracterul întâmplător al unei serii dinamice nu sunt totdeauna necesare. Deseori se poate identifica de la prima vedere existența unei variații sistematice. Calculul celor 2 coeficienți pot fi considerați ca 2 teste și se aplică în cazurile incerte și, în general, ele sunt suficiente pentru rezolvarea problemei. În continuare vom trata elaborarea unor modele care să reproducă comportamentul unor serii oscilatorii.

1.B. Analiza prin periodograme

Făcând o analogie simplă din fizica clasică, multe fenomene fizice oscilatorii (mareele, pendulul) pot fi reprezentate printr-o sumă de un număr anumit de armonici pure, fiecare

dintre ele fiind considerată ca o funcție sinusoidală sau cosinusoidală. Mișcarea pur oscilatorie este descrisă de o funcție de tipul:



, unde λ este lungimea de undă iar A amplitudinea.

Deseori un fenomen oscilator poate fi reprezentat prin suma unor asemenea termeni :



Lumina este și ea un fenomen de acest fel. Experiența clasică a lui Newton de descompunere a luminii albe prin mijlocirea unei prisme poate fi considerată o analiză a unui fenomen periodic complicat, o separare a acestuia în elemente simple, fiecare având propria culoare sau lungime de undă.

Exemplul modelării din fizică a procesului oscilator prin o serie armonică nu este totuși pe deplin suportat și de o serie dinamică economică. Modelul presupune că erorile vor afecta observația la un moment dat, dar nu influențează asupra mișcării ulterioare a sistemului. În economie, ca și în meteorologie, când intervine o perturbație, este mai plauzibil să presupui că efectele acesteia se integrează în mișcarea ulterioară a sistemului și devin o parte a acestuia.

Identificarea periodicității într-o serie dinamică cu ajutorul analizei armonice poate fi comparată cu reglarea unui aparat de radio pe un anumit post. Corelând o serie a cărei lungime de undă este cunoscută cu seria dată dacă cele 2 lungimi de undă se acordează, corelația va fi intensă. Comparând, în teste succesive, diferitele intensități de corelație, putem ajunge la cunoașterea lungimii de undă a seriei date.

Să exemplificăm:

- alegem lungimea de undă μ și calculăm sumele :

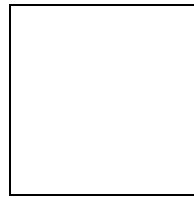


, S se numește intensitate. A și B sunt covarianțele seriei de care ne servim pentru primul test și ai cărei termeni sunt exprimați în valorile \sin și \cos . Să presupunem că

seria e dată de în care b este un termen necorelat cu perioada seriei pe care am ales-o pentru primul test. Atunci :

unde :

Expresia pentru B este similară, cu deosebirea că se înlocuiește cu respectivul cosinus. Pentru valori mari ale lui n , expresia are valori mici, dacă termenii în interiorul nu sunt valori mari, deci dacă $\alpha - \beta$ sau $\alpha + \beta$ nu sunt mici. În acest caz, neglijând termenul de ordinul $1/n$ găsim :



Întrucât pentru valori mici θ , $\sin\theta \approx \theta$, rezultă că : și deci putem scrie : $A^2+B^2=a^2$.

Astfel, S rămâne mic, dacă α este aproape egal cu β și deci perioada care a servit drept test este apropiată de perioada reală λ și S este egal cu constanta și dă amplitudinea termenului.

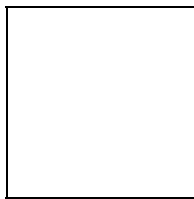
Reprezentarea grafică, în care valorile S sunt pe ordonată și valorile μ pe abscisă este o periodogramă iar analiza ca atare se numește analiza prin periodogramă.

Un rezultat important este că probabilitatea ca S^2 să depășească mărimea $4vk/n$ unde v dispersia seriei este e^{-k} . Rezultă de aici că probabilitatea ca cel puțin o perioadă dintr-un număr q de perioade alese pentru testare să întrecă $4vk/n$ este : $1-(1-e^{-k})^q$.

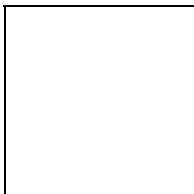
1.C. Corelograma

Graficul autocorelației, în care ρ_k este ordonată și k abscisa se numește corelogramă. Întrucât $\rho_k = \rho_{-k}$ graficul se alcătuiește numai pentru valorile nenegative ale lui k .

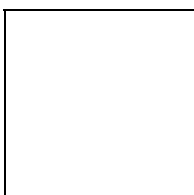
Întorcându-ne la exprimarea lui ρ_k în forma :



se poate demonstra că pentru seria :

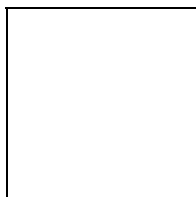


corelograma este dată de :



cu condiția ca ε să fie independent de termenii armonicii.

Astfel, oricărui termen cu amplitudinea A_j în seria inițială îi corespunde o undă cu amplitudinea :



în corelograma care nu are vârfurile teșite.

Teoretic, deci corelograma ar trebui să ne furnizeze o metodă care să ne permită să distingem oscilațiile armonice suprapuse de corelațiile autoregresiei. Într-un caz, oscilațiile din corelogramă nu se netezesc, pe când în celălalt caz ele suferă asemenea modificări.

În practică însă, pentru serii scurte, nu ne putem bizui prea mult pe corelograme, deoarece, pentru rațiuni ce țin de selecție, corelogramele autoregresive nu se netezesc. Totuși, examinarea corelogramei constituie deseori un bun început pentru o cercetare privind modelul care generează un anumit sistem.

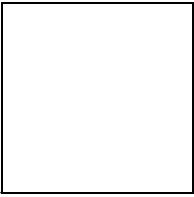
Capitolul 2. Analiza numerică a seriilor de timp.

2.1. Descompunerea seriilor de timp.

Schimbarea fenomenelor în timp este determinată de acțiunea unui complex de factori esențiali și neesențiali care exercită influențe simultane sau succesive cu intensități diferite, în același sens sau în sens opus. Factorii esențiali pot acționa în permanență asupra fenomenului, își pot modifica intensitatea acțiunii pe măsura modificării condițiilor în timp și imprimă fenomenului o anumită tendință generală de dezvoltare (trendul). Există și factori care acționează regulat, dar periodic, imprimând tendințe ciclice de dezvoltare (sezonalitatea). În sfârșit există și factori întâmplători care influențează fenomenele. Dacă $X(t)$ este valoarea variabilei X la momentul t atunci $X(t)=T(t)+S(t)+e(t)$ unde $T(t)$ reprezintă trendul, $S(t)$ sezonalitatea, iar $e(t)$ reziduul aleator la momentul t .

2.2. Tendința de evoluție. Trend.

Vom folosi criteriul analitic de determinare a trendului deci vom determina funcția f prin coeficienții săi astfel încât :

suma  să fie minimă.

2.2.1. Regresie liniară.

Se consideră funcția $f(x) = ax+b$. Condiția ca funcția f să fie cea optimă este ca derivatele parțiale ale lui S în raport cu coeficienții a și b să se anuleze. Aceste două condiții generează un sistem de două ecuații cu două necunoscute: a și b .

Deoarece regresia liniară este un caz particular de regresie polinomială și regresie Taylor o vom trata pe aceasta din urmă mai în detaliu.

De menționat că există o bogată literatură de specialitate care tratează în detaliu regresia liniară. Dintre lucrările care tratează în detaliu regresia liniară menționăm [7],[8],[9],[11],[15],[18].

2.2.2. Regresie prin serie parțială Taylor.

Cazul regresiei prin serie parțială Taylor se reduce la cazul regresiei liniare multiple dacă considerăm că variabilele Z_j sunt funcții ce depind de variabila $X=X(t)$. Astfel variabila aleatoare Z_1 o înlocuim cu variabila aleatoare $X_1=X$, variabila aleatoare Z_2 o înlocuim cu variabila aleatoare $X_2=X*X$ ș.a.m.d..

Regresia prin serie parțială Taylor se realizează presupunând că distribuția urmează o funcție de timp de tipul :

$$f(x)=a_0+a_1x+a_2x^2+\dots+a_{n-1}x^{n-1}+a_nx^n.$$

Minimizarea sumei pătratelor erorilor are loc așadar în cazul când f are această formă și minimizarea are loc prin calculul coeficienților.

Suma pătratelor erorilor este :

și ea este minimă când :

Aceste ecuații formează un sistem de $m+n+1$ ecuații cu $m+n+1$ necunoscute :

Dacă introducem Media $M(W)$ a unei variabile aleatoare W definită la regresia liniară multiplă avem :

Ca observație se poate remarca faptul că determinantul sistemului este simetric față de diagonala principală și această observație duce la reducerea numărului de calcule.

Acest sistem de ecuații se poate rezolva prin metoda reducerii Gauss și din rezolvarea sa rezultă coeficienții a_0, \dots, a_n .

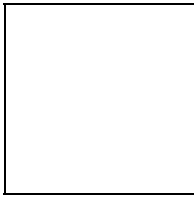
Metoda pe care am ales-o în acest caz a fost formarea de elemente nule sub diagonala principală a matricei extinse a sistemului ce conține ca și coloană suplimentară coloana termenilor liberi.

2.2.3. Regresie parabolică. Regresie hiperbolică.

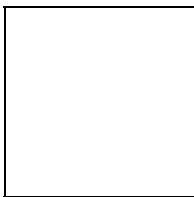
Cazul regresiei parabolice se reduce la cazul regresiei liniare multiple dacă considerăm că variabilele Z_j nu sunt independente ci sunt funcții ce depind de variabila x .

Astfel variabila aleatoare Z_1 o înlocuim cu variabila aleatoare $X_1=X$, variabila aleatoare Z_2 o înlocuim cu variabila aleatoare $X_2=X*X$ și la cazul regresiei Taylor cu indicii m și n aleși convenabil.

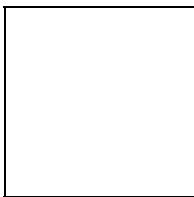
În cazul regresiei hiperbolice distingem 3 metode de soluționare :



În acest caz variabila Y se înlocuiește cu variabila $W=1/Y$ pentru a se determina a și b prin regresie liniară.



În acest caz variabilele X și Y își schimbă între ele poziția pentru a se determina a și ab prin regresie liniară.

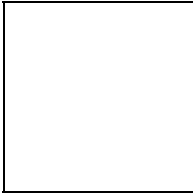


În acest caz $Z_1=X$, $Z_2=Y$, $Z_3=1$, $Y=-XY$ și aplicând regresia liniară multiplă obținem $b, a, ab-c$ de unde se obțin coeficienții a, b și c .

2.2.4. Cazul special al trendului polinomial.

Trendul polinomial se poate privi ca un caz particular de trend Taylor. Ținând seama că în acest caz este o problemă de decizie gradul polinomului va trebui să construim o metodă care să ne furnizeze acest grad în caz general.

Pentru a intui trendul va trebui să construim diagrama seriei dinamice. Dacă considerăm o familie de funcții polinomiale $\{f_i\}$, $i=1, \dots, k$ care aproximează trendul T atunci criteriul analitic de alegere a trendului este alegerea celei mai bune dintre aproximările $T(t)=f_i(t)$, $i=1, \dots, k$ prin calcularea sumelor :



și alegerea acelei aproximări pentru care suma S_i este minimă. Totuși acest model nu este satisfăcător deoarece odată cu creșterea gradului scade și suma S .

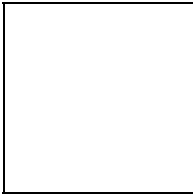
Dacă calculăm coeficientul de variație al lui Pearson atunci cea mai adecvată aproximare va fi aceea pentru care coeficientul de variație este minim.

Însă nici această metodă nu ne poate ajuta în acest caz din același motiv. Se formulează așadar problema gradului optim al polinomului. Această problemă va fi expusă pe larg paragraful următor.

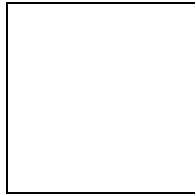
2.2.5. Determinarea ordinului tendinței polinomiale

Pentru determinarea ordinului tendinței polinomiale este necesar să se stabilească indicele n_0 al ultimului coeficient $b_{n_0} \neq 0$. În general se urmărește obținerea unui model de ordin cât mai mic.

Dacă însă ordinul modelului este prea mic, estimația tendinței este deviată în raport cu tendința reală, iar dacă ordinul tendinței este prea mare apar fluctuații ale estimației

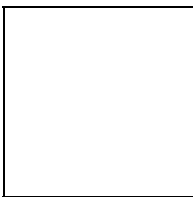


în jurul tendinței



Pentru determinarea ordinului modelului se folosesc proceduri de tip pas cu pas. În practică s-au răspândit 2 categorii de metode, bazate respectiv pe studiul estimației dispersiei reziduurilor și pe verificarea succesivă a ipotezei nule.

Procedura analizei estimației dispersiei reziduale se bazează pe faptul că estimația dispersiei reziduurilor :

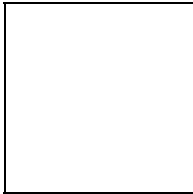


va avea valori mai mari când numărul de termeni este mai mic decât cel optim. Astfel funcția $\lambda_2(n)$ va prezenta un minim local sau va deveni plată în zona numărului corect de termeni n_0 ; procedura constă în calculul succesiv al estimațiilor b_0, b_1, \dots până când se obține minimum funcției $\lambda_2(n)$ adică $\lambda_2(n_0)$.

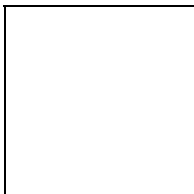
Procedura descrisă prezintă avantajul simplității, având însă dezavantajul slabei precizii. Dezavantajul se datorează faptului că minimum funcției $\lambda_2(n)$ nu este suficient de pronunțat. Așa că în zona de extrem curba este plată, sensibilitatea în raport cu n fiind scăzută. Aceasta se manifestă în faptul că, pentru $n_0 - 1, n_0, n_0 + 1, \dots$ se obțin valori apropiate ale estimației dispersiei reziduurilor $\lambda_2(n)$.

De obicei prin această procedură se alege valoarea maximă a ordinului din zona staționară $n = n_0 + q = n_q$ și se utilizează în continuare o procedură mai sensibilă cum ar fi procedura bazată pe utilizarea ipotezei nule.

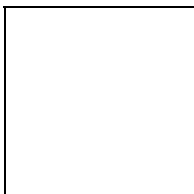
Procedura verificării ipotezei nule presupune cunoașterea faptului că $n_0 \leq n_q$. În aceste condiții procedura constă în verificarea ipotezei

H : 


în raport cu alternativa



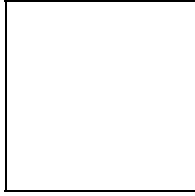
cu un grad de încredere ε utilizând intervalul de încredere pentru parametrul ai dat de relația :



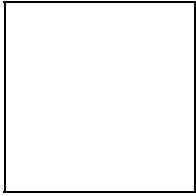
în care :



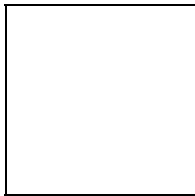
Ipoteza nulă



este respinsă cu un grad de încredere ε dacă ai se află în afara intervalului de încredere adică este satisfăcută condiția :



unde



reprezintă un punct de pe distribuția Student și se determină din tabele în funcție de ε și $v=N-nq-1$.

Dacă ipoteza nulă a fost admisă pentru nq , se repetă verificarea ipotezei nule pentru $nq-1$ și așa mai departe până la $n=n_0$, când ipoteza nulă este respinsă, adică este satisfăcută inegalitatea de mai sus.

2.2.6.Regresie neliniară. Metode iterative.

Dependența ce se urmărește în acest caz este $y = f(x;bj)+err$, $j=1,...,p$ unde f neliniară în b_j .

Metoda celor mai mici pătrate poate fi aplicată pentru estimarea para-metrilor oricărui tip de funcție de regresie chiar dacă nu e liniară de parametrii cu condiția ca erorile să nu fie aditive.

În cazul erorilor aditive relația de dependență stocastică nu mai poate fi făcută prin reducerea la modelul de regresie liniară multiplă.

Sistemul de ecuații ce rezultă din aplicarea metodei celor mai mici pătrate nu poate fi rezolvat direct și în acest caz se minimizează sumele pătratelor erorilor prin metode iterative.

Cu alte cuvinte, pentru a găsi minimumul unei funcții neliniare de mai multe variabile se aplică procedee iterative cu mai multe etape, plecând de la estimări inițiale cunoscute dintr-o experiență anterioară. Fie funcția de regresie depinzând de p parametri, și anume funcția neliniară de parametri b_0, \dots, b_p : $f(x; b_j) = M(Y|X) \quad j=0, \dots, p$.

Prezentăm obținerea estimărilor b_j folosind un algoritm în două etape.

Etapa1. Se pleacă de la valorile inițiale ale parametrilor b_0^*, \dots, b_p^* cărora le corespunde valoarea inițială a funcției de regresie $f^* = f(x; b_j^*)$, $j=0, \dots, p$. Se introduc variabilele independente x_1, \dots, x_p , care sunt derivatele funcției de regresie în raport cu fiecare dintre parametri b_j :

Suma pătratelor erorilor ce urmează să fie minimizată este :

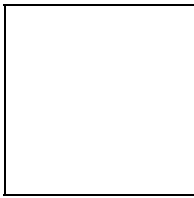
și valoarea ei minimă se obține aplicând metoda celor mai mici pătrate

$$B^{\wedge} = (XX')^{-1}XY.$$

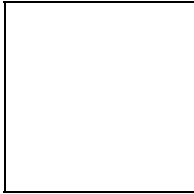
Minimumul sumei pătratelor erorilor poate fi scris :

unde :

Estimațiile parametrilor prin metoda celor mai mici pătrate duc la rezolvarea sistemului de ecuații normale :



a căror soluție este :



unde X este matricea formată cu elementele (x_{ij}) și R vectorul de componente $R_i, i=1, \dots, n$.

Etapa 2. Rezultă imediat că estimațiile parametrilor b_j sunt $b_j^{\wedge} = b_j^* + d_j, j=0,1,\dots,p$ sau $b^{\wedge} = b^* + d$.

Pentru discuția completă a modelelor de regresie neliniară propriu-zise se pot consulta lucrările [5],[9],[35].

2.2.7. Regresie exponențială și regresie logaritmică.

Un caz concret de regresie neliniară apare în cazul regresiei exponențiale de forma :

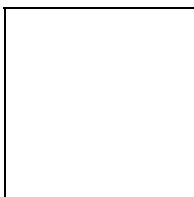
$$y = b_0 + b_1 e^{-b_2 x}$$

unde

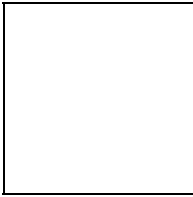
$$b_0, b_1, b_2$$

sunt parametrii necunoscuți ce urmează a fi estimați pe baza datelor specifice ale problemei de rezolvat.

Din acest motiv vom expune acest tip de regresie pe un caz particular : activitatea virusului Papova :



Pornind cu valorile inițiale ale parametrilor b_j^* , $j=0,1,2$ ca componente ale vectorului parametrilor :



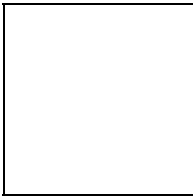
Se aplică procedeul în două etape expus anterior, calculând la sfârșit suma pătratelor erorilor pentru estimațiile parametrilor : $b_0^{(1)}$, $b_1^{(1)}$, $b_2^{(1)}$

Procedeul se repetă până când se observă o stabilizare a valorii sumei pătratelor erorilor. În cazul exemplului dat, după a treia iterație se obține o stabilizare a sumei pătratelor erorilor :

$$S(b_0^{(3)}, b_1^{(3)}, b_2^{(3)})=0.0024$$

care corespunde estimațiilor : $b_0^{(3)}$, $b_1^{(3)}$, $b_2^{(3)}$

componente ale vectorului estimațiilor :



În cazul regresiei logaritmice premisele de la care se pleacă sunt aceleași ca și în cazul regresiei exponențiale. Un studiu mai aprofundat al regresiei exponențiale în cazul expus se găsește în [12].

2.3. Sezonalitate. Periodicitate.

2.3.1. Seria de sezonalitate.

În general se admite că parametrul timp este discret. Deși această ipoteză simplificatoare satisface exigențele în majoritatea aplicațiilor, există situații în care timpul t parcurge o mulțime nenumărabilă. Distingem așadar serii dinamice cu parametru discret și serii dinamice cu parametru continuu. Pe de altă parte studiind o serie dinamică este posibil să obținem aceleași valori caracteristice (medie ,dispersie) sau valori foarte apropiate, indiferent de intervalul considerat. Aceasta revine la a spune că vectorii aleatori

$$X(t_1) , \dots , X(t_n)$$

nu își schimbă caracteristicile stocastice după un interval arbitrar de timp t sau, ceea ce în anumite condiții este echivalent, că vectorii :

$$X(t_1+t), \dots, X(t_n+t)$$

au aceeași repartiție cu vectorii :

$$X(t_1), \dots, X(t_n)$$

O astfel de serie dinamică se numește staționară în sens strict. Atunci când corespunde suficient de bine condițiilor din sistemul real ipoteza de staționaritate duce la simplificări considerabile în modelare.

Un exemplu tipic de serie dinamică staționară îl constituie fluctuațiile tensiunii într-o rețea electrică.

Desigur, existența unei tendințe sistematice exclude staționaritatea. Spunem atunci că avem de-a face cu o serie dinamică nestaționară. Adeseori sistemele evolutive complexe cum sunt cele economice sau biologice, manifestă astfel de tendințe și ca atare, nu li se poate aplica ipoteze staționare. În plus pot exista componente ciclice, componente sezoniere $S(t)$ și abateri aleatoare $e(t)$. Astfel, notând cu $T(t)$ tendința sistemul evolutiv poate fi investigat făcând apel la un model aditiv de forma :

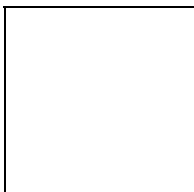
$$X(t) = T(t) + S(t) + e(t)$$

Totuși să observăm că acest model implică respectarea unor condiții de liniaritate și metodele de determinare a componentelor un sunt satisfăcătoare în toate cazurile. Teoria generală și aplicații consistente conțin lucrările [19] și [20].

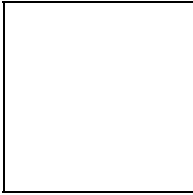
2.3.2. Construcția seriei de sezonalitate prin serii armonice.

Odată găsit trendul $T(t)$ vom căuta sezonalitate observând că $X(t) - T(t) = S(t) + e(t)$ deci sezonalitatea se va găsi eliminând trendul din seria dinamică.

Sezonalitatea $S(t)$ se poate aproxima prin suma unui număr de armonici :



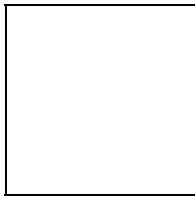
unde :



În cazul perioadei cunoscute T putem obține direct valorile coeficienților a_k și b_k cu ajutorul formulelor de mai sus. Apoi înlocuind în expresia lui $S(t)$ aflăm dependența sezonieră $S(t)$.
Un exemplu concret de aplicare îl constituie seria dinamică a livrărilor de gaz metan în perioada 1965-1971 exemplu expus pe larg în paragraful [3.4.1.]

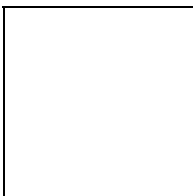
2.3.3. Folosirea FFT în construcția seriei de sezonalitate.

Fourier a dezvoltat mult teoria seriilor trigonometrice, aplicând-o la diferite probleme ale fizicii matematice. El a dat următoarea formulă :



unde funcția f e definită pe \mathbb{R} și verifică următoarele condiții :

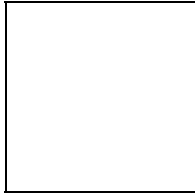
- 1.) Satisfacă în orice interval finit condițiile lui Dirichlet;
- 2.) Este absolut integrabilă într-un interval finit, adică integralele



formula lui Fourier se mai scrie câteodată și astfel :



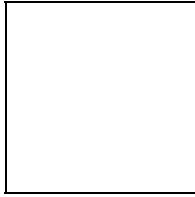
deoarece funcția $\cos(a-x)$ este pară. Deoarece $\sin(a-x)$ este impară de a , urmează că :



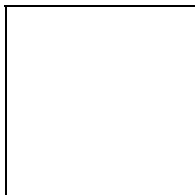
dacă ea există fiindcă toate elementele se anulează două câte două. Ținând seama de formula lui Euler :

$$e^{i(q-x)} = \cos(q-x) + i \sin(q-x)$$

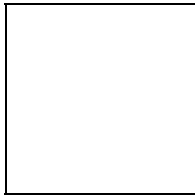
formula Fourier devine :



Dacă funcția $f(x)$ este pară : $f(x)=f(-x)$ atunci, scriind formula Fourier :

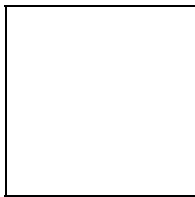


se observă că integrala care conține factorul $\sin ax$ se va anula, fiind integrala unei funcții impare, ale cărei elemente se anulează două câte două. Această formulă ia forma :



Formula poate fi aplicată la orice funcție, însă numai pentru valorile pozitive ale lui x , deoarece s-a presupus la deducerea ei că pentru valorile negative ale lui x este valabilă egalitatea $f(x)=f(-x)$.

Tot astfel dacă f impară : $f(-x)=-f(x)$ și $f(0)=0$, integrala care conține factorul $\cos ax$ se va anula :

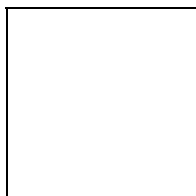


În problemele practice care ne interesează apare necesitatea de a dezvolta într-o serie trigonometrică o funcție, care este dată printr-un grafic sau pentru care se cunosc numai un șir

de valori particulare, corespunzătoare unor anumite valori ale argumentului din cuprinsul unei perioade. Aceste valori sunt de obicei echidistante.

Să presupunem că pentru funcția $y = f(x)$, a cărei perioadă este egală cu 1, se cunosc valorile Y_0, \dots, Y_{n-1} corespunzătoare valorilor $0, 1/n, 2/n, \dots, (n-1) \cdot 1/n$ ale variabilei independente x , unde este mai avantajos ca n să fie un număr par.

Am văzut că funcția y poate fi reprezentată prin seria :



sau notând $2\pi x/l = q$ atunci

$$y = A_0 + A_1 \cos q + A_2 \cos 2q + A_3 \cos 3q + \dots + B_1 \sin q + B_2 \sin 2q + B_3 \sin 3q + \dots$$

De obicei este suficient să se cunoască în această serie numai primii termeni, într-un număr nu prea mare, de exemplu până la argumentul $3q$ sau $4q$. În acest caz, fiecare din valorile cunoscute ale funcției dă posibilitatea să se scrie o ecuație corespunzătoare, iar sistemul acestor ecuații poate servi la determinarea coeficienților necunoscuți $A_0, A_1, A_2, A_3, \dots, B_1, B_2, B_3, \dots$.

Numărul acestor ecuații este de obicei cu mult mai mare decât numărul necunoscutelor și pentru rezolvarea ecuațiilor se folosește în plus metoda celor mai mici pătrate. Cu ajutorul acestei metode și pe baza sistemului de ecuații date, se formează un nou sistem de ecuații normale, la care numărul ecuațiilor este egal cu cel al necunoscutelor.

Regula formării ecuațiilor normale costă în următoarele :

- fiecare ecuație a sistemului dat se înmulțește cu coeficientul primei necunoscute și se adună apoi toate aceste ecuații, obținându-se prima ecuație a sistemului normal.
- tot așa pentru a obține a doua ecuație a sistemului normal se înmulțește fiecare din ecuațiile sistemului cu coeficientul celei de-a doua necunoscute și se adună apoi toate aceste ecuații.
- se procedează la fel pentru fiecare dintre necunoscute.

Să aplicăm această metodă în cazul nostru, mărginindu-ne pentru a simplifica scrierea la termenii până la $3q$. Sistemul ce servește la determinarea necunoscutelor $A_0, A_1, A_2, A_3, B_1, B_2, B_3$ cu notația $a = 2\pi/n$ va fi :

$$\begin{aligned} \theta = 0; & \quad Y_0 = A_0 + A_1 + A_2 + A_3 \\ \theta = \alpha; & \quad Y_1 = A_0 + A_1 \cos \alpha + A_2 \cos 2\alpha + A_3 \cos 3\alpha + B_1 \sin \alpha + B_2 \sin 2\alpha + B_3 \sin 3\alpha \\ \theta = 2\alpha; & \quad Y_2 = A_0 + A_1 \cos 2\alpha + A_2 \cos 4\alpha + A_3 \cos 6\alpha + B_1 \sin 2\alpha + B_2 \sin 4\alpha + B_3 \sin 6\alpha \\ & \dots \dots \dots \\ \theta = i\alpha; & \quad Y_i = A_0 + A_1 \cos i\alpha + A_2 \cos 2i\alpha + A_3 \cos 3i\alpha + B_1 \sin i\alpha + B_2 \sin 2i\alpha + B_3 \sin 3i\alpha \end{aligned}$$

.....

$$\theta = (n-1)\alpha; \quad Y_{n-1} = A_0 + A_1 \cos i\alpha + A_2 \cos 2i\alpha + A_3 \cos 3i\alpha + B_1 \sin i\alpha + B_2 \sin 2i\alpha + B_3 \sin 3i\alpha$$

Sistemul normal corespunzător este, prin urmare :

$$\sum_{i=0}^{n-1} Y_i = nA_0 + A_1 \sum_{i=0}^{n-1} \cos i\alpha + A_2 \sum_{i=0}^{n-1} \cos 2i\alpha + \dots + B_3 \sum_{i=0}^{n-1} \sin 3i\alpha$$

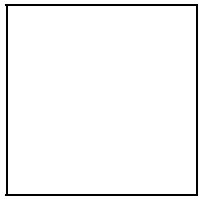
$$\sum_{i=0}^{n-1} Y_i \cos i\alpha = A_0 \sum_{i=0}^{n-1} \cos i\alpha + A_1 \sum_{i=0}^{n-1} \cos^2 i\alpha + A_2 \sum_{i=0}^{n-1} \cos i\alpha \cos 2i\alpha + \dots + B_3 \sum_{i=0}^{n-1} \cos i\alpha \sin 3i\alpha$$

$$\sum_{i=0}^{n-1} Y_i \cos 2i\alpha = A_0 \sum_{i=0}^{n-1} \cos 2i\alpha + A_1 \sum_{i=0}^{n-1} \cos 2i\alpha \cos i\alpha + A_2 \sum_{i=0}^{n-1} \cos^2 2i\alpha + \dots + B_3 \sum_{i=0}^{n-1} \cos 2i\alpha \sin 3i\alpha$$

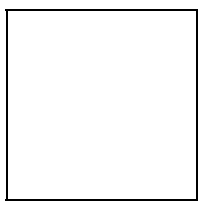
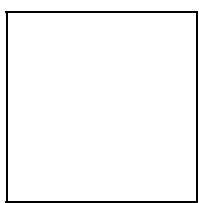
.....

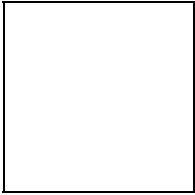
$$\sum_{i=0}^{n-1} Y_i \sin 3i\alpha = A_0 \sum_{i=0}^{n-1} \sin 3i\alpha + A_1 \sum_{i=0}^{n-1} \sin 3i\alpha \cos i\alpha + A_2 \sum_{i=0}^{n-1} \sin 3i\alpha \cos 2i\alpha + \dots + B_3 \sum_{i=0}^{n-1} \sin^2 3i\alpha$$

Se vede ușor că aceste ecuații pot fi mult simplificate, dacă folosim faptul că valorile argumentelor care figurează sub semnul sumei cresc în progresie aritmetică și parcurg o dată sau de mai multe ori un cerc întreg. Intr-adevăr, fiecare ecuație va conține numai o singură necunoscută, deoarece toate sumele cuprinse în ea, în afară de una singură, se vor anula. De exemplu pentru suma :

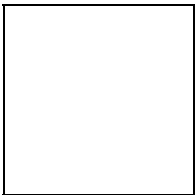
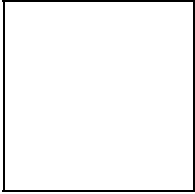


Avem următorul șir de egalități :

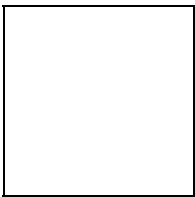




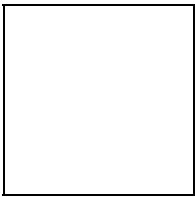
.....



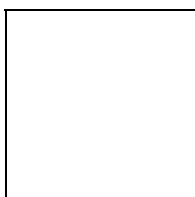
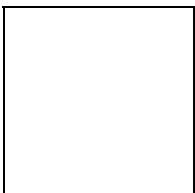
pe care adunându-le obținem :



Deoarece $na = 2p$, urmează că :



de unde rezultă $S_1=0$. În același mod se deduce că sunt valabile și relațiile :



Prin urmare sistemul normal de ecuații va fi :

... ..

de unde se află valorile coeficienților seriei Fourier $A_0, A_1, A_2, A_3, B_1, B_2, B_3$. Informații suplimentare referitoare la acest material se găsesc în [10].

2.3.4. Cazul perioadei necunoscute. Aflarea perioadei seriei.

În cazul unui proces periodic cu perioada T avem $X(t+T)=X(t)$ pentru orice t .

Dacă observăm acest proces în n momente de timp $t_0, t_0+1, \dots, t_0+n-1$ cu $n > T$, n suficient de

mare, seriile :

sunt teoretic identice. Înseamnă că raportul lor de corelație $r(X_{t_0}, X_{t_0+T})$ este 1. Ținând seama de această observație se poate determina perioada T .

Folosind autocorelația de rangul k $r_k = r(X_{t_0}, X_{t_0+k})$ unde $k=0,1,2,\dots$ conform observației precedente valoarea perioadei T este acel $k > 0$ pentru care $r_k = 1$.

Revenind la cazul real putem spune că seriile X_{t_0} și X_{t_0+T} nu sunt identice, astfel încât r_T nu va fi 1. Totuși metoda este funcțională în sensul că vom alege ca valoare a lui T acel k pentru care r_k și r_{2k} sunt maxime locale.

2.3.5. Intocmirea periodogramei. Extragerea informațiilor din periodogramă.

Graficul autocorelației, în care punctele (k, r_k) , $k \geq 0$ sunt pe grafic extremitățile segmentelor unei linii poligonale se numește corelogramă sau periodogramă. Periodogramă sugerează care sunt valorile posibile ale perioadei T . Astfel, se întocmește tabelul de autocorelație ce conține perechile (k, r_k) pentru $k=0, \dots, n$ unde n este numărul de măsurători efectuate la același interval de timp. Se construiește apoi periodogramă și perioada seriei dinamice va fi găsită ținând seama de observația făcută anterior că perioada T este acel k pentru care r_k și r_{2k} sunt maxime locale.

Dacă totuși nu dispunem de suficiente date astfel încât să se îndeplinească această condiție, funcția de corelație având o descreștere monotonă până la o valoare apropiată de -1 și apoi o creștere monotonă până la o valoare apropiată de $+1$ valori care se ating una la jumătatea intervalului de date și cea de-a doua la sfârșitul acestui interval putem spune că perioada coincide cu intervalul de timp pe care au fost efectuate observațiile.

O formă în care valorile din periodogramă tind spre 1 după un anumit timp ne sugerează că este vorba de un fenomen ce tinde să se comporte constant.

Dacă în periodogramă după un anumit timp apar periodic atât valori pozitive către $+1$ cât și valori negative mici către -1 în mod periodic acest lucru ne dă informații despre simetria comportării fenomenului respectiv.

O formă în care valorile din periodogramă sunt apropiate de 0 ne sugerează că este vorba de un fenomen care nu este nici periodic, nici staționar și nici nu tinde la staționaritate.

O metodă bună de a observa o serie dinamică se poate rezuma astfel :

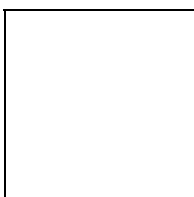
- se încarcă datele aplicației.
- se realizează reprezentarea grafică a punctelor $X=X(t)$ într-un sistem de coordonate cartezian (X, T) și se intuiește forma trendului.
- se calculează coeficienții trendului presupus și se reprezintă grafic. Se trasează grafic corelograma ce corespunde trendului respectiv care ne permite să sesizăm dacă este vorba despre un proces periodic sau nu.

- în cazul în care rezultatele constatate nu sunt satisfăcătoare se selectează alt tip de trend și pentru acesta se realizează din nou diagrama de corelație.
- se repetă acest proces până când rezultatele observațiilor sunt satisfăcătoare.
- cu trendul selectat se salvează valorile diferențelor $X(t)-T(t)$.
- pentru procesele periodice se determină perioada acestora cu ajutorul acestor date.
- se construiește seria numerică ce aproximează seria dinamică prin însușirea la trend a seriei de sezonalitate calculate cu ajutorul perioadei calculate anterior.

2.4. Evoluție în timp. Prognoză.

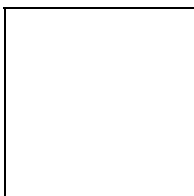
2.4.1. Tendința de stabilizare.

Pentru ca seria dinamică $X=X(t), t=1, \dots, t_n$ să fie complet definită este suficient să cunoaștem funcția de repartiție a șirului $X(t_1), \dots, X(t_n)$, deci să cunoaștem



pentru orice x_1, \dots, x_n reali.

Dacă funcția de repartiție rămâne neschimbată după un interval arbitrar de timp t adică :



atunci vom spune că seria dinamică este strict staționară.

Aceasta este o condiție extrem de severă care se întâlnește rar în forma pură. O condiție mai puțin restrictivă poate fi obținută făcând apel la noțiunea de covarianță :

$$\text{cov}(X, Y) = M((X - M(X)) * (Y - M(Y)))$$

definită prin :

$$\text{cov}(X(t), X(t+t)) = M((X(t) - M(X(t))) * (X(t+t) - M(X(t+t))))$$

În cazul în care media procesului este constantă ea poate fi considerată nulă deci :

$$M(X(t)) = M(X(t+t)) = 0,$$

pentru orice t și atunci :

$$g(t) := \text{cov}(X(t), X(t+t)) = M(X(t) * X(t+t))$$

Funcția $g(t) = g(-t)$ se numește Autocovarianța seriei dinamice $X(t)$, $t = t_1, \dots, t_n$.

Este evident că definirea unei astfel de mărimi are sens pentru procese staționare, ceea ce este atestat și de :

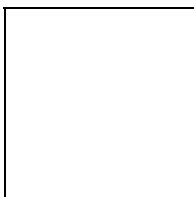
$$M(X(t)) = M(X(t+t)) = 0, \text{ pentru orice } t.$$

De menționat că conceptul de staționaritate implică o anumită regularitate statistică considerată în general incompatibilă cu existența unei tendințe (trend) a seriei dinamice. Și aceasta fie că este vorba de o tendință a mediei, dispersiei sau a ambelor.

De aceea cel mai puternic și eficient instrument de studiu și modelare a seriilor nestaționare rămâne funcția de autocorelație. Din considerente de facilitare în efectuarea calculelor funcția de autocorelație este definită ca o autocovarianță normalizată astfel încât la $t=0$ să fie egală cu unitatea :

$$r(t) = g(t) / g(0)$$

Pentru serii empirice la care termenii individuali se consideră în raport cu media seriei aceasta duce la formula :



utilizată des în aplicații. În aceste condiții funcția de autocorelație ce exprimă o anumită legătură internă a procesului, presupus staționar se constituie într-un instrument de calcul puternic, capabil să conducă la predicții.

2.4.2. Analiza Fourier. Reprezentarea Fourier a unui șir finit de date.

Tendința ciclică este aproximată de obicei prin serii Fourier trunchiate la un număr corespunzător n de termeni care reprezintă o combinație liniară de funcții trigonometrice de timp cu proprietăți de ortogonalitate :



Funcțiile care îndeplinesc condițiile de ortogonalitate :

sunt luate în acest caz funcții trigonometrice și . Coeficienții

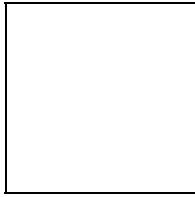
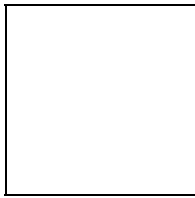
ai seriei reprezintă parametrii modelului care trebuie estimați.

Reprezentarea Fourier a unui șir finit de date.

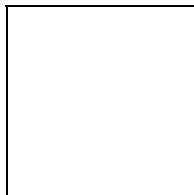
Dacă eșantionul de observații asupra seriei $y(1), \dots, y(N)$, se fac pe un interval N de timp

funcțiile se aleg ortogonale pe mulțimea $1, \dots, N$ iar frecvențele armonicilor care pot apare sunt respectiv i/N $i=0,1,2, \dots, n=N/2$, dacă $N=2n$ este par. Pulsațiile componentelor sunt $2\pi i/N$, $i=0,1, \dots, n$. Se presupune astfel că evoluția pe o perioadă a armonicii de bază cu frecvența cea mai joasă $1/N$ se încadrează pe întreaga lungime a eșantionului, adică această componentă are perioada cea mai lungă (egală cu N), iar armonica cu frecvența cea mai înaltă $1/2$ Hz are perioada egală cu dublul pasului de discretizare a timpului. În cazul când numărul observațiilor N este impar $n=(N-1)/2$. deci de-a lungul întregului interval de observare N se încadrează i perioade ale armonicii de frecvență i/N .

Proprietățile de ortogonalitate ale funcțiilor sinus și cosinus pentru aceste frecvențe sunt :



Reprezentarea prin serii Fourier trunchiate a unui șir finit de date $y(1), y(2), \dots, y(N)$ poate fi justificată intuitiv presupunând că șirul de numere $y(1), \dots, y(N)$ reprezintă valorile unei funcții continue de timp $y(t)$, eșantionată la momentele $t=1, 2, \dots, N$. Problema care se pune este să se găsească o funcție continuă $y(t)$ care să treacă prin punctele $y(1), \dots, y(N)$ în momentele de timp $t=1, 2, \dots, N$ și care să fie compusă dintr-o sumă de funcții trigonometrice ortogonale (sinusoidale și cosinusoidale) ca de exemplu :



în care $N=2n$.

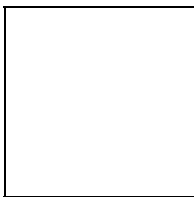
Acest model conține N constante a_i, b_i care pot fi determinate astfel încât $y(t)$ la momentele $t=1, 2, \dots, N$ să coincidă respectiv cu $y(1), \dots, y(N)$.

Valorile $y(1), \dots, y(N)$ sunt considerate coordonatele unui punct din spațiul N -dimensional al observațiilor. Valorile coeficienților c_1, \dots, c_n reprezintă coordonatele punctului în noul sistem tot N -dimensional.

Transformarea coordonatelor y în coordonate c se face prin intermediul unei matrice ortogonale de dimensiune $N \times N$ astfel încât $y = Lc$ iar $c = L^T y$ deci

$$L L^T = I$$

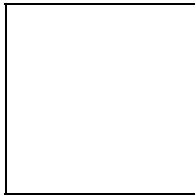
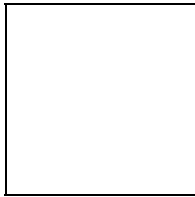
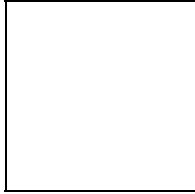
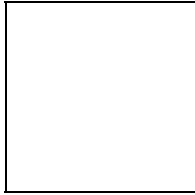
Particularitatea constă în alegerea unei matrice L ce are ca elemente funcții trigonometrice ortogonale alese corespunzător. O astfel de matrice pentru cazul $N = \text{par}$ poate fi :



Se observă că se respectă condiția de ortogonalitate $L L^T = I$. Folosind transformarea

$$c = L^T y$$

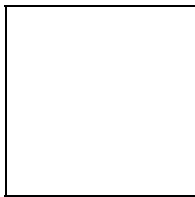
rezultă :



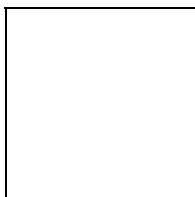
In acest caz din

$$Lc = Lc(LTy) = y$$

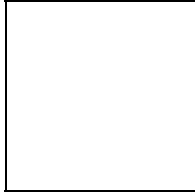
se obține :



Această dezvoltare constituie o reprezentare Fourier a șirului $y(1), \dots, y(N)$ cu coeficienții seriei c_2, \dots, c_n . Dacă N este impar matricea L este



iar reprezentarea Fourier corespunzătoare rezultată din $Lc = y$ este



2.4.3. Serii periodice. Reprezentarea Fourier a unui șir periodic de date.

Presupunem că șirul $y(1), \dots, y(N)$ are o evoluție periodică cu perioada n_0 : număr întreg, astfel încât pe intervalul $[1, N-n_0]$:

$$y(t+n_0)=y(t) \quad t=1,2,\dots,N-n_0$$

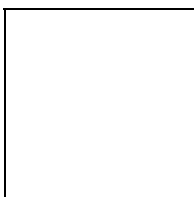
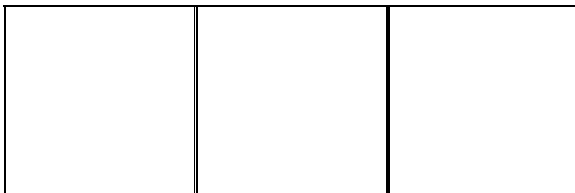
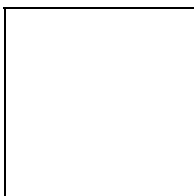
Dacă notăm cu $y_0 = [y(1)y(2)\dots y(n_0)]^T$, șirul periodic inițial, în cazul când $N-hn_0$, se poate scrie sub forma : $y_0 = [y(1)y(2)\dots y(n_0h)]^T$.

În acest mod matricea L_0 va trebui să aibă dimensiunea $n_0 \times n_0$ și ca atare :

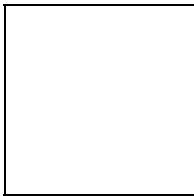
$$y_0 = L_0 x_0 \text{ și } x_0 = L_0^T y_0$$

Prin urmare seria Fourier de reprezentare va avea n_0 termeni.

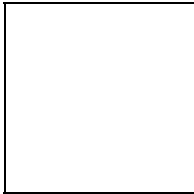
Coefficienții dezvoltării sunt :



În acest caz pentru $t=1,2, \dots, n_0$ va fi :



iar pentru $t > n_0$:



Periodicitatea poate proveni dintr-o tendință ciclică sau o tendință sezonieră. În funcție de acest fapt se aleg și metodele de investigare.

2.4.4. Estimarea parametrilor tendinței ciclice

Componenta ciclică $y_c(t)$ trebuie determinată pe baza observațiilor asupra seriei inițiale :

$$y_0(t) = y_c(t) + y(t) \quad t = 1, \dots, N.$$

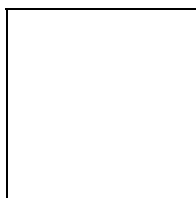
Componenta aleatoare o considerăm zgomot alb de medie 0. Presupunem că perioadele tendinței sunt cunoscute.

Se consideră că N este divizibil prin aceste perioade :

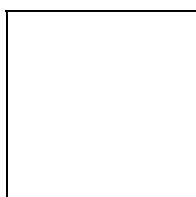
$$n_1, \dots, n_q \quad n_i = N/k_i, \quad i=1, \dots, q$$

în care k_1, \dots, k_q este o submulțime de numere întregi din șirul $1, 2, \dots, (N-1)/2$ pt. N impar și $N/2-1$ pt. n par iar în acest caz putem să includem și armonica cu perioadă 2.

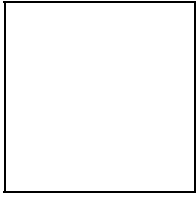
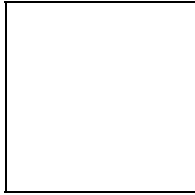
Pentru cazul când N este impar (fără armonica de perioadă 2) seria poate fi :



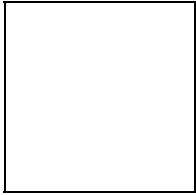
În cazul când N este par se poate include și armonica de frecvență $1/2$ Hz :



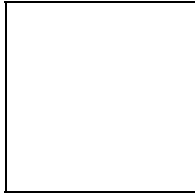
În notație matriceală, pentru N par :



iar pentru N impar :

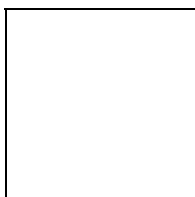
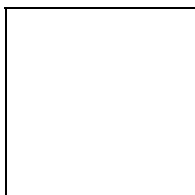
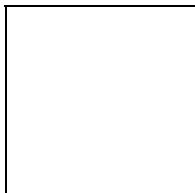


Estimațiile parametrilor se obțin folosind criteriul celor mai mici pătrate :

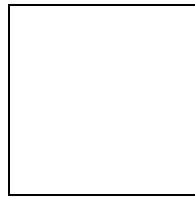


și am înțeles că c este vectorul format cu argumentele funcției dintre acolade care este chiar suma S a pătratelor erorilor care se minimizează în funcție de aceleași argumente c .

Având în vedere proprietățile de ortogonalitate ale funcțiilor $\varphi_i(t)$ rezultă că ecuațiile normale furnizate de criteriul celor mai mici pătrate sunt independente și estimațiile rezultă direct :

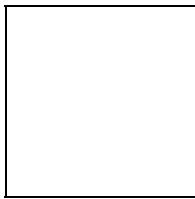


iar în cazul N-par se calculează și estimația :



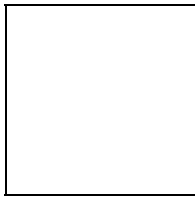
2.4.5. Modelul armonic al tendinței sezoniere

O funcție sinusoidală sau cosinusoidală se poate reprezenta ca o combinație liniară de sinusoidă și cosinusoidă. De exemplu :

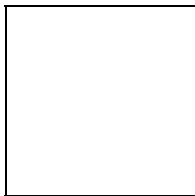


unde $a = R \sin \alpha$ și $b = -R \cos \alpha$

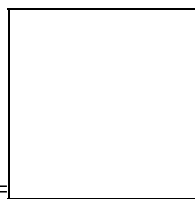
Ținând cont de aceasta, modelul tendinței ciclice pentru N impar va fi :



iar modelul pentru N par devine :

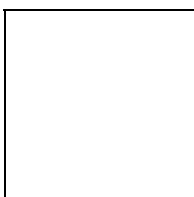


în care :

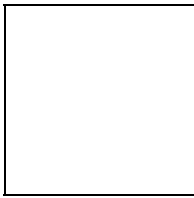


$R_0 = a_0 = M(y_0)$ și $R_{N/2} = a_{N/2}$, $R_i =$, $a_i = R \cos \alpha_i$, $b_i = R \sin \alpha_i$ $\alpha_i = \arctg(-b_i/a_i)$

Deoarece a_i și b_i sunt estimațiile parametrilor ce minimizează suma pătratelor :



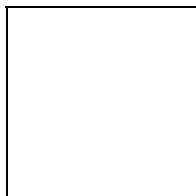
atunci și R_i și α_i sunt estimațiile ce minimizează suma pătratelor :



2.4.6. Teorema Parseval. Spectrograma Shuster.

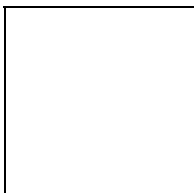
Media pătratică a mărimii $y_c(t)$ pe intervalul N de observație se mai numește în tehnică și puterea medie a semnalului $y_c(t)$.

Din relația de mai sus rezultă că media pătratică a tendinței este :



Mai comod este să se exprime puterea medie a componentei ciclice, adică a semnalului $y_c(t)$ centrat prin componenta R_0 care corespunde valorii medii a semnalului $y_c(t)$. Puterea medie a unui semnal centrat corespunde dispersiei semnalului.

Teorema Parseval. Dispersia semnalului $y_c(t)$ poate fi descompusă pe armonici :



Spectrograma Schuster se numește diagrama obținută prin reprezenta-rea puterii medii a fiecărei armonici, în dreptul frecvenței aferente armonicii. Uneori în locul spectrogramei se trasează periodogramă.

Periodograma Schuster este aceeași reprezentare grafică a puterii medii a fiecărei armonici în dreptul perioadei acesteia.

Spectrograma seriei inițiale pune în evidență armonicile semnificative ale tendinței ciclice.

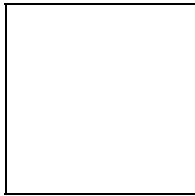
2.4.7. Metoda Whittaker-Robinson

Variațiile sezoniere apar datorită caracterului de reproducere periodică a activităților. În acest caz peste tendința generală lent variabilă în timp $y_T(t)$ se suprapune o componentă deterministă periodică $y_S(t)$ în afara celei aleatoare $y(t)$.

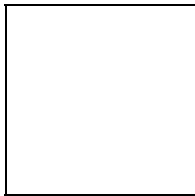
Dacă seria conține variații sezoniere numărul total de observații N asupra seriei se alege divizibil prin nS : perioada componentei sezoniere.

Metoda Whittaker-Robinson de căutare a perioadei componentei sezoniere constă în construirea unui tabel cu n coloane pe baza căruia se calculează estimarea dispersiei reziduurilor.

Deoarece :

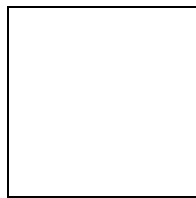


estimarea dispersiei se poate calcula cu relația :



în care $k = N/n$ este numărul de linii pline ale tabelului întocmit.

Calculul se repetă pentru $n=2,3, \dots$ în mod crescător până când se atinge prima valoare



minimă a estimației care este atinsă pentru $n = ns$.

Metoda prezentată prezintă avantajul că furnizează simultan estimațiile componentei sezoniere și ale perioadei acesteia.

de timp

Capitolul 3. Implementarea metodelor de analiza seriilor

de timp

3.13 Algoritmii implementati

SubAlgoritmul CitModSerie (n,X) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

X : lista de perechi de reali; { seria dinamica $X=X(t)$ }

Variabile locale :

i : intreg;

indice : intreg;

optiune : sir de caractere;

valoare : pereche de reali;

Start :

Citeste (optiune);

Daca optiune='citire' atunci

 Citeste (optiune);

 Daca optiune='fisier' atunci

 Citeste (nume);

 Citeste din fisierul numeÎn seria X;

 SfD;

 Daca optiune='direct' atunci

 Citeste (n);

 Pentru i=1,n executa

 Citeste (X[i]);

 SfP;

 Sfd;

 SfD;

Daca optiune='modificare' atunci

 Repeta

 Citeste (indice);

 Daca indice \leq 0 atunci IesiDinCiclu;


```
Citeste ( optiune );
Daca optiune='sterge' atunci
    Elimina pe X[indice] din lista X;
SfD;
Daca optiune='adauga' atunci
    Citeste ( valoare );
    AadaugaÎn lista X : X[indice]:=valoare;
SfD;
Daca optiune='modifica' atunci
    Citeste ( valoare );
    X[indice]:=valoare;
SfD;
```

pana cand indice \leq 0;

SfD;

SfS;

Subalgoritmul SalvSerie (n,X,nume) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

nume : fisier; { nume fisierÎn care salvam seria dinamica $X=X(t)$ }

X : lista de perechi de reali; { seria dinamica }

Variabile locale :

Y,Z : lista de perechi de reali;

optiune : sir de caractere;

Start :

Citeste (nume);

Citeste (optiune);

Daca optiune='adaugare' atunci

Citeste din fisierul nume in seria Y;

Aadauga seria X seriei Y;

Salveaza seriaY in fisierul nume;

SfD;

Daca optiune='interclasare' atunci

Interclaseaza seriile X si Y in Z;

Salveaza seria Z in fisierul nume;

SfD;

Daca optiune='initializare' atunci

Salveaza seria X in fisierul nume;

SfD;

SfS;

Subalgoritmul Covarianta (n,X,ix,Y,iy,z) este:

Parametrii:

n:intreg; {numarul termenilor seriei}

X:lista de reali; {vector 1 de date}

ix:intreg; {puterea vectorului 1 de date}

Y: lista de reali; {vector 2 de date}

iy:intreg; {puterea vectorului 2 de date}

z:real; {valoarea covariantei celor 2 vectori de date}

Variabile locale:

wx,wy,wz:reali;

i,j:intregi;

Start:

În valoare absoluta;

Schimba liniile j si i intre ele în $Z[mx0..mx1][mx0..mx1+1]$;

Cu noul $ZZ[i][i]$ imparte linia i la el;

Cu noul $ZZ[i][i]=1$ fa 0 în coloana i prin transformari
elementare asupra liniilor i+1,...,mx1;

SfP;

Pentru i=mx1,mx0,-1 executa

Obtine elementul $a[i]=Z[i][mx1+1]$ - EMBED Equation.2

SfP;

SfS;

Subalgoritmul FTrendTaylor (m0,m1,A,x,v) este :

Parametrii :

m_0, m_1 : intreg; { ordinul termenului inferior si ordinul termenului superior al polinomului de trend }

A : lista de reali; { coeficientii trendului polinomial }

x : real; { valoarea argumentului }

v : real; { valoarea intoarsa de functie }

Variabile locale :

i, j : intreg;

s : real;

Start :

v:=0;

Pentru i:=m₀,m₁ executa

s:=1;

Pentru j:=1,i,sgn(i) executa

s:=s(x;

SfP;

Daca i<0 atunci

s:=1/s;

SfD;

v:=v+s(A[i];

SfP;

SfS;

Subalgoritmul TrendTaylor(n,X,A) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

m_0, m_1 : intregi; { ordinul inferior si ordinul superior al functiei polinomiale de trend }

X : lista de perechi de reali; { seria dinamica }

A : lista de reali; { coeficientii de trend calculati }

Variabile locale :

XX,YY : lista de reali;

ZZ : matrice[m₀,...,m₁]([m₀,...,m₁,m₁+1] de reali;

i,j:intregi;

Start :

```

Mesaj(' Da gradul termenului inferior : ');
Citeste( m0 );
Mesaj(' Da gradul termenului superior : ');
Citeste( m1 );
Pentru i=0,n executa
    XX[i]:=X[1][i];
    YY[i]:=X[2][i];
SfP;
Pentru i=m0,m1 executa
    Pentru j=m0,m1 executa
        Cheama Covarianta(n,XX,i,XX,j,ZZ[i][j]);
    SfP;
    Cheama Covarianta(n,XX,i,YY,1,ZZ[i][m1+1]);
SfP;
Cheama RedGauss(ZZ,m0,m1,A);
Pentru i:=1,n executa
    Cheama FTrendTaylor(m0,m1,A,X[2][i],T[2][i]);
SfP;

```

SfS;

Subalgoritmul FunctieTrend (A,wA,int,rez) este :

Date :

Fct : vector de functii f : (((;

Parametrii :

A : lista de reali; { coeficienti folositi de functia de trend }

wA : intreg; { indice ce da functia de trend }

int : real; { valoarea x pentru care se calculeaza f(x) }

var rez : real; { intoarce valoarea functiei }

Start :

rez:=Fct[wA](int);

SfS;

Subalgoritmul SalvezTrend(n,X,wA) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

X : lista de perechi de reali; { seria de timp $X=X(t)$ }

A : lista de reali; { coeficientii trendului }

wA : intreg; { cod pentru tipul trendului }

Variabile locale :

XX,YY,ZZ : lista de reali;

i,j : intregi;

nume : sir de caractere;

Start :

Pentru i=1,n executa

XX[i]:=X[1][i];

YY[i]:=X[2][i];

Cheama FunctieTrend (A,wA,XX[i],ZZ[i]);

SfP;

Mesaj(' Da nume de fisier trend: ');

Citeste (nume);

Deschide pentru citire (nume);

Pentru i=1,n executa

ScrieÎn fisier

(nume , ' F(' + Caractere(XX[i]) + ')=' + Caractere(YY[i]-ZZ[i]));

SfP;

EndS;

Subalgoritmul IndSort (n,X,Y,c1,c2) este :

{ face sortarea dupa indici a lui X }

Date :

n : intreg; { numarul termenilor seriei }

X,Y : lista de perechi de reali; { seria dinamica si indicii sortati }

c1,c2 : intregi; { capetele intre care se face sortarea lui X }

Variabile locale :

i,j : intregi;

t : real;

Start :

Pentru i=c1,c2 executa

Y[1][i]:=i-c2+1;

Y[2][i]:=i-c2+1;

SfP;

Pentru $i=c1, c2-1$ executa

Pentru $j=i+1$ to $c2$ executa

Daca $X[2][i]>X[2][j]$ atunci

$t:=y[2][j];$

$Y[2][j]:=Y[2][i];$

$Y[2][i]:=t;$

SfD;

SfP;

Sfp; { Y [c1...c2] sirul indicilor lui X a.i. $X[2][Y[2][i]]$ crescator }

SfS;

Subalgoritmul AutoCorRang (n,X,ACR) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

X : lista de perechi de reali; { seria dinamica $X=X(t)$ }

ACR : matrice de reali; { autocorelatii de rang de subserii de lungimi
variabile dupa primul indice }

Variabile locale :

Y,Z : lista de perechi de reali;

i,j,k1,k2,k : intregi;

d : real;

Start :

Pentru $i:=1, n/2$ executa

Pentru $j:=1, n-i+1$ executa

$k1:=j;$

$k2:=j+i;$

Cheama IndSort(n,X,Z,1,i);

Cheama IndSort(n,X,Y,k1,k2);

$d:=0;$

Pentru $k=1, i$ executa

$d:=d+(Z[k]-Y[k+k1])((Z[k]-Y[k+k1]));$

SfP;

ACR[i][j]:=6(d²/(i³-i));

SfP;

SfP;

SfS;

Subalgoritmul Covar(n,,k,X,z) este :

Parametrii :

n,k : intregi;

X : lista de perechi de reali;

z : real;

Variabile locale :

wz : real;

i,j : intregi;

Start :

wz:=0.0;

Pentru i=1,n-k executa

wz:=wz+X[2][i](X[2][i+k]);

SfP;

z:=wz/n;

SfS;

Subalgoritmul Dispe(n1,n2,X,rez) este :

Parametrii :

n1,n2 : intregi; { indicii primului si ultimului termen al seriei

luat în calcul }

X : lista de perechi de reali; { seria dinamica }

rez : real; { valoarea dispersiei subseriei X(n1)...X(n2) }

Variabile locale :

wz : real;

i : intregi;

Start :

wz:=0;

Pentru i:=n1,n2 executa

wz:=wz+X[2][i];

SfP;

wz:=wz/(n2-n1+1);

rez:=0;

Pentru i:=n1,n2 executa

rez:=rez+(X[2][i]-wz)²;

SfP;

rez:=rez/(n2-n1+1);

SfS;

Subalgoritmul Autocorelatie(n,k,X,z) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

k : intreg; { ordinul autocorelatiei }

X : lista de perechi de reali; { seria dinamica }

z : real;

Variabile locale :

wz,wx,wy : reali;

Start :

Cheama Covar(n,k,X,wz);

Cheama Dispe(1,n-k,X,wx);

Chema Dispe(k,n,X,wy);

z:=wz/ $\sqrt{wx*wy}$;

SfS;

Subalgoritmul Periodograma(n,X,Y) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

X,Y : lista de perechi de reali; { seria dinamica si periodograma calculata }

Variabile locale :

A,B,S : reali;

i,j : intreg;

jx : real;

Start :

Pentru j:=1,2n executa

jx:=j/2;

A:=0;

B:=0;

Pentru i:=1,n executa

A:=A+X[2][i](cos(2(i/jx));

B:=B+X[2][i](sin(2(i/jx));

Sfp;

S:=A²+B²;

Y[2][j]:=S; { valorile servesc la reprezentarea grafica S=S(j/2) }

Y[1][j]:=j;

SfP;

SfS;

Subalgoritmul Perioada (n,X,p) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

X : lista de perechi de reali; { seria dinamica X=X(t) }

p : real; { perioada ce rezulta prin calcul }

Variabile locale :

Y : lista de reali;

i,k : intregi;

Start :

Pentru i:=1,n executa

Cheama Autocorelatie(n,i,X,Y[i]);

SfP;

i:=0;

Pentru k:=n/2,2,-1 executa

Daca

(Y[k-1]<Y[k])si(Y[k]<Y[k+1])si(Y[2k-1]<Y[2k])si(Y[2k]<Y[2k+1])

atunci

i:=k;

SfD;

SfP;

p:=i;

SfS;

Subalgoritmul SeriaSezonalitate (n,X,T,p,A,B) este :

Date :

n : intreg; { numarul termenilor seriei }

X : lista de perechi de reali;

T : lista de perechi de reali; { valorile calculate ale trendului în cele n pct }

p : real; { perioada calculata a seriei dinamice }

A,B : lista de reali; { coeficientii seriei de sezonalitate }

Variabile locale :

i,j : intreg;

Start :

Cheama Perioada(n,X,p);

Daca p=0 atunci

Pentru i:=1,n executa

A[i]:=0;

B[i]:=0;

SfP;

altfel

Pentru i:=1,n executa

A[i]:=0;

B[i]:=0;

Pentru j:=1,n executa

A[i]:=A[i]+2((X[2][j]-T[2][j])cos(2(iX[1][j]/p)/n);

B[i]:=B[i]+2((X[2][j]-T[2][j])sin(2(iX[1][j]/p)/n);

SfP;

SfP;

SfD;

SfS;

Subalgoritmul FunctiaSezonalitate (n,A,B,p,t,s) este :

Date :

n : intreg; { numarul termenilor seriei }

A,B : lista de reali; { coeficientii seriei de sezonalitate }

t : real; { momentul de timp }

s : real; { valoarea seriei de sezonalitate la momentul t }

p : real; { perioada seriei dinamice }

Variabile locale :

i : intreg;

Start :

s:=0;

Pentru i:=1,n executa

s:=s+A[i](cos(2(it/p)+B[i] (sin(2(it/p);

SfP;

SfS;

Subalgoritmul Corelograma (n,A,B,p,C) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

X : lista de perechi de reali; { seria dinamica }

T : lista de perechi de reali; { valorile trendului în punctele seriei }

A,B : lista de reali; { coeficientii seriei de sezonalitate }

p : real; { perioada seriei dinamice }

C : lista de reali; { corelograma }

Variabile locale :

Eps : lista de perechi reali;

dis,ro : real;

i : intreg;

Start :

Pentru i:=1,n executa

Eps[1][i]:=i;

Cheama FunctiaSezonalitate(n,A,B,p,X[1][i],C[i]);

Eps[2][i]:=X[2][i]-T[2][i]-C[i];

SfP;

Cheama Dispe(1,n,Eps,dis);

ro:=0;

Pentru i:=1,n executa

ro:=ro+A[i](A {i}+ B[i](B {i};

SfP;

ro:=ro+2(dis;

Pentru i:=1,n executa

$C[i]:=C[i]/ro;$

SfP;

SfS;

Subalgoritmul FRegHip1 (A,x,v) este :

Parametrii :

A : lista de reali; { coeficientii trendului polinomial }

x : real; { valoarea argumentului }

v : real; { valoarea intoarsa de functie }

Start :

$v:=1/(A[1](x+A[0]));$

SfS;

Subalgoritmul RegHip1 (n,X,A) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

X : lista de perechi de reali; { seria dinamica }

A : lista de reali; { coeficientii de trend calculati }

Variabile locale :

XX,YY : lista de reali;

ZZ : matrice[0,...,1]([0,...,1,2] de reali;

i,j:intregi;

Start :

Pentru $i=0,n$ executa

$XX[i]:=X[1][i];$

$YY[i]:=1/X[2][i];$

SfP;

$m0:=0;$

$m1:=1;$

Pentru $i=m0,m1$ executa

Pentru $j=m0,m1$ executa

Cheama Covarianta(n,XX,i,XX,j,ZZ[i][j]);

SfP;

Cheama Covarianta(n,XX,i,YY,1,ZZ[i][m1+1]);

SfP;

Cheama RedGauss(ZZ,m0,m1,A);

Pentru $i:=1,n$ executa

Cheama FTrendTaylor(m0,m1,A,X[2][i],T[2][i]);

T[2][i]:=1/T[2][i];

SfP;

SfS;

Subalgoritmul FRegHip2 (A,x,v) este :

Parametrii :

A : lista de reali; { coeficientii trendului polinomial }

x : real; { valoarea argumentului }

v : real; { valoarea intoarsa de functie }

Start :

v:=1/(A[1](x)-A[0]);

SfS;

Subalgoritmul RegHip2 (n,X,A) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

X : lista de perechi de reali; { seria dinamica }

A : lista de reali; { coeficientii de trend calculati }

Variabile locale :

XX,YY : lista de reali;

ZZ : matrice[0,...,1]([0,...,1,2] de reali;

i,j:intregi;

Start :

Pentru i=0,n executa

XX[i]:=X[2][i];

YY[i]:=1/X[1][i];

SfP;

m0:=0;

m1:=1;

Pentru i=m0,m1 executa

Pentru j=m0,m1 executa

Cheama Covarianta(n,XX,1,XX,1,ZZ[i][j]);

SfP;

Cheama Covarianta(n,XX,i,YY,1,ZZ[i][m1+1]);

SfP;

Cheama RedGauss(ZZ,m0,m1,A);

Pentru $i=1,n$ executa

Cheama FRegHip2($m_0,m_1,A,X[2][i],T[2][i]$);

SfP;

SfS;

Subalgoritmul FRegHip3 (A,x,v) este :

Parametrii :

A : lista de reali; { coeficientii trendului polinomial }

x : real; { valoarea argumentului }

v : real; { valoarea intoarsa de functie }

Start :

$v:=(A[1](A[2]-A[3])/(x+A[2])-A[1])$;

SfS;

Subalgoritmul RegHip3 (n,X,A) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

X : lista de perechi de reali; { se

A : lista de reali; { coeficientii de trend calculati }

Variabile locale :

$XX1,XX2,XX3,YY$: lista de reali;

ZZ : matrice $[1,\dots,3][1,\dots,3,4]$ de reali;

i,j :intregi;

Start :

Pentru $i=0,n$ executa

$XX1[i]:=X[1][i]$;

$XX2[i]:=X[2][i]$;

$XX3[i]:=1$;

$YY[i]:=X[1][i](X[2][i])$;

SfP;

Cheama Covarianta($n,XX1,1,XX1,1,ZZ[1][1]$);

Cheama Covarianta($n,XX1,1,XX2,1,ZZ[1][2]$);

Cheama Covarianta($n,XX1,1,XX3,1,ZZ[1][3]$);

Cheama Covarianta($n,XX2,1,XX2,1,ZZ[2][2]$);

Cheama Covarianta($n,XX2,1,XX3,1,ZZ[2][3]$);

Cheama Covarianta($n,XX3,1,XX3,1,ZZ[2][3]$);

$ZZ[2][1]:=ZZ[1][2]$;

```
ZZ[3][1]:=ZZ[1][3];
ZZ[3][2]:=ZZ[2][3];
Cheama Covarianta(n,XX1,1,YY,1,ZZ[1][4]);
Cheama Covarianta(n,XX2,1,YY,1,ZZ[2][4]);
Cheama Covarianta(n,XX3,1,YY,1,ZZ[3][4]);
Cheama RedGauss(ZZ,1,3,A);
```

Pentru i:=1,n executa

```
Cheama FRegHip3(1,3,A,X[2][i],T[2][i]);
```

SfP;

SfS;

Subalgoritmul TrendTaylor1(n,m0,m1,X,A) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

m0,m1 : intregi; { ordinul inferior si ordinul superior al
functiei polinomiale de trend }

X : lista de perechi de reali; { seria dinamica }

A : lista de reali; { coeficientii de trend calculati }

Variabile locale :

XX,YY : lista de reali;

ZZ : matrice[m0,...,m1]([m0,...,m1,m1+1] de reali;

i,j:intregi;

Start :

Pentru i=0,n executa

```
XX[i]:=X[1][i];
```

```
YY[i]:=X[2][i];
```

SfP;

Pentru i=m0,m1 executa

Pentru j=m0,m1 executa

```
Cheama Covarianta(n,XX,i,XX,j,ZZ[i][j]);
```

SfP;

```
Cheama Covarianta(n,XX,i,YY,1,ZZ[i][m1+1]);
```

SfP;

Cheama RedGauss(ZZ,m0,m1,A);

Pentru i:=1,n executa

```
Cheama FTrendTaylor(m0,m1,A,X[2][i],T[2][i]);
```

SfP;

Subalgoritmul OrdTrendPol (n,X,o) este :

Parametrii :

n : intreg; { numarul termenilor seriei }

X : lista de perechi de reali; { seria dinamica }

o : intreg; { valoarea rezultat a ordinului tendintei polinomiale }

Variabile locale :

lamb : lista de reali;

Start :

Pentru i:=0,n-2 executa

Cheama TrendTaylor1(n,0,i,X,A);

lamb[i]:=0;

Pentru j:=1,n executa

lamb[i]:=lamb[i]+(X[2][j]-T[2][j])((X[2][j]-T[2][j]));

SfP;

lamb[i]:=lamb[i]/(n-i-1);

SfP;

o:= -1;

Repeta

o:=o+1;

PanaCand (lamb[o+1]>lamb[o])sau(o=n-1);

SfS;

3.2.Texte sursa

Continutul fisierului I0.pas:

contine driverul grafic egavga.bgi sub forma de modul .obj

Continutul fisierului I1.pas:

uses graph,crt;

const

p2:array[0..3] of word=(1,2,4,8,18,32,64,128,256,512,1024,2048,4096,8192);


```

    {folosit pentru usurarea calculelor de Transformata Fourier}
type
  treal=extended; {tipul de data generic al clasei nr reale}
  nrcomplex=record
    x,y:treal;
  end;
  st6=string[6]; pop=^optiuni;
  optiuni=record sir:st6;nrr:byte;urr:pop;ann:pop;end;
  {tipul de data optiuni}
  pocr= record SizeIm: Word ;pim1,pim2,pim3,pim4,pim5: Pointer; end;
  {salvarea de imagini}
type
  sircor=array[0..100] of treal;
var
  fx:sircor;
var
  pop1,pop2,popc:pop;
type
  plist=^list;
  list=record;
    x,fdex:treal ;{tipul de data "serie dinamica"}
    u:plist;
  end;
  pt1=^t1;t1:array[0..30,0..30] of treal; {matricea reducerii gauss}
  fun:procedure(var xc:plist); {tip procedural generic al clasei}
  tfu:function(tx:treal):treal; {tip procedural generic al clasei}
var
  pf:integer; {fourier}
  cap0,capx,cap1:plist;
  code:integer; {global code error}
  xp:fun; {variabila de tip functie}
  xf:tfu; {variabila de tip procedura}

type
  pcoefT=^coefT;

```

```

coefT=record;
  c:real;
  u:pcoeft;
end; {tip de data coeficienti de functie}
var
coef:record;
  trd:pcoeft;
  sd:pcoeft;
  tff:plist;
end; {grupul tipurilor de functii folosite}
type
pp=record;
  mix,max,miy,may,mdx,mdy,dix,diy:treal;
  inx,sux,iny,suy,ryx,mxy:treal;
end; {grupul parametrilor statistici}
var
p:pp;
type
st7=string[7];
ptmenu7=^tmenu7;
tmenu7=record;
  help:string;
  ss:st7;
  xf:fun;
  xf0:tfu;
  a,u:ptmenu7;
end; {tip de data meniu in explorare}
var
ps00,ps01,ps10,ps20,ps21,ps30,ps31,ps40,ps41,psxx:ptmenu7;
var
sim:pecr;
rc:string[20];
T:real; {global perioada seriei dinamice}
var
cc0:plist;

```

```

    ppp,m,mm,j,L,q,K,K2:word;
    s:string;
    psi,putereapsi,eta:NrComplex,
procedure Trend(var xc:plist);
    {functie complexa de depistare si evaluare trend}
    {returneaza un pointer la o functie ce contine trendul fiecarui punct}

    { Action:  Evaluate Trend(Data) }
    { Data   :  Memory and Stack }
    { Perform:  Liniar-Polinomial-... }
begin
end;
procedure Echidist(var xc:plist;num:integer);
    {e important ca punctele sa fie echidistante}
    {construim num puncte echidistante prin curbe de ordin 2}
    { Action:  Reevaluate Data}
    { Data   :  Memory}
    {Perform:  Parabolic Interpolation}
begin
end;
function Perioada(xc:plist):treal;
{returneaza o functie constanta ce arata perioada cu o precizie jumata
din distanta dintre doua puncte: "T"}
    { Action:  Evaluate Period T}
    { Data   :  Memory}
    {Perform:  Calculate Array Autocorelation}
begin
    perioada:=t;
end;
function NrTotalPlist(xc:plist):integer;
var
    NrT:integer;
begin
    NrT:=0; whilexc<>nil do begin inc(NrT);xc:=xc^.u;end;

```

```

    NrTotalPlist:=NrT;
end; {numara cate elemente (puncte) se gasesc in lista xc:plist}
function PozitieInPlist(xc,zc:plist):integer;
var
    yc:plist;
    NrT:integer;
begin
    NrT:=0;
    while (xc<>zc^.u) do begin
        if xc=nil then begin PozitieInPlist:=0;exit;end;
        inc(NrT);
        xc:=xc^.u;
    end;
    PozitieInPlist:=NrT;
end; {intoarce pozitia elementului zc:plist in lista xc:plist}
function PointerInPlist(xc:plist;ic:integer):plist;
var
    NrT:integer;
begin
    if (ic<=1) or (xc=nil) then begin PointerInPlist:=xc;exit;end;
    NrT:=1;
    while (xc^.u<>nil) do begin
        inc(NrT);
        xc:=xc^.u;
        if NrT=ic then begin PointerInPlist:=xc;exit;end;
    end;
    PointerInPlist:=xc;
end; {introduce elementul de pe pozitia ic din lista xc:plist}
procedure AduPlistXsiFdeX(var pxn:plist;var xxx,fff:treal);
var
    xypi:plist;
begin
    xypi:=pxn;
    new(pxn);
    pxn^.u:=xypi;pxn^.x:=xxx;pxn^.fdex:=fff;

```

```

end; {adauga in lista dupa elementul pxn:plist punctul de coordonate xxx,fff}
procedure ElibPlistXsiFdeX(var pxn:plist);
var
  xypi:plist;
begin
  xypi:=pxn;
  while pxn<>nil do begin
    pxn:=pxn^.u;
    dispose(xypi);
    xypi:=pxn;
  end;
end; {goleste lista pxn}
procedure MovePlist(p11:plist;var p12:plist);
var p111,p112:plist;
begin
  p111:=p11;
  p112:=nil;
  while p111<>nil do begin
    adauplistxsifdex(p112,p111^.x,p111^.fdex);
    p111:=p111^.u;
  end;
  elibplistxsifdex(p11);
  p12:=p112;
end; {muta o lista cu inversarea capului cu coada}
function k22 ( k,p,q:word):word;var i,k2:word; kk:array[0..20] of word;
begin for i:=p downto p do begin kk[i] :=k div p2[i] ; k:=k mod p2[i] ;end;
  k2:=0; for i:=q to p do k2:=k2+kk[i] * p2[p+q-i] ; k22:=k2;
end; {descompunere binara folosita de algoritmul FFT}
function jjj (k,p:word):word; var i,j:word; kk:array[0..20] of word;
begin for i:=p down to 0 do begin kk[i] :=k div p2[i] ; k:=k mod p2[i] ;end;
  j:=0; for i:=0 to p do j:=j+kk[i] * p2[p-i]; jjj:=j;
end; {descompunere binara folosita de algoritmul FFT}
procedure InmulComplexa3 (a,b:plist; var c:nrcomplex); begin c.x:=a^.x*b^.x-
a^.fdex*b^.fdex;c.y:=a^.fdex*b^.x+a.x*b^.fdex;end;

```

```

procedure InmulComplexa2 (a,b:nrcomplex;var c:nrcomplex);begin c.x:=a.x*b.x-a.y*b.y;
c.y:=a.y*b.x+a.x*b.y; end;
procedure CalculeazaCoeficientiComplecsi;
var
  cc1,cc2,eps1,eps2,ccx:plist;
begin
  cc1:=nil;eps1:=nil;K:=0;
  m:=p2[ppp];mm:=m;q:=ppp;psi.x:=cos(pi/m);psi.y:=sin(pi/m);
  putereapsi.y:=0.0;
  ccx:=cap0;
  for j:=0 to 2*m-1 do begin
    putereapsi.x:=ccx^.fdex;
    adauplistxsifdex(cc1,putereapsi.x,putereapsi.y);
    ccx:=ccx^.u;
  end;
  MovePlist(cc1,cc1);
  putereapsi.x:=1.0;
  for j:=m downto 1 do begin
    InmulComplexa2(putereapsi,psi,putereapsi);
    eta.x := - putereapsi.x; eta.y := -putereapsi.y;
    adauplistxsifdex (eps1,eta.x,eta.y);
  end; eps2:=eps1;
  for j:=m downto 1 do begin
    eta.x:=-eps2^.x; eta.y:= -eps2^.fdex;
    adauplistxsifdex (eps1,eta.x,eta.y);
    eps2:=eps2^.u;
  end;
  putereapsi.y:=0.0;putereapsi.x:=1.0;
  adauplistxsifdex (eps1, putereapsi.x,putereapsi.y);
  for L:=1 to ppp+1 do begin
    while K<2*m-1 do begin
      for j:= 1 to mm do begin
        k2:=k22 (k,ppp,q);
        cc2:=PointerInPlist (cc1,K+mm+1); eps2:=PointerInPlist (cc1,K+1);

```

```

    InmulComplexa3(cc2,pointerinPlist(eps1,K2+1),eta);
    cc2^.x:=eps2^.x-eta.x; cc2^.fdex:=eps2^.fdex-eta.y;
    eps2^.x:=eps2^.x+eta.x;eps2^.fdex:=eps2^.fdex+eta.y;
    inc (K);
end; K:=K+mm;
end; K:=0; mm:=mm div 2 ; dec(q) ;
end;
eliblistxsifdex(eps1);
while K<2 *m-1 do begin
j:=jjj(k,ppp);
if j>K then begin
    cc2:=pointerInPlist(cc1, j+1);eps2:=pointerInPlist(cc1,k+1);
    eta.x:=cc2^.x ; eta.y:=cc2.fdex;
    cc2^.x:=eps2^.x;cc2^.fdex:=eps2^.fdex;
    eps2^.x:=eta.x;eps2^.fdex:=eta.y;
end; inc (K) ;
end;
cc0:=cc1;
end; {FFT Directa}
procedure CalculeazaCoeficientiReali;
var
    cc1:plist;
begin
    cc1:=cc0;
    cc^.x:=cc1^.x/m;cc1:=cc1^.u;
    for j =1 to m-1 do begin
        cc1^.x:=cos(-j*pi)*cc1^.x/m;
        cc1^.fdex:=cos(-j *pi)*cc1^.fdex/m;
        cc1:=cc1^.u;
    end;
    cc1^.x:=cos(-m*pi)*cc1^.x/m;
end; { FFT Inversa }
function FFR(x:treal):treal;
var
    k:word;

```

```

ff: treal;
cc2: plist;
a1, b1: treal;
begin
a1:=(2*pi-pi/m)/(p.max-p.mix);b1:= -pi-a1*p.mix;x:=a1*x+b1;
cc2:=cc0^.u;
ff:=cc0^.x/2;
for k:=1 to m-1 do begin
ff:=ff+cc2^.x*cos(k*x)+cc2^.fdex*sin(k*x);
cc2:=cc2^.u;
end;
ffr:=ff+cc2^.x*cos(m*x);

```



```
end; {Functia FFT}
```

Continutul fisierului 12.pas :

```
function readkey:string; {asteapta caracter}
var
  b:byte;
begin
  asm
    mov ah,07h
    int 21h
    mov b,a1
  end;
  readkey:=chr(b);
  if b=0 then begin
    asm
      mov ah,07h
      int 21h
      mov b,a1
    end;
    readkey[2]:=chr(b) ;
    readkey[0]:=chr(2);
  end;
end; {Asteapta caracter.Se comporta asemanator cu 'readkey' din unitul CRT}
function inkey:string; {cauta caracter}
label 12;
var b,b1,b2:byte;
begin
  asm
    mov ah,06h
    mov d1.0ffh
    int 21h
    mov b1,a1
    lahf
    and ah,40h
```

```

mov b,ah
jne I2
cmp b1,0
jne 12
mov ah,06h
mov d1,0ffh
int 21h
mov b2 , a1
12:
end;
if b=0 then begin
  inkey:=chr(b1);
  if b1=0 then inkey:=chr(b1)+chr(b2);
end else inkey:=' ' ;
end; {Cauta caracter in buffer.Daca nu gaseste intoarce sir vid}
function HaltSignaled:boolean;
var
  v:byte;
begin
  asm
  mov ah,12h
  int 16
  and ah,0Fh
  mov v, ah
end;
HaltSignaled:=v=15;
end; {procedura de iesire din Programul Principal rapida cu dealocarile de
  memorie corespunzatoare si salvarea datelor}
function ExitSignaled:boolean;
var
  v: byte;
begin
  asm
  mov ah,12h
  int 16h

```

```

    and ah, 0Ah
mov v, ah
end;
ExitSignaled:=v=10;
end;{procedura de iesire din Bucla cu Nivel de inglobare dinamic de ordinul 2 rapida cu
dealocarile de memorie corespunzatoare si salvarea datelor}
function BreakSignaled:boolean;
var
    v:= byte;
begin
    asm
    mov ah,12h
    int 16h
        and a1,03h
        mov v, a1
    ·end;
    BreakSignaled:=v=3;
end;{procedura de iesire din procedura curenta dealocarile de memorie
    corespunzatoare si salvarea datelor}
procedure wait(var zona:byte;msh ,msl:word);
{dupa msh*ffffh+msl microsec zona=128 altfel zona=0}
var p:pointer;
begin
    zona:=0;
    p:=@((@zona)^);
    asm
        mov cx,msh
        mov dx,mpl
        les bx,p
        mov al,0
        mov ah,83h
        int 15h
    end;
end;{genereaza o comanda de timp microprocesorului care va pune 128 la adresa zona dupa
expirarea timpului intre timp executindu-se alte instructiuni}

```

```

function citstr:string;
const
  stai=4;
var
  cstr, c1:=string;
  zo:byte;
  b:boolean;
begin
  cstr:= '
  outtext(' '); c1:= ' ';
  wait(zo,stai,0);
  b:=true;
  repeat
    c1:=inkey;
  if (haltsignaled) or (exitsignaled) or (breaksignaled) then exit;
  if c1 <> ' ' then begin
    if (ord(c1[1])=8) and (ord(cstr[0])>0) then begin
      bar (getx-16, gety-1, getx, gety+7);
      moveto(getx-16, gety);
      if ord(cstr[0])>0 then delete (cstr, ord(cstr[0]),1);
      outtext(' ');
    end;
    if not(ord(c1[1]) in [0,13,8]) then begin
      cstr:=cstr+c1;
      bar( getx-8 ,gety-1, getx,gety+7);
      moveto (getx-8,gety);
      outtext(c1);
      outtext(' ');
    end;
  end;
  if zo=128 then begin
    wait (zo,stai,0);
    if b then bar(getx-8,gety-1,getx,qety+7)
  else begin
    moveto(getx-8,gety);

```

```

    outtext('_ ');
end;
b:=not(b);
end;
until ord(c1[1])=13;
citstr:=cstr;
end; {citeste in mod grafic cu simularea cursorului din mod test un string}
function citre:treal;
var
ix, px ,py:integer;
scc: string;
scre:treal;
begin
px:=getx;
py:=gety;
repeat
scc:=citstr;
if (haltsignaled) or (exitsignaled) or (breaksignaled) then exit;
val(scc,scre,ix);
if ix<>0 then bar(px,py-1,px+8*ord(scc[0]),py+7);
moveto (px,py);
until ix=0;
citre:=scre;
end; {citeste in mod grafic cu simularea cursorului din mod test un treal}
function citint:integer;.
var
scint,code:integer;
px, py:integer;
scc:string;
begin
px:=getx; py:=gety;
repeat
scc:=cistr;
if (haltsignaled) or (exitsignaled) or (breaksignaled) then exit;
val(scc,scint,code);

```

```
    if code <>0 then bar (px,py-1,px+8*ord(scc[0]),py+7);
    moveto (px,py);
until code=0;
CitInt:=scint;
end; {citeste in mod grafic cu simularea cursorului din mod test un integer}
```

```

Seeii de tim . Pr nozã: AI ritmi. i im lementare.
procedure Mesaj (ii:integer;pw:pointer;me:string;poz:integer);
var
  pe: ^treal;
  ps:^string;
begin
  if ii=5 then begin me:=readkey;exit;end;
  pe:= pw; ps:=pw;
  setfillstyle(solidfill,9); bar(poz-1, getmaxy-14, getmaxx-1, getmaxy-1);
  setusercharsize(1,8,1,8); setcolor(11); settextstyle(0,0,1);
  moveto(poz, getmaxy-10);
  if ii in [0,1] then outtextxy(poz, getmaxy-10,me);
  moveto(poz+8*ord(me[0]), getmaxy-10);
  if ii=2 then ps^:=citstr;
  if ii=3 then pe^:=citre;
end; {gestioneaza mesajele si oeratiile I/O care se transmit pe ultima linie}
//////////////////////////pag70

```

3.4. Aplicații

Sursa Datelor

Institutul Național de Meteorologie și Hidrologie

Precizări Metodologice

Cercetarea proceselor și fenomenelor fizice din atmosfera terestră, care determină starea timpului și clima, are la bază măsurătorile și observațiile meteorologice ce se efectuează permanent sau periodic în diferite puncte de pe suprafața globului terestru și la diferite înălțimi în atmosfera liberă. Programul unitar al măsurătorilor și observațiilor precum și alte activități meteorologice se asigură prin colaborarea internațională în cadrul Organizației Meteorologice Mondiale.

Majoritatea observațiilor, măsurătorilor și determinărilor meteorologice de la stații se efectuează pe platforma meteorologică situată pe un teren deschis tipic pentru regiunea respectivă, cu dimensiunea standard 26×26 m.

Pentru cercetarea proceselor și fenomenelor atmosferice și pentru calculul diferiților parametrii meteorologici și climatici, este necesară raportarea observațiilor și măsurătorilor atât în spațiu cât și în timp.

Temperatura aerului se măsoară cu instrumente cu citire directă, psihometrul cu ventilație artificială, și cu aparate înregistratoare (termografe).

Măsurarea cantităților de apă ce provin din precipitații atmosferice sau care se depun din alți hidrometeori, se efectuează cu ajutorul pluviometrului, iar înregistrarea continuă a precipitațiilor (lichide) se face cu pluviograful.

Tabelul 1 : Temperatura aerului : media lunară în perioada 1991-1993
înregistrate la stația Târgu-Mureș

	1991	1992	1993
Ianuarie	-1.5	-5.7	-4.2
Februarie	-3.4	-3.1	-6.5
Martie	6.8	3.9	1.1
Aprilie	9.3	10.7	9.0
Mai	12.4	14.2	16.8
Iunie	18.4	18.0	18.3
Iulie	20.5	19.8	18.8
August	18.2	22.6	19.4
Septembrie	14.6	13.6	13.1
Octombrie	8.9	9.9	11.0
Noiembrie	4.1	4.3	-0.3
Decembrie	-3.6	-3.6	1.4

Tabelul 2 : Temperatura aerului : media lunară în perioada 1991-1993
înregistrate la stația Sibiu

	1991	1992	1993
Ianuarie	-2.6	-5.0	-3.4
Februarie	-3.8	-2.9	-6.1
Martie	5.4	3.7	1.3

Aprilie	8.2	10.3	8.7
Mai	11.3	13.1	15.7
Iunie	17.9	17.1	17.7
Iulie	19.6	19.2	18.8
August	17.5	21.9	19.0
Septembrie	14.2	13.5	13.0
Octombrie	8.5	9.7	10.8
Noiembrie	4.5	4.8	-0.5
Decembrie	-4.8	-3.9	1.5

Tabelul 3 : Temperatura aerului : media lunară în perioada 1991-1993
înregistrate la stația Vârfu Omu

	1991	1992	1993
Ianuarie	-9.8	-9.8	-10.3
Februarie	-12.6	-13.1	-12.2
Martie	-5.5	-9.3	-9.1
Aprilie	-5.9	-3.8	-4.8
Mai	-2.5	-1.0	0.9
Iunie	4.7	3.4	3.4
Iulie	6.7	5.1	5.2
August	4.8	9.4	6.8
Septembrie	3.1	1.3	2.9
Octombrie	-2.2	-1.3	2.7
Noiembrie	-3.6	-6.3	-6.4
Decembrie	-11.6	-9.8	-7.0

Tabelul 4 : Temperatura aerului : media lunară în perioada 1991-1993
înregistrate la stația Cluj-Napoca

	1991	1992	1993
--	------	------	------

Ianuarie	-2.2	-4.9	-3.3
Februarie	-2.8	-2.6	-5.5
Martie	6.0	3.6	1.4
Aprilie	8.3	10.0	8.3
Mai	11.2	13.5	16.3
Iunie	17.1	17.4	17.6
Iulie	20.0	19.2	18.0
August	17.6	22.2	18.9
Septembrie	14.4	13.4	12.6
Octombrie	8.4	9.3	10.7
Noiembrie	4.1	4.2	-0.7
Decembrie	-4.8	-3.5	1.2

Tabelul 5 : Precipitații atmosferice : cantități lunare în perioada 1991-1993
înregistrate în bazinul Mureș la stația Târgu-Mureș

	1991	1992	1993
Ianuarie	10.5	19.9	13.7
Februarie	12.2	8.0	19.2
Martie	7.4	5.8	47.5
Aprilie	16.8	25.4	56.9
Mai	100.3	40.3	18.2
Iunie	100.0	161.0	58.9
Iulie	102.7	49.4	92.6
August	55.6	19.0	53.9
Septembrie	36.7	55.5	86.4
Octombrie	60.4	56.8	23.9
Noiembrie	17.1	33.4	38.4
Decembrie	18.8	7.9	54.0

Tabelul 6 : Precipitații atmosferice : cantități lunare în perioada 1991-1993

înregistrate în bazinul Sibiu la stația Sibiu

	1991	1992	1993
Ianuarie	13.6	15.8	13.2
Februarie	23.5	20.9	32.3
Martie	14.0	5.0	56.7
Aprilie	24.5	62.2	53.5
Mai	145.8	66.0	51.2
Iunie	120.7	90.8	53.9
Iulie	225.5	80.4	89.8
August	49.2	44.4	113.1
Septembrie	40.1	32.6	86.9
Octombrie	35.3	57.5	19.1
Noiembrie	59.7	20.9	34.4
Decembrie	18.4	21.6	35.4

Tabelul 7 : Precipitații atmosferice : cantități lunare în perioada 1991-1993
înregistrate la stația Vârfu Omu

	1991	1992	1993
Ianuarie	25.8	15.4	13.2
Februarie	20.0	19.4	29.3
Martie	19.0	22.6	63.0
Aprilie	84.4	22.2	62.1
Mai	126.4	75.3	97.9
Iunie	230.1	168.4	107.7
Iulie	222.5	99.6	70.9
August	87.5	24.4	133.9
Septembrie	41.7	26.4	76.1
Octombrie	112.5	81.9	24.5
Noiembrie	22.8	54.5	80.9
Decembrie	26.7	37.3	75.7

Tabelul 8 : Precipitații atmosferice : cantități lunare în perioada 1991-1993

înregistrate în bazinul Cluj la stația Cluj-Napoca

	1991	1992	1993
Ianuarie	7.2	24.9	15.2
Februarie	9.8	11.4	21.9
Martie	7.4	4.9	53.9
Aprilie	24.5	33.9	65.3
Mai	152.5	28.8	30.0
Iunie	70.5	125.1	64.2
Iulie	108.2	76.3	74.7
August	85.7	36.4	55.5
Septembrie	41.0	53.7	74.3
Octombrie	74.7	65.6	16.5
Noiembrie	42.0	39.8	52.3
Decembrie	24.6	14.1	43.3

Funcția de trend rezultată este

$$T(t) = -4,20238 + 8,56241 \times 10^{-2} \times t$$

unde

$T(t)$ este temperatura medie în grade celsius

și

t este numărul lunii în raport cu luna ianuarie 1991 considerată ca luna întâi.

Rezultatul eliminării acestui trend se află calculat în fișierul temp2omu.pas

iar diagrama de autocorelație se prezintă astfel :

1	1.000	13	0.9357	25	0.8776
2	0.7636	14	0.7924	26	0.7891
3	0.4461	15	0.4367	27	0.4527
4	-0.0180	16	-0.1260	28	-0.1857
5	-0.4623	17	-0.5409	29	-0.6772
6	-0.7770	18	-0.8207	30	-0.8989
7	-0.9304	19	-0.9250	31	-0.9127
8	-0.7909	20	-0.7674	32	-0.7105
9	-0.4582	21	-0.4507	33	-0.9204

10	0.0300	22	0.0900	34	-0.1635
11	0.4512	23	0.4849	35	1.0000
12	0.7906	24	0.8322		

Seria de timp a temperaturilor medii înregistrate pe Vârful Omu este periodică, cu perioada de 12 luni.

Aplicația 4.

Tabelul 4 : Temperatura aerului : media lunară în perioada 1991-1993
înregistrate la stația Cluj-Napoca

	1991	1992	1993
Ianuarie	-2.2	-4.9	-3.3
Februarie	-2.8	-2.6	-5.5
Martie	6.0	3.6	1.4
Aprilie	8.3	10.0	8.3
Mai	11.2	13.5	16.3
Iunie	17.1	17.4	17.6
Iulie	20.0	19.2	18.0
August	17.6	22.2	18.9
Septembrie	14.4	13.4	12.6
Octombrie	8.4	9.3	10.7
Noiembrie	4.1	4.2	-0.7
Decembrie	-4.8	-3.5	1.2

Funcția de trend rezultată este

$$T(t) = 7,21047 + 5,25868 \times 10^{-2} \times t$$

unde

$T(t)$ este temperatura medie în grade celsius

și

t este numărul lunii în raport cu luna ianuarie 1991 considerată ca luna întâi.

Rezultatul eliminării acestui trend se află calculat în fișierul temp2clu.pas

iar diagrama de autocorelație se prezintă astfel :

1	1.000	13	0.9620	25	0.9235
2	0.8132	14	0.8241	26	0.8022
3	0.4362	15	0.4418	27	0.2911
4	-0.0461	16	-0.0952	28	-0.3920
5	-0.5055	17	-0.5135	29	-0.8054
6	-0.8287	18	-0.8179	30	-0.9058
7	-0.9601	19	-0.9502	31	-0.9242
8	-0.8125	20	-0.8357	32	-0.8728
9	-0.4481	21	-0.4779	33	-0.9481
10	0.0652	22	0.0500	34	-0.3010
11	0.4970	23	0.5202	35	-1.0000
12	0.8298	24	0.8275		

Seria de timp a temperaturilor medii înregistrate în municipiul Cluj-Napoca este periodică, cu perioada de 12 luni.

Aplicația 5.

Tabelul 5 : Precipitații atmosferice : cantități lunare în perioada 1991-1993
înregistrate în bazinul Mureș la stația Târgu-Mureș

	1991	1992	1993
Ianuarie	10.5	19.9	13.7
Februarie	12.2	8.0	19.2
Martie	7.4	5.8	47.5
Aprilie	16.8	25.4	56.9
Mai	100.3	40.3	18.2
Iunie	100.0	161.0	58.9
Iulie	102.7	49.4	92.6
August	55.6	19.0	53.9
Septembrie	36.7	55.5	86.4
Octombrie	60.4	56.8	23.9
Noiembrie	17.1	33.4	38.4

Decembrie 18.8 7.9 54.0

Funcția de trend rezultată este

$$P(t) = 38,66936 + 2,88890 \times 10^{-1} \times t$$

unde

$P(t)$ este cantitatea de precipitații medie, în l/m²

și

t este numărul lunii în raport cu luna ianuarie 1991 considerată ca luna întâi.

Rezultatul eliminării acestui trend se află calculat în fișierul prec2tgm.pas iar diagrama de autocorelație se prezintă astfel :

1	1.000	13	0.4657	25	0.2602
2	0.3426	14	0.4894	26	0.5309
3	0.0129	15	0.0400	27	0.6061
4	-0.1003	16	-0.0241	28	-0.0540
5	-0.2394	17	-0.2882	29	-0.1540
6	-0.2217	18	-0.3334	30	-0.7309
7	-0.4324	19	-0.3543	31	-0.4111
8	-0.4982	20	-0.6021	32	0.0719
9	-0.2099	21	-0.2654	33	0.0782
10	-0.0289	22	0.0740	34	-0.7151
11	0.1576	23	0.1346	35	1.0000
12	0.2875	24	0.3925		

Aplicația 6.

Tabelul 6 : Precipitații atmosferice : cantități lunare în perioada 1991-1993 înregistrate în bazinul Sibiu la stația Sibiu

	1991	1992	1993
Ianuarie	13.6	15.8	13.2

Februarie	23.5	20.9	32.3
Martie	14.0	5.0	56.7
Aprilie	24.5	62.2	53.5
Mai	145.8	66.0	51.2
Iunie	120.7	90.8	53.9
Iulie	225.5	80.4	89.8
August	49.2	44.4	113.1
Septembrie	40.1	32.6	86.9
Octombrie	35.3	57.5	19.1
Noiembrie	59.7	20.9	34.4
Decembrie	18.4	21.6	35.4

Funcția de trend rezultată este

$$P(t) = 58,60492 - 2,73088 \times 10^{-1} \times t$$

unde

$P(t)$ este cantitatea de precipitații medie, în l/m²

și

t este numărul lunii în raport cu luna ianuarie 1991 considerată ca luna întâi.

Rezultatul eliminării acestui trend se află calculat în fișierul prec2sib.pas iar diagrama de autocorelație se prezintă astfel :

1	1.000	13	0.4855	25	0.3984
2	0.4369	14	0.3779	26	0.7050
3	0.1982	15	0.1368	27	0.7018
4	-0.2547	16	-0.0543	28	0.0115
5	-0.2962	17	-0.3793	29	-0.3721
6	-0.3185	18	-0.4671	30	-0.7625
7	-0.3613	19	-0.6621	31	-0.5882
8	-0.4965	20	-0.4083	32	-0.3205
9	-0.4429	21	-0.1069	33	-0.6678
10	-0.0456	22	0.1119	34	0.5141
11	0.1441	23	0.2362	35	1.0000
12	0.5425	24	0.1295		

Aplicația 7.

Tabelul 7 : Precipitații atmosferice : cantități lunare în perioada 1991-1993
înregistrate la stația Vârfu Omu

	1991	1992	1993
Ianuarie	25.8	15.4	13.2
Februarie	20.0	19.4	29.3
Martie	19.0	22.6	63.0
Aprilie	84.4	22.2	62.1
Mai	126.4	75.3	97.9
Iunie	230.1	168.4	107.7
Iulie	222.5	99.6	70.9
August	87.5	24.4	133.9
Septembrie	41.7	26.4	76.1
Octombrie	112.5	81.9	24.5
Noiembrie	22.8	54.5	80.9
Decembrie	26.7	37.3	75.7

Funcția de trend rezultată este

$$P(t) = 74,91428 - 2,92664 \times 10^{-1} \times t$$

unde

$P(t)$ este cantitatea de precipitații medie, în l/m²

și

t este numărul lunii în raport cu luna ianuarie 1991 considerată ca luna întâi.

Rezultatul eliminării acestui trend se află calculat în fișierul prec2omu.pas
iar diagrama de autocorelație se prezintă astfel :

1	1.000	13	0.6019	25	0.3725
2	0.5483	14	0.3207	26	0.6347
3	0.0792	15	0.0544	27	0.3817

4	-0.0890	16	-0.2615	28	-0.3760
5	-0.2310	17	-0.4169	29	-0.5578
6	-0.3463	18	-0.3230	30	-0.3885
7	-0.4153	19	-0.3715	31	-0.2128
8	-0.4853	20	-0.4386	32	0.0436
9	-0.2510	21	-0.2865	33	0.3213
10	-0.2001	22	-0.1693	34	-0.9834
11	0.0456	23	0.3573	35	1.0000
12	0.4398	24	0.5139		

Aplicația 8.

Tabelul 8 : Precipitații atmosferice : cantități lunare în perioada 1991-1993
înregistrate în bazinul Cluj la stația Cluj-Napoca

	1991	1992	1993
Ianuarie	7.2	24.9	15.2
Februarie	9.8	11.4	21.9
Martie	7.4	4.9	53.9
Aprilie	24.5	33.9	65.3
Mai	152.5	28.8	30.0
Iunie	70.5	125.1	64.2
Iulie	108.2	76.3	74.7
August	85.7	36.4	55.5
Septembrie	41.0	53.7	74.3
Octombrie	74.7	65.6	16.5
Noiembrie	42.0	39.8	52.3
Decembrie	24.6	14.1	43.3

Funcția de trend rezultată este

$$P(t) = - 2,30785 \times 10^{-1} + 2,02059 \times 10^{-3} \times t$$

unde

$P(t)$ este cantitatea de precipitații medie, în l/m²

și

t este numărul lunii în raport cu luna ianuarie 1991 considerată ca luna întâi.

Rezultatul eliminării acestui trend se află calculat în fișierul prec2clu.pas

iar diagrama de autocorelație se prezintă astfel :

1	1.000	13	0.3935	25	0.1084
2	0.3148	14	0.5959	26	0.4745
3	0.1549	15	0.2485	27	0.2591
4	-0.0508	16	-0.1543	28	-0.2437
5	-0.3727	17	-0.4122	29	0.0886
6	-0.2665	18	-0.2922	30	-0.9034
7	-0.5031	19	-0.4881	31	-0.2045
8	-0.5732	20	-0.5243	32	-0.1145
9	-0.2169	21	-0.1648	33	-0.2262
10	-0.0644	22	0.0116	34	0.7416
11	0.1374	23	0.3869	35	-1.0000
12	0.4100	24	0.6598		

Aplicația 9.

Tabelul 9. Indicii de criminalitate în perioada 1985-1993 în județul Cluj

1985	412
1986	425
1987	385
1988	129
1989	255
1990	160
1991	263

1992	303
1993	366

Funcția de trend rezultată este

$$T(t) = 1,02601 \times 10^3 - 9,29274 \times 10^{-1} \times t$$

unde

$T(t)$ este cantitatea medie de gaz metan livrată în m³/loc

și

t este luna corespunzătoare în raport cu Ianuarie 1957 luată drept referință.

Rezultatul eliminării acestui trend se află calculat în fișierul gaz2metan.pas

iar diagrama de autocorelație se prezintă astfel :

1	1.0000	25	0.9752	49	0.9627	73	0.9603
2	0.8001	26	0.7738	50	0.7769	74	0.6862
3	0.4475	27	0.4263	51	0.4009	75	0.1713
4	-0.0298	28	-0.0713	52	-0.1182	76	-0.5114
5	-0.4934	29	-0.5083	53	-0.5507	77	-0.7936
6	-0.7932	30	-0.8111	54	-0.8444	78	-0.9216
7	-0.9177	31	-0.9167	55	-0.9199	79	-0.9781
8	-0.7841	32	-0.7620	56	-0.7614	80	-0.9743
9	-0.4569	33	-0.4330	57	-0.4047	81	-0.9613
10	0.0326	34	0.0501	58	0.1221	82	-0.9999
11	0.4700	35	0.4657	59	0.5223	83	-1.0000
12	0.8086	36	0.7981	60	0.8461		
13	0.9744	37	0.9641	61	0.9695		
14	0.7792	38	0.7790	62	0.7446		
15	0.4288	39	0.4222	63	0.3299		
16	-0.0544	40	-0.0663	64	-0.1769		
17	-0.5060	41	-0.4991	65	-0.5885		
18	-0.8022	42	-0.8088	66	-0.8569		
19	-0.9187	43	-0.9212	67	-0.9371		
20	-0.7756	44	-0.7774	68	-0.7707		
21	-0.4528	45	-0.4560	69	-0.3649		
22	0.0387	46	0.0692	70	0.2098		
23	0.4766	47	0.4796	71	0.5787		

Seria de timp a livrărilor de gaz metan este periodică, cu perioada de 12 luni.

Bibliografie

- [1]: Bloomfield,P : Fourier analysis of time series:an introduction, New York, J.Wiley,1976.
- [2]: Brigham,E.O. : The fast fourier transform,New Jersey, Englewood Cliffs, Prentice-Haal, 1972.
- [3]: Koopmans, L.H. : The spectral analysis of time series, New York, Academic Press, 1974.
- [4]: Lighthill,M.J. : Fourier analysis and generalized functions, Cambridge, Univ.Press, 1964.
- [5]: Box, G.E.P., Hunter, W.G., Hunter, J.I., Statistics for Experimenters to introduction to design, Data Analysis and Model Building
- [6]: Teodorescu,D. : Modele stohastice de optimizare, EA București,1982.
- [7]: Giurgiu,N., Dăscălescu,M. : Culegere de Programe pentru rezolvarea problemelor de regresie, corelație și tendința Seriilor Dinamice, ISE Cluj, uzul studentilor,1983.
- [8]: Urseanu,V. : Elemente de statistică matematică și aplicațiile ei, ES Bucuresti,1966.
- [9]: Mihoc,G., Urseanu V. : Modele de analiză statistică, ESP București,1982.
- [10]: Krilov,A.N. : Lecții de calcule prin aproximare, ediția a 6-a, Moscova,1954.
- [11]: Frențiu,M. : Geologie Matematică, UBB, Fac de MF, pentru uzul studentilor, Cluj, 1987.
- [12]: Cristea, A.L., Postelnicu, T., Ursianu Emiliana, Asupra relației doză-efect cu aplicații în virusologie, "Revue de Virusologie", București, 1980.

- [13]: M. Tiron : Analiza preciziei de estimare a funcțiilor aleatoare,ET București,1981.
- [14]: Reismann,G. : Die Kolokation, Vermessungstechnik, No 1,1979.
- [15]: Moritz,H. : Advanced physical geodesy,Herbert Wichman Verlag,1980.
- [16]: Gârlașu,Șt. : PrelucrareaÎn timp real a semnalelor fizice, Craiova, Scrisul Românesc, 1978.
- [17]: Bury,K.V. : Statistical ModelsÎn Applied Science John Wiley, New York,1975.
- [18]: Vodă,Gh.,V. : Gândirea Statistică-un mod de gândire al viitorului, E Albatros,1977.
- [19]: Mihoc,Gh., Bergthaller,C., Urseanu,V. : Procese stohastice, București, ESP, 1978.
- [20]:Yule, G.U., Kendall, M.G. : Introducere în teoria statisticii, București, ES, 1969.
- [21]: Frențiu,M. , Pârv,B. : Elaborarea Programelor : Metode și tehnici moderne, Editura Promedia, Cluj 1994.
- [22]: Frențiu,M., Boian,F. : Bazele informaticii. Limbajul Pascal, Litografia Univ. "Babeș-Bolyai", Cluj Napoca, 1992.
- [23]: Knuth,D. : Tratat de programarea calculatoarelor. Algoritmi fundamentali, Editura Tehnică, București, 1974.
- [24]: Pârv,B., Vancea,A. : Fundamentele limbajelor de programare, Litografia Univ. " Babeș-Bolyai ", Cluj-Napoca, 1992.
- [25]: ***** : Dicționar de informatică, Editura științifică și enciclopedică, Bucureși, 1981.
- [26]: Bjerhammar, Arne : Theory of errors on generalized matrix inverses, Amsterdam-London-New York, Elvister, 1973.
- [27]: Brandin,N.V., ș.a. : Osnovî eksperimentalnoi

- kosmiceskoi balistiki, Moscova, Maşinostroenie, 1974.
- [28]: Lapa,G.V. : Matematiceskie osnovî kibernetiki, Kiev, Vişsaia şkola, 1977.
- [29]: Tiron,M. : Teoria erorior şi metoda celor mai mici pătrate, Bucureşti, Editura Tehnică, 1972.
- [30]: Terişco,M. , Stoica,P. , Popescu,T. : Modelarea şi predicţia seriilor de timp, Editura Academiei RSR, Bucureşti, 1985.
- [31]: Terişco,M. , Stoica,P. : Identificarea şi estimarea parametrilor sistemelor, Editura Academiei RSR, Bucureşti, 1980.
- [32]: Stoica,P. , T. Söderström : A metod for the identification of linear systems using the generalized least-squares principle, IEEE Tans. Automat. Contr. , AC-22, 631-634 (1977).
- [33]: Markel,J.D. , Gray,A.H. : Linear prediction of speech, Springer Verlag, Berlin, 1976.
- [34]: Anderson,T.W. : The statistical Analysis of Time Series, John Wiley & Sons, New York, 1976.
- [35]: Miller,R.B. : Intermediate Business Statistics, Analysis of variance. Regresion and Time Series, Holt, Rinhart and Winston, 1977.
- [36]: Institutul de statistică : Anuarul de statistică al României, anul 1993, Editura Academiei, 1994.

Cuprins

1. Studiul dependenței de timp	pag. 1
1.1. Serii de timp	pag. 1
1.2. Teoria estimației	pag. 4
1.3. Estimarea parametrilor prin metoda celor mai mici pătrate	pag. 5
1.4. Legătura metodei celor mai mici pătrate cu celelalte metode de estimare	pag. 8
1.5. Analiza seriilor de timp prin metoda corelației	pag. 8
1.6. Corelația simplă. Regresia liniară. Coeficientul de regresie	pag. 9
1.7. Coeficientul de corelație. Covarianța	pag. 11
1.8. Raportul de corelație	pag. 13
1.9. Corelația rangurilor	pag. 15
1.A. Autocorelația	pag. 19
1.B. Analiza prin periodograme	pag. 20
1.C. Corelograma	pag. 22
2. Analiza numerică a seriilor de timp	
2.1. Descompunerea seriilor de timp	pag. 23
2.2. Tendința de evoluție. Trend	pag. 23
2.2.1. Regresie liniară	pag. 23
2.2.2. Regresie prin serie parțială Taylor	pag. 23
2.2.3. Regresie parabolică. Regresie hiperbolică	pag. 25
2.2.4. Cazul special al trendului polinomial	pag. 25
2.2.5. Determinarea ordinului tendinței polinomiale	pag. 26
2.2.6. Regresie Neliniară. Metode iterative	pag. 27
2.2.7. Regresie exponențială și regresie logaritmică	pag. 29
2.3. Sezonalitate. Periodicitate	
2.3.1. Seria de sezonality	pag. 30
2.3.2. Construcția seriei de sezonality prin serii armonice	pag. 31
2.3.3. Folosirea FFT în construcția seriei de sezonality	pag. 31
2.3.4. Cazul perioadei necunoscute. Aflarea perioadei seriei	pag. 35
2.3.5. Întocmirea și extragerea informațiilor din periodogramă	pag. 35
2.4. Evoluție în timp. Prognoză	
2.4.1. Tendința de stabilizare	pag. 36
2.4.2. Analiza Fourier. Reprezentarea Fourier a unui șir de date	pag. 37
2.4.3. Serii periodice. Reprezentarea Fourier a unui șir periodic de date	pag. 40
2.4.4. Estimarea parametrilor tendinței ciclice	pag. 41

2.4.5. Modelul armonic al tendinței sezoniere.....	pag. 42
2.4.6. Teorema Parseval. Spectrograma Shuster	pag. 43
2.4.7. Metoda Whittaker-Robinson	pag. 43
3. Implementarea metodelor de analiza seriilor de timp	pag. 44
3.1. Algoritmii implementați	pag. 44
3.2. Textele sursă ale programului	pag. 65
3.3. Memorarea, interfața, nucleul grafic, portabilitatea	pag. 91
3.4. Aplicații	pag. 95
4. Bibliografie.....	pag. 117
5. Cuprins	pag. 119