6.4. Transformata de înălbire

6.4.1. Introducere

În problemele reale de clasificare apare adeseori necesitatea diagonalizării simultane a două matrici diferite de corelaţie sau covarianţă.

Exemplul 6.4: O astfel de situaţie poate să apară, de exemplu, în aplicaţiile practice de clasificare unde avem două clase pentru eşantionul de date disponibil. În cazul în care trăsăturile vectorului aleator multidimensional de trăsături, x, sunt corelate este de dorit aplicarea unei transformări care să ducă datele din spaţiul iniţial al trăsăturilor într-un nou spaţiu, în care trăsăturile să fie, de exemplu, necorelate. Avantajul unei asemenea operaţii îl constituie, aşa după cum ştim, posibilitatea de a obţine performanţe de clasificare mai ridicate inclusiv folosind un clasificator de tipul celui de minimă distanţă. Obiectivul, în acest context, îl constituie găsirea acelei transformate care, aplicată tuturor datelor eşantionului (populaţiei), să ne conducă la o diagonalizare simultană a celor două matrici de covarianţă ce caracterizează, corespunzător, cele două clase. Un exemplu în acest sens este şi problema tratată în ***Subcapitolul** 7.2.4* în care avem un vector aleator bidimensional de trăsături, $x = [puls, presiune\ sistolic\breve{a}]^T$ şi două clase: clasa $\{obosit\}$ şi, respectiv, clasa $\{odihnit\}$.

În cele ce urmează discutăm cazul unui eşantion de date statistice caracterizat de două clase distincte.

Din **Figura 6.7,** în care se prezintă distribuţia celor două clase – prin intermediul a două elipse de concentrare –, se observă că este imposibil să obţinem diagonalizarea simultană a matricelor de corelaţie numai prin acţiuni simple de rotire a sistemului iniţial de coordonate (ne reamintim că un astfel de efect se poate obţine prin aplicarea, în mod direct asupra datelor eşantionului, a unei transformate unitare).

În cazul cel mai general, prin intermediul transformatei unitare putem diagonaliza, cel mai adesea, cel mult o singură matrice de covarianţă din cele două, cea de-a doua matrice de covarianţă nou obţinută nefiind şi ea una tot diagonală.

Soluţia, în acest caz, o reprezintă în fapt, aşa după cum vom vedea în continuare, aplicarea unei anumite *transformate* ce are ca efect <u>o combinaţie de rotiri şi de scalări ale sistemului iniţial de coordonate</u>. Pentru a determina această transformată particulară, avem nevoie să introducem, ca pas intermediar absolut necesar, definirea unei alte transformate, numită *de înălbire*, de care ne vom folosi în mod deosebit.

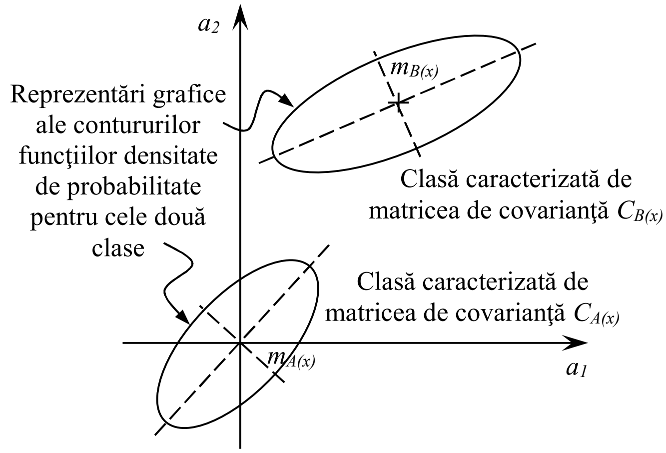

Figura 6.7. Exemplu de distribuţie a datelor în spaţiul iniţial al trăsăturilor (spaţiu caracterizat de existenţa a două clase).

6.4.2. Transformata de înălbire

Transformată de înălbire (sau „albire") a seturilor de date urmăreşte, în principal, următoarele două obiective:

(a) **Primul obiectiv** are în vedere decorelarea setului de date prin aplicarea unei transformate liniare (în particular, transformata unitară). Noua matrice de covarianţă rezultată după aplicarea acestei transformate va fi una diagonală.

(b) **Cel de-al doilea obiectiv** vizează o scalare, în mod corespunzător, a setului de date astfel încât varianţa fiecărei variabile aleatoare ce compune vectorul aleator de trăsături să fie egală cu unitatea. Noua matrice de covarianţă ce rezultă în urma aplicării acestei ultime transformări va fi, aşa după cum vom vedea, chiar matricea unitate.

Conform abordării prezentate în ***Subcapitolul 6.3.2*****,** după determinarea valorilor proprii, $\lambda_1, \lambda_2, \ldots \lambda_d$, şi, respectiv, a vectorilor proprii, $e_1, e_2, \ldots, e_d$, asociaţi acestor valori proprii, putem defini matricea vectorilor proprii astfel:

$$E = \begin{bmatrix} | & | & & | \\ e_1 & e_2 & \cdots & e_d \\ | & | & & | \end{bmatrix} \quad (6.107)$$

(a) Primul obiectiv îl vom atinge prin utilizarea transformatei $A = E^{*T}$:

$$y = E^{*T} x \qquad (6.108)$$

Matricea de covarianţă a noului vector aleator, y, devine una diagonală:

$$C_y = E^{*T} C_x E = \begin{bmatrix} - & e_1^{*T} & - \\ - & e_2^{*T} & - \\ & \vdots & \\ - & e_d^{*T} & - \end{bmatrix} C_x \begin{bmatrix} | & | & & | \\ e_1 & e_2 & \cdots & e_d \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_d \end{bmatrix} = \Lambda$$

(6.109)

În acest mod am obţinut, într-o primă fază, **decorelarea vectorului de trăsături**.

(b) Pentru atingerea celui de-al doilea obiectiv vom aplica – în cascadă cu transformata E^{*T} de mai sus – o nouă transformată, notată cu $\Lambda^{-½}$; această transformată va determina ca matricea de covarianţă a noului vector de variabile aleatoare să fie o matrice unitate.

Astfel, dacă matricea Λ este dată de relaţia (6.109) atunci vom avea:

$$\Lambda^{-1/2} = \begin{bmatrix} 1/\sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & 1/\sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sqrt{\lambda_d} \end{bmatrix} \qquad (6.110)$$

Prin aplicarea acestei transformate vectorului aleator y:

$$z = \Lambda^{-1/2} y \qquad (6.111)$$

vom obţine în spaţiul z următoarea matrice de covarianţă:

$$C_z = \Lambda^{-1/2} \cdot C_y \cdot \left(\Lambda^{-1/2}\right)^{*T} = \Lambda^{-1/2} \cdot \Lambda \cdot \Lambda^{-1/2} = I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \qquad (6.112)$$

În scrierea relaţiilor (6.112) s-a ţinut cont de o proprietate a valorilor proprii, şi anume aceea că: valoriile proprii sunt întotdeauna numere reale.

Concluzionând cele de mai sus, vom spune că ***transformata de înălbire*** este dată de o matrice A ce este egală cu:

$$A = \Lambda^{-1/2} E^{*T} \qquad (6.113)$$

unde notaţiile folosite sunt cele cu semnificaiile de mai sus.

În urma aplicării acestei transformate se obţine, simultan, atât decorelarea setului de date cât şi scalarea acestuia astfel încât, în final, matricea de covarianţă în spaţiul z să fie matricea unitate.

Observaţa 6.12: Transformata de înălbire *nu este o transformată ortogonală*; în concluzie, nu conservă distanţele:

$$\| z \|^2 \neq \| x \|^2 \quad (6.114)$$

Observaţa 6.13: După aplicarea transformatei de înălbire, **matricea** de covarianţă obţinută ($C_z = I$) **este invariantă la orice transformare ortonormală**. Pe baza acestei observaţii, în subcapitolul următor vom diagonaliza simultan matricile de covarianţă pentru două clase distincte în vederea decorelării ambelor clase.

Observaţa 6.14: Transformata de înălbire este o transformată generică care se poate aplica oricărei matrici A. În consecinţă, această transformată se poate aplica inclusiv matricelor de corelaţie precum şi celor de covarianţă.

6.5. Diagonalizări simultane ale matricilor de covarianţă aferente la două clase distincte

Pentru prezentarea acestei proceduri, de diagonalizare simultană a ambelor matrici, ne vom folosi de următoarele două matrici de covarianţă, notate în continuare cu $C_{A(x)}$ şi $C_{B(x)}$. Două exemple de conturuti caracteristice pentru aceste două matrici sunt reprezentate în **Figura 6.7**. Cu toate că în această reprezentare grafică discutăm cazul bidimensional, metoda de diagonalizare descrisă şi prezentată în formă analitică în paragrafele următoare este una generală, capabilă să lucreze cu vectori d-dimensionali.

Paşii de realizare ai procedurii de diagonalizare simultană sunt:

1. Găsirea vectorilor şi valorilor proprii pentru matricea $C_{A(x)}$, respectiv:

$$C_{A(x)} = E_A \Lambda_A E_A^{*T} \quad (6.115)$$

Reamintim că vectorii şi, respectiv, valorile proprii sunt elemente ale matricelor E_A şi, respectiv, Λ_A, din relaţia (6.115).

2. Aplicarea transformatei de înălbire pentru vectorul aleator x, de intrare:

$$T = \Lambda_A^{-1/2} E_A^{*T} \quad (6.116)$$

În aceste condiţii, noul vector, y, va fi calculat cu relaţia:

$$y = \Lambda_A^{-1/2} E_A^{*T} \cdot x \quad (6.117)$$

După cum s-a prezentat anterior, matricea $\Lambda_A^{-1/2}$ este inversa matricii diagonale $\Lambda_A^{1/2}$, ale cărei elemente de pe diagonala principală sunt date de radicalul valorilor proprii.

Aplicând transformata anterior prezentată, vom obţine pentru clasa A o matrice de covarianţă egală cu matricea unitate:

$$\begin{aligned} C_{A(y)} &= (\Lambda_A^{-1/2} E_A^{*T}) C_{A(x)} (\Lambda_A^{-1/2} E_A^{*T})^{*T} = \\ &= \Lambda_A^{-1/2} E_A^{*T} \cdot E_A \Lambda_A E_A^{*T} \cdot E_A \Lambda_A^{-1/2} = I \end{aligned} \quad (6.118)$$

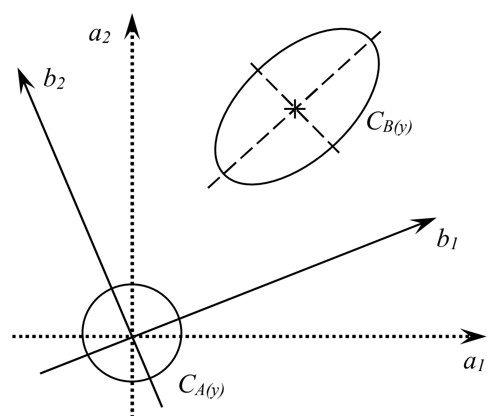

Figura 6.8. Distribuţiile claselor după aplicarea transformatei de înălbire

Prin aplicarea transformatei dată de relaţia (6.117), cunoscută sub numele de transformată de "înălbire", se obţine (atenţie! – doar pentru clasa A), simultan cu decorelarea trăsăturilor, şi o varianţă egală cu unitatea pentru toate trăsăturile. Reprezentarea grafică a unui contur al funcţiei densitate de probabilitate pentru clasa A este redat în **Figura 6.8** sub forma unui cerc (varianţe egale pe toate direcţiile), de rază unitară.

Pentru clasa B, în urma aplicării transformatei dată de relaţia (5.131), matricea de covarianţă, $C_{B(x)}$, devine:

$$C_{B(y)} = (\Lambda_A^{-1/2} E_A^{*T}) C_{B(x)} (\Lambda_A^{-1/2} E_A^{*T})^{*T} = \\ = \Lambda_A^{-1/2} E_A^{*T} C_{B(x)} E_A \Lambda_A^{-1/2} \quad (6.119)$$

Rezumând cele prezentate până aici, vom spune că efectul global al acestei prime transformări (de înălbire) aplicată tuturor datelor eşantionului rezidă, simultan:

- într-o rotire a sistemului iniţial de coordonate (aşa cum se prezintă şi în **Figura 6.8)** şi, respectiv,
- într-o scalare a datelor,

acestea având ca efect:

- pentru clasa A, diagonalizarea, într-o primă fază, şi, transformarea, în final, într-o matrice unitate, a matricei de covarianţă $C_x(A)$ iar,
- pentru clasa B, efectul este unul nedefinit (a se vedea şi reprezentarea grafică din **Figura 6.8)**.

Aceste efecte, diferenţiate pe clase, ale transformatei de înălbire aplicată eşantionului de date se datorează, bineînţeles, modului cum a fost construită aceasta, respectiv, utilizând valorile şi vectorii proprii ai clasei A; de aici, şi efectul ei specific obţinut doar asupra acestei clase.

3. Determinarea vectorilor şi valorilor proprii pentru matricea $C_{B(y)}$ în noul sistem de coordonate, y. Prin metodele cunoscute se determină matricea vectorilor proprii, E_B, şi matricea valorilor proprii, Λ_B, asociate clasei caracterizată de matricea de covarianţă $C_{B\,(y)}$.
4. Diagonalizarea matricei $C_{B\,(y)}$ este ultimul pas al procedurii. Acest ultim obiectiv se atinge prin aplicarea următoarei transformate vectorului, y, de intrare:

$$z = E_B^{*T} \cdot y \quad (6.120)$$

În acest caz, matricile de covarianţă corespunzătoare celor două clase devin:

$$C_{B(z)} = E_B^{*T} C_{B(y)} E_B = \Lambda_B \quad (6.121)$$

şi, respectiv,

$$C_{A(z)} = E_B^{*T} I\, E_B = I \quad (6.122)$$

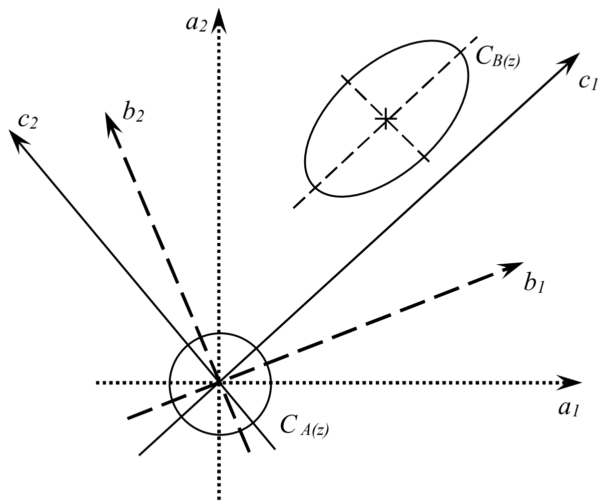

Figura 6.9. Rezultatul final al procesului de diagonalizare simultană

Prin această ultimă transformare sistemul de coordonate este rotit din nou, aşa după cum se poate observa şi pe exemplul din **Figura 6.9,** iar efectul final este acela de diagonalizare şi a matricei de covarianţă pentru cea de a doua clasă.

Pe baza paşilor descrişi anterior deducem uşor relaţia transformatei ce are ca efect transformarea setului original de date de o manieră similară ca cea prezentată mai sus. Astfel:

$$z = (E_{global})^{*T} \cdot x \qquad (6.123)$$

unde:

$$E_{global} = E_A \Lambda_A^{-1/2} E_B \qquad (6.124)$$

În spaţiul final de reprezentare a datelor, matricea de covarianţă $C_{A(x)}$ iniţială devine o matrice identitate, în timp ce matricea de covarianţă $C_{B(x)}$ iniţială devine o matrice diagonală, Λ_B, ambele caracterizând distribuţia elementelor claselor în noul plan, planul vectorului aleator z. Prin diagonalizarea simultană a matricelor de covarianţă a claselor, componentele vectorului aleator z au devenit necorelate pentru fiecare clasă în parte.

Aplicaţia 6.3: Utilizând aplicaţia din directorul „Diagonalizări simultane”, asociat acestui capitol – aplicaţie ce implementează algoritmul de diagonalizare simultană a matricilor de covarianţă ce caracterizează două clase distincte (algoritm prezentat în ***Subcapitolul 6.5***) – să se parcurgă următoarele cerinţe:

1. Încărcaţi setul de date „CL2clase.txt”. Porniţi procesul de diagonalizare simultană a matricilor de covarianţă ce caracterizează cele două distribuţii şi urmăriţi paşii prezentaţi anterior. Verificaţi, prin calcul direct, corectitudinea rezultatelor. Salvaţi setul de date final obţinut după aplicarea transformatei globale (6.124).
2. După parcurgerea Capitolului 7 al cărţii utilizaţi programul din directorul „Comparatie minDist-Mahalanobis-Bayes” pentru a analiza performanţele de clasifcare comparative pe cei trei clasificatori implementaţi pe seturile de date „CL2clase.txt” şi pe cel obţinut după aplicarea transformatei globale (6.124).

6.6. Transformata inferior-superior triunghiulară

Cele mai importante aplicaţii ale unei descompuneri (factorizări) inferior-superior triunghiulară constă în calculul determinantului şi al inversei unei matrici pătratice, respectiv, în rezolvarea sistemelor liniare de n ecuaţii cu n necunoscute. Ceea ce ne interesează însă pe noi este o altă utilitate a acestor descompuneri, utilitate ce rezidă din însăşi modul cum au fost definite acestea.

Astfel, într-unul din subcapitolele precedente s-a prezentat o tehnică de diagonalizare a matricilor de corelaţie şi de covarianţă cu ajutorul transformatei unitare. O altă tehnică de diagonalizare a acestor matrici este dată de metoda **descompunerii în matrici triunghiulare,** particularizată în cazul acestui capitol prin **descompunerea *Cholesky***. Această metodă de descompunere în matrici triunghiulare conduce către o transformată care poate fi interpretată, în funcţie de forma exactă a descompunerii, ca o operaţie de filtrare cauzală sau necauzală a secvenţei căreia i se aplică.

6.6.1. Descompunerea în matrici inferior-superior triunghiulare

O matrice pătratică $A \in R^{n \times n}$ nesingulară[8] admite o descompunere în matrici inferior-superior triunghiulare (descompunere *LU*[9]) dacă şi numai dacă toţi minorii săi principali, A_k (k=1,..., n-1) sunt nesingulari. În cazul în care o astfel de decompunere există (respectiv, matricea *A* îndeplineşte condiţiile de mai sus) atunci, descompunerea *LU* permite scrierea matricei *A* ca un produs dintre o matrice inferior triunghiulară *L* şi o matrice superior trunghiulară *U* astfel:

$$A = L \cdot U \quad (6.125)$$

În forma extinsă, relaţia de mai sus se scrie şi:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \underbrace{\begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}}_{L} \cdot \underbrace{\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}}_{U} \quad (6.126)$$

Factorizarea dată sub forma relaţiei (6.125), atunci când ea există, nu este unică, în literatura de specialitate fiind prezentate, printre alte descompuneri posibile, trei astfel de factorizări *LU* fundamentale, respectiv:

(a) factorizarea *Doolittle* – în care termenii diagonali ai matricei inferior triunghiulare, *L*, sunt toţi unitari (altfel spus, toate elementele de pe diagonala principală sunt egale cu 1, $l_{ii} = 1$, în timp ce toate elementele superioare diagonalei principale sunt egale cu zero); în acest caz, matricea *L* poartă numele de matrice inferior triunghiulară unitară;

(b) factorizarea *Crout* – în care toţi termenii de pe diagonala principală a matricei superior triunghiulare *U* sunt egali cu 1 ($u_{ii} = 1$); în acest caz matricea *U* se mai numeşte şi matrice superior triunghiulară unitară şi, ultima variantă,

(c) factorizarea *Cholesky*[10] – în care cele două matrici din relaţia (6.126) se află într-o relaţie de forma:

$$U = L^{*T} \quad (6.127)$$

În acest caz, $l_{ii} = u_{ii}$, pentru toţi i.

---

8 O matrice pătratică se numeşte nesingulară dacă determinantul ei este nenul.

9 În engleză, *lower-upper* decomposition.

10 Această ultimă descompunere a fost denumită după numele descoperitorului ei, respectiv, matematicianul francez *André-Louis Cholesky*.

Factorizarea *LU* este ***unică*** atunci când fie matricea *L*, fie matricea *U* din descompunerea *LU* sunt matrici inferior, respectiv, superior triunghiulare ***unitare***. În acest context, toate cele trei variante de descompunere prezentate mai sus sunt descompuneri unice.

Pe lângă descompunerea *LU* a unei matrici pătratice nesingulare există şi **descompunerea (factorizarea) *LDU*** a acesteia, dată de:

$$A = L \cdot D \cdot U \qquad (6.128)$$

unde: *L* este o matrice inferior triunghiulară unitară, *U* este o matrice superior triunghiulară unitară iar *D* este o matrice diagonală. Existenţa unei astfel de descompuneri presupune aceleaşi condiţii de îndeplinit de către matricea *A*, ca şi în cazul factorizării *LU*. În plus, din modul cum a fost definită rezultă imediat şi caracterul unic al factorizării *LDU*.

Observaţia 6.15: Dacă există o descompunere *LDU* pentru o matrice nesingulară, *A* (vezi relaţia (6.128)) atunci, ţinând cont de modul cum au fost definite mai sus cele trei descompuneri *LU* fundamentale, precum şi de faptul că toate acestea reprezintă descompuneri unice atunci, factorizarea *LDU* este echivalentă cu:

- o *factorizare Doolittle* astfel:

$$A = L \cdot (D \cdot U) = L \cdot U_{Doolittle} \qquad (6.129)$$

- o *factorizare Crout* astfel:

$$A = (L \cdot D) \cdot U = L_{Crout} \cdot U \qquad (6.130)$$

iar în cazul în care matricea *A* este simetrică şi pozitiv definită, cu

- o *factorizare Cholesky* astfel:

$$A = (L \cdot D_1) \cdot (D_1 \cdot U) = L_C \cdot L_C^{*T} \qquad (6.131)$$

unde matricea $D_1 = diag\left(\sqrt{d_1}, \sqrt{d_2}, \ldots, \sqrt{d_n}\right)$ este dată de relaţiile (6.127) şi, respectiv, (6.137). Modul cum se deduce această matrice va fi prezentat în cele ce urmează.

6.6.2. Descompunerea *Cholesky*

Dacă matricea $A \in R^{n \times n}$ este o matrice nesingulară şi, în plus, ea mai este şi **simetrică** şi **pozitiv definită** atunci aceasta poate fi descompusă folosind descompunerea *Cholesky* dată de relaţia:

$$A = L_C \cdot L_C^{*T} \quad (6.132)$$

unde L_C este o matrice inferior triunghiulară ale cărei elemente de pe diagonala principală sunt strict pozitive iar L_C^{*T} este matricea L_C transpusă şi complex conjugată.

Descompunerea *Cholesky* – definită, aşa cum am văzut mai sus, pentru matrici A simetrice şi pozitiv definite **– este o descompunere unică**. Cu alte cuvinte, pentru o matrice A îndeplinind condiţiile de mai sus va exista întotdeauna doar o singură matrice inferior triunghiulară care să aibă elementele de pe diagonala principală strict pozitiv definite şi care să satisfacă relaţia (6.132).

Inversa afirmaţiei precedente este, de asemenea, adevărată. Astfel, dacă avem o matrice A care se poate descompune sub forma dată de relaţia (6.132), în care L este o matrice inferior triunghiulară ce are toate elementele de pe diagonala principală mai mari strict decât zero atunci această matrice este simetrică şi pozitiv definită.

Întrucât toate matricile de covarianţă satisfac ambele condiţii cerute pentru o descompunere de tip *Cholesky*, fiind, în plus, Hermitian-simetrice, descompunerea acestora poate fi făcută într-o reprezentare *LDU* ca un produs de matrici de forma[11]:

$$C_x = L \cdot D_L \cdot L^{*T} \quad (6.133)$$

În relaţia de mai sus matricea L este o matrice inferior triunghiulară unitară iar matricea D_L este o matrice diagonală (diferită, în general, de matricea unitate). Se observă de aici uşor că şi matricea L^{*T} este, la rândul ei, o matrice superior triunghiulară unitară (respectiv, toate elementele ei diagonale sunt egale cu unitatea iar termenii aflaţi sub diagonala principală sunt toţi egali cu zero).

Observaţia 6.16: Tot din relaţia (6.133) se mai poate observa faptul că matricea L din descompunerea (6.133) este una diferită de matricea L_C a descompunerii de tip *Cholesky* aplicată aceleiaşi matrici de covarianţă (vezi relaţia generală (6.132)). De altfel, între matricile L_C şi L ale celor două descompuneri – (6.132) şi, respectiv, (6.133) – există următoarea relaţie clar definită:

[11] În mod similar, şi toate matricile de corelaţie pot fi descompuse într-o formă de tip *LU*, adică:

$$R_x = L \cdot D_L \cdot L^{*T}$$

În relaţia anterioară, în modul cel mai general, matricile L şi D_L sunt diferite de cele din relaţia (6.133), descompunerea în forma *LU*, fiind unică. Aceste două matrici, R_x şi C_x, au descompuneri identice doar atunci când ele sunt egale (respectiv, atunci când vectorul medie al vectorului aleator x este zero).

- să presupunem că scriem matricea C_x sub forma descompunerii *Cholesky* astfel:

$$C_x = L_C \cdot L_C^{*T} \qquad (6.134)$$

  şi totodată descompunem aceeaşi matrice C_x sub forma dată de relaţia (6.133):

$$C_x = L \cdot D_L \cdot L^{*T} \qquad (6.135)$$

- în relaţia (6.135) matricea D_L este dată de o matrice de forma:

$$D_L = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} = diag(d_1, d_2, ..., d_n) \quad (6.136)$$

- din relaţiile (6.134), (6.135) şi (6.136) se poate deduce uşor relaţia dintre matricile L şi L_C, respectiv:

$$L_C = L \cdot diag\left(\sqrt{d_1}, \sqrt{d_2}, \ldots, \sqrt{d_n}\right) \quad (6.137)$$

Observaţia 6.17: Având în vedere echivalenţa dintre factorizarea *LDU* şi factorizarea *Doolittle* dată de relaţia (6.129) şi, în plus, ţinând cont de descompunerea *LDU* a unei matrici de covarianţă (vezi relaţia (6.133)), obţinem:

$$D_L \cdot L^{*T} = U_{Doolittle} \quad (6.138)$$

Întrucât L^{*T} este o matrice *superior triunghiulară unitară*, rezultă din ecuaţia de mai sus că elementele de pe diagonala principală a matricei $U_{Doolittle}$ sunt chiar elementele de pe diagonala principală a matricei diagonale D_L.

Reamintim aici că ***descompunerea dată sub forma relaţiei*** (6.133) ***este şi ea, la rândul ei, unică***, fiind o descompunere de tip *LDU*, particularizată pentru matrici simetrice şi pozitiv definite.

Dacă în relaţia (6.133) înmulţim în stânga cu L^{-1} şi în partea dreaptă cu $(L^{*T})^{-1}$ obţinem:

$$L^{-1}\, C_x\, (L^{-1})^{*T} = D_L \quad (6.139)$$

De aici se poate observa că D_L este, de fapt, matricea de covarianţă, C_y, pentru un vector aleator y obţinut din vectorul aleator x cu ajutorul următoarei transformate:

$$y = L^{-1}\, x \quad (6.140)$$

Deoarece D_L este o matrice diagonală, componentele vectorului aleator y sunt necorelate iar elementele diagonale ale matricii D_L sunt însăşi momentele de ordin doi (varianţele) fiecărei trăsături în parte.

În timp ce transformata bazată pe vectorii proprii (transformata unitară) poate fi interpretată ca o rotire a sistemului de coordonate, transformata dată de relaţia (6.140) poate fi interpretată drept o transformare cauzală.

Problemă 6.13: Utilizând mediul de dezvoltare LabWindows CVI scrieţi un program care să descompună următoarea matrice de covarianţă într-un produs de matrici similar relaţiei (6.133):

$$C_x = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

Rezolvare: **(1).** ntr-o primă variantă, utilizând una funcţiile oferite de biblioteca „*Advanced analysis*” a mediului de dezvolare LabWindow CVI, mai exact funcţia *LU*, putem realiza descompunerea inferior-superior triunghiulară a matricei de mai sus (programul implementat îl găsiţi în directorul ***Descompunere LU*** asociat acestui capitol). Această funcţie acceptă drept intrare o matrice ce urmează a fi descompusă în forma dată de descompunerea *LU* de tip *Doolittle*. La terminarea execuţiei funcţiei *LU* rezultatul obţinut este întors în aceeaşi matrice în care au fost furnizate datele de intrare, respectiv: matricea *U* ocupă partea triunghiular superioară a matricei de intrare, inclusiv elementele de pe diagonala principală, în timp ce matricea *L* ocupă partea triunghiular inferioară a aceleiaşi matrici. În final, ţinând cont de particularitatea matricei *L* din descompunerea *Doolittle*, se trece la distribuirea corespunzătoare a elementelor matricei rezultante în cele două matrici ale descompunerii *LU*, respectiv:

$$C_x = \begin{bmatrix} 1/2 & 1/4 \\ 1/4 & 1/2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ 1/2 & 1 \end{bmatrix}}_{L} \cdot \underbrace{\begin{bmatrix} 1/2 & 1/4 \\ 0 & 3/8 \end{bmatrix}}_{U} \quad (6.141)$$

Din relaţia de mai sus se observă că descompunerea astfel obţinută este diferită faţă de cea prezentată în relaţia (6.133) – descompunere la care, de altfel, dorim să ajungem. Pentru aceasta ne vom folosi de faptul că elementele de pe diagonala principală a matricii D_L sunt identice cu elementele de pe diagonala principală a matricii U din descompunerea anterior prezentată (vezi ***Observaţia 6.17***). În acest caz, forma finală a descompunerii este dată de:

$$C_x = \begin{bmatrix} 1/2 & 1/4 \\ 1/4 & 1/2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ 1/2 & 1 \end{bmatrix}}_{L} \cdot \underbrace{\begin{bmatrix} 1/2 & 0 \\ 0 & 3/8 \end{bmatrix}}_{D_L} \cdot \underbrace{\begin{bmatrix} 1 & 1/2 \\ 0 & 1 \end{bmatrix}}_{L^{*T}} \qquad (6.142)$$

(2). Mediul de dezvoltare LabWindow CVI dispune, de asemenea, în propriile librării şi de o funcţie capabilă să efectueze descompunerea *Cholesky*. În această a doua abordare a problemei, utilizând funcţia de mai sus în cadrul programului din directorul ***Descompunere LU*** obţinem:

$$C_x = \begin{bmatrix} 1/2 & 1/4 \\ 1/4 & 1/2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0.7071 & 0 \\ 0.3536 & 0.6124 \end{bmatrix}}_{L_C} \cdot \underbrace{\begin{bmatrix} 0.7071 & 0.3536 \\ 0 & 0.6124 \end{bmatrix}}_{L_C^{*T}} \qquad (6.143)$$

Pentru a ajunge la forma (6.133) plecăm de la matricea L_C

$$L_C = \begin{bmatrix} l_{C1,1} & 0 & \cdots & 0 \\ l_{C2,1} & l_{C2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{Cn,1} & l_{Cn,2} & \cdots & l_{Cn,n} \end{bmatrix} \qquad (6.144)$$

şi ne folosim de relaţiile:

$$diag(L_C) = \begin{bmatrix} l_{C1,1} & 0 & \cdots & 0 \\ 0 & l_{C2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & l_{Cn,n} \end{bmatrix} \qquad (6.145)$$

$$D_L = diag(L_c) \cdot diag(L_c) = \begin{bmatrix} l_{C1,1}^2 & 0 & \cdots & 0 \\ 0 & l_{C2,2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & l_{Cn,n}^2 \end{bmatrix} \qquad (6.146)$$

şi

$$L = L_C \cdot [diag(L_C)]^{-1} \qquad (6.147)$$

Relaţiile (6.146) şi (6.147) rezultă în mod direct din (6.134), (6.135), (6.136) şi (6.137).

Particularizând relaţiile (6.146) şi (6.147) obţinem:

$$diag(L_c)=\begin{bmatrix}0.7071 & 0\\ 0 & 0.6124\end{bmatrix}$$

şi de aici:

$$D_L=\begin{bmatrix}0.4999904 & 0\\ 0 & 0.3750337\end{bmatrix}\cong\begin{bmatrix}0.5 & 0\\ 0 & 0.375\end{bmatrix}$$

În final, utilizând relaţia (6.147) obţinem:

$$L=\begin{bmatrix}0.7071 & 0\\ 0.3536 & 0.6124\end{bmatrix}\cdot\begin{bmatrix}1/0.7071 & 0\\ 0 & 1/0.6124\end{bmatrix}=\begin{bmatrix}1 & 0\\ 1/2 & 1\end{bmatrix}$$

Problemă 6.14: Se dă următoarea matrice de covarianţă:

$$C_x=\begin{bmatrix}3 & -3 & -6\\ -3 & 4 & 8\\ -6 & 8 & 20\end{bmatrix}$$

(a) Să se găsească – de această dată, printr-o metodă manuală de calcul – descompunerea *LU* a acestei matrici şi, respectiv, descompunerea dată de o relaţie similară cu relaţia (6.133).

(b) Folosind descompunerea *LU* să se determine matricea transformării A ($y = A \cdot x$) care va determina ca în noul spaţiu, cel al vectorului y de trăsături, componentele vectorului aleator y să fie necorelate.

(c) Determinaţi matricea de covarianţă ce caracterizează noul vector aleator de trăsături, y, obţinut la punctul **(b)**.

Rezolvare:

(a) ***Metoda clasică*** utilizată în descompunerea unei matrici nesingulare A în două matrici de tipul inferior-superior triunghiulare, $A = L\cdot U$, se bazează pe metoda de reducere ***Gauss***. În cazul utilizării acestei metode de reducţie, simultan aplicată şi asupra unei matrici identitate de aceeaşi dimensiune cu A, vom obţine în final, în acelaşi timp, atât matricea superior triunghiulară, U, cât şi matricea L^{-1} (inversa matricei inferior triunghiulară, L, a descompunerii *LU*).

Pentru obţinerea simultană şi a matricii L^{-1}, o dată cu obţinerea matricei U, vom pune în paralel cu matricea C_x şi matricea identitate – metoda reducerii *Gauss* aplicându-se, aşa cum am spus, simultan ambelor matrici:

$$\left[\begin{array}{ccc|ccc} 3 & -3 & -6 & 1 & 0 & 0 \\ -3 & 4 & 8 & 0 & 1 & 0 \\ -6 & 8 & 20 & 0 & 0 & 1 \end{array}\right] \quad (6.148)$$

Scopul final al metodei este acela de a reduce la zero toate elementele poziţionate sub diagonala principală a matricii C_x, matricea rezultantă fiind, în acest caz, matricea U superior triunghiulară a descompunerii LU. Operaţiile efectuate asupra matricii C_x vor fi efectuate simultan şi asupra matricei identitate, I; matricea ce rezultă în final în acest caz va fi matricea L^{-1} prezentată mai sus.

Metoda de reducere *Gauss*:

- Într-un prim pas vom reduce la zero elementele din prima coloană, mai exact acele elemente poziţionate sub primul element aflat pe diagonala principală a matricii C_x. Pentru a ne atinge acest obiectiv vom multiplica de fiecare dată prima linie cu o anumită valoare astfel încât în momentul scăderii acestei linii din liniile inferioare să rezulte o valoare zero în prima poziţie a liniei inferioare respective. Aceste elemente multiplicative, corespunzând diferitelor linii din matricea C_x, se numesc ***pivoţi***. Ca menţiune: elementele primei linii rămân nemodificate pe parcursul întregului procedeu. Aplicând acest algoritm relativ simplu obţinem pentru problema noastră:

$$\left[\begin{array}{ccc|ccc} 3 & -3 & -6 & 1 & 0 & 0 \\ -3 & 4 & 8 & 0 & 1 & 0 \\ -6 & 8 & 20 & 0 & 0 & 1 \end{array}\right] \xrightarrow{-1} \left[\begin{array}{ccc|ccc} 3 & -3 & -6 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ -6 & 8 & 20 & 0 & 0 & 1 \end{array}\right] \xrightarrow{-2} \left[\begin{array}{ccc|ccc} 3 & -3 & -6 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 2 & 8 & 2 & 0 & 1 \end{array}\right]$$

$$\underbrace{\qquad\quad}_{C_x} \quad \underbrace{\qquad}_{I}$$

(6.149)

Pivoţii utilizaţi pentru punerea pe zero a elementelor din prima coloană, poziţionate sub diagonala principală, au fost notaţi deasupra săgeţilor ce marchează diferitele transformări, aplicate simultan, celor două matrici, C_x şi I.

- Într-un pas următor (final, dealtfel în cazul nostru particular) se vor pune pe zero şi elementele din cea de a doua coloană poziţionate sub diagonala principală. Metoda folosită este una similară cu cea prezentată mai sus, cu observaţia că de această dată se va înmulţi linia doi cu pivoţii corespunzători liniilor următoare din matrice şi ulterior se va scădea corespunzător din fiecare linie inferioară în parte. Generalizând, în cazul unor matrici de dimensiuni mai mari, procedeul continuă în mod similar până la inclusiv penultima coloană a matricei de interes. În cazul nostru obţinem:

$$\left[\begin{array}{ccc|ccc} 3 & -3 & -6 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 2 & 8 & 2 & 0 & 1 \end{array}\right] \overset{2}{\longrightarrow} \left[\begin{array}{ccc|ccc} 3 & -3 & -6 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 4 & 0 & -2 & 1 \end{array}\right] \tag{6.150}$$

$$\underbrace{\qquad\qquad}_{U}\quad\underbrace{\qquad\qquad}_{L^{-1}}$$

În relaţia (6.150) observăm că, pe lângă matricea superior triunghiulară obţinem simultan şi inversa matricei inferior triunghiulare unitare, L^{-1}, care nu este altceva decât chiar matricea transformării din spaţiul iniţial al vectorului aleator x – caracterizat de matricea de covarianţă C_x – într-un alt spaţiu y, caracterizat de o matrice de covarianţă, C_y, dată de matricea diagonală D_L.

Matricea inferior triunghiulară unitară se obţine, în final, dintr-o matrice unitate, I, în care se inserează pivoţii obţinuţi în timpul metodei de *reducere Gauss*. Aceşti pivoţi vor fi inseraţi în acele poziţii specifice la care au contribuit în procesul de reducere la zero a diferitelor elemente ale matricii iniţiale C_x.

În aceste condiţii, şi ţinând cont şi de pivoţii folosiţi în pocedeul Gauss de reducţie, matricea inferior triunghiulară unitară calculată pentru matricea C_x iniţială va fi:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \tag{6.151}$$

În final, descompunerea matricei C_x în formatul LU vafi dată de:

$$C_x = \begin{bmatrix} 3 & -3 & -6 \\ -3 & 4 & 8 \\ -6 & 8 & 20 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix}}_{L} \cdot \underbrace{\begin{bmatrix} 3 & -3 & -6 \\ 0 & 1 & 2 \\ 0 & 0 & 4 \end{bmatrix}}_{U} \tag{6.152}$$

În continuare, pentru obţinerea descompunerii matricei C_x într-un format similar cu cel dat de relaţia (6.133) vom ţine cont de ***Observaţia 6.17***, respectiv: elementele de pe diagonala principală a matricei D_L sunt elementele de pe diagonala principală a matricei U rezultată în urma descompunerii matricei C_x, relaţia (6.152). Ştiind toate acestea putem în final scrie:

$$C_x = \begin{bmatrix} 3 & -3 & -6 \\ -3 & 4 & 8 \\ -6 & 8 & 20 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix}}_{L} \cdot \underbrace{\begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}}_{C_y = D_L} \cdot \underbrace{\begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}_{L^{*T}} \qquad (6.153)$$

(b) Pentru decorelarea trăsăturilor unui vector aleator x vom utiliza transformata $A = L^{-1}$:

$$y = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}}_{L^{-1}} \cdot x \qquad (6.154)$$

(c) Transformata A determină ca în noul spaţiu – cel al vectorului aleator de trăsături y –, trăsăturile să fie necorelate. Matricea de covarianţă în noul spaţiu de trăsături este, în acest caz, chiar matricea diagonală din descompunerea dată de relaţia (5.152):

$$C_y = D_L = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \qquad (6.155)$$

Aplicaţie 6.4: Utilizând programul aflat în directorul ***Gen_distrib_2D_Gauss*** asociat acestui capitol al cărţii:

1. Deteminaţi relaţiile matematice, respectiv, transformata, prin intermediul căreia puteţi să transformaţi cele două variabile aleatoare n_1 şi n_2 necorelate (ambele având o distribuţie de tipul $\mathcal{N}(0, 1)$) ale vectorului aleator x într-un vector aleator z caracterizat de o matrice de covarianţă C_z. În spaţiul vectorului aleator z va exista între trăsături o corelaţie ρ iar acestea vor avea varianţele σ_1^2 şi σ_1^2. După deducerea relaţiilor verificaţi corectitudinea implementării acestora în cadrul programului prezentat anterior.
2. Pentru următoarele valori ale varianţelor şi ale coeficientului de corelaţie între variabilele aleatoare ale vectorului z determinaţi matricile de covarianţă corespondente:

   i. $\sigma_1^2 = 4, \quad \sigma_2^2 = 9, \quad \rho = 0.2$,
   ii. $\sigma_1^2 = 4, \quad \sigma_2^2 = 9, \quad \rho = 0.6$,
   iii. $\sigma_1^2 = 4, \quad \sigma_2^2 = 9, \quad \rho = 0.8$,
   iv. $\sigma_1^2 = 4, \quad \sigma_2^2 = 9, \quad \rho = 1$.

3. Determinaţi pentru aceiaşi parametri prezentaţi la punctul 2 şi matricile transformării ($z = A \cdot x$).
4. Generaţi seriile de date, pe baza matricilor de covarianţă anterior determinate, şi analizaţi-le ulterior cu ajutorul programului din directorul ***Determinare Contur si Decorelare***.
5. Verificaţi corectitudinea rezultatelor obţinute.

Rezolvare:

1. Cerinţa primului punct este aceea de a găsi o transformată A:

$$z = A \cdot x \quad (6.156)$$

astfel încât, plecând din spaţiul x de trăsături caracterizat de o matrice de covarianţă unitară să ajungem în spaţiul z caracterizat de o matrice de covarianţă $C_z = \begin{bmatrix} cz_{11} & cz_{12} \\ cz_{21} & cz_{22} \end{bmatrix}$, ţinându-se cont de următoarele constrângeri: variabilele aleatoare vor fi corelate cu un coeficient de corelaţie ρ şi, în plus, acestea vor avea varianţele σ_1 şi σ_2.

Prin utilizarea relaţiilor (5.200), (5.201), şi (5.243):

$$c_{kl} = \mathrm{Cov}\,(x_k, x_l) = E\{\,(x_k - m_k)\,(x_l - m_l)^*\,\} \quad (6.157)$$

$$c_{kk} = E\{\,|\,x_k - m_k\,|^2\,\} = \sigma_{xk}^2 \quad (6.158)$$

$$\rho = \frac{Cov(x_k, x_l)}{\sigma_{x_k} \cdot \sigma_{x_l}} \quad (6.159)$$

putem scrie matricea de covarianţă ce caracterizează distribuţia spaţială a datelor date de vectorul aleator z sub forma:

$$C_z = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \quad (6.160)$$

Întrucât în cadrul acestei cărţi s-au prezentat câteva tehnici prin care s-a realizat decorelarea trăsăturilor unui vector aleator (transformata unitară, transformata LU, utilizarea descompunerii în valori singulare) şi scalarea seturilor de date astfel încât varianţa fiecărei variabile aleatoare ce compune vectorul aleator să fie egală cu unitatea (transformata de înălbire) ne vom folosi de toate aceste cunoştiinţe dobândite şi vom merge cu procesarea datelor în sens invers relaţiei (6.156). Cu alte cuvinte, vom determina o matrice B astfel încât:

$$x = B \cdot z \quad (6.161)$$

Deci, plecând de la matricea C_z dată de relaţia (6.160) trebuie să ajungem, în final, la o matrice de covarianţă în spaţiul x unitară. O dată determinată matricea B, transformata dorită A va fi dată de $A = B^{-1}$.

Pentru decorelarea setului iniţial de date vom utiliza descompunerea *Cholesky*. Prin intermediul acestei transformate, conform teoriei prezentate anterior şi folosind relaţia (6.133) matricea de covarianţă C_z se va descompune sub forma:

$$C_z = L \cdot D_L \cdot L^{*T} \qquad (6.162)$$

de unde obţinem mai departe,

$$L^{-1} \cdot C_z \cdot \left(L^{-1}\right)^{*T} = D_L \qquad (6.163)$$

Rezultă de aici că transformata $y = C \cdot x$ care ne duce din spaţiul z într-un spaţiu y al trăsăturilor necorelate (nu neapărat şi de varianţă unu) este transformata $C = L^{-1}$.

Aplicând metoda de reducere *Gauss* descompunem matricea C_z astfel:

$$C_z = \left[\begin{array}{cc|cc} \sigma_1^2 & \rho\sigma_1\sigma_2 & 1 & 0 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 & 0 & 1 \end{array}\right] \xrightarrow{\rho\frac{\sigma_2}{\sigma_1}} \left[\begin{array}{cc|cc} \sigma_1^2 & \rho\sigma_1\sigma_2 & 1 & 0 \\ 0 & \sigma_2^2 - \rho^2\sigma_2^2 & -\rho\dfrac{\sigma_2}{\sigma_1} & 1 \end{array}\right]$$

$$\qquad\qquad\qquad\qquad\qquad\qquad \underbrace{}_{U} \quad \underbrace{\phantom{-\rho\frac{\sigma_2}{\sigma_1} \; 1}}_{L^{-1}}$$

(6.164)

Prin utilizarea pivoţilor (pivotului în cazul nostru particular) se poate determina în mod direct matricea L:

$$L = \begin{bmatrix} 1 & 0 \\ \rho\dfrac{\sigma_2}{\sigma_1} & 1 \end{bmatrix} \qquad (6.165)$$

Descompunerea finală a matricii C_z va fi, în acest caz:

$$C_z = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ \rho\dfrac{\sigma_2}{\sigma_1} & 1 \end{bmatrix}}_{L} \cdot \underbrace{\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 - \rho^2\sigma_2^2 \end{bmatrix}}_{D_L} \cdot \underbrace{\begin{bmatrix} 1 & \rho\dfrac{\sigma_2}{\sigma_1} \\ 0 & 1 \end{bmatrix}}_{L^{*T}}$$

(6.166)

În urma aplicării transformatei $C = L^{-1}$ vom obţine o trecere din spaţiul z în spaţiul y:

$$z \xrightarrow{L^{-1}} y \quad (6.167)$$

În noul spaţiu de trăsături acestea sunt necorelate, dovada fiind dată de forma diagonală a matricei de covarianţă:

$$C_y = D_L = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2\left(1-\rho^2\right) \end{bmatrix} \quad (6.168)$$

Pentru ca matricea de covarianţă în spaţiul x să fie egală cu matricea unitate vom aplica o nouă transformată $\Lambda^{-½}$, transformată similară cu cea utilizată în cadrul subcapitolului *"Diagonalizări simultane ale matricelor de covarianţă a două clase distincte"*. Această matrice va fi de forma:

$$\Lambda^{-½} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{bmatrix} \quad (6.169)$$

În urma aplicării transformatei (6.169) vom trece din spaţiul y, caracterizat de matricea de covarianţă (6.168), în spaţiul z unde matricea de covarianţă va fi egală cu:

$$C_x = \Lambda^{-½} C_y \left(\Lambda^{-½}\right)^{*T} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{bmatrix} \cdot \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2\left(1-\rho^2\right) \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
(6.170)

Prin această aplicare, în cascadă, a celor două transformate am obţinut:

$$\underbrace{z}_{C_z = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}} \xrightarrow{L^{-1}} \underbrace{y}_{C_y = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2\left(1-\rho^2\right) \end{bmatrix}} \xrightarrow{\Lambda^{-1/2}} \underbrace{x}_{C_x = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}$$
(6.171)

De aici deducem relaţia matematică a transformatei care ne-ar duce din spaţiul vectorului aleator z în spaţiul vectorului aleator x, în condiţiile specificate de aplicaţie:

$$x = \Lambda^{-½} \cdot L^{-1} \cdot z \quad (6.172)$$

Pentru determinarea transformatei A definită în (6.156) vom înmulţi la stânga relaţia (6.172) mai întâi cu $\Lambda^{½}$, iar apoi vom înmulţi cu matricea L, în final obţinându-se:

$$z = \underbrace{L \cdot \Lambda^{½}}_{A} \cdot x \qquad (6.173)$$

Transformata care îndeplineşte în totalitate cerinţele problemei noastre este deci:

$$A = L \cdot \Lambda^{½} = \begin{bmatrix} 1 & 0 \\ \rho \dfrac{\sigma_2}{\sigma_1} & 1 \end{bmatrix} \cdot \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix} \qquad (6.174)$$

Relaţia (6.173) se poate scrie acum:

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad (6.175)$$

Dacă dorim ca distribuţia în planul z să fie poziţionată şi în vectorul mediu $m_z = [m_1, m_2]^T$ atunci avem:

$$\begin{cases} z_1 = m_1 + \sigma_1 \cdot x_1 \\ z_2 = m_2 + \rho \cdot \sigma_2 \cdot x_1 + \sqrt{1-\rho^2} \cdot \sigma_2 \cdot x_2 \end{cases} \qquad (6.176)$$

6.6.3. Interpretarea descompunerii LU şi aplicaţii

Deoarece în cazul unei descompuneri de forma $C_x = L{\cdot}D_L{\cdot}L^{*T}$ (vezi relaţia (6.133)) a unei matrici de covarianţă, matricea tansformării din spaţiul x în spaţiul y este L^{-1}, pentru a interpreta semnificaţia acestei transformate trebuie mai întâi să vedem ce formă va lua matricea transformării.

Observaţia 6.18: Înainte de a demonstra că transformata obţinută în urma unei descompuneri *Cholesky* este una cauzală sau nu vom arăta mai întâi faptul că, întrucât L este o matrice inferior triunghiulară unitară, matricea L^{-1} prezintă şi ea aceeaşi proprietate.

Demonstraţie: Pentru a demonstra această proprietate ne vom folosi de metoda inducţiei. Să presupunem că această proprietate este valabilă pentru matricile de ordin d–1. În acest caz pentru o matrice inferior triunghiulară de dimensiune d putem scrie:

$$L_d^{-1} \cdot L_d = I \qquad (6.177)$$

sau

$$\left[\begin{array}{c|c} L_{d-1}^{-1} & \beta_{d-1} \\ \hline \gamma_{d-1}^{T} & \alpha \end{array}\right]\left[\begin{array}{c|c} L_{d-1} & \begin{matrix}0\\0\\0\\\vdots\\0\end{matrix} \\ \hline l_{d-1}^{T} & 1 \end{array}\right]=\left[\begin{array}{ccccc|c} 1 & & & 0 & & 0 \\ & 1 & & & & 0 \\ & & 1 & \cdots & & 0 \\ & & & 1 & & \\ 0 & & \vdots & & \ddots & 0 \\ \hline 0 & 0 & 0 & \cdots & 0 & 1 \end{array}\right]$$

(6.178)

Deoarece produsul celor două matrici trebuie să fie egal cu matricea unitate rezultă că elementul α este egal cu 1. Mai mult, dacă ar exista elemente ale vectorului $\beta_{d\text{-}1}$ diferite de zero atunci unele dintre elementele din ultima coloană a matricii unitate ar fi şi ele diferite de zero. Dar, întrucât matricea I este matricea unitate (ea neputând avea elemente diferite de zero decât pe diagonala principală) rezultă că matricea inversă a matricii L_d nu poate avea elemente diferite de zero pe ultima coloană (excepţie făcând ultimul element care este egal cu 1); de aici rezultă că matricea L_d^{-1} este una inferior triunghiulară unitară. Deoarece proprietatea de matrice inferior triunghiulară unitară este trivială pentru cazul $d = 1$, rezultă că proprietatea enunţată mai sus este valabilă pentru orice valoare reală d.

În concluzie vom spune că, dacă **o matrice *A* este o matrice inferior triunghiulară unitară atunci inversa ei va fi tot o matrice inferior triunghiulară unitară**.

În mod similar cu argumentaţia anterioară se poate demonstra că, dacă A este o matrice superior triunghiulară unitară atunci inversa ei va fi tot o matrice de forma superior triunghiulară unitară.

Deoarece L^{-1} este o matrice inferior triunghiulară unitară, prin aplicarea acestei transformate unui vector aleator x, N-dimensional obţinem:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \\ \vdots \\ y_N \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ l_{2,1} & 1 & \cdots & 0 & \cdots & 0 \\ l_{3,1} & l_{3,2} & \ddots & & & \vdots \\ l_{k,1} & l_{k,2} & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & & \ddots & \vdots \\ l_{N,1} & l_{N,2} & \cdots & l_{N,k} & \cdots & 1 \end{bmatrix}}_{L^{-1}} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ \vdots \\ x_N \end{bmatrix} \qquad (6.179)$$

Din relaţia (6.179) se observă că orice componentă cu indicele k, y_k**,** a vectorului y este o funcţie de variabilele x_l pentru $l \le k$, după cum se prezintă în relaţia de mai jos:

$$y_k = x_k + l_{k,k-1} \cdot x_{k-1} + l_{k,k-2} \cdot x_{k-2} + \ldots + l_{k,2} \cdot x_2 + l_{k,1} \cdot x_1 \quad (6.180)$$

Dacă, de exemplu, elementele vectorului x ar fi eşantioane consecutive ale unei secvenţe aleatoare, atunci L^{-1} ar reprezenta o ***transformată cauzală***.

Aşa după cum am spus în introducerea acestui subcapitol, o transformată *LU* poate fi deosebit de utilă în situaţii cum ar fi: calculul determinanţilor unor matrici, calculul inverselor matricelor nesingulare etc.

1. Calculul determinanţilor matricelor de covarianţă

Toate matricile A inferior triunghiulare unitare au determinantul egal cu 1. Dacă descompunem matricea de covarianţă C_x sub forma (6.133) şi ţinem cont că o matrice de tipul L este inferior triunghiulară unitară atunci vom avea:

$$\left|C_x\right| = \left|L\right| \cdot \left|D_L\right| \cdot \left|L^{*T}\right| = \left|D_L\right| = \prod_{k=1}^{d} d_{kk} \quad (6.181)$$

unde d_{kk} sunt elementele de pe diagonala principală a matricii D_L. Ecuaţia (6.181) este similară cu relaţia (6.34). Deoarece d_{kk} sunt momentele de ordin doi ale componentelor y_k ale vectorului aleator y, relaţia (6.181) arată că determinantul matricei de covarianţă poate fi scris ca un produs al momentelor de ordin doi, respectiv, al varianţelor variabilelor aleatoare ce compun vectorul aleator y.

În cazul descompunerii matricei de corelaţie sau a celei de covarianţă sub forma dată de descompunerea *Cholesky* (6.134):

$$C_x = L_C \cdot L_C^{*T} \quad (6.182)$$

observăm din relaţia (6.146) că determinantul matricei de covarianţă (corelaţie) a unui set de date este egal cu produsul pătratelor elementelor de pe diagonala principală a matricei L_C, respectiv:

$$\left|C_x\right| = \left|D_L\right| = \prod_{k=1}^{d} l_{Ck,k}^2 \quad (6.183)$$

2. Calculul inversei unei matrici

Descompunerea în forma *LU* este utilizată, de asemenea, şi în calculul inversei unei matrici generice *A* prin intermediul relaţiei:

$$A^{-1} = U^{-1} \cdot L^{-1} \quad (6.184)$$

Algoritmul de calcul al inversei unei matrici prin intermediul transformatei *LU* este mai rapid decât cel standard, bazat pe algoritmul *Gauss-Jordan*, chiar dacă acest prim algoritm necesită la un loc: calculul transformatei *LU*, inversarea matricelor *L* şi *U*, şi, în final, înmulţirea matricelor inversate, U^{-1} şi L^{-1}. În acest sens trebuie să ţinem cont de faptul că forma triunghiulară a acestor matrici uşurează mult calculul inverselor matricelor *L* şi *U*.

Dacă matricea generică *A* – a cărei inversă dorim să o calculăm – este, în plus, şi simetrică şi pozitiv definită atunci putem utiliza algoritmul *Cholesky*. Dacă luăm în considerare doar factorii termenilor cu puterea cea mai ridicată din polinomul ce descrie numărul de înmulţiri necesare finalizării algoritmului vom observa că pentru descompunerea de tip *LU* sunt necesare $2 \cdot n^3/3$ înmulţiri în timp ce algoritmul de descompunere *Cholesky* necesită doar $n^3/3$ înmulţiri [Happonen, 2005]. În concluzie, în momentul în care dorim să calculăm inversa unei matrici simetrice şi pozitiv definite, aşa cum este cazul matricelor de corelaţie şi, respectiv, de covarianţă, vom descompune, mai întâi, matricea respectivă prin intermediul algoritmului *Cholesky* şi apoi vom calcula inversa produsului $L_C \cdot L_C^{*T}$. Pentru aceasta vom ţine cont suplimentar de faptul că $(A^T)^{-1} = (A^{-1})^T$, în acest mod fiind de ajuns să calculăm inversa doar pentru o matrice din cele două.

Diagonalizarea matricei de corelaţie prin descompunerea triunghiulară urmează o procedură identică, matricea de corelaţie (R_x) putând fi pur şi simplu substituită matricei de covarianţă C_x în toate ecuaţiile anterior prezentate. În acest caz componentele vectorului *y* definite de relaţia (6.140) ($y = L^{-1} \cdot x$) nu vor mai fi necorelate ci ele vor fi *ortogonale*.

În multe situaţii practice este mai util să se înlăture media din setul de date. În acest caz diagonalizarea matricei de corelaţie coincide cu diagonalizarea matricei de covarianţă (cele două matrici fiind identice) şi, ca rezultat final, trăsăturile noului vector obţinut prin transformata L^{-1} vor fi atât necorelate cât şi ortogonale.

6.6.4. Utilizarea metodei de factorizare QR

1. Introducere

Metoda de factorizare de tip QR este o altă metodă de tip "rădăcină pătrată" care poate fi utilizată pentru diagonalizarea oricărui tip de matrice, deci şi a matricei de corelaţie, respectiv, de covarianţă. Când această metodă este aplicată matricei de date, *X*, se pot pune în evidenţă aceleaşi relaţii, din punct de vedere al complexităţii şi încărcării computaţionale care există între descompunerea prin metoda SVD şi descompunerea unitară.

O ***descompunere QR*** a unei matrici reale *A*, cu $A \in R^{m \times n}$ şi $m \geq n$, este dată de relaţia:

$$A = Q \cdot R \quad (6.185)$$

În relaţia (6.185) matricea *Q*, de tip *m* x *n*, este o matrice ce are coloanele ortogonale în timp ce matricea *R* este o matrice pătratică, *n* x *n*, superior triunghiulară.

În mod similar putem obţine o altă descompunere, mai generală, tot de tip QR, dată în acest caz de produsul dintre o matrice unitară *Q* de dimensiune *m* x *m* şi o matrice superior triunghiulară, de dimensiune *m* x *n*.

2. Obţinerea transformatei QR prin metoda Gram–Schmidt

Pentru obţinerea descompunerii QR se cunosc mai mult metode precum teorema de ortogonalizare Gram–Schmidt [Horn, 1985], transformarea Householder [Cox, 1997], [Pothen, 1989] sau metoda de rotire Givens [Pothen, 1989]. Fiecare dintre aceste metode are avantajele şi dezavantajele ei.

În cadrul acestei cărţi se va prezenta doar metoda de ortogonalizare Gram–Schmidt, metodă cunoscută în general studenţilor şi persoanelor ce au tangenţă cu profilurile electric, electronic, calculatoare şi informatică. Metoda de ortogonalizare Gram–Schmidt, ce va fi utilizată în obţinerea factorizării QR, va fi particularizată în cele ce urmează pentru matricea setului de date, *X*, a unui vector aleator *x*.

În această situaţie se pleacă de la matricea setului de date (dată de relaţia (5.236)) pe care o rescriem sub forma:

$$X = \begin{bmatrix} | & | & & | \\ | & | & & | \\ x_1 & x_2 & & x_d \\ | & | & & | \\ | & | & & | \end{bmatrix} \quad (6.186)$$

Vom presupune, în mod implicit, că numărul de linii K al matricii X este cel puţin egal cu numărul de coloane d – condiţie satisfăcută în cazul problemelor reale în care dimensiunea unui vector aleator de trăsături este mai mică decât numărul vectorilor de trăsături ce compun setul de date –, şi, în plus, coloanele sunt independente. Această ultimă condiţie ne ajută în construirea, cu ajutorul procedurii Gram-Schmidt de ortogonalizare, a unui set de d vectori ortonormaţi q_1, q_2, …, q_d, ce vor avea acelaşi cardinal. Această procedură este descrisă în cele ce urmează.

Într-un prim pas se alege primul vector pe care ulterior îl ortonormăm. Acest vector q_1 este o versiune normalizată a primei coloane:

$$q_1' = x_1, \quad q_1 = q_1' \,/\, \| q_1' \| \quad (6.187)$$

Pentru determinarea următorului vector din mulţimea vectorilor ortonormaţi, ne folosim de cea de a doua coloană a matricei X din care s-a înlăturat contribuţia acesteia pe direcţia lui q_1:

$$q_2' = x_2 - \underbrace{\left(x_2^{*T} q_1\right) q_1}_{proiectia\ \ lui\ x_2\ \ pe\ directia\ q_1} \quad (6.188)$$

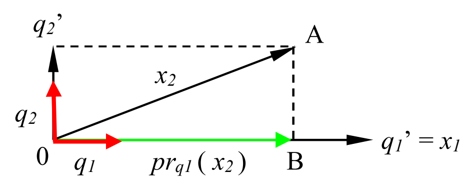

Figura 6.10. Reprezentarea grafică a principiului ortogonalizării

Normalizând se obţine în final şi al doilea vector ortonormat:

$$q_2 = q_2' \,/\, \| q_2' \| \quad (6.189)$$

Observaţia 6.19: Pentru a obţine proiecţia pe q_1 a vectorului x_2, notată cu $pr_{q1}(x_2)$ – **Figura 6.10**, ştim că:

$$\begin{aligned} pr_{q1}(x_2) &= \| OB \| = \| OA \| \cos(\alpha: x_2, q_1) \\ &= \| x_2 \| \cos(\alpha: x_2, q_1) \quad (6.190) \end{aligned}$$

dar

$$\cos(\alpha : x_2, q_1) = \frac{x_2^{*T} q_1}{\|x_2\| \cdot \|q_1\|} \quad (6.191)$$

Ştiind că vectorul q_1 face parte dintr-un set de vectori ortonormaţi (fiind, deci, de lungime 1) şi introducând relaţia (6.191) în (6.190) obţinem:

$$pr_{q1}(x_2) = x_2^{*T} q_1 \quad (6.192)$$

În general al l-lea vector din setul vectorilor ortonormaţi este obţinut din a l-a coloană a matricii X, văzută ca un vector, prin înlăturarea tuturor componentelor acestui vector de pe direcţiile vectorilor ortonormali anterior găsiţi:

$$q_l^{'} = x_l - \sum_{i=1}^{l-1} \left(x_l^{*T} q_i\right) q_i; \quad q_l = q_l^{'} / \left\| q_l^{'} \right\| \quad (6.193)$$

În ultima fază se normalizează şi acest vector:

$$q_l = q_l' \,/\, \| q_l' \| \quad (6.194)$$

Relaţia (6.193) este aplicată în continuare pentru orice index $l \leq d$.

Datorită modului de obţinere se observă că vectorii q_1 şi q_2 se găsesc în acelaşi plan definit de x_1 şi x_2, vezi **Figura 6.10**. În mod similar q_1, q_2, …, q_l se găsesc în acelaşi spaţiu definit de x_1, x_2, …, x_l. În concluzie putem scrie:

$$\underbrace{\begin{bmatrix} | & | & & | \\ | & | & & | \\ q_1 & q_2 & & q_d \\ | & | & & | \\ | & | & & | \end{bmatrix}}_{Q} = \underbrace{\begin{bmatrix} | & | & & | \\ | & | & & | \\ x_1 & x_2 & & x_d \\ | & | & & | \\ | & | & & | \end{bmatrix}}_{X} \cdot \underbrace{\begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1d} \\ 0 & \rho_{22} & \cdots & \rho_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \rho_{dd} \end{bmatrix}}_{R^{-1}} \quad (6.195)$$

În relaţia (6.195) R^{-1} este o matrice superior triunghiulară. Această ultimă ecuaţie poate fi rescrisă sub următoare formă, identică cu cea a unei factorizări QR:

$$X = QR = \begin{bmatrix} | & | & & | \\ | & | & & | \\ q_1 & q_2 & \cdots & q_d \\ | & | & & | \\ | & | & & | \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1d} \\ 0 & r_{22} & \cdots & r_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{dd} \end{bmatrix} \quad (6.196)$$

În relaţia (6.196) matricea R este, de asemenea, o matrice superior triunghiulară în special datorită implicaţiilor generate de relaţiile (6.178) şi (6.179) – relaţii ce dovedesc că o matrice triunghiulară de orice tip în urma aplicării inversei va deveni tot o matrice triunghiulară de acelaşi tip.

După cum am prezentat şi anterior, în introducerea acestui subcapitol, relaţia (6.196) este doar o posibilă formă a factorizării QR; în această factorizare matricea datelor este egală cu produsul dintre o matrice dreptunghiulară, ale cărei coloane sunt ortogonale şi o altă matrice pătratică superior triunghiulară. Această ecuaţie are interpretarea dată de observaţia că x_1 are aceeaşi direcţie cu q_1, x_1 şi x_2 se găsesc în planul determinat de q_1 şi q_2, iar $x_1, x_2, \dots, x_l$ se află în subspaţiul determinat de $q_1, q_2, \dots, q_l$ pentru $l \leq d$. Coeficienţii r_{ij} sunt daţi de:

$$r_{ij} = q_i^{*T} x_j \qquad (6.197)$$

În mod similar definiţiei şi relaţiei (6.185) putem obţine o altă descompunere tot de tip QR dar care în acest caz este formată dintr-o matrice unitară Q de dimensiune m x m şi o matrice m x n superior triunghiulară. Conform acestei descompuneri de tip QR se vor obţine de asemenea două matrici de dimensiuni mai mari decât cele din relaţia (6.196) dar care le vor conţine pe acele două – aceste matrici au fost prezentate anterior (matricea Q şi matricea R). Această descompunere ia forma:

$$X = \underbrace{[Q\; Q_0]}_{Q_1} \cdot \underbrace{\begin{bmatrix} R \\ 0 \end{bmatrix}}_{R_1} =$$

$$= \begin{bmatrix} | & | & & | & | & & | \\ | & | & & | & | & & | \\ q_1 & q_2 & \cdots & q_d & q_{d+1} & \cdots & q_K \\ | & | & & | & | & & | \\ | & | & & | & | & & | \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1d} \\ 0 & r_{22} & \cdots & r_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{dd} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \qquad (6.198)$$

În relaţia (6.198) Q_1 este o matrice de tipul K x K în timp ce matricea R_1 are dimensiunea K x d (evident că în situaţia problemelor reale avem întotdeauna $K > d$). În sistemul (6.198) coloanele $q_{d+1}, \dots, q_K$ sunt o mulţime de vectori ortonormali, ortogonali pe vectorii $q_1, \dots, q_N$. Astfel, primul termen al produsului este, de asemenea, o matrice unitară, în timp ce cea de a doua matrice (al doilea termen – R_1) este o matrice superior triunghiulară.

3. Obţinerea descompunerii LU prin factorizarea QR

Factorizarea QR poate fi utilizată pentru obţinerea descompunerii *Cholesky* fără a mai calcula matricea de corelaţie. Matrica de corelaţie va fi estimată cu ajutorul relaţiei (5.238). Produsul $X^{*T}X$ poate fi scris folosindu-ne de relaţia (6.196) şi de proprietatea matricii Q (matrice ortonormată, având coloanele ortogonale) în forma:

$$X^{*T}X = R^{*T}Q^{*T}Q\,R = R^{*T}R \quad (6.199)$$

Acest produs poartă numele de descompunerea *Cholesky*. În cazul în care $\hat{R}_x$ este definită în forma (5.238), atunci din (6.199) avem:

$$\hat{R}_x = \frac{1}{K}R^{*T}R \quad (6.200)$$

Descompunând matricea $\hat{R}_x$ în forma de mai jos:

$$\hat{R}_x = LD_LL^{*T} \quad (6.201)$$

şi ţinând cont ca descompunerea este unică, rezultă:

$$R^{*T} = \sqrt{K}LD_L^{1/2} \quad (6.202)$$

Toate aceste relaţii ce au drept obiectiv determinarea termenilor descompunerii (6.201) sunt oarecum similare demersului prezentat prin intermediul relaţiilor (6.134), (6.135), (6.136) şi (6.137).

Din relaţia (6.202) rezultă în mod direct relaţiile:

$$D_L^{1/2} = \frac{1}{\sqrt{K}}diag(R) \quad (6.203)$$

şi

$$L = \frac{1}{\sqrt{K}}R^{*T}D_L^{-1/2} \quad (6.204)$$

Evident că ulterior se poate calcula D_L cu:

$$D_L = \left(D_L^{1/2}\right)^2 \quad (6.205)$$

6.6.5. Descompunerea superior-inferior triunghiulară

Matricele de corelaţie sau de covarianţă pot fi factorizate într-o formă alternativă, printr-o descompunere superior-inferior triunghiulară. Această descompunere este de forma:

$$C_x = U \cdot D_U \cdot U^{*T} \quad (6.206)$$

unde U este o matrice superior triunghiulară unitara în timp ce matricea D_U este una diagonală. În general U nu este aceiaşi matrice cu L^{*T}, din descompunerea inferior-superior triunghiulară. În mod similar D_U este diferită de D_L.

Din relaţia (6.206) avem:

$$U^{-1} R_x (U^{-1})^{*T} = D_U \quad (6.207)$$

În această situaţie transformata utilizată pentru diagonalizarea matricii de covarianţă este:

$$y = U^{-1} x \quad (6.208)$$

care determină obţinerea unui vector cu componente ortogonale, iar momentele de ordin doi ale componentelor (varianţa) fiind chiar elementele de pe diagonala principală a matricii D_U. Deoarece U^{-1} este o matrice superior triunghiulară, această transformare este una necauzală a vectorului aleator x.