Cursul 1

Problematica Algebrei liniare. Generalităţi despre matrici. Aplicaţii

1.1 Problematica Algebrei liniare

[1] Algebra liniară este un domeniu al matematicii care studiaza metode de calcul matricial necesare pentru a extrage informaţie din date şi furnizează suportul teoretic, tehnici şi instrumente pentru dezvoltarea a numeroşi algoritmi.

Cursul introduce noţiunile şi metodele de bază ale algebrei liniare, relevante pentru Computer Science (CS).

Metodele matriciale pe care le studiem sunt folosite în grafica 2D şi 3D, gaming, animaţie, cautare pe WEB, extragerea informaţiei din documente, controlul mişcării roboţilor, recunoaşterea şi identificarea obiectelor, analiza reţelelor de calculatoare sau a reţelelor sociale (Twitter, Facebook, Google +, Linkedin), designul sistemelor de recomandare de produse şi continut internet, etc. Ori de cate ori sunăm pe cineva la telefon, facem poze cu telefonul mobil, procesam o imagine în Photoshop, jucam un video-game sau urmărim un film cu efecte digitale, folosim de fapt tehnologii bazate pe metode ale algebrei liniare.

Notaţiile şi modul de manipulare al matricilor pe care le folosim în cadrul cursului sunt tipice pentru literatura destinată CS.

1.2 Generalităţi despre matrici

Noţiunea de matrice este fundamentală în algebra liniară. În cadrul cursului vor interveni preponderent matrici de elemente reale şi vom aborda doar câteva noţiuni şi proprietăţi ale matricilor cu elemente complexe, respectiv în $\mathbb{Z}_2$, corpul claselor de resturi modulo 2.

Reamintim că $\mathbb{Z}_2$ este o mulţime de doua elemente (biţi), $\{0, 1\}$ şi operatiile de adunare şi inmulţire sunt definite astfel:

$+$	0	1
0	0	1
1	1	0

.	0	1
0	0	0
1	0	1

[1] Cursul este editat în LaTeX `http://en.wikipedia.org/wiki/LaTeX`, `http://www.latex-project.org/`

Notăm cu $\mathbb{R}^{m\times n}$, $\mathbb{C}^{m\times n}$ mulţimea matricilor de elemente reale, respectiv complexe şi cu $\mathbb{Z}_2^{m\times n}$ mulţimea matricilor binare (de elemente 0 şi 1), de tip $m \times n$, adică mulţimea matricilor de forma:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \text{unde } a_{ij} \in \mathbb{K}, \;\; \mathbb{K} = \mathbb{R}, \mathbb{C} \text{ sau } \mathbb{Z}_2 \tag{1.1}$$

Notăm coloana j a matricii A prin $\mathbf{A}[:, \mathbf{j}]$, iar linia i prin $\mathbf{A}[\mathbf{i}, :]$. Această notaţie este tipică în limbajul `Python`[2].

De exemplu,

$$A[:, 2] = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \text{iar} \quad A[3, :] = [a_{31}\; a_{32}\; \ldots a_{3n}]$$

Indexarea elementelor unei matrici, $A = (a_{ij})$, începe în matematică de la 1, în timp ce în limbajele evoluate de programare, `C/C++, Java, Python`, de la 0, adică:

$$A = \begin{bmatrix} a_{00} & a_{01} & \cdots & a_{0,n-1} \\ a_{10} & a_{11} & \cdots & a_{1,n-1} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m-1,0} & a_{m-1,1} & \cdots & a_{m-1,n-1} \end{bmatrix}. \tag{1.2}$$

De aceea trebuie acordată mare atentie implementării algoritmilor ce conţin ca date, matrici, şi prezentarea algoritmului este facută în notaţie matematică!

Oricărei matrici $A \in \mathbb{R}^{m\times n}$ $A = (a_{ij})$, i se asociază **transpusa** sa, care este o matrice de n linii şi m coloane, notată, A^T. Matricea A^T se obţine din A, scriind liniile lui A pe coloanele corespunzătoare din A^T.

Mai precis notând elementele matricii A^T cu $b_{k\ell}$, avem că $b_{k\ell} = a_{\ell k}$.

Exemplul 1. Transpusa matricii

$$A = \begin{bmatrix} -1 & 2 & 0 \\ 3 & 5 & -4 \\ 1 & -2 & 5 \\ 7 & -4 & 0 \end{bmatrix} \in \mathbb{R}^{4\times 3}$$

este matricea $A^T \in \mathbb{R}^{3\times 4}$, unde:

$$A^T = \begin{bmatrix} -1 & 3 & 1 & 7 \\ 2 & 5 & -2 & -4 \\ 0 & -4 & 5 & 0 \end{bmatrix}$$

[2] `https://en.wikipedia.org/wiki/Python_(programming_language)`

Proprietate a operatorului de transpunere:

$$(A^T)^T = A$$

În mod deosebit vom studia matrici pătratice, care au numărul liniilor egal cu numărul coloanelor. Notăm $A \in \mathbb{R}^{n\times n}$ o matrice pătratică cu elemente reale.

Definiţia 1.2.1 *O matrice pătratică $A \in \mathbb{R}^{n\times n}$, $A = (a_{ij})$, cu proprietatea că $A^T = A$ sau echivalent $a_{ij} = a_{ji}$, oricare ar fi $i, j = \overline{1, n}$ se numeşte matrice simetrică.*

Exemplul 2. Matricea următoare este o matrice simetrică:

$$A = \begin{bmatrix} -2 & 3 & 1 \\ 3 & 0 & -5 \\ 1 & -5 & 7 \end{bmatrix}$$

Observaţia 1.2.1 *În abordările ulterioare din curs, o matrice generală de tipul (1.1) va fi reprezentată de coloanele sale notate fie ca mai sus, folosind regula* `Python`*, fie notate $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, unde*

$$\mathbf{v}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \qquad j = \overline{1, n}$$

şi vom nota

$$A = [\mathbf{v}_1|\mathbf{v}_2| \ldots \mathbf{v}_n]$$

Spunem că matricea A se obţine prin **concatenarea** coloanelor $\mathbf{v}_1$, $\mathbf{v}_2, \ldots, \mathbf{v}_n$.

1.2.1 Exemple de matrici folosite în diverşi algoritmi din CS

În continuare dăm exemple de matrici ce le vom folosi pentru a ilustra diverse noţiuni, proprietăţi şi algoritmi pe care îi vom studia pe parcurs.

1. Matricea de adiacenţă/conectivitate a unei reţele.

O reţea de calculatoare sau o reţea socială (Facebook, de exemplu) constă dintr-un număr finit de noduri (calculatoare in primul caz şi persoane în al doilea) şi conexiunile dintre ele. Modelul matematic al unei reţele este graful. Un graf constă dintr-o pereche $G = (V, L)$, unde $V = \{1, 2, \ldots, m\}$ este mulţimea nodurilor (a entităţilor ce se interconectează), iar L mulţimea arcelor (linkurilor, conexiunilor dintre noduri). L este definit matematic ca o mulţime de perechi de noduri (i, j) între care există conexiune.

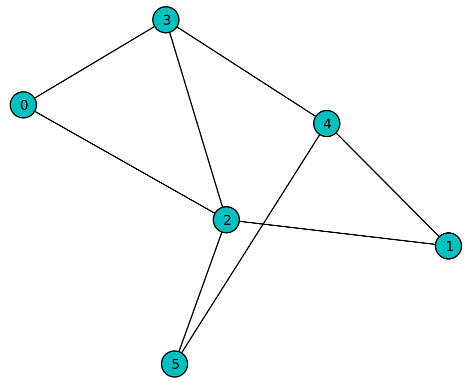

Fig.1.1: Graf neorientat ce ilustrează conexiunile într-o reţea.

Graful din figura Fig.1.1 are mulţimea nodurilor indexate astfel $V = \{0, 1, 2, 3, 4, 5\}$ şi mulţimea arcelor,

$$L = \{(0, 2), (0, 3), (1, 2), (1, 4), (2, 3), (2, 5), (3, 4), (4, 5)\}$$

Unei reţele (graf) cu m noduri i se asociază matricea de conectivitate sau adiacenţă, care este o matrice pătratică, $C \in \mathbb{R}^{m\times m}$, de elemente c_{ij}, unde:

$$c_{ij} = \begin{cases} 1 & \text{dacă există conexiune între nodurile i şi j, adică } (i, j) \in L \\ 0 & \text{dacă nodurile i şi j nu sunt conectate} \end{cases}$$

Un graf pentru care nu se precizează şi sensul conexiunii, adică arc de la i spre j, ci doar perechea de noduri conectate, se numeşte **graf neorientat**. Este graf neorientat graful asociat grupului tău de prieteni pe Facebook. Dacă persoana i este conectată cu j, atunci obligatoriu şi j este conectat cu i.

Exemplul 3. Considerăm o reţea socială formată din 6 persoane, codificate $0, 1, 2, 3, 4, 5$, şi ilustrată în graful neorientat din Fig.1.1: Matricea de conectivitate/adiacenţă a reţelei este:

$$C = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Matricea de adiacenţă a unui graf neorientat este simetrică, adică $a_{ij} = a_{ji}$, oricare ar fi $i, j = \overline{1, m}$.

Pe lângă grafuri neorientate există şi grafuri orientate, $G = (V, L)$, unde din nou V este mulţimea nodurilor, iar arcele, numite linkuri, sunt orientate. Un exemplu concludent este graful WEB. Dacă există link în pagina i spre pagina j, nu e obligatoriu ca şi pagina j să aibă link spre i. Scriem în acest caz că $(i, j) \in L$, dar $(j, i) \notin L$.

Să considerăm acum o reţea în care conexiunile dintre noduri sunt orientate. Nodurile sunt câteva pagini WEB, iar arcele orientate dintre acestea sunt hyperlinkurile. În figura 1.2 avem o reţea formată din 6 noduri şi linkuri între ele (sensul linkului nu este ilustrat printr-o săgeată, ci printr-un segment pe arc, in format bold, modalitate specifică de vizualizare a grafurilor orientate în `networkx` (`http://networkx.github.io/`), un pachet în `Python`).

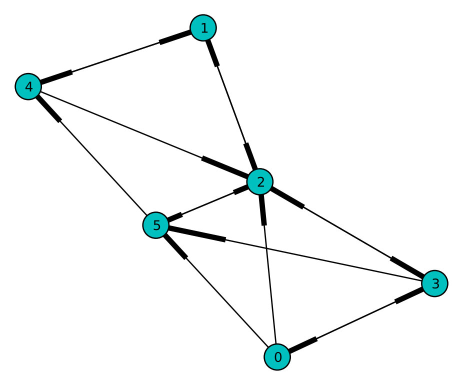

Fig.1.2: Graf orientat ce ilustrează linkurile dintre pagini într-o reţea WEB.

Matricea de adiacenţă a grafului, numită matricea hyperlink este:

$$C = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Observăm că aceasta nu mai este o matrice simetrică.

Pe parcursul cursului vom învăţa tehnici şi algoritmi pentru a extrage informaţie din matricea de conectivitate a unui graf neorientat, de a determina nodurile ”importante” din reţea, de a clusteriza, grupa nodurile ce au aceeasi intensitate a legăturilor intre ele, respectiv de a realiza ierarhizarea paginilor WEB în funcţie de autoritatea lor, înglobată în structura inlinkurilor. Mai precis vom prezenta teoria ce stă la baza algoritmului PageRank–Google, de ierarhizare a paginilor WEB.

2. Matricea term–by–document.

Information Retrieval (IR) este un subdomeniu în Computer science care dezvoltă tehnici de stocare a informaţiei dintr-o colecţie de documente în format electronic şi de selectare a documentelor relevante pentru o cerere (*query*) a unui utilizator. Sistemele bazate pe calculator de extragere a informaţiei relevante pentru necesităţile de informare ale utilizatorilor se numesc sisteme IR.

Orice motor de căutare are un sistem IR. Bibliotecile, portalurile diverselor ziare/reviste pun la dispoziţie o aplicaţie de căutare.

Documentele în format electronic sunt analizate şi se constituie o listă de termeni prin filtrare (nu se includ în listă conjuncţii, prepoziţii, semne de punctuaţie). Dacă m este numărul de termeni rezultaţi în urma procesului de filtrare a celor n documente din colecţie, atunci se analizează fiecare document $D_1, D_2, \dots, D_n$ şi documentului D_j i se asociază o matrice coloană $d_j \in \mathbb{R}^m$, ale cărui coordonate $p_{ij} \geq 0$ reprezintă ponderea termenului i în documentul D_j:

$$d_j = \begin{bmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{mj} \end{bmatrix} \in \mathbb{R}^m$$

d_j reprezintă conţinutul informaţiei din documentul D_j.

Matricea de m linii şi n coloane, $D = [d_1|d_2|\dots|d_n]$, se numeşte în IR, *term by document matrix*:

$$D = \begin{array}{c} \text{termen } 1 \rightarrow \\ \text{termen } 2 \rightarrow \\ \vdots \\ \text{termen } m \rightarrow \end{array} \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \dots & \vdots \\ p_{m1} & p_{m2} & \dots & p_{mn} \end{bmatrix}$$

Pe linia i sunt înregistrate ponderile termenului i în documentele $D_1, D_2, \dots, D_n$ (care de fapt sunt frecventele de apariţie ale termenului i în documente; detalii intr-un curs ulterior), iar pe coloane avem "vectorii" d_j ce conţin ponderile termenilor $1, 2, \dots m$, în documentul D_j, $j = \overline{1, n}$.

În cursul 4-5 vom învăţa în ce fel matricea term–by–document este folosită pentru a găsi documentele relevante pentru cererea unui utilizator.

3. Matricea user-produs folosită pentru a recomanda produse, conţinut internet, etc.

În paginile WEB destinate comerţului electronic (`e-commerce`) sau în paginile de închiriere online de filme, muzică, conţinut internet, se includ cel mai adesea şi sisteme de recomandare. Sistemul de recomandare, pe baza unor tehnici de inteligenţa artificială, face recomandări personalizate de produse sau servicii, în cursul unei interacţiuni live cu clienţii, într-un site WEB al producătorului (ofertantului de servicii).

Un astfel de sistem creează o bază de date a preferintelor consumatorilor pentru o mulţime de produse. Profilul unui consumator nou, sa-i zicem Radu, este comparat cu cel al utilizatorilor din baza de date pentru a descoperi afinităţi de consum cu unii dintre ei. Pe baza comparaţiilor i se fac recomandări şi lui Radu.

Metoda cel mai mult folosită este aşa numita filtrare colaborativă, bazată pe rating. Ea s-a dezvoltat datorită interesului utilizatorilor tineri în folosirea sistemelor de recomandare a filmelor, muzicii, nouăţilor.

La baza filtrării colaborative stă aşa numita matrice de recomandare. Notăm cu R matricea asociată unei comunităţi de m clienţi/useri şi n produse. Fiecare client acordă un rating (o notă) pe o scară fixată, $\{1, 2, 3, \dots, r_{max}\}$, produselor achiziţionate sau filmelor/cărţilor vizionate/citite, serviciilor hoteliere sau de turism la care a apelat. Astfel r_{ij} este ratingul clientului i pentru produsul j. Dacă clientul i nu a acordat un rating produsului j, în mod uzual se setează $r_{ij} = 0$. Matricea R este o matrice rară (cu multe ratinguri lipsă), pentru că un client nu a cumpărat toate produsele, nu a vizionat toate filmele, etc, deci nu a acordat tuturor un rating.

Pe baza observaţiilor înregistrate în matricea rating, R, considerată ca mulţime de antrenament, un sistem de recomandare generează estimatori $\hat{r}_{ij}$ pentru toate perechile $(i, j) = (client, produs)$ ce aveau întrări nule în R şi apoi face recomandări clienţilor pentru produse neconsumate încă, pe baza similarităţii preferinţelor lor cu ai altor useri.

Exemplu de matrice recomandare pentru un site ce închiriază filme:

Fie R matricea rating asociată unui număr de utilizatori, care au vizionat unele dintre cele 8 filme puse la dispoziţie:

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
Alex	0	0	4	0	3	0	5	0
Gigel	5	0	0	3	0	4	0	0
Ina	0	0	0	4	4	0	2	0
Radu	0	4	3	0	2	0	0	5
Cici	3	0	0	4	0	5	0	0
Răvz	0	0	5	0	3	0	3	0

In cursul 12 vom învăţa cum reuşeste sistemul de recomandare să determine estimatorii $\hat{r}_{ij}$ şi cum se generează recomandări de noi filme pentru utilizatori.

End Exemple

Produsul a două matrici. Reamintim că pentru orice două matrici A, B cu elemente din acelaşi corp $\mathbb{K}$, cu particularitatea că numărul de coloane al primei matrici coincide cu numărul de linii al celei de-a doua, adică A este de tip $m \times p$, iar B de tip $p \times n$, definim **matricea produs**, ca fiind matricea $C = AB$, de elemente $c_{ij} \in \mathbb{K}$, cu

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}, \tag{1.3}$$

adică:

$$\begin{bmatrix} c_{11} & \dots & c_{1j} & \dots & c_{1n} \\ \vdots & \dots & \vdots & \dots & \vdots \\ c_{i1} & \dots & \mathbf{c_{ij}} & \dots & c_{in} \\ \vdots & \dots & \vdots & \dots & \vdots \\ c_{m1} & \dots & c_{mj} & \dots & c_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1k} & \dots & a_{1p} \\ \vdots & \dots & \vdots & \dots & \vdots \\ \mathbf{a_{i1}} & \cdots & \mathbf{a_{ik}} & \cdots & \mathbf{a_{ip}} \\ \vdots & \dots & \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mk} & \dots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \dots & \mathbf{b_{1j}} & \dots & b_{1n} \\ \vdots & \dots & \vdots & \dots & \vdots \\ b_{k1} & \dots & \mathbf{b_{kj}} & \dots & b_{kn} \\ \vdots & \dots & \vdots & \dots & \vdots \\ b_{p1} & \dots & \mathbf{b_{pj}} & \dots & b_{pn} \end{bmatrix}$$

Fixând o coloană j, din relaţia (1.3) rezultă că elementele acesteia se exprimă astfel:

$$\begin{bmatrix} c_{1j} \\ c_{2,j} \\ \vdots \\ c_{mj} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{bmatrix}$$

sau concentrat, $C[:, j] = A \cdot B[:, j]$.

Astfel matricea produs $C = AB$ este reprezentată de coloanele sale astfel:

$$C = AB = [A \cdot B[:, 1] | A \cdot B[:, 2] | \cdots | A \cdot B[:, n]]$$

Relaţia dintre transpusa produsului şi produsele transpuselor a două matrici

Dacă $A \in \mathbb{K}^{m \times n}$, $B \in \mathbb{K}^{n \times p}$, atunci:

$$(AB)^T = B^T A^T$$

De interes special va fi produsul dintre o matrice $A \in \mathbb{K}^{m \times n}$ şi o matrice coloană $x \in \mathbb{R}^{n \times 1}$: $Ax = b$, $b \in \mathbb{R}^{m \times 1}$.

Exprimarea produsului $\mathbf{Ax}$

Produsul Ax, unde $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ este egal cu:

$$Ax = x_1 A[:, 1] + x_2 A[:, 2] + \cdots + x_n A[:, n]$$

Într–adevăr, folosind formula produsului avem:

$$b_i = \sum_{k=1}^{n} a_{ik} x_k = \sum_{k=1}^{n} x_k a_{ik}$$

ceea ce detaliat se exprimă astfel:

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = x_1 A[:, 1] + x_2 A[2, :] + \cdots + x_n A[:, n]$$

1.3 Produsul a două matrici exprimat ca suma de produse exterioare

Produsul unei matrici coloana, $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \in \mathbb{R}^{m\times 1}$ cu o matrice linie, $Y = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} \in \mathbb{R}^{1\times n}$ se numeşte produs exterior (*outer product* în limba engleză):

$$X \cdot Y = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_n \\ x_2y_1 & x_2y_2 & \dots & x_2y_n \\ \vdots & \vdots & \dots & \vdots \\ x_my_1 & x_my_2 & \dots & x_my_n \end{bmatrix} \in \mathbb{R}^{m\times n}$$

În continuare dăm o modalitate de calcul a produsului a două matrici folosită în Big Data `https://en.wikipedia.org/wiki/Big_data`.

1. Produsul a doua matrici $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times p}$, $C = AB$, se exprima ca suma produselor exterioare a perechilor de coloane din A cu liniile corespunzatoare din B, mai precis:

$$AB = A[:,1]B[1,:] + A[:,2]B[2,:] + \cdots A[:,n]B[n,:]$$

Demonstraţie: Notam cu C suma produselor exterioare:

$$C = A[:,1]B[1,:] + A[:,2]B[2,:] + \cdots A[:,n]B[n,:]$$

Vom arăta că un element arbitrar, c_{ij}, al matricii C este egal cu $a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$, adica cu elementul din linia i, coloana j a produsului de matrici AB, efectuat după regula de la liceu.

Dacă

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad \text{şi} \quad B = \begin{bmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1p} \\ \vdots & \dots & \vdots & \dots & \vdots \\ b_{k1} & \dots & b_{kj} & \dots & b_{kp} \\ \vdots & \dots & \vdots & \dots & \vdots \\ b_{n1} & \dots & b_{nj} & \dots & b_{np} \end{bmatrix},$$

atunci

$$C = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{i1} \\ \vdots \\ a_{m1} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1j} & \dots & b_{1p} \end{bmatrix} + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{i2} \\ \vdots \\ a_{m2} \end{bmatrix} \begin{bmatrix} b_{21} & b_{22} & \dots & b_{2j} & \dots & b_{2p} \end{bmatrix} +$$

$$\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{in} \\ \vdots \\ a_{mn} \end{bmatrix} \begin{bmatrix} b_{n1} & b_{n2} & \dots & b_{nj} & \dots & b_{np} \end{bmatrix}$$

Elementul c_{ij} al matricii C se obtine adunand elementele din linia i, coloana j a fiecărui produs exterior. Deci:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = (AB)_{ij}$$

$\square$

Exemplul 4. Să exprimăm produsul următoarelor două matrici ca o sumă de produse exterioare:

$$AB = \begin{bmatrix} -2 & 1 \\ 0 & 3 \\ -5 & -1 \end{bmatrix} \begin{bmatrix} -3 & 1 \\ 7 & -2 \end{bmatrix} =$$

$$\begin{bmatrix} -2 \\ 0 \\ -5 \end{bmatrix} \begin{bmatrix} -3 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} \begin{bmatrix} 7 & -2 \end{bmatrix} =$$

$$\begin{bmatrix} 6 & -2 \\ 0 & 0 \\ 15 & -5 \end{bmatrix} + \begin{bmatrix} 7 & -2 \\ 21 & -6 \\ -7 & 2 \end{bmatrix} = \begin{bmatrix} 13 & -4 \\ 21 & -6 \\ 8 & -3 \end{bmatrix}$$

Matrici particulare

• Matricea de tip $m \times n$ ce are toate elementele 0 se numeşte **matricea nulă** şi se notează cu O_{mn};

• Matricea pătratică de tip $n \times n$,

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

având 1 pe diagonala principală şi zero în rest, i.e. $a_{ii} = 1, \forall\ i = \overline{1,n}$ şi $a_{ij} = 0$, pentru orice $i \neq j$, se numeşte **matricea unitate**.

Un element arbitrar (generic) al matricii unitate se notează δ_{ij}, unde δ_{ij}, numit simbolul lui Kronecker, se defineşte astfel:

$$\delta_{ij} = \begin{cases} 1 & \text{dacă } i = j \\ 0 & \text{dacă } i \neq j \end{cases}$$

Oricare ar fi A o matrice pătratică de tip $n \times n$, avem $AI_n = I_nA = A$, adică I_n este elementul unitate faţă de înmulţirea matricilor pătratice.

Notăm cu $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ coloanele matricii unitate:

$$\mathbf{e}_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow\ j \\ \\ \\ \end{array}$$

Astfel avem reprezentarea $I_n = [\mathbf{e}_1|\mathbf{e}_2|\ldots|\mathbf{e}_n]$.

Matrici permutare

Reamintim că o permutare a mulţimii ordonate $(1, 2, \ldots, n)$ este o aplicatie bijectivă, $\pi : (1, 2, \ldots, n) \to (1, 2, \ldots, n)$ şi $\pi(i)$ se notează π_i, $i = \overline{1,n}$. Vom nota o permutare astfel:

$$\begin{pmatrix} 1 & 2 & \ldots & i & \ldots & n \\ \downarrow & \downarrow & \ldots & \downarrow & \ldots & \downarrow \\ \pi_1 & \pi_2 & \ldots & \pi_i & \ldots & \pi_n \end{pmatrix}$$

Exemplu:

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix},$$

adică $\pi_1 = 3, \pi_2 = 1, \pi_3 = 4, \pi_4 = 1$.

O notaţie şi mai simplă pentru o permutare π este $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$.

Definiţia 1.3.1 *Fie* $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ *o permutare a mulţimii* $(1, 2, \ldots, n)$. **Matricea permutare** *asociată permutării* π, *şi notată* P_π, *este matricea pătratică din* $\mathbb{R}^{n\times n}$ *ce se obţine din matricea unitate,* I_n, *prin permutarea liniilor lui* I_n *conform permutarii* π. *Cu alte cuvinte dacă notăm cu* $\ell_1, \ell_2, \ldots, \ell_n$ *liniile din* I_n, *atunci în liniile* $1, 2, \ldots, n$ *ale matricii* P_π *se scriu liniile* $\ell_{\pi_1}, \ell_{\pi_2}, \ldots, \ell_{\pi_n}$ *ale matricii unitate.*

Exemplul 5. Matricea permutare asociată permutării

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$$

se obţine din

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

scriind în linia 1 din P_π, linia 3 din I_4, în linia 2 din P_π, linia 1 din I_4, etc:

$$P_\pi = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Notând cu p_{ij} elementele matricii permutare $P_\pi \in \mathbb{R}^{n\times n}$ şi cu δ_{ij} elementele matricii unitate, I_n, avem următoarea relaţie dintre cele două matrici, şi anume elementele din linia i ale lui P_π sunt elementele din linia π_i ale matricii unitate, ceea ce revine la a scrie că:

$$p_{ij} = \delta_{\pi_i j}, \ \forall\, j = \overline{1, n}$$

Cum pe linia π_i din matricea I_n doar elementul $\delta_{\pi_i\pi_i}$ este egal cu 1, restul fiind zero, rezultă că pe linia i a matricii P_π este egal cu 1, doar elementul $p_{i\pi_i}$:

$$p_{ij} = \delta_{\pi_i j} = \begin{cases} 1 & \text{dacă } j = \pi_i \\ 0 & \text{dacă } j \neq \pi_i \end{cases} \tag{1.4}$$

Să evidenţiem câteva proprietăţi ale matricilor permutare:

▶ *Produsul dintre o matrice permutare* P_π *şi o matrice coloană este:*

$$P_\pi \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{\pi_1} \\ x_{\pi_2} \\ \vdots \\ x_{\pi_n} \end{bmatrix}$$

adică matricea P_π *are ca efect permutarea liniilor matricii coloană conform permutării* π.

Demonstraţie: Notăm cu Y matricea produs $P_\pi X$,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Din $Y = P_\pi X$, un element arbitrar al matricii produs, y_i, se exprimă conform relaţiei de înmulţire $P_\pi X$ astfel:

$$y_i = \sum_{k=1}^{n} p_{ik} x_k = p_{i1}x_1 + p_{i2}x_2 + \cdots + p_{in}x_n$$

Dar $p_{ik} = \delta_{\pi_i k}$, $\forall\, k = \overline{1, n}$. Înlocuind mai sus avem:

$$y_i = \underbrace{\delta_{\pi_i 1}}_{=0} x_1 + \underbrace{\delta_{\pi_i 2}}_{=0} x_1 + \cdots + \underbrace{\delta_{\pi_i \pi_i}}_{=1} x_{\pi_i} + \cdots + \underbrace{\delta_{\pi_i n}}_{=0} x_n = x_{\pi_i}$$

$\square$

Pentru ilustrare folosim matricea permutare definită în Exemplul (5) şi avem:

$$P_{(3,1,4,2)} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_3 \\ x_1 \\ x_4 \\ x_2 \end{bmatrix}$$

Consecinţă. *În particular produsul, $P_\pi A$, al unei matrici permutare P_π, de tip $n \times n$, cu o matrice A de tip $n \times m$, este matricea obţinută din A aplicând permutarea π liniilor sale.*

Într-adevăr, scriind coloanele matricii produs $P_\pi A$ avem că:

$$P_\pi A = [P_\pi A[:, 1] | P_\pi A[:, 2] | \ldots | P_\pi A[:, m]]$$

şi cum o matrice coloană $P_\pi A[:, j]$ este rezultatul aplicării permutării π liniilor din matricea coloană $A[:, j]$, rezultă că $P_\pi A$ are ca efect aplicarea permutării π liniilor matricii A.

Exemplul 6. Să calculăm $P_\pi A$, unde $\pi = (3, 1, 4, 2)$ iar $A \in \mathbb{R}^{4\times 2}$:

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} \\ a_{11} & a_{12} \\ a_{41} & a_{42} \\ a_{21} & a_{22} \end{bmatrix}$$

► O matrice permutare este inversabilă şi inversa ei coincide cu transpusa sa: $P_\pi^{-1} = P_\pi^T$.

Demonstraţie: **Opţional! Demonstraţia doar pentru cei interesaţi să–şi dezvolte raţionamentul.**

În loc să demonstrăm în prealabil că matricea P_π este nesingulară, demonstrăm direct că $P_\pi P_\pi^T = I_n$ ceea ce conduce la $P_\pi^T = P_\pi^{-1}$, deoarece inversa unei matrici A este unica matrice cu proprietatea că produsul la stânga şi la dreapta cu matricea A este matricea unitate. (demonstraţia pentru $P_\pi^T P_\pi = I_n$ se face similar cu demonstraţia $P_\pi P_\pi^T = I_n$).

Fie $P_\pi = (p_{ij})$, $i, j = \overline{1, n}$. Notăm cu $Q = P_\pi^T$, transpusa matricii P_π şi cu q_{ij} un element arbitrar al său. Astfel avem $q_{ij} = p_{ji}$, $i, j = \overline{1, n}$.

Din relaţia de determinare a unui element arbitrar al produsului $R = P_\pi Q$, $R = (r_{ij})$, avem:

$$r_{ij} = \sum_{k=1}^{n} p_{ik} q_{kj} = \sum_{k=1}^{n} \delta_{\pi_i k} q_{kj}$$

Dar dintre elementele $\delta_{\pi_i 1}, \delta_{\pi_i 2}, \dots, \delta_{\pi_i n}$ doar $\delta_{\pi_i \pi_i} = 1$ este nenul (adică pentru $k = \pi_i$), restul fiind zero şi prin urmare:

$$r_{ij} = \underbrace{\delta_{\pi_i \pi_i}}_{=1} \underbrace{q_{\pi_i j}}_{=p_{j\pi_i}} = p_{j\pi_i}$$

Dar din definitia matricii permutare avem că $p_{j\pi_i} = \delta_{\pi_j \pi_i}$. Elementul $\delta_{\pi_j \pi_i}$ este egal cu:

$$\delta_{\pi_j \pi_i} = \begin{cases} 1 & \text{dacă } \pi_i = \pi_j \;\Leftrightarrow\; i = j \\ 0 & \text{dacă } \pi_i \neq \pi_j \;\Leftrightarrow\; i \neq j \end{cases} = \delta_{ij}$$

Prin urmare $r_{ij} = \delta_{ij}$ şi deci $R = P_\pi P_\pi^T = I_n$, adică $P_\pi^T = P_\pi^{-1}$. □

End Opţional

► Fie $A \in \mathbb{R}^{m \times n}$ şi P_π o matrice permutare asociată permutării $\pi = (\pi_1, \pi_2, \dots, \pi_n)$. Matricea produs AP_π^T este o matrice ale cărei coloane se obţin aplicând permutarea π coloanelor lui A.

Demonstraţie: Deoarece ştim că permutarea liniilor unei matrici M, conform unei permutări π, se realizează calculând produsul $P_\pi M$, aplicăm această proprietate matricii A^T. Liniile lui A^T sunt coloanele lui A. Prin urmare matricea $B = P_\pi A^T$ are liniile lui A^T, dar permutate. Daca transpunem matricea B rezultă că B^T va avea coloanele lui A permutate:

$$B^T = (P_\pi A^T)^T = AP_\pi^T$$

□

Exemplul 7. Fie $A \in \mathbb{R}^{3\times 4}$ şi $P_{(3,1,4,2)}$.

$$AP_\pi^T = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} a_{13} & a_{11} & a_{14} & a_{12} \\ a_{23} & a_{21} & a_{24} & a_{22} \\ a_{33} & a_{31} & a_{34} & a_{32} \end{bmatrix}$$

O serie de alte matrici particulare vor fi definite şi studiate pe parcursul cursului.

1.4 Sisteme de ecuaţii liniare. Reducerea unei matrici la forma scară pe linii

O problemă fundamentală în algebra liniară este rezolvarea unui sistem de m ecuaţii algebrice, liniare, cu n necunoscute:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n &= b_m \end{aligned} \tag{1.5}$$

Coeficienţii a_{ij}, b_i, $i = \overline{1,m}$, $j = 1,n$, sunt numere reale, complexe sau din $\mathbb{Z}_2$.

Acestui sistem i se asociază matricea A:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

şi matricea prelungită $\overline{A}$ obţinută prin bordarea matricii A cu coloana termenilor liberi:

$$\overline{A} = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$

Mulţimea soluţiilor unui astfel de sistem poate fi:

1. Mulţimea vidă. În acest caz sistemul se numeşte **incompatibi**.

2. O mulţime cu un singur element (adică sistemul are o **unică soluţie** şi în acest caz sistemul este **compatibil determinat** sau o mulţime cu o **infinitate de** elemente (**soluţii**) şi sistemul se numeşte **compatibil nedeterminat**.

Numărul soluţiilor sistemului (1.5) depinde de rangul matricii A şi al matricii prelungite.

Rangul unei matrice $A \in \mathbb{K}^{m\times n}$ este cel mai mare ordin de determinant nenul ce se poate constitui din elementele de intersecţie a k linii, $i_1 < i_2 < \cdots < i_k$ şi k coloane, $j_1 < j_2 < \cdots < j_k$ ale matricii A, unde $k = 1, 2, \dots, \min(m, n)$.

Teoretic, s-ar calcula determinanţii de ordin $k = \min(m, n)$. Dacă cel puţin unul este diferit de zero, atunci rangul este $r = k$. Dacă toţi determinţii de ordin k sunt zero, se constituie determinanţii de ordin $k - 1$, etc, până se identifică un determinant diferit de zero şi ordinul lui dă rangul matricii.

Teorema Kronecker–Capelli. Sistemul (1.5) este compatibil dacă şi numai dacă rangul matricii A, a sistemului, coincide cu rangul matricii prelungite, $\overline{A} = [A|b]$ (obţinută prin bordarea matricii A, cu coloana termenilor liberi, b).