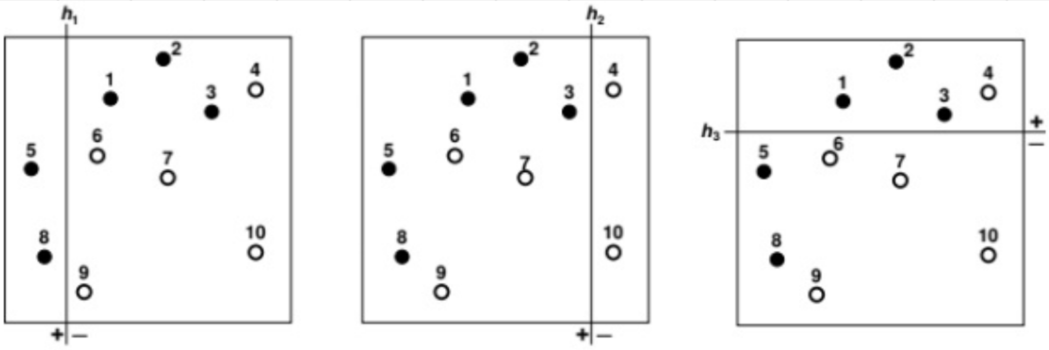
Ada Boost.

61. (Algoritmul AdaBoost: aplicare pe un set de date din $\mathbb{R}^2$)

Liviu Ciortuz, 2021, pornind de la un slide al lui Virgil Pavlu, Northeastern University, USA

Se consideră că aplicăm algoritmul AdaBoost pe dataset-ul din figura alăturată. (Pentru uşurinţa exprimării la calcule, am notat pe figură indicii instanţelor de antrenament, în imediata apropiere a acestora.)
Folosim convenţia noastră obişnuită de notare: simbolul • desemnează instanţe pozitive, iar simbolul ◦ instanţe negative.
La primele trei iteraţii ale algoritmului au fost selectaţi compaşii de decizie h_1, h_2 şi h_3 (în această ordine), aşa cum se indică în figurile de mai jos.



Obiectivul **acestei probleme este să determinăm dacă la sfârşitul celor trei iteraţii algoritmul AdaBoost reuşeşte să clasifice perfect toate instanţele de antrenament.**

a. Calculaţi

– distribuţiile probabiliste corespunzătoare celor 3 iteraţii, adică $D_1(x_i)$, $D_2(x_i)$ şi $D_3(x_i)$, pentru $i = 1, \ldots, 10$;
– pentru fiecare dintre cele 3 iteraţii ($t = 1, 2, 3$): eroarea ponderată la antrenare (ε_t) produsă de compasul de decizie h_t, precum şi ponderea (α_t) asociată ipotezei / compasului de decizie h_t.

Veţi completa tabelele următoare şi veţi indica succint(!) modul în care aţi procedat pentru a ajunge la rezultatele respective.

i	1	2	3	4	5	6	7	8	9	10
$D_1(x_i)$	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10
$D_2(x_i)$	1/6	1/6	1/6	1/14	1/14	1/14	1/14	1/14	1/14	1/14
$D_3(x_i)$	7/66	7/66	7/66	1/22	1/22	1/6	1/6	1/22	1/6	1/22

$$\varepsilon_1 = \frac{3}{10} \qquad \alpha_1 = \frac{1}{2} \ln \frac{\frac{7}{\cancel{10}}}{\frac{3}{\cancel{10}}} = \ln \sqrt{\frac{7}{3}}$$

$$D_2(x_i): \begin{cases} \dfrac{D_1\, i}{2 \cdot \varepsilon_1}, & i \in \{1,2,3\} \\[2ex] \dfrac{D_1\, i}{2 \cdot (1-\varepsilon_1)}, & i \notin \{1,2,3\} \end{cases}$$

$$1,2,3: \frac{\frac{1}{10}}{2 \cdot \frac{3}{10}} = \frac{1}{6} \qquad 4,5,6,\ldots 10: \frac{\frac{1}{\cancel{10}}}{2 \cdot \frac{7}{\cancel{10}}} = \frac{1}{14}$$

$\varepsilon_2 = \frac{3}{14}$ $\qquad \alpha_2 = \frac{1}{2} \ln \frac{\frac{11}{14}}{\frac{3}{14}} = \ln \sqrt{\frac{11}{3}}$

$$D_3(x_i): \begin{cases} \frac{D_2\, i}{2 \cdot \varepsilon_2} & ; \ i \in \{6, 7, 9\} \\ \frac{D_2\, i}{2 \cdot (1 - \varepsilon_2)} & ; \ i \notin \{6, 7, 9\} \end{cases}$$

$6, 7, 9: \ \frac{\frac{1}{14}}{2 \cdot \frac{3}{14}} = \frac{1}{6}$ $\qquad 4, 5, 8, 10: \ \frac{\frac{1}{14}}{2 \cdot \frac{11}{14}} = \frac{1}{22}$

$1, 2, 3: \ \frac{\frac{1}{\cancel{6}3}}{2 \cdot \frac{11}{\cancel{14}7}} = \frac{7}{66}$

$\varepsilon_3: \ \frac{3}{22}$ $\qquad \alpha_3 = \frac{1}{2} \ln \frac{\frac{19}{22}}{\frac{3}{22}} = \ln \sqrt{\frac{19}{3}}$

t	1	2	3
ε_t	$3/10$	$3/14$	$3/22$
α_t	$\ln\sqrt{\frac{7}{3}}$	$\ln\sqrt{\frac{11}{3}}$	$\ln\sqrt{\frac{19}{3}}$

***Atenţie*! Pentru a vă uşura munca, am completat noi câteva dintre elementele tabelului precedent. Bazaţi-vă pe valorile indicate de noi, ca să nu faceţi calcule laborioase, păstrând însă rigurozitatea / corectitudinea raţionamentelor.**

b. Folosind ipoteza combinată obţinută de algoritmul AdaBoost la finalul celei de-a treia iteraţii, stabiliţi:

i. eroarea la antrenare produsă (pentru calcularea ei, puteţi folosi tabelul de mai jos),

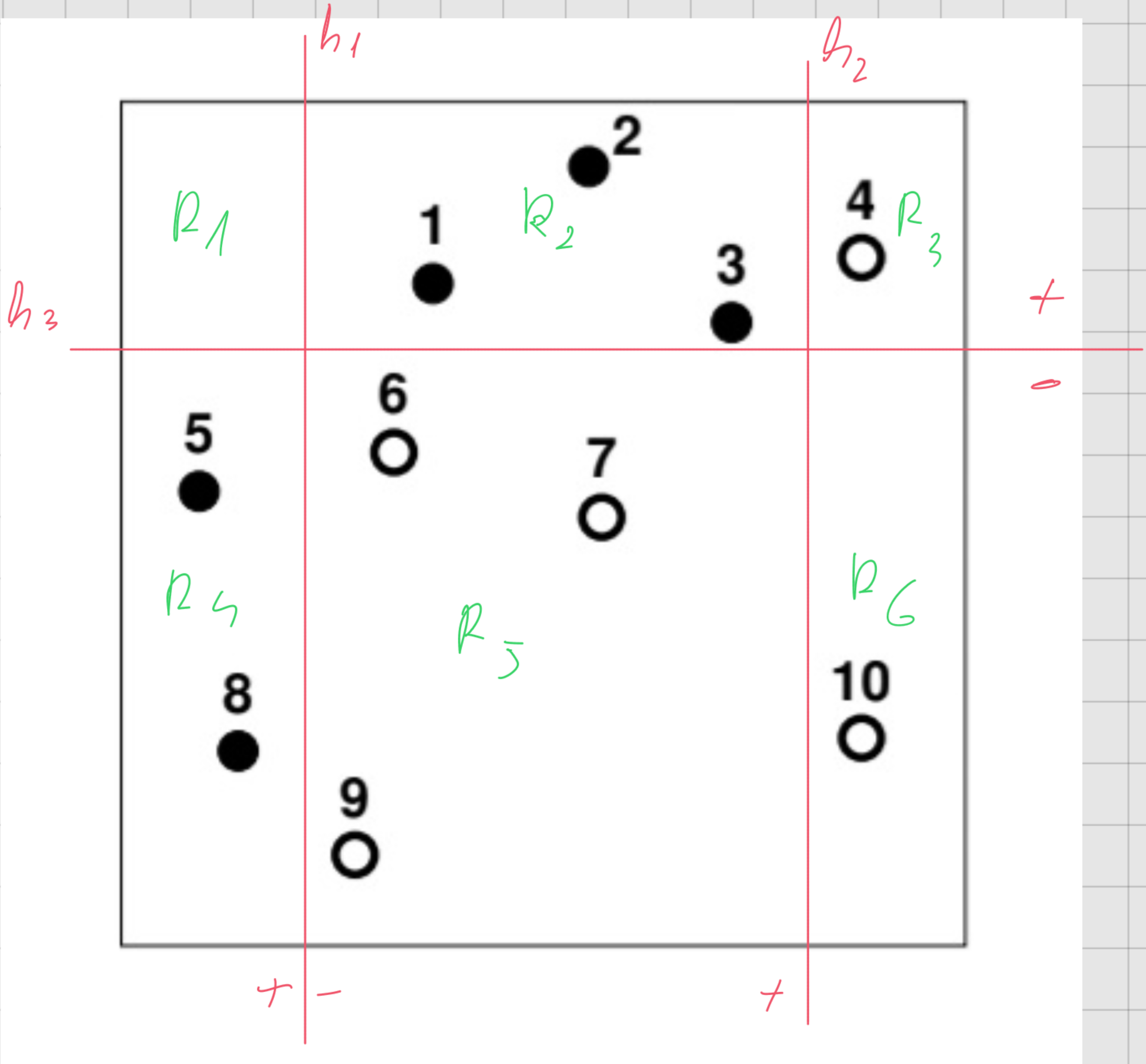
ii. zonele de decizie corespunzătoare acestui clasificator (veţi justifica modul în care aţi procedat!). Veţi indica aceste zone de decizie, precum şi graniţele de decizie, pe primul desen din enunţ.

b

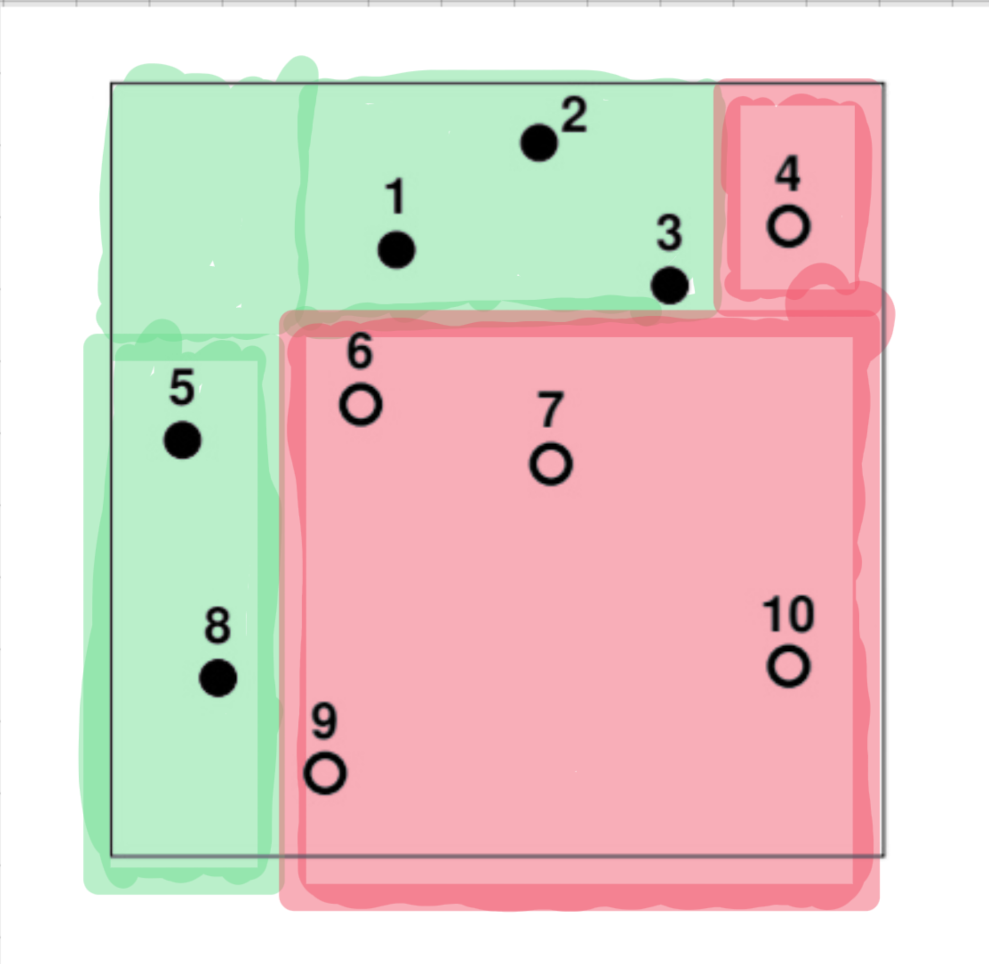
t	α_t	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
1	$\alpha_1 = 0.423$	$-$	$-$	$-$	$-$	$+$	$-$	$-$	$+$	$-$	$-$
2	$\alpha_2 = 0.640$	$+$	$+$	$+$	$-$	$+$	$+$	$+$	$+$	$+$	$-$
3	$\alpha_3 = 0.922$	$+$	$+$	$+$	$+$	$-$	$-$	$-$	$-$	$-$	$-$
	$H_3(x_i)$	$+$	$+$	$+$	$-$	$+$	$-$	$-$	$+$	$-$	$-$

$H_3(x) = \text{sign}(\alpha_1 \cdot h_1(x) + \alpha_2 h_2(x) + \alpha_3 h_3(x))$

err. de ent. $= 0$.



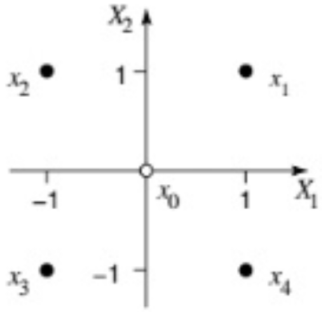
t	α_1	R_1	R_2	R_3	R_4	R_5	R_6
1	0.423	+	−	−	+	−	−
2	0.649	+	+	−	+	+	−
3	0.922	+	+	+	−	−	−
	$H_3(R_i)$	+	+	=	+	=	=



65. **(Algoritmul AdaBoost: aplicare pe un set de date din $\mathbb{R}^2$)**

□ • ○ ***CMU, 2010 fall, Aarti Singh, midterm, pr. 6.1-2***

În acest exerciţiu veţi aplica algoritmul AdaBoost folosind drept clasificatori „slabi“ compaşi de decizie pe setul de date de antrenament prezentat în figura alăturată. (Atenţie! Sunt patru instanţe pozitive şi una negativă.)



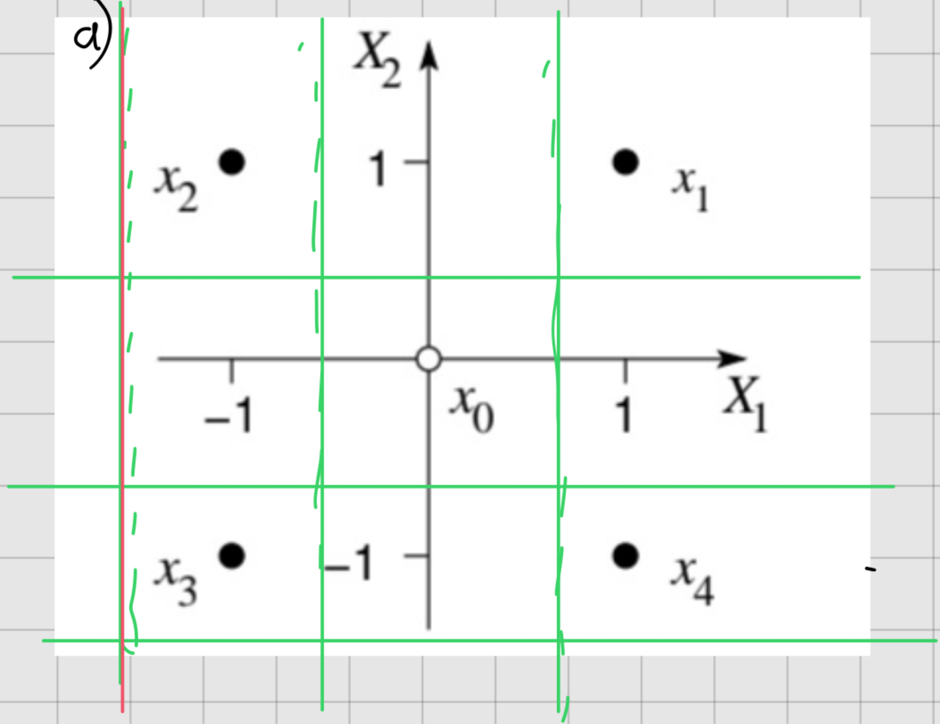
a. **Care dintre aceste exemple de antrenament vor avea probabilităţile ($D_t(i)$) mărite la sfârşitul primei iteraţii? Încercuiţi-le pe desen.**

b. **Cât de multe iteraţii vor fi necesare pentru a atinge eroare zero la antrenare? Justificaţi elaborând toate detaliile necesare.**

c. **Puteţi adăuga încă un exemplu (instanţă de antrenament) la setul de date de mai sus astfel încât algoritmul AdaBoost să obţină în [doar] două iteraţii eroare la antrenare zero? Dacă nu, explicaţi de ce nu este posibil aşa ceva.**

d. **Care credeţi că este motivul (principal) pentru care se folosesc clasificatori „slabi“ (şi nu clasificatori mai puternici / „tari“) în conjuncţie cu algoritmul AdaBoost?**

a)



X_1: n	-1.5	-0.5	0.5
err $(X_1 < n)$	4/5	2/5	3/5
err $(X_1 \geq n)$	1/5	3/5	2/5

X_2 $\quad n$	-1.3	-0.5	0.5
err $(X_2 < n)$	$4/5$	$2/5$	$3/5$
err $(X_2 \geq n)$	$1/5$	$3/5$	$2/5$

$$\varepsilon_1 = 1/5$$

$$\alpha_1 = \frac{1}{2} \ln \frac{1-\varepsilon_1}{\varepsilon_1} = \frac{1}{2} \ln \frac{\frac{4}{\cancel{5}}}{\frac{1}{\cancel{5}}} = \ln \sqrt{4} = \ln 2$$

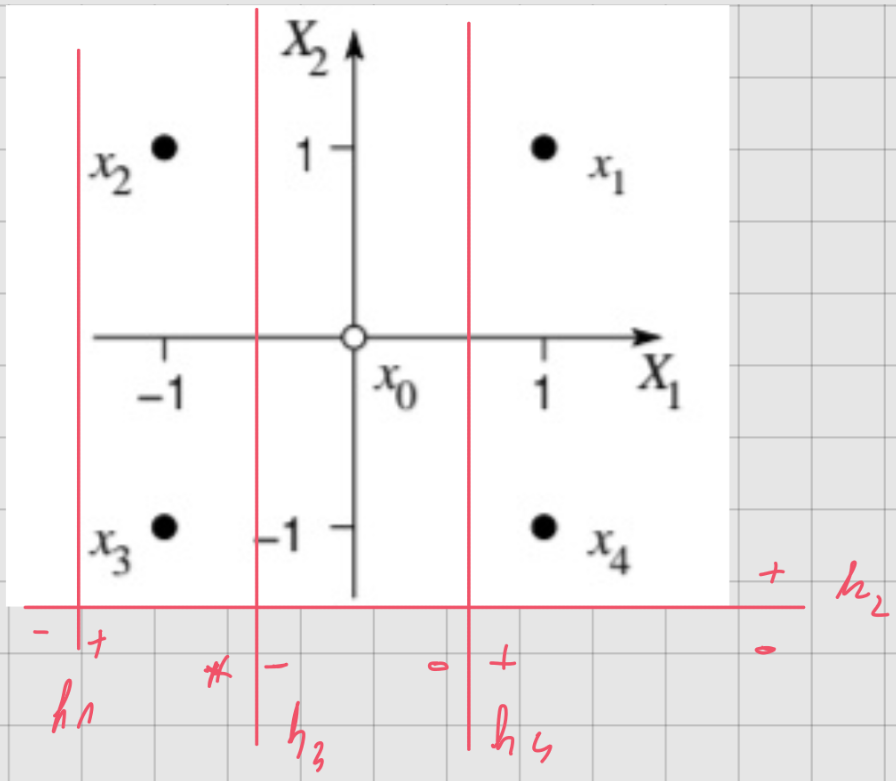
$h_1 = x_1 \geq -\frac{1}{5} \Rightarrow$ sign $\left(x_1 + \frac{1}{5}\right)$

$$D_2(i) : \begin{cases} \dfrac{D_1(i)}{2 \cdot \varepsilon_1} & , i \in O \\ \\ \dfrac{D_1(i)}{2 \cdot (1-\varepsilon_1)} & , i \notin O \end{cases}$$

$$x_0 = \frac{\frac{1}{\cancel{5}}}{2 \cdot \frac{1}{\cancel{5}}} = \frac{1}{2}$$

b) Mai este necesar de încă 3 iterații, încă 3 separatori, pt a putea clasifica corect cu err. la antrenare 0%

E * J



c)

Nu este posibil (din cauza inst. din mijloc) să obț. err la ant. 0 în 2 iterații chiar dacă am mai adăuga încă o instanță.

Datorită modului de așezare a inst. mereu pt primele 2 iterații vei folosi separatorii h_1, h_2 (separatorii exteriori), indiferent de adăugarea erorii și de poziția ei pe grafic.
instanțe

d)

Motivul principal este eficiența. Alg. adaboost este un alg iterativ, astfel dacă am folosi clasificatori puternici, atunci ar crește exponențial timpul de antrenare și iterare ceea ce l-ar face foarte prost în practică.

7.2.1 Clusterizare ierarhică

27. **(Clusterizare ierarhică aglomerativă, folosind similaritate de tip "single-", "complete-" şi "average-linkage")**

prelucrare de Liviu Ciortuz, după
■ • ○ *CMU, 2012 fall, T. Mitchell, Z. Bar-Joseph, HW4, pr. 2.a*

Tabelul de mai jos reprezintă matricea de distanţe pentru [o mulţime formată din] şase obiecte.

	A	B	C	D	E	F
A	0					
B	0.12	0				
C	0.51	0.25	0			
D	0.84	0.16	0.14	0		
E	0.28	0.77	0.70	0.45	0	
F	0.34	0.61	0.93	0.20	0.67	0

a. Aplicaţi algoritmul de clusterizare ierarhică aglomerativă pe aceste date, folosind mai întâi similaritate *single-linkage* şi apoi similaritate *complete-linkage*. La fiecare pas al algoritmului, rescrieţi în mod corespunzător matricea de distanţe. (De la o iteraţie la alta, se micşorează cu 1 numărul liniilor precum şi al coloanelor folosite.) La final, desenaţi dendrogramele rezultate.

Indicaţie: Înălţimea corespunzătoare fiecărui cluster non-singleton (adică, a fiecărui nod intern) din dendrogramă va fi considerată ca fiind egală cu distanţa (i.e., conform măsurii de similaritate) dintre cele două sub-clustere constitutive.

b. Dacă aţi lucrat corect, atunci cele două dendrograme obţinute la punctul a nu coincid [nici măcar] ca structură. Modificaţi două valori din matricea de distanţe dată mai sus, în aşa fel încât de data aceasta cele două dendrograme care se obţin să fie identice ca structură.

c. Procedaţi similar cu cerinţele de la punctul a, dar de această dată pentru *average-linkage*. La actualizarea matricei de distanţe (sau, de „proximitate") veţi ţine cont de formula:

$$\Delta(X \cup Y, Z) \stackrel{def.}{=} \frac{1}{(|X| + |Y|)\,|Z|} \sum_{x \in X \cup Y} \sum_{z \in Z} d(x, z)$$

$$\stackrel{calcul}{=} \frac{1}{|X| + |Y|} (|X|\,\Delta(X, Z) + |Y|\,\Delta(Y, Z)). \qquad (377)$$

unde X, Y şi Z sunt clustere disjuncte două câte două, iar notaţia $|X|$ desemnează cardinalul lui X (adică, numărul de elemente din X).

d. Demonstraţi formula enunţată la punctul c.

a)

	A	B	C	D	E	F
A	0					
B	0.12	0				
C	0.51	0.25	0			
D	0.84	0.16	0.14	0		
E	0.28	0.77	0.70	0.45	0	
F	0.34	0.61	0.93	0.20	0.67	0

	A/B	C	D	E	F
A/B	0				
C	0.25	0			
D	0.16	0.13	0		
E	0.28	0.7	0.45	0	
F	0.34	0.93	0.2	0.67	0

$d_{SL}(A|B, C) = \min(d_2(A, C), d_2(B, C)) =$

$d_{CL} = 0.51$

$= \min(0.51, 0.25) = 0.25$

$d_{SL}(A|B, D) = \min(d_2(A, D), d_2(B, D))$

$= \min(0.84, 0.16) = 0.16$

$d_{CL} = 0.84$

$d_{SL}(A|B, E) = \min(d_2(A,E), d_2(B,E)) = \min(0.28, 0.77) = 0.28$

$d_{SL}(A|B, F) = \min(d_2(A,F), d_2(B,F)) = \min(0.34, 0.61) = 0.34$

	A|B	C|D	E	F
A|B	0			
C|D	0.16	0		
E	0.28	0.45	0	
F	0.34	0.2	0.67	0

$d_{SL}(AB, C|D) = \min(d_2(AB,C), d_2(AB,D))$

$= \min(0.25, 0.16) = 0.16$

$d_{SL}(C|D, E) = \min(d_2(C,E), d_2(D,E)) =$

$= \min(0.7, 0.45) = 0.45$

$d_{SL}(C|D, F) = \min(d_2(C|F), d_2(D,F)) =$

$= \min(0.9, 0.2) = 0.2$

	AB|CD	E	F
AB|CD	0		
E	0.28	0	
F	0.2	0.67	0

$d_{SL}(AB|CD, E) = \min(d_2(AB,E),$

$d_2(CD,E)) = \min(0.28, 0.45) = 0.28$

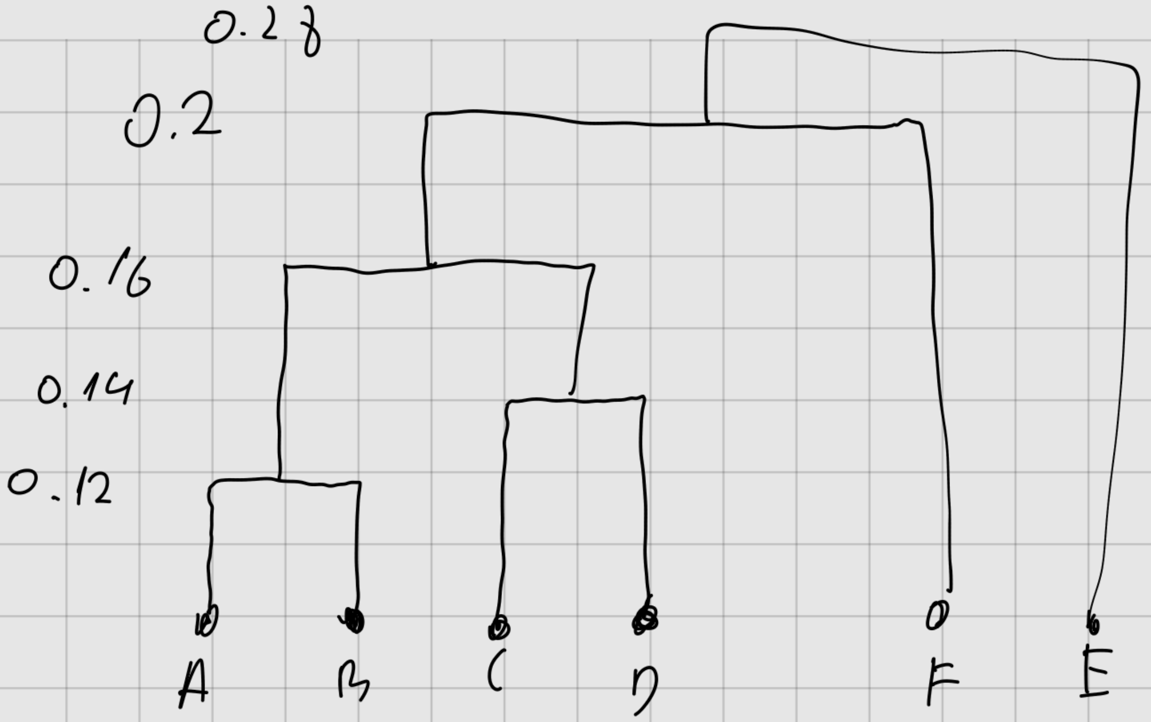
$d_{SL}(AB|CD, F) = \min(d_2(AB,F), d_2(CD,F)) =$

$= \min(0.34, 0.2) = 0.2$

	ABCD|F	E
ABCD|F	0	
E	0.28	0

$d_{SL}(ABCD|F, E) = \min(0.28, 0.67)$

$= 0.28$



Complete linkage

	A|B	C	D	E	F
A|B	0				
C	0.51	0			
D	18 0.84	0.13	0		
E	0.77	0.7	0.45	0	
F	0.61	0.93	0.2	0.67	0

	A|B	C|D	E	F
A|B	0			
C|D	~~0.84~~ 18	0		
E	0.77	0.7	0	
F	0.61	0.93	0.67	0

$$d_{CL}(AB, C|D) = d_{CL}(d_2(AB, C), d_2(AB, D)) = \max(0.51, 0.84) = 0.84$$

$$d_{CL}(C|D, E) = \max(0.7, 0.45) = 0.7$$

$$d_{CL}(C|D, F) = \max(0.93, 0.2) = 0.93$$

	AB|F	C|D	E
	0		

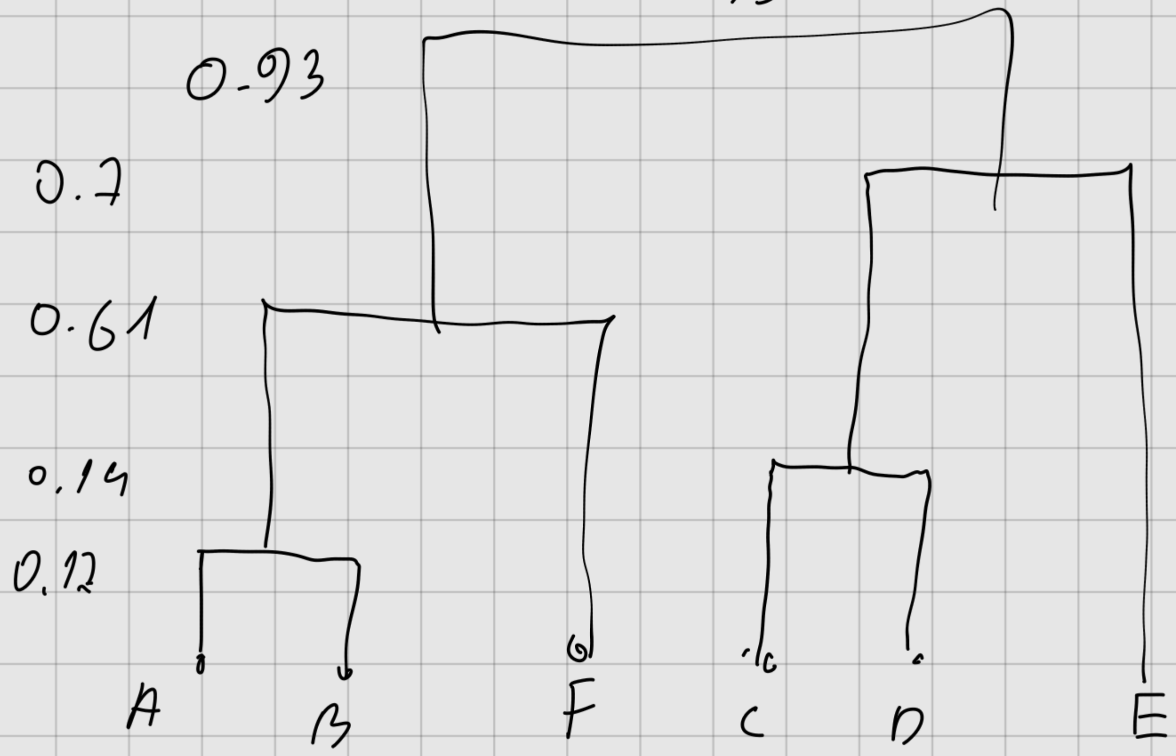
	0.93	0	
	0.77	0.7	0

$$d_{CL}(ABF, CD) = \max(0.84, 0.93) = 0.93$$

	ABF	CDE
ABF	0	
CDE	0.93	0

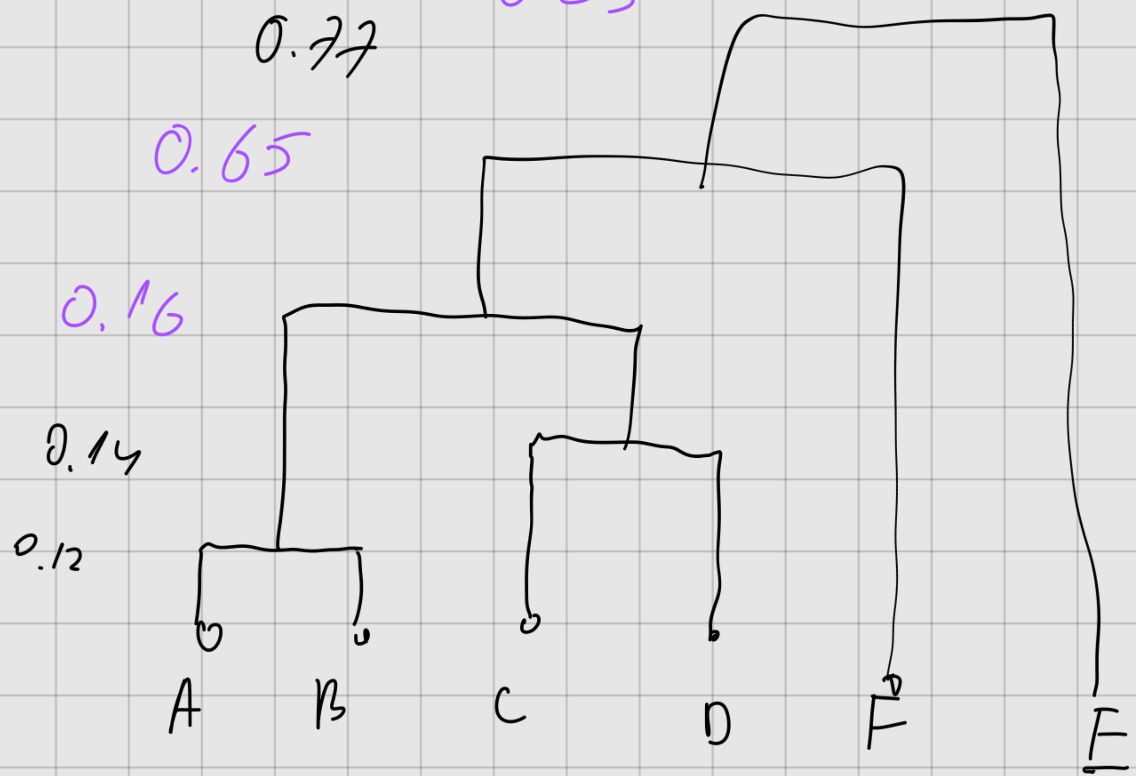
$$d_{CL}(ABF, E) = \max(0.77, 0.67) = 0.77$$

$$d_{CL}(ABF, CDE) = \max(0.93, 0.77) = 0.93$$



b)

	A	B	C	D	E	F
A	0					
B	0.12	0				
C	0.51	0.25	0			
D	~~0.84~~ 0.16	0.16	0.14	0		
E	0.28	0.77	0.70	0.45	0	
F	0.34	0.61	~~0.93~~ 0.65	0.20	0.67	0



	A/B	C	D	E	F
A/B	0				
C	0.51	0			
D	~~0.84~~ 0.16	0.14	0		
E	0.77	0.7	0.45	0	
F	0.61	~~0.93~~ 0.65	0.2	0.67	0

	A|B	C|D	E	F
A|B	0			
C|D	0.51	0		
E	0.77	0.7	0	
F	0.61	~~0.93~~ 0.65	0.67	0

	AB|CD	E	F
AB|CD	0		
E	0.77	0	
F	~~0.93~~ 0.65	0.67	0

Am obț. aceeași dendogramă.

d) . . . .

28. **(Clusterizare ierarhică aglomerativă, folosind similaritate de tip "single-linkage", "complete-linkage" şi "average-linkage"; aplicare pe date din $\mathbb{R}$)**

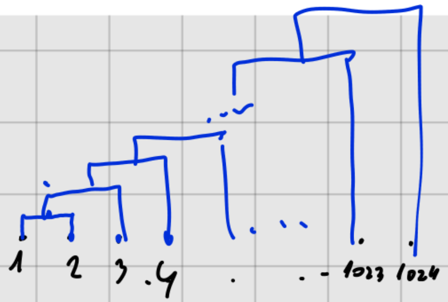
∗ *CMU, 2009 spring, Ziv Bar-Joseph, final exam., pr. 9.2*

Dorim să clusterizăm numerele de la 1 la 1024, folosind algoritmul de clusterizare ierarhică aglomerativă. Dacă la o iteraţie oarecare distanţele dintre două perechi de clustere au aceeaşi valoare, ne vom folosi de ordonarea numerică.[882]

Ex29.

Să se compare rezultatele obţinute cu trei variante ale funcţiei de similaritate discutate la curs — şi anume, "single-linkage", "complete-linkage" şi "average-linkage" —, specificându-se pentru fiecare dintre aceste variante care este numărul de elemente asignate fiecăruia dintre cele două clustere [care compun clusterul corespunzător nodului rădăcină] de la vârful dendrogramei rezultate.

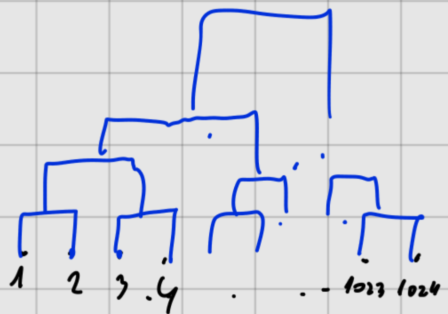
SL:



$C_2 = \{1024\}$

$C_1 = \{1, 2, \dots 1023\}$

CL:



$C_1 = \{1 \dots 512\}$

$C_2 = \{513 \dots 1024\}$

AL:

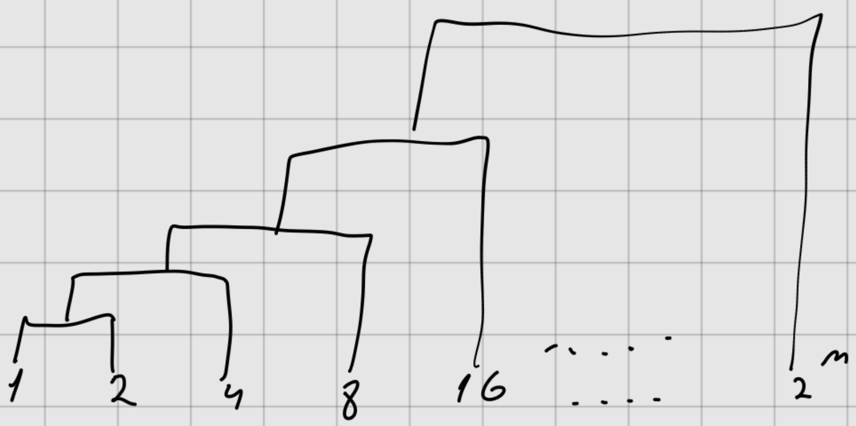
↳ răd. va fi compusă din 511 și 511 clustere

29. **(Clusterizare ierarhică aglomerativă: aplicare pe date din $\mathbb{R}$)**

□ • ○ *CMU, 2012 spring, Ziv Bar-Joseph, midterm, pr. 9.1*

Ne propunem să clusterizăm în manieră ierarhică aglomerativă instanţele $2^0, 2^1, 2^2, \ldots, 2^n$ (deci, în total, $n+1$ puncte), folosind distanţa euclidiană. Trasaţi dendrogramele (adică arborii de clusterizare ierarhică) care se obţin pentru fiecare dintre următoarele trei tipuri de funcţii de similaritate: single-, complete- şi average-linkage. Vom considera că fiecare nod neterminal din dendrogramă — adică rădăcina fiecărui cluster non-singleton —, are înălţimea egală cu valoarea măsurii de similaritate (adică, distanţa) dintre cele două clustere din care a fost obţinut.

Single-linkage



complete-linkage şi average linkage.

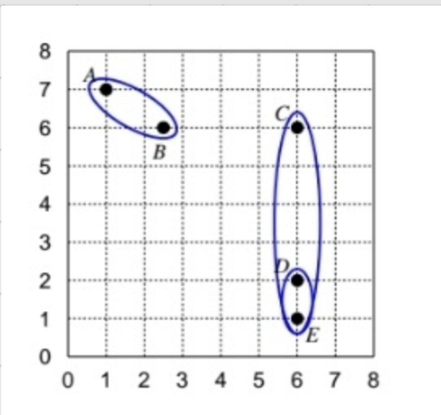
30. **(Clusterizare ierarhică aglomerativă: un exemplu simplu de aplicare, cu single-, complete- şi average-linkage, pe date din $\mathbb{R}^2$)**

prelucrare de Liviu Ciortuz, 2022, după
• ○ *CMU, 2021 fall, Aarti Signh, Recitation 9, pr. 3*

Considerăm punctele $A(1,7)$, $B(2.5,6)$, $C(6,6)$, $D(6,2)$ şi $E(6,1)$ din planul euclidian. Pe acest dataset veţi aplica algoritmul de clusterizare ierarhică aglomerativă (i.e., bottom-up), folosind pe rând (separat) funcţiile de similaritate single-linkage, complete-linkage şi average-linkage.

Care dintre aceste funcţii de similaritate va conduce după executarea a trei iteraţii consecutive la ierarhia aplatizată prezentată în figura de mai jos? (*Observaţie*: Pentru simplitate, nu am mai desenat şi elipsa corespunzătoare clusterului care include toate punctele de clusterizat.)

[882] Vedeţi de exemplu cum s-a procedat la problema 1.



$d_2(A,B) = \left\| \begin{pmatrix}1\\7\end{pmatrix} - \begin{pmatrix}2.5\\6\end{pmatrix} \right\| = \left\| \begin{pmatrix}1.5\\1\end{pmatrix} \right\| = \sqrt{2.25+1}$

↳ euclidiano

calc. ut. SL:

$= \sqrt{(1-2.5)^2+(7-6)^2} = \sqrt{3.25}$

$d_2(A,C) = \left\| \begin{pmatrix}1\\7\end{pmatrix} - \begin{pmatrix}6\\6\end{pmatrix} \right\| = \sqrt{25+1} = \sqrt{26}$

$d_2(B,C) = \left\| \begin{pmatrix}2.5\\6\end{pmatrix} - \begin{pmatrix}6\\6\end{pmatrix} \right\| = \sqrt{(3.5)^2} = \sqrt{12.25}$

$d_2(A,D) = \left\| \begin{pmatrix}1\\7\end{pmatrix} - \begin{pmatrix}6\\2\end{pmatrix} \right\| = \left\| \begin{pmatrix}5\\5\end{pmatrix} \right\| = \sqrt{50}$

$d_2(B,D) = \left\| \begin{pmatrix}2.5\\6\end{pmatrix} - \begin{pmatrix}6\\2\end{pmatrix} \right\| = \left\| \begin{pmatrix}3.5\\4\end{pmatrix} \right\| = \sqrt{28.25}$

$d_2(C,D) = \left\| \begin{pmatrix}6\\6\end{pmatrix} - \begin{pmatrix}6\\2\end{pmatrix} \right\| = \left\| \begin{pmatrix}0\\4\end{pmatrix} \right\| = \sqrt{16}$

$d_2(A,E) = \left\| \begin{pmatrix}1\\7\end{pmatrix} - \begin{pmatrix}6\\1\end{pmatrix} \right\| = \left\| \begin{pmatrix}5\\6\end{pmatrix} \right\| = \sqrt{25+36} = \sqrt{61}$

$d_2(B,E) = \left\| \begin{pmatrix}2.5\\6\end{pmatrix} - \begin{pmatrix}6\\1\end{pmatrix} \right\| = \left\| \begin{pmatrix}3.5\\5\end{pmatrix} \right\| = \sqrt{12.25+25} = \sqrt{37.25}$

$d_2(C,E) = \left\| \begin{pmatrix}6\\6\end{pmatrix} - \begin{pmatrix}6\\1\end{pmatrix} \right\| = \left\| \begin{pmatrix}0\\5\end{pmatrix} \right\| = \sqrt{25}$

$d_2(D,E) = \left\| \begin{pmatrix}6\\2\end{pmatrix} - \begin{pmatrix}6\\1\end{pmatrix} \right\| = \left\| \begin{pmatrix}0\\1\end{pmatrix} \right\| = \sqrt{1}$

	A	B	C	D	E
A	0				
B	$\sqrt{3.25}$	0			
C	$\sqrt{26}$	$\sqrt{12.25}$	0		
D	$\sqrt{50}$	$\sqrt{28.25}$	$\sqrt{16}$	0	
E	$\sqrt{61}$	$\sqrt{37.25}$	$\sqrt{25}$	$\sqrt{1}$	0

	A	B	C	D|E
A	0			
B	$\sqrt{3.25}$	0		
C	$\sqrt{26}$	$\sqrt{12.25}$	0	
D|E	$\sqrt{50}$	$\sqrt{28.25}$	$\sqrt{16}$	0

$d_{SL}(A, D|E) = \min(d_2(A,D), d_2(A,E)) =$
$= \min(\sqrt{50}, \sqrt{61}) = \sqrt{50}$

$d_{SL}(B, D|E) = \min(d_2(B,D), d_2(B,E)) =$
$= \min(\sqrt{28.25}, \sqrt{37.25}) = \sqrt{28.25}$

$d_{SL}(C, D|E) = \min(d_2(C,D), d_2(C|E)) =$
$= \min(\sqrt{16}, \sqrt{25}) = \sqrt{16}$

	A|B	C	D|E
A|B	0		
C	$\sqrt{12.25}$	0	
D|E	$\sqrt{28.25}$	$\sqrt{16}$	0

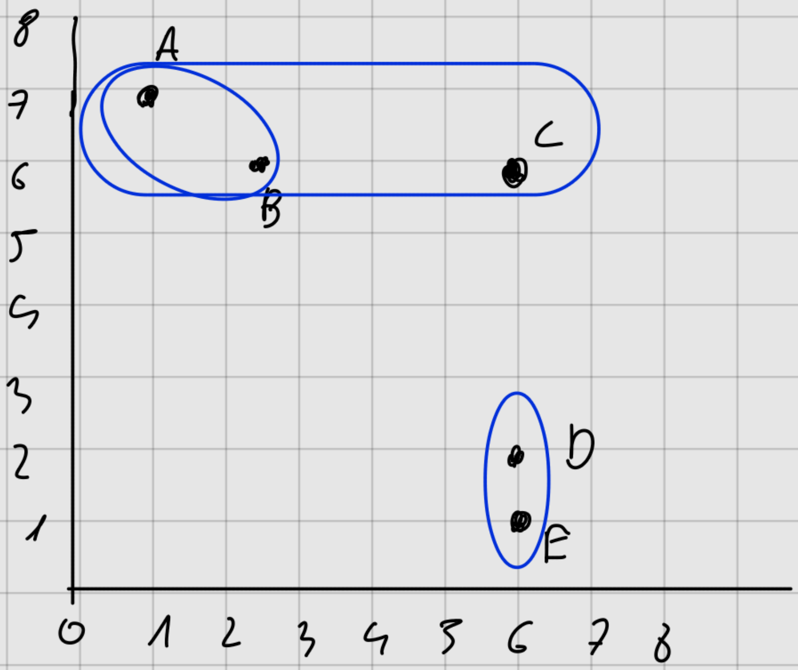
$d_{SL}(A|B, C) = \min(d_2(A,C), d_2(B,C)) =$
$\min(\sqrt{26}, \sqrt{12.25}) = \sqrt{12.25}$

$d_{SL}(A|B, D|E) = \min(d_2(A, D|E), d_2(B, D|E)) =$
$\min(\sqrt{50}, \sqrt{28.25}) = \sqrt{28.25}$

	A,B|C	D|E
A,B|C	0	
D|E	$\sqrt{16}$	0

$d_{SL}(A,B|C,\ D|E) = \min(d_2(A|B, D|E), d_2(C, D|E))$

$\min(\sqrt{28.25}, \sqrt{16}) = \sqrt{16}$



Nu conduce.

Complete linkage:

	A	B	C	D|E
A	0			
B	$\sqrt{2.25}$	0		
C	$\sqrt{26}$	$\sqrt{12.25}$	0	
D|E	$\sqrt{61}$	$\sqrt{37.25}$	$\sqrt{25}$	0

$d_{CL}(A, D|E) = \max(d_2(A,D), d_2(A|E)) =$

$\max(\sqrt{50}, \sqrt{61}) = \sqrt{61}$

$d_{CL}(B, D|E) = \max(d_2(B,D), d_2(B,E)) =$

$\max(\sqrt{28.25}, \sqrt{37.25}) = \sqrt{37.25}$

$d_{CL}(C, D|E) = \max(\sqrt{16}, \sqrt{25}) = \sqrt{25}$

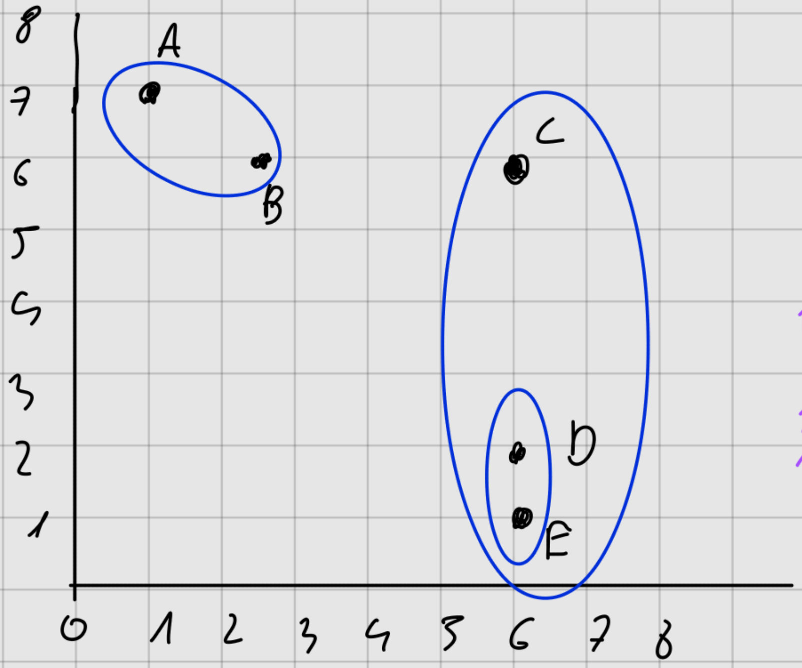
	A|B	C	D|E
A|B	0		
C	$\sqrt{26}$	0	
D|E	$\sqrt{61}$	$\sqrt{25}$	0

$d_{CL}(A|B, C) = \max(\sqrt{26}, \sqrt{12.25}) = \sqrt{26}$

$d_{CL}(A|B, D|E) = \max(\sqrt{61}, \sqrt{37.25}) = \sqrt{61}$

$d_{CL}(C, D|E) = \sqrt{25}$

	A|B	D,E|C
A|B	0	
D,E|C	...	0



corect, complete linkage ajunge după 3 iterații.

Average linkage

	A	B	C	D/E
A	0			
B	$\sqrt{2.25}$	0		
C	$\sqrt{26}$	$\sqrt{12.25}$	0	
D/E	7.44	5.70	4.5	0

$$d_{AL} = \frac{d_2(A, D/E)}{1} = \frac{d_2(A,D) + d_2(A,E)}{2} =$$

$$= \frac{\sqrt{50} + \sqrt{61}}{2} = 7.44$$

$$d_{AL} = \frac{d_2(B, D/E)}{1} = \frac{d_2(B,D) + d_2(B,E)}{2} =$$

$$= \frac{\sqrt{28.25} + \sqrt{37.25}}{2} = 5.70$$

$$d_{AL} = \frac{d_2(C,D) + d_2(C,E)}{2} = \frac{\sqrt{16} + \sqrt{25}}{2} = \frac{9}{2} = 4.5$$

	A/B	C	D/E
A/B	0		
C	4.29	0	
D/E	6.57	4.5	0

$$d_{AL}(A/B, C) = \frac{d_2(A,C) + d_2(B,C)}{2} = \frac{\sqrt{26} + \sqrt{12.25}}{2}$$

$$= 4.29$$

$$d_{AL}(A/B, D/E) = \frac{d_2(A/B, D) + d_2(A/B, E)}{2} =$$

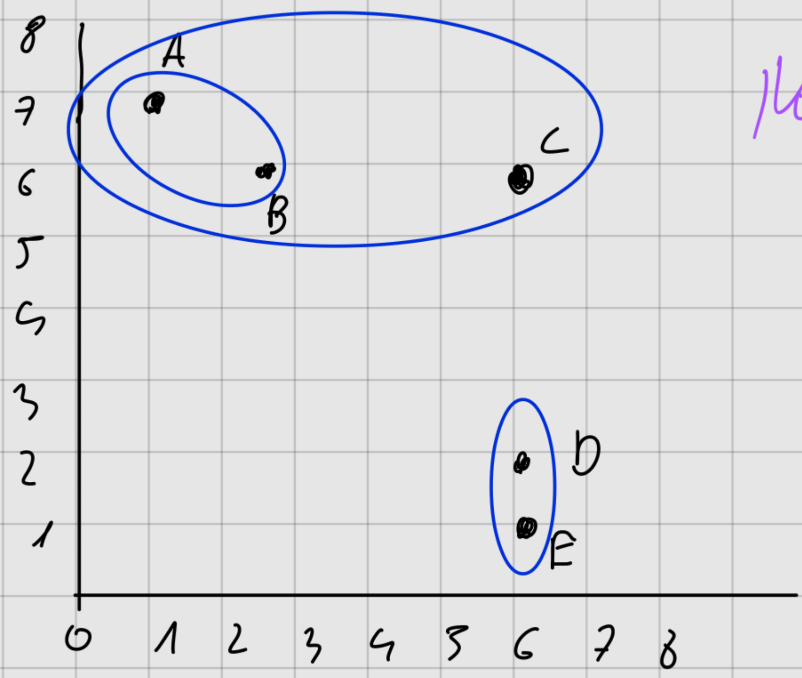
$$= \frac{d_2(A,D) + d_2(B,D) + d_2(A,E) + d_2(B,E)}{4} =$$

$$= \frac{\sqrt{50} + \sqrt{61} + \sqrt{28.25} + \sqrt{37.25}}{4} = 6.57$$

A,B|C D|E

A,B|C O

D|E ... O



Nu este totul!

Deci, singura singura funcție de similitate care coresp. cu fig. de jos este cea complete-linkage.

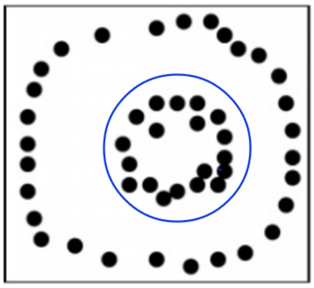
31. **(Clusterizare ierarhică aglomerativă: aplicare în manieră intuitivă pe date din $\mathbb{R}^2$)**

∗ *CMU, 2010 fall, Ziv Bar-Joseph, midterm, pr. 8.b*

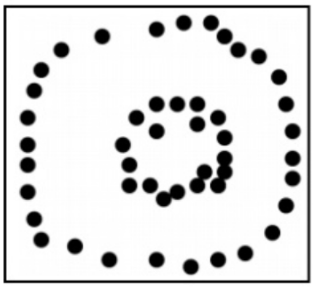
a. Folosind distanţa euclidiană, aplicaţi algoritmul de clusterizare ierarhică aglomerativă cu similaritate de tip "single-linkage" pe datele din figura alăturată.

Indicaţi pe desen care sunt instanţele care formează cele două clustere de la vârful dendrogramei.

Observaţie: Nu este necesar să construiţi efectiv dendrograma.



b. Care este rezultatul dacă se foloseşte similaritate de tip "average-linkage"?



33. **(Clusterizare ierarhică [aglomerativă]: o altă funcţie / măsură de similaritate între clustere: metrica / distanţa lui Ward)**

prelucrare făcută de Liviu Ciortuz, după ■ • ○ *CMU, 2010 fall, Aarti Singh, HW3, pr. 4.1*

În acest exerciţiu veţi analiza o modalitate alternativă pentru a defini distanţa dintre două clustere disjuncte, care a fost propusă de către Joe H. Ward în 1963.[883] O vom numi *metrica lui Ward.*

Această metrică defineşte distanţa dintre două clustere disjuncte X şi Y (ambele finite şi incluse în $\mathbb{R}^d$ cu $d \in \mathbb{N}^*$) ca fiind dată de creşterea *sumei pătratelor* distanţelor de la fiecare instanţă x_i la *centroidul* clusterului la care ea este asignată, atunci când reunim cele două clustere ca să formăm un cluster nou. Din punct de vedere formal, vom scrie:

$$\Delta(X,Y) = \sum_{x_i \in X \cup Y} \|x_i - \mu_{X\cup Y}\|^2 - \sum_{x_i \in X} \|x_i - \mu_X\|^2 - \sum_{x_i \in Y} \|x_i - \mu_Y\|^2 \qquad (378)$$

unde, spre exemplu, μ_X este centroidul („centrul de greutate") clusterului X, iar x_i este o instanţă generică dintr-un cluster [oarecare, fixat]. Prin definiţie, aici vom considera $\mu_X = \frac{1}{n_X}\sum_{x_i \in X} x_i$, unde n_X este numărul de elemente din X. Mărimea $\Delta(X,Y)$ din formula (378) poate fi interpretată ca reprezentând *costul pentru combinarea* celor două clustere X şi Y într-un singur cluster.[884]

a. Aduceţi expresia din partea dreaptă a formulei (378) la o formă mai simplă. (Redactaţi toţi paşii intermediari.) *Sugestie:* Formula obţinută trebuie să fie doar în funcţie de numărul de elemente din cele două clustere (n_X şi respectiv n_Y) şi de distanţa $\|\mu_X - \mu_Y\|^2$ dintre centroizii acestor clustere, μ_X şi respectiv μ_Y.

$\left\| \binom{1}{2} - \binom{3}{4} \right\| = \left\| \binom{2}{2} \right\| = \sqrt{5+5}$

$\|x - y\| = \sqrt{(x-y)^2} = |x-y|$

a) $\Delta(X,Y) = \sum_{x_i \in X\cup Y} \|x_i - \mu_{X\cup Y}\|^2 - \sum_{x_i\in X} \|x_i - \mu_X\|^2 - \sum_{x_i \in Y} \|x_i - \mu_Y\|^2$

$= \sum_{x_i \in X\cup Y} (x_i - \mu_{X\cup Y})^2 - \sum_{x_i \in X} (x_i - \mu_X)^2 - \sum_{x_i\in Y} (x_i - \mu_Y)^2 =$

$= \sum_{x_i \in X\cup Y} (x_i^2 - 2x_i \cdot \mu_{X\cup Y} + \mu_{X\cup Y}^2) - \sum_{x_i \in X} (x_i^2 - 2x_i\mu_X + \mu_X^2) - \sum_{x_i\in Y} (x_i^2 - 2x_i\mu_Y + \mu_Y^2) =$

$= \sum_{x_i \in X\cup Y} x_i^2 - 2\sum_{x_i\in X\cup Y} x_i \cdot \mu_{X\cup Y} + \sum_{X\cup Y} \mu_{X\cup Y}^2 - \cancel{\sum_X x_i^2} + 2\sum_X x_i\mu_X - \sum_X \mu_X^2 - \cancel{\sum_Y x_i^2} + 2\sum_Y x_i \mu_Y - \sum_Y \mu_Y^2 =$

$=$

$\hookrightarrow \cancel{\sum_{x_i\in X} x_i^2} + \cancel{\sum_{x_i \in Y} x_i^2}$

$= -2\mu_{X\cup Y} \sum_{x_i \in X} x_i - 2\mu_{X\cup Y} \sum_{x_i\in Y} x_i + \sum_{X\cup Y} \mu_{X\cup Y}^2 + 2\mu_X \sum_X x_i - \sum_X \mu_X^2 + 2\mu_Y \sum_Y x_i - \sum_Y \mu_Y^2 =$

$\mu_X \overset{def.}{=} \frac{\sum_X x_i}{n_X} \Rightarrow \sum_X x_i \overset{(1)}{=} n_X \cdot \mu_X$

$\mu_{X\cup Y} \overset{def}{=} \frac{1}{|X|+|Y|}\left(\sum_X x_i + \sum_Y x_i\right) \overset{(1)}{\longrightarrow}$ $\qquad |X| = n_X$

$\mu_{X\cup Y} \overset{(1)}{=} \frac{n_X\mu_X + n_Y\mu_Y}{n_X + n_Y}$

$$\boxed{\mu_{X\cup Y} \sum_{X \cup Y} x_i} \overset{(5),(1)}{=} \frac{(m_X\mu_X + m_Y\mu_Y)}{m_X+m_Y}(m_X\cdot\mu_X + m_Y\cdot\mu_Y) = \frac{(m_X\mu_X+m_Y\mu_Y)^2}{m_X+m_Y}$$

$$= -2\left(\frac{m_X\mu_X+m_Y\mu_Y}{m_X+m_Y}\right)m_X\mu_X - 2\left(\frac{m_X\mu_X+m_Y\mu_Y}{m_X+m_Y}\right)\cdot m_Y\mu_Y + \cancel{(m_X+m_Y)}\cdot\frac{(m_X\mu_X+m_Y\mu_Y)^2}{(m_X+m_Y)^{\cancel{2}}}$$

$$+\cancel{2}^{1} m_X\mu_X^2 - \cancel{m_X\cdot\mu_X^2} + \cancel{2}^{1} m_Y\mu_Y^2 - \cancel{m_Y\cdot\mu_Y^2} =$$

$$= -\cancel{2}^{1}\frac{(m_X\mu_X+m_Y\mu_Y)^2}{m_X+m_Y} + \cancel{\frac{(m_X\mu_X+m_Y\mu_Y)^2}{m_X+m_Y}} + m_X\mu_X^2 + m_Y\mu_Y^2 =$$

$$= -\frac{(m_X\mu_X+m_Y\mu_Y)^2}{m_X+m_Y} + m_X\mu_X^2 + m_Y\mu_Y^2 =$$

$$= \frac{-(m_X\mu_X+m_Y\mu_Y)^2 + (m_X+m_Y)(m_X\mu_X^2+m_Y\mu_Y^2)}{m_X+m_Y} =$$

$$= \frac{\cancel{-(m_X\mu_X)^2} - 2\cdot m_X\mu_X\cdot m_Y\mu_Y - \cancel{(m_Y\mu_Y)^2} + \cancel{(m_X\mu_X)^2} + m_Xm_Y\mu_Y^2 + m_Xm_Y\mu_X^2 + \cancel{(m_Y\mu_Y)^2}}{m_X+m_Y} = \frac{m_Xm_Y}{m_X+m_Y}(\mu_Y^2+\mu_X^2-2\mu_X\mu_Y) =$$

$$(\mu_X-\mu_Y)^2 = \|\mu_X-\mu_Y\|^2$$

$$\boxed{= \frac{m_Xm_Y}{m_X+m_Y}\|\mu_X-\mu_Y\|^2}$$

b) Urmărește ca atunci când creștem clusterele făcând merge, să le mențină cât de mici posibil.

34. (Clusterizare ierarhică aglomerativă: aplicare pe date din $\mathbb{R}^2$, folosind metrica lui Ward)

Liviu Ciortuz, 2018, folosind datele de la ■ ∘ ∗ Edinburgh, 2009 fall, C. Williams, V. Lavrenko, HW4, pr. 3

La problema 33 am prezentat o funcţie de similaritate numită *metrica lui Ward*. Potrivit acestei metrici, distanţa dintre două clustere disjuncte X şi Y se defineşte astfel:

$$\Delta(X,Y) = \sum_{x_i \in X \cup Y} \|x_i - \mu_{X \cup Y}\|^2 - \sum_{x_i \in X} \|x_i - \mu_X\|^2 - \sum_{x_i \in Y} \|x_i - \mu_Y\|^2 \qquad (379)$$

unde, spre exemplu, μ_X este centroidul [sau „centrul de greutate“ al] clusterului X, iar x_i este o instanţă generică dintr-un cluster [oarecare, fixat]. Prin definiţie, aici vom considera $\mu_X = \frac{1}{n_X}\sum_{x_i \in X} x_i$, unde n_X este numărul de elemente din X. (Similar sunt definiţi centroizii μ_Y şi $\mu_{X \cup Y}$.)

Se poate arăta (vedeţi tot problema 33) că

$$\boxed{\Delta(X,Y) = \frac{n_X n_Y}{n_X + n_Y}\|\mu_X - \mu_Y\|^2.} \qquad (380)$$

Observaţii:

1. Pentru perechi de clustere (X,Y) şi (X',Y') astfel încât $n_X = n_{X'}$ şi $n_Y = n_{Y'}$,[885] formula (380) arată că la clusterizare ierarhică folosind merica lui Ward [la o iteraţie oarecare] este „favorizată“ acea pereche pentru care centroizii (μ_X şi μ_Y, respectiv $\mu_{X'}$ şi $\mu_{Y'}$) sunt mai apropiaţi.
2. Invers, dacă $\|\mu_X - \mu_Y\| = \|\mu_{X'} - \mu_{Y'}\|$, atunci este favorizată perechea pentru care ponderea[886] (adică $\frac{n_X n_Y}{n_X + n_Y}$, respectiv $\frac{n_{X'} n_{Y'}}{n_{X'} + n_{Y'}}$) este mai mică.

Aceste două *observaţii* vă vor ajuta să simplificaţi / reduceţi foarte mult calculele pe care ar trebui să le faceţi pentru a rezolva următoarea cerinţă!

Aplicaţi algoritmul de clusterizare ierarhică aglomerativă pe setul de date de la problema 2,

$$A:(-4,-2), B:(-3,-2), C:(-2,-2), D:(-1,-2), E:(+1,-1)$$
$$F:(+1,+1), G:(+2,+3), H:(+3,+2), I:(+3,+4), J:(+4,+3)$$

utilizând [de această dată] metrica lui Ward.[887] Ca rezultat al clusterizării, veţi reprezenta *dendrograma* sub formă *aplatizată*, folosind elipse (şi indici) pentru a indica clusterele formate.

Coincide rezultatul obţinut aici cu vreunul din rezultatele de la problema 2?

Precizare: Dacă la o iteraţie a algoritmului de clusterizare distanţele (adică similarităţile) dintre două perechi de clustere au aceeaşi valoare, veţi considera că prioritatea la alcătuirea noului cluster este dictată de ordinea alfabetică.

	x_1	x_2
A	-4	-2
B	-3	-2
C	-2	-2
D	-1	-2
E	1	-1
F	1	1
G	2	3
H	3	2
i	3	4
j	4	3

	A	B	C	D	E	F	G	H	i	j
A	0									
B	$\frac{1}{2}$	0								
C	1	$\frac{1}{2}$	0							
D	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0						
E	$\frac{\sqrt{26}}{2}$	$\frac{\sqrt{17}}{2}$	$\frac{\sqrt{10}}{2}$	$\frac{\sqrt{5}}{2}$	0					
F	$\frac{\sqrt{34}}{2}$	$\frac{3}{2}$	$\frac{\sqrt{18}}{2}$	$\frac{\sqrt{13}}{2}$	1	0				
G	$\frac{\sqrt{61}}{2}$	$\frac{\sqrt{50}}{2}$	$\frac{\sqrt{41}}{2}$	$\frac{\sqrt{34}}{2}$	$\frac{\sqrt{17}}{2}$	$\frac{\sqrt{5}}{2}$	0			
H	$\frac{\sqrt{65}}{2}$	$\frac{\sqrt{52}}{2}$	$\frac{\sqrt{41}}{2}$	$\frac{\sqrt{32}}{2}$	$\frac{\sqrt{13}}{2}$	$\frac{\sqrt{5}}{2}$	$\frac{\sqrt{2}}{2}$	0		
i	$\frac{\sqrt{85}}{2}$	$\frac{\sqrt{72}}{2}$	$\frac{\sqrt{61}}{2}$	$\frac{\sqrt{52}}{2}$	$\frac{\sqrt{29}}{2}$	$\frac{\sqrt{13}}{2}$	$\frac{\sqrt{2}}{2}$	1	0	
j	$\frac{\sqrt{89}}{2}$	$\frac{\sqrt{75}}{2}$	$\frac{\sqrt{61}}{2}$	$\frac{\sqrt{50}}{2}$	$\frac{5}{2}$	$\frac{\sqrt{13}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	0

Fără radical

$$d_w(A,B) = \frac{1 \cdot 1}{1+1}\left\|\begin{vmatrix}-4\\-2\end{vmatrix} - \begin{vmatrix}-3\\-2\end{vmatrix}\right\|^2 = \frac{1}{2}\cdot\left\|\begin{vmatrix}-1\\0\end{vmatrix}\right\|^2 = \frac{1}{2}$$

$$d_w(A,C) = \frac{1}{2}\left\|\begin{vmatrix}-4\\-2\end{vmatrix} - \begin{vmatrix}-2\\-2\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-2\\0\end{vmatrix}\right\|^2 = \frac{1}{\cancel{2}}\cdot\cancel{\sqrt{4}} = 1$$

$$d_w(A,D) = \frac{1}{2}\cdot\left\|\begin{vmatrix}-4\\-2\end{vmatrix} - \begin{vmatrix}-1\\-2\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-3\\0\end{vmatrix}\right\|^2 = \frac{3}{2}$$

$$d_w(A,E) = \frac{1}{2}\left\|\begin{vmatrix}-4\\-2\end{vmatrix} - \begin{vmatrix}1\\-1\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-5\\-1\end{vmatrix}\right\|^2 = \frac{1}{2}\cdot\sqrt{26} = \frac{\sqrt{26}}{2}$$

$$d_w(A,F) = \frac{1}{2}\left\|\begin{vmatrix}-4\\-2\end{vmatrix} - \begin{vmatrix}1\\1\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-5\\-3\end{vmatrix}\right\|^2 = \frac{\sqrt{34}}{2}$$

$$d_w(A,G) = \frac{1}{2}\left\|\begin{vmatrix}-4\\-2\end{vmatrix} - \begin{vmatrix}2\\3\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-6\\-5\end{vmatrix}\right\|^2 = \frac{1}{2}\cdot\sqrt{36+25} = \frac{\sqrt{61}}{2}$$

$$d_w(A,H) = \frac{1}{2}\left\|\begin{vmatrix}-4\\-2\end{vmatrix} - \begin{vmatrix}3\\2\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-7\\-4\end{vmatrix}\right\|^2 = \frac{1}{2}\sqrt{49+16} = \frac{\sqrt{65}}{2}$$

$$d_w(A,I) = \frac{1}{2}\left\|\begin{vmatrix}-4\\-2\end{vmatrix} - \begin{vmatrix}3\\4\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-7\\-6\end{vmatrix}\right\|^2 = \frac{\sqrt{49+36}}{2} = \frac{\sqrt{85}}{2}$$

$$d_w(A,J) = \frac{1}{2}\left\|\begin{vmatrix}-4\\-2\end{vmatrix} - \begin{vmatrix}4\\3\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-8\\-5\end{vmatrix}\right\|^2 = \frac{\sqrt{64+25}}{2} = \frac{\sqrt{89}}{2}$$

$\vdots$

$$d_w(D,E) = \frac{1}{2}\left\|\begin{vmatrix}-1\\-2\end{vmatrix} - \begin{vmatrix}1\\1\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-2\\-1\end{vmatrix}\right\|^2 = \frac{\sqrt{5}}{2}$$

$$d_w(D,F) = \frac{1}{2}\left\|\begin{vmatrix}-1\\-2\end{vmatrix} - \begin{vmatrix}1\\1\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-2\\-3\end{vmatrix}\right\|^2 = \frac{\sqrt{13}}{2}$$

$$d_w(D,G) = \frac{1}{2}\left\|\begin{vmatrix}-1\\-2\end{vmatrix} - \begin{vmatrix}2\\3\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-3\\-5\end{vmatrix}\right\|^2 = \frac{\sqrt{34}}{2}$$

$$d_w(D,H) = \frac{1}{2}\left\|\begin{vmatrix}-1\\-2\end{vmatrix} - \begin{vmatrix}3\\2\end{vmatrix}\right\|^2 = \left\|\begin{vmatrix}-4\\-3\end{vmatrix}\right\|^2 \cdot \frac{1}{2} = \frac{\sqrt{32}}{2}$$

$$d_w(D,I) = \frac{1}{2}\left\|\begin{vmatrix}-1\\-2\end{vmatrix} - \begin{vmatrix}3\\4\end{vmatrix}\right\|^2 = \frac{\left\|\begin{vmatrix}-4\\-6\end{vmatrix}\right\|^2}{2} = \frac{\sqrt{52}}{2}$$

$$d_w(D,J) = \frac{1}{2}\left\|\begin{vmatrix}-1\\-2\end{vmatrix} - \begin{vmatrix}4\\3\end{vmatrix}\right\|^2 = \frac{\left\|\begin{vmatrix}-5\\-5\end{vmatrix}\right\|^2}{2} = \frac{\sqrt{50}}{2}$$

$\vdots$

$$d_w(F,G) = \frac{1}{2}\left\|\begin{vmatrix}1\\1\end{vmatrix} - \begin{vmatrix}2\\3\end{vmatrix}\right\|^2 = \frac{\left\|\begin{vmatrix}-1\\-2\end{vmatrix}\right\|^2}{2} = \frac{\sqrt{5}}{2}$$

$$d_w(F,H) = \frac{1}{2}\left\|\begin{vmatrix}1\\1\end{vmatrix} - \begin{vmatrix}3\\2\end{vmatrix}\right\|^2 = \frac{\left\|\begin{vmatrix}-2\\-1\end{vmatrix}\right\|^2}{2} = \frac{\sqrt{5}}{2}$$

$$d_w(F,I) = \frac{1}{2}\left\|\begin{vmatrix}1\\1\end{vmatrix} - \begin{vmatrix}3\\4\end{vmatrix}\right\|^2 = \frac{\left\|\begin{vmatrix}-2\\-3\end{vmatrix}\right\|^2}{2} = \frac{\sqrt{13}}{2}$$

$$d_w(F,J) = \frac{1}{2}\left\|\begin{vmatrix}1\\1\end{vmatrix} - \begin{vmatrix}4\\3\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-3\\-2\end{vmatrix}\right\|^2 = \frac{\sqrt{13}}{2}$$

$$d_w(G,H) = \frac{1}{2}\left\|\begin{vmatrix}2\\3\end{vmatrix} - \begin{vmatrix}3\\2\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{vmatrix}-1\\1\end{vmatrix}\right\|^2 = \frac{\sqrt{2}}{2}$$

$$d_w(G,I) = \frac{1}{2}\left\|\begin{vmatrix}2\\3\end{vmatrix} - \begin{vmatrix}3\\4\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{matrix}-1\\-1\end{matrix}\right\|^2 = \frac{\sqrt{2}}{2}$$

$$d_w(G,J) = \frac{1}{2}\left\|\begin{vmatrix}2\\3\end{vmatrix} - \begin{vmatrix}4\\3\end{vmatrix}\right\|^2 = \ldots = \frac{\sqrt{2}}{2}$$

$$d_w(H,I) = \frac{1}{2}\left\|\begin{vmatrix}3\\2\end{vmatrix} - \begin{vmatrix}3\\4\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{matrix}0\\-2\end{matrix}\right\|^2 = 1$$

$$d_w(H,J) = \frac{1}{2}\left\|\begin{vmatrix}3\\2\end{vmatrix} - \begin{vmatrix}4\\3\end{vmatrix}\right\|^2 = \frac{1}{2}\left\|\begin{matrix}-1\\-1\end{matrix}\right\|^2 = \frac{\sqrt{2}}{2}$$

$$d_w(I,J) = \frac{1}{2}\left|\begin{matrix}-1\\1\end{matrix}\right|^2 = \frac{\sqrt{2}}{2}$$

	A|B	C	D	E	F	G	H	i	j
A|B	0								
C	$\frac{2.25}{2}$	0							
D	,	$\frac{1}{2}$	0						
E	|	$\frac{\sqrt{17}}{2}$	$\frac{\sqrt{5}}{2}$	0					
F	|	$\frac{\sqrt{18}}{2}$	$\frac{\sqrt{13}}{2}$	1	0				
G	|	$\frac{\sqrt{61}}{2}$	$\frac{\sqrt{29}}{2}$	$\frac{\sqrt{17}}{2}$	$\frac{\sqrt{2}}{2}$	0			
H	|	$\frac{\sqrt{41}}{2}$	$\frac{\sqrt{32}}{2}$	$\frac{\sqrt{13}}{2}$	$\frac{\sqrt{5}}{2}$	$\frac{\sqrt{2}}{2}$	0		
i	|	$\frac{\sqrt{61}}{2}$	$\frac{\sqrt{52}}{2}$	$\frac{\sqrt{20}}{2}$	$\frac{\sqrt{13}}{2}$	$\frac{\sqrt{2}}{2}$	1	0	
j	|	$\frac{\sqrt{61}}{2}$	$\frac{\sqrt{50}}{2}$	$\frac{5}{2}$	$\frac{\sqrt{13}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	0

$$\mu_{A|B} = \frac{A+B}{2} = \frac{\begin{vmatrix}-4\\-2\end{vmatrix} + \begin{vmatrix}-3\\-2\end{vmatrix}}{2} = \begin{vmatrix}-3.5\\-2\end{vmatrix}$$

$$dw(A|B, C) = \frac{2}{3}\left\|\begin{vmatrix}-3.5\\-2\end{vmatrix} - \begin{vmatrix}-2\\-2\end{vmatrix}\right\|^2 = \frac{2}{3}\left\|\begin{vmatrix}-1.5\\0\end{vmatrix}\right\|^2 = \frac{\sqrt{2.25}}{2}$$

$dw(A|B, d) = \frac{2}{3}\|$

$dw(A|B, E) = \frac{2}{3}\|$

$dw(A|B, F) = \frac{2}{3}$

$dw(A|B, G) = \frac{2}{3}$

$dw(A|B, H) = \frac{2}{3}$

$dw(A|B, i) = \frac{2}{3}$

$dw(A|B, F) = \frac{2}{3}$

$dw(A|B, G) = \frac{2}{3}$

$dw(A, B, H) = \frac{2}{3}$

$dw(A|B, i) = \frac{2}{3}$

$dw(A|B, j) = \frac{2}{3}$

Ward → face. $\|\ldots - \ldots\|^2$

nu il mică.

	A|B	C|D	E	F	G	H	i	j
A|B	0							

C|D 4

E 14.5/3

	E	F	G	H	I	J
E	0					
F	1	0				
G	$\frac{\sqrt{17}}{2}$	$\frac{\sqrt{5}}{2}$	0			
H	$\frac{\sqrt{13}}{2}$	$\frac{\sqrt{5}}{2}$	$\frac{\sqrt{2}}{2}$	0		
I	$\frac{\sqrt{29}}{2}$	$\frac{\sqrt{13}}{2}$	$\frac{\sqrt{2}}{2}$	1	0	
J	$\frac{5}{2}$	$\frac{\sqrt{13}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	0

$$\mu_{C|D} = \frac{C+D}{2} = \frac{\begin{pmatrix}-3\\-4\end{pmatrix}}{2} = \begin{pmatrix}-1.5\\-2\end{pmatrix}$$

$$d_w(A|B, C|D) = \frac{2\cdot 2}{2+2} = \frac{4}{4} = 1 \cdot \left\| \begin{pmatrix}-3.5\\-2\end{pmatrix} - \begin{pmatrix}-1.5\\-2\end{pmatrix} \right\|^2 =$$

$$= \left\| \begin{pmatrix}-2\\0\end{pmatrix} \right\|^2 = 4$$

$$d_w(C|D, E) = \frac{2}{3} \left\| \begin{pmatrix}-1.5\\-2\end{pmatrix} - \begin{pmatrix}1\\-1\end{pmatrix} \right\|^2 = \frac{2}{3} \left\| \begin{pmatrix}-2.5\\-1\end{pmatrix} \right\|^2$$

$$= \frac{2}{3} \cdot 7.25 = \frac{14.5}{3}$$

	A|B	C|D	E	F	G|H	I	J
A|B	0						
C|D	4	0					
E		$\frac{14.5}{3}$	0				
F			1	0			
G|H				$\frac{5}{2}$	0		
I					$\frac{5}{3}$	0	
J						1	0

$$\mu_{G|H} = \begin{pmatrix}2.5\\2.5\end{pmatrix}$$

$$d_w(G|H, I) = \frac{2}{3} \left\| \begin{pmatrix}2.5\\2.5\end{pmatrix} - \begin{pmatrix}3\\4\end{pmatrix} \right\|^2 = \left\| \begin{pmatrix}-0.5\\-1.5\end{pmatrix} \right\|^2 \frac{2}{3} = \frac{5}{3}$$

$$d_w(G|H, J) = \text{—}\,''\,\text{—} \begin{pmatrix}5\\3\end{pmatrix} \Big\|^2 = \frac{2}{3} \left\| \begin{pmatrix}-1.5\\-0.5\end{pmatrix} \right\|^2 = \frac{5}{3}$$

	A|B	C|D	E|F	G|H	i	j
A|B	0					
C|D	4	0				
E|F		9.75	$\boxed{0}$			
G|H			10.25	0		
i			8.5	$\frac{5}{3}$	0	
j					1	

$$\mu_{E|F} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$d_W(C|D, E|F) = \frac{2\cdot 2}{2+2} \left\| \begin{pmatrix} -1.5 \\ -2 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\|^2 = 1 \cdot \begin{pmatrix} -2.5 \\ -2 \end{pmatrix}^2 =$$

$$= (6.25 + 4) = 10.25$$

d_W

$$d_W(E|F, G|H) = 1 \cdot \left\| \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 2.5 \\ 2.5 \end{pmatrix} \right\| = \begin{pmatrix} -1.5 \\ -2.5 \end{pmatrix}^2 = 2.25 + 6.25 =$$

$$= 8.5$$

	A|B	C|D	E|F	G|H	i|j
A|B	0				
C|D	4	0			
E|F		10.25	$\boxed{0}$		
G|H			8.5	0	
i|j				4.5	0

$$\mu_{i,j} = \begin{pmatrix} 3.5 \\ 3.5 \end{pmatrix}$$

$$d_W(G|H, i|j) = 1 \cdot \left\| \begin{pmatrix} 2.5 \\ 2.5 \end{pmatrix} - \begin{pmatrix} 3.5 \\ 3.5 \end{pmatrix} \right\| = \begin{pmatrix} -1.5 \\ -1.5 \end{pmatrix}^2 =$$

$$= 2.25 + 2.25 = 4.5$$

$\mu_{A|B} = \begin{pmatrix}-3.5\\-2\end{pmatrix}$ $\quad \mu_{C|D} = \begin{pmatrix}-1.5\\-2\end{pmatrix}$

$\mu_{E|F} = \begin{pmatrix}1\\0\end{pmatrix}$ $\quad \mu_{G|H} = \begin{pmatrix}2.5\\2.5\end{pmatrix}$

$\mu_{i|j} = \begin{pmatrix}3.5\\3.5\end{pmatrix}$

	A,B|C,D	E|F	G|H	i|j
A,B|C,D	0			
E|F	21.6	$\boxed{0}$		
G|H		8.5	0	
i|j			5.5	0

$$\mu_{AB|CD} = \frac{A+B+C+D}{4} = \frac{\begin{pmatrix}-10\\-8\end{pmatrix}}{4} = \begin{pmatrix}-2.5\\-2\end{pmatrix}$$

$$d_w(A,B|C,D,\ E|F) = \frac{4\cdot 2}{4+2}\left\|\begin{pmatrix}-2.5\\-2\end{pmatrix} - \begin{pmatrix}1\\0\end{pmatrix}\right\|^2 =$$

$$= \frac{8}{6}\left\|\begin{pmatrix}-3.5\\-2\end{pmatrix}\right\|^2 = 21.6$$

	A,B|C,D	E|F	G|H|i|j
A,B|C,D	0		
E|F	21.6	$\boxed{0}$	
G,H|i,j		17.3	0

$$\mu_{G,H|i,j} = \frac{\begin{pmatrix}12\\12\end{pmatrix}}{4} = \begin{pmatrix}3\\3\end{pmatrix}$$

$$d_w(G,H|i,j,\ E|F) = \frac{8}{6}\left\|\begin{pmatrix}3\\3\end{pmatrix} - \begin{pmatrix}1\\0\end{pmatrix}\right\|^2 = \left|\begin{pmatrix}2\\3\end{pmatrix}\right|^2 \cdot \frac{8}{6} =$$

$$= \frac{13\cdot\cancel{8}^{4}}{\cancel{6}_{3}} = \frac{52}{3} = 17.3$$

A, B | C, D   E, F | G H, i j

A, B | C, D   O

E, F | G H, i j   - - -   O