La proportion analogique dans les groupes. Application aux permutations et aux matrices.

Nelly Barbot, Laurent Miclet

Publication interne IRISA numéro 1914

La proportion analogique dans les groupes. Application aux permutations et aux matrices.

Nelly Barbot et Laurent Miclet

Projet Cordial

Résumé : Ce rapport de recherche étudie la notion de proportion analogique entre quatre éléments d'un groupe, en respectant les axiomes proposés par Lepage [1] et en suivant aussi la notion de factorisation, donnée comme fondamentale par Stroppa et Yvon pour définir la proportion analogique [2]. On montre d'abord qu'en général, l'équation analogique n'a pas de solution dans un groupe non commutatif. On s'intéresse ensuite aux conditions que doivent respecter trois éléments pour qu'il existe un quatrième en proportion analogique avec eux. Ces conditions sont présentées de différentes manières, que l'on démontre comme équivalentes. On définit ensuite la notion de dissemblance analogique à partir d'une distance sur le groupe.

On s'intéresse ensuite au cas particulier du groupe des permutations sur un ensemble à n éléments. On caractérise les conditions d'existence et on montre comment construire des proportions analogiques à partir de deux éléments et on en dénombre les cas possibles. On présente une distance sur le groupe et on en déduit une dissimilarité analogique entre quatre permutations.

Pour terminer, on étudie le groupe des matrices inversibles. On caractérise les conditions d'existence et on montre comment construire des proportions analogiques à partir de deux matrices. Quelques cas particuliers sont présentés. Enfin, on définit une dissimilarité analogique entre quatre matrices.

Mots clés : Proportion analogique, groupes, permutations, matrices inversibles.

(Abstract: pto)

IRISA, Campus universitaire de Beaulieu, 35042 Rennes Cedex (France)

Téléphone : (33) 02 99 84 71 00 — Télécopie : (33) 02 99 84 71 71 — www.irisa.fr

Abstract: In this report, we explore the notion of analogical proportion between four elements of a group, with respect to the axioms given by Lepage [1] and using also the notion of factorization, the foundation of analogical proportion according to Stroppa et Yvon [2]. We firstly show that the analogical equation has in general no solution in a non commutative group. We explore the conditions that three elements must respect to define a fourth as being in analogical proportion with them. These conditions are presented according to different forms, that we show to be equivalent. Finally, we define the notion of analogical dissimilarity, provided that a distance is given on the group.

We focus in the next section on the permutation group of a finite set. We characterize the existence of an analogical proportion and we show how to built a set of analogical proportion when two elements are given. The size of this set is computed. We present a distance on the permutation group and the consequent analogical dissimilarity between four permutations.

To finish, we study the group of inversible matrices. We characterize the existence of an analogical proportion and we show how to built a set of analogical proportion when two elements are given. Some particular cases are studied. An analogical dissimilarity between four matrices is defined.

Key-words: Analogical Proportion, Group, Permutation, Inversible Matrix.

Table des matières

1 Introduction **3**

2 La proportion analogique et l'apprentissage **4**
2.1 La proportion analogique 4
2.2 Sémantique et exemples 5
2.3 Application à l'apprentissage et à la génération 6

3 La proportion analogique dans les groupes **6**
3.1 Trois définitions équivalentes 6
3.2 Propriétés 8
3.3 Construction de quadruplets en analogie 10
3.4 Dissemblance analogique 10

4 Analogie entre permutations **13**
4.1 Définition 13
4.2 Analyse 13
4.3 Dissemblance analogique 16

5 Analogie entre matrices inversibles **18**
5.1 Définition 18
5.2 Recherche des proportions analogiques 19
5.3 Dissemblance analogique 22

6 Conclusion : quelques pistes pour l'utilisation en apprentissage **26**

1 Introduction

La notion d'analogie, en particulier de proportion analogique, a été étudiée comme une des modalités du raisonnement depuis Aristote [1, 3]. Il s'agit d'un raisonnement par généralisation, différent de l'induction, qui modélise une autre forme d'apprentissage. Son application en intelligence artificielle a été envisagée et testée très tôt, entre autres dans [4], mais ce n'est que depuis une vingtaine d'années qu'une forme opérationnelle particulière du raisonnement analogique s'est développée : le raisonnement à partir de cas (voir par exemple [5]). Récemment, un intérêt croissant s'est manifesté pour un point de vue différent sur l'analogie, la *proportion analogique*. Cette notion est en effet définie rigoureusement et son application dans des espaces de représentation de natures variées a été développée avec des résultats opérationnels intéressants. Son application à l'apprentissage et à la génération est conceptuellement simple ; en revanche, comme dans bien des domaines de l'intelligence artificielle, certains problèmes de complexité algorithmique restent à surmonter quand on l'applique.

Nous présentons dans le premier paragraphe la définition axiomatique de la proportion analogique et nous en montrons l'application à l'apprentissage supervisé et à la génération, en nous appuyant sur des exemples opérationnels. Dans le second paragraphe, nous étudions le cas général où les objets étudiés sont des éléments d'un groupe non commutatif. Dans le troisième, nous considérons le cas particulier des permutations, dans le quatrième celui des matrices. Pour conclure, nous nous intéressons à l'application potentielle de la proportion analogique dans de tels objets à l'apprentissage supervisé.

2 La proportion analogique et l'apprentissage

2.1 La proportion analogique

Définition 1 (Proportion analogique)**.** *Une proportion analogique sur un ensemble $\mathbb{E}$ est une relation sur $\mathbb{E}^4$ telle que, pour toute famille de quatre éléments A, B, C et D en relation dans cet ordre (ce qui est noté $A : B :: C : D$), on ait les deux équivalences :*

$$\begin{array}{lcl} A : B :: C : D & \Leftrightarrow & C : D :: A : B \\ A : B :: C : D & \Leftrightarrow & A : C :: B : D \end{array}$$

D'autre part, pour tout couple (A,B) d'éléments de $\mathbb{E}$, on doit avoir $A : B :: A : B$

On montre alors que cinq autres relations sont équivalentes à $A : B :: C : D$:

$$B : A :: D : C \qquad D : B :: C : A \qquad D : C :: B : A \qquad B : D :: A : C \qquad C : A :: D : B$$

Au total, ces axiomes donnent donc pour équivalentes huit relations entre quatre objets.

Définition 2 (Équation analogique)**.** *Soit A, B et C trois éléments de $\mathbb{E}$, on appelle* équation analogique *une expression du type $A : B :: C : X$ d'inconnue $X \in \mathbb{E}$.*

Une équation analogique peut avoir zéro, une seule ou plusieurs solutions.

On ajoute souvent aux axiomes qui définissent la proportion analogique celui du *déterminisme*, qui exige que l'équation suivante n'ait qu'une seule solution :

$$A : B :: A : X \quad \Rightarrow \quad X = B$$

Remarquons qu'il y a 24 façons d'arranger quatre objets en une équation analogique, qui se réduisent à trois classes d'équivalence dont les représentants sont par exemple :

$$A : B :: C : D \qquad B : A :: C : D \quad \text{et} \quad C : B :: A : D \ .$$

2.2 Sémantique et exemples

La relation $A : B :: C : D$ s'interprète classiquement comme « A est à B comme C est à D » et signifie que pour transformer A en B, il faut faire les mêmes opérations que pour transformer C en D.

Les domaines sur lesquels la théorie et la sémantique de la proportion analogique ont été étudiées sont les suivants :

$\{0,1\}^d$: Quatre objets définis par des vecteurs d'attributs binaires sont en proportion analogique lorsque, sur chaque coordonnée, on a l'une des 6 relations suivantes :

$$\mathit{0} : \mathit{0} :: \mathit{1} : \mathit{1} \ , \quad \mathit{1} : \mathit{1} :: \mathit{0} : \mathit{0} \ , \quad \mathit{1} : \mathit{0} :: \mathit{1} : \mathit{0} \ , \quad \mathit{0} : \mathit{1} :: \mathit{0} : \mathit{1} \ , \quad \mathit{0} : \mathit{0} :: \mathit{0} : \mathit{0} \ \text{ ou } \ \mathit{1} : \mathit{1} :: \mathit{1} : \mathit{1} \ .$$

$\mathbb{R}^d$: Quatre objets définis par des vecteurs d'attributs numériques sont en proportion analogique additive quand ils forment un parallélogramme, c'est à dire que sur chaque coordonnée on a
$$A_i : B_i :: C_i : D_i \ \Leftrightarrow \ A_i + D_i = B_i + C_i \, .$$

$\mathbb{R}^d$: Quatre objets définis par des vecteurs d'attributs numériques sont en proportion analogique multiplicative quand sur chaque coordonnée on a

$$A_i : B_i :: C_i : D_i \ \Leftrightarrow \ A_i \times D_i = B_i \times C_i \, .$$

$\Sigma^\star$: Pour la théorie et l'algorithmique de la proportion analogique entre mots, on peut se référer à [6, 7]. On a par exemple la proportion :

`survoler` est à `volais` comme `surcoter` est à `surcotais`.

Arbres ordonnés : L'étude de la proportion analogique sur les arbres a été entreprise dans [6, 8].

D'autre part, l'étude du raisonnement par proportion analogique en logique des propositions (et dans son interprétation en théorie des ensembles) et en logique floue a été abordée dans [1, 9].

Pour étendre ces études, nous proposons dans cet article d'étudier la proportion analogique dans les ensembles munis d'une structure de groupe (par exemple les matrices inversibles). Nous établissons des propriétés générales et nous donnons des algorithmes pour établir, par exemple, si quatre matrices sont en proportion analogique ou sinon pour mesurer de combien elles « ratent » cette proportion analogique.

2.3 Application à l'apprentissage et à la génération

Soit la séquence de lettres `cherchant`, dont on veut apprendre la supervision ou la classe [1] ; supposons que dans l'ensemble d'apprentissage se trouvent les trois séquences : `voler`, `volerai`, `chercher`, avec respectivement pour classes *infinitif*, *futur*, *infinitif*. On attribuera à `chercherai` la classe *futur*, par un raisonnement qui s'énonce informellement comme ceci. Puisque, dans l'univers des séquences

`voler` est à `volerai` comme `chercher` est à `chercherai`

la supervision de `cherchant` est donc la solution de l'équation (dans l'univers des classes) :

infinitif est à *futur* comme *infinitif* est à x d'où : $x =$*futur*

C'est en particulier sur ce principe qu'ont été réalisées des expériences de transcription orthographique-phonétique, de traduction, et d'apprentissage supervisé classique (voir [7, 10]). Cette technique permet aussi la génération d'objets, par exemple pour rééquilibrer des bases de données d'apprentissage de séquences, comme dans [11]. Dans ce cas, on prend trois objets a, b, c par exemple dans les classes ω_1, ω_2 et ω_1 et on génère des objets de la classe ω_2 en produisant des solutions exactes ou approchées à l'équation analogique $a : b :: c : x$.

3 La proportion analogique dans les groupes

3.1 Trois définitions équivalentes

Soit un $\mathbb{G}$ un ensemble muni d'une loi de composition interne $\otimes$, tel que $(\mathbb{G},\otimes)$ soit un groupe non commutatif. L'élément neutre est noté 1 et A^{-1} désigne l'inverse d'un élément A. Nous proposons dans ce paragraphe trois définitions différentes de la proportion analogique dans $\mathbb{G}$, avant de montrer qu'elles sont équivalentes. La première est basée sur la notion de factorisation, qui est, comme il a été montré dans [6], un principe algébrique unificateur pour définir la proportion analogique dans des univers variés. La seconde est une conséquence directe de la première (*a priori* plus restrictive). La troisième est une sorte de représentation sous forme de parallélogramme non commutatif.

Définition 3 (Propriété analogique $P1$)**.** *Soit A, B, C et D quatre éléments de $\mathbb{G}$, on dit que (A,B,C,D) vérifie la propriété analogique $P1$ s'il existe*

- *quatre éléments R, T, S et U de $\mathbb{G}$ tels que $A = R \otimes T$, $B = R \otimes U$, $C = S \otimes T$, $D = S \otimes U$,*
- *et quatre éléments R', T', S' et U' de $\mathbb{G}$ tels que $A = R' \otimes T'$, $B = S' \otimes T'$, $C = R' \otimes U'$ et $D = S' \otimes U'$.*

1. Le terme *supervision* est plus général que celui d'*étiquette* ou de *classe*, puisqu'il couvre aussi bien les cas de l'apprentissage de règles de classification, de la régression et le cas où la supervision est un objet structuré, comme une séquence.

Définition 4 (Propriété analogique $P2$)**.** *Soit A, B, C et D quatre éléments de $\mathbb{G}$, on dit que (A,B,C,D) vérifie la propriété analogique $P2$ si les deux égalités suivantes sont respectées*

$$C \otimes A^{-1} \otimes B \;=\; B \otimes A^{-1} \otimes C\,, \tag{1}$$
$$D \;=\; C \otimes A^{-1} \otimes B\,. \tag{2}$$

Définition 5 (Propriété analogique $P3$)**.** *Soit A, B, C et D quatre éléments de $\mathbb{G}$, on dit que (A,B,C,D) vérifie la propriété analogique $P3$ s'il existe un couple (P,Q) d'éléments de $\mathbb{G}$ tel que*

$$A \otimes P = B, \qquad C \otimes P = D, \qquad A \otimes Q = C \qquad \text{et} \qquad B \otimes Q = D\,.$$

Il est conseillé de se reporter à la figure 1 pour visualiser les trois définitions et avoir une intuition de leur équivalence. Celle-ci va être démontrée maintenant.

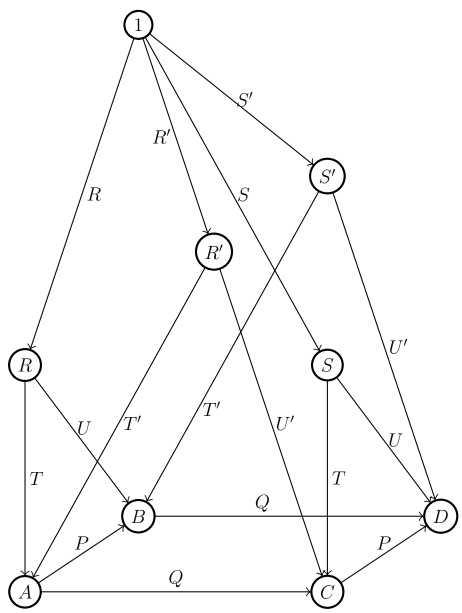

Fig. 1 – *Visualisation de l'équivalence entre trois définitions de la proportion analogique. L'arc orienté qui va du nœud A au nœud B avec l'étiquette P exprime la relation $A \otimes P = B$.*

Proposition 1. *Les trois propriétés $P1$, $P2$ et $P3$ sont équivalentes. Quatre éléments A, B, C et D de $\mathbb{G}$, tels que (A,B,C,D) vérifie l'une de ces propriétés, sont alors dits* en proportion analogique*, conformément à la définition 1, et l'on note*

$$A : B :: C : D \quad .$$

Démonstration :

1. Si (A,B,C,D) vérifie $P1$, alors il est immédiat de vérifier qu'il satisfait $P2$.
2. Si (A,B,C,D) vérifie $P2$, il vérifie également $P3$: il suffit de poser $Q = A^{-1} \otimes C$ et $P = A^{-1} \otimes B$.
3. Si (A,B,C,D) vérifie $P3$ alors il vérifie $P1$: les matrices $R = D \otimes Q^{-1}$, $S = D$, $T = P^{-1}$ et $U = 1$ respectent la première condition de $P1$ et $R' = D \otimes P^{-1}$, $S' = D$, $T' = Q^{-1}$ et $U' = 1$ respectent la seconde condition de $P1$.

$\blacksquare$

3.2 Propriétés

Dans ce paragraphe, nous énonçons des propriétés de la proportion analogique découlant directement de la structure de groupe de $\mathbb{G}$ et des définitions précédentes. Les démonstrations sont donc le plus souvent omises. La propriété ci-dessous rattache cette proportion analogique aux axiomes proposés par Lepage dans [1].

Proposition 2. *Si quatre éléments A, B, C et D de $\mathbb{G}$ sont en proportion analogique $A : B :: C : D$, alors on a de manière équivalente*

- $C : D :: A : B$,
- $A : C :: B : D$.

D'autre part, on a la propriété : $A : B :: A : D \quad \Rightarrow \; D = B$.

On a donc les huit formes équivalentes de la proportion analogique, en permutant les éléments à partir des deux règles ci-dessus. En conséquence, on peut énoncer le résultat suivant.

Proposition 3. *Si quatre éléments A, B, C et D de $\mathbb{G}$ sont en proportion analogique $A : B :: C : D$, alors on a de manière équivalente les quatre propriétés ci-dessous :*

- $A = B \otimes D^{-1} \otimes C = C \otimes D^{-1} \otimes B$
- $B = A \otimes C^{-1} \otimes D = D \otimes C^{-1} \otimes A$
- $C = A \otimes B^{-1} \otimes D = D \otimes B^{-1} \otimes A$
- $D = B \otimes A^{-1} \otimes C = C \otimes A^{-1} \otimes B$

Une conséquence immédiate de la proposition précédente est la stabilité de la proposition analogique par les opérations $\otimes$ à gauche, à droite, ainsi que par inversion.

Proposition 4. *Soit A, B, C, D et E cinq éléments de $\mathbb{G}$, les propriétés suivantes sont équivalentes*

- $A : B :: C : D$
- $A^{-1} : B^{-1} :: C^{-1} : D^{-1}$

– $E \otimes A : E \otimes B :: E \otimes C : E \otimes D$
– $A \otimes E : B \otimes E :: C \otimes E : D \otimes E$

La proposition ci-dessous offre une nouvelle définition équivalente à la proportion analogique. Elle se démontre aisément à partir de la propriété analogique $P3$.

Proposition 5 (Propriété analogique $P4$)**.** *Soit A, B, C et D quatre éléments de $\mathbb{G}$, on a $A : B :: C : D$ si et seulement s'il existe un couple (E,F) d'éléments de $\mathbb{G}$ vérifiant les quatre conditions :*

$$C = E \otimes A, \qquad D = E \otimes B, \qquad C = A \otimes F \qquad \textit{et} \qquad D = B \otimes F\,.$$

La proposition suivante, très proche de $P4$, permet de simplifier l'écriture des proportions analogiques.

Proposition 6. *Soit A et B deux éléments de $\mathbb{G}$, l'ensemble des couples (C,D) d'éléments de $\mathbb{G}$ vérifiant $A : B :: C : D$ est décrit par*

$$\{(E \otimes A, E \otimes B) \mid E \in \mathbb{G} \textit{ et } E \otimes P = P \otimes E\}$$

où $P = B \otimes A^{-1}$. De même, cet ensemble peut s'écrire

$$\{(A \otimes F, B \otimes F) \mid F \in \mathbb{G} \textit{ et } F \otimes Q = Q \otimes F\}$$

où $Q = A^{-1} \otimes B$.

Démonstration : soit A et B deux éléments de $\mathbb{G}$,

$$\begin{aligned}
&\big\{(C,D) \in \mathbb{G}^2 \mid \ A : B :: C : D \ \big\} \\
&\qquad = \big\{(C,D) \in \mathbb{G}^2 \ \big| D = C \otimes A^{-1} \otimes B = B \otimes A^{-1} \otimes C\big\} \\
&\qquad = \big\{(E \otimes A, D) \,\big|\, (E,D) \in \mathbb{G}^2 \text{ et } D = E \otimes B = B \otimes A^{-1} \otimes E \otimes A\big\} \\
&\qquad = \big\{(E \otimes A, E \otimes B) \,\big|\, E \in \mathbb{G} \text{ et } E \otimes B = B \otimes A^{-1} \otimes E \otimes A\big\} \\
&\qquad = \big\{(E \otimes A, E \otimes B) \,\big|\, E \in \mathbb{G} \text{ et } E \otimes B \otimes A^{-1} = B \otimes A^{-1} \otimes E\big\} \\
&\qquad = \{(E \otimes A, E \otimes B) \mid E \in \mathbb{G} \text{ et } E \otimes P = P \otimes E\}
\end{aligned}$$

On procède de même pour établir la seconde description de cet ensemble à l'aide des matrices F commutant avec $A^{-1} \otimes B$.

$\blacksquare$

3.3 Construction de quadruplets en analogie

Grâce à la proposition 6 et aux propriétés de symétrie décrites dans la proposition 2, la recherche des proportions analogiques $A : B :: C : D$ étant donné un couple (A,B) ou (C,D) ou (A,C) ou (B,D) revient à la résolution dans $\mathbb{G}$ d'une équation d'inconnue X de la forme

$$P \otimes X \;=\; X \otimes P \tag{3}$$

où P est un élément de $\mathbb{G}$ dont l'expression dépend du couple considéré. Les proportions analogiques sont alors, selon le couple initialement donné,

$$\begin{aligned}
A : B &:: (X \otimes A) : (X \otimes B) && \text{où} \quad P = B \otimes A^{-1},\\
(X \otimes C) : (X \otimes D) &:: C : D && \text{où} \quad P = D \otimes C^{-1},\\
A : (X \otimes A) &:: C : (X \otimes C) && \text{où} \quad P = C \otimes A^{-1},\\
(X \otimes B) : B &:: (X \otimes D) : D && \text{où} \quad P = D \otimes B^{-1}.
\end{aligned}$$

De même, il est aisé de vérifier que la recherche des proportions analogiques $A : B :: C : D$ suite à la donnée de (B,C) ou (A,D) se ramène à la résolution dans $\mathbb{G}$ d'une équation de la forme (3) de la manière suivante

$$\begin{aligned}
(X \otimes B) : B &:: C : \left(X^{-1} \otimes C\right) && \text{où} \quad P = B \otimes C^{-1},\\
A : (X \otimes A) &:: \left(X^{-1} \otimes D\right) : D && \text{où} \quad P = A \otimes D^{-1}.
\end{aligned}$$

Remarquons qu'une autre manière de construire des proportions analogiques en partant d'un couple d'éléments est de résoudre dans $\mathbb{G}$ d'une équation de la forme

$$P \otimes X \;=\; X \otimes Q \tag{4}$$

où P et Q sont deux éléments de $\mathbb{G}$ conjugués (i.e. il existe un élément M de $\mathbb{G}$ tel que $M \otimes P = Q \otimes M$). Par exemple, si l'on considère le couple (B,C), la recherche des éléments $A \in \mathbb{G}$ solutions de (1) correspond à la résolution de (4) avec $P = C \otimes B^{-1}$ et $Q = B^{-1} \otimes C$. Quant à la relation de conjugaison entre P et Q, elle est donnée par $P \otimes C = C \otimes Q$. La proportion analogique est ensuite complétée par le calcul de D à l'aide de la relation (2).

L'équation (3) étant un cas particulier de (4), nous nous intéresserons dans les paragraphes suivants à la résolution de (4) dans le groupe des permutations et des matrices inversibles.

3.4 Dissemblance analogique

Ce paragraphe a pour but de donner la définition et d'établir les propriétés d'une quantité réelle positive ou nulle qui mesure de combien quatre éléments d'un groupe « ratent » la proportion analogique. Cette notion a déjà été définie et utilisée en apprentissage d'objets à attributs binaires et nominaux, dans [12] et étudiée pour les séquences dans [7].

3.4.1 Définition et propriétés

On considère à présent le groupe $(\mathbb{G},\otimes)$ muni d'une distance Δ. On peut alors définir la notion de dissemblance analogique de la façon suivante.

Définition 6 (Dissemblance Analogique)**.** *Soit A, B, C et D quatre éléments de $\mathbb{G}$, la dissemblance analogique associée à (A,B,C,D) est définie par*

$$\begin{aligned} DA(A,B,C,D) &= \Delta(A^{-1}\otimes C,B^{-1}\otimes D)+\Delta(A^{-1}\otimes B,C^{-1}\otimes D) \\ &+ \Delta(C^{-1}\otimes A,D^{-1}\otimes B)+\Delta(B^{-1}\otimes A,D^{-1}\otimes C)\,. \end{aligned}$$

Les symétries dans l'expression de la dissemblance analogique permettent d'obtenir les propriétés suivantes.

Proposition 7. *Pour tout $(A,B,C,D)\in\mathbb{G}^4$, on a*

1. $DA(A,B,C,D)=0 \Leftrightarrow \ A:B::C:D\ $.
2. $DA(A,B,C,D)=DA(A,C,B,D)$.
3. $DA(A,B,C,D)=DA(C,D,A,B)$.
4. *La dissemblance analogique est stable par opération $\otimes$ à gauche et à droite.*
5. *En général,* $DA(A,B,C,D)\neq DA(B,A,C,D)$.

Démonstration : la première propriété est évidente, la seconde, la troisième et quatrième découlent de la symétrie de la formule. Quant à la dernière, elle exprime que l'ordre des éléments A, B, C et D importe dans une proportion analogique.

■

Afin de représenter géométriquement la dissemblance analogique entre quatre éléments A, B, C et D, on définit les éléments

$$\begin{aligned} A_1 &= C\otimes D^{-1}\otimes B \\ A_2 &= B\otimes D^{-1}\otimes C \\ D_1 &= B\otimes A^{-1}\otimes C \\ D_2 &= C\otimes A^{-1}\otimes B \end{aligned}$$

qui sont représentés sur la figure 2 et l'on a

$$\begin{aligned} DA(A,B,C,D) &= \Delta(B^{-1}\otimes D,B^{-1}\otimes D_1)+\Delta(C^{-1}\otimes D,C^{-1}\otimes D_2) \\ &+ \Delta(C^{-1}\otimes A,C^{-1}\otimes A_1)+\Delta(B^{-1}\otimes A,B^{-1}\otimes A_2)\,. \end{aligned}$$

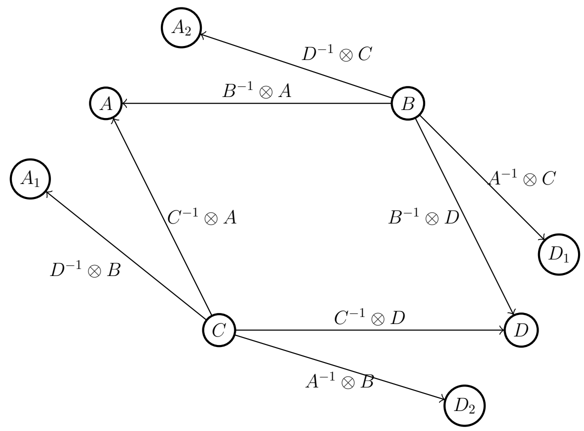

FIG. 2 – *La géométrie de la dissemblance analogique.*

La distance Δ peut avoir une propriété supplémentaire : l'invariance à droite et à gauche pour l'opération $\otimes$. Dans ce cas, pour trois éléments quelconques A, B et C, on a

$$\Delta(A,B) \;=\; \Delta(C \otimes A, C \otimes B) \;=\; \Delta(A \otimes C, B \otimes C)$$

et l'on peut écrire plus simplement

$$DA(A,B,C,D) \;=\; \Delta(D,D_1) + \Delta(D,D_2) + \Delta(A,A_1) + \Delta(A,A_2)\,.$$

3.4.2 Remarques

On peut noter qu'il y a au moins une autre définition possible du même type. Celle qui est proposée mesure une sorte de distance cumulée entre l'analogie exacte et la configuration réelle en prenant B et C comme points fixes et en construisant deux valeurs idéales pour A et deux pour D. On peut s'appuyer de manière alternative sur A et D comme points fixes. Si la distance est invariante par composition, le résultat est le même.

On peut noter aussi que l'inégalité triangulaire de la distance n'intervient pas dans la définition de la dissemblance analogique. La propriété (souhaitable)

$$DA(A,B,C,D) \;\leq\; DA(A,B,E,F) + DA(E,F,C,D) \tag{5}$$

semble cependant, si elle est vraie, difficile à démontrer. Nous conjecturons qu'elle est fausse dans le cas général. Cependant, la proposition suivante montre qu'elle peut se produire dans certains cas.

Proposition 8. *Si Δ est invariante à gauche et à droite pour l'opération $\otimes$ alors la dissemblance analogique vérifie l'« inégalité triangulaire » (5) pour tout six-uplet (A,B,C,D,E,F) d'éléments de $\mathbb{G}$.*

Démonstration : soit A,B,C,D,E et F six éléments de $\mathbb{G}$, on a grâce à l'invariance de Δ pour l'opération $\otimes$ à gauche et à droite

$$\begin{aligned} DA(A,B,C,D) &= \Delta(A^{-1}\otimes C,B^{-1}\otimes D)+\Delta(A^{-1}\otimes B,C^{-1}\otimes D) \\ &+ \Delta(C^{-1}\otimes A,D^{-1}\otimes B)+\Delta(B^{-1}\otimes A,D^{-1}\otimes C) \\ &= \Delta(B\otimes A^{-1},D\otimes C^{-1})+\Delta(A^{-1}\otimes B,C^{-1}\otimes D) \\ &+ \Delta(A\otimes B^{-1},C\otimes D^{-1})+\Delta(B^{-1}\otimes A,D^{-1}\otimes C) \end{aligned} \quad (6)$$

Δ étant une distance, elle vérifie l'inégalité triangulaire que l'on applique à chacun des termes ci-dessus de la façon suivante

$$\begin{aligned} \Delta(B\otimes A^{-1},D\otimes C^{-1}) &\le \Delta(B\otimes A^{-1},F\otimes E^{-1})+\Delta(F\otimes E^{-1},D\otimes C^{-1}) \\ \Delta(A^{-1}\otimes B,C^{-1}\otimes D) &\le \Delta(A^{-1}\otimes B,E^{-1}\otimes F)+\Delta(E^{-1}\otimes F,C^{-1}\otimes D) \\ \Delta(A\otimes B^{-1},C\otimes D^{-1}) &\le \Delta(A\otimes B^{-1},E\otimes F^{-1})+\Delta(E\otimes F^{-1},C^{-1}\otimes D) \\ \Delta(B^{-1}\otimes A,D^{-1}\otimes C) &\le \Delta(B^{-1}\otimes A,F^{-1}\otimes E)+\Delta(F^{-1}\otimes E,D^{-1}\otimes C) \end{aligned}$$

En sommant ces quatres dernières relations, on obtient, d'après l'écriture (6) de la dissemblance analogique entre quatre éléments,

$$DA(A,B,C,D) \le DA(A,B,E,F)+DA(E,F,C,D)\,.$$

■

4 Analogie entre permutations

4.1 Définition

L'ensemble des permutations $\mathcal{S}_n$ sur un ensemble fini à n éléments muni de l'opération de composition $\circ$ forme un groupe, appelé groupe symétrique d'ordre n. En conséquence, on peut énoncer la définition de la proportion analogique P_2 entre quatre éléments de $\mathcal{S}_n$.

Définition 7 (Proportion analogique entre permutations)**.** *Quatre permutations σ_1, σ_2, σ_3 et σ_4 sur un ensemble à n éléments sont en proportion analogique $\sigma_1 : \sigma_2 :: \sigma_3 : \sigma_4$ si les deux égalités suivantes sont respectées*

$$\sigma_2\circ\sigma_1^{-1}\circ\sigma_3 = \sigma_3\circ\sigma_1^{-1}\circ\sigma_2\,, \quad (7)$$

$$\sigma_4 = \sigma_2\circ\sigma_1^{-1}\circ\sigma_3\,. \quad (8)$$

4.2 Analyse

D'après le paragraphe 3.3, on s'intéresse à la résolution dans $\mathcal{S}_n$ de l'équation

$$\sigma_P\circ\sigma = \sigma\circ\sigma_Q \quad (9)$$

où σ_P et σ_Q sont des éléments conjugués de $\mathcal{S}_n$. Par exemple, pour σ_2 et σ_3 donnés, les solutions σ_1 de (7) vérifient (9) pour $\sigma_P=\sigma_3\circ\sigma_2^{-1}$ et $\sigma_Q=\sigma_2^{-1}\circ\sigma_3$.

4.2.1 Résolution et dénombrement

Comme les deux permutations σ_P et σ_Q sont conjuguées, leur décomposition en cycles à supports disjoints sont identiques : elles ont le même nombre k_1 de cycles de longueur 1, le même nombre k_2 de cycles de longueur 2, etc. Cette propriété est un résultat classique de la théorie des groupes de permutations (voir par exemple [13]). Par conséquent, la condition nécessaire et suffisante pour que σ soit solution de (9) est la suivante : elle doit transformer un cycle de σ_Q en un cycle de même longueur de σ_P en respectant l'ordre des éléments des cycles. En effet, soit σ une solution de (9) et

$$\left(i\,,\,\sigma_Q(i)\,,\,\ldots\,,\,\sigma_Q^{s-1}(i)\right)$$

un cycle de σ_Q de longueur s, où $1 \leq s \leq n$. L'équation (9) nous donne, pour tout $0 \leq j \leq s-1$,

$$\sigma_P\left(\sigma \circ \sigma_Q^j(i)\right) = \begin{cases} \sigma \circ \sigma_Q^{j+1}(i) & \text{si } j < s-1, \\ \sigma\,(i) & \text{si } j = s-1. \end{cases}$$

Autrement dit, si σ est solution de (9),

$$\left(\sigma(i)\,,\,\sigma \circ \sigma_Q(i)\,,\,\ldots\,,\,\sigma \circ \sigma_Q^{s-1}(i)\right)$$

est un cycle de longueur s de σ_P. La réciproque est également simple à vérifier.

On peut à présent s'intéresser au nombre de solutions de l'équation (9). On considère les décompositions de σ_P et σ_Q en cycles de support disjoint. Soit k_s le nombre de cycles de longueur s de σ_P (et de σ_Q), on a en particulier

$$\sum_{s=1}^{n} s k_s = n\,.$$

Pour s donné, on a k_s cycles de longueur s de σ_P à mettre en correspondance avec les k_s cycles de σ_Q de même longueur, soit $k_s!$ possibilités; pour un tel cycle de σ_P, le premier élément du cycle peut être envoyé sur n'importe lequel des s éléments du cycle de σ_Q correspondant, soit s possibilités, et ceci pour chacun des k_s cycles de longueur s, soit s^{k_s} possibilités. On dénombre ainsi $\prod_{s=1}^{n} k_s! s^{k_s}$ solutions à l'équation (9).

4.2.2 Un exemple de taille 6

Soit deux permutations σ_2 et σ_3 sur l'ensemble $\{1,\ldots,6\}$

$$\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 2 & 4 & 1 & 6 \end{pmatrix} = (1325)(4)(6)\,,$$

$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 5 & 6 & 4 & 1 \end{pmatrix} = (13546)(2)\,.$$

On recherche l'ensemble des couples (σ_1,σ_4) d'éléments de $\mathcal{S}_6$ tels que $\sigma_1 : \sigma_2 :: \sigma_3 : \sigma_4$. Pour cela, on pose $\sigma_P = \sigma_3 \circ \sigma_2^{-1}$ et $\sigma_Q = \sigma_2^{-1} \circ \sigma_3$ et on a

$$\sigma_P = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 3 & 6 & 2 & 1 \end{pmatrix} = (146)(25)(3)\,,$$

$$\sigma_Q = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 2 & 5 & 6 & 4 \end{pmatrix} = (1)(23)(456)\,.$$

Ces deux dernières permutations sont donc décomposées de la même manière en cycles à supports disjoints : un cycle de longueur 1, un cycle de longueur 2 et un cycle de longueur 3. Par conséquent, il y a six solutions σ_1 de l'équation (9) qui envoient le cycle (1) vers (3), (23) vers (25) et (456) vers (146). Ces six permutations sont décrites dans le tableau ci-dessous, chaque ligne correspondant à une solution.

i	1	2	3	4	5	6
$\sigma_1(i)$	3	2	5	1	4	6
	3	2	5	6	1	4
	3	2	5	4	6	1
	3	5	2	1	4	6
	3	5	2	6	1	4
	3	5	2	4	6	1

À l'aide de (8), on obtient alors les six couples (σ_1,σ_4) tels que $\sigma_1 : \sigma_2 :: \sigma_3 : \sigma_4$. Ces six proportions analogiques sont

$$(1354)(2)(6) : (1325)(4)(6) :: (13546)(2) : (13256)(4)$$

$$(135)(2)(46) : (1325)(4)(6) :: (13546)(2) : (13254)(6)$$

$$(1356)(2)(4) : (1325)(4)(6) :: (13546)(2) : (1325)(46)$$

$$(13254)(6) : (1325)(4)(6) :: (13546)(2) : (1356)(2)(4)$$

$$(1325)(46) : (1325)(4)(6) :: (13546)(2) : (1354)(2)(6)$$

$$(13256)(4) : (1325)(4)(6) :: (13546)(2) : (135)(2)(46)$$

et on peut également les représenter de manière plus visuelle, en séparant les groupes de colonnes en proportion analogique

i	1	2	3	4	5	6
$\sigma_1(i)$	3	2	5	1	4	6
$\sigma_2(i)$	3	5	2	4	1	6
$\sigma_3(i)$	3	2	5	6	4	1
$\sigma_4(i)$	3	5	2	4	6	1

i	1	2	3	4	5	6
$\sigma_1(i)$	3	2	5	6	1	4
$\sigma_2(i)$	3	5	2	4	1	6
$\sigma_3(i)$	3	2	5	6	4	1
$\sigma_4(i)$	3	5	2	1	4	6

i	1	2	3	4	5	6
$\sigma_1(i)$	3	2	5	4	6	1
$\sigma_2(i)$	3	5	2	4	1	6
$\sigma_3(i)$	3	2	5	6	4	1
$\sigma_4(i)$	3	5	2	6	1	4

i	1	2	3	4	5	6
$\sigma_1(i)$	3	5	2	1	4	6
$\sigma_2(i)$	3	5	2	4	1	6
$\sigma_3(i)$	3	2	5	6	4	1
$\sigma_4(i)$	3	2	5	4	6	1

i	1	2	3	4	5	6
$\sigma_1(i)$	3	5	2	6	1	4
$\sigma_2(i)$	3	5	2	4	1	6
$\sigma_3(i)$	3	2	5	6	4	1
$\sigma_4(i)$	3	2	5	1	4	6

i	1	2	3	4	5	6
$\sigma_1(i)$	3	5	2	4	6	1
$\sigma_2(i)$	3	5	2	4	1	6
$\sigma_3(i)$	3	2	5	6	4	1
$\sigma_4(i)$	3	2	5	6	1	4

4.3 Dissemblance analogique

Afin de définir la dissemblance analogique associée à un quadruplet de permutations, on définit dans le paragraphe suivante une distance entre permutations.

4.3.1 Distance entre permutations

Le résultat suivant est classique.

Proposition 9. *Soit $\sigma \in \mathcal{S}_n$, on note $\mathcal{Z}(\sigma)$ le nombre de cycles à support disjoint qui composent σ. On note également $\mathcal{T}(\sigma)$ le nombre minimum de transpositions dont la composition est σ. Par convention, si σ est l'identité alors $\mathcal{T}(\sigma) = 0$. On a*

$$\mathcal{Z}(\sigma) + \mathcal{T}(\sigma) \;=\; n\,.$$

Proposition 10. *Soit σ_1 et σ_2 deux élements de $\mathcal{S}_n$, on définit la quantité*

$$\begin{aligned}\Delta(\sigma_1,\sigma_2) &= n - \mathcal{Z}\left(\sigma_1^{-1}\circ\sigma_2\right)\\ &= \mathcal{T}\left(\sigma_1^{-1}\circ\sigma_2\right)\,.\end{aligned}$$

L'application Δ est une distance entre deux permutations.

Démonstration :

- D'après la proposition 9, Δ est une fonction à valeurs positives.
- Soit σ_1 et σ_2 deux éléments de $\mathcal{S}_n$ telles que $\Delta(\sigma_1,\sigma_2) = 0$, c'est-à-dire $\mathcal{Z}\left(\sigma_1^{-1}\circ\sigma_2\right) = n$. Comme $(\mathcal{S}_n,\circ)$ est un groupe, $\sigma_1^{-1}\circ\sigma_2$ est une permutation sur un ensemble à n éléments, composée de n cycles à support disjoint. Ces derniers ont donc leurs supports de longueur 1. Par conséquent, $\sigma_1^{-1}\circ\sigma_2$ est la permutation identité et on en déduit que $\sigma_1 = \sigma_2$.
- Δ est symétrique car pour tout $\sigma \in \mathcal{S}_n$, on a $\mathcal{Z}(\sigma^{-1}) = \mathcal{Z}(\sigma)$.
- Soit σ_1, σ_2 et σ_3 trois éléments de $\mathcal{S}_n$, on a

$$\begin{aligned}\Delta(\sigma_1,\sigma_2) &= \mathcal{T}\left(\sigma_1^{-1}\circ\sigma_2\right)\\ &\leq \mathcal{T}\left(\sigma_1^{-1}\circ\sigma_3\right) + \mathcal{T}\left(\sigma_3^{-1}\circ\sigma_2\right)\\ &= \Delta(\sigma_1,\sigma_3) + \Delta(\sigma_3,\sigma_2)\end{aligned}$$

car $\mathcal{T}\left(\sigma_1^{-1} \circ \sigma_2\right)$ est le nombre minimun de transpositions dont la composition est $\sigma_1^{-1} \circ \sigma_2$.

■

En particulier, $\mathcal{T}(\sigma) = n - \mathcal{Z}(\sigma)$ représente la distance entre une permutation σ et la permutation identité. Par ailleurs, Δ offre des propriétés d'invariance pour l'opération de composition décrites ci-dessous.

Proposition 11. *La distance Δ est invariante pour l'inversion, par composition à gauche et à droite. Autrement dit, quels que soient σ_1, σ_2 et σ_3 trois éléments de $\mathcal{S}_n$, on a*

$$\begin{aligned} \Delta(\sigma_1,\sigma_2) &= \Delta\left(\sigma_1^{-1},\sigma_2^{-1}\right) \\ &= \Delta(\sigma_3 \circ \sigma_1,\sigma_3 \circ \sigma_2) \\ &= \Delta(\sigma_1 \circ \sigma_3,\sigma_2 \circ \sigma_3)\,. \end{aligned}$$

Démonstration : soit trois éléments σ_1, σ_2 et σ_3 de $\mathcal{S}_n$, les permutations $\sigma_1^{-1} \circ \sigma_2$ et $\sigma_2 \circ \sigma_1^{-1}$ sont conjuguées et leurs décompositions font donc intervenir le même nombre de cycles à support disjoint. Ainsi,

$$\mathcal{Z}\left(\sigma_1^{-1} \circ \sigma_2\right) = \mathcal{Z}\left(\sigma_2 \circ \sigma_1^{-1}\right)$$

et on en déduit que

$$\Delta(\sigma_1,\sigma_2) = \Delta\left(\sigma_1^{-1},\sigma_2^{-1}\right)\,.$$

En outre,

$$\begin{aligned} \Delta(\sigma_1,\sigma_2) &= \mathcal{T}\left(\sigma_1^{-1} \circ \sigma_2\right) \\ &= \mathcal{T}\left(\sigma_1^{-1} \circ \sigma_3^{-1} \circ \sigma_3 \circ \sigma_2\right) \\ &= \mathcal{T}\left((\sigma_3 \circ \sigma_1)^{-1} \circ (\sigma_3 \circ \sigma_2)\right) \\ &= \Delta(\sigma_3 \circ \sigma_1,\sigma_3 \circ \sigma_2)\,. \end{aligned}$$

De même, $\sigma_3^{-1} \circ \left(\sigma_1^{-1} \circ \sigma_2\right) \circ \sigma_3$ est une permutation conjuguée de $\sigma_1^{-1} \circ \sigma_2$. Par conséquent,

$$\mathcal{Z}\left(\sigma_3^{-1} \circ \left(\sigma_1^{-1} \circ \sigma_2\right) \circ \sigma_3\right) = \mathcal{Z}\left(\sigma_1^{-1} \circ \sigma_2\right)$$

et l'on a

$$\Delta(\sigma_1 \circ \sigma_3,\sigma_2 \circ \sigma_3) = \Delta(\sigma_1,\sigma_2)\,.$$

■

On peut noter qu'en écrivant $\Delta(\sigma_1,\sigma_2) = n - \mathcal{Z}(\sigma_1^{-1} \circ \sigma_2)$, on démontre facilement qu'il existe un algorithme en complexité temporelle $\mathcal{O}(n)$ pour calculer $\Delta(\sigma_1,\sigma_2)$. Il suffit de calculer $\sigma_1^{-1} \circ \sigma_2$ (en temps linéaire en n), puis de trouver le nombre de cycles de $\sigma_1^{-1} \circ \sigma_2$ (également en temps linéaire en n).

4.3.2 La dissemblance analogique entre quatre permutations

On peut à présent introduire la notion de dissemblance analogique entre quatre permutations.

Définition 8. *Soit σ_1, σ_2, σ_3 et σ_4 quatre éléments de $\mathcal{S}_n$, la dissemblance analogique associée à $(\sigma_1,\sigma_2,\sigma_3,\sigma_4)$ est donnée par la quantité*

$$\begin{aligned} DA(\sigma_1,\sigma_2,\sigma_3,\sigma_4) &= \Delta(\sigma_1^{-1}\circ\sigma_3,\sigma_2^{-1}\circ\sigma_4)+\Delta(\sigma_1^{-1}\circ\sigma_2,\sigma_3^{-1}\circ\sigma_4) \\ &+ \Delta(\sigma_3^{-1}\circ\sigma_1,\sigma_4^{-1}\circ\sigma_2)+\Delta(\sigma_2^{-1}\circ\sigma_1,\sigma_4^{-1}\circ\sigma_3)\,. \end{aligned}$$

La dissemblance analogique entre quatre permutations se calcule également en temps linéaire en n. Elle est définie conformément aux principes donnés au paragraphe 3.4.1 et possède donc les propriétés énoncées dans la proposition 7.

Proposition 12. *Pour tout six-uplet $(\sigma_1,\ldots,\sigma_6)$ d'éléments de $\mathcal{S}_n$, on a*

$$DA(\sigma_1,\sigma_2,\sigma_5,\sigma_6) \leq DA(\sigma_1,\sigma_2,\sigma_3,\sigma_4)+DA(\sigma_3,\sigma_4,\sigma_5,\sigma_6)\,.$$

Démonstration : le résultat découle directement des propositions 8 et 11.

5 Analogie entre matrices inversibles

Les matrices à valeurs complexes, inversibles et de rang n forment un groupe $\mathcal{GL}_n(\mathbb{C})$ pour la multiplication matricielle, notée ici par simple concaténation. L'élément neutre est la matrice unité I.

5.1 Définition

La propriété analogique $P2$ donne la définition suivante de la proportion analogique entre matrices inversibles.

Définition 9 (Proportion analogique entre matrices)**.** *Quatre matrices inversibles A, B, C et D sont en proportion analogique $A : B :: C : D$ si les deux conditions suivantes sont satisfaites*

$$CA^{-1}B = BA^{-1}C\,, \qquad (10)$$

$$D = CA^{-1}B\,. \qquad (11)$$

5.2 Recherche des proportions analogiques

D'après le paragraphe 3.3, pour un couple de matrices donné, la recherche de toutes les proportions analogiques associées nécessite la résolution dans $\mathcal{GL}_n(\mathbb{C})$ d'une équation de la forme

$$PX = XQ \tag{12}$$

où P et Q sont des matrices inversibles et semblables (i.e. conjuguées) à valeurs complexes. Dans ce paragraphe, on présente une méthode de résolution de cette équation, extraite de [14].

5.2.1 Résolution

On rappelle que deux matrices P et Q semblables ont, dans $\mathcal{M}_n(\mathbb{C})$, la même écriture en blocs de Jordan à l'ordre des blocs près. En particulier, comme l'on considère les matrices P et Q inversibles, elles ont les mêmes valeurs propres, toutes non nulles, de même indice et de même ordre de multiplicité. On désigne cette réduction par la matrice J et les matrices de changements de bases U et $V \in \mathcal{GL}_n(\mathbb{C})$ telles que

$$\begin{aligned} P &= UJU^{-1} \\ Q &= VJV^{-1}\,. \end{aligned}$$

La matrice J est de la forme diagonale par blocs $J = diag\,(J_{\lambda_1}, \ldots, J_{\lambda_r})$ où $\lambda_1, \ldots, \lambda_r$ sont les valeurs propres de P, non nécessairement distinctes, et chaque bloc de Jordan J_{λ_i} est une matrice carrée de taille $p_i \times p_i$ de la forme

$$J_{\lambda_i} = \begin{pmatrix} \lambda_i & 1 & 0 & & \ldots & 0 \\ 0 & \lambda_i & 1 & 0 & & \\ & & \ddots & \ddots & \ddots & \\ \vdots & & & \lambda_i & 1 & 0 \\ & & & & \lambda_i & 1 \\ 0 & & & & 0 & \lambda_i \end{pmatrix}$$

On remarque que la recherche des solutions $X \in \mathcal{GL}_n(\mathbb{C})$ de l'équation (12) est équivalente à la résolution, pour $\Delta \in \mathcal{GL}_n(\mathbb{C})$, de

$$J\Delta = \Delta J \tag{13}$$

à l'aide du changement de variables $X = U\Delta V^{-1}$. Les solutions de (12) et de (13) sont décrites dans la proposition ci-dessous.

Proposition 13. *Les solutions de l'équation $PX = XQ$ dans $\mathcal{GL}_n(\mathbb{C})$ s'écrivent sous la forme $U\Delta V^{-1}$ avec $\Delta \in \mathcal{GL}_n(\mathbb{C})$ se décomposant par blocs*

$$\Delta = \begin{pmatrix} \Delta_{11} & \dots & \Delta_{1r} \\ \vdots & \ddots & \vdots \\ \Delta_{r1} & \dots & \Delta_{rr} \end{pmatrix}$$

où chaque matrice extraite $\Delta_{ij} \in \mathcal{M}_{p_i p_j}(\mathbb{C})$ est telle que

- *si $\lambda_i \neq \lambda_j$ alors Δ_{ij} est la matrice nulle,*
- *sinon Δ_{ij} est triangulaire supérieure de la forme*

$$\Delta_{ij} = \begin{pmatrix} 0 & \dots & 0 & a_1^{ij} & a_2^{ij} & \dots & a_{m_{ij}}^{ij} \\ & & & 0 & a_1^{ij} & \ddots & \vdots \\ \vdots & & \vdots & \vdots & \ddots & \ddots & a_2^{ij} \\ 0 & \dots & 0 & 0 & \dots & 0 & a_1^{ij} \\ 0 & \dots & 0 & 0 & \dots & & 0 \\ \vdots & & \vdots & \vdots & & & \vdots \\ 0 & \dots & 0 & 0 & \dots & & 0 \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m_{ij} \textit{ lignes} \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} p_i - m_{ij} \textit{ lignes} \end{array}$$

$$\underbrace{\hphantom{0 \dots 0}}_{p_j - m_{ij}\ \textit{col.}} \underbrace{\hphantom{a_1^{ij}\ a_2^{ij}\ \dots\ a_{m_{ij}}^{ij}}}_{m_{ij}\ \textit{col.}}$$

où $m_{ij} = \min\{p_i,p_j\}$ et $a_1^{ij},\dots,a_{m_{ij}}^{ij}$ complexes. En particulier, si $i = j$, $a_1^{ii} \neq 0$.

Démonstration : voir [14] pages 220-221.

5.2.2 Un exemple

On considère deux matrices A et B inversibles, telles que la réduction en blocs de Jordan de $P = BA^{-1}$, à l'aide de la matrice de changement de bases U, s'écrit

$$J = \left(\begin{array}{c|c|ccc|cc} \lambda & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & \lambda & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \beta & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \beta \end{array}\right)$$

où $(\lambda,\beta) \in (\mathbb{C}^*)^2$. On distingue sur la diagonale quatre blocs de Jordan délimités par les traits en pointillés. Les matrices inversibles X commutant avec P sont les matrices $U\Delta U^{-1}$, pour toute matrice inversible Δ commutant avec J. D'après la proposition 13, ces matrices Δ sont décrites par

$$\Delta = \left(\begin{array}{c:c:ccc:cc} a & b & 0 & 0 & c & 0 & 0 \\ \hdashline d & e & 0 & 0 & f & 0 & 0 \\ \hdashline g & h & i & j & k & 0 & 0 \\ 0 & 0 & 0 & i & j & 0 & 0 \\ 0 & 0 & 0 & 0 & i & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & l & m \\ 0 & 0 & 0 & 0 & 0 & 0 & l \end{array}\right)$$

où les nombres complexes $a,\ldots,m$ vérifient $i \neq 0$, $l \neq 0$ et $ae - bd \neq 0$, afin de garantir l'inversibilité des solutions Δ et X.

D'après la proposition 6, les proportions analogiques $A : B :: C : D$ définies à partir de A et B sont donc de la forme

$$A : B :: U\Delta U^{-1} : U\Delta J U^{-1} \ .$$

Par ailleurs, on peut remarquer, d'après la proposition 4, que ces proportions analogiques sont équivalentes à

$$I : J :: \Delta : \Delta J \ .$$

5.2.3 Un second exemple

Soit deux éléments A et B de $\mathcal{GL}_n(\mathbb{C})$ telles que la matrice $P = BA^{-1}$ admet n valeurs propres distinctes. La matrice de Jordan J associée à P est une matrice diagonale composée des n valeurs propres de P, toutes non nulles. Les matrices inversibles Δ qui commutent avec J sont les matrices Δ diagonales dont les éléments diagonaux sont également non nuls. Par conséquent, d'après les propositions 6 et 13, les proportions analogiques $A : B :: C : D$ sont de la forme

$$A : B :: EA : EB$$

où E est une matrice inversible et diagonalisable dans la même base de vecteurs propres que P. On peut également dire de manière équivalente que E est un polynôme à coefficients complexes en P, i.e. de la forme

$$E = a_m P^m + a_{m-1} P^{m-1} + \ldots + a_1 P + a_0 I \, .$$

Comme cas particulier de cet exemple, on considère un quadruplet (A,B,C,D) de matrices inversibles et diagonalisables dans une même base de vecteurs propres. Il existe une matrice $W \in \mathcal{GL}_n(\mathbb{C})$ telle que

$$A = WJ_AW^{-1} \qquad B = WJ_BW^{-1} \qquad C = WJ_CW^{-1} \qquad D = WJ_DW^{-1}$$

où J_A, J_B, J_C et J_D sont des matrices diagonales. On a alors les équivalences

$$\begin{aligned} A : B :: C : D \quad &\Leftrightarrow \quad J_A : J_B :: J_C : J_D \\ &\Leftrightarrow \quad J_A J_B^{-1} = J_C J_D^{-1}. \end{aligned}$$

Ainsi, la proportion analogique entre matrices inversibles et diagonalisables dans une même base de vecteurs est équivalente à la proportion analogique multiplicative entre les éléments diagonaux correspondants de même position, i.e. entre les valeurs propres correspondantes à un même vecteur propre.

5.2.4 Remarque

Soit A et B deux éléments de $\mathcal{GL}_n(\mathbb{C})$, on peut remarquer que l'équation d'inconnue $C \in \mathcal{GL}_n(\mathbb{C})$

$$CA^{-1}B \,=\, BA^{-1}C$$

admet comme solutions particulières $C = A$ et $C = B$. On cherche à résoudre cette équation sur un sous-ensemble de l'espace vectoriel $\mathcal{M}_n(\mathbb{C})$ afin de déterminer les proportions analogiques associées à A et B. Dans ce dernier, cette équation est linéaire et on en déduit que l'ensemble des combinaisons linéaires de A et B sont également solutions. Par conséquent, toutes les matrices inversibles de la forme

$$\alpha\,(B - \lambda A)$$

sont solutions dans $\mathcal{GL}_n(\mathbb{C})$; ces matrices sont inversibles si et seulement si α est un nombre complexe non nul et λ n'est pas une valeur propre de $P = BA^{-1}$.

5.3 Dissemblance analogique

On peut, cela a été présenté dans le paragraphe 3.4, définir la dissemblance analogique d'un quadruplet (A,B,C,D) de matrices inversibles à partir d'une distance Δ entre matrices par

$$\begin{aligned} DA(A,B,C,D) &= \Delta(B^{-1}D,A^{-1}C) + \Delta(C^{-1}D,A^{-1}B) \\ &+ \Delta(C^{-1}A,D^{-1}B) + \Delta(B^{-1}A,D^{-1}C)\,. \end{aligned}$$

La dissemblance analogique vérifie alors les propriétés énoncées dans la proposition 7. Cependant, l'invariance de Δ pour la multiplication à gauche et à droite n'est généralement pas satisfaite.

5.3.1 Exemple de la distance issue de la norme $\|.\|_2$

On considère la distance Δ entre matrices issue de la norme subordonnée à la norme $\|.\|_2$: soit $A \in \mathcal{GL}_n(\mathbb{C})$ et $y \in \mathbb{C}^n$, on rappelle que

$$\begin{aligned} \|y\|_2 &= y^* y \\ &= \sum_{i=1}^{n} |y_i|^2\ , \end{aligned}$$

où $y^* = \overline{y}^T$ et

$$\begin{aligned}\|A\|_2 &= \sup_{\substack{x\in\mathbb{C}^n\\ \|x\|_2=1}} \|Ax\|_2\\ &= \sqrt{\rho(A^*A)}\,.\end{aligned}$$

$A^* = \overline{A}^T$ désigne la transposée de la matrice conjuguée de A (pour la conjugaison des coefficients complexes qui la composent) et $\rho(A^*A)$ correspond à la plus grande grande valeur propre de A^*A.

On définit la distance Δ entre deux matrices A et $B \in \mathcal{GL}_n(\mathbb{C})$ par

$$\Delta(A,B) = \|A-B\|_2\,.$$

Cette distance est invariante pour la multiplication matricielle à gauche et à droite par une matrice unitaire [2]. Autrement dit, pour toutes matrices A, B et $C \in \mathcal{GL}_n(\mathbb{C})$, avec C matrice unitaire, on a

$$\Delta(CA,CB) = \Delta(AC,BC) = \Delta(A,B)\,. \tag{14}$$

5.3.2 Exemple des isométries du plan

On note $\mathcal{O}_2(\mathbb{R})$ l'ensemble des matrices orthogonales de rang 2 à coefficients réels. L'ensemble $\mathcal{O}_2(\mathbb{R})$ muni de la multiplication matricielle est un sous-groupe de $\mathcal{GL}_2(\mathbb{R})$. Il est appelé le groupe des isométries du plan et est composé de deux types de matrices :

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}$$

et

$$S_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta)\\ \sin(\theta) & -\cos(\theta)\end{pmatrix}$$

pour tout réel θ. La matrice R_θ représente la rotation de centre (0,0) et d'angle θ et S_θ représente la symétrie orthogonale par rapport à la droite vectorielle de vecteur directeur $(\cos(\theta/2), \sin(\theta/2))$, dans la base canonique de $\mathbb{R}^2$.

Afin d'établir les proportions analogiques entre des éléments de $\mathcal{O}_2(\mathbb{R})$, on énumère les différents produits possibles entre deux éléments de $\mathcal{O}_2(\mathbb{R})$: soit deux nombres réels β et θ, on a

$$R_\beta R_\theta = R_{\beta+\theta}, \qquad R_\beta S_\theta = S_{\beta+\theta}, \qquad S_\beta R_\theta = S_{\beta-\theta}, \qquad S_\beta S_\theta = R_{\beta-\theta}\,. \tag{15}$$

La recherche des proportions analogiques entre éléments de $\mathcal{O}_2(\mathbb{R})$ est par conséquent aisée et ne nécessite pas de réduction matricielle en blocs de Jordan.

2. On rappelle qu'une matrice $C \in \mathcal{M}_n(\mathbb{C})$ (respectivement $\mathcal{M}_n(\mathbb{R})$) est dite unitaire (respectivement orthogonale) si $CC^* = C^*C = I$.

Proposition 14. *Les proportions analogiques entre éléments de $\mathcal{O}_2(\mathbb{R})$ sont représentées ci-dessous, chaque proportion analogique pouvant se décliner sous des formes équivalentes décrites dans le paragraphe 2.1. Pour tous réels α, β et θ et tout entier relatif k, on a*

1. $R_\alpha : R_\beta :: R_\theta : R_{\theta+\beta-\alpha}$
2. $R_\alpha : S_\beta :: S_{\beta+k\pi} : R_{\alpha+k\pi}$
3. $S_\alpha : R_\beta :: S_{\alpha+k\pi} : R_{\beta+k\pi}$
4. $S_\alpha : S_\beta :: S_{\alpha+\theta} : S_{\beta+\theta}$

Démonstration : on considère le premier type de proportions analogiques énoncé dans la proposition, l'établissement des autres cas s'effectuant de manière similaire. Soit α et β deux réels, on pose $A = R_\alpha$ et $B = R_\beta$ et l'on cherche l'ensemble des couples (C,D) d'éléments de $\mathcal{O}_2(\mathbb{R})$ tels que l'on ait $A : B :: C : D$. Pour cela, on pose $P = BA^{-1}$, i.e. $P = R_{\beta-\alpha}$ et on doit déterminer, d'après le paragraphe 3.3, les éléments $X \in \mathcal{O}_2(\mathbb{R})$ qui commutent avec P :

– si $X = R_\theta$, on a d'après (15)

$$XP = PX = R_{\beta-\alpha+\theta}$$

quelle que soit la valeur de θ. Dans ce cas, on obtient les proportions analogiques

$$\forall\theta \in \mathbb{R}, \quad R_\alpha : R_\beta :: R_\theta R_\alpha : R_\theta R_\beta$$

c'est-à-dire

$$\forall\theta \in \mathbb{R}, \quad R_\alpha : R_\beta :: R_\theta : R_{\theta+\beta-\alpha}$$

ce qui correspond aux proportions analogiques de type 1 décrites dans la proposition.

– si $X = S_\theta$, on a

$$\begin{aligned} XP &= S_{\theta-(\beta-\alpha)}\,, \\ PX &= S_{\theta+(\beta-\alpha)}\,. \end{aligned}$$

Ainsi, X commute avec P si et seulement si

$$\theta - (\beta - \alpha) = \theta + (\beta - \alpha) \text{ modulo } 2\pi$$

i.e. $\beta = \alpha$ modulo π. Dans ce cas, il existe $k \in \mathbb{Z}$ tel que $\beta = \alpha + k\pi$ et l'on obtient les proportions analogiques

$$\forall\theta \in \mathbb{R}, \quad R_\alpha : R_{\alpha+k\pi} :: S_\theta R_\alpha : S_\theta R_{\alpha+k\pi}$$

soit encore

$$\forall\theta \in \mathbb{R}, \quad R_\alpha : R_{\alpha+k\pi} :: S_{\theta-\alpha} : S_{\theta-\alpha-k\pi}$$

i.e.

$$\forall\theta \in \mathbb{R}, \quad R_\alpha : R_{\alpha+k\pi} :: S_\theta : S_{\theta+k\pi}$$

car $S_{\theta-k\pi} = S_{\theta+k\pi}$. D'après la définition 1, ces proportions analogiques correspondent aux proportions de type 3.

■

Par ailleurs, on a, pour tout réel θ,

$$\begin{aligned}\Delta(I,R_\theta) &= 2\left|\sin\left(\frac{\theta}{2}\right)\right|\,,\\ \Delta(I,S_\theta) &= 2\,.\end{aligned}$$

Ainsi, à l'aide des relations (14) et (15), la distance entre deux matrices orthogonales se calcule simplement :

$$\begin{aligned}\Delta(R_\beta,R_\theta) &= \|R_\beta - R_\theta\|_2\\ &= \left\|I - R_\theta\,(R_\beta)^{-1}\right\|_2\\ &= \|I - R_\theta R_{-\beta}\|_2\\ &= \|I - R_{\theta-\beta}\|_2\\ &= 2\left|\sin\left(\frac{\theta-\beta}{2}\right)\right|\,,\end{aligned}$$

$$\begin{aligned}\Delta(S_\beta,R_\theta) &= \|S_\beta - R_\theta\|_2\\ &= \left\|I - R_\theta\,(S_\beta)^{-1}\right\|_2\\ &= \|I - R_\theta S_\beta\|_2\\ &= \|I - S_{\theta+\beta}\|_2\\ &= 2\,,\end{aligned}$$

$$\begin{aligned}\Delta(S_\beta,S_\theta) &= \|S_\beta - S_\theta\|_2\\ &= \left\|I - S_\theta\,(S_\beta)^{-1}\right\|_2\\ &= \|I - S_\theta S_\beta\|_2\\ &= \|I - R_{\theta-\beta}\|_2\\ &= 2\left|\sin\left(\frac{\theta-\beta}{2}\right)\right|\,.\end{aligned}$$

Dans cet exemple, la dissemblance analogique entre quatre matrices se calcule aisément et vérifie, d'après la proposition 8, l'« inégalité triangulaire ».

6 Conclusion : quelques pistes pour l'utilisation en apprentissage

Les études sur la proportion analogique peuvent sembler encore loin de produire des algorithmes compétitifs, il n'empêche qu'elles ont déjà mené à des résultats prometteurs, en particulier sur des données structurées. Ce rapport a pour but essentiel de formaliser cette notion dans deux types de données très particuliers : les permutations et les matrices. Dans le premier cas, les applications sont sans doute à chercher dans l'ordonnancement ou la comparaison analogique de préférences, mais les auteurs n'ont pas cherché à approfondir ce point.

En ce qui concerne les matrices, les applications sont plus familières. Prenons en effet un ensemble d'apprentissage composé de vecteurs de $\mathbb{R}^d$, supervisés par C classes. Chaque classe ω_i peut séparément être modélisée comme un mélange de sous-classes elles-mêmes représentées par des distributions gaussiennes :

$$p(x|\omega_i) = \sum_{j=1}^{K_i} \lambda_{ij} \mathcal{N}(\mu_{ij}, Qij) \, .$$

L'apprentissage des paramètres μ_{ij}, Qij et λ_{ij} se fait classiquement par l'algorithme EM. La structure analogique de cet ensemble d'apprentissage peut alors être explorée en cherchant les quadruplets de distributions dont les centres et les matrices de covariance sont à faible dissemblance analogique.

Ensuite, une technique d'apprentissage paresseux, comme celle décrite dans [7] peut être employée. On cherche l'ensemble des triplets dont la dissemblance analogique avec l'objet à classer est inférieure à un seuil. La classe de cet objet peut alors être déterminée à partir des triplets de sous-classes et des relations d'analogie entre ces sous-classes.

En quoi une telle méthode peut-elle rivaliser avec une simple décision bayésienne ? Les expériences menées tendent à montrer que des résultats compétitifs peuvent être obtenus sur des petites bases de données d'apprentissage, à condition de tenir compte de la structure analogique des classes. En quelque sorte, la quantité brute est remplacée par une information relationnelle sur la géométrie relative des classes (ou des sous-classes).

Cela reste à valider expérimentalement. Nous étudions en ce moment des bases de données de chiffres manuscrits, représentées par des vecteurs. Nous cherchons en particulier si la relation d'analogie entre matrices définie dans cet article y possède la sémantique attendue, c'est-à-dire si, par exemple, sont (ou ne sont pas) en proportion analogique les quatre nuages (modélisés par des distributions gaussiennes) des vecteurs correspondant à :

- la lettre « a » écrite par le scripteur « A »
- la lettre « b » écrite par le scripteur « A »
- la lettre « a » écrite par le scripteur « B »
- la lettre « b » écrite par le scripteur « B »

Si tel est le cas en général, les calculs algébriques de l'analogie sur les matrices seront un modèle utile pour explorer l'information relationnelle analogique des classes.

Références

[1] Y. Lepage, *De l'analogie rendant compte de la commutation en linguistique*, Habilitation à Diriger les Recherches, Université Joseph Fourier, Grenoble I, 2003.

[2] N. Stroppa and F. Yvon, *Formal models of analogical proportions*, Rapport de Recherche 2006-D008, École Nationale Supérieure des Télécommunications, 2006.

[3] K. Holyoak, *Analogy*, The Cambridge handbook of thinking and reasoning, chapter 6, Cambridge University Press, 2005.

[4] T. Evans, *A heuristic program to solve geometry analogy problems*, Semantic Information Processing, MIT Press, Cambridge, 2008.

[5] A. Aamodt and E. Plaza, *Case-based reasoning: foundational issues, methodological variations, and system approaches*, Artificial Intelligence Communications, Volume 7, number 1, pages 39–59, 1994.

[6] N. Stroppa and F. Yvon, *Analogical learning and formal proportions: definitions and methodological issues*, Rapport de Recherche 2005-D004, École Nationale Supérieure des Télécommunications, 2005.

[7] L. Miclet and S. Bayoudh and A. Delhay, *Analogical dissimilarity: definition, algorithms and two experiments in machine learning*, Journal of Artificial Intelligence Research, Volume 32, pages 793–824, 2008.

[8] A. Ben Hassena, *Proportion analogique entre arbres : une étude bibliographique*, Rapport de Recherche IRISA, numéro 1912, 2008.

[9] L. Miclet and H. Prade, *Logical definition of analogical proportion and its fuzzy extensions.*, NAFIPS Conference, New York, 2008

[10] F. Yvon, *Paradigmatic cascades: a linguistically sound model of pronunciation by analogy*, Proceedings of the 35th annual meeting of the Association for Computational Linguistics (ACL), pages 248–435, Madrid, Spain, 1997.

[11] S. Bayoudh and H. Mouchère and L. Miclet and E. Anquetil, *Learning a classifier with very few examples: analogy based and knowledge based generation of new examples for character recognition.*, Proceedings of the 18th European Conference on Machine Learning, Volume 18, Springer Verlag LNAI 4701, Warsaw, Poland, 2007.

[12] S. Bayoudh and L. Miclet and A. Delhay, *Learning by analogy : a classification rule for binary and nominal data*, Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI), Volume 20, pages 678–683, Hyderabad, India, 2007.

[13] `http://fr.wikipedia.org/wiki/Groupe_sym%C3%A9trique`, 2008.

[14] F. R. Gantmacher, *The theory of matrices*, AMS Chelsea Publishing, 2000.