Action de groupes sur des espaces de matrices-Leçon 150

"Oui, cela était autrefois ainsi ; mais nous avons changé tout cela, et nous faisons maintenant la médecine d'une méthode toute nouvelle."

Le médecin malgré lui, (1666), Molière.

0. Le 6 minutes

Les actions de groupes sur des espaces de matrices illustrent une méthode uniforme pour les problèmes de classification que l'on rencontre en mathématiques. En effet, un groupe, en agissant, partitionne en orbites l'ensemble sur lequel il agit, avec, dans le cas des espaces de matrices, une possibilité d'avoir des actions linéaires. La nature de la classification dépendra alors du choix du groupe agissant : groupe linéaire, pour des classifications linéaires ; groupe affine, pour des classifications affines ; le groupe $\mathrm{O}_n(\mathbb{R})$, pour des classifications euclidiennes ; et enfin, les classifications projectives, avec le groupe projectif. Chaque orbite se verra munie d'un classifiant (invariant total) et souvent d'une matrice de forme normale.

Dans les espaces de matrices, le problème de classification provient principalement du problème de changement de base. En effet, on se sert des matrices pour coder des objets (applications linéaires, endomorphismes, formes quadratiques, représentations), mais ce codage dépend de façon drastique d'une base ; il faut alors gérer le problème de changement de base.

Dans un premier temps, les problématiques sont les suivantes : décrire les actions, décrire les classifiants, trouver des algorithmes pour calculer les classifiants, déterminer les formes normales. Dans un deuxième temps, on va pouvoir, si le corps est $\mathbb{R}$ ou $\mathbb{C}$, mettre une topologie (normique) sur l'espace de matrices, puis, chercher à décrire les adhérences d'orbites. Si le corps est fini, on peut chercher les cardinaux de chaque orbite. Enfin, on peut, dans un troisième temps, s'intéresser à des problèmes de descente, c'est-à-dire comment passer de la classification sur un corps $\mathbb{L}$ à un sous-corps $\mathbb{K}$.

Le premier exemple édifiant que tout le monde peut comprendre est l'action de Steinitz. Une même application linéaire va être codée dans deux paires de bases différentes, $(\underline{e}, \underline{f})$ et $(\underline{e}', \underline{f}')$, et les matrices respectives vont vérifier

$A' = P^{-1}AQ$ où P, resp Q, désigne la matrice de passage de $\underline{f}$ à $\underline{f}'$, resp. $\underline{e}$ à $\underline{e}'$. Le classifiant est le rang, qui se calcule grâce au pivot de Gauss sur les lignes (à gauche) et sur les colonnes (à droite). La matrice de forme normale de rang r est la matrice avec r "1" sur sa diagonale, et des zéros ailleurs. Il n'y a ici pas d'obstruction de descente, puisque le rang est indépendant du corps de base (on rappelle que le rang est égal à la taille du plus grand mineur non nul), et donc deux matrices sur $\mathbb{K}$ sont équivalentes sur $\mathbb{L}$ si et seulement si elles le sont sur $\mathbb{K}$. De plus, si O_r est l'orbite des matrices de rang r sur $\mathbb{R}$ ou $\mathbb{C}$, alors son adhérence est donnée par la réunion des $O_{r'}$, où $0 \leq r' \leq r$. Le cardinal d'une orbite sur un corps fini est un exercice facile ; il utilise le cardinal du groupe linéaire et le cardinal d'un stabilisateur.

On peut traîter le cas de l'action par conjugaison de $\mathrm{GL}_n(\mathbb{C})$ sur les matrices diagonalisables sur $\mathbb{C}$, et même sur les matrices nilpotentes. Dans le second cas, on tombe, pour les formes normales, sur les réduites de Jordan. Le premier cas a sa petite spécificité : les orbites sont toutes fermées et cela constitue même une caractérisation des matrices diagonalisables. Dans tous les cas, on n'a pas d'obstruction de descente : deux matrices carrées sur $\mathbb{K}$ sont $\mathbb{L}$-semblables si et seulement si elles sont $\mathbb{K}$-semblables.

On peut ensuite traiter le cas de l'action de $\mathrm{GL}_n(\mathbb{K})$ par congruence sur l'espace $\mathcal{S}_n(\mathbb{K})$ des matrices symétriques. Ici, les choses dépendent de façon drastique du corps de base. On traite en général les cas où le corps de base est $\mathbb{C}$ (invariant=rang), $\mathbb{R}$ (invariant=signature, par le théorème de Sylvester), et $\mathbb{F}_q$ (invariant=discriminant). L'algorithme dominant est la fameuse méthode de Gauss (encore lui !).

Il y a aussi l'action à gauche de $\mathrm{GL}_n(\mathbb{K})$ sur l'espace $\mathcal{M}_{n,m}(\mathbb{K})$. En effet quand on veut résoudre le système linéaire $AX = Y$, on intervient par combinaisons linéaires sur les lignes et donc, uniquement à gauche sur $A \in \mathcal{M}_{n,m}(\mathbb{K})$. On effectue un algorithme de pivot, mais uniquement à gauche (le pivot à droite correspondrait à des changements de variables). L'action $P \cdot A = PA$ doit être comprise. Les formes normales sont alors les matrices échelonnées réduites : pour tout A, il existe une unique matrice échelonnée réduite E telle que $PA = E$ pour un P dans $\mathrm{GL}_n(\mathbb{K})$.

La théorie des représentations peut rentrer dans le cadre de la leçon, mais à un niveau assez élevé. Se donner une représentation complexe d'un groupe fini G d'ordre n revient à se donner n matrices $A_g \in \mathcal{M}_m(\mathbb{C})$, $g \in G$, qui vérifient les même relations que dans le groupe : $gh = k$ implique $A_gA_h = A_k$. On peut se demander s'il existe une matrice de passage P, si les PA_gP^{-1} sont réelles pour tout g. Un superbe résultat dans le cas d'une représentation irréductible est donné par l'indicatrice de Frobenius-Schur.

Pièges classiques :

1. Il faut éviter de faire un catalogue ennuyeux. Pour cela, il est bon d'insister sur les spécificités de chaque action.
2. Les changements de bases fournissent des actions naturellement à droite. Par exemple, deux matrices A et B sont semblables si $B = P^{-1}AP$, où P est la matrice de passage de la base pour A vers la base pour B. Mais il faut remplacer l'action $P \cdot A = P^{-1}AP$ par $P \cdot A = PAP^{-1}$, afin d'avoir une action à gauche (convention oblige).
3. Toutes ces matrices de passage sont intéressantes quand on veut résoudre une équation ou un système d'équations. La matrice P permet de passer de la matrice A à une matrice plus simple B, et la matrice P correspond à un changement de variables. Il faut toujours avoir en tête le principe de translation quand on passe du monde des matrice au monde de la géométrie. Si par exemple $B = P^{-1}AP$, alors $\ker B$ est l'image par P^{-1} de $\ker A$. On a des propriétés analogues pour des sous-espaces propres, les sous-espaces caractéristiques...
4. De même quand on a une forme quadratique q, dont on veut connaître le cône isotrope $\mathcal{C}_q$, c'est-à-dire, plus simplement, quand on veut résoudre l'équation $q(u) = 0$ (homogène de degré 2). Dans ce cas, on se ramène à la situation matricielle ${}^tXAX = 0$, puis ${}^t(PX)B(PX) = 0$, où B est congruente à A. On a alors avec des notations évidentes $X \in \mathcal{C}_A$ si et seulement si $PX \in \mathcal{C}_B$, ce qui donne $\mathcal{C}_A = P^{-1}\mathcal{C}_B$ qui est un avatar du principe de translation.

I. Les fondamentaux

1. L'action de Steinitz. Deux matrices sont équivalentes si et seulement si elles ont même rang. L'adhérence sur $\mathbb{R}$ ou $\mathbb{C}$ des matrices de rang r est l'ensemble des matrices de rang inférieur (ou égal) à r. C'est la semi-continuité inférieure du rang.
2. L'action par conjugaison de $\mathrm{GL}_n(\mathbb{K})$ sur les matrices de $\mathcal{M}_n(\mathbb{K})$. Il s'agit de ce que l'on appelle communément la réduction des matrices. Tous les outils que l'on utilise en réduction sont des invariants partiels. Par exemple le polynôme caractéristique : si A est B sont semblables, alors $\chi_A = \chi_B$. La leçon est un bonne occasion pour réviser l'ensemble des invariants partiels ; pêle-mêle : le déterminant, la trace, le rang, la multiplicité géométrique d'une valeur propre, le polynôme minimal, et le spectre (avec multiplicité et modulo permutation) !
3. Le noyau est un joli invariant total pour l'action à gauche : soit A, B dans $\mathcal{M}_{n,m}(\mathbb{K})$, alors il existe P dans $\mathrm{GL}_n(\mathbb{K})$ tel que $PA = B$ si et seulement si $\ker A = \ker B$.
4. Si A est une matrice symétrique de $\mathcal{S}_n(\mathbb{K})$ (en caractéristique différente de 2), alors il existe une matrice diagonale D telle que ${}^tPAP = D$ pour un P dans $\mathrm{GL}_n(\mathbb{K})$. On peut le voir par exemple avec la méthode de Gauss. Cependant, on ne peut pas s'arrêter là, car cette diagonale n'est pas un invariant, même à permutation près. Sur $\mathbb{C}$, on montre qu'un invariant total est le rang, et sur $\mathbb{R}$ il est donné par la signature ; c'est le théorème de Sylvester.
5. Sur $\mathbb{R}$, on peut affiner l'action de congruence en une action de $\mathrm{O}_n(\mathbb{R})$ par congruence sur les matrices symétriques. Or, si $P \in \mathrm{O}_n(\mathbb{R})$, ${}^tPSP = P^{-1}SP$, ce qui signifie que l'action de congruence et l'action de conjugaison coïncident. Il ne faut pas passer à côté du théorème spectral qui joue sur ces deux tableaux !
6. Deux matrices carrées sur $\mathbb{R}$ sont $\mathbb{C}$-semblables si et seulement si elles sont $\mathbb{R}$ semblables. Ceci est équivalent à dire que si $A \in \mathcal{M}_n(\mathbb{R}) \subset \mathcal{M}_n(\mathbb{C})$ et $\mathcal{O}_A(\mathbb{C})$, resp. $\mathcal{O}_A(\mathbb{R})$, est l'ensemble des matrices de $\mathcal{M}_n(\mathbb{C})$, resp. $\mathcal{M}_n(\mathbb{R})$, $\mathbb{C}$-semblables, resp. $\mathbb{R}$-semblable,s à A, alors $\mathcal{M}_n(\mathbb{R}) \cap \mathcal{O}_A(\mathbb{C}) = \mathcal{O}_A(\mathbb{R})$.

II. Pour aller plus loin

1. Il est bon de donner la classification des formes quadratiques sur $\mathbb{F}_q$, [1, Théorème V-1.4.1 (iii)].
2. Un groupe topologique transmet à ses orbites ses propriétés topologiques (connexité, compacité). On montre que les orbites de rang r sur

$\mathcal{M}_n(\mathbb{C})$ forment un connexe car $\mathrm{GL}_n(\mathbb{C})$ est connexe, [1, Proposition II-2.3.1]. En revanche, ce résultat est faux sur $\mathbb{R}$, il faut prendre $r < n$ et utiliser le fait que $\mathrm{GL}_n(\mathbb{R})^+$ est connexe, [1, Proposition II-4.2.5].

On peut partir gaiement vers les cardinaux des orbites sur $\mathbb{F}_q$, [1, I-3.6]. On calculera les cardinaux des ensembles suivants : matrices de rang r, matrices diagonalisables, matrices nilpotentes, formes quadratiques de rang r sur $\mathbb{F}_q$, [2, Annexe VIII-A].

3. Plus fort encore, l'action à gauche de $\mathrm{GL}_n(\mathbb{F}_q)$ sur l'espace des matrices fournit une jolie formule polynomiale pour les grassmanniennes, puis, la formule du triple produit de Jacobi, en passant par la formule du binôme quantique, [3, IV-1.3, IV-1.10].
4. Classification affine et euclidienne des coniques, [1, V-6].
5. On a parlé des fonctions de $\mathcal{M}_n(\mathbb{C})$ dans $\mathbb{C}$ qui sont invariantes, comme le déterminant ($\det(PAP^{-1}) = \det(A)$), la trace ($\mathrm{tr}(PAP^{-1}) = \mathrm{tr}(A)$). D'ailleurs, toutes les fonctions qui se retrouvent comme coefficient du polynôme caractéristique sont forcément invariantes puisque le polynôme caractéristique est un invariant de similitude. En fait, on peut montrer, à l'aide des relations coefficients-racines, que ces fonctions engendrent l'algèbre des fonctions polynomiales invariantes ; c'est un théorème de Harish-Chandra, [1, Exercice III-D.29].
6. Pour tout n soit J_n le bloc de Jordan de taille n. Toute matrice nilpotente est semblable à une unique matrice diagonale par blocs de Jordan $\mathrm{diag}(J_{\lambda_1}, \cdots, J_{\lambda_k})$ avec $\lambda_1 + \cdots + \lambda_k = n$ et $\lambda_1 \geq \cdots \geq \lambda_k$ (on dit que les λ_i forment une partition de n), [1, Théorème III-2.5.1]. L'étude de l'adhérence des classes de similitude met en évidence un ordre sur l'ensemble des partitions de n, [1, Théorème III-2.9.1], (malheureusement la preuve ne tient pas en quinze minutes !).
7. Partir sur la théorie des représentations des groupes finis. Si G est un groupe fini, une représentation complexe de G est un morphisme de G dans $\mathrm{GL}_n(\mathbb{C})$. Si on a deux représentations de G de degré n et m , de caractère respectif χ_n et χ_m, alors on a un morphisme de G vers $\mathrm{GL}_n \times \mathrm{GL}_m$. En le composant avec l'action de Steinitz, on obtient une nouvelle représentation de degré nm sur l'espace des matrices de caractère $\overline{\chi}_n \chi_m$, [2, Lemme XIII-2.3.1 (iv)]. Ceci permet le théorème fondamental d'orthogonalité des caractères, [2, Corollaire XIII-2.5.6, Théorème XIII-2.6.1].
8. Toujours dans la théorie des représentations, mais plus difficile cette fois, il y a l'indicatrice de Frobenius-Schur qui dit que si une représentation complexe de caractère χ est irréductible, alors, elle se réalise sur

un espace réel si et seulement si la moyenne des $\chi(g^2)$, $g \in G$ vaut 1, [2, Proposition XIV-2.2.1], (pareil, la preuve complète ne tient pas en quinze minutes).

III. Questions classiques (ou pas) du jury

1. Montrer que deux matrices de projection sont équivalentes si et seulement si elles sont semblables.

   Il suffit de montrer l'implication. Les matrices de projections sont annulées par le polynôme scindé simple $X^2 - X$, et à ce titre, sont diagonalisables avec pour seules valeurs possibles 0 et 1. Le rang détermine donc la dimension du noyau, donc, la multiplicité de 0, donc, la multiplicité de 1. En conclusion, si deux matrices de projection sont équivalentes, elles ont même rang, donc même spectre, et comme elles sont diagonalisables, elles sont semblables.

2. Montrer que A est nilpotente si et seulement si A est semblable à $2A$, [1, Exercice III-D.13]. Dans le même genre, montrer que sur $\mathbb{C}$, A est nilpotente si et seulement si la matrice nulle est dans l'adhérence de la classe de similitude de A.

   Soit (A_m) une suite de matrices semblables qui tend vers 0. Alors le polynôme caractéristique des A_m est constant égal à χ_A. De plus, par continuité de la fonction $M \mapsto \chi_M$, χ_A est le polynôme caractéristique de la fonction nulle. Donc, $\chi_A = X^n$ et A est nilpotente par Cayley-Hamilton. Réciproquement, si A est nilpotente, on se ramène à A strictement triangulaire supérieure et on conjugue par $P := \mathrm{diag}(1, \varepsilon, \cdots, \varepsilon^{n-1})$, pour faire tendre ϵ vers 0.

3. Montrer que les uniques classes de similitudes bornées pour l'espace normé $\mathcal{M}_n(\mathbb{C})$ sont les classes d'homothéties, [1, Exercice III-D.23].

4. Montrer la connexité de l'ensemble des matrices complexes de rang inférieur à r, [1, Proposition I-1.4.1].

5. Donner une matrice (non scalaire !) de $\mathcal{M}_n(\mathbb{K})$ telle que sa classe de similitude ne contient pas de matrice compagnon.

   Une matrice compagnon a la propriété que son polynôme minimal est égal à son polynôme caractéristique, et ceci est alors valable pour toute matrice de sa classe de similitude. Il suffit de prendre une matrice diagonalisable avec multiplicité algébrique d'une valeur propre.

6. Soit S une matrice symétrique réelles et O_S sa classe de congruence dans $\mathbb{C}$, c'est-à-dire l'ensemble des ${}^tPSP$, $P \in \mathrm{GL}_n(\mathbb{C})$. Décrire $O_S \cap \mathcal{M}_n(\mathbb{R})$ en terme de signatures, [1, V-1.4.8].

7. Soit q la puissance d'un nombre premier impair, et A une matrice symétrique inversible sur $\mathbb{F}_q$ de taille n. Montrer qu'il existe $P \in \mathrm{GL}_n(\mathbb{F}_{q^2})$ telle que $A =^t PP$.

   On veut montrer que A est congruente à l'identité, donc que son discriminant est un carré de $\mathbb{F}_{q^2}^*$ (car A est inversible). On calcule donc le symbole de Legendre $\det(A)^{(q^2-1)/2}$ *qui vaut* 1 *car* $\det(A)^{(q-1)/2} = \pm 1$ *et $q+1$ est pair.*

8. (Variante) Soit q la puissance d'un nombre premier impair, et A, B deux matrices symétriques inversibles sur $\mathbb{F}_q$ de taille n. Montrer que si elles sont $\mathbb{F}_{q^m}$-congruentes, avec m impair, elles sont $\mathbb{F}_q$-congruentes.

   On fait comme précédemment, en notant que $q^m - 1 = (q-1)(q^{m-1} + \cdots + 1)$*, le second facteur étant impair.*

IV. Les développements

1. Description de l'adhérence de l'ensemble des matrices de rang r (voir plus haut), [1, Proposition I-1.4.1]. Niveau : 3/5, Originalité : 2/5.
2. Classification des formes quadratiques sur $\mathbb{C}$, $\mathbb{R}$ et $\mathbb{F}_q$, [1, Théorème V-1.4.1]. Niveau : 3/5, Originalité : 3/5.
3. Loi de réciprocité quadratique par les formes quadratiques, [1, Théorème V-C.5]. Niveau : 4/5, Originalité : 3/5.
4. Une matrice est diagonalisable dans $\mathbb{C}$ si et seulement si sa classe de conjugaison est fermée, [1, Proposition III-1.7]. Niveau : 3/5, Originalité : 3/5.
5. Deux matrices de $\mathcal{M}_n(\mathbb{K})$ $\mathbb{L}$-semblables sont $\mathbb{K}$-semblables, [1, Corollaire III-5.14]. Niveau : 4/5, Originalité : 4/5.
6. Le théorème de Harish-Chandra (voir plus haut), [1, Exercice III-D.29]. Niveau : 4/5, Originalité : 5/5.
7. Adhérence de l'orbite de congruence d'une matrice symétrique réelle [1, Proposition V-1.5.2]. Niveau : 4/5, Originalité : 5/5.
8. Décomposition en cellules de la grassmannienne et calcul du cardinal sur $\mathbb{F}_q$ comme un polynôme en q, voir [2, IV-2.6.1], puis [3, Corollaire IV-1.5], Niveau : 4/5, Originalité : 5/5.

Références

[1] Philippe Caldero et Jérôme Germoni. *Nouvelles Histoires Hédonistes de Groupes et de Géométries.* Calvage et Mounet, 2017.

[2] Philippe Caldero et Jérôme Germoni. *Nouvelles Histoires Hédonistes de Groupes et de Géométries, tome second.* Calvage et Mounet, 2018.

[3] Philippe Caldero et Jérôme Germoni. *Histoires Hédonistes de Groupes et de Géométries-Tome 2.* Calvage et Mounet, 2015.