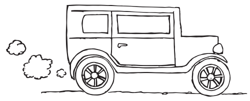

Chapitre 2

Formes quadratiques

You know, there are a million ways to die in the west, Clinch. There's, uh, famine, disease, gunfights... And, uh, wild animals. You know, like snakes. And, you know, the funny thing is, you don't even have to get bitten.

A Million Ways to Die in the West, Seth MacFarlane (2014)

2.1 Prérequis sur les formes quadratiques

Toutes les preuves sont dans [3, Chap. V, Annexes A et B].

Formes quadratiques, diverses approches

Les formes quadratiques peuvent être abordées de différentes façons : par les fonctions polynômes, par les formes bilinéaires symétriques, par les matrices, et même, dans une certaine mesure, par leurs cônes isotropes.

Commençons par une première définition. Soit E un espace vectoriel de dimension finie n sur un corps $\mathbb{K}$. Si l'on fixe une base de E, on peut assimiler l'espace E à l'espace vectoriel $\mathbb{K}^n$. Une *forme quadratique* sur E est alors tout simplement une fonction polynôme homogène de degré 2 sur $\mathbb{K}^n$ à coefficients dans $\mathbb{K}$. On fera tout de même attention à vérifier que cette définition ne dépend pas de la base choisie.

Si la caractéristique de $\mathbb{K}$ est différente de 2, ce que l'on supposera par la suite, la théorie part du bon pied. En effet, dans ce cas, il est équivalent de définir une forme quadratique q sur E et une *forme bilinéaire symétrique* b sur $E \times E$ *via* les formules

$$b(x,y) = \frac{q(x+y) - q(x) - q(y)}{2}, \quad q(x) = b(x,x).$$

L'étude des formes quadratiques se ramène alors à celle des formes bilinéaires symétriques. Cela permet, entre autres, de définir la notion d'orthogonalité : deux vecteurs x et y de E sont dit q-orthogonaux si $b(x,y) = 0$. On définit également dans la partie 2.1 l'orthogonal d'un sous-espace de E, ou d'un sous-ensemble.

La forme bilinéaire associée à q permet en outre de définir une application linéaire ψ_b entre E et son dual E^* ; explicitement :

$$\forall (x,y) \in E \times E, \quad \langle \psi_b(x), y \rangle = b(x,y),$$

où E^* désigne le dual de E et $\langle \cdot, \cdot \rangle : E^* \times E \to \mathbb{K}$ désigne ici l'évaluation (d'une forme linéaire en un vecteur).

Grâce à cette application linéaire, nous pouvons faire le lien avec les matrices ! On peut en effet définir la matrice (symétrique) de la forme q (ou de la forme b) dans une base fixée $\underline{e}$ par

$$\mathrm{mat}_{\underline{e}}(q) = \mathrm{mat}_{\underline{e},\underline{e}^*}(\psi_b) = \big(b(e_j, e_i)\big)_{1 \leq i,j \leq n},$$

où $\underline{e}^*$ désigne la base duale de $\underline{e}$.

Si x et y dans E s'écrivent respectivement X et Y en colonne dans la base $\underline{e}$, alors on obtient les formules :

$$q(x) = {}^tXA_qX, \; b(X,Y) = {}^tXA_qY,$$

où A_q désigne la matrice de q dans $\underline{e}$.

Plus précisément, si l'on note $A_q = (a_{ij})$, alors q s'écrit

$$q(x) = \sum_{i=1}^n a_{ii}x_i^2 + \sum_{1 \leq i < j \leq n} a_{ij}x_ix_j, \text{ avec } x = \sum_{i=1}^n x_ie_i.$$

Cela signifie que l'écriture de q en tant que fonction polynomiale est *unique*, puisque la matrice A_q est entièrement déterminée par b, et que b est entièrement déterminée par q (en caractéristique différente de 2). Dit autrement, on peut identifier polynôme et fonction polynôme en degré homogène 2 (c'est-à-dire, pour les formes quadratiques).

Il convient également de définir le noyau et le rang d'une forme quadratique. On appelle *noyau* de b le noyau du morphisme associé ψ_b :

$$\mathrm{Ker} b = \mathrm{Ker} \psi_b = \{x \in E \,:\, \forall y \in E, \; b(x,y) = 0\}.$$

On dit que b est *non dégénérée* si son noyau est réduit à $\{0\}$. Dans ce cas, la matrice A_q est inversible.

Plus généralement, on appelle *rang* de q, le rang de l'application linéaire ψ_b. C'est aussi le rang de la matrice A_q dans n'importe quelle base.

Pour terminer, à une forme quadratique, on peut associer le cône isotrope

$$\mathcal{C}_q = \{x \in E, \, q(x) = 0\},$$

que l'on ne confondra pas avec le noyau de la forme !

On verra dans l'exercice 2.2.3, que si le cône isotrope est non réduit à zéro, il caractérise q à scalaire près[1].

[1] L'idée est que la q-orthogonalité se voit sur le cône isotrope.

Classification : les trois états de la congruence

On dira que deux formes quadratiques q et q' sont *congruentes* s'il existe φ dans $\mathrm{GL}(E)$ tel que $q' = q \circ \varphi$. La congruence fournit clairement une relation d'équivalence dans l'ensemble des formes quadratiques, et l'on entend par « classification des formes quadratiques » l'étude des classes d'équivalence pour cette relation.

Si l'on munit E d'une base $\underline{e}$, on peut voir la congruence de façon matricielle :

$$A_{q'} = {}^tPA_qP, \quad \text{avec } P := \mathrm{mat}_{\underline{e}}(\varphi).$$

On dira donc que deux matrices (symétriques) A et B sont *congruentes* s'il existe une matrice P de $\mathrm{GL}_n(\mathbb{K})$ telle que $B = {}^tPAP$.

On peut considérer la congruence matricielle de deux façons :

• soit $A = \mathrm{mat}_{\underline{e}}(q)$ et $B = \mathrm{mat}_{\underline{e}}(q')$, avec q' et q congruentes, et dans ce cas P est la matrice d'un automorphisme φ tel que $q' = q \circ \varphi$;

• soit $A = \mathrm{mat}_{\underline{e}}(q)$ et $B = \mathrm{mat}_{\underline{e}'}(q)$, et dans ce cas P est la matrice de passage de la base $\underline{e}$ vers la base $\underline{e}'$.

Enfin, la notion de congruence se transfère sur les cônes isotropes puisque, si $q' = q \circ \varphi$, alors

$$\mathcal{C}_{q'} = \varphi^{-1}(\mathcal{C}_q).$$

On classifiera ainsi à l'envi les cônes isotropes, les formes quadratiques, les matrices symétriques. On pourra aussi décider d'affiner la classification en demandant que φ, resp. P, appartiennent à un sous-groupe fixé G de $\mathrm{GL}(E)$, resp. $\mathrm{GL}_n(\mathbb{K})$. Sur $\mathbb{R}$, le groupe pourra être le groupe orthogonal (classification euclidienne), ou le groupe des similitudes.

Invariants pour la classification

La question que l'on se pose à présent est la suivante : qu'est-ce que deux formes quadratiques, ou deux matrices symétriques, congruentes, ont en commun ? Inversement, comment reconnaître deux formes quadratiques, ou deux matrices, congruentes ? On attaque ici le problème profond de la recherche d'invariants de congruence.

C'est sous forme matricielle que l'on voit le mieux les invariants, puisque les matrices se rapportent aux applications linéaires que l'on connaît très bien.

Le rang est déjà un premier invariant intéressant. En effet, comme ${}^tPAP$ et A sont deux matrices équivalentes, elles ont même rang.

Le déterminant, qui était un invariant de similitude, n'est en revanche pas un invariant de congruence, puisque $\det({}^tPAP) = \det(A)\det(P)^2$. Toutefois, de cette égalité, on peut récupérer un invariant utile : le *discriminant* d'une forme quadratique ; il s'agit du déterminant de $\mathrm{mat}_{\underline{e}}(q)$ à un facteur carré près, qui est bien un invariant, car il ne dépend pas de la base $\underline{e}$ choisie. Par exemple, sur $\mathbb{R}$, le discriminant est juste le signe du déterminant (dans le sens $-$, 0 ou $+$). Sur $\mathbb{C}$, c'est toujours un invariant, mais il est trivial (0 ou 1) car tout complexe est un carré.

La notion de positivité sur le corps $\mathbb{R}$, due au fait que c'est un corps ordonné, est intéressante à exploiter dans le cadre de la classification des formes quadratiques sur $\mathbb{R}$: à une forme quadratique q, on associe sa signature (s, t) où s, resp. t, désigne le maximum de la dimension d'un sous-espace F de E tel que $q|F$ est définie positive, resp. définie négative.

Rang, discriminant et signature sont les trois invariants de congruence à disposition de l'étudiant en mastère. Mais comment se calculent-ils ? En fait, comme ils sont invariants, ils peuvent se calculer modulo changement de base. C'est la méthode de Gauss, qui fournit l'outil de calcul le plus efficace, voir [3, V-1.3.1] pour une description. Cette méthode nous permet d'écrire q sous la forme

$$q(x) = \sum_{i=1}^{r} \alpha_i \ell_i(x)^2,\ \alpha_i \in \mathbb{K}^*,$$

où r désigne le rang de q et $(\ell_i)_{1 \leq i \leq r}$ est une famille libre de formes linéaires sur E. Cela a une conséquence matricielle : si l'on complète la famille libre $(\ell_i)_{1 \leq i \leq r}$ en une base de E^*, alors dans la base (anté)-duale de E, la matrice de q est $\operatorname{mat}(q) = \operatorname{diag}(\alpha_1, \ldots, \alpha_r, 0, \ldots, 0)$.

Cette écriture a plusieurs avantages : tout d'abord, elle permet de visualiser la valeur du rang r de la matrice, et donc, de la forme quadratique. Le discriminant vaut, quant à lui, 0 si $r < n$ et $\prod_i \alpha_i$ (modulo les carrés de $\mathbb{K}$) si $r = n$. Enfin, sur $\mathbb{R}$, la signature (s, t) se voit facilement : s, resp. t, est le nombre de réels α_i positifs, resp. négatifs.

Pour l'instant, nous avons juste exhibé des invariants de congruence, c'est-à-dire des objets que l'on associe à une matrice symétrique, et qui sont égaux pour deux matrices congruentes. Maintenant, nous avons besoin, si possible, d'invariants *totaux* de congruence, c'est-à-dire des objets qui soient égaux si et seulement si les deux matrices sont congruentes.

Il s'agit là d'un problème plus difficile, et la réponse dépend drastiquement du corps de base ! Voici un tableau donnant l'invariant total, ainsi que la forme normale associée, en fonction du corps considéré. Précisons que la forme normale d'une classe de congruence donnée (déterminée par l'invariant total) désigne l'élément « le plus simple possible », tant qu'à faire, le plus proche de l'identité possible, de l'orbite en question. Rappelons que toute matrice d'une orbite est congruente à tous les éléments appartenant à la même orbite. Cette remarque permet de voir que l'on peut, dans la théorie, se contenter de la forme normale, plus facile à considérer, de chaque orbite pour pouvoir établir les résultats relatifs à l'étude des formes quadratiques[2].

Corps $\mathbb{K}$	Invariant total	Forme normale
$\mathbb{C}$	Rang : $r \in \{0, \ldots, n\}$	$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$
$\mathbb{R}$	Signature : (s, t) ; $s + t \leqslant n$	$\begin{pmatrix} I_s & 0 & 0 \\ 0 & -I_t & 0 \\ 0 & 0 & 0 \end{pmatrix}$
$\mathbb{F}_q$	Discriminant : $\overline{\delta} \in \mathbb{K}^*/(\mathbb{K}^*)^2$	$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \delta \end{pmatrix}$

[2]La dernière ligne ne concerne que le cas non dégénéré.

Trois états du théorème de Sylvester

La deuxième ligne du tableau précédent, correspondant au corps $\mathbb{R}$, est une reformulation du théorème de Sylvester. Ce théorème est un théorème central (avec le théorème spectral) dans l'étude des formes quadratiques du programme de licence et mastère. Il n'est pas inutile de s'y pencher un peu. Nous allons voir trois déclinaisons de ce théorème selon l'approche proposée des formes quadratiques. Voir chapitre 3, pour une approche par les actions de groupes.

Théorème 2.1.1. [Théorème d'inertie de Sylvester ; version 1]

Soit q une forme quadratique sur un espace vectoriel réel *de dimension n. Il existe deux entiers s et t, $0 \leq s, t$, $s + t \leq n$ et une base $\mathbf{e} = (e_1, \dots, e_n)$ de E tels que :*

$$q(x) = \sum_{i=1}^{s} x_i^2 - \sum_{j=1}^{t} x_{s+j}^2, \quad \textit{pour tout vecteur } x = \sum_{i=1}^{n} x_i e_i.$$

De plus, les entiers s et t ne dépendent que de la forme quadratique q, et non d'un choix de la base :

$$s + t = \operatorname{rg} q, \quad s = \max\{\dim F : F \textit{ s.e.v. ; } q|_F \textit{ est définie positive}\}.$$

De façon équivalente, on peut l'exprimer en termes matriciels, comme suit.

Théorème 2.1.2. [Théorème d'inertie de Sylvester ; version 2]

Soit n un entier naturel et soit A une matrice symétrique $n \times n$ réelle. Il existe deux entiers s et t, $0 \leq s, t$, $s + t \leq n$, et une matrice inversible P telle que l'on ait, par blocs :

$${}^tPAP = \begin{pmatrix} \mathrm{I}_s & 0 & 0 \\ 0 & -\mathrm{I}_t & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

De plus, les entiers s et t ne dépendent que de A, et non pas de P.

Voici une version plus géométrique du théorème de Sylvester dont on trouvera une idée de preuve à la suite de l'exercice 2.2.3, voir remarque 2.2.5. Il s'agit ici de reconnaître la signature (modulo permutation) d'une forme quadratique à partir du cône isotrope qu'elle définit.

Théorème 2.1.3. [Théorème d'inertie de Sylvester ; version 3]

Soit n un entier naturel et E un espace vectoriel réel de dimension n. Notons $\mathcal{C}_q \subset E$ le cône isotrope associé à la forme quadratique q. On suppose que q, resp. q', est une forme quadratique de signature (s, t), resp. (s', t'). Les deux assertions sont équivalentes :

(i) *il existe φ dans $\mathrm{GL}(E)$ telle que $\mathcal{C}_{q'} = \varphi(\mathcal{C}_q)$,*

(ii) *$(s', t') = (s, t)$ ou $(s', t') = (t, s)$.*

On termine cette partie sur le théorème spectral, qui peut être vu comme un raffinement du théorème de Sylvester où la $\mathrm{GL}_n(\mathbb{R})$-congruence est remplacée par la $\mathrm{O}_n(\mathbb{R})$-congruence. La signature ne suffit plus pour assurer cette congruence plus fine. L'invariant total est alors donné par l'ensemble ordonné des valeurs propres (réelles) de la matrice. Le moyen le plus simple de le voir est le suivant : deux matrices symétriques réelles sont $\mathrm{O}_n(\mathbb{R})$-congruentes si et seulement si elles ont même ensemble de valeurs propres, avec mutiplicités, voir l'exercice 2.4.1 et 2.4.2, ainsi que la remarque 2.4.3.

Orthogonalité et isotropie

L'orthogonalité est un outil naturel et puissant tant dans la réduction que dans l'étude des formes quadratiques. Elle permet de faire des récurrences et surtout d'obtenir des formes normales (par exemple diagonales par blocs) pour les classes de similitude ou pour les classes de congruence de matrices.

Étant donné une forme bilinéaire symétrique b sur un espace E, on appelle *orthogonal* d'une partie P et l'on note $P^\perp$ l'ensemble :

$$P^\perp = \{x \in E \,:\, \forall y \in P,\ b(x,y) = 0\}.$$

Si $F \subset E$, son orthogonal $F^\perp$ est un sous-espace vectoriel. Il pourra être bon de le voir comme un noyau : il s'agit du noyau de l'application linéaire de E dans F^* donnée par $r_F \circ \psi_b$, où r_F désigne le morphisme de E^* dans F^* qui, à une forme linéaire, associe sa restriction à F.

Le noyau de la restriction de b au sous-espace F est :

$$\operatorname{Ker} b|_{F\times F} = \{x \in F \,:\, \forall y \in F,\ b(x,y) = 0\} = F \cap F^\perp.$$

On dira que le sous-espace F est *régulier* si la restriction de b à F est non dégénérée, autrement dit si $F \cap F^\perp = \{0\}$. Dans le cas contraire, le sous-espace F est dit *isotrope*. On dit que le sous-espace F est *totalement isotrope* si la restriction de b à F est nulle, c'est-à-dire si $F \subset F^\perp$.

C'est avec l'orthogonalité que commencent les difficultés pour les étudiants qui s'étaient confortablement installés dans la routine des espaces euclidiens. Effectivement, tout sous-espace F d'un espace euclidien (E,q) hérite de la structure euclidienne : puisque la forme q vérifie $q(x) > 0$ pour x non nul de E, la propriété se transmet à F et donc, F et $F^\perp$ sont en somme directe.

En revanche, si la forme q n'est plus euclidienne, mais seulement non dégénérée (E est dit *régulier* pour la forme), et si F est un sous-espace, alors F peut intersecter son orthogonal $F^\perp$ non trivialement, et dans ce cas, F n'est pas régulier pour cette même forme. En gros, il faut bien se rappeler que la structure euclidienne est « héréditaire » (c'est-à-dire se transmet de E à F), mais que la régularité ne l'est pas ! L'orthogonal admet toujours (dans le cas q non dégénérée) la dimension attendue, c'est-à-dire la codimension de F, mais il peut ne pas être en somme directe avec F : on a alors une intersection $F \cap F^\perp$ non triviale. Cette intersection n'est autre que le *noyau* de la forme $q|_F$, restriction de q à F.

Voici une compilation de résultats à bien connaître, voir [3, V-A.1.14].

Lemme 2.1.4. *Soit b une forme bilinéaire sur E et F un sous-espace de E. Alors :*

(i) $\dim E + \dim(F \cap \operatorname{Ker} b) = \dim F + \dim F^\perp$*, et en particulier :* $\dim F^\perp \geq \dim E - \dim F$ *;*

(ii) *si b est non dégénérée,* $\dim E = \dim F + \dim F^\perp$ *;*

(iii) *si F est régulier, on a :* $E = F \oplus F^\perp$ *;*

(iv) *si b est non dégénérée,* $(F^\perp)^\perp = F$.