Chapitre 10

Espace dual et Formes Bilinéaires

10.1 Espace Dual

Définitions. Le *dual* d'un K-espace vectoriel V est l'espace vectoriel des applications linéaires définies sur V à valeurs dans le corps K. On le note

$$V^* = \mathcal{L}(V, K).$$

Un élément de V^* s'appelle un *covecteur* de V ou une *forme linéaire* sur V.

Exemples. 1.) Toute forme linéaire sur K^n est une application $\varphi : K^n \to K$ qui peut s'écrire

$$\varphi(x) = \varphi(x_1, \dots, x_n) = a_1 x_1 + \dots, +a_n x_n, \qquad (a_i \in K).$$

2.) La *trace* défini une forme linéaire sur l'espace vectoriel $M_n(K)$ des matrices carrées de taille $n \times n$ sur K.
3.) Si $V = C^0([a, b])$ est l'espace des fonctions continues sur l'intervalle $[a, b]$ et $x_0 \in [a, b]$, alors l'*évaluation en* x_0 défini une forme linéaire sur V^*. On la note

$$\delta_{x_0} : C^0([a, b]) \to \mathbb{R}, \qquad \delta_{x_0}(g) = g(x_0).$$

On dit parfois que le covecteur δ_{x_0} est la *masse de Dirac* concentrée au point x_0.
4.) Une autre forme linéaire sur $C^0([a, b])$ est l'intégration :

$$I_{[a,b]} : C^0([a, b]) \to \mathbb{R}, \qquad I_{[a,b]}(g) = \int_a^b g(x)dx.$$

Proposition 10.1.1. *Tout espace vectoriel de dimension finie est isomorphe à son dual.*

Preuve. Soit V un espace vectoriel de dimension finie sur le corps K et V^* son dual. Alors Les espaces V et V^* ont même dimension car

$$\dim(V^*) = \dim(\mathcal{L}(V, K)) = \dim(V) \cdot \dim(K) = \dim(V).$$

Ces deux espaces vectoriels sont donc isomorphes.

□

Remarque. Observons que cette démonstration ne repose pas sur un argument direct mais elle utilise la théorie de la dimension. Pour un espace vectoriel de dimension infinie l'argument ne marche pas et on peut démontrer qu'*un espace vectoriel de dimension infinie n'est jamais isomorphe à son dual.*

La proposition suivante complète la précédente.

Proposition 10.1.2. *Soit $\mathcal{B} = \{v_1, \dots v_n\}$ une base de l'espace vectoriel V. Notons $\varphi_i \in V^*$ le covecteur défini par*

$$\varphi_i(v_j) = \delta_{ij} = \begin{cases} 1, & si\ i = j, \\ 0, & si\ i \neq j. \end{cases} \tag{10.1}$$

Alors $\mathcal{B}^* = \{\varphi_1, \dots \varphi_n\}$ est une base de V^*.

Preuve. On remarque que

$$\mathrm{Card}(\mathcal{B}^*) = \mathrm{Card}(\mathcal{B}) = n = \dim(V^*),$$

il suffit de prouver que $\mathcal{B}^*$ est une famille libre. Supposons que $\sum_{i=1}^n \lambda_i \varphi_i = 0$, alors on a pour tout j

$$0 = \left(\sum_{i=1}^n \lambda_i \varphi_i\right)(v_j) = \sum_{i=1}^n \lambda_i \varphi_i(v_j) = \sum_{i=1}^n \lambda_i \delta_{ij} = \lambda_j.$$

Donc $\lambda_j = 0$ pour tout j. On a montré que $\mathcal{B}^*$ est une famille libre de V^*, et c'est donc une base puisque son cardinal est égal à la dimension de V.

□

Définition. La base $\mathcal{B}^*$ s'appelle la *base duale* de $\mathcal{B}$. On note parfois v_i^* le covecteur φ_i et on dit que v_i^* est le covecteur dual (ou la forme linéaire duale) au vecteur de base v_i. L'équation (10.1) s'appelle la *relation de dualité* entre les deux bases.

Lorsque $\{e_1, \dots, e_n\}$ est la base canonique de K^n, la base duale est notée $\{\varepsilon_1, \dots, \varepsilon_n\} \subset (K^n)^*$ et la relation de dualité s'écrit

$$\varepsilon_i(e_j) = \delta_{ij}.$$

Problème. Pour illustrer ces notions considérons le problème suivant : Soit $\{v_1, \dots, v_n\}$ une base quelconque de $\mathbb{R}^n$ et notons $\theta_1, \dots, \theta_n \subset (\mathbb{R}^n)^*$ la base duale. On demande de déterminer la matrice de transition de la base duale canonique $\{\varepsilon_i\}$ vers la base $\{\theta_i\}$ à partir de la matrice de transition de la base canonique $\{e_i\}$ vers la base $\{v_i\}$.

Solution. Notons P la matrice de transition de la base $\{e_i\}$ vers la base $\{v_i\}$ et P' la matrice de transition de la base $\{\varepsilon_i\}$ vers la base $\{\theta_i\}$. Rappelons que par définition

$$v_j = \sum_{k=1}^n p_{kj} e_k, \qquad \text{et} \qquad \theta_i = \sum_{l=1}^n p'_{li} \varepsilon_l.$$

Par définition de ε_l on a donc

$$\theta_i(e_k) = \sum_{l=1}^n p'_{li} \varepsilon_l(e_k) = \sum_{l=1}^n p'_{li} \delta_{lk} = p'_{ki},$$

et avec la relation de dualité entre les bases $\{\theta_i\}$ et $\{v_j\}$, nous obtenons.

$$\delta_{ij} = \theta_i(v_j) = \theta_i\left(\sum_{k=1}^{n} p_{kj}e_k\right) = \sum_{k=1}^{n} p_{kj}\theta_i(e_k) = \sum_{k=1}^{n} p_{kj}p'_{ki}.$$

Cette relation s'écrit matriciellement $P^\top P' = \mathrm{I}_n$, la matrice cherchée P' est donc la matrice inverse de la transposée de P.

$$P' = \left(P^\top\right)^{-1} = \left(P^{-1}\right)^\top.$$

Cette matrice s'appelle la *matrice contragrédiente* de P. (matrice)

Exemple. Pour trouver la base duale $\mathcal{B}^* \subset (K^2)^*$ de la base $\mathcal{B} = \{v, w\} = \{(2,1), (-1,1)\} \subset K^2$ on cherche la matrice de changement de base P et sa contragrédiente P' :

$$P = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \qquad \text{et} \qquad P' = (P^{-1})^\top = \frac{1}{3}\begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}.$$

Notons cette base duale $\mathcal{B}^* = \{\varphi, \psi\}$, alors ces covecteurs sont donnés par

$$\varphi = \frac{1}{3}(\varepsilon_1 + \varepsilon_2) \quad \text{et} \quad \psi(x,y) = \frac{1}{3}(-\varepsilon_1 + 2\varepsilon_2).$$

Ces covecteurs sont donc les fonctions $K^2 \to K$ telles que

$$\varphi(x,y) = \frac{1}{3}(x+y) \quad \text{et} \quad \psi(x,y) = -\frac{1}{3}x + \frac{2}{3}y.$$

On vérifie facilement que $\varphi(v) = \psi(w) = 1$ et $\varphi(w) = \psi(v) = 0$.

Proposition 10.1.3. *Soit $\mathcal{B} = \{v_1, \dots . v_n\}$ une base de l'espace vectoriel V et $\mathcal{B}^* = \{\varphi_1, \dots . \varphi_n\}$ la base duale de V^*, alors on a les propriétés suivantes :*

(a) Tout vecteur $x \in V$ s'écrit

$$x = \sum_{i=1}^{n} \varphi_i(x)v_i.$$

(b) Tout covecteur $\psi \in V^*$ s'écrit

$$\psi = \sum_{i=1}^{n} \psi(v_i)\varphi_i.$$

Preuve. (a) Développons le vecteur x dans la base $\mathcal{B}$, on a $x = \sum_{j=1}^{n} x_j v_j$, donc

$$\varphi_i(x) = \varphi_i\left(\sum_{j=1}^{n} x_j v_j\right) = \sum_{j=1}^{n} x_j \varphi_i(v_j) = \sum_{j=1}^{n} x_j \delta_{ij} = x_i,$$

par conséquent $x = \sum_{i=1}^{n} x_i v_i = \sum_{i=1}^{n} \varphi_i(x) v_i$.

(b) Développons le covecteur ψ dans la base duale $\mathcal{B}^*$, on a $\psi = \sum_{j=1}^{n} \xi_j \varphi_j$, donc

$$\psi(v_i) = \left(\sum_{j=1}^{n} \xi_j \varphi_j\right)(v_i) = \sum_{j=1}^{n} \xi_j \varphi_j(v_i) = \sum_{j=1}^{n} \xi_j \delta_{ij} = \xi_i,$$

et donc $\psi = \sum_{i=1}^n \xi_i \varphi_i = \sum_{i=1}^n \psi(v_i)\varphi_i$.

□

Définition. Soient V, W deux espaces vectoriels sur le corps K et $f : V \to W$ une application linéaire. L'*application duale* $f^* : W^* \to V^*$ est l'application linéaire définie par

$$f^*(\psi) = \psi \circ f.$$

Ainsi si $v \in V$ et $\psi \in W^*$, alors $f^*(\psi)(v) = \psi(f(v)) \in K$. Il est facile de vérifier que f^* est linéaire :

$$f^*(\alpha_1\psi_1 + \alpha_2\psi_2) = \alpha_1 f^*(\psi_1) + \alpha_2 f^*(\psi_2).$$

Théorème 10.1.4. *Soit $f \in \mathcal{L}(V, W)$ une application linéaire entre les espaces vectoriels V et W. Donnons-nous des bases $\mathcal{B} = \{v_1, \dots, v_n\}$ et $\mathcal{B}' = \{w_1, \dots, w_m\}$ de V et W respectivement. Si A est la matrice de f dans ces bases, alors la matrice de l'application duale $f^* \in \mathcal{L}(W^*, V^*)$ dans les bases duales $(\mathcal{B}')^*$ et $\mathcal{B}^*$ est la matrice transposée $A^\top$ de A :*

Remarque. Observons que A est une matrice de taille $m \times n$ ($m = \dim W$ et $n = \dim V$) et $A^\top$ est une matrice de taille $n \times m$, ce qui est compatible avec le fait que f est une application de W^* dans V^* et $n = \dim V^*$ $m = \dim W^*$.

Preuve. Notons $\mathcal{B}^* = \{\varphi_1, \dots, \varphi_n\} \subset V^*$ la base duale de $\mathcal{B}$ et $\mathcal{B}'^* = \{\psi_1, \dots, \psi_m\} \subset W^*$ la base duale de $\mathcal{B}'$. Rappelons que la matrice $A = (a_{ij})$ de f dans les bases $\mathcal{B}$, $\mathcal{B}'$ est définie par la relation

$$f(v_j) = \sum_{i=1}^m a_{ij} w_i.$$

Calculons $f^*(\psi_i)(v_j)$ en utilisant la définition de f^* et la propriété $\psi_i(w_k) = \delta_{ik}$:

$$f^*(\psi_i)(v_j) = \psi_i(f(v_j)) = \psi_i\left(\sum_{k=1}^m a_{kj} w_k\right) = \sum_{k=1}^m a_{kj}\psi_i(w_k) = \sum_{k=1}^m a_{kj}\delta_{ik} = a_{ij}.$$

La proposition précédente implique alors

$$f^*(\psi_i) = \sum_{j=1}^n f^*(\psi_i)(v_j)\varphi_j = \sum_{j=1}^n a_{ij}\varphi_j = \sum_{j=1}^n (a_{ji})^\top \varphi_j.$$

Ce qui montre que la matrice de f^* dans les bases duales est la matrice $A^\top$.

□

Corollaire 10.1.5. *Si V et W sont des espaces vectoriels de dimension finie, alors pour tout $f \in \mathcal{L}(V, W)$ on a $\operatorname{rang}(f^*) = \operatorname{rang}(f)$.*

Preuve. Choisissons des bases $\mathcal{B}$ et $\mathcal{B}'$ de V et W, et notons $A = M_{\mathcal{B}'\mathcal{B}}(f)$ la matrice de f dans ces bases. Alors $A^\top$ est la matrice de f^* dans les bases duales et on a donc

$$\operatorname{rang}(f^*) = \operatorname{rang}(A^\top) = \operatorname{rang}(A) = \operatorname{rang}(f).$$

□

10.1.1 Interpolation de Lagrange

Le *problème de l'interpolation* est le suivant : Soit $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction quelconque. La fonction n'est pas connue mais on dispose d'un nombre fini de mesures (ou d'observations) qui nous donnent un nombre fini de valeurs, disons

$$\varphi(a_1) = b_1,\ \varphi(a_2) = b_2, \dots, \varphi(a_n) = b_n, \tag{10.2}$$

et on voudrait, à partir de cette information, reconstruire la fonction φ. Ce problème n'a pas de solution unique en général, mais nous allons montrer qu'il est uniquement résoluble dans l'espace vectoriel des polynômes de degrés $\leq n-1$.

Pour résoudre ce problème, il est utile de considérer l'espace dual de l'espace vectoriel des polynômes. Rappelons que le *covecteur d'évaluation en* a est la forme linéaire $\delta_a : \mathbb{R}[x] \to \mathbb{R}$ définie par

$$\delta_a(p) = p(a),$$

pour tout polynôme $p(x)$. Pour résoudre le problème de l'interpolation dans l'espace $\mathcal{P}_{n-1} \subset \mathbb{R}[x]$ des polynômes de degré $\leq n-1$, on cherche d'abord n polynômes $\varphi_1, \dots \varphi_n \in \mathcal{P}_{n-1}$ tels que

$$\delta_{a_j}(\varphi_i) = \varphi_i(a_j) = \delta_{ij}.$$

Supposons que ces polynômes ont été construits, alors la solution du problème (10.2) est clairement donnée par

$$\varphi(x) = \sum_{i=1}^{n} b_i \varphi_i(x).,$$

en effet on a

$$\varphi(a_j) = \sum_{i=1}^{n} b_i \varphi_i(a_j) = \sum_{i=1}^{n} b_i \delta_{ij} = b_j.$$

Or la construction des polynômes $\varphi_i \in \mathcal{P}_{n-1}$ est élémentaire, il suffit de poser

$$\varphi_i(x) = \prod_{j \neq i} \frac{x - a_j}{a_i - a_j}.$$

Par conséquent la solution du problème d'interpolation (10.2) est donnée explicitement par la formule

$$\varphi(x) = \sum_{i=1}^{n} b_i \prod_{j \neq i} \frac{x - a_j}{a_i - a_j}.$$

Cette formule s'appelle la *formule d'interpolation de Lagrange*. Par exemple si $n = 3$ cette formule d'écrit

$$\varphi(x) = \frac{b_1(x-a_2)(x-a_3)}{(a_1-a_2)(a_1-a_3)} + \frac{b_2(x-a_1)(x-a_3)}{(a_2-a_1)(a_2-a_3)} + \frac{b_3(x-a_1)(x-a_2)}{(a_3-a_1)(a_3-a_2)}.$$

Remarque. La construction de Lagrange met en évidence le fait que les formes linéaires $\{\delta_{a_1}, \dots, \delta_{a_n}\}$ forment une base de $\mathcal{P}_{n-1}^*$, et que $\{\varphi_1, \dots, \varphi_n\} \subset \mathcal{P}_{n-1}$ est la base duale.

10.2 Couplage entre deux espaces vectoriels

Définition. Un *couplage*[1] entre deux espaces vectoriels V et W sur un corps K est une application :

$$\beta : V \times W \to K,$$

qui est bilinéaire, c'est-à-dire linéaire en chaque variable :

$$\beta(x, \mu_1 y_1 + \mu_2 y_2) = \mu_1 \beta(x, y_1) + \mu_2 \beta(x, y_2), \quad \beta(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 \beta(x_1, y) + \lambda_2 \beta(x_2, y).$$

La condition de bilinéarité peut également s'écrire :

$$\beta\left(\sum_{i=1}^{m} \lambda_i x_i, \sum_{j=1}^{n} \mu_j y_j\right) = \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_i \mu_j \beta(x_i, y_j).$$

Exemples 1. L'intégration définit un couplage entre $V = \mathbb{R}[x]$ et $W = C^0([a, b])$:

$$I : V \times W \to \mathbb{R}, \qquad I(p, f) = \int_a^b p(x) f(x) dx.$$

2. On note ℓ_1 l'espace vectoriel des *suites réelles absolument sommables*, qui est défini par

$$\ell_1 = \{\xi = (x_k)_{k=1}^{\infty} \mid x_k \in \mathbb{R}, \sum_{k=1}^{\infty} |x_k| < \infty\}.$$

On note aussi ℓ_∞ l'espace vectoriel des suites réelles bornées

$$\ell_\infty = \{\xi = (x_k)_{k=1}^{\infty} \mid x_k \in \mathbb{R}, \sup_k |x_k| < \infty\}.$$

Alors le *couplage de sommation* est défini par

$$\sigma : \ell_1 \times \ell_\infty \to \mathbb{R}, \qquad \sigma(\xi, \eta) = \sum_{k=1}^{\infty} x_k y_k.$$

3. L'exemple qui suit est une variante de l'exemple précédent. On fixe $p \in (0, \infty)$ et on note ℓ_p l'espace vectoriel des *suites p-sommables* :

$$\ell_p = \{\xi = (x_k)_{k=1}^{\infty} \mid x_k \in \mathbb{R}, \sum_{k=1}^{\infty} |x_k|^p < \infty\}.$$

Alors le couplage de sommation $\sigma : \ell_p \times \ell_q \to \mathbb{R}$ est bien défini à condition que $1/p + 1/q = 1$ (c'est une conséquence de l'*inégalité de Hölder*, démontrée au cours d'analyse).

4. Un couplage entre les espaces de matrices $M_{m,n}(K)$ et $M_{n,m}(K)$ est défini par la trace du produit matriciel :

$$\begin{array}{ccc} M_{m,n}(K) \times M_{n,m}(K) & \to & K \\ (A, B) & \mapsto & \mathrm{Trace}(A \cdot B) \end{array}$$

1. 'couplage' se dit 'pairing' en anglais.

5. Tout espace vectoriel V admet un couplage avec son dual :

$$\begin{array}{ccc} V^* \times V & \to & K \\ (\phi, v) & \mapsto & \phi(v). \end{array}$$

On l'appelle le *couplage canonique de V avec son dual.* Ce couplage est universel (il est défini pour tout espace vectoriel) et ne dépend pas du choix d'une base, ni d'aucun autre choix.

Couplage et dualité

A tout couplage $\beta : V \times W \to K$ entre deux K-espaces vectoriels on peut associer une application linéaire entre chacun des espaces vectoriels et le dual de l'autre. Ces applications linéaires

$$\beta_g : V \to W^* \quad \text{et} \quad \beta_d : W \to V^*$$

sont définies de la façon suivante :

$$\beta_g(v) \in W^* \text{ est le covecteur tel que } \beta_g(v)(w) = \beta(v, w) \text{ pour tout } w \in W.$$

De même

$$\beta_d(w) \in V^* \text{ est le covecteur tel que } \beta_d(w)(v) = \beta(v, w) \text{ pour tout } v \in V.$$

Les lettres 'g' et 'd' signifient que β_g agit sur la variable de gauche et β_d agit sur la variable de droite.

Définition 10.2.1. Le couplage $\beta : V \times W \to K$ est *non dégénéré* si

$$\forall v \in V, \quad \text{on a } [(\beta(v, y) = 0 \ \forall y \in W) \Leftrightarrow v = 0]$$

et

$$\forall w \in W, \quad \text{on a } [(\beta(x, w) = 0 \ \forall x \in V) \Leftrightarrow w = 0].$$

Lemme 10.2.2. *$\beta : V \times W \to K$ est non dégénéré si et seulement si β_g et β_d sont injectives.*

Preuve. La première condition de la définition précédente dit exactement que $\mathrm{Ker}(\beta_g) = \{0\}$ et la deuxième condition dit que $\mathrm{Ker}(\beta_d) = \{0\}$.

□

Corollaire 10.2.3. *Soit $\beta : V \times W \to K$ un couplage entre deux espaces vectoriels de dimension finies. Alors β est non dégénéré si et seulement si β_g et β_d sont des isomorphismes.*

Preuve. Observons d'abord que si β_g est injective, alors $\dim(V) \leq \dim(W^*) = \dim(W)$ et si β_d est injective, alors $\dim(W) \leq \dim(V^*) = \dim(V)$, par conséquent, si β est non dégénéré, alors $\dim(V) = \dim(W)$. On conclut la preuve en rappelant qu'une application linéaire injective entre deux espaces vectoriels de même dimension finie est un isomorphisme.

□

La notation "⟨bra|ket⟩" de Dirac

Si $\beta : V \times W \to K$ est un couplage entre deux K-espaces vectoriels, il est commode de noter

$$\langle v \mid w \rangle_\beta = \beta(v, w),$$

ou simplement $\langle v \mid w \rangle$ lorsque le couplage β a été fixé. Les homomorphismes définis précédemment peuvent alors s'écrire de façon plus concise

$$\beta_g(v) = \langle v \mid \cdot \rangle \in W^* \quad \text{et} \quad \beta_d(w) = \langle \cdot \mid w \rangle \in V^*.$$

Le vecteur $v \in V$ est alors vu comme un covecteur de W (i.e. un élément de W^*) et $w \in W$ est vu comme un covecteur de V (i.e. un élément de V^*). En mécanique quantique on utilise souvent la variante suivante de cette notation :

$$\beta_g(v) = \langle v \mid \quad \text{et} \quad \beta_d(w) = \mid w \rangle.$$

10.3 Formes bilinéaires sur un espace vectoriel

Définition. Une *forme bilinéaire* sur un K-espace vectoriel V est une application

$$g : V \times V \to K$$

qui est bilinéaire. Une forme bilinéaire est donc un couplage de V avec lui même. La bilinéarité signifie que g est linéaire en chacune de ses deux variables, ce qu'on peut aussi écrire sous la forme

$$g\left(\sum_{i=1}^{m} \lambda_i x_i, \sum_{j=1}^{n} \mu_j y_j\right) = \sum_{i,j=1}^{m} \lambda_i \mu_j g(x_i, y_j).$$

Exemples 1. Le produit scalaire standard sur $\mathbb{R}^n$ défini par

$$x \cdot y = x_1 y_1 + \cdots + x_n y_n$$

est une forme bilinéaire sur $\mathbb{R}^n$.

2. On définit une forme bilinéaire sur l'espace des $m \times n$ matrices sur le corps K par la formule

$$\begin{array}{ccc} M_{m,n}(K) \times M_{m,n}(K) & \to & K \\ (A, B) & \mapsto & \mathrm{Trace}(A^\top \cdot B) \end{array}$$

3. Si $C \in M_n(K)$ est une matrice carrée quelconque, on peut lui associer une forme bilinéaire sur K^n définie par

$$g(x, y) = \sum_{i,j=1}^{n} c_{ij} x_i y_j.$$

Définition. Soit g une forme bilinéaire définie sur un espace vectoriel V de dimension $n < \infty$, et soit $\mathcal{B} = \{v_1, \dots v_n\}$ une base de V. La matrice

$$G = (g_{ij}) \in M_n(K) \ \text{ définie par } g_{ij} = g(v_i, v_j)$$

s'appelle la *matrice de Gram*[2] de g relativement à la base $\mathcal{B}$.

Exemple. La matrice de Gram de la forme bilinéaire g définie sur $\mathbb{R}^2$ par

$$g(x,y) = ax_1y_1 + bx_1y_2 + cx_2y_1 + dx_2y_2 \quad \text{est} \quad G = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

On fera attention à ne pas confondre une matrice de Gram avec la matrice d'un endomorphisme. Dans les deux cas il s'agit d'une matrice carrée, mais leur signification est très différente. L'interprétation de la matrice de Gram vient de la proposition suivante :

Proposition 10.3.1. *Si* $x = \sum_{i=1}^n x_i v_i$ *et* $y = \sum_{j=1}^n y_j v_j$*, alors*

$$g(x,y) = \sum_{i,j=1}^n g_{ij} x_i y_j. \tag{10.3}$$

La preuve est une application immédiate de la bilinéarité de g.

□

Remarque. Si $\mathcal{B} = \{v_1, \dots, v_n\}$ est une base de V et si $X \in K^n$ et $Y \in K^n$ sont les vecteurs colonnes associés respectivement aux vecteurs $x = \sum_{i=1}^n x_i v_i \in V$ et $y = \sum_{j=1}^n y_j v_j \in V$:

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

alors

$$g(x,y) = X^\top G Y = (x_1 \cdots x_n) \cdot G \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \tag{10.4}$$

où G est la matrice de Gram de la forme bilinéaire g dans la base $\mathcal{B}$.

Corollaire 10.3.2. *Si* $\mathcal{B} = \{v_1, \dots, v_n\}$ *et* $\mathcal{B}' = \{v_1', \dots, v_n'\}$ *sont deux bases de* V*, et si* P *est la matrice de changement de base (i.e.* $P = M_{\mathcal{B}\mathcal{B}'}(\mathrm{Id}_V)$*), alors les matrices de Gram de la forme bilinéaire* g *dans ces deux bases sont reliées par*

$$G' = P^\top G P.$$

En particulier les matrices G *et* G' *ont le même rang.*

Définition. On appellera *rang* de la forme bilinéaire g le rang de sa matrice de Gram dans une base quelconque.

Preuve. Rappelons que si $x = \sum_{i=1}^n x_i v_i = \sum_{i=1}^n x_i' v_i'$, alors les vecteurs colonnes X et X' correspondants sont reliés par

$$X = PX'.$$

2. Jørgen Pedersen Gram, mathématicien danois 1850–1916.

Par conséquent nous avons d'une part

$$g(x, y) = X^\top GY = (PX')^\top G(PY') = (X'^\top P^\top)G(PY') = X'^\top(P^\top GP)Y';$$

et d'autre part $g(x, y) = X'^\top G'Y'$ pour tous $x, y \in K^n$. Ceci implique $G' = P^\top GP$.

□

Définition. Deux matrices carrées G_1, G_2 sont dites *congruentes* s'il existe une matrice inversible P telle que

$$G_2 = P^\top G_1 P.$$

Exercice. Montrer que la relation de congruence est une relation d'équivalence.

Attention de ne pas confondre la relation de congruence $G \sim P^\top GP$ avec la relation de similitude $G \sim P^{-1}GP$. Deux matrices sont congruentes si et seulement si elles représentent la même forme bilinéaire dans des bases différentes alors que deux matrices sont semblables si et seulement si elles représentent le même endomorphisme dans des bases différentes.

Définition. Le *produit tensoriel* de deux co-vecteurs $\phi, \psi \in V^*$ est la forme bilinéaire $\phi \otimes \psi : V \times V \to K$ définie par la formule

$$(\phi \otimes \psi)(x, y) = \phi(x)\psi(y).$$

Proposition 10.3.3. *Si $\mathcal{B} = \{v_1, \dots, v_n\}$ est une base de V et $\mathcal{B}^* = \{\varphi_1, \dots, \varphi_n\} \subset V^*$ est la base duale, alors toute forme bilinéaire $g : V \times V \to K$ s'écrit*

$$g = \sum_{i,j=1}^{n} g_{ij}\, \varphi_i \otimes \varphi_j,$$

où $G = (g_{ij})$ est la matrice de Gram de g dans la base $\mathcal{B}$.

Preuve. Notons $h = \sum_{i,j=1}^{n} g_{ij}\, \varphi_i \otimes \varphi_j$. Il faut montrer que $h = g$. Or par définition du produit tensoriel on a

$$h(v_\mu, v_\nu) = \sum_{i,j=1}^{n} g_{ij}\; \varphi_i(v_\mu)\varphi_j(v_\nu) = g_{\mu\nu} = g(v_\mu, v_\nu),$$

car $\varphi_i(v_\mu) = \delta_{\mu,i}$ et $\varphi_j(v_\nu) = \delta_{\nu,j}$. Ceci montre que g et h coïncident sur la base $\mathcal{B}$, donc $g = h$ par bilinéarité.

□

Corollaire 10.3.4. *L'ensemble des formes bilinéaires sur un espace vectoriel V est un espace vectoriel de dimension n^2 (si $n = \dim V$) et*

$$\{\varphi_i \otimes \varphi_j \mid 1 \le i, j \le n\}$$

est une base de cet espace vectoriel.

La proposition précédente nous dit que la matrice de Gram (g_{ij}) représente les composantes de la forme bilinéaire g dans la base $\{\varphi_i \otimes \varphi_j\}$.

10.4 Formes bilinéaires symétriques et antisymétriques

Soit V un espace vectoriel sur un corps K. On suppose, dans ce paragraphe et le suivant, que K n'est pas de caractéristique 2 (c'est-à-dire $1+1 \neq 0$ dans K).

Définition. Une forme bilinéaire $\alpha : V \times V \to K$ est dite *symétrique* si $\alpha(y,x) = \alpha(x,y)$ pour tous $x, y \in V$. Elle est *antisymétrique*[3] si $\alpha(y,x) = -\alpha(x,y)$ pour tous $x, y \in V$.

On observe que toute forme bilinéaire α sur V s'écrit de façon unique comme somme d'une forme bilinéaire symétrique et d'une forme bilinaire antisymétrique. On peut en effet écrire $\alpha(v,w) = \phi(v,w) + \psi(v,w)$ avec $\phi(v,w) := \frac{1}{2}(\alpha(v,w) + \alpha(w,v))$ et $\psi(v,w) := \frac{1}{2}(\alpha(v,w) - \alpha(w,v))$. Il est clair que ϕ est bilinéaire et symétrique et ψ est bilinéaire et antisymétrique.

Lorsque V est de dimension finie, on peut relier ces notions à la matrice de Gram. En effet la matrice de Gram G d'une forme bilinéaire dans une base quelconque est une matrice symétrique (i.e. $G^\top = G$) si et seulement si la forme bilinéaire est symétrique et elle est antisymétrique (i.e. $G^\top = -G$) si et seulement si la forme bilinéaire est antisymétrique.

Théorème 10.4.1. *Soit β une forme bilinéaire symétrique sur un K-espace vectoriel V de dimension finie. Alors il existe une base $\{v_1, \dots, v_n\}$ de V telle que $\beta(v_i, v_j) = 0$ si $i \neq j$.*

Définition. Une telle base est dite *orthogonale pour le forme bilinéaire β* (ou β-orthogonale).

Preuve. On raisonne par récurrence sur $n = \dim(V)$. Si $n = 1$, il n'y a rien à démontrer. Supposons donc que le théorème est démontré pour tout espace vectoriel de dimension $(n-1)$, et soit $\beta : V \times V \to K$ une forme bilinéaire symétrique sur un espace vectoriel V de dimension n.

Nous affirmons d'abord que si $\beta(v,v) = 0$ pour tout $v \in V$, alors on a aussi $\beta(u,v) = 0$ pour tous $u, v \in V$. Cela découle par exemple du raisonnement suivant :

$$0 = \beta(u+v, u+v) = \beta(u,u) + \beta(u,v) + \beta(v,u) + \beta(v,v) = \beta(u,v) + \beta(v,u) = 2\beta(u,v),$$

donc $\beta(u,v) = 0$. Dans ce cas toute base est orthogonale. Supposons donc qu'il existe $v_1 \in V$ tel que $\beta(v_1, v_1) \neq 0$ et définissons

$$W := \{w \in V \mid \beta(v_1, w) = 0\}.$$

Alors W est un sous-espace vectoriel de V de dimension $n-1$ (c'est le noyau du covecteur non nul $x \mapsto \beta(v_1, x)$). Par hypothèse de récurrence, il existe une base β-orthogonale de W, que nous notons $\{v_2, \dots, v_n\} \subset W$. Il est clair que $\beta(v_1, v_j) = 0$ pour tout $j = 2, \dots, n$ par définition de W et $\beta(v_i, v_j) = 0$ pour tous $i, j = 2, \dots, n$ par choix des vecteurs w_j. On a donc obtenu une base β-orthogonale $\{v_1, \dots, v_n\}$ de V.

$\square$

Remarque. Soit $\{v_1, \dots, v_n\}$ une base β-orthogonale de V et $x, y \in V$, alors

$$\beta(x,y) = \sum_{i=1}^{n} \alpha_i x_i y_i, \quad \text{avec } \alpha_i = \beta(v_i, v_i),$$

3. En anglais on dit *skew symmetric*.

où les x_i, y_i sont les composantes de x et y dans cette base. La matrice de Gram de β dans cette base est donc la matrice diagonale

$$B = (\beta(v_i, v_j)) = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$$

Corollaire 10.4.2. *Soit β une forme bilinéaire symétrique sur un K-espace vectoriel V de dimension finie. Alors il existe des formes linéaires $\phi_1, \dots, \phi_r \in V^*$ linéairement indépendantes et des scalaires non nuls $\alpha_1, \dots \alpha_r \in K$ tels que*

$$\beta = \sum_{i=1}^{r} \alpha_i \phi_i \otimes \phi_i, \quad \text{i.e.} \quad \beta(x,y) = \sum_{i=1}^{r} \alpha_i \phi_i(x)\phi_i(y) \quad \text{pour tous } x, y \in V.$$

De plus r est le rang de la matrice de Gram de β. Ce rang est donc indépendant de la base orthogonale choisie.

Preuve. Soit $\{v_1, \dots, v_n\}$ une base β-orthogonale de V et notons $\alpha_i = \beta(v_i, v_i)$. On a alors $\beta(v_i, v_j) = \alpha_i \delta_{ij}$. Soit maintenant $\{\phi_1, \dots \phi_n\} \subset V^*$ la base duale de $\{v_1, \dots, v_n\}$, alors on a

$$\beta = \sum_{i=1}^{n} \alpha_i \phi_i \otimes \phi_i.$$

Il suffit en effet de vérifier cette égalité sur des vecteurs de bases :

$$\left(\sum_{i=1}^{n} \alpha_i \phi_i \otimes \phi_i \right) (v_j, v_k) = \sum_{i=1}^{n} \alpha_i \phi_i(v_j)\phi_i(v_k) = \sum_{i=1}^{n} \alpha_i \delta_{ij}\delta_{ik} = \alpha_j \delta_{jk} = \beta(v_j, v_k).$$

Quitte à réordonner les vecteurs de base, on peut supposer que $\alpha_i \neq 0$ si et seulement si $1 \leq i \leq r$, on a donc finalement

$$\beta = \sum_{i=1}^{r} \alpha_i \phi_i \otimes \phi_i.$$

□

10.5 Formes quadratiques

Dans ce paragraphe, on suppose que V est un espace vectoriel de dimension finie sur un corps K de caractéristique $\neq 2$.

Définition. Une *forme quadratique* sur V est une application $Q : V \to K$ pour laquelle il existe une forme bilinéaire symétrique $\beta : V \times V \to K$ telle que

$$Q(v) = \beta(v, v).$$

Lemme 10.5.1. *La forme bilinéaire symétrique β est déterminé par Q de façon unique. Plus précisément, on a la* formule de polarisation *:*

$$\beta(v, w) = \frac{1}{4}(Q(v+w) - Q(v-w)). \tag{10.5}$$

Remarque. Les formules suivantes permettent également de retrouver la forme bilinéaire symétrique β à partir de la forme quadratique Q :

$$\beta(v, w) = \frac{1}{2}(Q(v+w) - Q(v) - Q(w)) \tag{10.6}$$

$$\beta(v, w) = \frac{1}{2}(Q(v) + Q(w) - Q(v-w)) \tag{10.7}$$

Les formules (10.5), (10.6) et (10.7) s'appellent les *formules de polarisation* De la forme quadratique Q.

Lorsque $V = K^n$, une forme quadratique Q sur $V = K^n$ s'écrit

$$Q(x) = \beta(x, x) = \sum_{i,j=1}^{n} b_{ij} x_i x_j, \tag{10.8}$$

où $b_{ij} = \beta(e_i, e_j)$. Ainsi une forme quadratique sur K^n n'est rien d'autre qu'un *polynôme homogène de degré 2 en n variables.*

Le corollaire 10.4.2 peut se reformuler pour les formes quadratiques de la façon suivante :

Théorème 10.5.2. *Soit Q une forme quadratique sur un K-espace vectoriel V de dimension finie. Alors il existe des formes linéaires $\phi_1, \ldots, \phi_r \in V^*$ linéairement indépendantes et des scalaires $\alpha_1, \ldots \alpha_r \in K$ non nuls tels que*

$$Q = \sum_{i=1}^{r} \alpha_i \phi_i^2, \quad \textit{i.e.} \quad Q(x) = \sum_{i=1}^{r} \alpha_i \phi_i(x)^2, \ \textit{pour tout } x \in V.$$

De plus l'entier r ne dépend que de la forme quadratique Q.

Définitions. 1.) L'entier r s'appelle le *rang* de la forme quadratique Q, observons que nécessairement $r \leq \dim(V^*) = \dim(V)$.
2.) On dit que la forme quadratique Q est *non dégénérée*, si elle est de rang maximal, i.e. si $r = \dim(V)$.
3) La matrice de Gram de la forme quadratique Q par rapport à une base donnée est par définition la matrice de Gram de la forme bilinéaire symétrique associée.

Remarque. On dit qu'une base $\{v_1, ..., v_n\}$ de V *orthogonalise* la forme quadratique Q si dans cette base on a

$$Q(x) = \sum_{i=1}^{n} \alpha_i x_i^2.$$

Dans ce cas on peut constater directement que la matrice de Gram de la forme bilinéaire symétrique associée β est une matrice diagonale. On a en effet

$$\beta(v_i, v_j) = \frac{1}{4}\left(Q(v_i + v_j) - Q(v_i - v_j)\right) = \frac{1}{4}\left((\alpha_i + \alpha_j) - (\alpha_i + \alpha_j)\right) = 0, \quad \text{si } i \neq j,$$

et

$$\beta(v_i, v_i) = \frac{1}{4}\left(Q(v_i + v_i) - Q(v_i - v_i)\right) = \frac{1}{4}Q(2v_i) = \alpha_i.$$

Cela signifie que la matrice de Gram de β dans la base $\{v_i\}$ est la matrice diagonale dont les coefficients sont $\beta_{ij} = \alpha_i \delta_{ij}$.

Remarque. Une forme quadratique sur K^n est un polynôme homogène de n-variables à coefficients dans le corps K. Orthogonaliser cette forme quadratique revient à faire un changement de variables qui l'exprime comme somme pondérée de carrés.

10.5.1 Réduction d'une forme quadratique à une somme de carré (méthode de complétion des carrés de Gauss)

La méthode élémentaire suivante, qu'on attribue à Gauss, permet de construire un changement linéaire de variables qui orthogonalise une forme quadratique donnée. Soit $Q(x) = \sum_{i,j=1}^{n} b_{ij} x_i x_j$ une forme quadratique non nulle à n variables. Plusieurs cas peuvent se présenter :

(i) Si Q contient le terme x_1^2, alors cette forme quadratique peut s'écrire sous la forme

$$Q(x_1, \ldots, x_n) = a\left(x_1^2 + 2\sum_{i=2}^{n} b_i x_1 x_i\right) + \widehat{Q}_1(x_2, \ldots, x_n),$$

où $a \neq 0$ et $\widehat{Q}_1$ est une forme quadratique en $(n-1)$ variables (qui ne contient pas la variable x_1). L'idée est alors d'ajouter le terme $a\left(\sum_{i=2}^{n} b_i x_i\right)^2$ pour compléter le carré de la partie de Q qui contient x_1, puis de soustraire ce terme (pour conserver l'égalité). On écrit donc

$$\begin{aligned} Q(x_1, \ldots, x_n) &= a\left(x_1 + \sum_{i=2}^{n} b_i x_i\right)^2 - a\left(\sum_{i=2}^{n} b_i x_i\right)^2 + \widehat{Q}_1(x_2, \ldots, x_n) \\ &= a\left(x_1 + \sum_{i=2}^{n} b_i x_i\right)^2 + \widehat{Q}_2(x_2, \ldots, x_n), \end{aligned}$$

où $\widehat{Q}_2$ est la forme quadratique en $(n-1)$ variables définie par

$$\widehat{Q}_2(x_2, \ldots, x_n) = \widehat{Q}_1(x_2, \ldots, x_n) - a\left(\sum_{i=2}^{n} b_i x_i\right)^2$$

(ii) Si Q ne contient pas le terme x_1^2, mais qu'il contient un terme x_j^2 avec $j \geq 2$, alors on procède comme dans le cas (i) mais avec le terme x_j .

(iii) Si Q ne contient aucun terme carré, alors c'est une somme de terme mixtes $x_i x_j$. Dans ce cas peut utiliser l'identité

$$x_i x_j = \frac{1}{4}\left((x_i + x_j)^2 - (x_i - x_j)^2\right),$$

qui nous ramène au cas précédent.

On itère le procédé jusqu'à ce que la forme $Q(x)$ apparaisse comme somme de carrés. Voyons quelques exemples concrets.

Exemples.
1. La forme quadratique sur $\mathbb{R}^2$ définie par

$$Q_1(x_1, x_2) = 2x_1^2 - 5x_2^2 + 4x_1x_2$$

peut se réduire ainsi

$$\begin{aligned} Q_1(x_1, x_2) &= 2x_1^2 - 5x_2^2 + 4x_1x_2 \\ &= 2(x_1^2 + 2x_1x_2 + x_2^2) - 2x_2^2 - 5x_2^2 \\ &= 2(x_1 + x_2)^2 - 7x_2^2. \end{aligned}$$

2. La forme quadratique

$$Q_3(x, y, z) = x^2 + 2xy + 10y^2 - 6yz + 6z^2$$

peut s'écrire de la façon suivante comme somme de carrés

$$Q_3 = (x + y)^2 + (3y - z)^2 + 5z^2$$

Les étapes pour cet exemple sont :

$$\begin{aligned} Q_3(x, y, z) &= x^2 + 2xy + 10y^2 - 6yz + 6z^2 \\ &= (x + y)^2 + 9y^2 - 6yz + 6z^2 \\ &= (x + y)^2 + (3y - z)^2 + 5z^2 \end{aligned}$$

3. La forme quadratique sur $\mathbb{R}^3$ définie par

$$Q_2(x, y, z) = 6x^2 + 12xy - 12xz + 7y^2 - 8yz + 10z^2$$

peut se réduire ainsi

$$\begin{aligned} Q_2(x, y, z) &= 6(x + y - z)^2 - 6(y - z)^2 + 7y^2 - 8yz + 10z^2 \\ &= 6(x + y - z)^2 - 6(y^2 - 2yz + z^2) + 7y^2 - 8yz + 10z^2 \\ &= 6(x + y - z)^2 + y^2 + 4yz + 4z^2 \\ &= 6(x + y - z)^2 + (y + 2z)^2. \end{aligned}$$

Notons que cette méthode ne donne pas une façon unique d'écrire une forme quadratique comme somme de carrés, car elle dépend de l'ordre donné aux variables.