PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 15 (1995), pp. 469–492

UNE INTRODUCTION À LA THÉORIE GÉNÉRALE DE L'APPROXIMATION QUADRATIQUE D'UNE APPLICATION LINÉAIRE

PAR

JEAN-JACQUES TÉCHENÉ (PAU, FRANCE)

Abstract. The linear statistical inference constitutes one of the more usual frameworks for quadratic approximation of linear mappings. Generally, the set of observations is a Euclidean space $\boldsymbol{R}^d$, and the selected risk function is the trace of expectation of a quadratic loss function. In fact, most common notions which one can manipulate there (Gauss–Markov's estimation for example) are totally independent of all but linear structures on the space E of the observations.

We would like to introduce here the general problem of the quadratic approximation of a linear mapping of a finite-dimensional vector space E into another finite-dimensional vector space F, including that of the admissibility of solutions, without any other hypothesis other than the spaces E and F each being placed in separate duality with another vector space. We show how the positivity of the considered operators and the provision of a Euclidean structure on the image space F is sometimes necessary to assure the non-emptiness of the set of solutions. We would like to end by giving, within this general framework, a proof of an extension of a fundamental L. R. LaMotte [5] theorem based on the Hahn–Banach theorem.

1. DUALITÉ ET CONVEXITÉ EN DIMENSION FINIE

Notations. Etant donnés une application T d'un ensemble E dans un ensemble F et une application S de F dans un autre ensemble G, l'application composée de T par S qui à tout u de E associe l'élément $S(Tu)$ de F sera notée $S \circ T$ et l'image $T(E)$ de T, $\operatorname{Im} T$. Lorsque E et F sont deux espaces vectoriels sur un même corps (ici $\boldsymbol{R}$), nous noterons $\mathscr{L}(E, F)$ [resp. $\mathscr{L}(E)$] l'espace des applications linéaires de E dans F [resp. E]. Pour chaque $T \in \mathscr{L}(E, F)$, nous désignerons par $\operatorname{Ker} T$ le noyau $T^{-1}(\{0\})$ de T.

1.1. Sous-espaces affines d'un espace vectoriel. Nous nommerons *sous-espace affine* d'un espace vectoriel (ici *réel*) E l'image L par une translation d'un sous-espace vectoriel unique Z de E, appelé *direction* de L. Si v est le vecteur de cette translation, on a donc $L = \{v\} + Z$, ce que l'on note plus simplement $L = v + Z$. Cette écriture ne dépend pas du choix de v dans L. On sait que deux sous-espaces affines de E, $L = v + Z$ et $L' = v' + Z'$, de directions respectives Z et Z', ont un point commun si, et seulement si, $v - v' \in Z + Z'$, et que leur intersection est alors un sous-espace affine de E de direction $Z \cap Z'$. Soulignons en outre la propriété caractéristique suivante: pour qu'un sous-ensemble non vide L de E soit un sous-espace affine de E, il faut et il suffit que, quels que soient les points distincts u et v de L, la droite joignant u et v soit contenue dans L, autrement dit que

$$(\forall \lambda \in K) \ \big(\forall (u, v) \in L^2\big) \ \lambda u + (1 - \lambda) v \in L.$$

1.2. Formes bilinéaires et dualité. Soient (E, E') un couple d'espaces vectoriels réels et φ une forme bilinéaire sur $E \times E'$. Nous dirons alors que E et E' *sont en dualité relativement à* φ ou que φ *est une dualité sur* $E \times E'$. Lorsque $E' = E$, nous dirons plus simplement que φ *est une dualité sur* E. Rappelons que toute forme bilinéaire φ sur E se décompose d'une manière unique en $\varphi = \varphi_1 + \varphi_2$, où φ_1 est symétrique et φ_2 antisymétrique. Toute application q de E dans $\boldsymbol{R}$ telle qu'il existe une forme bilinéaire symétrique φ sur E vérifiant, pour chaque $x \in E$, $q(x) = \varphi(x, x)$, est appelée *forme quadratique sur* E. La dualité ψ sur E définie par

$$\psi(x, y) = \tfrac{1}{2}[q(x+y) - q(x) - q(y)]$$

est nommée *forme polaire de* q, et l'identité

$$2\psi(x, y) = q(x+y) - q(x) - q(y),$$

identité de polarisation.

Nous dirons que deux espaces vectoriels réels E et E' sont en *dualité séparante* relativement à une forme bilinéaire φ (sur $E \times E'$) si φ est non dégénérée, autrement dit si, pour tout x non nul de E, il existe $y \in E'$ tel que $\varphi(x, y) \neq 0$ et si, pour tout y non nul de E', il existe $x \in E$ tel que $\varphi(x, y) \neq 0$. Pour tout espace vectoriel réel E, de dual E^* (espace des formes linéaires sur E), le couple (E^*, E) est clairement en dualité séparante relativement à *la forme bilinéaire canonique* sur $E^* \times E$, notée $\langle \cdot, \cdot \rangle$, qui à tout (y^*, x) de $E^* \times E$ associe le réel $\langle y^*, x \rangle = y^*(x)$. Evidemment, un espace vectoriel réel de *dimension finie* ne peut être en dualité séparante qu'avec un espace vectoriel E' de *même dimension* et E' est alors canoniquement isomorphe à E^*. Dans ce cas, nous dirons que (E, E') est un *système dual* (sans autre précision) et noterons, si aucune confusion n'est à craindre, par $\langle x, y \rangle$ la valeur en (x, y) de la forme bilinéaire φ sur $E \times E'$, et donc logiquement par $Z^{\perp}$ [resp. $Z'^{\perp}$] l'orthogonal dans E' [resp. dans E] de toute partie Z [resp. Z'] de E [resp. E']. Lorsque E' est l'espace E^*, il est entendu que la forme bilinéaire alors mise en jeu est la forme bilinéaire canonique sur $E \times E^*$.

Soient (E, E') et (F, F') deux systèmes duaux et ψ une forme bilinéaire sur $E \times F$. Pour tout $x \in E$, l'application $\psi(x, \cdot)$: $y \to \psi(x, y)$ est une forme linéaire sur F et donc il existe un élément unique $z \in F'$ tel que $\psi(x, y) = \langle x, z \rangle$ pour tout $y \in F$. Il est immédiat que M: $x \to z$ est l'unique application linéaire de E

dans F' telle que, pour chaque (x, y) de $E \times F$, $\psi(x, y) = \langle y, Mx \rangle$ et que l'application $\psi \to M$ est un isomorphisme de l'espace vectoriel $\mathscr{B}(E \times F)$ des formes bilinéaires sur $E \times F$ sur l'espace vectoriel $\mathscr{L}(E, F')$. Nous identifierons l'espace $\mathscr{B}(E, F)$ à l'espace $\mathscr{L}(E, F')$ et chaque forme bilinéaire ψ sur $E \times F'$ à son application linéaire associée M.

Pour tout système dual (E, E'), cette identification nous amène à appeler *tenseur sur* E toute application linéaire M de E dans E', à dire que M est *symétrique* si nous avons $\langle x, My \rangle = \langle y, Mx \rangle$ pour tout $(x, y) \in E^2$, *positif* [resp. *strictement positif*] sur un sous-espace vectoriel Z de E si $\langle x, Mx \rangle \geqslant 0$ pour tout $x \in Z$ [resp. > 0 pour tout x de $Z \backslash \{0\}$] et à confondre légitimement les notions de forme quadratique et de tenseur symétrique sur E. Nous définirions de même un tenseur sur E'. Nous désignerons respectivement par $\mathscr{L}_s(E, E')$ et $\mathscr{L}_s^+(E, E')$ l'ensemble des tenseurs symétriques sur E et l'ensemble des tenseurs symétriques et positifs sur E.

Considérons maintenant deux systèmes duaux (E, E') et (F, F') (nous noterons indistinctement $\langle \cdot, \cdot \rangle$ les deux dualités associées) et T une application linéaire de E dans F. L'application φ de $E \times F'$ dans $\boldsymbol{R}$ qui à tout couple (x, y) associe le réel $\langle Tx, y \rangle$ est une forme bilinéaire sur $E \times F'$, de transposée ${}^t\varphi$: $(y, x) \to \langle Tx, y \rangle$. Il existe ainsi une application linéaire et une seule de F' dans E', notée ${}^tT$ et appelée *transposée* de T, telle que, pour chaque $(x, y) \in E \times F'$, $\langle Tx, y \rangle = \langle x, {}^tTy \rangle$. Quand $E' = E^*$ et $F' = F^*$, la transposée de T n'est autre que l'application qui à tout $y^* \in F^*$ associe l'élément $y^* \circ T$ de E^*. Nous supposons connues les propriétés fondamentales de la transposition, notamment *le théorème de Farkas*: pour toutes parties C et D de E et F respectivement, nous avons

$$[T(C)]^\perp = ({}^tT)^{-1}(C^\perp) \quad \text{et} \quad {}^tT(D^\perp) \subset [T^{-1}(D)]^\perp,$$

l'égalité ayant lieu, en particulier, si D est un sous-espace vectoriel de F (et donc $\operatorname{Im} {}^tT = (\operatorname{Ker} T)^\perp$).

1.3. Produits tensoriels en dimension finie. Considérons un système dual (E, E'), un espace vectoriel réel F et un couple (Z, W) de sous-espaces vectoriels de E' et F respectivement. Pour tout élément (u, y) de $E' \times F$, soient $u \otimes y$ l'application linéaire élémentaire $x \to \langle x, u \rangle y$ de E dans F et $Z \otimes W$ le sous-espace vectoriel de $\mathscr{L}(E, F)$ engendré par $\{u \otimes y$: $(u, y) \in Z \times W\}$. L'application φ qui à (u, y) associe $u \otimes y$ est une application bilinéaire de $Z \times W$ dans $Z \otimes W$ et $(Z \otimes W, \varphi)$ est un couple *universel* dans le sens où $\varphi(Z \times W)$ engendre linéairement $Z \otimes W$ et où, pour toutes bases $\{u_i$: $i \in I\}$ de Z et $\{y_j$: $j \in I\}$ de W, $\{\varphi(u_i, y_j)$: $(i, j) \in I \times J\}$ est une base de l'espace $Z \otimes W$ (vérification laissée au lecteur). Pour tout espace vectoriel réel H et toute application bilinéaire ψ de $Z \times W$ dans H, il existe donc une seule application linéaire T de $Z \otimes W$ dans H telle que $\psi = T \circ \varphi$, d'où il résulte que $Z \otimes W$ apparaît comme *un* produit tensoriel de Z par W (cf. [2] par exemple). C'est pourquoi, nous nommerons *produit tensoriel de Z par W*, et noterons $Z\otimes W$, le sous-espace vectoriel de $\mathscr{L}(E, F)$ engendré par l'ensemble $\{u\otimes y\colon (u, y)\in Z\times W\}$. Les éléments de $Z\otimes W$ de la forme $u\otimes y$ seront dits *décomposables*.

Etant donnés quatre systèmes duaux (E, E'), (F, F'), (G, G') et (H, H'), les règles opératoires de ces produits tensoriels (laissées à la diligence du lecteur) sont les suivantes:

(i) Quels que soient les sous-espaces vectoriels Z et Z' de E', W et W' de F,

$$(Z\cap Z')\otimes(W\cap W') = (Z\otimes W)\cap(Z'\otimes W').$$

(ii) Quel que soit $(u, y)\in E'\times F$, ${}^t(u\otimes y) = y\otimes u$.

(iii) Quels que soient $(A, B)\in\mathscr{L}(E', H')\times\mathscr{L}(F, G)$ et $(u, y)\in E'\times F$,

$$Au\otimes By = B\circ(u\otimes y)\circ{}^tA.$$

PROPOSITION 1.1. *Etant donnés un système dual (E, E') de dimension finie et un espace vectoriel réel F, on a $\mathscr{L}(E, F) = E'\otimes F$ et, plus généralement, pour tout sous-ensemble L de E et tout sous-espace vectoriel W de F:*

$$\{T\in\mathscr{L}(E, F)\colon L\subset \operatorname{Ker} T \text{ et } \operatorname{Im} T\subset W\} = L^{\perp}\otimes W.$$

Preuve. Soient T une application linéaire de E dans F, $\{f_1, \ldots, f_q\}$ une base de $\operatorname{Im} T$ et, pour tout $j\in\{1, \ldots, q\}$, u_j^* la forme linéaire sur E qui à chaque $x\in E$ associe la composante du vecteur Tx sur $\mathbf{R}f_j$. Il suit clairement

$$(\forall x\in E)\ Tx = \sum_{j=1}^{q} u_j^*(x)f_j = \sum_{j=1}^{q}\langle x, u_j\rangle f_j,$$

où u_j est l'unique élément de E' tel que $u_j^*(x) = \langle x, u_j\rangle$ pour tout $x\in E$; il en découle que $\mathscr{L}(E, F) = E'\otimes F$, car E étant de dimension finie, l'ensemble des applications linéaires de rang fini de E dans F coïncide avec $\mathscr{L}(E, F)$.

Maintenant, pour chaque $(u, y)\in L^{\perp}\times W$, il est immédiat que $u\otimes y$ appartient à l'ensemble $\Delta = \{T\in\mathscr{L}(E, F)\colon L\subset\operatorname{Ker} T \text{ et } \operatorname{Im} T\subset W\}$ et donc que $L^{\perp}\otimes W\subset\Delta$. Réciproquement, soit $T=\sum_{j=1}^{q}u_j\otimes f_j$ un élément de Δ, où $\{f_1, \ldots, f_q\}$ est une famille libre de F (la preuve de la relation $\mathscr{L}(E, F) = E'\otimes F$ montre clairement que toute application linéaire de E dans F peut s'écrire ainsi). L'indépendance linéaire des vecteurs $f_1, \ldots, f_q$ entraîne alors que, quel que soit $x\in L$, la relation $Tx=0$ équivaut à $\langle x, u_j\rangle = 0$ pour tout $j\in\{1, \ldots, q\}$; on en déduit que:

$$L\subset\operatorname{Ker} T\Leftrightarrow(\forall j\in\{1, \ldots, q\})\ u_j\in L^{\perp}$$

et la propriété cherchée en résulte immédiatement. ∎

Exprimant que chaque vecteur f_k d'une base de $\operatorname{Im} T$ s'écrit comme combinaison linéaire unique de vecteurs de toute base de W, on obtient plus précisément le

COROLLAIRE 1.2. *Pour toute base* $\{y_j\colon j\in J\}$ *de* W, *chaque élément* T *de* $Z\otimes W$ *s'écrit d'une manière et d'une seule sous la forme* $T=\sum_{j\in J} u_j\otimes y_j$, *où* $u_j\in Z$ *pour tout* $j\in J$.

Considérons maintenant trois systèmes duaux (E,E'), (F,F') et (G,G'). Compte tenu de la proposition 1.1 précédente, le théorème de factorisation des applications linéaires ([2], p. 13) conduit alors aussitôt aux résultats suivants:

COROLLAIRE 1.3. *Soit* Z *un sous-espace vectoriel de* E'. *Pour toute application linéaire* R *de* E *dans* G *telle que* $\operatorname{Ker} R = Z^{\perp}$, *on a* $Z\otimes F=\{S\circ R\colon S\in\mathscr{L}(G,F)\}$.

COROLLAIRE 1.4. *Soient* T *et* A *deux applications linéaires de* G *dans* E *et de* G *dans* F *telles que* $\operatorname{Ker} T\subset \operatorname{Ker} A$. *Alors l'ensemble* $\Gamma=\{X\in\mathscr{L}(E,F)\colon X\circ T=A\}$ *est un sous-espace affine de* $\mathscr{L}(E,F)$ *de direction* $\Delta=(\operatorname{Ker}{}^tT)\otimes F$.

En effet, Δ est l'ensemble des X de $\mathscr{L}(E,F)$ tels que $\operatorname{Im} T\subset \operatorname{Ker} X$, d'où le résultat d'après le corollaire 1.3, car $(\operatorname{Im} T)^{\perp}=\operatorname{Ker}{}^tT$. Enfin, la démonstration de la proposition suivante n'est qu'une simple vérification laissée au lecteur:

PROPOSITION 1.5. *Soient* (E,E') *et* (F,F') *deux systèmes duaux. Les espaces vectoriels* $\mathscr{L}(E,F)$ *et* $\mathscr{L}(E',F')$ *sont alors en dualité séparante relativement à la forme bilinéaire* Ψ *définie* [*sur une famille génératrice de* $\mathscr{L}(E,F)\times\mathscr{L}(E',F')$] *par*

$$\Psi(u\otimes y,\, x\otimes v)=\langle x,u\rangle\langle y,v\rangle$$

pour tous couples (u,x) *et* (v,y) *respectivement de* $E'\times E$ *et* $F'\times F$. *De plus, pour chaque couple* (Z,W) *de sous-espaces vectoriels de* E' *et* F *respectivement, l'orthogonal de* $Z\otimes W$ *relativement à* Ψ *dans* $\mathscr{L}(E',F')$ *est* $Z^{\perp}\otimes F'+E\otimes W^{\perp}$.

Les éléments des espaces $\mathscr{L}(E,F)$ et $\mathscr{L}(E',F')$ étant notés par les lettres latines majuscules, ou par un produit tensoriel de vecteurs s'il s'agit d'éléments décomposables, nous noterons, sans risque de confusion possible, $\langle A,B\rangle$ la valeur en (A,B) de la forme bilinéaire Ψ.

Etant donné un système dual (E,E'), la relation $\mathscr{L}(E)=E'\otimes E$ montre qu'il existe une forme linéaire et une seule sur $\mathscr{L}(E)$, appelée *trace* et notée tr, telle que:

$$(\forall (u,x)\in E'\times E)\ \operatorname{tr}(u\otimes x)=\langle x,u\rangle.$$

Comme, pour tout (u,x) et (v,y) de $E'\times E$, $(u\otimes x)\circ(v\otimes y)=\langle y,u\rangle\, u\otimes y$, il est clair que l'on a

$$\operatorname{tr}[(u\otimes x)\circ(v\otimes y)]=\operatorname{tr}[(v\otimes y)\circ(u\otimes x)],$$

dont il résulte aussitôt que, d'une part, pour tout couple (A,B) d'endomorphismes de E, on a $\operatorname{tr}(A\circ B)=\operatorname{tr}(B\circ A)$ et $\operatorname{tr}{}^tA=\operatorname{tr} A$, et d'autre part, que si $F'=F$, la dualité Ψ sur $\mathscr{L}(E,F)\times\mathscr{L}(E',F)$ définie par la proposition 1.5 n'est autre que la forme bilinéaire telle que:

$$(\forall (S, T) \in \mathscr{L}(E, F) \times \mathscr{L}(E', F)) \ \langle S, T \rangle = \mathrm{tr}(S \circ {}^t T) = \mathrm{tr}(T \circ {}^t S).$$

Enfin (propriété bien connue laissée au lecteur), si E est euclidien, la restriction de la trace à l'ensemble $\mathscr{L}_s^+(E)$ des tenseurs symétriques et positifs sur E est injective.

1.4. La génération des cônes convexes. Rappellons que l'on dit qu'une partie C d'un espace vectoriel réel E est un *cône* lorsque C est stable pour toutes les homothéties vectorielles de rapport strictement positif. Si l'origine 0 appartient à C, le cône C est dit *pointé*. Un cône convexe *saillant* est un cône convexe pointé qui ne contient aucune droite vectorielle. On sait que, pour qu'une partie C de E soit un cône convexe, il faut et il suffit que l'on ait $C+C \subset C$ et $\lambda C \subset C$ pour tout $\lambda > 0$, et que, si C est un cône convexe, l'ensemble $C-C$ est le sous-espace vectoriel de E engendré par C. Le cône convexe engendré par une partie A de E est donc l'ensemble des combinaisons linéaires $\sum_{i \in I} \lambda_i x_i$, où $\{x_i: i \in I\}$ est une famille finie non vide quelconque d'éléments de A, et où $\lambda_i > 0$ pour tout $i \in I$.

Par ailleurs, si C est un cône convexe pointé de E, le plus grand sous-espace vectoriel de E contenu dans C est l'ensemble (non vide) $V = C \cap (-C)$: en effet, $\lambda V = V$ pour tout réel $\lambda \neq 0$, et de plus

$$V + V \subset (C+C) \cap [-(C+C)] \subset C \cap (-C) = V,$$

ce qui montre que V est un sous-espace vectoriel de E, qui contient à l'évidence tout sous-espace vectoriel de E inclus dans C. Il s'ensuit qu'un cône convexe C est saillant si, et seulement si, $C \cap (-C) = \{0\}$.

Nous appellerons *base* d'un cône convexe pointé C toute partie convexe B de E, s'il en existe, ne contenant pas l'origine 0 et telle que $C = \mathbf{R}^+ B = \{\lambda x: \lambda \in \mathbf{R}^+, x \in B\}$. Nous laissons au lecteur le soin de vérifier que, si un cône convexe pointé C admet une base B, alors C n'est autre que le cône convexe engendré par B au sens où C est l'ensemble des combinaisons linéaires positives $\sum_{i \in I} \lambda_i x_i$, où $\{x_i: i \in I\}$ est une famille finie non vide quelconque d'éléments de B et où $\lambda_i \geqslant 0$ pour tout $i \in I$. Le résultat suivant a été démontré par L. R. LaMotte dans un cadre plus restreint ([5], p. 248):

PROPOSITION 1.6. *Soit E un espace vectoriel réel topologique séparé de dimension finie. Alors*:

(i) *Tout cône convexe fermé saillant de E admet une base compacte.*

(ii) *Pour tout cône convexe fermé C de E, il existe une partie convexe compacte A ne contenant pas l'origine* 0, *appelée socle de C, telle que*

$$C = \mathbf{R}^+ A + V \quad \text{et} \quad \mathbf{R}^+ A \cap V = \{0\},$$

où V est le plus grand sous-espace vectoriel de E contenu dans C.

Preuve. (i) On sait que l'unique topologie vectorielle séparée sur un espace vectoriel réel de dimension finie E (dite canonique) peut être définie par une norme arbitraire $\|\cdot\|$ sur E (cf. [1]). Soit C un cône convexe fermé saillant de E et considérons la sphère-unité S de E. Alors l'ensemble $A = C \cap S$ est un fermé de S et donc, comme S, un compact de E. Son enveloppe convexe $\mathscr{C}(A)$ est donc aussi compacte dans E. De plus, l'origine 0 n'appartient pas à $\mathscr{C}(A)$. En effet, si l'origine appartenait à $\mathscr{C}(A)$, il existerait une combinaison convexe $\sum_{i=1}^{n} \lambda_i x_i$ d'au moins deux points $x_1, x_2, \ldots, x_n$ de A (car $0 \notin A$) qui serait nulle, et on aurait (en supposant par exemple que $\lambda_1 \neq 0$)

$$x_1 = -(\sum_{i=2}^{n} \lambda_i x_i)/\lambda_1.$$

Autrement dit, comme $x_1 \in A$ et que $0 \notin A$, $\boldsymbol{R}x_1 \neq \{0\}$ et $-x_1$ appartiendrait également au sous-ensemble A, d'où l'on déduirait que $\boldsymbol{R}x_1 \subset C$, car C est un cône convexe de E (et donc $C+C \subset C$ et $\lambda C \subset C$). Il en résulterait que C, cône convexe saillant par hypothèse, contiendrait un sous-espace vectoriel de E autre que $\{0\}$, ce qui est impossible car $C \cap (-C) = \{0\}$. De plus, pour tout x non nul de C, on a $\|x\|^{-1} x \in A$ et donc C est contenu dans le cône convexe fermé engendré par $\mathscr{C}(A)$. Comme C est un cône convexe contenant $\mathscr{C}(A)$, on a l'inclusion opposée, et par suite $\mathscr{C}(A)$ est bien une base compacte de C.

(ii) Soit C un cône convexe fermé de E. Remarquous que, pour tout supplémentaire W de $V = C \cap (-C)$ dans E, on a $C = C \cap W + V$. En effet, pour tout x de C, on peut écrire $x = v + w$ avec $v \in V$ et $w \in W$; comme alors $w = x - v \in C + V$ et que $C + V \subset C + C \subset C$, il vient $w \in C \cap W$ et donc $C \subset C \cap W + V$. L'inclusion opposée est triviale. L'application de (i) au cône convexe fermé saillant $C \cap W$ donne alors immédiatement (ii). ■

A noter que cette proposition est généralement fausse en dimension infinie car sa démonstration repose sur la compacité de la boule-unité.

2. ORTHOGONALITÉ ET PROJECTION RELATIVEMENT À UN TENSEUR SYMÉTRIQUE

2.1. L'orthogonalité relativement à un tenseur symétrique. Nous considèrons une fois pour toutes, dans ce paragraphe, un système dual (E, E'). Etant donnés un tenseur symétrique M sur E et un sous-ensemble A de E, il est immédiat que $[M(A)]^{\perp}$ est l'orthogonal de A dans E relativement au tenseur M (c'est-à-dire relativement à la forme bilinéaire sur E qui à tout (x, y) associe $\langle x, My \rangle$). Comme M est symétrique, alors en vertu du théorème de Farkas (§ 1.1), $[M(A)]^{\perp}$ est identique à $M^{-1}(A^{\perp})$. Afin de distinguer cette orthogonalité de celle relative à la dualité $\langle \cdot, \cdot \rangle$ sur $E \times E'$, nous dirons que $[M(A)]^{\perp}$ est le sous-espace (vectoriel) *M-orthogonal* de A dans E; nous le noterons $A^{\perp_M}$. Lorsque M est un tenseur symétrique sur E', il est clair que l'on peut définir de la même manière le sous-espace M-orthogonal d'un sous-ensemble A' de E'.

PROPOSITION 2.1. *Soient M un tenseur symétrique sur E et Z un sous-espace vectoriel de E. Si M est positif sur Z, les relations suivantes sont équivalentes et vérifiées*:

(i) $Z^\perp \cap M(Z) = \{0\}$,

(ii) $E = Z + Z^{\perp M}$,

(iii) $\operatorname{Im} M \subset Z^\perp + M(Z)$.

Preuve. Soit M_Z la restriction de M au sous-espace Z. Comme M est positif sur Z, l'inégalité de Cauchy–Schwarz entraîne aussitôt que

$$\operatorname{Ker} M_Z = \{u \in Z \colon \langle u, Mu \rangle = 0\}.$$

Il suit que tout $u = Mz \in Z^\perp \cap M(Z)$, avec $z \in Z$, est tel que $0 = \langle z, u \rangle = \langle z, Mz \rangle$, donc nécessairement nul, car alors $z \in \operatorname{Ker} M_Z$, ce qui prouve (i).

L'équivalence entre (i) et (ii) s'obtient immédiatement par passage aux orthogonaux, étant donné que $Z^{\perp M} = [M(Z)]^\perp$. Par ailleurs, la relation $M^{-1}[M(Z)] = Z + \operatorname{Ker} M$ implique que:

$$Z^\perp \cap M(Z) = \{0\} \Leftrightarrow Z \cap M^{-1}(Z^\perp) \subset \operatorname{Ker} M,$$

où $M^{-1}(Z^\perp) = [M(Z)]^\perp$. L'équivalence entre (i) et (iii) s'en déduit par passage aux orthogonaux. ■

On notera que la relation (iii) entraîne que $Z^\perp + \operatorname{Im} M \subset Z^\perp + M(Z)$, dont on déduit aussitôt que $Z^\perp + \operatorname{Im} M = Z^\perp + M(Z)$, et par suite que $Z \cap \operatorname{Ker} M = Z \cap Z^{\perp M}$. D'où également la

PROPOSITION 2.2. *Pour tout sous-espace vectoriel Z de E et tout tenseur symétrique M sur E, positif sur Z, on a*:

(i) $Z^\perp + \operatorname{Im} M = Z^\perp \oplus M(Z)$,

(ii) $Z \cap \operatorname{Ker} M = Z \cap Z^{\perp M}$,

(iii) $Z^\perp + \operatorname{Im} M = E$, *si M est strictement positif sur Z.*

2.2. Projection relativement à un tenseur symétrique. Soient Z un sous-espace vectoriel de E, M un tenseur symétrique sur E, R une application d'un ensemble $\mathscr{E}$ dans E' et D une application de $\mathscr{E}$ dans E. Pour chaque $v \in \mathscr{E}$, considérons le programme quadratique:

$$(\boldsymbol{P}_v) \qquad \operatorname{Min} \langle u, Mu - 2Rv \rangle \colon u \in Dv + Z.$$

THÉORÈME 2.3. *Pour que le programme $(\boldsymbol{P}_v)$ admette au moins une solution pour chaque $v \in \mathscr{E}$, il faut et il suffit que M soit positif sur Z et que* $\operatorname{Im}(M \circ D - R)$ *soit contenu dans $Z^\perp + \operatorname{Im} M$. L'ensemble des solutions de $(\boldsymbol{P}_v)$ est alors un sous-espace affine de E de direction $Z \cap \operatorname{Ker} M$.*

Preuve. Pour tout $v \in \mathscr{E}$, le gradient en $a \in E$ de l'application q: $u \to \langle u, Mu-2Rv \rangle$ est $\nabla q(a) = 2(Ma-Rv)$. Donc, pour que $\bar{v}$ est une solution de (P_v), il faut et il suffit qu'il existe $x \in Z$ tel que l'on ait à la fois $\bar{v} = x+Dv$ et $M\bar{v}-Rv \in Z^{\perp}$. La condition

$$\mathrm{Im}(M \circ D - R) \subset Z^{\perp} + M(Z)$$

est donc nécessaire pour que (P_v) admette une solution pour *chaque* $\bar{v} \in \mathscr{E}$. Par ailleurs, l'identité de polarisation (§ 1.1) appliquée au tenseur M conduit, pour tout $(u, z) \in E \times E$, à l'identité:

(1) $\langle u+z, M(u+z)-2Rv \rangle - \langle u, Mu-2Rv \rangle = 2\langle z, Mu-Rv \rangle + \langle z, Mz \rangle$.

Or, si $\bar{v}$ est une solution de (P_v), nous avons à la fois, quel que soit $z \in Z$:

$$\langle \bar{v}, M\bar{v}-2Rv \rangle \leqslant \langle \bar{v}+z, M(\bar{v}+z)-2Rv \rangle \quad \text{et} \quad \langle z, M\bar{v}-Rv \rangle = 0,$$

d'où l'on tire que M est positif sur Z, et par suite, d'après la proposition 2.2, que:

$$Z^{\perp} + M(Z) = Z^{\perp} + \mathrm{Im}\, M.$$

Réciproquement, si $\mathrm{Im}(M \circ D - R) \subset Z^{\perp} + M(Z)$, alors, pour tout $v \in \mathscr{E}$, il existe $x \in Z$ tel que l'on ait $M(Dv-x)-Rv \in Z^{\perp}$; l'identité (1), avec $u = Dv-x$, donne alors:

$$(\forall z \in Z) \ \langle u+z, M(u+z)-2Rv \rangle - \langle u, Mu-2Rv \rangle = \langle z, Mz \rangle,$$

ce qui prouve aussitôt, si de plus M est positif sur Z, que $Dv-x$ est une solution du programme (P_v). L'ensemble des solutions est alors l'ensemble des $\bar{v} \in Dv+Z$ tels que l'on ait $M\bar{v} \in Z^{\perp}+Rv$. C'est clairement (cf. § 1.1) un sous-espace affine de E, contenu dans le sous-espace affine $Dv+Z$, de direction (proposition 2.2):

$$Z \cap M^{-1}(Z^{\perp}) = Z \cap Z^{\perp M} = Z \cap \mathrm{Ker}\, M,$$

ce qu'il fallait montrer. ∎

Il en résulte immédiatement que le programme (P_v) admet une solution unique pour *chaque* v de $\mathscr{E}$ si, et seulement si, le tenseur M est strictement positif sur Z, ou ce qui est équivalent si, et seulement si, $Z^{\perp M}$ est un supplémentaire de Z dans E. De plus, on notera que si D est une application constante d'image $\{d\}$ et si $R = M$ (avec $\mathscr{E} = E$), alors le programme (P_v) est identique au programme:

$$\mathrm{Min}\langle v-u, M(v-u) \rangle \colon u \in d+Z,$$

qui n'est autre que la recherche des "projections" de v sur le sous-espace affine $d+Z$ (indépendant de v) au sens du tenseur M. L'étude d'un tel programme, par la famille des "projecteurs" qui lui sont associés, constitue en particulier le fondement de l'interprétation géométrique de l'estimation de Gauss–Markov (cf. [7], [8] ou [9]).

3. L'APPROXIMATION QUADRATIQUE D'UNE APPLICATION LINÉAIRE

C'est relativement à deux préordres fondamentaux que nous allons rappeler qu'est classiquement envisagée l'approximation quadratique d'une application linéaire.

3.1. Les relations d'ordre de Loewner et de préordre de Schur. Soit (E, E') un système dual. Nous appellons *relation d'ordre de Loewner* sur $\mathscr{L}(E', E)$ et noterons $\preccurlyeq_{\mathrm{L}}$, la relation d'ordre partiel sur $\mathscr{L}(E', E)$ définie par

$$M_1 \preccurlyeq_{\mathrm{L}} M_2 \Leftrightarrow M_2 - M_1 \in \mathscr{L}_s^+(E', E);$$

nous nommerons *préordre de Schur* sur $\mathscr{L}(E)$ et désignerons par $\preccurlyeq_{\mathrm{S}}$ la relation de préordre total sur $\mathscr{L}(E)$ telle que:

$$M_1 \preccurlyeq_{\mathrm{S}} M_2 \Leftrightarrow \mathrm{tr}(M_2 - M_1) \geqslant 0.$$

Lorsque $E' = E$ et que la dualité $\langle \cdot, \cdot \rangle$ est positive sur E, il est immédiat que la trace conserve l'ordre de Loewner, c'est-à-dire que $M_1 \preccurlyeq_{\mathrm{L}} M_2$ entraîne $M_1 \preccurlyeq_{\mathrm{S}} M_2$ pour tout couple (M_1, M_2) d'éléments de $\mathscr{L}(E)$. La réciproque est vraie. En effet, pour tout $x \in E$, on a $x \otimes x \succcurlyeq_{\mathrm{L}} 0$ et la positivité de la dualité $\langle \cdot, \cdot \rangle$ sur E est une condition *nécessaire* pour que l'on ait $\mathrm{tr}(x \otimes x) = \langle x, x \rangle \geqslant 0$ pour chaque $x \in E$.

3.2. L'approximation au sens de l'ordre de Loewner. Dans toute la suite de cette section, on désigne par (E, E') et (F, F') deux systèmes duaux, par M et N deux tenseurs symétriques sur E' (et à valeurs dans E) et par B une application linéaire de E dans F. Pour chaque application linéaire T de E dans F, nous appellerons *tenseur de l'erreur quadratique associée à T* (erreur commise en "approchant" B par T) le tenseur sur F', noté $\varrho(M, N, T)$, défini par:

$$\varrho(M, N, T) = T \circ M \circ {}^tT + (T-B) \circ N \circ {}^t(T-B).$$

Pour tout sous-espace affine Γ de $\mathscr{L}(E, F)$, on notera $\Gamma_{MN}(B)$ l'ensemble, éventuellement vide, des éléments $\bar{B}$ de Γ tels que $\varrho(M, N, \bar{B}) \preccurlyeq_{\mathrm{L}} \varrho(M, N, C)$ pour tout $C \in \Gamma$, autrement dit l'ensemble des solutions du programme:

$$(\mathscr{P}_1) \qquad (\forall v \in F') \ \mathrm{Min}\langle v, \varrho(M, N, C)v\rangle\colon \ C \in \Gamma.$$

Pour qu'un élément $\bar{B}$ de Γ appartienne à $\Gamma_{MN}(B)$, il faut et il suffit donc que, pour tout $C \in \Gamma$ et tout $v \in F'$, on ait:

$$\begin{aligned}&\langle {}^t\bar{B}v, M \circ {}^t\bar{B}v\rangle + \langle {}^t(\bar{B}-B)v, N \circ {}^t(\bar{B}-B)v\rangle \\ &\qquad\qquad \leqslant \langle {}^tCv, M \circ {}^tCv\rangle + \langle {}^t(C-B)v, N \circ {}^t(C-B)v\rangle,\end{aligned}$$

autrement dit, après un calcul immédiat, que, pour tout $C \in \Gamma$ et tout $v \in F'$:

$$\langle {}^t\bar{B}v, (M+N) \circ {}^t\bar{B}v - 2N \circ {}^tBv\rangle \leqslant \langle {}^tCv, (M+N) \circ {}^tCv - 2N \circ {}^tBv\rangle.$$

Comme par ailleurs, pour chaque $v \in F'$, l'ensemble $\{{}^tCv\colon C \in \Gamma\}$ est un sous-espace affine de E' (vérification immédiate), on en déduit la

PROPOSITION 3.1. *Soit Γ un sous-espace affine de $\mathscr{L}(E, F)$. Pour qu'un élément $\bar{B}$ de Γ appartienne à $\Gamma_{MN}(B)$, il faut et il suffit que, pour tout $v \in F'$, $u = {}^t\bar{B}v$ soit une solution du programme quadratique*:

$$\text{(2)} \qquad \operatorname{Min}\langle x, (M+N)x - 2N \circ {}^tBv\rangle\colon\ x \in \{{}^tCv\colon\ C \in \Gamma\}.$$

Dans les applications, la direction de Γ est le produit tensoriel d'un sous-espace Z de E' par un sous-espace W de F. La direction Z_v du sous-espace affine $L_v = \{{}^tCv\colon C \in \Gamma\}$ de E' est alors le sous-espace vectoriel de E' engendré par la famille $\mathscr{F}_v$ telle que:

$$\mathscr{F}_v = \{y \otimes z(v)\colon (z, y) \in Z \times W\} = \{\langle v, y\rangle z\colon (z, y) \in Z \times W\},$$

et donc $Z_v = Z$ si $W \not\subset (\boldsymbol{R}v)^\perp$, ou $\{0\}$ sinon. De cette propriété, on déduit le

THÉORÈME 3.2. *Soient Z et W deux sous-espaces vectoriels de E' et F respectivement, $\mathscr{E}$ le complémentaire de $W^\perp$ dans F' et Γ un sous-espace affine de $\mathscr{L}(E, F)$ de direction $Z \otimes W$. Alors*:

(i) *Une condition nécessaire et suffisante pour que l'ensemble $\Gamma_{MN}(B)$ ne soit pas vide est que le tenseur $M+N$ soit positif sur Z et que, C désignant un élément arbitraire de C et posant $R = C \circ (M+N) - B \circ N$, on ait*

$$Z \cap \operatorname{Ker}(M+N) \subset R^{-1}(\mathscr{E}^\perp).$$

L'ensemble $\Gamma_{MN}(B)$ est un singleton si, et seulement si, le tenseur $M+N$ est strictement positif sur Z.

(ii) *Lorsque $\Gamma_{MN}(B) \neq \emptyset$, pour qu'un élément $\bar{B}$ de Γ appartienne à $\Gamma_{MN}(B)$, il faut et il suffit que $Z \subset \operatorname{Ker}[\bar{B} \circ (M+N) - B \circ N]$ et $\Gamma_{MN}(B)$ est alors un sous-espace affine de $\mathscr{L}(E, F)$ de direction $[Z \cap \operatorname{Ker}(M+N)] \otimes W$.*

Preuve. (i) Soit v un élément non nul de F'. La direction de $L_v = \{{}^tCv\colon C \in \Gamma\}$ étant $\{0\}$ ou Z selon que v appartient ou n'appartient pas à $W^\perp$, la proposition 3.1 nous renvoie au théorème 2.3 (ou à un prolongement immédiat de ce théorème) dans lequel on remplace M par $M+N$, R par $N \circ {}^tB$ puis D par ${}^tC$: pour que le programme (2) admette, pour chaque $v \in F'$, au moins une solution, il faut et il suffit que $M+N$ soit positif sur Z et que, C désignant un élément arbitraire de Γ, on ait

$$[(M+N) \circ {}^tC - N \circ {}^tB](\mathscr{E}) \subset Z^\perp + \operatorname{Im}(M+N),$$

autrement dit, avec les notations de ce théorème, ${}^tR$ étant linéaire,

$${}^tR[\operatorname{vect}(\mathscr{E})] \subset Z^\perp + \operatorname{Im}(M+N),$$

où $\operatorname{vect}(\mathscr{E})$ désigne le sous-espace vectoriel de F' engendré par $\mathscr{E}$; par passage aux orthogonaux, il revient à écrire (comme $[\operatorname{vect}(\mathscr{E})]^\perp = \mathscr{E}^\perp$, cf. § 1.2) que

$$Z \cap \operatorname{Ker}(M+N) \subset R^{-1}(\mathscr{E}^\perp).$$

La deuxième partie de cette assertion se déduit clairement de la proposition 2.2: si le tenseur $M+N$ est positif sur Z, il est strictement positif sur Z si, et seulement si, sa restriction à Z est injective, ce qui revient à dire que $Z \cap \operatorname{Ker}(M+N) = \{0\}$, et alors le programme (2) admet une et une seule solution.

(ii) Posons $R = \bar{B} \circ (M+N) - B \circ N$, où $\bar{B} \in \Gamma$. Le gradient au point $a \in E'$ de l'application q qui à tout $x \in E'$ associe le réel $\langle x, (M+N)x - 2N \circ {}^t Bv \rangle$ est

$$\nabla q(a) = 2[(M+N)a - N \circ {}^t Bv].$$

Donc, si les conditions d'existence (i) sont remplies, alors, quel que soit $v \in F'$, $u = {}^t \bar{B} v$ est une solution de (2) si, et seulement si, $\operatorname{Im} {}^t R \subset Z^\perp$, d'où la première partie de (ii) par passage aux orthogonaux. Comme l'ensemble des éléments décomposables engendre le produit tensoriel $Z \otimes W$ et que $W \neq \{0\}$, on notera ici que la condition $\operatorname{Im} {}^t R \subset Z^\perp$ est équivalente à $(z \otimes y) \circ {}^t R = 0$ pour tout $(z, y) \in Z \times W$, car $(z \otimes y) \circ {}^t R = Rz \otimes y$.

De plus, d'après le théorème 2.3, C désignant un point arbitraire de Γ, l'ensemble des solutions du programme (pour chaque $v \in F'$):

$$\operatorname{Min} \langle x, (M+N)x - 2N \circ {}^t Bv \rangle: \ x \in {}^t Cv + Z$$

est, s'il n'est pas vide, un sous-espace affine de E de direction $Z \cap \operatorname{Ker}(M+N)$, ce qui prouve, en vertu de la proposition 3.1, la deuxième partie de l'assertion (ii). ∎

Remarque 1. Dans les applications (cf. [5] par exemple), l'orthogonal $Z^\perp$ de Z est l'image d'une application linéaire donnée T d'un espace vectoriel réel G et à valeurs dans E, et

$$\Gamma = \{X \in \mathscr{L}(E, F): \ X \circ T = A\},$$

où A est une application linéaire de G dans F, également fixée, telle que $\operatorname{Ker} T \subset \operatorname{Ker} A$ (cette dernière condition assurant la non vacuité de Γ, voir aussi le corollaire 1.4). Si l'on note par G' un espace vectoriel en dualité séparante avec G, l'ensemble $\Gamma_{MN}(B)$ est clairement l'ensemble des solutions du système d'équations suivant, d'inconnue X dans $\mathscr{L}(E, F)$:

$$X \circ T = A, \qquad X \circ (M+N) \circ {}^t T = B \circ N \circ {}^t T.$$

Remarque 2. On observera que lorsque E' et F' sont les duaux E^* et F^* de E et F respectivement, tous les résultats mis en évidence par le théorème 3.2 précédent ne mettent en jeu que la seule structure d'espace vectoriel de E et de F.

Parmi les nombreuses conséquences du théorème 3.2, nous retiendrons tout de suite le corollaire suivant qui précise l'effet d'une application linéaire sur les éléments de l'ensemble $\Gamma_{MN}(B)$. On désigne par (G, G') un autre système dual, par Z un sous-espace vectoriel de E' et par I_E l'application identique de E:

COROLLAIRE 3.3. *Soient $\Gamma = A + Z \otimes F$ un sous-espace affine de $\mathscr{L}(E, F)$ et T une application linéaire de F dans G. L'ensemble $\Gamma' = \{T \circ C\colon C \in \Gamma\}$ est un sous-espace affine de $\mathscr{L}(E, G)$ inclus dans $\Gamma'' = T \circ A + Z \otimes G$ et de direction $Z \otimes \operatorname{Im} T$. De plus, si $\Gamma_{MN}(B)$ n'est pas vide:*

$$\Gamma'_{MN}(T \circ B) = \{T \circ \bar{B}\colon \bar{B} \in \Gamma_{MN}(B)\} \subset \Gamma''_{MN}(T \circ B).$$

En particulier, si l'on pose $\Lambda = I_E + Z \otimes E$, alors, lorsque $B \in \Gamma$, on a

$$\{B \circ Q\colon Q \in \Lambda_{MN}(I_E)\} \subset \Gamma_{MN}(B),$$

et cette inclusion est une égalité si, et seulement si, B est surjective.

Preuve. Il est clair que l'ensemble $\Gamma' = \{T \circ C\colon C \in \Gamma\}$ est un sous-espace affine de $\mathscr{L}(E, G)$ de direction $\Delta' = \{T \circ U\colon U \in Z \otimes F\}$. Comme $\{u \otimes y\colon (u, y) \in Z \times F\}$ engendre linéairement $Z \otimes F$, Δ' est engendré par la famille $\{T \circ (u \otimes y)\colon (u, y) \in Z \times F\}$, d'où l'on tire que $\Delta' = Z \otimes \operatorname{Im} T$, car on a $T \circ (u \otimes y) = u \otimes Ty$.

Supposons $\Gamma_{MN}(B) \neq \emptyset$. Soit $\bar{B} \in \Gamma_{MN}(B)$ et posons $R = \bar{B} \circ (M+N) - B \circ N$. Comme alors $\bar{B} \in \Gamma$, $Z \subset \operatorname{Ker} R$ et que $\operatorname{Ker} R \subset \operatorname{Ker}(T \circ R)$, on a $T \circ \bar{B} \in \Gamma'$ et $Z \subset \operatorname{Ker}(T \circ R)$. La première partie de cette proposition est alors une conséquence du théorème 3.2, car $T \circ \bar{B}$ appartient simultanément aux sous-espaces affines $\Gamma'_{MN}(T \circ B)$ et $\Gamma''_{MN}(T \circ B)$ de $\mathscr{L}(E, G)$ de directions $[Z \cap \operatorname{Ker}(M+N)] \otimes \operatorname{Im} T$ et $[Z \cap \operatorname{Ker}(M+N)] \otimes G$ respectivement, ces deux ensembles ne coïncidant que si T est surjective. La deuxième partie en découle par un aménagement évident laissé au lecteur. ∎

Enfin, le corollaire suivant étend sensiblement divers résultats publiés notamment dans [5] (pp. 246–247) ou [3] (pp. 15–17):

COROLLAIRE 3.4. *Si les tenseurs M et N sont positifs sur Z et si $\Gamma = B + Z \otimes F$, alors $\Gamma_{MN}(B)$ n'est pas vide et, pour tous C de $B + (Z \cap \operatorname{Ker} N) \otimes F$ et $\bar{B}$ de $\Gamma_{MN}(B)$, on a:*

(i) $\varrho(M, N, \bar{B}) = \bar{B} \circ M \circ {}^tC = C \circ N \circ {}^tC - \bar{B} \circ N \circ {}^tC$;

(ii) $Z \cap \operatorname{Ker} N \subset \operatorname{Ker}(\bar{B} \circ M)$;

(iii) $\bar{B} \circ M \circ {}^t\bar{B} \preccurlyeq_{\mathrm{L}} \bar{B} \circ M \circ {}^tC \preccurlyeq_{\mathrm{L}} C \circ M \circ {}^tC$.

Preuve. Comme les tenseurs M et N sont positifs sur Z et que *B appartient à Γ*, la non vacuité de $\Gamma_{MN}(B)$ est équivalente à la relation

$$Z \cap \operatorname{Ker}(M+N) \subset \operatorname{Ker}(B \circ M)$$

qui est immédiate, car, sous de telles hypothèses, il est clair que:

$$Z \cap \operatorname{Ker}(M+N) = Z \cap \operatorname{Ker} M \cap \operatorname{Ker} N \subset \operatorname{Ker} M \subset \operatorname{Ker}(B \circ M).$$

(i) Soient $\bar{B} \in \Gamma_{MN}(B)$ et $C \in B + (Z \cap \operatorname{Ker} N) \otimes F$. Compte tenu de la relation $\operatorname{Im} {}^t(B - C) \subset Z$ (proposition 1.1), l'assertion (ii) du théorème 3.2 précédent entraîne

$$\bar{B}\circ(M+N)\circ{}^t(B-C)=B\circ N\circ{}^t(B-C),$$

où $\operatorname{Im}{}^t(B-C)\subset \operatorname{Ker} N$, dont on déduit que $\bar{B}\circ M\circ{}^tB=\bar{B}\circ M\circ{}^tC$. Par ailleurs, $\bar{B}\in\Gamma$, d'où $\operatorname{Im}{}^t(B-\bar{B})\subset Z$ (proposition 1.1), et donc on a également

$$\bar{B}\circ(M+N)\circ{}^t(B-\bar{B})=B\circ N\circ{}^t(B-\bar{B}),$$

ce qui en développant, donne:

$$\begin{aligned}\bar{B}\circ M\circ{}^tB&=\bar{B}\circ M\circ{}^t\bar{B}-\bar{B}\circ N\circ{}^t(B-\bar{B})+B\circ N\circ{}^t(B-\bar{B})\\&=\bar{B}\circ M\circ{}^t\bar{B}+(B-\bar{B})\circ N\circ{}^t(B-\bar{B})=\varrho(M,N,\bar{B}).\end{aligned}$$

Enfin, pour tout $T\in\mathscr{L}(E,F)$, on a $\varrho(M,N,T)=\varrho(N,M,B-T)$, ce qui implique aussitôt, à l'aide de l'égalité précédente, que:

$$\begin{aligned}\varrho(M,N,\bar{B})&=(B-\bar{B})\circ N\circ{}^tB=B\circ N\circ{}^tB-\bar{B}\circ N\circ{}^tB\\&=C\circ N\circ{}^tC-\bar{B}\circ N\circ{}^tC.\end{aligned}$$

(ii) Pour tout $C\in B+(Z\cap\operatorname{Ker}N)\otimes F$ tel que $\operatorname{Im}{}^t(B-C)=Z\cap\operatorname{Ker}N$, on a donc, d'après (i), $\bar{B}\circ M\circ{}^t(B-C)=0$, d'où il résulte que $Z\cap\operatorname{Ker}N\subset\operatorname{Ker}(\bar{B}\circ M)$.

(iii) Puisque $\varrho(M,N,\bar{B})=\bar{B}\circ M\circ{}^tC$, que $C\in\Gamma$ et que $\bar{B}\in\Gamma_{MN}(B)$, on a

$$\bar{B}\circ M\circ{}^tC\preccurlyeq_{\mathrm{L}} C\circ M\circ{}^tC.$$

Par ailleurs, $\operatorname{Im}{}^t(B-\bar{B})\subset Z$, et comme le tenseur N est supposé positif sur Z, il est clair que $(B-\bar{B})\circ N\circ{}^t(B-\bar{B})\succcurlyeq_{\mathrm{L}}0$, d'où l'on tire:

$$\bar{B}\circ M\circ{}^t\bar{B}=\varrho(M,N,\bar{B})-(B-\bar{B})\circ N\circ{}^t(B-\bar{B})\preccurlyeq_{\mathrm{L}}\bar{B}\circ M\circ{}^tC,$$

ce qui achève la preuve. ∎

On notera que l'assertion (iii) entraîne que $\Gamma_{MN}(C)=\Gamma_{MN}(B)$ pour tout élément C du sous-espace affine $B+(Z\cap\operatorname{Ker}N)\otimes F$ de Γ.

3.3. Approximation au sens du préordre de Schur. Nous conservons dans ce paragraphe les hypothèses et notations précédentes mais supposons en outre que $F'=F$ (de manière à ce que l'on puisse définir le préordre de Schur sur $\mathscr{L}(F)$). Pour tout sous-espace affine Γ de $\mathscr{L}(E,F)$, nous noterons $\tilde{\Gamma}_{MN}(B)$ l'ensemble, éventuellement vide, des éléments $\bar{B}\in\Gamma$ solutions du programme:

$$(\mathscr{P}_2)\qquad \operatorname{Min}\operatorname{tr}[\varrho(M,N,C)]:\ C\in\Gamma.$$

Soulignons que, si la dualité $\langle\cdot,\cdot\rangle$ sur F est positive, la trace sur $\mathscr{L}(F)$ conserve l'ordre de Loewner, et donc dans ce cas $\Gamma_{MN}(B)\subset\tilde{\Gamma}_{MN}(B)$.

Théorème 3.5. *Soient Z et W deux sous-espaces vectoriels respectivement de E' et F, Γ un sous-espace affine de $\mathscr{L}(E,F)$ de direction $Z\otimes W$. Alors*:

(i) *Une condition nécessaire et suffisante pour que* $\tilde{\Gamma}_{MN}(B)$ *ne soit pas vide est que, d'une part, la restriction de la dualité sur* F *au sous-espace* W *et la restriction du tenseur* $M+N$ *au sous-espace* Z *soient de même signe constant et que, d'autre part,* C *désignant un élément arbitraire de* Γ *et posant* $R = C\circ(M+N)-B\circ N$, *on ait*

$$R[Z\cap\mathrm{Ker}(M+N)] \subset W^{\perp}.$$

L'ensemble $\tilde{\Gamma}_{MN}(B)$ *est un singleton si, et seulement si, les restrictions de la dualité sur* F *au sous-espace* W *et de* $M+N$ *au sous-espace* Z *sont de même signe constant, la restriction de* $M+N$ *à* Z *étant de plus injective.*

(ii) *Supposons* $\tilde{\Gamma}_{MN}(B) \neq \emptyset$. *Pour qu'un élément* $\bar{B}$ *de* Γ *appartienne à* $\tilde{\Gamma}_{MN}(B)$, *il faut et il suffit que* $[\bar{B}\circ(M+N)-B\circ N](Z) \subset W^{\perp}$ *et* $\tilde{\Gamma}_{MN}(B)$ *est alors un sous-espace affine de* $\mathscr{L}(E, F)$ *de direction* $[Z\cap\mathrm{Ker}(M+N)]\otimes W$.

Preuve. (i) Posons $\Delta = Z\otimes W$ et considérons les tenseurs symétriques Φ et Φ' sur $\mathscr{L}(F, E)$ relativement à la dualité Ψ sur $\mathscr{L}(F, E)\times\mathscr{L}(F, E')$ définie par $\Psi(S, T) = \mathrm{tr}(S\circ{}^tT)$ (proposition 1.5) tels que $\Phi(T) = (M+N)\circ T$ et $\Phi'(T) = N\circ T$. L'ensemble $\tilde{\Gamma}_{MN}(B)$ apparaît alors comme l'ensemble des solutions du programme

$$\mathrm{Min}\,\Psi[{}^tS, \Phi({}^tS)-2\Phi'({}^tB)]:\ S\in\Gamma,$$

autrement dit encore du programme

$$\text{(3)}\qquad \mathrm{Min}\,\Psi[T, \Phi(T)-2\Phi'({}^tB)]:\ T\in{}^tC+W\otimes Z,$$

où C est un élément arbitraire (mais fixé) de Γ. D'après le théorème 2.3 appliqué à la dualité séparante Ψ, un tel programme admet une solution si, et seulement si, le tenseur Φ est positif sur $W\otimes Z$, ce qui s'écrit

$$\text{(4)}\qquad (\forall D\in\Delta)\ \mathrm{tr}[D\circ(M+N)\circ{}^tD] \geqslant 0,$$

et si

$$\text{(5)}\qquad (M+N)\circ{}^tC-N\circ{}^tB\in(W\otimes Z)^{\perp}+\mathrm{Im}\,\Phi.$$

Or (proposition 1.1), $\mathscr{L}(F, E)$ coïncide avec $F\otimes E$, donc Φ n'est autre que l'endomorphisme de $F\otimes E$ entièrement défini par

$$\Phi:\ y\otimes x\to(M+N)\circ(y\otimes x) = y\otimes[(M+N)x]$$

et $\mathrm{Im}\,\Phi = F\otimes\mathrm{Im}(M+N)$. Il en résulte, en particulier, que la condition (4) entraîne

$$\big(\forall(z, x)\in Z\times W\big)\ \Psi[z\otimes x, \Phi(z\otimes x)] = \langle x, x\rangle\langle z, (M+N)z\rangle \geqslant 0,$$

d'où il découle que la restriction de la dualité $\langle\cdot, \cdot\rangle$ sur F au sous-espace W et la restriction de $M+N$ au sous-espace Z doivent être de *même signe* (constant) pour que $\tilde{\Gamma}_{MN}(B) \neq \emptyset$. Cette condition est également suffisante. En effet, soit

$D \in Z \otimes W$; d'après le corollaire 1.2, pour toute base $\{y_i \colon i \in I\}$ de W (que l'on peut choisir orthogonale relativement à la composante symétrique de la dualité $\langle \cdot, \cdot \rangle$ sur F), il existe une famille (unique) $\{z_i \colon i \in I\}$ de Z telle que $D = \sum_{i \in I} z_i \otimes y_i$. Par suite, d'après la proposition 1.5, il vient:

$$\begin{aligned} \Psi[{}^tD, \Phi({}^tD)] &= \Psi\big[\sum_{i \in I} y_i \otimes z_i, (M+N) \circ \sum_{i \in I} y_i \otimes z_i\big] \\ &= \Psi\big[\sum_{i \in I} y_i \otimes z_i, \sum_{i \in I} y_i \otimes (M+N) z_i\big] \\ &= \sum_{i \in I} \langle y_i, y_i \rangle \langle z_i, (M+N) z_i \rangle + \sum_{i \neq j} \langle y_i, y_j \rangle \langle z_i, (M+N) z_j \rangle \\ &= \sum_{i \in I} \langle y_i, y_i \rangle \langle z_i, (M+N) z_i \rangle \geqslant 0. \end{aligned}$$

De plus, toujours d'après la proposition 1.5, l'orthogonal $(W \otimes Z)^{\perp \Psi}$ de $W \otimes Z$ par rapport à la dualité Ψ est le sous-espace $W^{\perp} \otimes E + F \otimes Z^{\perp}$ de $\mathscr{L}(F, E)$; on en déduit:

$$\begin{aligned} (W \otimes Z)^{\perp \Psi} + \operatorname{Im} \Phi &= W^{\perp} \otimes E + F \otimes [Z^{\perp} + \operatorname{Im}(M+N)] \\ &= W^{\perp} \otimes E + F \otimes [Z \cap \operatorname{Ker}(M+N)]^{\perp} \\ &= \{[W \otimes E] \cap [F \cap \operatorname{Ker}(M+N)]\}^{\perp}. \end{aligned}$$

On peut remarquer que l'on a:

$$\begin{aligned} [W \otimes E] \cap [F \cap \operatorname{Ker}(M+N)] &= [W \cap F] \otimes \{E \cap [Z \cap \operatorname{Ker}(M+N)]\} \\ &= W \otimes [Z \cap \operatorname{Ker}(M+N)]. \end{aligned}$$

Il en résulte clairement que si l'on pose $R = C \circ (M+N) - B \circ N$, la condition (5) ci-dessus est équivalente à:

$$(\forall y \in W) \; \big(\forall z \in Z \cap \operatorname{Ker}(M+N)\big) \; \operatorname{tr}(Rz \otimes y) = 0,$$

autrement dit à la relation

$$R[Z \cap \operatorname{Ker}(M+N)] \subset W^{\perp}.$$

La première partie de l'assertion (i) est ainsi démontrée. La deuxième partie est une conséquence immédiate du théorème 2.3: pour que $\tilde{\Gamma}_{MN}(B)$ ne soit pas vide, il faut et il suffit que le tenseur Φ soit strictement positif sur $W \otimes Z$, ce qui se traduit bien, d'après la preuve précédente, par la condition énoncée.

(ii) D'après le théorème 2.3, pour qu'un élément $\bar{B} \in \Gamma$ soit une solution du programme (3), il faut et il suffit que:

$$(M+N) \circ {}^t\bar{B} - N \circ {}^tB \in (W \otimes Z)^{\perp \Psi},$$

c'est-à-dire, si l'on pose $R = \bar{B} \circ (M+N) - B \circ N$, que:

$$\big(\forall (z, y) \in Z \times W\big) \; \operatorname{tr}(Rz \otimes y) = 0.$$

Cette propriété est clairement équivalente à la relation $R(Z) \subset W^{\perp}$. De plus, l'ensemble des solutions de (3), s'il n'est pas vide, est, toujours d'après le théorème 2.3 appliqué à ce programme, un sous-espace affine de Γ dont la direction est l'ensemble des éléments D de $\mathscr{L}(E', F)$ tels que ${}^tD \in (W \otimes Z) \cap \operatorname{Ker} \Phi$. Comme:

$$\operatorname{Ker} \Phi = \{T \in \mathscr{L}(F, E) \colon \operatorname{Im} T \subset \operatorname{Ker}(M+N)\}$$

et que (cf. proposition 1.1):

$$W \otimes Z = \{T \in \mathscr{L}(F, E) \colon \operatorname{Im} T \subset Z \text{ et } \operatorname{Im} {}^tT \subset W\},$$

le sous-espace $[Z \cap \operatorname{Ker}(M+N)] \otimes W$ est bien la direction de $\tilde{\Gamma}_{MN}(B)$, ce qui achève la preuve. ∎

D'après l'assertion (i) de chacun des théorèmes 3.2 et 3.5, on voit que si la dualité $\langle \cdot, \cdot \rangle$ est positive sur le sous-espace W, $\Gamma_{MN}(B)$ peut être vide sans que $\tilde{\Gamma}_{MN}(B)$ le soit, mais si $\Gamma_{MN}(B)$ n'est pas vide, il en est de même pour $\tilde{\Gamma}_{MN}(B)$ et alors $\tilde{\Gamma}_{MN}(B)$ coïncide avec $\Gamma_{MN}(B)$, de telle sorte que l'on a le théorème fondamental suivant *d'invariance quadratique*:

THÉORÈME 3.6. *Soient Z et W deux sous-espaces vectoriels respectivement de E' et F, Γ un sous-espace affine de $\mathscr{L}(E, F)$ de direction $Z \otimes W$.*

(i) *Si $\Gamma_{MN}(B) \neq \emptyset$ il faut et il suffit que la restriction de la dualité sur F au sous-espace W soit positive pour que $\tilde{\Gamma}_{MN}(B) \neq \emptyset$, et alors les ensembles $\tilde{\Gamma}_{MN}(B)$ et $\Gamma_{MN}(B)$ coïncident.*

(ii) *Si la restriction de la dualité sur F au sous-espace W est positive, pour que les programmes $(\mathscr{P}_1)$ et $(\mathscr{P}_2)$ soient équivalents, c'est-à-dire en raison de* (i), *pour que les conditions de non vacuité de $\tilde{\Gamma}_{MN}(B)$ et $\Gamma_{MN}(B)$ soient identiques, il faut et il suffit que $W = F$.*

CONCLUSION. On notera que le caractère éventuellement positif de la dualité $\langle \cdot, \cdot \rangle$ sur F ne concerne véritablement que sa composante symétrique, elle-même alors non dégénérée. D'un point de vue pratique qui n'affaiblit pas vraiment la généralité, si le tenseur $M+N$ est positif sur Z et si $W = F$, il est alors *nécessaire* (et bien entendu suffisant) de supposer que l'espace F est *euclidien* pour que l'ensemble $\tilde{\Gamma}_{MN}(B)$ ne soit pas vide. Les propriétés relatives à $\Gamma_{MN}(B)$, notamment la stabilité par linéarité s'étendent alors, en termes absolument identiques, à l'ensemble $\tilde{\Gamma}_{MN}(B)$.

4. L'ADMISSIBILITÉ EN APPROXIMATION QUADRATIQUE D'APPLICATIONS LINÉAIRES

Dans cette section, on conserve les notations de la section 3 précédente. On considère un système dual (E, E'), un espace *euclidien* F, un sous-espace vectoriel Z de E' non réduit à $\{0\}$ et une application linéaire B de E dans F. On désigne par Γ un sous-espace affine de $\mathscr{L}(E, F)$ de direction

$Z \otimes F$. Nous utiliserons une seule et même notation pour désigner les ensembles $\tilde{\Gamma}_{MN}(B)$ et $\Gamma_{MN}(B)$ qui coïncident sous ces hypothèses (théorème 3.6): la notation $\Gamma_{MN}(B)$.

Chaque espace vectoriel réel de dimension finie mis en jeu est naturellement supposé muni de la topologie canonique. Nous désignons une fois pour toutes par Σ un *cône convexe fermé* de l'espace $\mathscr{L}_s(E', E) \times \mathscr{L}_s(E', E)$ formé de couples (M, N) de tenseurs *positifs sur* Z et non réduit à la famille des couples (M, N) de tenseurs s'annulant sur Z, par Ξ un sous-espace vectoriel de $\mathscr{L}_s(E', E) \times \mathscr{L}_s(E', E)$ contenant Σ et par Ξ^+ le cône saillant de Ξ constitué des couples de tenseurs positifs sur Z. A noter que, sous de telles hypothèses, l'ensemble Σ ne peut pas être un espace vectoriel.

4.1. Le concept d'approximation quadratique admissible. Pour tout sous-ensemble non vide Θ de $\mathscr{L}_s(E', E) \times \mathscr{L}_s(E', E)$, on définit clairement sur $\mathscr{L}(E, F)$ une relation de préordre partiel $\preccurlyeq$ en posant:

$$T_1 \preccurlyeq T_2 \Leftrightarrow (\forall (M, N) \in \Theta) \ \varrho(M, N, T_1) \preccurlyeq_S \varrho(M, N, T_2)$$

et notera par $\prec$ le préordre strict déduit de $\preccurlyeq$. Tout élément minimal de Γ relativement au préordre $\preccurlyeq$ sur $\mathscr{L}(E, F)$, s'il en existe, sera appelé *approximation Θ-admissible de B dans Γ.*

A noter qu'il revient au même de considérer le préordre $\preccurlyeq$ relativement à l'ensemble Θ ou relativement au cône convexe fermé $[\Theta]$ engendré par Θ (le cône convexe fermé engendré par Θ est l'adhérence de l'ensemble des combinaisons linéaires à coefficients strictement positifs d'éléments de Θ, voir [1] par exemple). Donc les notions d'approximation Θ-admissible et d'approximation $[\Theta]$-admissible] dans Γ sont identiques.

Proposition 4.1. *Soit (M, N) un couple de Θ tel que $\Gamma_{MN}(B) \neq \emptyset$. Pour qu'un élément A de $\Gamma_{MN}(B)$ soit une approximation Θ-admissible de B dans Γ, il faut et il suffit que A soit une approximation Θ-admissible de B dans $\Gamma_{MN}(B)$.*

En effet, s'il existe $T \in \Gamma$ tel que $T \prec A$, alors $T \in \Gamma_{MN}(B)$, ce qui prouve que la condition est suffisante. La condition nécessaire est triviale puisque $\Gamma_{MN}(B) \subset \Gamma$. ∎

4.2. La notion d'élément singulier. A l'instar de LaMotte ([5], p. 246), nous dirons qu'un couple (M, N) de Ξ est *singulier* si $\Gamma = \Gamma_{MN}(B)$ et désignerons par $\mathscr{S}(B, \Gamma)$ l'ensemble des éléments singuliers de Ξ. Compte tenu des assertions (i) et (ii) du théorème 3.2, il est immédiat que:

Proposition 4.2. *L'ensemble $\mathscr{S}(B, \Gamma)$ est le sous-espace vectoriel de Ξ tel que*:

$$\mathscr{S}(B, \Gamma) = \{(M, N) \in \Xi\colon \ Z \subset \operatorname{Ker}(M+N) \ \text{ et } \ Z \subset \operatorname{Ker}[C \circ (M+N) - B \circ N]\},$$

où C est un élément arbitraire de Γ.

Nous laissons au lecteur le soin de vérifier lui-même l'identité:

(6) $\varrho(M, N, C+U)-\varrho(M, N, C) = U\circ {}^tR+R\circ {}^tU+U\circ(M+N)\circ {}^tU,$

où $R = C\circ(M+N)-B\circ N$. Considérons un couple (M_0, N_0) de $\mathscr{S}(B, \Gamma)$. Comme la restriction de M_0+N_0 au sous-espace Z est nulle (proposition 4.2), il résulte clairement de l'identité (6) précédente que, pour tout $(M, N)\in\Xi$ et tout $(C, U)\in\Gamma\times Z\otimes F$, on a:

$$\varrho(M+M_0, N+N_0, C+U)-\varrho(M+M_0, N+N_0, C) = \varrho(M, N, C+U)-\varrho(M, N, C).$$

Par suite, si $\mathfrak{I}$ désigne le cône convexe fermé engendré par $\Sigma+\mathscr{S}(B, \Gamma)$ (autrement dit tout simplement l'adhérence de $\Sigma+\mathscr{S}(B, \Gamma)$ car Σ est un cône convexe, cf. [1]), la Σ-admissibilité dans Γ équivaut à la $\mathfrak{I}$-admissibilité dans Γ. On notera encore que, pour tout cône convexe fermé Σ' de $\mathscr{L}_s(E', E)\times\mathscr{L}_s(E', E)$ tel que $\mathfrak{I}$ soit l'adhérence du cône convexe $\Sigma'+\mathscr{S}(B, \Gamma)$, la Σ-admissibilité dans Γ équivaut encore à la Σ'-admissibilité dans l'ensemble Γ. Soulignons le résultat plus précis suivant tiré de [5] (p. 250):

Proposition 4.3. *Supposons $\Sigma \not\subset \mathscr{S}(B, \Gamma)$. Soient $\mathfrak{I}$ l'adhérence de $\Sigma+\mathscr{S}(B, \Gamma)$, $\mathscr{S}'$ un supplémentaire de $\mathscr{S}(B, \Gamma)$ dans Ξ et $\Sigma' = \mathfrak{I}\cap\mathscr{S}'$. Alors Σ' n'est pas un espace vectoriel et l'ensemble des approximations Σ-admissibles de B dans Γ coïncide avec l'ensemble des approximations Σ'-admissibles de B dans Γ.*

Preuve. Considérons un supplémentaire $\mathscr{S}'$ de $\mathscr{S}(B, \Gamma)$ dans Ξ et le cône convexe fermé $\Sigma' = \mathfrak{I}\cap\mathscr{S}'$. Pour tout $(M, N)\in\mathfrak{I}\backslash\mathscr{S}(B, \Gamma)$, on a clairement:

$$(\forall x\in Z)\ \langle x, (M+N)x\rangle = \langle x, (M'+N')x\rangle,$$

où (M', N') est la projection de (M, N) sur $\mathscr{S}'$ parallèlement à $\mathscr{S}(B, \Gamma)$; comme par hypothèse il existe un couple $(M, N)\in\Sigma$ (et n'appartenant pas à $\mathscr{S}(B, \Gamma)$) tel que la restriction de $M+N$ au sous-espace Z ne soit pas nulle, Σ' n'est donc pas un sous-espace vectoriel de Ξ (sinon $-(M', N')\in\Sigma'$ et $\langle x, (M+N)x\rangle = 0$ pour tout $x\in Z$).

Par ailleurs, pour toute partie $\mathscr{E}$ de Ξ, la projection de $\mathscr{E}$ sur $\mathscr{S}'$ parallèlement à $\mathscr{S}(B, \Gamma)$ est

$$[\mathscr{E}+\mathscr{S}(B, \Gamma)]\cap\mathscr{S}' \quad \text{et} \quad \mathscr{E}+\mathscr{S}(B, \Gamma) = \mathscr{E}+[\mathscr{E}+\mathscr{S}(B, \Gamma)]\cap\mathscr{S}'.$$

Comme on a $\mathscr{S}(B, \Gamma)\subset\mathfrak{I}$, la projection de $\mathfrak{I}$ sur $\mathscr{S}'$ parallèlement à $\mathscr{S}(B, \Gamma)$ est donc l'ensemble

$$[\mathfrak{I}+\mathscr{S}(B, \Gamma)]\cap\mathscr{S}' = \mathfrak{I}\cap\mathscr{S}',$$

d'où l'on tire $\Sigma'+\mathscr{S}(B, \Gamma) = \mathfrak{I}$, avec clairement $\Sigma'\cap\mathscr{S}(B, \Gamma) = \{0\}$, ce qui achève la preuve. ■

4.3. Le théorème de L. R. LaMotte. Extension significative des travaux originaux de Olsen et al. ([6], voir aussi [4]), le théorème suivant, dû à LaMotte ([5], p. 249) dans un cadre plus restrictif (et prouvé à l'aide d'arguments exclusivement matriciels) constitue une approche déterminante de l'admissibilité en approximation quadratique d'applications linéaires. En effet, Klonecki et Zontek [3] ont réussi à en déduire efficacement la structure des estimateurs linéaires admissibles relativement à un modèle classique de Gauss–Markov (voir aussi [10]). Compte tenu des hypothèses plus générales et de la terminologie retenues dans cet article, c'est le

THÉORÈME 4.1. *Pour que $A \in \Gamma$ soit une approximation Σ-admissible de B dans Γ, il faut et il suffit que $\Sigma \subset \mathscr{S}(B, \Gamma)$ ou qu'il existe $(M, N) \in \mathfrak{I} \backslash \mathscr{S}(B, \Gamma)$ tel que $A \in \Gamma_{MN}(B)$ et que A soit une approximation Σ-admissible de B dans $\Gamma_{MN}(B)$.*

Sa démonstration repose essentiellement sur le théorème ci-dessous, également démontré pour la première fois par LaMotte ([5], p. 249) lorsque E est muni d'une structure euclidienne. D'un point de vue mathématique, cette hypothèse n'est pas nécessaire. Nous allons en donner une preuve plus générale fondée sur le théorème de Hahn–Banach, prolongeant en quelque sorte une idée de Olsen et al. ([6], pp. 883–884) et n'utilisant que la seule structure vectorielle de E. Pour tout cône convexe $\mathscr{C}$ de $\mathscr{L}_s(E', E) \times \mathscr{L}_s(E', E)$, $\mathscr{C}^*$ désigne le cône épointé $\mathscr{C} \backslash \{(0, 0)\}$:

THÉORÈME 4.2. *Soit $\mathscr{C}$ un cône convexe fermé de Ξ contenu dans l'adhérence de $\Xi^+ + \mathscr{S}(B, \Gamma)$, ne formant pas un sous-espace vectoriel, tel que $\mathscr{C} \cap \mathscr{S}(B, \Gamma) = \{0\}$. S'il existe $A \in \Gamma$ tel que, quel que soit $(M, N) \in \mathscr{C}^*$, $A \notin \Gamma_{MN}(B)$, alors il existe $\bar{A} \in \Gamma$ vérifiant à la fois:*

(i) $(\forall (M, N) \in \mathscr{C})$ $\operatorname{tr}[\varrho(M, N, \bar{A})] \leqslant \operatorname{tr}[\varrho(M, N, A)]$,

(ii) $(\forall (M, N) \in \mathscr{C} \backslash \Delta)$ $\operatorname{tr}[\varrho(M, N, \bar{A})] < \operatorname{tr}[\varrho(M, N, A)]$, *où Δ est le plus grand sous-espace vectoriel de Ξ contenu dans $\mathscr{C}$.*

Preuve. Observons pour commencer que, pour chaque couple (M, N) de $\mathscr{C}$, la restriction de $M+N$ au sous-espace Z est positive. En effet, le cône convexe $\Xi^+ + \mathscr{S}(B, \Gamma)$ — et donc aussi son adhérence — est contenu dans le cône convexe fermé formé des couples $(M, N) \in \Xi$ tels que la restriction de $M+N$ au sous-espace Z soit positive (proposition 4.2). Notons aussi que la condition $\mathscr{C} \cap \mathscr{S}(B, \Gamma) = \{0\}$ est nécessaire pour qu'il existe un élément $A \in \Gamma$ tel que, pour tout $(M, N) \in \mathscr{C}^*$, $A \notin \Gamma_{MN}(B)$.

Par ailleurs, comme $F' = F$, les espaces vectoriels $\mathscr{L}(F, E)$ et $\mathscr{L}(F, E')$ sont mis en dualité séparante par la forme bilinéaire Ψ sur $\mathscr{L}(F, E') \times \mathscr{L}(F, E)$ telle que (proposition 1.5):

$$\Psi(S, T) = \operatorname{tr}({}^tT \circ S) = \operatorname{tr}({}^tS \circ T).$$

Afin d'éviter toute confusion, nous noterons $\mathscr{A}^{\perp\Psi}$ l'orthogonal dans $\mathscr{L}(F, E')$ de toute partie $\mathscr{A}$ de $\mathscr{L}(F, E)$ relativement à Ψ.

Enfin, le cône convexe fermé $\mathscr{C}$ s'écrit $\mathscr{C} = \boldsymbol{R}^+\Omega+\Delta$, où Ω est un convexe compact ne contenant pas l'origine (*socle* de $\mathscr{C}$) et disjoint de $\Delta = \mathscr{C}\cap(-\mathscr{C})$ qui est le plus grand sous-espace vectoriel contenu dans $\mathscr{C}$ (proposition 1.6). Comme $\mathscr{C}$ n'est pas un sous-espace vectoriel de $\mathscr{L}_s(E', E)\times\mathscr{L}_s(E', E)$, on a donc $\mathscr{C} \neq \Delta$.

Considérons l'application linéaire

$$\varphi\colon (M, N) \to (M+N)\circ {}^tA - N\circ {}^tB$$

de Ξ dans $\mathscr{L}(F, E)$. De l'équivalence (théorème 3.2):

$$A\in\Gamma_{MN}(B) \Leftrightarrow \varphi(M, N)\in F\otimes Z^\perp$$

il résulte, par contraposition, que supposer $A\notin\Gamma_{MN}(B)$ pour chaque (M, N) de $\mathscr{C}^*$ revient à écrire que

$$\varphi(\mathscr{C}^*)\cap(F\otimes Z^\perp) = \emptyset,$$

dont on déduit que la restriction de φ à l'ensemble $\mathscr{C}$ est *injective* (car s'annule seulement à l'origine). Il suit clairement que $\varphi(\Delta)$ est le plus grand sous-espace vectoriel inclus dans le cône convexe $\varphi(\mathscr{C}) = \boldsymbol{R}^+\varphi(\Omega)+\varphi(\Delta)$, que $\varphi(\Omega)\cap\varphi(\Delta) = \emptyset$ et que $\varphi(\Omega)$ est un convexe compact (φ est ici continue) de $\mathscr{L}(F, E)$ disjoint de $F\otimes Z^\perp+\varphi(\Delta)$. D'après le théorème de Hahn–Banach (cf. Bourbaki [1]), il existe donc un hyperplan vectoriel $\mathscr{H}$ de $\mathscr{L}(F, E)$ contenant le sous-espace $F\otimes Z^\perp+\varphi(\Delta)$ et disjoint de $\varphi(\Omega)$. Autrement dit, il existe $U'\in[F\otimes Z^\perp+\varphi(\Delta)]^{\perp\Psi}$, où (proposition 1.5):

$$[F\otimes Z^\perp+\varphi(\Delta)]^{\perp\Psi} = (F\otimes Z)\cap[\varphi(\Delta)]^{\perp\Psi},$$

tel que, si l'on pose $U = {}^tU'$ (application linéaire de E dans F), on ait:

$$\varphi(\Omega) \subset \{T\in\mathscr{L}(F, E)\colon\ \Psi({}^tU, T) < 0\}.$$

Il en résulte que $\Psi[{}^tU, (M+N)\circ {}^tU] > 0$ pour chaque $(M, N)\in\Omega$. En effet, pour chaque couple $(M, N)\in\mathscr{C}$, le tenseur $M+N$ est positif sur Z. Donc, s'il existait $(M, N)\in\Omega$ tel que

$$\Psi[{}^tU, (M+N)\circ {}^tU] \leqslant 0,$$

on aurait

$$\Psi[{}^tU, (M+N)\circ {}^tU] = \mathrm{tr}[U\circ(M+N)\circ {}^tU] = 0$$

et on en déduirait (car $U\in F\otimes Z$ et donc $\mathrm{Im}\,{}^tU \subset Z$) que

$$U\circ(M+N)\circ {}^tU = 0.$$

Or (M, N) est la limite de la somme de deux suites $((M_1^{(n)}, N_1^{(n)}))$ et $((M_0^{(n)}, N_0^{(n)}))$ respectivement de Ξ^+ et $\mathscr{S}(B, \Gamma)$. Comme

$$U \circ (M_0^{(n)} + N_0^{(n)}) \circ {}^tU = 0 \quad \text{et}$$

$$\Psi[{}^tU, \varphi(M_0^{(n)}, N_0^{(n)})] = 0 \quad \text{pour chaque } n \in \boldsymbol{N}$$

et que φ est linéaire (donc ici continue), on aurait:

$$\lim_{n\to\infty} U \circ (M_1^{(n)} + N_1^{(n)}) \circ {}^tU = 0 \quad \text{et}$$

$$\Psi[{}^tU, \varphi(M, N)] = \lim_{n\to\infty} \Psi[{}^tU, \varphi(M_1^{(n)}, N_1^{(n)})].$$

Il en découlerait que les suites de tenseurs positifs $(U \circ M_1^{(n)} \circ {}^tU)$ et $U \circ N_1^{(n)} \circ {}^tU$ convergeraient vers 0 et donc (les tenseurs $M_1^{(n)}$ et $N_1^{(n)}$ étant positifs sur Z pour tout $n \in \boldsymbol{N}$, et $\operatorname{Im}{}^tU$ étant inclus dans Z) que:

$$\lim_{n\to\infty} U \circ M_1^{(n)} = \lim_{n\to\infty} U \circ N_1^{(n)} = 0,$$

d'où $\Psi[{}^tU, \varphi(M, N)] = 0$, ce qui contredirait la définition de U.

Posons $C = A + \lambda U$ avec $\lambda \in \boldsymbol{R}^*$, et $P(\lambda, M, N) = \operatorname{tr}[\varrho(M, N, C) - \varrho(M, N, A)]$. Comme $\operatorname{Im}{}^tU \subset Z$, on a $C \in \Gamma$, et compte tenu de la relation (6), un calcul élémentaire donne:

$$P(\lambda, M, N) = 2\lambda\Psi[{}^tU, \varphi(M, N)] + \lambda^2\Psi[{}^tU, (M+N) \circ {}^tU].$$

Pour chaque $(M, N) \in \Omega$, $P(\lambda, M, N)$ est un polynôme en λ, minimum pour

$$\lambda = \lambda_0 = -\frac{\Psi[{}^tU, \varphi(M, N)]}{\Psi[{}^tU, (M+N) \circ {}^tU]},$$

et le minimum atteint

$$\zeta = -\frac{(\Psi[{}^tU, \varphi(M, N)])^2}{\Psi[{}^tU, (M+N) \circ {}^tU]}$$

est strictement négatif. L'application f de Ω dans $\boldsymbol{R}$ définie par $f(M, N) = \lambda_0$ est continue et strictement positive sur le compact Ω. La borne inférieure m dans $\boldsymbol{R}$ de l'ensemble $f(\Omega)$ est donc *strictement positive* et, pour chaque $(M, N) \in \Omega$, on a

$$\frac{P(m, M, N)}{m} \leqslant 2\Psi[{}^tU, \varphi(M, N)] + f(M, N)\Psi[{}^tU, (M+N) \circ {}^tU],$$

où

$$2\Psi[{}^tU, \varphi(M, N)] + f(M, N)\Psi[{}^tU, (M+N) \circ {}^tU] = \frac{P(f(M, N), M, N)}{f(M, N)} < 0,$$

ce qui, en posant $\bar{A} = A + mU$, se traduit par

$$(\forall (M, N) \in \Omega) \ \operatorname{tr}[\varrho(M, N, \bar{A}) - \varrho(M, N, A)] < 0.$$

Considérons maintenant un couple $(M, N) \in \mathscr{C}$. On a $(M, N) = (M_1, N_1) + (M_2, N_2)$ avec $(M_1, N_1) \in \boldsymbol{R}^+ \Omega$ et $(M_2, N_2) \in \Delta$, et la linéarité de φ entraîne que

$$\begin{aligned} \Psi[{}^tU, \varphi(M, N)] &= \Psi[{}^tU, \varphi(M_1, N_1)] + \Psi[{}^tU, \varphi(M_2, N_2)] \\ &= \Psi[{}^tU, \varphi(M_1, N_1)], \end{aligned}$$

car ${}^tU \in [\dot{\varphi}(\Delta)]^{\perp \Psi}$. D'autre part, $-(M_2, N_2) \in \Delta$ et donc $M_2 + N_2$ s'annule sur Z, d'où il découle immédiatement (car $\operatorname{Im} {}^tU \subset Z$) que

$$\operatorname{tr}[\varrho(M_2, N_2, \bar{A}) - \varrho(M_2, N_2, A)] = 0$$

et donc que

$$\operatorname{tr}[\varrho(M, N, \bar{A}) - \varrho(M, N, A)] = \operatorname{tr}[\varrho(M_1, N_1, \bar{A}) - \varrho(M_1, N_1, A)],$$

ce qui permet clairement de conclure, A et $\bar{A}$ ayant même tenseur d'erreur quadratique pour tout $(M, N) \in \Delta$. ∎

Sous les hypothèses de ce théorème, il s'ensuit clairement que A n'est pas une approximation $\mathscr{C}$-admissible de B dans Γ. Par contraposition, pour qu'un élément A de Γ soit une approximation $\mathscr{C}$-admissible de B dans Γ, il faut qu'il existe $(M, N) \in \mathscr{C}^*$ tel que $A \in \Gamma_{MN}(B)$, et la relation $\Gamma_{MN}(B) \subset \Gamma$ implique alors aussitôt que A est une approximation $\mathscr{C}$-admissible de B dans $\Gamma_{MN}(B)$. Par ailleurs (proposition 4.1) toute approximation $\mathscr{C}$-admissible de B dans $\Gamma_{MN}(B)$ est aussi $\mathscr{C}$-admissible dans Γ. On peut maintenant donner la preuve du théorème 4.1.

Preuve. Lorsque $\Sigma \not\subset \mathscr{S}(B, \Gamma)$ et si $\mathfrak{I}$ désigne l'adhérence de $\Sigma + \mathscr{S}(B, \Gamma)$, alors pour tout supplémentaire $\mathscr{S}'$ de $\mathscr{S}(B, \Gamma)$ dans Ξ, $\Sigma' = \mathfrak{I} \cap \mathscr{S}'$ est un cône convexe fermé qui vérifie les hypothèses du théorème 4.2 et la Σ-admissibilité équivaut à la Σ'-admissibilité (proposition 4.2). De plus, si $\Sigma \subset \mathscr{S}(B, \Gamma)$, tous les éléments de Γ ont même tenseur d'erreur quadratique pour tout $(M, N) \in \Sigma$ et chaque élément de Γ est alors une approximation Σ-admissible de B dans Γ. Le théorème 4.1 en résulte aussitôt.

Soulignons que la formulation de ce théorème n'exige pas que Σ soit un cône convexe fermé à condition que $\mathfrak{I}$ désigne alors le cône convexe fermé engendré par l'ensemble $\Sigma + \mathscr{S}(B, \Gamma)$.

RÉFÉRENCES

[1] N. Bourbaki, *Eléments de mathématique: Livre V, Espaces vectoriels topologiques*, Chap. 1–2, Actualités Sci. Indust. n° 1189 (2° éd.), Hermann, Paris 1966.

[2] L. Chambadal et J. L. Ovaert, *Algèbre linéaire et tensorielle*, Dunod, Paris 1968.

[3] W. Klonecki et S. Zontek, *On the structure of admissible linear estimators*, J. Multivariate Anal. 24 (1988), pp. 11–30.

[4] L. R. LaMotte, *On admissibility and completeness of linear estimation in a general linear model*, J. Amer. Statist. Assoc. 72 (1977), pp. 438–441.

[5] — *Admissibility in linear estimation*, Ann. Statist. 10 (1982), pp. 245–256.

[6] A. Olsen, J. Seely et D. Birkes, *Invariant quadratic unbiased estimation for two variance components*, ibidem 4 (1976), pp. 878–890.

[7] J. L. Philoche, *A propos du théorème de Gauss–Markov*, Ann. Inst. H. Poincaré 4 (1971), pp. 271–281.

[8] C. R. Rao, *Projections under semi-norms and generalized Moore–Penrose inverses*, Linear Algebra Appl. 9 (1974), pp. 155–167.

[9] — *Projectors, generalized inverses and the BLUE's*, J. Roy. Statist. Soc. Ser. B 24 (1974), pp. 152–158.

[10] J.-J. Téchené, *L'admissibilité en approximation quadratique d'applications linéaires*, Preprint, Laboratoire de Mathématiques Appliquées, U.R.A. – C.N.R.S. 1204, Université de Pau, 1993.

[11] — *Logique des moindres-carrés et inférence statistique*, Thèse d'Etat, Université de Pau, 1994.

Laboratoire de Mathématiques Appliquées – U.R.A. – C.N.R.S. 1204
I.P.R.A. – Université de Pau et des Pays de l'Adour
Avenue de l'Université – 64000 Pau, France

Received on 14.12.1993;
revised version on 8.11.1994